UNIVERSITÉ DE LORRAINE | BIBLIOTHÈQUES UNIVERSITAIRES

AVERTISSEMENT
DROIT D'AUTEUR

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document. Toute contrefaçon, plagiat ou reproduction illicite encourt une poursuite pénale.

Contact bibliothèque : ddoc-theses-contact@univ-lorraine.fr
(Cette adresse ne permet pas de contacter les auteurs)

Code de la Propriété Intellectuelle. articles L 122. 4
Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

UNIVERSITÉ DE LORRAINE
34 COURS LÉOPOLD — CS 25233
54052 NANCY CEDEX
TÉL : 03 54 50 54 00
CONTACT@UNIV-LORRAINE.FR
WWW.UNIV-LORRAINE.FR

École doctorale IAEM Lorraine

Optimisation et Approches d'Apprentissage Automatique pour la Planification Intelligente de la Maintenance

THÈSE

présentée et soutenue publiquement le 14 octobre 2025

pour l'obtention du

Doctorat de l'Université de Lorraine

(Automatique, Traitement du signal et des images, Génie informatique)

par

Ryan O'Neil

Composition du jury

Président :	Dr. Lyes Benyoucef	Professeur, Aix-Marseille Université
Rapporteurs :	Dr. Bernard Penz	Professeur, Institut Polytechnique de Grenoble
	Dr. Bruno Castanier	Professeur, Université d'Angers
Examinateurs :	Dr. Ayse Akbalik	Professeur Associé, Université de Lorraine
	Dr. Ahmed Saif	Professeur Associé, Dalhousie University
Invité :	Dr. Uday Venkatadri	Professeur, Dalhousie University
Encadrants :	Dr. Abdelhakim Khatab	Professeur, Université de Lorraine
	Dr. Claver Diallo	Professeur, Dalhousie University

Laboratoire de Génie Informatique, de Production et de Maintenance

Remerciements

I would like to express my gratitude to my supervisors, Dr. Claver Diallo and Dr. Abdelhakim Khatab, for their unwavering support and guidance throughout my PhD journey. Their dedication and expertise have been a constant source of inspiration and motivation. Looking back at the person I was when I first began graduate school, compared to the confident researcher and academic I am today, I can truly say that I would not have achieved this growth without their supervision. I could not have asked for more supportive and understanding mentors who have been true role models and whom I will strive to emulate should I choose to pursue an academic career.

I sincerely thank my committee members, Drs. Ayse Akbalik, Bernard Penz, Ahmed Saif, and Uday Venkatadri. Their expertise, insightful feedback, and willingness to offer guidance and support have been invaluable in shaping my research and bringing it to completion. I would also like to extend my deep appreciation to my external referees, Drs. Bruno Castanier, and Bernard Penz for accepting this role and for their thoughtful questions and valuable feedback. My deep appreciation also goes to my external examiner, Dr. Lyes Benyoucef, for accepting the role of president of the jury. Moreover, I would like to thank Université de Lorraine for their support during the joint PhD program, and the Laboratoire de Génie Informatique, de Production et de Maintenance (LGIPM) and Dr. Nidhal Rezg and Dr. Sofiene Dellagi for hosting me during my time in Metz, France.

Lastly, to my parents, the ones who believed in me the most and supported me through everything. I would not be where I am today without your love and support. I can never fully repay you for the time and effort you have invested in me, but I will spend the rest of my life striving to make you proud.

Contents

List of Tables v

List of Figures vii

Chapter 1 Introduction 1
- 1.1 Research Themes . . . 2
  - 1.1.1 Theme 1: Joint Selective Maintenance and Workforce Routing . . . 3
  - 1.1.2 Theme 2: Predictive Selective Maintenance . . . 3
  - 1.1.3 Theme 3: Resource-Constrained Fleet Selective Maintenance . . . 4
  - 1.1.4 Theme 4: Joint Selective Maintenance and Mission Abort Decisions . . . 4
- 1.2 Dissertation Structure . . . 5

Chapter 2 Joint Selective Maintenance and Workforce Routing 7
- 2.1 Optimal Joint Maintenance and Orienteering Strategy . . . 7
  - 2.1.1 Introduction . . . 7
  - 2.1.2 Literature Review . . . 9
  - 2.1.3 System Description and Reliability Computation . . . 13
  - 2.1.4 Maintenance Modeling, and Total Maintenance Time and Cost Computation . . . 16
  - 2.1.5 Mathematical Formulations . . . 17
  - 2.1.6 CG Approach for Large-Scale JSMOP . . . 21
  - 2.1.7 Numerical Experiments . . . 26
  - 2.1.8 Conclusions and Future Research Directions . . . 42
- 2.2 Enhancing Critical Network Infrastructure Resilience . . . 43
  - 2.2.1 Introduction . . . 43
  - 2.2.2 System Description and Reliability Computation . . . 47
  - 2.2.3 Model Formulation . . . 52
  - 2.2.4 Solution Methods . . . 54
  - 2.2.5 Numerical Experiments . . . 59
  - 2.2.6 Conclusions and Future Research Directions . . . 77

Chapter 3 Predictive Selective Maintenance 79
- 3.1 Optimal Predictive Selective Maintenance for Fleets of Mission-Oriented Systems . . . 79
  - 3.1.1 Introduction . . . 79
  - 3.1.2 Literature Review . . . 81
  - 3.1.3 System Description . . . 84

3.1.4 PdSM Optimization Approach . . . 86
3.1.5 Numerical Experiments . . . 92
3.1.6 Conclusions and Future Research Directions . . . 106
3.2 Optimizing Predictive Maintenance and Mission Assignment . . . 109
3.2.1 Introduction . . . 109
3.2.2 Problem Description . . . 111
3.2.3 The Deterministic Formulation . . . 115
3.2.4 Handling Uncertainty in Maintenance and Break Durations . . . 117
3.2.5 Case Study: Military Aircraft Fleet Maintenance . . . 118
3.2.6 Conclusions and Future Research Directions . . . 127

Chapter 4 Optimizing Resource-Constrained Fleet Selective Maintenance With Asynchronous Maintenance Breaks 129
4.1 Introduction . . . 129
4.2 Literature Review . . . 130
4.2.1 The SMP . . . 130
4.2.2 The FSMP . . . 132
4.2.3 Resource Constraints in Maintenance Scheduling . . . 133
4.2.4 Motivation and Contributions . . . 134
4.3 System Description and Assumptions . . . 134
4.3.1 Notation . . . 134
4.3.2 Assumptions . . . 135
4.3.3 Fleet Description . . . 135
4.3.4 Maintenance Modeling and Reliability Computation . . . 136
4.4 FSM Optimization Model and Solution Method . . . 137
4.4.1 Bi-Objective Nonlinear Formulation . . . 137
4.4.2 Model Reformulation . . . 138
4.5 Results and Discussions . . . 140
4.5.1 Experiment 1: Validation . . . 140
4.5.2 Experiment 2: Impact of Overlap Between SMB Windows . . . 141
4.5.3 Experiment 3: Application to a Coal Transportation System Case Study . . . 142
4.5.4 Experiment 4: Impact of SMB Duration on Fleet Reliability . . . 146
4.5.5 Managerial Insights . . . 149
4.6 Conclusions and Future Research Directions . . . 150

Chapter 5 Joint Selective Maintenance and Mission-Abort Decisions 151
5.1 Introduction . . . 151
5.2 System Description . . . 153
5.2.1 Notation . . . 153
5.2.2 System Description and MAP . . . 154
5.3 Reliability Computation and Maintenance Modeling . . . 155
5.3.1 Computation of the MSP . . . 156
5.3.2 Computation of the SSP . . . 156

5.3.3 Maintenance Modeling . . . 157
5.4 Formulation of the Mathematical Optimization Problem . . . 158
5.4.1 Solution Method . . . 158
5.5 Numerical Experiments . . . 159
5.5.1 Experiment 1: Validation . . . 160
5.5.2 Experiment 2: Impact of Initial System Age . . . 162
5.5.3 Experiment 3: Impact of Reliability Thresholds . . . 164
5.5.4 Experiment 4: Impact of the Break and Rescue Durations . . . 168
5.5.5 Summary of Insights . . . 171
5.6 Conclusions and Future Research Directions . . . 171

Chapter 6 Conclusion **172**

Appendixs **175**

Appendix A Supplementary Material for Section 2.1 **175**
A.1 Sequential optimization approach . . . 175
A.1.1 BIP formulation of the SMP . . . 175
A.1.2 MILP formulation of the crew routing problem . . . 176
A.2 Component and System Data . . . 177
A.3 Details of the optimal maintenance actions obtained for experiment set #5 (2.1.7) . . . 179

Appendix B Supplementary Material for Section 2.2 **181**
B.1 Network and component data used in experiments #1 and #2 . . . 181
B.2 Network and component data used in experiments #3 . . . 181

Bibliography **184**

List of Tables

2.1 Traveling time (hours) between turbines and maintenance base . . . 28
2.2 Traveling cost ($) between turbines and maintenance base . . . 28
2.3 Elapsed downtime t_i^d of failed OWTs . . . 28
2.4 Results of validation experiment with k-out-of-n:G subsystems . . . 29
2.5 Comparison of computation times based on the values of M . . . 29
2.6 Comparison of optimal values: case of Experiment #3 . . . 30
2.7 Patterns and resulting total time and cost: case of the joint optimization approach (Experiment #3) . . . 31
2.8 Patterns and resulting time and cost: case of the sequential optimization approach: case of Experiment #3 . . . 31
2.9 Repair crew assignment and routing obtained by JSMOP: case of Experiment #3 . . . 31
2.10 Repair crew assignment and routing obtained by the sequential approach: case of Experiment #3 . . . 32
2.11 Impact of reliability and multiple imperfect maintenance levels: case of Experiment #4 . . . 32
2.12 Results obtained for $T^{max} = 12$ and varying values of R_0: case of Experiments #5 . . . 33
2.13 Maintenance results obtained for $R_0 = 97\%$ and $T^{max} = 12$: case of Experiments #5 . . . 33
2.14 Repair crew assignment and routing for $R_0 = 97\%$ and $T^{max} = 12$: case of Experiments #5 33
2.15 Results when the shift length is varied: case of Experiments #5 . . . 34
2.16 Repair crew data: case of Experiments #6 . . . 36
2.17 JSMOP results for varying values of target reliability: case of Experiments #6 . . . 36
2.18 The 28 problem instances used in Experiment #7 . . . 38
2.19 Repair crew data: case of problem sinstances 17 to 28 if Experiment #7 . . . 38
2.20 Comparison of solution approaches for problem instances 1 to 16: case of Experiment #7 39
2.21 Comparison of solution approaches for problem instances 17 to 28: case of Experiment #7 40
2.22 Results for large-scale wind farm comprised of 80 OWTs: case of Experiment #8 . . . 40
2.23 Repair crew assignment and routing for $R_0 = 99\%$ and $T^{max} = 75$ hours: case of Experiments #8 . . . 41
2.24 Performance evaluation of two-phase approach. . . . 67
2.25 Maintenance actions selected and repair crew routing for $R_0 = 90.0\%$: case of the set of Experiments #2. . . . 67
2.26 Maintenance actions selected and repair crew routing for $R_0 = 92.5\%$: case of the set of Experiments #2. . . . 67
2.27 Comparison of Model 1 and Model 2. . . . 71
2.28 Comparative analysis of solution methods. . . . 76
2.29 Comparative analysis of genetic algorithm and Lagrangian decomposition. . . . 77

3.1 Comparison of proposed method for RUL point prediction with the SOTA. . . . 96
3.2 Comparison of prediction metrics at different monitoring points. . . . 97
3.3 Comparison of α-Coverage (α-C) and α-Mean Width (α-MW). . . . 99
3.4 Engine data used for case study. . . . 104
3.5 Results for varying values of $\mathcal{R}_0$ and Π_{ijk} with $D = 15$ hours. . . . 105
3.6 Impact of the break and mission duration on maintenance plans. . . . 107
3.7 Engine assignment. . . . 119

3.8 Component data for standard components. . . . 120
3.9 Maintenance duration parameters for different components and maintenance levels. . . . 120
3.10 Mission profiles. . . . 120
3.11 Impact of dropout rate on prediction accuracy. . . . 121
3.12 Comparison of Deterministic and Stochastic Solutions. Mean and standard deviation (in brackets) reported for CVaR approach. . . . 125

4.1 Results for different values of C_0 when $D_1 = 5$. . . . 140
4.2 Input data for experiment #2. . . . 141
4.3 Results for models 1 and 2: case of experiment #2. . . . 144
4.4 Input data for experiment #3. . . . 145
4.5 Fleet reliability R_F^* (%) as Δ_D% varies. . . . 147

5.1 Maintenance data. . . . 161
5.2 Initial age and state of components. . . . 161
5.3 Detailed results: case of experiment #1. . . . 161
5.4 Detailed maintenance plans: case of experiment #1. . . . 162
5.5 Detailed SM plans for MRUL=35 and $C^{mf} = 2000$: case of experiment #2. . . . 165
5.6 Comparison of JMA-SMP and NMA-SMP: case of experiment #4. . . . 169
5.7 Detailed SM plan: case of experiment #4. . . . 170

A.1 Reliability data for OWT i . . . 178
A.2 Component status and age . . . 178
A.3 Maintenance actions performed for $R_0 = 86\%$: case of Experiments #5 . . . 179
A.4 Maintenance actions performed for $R_0 = 90\%$: case of Experiments #5 . . . 179
A.5 Maintenance actions performed for $R_0 = 94\%$: case of Experiments #5 . . . 179
A.6 Maintenance actions performed for $R_0 = 98\%$: case of Experiments #5 . . . 180
A.7 Maintenance actions performed for $R_0 = 99\%$: case of Experiments #5 . . . 180

B.1 Component information for networks 1 and 2. . . . 181
B.2 Travel time (hours) between components and O&M base (Node 0) for network 1. . . . 182
B.3 Travel time (hours) between components and O&M base (Node 0) for network 2. . . . 182
B.4 Component information for WDN. . . . 183
B.5 Travel time (hours) between components and O&M base (Node 0) for WDN. . . . 183

List of Figures

1.1 Thesis contributions. 6

2.1 Illustration of an offshore wind farm 8
2.2 Reliability structure of the OWT at location i 15
2.3 Reliability structure of the OWT at location i 27
2.4 Illustration of paths selected for each repair crew 34
2.5 Trade-off between reliability and total cost 35
2.6 Trade-off between shift duration and total cost. 35
2.7 New extended resilience framework adapted from Fan et al. (2022). 48
2.8 Superimposed infrastructure and routing networks (adapted from Morshedlou et al. (2018)). 50
2.9 Infrastructure and routing network of the illustrative example. 51
2.10 Series structure of d-MC parallel structures: Case of the network in Figure 2.9. 52
2.11 Encoding of a solution. 56
2.12 Networks #1 and #2 considered for the first set of experiments. 61
2.13 SRI and total routing and maintenance time for varying values of T. Case of experiment #1.1. 62
2.14 Restoration curves and route selection for varying values of T. Case of experiment #1.1. . 63
2.15 Restoration curves for varying values R_0. Case of experiment #1.2. 64
2.16 Restoration curves for varying values K. Case of experiment #1.3. 65
2.17 Impact of $|\mathcal{T}|$ on network resilience and computational time of the two-phase approach. Case of experiment #1.4. 66
2.18 Restoration curves and network reliability for varying values of R_0: case of the set of Experiments #2. 68
2.19 Water distribution network presented in Kansal et al. (1995). 69
2.20 Transformed water distribution network. 70
2.21 Restoration curves for varying disaster scenarios: case of the of Experiments #3. 72
2.22 Route selection for disaster scenarios #1-#3. 73
2.23 Route selection for disaster scenarios #4-#6. 74
2.24 Network considered for fourth set of experiments. 75
2.25 Evolution of genetic algorithm and Lagrangian decomposition. 76

3.1 Flowchart of the proposed PdSM framework for fleets of MOS. 86
3.2 Clustering results for multiple training and testing engines. 95
3.3 RUL interval estimates for multiple test engines. 98
3.4 RUL *pdf* plots for multiple test engines. 98
3.5 Failure probability curves when $U = 40$ cycles for multiple test engines. 99
3.6 Comparing methods for reliability assessment for $U = 20$ cycles and varying values of $\mathcal{R}_0$. Case of experiment #1. 102
3.7 Comparing methods for reliability assessment for $\mathcal{R}_0 = 0.95$ and $D = 10$ hours for varying values of Π_{ijk}. Case of experiment #2. 103
3.8 Component downtime and total cost for varying values of Π_{ijk}. 106
3.9 Total cost for varying values of U and D with $\Pi_{ijk} = 4$/cycle and $\mathcal{R}_0 = 0.950$. 107
3.10 Aircraft RBD considered for the case study. 118
3.11 RUL interval estimate for test engine 24 for varying dropout rates. 122

3.12 RUL distribution plots for test engine 90 for varying dropout rates. . . . 123
3.13 Impact of reliability threshold and dropout rate on early replacements and system failures for different mission lengths. . . . 124
3.14 Probability density functions of the three break durations considered. . . . 124
3.15 Mean and standard deviation of penalty cost, variable cost, and fixed cost for varying values of mean break duration. . . . 125
3.16 Constraint satisfaction probability for varying values of mean break duration. . . . 125
3.17 Impact of standard deviation of break duration on mission selection probability and objective function value. . . . 126
3.18 Average variable cost, fixed cost, and overtime for varying values of α. . . . 127
3.19 Average maintenance durations for varying values of α. . . . 127

4.1 Illustration of the synchronous and asynchronous maintenance break cases. . . . 134
4.2 Alternating sequence of missions and scheduled breaks for the fleet of n systems. . . . 136
4.3 Gantt charts of maintenance plans for 75%, 50%, and 0% overlaps. . . . 143
4.4 Real-world coal conveyor system (Liu and Huang, 2010; Khatab et al., 2022). . . . 145
4.5 Trends of fleet reliability R_F^* (%) and total Cost as the number of resources increase. . . . 146
4.6 Impact of the number of resources and available budget (C_0) on R_F^*. . . . 147
4.7 Break duration structure for experiment #4. . . . 148
4.8 Maintenance times by system and end date percent difference Δ_D. . . . 148
4.9 Trends of R_F^* by available resources and end date percent difference Δ_D. . . . 149

5.1 RBD of the system employing a mission-abort policy. . . . 155
5.2 Description and features of UAV system considered. . . . 160
5.3 Impact of V on expected costs: case of experiment #1. . . . 162
5.4 Cost components for varying values of MRUL: case of experiment #2. . . . 164
5.5 Cost breakdown for varying values of C^{sl}: case of experiment #2. . . . 165
5.6 V and total maintenance time per block for varying values of MRUL: case of experiment #2. . . . 166
5.7 Maintenance and penalty costs for varying MRUL and D ($C^{mf} = 2000$, $C^{sl} = 15000$): case of experiment #2. . . . 166
5.8 V and total maintenance time for varying D ($C^{mf} = 2000$, $C^{sl} = 15000$): case of experiment #2. . . . 167
5.9 Results for varying MSP_0: case of experiment #3. . . . 168
5.10 Results obtained for varying SSP_0: case of experiment #3. . . . 168
5.11 Results obtained for varying values of Δ: case of experiment #4. . . . 169

Chapter 1

Introduction

Modern life depends upon complex and highly interconnected production, distribution, and logistics systems for energy, goods, and services. These systems are designed to operate with little interruption, extraordinarily high reliability, and high readiness to better respond to disruptions caused by unforeseen events such as weather, natural disasters, and pandemics. Their failure can be extremely costly, making it increasingly crucial for organizations to efficiently manage and maintain equipment and machinery, enhance asset availability, and reduce costs.

The emergence of Industry 4.0, which encompasses next-generation manufacturing technology and smart factories, integrates embedded systems, sensors, machine-to-machine communication, IIOT, and CPS to bridge the gap between the virtual and physical realms (Xu et al., 2018; Kusiak, 2018). These advancements hold the potential to enhance the production and distribution of personalized goods at costs comparable to mass production. The integration of supply chains, production systems, and logistic networks results in fully integrated and responsive systems (Kaynak, 2007; Ivanov et al., 2016). However, despite these advancements, ensuring the continuous operation of such networks remains critical; their failures can be highly disruptive and costly. Therefore, monitoring system health, predicting failures, and implementing proactive maintenance strategies are mandatory to sustaining high performance across these integrated networks.

Many production, distribution, and logistics systems operate through alternating missions and breaks. These systems are referred to as MOS and are commonly found across various industries, including defense, petrochemistry, transportation, and logistics. SM is widely applied in such contexts to assess and enhance system reliability or RUL. With the aim of ensuring optimal performance of these systems during missions, it becomes imperative to carry out maintenance during the scheduled breaks. Due to limited resources such as cost, time, spare parts, or number of maintenance technicians, it is usually not possible to perform all desirable maintenance actions. The SMP seeks to answer the question of which maintenance actions to perform on what components such that either (i) the system reliability is maximized for the upcoming mission or (ii) the system meets a specified reliability target for the upcoming mission at the lowest overall cost. There are numerous industrial and military applications where SM can be implemented. For instance, computer systems are commonly used throughout the day but can be maintained outside of business hours. Fleets of delivery vehicles are dispatched during the day but are available in the evenings for maintenance. Production systems are expected to operate continuously on weekdays but undergo maintenance shut-downs on weekends. During military operations, maintenance is performed between combat missions.

Since its inception in 1998 by Rice et al. (1998), the SMP has garnered considerable interest from both academic researchers and industrial practitioners. Nevertheless, there remain several crucial real-life scenarios that require attention in order to substantially enhance understanding of SM modeling and optimization. The present research work aims to address four pivotal gaps as outlined below:

1. Existing SMP models assume that the maintenance decisions are made and performed on systems which are at the same location and have access to ample repair crews. This is not always the case

since industrial systems can be geographically distributed and their maintenance may require maintenance crews to travel. The first theme of this research looks to develop joint SM and repair crew routing models that can be applied to fleets of systems or components of networked infrastructure that are geographically dispersed.

2. The majority of SMP models proposed in the existing literature merely rely on statistical models for the computation of component reliability. The emergence of new technologies like wireless sensors and IIOT, along with the advancements in AI, has popularized data-driven approaches based on DL algorithms for reliability assessment and RUL prediction. The objective of the second research theme is indeed to develop a PdSM framework that adopts DL approaches to predict component reliability.
3. Limited SMP models in the literature consider fleets of multi-component systems and the inherent constraints on maintenance capacities when planning maintenance tasks. Although the joint SMP and repair person assignment problem has received attention in recent years, repair persons are only one of many maintenance resources that must be considered. Other capacities, such as specialty tools and repair facilities, also need to be incorporated into the maintenance planning problem. Additionally, SMP models for fleets of systems often rely on the restrictive assumption that missions and breaks begin simultaneously for all systems and that upcoming missions are identical, requiring only a single system for successful completion. The third theme of this research aims to develop joint SM and resource scheduling optimization models for fleets of systems with asynchronous maintenance break durations.
4. No existing studies on SM have explored systems that implement a MAP. In these systems, a mission can be aborted, and a rescue procedure must then be initiated to safeguard lives or valuable equipment when the risks associated with continuing the mission reach a critical level. The fourth and final theme of this research delves into the SMP applied to systems incorporating mission abort policies, where the trade-offs between optimizing two metrics, namely the MSP and SSP, will be explored.

In reliability theory, a system is typically represented by the logical configuration of its components required to perform a specified system function. This representation of the system will be adopted throughout this thesis. It is assumed that complete access to all necessary modeling and evaluation data is available. While this may not always hold in all applications, the developed models aim to provide a general framework that can be adapted to various mission-oriented settings.

In the present dissertation, four themes are explored, each one addressing one of the above mentioned knowledge gaps. Each theme is developed in a dedicated self-contained chapter that has its own literature review, introduction and results. Regarding the first theme, a paper introducing the problem of joint SM and orienteering was published in *Computers & Operations Research*. A second paper proposing a joint maintenance and repair crew routing model to enhance network resilience after disruptive events has been published in *Reliability Engineering & System Safety*. In terms of the second theme, a paper proposing a novel PdSM strategy was presented at the 10th *Manufacturing Modeling, Management and Control* conference held in Nantes, France. A paper significantly extending and improving upon this work has been published in the *International Journal of Production Research*. A paper developing a novel data-driven FSM model that jointly optimizes system-to-mission assignment, component and maintenance level selection, and repair task allocation has been submitted to *Autonomous Intelligent Systems* journal. As per the third theme, a paper considering asynchronous maintenance breaks and resource-constrained maintenance planning for fleets of multi-component systems has been published in the *International Journal of Production Research*. Finally, in terms of the fourth theme, a paper introducing a novel model to jointly optimize SM scheduling and mission-abort decisions for mission-critical systems has been resubmitted to *Reliability Engineer & System Safety* after revisions were requested. In the following section, the motivations and objectives of each research theme are summarized.

1.1 Research Themes

In this section, the objectives and motivations of each of the above listed research themes will be presented as well as the key research questions they aim to address.

1.1.1 Theme 1: Joint Selective Maintenance and Workforce Routing

The first theme investigated in this dissertation centers around the joint optimization of maintenance and workforce routing decisions. The challenge involves jointly selecting system components for maintenance, determining the appropriate level of maintenance to carry out, and optimizing the routing of multiple maintenance crews. Traditionally, these decisions have been made independently, leading to sub-optimal outcomes and sub-optimal resource utilization. Considering them simultaneously can potentially lead to significant improvements in operational efficiency, cost reduction, and overall system performance.

The overarching objective of the first theme is to develop optimization models that encompass the joint optimization of maintenance and workforce routing decisions. By leveraging techniques from operations research, optimization algorithms capable of efficiently solving the large-scale and complex problem instances inherent in real-world scenarios will be developed. The contributions of this theme extend beyond theoretical advancements. By jointly optimizing maintenance and workforce routing decisions, we aim to provide practical decision support tools that can be implemented in diverse industries such as manufacturing, transportation, energy, and telecommunications. The development of effective optimization approaches will indeed allow organizations to make informed decisions, leading to improved resource allocation, enhanced operational performance, and ultimately, increased competitiveness.

This theme aims to answer the following questions:

- Are there benefits in terms of operational costs or system performance in jointly optimizing maintenance and workforce routing decisions?
- How can maintenance decisions and workforce routing be effectively integrated to improve operational efficiency?
- What are the key factors and constraints that need to be considered in jointly optimizing maintenance and workforce routing decisions?
- What large-scale optimization algorithms and metaheuristics can capture the complexity of the joint optimization problem and provide efficient solutions?
- How can the joint optimization approach be adapted and applied to different industries and operational contexts?

1.1.2 Theme 2: Predictive Selective Maintenance

Industry 4.0 is the latest trend in automation that employs sensors, CPS, and IIOT to connect the physical and virtual world. As a result of implementing these innovative technologies, predictive maintenance is increasingly being recognized as an appealing maintenance strategy in numerous industries. A key element of predictive maintenance is prognostics, which deals with fault prediction before it occurs (Jardine et al., 2006). There are two main prediction types in prognostics. The first is RUL prediction, and the second deals with predicting the probability that a component operates without failure up to some future time. Recently, DL algorithms have made tremendous strides in performance and have improved the state of the art in prognostics and predictive maintenance (Namuduri et al., 2020). Effective predictive maintenance plans have the capability of reducing maintenance costs, extending equipment life, reducing machine downtime, and improving production. Prognostic methodologies encompass traditional ML techniques, including logistic regression, support vector machines, and ensemble methods, as well as deep DL techniques, such as multilayer perceptron models, plCNN, and plRNN. DL algorithms require tens to hundreds of thousands of training samples to achieve a desirable level of accuracy. IIOT sensors can produce massive amounts of data, thus enabling DL algorithms to be applied. ML and DL methods can identify a relationship between sensor values and dependent variables such as time of failure, future sensor value, or the detection of anomalies.

The second theme of this dissertation focuses on the development of cutting-edge data-driven predictive SM strategies. The aim is to leverage the power of both DL techniques combined with optimization approaches to identify optimal maintenance decisions for MOS. This strategy, referred to as PdSM, draws inspiration from the field of predictive maintenance for machine prognostics. The concept of predictive SM has gained attention due to its potential to significantly enhance the reliability and performance of MOS. Building upon the work of Hesabi et al. (2022), who introduced the first predictive SM framework, the

aim of this research theme is to further advance this approach to address the limitations and challenges associated with complex systems. This theme also considers the impact of component degradation level on the time required to perform maintenance. The optimization model developed not only identifies the optimal set of components to be maintained, but also jointly provides optimal repairperson maintenance task assignment decisions. An efficient solution method based on a two-phase decomposition approach is developed. Finally, rather than using a probabilistic classifier to predict component reliability, a DL algorithm combined with MCD is used. The advantages of this approach will be highlighted and fully discussed.

This research theme aims to answer the following questions:
- What are the limitations associated with existing predictive SM frameworks? How can these limitations be overcome?
- How can DL techniques be effectively utilized for reliability prediction in the context of PdSM?
- How can the uncertainty in RUL and reliability prediction be addressed?
- How can the impact of component degradation levels on maintenance time requirements be quantified and incorporated into the maintenance planning process?

1.1.3 Theme 3: Resource-Constrained Fleet Selective Maintenance

The third theme of this thesis focuses on extending the classical FSMP to include detailed resource scheduling and systems with asynchronous maintenance breaks. All existing FSMP models merely rely on the restrictive assumption according to which all systems in the fleet operate synchronously, meaning that maintenance tasks can only be decided and executed when all these systems have completed the current mission and are all available simultaneously (Al-Jabouri et al., 2024b). However, in real-world operations, fleets consist of multicomponent systems with staggered mission schedules, requiring asynchronous maintenance planning. Furthermore, planning maintenance for a fleet requires detailed scheduling and maintenance capacity allocation, yet no existing SMP models explicitly account for the primary resources needed to execute maintenance tasks, such as tools, facilities, and repair personnel. This research addresses these limitations by proposing a generalized resource-constrained FSMP model, which introduces asynchronous maintenance planning and explicit resource allocation. The model establishes a theoretical connection between the SMP and the RCPSP, providing a more flexible and realistic approach to fleet maintenance planning.

This research theme aims to answer the following questions:
- How can asynchronous maintenance breaks and resource constraints be effectively modeled in a SM framework?
- What are the impacts of asynchronous break durations on maintenance selection decisions and achievable fleet reliability?
- What managerial insights can be derived from this model, and how can it support decision-making for maintenance managers?

1.1.4 Theme 4: Joint Selective Maintenance and Mission Abort Decisions

In the realm of SM models, the focus has mainly been on maximizing mission success probability (MSP), a vital measure when evaluating the performance of MOS. However, certain MOS, such as aircraft, spacecraft, and unmanned aerial vehicles (UAVs), place a higher emphasis on system survival than mission success. The loss of these systems can result in significant economic, environmental hazards, and even human life consequences. To mitigate such risks, mission abort policies (MAPs) have been introduced, allowing for the initiation of rescue procedures to prevent system failures. In this context, this research theme of the thesis aims to address the challenges and opportunities presented by systems with plMAP in SM setting.

In systems operating under MAPs, the decision to abort a mission and initiate a rescue procedure is influenced by various factors. The number of failed components (Levitin et al., 2017) and the number of external shocks (Levitin and Finkelstein, 2018) impacting the system can all serve as triggers for mission

termination. This research aims to develop optimization models that incorporate these factors, providing insights into the allocation of maintenance resources to enhance both mission success probability (MSP) and SSP metrics. By considering the trade-offs between these competing objectives, one can optimize the decision-making process and ensure the optimal allocation of maintenance resources to mission-critical and safety-critical systems.

The primary objective of this research theme consists of bridging the gap in knowledge regarding SM for systems operating under MAPs. While existing literature has extensively explored the optimization of maintenance decisions for MSP, this research will be the first to consider systems with mission abort capabilities within the SM framework. By developing SM optimization models tailored to these unique systems, we can unlock significant advancements in preserving system survival and mitigating risks. The research outcomes will have far-reaching implications for industries relying on mission-critical systems, ranging from aviation and space exploration to UAV operations.

This research theme aims to answer the following questions:
- How do we define the mission success and system survivability probabilities?
- How can the selection and allocation of maintenance resources, including time, budget, personnel, and spare parts be optimized to maximize both mission success and system survivability probabilities?
- What are the key parameters and factors that should be considered in the development of SM models for MOS with mission abort policies?
- What are the potential trade-offs between maintenance costs, MSP, and SSP in MOS?

1.2 Dissertation Structure

As it was mentioned previously, this dissertation is a thesis by article and is composed of six manuscripts (four published, one resubmitted after revisions and one submitted for publication). Chapter 2 presents joint models that optimize SM and workforce routing decisions. Two applications are explored: (1) maintenance service providers tasked with servicing plOWT from an operation and maintenance support port, and (2) joint maintenance and routing decisions to enhance network resilience post-disruption. Chapter 3 introduces a predictive SM framework leveraging DL. A DNN with MCD is used for reliability assessment, while a fuzzy clustering algorithm identifies components' health states. Numerical experiments demonstrate the benefits of this data-driven prognostic approach. Chapter 4 extends the classical FSM model by incorporating asynchronous maintenance breaks and detailed maintenance resource constraints. Chapter 5 develops a joint SM and mission abort decision-making model applicable to mission-oriented and mission-critical systems. Extensive experiments demonstrate that jointly optimizing mission-abort decisions and maintenance planning enhances system survivability compared to cases without an abort policy.

From a theoretical perspective, this thesis introduces novel mathematical models and solution methods. From a practical perspective, it significantly extends the applicability of the SMP. This thesis is timely: it addresses maintenance and reliability requirements for large multicomponent systems, advances the state of the art in modeling such systems, and uses emerging methods for solution such as data analytics and AI. Figure 1.1 summarizes the contributions made in the thesis by addressing key research gaps across four primary areas, including system, reliability, degradation, and maintenance modeling, as well as fleet-level planning.

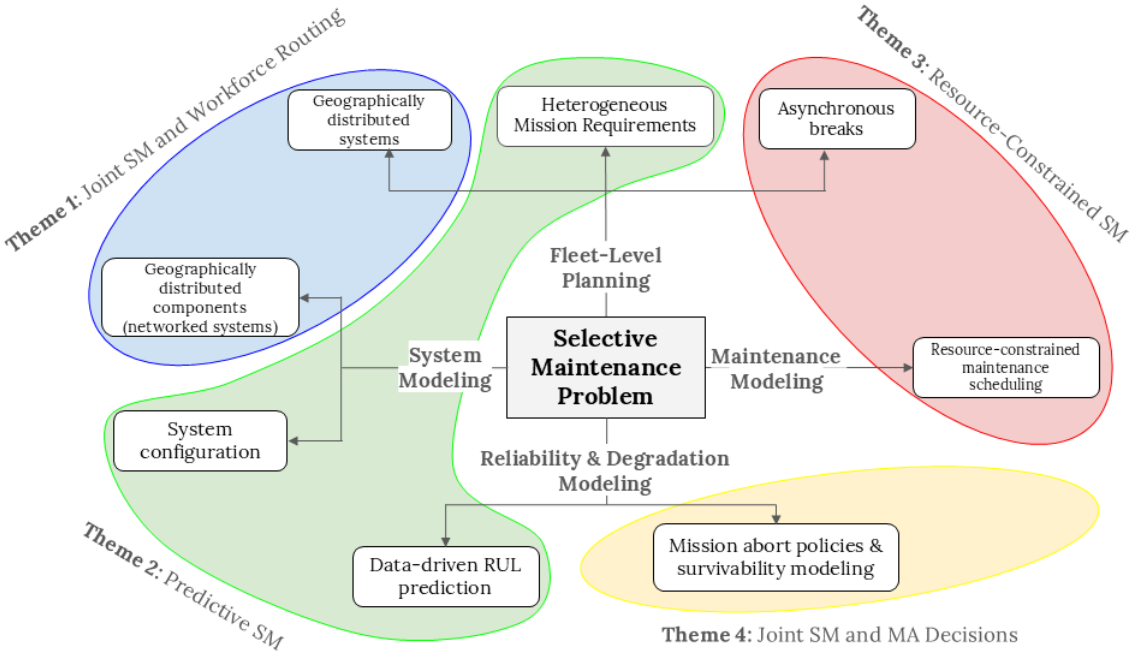


Figure 1.1 – Thesis contributions.

Chapter 2

Joint Selective Maintenance and Workforce Routing

This chapter introduces and solves the JSMWRP by exploring two key applications. In Section 2.1, the maintenance of offshore wind farms is examined, where service providers dispatch maintenance crews from an operation and maintenance (O&M) support port to service plOWT. In Section 2.2, a post-disruption network restoration model is developed, where joint maintenance and routing decisions are optimized to enhance network resilience following a disruptive event.

2.1 Optimal Joint Maintenance and Orienteering Strategy for Complex Mission-Oriented Systems: A Case Study in Offshore Wind Energy

This section presents the formulation and solution of the joint maintenance and OP, with a focus on its application to offshore wind farms. The quest for sustainable energy production is fueling the growth of offshore wind electricity generation. Energy producing plOWT are typically dispersed across several remote wind farms and must be maintained and operated with high reliability levels for long time-periods separated by scheduled maintenance rotations. Due to resource constraints such as travel time, cost, and availability of repair crews, only a subset of turbines and their components can be selected for maintenance operations during maintenance trips. This section proposes a novel joint maintenance and orienteering framework to address the selection of turbines to visit, the components to maintain, the maintenance levels to be performed, the assignment of repair crews, and their routing with the goal of minimizing total cost while satisfying a minimum required reliability threshold during the next operating mission until the next maintenance rotation. A mixed-integer linear programming optimization model is developed and fully discussed. To solve the problem for large-scale instances, a CG method based on Dantzig-Wolfe decomposition and labeling algorithm is proposed. Several numerical experiments demonstrate the validity of the proposed model and the benefit of jointly optimizing maintenance and orienteering decisions. The results show that inclusion of detailed reliability and multiple maintenance levels allows better maintenance and routing decisions. Numerical results also show trade-offs between reliability requirement, shift duration and total cost.

2.1.1 Introduction

Wind turbines have the proven potential in supplying clean energy across the world in a sustainable and cost-effective manner. Due to the higher wind speeds, unoccupied sea surface, and reduced noise pollution, offshore wind power has emerged as a promising renewable energy source and has seen rapid growth in the last decade. In some areas of the globe, large unoccupied sea surface areas are available to develop wind farms. In other areas such as North-West Europe, less space is available and many maritime stakeholders (*e.g.,* container shipping, leisure industry, fishing communities, oil & gas producers) operate and must share the same space and navigation corridors. These constraints compounded with weather uncertainties add to the complexity of selecting a support port location, scheduling and routing OWT

maintenance crews.

According to the International Energy Agency (IEA), offshore wind electricity generation increased by 32% in 2017 to reach 56 TWh and is projected to reach around 606 TWh in 2030 (IEA, 2020). One of the largest cost components of an offshore wind farm is operation and maintenance (O&M) as it accounts for approximately 25-30% of the total life cycle costs Röckmann et al. (2017). It is of importance to develop optimal O&M plans for offshore wind farms so that the produced energy can be sold at a competitive price.

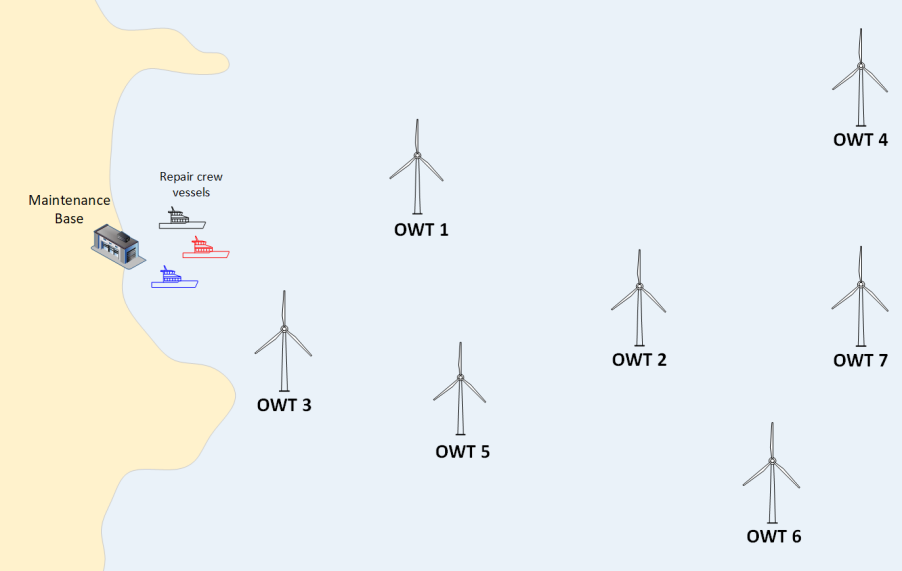


Figure 2.1 – Illustration of an offshore wind farm

On average, OWTs are expected to operate at a relatively high reliability level for six-month periods/rotations separated by scheduled maintenance breaks. To ensure that OWTs continue to operate at the expected reliability level, maintenance actions must be performed during scheduled breaks coinciding with technician visits. It has been approximated that OWTs fail on average 8.3 times per year Carroll et al. (2016). During a given maintenance rotation there may be a set of turbines that have failed and also a set of turbines that have been scheduled for maintenance. Due to time and weather constraints, it may not be feasible for maintenance technicians to visit all turbines during a visit. Therefore, it is critical to select the subset of turbines to visit, the maintenance actions that should be performed at each turbine as to meet or exceed the expected reliability level, and also determine the optimal routes for the different crews of technicians (orienteering problem). The maintenance strategy used, known as SM, has been shown to be effective in addressing MOS (Rice et al., 1998). Currently, the decisions of which maintenance actions to perform and how the repair crews are routed are made sequentially: the SMP first and then the OP.

The objectives of this section are to develop a new mathematical model to solve the JSMOP for moderately large systems. The aim of the proposed model is to identify the components and corresponding maintenance levels to be performed on the visited locations/systems (*e.g.,* turbines) to meet a required reliability level while also determining the optimal route for multiple repair crews. The objective function of the proposed model will minimize the loss in revenue due to downtime of failed systems, travel cost between systems, maintenance cost, and a penalty cost incurred for not performing maintenance on a system, either PM or CM. Although the traditional OP aims to maximize the reward collected, the proposed minimization objective function is equivalent to its maximization counterpart that would maximize the reward from energy produced from maintained wind turbines. The JSMOP is also applicable to the maintenance of a wide spectrum of complex, large, and critical systems geographically distributed and maintained from a pool of crews such as bridges, power-stations, mines, petrochemical plants, and

windmills.

The remainder of this section is structured as follows. Subsection 2.1.2 is a review of relevant papers dealing with the SMP, OP, routing and scheduling of maintenance technician problems, and combined maintenance and routing problems. In Subsection 2.1.3, the notation used, the main working assumptions considered and a description of the system investigated are given. This Subsection also presents reliability computations for an OWT. Subsection 2.1.4 describes the maintenance modeling and total maintenance time and cost derivation for a single OWT. Subsection 2.1.5 presents the mathematical formulation of the joint optimization approach to deal with the combination of the SM and the orienteering problems. Subsection 2.1.6 presents the decomposition method used to solve the JSMOP for large-scale instances. Subsection 2.1.7 presents several sets of numerical experiments carried out to validate the proposed optimization models and solution approach. These experiments demonstrate indeed the benefit of jointly optimizing the SM and orienteering decisions. Conclusions and future research extensions are presented in Subsection 2.1.8.

2.1.2 Literature Review

This review section deals with three topics that are closely related to the proposed optimization model, namely the SM, the orienteering, and the maintenance technician problems.

The SMP

MOS such as military and industrial systems are required to perform a series of missions with finite breaks between missions. Examples of these systems include aerospace equipment, weapons, computers, production and manufacturing lines, and energy production systems. To maintain these systems and ensure they continue to operate at an acceptable performance level during the missions, maintenance actions must be performed during the scheduled break intervals. The quality of maintenance actions that can be performed varies from mR, applied only to failed components and restoring them to the ABAO state, to perfect replacement, which returns components to the AGAN state. IM offers an intermediate level of restoration between these two extremes. Due to resource constraints such as time, cost, and repair person availability, only a subset of system components can be selected for maintenance operations. The SMP aims to identify the optimal subset of components and maintenance actions that will result in meeting the specified system reliability or maximizing the system reliability for the upcoming mission(s) under resource constraints.

The original SM model introduced by Rice et al. (1998) dealt with a series-parallel system with constant failure rate components and perfect repair of failed components. Due to the nonlinear objective function, an enumeration method was used to find the optimal solution. In the intervening years since Rice et al. (1998) proposed the first SMP model, many researchers have expanded upon their work. These studies have included:

— complex system configurations Cassady et al. (2001b); Diallo et al. (2018),
— multistate systems Zhang et al. (2019); Liu and Huang (2010); Pandey et al. (2013),
— component dependence Xu et al. (2016); Dao et al. (2014),
— fleet level SM Khatab et al. (2020); Schneider and Cassady (2015),
— multimission Chaabane et al. (2020); Liu et al. (2022b),
— stochastic break and/or mission duration Liu et al. (2022b, 2018); Khatab et al. (2017a),
— condition-based SMP Khatab et al. (2018a),
— multiple repair channels Liu et al. (2022b); Diallo et al. (2019); Khatab et al. (2018b),
— uncertainty and robustness Jiang and Liu (2020a,b), and
— literature reviews Xu et al. (2015); Cao et al. (2018); Al-Jabouri et al. (2024b).

The general goal of any SMP model is usually to maximize the reliability R of the system under consideration during the mission(s) following one or several intermission breaks without exceeding the total budget C_0. A variant is to minimize the total maintenance cost C that guarantees a minimum required reliability level R_0 during the subsequent mission(s). The total duration of the maintenance actions T should not exceed the length of the intermission break(s) T_0. These formulations take the general forms

below.

Max R		Min C
s.t.		s.t.
$C \leq C_0$	or	$R \geq R_0$
$T \leq T_0$		$T \leq T_0$

The original SMP has been shown to be NP-hard Rice (1999). Therefore, all the extended variants of the SMP are also NP-hard. Hence, the large majority of papers dealing with SMP use full enumeration method for very small sized problems Cassady et al. (2001b,a); Rice et al. (1998). To deal with large sized SMP, nonlinear solvers or heuristics have been used extensively to find near-optimal solutions Lust et al. (2009); Pandey et al. (2013); Cao et al. (2016); Dao and Zuo (2015); Ikonen et al. (2020). Sharma et al. (2017) combined simulation and GA to optimize spare parts forecasting and SM decisions. Cao et al. (2017) used a simulation approach to maximize system availability for an SMP.

Diallo et al. (2018) introduced the first SMP model for serial k-out-of-n systems. The k-out-of-n system is a more complex structure than those previously considered and is a generalization of both the parallel and series configurations. A two-phase approach is developed which eliminates the need for solving a nonlinear objective function as the approach converts the SMP to a multidimensional multiple choice knapsack problem (MMKP). The first phase of the approach involves determining all possible combinations (patterns) of components and maintenance levels, while the second phase consists of solving the resulting BIP.

All SMP papers assume that the maintenance decisions are made and performed on systems which are all at the same location and have access to ample repair crews. This is not always the case since industrial systems can be geographically distributed and their maintenance may require traveling maintenance crews under transportation time and cost constraints. Hence, the need for the proposed joint SM and orienteering problems. A brief review of the OP is presented in the following subsection.

The orienteering problem

The OP can be defined on an undirected graph where the start point (vertex 1), and end point (vertex n) are fixed, and a score is given to the remaining $n-2$ vertices. The aim is to identify a path through a subset of these vertices that maximizes the total path score collected given a time constraint Tsiligirides (1984); Golden et al. (1987); Chao et al. (1996). The OP has been shown to be NP-hard Golden et al. (1987). Vansteenwegen et al. (2011) provide a mathematical model for the basic OP, as well as models for different extensions such as the team OP, and the OP with time windows. The time dependent OP (TDOP) is a variant of the original OP in which the travel time between two vertices depends on the leaving time of the first vertex. This variant was originally introduced in Fomin and Lingas (2002) and has been extended in Verbeeck et al. (2017) to include time windows. Another variant of the OP that has been studied is the capacitated TOP (CTOP) where customers have demands and vehicles have limited cargo capacity. The goal of the CTOP is to determine a path for multiple vehicles through a subset of vertices that maximizes profit collected while satisfying time limits and capacity constraints. The CTOP was introduced by Archetti et al. (2009) and extended to allow for incomplete service (CTOP-IS) Archetti et al. (2013). Afsar and Labadie (2013) extend the classic TOP by considering the profit collected at each customer vertex as decreasing with time.

Due to the uncertain nature of travel and service times, in addition to collected profits, some researchers have studied the OP with stochastic elements. Ilhan et al. (2008) present the OP with scores that are normally distributed. The objective of the formulation is to maximize the probability that the profit collected in a given time limit exceeds a target profit level. Campbell et al. (2011) introduce an interesting variant of the OP in which both travel and service times are stochastic (OPSTS). The goal of the OPSTS is to identify a subset of customers and the order according to which they will be visited to maximize the expected benefit. Zhang et al. (2014) present the stochastic OP with time windows where

the wait time at each customer is modeled as a random variable.

The joint SM and OP presents unique features that differentiate it from the OP problems presented above. In particular, the time spent at a location depends on the required minimum reliability targets, the skills of the repair persons, the components selected to undergo maintenance and the maintenance levels to be carried out. Furthermore, the time spent at a repair location affects the ability to visit and complete other repairs later in the trip. The sojourn time at a location is a result of optimal trade-offs between multiple factors, namely reliability, time, and cost.

Combined routing and maintenance technician problems

Numerous studies have focused on the combined routing and scheduling of maintenance technicians. Tang et al. (2007) model the planned maintenance scheduling problem as a multiple tour maximum collection problem with time-dependent rewards. Kovacs et al. (2012) define the service technician routing and scheduling problem (STRSP) with and without team building and provide mathematical optimization models. The objective of their proposed models is to minimize the sum of the total routing and outsourcing costs. Goel and Meisel (2013) study the combined routing and scheduling problem for the maintenance of electricity networks. An optimization model is presented with the goal of finding an assignment of maintenance tasks to workers and a schedule for performing the tasks such that the downtime and travel costs are minimized. Mathlouthi et al. (2018) consider a multi-attribute TRSP and propose a MILP to address the problem.

There have been a number of studies on the optimization of routing and scheduling of vessels to perform maintenance at offshore wind farms. Stålhane et al. (2015) proposed an arc-flow model with the objective of finding a route for multiple repair crews through a subset of potential plOWT that require maintenance while minimizing the overall cost of travel, maintenance, and loss in revenue due to downtime. The objective function also includes a penalty cost that is incurred when a scheduled maintenance action is not performed. It is assumed in this section that the fleet of service vessels is heterogeneous and the planning horizon is a 12-hour workday. Dai et al. (2015) developed a mathematical model with the aim of routing a heterogeneous fleet of service vessels to different OWTs while minimizing travel cost and a penalty cost incurred per day for delaying maintenance tasks on a turbine. This study considers a planning horizon of multiple days.

The major drawback of these studies is that they merely focused on the routing and scheduling of maintenance technicians and vessels where the maintenance actions to be performed are assumed known and the problem is simply treated as a routing problem. López-Santana et al. (2016) present a two-step iterative approach for the combined maintenance and routing optimization problem for a set of geographically distributed machines. Jia and Zhang (2020) develop a bi-objective optimization approach for the joint optimization of maintenance planning and workforce routing for a networked infrastructure. The objective of the proposed model is to minimize the total maintenance and travel costs while satisfying reliability requirements.

Irawan et al. (2017) studied the maintenance routing and scheduling problem for multiple wind farms and several O&M bases taking into account several constraints such as weather condition, vessels availability and the number of repair crews available in each O&M support port. An integer linear programming optimization model is developed to derive the optimal route and maintenance decisions minimizing the total maintenance cost. This problem is also investigated by Raknes et al. (2017) where one or more wind farms share the same vessel fleet with one onshore depot. A simulation-based solution method is used for maintenance scheduling, and crew transfer vessel (CTV) routing. Schrotenboer et al. (2018) also studied the routing problem for wind farms sharing technicians over a multi-period time horizon. The problem addressed consists to determine, for each period, the vessel routes to deliver and pick up technicians. Schrotenboer et al. (2019) developed an exact solution method based on a branch-and-price-and-cut algorithm to solve the problem initially investigated by Irawan et al. (2017). Allal et al. (2021) developed a multi-agent-based model for maintenance routing optimization of offshore wind farms. A simulation-optimization approach is used to derive the best routing and maintenance de-

cisions minimizing cost under OWTs availability levels, weather conditions, and maintenance resources constraints. In a more recent work, Irawan et al. (2022) developed an integrated continuous location and maintenance routing optimization problem for an offshore wind farm where maintenance activities use a service operation vessel (SOV) and a safe transfer boat (STB) which serves to transfer technicians and equipment to the turbines. The optimization problem is to jointly determine the best locations for the SOV at sea, and to find the best routes of the STB to transfer the technicians to the turbines.

Maintenance and transport activities performed offshore face various types of uncertainty which can affect maintenance performance, force the adjustment of the maintenance decisions or even yield inappropriate decisions (Li et al., 2022). These uncertainties are related to statistical variability related to component reliability estimations, weather conditions, maintenance resources and spare part availability, etc. Such uncertainties have been accounted for in the offshore routing and maintenance scheduling problem. Irawan et al. (2021) developed a simulation-based optimization model to find the best routes for the CTV to maintain the OWTs in order to minimize the total cost under uncertainty. Their model include uncertainty related for example to travel, maintenance, and transfer times. Schrotenboer et al. (2020a) addressed maintenance and routing problem of offshore wind farms under uncertainty related to weather condition and maintenance tasks. A two-stage stochastic mixed integer programming model is proposed to determine an optimal maintenance plan performed by a maintenance provider subjected to minimum service requirements. Sample Average Approximation solution method is used to solve the resulting stochastic problem.

Gaps and contributions

Contrary to most papers dealing with the routing and scheduling of maintenance technicians (previous subsection), the proposed approach focuses simultaneously on the routing, the assignment of maintenance crews, and also on the detailed selection of maintenance actions, levels and components. This is a level of detail and decision making needed in real service settings.

To help maintenance decision makers resolve the SMP in offshore wind farms industry setting, the present work develops a new approach where the classical SMP is extended to include both repairpersons assignment and routing problems. The resulting problem is called the JSMOP where SM decisions and their assignment to repairpersons, in addition to repairpersons routing decisions are jointly accounted for. We are given an offshore wind farm composed of several OWTs in addition to an O&M base where the repair crews start and end their routes. Each OWT is composed of multiple components linked according to a given RBD. Each OWT component has a list of eligible maintenance actions ranging from mR, through intermediate IM actions, to replacement. Repairpersons are available to perform maintenance actions when necessary. OWTs must be maintained and operated with high reliability for long periods of time separated by scheduled maintenance rotations. To meet the pre-specified required performance level, maintenance activities are carried on the OWTs components during scheduled maintenance rotations. Because of the resource constraints related to travel time, cost, and availability of repair crews, not all turbines and their components are likely to be visited and maintained during maintenance trips. With the goal of minimizing the total cost while satisfying a minimum required reliability threshold during the next operating mission, the JSMOP as new variant of the SMP is developed to jointly make six decisions: (1) Select the set of OWTs to visit, (2) Select the set of components to maintain for each selected OWT, (3) Select the maintenance level to be performed on each selected component of each selected OWT, (4) Determine the number of repair crews to hire/utilize, (5) Assign a set of maintenance duties to each repair crew, and (6) Determine the routing for each repair crew. The resulting JSMOP is formulated as a mixed integer linear programming optimization model whose main components and characteristics are computed and fully discussed. To efficiently solve the JSMOP for large-scale instances within reasonable computation times, a CG method combining the Dantzig-Wolfe decomposition and labeling algorithm is proposed. Finally, the contributions of the present work can be summarized as follows:

— Develop a novel integrated SM optimization model where SM, repair crew assignment and routing decisions are jointly made. This contribution overcomes the main drawback of the sequential approaches where maintenance and routing decisions are made separately one after the other. Indeed, the results obtained from the experiments conducted show that the joint formulation

approach outperforms the sequential approach without adding significant computation burden for small and moderately large systems.
— Develop a novel optimization model where detailed reliability and IM modeling are integrated into the mathematical formulation. This contribution allows the proposed JSMOP to effectively address realistic and practical industrial scenarios. Furthermore, as it will be demonstrated, including detailed reliability and multiple IM options provides better joint maintenance and routing decisions.
— Develop an efficient CG solution method to deal with large-scale instances of the problem. This contribution will allow the proposed approach to be used by the decision makers in real service settings. As will be shown later, the proposed solution method is able to find near optimal solutions for large-scale wind farms in reasonable computation times.
— Provide managerial insights obtained from the extensive numerical experiments conducted. Managerial insights from both research and industrial point of views are presented in the form of contributions to the state-of-the-art of the SMP, guidelines to help the decision maker to perform what-if-analyses under maintenance resources availability, and the trade-offs between reliability requirements, maintenance duration and costs.

2.1.3 System Description and Reliability Computation

Before presenting the general setting of the proposed JSMOP modeling and optimization approach, we first give the system of notation used and the main working assumptions made in this section.

Notation

$\mathcal{W}$	Set of OWTs/locations, $\mathcal{W} = \{1, 2, ..., N\}$
$\mathcal{Q}$	Set of maintenance base (start and finish), $\mathcal{Q} = \{0, N+1\}$
$\mathcal{N}$	Set of nodes defined as $\mathcal{N} = \mathcal{W} \cup \mathcal{Q}$, with index i
$\mathcal{A}$	Set of undirected arcs $(i, i^{'})$ where $i, i^{'} \in \mathcal{N}$
$t_{ii'}(c_{ii'})$	Time (cost) to traverse arc $(i, i^{'})$
$\mathcal{R}$	Set of repair crews with index r
$\mathcal{S}_i$	Set of subsystems S_{ij} in OWT i, $\mathcal{S}_i = \{1, 2, ..., N_i\}$ with index j
N_{ij}	Number of components in subsystem j of OWT i
K_{ij}	Minimum number of components that must be functioning in subsystem j of OWT i
E_{ijk}	The kth component of subsystem j of OWT i
$A_{ijk}(B_{ijk})$	Effective age of component E_{ijk} after (before) maintenance
$v_{ijk}(u_{ijk})$	Operational status of component E_{ijk} after (before) maintenance
$t^c_{ijkl}(c^c_{ijkl})$	Duration (cost) of CM action of level l on component E_{ijk}
$t^p_{ijkl}(c^p_{ijkl})$	Duration (cost) of PM action of level l on component E_{ijk}
$\mathcal{P}_{ij}$	Set of patterns generated for subsystem j of OWT i, $\mathcal{P}_{ij} = \{1, \cdots, \mathcal{P}_{ij}\}$
T^{max}	Maximum working time (max maintenance trip duration)
$T_{ijp}(C_{ijp})$	Time (cost) to perform maintenance plan from pattern p for subsystem j in the ith OWT
c^d_i	Downtime cost rate for non-operating OWT i
t^d_i	Elapsed downtime of failed OWT i at the beginning of the maintenance trip; $t^d_i = 0$ if OWT i is working at the start of the trip
π^c_i (π^p_i)	Penalty cost per failed (functioning) OWT i that is not visited/maintained
R^c_{ijk}	Conditional reliability of component E_{ijk}
R_{ijp}	Reliability of subsystem j of OWT i when pattern p is selected
R_i	Reliability of OWT i
R_0	Required minimum OWT reliability level

Main working assumptions

1. Each OWT is comprised of multiple subsystems arranged in a series configuration. Each subsystem is made up of multiple binary components (meaning the components and OWT as a whole can either be functioning or failed). This is a reasonable and common assumption within the SMP literature (Diallo et al., 2018; Jiang and Liu, 2020a).
2. All OWTs are identical in terms of components technologies, configuration (RBD), and maintenance cost and time. This is a reasonable assumption but could be easily relaxed as the formulation is general enough and the maintenance generation pattern has been proven to handle all types of reliability structures and configurations (Diallo et al., 2019; O'Neil et al., 2022b).
3. Each maintenance action requires exactly one repair crew. This is reasonable as repair crew can have multiple repairpersons as needed as long as work is not hampered.
4. Each service vessel carries a single repair crew and their equipment and spare parts. Service vessels are required to be present while maintenance is being performed and have the necessary capacity to carry any required parts/tools. There is no waiting time when maintenance is completed at a location. In the case of large wind farms in deep waters or far from the O&M support port, an efficient combination of service operations vessel (SOV) and safe transfer boats (STB) can be used to transport crews (Irawan et al., 2022). In this case, the SOV becomes the equivalent of the maintenance base.
5. All crews have the skills to perform any required maintenance action. This is a reasonable and common assumption within the SMP literature (Diallo et al., 2018; Jiang and Liu, 2020a).
6. At the start of the maintenance trip/rotation, OWT component statuses as well as elapsed down times, and component statuses and ages are assumed known. This is also a reasonable and common assumption given that such data are mandatory to start the decision making process (Irawan et al., 2022; Stålhane et al., 2015; López-Santana et al., 2016).
7. Travel times are negligible compared to component lifetimes so that component statuses do not change between maintenance crew departure and arrivals at OWT locations. This is reasonable as travel time is in hours while turbines take many days or months to fail (Carroll et al., 2016).

System description

The system investigated is an offshore wind farm composed of N OWTs in addition to a maintenance base as illustrated in Figure 2.1. The system is viewed as a network modeled as an undirected graph $\mathcal{G} = (\mathcal{N}, \mathcal{A})$ where $\mathcal{N} = \mathcal{W} \cup \mathcal{Q}$ is the set of nodes, and $\mathcal{A}$ is the set of arcs. A node $i \in \mathcal{N}$ can refer either to an OWT or to the maintenance base where the repair crews must start and end their trips. Nodes numbered from 1 to N (*i.e.,* nodes in $\mathcal{W}$) refer to OWT locations, while nodes 0 and $N+1$ (*i.e.,* nodes in $\mathcal{Q}$) both refer to the same maintenance base location. For two nodes $i, i^{'} \in \mathcal{N}$, an arc $(i, i^{'}) \in \mathcal{A}$ is a link between i and $i^{'}$, respectively, the origin and the destination nodes. Each arc $(i, i^{'}) \in \mathcal{A}$ is characterized by the amount of time $t_{ii'}$ and cost $c_{ii'}$ incurred to traverse it.

The RBD of the ith OWT is composed of N_i GA(K_{ij},N_{ij}) ($j = 1, \ldots, N_i$) subsystems in series as depicted on Figure 2.2. The notation GA(k, n) refers to the k-out-of-n:G configuration and specifies that the system is functioning if at least k among the n components are functioning.

At the start of a given maintenance rotation/shift, there may be a number of failed OWTs, in addition to other components from other OWTs that have been scheduled for either PM or CM. At the start of the rotation, the exact age B_{ijk} as well as the status of every component E_{ijk} are assumed to be known. The status of component E_{ijk} before and after the maintenance visit is defined by two binary variables,

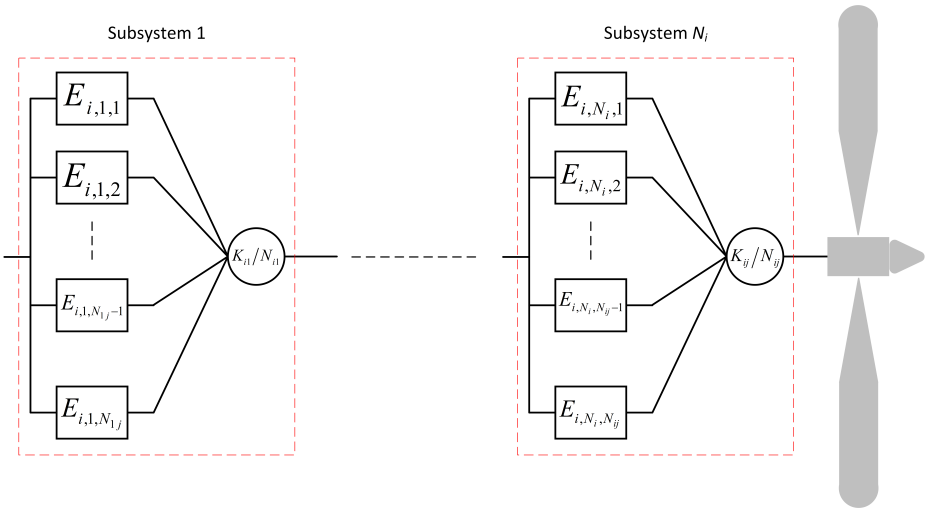


Figure 2.2 – Reliability structure of the OWT at location i

respectively, u_{ijk} and v_{ijk} as follows:

$$u_{ijk} = \begin{cases} 1 & \text{if component } E_{ijk} \text{ is functioning at the start of the rotation,} \\ 0 & \text{otherwise.} \end{cases} \tag{2.1}$$

$$v_{ijk} = \begin{cases} 1 & \text{if component } E_{ijk} \text{ is functioning after the end of the rotation,} \\ 0 & \text{otherwise.} \end{cases} \tag{2.2}$$

At the start of a given rotation, the operating state of the OWT i is also known and defined by a binary variable U_i as:

$$U_i = \begin{cases} 1 & \text{if OWT } i \text{ is functioning at the start of the rotation,} \\ 0 & \text{otherwise.} \end{cases} \tag{2.3}$$

OWT reliability computation

It is required that each OWT successfully operate the subsequent mission of duration D at a predetermined required minimum reliability level R_0. The mission refers to the time period between two consecutive scheduled maintenance rotations. In the present work, the reliability R_i of the ith OWT is defined by the probability of this OWT to operate during the next mission of duration D without failure. To compute this reliability, we first compute the respective reliabilities R^c_{ijk} and R_{ij} of component E_{ijk} and the subsystem S_{ij}. Without loss of generality, it is assumed that component E_{ijk} has Weibull distributed lifetimes with the shape and scale parameters β_{ijk} and η_{ijk}, respectively.

Given that component E_{ijk} has an age A_{ijk} at the end of the maintenance rotation (*i.e.,* at the start of the next mission), the probability that it will survive the next mission of duration D is defined by the conditional reliability function as:

$$R^c_{ijk} = \exp\left[\left(\frac{A_{ijk}}{\eta_{ijk}}\right)^{\beta_{ijk}} - \left(\frac{A_{ijk}+D}{\eta_{ijk}}\right)^{\beta_{ijk}}\right]. \tag{2.4}$$

The reliability R_{ij} of the jth subsystem of the ith OWT is obtained from the exact formulation proposed

in Arulmozhi (2002):

$$\mathcal{R}_{ij} = \sum_{e_{K_{ij}}=1}^{N_{ij}} \sum_{e_{K_{ij}-1}=1}^{e_{K_{ij}}-1} \cdots \sum_{e_1=1}^{e_2-1} \left(\prod_{v\in\{e_1,\cdots,e_{K_{ij}}\}} \mathcal{R}^c_{ijv} \right) \left(\prod_{\substack{u=1 \\ u\notin\{e_1,\cdots,e_{K_{ij}}\}}}^{e_{K_{ij}}} (1-\mathcal{R}^c_{iju}) \right). \tag{2.5}$$

To compute the reliability R_{ij} of each subsystem S_{ij}, the algorithm proposed by Kuo and Zuo (2003) and exploited in Diallo et al. (2018) is implemented. Because the subsystems for each OWT are assumed, without loss of generality, arranged in a serial configuration, the reliability of the ith OWT is then computed as:

$$R_i = \prod_{j\in\mathcal{S}_i} R_{ij}. \tag{2.6}$$

2.1.4 Maintenance Modeling, and Total Maintenance Time and Cost Computation

When an OWT i is visited by a repair crew, the components to maintain and the maintenance actions to be performed on the selected components are known. For each component E_{ijk}, there are $L_{ijk}+1$ maintenance levels $l \in \{0,\cdots,L_{ijk}\}$ that can be selected. These maintenance levels include DN, mR, IM, and replacement. The DN case ($l=0$) refers to no maintenance being performed on the component, the mR ($l=1$) case if selected will return a failed component to its age just before failure. The replacement level ($l=L_{ijk}$) resets the component's age to 0 (*i.e.,* component renewal), while an IM level $1<l<L_{ijk}$ if selected will return the component's age between that obtained after mR and replacement. It should be noted that only failed components are eligible to mR.

Commonly used IM models in the literature are: age reduction Malik (1979), hazard rate adjustment Nakagawa (1988a), and hybrid hazard rate Lin et al. (2000b) models. Without loss of generality, the age reduction approach is adopted here for IM modeling. According to this model, each IM level l available for component E_{ijk} is characterized by an age reduction coefficient θ_{ijkl} ($0 \le \theta_{ijkl} \le 1$). Therefore, when the IM is performed on E_{ijk}, its effective age B_{ijk} is reduced and becomes:

$$A_{ijk} = \theta_{ijkl} \times B_{ijk}. \tag{2.7}$$

There are durations and costs associated with PM and CM maintenance levels. When performed on the component E_{ijk}, a PM of level l requires t^p_{ijkl} units of time and costs c^p_{ijkl} monetary units. Similarly, a CM of level l when carried out on the same component consumes t^c_{ijkl} unit of times and costs c^c_{ijkl} monetary units.

The total time required to maintain the ith OWT and the corresponding total cost incurred are computed as:

$$T_i = \sum_{j=1}^{N_i} \sum_{k=1}^{N_{ij}} \left(\sum_{l=1}^{L_{ijk}} t^c_{ijkl} \cdot (1-u_{ijk}) \cdot \xi_{ijkl} + \sum_{l=2}^{L_{ijk}} t^p_{ijkl} \cdot u_{ijk} \cdot \xi_{ijkl} \right) \tag{2.8}$$

$$C_i = \sum_{j=1}^{N_i} \sum_{k=1}^{N_{ij}} \left(\sum_{l=1}^{L_{ijk}} c^c_{ijkl} \cdot (1-u_{ijk}) \cdot \xi_{ijkl} + \sum_{l=2}^{L_{ijk}} c^p_{ijkl} \cdot u_{ijk} \cdot \xi_{ijkl} \right) \tag{2.9}$$

where ξ_{ijkl} is a binary variable defined as:

$$\xi_{ijkl} = \begin{cases} 1 & \text{if maintenance level } l \text{ is selected to be performed} \\ & \text{on component } E_{ijk}, \\ 0 & \text{otherwise.} \end{cases} \tag{2.10}$$

For each OWT that has failed, downtime costs are incurred from the instant of OWT failure detection until this OWT is brought back to a functioning state through CM actions. In addition, when an OWT

undergoes PM, a downtime cost is incurred from the instant a repair crew arrives on-site and ends when the scheduled SM plan is achieved. If the failed OWT i (*i.e.,* $U_i = 0$ at the start of the rotation) is not visited during the maintenance rotation then a penalty cost π_i^c is incurred. This penalty cost is the sum of lost revenue caused by the OWTs downtime, the cost of finding and dispatching a new repair crew. An average penalty cost based on historical data was provided by the service provider. Similarly, if the functioning OWT i (*i.e.,* $U_i = 1$ at the start of the rotation) is not visited during the maintenance rotation then a penalty cost π_i^p is incurred.

The actual penalty cost can be difficult and complex to estimate. In the literature, several factors and estimation methods are considered. For example, the penalty cost represents the incentives to perform the maintenance tasks and translates the priority of visiting the wind turbines Irawan et al. (2017). This penalty cost is also said to be the revenue lost when a turbine must be derated (*i.e.,* set to operate at reduced performance) under condition-based maintenance monitoring. Furthermore, as pointed out in Irawan et al. (2017), the estimation of the penalty costs depends on the preferences and perspective of the decision maker (wind farm owner/operator or an O&M service provider), and the electricity price of wind energy.

2.1.5 Mathematical Formulations

It is required that OWTs operate at a predetermined reliability level for the upcoming mission of duration D, where D refers to the time until the next scheduled maintenance rotation. Thus, the maintenance actions selected for a given OWT must result in a probability of successfully completing the next mission that is equal or greater than the required minimum reliability level R_0. The goal of the proposed JSMOP is to identify the subset of OWTs that should be visited, the corresponding components to be maintained, the levels of maintenance to be carried out, in addition to the appropriate routing that should be taken for multiple service vessels transporting the repair crews to minimize the total cost. It should be noted that, given the limited time T^{max} allotted to the routing maintenance, the total maintenance time T_i, required to maintain the ith OWT, must be less than or equal to an upper bound T_{0i} such that:

$$T_{0i} = T^{max} - (t_{0i} + t_{i,N+1}), \tag{2.11}$$

where t_{0i} is the required transportation time to reach OWT i from the maintenance base, while $t_{i,N+1}$ is the time required to return back.

Maintenance pattern definition and generation algorithm

A maintenance pattern p is defined as a combination of components and related maintenance levels to be performed, which results in a discrete reliability value R_{ijp} for the jth GA(K_{ij},N_{ij}) subsystem of the ith OWT. The pattern generation algorithm has been demonstrated to be extremely powerful as it removes the need to solve a nonlinear objective function or nonlinear constraints Diallo et al. (2018). The overall principle of pattern generation is given by Algorithm 1. The pattern generation algorithm consists of the following main steps:

— Generate all patterns for a given subsystem;
— Select only valid patterns. A valid pattern is one for which the corresponding maintenance time and subsystem reliability satisfy the maximum duration and minimum reliability constraints, respectively;
— Return all valid maintenance patterns.

The list of patterns obtained from the algorithm can then be used as input data for the optimization phase. The output of the algorithm is a dataset containing i, j, p, R_{ijp}, T_{ijp}, and C_{ijp} for each pattern for each subsystem in each OWT.

To illustrate how the pattern generation works, a small GA$(1, 2)$ subsystem is considered with two levels of maintenance: DN ($l = 0$) and Replacement ($l = 1$). If at the start of the rotation maintenance both components are still functioning then the following list of four patterns will be generated: $(0, 0)$, $(1, 0)$, $(0, 1)$, and $(1, 1)$. Pattern $(1, 0)$ means that component 1 is replaced and no maintenance action is performed on component 2, while pattern $(1, 1)$ means that both components are replaced. If instead

Algorithm 1 Compute $R_{ijp}, C_{ijp}, T_{ijp}$ for all valid patterns for subsystem j in the OWT i

Input data: N, N_i, K_{ij}, N_{ij}, c^c_{ijkl}, t^c_{ijkl}, c^p_{ijkl}, t^p_{ijkl}, B_{ijk}, θ_{ijkl}, β_{ijk}, η_{ijk}, $\mathcal{R}_0$, T_{0i}
Initialize: $i = 1$
while $i \leq N$ **do**
  Initialize: $j = 1$
  while $j \leq N_i$ **do**
    – Generate an integer numbered set P_{ij} of all valid combination/patterns of components and their PM or CM levels such that at least K_{ij} components will be working after maintenance.
    – Find the cardinality P_{ij} of the set $\mathcal{P}_{ij}$: $P_{ij} = |\mathcal{P}_{ij}|$.
    Initialize: $p = 1$
    while $p \leq P_{ij}$ **do**
      – Calculate related maintenance duration T_{ijp} by summing up all the individual durations.
      if $T_{ijp} \leq T_{0i}$ **then**
        – Compute $\mathcal{R}^c_{ijk}$ the conditional reliability for all components E_{ijk} in the current pattern p using Equation (2.4).
        – Compute R_{ijp} the reliability of the subsystem j under the current pattern p using the algorithm developed in Kuo and Zuo (2003).
        if $R_{ijp} \geq R_0$ **then**
          – Calculate related maintenance cost C_{ijp} by summing up all the individual costs.
          – Store values of $i, j, p, R_{ijp}, C_{ijp}, T_{ijp}$.
        end if
      else
        – Remove current pattern p from the set $\mathcal{P}_{ij}$. (all patterns above p get shifted down by one position after the removal of p.)
        – Update $p = p - 1$ (to account for the removed pattern).
        – Update $P_{ij} = |\mathcal{P}_{ij}|$.
      end if
      $p = p + 1$.
    end while
    $j = j + 1$.
  end while
  $i = i + 1$.
end while

at the start of the shift maintenance both component are failed, only three patterns will be generated: $(1, 0)$, $(0, 1)$, and $(1, 1)$. Here, the pattern $(0, 0)$ is not eligible because both components are down and at least one of them must be replaced.

Once the maintenance patterns have been generated for each OWT i, the required maintenance time T_i and cost C_i to perform the corresponding SM plan can be obtained using Eqs. (2.8) and (2.9). They can also equivalently be derived as functions of patterns as follows:

$$T_i = \sum_{j \in \mathcal{S}_i} \sum_{p \in \mathcal{P}_{ij}} T_{ijp} \cdot s_{ijp} \tag{2.12}$$

$$C_i = \sum_{j \in \mathcal{S}_i} \sum_{p \in \mathcal{P}_{ij}} C_{ijp} \cdot s_{ijp}, \tag{2.13}$$

where s_{ijp} is a binary variable defined as:

$$s_{ijp} = \begin{cases} 1 & \text{if pattern } p \text{ is selected for subsystem } j \text{ of OWT } i, \\ 0 & \text{otherwise.} \end{cases} \tag{2.14}$$

JSMOP

The proposed JSMOP allows simultaneous decision-making of the optimal routes for multiple repair crews together with the optimal SM plan to be performed at the level of each visited OWT. In contrast to the classic OP that aims to maximize the reward collected at each vertex (node), the objective function of the proposed JSMOP formulation will be to minimize the total cost resulting from maintenance activities, repair crews' duties and transportation. In this section, fixed and variable labor costs of repair crews are accounted for in the maintenance cost. The resulting JSMOP provides joint optimal solutions

to three optimization problem, namely the SMP, the repair crews assignment problem, and the vessels routing problem. Let c_r^f and c_r^v represent the respective fixed cost and the variable cost rate of the repair crew $r \in \mathcal{R}$. To distinguish between repair crews' qualification levels, repair crew r is characterized by its speed ϵ_r ($\epsilon_r > 0$) to perform a maintenance action. For example, $\epsilon_r = 1$ means that the repair crew r carries out the maintenance action in the standard/baseline time duration, while $\epsilon_r = 0.5$ means that the repair crew is capable of performing the same maintenance action in half the standard duration. It is assumed that faster repair crews have higher costs.

To formulate the JSMOP, the following decision variables are introduced:

$$x_{ii'r} = \begin{cases} 1 & \text{if arc } (i, i^{'}) \text{ is traversed by repair crew } r, \\ 0 & \text{otherwise.} \end{cases} \tag{2.15}$$

$$z_{ijpr} = \begin{cases} 1 & \text{if pattern } p \text{ is selected for subsystem } j \text{ in OWT } i, \\ & \text{and performed by repair crew } r, \\ 0 & \text{otherwise.} \end{cases} \tag{2.16}$$

$$y_r = \begin{cases} 1 & \text{if repair crew } r \text{ is used,} \\ 0 & \text{otherwise} \end{cases} \tag{2.17}$$

In addition to the above decision variable, two other variables O_{ir} and I_{ir} are introduced. O_{ir} tracks the arrival time of repair crew r at the ith OWT. The variable I_{ir} refers to the standard decision variable used in the MTZ subtour elimination constraint, and identifies the order in which node i is visited in the travel path of repair crew r.

The mathematical JSMOP is then formulated as follows:

$$\begin{aligned}
\max \quad & \sum_{i\in\mathcal{W}} (\pi_i^c(1-U_i) + \pi_i^p \cdot U_i) \cdot \left(1 - \sum_{i'\in\mathcal{N}\setminus\{0\}} \sum_{r\in\mathcal{R}} x_{ii'r}\right) \\
& + \sum_{i\in\mathcal{W}} c_i^d \cdot (1-U_i) \cdot \left(t_i^d + \sum_{r\in\mathcal{R}} O_{ir} + \sum_{j\in\mathcal{S}_i} \sum_{p\in\mathcal{P}_{ij}} \sum_{r\in\mathcal{R}} T_{ijp} \cdot z_{ijpr} \cdot \epsilon_r\right) \\
& + \sum_{i\in\mathcal{W}} c_i^d \cdot U_i \cdot \left(\sum_{j\in\mathcal{S}_i} \sum_{p\in\mathcal{P}_{ij}} \sum_{r\in\mathcal{R}} T_{ijp} \cdot z_{ijpr} \cdot \epsilon_r\right) \\
& + \sum_{i\in\mathcal{W}} \sum_{j\in\mathcal{S}_i} \sum_{p\in\mathcal{P}_{ij}} \sum_{r\in\mathcal{R}} C_{ijp} \cdot z_{ijpr} \\
& + \sum_{i\in\mathcal{N}\setminus\{N+1\}} \sum_{i'\in\mathcal{N}\setminus\{0\}} \sum_{r\in\mathcal{R}} x_{ii'r} \cdot c_{ii'} + \sum_{r\in\mathcal{R}} \left(O_{N+1,r} \cdot c_r^v + y_r \cdot c_r^f\right)
\end{aligned} \tag{2.18}$$

s.t.:

$$\sum_{i\in\mathcal{N}\setminus\{0\}} x_{0ir} = 1, \quad \forall r \in \mathcal{R} \tag{2.19}$$

$$\sum_{i\in\mathcal{N}\setminus\{N+1\}} x_{i,N+1,r} = 1, \quad \forall r \in \mathcal{R} \tag{2.20}$$

$$\sum_{i\in\mathcal{N}\setminus\{N+1\}} x_{ii'r} = \sum_{i''\in\mathcal{N}\setminus\{0\}} x_{i'i''r}, \quad \forall r \in \mathcal{R}; \forall i^{'} \in \mathcal{W} \tag{2.21}$$

$$\sum_{i\in\mathcal{N}\setminus\{N+1\}} \sum_{r\in\mathcal{R}} x_{ii'r} \leq 1, \quad \forall i^{'} \in \mathcal{W} \tag{2.22}$$

$$\sum_{i \in \mathcal{N} \setminus \{N+1\}} x_{ii'r} \leq 1, \quad \forall i^{'} \in \mathcal{W}; \forall r \in \mathcal{R} \tag{2.23}$$

$$y_r \geq x_{ii'r}, \quad \forall i \in \mathcal{N} \setminus \{N+1\}; \forall i^{'} \in \mathcal{W}; \forall r \in \mathcal{R} \tag{2.24}$$

$$O_{ir} \geq 0, \quad \forall i \in \mathcal{N}; \forall r \in \mathcal{R} \tag{2.25}$$

$$O_{i'r} \leq T^{max} \cdot \sum_{i \in \mathcal{N} \setminus \{N+1\}} x_{ii'r}, \quad \forall i^{'} \in \mathcal{N} \setminus \{0\}; \forall r \in \mathcal{R} \tag{2.26}$$

$$O_{0r} = 0, \quad \forall r \in \mathcal{R} \tag{2.27}$$

$$O_{0r} + t_{0i'} - O_{i'r} - M_{0i'} \cdot (1 - x_{0i'r}) \leq 0, \quad \forall i^{'} \in \mathcal{N} \setminus \{0\}; \forall r \in \mathcal{R} \tag{2.28}$$

$$O_{ir} + t_{ii'} + \sum_{j \in \mathcal{S}_i} \sum_{p \in \mathcal{P}_{ij}} T_{ijp} \cdot z_{ijpr} \cdot \epsilon_r - O_{i'r} - M_{ii'} \cdot (1 - x_{ii'r}) \leq 0, \quad \forall i \in \mathcal{W}; \forall i^{'} \in \mathcal{N} \setminus \{0\}; \forall r \in \mathcal{R} \tag{2.29}$$

$$2 \leq I_{i'r} \leq |\mathcal{N}|, \quad \forall i^{'} \in \mathcal{N} \setminus \{0\}; \forall r \in \mathcal{R} \tag{2.30}$$

$$I_{ir} - I_{i'r} + 1 \leq (|\mathcal{N}| - 1) \cdot (1 - x_{ii'r}), \quad \forall i \in \mathcal{W}; \forall i^{'} \in \mathcal{N} \setminus \{0\}; \forall r \in \mathcal{R} \tag{2.31}$$

$$\sum_{p \in \mathcal{P}_{ij}} z_{ijpr} = \sum_{i' \in \mathcal{N} \setminus \{0\}} x_{ii'r}, \quad \forall i \in \mathcal{W}; \forall j \in \mathcal{S}_i; \forall r \in \mathcal{R} \tag{2.32}$$

$$\sum_{j \in \mathcal{S}_i} \sum_{p \in \mathcal{P}_{ij}} \ln(R_{ijp}) \cdot z_{ijpr} \geq \ln(\mathcal{R}_0), \quad \forall i \in \mathcal{W}; \forall r \in \mathcal{R} \tag{2.33}$$

$$x_{ii'r} \in \{0, 1\}, \quad \forall i, i^{'}, r \tag{2.34}$$

$$y_r \in \{0, 1\}, \quad \forall r \tag{2.35}$$

$$z_{ijpr} \in \{0, 1\}, \quad \forall i, j, p, r \tag{2.36}$$

$$I_{ir} \geq 0, O_{ir} \geq 0 \quad \forall i, r \tag{2.37}$$

The objective function (2.18) minimizes the total cost of performing maintenance on the distributed systems. The first term represents the penalty cost incurred when maintenance is not performed on failed or functioning turbines. The second term represents the increasing penalty for failed turbines. This penalty cost runs from the instant when the OWT failure is detected until it is put back into operation. The third term is the downtime cost associated with performing PM. The fourth term computes the cost of performing the SM plan, while the fifth term is the transportation cost. The last term of the objective function refers to the summation of the fixed cost and variable labor cost for hiring and using the repair crews. Constraints (2.19) and (2.20) ensure that the route for each service vessel starts and ends at the maintenance base. Constraints (2.21) ensure the connectivity of the route, and constraints (2.22) ensure that each vertex is visited at most once. Constraints (2.23) ensure that each repair crew visits a node at most once. Constraints (2.24) ensure that repair crews are available before they can be assigned to maintenance routes. Constraints (2.25) state that arrival times at OWTs are nonnegative. Constraints (2.26) ensure that crews arrive at OWTs on their paths before the maintenance rotation time T^{max} has elapsed. Constraints (2.27) state that repair crews are at the maintenance base ($i = 0$) at time 0. Constraints (2.28) define the arrival times at their first OWT when repair crews leave the maintenance base and Constraints (2.29) define the arrival times at the other OWTs. Constraints (2.30) – (2.31) prevent subtours. $M_{ii'}$ are large positive numbers (*i.e.,* big M values). Constraints (2.32) ensure that a single maintenance pattern is selected for every subsystem of each OWT, however no maintenance pattern will be selected for OWTs that are not visited. Constraints (2.33) ensure that, after maintenance actions are performed, the reliability of visited OWTs must be greater than or equal to the required minimum reliability level. The term on the left-hand side of the inequality is the linearization of the OWTs reliability as given by Eq. (2.6). Because z_{ijpr} is a binary variable, Eq. (2.6) becomes:

$$R_i = \prod_{j \in \mathcal{S}_i} \left(\sum_{p \in \mathcal{P}_{ij}} \sum_{r \in \mathcal{R}} R_{ijp} \cdot z_{ijpr} \right) \tag{2.38}$$

After applying the natural logarithm, which is a monotonic function, to both sides of the above

equation, the result in constraints (2.33) is obtained:

$$\ln(R_i) = \sum_{j \in \mathcal{S}_i} \sum_{p \in \mathcal{P}_{ij}} \sum_{r \in \mathcal{R}} \ln(R_{ijp}) \cdot z_{ijpr} \tag{2.39}$$

Constraints (2.34) – (2.37) define the binary decision variables $x_{ii'r}$, y_r and z_{ijpr} used in the formulation.

2.1.6 CG Approach for Large-Scale JSMOP

This section presents a CG-based solution approach for solving the JSMOP for large-scale wind farms. This approach relies on Dantzig-Wolfe decomposition. Accordingly, the original JSMOP is reformulated as a master problem and multiple subproblems that are iteratively solved by means of CG.

Let constraints (2.19)-(2.21) and (2.23)-(2.37) define the feasible region of the SPs and let $\mathcal{D}^r = \{(x^r, y^r, z^r, O^r, I^r) | (x^r, y^r, z^r, O^r, I^r)$ satisfies $(2.19)-(2.21)$ and $(2.23)-(2.37)\}$ be the feasible bounded domain of the SP associated with each repair crew $r \in \mathcal{R}$. Let $\mathcal{P}^r$ be the set of extreme points of $\mathcal{D}^r$. Each extreme point $p \in \mathcal{P}^r$ represents a feasible route and maintenance patterns to be performed on visited turbines for repair crew r.

Before defining the MP, the objective function in Eq. (2.18) is first exposed as an equivalent maximization problem by eliminating constant terms (*i.e.,* terms independent of $r \in \mathcal{R}$). This is a crucial step in applying the decomposition technique. The new objective function can then be written as follows:

$$\begin{aligned}
&\sum_{i \in \mathcal{W}} (\pi_i^c(1-U_i) + \pi_i^p \cdot U_i) + \sum_{i \in \mathcal{W}} c_i^d \cdot (1-U_i) \cdot t_i^d - \\
&\max \Bigg\{ \sum_{i \in \mathcal{W}} (\pi_i^c(1-U_i) + \pi_i^p \cdot U_i) \cdot \sum_{i' \in \mathcal{N} \setminus \{0\}} \sum_{r \in \mathcal{R}} x_{ii'r} - \\
&\quad \sum_{i \in \mathcal{W}} c_i^d \cdot (1-U_i) \cdot \left(\sum_{r \in \mathcal{R}} O_{ir} + \sum_{j \in \mathcal{S}_i} \sum_{p \in \mathcal{P}_{ij}} \sum_{r \in \mathcal{R}} T_{ijp} \cdot z_{ijpr} \cdot \epsilon_r \right) - \\
&\quad \sum_{i \in \mathcal{W}} c_i^d \cdot U_i \cdot \left(\sum_{j \in \mathcal{S}_i} \sum_{p \in \mathcal{P}_{ij}} \sum_{r \in \mathcal{R}} T_{ijp} \cdot z_{ijpr} \cdot \epsilon_r \right) - \\
&\quad \sum_{i \in \mathcal{W}} \sum_{j \in \mathcal{S}_i} \sum_{p \in \mathcal{P}_{ij}} \sum_{r \in \mathcal{R}} C_{ijp} \cdot z_{ijpr} - \\
&\quad \sum_{i \in \mathcal{N} \setminus \{N+1\}} \sum_{i' \in \mathcal{N} \setminus \{0\}} \sum_{r \in \mathcal{R}} x_{ii'r} \cdot c_{ii'} - \\
&\quad \sum_{r \in \mathcal{R}} \left(O_{N+1,r} \cdot c_r^v + y_r \cdot c_r^f \right) \Bigg\}
\end{aligned} \tag{2.40}$$

Thus, the objective function reduces to maximizing the total reward of performing maintenance on the distributed wind turbines. The reward is a function of penalty costs collected as well as downtime, maintenance, travel, and repair crew expenses. To formulate the MP, additional parameters τ_{ipr}, and ρ_{pr}, and a decision variable λ_{pr} are introduced. The parameter ρ_{pr} is defined as the reward of solution $p \in \mathcal{P}^r$, while τ_{ipr} and λ_{pr} are defined as:

$$\tau_{ipr} = \begin{cases} 1 & \text{if solution } p \in \mathcal{P}^r \text{ includes node } i, \\ 0 & \text{otherwise.} \end{cases} \tag{2.41}$$

$$\lambda_{pr} = \begin{cases} 1 & \text{if solution } p \in \mathcal{P}^r \text{ is selected,} \\ 0 & \text{otherwise} \end{cases} \tag{2.42}$$

The master problem can now be formulated as follows:

$$\max \quad \sum_{r \in \mathcal{R}} \sum_{p \in \mathcal{P}^{r'}} \lambda_{pr} \cdot \rho_{pr} \tag{2.43}$$

s.t:

$$\sum_{r \in \mathcal{R}} \sum_{p \in \mathcal{P}^{r'}} \lambda_{pr} \cdot \tau_{ipr} \leq 1, \quad \forall i \in \mathcal{I} \qquad (\mu_i) \tag{2.44}$$

$$\sum_{p \in \mathcal{P}^{r'}} \lambda_{rp} = 1, \quad \forall r \in \mathcal{R} \qquad (\theta_r) \tag{2.45}$$

$$\lambda_{rp} \geq 0, \quad \forall r \in \mathcal{R}; \forall p \in \mathcal{P}^{r'} \tag{2.46}$$

The objective function (Eq. 2.43) seeks to maximize the total reward. Constraints (2.44) ensure that each turbine is visited at most once and constraints (2.45) enforce a route to be selected for each of the repair crews. The MP formulation is initialized with only a small subset ($\mathcal{P}^{r'} \subset \mathcal{P}^r$) of feasible columns. By solving the MP, the dual variables (μ_i, θ_r) can be extracted and used in the pricing SPs to identify whether there are any columns (*i.e.,* routes and maintenance patterns) that should be added to the set $\mathcal{P}^{r'}$.

In a CG framework, the SPs are used to identify the route and maintenance patterns that maximize the reduced cost. The formulation of the SPs rely on the same variables $x_{ii'r}$, y_r, z_{ijpr}, O_{ir}, and I_{ir} that have been presented in Section (2.1.5) where the index r is fixed to indicate the repair crew that is being considered. The formulation is as follows:

$$\begin{aligned} \max \quad & \sum_{i \in \mathcal{W}} (\pi_i^c (1 - U_i) + \pi_i^p \cdot U_i) \cdot \sum_{i' \in \mathcal{N} \setminus \{0\}} x_{ii'r} \\ & - \sum_{i \in \mathcal{W}} c_i^d \cdot (1 - U_i) \cdot \left(O_{ir} + \sum_{j \in \mathcal{S}_i} \sum_{p \in \mathcal{P}_{ij}} T_{ijp} \cdot z_{ijpr} \cdot \epsilon_r \right) \\ & - \sum_{i \in \mathcal{W}} c_i^d \cdot U_i \cdot \left(\sum_{j \in \mathcal{S}_i} \sum_{p \in \mathcal{P}_{ij}} T_{ijp} \cdot z_{ijpr} \cdot \epsilon_r \right) \\ & - \sum_{i \in \mathcal{W}} \sum_{j \in \mathcal{S}_i} \sum_{p \in \mathcal{P}_{ij}} C_{ijp} \cdot z_{ijpr} - \sum_{i \in \mathcal{N} \setminus \{N+1\}} \sum_{i' \in \mathcal{N} \setminus \{0\}} x_{ii'r} \cdot c_{ii'} \\ & - \left(O_{N+1,r} \cdot c_r^v + y_r \cdot c_r^f \right) - \sum_{i \in \mathcal{W}} \sum_{i' \in \mathcal{N} \setminus \{0\}} \pi_i \cdot x_{ii'} - \theta_r \end{aligned} \tag{2.47}$$

s.t:

$$\sum_{i \in \mathcal{N} \setminus \{0\}} x_{0ir} = 1 \tag{2.48}$$

$$\sum_{i \in \mathcal{N} \setminus \{N+1\}} x_{i,N+1,r} = 1 \tag{2.49}$$

$$\sum_{i \in \mathcal{N} \setminus \{N+1\}} x_{ii'r} = \sum_{i'' \in \mathcal{N} \setminus \{0\}} x_{i'i''r}, \quad \forall i' \in \mathcal{W} \tag{2.50}$$

$$\sum_{i \in \mathcal{N} \setminus \{N+1\}} x_{ii'r} \leq 1, \quad \forall i' \in \mathcal{W} \tag{2.51}$$

$$y_r \geq x_{ii'r}, \quad \forall i \in \mathcal{N} \setminus \{N+1\}; \forall i' \in \mathcal{W} \tag{2.52}$$

$$O_{ir} \geq 0, \quad \forall i \in \mathcal{N} \tag{2.53}$$

$$O_{i'r} \leq T^{max} \cdot \sum_{i \in \mathcal{N} \setminus \{N+1\}} x_{ii'r}, \quad \forall i' \in \mathcal{N} \setminus \{0\} \tag{2.54}$$

$$O_{0r} = 0 \tag{2.55}$$

$$O_{0r} + t_{0i'} - O_{i'r} - M_{0i'} \cdot (1 - x_{0i'r}) \leq 0, \quad \forall i' \in \mathcal{N} \setminus \{0\} \tag{2.56}$$

$$O_{ir} + t_{ii'} + \sum_{j \in \mathcal{S}_i} \sum_{p \in \mathcal{P}_{ij}} T_{ijp} \cdot z_{ijpr} \cdot \epsilon_r - O_{i'r} - M_{ii'} \cdot (1 - x_{ii'r}) \leq 0, \quad \forall i \in \mathcal{W}; \forall i' \in \mathcal{N} \setminus \{0\} \tag{2.57}$$

$$2 \leq I_{i'r} \leq |\mathcal{N}|, \quad \forall i' \in \mathcal{N} \setminus \{0\} \tag{2.58}$$

$$I_{ir} - I_{i'r} + 1 \leq (|\mathcal{N}| - 1) \cdot (1 - x_{ii'r}), \quad \forall i \in \mathcal{W}; \forall i' \in \mathcal{N} \setminus \{0\} \tag{2.59}$$

$$\sum_{p \in \mathcal{P}_{ij}} z_{ijpr} = \sum_{i' \in \mathcal{N} \setminus \{0\}} x_{ii'r}, \quad \forall i \in \mathcal{W}; \forall j \in \mathcal{S}_i \tag{2.60}$$

$$\sum_{j \in \mathcal{S}_i} \sum_{p \in \mathcal{P}_{ij}} \ln(R_{ijp}) \cdot z_{ijpr} \geq \ln(R_0), \quad \forall i \in \mathcal{W} \tag{2.61}$$

$$x_{ii'r} \in \{0,1\}, \quad \forall i, i' \tag{2.62}$$

$$y_r \in \{0,1\} \tag{2.63}$$

$$z_{ijpr} \in \{0,1\}, \quad \forall i, j, p \tag{2.64}$$

$$I_{ir} \geq 0, O_{ir} \geq 0, \quad \forall i, r \tag{2.65}$$

Solving the subproblems

The SPs are complex optimization problems that would still be hard to solve for large instances using basic commercial solvers. Thus, in the present work an efficient multi-tier approach is developed to generate new columns. The approach relies on the observation that it is better to visit as many turbines as possible with all the other constraints satisfied. Accordingly, when the penalty costs (*i.e.,* π_i^p and π_i^c) are significantly higher than travel, maintenance, and downtime costs, the approach can lead to near optimal solutions as will be shown in the numerical experiments.

The first step of the solution approach consists in solving the SMP for each turbine with the objective of identifying the maintenance actions that minimize intervention time. Indeed, selecting the maintenance actions that minimize maintenance time will allow the maximum number of turbines to be visited within the working time T^{max}. For each OWT i, the corresponding BIP formulation for the minimization of the overall maintenance time is as follows:

$$\min \sum_{j \in \mathcal{S}_i} \sum_{p \in \mathcal{P}_{ij}} T_{ijp} \cdot s_{ijp} \tag{2.66}$$

s.t:

$$\sum_{p \in \mathcal{P}_{ij}} s_{ijp} = 1, \quad \forall j \in \mathcal{S}_i \tag{2.67}$$

$$\sum_{j \in \mathcal{S}_i} \sum_{p \in \mathcal{P}_{ij}} \ln(R_{ijp}) \cdot s_{ijp} \geq \ln(R_0) \tag{2.68}$$

$$s_{ijp} \in \{0,1\}, \quad \forall j, p \tag{2.69}$$

In the above optimization model, constraints (2.67) ensure that a single maintenance pattern is selected for each subsystem j of OWT i. Constraint (2.68) ensures that the reliability of OWT i must be greater than or equal to the target reliability. Once the above SMP model has been solved for each OWT i, the

resulting maintenance time (T_i and cost C_i) can be computed from Equations (2.12) and (2.13), and then used to identify the optimal route for each repair crew r as it will be discussed in the following section.

To solve the routing portion of the problem and identify promising paths, the forward labeling algorithm used in Stålhane et al. (2015) is adapted and shown in Algorithm 2. In this algorithm, a partial path is represented by a 5-tuple $(\eta, rc, t, \Omega, \phi)$ called a label. Roughly speaking, a label $\zeta = (\eta(\zeta), rc(\zeta), t(\zeta), \Omega(\zeta), \phi(\zeta))$ where $\eta(\zeta)$, $rc(\zeta)$, $t(\zeta)$, $\Omega(\zeta)$, and $\phi(\zeta)$ refer, respectively, to the last visited node, the accumulated reduced cost, the accumulated time, the set of turbines and the number of failed turbines visited on the partial path. The first forward label is given by $(0, 0, 0, \emptyset, 0)$, a route that is just visiting the maintenance base. Let $\mathcal{P}_\zeta$ and $\bar{\mathcal{P}}_\zeta$ be the set of processed and unprocessed labels, respectively. Initially, $\bar{\mathcal{P}}_\zeta$ contains only the first forward label while $\mathcal{P}_\zeta$ is empty. While the set $\bar{\mathcal{P}}_\zeta$ is non-empty, a label ζ is selected using the $GetNextLabel()$ function and extended along all feasible arcs $(\eta(\zeta), j) \in \mathcal{A}$. An extension $(\eta(\zeta), j)$ is feasible if the following constraints hold:

$$j \notin \Omega(\zeta) \tag{2.70}$$
$$\eta(\zeta) \neq N + 1 \tag{2.71}$$
$$T^{max} \geq t(\zeta) + t_{\eta(\zeta),j} + T_j \tag{2.72}$$

In the above equations, constraints (2.70) and (2.71) ensure that each vertex is visited at most once, and enforce routes to finish at the maintenance base. Constraint (2.72) ensures that the maximum working time is never exceeded. Recall that $t_{\eta(\zeta),j}$ is the travel time from node $\eta(\zeta)$ to node j. If an extension is found to be feasible, a new label $\zeta' = (\eta(\zeta'), rc(\zeta'), t(\zeta'), \Omega(\zeta'), \phi(\zeta'))$ is created and its attributes are computed as follows:

$$\eta(\zeta') = j. \tag{2.73}$$

$$rc(\zeta') = \begin{cases} rc(\zeta) + \left(\pi_j^c(1 - U_j) + \pi_j^p \cdot U_j\right) \\ \qquad - c_j^d \cdot (1 - U_j) \cdot \left(t(\zeta) + t_{\eta(\zeta),j} + T_j \cdot \epsilon_r\right) \\ \qquad - c_j^d \cdot U_j \cdot (T_j \cdot \epsilon_r) - c_r^v \cdot \left(t_{\eta(\zeta),j} + T_j\right) - c_r^f \\ \qquad - C_j - c_{\eta(\zeta),j} - \pi_j - \theta_r \\ \qquad\qquad \text{if } \eta(\zeta) = 0 \text{ and } j \neq N + 1, \\ rc(\zeta) + \left(\pi_j^c(1 - U_j) + \pi_j^p \cdot U_j\right) \\ \qquad - c_j^d \cdot (1 - U_j) \cdot \left(t(\zeta) + t_{\eta(\zeta),j} + T_j \cdot \epsilon_r\right) \\ \qquad - c_j^d \cdot U_j \cdot (T_j \cdot \epsilon_r) - c_r^v \cdot \left(t_{\eta(\zeta),j} + T_j\right) \\ \qquad - C_j - c_{\eta(\zeta),j} - \pi_j - \theta_r \\ \qquad\qquad \text{if } \eta(\zeta) \neq 0 \text{ and } j \neq N + 1, \\ rc(\zeta) - c_r^v \cdot t_{\eta(\zeta),j} \qquad \text{if } j = N + 1. \end{cases} \tag{2.74}$$

$$t(\zeta') = \begin{cases} t(\zeta) + t_{\eta(\zeta),j}, & \text{if } j = N + 1\ , \\ t(\zeta) + t_{\eta(\zeta),j} + T_j, & \text{otherwise.} \end{cases} \tag{2.75}$$

$$\Omega(\zeta') = \Omega(\zeta) \cup \{j\}. \tag{2.76}$$

$$\phi(\zeta') = \phi(\zeta) + (1 - U_j). \tag{2.77}$$

The new label ζ' is compared to already existing labels based on a dominance criterion which is used to eliminate paths that will likely not appear in the optimal solution, and to avoid the full enumeration of

all feasible paths. In this work, we let label ζ_1 dominates label ζ_2 if and only if:

$$\eta(\zeta_1) = \eta(\zeta_2) \tag{2.78}$$
$$|\Omega(\zeta_1)| = |\Omega(\zeta_2)| \tag{2.79}$$
$$\phi(v_1) = \phi(\zeta_2) \tag{2.80}$$
$$t(\zeta_1) \leq t(\zeta_2) \tag{2.81}$$
$$rc(\zeta_1) \geq rc(\zeta_2) \tag{2.82}$$

The first three parts of the dominance criterion state that the two paths must have the same last node, have visited the same number of OWTs, and have visited the same number of failed OWTs. Condition (2.81) states that the accumulated time of ζ_1 must be less than or equal to the accumulated time of ζ_2. Finally, condition (2.82) states that the accumulated reduced cost of ζ_1 must be greater than or equal to the accumulated reduced cost of ζ_2. The output of Algorithm 2 is the label ζ that contains the path $\Omega(\zeta)$ with the largest reduced cost denoted as $\mathcal{W}_\zeta$ while its corresponding total maintenance time and total travel cost are denoted as T_ζ and C_ζ, respectively. This label ζ, $\mathcal{W}_\zeta$, T_ζ and C_ζ are passed to the final stage of the approach where the maintenance decisions are optimized.

Algorithm 2 Route Generator

Input data: Undirected graph $\mathcal{G} = (\mathcal{N}, \mathcal{A})$
Initialize: $\bar{\mathcal{P}}_\zeta = \{(0, 0, 0, \emptyset, 0)\}$, $\mathcal{P}_\zeta = \emptyset$
while $\bar{\mathcal{P}}_\zeta \neq \emptyset$ **do**
  $\zeta = GetNextLabel(\bar{\mathcal{P}}_\zeta)$
  for each feasible extension of $\zeta \rightarrow \zeta'$ **do**
    if no label in $\mathcal{P}_\zeta$ dominates ζ' **then**
      $\mathcal{P}_\zeta = \mathcal{P}_\zeta \cup \{\zeta'\}$.
      $\bar{\mathcal{P}}_\zeta = \bar{\mathcal{P}}_\zeta \cup \{\zeta'\}$.
      Remove labels in $\bar{\mathcal{P}}_\zeta$ and $\bar{\mathcal{P}}_\zeta$ dominated by ζ'.
    end if
  end for
end while
Initialize: $\zeta = (0, 0, 0, \emptyset, 0)$, $rc_0 = 0$.
for $p \in \bar{\mathcal{P}}_\zeta$ **do**
  if $\eta(p) = N + 1$ **then**
    if $rc(p) \geq rc_0$ **then**
      $\zeta = p$.
      $rc_0 = rc(p)$.
    end if
  end if
end for
return ζ

The final stage of the SP solution approach optimizes the maintenance decisions. Based on the path $\mathcal{W}_\zeta$ found through Algorithm 2 for repair crew r, this stage of the approach aims to identify a subset of maintenance actions that will further improve the reduced cost. It should be noted that when the path is fixed, the variables $x_{ii'r}$ and I_{ir} become parameters, whereas z_{ijpr} and O_{ir} remain the only decision variables to solve. Let $\mathcal{N}_\zeta$ be the set of nodes including the maintenance base, *i.e.,* $\mathcal{N}_\zeta = \mathcal{W}_\zeta \cup \mathcal{Q}$. The SP optimization model can then be formulated as follows:

$$
\begin{aligned}
\max \quad & \sum_{i \in \mathcal{W}_\zeta} (\pi_i^c (1 - U_i) + \pi_i^p \cdot U_i) \\
& - \sum_{i \in \mathcal{W}_\zeta} c_i^d \cdot (1 - U_i) \cdot \left(O_{ir} + \sum_{j \in \mathcal{S}_i} \sum_{p \in \mathcal{P}_{ij}} T_{ijp} \cdot z_{ijpr} \cdot \epsilon_r \right) \\
& - \sum_{i \in \mathcal{W}_\zeta} c_i^d \cdot U_i \cdot \left(\sum_{j \in \mathcal{S}_i} \sum_{p \in \mathcal{P}_{ij}} T_{ijp} \cdot z_{ijpr} \cdot \epsilon_r \right) \\
& - \sum_{i \in \mathcal{W}_\zeta} \sum_{j \in \mathcal{S}_i} \sum_{p \in \mathcal{P}_{ij}} C_{ijp} \cdot z_{ijpr} - C_\zeta \\
& - \left(O_{N+1,r} \cdot c_r^v + c_r^f \right) - \sum_{i \in \mathcal{W}_\zeta} \mu_i - \theta_r \qquad (2.83)
\end{aligned}
$$

s.t:

$$0 \leq O_{ir} \leq T^{max}, \quad \forall i \in \mathcal{N}_\zeta \qquad (2.84)$$

$$O_{0r} = 0 \qquad (2.85)$$

$$O_{0r} + t_{0i'} - O_{i'r} - M_{0i'} \cdot (1 - x_{0i'r}) \leq 0, \quad \forall i^{'} \in \mathcal{W}_\zeta \qquad (2.86)$$

$$
\begin{aligned}
& O_{ir} + t_{ii'} + \sum_{j \in \mathcal{S}_i} \sum_{p \in \mathcal{P}_{ij}} T_{ijp} \cdot z_{ijpr} \cdot \epsilon_r \qquad (2.87) \\
& \quad - O_{i'r} - M_{ii'} \cdot (1 - x_{ii'r}) \leq 0, \quad \forall i \in \mathcal{W}_\zeta; \forall i^{'} \in \mathcal{W}_\zeta
\end{aligned}
$$

$$\sum_{p \in \mathcal{P}_{ij}} z_{ijpr} = 1, \quad \forall i \in \mathcal{W}_\zeta; \forall j \in \mathcal{S}_i \qquad (2.88)$$

$$\sum_{j \in \mathcal{S}_i} \sum_{p \in \mathcal{P}_{ij}} \ln(R_{ijp}) \cdot z_{ijpr} \geq \ln(R_0), \quad \forall i \in \mathcal{W}_\zeta \qquad (2.89)$$

$$z_{ijpr} \in \{0, 1\}, \quad \forall i, j, p \qquad (2.90)$$

$$O_{ir} \geq 0 \quad \forall i \qquad (2.91)$$

In the above optimization model, the index r is fixed to indicate the repair crew that is being considered. The objective function (2.83) maximizes the reduced cost. Constraints (2.84)–(2.87) update the arrival time at each visited node. Constraints (2.88) ensure that a maintenance pattern is selected for each visited OWT. Constraints (2.89) guarantee that, after maintenance actions are performed, the reliability of visited OWTs is greater than or equal to the required minimum reliability level. Constraints (2.90) and (2.91) define the decision variables z_{ijpr} and O_{ir}, respectively.

2.1.7 Numerical Experiments

In this section, eight sets of numerical experiments are conducted using the case and data from a maintenance firm providing services to an offshore wind farm. The first set is a validation experiment to ensure the proposed extended model obtains the same results as previously published articles. The second experiment demonstrates the benefit of selecting an appropriately scaled value for the big M values used in the formulation in constraints (2.27) and (2.28). The third experiment compares the proposed joint model and the sequential model and the fourth experiment demonstrates the importance of considering multiple maintenance levels. The fifth and sixth experiments demonstrate additional features and advantages of the JSMOP optimization model. Finally, the seventh and eighth experiments demonstrate the ability of the proposed solution method to deal with JSMOP of large and practical size. Please refer to Appendix A.1 for a detailed presentation of the sequential SM and OP approach.

All experiments are run on a Intel™i5 2.9GHz desktop computer with 12GB of RAM running Windows 10™. All algorithms were coded in Python 3.8. The optimization runs were carried out by Gurobi 9.1 using gurobipy.

Case study

Experiments #2 to #6 are carried out using an offshore wind farm composed of $N = 7$ identical OWTs as illustrated in Figure 2.1. As depicted in Figure 2.3, the RBD of OWT i ($i = 1, \cdots, 7$) is composed of $S_i = 4$ k-out-of-n:G subsystems GA$(2, 5)$, GA$(2, 4)$, GA$(1, 3)$ and GA$(1, 2)$ in series. The shape and scale parameters, and operational states of each OWT components are given in the data table in Appendix A.2 (Tables A.1 and A.2). Table A.1 summarizes the reliability information of components and lists also the cost and duration for each maintenance level. Table A.2 gives for each component its corresponding status and age. According to Table A.2, OWTs in locations 3, 5 and 7 are failed at the start of maintenance rotation. To carry out maintenance activities, there are $R = 3$ identical repair crews with fixed and variable costs set to \$150 and \$25 respectively. The required travel times and costs between OWT locations and the maintenance base are respectively given in Tables 2.1 and 2.2. The values of elapsed down times t_i^d are provided in Table 2.3. The other parameters are: $\pi_i^c = \$12,500$, $\pi_i^p = \$10,000$, and $c_i^d = \$162$/hour. The duration D of the next mission (*i.e.,* the period that OWTs will operate until the next maintenance rotation) is of six months (*i.e.,* $D = 6$ months). Typically, each vessel/safe transfer boat (STB) capacity ranges from 5 to 10 maintenance technicians with crews of two to five repairpersons. These data ranges are in line with that used by Irawan et al. (2017) and Irawan et al. (2022). The problem as described deals with routine corrective and preventive maintenance actions that do not require special or heavy load lifting equipment (similar to the minor maintenance actions considered by Irawan et al. (2022)). Maintenance technicians will generally test, inspect, swap modules, conduct non-heavy repairs, non-major welding, alignment, etc. These levels of maintenance actions are normally performed on components from subsystems such as gearbox, yaw drive/motor, air brake, power electronic switch, oil heater/cooler.

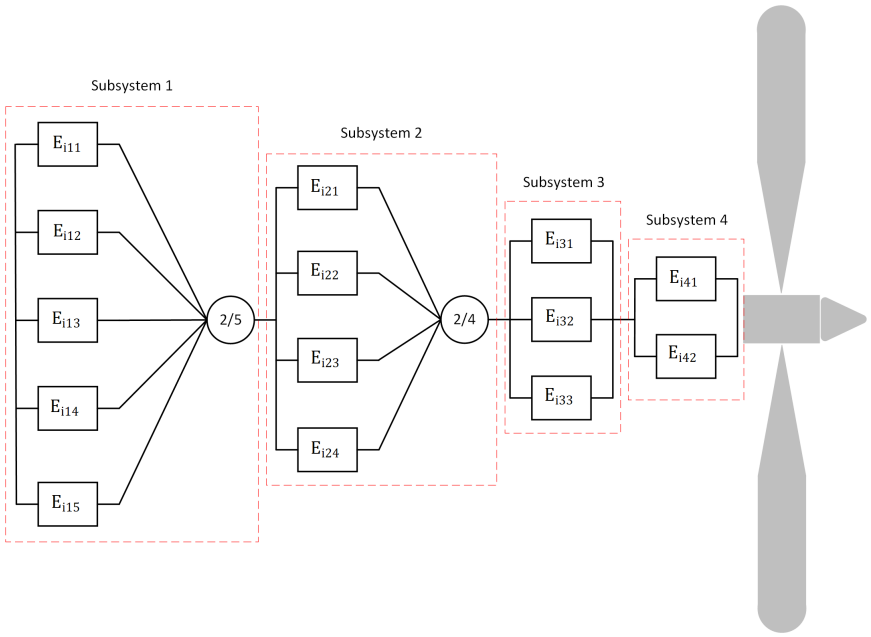


Figure 2.3 – Reliability structure of the OWT at location i

For all numerical experiments, four potential maintenance levels for failed components will be considered: DN ($l = 0$), mR ($l = 1$), IM ($l = 2$) and replacement ($l = 3$). The same maintenance levels will be considered for functioning components with the exception of mR. It will be assumed that the IM level will reduce the age of the component by half ($\theta_{ijk2} = 0.5$). Except for Experiment #6, all the repair crews are assumed to be of standard qualification with $\epsilon_r = 1$, $\forall r \in \mathcal{R}$.

Table 2.1 – Traveling time (hours) between turbines and maintenance base

	End vertex							
Start vertex	0	1	2	3	4	5	6	7
0	0	0.47	0.61	0.54	0.69	0.59	0.76	0.71
1	0.47	0	0.15	0.12	0.25	0.12	0.27	0.25
2	0.61	0.15	0	0.20	0.17	0.10	0.15	0.12
3	0.54	0.12	0.20	0	0.34	0.10	0.29	0.32
4	0.69	0.25	0.17	0.34	0	0.27	0.25	0.12
5	0.59	0.12	0.10	0.10	0.27	0	0.20	0.22
6	0.76	0.27	0.15	0.29	0.25	0.20	0	0.12
7	0.71	0.25	0.12	0.32	0.12	0.22	0.12	0

Table 2.2 – Traveling cost ($) between turbines and maintenance base

	End vertex							
Start vertex	0	1	2	3	4	5	6	7
0	0	17.60	22.84	20.22	25.84	22.10	28.46	26.59
1	17.60	0	5.62	4.49	9.36	4.49	10.11	9.36
2	22.84	5.62	0	7.49	6.37	3.75	5.62	4.49
3	20.22	4.49	7.49	0	12.73	3.75	10.86	11.98
4	25.84	9.36	6.37	12.73	0	10.11	9.36	4.49
5	22.10	4.49	3.75	3.75	10.11	0	7.49	8.24
6	28.46	10.11	10.86	10.86	9.36	7.49	0	4.49
7	26.59	9.36	11.98	11.98	4.49	8.24	4.49	0

Table 2.3 – Elapsed downtime t_i^d of failed OWTs

OWT	t_i^d(hours)
3	48
5	10
7	2

Set of experiments #1: validation of the JSMOP model

This set of experiments considers the complex system comprised of three GA(k, n) subsystems GA$(2, 5)$, GA$(3, 8)$ and GA$(4, 10)$ from experiment set #5 in Diallo et al. (2018). The goal is to show that the proposed JSMOP model finds the same optimal maintenance decisions as in Diallo et al. (2018). Given that the JSMOP is a more general model than the one in Diallo et al. (2018), several parameters are set to zero to allow for a valid comparison to take place. In particular, we set $i = 1$ because there is only one equipment in the reference experiment and all routing times and costs are set to 0. The proposed JSMOP is solved for a required minimum reliability level $R_0 = 0.70$, and different values of T^{max}. The overall results obtained are reported in Table 2.4 along with the optimal results from Diallo et al. (2018). From Table 2.4, one can indeed conclude that the proposed JSMOP model reaches the same solutions as in Diallo et al. (2018).

Table 2.4 – Results of validation experiment with k-out-of-n:G subsystems

T^{max}	Results from JSMOP			From Diallo et al. (2018)		
	$\mathcal{Z}^*$ (\$)	T^* (hr)	$\mathcal{R}^*$ (%)	$\mathcal{Z}^*$ (\$)	T^* (hr)	$\mathcal{R}^*$ (%)
100	147	64	70.84	147	64	70.84
60	153	59	71.56	153	59	71.56
56	154	56	70.18	154	56	70.18

Set of experiments #2: selecting the value of parameter M in JSMOP

For computational efficiency, M should be as small as possible while ensuring that constraints (2.27) and (2.28) are enforced. When $x_{ii'r} = 1$, the value of $M_{ii'}$ is irrelevant for constraint (2.28). However, when $x_{ii'r} = 0$, constraint (2.28) become:

$$O_{ir} + t_{ii'} + \sum_{j \in \mathcal{S}_i} \sum_{p \in \mathcal{P}_{ij}} T_{ijp} \cdot z_{ijpr} \cdot \epsilon_r \leq O_{i'r} + M_{ii'} \tag{2.92}$$

Even in the case where $O_{i'r} = 0$, it is sufficient to set $M_{ii'} \geq T^{max}$ for equation (2.92) to be valid. For illustration purposes, several experiments were conducted where the JSMOP is solved using both a large value ($M = 100,000$) and $M = T^{max}$. The results of the different experiments with varying target reliabilities and shift lengths are shown in Table 2.5. These results clearly show that the computation time (CPUt) reduces significantly for all problem instances when the value of M is set to $M = T^{max}$.

Table 2.5 – Comparison of computation times based on the values of M

R_0(%)	T^{max}	$M = T^{max}$		$M = 100,000$	
		$\mathcal{Z}^*$ (\$)	CPUt (s)	$\mathcal{Z}^*$ (\$)	CPUt (s)
96.0	15	19,656	**32.69**	19,656	132.98
98.0	15	41,341	**50.92**	41,341	383.84
98.0	18	31,818	**63.68**	31,818	501.92
99.5	25	64,692	**2.27**	64,692	392.22

Set of experiments #3: comparison between the JSMOP formulation and the sequential model

In this set of experiments, the sequential and joint SMOP optimization approaches are compared. The same OWT farm and its corresponding input data as described in subsection 2.1.7 are considered. The optimization problems resulting from both the sequential and joint formulations are solved for varying required minimum reliability levels when $T^{max} = 28$ hours and $D = 11$ months. The fixed

and variable costs for the repair crews are set to \$0 as repair crews are not considered for the current experiments. The results obtained are reported in Table 2.6 which gives, for each value of the target reliability, the optimal cost $\mathcal{Z}^*$ and the total number N^* of the OWTs visited by the repair crews. From these results, it is clear that the proposed joint optimization approach achieves equal or in most cases better solutions than the sequential optimization approach. Indeed, the joint approach allows not only to select the best combination of maintenance patterns to meet the target reliability at the lowest cost, but also to select more expensive patterns that either allow visiting more OWTs or lead to a reduction in OWT maintenance time which in turn may reduce the overall incurred cost. More important than the cost savings, is the fact that the joint model allows in some cases to visit more turbines. More turbines maintained means higher reliability and higher OWT availability to produce more electricity which has more economical as well as environmental benefits. This is indeed the most important benefit to the service provider.

Table 2.6 – Comparison of optimal values: case of Experiment #3

	Joint approach (JSMOP)			Sequential approach			JSMOP Cost
$R_0(\%)$	$\mathcal{Z}^*$ (\$)	N^*	CPUt (s)	$\mathcal{Z}^*$ (\$)	N^*	CPUt (s)	reduction (%)
60.0	**16,185**	7	2.03	16,322	7	0.07	0.8
65.0	**17,526**	7	2.20	17,755	7	0.03	1.3
70.0	**19,137**	7	1.88	19,391	7	0.05	1.3
75.0	**20,636**	7	1.77	21,128	7	0.05	2.3
80.0	**23,301**	7	1.31	23,535	7	0.11	1.0
84.5	**29,192**	**7**	214.82	35,149	6	0.50	16.9
85.0	35,425	6	367.84	35,456	6	0.51	0.1
87.0	**37,437**	**6**	686.31	46,320	5	0.52	19.2
90.0	55,575	4	6.23	55,575	4	0.05	0.0
91.0	**56,629**	**4**	7.58	63,075	3	0.06	10.2
95.0	**66,984**	3	2.51	67,091	3	0.07	0.2

As expected, the results in Table 2.6 show that as the required minimum reliability R_0 increases, the total repair time and cost increase. As a result, the total number N^* of OWTs visited per maintenance rotation decrease while the total cost incurred increase. For all R_0 values, the joint approach achieves a lower or equal objective function value (cost) with at least as many or more OWT visits.

As expected, the computation times are in most cases slightly higher with the joint approach than with the sequential as the JSMOP is more complex. However, the results in Table 2.6 show that the benefits of taking a little bit more time are worth it. The joint model can quickly be run few hours before crews are assigned their routes and depart.

For the joint and sequential optimization approaches, when the required minimum reliability level is $R_0 = 91\%$, Tables 2.7 and 2.8 give the respective maintenance patterns selected and the resulting total maintenance cost C_i and time T_i for each turbine i. For example, from the results from Table 2.7, the first row shows the maintenance pattern selected for the first turbine. Accordingly, components E_{112}, E_{114}, and E_{122} are selected for mR (maintenance level $l = 1$), component E_{124} is selected for IM (maintenance level $l = 2$), and components E_{115}, E_{132} and E_{141} undergo a replacement (maintenance level $l = 3$). The remaining components are not maintained. This selected maintenance pattern results in a total maintenance time $T_1 = 14.75$ hours and total maintenance cost $C_1 = \$895$. In Table 2.7, turbines 2, 3 and 6 have a corresponding total maintenance time and a cost equal to zero meaning that they are not visited during the maintenance rotation. As for the sequential approach, Table 2.8 shows the maintenance patterns selected for all OWTs followed by the decisions of which OWTs to visit and the routes that the repair crews should take.

Table 2.7 – Patterns and resulting total time and cost: case of the joint optimization approach (Experiment #3)

OWT i	Patterns selected	Cost C_i ($)	Time T_i (hr)
1	(0, 1, 0, 1, 3), (0, 1, 0, 2), (0, 3, 0), (3, 0)	895	14.75
2	(0, 0, 0, 0, 0), (0, 0, 0, 0), (0, 0, 0), (0, 0)	0	0.00
3	(0, 0, 0, 0, 0), (0, 0, 0, 0), (0, 0, 0), (0, 0)	0	0.00
4	(0, 1, 1, 0, 0), (0, 3, 0, 1), (0, 0, 0), (3, 0)	625	10.50
5	(0, 1, 0, 0, 3), (0, 2, 0, 2), (0, 0, 3), (3, 2)	1045	19.25
6	(0, 0, 0, 0, 0), (0, 0, 0, 0), (0, 0, 0), (0, 0)	0	0.00
7	(2, 2, 1, 0, 0), (0, 3, 0, 2), (0, 2, 0), (3, 3)	1115	19.50

Table 2.8 – Patterns and resulting time and cost: case of the sequential optimization approach: case of Experiment #3

OWT i	Patterns selected	Cost C_i ($)	Time T_i (hr)	Visited
1	(0, 2, 0, 3, 0) (0, 1, 0, 2) (0, 3, 0) (3, 0)	895	15.25	
2	(3, 1, 0, 0, 0) (0, 3, 2, 0) (0, 0, 0) (3, 2)	925	17.00	
3	(0, 2, 0, 0, 3) (0, 3, 0, 2) (0, 3, 0) (2, 3)	1145	20.50	✓
4	(0, 2, 1, 0, 0) (0, 2, 0, 2) (0, 0, 0) (3, 0)	625	11.75	
5	(0, 2, 0, 0, 0) (0, 2, 2, 2) (3, 0, 0) (3, 2)	1020	19.50	✓
6	(0, 1, 0, 0, 3) (0, 0, 0, 3) (0, 3, 0) (3, 0)	920	16.00	
7	(2, 2, 1, 0, 0) (0, 3, 0, 2) (0, 2, 0) (3, 3)	1115	19.50	✓

Now, if one considers from Table 2.7 the particular patterns selected for turbines 1, 4, and 5, it results that the joint optimization approach suggests a combination of maintenance patterns that are more expensive but however require less time to be performed. This allows for more turbines to be visited within the maintenance rotation which in turns leads to a reduction of the penalty cost in particular, and in the overall cost in general. This clearly demonstrates the benefit of jointly solving the SMP and routing problem rather than making the decisions separately.

Table 2.9 – Repair crew assignment and routing obtained by JSMOP: case of Experiment #3

Repair crews	Route	Working time (hr)	Travel cost ($)
RC1	[0, 7, 0]	20.92	53.2
RC2	[0, 1, 4, 0]	26.66	52.8
RC3	[0, 5, 0]	20.43	44.2
	Total:	68.01	150.2

Tables 2.9 and 2.10 indicate the routes, working time, and the required travel costs for the three repair crews under the joint and sequential approaches. The joint approach leads to a higher but rational utilization of the crews which translates into more working time and travel costs.

Table 2.10 – Repair crew assignment and routing obtained by the sequential approach: case of Experiment #3

Repair crews	Route	Working time (hr)	Travel cost ($)
RC1	[0, 3, 0]	21.58	40.4
RC2	[0, 5, 0]	20.68	44.2
RC3	[0, 7, 0]	20.92	53.2
	Total:	63.18	137.8

Set of experiments #4: impact of reliability and multiple imperfect maintenance levels

This set of experiments aims to demonstrate the importance and the added value of considering multiple IM levels when dealing with the joint maintenance and OP. To do so, three scenarios are considered and their corresponding maintenance and routing decisions are compared. In the first scenario (Scenario #1), only mR of failed components is available, while in the second scenario (Scenario #2), only replacement of failed components is available. The third scenario (Scenario #3) corresponds to the proposed JSMOP where multiple maintenance levels are allowed. These scenarios use the same OWT farm and its input data as described in Section 2.1.7 except that the list of maintenance levels of the first and second scenarios are, respectively, reduced to mR and replacement options. The duration allotted to the scheduled break is set at $T^{max} = 25$ hours, and the next mission duration is set at $D = 11$ months. The JSMOP corresponding to each scenario is then solved for two values of the required minimum reliability threshold $R_0(\%) \in \{60, 75\}$. The optimal results obtained are reported in Table 2.11. For each value of the target reliability R_0, Table 2.11 gives the optimal cost $\mathcal{Z}^*$ and the total number N^* of OWTs visited by the repair crews.

Table 2.11 – Impact of reliability and multiple imperfect maintenance levels: case of Experiment #4

	Scenario #1		Scenario #2		Scenario #3 (JSMOP)	
$\mathcal{R}_0(\%)$	$\mathcal{Z}^*$ ($)	N^*	$\mathcal{Z}^*$ ($)	N^*	$\mathcal{Z}^*$ ($)	N^*
60.0	46,970	4	24,095	7	16,185	7
75.0	77,852	1	52,644	4	20,638	7

From the results reported in Table 2.11, the following observations can be made. When the target reliability is set at $R_0 = 60\%$, not all OWTs are visited under Scenario #1. The number of visited OWT reduces to one when the reliability threshold is increased to $R_0 = 75\%$. This result is mainly due to the fact that performing mR only is not sufficient to restore some OWTs reliability to meet the required level. Under Scenario #2, when the target reliability is set at $R_0 = 60\%$, all 7 OWTs are visited but the resulting total cost incurred is increased since the replacement cost of a failed component is naturally higher than its mR. However, under Scenario #2, when the reliability threshold R_0 is increased to 75%, the number of visited OWT goes down to 4. This result is due to the fact that the shift duration is limited and replacements of failed components are time consuming. Now, dealing with Scenario #3 which corresponds to the proposed JSMOP, one can observe that, for both required minimum reliability levels, all OWTs are visited and the total cost is less than that resulting from both Scenarios #1 and #2. Thereby, it can be concluded that including detailed reliability and multiple IM levels, the JSMOP provides better reliability achievements at a minimum cost.

Set of experiments #5: impact of target reliability and trip length

To show how the target reliability impacts the SM plan and repair crews paths assignment, the JSMOP is solved for a limited duration $T^{max} = 16$ hours while the target reliability R_0 is varied from 96% to 99% for the same OWT farm and its corresponding input data as described in subsection 2.1.7. The results obtained are depicted in Table 2.12. This table shows the optimal number N^* of the OWTs visited, the total cost incurred $\mathcal{Z}^*$, in addition to the CPU times. From these results, one may observe

that when the reliability target R_0 increases, the number N^* of OWTs that are visited decreases while the resulting total cost $\mathcal{Z}^*$ increases.

Table 2.12 – Results obtained for $T^{max} = 12$ and varying values of R_0: case of Experiments #5

R_0(%)	N^*	$\mathcal{Z}^*$	CPUt (s)
96	7	19,650	37.5
97	6	29,774	500.8
98	6	32,934	233.3
99	3	61,227	13.4

For the particular case where $R_0 = 97\%$ and $T^{max} = 16$ hours, Table 2.13 lists the maintenance patterns selected for each OWT, as well as the resulting cost C_i (\$), time T_i (hr), and OWT reliability R_i (%). In addition, for each one of the repair crews, Table 2.14 indicates the route, working time, and the required travel cost. Based on the selected repair crew routes it can be seen that the initially failed OWTs are visited as early as possible if not immediately. The three available repair crews each starts their maintenance duties by visiting a failed OWT: the first repair crew ($r = 1$) first visits OWT 7 before completing their route via OWT 4; repair crew $r = 2$ first visits the failed OWT 5, then goes to OWT 6 and returns to the maintenance base, and the last repair crew visits OWT 3 followed by OWT 1. OWT 2 is the only one that was not selected for maintenance despite its reliability which is lower than $R_0 = 97\%$. As consequence, a penalty cost will be incurred. The path selected for each repair crew is illustrated in Figure (2.4).

Table 2.13 – Maintenance results obtained for $R_0 = 97\%$ and $T^{max} = 12$: case of Experiments #5

OWT i	Patterns selected	Cost (\$)	Time (hr)	$\mathcal{R}^*$(%)
1	(0, 1, 0, 1, 0), (0, 1, 0, 0), (0, 2, 0), (0, 0)	290	4.25	97.5
2	(0, 0, 0, 0, 0), (0, 0, 0, 0, 0), (0, 0, 0), (0, 0)	–	–	85.3
3	(0, 1, 0, 0, 1), (0, 1, 0, 2), (0, 2, 0), (2, 2)	565	9.75	97.1
4	(0, 0, 1, 0, 0), (0, 1, 0, 1), (0, 0, 0), (0, 0)	200	2.75	97.2
5	(0, 1, 0, 0, 0), (0, 1, 0, 1), (0, 0, 0), (2, 1)	350	6.00	97.6
6	(0, 1, 0, 0, 1), (0, 0, 0, 0), (0, 1, 0), (2, 0)	280	4.50	97.1
7	(1, 1, 1, 0, 0), (0, 1, 0, 0), (0, 1, 0), (1, 3)	580	9.00	97.4

Table 2.14 – Repair crew assignment and routing for $R_0 = 97\%$ and $T^{max} = 12$: case of Experiments #5

Repair crews	Route	Working time (hr)	Travel cost (\$)
RC1	[0, 7, 4, 0]	13.27	56.92
RC2	[0, 5, 6, 0]	12.05	58.05
RC3	[0, 3, 1, 0]	15.13	42.31

Now, to investigate the impact of the limited duration T^{max} allotted to the maintenance rotation on the SM plan and repair crews routes assignment, the JSMOP is solved for a required minimum reliability $R_0 = 98\%$ and varied values of the shift duration T^{max}. The benefit of such an investigation can be justified from the fact that, in real-world wind farm maintenance setting, shift lengths are frequently shortened or interrupted due to operating environment conditions (*e.g.,* poor weather conditions).

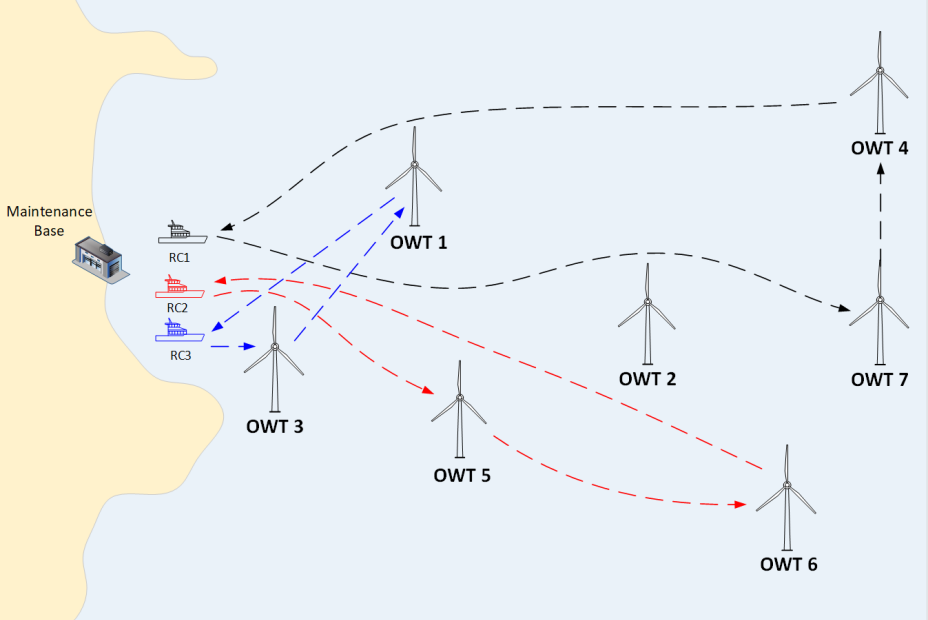


Figure 2.4 – Illustration of paths selected for each repair crew

The results obtained for values of $T^{max} \in \{14, 16, 18, 20\}$ are reported in Table 2.15. For each value of T^{max}, this table gives the optimal number N^* of the OWT to visit, the total cost $\mathcal{Z}^*$ incurred, in addition to the CPU time. From these results, it is observed that as the shift length is shortened the optimal number N^* of visited OWTs decreases while the total cost $\mathcal{Z}^*$ increases.

Table 2.15 – Results when the shift length is varied: case of Experiments #5

T^{max}	Visited nodes	$\mathcal{Z}^*$	CPUt (s)
14	5	41,345	237.5
16	6	32,934	233.3
18	6	31,818	356.7
20	7	23,412	78.2

To summarize, combining the results reported in Tables 2.12 and 2.15, one may observe that there are trade-offs between minimum reliability requirement R_0, maintenance/shift time window T^{max} and total cost $\mathcal{Z}$ as illustrated in Figures 2.5 and 2.6. Indeed, on one hand (see Figure 2.5), for a fixed shift time, if the maintenance decision-maker increases the required minimum reliability, the total cost incurred increased while the number of OWTs visited decreased. This means that the repair crews spend more time in performing high maintenance levels on a restricted number of OWTs which increases the maintenance and penalty costs. On the other hand (see Figure 2.6), if the maintenance decision-maker has more available time flexibility, in this situation, by increasing the shift duration, the number of visited OWTs increases, while the total cost decreases. In fact, increasing the number of visited OWTs reduces penalty cost. Because maintenance costs are typically lower than the penalty costs, when the number of OWTs visited increases, the total maintenance cost increases but the overall total cost decreases since higher penalty costs are avoided. This is similar to results observed by Irawan et al. (2017): good (optimal) solutions are the ones that can avoid the penalty cost.

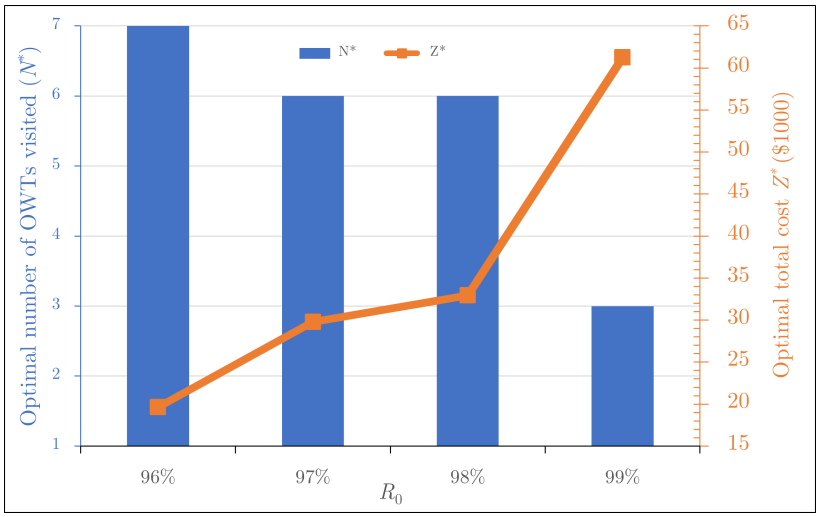


Figure 2.5 – Trade-off between reliability and total cost

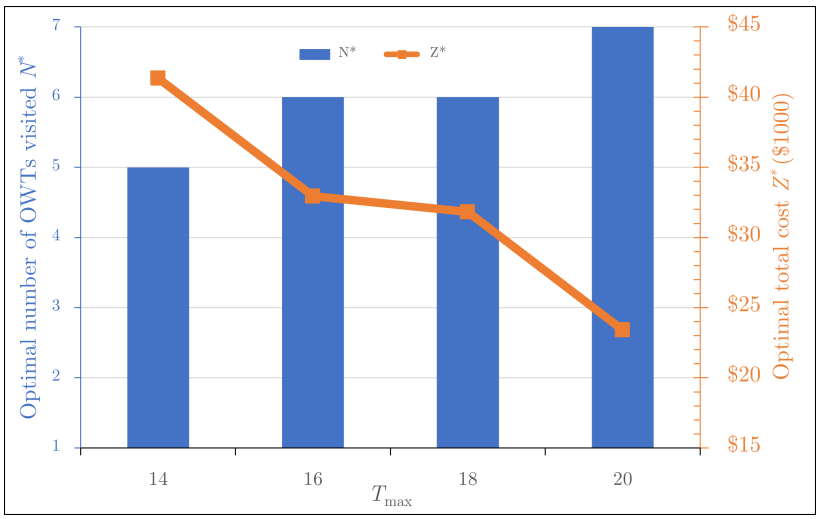


Figure 2.6 – Trade-off between shift duration and total cost.

Set of experiments #6: selecting the optimal number of repair crews

For this set of experiments, the problem described in experiment #2 is extended to include $R = 3$ potential repair crews with different skill levels and costs. For each repair crew $r \in \{1, 2, 3\}$, Table 2.16 shows the repair crews speed coefficients ϵ_r, the variable cost c_r^v and the fixed cost c_r^f. From these data, repair crew 1 ($r = 1$) is characterized by speed coefficient $\epsilon_1 = 0.5$ which means that the repair crew is capable of carrying out a maintenance action in half the required standard duration. With a speed coefficient $\epsilon_2 = 1$, the second repair crew has standard/baseline qualification. The third repair crew is the least skilled with a speed coefficient $\epsilon_2 = 1.2$ which means that it requires 20% more time than repair crew 2 to perform the same maintenance action.

Table 2.16 – Repair crew data: case of Experiments #6

Repair crew r	ϵ_r	c_r^v	c_r^f
1	0.50	40	275
2	1.00	30	200
3	1.20	20	125

The JSMOP is solved for a limited shift duration $T^{max} = 15$ hours while the required minimum reliability R_0 varies from 86% to 99%. The optimal results obtained, namely the number N^* of visited OWTs, the total cost incurred $\mathcal{Z}^*$, the repair crews used, their routes and the CPU times are depicted in Table 2.17. From this table, the following observations can be made:

Table 2.17 – JSMOP results for varying values of target reliability: case of Experiments #6

		Crews used							
R_0	N	#1	#2	#3	Routes	m	T_m	$\mathcal{Z}^*$ (\$)	CPUt(s)
86%	7	✓			[0,3,5,7,4,6,2,1,0]	15	7.125	13,555	94.3
90%	7	✓			[0,3,7,4,6,2,1,0]	16	7.625	14,380	79.7
				✓	[0,5,0]	3	3.600		
94%	7	✓			[0,3,7,4,6,2,1,0]	22	12.125	16,335	112.7
			✓		[0,5,0]	4	4.000		
98%	6	✓			[0,7,4,2,1,0]	20	12.875	30,059	387.6
			✓		[0,3,0]	7	11.750		
				✓	[0,5,0]	5	8.700		
99%	4	✓			[0,7,6,0]	13	11.625	52,041	87.7
			✓		[0,5,0]	6	12.250		
				✓	[0,4,0]	5	8.700		

m: number of maintenance actions carried out by repair crew
T_m: duration in hours of maintenance actions carried out by repair crew

— When the required minimum reliability $R_0 = 86\%$, repair crew #1 (the fastest crew) performs 15 maintenance actions on all 7 turbines for a total maintenance time of 7.125 hours. According to the detailed experiment results in A.3, the 15 maintenance actions are spread as follows: 14 mR and one IM.
— When R_0 increases to 90%, repair crew #1 is no longer capable of visiting all turbines as more maintenance activities are required meaning spending more time at the turbines. Therefore, the model adds repair crew #3 to perform the maintenance actions required at turbine 5. In total, 19

maintenance actions (18 mR and one IM) are conducted and take 11.225 hours (See A.3).
— When R_0 is further increased to 94%, repair crews #1 and #3 are no longer capable of visiting all turbines as more high level maintenance activities are required meaning spending more time at the turbines. Therefore, the model replaces repair crew #3 with the faster repair crew #2 to perform the maintenance actions required at turbine 5. In total, 26 maintenance actions (23 mR, two IM, and one Replacement) are conducted and take 16.125 hours. It should be noted that the number of maintenance actions performed increases as well as their quality (level): two IM and one replacement are carried out.
— When R_0 reaches to 98%, all three repair crews are used. However, the reliability requirement is so high that there is not enough time to perform the required maintenance actions needed. Thus, turbine 6 is not visited hence incurring a penalty cost. In this case, 32 maintenance actions (21 mR, eight IM, and three replacements) are performed for a total maintenance time of 33.325 hours. Similar model behavior is observed for $R_0 = 99\%$. The reliability requirement far exceeds the capacity of the repair crews for the maintenance rotation time available. Thus, turbines 1, 2, and 3 are not visited and turbine 3 remains down. In this case, 24 maintenance actions (9 mR, 12 IM, and three replacements) are performed for a total maintenance time of 32.575 hours.
— From these results, one can conclude that as the target reliability is increased the number of repair crews utilized is increased because it takes longer to carry out the maintenance actions needed to guarantee increasing reliability. In general, the computation time also increases as the reliability requirements are increased. When the repair crew capacity is reached, the number of turbines visited drops and penalties are incurred.

The sets of experiments carried out above clearly show that the proposed joint formulation of the SMP and OP is efficient and yields valid maintenance and routing decisions. The joint formulation is superior to the sequential approach without adding significant computation burden for small and moderately large systems.

Set of experiments #7: performance evaluation of the proposed CG algorithm

This set of numerical experiments is used to demonstrate the ability of the CG approach to achieve optimal or near optimal solutions for different problem instances. For each problem instance, Table 2.18 displays the size of the wind farm $|\mathcal{W}|$, the number of repair crews $|\mathcal{R}|$, the maximum shift duration T^{max} (given in hours), and the required minimum reliability R_0. For all instances the other parameters are: $\pi_i^c = \$15,000$, $\pi_i^p = \$12,500$, and $c_i^d = \$75$/hour. For problem instances 1 to 16, repair crews are standard (*i.e.*, $\epsilon_r = 1, \quad \forall r \in \mathcal{R}$) and the fixed and variable cost are set to \$300 and \$75, respectively. For problem instances 17 to 28, repair crews of different qualification levels are considered. Their corresponding speed ϵ_r, fixed cost c_r^f, and variable cost rate c_r^v are provided in Table 2.19.

Table 2.20 compares the exact optimal solution $\mathcal{Z}^*$ obtained from formulation (2.18)-(2.36) and the solution $\mathcal{Z}_{CG}$ obtained with the proposed CG approach for problem instances 1 to 16. The gap between the two solutions is provided and calculated by the ratio $\frac{\mathcal{Z}_{CG}-\mathcal{Z}^*}{\mathcal{Z}^*}$. A positive gap indicates the exact method is able to find a better solution than the CG approach. Conversely, a negative gap means that the best solution obtained by the exact method within the defined time limit was worse than the solution obtained using the proposed CG method. The results from Table 2.20 clearly show that the CG approach is capable of achieving high-quality solutions in a short amount of time. For small instances of no more than 6 OWTs, the exact method is preferred as it is able to quickly find optimal solutions. However, as the problem size increases, the exact method becomes time consuming even for reaching a solution of acceptable quality. This justifies the need to implement the proposed CG solution method to deal with large number of OWTs.

For problem instances where a wind farm is composed of 10 and 15 OWTs, the CG method is able to find near optimal solutions as the largest gap is 0.06%. For problem instances 13 to 16, where the offshore wind farm is composed of 40 OWTs, the best exact solution found after 3600 seconds is reported in Table 2.20. For these four problem instances, the proposed CG method was able to achieve better solutions in more reasonable computation time.

The results obtained by the exact and the CG methods are reported in Table 2.21 for problem

Table 2.18 – The 28 problem instances used in Experiment #7

Instances	$\|\mathcal{W}\|$	$\|\mathcal{R}\|$	T^{max}	R_0	Instances	$\|\mathcal{W}\|$	$\|\mathcal{R}\|$	T^{max}	R_0
1	6	2	50	0.985	15	40	6	25	0.980
2	6	2	25	0.985	16	40	6	20	0.980
3	6	2	25	0.975	17	6	3	20	0.985
4	6	2	20	0.965	18	6	3	20	0.980
5	10	4	100	0.990	19	6	3	25	0.995
6	10	4	75	0.990	20	6	3	25	0.990
7	10	4	50	0.990	21	10	4	25	0.995
8	10	4	30	0.990	22	10	4	25	0.990
9	15	5	40	0.980	23	10	4	25	0.985
10	15	5	30	0.980	24	10	4	25	0.980
11	15	5	20	0.980	25	15	5	30	0.995
12	15	5	15	0.980	26	15	5	30	0.990
13	40	6	35	0.980	27	15	5	30	0.985
14	40	6	30	0.980	28	15	5	30	0.980

Table 2.19 – Repair crew data: case of problem sinstances 17 to 28 if Experiment #7

Repair crew r	ϵ_r	c_r^v	c_r^f
1	0.75	100	500
2	0.85	80	400
3	1.00	75	300
4	1.15	60	200
5	1.25	50	150

Table 2.20 – Comparison of solution approaches for problem instances 1 to 16: case of Experiment #7

	Exact method		CG method		
Instances	$\mathcal{Z}^*$(\$)	CPUt (s)	$\mathcal{Z}_{CG}$ (\$)	CPUt (s)	Gap (%)
1	42,990	16.40	43,012	1.66	0.05
2	48,316	15.99	48,316	2.53	0.00
3	32,441	6.96	32,481	2.49	0.12
4	28,041	3.42	28,043	3.07	0.01
5	91,669	1800.00*	91,720	9.84	0.06
6	91,673	1800.00*	91,673	8.14	0.00
7	91,676	1800.00*	91,708	6.20	0.03
8	94,941	1800.00*	94,953	2.24	0.01
9	93,461	1800.00*	93,487	17.01	0.03
10	93,494	1800.00*	93,513	9.59	0.02
11	104,104	1800.00*	104,113	3.63	0.01
12	118,320	1800.00*	118,303	2.91	-0.01
13	320,100	3600.00[a]	307,088	446.56	-4.07
14	340,428	3600.00[a]	327,347	169.78	-3.84
15	360,489	3600.00[a]	353,772	43.85	-1.86
16	389,545	3600.00[a]	381,570	12.29	-2.05

* Calculations interrupted after 1800 seconds and best solution reported.
[a] Calculations interrupted after 3600 seconds and best solution reported.

instances 17 to 28. For problem instances 22 to 28, Table 2.21 gives the best exact solution found after 1800 seconds. For problem instances 17 to 21, the proposed CG approach finds the optimal solutions within a competitive computation time. Furthermore, the CG approach finds solutions with equivalent or better quality than the exact method for instances 22, 25, and 26.

Set of experiments #8: large-scale wind farm

This last set of experiments aims to demonstrate how the proposed CG approach is able to deal with JSMOP for large-scale wind farms. This experiment investigates an offshore wind farm composed of 80 OWTs for which maintenance actions are carried out by 15 standard repair crews whose fixed cost c_r^f, and variable cost rate c_r^v are set to \$500 and \$100, respectively. The other parameters are: π_i^c = \$15,000, π_i^p = \$12,500, and c_i^d = \$75/hour.

For $R_0 = 99\%$ and different values of the maximum allotted shift duration T^{max}, the JSMOP is solved by the exact and the CG solution methods. The results obtained are shown in Table 2.22 where the exact solutions reported are those obtained when calculation is interrupted after 7200 seconds. Table 2.22 also gives the number of OWTs visited N^*, the CPU time, the number of columns (Cols) generated through the CG solution method, in addition to the gap between the exact and CG solutions. For three out of the four instances, results in Table 2.22 show that the CG method provides approximate solutions of better quality than the exact method in less computation time.

From the result of Table 2.22, one may also observe that the computation time decreases when the maximum allotted shift duration T^{max} decreases. This result is mainly due to the labeling algorithm used to generate potential routes for repair crews. Indeed, as the shift length decreases the length of a feasible route also decreases, which in turn causes the labeling algorithm to converge rapidly to an optimal route.

For the particular case where $T^{max} = 75$ hours, Table 2.23 indicates the route, the working time, and the travel cost for each of the repair crews. For this instance, 14 repair crews are sufficient to visit all 80

Table 2.21 – Comparison of solution approaches for problem instances 17 to 28: case of Experiment #7

	Exact method		CG method		
Instances	$\mathcal{Z}^*$(\$)	CPUt (s)	$\mathcal{Z}_{CG}$ (\$)	CPUt (s)	Gap (%)
17	74,291	2.48	74,291	0.18	0.00
18	56,752	13.92	56,752	1.17	0.00
19	42,914	59.33	42,914	3.14	0.00
20	35,640	17.14	35,640	6.19	0.00
21	126,733	1233.83	126,733	1.85	0.00
22	942,34	1800.00*	94,234	7.72	0.00
23	74,357	1800.00*	74,428	11.26	0.09
24	64,479	1800.00*	64,750	14.77	0.42
25	184,545	1800.00*	184,434	4.72	-0.06
26	138,949	1800.00*	138,150	21.28	-0.57
27	108,294	1800.00*	108,311	26.57	0.02
28	92,328	1800.00*	92,356	57.82	0.03

* Calculations interrupted after 1800 seconds and best solution reported.

Table 2.22 – Results for large-scale wind farm comprised of 80 OWTs: case of Experiment #8

	Exact method			CG method				
T^{max}	$\mathcal{Z}^*$(\$)	N^*	CPUt (s)	$\mathcal{Z}_{CG}$ (\$)	N	CPUt (s)	Cols.	Gap (%)
75	720,230	80	7200.00*	720,242	80	4165.08	902	0.00
60	729,265	71	7200.00*	722,138	75	1582.10	728	-0.98
50	747,009	62	7200.00*	736,697	67	727.95	562	-1.38
40	787,786	49	7200.00*	766,591	55	175.31	397	-2.69

* Calculations interrupted after 7200 seconds and best solution reported.

OWTs within the shift duration.

Table 2.23 – Repair crew assignment and routing for $R_0 = 99\%$ and $T^{max} = 75$ hours: case of Experiments #8

Repair crews	Route	Working time (hr)	Travel cost ($)
RC1	[0, 15, 4 , 74, 64, 44, 34, 54, 24, 14, 0]	71.52	94.00
RC2	[0, 65, 6 , 46, 56, 16, 36, 66, 0]	70.41	73.00
RC3	[0, 39, 26, 1 , 76, 71, 61, 31, 0]	74.96	93.00
RC4	[0, 41, 51, 52, 11, 21, 72, 12, 0]	74.94	71.00
RC5	[0, 32, 22, 2 , 60, 62, 42, 0]	67.02	93.00
RC6	[0, 30, 80, 28, 18, 78, 50, 0]	66.75	61.00
RC7	[0, 40, 8 , 10, 58, 70, 68, 0]	66.93	99.00
RC8	[0, 48, 55, 38, 20, 45, 5 , 0]	70.64	96.00
RC9	[0, 3 , 25, 75, 35, 7 , 0]	67.55	71.00
RC10	[0, 37, 17, 77, 49, 67, 0]	69.46	93.00
RC11	[0, 19, 79, 29, 57, 69, 0]	69.58	61.00
RC12	[0, 27, 47, 59, 9 , 53, 0]	71.89	99.00
RC13	[0, 73, 23, 13, 63, 0]	64.91	96.00
RC14	[0, 33, 43, 0]	26.93	99.00

From the above results, it can be concluded that the proposed CG solution method can solve the JSMOP for large size offshore wind farms, and provide near joint optimal maintenance and routing decisions in a reasonable computation time.

The sets of experiments carried out above clearly show that the proposed joint formulation of the SMP and OP is efficient and yields valid maintenance and routing decisions. The joint formulation is superior to the sequential approach without adding significant computation burden for small and moderately large systems. The experiments also demonstrate the efficiency of the proposed CG solution method to find near optimal solutions for large-scale problem instances.

Managerial insights

From a research point of view, this is the first work to integrate the repairpersons assignment, detailed reliability and multiple IM levels, and routing decisions to the classical SMP. A CG method combining a Dantzig-Wolfe decomposition with a labeling algorithm is used to efficiently solve the JSMOP for large-scale instances. This novel combination was shown to be able to tackle complex and realistic JSMOPs.

From the industrial and applications point of view, the present work offers maintenance decision-makers a formal framework to deal with the joint SM and routing decisions for realistic offshore wind farms. Instead of sequentially optimizing maintenance and then routing decisions which is suboptimal, the proposed model and CG solution method will help maintenance managers to make their decisions in one step in reasonable computation times. Furthermore, because of low computation times, the maintenance manager can change key problem parameters and then perform what-if analyses.

Results from Experiment #3 show that the joint approach leads to a higher but rational utilization of the crews which translates into more working time and travel costs. The results obtained from the numerical experiments conducted (Experiment #4) showed that resorting to either perfect or plmR, as most existing models do, may lead to situations where fewer OWTs are visited and maintained thus forcing the decision maker to schedule additional rotations or request additional budget. However, when detailed reliability and multiple IM levels are accounted for, it is possible to avoid such uncomfortable situations.

Experiment #5 reveals the important trade-offs between reliability requirement, shift duration and total cost. For a fixed shift duration, when the required minimum reliability increases, the total cost incurred increases while the number of OWTs visited decreases meaning that the repair crews spend more time performing high maintenance levels on a limited number of OWTs which increases the maintenance and penalty costs. If the maintenance decision-maker has more available time flexibility, and increases the shift duration, the number of visited OWTs increases, while the total cost decreases.

2.1.8 Conclusions and Future Research Directions

This section introduced a mathematical formulation to optimally solve the joint SM and OP. The pattern generation algorithm developed by Diallo et al. (2018) was extended to generate all feasible maintenance patterns and eliminate the need for solving a nonlinear formulation. The proposed model was applied to a real-world problem of selecting the maintenance actions to be performed on a set of OWT and also determining the optimal route for several repair crews. Multiple numerical experiments were run to show that the joint model produces cheaper overall solutions than the traditional approach of sequentially solving the SMP followed by the resulting routing problem. Furthermore, the joint model allows in some cases to visit and maintain more turbines which implies economical and environmental benefits. The experiments also demonstrated key characteristics and properties of the model.

Future extensions of the work produced in this section include eliminating some of the assumptions made regarding the way in which maintenance operations are conducted at offshore wind farms. The assumptions that service vessels are required to be present while maintenance is being performed and that they can carry only a single repair crew is not necessarily realistic. It may also be somewhat unrealistic to assume that service vessels have the necessary capacity to carry any spare parts/tools that may be required for maintenance. Developing a model that drops the described assumptions may provide a more accurate and realistic representation of the problem.

Reliability was used as the desired performance indicator for turbines, however availability is another common indicator that is often used in practice. An innovative and novel extension of the SMP and joint JSMOP could be to develop an optimization model with an objective of selecting a subset of components and maintenance actions to be performed on the selected components in which the system availability over the upcoming mission (mean availability) meets or exceeds an availability target. The mean system availability refers to the percentage of time that a system is operable over a mission length and in many industrial applications would be of more importance than the system reliability. However, availability is much more complex performance indicator than system reliability.

In light of the discussion around trade-offs between reliability, time and cost, it is suggested to further investigate the implications of such trade-offs through a formal multi-objective optimization framework. Indeed, it would be more beneficial to decision-makers to quantify the Pareto-frontier between these three objectives. Furthermore, it would allow sensitivity analyses to be conducted based on the preferences of the decision-makers which vary depending on energy production requirements, weather disruption affecting shift duration, and budget available.

One last area of future research is the solving of the JSMOP for very large instances of the problems with hundreds of systems comprised of hundreds of components. This would require the development of decomposition methods, the use of CG to reduce the number of patterns generated, and matheuristics based on recent developments in the solution of the vehicle routing problem with profits.

2.2 Enhancing Critical Network Infrastructure Resilience Through Optimal Post-Disruption Maintenance and Routing decisions

plCNI such as power grids, telecommunications, and transportation are considered critical for both the development and functioning of modern-day societies. Due to the complexity and interconnectedness of these network infrastructures, potential stress and disaster events can have serious impacts and cause wide-spread disruption and network components failure. Developing a post-disruption restoration plan is crucial in improving network resilience and rebounding a network back to a nominal or target functioning state. This section deals with the problem of enhancing network resilience in the aftermath of random stress events. To restore the network to nominal functioning, maintenance actions are performed on its components. To optimize the network resilience after a stress event, a Mixed Integer Non-Linear optimization model is developed. This model aims to ensure efficient restoration, while minimizing the impact of disruptions on critical infrastructures' performance. Unlike many existing network resiliency optimization models, the present model considers traveling time and allows joint decisions related to maintenance level selection, and repair crews' maintenance tasks assignment and routing. Efficient solution methods are implemented and compared to solve the resulting optimization problem. Several numerical experiments are conducted to demonstrate key characteristics of the proposed optimization model and its ability to provide optimal and valid maintenance, repairpersons assignment and routing decisions for real-world plCNI.

2.2.1 Introduction

Modern life has evolved to depend upon complex and highly interconnected production and distribution networks for energy, goods, and services. plCNI such as power grids, rail and road transportation, telecommunication, water distribution, and oil/gas are considered critical for the economic development and social well-being of modern societies (Zhang et al., 2011). The failure or malfunction of such networks has the potential to negatively impact the lives of many. Throughout this section, the term CNI will be used interchangeably with network to refer to these complex systems. For example, the Northeast blackout of 2003, a massive power outage, affected approximately 55 million people in the United States and Canada (Minkel, 2008). Similarly, the Christchurch earthquake in 2011 disrupted sixty percent of electricity and water supplies (Eidinger and Tang, 2012; Jacques et al., 2014). In the highly interdependent and complex operational environment, natural disasters, human-made incidents, and technical accidents are unfortunately unavoidable and can lead to failures in network components, causing a decline in the network performance and substantial economic losses. In this context, the terms disruption or stress events will be used to refer to these natural disasters, human-made incidents, or technical accidents. Therefore, it is essential to devise comprehensive disaster response plans for CNIs, ensuring their ability to absorb and swiftly recover from any stress event, enhancing overall resilience and minimizing the impact on society.

Given the crucial role they play, CNIs not only need to operate with extraordinarily high reliability, but also with high readiness to respond to unforeseen events. Resiliency is typically described as a network's ability to effectively absorb, adapt, and rapidly recover from disruptive events (Hosseini et al., 2016). According to the multidisciplinary center for earthquake engineering research (MCEER) there are four properties, namely robustness, redundancy, resourcefulness, and rapidity that can be used to define and measure network resilience (Renschler et al., 2010). Similar properties have also been proposed by Woods (2015):

- **Robustness:** Managing increasing stressors, complexity, and challenges with limited-to-no impact on normal activities.
- **Rebound:** Returning to normal activities after a stress event.
- **Extensibility:** Extending system performance or capabilities to respond to surprise events that challenge current activities.
- **Adaptability:** Managing trade-offs to build adaptive capacity to continuously evolving contexts.

The present work aims to develop a novel joint maintenance and repair crews assignment and routing decision-making approach to improve and enhance networks resilience through rebound. After a disruption, resilience as rebound focuses on "bouncing back" or returning network performance to a normal

or acceptable level. This bounce back is achieved through maintaining or repairing components of the network that have failed, and allocating resources to perform the repairs or installations. The recovery and maintenance decisions that are made have a significant impact on the resilience of the network and its ability to rapidly return to the normal operating conditions.

In the realm of network restoration strategies following disruptions, numerous optimization approaches have been put forth. These optimization models primarily aim to maximize network resilience or to rapidly re-establish system nominal functionality. A mathematical model to optimize the resilience of a transportation network under disaster events is proposed in Liao et al. (2018) where the optimization model considers both preparedness and recovery activities, and the objective function aims to maximize the connectivity of the network after the stressful event. In González et al. (2016), an Interdependent Network Design Problem (INDP) to identify the reconstruction strategy for a partially destroyed system of infrastructure networks at a minimum cost is introduced, and a MILP is then developed to solve the problem. In Nurre et al. (2012), a novel Integrated Network Design and Scheduling Problem (INDSP) to model the restoration of infrastructures after an extreme stress event is presented. The resulting optimization problem is formulated as an integer program, and its application on realistic datasets demonstrates how it can be used in both real-time restoration activities and long-term scenario planning activities. In Nayak and Turnquist (2016), an optimization model is formulated to minimize the total recovery cost after a disruption in a water distribution network. The total cost is a function of unmet demand during the maintenance downtime, and the repair cost itself. A resilience approach based on a graph CNN-integrated deep reinforcement learning model was proposed by Fan et al. (2022) to make repair decisions to improve the resilience of a water distribution network after an earthquake. The resilience framework used is based on three stages: preparedness (stage I), response (stage II) and recovery (stage III). The objective of the proposed model is to maximize the SRI, which is defined as the area under a performance curve. A multi-objective restoration optimization model that aims to maximize the resilience of interdependent infrastructure networks while minimizing the total cost associated with the restoration process is developed in Almoghathawi et al. (2019). The authors in Kong et al. (2021) propose an approach for developing both pre-disaster and post-disaster infrastructure resilience improvement strategies for natural hazards. These measures include hardening of critical components, setting cross-system buffers, and restoring/repairing failed/damaged components. Finally, several papers (Ghorbani-Renani et al., 2021; Wu et al., 2022; Ouyang, 2017) have proposed tri-level defender-attacker-defender (DAD) models with the objective of network resilience enhancement.

In Iannacone et al. (2022), the authors proposed a unified formulation for modeling the recovery of infrastructure systems whose components are subjected to possible degradation interactions and operate under multiple time-varying processes (covariates). The proposed formulation has been implemented to deal with the time-varying reliability and resilience of a potable water distribution network. The authors in Sharma et al. (2020) developed a mathematical framework for large-scale interdependent infrastructure resilience modeling and optimization. As in Iannacone et al. (2022), resilience metrics are proposed to capture the temporal and spatial variations of the recovery process. The proposed approach is validated on an electric power infrastructure. Liu et al. Liu et al. (2020a) developed a resilience approach for water distribution networks under seismic activity risks where pipe lifetimes are exponentially distributed, and Monte Carlo simulation is used to compare three pipe recovery strategies. Reference Hao et al. (2023) developed a resilience optimization approach for a high speed train modeled as an interdependent machine electricity–communication network (IMECN), and used a tabu search algorithm to derive the best recovery strategy. Zhou et al. (2023a) developed a chance constrained optimization model for a joint hardening and restoration of a network system subjected to cascading failures and network dependency. The objective is to minimize the total cost induced by hardening, restoration and economic loss induced by cascading failures. Dui et al. Dui et al. (2024) also dealt with cascading failures and proposed a resilience optimization model for unmanned vehicle distribution networks. In Cai et al. (2021), the authors proposed a modeling and optimization methodology for system resilience by combining Markov chain and dynamic Bayesian networks to evaluate the resilience of the system subjected to various external disaster phenomena. Reference Dui et al. (2021) developed a resilience optimization model for maritime transportation systems where ports and routes are both failure prone. The best recovery/repair sequence of failed ports and routes is carried out according to different importance measures Dui et al. (2023).

In another work, Zhang et al. (2022a), dealing with a critical system whose components lifetimes are exponentially distributed, the authors proposed two importance measures to identify the importance of components for maintenance. These measures are also used to select the best maintenance sequence which allows to achieve the best overall efficiency of system recovery. However, maintenance time and cost are not accounted for.

The problem of identifying optimal post-disruption recovery decisions under uncertainty has received substantial coverage in the literature. The post-disruption restoration problem for critical infrastructures where repair time and available resources are stochastic is studied in Fang and Sansavini (2019). A two-stage stochastic optimization model is proposed to minimize the system resiliency loss. A restoration framework to improve the resiliency of a power distribution network is proposed by Wu and Wang (2021). Potential uncertainties in the recovery process are accounted for and a stochastic optimization model is then developed. The aim is to minimize the ratio of the area under the restoration curve to the area under the target performance curve. To solve the problem, a decomposition method in the form of a modified Lagrangian dual decomposition is used. The authors in Fang and Sansavini (2016) propose a credibility-based fuzzy mixed integer programming approach to deal with network restoration planning under uncertainty. Their model is applied to the Swiss high voltage electric power transmission network, and the results obtained demonstrate efficient restoration solutions. Dealing with repair time uncertainty, Sanci and Daskin (2019) propose a two-stage stochastic programming model integrating decisions relating to the preposition of facility location, distribution of relief items, and the restoration of damaged network arcs.

A common shortcoming of the papers discussed above is that they do not consider the repair crews' routing decisions. However, when developing disaster response plans, it is worth noting that travel time of the repair crews can significantly impact maintenance and routing decisions especially when network components are geographically dispersed. The optimal sequence in which a maintenance plan should be performed will be affected by the amount of time spent by repair crews to travel between the network components (O'Neil et al., 2023). The authors in Li et al. (2019b) introduce a combined optimization strategy that concurrently tackles the scheduling of restoration activities for each affected component, the allocation of restoration tasks among the available technician teams, and the mapping of routes for each of these teams. An ant colony algorithm is used to solve the resulting optimization model. The Power Restoration Traveling Repairman Problem (PRTRP), consisting in optimal routing of technical crews to damaged components in a power system while minimizing the total service disruption time, is introduced by Çavdar et al. (2022). To solve the resulting optimization problem, a bi-directional labeling algorithm is developed. In Morshedlou et al. (2018), an optimization model is proposed to maximize the network resilience by optimally routing multiple repair crews to damaged network components. An efficient feasibility algorithm is developed to generate a strong initial solution. A relaxed MILP model and solution approach that is capable of finding near-optimal feasible solutions for large scale problems are developed in Morshedlou et al. (2021). With the objective of maximizing the resilience of transportation networks, Zhang et al. (2022b) tackle a joint optimization problem for scheduling inspection routes and restoration. The restoration of CNIs when the restoration and travel times are stochastic is addressed by Alkhaleel et al. (2022b). Two-stage stochastic optimization models are proposed to schedule repair activities for a disrupted network with the objective of maximizing the network resilience. A two-stage stochastic optimization model that addresses both the pre-disaster positioning of restoration teams and the subsequent post-disaster routing to failed network components was recently developed by Jia et al. (2023). Recent literature reviews on infrastructure resilience can be found in (Doorn et al., 2019; Sathurshan et al., 2022; Wang et al., 2024).

The major drawback of the references discussed above is that they rely on a restrictive assumption, according to which maintenance actions to be performed on the network components are assumed known, meaning that the maintenance plan is assumed to be an input data to the optimization problems. However, in many industrial settings, multiple maintenance levels ranging from mR to replacement can be available for each failed component/system of the network. In this situation, selecting the appropriate maintenance level to carry out on each damaged component/system is a crucial decision with which the maintenance manager is often faced. Indeed, maintenance resources such as time and budget are usually

limited and thus need to be rationally exploited. In this context, a two-stage restoration model with the goal of minimizing unsatisfied demand, repair, and flow costs is proposed by Alkhaleel et al. (2022a). The authors consider the possibility of restoring components under perfect and imperfect repair modes, where each repair mode is related proportionally to the restored capacity. However, their approach fails to consider the impact of repair modes on the component and network reliability. To fill this research gap, the present work extends the existing network resilience literature by developing a novel integrated network resilience optimization model where optimal decisions related to maintenance level selection, repair crew maintenance task assignment and routing are simultaneously made. The choice of maintenance level for each failed component carries significant implications not only for the reliability of the component but also for the overall network reliability. Furthermore, contrary to the extant literature, our approach offers a fourth stage called the Durability/Viability stage whose crucial function is to ensure that the recovered functionality of the system is maintained for a predetermined mission duration as it would be poor performance to re-establish nominal capacity for only a brief time period. Our approach not only focuses on identifying maintenance actions that enhance the resilience of a network infrastructure but also ensures that the network operates at a specified reliability target over an extended period following a disruptive event. Guaranteeing that the network can consistently operate with a required minimum reliability reduces its failure risk, and the subsequent far-reaching consequences that encompass economic, safety, and environmental aspects. While several previous works have explored the joint optimization of detailed maintenance level selection and routing decisions (Li and Zhang, 2020; Jia and Zhang, 2020; O'Neil et al., 2023), none of these studies encompass the aspect of maintenance planning in the aftermath of disruptive events. The optimization model proposed in the present work is formulated as a MINLP, which is then reformulated as a MILP combined with a novel solution method to find optimal solutions. Two heuristic algorithms are proposed to deal with disruptions impacting a large number of network components. Multiple numerical experiments are conducted to demonstrate the effectiveness of the proposed approach in selecting optimal recovery actions, as well as the efficiency of the solution methods.

The main contributions are summarized as follows:

- A new resilience framework is proposed that extends the classical 3-stage framework. The goal of the novel fourth stage is to ensure that the recovered network can sustain its performance for an extended period of time after recovery. Through numerical experiments, the benefits of considering the fourth stage are highlighted.
- A novel MINLP resilience optimization model is developed. It extends the existing network resilience literature by finding the joint optimal decisions for maintenance level selection, repair crew task assignment, and routing.
- An exact two-phase solution method is devised to identify optimal maintenance plans, repair crew routing, and task assignment decisions for small and moderate-sized networks. For disruptions impacting large networks with many component failures, two heuristic algorithms are proposed: one based on an evolutionary algorithm and the other on Lagrangian decomposition.
- A detailed analysis of the recovery solutions are conducted through several numerical experiments. The three examined solution methods are compared and show the accuracy and the added value of the proposed approaches. Indeed, the present work provides the decision-maker with an effective and efficient maintenance resource allocation and routing decisions such that not only the critical system network can efficiently recover its nominal performance level, but also is able to operate for a sustained period of time after recovery.

The remainder of the section is structured around 5 additional subsections. In Subsection 2.2.2, the notation used and a description of the system investigated are provided. This Subsection also gives the main working assumptions considered, presents the method for computing the reliability of the network infrastructure, and the IM model used. Subsection 2.2.3 presents the mathematical formulation. Subsection 2.2.4 provides a detailed description of the two-phase solution approach used to reformulate the complex MINLP into a MILP, the GA, and the Lagrangian decomposition approach used to handle large-scale problems. Subsection 2.2.5 presents several sets of numerical experiments carried out to illustrate key properties of the model and to demonstrate its ability in selecting optimal maintenance and routing decisions for achieving resilience in real-world infrastructures. Conclusions and future research directions are presented in Subsection 2.2.6.

2.2.2 System Description and Reliability Computation

In this section, the main notation system is first presented. Then the proposed extended resilience framework including the newly introduced fourth phase is presented and discussed. Next, the infrastructure system under consideration is then described. The main working assumptions are presented before the network reliability computation.

Notation

The following notation system is used in the modeling of the problem.

Infrastructure network notation

d	Network demand
s	Source node
t	Sink node
N	Number of nodes
W	Number of arcs/edges
$\mathcal{G}$	Directed graph of the infrastructure network
$\mathcal{N}$	Set of nodes, $\mathcal{N} = \{1, \cdots, N\}$
$\mathcal{A}$	Set of arcs/edges between nodes u and v, $\mathcal{A} = \{(u,v)\|(u,v) \in \mathcal{N}\}$
$\mathcal{A}^F$	Set of damaged arcs/edges, $\mathcal{A}^F \subset \mathcal{A}$
$\mathcal{I}$	Index set of arcs/edges, $\mathcal{I} = \{1, \cdots, W\}$
$\mathcal{I}^F$	Index set of damaged arcs/edges, $\mathcal{I}^F \subset \mathcal{I}$

Routing network notation

$\overline{\mathcal{G}}$	Undirected graph of the routing network
$\overline{\mathcal{N}}^F$	Set of nodes corresponding to damaged network arcs, $\overline{\mathcal{N}}^F = \mathcal{I}^F$
$\overline{\mathcal{N}}^D$	Set of maintenance depots, $\overline{\mathcal{N}}^D = \{0, W+1\}$
$\overline{\mathcal{N}}$	Set of nodes, $\overline{\mathcal{N}} = \overline{\mathcal{N}}^F \cup \overline{\mathcal{N}}^D$
$\overline{\mathcal{A}}$	Set of edges, $\overline{\mathcal{A}} = \{(i,j)\|(i,j) \in \overline{\mathcal{N}}\}$
$\mathcal{K}$	Set of repair crews, $\mathcal{K} = \{1, \cdots, K\}$

Parameters

ΔT	Time step used to discretize the restoration horizon
$\mathcal{T}$	Set of discrete time periods in planing/restoration horizon, $\mathcal{T} = \{0, \Delta T, 2\Delta T, \cdots, T\}$
R_0	Required minimum post-restoration reliability level
C_{uv}	Maximum flow on component/arc $(u,v) \in \mathcal{A}$
c_{ij}	Travel time on arc $(i,j) \in \overline{\mathcal{A}}$
δ_{uvi}	Binary parameter $\delta_{uvi} = 1$ if $i \in \mathcal{I}$ is the index of arc $(u,v) \in \mathcal{A}$, $\delta_{uvi} = 0$ otherwise
$B_i(A_i)$	Effective age of component $i \in \mathcal{I}$ before (after) maintenance
$\mathcal{L}_i$	Set of maintenance levels available for component $i \in \mathcal{I}$
M	A sufficiently large positive number (*i.e.,* big M)

Decision variables

$f_{uv\tau}$	Continuous variable used to track the flow on arc (u,v) at time τ
$z_{uv\tau}$	Binary variable equal to 1 if component (u,v) is functioning at time τ
$r_{i\tau}$	Binary variable equal to 1 if damaged component i is visited and repaired by time τ
x_{ijk}	Binary variable equal to 1 if maintenance crew k traverses arc (i,j)
O_{ik}	Continuous variable used to track the arrival time of maintenance crew k at node i
$\widehat{O}_i$	Time to perform maintenance on damaged component i.
y_{il}	Binary variable equal to 1 if maintenance level l is performed on damaged component i

For clarity, locally required notation will be introduced as needed.

New resilience framework

Figure 2.7 depicts the proposed new resilience framework which extends the 3-stage framework (Preparedness, Response, Recovery) proposed in Fan et al. (2022) by adding the newly introduced stage called "Durability/Viability". The goal of this fourth stage is to ensure that a certain performance criterion of the recovered infrastructure is maintained above a given threshold for a given time period which hereafter referred to as a mission length. From resiliency quality point of view, the Durability/Viability stage is an important one as it ensures that after recovery the service provided is sustained for a sufficiently long time period for the populations or customers affected by the initial service disruption. It would indeed be adding insult to injury for the populations to experience subsequent drop in performance or further interruptions shortly after the outages caused by the initial disruptive event.

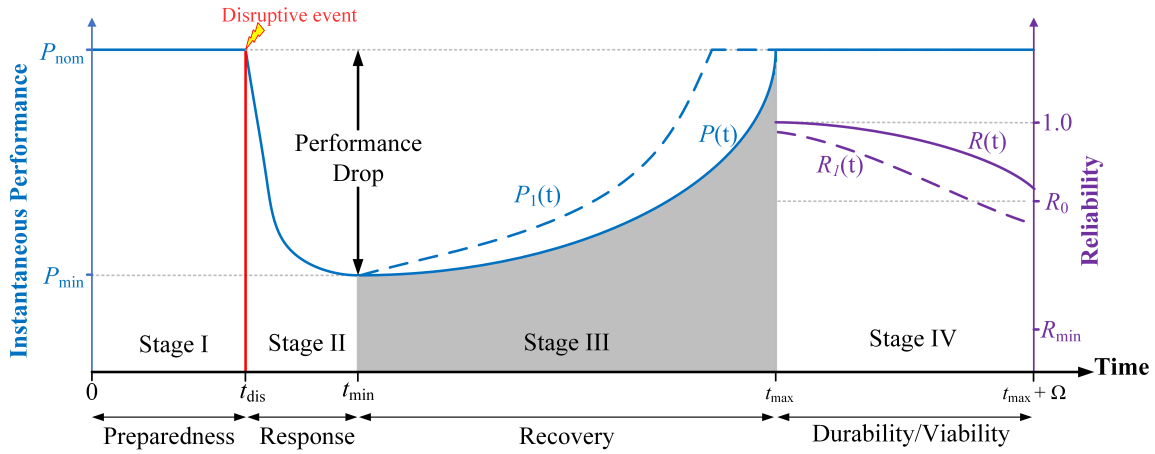


Figure 2.7 – New extended resilience framework adapted from Fan et al. (2022).

The instantaneous performance function $P(t)$, solid blue line, measures the output of the infrastructure in a metric appropriate for the system (*e.g.,* water availability or quality for a water distribution network Cimellaro et al. (2016)). During the pre-disturbance phase (Stage I – Preparedness), the infrastructure performs at its nominal performance value $P(t) = P_{\text{nom}}$. The occurrence of a disruption at time t_{dis} causes a steep drop in performance from P_{nom} to P_{min} in a phase known as Stage II – Response. Stage III – Recovery begins at t_{min} when performance is at its minimum and ends at t_{max} when performance again reaches its nominal value P_{nom}. During Stage III, the SRI is a key performance indicator typically used to quantify the resilience of a system in general, and for a network in particular (Fan et al., 2022). The SRI is mathematically defined as:

$$\text{SRI} = \int_0^T P(t + t_{min})dt. \tag{2.93}$$

where $T = t_{max} - t_{min}$ is the restoration planning horizon. A larger area under the resilience curve generally implies a smaller portion of the performance loss caused by the disruption, meaning that the system is more resilient against disruptive events (Wu and Wang, 2021). Figure 2.7 also depicts a second performance function $P_1(t)$, dashed blue line, reaching P_{nom} faster than $P(t)$, thus being more resilient. If the goal is to select the most resilient profile in the short-term, then $P_1(t)$ is better than $P(t)$. However as explained earlier, such a myopic decision has drawbacks because reaching the nominal performance quickly may cause poor durability in the next stage.

Stage IV – Durability/Viability is characterized by the ability of the recovered infrastructure to satisfy its intended function which is defined in this section as the probability $R_s(d, \Omega)$ that the flow in

the network will satisfy the network demand d for a given mission length Ω. The system is said to be durable/viable if $R_s(d, \Omega)$ is equal or greater than a required minimum post-restoration reliability level R_0. The choice of recovery operations and their quality of execution influence the profile of $R_s(d, \Omega)$ as well as the recovery costs and how long the recovered system will remain viable (*i.e.,* $R_s(d, \Omega) \geq R_0$). It should be noted that the durability/viability stage can be characterized by measures other than reliability such as network availability, dependability, performability or any other appropriate performance index (Trivedi et al., 2009; Chari et al., 2014). In Figure 2.7, $R(t)$ and $R_1(t)$ are the resulting reliability (durability) functions for performance function $P(t)$ and $P_1(t)$ respectively. $P_1(t)$ is more resilient than $P(t)$ during Stage III, but its corresponding reliability $R_1(t)$ during Stage IV is much lower than $R(t)$ as it traded off future reliability for quicker low-level maintenance repairs to reach nominal performance faster. There is thus, a sweet spot between immediate resilience and long-term viability/durability.

The profile and trends of the performance and reliability functions during stages III and IV are dependent on the recovery, routing, and maintenance decisions and actions that are implemented during the recovery phase. Through the optimal routing of repair crews and selection of maintenance actions, the area under the resilience curve can be maximized while satisfying the post-restoration minimum reliability level R_0. Spending more time and performing higher level maintenance activities may delay the restoration time at a higher cost but it has the potential of providing better durability/viability. Therefore, these trade-offs will be investigated to derive managerial insights for decision makers.

System description

The system considered is a 2-terminal CNI with its known source and sink nodes denoted as s and t respectively. In line with Morshedlou et al. (2018), two interrelated infrastructure and routing networks are considered. The former shows how components are connected to perform the intended function of the infrastructure. The latter network represent the physical routing network (*e.g.,* roads, paths) used by repair crews to navigate between the maintenance depot(s) and infrastructure components. A superimposed visual representation of the relationship between the infrastructure and the routing networks is shown in Figure 2.8. The infrastructure network is modeled as a directed graph $\mathcal{G} = (\mathcal{N}, \mathcal{A})$ where $\mathcal{N} = \{1, \ldots, N\}$ is the set of nodes and $\mathcal{A} = \{(u, v) | u, v \in \mathcal{N}\}$ is the set of arcs. The source and sink nodes are numbered as 1 and N, respectively. The set $\mathcal{I} = \{1, \ldots, W\}$ is the index set of arcs. The arcs in the infrastructure network $\mathcal{G}$ are referred to as network components. These arcs (components) are failure-prone, while the nodes are perfectly reliable.

When a stress event occurs, a subset $\mathcal{A}^F \subset \mathcal{A}$ of arcs fails, leading to a degradation in the network performance. The network's performance refers to its ability to meet the demand from the source to the sink nodes. To restore the network's nominal functionality, maintenance actions must be carried out on the network components. Repair crews are routed to the failed network components to perform the needed repairs aiming to increase the network flow and meet the total demand. By successfully restoring a sufficient number of failed components, the network can reach a required minimum performance threshold and recover from the disruption, hence "bouncing back."

Similar to the infrastructure network, the routing network is modeled as an undirected graph $\overline{\mathcal{G}} = (\overline{\mathcal{N}}, \overline{\mathcal{A}})$ where the set $\mathcal{A}^F$ of failed arcs in the infrastructure network corresponds to the set $\overline{\mathcal{N}}^F \subset \overline{\mathcal{N}}$ of nodes in the routing network $\overline{\mathcal{G}}$. This routing network also includes a set of maintenance depots denoted as $\overline{\mathcal{N}}^D \subset \overline{\mathcal{N}}$ from which the maintenance crews start and end their routes/trips. Accordingly, the set $\overline{\mathcal{N}}$ of nodes in the routing network graph $\overline{\mathcal{G}}$ is defined as $\overline{\mathcal{N}} = \overline{\mathcal{N}}^F \cup \overline{\mathcal{N}}^D$. The time spent by maintenance crews to traverse arc $(i, j) \in \overline{\mathcal{A}}$ is denoted by c_{ij}. A maintenance crew arriving at a failed component/node $i \in \overline{\mathcal{N}}^F$ performs the maintenance level determined for that component. Higher maintenance levels require more time to complete but naturally result in higher component reliability.

After restoration from a disruptive event, it is required that the network must operate at some required minimum reliability R_0 for a defined period of time Ω referred to as the mission. This will ensure that the restored network is durable/viable, and therefore it will not only satisfy the required demand but also do so with a high probability for an extended period of time. The following subsections present the

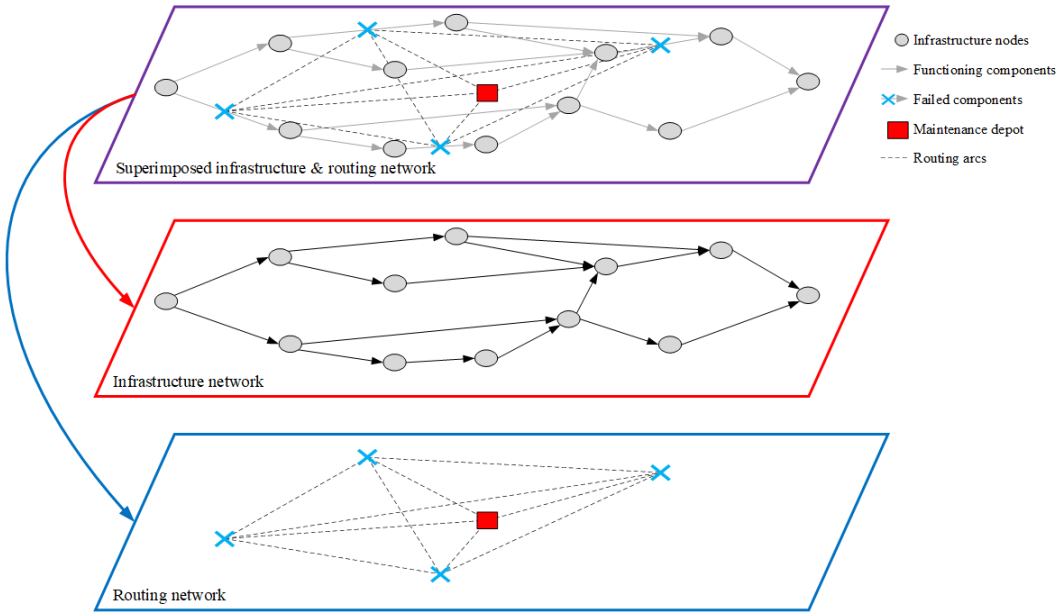


Figure 2.8 – Superimposed infrastructure and routing networks (adapted from Morshedlou et al. (2018)).

main working assumptions made, followed by a detailed description of how the individual component and system reliabilities of the network infrastructure are computed.

Main working assumptions

1. Only failed network components are considered for maintenance. Opportunistic maintenance of functioning components is not accounted for. This is a reasonable assumption given that under the conditions of a disruptive event the goal is usually to restore service and maximum coverage as soon as possible. This is a common assumption made in the literature (Morshedlou et al., 2018; Fang and Sansavini, 2019; Jia et al., 2023).
2. Each maintenance action requires exactly one repair crew. This is usually the case in practice. If several crews are combined to tackle a failed component then the combined crews can still be considered as one team and the model would remain valid. This is a common assumption made in the literature (Alkhaleel et al., 2022b; Zhang et al., 2022b; Jia et al., 2023).
3. All crews have the skills required to perform any required maintenance action. This is a reasonable assumption both in practice and in fundamental models (O'Neil et al., 2023).
4. At the start of the recovery process, component statuses and ages are known. In practice, the implementation of live condition monitoring systems and proper maintenance management systems give relatively easy access to such data. Furthermore, mobile inspection teams sent ahead of the repair crews provide status information needed to run the recovery model.

Component reliability and demand minimal cuts

After performance restoration, the infrastructure network is required to be viable over the duration Ω of Stage IV. Recall that in this section, the system is said to be durable/viable if $R_s(d, \Omega)$ is equal or greater than a required minimum post-restoration reliability level R_0. The reliability of the network can be expressed as a function of the individual component/arc reliabilities and the minimal cuts at level d (d-MCs). The network components are binary (*i.e.,* either functioning or failed), non-identical, s-independent (*i.e.,* components are mutually stochastically independent), and repairable. Without loss of generality, each component $i \in \mathcal{I}$ has a Weibull distributed lifetime with the respective shape and scale parameters β_i and η_i. Given that component i has an effective age A_i after being maintained during the

restoration period, its conditional reliability $R_i(\Omega; A_i, \beta_i, \eta_i)$ for the viability phase of duration Ω is:

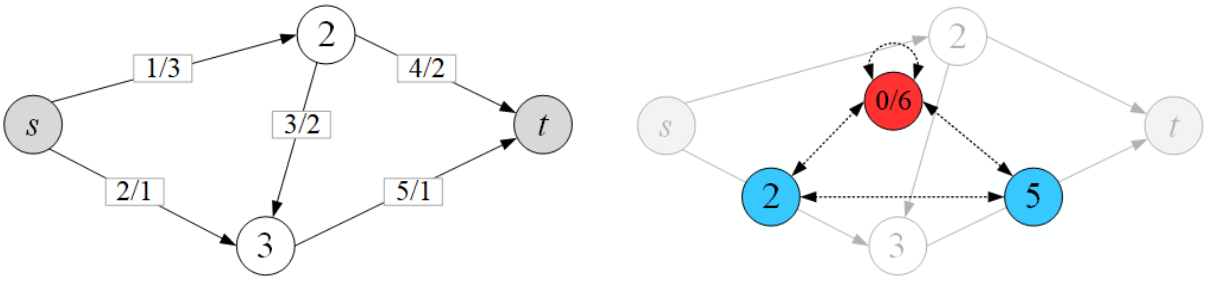
$$R_i^c(\Omega; A_i, \beta_i, \eta_i) = \exp\left(\left(\frac{A_i}{\eta_i}\right)^{\beta_i} - \left(\frac{A_i + \Omega}{\eta_i}\right)^{\beta_i}\right). \tag{2.94}$$

Next, the concepts of MC and d-MCs are used to compute the CNI reliability. Let us first recall that in reliability theory, a cut set is a set of network components such that failure of all its components induce system failure. A MC is defined as a cut set for which system failure cannot occur if one of its components is removed. The d-MCs refer to a set of network components whose simultaneous failures can guarantee that the network's maximal flow will be less than the required demand d, and if one component is removed from the set, the maximal network flow can no longer be guaranteed to be less than d (Li and Zhang, 2020). The CNI can equivalently be expressed as a series structure of the d-MC parallel structures.

To identify the d-MCs of the CNI, we use the element substitution algorithm from Gebre and Ramirez-Marquez (2007) to first find the MCs, then apply the algorithm from Soh and Rai (2005) to identify the d-MCs. Based on the computed component conditional reliabilities $R_i^c(\Omega; A_i, \beta_i, \eta_i)$ and the identified d-MCs, the network reliability can be evaluated using the recursive sum of disjoint products (RSDP) algorithm (Zuo et al., 2007). For relatively small networks, the RSDP algorithm efficiently computes network reliability. However for larger networks, the Linear-Quadratic (LQ) and Esary-Proschan (ESP), which provide efficient approximations (Jin and Coit, 2003), will be used.

To illustrate the relationship between the infrastructure and routing networks, as well as the computation of the network reliability, let us consider an illustrative CNI composed of 5-component/arcs as depicted in Figure 2.9a. Each arc is annotated with two numbers using the following format #/#, where the first number (#) is the arc index and the second is the arc's flow capacity. The target demand is $d = 3$ units, the maximum flow in the network. The number of nodes $N = 4$ (the source s is numbered as node 1 and the sink t as node N). Thus, the sets $\mathcal{N}$ and $\mathcal{A}$ are respectively computed as $\{1, 2, 3, 4\}$ and $\{(1, 2), (1, 3), (2, 3), (2, 4), (3, 4)\}$. The arcs set $\mathcal{A}$ has $W = 5$ arcs/edges, hence the arc index set is defined as $\mathcal{I} = \{1, 2, 3, 4, 5\}$. In this example, a stress event has caused the failure of arcs with index 2 and 5, thus the set of failed arcs $\mathcal{A}^F = \{(1, 3), (3, 4)\}$ and the index set of failed arcs $\mathcal{I}^F = \{2, 5\}$. In terms of the routing network (which is displayed in Figure 2.9b), the node and arc sets are defined as $\overline{\mathcal{N}} = \{0, 2, 5, 6\}$ and $\overline{\mathcal{A}} = \{(0, 2), (2, 0), (0, 5), (5, 0), (0, 6), (2, 5), (5, 2), (2, 6), (5, 6)\}$, where node 0 and node $W + 1 = 6$ correspond to the same maintenance base where repair crews must start and end their trips.

To compute the reliability of the infrastructure for a mission length of Ω time units, the MCs and d-MCs of the network must first be identified. Using the algorithm in Gebre and Ramirez-Marquez (2007), 4 MCs are obtained: $\mathbf{MC}_1 = \{1, 2\}$, $\mathbf{MC}_2 = \{1, 5\}$, $\mathbf{MC}_3 = \{4, 5\}$, and $\mathbf{MC}_4 = \{2, 3, 4\}$. The following d-MCs are found through the algorithm proposed in Soh and Rai (2005): $\mathcal{C}_1 = \{1\}$, $\mathcal{C}_2 = \{2, 3\}$, $\mathcal{C}_3 = \{4\}$, and $\mathcal{C}_4 = \{5\}$. The equivalent series parallel representation of the CNI is shown in Figure 2.10. For this simple example, a tractable expression for the CNI reliability could be derived.



(a) Infrastructure Network. (b) Routing Network.

Figure 2.9 – Infrastructure and routing network of the illustrative example.

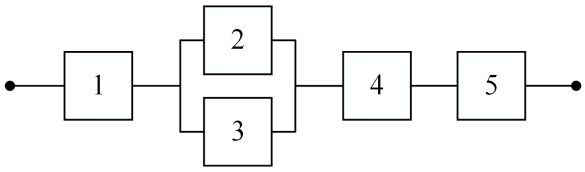


Figure 2.10 – Series structure of d-MC parallel structures: Case of the network in Figure 2.9.

Maintenance modeling

Maintenance crews scheduled to travel to failed component $i \in \overline{\mathcal{N}}^F$ can perform one maintenance level l out of L_i levels available ($l \in \mathcal{L}_i = \{1, \cdots, L_i\}$). These maintenance levels include mR, IM, and replacement (R). Without loss of generality, it is adopted that the lower the level l the lower the maintenance quality is (*i.e.,* the resulting improvement of the reliability is lower). The mR case ($l = 1$) if selected will return a failed component to a working condition while maintaining its age to what it was just before the failure. The replacement level ($l = L_i$) resets the component's age to 0 (*i.e.,* component renewal), while an IM level l ($1 < l < L_i$) if selected will return the components age between that obtained after mR and replacement. Without loss of generality, here the age reduction approach is adopted for IM modeling Malik (1979). Accordingly, to each maintenance level l corresponds an age reduction coefficient θ_l ($0 \leq \theta_l \leq 1$) such that if $l < l'$ then $\theta_l > \theta_{l'}$ $\forall l, l' \in \mathcal{L}_i$. For mR and replacement maintenance levels, the respective age reductions are $\theta_1 = 1$ and $\theta_{L_i} = 0$. Therefore, when a maintenance level l is performed on component i, its effective age B_i is reduced to $A_i = \theta_l \cdot B_i$. There are $k \in \mathcal{K} = \{1, \cdots, K\}$ repair crews available to be assigned to perform the maintenance actions. When a maintenance action of level l is carried out by a repair crew k on component i, it takes t_{il} time units. It is reasonable to assume that higher maintenance levels take more time and cost more.

The following section presents and fully discusses the mathematical optimization model developed to maximize the network resilience while meeting a required minimum reliability threshold R_0 for the next period/mission following a stress event.

2.2.3 Model Formulation

After a stress event, the objective of the proposed optimization model is to maximize the SRI through re-establishing connectivity between source and sink nodes by visiting and maintaining failed network components/arcs. To do so, the restoration horizon $T = t_{\max} - t_{\min}$ is discretized into a finite number of equal time periods of width ΔT thus defining a set of discrete time periods $\mathcal{T} = \{0, \Delta T, 2\Delta T, \ldots, T\}$. The objective function of the proposed optimization model is to maximize the amount D_τ of demand (*i.e.,* maximum network flow) that can be supplied at time $\tau \in \mathcal{T}$ as given by equation (2.95a) in the following formulation:

$$\max \quad \mathcal{Z} = \sum_{\tau \in \mathcal{T}} D_\tau. \tag{2.95a}$$

s.t.:

$$\sum_{v:(u,v)\in\mathcal{A}} f_{uv\tau} - \sum_{v:(v,u)\in\mathcal{A}} f_{vu\tau} = \begin{cases} D_\tau & \text{if } u = 1, \\ 0 & \text{if } u \neq 1 \text{ and } u \neq N, \\ -D_\tau & \text{if } u = N. \end{cases} \quad \forall u \in \mathcal{N}; \forall \tau \in \mathcal{T} \tag{2.95b}$$

$$f_{uv\tau} \leq C_{uv}, \quad \forall (u,v) \in \mathcal{A} \setminus \{\mathcal{A}^F\}; \forall \tau \in \mathcal{T} \tag{2.95c}$$

$$f_{uv\tau} \leq C_{uv} z_{uv\tau}, \quad \forall (u,v) \in \mathcal{A}^F; \forall \tau \in \mathcal{T} \tag{2.95d}$$

$$\sum_{j \in \overline{\mathcal{N}} \setminus \{0\}} x_{0jk} = 1, \quad \forall k \in \mathcal{K} \tag{2.95e}$$

$$\sum_{i \in \overline{\mathcal{N}} \setminus \{W+1\}} x_{i,W+1,k} = 1, \quad \forall k \in \mathcal{K} \tag{2.95f}$$

$$\sum_{i\in\overline{\mathcal{N}}\setminus\{W+1\}} x_{iwk} = \sum_{j\in\overline{\mathcal{N}}\setminus\{0\}} x_{wjk}, \quad \forall w \in \overline{\mathcal{N}}^F; \forall k \in \mathcal{K} \tag{2.95g}$$

$$\sum_{i\in\overline{\mathcal{N}}\setminus\{W+1\}} \sum_{k\in\mathcal{K}} x_{ijk} = 1, \quad \forall j \in \overline{\mathcal{N}}^F \tag{2.95h}$$

$$O_{jk} \leq T, \quad \forall j \in \overline{\mathcal{N}}; \forall k \in \mathcal{K} \tag{2.95i}$$

$$O_{0k} + c_{0j} - O_{jk} - M\left(1 - x_{0jk}\right) \leq 0, \quad \forall j \in \overline{\mathcal{N}} \setminus \{0\}; \forall k \in \mathcal{K} \tag{2.95j}$$

$$O_{ik} + c_{ij} + \widehat{O}_i - O_{jk} - M(1 - x_{ijk}) \leq 0, \quad \forall i \in \overline{\mathcal{N}}^F; \forall j \in \overline{\mathcal{N}} \setminus \{0\}; \forall k \in \mathcal{K} \tag{2.95k}$$

$$\widehat{O}_i \geq \sum_{l\in\mathcal{L}_i} y_{il} t_{il}, \quad \forall i \in \overline{\mathcal{N}}^F \tag{2.95l}$$

$$\sum_{\tau\in\mathcal{T}} r_{i\tau}\tau \geq O_{ik} + \widehat{O}_i, \quad \forall i \in \overline{\mathcal{N}}^F; \forall k \in \mathcal{K} \tag{2.95m}$$

$$\sum_{\tau\in\mathcal{T}} r_{i\tau} = 1, \quad \forall i \in \overline{\mathcal{N}}^F \tag{2.95n}$$

$$\sum_{l\in\mathcal{L}_i} y_{il} = 1, \quad \forall i \in \overline{\mathcal{N}}^F \tag{2.95o}$$

$$z_{uv\tau} \leq \sum_{\tau'\in\mathcal{T}} r_{i\tau'} + M(1 - \delta_{uvi}), \quad \forall i \in \overline{\mathcal{N}}^F; \forall (u,v) \in \mathcal{A}^F; \forall \tau \in \mathcal{T} \tag{2.95p}$$

$$R_i^c = \exp\left(\left(\frac{A_i \sum_{l\in\mathcal{L}_i} y_{il}\theta_l}{\eta_i}\right)^{\beta_i} - \left(\frac{A_i \sum_{l\in\mathcal{L}_i} y_{il}\theta_l + \Omega}{\eta_i}\right)^{\beta_i}\right), \quad \forall i \in \overline{\mathcal{N}}^F \tag{2.95q}$$

$$R_i^c = \exp\left(\left(\frac{B_i}{\eta_i}\right)^{\beta_i} - \left(\frac{B_i + \Omega}{\eta_i}\right)^{\beta_i}\right), \quad \forall i \in \mathcal{I} \setminus \overline{\mathcal{N}}^F \tag{2.95r}$$

$$R_s(d, \Omega) \geq R_0 \tag{2.95s}$$

$$z_{uv\tau} \in \{0,1\}, \quad \forall (u,v), \tau \tag{2.95t}$$

$$r_{i\tau} \in \{0,1\}, \quad \forall i, \tau \tag{2.95u}$$

$$x_{ijk} \in \{0,1\}, \quad \forall i, j, k \tag{2.95v}$$

$$y_{il} \in \{0,1\}, \quad \forall i, l \tag{2.95w}$$

$$f_{uv\tau} \geq 0, \quad \forall u, v, \tau \tag{2.95x}$$

$$O_{ik} \geq 0, \quad \forall i, k \tag{2.95y}$$

$$\widehat{O}_i \geq 0, \quad \forall i. \tag{2.95z}$$

Constraints (2.95b) are the flow conservation constraints, and constraints (2.95c) ensure that the flow on functioning components is less than or equal to the maximum capacity. Constraints (2.95d) ensure the flow on failed components must be 0 until visited and maintained by a repair crew. Constraints (2.95e) and (2.95f) ensure that the route for each repair crew begins and ends at the maintenance depot, respectively. Constraints (2.95g) ensure the connectivity of the route (*i.e.,* if repair crew k enters node w, they must leave it as well). Constraints (2.95h) ensure that each failed component is visited and maintained exactly once. Constraints (2.95i) ensure that the repair crews return to the maintenance depot before the end of the recovery period T. Constraints (2.95j) define the arrival time at the first failed component, Constraints (2.95k) define the arrival times at other failed components, and Constraints (2.95l) define the time spent maintaining each component. Constraints (2.95m) determine the time period when failed components are visited by a repair crew and restored back to the functioning state. Constraints (2.95n) ensure that each failed component is repaired by time τ, and Constraints (2.95o) ensure that a single maintenance action is selected for each visited component. Constraints (2.95p) are used to determine the state of failed components at each time interval. Constraints (2.95q) and (2.95r) define the reliability of failed and functioning components respectively, and constraint (2.95s) ensures the system reliability is greater than or equal to the required minimum reliability level. Finally, Constraints (2.95t)-(2.95w) are binary variable restrictions.

2.2.4 Solution Methods

The optimization model presented above is challenging to solve optimally due to the nonlinear nature of the network reliability expression, making it necessary to develop efficient solution methods. To address this complexity, three solution approaches are proposed: an exact reformulation, an evolutionary algorithm, and a customized algorithm based on Lagrangian decomposition. These methods will be described in detail in the following subsections.

Exact reformulation

The nonlinear formulation (2.95) is reformulated using a two-phase approach adapted from Diallo et al. (2018). Phase I generates a comprehensive set of feasible maintenance patterns for the network. Phase II uses the patterns to formulate the problem as a MILP that maximizes the resilience of the CNI. This two-phase approach guarantees globally optimal solutions. A maintenance pattern $\boldsymbol{p}$ $(p_i|i \in \mathcal{I}^F)$ for the network is a vector of length $|\mathcal{I}^F|$, where element p_i represents the selected maintenance level for component i. The generation of maintenance patterns follows an approach known as the Cartesian combination. This method is employed when multiple maintenance actions are available for each failed component. This method explores all combinations of actions, ensuring a comprehensive search of all possible configurations of maintenance actions assigned to failed network components. For each generated pattern, we use the following LQ approximation to estimate the network reliability achieved. This approximation has demonstrated high accuracy (Jin and Coit, 2003). For a CNI with S d-MCs indexed as C_γ $(\gamma = 1, \ldots, S)$:

$$R_s(d, \Omega) \approx 1 - \sum_{\gamma=1}^{S} Q_\gamma + \sum_{\gamma=1}^{S} \Psi_\gamma \cdot Q_\gamma^2 \tag{2.96}$$

where $Q_\gamma = \prod_{i \in \mathcal{C}_\gamma} (1 - R_i^c(\Omega; A_i, \beta_i, \eta_i))$ is the unreliability of d-MC C_γ, and $\Psi_\gamma = \sum_{\gamma'=1}^{S} \min\left(\frac{Q_\gamma}{Q_{\gamma'}}, \frac{Q_{\gamma'}}{Q_\gamma}\right) - 1$.

To illustrate the principle of the maintenance pattern generation procedure, let us consider the infrastructure network shown in Figure 2.9a, where arcs are numbered from 1 to 5. Let us assume the occurrence of a disaster event causing the failure of arcs 1 and 2, and two common CM levels are available: mR ($l = 1$) and replacement ($l = 2$). In this case, four possible maintenance patterns are possible: $\boldsymbol{p}^1 = [1, 1]^\top$, $\boldsymbol{p}^2 = [1, 2]^\top$, $\boldsymbol{p}^3 = [2, 1]^\top$, and $\boldsymbol{p}^4 = [2, 2]^\top$, where $\top$ is the matrix transpose operator. Each pattern represents a unique combination of maintenance actions. For instance, the first pattern indicates that mR is selected for both failed components/arcs, while the third pattern indicates replacing component 1 and performing mR on component 2.

After generating all feasible patterns, the ones that are dominated by others are identified and then eliminated. A feasible pattern $\boldsymbol{p}^1(p_i^1|i \in \mathcal{I}^F)$ dominates feasible pattern $\boldsymbol{p}^2(p_i^2|i \in \mathcal{I}^F)$ if $p_i^1 < p_i^2$, $\forall i \in \mathcal{I}^F$ which means that the maintenance levels selected for each component under pattern $\boldsymbol{p}^1$ are lower hence take less time to complete than the levels selected for the corresponding components in $\boldsymbol{p}^2$. An illustration of the dominance scheme is presented using the network example of Figure 2.9. If feasible pattern $\boldsymbol{p}^1 = [1, 1]^\top$ already achieves the desired network reliability R_0, then considering the other generated patterns would be unnecessary given that these other three patterns take longer to execute. For example, selecting pattern $\boldsymbol{p}^4$ will result in an increase in the downtime of failed network components compared to pattern $\boldsymbol{p}^1$, based on the assumption that higher maintenance levels take longer to execute. Thus, if both patterns achieve the desired reliability R_0, clearly pattern $\boldsymbol{p}^1$ is preferable as the network will be restored faster. The elimination of obviously non-optimal solutions aims to enhance the efficiency of the solution algorithm. To each non-dominated pattern $h \in \mathcal{H} = \{1, \ldots, H\}$ corresponds a vector $\boldsymbol{\Gamma}^h$ of length $|\mathcal{I}^F|$ composed of the time γ_i^h required to maintain each failed component $i \in \mathcal{I}^F$. The output of the first phase is a dataset containing the non-dominated patterns $\boldsymbol{p}^h$, the corresponding vectors of maintenance times $\boldsymbol{\Gamma}^h$ for each non-dominated pattern $h \in \mathcal{H}$. Algorithm (3) outlines how all feasible patterns are generated.

The aim of the second phase of the approach is to select a non-dominated maintenance pattern and assign a maintenance crew to each maintenance task such that the network resilience is maximized. The

Algorithm 3 Generation of non-dominated maintenance patterns

Input data: t_{il}, β_i, η_i, θ_l, R_0, d-MCs
Initialize: $\overline{\mathcal{H}} = \emptyset$
– Generate the set $\mathcal{H}$ of all valid maintenance patterns.
– Find the cardinality H of the set $\mathcal{H}$: $H = |\mathcal{H}|$.
Initialize: $h = 0$
while $h \leq H$ **do**
  if no pattern in $\overline{\mathcal{H}}$ dominates $\boldsymbol{p}^h$ **then**
    – Compute the network reliability $R_s^h(d, \Omega)$ using (2.96) and the vector $\boldsymbol{\Gamma}^h(\gamma_i^h | i \in \mathcal{I}^F)$ of maintenance durations of pattern h.
    if $(R_s^h \geq R_0)$ **and** $\left(\sum_{i=1}^{|\mathcal{I}^F|} \gamma_i^h \leq T \right)$ **then**
      – $\overline{\mathcal{H}} = \overline{\mathcal{H}} \cup \{\boldsymbol{p}^h\}$.
      – store $\boldsymbol{p}^h$, $\boldsymbol{\Gamma}^h$.
    end if
  end if
  – $h = h + 1$.
end while
– $H := |\overline{\mathcal{H}}|$.

following new decision variable λ^h is introduced:

$$\lambda^h = \begin{cases} 1 & \text{if pattern } h \text{ is selected,} \\ 0 & \text{otherwise.} \end{cases} \tag{2.97}$$

The original optimization problem (2.95) can then be reformulated as a MILP as follows:

$$\max \quad \mathcal{Z} = \sum_{\tau \in \mathcal{T}} D_\tau \tag{2.98a}$$

s.t.: (2.95b) – (2.95j)

$$O_{ik} + c_{ij} + \widehat{O}_i - O_{jk} - M \cdot (1 - x_{ijk}) \leq 0, \quad \forall i \in \overline{\mathcal{N}}^F; \forall j \in \overline{\mathcal{N}} \setminus \{0\}; \forall k \in \mathcal{K} \tag{2.98b}$$

$$\widehat{O}_i \geq \sum_{h \in \mathcal{H}} \lambda^h \gamma_i^h, \quad \forall i \in \overline{\mathcal{N}}^F \tag{2.98c}$$

$$\sum_{\tau \in \mathcal{T}} r_{i\tau} \cdot \tau \geq O_{ik} + \widehat{O}_i, \quad \forall i \in \overline{\mathcal{N}}^F; \forall k \in \mathcal{K} \tag{2.98d}$$

$$\sum_{\tau \in \mathcal{T}} r_{i\tau} = 1, \quad \forall i \in \overline{\mathcal{N}}^F \tag{2.98e}$$

$$\sum_{h \in \mathcal{H}} \lambda^h = 1 \tag{2.98f}$$

$$z_{uv\tau} \leq \sum_{\tau' \in \mathcal{T}} r_{i\tau'} + M(1 - \delta_{uvi}), \quad \forall i \in \overline{\mathcal{N}}^F; \forall (u,v) \in \mathcal{A}^F; \forall \tau \in \mathcal{T} \tag{2.98g}$$

$$z_{uv\tau} \in \{0,1\}, \quad \forall (u,v), \tau \tag{2.98h}$$

$$r_{i\tau} \in \{0,1\}, \quad \forall i, \tau \tag{2.98i}$$

$$x_{ijk} \in \{0,1\}, \quad \forall i, j, k \tag{2.98j}$$

$$\lambda^h \in \{0,1\}, \quad \forall h \tag{2.98k}$$

$$f_{uv\tau} \geq 0, \quad \forall u, v, \tau \tag{2.98l}$$

$$O_{ik} \geq 0, \quad \forall i, k \tag{2.98m}$$

$$\widehat{O}_i \geq 0, \quad \forall i \tag{2.98n}$$

Constraint (2.98f) is added to ensure a single maintenance pattern is selected. Furthermore, Constraints (2.98c) have been revised to incorporate the new decision variable λ^h.

Genetic algorithm

While the exact reformulation can provide optimal solutions, it requires the generation of all feasible maintenance patterns at the outset. In the context of restoration planning for disruptive events that cause widespread network component failures, this approach may not be practical. Additionally, the exact reformulation introduces additional complexity beyond that of the classic VRP, as it simultaneously determines both the optimal maintenance actions and the routing for multiple repair crews. Given these challenges, the exact reformulation is computationally intractable for large-scale instances. Therefore, a more scalable solution method is needed to efficiently handle large networks. The present work proposes an evolutionary approach to solve the problem. Evolutionary algorithms have indeed been shown to perform well on similar combinatorial optimization problems, and are capable of providing good-quality solutions within reasonable computational time.

The GA is a common and very powerful metaheuristic that has been applied successfully to a wide range of combinatorial optimization problems (McCall, 2005; Talbi, 2009). The GA is inspired by the process of natural selection and is initialized with a random population of N_p individuals. Each individual has a fitness value and represents a potential solution to the optimization problem. The fitness value is determined by the fitness function which is defined by the objective function (2.95a). An individual in the population is represented by two solution vectors: the routing vector $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_{|\mathcal{I}^F|+K-1})$ and the maintenance level vector $\boldsymbol{\kappa} = (\kappa_1, \ldots, \kappa_{|\mathcal{I}^F|})$. The vector $\boldsymbol{\delta}$ is composed of K modules/subvectors, each of which corresponds to the route of a specific repair crew. Each module begins with the identifier of the repair crew, followed by the failed components that the crew will service, listed in the order they will be visited. The complete routing vector is formed by combining all K modules. The maintenance level vector $\boldsymbol{\kappa}$ is comprised of $|\mathcal{I}^F|$ elements, where $|\mathcal{I}^F|$ is the total number of failed network components. The ith element in this vector indicates the maintenance level to be performed on the corresponding failed network component. The elements of $\boldsymbol{\kappa}$ are ordered according to the failed components in $\mathcal{I}^F$. The initial population is generated by assigning random values to elements of the solution vectors. To provide an example of the solution encoding, consider again the small network presented in Figure 2.9a. In this disaster scenario, arcs with indices 1, 2, and 3 have failed. Assuming two available repair crews and three possible maintenance levels, Figure 2.11 displays the encoding for a possible solution. In this solution, the first repair crew visits component 1 followed by component 3, while the second repair crew visits component 2. The vector $\boldsymbol{\kappa}$ indicates that the second maintenance level is selected for component 1, the first maintenance level is selected for component 3, and the third maintenance level is selected for component 2.

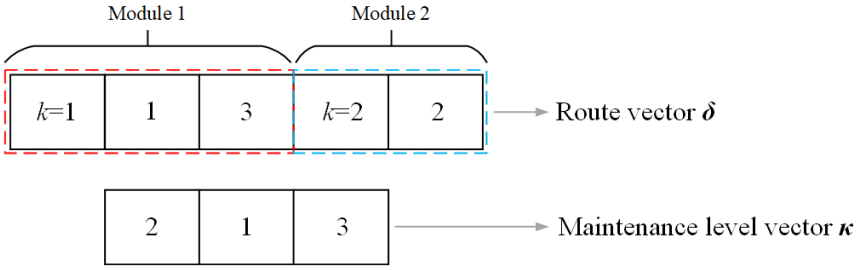


Figure 2.11 – Encoding of a solution.

The elitism ranking procedure is used, where population individuals are ranked based on an increasing order of their fitness values (*i.e.,* higher fitness values imply better solutions as we are dealing with a maximization problem). A fraction of the best ranked solutions (*i.e.,* elites) are selected to move on to the next generation. There are n_{elites} elites in the population. In addition to the elitism operator, crossover and mutation are also key operations performed within each iteration or generation of the GA. The crossover

operator promotes diversity in the solutions population, while the mutation operator prevents premature convergence to local optima. In the present work, order based crossover (Ochelska-Mierzejewska et al., 2021) is applied to the routing vector and uniform crossover is applied to the maintenance level vector. After performing a crossover operation, a mutation operation is applied to the offspring. This involves randomly altering the offspring with a given probability p_{mut}. The simple swap mutation is used on the routing vector, which randomly selects two positions and exchanges their values. For the maintenance vector, we iterate through the list and, with a given probability p_{mut}, we attempt to modify the selected maintenance level. The number of children generated in each generation is denoted as n_{children}. To increase the diversity of individuals in the population, a number of n_{new} new solutions are randomly generated and injected into the next generation. The algorithm terminates when one of three stopping criteria is met: (1) a given number n_{gen} of generations is reached, (2) there is no improvement for a number n_{stagnant} of consecutive generations, or (3) a specified time limit is reached. Within the GA the reliability of the CNI is computed using the LQ approximation introduced in (2.96).

Lagrangian decomposition

The second solution algorithm proposed for restoration planning after stress events impacting a large number of network components is based on Lagrangian decomposition. This approach has been applied to similar problems (Jia et al., 2023) and has achieved excellent performance. The Esary-Proschan (ESP) lower bound (Jin and Coit, 2003) approximation estimating a lower bound of the network reliability is used for the Lagrangian decomposition-based solution method. Its main advantage is that it allows to reformulate Constraint (2.95s) as a set of convex constraints. As introduced previously, there are S d-MCs indexed as C_γ ($\gamma = 1, \ldots, S$). The ESP approximation of the CNI reliability $R_s(d, \Omega)$ is then computed as:

$$\begin{aligned} R_s(d,\Omega) &\geq \prod_{\gamma=1}^{S} \left(1 - \prod_{i \in \mathcal{C}_\gamma} \left(1 - R_i^c(\Omega; A_i, \beta_i, \eta_i)\right) \right) \\ &\geq \prod_{\gamma=1}^{S} \left(1 - Q_\gamma\right) \\ &\geq \prod_{\gamma=1}^{S} \overline{Q}_\gamma \end{aligned} \tag{2.99}$$

where Q_γ and $\overline{Q}_\gamma$ are the respective reliability and unreliability corresponding the d-MC $\gamma \in \{1, \ldots, S\}$ with $Q_\gamma = 1 - \overline{Q}_\gamma$.

According to this approximation, a lower bound of the network reliability can be determined as the reliability of an equivalent series system of the d-MC parallel structures. This approach is particularly suitable for networks with relatively high reliability. Recently, efficient methods have been developed to optimize the reliability of series-parallel systems by using exponential cones (Al-Jabouri et al., 2023b). Here, by letting $q_\gamma = \ln(Q_\gamma)$, $\overline{q}_\gamma = \ln(\overline{Q}_\gamma)$, and by combining exponential cones together with the ESP network reliability approximation, Constraint (2.95s) can be reformulated as the following set of convex

constraints:

$$q_\gamma = \sum_{i \in \mathcal{C}_\gamma^F} \sum_{l \in \mathcal{L}_i} \ln(F_{il}) \cdot y_{il} + \sum_{i \in \mathcal{C}_\gamma \setminus \mathcal{C}_\gamma^F} \ln(F_i), \qquad \forall \gamma \in \{1, \dots, S\} \tag{2.100a}$$

$$Q_\gamma \geq e^{q_\gamma}, \qquad \forall \gamma \in \{1, \dots, S\} \tag{2.100b}$$

$$\overline{Q}_\gamma + Q_\gamma \leq 1, \qquad \forall \gamma \in \{1, \dots, S\} \tag{2.100c}$$

$$\overline{Q}_\gamma \geq e^{\overline{q}_\gamma}, \qquad \forall \gamma \in \{1, \dots, S\} \tag{2.100d}$$

$$\sum_{\gamma=1}^{S} \overline{q}_\gamma \geq \ln(R_0) \tag{2.100e}$$

$$Q_\gamma, \overline{Q}_\gamma \geq 0, \qquad \forall \gamma \in \{1, \dots, S\} \tag{2.100f}$$

$$q_\gamma, \overline{q}_\gamma \leq 0, \qquad \forall \gamma \in \{1, \dots, S\} \tag{2.100g}$$

where $\mathcal{C}_i^F$ refers to the set of failed components in the γth d-MC, and F_{il} refers to the unreliability of component i if maintenance level l is selected, and can be expressed as $F_{il} = 1 - R_i^c(\Omega; A_i \cdot \theta_l, \beta_i, \eta_i)$. Constraints (2.100a) - (2.100d) define the reliability and unreliability of the γth d-MC, and Constraint (2.100e) ensures that the network reliability is greater than or equal to the required minimum reliability level. Constraints (2.100b) and (2.100d) can be expressed as exponential conic constraints, taking the form of $(Q_\gamma, 1, q_\gamma) \in \mathcal{K}_{\text{exp}}$ and $(\overline{Q}_\gamma, 1, \overline{q}_\gamma) \in \mathcal{K}_{\text{exp}}$ for $\gamma \in \{1, \dots, S\}$.

The Lagrangian decomposition approach is discussed next. By relaxing constraints (2.95l) and (2.95m), the problem can be decomposed into several manageable subproblems. Simply ignoring the constraints would result in a weak upper bound, so instead, we allow the constraints to be violated but impose a penalty that is linearly proportional to the amount of violation. This is achieved by introducing non-negative Lagrangian multipliers $\boldsymbol{\pi}$ and $\boldsymbol{\mu}$, leading to the formulation of the following Lagrangian subproblem:

$$[\mathcal{LSP}] : \mathcal{L}(\boldsymbol{\pi}, \boldsymbol{\mu}) = \max \quad \widehat{D} + \sum_{i \in \overline{\mathcal{N}}^F} \pi_i \left(\widehat{O}_i - \widehat{y}_i \right) + \sum_{i \in \overline{\mathcal{N}}^F} \sum_{k \in \mathcal{K}} \mu_{ik} \left(\widehat{r}_i - O_{ik} - \widehat{O}_i \right) \tag{2.101a}$$

$$\text{s.t.:} \quad (2.95\text{b}) - (2.95\text{k}), (2.95\text{n}) - (2.95\text{z})$$

$$\widehat{D} = \sum_{\tau \in \mathcal{T}} D_\tau \tag{2.101b}$$

$$\widehat{r}_i = \sum_{\tau \in \mathcal{T}} r_{it} \cdot \tau, \qquad \forall i \in \overline{\mathcal{N}}^F \tag{2.101c}$$

$$\widehat{y}_i = \sum_{l \in \mathcal{L}_i} y_{il} \cdot t_{il}, \qquad \forall i \in \overline{\mathcal{N}}^F \tag{2.101d}$$

The Lagrangian subproblem can then be decomposed into three subproblems. The first referred to as $\mathcal{LSP}_1$ relates to decision variables $f_{uv\tau}$, z_{uvt}, and r_{it}, and is formulated as follows.

$$[\mathcal{LSP}_1] : \max \quad \widehat{D} + \sum_{i \in \overline{\mathcal{N}}^F} \sum_{k \in \mathcal{K}} \widehat{r}_i \cdot \mu_{ik} \tag{2.102a}$$

$$\text{s.t.:} \quad (2.95\text{b}) - (2.95\text{d}), (2.95\text{n}), (2.95\text{p}), (2.95\text{t}) - (2.95\text{u}), (2.95\text{x})$$

$$\widehat{D} = \sum_{\tau \in \mathcal{T}} D_\tau \tag{2.102b}$$

$$\widehat{r}_i = \sum_{\tau \in \mathcal{T}} r_{it} \cdot \tau, \qquad \forall i \in \overline{\mathcal{N}}^F \tag{2.102c}$$

The second subproblem $\mathcal{LSP}_2$ (2.103) deals with the selection of maintenance actions:

$$[\mathcal{LSP}_2]: \min \quad \sum_{i \in \overline{\mathcal{N}}^F} \widehat{y}_i \cdot \pi_i \tag{2.103a}$$

$$\text{s.t.:} \quad (2.95\text{o}), (2.95\text{w}), (2.100\text{a}) - (2.100\text{g})$$

$$\widehat{y}_i = \sum_{l \in \mathcal{L}_i} y_{il} \cdot t_{il}, \qquad \forall i \in \overline{\mathcal{N}}^F \tag{2.103b}$$

The third and final subproblem $\mathcal{LSP}_3$ (2.104) deals with the routing of repair crews. As this problem is still challenging to deal with, we use a GA where the encoding of the repair crew routes is done the same way as described in subsection 2.2.4. Subproblem $\mathcal{LSP}_3$ is formulated as:

$$[\mathcal{LSP}_3]: \max \quad \sum_{i \in \overline{\mathcal{N}}^F} \widehat{O}_i \cdot \pi_i - \sum_{i \in \overline{\mathcal{N}}^F} \sum_{k \in \mathcal{K}} \left(O_{ik} + \widehat{O}_i \right) \cdot \mu_{ik} \tag{2.104}$$

$$\text{s.t.:} \quad (2.95\text{e}) - (2.95\text{k}), (2.95\text{v}), (2.95\text{y}) - (2.95\text{z})$$

Solving $\mathcal{L}(\boldsymbol{\pi}, \boldsymbol{\mu})$ for any value of $\boldsymbol{\pi}$ and $\boldsymbol{\mu}$ provides an upper bound. To find the best (lowest) upper bound, we solve the Lagrangian dual problem $\min_{\boldsymbol{\pi}, \boldsymbol{\mu}} \mathcal{L}(\boldsymbol{\pi}, \boldsymbol{\mu})$, which can be formulated into the following large-scale LP:

$$[\mathcal{LMP}]: \min \quad \zeta_1 + \zeta_2 + \zeta_3 \tag{2.105a}$$

$$\text{s.t.:} \quad \zeta_1 - \sum_{i \in \overline{\mathcal{N}}^F} \sum_{k \in \mathcal{K}} \mu_{ik} \cdot r_i^h \geq D^h, \qquad \forall h \in \mathcal{H}_1 \tag{2.105b}$$

$$\zeta_2 + \sum_{i \in \overline{\mathcal{N}}^F} \pi_i \cdot \widehat{y}_i^h \geq 0, \qquad \forall h \in \mathcal{H}_2 \tag{2.105c}$$

$$\zeta_3 + \sum_{i \in \overline{\mathcal{N}}^F} \sum_{k \in \mathcal{K}} \mu_{ik} \left(O_{ik}^h + \widehat{O}_i^h \right) - \sum_{i \in \overline{\mathcal{N}}^F} \pi_i \cdot \widehat{O}_i^h \geq 0, \qquad \forall h \in \mathcal{H}_3 \tag{2.105d}$$

$$\pi_i, \mu_{ik} \geq 0 \tag{2.105e}$$

In the above optimization model, $\mathcal{H}_1$, $\mathcal{H}_2$, and $\mathcal{H}_3$ represent the index sets of extreme points corresponding to the first, second, and third subproblems, respectively. Given the exponential size of the sets $\mathcal{H}_1$, $\mathcal{H}_2$, and $\mathcal{H}_3$, we implement Kelley's cutting plane algorithm to solve the $\mathcal{LMP}$ (Kelley, 1960). This algorithm iterates between solving a relaxed version of the master problem, using subsets $\mathcal{H}_1^{'}$, $\mathcal{H}_2^{'}$, and $\mathcal{H}_3^{'}$ of extreme points, to obtain new trial Lagrangian multipliers. Solving this relaxed master problem also provides a lower bound (LB). Then, the Lagrangian subproblem is solved using the trial multipliers to obtain new extreme points and an upper bound (UB). These extreme points are used to generate new cuts (rows) for the $\mathcal{LMP}$, which is then re-solved. The algorithm terminates as soon as one of the following stopping criteria is met: (1) the difference between the UB and LB is sufficiently small, or (2) a specified time limit is reached. In each iteration, a maintenance plan and route assignment for each repair crew are generated. However, simply combining the generated routes and maintenance plan may lead to infeasible solutions, as the plans could exceed the restoration horizon T. To find a feasible solution, we apply the well-known simulated annealing algorithm to adjust the generated routes, aiming to identify a high-quality feasible solution.

2.2.5 Numerical Experiments

In this section, four sets of numerical experiments are conducted and fully discussed to illustrate key properties of the proposed resilience framework and optimization model, and their ability to provide accurate joint optimal maintenance and routing decisions for real world CNIs. The first set of experiments is used to demonstrate the validity and ability of the proposed model to provide joint optimal routing and maintenance decisions for different stress events impacting small to moderately large networks. This set of experiments will also highlight the impact that different model parameters have on restoration plans. The second set of experiments highlights the critical role of incorporating the Durability/Viability

stage in planning restoration activities. A real-world water distribution network (WDN) is investigated in the third set of experiments. In this set of experiments, a different optimization approach that seeks to minimize the time it takes to restore a network will also be presented. The two formulations are compared, and the benefit of maximizing resiliency will be discussed. Finally, the fourth set of experiments presents a comparative analysis of the various solution approaches applied to significantly larger problem instances, where numerous network components have failed. The objective is to demonstrate the effectiveness of the proposed approaches in delivering high-quality solutions for relatively large and complex problem instances.

For all experiments, three potential maintenance levels will be considered: mR ($l = 1$), IM ($l = 2$) and replacement ($l = 3$). The IM level reduces the component age by half ($\theta_2 = 0.5$). The target demand d for all networks considered is set to the maximum flow from source to sink. All experiments are run on a Intel i5 2.9GHz desktop computer with 12GB of RAM running Windows 10. All algorithms were coded in Python 3.12 and the optimization runs were carried out by Gurobi 11.0 using gurobipy. The second subproblem $\mathcal{LSP}_2$ in the Lagrangian decomposition approach was solved using the Mosek Fusion API for Python.

First set of experiments: Optimization model validation

This set of numerical experiments aims to highlight key properties of the proposed CNI resilience optimization model and derive insights into decision-making across a diverse set of randomly generated stress events. The impact that different model parameters have on maintenance and routing decisions, as well as the efficiency of the two-phase solution method will also be investigated. The randomly generated stress events range in severity from relatively few failed components to critical disruptions that cause widespread failures. This set of experiments aims not only to demonstrate the ability of the proposed optimization model to provide valid and reasonable post-disruption restoration plans in a timely manner but, through analyzing the various scenarios, it also aims to uncover the interaction between network resilience, durability (*i.e.,* reliability), maintenance level selection, and routing decisions.

The networks considered for the first set of experiments were randomly generated (Figure 2.12). Both networks have a known source node s and sink node t and multiple transshipment nodes. The arc index and its capacity in CMD are given on each network arc. The first network has a total of 12 components/arcs. When all components function, the maximum achievable network flow is 49 CMD . For this network, When the demand is set at $d = 49$ CMD , there are $S = 10$ d-MCs listed as: $\{1\}, \{2\}, \{3\}, \{5\}, \{6\}, \{7\}, \{9\}, \{10\}, \{11\}$, and $\{12\}$. The second network examined comprises 21 components and exhibits a higher capacity. With all components operating, this network attains a maximum flow capacity of 67 CMD. For this second network, when the demand d is at its maximum (*i.e.,* $d = 67$ CMD), there are $S = 20$ d-MCs: $\{1\}, \{2\}, \{3\}, \{4, 6\}, \{5\}, \{7, 8\}, \{7, 9\}, \{7, 10\}, \{8, 9\}, \{8, 14\}, \{9, 16\}$, $\{11\}, \{12\}, \{13\}, \{15\}, \{17\}, \{18\}, \{19\}, \{20\}$, and $\{21\}$. The component reliability and maintenance data considered are generated as follows: $\eta_i \in \{15, 16, 17, 18, 19, 20\}$, $\beta_i \in \{3.0, 3.1, 3.2, 3.3, 3.4, 3.5\}$, $B_i \sim \mathcal{U}[5.0, 10.0]$, $t_{i1} \sim \mathcal{U}[1.5, 2.5]$, $t_{i2} \sim \mathcal{U}[4.0, 5.0]$, and $t_{i3} \sim \mathcal{U}[8.5, 10.0]$, where $\mathcal{U}$ refers to the uniform probability distribution. The travel time matrix for the two networks as well as the component and maintenance data used are provided in Appendix B. For all the experiments conducted, a mission length of $\Omega = 0.50$ is considered.

The initial experiment (# 1.1) conducted explores the impact of the length of the recovery stage on resilience (*i.e.,* how does the timeframe for restoration activities shape a network's resilience after a disruptive event?) Understanding the impact of the recovery horizon is crucial for decision makers, as the time available for executing restoration activities directly affects response efficacy. Shorter planning horizons will require efficient restoration plans that can be completed quickly, whereas longer planning periods allow for more time-intensive decisions that further enhance network resilience. For this experiment, we randomly generate stress events affecting the first network, causing the failure of $|\mathcal{I}^F| = 6$ components. The recovery horizon T is varied over the range $T \in \{75, 80, 85, 90, 100\}$. For four unique stress events, Figure 2.13 displays the SRI as well as the total maintenance and routing time obtained. From these results, it is shown that increasing the recovery horizon tends to lead to enhanced network

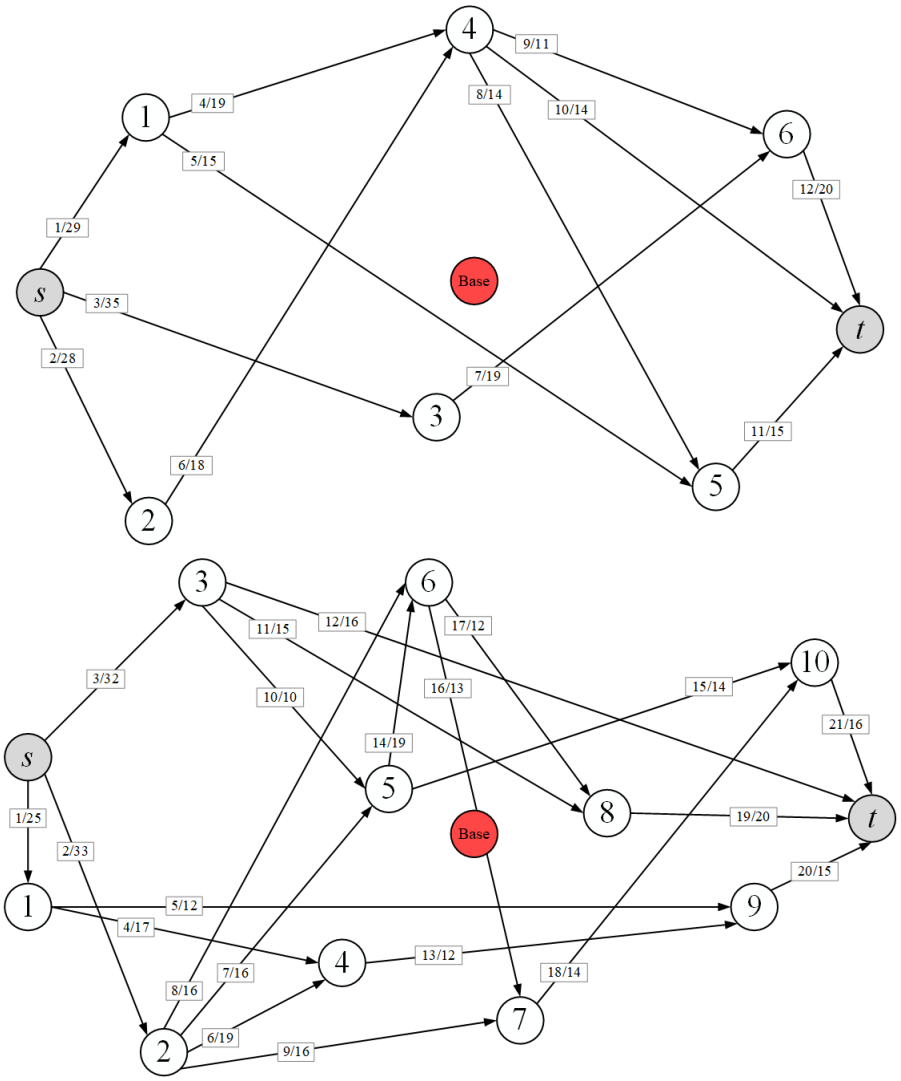


Figure 2.12 – Networks #1 and #2 considered for the first set of experiments.

resilience. This is indeed due to the fact that shorter planning horizons require the selection of more efficient restoration plans. In other words, for lower values of T, the challenge consists of identifying an efficient combination of routing and maintenance decisions to restore the failed components within the planning horizon while meeting the required minimum post-restoration viability threshold R_0.

It is worth noting that a time-efficient restoration plan does not necessarily guarantee a resilient response. To illustrate this observation, let us consider the specific stress event causing the failure of components $\mathcal{I}^F = \{1, 3, 5, 9, 10, 11\}$ (stress event 1). For two different planning horizons $T = 80$ and $T = 100$ hours, Figure 2.14 showcases the system's performance evolution over time as well as the routes selected. Here, it is interesting to note that for both planning horizons, the initial component visited and restored is component 3. This selection likely stems from the significant impact that repairing component 3 has on the network flow, increasing it from 0 CMD to 19 CMD. Beyond this point, the routing differs significantly, emphasizing the interaction between the significance/importance of components and the timing of their restoration levels. For $T = 80$ hours, the recovery is achieved in 79 hours including 56 hours of travel time by giving priority to actions that do not require long travel times even if they do not yield immediate big jumps in performance as depicted by the orange curve on the left. For $T = 100$ hours, the recovery is achieved in about 99 hours including 77 hours of travel by focusing on getting large performance jumps as early as possible even if it means traveling longer distances.

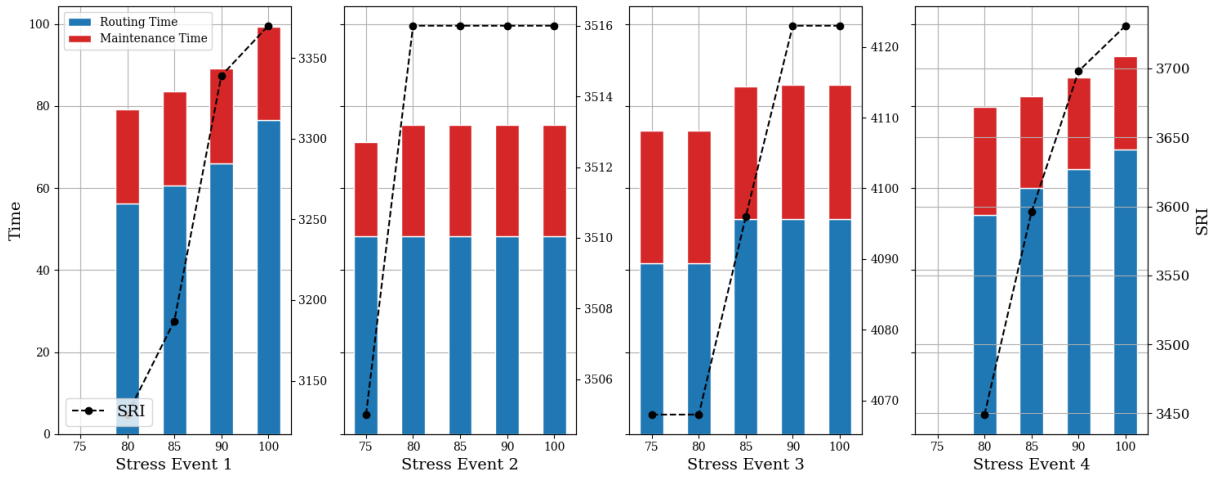
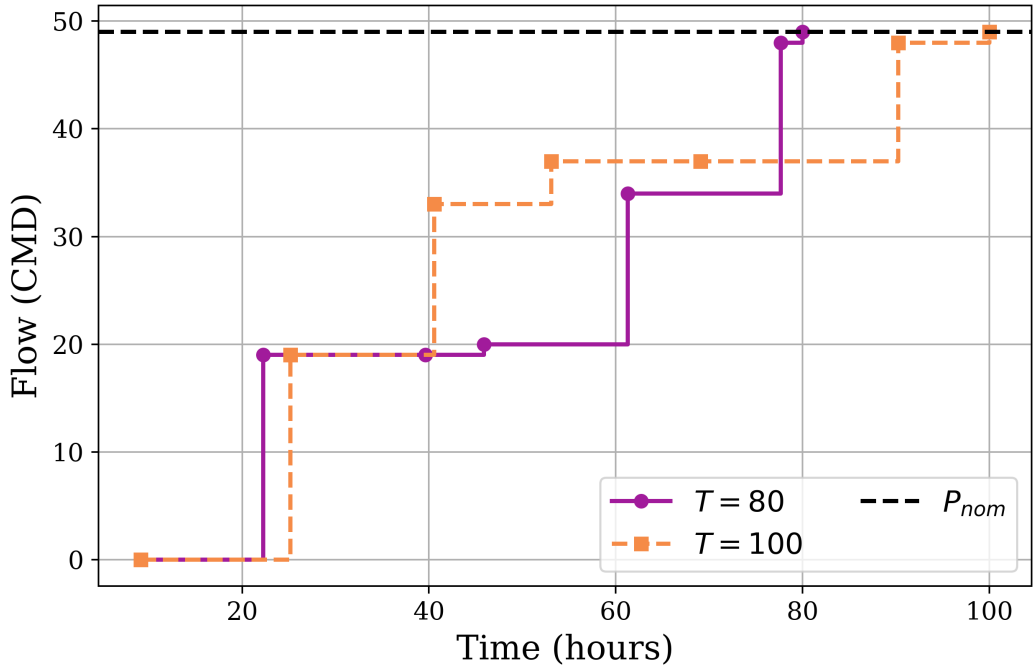


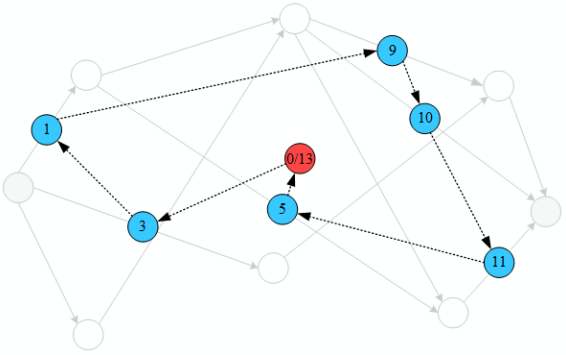
Figure 2.13 – SRI and total routing and maintenance time for varying values of T. Case of experiment #1.1.

The objective of the second experiment (# 1.2) is to explore the impact of the target network durability/viability threshold R_0 on maintenance and routing decisions, and the network's overall ability to recover after disruption. For this experiment, the first network of Figure (2.12) is considered and four stress events are randomly generated causing the failure of $|\mathcal{I}^F| = 8$ components. The reliability threshold is varied over the range $R_0 \in \{0.860, 0.900, 0.940\}$. For each stress event generated and each reliability threshold R_0, Figure 2.15 displays the resilience curve and shows how the system's performance evolves over time. From these results one may observe that by relaxing the required minimum reliability requirement, less time is spent performing maintenance on failed components, and thus the network can "bounce-back" faster. However, decreasing the required minimum reliability increases the risk of mission failure, which may in turn imply serious economic, safety and environmental impacts. Consequently, while the relaxation of the minimum reliability level yields improvements in terms of network resilience and recovery time, it also involves increased vulnerability to post-recovery failure and decreased durability.

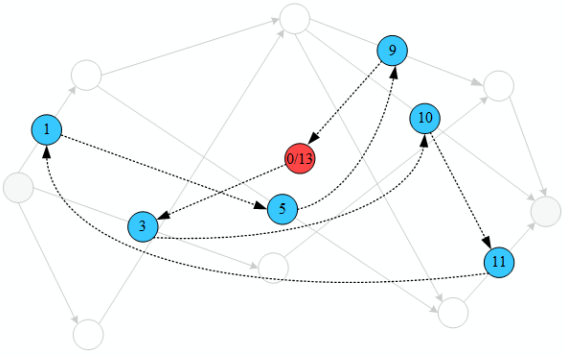
The third experiment (# 1.3) is conducted on the second network of Figure 2.12, subjected to four randomly generated stress events, each resulting in the failure of $|\mathcal{I}^F| = 10$ network components. To show how the number of available repair crews impacts both the network resilience and the recovery,



(a) Restoration curve for $T = 80$ and $T = 100$ (hrs).

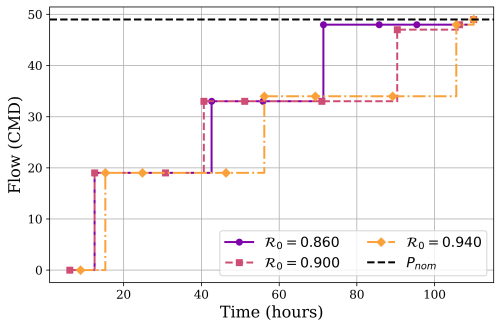


(b) Route selected for $T = 80$ (hrs).

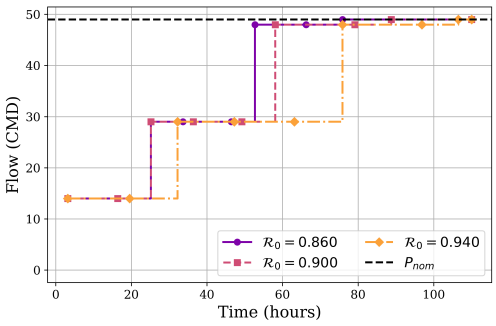


(c) Route selected for $T = 100$ (hrs).

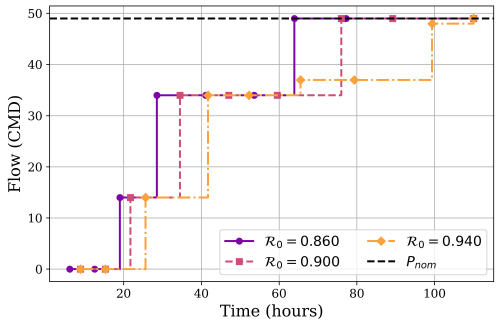
Figure 2.14 – Restoration curves and route selection for varying values of T. Case of experiment #1.1.



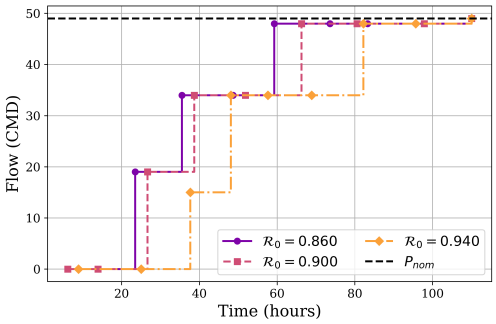
(a) Restoration curve: Stress event #1.



(b) Restoration curve: Stress event #2.



(c) Restoration curve: Stress event #3.



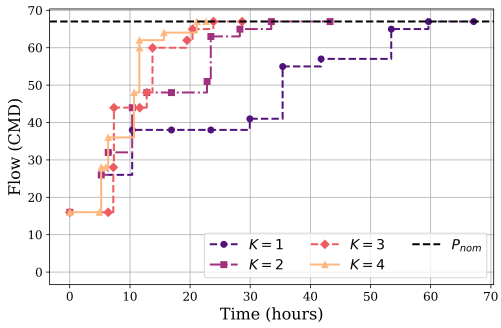
(d) Restoration curve: Stress event #4.

Figure 2.15 – Restoration curves for varying values R_0. Case of experiment #1.2.

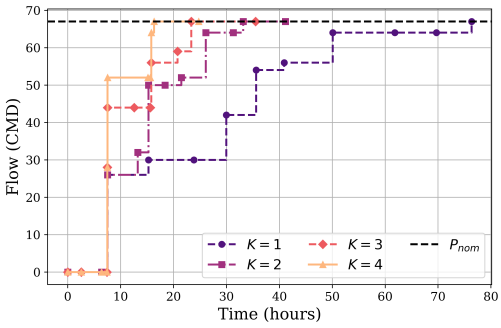
the optimization model is solved for a varying number of repair crews. The overall results obtained are presented in Figure 2.16. As expected, these results indicate that increasing the number of available repair crews leads to enhanced network resilience and improved recovery capabilities. These results highlight the importance and impact of repair crew availability on a network's ability to efficiently return to its nominal operating mode. From Figure 2.16, one may also observe that increasing the number of repair crews beyond a certain point leads to only minor improvements in the network resilience and recovery time. Across all disaster scenarios, increasing the number of repair crews from 3 to 4 results in only a slight improvement in network resilience. Such analysis may in fact help the decision-maker to efficiently manage the maintenance resources available when faced with stress events.

The fourth experiment (# 1.4) investigates the influence of the number of discrete time intervals on restoration and recovery decisions. In this experiment, the second network of Figure 2.12 is once again revisited. For different numbers of discrete time intervals, the proposed optimization model is solved, and the results obtained are displayed in Figure 2.17. The following observations can be made: as the value of $|\mathcal{T}|$ is increased, it is possible to find a more refined and improved restoration plan. The reason behind this observation lies in the enhanced precision that finer divisions of the restoration planning period provide. Indeed, this finer level of detail allows the algorithm to make more accurate and informed decisions. However, it is important to note that beyond a certain point, increasing the value of $|\mathcal{T}|$ does not lead to further solution improvements. Additionally, the computational time increases as the number $|\mathcal{T}|$ of discrete time intervals increases. The overall observations made highlight the importance of finding an appropriate value for $|\mathcal{T}|$ that balances solution quality and computational efficiency.

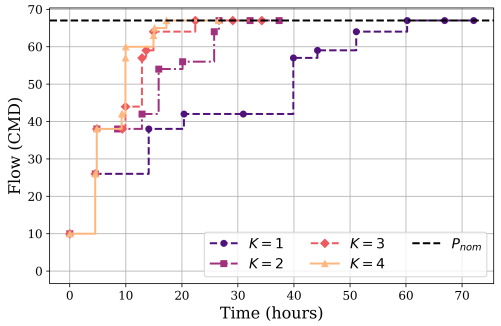
To conclude this first set of experiments, different problem instances are solved to highlight the performance of the two-phase solution method. The results obtained are reported in Table 2.24. For a stress event impacting less than 10 network components, the two-phase solution approach is able to quickly find optimal solutions with the longest computational time being less than 10 minutes. However,



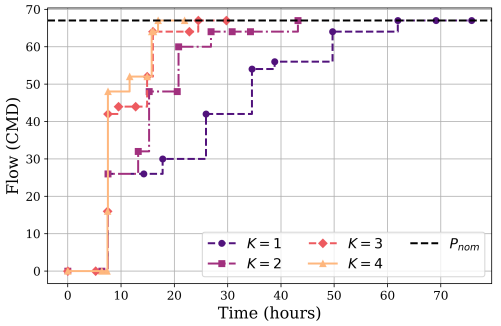
(a) Restoration curve: Stress event #1.



(b) Restoration curve: Stress event #2.



(c) Restoration curve: Stress event #3.



(d) Restoration curve: Stress event #4.

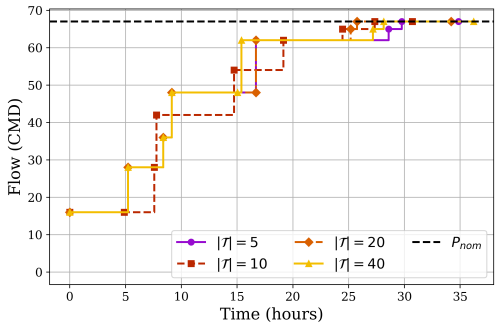
Figure 2.16 – Restoration curves for varying values K. Case of experiment #1.3.

for instances with 10 affected components, the 2-phase method is unable to prove optimality within the defined time limit. However, it is important to note that the optimality gap remains relatively low in instances where the solver was terminated before reaching an optimal solution. For small-size instances, phase I generates the set of non-dominated maintenance patterns in less than a second. However, there is a significant increase in computational time for larger instances with 10 failed components. This outcome highlights a drawback of this exact solution approach, as generating all feasible patterns for large networks may not be practical. Therefore, the two-phase solution method can still be used to provide exact solutions only for small up to moderate sized problems.

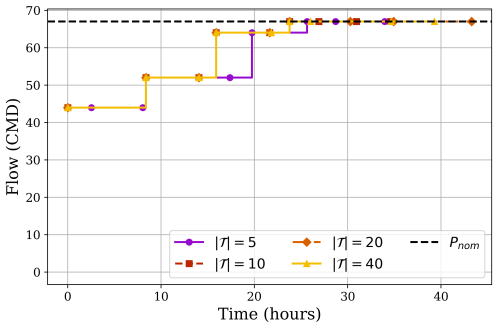
Second set of experiments: Importance of considering the Durability/Viability Stage when planning restoration activities

This second set of experiments aims to demonstrate the importance of incorporating the proposed Durability/Viability stage when planning restoration activities for CNI. For these experiments, the first network of Figure 2.12 is revisited, with a restoration planning period of $T = 120$ hours which is equally discretized to $|\mathcal{T}| = 30$ time intervals. The network suffered from the occurrence of a disruption leading to a set of failed components $\mathcal{I}^F = \{1, 2, 3, 5, 6, 7, 10, 12\}$ which reduces the overall network flow from 49 CMD to 0.

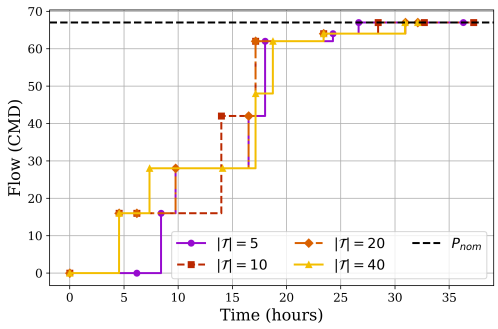
The optimization model is then solved for two cases. Case #1 refers to the proposed approach where the fourth stage (Durability/Viability) is accounted for with the subsequent mission duration of $\Omega = 0.75$ and different required minimum reliability thresholds R_0. Case #2 refers to the existing approaches where the decision-maker focuses solely on maximizing the SRI without considering long-term performance or viability/durability (*i.e.,* without Stage IV). In this case, and from a maintenance point of view, only mR are available as maintenance options. The results obtained are reported in Figure 2.18 which illustrates the restoration curves and the network reliability over the post-recovery period. The detailed maintenance



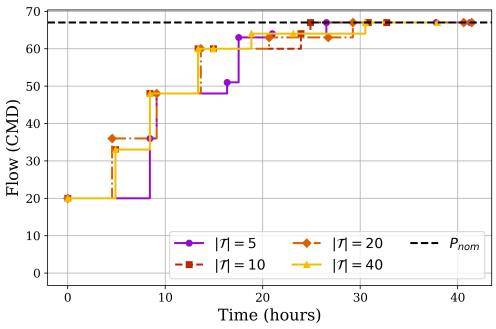
(a) Restoration curve: Stress event #1.



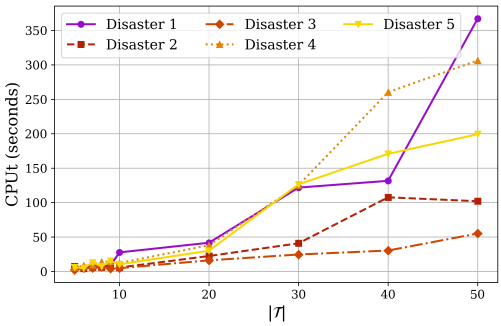
(b) Restoration curve: Stress event #2.



(c) Restoration curve: Stress event #3.



(d) Restoration curve: Stress event#4.



(e) Impact of $|\mathcal{T}|$ on solution time.

Figure 2.17 – Impact of $|\mathcal{T}|$ on network resilience and computational time of the two-phase approach. Case of experiment #1.4.

Table 2.24 – Performance evaluation of two-phase approach.

							Exact Two-Phase				
Network ID	$\|\mathcal{I}^F\|$	Stress Event	R_0	T (hours)	$\|\mathcal{T}\|$	K	Objective value	Gap (%)	Phase I - CPUt (s)	Phase II - CPUt (s)	$\|\mathcal{H}\|$
		1	0.900	50	15	1	353	0.00	0.19	1.92	33
		2	0.900	50	15	1	471	0.00	0.21	1.60	34
1	6	3	0.900	50	15	1	580	0.00	0.14	3.03	40
		4	0.900	50	15	1	458	0.00	0.18	3.44	42
		5	0.900	50	15	1	544	0.00	0.22	1.64	29
		1	0.940	80	30	1	978	0.00	2.12	117.62	249
		2	0.940	80	30	1	1084	0.00	0.37	303.18	65
1	8	3	0.940	80	30	1	999	0.00	0.12	289.82	25
		4	0.940	80	30	1	594	0.00	0.32	32.04	8
		5	0.940	80	30	1	933	0.00	0.37	53.00	59
		1	0.920	40	15	2	514	0.00	2.40	111.99	190
		2	0.920	40	15	2	499	0.00	0.62	81.36	98
1	8	3	0.920	40	15	2	502	0.00	0.15	32.53	37
		4	0.920	40	15	2	503	0.00	1.78	124.33	120
		5	0.920	40	15	2	474	0.00	0.15	354.62	47
		1	0.875	40	15	2	837	0.00	2.72	19.84	99
		2	0.875	40	15	2	941	0.00	1.98	21.01	10
2	8	3	0.875	40	15	2	952	0.00	1.82	9.81	17
		4	0.875	40	15	2	727	0.00	1.90	67.17	27
		5	0.875	40	15	2	876	0.00	2.29	32.82	40
		1	0.900	60	20	2	1087	7.97	64.33	600.00*	296
		2	0.900	60	20	2	1124	7.64	83.03	600.00*	183
2	10	3	0.900	60	20	2	1003	11.76	137.00	600.00*	484
		4	0.900	60	20	2	1100	9.00	101.03	600.00*	376
		5	0.900	60	20	2	1004	11.37	70.11	600.00*	359
		1	0.900	60	20	3	1165	2.23	26.45	600.00*	296
		2	0.900	60	20	3	1215	0.00	76.21	466.69	183
2	10	3	0.900	60	20	3	1115	1.61	129.08	600.00*	484
		4	0.900	60	20	3	1180	1.44	95.71	600.00*	376
		5	0.900	60	20	3	1122	1.13	85.75	600.00*	359

plan selected for each case and the optimal route assignment are reported in Tables 2.25 and 2.26.

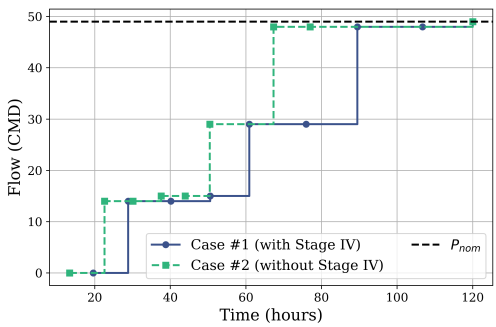
	Maintenance Actions	Route
Case #1 (with Stage IV)	E_1 - R, E_2 - R, E_3 - mR, E_5 - IM, E_6 - IM, E_7 - R, E_{10} - mR, E_{12} - IM	[0, 1, 3, 5, 10, 12, 7, 6, 2, 13]
Case #2 (without Stage IV)	E_1 - mR, E_2 - mR, E_3 - mR, E_5 - mR, E_6 - mR, E_7 - mR, E_{10} - mR, E_{12} - mR	[0, 1, 3, 5, 7, 12, 10, 6, 2, 13]

Table 2.25 – Maintenance actions selected and repair crew routing for $R_0 = 90.0\%$: case of the set of Experiments #2.

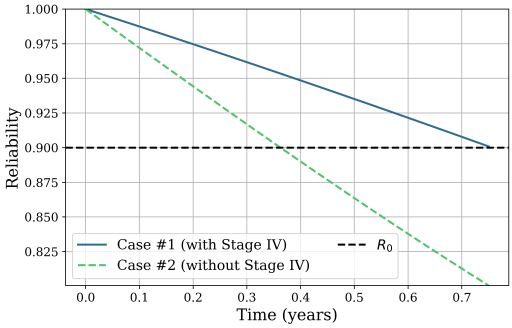
	Maintenance Actions	Route
Case #1 (with Stage IV	E_1 - R, E_2 - R, E_3 - IM, E_5 - R, E_6 - IM, E_7 - R, E_{10} - IM, E_{12} - R	[0, 1, 2, 3, 6, 10, 12, 7, 5, 13]
Case #2 (without Stage IV)	E_1 - mR, E_2 - mR, E_3 - mR, E_5 - mR, E_6 - mR, E_7 - mR, E_{10} - mR, E_{12} - mR	[0, 1, 3, 5, 7, 12, 10, 6, 2, 13]

Table 2.26 – Maintenance actions selected and repair crew routing for $R_0 = 92.5\%$: case of the set of Experiments #2.

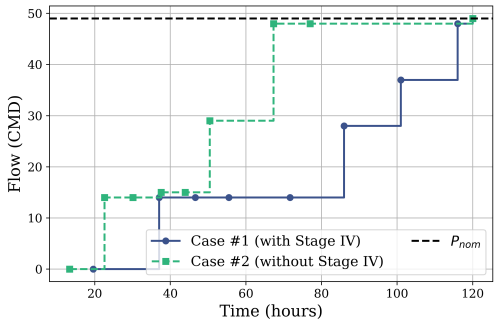
From the results displayed in Figure 2.18, the following observations can be made: considering only mR will obviously lead to the best performance in terms of network recovery speed. Indeed, mR is the least time-consuming maintenance level and thus less time is spent maintaining failed network components resulting then in a larger area under the curve and a faster time to full-rebound. However, performing mR only will likely lead to poor reliability over the defined subsequent post-recovery period/mission. This results from the fact that mR will only bring a failed component back to its functioning state with no impact on its effective age or reliability. These results are clearly and concretely demonstrated in the curves of Figure 2.18, where it is shown that performing mR only results in a system reliability of approximately 80.0%, which is lower than the target reliability values of $R_0 = 90.0\%$ and $R_0 = 92.5\%$. This result demonstrates the trade-off between network recovery speed and its Durability/Viability. While performing only mR results in faster recovery times and a more resilient response, it compromises the overall system reliability, failing to meet the target reliability levels, which in turn may induce disastrous economic, human and environmental consequences. In contrast, incorporating higher maintenance levels,



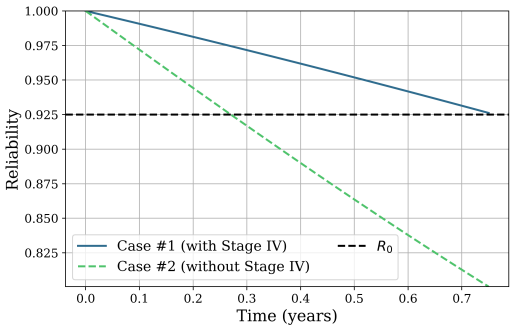
(a) Restoration curves for $R_0 = 0.900$.



(b) Reliability curves for $R_0 = 0.900$.



(c) Restoration curves for $R_0 = 0.925$.



(d) Reliability curves for $R_0 = 0.925$.

Figure 2.18 – Restoration curves and network reliability for varying values of R_0: case of the set of Experiments #2.

as outlined in the proposed resilience framework, may indeed take longer recovery time but ensures higher reliability and better long-term performance of the network during the post-recovery phase. The overall results discussed here highlight the importance of balancing short-term recovery with long-term sustainability when planning restoration activities.

Third set of experiments: Application to a large water distribution network

In this set of experiments, the ability of the mathematical model to solve real-world problems is illustrated by considering the water distribution network (WDN) presented in Kansal et al. (1995) as shown in Figure 2.19. The nodes represent pumping stations and the arcs denote pipelines, with the thicker lines representing the backbone of the network and the thinner lines representing the branches. The pumping stations on the backbone are both transshipment nodes and demand nodes used to supply water to customers on the branches of the network. Similar to Jia and Zhang (2020), we only consider the backbone, resulting then in the network shown in Figure 2.20. The transformed network contains 18 nodes, among which node 1 is the supply node, and all the other nodes, except node 18, represent pumping stations. Node 18 is added to transform the k-terminal network into a two-terminal network by connecting it to each pumping station. The two numbers above each solid arc represent the corresponding arc index and capacity, respectively. The numbers in blue above the dotted arcs represent the demand at the corresponding node. When all components are functioning the maximum network flow is equal to 14300 CMD. The d-MCs under the given demand ($d = 14300$ CMD) of this network are enumerated as $\{1\}, \{2\}, \{3\}, \{5\}, \{6\}, \{7\}, \{9\}, \{10\}, \{11\}, \{12\}, \{13\}, \{14\}, \{15\}, \{16\}, \{17\}, \{18\}$, $\{19\}, \{20\}, \{21\}, \{22\}, \{23\}, \{24\}, \{25\}, \{26\}, \{27\}, \{28\}, \{29\}, \{30\}, \{31\}, \{32\}, \{33\}$, $\{34\}, \{35\}, \{36\}$. Each arc/component is a d-MC of the network, because of this we can compute the reliability as the product of the individual arc reliabilities.

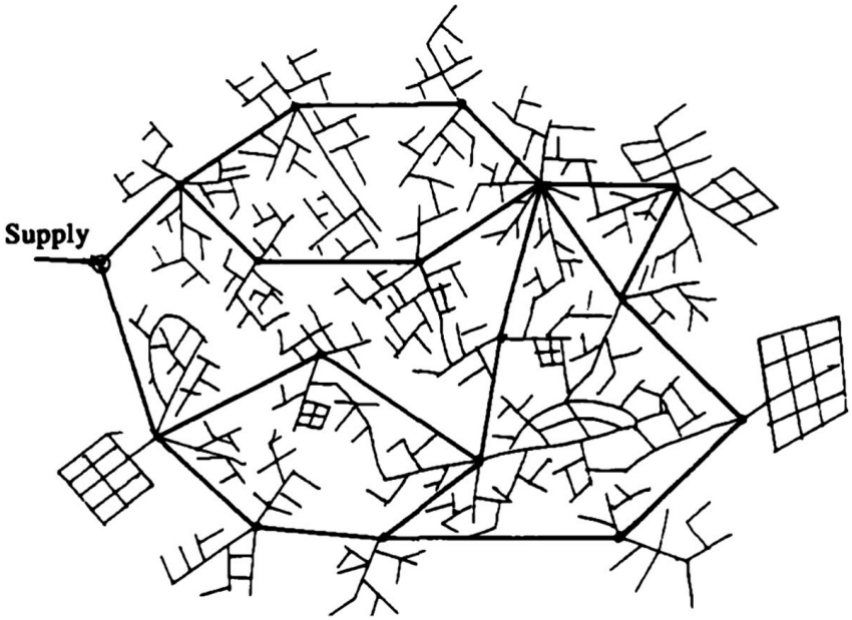


Figure 2.19 – Water distribution network presented in Kansal et al. (1995).

Although this work is focused on maximizing resiliency, there are other interesting and practical metrics to consider when planning the restoration of a critical infrastructure. In fact, minimizing the time to reach full resilience is another interesting objective to consider when planning restoration activities. The term "Time to Full Resilience" (TFR) refers to the point in time at which the network is able to fulfill the target demand d. The term "Time to Full Resilience" was used by Almoghathawi et al. (2019). In the present set of experiments, a new optimization model (Model 2) is formulated to minimize the

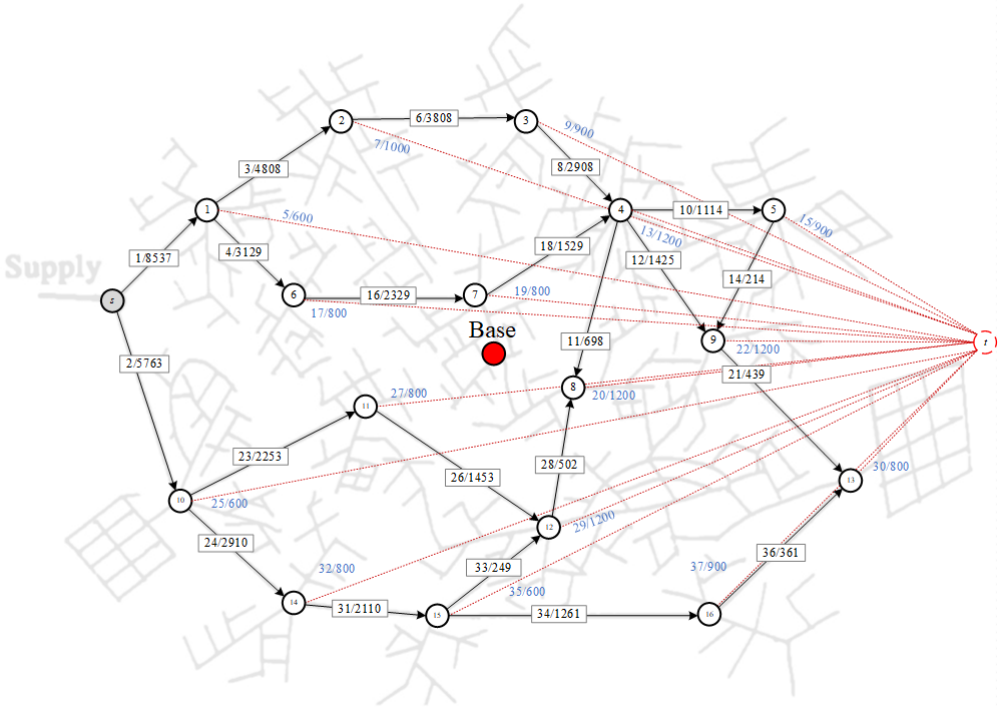


Figure 2.20 – Transformed water distribution network.

TFR. This model is solved in two parts, we first identify the set of feasible maintenance patterns using Algorithm (3), we then solve the following optimization model:

$$\max \quad \sum_{\tau \in \mathcal{T}} \alpha_\tau \tag{2.106a}$$

s.t.

Eq. (2.95b) – (2.95j)

Eq. (2.98b) – (2.98k)

$$D_\tau \geq d - M \cdot (1 - \alpha_\tau), \quad \forall \tau \in \mathcal{T} \tag{2.106b}$$

$$\alpha_\tau \in \{0, 1\}, \quad \forall \tau \tag{2.106c}$$

In the above optimization model, the objective function maximizes the number of discrete time intervals during which the network can successfully fulfill the target demand d. Constraint (2.106b) determines the time intervals when the network flow D_τ has the capacity to meet network demand d.

For this set of experiments, six different disaster scenarios will be studied as outlined in Table 2.27. For all scenarios it will be assumed that the mission length is $\Omega = 0.25$ years. For both Models 1 and 2, Figure 2.21 displays the resilience curve corresponding to each disaster scenario. Detailed results, encompassing the maintenance pattern p^* selected, time to reach full resilience (TFR), and the computational duration of the optimization phase are provided in Table 2.27.

What can be noticed immediately is that the solutions obtained considering Model 2 result in faster network recovery compared to the solutions obtained from Model 1. However, it can also be seen that Model 1 produces solutions that are better in terms of resilience. The area beneath the performance curve when considering Model 1 is larger, and in many cases the area is significantly larger. A greater area beneath the resilience curve signifies a smaller portion of performance loss due to the disruption. In general, Model 1 tends to favor gradual restoration of functionality whereas model 2 will focus on getting big jumps in functionality. Model 1 is interesting for providing services to consumers/populations where

Table 2.27 – Comparison of Model 1 and Model 2.

$\mathcal{I}^F$	$\mathcal{R}_0$	T	K	Model 1			Model 2		
				p^*	TFR	CPUt (secs)	p^*	TFR	CPUt (secs)
$\{3, 4, 6, 8, 16, 18\}$	0.900	50	1	(1, 2, 3, 2, 3, 3)	38.0	3.4	(2, 2, 3, 3, 2, 2)	32.3	0.8
$\{2, 23, 24, 26, 31, 33\}$	0.850	40	1	(1, 1, 1, 1, 1, 2)	22.9	1.5	(1, 1, 1, 1, 2, 1)	21.4	1.0
$\{1, 2, 3, 4, 6, 8, 16, 23\}$	0.880	25	2	(1, 1, 1, 1, 2, 2, 2, 1)	18.9	257.9	(1, 1, 1, 1, 2, 2, 2, 1)	17.4	51.2
$\{10, 11, 12, 14, 21, 28, 34, 36\}$	0.850	20	2	(1, 1, 1, 2, 1, 1, 1, 1)	16.6	51.3	(1, 1, 1, 1, 1, 2, 1, 1)	16.5	13.1
$\{1, 2, 10, 11, 12, 14, 21, 28, 34, 36\}$	0.910	30	3	(2, 3, 3, 3, 3, 3, 3, 3, 2, 3)	26.4	1485.7	(3, 3, 2, 3, 3, 3, 3, 2, 3, 3)	22.7	1800*
$\{3, 4, 6, 8, 16, 18, 23, 24, 26, 31\}$	0.875	20	3	(1, 1, 1, 1, 3, 3, 1, 1, 1, 2)	16.0	1800*	(1, 1, 1, 2, 2, 1, 2, 1, 2, 1)	14.0	1177.2

* Calculations interrupted after 1800 seconds.

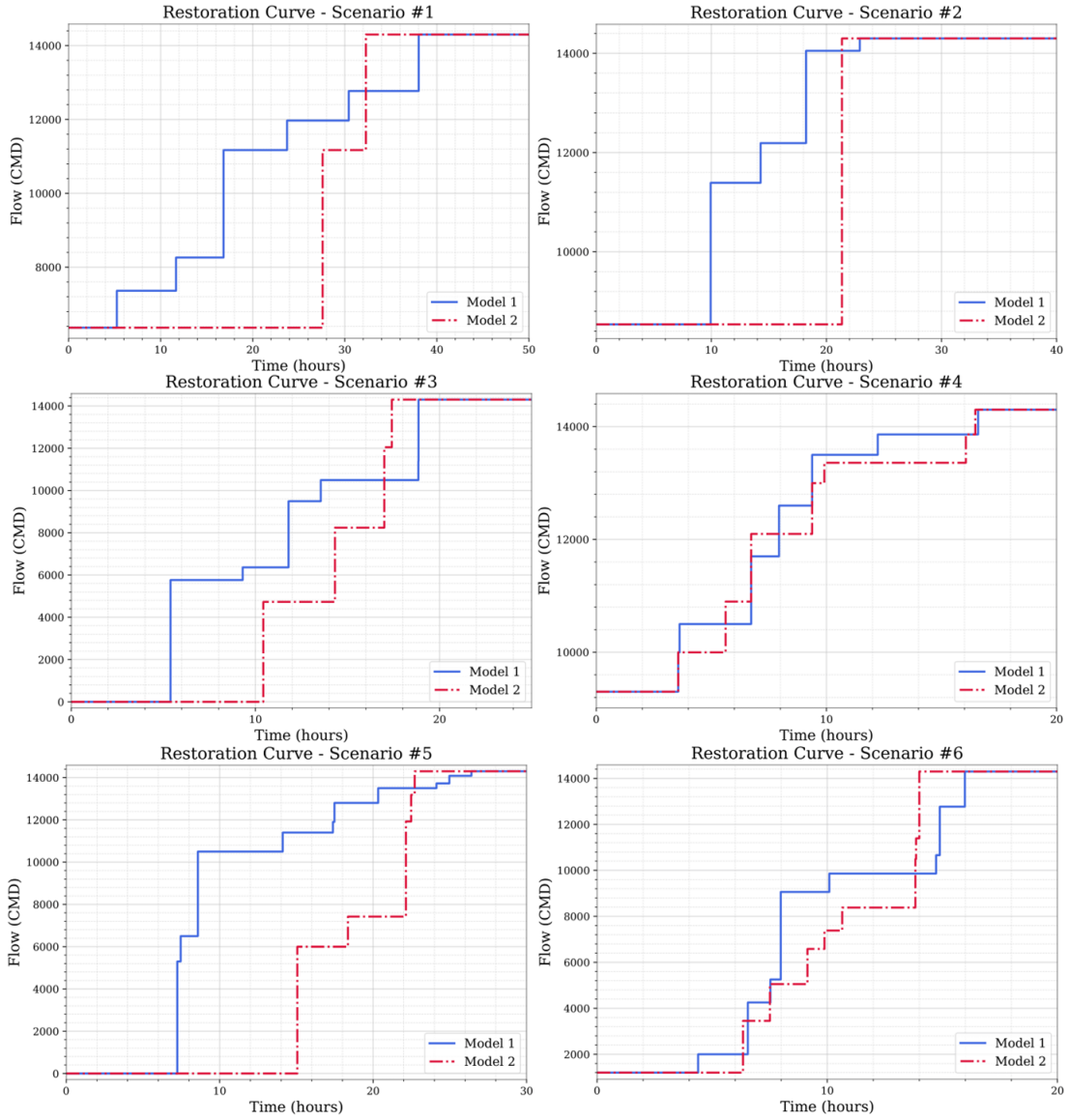


Figure 2.21 – Restoration curves for varying disaster scenarios: case of the of Experiments #3.

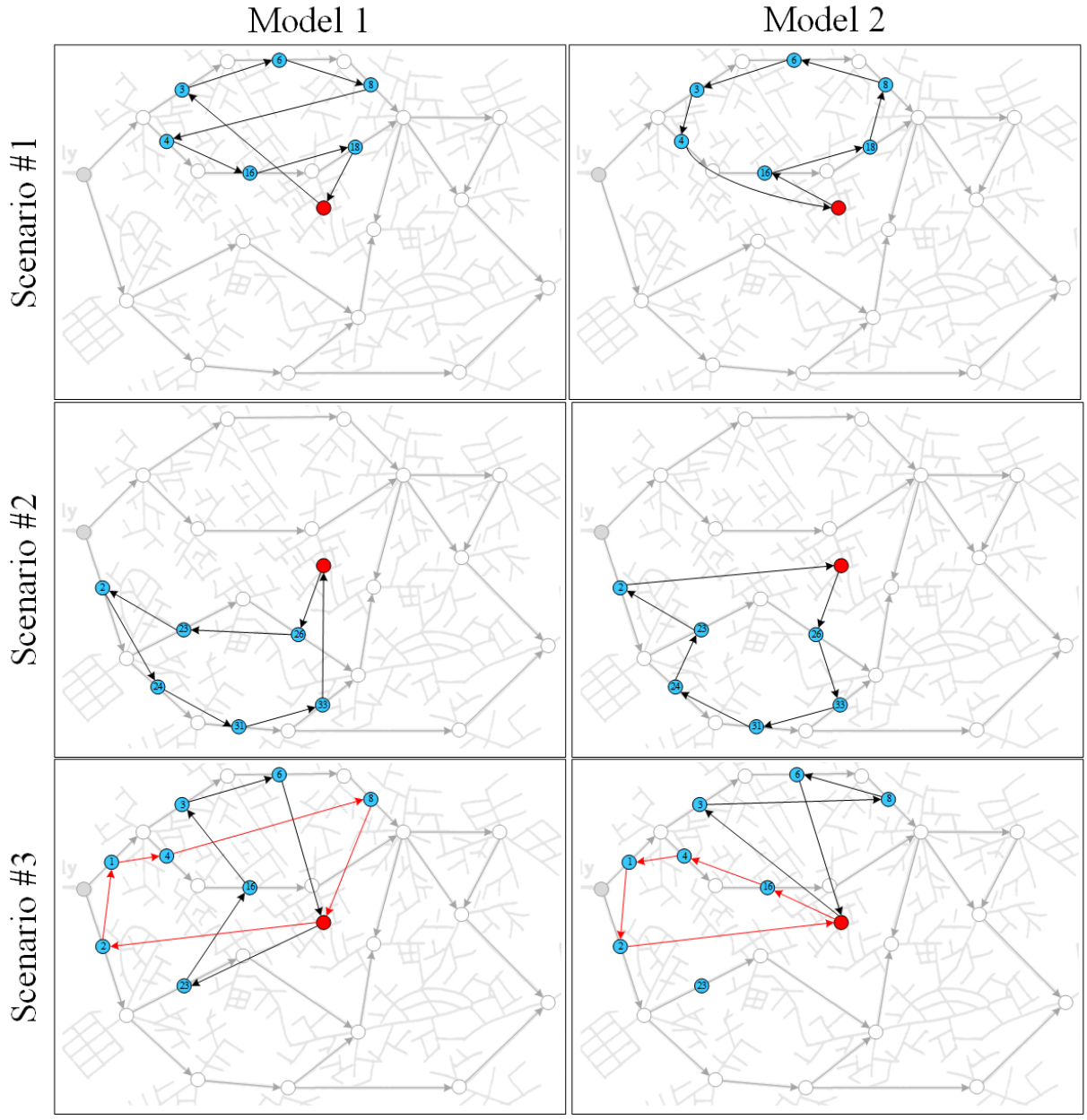


Figure 2.22 – Route selection for disaster scenarios #1-#3.

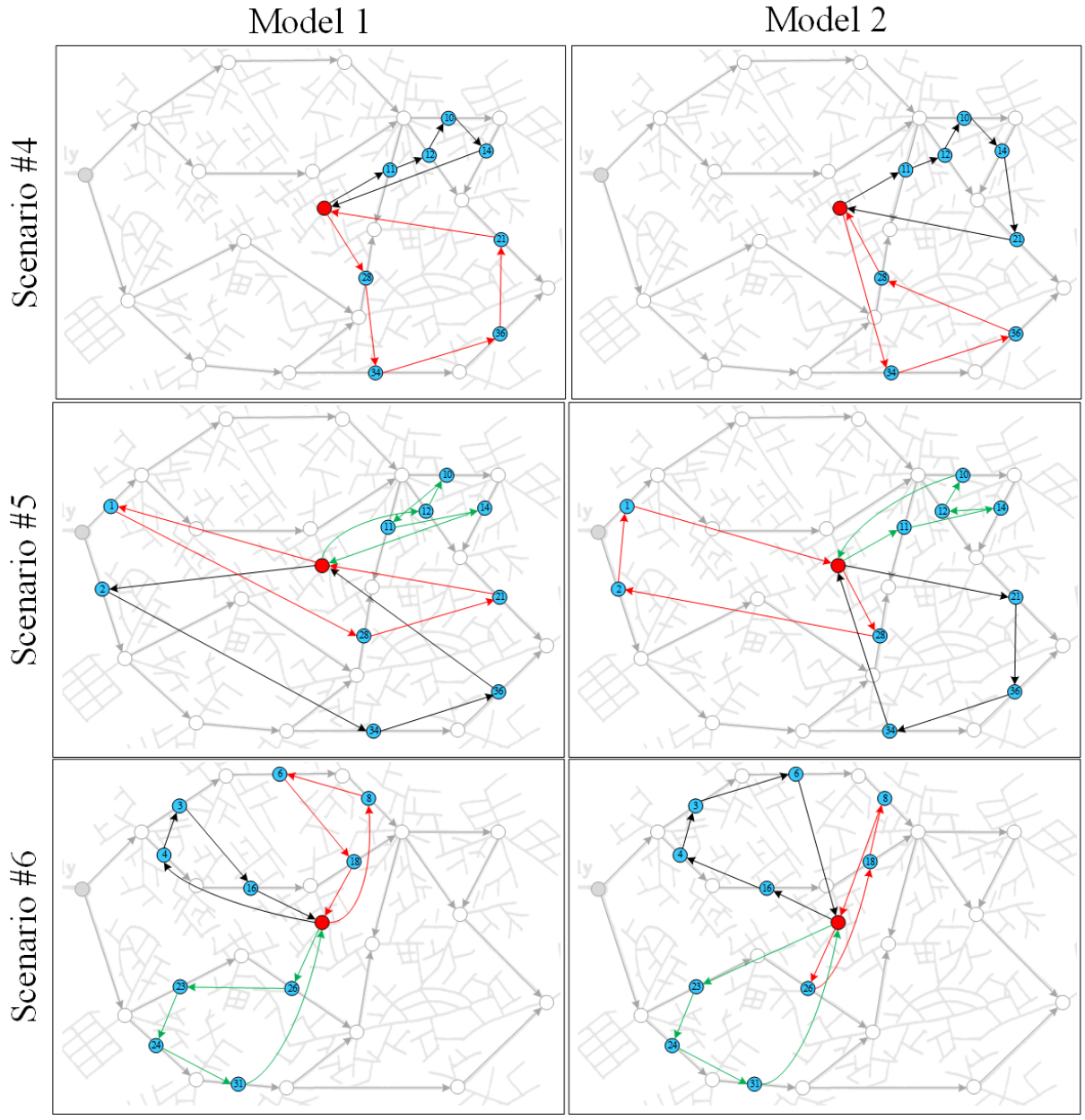


Figure 2.23 – Route selection for disaster scenarios #4-#6.

it is important to provide critical water or power gradually unlike Model 2 that would starve them for longer periods of time and then suddenly a large proportion will get service. This is clearly demonstrated when we examine the second disaster scenario. Using Model 2, the route selected visits component 2 last, resulting in no improvements in the network flow will until this critical component (connected directly to the source) is repaired. Model 1 would be useful for systems such as nuclear plants or large industrial plants where restoring a fraction of a service would not be sufficient to safely restart operation.

Figures 2.22 and 2.23 display the detailed route selections for all six disaster scenarios. From the results shown in these figures, several observations and insights can be made. Firstly, the observed route selections prioritize "important" components over mere efficiency in terms of distance/time traveled. While it might be expected that the shortest paths would be favorable, the choices made reveal a deliberate consideration of certain network components that are of high importance. The restoration of these important components emerges as a priority, as their restoration can considerably impact the flow of the network. This strategic emphasis, at times, takes precedence over the pursuit of optimal distance reduction. Also, when selecting the repair crew routing, the time required to maintain failed network components as well as the planning horizon play a crucial role in the feasibility of a route and also the optimal path that should be taken. Secondly, a principle of proximity-based decision-making can be seen. Within some of the selected routes, there is a tendency for repair crews to address failed components located in close geographical proximity. This tendency aligns with the time-constrained nature of the restoration process, where the need for swift recovery compels the logical grouping of closely situated repairs.

Fourth set of experiments: Comparative analysis of solution methods

The objective of the fourth and final set of experiments is to perform a comparative analysis of the three solution methods: the 2-phase approach, GA and Lagrangian decomposition. This set of experiments considers the network displayed in Figure 2.24, which includes a known source node s, sink node t, and multiple transshipment nodes. The arc index and capacity (in cubic meters per day, CMD) are indicated on each network arc. The network comprises a total of 31 components/arcs. When all components are functioning, the maximum achievable network flow is 112 CMD. With a demand set at $d = 112$ CMD, there are $S = 43$ d-MCs. The component data and travel time matrix are available at https://data.mendeley.com/datasets/2rd628xtbx/1.

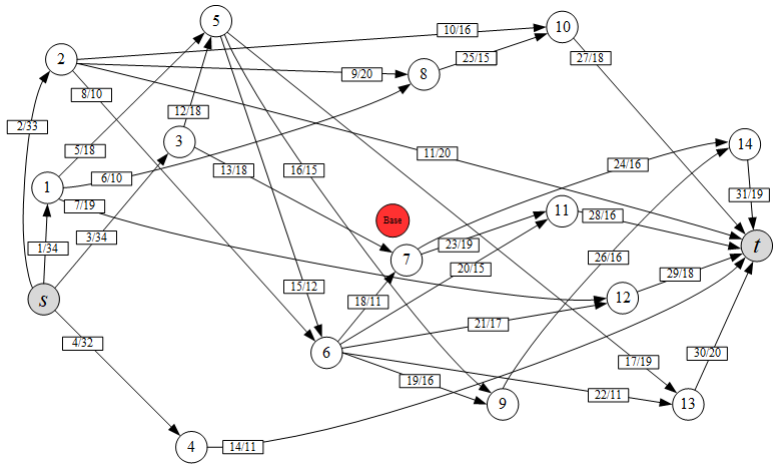


Figure 2.24 – Network considered for fourth set of experiments.

In this set of experiments, several stress events are randomly generated, resulting in the failure of $|\mathcal{I}^F| = 12$ components. The other relevant parameters are set as follows: $R_0 = 0.875$, $D = 75$, $|\mathcal{T}| = 25$, $\Omega = 0.50$, and $K = 3$. For the GA hyperparameters, we use: $n_{\text{gen}} = 5000$, $n_{\text{pop}} = 100$, $n_{\text{elites}} = 1$, $n_{\text{children}} = 70$, $p_{\text{mut}} = 0.075$, and $n_{\text{stagnant}} = 2000$. In the Lagrangian decomposition, the hyperparameters for the GA used to solve the third subproblem are: $n_{\text{gen}} = 500$, $n_{\text{pop}} = 100$, $n_{\text{elites}} = 1$, $n_{\text{children}} = 70$,

$p_{\text{mut}} = 0.100$, and $n_{\text{stagnant}} = 100$.

Table 2.28 displays the results for the three solution methods developed. Both the GA and Lagrangian decomposition solution methods are terminated after 1200 seconds, while a time limit of 1200 seconds is also set for phase II of the two-phase solution approach. In light of the results obtained, the following key observations can be made. The two-phase approach yields the best performance in terms of the objective function value. However, the other solution methods produce solutions that are quite comparable. For some randomly generated stress events, the two-phase approach takes a significant amount of time to generate feasible and non-dominated maintenance patterns. For instance, in the first stress event, it takes almost 30 minutes to generate all feasible patterns. When planning restoration for networks with widespread component failure, generating all patterns may be impractical.

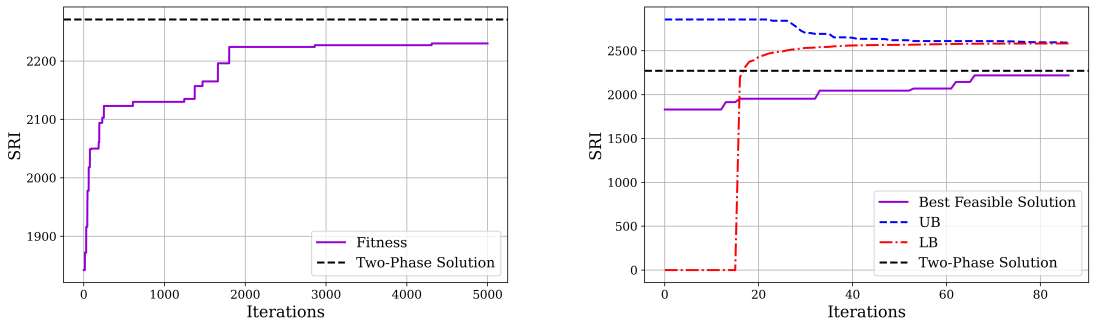
For the first three stress events, the Lagrangian decomposition solution method produces better solutions than the GA. However, in the case of the last stress event, the Lagrangian decomposition method fails to find a feasible solution. This is mainly due to the lower bound of the network reliability provided by the ESP approximation. As pointed out in Jin and Coit (2003), The LQ approximation method, used in both the two-phase approach and the GA, does not provide a lower bound but rather an estimate of the network reliability. Therefore, the LQ approximation method may, in some cases, overestimate/underestimate the network reliability. Is is worth noting that, in our work, one may resort to simulation to compute network reliability. However, doing so would significantly increase computational time, as the proposed solution methods require numerous reliability evaluations.

Table 2.28 – Comparative analysis of solution methods.

		Exact Two-Phase				Genetic Algorithm		Lagrangian Decomposition		
$\lvert\mathcal{I}^F\rvert$	Stress Event	Obj. val	Gap (%)	Phase I - CPUt	Phase II - CPUt	Obj. val	CPUt	Obj. val	Iter	CPUt
12	1	2271	5.76	1722	1200*	2230	1200*	2253	80	1200*
	2	2031	11.30	177	1200*	1911	1200*	2015	150	1200*
	3	2123	6.64	953	1200*	2093	1200*	2043	111	1200*
	4	2263	2.39	334	1200*	2176	1005	–	–	–

Iter: Number of iterations of the Lagrangian decomposition algorithm.

For the first stress event, the evolution of the GA and Lagrangian decomposition over time is shown in Figure 2.25. Figure 2.25a shows the best fitness value/objective function value obtained for each generation using the GA. We observe that the fitness value improves with the number of generations, as the crossover and mutation operations lead to progressively better solutions. Figure 2.25b illustrates the upper and lower bounds, along with the best feasible solution obtained for the Lagrangian decomposition approach across generations. We can see that the upper and lower bounds converge after numerous iterations, with the best feasible solution plateauing at around 50 iterations.



(a) Genetic algorithm.

(b) Lagrangian decomposition.

Figure 2.25 – Evolution of genetic algorithm and Lagrangian decomposition.

Table 2.29 provides a comparative analysis of the GA and Lagrangian decomposition solution approaches on larger problem instances. The two-phase solution approach was excluded from this analysis due to the substantial time required to generate feasible maintenance patterns. Stress events causing the failure of 15 and 20 network components were considered, with time limits set at 1500 seconds for 15 failures and 1800 seconds for 20 failures. The results highlight the advantages of the Lagrangian decomposition approach, which identifies better restoration plans compared to the GA when run for approximately the same amount of time. Table 2.29 also presents the true network reliability achieved. We refer to this as the "true" reliability because the solution methods rely on approximations to estimate reliability. For the final maintenance plan identified using both the GA and Lagrangian Relaxation, the Monte Carlo simulation approach proposed by Ramirez-Marquez and Coit (2005) was used to compute the "true" reliability based on the selected maintenance actions. The values shown in the table represent the mean network reliability ($\bar{R}_T$) and its standard deviation (σ_{R_T}) across 10 simulation runs. Both the GA and Lagrangian decomposition approaches result in true network reliabilities that are very close to the target values.

Table 2.29 – Comparative analysis of genetic algorithm and Lagrangian decomposition.

						Genetic Algorithm				Lagrangian Decomposition				
$\|\mathcal{I}^F\|$	Stress Event	T	$\|\mathcal{T}\|$	R_0	K	Obj. val	CPUt (s)	$\bar{R}_T$	σ_{R_T}	Obj. val	Iter	CPUt (s)	$\bar{R}_T$	σ_{R_T}
15	1	75	25	0.875	3	2117	1500	0.873	1.86E-3	2213	103	1500	0.875	3.43E-3
		90	30	0.900	3	2470	1500	0.898	2.72E-3	2629	88	1500	0.903	3.34E-3
	2	75	25	0.875	3	2055	1500	0.878	3.65E-3	2114	112	1500	0.875	3.05E-3
		90	30	0.900	3	2329	785	0.899	2.45E-3	2504	127	1500	0.899	2.26E-3
20	3	110	40	0.900	3	2819	1800	0.899	2.30E-3	2955	81	1800	0.900	2.19E-3
		110	40	0.900	4	3374	1800	0.898	3.12E-3	3496	107	1800	0.903	2.97E-3
	4	110	40	0.900	3	3225	1800	0.899	2.75E-3	3453	84	1800	0.899	3.71E-3
		110	40	0.900	4	3526	1800	0.899	2.46E-3	3661	102	1800	0.898	1.60E-3

R_T: Network reliability computed using Monte Carlo simulation (Ramirez-Marquez and Coit, 2005).
Iter: Number of iterations of the Lagrangian decomposition algorithm.

2.2.6 Conclusions and Future Research Directions

This section introduced a novel comprehensive framework and a mathematical model aimed at optimizing network resilience in the aftermath of disasters or stress events. The proposed framework presented the concept of durability/viability in the post-recovery stage. The proposed optimization model simultaneously decides the maintenance actions to be performed on failed network components and the optimal routing of multiple repair crews. The model proposed is novel and innovative as it aims to identify an optimal restoration strategy while accounting for the network's performance target R_0 and the availability of multiple maintenance options for failed network components. For smaller problems, a pattern generation algorithm was employed to generate all feasible maintenance patterns and convert the MINLP into a more manageable MILP. For larger problems, two heuristic algorithms were developed: a tailored evolutionary algorithm and a custom Lagrangian decomposition approach. Multiple numerical experiments were run to demonstrate key characteristics of the model and to highlight its value and importance. The model was applied to a real-world water distribution network and it was shown to be capable of efficiently providing optimal maintenance and routing decisions. A second formulation with an objective of minimizing the "Time to Full Resilience" (TFR) was presented and compared to the first. These comparisons demonstrated the advantage of maximizing network resilience. Both heuristic algorithms were shown to provide high-quality solutions within a reasonable computational time, with the Lagrangian decomposition approach typically yielding superior results.

Maintenance managers can gain valuable insights through the optimization model presented in this study. For instance, managers can use the model to facilitate the decision-making process regarding the assignment of maintenance actions to failed components. This decision is based on a careful analysis of the trade-off between the time needed to achieve full resilience and the overall system performance requirement during the mission. Selecting the appropriate maintenance level to carry out on each damaged

component is a crucial decision the maintenance manager must make. The integration of this decision into the proposed optimization model is what makes this work innovative and highly applicable to the restoration activities of real-world plCNI. Moreover, the model can be used to assess and decide on the required number of work crews by comparing recovery trajectories under different disaster scenarios. Such insights enable managers to strategically plan and allocate resources to enhance network resilience and minimize the impact of disruptions.

This work presents many opportunities for further extensions and refinements. The heuristic solution approaches proposed are shown to provide high-quality solutions for large problem instances; however, they can be improved. The Lagrangian decomposition method, which performed best for larger instances, could be enhanced by developing an exact and efficient method to solve the third subproblem. Currently, a metaheuristic is used for this subproblem, which cannot guarantee global optimality. Additionally, finding better initial Lagrangian multipliers could improve the algorithm's convergence. Furthermore, the current approach relies on simulated annealing to find feasible solutions by adjusting the generated routes at each iteration. Developing a heuristic that simultaneously adjusts both routes and maintenance plans could lead to superior solutions. In cases where the ESP lower bound results in poor network reliability estimates, the Lagrangian decomposition method may perform poorly or fail to find a feasible solution. In such instances, using the GA is recommended. Recently, hybrid metaheuristics, which combine multiple algorithms, have shown excellent performance on similar problems and represent another promising area for future research.

This work focused on the optimization of maintenance and routing decisions after a disruptive event (*i.e.,* reactive actions). An interesting and important extension of the work done would be to extend the proposed models to encompass proactive measures, such as hardening decisions and pre-positioning of maintenance crews, as it would provide a more comprehensive perspective on enhancing network resilience. Recognizing and accounting for the inherent uncertainty in restoration activities through the development of stochastic and robust optimization models is another area of future work. In reality, the exact state of a failed network component cannot be known until a maintenance crew arrives and assesses the damage. Developing an optimization approach that treats maintenance and travel time, and maintenance quality as stochastic parameters would be of great interest for industrial practitioners. A second limiting assumption made in this work is that only a single repair crew is needed to repair failed components. It would be interesting to consider the case where some failed components require more than a single repair crew. Other extensions of this work include considering repair crew fatigue and workload balancing. In practice, it would be very common to see repair crews fatigue over time and take longer to perform the selected maintenance plan. In some cases, disruptions can be caused by malicious and ill-intentioned attackers, which can lead to equipment failure. The proposed model in this section is applicable to the case of post-disruption restoration where the attack has been well-identified and interrupted, and the goal is then to recover CNI performance. This is a single period problem with no opportunity for a back-and-forth game to take place between attacker and protector. We are recommending the multi-period game as an extension. Finally, extending the work to CNIs with multiple source and sink nodes is also an important research avenue.

Chapter 3

Predictive Selective Maintenance

This chapter introduces a data-driven FSM framework that leverages DNN to predict component reliability. In Section 3.1, the core framework is presented, which is further extended in Section 3.2 to accommodate fleets performing unique mission types under diverse operational requirements and uncertain maintenance and break durations.

3.1 Optimal Predictive Selective Maintenance for Fleets of Mission-Oriented Systems

In many settings, fleets of assets must perform series of missions with in-between finite breaks. For such fleets, a widely used maintenance strategy is the FSM. Under resource constraints, the FSM problem selects an optimal subset of feasible maintenance actions to be performed on a subset of components to minimize the maintenance cost while ensuring high system reliability during the upcoming mission. The majority of extant FSMP models are focused on traditional physics-based reliability models. With recent advances in ML and DL algorithms, data-driven methods have shown accuracy in predicting RUL. This section proposes a predictive FSM strategy for fleets of complex and large multicomponent systems. It relies on a concurrent ML/DL and optimization approach where a clustering algorithm is used to determine the health states of components and a probabilistic RUL prognostics model is used for component reliability assessment. An optimization model is developed to solve the predictive FSM problem to ensure high reliability of all systems within the fleet. An efficient two-phase solution approach is developed to solve this complex optimization problem. Numerical experiments show the validity of the approach and highlight the improved maintenance plans achieved.

3.1.1 Introduction

The functioning of contemporary society relies heavily on complex and interdependent networks for the production and distribution of goods, energy, and services. These networks must be continuously monitored and proactively maintained to ensure minimal disruptions, exceptional dependability, and a high degree of preparedness to respond to unforeseen events such as pandemics, natural calamities, and weather disturbances. Many of these production and distribution assets can be classified as mission-oriented, meaning they are designed to operate in cycles, alternating between missions and maintenance breaks. SM is a strategy that is applicable to these MOS that aims to identify an optimal set of maintenance actions to be performed on the different system components. The ultimate goal of SM is usually to maximize the system's reliability for the upcoming mission or to minimize maintenance expenses (Diallo et al., 2018).

Industry 4.0 is the latest trend in automation that employs sensors, CPS, and the IIoT to connect the physical and virtual world. As a result of implementing these innovative technologies, predictive maintenance is increasingly being recognized as an appealing maintenance strategy in numerous industries. A key element of predictive maintenance is prognostics, which deals with fault prediction before it occurs

(Jardine et al., 2006). There are two main prediction types in prognostics. The first is RUL prediction, and the second deals with predicting the probability that a component operates without failure up to some future time, *i.e.,* its reliability. Recently, DL algorithms have made tremendous strides in performance and have improved the state of the art in prognostics and predictive maintenance (Namuduri et al., 2020). Effective predictive maintenance plans have the capability of reducing maintenance cost, extending equipment life, reducing machine downtime, and improving production. Prognostic methodologies encompass traditional ML techniques, including logistic regression, support vector machines, and ensemble methods, as well as DL techniques, such as multilayer perceptron models, plCNN, and plRNN. DL algorithms require tens to hundreds of thousands of training samples to achieve a desirable level of accuracy. IIoT sensors can produce massive amounts of data thus enabling DL algorithms to be applied. ML and DL methods can identify a relationship between sensor values and dependent variables such as time of failure, future sensor value, or the detection of anomalies.

Almost all SMP models proposed in the literature are based on statistical approaches for reliability assessment. Components are modeled using lifetime distributions and the hazard rate function is used to describe the aging process. A limitation of using lifetime distributions is the assumption that the failure rate of a system depends only on time and a limited number of environmental factors/co-variates. However, many industrial systems degrade or deteriorate with both time, usage and multiple other factors. It is therefore more realistic to base components failure on their health condition (Khatab et al., 2018a). The application of new technologies such as wireless sensors and IIoT, and the recent advancements in AI, have led to data-driven methods based on DL algorithms to become very popular for RUL prediction and reliability assessment. Two major advantages of utilizing these data-driven approaches are: (i) the current health condition of the component is considered and (ii) they do not rely on knowing the physical nature of the degradation mechanism model. These data-driven approaches utilize data collected from similar systems to train a model, from which the RUL or reliability can be predicted (An et al., 2018).

In this section, a data-driven predictive SM strategy is proposed for fleets of MOS, where a concurrent DL and optimization approach is developed to identify optimal maintenance decisions. The strategy is referred to as PdSM because a DL approach is utilized to predict a RUL distribution, an approach that is commonly used in predictive maintenance for machine prognostics. The first PdSM framework was introduced by Hesabi et al. (2022) where a probabilistic classifier is used for reliability prediction and a basic SM optimization model is used to identify maintenance decisions. The authors consider a relatively small series-system and utilize a full enumeration approach to solve the optimization problem. For large and complex systems comprised of many components and subsystems, their proposed solution method would be unable to provide optimal decisions within a reasonable time frame. This highlights the need to develop a more intelligent and efficient solution approach. In the work of Hesabi et al. (2022), it is assumed that maintenance actions can restore a component to a previous state where the RUL is higher, and more expensive maintenance actions lead to a better RUL enhancement. In practice, the level of RUL improvement of a given maintenance action is an extremely difficult if not impossible quantity to determine.

The goal of this section is to extend and improve upon the PdSM approach proposed by Hesabi et al. (2022). Rather than a basic series system, we consider fleets of systems each composed of multiple components connected according to complex reliability structures. The present work also considers the impact of component degradation level on the time required to perform maintenance. Such a consideration is necessary to meet practical maintenance plans requirements. Unlike the work of Hesabi et al. (2022), the optimization model developed here not only identifies the optimal set of components to be maintained, but also jointly provides optimal maintenance task assignment decisions for repairmen. To overcome the limitations of the full enumeration solution method used in (Hesabi et al., 2022), the present section develops an efficient solution method based on a two-phase decomposition approach. Finally, rather than using a probabilistic classifier to predict component reliability, a DL algorithm combined with Monte Carlo Dropout MCD is used. The advantages of this approach will be highlighted and discussed later in this section.

This work is also a major extension of its preliminary version published in O'Neil et al. (2022a).

Compared to O'Neil et al. (2022a), the problem is extended to a fleet of MOS, a new prediction framework is proposed, new DL methods are applied and a new BIP model is developed. The remainder of this section is structured around five additional subsections. Subsection 3.1.2 provides a review examining papers related to selective and predictive maintenance, with an emphasis on the utilization of DL techniques for the purpose of prognostics. The system investigated is presented and fully described in Subsection 3.1.3 along with the notation and the main working assumptions used. Subsection 3.1.4 presents the concurrent DL and optimization approach used to solve the fleet predictive SM strategy. The clustering algorithm used for health state identification and the probabilistic DL classifier used for reliability assessment are both discussed. The FSMO model and the solution method are also presented in Subsection 3.1.4. In Subsection 3.1.5, various numerical experiments are carried out to validate the ML and DL models and algorithms, and demonstrate the effectiveness and the accuracy of the proposed PdSM approach. This Subsection also includes a description of the dataset used to train and test the ML and DL algorithms, the data preprocessing steps, and the evaluation metrics used to evaluate the effectiveness of the models. Conclusions and future research extensions are presented in Subsection 3.1.6.

3.1.2 Literature Review

This review section deals with three topics related to the objective of the proposed maintenance strategy: the SMP, DL methods for prognostics and state of the art (SOTA) data-driven maintenance strategies.

The SMP

Many military and industrial systems are designed to complete a sequence of missions with finite breaks in between. Maintenance is necessary during these breaks to ensure these systems perform acceptably during missions. Maintenance actions range from mR for failed components to perfect replacement. Resource constraints, such as time, cost, and availability of repairpersons, dictate that only a subset of system components can be selected for maintenance. The goal of the SMP is to determine the optimal subset of components and maintenance actions that will minimize maintenance cost subject to a reliability requirement being met for the upcoming mission(s), given these resource constraints.

The original SM model introduced in Rice et al. (1998) dealt with an elementary series system comprised of components with exponentially distributed lifetimes. The only maintenance action considered was the replacement of failed components. An enumeration approach was devised to identify the optimal maintenance decisions that would maximize the system reliability for the upcoming mission. Many researchers have since expanded upon the original SMP to include complex system configurations (Cassady et al., 2001b; Diallo et al., 2018), multistate systems (Liu and Huang, 2010; Pandey et al., 2013), component dependence and failure propagation (Xu et al., 2016; Dao and Zuo, 2017a; Maaroufi et al., 2013), fleet level SM (Khatab et al., 2020; Schneider and Cassady, 2015), multiple repair channels (Khatab et al., 2018b; Diallo et al., 2019), stochastic mission, break, and maintenance durations (Khatab et al., 2017a,b; Liu et al., 2018; Al-Jabouri et al., 2024a), stochastic maintenance quality (Khatab and Aghezzaf, 2016; Duan et al., 2018), and joint SM and orienteering (O'Neil et al., 2023).

There have been significant improvements in the solution methods and approaches used to solve the SMP. Rajagopalan and Cassady (2006) proposed refinements to the total enumeration approach developed by Rice et al. (1998). Pandey et al. (2013) utilize a differential evaluation approach as a solution method, multiple studies (Dao et al., 2014), Chaabane et al. (2020) use plGA as a solution technique, and O'Neil et al. (2022b) propose a hybrid CG and GA. Al-Jabouri et al. (2023b) propose a branch-and-price approach to solve large-scale instances of the SMP. In addition, several papers have used reinforcement learning to handle this complex optimization problem (Achamrah and Attajer, 2024; Liu and Qian, 2022; Liu et al., 2020b). Most of the papers reviewed thus far focus on simple series-parallel system architectures, although more complicated structures, such as k-out-of-n:G, are common in various industrial applications. Diallo et al. (2018) is the first to examine the SMP in the context of serial k-out-of-n:G systems. The authors propose two new nonlinear SMP formulations. To solve the resulting SM optimization problems, a two-phase approach using a BIP model is developed: the first phase generates all feasible maintenance patterns for each k-out-of-n:G subsystem, while the second phase solves a multidimensional

multiple-choice knapsack problem to select the optimal mix of maintenance patterns. Recent literature reviews on the SMP can be found in Al-Jabouri et al. (2022) and Al-Jabouri et al. (2024b).

The majority of SMP models proposed in the literature use probability distributions to model component lifetimes, where the hazard rate function is used to describe the aging process. A limitation of using lifetime distributions is the assumption that the failure rate of a system depends only on time, when in reality many systems degrade with both time and usage. Khatab et al. (2018a) introduced the first SM strategy that considers the actual health condition of components, and refer to the strategy as condition-based SM. The authors assume that each component degrades following a stochastic process and the actual health condition of a component is revealed only through periodic inspections performed during a mission. Hesabi et al. (2022) introduce a predictive SM framework, where system components are assumed to be continuously monitored by multiple sensors. A trained DL algorithm is used to predict the mission failure probability of each component. Meanwhile, only a simple series system is considered and a basic solution strategy is used to solve the optimization problem. As was previously stated, for large and complex systems, the full enumeration approach proposed by Hesabi et al. (2022) would be unable to provide solutions within a reasonable time frame. Thus, there is a need to develop predictive SM approaches that are capable of efficiently solving systems of practical size.

Data-driven maintenance strategies

Prognostics is a central aspect of predictive maintenance and usually deals with the prediction of RUL. Different techniques have been proposed for RUL prediction such as physics based, data-driven, and hybrid methods. In recent years, however, there has been a surge in the use of data-driven methods based on DL techniques. These methods have shown exceptional performance in their ability to learn patterns and useful information from multi-variate sensor data and text data from maintenance logs to make accurate predictions (Zhang et al., 2024b; Fu et al., 2024; Usuga-Cadavid et al., 2022). This review section will focus on data-driven prognostic models and the SOTA data-driven maintenance strategies.

Li et al. (2018) propose a new data-driven approach for prognostics that implements a deep CNN for RUL prediction of turbofan engines. The DCNN is compared to other popular approaches like the multi-layer perceptron (MLP) and the LSTM network. The proposed DCNN was shown to outperform the other algorithms in terms of both root-mean-square error (RMSE) and a popular scoring function employed by the International Conference on Prognostics and Health Management Data Challenge. Yang et al. (2019) propose a double-CNN architecture for the purpose of RUL prediction. The first CNN is used to identify the incipient fault point and the second is used for RUL prediction. Li et al. (2020) propose a multi-scale CNN (MS-CNN) and multiple recent studies implement temporal convolutions with attention mechanisms (Zhang et al., 2024b; Fu et al., 2024).

The application of RNNs for RUL prediction has also been studied extensively in the literature. Zhang et al. (2018) utilize a LSTM neural network for RUL prediction of Lithium-Ion batteries. The experiment results showed that the LSTM generally predicts more accurate RULs than simple RNNs and support vector machines. There have also been studies that have used bi-directional LSTMs Zhang et al. (2018) and convolutional bi-directional LSTMs Zhao et al. (2017) for RUL prediction. Su et al. (2020) propose an integrated Variational Autoencoder (VAE) and time-window-based sequence neural network (twSNN) approach for RUL prediction. The VAE is used to reduce data dimension and extract meaningful data features; the twSNN is then used to predict the RUL. Song et al. (2022) define the problem of RUL estimation as a bi-level prediction problem. The lower-level uses an LSTM to predict the time-series sensor data in the next several time-steps. The upper-level takes in the predicted future time-series data as input and uses an LSTM to estimate the RUL. The proposed method was shown to outperform state-of-the-art DL and ML algorithms. Ruan et al. (2021) compare CNNs, LSTMs, and a combined CNN-LSTM architecture for predicting RUL of turbofan engines. The results obtained show that the LSTM network with 2 LSTM layers and 2 dense layers performs the best in terms of RUL prediction. Other studies that propose hybrid approaches incorporating multiple DL algorithms can be found in Niu et al. (2019), Li et al. (2019a) and Ren et al. (2020).

The aforementioned studies are mainly focused on RUL point prediction. However, the focus of the present work lies in the application of the prognostics results to guide and inform decisions related to SM planning. Dealing with the recent data-driven maintenance approaches, Nguyen and Medjaher (2019) use an LSTM classifier to predict the probability that the RUL of a system falls within certain time intervals. These reliability predictions are then used to optimize maintenance and spare parts ordering decisions. Zhou et al. (2023b) propose a PdSM strategy for multi-state series-parallel systems using an adaptive C-Transformer for RUL prediction, and a CNN for reliability prediction. Dehghan Shoorkand et al. (2023) and Shoorkand et al. (2024) develop integrated predictive maintenance and production planning frameworks that leverage DL algorithms for component reliability prediction and mathematical programming.

All the papers mentioned above rely on training a neural network for the task of predicting failure probability within a given time-window or mission length. The main drawback of these approaches is the need to train different models for varying window-sizes or mission lengths, resulting in a lack of flexibility in handling different operational scenarios without having to retrain a new model. In the literature, several works have been proposed for data-driven maintenance using RUL predictions instead of reliability predictions (Chen et al., 2021b; de Pater et al., 2022; Zhang and Zhang, 2020). However, a shortcoming of these works resides in their inability to consider the uncertainty inherent to RUL predictions. Chen et al. (2022) propose an approach for RUL uncertainty quantification using a Local Uncertainty Estimation (LUE) model together with a Bi-LSTM. This approach allows to construct a RUL prediction interval from which a RUL probability distribution is derived and used to deal with maintenance optimization. Zeng and Liang (2023) propose a predictive maintenance framework that implements probabilistic RUL prognostics into maintenance planning using a Gaussian distribution to model prediction uncertainties. Rather than training a model solely for RUL point prediction, the model proposed in Zeng and Liang (2023) is designed to predict the mean and standard deviation of a Gaussian distribution, thus providing a distribution of the RUL. An integer program is then proposed to optimize the maintenance planning problem. Recently, several other works (Zhuang et al., 2023; Lee and Mitici, 2023; Mitici et al., 2023) have implemented MCD to predict the RUL probability distribution. MCD is a computationally efficient method for estimating the uncertainty in neural network outputs (Abdar et al., 2021). Thus, a DL algorithm combined with MCD will be used to predict the RUL probability distribution. MCD is known to be biased but Puzyrev and Duuring (2025) have shown that the bias is negligible for large scale datasets as is the case here.

Contributions

The main contributions are as follows:

— A novel PdSM approach is developed for a fleet of systems with multiple components connected via complex reliability architectures. The impact of component degradation on maintenance time is accounted for. The approach is designed to jointly provide optimal maintenance actions and repairperson task assignments decisions.

— Rather than using a probabilistic classifier for reliability prediction as in Hesabi et al. (2022), an advanced DL algorithm with MCD is implemented. This reliability prediction approach holds several key advantages including more accurate predictions in terms of mission success probability and mission length variability. Furthermore, the ability to predict a RUL distribution allows for the development of a more realistic and practical data-driven SM model as it takes into account the impact of component failure in the optimization process.

— In contrast to the full enumeration solution method used in Hesabi et al. (2022), a more efficient solution method is proposed on the basis of decomposition techniques.

— In contrast to most existing data-driven maintenance approaches, the approach developed in this section leverages a predicted RUL distribution to compute both component reliability and expected downtime during the upcoming mission. The predicted reliability of components is used to accurately determine the overall system reliability. Indeed, calculating the expected component downtime allows to guide the maintenance selection process toward components that are likely to fail early in the mission, thus contributing to an improved fleet performance.

— This section provides a comparative analysis of different methods used for reliability prediction.

This comparison focuses specifically on different recent DL techniques as well as traditional statistical approaches. The advantage of using DL models to make predictions based on the current condition is clearly demonstrated. The results also highlight the advantages and disadvantages of using predicted RUL distributions opposed to point predictions or classification models for reliability assessment.

— Finally, the proposed predictive models are general and can indeed be applied to any MOS such as production lines in manufacturing, turbines in hydroelectricity generation, and turbofan engines in aerospace to name just a few.

3.1.3 System Description

The following notation system is used.

I	Number of systems within the fleet
i	Index of systems, $i = 1, ..., I$
N_i	Number of subsystems within system i
j	Index of subsystems in system i, $j = 1, ..., N_i$
M_{ij}	Number of components in subsystem j of system i
k	Index of components in subsystem j of system i, $k = 1, ..., M_{ij}$
E_{ijk}	The kth component of subsystem j of system i
Π_{ijk}	The penalty cost per unit time for component E_{ijk}
L_{ijk}	The highest maintenance level available for component E_{ijk}
l	Index of maintenance levels for E_{ijk}, $l = 1, ..., L_{ijk}$
Q	Number of repairpersons
r	Index of repairpersons, $r = 1, ..., Q$
t^p_{ijkl}	Time to perform PM of level l on component E_{ijk}
t^c_{ijkl}	Time to perform CM of level l on component E_{ijk}
c^p_{ijkl}	Cost of PM of level l on component E_{ijk}
c^c_{ijkl}	Cost of CM of level l on component E_{ijk}
c^f_r	Fixed cost of hiring or using repairperson r
c^v_r	Variable cost rate of using repairperson r
D	Maintenance break duration
U	Mission duration
$\mathcal{R}_{ijk}$	Reliability of component E_{ijk} during the next mission
$\mathcal{R}_{ij}$	Reliability of subsystem j of system i during the next mission
$\mathcal{R}_i$	Reliability of system i during the next mission

Main working assumptions

1. The systems and components degrade with both age and usage (Khatab et al., 2018a).
2. During the maintenance break, the systems are assumed to be switched off and therefore not experiencing any degradation. This is reasonable and is commonly used (Diallo et al., 2018).
3. Maintenance activities are allowed only during the break duration. This is also reasonable and commonly assumed (Khatab et al., 2018a, 2020)
4. After a maintenance action is performed, the component is brought back to an AGAN state.
5. The time and cost of a maintenance action are tightly dependent on the degradation level of the component to be maintained. A maintenance action carried out on a component with higher degradation level will induce a higher cost and require longer time.
6. Individual components contribute to the overall performance of the system. Accordingly, the failure of a component is penalized at a cost proportional to its duration as such failures lead to a decrease in system productivity or efficiency. This is a reasonable assumption that aligns with existing literature (Khatab et al., 2018a).

Fleet description and reliability computation

A fleet of I multi-component systems is investigated in which components and systems are binary. Each system i ($i = 1, \cdots, I$) is comprised of N_i subsystems in series, and each subsystem is composed of M_{ij} statistically independent components E_{ijk} ($j = 1 \cdots N_i; k = 1 \cdots M_{ij}$). We consider K_{ij}-out-of-M_{ij}:G subsystems, which implies that for the subsystem to operate correctly, at least K_{ij} out of the M_{ij} constituent components must be in working order. Finally, we assume components degrade/deteriorate with both operational time and usage.

The fleet of systems being examined is required to operate in sequences, alternating between active missions and maintenance breaks. It is assumed that the fleet has recently accomplished a mission and is entering the first maintenance break. During the break, all systems of the fleet are switched off for a period D, during which maintenance actions can be performed on their components. There is a time and cost associated with performing maintenance actions. To monitor the degradation of components, a common set of S sensors are available. New sensor measurements are recorded after the completion of each operating cycle. An operating cycle may refer to a defined period of time or a set of operations that must be executed. For example, the lifetime of a jet engine can be measured in flight cycles, where a flight cycle refers to all operational activities from take-off to landing. The overall degradation measurements corresponding to component E_{ijk} are represented as a matrix $\mathbf{G}_{ijk}$:

$$\mathbf{G}_{ijk} = \begin{bmatrix} \boldsymbol{g}^1 \\ \boldsymbol{g}^2 \\ \vdots \\ \boldsymbol{g}^{T_{ijk}} \end{bmatrix} = \begin{bmatrix} g_1^1 & g_2^1 & \cdots & g_S^1 \\ g_1^2 & g_2^2 & \cdots & g_S^2 \\ \vdots & \vdots & & \vdots \\ g_1^{T_{ijk}} & g_2^{T_{ijk}} & \cdots & g_S^{T_{ijk}} \end{bmatrix},$$

where $\boldsymbol{g}^t$ is a vector of all sensor measurements during cycle t ($t = 1, \ldots, T_{ijk}$), and T_{ijk} refers to the number of cycles that component E_{ijk} has completed.

When the fleet starts the maintenance break, the sensor data and the number of cycles completed by components are available. The state of each component S_{ijk} is known and defined as follows:

$$S_{ijk} = \begin{cases} 1 & \text{if component } E_{ijk} \text{ is functioning at the start of the break,} \\ 0 & \text{otherwise.} \end{cases} \tag{3.1}$$

By first training a DL algorithm on the historical sensor data, the RUL of each component E_{ijk} can be evaluated and compared with the next mission duration U. Accordingly, the probability $\mathcal{R}_{ijk}$ of component E_{ijk} to successfully achieve the next mission can be determined; this probability is the reliability of component E_{ijk} for the subsequent mission. Similarly, the reliability $\mathcal{R}_i$ of system i is also defined as the probability that it can successfully complete the upcoming mission. The reliability $\mathcal{R}_i$ depends on both the system structure and the corresponding component reliability functions $\mathcal{R}_{ijk}$. Given that systems are comprised of N_i k-out-of-n:G subsystems in series, the reliability $\mathcal{R}_i$ is then computed as:

$$\mathcal{R}_i = \prod_{j=1}^{N_i} \left[\sum_{e_{k_{ij}}=1}^{M_{ij}} \sum_{e_{k_{ij}-1}=1}^{e_{k_{ij}}-1} \cdots \sum_{e_1=1}^{e_2-1} \left(\prod_{v \in \{e_1, \cdots, e_{k_{ij}}\}} \mathcal{R}_{ijv} \right) \left(\prod_{\substack{u=1 \\ u \notin \{e_1, \cdots, e_{k_{ij}}\}}}^{e_{k_{ij}}} (1 - \mathcal{R}_{iju}) \right) \right]. \tag{3.2}$$

The exact formulation of the k-out-of-n:G subsystems was first presented by Arulmozhi (2002). Efficient algorithms like the one proposed by Kuo and Zuo (2003) exist to solve this complex equation.

In this section, we consider that a component subjected to a maintenance action becomes AGAN. There are q repair crews to whom maintenance actions are assigned with fixed cost c_r^f and variable cost rate c_r^v ($r = 1, \cdots, Q$). A maintenance action carried out on component E_{ijk} requires t_{ijk} units of time. In the proposed PdSM approach, maintenance time is considered as dependent on the component degradation level. This is a reasonable assumption that aligns with previous work (Khatab et al., 2018a;

Tai and Chan, 2010). Components that are in more degraded states are likely to require more thorough inspections and repairs/replacements. The effective maintenance time estimation will be further discussed in the subsequent section.

3.1.4 PdSM Optimization Approach

In this section, the proposed predictive SM approach is developed and fully discussed. The general principle of this approach is summarized in Figure 3.1 and mainly relies on three phases: offline, online, and maintenance decision making phases. During the offline phase, a FCM algorithm and Bi-LSTM are trained on historical components run-to-failure data. Before the training procedure for the unsupervised and supervised algorithms are initiated, data preprocessing and denoising steps are conducted.

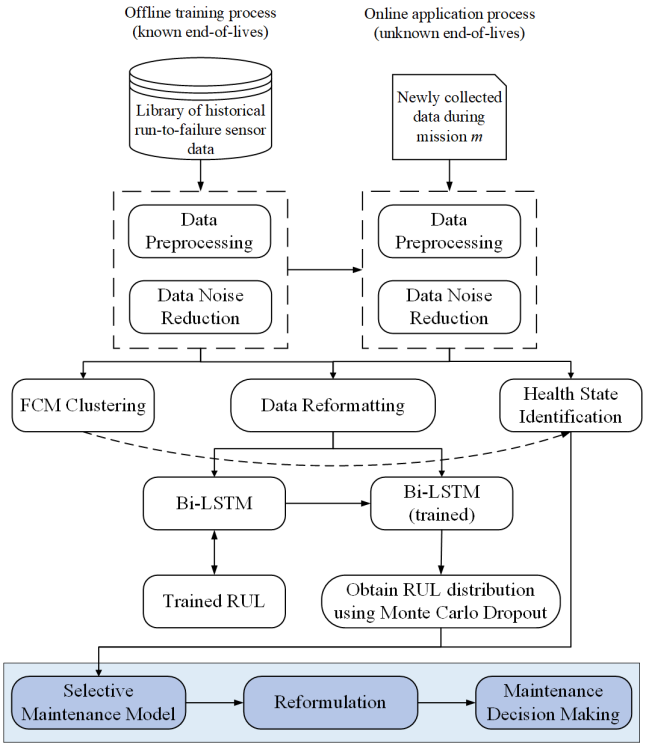


Figure 3.1 – Flowchart of the proposed PdSM framework for fleets of MOS.

After the completion of the preprocessing steps, the unsupervised clustering algorithm can be trained on the processed data. The FCM clustering algorithm is used to partition the historical run-to-failure data into discrete health states. The output of the clustering algorithm are cluster centroids which can be used to identify the health state of components that are currently in operation. This is a crucial step in the proposed predictive SM approach as maintenance time depend on component's degradation level.

Before initiating the training process for the DL model, the historical components run-to-failure dataset is reformatted into 3-dimensional tensors of the form (N_s, N_{ts}, N_f), referring to the number of samples, the time sequence dimension, and the number of selected features, respectively. Unlike probabilistic classifier-based approaches (Hesabi et al., 2022), the approach developed in this section adopts a DL model with MCD to predict a RUL distribution. Reliability computation involves conducting $\mathcal{M}$ forward passes of the trained network and generating $\mathcal{M}$ unique RUL predictions. The probability that a component will successfully complete the upcoming mission of length U can then be determined by assessing the proportion of these predictions that are greater than U.

The second phase of the proposed approach (*i.e.,* the online phase) deals with the application of the

trained DL model and clustering algorithm for online prognostics and health state identification. At the start of the maintenance break, the collected real-time monitoring data for each component are used as input to the trained DL model to predict a RUL distribution. The predicted distribution is used to compute the reliability and the expected down time of each component for the next mission. The centroids identified through training the FCM clustering algorithm serve to identify the current health state of each system component. It should be noted that the second phase does not repeat the time-consuming training process which reduces the time of failure prediction and health state identification.

The final phase of the proposed PdSM approach utilizes the predicted failure probability and expected downtime along with health state for each component as input data for maintenance decision making. Reliability, expected downtime and degradation levels of components are used as input to a FSMO model to select the components that should be maintained in a fleet of heterogeneous systems with the objective of minimizing the total maintenance cost while ensuring that each system in the fleet operates at a required minimum reliability level. The following subsections provide a detailed description of the three phases of the proposed PdSM approach.

FCM clustering for health state identification

The FCM clustering algorithm is implemented in this work to determine the hidden structure within the historical component failure data representing states of degradation, ranging from normal to near failure. FCM is trained in an unsupervised way and aims to find similarities within the data. In the literature, FCM has been successfully implemented for the purpose of health state division (Chen et al., 2021a,b; Wang et al., 2019; Javed et al., 2015). The main advantage of FCM algorithm over hard clustering algorithms is its ability to allow gradual memberships of data points to clusters measured as degrees in the interval $[0, 1]$. This gives the flexibility to express the data points as belonging to more than one cluster (Bora et al., 2014). After training the FCM algorithm on historical component run-to-failure data, the identified cluster centroids can then be used to estimate components health state membership.

The objective of the following FCM algorithm is to find the optimal cluster centers and components membership degrees such that the following constrained optimization problem is minimized:

$$\min \quad \mathcal{J} = \sum_{c=1}^{\mathcal{C}} \sum_{i=1}^{\mathcal{M}} \sum_{j=1}^{T_i} u_{cij}^m \cdot d^2\left(\boldsymbol{g}_i^j, w_c\right) \tag{3.3}$$

$$\text{s.t.} \quad \sum_{c=1}^{\mathcal{C}} u_{cij} = 1, \quad \forall i, j \tag{3.4}$$

$$u_{cij} \in [0, 1], \quad \forall c, i, j \tag{3.5}$$

where m is the degree of fuzziness, u_{cij} is the membership of sample $\boldsymbol{g}_i^j$ to cluster w_c, $\mathcal{M}$ is the number of training components, T_i is the time to failure (in cycles) of training component i, $\mathcal{C}$ is the number of clusters, and $d^2(\boldsymbol{g}_i^j, w_c)$ is the euclidean distance between sample $\boldsymbol{g}_i^j$ and cluster w_c. Elements of vector $\boldsymbol{g}_i^j$ refer to all sensor measurements for training component i during cycle j. The above optimization problem is solved by converting it into an unconstrained problem using Lagrange multipliers. An iterative approach is then used where firstly, the cluster centers are fixed to find the membership degrees; secondly, the membership degrees are fixed to find the cluster centers (Benaichouche et al., 2013). The two steps are repeated until a convergence criteria is met.

Once the optimal cluster centroids have been found through training, the membership degree of operating component E_{ijk} to each cluster (or health state) is computed using equation 3.6.

$$u_{c,E_{ijk}} = \frac{1}{\sum_{v=1}^{\mathcal{C}} \left(\frac{d^2\left(\boldsymbol{g}^{T_{ijk}}, w_c\right)}{d^2\left(\boldsymbol{g}^{T_{ijk}}, w_v\right)} \right)^{\frac{2}{m-1}}}, \quad \forall c \in \{1 \cdots \mathcal{C}\}. \tag{3.6}$$

The expected times to perform a PM (t^p_{ijkl}) and a CM (t^c_{ijkl}) action on component E_{ijk} are then computed as:

$$t^p_{ijkl} = \sum_{c=1}^{\mathcal{C}} u_{c,E_{ijk}} \cdot t^p_{cl} \tag{3.7}$$

and

$$t^c_{ijkl} = \sum_{c=1}^{\mathcal{C}} u_{c,E_{ijk}} \cdot t^c_{cl} \tag{3.8}$$

where t^p_{cl} and t^c_{cl} are respectively the expected times to perform a PM and a CM of level l on components whose health states belong to cluster w_c.

Probabilistic DL approach for RUL and reliability prediction

The LSTM network is a type of RNN that is designed to overcome the vanishing and exploding gradient problem frequently encountered when training standard RNNs. LSTMs are particularly well-suited for processing sequential or time-series data, where the network needs to retain information about past events to make accurate predictions. At the core of an LSTM network are "memory cells", which are connected via "gates" that allow the cells to selectively add or remove information. In an LSTM, there are three types of gates:

— The forget gate, comprised of the first sigmoid layer in each unit, determines what information from the previous cell state should be discarded. It does this by taking as input the previous hidden state h_{t-1} and the current input g_t, passing them through a sigmoid function, and generating a vector f_t of values between 0 and 1. The vector f_t is then element-wise multiplied by the previous cell state to "forget" the irrelevant information. The vector f_t is given by:

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, g_t] + b_f\right). \tag{3.9}$$

— The input gate, formed by the subsequent sigmoid and *tanh* layers, determine the new information that will be added to the cell state. The sigmoid layer produces a vector i_t of values between 0 and 1 to determine which values in the cell to update. The *tanh* layer produces a vector s_t of new candidate values. These two vectors can be mathematically expressed as:

$$i_t = \sigma\left(W_i \cdot [h_{t-1}, g_t] + b_i\right) \tag{3.10}$$

$$s_t = tanh\left(W_s \cdot [h_{t-1}, g_t] + b_s\right). \tag{3.11}$$

— Finally, the output gate determines what information should be provided as input to the next hidden state. It takes as input the current hidden state and the current input, passes them through a sigmoid function to generate a vector o_t of values between 0 and 1. The vector o_t is multiplied by the hyperbolic tangent of the current cell state to produce the new hidden state h_t. The vector o_t and h_t are calculated as follows:

$$o_t = \sigma\left(W_o \cdot [h_{t-1}, g_t] + b_o\right). \tag{3.12}$$

$$h_t = o_t \cdot tanh(c_t). \tag{3.13}$$

The Bi-LSTM network, which comprises two LSTM cells, is an extended version of the standard LSTM. It has demonstrated exceptional performance in predicting RUL (Chen et al., 2022; Zhang et al., 2018; Zhao et al., 2017), as well as other tasks involving sequential and time series data, such as natural language processing. The main advantage of using plBi-LSTM instead of standard RNNs is that they can learn the temporal relationships in both the forward and backward directions simultaneously. The hidden state and output result for the Bi-LSTM at time t are computed according to the following equations:

$$\overrightarrow{h}_t = \text{LSTM}\left(g_t, \overrightarrow{h}_{t-1}\right) \tag{3.14}$$

$$\overleftarrow{h}_t = \text{LSTM}\left(g_t, \overleftarrow{h}_{t+1}\right) \tag{3.15}$$

$$y_t = W_{\overrightarrow{h}y}\overrightarrow{h}_t + W_{\overleftarrow{h}y}\overleftarrow{h}_t + b_y. \tag{3.16}$$

The Bi-LSTM is followed by multiple fully connected dense layers, where the last one is comprised of a single neuron used to output the predicted RUL. Similar to the LSTM, the Bi-LSTM network is trained using the back-propagation through time (BPTT) algorithm. The loss function to minimize is the mean squared error with L2 regularization:

$$\mathcal{L} = \frac{1}{\mathcal{N}} \sum_{n=1}^{\mathcal{N}} (y_n - \widehat{y}_n)^2 + \mu \cdot ||\mathbf{W}||_2^2. \tag{3.17}$$

where $\mathcal{N}$ is the number of samples, y_n and $\widehat{y}_n$ are the true and predicted targets, $\mathbf{W}$ are the weights of the neural network and μ is a hyperparameter that controls the strength of the regularization.

The above model is designed for RUL point prediction, however, to account for uncertainty in the prediction, MCD is implemented. Typically, dropout is a regularization technique that is applied only during the training phase of a neural network. MCD is an approach that extends dropout to the inference phase (Gal and Ghahramani, 2016). During inference, rather than using the trained model to make a single prediction, MCD involves performing $\mathcal{M}$ forward passes with dropout applied each time. This produces $\mathcal{M}$ different RUL predictions. Let $\left(\widehat{\text{RUL}}_{ijk}^{1}, \ldots, \widehat{\text{RUL}}_{ijk}^{\mathcal{M}}\right)$ represent the $\mathcal{M}$ RUL predictions obtained for component E_{ijk} and let $\widehat{F}_{\text{RUL}_{ijk}}(x)$ represent the empirical distribution of the RUL of E_{ijk}. Accordingly, the reliability $\mathcal{R}_{ijk}(U)$ and the expected downtime Δ_{ijk} of E_{ijk} for a mission of length U can then be computed as follows:

$$\mathcal{R}_{ijk}(U) = \frac{1}{\mathcal{M}} \sum_{m=1}^{\mathcal{M}} \mathbb{1}_{\left(\widehat{\text{RUL}}_{ijk}^{m} > U\right)}, \tag{3.18}$$

$$\begin{aligned}
\Delta_{ijk} &= \int_0^U (U - x)\, \mathrm{d}\widehat{F}_{\text{RUL}_{ijk}}(x) \\
&= U \int_0^U \mathrm{d}\widehat{F}_{\text{RUL}_{ijk}}(x) - \int_0^U x\ \mathrm{d}\widehat{F}_{\text{RUL}_{ijk}}(x) \\
&= U \frac{\sum_{m=1}^{\mathcal{M}} \mathbb{1}_{\left(\widehat{\text{RUL}}_{ijk}^{m} \leq U\right)}}{\mathcal{M}} - \frac{\sum_{m=1}^{\mathcal{M}} \widehat{\text{RUL}}_{ijk}^{m} \cdot \mathbb{1}_{\left(\widehat{\text{RUL}}_{ijk}^{m} \leq U\right)}}{\mathcal{M}},
\end{aligned} \tag{3.19}$$

where $\mathbb{1}_{(A)}$ is the indicator function taking a value of 1 if event A occurs and 0 otherwise.

Fleet predictive SM optimisation model and solution method

Due to resource limitations such as time and budget, it is not always possible to perform all desirable maintenance actions during the break. Thus, leading to the FSMO problem which corresponds to the third phase of the proposed predictive SM approach (see Figure 3.1). The FSMO model uses predicted health state, reliability and expected downtime of components provided by the Bi-LSTM and FCM algorithms respectively to jointly select the maintenance actions to be performed on selected components, and the assignment of repair crews to selected maintenance actions. The objective of the optimization model is to minimize the grand total maintenance cost while ensuring the fleet can operate at a target reliability for the upcoming mission. Before defining the optimization model, the following decision variables and notation are introduced:

$$x_{ijkrl} = \begin{cases} 1 & \text{if repairperson } r \text{ performs maintenance level } l \text{ on } E_{ijk}, \\ 0 & \text{otherwise.} \end{cases} \tag{3.20}$$

$$z_r = \begin{cases} 1 & \text{if repaircrew } r \text{ is hired,} \\ 0 & \text{otherwise.} \end{cases} \tag{3.21}$$

Two levels of maintenance for each component are considered in this work (*i.e.,* $L_{ijk} = 2$). The $l = 1$ case refers to DN and the $l = 2$ case refers to a maintenance action that will restore a component to

the AGAN state. In addition to the above decision variables, we denote by $R_{ijkl}(U)$ the reliability of component E_{ijk} if it undergoes maintenance level l. Accordingly, $R_{ijk1}(U)$ is given by Equation (3.22), while $R_{ijk2}(U)$ is the reliability of a brand new component.

$$R_{ijk1}(U) = \begin{cases} \mathcal{R}_{ijk}(U) & \text{if } S_{ijk} = 1, \\ 0 & \text{otherwise,} \end{cases} \tag{3.22}$$

where $\mathcal{R}_{ijk}(U)$ is given by Equation (3.18).

It is reasonable to assume that components within a system contribute to the overall system performance, as highlighted by Khatab et al. (2018a). Thus, a penalty cost due to components' downtime is included in the objective function. This cost is, without loss of generality, assumed to be proportional to the failure duration. Let Δ_{ijkl} represents the expected downtime of component E_{ijk} when maintenance level l is selected. Δ_{ijk1} is given by Equation (3.23) and Δ_{ijk2} is the expected downtime of a brand new component.

$$\Delta_{ijk1} = \begin{cases} \Delta_{ijk} & \text{if } S_{ijk} = 1, \\ U & \text{otherwise,} \end{cases} \tag{3.23}$$

where Δ_{ijk} is given by Equation (3.19).

The FSMO model can be formulated as the following mixed-integer non-linear optimization problem:

$$\begin{aligned} \min \ \mathcal{Z} = & \sum_{i=1}^{I}\sum_{j=1}^{N_i}\sum_{k=1}^{M_{ij}}\sum_{r=1}^{Q}\sum_{l=1}^{L_{ijk}} \Pi_{ijk}\cdot\Delta_{ijkl}\cdot x_{ijkrl} + \\ & \sum_{i=1}^{I}\sum_{j=1}^{N_i}\sum_{k=1}^{M_{ij}}\sum_{r=1}^{Q}\sum_{l=1}^{L_{ijk}} \left(c^p_{ijkl} + c^v_r\cdot t^p_{ijkl}\right)\cdot S_{ijk}\cdot x_{ijkrl} + \\ & \sum_{i=1}^{I}\sum_{j=1}^{N_i}\sum_{k=1}^{M_{ij}}\sum_{r=1}^{Q}\sum_{l=1}^{L_{ijk}} \left(c^c_{ijkl} + c^v_r\cdot t^c_{ijkl}\right)\cdot (1-S_{ijk})\cdot x_{ijkrl} + \\ & \sum_{r=1}^{Q} c^f_r\cdot z_r \end{aligned} \tag{3.24}$$

$$\text{s.t.} \quad \sum_{i=1}^{I}\sum_{j=1}^{N_i}\sum_{k=1}^{M_{ij}}\sum_{l=1}^{L_{ijk}} \left(t^p_{ijkl}\cdot S_{ijk} + t^c_{ijkl}\cdot(1-S_{ijk})\right)\cdot x_{ijkrl} \le D\cdot z_r, \quad \forall r \tag{3.25}$$

$$\begin{aligned} \mathcal{R}_i(U) = & \prod_{j=1}^{N_i}\left(\sum_{e_{k_{ij}}=1}^{M_{ij}}\sum_{e_{k_{ij}-1}=1}^{e_{k_{ij}}-1}\cdots\sum_{e_1=1}^{e_2-1}\left(\prod_{v\in\{e_1,\cdots,e_{k_{ij}}\}}\sum_{r=1}^{Q}\sum_{l=1}^{L_{ijv}} R_{ijvl}(U)\cdot x_{ijvrl}\right)\right)\cdot \\ & \left(\prod_{\substack{u=1 \\ u\notin\{e_1,\cdots,e_{k_{ij}}\}}}^{e_{k_{ij}}}(1-\sum_{r=1}^{Q}\sum_{l=1}^{L_{iju}} R_{ijul}(U)\cdot x_{ijurl})\right), \quad \forall i \end{aligned} \tag{3.26}$$

$$\mathcal{R}_i(U) \ge \mathcal{R}_0, \quad \forall i \tag{3.27}$$

$$\sum_{r=1}^{Q}\sum_{l=1}^{L_{ijk}} x_{ijkrl} = 1, \quad \forall i,j,k \tag{3.28}$$

$$x_{ijkrl} \in \{0,1\}, \quad \forall i,j,k,r,l \tag{3.29}$$

$$z_r \in \{0,1\}, \quad \forall k \tag{3.30}$$

In the above optimization model, the objective function seeks to minimize the expected predicted grand total cost. The first term represents the expected penalty cost of component failure. The second and third terms represent the cost of performing PM and CM maintenance, respectively. These costs include the variable repair crew cost and the fixed maintenance cost. The fourth term represents the fixed cost for

hiring repairpersons. Constraints (3.25) ensure that the time to perform maintenance for each repairperson does not exceed the break duration. Constraints (3.26) and (3.27) ensure that the target reliability level is met for each system. Constraints (3.28) ensure that each component is replaced only once if it is selected to undergo maintenance. Finally, Constraints (3.29) and (3.30) are binary variable restrictions.

The optimization model presented above is complex and requires the development of an efficient algorithm to provide optimal solutions. To solve the complex problem, an adaptation of the two-phase approach initiated by Diallo et al. (2018) and used in Khatab et al. (2020) is developed. The first phase of the approach consists in generating all feasible maintenance patterns for each subsystem, and the second phase consists in solving a MILP. A maintenance pattern $\boldsymbol{p}$ for subsystem j in system i is defined as a vector of length M_{ij} where the kth element p_k^h takes a value of 1 if component k is selected for maintenance and 0 otherwise. For a small two-component parallel subsystem, the four possible maintenance patterns would be: $\boldsymbol{p}^1 = [0,0]^\top, \boldsymbol{p}^2 = [1,0]^\top, \boldsymbol{p}^3 = [0,1]^\top, \boldsymbol{p}^4 = [1,1]^\top$ where $[.]^\top$ is the transpose operator. The first pattern implies no maintenance is selected, pattern two implies only the first component is maintained, and so on. For each pattern generated, the vector $\boldsymbol{t}_{ij}^h$ of length M_{ij} containing the time t_{ijk}^h to maintain each component E_{ijk} $(k = 1, \cdots, M_{ij})$ is stored, as well as the subsystem reliability $\mathcal{R}_{ij}^h(U)$ achieved and the total cost $\mathcal{C}_{ij}^h$ excluding repair crew variable and fixed cost. The output of the first phase is a dataset containing the following parameters: i, j, $\boldsymbol{p}^h$, $\boldsymbol{t}_{ij}^h$, $\mathcal{R}_{ij}^h(U)$, $\mathcal{C}_{ij}^h$, $\mathcal{H}_{ij}$ for each pattern h $(h = 1 \cdots \mathcal{H}_{ij})$ of each subsystem j in each system i. Algorithm (4) outlines how all feasible patterns are generated.

Algorithm 4 Compute $\mathcal{R}_{ij}^h(U)$, $\mathcal{C}_{ij}^h$, $\boldsymbol{t}_{ij}^h$, p^h for all valid patterns for subsystem j in system i

Input data: I, N_i, M_{ij}, t_{ijkl}, R_{ijk1}, $\mathcal{R}_0$, D, L_{ijk}
Initialize: $i = 1$
while $i \leq I$ **do**
  Initialize: $j = 1$
  while $j \leq N_i$ **do**
    – Generate all valid maintenance patterns for subsystem j in system i and store them in the set H_{ij}.
    – Find the cardinality of the set H_{ij}: $\mathcal{H}_{ij} = |H_{ij}|$.
    Initialize: $h = 1$
    while $h \leq \mathcal{H}_{ij}$ **do**
      – Compute $\mathcal{C}_{ij}^h$ the total cost of subsystem j excluding repair crew fixed and variable cost.
      – Compute $\mathcal{R}_{ij}^h(U)$ the reliability of subsystem j using the algorithm proposed by Kuo and Zuo (2003).
      if $\mathcal{R}_{ij}^h(U) \geq \mathcal{R}_0$ **then**
        – Store i, j, h, $\mathcal{R}_{ij}^h(U)$, $\mathcal{C}_{ij}^h$, $\boldsymbol{t}_{ij}^h$, $\boldsymbol{p}^h$.
      else
        – Remove current pattern h from the list H_{ij} (all patterns above h get shifted down by one position after the removal of h).
        – Update $h = h - 1$ (to account for the removed pattern).
        – Update $\mathcal{H}_{ij} = |H_{ij}|$.
      end if
      $h = h + 1$.
    end while
    $j = j + 1$.
  end while
  $i = i + 1$.
end while

The aim of the second phase of the approach is to select a maintenance pattern for each subsystem and assign a repairperson to each maintenance task such that the total cost is minimized. The following

new decision variables λ_{ij}^h, w_{ijkr}, and τ_{ijkr} are introduced.

$$\lambda_{ij}^h = \begin{cases} 1 & \text{if pattern } h \text{ is selected for subsystem } j \text{ in system } i, \\ 0 & \text{otherwise.} \end{cases} \tag{3.31}$$

$$w_{ijkr} = \begin{cases} 1 & \text{if repairperson } r \text{ performs maintenance on } E_{ijk}, \\ 0 & \text{otherwise.} \end{cases} \tag{3.32}$$

τ_{ijkr} is the total time spent performing maintenance on component E_{ijk} by repairperson r. With the above notation, the MILP formulation of the FSMO problem is as follows:

$$\min \ \mathcal{Z} = \sum_{i=1}^{I}\sum_{j=1}^{N_i}\sum_{h=1}^{\mathcal{H}_{ij}} \mathcal{C}_{ij}^h \cdot \lambda_{ij}^h + \sum_{i=1}^{I}\sum_{j=1}^{N_i}\sum_{k=1}^{M_{ij}}\sum_{r=1}^{Q} c_r^v \cdot \tau_{ijkr} + \sum_{r=1}^{Q} c_r^f \cdot z_r \tag{3.33}$$

$$\text{s.t.} \quad \sum_{h=1}^{\mathcal{H}_{ij}} \lambda_{ij}^h = 1, \quad \forall i,j \tag{3.34}$$

$$\sum_{j=1}^{N_i}\sum_{h=1}^{\mathcal{H}_{ij}} \ln\left(\mathcal{R}_{ij}^h(U)\right) \cdot \lambda_{ij}^h \geq \ln(\mathcal{R}_0), \quad \forall i \tag{3.35}$$

$$\sum_{r=1}^{Q} w_{ijkr} = \sum_{h=1}^{\mathcal{H}_{ij}} \lambda_{ij}^h \cdot p_k^h, \quad \forall i,j,k \tag{3.36}$$

$$\tau_{ijkr} \leq \sum_{h=1}^{\mathcal{H}_{ij}} \lambda_{ij}^h \cdot t_{ijk}^h, \quad \forall i,j,k,r \tag{3.37}$$

$$\tau_{ijkr} \geq \sum_{h=1}^{\mathcal{H}_{ij}} \lambda_{ij}^h \cdot t_{ijk}^h - (1 - w_{ijkr}) \cdot D, \quad \forall i,j,k,r \tag{3.38}$$

$$\sum_{i=1}^{I}\sum_{j=1}^{N_i}\sum_{k=1}^{M_{ij}} \tau_{ijkr} \leq D \cdot z_r, \quad \forall r \tag{3.39}$$

$$\tau_{ijkr} \leq D \cdot w_{ijkr}, \quad \forall i,j,k,r \tag{3.40}$$

$$\lambda_{ij}^h \in \{0,1\}, \quad \forall i,j,h \tag{3.41}$$

$$w_{ijkr} \in \{0,1\}, \quad \forall i,j,k,r \tag{3.42}$$

$$z_r \in \{0,1\}, \quad \forall r \tag{3.43}$$

$$\tau_{ijkr} \geq 0, \quad \forall i,j,k,r \tag{3.44}$$

In the above optimization model, the objective function (3.33) seeks to minimize the grand total cost. Constraints (3.34) ensure that exactly one maintenance pattern is selected for each subsystem. Constraint (3.35) warrants that the target reliability level is met for each system. Constraints (3.36) ensure that a repairperson is assigned to each maintenance task in the selected pattern. Constraints (3.37)–(3.38) are used to track the maintenance time for each repairperson. Constraints (3.39) ensure that repairperson is hired (fixed cost paid) before they can perform any maintenance action while constraints (3.40) guarantee that the time to perform maintenance for each repairperson does not exceed the break duration. Finally, constraints (3.41)–(3.44) are binary variables restrictions and non negativity constraints.

3.1.5 Numerical Experiments

In this section, multiple numerical experiments are conducted to demonstrate the efficiency and the added value of the proposed predictive SM approach and its ability to deal with maintenance decisions for fleets of multi-component systems. All experiments are run on an Intel^{TM} i5 2.9GHz desktop computer with 12GB of RAM running Windows 10^{TM}. All algorithms were coded in Python 3.8. The optimization runs were carried out by Gurobi 9.1 using gurobipy. All DL algorithms were coded in Python 3.8 using

TensorFlow. The DL algorithms were implemented in the Google Colaboratory environment using the T4 GPU. The dataset used, preprocessing steps, and evaluation metrics will first be presented. Then after, three sets of experiments are provided and the numerical results obtained are fully discussed.

Dataset, preprocessing, and evaluation metrics

The NASA C-MAPSS dataset is used to train the FCM algorithm, Bi-LSTM and to validate the predictive SM strategy. This dataset provides degradation trajectories under 4 different operational conditions and fault modes for turbofan engines. Because these different sets consider different operational conditions and fault modes they cannot be jointly studied. In this section, the first set called FD001, which considers a single fault mode, is selected. This set provides run to failure information for 100 engines and includes 3 operational settings and 21 sensor values for each operating cycle. An operating cycle refers to a single flight of approximately 90 minutes which includes both ascent to cruise at 35000 ft and descent back to sea level. A detailed description of the 21 sensor variables can be found in Saxena et al. (2008). In total, 20631 cycles are available for training. Each engine starts with different degrees of initial wear and manufacturing variation. The operational settings and sensor values are provided for every cycle up until engine failure. A second test set is also provided that includes truncated time series of various lengths prior to failure for 100 engines and the actual RULs.

Data preprocessing of the NASA C-MAPSS FD001 dataset was carried out according to the following steps:

— **Basic data visualization**. The first step taken was to visualize the sensor readings over the lifetime of the engines. This reveals that some sensor readings remain constant over time and add no value as inputs to the ML/DL algorithms. In particular, data collected from sensors 1, 5, 6, 10, 16, 18, and 19 were not considered when training the FCM clustering algorithm or Bi-LSTM.

— **Data labeling**. A column was added to both the training and testing sets that specified the RUL after completion of each cycle. This column provides the labels that will be used to train and test the models.

— **Input data scaling**. It is widely acknowledged that scaling input data for neural networks leads to more efficient training. The min-max normalization approach was used to scale the input of the DL algorithm to be in the range $[0, 1]$. The min-max normalization of data is performed according to the following formula:

$$\breve{x}_j^i = \frac{x_{ij} - x_j^{min}}{x_j^{max} - x_j^{min}}, \tag{3.45}$$

where x_j^i denotes the ith data measure of the jth sensor, $\breve{x}_j^i$ is the scaled value of x_j^i, x_j^{max} and x_j^{min} are the respective maximum and minimum raw measurements from the jth sensor.

— **Data reformatting**. In preparation for fitting the Bi-LSTM, the training and testing data was reformatted into 3-dimensional tensors of the form (N_s, N_{ts}, N_f). Recall that N_s refers to the number of samples, N_{ts} denotes the time sequence dimension, and N_f refers to the number of selected features.

— **Noise reduction**. To prepare the sensor data for the unsupervised clustering, denoising is carried out using the moving average technique. Similar to Wang et al. (2019), we use a window size of 20. Denoising is a crucial aspect of sensor data processing as it seeks to minimize the background noise in the acquired sensor measurements.

The following three metrics are used to evaluate the point prediction performance of the proposed Bi-LSTM: root mean squared error (RMSE), scoring function (SC), and accuracy (AC). RMSE is widely used for assessing point prediction accuracy, while SC and AC are frequently applied to the C-MAPSS dataset (Zhuang et al., 2023). Smaller values of RMSE and SC, along with higher values of AC are preferred. Assuming a total of $\mathcal{N}$ test examples with $\mathcal{M}$ predictions for each test example $n \in \{1, \ldots \mathcal{N}\}$, we define $d_n = \left(\frac{1}{\mathcal{M}} \sum_{m=1}^{\mathcal{M}} \widehat{y}_{nm}\right) - y_n$ as the difference between the average prediction and the target.

The mathematical expressions for the three evaluation metrics are as follows:

$$\text{RMSE} = \sqrt{\frac{\sum_{n=1}^{\mathcal{N}} (d_n)^2}{\mathcal{N}}}. \tag{3.46}$$

$$\text{SC} = \sum_{n=1}^{\mathcal{N}} s_n, \quad \text{where} \quad s_n = \begin{cases} \exp\left(\frac{-d_n}{13}\right) - 1 & \text{if } d_n \leq 0, \\ \exp\left(\frac{d_n}{10}\right) - 1 & \text{if } d_n \geq 0. \end{cases} \tag{3.47}$$

$$\text{AC} = \frac{100}{\mathcal{N}} \sum_{n=1}^{\mathcal{N}} a_n, \quad \text{where} \quad a_n = \begin{cases} 1 & \text{if } d_n \in [-13, 10], \\ 0 & \text{if } d_n \notin [-13, 10]. \end{cases} \tag{3.48}$$

The above metrics are valuable in assessing point prediction performance. However, as pointed out earlier, it is also important to assess the performance and quality of the predicted distribution. For this, the α-Coverage (α-C) and α-Mean Width (α-MW) metrics are used (de Pater and Mitici, 2022). The α-Coverage metric evaluates the reliability of the predicted distribution and is defined as follows:

$$\alpha\text{-C} = \frac{1}{\mathcal{N}} \sum_{n=1}^{\mathcal{N}} \mathbb{I}(\alpha)_n, \quad \text{where} \quad \mathbb{I}(\alpha)_n = \begin{cases} 1 & \text{if } y_n \in [\widehat{y}_n^{0.5-0.5\alpha}, \widehat{y}_n^{0.5+0.5\alpha}], \\ 0 & \text{otherwise.} \end{cases} \tag{3.49}$$

Here, for predicted values corresponding to test instance n, $\widehat{y}_n^k$ denotes the kth percentile and $[\widehat{y}_n^{0.5-0.5\alpha}, \hat{y}_i^{0.5+0.5\alpha}]$ represents the α percent confidence interval around the median. Ideally, α-C is very close to α, as this indicates more reliable predictions. The second metric, α-MW, defines the mean width of the α confidence interval. Obviously, tighter intervals are preferred as this implies sharper predicted distributions. The α-MW metric is defined as follows:

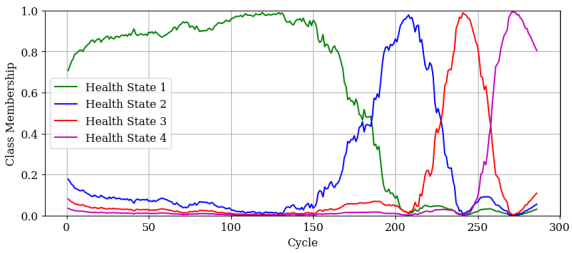
$$\alpha\text{-MW} = \frac{1}{\mathcal{N}} \sum_{n=1}^{\mathcal{N}} \left(\widehat{y}_n^{0.5+0.5\alpha} - \widehat{y}_n^{0.5-0.5\alpha}\right) \tag{3.50}$$

Results and analysis of FCM clustering and Bi-LSTM

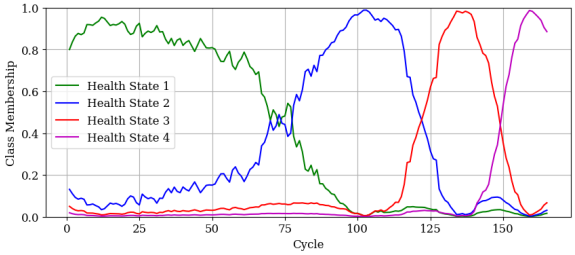
The FCM clustering algorithm presented in Section 3.1.4 is used to partition the training examples from the C-MAPSS dataset into discrete health states. Similar to (Chen et al., 2021a), the number of clusters C is set to 4, and the degree of fuzziness m is set to 2.00. The clustering results for four of the training engines is provided in Figure 3.2. What can be seen is that the FCM algorithm has been able to identify four very clear health states for the training engines ranging from normal, mild degradation, moderate degradation, and severe degradation or near to failure. Three of the four training engines begin in the first health state and progressively degrade and deteriorate over operational time until failure. One may observe that training engine 81 begins in the second identified health state and progressively degrades into states 3 and 4. This may be an indication that both health states 1 and 2 are similar in terms of performance and degradation level. After training the FCM algorithm, the cluster centroids can be used to identify the health states of previously unseen observations. Figure 3.2 also displays the results of the clustering algorithm on four of the testing engines that have not been used in training. These results seem to indicate that the algorithm has been able to generalize well to the testing set as the four health states have been preserved. Test engine 27 is quite healthy as it has only reached health state 2, where test engine 37 is more degraded as it has entered the third health state. The RUL of test engines 27 and 37 are 66 and 21 cycles respectively. Test engines 76 and 81 are clearly much more degraded as they have entered the fourth health state and are nearing failure. The actual RULs of these two engines are 10 and 8 cycles respectively.

In terms of the DL model used for RUL prediction, initial hyperparameters were informed by reviewing similar models proposed in the literature (Zhuang et al., 2023). The final network structure is defined by the following hyperparameter values:

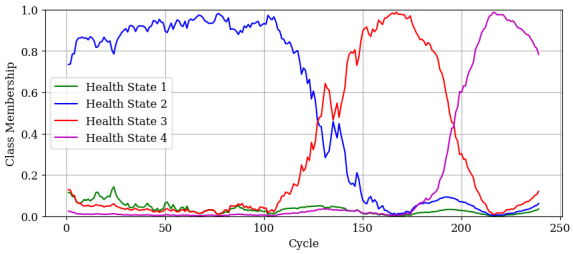
— 2 stacked Bi-LSTM layers each consisting of 20 hidden units,



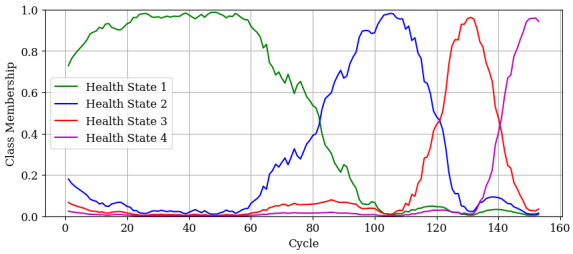
(a) Clustering results for training engine 2.



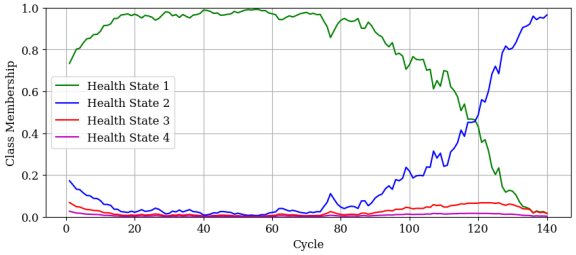
(b) Clustering results for training engine 74.



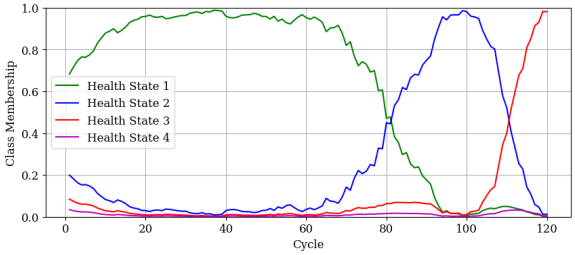
(c) Clustering results for training engine 81.



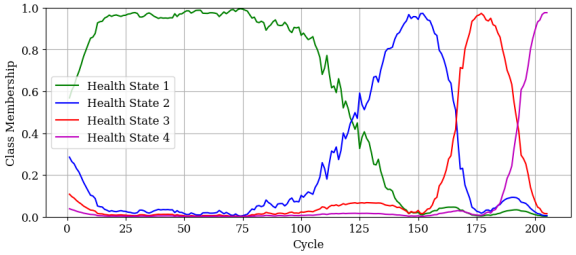
(d) Clustering results for training engine 90.



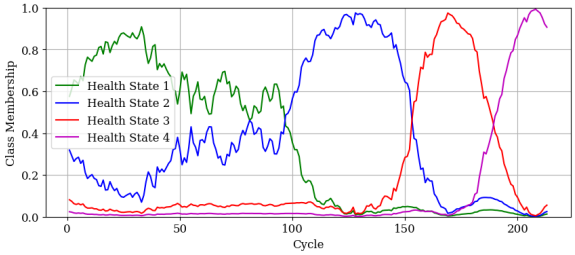
(e) Clustering results for test engine 27.



(f) Clustering results for test engine 37.



(g) Clustering results for test engine 76.



(h) Clustering results for test engine 81.

Figure 3.2 – Clustering results for multiple training and testing engines.

— 100 neurons for the fully connected dense layer,
— dropout rate set to 0.5,
— learning rate value of 0.0005,
— batch size of 128,
— number of training epochs set to 100,
— parameter μ controlling the strength of regularization set to $1e-3$, and
— the number of Monte Carlo simulations $\mathcal{M}$ set to 500.

To ensure efficiency while preventing overfitting, an early stopping procedure is implemented. Furthermore, a learning rate decay strategy is used to improve the convergence to a local minima and reduce oscillation. If no improvement to the validation loss is observed in 10 epochs, the learning rate is reduced by a factor of 0.5. Consistent with previous works (Li et al., 2018; Xia et al., 2021; Mitici et al., 2023), a piecewise linear target function is adopted, where the maximum RUL for all training and test engines is set to 125. A sliding window of size of $N_{ts} = 50$ is also used. To deal with test instances with no more than 50 historical flight cycles, the missing flight cycles are padded with the first sensor recordings.

Table 3.1 displays the results obtained from the proposed Bi-LSTM models across the three RUL point prediction metrics considered. This table also reports the results of various SOTA models. From the overall results obtained, one may indeed observe that the proposed predictive RUL model with MCD is very competitive and even outperforms some SOTA models. The proposed predictive RUL model shows high accuracy, and is among the best for RMSE and the scoring function.

Table 3.1 – Comparison of proposed method for RUL point prediction with the SOTA.

	RMSE	SC	AC
BiLSTM-ED (Yu et al., 2019)	14.74	273.00	57%
DBN-IPF (Peng et al., 2019)	13.11	314.00	63%
SBI (Yu et al., 2020)	13.58	228.00	67%
1-FCLCNN-LSTM (Peng et al., 2021)	**11.17**	**204.00**	–
PGRU (Zeng and Liang, 2023)	12.39	–	–
C-Transformer (Zhou et al., 2023b)	13.79	475.46	–
Bi-LSTM (Zhuang et al., 2023)	12.70	234.90	70%
Multi-channel CNN (Lee and Mitici, 2023)	11.81	–	–
CNN (Mitici et al., 2023)	12.42	–	–
BGT (Xiang et al., 2024)	12.09	262.67	–
Bi-LSTM (this work)	**11.58**	230.00	**73%**
Bi-LSTM-MCD (this work)	11.64	**214.85**	**74%**

For selected test engines, Figures 3.3, 3.4, and 3.5 display the predicted RUL, the empirical RUL probability density function (*pdf*), and the failure probability, respectively. The latter is computed by assessing the proportion of predictions that are less than the specified mission length. The following key observations can be made.

— From Figure 3.3, it appears that the Bi-LSTM is capable of making accurate RUL predictions. One can note that the accuracy of the model seemingly increases as the true RUL approaches 0. For test engines 24, 56, and 100, it can be seen that when the true RUL is at most 50 cycles, the model can almost perfectly predict the true remaining life. The information presented in Table 3.2 further supports this observation. This table shows the RMSE and AC for varying monitoring points along with the α-C and α-MW metrics for two values of α. The metrics are computed for test engines that have RULs of 75, 50, and 25 cycles. The results obtained show that as the true RUL decreases, RMSE decreases while the AC increases. This further demonstrates the enhanced ability of the model to accurately predict RUL as the engine approaches failure. This enhanced predictive ability may be a result of clear degradation trends emerging. It should be noted that the number of samples available for the three different monitoring points are 43, 33, and 19 respectively.

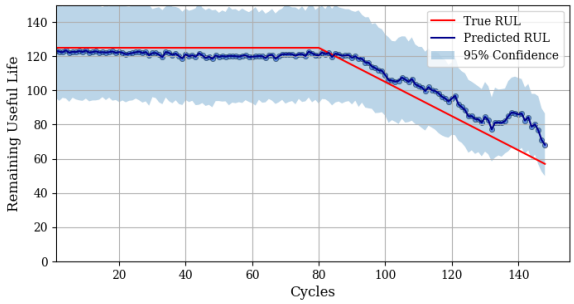
— Still dealing with Figure 3.3, one may observe that as the true RUL decreases, the size of the 95% confidence interval reduces. Looking at the RUL *pdf* curves displayed in Figure 3.4, the same trend emerges, and the standard deviation of the predicted RUL distribution increases as the true RUL value increases. Referring to the results reported in Table 3.2, one may observe that the uncertainty is overestimated for the test engines with 50 and 25 cycles remaining, and a drastic underestimation in the uncertainty occurs when dealing with test engines with 75 cycles remaining. Results also show that α-MW decreases as the true RUL decreases, emphasizing the sharper predictions achieved for engines with smaller remaining lifetimes.
— Figure 3.4 demonstrates the importance of considering the RUL distribution rather than the RUL point prediction when dealing with maintenance decisions. To illustrate, let us consider the example of test engine 36, and assume that the upcoming mission length $U = 20$ cycles. In this case, as the mean predicted RUL of 23.2 is greater than 20, a maintenance decision based solely on the RUL point prediction would result in no maintenance being performed. This decision would lead to engine failure as the true RUL is only 19 cycles. However, the RUL distribution shows that the probability of mission success is approximately only 50%. In this case, if a moderately high reliability target was set, the engine would be selected for maintenance, ultimately preventing failure. This highlights the need to incorporate the RUL distribution into the maintenance decision process.
— The failure curves in Figure 3.5 show that as the true RUL approaches the mission length U, the predicted failure probability increases significantly. In most cases, the probability is close to 1 when the RUL value is approaching the mission length U. This indicates that the proposed approach for failure prediction is accurate and behaving as expected. Ideally, if the model could perfectly predict RUL, the failure curves would exactly match the ideal failure curves. However, imperfection is inherent in every prediction model used, which introduces some level of non-smoothness in the failure curves and inaccuracy in the failure prediction.
— Dealing with the prediction of the RUL distribution, the quality of the Bi-LSTM model is discussed by comparing it to the CNN model used in (Mitici et al., 2023). This comparison is based on the α-C and α-MW metrics as reported in Table 3.3 for $\alpha = \{0.5, 0.9, 0.95\}$. The α-C values are very close to the α values. When $\alpha = 0.5$, the proposed Bi-LSTM model slightly overestimates the uncertainty. However, it slightly underestimating uncertainty in the case where $\alpha = 0.95$. For all values of α, the results obtained also show that the α-MW values obtained with the proposed model are smaller (*i.e.,* better) than those obtained for the CNN model in (Mitici et al., 2023).

Table 3.2 – Comparison of prediction metrics at different monitoring points.

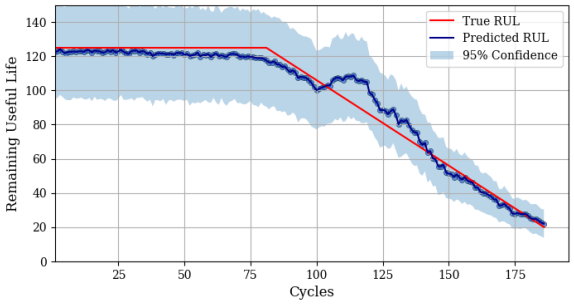
		Monitoring point with 75 cycles left	Monitoring point with 50 cycles left	Monitoring point with 25 cycles left
	RMSE	17.33	6.71	5.07
	AC	56%	94%	100%
$\alpha = 0.90$	α-C	0.70	0.91	0.95
	α-MW	33.50	24.26	15.06
$\alpha = 0.95$	α-C	0.74	0.97	1.00
	α-MW	39.89	28.65	17.92

Experiment #1: Comparison of reliability prediction methods

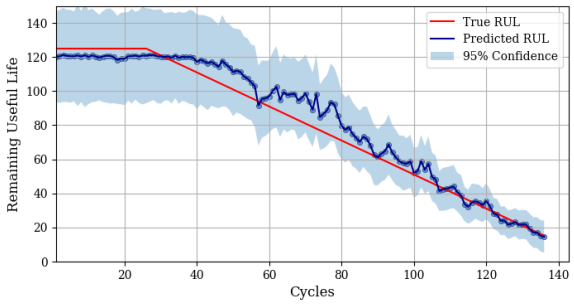
This set of numerical experiments intends to compare the proposed approach for reliability assessment with traditional statistical approaches and other data-driven methods from the literature. The following methods and models are compared with the perfect RUL prognostics model as a benchmark.
— Bi-LSTM with MCD (Bi-LSTM-MCD), the proposed approach,
— An approach that leverages the RUL point prediction (P.RUL) of the Bi-LSTM,
— The best fitting statistical distribution,
— A probabilistic classifier for reliability prediction (Bi-LSTM-PC), and



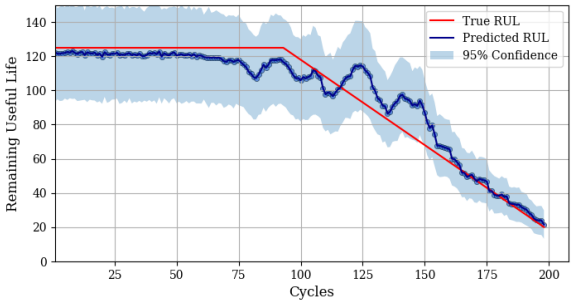
(a) RUL interval estimate for test engine 21.



(b) RUL interval estimate for test engine 24.

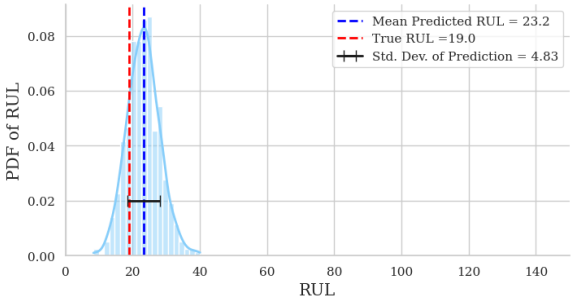


(c) RUL interval estimate for test engine 56.

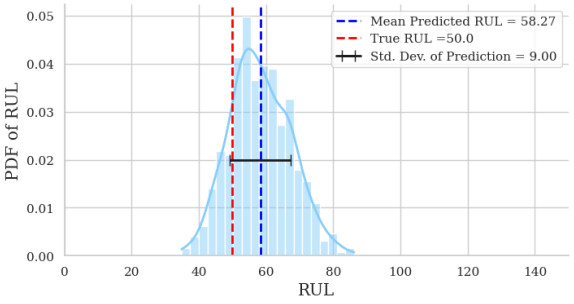


(d) RUL interval estimate for test engine 100.

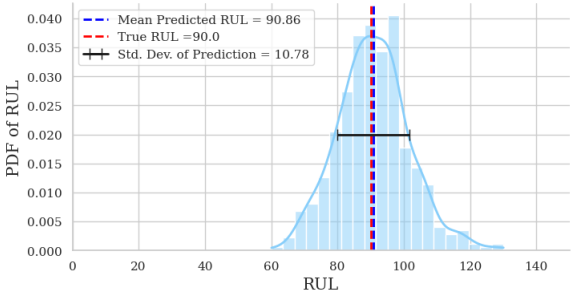
Figure 3.3 – RUL interval estimates for multiple test engines.



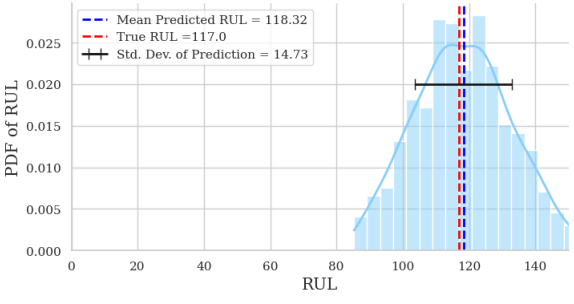
(a) RUL *pdf* plot for test engine 36.



(b) RUL *pdf* plot for test engine 72.

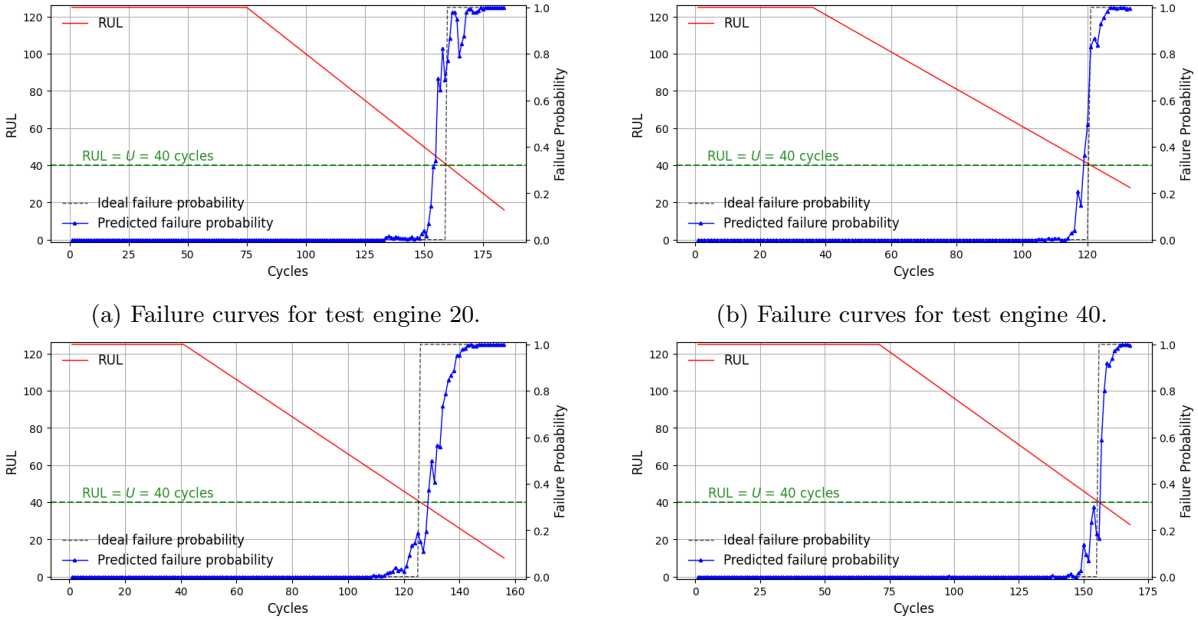


(c) RUL *pdf* plot for test engine 80.



(d) RUL *pdf* plot for test engine 99.

Figure 3.4 – RUL *pdf* plots for multiple test engines.



(a) Failure curves for test engine 20.

(b) Failure curves for test engine 40.

(c) Failure curves for test engine 42.

(d) Failure curves for test engine 64.

Figure 3.5 – Failure probability curves when $U = 40$ cycles for multiple test engines.

Table 3.3 – Comparison of α-Coverage (α-C) and α-Mean Width (α-MW).

		Bi-LSTM-MCD (this work)	CNN (Mitici et al., 2023)
$\alpha = 0.50$	α-C	0.52	0.54
	α-MW	12.72	16.30
$\alpha = 0.90$	α-C	0.90	0.91
	α-MW	30.99	39.20
$\alpha = 0.95$	α-C	0.92	0.95
	α-MW	36.86	46.4

— A perfect/true RUL (T.RUL) prognostics model (Benchmark).

This first set of experiments considers a single system comprised of one turbo fan engine. To analyze the performance of the above listed methods for reliability assessment over multiple instances, simulations are run for the entire life cycle of all 100 training engines in the FD001 dataset. Each engine is put into operation at time 0 and performs a mission followed by a maintenance break. The proposed optimization model is used to determine if maintenance is required during the break duration. If no maintenance is selected, the engine is sent off for another mission, and the process is repeated until the engine is replaced or fails during a mission. This procedure is performed for all 100 engines, which are divided into 5 folds, each composed of 80 training engines and 20 testing engines. It is worth noting that each testing set is totally unique, meaning that the 5 testing sets comprise all 100 engines. To ensure a fair comparative study, for each training set, a different DL model is trained and used to predict the RUL distribution of the test engines in the corresponding test set.

The best fitting statistical distribution is found to be the Weibull distribution with a shape and scale parameter of $\beta = 4.39$ and $\eta = 224$ respectively, similar to the parameters found in previous papers (Shoorkand et al., 2024; Hesabi et al., 2022). For the probabilistic classifier, an architecture similar to the one discussed in Section 3.1.5 is used, with the exception of the dropout rate reduced to 0.2, and the initial learning rate set to 0.001. The final layer of the classifier is comprised of two neurons, where the *softmax* activation function is employed to determine the probability that the component belongs to one of the two following classes:

— **Class 0:** RUL of the component is greater than the mission length U.

— **Class 1:** RUL of the component is less than or equal to the mission length U.

The probability that a component belongs to Class 0 can be interpreted as its reliability. Here, the categorical cross entropy loss function is minimised during the training of the probabilistic classifier.

To compare the methods listed above under similar conditions, all costs excluding maintenance cost are set to 0, while the reliability target $\mathcal{R}_0$ is varied. Furthermore, the maintenance cost is set to a very small value to perform maintenance only when necessary, and sufficient time is available to perform maintenance actions. The comparison uses three indicators: the number of early repairs, the number of system failures, and the number of missions completed. The number of early repairs refers to the number of times that a repair was performed when it was not necessary. An early repair is recorded when a replacement is selected despite the true RUL of a given component being greater than the upcoming mission length U. The decision to perform maintenance depends on both the predicted reliability and the target reliability threshold. The experiments are carried out for $U = 20$ cycles.

The first immediate observation that can be made from the results in Figure 3.6 is the significant improvement achieved by the DL approaches. Another observation is that as the target reliability increases, both the number of system failures and the number of completed missions decrease, while the number of early repairs increases. This is an expected outcome since an increase in the target reliability implies a heightened emphasis on failure prevention.

Utilizing a statistical approach (Weibull distribution) results in significantly more early replacements, and far fewer completed missions in most cases. These results highlight the advantage of data-driven methods over distribution-based approaches for reliability prediction. The data-driven approaches are indeed more accurate as they base their predictions on the current health-condition of an engine.

Another important observation to point out is the seemingly better prediction made by the probabilistic classifier compared to the regression model implementing MCD. In fact, for many of the instances solved, the probabilistic classifier for reliability prediction results in fewer early repairs, and higher number of successful missions. In other words, the probabilistic classifier seems to better utilize the full operational life of the engines. However, what must also be emphasized is the difference in the number of system failures observed when the reliability target is set to a high value (*i.e.,* greater than or equal to 95%). Even with extremely high reliability targets of 95.0% and 99.5%, the probabilistic classifier still allows several system failures to occur. This means that the model is predicting an extraordinarily high likelihood of mission success, when, in fact, the system will fail during the upcoming mission. In the

context of mission-critical systems, whose failure could be catastrophic, the Bi-LSTM-MCD should be preferred to the Bi-LSTM-PC. Conversely, the Bi-LSTM-PC should be preferred if failures are not costly.

The results displayed in Figure 3.6 also highlight the benefit of using the RUL distribution rather than the point prediction. The RUL point prediction yields fewer early repairs and more missions are completed. However, this approach leads to many more system failures.

Experiment #2: Application to fleets of large multi-component systems

This set of numerical experiments solves the FSMP for a fleet of 5 large multi-component systems. Each system is comprised of 3 subsystems consisting of 5 turbofan engines, where at least 3 of the 5 engines must function in order for the subsystem to function. Thus, the reliability structure of each subsystem is a 3-out-of-5:G. To study the long-run performance of the proposed FSM approach, several breaks and missions are replicated: an engine from the FD001 training is randomly assigned to each component within each system. At the start of the first maintenance break, all engines are considered to be brand new. The fleet performs a mission of length $U = 30$ cycles, after which the proposed maintenance optimization model is used to determine what maintenance actions to perform, the number of repairpersons to hire and utilize, in addition to the assignment of these repairpersons to maintenance tasks. After the maintenance break, the system is put into operation to perform another mission of the same length $U = 30$ cycles. This process is repeated for 100 mission and break cycles. During a break, if an engine is selected for maintenance, it is replaced by an engine which is randomly drawn from the training set. The entire process is repeated 5 times and the mean and standard deviations are reported for multiple metrics in Figure 3.7. These metrics are reported for the different methods used for reliability prediction.

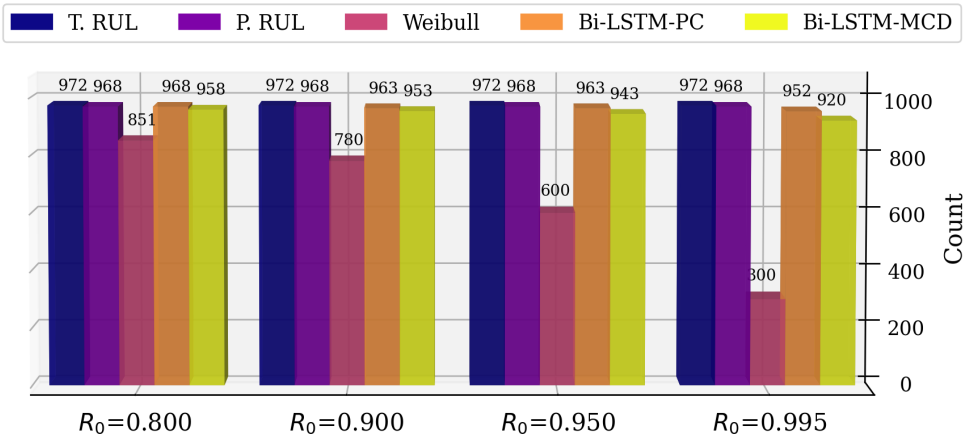
For this experiment, 10 repairpersons are available with fixed and variables costs of $c_r^f = 5$ monetary units (m.u.) and $c_r^v = 1$ m.u./hr, respectively. The time and cost of PM for all health states is assumed to be the same and set to $c_{ijk2}^p = 2$ m.u. and $t_{ijk2}^p = 1$ hour. Additionally, the time and cost of CM actions are set to $c_{ijk2}^c = 5$ m.u. and $t_{ijk2}^c = 2$ hours. Finally, the break duration is set to 10 hours. The penalty cost Π_{ijk} is the same for all components and the target reliability threshold $\mathcal{R}_0$=0.95.

The results obtained are summarized in Figure 3.7, which displays the average number of system failures, the average total cost per break, the average component downtime per mission, and the average number of CM actions performed per maintenance break. An advantage of using the Bi-LSTM-MCD can be seen immediately from the histograms corresponding to the average number of system failures, and the resulting grand total cost. Compared to the probabilistic classifier or the method using the RUL point prediction, the Bi-LSTM-MCD method is very competitive as it provides cost effective decisions while reducing system failures. For all reliability prediction methods considered, the number of system failures decreases as the component downtime penalty cost increases. This is due to the fact that when the penalty cost increases, components are more likely to be selected for maintenance, which then leads to an increase in their reliability and therefore a reduction in their downtime. The component penalty cost "takes over" the objective function and the primary goal is to reduce component downtime.

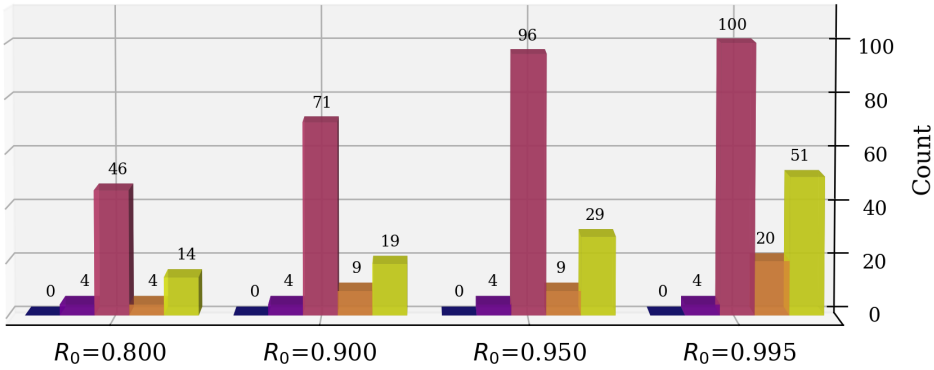
Dealing with the particular case where no penalty cost is incurred for component failure during missions, a very small number of CM are performed, resulting in high downtime. This is due to the fact that CM is typically costlier than PM, thus it is unlikely that the model will ever elect to perform a CM action. However, in a context where individual components contribute to the overall efficiency and global performance of the system, it is undesirable and impractical to allow a component to remain in the failed state for an extended period of time. Thus, it is required to explicitly account for component downtime in the maintenance optimization problems in general, and in the SMP in particular.

Illustrative case study: multi-engine aircraft

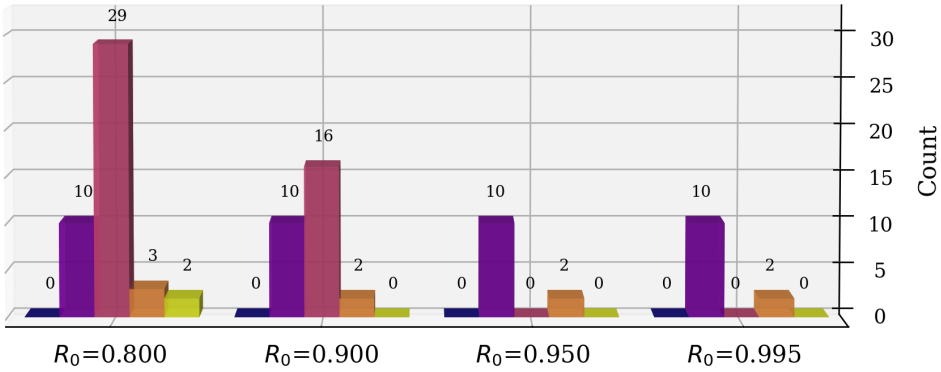
This set of experiments illustrates how the proposed predictive SM approach can deal with joint maintenance and repairpersons assignment decisions for a fleet composed of several multi-engine aircraft. The fleet investigated is composed of 6 aircraft ($i = 1, \ldots, 6$) each equipped with 4 turbofan engines



(a) Number of completed missions.



(b) Number of early repairs.



(c) Number of system failures.

Figure 3.6 – Comparing methods for reliability assessment for $U = 20$ cycles and varying values of $\mathcal{R}_0$. Case of experiment #1.

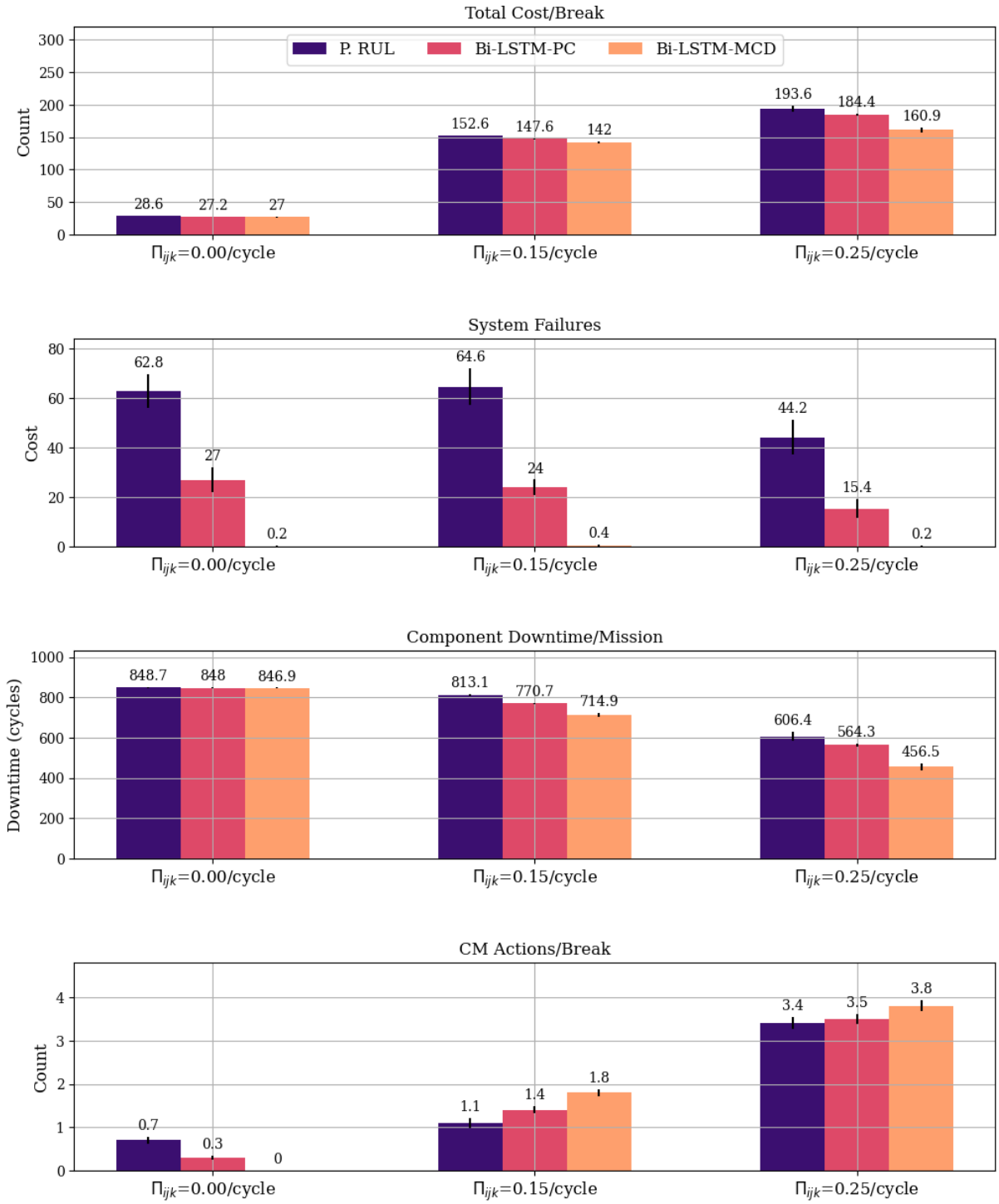


Figure 3.7 – Comparing methods for reliability assessment for $\mathcal{R}_0 = 0.95$ and $D = 10$ hours for varying values of Π_{ijk}. Case of experiment #2.

($k = 1, \ldots, 4$). An aircraft will successfully operate the upcoming mission if at least two of the four engines survive the mission duration. Thus, the reliability structure of an aircraft's engine system is a 2-out-of-4:G structure where each component E_{i1k} is assigned the sensor values of an engine from the test set of the NASA-CMAPSS FD001 dataset. The duration of the next mission is set at $U = 40$ cycles. To perform maintenance activities, $Q = 15$ repairpersons are available with fixed cost and variable cost rate of $c_r^f = 50$ m.u. and $c_r^v = 5$ m.u./hr, respectively. Table 3.4 provides the engine data for each aircraft. This data includes the number of cycles that each engine has completed, health state membership, its corresponding true and predicted RUL, predicted reliability $\mathcal{R}_{i1k1}$ and predicted downtime Δ_{i1k1}. Aircraft 3, 4, and 5 will fail during the next mission as at least 3 of the engines equipped have true remaining lifetimes that are less than the 40 cycle mission length. The average maintenance time corresponding to health states HS1 (normal), HS2 (mild degradation), HS3 (moderate degradation), and HS4 (near failure) are 4, 5, 8, and 10 hours respectively and the maintenance cost $c_{ijk2}^p = 30$ m.u.

Table 3.4 – Engine data used for case study.

Components	Engine Assignment	Health State Membership	True RUL (cycles)	Predicted RUL (cycles)	$\mathcal{R}_{i1k1}$	Δ_{i1k1}
E_{111}	17	[0.05, 0.57, 0.35, 0.03]	50	54	0.98	0.07
E_{112}	18	[0.01, 0.06, 0.91, 0.02]	28	33	0.11	7.59
E_{113}	20	[0.00, 0.01, 0.02, 0.96]	16	13	0.00	26.85
E_{114}	21	[0.29, 0.63, 0.06, 0.01]	57	68	1.00	0.01
E_{211}	24	[0.02, 0.09, 0.86, 0.03]	20	22	0.00	18.07
E_{212}	27	[0.02, 0.96, 0.02, 0.00]	66	66	1.00	0.00
E_{213}	31	[0.00, 0.01, 0.02, 0.96]	8	11	0.00	28.95
E_{214}	32	[0.03, 0.93, 0.03, 0.01]	48	45	0.76	0.92
E_{311}	34	[0.01, 0.01, 0.03, 0.95]	7	8	0.00	31.28
E_{312}	35	[0.01, 0.02, 0.04, 0.93]	11	12	0.00	28.26
E_{313}	46	[0.01, 0.04, 0.92, 0.03]	47	38	0.38	3.28
E_{314}	49	[0.00, 0.01, 0.03, 0.96]	21	25	0.00	14.7
E_{411}	52	[0.04, 0.38, 0.55, 0.03]	29	30	0.00	10.02
E_{412}	56	[0.00, 0.01, 0.02, 0.97]	15	14	0.00	25.82
E_{413}	58	[0.02, 0.07, 0.82, 0.09]	37	37	0.29	3.44
E_{414}	62	[0.05, 0.44, 0.47, 0.03]	54	46	0.85	0.49
E_{511}	68	[0.00, 0.00, 0.00, 0.99]	8	12	0.00	28.27
E_{512}	77	[0.02, 0.07, 0.80, 0.11]	34	29	0.01	10.00
E_{513}	81	[0.01, 0.02, 0.06, 0.91]	8	9	0.00	30.84
E_{514}	84	[0.00, 0.01, 0.02, 0.97]	38	34	0.14	7.06
E_{611}	91	[0.01, 0.05, 0.90, 0.04]	58	62	1.00	0.00
E_{612}	94	[0.02, 0.07, 0.83, 0.09]	55	52	0.97	0.13
E_{613}	98	[0.03, 0.24, 0.69, 0.03]	59	80	1.00	0.00
E_{614}	100	[0.02, 0.08, 0.86, 0.04]	20	21	0.00	18.3

To show how the required minimum reliability level impacts the maintenance and repairpersons assignment decisions within a FSMP, this experiment is conducted in the case where the break duration $D = 15$ hours while the required minimum reliability $\mathcal{R}_0$ and component penalty cost Π_{ijk} are varied. The latter is set to be the same for all engines/components for simplicity. The results obtained are reported in Table 3.5. This table gives the optimal predicted objective function value $\mathcal{Z}^*$, the true grand total cost $\mathcal{Z}'$, the total number NC^* of components maintained, the number Q^* of repairpersons hired and utilized, the total maintenance time (MT), in addition to the total component downtime (DT). The total maintenance time is computed by summing the individual repairperson maintenance times. It is worth noting here that because components downtimes are computed using the predicted RUL distribution opposed to the true RUL, it results in a difference between the predicted optimal objective function and true grand total cost.

When the downtime penalty cost is 0, it can be seen that as the required minimum reliability is increased, the total maintenance time and total number of maintenance actions increase. An increase in the number of repairpersons utilized when $\mathcal{R}_0$ increases from 0.950 to 0.995 can also be observed. The case where $\Pi_{ijk} = 4$ m.u./cycle shows that increasing the reliability target results in an increase in the total maintenance time, number of maintenance actions, and number of repairpersons utilized. Both increases and decreases in the total component downtime can be seen as the reliability target is increased, highlighting the complex interaction between penalty, maintenance, and crewing costs. When Π_{ijk} is set to 8 m.u./cycle and 16 m.u./cycle, we see the respective maintenance plans remain fixed for all reliability targets. This indicates that the reliability constraint is not tight, *i.e.,* the maintenance plan is not constrained by the target reliability threshold.

The penalty cost per cycle Π_{ijk} clearly influences maintenance decisions related to the number of repair crews to hire/utilize and the components to maintain. When $\mathcal{R}_0 = 0.95$, Figure 3.8 displays both the total cost as well as the total component downtime for increasing values of Π_{ijk}. It is evident that as Π_{ijk} increases, there is an emphasis on ensuring that no components fail during the subsequent mission, effectively minimizing downtime. However, achieving minimal downtime requires increases in maintenance cost, fixed and variable crew costs, ultimately leading to an increase in grand total cost.

When $\mathcal{R}_0 = 0.75$ and $\Pi_{ijk} = 0$ m.u./cycle, the following components are selected for maintenance: $E_{3,1,2}$, $E_{3,1,3}$, $E_{4,1,1}$, $E_{5,1,2}$, and $E_{5,1,3}$, resulting in a grand total cost of 560. It is interesting to note that when the fleet reliability target is set to a relatively low value of 0.75, all aircraft that would have failed during the mission are selected to be maintained. As the reliability target is increased, additional aircraft equipped with engines having RULs nearing 40 cycles are also selected for maintenance. For fleets comprised of critical assets where the failure of one system would have catastrophic consequences, the decision maker may set a high reliability target. This will indeed reduce the risk of mission failure, but at the expense of maintenance cost.

Table 3.5 – Results for varying values of $\mathcal{R}_0$ and Π_{ijk} with $D = 15$ hours.

Π(m.u./cycle)	$\mathcal{R}_0$	$\mathcal{Z}^*$(m.u.)	$\mathcal{Z}'$(m.u.)	$\frac{\mathcal{Z}^*-\mathcal{Z}'}{\mathcal{Z}'} \cdot 100$ (%)	NC^*	Q^*	MT (hours)	DT (cycles)
0	0.750	560	560	0.00	5	4	42.0	222
	0.850	616	616	0.00	6	4	47.1	222
	0.950	678	678	0.00	7	4	53.6	222
	0.995	789	789	0.00	8	5	59.8	222
4	0.750	1245	1242	0.24	5	5	49.2	149
	0.850	1259	1243	1.29	6	6	59.1	117
	0.950	1315	1300	1.15	7	6	63.6	118
	0.995	1337	1334	0.22	8	7	73.5	94
8	0.750	1587	1542	2.92	10	9	91.2	42
	0.850	1587	1542	2.92	10	9	91.2	42
	0.950	1587	1542	2.92	10	9	91.2	42
	0.995	1587	1542	2.92	10	9	91.2	42
16	0.750	1820	1654	10.04	13	12	116.7	5
	0.850	1820	1654	10.04	13	12	116.7	5
	0.950	1820	1654	10.04	13	12	116.7	5
	0.995	1820	1654	10.04	13	12	116.7	5

The purpose of the present experiment is to demonstrate the impact of the scheduled break duration D on the maintenance plans. The FSMO problem is solved for each break duration $D \in \{10, 15, 20, 25\}$, while the required minimum reliability level $\mathcal{R}_0 = 0.95$. For each problem instance, the optimal grand total cost $\mathcal{Z}'$, the total number NC^* of components maintained, the number Q^* of repairpersons hired and utilized, and the total downtime (DT) are provided in Table 3.6. This table also displays the results for the perfect RUL predictive model (T.RUL), which serves as the ideal scenario.

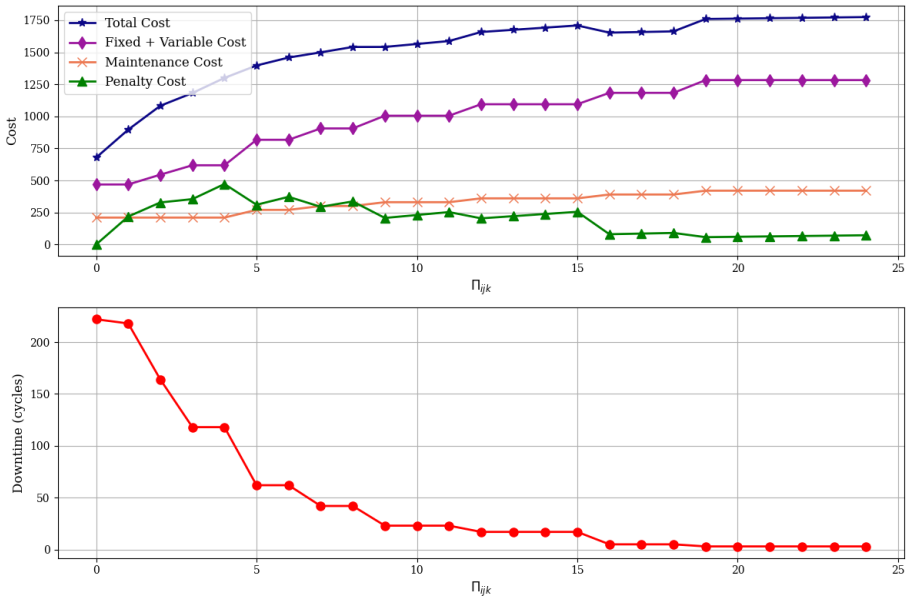


Figure 3.8 – Component downtime and total cost for varying values of Π_{ijk}.

As expected, when Π_{ijk}=0 m.u./cycle, increasing the maintenance break duration results in fewer repairpersons being hired to complete the maintenance duties. Table 3.6 also depicts the gap between the optimal grand total cost values resulting from the two reliability prediction methods. The Bi-LSTM-MCD model suggests performing 7 maintenance actions, while the T.RUL method recommends only 4 maintenance actions to meet the required minimum reliability $\mathcal{R}_0 = 0.95$. This once again highlights the tendency of the Bi-LSTM-MCD model to select more maintenance actions for high reliability thresholds to protect against failures.

Results from Table 3.6 show that the gap in grand total cost decreases as the downtime penalty cost rate Π_{ijk} increases. The same trend is observed for the optimal number NC^* of components to be maintained, and the number Q^* of the repair crews hired and utilized. One trend emerging for larger values of Π_{ijk} is the decreasing downtime as the break duration is increased. This is an expected result as increasing the break duration allows for more maintenance tasks to be performed, ultimately leading to a reduction in total component downtime.

Finally, Figure 3.9 depicts the relationship between total cost, break duration and mission length. It shows that the total cost increases as the length of the mission is increased, and as the break duration is decreased. The first observation is due to the fact that when the mission length is increased, more components need to be maintained to reduce components downtime which in turn increase the probability of mission success. Regarding the second observation, it is due to the fact that when the break duration is shortened, more repairpersons are required to perform the maintenance duties.

The results obtained and discussed above show that the proposed integrated DL and optimization method yields reasonable and valid maintenance and repairperson assignment decisions.

3.1.6 Conclusions and Future Research Directions

In this section, a FSM strategy is proposed that utilizes recent advancements in AI for reliability assessment. An integrated DL and optimization approach is used to identify optimal maintenance decisions. This section improved upon the original data-driven SM strategy proposed by Hesabi et al. (2022), specifically by considering fleets of large multicomponent systems and by developing and efficient solution method. This section considered the impact that degradation has on maintenance time and also dealt with the assignment of multiple repairpersons to maintenance tasks. A Bi-LSTM network implementing

Table 3.6 – Impact of the break and mission duration on maintenance plans.

		Bi-LSTM-MCD				T.RUL				
Π (m.u./cycle)	D (hours)	$\mathcal{Z}'$(m.u.)	NC^*	Q^*	DT (cycles)	$\mathcal{Z}'$(m.u.)	NC^*	Q^*	DT (cycles)	Gap(%)
0	10	828	7	7	222	491	4	4	222	68.84
	15	678	7	4	222	441	4	3	222	53.74
	20	628	7	3	222	391	4	2	222	60.61
	25	628	7	3	222	391	4	2	222	60.61
4	10	1313	7	7	106	1229	4	4	178	6.83
	15	1300	7	6	118	1229	4	4	178	5.78
	20	1146	8	4	82	1093	6	3	117	4.85
	25	1135	9	4	62	1093	6	3	117	3.84
8	10	1592	10	10	42	1543	10	10	34	3.18
	15	1542	10	9	42	1519	11	10	23	1.51
	20	1342	10	5	42	1292	12	6	11	3.87
	25	1269	11	5	23	1266	11	5	22	0.24
16	10	1704	13	13	5	1680	12	12	11	1.43
	15	1654	13	12	5	1630	12	11	11	1.47
	20	1451	14	7	3	1380	12	6	11	5.14
	25	1401	14	6	3	1354	13	6	5	3.47

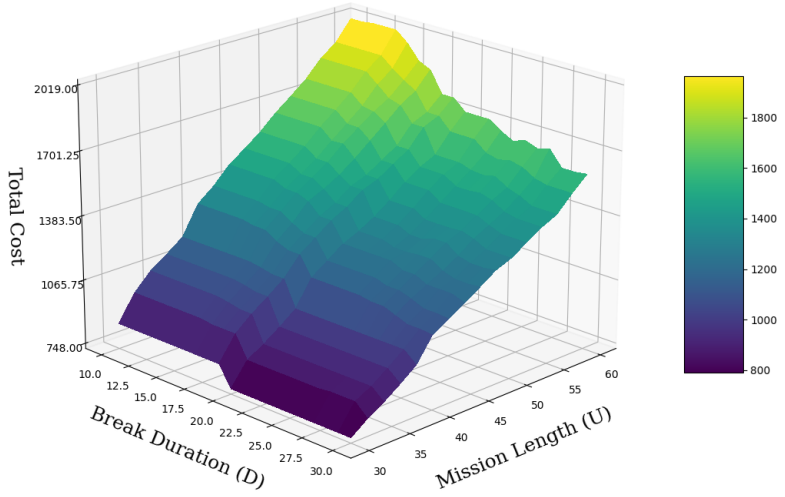


Figure 3.9 – Total cost for varying values of U and D with $\Pi_{ijk} = 4$/cycle and $\mathcal{R}_0 = 0.950$.

MCD was used to predict a RUL distribution from which component and system reliability could be derived. A comparison of several state-of-the-art DL algorithms for the purpose of reliability assessment was also provided. The results from the comparative study indicated that the proposed approach was preferable compared to methods leveraging RUL point predictions and probabilistic classifiers. The Bi-LSTM network was able to achieve a very high RMSE and accuracy indicating its ability to accurately predict RUL. The Bi-LSTM also showed excellent results in metrics used for assessing the quality of the predicted RUL distribution such as α-Coverage and α-Mean Width.

In this work we consider perfect repair of components as the only maintenance action. A way of extending and improving upon this work would be to consider several IM levels, where a component is brought back to a state that is between perfect replacement and DN. Modeling IM and its impact when using data-driven approaches for reliability assessment is a topic worthy of being studied. Another important extension of the proposed framework is the development of a maintenance optimization model that takes into consideration multiple future missions. Currently, our approach only considers the upcoming mission; however, this does not necessarily produce cost-effective maintenance plans in the long run. Thus, it would be of great benefit to extend the proposed framework to the multi-mission context.

Due to time restrictions, the number of iterations or combinations of hyperparameters used in the random search was limited to 25. By trying more combinations of hyperparameters, we would have stronger evidence that the selected hyperparameters are optimal for model performance. Areas of future work would include implementing data reduction methods such as Principle Component Analysis (PCA), Autoencoders (AE), and Variational Autoencoders (VAEs). AEs and VAEs have the ability to learn compressed representations of the data and extract meaningful features. It would also be interesting to study more complex DL architecture for the task of predicting RUL, such as transformers and hybrid DL approaches. Hybrid approaches that utilize multiple DL algorithms have shown promise in RUL prediction. Another area of future research would be studying deep clustering algorithms for the purpose of health state identification. Finally, what has yet to be studied in the context of predictive SM is the development of multimodal learning models for the task of reliability prediction. The development of a multimodal learning model that combines not only time-series sensor data, but also maintenance reports, thermal imaging, and machine video to make reliability predictions has the potential to significantly improve prediction accuracy.

3.2 Optimizing Predictive Maintenance and Mission Assignment to Enhance Fleet Readiness Under Uncertainty

In many industrial settings, fleets of assets are required to operate through alternating missions and breaks. FSM is widely used in such contexts to improve the fleet performance. However, existing FSM models assume that upcoming missions are identical and require only a single system configuration for completion. Additionally, these models typically assume that all missions must be completed, overlooking resource constraints that may prevent readying all systems within the available break duration. This makes mission prioritization and assignment a necessary consideration for the decision-maker. This work proposes a novel FSM model that jointly optimizes system to mission assignment, component and maintenance level selection, and repair task allocation. The proposed framework integrates analytical models for standard components and DNN for sensor-monitored ones, enabling a hybrid reliability assessment approach that better reflects real-world multi-component systems. To account for uncertainties in maintenance and break durations, a chance-constrained optimization model is developed to ensure that maintenance is completed within the available break duration with a specified confidence level. The optimization model is reformulated using the Conditional Value-at-Risk (CVaR) approximation. A case study of military aircraft fleet is investigated to demonstrate the accuracy and added value of the proposed approach.

3.2.1 Introduction

Many engineered assets and systems are designed to operate according to a sequence of alternating missions and breaks. These MOS are ubiquitous in several industrial applications such as air transportation, military equipment, nuclear plants, spacecraft, and unmanned autonomous vehicles. A key metric used to assess their performance is the mission success probability, defined as the system's ability to successfully complete a defined mission of a specific profile under given operational conditions. To ensure the that mission success probability meets a required level for the upcoming mission, maintenance actions are usually carried out on the system/components during scheduled maintenance breaks. However, due to the limited maintenance resources (*e.g.,* time, budget, materials, spare parts, crews) only a subset of the desired maintenance tasks can be performed. In the literature, this resource-constrained maintenance problem is known as the SMP. The primary objective is to identify the optimal subset of components and maintenance actions that either minimize maintenance costs while ensuring a required level of the system reliability to operate the next mission, or maximize the system reliability subjected to maintenance budget and/or break duration constraints.

The original SM model was introduced in Rice et al. (1998) with the aim of maximizing mission success probability using a full enumeration algorithm. Many researchers have since expanded upon this original work to include complex system configurations (Diallo et al., 2018), multistate systems (Pandey et al., 2013; Ma et al., 2023), component dependence (Zhang et al., 2024a; Xu et al., 2016; Dao and Zuo, 2017a), uncertainties related to maintenance, break, and mission durations (Khatab et al., 2017a; Liu et al., 2018; Jiang and Liu, 2020b), uncertainties related to maintenance quality (Khatab and Aghezzaf, 2016; Al-Jabouri et al., 2023a), FSM (Khatab et al., 2022; Amjadian et al., 2024), and multiple repair channels (Diallo et al., 2019; Chaabane et al., 2020). Significant advances have also been made in the solution methods used to solve the SMP. Rajagopalan and Cassady (2006) refined to the total enumeration approach developed in Rice et al. (1998). Multiple authors have utilized metaheuristics such as Tabu search Lust et al. (2009), GA (Chaabane et al., 2020), and ant colony optimization (Liu et al., 2018). Diallo et al. (2018) introduced a two-phase approach that eliminates the need for solving a nonlinear objective function as the approach converts the problem to a multidimensional multi-choice knapsack problem. More recently, O'Neil et al. (2023) introduced a CG approach to solve the multi-mission SMP, where a metaheuristic was employed to solve the subproblems. Al-Jabouri et al. (2023b) developed a branch-and-price framework to solve large-scale single-mission SMP problems where the subproblems were solved exactly by reformulating the problem using exponential cones.

Many existing SMP papers treat maintenance and break durations as deterministic parameters that are known and constant. Nevertheless, there are many applications in which the precise duration of the

maintenance actions and break durations cannot be exactly determined but would be characterized by a random value (Khatab et al., 2017a; Wang et al., 2023a; Al-Jabouri et al., 2024a). For instance, in military aircraft fleets, the unexpected start of the next mission/battle introduces uncertainties in the available maintenance time Liu et al. (2018). Due to the variability in repair crew experience, among many other factors, the time to perform maintenance actions may also be nondeterministic. Khatab et al. (2017a) studied the SMP when durations of the maintenance actions, scheduled breaks, and missions are stochastic by proposing a nonlinear stochastic optimization problem with a chance constraint to ensure the probability of completing the maintenance actions is greater than some predefined service level. Liu et al. (2018) studied the SMP with stochastic maintenance and mission durations for multi-state systems. Their model aims to identify the subset of maintenance actions to perform, as well as the sequencing of the actions such that the system reliability is maximized. The saddlepoint approximation is utilized to approximate the multi-dimensional convolution that arises when evaluating the system reliability (O'Neil et al., 2024). Liu et al. (2022b) dealt with a multi-mission SMP for multicomponent systems where the durations of missions, maintenance actions, and breaks are stochastic. Yin et al. (2023) investigated the SMP with stochastic duration of maintenance tasks and breaks. The authors used the expected cumulative performance to calculate system reliability. More recently, Al-Jabouri et al. (2024a) studied the SMP under uncertain maintenance durations, and employed distributionally robust chance-constraints to handle such uncertainties.

An interesting and very important extension of the classical SMP that has received much attention is the FSM. In the FSMP, maintenance decisions must be determined for each component within a fleet of systems. Cassady et al. (2003) first introduced the FSMP and studied its application to a fleet of military aircraft. The authors proposed an optimization model aimed at maximizing fleet reliability, which is defined as the probability that all systems successfully complete the upcoming mission. Schneider and Cassady (2004) extended the work of Cassady et al. (2003) by proposing a cost-based formulation and considering the possibility of canceling upcoming missions. Lan et al. (2017) introduced a multi-objective FSMP with three unique objective functions: maximize fleet reliability, minimize total maintenance cost, and minimize gap time. The proposed model was applied to a truck fleet carrying out long-distance highway transportation missions. Khatab et al. (2022) and Chen et al. (2023b) extended the FSMP problem to include repairperson assignment decisions and Zhang et al. (2023b) proposed a resource-constrained fleet maintenance optimization model with the objective of maximizing the expected fleet readiness under stochastic break durations.

Despite the above-mentioned extensions, several key limitations remain in the existing FSMP literature. First, all FSMP models rely on a restrictive assumption according to which missions are identical and require a single system configuration for successful completion. However, this assumption does not hold in many real-world scenarios. For example, in a military aircraft fleet, missions often require different system configurations and have different durations in accordance with mission goals Shahmoradi-Moghadam et al. (2021). Nearly all proposed models assume that all upcoming missions must be completed. However, due to resource limitations, it may not be possible to ready all systems within the available break duration, making mission prioritization and assignment a necessary consideration. Additionally, all FSMP models mentioned above fail to consider the inherent uncertainties in both maintenance and break durations. Finally, the FSMP models that have been discussed thus far rely on statistical/model-based methods for system reliability computation. However, with the recent advances in data collection with sensors and the advent of IIoT, data-driven methods relying on DNN have become increasingly popular for the remaining useful life (RUL) prediction and reliability assessment Namuduri et al. (2020). These methods have achieved state-of-the-art performance in terms of prognostics and predictive maintenance. Although data-driven SM strategies have recently been studied, the few existing works, *e.g.,* O'Neil et al. (2024); Hesabi et al. (2022), assume that all system components are sensor monitored. In practice, however, many industrial systems are composed of both standard components with known lifetime distributions, and critical components which are sensor-monitored. Thus, a hybrid approach in needed for the reliability modeling and assessment of such systems. These limitations motivate the need for the development of a generalized FSMP framework that incorporates stochastic maintenance and break durations, mission prioritization and assignment, and hybrid reliability assessment techniques.

Aiming to help maintenance decision-makers resolve practical issues arising in the implementation of the FSMP, this section develops a new approach where the classical FSMP is extended to include flexible system configuration adapted to multiple mission types with unique operational demands, reliability requirements, minimum system assignment thresholds, and associated penalties. Maintenance activities are permitted during the break duration to ensure that systems assigned to specific missions meet the pre-specified performance requirements. Unlike existing studies, the proposed approach introduces a new FSMP model that jointly determines four key decisions: (1) the assignment of systems to mission types, (2) the selection of components within each system to maintain, (3) the corresponding maintenance levels to be performed on the selected components, and (4) the assignment of maintenance tasks to repairpersons. Inevitable uncertainties in maintenance and break durations are also explicitly accounted for. A chance-constrained optimization model is then developed and reformulated using the Conditional Value-at-Risk (CVaR) approximation. A fleet of identical multi-component systems is considered, where components are of two types: standard and critical. The first type of components (*i.e.,* standard) are assumed to follow known lifetime distributions, allowing for the analytical reliability calculations. The second type of components (*i.e.,* critical) are sensor-monitored with reliability estimates obtained using a DNN with MCD, a technique that has demonstrated outstanding performance in predictive maintenance (Lee and Mitici, 2023; Zhuang et al., 2023; O'Neil et al., 2024).

A preliminary version of the present work appeared in a conference paper (O'Neil et al., 2024) where the SMP is solved for a single MOS operating under uncertain maintenance and break durations. The resulting optimization model maximized the system reliability for the next mission. The present work is a significant extension of (O'Neil et al., 2024) as it expands the complexity of the model and analyses. Major contributions of the extended work include: i) an expanded literature review covering recent advances in SMP; ii) an extension from a single MOS to a fleet of several MOS, where systems are operating under break duration uncertainties; iii) systems are composed of two types of components and DNN are used to predict the RUL of sensor-monitored components; v) a new FSMP formulation and efficient solution method are presented; vi) several numerical experiments are conducted to highlight the proposed approach.

This section is organized as follows. Subsection 3.2.2 provides a comprehensive description of the system and the problem under investigation. This section also develops the reliability and maintenance models for both standard and sensor-monitored components. In Subsection 3.2.3, the novel deterministic formulation of the FSMP is introduced, whereas Subsection 3.2.4 presents the chance-constrained formulation and its CVaR approximation. The mathematical model is validated through a case study involving the maintenance of a fleet of military aircraft. The numerical results are analyzed and extensively discussed in Subsection 3.2.5. Finally, Subsection 3.2.6 concludes the section by summarizing the key findings and identifying promising directions for future research.

3.2.2 Problem Description

In this section, the problem considered is described and the main working assumptions are presented. First, the notation system is introduced.

Notation

The following notation system is used in the modeling of the problem.

Sets:

$\mathcal{I}$	Set of systems, $\mathcal{I} = \{1, \dots, I\}$, with index i.
$\mathcal{J}$	Set of subsystems per system. $\mathcal{J} = \{1, \dots, J\}$, with index j.
$\mathcal{J}_m$	Set of subsystems required to perform mission type m; $\mathcal{J}_m \subseteq \mathcal{J}$.
$\mathcal{K}_j$	Set of all components in subsystem j. $\mathcal{K}_j = \{1, \dots, J_j\}$, with index k.
$\mathcal{K}_j^{\mathrm{s}}$	Set of standard components in subsystem j. $\mathcal{K}_j^{\mathrm{s}} \subseteq \mathcal{K}_j$, with index k.
$\bar{\mathcal{K}}_j^{\mathrm{s}}$	Set of sensor-monitored components in subsystem j. $\bar{\mathcal{K}}_j^{s} = \mathcal{K}_j \setminus \mathcal{K}_j^{s}$.
$\mathcal{L}_{ijk}$	Set of maintenance levels available for component E_{ijk} $\mathcal{L}_{ijk} = \{1, \dots, L_{ijk}\}$, with index l

$\mathcal{Q}$	Set of repairpersons , $\mathcal{Q} = \{1, \ldots, Q\}$, with index q.
$\mathcal{M}$	Set of mission types, $\mathcal{M} = \{1, \ldots, M\}$, with index m.

Parameters:

I	Number of systems within the fleet
J	Number of subsystems per system.
J_j	Number of components in subsystem j
E_{ijk}	The kth component in subsystem j of system i
B_{ijk}	Age of component E_{ijk} at the start of the break
S_{ijk}	Status of E_{ijk} at the start of the break
L_{ijk}	Number of maintenance levels available for component E_{ijk}
D	Duration of maintenance break
Q	Maximum number of repair persons available
c_q^f	Fixed cost incurred for using repairperson q
t_{ijkl}^p	Time to carry out PM level l on component E_{ijk}
t_{ijkl}^c	Time to carry out CM level l on component E_{ijk}
M	Number of mission types
U_m	Length of mission type m
ρ_m	Penalty incurred for not completing mission type m
$R_{jm}^{\min}$	Minimum required reliability for subsystem j for mission type m

Decision variables:

χ_m	Binary variable equal to 1 if mission type m is completed
γ_q	Binary variable equal to 1 if repair person q is utilized
z_{im}	Binary variable equal to 1 if system i is assigned to mission type m
x_{ijklq}	Binary variable equal to 1 if repairperson q performs maintenance level l on component E_{ijk}

Main working assumptions

1. Components and systems are binary state. This assumption is appropriate for a wide range of systems and is commonly used in the literature (Rice, 1999; Chaabane et al., 2020; Amjadian et al., 2024).
2. During the maintenance break, the systems are assumed to be switched off and therefore not experiencing any degradation. This is a reasonable and commonly used assumption (Diallo et al., 2019; O'Neil et al., 2024).
3. Maintenance activities are allowed only during the break duration. This is the definition of the SMP (Rice, 1999; Cassady et al., 2003; Diallo et al., 2018).
4. Maintenance and break durations are stochastic quantities with known probability distributions. This is reasonable and commonly used (Jiang and Liu, 2020b; Jin et al., 2023; Liu et al., 2018).

Problem description

The FSMP addressed in this work deals with the optimization of the SM planning, repairperson scheduling, as well as system-to-mission assignment. This problem is inspired by the maintenance scheduling challenges encountered in maintaining military aircraft fleets. Specifically, it is assumed that a fleet of systems has just entered a maintenance break, providing the opportunity to perform necessary maintenance actions to get the fleet ready for a set of various upcoming missions. During this break, the goal is to determine the assignment of systems to mission types, the set of components within each system should be maintained, the required number of repairpersons needed to carry out the selected maintenance plan, and the assignment of maintenance tasks to repairpersons. The objective is to minimize the grand total cost, which includes penalties for unselected missions as well as fixed and variable costs related to selected maintenance actions. Maintenance plays an important role in ensuring systems are mission-ready by addressing reliability requirements. A system can only be assigned to a mission provided that

its reliability meets the required threshold. Each mission type has unique operational demands, reliability requirements, minimum system assignment thresholds, and associated penalties. However, due to resource limitations, such as the availability of repairpersons and time constraints, only a subset of all desirable maintenance actions can be performed. In the FSMP addressed, maintenance and break durations are stochastic, following known probability distributions. The uncertainty in task and break durations impacts planning by introducing the risk of potentially not completing the required maintenance actions during the break and thus delaying mission readiness. Effective maintenance planning must account for this uncertainty to ensure cost-effective, and robust solutions.

Each mission type $m \in \mathcal{M}$ is characterized by a profile captured by the vector $(\rho_m, U_m, \zeta_m, \mathcal{J}_m, \mathcal{R}_m)$, where ρ_m is the penalty incurred for not completing mission m, U_m is the mission duration, and ζ_m is the minimum number of systems required to successfully operate mission m. The parameter set $\mathcal{J}_m$ corresponds to the subset of subsystems or functions that are required to perform mission m. For example, in the context of a military aircraft fleet, a logistics support mission would likely not require the functionality of weapon systems, whereas a combat mission certainly would. Consequently, a system assigned to a logistics support mission would not require maintenance on those subsystems. Finally, the parameter $\mathcal{R}_m$ defines the set of minimum reliability requirements for all subsystems in the subset $\mathcal{J}_m$. To perform the different mission types $m \in \mathcal{M}$, a fleet of I identical multi-component MOS is available. Each MOS i is comprised of J subsystems in series, and each subsystem is composed of J_j statistically independent components E_{ijk} in parallel. A subsystem j functions if at least 1 out of its J_j components is functioning. At the start of the current maintenance break, each component E_{ijk} is characterized by its current effective age φ_{ijk}, while its status is determined by the binary state variable u_{ijk} defined as follows:

$$u_{ijk} = \begin{cases} 1 & \text{if } E_{ijk} \text{ is functioning at the start of the break,} \\ 0 & \text{otherwise.} \end{cases} \tag{3.51}$$

The components within each system are divided into two sets based on the availability of reliability information. First, there are standard components, defined by the set $\mathcal{K}_j^{\mathrm{s}}$, which have lifetime distributions following known probability distributions, allowing their reliability to be computed analytically. Conversely, sensor-monitored components, defined by the set $\bar{\mathcal{K}}_j^{\mathrm{s}}$, are equipped with sensors that provide real-time condition monitoring data (*e.g.,* vibration, temperature, wear). These data are used to predict the RUL of each component. Data-driven approaches relying on DNN have shown exceptional performance in predicting the probability of failure of such components (Namuduri et al., 2020). In the present work, a DNN with MCD is used to estimate the probability that these sensor-monitored components will successfully operate the assigned mission. A detailed description of the reliability computation for both standard and sensor-monitored components is provided in the next subsection.

Reliability computation

System i can be assigned to mission m if and only if each subsystem j of its j_i subsystems meet the required subsystem reliability threshold $R_{jm}^{\min}$ for that mission type. For each mission m, specific reliability requirements $\mathcal{R}_m$ exist for different functions or subsystems $\mathcal{J}_m$. To compute the reliability of each subsystem $j \in \mathcal{J}_m$, the reliability of each component E_{ijk} must first be determined. If E_{ijk} is a standard component with known lifetime distributions and effective age ψ_{ijk} at the end of the break, then its reliability $\mathcal{R}_{ijk}(t|\psi_{ijk})$ is computed as:

$$\mathcal{R}_{ijk}(t|\psi_{ijk}) = \frac{\mathcal{R}_{ijk}(\psi_{ijk} + t)}{\mathcal{R}_{ijk}(\psi_{ijk})}, \tag{3.52}$$

where $\mathcal{R}_{ijk}(\cdot)$ is the unconditional reliability function of component E_{ijk}. Without loss of generality, in this work, the lifetime of a standard component E_{ijk} is assumed to be governed by a Weibull distribution, with β_{ijk} and η_{ijk} being its respective shape and scale parameters. Therefore, its unconditional reliability function is $\mathcal{R}_{ijk}(t) = \exp\left[-\left(\frac{t}{\eta_{ijk}}\right)^{\beta_{ijk}}\right]$.

For sensor-monitored components, a common set of S sensors are available to monitor components' degradations, where new sensor measurements are recorded after the completion of each operating cycle.

For a given component E_{ijk} that has operated for φ_{ijk} cycles, the overall degradation measurement can be stored in a matrix $\mathbf{G}_{ijk}$ as follows:

$$\mathbf{G}_{ijk} = \begin{bmatrix} \boldsymbol{g}^1 \\ \boldsymbol{g}^2 \\ \vdots \\ \boldsymbol{g}^{\varphi_{ijk}} \end{bmatrix} = \begin{bmatrix} g_1^1 & g_2^1 & \cdots & g_S^1 \\ g_1^2 & g_2^2 & \cdots & g_S^2 \\ \vdots & \vdots & & \vdots \\ g_1^{\varphi_{ijk}} & g_2^{\varphi_{ijk}} & \cdots & g_S^{\varphi_{ijk}} \end{bmatrix},$$

where $\boldsymbol{g}^c$ is a vector of all sensor measurements collected during cycle c ($c = 1, \ldots, \varphi_{ijk}$). To compute the reliability of such components, a DNN with MCD is employed. The DNN, denoted as $f(\cdot;\theta)$, is parameterized by weights and biases θ and aims to learn a mapping between current sensor measurements and the true RUL. To learn this mapping, the network is trained on historical degradation data, where the true RUL values are known. Let the training set be defined as $\{(\mathbf{G}_\lambda, RUL_\lambda)\}_{\lambda=1}^{\Lambda}$, where $\mathbf{G}_\lambda$ is a matrix of historical sensor recordings and RUL_λ is the true remaining lifetime. The training process involves identifying the set of learnable weights and biases θ that minimize the following loss function:

$$\min_{\theta} \quad \Xi = \frac{1}{\Lambda} \sum_{\lambda=1}^{\Lambda} \left(RUL_\lambda - f(\mathbf{G}_\lambda;\theta)\right)^2 + \mu \cdot \Big|\Big|\theta\Big|\Big|_2^2. \tag{3.53}$$

The loss function being minimized is the mean squared error with L2 regularization, a typical loss function used for regression tasks. Once trained, the model is capable of predicting the RUL of components currently in operation. Let E_{ijk} be a sensor-monitored component with sensor measurements $\mathbf{G}_{ijk}$, the predicted RUL can then be obtained as $\widehat{RUL}_{ijk} = f(\mathbf{G}_{ijk};\theta)$. If the trained model were capable of predicting the RUL with no error, then the probability of component E_{ijk} completing the upcoming mission would be known with certainty. However, in practice, developing a perfect RUL predictive model is not possible due to model limitations, noisy data, and other factors. To account for uncertainty in the prediction, MCD is implemented. Typically, dropout is a regularization technique that is applied only during the training phase of a neural network. MCD is an approach that extends dropout to the inference phase (Gal and Ghahramani, 2016). During inference, rather than using the trained model to make a single prediction, MCD involves performing Ω forward passes with dropout applied each time. This produces Ω different RUL predictions. Let $\left(\widehat{\mathrm{RUL}}_{ijk}^{1}, \ldots, \widehat{\mathrm{RUL}}_{ijk}^{\Omega}\right)$ represent the Ω RUL predictions obtained for component E_{ijk}. Accordingly, the reliability of component E_{ijk} after completing φ_{ijk} cycles is computed as follows:

$$\mathcal{R}_{ijk}^{G}(t|\mathbf{G}_{ijk}) = \frac{1}{\Omega} \sum_{\omega=1}^{\Omega} \mathbb{1}\Big(\widehat{\mathrm{RUL}}_{ijk}^{\omega} > t\Big), \tag{3.54}$$

where $\mathbb{1}(\cdot)$ is the indicator function taking a value of 1 if the condition in the brackets is true, and 0 otherwise.

Given the individual component reliabilities, the reliability of subsystem $j \in \mathcal{J}$ of system $i \in \mathcal{I}$ is given by (3.55) as its configuration is parallel:

$$\mathcal{R}_{ij}(t) = 1 - \prod_{k \in \mathcal{K}_j^{\mathrm{s}}} \left(1 - \mathcal{R}_{ijk}\left(t|\psi_{ijk}\right) v_{ijk}\right) \cdot \prod_{k \in \bar{\mathcal{K}}_j^{\mathrm{s}}} \left(1 - \mathcal{R}_{ijk}^{G}\left(t|\mathbf{G}_{ijk}\right) v_{ijk}\right). \tag{3.55}$$

Maintenance modeling

During the intermission break, each failed component E_{ijk} can be subjected to a CM of level $l \in \{0, \ldots, L_{ijk}\}$. The lowest maintenance level $l = 0$ corresponds to the DN case, while the highest level $l = L_{ijk}$ corresponds to the perfect corrective repair or AGAN, and level $l = 1$ corresponds to mR, which brings the component to an ABAO condition. Intermediate values of l ($1 < l < L_{ijk}$) represent IM actions that bring the component back to a condition between AGAN and ABAO. Similarly, a working component can be subjected to a PM of level $l \in \{0, \ldots, L_{ijk}\}$. PM levels $l = 0$ and $l = L_{ijk}$ correspond to the DN and perfect preventive maintenance (AGAN) cases, respectively. Following the work

by (Chaabane et al., 2020), there is no mR equivalent for a PM, and the maintenance option $l = 1$ is assumed to be equivalent to the DN case. Intermediate values of l ($0 < l < L_{ijk}$) represent IM actions.

For standard components, the age reduction approach is used to model the impact of maintenance actions Malik (1979). When component E_{ijk} is subjected to maintenance level l, its age φ_{ijk} is reduced by an age reduction factor, thereby increasing its reliability. The age reduction coefficients are defined separately for corrective and preventive maintenance actions as follows:

— Corrective maintenance: the age reduction factor is μ^c_{ijkl}, where $0 \leq \mu^c_{ijkl} \leq 1$, and the maintenance action requires a duration of t^c_{ijkl}.
— Preventive maintenance: the age reduction factor is μ^p_{ijkl}, where $0 \leq \mu^p_{ijkl} \leq 1$, and the maintenance action requires a duration of t^p_{ijkl}.

The time required to perform maintenance level l on standard component E_{ijk} can be formulated as:

$$t_{ijkl} = t^p_{ijkl} \cdot u_{ijk} + t^c_{ijkl} \cdot (1 - u_{ijk}) \tag{3.56}$$

The reliability r_{ijklm} of component E_{ijk} after maintenance, given that maintenance level l has been selected and system i has been assigned to mission m, is defined as:

$$r_{ijklm} = \mathcal{R}_{ijk}(U_m|\psi_{ijk}) \cdot v_{ijkl} \tag{3.57}$$

where $v_{ijkl} = \mathbb{1}(u_{ijk} = 1 \vee l > 0)$ is the status of component E_{ijk} at the end of the break, and ψ_{ijkl} is the effective age of E_{ijk} at the end of the break. This quantity is defined as:

$$\psi_{ijkl} = \varphi_{ijk} \left(u_{ijk} \cdot \mu^p_{ijkl} + (1 - u_{ijk}) \cdot \mu^c_{ijkl} \right). \tag{3.58}$$

By analogy to the age reduction factor used to model IM, proposed by Malik (1979), we propose an RUL increase factor Δ_{ijkl} to model IM for sensor-monitored components. When sensor-monitored component E_{ijk} is subjected to maintenance level l, its RUL is extended by Δ_{ijkl}. The factor Δ_{ijkl} increases with respect to the maintenance levels. Thus, the reliability r_{ijklm} of component E_{ijk}, given that maintenance level l has been selected and system i has been assigned to mission m, is defined as:

$$r_{ijklm} = \frac{1}{\Omega} \sum_{\omega=1}^{\Omega} \mathbb{1}\left(\widehat{\mathrm{RUL}}^{\omega}_{ijk} + \Delta_{ijkl} > U_m \right), \tag{3.59}$$

In practice, a dataset containing sensor data across multiple operating cycles of a machine undergoing maintenance and ML techniques can be used to predict how different maintenance actions influence RUL. By analyzing historical sensor trends before and after maintenance, such models can provide data-driven estimates of Δ_{ijkl}. The time required to perform maintenance level l on sensor-monitored component E_{ijk} follows Equation (3.56).

3.2.3 The Deterministic Formulation

The fleet of systems under study is assumed to have just completed a mission and is entering a maintenance break. Due to limited resources, maintenance actions can be performed only on a subset of components across all systems. The goal of the proposed optimization model is to jointly select the components to be maintained, determine the maintenance levels for the selected components, assign repairpersons to maintenance tasks, and assign systems to mission types so that the total expected cost $\mathcal{Z}$ is minimized. The resulting optimization problem is formulated as the following MINLP:

$$\min \ \mathcal{Z} = \sum_{m\in\mathcal{M}} \rho_m\,(1-\chi_m) + \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}_j}\sum_{l\in\mathcal{L}_j}\sum_{q\in\mathcal{Q}} c^v\, t_{ijkl}\, x_{ijklq} + \sum_{q\in\mathcal{Q}} c^f\,\gamma_q \tag{3.60a}$$

$$\text{s.t.} \sum_{m\in\mathcal{M}} z_{im} \le 1, \qquad \forall i \in \mathcal{I} \tag{3.60b}$$

$$\sum_{i\in\mathcal{I}} z_{im} \ge \zeta_m\,\chi_m, \qquad \forall m \in \mathcal{M} \tag{3.60c}$$

$$\sum_{l\in\mathcal{L}_{ijk}}\sum_{q\in\mathcal{Q}} x_{ijklq} = 1, \qquad \forall i \in \mathcal{I}, \forall j \in \mathcal{J}, \forall k \in \mathcal{K}_j \tag{3.60d}$$

$$1 - \prod_{k\in\mathcal{K}_j}\left(1 - \sum_{l\in\mathcal{L}_{ijk}}\sum_{q\in\mathcal{Q}} x_{ijklq}\, r_{ijklm}\right) \ge R^{\min}_{jm}\cdot z_{im}, \quad \forall i, \forall m, \forall j \tag{3.60e}$$

$$\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}_j}\sum_{l\in\mathcal{L}_{ijk}\setminus\{0\}} x_{ijklq}\, t_{ijkl} \le D\,\gamma_q, \qquad \forall q \in \mathcal{Q} \tag{3.60f}$$

$$\chi_m \in \{0,1\} \tag{3.60g}$$

$$\gamma_q \in \{0,1\} \tag{3.60h}$$

$$z_{im} \in \{0,1\} \tag{3.60i}$$

$$x_{ijklq} \in \{0,1\} \tag{3.60j}$$

The objective function (3.60a) minimizes the total expected cost. The first term is the cost associated with unselected mission types and the second and third terms account for the cost of performing the selected maintenance actions and the fixed cost of hiring the necessary repairpersons. Constraints (3.60b) ensure that each system is assigned to no more than one mission type, while Constraints (3.60c) enforce the requirement that a minimum number of systems must be assigned to a mission type to be successful and thus not incur the penalty cost. Constraints (3.60d) enforce a single maintenance action to be selected for each component. Constraints (3.60e) ensure that if a system is assigned to a mission type, the reliability of the required subsystems must exceed a defined threshold (*i.e.,* minimum reliability). Finally, Constraints (3.60f) guarantee that the selected maintenance actions can be performed within the break duration, and Constraints (3.60g)-(3.60j) are binary variable restrictions.

The above formulation is difficult to solve optimally especially for large fleets composed of systems with many subsystems and components due to the nonlinear reliability expression in Constraints (3.60e). However, these Constraints (3.60e) can equivalently be reformulated as linear constraints as shown through the following transformation:

$$\prod_{k\in\mathcal{K}_j}\left(1 - \sum_{l\in\mathcal{L}_{ijk}}\sum_{q\in\mathcal{Q}} x_{ijklq}\cdot r_{ijklm}\right) \le 1 - R^{\min}_{jm}\cdot z_{im}, \forall i \in \mathcal{I}, \forall m \in \mathcal{M}, \forall j \in \mathcal{J}_m$$

$$\ln\left(\prod_{k\in\mathcal{K}_j}\left(1 - \sum_{l\in\mathcal{L}_{ijk}}\sum_{q\in\mathcal{Q}} x_{ijklq}\cdot r_{ijklm}\right)\right) \le \ln\left(1 - R^{\min}_{jm}\cdot z_{im}\right), \forall i, \forall m, \forall j$$

$$\sum_{k\in\mathcal{K}_j}\ln\left(1 - \sum_{l\in\mathcal{L}_{ijk}}\sum_{q\in\mathcal{Q}} x_{ijklq}\cdot r_{ijklm}\right) \le \ln\left(1 - R^{\min}_{jm}\cdot z_{im}\right), \forall i, \forall m, \forall j$$

Due to the binary nature of the decision variables x_{ijklq} and z_{im}, they can be factored out of the logarithm expressions leading to the linearization of Constraints (3.60e) as:

$$\sum_{k\in\mathcal{K}_j}\sum_{l\in\mathcal{L}_{ijk}}\sum_{q\in\mathcal{Q}} \ln\left(1 - r_{ijklm}\right)\cdot x_{ijklq} \le \ln\left(1 - R^{\min}_{jm}\right)\cdot z_{im}, \quad \forall i, \forall m, \forall j \tag{3.61}$$

In the following section, the methods used to deal with uncertainties in maintenance and break durations is presented and fully discussed.

3.2.4 Handling Uncertainty in Maintenance and Break Durations

Variability and uncertainty in maintenance durations are inevitable due to factors such as the operating environment, the complexity of repair tasks, and the skill level of repair personnel. Furthermore, in many applications, it would be extremely difficult or even impossible to precisely forecast the exact duration of the maintenance break. For example, in military aircraft fleets, the unexpected start of the next mission/battle introduces uncertainties in the available maintenance time window Liu et al. (2018). Ignoring these potential uncertainties can result in a lack of readiness and exposure to additional risk. Rather than treating maintenance and break durations as deterministic parameters, in the present work, these durations are considered as stochastic quantities with known probability distributions. Thus, in the following, the time required to perform maintenance level l on component E_{ijk} is represented by the random variable $\tilde{t}_{ijkl}$ with $\mathbb{E}[\tilde{t}_{ijkl}]$ as its expectation, while the uncertain break duration is represented by the random variable $\tilde{D}$. The stochastic version of the proposed FSMP is then formulated using chance constraints as follows:

$$\min \ \mathcal{Z} = \sum_{m\in\mathcal{M}} \rho_m \left(1-\chi_m\right) + \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}_j}\sum_{l\in\mathcal{L}_j}\sum_{q\in\mathcal{Q}} c^v\, \mathbb{E}[\tilde{t}_{ijkl}]\, x_{ijklq} + \sum_{q\in\mathcal{Q}} c^f\, \gamma_q \tag{3.62a}$$

$$\text{s.t.} \quad \mathbb{P}\left(\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}_j}\sum_{l\in\mathcal{L}_j\setminus\{0\}} \tilde{t}_{ijkl}\, x_{ijklq} \le \tilde{D}\,\gamma_q\right) \ge 1-\alpha, \quad \forall q\in\mathcal{Q} \tag{3.62b}$$

$$(3.60\text{b}) - (3.60\text{d})$$

$$(3.61)$$

$$(3.60\text{g}) - (3.60\text{j})$$

The chance-constrained formulation above ensures that the total maintenance time for each repairperson $q \in \mathcal{Q}$ does not exceed the allocated break duration with a specified confidence level $1-\alpha$. The remaining constraints are those previously described.

Let $\mathbf{x_q}$ denote the vector of all maintenance actions assigned to repairperson q. For such a vector, we associate the random variable Y_q defined as:

$$Y_q := \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}_j}\sum_{l\in\mathcal{L}_j\setminus\{0\}} \tilde{t}_{ijkl}\, x_{ijklq} - \tilde{D} \tag{3.63}$$

If furthermore we let $F_{Y_q}(y) = \mathbb{P}\left(Y_q \le y\right)$ be the cumulative distribution function (CDF) of Y_q, it follows that the maintenance task assignment $\mathbf{x_q}$ to repairperson q satisfies the chance constraints of (3.62b) if $F_{Y_q}(0) \ge 1-\alpha$. However, Y_q is the sum of several random variables, making the computation of $F_{Y_q}(y)$ intractable analytically due to the required multi-dimensional integral convolutions. To address this challenge, the well-known Conditional Value-at-Risk (CVaR) approximation technique is used. This method indeed transforms the probabilistic constraints into deterministic approximations that are computationally more tractable. When the number of realizations N of the random parameters is finite, the probabilistic constraints (3.62b) are approximated using the CVaR constraints:

$$\inf_{t_q}\left\{ t_q + \alpha^{-1}\sum_{n=1}^{N} p_n \left[\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}_j}\sum_{l\in\mathcal{L}_j\setminus\{0\}} t_{ijkl,n}\, x_{ijklq} - D_n\,\gamma_q - t_q\right]_+ \right\} \le 0, \quad \forall q\in\mathcal{Q} \tag{3.64}$$

where $[\nu]_+ = \max(\nu, 0)$ denotes the positive part of a real number ν, $t_{ijkl,n}$ and D_n are the nth samples of t_{ijkl} and D, respectively, and p_n is the nth corresponding weight. Parameter α is the confidence level specified as the risk measure to model the decision makers' risk preferences regarding the randomness of the maintenance time and the break duration. By introducing auxiliary variables $u_{q,n}$ and t_q, the

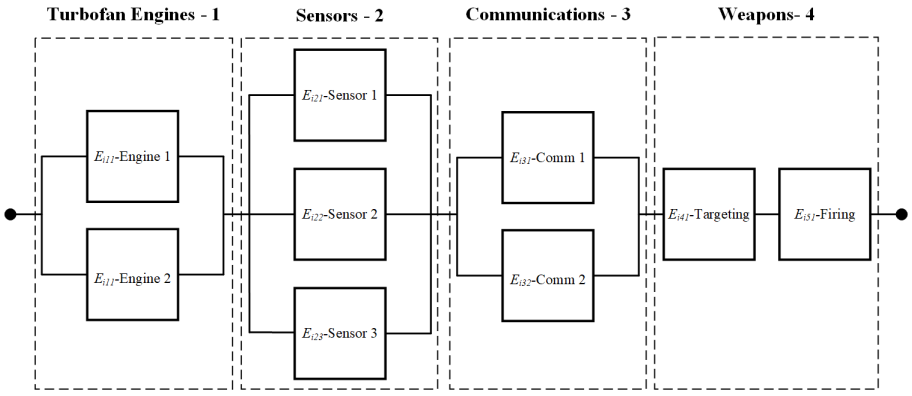


Figure 3.10 – Aircraft RBD considered for the case study.

constraints can be reformulated as:

$$t_q + \alpha^{-1} \sum_{n=1}^{N} p_n \, u_{q,n} \leq 0, \qquad \forall q \in \mathcal{Q} \tag{3.65a}$$

$$u_{q,n} \geq \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}_i} \sum_{k \in \mathcal{K}_j} \sum_{l \in \mathcal{L}_{ij} \setminus \{0\}} t_{ijkl,n} \, x_{ijklq} - D_n \, \gamma_q - t_q, \quad \forall q \in \mathcal{Q}, \forall n \in \{1, \ldots, N\} \tag{3.65b}$$

$$u_{q,n} \geq 0, \qquad \forall q \in \mathcal{Q}, \forall n \in \{1, \ldots, N\} \tag{3.65c}$$

Regarding the derivation of the CVaR constraints, the reader can find more details in Elçi and Noyan (2018). The CVaR approximation is more conservative than other methods like sample average approximation (SAA); however, it offers computational efficiency as only continuous variables are introduced.

3.2.5 Case Study: Military Aircraft Fleet Maintenance

This section illustrates how the proposed optimization model can deal with joint maintenance, repairpersons assignment decisions, and assignment of systems to mission types for a fleet composed of $I = 4$ military aircraft. According to the reliability block diagram in Figure 3.10, each aircraft is comprised of four key subsystems: propulsion (engines), sensors and communication, and combat systems (targeting, offensive weapons). It is assumed that the engines are sensor-monitored while all other components are standard with known lifetime distributions. The NASA C-MAPSS dataset is used to train the DNN for the engines. This dataset provides degradation trajectories under 4 different operational conditions and fault modes for turbofan engines. Because these different sets consider different operational conditions and fault modes, they cannot be jointly studied. In this work, the first set called FD001, which considers a single fault mode, is selected. This set provides run-to-failure information for 100 engines and includes 3 operational settings and 21 sensor values for each operating cycle. The detailed description of the 21 sensor variables can be found in Saxena et al. (2008). In total, 20,631 cycles are available for training. Each engine starts with a different degree of initial wear and manufacturing variation. The operational settings and sensor values are provided for every cycle up until engine failure. A second test set is also provided that includes truncated time series of various lengths before failure for 100 engines and actual RULs.

The engine data from FD001 is randomly assigned to aircraft engines to simulate a diverse range of initial health conditions. The engine assignment, RUL point prediction, and predicted reliability can be found in Table 3.7. Table 3.8 gives the parameters of the lifetime distributions of the standard components. Let us recall that, for the sensor-monitored engines, a single maintenance action is available, which restores the engines to the AGAN state. For standard components, a list of four potential maintenance levels is available, including DN, mR (for failed components only), IM, and perfect repair. All maintenance durations are assumed to follow a Gamma distribution, where the respective shape and scale parameters ξ_{sh} and ξ_{sc} are reported in Table 3.9. Table 3.10 provides a detailed description of the different mission

types considered, including the penalty for unassigned mission, duration, and other relevant input data. It is important to note that operational data for sensor-monitored components is recorded in cycles, while the lifetimes of standard components are measured in operating hours. Accordingly, Table 3.10 provides the mission durations in both operating hours and cycles.

Table 3.7 – Engine assignment.

Components	Test Engine ID	True RUL (cycles)	Predicted RUL (cycles)	R_{ijk1} ($U_1 = 10$)	R_{ijk2} ($U_2 = 25$)	R_{ijk3} ($U_3 = 40$)
$E_{1,1,1}$	17	50	54	1.000	1.000	0.994
$E_{1,1,2}$	18	28	32	1.000	0.966	0.014
$E_{2,1,1}$	20	16	13	0.912	0.000	0.000
$E_{2,1,2}$	21	57	68	1.000	1.000	1.000
$E_{3,1,1}$	24	20	20	1.000	0.054	0.000
$E_{3,1,2}$	27	66	67	1.000	1.000	1.000
$E_{4,1,1}$	31	8	8	0.214	0.000	0.000
$E_{4,1,2}$	32	48	47	1.000	1.000	0.912

The following subsections will first present the results of the DNN with MCD for reliability prediction, followed by multiple experiments to demonstrate the impact of stochastic durations and the risk confidence level parameter α on the maintenance plan and the overall assignment decisions.

Performance Analysis of DNN with MCD

In this work, the DL architecture implemented for RUL prediction is a RNN, specifically a Bi-LSTM network. LSTM are a type of RNN designed to overcome the vanishing and exploding gradient problem frequently encountered in standard RNN. Bi-LSTM comprise two LSTM cells and have demonstrated exceptional performance in predicting remaining useful life (O'Neil et al., 2024; Zhuang et al., 2023). The main advantage of using Bi-LSTM instead of standard RNN is that they can learn the temporal relationships in the forward and backward directions simultaneously. For detailed information on the inner workings of the LSTM cells, the reader is referred to Zhang et al. (2018). Before training the Bi-LSTM for RUL prediction, the NASA C-MAPSS FD001 dataset was pre-processed and cleaned according to the following steps:

— **Data labeling**. A column was added to both the training and testing sets that specified the RUL after completion of each cycle. This column provides the labels that will be used to train and test the models.

— **Input data scaling**. It is widely acknowledged that scaling input data for neural networks leads to more efficient training. The min-max normalization approach was used to scale the input of the DL algorithm to be in the range $[0, 1]$. The min-max normalization of data is performed according to the following formula:

$$\check{x}_j^i = \frac{x_{ij} - x_j^{min}}{x_j^{max} - x_j^{min}}, \tag{3.66}$$

where x_j^i denotes the ith data measure of the jth sensor, $\check{x}_j^i$ is the scaled value of x_j^i, x_j^{max} and x_j^{min} are the respective maximum and minimum raw measurements from the jth sensor.

— **Data reformatting**. In preparation for fitting the Bi-LSTM, the training and testing data was reformatted into 3-dimensional tensors of the form (N_s, N_{ts}, N_f). N_s refers to the number of samples, N_{ts} denotes the time sequence dimension, and N_f refers to the number of selected features.

For the Bi-LSTM, the hyperparameters proposed in O'Neil et al. (2024) were adopted. The following three commonly used metrics are exploited to evaluate the proposed Bi-LSTM performance for point prediction: root mean squared error (RMSE), scoring function (SC), and accuracy (AC). Smaller values of RMSE and SC, along with higher values of AC are preferred. Assuming a total of Λ test examples with Ω predictions for each test example $\lambda \in \{1, \dots \Lambda\}$, we define $d_\lambda = \left(\frac{1}{\Omega}\sum_{\omega=1}^{\Omega} \widehat{\mathrm{RUL}}_{\lambda\omega}\right) - \mathrm{RUL}_\lambda$ as the

Table 3.8 – Component data for standard components.

	Initial age (φ_{ijk})	Initial state (u_{ijk})	Weibull shape parameter (β_{ijk})	Weibull scale parameter (η_{ijk})
$E_{1,2,1}$	19	1	1.8	30
$E_{1,2,2}$	30	0	1.8	30
$E_{1,2,3}$	17	1	1.8	30
$E_{1,3,1}$	24	0	2.4	40
$E_{1,3,2}$	25	0	2.4	40
$E_{1,4,1}$	42	1	2.2	45
$E_{1,5,1}$	20	0	2.5	30
$E_{2,2,1}$	22	0	1.8	30
$E_{2,2,2}$	20	0	1.8	30
$E_{2,2,3}$	16	1	1.8	30
$E_{2,3,1}$	39	0	2.4	40
$E_{2,3,2}$	24	0	2.4	40
$E_{2,4,1}$	22	1	2.2	45
$E_{2,5,1}$	17	1	2.5	30
$E_{3,2,1}$	21	0	1.8	30
$E_{3,2,2}$	17	1	1.8	30
$E_{3,2,3}$	24	0	1.8	30
$E_{3,3,1}$	26	0	2.4	40
$E_{3,3,2}$	22	1	2.4	40
$E_{3,4,1}$	38	1	2.2	45
$E_{3,5,1}$	18	0	2.5	30
$E_{4,2,1}$	25	0	1.8	30
$E_{4,2,2}$	21	1	1.8	30
$E_{4,2,3}$	24	0	1.8	30
$E_{4,3,1}$	31	0	2.4	40
$E_{4,3,2}$	27	0	2.4	40
$E_{4,4,1}$	43	1	2.2	45
$E_{4,5,1}$	19	1	2.5	30

Table 3.9 – Maintenance duration parameters for different components and maintenance levels.

E_{ijk}	t^p_{ij2}		t^p_{ij3}		t^c_{ij1}		t^c_{ij2}		t^c_{ij3}	
	ξ_{sc}	ξ_{sh}	ξ_{sc}	ξ_{sh}	ξ_{sc}	ξ_{sh}	ξ_{sc}	ξ_{sh}	ξ_{sc}	ξ_{sh}
E_{i1k}	–	–	41.4	2.56	–	–	–	–	44.4	2.56
E_{i2k}	20.5	2.28	31.4	2.78	8.9	2.20	23.5	2.28	34.4	2.78
E_{i3k}	18.1	2.17	35.6	2.67	13.0	1.84	21.1	2.17	38.6	2.67
E_{i4k}	15.9	1.71	32.1	2.21	14.7	2.29	18.9	2.32	35.1	2.82
E_{i5k}	16.7	1.72	33.7	2.22	16.8	1.88	25.7	1.94	37.7	2.44

Table 3.10 – Mission profiles.

m	Mission Description	ρ_m(\$)	U_m (hours/cycles)	ζ_m	$\mathcal{J}_m$	$\mathcal{R}_m$
1	Combat	80,000	2 / 10	2	$\{1, 2, 3, 4, 5\}$	$\{0.995, 0.990, 0.990, 0.950, 0.950\}$
2	Reconnaissance	60,000	5 / 25	1	$\{1, 2, 3\}$	$\{0.995, 0.990, 0.950\}$
3	Logistics Support	40,000	8 / 40	1	$\{1, 2, 3\}$	$\{0.995, 0.990, 0.900\}$

difference between the average prediction and the target. The mathematical expression corresponding to each of the above valuation metric is:

$$\text{RMSE} = \sqrt{\frac{\sum_{\lambda=1}^{\Lambda} (d_\lambda)^2}{\Lambda}}. \tag{3.67}$$

$$\text{SC} = \sum_{\lambda=1}^{\Lambda} s_\lambda, \quad \text{where} \quad s_\lambda = \begin{cases} \exp\left(\frac{-d_\lambda}{13}\right) - 1 & \text{if } d_\lambda \leq 0, \\ \exp\left(\frac{d_\lambda}{10}\right) - 1 & \text{if } d_\lambda \geq 0. \end{cases} \tag{3.68}$$

$$\text{AC} = \frac{100}{\Lambda} \sum_{\lambda=1}^{\Lambda} a_\lambda, \quad \text{where} \quad a_\lambda = \begin{cases} 1 & \text{if } d_\lambda \in [-13, 10], \\ 0 & \text{if } d_\lambda \notin [-13, 10]. \end{cases} \tag{3.69}$$

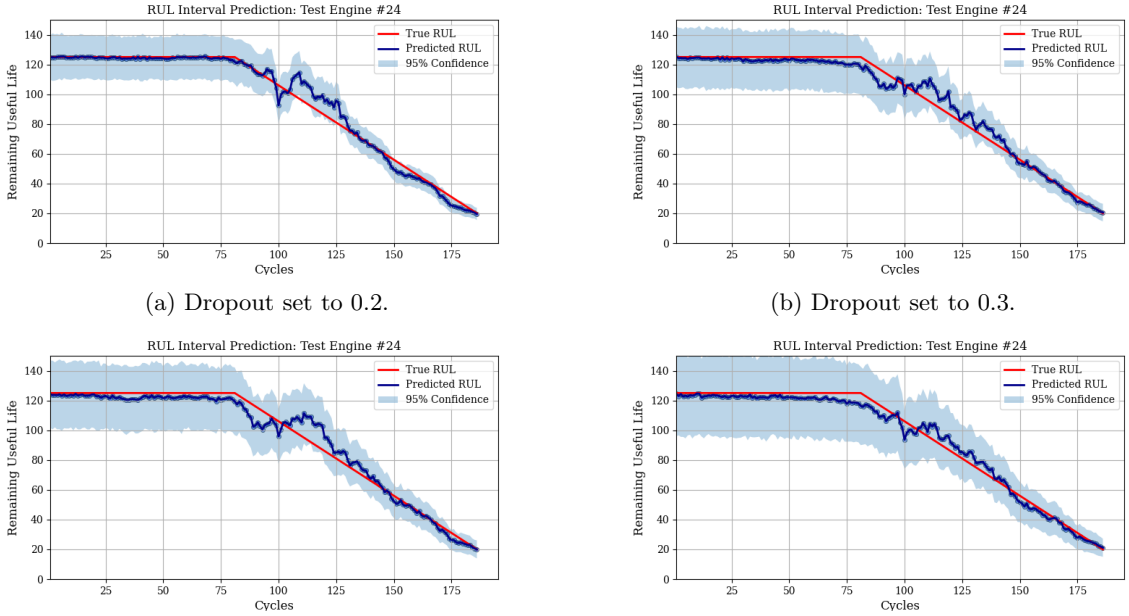
The same early stopping, learning rate decay strategy, and piecewise target function proposed in O'Neil et al. (2024) are also implemented. Furthermore, the impact of the dropout rate on prediction accuracy and the predicted RUL distribution is analyzed. For varying dropout rates, Table 3.11 displays the results obtained across multiple point prediction metrics. Interestingly, models with different dropout rates demonstrate comparable performance across the different metrics considered. However, a dropout rate of 0.2 results in the worst performance across the three metrics considered for the full dataset, likely due to overfitting to the training data. For all dropout values, it can be seen that the accuracy of the models improves significantly as the number of remaining operating cycles approaches 0. This is an important observation which indicates that the models become more accurate as the engine approaches failure. All models achieve an RMSE below 5 when considering monitoring points of 25 remaining cycles, in contrast to a significantly higher RMSE of approximately 18 when considering monitoring points of 75 remaining cycles.

MCD is utilized in this work to estimate the uncertainty of RUL predictions. By performing multiple stochastic forward passes through the model during inference, MCD generates a distribution of predictions. Figure 3.11 displays the mean of the MCD predictions (blue line) and 95% confidence interval (shaded region) for test engine 24, and for varying dropout rates. As one may observe, a dropout rate of 0.2 produces narrower confidence intervals, indicating reduced uncertainty, whereas a dropout rate of 0.5 results in broader intervals. This trade-off highlights the impact of the dropout rate on uncertainty quantification, with higher dropout rates offering more robust confidence estimates. Figure 3.12 further emphasizes the impact of the dropout rate on the width of the predicted RUL distribution. This figure displays the kernel distribution for test engine 90 considering multiple monitoring points for varying dropout values. Since MCD is used to compute the probability of successfully completing the upcoming mission, higher dropout rates may lead to unnecessary maintenance tasks due to the wider predicted RUL distribution. In contrast, lower dropout rates may increase the risk of component failure as the predicted RUL distribution becomes narrow.

Table 3.11 – Impact of dropout rate on prediction accuracy.

	Full dataset			Monitoring point with 75 cycles left		Monitoring point with 50 cycles left		Monitoring point with 25 cycles left	
Dropout	RMSE	SC	AC	RMSE	AC	RMSE	AC	RMSE	AC
0.2	13.10	310.24	68.00	20.14	53.49	6.12	87.89	3.46	100.00
0.3	11.71	246.15	76.00	18.21	58.14	7.00	90.91	3.85	100.00
0.4	11.56	278.07	69.00	18.76	53.49	7.27	90.91	3.46	100.00
0.5	11.74	254.53	72.00	18.47	53.49	6.70	87.87	4.23	100.00

To clearly demonstrate the impact of the dropout rate on maintenance decisions, Figure 3.13 displays the number of early replacements and system failures for the engines in the CMAPSS FD001 test set. These metrics are computed as follows:



(a) Dropout set to 0.2.

(b) Dropout set to 0.3.

(c) Dropout set to 0.4.

(d) Dropout set to 0.5.

Figure 3.11 – RUL interval estimate for test engine 24 for varying dropout rates.

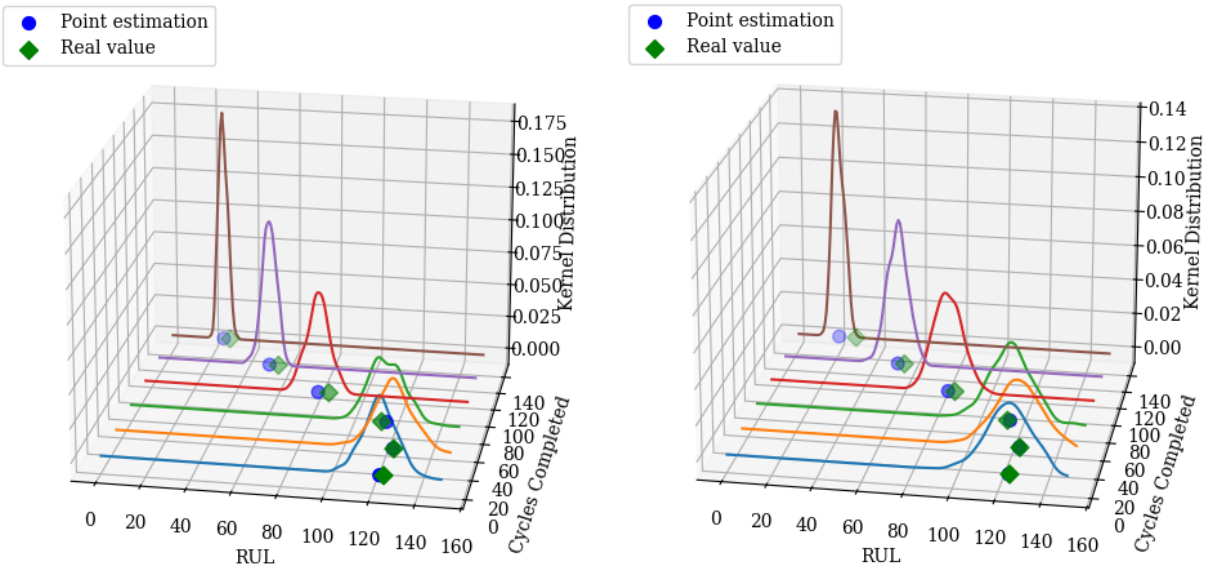
— If the predicted engine reliability is lower than the target reliability, but the engine survives the upcoming mission of duration U (*i.e.,* the true RUL is greater than U), this is counted as an early replacement.
— If the predicted engine reliability is higher than the target reliability, but the engine fails during the upcoming mission of duration U (*i.e.,* the true RUL is less than U), this is counted as a failure of the engine.

The above metrics are computed for all cycles completed across all 100 test engines and the results obtained are plotted in Figure 3.13. These results highlight the trade-off between early replacements and system failures, as well as the improved prediction performance for shorter mission durations. Generally, higher dropout values result in fewer system failures, but an increase in early replacements. In the case of exceptionally high reliability targets (99%), dropout values above 0.2 completely eliminate system failures. As expected, increasing the reliability requirement leads to fewer system failures but also more early replacements. Therefore, decision-makers have to choose an appropriate balance based on risk tolerance and cost constraints. Selecting high dropout values and high reliability requirements will result in an aggressive plan that likely leads to very few failures with high maintenance costs. In contrast, riskier maintenance strategies (*i.e.,* low dropout rate and/or low reliability thresholds) will likely result in reduced maintenance costs but increased mission failure probability. For the remaining numerical experiments, a dropout rate of 0.3 is selected as it provides the fewest early replacements while ensuring no system failures.

Impact of Stochastic Parameters

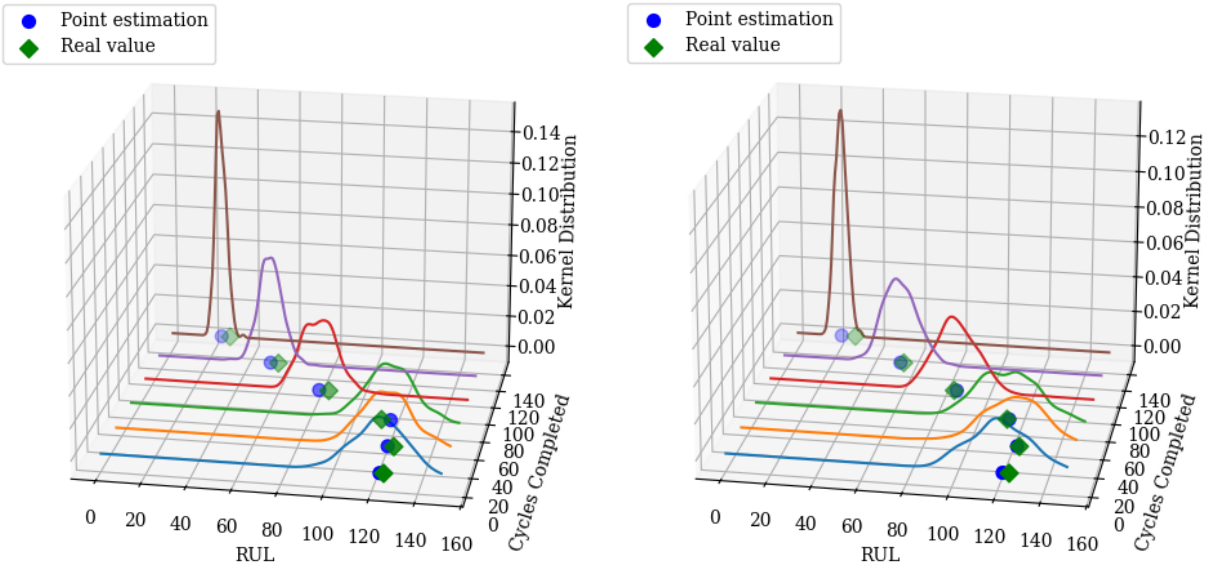
This set of experiments aims to demonstrate the importance of considering the stochastic nature of maintenance and break durations. In these experiments, the break duration is modeled as a truncated normal distribution with lower and upper bounds of 50 and 350 hours, respectively, a standard deviation of 50 hours, and a mean of 200 hours. There are $Q = 4$ available repairpersons to carry out the selected maintenance levels, with the respective fixed and variable costs of $c^f = \$2500$ and $c^v = \$50$. Finally, the required probability of completing the selected maintenance actions is set to $1 - \alpha = 90\%$.

To illustrate the impact of the mean duration of the break on maintenance and assignment decisions, three mean break durations are considered: $\mu = 150$, $\mu = 200$, and $\mu = 250$ hours. The probability density



(a) Dropout set to 0.2.

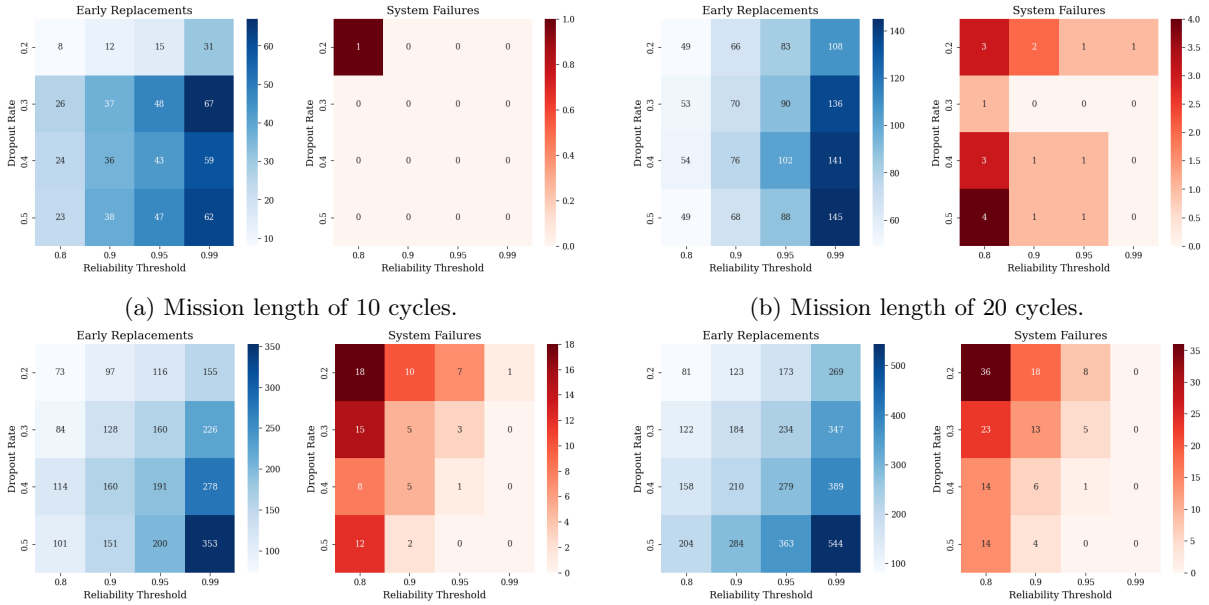
(b) Dropout set to 0.3.



(c) Dropout set to 0.4.

(d) Dropout set to 0.5.

Figure 3.12 – RUL distribution plots for test engine 90 for varying dropout rates.



(a) Mission length of 10 cycles.

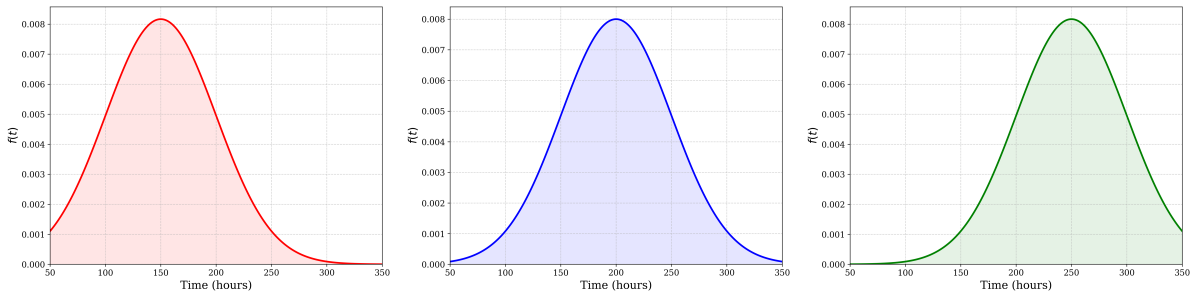
(b) Mission length of 20 cycles.

(c) Mission length of 30 cycles.

(d) Mission length of 40 cycles.

Figure 3.13 – Impact of reliability threshold and dropout rate on early replacements and system failures for different mission lengths.

functions (*pdf*) of the three break durations are shown in Figure 3.14. The number of realizations of the random maintenance and break durations generated is set to $N = 500$, and the optimization model is solved 10 times for each instance to account for variability in the random variable realizations. For each value of the break mean duration, Figure 3.15 displays the mean and standard deviation of the penalty cost, variable cost, and fixed cost. The results obtained show that increasing the mean break duration leads to a reduction in penalty cost, as more missions can be initiated. Indeed, a longer break duration allows for additional maintenance tasks while ensuring that the constraint satisfaction probability is met. The observed increase in average variable cost results from additional maintenance required to ensure the readiness of more aircraft, enabling the initiation of more missions. Figure 3.16 displays the true constraint satisfaction probability for the four available repairpersons. This probability is calculated by generating a large number of realizations for the durations of both selected maintenance actions and the break. The constraint satisfaction probability is then computed as the ratio of the number of scenarios with no constraint violations to the total number of realizations generated. It can be seen that for all mean break durations, no repairperson violates the required constraint satisfaction probability of $1 - \alpha = 0.9$.



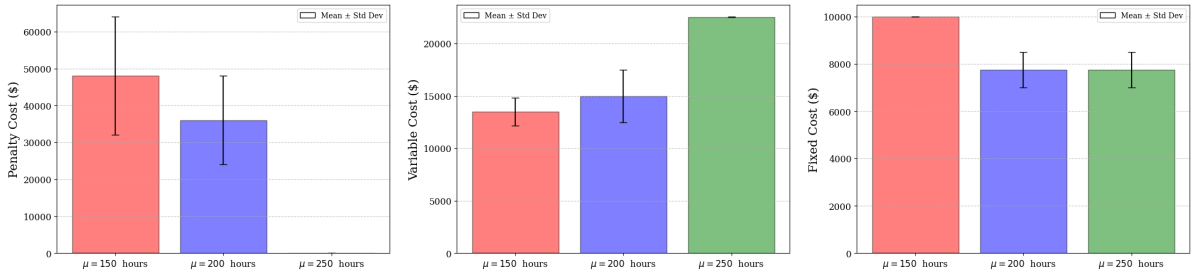
(a) $\mu = 150$.

(b) $\mu = 200$.

(c) $\mu = 250$.

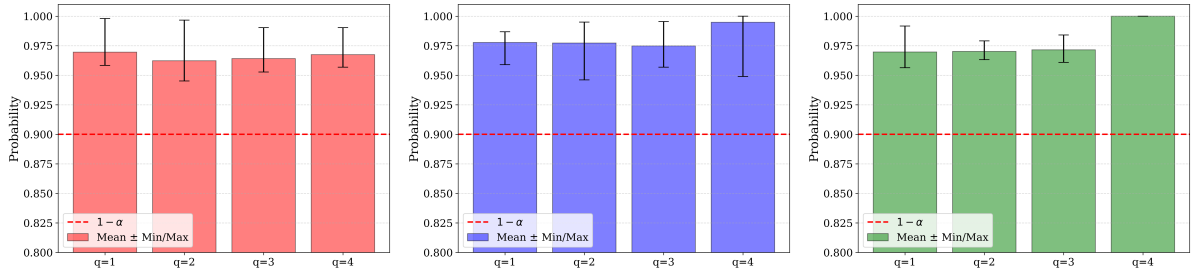
Figure 3.14 – Probability density functions of the three break durations considered.

To highlight the importance of accounting for the stochastic nature of maintenance and break dura-



(a) Penalty cost. (b) Variable cost. (c) Fixed cost.

Figure 3.15 – Mean and standard deviation of penalty cost, variable cost, and fixed cost for varying values of mean break duration.



(a) $\mu = 150$ (b) $\mu = 200$ (c) $\mu = 250$

Figure 3.16 – Constraint satisfaction probability for varying values of mean break duration.

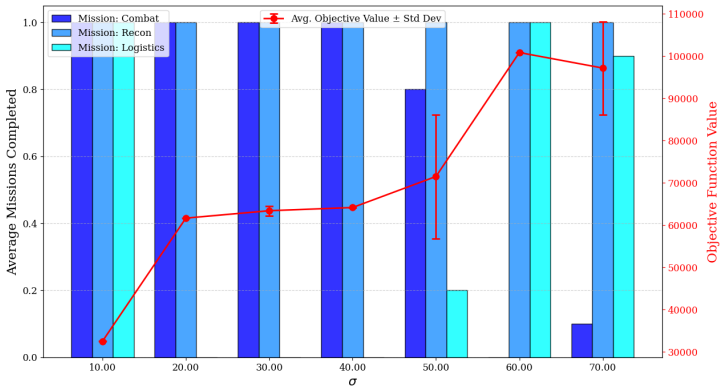
tions, the solutions obtained using only the average values for these durations are reported in Table 3.12 for different mean break durations. This table gives the penalty, variable cost, fixed cost, the number of missions initiated (M^*), the number of repairpersons hired and utilized (Q^*), the probability of true constraint satisfaction for multiple repairpersons, the average overtime and the computational time. The average overtime represents the expected additional time required to complete the selected maintenance tasks. Considering only average values results in overly optimistic solutions that have low probabilities of being completed within the stochastic break duration. In some cases, the probability of completing the assigned maintenance tasks is around or even below 50%. It is also important to highlight the significant increase in average overtime when considering only the mean values of maintenance and break durations. In the context of a military aircraft fleet during wartime, requiring substantial additional time to complete maintenance tasks could delay critical operations and put lives at risk.

Table 3.12 – Comparison of Deterministic and Stochastic Solutions. Mean and standard deviation (in brackets) reported for CVaR approach.

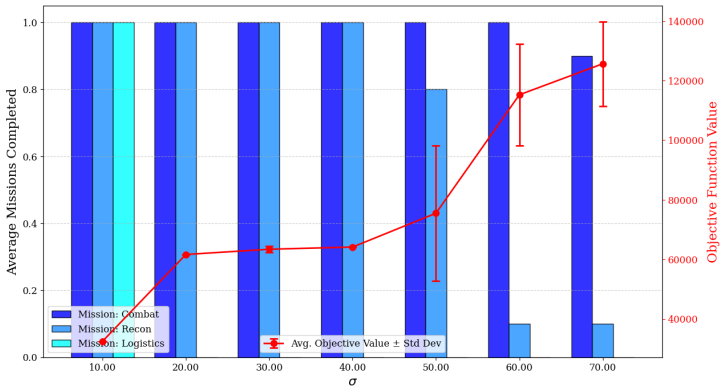
		Penalty ($\$ \times 10^3$)	Var. Cost ($\$ \times 10^3$)	Fixed Cost ($\$ \times 10^3$)	M^*	Q^*	Constraint Satisfaction Probability (%) $q=1$	$q=2$	$q=3$	$q=4$	Avg. Overtime (hr)	CPU Time (s)
$\mu = 150$	Determ.	0	22.5	7.5	3.0	3.0	50.7	53.3	49.6	–	19.79	0.1
	CVaR	48(16)	13.5(1.3)	10.0(0.0)	2.0(0.0)	4.0(0.0)	96.9(1.1)	96.2(1.7)	96.3(1.0)	96.7(1.2)	0.55(0.16)	19.7(34.7)
$\mu = 200$	Determ.	0	22.5	7.5	3.0	3.0	60.0	79.0	96.5	–	15.17	0.2
	CVaR	36(12)	14.9(2.5)	7.8(0.8)	2.1(0.3)	3.1(0.3)	97.8(0.9)	97.7(1.4)	97.5(1.0)	99.5(1.5)	0.62(0.23)	107.4(129.6)
$\mu = 250$	Determ.	0	22.5	7.5	3.0	2.0	56.3	76.7	–	–	17.68	0.1
	CVaR	0(0)	22.5(0)	7.8(0.8)	3.0(0.0)	3.0(0.0)	97.0(1.0)	97.0(0.6)	97.1(0.8)	–	0.76(0.14)	13.9(2.3)

Finally, the impact of the standard deviation of the break duration on maintenance plan, system-to-mission assignment, and the objective function value is studied using a truncated normal distribution with a mean of 150 hours. The standard deviation is varied from 10 to 70 hours and the results are displayed in Figure 3.17. This figure illustrates how the probability of selecting each mission type changes over the 10 optimization runs as the standard deviation of the break increases. The results are reported for

two values of ρ_1, the penalty assigned to the combat mission. The bar chart represents the probability of initiating combat, reconnaissance, and logistics missions, while the line plot shows the corresponding objective function value. Generally, a decreasing trend in the objective function value can be seen as the break duration standard deviation increases. A higher standard deviation of the break duration limits the number of maintenance actions each repairperson can perform while ensuring the desired constraint satisfaction probability. Consequently, more repairpersons must be hired, increasing fixed costs, or fewer mission types can be selected. Another interesting observation is the shift in mission type selection as the standard deviation increases. For $\rho_1 = \$80,000$, the probability of selecting the combat mission decreases as the standard deviation increases. Although the combat mission carries the highest penalty, it requires two aircraft for successful completion and also requires the functionality of the weapons subsystem. As uncertainty increases, the model favors selecting the reconnaissance and logistics support missions, whose combined penalty exceeds that of the combat mission alone while requiring the same number of operational systems. However, when $\rho_1 = \$120,000$, the model tends to select the combat mission, as its penalty surpasses the combined penalty of the other two mission types.



(a) $\rho_1 = \$80,000$



(b) $\rho_1 = \$120,000$

Figure 3.17 – Impact of standard deviation of break duration on mission selection probability and objective function value.

Impact of Constraint Satisfaction Probability $1 - \alpha$

This final experiment aims to explore the impact of the constraint satisfaction probability $1 - \alpha$. To do so, a truncated normal distribution with mean 200 hours, standard deviation of 50 hours, and lower and upper bounds of 50 and 350 hours is once again considered. In this set of experiments, $Q = 8$ repairpersons are available to carry out the selected maintenance actions, while all other input data remain unchanged. The problem is solved for varying values of α and the results are displayed in Figure 3.18. For all values of α, all mission types are selected; however, more repairpersons are required to

perform the necessary maintenance tasks for smaller values of α. This is because smaller values of α imply a higher required probability of constraint satisfaction, meaning fewer maintenance tasks can be assigned to each repairperson. Additionally, as the constraint satisfaction probability is relaxed (*i.e.*, α increases), expected overtime increases. This represents an important trade-off that decision-makers must consider. To further illustrate the impact of α on maintenance assignment, Figure 3.19 shows the average maintenance time per repairperson, along with the minimum and maximum values, for varying α. It is evident that smaller values of α lead to lower average maintenance times per repairperson.

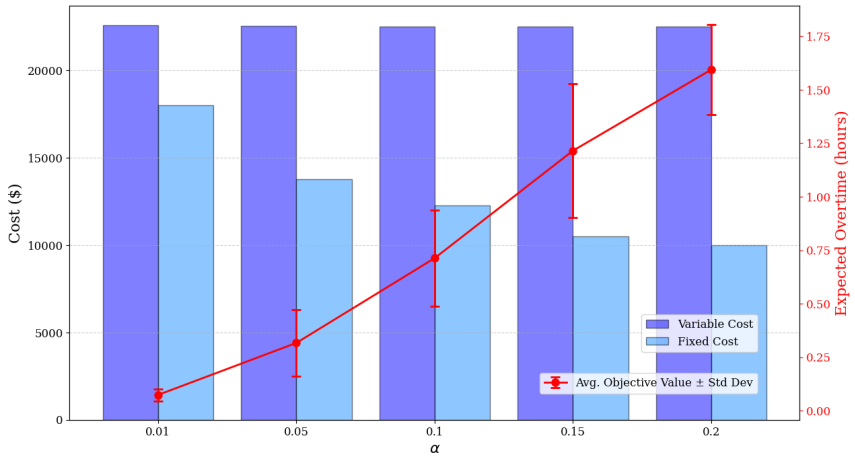


Figure 3.18 – Average variable cost, fixed cost, and overtime for varying values of α.

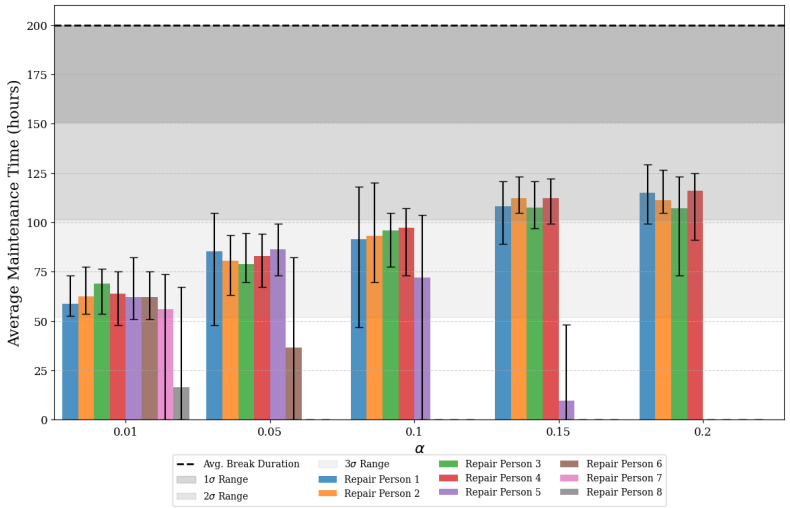


Figure 3.19 – Average maintenance durations for varying values of α.

3.2.6 Conclusions and Future Research Directions

This section presented a novel mathematical model to jointly optimize fleet maintenance planning, repairperson scheduling, and system-to-mission assignment to minimize the total expected cost incurred. The fleet consists of identical multi-component systems, which undergo maintenance during the break duration to ensure readiness for a set of unique upcoming missions. The resulting decision problem involves: (i) selecting the components within each system to maintain and their corresponding maintenance levels, ensuring reliability requirements are met, (ii) assigning maintenance tasks to repairpersons, and (iii) allocating systems to mission types. The systems include both standard and sensor-monitored

components. Standard components follow known lifetime distributions, allowing reliability values to be computed analytically. Sensor-monitored components are continuously tracked by sensors, and a data-driven approach leveraging a DNN and MCD was used for accurate reliability prediction. To address the complex optimization problem while accounting for uncertainties in maintenance and break durations, a chance-constrained optimization model was proposed. The problem was then reformulated as a large-scale mixed-integer linear program (MILP) using the conditional value-at-risk (CVaR) approximation. Extensive numerical experiments validated the proposed joint model and highlighted the importance of accounting for the inherent uncertainties associated with maintenance and break durations.

This research opens numerous promising avenues for future investigation. The proposed chance-constrained optimization model tends to result in conservative maintenance plans, where repairpersons are likely to complete their assigned maintenance actions well within the break duration. This is particularly true when a high constraint satisfaction probability is required. While this ensures feasibility, it may lead to the hiring of more repairpersons than necessary. To address this, it would be valuable to explore two-stage stochastic optimization models, or other sequential decision methods. However, developing a two-stage optimization model would require the development of efficient solution methods. Finally, the application of various data-driven methods for RUL prediction and uncertainty quantification is of great interest. The authors are also exploring domain adaptability, where an RUL predictive model trained in one domain is applied to a system in a different domain using techniques such as transfer learning. This is a crucial research area, as it is not common in many organizations to have a large amount of labeled run-to-failure data.

Chapter 4

Optimizing Resource-Constrained Fleet Selective Maintenance With Asynchronous Maintenance Breaks

4.1 Introduction

Maintenance programs have helped to minimize downtime and energy consumption, enhance system dependability and availability, and extend asset RUL in industrial and domestic settings. Technological progress has led to increased complexity in engineered systems through the combination of numerous components and subsystems (An et al., 2021). Examples of these complex assets and systems vary from single systems such as automobiles, trucks, unmanned drones and urban train transit vehicles to expansive systems such as wind turbines, navy ships and petrochemical plants. These repairable multi-component systems often need to meet operational targets over consecutive mission cycles. In order to achieved the required operational targets, it is essential to conduct PM and/or CM during SMB intervals between successive missions to ensure high overall equipment efficiency, increased system safety, and reduced energy consumption (Liu et al., 2023; Khatab et al., 2018a).

According to existing surveys, effective maintenance planning can lead to savings of 20-30% in total operational expenses (Syan and Ramsoobag, 2019; An et al., 2023). Reliability Centered Maintenance (RCM) (Geisbush and Ariaratnam, 2023), as a maintenance approach, was initially successfully implemented in the airline industry (Guo et al., 2016). Since then, its applicability has been extended to cover other industrial sectors (Tripathi and Hari Prasad, 2024; Braaksma et al., 2013). A recent literature review on RCM current industrial practices is available in Geisbush and Ariaratnam (2023). However, carrying out perfect and high-quality PM and CM maintenance operations for every component during the intermission maintenance break is often unattainable owing to restrictions on maintenance resources such as working hours, budget, equipment, and personnel. Given these constraints, it is imperative for stakeholders to determine which components to maintain, the extent/degree of PM and CM maintenance activities for those chosen components, and the assignment of resources. Rice et al. (1998) coined the term SMP to describe this optimization issue.

Although the SMP has been significantly researched (Al-Jabouri et al., 2022; Liu et al., 2023), most of the research primarily focuses on single systems, with maintenance considerations being at the component level of these systems. Yet, numerous goods production and service sectors often operate with multiple identical systems in a fleet. To illustrate, in the transport sector, fleets consist of a variety of vehicles such as trains, buses, and airplanes. This means that all systems in a fleet share the available maintenance resources, necessitating SM decisions for every component across the entire fleet. Given the often-limited nature of these resources, achieving optimal performance for every system in the fleet for upcoming missions can be challenging. Therefore, the objective of FSM is to identify specific maintenance tasks to be performed on selected components within each system of the fleet. The primary aim is often

to increase the likelihood that every system in the fleet will successfully complete the upcoming mission (Khatab et al., 2022). This particular SMP variant is called the FSMP, a concept initially presented by Cassady et al. (2003).

Existing FSMP models rely on the restrictive and unrealistic assumption in which all systems in the fleet operate synchronously, meaning that maintenance tasks can only be decided and executed only when all these systems have completed the current mission (Al-Jabouri et al., 2024b) and are all available simultaneously. No paper dealing with the FSMP addresses the general case of asynchronous maintenance breaks. Our primary goal is to address this asynchronous break gap by proposing a novel generalized resource-constrained model of the SMP. In this context, the proposed model seeks to optimize maintenance activities within these fleet systems, ensuring efficient resource allocation without compromising reliability and operational continuity. By introducing the FSMP with asynchronous breaks, this chapter effectively bridges the gap between the traditional FSMP and the RCPSP. The secondary goal is to optimize the trade-offs between maintenance scheduling, resource availability and achieved mission reliability. The third goal is to derive managerial insights for maintenance schedulers dealing with efficient maintenance planning for fleets of complex multicomponent systems under resource constraints.

This chapter is a major extension of the preliminary work published in Patel et al. (2023). Compared to Patel et al. (2023), the problem is significantly extended to include a fleet of MOS, each system is allowed to undergo maintenance according to its own break window, and multiple resource types are considered with their own availability limitations. Such resource types could be repair persons, tools or spare parts. We use the following outline. In Section 4.2, the relevant literature is reviewed. Section 4.3 describes the system under consideration and underlying assumptions, while Section 4.4 develops a novel mathematical formulation of the FSMP and a corresponding optimal solution method is proposed. In Section 4.5, different experiments are performed and their results analyzed and discussed. Section 4.6 concludes the chapter, where findings are summarized and future research directions presented.

4.2 Literature Review

As the resource-constrained FSMP combines the SMP with concepts of the RCPSP, this section will present a literature review of both areas.

4.2.1 The SMP

The SMP has seen significant research growth since it inception in Rice et al. (1998). Over the last two decades, a sizable body of literature has extended the SMP in numerous avenues. For example, Maillart et al. (2009) studied the SMP over extended horizons, revealing minimal differences between single and multi-mission policies, thus directing focus to single-mission problem extensions. Liu and Huang (2010) integrated IM into SMP, specifically for multi-state systems (MSS), employing the Kijima IM models for modeling and plGA for optimization . Pandey et al. (2013) used the hybrid hazard rate approach to model the IM for the SMP. Their paper was pivotal in linking resource allocation with maintenance levels. Chen et al. (2012) developed an SMP optimization model with load-sharing in multi-state systems and solved it with a GA. Dao and Zuo (2015) explored the SMP for series MSS with stochastic dependent components, developing a profit maximization model. Zhao and Zeng (2016) introduced a stochastic model for two-state parallel systems employing a hybrid intelligent optimization algorithm. This represented a significant shift towards acknowledging the uncertainty in maintenance and mission duration. Pandey et al. (2016) proposed a PM model with a finite planning horizon using a hybrid IM model. Khatab and Aghezzaf (2016) addressed the optimization of SMP in multi-component systems with stochastic maintenance action quality reflecting the real-world variability in maintenance outcomes. Das Adhikary et al. (2016) presented a multi-objective model for continuously operating series systems aiming to optimize availability and maintenance cost using a GA. Khatab et al. (2017a) extended the SMP to accommodate stochastic mission and break durations by employing a non-linear stochastic program. This study underscored the importance of considering stochastic elements in maintenance planning. Diallo et al. (2017, 2019) introduced a novel variant of SMP integrating repair task assignment to multiple repair persons leading to a significant step towards addressing the complexity of real-world SMP scenarios. O'Neil et al.

(2023) introduced a joint maintenance and orienteering framework for offshore wind farms, proposing a linear programming optimization model and a column-generation based solution method. They applied CG to solve the SMP by starting with a limited set of variables and iteratively adding new columns to improve the objective function. This approach effectively minimizes maintenance and routing costs while meeting reliability requirements, handling the large-scale nature of the problem efficiently. Dao and Zuo (2017a) studied SM in series MSS under variable loading conditions, proposing a load-dependent degradation model and Monte-Carlo simulation for reliability evaluation. Liu et al. (2018) developed a new model for MSS under stochastic maintenance durations and optimized the sequence of maintenance actions using ant colony optimization.

Shahraki et al. (2020) presented an SM optimization problem for systems with stochastically dependent components. It introduced a novel approach capturing interactions between components and considered various maintenance actions including perfect and stochastic IM. Zhang et al. (2024a) addressed the SMP for MOS operating several missions under s-dependent competing risks, IM and limited repairpersons. Copula functions are used to capture the s-dependent competing risks between components, and their unknown parameters are estimated by a two-stage maximum likelihood estimation method. Ruiz et al. (2020) developed a novel SM framework for multi-component systems with multiple failure modes. This research distinguished between soft and hard failures and optimized system performance under IM. Jiang and Liu (2020b) proposed a multimission SM strategy with multiple optional maintenance actions and uncertainties related to operation time and future mission duration and Jiang and Liu (2020a) addressed the robust SMP under imperfect observations. Sun and Sun (2021) proposed an SM model for a transportation MSS emphasizing the sequencing of maintenance actions to maximize system reliability within the constraints of transportation volume requirements and a limited budget. Li et al. (2021) tackled SM for series MSS, considering maintenance quality uncertainty and failure effects. This study established an approach to measure failure effects under different states, providing insights into the impacts of maintenance quality uncertainty on maintenance decisions. Tambe (2022) presented a SM decision model for a multi-component system, utilizing the simulated annealing algorithm for optimization. The model minimize s the total cost of maintenance while meeting the required availability within a specified time window.

More recent SMP contributions have trended towards more sophisticated models and solution approaches, thus reflecting an acute awareness of the complexities and uncertainties inherent in modern industrial and military systems. Al-Jabouri et al. (2023a) developed a robust optimization SMP model for series–parallel systems under maintenance quality uncertainty. This study highlighted the importance of considering maintenance quality uncertainty in SM plans. Ma et al. (2023) proposed an integrated SM and task assignment model for MSS, formulating a max-min optimization model resolved via a cooperative GA. Yin et al. (2023) formulated a new SM model with multiple heterogeneous repair channels, considering uncertainties in maintenance tasks and break durations. They solved the optimization problem using a hybrid double-loop algorithm and ant colony optimization method. Al-Jabouri et al. (2023b) presented branch-and-price algorithms for large-scale mission-oriented maintenance planning problems, demonstrating computational efficiency and solution quality. Guo et al. (2023) developed an availability model for multi-component k-out-of-n:G warm-standby systems in series, considering the rule of suspended animation. Their approach effectively solved complex situations and verified model accuracy and algorithm efficiency. O'Neil et al. (2023) introduced a joint maintenance and orienteering framework for offshore wind farms, proposing a linear programming optimization model and a column-generation based solution method. Al-Jabouri et al. (2024a) studied distributionally-robust chance-constrained optimization of SM under uncertain repair duration with piecewise linear approximations. Finally, AI and ML methods have been recently used to model and solve the SMP (Achamrah and Attajer, 2024; Hesabi et al., 2022).

Through these studies, the SMP has been significantly advanced, covering a wide range of system features, optimization model characteristics, and solution approaches. The literature demonstrates a strong focus on enhancing system reliability, optimizing maintenance costs, and addressing the practical complexities of SM in varied industrial and military environments. However, these developments have been observed for single systems. The next subsection presents a literature review of references dealing

with the SMP for multiple systems (*i.e.,* FSMP).

4.2.2 The FSMP

The FSMP has received increased attention due to its vital role in assuring the operational readiness and efficiency of fleet systems in a variety of industries, including transportation and defense. In the FSMP, maintenance decisions need to be determined for every component within the fleet, introducing an additional layer of complexity to the optimization procedure. Given that under FSMP, multiple systems operate separate missions independently, stakeholders must define what constitutes the successful completion of the fleet's mission. In general, if stakeholders decide that k systems must successfully complete their missions out of the n systems in the fleet for the fleet's mission to be deemed successful, then the fleet reliability structure can be defined as a k-out-of-n configuration. For example, if stakeholders define the fleet mission success as the mission completion of all systems, then $k = n$ and we have a n-out-of-n or series configuration (O'Neil et al., 2024; Chen et al., 2023b). Conversely, if the stakeholders decide that the success of one system is sufficient to deem the fleet's mission a success, then $k = 1$ and the fleet's reliability structure is a 1-out-of-n or parallel configuration. Cassady et al. (2003) were pioneers in integrating maintenance planning with sortie scheduling, formulating an optimization model for FSM under dynamic mission profiles. This work laid a foundation for subsequent research that sought to balance mission requirements with limited maintenance resources. Schneider and Cassady (2004) studied the reliability maximization of fleets, employing a simulation-optimization model. Their approach revealed the complex interplay between maintenance activities and fleet performance over sequential mission sets, thus highlighting the significance of strategic maintenance planning. Meanwhile Cassady et al. (2005) broadened the scope to multi-state SM decisions addressing scenarios with multiple system states and varied missions. Schneider and Cassady (2015) extended the original FSM model to include cost-based optimization , offering insights into managing future missions with potential cancellations. This model provided a more nuanced understanding of the trade-offs involved in maintenance decisions under varying mission scenarios. Chen et al. (2023a) introduced a two-stage fleet maintenance optimization method, focusing on balancing mission requirements with maintenance cost and time, thus offering a practical approach to maintenance scheduling in fleet systems.

Lan et al. (2017) presented an FSM model for highway transportation, considering stochastic repair quality. Their work brought to light the importance of maintenance time windows in ensuring mission reliability and cost-effectiveness. Feng et al. (2017) proposed a heuristic hybrid game approach for fleet condition-based maintenance (CBM) planning, emphasizing the dynamic nature of fleet maintenance decision-making in military scenarios. Yang et al. (2018) tackled the FSMP under phased missions and short breaks, introducing a heuristic game framework. This approach underscored the need for frequency-based maintenance optimization in scenarios with rapid mission turnovers. Khatab et al. (2022) optimized the integrated imperfect FSMP, highlighting the complexity added by multiple repair channels and the variable skill sets of repairpersons. Zhang et al. (2023b) developed a heuristic framework for scheduling the maintenance of military aircraft fleets, addressing the challenges posed by limited maintenance capacities and the unpredictability of mission schedules. Sun et al. (2023) proposed a comprehensive life-cycle simulation framework for aero engine fleet management, offering an effective decision support tool that accounts for varying environmental and operating conditions. Sieb et al. (2023) identified challenges in maintenance planning for on-demand Urban Air Mobility fleets. Their agent-based simulations shed light on the interplay between operation and maintenance in novel aircraft architectures, suggesting a combined approach to optimize maintenance costs. Tseremoglou and Santos (2024) addressed CBM scheduling for aircraft fleets under partial observability using a deep reinforcement learning approach. Their model, which efficiently integrated health condition predictions and stochastic task arrivals, demonstrated significant maintenance cost reductions. Lastly, Chen et al. (2023b) contributed to the joint optimization of sequential FSMP and repair crew assignment for multi-state manufacturing systems. They used a tailored GA to solved the resulting optimization model. Their work emphasized the need to consider uncertain maintenance duration and flow dependency. Despite the extensive research carried out on FSMP, there remains a critical aspect that has been largely ignored: the exploration of asynchronous break scenarios under limited resource conditions, which more accurately represent real-world fleet operations.

4.2.3 Resource Constraints in Maintenance Scheduling

The RCPSP, a subset of scheduling issues, primarily addresses scenarios in which a workforce is tasked with completing activities within a planning horizon while adhering to specific start times and due dates (Cavalcante et al., 2013). The disaggregated discrete-time formulation for the RCPSP proposed by Christofides et al. (1987) has proven to be efficient in solving problems of moderate size and is widely used in project management. The RCPSP finds extensive applications across various domains, notably in project management, production planning, scheduling, and maintenance management (Pellerin et al., 2020). The planning of maintenance for complex technical systems often evolves into a challenge of multi-objective optimization (MOO), as highlighted by Zio (2009). This MOO approach is prevalently employed to address design and organizational trade offs in many engineering fields (Coit and Zio, 2019). Keizer et al. (2017) have described the concept of resource dependence, which arises when various components like spare parts, tools, or personnel are collectively needed for operational tasks. Furthermore, elements such as repair tools, spare parts (Nguyen et al., 2017), and budget availability (Mild and Salo, 2009) have been identified as typical examples of resource dependencies in maintenance scenarios (De Jonge and Scarf, 2020).

Awad and Ertem (2017) tackled the challenge of scheduling PM activities in petrochemical plants using stochastic modeling to minimize tardiness. The model considered maintenance crew skill-set constraints, a specific time horizon, and budget limitations. The research emphasized the importance of both direct and indirect maintenance times, including disruptions like weather or site availability, in refining the scheduling process. Gu et al. (2022) optimized railway maintenance scheduling by determining optimal train arrival times at service facilities. A nonlinear model and a linear programming model were developed to address this complex problem, taking into account multiple variables such as maintenance duration, time since last service, total fleet size, and capacity constraints at maintenance centers. Jiu et al. (2013) investigated equipment maintenance scheduling considering finite intermediate storage, train transportation, and equipment-specific maintenance time preferences. They minimized a weighted sum of maintenance time aversion, changeover times, and train waiting time. A mixed integer linear programming model was developed and solved using a hybrid GA. Witteman et al. (2021) introduced a novel bin-packing approach for solving the aircraft maintenance task allocation problem, framing it as a time-constrained variable-sized bin-packing problem. This approach considered due dates and task repetitions, dividing the planning horizon into variable-sized bins for task allocation. Ertem et al. (2022) focused on workers-constrained shutdown scheduling, considering skill flexibility. They developed two mixed-integer models with solution algorithms, allowing higher-skilled workers to perform tasks of lower skill levels. Their heuristics showed significant reductions in shutdown completion times and were within close deviations from optimally solved instances, highlighting the efficiency of flexible worker utilization. Urbani et al. (2023) developed a bi-objective maintenance scheduling model for networked systems with limited resources. Their model, representing the system as a directed graph, aimed at minimizing maintenance costs and maximizing productivity. The study also provided insights into the impact of adding maintenance personnel on system productivity and resource idle time. Fekri et al. (2023) explored PM order scheduling as a multi-skilled resource-constrained flow shop problem. They formulated a MILP model with dual objectives to minimize the completion time and idle time of human resources. The model's efficiency was demonstrated through a GA, emphasizing the necessity of considering human resources in maintenance scheduling. Wang et al. (2023b) investigated scheduling issues in marine equipment maintenance by proposing a preemptive multi-skill resource-constrained project scheduling model, employing an improved moth-flame optimization algorithm. Finally, Zhang et al. (2023c) introduced a model to schedule carrier-based aircraft maintenance tasks under time-critical and resource-constrained environments. Their model was solved by an improved non-dominated sorting genetic algorithm II (NSGA-II) to find effective strategies for adapting to dynamic uncertainties and enhance fleet availability.

As highlighted in the papers discussed above, multiple works have considered resource constraints when scheduling maintenance activities. Unfortunately, few studies on FSMP have investigated resource constraints and even fewer have considered multiple resource types.

4.2.4 Motivation and Contributions

As pointed out earlier, the main common shortcoming of current FSMP models is that they all assume that individual systems in the fleet are available and/or eligible for maintenance activities at the same time (*i.e.,* the break period is synchronous). In reality, this is seldom the case. For example, long-haul trucks reach maintenance depots at different times. Navy ships return from missions and access dry-docks in different work periods. Production lines and workstations undergo maintenance during different shifts or days. Thus, it is crucial to model the FSM under asynchronous break windows. Figure 4.1 shows the synchronous and asynchronous maintenance break cases for a fleet with three systems. For the synchronous case, all three SMB intervals start at the same time (*i.e.,* $ES_1 = ES_2 = ES_3$) and overlap completely. For the asynchronous case shown, the breaks start at three different times (*i.e.,* $ES_1 \neq ES_2 \neq ES_3$) and have 50% overlap between consecutive systems.

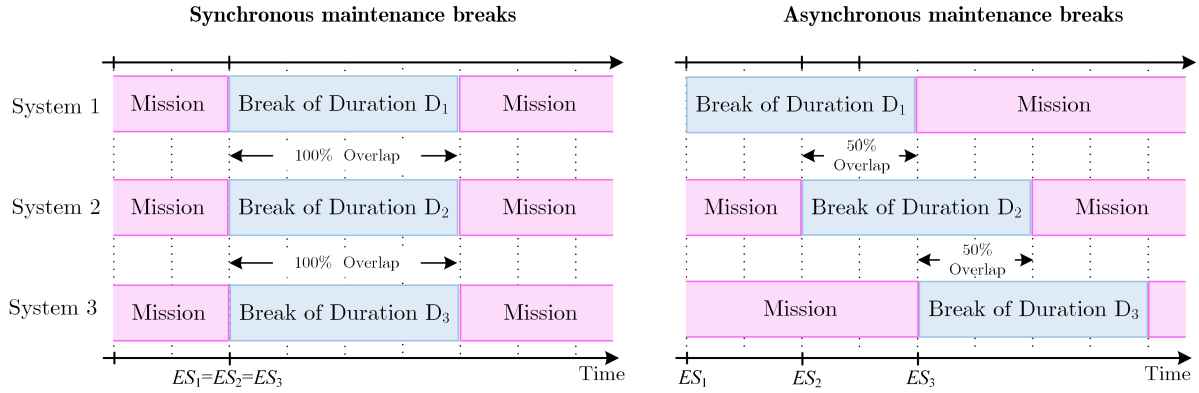


Figure 4.1 – Illustration of the synchronous and asynchronous maintenance break cases.

The major contributions of this chapter are:

— Developing a novel discrete-time formulation for the resource-constrained FSMP with asynchronous breaks, denoted as ARCFMSP.

— Using the proposed model to investigate the trade-offs between maintenance scheduling, resource availability and achieved mission reliability.

— Deriving novel managerial insights for maintenance schedulers dealing with maintenance planning for fleets of complex multicomponent systems under resource constraints.

4.3 System Description and Assumptions

4.3.1 Notation

I	Number of systems within the fleet
$\mathcal{I}$	Set of systems within the fleet indexed by i, $i \in \mathcal{I} = \{1, \ldots, I\}$
N_i	Number of subsystems within system i
$\mathcal{N}_i$	Set of subsystems in system i indexed by j, $j \in \mathcal{N}_i = \{1, \ldots, N_i\}$
M_{ij}	Number of components in subsystem j of system i
$\mathcal{M}_{ij}$	Set of parts in subsystem j of system i indexed by k, $k \in \mathcal{M}_{ij} = \{1, \ldots, M_{ij}\}$
E_{ijk}	The kth component/part in subsystem j of system i
B_{ijk}	Age of component E_{ijk} at the start of the SMB of system i
S_{ijk}	Status of E_{ijk} at the start of the SMB of system i
L_{ijk}	Number of maintenance levels available for component E_{ijk}
$\mathcal{L}_{ijk}$	Set of maintenance levels available for E_{ijk} indexed by $l \in \mathcal{L}_{ijk} = \{1, \ldots, L_{ijk}\}$
ES_i	Earliest start time for maintenance operations on system i

D_i	Duration of the SMB for system i in break time units (b.t.u.), *e.g.,* hours or days
$\mathcal{D}_i$	Set of discretised time intervals in the SMB of system i indexed by $\tau \in \mathcal{D}_i = \{ES_i, \ldots, ES_i + D_i\}$
Q	Number of maintenance resource types
$\mathcal{Q}$	Set of maintenance resource types indexed by q, $q \in \mathcal{Q} = \{1, \ldots, Q\}$
C_0	Total maintenance budget available in monetary units (m.u.), *e.g.,* \$, ¥, and £
Γ_q	Total number of resources of type q available
r_{ijklq}	Number of resources of type q needed to perform maintenance level l on E_{ijk}
c_q^f	Fixed cost incurred per resource of type q used
c_q^v	Variable cost per unit of time incurred per resource of type q used
W	Mission length in life time units (l.t.u.), *e.g.,* months or years
R_0	Minimum required overall system reliability level
R_{ijkl}	Reliability of component E_{ijk} under maintenance level l
R_{ij}	Reliability of subsystem j in system i
R_i	Reliability of system i
R_F	Reliability of the fleet of systems
t_{ijkl}^p	Time to carry out PM level l on component E_{ijk}
t_{ijkl}^c	Time to carry out CM level l on component E_{ijk}
ΔT	Unit time step/period used to discretise the planning horizon

4.3.2 Assumptions

The following working assumptions are made in the formulation of the resource-constrained FSMP .

— During breaks, system components do not degrade or age (*i.e.,* the age of a component is operation-dependent). This is a reasonable assumption for most systems, whose break durations are much shorter than their mission lengths (Al-Jabouri et al., 2023b; Diallo et al., 2019; O'Neil et al., 2024).

— The time horizon is discretized into unit time periods of length ΔT. In most maintenance operations, there is a minimal maintenance duration used for accounting such as 15-minute blocks for normal activities or even daily blocks for long refit/overhaul projects. The duration of all other maintenance activities are multiples of this unit time period ΔT.

— By definition of the SMP, maintenance activities occur only within the break windows (Cassady et al., 2003; Rice et al., 1998; Diallo et al., 2018).

4.3.3 Fleet Description

The fleet under consideration consists of I systems, where each system follows a general series-parallel configuration. The mission of the fleet is a success if and only if all I systems survive the mission. Thus, the fleet reliability structure is a I-out-of-I system (*i.e.,* series configuration). System i consists of N_i subsystems in series, where each subsystem is comprised of M_{ij} components arranged in parallel. Maintenance activities are time-sensitive, with ES_i representing the earliest possible start time for maintenance on system i. When system i from the fleet enters its corresponding SMB of duration D_i, the age B_{ijk} and the state S_{ijk} of its components E_{ijk} are known, and the state parameter S_{ijk} is defined as:

$$S_{ijk} = \begin{cases} 1, & \text{if component } E_{ijk} \text{ is working at the start of the SMB of system } i \\ 0, & \text{otherwise} \end{cases} \tag{4.1}$$

For system i, maintenance tasks must be carried out during the corresponding allocated SMB window of length D_i. At the fleet level, the maintenance tasks are performed during a planning horizon of length T computed as:

$$T = \max_{i \in \mathcal{I}} \{ES_i + D_i\} - \min_{i \in \mathcal{I}} \{ES_i\}. \tag{4.2}$$

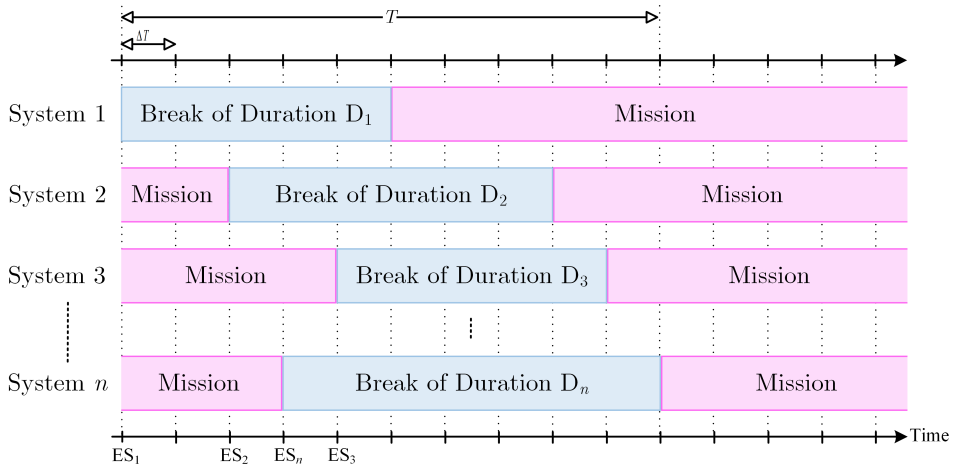


Figure 4.2 – Alternating sequence of missions and scheduled breaks for the fleet of n systems.

The planning horizon T is segmented into a finite number of equal time periods of length ΔT (Figure 4.2), thus defining a set of discrete time periods $\mathcal{T} = \{0, \Delta T, 2\Delta T, \ldots, T\}$.

4.3.4 Maintenance Modeling and Reliability Computation

For each component E_{ijk}, a set $\mathcal{L}_{ijk} = \{1, \cdots, L_{ijk}\}$ of maintenance levels is available. A maintenance level $l \in \mathcal{L}_{ijk}$ refers to a maintenance action such as DN, mR, IM, and replacement (R). The DN option ($l = 1$), if chosen, means that no maintenance is performed on the component. The mR option ($l = 2$) restores a failed component back to a functioning state without altering its age just before failure. The Replacement option ($l = L_{ijk}$) completely renews the component by resetting its age to zero. An IM level, $2 < l < L_{ijk}$ improves the component's age to a point between the outcomes of mR and replacement. Following the age reduction approach for IM modeling suggested by Malik (1979), each maintenance level l performed yields an age reduction coefficient $\theta_l \in [0, 1]$, and for any $l, l' \in [0, L_{ijk}]$, $\theta_l \geq \theta_{l'}$ if $l < l'$. For maintenance levels DN ($l = 1$) and replacement ($l = L_{ijk}$), the corresponding age reduction coefficients are $\theta_1 = 1$ and $\theta_{L_{ijk}} = 0$, respectively. In cases where a component has failed, the age reduction coefficient for mR ($l = 2$) is $\theta_2 = 1$. The following binary decision variable is introduced to model the selection of maintenance level l to be carried out on component E_{ijk}:

$$y_{ijkl} = \begin{cases} 1 & \text{if maintenance level } l \text{ is selected for component } E_{ijk}, \\ 0 & \text{otherwise.} \end{cases} \tag{4.3}$$

When a PM or CM maintenance action of level l is carried out on component E_{ijk} (*i.e.*, $y_{ijkl} = 1$), it reduces the component's effective age from B_{ijk} to $\theta_l B_{ijk}$, and consumes r_{ijklq} resources of type q. Furthermore, it requires t_{ijkl} time units and incurs a cost of c_{ijkl} monetary units such that:

$$t_{ijkl} = t^p_{ijkl} S_{ijk} + t^c_{ijkl} \left(1 - S_{ijk}\right) \tag{4.4}$$

$$c_{ijkl} = t_{ijkl}\, r_{ijklq}\, c^v_q \tag{4.5}$$

PM and CM maintenance actions play a pivotal role in increasing the reliability of individual components, thereby enhancing the overall reliability of each system which in turn increases the overall fleet reliability. By decreasing the effective age of components, notable reliability improvements can be achieved. In this study, without loss of generality, the lifetime of each component E_{ijk} is assumed to follow a Weibull distribution with shape and scale parameters β_{ijk} and η_{ijk}, respectively. For a component E_{ijk} with an effective age B_{ijk} at the start the SMB of system i, its conditional reliability R_{ijkl} for the upcoming mission of duration W, given that maintenance level l has been selected (*i.e.*, $y_{ijkl} = 1$), is calculated as follows:

$$R_{ijkl} = \exp\left[\left(\frac{\theta_l \cdot B_{ijk}}{\eta_{ijk}}\right)^{\beta_{ijk}} - \left(\frac{\theta_l \cdot B_{ijk} + W}{\eta_{ijk}}\right)^{\beta_{ijk}}\right] \cdot \mathbb{I}(l), \tag{4.6}$$

where $\mathbb{I}(l)$ is an indicator function defined as:

$$\mathbb{I}(l) = \begin{cases} S_{ijk} & \text{if } l = 1, \\ 1 & \text{otherwise.} \end{cases} \tag{4.7}$$

Equations (4.6) and (4.7) above guarantee that the calculated component reliability is set to 0 if the DN maintenance level ($l = 1$) is chosen for a failed component. Given that the architecture of each system in the fleet is series-parallel, then the reliability R_i of system i is computed as:

$$R_i = \prod_{j \in \mathcal{N}_i} \left(1 - \prod_{k \in \mathcal{M}_{ij}} \left(1 - \sum_{l \in \mathcal{L}_{ijk}} R_{ijkl} \cdot y_{ijkl} \right) \right). \tag{4.8}$$

Finally, given that the fleet successfully operates the upcoming mission if all its systems survive the mission (I-out-of-I), the reliability R_F of the fleet is then:

$$R_F = \prod_{i \in \mathcal{I}} R_i. \tag{4.9}$$

4.4 FSM Optimization Model and Solution Method

In this section, the ARCFMSP is formulated as a bi-objective optimization model to maximize the fleet reliability, and to minimize the grand total cost for a fleet comprising I multi-component systems nearing the end of their current missions and about to enter their respective SMBs. Notably, these SMB are asynchronous and can overlap, leading decision-makers to face complex managerial situations where systems compete for the same resources.

4.4.1 Bi-Objective Nonlinear Formulation

For modeling purposes, the following decision variables $x_{ijkl\tau}$ and z_q are introduced:

$$x_{ijkl\tau} = \begin{cases} 1 & \text{if maintenance level } l \text{ performed on } E_{ijk} \text{ is started at time } \tau \in \mathcal{D}_i \\ 0 & \text{otherwise.} \end{cases} \tag{4.10}$$

$$z_q = \text{integer maximum number of resource type } q \text{ used } (z_q \in \mathbb{Z}^+). \tag{4.11}$$

The studied ARCFMSP aims to jointly determine three key decisions: i) which components of each system to maintain, ii) what PM or CM maintenance level to carry out for each selected component, iii) when to perform the PM or CM maintenance actions while satisfying the resource requirements and availability. Furthermore the PM and CM maintenance actions must all fit within the break windows and satisfy reliability requirements and budget available. The proposed bi-objective optimization model

aiming to maximize fleet reliability and minimize maintenance cost is formulated as follows.

$$\max R_F = \prod_{i\in\mathcal{I}}\left(\prod_{j\in\mathcal{N}_i}\left(1-\prod_{k\in\mathcal{M}_{ij}}\left(1-\sum_{l\in\mathcal{L}_{ijk}} R_{ijkl}\cdot y_{ijkl}\right)\right)\right) \tag{4.12a}$$

$$\min TC = \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{N}_i}\sum_{k\in\mathcal{M}_{ij}}\sum_{l\in\mathcal{L}_{ijk}}\sum_{q\in\mathcal{Q}} t_{ijkl}\, r_{ijklq}\, c_q^v\, y_{ijkl} + \sum_{q\in\mathcal{Q}} c_q^f\, z_q \tag{4.12b}$$

$$\text{s.t.} \sum_{l\in\mathcal{L}_{ijk}} y_{ijkl} = 1, \qquad \forall i,j,k \tag{4.12c}$$

$$\sum_{\tau\in\mathcal{D}_i} x_{ijkl\tau} = y_{ijkl}, \qquad \forall i,j,k,l\neq 0 \tag{4.12d}$$

$$\sum_{\tau\in\mathcal{D}_i} x_{ijkl\tau}\,(\tau + t_{ijkl}) \leq ES_i + D_i, \qquad \forall i,j,k,l \tag{4.12e}$$

$$\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{N}_i}\sum_{k\in\mathcal{M}_{ij}}\sum_{l\in\mathcal{L}_{ijk}}\sum_{q\in\mathcal{Q}} t_{ijkl}\, r_{ijklq}\, c_q^v\, y_{ijkl} + \sum_{q\in\mathcal{Q}} c_q^f\, z_q \leq C_0 \tag{4.12f}$$

$$\prod_{i\in\mathcal{I}}\left(\prod_{j\in\mathcal{N}_i}\left(1-\prod_{k\in\mathcal{M}_{ij}}\left(1-\sum_{l\in\mathcal{L}_{ijk}} R_{ijkl}\cdot y_{ijkl}\right)\right)\right) \geq R_0 \tag{4.12g}$$

$$\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{N}_i}\sum_{k\in\mathcal{M}_{ij}}\sum_{l\in\mathcal{L}_{ijk}}\sum_{b=\max\{ES_i,\tau-t_{ijkl}+\Delta T\}}^{\tau} x_{ijklb}\, r_{ijklq} \leq z_q, \qquad \forall q,\tau \tag{4.12h}$$

$$z_q \leq \Gamma_q, \qquad \forall q \tag{4.12i}$$

$$z_q \leq \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{N}_i}\sum_{k\in\mathcal{M}_{ij}}\sum_{l\in\mathcal{L}_{ijk}} y_{ijkl}\cdot r_{ijklq}, \qquad \forall q \tag{4.12j}$$

$$y_{ijkl}, x_{ijkl\tau} \in \{0,1\}, \qquad \forall i,j,k,l,\tau \tag{4.12k}$$

$$z_q \in \mathbb{Z}^+, \qquad \forall q \tag{4.12l}$$

The first objective function in Equation (4.12a) maximizes the overall fleet reliability and the second objective function in Equation (4.12b) minimizes the grand total maintenance cost. Constraints (4.12c) ensure that each component is subjected to only one maintenance action. Constraints (4.12d) guarantee that each maintenance action selected for system i takes place within its SMB window defined as $[ES_i, ES_i + D_i]$. Constraints (4.12e) mandate that, for a system, all maintenance activities cannot last longer than the end of its break period. Constraint (4.12f) assures that the total maintenance cost (*i.e.,* sum of variable and fixed maintenance costs) does not exceed the total budget C_0 available and constraint (4.12g) is the threshold fleet reliability constraint. Constraints (4.12h) and (4.12i), adapted from the disaggregated discrete-time formulation of the RCPSP (Christofides et al., 1987), ensure that the PM or CM maintenance activities starting at any given time do not use more of each resource q than is available in that same time period. These constraints apply to renewable resources such as crews, tools and repair bays/channels, which are released and available after the activities using them are completed. Constraints (4.12j) define the maximum number of resource type q used if PM/CM maintenance level l is assigned to component E_{ijk}. Lastly, the decision variable domains are set by constraints (4.12k) and (4.12l).

4.4.2 Model Reformulation

The optimization model (4.12) presented above is nonlinear due to the fleet reliability function $R_F = \prod_{i\in\mathcal{I}} R_i$. First, the product operation is linearized by taking the natural logarithm to yield $\ln(R_F) = \sum_{i\in\mathcal{I}} \ln(R_i)$. Next, a piecewise linear approximation (PWLA) as in (Al-Jabouri et al., 2023b) is used to linearize the non-linear subsystem reliability functions R_i. The linearization approximation described below uses SOS2 constraints (Beale and Forrest, 1976). This approximation involves expressing the natural logarithm of each reliability R_{ij} of the jth subsystem of system i as $\ln(R_{ij}) = \ln(1 - F_{ij})$, with $F_{ij} = 1 - R_{ij}$ as the failure time distribution of that subsystem. Combined with the fact that

$\sum_{l\in\mathcal{L}_{ijk}} y_{ijkl} = 1$ from Equation (4.12c), one has:

$$F_{ij} = 1 - R_{ij} \tag{4.13}$$

$$= \prod_{k\in\mathcal{M}_{ij}} \left(1 - \sum_{l\in\mathcal{L}_{ijk}} y_{ijkl}\cdot R_{ijkl}\right) \tag{4.14}$$

$$= \prod_{k\in\mathcal{M}_{ij}} \left(\sum_{l\in\mathcal{L}_{ijk}} y_{ijkl}\cdot F_{ijkl}\right) \tag{4.15}$$

$$\ln(F_{ij}) = \sum_{k\in\mathcal{M}_{ij}} \left(\sum_{l\in\mathcal{L}_{ijk}} y_{ijkl}\cdot \ln(F_{ijkl})\right) \tag{4.16}$$

The two nonlinear functions $u(F_{ij}) = \ln(F_{ij})$ and $v(F_{ij}) = \ln(1 - F_{ij})$ can be approximated by their respective piecewise-linear functions $\hat{u}$ and $\hat{v}$. Using a set of P breakpoints $\{\hat{F}_{ijp}\}$ ($p \in \mathcal{P} = \{1,\ldots,P\}$) for the nonlinear functions $v(F_{ij})$ and $u(F_{ij})$, and the Special-Ordered Set of type 2 (SOS-2) variables $\lambda_{ijp} \geq 0$ and $w_{ijp} \in \{0,1\}$, the piecewise linear reformulation of the ARCFMSP model in (4.12) is given by:

$$\max \quad R_F = \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{N}_i}\sum_{p\in\mathcal{P}} \ln(1-\hat{F}_{ijp})\,\lambda_{ijp} \tag{4.17a}$$

$$\min \quad TC = \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{N}_i}\sum_{k\in\mathcal{M}_{ij}}\sum_{l\in\mathcal{L}_{ijk}}\sum_{q\in\mathcal{Q}} t_{ijkl}\, r_{ijklq}\cdot c_q^v\, y_{ijkl} + \sum_{q\in\mathcal{Q}} c_q^f\, z_q \tag{4.17b}$$

$$\text{s.t.} \quad \sum_{p\in\mathcal{P}} \ln(\hat{F}_{ijp})\,\lambda_{ijp} = \sum_{k\in\mathcal{M}_{ij}}\sum_{l\in\mathcal{L}_{ijk}} y_{ijkl}\,\ln(F_{ijkl}), \qquad \forall i,j \tag{4.17c}$$

$$\sum_{p\in\mathcal{P}} \lambda_{ijp} = 1, \qquad \forall i,j \tag{4.17d}$$

$$\lambda_{ijp} \leq w_{ijp}, \qquad \forall i,j,p \tag{4.17e}$$

$$\sum_{p\in\mathcal{P}} w_{ijp} \leq 2, \qquad \forall i,j \tag{4.17f}$$

$$w_{ijp} + w_{ij(p+q)} \leq 1, \qquad \forall i,j,p\in\{1,\ldots,P-2\}, q\in\{2,\ldots,P-p\} \tag{4.17g}$$

$$\lambda_{ijp} \geq 0, \qquad \forall i,j,p \tag{4.17h}$$

$$w_{ijp} \in \{0,1\}, \qquad \forall i,j,p \tag{4.17i}$$

$$\text{Equations (4.12c)} - \text{(4.12l)}. \tag{4.17j}$$

Equation (4.17a) approximates the natural logarithm of the fleet reliability and constraints (4.17c) define the natural logarithm of the unreliability of each subsystem within each system. Equation (4.17b) is the second objective function: total maintenance cost. Constraints (4.17d)–(4.17g) enforce the SOS-2 conditions (Hummeltenberg, 1984), which require that at most two variables in the specified ordered list are allowed to take a non-zero value (4.17f), and those non-zero variables must be contiguous in the list (4.17g). Constraints (4.17h) and (4.17i) define the domains of the SOS-2 variables.

A hierarchical or lexicographic approach is taken to deal with the obtained bi-objective model. The objectives are prioritized and optimized sequentially. The process begins with the highest-priority objective and proceeds to the other objective. When optimizing for one objective, only solutions that do not degrade the objective values of higher-priority objectives are considered (Gurobi Optimization, 2024). This method ensures that improvements in lower-priority goals never come at the expense of more important ones. The general procedure for this approach is described below.

— **Priority given to reliability maximization (Sequence 1):** Formulation (4.17a) – (4.17j) is solved, ignoring the second objective function in Equation (4.17b), to find the optimal reliability value R_F^*. Then R_0 is set to R_F^*, and formulation (4.17b) – (4.17j) is solved, leading to the determination of TC^*.

— **Priority given to maintenance cost minimization (Sequence 2):** Formulation (4.17b) – (4.17j) is solved to find the optimal maintenance cost TC^*. Then C_0 is set to TC^*, and formulation (4.17a) – (4.17j) is solved, ignoring the second objective function in Equation (4.17b), leading to the computation of R_F^*.

4.5 Results and Discussions

This section presents numerical experiments to illustrate that the proposed ARCFMSP model is effective at delivering valid and appropriate maintenance decisions. Four numerical experiments are conducted and fully discussed. The first experiment validates the proposed model by comparing its results to those obtained by prior studies in the literature. In the second experiment, the impact of the overlap of the SMB windows on maintenance schedule and reliability achieved is investigated. The third experiment evaluates the models' performance for a real-world example. Lastly, the fourth experiment explores how variations in SMB duration affect the fleet reliability as resource availability varies.

For all experiments, the optimization runs were carried out by the commercial solver Gurobi 9.5 using gurobipy. All experiments are run on a 12th Gen Intel(R) Core(TM) i7 2.60GHz laptop computer with 16GB of RAM running Windows 11™. To simplify the notation, monetary and time units are ignored.

4.5.1 Experiment 1: Validation

This first experiment validates the model by comparing its results with those reported by Cassady et al. (2001b), Pandey et al. (2013) and Khatab et al. (2018b). A single ($I = 1$) 2-by-2 identical four-component series-parallel system configuration is used for this comparison. It consists of two ($N_1 = 2$) subsystems, each with two ($M_{1j} = 2$) parallel components ($j = 1, 2$) arranged in parallel. For these components, two maintenance levels are available: $l = 1$ for DN, and $l = 2$ for Replacement (R). Similar to Khatab et al. (2018b), we set $D_1 = 5$, c_q^f=50 and c_q^v=2. The other parameters are set as follows: $\eta_{111} = \eta_{112} = 15$, $\eta_{121} = \eta_{122} = 20$, $\beta_{111} = \beta_{112} = 1.5$, $\beta_{121} = \beta_{122} = 3$, $S_{111} = 1$, $S_{112} = 1$, $S_{121} = 0$, $S_{122} = 1$, $B_{111} = 15$, $B_{112} = 20$, $B_{121} = 8$, $B_{122} = 15$, $t_{1112}^p = 5$, $t_{1122}^p = 5$, $t_{1212}^c = 2$, and $t_{1222}^p = 4$. The model is then solved for different values of C_0 and R_0 is set to 0. Priority is given to reliability maximization and 600 breakpoints are used for the PWLA as it offers a good trade-off between accuracy and computation times (Al-Jabouri et al., 2023b, 2024a). The planning period is discretized into 1-unit time periods ($\Delta T = 1$) and the outcomes of the experiment are displayed in Table 4.1. This table gives, for each value of the maintenance budget, the maximal achievable fleet reliability, the total number of resources consumed, the total maintenance time, in addition to the optimal SM plan and its corresponding total grand cost. The results match exactly with the ones reported in Khatab et al. (2018b).

Table 4.1 – Results for different values of C_0 when $D_1 = 5$.

C_0	R_F^* (%)	Resources used	TC^* ($)	Maintenance time (hr)	Component-maintenance level	
					DN	Replace (R)
200	85.89	3	174	12	E_{122}	E_{112};E_{121};E_{122}
160	77.53	2	114	7	E_{111}; E_{122}	E_{121}; E_{112}
100	59.71	1	54	2	E_{111}; E_{112};E_{122}	E_{121}
50	20.75	0	0	0	E_{111}; E_{112};E_{121}; E_{122}	—

Observations from Table 4.1 indicate, as one may expect, that a reduction in the available maintenance budget leads to a decrease in the number of resources, the total maintenance time, and the number of components replaced, resulting in a lower maximum reliability. Specifically, when the budget C_0 is reduced from 200 to 100, the count of resources used falls from 3 to 1, and concurrently, the number of components replaced diminishes from 3 to 1. Consequently, this budget reduction causes the maximum achievable reliability to plummet from 85.89% to 59.71%. A budget of $C_0 = 50$ is too low to carry out any maintenance, resulting in a low reliability of 20.75%.

4.5.2 Experiment 2: Impact of Overlap Between SMB Windows

The primary contribution of the model introduced in this chapter is its ability to accommodate the asynchronous start times of PM and CM maintenance activities. The current experiment aims to investigate the impact of varying SMB windows overlap on maintenance and resources assignment decisions. For this purpose, a fleet comprising three ($I = 3$) systems is examined, where each system i ($i = 1, 2, 3$) features four ($N_i = 4$) subsystems in series, with each subsystem j ($i = 1, \cdots, N_i$) containing four ($M_{ij} = 4$) parallel components.

To investigate the effects of SMB overlap on the on maintenance and resources assignment decisions, the case with 100% overlap (*i.e.,* synchronous SMB) is set as the baseline as it is the only setting modeled in the extant literature. The amount of overlap is then reduced to 75%, 50%, 25% and 0% (*i.e.,* complete asynchronous). In this experiment, a single type of resource is utilized, with the mission length set to $W = 6$, and the SMB durations $D_i = 12$. Additionally, 6 CM levels (L_{ijk}=6) are available for failed components: $l = 1$ for DN, $l = 2$ for mR, $l = 3$ for IM action 1 (IM1) with $\theta_3 = 0.75$, $l = 4$ for IM action 2 (IM2) with $\theta_4 = 0.50$, $l = 5$ to IM action 3 (IM3) with $\theta_5 = 0.25$, and l_6 for replacement (R). There are 5 PM levels for working components: $l = 1$ for DN, $l = 3$ for IM1 with $\theta_3 = 0.75$, $l = 4$ for IM2 with $\theta_4 = 0.5$, $l = 5$ for IM3 with $\theta_5 = 0.25$, and l_6 for replacement (R). Functioning components cannot undergo mR. There is $Q = 1$ resource type and a maximum of $\Gamma_1 = 25$ resources available per period with $c_1^f = 50$ and $c_1^v = 4$. DN does not require any resources. Exactly one resource is required for any maintenance level l for components E_{111}, E_{112}, E_{113}, E_{122}, E_{123}, E_{131}, E_{132}, E_{143}, E_{212}, E_{214}, E_{221}, E_{223}, E_{234}, E_{311}, E_{312}, E_{321}, E_{322}, E_{323}, E_{332}, E_{333}, E_{342}, and E_{344}. Maintenance of all other components require two resources. For functioning components, we have: $t^p_{ijk1} = 0$, $t^p_{ijk3} = 2$, $t^p_{ijk4} = 4$, $t^p_{ijk5} = 6$, and $t^p_{ijk6} = 9$. Recall that $l = 2$ is not available for working components. For failed components, we have: $t^c_{ijk1} = 0$, $t^c_{ijk2} = 2$, $t^c_{ijk3} = 3$, $t^c_{ijk4} = 6$, $t^c_{ijk5} = 9$, and $t^c_{ijk6} = 14$. Additional lifetime parameter values for components relevant to this experiment are detailed in Table 4.2. All parameter values are either taken from the existing literature (Schneider and Cassady, 2004; Diallo et al., 2019; Khatab et al., 2022) or arbitrarily set while being reasonable when no similar features or parameters exist in the extant literature.

Table 4.2 – Input data for experiment #2.

System 1					System 2					System 3				
E_{ijk}	η_{ijk}	β_{ijk}	B_{ijk}	S_{ijk}	E_{ijk}	η_{ijk}	β_{ijk}	B_{ijk}	S_{ijk}	E_{ijk}	η_{ijk}	β_{ijk}	B_{ijk}	S_{ijk}
E_{111}	17	2.10	14	1	E_{211}	12	2.42	11	0	E_{311}	17	3.31	12	1
E_{112}	12	2.27	10	0	E_{212}	15	2.95	12	1	E_{312}	18	2.94	11	0
E_{113}	11	3.43	6	1	E_{213}	13	2.15	5	1	E_{313}	19	3.48	14	1
E_{114}	15	2.37	11	1	E_{214}	10	3.12	9	1	E_{314}	10	2.68	15	1
E_{121}	12	2.73	7	0	E_{221}	14	3.28	5	1	E_{321}	19	2.25	14	1
E_{122}	17	2.11	15	1	E_{222}	14	1.83	12	1	E_{322}	14	1.83	12	1
E_{123}	12	2.46	10	1	E_{223}	13	3.47	7	1	E_{323}	16	1.88	14	0
E_{124}	10	2.59	5	1	E_{224}	16	1.92	6	1	E_{324}	16	1.90	14	1
E_{131}	13	1.66	12	1	E_{231}	17	3.23	13	1	E_{331}	10	2.82	14	1
E_{132}	16	3.43	12	1	E_{232}	20	2.99	8	1	E_{332}	16	2.15	15	1
E_{133}	18	3.48	9	1	E_{233}	11	2.75	15	1	E_{333}	11	2.71	10	0
E_{134}	11	1.83	15	1	E_{234}	19	2.32	11	0	E_{334}	18	2.52	10	1
E_{141}	12	2.43	14	1	E_{241}	11	3.04	6	0	E_{341}	19	3.26	5	1
E_{142}	17	2.90	13	1	E_{242}	14	2.22	11	0	E_{342}	17	2.72	12	0
E_{143}	10	3.24	12	1	E_{243}	17	2.89	13	1	E_{343}	15	3.15	10	0
E_{144}	15	1.69	6	1	E_{244}	16	2.38	10	1	E_{344}	16	1.89	11	1

The results obtained are presented in Table 4.3, which gives for each overlap percentage considered the optimal fleet reliability obtained R^*_F, the number of resources used, the variable costs, fixed costs, the grand total cost, the distribution of PM/CM maintenance times per system, the total maintenance

times, the distribution of PM/CM maintenance levels carried out, and the total number of maintenance actions performed.

When fleet reliability is prioritized and a maintenance budget of $C_0 = \$1200$ is set, the results show that as the overlap decreases, the fleet reliability achieved increases. For example, when the overlap decreases from 100% to 50%, the fleet reliability increases from 81.71% to 90.74% with the number of replacements (R) edging from 9 to 13. This is due to the fact that decreasing overlap means that there is less competition for the resources, thus leading to more maintenance actions performed overall. In addition, more high-quality maintenance actions, which tend to take longer and cost more, are also performed.

Another noteworthy result observed is the reduction in fixed costs alongside an increase in variable costs. This indicates that in industries with high fixed costs due to specialized resources, the model can provide reasonable decisions. In such industries, if stakeholders (*i.e.,* maintenance managers or schedulers) can minimize the overlapping of activities, they can avoid over-utilising a specific resource, thereby lowering total costs without compromising reliability levels.

Figure 4.3 is a Gantt chart representation of the maintenance plans for the 75%, 50%, and 0% overlap cases. Each rectangle on the chart represents the maintenance action conducted. Each rectangle (or maintenance action) is labeled $ijk-l$ denoting the component identification ijk and the maintenance level l performed. Replacements, the highest maintenance level available, denoted by R are shown in darker shades to make them standout. It can be seen in accordance with the results in Table 4.3 that the number of replacements increases from 11 to 13 when the overlap decreases from 75% to 50%. Similarly, the number of replacements increases from 11 to 14 when the overlap decreases from 75% to 0%.

The outcomes when maintenance cost is prioritized are also detailed in Table 4.3 when a reliability target $R_0 = 96\%$ is set. The findings indicate that decreasing the overlap between SMB windows leads to a notable reduction in the grand total cost TC^*, while the achievable reliability remains relatively stable. This is an expected behavior as less competition for resources results in needing less resources. Hence, the number of resources used decreases as well as the total maintenance costs. For example, the total count of maintenance actions decreases from 35 to 27 and the number of resources used drops from 24 to 9 when the break overlap is reduced from 100% to 0%.

4.5.3 Experiment 3: Application to a Coal Transportation System Case Study

This experiment aims to demonstrate the model's real-world applicability and its alignment with industrial settings. For this purpose, we consider a fleet composed of $I = 3$ identical coal transportation systems that supply a boiler in a plant, reflecting a practical scenario. This example is inspired by the case study presented by Khatab et al. (2022). Figure 4.4 displays the RBD of each system in this fleet, comprising $N_i = 5$ subsystems ($i = 1, 2, 3$): two conveyors, two feeders, and a stacker-reclaimer. The parameter values for each system component is given in Table 4.4.

To study the effects of different resources on the maintenance and resources assignment decisions, $Q = 3$ types of resources are available, with their respective fixed costs set at 50, 75, and 20, and their variable cost rates are set at 4, 8, and 2 respectively. Additionally, the SMB duration D_i for each system i ($i = 1, 2, 3$) is considered to be 10, 12, and 14, respectively, and their respective earliest start times ES_i are set at 0, 5, and 2. The mission length is set to $W = 50$. Four CM levels are available for failed components: $l = 1$ (DN), $l = 2$ (mR), $l = 3$ (IM1) and $l = 4$ (R). For working components, three PM levels are considered: $l = 1$ (DN), $l = 3$ (IM1), and $l = 4$ (R).

The bi-objective formulation is solved using the hierarchical approach described previously using Sequence 1 (priority given to reliability maximization) and compared to the single objective formulation maximizing only fleet reliability R_F. For $C_0 = 2500$ and various maintenance resource amounts, the results obtained are depicted in Figure 4.5. These values indicate that an increase in the number of available resources from 2 to 12 increases the fleet reliability from around 83% to about 91% because more resources mean less competition and more maintenance performed. The proposed hierarchical approach

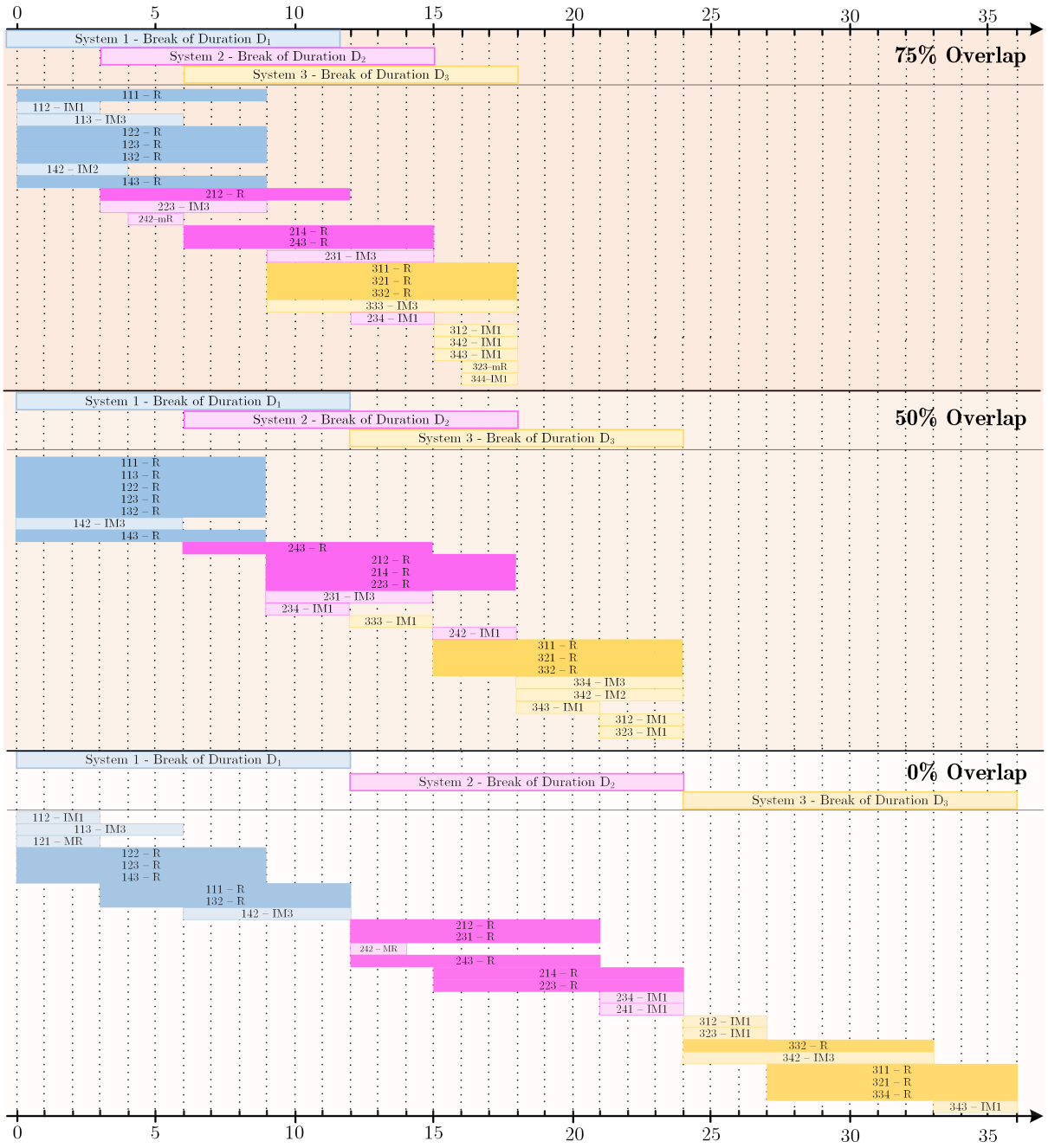


Figure 4.3 – Gantt charts of maintenance plans for 75%, 50%, and 0% overlaps.

Table 4.3 – Results for models 1 and 2: case of experiment #2.

Priority given to reliability maximization – Sequence 1

Overlap	R_F^*	Resources	Variable	Fixed	TC^*	M. Time (hr)			Total. M	Maint. levels performed						Nb. of Maint.
(%)	(%)	used	Cost ($)	Cost ($)	($)	S1	S2	S3	Time (hr)	DN	mR	IM1	IM2	IM3	R	Actions
100	81.71	12	576	600	1176	48	36	39	123	28	0	8	0	3	9	20
75	88.29	10	700	500	1200	58	44	49	151	24	2	6	1	4	11	24
50	90.74	8	768	400	1168	60	48	51	159	25	0	6	1	3	13	23
25	92.50	7	840	350	1190	63	51	54	168	24	0	7	0	3	14	24
0	92.77	7	848	350	1198	62	63	54	169	23	2	6	0	3	14	25

Priority given to maintenance cost minimization – Sequence 2

Overlap	TC^*	Variable	Fixed	R_F^*	Resources	M. Time (hr)			Tot. Maint.	Maint. levels performed						Nb. of Maint.
(%)	($)	Cost($)	Cost($)	(%)	used	S1	S2	S3	time(hr)	DN	mR	IM1	IM2	IM3	R	Actions
100	2316	1116	1200	96.00	24	75	65	80	220	13	0	14	0	4	17	35
75	1900	1100	800	96.01	16	73	64	78	215	16	2	6	3	5	16	32
50	1688	1088	600	96.03	12	73	63	80	216	18	2	4	3	4	17	30
25	1580	1080	500	96.02	10	73	65	78	216	20	1	3	2	4	18	28
0	1542	1092	450	96.02	9	77	65	77	219	18	2	5	0	6	17	30

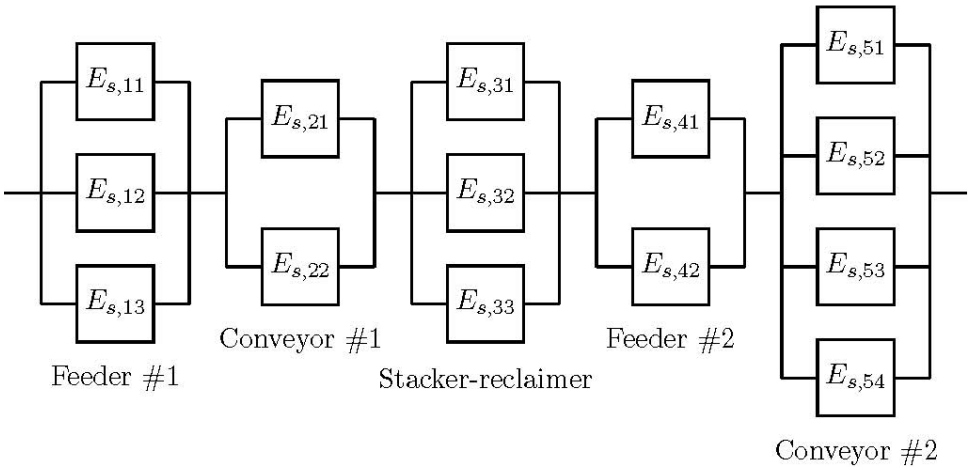


Figure 4.4 – Real-world coal conveyor system (Liu and Huang, 2010; Khatab et al., 2022).

Table 4.4 – Input data for experiment #3.

					$l=2$ (mR)				$l=3$ (IM1)				$l=4$ (R)			
E_{ijk}	η_{ijk}	β_{ijk}	B_{ijk}	S_{ijk}	t_{ijk2}	r_{ijk21}	r_{ijk22}	r_{ijk23}	t_{ijk3}	r_{ijk31}	r_{ijk32}	r_{ijk33}	t_{ijk4}	r_{ijk41}	r_{ijk42}	r_{ijk43}
E_{i11}	150	1.5	35	1	–	–	–	–	2	1	1	1	3	1	1	1
E_{i12}	228	2.4	24	0	1	2	0	1	2	2	0	1	5	2	0	1
E_{i13}	168	1.6	45	0	1	2	0	1	2	2	0	1	6	2	0	1
E_{i21}	240	2.6	35	0	1	1	1	1	2	1	1	1	6	1	1	1
E_{i22}	168	1.8	28	1	–	–	–	–	1	2	1	0	2	2	1	0
E_{i31}	204	2.4	36	1	–	–	–	–	2	1	1	0	3	1	1	0
E_{i32}	156	2.5	44	0	1	1	1	1	2	1	1	1	6	1	1	1
E_{i33}	168	2.0	28	0	1	2	0	1	2	2	0	1	6	2	0	1
E_{i41}	156	1.2	38	1	–	–	–	–	1	2	0	1	2	2	0	1
E_{i42}	210	1.4	16	0	1	2	1	0	2	2	1	0	6	2	1	0
E_{i51}	240	2.8	30	0	1	1	0	0	2	1	0	0	6	1	0	0
E_{i52}	210	1.5	22	1	–	–	–	–	1	2	1	1	2	2	1	1
E_{i53}	180	2.4	38	1	–	–	–	–	2	1	1	1	3	1	1	1
E_{i54}	270	2.2	35	0	1	2	0	0	2	2	0	0	6	2	0	0

(Sequence 1) guarantees the minimal grand total cost for the optimal fleet reliability. The trends from Figure 4.5 show that the fleet reliability increases significantly from 83% to 90.54% when the number of resources increases from 2 to 8. After 8 resources, the fleet reliability hits a plateau with diminishing returns, as other constraints such as time and budget become limiting factors. Indeed, when the number of resources increases from 8 to 12, the fleet reliability only increases from 90.54% to 90.95%. A point is reached where the addition of more available resources does not translate to higher fleet reliability. Other factors such as SMB duration and maintenance budget must be considered if further improvement in reliability is desired.

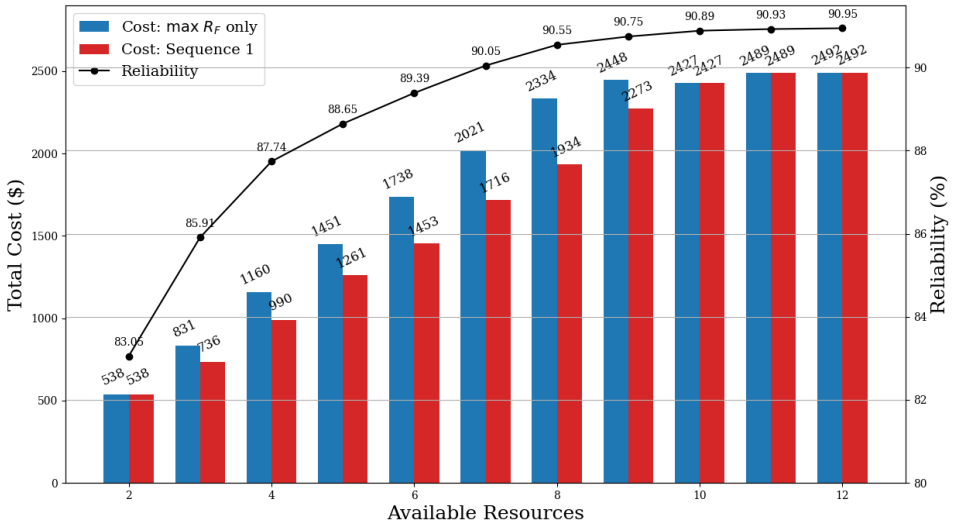


Figure 4.5 – Trends of fleet reliability R_F^* (%) and total Cost as the number of resources increase.

Given the importance of the available budget in activity planning, the problem was analyzed with varying C_0 values across different resource scenarios, with findings presented in Figure 4.6. Notably, the results reveal that for each budget level C_0, increasing the number of available resources increases the fleet reliability to a certain point. For instance, with C_0=1000, an increase in available resources beyond five does not further improve R_F^*. This observation is consistent across all C_0 levels, highlighting a key consideration for resource allocation. This indicates that augmenting the number of available resources, in conjunction with a specific C_0 up to a certain threshold can improve R_F^*. To achieve higher levels of R_F^*, it is essential to adjust the budget and number of available resources together.

4.5.4 Experiment 4: Impact of SMB Duration on Fleet Reliability

This set of experiments is designed to further investigate the critical roles of SMB duration and break overlap. These factors have a major influence on maintenance and resource assignment decisions, which in turn may greatly impact the overall fleet reliability that can be achieved as well-demonstrated in experiments #2 and #3. For experiment #4, the fleet and data presented in experiment #3 are used. The SMB duration of the second system is kept constant at $D_2 = T$, while the duration of the first system is reduced by Δ_D% (*i.e.,* $D_1 = T\,(1 - \Delta_D)$) and the duration of the third system is increased by Δ_D% (*i.e.,* $D_3 = T\,(1 + \Delta_D)$), so that the average break duration is set at T as shown on Figure 4.7. In what follows, Δ_D% is referred to as the end date percent difference (EDPD). In the first experiment, the average SMB duration is set to $T = 20$ while ΔD is varied from 0% to 100%, where the particular value $\Delta_D = 0$% refers to the synchronous SMB configuration. When 5 resources are available, the results obtained are reported in Table 4.5.

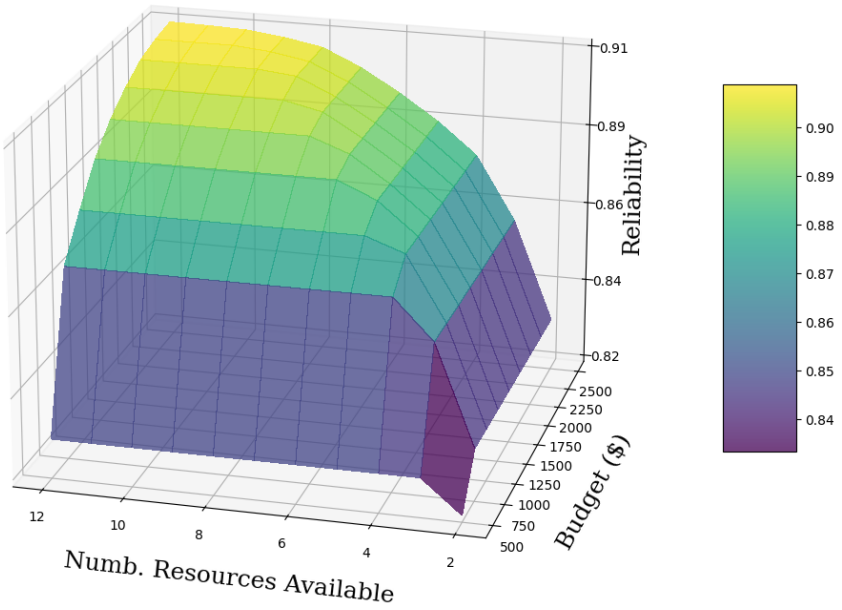


Figure 4.6 – Impact of the number of resources and available budget (C_0) on R_F^*.

The fleet reliability for the baseline synchronous SMB case is 89.31%. As Δ_D increases, the optimal fleet reliability increases until reaching a maximum of at least 90.38% for $\Delta_D \in [40\%, 50\%]$. Beyond this range, further increases in Δ_D lead to a decline in fleet reliability, reaching a minimum of 35.33% at $\Delta_D = 100\%$. Initially, when Δ_D increases, there is more spread between SMB windows which results in less competition for resources, thus yielding more maintenance actions and hence increased fleet reliability. Looking at Figure 4.8, the total time spent maintaining each system generally increases with Δ_D when $\Delta_D \leq 50\%$. Past this point, while the trend remains consistent for systems 2 and 3, the maintenance time for system 1 significantly decreases as Δ_D exceeds 50%, due to the reduction of its corresponding SMB duration. This reduction in maintenance time negatively impacts the reliability of system 1, contributing to an overall reduction in fleet reliability. The key managerial insight here, is that it can be beneficial for the maintenance decision maker to appropriately spread the end dates of the maintenance breaks to increase the fleet reliability with the same resources. Substantial improvements in fleet reliability can be achieved by adjusting mission start times of different systems within the fleet, advancing some and delaying others.

Table 4.5 – Fleet reliability R_F^* (%) as Δ_D% varies.

Δ_D (%)	0	10	20	30	40	50	60	70	80	90	100
D_1	20	18	16	14	12	10	8	6	4	2	0
D_2	20	20	20	20	20	20	20	20	20	20	20
D_3	20	22	24	26	28	30	32	34	36	38	40
R_F^* (%)	89.31	89.74	90.06	90.25	**90.38**	**90.38**	90.34	90.06	89.52	87.37	35.33

To examine the impact of variations in the number of available resources, the ARCFMSP was solved and fully analyzed across different levels of available resources. All parameters remain the same except the average break duration is set to $T = 10$ b.t.u. The outcomes, as depicted in Figure 4.9, show the profile

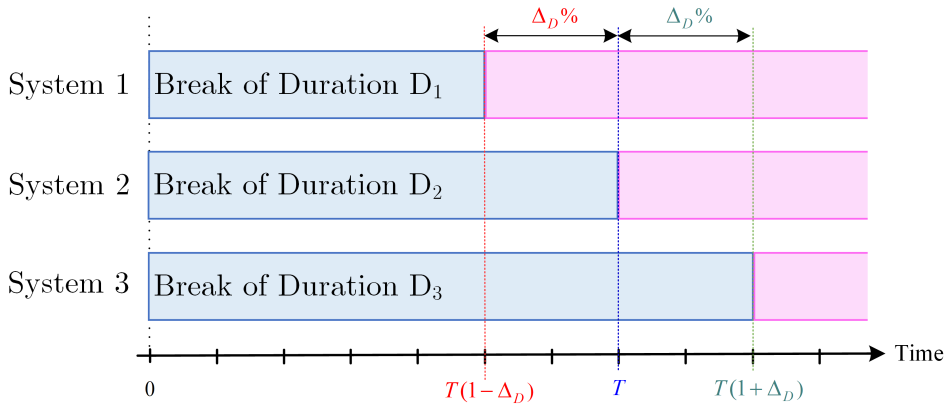


Figure 4.7 – Break duration structure for experiment #4.

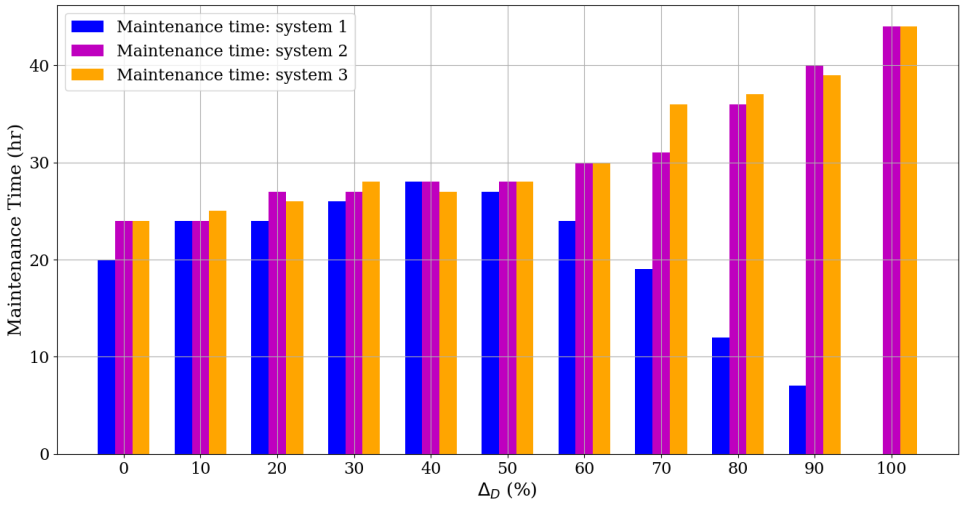


Figure 4.8 – Maintenance times by system and end date percent difference Δ_D.

of the fleet reliability R_F^* by the number of available resources (*i.e.,* 6, 12 and 18) for varying end date percent difference Δ_D. The dotted lines indicate the baseline fleet reliability obtained for each resource level when $\Delta_D = 0$. The results indicate that as the number of available resources grows, the improvement in fleet reliability from the baseline case (*i.e.,* $\Delta_D = 0$) obtained by increasing the EDPD Δ_D diminishes. In other words, the impact of increasing Δ_D is less significant when more resources are available. When resources are limited, increasing Δ_D allows the maintenance tasks to be staggered (*i.e.,* assign most maintenance resources to system 1, then system 2, then system 3), which helps in allocating the scarce resources across systems. When resources are plentiful, the staggered approach becomes less important because many maintenance tasks can be performed in parallel on multiple systems. Additionally, at every resource level, extending the EDPD and modifying break durations consistently yield reliability benefits. However, for every resource level, there is a point beyond which increasing Δ_D will result in a reduction in fleet reliability. This is due to the fact that setting Δ_D to large values implies a short break duration for system 1, thus little maintenance can be completed on this particular system, resulting in only small improvements in system reliability. Furthermore, these findings offer another crucial insight for managers, highlighting that in sectors with a constrained pool of resources, augmenting the planning horizon can significantly boost reliability, even with a fixed resource count and budget. In practical terms, increasing the EDPD means that the maintenance planner/scheduler should negotiate with stakeholders to set break end dates that are spread over the planning horizon. The benefits are more pronounced when resources are tight.

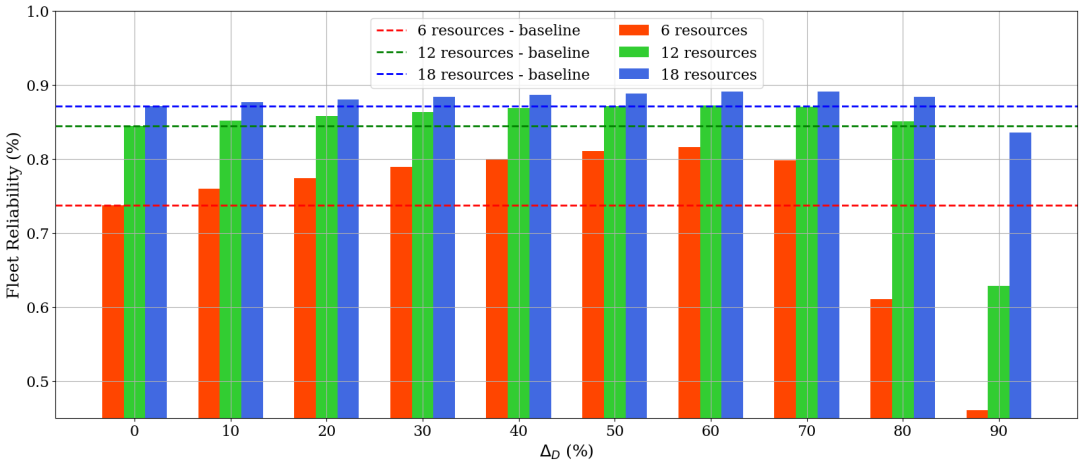


Figure 4.9 – Trends of R_F^* by available resources and end date percent difference Δ_D.

4.5.5 Managerial Insights

The experimental investigations detailed in this section conclusively demonstrate the effectiveness and adaptability of the proposed model, with a particular emphasis on optimizing maintenance strategies through asynchronous SMB windows for complex systems. The experiments have illustrated how strategic adjustments in maintenance scheduling, especially the asynchronous initiation of maintenance activities, significantly improve system reliability. Key findings reveal that increasing the number of available resources and extending the planning horizon, when applied within an asynchronous framework, positively impact reliability outcomes. The model's ability to adapt to essential parameter changes, such as SMB durations and number of available resource, further underscores its practical utility for industries facing resource constraints. These findings highlight the model's real-world applicability and provide a robust framework for managers aiming to improve maintenance planning and achieve desired reliability levels more efficiently. The proposed model is more general than previous FSMP models and is applicable to small and large fleets ranging from multiple production lines to fleet of military systems, from fleet of city buses or delivery drones to navy frigates.

4.6 Conclusions and Future Research Directions

This chapter presented a novel and significant extension of the FSMP by considering asynchronous maintenance breaks and resource-constrained maintenance planning. The proposed FSMP formulation for asynchronous breaks was shown to be more general and versatile in adapting to a broad spectrum of operational constraints and resource scarcity. The multiple numerical experiments conducted have highlighted the trade-offs between the timing and quality levels of maintenance activities and the consumption of resources that maintenance planners can make to obtain the best system performance for the budget and maintenance windows available. Two novel insights were derived. The first insight is that in the context of the FSMP, asynchronous scheduling of maintenance can yield better reliability than the synchronous case that is proposed in the current literature. The second insight is that it can be beneficial for the maintenance decision maker to appropriately spread the end dates of the maintenance breaks to increase the fleet reliability with the same resources. Substantial improvements in fleet reliability can be achieved by adjusting mission start times of different systems within the fleet, advancing some while delaying others.

Although the presented FSMP model is demonstrated within a series-parallel system context, it can be adapted to work with more complex reliability configurations such as k-out-of-n structures. The fleet reliability was modeled as a series configuration. The more general k-out-of-n configuration could be investigated as an extension. It is worth investigating the cases where the resources are non-renewable as well as the balanced utilization rate of renewable resources. Another extension that could render the model even more reflective of real-world scenarios is the incorporation of component aging. Additionally, integrating a system's availability represents another potential avenue for future research. Pursuing these improvements, however, may complicate the model, necessitating the adoption of diverse decomposition algorithms for effective solutions. Finally, an interesting research avenue is to include the overlap interval as an optimization variable. In the current work, the overlap percentages are known to the optimization model. It would be interesting to determine the optimal overlap percentages based on the resources available.

Chapter 5

Joint Selective Maintenance and Mission-Abort Decisions

5.1 Introduction

Many engineered assets and systems are designed to perform critical missions. These mission-critical systems encompass a broad spectrum of applications such as air transportation, military systems, nuclear plants, spacecraft, and unmanned autonomous devices. A key metric used to assess the performance of these systems is the MSP, defined as the probability of completing a defined mission of a specific duration under given operational conditions. To maximize the MSP for upcoming missions, maintenance actions are usually carried out on the system/components during scheduled inter-mission breaks. However, due to limitations in maintenance resources (*i.e.,* time, budget, materials, equipment, and crews) only a subset of the desired maintenance tasks can be performed. In the literature, this resource-constrained maintenance problem is known as the SMP. SM has been demonstrated to be suitable for MOS, *i.e.,* systems that operate according to a sequence of alternating missions and breaks. Its main objective is to identify the optimal subset of components and maintenance actions that minimize maintenance costs while ensuring that the system operates at a required level of reliability or MSP.

A substantial literature on SMP has spawned from the original work on SM by Rice Rice (1999). Researchers have since expanded upon the original SMP to include complex system configurations (Diallo et al., 2018), multistate systems (Ma et al., 2023; Pandey et al., 2013), component dependence (Dao and Zuo, 2017b; Feng et al., 2024; Xu et al., 2016), fleet level SM (Amjadian et al., 2024; Khatab et al., 2022; O'Neil et al., 2024), multiple repair channels (Chaabane et al., 2020; Diallo et al., 2019; Zhang et al., 2024a), uncertain mission, break, and maintenance durations (Al-Jabouri et al., 2024a; Jiang and Liu, 2020b; Jin et al., 2023; Liu et al., 2018; Yin et al., 2023), uncertain mission conditions (Ghorbani et al., 2024), uncertain maintenance quality (Al-Jabouri et al., 2023a; Khatab and Aghezzaf, 2016; Duan et al., 2018), and joint SM and workforce routing (O'Neil et al., 2023). For more details on the SMP, the reader is referred to the recent surveys in Al-Jabouri et al. (2022, 2024b).

All literature reviews conducted to date show that existing SM models are solely dedicated to MSP assessment and improvement. However, in many real-world situations, the survival of mission-critical systems, such as aircraft, spacecraft, unmanned aerial vehicles (UAVs), submarines, and chemical reactors, is more important than the completion of the mission due to the severe or catastrophic consequences of their failures Levitin et al. (2021c). In the airline industry, for example, failure of a key aircraft system during a flight may cause catastrophic economic, environmental, and social consequences. In such situations, resorting to a MAP indeed contributes to improving the system's survivability and reducing the magnitude of the risks that may arise from continuing the mission. A mission is aborted when specific conditions are satisfied, and is followed by a rescue/recovery procedure. In general, a MAP will lead to a reduction in MSP but an increase in SSP (Levitin et al., 2018a). Therefore, an optimal MAP shall achieve a balance between the two metrics (mission success versus system survival).

The optimization of MAP for mission-critical systems has recently received much attention in the literature. One of the first works on MAP was published in (Myers, 2009), which proposed a MAP optimization approach for a mission-critical k-out-of-n:G systems, where the number of failed redundant components is taken as the condition to abort the mission and to initiate the rescue procedure. Since then, a variety of MAP optimization models have been proposed based on different triggering conditions. MAP models where both the amount of completed mission work and the number of failed components are accounted for are developed by Levitin et al. (2017, 2019). Focusing on a warm-standby system, they jointly optimized the mission-abort decisions and the component activation sequence. Several studies have investigated the MAP for multistate systems where the level of component degradation is crucial in triggering mission abortion. For example, the works in Cheng et al. (2023, 2024); Levitin et al. (2021a); Zhao et al. (2021) assumed that component degradation is totally or partially revealed through inspections and propose optimization approaches to determine mission-abort decisions and/or inspection times. In some other works, *e.g.,* (Liu and Yang, 2023; Liu et al., 2024; Qiu and Cui, 2019; Yang et al., 2022; Zhao et al., 2023a), it is assumed that system components are equipped with sensors, allowing real-time health data monitoring. Other researchers have investigated MAP triggered by the number of external shocks a system has experienced. This policy was initially studied by Levitin and Finkelstein (2018) and has since been extended in several directions to include overlapping primary and rescue subsystems (Levitin et al., 2020), multi-state systems (Levitin et al., 2021a), balanced systems (Wu et al., 2021), random rescue duration (Levitin et al., 2023), resource-constrained missions (Levitin et al., 2024f), and self-adaptive MAP (Levitin et al., 2024b). Recent studies have focused on MAP applied to multi-attempt mission systems (Levitin et al., 2024a,c,d,e; Meng et al., 2024a,b).

Significant research has also been conducted on jointly optimizing mission-abort decisions and other factors that impact both mission success and system survival. For instance, Levitin et al. (2018b) jointly optimized mission-abort and component loading decisions for a 1-out-of-N:G warm standby system, while the abort decision is based on mission progress and component failures. Focusing on UAV applications, Zhu et al. (2021) developed a MAP approach to jointly determine target assignments, routing, and mission-abort decisions under external random shocks. Their approach is applied to the case of UAVs executing targeting operations. Dealing with UAVs in transportation, Zhao et al. (2024) proposed a joint optimization approach that identifies the optimal loading level, abort threshold, and location of rescue sites such that survivability is maximized and total expected cost minimized. In a recent study, Zhao et al. (2023b) proposed a MAP model to jointly optimize mission-abort decisions and system structure for a k-out-of-n:G warm standby system.

Despite the increasing number of studies dealing with MAP, very few references considered the impact of maintenance on mission-abort decisions. Only two papers, (Levitin et al., 2021b; Liu et al., 2022a), addressed the joint mission-abortion and maintenance decisions. In Levitin et al. (2021b), the problem of jointly optimizing maintenance and mission-abort decisions for a dual-unit standby system is investigated. During the mission, one component must be operating while the other can undergo maintenance. If a component fails, the primary mission may be aborted depending on the amount of work remaining. Multiple PM levels are considered to develop an optimization model that maximizes MSP while ensuring that the SSP meets a required minimum level. This work relies on the assumption that maintenance actions can be performed during the mission, making the approach inapplicable to many mission-critical systems that cannot be properly maintained during a mission such as UAVs, aircraft, and military equipment, to name a few. Furthermore, the limited maintenance time window is not considered. The second attempt, presented in Liu et al. (2022a), investigates the problem of jointly optimizing task termination and PM decisions for phased-task systems (*i.e.,* systems operating a mission composed of several phases). At the end of each phase, a decision is made to either terminate the task or perform PM to reduce component degradation. Several maintenance levels ranging from no action to complete replacement are considered. This problem is formulated as a MDP to identify the optimal PM level and the number of phases to carry out. The main limitations of Liu et al. (2022a) include the fact that the system is reduced to a single component and that task termination decisions can only be made at the end of each phase regardless of system failure during the phase. Another drawback is that the duration of the rescue procedure is negligible, thus causing the rescue procedure to be always successful. Additionally, while the cost of performing

maintenance is included in the optimization model, other resource constraints such as time and available repairpersons, which often need to be considered when planning maintenance tasks, are not accounted for.

Aiming to help maintenance decision-makers resolve practical issues arising in the implementation of the SMP, this chapter develops a new approach where the classical SMP is extended to include mission-abort policies. In this work, both SM and mission abortion decisions are integrated into a single joint optimization model. A multi-component mission-critical system that must operate a sequence of alternating missions and scheduled breaks is considered. Conditions under which a mission can be aborted and the rescue procedure activated are investigated. To improve the MSP and system survivability, a list of maintenance actions ranging from mR, through intermediate IM actions, to perfect repair (AGAN) are available for each component. To meet a pre-specified required performance level for the next mission, maintenance activities are allowed only during the break. However, due to the limited break duration and maintenance budget, not all components can be maintained. In contrast to the existing literature, a new integrated SMP and MAP is developed to jointly determine three decisions: (1) the selection of the components to be maintained, (2) the selection of the maintenance levels to be performed on those components, and (3) the time to abort a mission and activate the rescue procedure. The resulting integrated SMP and MAP problem is formulated as a nonlinear binary optimization model whose main components and characteristics are derived and fully discussed. Given the complex nature of the resulting problem, realistic-size instances cannot be solved using exact optimization methods. Hence, a GA is developed to solve it.

In summary, the present work advances the state-of-the-art of both SM and MAP literature. It is indeed the first to jointly address MAP and SM. Unlike all previously proposed SMP models, the proposed approach considers the possibility of aborting the primary mission and initiating a rescue procedure to enhance system survivability. The main contributions are:

1. A MINLP formulation that jointly integrates mission-abort and SM decisions for mission-critical multi-component systems is proposed. Unlike previous models dealing with MAP, the model developed in this chapter fully presents and discusses the optimization variables, parameters, and constraints.
2. Unlike other works on the joint optimization of mission-abort and maintenance decisions, we consider large systems comprising multiple subsystems and components. Additionally, available resources for maintenance tasks are directly accounted for in the optimization process.
3. Extensive experimentation including sensitivity analyses of key parameters are conducted to demonstrate the added value of jointly optimizing SM and mission-abort decisions. The insights gained from the numerical experiments can help guide decision makers in allocating maintenance resources effectively and efficiently setting optimal mission-abort policies.

This chapter is organized as follows. First, a comprehensive description of the system and problem under investigation is provided in Section 5.2. Section 5.3 develops the reliability and maintenance models for MSP and SSP. In Section 5.4, the novel combined mathematical formulation for optimizing the joint MAP and SMP is proposed along with a solution method. The mathematical model and solution approach are validated through extensive experiments in Section 5.5, where the results obtained are analyzed and extensively discussed. Finally, Section 5.6 concludes the chapter by summarizing the key findings and identifying promising directions for future research.

5.2 System Description

In this section, the system under consideration is described and the mission-abort policy followed is presented. First, the notation system is introduced.

5.2.1 Notation

The following notation system is used.

i Index of reliability blocks, $i = 1, 2, 3$

N_i Number of subsystems in block i, indexed by $j = 1, \ldots, N_i$

N_{ij}	Number of components in subsystem j of block i, indexed by $k = 1, \ldots, N_{ij}$
E_{ijk}	The kth component in subsystem j of block i
L_{ijk}	The highest maintenance level available for component E_{ijk}
l	Index of maintenance levels for E_{ijk}, $l = 1, \ldots, L_{ijk}$
$t^p_{ijkl}(c^p_{ijkl})$	Time (cost) to perform PM level l on component E_{ijk}
$t^c_{ijkl}(c^c_{ijkl})$	Time (cost) to perform CM level l on component E_{ijk}
$\varphi_{ijk}(\psi_{ijk})$	Effective age of component E_{ijk} at the start (end) of the maintenance break
$u_{ijk}(v_{ijk})$	Status of component E_{ijk} at the start (end) of the maintenance break
D	Maintenance break duration
T	Mission duration
V	Mission-abort decision threshold ($V < T$)
$\mathcal{R}_{ijk}(t)$	Reliability of component E_{ijk} for a mission of length t
$\mathcal{R}_{ij}(t)$	Reliability of subsystem j of block i for a mission of length t
$\mathcal{R}_i(t)$	Reliability of block i for a mission of length t
C^{mf}	Penalty cost of mission failure
C^{sl}	Penalty cost of system loss
MSP	MSP
SSP	SSP
MSP_0	Required minimum MSP
SSP_0	Required minimum SSP

5.2.2 System Description and MAP

This work investigates the risk control and SM planning for mission-critical systems operating consecutive missions interspersed with scheduled breaks for maintenance. For such systems, failure or malfunction can lead to significant economic losses or even loss of life. To reduce the likelihood of such failures, maintenance actions can be performed during break periods to enhance the reliability of individual system components. Additionally, during the mission, system operators can decide to terminate the mission and initiate a rescue procedure to protect lives or valuable assets when the risks of continuing the mission become too high. The critical decision problem involves selecting the components to maintain, determining the corresponding maintenance levels, and identifying the point at which the primary mission should be aborted while balancing the costs of system failure, mission failure, and maintenance. In this work, mission elapsed time is used to set the mission-abort decision threshold V. In other applications, other criteria such as number of component failures, number of shocks sustained, degradation level can be used to set V.

The general system architecture considered is represented by its corresponding standby RBD composed of $I = 3$ main blocks. Block 1 is in series with a cold stand-by structure composed of block 2 (online/operating), block 3 (reserve), and a reliable (*i.e.,* perfect) sensing and switching component (SSW), as depicted in Figure 5.1. Each block i ($i = 1, 2, 3$) is comprised of a series of N_i subsystems. Each subsystem $j = 1$ ($j = 1, \ldots, N_i$) is composed of N_{ij} components (E_{ijk}, $k = 1, \cdots, N_{ij}$) connected in parallel. Figure 5.1 shows the RBD representing the system and the problem setting. After completing a mission, the system enters a scheduled break period of length D during which maintenance activities can be performed. After the break, the system must perform a mission of length T with mission-abort decision threshold V. At any time during the mission, if block 2 fails, then the sensing and switching component switches from block 2 to block 3 to maintain the operation of the system. If block 2 were to fail at any time $t \leq V$ during the mission, then the mission would be aborted and a rescue procedure immediately initiated (*i.e.,* mission interrupted and system initiates safe shutdown or return to base). Beyond V, if block 2 fails, then the mission continues with blocks 1 and 3. For a rescue to be successful, both blocks 1 and 3 must survive for a duration $\theta(t)$, the length of the rescue procedure initiated at time t. The model proposed in this work aims to jointly optimize maintenance decisions (to be formally introduced next) and the mission-abort decision threshold V to minimize the expected total cost.

At the end of the current mission (*i.e.,* the start of the break), each component E_{ijk} is characterized by its current effective age φ_{ijk}, while its status is determined by the binary state variable u_{ijk} defined

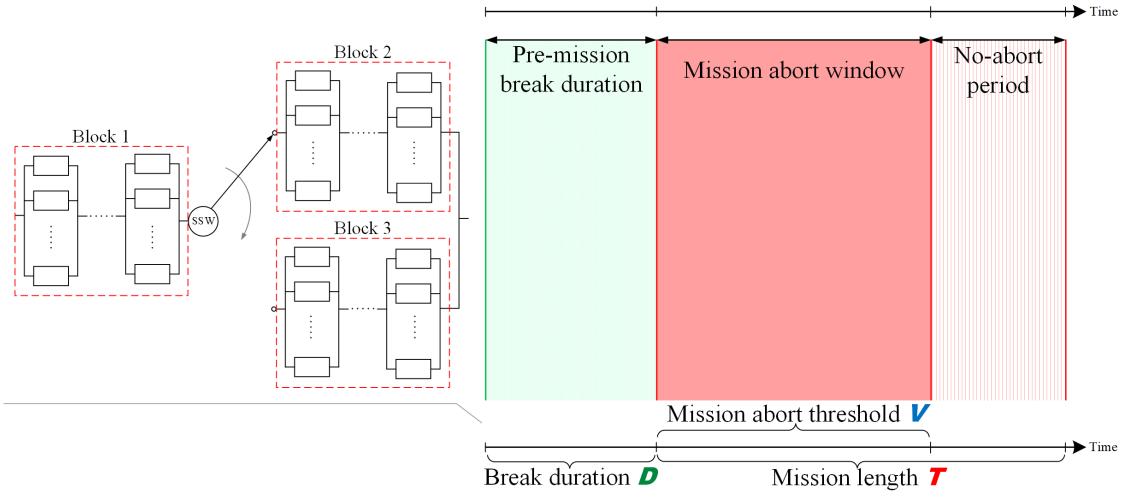


Figure 5.1 – RBD of the system employing a mission-abort policy.

as follows:

$$u_{ijk} = \begin{cases} 1 & \text{if } E_{ijk} \text{ is functioning at the start of the break,} \\ 0 & \text{otherwise.} \end{cases} \tag{5.1}$$

At the end of the break, each component E_{ijk} is also described by its effective age ψ_{ijk}, and its status is given by a binary state variable v_{ijk}.

$$v_{ijk} = \begin{cases} 1 & \text{if } E_{ijk} \text{ is functioning at the end of the break,} \\ 0 & \text{otherwise.} \end{cases} \tag{5.2}$$

By definition of the SMP, maintenance activities can only take place during the intermission break windows. Furthermore, it is assumed that components do not degrade or age during the maintenance break (*i.e.,* the age of a component is operation-dependent). This is a reasonable assumption for most systems whose break durations are much shorter than their mission lengths.

5.3 Reliability Computation and Maintenance Modeling

The optimization approach developed in this chapter aims to minimize the expected total cost while providing a required level of MSP and SSP. To compute MSP and SSP, we first define the conditional reliability of each component E_{ijk} for a duration t given its effective initial age ψ_{ijk}:

$$\mathcal{R}_{ijk}(t|_{\psi_{ijk}}) = \frac{\mathcal{R}_{ijk}(\psi_{ijk}+t)}{\mathcal{R}_{ijk}(\psi_{ijk})}, \tag{5.3}$$

where $\mathcal{R}_{ijk}(t)$ is the unconditional reliability function of component E_{ijk}.

The conditional reliability of each block i is given by (5.4) as its configuration is *series-parallel*:

$$\mathcal{R}_i(t|\mathbf{\Psi}_i) = \prod_{j=1}^{N_i} \left(1 - \prod_{k=1}^{N_{ij}} \left(1 - \mathcal{R}_{ijk}\left(t|_{\psi_{ijk}}\right) \cdot v_{ijk} \right) \right), \tag{5.4}$$

where $\mathbf{\Psi_i}$ is a vector containing the ages of each component in block i after maintenance.

Without loss of generality, the lifetime of a component E_{ijk} is assumed to be governed by a Weibull distribution, with β_{ijk} and η_{ijk} being its shape and scale parameters, respectively. Thus, the unconditional component reliability function is $\mathcal{R}_{ijk}(t) = \exp\left[-\left(\frac{t}{\eta_{ijk}}\right)^{\beta_{ijk}} \right]$.

5.3.1 Computation of the MSP

The mission will be successfully completed if: a) both blocks 1 and 2 survive the mission of length T, or b) block 1 survives the mission while block 2 fails during the mission at time $t > V$ and block 3 is switched on and survives the remaining mission of length $T - t$. The probability $MSP(T, V, \mathbf{\Psi}_1, \mathbf{\Psi}_2, \mathbf{\Psi}_3)$ of successfully completing the mission can then be expressed as follows:

$$\begin{aligned} MSP(T, V, \mathbf{\Psi}_1, \mathbf{\Psi}_2, \mathbf{\Psi}_3) &= \mathcal{R}_1(T|_{\mathbf{\Psi}_1})\mathcal{R}_2(T|_{\mathbf{\Psi}_2}) \\ &\quad + \int_V^T \mathcal{R}_1(t|_{\mathbf{\Psi}_1})\mathcal{R}_1(T-t|_{\mathbf{\Psi}_1+t})\mathcal{R}_3(T-t|_{\mathbf{\Psi}_3}) f_2(t|_{\mathbf{\Psi}_2})\,\mathrm{d}t \\ &= \mathcal{R}_1(T|_{\mathbf{\Psi}_1})\mathcal{R}_2(T|_{\mathbf{\Psi}_2}) \\ &\quad + \mathcal{R}_1(T|_{\mathbf{\Psi}_1}) \int_V^T \mathcal{R}_3(T-t|_{\mathbf{\Psi}_3}) f_2(t|_{\mathbf{\Psi}_2})\,\mathrm{d}t, \end{aligned} \tag{5.5}$$

where $f_2\left(t|_{\mathbf{\Psi}_2}\right)$ is the conditional lifetime probability density function of block 2 given that the initial ages of its components are $\mathbf{\Psi}_2$.

5.3.2 Computation of the SSP

The SSP (SSP) can be computed by considering the probability of two mutually exclusive events. The first event being that the mission is successfully completed with a probability given by Equation (5.5). The second event, whose probability will be denoted as $RSP(T, \mathbf{\Psi}_1, \mathbf{\Psi}_2, \mathbf{\Psi}_3)$, is the successful initiation and completion of the rescue procedure. The initiation of the rescue procedure is successful if block 2 fails at time $t \leq V$ while block 1 remains operational up to that time t. Hence, the probability of successfully initiating the rescue procedure at time t can be calculated as $\mathcal{R}_1(t|_{\mathbf{\Psi}_1}) \cdot f_2(t|_{\mathbf{\Psi}_2})\,\mathrm{d}t$. Given that the rescue mission has a duration of $\theta(t)$, the probability of both successfully initiating the rescue procedure at time t and surviving for an additional $\theta(t)$ is given by $\mathcal{R}_1(\theta(t)|_{\mathbf{\Psi}_1+t}) \cdot \mathcal{R}_1(t|_{\mathbf{\Psi}_1}) \cdot \mathcal{R}_3(\theta(t)|_{\mathbf{\Psi}_3}) \cdot f_2(t|_{\mathbf{\Psi}_2})\,\mathrm{d}t$. Integrating over the interval $[0, V]$ yields the following expression for the probability of successfully aborting the mission and rescuing the system.

$$\begin{aligned} RSP(V, \mathbf{\Psi}_1, \mathbf{\Psi}_2, \mathbf{\Psi}_3) &= \int_0^V \mathcal{R}_1(\theta(t)|_{\mathbf{\Psi}_1+t})\mathcal{R}_1(t|_{\mathbf{\Psi}_1})\mathcal{R}_3(\theta(t)|_{\mathbf{\Psi}_3}) f_2(t|_{\mathbf{\Psi}_2})\,\mathrm{d}t \\ &= \int_0^V \mathcal{R}_1(\theta(t)+t|_{\mathbf{\Psi}_1})\mathcal{R}_3(\theta(t)|_{\mathbf{\Psi}_3}) f_2(t|_{\mathbf{\Psi}_2})\,\mathrm{d}t. \end{aligned} \tag{5.6}$$

Finally, since the successful mission completion and system rescue are mutually exclusive, the SSP is then given by:

$$SSP(T, V, \mathbf{\Psi}_1, \mathbf{\Psi}_2, \mathbf{\Psi}_3) = MSP(T, V, \mathbf{\Psi}_1, \mathbf{\Psi}_2, \mathbf{\Psi}_3) + RSP(V, \mathbf{\Psi}_1, \mathbf{\Psi}_2, \mathbf{\Psi}_3). \tag{5.7}$$

Both MSP and SSP require the conditional probability density function (*pdf*) of block 2. To derive this expression, first recall that the conditional reliability of block 2 is given as:

$$\begin{aligned} \mathcal{R}_2(t|\mathbf{\Psi}_i) &= \prod_{j=1}^{N_2} \left(1 - \prod_{k=1}^{N_{2j}} \left(1 - \mathcal{R}_{2jk}(t|_{\psi_{2jk}}) \cdot v_{ijk}\right)\right) \\ &= \prod_{j=1}^{N_2} \mathcal{R}_{2j}(t|_{\mathbf{\Psi}_2}) \end{aligned} \tag{5.8}$$

where $\mathcal{R}_{2j}(t|_{\mathbf{\Psi}_2}) = \left(1 - \prod_{k=1}^{N_{2j}} \left(1 - \mathcal{R}_{2jk}(t|_{\psi_{2jk}})\right)\right)$ is the conditional reliability of the jth subsystem of block 2. Note that parameters v_{2jk} of components status are dropped for simplicity of notation and avoid cumbersome formulae. Applying the logarithm on each side of Equation (5.8), and computing the first derivative we get:

$$\frac{1}{\mathcal{R}_2(t|\mathbf{\Psi}_2)} \times \frac{\mathrm{d}\mathcal{R}_2(t|_{\mathbf{\Psi}_2})}{\mathrm{d}t} = \sum_{j=1}^{N_2} \frac{1}{\mathcal{R}_{2j}(t|_{\mathbf{\Psi}_2})} \times \frac{\mathrm{d}\mathcal{R}_{2j}(t|_{\mathbf{\Psi}_2})}{\mathrm{d}t}, \tag{5.9}$$

which is equivalent to:

$$f_2(t|_{\mathbf{\Psi}_2}) = \mathcal{R}_2(t|_{\mathbf{\Psi}_2}) \sum_{j=1}^{N_2} \frac{f_{2j}(t|_{\mathbf{\Psi}_2})}{\mathcal{R}_{2j}(t|_{\mathbf{\Psi}_2})}. \tag{5.10}$$

In (5.10) above, $f_{2j}(t|_{\mathbf{\Psi}_2})$ is the conditional *pdf* of the jth subsystem of block 2. To compute this, recall that

$$\begin{aligned} \mathcal{R}_{2j}(t|_{\mathbf{\Psi}_2}) &= 1 - \mathcal{F}_{2j}(t|_{\mathbf{\Psi}_2}) \\ &= 1 - \prod_{k=1}^{N_{2j}} \mathcal{F}_{2jk}(t|_{\psi_{2jk}}), \end{aligned} \tag{5.11}$$

where $\mathcal{F}_{2j}(t|_{\mathbf{\Psi}_2}) = \prod_{k=1}^{N_{2j}} \mathcal{F}_{2jk}(t|_{\psi_{2jk}})$ and $f_{2jk}(t|_{\psi_{2jk}})$ are respectively the conditional CDF and *pdf* of component E_{2jk}. Applying the logarithm on each side of $\mathcal{F}_{2j}(t|_{\mathbf{\Psi}_2}) = \prod_{j=1}^{N_2} \mathcal{F}_{2jk}(t|_{\psi_{2jk}})$, and computing the first derivative we get:

$$f_{2j}(t|_{\mathbf{\Psi}_2}) = \mathcal{F}_{2j}(t|_{\mathbf{\Psi}_2}) \sum_{k=1}^{N_{2j}} \frac{f_{2jk}(t|_{\psi_{2jk}})}{\mathcal{F}_{2jk}(t|_{\mathbf{\Psi}_{2jk}})}. \tag{5.12}$$

Substituting (5.12) into 5.10, yields:

$$f_2(t|_{\mathbf{\Psi}_2}) = \mathcal{R}_2(t|_{\mathbf{\Psi}_2}) \sum_{j=1}^{N_2} \frac{F_{2j}(t|_{\mathbf{\Psi_2}})}{\mathcal{R}_{2j}(t|_{\mathbf{\Psi}_2})} \sum_{k=1}^{N_{2j}} \frac{f_{2jk}(t|_{\psi_{2jk}})}{1 - \mathcal{R}_{2jk}(t|_{\psi_{2jk}})} \tag{5.13}$$

5.3.3 Maintenance Modeling

During the intermission break, each failed component E_{ijk} with age φ_{ijk} can be subjected to a CM of level $l \in \{0, \ldots, L_{ijk}\}$. The lowest maintenance level $l = 0$ corresponds to the DN case, while the highest level $l = L_{ijk}$ corresponds to the perfect corrective repair or AGAN case, and level $l = 1$ corresponds to mR, which brings the component to an ABAO condition. Intermediate values of l $(1 < l < L_{ijk})$ represent IM actions that bring the component back to a condition between AGAN and ABAO. The age reduction approach is used to model IM Malik (1979). Accordingly, when a CM level l is performed on E_{ijk}, its age φ_{ijk} is reduced by a factor μ_{ijkl}, where $0 \leq \mu_{ijkl} \leq 1$, for a cost c^c_{ijkl} and a duration t^c_{ijkl}.

Similarly, a working component with age φ_{ijk} at the beginning of the break can also be subjected to a PM of level $l \in \{0, 2, \ldots, L_{ijk}\}$. PM levels $l = 0$ and $l = L_{ijk}$ correspond to the DN and perfect preventive maintenance (AGAN) cases, respectively. There is no mR equivalent for a PM, therefore the maintenance option $l = 1$ is not available for functioning components. Intermediate values of l $(0 < l < L_{ijk})$ represent IM actions that rejuvenate the component by reducing its age by a factor $0 \leq \gamma_{ijkl} \leq 1$. If the highest maintenance level $l = L_{ijk}$ is performed on a working component E_{ijk}, then $\gamma_{ijkl} = 0$, meaning that the component's age resets to 0. If the DN $(l = 0)$ option is selected, then $\gamma_{ijkl} = 1$, meaning that the age of the component does not change. Each PM action of level l incurs a cost c^p_{ijkl} and a has duration t^p_{ijkl}.

The effective age ψ_{ijk} of component E_{ijk} at the end of the break can then be expressed as a function of its initial status u_{ijk}, initial age φ_{ijk}, and the maintenance level l performed; thus, yielding:

$$\psi_{ijk} = \varphi_{ijk} \cdot \left(u_{ijk} \cdot \sum_{\substack{l=0 \\ l \neq 1}}^{L_{ijk}} \gamma_{ijkl} \cdot x_{ijkl} + (1 - u_{ijk}) \cdot \sum_{l=1}^{L_{ijk}} \mu_{ijkl} \cdot x_{ijkl} \right), \tag{5.14}$$

where x_{ijkl} is a binary decision variable introduced to model the selection of the maintenance level l for component E_{ijk}:

$$x_{ijkl} = \begin{cases} 1 & \text{if maintenance level } l \text{ is selected for component } E_{ijk}, \\ 0 & \text{otherwise.} \end{cases} \tag{5.15}$$

Similarly, the status v_{ijk} of component E_{ijk} at the end of the maintenance break is computed as:

$$v_{ijk} = u_{ijk} + \sum_{l=1}^{L_{ijk}} (1 - u_{ijk})\, x_{ijkl}. \tag{5.16}$$

5.4 Formulation of the Mathematical Optimization Problem

The mission-critical system under study is assumed to have just completed a mission and is entering a maintenance break. Due to limited resources, maintenance actions can be performed only on a subset of systems components. The goal of the proposed optimization model is then to jointly select the set of components to be maintained, the maintenance levels to be performed on the selected components, and the mission-abort threshold V, such that the total expected cost (TEC) is minimized. TEC is the sum of maintenance, mission failure, and system loss costs. The resulting optimization problem is a MINLP, formulated as follows:

$$\begin{aligned}
\min\ TEC = & \sum_{i=1}^{3}\sum_{j=1}^{N_i}\sum_{k=1}^{N_{ij}} \left(\sum_{l=2}^{L_{ijk}} c^{p}_{ijkl} \cdot u_{ijk} \cdot x_{ijkl} + \sum_{l=1}^{L_{ijk}} c^{c}_{ijkl} \cdot (1 - u_{ijk}) \cdot x_{ijkl} \right) \\
& + C^{mf} \cdot (1 - MSP(T, V, \boldsymbol{\Psi}_1, \boldsymbol{\Psi}_2, \boldsymbol{\Psi}_3)) \\
& + C^{sl} \cdot (1 - SSP(T, V, \boldsymbol{\Psi}_1, \boldsymbol{\Psi}_2, \boldsymbol{\Psi}_3)) && \text{(5.17a)} \\
\text{s.t.} & \\
& MSP(T, V, \boldsymbol{\Psi}_1, \boldsymbol{\Psi}_2, \boldsymbol{\Psi}_3) \geq MSP_0, && \text{(5.17b)} \\
& SSP(T, V, \boldsymbol{\Psi}_1, \boldsymbol{\Psi}_2, \boldsymbol{\Psi}_3) \geq SSP_0 && \text{(5.17c)} \\
& \sum_{i=1}^{3}\sum_{j=1}^{N_i}\sum_{k=1}^{N_{ij}} \left(\sum_{l=2}^{L_{ijk}} t^{p}_{ijkl} \cdot u_{ijk} \cdot x_{ijkl} + \sum_{l=1}^{L_{ijk}} t^{c}_{ijkl} \cdot (1 - u_{ijk}) \cdot x_{ijkl} \right) \leq D && \text{(5.17d)} \\
& \sum_{\substack{l=0 \\ l \neq 1}}^{L_{ijk}} u_{ijk} \cdot x_{ijkl} + \sum_{l=0}^{L_{ijk}} (1 - u_{ijk}) \cdot x_{ijkl} = 1, \qquad \forall i, j, k && \text{(5.17e)} \\
& \psi_{ijk} = \varphi_{ijk} \cdot \left(u_{ijk} \cdot \sum_{\substack{l=0 \\ l \neq 1}}^{L_{ijk}} \gamma_{ijkl} \cdot x_{ijkl} + (1 - u_{ijk}) \cdot \sum_{l=1}^{L_{ijk}} \mu_{ijkl} \cdot x_{ijkl} \right), \qquad \forall i, j, k && \text{(5.17f)} \\
& v_{ijk} = u_{ijk} + \sum_{l=1}^{L_{ijk}} (1 - u_{ijk}) \cdot x_{ijkl}, \qquad \forall i, j, k && \text{(5.17g)} \\
& x_{ijkl} \in \{0, 1\}, \qquad \forall i, j, k, l && \text{(5.17h)} \\
& V \geq 0. && \text{(5.17i)}
\end{aligned}$$

In the above optimization model, the objective function (5.17a) seeks to minimize the expected total cost TEC. Constraints (5.17b) and (5.17c) ensure the MSP and SSP meet their respective required level. Constraint (5.17d) ensures that the maintenance break duration is not exceeded. Constraints (5.17e) guarantee that a single maintenance action is selected for each component and constraints (5.17f) and (5.17g) are used to compute the component effective age and status after maintenance. Lastly, constraints (5.17h) impose binary variable restrictions, while constraint (5.17i) ensures non-negativity of the mission-abort threshold V.

5.4.1 Solution Method

The set of decisions sought by the optimization problem formulated in the previous section includes both SM and mission abortion decisions. This optimization problem is nonlinear due to the system survival and mission success probability formulas. Additionally, the integrals involved in the MSP and SSP calculations are intractable, and a closed-form expression cannot be obtained even for basic reliability structures or component lifetime distributions. The lack of an analytical solution for these integrals, combined with the fact that the SMP belongs to the class of NP-hard problems (Rice, 1999), makes the

optimization problem highly difficult to solve exactly. Thus, this work adopts a metaheuristic approach as a solution method. Metaheuristic-based solution methods have been shown to perform well on similar combinatorial optimization problems, and to provide good-quality solutions within reasonable computational times (Jiang and Liu, 2020b; Liu et al., 2018; Pandey et al., 2013; Zhang et al., 2024a).

The GA is a common and very powerful metaheuristic that has been applied successfully to a wide range of combinatorial optimization problems (Talbi, 2009). GA is inspired by the process of natural selection and is initialized with a random population of N_p individuals or solutions. Each individual S has a fitness value and represents a potential solution to the optimization problem. An individual S in the population is represented by a single solution vector comprised of $N_c + 1$ elements, where $N_c = \sum_{i=1}^{3} \sum_{j=1}^{N_i} N_{ij}$ is the total number of components in the system. The first N_c genes (elements of the solution vector) take values from the sets $\{0, \ldots, L_{ijk}\}$ and indicate the maintenance level to be performed on each component E_{ijk}. The final element of the solution vector takes a continuous value and represents the mission-abort threshold V. The fitness value for each individual S is calculated using a fitness function $f(S)$, which is the sum of the previously defined objective function (5.17a) and additional penalty terms to account for constraint violations:

$$f(S) = f_{\mathrm{O}}(S) + \pi_{\mathrm{D}} f_{\mathrm{D}}(S) + \pi_{\mathrm{MSP}} f_{\mathrm{MSP}}(S) + \pi_{\mathrm{SSP}} f_{\mathrm{SSP}}(S) \tag{5.18}$$

where $f_{\mathrm{O}}(S)$ is the objective function (5.17a), and $f_{\mathrm{D}}(S)$, $f_{\mathrm{MSP}}(S)$, $f_{\mathrm{SSP}}(S)$ are the violations of the break duration (5.17d), MSP (5.17b), and SSP (5.17c) respectively. $\pi_{\mathrm{D}}(S)$, $\pi_{\mathrm{MSP}}(S)$, $\pi_{\mathrm{SSP}}(S)$ are penalty factors corresponding to each constraint violation.

The initial population is generated by assigning random values to elements of the solution vector. The elitism ranking procedure is used, where population individuals are ranked based on an increasing order of their fitness values (*i.e.,* lower fitness values imply better solutions as we are dealing with a minimization problem). The best N_e out of the N_p solutions are selected to move on to the next generation. N_e describes the number of elites in the population.

In addition to the elitism operator, crossover and mutation are key operations performed at each iteration or generation of the GA. Crossover operator promote diversity in the solutions population, while mutation operator helps preventing premature convergence to local optima. In this work, uniform crossover is applied. Accordingly, the first N_c genes in the offspring are chosen randomly from one of the two parents with equal probability. The final gene is computed as a convex combination of the corresponding genes from both parents. After performing a crossover operation, a mutation operation is applied to the offspring. This involves randomly altering a gene in the solution vector with a small probability. A gene representing the maintenance level or the mission-abort threshold V can be slightly perturbed within the allowable range to explore new solutions. The number of children generated in each generation is denoted as N_c. To increase the diversity of individuals in the population, N_n new solutions are randomly generated and injected into the next generation.

5.5 Numerical Experiments

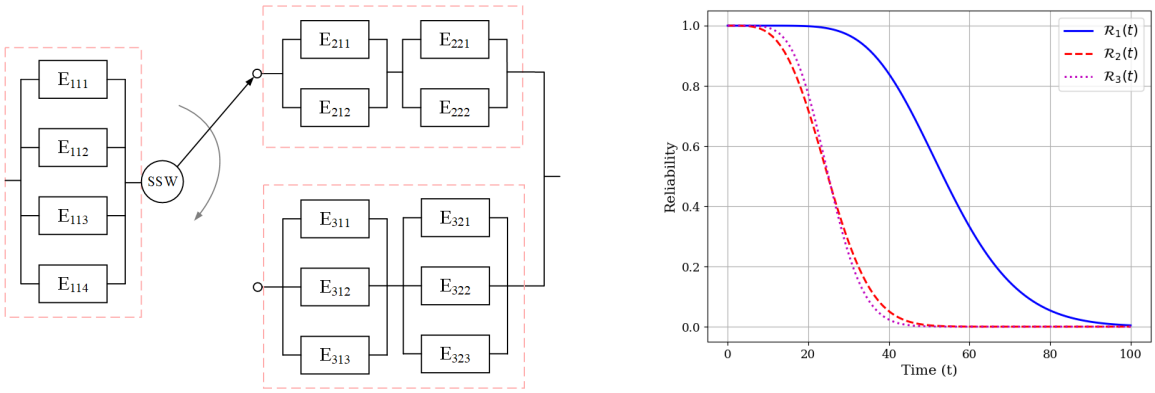
To validate the proposed novel joint SMP and MAP model for mission-critical systems, this section presents four sets of numerical experiments. These experiments also aim at offering insights into key model parameters, and illustrating the efficiency of the solution method. The first set aims to validate the proposed approach and to show how the mission abortion policy is crucial for SMP in mission-critical systems. The other experiments are sensitivity analyses with respect to key parameters of the joint MAP and SMP.

The numerical experiments are conducted on a mission-critical UAV used to perform line inspection missions on high-voltage power grids as described in Yang et al. (2022). Its RBD is composed of three blocks, each composed of identical components, as depicted in Figure 5.2a. The first block is comprised of a single subsystem with four components in parallel. Block 2 is composed of 2 subsystems each consisting of 2 components in parallel, while block 3 is comprised of two subsystems each consisting of three

components in parallel. To simplify notation, monetary and time units are omitted.

For each block i ($i = 1, 2, 3$), the corresponding shape and scale parameters of the Weibull distribution of its components' lifetimes are set to $\beta_{1jk} = 2.1$, $\eta_{1jk} = 40$, $\beta_{2jk} = 2.1$, $\eta_{2jk} = 28$, and $\beta_{3jk} = 2.1$, $\eta_{3jk} = 24$. It should be noted that the shape and scale parameters for individual components are similar to those used in other studies (Diallo et al., 2018) and were selected to ensure that block 1 is the most reliable, and blocks 2 and 3 have comparable reliabilities. When all UAV components are brand new, the reliability curves for the three blocks are plotted in Figure 5.2b.

Once the considered UAV reaches the area of inspection, it can return to the base in Δ time units in rescue mode. Therefore, at any time t, the length of the rescue duration is defined as $\theta(t) = \min\{t, \Delta\}$. In other words, if the current mission time t is below Δ, the rescue duration is equal to t, whereas if the current mission time exceeds Δ, the rescue duration remains at Δ time units as this represents the time for the UAV to return to the base from the location of the inspections. The proposed time-dependent rescue duration is a formalization of rescue duration functions that have been used and studied in the literature to date (Yang et al., 2022).



(a) RBD considered for numerical experiments.

(b) Reliability values for the three blocks.

Figure 5.2 – Description and features of UAV system considered.

For failed components, 5 CM levels (L_{ijk}=4) are available: $l = 0$ for DN, $l = 1$ for mR, $l = 2$ for IM action 1 (IM1) with age reduction coefficient $\mu_3 = 0.70$, $l = 3$ for IM action 2 (IM2) with $\mu_4 = 0.30$, and $l = 4$ for perfect repair (R). There are 4 PM levels for working components: $l = 0$ for DN, $l = 2$ for IM1 with an age reduction coefficient $\gamma_3 = 0.70$, $l = 3$ for IM2 with $\gamma_4 = 0.30$, and $l = 4$ for perfect preventive repair (R). The times and costs to perform the different maintenance levels are provided in Table 5.1. For all experiments, the length of the upcoming mission is set to $T = 24$.

The GA was implemented in Python 3.12. The experiments were conducted on a 12th Gen Intel(R) Core(TM) i7 2.60GHz laptop with 16GB of RAM running Windows 11™.

5.5.1 Experiment 1: Validation

To demonstrate the importance of incorporating mission-abort decisions into the classical SM model, this set of experiments provides a comparison between the results obtained from the proposed joint optimization model (JMA-SMP) and the classical SM optimization approach, hereinafter called "No Mission-Abort SMP" (NMA-SMP). The NMA-SMP corresponds to classical SM models that rely on MSP and have no abortion policy or rescue procedure. In fact, the NMA-SMP is a special case of the proposed JMA-SMP when the mission-abort decision threshold $V = 0$. In this case, the SSP must equal the MSP as there is no opportunity to initiate a system rescue procedure. For this set of experiments, the following parameter values are considered: $D = 50$, $C^{mf} = 2500$, $C^{sl} = 12500$, $MSP_0 = 0.80$,

Table 5.1 – Maintenance data.

	$l=1$ (mR)		$l=2$ (IM1)				$l=3$ (IM2)				$l=4$ (R)			
E_{ijk}	t^c_{ijk1}	c^c_{ijk1}	t^p_{ijk2}	c^p_{ijk2}	t^c_{ijk2}	c^c_{ijk2}	t^p_{ijk3}	c^p_{ijk3}	t^c_{ijk3}	c^c_{ijk3}	t^p_{ijk4}	c^p_{ijk4}	t^c_{ijk4}	c^c_{ijk4}
E_{111}	1.75	71.0	2.75	114.0	3.75	149.0	4.25	144.0	5.75	193.0	6.5	175.0	8.25	241.0
E_{112}	1.25	63.0	2.0	113.0	2.75	151.0	3.25	148.0	5.0	204.0	5.0	183.0	6.25	221.0
E_{113}	1.25	74.0	2.25	118.0	3.25	161.0	3.5	152.0	4.25	205.0	6.25	199.0	10.0	269.0
E_{114}	1.5	75.0	2.5	116.0	3.0	155.0	3.75	150.0	4.75	188.0	6.75	189.0	9.5	229.0
E_{211}	2.0	63.0	3.25	103.0	4.0	135.0	5.5	138.0	8.0	180.0	8.5	177.0	13.5	233.0
E_{212}	2.0	64.0	3.75	103.0	5.75	139.0	5.75	140.0	9.0	185.0	9.25	173.0	14.0	220.0
E_{221}	1.75	66.0	2.5	118.0	3.5	155.0	4.0	164.0	5.5	205.0	6.75	217.0	10.5	285.0
E_{222}	1.25	66.0	2.0	112.0	2.5	155.0	3.25	146.0	4.75	204.0	5.75	186.0	7.25	228.0
E_{311}	2.25	74.0	3.75	114.0	4.5	137.0	5.75	154.0	7.0	192.0	10.0	213.0	15.5	279.0
E_{312}	1.75	78.0	3.0	117.0	4.5	146.0	4.5	147.0	6.75	188.0	7.5	205.0	11.5	277.0
E_{313}	2.0	60.0	3.0	103.0	4.25	124.0	4.5	126.0	5.5	176.0	6.75	154.0	9.75	191.0
E_{321}	1.5	70.0	2.5	115.0	3.5	157.0	4.0	151.0	5.0	182.0	7.25	207.0	9.0	269.0
E_{322}	2.0	64.0	3.25	111.0	4.75	139.0	5.25	155.0	6.5	197.0	8.75	198.0	12.75	246.0
E_{323}	2.0	57.0	3.5	101.0	4.5	123.0	5.5	137.0	8.25	186.0	8.5	182.0	13.25	252.0

$SSP_0 = 0.90$, $\Delta = 4$ and the initial component ages and states provided in Table 5.2 are used.

Table 5.2 – Initial age and state of components.

	E_{111}	E_{112}	E_{113}	E_{114}	E_{211}	E_{212}	E_{221}	E_{222}	E_{311}	E_{312}	E_{313}	E_{321}	E_{322}	E_{323}
u_{ijk}	1	1	1	0	0	1	0	1	0	0	1	0	1	1
φ_{ijk}	18	18	9	12	15	16	16	10	14	8	11	14	11	16

The JMA-SMP is solved for different values of the mission-abort decision threshold V, and the results obtained are displayed in Figure 5.3. This figure shows how various cost components of the best solution vary with V. Setting $V = 0$ (corresponding to the NMA-SMP) leads to a relatively poor solution. The total expected cost decreases as V increases until a minimum point, after which it begins to rise again. Setting different values of V results in different maintenance plans being selected, as indicated by the varying maintenance costs as V increases. This result emphasizes the importance of jointly optimizing the mission-abort decision threshold V and the maintenance plan. Optimizing both simultaneously, leads to a cost-effective maintenance plan while balancing the trade-offs between MSP and SSP.

Table 5.3 displays the total cost, mission-abort decision threshold, and the resulting reliability values for the JMA-SMP and NMA-SPM. From this table, one may observe that given the significantly higher cost of system loss, the JMA-SMP trades off MSP for higher SSP. When the mission-abort policy is not accounted for, the MSP is higher by 12.3%, while the system loss probability is 4.1% higher, which may lead to undesirable economic losses, environment damages and safety risks. The SM plan suggested by the JMA-SMP, its resulting total maintenance time (TMT) and cost (TMC) are given in Table 5.4. It can be seen that the two models lead to the selection of unique SM plans, with the NMA-SPM selecting a more expensive and costly SM plan. The JMA-SMP can reduce penalty costs by selecting more maintenance tasks and altering the mission-abort threshold V. In contrast, the NMA-SPM can only reduce penalty costs by selecting more maintenance tasks and higher maintenance levels. This is likely why a more expensive SM plan is selected for the NMA-SMP.

Table 5.3 – Detailed results: case of experiment #1.

	TEC	V	MSP	SSP
JMA-SMP (proposed)	2222	14.29	0.801	0.965
NMA-SMP	2555	0	0.924	0.924

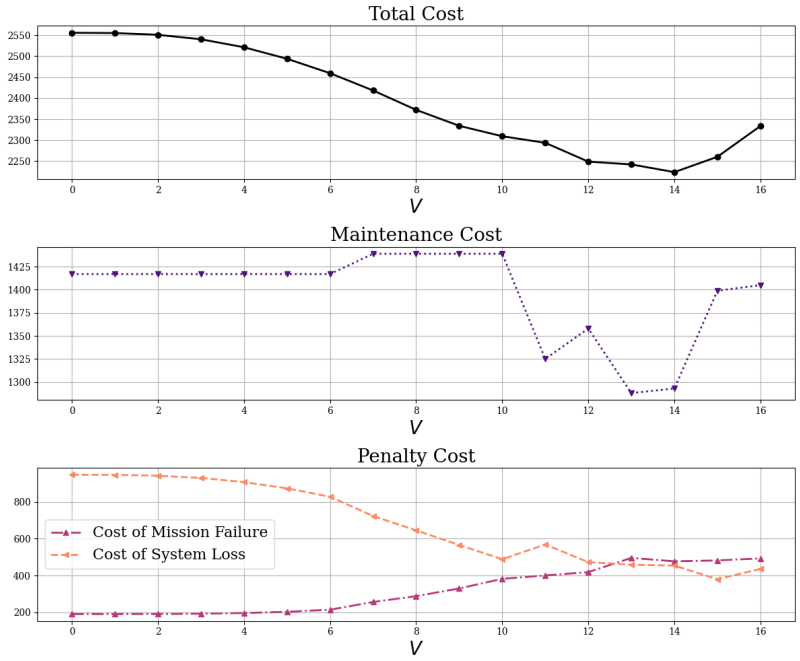


Figure 5.3 – Impact of V on expected costs: case of experiment #1.

Table 5.4 – Detailed maintenance plans: case of experiment #1.

JMA-SMP

Block i	SM plan	TMT	TMC
1	E_{111}–R, E_{112}–R, E_{113}–DN, E_{114}–mR		
2	E_{211}–IM2, E_{212}–R, E_{221}–R, E_{222}–DN	46.25	1293
3	E_{311}–mR, E_{312}–mR, E_{313}–DN, E_{321}–mR, E_{322}–DN, E_{323}–DN		

NMA-SPM

Block i	SM plan	TMT	TMC
1	E_{111}–R E_{112}–DN E_{113}–DN E_{114}–IM2		
2	E_{211}–mR E_{212}–R E_{221}–mR E_{222}–R	49.75	1426
3	E_{311}–mR E_{312}–mR E_{313}–R E_{321}–R E_{322}–DN E_{323}–DN		

5.5.2 Experiment 2: Impact of Initial System Age

This set of experiments aims to investigate the impact of system age on maintenance and mission-abort decision-making. The relevant parameters used for this experiment are defined as follows: $\Delta = 8$, $D = 100$, $C^{mf} = 2000$, and all components have the same age A_0 and still functioning at the start of the maintenance break (*i.e.,* $u_{ijk} = 1$, $\varphi_{ijk} = A_0$, $\forall i, j, k$). For this analysis, the optimization problem is relaxed by setting the threshold constraints (5.17b) and (5.17c) related respectively to mission success and system survivability probabilities to 0 (*i.e.,* the optimization model can freely choose MSP and SSP with no minimum values). The problem is then solved for different values of the MRUL of the system and the penalty cost C^{sl}. The MRUL is defined mathematically as follows:

$$\text{MRUL} = \frac{1}{\mathcal{R}_s(A_0)} \int_{A_0}^{\infty} \mathcal{R}_s(t)\, dt, \tag{5.19}$$

where $\mathcal{R}_s(t)$ is the system reliability function and A_0 is to the current age of the system. Since the system consists of block 1 in series with blocks 2 and 3 in cold standby, the system reliability expression is defined as (Kuo and Zuo, 2003):

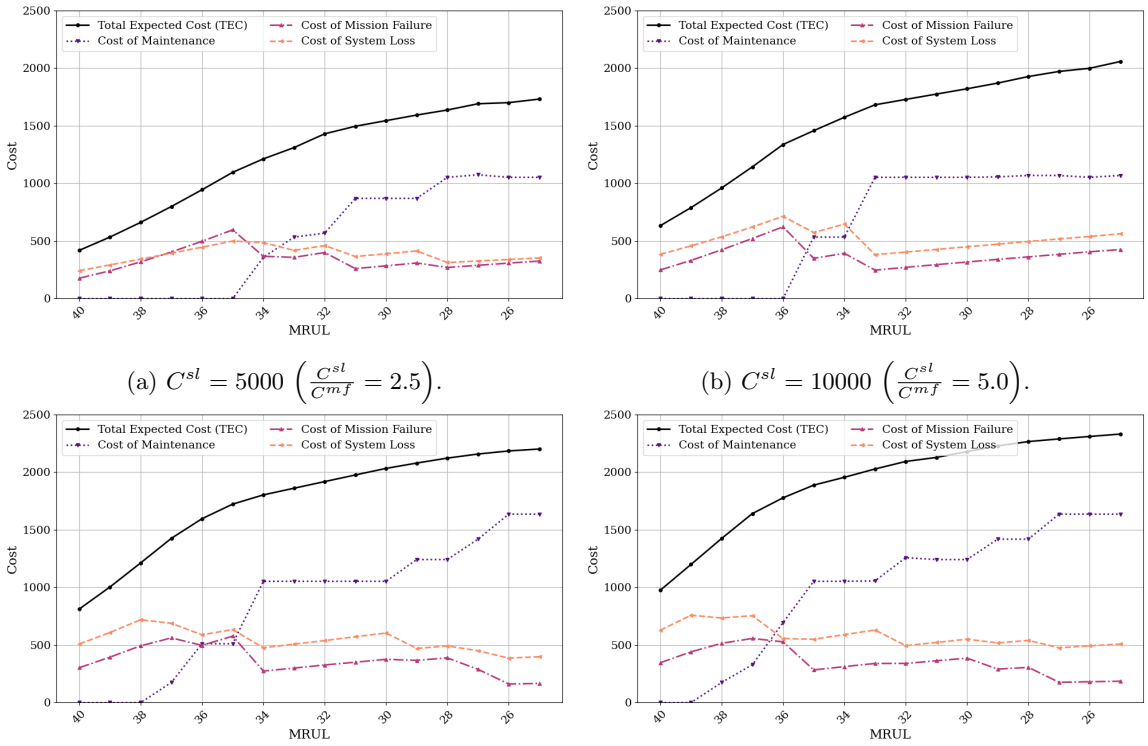
$$\mathcal{R}_s(t) = \mathcal{R}_1(t) \left(\mathcal{R}_2(t) + \int_0^t \mathcal{R}_3(t-\tau)\, f_2(\tau)\, d\tau \right). \tag{5.20}$$

Figures 5.4 – 5.8 display the different cost components and the maintenance time as the system MRUL and penalty cost C^{sl} are varied. The following key observations can be made. From Figure 5.4, we observe an increase in the expected total cost as the MRUL decreases. This is due to the fact that lower values of MRUL imply older system components, which leads to higher maintenance costs and/or increased penalty costs due to reduced MSP and SSP. From Figure 5.4, one may also observe that higher system loss costs C^{sl} lead to maintenance actions being performed on even relatively newer systems (*i.e.,* systems that have high MRUL). This result is mainly due to the fact that as the system loss becomes costly, more maintenance actions are carried out to increase the system survivability probability. For lower values of C^{sl}, the cost of maintaining newer systems outweighs the cost benefit gained from improving SSP. However, for higher values of C^{sl}, improving SSP can lead to significant cost savings.

The JMA-SMP is also solved for varying values of the system loss cost C^{sl}, the MRUL and the mission failure cost C^{mf}. The results obtained are reported in Figure 5.5 and Table 5.5. Figure 5.5 displays the different cost components for fixed values of MRUL and C^{mf}, while Table 5.5 provides for each block i ($i = 1, 2, 3$) the detailed SM plans in addition to the corresponding total maintenance time (TMT_i) and cost (TMC_i). This table gives also the mission-abort threshold V, and the resulting best MSP and SSP values. From these results, the following insights can be made.

- The maintenance cost does not necessarily increase as the penalty cost C^{sl} rises. Indeed, the model aims to balance maintenance and penalty costs by adjusting both maintenance decisions and the abort threshold V. For higher values of C^{sl}, the best solution may involve maintaining or even reducing maintenance expenses and accepting higher costs from mission failures and system losses as a trade-off.
- Increasing maintenance spending will not result in a decrease of the penalty costs. Beyond a certain point, the cost of increased maintenance actions may not be justified due to diminishing returns. As a result, the model may reduce maintenance costs and adjust the abort threshold V to manage penalty costs effectively.
- Generally, we observe a non-decreasing trend in mission failure cost as C^{sl} is increased. This is likely a result of the increasing trend of V as this penalty cost is increased. Increasing the abort threshold will result in a decrease in MSP and an increase in SSP.
- In the case where $MRUL = 30$, increasing C^{sl} from 10000 to 12000 results in a reduction in V and an increase in maintenance cost. However, the mission failure probability still increases. To understand this, one needs to examine the specific maintenance actions selected, as shown in Table 5.5. Although the overall maintenance cost rises, the amount of maintenance performed on block 2 decreases when C^{sl} increases from 10000 to 12000. Since block 2 is critical for mission success, this reduction in maintenance actions leads to a lower MSP, thereby increasing the mission failure probability. This example highlights how the allocation of maintenance resources to blocks significantly impacts both MSP and SSP.

For varying values of the MRUL and system loss cost C^{sl}, Figure 5.6 displays the total maintenance time TMT_i per block i ($i = 1, 2, 3$), the total maintenance cost TMC_i per block i as well as the mission-abort threshold V. In general, the numerical results obtained show an increase in maintenance time as the MRUL decreases. This result is due to the fact that more maintenance actions are required to improve the aging system components. Another interesting observation is related to the behavior of the mission-abort decision threshold V. While V generally increases as the MRUL decreases, we observe a sudden drop in V at a certain point for all values of C^{sl}. This drop in V coincides with a point where significant maintenance actions are selected. The explanation for this phenomenon requires an understanding of how V impacts the mission success and system survival probabilities. A decrease in V leads to an increase in MSP, but also a decrease in SSP. To compensate for this drop in survival probability, more maintenance actions are carried out on system components. This highlights the model's ability to balance maintenance costs with the expected costs resulting from mission failure and system loss. This trade-off is ensured by



(a) $C^{sl} = 5000$ $\left(\frac{C^{sl}}{C^{mf}} = 2.5\right)$.

(b) $C^{sl} = 10000$ $\left(\frac{C^{sl}}{C^{mf}} = 5.0\right)$.

(c) $C^{sl} = 15000$ $\left(\frac{C^{sl}}{C^{mf}} = 7.5\right)$.

(d) $C^{sl} = 20000$ $\left(\frac{C^{sl}}{C^{mf}} = 10.0\right)$.

Figure 5.4 – Cost components for varying values of MRUL: case of experiment #2.

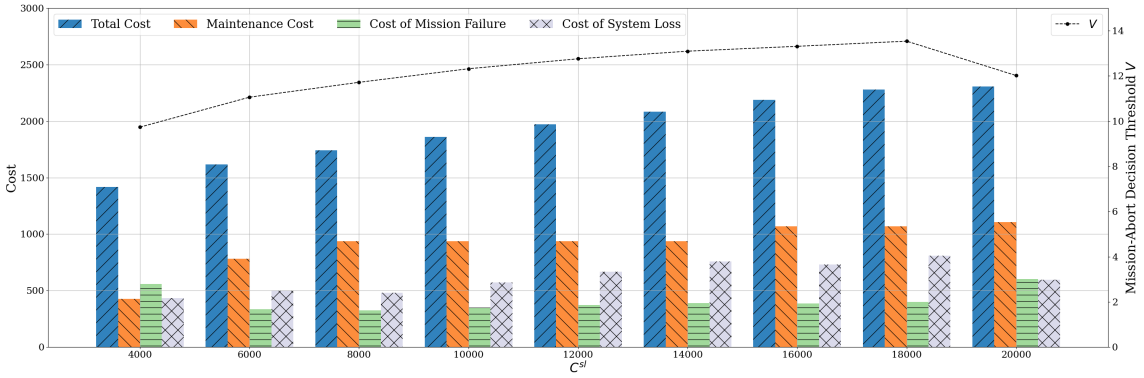
jointly optimizing the mission-abort decision threshold and the SM plan.

Figure 5.7 shows how the maintenance and penalty costs behave when the system MRUL and break duration D are varied. Notably, there is an opposite relationship between these two costs. As the break duration increases, the maintenance cost rises, particularly for more degraded systems (*i.e.,* those with lower MRUL). However, this increase in maintenance costs results in an increase in system reliability, which in turn leads to substantial reductions in penalty costs, ultimately lowering the expected total cost. It is also important to note that systems with lower MRUL incur higher total expected costs.

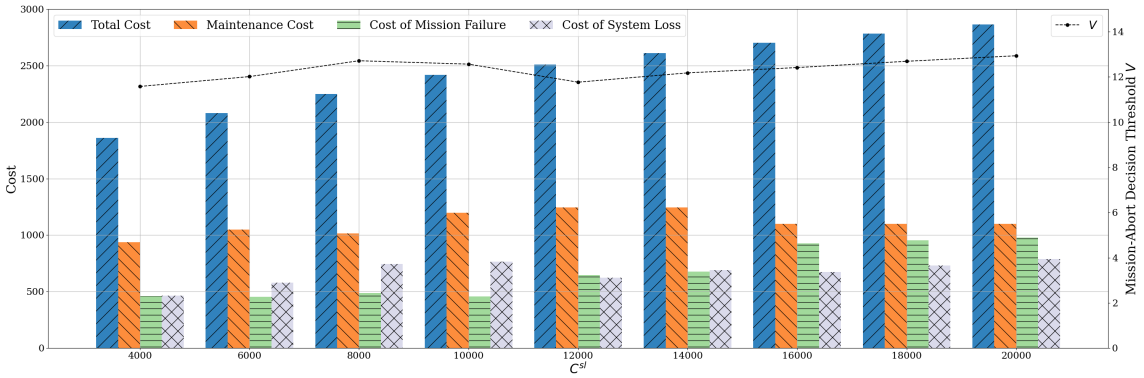
For different values of MRUL, Figure 5.8 displays the total maintenance time (TMT_i), along with the total maintenance time per block i and the mission-abort decision threshold V. From this figure, the following observations can be made. For systems with relatively high MRUL, there is a point at which extending the break duration no longer affects the selected maintenance plan or mission-abort threshold. At this stage, the break duration is no longer a binding constraint, as performing additional or higher levels of maintenance is not cost-effective. In general, a decreasing trend in V is observed as the break duration increases.

5.5.3 Experiment 3: Impact of Reliability Thresholds

This set of experiments aims to study the impact of target reliability levels on the mission-abort decision threshold and SM plan. The initial component ages and states used are identical to those displayed in Table 5.2, and the other relevant parameters are $\Delta = 8$, $D = 100$, $C^{mf} = 2500$, and $C^{sl} = 12500$. To study the impact of reliability thresholds MSP_0 and SSP_0, the proposed optimization model is solved by setting one of the thresholds to 0 (relaxed) while varying the other. Figure 5.9 displays the results when $SSP_0 = 0$ and MSP_0 is varied. It shows that increasing the MSP_0 results in an increase in the



(a) $MRUL = 35$ and $C^{mf} = 2000$.



(b) $MRUL = 30$ and $C^{mf} = 2000$.

Figure 5.5 – Cost breakdown for varying values of C^{sl}: case of experiment #2.

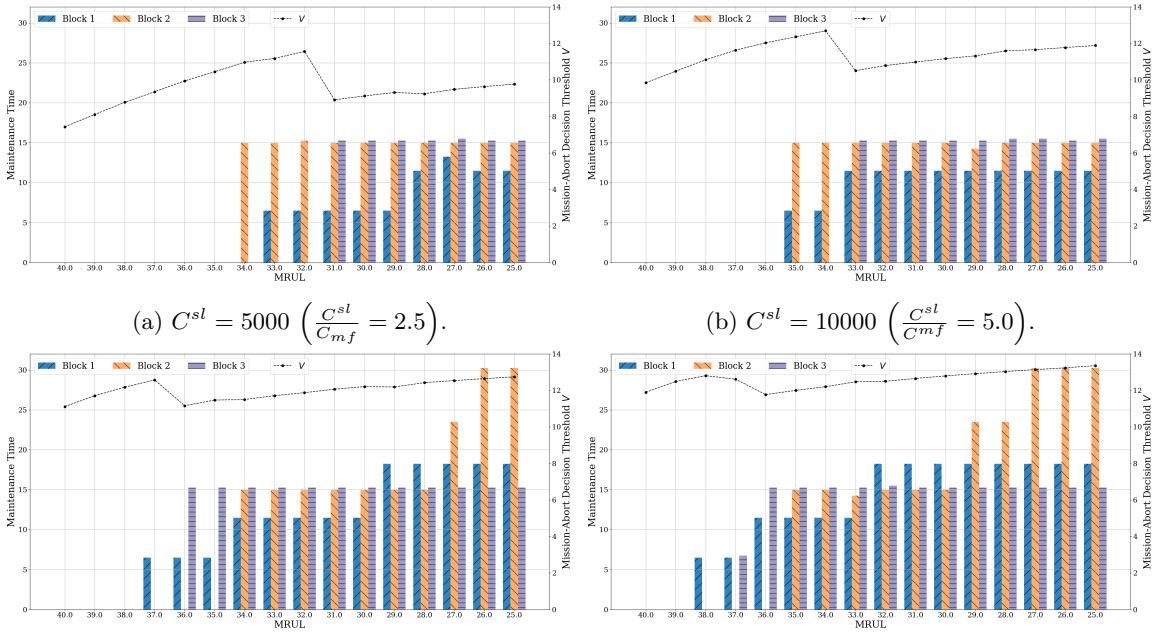
Table 5.5 – Detailed SM plans for MRUL=35 and $C^{mf} = 2000$: case of experiment #2.

$C^{sl} = 10000$

Block i	SM plan	TMT_i	TMC_i	V	MSP	SSP
1	E_{111}-DN E_{112}-R E_{113}-DN E_{114}-mR	6.50	258			
2	E_{211}-mR E_{212}-R E_{221}-mR E_{222}-R	18.75	488	12.57	0.772	0.923
3	E_{311}-IM2 E_{312}-mR E_{313}-DN E_{321}-IM2 E_{322}-DN E_{323}-DN	13.75	452			

$C^{sl} = 12000$

Block i	SM plan	TMT_i	TMC_i	V	MSP	RSP
1	E_{111}-DN E_{112}-R E_{113}-DN E_{114}-IM2	9.75	371			
2	E_{211}-mR E_{212}-DN E_{221}-mR E_{222}-R	9.50	315	11.77	0.677	0.948
3	E_{311}-mR E_{312}-mR E_{313}-R E_{321}-mR E_{322}-DN E_{323}-R	20.75	558			



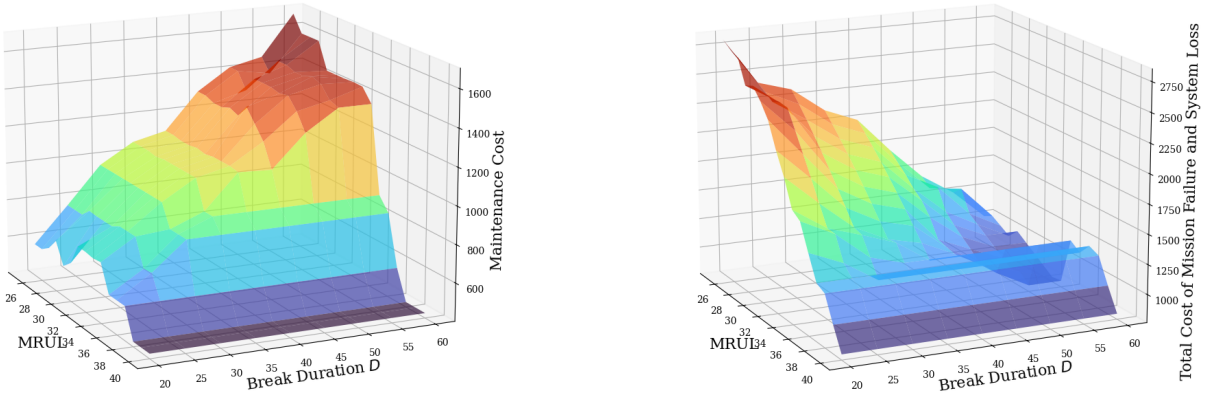
(a) $C^{sl} = 5000$ $\left(\frac{C^{sl}}{C^{mf}} = 2.5\right)$.

(b) $C^{sl} = 10000$ $\left(\frac{C^{sl}}{C^{mf}} = 5.0\right)$.

(c) $C^{sl} = 15000$ $\left(\frac{C^{sl}}{C^{mf}} = 7.5\right)$.

(d) $C^{sl} = 20000$ $\left(\frac{C^{sl}}{C^{mf}} = 10.0\right)$.

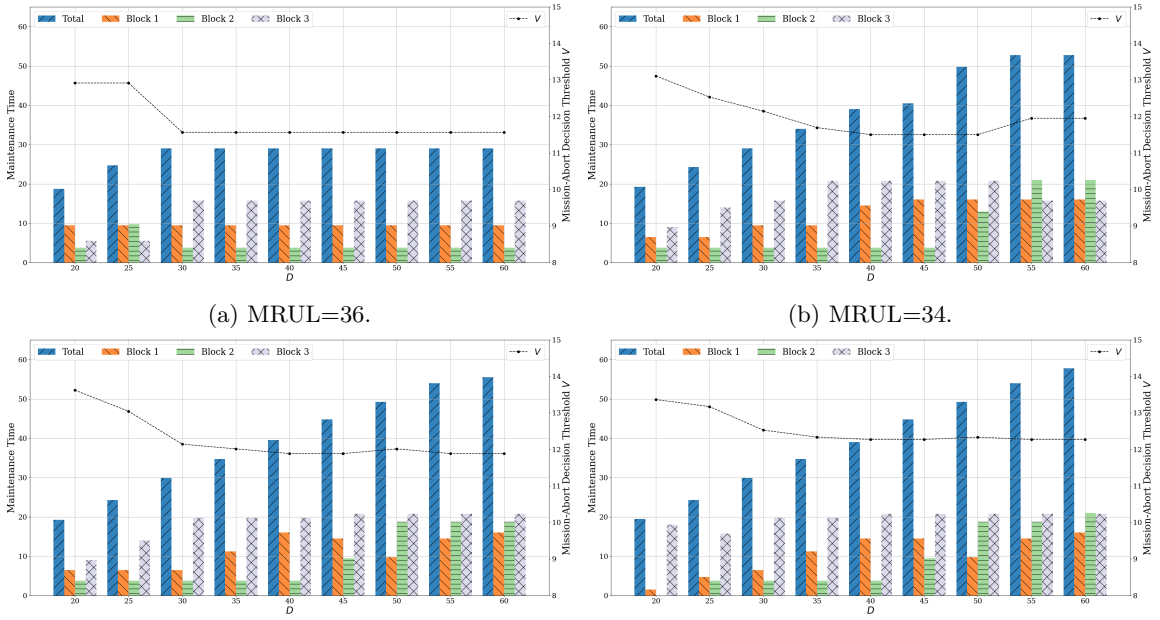
Figure 5.6 – V and total maintenance time per block for varying values of MRUL: case of experiment #2.



(a) Maintenance cost.

(b) Penalty cost.

Figure 5.7 – Maintenance and penalty costs for varying MRUL and D ($C^{mf} = 2000$, $C^{sl} = 15000$): case of experiment #2.



(a) MRUL=36.

(b) MRUL=34.

(c) MRUL=32.

(d) MRUL=30.

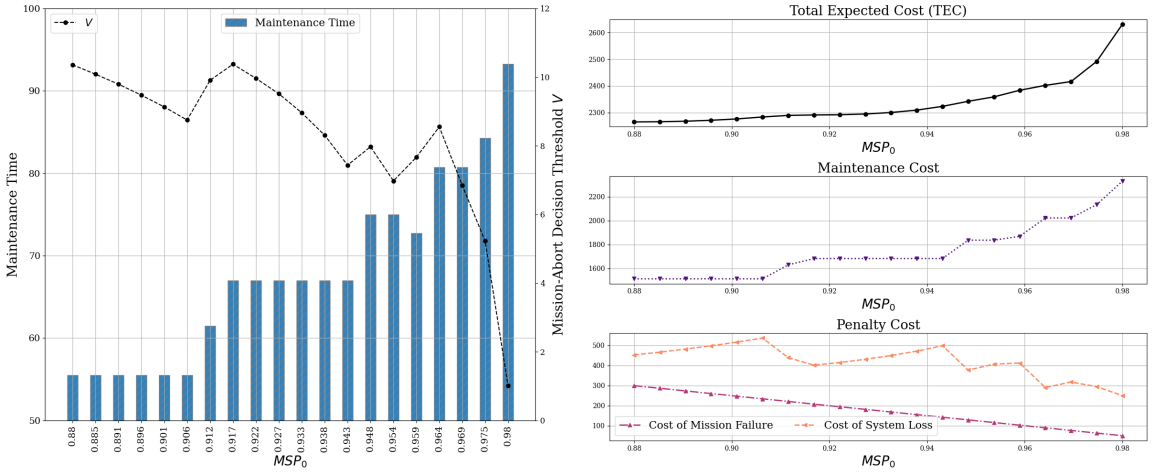
Figure 5.8 – V and total maintenance time for varying D ($C^{mf} = 2000$, $C^{sl} = 15000$): case of experiment #2.

total expected cost. This cost increases drastically for high values of MSP_0, while its increase is much less pronounced for smaller values. Although increasing the mission reliability threshold results in an increase in MSP and a corresponding decrease in the penalty cost due to mission failure, the cost of performing the necessary maintenance tasks to achieve this high level of MSP outweighs the benefits, leading to an overall increase in expected total cost. Increasing MSP_0 from 0.880 to 0.980 resulted in a 16.2% increase in the total expected cost and a 68.0% increase in the total maintenance time. Another interesting observation that can be made is the decreasing trend in V as the mission reliability threshold increases. Reducing the value of V decreases the chance of rescuing the system but increases the likelihood of mission success. Therefore, it is expected that setting exceptionally high mission success thresholds results in the selection of smaller values of V. However, V is not a strictly decreasing function in threshold MSP_0. The MSP can be improved not only by decreasing V but also by performing additional maintenance actions. As can be seen in Figure 5.9a, points where V grows abruptly coincide with instances where more maintenance actions are performed.

The results obtained when $MSP_0 = 0$ and SSP_0 is varied are depicted in Figure 5.10. Similar to the previous case discussed above, one may also observe that increasing SSP_0 results in an increase in the total expected cost. Furthermore, by increasing SSP_0, the overall SSP increases while the penalty cost due to system failure decreases. However, the cost of performing the necessary maintenance tasks to achieve this high level of SSP leads naturally to an overall increase in the total expected cost. Increasing SSP_0 from 0.950 to 0.990 resulted in a 19.8% increase in the total expected cost and a 63.1% increase in the total maintenance time. In general, the penalty cost due to mission failure remains relatively constant for increasing values of SSP_0. In contrast to the previous case, where the mission reliability threshold MSP_0 was varied, the results in Figure 5.10 show an increasing trend in V as the system survival threshold SSP_0 increases. However, V is not a strictly increasing function in SSP_0, and this behavior can be explained similarly to the previous case. This indicates that while higher mission reliability thresholds tend to result in smaller values of V, higher system survival thresholds lead to larger values of V.

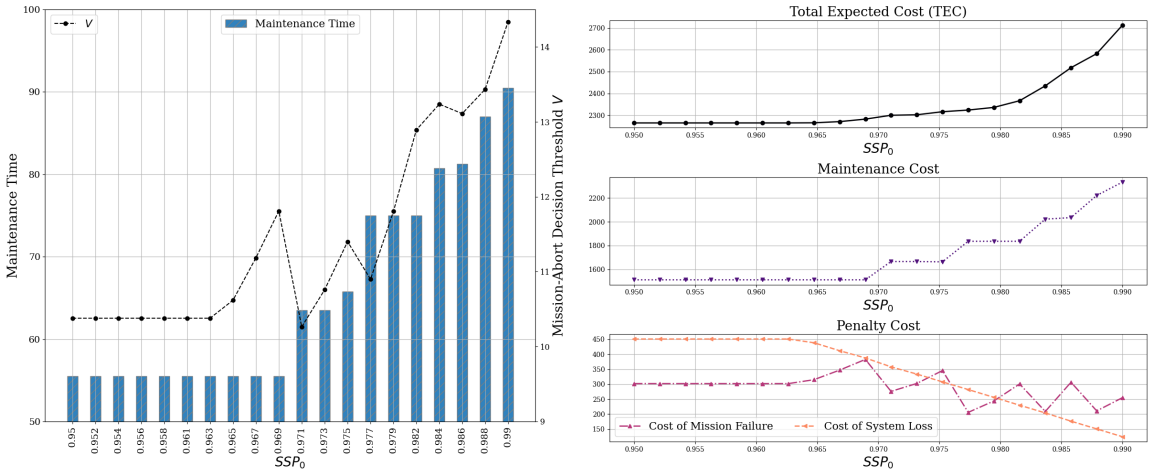
To conclude this set of experiments, it is important to include the minimum reliability threshold constraints (5.17b) and (5.17c) in the formulation of the JMA-SMP as they allow the decision-maker to set certain requirements on the MSP and SSP. If these constraints were relaxed, the optimal solution

might result in such a low MSP that initiating the mission would not be worthwhile. Additionally, a decision-maker may want to set a very high SSP_0 to ensure, with high confidence, that the system will not be lost. A decision-maker can perform a similar analysis to what was done in this experiment to identify a balance between the total expected cost TEC and minimum reliability levels.



(a) V and total maintenance time. (b) Cost components.

Figure 5.9 – Results for varying MSP_0: case of experiment #3.



(a) V and total maintenance time. (b) Cost components.

Figure 5.10 – Results obtained for varying SSP_0: case of experiment #3.

5.5.4 Experiment 4: Impact of the Break and Rescue Durations

This set of experiments investigates the impact of the break duration and parameter Δ on the maintenance and mission-abortion decisions. It also aims to further highlight the importance of incorporating mission-abort decisions into the classical SM model. The following parameter values are used: $C^{mf} = 2500$, $C^{sl} = 12500$, $MSP_0 = 0.80$, $SSP_0 = 0.90$. Additional parameters values can be found in Table 5.2.

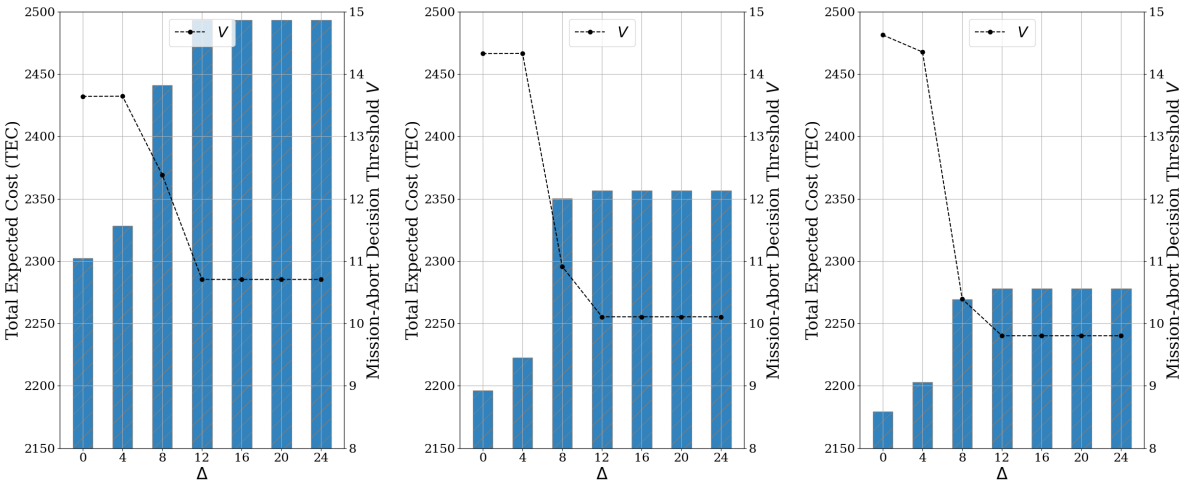
Table 5.6 presents the overall results obtained for the JMA-SMP and NMA-SMP for different values of the break duration D and the rescue duration parameter Δ. Table 5.7 shows the detailed SM plans

for the joint optimization approach, along with the total maintenance time and cost. The results clearly demonstrate the benefits of implementing a mission-abort procedure. Indeed, one may observe that in all instances, the JMA-SMP produces solutions with lower expected total costs. It is important to emphasize that when the break duration $D = 40$, the NMA-SMP fails to produce a maintenance plan that meets the required reliability levels. Additionally, it is interesting to note that in all instances, the NMA-SMP results in higher MSP and lower SSP compared to the JMA-SMP. This is mainly due to the fact that the NMA-SMP is equivalent to setting the mission-abort decision threshold V to 0. These observations confirm the results obtained within the validation example.

Now, focusing on the results obtained from the JMA-SMP, one can observe that increasing Δ seemingly leads to an increase in the expected total cost. Furthermore, when Δ increases, there is also a reduction in the mission-abort decision threshold V. This intuitive result comes from the fact that a longer rescue procedure (due to the increase in Δ) reduces the SSP, and lowering the abort threshold V compensates for this reduction. Figure 5.11 displays the total expected cost TEC and mission-abort decision threshold V for varying values of Δ and break durations. The results indicate that increasing the rescue duration Δ raises the total cost up to a certain point, beyond which further increases in Δ have no impact on the SM plan, the mission-abort threshold, or total expected cost. For relatively high values of Δ ($\Delta \geq 12$ in this experiments), the model selects $V \leq \Delta$, implying $\theta(t) = t$, meaning that the rescue duration will always equal the elapsed mission time at the failure of block 2. Notably, the same maintenance and mission-abort strategy remains optimal for all values of $\Delta \geq 12$ for all break durations considered. The impact of the break duration can once again be observed. Indeed, increasing the break duration up to a certain point leads to better SM plans in terms of reducing the total expected cost. Longer break durations allow for more maintenance to be completed as shown in Table 5.7, resulting in increases in both MSP and SSP.

Table 5.6 – Comparison of JMA-SMP and NMA-SMP: case of experiment #4.

		$\Delta = 4$				$\Delta = 8$				$\Delta = 12$			
		C	V	MSP	SSP	C	V	MSP	SSP	C	V	MSP	SSP
$D = 40$	JMA–SMP	2328	13.65	0.800	0.943	2441	12.38	0.800	0.933	2493	10.71	0.815	0.934
	NMA–SMP	–	–	–	–	–	–	–	–	–	–	–	–
$D = 50$	JMA–SMP	2222	14.33	0.800	0.966	2350	10.91	0.827	0.959	2356	10.11	0.868	0.955
	NMA–SMP	2555	0.00	0.924	0.924	2555	0.00	0.924	0.924	2555	0.00	0.924	0.924
$D = 60$	JMA–SMP	2202	14.35	0.824	0.967	2269	10.39	0.878	0.964	2278	9.80	0.889	0.961
	NMA–SMP	2462	0.00	0.937	0.937	2462	0.00	0.937	0.937	2462	0.00	0.937	0.937



(a) $D = 40$. (b) $D = 50$. (c) $D = 60$.

Figure 5.11 – Results obtained for varying values of Δ: case of experiment #4.

Table 5.7 – Detailed SM plan: case of experiment #4.

$D = 40$

Δ	Block 1 SM plan	Block 2 SM plan	Block 3 SM plan	MT	MC
4	E_{111}-DN E_{112}-R E_{113}-DN E_{114}-mR	E_{211}-IM2 E_{212}-R E_{221}-R E_{222}-DN	E_{311}-mR E_{312}-mR E_{313}-DN E_{321}-mR E_{322}-DN E_{323}-DN	39.75	1118
8	E_{111}-R E_{112}-DN E_{113}-DN E_{114}-mR	E_{211}-mR E_{212}-R E_{221}-R E_{222}-DN	E_{311}-mR E_{312}-mR E_{313}-DN E_{321}-IM2 E_{322}-DN E_{323}-DN	38.75	1105
12	E_{111}-R E_{112}-R E_{113}-DN E_{114}-mR	E_{211}-mR E_{212}-R E_{221}-IM2 E_{222}-DN	E_{311}-mR E_{312}-mR E_{313}-DN E_{321}-IM2 E_{322}-DN E_{323}-DN	38.75	1208

$D = 50$

Δ	Block 1 SM plan	Block 2 SM plan	Block 3 SM plan	MT	MC
4	E_{111}-R E_{112}-R E_{113}-DN E_{114}-mR	E_{211}-IM2 E_{212}-R E_{221}-R E_{222}-DN	E_{311}-mR E_{312}-mR E_{313}-DN E_{321}-mR E_{322}-DN E_{323}-DN	46.25	1293
8	E_{111}-R E_{112}-R E_{113}-DN E_{114}-mR	E_{211}-mR E_{212}-R E_{221}-IM2 E_{222}-DN	E_{311}-mR E_{312}-mR E_{313}-IM2 E_{321}-mR E_{322}-DN E_{323}-R	48.25	1404
12	E_{111}-R E_{112}-R E_{113}-DN E_{114}-mR	E_{211}-mR E_{212}-R E_{221}-mR E_{222}-R	E_{311}-IM2 E_{312}-mR E_{313}-DN E_{321}-R E_{322}-DN E_{323}-DN	49.5	1460

$D = 60$

Δ	Block 1 SM plan	Block 2 SM plan	Block 3 SM plan	MT	MC
4	E_{111}-R E_{112}-R E_{113}-DN E_{114}-mR	E_{211}-R E_{212}-R E_{221}-R E_{222}-DN	E_{311}-mR E_{312}-mR E_{313}-DN E_{321}-mR E_{322}-DN E_{323}-DN	51.75	1346
8	E_{111}-R E_{112}-R E_{113}-DN E_{114}-mR	E_{211}-R E_{212}-DN E_{221}-R E_{222}-DN	E_{311}-mR E_{312}-mR E_{313}-R E_{321}-mR E_{322}-DN E_{323}-R	57.75	1509
12	E_{111}-R E_{112}-R E_{113}-DN E_{114}-mR	E_{211}-R E_{212}-DN E_{221}-R E_{222}-DN	E_{311}-mR E_{312}-mR E_{313}-R E_{321}-mR E_{322}-DN E_{323}-R	57.75	1509

5.5.5 Summary of Insights

From a research point of view, this is the first work to integrate SM with mission-abort decision-making. Previous SMP models have focused solely on maximizing MSP, but for many mission-critical systems, implementing a mission-abort procedure can improve system survivability (SSP) and reduce total expected cost (TEC). In this work, we allow mission abortion to initiate a rescue procedure, which results in increased system survivability. A general standby system structure is considered, with mission-abort triggered based on elapsed mission time at the point of mission failure. Novel expressions have been derived for the MSP and SSP for the general system under consideration.

From the industrial and applications point of view, the proposed model can substantially reduce TEC while ensuring higher system survivability, which is valuable for mission-critical systems. The experimentation detailed above conclusively demonstrate the superiority of the proposed model compared to classical SMP models which only consider MSP. The sensitivity analysis performed highlighted key properties of the proposed optimization model and revealed insights important for decision-makers. The impact of the initial system condition, target reliability thresholds, break duration, and rescue procedures was thoroughly investigated, analyzed and fully discussed. Important key findings revealed that extending the break duration usually allowed for improvements in TEC, however, for fairly new systems there may be little to no benefit in extending the break duration. Increasing MSP_0 and SSP_0 results in an increase in TEC, but in some cases, this increase is minimal, meaning that mission success and survivability probabilities can be improved with only slight cost increases. The analysis also demonstrated the interaction between V and the SM plan, as well as how these decisions impact MSP and SSP. Finally, it was shown that increasing the mission rescue duration parameter Δ leads to a rise in total expected cost. Beyond a certain point, however, increasing Δ results in no further increase in cost or change in the maintenance plan and mission-abort decisions.

5.6 Conclusions and Future Research Directions

This chapter introduced an innovative model aiming to jointly optimize mission-abort decisions and maintenance scheduling for mission-critical systems. The 3-block system considered undergoes maintenance during intermission breaks. During a mission, if a key block fails before a mission-abort decision threshold, the mission is aborted and a rescue operation launched. If the failure occurs after the decision threshold, then the mission is continued. A GA is designed to solve the resulting mixed integer nonlinear optimization model. Numerous experiments are conducted to validate the proposed joint model and show the benefits of jointly optimizing the abort decision and the maintenance plan. Moreover, it is shown that the inclusion of the mission-abort decision threshold provides a reasonable MSP and a higher system survivability than the case where there is no abort policy.

This research opens numerous promising avenues for future investigation. Beyond the current findings, potential extensions could investigate the case where there are multiple mission-abort decision thresholds: V_1 to abort and return to base or stop production, and V_2 to continue the mission/production with a lower performance target. The abort policy is currently based on elapsed mission time. An interesting extension would be to consider elapsed mission-time and external shocks to the system, and whichever reaches its threshold first would trigger the mission-abort decision. Additionally, investigating the joint optimal design of the mission-critical system, maintenance plan, and mission-abort policy would allow substantial benefits by considering three phases of the system's life. Finally, this work could be significantly improved through the development of an exact and efficient solution method to solve the problem. Finding closed-form expressions for the integrals involved in the reliability computations would allow global optimization solvers to be applied. Exploring the use of integral approximations to reformulate the problem is another interesting approach to improve solution efficiency.

Chapter 6

Conclusion

This PhD thesis contributes to the field of SM both theoretically and in application, addressing critical challenges in real-world maintenance operations and planning. SM plays a vital role in industries with large-scale maintenance operations, such as aerospace, mining, transportation, manufacturing, petrochemical plants, and offshore energy systems. The contributions of this dissertation are structured around four interrelated research themes, each tackling a key limitation of existing SM models.

The first research theme introduces and optimally solves the JSMWRP. Traditional SM models assume that system components are co-located and that repair resources are readily available, an assumption that is often unrealistic in networked systems such as power grids, water distribution networks, and offshore wind farms. Similarly, FSM models often assume that all systems within a fleet are located in a single facility. However, in industries such as maritime shipping, offshore energy production, and military aviation, fleets are geographically dispersed, requiring repair personnel to travel between sites under transportation and resource constraints. The proposed framework simultaneously determines which systems to visit, which components to maintain, the level of maintenance to perform, and the optimal assignment and routing of repair crews, with the objective of optimizing some performance metric while ensuring a required system reliability threshold. This research explores two key applications of the JSMWRP:

1. Offshore wind farm maintenance where service providers must dispatch maintenance crews from an operation and maintenance (O&M) support port to service OWTs. A CG algorithm is developed to solve the resulting optimization problem, utilizing a dynamic programming labeling algorithm for the subproblems.
2. Post-disruption network restoration where joint maintenance and routing decisions are optimized to enhance network resilience following a disruptive event. A Lagrangian relaxation approach is used to decompose the original problem into three manageable subproblems, enabling efficient solution strategies.

A promising direction for future research is to extend the JSMWRP by considering interval availability as a performance metric instead of reliability. In many industries, availability is a more relevant performance indicator, as it directly correlates with key operational metrics such as production and uptime. Additionally, the solution methods proposed for both applications are heuristic-based, meaning they do not guarantee optimality. An important avenue for future work would be the development of exact solution methods to improve solution quality and provide optimality guarantees. One potential direction is to refine the two-phase approach developed by (Diallo et al., 2018), incorporating more intelligent maintenance pattern generation. A final research direction involves incorporating opportunistic maintenance, where accessing one subsystem provides the opportunity to maintain multiple components, thereby optimizing intervention time and resource usage.

With the rise of CPS, IIoT, and ML, predictive maintenance has become increasingly relevant. SM models assume that components follow known lifetime distributions, which fails to leverage real-time sensor data. The second research theme addresses this gap by introducing a predictive SM framework that

integrates DNNs for reliability prediction. The framework explores multiple data-driven approaches for reliability estimation, including probabilistic classifiers, regression models using only point predictions, and regression models with MCD for uncertainty quantification. Compared to traditional statistical methods, the proposed approach enables more accurate reliability predictions, leading to better-informed maintenance decisions and improved system performance. To demonstrate its applicability, this theme presents a case study on a fleet of military aircraft, where some components, such as turbofan engines, are sensor-monitored, while others follow known lifetime distributions. The fleet must execute a set of upcoming missions, each with unique operational demands. Leveraging both data-driven and statistical reliability estimates, an optimization model is developed to maximize the number of successfully executed missions.

An important avenue for future research is domain adaptability, which involves applying an RUL predictive model trained in one domain to a different system using techniques such as transfer learning. This is particularly relevant given the scarcity of labeled run-to-failure data in many industries, making it essential to develop models that can generalize across different operational environments. Additionally, this research has primarily focused on leveraging data-driven methods for accurate parameter predictions, which are then used within traditional operations research (OR) frameworks for decision-making. A promising direction is the direct integration of data-driven decision-making techniques into OR models. Recent advances in reinforcement learning (RL), specifically deep reinforcement learning (DRL), have demonstrated remarkable success in solving large combinatorial optimization problems. Applying RL-based approaches to large-scale maintenance planning for multi-component systems represents a compelling future direction.

SM decisions are typically made without consideration of multiple resource capacities, such as repair personnel, tools, spare parts, and facilities. The third theme of this dissertation extends the classical SM problem by integrating resource constraints and developing a theoretical link between the SMP and the Resource-Constrained Project Scheduling Problem (RCPSP). The proposed formulation considers a fleet of assets and relaxes the assumption that all systems in the fleet operate synchronously and are available for maintenance simultaneously. In real-world applications, fleets consist of multi-component systems with staggered mission schedules, requiring asynchronous maintenance planning. Through several numerical experiments, the proposed model was shown to be more general and versatile in adapting to a broad spectrum of operational constraints and resource scarcity. The benefit of staggering maintenance breaks and spreading mission start/end times on fleet reliability was demonstrated. This research opens several promising directions for future work. Uncertainty in mission schedules, break durations, and maintenance times is a critical factor in maintenance planning. Addressing these uncertainties through stochastic and robust optimization could enhance decision-making. A particularly valuable direction is the development of two-stage optimization models, where the first stage determines maintenance selection and sequencing, while the second stage optimizes task start times under uncertainty. Formulating the problem in this way would allow decision-makers to maximize expected fleet reliability for upcoming missions while avoiding the overly conservative solutions often associated with single-stage or chance-constrained approaches. The current model determines the optimal set of components to maintain, the appropriate maintenance levels for each, and the execution sequence of these tasks, while respecting resource limitations. However, implementing the selected maintenance plan requires sufficient availability of spare parts and other necessary resources. An important direction for future research is the integration of spare parts procurement and inventory management into the optimization model.

The final contribution of this thesis is the integration of MAP into the SM decision-making framework. Traditional SM models focus solely on MSP, or in other words, the probability that the system will be capable of performing the upcoming mission without failure. However, in many real-world situations, the survival of mission-critical systems is more important than the completion of the mission due to the severe consequences of their failures. In such situations, the system operator(s) may choose to abort the primary mission when specific conditions are satisfied, and initiate a rescue procedure. In general, implementing a MAP will lead to a reduction in MSP but an increase in SSP. The fourth theme of this dissertation develops an optimization model that simultaneously determines the components to maintain, the corresponding maintenance levels, and the mission abort policy such that the total expected cost is minimized, which is a function of mission failure, system loss, and maintenance cost. From a research

point of view, this is the first paper to integrate SM with mission-abort decision-making. From the industrial and applications point of view, the proposed model can substantially reduce the total expect cost while ensuring high survivability. In terms of future research directions, potential extensions include investigating mission abort policies considering both internal and external degradation processes, opposed to the current model that only considers elapsed mission time (*i.e.,* internal degradation). The idea of expanding the decision domain to allow reduced system performance in exchange for increased mission success and system survival probability is another compelling research direction. Finally, investigating the joint optimal design of the mission-critical system, maintenance plan, and mission-abort policy would allow substantial benefits by considering three phases of the system's life.

Appendix A

Supplementary Material for Section 2.1

A.1 Sequential optimization approach

When dealing with the sequential optimization approach of the SMP and the OP, the decisions of what maintenance plan to be performed at the level of each OWT and the routing of the different repair crews are made separately. This is equivalent to first solving the SMP for all OWTs with the goal of minimizing cost subject to time constraints and reliability targets, and then solving the routing problem. The maintenance patterns that are selected for OWTs during the SMP phase are passed as input data to the routing optimization process where the decision of which OWT to visit and the path that each repair crew must take are provided. In what follows, the sequential optimization approach of the SMP and crew routing problem is presented.

A.1.1 BIP formulation of the SMP

The goal of the SMP is to select the maintenance plan to be performed within the time and reliability constraints such that the maintenance cost is minimized Rice et al. (1998). To formulate the corresponding optimization problem, the following decision variable is introduced:

$$s_{ijp} = \begin{cases} 1 & \text{if pattern } p \text{ is selected for subsystem } j \text{ in OWT } i, \\ 0 & \text{otherwise.} \end{cases} \tag{A.1}$$

The BIP formulation for the minimization of the overall maintenance cost that must be solved for each OWT i is as follows:

$$\text{Min } \mathcal{Z}_i = \sum_{j \in \mathcal{S}_i} \sum_{p \in \mathcal{P}_{ij}} C_{ijp} \cdot s_{ijp} \tag{A.2}$$

s.t:

$$\sum_{p \in \mathcal{P}_{ij}} s_{ijp} = 1, \quad \forall j \in \mathcal{S}_i \tag{A.3}$$

$$\sum_{j \in \mathcal{S}_i} \sum_{p \in \mathcal{P}_{ij}} T_{ijp} \cdot s_{ijp} \leq T_{0i} \tag{A.4}$$

$$\sum_{j \in \mathcal{S}_i} \sum_{p \in \mathcal{P}_{ij}} \ln(R_{ijp}) \cdot s_{ijp} \geq \ln(R_0) \tag{A.5}$$

$$s_{ijp} \in \{0, 1\}, \quad \forall j, p \tag{A.6}$$

In the above optimization model, constraints (A.3) ensure that a single maintenance pattern is selected for each subsystem j of OWT i. Constraint (A.4) ensures that the maintenance time at the OWT i does not exceed the allotted time window T_{0i} computed as the maximum working time T^{max} minus the sum of travel times $t_{0i} + t_{i,N+1}$ (see Eq. 2.11). Constraint (A.5) ensures that the reliability of an OWT must be greater than or equal to the target reliability. The term on the left-hand side of the inequality is the

linearization of the OWTs reliability as given by Eq. (2.6). Because s_{ijp} is a binary variable, Eq. (2.6) becomes:

$$R_i = \prod_{j \in \mathcal{S}_i} \left(\sum_{p \in \mathcal{P}_{ij}} R_{ijp} \cdot s_{ijp} \right) \tag{A.7}$$

After applying the natural logarithm which is a monotonic function to both sides of the above equation, the result in constraint (A.5) is obtained:

$$\ln(R_i) = \sum_{j \in \mathcal{S}_i} \sum_{p \in \mathcal{P}_{ij}} \ln(R_{ijp}) \cdot s_{ijp} \tag{A.8}$$

A.1.2 MILP formulation of the crew routing problem

Once the SMP has been first solved for each OWT i, the required maintenance time T_i and cost C_i to perform the corresponding SM plan can be obtained using Eqs. (2.8) and (2.9). They can also equivalently be derived as functions of patterns as follows:

$$T_i = \sum_{j \in \mathcal{S}_i} \sum_{p \in \mathcal{P}_{ij}} T_{ijp} \cdot s_{ijp} \tag{A.9}$$

$$C_i = \sum_{j \in \mathcal{S}_i} \sum_{p \in \mathcal{P}_{ij}} C_{ijp} \cdot s_{ijp} \tag{A.10}$$

At the second step of the sequential optimization approach, the output data resulting from the SMP solutions are used to model and solve the orienteering problem. The objective of the vessel routing problem is to find a route for multiple repair vessels starting and ending at the maintenance base (nodes in $\mathcal{Q} = \{0, N+1\}$), such that the sum of the downtime costs, travel costs, and costs associated with not performing maintenance actions within the working time is minimized. To develop the corresponding mathematical formulation, the following decision variable is introduced:

$$x_{ii'r} = \begin{cases} 1 & \text{if arc } (i, i') \text{ is traversed by repair crew } r, \\ 0 & \text{otherwise.} \end{cases} \tag{A.11}$$

In addition to the above decision variable, two other variables O_{ir} and I_{ir} are introduced. O_{ir} tracks the arrival time of repair crew r at the ith OWT. I_{ir} identifies the order in which node i is visited in the travel path of repair crew r (it is the standard decision variable used in the MTZ subtour elimination constraint).

The routing optimization problem is formulated as an MILP as follows. The objective function (A.12) minimizes the total cost resulting from repair crews maintenance activities and transportation over the distributed wind turbine farm. The first term represents the penalty cost incurred when maintenance is not performed on failed or functioning OWTs. The second term represents the increasing penalty as time passes for failed systems. The third term represents the downtime cost associated with performing PM. The fourth term represents the transportation cost, while the fifth term represents the cost of performing the selected maintenance actions. Constraints (A.13) and (A.14) ensure that the route for each service vessel starts and ends at the maintenance base. Constraints (A.15) ensure the connectivity of the route, and constraints (A.16) ensure that each vertex (OWT) is visited at most once by all repair crews. Constraints (A.17) state that arrival times at OWTs are nonnegative. Constraints (A.18) ensure that crews arrive at OWTs on their paths before the maintenance rotation time T^{max} has elapsed. Constraints (A.19) state that repair crews are at the maintenance base ($i = 0$) at time 0. Constraints (A.20) define the arrival times at their first OWT when repair crews leave the maintenance base. Constraints (A.21) define the repair crews' arrival times at subsequent OWTs after completing maintenance and leaving the previous OWT. $M_{ii'}$ are real numbers of high values. Inequalities (A.22) and (A.23) are the standard MTZ subtour elimination constraints. Constraints (A.24) are binary variables restrictions.

$$\text{Min } \mathcal{Z} = \sum_{i \in \mathcal{W}} (\pi_i^c (1 - U_i) + \pi_i^p \cdot U_i) \cdot \left(1 - \sum_{i' \in \mathcal{N} \setminus \{0\}} \sum_{r \in \mathcal{R}} x_{ii'r} \right)$$

$$+\sum_{i\in\mathcal{W}}\cdot c_i^d\cdot(1-U_i)\cdot\left(t_i^d+\sum_{r\in\mathcal{R}}O_{ir}+T_i\cdot\sum_{i'\in\mathcal{N}\setminus\{0\}}\sum_{r\in\mathcal{R}}x_{ii'r}\right)$$

$$+\sum_{i\in\mathcal{W}}c_i^d\cdot U_i\cdot\left(T_i\cdot\sum_{i'\in\mathcal{N}\setminus\{0\}}\sum_{r\in\mathcal{R}}x_{ii'r}\right)+\sum_{i\in\mathcal{N}\setminus\{N+1\}}\sum_{i'\in\mathcal{N}\setminus\{0\}}\sum_{r\in\mathcal{R}}x_{ii'r}\cdot c_{ii'}$$

$$+\sum_{i\in\mathcal{W}}C_i\cdot\left(\sum_{i'\in\mathcal{N}\setminus\{0\}}\sum_{r\in\mathcal{R}}x_{ii'r}\right) \tag{A.12}$$

s.t:

$$\sum_{i\in\mathcal{N}\setminus\{0\}}x_{0ir}=1,\quad\forall r\in\mathcal{R} \tag{A.13}$$

$$\sum_{i\in\mathcal{N}\setminus\{N+1\}}x_{i,N+1,r}=1,\quad\forall r\in\mathcal{R} \tag{A.14}$$

$$\sum_{i\in\mathcal{N}\setminus\{N+1\}}x_{ii'r}=\sum_{i''\in\mathcal{N}\setminus\{0\}}x_{i'i''r},\quad\forall r\in\mathcal{R};\forall i^{'}\in\mathcal{W} \tag{A.15}$$

$$\sum_{i\in\mathcal{N}\setminus\{N+1\}}\sum_{r\in\mathcal{R}}x_{ii'r}\leq 1,\quad\forall i^{'}\in\mathcal{W} \tag{A.16}$$

$$O_{ir}\geq 0,\quad\forall i\in\mathcal{N};\forall r\in\mathcal{R} \tag{A.17}$$

$$O_{i'r}\leq T^{max}\cdot\sum_{i\in\mathcal{N}\setminus\{N+1\}}x_{ii'r},\quad\forall i^{'}\in\mathcal{N}\setminus\{0\};\forall r\in\mathcal{R} \tag{A.18}$$

$$O_{0r}=0,\quad\forall r\in\mathcal{R} \tag{A.19}$$

$$O_{0r}+t_{0i'}-O_{i'r}-M_{0i'}\cdot(1-x_{0i'r})\leq 0,\quad\forall i^{'}\in\mathcal{N}\setminus\{0\};\forall r\in\mathcal{R} \tag{A.20}$$

$$O_{ir}+t_{ii'}+T_i-O_{i'r}-M_{ii'}\cdot(1-x_{ii'r})\leq 0,\quad\forall i\in\mathcal{W};\forall i^{'}\in\mathcal{N};\forall r\in\mathcal{R} \tag{A.21}$$

$$2\leq I_{ir}\leq|\mathcal{N}|,\quad\forall i\in\mathcal{N}\setminus\{0\};\forall r\in\mathcal{R} \tag{A.22}$$

$$I_{ir}-I_{i'r}+1\leq(|\mathcal{N}|-1)\cdot(1-x_{ii'r}),\quad\forall i\in\mathcal{W};\forall i^{'}\in\mathcal{N}\setminus\{0\};\forall r\in\mathcal{R} \tag{A.23}$$

$$x_{ii'r}\in\{0,1\},\quad\forall i,i^{'},r \tag{A.24}$$

A.2 Component and System Data

All OWTs are identical, thus their reliability information is identical and summarized in Table A.1. Components status and age are provided in Table A.2.

Table A.1 – Reliability data for OWT i

	η_{ijk}	β_{ijk}	t^c_{ijk1}	t^c_{ijk2}	t^c_{ijk3}	t^p_{ijk2}	t^p_{ijk3}	c^c_{ijk1}	c^c_{ijk2}	c^c_{ijk3}	c^p_{ijk2}	c^p_{ijk3}
E_{i11}	26	2	0.75	2	3.5	2	3.5	50	100	200	100	200
E_{i12}	26	2	0.75	2	3.5	2	3.5	50	100	200	100	200
E_{i13}	26	2	0.75	2	3.5	2	3.5	50	100	200	100	200
E_{i14}	26	2	0.75	2	3.5	2	3.5	50	100	200	100	200
E_{i15}	26	2	0.75	2	3.5	2	3.5	50	100	200	100	200
E_{i21}	28	3.5	1	2.5	4	2.5	4	75	125	225	125	225
E_{i22}	28	3.5	1	2.5	4	2.5	4	75	125	225	125	225
E_{i23}	28	3.5	1	2.5	4	2.5	4	75	125	225	125	225
E_{i24}	28	3.5	1	2.5	4	2.5	4	75	125	225	125	225
E_{i31}	26	4	0.75	1.75	3.75	1.75	3.75	80	115	220	115	220
E_{i32}	26	4	0.75	1.75	3.75	1.75	3.75	80	115	220	115	220
E_{i33}	26	4	0.75	1.75	3.75	1.75	3.75	80	115	220	115	220
E_{i41}	25	3	1	2.25	4	2.25	4	50	100	225	100	225
E_{i42}	25	3	1	2.25	4	2.25	4	50	100	225	100	225

Table A.2 – Component status and age

E_{ijk}	u_{ijk}	B_{ijk}	E_{ijk}	u_{ijk}	B_{ijk}	E_{ijk}	u_{ijk}	B_{ijk}	E_{ijk}	u_{ijk}	B_{ijk}
E_{111}	1	10	E_{211}	1	12	E_{311}	1	11	E_{411}	1	6
E_{112}	0	8	E_{212}	0	6	E_{312}	0	16	E_{412}	0	8
E_{113}	1	6	E_{213}	1	8	E_{313}	1	7	E_{413}	0	6
E_{114}	0	10	E_{214}	1	12	E_{314}	1	10	E_{414}	1	8
E_{115}	1	12	E_{215}	1	12	E_{315}	0	16	E_{415}	1	4
E_{121}	1	10	E_{221}	1	8	E_{321}	1	12	E_{421}	1	8
E_{122}	0	4	E_{222}	1	18	E_{322}	0	16	E_{422}	0	16
E_{123}	1	12	E_{223}	0	10	E_{323}	1	12	E_{423}	1	6
E_{124}	1	10	E_{224}	1	9	E_{324}	1	10	E_{424}	0	10
E_{131}	1	12	E_{231}	1	7	E_{331}	1	12	E_{431}	1	9
E_{132}	0	16	E_{232}	1	6	E_{332}	0	16	E_{432}	1	7
E_{133}	1	14	E_{233}	1	12	E_{333}	1	14	E_{433}	1	10
E_{141}	1	10	E_{241}	1	14	E_{341}	0	10	E_{441}	1	8
E_{142}	1	5	E_{242}	1	12	E_{342}	0	12	E_{442}	1	6
E_{511}	1	10	E_{611}	1	12	E_{711}	0	7			
E_{512}	0	5	E_{612}	0	8	E_{712}	0	8			
E_{513}	1	8	E_{613}	1	4	E_{713}	0	5			
E_{514}	1	9	E_{614}	1	8	E_{714}	1	8			
E_{515}	1	11	E_{615}	0	12	E_{715}	1	14			
E_{521}	1	12	E_{621}	1	10	E_{721}	1	15			
E_{522}	0	9	E_{622}	1	8	E_{722}	0	16			
E_{523}	1	12	E_{623}	1	10	E_{723}	1	10			
E_{524}	0	8	E_{624}	1	20	E_{724}	1	10			
E_{531}	1	18	E_{631}	1	10	E_{731}	1	10			
E_{532}	1	8	E_{632}	0	8	E_{732}	0	9			
E_{533}	1	14	E_{633}	1	16	E_{733}	1	13			
E_{541}	0	8	E_{641}	1	12	E_{741}	0	14			
E_{542}	0	8	E_{642}	1	8	E_{742}	0	15			

A.3 Details of the optimal maintenance actions obtained for experiment set #5 (2.1.7)

Table A.3 – Maintenance actions performed for $R_0 = 86\%$: case of Experiments #5

OWT i	Patterns selected	$\mathcal{R}^*(\%)$
1	(0, 0, 0, 1, 0), (0, 0, 0, 0), (0, 0, 0), (0, 0)	88.61
2	(0, 1, 0, 0, 0), (0, 0, 0, 0), (0, 0, 0), (0, 0)	87.81
3	(0, 1, 0, 0, 1), (0, 0, 0, 0), (0, 0, 0), (1, 1)	86.54
4	(0, 0, 1, 0, 0), (0, 0, 0, 0), (0, 0, 0), (0, 0)	86.32
5	(0, 0, 0, 0, 0), (0, 0, 0, 1), (0, 0, 0), (1, 1)	91.11
6	(0, 1, 0, 0, 0), (0, 0, 0, 0), (0, 0, 0), (0, 0)	90.88
7	(1, 0, 1, 0, 0), (0, 0, 0, 0), (0, 0, 0), (1, 2)	87.37

Table A.4 – Maintenance actions performed for $R_0 = 90\%$: case of Experiments #5

OWT i	Patterns selected	$\mathcal{R}^*(\%)$
1	(0, 1, 0, 1, 0), (0, 0, 0, 0), (0, 0, 0), (0, 0)	90.44
2	(0, 1, 0, 0, 0), (0, 0, 1, 0), (0, 0, 0), (0, 0)	92.01
3	(0, 0, 0, 0, 1), (0, 1, 0, 0), (0, 1, 0), (1, 1)	90.18
4	(0, 0, 0, 0, 0), (0, 0, 0, 1), (0, 0, 0), (0, 0)	91.55
5	(0, 0, 0, 0, 0), (0, 0, 0, 1), (0, 0, 0), (1, 1)	91.11
6	(0, 1, 0, 0, 0), (0, 0, 0, 0), (0, 0, 0), (0, 0)	90.88
7	(1, 1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 0), (2, 1)	90.55

Table A.5 – Maintenance actions performed for $R_0 = 94\%$: case of Experiments #5

OWT i	Patterns selected	$\mathcal{R}^*(\%)$
1	(0, 1, 0, 0, 0), (0, 1, 0, 0), (0, 1, 0), (0, 0)	94.93
2	(0, 1, 0, 0, 0), (0, 0, 1, 0), (0, 0, 0), (2, 0)	95.96
3	(0, 1, 0, 0, 1), (0, 1, 0, 0), (0, 1, 0), (1, 2)	94.69
4	(0, 1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 0), (0, 0)	94.38
5	(0, 0, 0, 0, 0), (0, 1, 0, 1), (0, 0, 0), (1, 1)	94.49
6	(0, 1, 0, 0, 1), (0, 0, 0, 0), (0, 1, 0), (0, 0)	95.24
7	(1, 1, 1, 0, 0), (0, 1, 0, 0), (0, 0, 0), (3, 0)	94.45

Table A.6 – Maintenance actions performed for $R_0 = 98\%$: case of Experiments #5

OWT i	Patterns selected	$\mathcal{R}^*(\%)$
1	(0, 1, 0, 1, 0), (0, 1, 0, 0), (0, 2, 0), (2, 0)	98.32
2	(0, 1, 0, 0, 0), (0, 0, 1, 0), (0, 0, 0), (0, 3)	98.12
3	(0, 1, 0, 0, 1), (0, 2, 0, 2), (0, 2, 0), (1, 3)	98.06
4	(0, 1, 1, 0, 0), (0, 1, 0, 1), (0, 0, 0), (0, 0)	98.07
5	(0, 1, 0, 0, 0), (0, 1, 0, 1), (0, 0, 0), (2, 2)	98.02
6	(0, 0, 0, 0, 0), (0, 0, 0, 0), (0, 0, 0), (0, 0)	85.17
7	(1, 1, 1, 0, 0), (0, 1, 0, 2), (0, 1, 0), (3, 1)	98.02

Table A.7 – Maintenance actions performed for $R_0 = 99\%$: case of Experiments #5

OWT i	Patterns selected	$\mathcal{R}^*(\%)$
1	(0, 0, 0, 0, 0), (0, 0, 0, 0), (0, 0, 0), (0, 0)	81.98
2	(0, 0, 0, 0, 0), (0, 0, 0, 0), (0, 0, 0), (0, 0)	85.34
3	(0, 0, 0, 0, 0), (0, 0, 0, 0), (0, 0, 0), (0, 0)	0
4	(0, 1, 1, 0, 0), (0, 2, 0, 1), (0, 0, 0), (2, 0)]	99.03
5	(0, 1, 0, 0, 0), (0, 2, 0, 1), (2, 0, 0), (2, 3)	99.00
6	(0, 1, 0, 0, 2), (0, 0, 0, 2), (0, 2, 0), (3, 0)	99.05
7	(1, 1, 1, 0, 0), (0, 2, 0, 2), (0, 2, 0), (2, 3)	99.03

Appendix B

Supplementary Material for Section 2.2

B.1 Network and component data used in experiments #1 and #2

Table B.1 – Component information for networks 1 and 2.

Component	η_i	β_i	B_i	t_{i1}^m (hours)	t_{i2}^m (hours)	t_{i3}^m (hours)
1	15	3.2	6.19	2.32	5.85	8.56
2	16	3.1	5.20	1.70	5.49	9.28
3	20	3.4	6.79	1.98	5.05	9.71
4	17	3.4	5.61	2.35	5.36	8.56
5	15	3.2	6.46	1.73	5.57	9.89
6	19	3.4	6.01	2.49	5.13	8.97
7	16	3.2	8.97	2.35	5.09	8.88
8	19	3.4	6.18	2.46	5.83	9.68
9	20	3.1	9.42	1.71	5.50	9.21
10	19	3.0	5.51	2.20	5.54	9.37
11	16	3.4	8.90	2.37	5.56	8.86
12	19	3.1	7.28	1.60	5.48	9.43
13	16	3.3	5.85	2.37	5.51	8.82
14	20	3.2	6.38	2.05	5.63	9.17
15	19	3.5	8.91	2.01	5.70	9.77
16	18	3.0	9.38	2.00	5.04	9.54
17	16	3.5	6.21	2.02	5.34	9.89
18	20	3.5	6.85	2.02	5.11	9.75
19	16	3.0	6.66	1.54	5.65	8.66
20	17	3.1	7.45	1.92	5.14	8.75
21	19	3.4	5.38	2.49	5.14	9.06

B.2 Network and component data used in experiments #3

In the case of experiment #3, the component information and travel time matrix are provided below.

	To												
From	0	1	2	3	4	5	6	7	8	9	10	11	12
0	0.0	11.0	11.7	7.2	9.6	1.5	4.6	3.8	3.4	9.1	6.7	10.1	9.9
1	11.0	0.0	8.2	7.3	8.0	10.1	6.6	14.7	14.0	15.6	15.9	21.0	20.3
2	11.7	8.2	0.0	4.5	14.7	10.2	8.	15.3	15.1	19.8	18.3	19.2	21.4
3	7.2	7.3	4.5	0.0	11.5	5.7	3.7	10.8	10.6	15.4	13.8	15.4	17.0
4	9.6	8.0	14.7	11.5	0.0	9.8	8.2	11.7	10.9	8.9	10.8	19.0	15.7
5	1.5	10.1	10.2	5.7	9.8	0.0	3.5	5.1	4.8	10.4	8.2	10.8	11.3
6	4.6	6.6	8.0	3.7	8.2	3.5	0.0	8.4	7.9	11.8	10.6	14.3	14.4
7	3.8	14.7	15.3	10.8	11.7	5.1	8.4	0.0	0.9	7.8	4.2	7.2	6.2
8	3.4	14.0	15.1	10.6	10.9	4.8	7.9	0.9	0.0	7.2	3.8	8.1	6.5
9	9.1	15.6	19.8	15.4	8.9	10.4	11.8	7.8	7.2	0.0	4.1	13.9	8.1
10	6.7	15.9	18.3	13.8	10.8	8.2	10.6	4.2	3.8	4.1	0.0	9.8	4.9
11	10.1	21.0	19.2	15.4	19.0	10.8	14.3	7.2	8.1	13.9	9.8	0.0	7.2
12	9.9	20.3	21.4	17.0	15.7	11.3	14.4	6.2	6.5	8.1	4.9	7.2	0.0

Table B.2 – Travel time (hours) between components and O&M base (Node 0) for network 1.

	To																					
From	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
0	0.0	11.3	11.5	11.1	10.0	6.2	11.9	7.7	4.8	12.2	8.0	5.6	5.9	7.2	5.6	5.6	1.1	4.6	5.7	7.4	10.6	10.6
1	11.3	0.0	3.7	8.8	5.5	11.0	9.1	6.6	6.7	12.3	11.1	12.3	15.4	14.2	12.5	16.0	12.0	14.4	16.6	18.6	21.2	21.9
2	11.5	3.7	0.0	12.1	2.5	9.1	5.5	4.6	6.8	9.0	13.5	14.1	16.5	12.4	14.2	16.9	11.8	15.4	15.9	18.1	20.2	22.0
3	11.1	8.8	12.1	0.0	12.9	14.9	16.6	12.4	9.4	19.0	4.6	7.6	11.2	17.4	7.8	12.6	12.1	10.9	16.6	17.8	21.4	19.3
4	10.0	5.5	2.5	12.9	0.0	6.8	3.7	2.5	5.6	6.8	13.5	13.4	15.5	10.0	13.6	15.6	10.1	14.3	13.9	16.1	17.9	20.2
5	6.2	11.0	9.1	14.9	6.8	0.0	6.8	4.6	5.6	6.2	13.2	11.6	12.0	3.3	11.6	11.4	5.6	10.7	7.6	10.0	11.2	14.4
6	11.9	9.1	5.5	16.6	3.7	6.8	0.0	4.7	8.4	3.6	16.8	16.3	17.7	9.4	16.3	17.6	11.7	16.5	14.5	16.8	17.8	21.2
7	7.7	6.6	4.6	12.4	2.5	4.6	4.7	0.0	3.8	6.6	12.1	11.6	13.3	7.8	11.7	13.3	7.7	12.1	11.4	13.7	15.6	17.7
8	4.8	6.7	6.8	9.4	5.6	5.6	8.4	3.8	0.0	10.0	8.4	7.9	9.8	8.4	8.0	10.1	5.3	8.7	9.9	11.9	14.6	15.4
9	12.2	12.3	9.0	19.0	6.8	6.2	3.6	6.6	10.0	0.0	18.4	17.3	18.1	7.5	17.4	17.6	11.7	16.8	13.3	15.6	15.8	20.2
10	8.0	11.1	13.5	4.6	13.5	13.2	16.8	12.1	8.4	18.4	0.0	3.2	6.7	15.1	3.3	8.1	9.0	6.5	13.0	13.7	17.6	14.8
11	5.6	12.3	14.1	7.6	13.4	11.6	16.3	11.6	7.9	17.3	3.2	0.0	3.6	12.9	0.2	5.0	6.6	3.3	10.0	10.6	14.5	11.7
12	5.9	15.4	16.5	11.2	15.5	12.0	17.7	13.3	9.8	18.1	6.7	3.6	0.0	12.3	3.5	1.6	6.4	1.2	8.0	7.9	11.9	8.1
13	7.2	14.2	12.4	17.4	10.0	3.3	9.4	7.8	8.4	7.5	15.1	12.9	12.3	0.0	12.9	11.4	6.3	11.2	5.9	8.1	8.4	12.8
14	5.6	12.5	14.2	7.8	13.6	11.6	16.3	11.7	8.0	17.4	3.3	0.2	3.5	12.9	0.0	4.8	6.6	3.2	9.9	10.5	14.4	11.5
15	5.6	16.0	16.9	12.6	15.6	11.4	17.6	13.3	10.1	17.6	8.1	5.0	1.6	11.4	4.8	0.0	5.9	1.7	6.6	6.3	10.3	6.7
16	1.1	12.0	11.8	12.1	10.1	5.6	11.7	7.7	5.3	11.7	9.0	6.6	6.4	6.3	6.6	5.9	0.0	5.2	4.8	6.6	9.7	10.2
17	4.6	14.4	15.4	10.9	14.3	10.7	16.5	12.1	8.7	16.8	6.5	3.3	1.2	11.2	3.2	1.7	5.2	0.0	7.2	7.4	11.4	8.4
18	5.7	16.6	15.9	16.6	13.9	7.6	14.5	11.4	9.9	13.3	13.0	10.0	8.0	5.9	9.9	6.6	4.8	7.2	0.0	2.3	4.9	6.9
19	7.4	18.6	18.1	17.8	16.1	10.0	16.8	13.7	11.9	15.6	13.7	10.6	7.9	8.1	10.5	6.3	6.6	7.4	2.3	0.0	4.0	4.6
20	10.6	21.2	20.2	21.4	17.9	11.2	17.8	15.6	14.6	15.8	17.6	14.5	11.9	8.4	14.4	10.3	9.7	11.4	4.9	4.0	0.0	6.8
21	10.6	21.9	22.0	19.3	20.2	14.4	21.2	17.7	15.4	20.2	14.8	11.7	8.1	12.8	11.5	6.7	10.2	8.4	6.9	4.6	6.8	0.0

Table B.3 – Travel time (hours) between components and O&M base (Node 0) for network 2.

Table B.4 – Component information for WDN.

Component	η_i	β_i	B_i	t_{i1}^m (hours)	t_{i2}^m (hours)	t_{i3}^m (hours)
1	15	3.4	5.46	2.02	3.74	5.01
2	16	3.4	7.32	1.99	3.73	5.18
3	18	3.2	5.46	1.94	3.26	5.40
4	20	3.1	7.56	1.59	3.50	5.08
6	16	3.2	9.80	1.51	3.60	4.73
8	16	3.1	8.97	2.17	3.66	5.46
10	16	3.1	5.89	1.75	3.45	5.75
11	17	3.0	7.52	2.06	3.00	4.82
12	15	3.1	6.45	2.47	3.60	5.26
14	18	3.5	9.11	2.03	3.10	4.78
16	15	3.3	8.61	2.05	3.12	5.20
18	17	3.5	9.16	1.55	3.62	5.90
21	17	3.5	8.56	2.26	3.59	4.55
23	15	3.0	5.58	1.69	3.94	5.61
24	19	3.4	5.61	1.94	3.71	5.22
26	16	3.5	5.59	1.66	3.19	4.77
28	18	3.5	6.49	2.12	3.65	5.24
31	15	3.1	5.41	2.35	3.20	5.49
33	16	3.4	5.02	1.77	3.17	4.74
34	18	3.3	6.83	1.72	3.60	5.34
36	19	3.2	8.31	1.99	3.58	5.19

	To																					
From	0	1	2	3	4	6	8	10	11	12	14	16	18	21	23	24	26	28	31	33	34	36
0	0.0	3.5	3.4	3.3	2.8	3.2	2.8	2.9	1.5	2.2	2.8	1.3	1.6	3.0	2.4	3.7	1.4	1.5	3.9	3.0	3.9	4.0
1	3.5	0.0	1.9	1.7	0.9	3.3	4.3	5.4	4.5	5.1	5.8	2.3	3.9	6.5	2.8	4.2	4.1	4.9	5.5	5.5	6.6	7.4
2	3.4	1.9	0.0	3.3	2.2	4.7	5.3	6.0	4.8	5.5	6.1	2.7	4.4	6.3	1.4	2.4	3.2	4.3	3.9	4.2	5.3	6.6
3	3.3	1.7	3.3	0.0	1.2	1.7	3.0	4.2	3.8	4.2	4.9	2.0	3.0	6.0	3.8	5.4	4.4	4.9	6.4	6.0	7.0	7.3
4	2.8	0.9	2.2	1.2	0.0	2.5	3.4	4.5	3.6	4.2	4.9	1.5	3.0	5.7	2.7	4.2	3.5	4.2	5.3	5.1	6.1	6.7
6	3.2	3.3	4.7	1.7	2.5	0.0	1.5	2.9	2.9	3.1	3.7	2.4	2.1	5.2	4.7	6.3	4.6	4.7	7.0	6.2	7.1	6.8
8	2.8	4.3	5.3	3.0	3.4	1.5	0.0	1.3	1.8	1.7	2.2	2.5	1.2	3.8	4.8	6.3	4.2	3.9	6.7	5.7	6.4	5.7
10	2.9	5.4	6.0	4.2	4.5	2.9	1.3	0.0	1.4	0.9	1.0	3.3	1.6	2.8	5.3	6.6	4.2	3.5	6.6	5.4	5.9	4.8
11	1.5	4.5	4.8	3.8	3.6	2.9	1.8	1.4	0.0	0.7	1.3	2.3	0.8	2.3	3.9	5.2	2.7	2.1	5.2	4.0	4.7	3.9
12	2.2	5.1	5.5	4.2	4.2	3.1	1.7	0.9	0.7	0.0	0.7	2.9	1.2	2.1	4.6	5.8	3.3	2.6	5.8	4.5	5.1	4.0
14	2.8	5.8	6.1	4.9	4.9	3.7	2.2	1.0	1.3	0.7	0.0	3.6	1.9	1.8	5.2	6.3	3.8	2.9	6.2	4.8	5.2	3.9
16	1.3	2.3	2.7	2.0	1.5	2.4	2.5	3.3	2.3	2.9	3.6	0.0	1.7	4.2	2.4	3.9	2.4	2.8	4.7	4.0	5.1	5.3
18	1.6	3.9	4.4	3.0	3.0	2.1	1.2	1.6	0.8	1.2	1.9	1.7	0.0	3.1	3.8	5.2	3.0	2.7	5.5	4.5	5.3	4.7
21	3.0	6.5	6.3	6.0	5.7	5.2	3.8	2.8	2.3	2.1	1.8	4.2	3.1	0.0	5.1	5.8	3.3	2.2	5.2	3.7	3.8	2.1
23	2.4	2.8	1.4	3.8	2.7	4.7	4.8	5.3	3.9	4.6	5.2	2.4	3.8	5.1	0.0	1.6	1.8	3.0	2.7	2.8	3.9	5.2
24	3.7	4.2	2.4	5.4	4.2	6.3	6.3	6.6	5.2	5.8	6.3	3.9	5.2	5.8	1.6	0.0	2.5	3.6	1.6	2.5	3.5	5.4
26	1.4	4.1	3.2	4.4	3.5	4.6	4.2	4.2	2.7	3.3	3.8	2.4	3.0	3.3	1.8	2.5	0.0	1.2	2.5	1.6	2.7	3.5
28	1.5	4.9	4.3	4.9	4.2	4.7	3.9	3.5	2.1	2.6	2.9	2.8	2.7	2.2	3.0	3.6	1.2	0.0	3.2	1.9	2.6	2.5
31	3.9	5.5	3.9	6.4	5.3	7.0	6.7	6.6	5.2	5.8	6.2	4.7	5.5	5.2	2.7	1.6	2.5	3.2	0.0	1.5	2.1	4.2
33	3.0	5.5	4.2	6.0	5.1	6.2	5.7	5.4	4.0	4.5	4.8	4.0	4.5	3.7	2.8	2.5	1.6	1.9	1.5	0.0	1.1	2.9
34	3.9	6.6	5.3	7.0	6.1	7.1	6.4	5.9	4.7	5.1	5.2	5.1	5.3	3.8	3.9	3.5	2.7	2.6	2.1	1.1	0.0	2.3
36	4.0	7.4	6.6	7.3	6.7	6.8	5.7	4.8	3.9	4.0	3.9	5.3	4.7	2.1	5.2	5.4	3.5	2.5	4.2	2.9	2.3	0.0

Table B.5 – Travel time (hours) between components and O&M base (Node 0) for WDN.

Bibliography

Abdar, M., Pourpanah, F., Hussain, S., Rezazadegan, D., Liu, L., Ghavamzadeh, M., Fieguth, P., Cao, X., Khosravi, A., Acharya, U. R., et al. (2021). A review of uncertainty quantification in deep learning: Techniques, applications and challenges. *Information fusion*, 76:243–297.

Achamrah, F. E. and Attajer, A. (2024). Multi-objective reinforcement learning-based framework for solving selective maintenance problems in reconfigurable cyber-physical manufacturing systems. *International Journal of Production Research*, 62(10):3460–3482.

Afsar, H. M. and Labadie, N. (2013). Team orienteering problem with decreasing profits. *Electronic Notes in Discrete Mathematics*, 41:285–293.

Al-Jabouri, H., Saif, A., and Diallo, C. (2023a). Robust selective maintenance optimization of series–parallel mission-critical systems subject to maintenance quality uncertainty. *Computational Management Science*, 20(1):29.

Al-Jabouri, H., Saif, A., Diallo, C., and Khatab, A. (2023b). Branch-and-price algorithms for large-scale mission-oriented maintenance planning problems. *Computers & Operations Research*, 153:106191.

Al-Jabouri, H., Saif, A., Diallo, C., and Khatab, A. (2024a). Distributionally-robust chance-constrained optimization of selective maintenance under uncertain repair duration. *Expert Systems with Applications*, 239:122303.

Al-Jabouri, H., Saif, A., Khatab, A., Diallo, C., and Venkatadri, U. (2022). Selective maintenance optimization: a condensed critical review and future research directions. *IFAC-PapersOnLine*, 55(10):1213–1218.

Al-Jabouri, H., Saif, A., Khatab, A., Diallo, C., and Venkatadri, U. (2024b). A critical review of selective maintenance for mission-oriented systems: challenges and a roadmap for novel contributions. *International Journal of Production Research*, 62(13):4980–5015.

Alkhaleel, B. A., Liao, H., and Sullivan, K. M. (2022a). Model and solution method for mean-risk cost-based post-disruption restoration of interdependent critical infrastructure networks. *Computers & Operations Research*, 144:105812.

Alkhaleel, B. A., Liao, H., and Sullivan, K. M. (2022b). Risk and resilience-based optimal post-disruption restoration for critical infrastructures under uncertainty. *European Journal of Operational Research*, 296(1):174–202.

Allal, A., Sahnoun, M., Adjoudj, R., Benslimane, S., and Mazar, M. (2021). Multi-agent based simulation-optimization of maintenance routing in offshore wind farms. *Computers and Industrial Engineering*, 157:107342.

Almoghathawi, Y., Barker, K., and Albert, L. A. (2019). Resilience-driven restoration model for interdependent infrastructure networks. *Reliability Engineering & System Safety*, 185:12–23.

Amjadian, A., O'Neil, R., Khatab, A., Chen, J., Venkatadri, U., and Diallo, C. (2024). Optimising resource-constrained fleet selective maintenance with asynchronous maintenance breaks. *International Journal of Production Research*, pages 1–23.

An, D., Choi, J.-H., and Kim, N. H. (2018). Prediction of remaining useful life under different conditions using accelerated life testing data. *Journal of Mechanical Science and Technology*, 32(6):2497–2507.

An, Y., Chen, X., Gao, K., Zhang, L., Li, Y., and Zhao, Z. (2023). A hybrid multi-objective evolutionary algorithm for solving an adaptive flexible job-shop rescheduling problem with real-time order acceptance and condition-based preventive maintenance. *Expert Systems with Applications*, 212:118711.

An, Y., Chen, X., Li, Y., Zhang, J., and Jiang, J. (2021). Flexible job-shop scheduling and heterogeneous repairman assignment with maintenance time window and employee timetable constraints. *Expert Systems with Applications*, 186:115693.

Archetti, C., Bianchessi, N., and Speranza, M. G. (2013). The capacitated team orienteering problem with incomplete service. *Optimization letters*, 7(7):1405–1417.

Archetti, C., Feillet, D., Hertz, A., and Speranza, M. G. (2009). The capacitated team orienteering and profitable tour problems. *Journal of the Operational Research Society*, 60(6):831–842.

Arulmozhi, G. (2002). Exact equation and an algorithm for reliability evaluation of k-out-of-n: G system. *Reliability Engineering & System Safety*, 78(2):87–91.

Awad, M. and Ertem, M. (2017). Stochastic scheduling of workforce-constrained preventive maintenance activities in petroleum plants. In *2017 Annual Reliability and Maintainability Symposium (RAMS)*, pages 1–5, Orlando, Florida, USA. IEEE.

Azadeh, A., Asadzadeh, S., Salehi, N., and Firoozi, M. (2015). Condition-based maintenance effectiveness for series–parallel power generation system—a combined markovian simulation model. *Reliability Engineering & System Safety*, 142:357–368.

Beale, E. and Forrest, J. J. (1976). Global optimization using special ordered sets. *Mathematical Programming*, 10:52–69.

Benaichouche, A. N., Oulhadj, H., and Siarry, P. (2013). Improved spatial fuzzy c-means clustering for image segmentation using pso initialization, mahalanobis distance and post-segmentation correction. *Digital signal processing*, 23(5):1390–1400.

Blazewicz, J., Lenstra, J. K., and Kan, A. R. (1983). Scheduling subject to resource constraints: classification and complexity. *Discrete applied mathematics*, 5(1):11–24.

Bora, D. J., Gupta, D., and Kumar, A. (2014). A comparative study between fuzzy clustering algorithm and hard clustering algorithm. *arXiv preprint arXiv:1404.6059*.

Braaksma, A., Klingenberg, W., and Veldman, J. (2013). Failure mode and effect analysis in asset maintenance: a multiple case study in the process industry. *International Journal of Production Research*, 51(4):1055–1071.

Cai, B., Zhang, Y., Wang, H., Liu, Y., Ji, R., Gao, C., Kong, X., and Liu, J. (2021). Resilience evaluation methodology of engineering systems with dynamic-bayesian-network-based degradation and maintenance. *Reliability Engineering & System Safety*, 209:107464.

Campbell, A. M., Gendreau, M., and Thomas, B. W. (2011). The orienteering problem with stochastic travel and service times. *Annals of Operations Research*, 186(1):61–81.

Cao, W., Jia, X., Hu, Q., Song, W., and Ge, H. (2017). Selective maintenance for maximising system availability: a simulation approach. *International Journal of Innovative Computing and Applications*, 8(1):12–20.

Cao, W., Jia, X., Hu, Q., Zhao, J., and Wu, Y. (2018). A literature review on selective maintenance for multi-unit systems. *Quality and Reliability Engineering International*, 34(5):824–845.

Cao, W., Song, W., Hu, Q., and Du, Y. (2016). An exact method for solving selective maintenance problems considering imperfect maintenance. In *Intelligent Networking and Collaborative Systems (INCoS), 2016 International Conference on*, pages 522–526. IEEE.

Carroll, J., McDonald, A., and McMillan, D. (2016). Failure rate, repair time and unscheduled O&M cost analysis of offshore wind turbines. *Wind Energy*, 19(6):1107–1119.

Cassady, C., Pohl, E. A., Mason, S. J., and Yeung, T. G. (2005). Multi-state selective maintenance decisions. Technical report, Arkansas University Fayetteville.

Cassady, C. R., Mason, S. J., Ormon, S., Schneider, K., Rainwater, C., Carrasco, M., and Honeycutt, J. (2003). Fleet-level selective maintenance and aircraft scheduling. Technical report, University of Arkansas/Air Force Research Laboratory.

Cassady, C. R., Murdock Jr, W. P., and Pohl, E. A. (2001a). Selective maintenance for support equipment involving multiple maintenance actions. *European Journal of Operational Research*, 129(2):252–258.

Cassady, C. R., Pohl, E. A., and Murdock, W. P. (2001b). Selective maintenance modeling for industrial systems. *Journal of quality in maintenance engineering*.

Cavalcante, V. F., Cardonha, C. H., and Herrmann, R. G. (2013). A resource constrained project scheduling problem with bounded multitasking. *IFAC Proceedings*, 46(24):433–437.

Çavdar, B., He, Q., and Qiu, F. (2022). Repair crew routing for power distribution network restoration. *arXiv preprint arXiv:2204.04848*.

Chaabane, K., Khatab, A., Diallo, C., Aghezzaf, E.-H., and Venkatadri, U. (2020). Integrated imperfect multimission selective maintenance and repairpersons assignment problem. *Reliability Engineering & System Safety*, 199:106895.

Chang, P.-C., Huang, D.-H., Lin, Y.-K., and Nguyen, T.-P. (2021). Reliability and maintenance models for a time-related multi-state flow network via d-mc approach. *Reliability Engineering & System Safety*, 216:107962.

Chao, I.-M., Golden, B. L., and Wasil, E. A. (1996). A fast and effective heuristic for the orienteering problem. *European journal of operational research*, 88(3):475–489.

Chari, N., Diallo, C., and Venkatadri, U. (2014). State of the art on performability across the sustainable value chain. *International Journal of Performability Engineering*, 10(6):543.

Chen, C., Liu, Y., and Huang, H.-Z. (2012). Optimal load distribution for multi-state systems under selective maintenance strategy. In *2012 International conference on quality, reliability, risk, maintenance, and safety engineering*, pages 436–442, Chengdu. IEEE.

Chen, C., Lu, N., Jiang, B., Xing, Y., and Zhu, Z. H. (2021a). Prediction interval estimation of aeroengine remaining useful life based on bidirectional long short-term memory network. *IEEE Transactions on Instrumentation and Measurement*, 70:1–13.

Chen, C., Shi, J., Lu, N., Zhu, Z. H., and Jiang, B. (2022). Data-driven predictive maintenance strategy considering the uncertainty in remaining useful life prediction. *Neurocomputing*, 494:79–88.

Chen, C., Zhu, Z. H., Shi, J., Lu, N., and Jiang, B. (2021b). Dynamic predictive maintenance scheduling using deep learning ensemble for system health prognostics. *IEEE Sensors Journal*, 21(23):26878–26891.

Chen, Q., Wang, P., Yang, L., Wang, J., and Yi, X. (2023a). Mission and reliability driven fleet-level selective maintenance planning and scheduling two-stage method. *Applied Sciences*, 13(15):8600.

Chen, Z., Chen, Z., Zhou, D., and Pan, E. (2023b). Joint optimization of fleet-level sequential selective maintenance and repairpersons assignment for multi-state manufacturing systems. *Computers & Industrial Engineering*, 182:109411.

Cheng, G., Li, L., Shangguan, C., Yang, N., Jiang, B., and Tao, N. (2023). Optimal joint inspection and mission abort policy for a partially observable system. *Reliability Engineering & System Safety*, 229:108870.

Cheng, G., Li, L., Zhang, L., Yang, N., Jiang, B., Shangguan, C., and Su, Y. (2022). Optimal joint inspection and mission abort policies for degenerative systems. *IEEE Transactions on Reliability*, 72(1):137–150.

Cheng, G., Shen, J., Wang, F., Li, L., and Yang, N. (2024). Optimal mission abort policy for a multi-component system with failure interaction. *Reliability Engineering & System Safety*, 242:109791.

Christofides, N., Alvarez-Valdés, R., and Tamarit, J. M. (1987). Project scheduling with resource constraints: A branch and bound approach. *European Journal of Operational Research*, 29(3):262–273.

Cimellaro, G., Tinebra, A., Renschler, C., and Fragiadakis, M. (2016). New resilience index for urban water distribution networks. *Journal of Structural Engineering*, 142(8):C4015014.

Coit, D. W. and Zio, E. (2019). The evolution of system reliability optimization. *Reliability Engineering & System Safety*, 192:106259.

Dai, L., Stålhane, M., and Utne, I. B. (2015). Routing and scheduling of maintenance fleet for offshore wind farms. *Wind Engineering*, 39(1):15–30.

Dao, C. D. and Zuo, M. J. (2015). Selective maintenance for multistate series systems with s-dependent components. *IEEE Transactions on Reliability*, 65(2):525–539.

Dao, C. D. and Zuo, M. J. (2017a). Optimal selective maintenance for multi-state systems in variable loading conditions. *Reliability engineering & system safety*, 166:171–180.

Dao, C. D. and Zuo, M. J. (2017b). Selective maintenance of multi-state systems with structural dependence. *Reliability engineering & system safety*, 159:184–195.

Dao, C. D., Zuo, M. J., and Pandey, M. (2014). Selective maintenance for multi-state series–parallel systems under economic dependence. *Reliability Engineering & System Safety*, 121:240–249.

Das Adhikary, D., Bose, G. K., Jana, D. K., Bose, D., and Mitra, S. (2016). Availability and cost-centered preventive maintenance scheduling of continuous operating series systems using multi-objective genetic algorithm: A case study. *Quality Engineering*, 28(3):352–357.

De Jonge, B. and Scarf, P. A. (2020). A review on maintenance optimization. *European Journal of Operational Research*, 285(3):805–824.

de Pater, I. and Mitici, M. (2022). Novel metrics to evaluate probabilistic remaining useful life prognostics with applications to turbofan engines. In *PHM society European conference*, volume 7, pages 96–109.

de Pater, I., Reijns, A., and Mitici, M. (2022). Alarm-based predictive maintenance scheduling for aircraft engines with imperfect remaining useful life prognostics. *Reliability Engineering & System Safety*, 221:108341.

Dehghan Shoorkand, H., Nourelfath, M., and Hajji, A. (2023). A deep learning approach for integrated production planning and predictive maintenance. *International Journal of Production Research*, pages 1–20.

Diallo, C., Khatab, A., Venkatadri, U., and Aghezzaf, E.-H. (2017). A joint selective maintenance and multiple repair-person assignment problem. In *International Conference on Industrial Engineering and Systems Management-IESM 2017*, pages 317–322, Saarbrucken, Germany. Wirtschaftswissenschaften HTW SAAR.

Diallo, C., Venkatadri, U., Khatab, A., and Liu, Z. (2018). Optimal selective maintenance decisions for large serial k-out-of-n: G systems under imperfect maintenance. *Reliability Engineering & System Safety*, 175:234–245.

Diallo, C., Venkatadri, U., Khatab, A., Liu, Z., and Aghezzaf, E.-H. (2019). Optimal joint selective imperfect maintenance and multiple repairpersons assignment strategy for complex multicomponent systems. *International Journal of Production Research*, 57(13):4098–4117.

Doorn, N., Gardoni, P., and Murphy, C. (2019). A multidisciplinary definition and evaluation of resilience: The role of social justice in defining resilience. *Sustainable and Resilient Infrastructure*, 4(3):112–123.

Duan, C., Deng, C., Gharaei, A., Wu, J., and Wang, B. (2018). Selective maintenance scheduling under stochastic maintenance quality with multiple maintenance actions. *International Journal of Production Research*, 56(23):7160–7178.

Dui, H., Liu, M., Song, J., and Wu, S. (2023). Importance measure-based resilience management: Review, methodology and perspectives on maintenance. *Reliability Engineering & System Safety*, 237:109383.

Dui, H., Zhang, H., Dong, X., and Zhang, S. (2024). Cascading failure and resilience optimization of unmanned vehicle distribution networks in iot. *Reliability Engineering & System Safety*, 246:110071.

Dui, H., Zheng, X., and Wu, S. (2021). Resilience analysis of maritime transportation systems based on importance measures. *Reliability Engineering & System Safety*, 209:107461.

Eidinger, J. and Tang, A. (2012). Christchurch, new zealand earthquake sequence of mw 7.1 september 04, 2010 mw 6.3 february 22, 2011 mw 6.0 june 13, 2011: Lifeline performance. *American Society of Civil Engineers, Reston, VA*.

Elçi, Ö. and Noyan, N. (2018). A chance-constrained two-stage stochastic programming model for humanitarian relief network design. *Transportation research part B: methodological*, 108:55–83.

Ertem, M., As' ad, R., Awad, M., and Al-Bar, A. (2022). Workers-constrained shutdown maintenance scheduling with skills flexibility: Models and solution algorithms. *Computers & Industrial Engineering*, 172:108575.

Fan, X., Zhang, X., and Yu, X. (2022). A graph convolution network-deep reinforcement learning model for resilient water distribution network repair decisions. *Computer-Aided Civil and Infrastructure Engineering*.

Fang, F., Zhao, Z.-J., Huang, C., Zhang, X.-Y., Wang, H.-T., and Yang, Y.-J. (2019). Application of reliability-centered maintenance in metro door system. *IEEE Access*, 7:186167–186174.

Fang, Y. and Sansavini, G. (2016). Optimum post-disruption restoration for enhanced infrastructure network resilience: a fuzzy programming approach. In *Proc. Eur. Saf. Rel. Conf*, pages 587–594.

Fang, Y.-P. and Sansavini, G. (2019). Optimum post-disruption restoration under uncertainty for enhancing critical infrastructure resilience. *Reliability Engineering & System Safety*, 185:1–11.

Fekri, M., Heydari, M., and Mazdeh, M. (2023). Two-objective optimization of preventive maintenance orders scheduling as a multi-skilled resource-constrained flow shop problem. *Decision Science Letters*, 12(1):41–54.

Feng, Q., Bi, X., Zhao, X., Chen, Y., and Sun, B. (2017). Heuristic hybrid game approach for fleet condition-based maintenance planning. *Reliability Engineering & System Safety*, 157:166–176.

Feng, X., Chen, X., Zhang, L., and An, Y. (2024). Optimization of selective maintenance problem with stochastic durations in mission-oriented system subjecting to s-dependent competing risks. *Expert Systems with Applications*, 245:123068.

Fomin, F. V. and Lingas, A. (2002). Approximation algorithms for time-dependent orienteering. *Information Processing Letters*, 83(2):57–62.

Frederick, D. K., DeCastro, J. A., and Litt, J. S. (2007). User's guide for the commercial modular aero-propulsion system simulation (c-mapss). Technical report, NASA.

Fu, S., Lin, L., Wang, Y., Guo, F., Zhao, M., Zhong, B., and Zhong, S. (2024). Mca-dtcn: a novel dual-task temporal convolutional network with multi-channel attention for first prediction time detection and remaining useful life prediction. *Reliability Engineering & System Safety*, 241:109696.

Gal, Y. and Ghahramani, Z. (2016). Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In *international conference on machine learning*, pages 1050–1059. PMLR.

Gebre, B. A. and Ramirez-Marquez, J. E. (2007). Element substitution algorithm for general two-terminal network reliability analyses. *IIE transactions*, 39(3):265–275.

Geisbush, J. and Ariaratnam, S. (2023). Reliability centered maintenance (rcm): literature review of current industry state of practice. *Journal of Quality in Maintenance Engineering*, 29(2):313–337.

Ghorbani, M., Nourelfath, M., and Gendreau, M. (2024). Stochastic programming for selective maintenance optimization with uncertainty in the next mission conditions. *Reliability Engineering & System Safety*, 241:109624.

Ghorbani-Renani, N., González, A. D., and Barker, K. (2021). A decomposition approach for solving tri-level defender-attacker-defender problems. *Computers & Industrial Engineering*, 153:107085.

Goel, A. and Meisel, F. (2013). Workforce routing and scheduling for electricity network maintenance with downtime minimization. *European Journal of Operational Research*, 231(1):210–228.

Golden, B. L., Levy, L., and Vohra, R. (1987). The orienteering problem. *Naval Research Logistics (NRL)*, 34(3):307–318.

González, A. D., Dueñas-Osorio, L., Sánchez-Silva, M., and Medaglia, A. L. (2016). The interdependent network design problem for optimal infrastructure system restoration. *Computer-Aided Civil and Infrastructure Engineering*, 31(5):334–350.

Gu, H., Lam, H. C., and Zinder, Y. (2022). Planning rolling stock maintenance: Optimization of train arrival dates at a maintenance center. *Journal of Industrial and Management Optimization*, 18(2):747–772.

Guo, J., Li, Z., and Wolf, J. (2016). Reliability centered preventive maintenance optimization for aircraft indicators. In *2016 Annual Reliability and Maintainability Symposium (RAMS)*, Tucson, United States.

Guo, L., Li, R., Wang, Y., Yang, J., Liu, Y., Chen, Y., and Zhang, J. (2023). Availability for multi-component k-out-of-n: G warm-standby system in series with shut-off rule of suspended animation. *Reliability Engineering & System Safety*, 233:109106.

Gurobi Optimization, L. (2024). Multiple objectives. Accessed: 2024-07-06.

Hao, Y., Jia, L., Zio, E., Wang, Y., Small, M., and Li, M. (2023). Improving resilience of high-speed train by optimizing repair strategies. *Reliability Engineering & System Safety*, 237:109381.

Heil, J., Hoffmann, K., and Buscher, U. (2020). Railway crew scheduling: Models, methods and applications. *European journal of operational research*, 283(2):405–425.

Hesabi, H., Nourelfath, M., and Hajji, A. (2022). A deep learning predictive model for selective maintenance optimization. *Reliability Engineering & System Safety*, 219:108191.

Hoseinie, S. H., Kumar, U., and Ghodrati, B. (2016). Reliability centered maintenance (rcm) for automated mining machinery. Technical report, Luleå University of Technology, Operation, Maintenance and Acoustics. Godkänd; 2016; Bibliografisk uppgift: VINNOVA SIP-STRIM; 20160316 (hadhos).

Hosseini, S., Barker, K., and Ramirez-Marquez, J. E. (2016). A review of definitions and measures of system resilience. *Reliability Engineering & System Safety*, 145:47–61.

Hummeltenberg, W. (1984). Implementations of special ordered sets in mp software. *European Journal of Operational Research*, 17(1):1–15.

Iannacone, L., Sharma, N., Tabandeh, A., and Gardoni, P. (2022). Modeling time-varying reliability and resilience of deteriorating infrastructure. *Reliability Engineering & System Safety*, 217:108074.

IEA (2020). Offshore wind. https://www.iea.org/reports/offshore-wind.

Ikonen, T. J., Mostafaei, H., Ye, Y., Bernal, D. E., Grossmann, I. E., and Harjunkoski, I. (2020). Large-scale selective maintenance optimization using bathtub-shaped failure rates. *Computers & Chemical Engineering*, 139:106876.

Ilhan, T., Iravani, S. M., and Daskin, M. S. (2008). The orienteering problem with stochastic profits. *Iie Transactions*, 40(4):406–421.

Iori, M., De Lima, V. L., Martello, S., Miyazawa, F. K., and Monaci, M. (2021). Exact solution techniques for two-dimensional cutting and packing. *European Journal of Operational Research*, 289(2):399–415.

Irawan, C. A., Eskandarpour, M., Ouelhadj, D., and Jones, D. (2021). Simulation-based optimisation for stochastic maintenance routing in an offshore wind farm. *European Journal of Operational Research*, 289(3):912–926.

Irawan, C. A., Ouelhadj, D., Jones, D., Stålhane, M., and Sperstad, I. B. (2017). Optimisation of maintenance routing and scheduling for offshore wind farms. *European Journal of Operational Research*, 256(1):76–89.

Irawan, C. A., Salhi, S., and Chan, H. K. (2022). A continuous location and maintenance routing problem for offshore wind farms: Mathematical models and hybrid methods. *Computers & Operations Research*, 144:105825.

Ivanov, D., Dolgui, A., Sokolov, B., Werner, F., and Ivanova, M. (2016). A dynamic model and an algorithm for short-term supply chain scheduling in the smart factory industry 4.0. *International Journal of Production Research*, 54(2):386–402.

Jabouri, H., Saif, A., Diallo, C., and Khatab, A. (2021). Optimal selective maintenance decisions for integrated imperfect large-scale selective maintenance and repairpersons assignment problem using a column generation approach. Technical report, Dalhousie University, Halifax, NS, Discussion paper.

Jacques, C. C., McIntosh, J., Giovinazzi, S., Kirsch, T. D., Wilson, T., and Mitrani-Reiser, J. (2014). Resilience of the canterbury hospital system to the 2011 christchurch earthquake. *Earthquake spectra*, 30(1):533–554.

Jardine, A. K., Lin, D., and Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical systems and signal processing*, 20(7):1483–1510.

Javed, K., Gouriveau, R., and Zerhouni, N. (2015). A new multivariate approach for prognostics based on extreme learning machine and fuzzy clustering. *IEEE transactions on cybernetics*, 45(12):2626–2639.

Jia, C. and Zhang, C. (2020). Joint optimization of maintenance planning and workforce routing for a geographically distributed networked infrastructure. *IISE Transactions*, 52(7):732–750.

Jia, C., Zhang, C., Li, Y.-F., and Li, Q.-L. (2023). Joint pre-and post-disaster planning to enhance the resilience of critical infrastructures. *Reliability Engineering & System Safety*, 231:109023.

Jiang, T. and Liu, Y. (2020a). Robust selective maintenance strategy under imperfect observations: A multi-objective perspective. *IISE Transactions*, 52(7):751–768.

Jiang, T. and Liu, Y. (2020b). Selective maintenance strategy for systems executing multiple consecutive missions with uncertainty. *Reliability Engineering & System Safety*, 193:106632.

Jin, H., Song, X., and Xia, H. (2023). Optimal maintenance strategy for large-scale production systems under maintenance time uncertainty. *Reliability Engineering & System Safety*, 240:109594.

Jin, T. and Coit, D. W. (2003). Approximating network reliability estimates using linear and quadratic unreliability of minimal cuts. *Reliability Engineering & System Safety*, 82(1):41–48.

Jiu, S., Zhou, Z., and Liu, J. (2013). The equipment maintenance scheduling problem in a coal production system. *International Journal of Production Research*, 51(17):5309–5336.

Kansal, M. L., Kumar, A., and Sharma, P. (1995). Reliability analysis of water distribution systems under uncertainty. *Reliability Engineering & System Safety*, 50(1):51–59.

Kaynak, O. (2007). The exhilarating journey from industrial electronics to industrial informatics. In *2007 2nd IEEE Conference on Industrial Electronics and Applications*, pages xli–xlii. IEEE.

Keizer, M. C. O., Flapper, S. D. P., and Teunter, R. H. (2017). Condition-based maintenance policies for systems with multiple dependent components: A review. *European Journal of Operational Research*, 261(2):405–420.

Kelley, Jr, J. E. (1960). The cutting-plane method for solving convex programs. *Journal of the society for Industrial and Applied Mathematics*, 8(4):703–712.

Keynia, F., Mirhosseini, M., Heydari, A., and Fekih, A. (2022). A budget allocation and programming-based rcm approach to improve the reliability of power distribution networks. *Energy Reports*, 8:5591–5602.

Khatab, A. and Aghezzaf, E.-H. (2016). Selective maintenance optimization when quality of imperfect maintenance actions are stochastic. *Reliability engineering & system safety*, 150:182–189.

Khatab, A., Aghezzaf, E. H., Diallo, C., and Djelloul, I. (2017a). Selective maintenance optimisation for series-parallel systems alternating missions and scheduled breaks with stochastic durations. *International Journal of Production Research*, 55(10):3008–3024.

Khatab, A., Aghezzaf, E.-H., Djelloul, I., and Sari, Z. (2017b). Selective maintenance optimization for systems operating missions and scheduled breaks with stochastic durations. *Journal of manufacturing systems*, 43:168–177.

Khatab, A., Diallo, C., Aghezzaf, E.-H., and Venkatadri, U. (2018a). Condition-based selective maintenance for stochastically degrading multi-component systems under periodic inspection and imperfect maintenance. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 232(4):447–463.

Khatab, A., Diallo, C., Aghezzaf, E.-H., and Venkatadri, U. (2020). Optimization of the integrated fleet-level imperfect selective maintenance and repairpersons assignment problem. *Journal of Intelligent Manufacturing*, pages 1–16.

Khatab, A., Diallo, C., Aghezzaf, E.-H., and Venkatadri, U. (2022). Optimization of the integrated fleet-level imperfect selective maintenance and repairpersons assignment problem. *Journal of Intelligent Manufacturing*, 33:703–718.

Khatab, A., Diallo, C., Venkatadri, U., Liu, Z., and Aghezzaf, E.-H. (2018b). Optimization of the joint selective maintenance and repairperson assignment problem under imperfect maintenance. *Computers & Industrial Engineering*, 125:413–422.

Kong, J., Zhang, C., and Simonovic, S. P. (2021). Optimizing the resilience of interdependent infrastructures to regional natural hazards with combined improvement measures. *Reliability Engineering & System Safety*, 210:107538.

Kovacs, A. A., Parragh, S. N., Doerner, K. F., and Hartl, R. F. (2012). Adaptive large neighborhood search for service technician routing and scheduling problems. *Journal of scheduling*, 15(5):579–600.

Kuo, W. and Zuo, M. J. (2003). *Optimal reliability modeling: principles and applications*. John Wiley & Sons.

Kusiak, A. (2018). Smart manufacturing. *International Journal of Production Research*, 56(1-2):508–517.

Lan, P., Lin, M., and Naichao, W. (2017). A fleet-level selective maintenance model for long-distance highway transportation considering stochastic repair quality. In *2017 2nd International Conference on System Reliability and Safety (ICSRS)*, pages 348–353, Milan, Italy. IEEE.

LeCun, Y., Haffner, P., Bottou, L., and Bengio, Y. (1999). Object recognition with gradient-based learning. In *Shape, contour and grouping in computer vision*, pages 319–345. Springer.

Lee, J. and Mitici, M. (2023). Deep reinforcement learning for predictive aircraft maintenance using probabilistic remaining-useful-life prognostics. *Reliability Engineering & System Safety*, 230:108908.

Levitin, G. and Finkelstein, M. (2018). Optimal mission abort policy for systems operating in a random environment. *Risk Analysis*, 38(4):795–803.

Levitin, G., Finkelstein, M., and Dai, Y. (2018a). Mission abort policy balancing the uncompleted mission penalty and system loss risk. *Reliability Engineering & System Safety*, 176:194–201.

Levitin, G., Finkelstein, M., and Dai, Y. (2020). Mission abort policy optimization for series systems with overlapping primary and rescue subsystems operating in a random environment. *Reliability Engineering & System Safety*, 193:106590.

Levitin, G., Finkelstein, M., and Xiang, Y. (2021a). Optimal inspections and mission abort policies for multistate systems. *Reliability Engineering & System Safety*, 214:107700.

Levitin, G., Xing, L., and Dai, Y. (2017). Mission abort policy in heterogeneous nonrepairable 1-out-of-n warm standby systems. *IEEE Transactions on Reliability*, 67(1):342–354.

Levitin, G., Xing, L., and Dai, Y. (2018b). Co-optimization of state dependent loading and mission abort policy in heterogeneous warm standby systems. *Reliability Engineering & System Safety*, 172:151–158.

Levitin, G., Xing, L., and Dai, Y. (2021b). Joint optimal mission aborting and replacement and maintenance scheduling in dual-unit standby systems. *Reliability Engineering & System Safety*, 216:107921.

Levitin, G., Xing, L., and Dai, Y. (2021c). Mission aborting in n-unit systems with work sharing. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 52(8):4875–4886.

Levitin, G., Xing, L., and Dai, Y. (2023). Optimal aborting policy for shock exposed missions with random rescue time. *Reliability Engineering & System Safety*, 233:109094.

Levitin, G., Xing, L., and Dai, Y. (2024a). Multi-attempt missions with multiple rescue options. *Reliability Engineering & System Safety*, 248:110168.

Levitin, G., Xing, L., and Dai, Y. (2024b). A new self-adaptive mission aborting policy for systems operating in uncertain random shock environment. *Reliability Engineering & System Safety*, 248:110184.

Levitin, G., Xing, L., and Dai, Y. (2024c). Optimal attempt scheduling and aborting in heterogenous system performing asynchronous multi-attempt mission. *Reliability Engineering & System Safety*, 251:110335.

Levitin, G., Xing, L., and Dai, Y. (2024d). Optimal system loading and aborting in additive multi-attempt missions. *Reliability Engineering & System Safety*, 251:110315.

Levitin, G., Xing, L., and Dai, Y. (2024e). Optimal task aborting and sequencing in time constrained multi-task multi-attempt missions. *Reliability Engineering & System Safety*, 241:109702.

Levitin, G., Xing, L., and Dai, Y. (2024f). Optimizing time-varying performance and mission aborting policy in resource constrained missions. *Reliability Engineering & System Safety*, 245:110011.

Levitin, G., Xing, L., and Luo, L. (2019). Influence of failure propagation on mission abort policy in heterogeneous warm standby systems. *Reliability Engineering & System Safety*, 183:29–38.

Li, H., Zhao, W., Zhang, Y., and Zio, E. (2020). Remaining useful life prediction using multi-scale deep convolutional neural network. *Applied Soft Computing*, 89:106113.

Li, J., Li, X., and He, D. (2019a). A directed acyclic graph network combined with cnn and lstm for remaining useful life prediction. *IEEE Access*, 7:75464–75475.

Li, M., Jiang, X., Carroll, J., and Negenborn, R. R. (2022). A multi-objective maintenance strategy optimization framework for offshore wind farms considering uncertainty. *Applied Energy*, 321:119284.

Li, W. and Zhang, C. (2020). A bi-objective optimization approach for the maintenance planning of networked systems. *Quality and Reliability Engineering International*, 36(4):1364–1385.

Li, X., Ding, Q., and Sun, J.-Q. (2018). Remaining useful life estimation in prognostics using deep convolution neural networks. *Reliability Engineering & System Safety*, 172:1–11.

Li, X., Ran, Y., Zhang, G., and Yu, H. (2021). Selective maintenance of multi-state series systems considering maintenance quality uncertainty and failure effects. *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering*, 235(5):1363–1374.

Li, Y., Zhang, C., Jia, C., Li, X., and Zhu, Y. (2019b). Joint optimization of workforce scheduling and routing for restoring a disrupted critical infrastructure. *Reliability Engineering & System Safety*, 191:106551.

Liao, T.-Y., Hu, T.-Y., and Ko, Y.-N. (2018). A resilience optimization model for transportation networks under disasters. *Natural hazards*, 93(1):469–489.

Lin, D., Zuo, M. J., and Yam, R. C. (2000a). General sequential imperfect preventive maintenance models. *International Journal of reliability, Quality and safety Engineering*, 7(03):253–266.

Lin, D., Zuo, M. J., and Yam, R. C. M. (2000b). General sequential imperfect preventive maintenance models. *International Journal of reliability, Quality and safety Engineering*, 7(3):253–266.

Liu, B., Huang, H., and Deng, Q. (2022a). On optimal condition based task termination policy for phased task systems. *Reliability Engineering & System Safety*, 221:108338.

Liu, L. and Yang, J. (2023). A dynamic mission abort policy for the swarm executing missions and its solution method by tailored deep reinforcement learning. *Reliability Engineering & System Safety*, 234:109149.

Liu, L., Yang, J., Kong, X., and Xiao, Y. (2022b). Multi-mission selective maintenance and repairpersons assignment problem with stochastic durations. *Reliability Engineering & System Safety*, 219:108209.

Liu, L., Yang, J., and Yan, B. (2024). A dynamic mission abort policy for transportation systems with stochastic dependence by deep reinforcement learning. *Reliability Engineering & System Safety*, 241:109682.

Liu, W., Song, Z., Ouyang, M., and Li, J. (2020a). Recovery-based seismic resilience enhancement strategies of water distribution networks. *Reliability Engineering & System Safety*, 203:107088.

Liu, Y., Chen, Y., and Jiang, T. (2018). On sequence planning for selective maintenance of multi-state systems under stochastic maintenance durations. *European Journal of Operational Research*, 268(1):113–127.

Liu, Y., Chen, Y., and Jiang, T. (2020b). Dynamic selective maintenance optimization for multi-state systems over a finite horizon: A deep reinforcement learning approach. *European Journal of Operational Research*, 283(1):166–181.

Liu, Y. and Huang, H.-Z. (2010). Optimal selective maintenance strategy for multi-state systems under imperfect maintenance. *IEEE Transactions on Reliability*, 59(2):356–367.

Liu, Y., Huang, H.-Z., and Jiang, T. (2023). *Selective Maintenance Modelling and Optimization: Basic Methods and Some Recent Advances*. Springer.

Liu, Y. and Qian, X. (2022). Selective maintenance optimization with stochastic break duration based on reinforcement learning. *Eksploatacja i Niezawodność*, 24(4).

López-Santana, E., Akhavan-Tabatabaei, R., Dieulle, L., Labadie, N., and Medaglia, A. L. (2016). On the combined maintenance and routing optimization problem. *Reliability Engineering & System Safety*, 145:199–214.

Lust, T., Roux, O., and Riane, F. (2009). Exact and heuristic methods for the selective maintenance problem. *European journal of operational research*, 197(3):1166–1177.

Ma, W., Zhang, Q., Xiahou, T., Liu, Y., and Jia, X. (2023). Integrated selective maintenance and task assignment optimization for multi-state systems executing multiple missions. *Reliability Engineering & System Safety*, 237:109330.

Maaroufi, G., Chelbi, A., and Rezg, N. (2013). Optimal selective renewal policy for systems subject to propagated failures with global effect and failure isolation phenomena. *Reliability Engineering & System Safety*, 114:61–70.

Maillart, L. M., Cassady, C. R., Rainwater, C., and Schneider, K. (2009). Selective maintenance decision-making over extended planning horizons. *IEEE Transactions on Reliability*, 58(3):462–469.

Malik, M. A. K. (1979). Reliable preventive maintenance scheduling. *AIIE transactions*, 11(3):221–228.

Mathlouthi, I., Gendreau, M., and Potvin, J.-Y. (2018). Mixed integer linear programming for a multi-attribute technician routing and scheduling problem. *INFOR: Information Systems and Operational Research*, 56(1):33–49.

McCall, J. (2005). Genetic algorithms for modelling and optimisation. *Journal of computational and Applied Mathematics*, 184(1):205–222.

Meng, S., Xing, L., and Levitin, G. (2024a). Activation delay and aborting policy minimizing expected losses in consecutive attempts having cumulative effect on mission success. *Reliability Engineering & System Safety*, 247:110078.

Meng, S., Xing, L., and Levitin, G. (2024b). Optimizing component activation and operation aborting in missions with consecutive attempts and common abort command. *Reliability Engineering & System Safety*, 243:109842.

Mild, P. and Salo, A. (2009). Combining a multiattribute value function with an optimization model: An application to dynamic resource allocation for infrastructure maintenance. *Decision Analysis*, 6(3):139–152.

Minkel, J. (2008). The 2003 northeast blackout–five years later. *Scientific American*, 13.

Mitici, M., de Pater, I., Barros, A., and Zeng, Z. (2023). Dynamic predictive maintenance for multiple components using data-driven probabilistic rul prognostics: The case of turbofan engines. *Reliability Engineering & System Safety*, 234:109199.

Morshedlou, N., Barker, K., González, A. D., and Ermagun, A. (2021). A heuristic approach to an interdependent restoration planning and crew routing problem. *Computers & Industrial Engineering*, 161:107626.

Morshedlou, N., González, A. D., and Barker, K. (2018). Work crew routing problem for infrastructure network restoration. *Transportation Research Part B: Methodological*, 118:66–89.

Myers, A. (2009). Probability of loss assessment of critical k-out-of-n: G systems having a mission abort policy. *IEEE Transactions on Reliability*, 58(4):694–701.

Nakagawa, T. (1988a). Sequential imperfect preventive maintenance policies. *IEEE Transactions on Reliability*, 37(3):295–308.

Nakagawa, T. (1988b). Sequential imperfect preventive maintenance policies. *IEEE Transactions on Reliability*, 37(3):295–298.

Namuduri, S., Narayanan, B. N., Davuluru, V. S. P., Burton, L., and Bhansali, S. (2020). Deep learning methods for sensor based predictive maintenance and future perspectives for electrochemical sensors. *Journal of The Electrochemical Society*, 167(3):037552.

Nayak, M. A. and Turnquist, M. A. (2016). Optimal recovery from disruptions in water distribution networks. *Computer-Aided Civil and Infrastructure Engineering*, 31(8):566–579.

Nguyen, K.-A., Do, P., and Grall, A. (2017). Joint predictive maintenance and inventory strategy for multi-component systems using birnbaum's structural importance. *Reliability Engineering & System Safety*, 168:249–261.

Nguyen, K. T. and Medjaher, K. (2019). A new dynamic predictive maintenance framework using deep learning for failure prognostics. *Reliability Engineering & System Safety*, 188:251–262.

Niu, J., Liu, C., Zhang, L., and Liao, Y. (2019). Remaining useful life prediction of machining tools by 1d-cnn lstm network. In *2019 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 1056–1063. IEEE.

Nurre, S. G., Cavdaroglu, B., Mitchell, J. E., Sharkey, T. C., and Wallace, W. A. (2012). Restoring infrastructure systems: An integrated network design and scheduling (inds) problem. *European journal of operational research*, 223(3):794–806.

Ochelska-Mierzejewska, J., Poniszewska-Marańda, A., and Marańda, W. (2021). Selected genetic algorithms for vehicle routing problem solving. *Electronics*, 10(24):3147.

O'Neil, R., Diallo, C., and Khatab, A. (2022a). A novel predictive selective maintenance strategy using deep learning and mathematical programming. *IFAC-PapersOnLine*, 55(10):1207–1212.

O'Neil, R., Diallo, C., Khatab, A., and Aghezzaf, E.-H. (2022b). A hybrid column-generation and genetic algorithm approach for solving large-scale multimission selective maintenance problems in serial k-out-of-n: G systems. *International Journal of Production Research*, pages 1–17.

O'Neil, R., Diallo, C., Khatab, A., and Aghezzaf, E.-H. (2023). A hybrid column-generation and genetic algorithm approach for solving large-scale multimission selective maintenance problems in serial k-out-of-n: G systems. *International Journal of Production Research*, 61(9):3071–3087.

O'Neil, R., Khatab, A., and Diallo, C. (2024). Optimal predictive selective maintenance for fleets of mission-oriented systems. *International Journal of Production Research*, pages 1–28.

Ouyang, M. (2017). A mathematical framework to optimize resilience of interdependent critical infrastructure systems under spatially localized attacks. *European Journal of Operational Research*, 262(3):1072–1084.

O'Neil, R., Diallo, C., and Khatab, A. (2024). Optimal selective maintenance for complex systems under stochastic maintenance and break durations. *IFAC-PapersOnLine*, 58(19):1252–1257. 18th IFAC Symposium on Information Control Problems in Manufacturing INCOM 2024.

O'Neil, R., Khatab, A., Diallo, C., and Venkatadri, U. (2023). Optimal joint maintenance and orienteering strategy for complex mission-oriented systems: A case study in offshore wind energy. *Computers & Operations Research*, 149:106020.

Pandey, M., Zuo, M. J., and Moghaddass, R. (2016). Selective maintenance scheduling over a finite planning horizon. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 230(2):162–177.

Pandey, M., Zuo, M. J., Moghaddass, R., and Tiwari, M. (2013). Selective maintenance for binary systems under imperfect repair. *Reliability Engineering & System Safety*, 113:42–51.

Paraskevopoulos, D. C., Laporte, G., Repoussis, P. P., and Tarantilis, C. D. (2017). Resource constrained routing and scheduling: Review and research prospects. *European Journal of Operational Research*, 263(3):737–754.

Patel, J., O'Neil, R., Khatab, A., Diallo, C., and Venkatadri, U. (2023). Optimal selective maintenance strategy with repairperson workload balancing. In *2023 IEEE International Conference on Networking, Sensing and Control (ICNSC)*, volume 1, pages 1–6, Marseille, France.

Pellerin, R., Perrier, N., and Berthaut, F. (2020). A survey of hybrid metaheuristics for the resource-constrained project scheduling problem. *European Journal of Operational Research*, 280(2):395–416.

Peng, C., Chen, Y., Chen, Q., Tang, Z., Li, L., and Gui, W. (2021). A remaining useful life prognosis of turbofan engine using temporal and spatial feature fusion. *Sensors*, 21(2):418.

Peng, K., Jiao, R., Dong, J., and Pi, Y. (2019). A deep belief network based health indicator construction and remaining useful life prediction using improved particle filter. *Neurocomputing*, 361:19–28.

Puzyrev, V. and Duuring, P. (2025). Uncertainty quantification of geochemical data imputation using monte carlo dropout. *Journal of Geochemical Exploration*, page 107695.

Qiu, Q. and Cui, L. (2019). Gamma process based optimal mission abort policy. *Reliability Engineering & System Safety*, 190:106496.

Rajagopalan, R. and Cassady, C. R. (2004). Solving selective maintenance problems. In *IIE Annual Conference. Proceedings*, pages 1–6, Norcross. Institute of Industrial and Systems Engineers (IISE).

Rajagopalan, R. and Cassady, C. R. (2006). An improved selective maintenance solution approach. *Journal of Quality in Maintenance Engineering*.

Raknes, N. T., Ødeskaug, K., Stålhane, M., and Hvattum, L. M. (2017). Scheduling of maintenance tasks and routing of a joint vessel fleet for multiple offshore wind farms. *Journal of Marine Science and Engineering*, 5(1).

Ramirez-Marquez, J. E. and Coit, D. W. (2005). A monte-carlo simulation approach for approximating multi-state two-terminal reliability. *Reliability Engineering & System Safety*, 87(2):253–264.

Ravaghi Ardabili, H. A., Haghifam, M.-R., and Abedi, S. M. (2021). A probabilistic reliability-centred maintenance approach for electrical distribution networks. *IET Generation, Transmission & Distribution*, 15(6):1070–1080.

Ren, L., Dong, J., Wang, X., Meng, Z., Zhao, L., and Deen, M. J. (2020). A data-driven auto-cnn-lstm prediction model for lithium-ion battery remaining useful life. *IEEE Transactions on Industrial Informatics*, 17(5):3478–3487.

Renschler, C. S., Frazier, A. E., Arendt, L. A., Cimellaro, G. P., Reinhorn, A. M., and Bruneau, M. (2010). *A framework for defining and measuring resilience at the community scale: The PEOPLES resilience framework*. MCEER Buffalo.

Rice, W., Cassady, C., and Nachlas, J. (1998). Optimal maintenance plans under limited maintenance time. In *Proceedings of the seventh industrial engineering research conference*, pages 1–3.

Rice, W. F. (1999). *Optimal selective maintenance decisions for series systems*. PhD thesis, Mississippi State University. Department of Industrial Engineering.

Röckmann, C., Lagerveld, S., and Stavenuiter, J. (2017). Operation and maintenance costs of offshore wind farms and potential multi-use platforms in the dutch north sea. In *Aquaculture Perspective of Multi-Use Sites in the Open Ocean*, pages 97–113. Springer, Cham.

Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6):386.

Ruan, D., Wu, Y., and Yan, J. (2021). Remaining useful life prediction for aero-engine based on lstm and cnn. In *2021 33rd Chinese Control and Decision Conference (CCDC)*, pages 6706–6712. IEEE.

Ruiz, C., Pohl, E., and Liao, H. (2020). Selective maintenance of multi-component systems with multiple failure modes. In *2020 Annual Reliability and Maintainability Symposium (RAMS)*, pages 1–6, Palm Springs, USA. IEEE.

Sanci, E. and Daskin, M. S. (2019). Integrating location and network restoration decisions in relief networks under uncertainty. *European Journal of Operational Research*, 279(2):335–350.

Sathurshan, M., Saja, A., Thamboo, J., Haraguchi, M., and Navaratnam, S. (2022). Resilience of critical infrastructure systems: a systematic literature review of measurement frameworks. *Infrastructures*, 7(5):67.

Saxena, A., Goebel, K., Simon, D., and Eklund, N. (2008). Damage propagation modeling for aircraft engine run-to-failure simulation. In *2008 international conference on prognostics and health management*, pages 1–9. IEEE.

Schneider, K. and Cassady, C. R. (2004). Fleet performance under selective maintenance. In *Annual symposium reliability and maintainability, 2004-RAMS*, pages 571–576, Los Angeles, USA. IEEE.

Schneider, K. and Cassady, C. R. (2015). Evaluation and comparison of alternative fleet-level selective maintenance models. *Reliability Engineering & System Safety*, 134:178–187.

Schrotenboer, A. H., uit het Broek, M. A., Jargalsaikhan, B., and Roodbergen, K. J. (2018). Coordinating technician allocation and maintenance routing for offshore wind farms. *Computers & Operations Research*, 98:185–197.

Schrotenboer, A. H., Ursavas, E., and Vis, I. F. (2020a). Mixed integer programming models for planning maintenance at offshore wind farms under uncertainty. *Transportation Research Part C: Emerging Technologies*, 112:180–202.

Schrotenboer, A. H., Ursavas, E., and Vis, I. F. (2020b). Mixed integer programming models for planning maintenance at offshore wind farms under uncertainty. *Transportation Research Part C: Emerging Technologies*, 112:180–202.

Schrotenboer, A. H., Ursavas, E., and Vis, I. F. A. (2019). A branch-and-price-and-cut algorithm for resource-constrained pickup and delivery problems. *Transportation Science*, 53(4):1001–1022.

Serradilla, O., Zugasti, E., Rodriguez, J., and Zurutuza, U. (2022). Deep learning models for predictive maintenance: a survey, comparison, challenges and prospects. *Applied Intelligence*, pages 1–31.

Shahmoradi-Moghadam, H., Safaei, N., and Sadjadi, S. J. (2021). Robust maintenance scheduling of aircraft fleet: A hybrid simulation-optimization approach. *IEEE Access*, 9:17854–17865.

Shahraki, A. F., Yadav, O. P., and Vogiatzis, C. (2020). Selective maintenance optimization for multi-state systems considering stochastically dependent components and stochastic imperfect maintenance actions. *Reliability Engineering & System Safety*, 196:106738.

Shamayleh, A., Awad, M., and Abdulla, A. (2020). Criticality-based reliability-centered maintenance for healthcare. *Journal of Quality in Maintenance Engineering*, 26(2):311–334.

Sharma, N., Tabandeh, A., and Gardoni, P. (2020). Regional resilience analysis: A multiscale approach to optimize the resilience of interdependent infrastructure. *Computer-Aided Civil and Infrastructure Engineering*, 35(12):1315–1330.

Sharma, P., Kulkarni, M. S., and Yadav, V. (2017). A simulation based optimization approach for spare parts forecasting and selective maintenance. *Reliability Engineering & System Safety*, 168:274–289.

Shoorkand, H. D., Nourelfath, M., and Hajji, A. (2024). A hybrid cnn-lstm model for joint optimization of production and imperfect predictive maintenance planning. *Reliability Engineering & System Safety*, 241:109707.

Sieb, P., Michelmann, J., Flöter, F., and Wicke, K. (2023). Identifying challenges in maintenance planning for on-demand uam fleets using agent-based simulations. *CEAS Aeronautical Journal*, 14(3):637–660.

Soh, S. and Rai, S. (2005). An efficient cutset approach for evaluating communication-network reliability with heterogeneous link-capacities. *IEEE Transactions on reliability*, 54(1):133–144.

Song, T., Liu, C., Wu, R., Jin, Y., and Jiang, D. (2022). A hierarchical scheme for remaining useful life prediction with long short-term memory networks. *Neurocomputing*, 487:22–33.

Stålhane, M., Hvattum, L. M., and Skaar, V. (2015). Optimization of routing and scheduling of vessels to perform maintenance at offshore wind farms. *Energy Procedia*, 80:92–99.

Su, C., Li, L., and Wen, Z. (2020). Remaining useful life prediction via a variational autoencoder and a time-window-based sequence neural network. *Quality and Reliability Engineering International*, 36(5):1639–1656.

Sun, J., Sun, Y., Chen, D., and Liu, H. (2023). An integrated life-cycle maintenance modeling and simulation framework for aero engine fleet management decision support. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 237(3):604–616.

Sun, Y. and Sun, Z. (2021). Selective maintenance on a multi-state transportation system considering maintenance sequence arrangement. *IEEE Access*, 9:70048–70060.

Syan, C. S. and Ramsoobag, G. (2019). Maintenance applications of multi-criteria optimization: A review. *Reliability Engineering & System Safety*, 190:106520.

Tai, A. H. and Chan, L.-Y. (2010). Maintenance models for a continuously degrading system. *Computers & industrial engineering*, 58(4):578–583.

Talbi, E.-G. (2009). *Metaheuristics: from design to implementation*. John Wiley & Sons.

Tambe, P. P. (2022). Selective maintenance optimization of a multi-component system based on simulated annealing algorithm. *Procedia Computer Science*, 200:1412–1421.

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

Tripathi, A. and Hari Prasad, M. (2024). Rcm based optimization of maintenance strategies for marine diesel engine using genetic algorithms. *International Journal of System Assurance Engineering and Management*, pages 1–19.

Trivedi, K. S., Kim, D. S., Roy, A., and Medhi, D. (2009). Dependability and security models. In *2009 7th International Workshop on Design of Reliable Communication Networks*, pages 11–20. IEEE.

Tseremoglou, I. and Santos, B. F. (2024). Condition-based maintenance scheduling of an aircraft fleet under partial observability: A deep reinforcement learning approach. *Reliability Engineering & System Safety*, 241:109582.

Tsiligirides, T. (1984). Heuristic methods applied to orienteering. *Journal of the Operational Research Society*, 35(9):797–809.

Urbani, M., Brunelli, M., and Punkka, A. (2023). An approach for bi-objective maintenance scheduling on a networked system with limited resources. *European Journal of Operational Research*, 305(1):101–113.

Usuga-Cadavid, J. P., Lamouri, S., Grabot, B., and Fortin, A. (2022). Using deep learning to value free-form text data for predictive maintenance. *International Journal of Production Research*, 60(14):4548–4575.

Vansteenwegen, P., Souffriau, W., and Van Oudheusden, D. (2011). The orienteering problem: A survey. *European Journal of Operational Research*, 209(1):1–10.

Verbeeck, C., Vansteenwegen, P., and Aghezzaf, E.-H. (2017). The time-dependent orienteering problem with time windows: a fast ant colony system. *Annals of Operations Research*, 254(1):481–505.

Wang, C., Lu, N., Cheng, Y., and Jiang, B. (2019). A data-driven aero-engine degradation prognostic strategy. *IEEE transactions on cybernetics*, 51(3):1531–1541.

Wang, H.-P., Duan, F.-H., Wang, X.-L., and He, Y.-L. (2023a). Selective maintenance of multistate systems considering the random uncertainty of the system mission period and mission breaks. *Arabian Journal for Science and Engineering*, 48(5):7059–7075.

Wang, P., Lu, S., Cheng, H., Liu, L., and Pei, F. (2023b). Preemptive multi-skill resource-constrained project scheduling of marine power equipment maintenance tasks. *Journal of Intelligent & Fuzzy Systems*, pages 1–20.

Wang, Y., Zhao, O., and Zhang, L. (2024). Multiplex networks in resilience modeling of critical infrastructure systems: A systematic review. *Reliability Engineering & System Safety*, page 110300.

Witteman, M., Deng, Q., and Santos, B. F. (2021). A bin packing approach to solve the aircraft maintenance task allocation problem. *European Journal of Operational Research*, 294(1):365–376.

Woods, D. D. (2015). Four concepts for resilience and the implications for the future of resilience engineering. *Reliability Engineering & System Safety*, 141:5–9.

Wu, C., Zhao, X., Qiu, Q., and Sun, J. (2021). Optimal mission abort policy for k-out-of-n: F balanced systems. *Reliability Engineering & System Safety*, 208:107398.

Wu, J. and Wang, P. (2021). Risk-averse optimization for resilience enhancement of complex engineering systems under uncertainties. *Reliability Engineering & System Safety*, 215:107836.

Wu, Y., Chen, Z., Dang, J., Chen, Y., Zhao, X., and Zha, L. (2022). Allocation of defensive and restorative resources in electric power system against consecutive multi-target attacks. *Reliability Engineering & System Safety*, 219:108199.

Xia, J., Feng, Y., Lu, C., Fei, C., and Xue, X. (2021). Lstm-based multi-layer self-attention method for remaining useful life estimation of mechanical systems. *Engineering Failure Analysis*, 125:105385.

Xiang, F., Zhang, Y., Zhang, S., Wang, Z., Qiu, L., and Choi, J.-H. (2024). Bayesian gated-transformer model for risk-aware prediction of aero-engine remaining useful life. *Expert Systems with Applications*, 238:121859.

Xu, L. D., Xu, E. L., and Li, L. (2018). Industry 4.0: state of the art and future trends. *International Journal of Production Research*, 56(8):2941–2962.

Xu, Q.-z., Guo, L.-m., Shi, H.-p., and Wang, N. (2016). Selective maintenance problem for series–parallel system under economic dependence. *Defence technology*, 12(5):388–400.

Xu, Q.-Z., Guo, L.-M., Wang, N., and Fei, R. (2015). Recent advances in selective maintenance from 1998 to 2014. *Journal of Donghua University (English Edition)*, 32(6):986–994.

Yang, B., Liu, R., and Zio, E. (2019). Remaining useful life prediction based on a double-convolutional neural network architecture. *IEEE Transactions on Industrial Electronics*, 66(12):9521–9530.

Yang, D., Wang, H., Feng, Q., Ren, Y., Sun, B., and Wang, Z. (2018). Fleet-level selective maintenance problem under a phased mission scheme with short breaks: A heuristic sequential game approach. *Computers & Industrial Engineering*, 119:404–415.

Yang, L., Chen, Y., Qiu, Q., and Wang, J. (2022). Risk control of mission-critical systems: Abort decision-makings integrating health and age conditions. *IEEE Transactions on Industrial Informatics*, 18(10):6887–6894.

Yin, M., Liu, Y., Liu, S., Chen, Y., and Yan, Y. (2023). Scheduling heterogeneous repair channels in selective maintenance of multi-state systems with maintenance duration uncertainty. *Reliability Engineering & System Safety*, 231:108977.

Yu, W., Kim, I. Y., and Mechefske, C. (2019). Remaining useful life estimation using a bidirectional recurrent neural network based autoencoder scheme. *Mechanical Systems and Signal Processing*, 129:764–780.

Yu, W., Kim, I. Y., and Mechefske, C. (2020). An improved similarity-based prognostic algorithm for rul estimation using an rnn autoencoder scheme. *Reliability Engineering & System Safety*, 199:106926.

Zeng, J. and Liang, Z. (2023). A dynamic predictive maintenance approach using probabilistic deep learning for a fleet of multi-component systems. *Reliability Engineering & System Safety*, page 109456.

Zhang, C., Chen, R., Wang, S., Dui, H., and Zhang, Y. (2022a). Resilience efficiency importance measure for the selection of a component maintenance strategy to improve system performance recovery. *Reliability Engineering & System Safety*, 217:108070.

Zhang, C., Ramirez-Marquez, J. E., and Sanseverino, C. M. R. (2011). A holistic method for reliability performance assessment and critical components detection in complex networks. *IIE Transactions*, 43(9):661–675.

Zhang, L., Chen, X., Khatab, A., An, Y., and Feng, X. (2023a). Joint optimization of selective maintenance and repairpersons assignment problem for mission-oriented systems operating under s- dependent competing risks. *Reliability Engineering & System Safety*, page 109796.

Zhang, L., Chen, X., Khatab, A., An, Y., and Feng, X. (2024a). Joint optimization of selective maintenance and repairpersons assignment problem for mission-oriented systems operating under s-dependent competing risks. *Reliability Engineering & System Safety*, 242:109796.

Zhang, L. and Zhang, J. (2020). A data-driven maintenance framework under imperfect inspections for deteriorating systems using multitask learning-based status prognostics. *IEEE Access*, 9:3616–3629.

Zhang, Q., Liu, Q., and Ye, Q. (2024b). An attention-based temporal convolutional network method for predicting remaining useful life of aero-engine. *Engineering Applications of Artificial Intelligence*, 127:107241.

Zhang, Q., Liu, Y., Xiahou, T., and Huang, H.-Z. (2023b). A heuristic maintenance scheduling framework for a military aircraft fleet under limited maintenance capacities. *Reliability Engineering & System Safety*, 235:109239.

Zhang, S., Ohlmann, J. W., and Thomas, B. W. (2014). A priori orienteering with time windows and stochastic wait times at customers. *European Journal of Operational Research*, 239(1):70–79.

Zhang, X., Chen, J., Han, B., and Li, J. (2019). Multi-mission selective maintenance modelling for multistate systems over a finite time horizon. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 233(6):1040–1059.

Zhang, Y., Li, C., Su, X., Cui, R., and Wan, B. (2023c). A baseline-reactive scheduling method for carrier-based aircraft maintenance tasks. *Complex & Intelligent Systems*, 9(1):367–397.

Zhang, Y., Xiong, R., He, H., and Pecht, M. G. (2018). Long short-term memory recurrent neural network for remaining useful life prediction of lithium-ion batteries. *IEEE Transactions on Vehicular Technology*, 67(7):5695–5705.

Zhang, Z., Ji, T., and Wei, H.-H. (2022b). Dynamic emergency inspection routing and restoration scheduling to enhance the post-earthquake resilience of a highway–bridge network. *Reliability Engineering & System Safety*, 220:108282.

Zhao, J. and Zeng, J.-C. (2016). Maintenance strategy for stochastic selective maintenance of a two-state system. *International Journal of Wireless and Mobile Computing*, 11(4):302–308.

Zhao, R., Yan, R., Wang, J., and Mao, K. (2017). Learning to monitor machine health with convolutional bi-directional lstm networks. *Sensors*, 17(2):273.

Zhao, X., Li, R., Cao, S., and Qiu, Q. (2023a). Joint modeling of loading and mission abort policies for systems operating in dynamic environments. *Reliability Engineering & System Safety*, 230:108948.

Zhao, X., Liu, H., Wu, Y., and Qiu, Q. (2023b). Joint optimization of mission abort and system structure considering dynamic tasks. *Reliability Engineering & System Safety*, 234:109128.

Zhao, X., Lv, Z., Qiu, Q., and Wu, Y. (2023c). Designing two-level rescue depot location and dynamic rescue policies for unmanned vehicles. *Reliability Engineering & System Safety*, 233:109119.

Zhao, X., Sun, J., Qiu, Q., and Chen, K. (2021). Optimal inspection and mission abort policies for systems subject to degradation. *European Journal of Operational Research*, 292(2):610–621.

Zhao, X., Wang, X., Dai, Y., and Qiu, Q. (2024). Joint optimization of loading, mission abort and rescue site selection policies for uav. *Reliability Engineering & System Safety*, 244:109955.

Zhao, Z., Xiao, B., Wang, N., Yan, X., and Ma, L. (2019). Selective maintenance optimization for a multi-state system considering human reliability. *Symmetry*, 11(5):652.

Zhou, J., Coit, D. W., Felder, F. A., and Tsianikas, S. (2023a). Combined optimization of system reliability improvement and resilience with mixed cascading failures in dependent network systems. *Reliability Engineering & System Safety*, 237:109376.

Zhou, K.-L., Cheng, D.-J., Zhang, H.-B., Hu, Z.-t., and Zhang, C.-Y. (2023b). Deep learning-based intelligent multilevel predictive maintenance framework considering comprehensive cost. *Reliability Engineering & System Safety*, 237:109357.

Zhu, H., Liu, F., Shao, X., Liu, Q., and Deng, Y. (2011). A cost-based selective maintenance decision-making method for machining line. *Quality and Reliability Engineering International*, 27(2):191–201.

Zhu, X., Zhu, X., Yan, R., and Peng, R. (2021). Optimal routing, aborting and hitting strategies of uavs executing hitting the targets considering the defense range of targets. *Reliability Engineering & System Safety*, 215:107811.

Zhuang, L., Xu, A., and Wang, X.-L. (2023). A prognostic driven predictive maintenance framework based on bayesian deep learning. *Reliability Engineering & System Safety*, 234:109181.

Zio, E. (2009). Reliability engineering: Old problems and new challenges. *Reliability Engineering & System Safety*, 94(2):125–141.

Zuo, M. J., Tian, Z., and Huang, H.-Z. (2007). An efficient method for reliability evaluation of multistate networks given all minimal path vectors. *IIE transactions*, 39(8):811–817.

Résumé

La vie moderne repose sur des systèmes de production, de distribution et de logistique complexes et fortement interconnectés pour la production de l'énergie, des biens et des services. Ces systèmes sont conçus pour fonctionner avec un minimum d'interruptions, garantissant une fiabilité exceptionnelle et une grande réactivité face aux perturbations causées par des événements imprévus tels que les conditions météorologiques extrêmes, les catastrophes naturelles et les pandémies. Leur défaillance peut s'avérer extrêmement coûteuse, rendant de plus en plus crucial pour les organisations de gérer efficacement leurs équipements, d'améliorer la disponibilité de leurs actifs et de réduire leurs coûts opérationnels. Bon nombre de ces systèmes fonctionnent selon un cycle alterné de phases de fonctionnement (mission) et d'arrêt programmé. La maintenance sélective (MS) est largement utilisée dans ces contextes et vise à optimiser les performances des actifs lors de la mission suivante en allouant des ressources de maintenance limitées.

Cette thèse explore quatre thèmes clés, chacun abordant des lacunes critiques dans la modélisation et l'application de la maintenance sélective (MS). Le premier thème explore l'application de la MS aux flottes de systèmes géographiquement distribués et aux systèmes en réseau comportant des composants géographiquement dispersés. Les modèles développés dans ce thème ont une large applicabilité et nécessitent le développement d'algorithmes de décomposition adaptés. Le deuxième thème développe un cadre de MS prédictif basé sur les données, où des réseaux neuronaux profonds sont utilisés pour la prédiction de la durée de vie résiduelle (RUL). Afin de quantifier l'incertitude des prédictions, une approche bayésienne est appliquée pour générer une distribution de la RUL, contrairement à une prédiction ponctuelle. Le troisième thème établit un lien théorique entre la planification de la maintenance sélective et le problème d'ordonnancement de projet sous contraintes de ressources. Ce lien conduit au développement d'un modèle de maintenance plus pratique, capable de gérer la planification de la maintenance sous contraintes de ressources et avec des capacités de maintenance multiples. Enfin, le quatrième thème explore l'application de la maintenance sélective à des systèmes critiques mettant en œuvre des politiques d'interruption et d'abandon d'une mission. Un modèle d'optimisation est développé afin de trouver un équilibre entre le succès de la mission et la survie du système. Les modèles développés sont appliqués à plusieurs exemples et études de cas illustratifs, démontrant ainsi leur applicabilité et leur valeur ajoutée.

Mots-clés: Maintenance intelligente, Apprentissage automatique, Planification de la maintenance, Abandon de mission, Optimisation, Maintenance sélective prédictive, Résilience, Ingénierie de la fiabilité, Routage des techniciens de maintenance, Maintenance de flotte, Pronostic fondé sur les données.

Abstract

Modern life relies on complex, highly interconnected production, distribution, and logistics systems for energy, goods, and services. These systems are designed to operate with minimal interruption, ensuring exceptionally high reliability and rapid responsiveness to disruptions caused by unforeseen events such as extreme weather, natural disasters, and pandemics. Their failure can be extremely costly, making it increasingly crucial for organizations to efficiently manage equipment, enhance asset availability, and reduce operational costs. Many of these systems follow an alternating cycle of missions and breaks. Selective maintenance (SM) is widely used in these contexts and aims to optimize asset performance during the next mission by allocating limited maintenance resources.

This dissertation explores four key themes, each addressing critical knowledge gaps in the modeling and application of SM. The first theme explores the application of SM to fleets of geographically distributed systems and networked systems with geographically dispersed components. The models developed in this theme have far-ranging applicability and require the development of tailored decomposition algorithms. The second theme develops a data-driven predictive SM framework where deep neural networks are utilized for remaining useful life (RUL) prediction. To quantify uncertainty in predictions, a

Bayesian approach is applied to generate a distribution of RUL opposed to a single point prediction. The third theme establishes a theoretical link between the SMP and the resource-constrained project scheduling problem (RCPSP). This connection leads to the development of a more practical maintenance model that is able to handle maintenance planning under resource constraints and multiple maintenance capacities. Finally, the fourth theme explores the application of SM to mission-critical systems implementing mission abort policies. An optimization model is developed that seeks to strike a balance between mission success and system survival probability. The developed models are applied to several illustrative examples and case studies, demonstrating their applicability and added value.

Keywords: Intelligent Maintenance, Machine Learning, Maintenance Planning, Mission Abort, Optimization, Predictive Selective Maintenance, Resilience, Reliability Engineering, Maintenance Technician Routing, Fleet Maintenance, Data-driven Prognostics

Résumé en français de la thèse de doctorat

La vie moderne dépend de systèmes complexes et fortement interconnectés de production, de distribution et de logistique pour l'énergie, les biens et les services. Ces systèmes sont conçus pour fonctionner avec peu d'interruptions, une fiabilité extrêmement élevée et un haut niveau de disponibilité afin de mieux répondre aux perturbations causées par des événements imprévus tels que les conditions météorologiques, les catastrophes naturelles et les pandémies. Leur défaillance peut être extrêmement coûteuse, ce qui rend de plus en plus crucial pour les organisations de gérer et d'entretenir efficacement les équipements et les machines, d'améliorer la disponibilité des actifs et de réduire les coûts.

L'émergence de l'Industrie 4.0, qui englobe les technologies manufacturières de nouvelle génération et les usines intelligentes, intègre des systèmes embarqués, des capteurs, la communication machine-a-machine, l'IIOT et les systèmes cyberphysiques (CPS) afin de combler le fossé entre les mondes virtuel et physique (Xu et al., 2018; Kusiak, 2018). Ces avancées offrent le potentiel d'améliorer la production et la distribution de biens personnalisés à des coûts comparables à ceux de la production de masse. L'intégration des chaînes d'approvisionnement, des systèmes de production et des réseaux logistiques aboutit à des systèmes entièrement intégrés et réactifs (Kaynak, 2007; Ivanov et al., 2016). Cependant, malgré ces avancées, assurer le fonctionnement continu de tels réseaux demeure essentiel; leurs défaillances peuvent, en effet, être très perturbatrices et coûteuses. Par conséquent, la surveillance de l'état des systèmes, la prédiction des défaillances et la mise en œuvre de stratégies de maintenance proactives sont indispensables pour maintenir des performances élevées dans ces réseaux intégrés.

De nombreux systèmes de production, de distribution et de logistique fonctionnent selon une séquence alternée de missions et de pauses ou d'arrêts programmés. Ces systèmes, appelés systèmes orientés mission (MOS), sont courants dans divers secteurs, notamment la défense, la pétrochimie, les transports et la logistique. La maintenance sélective (MS) est largement appliquée dans ces contextes afin d'évaluer et d'améliorer la fiabilité du système ou sa durée de vie résiduelle (RUL). Dans le but d'assurer des performances optimales de ces systèmes pendant les missions, il devient impératif d'effectuer des opérations de maintenance durant les pauses planifiées. En raison de ressources limitées telles que les coûts, le temps, les pièces de rechange ou le nombre de techniciens de maintenance, il n'est généralement pas possible d'effectuer toutes les actions de maintenance souhaitées. Le problème de la maintenance sélective (SMP) vise alors à répondre aux questions suivantes : quelles actions de maintenance doivent être réalisées et sur quels composants afin que (i) la fiabilité du système soit maximisée pour la mission à venir, ou encore (ii) le système atteigne un niveau de fiabilité spécifié pour la mission à venir au coût global le plus faible possible. Il existe de nombreuses applications industrielles et militaires où la MS peut être mise en œuvre. Par exemple, les systèmes informatiques sont couramment utilisés tout au long de la journée mais peuvent être entretenus en dehors des heures de travail. Les flottes de véhicules de livraison sont déployées durant la journée mais sont disponibles le soir pour la maintenance. Les systèmes de production sont censés fonctionner en continu pendant les jours ouvrables, mais subissent des arrêts de maintenance durant les fins de semaine. Lors des opérations militaires, la maintenance est effectuée entre les missions de combat.

Depuis son introduction en 1998 par Rice et al. (1998), le SMP a suscité un intérêt considérable tant chez les chercheurs universitaires que chez les industriels. Néanmoins, plusieurs situations réelles cruciales nécessitent encore une attention particulière afin d'améliorer substantiellement la compréhension de la modélisation et de l'optimisation de la MS. Le présent travail de recherche vise en effet à combler quatre lacunes majeures, telles que décrites ci-dessous :

1. Les modèles existants de SMP supposent que les décisions de maintenance sont prises et exécutées sur des systèmes situés au même endroit et disposant d'un nombre suffisant d'équipes de maintenance. Ce n'est pas toujours le cas, puisque les systèmes industriels peuvent être géographiquement dispersés

et leur maintenance peut nécessiter le déplacement des équipes de maintenance d'un endroit à un autre. Par conséquent, le premier thème de ces travaux de thèse vise à développer des modèles conjoints de MS et de routage des équipes de maintenance pouvant être appliqués à des flottes de systèmes ou à des composants d'infrastructures en réseau géographiquement dispersées.

2. La majorité des modèles de SMP proposés dans la littérature existante reposent uniquement sur des modèles statistiques pour le calcul de la fiabilité des composants. Cependant, l'émergence de nouvelles technologies telles que les capteurs sans fil et l'IIOT, ainsi que les avancées en IA, a popularisé les approches basées sur les données utilisant des algorithmes de Deep learning (DL) pour l'évaluation de la fiabilité et la prédiction de la durée de vie résiduelle (RUL). L'objectif du deuxième thème de recherche est précisément de développer un cadre de maintenace prédictive et sélective (PdSM) utilisant des approches de DL pour prédire la fiabilité des composants.

3. Peu de modèles de SMP dans la littérature considèrent des flottes de systèmes multicomposants ainsi que les contraintes inhérentes aux capacités de maintenance lors de la planification des tâches de maintenance. Bien que le problème conjoint de SMP et d'affectation des réparateurs ait reçu une attention particulière ces dernières années, les réparateurs ne représentent qu'une des nombreuses ressources de maintenance à considérer. D'autres capacités, telles que les outils spécialisés et les installations de réparation, doivent également être intégrées au problème de planification de la maintenance. De plus, les modèles de SMP pour les flottes de systèmes reposent souvent sur l'hypothèse restrictive selon laquelle les missions et les pauses programmées commencent simultanément pour tous les systèmes et que les missions à venir sont identiques, ne nécessitant qu'un seul système pour réussir. Le troisième thème de cette recherche vise alors à développer des modèles d'optimisation conjoints de MS et de planification des ressources pour des flottes de systèmes avec des durées de pauses programmées asynchrones.

4. Aucune étude existante sur la MS n'a exploré les systèmes mettant en œuvre une politique d'abandon de mission (MAP). Dans ces systèmes, une mission peut être interrompue et une procédure de sauvetage doit alors être initiée afin de protéger des vies humaines ou des équipements de valeur lorsque les risques associés à la poursuite de la mission atteignent un niveau critique. C'est dans ce contexte que le quatrième et dernier thème de ces travaux de recherche explore le SMP appliqué aux systèmes intégrant des politiques d'abandon de mission, où les compromis entre l'optimisation de deux mesures seront étudiés et analysés, à savoir la probabilité d'accomplir la mission (MSP) et la probabilité de survie du système (SSP).

En théorie de la fiabilité, un système est généralement représenté par la configuration logique de ses composants nécessaires à l'exécution d'une fonction donnée. Cette représentation du système sera adoptée tout au long de cette thèse. Il est supposé qu'un accès complet à toutes les données nécessaires à la modélisation et à l'évaluation est disponible. Bien que cela ne soit pas toujours le cas dans toutes les applications, les modèles développés visent à fournir un cadre général pouvant être adapté à divers contextes orientés mission.

Dans le présent manuscrit de thèse, quatre thèmes sont explorés, chacun abordant l'une des lacunes mentionnées précédemment. Chaque thème est développé dans un chapitre autonome comprenant sa propre revue de littérature, son introduction et ses résultats. Concernant le premier thème, un article introduisant le problème conjoint de MS et de routage (*orienteering*) des réparateurs a été publié dans *Computers & Operations Research*. Un deuxième article proposant un modèle conjoint de maintenance et de routage des équipes de réparation afin d'améliorer la résilience des réseaux après l'occurrence des événements perturbateurs a été publié dans *Reliability Engineering & System Safety*. Concernant le deuxième thème, un article proposant une nouvelle stratégie de maintenance prédictive et sélective (PdSM) a été présenté lors de la 10e conférence *Manufacturing Modeling, Management and Control* tenue à Nantes, en France. Un article étendant et améliorant significativement ce travail a été ensuite publié

dans l'*International Journal of Production Research*. Un article développant un nouveau modèle de MS pour une flotte (FSM) basé sur les données, optimisant conjointement l'affectation système-mission, la sélection des composants et des niveaux de maintenance, ainsi que l'affectation des tâches de réparation, a été soulis à la revue *Autonomous Intelligent Systems*. Pour le troisième thème, un article portant sur les pauses programmées de maintenance asynchrones et la planification de maintenance sous contraintes de ressources pour des flottes de systèmes multicomposants a été publié dans l'*International Journal of Production Research*. Enfin, concernant le quatrième thème, une version révisée d'un article introduisant un nouveau modèle optimisant conjointement la planification de la MS et les décisions d'abandon de mission pour des systèmes critiques a été soumise à la revue *Reliability Engineering & System Safety*. La section suivante résume les motivations et les objectifs de chacun des thèmes de recherche abordés dans nos travaux de thèse .

1 Thèmes de recherche

Dans cette section, les objectifs et les motivations de chacun des thèmes de recherche mentionnés précédemment seront présentés, ainsi que les principales questions de recherche auxquelles nous avons proposé des réponses.

1.1 Thème 1 : Maintenance sélective conjointe et routage des équipes de maintenance

Le premier thème étudié dans ces travaux de thèse porte sur l'optimisation conjointe des décisions de maintenance et de routage des équipes de maintenance. Le défi consiste à sélectionner conjointement les composants du système à maintenir, à déterminer le niveau de maintenance approprié à effectuer et à optimiser le routage de multiples équipes de maintenance. Traditionnellement, ces décisions étaient prises indépendamment les unes des autres, ce qui conduisait à des résultats sous-optimaux et à une utilisation inefficace des ressources. Les considérer simultanément peut potentiellement entraîner des améliorations significatives en matière d'efficacité opérationnelle, de réduction des coûts et de performance globale des systèmes.

L'objectif général de ce premier thème est de développer des modèles d'optimisation englobant l'optimisation conjointe des décisions de maintenance et de routage des équipes de maintenance. En s'appuyant sur des techniques de recherche opérationnelle, des algorithmes d'optimisation capables de résoudre efficacement les problèmes complexes et de grande taille rencontrés dans des contextes réels sont développés. Les contributions de ce thème vont au-delà des avancées théoriques. En optimisant conjointement les décisions de maintenance et le routage des équipes, nous cherchons à fournir des outils pratiques d'aide à la décision pouvant être appliqués dans divers secteurs tels que la fabrication, le transport, l'énergie et les télécommunications. Le développement d'approches d'optimisation efficaces permet aux organisations de prendre des décisions éclairées, menant à une meilleure allocation des ressources, à une amélioration des performances opérationnelles et, ultimement, à une compétitivité accrue.

Ce thème vise à répondre aux questions suivantes :

- Existe-t-il des avantages en termes de coûts opérationnels ou de performance des systèmes lorsqu'on optimise conjointement les décisions de maintenance et le routage des équipes de maintenance ?
- Comment les décisions de maintenance et de routage des équipes peuvent-elles être intégrées efficacement afin d'améliorer l'efficacité opérationnelle ?
- Quels sont les principaux facteurs et contraintes à considérer dans l'optimisation conjointe des décisions de maintenance et de routage des équipes de maintenance ?

- Quels algorithmes d'optimisation à grande échelle et quelles métaheuristiques permettraient de capturer la complexité du problème d'optimisation conjointe et de fournir des solutions efficaces ?
- Comment cette approche d'optimisation conjointe peut-elle être adaptée et appliquée à différents secteurs industriels et contextes opérationnels ?

1.2 Thème 2 : Maintenance sélective prédictive

L'Industrie 4.0 représente la tendance la plus récente en matière d'automatisation et utilise des capteurs, des CPS et l'IIOT afin de connecter les mondes physique et virtuel. Grâce à la mise en œuvre de ces technologies innovantes, la maintenance prédictive est de plus en plus reconnue comme une stratégie de maintenance attrayante dans de nombreux secteurs industriels. Un élément clé d'une telle stratégie de maintenance est le pronostic, qui concerne la prédiction des défaillances avant qu'elles ne surviennent (Jardine et al., 2006). En pronostic, il existe deux principaux types de prédiction. Le premier concerne la prédiction du RUL, tandis que le second consiste à prédire la probabilité qu'un composant fonctionne sans défaillance jusqu'à un instant futur donné. Récemment, les algorithmes de DL ont réalisé des avancées considérables en termes de performance et ont permis d'améliorer l'état de l'art dans les domaines du pronostic et de la maintenance prédictive (Namuduri et al., 2020). Des plans de maintenance prédictive efficaces permettent en effet de réduire les coûts de maintenance, de prolonger la durée de vie des équipements, de diminuer les temps d'arrêt des machines et d'améliorer la production. Les méthodologies de pronostic incluent des techniques traditionnelles de machine learning (ML), telles que la régression logistique, les machines à vecteurs de support (SVM) et les méthodes d'ensemble, ainsi que des techniques avancées de DL, comme les modèles perceptrons multicouches, les plCNN et les plRNN. Les algorithmes de DL nécessitent des dizaines, voire des centaines de milliers d'échantillons d'apprentissage pour atteindre un niveau de précision satisfaisant. Les capteurs IIOT peuvent produire des quantités massives de données, permettant ainsi l'application des algorithmes de DL. Les méthodes de ML et de DL permettent d'identifier des relations entre les valeurs des capteurs et des variables dépendantes telles que le temps avant défaillance, les futures valeurs des capteurs ou encore la détection d'anomalies.

Le deuxième thème de cette thèse porte sur le développement de stratégies avancées de maintenance prédictive et sélective (PdSM) basées sur les données. L'objectif est de tirer parti de la puissance des techniques de DL combinées à des approches d'optimisation afin d'identifier les décisions de maintenance optimales pour les systèmes orientés mission (MOS). Cette stratégie de maintenance s'inspire du domaine de la maintenance prédictive appliquée au pronostic des machines. Le concept de PdSM a suscité un intérêt important en raison de son potentiel à améliorer considérablement la fiabilité et les performances des MOS. En s'appuyant sur les travaux de Hesabi et al. (2022), qui ont introduit le premier cadre de PdSM, ce thème de recherche vise à faire progresser davantage cette approche afin de répondre aux limites et aux défis associés aux systèmes complexes. Ce thème considère également l'impact du niveau de dégradation des composants sur le temps requis pour effectuer les opérations de maintenance. Le modèle d'optimisation développé permet non seulement d'identifier l'ensemble optimal de composants à maintenir, mais également de déterminer conjointement l'affectation optimale des tâches de maintenance aux réparateurs. Une méthode de résolution efficace, basée sur une approche de décomposition, est développée. Enfin, plutôt que d'utiliser un classificateur probabiliste pour prédire la fiabilité des composants, un algorithme de DL combiné à la technique de simulation de Monté Carlo (MCD) est employé. Les avantages de cette approche sont mis en évidence et discutés en détail.

Ce thème de recherche vise à répondre aux questions de recherche suivantes :

- Quelles sont les limites associées aux cadres existants de PdSM et comment peuvent-elles être surmontées ?

- Comment les techniques de DL peuvent-elles être utilisées efficacement pour la prédiction de la fiabilité dans le contexte de la PdSM ?
- Comment l'incertitude liée à la prédiction du RUL et de la fiabilité peut-elle être prise en compte ?
- Comment l'impact des niveaux de dégradation des composants sur les besoins en temps de maintenance peut-il être quantifié et intégré au processus de planification de la maintenance ?

1.3 Thème 3 : Maintenance sélective de flotte sous contraintes de ressources

Le troisième thème de cette thèse porte sur l'extension du modèle classique du problème de la MS appliqué à une flotte de systèmes (FSMP) afin d'inclure une planification détaillée des ressources ainsi que des systèmes présentant des pauses de maintenance asynchrones. Tous les modèles existants de FSMP reposent sur l'hypothèse restrictive selon laquelle tous les systèmes de la flotte fonctionnent de manière synchrone, ce qui signifie que les tâches de maintenance ne peuvent être décidées et exécutées qu'une fois que tous les systèmes ont terminé la mission en cours et sont simultanément disponibles (Al-Jabouri et al., 2024). Cependant, dans les situations réelles, les flottes sont composées de systèmes multicomposants avec des calendriers de mission décalés, nécessitant une planification de maintenance asynchrone. De plus, la planification de la maintenance d'une flotte exige une planification détaillée et une allocation précise des capacités de maintenance, alors qu'aucun modèle existant de SMP ne prend explicitement en compte les principales ressources nécessaires à l'exécution des tâches de maintenance, telles que les outils, les installations et le personnel de réparation. Ce thème de recherche vise à remédier à ces limites en proposant un modèle généralisé de FSMP sous contraintes de ressources, introduisant une planification de maintenance asynchrone et une allocation explicite des ressources. Le modèle établit un lien théorique entre le SMP et le problème de gestion de projet à contraintes de ressources (RCPSP), offrant ainsi une approche plus flexible et réaliste de la planification de la maintenance des flottes.

Ce thème de recherche vise à répondre aux questions de recherche scientifique suivantes :

- Comment les pauses de maintenance asynchrones et les contraintes de ressources peuvent-elles être modélisées efficacement dans un cadre de la MS ?
- Quels sont les impacts des durées de pauses asynchrones sur les décisions de sélection de taches de maintenance et sur la fiabilité de la flotte ?
- Quels enseignements en management peuvent être tirés de ce modèle et comment peut-il soutenir la prise de décision des gestionnaires de maintenance ?

1.4 Thème 4 : Maintenance sélective conjointe et décisions d'abandon de mission

Dans le domaine des modèles de MS, l'accent a principalement été mis sur la maximisation de la probabilité de succès de mission (MSP), une mesure essentielle pour évaluer les performances des MOS. Cependant, certains MOS, tels que les aéronefs, les engins spatiaux et les véhicules aériens sans pilote (UAV), accordent davantage d'importance à la survie du système qu'au succès de la mission. La perte de ces systèmes peut entraîner des conséquences économiques importantes, des risques environnementaux et même des pertes humaines. Afin de réduire ces risques, des politiques d'abandon de mission (MAP) ont été introduites, permettant l'initiation de procédures de sauvetage pour prévenir les défaillances des systèmes. Dans ce contexte, ce thème de recherche de la thèse vise à aborder les défis et les opportunités présentés par les systèmes intégrant des MAP dans un cadre de MS.

Dans les systèmes fonctionnant sous des MAP, la décision d'abandonner une mission et d'initier une procédure de sauvetage est influencée par divers facteurs. Le nombre de composants défaillants (Levitin

et al., 2017) et le nombre de chocs externes (Levitin and Finkelstein, 2018) affectant le système peuvent tous servir de déclencheurs à l'interruption de la mission. Notre objectif de recherche vise à développer des modèles d'optimisation intégrant ces facteurs afin de fournir des indications sur l'allocation des ressources de maintenance permettant d'améliorer à la fois la probabilité de succès de mission (MSP) et la probabilité de survie du système (SSP). En tenant compte des compromis entre ces objectifs concurrentiels, il devient possible d'optimiser le processus décisionnel et d'assurer une allocation optimale des ressources de maintenance pour les systèmes critiques du point de vue profil de mission.

L'objectif principal de ce thème de recherche consiste à combler le manque de connaissances concernant la MS appliquée aux systèmes fonctionnant sous des MAP. Alors que la littérature existante s'est largement concentrée sur l'optimisation des décisions de maintenance pour maximiser la MSP, notre travail de recherche est le premier à considérer les systèmes dotés de capacités d'abandon de mission dans le cadre de la MS. En développant des modèles d'optimisation de MS adaptés à ces systèmes particuliers, il sera possible de réaliser des avancées significatives dans la préservation de la survie des systèmes et dans la réduction des risques. Les résultats de cette recherche auront des implications importantes pour les industries reposant sur des systèmes critiques, allant de l'aviation et de l'exploration spatiale aux opérations impliquant des UAV.

Ce thème de recherche vise à répondre aux questions suivantes :

- Comment définir les probabilités de succès de mission et de survie du système ?
- Comment la sélection et l'allocation des ressources de maintenance, notamment le temps, le budget, le personnel et les pièces de rechange, peuvent-elles être optimisées afin de maximiser à la fois les probabilités de succès de mission et de survie du système ?
- Quels sont les principaux paramètres et facteurs à considérer dans le développement de modèles de MS pour les MOS intégrant des politiques d'abandon de mission ?
- Quels sont les compromis potentiels entre les coûts de maintenance, la MSP et le SSP dans les MOS ?

2 Structure de la thèse

Comme mentionné précédemment, cette thèse réalisée par articles est composée de six manuscrits (quatre publiés, un resoumis après révisions et un soumis pour publication). Le chapitre 2 présente des modèles conjoints optimisant les décisions de MS et le routage des équipes de maintenance. Deux applications sont étudiées : (1) les fournisseurs de services de maintenance chargés de l'entretien d'une flotte de turbines offshore (OWT) à partir d'un port de soutien aux opérations et à la maintenance, et (2) les décisions conjointes de maintenance et de routage visant à améliorer la résilience des réseaux après une perturbation. Le chapitre 3 introduit un cadre de PdSM exploitant le DL. Un réseau de DN (DNN) combiné à la technique de simulation de Monté Carlo (MCD) est utilisé pour l'évaluation de la fiabilité, tandis qu'un algorithme de classification floue permet d'identifier les états de santé des composants. Les expériences numériques démontrent les avantages de cette approche de pronostic basée sur les données. Le chapitre 4 étend le modèle classique FSM en intégrant des pauses de maintenance asynchrones ainsi que des contraintes détaillées sur les ressources de maintenance. Le chapitre 5 développe un modèle conjoint de MS et de prise de décision d'abandon de mission applicable aux systèmes orientés mission et aux systèmes critiques. Des résultats issus d'expériences numériques démontrent que l'optimisation conjointe des décisions d'abandon de mission et de la planification de la maintenance améliore la "survivabilité" des systèmes comparativement aux cas sans politique d'abandon.

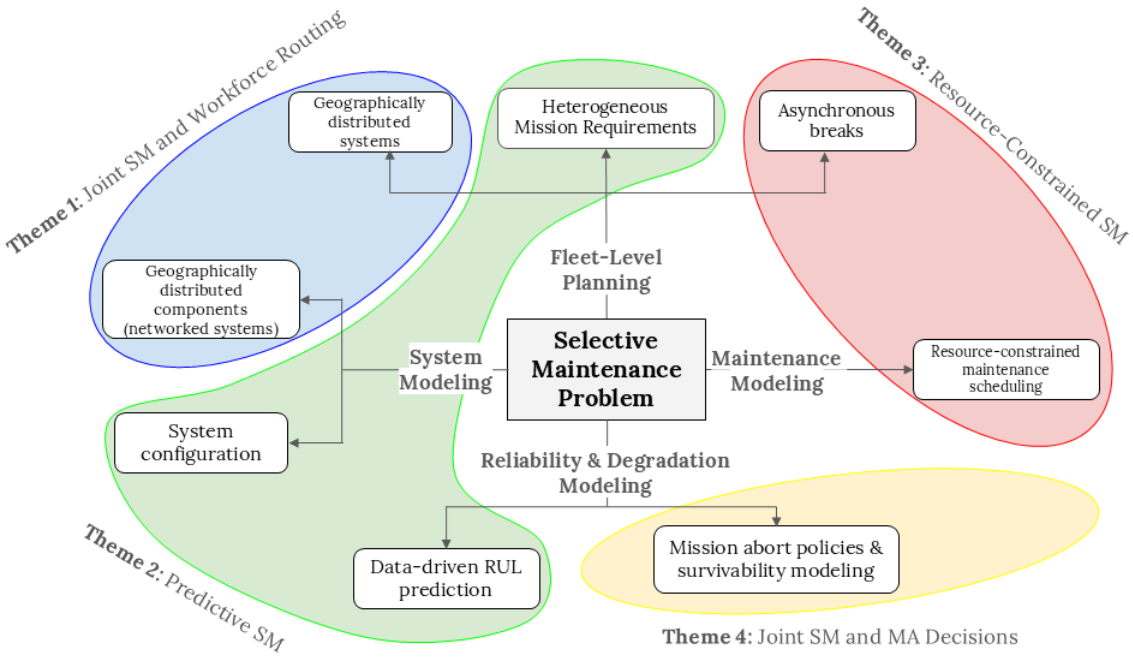


Figure 1: Contributions de la thèse.

D'un point de vue théorique, cette thèse introduit de nouveaux modèles mathématiques et de nouvelles méthodes de résolution. D'un point de vue pratique, elle élargit considérablement le champ d'application du problème de la maintenance sélective (SMP). Cette thèse est particulièrement pertinente puisqu'elle répond aux besoins de maintenance et de fiabilité des grands systèmes multicomposants, fait progresser l'état de l'art en matière de modélisation de tels systèmes et utilise des méthodes émergentes de résolution telles que l'analyse de données et l'intelligence artificielle. La figure 1 résume les contributions réalisées dans cette thèse en abordant plusieurs lacunes importantes des travaux de recherche existants dans quatre domaines principaux, notamment la modélisation des systèmes, la fiabilité, la dégradation et la maintenance, ainsi que la planification à l'échelle des flottes de systèmes multicomposants.

References

Al-Jabouri, H., Saif, A., Khatab, A., Diallo, C., and Venkatadri, U. (2024). A critical review of selective maintenance for mission-oriented systems: challenges and a roadmap for novel contributions. *International Journal of Production Research*, 62(13):4980–5015.

Hesabi, H., Nourelfath, M., and Hajji, A. (2022). A deep learning predictive model for selective maintenance optimization. *Reliability Engineering & System Safety*, 219:108191.

Ivanov, D., Dolgui, A., Sokolov, B., Werner, F., and Ivanova, M. (2016). A dynamic model and an algorithm for short-term supply chain scheduling in the smart factory industry 4.0. *International Journal of Production Research*, 54(2):386–402.

Jardine, A. K., Lin, D., and Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical systems and signal processing*, 20(7):1483–1510.

Kaynak, O. (2007). The exhilarating journey from industrial electronics to industrial informatics. In *2007 2nd IEEE Conference on Industrial Electronics and Applications*, pages xli–xlii. IEEE.

Kusiak, A. (2018). Smart manufacturing. *International Journal of Production Research*, 56(1-2):508–517.

Levitin, G. and Finkelstein, M. (2018). Optimal mission abort policy for systems operating in a random environment. *Risk Analysis*, 38(4):795–803.

Levitin, G., Xing, L., and Dai, Y. (2017). Mission abort policy in heterogeneous nonrepairable 1-out-of-n warm standby systems. *IEEE Transactions on Reliability*, 67(1):342–354.

Namuduri, S., Narayanan, B. N., Davuluru, V. S. P., Burton, L., and Bhansali, S. (2020). Deep learning methods for sensor based predictive maintenance and future perspectives for electrochemical sensors. *Journal of The Electrochemical Society*, 167(3):037552.

Rice, W., Cassady, C., and Nachlas, J. (1998). Optimal maintenance plans under limited maintenance time. In *Proceedings of the seventh industrial engineering research conference*, pages 1–3.

Xu, L. D., Xu, E. L., and Li, L. (2018). Industry 4.0: state of the art and future trends. *International Journal of Production Research*, 56(8):2941–2962.