# Short Duration Job Scheduling and Assignment using Staged Mixed Integer Programs

by

Christina Michaels

B.S., Chemical Engineering
Columbia University, 2017

Submitted to the MIT Sloan School of Management and
Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degrees of

Master of Business Administration

and

Master of Science in Electrical Engineering and Computer Science

in conjunction with the Leaders for Global Operations program
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
May 2022



Author . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
MIT Sloan School of Management and
Department of Electrical Engineering and Computer Science
May 6, 2022

Certified by . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Georgia Perakis
William F. Pounds Professor, MIT Sloan School of Management, Co-Director of Operations Research Center
Thesis Supervisor

Certified by . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Saurabh Amin
Robert N. Noyce Career Development Associate Professor of Civil and Environmental Engineering
Thesis Supervisor

Accepted by . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Maura Herson
Assitant Dean, MBA Program
MIT Sloan School of Management

Accepted by . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Leslie A. Kolodziejski
Professor of Electrical Engineering and Computer Science
Chair, Department Committee on Graduate Students

**Short Duration Job Scheduling and Assignment using Staged Mixed Integer Programs**

by

Christina Michaels

Submitted to the MIT Sloan School of Management and
Department of Electrical Engineering and Computer Science
on May 6, 2022, in partial fulfillment of the
requirements for the degrees of
Master of Business Administration
and
Master of Science in Electrical Engineering and Computer Science
in conjunction with the Leaders for Global Operations Program

## Abstract

As part of large-scale digital transformation efforts, Atlantic Utility's electric field force recently introduced a mobile work dispatch solution aimed at reducing inefficiencies associated with daily work. The application retired many of the manual, paper-based processes previously employed by field crews and supervisors to complete daily short-cycle (<6 hrs) jobs; it also introduced new capabilities that allow supervisors to review accumulated jobs in their operational region and strategize for their completion. Current operations find supervisors left with a long list of jobs to sift through when attempting to make daily work assignments and when scheduling work for one or more days in the future. Application users must manually identify jobs to schedule or assign while considering the distance to the job, required completion date, duration, and other factors. These factors contribute to the job priority level but without a simple way to aggregate these considerations into a clear set of prioritized jobs, supervisors are challenged to identify which work packets are highest priority and should be completed first. Daily scheduling and assignment is further complicated by the trade-off supervisors face when determining how to balance reduction of accumulated historical jobs with new jobs coming in at variable rates.

This thesis formulates two proof-of-concept mixed integer programs that perform staged scheduling and assignment of short duration jobs. The objective functions include use of a metric indicative of a job's total number of days past due or coming due. In this way, the formulations incorporate the real world trade-off supervisors face between historical and newly-created jobs subject to constraints on daily crew availability, increasing their utility as a future in-app aid for supervisors. Results of the scheduling stage over 2- to 6-day planning horizons indicate increased backlog reduction in comparison to naive or random strategies. Variation of user-defined inputs shows the scheduling formulation can be tuned to prioritize either jobs past due or those coming due in greater proportions, subject to the preferences of individual supervisors. When using both scheduling and assignment stages in sequence, results over 1- and 3-month simulated trials show consistently better performance in reducing job accumulation in comparison to historical records observed across operational barns of varying sizes.

These results provide justification for a full operational pilot and recommendations for how to deploy production-ready algorithms are included in this thesis. They also suggest that greater improvement in barn operations is possible without assumption of increased crew capacity. Use of the staged formulations in the mobile work dispatch solution could introduce greater uniformity in how short duration jobs across the Atlantic Utility network are prioritized and completed, and may lead to enhanced customer service. These improvements could be realized through incorporation of these formulations as an automatic in-app aid to supervisors and field crews. Further, application of the staged approach to workforce allocation can be considered in industries outside utilities including those that involve logistics and delivery operations.

Thesis Supervisor: Georgia Perakis
Title: William F. Pounds Professor, MIT Sloan School of Management, Co-Director of Operations Research Center

Thesis Supervisor: Saurabh Amin
Title: Robert N. Noyce Career Development Associate Professor of Civil and Environmental Engineering

# Acknowledgments

First, I'd like to thank the OnMyWay team for the opportunity to work on such a valuable and interesting project. Chris Kelly, Mey Ngu, Chris Annibale, Kathy Jackson, Eric Doyle, Erik Pola, Mark Willard, Aisling Jensen, and Ben Michaud, I sincerely appreciate your willingness to include me as part of your team and your support throughout each stage of my internship and thesis. I'd also like to thank Dawn Angell, Andrew Beliveau, Charles Kyle Bentley, Jean Dubois, Nicholas Maiello, and David Tse for your consistent guidance and encouragement as I worked to define my project, collect data, write code, and develop results. Thank you also to the all the field workers, supervisors, and resource coordinators that allowed me to safely perform interviews and collect critical data on their experiences in the field.

I'd also like to thank my MIT academic advisors Georgia Perakis and Saurabh Amin for their invaluable guidance, mentorship, and technical direction. Thank you for being so generous with your time and knowledge; your instruction was crucial to the development of this work.

Additionally, Monica Kabel and Heather Hausladen, thank you for mentoring me throughout the internship and for helping me navigate questions around my project scope. I greatly appreciate your guidance and advice.

I'd also like to thank my fellow LGO classmates for your friendship, technical advice, and personal understanding. I feel so lucky to be a part of the LGO program and community.

To my family – I wouldn't be here today without your support and encouragement. Thank you for sticking with me every step of the way. I am eternally grateful for your love.

THIS PAGE INTENTIONALLY LEFT BLANK

# Contents


**1 Introduction** **13**

1.1 Overview of Atlantic Utility . . . 13

1.1.1 EBU Field Operations . . . 15

1.1.2 Overview of OMW Application . . . 17

1.2 Problem Statement . . . 19

1.2.1 Organizational Context . . . 19

1.2.2 Generalized Problem Statement . . . 20

1.3 Project Approach . . . 20

1.4 Contributions . . . 22

1.5 Scope and Limitations . . . 23

1.6 Thesis Outline . . . 24

**2 Literature Review** **25**

2.1 Optimal Workforce Allocation . . . 25

2.2 Objectives . . . 28

**3 Data** **31**

3.1 Supervisor and Field Crew Primary Interviews . . . 32

3.1.1 Job Scheduling Criteria . . . 32

3.1.2 Job Assignment Process and Prioritization . . . 34

3.1.3 Challenges with Manual Scheduling and Assignment . . . 35

3.2 Dataset Characteristics . . . 36

3.2.1 Job Volume Varies by Region . . . 36

3.2.2 Job Accumulation Over Time . . . . . . . . . . . . . . . . . . 37
3.3 Data Availability . . . . . . . . . . . . . . . . . . . . . . . . . . 39

**4 Model Formulation 41**
4.1 Rationale for Staged Optimizations . . . . . . . . . . . . . . . . . 41
4.2 Optimization Objective . . . . . . . . . . . . . . . . . . . . . . . 42
4.3 Schedule Model Formulation . . . . . . . . . . . . . . . . . . . . 43
4.3.1 Assumptions . . . . . . . . . . . . . . . . . . . . . . . . 43
4.3.2 Formulation . . . . . . . . . . . . . . . . . . . . . . . . 43
4.4 Assignment Model Formulation . . . . . . . . . . . . . . . . . . . 47
4.4.1 Assumptions . . . . . . . . . . . . . . . . . . . . . . . . 47
4.4.2 Formulation . . . . . . . . . . . . . . . . . . . . . . . . 48

**5 Results 51**
5.1 Structure . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 51
5.2 Use of Open-Source Software . . . . . . . . . . . . . . . . . . . . 52
5.3 Scheduling Optimization Results . . . . . . . . . . . . . . . . . . 52
5.3.1 Scheduling Strategies . . . . . . . . . . . . . . . . . . . . 53
5.3.2 Strategy Comparison . . . . . . . . . . . . . . . . . . . . 53
5.3.3 Impact of Coming Due Parameter . . . . . . . . . . . . . . 57
5.4 Staged Optimization Results: Scheduling and Assignment . . . . . . . 59
5.5 Results Summary . . . . . . . . . . . . . . . . . . . . . . . . . . 63

**6 Recommendations for Operationalization 65**
6.1 Continuous Integration and Continuous Delivery . . . . . . . . . . . . 66
6.2 Trade-off between In-House and External solutions . . . . . . . . . . 68

**7 Conclusions and Future Work 71**
7.1 Future Work . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 72

**A Tables 75**

**B Figures 79**

# List of Figures


1-1 EBU Fieldwork Types . . . . . . . . . . . . . . . . . . . . . . . . . . . 15
1-2 Short- and long-cycle job assignment prior to introduction of OMW . 17
1-3 Digital capabilities with OMW for short-cycle service work . . . . . . 19
1-4 Staged Optimization Approach: Scheduling and Assignment . . . . . 21

3-1 Short-Cycle Job Filtering (Variable Length Planning Horizon) . . . . 33
3-2 Regional Job Accumulations . . . . . . . . . . . . . . . . . . . . . . . 37
3-3 Correlation Between Open Work Packets and Average Overdue [1] . . . 38
3-4 Correlation Between OMW Users and Open Work Packets . . . . . . 38
3-5 Job Inflow and Outflow Per Barn Throughout OMW Adoption . . . . 40

4-1 Staged Optimization Flow . . . . . . . . . . . . . . . . . . . . . . . . 42

5-1 Scheduling Strategy Comparison Over 2- to 6-Day Planning Horizon . 55
5-2 Impact of Coming Due Multiplier - Barn A (MA Region 1) . . . . . . 57
5-3 Impact of Coming Due Multiplier - Barn D (NY Region 4) . . . . . . 58
5-4 Impact of Coming Due Multiplier - Barn E (NY Region 5) . . . . . . 58
5-5 Backlog Comparison Over 1-Month Historical Simulation . . . . . . 60
5-6 Open Job Comparison Over 3-Month Historical Simulation . . . . . 62

B-1 Representative Short-Cycle Job Assignment Process . . . . . . . . . . 80
B-2 Backlog Comparison Over 3-Month Historical Simulation . . . . . . 81

THIS PAGE INTENTIONALLY LEFT BLANK

# List of Tables

3.1 Representative barns used for backlog and job accumulation analysis.[2] 39

4.1 Notation used in scheduling optimization. . . . . . . . . . . . . . . . . 44
4.2 Notation used in assignment optimization. . . . . . . . . . . . . . . . . 48

5.1 Jobs Completed Over a Single Historical Month . . . . . . . . . . . . . 61

A.1 No. of Customer Appointments Scheduled (A) . . . . . . . . . . . . . 75
A.2 No. of Jobs Coming Due Scheduled (A) . . . . . . . . . . . . . . . . . 75
A.3 No. of Customer Appointments Scheduled (B) . . . . . . . . . . . . . 75
A.4 No. of Jobs Coming Due Scheduled (B) . . . . . . . . . . . . . . . . . 76
A.5 No. of Customer Appointments Scheduled (C) . . . . . . . . . . . . . 76
A.6 No. of Jobs Coming Due Scheduled (C) . . . . . . . . . . . . . . . . . 76
A.7 No. of Customer Appointments Scheduled (D) . . . . . . . . . . . . . 76
A.8 No. of Jobs Coming Due Scheduled (D) . . . . . . . . . . . . . . . . . 76
A.9 No. of Customer Appointments Scheduled (E) . . . . . . . . . . . . . 77
A.10 No. of Jobs Coming Due Scheduled (E) . . . . . . . . . . . . . . . . . 77
A.11 No. of Customer Appointments Scheduled (F) . . . . . . . . . . . . . 77
A.12 No. of Jobs Coming Due Scheduled (F) . . . . . . . . . . . . . . . . . 77

THIS PAGE INTENTIONALLY LEFT BLANK

# Chapter 1

# Introduction

Chapter 1 provides an overview of Atlantic Utility operations, defines the problem statement that is addressed, and details the research approach. The chapter concludes with an outline of the thesis chapters and a discussion of research contributions.

## 1.1 Overview of Atlantic Utility

Atlantic Utility is a large investor-owned energy company in the northeast United States with operations in Massachusetts and New York. With direct business-to-consumer networks, the company provides generation, transmission, and distribution to millions of electric and natural gas customers; it also maintains several hundred transmission and distribution substations and well as thousands of circuit miles of overhead and underground electrical lines.

Atlantic Utility's vision is to be at the heart of a clean and affordable energy future; a vision it achieves in part through the strategic introduction of new digital technologies that improve internal operations and enhance its end users' experience with company products and services. The utility is investing in external solutions that provide customers enhanced access to digital information that introduces more digital ways of engaging as well as internal tools to enhance the productivity and engagement of employees. This increasing investment in new technology and software has been observed across the North American utilities industry as part of the larger "smart grid" transition in which new applications, sensors, meters, and data management technologies are used to reduce operational costs and improve customer experience [6].

The Electric Business Unit (EBU) within Atlantic Utility is responsible for the operation of transmission facilities and distribution networks across Massachusetts and New York jurisdictions. As part of these responsibilities, the EBU maintains control centers across both regions, supports emergency planning, and manages electric assets like poles, wires, and feeder lines.

Recent digital initiatives in the EBU are in line with the focus of the larger company and center on the introduction of innovative technologies and agile development methods to revamp field force operations and improve operational efficiency. Organizational objectives for these initiatives have three main areas of focus:

- **Replace existing processes** - The EBU is analyzing existing organizational processes to identify opportunities to improve field force and business unit efficiency; this extends to the use of agile as opposed to waterfall methods of development.

- **Introduce new technology** - The EBU has introduced new devices such as handheld tablets throughout its operations and is adopting advancements in AI, machine learning, and digital workflow automation.

- **Promote a cultural shift** - The EBU, and Atlantic Utility generally, intends to drive widespread cultural change across the organization by modifying the mindsets and behaviors of the field force to encourage swift adoption of new processes and technology.

The subject of this work addresses all three areas of focus in a concentrated setting. *Specifically, we focus on an initiative that replaces the existing process of service work distribution, introduces new technology in the form of a custom software application, and promotes a cultural shift to ensure widespread adoption.*

### 1.1.1 EBU Field Operations

The EBU covers New England and New York territories within Atlantic Utility and orchestrates fieldwork operations through over 65 service barns, each covering variably-sized territories and ranging from urban to rural settings. Barns are locally managed by resource coordinators and supervisors, who work to schedule incoming and existing work as well as assign available field crews to complete work on a daily, weekly, and monthly basis. Barns employ between 1 and 6 supervisors, depending on 1) their territory size, 2) the number of field force workers assigned to that territory, and 3) the amount of electric field work in the area; on average, each supervisor manages about 15 line workers amounting to about 1,200 line workers in total across all regions.

Electric field work can be of several types as shown in Figure 1-1. Examples of short cycle work include the installation of a new service connections, the removal of existing service connections, or the installation of new residential meters. Long cycle work is typically associated with larger construction or maintenance projects and can require several days to multiple weeks worth of time for assigned crews.

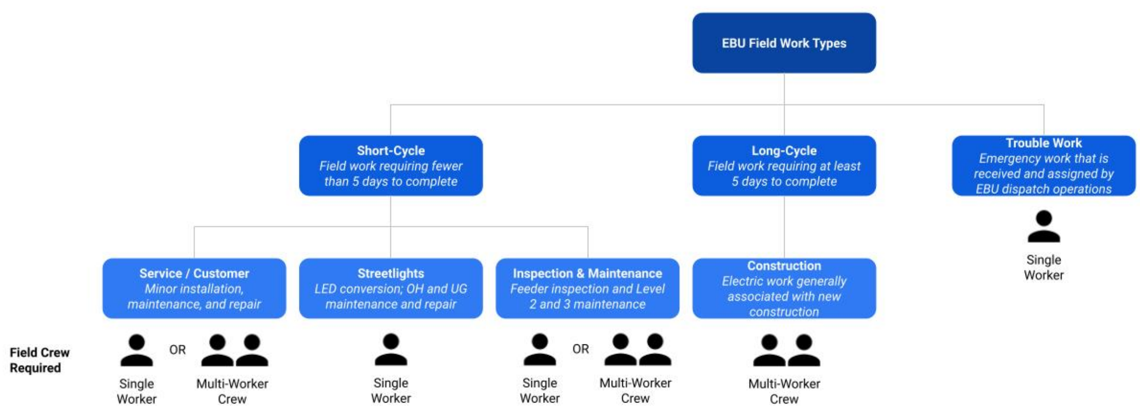


Figure 1-1: EBU Fieldwork Types

Completion of short- and long-cycle work can also be disrupted by the occurrence of trouble work. Trouble (or emergency) work can occur at any time for a variety of reasons including downed power lines, vehicle accidents involving power lines, or unexpected equipment failure. When it occurs, trouble work is of immediate priority and requires at least the attention of a single trouble worker but can require multiple crews depending on its severity.

### Crew Types

EBU field crews are of variable membership and size. The number of field force workers assigned to a particular crew depends on the work the crew will perform as well as the tenure of the field force worker. Multi-person crews always include at least one crew lead at a senior first class level. Crews' responsibilities include receiving job assignments from a supervisor, obtaining warehouse or stockroom supplies, driving to job sites, interacting with service customers, completing jobs, and picking up or dropping off work trucks.

The EBU operates with several different field crew types when performing short-cycle work:

- **Overhead (OH) Line Crew** - Typically a 2-person crew but may expand in size to include 3 - 5 field force workers; works most frequently on complex services and construction jobs but may also perform simple service or streetlight work.

- **Underground (UG) Line Crew** - A crew ranging in size from 3 - 5 field force workers that exclusively performs underground streetlight jobs; despite the focus on only a single job type, UG line crews can complete thousands of jobs per year.

- **Loop / Service Crew** - Typically a 2-person crew but may work on jobs requiring only a single person; works most frequently on simple and complex service jobs.

- **Trouble Worker / One-Person Crew (OPC)** - An exclusively 1-person crew that performs trouble work as it is occurs and may perform simple services or streetlight jobs in between trouble calls.

One of the EBU's development initiatives specifically focuses on work typically performed by loop / service and trouble worker / OPC crews. The OnMyWay application simplifies the manual, labor-intensive job assignment process performed by supervisors into a single, easily-traceable digital application that can be used by field crews as well. The EBU began development of OMW in January 2020.

### 1.1.2 Overview of OMW Application

OMW is a custom application that began development in January 2020 and is the focus of this thesis. *The application targets the scheduling, assignment, and completion of field work and digitizes a portion or all of these processes, which were previously entirely manual.*

#### History and Motivation

Short- and long-cycle job assignment throughout EBU service barns was previously a tedious and entirely manual process. Figure 1-2 maps the original procedure by which crews would be assigned daily and ongoing work by supervisors.

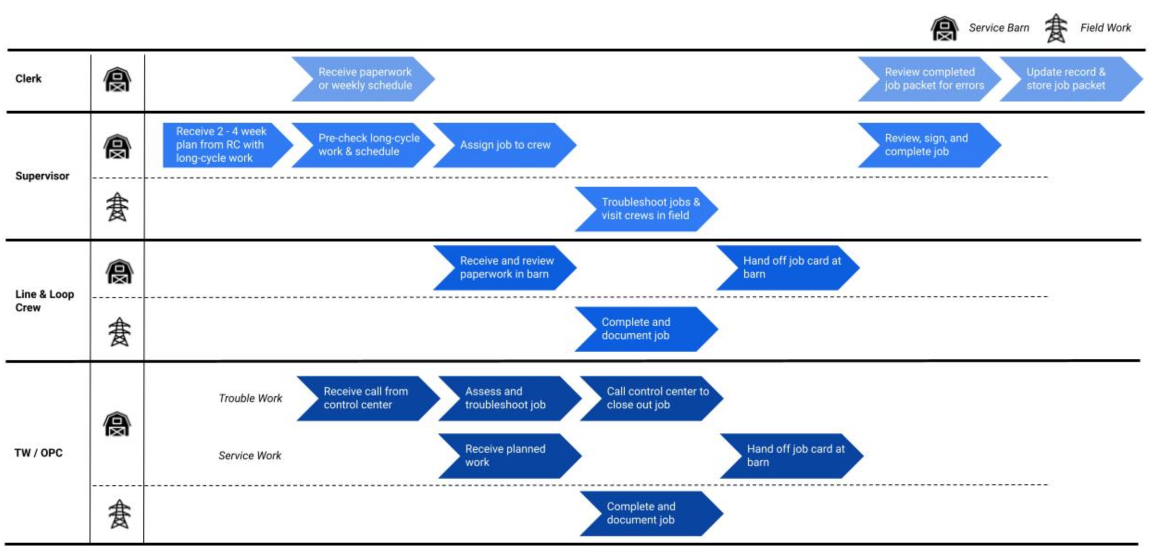


Figure 1-2: Short- and long-cycle job assignment prior to introduction of OMW

As shown, supervisors distributed paper work packets in single units or small sets to crews performing daily work. This was especially tedious for loop crews performing short-cycle work for which the completion of multiple jobs required multiple trips between the service barn and job site in a single day. The same inconvenience was true for trouble workers and one-person crews performing service work. *Additionally, with the distribution and collection of physical work packets, supervisors within barns were faced with daily challenges to ascertain which jobs were completed, which were outstanding, and which were highest priority of the set.*

In line with business objectives referenced in section 1.1, the EBU invested in the development of OMW which addresses two major frictions observed in EBU operations:

- Job scheduling and assignment was highly manual leading to:
  - *Time wasted receiving and prioritizing jobs and,*
  - *Limited visibility into overall job backlog and availability.*
- Paper was used to capture job assignment, close-out, and metering data leading to:
  - Inefficient work assignment via physical paper packets,
  - Redundant travel between job sites and barns,
  - Significant time needed to file and scan physical work packets, and
  - Exposure to error through manual processes.

### OMW Features

The OMW application addresses these frictions by providing a centralized portal for supervisors and field force workers to schedule, assign, and close out work. *Its initial focus is on short-cycle service jobs each requiring fewer than 6 hours of estimated duration*; the application does not include long-cycle or emergent trouble work. With this focus, the primary users are loop / service crews and trouble workers / one-person crews, who each perform simple service work during any given week.

Critically, the application contains details of all service jobs requiring fewer than 6 hours to complete for each barn. It also allows supervisors to assign jobs digitally and for field crews to claim jobs independently. Figure 1-3 details which actions can be performed digitally through the application.

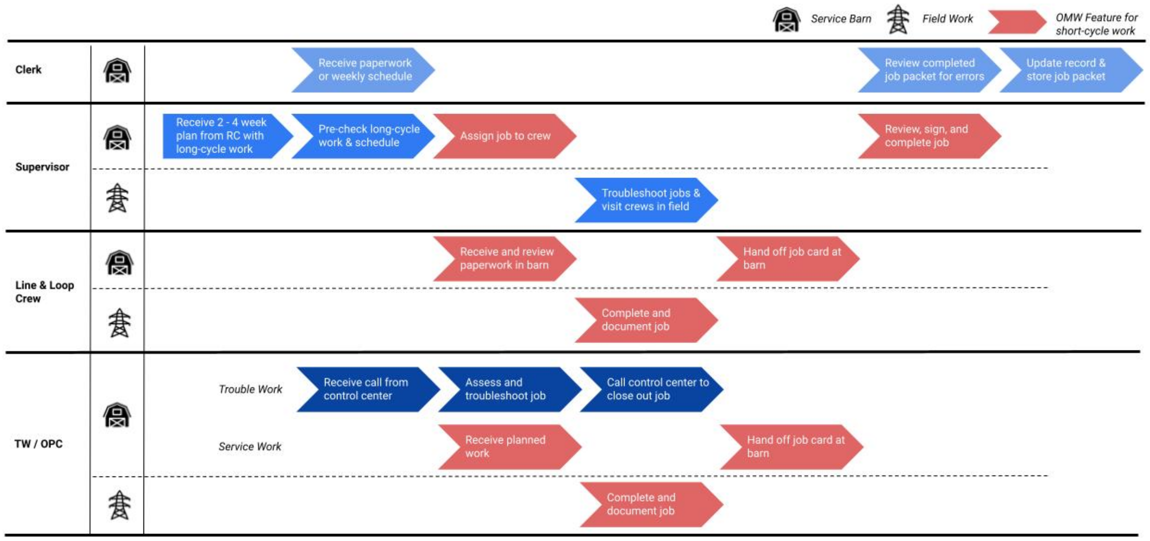


Figure 1-3: Digital capabilities with OMW for short-cycle service work

## 1.2 Problem Statement

### 1.2.1 Organizational Context

The transition to OMW addressed inefficiencies associated with the distribution of physical work packets – allowing supervisors to assign jobs to crews in the field without needing to make additional trips to the barn – and allowed for greater visibility into the overall job backlog per barn. With these process changes and with increased data visibility, it is possible to further alleviate the challenge faced by supervisors in the timely management of the set of service jobs such that there is reduced risk of jobs becoming overdue and increased confidence that those of highest priority are completed with preference.

Recognizing jobs have differing due dates, durations, and priorities that must be considered in tandem with constraints in the number and availability of field crews, we seek to address the operational challenge of managing a variably-sized backlog at a barn level and to introduce uniformity in how jobs are prioritized across the barn network. Example criteria that must be considered when scheduling and assigning service jobs include:

- Overdue jobs versus jobs coming due

- Date-sensitive customer appointments
- Daily crew availability (number of crews available)
- Crew capacity

### 1.2.2 Generalized Problem Statement

We generalize this context within and outside the utilities industry to relate to operational processes that manage short-cycle service work that is created at a variable rate and accumulates over time, and which is subject to capacity constraints. We seek to investigate whether:

1. Staged optimization formulations for job scheduling and assignment can reduce a service set of variable length across operational units with differing capacity and, whether
2. These formulations can be introduced with input parameters such that there is user-directed influence over which service jobs are scheduled and assigned.

To address these questions, we hypothesize that identifying criteria to determine job priority and generating two stages of optimization formulations can provide a predictable strategy to reduce a set of open service jobs.

## 1.3 Project Approach

We approach the thesis from the perspective of an individual operational unit, or barn, that manages a set of accumulated service jobs that are either already overdue or coming due. We approach the problem in this way because 1) scheduling and assignment of jobs is performed on a per barn basis and 2) there is limited resource sharing between barns, apart from trouble workers and one-person crews, meaning that each is subject to largely localized crew constraints.

Each barn is assumed to have two related objectives that can be completed in stages: job scheduling and crew assignment. This approach is guided by prior research within the utilities industry, as will be discussed in Chapter 2. Staging scheduling and assignment operations allows for flexibility in how the optimization formulations are used – for example, if supervisors in a barn use only scheduling suggestions while continuing to assign jobs manually based on additional constraints – and avoids the assumed intractability of combined optimizations. Figure 1-4 provides a high-level overview of how the staged optimizations can be used independently or in sequence.

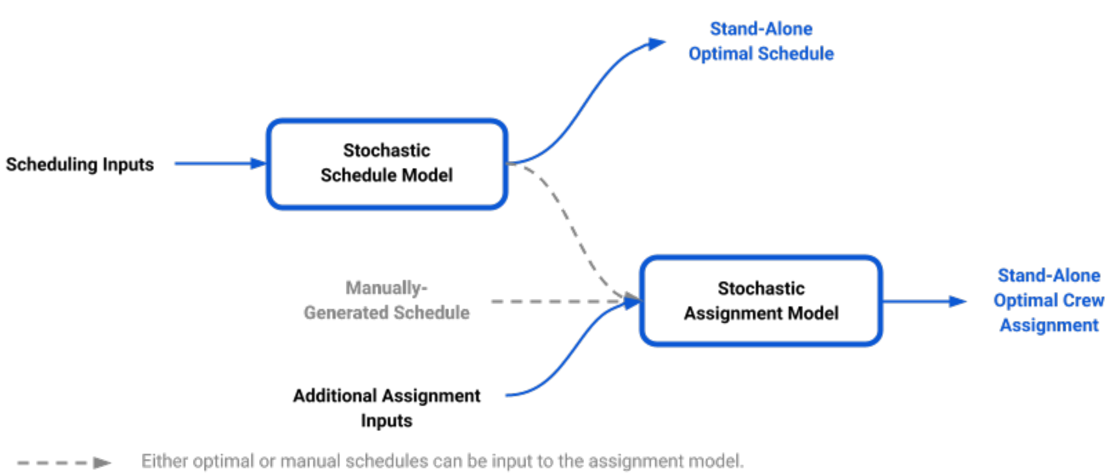


Figure 1-4: Staged Optimization Approach: Scheduling and Assignment

The two objectives are described in stages below:

- **Scheduling** - Scheduling service jobs to days within a variable-length planning horizon and,

- **Assignment** - Assigning crews to jobs scheduled on a given day within the planning horizon.

A third objective related to crew self-assignment of jobs based on their real-time location in the field was investigated as part of this work but is not discussed in this thesis. As will be discussed in section 1.5, optimization of crew routes between jobs was out of scope for this project and was not incorporated in the approach.

## 1.4 Contributions

This thesis presents an application of proof-of-concept mixed integer optimization algorithms to the scheduling and assignment processes of daily electric utility service work, though use cases are generalizable outside the utilities industry. This technique has previously been applied to the gas distribution side of the Atlantic Utility business, which will be discussed in Chapter 2. This thesis helps to further progress efforts to optimize workforce allocation by:

1. Formulating a standardized metric – number of days overdue – with which to assess the set of service jobs belonging to a localized service barn, or any regional service area, and applying it to the workforce allocation problem (scheduling and assignment);

2. Applying mixed integer optimization formulations for scheduling and assignment in stages to show that rapid, cumulative backlog reduction is possible; and,

3. Presenting analysis that reflects general improvement in backlog management with consideration of relevant trade-offs and without the assumption of increased daily crew capacity.

Application of mixed integer optimization formulations to variable-length field work performed on a daily basis can produce improvements in key workforce efficiency and productivity metrics such as those that measure the size of the backlog and the utilization of field crews. These algorithms can be considered industry agnostic and may be applicable to areas where a field force performs frequent, daily work of variable length that has associated priority levels (e.g. product delivery industries, home maintenance evaluations and services, real estate viewings and assessments).

## 1.5 Scope and Limitations

The project scope focuses on the application of mixed integer programs to singular service barn backlogs. Network-level job assignment in which multiple supervisors are assumed to assess multiple jobs across multiple barns is out-of-scope for this thesis. This network-level orchestration was also not observed in the field and therefore the per barn focus is assumed to be a valid scope.

Additionally, long-cycle and trouble jobs – those which are not included in the OMW application – are not considered in this work. Scheduling and assignment for these jobs require extended resource planning for the former and prediction of trouble work for the latter. However, these job types are strong candidates to extend this project.

Primary interviews with field force workers and supervisors indicated that route guidance between jobs was not a desired feature in the OMW application; for this reason, this thesis does not investigate individual routing optimization between service jobs.

Several limitations are also relevant to note for this project. First, limited tracking of job completion rates or of historical barn backlogs prevented the comparison to a standardized productivity baseline for the algorithms presented here. As a result, findings are presented with respect to relative reduction in current barn backlogs or in historical barn performance since the adoption of OMW.

Second, as availability of some relevant data is limited, several simplifying assumptions are made that allow for estimation of crew availability and job completion dates. These assumptions and the use of realistic, mock data are described in section 3.3.

Third, as the introduction of the OMW application is fairly recent, adoption across the EBU network has varied significantly. However it is assumed in this thesis that supervisors and field crews use OMW as intended, with use of the crew assignment features in the application, as opposed to manually performing those functions.

## 1.6 Thesis Outline

Chapter 2 provides an overview of relevant background information for this thesis. The background focuses on prior studies performed at Atlantic Utility relating to workforce allocation within the natural gas distribution unit of the company, which is subject to similar daily work requirements.

Chapter 3 presents the analysis of newly available service barn data as a result of the recent OMW introduction. It includes a discussion of findings from primary field interviews conducted as part of this thesis.

Chapter 4 focuses on the formulation of the proof-of-concept mixed integer optimization algorithms developed to schedule and assign service jobs. It also discusses assumptions made as a result of data availability limitations.

Chapter 5 shares results generated through application of the algorithm results to realistic scenarios including a comparison of optimal results to historical barn performance. It also describes how the optimization models can be used in staged phases.

Chapter 6 provides recommendations on how the optimization algorithms could be utilized in real-time through the OMW application. It also discusses the potential trade-off between use of an in-house solution versus external application to allow real-time use of the algorithms in the field.

Chapter 7 summarizes the key conclusions of this work and suggests recommendations for future extensions of the algorithms presented here.

# Chapter 2

# Literature Review

This chapter aims to contextualize this research in relation to two primary areas of focus: common approaches to achieving optimal workforce allocation and objective metrics used to evaluate it. Prior studies both within the utilities industry, including at Atlantic Utility itself, as well as an example outside the utilities context are discussed.

## 2.1 Optimal Workforce Allocation

Multiple past studies relating to optimal workforce allocation in the utilities industry have taken a staged approach to daily and weekly planning in which scheduling work within a fixed or variable time horizon is accomplished separately from assignment of resources and field crews. Prior research at Atlantic Utility performed by graduate researcher Siddharth Balwani in 2012 focuses on objectives similar to that of this thesis and that relate to resource planning optimization within the gas operations business [2]. Balwani investigates the short-term, daily planning process to allocate crews to pipeline construction and maintenance jobs; a process that is non-standardized and can vary significantly between operational yards. Operations are also subject to emergent, unplanned work in the form of gas leaks of varying levels of severity. Originally, the thesis develops a combined stochastic optimization model which takes as input all jobs and their respective deadlines as well as crews and their respective daily or weekly availability; the result is expected to be a full schedule and assignment of jobs and crews. However, the model is demonstrated to be impractical as a result of the "large number of integer decision variables" and the potential for model results to be sub-optimal based on sensitivity to variable inputs [2]. As a result of the intractability of a combined scheduling and assignment optimization formulation, which can cause undesirable delays in producing results, Balwani introduces the aggregate approach in which scheduling and assignment steps are staged in sequential stochastic optimization algorithms.

This thesis is framed in a similar context to that of Balwani's, including the need to plan for multiple operational units that manage long- and short-cycle work and which vary in size and demand based on their region. The problem addressed in thesis also reflects the lack of standardization between operational units that was also observed by Balwani; this highlights a shared motivation for both works being the introduction of more uniform job scheduling and assignment processes. However constraints around gas leakages that require immediate attention impact Balwani's formulation more significantly as there is an assumption that any crew could be called to address a gas leak. In our case, dedicated trouble workers are assumed to have intentionally reduced capacity for short-cycle work and a separate workforce (loop crews) present a more reliable capacity within any given planning horizon.

A continuation of Balwani's research was pursued in a 2014 Management Science publication by Angalakudati et al [1]. It further explores the decomposition of the operational resource allocation challenge through formalizing two phases of the problem: the job scheduling phase and the crew assignment phase. The former is based on a heuristic derived from a linear programming relaxation formulation which assumes a deterministic number of gas leak emergencies each day; in this case, the average number of emergencies is assumed. This allows for the optimization to be solved as a mixed integer program with significantly fewer decision variables and constraints and reduces the critical operational time needed for algorithm results [1]. The crew assignment phase is constrained to a single day time horizon and is solved as a stochastic mixed integer program with the allowance for overtime.

Angalakudati's work provides additional basis for the staged optimization approach adopted in this thesis. Similarly, we assume a deterministic availability of trouble workers which is also based on the average gas leak assumption from Angalakudati et al.

Similar approaches are observed in the utilities industry outside Atlantic Utility and the US. Research conducted by Ferreira et al. at the Energisa group in Brazil considers the optimal allocation of operational depots, the tactical sizing of regional crews of electricians, and the optimal work schedule of these crews [5]. While the study focuses specifically on the generation of spatio-temporal demand forecasts that serve as input to each of these modules, two topics of note are particularly relevant to this work.

First, Ferreira uses historical and current service order demand to guide forecasting under variable length planning horizons. Service orders are grouped into three categories: emergency, scheduled regulated, and scheduled non-regulated [5]. Emergency orders "must be handled as soon as possible" while regulated orders are allowed several days to complete and unregulated orders incur no penalties to the utility; all orders provide basic information about the job required including a description of the job, execution time, and what type of service is required [5]. These are similar to the trouble work, short- and long-cycle work categories discussed in section 1.1.1 and which provide similar input to the algorithms developed in this thesis. It is important to note that Ferreira uses big data algorithms applied to service order data containing location and technical work descriptions. This data is available from the eleven facilities that were studied and therefore Ferreira's work requires significantly more data as input than is available in this thesis.

This leads to the second relevant similarity in Ferreira's work relating to the use of separate algorithms to tackle different operational objectives including the identification of depot locations, sizing of regional crews, and scheduling of crew work. This is similar to Balwani et al.'s approach as well and suggests a precedent within the utilities industry for approaching workforce allocation in distinct stages.

On the basis of this prior work, it is not novel to apply a staged decomposition of scheduling and assignment algorithms to the challenge of optimal workforce allocation in industry. This thesis applies this staged approach with the intention to use algorithms in sequence to produce a comprehensive, suggested work plan for supervisors. Additionally, we focus on daily or near-term planning horizons without the allowance for overtime, presenting more constrained scenarios than in prior literature.

## 2.2 Objectives

Prior research also suggests feasible objective metrics by which to achieve optimal allocation. We will first discuss several metrics observed in prior literature and then comment on their relation to this thesis.

Angalakudati et al. further Balwani's original metric in which the reduction of overtime hours worked by maintenance crews is minimized. This meets a critical objective for Atlantic Utility to reduce overtime hours that present a "significant opportunity for achieving lower costs and better deadline compliance [1]." Importantly, overtime is known to be the result of controllable factors such as scheduling processes and uncontrollable factors like uncertainty associated with trouble work. The controllable factors present an opportunity to reduce overtime hours by introducing job scheduling and planning, where previously no standardized procedure existed.

Related work performed at Atlantic Electric by graduate researcher Sean Whipple in 2014 presents a different objective metric. Whipple focuses on the allocation of electric repair crews during emergency response situations and in contrast to Balwani and Angalakudati, investigates major storm events that are infrequent; these events were of interest as they are difficult to predict but have significant monetary and service implications [14]. The objective metric used in Whipple's thesis is the minimization of worst case repair time needed during emergency outages. Utility regulators impose costs associated with downtime and with the time needed to return to service; minimizing this metric ensures that repair plans meet regulator requirements.

Outside the utilities industry, optimal workforce allocation is also achieved through more direct measures of cost. Schaefer et al. investigates the scheduling of airline crews with consideration of unexpected disruptions to simulate more realistic operational scenarios. The paper investigates what makes a "good" schedule as it recognizes that airlines' traditional use of planned cost, assuming no disruptions, is rarely an accurate metric [12]. Schaefer instead proposes the use of an operational cost metric in which the total cost of a schedule is the summation of its constituent pairing, or crew trips, that compose it. The thesis reports that the operational cost metric used to evaluate crew schedules performs better when compared against real operations that experience disruptions and is more realistic against other relevant metrics such as on-time performance.

We can see a variety of metrics used to assess optimal workforce allocation and scheduling ranging from a reduction in overtime hours, restoration time, and operational costs. These metrics can have a direct or indirect impact on bottom line cost to the business. While Balwani and Angalakudati focus on the minimization of overtime hours allowing for a direct measurement of cost savings, we investigate a metric with an indirect impact on EBU operations by focusing on the reduction of the cumulative backlog of any given operational barn. Similar to Whipple's work, this metric has the potential to reduce regulatory constraints and ensure customer service expectations are met.

THIS PAGE INTENTIONALLY LEFT BLANK

# Chapter 3

# Data

This thesis is intended to aid supervisors in the scheduling and assignment of short-cycle jobs. To do this, it is critical to understand what criteria are used to determine job priority and what constraints are typically imposed on supervisors when performing these functions. Two primary sources of data were used in this thesis to understand these criteria and constraints: 1) firsthand accounts from field crews and supervisors regarding job scheduling and assignment processes and 2) historical barn data acquired through the OMW application database. The former provides insight on the criteria used to prioritize jobs for scheduling and assignment as well as the challenges associated with executing these actions manually. The latter provides metrics that can be used to assess the accumulation of jobs in a given barn. Data from the OMW application database provides a historical view of job completion rates since the application's adoption in early 2020. As the application is in its early stages, some data relating to specific field crew assignment and completion of jobs is not available. However, barn-level data provides insight into the variability of job volume across regions and how jobs accumulate over time.

Data associated with short-cycle jobs have several key attributes which have been identified as contributing to a job's priority. These include the job's required date, customer appointment status, and estimated duration. Definitions for these terms are as follows:

- **Required Date**: The date by which a service job should be completed, typically five days upon its import into the OMW application.
- **Customer Appointment Status**: The boolean distinction whether a job requires a specific appointment with a customer that has an associated date and time for completion; customer appointments can only be completed within the specified window.
- **Estimated Duration**: The estimated man hours required to complete the service job.

## 3.1 Supervisor and Field Crew Primary Interviews

A qualitative data gathering effort was conducted as part of this thesis which included over 30 interviews with EBU supervisors, field workers, and resource coordinators across 8 service barns in New England and New York. The objective was to understand how short-cycle jobs are prioritized, what challenges exist with manual job scheduling and assignment, and which metrics are regularly reviewed to evaluate barn performance.

### 3.1.1 Job Scheduling Criteria

We generated Figure 1-3 after performing the primary interviews with field crews, resource coordinators, and supervisors. It describes several filtering steps used determine how short-cycle jobs are scheduled. Job prioritization is usually at the supervisor's discretion – some may prioritize those jobs that are past due more than those that are coming due, or vice versa. As a result, there is no uniform criteria with which short-cycle jobs are prioritized in a daily or weekly schedule; however, data collected from interviews allowed for the creation of a general decision-making flow for scheduling jobs as shown in Figure 3-1.

From this we can see several criteria which impact the priority for each short-cycle job including two critical elements: the customer appointment status and whether

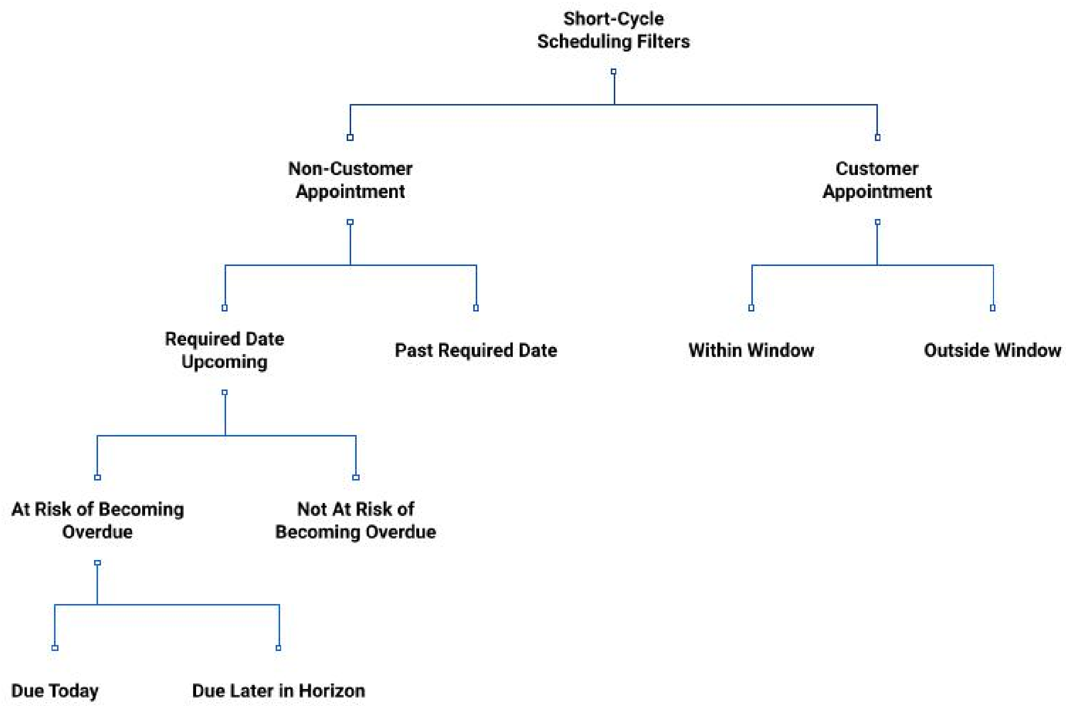


Figure 3-1: Short-Cycle Job Filtering (Variable Length Planning Horizon)

the job is past its required date. Assuming a variable length horizon, we propose a general ascending order of importance by job type:

1. Customer Appointments Within Window
2. Non-Customer Appointments Due Today (Start of Planning Horizon)
3. Non-Customer Appointments that are Past the Required Date
4. Non-Customer Appointments with Upcoming Required Dates at Risk of Becoming Overdue
5. Non-Customer Appointments with Upcoming Required Dates Not at Risk of Becoming Overdue

Short-cycle customer appointments have a specified date and time at which they must be completed. As a result, they are of highest priority to satisfy in any given scheduling horizon. Those that are outside the window are not prioritized as they should be completed only on their scheduled date.

Non-customer appointments can range in type. Assuming the start date of the planning horizon is eligible for jobs to be scheduled, those which are due on the starting day should be prioritized over those that coming due later in the horizon. Similarly, jobs that are already past their required date should be completed before those due later in the window.

This filtering allows some generalizations to be drawn about how jobs should be scheduled. First, customer appointments are of highest priority to be scheduled on their assigned date. Second, a trade-off exists between those jobs that are past their required date and those that are coming due. Supervisors must act under resource and time constraints to make this daily trade-off. This filtering and these generalizations about job priority are a primary reference for the following Chapter 4 that discusses scheduling optimization model formulation.

### 3.1.2 Job Assignment Process and Prioritization

Once a job has been scheduled for a particular day, the job should be assigned to a crew for completion. Note that each barn has a variable number of crews that perform short-cycle work and emergent or long-cycle work may require the attention of multiple crews unexpected. As a result, crew availability can constrain the number of jobs that can be completed in one day, even if the job has already been scheduled.

The assignment of crews to jobs is subject to most of the same criteria that is used in the scheduling process with the addition of the consideration of daily crew availability. The assignment process is performed in the following high-level steps:

1. **Review of job and crew data:** Supervisors review available data at the start of the assignment process. They can see details of the jobs themselves as well as which crews are available for the day.

2. **Prioritize jobs for assignment:** Supervisors consider several attributes of the job to determine priority including the customer appointment status, whether the job is overdue and by how much, whether the job is coming due, and whether the crews available have enough time to complete the job; this prioritization can

vary depending on the supervisor and barn but is generally expected to follow the criteria shown in Figure 3-1. This is assumed to be the most time-intensive step in the job assignment process and is the focus of this thesis.

3. **Assign crews to scheduled jobs:** Supervisors assign scheduled jobs to crews based on the crew availability in the OMW application.

4. **Crews manage assigned jobs:** Crews can complete jobs to which they are assigned or mark them as obstructed, otherwise.

This process is shown in Appendix B, Figure B-1. These steps are also the basis for the optimization formulations presented in Chapter 4.

### 3.1.3 Challenges with Manual Scheduling and Assignment

As shown, scheduling and assignment processes are closely linked and require consideration of multiple attributes of a single job. In addition to long-cycle and trouble work planning, supervisors must manage a growing number of short-cycle jobs with varying requirements and priorities. As opposed to long-cycle work which can be planned in advance, short-cycle service work is quickly cycled as regulatory requirements impose a 5-day window in which to complete them; this window frequently corresponds to the required date attribute of the job. As previously discussed, there is also the need for customer appointments to be completed on a particular day and time. In addition, daily crew availability is variable meaning that supervisors must perform a daily re-assessment of jobs to be scheduled and completed. This may lead to significant time spent on these processes which could otherwise be devoted to alternative supervisor activities like performing pre-checks on long-cycle work.

These challenges can also lead to an accumulation of jobs within a given barn. Sections 3.2.1 and 3.2.2 discuss this accumulation which can further complicate the supervisor's ability to perform optimal job scheduling and assignment manually.

In total, this thesis seeks to address these challenges at a barn level. Specifically, we address the accumulation of jobs in a given barn that must be completed in tandem with new short-cycle work coming in. Outside of this focus, there are also potential network-level benefits with the introduction of algorithms used across barns that could help with uniformity and predictability of overall short-cycle job completion.

## 3.2 Dataset Characteristics

OMW made data relating to barn backlogs and service job requirements more accessible. 20 service barns were chosen for this analysis as they represent a variety of barn types and sizes within the Atlantic Utility network; this set is more than 25% of the total set of barns in the EBU. The analysis in this section confirms that while job accumulation varies in magnitude across regions, it is prevalent across the EBU network, and the lack of uniformity in job inflow and outflow makes it challenging for individual supervisors to consistently manage their backlog. These barns are also the focus of the results presented in Chapter 5.

### 3.2.1 Job Volume Varies by Region

Figure 3-2 shows the magnitude of job accumulation in the 20 barns of interest across New England and New York regions. Specific data values are proprietary to Atlantic Utility and are redacted from the analysis. However, we observe a clear spread both within and across regions in terms of the number of jobs individual barns currently have in their queues and in general we expect barns within both regions to have some accumulation of jobs.

In addition to reviewing the number of open jobs in queue, it is also informative to observe barns from a perspective of the average number of overdue work packets in queue. As discussed in the previous section, overdue jobs introduce a trade-off for supervisors that have to balance completion of existing jobs with those that are newly created. Figure 3-3 presents this view in which there is a positively correlated relationship between barns with more open work packets that also have a larger number of work packets that are past their required date.

[1]Excluding barn outlier with close to 400 open work packets.

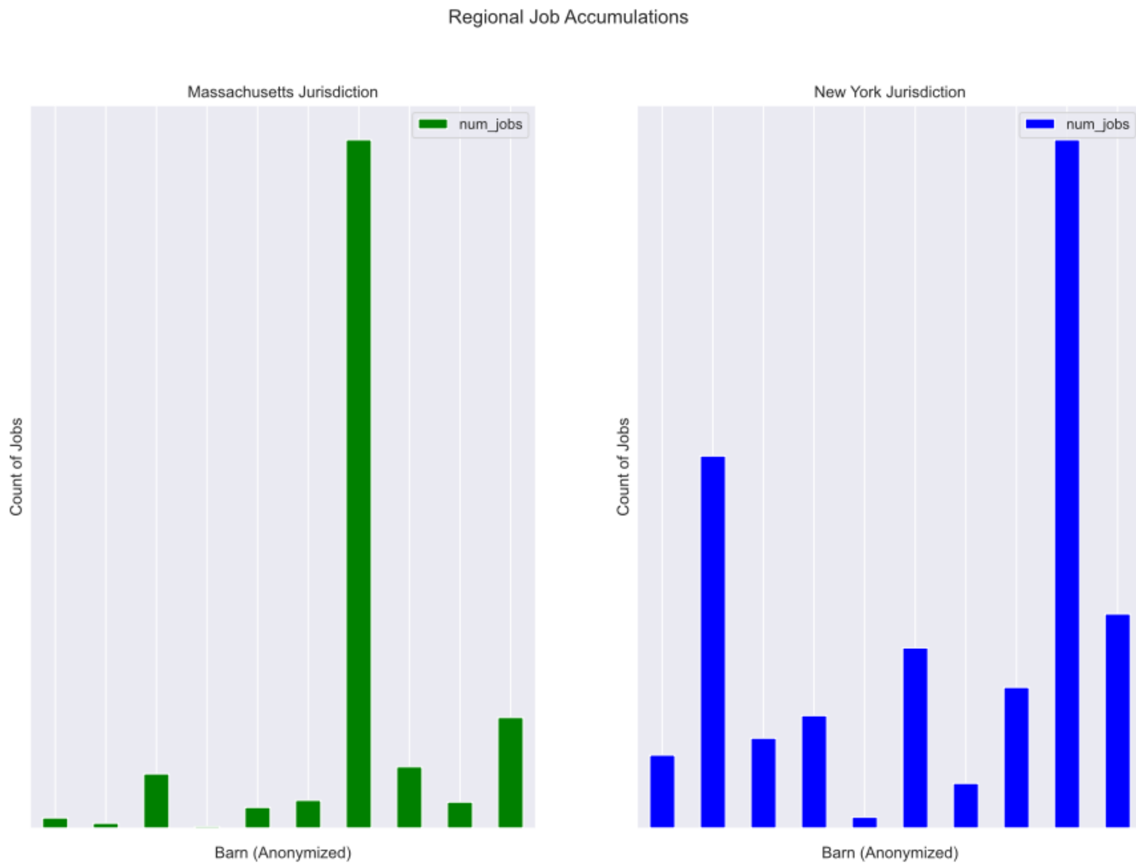


Figure 3-2: Regional Job Accumulations

One potential explanation is that barns with more work packets experience resource limitation. While data relating to specific daily crew availability is not available in OMW, Figure 3-4 presents a proxy for barn resource size by looking at the number of OMW users assigned to the barn. The figure shows that number of users in a barn does not appear to affect the magnitude of overdueness of its backlog as barns with the full range of users have similar numbers of jobs that are past due. This suggests that regardless of the number of users in OMW, barns of varying sizes can struggle with job accumulation.

### 3.2.2 Job Accumulation Over Time

Figure 3-5 shows aggregate job accumulation over time. Of the 20 service barns analyzed in this thesis, collectively representing about more than 25% of the total

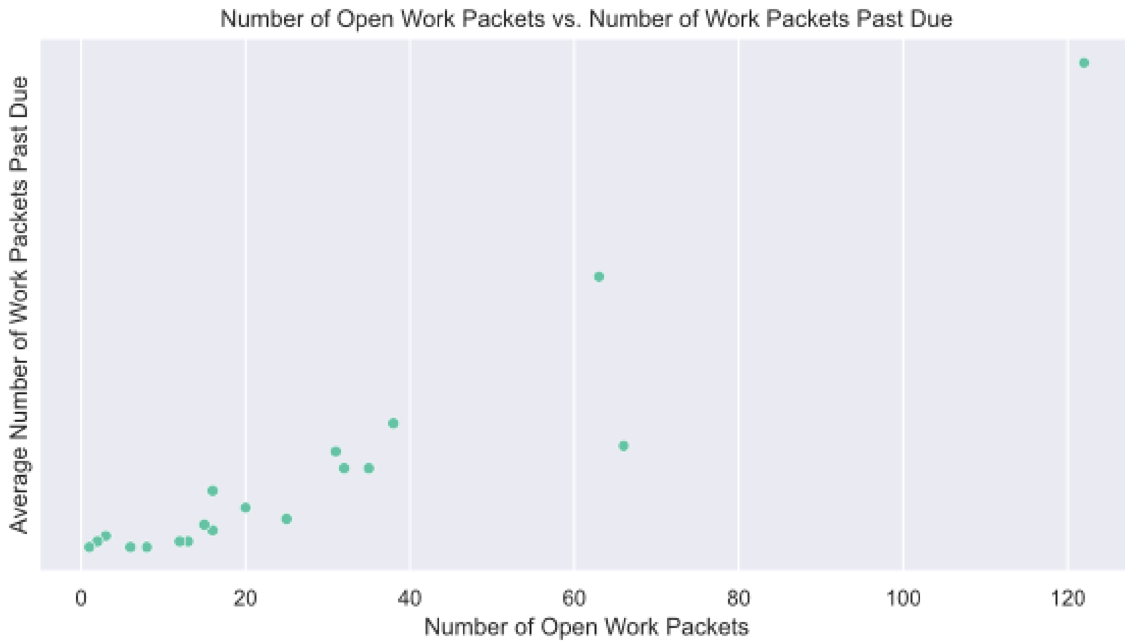


Figure 3-3: Correlation Between Open Work Packets and Average Overdue [1]

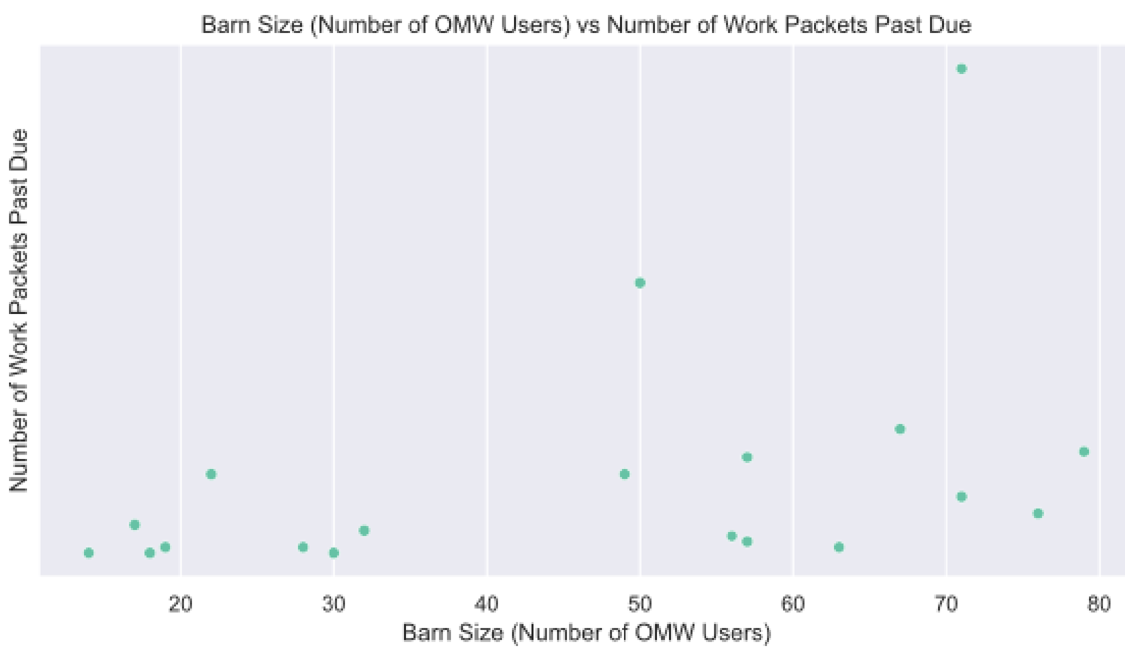


Figure 3-4: Correlation Between OMW Users and Open Work Packets

set of barns, 6 are presented here for comparison as they provide a perspective on each of the barn regions within the EBU. These barns are among the largest in their respective regions and varying job volumes.

| Barn Identifier | State | Region | No. Service Jobs | No. App Users |
|---|---|---|---|---|
| A | MA | 1 | 391 | 71 |
| B | MA | 1 | 63 | 50 |
| C | MA | 2 | 35 | 49 |
| D | NY | 3 | 122 | 71 |
| E | NY | 4 | 66 | 79 |
| F | NY | 5 | 20 | 76 |

Table 3.1: Representative barns used for backlog and job accumulation analysis.[2]

Figure 3-5 shows job inflow and outflow since the adoption of OMW across barns. Discrepancies between inflow and outflow indicate areas during which jobs accumulate. While the green shaded regions show areas where the accumulation is assumed to be reduced, the pattern of oscillating accumulation followed by reduction is consistent across barns. This lack of uniformity in job inflow and completion makes high-priority job requirements difficult for supervisors to predict on any given day.

These graphs also only show the inflow and outflow of jobs, they do not indicate the cumulative number of days overdue of the entire backlog. It is possible that even during periods of job reduction, those jobs which are highest priority – meaning they are already overdue or imminently coming due – may not be completed.

In total, data from the OMW application database reaffirms assumptions around the difficulty of manually scheduling and assigning jobs. It also suggests a lack of predictability within barns and across the EBU network which can be addressed with the introduction of standardized scheduling and assignment processes.

## 3.3 Data Availability

The OMW application is a recent introduction to EBU operations and as a result, some data is not yet available for analysis. Specifically, an estimation of how much time supervisors or crews spend on job scheduling and assignment would allow for an estimation of labor cost associated with these actions. As this data is not currently available, alternative estimations of job completion are used in this thesis and will be

[2]Data obtained between Dec. 2021 - Jan. 2022.


Figure 3-5: Job Inflow and Outflow Per Barn Throughout OMW Adoption

discussed in the following chapter.

Additionally, requirements relating to job materials, minimum crew size, crew complements, or daily crew availability are monitored separately outside the OMW application and were not available for analysis in this thesis.

# Chapter 4

# Model Formulation

This chapter describes the set of mixed integer optimization algorithms that apply the prioritization criteria described in Chapter 3 to allow for optimal scheduling of service jobs to days within a variable length planning horizon and assignment of the service jobs to available crews within a given barn.

Model formulation also builds upon past work by allowing for flexibility in the planning horizon for job scheduling – the user can request jobs to be scheduled for 1 day or any variable number of days – as well as introduction of a trade-off in prioritizing jobs that are overdue versus those that are coming due [14].

## 4.1 Rationale for Staged Optimizations

As discussed in Chapter 2, prior research both at Atlantic Utility and within the utilities industry have taken a staged approach to the scheduling of service work and assignment of crew resources. A similar approach is applied in this work in which algorithms are intended to be used sequentially. First, the scheduling optimization model will identify to which days a given service job should be scheduled taking into account its required date and estimated duration. Second, the assignment optimization model will take as input the optimal schedule as well as data on daily crew availability to produce a result that assigns crews to complete service jobs. Figure 4-1 describes the staged approach to model formulation.

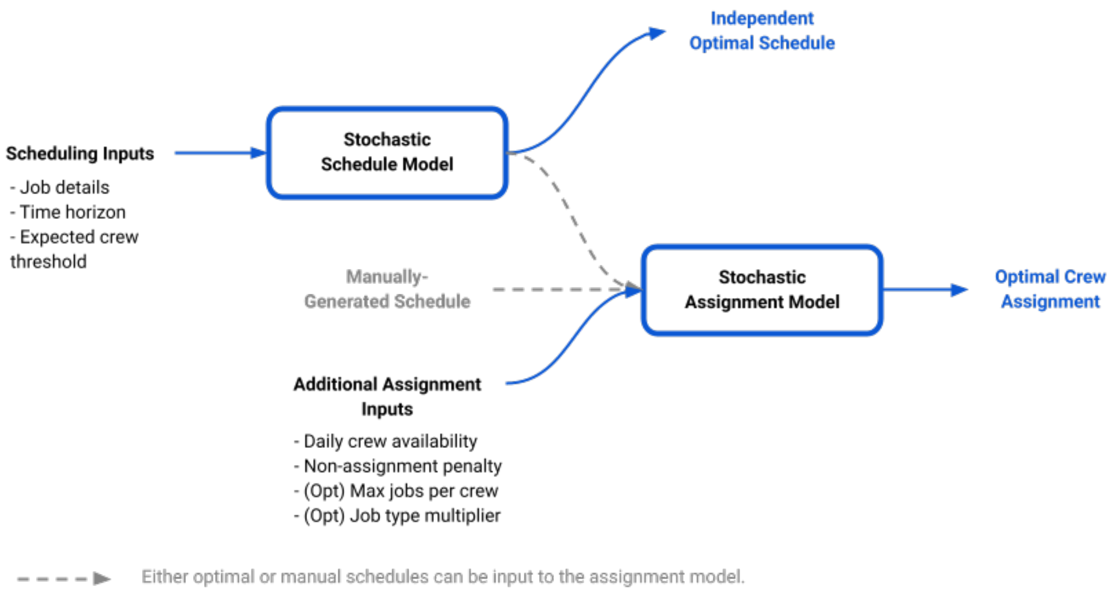


Figure 4-1: Staged Optimization Flow

The staged approach also has the added benefit of allowing supervisors to discretely produce the results of scheduling and assignment separately. This can be helpful if a manual modification needs to be made to the schedule prior to its input to the assignment model, which may occur on days in which extreme weather events occur or emergent trouble work requires multiple resources.

## 4.2 Optimization Objective

The optimization objectives applied in this thesis measure, in part, the cumulative number of days overdue of all service jobs in a given barn. Internally monitored metrics and primary interviews with field force workers indicated that service jobs which are most overdue are some of the highest priority jobs to complete. Overdue jobs present a risk of missing service level commitments or providing less efficient customer service. In addition, the trade-off between overdue jobs and those that are coming due is introduced through the use of penalties as described in the next section.

Future extensions of this work might investigate alternative objective metrics such as those more directly related to labor costs through the reduction in time spent on scheduling and assignment tasks.

## 4.3 Schedule Model Formulation

The objective of the schedule optimization formulation is to maximize the scheduling of jobs that are overdue or that would become overdue within the planning horizon; these jobs will be scheduled in preference over those that are not yet overdue or those that are due outside the planning horizon.

Similarly to Angalakudati's assumption regarding average daily emergencies, we assume a deterministic trouble work availability that can be explicitly entered by the user at run time [1]. This availability is otherwise determined to be the general average time that a trouble worker / one-person crew would devote to short-cycle service work on a daily basis.

### 4.3.1 Assumptions

The following assumptions are made to model short-cycle service job scheduling:

1. Short-cycle service jobs with less than 6 hours of estimated duration are used as input.

2. Daily job schedules are capped by an estimated, variable threshold. The threshold default is an estimation of how many crews are available on any given day multiplied by the number of on-duty shift hours per crew; this estimate can be modified by manual data input.

3. Trouble work is assumed to be deterministic as it is entered at run time by the user.

### 4.3.2 Formulation

Notation for the schedule optimization is shown in Table 4.1. We denote the decision variable as $X_{tj}$ which is a binary variable that takes a value of 1 if job $j$ was scheduled on day $t$ and a value of 0 otherwise.

The full model is shown mathematically as:

| Notation | Brief Definition |
|---|---|
| $t$ | Day index |
| $j$ | Job index |
| $A_j$ | Customer appointment status of job $j$ |
| $c_j$ | Days until job $j$ is due (e.g. $c_a = 5$ implies job $a$ is due in 5 days) |
| $D_j$ | Days until customer appointment; 0 if not |
| $e_j$ | Days overdue for job $j$ at the start of horizon |
| $H$ | Daily threshold; cumulative daily expected work hours |
| $l_j$ | Time taken to complete job $j$ (in hours) |
| $T$ | Planning horizon |
| $U$ | Multiplier for scheduling jobs that are coming due |
| $p$ | Percentage penalty for non-scheduling |
| $\mu_e$ | Average number of days past due of the set of accumulated jobs |

Table 4.1: Notation used in scheduling optimization.

**Objective:** Maximize scheduling of jobs that are overdue or would become overdue in the planning horizon.

$$\text{maximize } \sum_t^T \sum_j \begin{cases} (e_j + t) * X_{tj} - p * (1 - X_{tj}) * (e_j + T) & e_j > 0 \\ U * \mu_e * (c_j + 1) * X_{tj} & e_j = 0 \end{cases}$$

subject to the constraints:

1. Each job $j$ is scheduled on at most 1 day:

$$\sum_j \sum_t^T X_{tj}$$

2. Total number of jobs scheduled cannot exceed $H$ hours per day:

$$\sum_t^T \sum_j X_{tj} * L_j \leq H$$

3. A customer appointment must be assigned on the day of the appointment:

$$\sum_j \sum_t^T X_{tj} \geq 1 \quad \text{such that } D_j = t$$

4. Binary constraints for job assignments:

$$X_{tj} \in 0, 1 \quad \forall t, j$$

Upon starting the scheduling horizon, jobs can be 1) already past their required date 2) coming due within the planning horizon or 3) coming due outside the planning horizon. The three terms of the objective account for all possible states of a service job.

Objective terms that apply when $e_j > 0$ are those that apply to overdue jobs. Scheduling a job within the horizon indicates that $X_{tj} = 1$ and the objective term $(e_j + t) * X_{tj}$ applies. This term is equivalent to the current number of days overdue of the job added to the number of days into the horizon in which it is assigned. If an overdue job is not scheduled within the horizon, the maximization objective is deducted through the $(1 - X_{tj}) * (e_j + T)$ term. An additional, optional penalty multiplier can be applied to this term if a supervisor or barn wants to reduce the number of jobs that are not scheduled. In this case, the objective is most penalized when jobs that are significantly overdue are not scheduled.

The objective term that applies when $e_j = 0$ is relevant for jobs that are coming due. The term acts so that jobs which are coming due can also positively contribute to the objective. The $U * \mu_e$ term acts as a multiplier for jobs which are coming due that is based on the average days overdue of the existing backlog. The reason for this multiplier is that jobs which are coming due would have $e_j = 0$ so they would not contribute to the objective in any other term. Additionally, the magnitude of $c_j$ is expected to typically range from 1 - 5 days, unless the job is a customer appointment, and therefore the individual magnitude of overdue jobs, $e_j$, would likely outweigh that. The multiplier allows coming due jobs and those that are past due to be comparably weighted in the objective. It is important to note that this term can be modified to use any value for $\mu_e$ including the maximum or minimum number of days overdue; this flexibility is investigated in the next chapter. This term is also subject to supervisor discretion and allows for the trade-off discussed in the next section, 4.3.2.

The inverse case in which $X_{tj} = 0$ does not have to be accounted for as constraint 4 ensures that all jobs coming due within the planning horizon are assigned.

### Model Sensitivity to Supervisor Trade-offs

As previously discussed, supervisors face a trade-off between scheduling jobs which are coming due and those which are already overdue. The optimization model presented here incorporates a variable input multiplier, represented as $U$ where $U \geq 0$, that allows supervisors to increase the model's sensitivity to either of those two job categories. As the percentage increases, the weight of jobs which are coming due increases in the objective and incentivizes greater scheduling. The expected behavior of the multiplier is as follows:

- $\mathbf{U = 0}$: Jobs which are coming due are not included in the objective. They are only scheduled if there is available capacity outside of that used for overdue jobs.
- $\mathbf{U \leq 1}$: Jobs which are coming due are not as prioritized as those which are overdue but are given a weight that is a percentage of the average number of days overdue of accumulated jobs.
- $\mathbf{U > 1}$: Jobs which are coming due are prioritized over those which are overdue. They are given a weight that is some percentage greater than the average number of days overdue of accumulated jobs.

This is an important component of the model as it allows for a more realistic representation of job scheduling conditions and provides supervisors a lever with which to influence model results depending on the specific conditions within their barn.

### Model Feasibility

During initial development, the model presented infeasible solutions as a result of customer appointments that had due dates prior to the current date. These jobs caused optimization results to be infeasible due to constraint 3 which requires customer appointments be completed on the day the date of customer request. To avoid this infeasibility, a pre-filtering step is performed such that only customer appointments which have not yet come due are subject to constraint 3. Those which are already past their due date are subject to the same constraints as other overdue jobs.

Cases in which the model becomes infeasible include those in which the assumed crew capacity for the scheduling window is insufficient to complete the entire set of jobs. This capacity is further constrained by the assumption of no overtime, as discussed in prior sections. In the case of infeasibility, the model will allow for no jobs to be scheduled and in a productionalized system, it will indicate to the supervisor that there is not enough capacity to meet job demand.

**Model Output**

After the first optimization phase scheduling has been completed, the model returns a list of jobs scheduled to be completed for every day in the planning horizon.

## 4.4 Assignment Model Formulation

The objective of the assignment optimization formulation is to maximize assignment of jobs that are most overdue across the variable length planning horizon subject to a penalty for every job that is not assigned to be completed.

### 4.4.1 Assumptions

The following assumptions are made to model short-cycle service job assignment:

1. Short-cycle service jobs with less than 6 hours of estimated duration are used as input.

2. A schedule, generated via the schedule optimizer or manually, is pre-defined with short-cycle jobs assigned to days within the planning horizon.

3. Daily crew availability is deterministic and known throughout the planning horizon.

4. Estimated man hours and other job data is assumed to be accurate.

5. The daily uncertainty of trouble work is reflected in the reduced availability of crews, which is input by the user at run-time.

For the purposes of this thesis, it is also assumed that no overtime is allowed as it is not typically used for short-cycle jobs in the field.

### 4.4.2 Formulation

We denote the decision variable as $Z_{tkj}$ which is a binary variable that takes a value of 1 if job $j$ was assigned to crew $k$ on day $t$ and a value of 0 otherwise.

| Notation | Brief Definition |
|---|---|
| $t$ | Day index |
| $k$ | Crew index |
| $j$ | Job index |
| $T$ | Planning horizon |
| $m_j$ | Priority multiplier of job j by job type |
| $s_j$ | Crew size minimum of job j (default 1) |
| $o_j$ | Number of days overdue for job j given its position in the schedule |
| $s_c$ | Size of crew c |
| $l_j$ | Time taken to complete job j (in hours) |
| $h_k$ | Available hours per day of crew k |
| $p$ | Penalty incurred for not assigning a job |
| $w$ | Maximum number of jobs that can be assigned to a crew k on a given day |

Table 4.2: Notation used in assignment optimization.

The full model is shown mathematically as:

**Objective:** Maximize completion of jobs that are most overdue across the planning horizon subject to a penalty for every job that is not assigned.

$maximize \sum_t \sum_k \sum_j o_j * m_j * Z_{tkj} - (1 - Z_{tkj}) * m_j * p$

subject to the constraints:

1. Each job is assigned to at most 1 crew:

$$\sum_j \sum_t^T \sum_k Z_{tkj} \leq 1$$

2. Crews cannot work over the number of hours in a shift (no overtime):

$$\sum_k \sum_t^T \sum_j l_j * Z_{tkj} \geq h_k$$

3. Crews cannot be assigned more than $w$ jobs in a day:

$$\sum_t^T \sum_k \sum_j Z_{tkj} \leq w$$

4. Crews cannot be assigned to work on a job j if they do not meet the minimum crew size requirement, $s_j$:

$$\sum_t^T \sum_k \sum_j Z_{tkj} * s_c \geq s_j$$

5. Binary constraints for job assignments:

$$Z_{tkj} \in 0, 1 \quad \forall t, k, j$$

Reviewing each term in the objective, the first $o_j * m_j * Z_{tkj}$ is applied when a job has been assigned and $Z_{tkj} = 1$. In this case, the objective seeks to maximize assignment of jobs with the greatest magnitude $o_j$ and $m_j$ which correspond to the number of days overdue of the job and the job type multiplier, respectively. In practice, job type multipliers are not well known however their magnitude can be tuned through primary field studies.

The second term of the objective applies when a job is not assigned and introduces the penalty associated with non-assignment of jobs from the backlog. This is considered a penalty term as it is decreasing the maximization in the event that a job is not assigned. The prioritization of overdue jobs should not come at the expense of letting upcoming jobs become overdue themselves, which reflects the trade-off that was previously discussed and which supervisors face daily. The penalty term in the objective incentivizes the assignment of jobs that do not contribute to the overall state of the backlog but that cannot be perpetually de-prioritized in favor of overdue job completion, as would be the case if only the first term in the objective was retained. While the penalty can be selected at run-time – allowing supervisors to tune how sensitive they need to be about jobs coming due – the default value is equivalent to the most overdue job. Therefore, the non-assignment of coming due jobs is equivalent in magnitude to the job that has the largest number of days overdue in the backlog.

This approach is similar to that taken in Schaefer's work discussed in Chapter 2.1. Schaefer et al also applies a penalty to poor crew pairings after searching a penalty space that is at first assumed to be 0 and modified as the solution moves from the deterministic plan.

Finally, constraint 3 was added as a heuristic determined by supervisors. In certain barns and regions, it is infeasible to complete more than a certain number of service jobs in a day and the assignment of such would make the model impractical. As a result, constraint 3 allows for input by the user at run-time to indicate practical job limits for a particular barn.

### Impact of Uncertainty in Job Assignment

When it occurs, trouble work is of immediate priority and requires at least the attention of a single trouble worker / one-person crews. As a result, one-person crews are assumed to have significantly reduced availability to complete service jobs.

This reduced availability can be reflected as a variable input to the assignment optimization. Along with daily estimated availability of all crews performing service work, the estimated man hours available for a one-person crew can be individually input allowing supervisors to indicate that as few as 1 hour per day of a trouble worker / one-person crew's time can be devoted to service work. Additionally, it is important to note that given the starting assumption of reduced availability, modeling trouble worker availability through a probabilistic distribution would likely produce immaterial changes in the outcome of scheduling results.

### Model Output

After the second optimization phase, scheduling and assignment have been completed and the model returns a pairing of single crews to sets of jobs on each day in the planning horizon.

# Chapter 5

# Results

Application of the mixed integer programs described in Chapter 4 can be viewed both against a historical baseline and through the comparison to alternative scheduling strategies. Overall, historical simulation of optimal scheduling and assignment produces an improved job accumulation state for barns of varying size and region, on the basis of the total number of days overdue of their backlog. In addition, optimal scheduling allows supervisors to manage the trade-off between overdue jobs and those coming due through the tuning of an objective parameter relating to those two metrics. We note that historical daily crew availability data is not available and assumptions have been made to ensure simulations of barn performance are realistic.

## 5.1 Structure

We evaluate performance of the two optimization formulations presented in Chapter 4 through two primary analyses:

- **Scheduling Stage:** We begin with an analysis of the scheduling formulation alone and compare it to two alternative scheduling strategies. This analysis allows us to determine whether the optimal scheduling formulation can perform as good or better than randomized or minimally-directed schedules whilst subject to realistic job prioritization criteria. In addition, we present the sensitivity of the scheduling stage to the coming due parameter with three example barns.

- **Sequence of Scheduling and Assignment Stages:** Use of the scheduling and assignment stages in sequence allow for a comparison to historical barn backlog variation over 1- and 3-month horizons. We observe how the optimal sequence quickly reduces the cumulative job backlog for service barns and maintains this reduction throughout these periods.

Unless otherwise stated, performance is measured by the reduction in cumulative number of days overdue of the backlog for a given barn.

## 5.2 Use of Open-Source Software

A public, open-source combinatorial optimization programming software (Google OR-tools) is used in the implementation of the optimization formulations [10]. This software is available without a license and without additional costs to the organization. It also provides extensive documentation to aid in future organizational support of the algorithms. The code was also created using freely-available software languages in order to allow for the continued support and potential expansion of these models within the organization.

## 5.3 Scheduling Optimization Results

As reported from the primary interviews discussed in section 3.1, schedules for short-cycle service jobs are typically produced for a shorter time horizon than for long-cycle jobs. Schedules may be produced for as few as 2 days out or for slightly longer, up to a week. Outside of these windows, it is difficult to predict the inflow of work (see Figure 3-5) and therefore to produce a reliable schedule.

As a result, time horizons of varying length up to 6 days can be compared to see how job accumulation varies in relation to other requisites like customer appointments and jobs coming due.

### 5.3.1 Scheduling Strategies

To evaluate the scheduling optimization results using the preferred parameter settings, two comparison strategies are developed to indicate how the model performs: a randomized schedule and a minimally-directed, naive schedule.

- **Random Strategy:** The use of randomized job allocation to days within a given planning horizon assuming a certain crew capacity.
- **Naive Strategy:** The use of a pre-sorted job list that prioritizes only those jobs which are most overdue to schedule first within a given planning horizon.
- **Optimal Strategy:** The use of the scheduling optimization algorithm to schedule jobs within a given planning horizon.

The naive strategy is reflective of the ability to sort jobs within the OMW application by those which are most overdue. Supervisors currently have this function and if effective, this strategy would represent a smaller upfront effort to operationalize than those of the optimization algorithms. Note that depending on the tuning of the trade-off terms discussed in the prior section, the naive schedule can be expected to outperform the optimal strategy from solely a backlog reduction perspective.

All schedule strategies use an arbitrary assignment approach wherein once the schedule is set, the jobs assigned on a particular day are random. Arbitrary assignment is critical to ensure that the scheduling strategies are fairly assessed. In this case, assignment is performed with a random accumulation of jobs that in total require less than 7 hours of estimated duration; this should be considered sub-optimal assignment given the random nature and any variation in the reduction of the job backlog from the perspective of the number of days overdue can be attributed to the scheduling strategy.

### 5.3.2 Strategy Comparison

Figure 5-1 shows a comparison of the three strategies discussed in the prior section using the six representative barns of interest first described in Table 3.1. Y-axis values are proprietary to Atlantic Utility and have been redacted. These trends are observed between Dec. 2021 - Jan. 2022. Appendix A includes tables that show the number of customer appointments satisfied and the percentage of jobs coming due that were scheduled with these varying strategies. Appendix tables are referenced throughout this section.

The metric of interest in these results is the cumulative number of days overdue of jobs assigned to a particular barn. In these results, an upward trend in backlog reduction is both expected and desired as the time horizon is increased from 2 days to 6 days. Increases in backlog reduction indicate that cumulative "overdueness" of the jobs assigned to a particular barn have decreased.

Comparing the optimal scheduling results with those of the naive and random schedules, a generally consistent behavior is observed: optimal results are typically more favorable than those in either alternative scheduling strategy, particularly in accommodating the scheduling of customer appointments and upcoming jobs. For example, Barn B in Figure 5-1 shows that the optimal solution presents a consistently decreasing backlog in tandem with scheduling 100% of jobs coming due in planning horizon windows of 4 days or more, shown in Table A.4. These results represent a relatively small barn in which the number of jobs coming due does not exceed 6, cumulatively requiring fewer than 4 hours in total; however, these results are encouraging as they indicate that job accumulation can be addressed without requiring supervisors to compromise to allow new jobs to become overdue.

In several scenarios, the optimal solution does not consistently exceed the naive strategy. Several explanations are presented in these cases and relevant tables are reproduced below:

- **Barn A**: Reviewing the daily jobs in this Region 1 barn, the extension of the time horizon from 2 to 3 days introduces 40% more jobs coming due in the latter time horizon than the former. Further the optimal strategy does not fall below 10% assignment of jobs coming due, rising so far as 41.2% in the 6 day time horizon, reference Table A.2. If the coming due multiplier is reduced, we would expect the optimal solution to approach the naive schedule. The sensitivity can

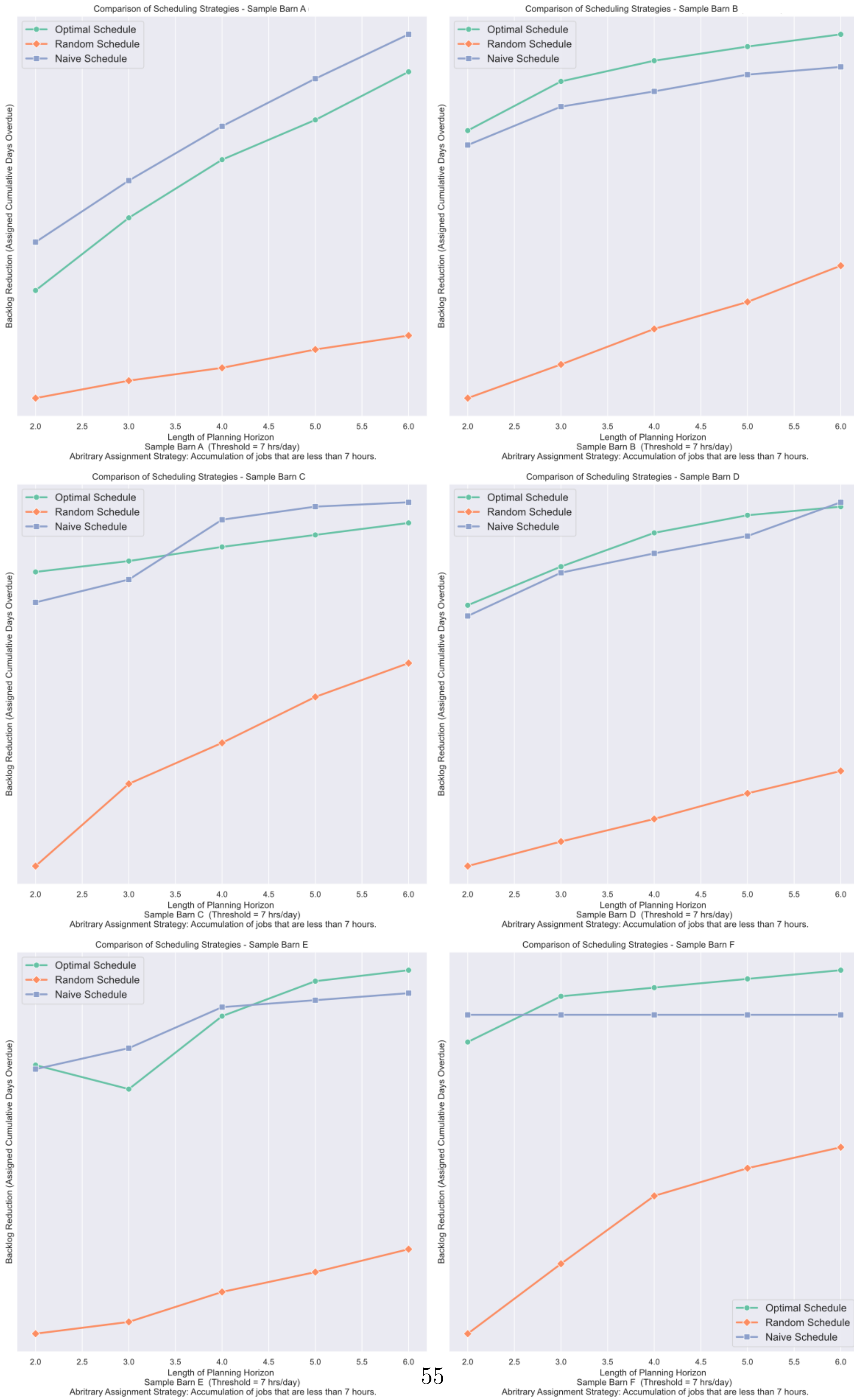


Figure 5-1: Scheduling Strategy Comparison Over 2- to 6-Day Planning Horizon

be tuned according to barn preference such that more or less jobs coming due are scheduled.

- **Barn C**: In this case, the barn is able to satisfy 100% of the jobs coming due within planning horizons of 3 days or greater as shown in Table A.6. As a result of this prioritization of jobs coming due, the optimal schedule falls slightly below the naive schedule in terms of backlog reduction. Again this trade-off is desired and the behavior of the model in this way is preferable to barn supervisors.

- **Barn E**: The optimal solution for this Region 3 barn is able to satisfy all customer appointments that are coming due as well as over 80% of all jobs coming due across 2- to 6-day time horizons as shown in Table A.9. The increase in the backlog from the 2-day to 3-day time horizon is due to the increase of 6 additional jobs coming due and 2 additional customer appointments required to be scheduled within the window.

It is important to note that in almost all cases, the naive strategy is unable to incorporate jobs that are coming due in tandem with scheduling overdue jobs. In cases where the naive appears superior to the optimal strategy, there is a supervisor preference for this optimal model behavior as it indicates a trade-off between scheduling overdue and coming due jobs.

In several cases, optimal results are shown not to satisfy all customer appointments, like the performance of Barn E shown in Table A.9 for example. Every instance of this behavior is the result of the customer appointments that are already past their due date. In these cases, constraint 3 in the scheduling optimization cannot be applied; instead, these customer appointments are treated the same as other accumulated jobs.

In total, while the application of the optimal strategy does not result in consistent reduction of accumulated jobs in all barns, it does allow for the concurrent prioritization of customer appointments, jobs coming due, and jobs which are overdue. This presents critical opportunities for supervisors who would not be required to manage competing priorities manually. As was shown in the naive schedule, the simple prioritization of overdue jobs risks the completion of other simple service jobs.

### 5.3.3 Impact of Coming Due Parameter

The trade-off parameter in the scheduling optimization provides supervisors flexibility in how job scheduling is performed. This parameter was first discussed in section 4.3.2 and represents the multiplier assigned to jobs coming due which can increase their contribution to the objective function. The greater this multiplier, the greater the optimization is incentivized to schedule upcoming jobs and prefers the assignment of these jobs over those that are overdue. Figures 5-2, 5-3, and 5-4 demonstrate how varying the coming due multiplier affects the scheduling optimization results for several barns.

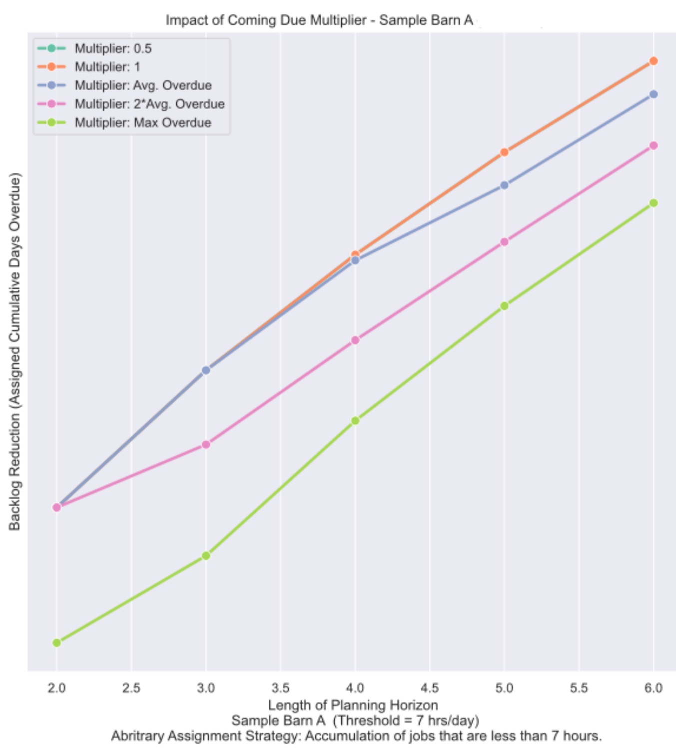


Figure 5-2: Impact of Coming Due Multiplier - Barn A (MA Region 1)

This allows us to see how the model behaves as overdue jobs and those that are coming due are weighted differently. As expected, lower multipliers on jobs coming due allow for greater backlog reduction. As multipliers increase, the backlog reduction is smaller in magnitude.

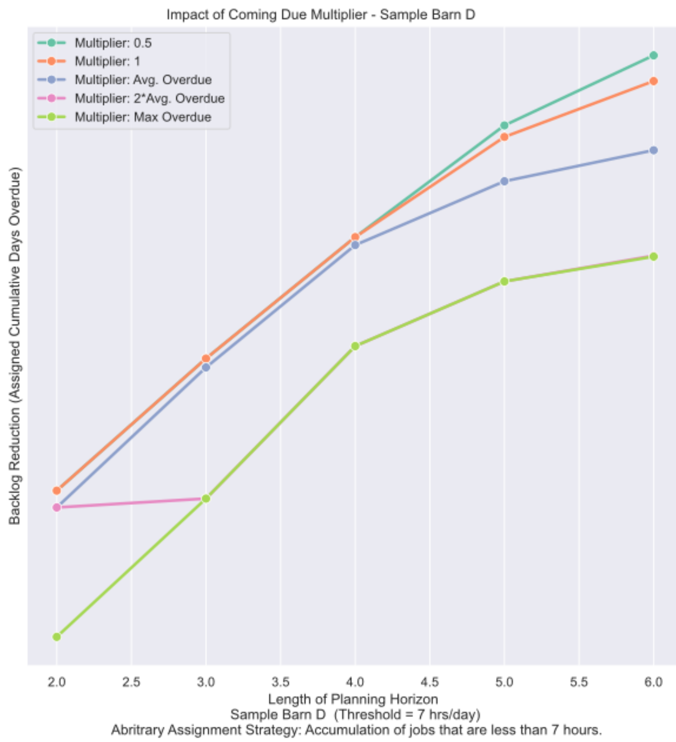


Figure 5-3: Impact of Coming Due Multiplier - Barn D (NY Region 4)

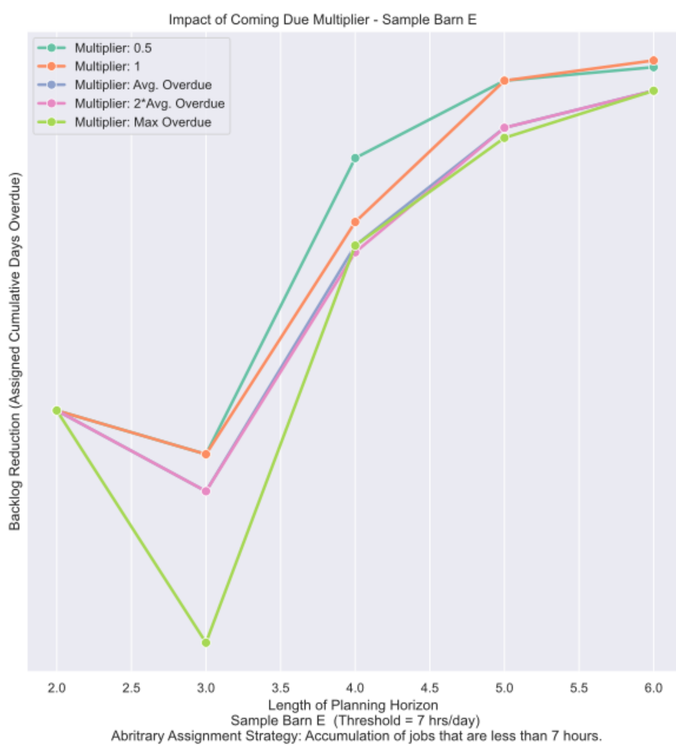


Figure 5-4: Impact of Coming Due Multiplier - Barn E (NY Region 5)

In some cases, a greater multiplier impact is observed over longer planning horizons. It is possible this behavior is the result of a small percentage of highly overdue jobs dominating the optimal solution over shorter time horizons to dictate the initial scheduling. As time horizons increase and the magnitude of overdue jobs and coming due jobs are of comparable value, the multiplier is able to sway scheduling results more heavily. This reasoning would also explain why initial periods in the staged optimization results, reference section 5.4, see a sharp decrease at the start of the month, as very overdue jobs dictate scheduling.

As the daily operations of individual barns may differ, we recognize these parameters are subject to change at the discretion of the barn supervisors. For the analysis in this chapter, we apply an intermediate value for the coming due multiplier: the average overdue value for accumulated jobs.

## 5.4 Staged Optimization Results: Scheduling and Assignment

Extending the scheduling phase to include assignment, it is possible to simulate how the accumulated jobs in a barn would have changed in comparison to the historical performance over one or more months. Figure 5-5 presents these results for the six barns of interest; once again, the Y-axis values are proprietary to Atlantic Utility. The historical performance is observed through job creation and completion dates; optimal results are simulated by iteratively running staged scheduling and assignment beginning with open jobs on Sept. 1, 2021. It is assumed that optimal results are followed on each subsequent day of the month. Reviewing the historical performance with perfect hindsight is similar to the analysis performed in Angalakudati et al's work [1].

Due to limited historical data availability regarding crew capacity, several assumptions were made to perform this simulation. The number of crews and their daily availability was approximated to be close to 10% to that which was performed histor-

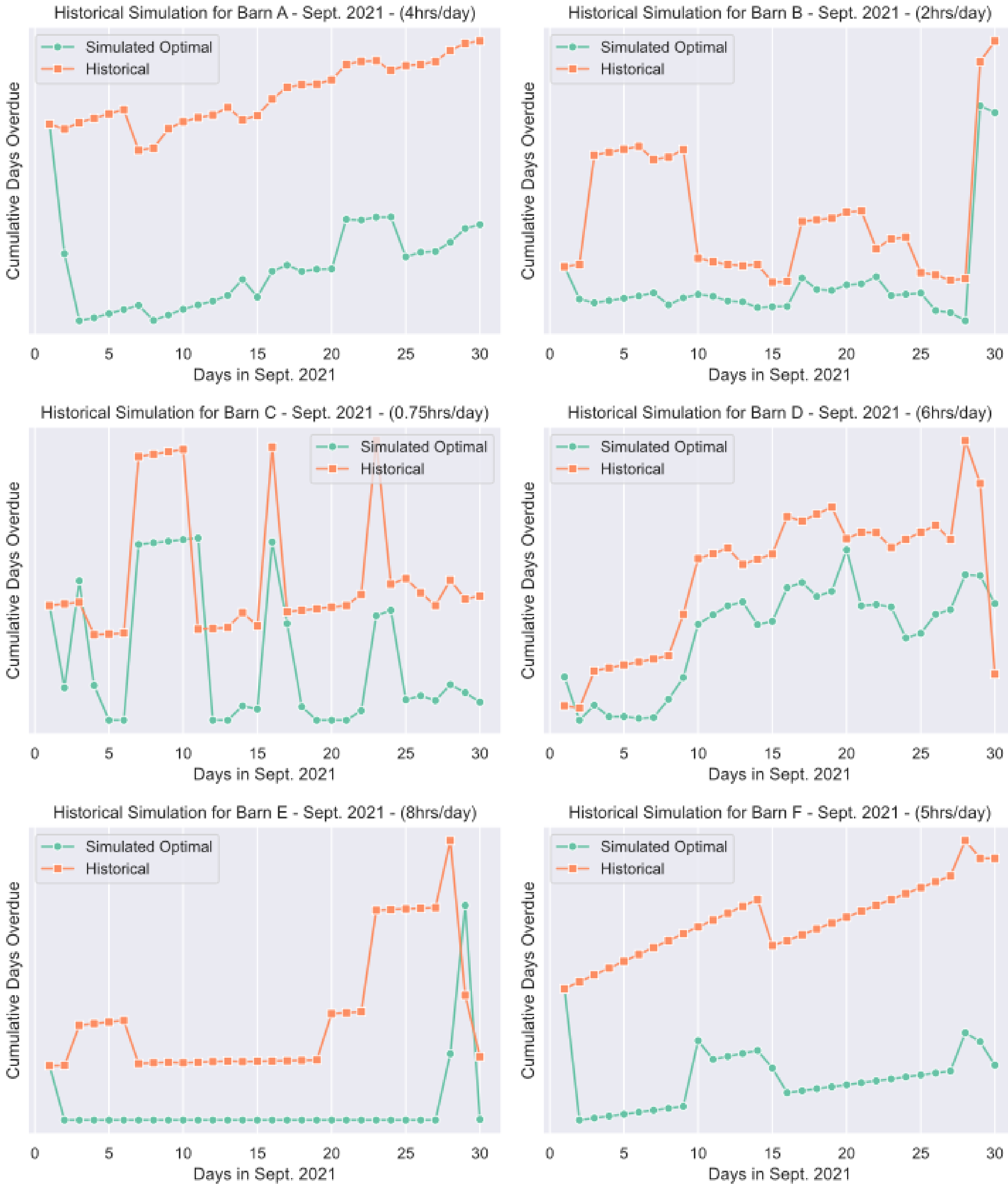


Figure 5-5: Backlog Comparison Over 1-Month Historical Simulation

ically, as shown in Table 5.1. For example, if a barn was historically able to complete 100 jobs in a month, the simulation would apply an appropriate number of crews such that between 90 and 110 jobs were completed. It is not possible to exactly match the number of jobs completed historically as the optimal simulation may not select jobs with equivalent estimated duration.

| Barn | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| Historical | 184 | 81 | 25 | 99 | 62 | 78 |
| Simulated | 184 | 79 | 24 | 102 | 56 | 72 |
| % Difference | 0% | 2.5% | 4% | 10.1% | 9.6% | 7.7% |

Table 5.1: Jobs Completed Over a Single Historical Month

Additionally, though it is not necessary to set the scheduling threshold to be larger than the crew availability, in practice it is likely that supervisors will allot more work than is able to be completed by field crews; this could be the result of jobs requiring longer to complete than their originally estimated duration or as a result of trouble workers being called away to complete unscheduled work. Simulation results incorporate this behavior and assume crew availability is always less than the scheduling threshold.

Figure 5-5 shows a consistent pattern in which optimal results produce a significantly reduced backlog accumulation than the historical performance. This suggests that the staged optimizations are able to begin the month with the same job accumulation and handle equivalent inflow of jobs while reducing the cumulative backlog of jobs in any given barn. In most cases, a steep decline is observed initially – jobs which are significantly overdue are assigned first – and then the subsequent results outperform those historically. This is particularly interesting as these simulations are not assuming any increase in capacity among barns, as shown in Table 5.1. We can further observe how similarly the staged optimizations are able to match historical job completion in Figure 5-6 which is presented over 3 historical months (September - November 2021).

These results suggest that the optimization formulations are targeting *different jobs* in comparison to those historically chosen in order to produce cumulative backlog improvements. This behavior is opposed to simply assigning crews to complete more jobs, which may be infeasible given current capacity limitations. Figure 5-6 shows

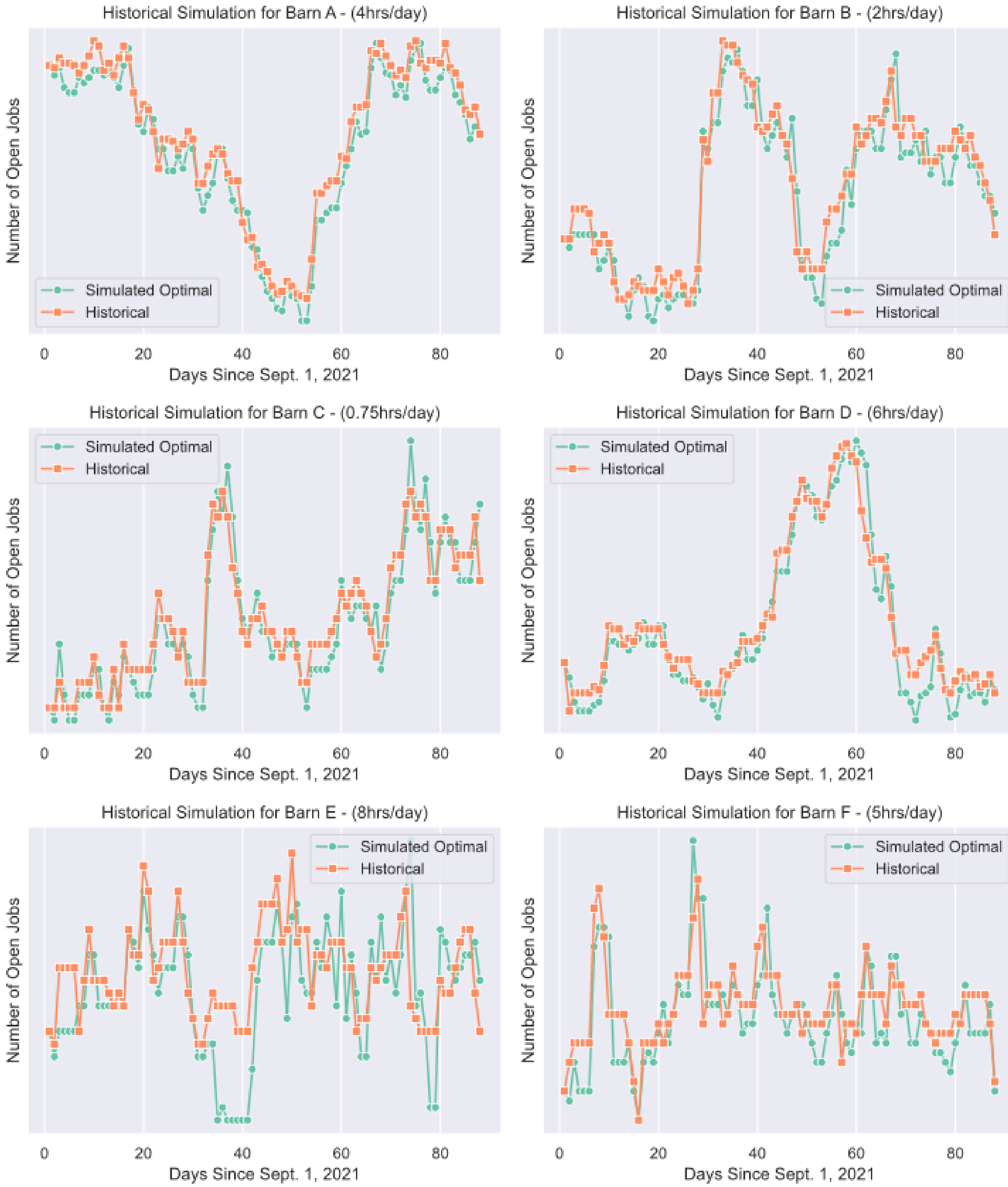


Figure 5-6: Open Job Comparison Over 3-Month Historical Simulation

the same metric – number of open jobs – as presented in Figure 3-5 and we can see that the variability in job completion totals remains. This is expected given that the objective metrics evaluate the number of days overdue as opposed to the number of jobs completed.

Continuing to observe optimal results over a 3 month period, we see that in some cases, the variability in barn backlog is very similar to that of historical performance as shown in Appendix Figure B-2. This could be the result of the inflow of jobs that are already past their due date causing spikes in the cumulative backlog metric. Reducing spikes in the backlog, and allowing greater predictability in backlog reduction, is a potential extension topic for this work.

## 5.5 Results Summary

Proof-of-concept results of the first scheduling phase and combined phases indicate a promising opportunity to reduce the cumulative backlog in a given barn. If these programs were included in the OMW application, supervisors would have the option to solely schedule or sequentially schedule and assign work with the assurance that those jobs which are highest priority – those past due and coming due – would be given the greatest weight in the results. Parameters for the non-assignment penalty and for the weight of jobs coming due are variable and we show that reasonable defaults can also be used to reduce manual intervention in the optimizations.

THIS PAGE INTENTIONALLY LEFT BLANK

# Chapter 6

# Recommendations for Operationalization

Results of the staged optimizations warrant additional analysis and Atlantic Utility is investigating the launch of a pilot within the OMW application to observe their impact on field operations. While development of the proof-of-concept algorithms is a prerequisite to pursuing a pilot, additional consideration of how to deploy these optimizations for real-time usage is needed.

Organizational use of optimization algorithms is common across industries including energy, telecommunications, aviation, logistics, and finance [11]; however, the deployment of these algorithms and their use in daily operations can be challenging as teams need to consider real-time performance, algorithm maintenance, and testing. With an extended lifecycle management process, these applications can require software infrastructure and significant computing resources as well as widespread cultural adoption and the development of new skills across teams [15].

Despite these challenges however, effective deployment of complex algorithms for organizational use, including many machine learning and optimization algorithms, can be achieved through the application of continuous integration (CI) and continuous delivery (CD) principles – frequently referred to as "CI/CD". Aggregated research performed by Bolscher et al. provides numerous examples of the potential benefits of these principles including "higher deployment frequency, shorter lead time, improved automated testing, feedback gathering and problem solving, fewer escalations...more public facing services and an increased velocity [3]." In fact, the application of these principles is known to be growing from a "niche to a mainstream strategy employed by 25% of Global 2000 organizations" and suggests that organizations face a competitive risk if they opt for non-adoption [3].

This chapter first reviews CI/CD principles and describes what benefits they provide to organizations that employ them to operationalize machine learning and optimization algorithms. The chapter concludes with a discussion of trade-offs between the use of in-house custom software and external third-party software for implementation.

## 6.1 Continuous Integration and Continuous Delivery

CI/CD principles derive from a set of software development and IT operations practices known as DevOps. DevOps is a "practice that aims at merging development, quality assurance, and operations (deployment and integration) into a single, continuous set of processes" and is part of the larger philosophy of agile software development [9]. While the process is typically considered continuous and iterative, as opposed to having defined beginning and ending point, primary steps include [9]:

- **Planning** releases, features, and sprints;
- **Development** of code by engineers in small blocks that is pushed to the central repository multiple times a day;
- **Testing**, usually by a quality assurance team, that produces automated testing for committed code and maintains version control;
- **Deployment** to a production or lower environment server in a way that does not negatively affect existing features; and,

- **Monitoring** for deficiencies in product functioning and collection of feedback to assist with future planning.

Continuous integration refers to development and testing steps when code is being pushed and merged into the shared repository. Merging developers' code into shared branches multiple times per day, per week, and per release cycle ensures that parallel development by one engineer does not conflict with that of another [4]. Continuous delivery refers to the combination of development, testing, and deployment steps such that automation between these steps can be introduced and pipelines can be created [9].

The application of DevOps, including CI/CD principles, to machine learning algorithms is termed "MLOps" [7]. While some of these components are the same for any software system deployment, machine learning and optimization algorithms have more specialized requirements, as listed below. Additionally, though this term refers to deployment of machine learning algorithms, the general strategy associated with optimization algorithm deployment is similar. Several key differences between DevOps and MLOps include [7]:

- **Software Engineer Experience:** The experimental nature of the algorithm requires parameter tuning and analysis of results in the field indicate that software engineers with experience in both developing mixed integer programs and deploying production-class services be available on the team to maintain the algorithm.

- **Multi-step Pipeline:** As the algorithm is intended to be iteratively improved as it is tuned and as additional data becomes available, there is a need for it to be deployed via a multi-step pipeline that can automate steps to easily implement changes and re-deploy the model. This is different from other production-ready software which is typically expected to have fewer structural changes.

- **Potential for Reduced Performance:** Sub-optimal coding or necessary computation time required to produce model results can impact the use of the model and deviate from operational standards for software deployment.

As a result of these differences, CI/CD principles are also modified such that continuous integration also refers to testing the model results and validating performance [7]. Continuous development requires the use of a deployment pipeline and the use of a service that can run and support the optimization model. Integrating new code now extends to validation of data schema that are used as input to the model and parameter testing.

In sum, the proof-of-concept optimization code only accounts for one portion of the components required to implement a real-world optimization solution. Additional components include data collection, process management, configuration, automation, testing, and debugging [7]. Industry standards for these practices make use of CI/CD principles that recognize the differences between traditional software deployment and that required for optimization or machine learning programs. These principles and the challenges associated with deployment are a significant consideration for organizations looking to employ these algorithms and when considering whether to use in-house tools or to license external solutions. This thesis recommends the use of CI/CD principles in the deployment and maintenance of the discussed optimization algorithms within any organization.

## 6.2 Trade-off between In-House and External solutions

The build-or-buy decision in an organization is typically guided by cost considerations in terms of time, computational resources, and headcount [13].

Organizations that choose to build in-house DevOps solutions should expect significant time to be required to complete these projects as there is a requisite for hardware resources, potential licensing integrations with data science tools or optimization software like Gurobi, monitoring instrumentation, and cluster maintenance [13]. Building an in-house machine learning platform could take from between three months to over a year of work, during which time the opportunity cost of not having the algorithm fully deployed could be significant [13]. In addition, sustained maintenance of not just the algorithm but also the platform would then be required and challenges with scalability across the organization are anticipated.

Computational resources are also a significant expense when building an in-house DevOps solution. Depending on the size and integration needs of the system, the cost of servers, establishing the network of connected devices, and potentially hosting the on-premise network in physical facilities are estimated to be 5 - 20% of total costs during the lifecycle of the project [8].

As described above, workforce requirements to build and maintain a DevOps solution can be challenging to satisfy; organizations have identified the need for a team of at least 10 experienced software engineers or data scientists with knowledge of Cloud-based platforms, frontend development, backend databases, networking, and testing [13]. Finally, it can be challenging to even identify reference sources or existing literature to recognize the relationships between software architecture, DevOps, and complex applications [3] – further straining organizations that do not specialize in software development and deployment.

Instead, use of a ready-to-use platform can alleviate some of the challenges associated with developing and deploying these applications as they "provide end-to-end lifecycle management...[and] a generic solution for multiple development scenarios that need to accomplish difference ML tasks [15]." The setup time is significantly reduced when using established, extensible platforms as there is little to no server setup or network configuration [15]. In addition, computation time can be quite minimal with CI/CD tools established for algorithm deployment: Zhou et al. presents a case study of available tools used to build an ML platform including Kubernetes, Kubeflow, Gitea, and Drone in which all CI/CD pipelines were completed in less than 1 minute. Operationalization of algorithms face a similar requirement as they will likely be used by non-technical employees, and therefore need to have a level of performance that does not degrade the end-user experience.

Zhou also discusses the reduction in computing resources needed in accomplishing steps of the the CI/CD pipeline when using tools like Kubeflow as they ensure a high GPU (Graphic Processing Unit) utilization, making the most use of available resources. Assuming GPU does not become a bottleneck in the process, Zhou reports that time and resource consumption do not vary significantly based on the model parameters. This is an attractive feature as it indicates that "design, construction, and implementation of the models" is the biggest determination of performance and which can be aided by the use of existing CI/CD tools [15].

Finally, it is expected that fewer employees will be needed throughout the lifetime of the project when using an existing CI/CD platform as much of the basic configuration and features are provided out-of-the-box for tools like Kubeflow. Available documentation and the ability to extend these products across business units within an organization also make these attractive tools. This thesis recommends the investment in an existing CI/CD platform tool to reduce initial upfront configuration and hardware costs, limit the time in which algorithms are not production ready, and ensure optimal usage of available computing resources.

# Chapter 7

# Conclusions and Future Work

This thesis provides justification for a larger pilot within the OMW application to investigate the impacts of staged, mixed integer programs on field force operations in the Electric Business Unit at Atlantic Utility. Service work variability exists across the field force network of barns – both in terms of the number of service jobs coming in and the capacity of field crews available – and induces job accumulation as a result of unpredictability in inflow and outflow. Without a consistent way to tackle this accumulation, while concurrently prioritizing new jobs coming in, there is the potential that jobs are completed in a sub-optimal order allowing overdue jobs to become more overdue or new jobs to pass their due date. The staged optimization approach provides this consistency that can be applied across the EBU network of barns while still allowing for barn-specific tuning of parameters.

Results of the first phase indicate that optimal scheduling can outperform naive strategies based on the magnitude of individual barn backlog reduction, measured in days. Though naive strategies like job sorting could be implemented with current functionality in the OMW application, these do not consistently produce as positive a performance in reducing a given barn's backlog. Primary research conducted with members of the EBU shows evidence of a trade-off that supervisors encounter, requiring them to decide how to prioritize jobs that are past due versus those that are coming due. Optimal scheduling can effectively incorporate this trade-off through use of a variable penalty guided by user input; over increasing time horizons, these schedules also showed increasing ability to incorporate a balance of both overdue and coming due jobs into daily schedules.

Combining scheduling and assignment phases shows significant improvement in barn backlog reduction over a simulated historical month in comparison to true performance. The combined use of these programs can significantly reduce job accumulations in barns with the dual benefit of ensuring uniformity in how barns tackle existing job accumulations and reducing the amount of time and potentially suboptimal schedules generated with manual preparation.

While current data limitations prevented investigation into specific monetary metrics that might be affected through optimal scheduling and assignment, there is significant opportunity to incorporate these valuations into the objective functions of these programs. For example, comparison of the time spent performing scheduling or assignment actions between supervisors in barns using OMW and those not yet using the tool could allow for an estimate of time savings against baseline performance. The results of this thesis recommend the continuation of this proof-of-concept work to more directly quantify the effects of optimal scheduling and assignment on barn performance and crew utilization.

## 7.1 Future Work

There are multiple opportunities to extend this work through a field pilot of the mixed integer formulations developed in this thesis. Future studies may investigate the incorporation of additional data sources including crew composition, job material requirements, and job description or type into the priority determination of specific jobs. Additionally, the focus of this thesis is specific to individual barn optimization, however network-level views of on-time service job completion rates or of regulatory compliance can be pursued through an optimization lens. Finally, the relaxation of some assumptions made in these programs could introduce additional robustness including the allowance for some crew overtime and the incorporation of the uncertainty of some trouble work.

Outside of the utilities industry, applications of the sequence of scheduling and assignment stages can be considered in logistics and delivery settings. This environment can share similar short-term delivery expectations with the need to weigh trade-offs between on-time and overdue deliveries. We could envision similar daily or weekly planning scenarios in which deliveries have specific due dates that must be maintained to meet customer expectations but that are of variable distance and have minimum crew availability requirements. These scenarios are similar to the work described in this thesis and present opportunities for extension of this work to additional fields of study.

THIS PAGE INTENTIONALLY LEFT BLANK

# Appendix A

# Tables

| | T = 2 | T = 3 | T = 4 | T = 5 | T = 6 |
|---|---|---|---|---|---|
| Data | 3 | 3 | 3 | 3 | 3 |
| Random | 0 | 0 | 0 | 0 | 0 |
| Naive | 0 | 0 | 0 | 0 | 0 |
| Optimal | 2 | 2 | 2 | 2 | 2 |

Table A.1: No. of Customer Appointments Scheduled (A)

| | T = 2 | T = 3 | T = 4 | T = 5 | T = 6 |
|---|---|---|---|---|---|
| Random | 10.0% | 0% | 6.7% | 6.2% | 0% |
| Naive | 0% | 0% | 0% | 0% | 0% |
| Optimal | 20.0% | 14.3% | 20.0% | 37.5% | 41.2% |

Table A.2: No. of Jobs Coming Due Scheduled (A)

| | T = 2 | T = 3 | T = 4 | T = 5 | T = 6 |
|---|---|---|---|---|---|
| Data | 0 | 0 | 0 | 0 | 0 |
| Random | 0 | 0 | 0 | 0 | 0 |
| Naive | 0 | 0 | 0 | 0 | 0 |
| Optimal | 0 | 0 | 0 | 0 | 0 |

Table A.3: No. of Customer Appointments Scheduled (B)

| | T = 2 | T = 3 | T = 4 | T = 5 | T = 6 |
|---|---|---|---|---|---|
| Random | None | None | 100% | 100% | 100% |
| Naive | None | None | 0% | 0% | 33.3% |
| Optimal | None | None | 0% | 0% | 0% |

Table A.4: No. of Jobs Coming Due Scheduled (B)

| | T = 2 | T = 3 | T = 4 | T = 5 | T = 6 |
|---|---|---|---|---|---|
| Data | 4 | 4 | 4 | 4 | 4 |
| Random | 1 | 1 | 3 | 3 | 2 |
| Naive | 1 | 2 | 3 | 4 | 4 |
| Optimal | 0 | 0 | 0 | 0 | 0 |

Table A.5: No. of Customer Appointments Scheduled (C)

| | T = 2 | T = 3 | T = 4 | T = 5 | T = 6 |
|---|---|---|---|---|---|
| Random | None | 71.4% | 60.0% | 52.9% | 82.4% |
| Naive | None | 0% | 0% | 0% | 35.3% |
| Optimal | None | 100% | 100% | 100% | 100% |

Table A.6: No. of Jobs Coming Due Scheduled (C)

| | T = 2 | T = 3 | T = 4 | T = 5 | T = 6 |
|---|---|---|---|---|---|
| Data | 4 | 4 | 4 | 4 | 4 |
| Random | 1 | 2 | 1 | 1 | 2 |
| Naive | 1 | 2 | 4 | 4 | 4 |
| Optimal | 0 | 0 | 0 | 0 | 0 |

Table A.7: No. of Customer Appointments Scheduled (D)

| | T = 2 | T = 3 | T = 4 | T = 5 | T = 6 |
|---|---|---|---|---|---|
| Random | 0% | 0% | 66.7% | 41.7% | 36.8% |
| Naive | 0% | 0% | 0% | 0% | 0% |
| Optimal | 50.0% | 33.3% | 33.3% | 75.0% | 89.5% |

Table A.8: No. of Jobs Coming Due Scheduled (D)

| | T = 2 | T = 3 | T = 4 | T = 5 | T = 6 |
|---|---|---|---|---|---|
| Data | 7 | 9 | 9 | 9 | 10 |
| Random | 1 | 6 | 3 | 8 | 12 |
| Naive | 2 | 3 | 4 | 4 | 4 |
| Optimal | 3 | 4 | 6 | 6 | 6 |

Table A.9: No. of Customer Appointments Scheduled (E)

| | T = 2 | T = 3 | T = 4 | T = 5 | T = 6 |
|---|---|---|---|---|---|
| Random | 0% | 22.2% | 0% | 25.0% | 33.3% |
| Naive | 0% | 0% | 0% | 0% | 0% |
| Optimal | 100% | 88.9% | 90.0% | 83.3% | 80.0% |

Table A.10: No. of Jobs Coming Due Scheduled (E)

| | T = 2 | T = 3 | T = 4 | T = 5 | T = 6 |
|---|---|---|---|---|---|
| Data | 8 | 8 | 9 | 9 | 9 |
| Random | 3 | 5 | 9 | 11 | 9 |
| Naive | 5 | 8 | 9 | 9 | 9 |
| Optimal | 5 | 8 | 9 | 9 | 9 |

Table A.11: No. of Customer Appointments Scheduled (F)

| | T = 2 | T = 3 | T = 4 | T = 5 | T = 6 |
|---|---|---|---|---|---|
| Random | 25.0% | 50.0% | 66.7% | 66.7% | 83.3% |
| Naive | 0% | 16.7% | 33.3% | 50.0% | 100% |
| Optimal | 50.0% | 66.7% | 75.0% | 75.0% | 75.0% |

Table A.12: No. of Jobs Coming Due Scheduled (F)

THIS PAGE INTENTIONALLY LEFT BLANK

# Appendix B

# Figures

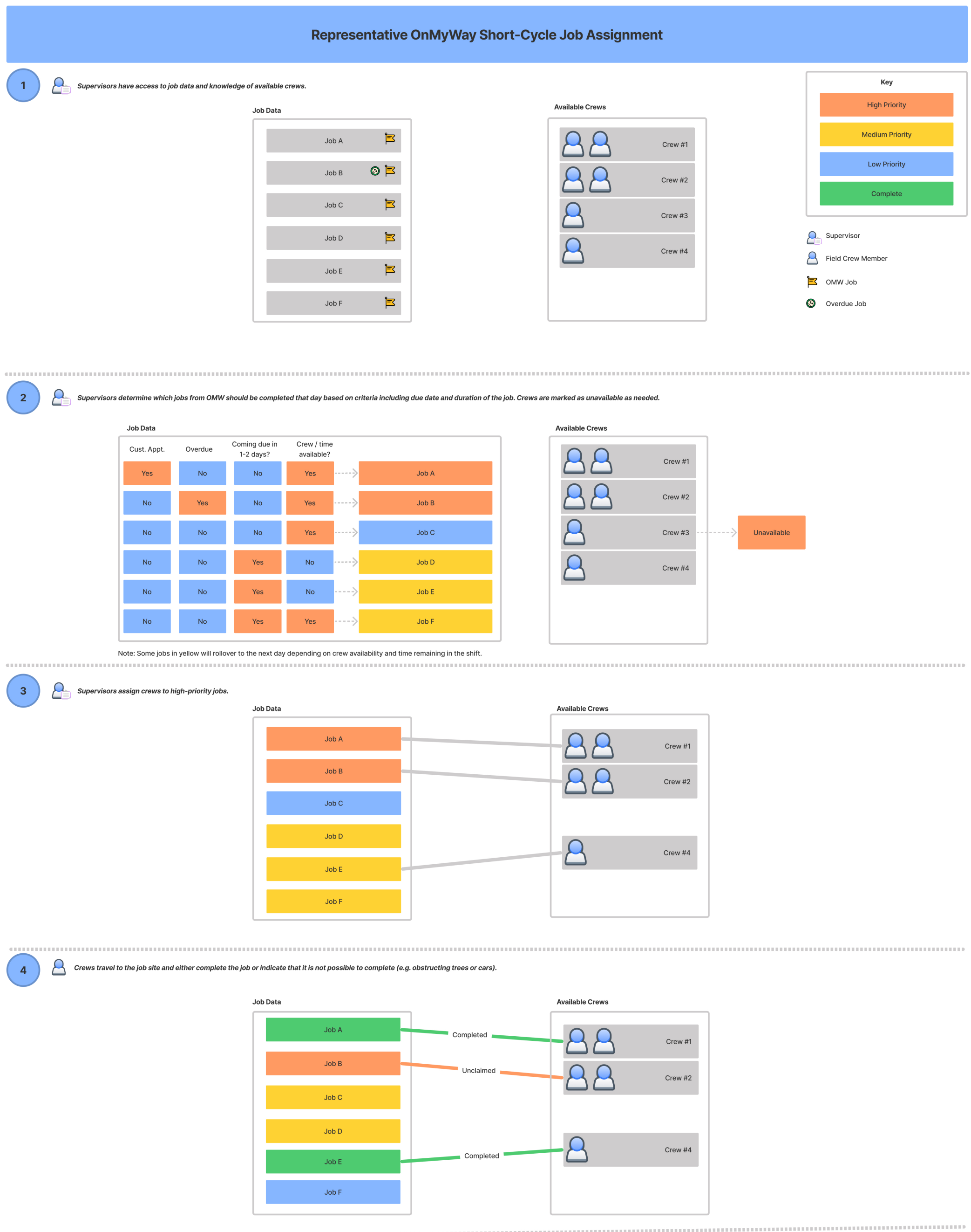


Figure B-1: Representative Short-Cycle Job Assignment Process

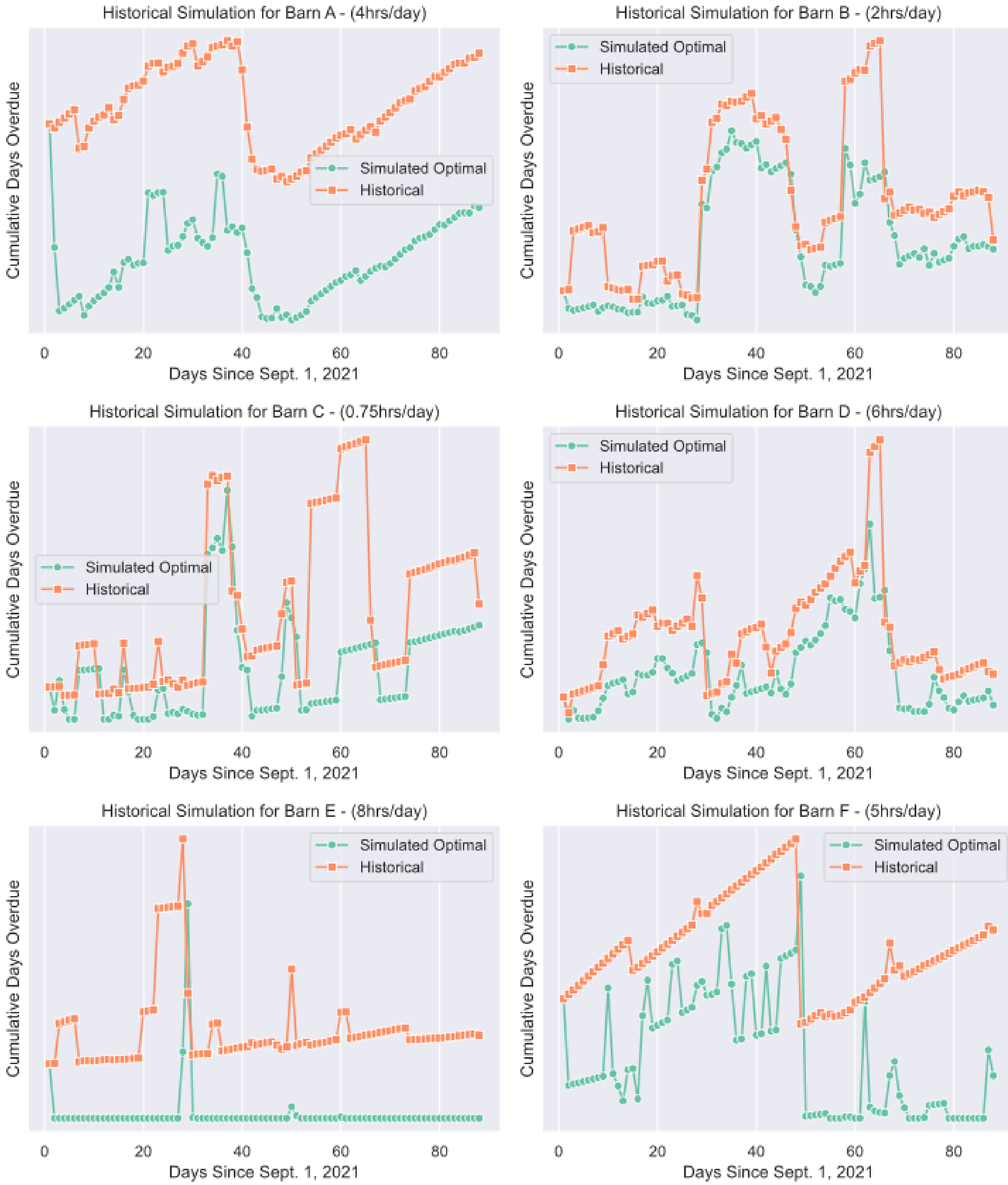


Figure B-2: Backlog Comparison Over 3-Month Historical Simulation

THIS PAGE INTENTIONALLY LEFT BLANK

# Bibliography


[1] Mallik Angalakudati, Siddharth Balwani, Jorge Calzada, Bikram Chatterjee, Georgia Perakis, Nicolas Raad, and Joline Uichanco. Business analytics for flexible resource allocation under random emergencies. *Management Science*, 60(6):1552–1573, 2014.

[2] Siddharth Balwani. Operational efficiency through resource planning optimization and work process improvement. Master's thesis, Massachusetts Institute of Technology, Electrical Engineering and Computer Science Department, June 2012.

[3] Robin Bolscher and Maya Daneva. Designing Software Architecture to Support Continuous Delivery and DevOps: A Systematic Literature Review. In VanSinderen, M and Maciaszek, L, editor, *ICSOFT: PROCEEDINGS OF THE 14TH INTERNATIONAL CONFERENCE ON SOFTWARE TECHNOLOGIES*, pages 27–39, 2019. 14th International Conference on Software Technologies (ICSOFT), Prague, CZECH REPUBLIC, JUL 26-28, 2019.

[4] Continuous integration. `https://en.wikipedia.org/wiki/Continuous_integration`, Jan 2022.

[5] Vitor Hugo Ferreira, Rubens Lucian da Silva Correa, Angelo Cesar Colombini, Márcio Zamboti Fortes, Flávio Luis de Mello, Fernando Carvalho Cid de Araujo, and Natanael Rodrigues Pereira. Big data analytics for spatio-temporal service orders demand forecasting in electric distribution utilities. *Energies (Basel)*, 14(23):7991–, 2021.

[6] Marc Goldsmith. The merger of technology talent. *Mechanical Engineering*, 139(1):34 – 39, 2017.

[7] Mlops: Continuous delivery and automation pipelines in machine learning | google cloud. `https://cloud.google.com/architecture/mlops-continuous-delivery-and-automation-pipelines-in-machine-learning`, Jul 2020.

[8] Ville-Veikko Helppi. The costs of in-house vs. cloud-based development environments. `https://bitbar.com/blog/the-costs-of-in-house-vs-cloud-based-development-environments/`, Jan 2020.

[9] Ioannis Karamitsos, Saeed Albarhami, and Charalampos Apostolopoulos. Applying devops practices of continuous automation for machine learning. *Information*, 11(7), 2020.

[10] About or-tools | google developers. `https://developers.google.com/optimization/introduction/overview?hl=US`. [Accessed: 01.12.2022].

[11] Edward Rothberg. Council post: Should your company be using mathematical optimization? ask yourself these four questions to find out. https://www.forbes.com/sites/forbestechcouncil/2020/07/07/should-your-company-be-using-mathematical-optimization-ask-yourself-these-four-questions-to-find-out/?sh=163897f867ec, Jul 2020.

[12] Andrew J. Schaefer, Ellis L. Johnson, Anton J. Kleywegt, and George L. Nemhauser. Airline crew scheduling under uncertainty. *Transportation Science*, 39(3):340–348, 2005.

[13] Build internally vs buy? a fundamental decision in machine learning operations (mlops). `https://spell.ml/blog/build-vs-buy-machine-learning-platform-mlops-XlasqREAACQA5Jhl`, Jan 2020.

[14] Sean David Whipple. Predictive storm damage modeling and optimizing crew response to improve storm response operations. Master's thesis, Massachusetts Institute of Technology, Electrical Engineering and Computer Science Department, June 2014.

[15] Yue Zhou, Yue Yu, and Bo Ding. Towards mlops: A case study of ml pipeline platform. In *2020 International Conference on Artificial Intelligence and Computer Engineering (ICAICE)*, pages 494–500, 2020.