Human-centered approaches to system level design with applications to desalination

by

Bo Yang Yu

B.A.Sc., Mechatronics Engineering, University of Waterloo, 2009
M.A.Sc., Mechanical Engineering, University of Waterloo, 2010

Submitted to the Department of Mechanical Engineering in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2015

© Massachusetts Institute of Technology 2015. All rights reserved.

Signature redacted

Author ................

Department of Mechanical Engineering

November 19, 2014

Signature redacted

Certified by ........

Maria C. Yang
Associate Professor of Mechanical Engineering
Thesis Supervisor

Signature redacted

Accepted by ........................................

David Hardt
Professor of Mechanical Engineering
Chairman, Committee on Graduate Students
Human-centered approaches to system level design with applications to desalination

by

Bo Yang Yu

Submitted to the Department of Mechanical Engineering on November 17, 2014, in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Engineering

Abstract

The goal of this thesis is to better understand the design practice employed by the desalination industry from a systems-level viewpoint and offer recommendations to improve the process of design. A human-centered design research approach is used, in which industry practitioners were interviewed about the design strategies they employ for industrial desalination systems.

A common theme from the interviews is that long term performance of desalination systems needs to be emphasized during the design process. Based on this observation, a novel design framework is proposed that incorporates health monitoring and maintenance in the design stage. The proposed framework is shown to make the design process more effective and can result in more optimal design over the lifecycle of a plant.

The interviews also suggest open questions about how designers use computational tools for the design of desalination systems. An investigation of how designers respond to the complexity of desalination parameter design problems was conducted. The behavior of designers during a series of simulated design processes involving seawater reverse osmosis (SWRO) plants was observed. The experiments revealed that desalination knowledge seemed to lead to better performance, but the results also showed that subjects with limited desalination knowledge could perform worse than subjects with no desalination knowledge. It was also observed that human designers had difficulties understanding the sensitivity of coupled variables, which can lead to poor performance on parameter design problems. Additionally, the top-performance-ranked subjects were observed to behavior very differently from the bottom-ranked. The problem-solving profiles of best performing subjects resembled a well-tuned simulated annealing algorithm while the worst performing subjects used a pseudo random search strategy.

Thesis Supervisor: Maria C. Yang
Title: Associate Professor of Mechanical Engineering
Acknowledgments

Words cannot describe how blessed I am to have Dr. Maria Yang as my supervisor over the three years I spent in her Ideation Lab. Not only is she incredibly knowledgable, she is the most optimistic person I have ever had a chance to meet. Her positive energy has made this journey an enjoyable one even during the occasional tough times.

I was also very fortunate to have Dr. Oli de Weck, Dr. Dan Frey and Dr. John Lienhard V to be on my doctoral committee. Their professionalism, experience and creativity were invaluable to the development of this thesis and also my personal growth as a researcher.

I would like to offer my gratitude to every member of the Ideation Lab for creating a family-style environment that is both scientific and artistic, while always keeping the areas surrounding my desk space super clean and free of obstacles; my UROPs, members of the CAD lab, Lienhard’s lab, Mitsos’ lab, the bio-instrumentation lab, and collaborators in KFUPM and CCES for their continued help during my study. Special thanks to the Center for Clean Water and Clean Energy at MIT and KFUPM (project number R13-CW-10) and National Science and Engineering Council of Canada for funding this work, and Dr. Hadjiconstantinou for providing me with a TA opportunity in my first semester at MIT.

Special thanks to all my friends and family for the fun and warmth. To my parents for making the drive from Toronto every year to bring me delicious food. And finally, to the love of my life Yili, who gave up her chance to a Ph.D so that I could see her beautiful face everyday, thank you for all of your love.
## Contents

### 1 Introduction

1.1 Motivation .......................................................... 15
  1.1.1 Human Centered Design Approach ........................... 16
  1.1.2 Maintenance Strategies ....................................... 17
1.2 Background on Desalination Technology and Maintenance ...... 18
1.3 Thesis Contribution .................................................. 20
  1.3.1 Human-centered approach to design process research ....... 20
  1.3.2 Framework for maintenance in design stage ................. 20
  1.3.3 Characterization of desalination designer behaviors ..... 21
1.4 Thesis Organization .................................................. 21

### 2 Understanding Challenges of Designing Desalination Systems 23

2.1 Motivation ............................................................ 23
2.2 Methodology ........................................................... 23
2.3 Results ................................................................. 24
  2.3.1 Municipal Water/Energy System ............................... 24
  2.3.2 Industrial Water Treatment Systems ......................... 27
  2.3.3 Use of Numerical Optimization ............................... 27
  2.3.4 Design for Lifecycle Optimality ............................. 28
  2.3.5 Other Themes .................................................. 29
2.4 Summary ............................................................... 30
3  A Maintenance Focused Approach to Complex System Design  33

3.1 Motivation ................................................. 33
3.2 Background ................................................. 34
  3.2.1 Maintenance ......................................... 34
3.2.2 System Design and Optimization ....................... 35
3.2.3 Design-Maintenance Integration ....................... 36
3.3 Framework ................................................. 37
  3.3.1 Problem Formulation ................................. 37
3.3.2 Objective Function Definitions ...................... 39
3.3.3 Computational Complexity ......................... 40
3.3.4 Framework Summary .................................. 41
3.4 Power Plant Condenser Case Study ...................... 42
  3.4.1 System Modeling ..................................... 43
3.4.2 Traditional Design Method (LMTD Method) .......... 49
3.4.3 Optimal Design Method ................................ 51
3.5 Results ................................................... 51
  3.5.1 Single Objective Optimization ..................... 51
3.5.2 Parameter Sensitivity ................................ 54
3.5.3 Multiple Objective Optimization ................... 56
3.6 Conclusions & Future Work ............................... 57

4 Experimental design task: Development of reverse osmosis plant model 61

4.1 Motivation ................................................ 61
4.2 Background in Reverse Osmosis Modeling ............... 61
4.3 Reverse Osmosis System Model Development ............ 63
  4.3.1 Reverse Osmosis Unit Model ....................... 64
4.3.2 Flow Structure (FS) Model .......................... 72
4.3.3 Intake & Pre-treatment (IP) Model .................. 76
4.3.4 Lifecycle Model ....................................... 78
4.3.5 System Integration ........................................... 78
4.4 Model Verification ................................................. 80
  4.4.1 RO membrane model ........................................ 81
  4.4.2 System level performance .................................... 83
4.5 Summary & Future Applications ................................. 85

5 Designer behaviors in parameter-based Reverse Osmosis design 87
  5.1 Motivation ....................................................... 87
    5.1.1 Related Work ............................................. 89
    5.1.2 Research Question ....................................... 90
  5.2 Experimental Methods ......................................... 91
    5.2.1 Design Problem and User Interface ...................... 92
    5.2.2 Coupling Effect ......................................... 93
    5.2.3 Design Tasks & Complexities ............................ 94
    5.2.4 Computer Graphic User Interface ...................... 95
    5.2.5 Experimental Procedure ................................ 100
  5.3 Results & Analysis ............................................ 101
    5.3.1 Test Subject Demographics .................. 101
    5.3.2 Qualitative Observations ............................ 102
    5.3.3 Quantitative Analysis ............................... 105
    5.3.4 Detailed Analysis of Strategies .................. 115
    5.3.5 Similarities to Simulated Annealing .............. 127
  5.4 Conclusions & Future Work .................................. 134

6 Conclusions ......................................................... 137
  6.1 Summary ....................................................... 137
  6.2 Suggestions for Future Work ................................ 139

A Designer’s Behaviors Experimental Questionnaire 143

B Additional Figures 147
### List of Figures

1-1 World desalination capacity and large-scale plants .......................... 16
1-2 Black-box representation of a desalination plant ............................. 18

2-1 Municipal desalination plant design process and stakeholder relationships 26

3-1 Proposed framework for maintenance integrated at design stage ........... 38
3-2 Model Schematics of Rankine Cycle ............................................. 43
3-3 Model Schematic of Condenser ..................................................... 45
3-4 Model Evaluation Process ............................................................ 48
3-5 Lifecycle Cost Breakdown, Traditional Design vs Optimal Design ... 53
3-6 Lifecycle cost breakdown: 2× Condenser Cost .................................. 55
3-7 Lifecycle cost breakdown: $r = 0.05$ ............................................. 56

4-1 Example of reverse osmosis flexible superstructure ........................... 63
4-2 System Diagram of Reverse Osmosis Plant ........................................ 64
4-3 Schematic of reverse osmosis unit .................................................. 65
4-4 Illustration of transport phenomenon at RO membrane ....................... 66
4-5 Reverse osmosis flexible flow structure .......................................... 73
4-6 N² Diagram of reverse osmosis system model .................................... 78
4-7 RO Model verification, varying feed flow rate and pressure ............... 81
4-8 RO Model verification, varying membrane type and pressure vessels .... 82
4-9 RO Model verification, salt concentration and recovery ratio .............. 82

5-1 User interface of ROSA, a commercial reverse osmosis design software 88
5-2 Factors affecting experimental outcome .......................................... 91
5-3 2-Pass SWRO process used in design problem .......................... 92
5-4 SWRO design problem parameter-based DSM .......................... 93
5-5 Graphic user interface for the test problem ............................ 98
5-6 Box plot of problem times overlaid with exponential relationship ... 107
5-7 Box plot of problem Iterations overlaid with exponential relationship 108
5-8 Order effect, problem 10x5 .............................................. 109
5-9 Order effects in time measurements ...................................... 110
5-10 Order effects in iterations measurements ............................... 112
5-11 Distribution of subject knowledge, fast vs slow group ............... 114
5-12 Design space, target region, and problem solving process .......... 115
5-13 Sample plots of distance-to-target measurements for the 3x3 problem 117
5-14 The phenomenon of subjects consistently go away from target .... 118
5-15 Number of critical input parameters vs initial jump, problem 10x5 122
5-16 Iterations vs number of critical input parameters ........................ 123
5-17 Sample plot of step-size measurements during experiment .......... 123
5-18 Comparison of performances when changing multiple parameters ... 127
5-19 Comparison to simulated annealing: subject’s temperature profile .. 129
5-20 Comparison to simulated annealing: subject’s search neighborhood . 131
5-21 Simulated annealing iterations vs average step-size .................... 132
5-22 Iterations vs average step-size, comparing simulated annealing to test subjects ................................................................. 133

6-1 Improvement to the desalination system design process ............... 139

B-1 Iterations vs tendency to go away from target region .................. 148
B-2 Iterations vs initial jump .................................................. 149
B-3 Iterations vs step-size kurtosis .......................................... 150
B-4 Iterations vs average step-size .......................................... 151
B-5 Iterations vs tendency to change multiple parameters each iteration . 152
List of Tables

1.1 Techniques in human centered design process .................................. 17
3.1 List of user-defined parameters ......................................................... 49
3.2 Results of traditional design vs optimal design ................................. 52
3.3 Results of Optimal Condenser Design ............................................... 53
3.4 Optimal Condenser Design: 2× Condenser Cost .................................. 55
3.5 Optimal Condenser Design: Discount Rate $r = 0.05$ .......................... 55
4.1 Library of reverse osmosis membranes ............................................. 69
4.2 Constraints on reverse osmosis membrane elements ............................ 70
4.3 Number of variables in subsystems .................................................. 79
4.4 Summary of input & output variables .............................................. 80
4.5 Design parameters of existing reverse osmosis plant case example ........ 84
4.6 Model prediction vs plant proposal values ....................................... 84
5.1 List of test problems ................................................................. 95
5.2 Constants & target values for each test problem .............................. 96
5.3 Upper and lower limits of GUI inputs ............................................ 99
5.4 Technical experiences of participants ............................................. 102
5.5 Problem completion rate ........................................................... 106
5.6 Time taken to complete each problem (minutes) ............................. 106
5.7 Iterations taken to complete each problem ..................................... 108
5.8 Order effects in time measurements ............................................. 111
5.9 Order effects in iterations measurements ..................................... 111
5.10 Performance ranking based on technical experiences of subjects . . . 113
5.11 Going away from target ............................................. 120
5.12 Correlation coefficients between initial jump and iterations ........ 121
5.13 Correlation coefficients between step-size kurtosis and iterations . . 124
5.14 Correlation coefficients between average step-size and iterations . . 125
5.15 Correlation coefficients between tendency to change multiple parame-
ters per iteration and iterations ........................................ 126
5.16 Similarities between simulated annealing and human subjects ...... 128
5.17 Correlation between likelihood of accepting worse design and performance 130
5.18 Correlation between step-size at each progress interval and performance 131
Chapter 1

Introduction

1.1 Motivation

Water, one of the fundamental building blocks of humanity, is at a serious risk of running out in the civilized world. Even though water is the most abundant compound on the Earth’s surface, over 99% of it is salt water, and the remaining 1% that is fresh water is extremely unevenly distributed across continents. One and a quarter billion people already live in regions with physical water scarcity, with another half billion approaching this situation [1]. In addition, population growth and industrial development will only make the current water shortage more difficult to address in the future.

Desalination converts salt water into fresh water, and is widely believed to be a promising option to alleviate the water scarcity problem, despite the high cost, stringent maintenance requirements and massive energy consumption associated with desalination technologies. In the past few decades, the world’s total desalination capacity has been growing at a very fast rate, as are the number of large-scale desalination plants being constructed around the world (Figure 1-1), with many more high capacity systems being planned in the near future.

Large-scale desalination plants are complex systems that typically take several years to plan, design and construct, and have long life times of around 30 years. A number of large-scale plants have faced substantial design issues in recent years, re-
Figure 1-1: World desalination capacity and large-scale plants (capacity greater than 100,000 m³/day) [2, 3, 4].

resulting in large delays of project delivery, cost overruns, and degraded performances. Some of these include the Tampa Bay desalination plant [5, 6], Addur desalination plant in Bahrain [7], and the Victoria desalination plant in Australia [8]. The combination of water scarcity, increase in the number of large-scale plants, and the number of problematic plants motivated the following question: is the current design process effective enough to handle the increasing complexity of desalination systems?

1.1.1 Human Centered Design Approach

One branch of design research considers how to best support designers in their work, and one strategy to conduct this research is to start with understanding designer behavior and practice. Current design practice in the desalination industry may not address all needs of different stakeholders, and there may be "pain-points" in how designers accomplish their work. This dissertation takes a human centered approach that focuses on both the designer and the design process to find the areas of potential improvement in current design practices in the desalination industry, and suggest appropriate strategies to address them.

Successful design approaches need to focus on ensuring that a real need is ad-
dressed in the design of an artifact/service. "Human-centered design" were first developed as a process in which the needs of end-users drive how a design is conceived [9, 10]. Table 1.1 lists a few commonly used techniques in human-centered design. The involvement of human users assures that the outcome of the design will be satisfying for them.

Table 1.1: Techniques in human centered design process [10, 11]

<table>
<thead>
<tr>
<th>Technique</th>
<th>Method</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interviews</td>
<td>Talk to a user using a set of prepared questions</td>
<td>Collect detailed data related to needs of users</td>
</tr>
<tr>
<td>Focus groups</td>
<td>Have a discussion with a group of users</td>
<td>Obtain multiple points of view from different stakeholders</td>
</tr>
<tr>
<td>Field study</td>
<td>Observe users in their real environment</td>
<td>Collect data to design based on reality, not assumptions</td>
</tr>
</tbody>
</table>

1.1.2 Maintenance Strategies

One critical characteristic of complex engineering systems is the inevitable degradation and failure that occur to functional components over time, such as the aging of batteries and wearing out of bearings. Proper maintenance of degrading components over time is critical to the efficiency and longevity of engineering systems. For many industrial systems such as refineries, power plants, and desalination plants, maintenance cost can range between 10% to 30% of the total operational budget, and is the second highest trailing only the energy cost [12, 13, 14, 15].

The uncertain nature of degradation and failure means that traditional maintenance strategies can either lead to expensive and inconvenient down time or a waste of resources [16]. The emergence of prognostics and health management (PHM), which is the diagnosis of health conditions based on sensory signals and modeling/prediction of remaining useful life, has led to condition based maintenance (CBM) approaches, which tries to minimize both system downtime and maintenance cost simultaneously.

Despite the importance of maintenance, there has not been much focus on finding
the optimal maintenance strategy during the design stage. This was a view shared in interviews of practitioners in the desalination industry. This dissertation takes a systems level design approach to the issue of designing the optimal maintenance strategy from the early design stage.

1.2 Background on Desalination Technology and Maintenance

This section is a brief introduction to the industrial desalination process that will help contextualize the system-level view of this thesis. Figure 1-2 is a simplified representation of a desalination system. Salt water is introduced to the system through the feed stream, which is then split into the permeate stream that contains fresh water, and the brine stream, which contains more concentrated salt water. Of course, energy is required by the system to perform this process. The salinity – usually referred to as total dissolved solids (TDS) – of the feed stream can range from 5,000 parts-per-million (ppm) in industrial waste water desalination to 45,000 ppm in seawater desalination plants in the Middle East, to 200,000 ppm or more in specialty applications, such as the oil and gas industry [17].

Figure 1-2: Black-box representation of a desalination plant
Existing desalination technologies generally fall into two categories: thermal desalination and membrane desalination [18]. Common thermal desalination processes, such as multi-stage flash (MSF) and multi-effect distillation (MED), use energy to evaporate water and subsequently condense it again to produce fresh water. Thermal desalination plants require a significant amount of thermal energy for evaporation, and also a considerable amount of electric energy for pumping [14]. Since the cost of thermal processes is fairly insensitive to feed water conditions, thermal desalination are still the most common desalination processes in the Middle East [19], where seawater salinity is very high while fuel availability is abundant.

Membrane based desalination technologies are based on the filtration of seawater to remove salt. The most common type of membrane process is reverse osmosis [18, 20], often abbreviated as RO. In an RO process, pressurized seawater flows across a semipermeable membrane, and the pressure forces water molecules to diffuse through the membrane, producing fresh water. Recent advancements in membrane technology have drastically reduced the cost and energy consumption of RO systems [21], and the desalination industry has witnessed a boom in the numbers of new RO plants constructed around the world [20]. RO also has its drawbacks: its energy consumption is heavily affected by the feed water salinity, pH value, and temperature; silt, algae and bacteria in the seawater can accumulate on the membrane, also known as membrane fouling, and the pre-treatment process that removes them requires substantial capital cost and energy consumption [20].

Due to the large energy consumption of desalination systems, it is potentially beneficial to construct them alongside power plants, known as power-water cogeneration plants. Thermal desalination plants can then use the waste heat from the power plant in the seawater evaporation process. RO plants can be used for load-balancing of the power plant [22], while the power plant can be used to regulate RO feed water temperature [23].

Operational cost is a significant expense for desalination plant owners. The majority of the operational cost is due to the phenomenon of fouling. Fouling can occur on RO membranes, in heat exchangers of thermal plants, and also condensers in power
plants for cogeneration systems. Fouling reduces the efficiency of heat/mass transport, leads to high energy consumption, and requires the plant to be shut down and the fouled surfaces to be cleaned regularly. Optimal planning of maintenance cycles in desalination plants is a major factor to lowering plants’ lifecycle costs.

1.3 Thesis Contribution

In this dissertation, a novel approach that combines the needs of designers is used for desalination design research, with the aim of improving both the design of desalination systems and also the process of designing. Contributions would be made in the following three areas.

1.3.1 Human-centered approach to design process research

This thesis takes a human-centered approach which is common in other areas of engineering, from product design, to automotive, and software engineering. In this thesis the designers are the focus of the study, and human-centered approach gathers needs that are used to re-design the design process for the designers. A key strategy in this approach is to observe and/or interview stakeholders. In the first part of this thesis, four industry experts in the design of desalination systems were interviewed.

1.3.2 Framework for maintenance in design stage

A common theme from the interviews and other discussions is that the initial design is important, but long term performance is critical because desalination systems have a lifetime of about 20 to 30 years. To that end, a new design framework is proposed that combines the modeling of degradation of components with maintenance of the system over the life of the plant and design decisions in the design stage. This novel design framework makes the overall design process more efficient and can result in designs that are more optimal over its lifecycle.
1.3.3 Characterization of desalination designer behaviors

Keeping to the user-centered approach, the behavior of designers was observed during a design experiment. The aim was to investigate more deeply how novice as well as more experienced designers of desalination systems respond to the challenges of parameter-based desalination design problems, and explore the efficiency and strategies of human designers at solving problems commonly found in commercial desalination design software. This portion of the thesis involves a controlled experiment using a detailed RO model and custom developed interactive interface, and observes how increases in parameter numbers can change designer’s ability to find design solutions.

1.4 Thesis Organization

There are six chapters in this dissertation. Chapter 2 presents the interview process and results of discussions with desalination practitioners. Chapter 3 explains the design framework that incorporates maintenance strategies in the design stage, with a case example of fouling in heat exchangers. Both chapters 4 and 5 are dedicated to the designer behavior experiments, with chapter 4 describing the development of a detailed RO model to simulate the design environment, and chapter 5 focusing on the experiment and analysis of the results. Chapter 6 concludes this dissertation and suggests areas of future investigations.
Chapter 2

Understanding Challenges of Designing Desalination Systems

2.1 Motivation

This chapter presents results from interviews conducted with desalination industry practitioners from several different organizations. The interviews focus on understanding the real world planning and decision making process used in desalination projects. The intent is to both understand the industry practice, and find areas of the design process that could be improved.

2.2 Methodology

Four practitioners with experiences in designing desalination systems were interviewed for this study. The candidates possessed a variety of industry backgrounds, including design engineers from the desalination and waste water treatment department of large corporations and also start-up companies, and consultants with over 20 years of experiences in power and water industry. Collectively, the interviewees had experiences in reverse osmosis, electrodialysis, hybrid thermal desalination technologies as well as renewable energy powered desalination.

Three of the candidates were located within driving distances, and in-person in-
terviews were arranged, while phone interviews were requested for one candidate from another parts of the country.

Each interview consisted of an hour of discussion on the desalination design process. The key questions are listed below:

1. From your own experience, can you give us a brief description of the key steps in designing a desalination plant?

2. How much is innovation valued in the design of plants?

3. Do you think the current design process results in optimal plant designs? How much more improvements can be had if more numerical optimization methods are used?

4. What do you think is the most critical area that need improvement in the design process?

At the beginning of the interview an informed consent form was provided to the interviewees. Notes were taken during the interviews and analyzed afterward to summarize the design process being described, as well as any consistent topics emerged from the interviews. Select quotes and themes from the interviews are presented in the results section. These concepts were incorporated into and informed the studies described in later chapters.

2.3 Results

2.3.1 Municipal Water/Energy System

All interviewees had experience working on municipality water treatment projects, and when they talked about their design process, this experience came through as having a heavy focus on the relationship between the client (government) and the developer (the design firm). The design process mentioned by the interviewees is summarized below, and it roughly follows the conceptual-preliminary-detailed design phases commonly found in the design of complex engineering systems.
Desalination plants that are built to provide fresh water to a local community usually follow the process described below. Such plants are usually initiated by the client to serve a water shortage problem. The client would hire a team of consultants to help them conduct a conceptual design. The first step of conceptual design is to evaluate water demand. Since large scale desalination plants have large energy requirements, the electricity availability of the region may also need to be evaluated to decide whether the desalination plant can be constructed as a stand-alone plant, or be part of a water-electricity cogeneration plant. It is also very important to understand how much variation there will be in water demand over time to determine how many individual desalination trains are needed for production flexibility. Based on the demand information, the size of the plant can be determined, a location for the plant will be selected, and a technology would be chosen for the conceptual design of the plant.

A feasibility study would then be conducted for this location to analyze the feed water quality, and determine realistic targets for capital cost, operating cost, and efficiency. The client would then start a bidding process by sending out a request for proposals (RFP). Competing desalination design firms, which are in charge of the systems engineering aspects of the project such as project finance, equipment procurement and scheduling, submit proposals to the client, providing their estimate for the cost of the project.

Once a developer is selected based on their proposal, the developer must create a preliminary design of the plant for the client for review and negotiation. The majority of the design requirements will be frozen at this stage of the design process, such as type of technology (membrane or thermal), requirements on the quantity and quality of water, capital cost, operational cost, efficiency, and any additional detailed requirements. The design requirements are very rarely changed after this point.

After all design requirements are frozen, the developer will perform the detailed engineering design of the plant, and may hire contractors to help with specific portions of the design process. In the case of a reverse osmosis plant, the detailed design process would include determining the exact flow rate, number of RO passes and
stages, recovery ratio, number of membrane elements, specifications of pumps, as well as designing the architectural layout and selecting the electrical and instrumentation equipments. Design evaluation software is used at this stage. The desalination process design is completed first, while the architectural, electrical and instrumental design are done after the process design is completed. The detailed design would again need to be reviewed and accepted by the client, before the procurement of equipment, and finally the construction of the physical plant, which very often is done by a different company than the designer. The entire cycle from conceptual design to the completion of construction may take a few years. Figure 2-1 is a graphic representation of the stakeholder relationships during the design of municipal desalination plants, and also show the break-down of the three design phases.

Figure 2-1: Municipal desalination plant design process and stakeholder relationships

After the completion of the plant, the developer must demonstrate that the plant meets all requirements such as operational cost, production capacity and quality. Depending on the contractual agreement, the ownership of the plant may be different.
Historically, the most common contract structure is the engineering, procurement, construction (EPC) structure, under which the plant is owned by the client upon completion, and the developer must provide at least one year of guarantee on the operation of the plant. Recently, the build-own-operate-transfer (BOOT) structure has been used in many municipal projects to reduce the risk faced by municipalities during plant operation. Under the BOOT structure the developer owns and operates the plant for a number of years before transferring the plant’s ownership to the client.

2.3.2 Industrial Water Treatment Systems

Some of the interviewees also talked about the design of industrial water treatment systems utilizing reverse osmosis. Compared to designing for municipalities, where there is a set of critical water quality requirements that are non-negotiable, the major design objective for industrial water treatment is to reduce the cost of ownership: the capital and operational cost of the water treatment must be as low as possible, while the water quality requirement is less important. The permeate total dissolved solid (TDS) limit of an industrial system could be as high as 1000 ppm, compared to 100 - 500 ppm for municipal systems.

The sale cycle of industrial water treatment systems are also much shorter, on the order months compared to years for municipal projects. The reason for this are multi-fold. Industrial systems are typically smaller, and there are fewer stakeholders involved in the project, and usually the engineering company provides the design construction and support for the client. In comparison, municipality projects often involve the government, developer company, several design consultants, construction companies, financial institutions, power companies, environmental agencies, and community members.

2.3.3 Use of Numerical Optimization

Numerical optimization can be a powerful tool during the detailed design stage of complex systems. Numerical optimization can quickly evaluate a large number of
design alternatives and help designers find more optimal designs. When asked about whether numerical optimization tools were used in their design process, all interviewees responded with the answer "No, we don't." The design process was basically tweaking the design parameters in a 'trial-and-error" or "empirical approach", largely based on past experiences, to find a design that meets all requirements. Design evaluation software is used during the design process to simulate the performance of designs but does not perform optimization. ROSA (Reverse Osmosis System Analysis) design software, provided for free by Dow Water & Process Solutions [24], was mentioned by more than one of the interviewees as the software they use in their company.

One interviewee felt that the current design process is already producing "pretty optimal mechanical processes", as long as feed water properties are fully understood, and the most experienced designers are in charge of key decision making. Another interviewee made similar comments, stating that experienced designers are very "cost conscious" while making design decisions, and could only expect numerical optimization to make "incremental improvements" but not revolutionary changes to the industry.

It was also mentioned that all published records of desalination system design optimization are in the academia, and that there is no industrial level numerical optimization tool available to the public. One interviewee felt that numerical optimization might be more relevant to municipal seawater desalination, since the composition of seawater around the world is largely the same, making it easier to setup a numerical analysis, but for brackish water and industrial waste water, the feed water composition can vary drastically from project to project, making it difficult to come up with a generalized optimization tool.

### 2.3.4 Design for Lifecycle Optimality

Another trend in the desalination industry that arose in the interviews was the emphasis on lifecycle management. For industrial water treatment systems, the cost or saving of the system over its lifecycle is the biggest deciding factor in whether a design is chosen. Municipality projects had less emphasis on lifecycle cost in the past,
since under the EPC contract structure there is less incentive for the developer to
design for optimal long term operation. Under the BOOT structure that’s becoming
widely used today, the operational and maintenance risks are allocated to the project
developer company [25, 26], and thus resulting in higher incentives for developers to
optimize the plant performance over its lifecycle.

One major factor affecting lifecycle performance is fouling and scaling, which
is the accumulation of foreign material on transport surfaces. Fouling affects both
reverse osmosis desalination plants in the form of membrane bio-fouling, and also
thermal distillation plants, where scale formations and bio-fouling can occur in heat
exchangers and condensers. Fouling results in higher pumping pressure for reverse
osmosis systems and lower heat transfer rate for thermal systems, resulting in de-
creased plant efficiency over time. Periodic back-wash and chemical treatments can
slow down the fouling and scaling process, but eventually the plant needs to be shut
down for in-depth cleaning of fouled surfaces. The cost of fouling, both energy cost
and maintenance cost, contribute to a large portion of the lifecycle cost, and is an area
that could be improved in the future. Optimal maintenance strategies that maximize
plants’ lifecycle performance was identified by an interviewee as a critical area for
future improvement.

2.3.5 Other Themes

Creativity and Innovation in Water systems

In general, the interviewees expressed that innovation is not a major contributor to the
design of desalination systems, especially in municipal desalination systems. Because
of the large number of stakeholders involved in the projects, local governments tend
to be more conservative and unwilling to accept new technologies that may have
higher risks. Therefore engineering companies usually start with a seed design that
has been successfully implemented in the past, and modify the design to suit the
design requirements provided by their client.
Water-Energy Nexus

The importance of energy in desalination was commonly mentioned throughout the interviews. All current desalination plants, especially seawater desalination, are major energy consumers, and the varying water demand could put significant burden on the electricity grid, and make it challenging to use renewable energy sources such as solar power. Adding water/energy storage could alleviate some of the energy problem, but adds additional levels of complexity to the plant designers. Another common approach, especially in the Middle East, due to the abundance of fuel in the region, is to build water-energy cogeneration plants, so that the desalination plant can share intake structures with the power plant condenser, waste heat from the power plant can be used by thermal distillation processes, and reduce electricity demand directly from the grid. A few of the interviewees suggested that in-depth analysis of how different energy production and water production technologies could work together in a hybrid energy-water system should be the focus of future desalination research.

2.4 Summary

This chapter summarized the interviews that were conducted with practitioners who are experienced with both municipal desalination plants and industrial wastewater treatment systems. The design process used by both municipal projects and industrial projects were solicited and outlined.

A number of themes were identified from the interviews. Two of the themes that stood out the most were the importance of lifecycle costs and the "trial-and-error" approach used by designers. Recently, there has been interest in research areas focusing on predictive maintenance, which utilizes artificial intelligence methods to predict future degradation in systems in order to schedule maintenance activities at the most optimal time. With increasing interest in reducing lifecycle cost in desalination plants, one part of the design process that could be improved is to consider the effect of maintenance during the design stage, investigating how optimal maintenance strategies may affect design choices in a desalination plant, and vice-versa.
It was interesting to find that desalination processes, especially reverse osmosis processes, are designed using a parameter-design approach, where a number of input parameters are controlled to make changes in some output parameters. Parameter-design problems are common in many different engineering design fields and have been studied in the past. Interestingly, past research has indicated that human designers perform poorly at solving parameter-design problems [27, 28], and struggle even more as the number of variables increase even modestly. Understanding how desalination designers solve desalination based parameter design problems, identifying what design strategies are more efficient and what are not, could significantly improve the design of desalination systems, as well as any parameter-design problems in general.

The next three chapters will present research based on these two themes identified from the interviews. Chapter 3 will discuss a novel design framework that considers the effect of maintenance on design decisions. Chapters 4 and 5 present a controlled experiment that simulate the design process of a reverse osmosis plant, and observe the designers’ behaviors during the design process.
Chapter 3

A Maintenance Focused Approach to Complex System Design

3.1 Motivation

One of the major themes identified in Chapter 2 was the importance of optimizing the lifecycle cost in desalination plants, which is also critical to other types of large-scale engineering systems. Due to the long lifecycle of engineering systems, the evaluation of cost and performance must go beyond the initial design and construction stage, to include the long-term variations in system condition [29, 30].

Degradation of system components is one major factor that can contribute to system lifecycle performance variations. Component fatigue, fouling, corrosion and fractures can all lead to performance losses or failure of required system functions. Maintenance is a critical aspect of system operation: degrading components need to be cleaned or repaired, and failed components need to be replaced to restore performance. However, the uncertain nature of degradation and failures can make maintenance scheduling a challenge [31]. Recent advances in maintenance strategy focus on prognostics and health management (PHM) methods [32, 33, 34] to detect, diagnose, and predict system degradation and failures, and condition based maintenance (CBM) [35, 36, 37] that utilizes PHM information to maximize system availability and minimize operational costs [38].
Traditionally, maintenance optimization are not considered during the design stage of complex engineering systems, and the consideration of component degradation is usually limited to safety-factors and redundancies. However, maintenance strategies developed independently from the design stage may result in unexpected lifecycle outcomes. In many industrial components, physical design parameters and the mode of degradation are highly coupled. Design choices made on the system can influence degradation, which can affect the maintenance pattern, and deviate the system performance from designed specifications. Current research in relevant areas has focused on integrating reliability analysis and probability of system failure in the design stage [39, 40], but only a few had attempted to fully address the interactions between physical degradation process, design decisions, maintenance strategy, and system performance [41, 38].

This work aims to exploit the potential of capturing the complex relationships between design, degradation, and maintenance, and determines the optimal design and maintenance strategies. A framework is proposed that focuses on capturing design-maintenance interactions and system modeling of uncertainties, to provide a deeper understanding of the trade-offs between design and maintenance decisions.

3.2 Background

This section first provides a brief overview of relevant fields in maintenance and design optimization, and then a literature survey to illustrate the existing gap in the literature.

3.2.1 Maintenance

Maintenance usually involves cleaning, repair, or replacement of degraded and aging components. The oldest and most common maintenance strategy is corrective maintenance (CM), or "fix it when it breaks" [13]. Corrective maintenance does not require any system analysis or planning effort, but at the same time the operator has no control over the occurrence of downtime.
Preventive maintenance (PM) is also a commonly used strategy in industry [42], where maintenance and repairs are performed at pre-established intervals. This strategy can effectively prevent unexpected downtimes, but since the fixed maintenance interval ignores the stochastic nature of degradation, very often maintenance is performed unnecessarily, leading to high operational cost.

Condition based maintenance (CBM) is an advanced maintenance strategy aimed at balancing maintenance cost, which is high in PM, with downtime cost, which is high in CM [42, 12]. In CBM, system components are continuously monitored by various sensors to measure the state of the component, and computer algorithms estimate the "health" of the component, and predict the remaining useful life (RUL) using either physics-based models or data-driven methods. Prognostics and health monitoring (PHM) is the research area that focuses on modeling RUL distribution and reliability in a number of different areas including electronic components [43], civil infrastructure systems [44, 45], airplane maintenance [41], and battery technologies [46]. Information provided by PHM technologies are used in CBM to schedule maintenance accordingly.

3.2.2 System Design and Optimization

Complexity in engineering systems make design optimization challenging. Interactions between different subsystems mean that the aggregation of optimal subsystems do not necessarily guarantee system level optimality, and system level modeling is necessary for design optimization. The multidisciplinary nature of systems lead to multiple competing objectives during design evaluation. Multidisciplinary design optimization (MDO) is a broad research area that aims to involve several disciplinary models into a single system-level optimization loop [47]. The general formulation of MDO is as follows [48, 49, 50]

\[
\begin{align*}
\text{minimize} & \quad F(x_{cs}, y) \\
\text{subject to} & \quad y_i = Y_i(x_{cs}, x_i, y_{cj}), \quad i, j = 1, 2, \ldots, s \\
& \quad g_k(x, y) \leq 0
\end{align*}
\] (3.1)
where $F$ is some objective such as cost or performance, $\mathbf{y}$ is a vector output from the corresponding subsystems and disciplines; $\mathbf{y}_i$ is a vector output of subsystem $i$ modeled by $Y_i$; $\mathbf{y}_{cj}$ is a vector output from other disciplines $j$; $\mathbf{x}$ is a vector design variables including the system design variables $\mathbf{x}_s$ and subcomponent design variables $\mathbf{x}_i$; $s$ is the number of subcomponents and disciplines, and $g_k$ is a set of constraints.

3.2.3 Design-Maintenance Integration

Despite substantial research in design optimization and maintenance strategy optimization, there has been very little work that focuses on the integrated optimization of design with maintenance. Bodden et al. conducted a study that considers prognostics and health management as a design variable in air vehicle conceptual design. In this work, the redundancy in air vehicles could be reduced with some knowledge of RUL [41]. Wang et al. presented an optimal design approach accounting for reliability, maintenance and also warranty [35]. Youn et al. proposed a framework for resilience-driven design of complex systems which integrates PHM into the design process using a reliability-based design optimization strategy [38]. Kurtoglu and Tumer developed a fault identification and propagation framework for evaluating failure in the system in the early design stage [39].

Related research can also be found in disciplines outside mechanical engineering: Camci explored maintenance scheduling with prognostic information which considered the probabilistic nature of prognostics information and its effect on maintenance scheduling [42]. Santander and Sanchez-Silva studied design and maintenance optimization for large infrastructure systems. By applying reliability-based optimization using a deterministic system model, they found that inefficient maintenance policy leads to the optimization algorithm to converge to a design with higher degrees of redundancy [51]. Monga and Zuo considered both maintenance and warranty in optimal system design in their work. They compared selected system configurations with different failure rate functions, though no predictive maintenance was considered in this study [52]. The above-mentioned work mostly focuses on the integration of design and system reliability, but does not explicitly consider the physical degradation
process, nor the causal relationship between design decisions and degradation.

The work done by Caputo et al. on joint economic optimization of heat exchanger design and maintenance policy considers the interaction between design decisions of a heat exchanger and its degradation (fouling), but this study only examines the traditional maintenance strategy, and does not consider uncertainty associated with degradation [53]. Honda and Antonsson proposed the notion of grayscale reliability to capture system performance degradation and the time dependency of reliability. They also studied design choices and their effects on system degradation. However, their study does not consider the effects of maintenance on system degradation [54].

There are no comprehensive studies in previous literature that link degradation to design variables, considers both advanced maintenance strategies, and uncertainties in degradation. The approach proposed in this paper seeks to fill this gap by addressing all three issues.

3.3 Framework

3.3.1 Problem Formulation

The proposed framework for integrating design and maintenance is based on an MDO approach, and the overall problem formulation is shown below:

\[
\begin{align*}
\text{minimize } & \{C_L, -A\} = F(\tau_m, E_s, x_{cs}) \\
\text{subject to } & \tau_m = f_o(D(t), x_m) \\
& E_s = f_{cs}(D(t), x_{cs}, x_m) \\
& D(t) = f_D(x_{cs}) \\
& g_{cs}(x_{cs}, x_m) \leq 0
\end{align*}
\] (3.2)

where \(C_L\) is the lifecycle cost of the system, and \(A\) is the availability or other metrics related to availability (such as mean time between maintenance). These are the objectives that need to be minimized and maximized in the design optimization, and are influenced by the variables of \(\tau_m\), \(E_s\), and \(x_{cs}\). \(\tau_m\) is a sequence of times indicating
maintenance occurrences during system lifecycle, and is computed by the operational subsystem model $f_o$, based on component degradation profiles $D(t)$ and maintenance variables $x_m$. $E_s$ is the system lifecycle performance, typically an efficiency or output measure, that is calculated by the physical system models $f_{cs}$ based on degradation profiles, system design variables $x_{cs}$, and maintenance variables. The component degradation profiles $D(t)$ are generated by degradation subsystem models $f_D$ based on system design variables. $g_{cs}$ are the system level constraints that must be satisfied.

A diagram of the problem formulation is shown in Figure 3-1. There are two parallel divisions in the system model. The first division is the system design division, that contains the forms and functions of the physical system and its subcomponents and disciplines. The system model could be highly complicated with many interactions, while in this framework, the system model is generalized into $f_{cs}$. The second division contains the degradation models $f_D$ and the operational model $f_o$, and also any other subsystems related to operation. It was assumed the physical system can be modeled deterministically, while the degradation models are uncertain, and requires uncertainty evaluation methods such as Monte-Carlo simulation.

The degradation models generate profiles that indicate how each component degrades overtime. These degradation-vs-time curves are used by the system model
to simulate how the system performance will vary over this time period, the performance-vs-time profiles $P_s$ come into the operational model that computes the maintenance pattern $\tau_m$. The operational model is simulated for $N_k$ years, the total duration of the system lifecycle, with different degradation profiles generated over the lifecycle to simulate the uncertainty in degradation.

### 3.3.2 Objective Function Definitions

This framework focuses on two design objectives, the lifecycle cost and system availability. Other objectives may be considered, but cost and availability are representative for comparing optimality of system designs and maintenance strategies. The lifecycle cost $C_L$ is defined as the total present cost of the system, and contains:

$$C_L = C_C + C_M + C_E$$

where $C_C$ is the capital cost associated with system development. If any portion of the capital cost occurs later in the system lifecycle, then the cost needs to be discounted to the present value with the discount rate $r$. $C_M$ is the maintenance cost, which is the sum of all maintenance costs:

$$C_M = \sum_{i=1}^{N_k} \frac{C_{M,i}}{(1+r)^i}$$

and $C_E$ is the efficiency cost, which is the cost due to degradation, such as the extra fuel needed:

$$C_E = \sum_{i=1}^{N_k} \frac{C_{E,i}}{(1+r)^i}$$

$$C_{E,i} = \int_{t_{i-1}}^{t_i} \mathcal{L}(\eta(\tau), \eta_d) \, d\tau$$

where $\eta$ is the actual system efficiency (between time $t_{i-1}$ and $t_i$), $\eta_d$ is the designed system efficiency without degradation, and $\mathcal{L}$ is a function that computes the mone-
tary loss due to degradation, dependent on the physical system.

The second objective is availability, which is generally defined in the literature as:

$$A = \frac{E(\text{Up Time})}{E(\text{Up Time}) + E(\text{Down Time})}$$  \hspace{1cm} (3.7)

where $E()$ is the expected value operator, Up Time is the times that the system is operational, and Down Time is the times the system spent idling, including due to failure and during maintenance. In certain cases, the time taken to perform maintenance may be negligible, then an alternative measurement, mean time between maintenance (MTBM), can be used instead.

In general, the lifecycle cost of the system should be minimized, while the system availability should be maximized. Multi-objective optimization (MOO) methods can be used to find the Pareto-optimal designs. Alternatively, one of the two objectives can be treated as an intermediate variable, and single-objective optimization methods can be used.

### 3.3.3 Computational Complexity

Since there is significant uncertainty associated with the lifecycle analysis, uncertainty quantification methods such as Monte-Carlo simulation is required to produce many different lifecycle simulations to fully characterize the effects of uncertainty. The need for Monte-Carlo simulation significantly increases the computational complexity of the framework.

Comparing the two parallel subdivisions of the framework, the degradation and maintenance subdivisions contains all the stochastic models, such as the component degradation models. The physical MDO model can be assumed to be completely deterministic, and can be treated as a black box in the Monte-Carlo simulation. The MDO model only need to be evaluated a few times to generate a surrogate model, such as a look-up table, and the Monte-Carlo lifecycle simulations can use the surrogate model instead of calling the full system model to reduce computing complexity.

Despite the reduction of computing complexities, the computing requirement is
still significant, and thus balancing between model fidelity and complexity is a major challenge. For a multi-disciplinary system, domain specific models are usually high fidelity with very low discrepancies with reality but require significant computational time (on the order of hours or days). Furthermore, high fidelity discipline-based models are usually represented using different software tools, making the data transfer between models complicated. Thus, high-fidelity models are not the best option for use in this framework. A common approach for model complexity reduction is to generate low-fidelity models from high fidelity models using metamodeling methods such as Kriging or response surface method [49]. Low fidelity models can then be evaluated very quickly, but can have very high discrepancy.

A mid-fidelity model is a simplified representation of a system which captures the essence of different domains by using first order approximations of physics based models [55]. Because they are physics-based, no special software is needed, and allows simple integration of different domains and subsystems. A mid-fidelity model has the advantages of short simulation time on the order of several seconds. Mid-fidelity models are commonly used in the early design phases to identify promising design strategies. Therefore, it is recommended that mid-fidelity models to be used whenever possible.

3.3.4 Framework Summary

Below are the proposed steps for setting up the integrated design and maintenance optimization problem:

1. Identify key components and their degradation modes that contribute to system performance loss and require regular maintenance services.

2. Determine the relationship between physical parameters and degradation.


4. Construct the system model. Setup domain models for the system, incorporate
the degradation relationship, simulate over the life time operation of the system and compute objective functions.

3.4 Power Plant Condenser Case Study

This section illustrates the proposed approach through a case example of power plant condenser design. The case study will compare different maintenance strategies in the design optimization of power plant condenser, with consideration of condenser degradation and maintenance. The results will demonstrate the interactions between maintenance strategies and design decisions. It is expected that an advanced maintenance strategy that is based on prognostics of future degradation will result in lower lifecycle cost, and potentially reduce design redundancy.

In a steam power plant, the condenser is needed at the exit of the low-pressure turbine to condense the exiting steam into liquid. Condensers are shell-and-tube heat exchangers. The steam flows through the shell side, which is usually kept at a very low pressure to achieve higher cycle efficiency. The heat ejected from the steam condensation is carried away by cooling water in the tube side of the condenser.

Fouling is the major degradation mode of a condenser. Fouling is the build-up of foreign materials inside the tubes due to bio-particles and inorganic salt in the cooling water. Fouling causes high thermal resistance in the condenser (commonly measured as fouling resistance with units of $[\text{m}^2\text{K}/\text{kW}]$), which increases the shell side pressure and ultimately reduces plant efficiency. It is recognized as one of the biggest problems associated with efficiency loss in power plants [56].

The build-up of fouling resistance in a condenser usually follows an asymptotic curve [57]. The asymptotic values of fouling and the rate of build-up are highly stochastic. Over the past fifty years much research has focused on finding the underlying physics that govern fouling. The results have suggested that the amount of fouling and the rate of fouling are proportional to temperature, and inversely proportional to the cooling water flow rate, assuming unchanging water quality and tube material [58, 59].

42
The maintenance of a condenser is performed offline, commonly during a scheduled power plant outage. Specialized scrapers are shot through the tubes with pressurized water to physically remove built-up foulant. Cleaning can usually be completed within 48 hours before returning the condenser to its original condition.

Following the proposed framework, a condenser design optimization case study was conducted. Three different maintenance strategies were considered, and the details of the strategies and the overall system model would be described in detail in the next subsection.

3.4.1 System Modeling

Power Model

For simplicity, a Rankine cycle with a single stage turbine is used. The power cycle only has five components: the boiler, the turbine, the condenser, the cooling water pump, and the feed water pump, as shown in Figure 3-2.

![Figure 3-2: Model Schematics of Rankine Cycle](image)

There are six nodes in the Rankine cycle numbered 1 to 6 as shown by the subscripts in Figure 3-2. $P_i, T_i, h_i, x_i$ are the pressure, temperature, enthalpy and steam
quality of the fluid at node \(i\) respectively. \(\dot{Q}_i\) and \(\dot{Q}_o\) stand for the heat [kW] input by the boiler and heat rejected by the condenser, \(\dot{W}_t\), \(\dot{W}_{fp}\), and \(\dot{W}_{cp}\) are the work [kW] extracted through the turbine, input to the feed water pump, and input to the coolant pump respectively. \(\dot{m}_s\) and \(\dot{m}_c\) are the mass flow rate [kg/s] of the steam/water circulated through the power cycle and cooling water through the condenser.

The power cycle model computes the cycle efficiency \(\eta\) as its output:

\[
\eta = \frac{\dot{W}_o}{\dot{Q}_i}
\]  \hspace{1cm} (3.8)

where \(\dot{W}_o\) is the net power output of the plant. Losses associated with piping friction, and inefficiency associated with pumps and turbines were neglected. The water/steam properties at each node are either defined by the designer (boiler specification: \(P_3\), \(T_3\), turbine back pressure specification: \(P_4\)), by the physical environment (\(T_5\)), or determined through a steam table [60].

**Condenser Model**

The condenser model computes the condenser heat duty based on the inlet conditions and geometry using the log-mean-temperature-difference (LMTD) method [61]. Where the condenser heat load \(\dot{Q}_c\) is equal to:

\[
\dot{Q}_c = U \cdot \Delta T_m \cdot S
\]  \hspace{1cm} (3.9)

where \(U\) is the condenser’s overall heat transfer coefficient with unit of \([\text{W}/(\text{m}^2\text{K})]\), \(\Delta T_m\) is the log-mean-temperature-difference (LMTD), and \(S\) is the overall surface area of the condenser in \([\text{m}^2]\).

It was assumed that the condenser was a one pass X-type shell and tube heat exchanger made of 70-30 copper-nickel alloy, with the terminal conditions listed according to Figure 3-3, where \(P_{atm}\) is the atmospheric pressure, \(\Delta P\) is the pressure drop in the coolant side of the condenser, and \(T_{sat}\) is the saturation temperature of steam at pressure \(P_4\).
An iterative process computes the condenser heat load $\dot{Q}_c$ and coolant pressure loss $\Delta P$ based on the condenser geometry (number of tubes $N_t$ and tube length $L$) and the condenser fouling resistance $R_f$, measured in $[m^2K/W]$. Details of the model can be found in reference [61].

Fouling Model

Fouling in the condenser is measured by fouling resistance $R_f$. Fouling is a complicated phenomenon that has been studied extensively. A general consensus is that the overall fouling trend can be represented by this asymptotic model: $[58, 56, 62, 63]$

$$R_f = R^*_f \left( 1 - e^{-\beta t} \right)$$

(3.10)

where $R^*_f$ is an asymptotic fouling resistance value related to the cooling water velocity $v_c$, tube inner diameter $D_i$, and the kind of foulant, while $\beta$ [year$^{-1}$] is the reciprocal of the time constant. Empirical correlations used for $R^*_f$ with calcium carbonate fouling and $\beta$ are [53]:

$$R^*_f = \frac{0.101}{v_c^{1.33}D_i^{0.23}}$$

(3.11)

$$\beta = 0.0016v_c^{-0.35}$$

(3.12)

Since fouling is a stochastic process, the asymptotic fouling resistance $R^*_f$ can change randomly over time $[56, 63]$. In this study, $R^*_f$ is treated as a random variable.
that takes on a new value each time the condenser is cleaned. The distribution of \( R_t^* \) is assumed to be log-normal: 
\[
R_t^* = LN(\mu, \sigma),
\]
with mean \( \mu \) defined by Equation 3.11, and \( \sigma / \mu = 0.3 \) [56].

**Maintenance Models**

Three types of maintenance strategies were considered in this study: 1) fixed interval, 2) corrective, 3) predictive.

Power plants are usually shut down annually during the spring when energy demand is the lowest. For this study, it is assumed the condenser is inspected and cleaning decisions are made during the annual downtime. A fixed interval strategy follows a preset schedule to maintain the condenser every \( N \) years. Both corrective and predictive strategies are based on monitoring the fouling resistance and comparing to a threshold value \( R_{fs} \). In corrective maintenance, cleaning happens if current \( R_t > R_{fs} \). In preventive maintenance, a remaining useful life (RUL) is estimated each year by predicting when \( R_t \) will exceed \( R_{fs} \), and schedule cleaning if \( RUL < 1 \) year. In this study, the prediction algorithm is assumed to be perfect.

**Operation & Economics Model**

A 50 year lifetime is assumed for the power plant. The lifecycle cost \( C_L \) consists of the capital cost \( C_C \), maintenance cost \( C_M \) and efficiency cost \( C_E \) as indicated previously in Equation 3.3.

For simplicity, the capital cost only included the condenser’s cost, which is proportional to the amount of materials used in the condenser tubes.

\[
C_C = P_m \rho_m n_l L \frac{\pi}{4} (D_o^2 - D_i^2)
\]

(3.13)

where \( P_m \) is the price of raw material in [\$/kg], and \( \rho_m \) is the density of material in [kg/m\(^3\)]. Capital cost associated with implementing maintenance strategies, such as sensors and computing equipments, are not considered in this study for simplicity.

The maintenance cost is associated with the physical cleaning of the condenser
tubes. The annual cleaning cost $C_{M,i}$ is either zero when there is no cleaning during the year, or:

$$C_{M,i} = P_L + P_s N_t$$

when cleaning is performed. The first part is the fixed maintenance cost where $P_L$ is the cost of labor, and the second part is the cleaning equipment cost and depends on how many tubes are in the condenser, and $P_s$, which is the price of each scraper.

The efficiency cost is an estimate of the productivity lost due to fouling. It was assumed the plant has zero efficiency cost when it operates at the designed efficiency $\eta_d$. Efficiency cost increases proportionally to the amount of extra fuel consumed as a result of fouling. The annual efficiency cost $C_{E,i}$ is computed:

$$C_{E,i} = \frac{P_f}{HV} \int_{t_{i-1}}^{t_i} W_o(\tau) \left( \frac{1}{\eta(\tau)} - \frac{1}{\eta_d} \right) d\tau$$

where $P_f$ is the price of fuel in [\$/kg], HV is the heating value of the fuel in [kJ/kg], $W_o$ is the actual instantaneous power output during the year $i$, and $\eta$ is the actual instantaneous thermal efficiency during the year $i$.

The annual maintenance cost $C_{M,i}$ and efficiency cost $C_{E,i}$ are discounted to present value using a discount rate $r$ to obtain $C_M$ and $C_E$, according to Equations 3.4 and 3.5.

Model Integration

According to the framework structure, the system model is divided into a deterministic model, which include the power and condenser models, and a probabilistic model, including the fouling and operation models.

For simplicity, boiler temperature and pressure are kept as constant parameters, and the only design variables considered are the number of condenser tubes $N_t$, and the condenser tube length $L$. Constraints in the power-condenser model include the minimum turbine back pressure $P_{4,i}$ and the maximum condenser coolant flow rate.
Specify Design Variables

- $N_t$, $L$, $R_{fs}$, $N$

Power-Condenser Model

- $\eta = f_{\text{sur}}(R_f)$

Monte-Carlo Simulation

- Generate $R_{f,i}(t)$

- $\eta_j(t) = f_{\text{sur}}(R_{f,i}(t))$

Next Year
- $j = j + 1$

End of MC?
- No
- End of Lifecycle?
- Yes

Next Iteration
- No

End of Lifecycle?
- Yes

End of Year
- Maintenance?
- Yes or No

End of Year
- Compute Annual Cost
- $C_{E,j}$
- $C_{M,j}$

End of Year
- Anayze Results

End of Year

Figure 3-4: Model Evaluation Process
The deterministic system model has an input-output relationship of:

\[ \eta = f_{sys}(N_t, L, R_{ts}) \]  

(3.16)

In the probabilistic subsystem models, the operational model uses a Monte-Carlo simulation to simulate the lifecycle of the power plant randomly multiple times. The simulation process is shown in the flow chart in Figure 3-4.

Some designer-defined parameter values used in the modeling process are tabulated in Table 3.1. Pricing information is obtained from relevant vendor websites.

Table 3.1: List of user-defined parameters

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Value</th>
<th>Unit</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>(P_2)</td>
<td>80</td>
<td>bar</td>
<td>Feed water pressure</td>
</tr>
<tr>
<td>(T_3)</td>
<td>480</td>
<td>°C</td>
<td>Boiler temperature</td>
</tr>
<tr>
<td>(m_{c,l})</td>
<td>3300</td>
<td>kg/s</td>
<td>Maximum coolant flow rate</td>
</tr>
<tr>
<td>(P_{4,t})</td>
<td>0.1</td>
<td>bar</td>
<td>Minimum steam side pressure</td>
</tr>
<tr>
<td>(W_o)</td>
<td>100</td>
<td>MJ</td>
<td>Rated output power of plant</td>
</tr>
<tr>
<td>(T_5)</td>
<td>25</td>
<td>°C</td>
<td>Coolant inlet temperature</td>
</tr>
<tr>
<td>(D_o)</td>
<td>19</td>
<td>mm</td>
<td>Condenser tube O.D.</td>
</tr>
<tr>
<td>(D_i)</td>
<td>17</td>
<td>mm</td>
<td>Condenser tube I.D.</td>
</tr>
<tr>
<td>(P_m)</td>
<td>20</td>
<td>$/kg</td>
<td>Condenser tube price</td>
</tr>
<tr>
<td>(P_L)</td>
<td>20000</td>
<td>$</td>
<td>Condenser cleaning labor cost</td>
</tr>
<tr>
<td>(P_s)</td>
<td>7</td>
<td>$/tube</td>
<td>Cleaning equipment cost</td>
</tr>
<tr>
<td>(P_f)</td>
<td>0.9</td>
<td>$/kg</td>
<td>Power plant fuel price</td>
</tr>
<tr>
<td>HV</td>
<td>25</td>
<td>MJ/kg</td>
<td>fuel heating value</td>
</tr>
</tbody>
</table>

3.4.2 Traditional Design Method (LMTD Method)

Traditionally, the LMTD design process of choosing a condenser is based on the following three-step process [61]:

1. Determine the condenser heat load \(Q_c\) from the power plant’s requirements

2. Select an operating fouling resistance based on empirical correlations

3. Determine the overall heat transfer coefficient \(U\), the LMTD, and consequently calculate the condenser area, from which \(L\) and \(N_t\) are determined.
The process is similar to adding design redundancy into the condenser. The fouling resistance for different fluids are tabulated in heat exchanger design handbooks. For example, the Tubular Exchanger Manufacturers Association (TEMA) lists fouling resistances for seawater to be $0.176 \; [m^2 K/kW]$ and for brackish water to be $0.352 \; [m^2 K/kW]$.

Once the fouling resistance is selected, the condenser design variables, $N_i$ and $L$ can be determined using the LMTD method. For this case example, fouling resistance was chosen to be $0.3 \; [m^2 K/kW]$, in between the values of seawater and brackish water. Using the traditional design method, the condenser would require 7280 tubes and a tube length of 14.7m.

Based on this specific condenser design, the three different maintenance strategies can be analyzed independently by simulating the lifecycle with uncertainties, and a simple maintenance optimization can be solved to find the most optimal maintenance variable that results in the minimum lifecycle cost ($C_L$) of system. The problem definition is shown below:

$$\text{minimize } \bar{C}_L = F_{d|N_i,L}(\text{Stgy}, N, R_{fs})$$

subject to $g_{|N_i,L}(\text{Stgy}, N, R_{fs}) \leq 0$

where $\bar{C}_L$ is the mean value of the lifecycle cost from the Monte-Carlo system model simulation, Stgy indicates one of the three maintenance strategies being considered: fixed interval, for which $N$ is the maintenance interval in number of years, corrective and preventive, for which $R_{fs}$ is the fouling resistance threshold. $F_d$ and $g$ are the system models described in the section above.

The problem formulation is a mixed-integer nonlinear programming problem. However, since only three different strategies are being considered, they can be analyzed independently, and the other integer variable $N$ is only associated with the fixed interval strategy and can be handled through an exhaustive search, and thus reduces the problem down to a set of continuous nonlinear problems.
3.4.3 Optimal Design Method

The proposed framework allows an optimal condenser design to be found directly through optimization of the full system model. The problem definition is:

\[
\begin{align*}
\text{minimize} & \quad \tilde{C}_L = F_s(N_t, L, \text{Stgy}, N, R_{fs}) \\
\text{subject to} & \quad g(N_t, L, \text{Stgy}, N, R_{fs}) \leq 0
\end{align*}
\]  

(3.18)

Similar to the maintenance optimization problem described in the previous section, this mixed-integer nonlinear programming problem can also be reduced into several continuous nonlinear problems by independently analyzing the three different maintenance strategies.

All of the programming and simulation for this case example were done in MATLAB 2011a. Built-in optimization algorithms including "ga" (genetic algorithm) and "fmincon" (gradient based methods) were used as comparisons.

3.5 Results

This section describes the results from the case study detailed above, to demonstrate the proposed design-maintenance optimization framework, and how it can be used to improve designer’s understanding of trade-offs between different subsystems and disciplines. Comparisons between traditional and optimal design methods, and companions between different maintenance strategies were made.

3.5.1 Single Objective Optimization

The maintenance optimization problem for the traditional design approach described in Equation 3.17 and the optimal design problem described in Equation 3.18 were both solved. The design that resulted in the lowest lifecycle cost was found, and the mean time between maintenance (MTBM) for each design was also computed, which is a measure of the availability of the system. The lifecycle cost and MTBM for each maintenance strategy and the two different design approaches are plotted in
Table 3.2: Results of traditional design vs optimal design

<table>
<thead>
<tr>
<th></th>
<th>(a) traditional</th>
<th>(b) optimal</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_L \ [$M]$</td>
<td>5.00</td>
<td>4.19</td>
</tr>
<tr>
<td>MTBM [yr]</td>
<td>2</td>
<td>2.50</td>
</tr>
</tbody>
</table>

Comparing the three different maintenance strategies, the fixed interval strategy resulted in the worst (highest) lifecycle cost, and also worst (lowest) MTBM. Using maintenance strategies based on fouling-monitoring (corrective and predictive) reduced lifecycle cost, and also extended mean time between maintenance. This is to be expected since the fixed interval strategy ignores any uncertainty associated with degradation, while monitoring-based strategies adapt the maintenance schedule to the actual fouling trend. A preventive maintenance strategy resulted in slightly lower cost and longer MTBM compared to corrective maintenance, because the predictive capability allows an operation to be cut short if the fouling resistance develops rapidly, while it still allow long periods of operation if fouling development is minimal.

Comparing traditional and optimal approaches to the condenser design, the optimal design approach reduced the total lifecycle cost significantly for the fixed-interval maintenance case, while cost reduction in the corrective and preventive maintenance cases were not as large. However, the mean time between maintenance (MTBM) value was improved significantly.

Figure 3-5 shows the break down of the lifecycle cost into capital, cleaning and efficiency costs. In the traditional design approach, the capital cost for all three strategies were the same, because the same condenser design was considered. The fixed interval strategy however had significantly higher cleaning cost and efficiency cost, while the efficiency costs between corrective and predictive maintenance are similar, predictive maintenance has a lower cleaning cost, since the longer MTBM value results in fewer maintenance operations.

With the optimal design approach, cleaning cost was reduced for all three maintenance strategies by using a condenser with fewer numbers of tubes as shown in the
Figure 3-5: Lifecycle Cost Breakdown, Traditional Design vs Optimal Design

Table 3.3: Results of Optimal Condenser Design

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(N_t)</td>
<td>7280</td>
<td>5840</td>
<td>6650</td>
<td>6600</td>
</tr>
<tr>
<td>(L [m])</td>
<td>14.7</td>
<td>17.9</td>
<td>16.5</td>
<td>15.9</td>
</tr>
<tr>
<td>Area ([m^2])</td>
<td>6380</td>
<td>6240</td>
<td>6550</td>
<td>6270</td>
</tr>
</tbody>
</table>

tabulated condenser designs in Table 3.3. Optimization with a fixed interval strategy resulted in significant cost reductions, but the overall cost was still more than designs with monitoring-based maintenance methods. It is interesting to note that the optimal design with corrective maintenance had similar cleaning and efficiency costs compared to the design with predictive maintenance, but the capital cost was higher.
for design with corrective maintenance as a larger condenser was needed (longer tubes and higher area, shown in Table 3.3). This was consistent with results found in similar studies, indicating design redundancy can be replaced by advanced maintenance strategies [38, 41]

### 3.5.2 Parameter Sensitivity

A large number of design parameters were involved in the modeling process, thus understanding how the optimal design changes when the parameter values are changed is a critical step in the design optimization process. In this section, the optimal design approach was re-evaluated under different parameter values, to compare to the optimal design found with the default parameter values.

**Condenser Cost**

The cost of condenser tubes was increased from the default value of $20/kg to $40/kg, and optimization analysis is performed again to find the optimal condenser design with the new parameter value. The results are shown in Table 3.4, and shows that the total lifecycle costs increased roughly 25% from the default parameter values, the MTBM values were lower compared to the default designs, and the condenser area is smaller compared to the default designs.

The increase in condenser capital cost resulted in the optimization algorithm picking a smaller condenser, and reduced the average cleaning interval to compensate for the smaller condenser. It can also be noted that the advantage of predictive maintenance over corrective maintenance appeared to increase, as the MTBM value for the design with predictive maintenance was noticeably longer than corrective maintenance.

**Discount Rate**

In the default analysis no discount rate was considered, and costs incurred anytime during the system lifecycle were considered the same. Table 3.5 shows the condenser
Table 3.4: Optimal Condenser Design: 2× Condenser Cost

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_L$ [M$\text{J}$]</td>
<td>5.64</td>
<td>5.09</td>
<td>5.06</td>
</tr>
<tr>
<td>MTBM [yr]</td>
<td>2</td>
<td>2.43</td>
<td>2.77</td>
</tr>
<tr>
<td>$N_t$</td>
<td>5690</td>
<td>6200</td>
<td>6640</td>
</tr>
<tr>
<td>$L$ [m]</td>
<td>16.0</td>
<td>14.4</td>
<td>14.1</td>
</tr>
<tr>
<td>Area [m$^2$]</td>
<td>5430</td>
<td>5330</td>
<td>5430</td>
</tr>
</tbody>
</table>

Optimization Design Approach [$M$]

- **Fixed Interval**
- **Corrective**
- **Predictive**

Figure 3-6: Lifecycle cost breakdown: 2× Condenser Cost

design found by the optimization algorithm with a discount rate of $r = 5\%$ applied to the model. As expected, the net present lifecycle cost decreased significantly, especially the cleaning and efficiency cost, due to discounting of costs over the lifecycle.

Table 3.5: Optimal Condenser Design: Discount Rate $r = 0.05$

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_L$ [M$\text{J}$]</td>
<td>2.28</td>
<td>2.09</td>
<td>2.06</td>
</tr>
<tr>
<td>MTBM [yr]</td>
<td>2</td>
<td>2.44</td>
<td>2.42</td>
</tr>
<tr>
<td>$N_t$</td>
<td>5580</td>
<td>6230</td>
<td>6210</td>
</tr>
<tr>
<td>$L$ [m]</td>
<td>15.4</td>
<td>13.9</td>
<td>13.3</td>
</tr>
<tr>
<td>Area [m$^2$]</td>
<td>5130</td>
<td>5170</td>
<td>4910</td>
</tr>
</tbody>
</table>

Three effects were noticeable when discount is considered, compared to the default case. The MTBM values for designs with corrective and predictive maintenance had decreased from $\sim 3$ years to $\sim 2.4$ years, the capital cost decreased from $1\text{M}$ to about $0.8\text{M}$, and the optimal condenser became smaller with area in the 5000$m^2$ range, compared to 6000$m^2$ range of the default setting. It is also noticeable that corrective
and predictive maintenance were almost equally effective, with similar lifecycle costs and MTBM values. This result was surprising but reasonable. The discounting makes the cost of maintenance and lost efficiency lower compared to the capital cost, and the major advantage of predictive maintenance is in reducing unnecessary maintenance while keep efficiency high.

Two sensitivity analyses are presented in this section to demonstrate that, by varying the model parameter values, designers can gain insights into how maintenance strategies and design variables interact, and some of the interactions may be non-obvious. This system level approach of modeling physical design and maintenance operation concurrently has the ability to reveal these non-obvious insights about the system. Similar studies can be performed on other parameters and variables of the model to obtain additional information.

### 3.5.3 Multiple Objective Optimization

In this section multi-objective optimization (MOO) was performed to find the Pareto-optimal designs that minimizes lifecycle cost and maximizes MTBM. The problem
formulation is described below:

\[
\text{minimize } \{C_L, -\text{MTBM}\} = F_s(N_t, L, \text{Stgy}, N, R_{fs})
\]

subject to \( \text{MTBM} \leq 5 \) \hspace{1cm} (3.19)

\[
g(N_t, L, \text{Stgy}, N, R_{fs}) \leq 0
\]

The maximum mean time between maintenance (MTBM) value is artificially limited to 5 years for this study. MATLAB build-in multi-objective genetic algorithm "gamultiobj" was used.

Figure 3-8 shows the Pareto-fronts for the optimization results with corrective maintenance and predictive maintenance. Fixed interval maintenance results were significantly higher (by over $1M) compared to the corrective and predictive strategies and are not shown in the figure. The predictive maintenance designs completely dominated the corrective maintenance designs. It is interesting to note that the Pareto-front of predictive strategy is very similar in shape compared to the corrective strategy: one is simply a shifted version of another, and the predictive strategy result in longer MTBM and lower cost. Although not shown here, the fixed interval strategy however produced a completely different shape Pareto front: lifecycle cost equal to $5.6M at MTBM equal to 2 years, and increasing almost linearly to $7.6M at MTBM equal to 5 years. This suggests that predictive and corrective strategies are better at handling systems requiring long MTBM values.

It should be noted that the Pareto-fronts generated through multi-objective genetic algorithm are not the analytical Pareto-fronts, and the shapes of the fronts are sensitive to model parameters variations. Despite these short-comings, the Pareto-fronts can still provide valuable insights for the designers when making decisions on selecting a maintenance strategy.

### 3.6 Conclusions & Future Work

This study was motivated by the need for optimizing lifecycle cost of complex engineering systems, especially desalination plants, as suggested by the interviews with
desalination designers reported in Chapter 2. A novel design framework is demonstrated for integrating maintenance in the design stage by considering the effect of system design parameters on the physics of component degradation. The framework is capable of revealing non-obvious trade-offs between design and maintenance decision variables in the preliminary design stage, and allow designers to better understand the system level interactions, and select designs that are more optimal over their lifecycles.

A case study of a condenser design was used to evaluate the significance of different maintenance strategies and their effects on system design decisions. Three maintenance policies were considered: fixed maintenance interval, corrective maintenance based on degradation threshold, and predictive maintenance based on the prediction of future degradation. Single objective optimization, sensitivity analyses, and multi-objective optimization were conducted on the case study, and the results found show that by concurrently optimizing both the system design variables and the maintenance variables, the redundancy in the design can be reduced when an advanced maintenance strategy is considered. However, depending on the values of system parameters, the effects of maintenance policies on design may vary in magnitude.
Two major assumptions were made in this work to simplify the analysis. First, the costs associated with implementing different maintenance strategies are neglected. In reality, there would be additional capital and operational costs associated with the necessary sensors for degradation monitoring, and computing equipment for making degradation predictions. However, the lifecycle cost analysis presented here can still reveal valuable information about the different maintenance strategies, and can help designers decide whether the benefits from advanced maintenance strategies can justify their added costs. The second major assumption is that the algorithm used in predictive maintenance can perfectly predict the future. Any uncertainties associated with the prediction algorithm would reduce the advantages of predictive maintenance observed in this study. Future research should consider simulations with different values of prediction uncertainties to understand their effects.

The proposed framework takes a system level approach by integrating the maintenance strategies and lifecycle analysis with the design process, and thus significantly increased the problem complexity. Efforts to reduce complexity included decoupling the uncertain degradation and maintenance models from the deterministic system models, and implementing mid-fidelity physics based models. Future research efforts should look into the effects of having multiple degrading components.

Future study should also focus on exploring the design-maintenance space by including the different uncertain parameters, such as energy and resource pricing to evaluate the sensitivity of a design to maintenance policy.
Chapter 4

Experimental design task:
Development of reverse osmosis plant model

4.1 Motivation

As described in the introduction chapter, the second phase of the thesis work will focus on understanding how designers manage a large number of parameters during the design of desalination systems. To accomplish this goal, a design simulation environment needed to be developed to experimentally test designers’ behaviors during the design process. The first step to create such a simulation environment is to develop a detailed model of a complete reverse osmosis plant, then implement the model with a functional user interface. This chapter will focus on the development of the reverse osmosis plant model, while the next chapter will describe the human behavior experiment that uses the model described here.

4.2 Background in Reverse Osmosis Modeling

A number of different studies have examined different aspects/approaches to RO modeling, as well as the application of numerical optimization to RO models. Different
models exist for RO membranes, and most employ the solution-diffusion model. In the solution diffusion model, water molecules and salt ions are assumed to dissolve into the RO membrane and then diffuse through the membrane due to a concentration gradient [64]. The solution diffusion model can be applied to hollow fiber membranes [65] as well as spiral-wound membranes [66]. Some use a differential model approach to simulate variations along the length of a membranes [65], while a majority take a lumped parameter approach, assuming uniform properties over a membrane element [66, 15]. Studies have also investigated the accuracies of solution diffusion models that neglect concentration polarization and pressure drops in the membrane [67]. The lumped parameter version of the solution diffusion model is widely used in optimization studies of RO plants, [68, 69]. Some use a simplified version of the model that neglects pressure drops along the membrane [70], or neglect concentration polarization [71].

Alternatively, RO plant models can also be constructed through data-driven modeling approaches. Leidner, et al. analyzed the effects of changing water quality, daily operations schedule and input prices through regression analysis of pilot plant operational data [72]. Libotean, et al. applied a neural network-based modeling approach to forecast RO plant performance and operational diagnostics. Surrogate models are also used in some optimization studies to reduce numerical complexity [73].

There are a large pool of RO cost models in the literature. Karagiannis reviewed several publications regarding water desalination costs and found that for plant capacities between 12,000 and 60,000 m³/day, the water cost is in the range of $0.44 to $1.62/m³ [74]. Reddy and Ghaffour compiled historical price data of membrane desalination and thermal desalination, and made future cost predictions. The general consensus is that there is high variability in water cost, and that the actual cost is highly site specific [74]. Most studies of RO optimization in the literature include an economics model, usually consisting of cost estimations of RO system components, such as intake and pre-treatment [75], pumps and energy recovery devices [69, 76].

Research has likewise considered ways to find a cost-effective RO network configuration, which is a representation of how RO modules, pumps and energy recovery
devices are connected. RO super-structures have been demonstrated to successfully represent different types of RO flow structures, such as single-pass structure and a 2-pass structure. Figure 4-1 shows the superstructure documented by Vince, et al. [69]. A review of various superstructure approaches are summarized in [77].

![Diagram of reverse osmosis flexible superstructure](image)

Figure 4-1: Example of reverse osmosis flexible superstructure proposed by Vince, et al. [69]

RO membrane fouling and replacement are modeled [75, 78] and included in RO optimization studies [79]. Boron is a critical element in seawater that could be difficult to remove by the RO process, and the modeling of boron removal by RO membranes are also documented [80]. Since RO systems are generally large energy consumers, models of renewable energy powered RO systems are also reported in the literature [81, 82]. Other area of RO modeling has focused on risk management [83] and optimal operations [84].

### 4.3 Reverse Osmosis System Model Development

The model developed in this study will be calibrated using actual cost and capacity values of a 30,000m³/day capacity seawater reverse osmosis plant currently operating in the Middle East. Plant data are drawn from design documents of the plant, but due to confidentiality reasons the name and location of the plant cannot be revealed. The system boundary of a reverse osmosis desalination system is defined to include
the intake and pre-treatment system, the RO process, RO membranes and pressure vessels, and the life-cycle analysis. The energy required for the RO plant is assumed to be drawn from the electric grid. The post-treatment process and water distribution system are not considered.

Figure 4-2 shows the layout of the reverse osmosis plant being modeled. Three different subsystems can be identified based on the plant layout: the intake & pre-treatment (IP) subsystem, which deals with preparing the feed seawater for desalination, the RO process flow structure (FS) subsystem, which is focused on the selection of pumps, energy recovery devices (ERD) and how they are connected, and the RO pressure vessel (RO) subsystem that involves selecting RO membrane and designing the pressure vessel rack. Life-cycle analysis (LC) considers the finance, operation and maintenance of the plant. This figure shows a single-pass single-stage flow structure, though the model developed will have a flexible flow structure that can accommodate both single-pass and two-pass flow structures.

4.3.1 Reverse Osmosis Unit Model

The RO subsystem involves the design of the pressure vessel rack, and the selection of the membrane. An RO pressure vessel rack consists of a number of identical pressure vessels \( N_{pv} \) connected in parallel, and each pressure vessel contains several RO membrane elements \( N_{memb} \) connected in series, as shown in Figure 4-3. A pressure
vessel rack can house hundreds of pressure vessels, while each pressure vessel can contain up to 8 membrane elements. Seawater feed flows into the pressure vessels at one end and into the first RO membrane. The concentrated brine from the first RO membrane then feed into the second RO membrane and so forth. The brine from the last membrane element leaves the pressure vessel as the brine stream shown in Figure 4-3. Each of the RO membranes inside the pressure vessel produces a permeate flow of different concentrations. For brackish water reverse osmosis or low salinity seawater reverse osmosis, all permeate flows are grouped into one permeate stream coming out of the pressure vessel, while for high salinity seawater reverse osmosis, the first membrane elements (referred to as front elements) produce lower salinity permeate compared to the later elements (back elements), and thus front elements’ permeate flows are grouped together into permeate stream #1, while the back element permeate flows are grouped together into permeate stream #2 [85]. The element position at which this grouping occurs is defined as variable "psplit".

![Figure 4-3: Schematic of reverse osmosis unit, including the pressure vessels, membranes, feed, brine, and permeate streams](image)

Each RO membrane element is the standard spiral wound polyamide membrane. Essentially, the membrane element consists of two large membrane sheets, where seawater flows on the outside of the sheets and permeate flows in between. The
sheets are rolled up to reduce the membrane’s footprint and support spacers are placed in between the membranes. The simplified transport phenomenon at one of the membrane sheet is shown in Figure 4-4.

Figure 4-4: Illustration of transport phenomenon at RO membrane. Grey line in the middle of the figure is the RO membrane. The colors green, blue, and purple indicate salt concentrations from low to high. White dots are representation of salt molecules.

The water flow rate, pressure, and concentration vary along the length of the membrane, from the feed terminal to the brine terminal. Due to concentration polarization, the salt concentration next to the membrane will be higher compared to the concentration further away from the membrane. Water molecules and salt molecules diffuse through the RO membrane when pressure is applied. Since the membrane has a higher water permeability compared to salt permeability, the permeate on the other side of the membrane will have a much lower salt concentration. For simplicity, a lumped-parameter model is used to describe the membrane element in this study, and a membrane element is characterized by the following parameters: flow rate, pressure, and concentration at each of the feed, brine, and permeate terminals \((Q_f, P_f, c_f, Q_b, P_b, c_b, Q_p, P_p, c_p)\), the salt concentration at the membrane wall \(c_w\), the total membrane surface area \(S\), the membrane water permeability \(A\) and salt permeability \(B\).
Solution-Diffusion Model

According to the solution-diffusion model [86, 69, 87, 88], the flux of water \((J_v)\) and salt \((J_s)\) across a semipermeable membrane are governed by:

\[
J_v = A F_A(t) K_{TCF}(\Delta P - \Delta\pi) \tag{4.1}
\]

\[
J_s = B F_B(t) K_{TCF}(c_w - c_p) \tag{4.2}
\]

where \(F_A(t)\) and \(F_B(t)\) are the membrane fouling factors over time, with \(F = 1\) representing a clean RO membrane, and \(F < 1\) when fouling has developed. \(K_{TCF}\) is a temperature correction factor for the feed water temperature \(T\) and equals:

\[
K_{TCF} = \begin{cases} 
\exp \left( \frac{2640}{298} \left( \frac{1}{T} - \frac{1}{298} \right) \right) & \text{if } T \geq 298 \\
\exp \left( \frac{3020}{298} \left( \frac{1}{T} - \frac{1}{298} \right) \right) & \text{if } T < 298
\end{cases} \tag{4.3}
\]

\(\Delta P\) is the pressure differences across the membrane and \(\Delta\pi\) is the differences in osmotic pressure across the membrane. \(\Delta P\) is calculated by assuming a linear pressure gradient between the feed terminal and the brine terminal, where the pressure drop \((P_{\text{drop}})\) is caused by hydrodynamic resistances along the membrane [89].

\[
\Delta P = P_f - 0.5 P_{\text{drop}} - P_p \tag{4.4}
\]

\[
P_{\text{drop}} = P_f - P_b = 9.5 \times 10^8 \left( \frac{Q_f + Q_b}{2 \rho} \right)^{1.7} \tag{4.5}
\]

where \(\rho\) is the density of seawater. The osmotic pressure difference is estimated by assuming pure sodium chloride (NaCl) solution for seawater, using equation:

\[
\Delta\pi = \frac{2\phi RT \rho}{M_{\text{NaCl}}} (c_w - c_p) \tag{4.6}
\]
where $\phi$ is the osmotic coefficient, $R$ is the universal gas constant, and $M_{NaCl}$ is the molar mass of sodium chloride. The membrane wall concentration can be calculated using the concentration polarization estimation:

$$c_w - c_p = \left( \frac{c_f + c_b}{2} - c_p \right) e^{0.7r_m}$$  \hspace{1cm} (4.7)

$$r_m = \frac{Q_p}{Q_f}$$  \hspace{1cm} (4.8)

where $r_m$ is the membrane recovery ratio.

Based on the salt and water flux values, the permeate flow rate and concentration can be calculated:

$$Q_p = (J_v + J_s M_{NaCl})S$$  \hspace{1cm} (4.9)

$$c_p = \frac{J_s}{J_v}$$  \hspace{1cm} (4.10)

Fluid mass and salt mass are conserved within the membrane element, yielding two additional equations:

$$Q_b = Q_f - Q_p$$  \hspace{1cm} (4.11)

$$Q_b c_b = Q_f c_f - Q_p c_p$$  \hspace{1cm} (4.12)

Equations 4.1 to 4.12 completely describes the lumped-parameter model of the RO membrane element. Among the variables associated with the model, $A$, $B$, $S$, $T$, $\rho$, $\phi$, $R$, and $M_{NaCl}$ are either constants or can be found in a seawater property table, and the feed water conditions $Q_f$, $c_f$, permeate size pressure $P_b$, and one of $P_f$ or $r_m$ are given based on design choices. Assume constant fouling factors $F_A$ and $F_B$, which leaves 12 unknowns for 12 equations.

An iterative process is needed to solve the membrane element solution-diffusion
model. For a given feed pressure, first guess the value of permeate flow rate $Q_p$, and assume $c_w = c_f$ and $J_s = 0$ initially, then solve the entire set of equations and update the values of assumed variables at the end of each iteration, and stop when the values of $Q_p$ have converged to a relative tolerance <0.5%. At the end, the solver outputs $Q_p$, $c_p$, $Q_b$, $P_b$, $c_b$, and $r_m$. Membrane elements placed in a pressure vessel are solved sequentially, with the output of one element becoming the input of the next.

**Membrane Library**

Different models of membranes can have different cost, surface area, and permeability values. Typically, designers have a large number of different RO membranes to select from. A small library of RO membranes is assembled in this model development effort, which consists four seawater membranes and one brackish water membranes. The attributes of the membranes are obtained from Dow Water & Process Solutions and listed in Table 4.1. Additional membranes from different companies could be added in the future for a more complete collection of membranes.

**Table 4.1: Library of reverse osmosis membranes**

<table>
<thead>
<tr>
<th>$I_{RO}$</th>
<th>Membrane Model</th>
<th>Salt Reject</th>
<th>Boron Reject</th>
<th>$A$</th>
<th>$B$</th>
<th>$S$</th>
<th>Cost [90]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SW30ULE-440i</td>
<td>99.7</td>
<td>89</td>
<td>$5.03 \times 10^{-4}$</td>
<td>$3.50 \times 10^{-5}$</td>
<td>41</td>
<td>1200</td>
</tr>
<tr>
<td>2</td>
<td>SW30XLE-440i</td>
<td>99.7</td>
<td>88</td>
<td>$4.12 \times 10^{-4}$</td>
<td>$1.93 \times 10^{-5}$</td>
<td>41</td>
<td>1000</td>
</tr>
<tr>
<td>3</td>
<td>SW30HRLE-440i</td>
<td>99.8</td>
<td>92</td>
<td>$3.39 \times 10^{-4}$</td>
<td>$1.59 \times 10^{-5}$</td>
<td>41</td>
<td>930</td>
</tr>
<tr>
<td>4</td>
<td>SW30XHR-440i</td>
<td>99.8</td>
<td>93</td>
<td>$2.72 \times 10^{-4}$</td>
<td>$1.16 \times 10^{-5}$</td>
<td>41</td>
<td>920</td>
</tr>
<tr>
<td>5</td>
<td>BW30HR-440i</td>
<td>99.7</td>
<td>83</td>
<td>$1.10 \times 10^{-3}$</td>
<td>$3.38 \times 10^{-5}$</td>
<td>41</td>
<td>1000</td>
</tr>
</tbody>
</table>

The variable $I_{RO}$ is an integer defining the membranes in this library. Note that membrane suppliers do not provide the actual permeability values $A$ and $B$. Instead, the rated flow rate and salt rejection are provided at a standard flow condition, typically 32,000 ppm feed concentration, 55 bar pressure, and 8% membrane recovery at 25°C. The permeability values can be estimated based on this information by solving Equations 4.1 to 4.12.
Table 4.2: Constraints on reverse osmosis membrane elements [91]

<table>
<thead>
<tr>
<th></th>
<th>Seawater ultrafiltration pretreatment</th>
<th>Seawater conventional pretreatment</th>
<th>RO Permeate 2nd pass</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>maximum feed flow rate</strong> $Q_f$ ($m^3/h$)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BW membranes</td>
<td>-Not Recommended-</td>
<td>-Not Recommended-</td>
<td>17</td>
</tr>
<tr>
<td>SW membranes</td>
<td>18</td>
<td>15</td>
<td>18</td>
</tr>
<tr>
<td><strong>minimum brine flow rate</strong> $Q_b$ ($m^3/h$)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BW membranes</td>
<td>-Not Recommended-</td>
<td>-Not Recommended-</td>
<td>2.3</td>
</tr>
<tr>
<td>SW membranes</td>
<td>3.0</td>
<td>3.4</td>
<td>2.3</td>
</tr>
<tr>
<td><strong>maximum feed pressure</strong> $P_f$ (bar)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BW membranes</td>
<td>-Not Recommended-</td>
<td>-Not Recommended-</td>
<td>41</td>
</tr>
<tr>
<td>SW membranes</td>
<td>83</td>
<td>83</td>
<td>83</td>
</tr>
<tr>
<td><strong>maximum membrane recovery</strong> $r_m$ (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BW &amp; SW</td>
<td>16</td>
<td>14</td>
<td>30</td>
</tr>
</tbody>
</table>

In some advanced designs, two different types of membranes may be used in a single pressure vessel to optimize performance and cost, since the front elements in the pressure vessel can operate under vastly different feed conditions compared to the back elements. In this model it was assumed that only a single type of membrane is used for each pass/stage of RO racks.

**Constraints**

Several design constraints must be satisfied by all the RO membranes in the pressure vessel. These include the maximum feed pressure ($P_f$) that could be applied, the maximum feed flow rate ($Q_f$), the minimum brine flow rate ($Q_b$), and the maximum membrane recovery ($r_m$). The constraints depend on the quality of the feed water. Different pre-treatment technologies would have an effect on the constraints, and second pass membranes have a different set of constraints compared to the first pass membranes. The constraint values are available from Dow Water & Process Solutions, and are tabulated in Table 4.2.
Subsystem Level Interface

The RO subsystem takes inputs from both the intake & pre-treatment subsystem and also the flow structure subsystem. The overall mathematical formulation of the RO subsystem is:

\[
[CC_{PV,i}, CC_{memb,i}, X_{FS}] = f_{RO}(r_i, N_{PV,i}, N_{memb,i}, I_{RO,i}, psplit_i, X_{RO}, I_{pre})
\]

\[
X_{FS} = \{P_{f,i}, (Q_{p1,i}, c_{p1,i}, b_{p1,i}), (Q_{p2,i}, c_{p2,i}, b_{p2,i}), (Q_{b,i}, P_{b,i}, c_{b,i}, b_{b,i})\}
\]

\[
X_{RO} = \{Q_{f,i}, c_{f,i}, b_{f,i}\}
\]

where \( r_i \) is the overall recovery ratio of the entire RO pass, \( N_{PV,i} \) is the number of pressure vessels in parallel, \( N_{memb,i} \) is the number of RO membranes inside a single pressure vessel, \( I_{RO,i} \) is the type of RO membranes selected from the membrane library. All three of these variables are integer variables. All pressure vessels are identical and are connected in parallel. In each pressure vessel, \( N_{memb,i} \) is the number of identical RO membranes connected in series. \( psplit \) defines how the permeate stream is divided into front and back streams. \( psplit = N_{memb} \) represent all permeate flow in the front stream (single pass system), and \( psplit = 0 \) represent all permeate flow in the back stream (2-pass system with no permeate split). These variables are the typical system level design parameters defined by desalination process engineers.

\( X_{RO} \) contains input to the RO subsystem that is provided by the flow-structure subsystem (described in Equation 4.19). where \( Q_{f,i}, c_{f,i} \) and \( b_{f,i} \) are the flow rate, salt and boron concentration of the feed water to the RO unit.

\( CC_{PV,i} \) and \( CC_{memb,i} \) are the capital costs of pressure vessels and membranes, respectively. The capital cost is calculated by multiplying the unit cost of pressure vessel or RO membrane by the total number of pressure vessels/membranes. The unit cost of the pressure vessel is assumed to be $1,000.

\( X_{FS} \) contains all parameters that are fed to the flow-structure. \( P_{f,i} \) is the feed pressure that is required to achieve the recovery ratio \( r_i \) specified in the input. Since the solution-diffusion model described in Equations 4.1 to 4.12 requires \( P_{f} \) as an input, a
non-linear solver is needed to find the value of $P_{f,i}$. $(Q_{p1,i}, c_{p1,i}, b_{p1,i}), (Q_{p2,i}, c_{p2,i}, b_{p2,i})$ are the flow rate, salt and boron concentrations of the front permeate stream (subscript p1) and the back permeate stream (subscript p2). $Q_{b,i}, P_{b,i}, c_{b,i}, b_{b,i}$ are the flow rate, pressure, salt and boron concentrations of the brine stream.

For a two pass flow structure, the RO subsystem would consist pressure vessels and membranes from both passes, and hence an index $i$ is associated the equation.

4.3.2 Flow Structure (FS) Model

The FS subsystem involves configuring the connectivity between pumps, pressure vessels, and energy recovery devices. The majority of RO optimization work in the past has focused on this area, and many novel approaches have been developed, especially super-structure optimization [68, 69]. In this study, a flexible two-pass single-stage flow structure will be implemented. The structure is flexible to also represent a single-pass single-stage flow structure without any modifications, and can also incorporate some of the advanced design techniques, such as permeate splitting and brine recirculation. Figure 4-5 is an illustration of the flow structure.

Feed water coming from the intake & pre-treatment stage first flows into the first RO pass ($Q_{f,1}$), becomes pressurized through the high-pressure pump (HPP1), and flows into the first pass RO unit, which consists of hundreds of individual membrane elements. There are two (or three) output streams of the RO unit, the brine stream ($Q_{b,1}$), which flows through the energy recovery device (ERD) to boost the feed water pressure by recovering the residual pressure in the brine stream, which then goes to brine disposal. The permeate stream can be further divided into two separate streams, permeate 1 ($Q_{p1,1}$) and permeate 2 ($Q_{p2,1}$). Permeate 1 flows directly to the product, while permeate 2 flows to the second pass and becomes its feed source ($Q_{f,2}$). Recent designs of reverse osmosis plants usually split the permeate directly inside the pressure vessel. A second high pressure pump (HPP2) pressurizes the feed water to go through the second pass RO unit. The permeate from the second pass ($Q_{p,2}$) does not need to be split and combined with permeate stream #1 from the first pass to form the product. Part of the brine ($Q_{b,2}$) can be recombined with the original feed water.
Figure 4-5: Reverse osmosis flexible flow structure
to increase the overall flow rate, the ratio of recirculation is defined by $R_{\text{bb}}$. It is also possible to blend some of the feed directly with the product to increase production rate ($Q_\#$), but this practice is only used in brackish water reverse osmosis when the feed water has a low salinity. It is generally not recommended for seawater reverse osmosis. The pressure, flow rate, salt and boron concentration of each stream is either calculated based on mass conservation, or results of RO unit subsystem calculations.

**Capital Cost & Energy Consumption**

The high-pressure pump (HPP) raises the pressure of the feed seawater from atmospheric pressure to the required RO feed pressure. The efficiency of the high-pressure pumps and motor unit is assumed to be $\eta_{\text{HPP}} = 76\%$. The capital cost of a high pressure pump depends on the flow rate and pump pressure [68]:

$$CC_{\text{pump}} = 51 (Q_{\text{pump}} \Delta P_{\text{pump}})^{0.96} \quad (4.14)$$

The energy consumption of a pump is:

$$W_{\text{pump}} = \frac{Q_{\text{pump}} \Delta P_{\text{pump}}}{\eta_{\text{pump}}} \quad (4.15)$$

The booster pump (BP) is used to raise the pressure of the feed stream that flows through the energy recovery device (ERD). The booster pump requires a smaller pressure differential compared to the high-pressure pump. The efficiency of the booster pump is assumed to be $\eta_{\text{BP}} = 80\%$. The method for calculating its capital cost and energy consumption is the same as the high-pressure pump.

The ERD uses the residual pressure of the brine stream to pressurize the feed stream through an isobaric process, where the flow rate is the same between the feed and brine streams. The efficiency of the energy recovery device is defined as [92]:

74
Currently, the majority of isobaric energy recovery devices are manufactured by Energy Recovery Inc, and they report an efficiency of $\eta_{ERD} = 97\%$ for their products. The capital cost of the energy recovery device, also provided by Energy Recovery Inc., is directly proportional to the flow rate:

$$CC_{ERD} = 439Q_{ERD} \tag{4.17}$$

The total energy consumption of the reverse osmosis process is calculated in units of energy per volume of water produced (kWh/m$^3$)

$$EC_{FS} = \frac{\Sigma W_{pump}}{Q_{prod}} \tag{4.18}$$

**Subsystem Interface**

The FS subsystem is heavily interconnected with the RO subsystem, and depends on several parameters from the RO subsystem as inputs, including feed pressure required to operate RO unit, permeate and brine stream flow rate, salt and boron concentrations. The FS subsystem also outputs feed water flow rate, salt and boron concentration to the RO subsystem. The mathematical formulation for the FS subsystem is:

$$[CC_{FS}, EC_{FS}, Q_{prod}, c_{prod}, b_{prod}, X_{RO}] = f_{FS}(Q_{in}, c_{in}, b_{in}, N_{train}, Q_{ff}, R_{bb}, X_{FS}) \tag{4.19}$$

where $Q_{in}$, $c_{in}$ and $b_{in}$ are the flow rate, salt and boron concentration of the feed water entering the RO pressure vessel rack, $N_{train}$ is the number of individual trains,
$X_{FS}$ are the parameters supplied by the RO subsystem described in Equation 4.14.

$CC_{FS}$ are the total capital costs of the pumps and energy recovery devices. Costs of piping are neglected in this version of the model but can be easily implement in a future iteration. $Q_{prod}$, $c_{prod}$, $b_{prod}$ are the flow rate, salt and boron concentration of the product water. $X_{RO}$ are parameters being passed to the RO subsystems.

Solving the outputs of the flow-structure subsystem is straight-forward if brine recirculation ($R_{bb}$) is zero. Simply start from the feed water junction and proceed down-stream, using mass-conservation to solve for the flow-rates and ion concentrations in each of the streams. Once the properties of all streams are computed, the power consumption and capital costs of components can be calculated. If brine recirculation is not zero, then an iterative process is required to solve the flow-structure, by assuming brine flow is zero initially.

### 4.3.3 Intake & Pre-treatment (IP) Model

The IP subsystem covers the design choices associated with the intake technology and pre-treatment technology. For this version of the model, two intake structures were considered: deep water intake or shallow water intake. Two different pre-treatment processes are possible, conventional pre-treatment with cartridge filters or ultrafiltration pre-treatment [6, 93]. The capital cost, operational cost and energy consumption of the intake and pre-treatment system depends on the combination of intake and pre-treatment choices [93, 94], as well as the intake seawater flowrate [75]. The mathematical formulation for the IP subsystem interface is:

$$[CC_{IP}, EC_{IP}, OC_{IP}] = f_{IP}(Q_{in}, I_{intake}, I_{pre})$$  \hspace{1cm} (4.20)

where CC, EC, OC stand for capital cost, energy consumption, and operational cost, $Q_{in}$ is the intake seawater flowrate, while $I_{intake}$ and $I_{pre}$ are integer variables indicating the choices of technology in intake and pre-treatment respectively.

Costs of intake structures depends largely on the geography of the plant location, and does not vary significantly with intake capacity. For intake flow rate of less than
150,000m³/day, the capital cost of the two different intake structures are obtained from the existing plant design document:

\[
\text{Shallow open: } C_{\text{intake}} = \$6.6 \text{ Million} \quad (4.21) \\
\text{Deep open: } C_{\text{intake}} = \$15.6 \text{ Million} \quad (4.22)
\]

Since the shallow open intake collects seawater closer to shore, the water quality can be significantly worse than deep open intake. Therefore, conventional pre-treatment cannot be used with shallow open intake, and is only compatible with deep open intake. Ultrafiltration pre-treatment is compatible with both shallow and deep open intake. The capital cost of the pre-treatment system is modeled based on a power law [75]:

\[
\text{Conventional: } C_{\text{pre}} = 1926Q_{\text{in}}^{0.8} \\ \\
\text{Ultrafiltration: } C_{\text{pre}} = 2769Q_{\text{in}}^{0.8} \quad (4.24)
\]

The operational cost also differs depending on the choice of pre-treatment. The following values are gathered from literature and actual plant data in the Middle East, and include both chemical consumption and cartridge/filter replacement [93]:

\[
\text{Conventional: } O_{\text{CIP}} = \$0.0118/m^3 \\ \\
\text{Ultrafiltration: } O_{\text{Cpre}} = \begin{cases} 
\$0.0398/m^3, \text{ shallow open intake} \\
\$0.0226/m^3, \text{ deep open intake}
\end{cases} \quad (4.26)
\]

The energy consumption is assumed to be the same for different pre-treatment technologies:

\[
E_{\text{CIP}} = 0.15 kWh/m^3 \quad (4.27)
\]
4.3.4 Lifecycle Model

The life-cycle (LC) subsystem takes inputs from all subsystems and computes several system-level performance outputs. Capital costs of the intake & pre-treatment subsystem, flow structure subsystem, and RO units are combined to find the total reverse osmosis system capital cost. The yearly operating cost of the system is also found based on operating costs of the subsystems. Assuming the plant has a fixed production capacity factor, and amorting the capital and operating costs over the lifecycle of the plant, a total water price (TWP) in $/m³ can be calculated. The plant is assumed to have a lifecycle of 25 years, the discount rate is 5%, and the electricity cost is $.03/kWh. The mathematical formulation of the subsystem is:

\[
[TWP, CC, EC, OC] = f_{LC}(CC_{RO}, CC_{memb}, CC_{FS}, CC_{IP}, EC_{FS}, EC_{IP}, OC_{IP}, Q_{in}, Q_{prod}, N_{train})
\]

(4.28)

4.3.5 System Integration

![Diagram of reverse osmosis system model. Input variables are indicated in red, output variables indicated in blue. Shared variables are black.]

Figure 4-6: N² Diagram of reverse osmosis system model. Input variables are indicated in red, output variables indicated in blue. Shared variables are black.
Figure 4-6 shows the $N^2$ diagram of the reverse osmosis systems. There are a total of fifteen input variables to the system, and seven objective variables as output, listed in Table 4.4. The diagram suggests that there is a strong coupling between the subsystems FS and RO as indicated by the feedback loop between them, and only a weak connection between the rest of the subsystems. The IP, FS, and RO subsystems all have outputs that feed into the life-cycle subsystem, forming a hierarchy of three lower level subsystems reporting to a central system administrator. The variables associated with each subsystem and their connectivity are shown in Table 4.3.

Table 4.3: Number of variables in subsystems

<table>
<thead>
<tr>
<th></th>
<th>lower level</th>
<th>upper level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IP</td>
<td>FS</td>
</tr>
<tr>
<td>input variables</td>
<td></td>
<td></td>
</tr>
<tr>
<td>design variable (shared)</td>
<td>3 (2)</td>
<td>4 (2)</td>
</tr>
<tr>
<td>coupling variable</td>
<td>0</td>
<td>19</td>
</tr>
<tr>
<td>total</td>
<td>3</td>
<td>23</td>
</tr>
<tr>
<td>output variables</td>
<td></td>
<td></td>
</tr>
<tr>
<td>shared with upper level</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>shared with lower level</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>objective (shared)</td>
<td>0</td>
<td>3 (1)</td>
</tr>
<tr>
<td>total</td>
<td>3</td>
<td>11</td>
</tr>
</tbody>
</table>

*outputs of lower level subsystems

The strong coupling between flow structure and RO subsystem is indicated by the large number of variables shared between them. The number of variables passed between the two subsystems depends on what type of flow structure is used. For a single-pass flow structure, the number of shared variables is reduced by half. Combining the flow structure and RO into a single large subsystem can eliminate the intermediate variables, at the cost of creating one large subsystem that can be come slightly more difficult to comprehend.

The entire reverse osmosis system model is implemented in MATLAB 2013b. A Matlab class structure is created that has all the model parameters as private class properties. Each subsystem is implemented as a stand-alone function of the class and can be evaluated independently. To analyze a single design of the plant model, an
iterative multi-disciplinary analyzer (MDA) is required. The MDA would evaluate
the IP subsystem first, then an iterative process is needed to evaluate the flow-structure
and RO unit completely, because of the shared variables between the two sub-system.
Once the flow-structure and RO subsystems converged, the lifecycle subsystem can be
evaluated to finish the multi-disciplinary analysis.

Table 4.4: Summary of input & output variables

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Bounds</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Q_{in}$</td>
<td>Feed water flow rate</td>
<td>50,000 to 100,000 m³/day</td>
</tr>
<tr>
<td>$I_{intake}$</td>
<td>Type of intake structure</td>
<td>shallow vs deep</td>
</tr>
<tr>
<td>$I_{pre}$</td>
<td>Type of pre-treatment</td>
<td>conventional / ultrafiltration</td>
</tr>
<tr>
<td>$N_{train}$</td>
<td>Number of RO trains</td>
<td>up to 6</td>
</tr>
<tr>
<td>$Q_{ff}$</td>
<td>Permeate blending flow rate</td>
<td>less than 30 m³/day</td>
</tr>
<tr>
<td>$R_{bb}$</td>
<td>Brine recirculation ratio</td>
<td>0 to 100% of brine flow rate</td>
</tr>
<tr>
<td>$N_{PV,1}$</td>
<td>Number of pressure vessel (PV)</td>
<td>up to 100</td>
</tr>
<tr>
<td>$N_{PV,2}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$N_{memb,1}$</td>
<td>Number of membranes per PV</td>
<td>up to 8</td>
</tr>
<tr>
<td>$N_{memb,2}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$I_{RO,1}$</td>
<td>Type of membrane</td>
<td>5 types to choose from</td>
</tr>
<tr>
<td>$I_{RO,2}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r_{1, 2}$</td>
<td>Recovery ratio of RO pass 1 &amp; 2</td>
<td>30% to 95%</td>
</tr>
<tr>
<td>psplit</td>
<td>Position to split permeate stream</td>
<td>0 to 8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Output Variables</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$Q_{p}$</td>
<td>Product water flow rate</td>
</tr>
<tr>
<td>$c_{p}$</td>
<td>Product water salt concentration</td>
</tr>
<tr>
<td>$b_{p}$</td>
<td>Product water boron concentration</td>
</tr>
<tr>
<td>TWP</td>
<td>Total water price ($/m³)</td>
</tr>
<tr>
<td>CC</td>
<td>System capital cost ($)</td>
</tr>
<tr>
<td>EC</td>
<td>System energy consumption (kWh/m³)</td>
</tr>
<tr>
<td>OC</td>
<td>System operating cost ($/day)</td>
</tr>
</tbody>
</table>

4.4 Model Verification

Model verification is done on two levels. First, verify the solution-diffusion model
used to simulate the reverse osmosis membranes by comparing it against a commer-
cial software package, then verify the system level model by comparing design data
obtained from existing reverse osmosis plant to the model output.

### 4.4.1 RO membrane model

The RO membrane model described in Equations 4.1 to 4.12 was compared to the Reverse Osmosis System Analysis (ROSA) program provided by Dow Water & Process Solutions [24].

Figure 4-7 shows comparisons of RO unit recovery ratio computed by both ROSA software and the custom developed RO membrane model. The RO unit used for comparison consisted of 65 pressure vessels, 7 membranes elements per vessel, and membrane type SW30HRLE-440i. In Figure 4-7a, the feed flow rate is varied while the feed pressure is kept constant at 70 bar, and in Figure 4-7b the feed pressure is varied while feed flow rate is kept constant at 500 m\(^3\)/h. The results show some slight discrepancy between ROSA results and model results, but the difference is small (around 1% value difference in recovery ratio for all flow rates and feed pressures tested), and a consistent trend between them.

Figure 4-8 shows comparison of RO unit recovery ratio when membrane type (Figure 4-8a) and number of pressure vessels (Figure 4-8b) are varied. Feed pressure
Figure 4-8: RO Model Verification. (a) different types of RO membrane element. (b) varying number of pressure vessels were kept constant at 70 bar, and feed flow rate kept constant at 500 m$^3$/h. Results show consistent trends with slight variations in numerical values.

Figure 4-9: RO Model verification, salt concentration and recovery ratio

Figure 4-9 shows the permeate salt concentration predicted by ROSA vs cus-
tom developed model. The salt concentration predicted by ROSA is roughly 50 ppm lower than the model results, but both follow an identical trend. The reason for this discrepancy is mainly because the unavailability of official permeability values. Since the permeability values were back calculated from the typical salt-rejection ratio published by membrane suppliers, which were only accurate to 3 significant digits. Doing this back calculation would inevitably introduce error in the permeability values due to the lack of significant digits, and was reflected in the salt concentration of the permeate. The consistency in the trends of the results also suggest that the differences are due to different values of constants being used. The lumped-parameter model used in the custom built model add another level of inconsistency when compared to ROSA.

Since the trend for the permeate salt concentration was consistent with ROSA, it was assumed that the discrepancy would not drastically change how the system is designed.

4.4.2 System level performance

Table 4.5 lists the design parameters reported in the design document of the Middle East desalination plant that was used as a reference in this section.

The plant is a 2-pass seawater reverse osmosis plant. The seawater at the proposed site has a TDS of 45000 ppm and temperatures between 15°C and 30°C, and analyses for both conditions were available. The 15°C case was used for model verification purposes. Table 4.6 shows the comparison of the system level predictions of the model vs the design proposal.

Since this proposed design was used to calibrate some of the model parameters, there is very good agreement between the model prediction and the design proposal. The major discrepancies come from the capital costs and energy consumptions not considered in the model, such as post-treatment cost (roughly $1M), civil and electrical costs (roughly $13M). The unit price of membrane elements assumed in the model are based on online prices, which would be higher than the whole-sale prices directly between vendors and contractors.
Table 4.5: Design parameters of existing reverse osmosis plant case example

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>feed water flow rate</td>
<td>75,000 m$^3$/day</td>
</tr>
<tr>
<td>intake structure</td>
<td>deep open</td>
</tr>
<tr>
<td>pre-treatment</td>
<td>conventional</td>
</tr>
<tr>
<td>number of trains</td>
<td>6</td>
</tr>
<tr>
<td><strong>first pass</strong></td>
<td></td>
</tr>
<tr>
<td>number of pressure vessels</td>
<td>65</td>
</tr>
<tr>
<td>membranes / vessel</td>
<td>7</td>
</tr>
<tr>
<td>feed pressure</td>
<td>70 bar</td>
</tr>
<tr>
<td><strong>second pass</strong></td>
<td></td>
</tr>
<tr>
<td>number of pressure vessels</td>
<td>14</td>
</tr>
<tr>
<td>membranes / vessel</td>
<td>7</td>
</tr>
<tr>
<td>feed pressure</td>
<td>14 bar</td>
</tr>
<tr>
<td>brine recirculation</td>
<td>100%</td>
</tr>
</tbody>
</table>

Table 4.6: Model prediction vs plant proposal values

<table>
<thead>
<tr>
<th></th>
<th>design proposal</th>
<th>model prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>water production</td>
<td>30,000 m$^3$/day</td>
<td>30,600</td>
</tr>
<tr>
<td>system recovery</td>
<td>40%</td>
<td>40%</td>
</tr>
<tr>
<td>capital cost (CC)</td>
<td>$M 50.94</td>
<td>$M 39.2</td>
</tr>
<tr>
<td>CC, intake</td>
<td>$M 15.4</td>
<td>$M 15.6</td>
</tr>
<tr>
<td>CC, pre-treatment</td>
<td>$M 15.36</td>
<td>$M 15.68</td>
</tr>
<tr>
<td>CC, RO train</td>
<td>$M 5.98</td>
<td>$M 8.3</td>
</tr>
<tr>
<td>CC, other</td>
<td>$M 14.2</td>
<td>-</td>
</tr>
<tr>
<td>energy consumption</td>
<td>kWh/m$^3$ 4.1</td>
<td>3</td>
</tr>
<tr>
<td>operational cost</td>
<td>$/day -</td>
<td>$/day 11,300</td>
</tr>
</tbody>
</table>

This model was developed with the intent of simulating design processes so that experiments can be conducted with design participants. Thus very precise predictions of plant performances over a large range of operating parameters were not necessary, and making this model adequate for its purpose.
4.5 Summary & Future Applications

This chapter described the details of a system level, multi-disciplinary model of a reverse osmosis plant. The modeling process followed a multi-disciplinary design optimization approach of identifying subsystems, mapping interactions between subsystems, and identifying multiple objectives. Four subsystems were identified: intake & pre-treatment, flow structure, reverse osmosis membrane unit, and lifecycle. Each of the subsystems were described in details, including the assumptions, governing equations, constraints, and interfaces with other subsystems. The system model was implemented in MATLAB and verified by comparing to existing reverse osmosis software packages and using data from a real plant for calibration.

The level of model fidelity is not high enough to give exact predictions of performances and costs for a wide range of plant sizes, but provides trade-off information among the input and output parameters. In the next chapter, this model will be used in the implementation of a reverse osmosis design process simulation.

The flow-structure subsystem proposed a flexible representation of reverse osmosis flow structure. The current implementation of the model only considered a double-pass-single-stage structure, it can easily be extended to a multiple-pass-multiple-stage structure. Optimization techniques can be applied to find non-obvious flow structure that are optimal for plants with special requirements.

Future studies should also focus on increasing model fidelity by including water distribution subsystem and energy source subsystems, and eventually create a regional desalination network system model that include multiple desalination plants and power plants that service the water and energy demand of a local community.
Chapter 5

Designer behaviors in parameter-based Reverse Osmosis design

One goal of this thesis research is to link the design process with the way designers actually design. A key challenge of complex systems design is that of comprehension. By nature, complex systems deal with a massive number of design parameters, and human designers must make sense of them. Previous chapters have focused on how complex systems can be optimized computationally, but this chapter explores how designers manage varying scales of complexity during design process.

5.1 Motivation

Interviews conducted with practitioners in the desalination industry had revealed that a common process used to design a reverse osmosis plant involves an engineer/designer tuning design parameters in a trial-and-error fashion. Typically, the engineer/designer starts with a set of guidelines, or a previous design, and adjust design parameters until all system requirements were satisfied.

To give context to this, figure 5-1 shows the user interface of a widely used reverse osmosis design evaluation software program named Reverse Osmosis System Analysis
Figure 5-1: User interface of ROSA, a commercial reverse osmosis design software (ROSA), available for free from Dow Water & Process Solutions [24]. The software allows designers to specify several input parameters, and then generate a report that shows the performance of the design in terms of cost, water production and energy consumption, as well as any design constraint violations.

The process of designing a reverse osmosis system using software packages can be classified as a parameter design process, in which the designer uses a set of input parameters to make changes in a set of output parameters [27]. A parameter design problem is a well defined problem, with clearly identified design parameters (the input parameters), objectives and constraints (the output parameters). Parameter design problems can either be coupled or uncoupled. Uncoupled problems have a one-to-one mapping of the input parameters to the output parameters, whereas in the coupled problems, multiple input parameters can affect an output parameter simultaneously. In a reverse osmosis design problem, the input parameters are recovery ratios, feed flow rate, and numbers of reverse osmosis (RO) membranes. Examples of output parameters include permeate TDS, energy consumption, and capital cost, among others.
Past research in the area of parameter design problems have shown that, due to cognitive limitations, human designers are very inefficient at solving generic parameter design problems even with as few as 5 variables [27, 28, 95], however, the reverse osmosis process design problem typically have more than 10 variables. This can be an area of significant concern for the desalination industry: What are ways to make the engineers and designers of RO systems more efficient and effective at their work? This chapter will describe a set of human subject experiments to evaluate how designers approach technical parameter design problems. The experiments would investigate how participants address design parameter problems for a range of problem scales, to understand various approaches that were used.

5.1.1 Related Work

A few studies have been conducted in the past decade related to humans’ abilities to solve design parameter design problems. Hirschi and Frey were one of the first to conduct parameter design experiments on human subjects [27]. They performed a series of experiments in which human subjects were asked to solve generic mathematical parameter design problems that ranged from 2-input-2-output (2x2) variables to 5-input-5-output (5x5) variables, using a custom-built computer user interface. They found that the time taken to solve parameter design problems was on the order of $O(n)$ for uncoupled problems, and $O(n^{3.4})$ for coupled problems, where $n$ is the number of input variables. They also found that coupled problems with more than 4 variables were extremely difficult and frustrating for the test subjects to solve. Grogan reported a study in his PhD thesis, where he compared solving parameter design problems individually vs collaboratively, using test problems – similar to Hirschi and Frey’s experiment – that have no technical context. His results were in agreement with the results of Hirschi and Frey [95]. Flager, et al. conducted similar research using parameter design problems specific to the building design domain [28]. Their results indicated that the design solution quality found by their test subjects decreased with increase in problem scale, following an exponential relationship similar to that reported by Hirschi and Frey. Austin-Breneman, et al. conducted a set of collabora-
tive design optimization experiments where human subjects were asked to optimize a simplified satellite design system, and these experiments illustrated humans can be inefficient at optimization and have little system level awareness [96].

All of these past studies have unanimously suggested that human designers are not efficient at solving coupled parameter design problems. The challenges of making sense of coupled problems made people struggle significantly with problems having more than 4 variables, and the strategies people used were often sub-optimal.

The work presented in this thesis adds to the existing literature by looking at problems specifically in the context of desalination systems, how subject’s technical knowledge play a role in their performance, and focus on strategies and characteristics of successful approaches.

5.1.2 Research Question

The work in this chapter will focus on understanding the following question:

1. How human designers make sense of the complexity of reverse osmosis design problems?

2. How domain knowledge is used in parameter design problems?

3. What are efficient strategies to solve parameter design problems?

The expectation is that there will be some common characteristics that will improve the efficiency of solving parameter design problems, such as methods to simplify the problem and better solution finding techniques. Technical knowledge about the problem domain should also have an impact on the ability to solve problems.

The results of this research could reveal some of the limitations of human designers and also the strategies that make the parameter design process more efficiently, and could have a significant impact on the developing of computer-aided engineering design tools and the curriculum of design education.
5.2 Experimental Methods

In this study, human subjects were asked to design a reverse osmosis system using various sets of design parameters. Figure 5-2 shows a relationship map of the major factors that could affect the output of this study.

![Figure 5-2: Factors affecting experimental outcome](image)

The experiment was designed to measure test subjects' ability to design RO processes by presenting a set of sample design problems for the test subjects to solve. The problem complexity and the efficiency of test subjects would presumably both affect the outcome of the experiment. Problem complexity could be further affected by factors including the degree of coupling, problem scale, the type of design task, the initial condition, time limit, and subjects' mental focus. Factors that could affect the efficiency of test subjects include their knowledge level about the test problem domain, the user interface of the experiment, any external aids available, learning effect between multiple problems, and also the mental focus of test subjects. Of these 11 factors, the problem scale and subject's knowledge were independent variables that would be tested, and is highlighted in red in Figure 5-2. Test subjects' mental focus might be affected by a number of uncontrollable factors and needed to be monitored through the experiment. The rest of the factors could be controlled through proper design of the experimental procedure.
5.2.1 Design Problem and User Interface

The design problem used for the experiment was a 2-pass seawater reverse osmosis (SWRO) process similar to existing reverse osmosis plants. Seawater intake, pre-treatment, post-treatment and maintenance were not considered so the test subject only needed to focus on the reverse osmosis process itself.

Figure 5-3: 2-Pass SWRO process used in design problem

The reverse osmosis numerical model developed in Chapter 4 was used to perform the design evaluation in the computer GUI. Figure 5-3 depicts the 2-pass SWRO process used in the design problem. There were 10 input parameters highlighted in yellow: the feed flow rate, permeate blending flow rate, brine recirculation percentage, recovery ratio, number of pressure vessels, and membranes per pressure vessel for both RO passes, and the permeate split of the first pass RO. Other parameters, such as the feed water TDS and boron concentration, type of RO membranes, efficiency of pumps and energy recovery devices (ERD) were assumed to be constants. There were 5 performance parameters that were the output of the program highlighted in green: the product flow rate, product TDS, product boron concentration, capital
cost, and energy consumption. In addition to the 5 performance parameters, there were also 4 design constraints for membranes in each of the RO passes: the maximum membrane feed flow rate, the minimum membrane brine flow rate, the maximum membrane recovery ratio, and the maximum feed pressure. These parameters are among the most critical parameters to a reverse osmosis plant. The design problem was a non-linear mix-integer problem with non-linear constraints.

5.2.2 Coupling Effect

To evaluate whether the input-output parameters of the design problem are coupled, a parameter-based design structure matrix (DSM) [97] was constructed. Also known as low-level schedule DSM, parameter-based DSM are effective for integrating low-level design processes based on physical design parameter relationships [98, 99]. The parameter-based DSM for this design problem was shown in Figure 5-4.

![Figure 5-4: SWRO design problem parameter-based DSM](image-url)
The DSM was constructed by analyzing the physical model of the SWRO process to map out all dependencies between the input and output parameters. In the DSM, the input parameters are highlighted in yellow, the output parameters are in green, and red highlighted cells indicate dependencies between the variables in the corresponding row and column of that cell. This DSM showed that every output parameter is dependent on several input parameters, and thus making this design problem a highly coupled problem.

5.2.3 Design Tasks & Complexities

Two different approaches to the experiment tasks were considered. One was the satisficing approach [100], where the subjects would be asked to find any design that meets a set of target performances, and the time or iterations it takes them would be measured. Another was the optimization approach, where the subjects would be asked to minimize one or more performance parameters given a time limit, and solution quality would be measured. Since the optimization approach could be heavily influenced by test subject’s familiarity with numerical optimization, it was decided that optimization was not the best option for this study, and thus the satisficing approach was chosen.

Experiments conducted in previous research had used test problems that had the same number of input parameters and output parameters, and described the scale of the problem as 2x2, 3x3, etc. Five different problems were selected for this study based on a similar guideline and are shown in Table 5.1. Each output parameter was assigned a target value which the test subject must achieve, effectively turning the output parameter into constraints. Any unused parameter would be assigned a very generous target that would always be met, and any unused input parameters would be kept constant.

The constants, target values, and numerical complexity of each test problem are shown in Table 5.2. The values of the output targets for each problem were selected to ensure diversity between the problems while at the same time the numerical complexities of all problems were kept consistent. To benchmark a problem’s numerical
Table 5.1: List of test problems

<table>
<thead>
<tr>
<th>Problem Scale</th>
<th>Input Parameters</th>
<th>Output Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>2x2</td>
<td>recovery, pass 1 (r₁) recovery, pass 2 (r₂)</td>
<td>product flow rate (qₚ) energy consumption (EC)</td>
</tr>
<tr>
<td>3x3</td>
<td>feed flow rate (qₐ) recovery, pass 1 (r₁) recovery, pass 2 (r₂)</td>
<td>product flow rate (qₚ) product TDS (cₒ) energy consumption (EC)</td>
</tr>
<tr>
<td>4x4</td>
<td>feed flow rate (qₐ) recovery, pass 1 (r₁) permeate split (split) recovery, pass 2 (r₂)</td>
<td>product flow rate (qₚ) product TDS (cₒ) energy consumption (EC) capital cost (CC)</td>
</tr>
<tr>
<td>5x5</td>
<td>feed flow rate (qₐ) recovery, pass 1 (r₁) # of pressure vessels, pass 1 (npv₁)</td>
<td>all output parameters</td>
</tr>
<tr>
<td></td>
<td>recovery, pass 2 (r₂) # of pressure vessels, pass 2 (npv₂)</td>
<td></td>
</tr>
<tr>
<td>10x5</td>
<td>all input parameters</td>
<td>all output parameters</td>
</tr>
</tbody>
</table>

complexity, the input design space was sampled using a latin-hypercube sampling method and then evaluated to find designs that meet the output targets. The average number of latin-hypercube samples required to find one satisficing design was used to represent the problem complexity, and is shown in the last row of Table 5.2. The numerical complexities of problems from 2x2 to 5x5 roughly follow an exponential relationship, except for the 10x5 problem, which was made slightly easier to ensure the latin-hypercube sampling could arrive at a target design in a reasonable time.

5.2.4 Computer Graphic User Interface

The design problem was presented to the test subjects in a form of graphic user interface. The interface could be either an existing software package (such as ROSA), or a customized software program. The advantages of using an existing software package were that some test subjects may already be familiar with how to navigate them, and commercial software are usually stable and bug-free. But these software packages would have many features that were not useful to the experiment and could
Table 5.2: Constants & target values for each test problem

<table>
<thead>
<tr>
<th>feed water quality</th>
<th>2x2</th>
<th>3x3</th>
<th>4x4</th>
<th>5x5</th>
<th>10x5</th>
</tr>
</thead>
<tbody>
<tr>
<td>feed TDS (ppm)</td>
<td>$c_i$</td>
<td>45000</td>
<td>46000</td>
<td>44000</td>
<td>35000</td>
</tr>
<tr>
<td>feed boron conc. (ppm)</td>
<td>$b_i$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>input parameters</th>
<th>2x2</th>
<th>3x3</th>
<th>4x4</th>
<th>5x5</th>
<th>10x5</th>
</tr>
</thead>
<tbody>
<tr>
<td>feed flow rate (m³/h)</td>
<td>$q_{in}$</td>
<td>615</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>permeate blending (m³/h)</td>
<td>$q_{fr}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>recovery, pass 1 (%)</td>
<td>$r_1$</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td># of pressure vessels, pass 1</td>
<td>$n_{pv1}$</td>
<td>65</td>
<td>65</td>
<td>65</td>
<td>*</td>
</tr>
<tr>
<td>membranes / vessel, pass 1</td>
<td>$n_{mb1}$</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>permeate split</td>
<td>$p_{split}$</td>
<td>2</td>
<td>3</td>
<td>*</td>
<td>3</td>
</tr>
<tr>
<td>recovery, pass 2 (%)</td>
<td>$r_2$</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td># of pressure vessels, pass 2</td>
<td>$n_{pv2}$</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>*</td>
</tr>
<tr>
<td>membranes / vessel, pass 2</td>
<td>$n_{mb2}$</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>brine recirculation rate (%)</td>
<td>$r_{bb}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>target output parameters</th>
<th>2x2</th>
<th>3x3</th>
<th>4x4</th>
<th>5x5</th>
<th>10x5</th>
</tr>
</thead>
<tbody>
<tr>
<td>target product flow rate, ± 3% (m³/h)</td>
<td>$q_p$</td>
<td>208*</td>
<td>278*</td>
<td>208*</td>
<td>243*</td>
</tr>
<tr>
<td>max product TDS (ppm)</td>
<td>$c_o$</td>
<td>300</td>
<td>150*</td>
<td>200*</td>
<td>150*</td>
</tr>
<tr>
<td>max product boron conc. (ppm)</td>
<td>$b_i$</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5*</td>
</tr>
<tr>
<td>max energy consumption (kWh/m³)</td>
<td>EC</td>
<td>3.350*</td>
<td>3.460*</td>
<td>3.216*</td>
<td>2.858*</td>
</tr>
<tr>
<td>max capital cost ($M)</td>
<td>CC</td>
<td>3</td>
<td>3</td>
<td>1.7*</td>
<td>1.8*</td>
</tr>
</tbody>
</table>

| numerical complexity | $33e^2$ | $25e^3$ | $28e^4$ | $31e^5$ | $30e^{6.8}$ |

*active parameters
become distractions, would not allow customization, and would not collect experimental data automatically. Custom built software could be constructed to interact with users, change problem complexity dynamically, and collect data automatically. The information shown to the users could be controlled, and made simple enough that subjects with varying levels of desalination knowledge could be tested. The only drawback for using a custom built user interface was the need for development and debugging. Based on these advantages, a custom graphic user interface (GUI) was developed for this study.

The GUI software program was created in MATLAB to interact with the RO numerical model from Chapter 4. The GUI took some design cues from ROSA. A screenshot of the GUI is shown in Figure 5-5. All input parameters were listed in the top-left panel of the GUI as either sliders, text-boxes, or drop-down menus. Each input had an upper and lower limit, as shown in Table 5.3. Parameters that only range from 1 to 8 were controlled using a single drop-down menu, while the other parameters were controlled by a slider and a text box, either of which could modify the parameter's value. All output parameters were listed in the bottom-right panel, with each parameter corresponding to a text box followed by another text box to its right to indicate the target or constraint values that must be satisfied. The outputs were divided into performance constraints, which were the output parameters listed in Table 5.2, and design constraints, which were limitations imposed on the RO membranes. The top-right panel had an image of the 2-pass SWRO process, the feed water quality (highlighted in green), and intermediate variables could be used to get a sense of the test problem. The bottom-left panel showed additional information about the experiment itself, including the time remaining and current progress.

The GUI would present the five problems to the test subjects in pseudo-random order. There were 4 different possible sequences for problems from 2x2 to 5x5: order 1: 2–3–4–5; order 2: 3–2–5–4; order 3: 4–5–2–3; order 4: 5–4–3–2. This sequences introduced some randomness to the problem order the problem, at the same time limited the effect of different sequences could have on the subjects so that all subjects that solved a particular problem last would have gone through the same problems.
Figure 5-5: Graphic user interface for the test problem
Table 5.3: Upper and lower limits of GUI inputs

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Lower Limit</th>
<th>Upper Limit</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>feed flow rate</td>
<td>400</td>
<td>1000</td>
<td>600</td>
</tr>
<tr>
<td>permeate blending</td>
<td>0</td>
<td>2.5</td>
<td>2.5</td>
</tr>
<tr>
<td>recovery, pass 1</td>
<td>30%</td>
<td>60%</td>
<td>30%</td>
</tr>
<tr>
<td>recovery, pass 2</td>
<td>40%</td>
<td>90%</td>
<td>50%</td>
</tr>
<tr>
<td># of pressure vessels, pass 1</td>
<td>1</td>
<td>200</td>
<td>200</td>
</tr>
<tr>
<td># of pressure vessels, pass 2</td>
<td>1</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>membranes / vessel &amp; permeate split</td>
<td>1</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>brine recirculation</td>
<td>0</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

first. One exception was the 10x5 problem, which would either be presented second or last, so that a subject would not be overwhelmed by an intimidating problem as the first problem. The order of the problems were recorded and would be used to determine if there was any ordering effect. Active input parameters for each problem were highlighted in blue, the rest of the inputs were gray and inactive. The design was only evaluated when the "calculate" button (located in the top-right panel) was pressed, and the simulation usually took less than 1 second to complete. Each click of the "calculate" button was considered one design iteration. The output parameters were then updated, and any un-satisfied constraints would be highlighted in red. A count-down timer was displayed in the bottom-left panel that would turn yellow at the 5 minute mark, and turn red when time expired. The test subjects were allowed to perform one more calculation after the time had expired, then had to move on to the next problem. When a design that satisfied all constraints was found, a message would display to inform the subject that he had successfully solved the problem, the "calculate" button would be disabled, the "next" button would be enabled, and the timer would be stopped. Clicking on the "next" button would bring the subject to the next problem. Because participation was voluntary, test subjects were allowed to skip any problem at any point of the experiment, however, once a problem was skipped, the subjects could not come back to this problem anymore. Subjects were also informed that they could stop the experiment at anytime.
The time limit for each problem was 20 minutes, except for the 10x5 problem, which had a limit of 30 minutes. There were no limits on the number of design iterations a test subject could perform within the time limit. Initially, all input variables were set to their lower limits, as shown in the first column of Table 5.3. Three pilot studies were conducted prior to the official experiments to ensure the time limit and problem difficulty level were adequate.

5.2.5 Experimental Procedure

Test subjects were recruited from the Mechanical Engineering department of MIT, including undergraduate, graduate students and post-doctoral researchers. They were invited to a pre-arranged office desk space, with a PC running the experiment GUI on MATLAB 2013b. The subjects were first provided informed consent, and filled out a background questionnaire that collected demographic information, and their past experiences (research, work, or classes taken) in desalination related fields and systems engineering fields.

A short introduction was given to the design task. The presentation consisted of an overview of the graphic user interface tool, explanation of how to change the design parameters to calculate a new design, the objective of the problems and time limit, followed by a short presentation of the 2-pass seawater reverse osmosis process being designed, how the parameters relate to the process and definitions of terminologies. The subjects would then perform the design tasks uninterrupted. Pen, paper and a calculator were provided as optional external aids. On each click of the "calculate" button, a large set of data was recorded, including all current input and output parameter values, number of constraints violated, and time elapsed. There was also a screen capture program installed on the work station PC and a video camera recorded the behavior of the test subjects. An experimenter sat behind the test subject and made notes of any observations.

After the design tasks were completed, there was a debriefing interview. The subjects were asked whether they felt difficulty or frustration with the experiment, whether they felt the time limit was adequate, whether the design task was similar
to anything they did. Any noteworthy behavior observed by the experimenter was discussed, and comments and suggestion about the experiments were elicited from the test subjects. The overall experiment took between 1 to 2 hours to complete.

A copy of each of the background questionnaire and debriefing interview can be found in Appendix A.

5.3 Results & Analysis

Results are grouped into the following three sub-sections: qualitative observations, high level quantitative analysis, detailed analysis of strategies. The qualitative observation section focuses on the observations made during experiments, without any statistical analysis. The high level analysis looks at several performance metrics and the overall trend of test subjects. The detailed analysis of strategies takes a closer look at the characteristics of each test subject during the problem solving process.

5.3.1 Test Subject Demographics

In total, 22 test subjects were recruited and tested. 19 of the subjects were in the age range of 20-29, and 3 were in the range or 30-39. The gender ratio was 17 male to 5 female. There were 2 undergraduate students, 3 masters students, 14 PhD students, and 3 post-docs.

Nine of the subjects had no previous work or research experience with desalination, 8 subjects had worked or did research in desalination for 2 years or less, and 5 subjects had 3 or more years of experiences. Meanwhile, 9 of the 22 subjects had taken at least one desalination class in the past. Four of the 22 subjects have had at least 1 year of work or research experience in the field of systems engineering.

Based on their technical experiences in desalination and systems engineering, the subjects could be divided into four mutually exclusive groups: 1. subjects with strong desalination knowledge; 2. subjects with systems engineering knowledge; 3. subjects with some desalination knowledge; 4. subjects without any relevant knowledge. The grouping metrics were shown in Table 5.4. It was assumed that if a subject had
Table 5.4: Technical experiences of participants

<table>
<thead>
<tr>
<th>Group</th>
<th>Experience Level</th>
<th>Number of Subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>strong desalination</td>
<td>3+ years desalination, and taken desalination classes</td>
<td>5</td>
</tr>
<tr>
<td>knowledge</td>
<td></td>
<td></td>
</tr>
<tr>
<td>systems engineering</td>
<td>1+ years systems engineering</td>
<td>4</td>
</tr>
<tr>
<td>knowledge</td>
<td></td>
<td></td>
</tr>
<tr>
<td>some desalination</td>
<td>≤ 2 years desalination, or taken desalination class, and no systems knowledge</td>
<td>7</td>
</tr>
<tr>
<td>knowledge</td>
<td></td>
<td></td>
</tr>
<tr>
<td>no relevant knowledge</td>
<td>none</td>
<td>6</td>
</tr>
</tbody>
</table>

3 years or more experiences working/doing research in the field of desalination, and assume 8 hours work days, then they would have accumulated roughly 10,000 hours and could be considered an expert in the field [101].

5.3.2 Qualitative Observations

Use of External Aids

When dealing with even moderately complex problems, it was expected that participants might turn to external aids to help keep track of parameter values for each iterations. For this reason, a pen, paper and a calculator were provided.

Ten out of the 22 subjects were observed to make notes on the paper provided, and 3 subjects used a calculator, however 2 of these 3 subjects only used the calculator once during the entire experiment, and it wasn't clear if the calculator helped these subjects in solving the problem. The only subject that used the calculator extensively were estimating the feed flow rate based on the product flow rate constraints and recovery ratios. In short, about half of the participants used external aids to help track their progress, which meant that the other half used some other strategy, possible memorization, or not keeping track at all.
Problem Difficulty

All subjects appeared to be fully engaged during the experiment. Some subjects appeared to be frustrated at some point during the experiment, but did not lose mental focus. No subject elected to skip any problem.

Only 11 subjects were able to solve all five problems in the given time limit. The 10x5 problem seemed to give subjects the most trouble, especially if it was presented in the second problem. The 3x3 problem also appeared to be difficult for a number of subjects, it was very common to see a subject get all the constraints satisfied except for the product TDS constraint, and then spend a very long time and still unable to satisfy the last constraint.

Sensitivity Testing

It was very common for a test subject to test the sensitivity of variables at the beginning of a problem. Two different methods of testing variable sensitivity were observed: a subject may make a change in one input, and then looks at how the output parameters changed, and then revert the changes made in that input back to the original value, and then move on to change another input; a subject may also change the value of one input over several iterations, going from the lower limit of the input to its upper limit, set the input to some value in between, and then move on to different input variable. 17 of the 22 subjects were observed performing these sensitivity testing procedures at some point of the experiment.

Input Parameters

It was overwhelmingly obvious that most of the time a test subject would only change one parameter in each design iteration. The values of changes made to input parameters each iteration can vary significantly between test subjects, and it appears that subjects who make relatively small changes at each iteration (clicking on the sliders a few times vs dragging the slider) tend to take longer to solve a problem.

Some subjects clearly had a sense of what the values of input parameters should
be, for example some subject would never bring the first pass recovery ratio above 50%, or setting all input parameters on the first iteration to values typical of reverse osmosis plants. There were also some subjects who did not have any sense of expected input parameter values at the beginning of the experiment, but learned over the course of the experiment and were able to guess the correct ranges input parameters and set them accordingly on the first iterations of later problems.

**Fixation on Constraints**

It was also observed that some subjects were fixated on reducing the number of constraint violations, or the number of "red boxes" rather than thinking about the problem as a whole. If the subject changed a single input parameter and resulted in multiple constraints becoming violated, the subject would immediately revert any changes to reduce the number of constraint violations.

**Debriefing Interview**

Only 4 subjects felt that the experiment was easy, 4 subjects felt the experiment was very difficult, and the rest felt the difficulty was moderate. Some subjects mentioned that a few of the problems were difficult, the 10x5 problem were mentioned to be difficult by 6 subjects, 3x3 were mentioned by 4 subjects, and 5x5 were mentioned by 3 subjects.

Although about half of the subjects were not able to complete all 5 problems in the given time limit, only 2 subjects felt that the time limit was too short and wanted to have more time.

When asking whether they were frustrated during any part of the experiment, 18 subjects felt some level of frustration during the experiment. Some subjects mentioned which one of the problems frustrated them: the 10x5 problem were mentioned 6 times, 5x5 were mentioned 5 times, 3x3 were mentioned 4 times, and 4x4 were mentioned once. Subjects who thought a particular problem was difficult did not necessarily think that problem was frustrating.

When asked if the design problems were similar to anything they did in the past,
12 subjects felt the experience was similar to other design tasks they did in the past, including PID controller tuning, power plant design, sizing of heating-cooling systems, curve fitting, and structural design tuning. Surprisingly, 10 subjects said during the interview that the experiment was fun and enjoyable, with 2 subjects explicitly saying that it felt like a video game.

When asked what approaches were used by the subjects when solving the problem, 9 subjects said they had a sense of what values some of the input parameters should be based on their technical knowledge, 3 subjects said they did not have prior knowledge about the parameters’ values, but had a sense toward the end of the experiment. 1 subject said he guessed the correct range of the parameters by looking at the upper and lower limits of slider controllers. 3 subjects mentioned explicitly that their approach focused on reducing the constraints, or "eliminating the red boxes", which was consistent with our observation of subjects trying to avoid constraint violations. Some other key words that were mentioned by multiple subjects included "make a good initial jump toward the solution area", "trial and error", "solve first pass and second pass independently".

5.3.3 Quantitative Analysis

Completion Rate

Table 5.5a shows how often a test subject completed the test problems in the given time limit. 11 subjects were able to finish all 5 problems in the given time, 19 subjects finished at least 4 problems, and everyone finished at least 2 problems. The average completion rate for the subjects is 86%, or 4.3 problems.

Looking at the completion rate for each problem in Table 5.5b, there was a visible dip in the completion rate of the 3x3 problem, compared to the 2x2 and 4x4 problems. This was consistent with the observations during the experiment, the reason the 3x3 problem was more difficult compared to 2x2 and 4x4 would be discussed in section 5.3.4.

Since some of the subjects were not able to complete some test problems, the time
Table 5.5: Problem completion rate

(a) Problems Completed by Subjects

<table>
<thead>
<tr>
<th>Problems Completed</th>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>≤1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Subjects</td>
<td>11</td>
<td>8</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

(b) Completion Rate per Problem

<table>
<thead>
<tr>
<th>Problem Difficulty</th>
<th>2x2</th>
<th>3x3</th>
<th>4x4</th>
<th>5x5</th>
<th>10x5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Completion Rate</td>
<td>100%</td>
<td>77.3%</td>
<td>95.5%</td>
<td>86.4%</td>
<td>72.7%</td>
</tr>
</tbody>
</table>

and iterations measurements collected would be right-censored in nature. Therefore, only a few statistical analyses are suitable for this study. The log-rank test based on the Kaplan-Meier procedure could be used to compare significant differences between two populations [102, 103]. A modified Spearman’s correlation test could be used to find correlation coefficients involving censored measurements, where any censored measurements would be multiplied by a constant factor of 2 before being used in the correlation analyses [104]. Comparison of uncensored data were done using Wilcoxon rank sum test, which is a non-parametric test that does not assume a prior distribution of data points.

**Completion Time**

Table 5.6: Time taken to complete each problem (minutes)

<table>
<thead>
<tr>
<th></th>
<th>2x2</th>
<th>3x3</th>
<th>4x4</th>
<th>5x5</th>
<th>10x5</th>
</tr>
</thead>
<tbody>
<tr>
<td>median</td>
<td>1.4</td>
<td>7.1</td>
<td>5.1</td>
<td>13.4</td>
<td>18.9</td>
</tr>
<tr>
<td>minimum</td>
<td>0.7</td>
<td>2.2</td>
<td>0.6</td>
<td>2.3</td>
<td>1.1</td>
</tr>
<tr>
<td>maximum</td>
<td>10</td>
<td>20*</td>
<td>20*</td>
<td>20*</td>
<td>30*</td>
</tr>
</tbody>
</table>

*problem incomplete

The time taken to solve each test problem was summarized in Table 5-6. Previous research had suggested that there was an exponential relationship between problem
scale (number of input parameters) and task time [27] described in Equation 5.1:

\[ t(n) = a_0 \cdot a_1^n \]  

\[ \text{linear regression } \Rightarrow \ln(t(n)) = b_0 + b_1 \cdot t \]  

The time measurements from problems 2x2 to 5x5 were plotted in Figure 5-6 using box plots with the result of regression analysis of an exponential model overlaid on top. The 10x5 problem was not included in the analysis because its numerical complexity was designed to be significantly simpler compared to the rest (indicated in Table 5.2).

Figure 5-6: Box plot of problem times overlaid with exponential relationship. Boxes bound the 1st and 3rd quartiles, whiskers bound extremes within 1.5 times the interquartile range.

Model parameter \( b_1 = 0.505, (a_1 = e^{0.505} = 1.66) \) was found to be significant: \( t(82) = 6.36, p < 0.001 \). However the adjusted \( R^2 \) value was only 0.322, compared to \( R^2 > 0.7 \) found in previous studies, suggesting that subjects’ performances did not follow an exponential relationship found in previous studies. The 3x3 problem appeared to be more difficult compared to the 4x4 problem. Average time taken to solve problem 3x3 and problem 4x4 were compared using the log-rank test and a
p-value of 0.05 was found. Detailed analysis revealed that the coupling effects in the 3x3 problem affected subjects' performances differently compared to the 4x4 problem. This phenomenon would be discussed in detail in Section 5.3.4.

**Problem Iterations**

Table 5.7: Iterations taken to complete each problem

<table>
<thead>
<tr>
<th></th>
<th>2x2</th>
<th>3x3</th>
<th>4x4</th>
<th>5x5</th>
<th>10x5</th>
</tr>
</thead>
<tbody>
<tr>
<td>median</td>
<td>11</td>
<td>70</td>
<td>32</td>
<td>96</td>
<td>93</td>
</tr>
<tr>
<td>minimum</td>
<td>5</td>
<td>12</td>
<td>2</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>maximum</td>
<td>31</td>
<td>172*</td>
<td>162</td>
<td>177*</td>
<td>179*</td>
</tr>
</tbody>
</table>

*problem incomplete

The number of iterations needed to complete each problem were summarized in Table 5.7. Regression analysis was applied to results from problem 2x2 to 5x5, using an exponential model similar to Equation 5.1. The results were plotted in Figure 5-7.

![Iterations taken to solve](image)

Figure 5-7: Box plot of problem Iterations overlaid with exponential relationship. Boxes bound the 1st and 3rd quartiles, whiskers bound extremes within 1.5 times the interquartile range.

Similar to the problem time measurement, model parameter $a_1 = 1.65$ was found to be significant: $t(82) = 6.52$, $p < 0.001$, and adjusted $R^2 = 0.334$. Log rank test comparison of the 3x3 problem and 4x4 showed a p-value of 0.09.
Based on the analysis of completion rate, completion time, and iterations suggest that the difficulty of test problems did not increase monotonically with the problem scale: the 3x3 problem seemed to be more difficult compared to the 4x4 problem.

**Effect of Problem Order**

Since the five test problems were presented to the test subjects in random order, it is possible to evaluate the effect of problem order to subject’s performance. Two different mechanisms may introduce an order effect: the learning effect and the fatigue effect. The learning effect can occur when test subjects became more familiar with the problem dynamics or user interface over time, and the result could be an increase in performance toward the end of the experiment. The fatigue effect is when test subjects grow tired and slow down towards the end of the experiment.

The 10x5 problem was presented to the test subjects as either the 2nd problem or the last problem. Figure 5-8 showed both time and iterations measurements for the 10x5 problem vs the order they were presented to the subjects.

![Box plots](image)

**Figure 5-8:** Order effect, problem 10x5. Boxes bound the 1st and 3rd quartiles, whiskers bound extremes within 1.5 times the interquartile range. Black crosses show subjects who did not finish the problem.

Log rank tests compared the average time and average iterations at different prob-
lem order. The test suggested that there is a statistically significant difference in the time measurement ($p = 0.03$), and slightly less significant in the iterations measurement ($p = 0.05$).

The time vs problem order plots of the four problems from 2x2 to 5x5 were shown in Figure 5-9. The y-axis showed the time required to solve, in units of minutes, and the x-axis showed the relative order the problem was presented to the subject, neglecting problem 10x5.

![Figure 5-9](image)

Figure 5-9: Order effects in time measurements, y-axis shows time to solve in minutes, x-axis shows relative order of problem. Boxes bound the 1st and 3rd quartiles, whiskers bound extremes within 1.5 times the interquartile range. Black crosses show subjects who did not finish the problem.

To evaluate whether the learning effect was significant, the average time of the first problem was compared to the average time of the second and third problem using
Table 5.8: Order effects in time measurements

<table>
<thead>
<tr>
<th></th>
<th>2x2</th>
<th>3x3</th>
<th>4x4</th>
<th>5x5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Learning Effect</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_1(t) - \mu_{2,3}(t)$</td>
<td>0.8 min.</td>
<td>4.8 min.</td>
<td>6.0 min.</td>
<td>6.5 min.</td>
</tr>
<tr>
<td>p-value</td>
<td>0.67</td>
<td>0.36</td>
<td>0.15</td>
<td>0.02</td>
</tr>
<tr>
<td><strong>Fatigue Effect</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_4(t) - \mu_{2,3}(t)$</td>
<td>-0.7 min.</td>
<td>-1.1 min.</td>
<td>0.19 min.</td>
<td>2.4 min.</td>
</tr>
<tr>
<td>p-value</td>
<td>0.41</td>
<td>0.85</td>
<td>0.81</td>
<td>0.14</td>
</tr>
</tbody>
</table>

Table 5.9: Order effects in iterations measurements

<table>
<thead>
<tr>
<th></th>
<th>2x2</th>
<th>3x3</th>
<th>4x4</th>
<th>5x5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Learning Effect</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_1(i) - \mu_{2,3}(i)$</td>
<td>-4</td>
<td>1</td>
<td>43</td>
<td>34</td>
</tr>
<tr>
<td>p-value</td>
<td>0.26</td>
<td>0.53</td>
<td>0.17</td>
<td>0.06</td>
</tr>
<tr>
<td><strong>Fatigue Effect</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_4(i) - \mu_{2,3}(i)$</td>
<td>-0.8</td>
<td>1</td>
<td>-2.5</td>
<td>34</td>
</tr>
<tr>
<td>p-value</td>
<td>0.93</td>
<td>0.72</td>
<td>0.75</td>
<td>0.12</td>
</tr>
</tbody>
</table>

the log rank test. The fatigue effect was explored in a similar fashion by comparing the average of the fourth problem to the second and third problem. The p-values of the tests were shown in Table 5.8.

Only the 5x5 problem showed a significant learning effect: subjects who were given the 5x5 problem first were on average 6.5 minutes slower. The fatigue effect was found to be non existent.

Figure 5-10 showed the iterations vs order of problem for 2x2 to 5x5. The y-axis shows the iterations required to solve, and the x-axis shows the relative order the problem is presented to the subject. The learning and fatigue effects were evaluated using the same method above, and the results are shown in Table 5.9.

Results in Table 5.9 suggested that learning effect was weaker in the iteration measurements compared to the time measurements indicated by higher p-values. Fatigue effect in iteration measurements was also insignificant, similar to time measurements.

The reason that learning effect was weaker in iterations measurements compared to in time measurements is most likely due to test subject’s familiarity with the testing environment. At the beginning of the experiment a test subject was likely to spend
Figure 5-10: Order effects in iterations measurements, y-axis shows iterations taken, x-axis shows relative order of problem. Boxes bound the 1st and 3rd quartiles, whiskers bound extremes within 1.5 times the interquartile range. Black crosses show subjects who did not finish the problem.

more time learning the computer interface, which would reflect in longer time but not more iterations. There could be additional variances in the time measurement unrelated to problem solving, for example using the sliders for input control could be faster compared to typing numbers into a text box. Based on these concerns, future analyses would focus on iterations measurements.

**Knowledge Levels and Performances Ranking**

To calculate an overall ranking of test subjects by their performances, the iterations taken to solve each test problem were first ranked individually. Subjects who were
tied are assigned equal rankings, and subjects who did not complete a problem were considered to be tied for last. Then, a weighted sum of the rankings of individual problems were calculated to determine the overall ranking of each subject. The weights are proportional to the median iterations taken to solve each problem, as indicated in Table 5.7, so that the rankings of more difficult problems were given a higher weight.

In Section 5.3.1 it was mentioned that the 22 test subjects could be divided into 4 mutually exclusive groups based on their technical knowledge level. Table 5.10 shows the overall ranking of the subjects in each group, while Figure 5-11 shows a graphical distribution of subjects’ knowledge level and their performance rankings. The top 11-ranked subjects are considered to belong to the fast group, and the slow group consists the bottom 11-ranked subjects.

Table 5.10: Performance ranking based on technical experiences of subjects

<table>
<thead>
<tr>
<th>Group</th>
<th>Number of Subjects</th>
<th>Rankings of Subjects</th>
<th>Median Ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>strong desalination knowledge</td>
<td>5</td>
<td>1, 3, 4, 5, 11</td>
<td>4</td>
</tr>
<tr>
<td>systems engineering knowledge</td>
<td>4</td>
<td>2, 7, 12, 15</td>
<td>9.5</td>
</tr>
<tr>
<td>some desalination knowledge</td>
<td>7</td>
<td>9, 13, 14, 16, 17, 21, 22</td>
<td>16</td>
</tr>
<tr>
<td>no relevant knowledge</td>
<td>6</td>
<td>6, 8, 10, 18, 19, 20</td>
<td>14</td>
</tr>
</tbody>
</table>

Subjects with strong desalination knowledge ranked higher compared to subjects without strong desalination knowledge (Wilcoxon rank sum test, p-value = 0.03), confirming the expectation that for parameter design problems with technical context, having strong domain knowledge improved subject’s efficiency. Surprisingly, subjects with some desalination knowledge ranked lower compared to the rest of the group (Wilcoxon rank sum test, p-value = 0.05), this suggested that these subjects had enough knowledge to know the concepts of the system, but not deep enough to make good design choices. The next section will focus on the methods and characteristics
Summary of Results Overview

In this section the following experimental measurements: 1) completion rate, 2) completion time, and 3) iterations taken to solve a test problem were presented. It was found that problem difficulty (longer completion time and more iterations) did not correspond to a simple exponential relationship with the number of problem variables as suggested in previous studies, and that the 3x3 problem seemed to be more difficult than the 4x4 problem.

Learning effect was significant on the large-scale problems (5x5 and 10x5), but had a less significant affect on the iterations measurement compared to the time measurement. Fatigue effect could not be measured with statistical significance.

Analysis of subjects' domain knowledge and their performance showed that test subjects with strong desalination knowledge (3 or more years of experiences) ranked
the highest in performance, while subjects with little desalination knowledge (between 0 and 2 years of experience) had the worst performance compared to all test subjects, including people with no prior desalination knowledge.

5.3.4 Detailed Analysis of Strategies

This section describes detailed analyses of some of the strategies and characteristics observed during the experiment. Including subjects' ability to jump from the initial design point to an area close to the target, a phenomenon that was observed where the subjects consistently go away from the target, the effect of step-sizes, and also the effect of changing multiple variables in one iteration.

Distance-to-Target

Figure 5-12: Design space, target region, and problem solving process.

Figure 5-12 is a notional figure of the design space and the components of search used by the participants. Only 2 input parameters are shown but the figure is generalizable to multiple input parameters. The green color dash-line represented the target region containing all possible designs that satisfy all constraints. The designer must find this target region through design iterations. The solid red circle shows the initial design point presented to the test subject, and the hollow red circles represent
the design points at each iteration. The distance-to-target metric is defined as the shortest distance in the design space from the current design point to the target region. This distance was computed post facto based on the experiment log, and thus not presented to the subjects during the experiment. The target region for each test problem was found through latin-hypercube sampling of the design space (300,000 samples for problems 2x2 to 5x5, 3,000,000 samples for 10x5 problem) to find a set of designs that satisfied all constraints, plus the solutions found by all test subjects during experiment.

The Manhattan distance was used for all distance measurements in this study. Manhattan distance was selected over others such as Euclidean distance because most subjects only changed one variable at each iteration, and thus the Manhattan distance was most similar to how the human mind navigate the design space in this particular experiment. Input variables were normalized to values between 0 and 1, using the GUI input range shown in Table 5.3 as normalization factors.

Four sample plots of distance-to-target during experiment were shown in Figure 5-13, which showed the process of solving the 3x3 problem from four different subjects. These plots showed the typical trends observed throughout the subjects and problems. The fast subjects either followed the trend shown in Figure 5-13a and slowly moved toward the target region, or similar to Figure 5-13b by making a large initial jump and then searching around before converging to the target. Subjects who were slower may also take a large initial jump or move slowly toward the target, but the slow subjects were usually observed to go away from the target region in extended and repeated periods, such as the multiple peaks shown in Figure 5-13c and the upward trend observed at the end of Figure 5-13d. It seemed that these subjects were not aware of their relative position from the target region.

**Going Away From Target**

The subject consistently going away from the target region over several iterations was a phenomenon commonly observed in the test subjects. Figure 5-14 illustrates an example of a subject going away from the target region, using screen recordings
Figure 5-13: Sample plots of distance-to-target measurements for the 3x3 problem. y-axis is the normalized distance-to-target measurements, x-axis is time elapsed during experiment in minutes. (a) and (b) show subjects who are relatively faster with red lines, and (c) and (d) show subjects who are slower with blue lines.

from the actual experiment. In Figure 5-14a the subject was already fairly close to the target region but needed to reduce the product TDS from 156ppm to below 150ppm. The subject then increased the first pass recovery ratio over several iterations, which brought product TDS down to 149ppm, but resulted in a high product flow rate that violated its constraint, as shown in Figure 5-14b. The subject then lowered the feed flow rate in the next few iterations, which brought the product flow rate down into the acceptable range, but the product TDS constraint was again violated (Figure 5-14c), with TDS equaled to 173ppm, higher than the value of 156ppm at the beginning.

Seventeen of the 22 subjects were observed to go away from target at least once during the experiment. This behavior was very easily spotted in the 3x3 problem, but could occur in other problems as well. Some subjects would recognize that they went in the wrong direction and then reverse the values, however, many would not recognize this and continued to do it multiple times. There were 2 subjects during the debriefing
Figure 5-14: The phenomenon observed where subjects consistently go away from target. Subject tries to reach the target value of TDS<150ppm.
interview stated that they did not notice "TDS was actually getting higher". Subjects not remembering the previous states of design was a common observation. Several other subjects expressed the desire to have a sensitivity indicator to show which way the parameters have changed between iterations, or a list of designs from previous iterations.

The going-away-from-target phenomenon occurred in the 3x3 problem when product TDS was the only constraint that was still violated. The correct course of action should be to increase feed flow rate and reduce first pass recovery ratio. Most subjects' response at this point would be to first try to decrease TDS to an acceptable level, and deal with other constraints afterward. Increasing first pass recovery, or increasing feed flow rate would both decrease the TDS value, however, most subjects would choose the incorrect action to increase the first pass recovery. A plausible explanation is that people have an incorrect one-to-one mapping of feed flow rate to product flow rate in their mind, since intuition suggest that feed flow rate would affect product flow rate, and this association makes subjects unwilling to change the feed flow rate unless the product flow rate constraint is violated.

There were several possible reasons for why some subjects fail to notice that they were going away from the target region after repeatedly going in the wrong direction. One was that because there were more than 2 input parameters, the other parameters might be a distraction to the subjects that led them to try to modify other parameters. Another was that the subjects may not remember values of output parameters from several iterations ago, possibly because they focused on reducing constraint violations (eliminating "red boxes"), and also the limit capacity of short-term memory [27]. It might also be that subjects had some an incorrect mental model of the system, leading them to de-couple the variables incorrectly, or apply the wrong sensitivity information. This might explain why subjects with some desalination knowledge seem performance the worst.

In short, the phenomenon of going away from target happened when the subject failed to understand the combined effect of two input parameters on the system performance, and also failed to notice their mistake over multiple design iterations.
Table 5.11: Going away from target

<table>
<thead>
<tr>
<th></th>
<th>2x2</th>
<th>3x3</th>
<th>4x4</th>
<th>5x5</th>
<th>10x5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>average tendency of going away</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>minimum</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>mean</td>
<td>9.7%</td>
<td>30.3%</td>
<td>22.3%</td>
<td>32.4%</td>
<td>32.9%</td>
</tr>
<tr>
<td>maximum</td>
<td>44.4%</td>
<td>51.9%</td>
<td>51.0%</td>
<td>50.8%</td>
<td>51.2%</td>
</tr>
<tr>
<td><strong>performance: average iterations taken</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>70</td>
<td>32</td>
<td>96</td>
<td>93</td>
</tr>
<tr>
<td><strong>correlations to performance</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spearman’s $\rho$</td>
<td>0.69</td>
<td>0.79</td>
<td>0.85</td>
<td>0.77</td>
<td>0.89</td>
</tr>
<tr>
<td>p-value</td>
<td>$&lt; 0.001$ for all analyses</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This going-away-from-target phenomenon could be quantified by taking the distance-to-target data log for each test subject, applying a moving-average filter to remove short-term fluctuations, and then calculate the percentage of iterations each subject spent going away from the target.

Table 5.11 shows the compilation of subject’s tendencies to go away from target. For each test problem, the ratio of iterations spent going away from target (tendency of going away) ranged from 0 to 50%, while the averages seemed to follow the total iterations taken for each problem. This was consistent with the observation that the 3x3 problem was more difficult compared to the 4x4 problem. The nature of the 3x3 problem indicated that if a subject did not realize that he/she was going away from target, they would not be able to solve the problem because changing the third parameter would not lower the TDS value enough, whereas in the 4x4 problem, the additional input parameter could have a large impact on TDS, thus allowing more freedom in the design space.

Correlation analyses were performed between the percentage of iterations spent going away and total iterations taken to solve each problem, and found that there was a strong correlation (last two rows of Table 5.11), suggest that subjects who spent more iterations going away from target tend to take more iterations to solve a problem, which was intuitively obvious. Plots of the correlation analyses were shown in Figure B-1 in Appendix B.
**Initial Jump**

Initial jump is the ability to quickly move from the given initial design point to an area of the design space that is close to the target region. This ability depended on the subject’s problem solving strategy, and could also be affected by subject’s prior knowledge. Theoretically, a large initial jump would bring the subject much closer to the target region in very few iterations, and should reduce the total iterations needed to solve a problem.

To test the effect of initial jump, the first 5% of iterations of each problem for every test subject were analyzed. The reductions in distance-to-target after the first 5% of iterations was obtained, and plotted against the iterations taken to solve each problem. Correlation analyses were performed to find any relationships between initial jump and iterations, both Pearson’s product-moment correlation coefficient and Spearman’s rank correlation coefficient were computed and showed consistent results. The results of the correlation analyses were shown in Table 5.12, the plots of iterations vs initial jump for all 5 problems could be found in Figure B-2 in Appendix B.

<table>
<thead>
<tr>
<th></th>
<th>2x2</th>
<th>3x3</th>
<th>4x4</th>
<th>5x5</th>
<th>10x5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spearman’s $\rho$</td>
<td>-0.02</td>
<td>-0.34</td>
<td>0.15</td>
<td>-0.03</td>
<td>0.57</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.92)</td>
<td>(0.12)</td>
<td>(0.52)</td>
<td>(0.88)</td>
<td>(0.01)</td>
</tr>
</tbody>
</table>

Only the 10x5 problem showed signs of correlation between initial jump and performance. This was different from the initial expectation but consistent with the observations of distance-to-target data logs: a test subject could have a good initial jump, but could still perform poorly if the subject consistently went away from the target. The correlation found in the 10x5 problem suggested it may be possible that a "good" initial jump could reduce the number of input parameters that the subject needed to be concerned of.

It was assumed that if an input parameter was manipulated no more than once after the first 5% of iterations, then this input parameter would not be considered as critical to the problem solving process. The number of critical input parameters for
the 10x5 problem were calculated for each test subject, and was plotted against the quality of initial jump in Figure 5-15.

![Graph showing the relationship between initial jump and number of critical parameters.](image)

**Figure 5-15:** Relationship between initial jump and number of critical input parameters, problem 10x5. Black arrow shows a "bad" initial jump (less negative values) correlates to more critical parameters.

There was a correlation between initial jump and number of critical parameters, (Spearman’s correlation coefficient of $\rho = 0.56$, $p$-value = 0.01), suggesting that a good initial jump (more negative values of reduction to distance-to-target) relates to the test subject manipulating fewer input parameters. This result was consistent with the hypothesis stated in the above paragraph.

Figure 5-16 showed subjects’ performances measured in iterations plotted against the number of critical input parameters. A strong correlation (Spearman’s correlation coefficient $\rho = 0.72$, $p < 0.001$) exist between number of critical parameters and total iterations. This was intuitive since reducing the number of variables in a problem could reduce the complexity of a problem.

**Step-Size**

Step-size is the distance between designs from consecutive iterations, shown as red arrows in Figure 5-12. Two sample step-size data logs are plotted in Figure 5-17.
Figure 5-16: Relationship between iterations and number of critical input parameters, problem 10x5. Black arrow shows more critical parameters correlates to longer problem solving.

Figure 5-17: Sample plot of step-size measurements during experiment. A fast subject is shown on the top plot, and a slow subject is shown on the bottom plot.
A trend was noticed in the step-size plots: slow subjects seemed to have inconsistent step-sizes, making lots of large jumps throughout the problem solving process, similar to what is shown on in the bottom plot of Figure 5-17, compared to a more consistent step-size for fast subjects. The kurtosis of the step-size data was calculated for each test subject and compared to the iterations measurement. Kurtosis is the fourth moment of the data and is a measurement of tail-distribution, in this case, it was a measure of how often a subject took very large steps compared to their average step size. The results of correlation analyses between step-size kurtosis and iterations were shown in Table 5.13, plots of the data could be found in Figure B-3 in Appendix B.

Table 5.13: Correlation coefficients between step-size kurtosis (tail distribution) and iterations

<table>
<thead>
<tr>
<th></th>
<th>2x2</th>
<th>3x3</th>
<th>4x4</th>
<th>5x5</th>
<th>10x5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spearman’s ( \rho )</td>
<td>0.72</td>
<td>0.71</td>
<td>0.69</td>
<td>0.25</td>
<td>0.71</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(&lt;0.01)</td>
<td>(&lt;0.01)</td>
<td>(&lt;0.01)</td>
<td>(0.28)</td>
<td>(&lt;0.01)</td>
</tr>
</tbody>
</table>

A strong correlation was found between step-size kurtosis and iterations in 4 of the 5 test problems, with the exception of the 5x5 problem. This suggests that subjects who were slow at solving the design tasks tended to take inconsistent step-sizes. The reason for this could be that taking random large steps during the design process could be a source of confusion for the test subjects, however, it could also be that when a test subject was having trouble finding the target region, it led to the subject making random large jumps to reset input parameters or jump to a different area of the design space. It’s difficult to conclude on the causality of this relationship without further studies.

Another trend that was observed during the experiment and noted in Section 5.3.2 was subjects who make relatively small changes to the input parameters at each iteration tend to take longer to solve a problem. The mean of the step-size data was calculated for each test subject and compared to the iterations measurement. The results of correlation analyses between average step-size and iterations are shown in
Table 5.15, plots of the data could be found in Figure B-4 in Appendix B.

Table 5.14: Correlation coefficients between average step-size and iterations

<table>
<thead>
<tr>
<th></th>
<th>2x2</th>
<th>3x3</th>
<th>4x4</th>
<th>5x5</th>
<th>10x5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spearman’s $\rho$</td>
<td>-0.52</td>
<td>-0.66</td>
<td>-0.57</td>
<td>-0.37</td>
<td>-0.41</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.02)</td>
<td>(&lt;0.01)</td>
<td>(0.01)</td>
<td>(0.1)</td>
<td>(0.07)</td>
</tr>
</tbody>
</table>

Significant negative correlations was found between average step-size and iterations in 3 of the 5 test problems, with the exception of problems 5x5 and 10x5. The negative correlation indicated that subjects taking small step-sizes tend to take more iterations to solve a problem. This was consisted with the observations, and is intuitively obvious: to travel the same distance from the initial design to target region, taking larger steps each iteration means it would take less iterations to find the solution. Another mechanism that could contribute to this correlation was that making small changes to input parameters means the output parameters will also change very little, making it more difficult to notice the sensitivity of the parameters.

**Parameters Changed Each Iteration**

The third characteristic observed was the number of input parameters changed at each iteration. The one-factor-at-a-time design approach was observed to be the most commonly used approach by test subjects. Changing multiple parameters in one iteration could be a potentially faster approach and less likely to arrive at a local optimum [105], but changing multiple parameters at once could be more difficult for the subjects to infer parameter sensitivities, especially when the parameters were coupled.

Test subjects’ tendency to change multiple parameters per iteration could be obtained through analysis of subject’s experiment data to measure the ratio of iterations that 2 or more parameters were changed. On average, a subject changes 2 or more input parameters on only 7.8% of his/her design iterations, ranging from as few as 1.3% for one subject to as much as 16.6% for another subject. Correlation analyses were performed between subjects’ tendencies to change multiple parameters and
number of iterations to solve a test problem, the results were shown in Table 5.15, and plots of the data could be found in Figure B-5 in Appendix B.

Table 5.15: Correlation coefficients between tendency to change multiple parameters per iteration and iterations

<table>
<thead>
<tr>
<th></th>
<th>2x2</th>
<th>3x3</th>
<th>4x4</th>
<th>5x5</th>
<th>10x5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spearman’s $\rho$</td>
<td>-0.22</td>
<td>0.13</td>
<td>-0.23</td>
<td>-0.28</td>
<td>-0.43</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.34)</td>
<td>(0.56)</td>
<td>(0.33)</td>
<td>(0.22)</td>
<td>(0.05)</td>
</tr>
</tbody>
</table>

Based on the results obtained, no significant correlation could be found between subjects’ performances and their tendency to change multiple parameters in one iteration. An alternative approach was taken to compare between subjects from the fast group and slow group and their performance when changing multiple parameters.

For every iteration that 2 or more variables were changed, the change in distance-to-target on that iteration was recorded. The median value of changes in distance-to-target each test subject could achieve was plotted in Figure 5-18, and comparing subjects from the fast and slow groups. Higher negative values indicated the subjects performed better – they were able to move closer to the target region – when changing 2 or more variables.

The results showed that subjects who were faster performed better when changing multiple parameters in one iteration compared to the slower subjects (Wilcoxon rank sum test, p-value = 0.02). This could be an indication that subjects with better understanding of the problem technical context is capable of changing multiple parameters in one iteration to their advantage.

Summary of Design Strategies

In this section three detailed metrics were introduced to analyze strategies used in solving the test problems. The observation of subjects consistently going away from the target design revealed that subjects tended to have an incorrect understanding of how multiple input parameters could affect output parameters. Additionally, subjects’ abilities to make good initial jumps toward the target region were also evaluated, and
found that on large-scale problems, making a good initial jump could lead to solving the design problem faster by reducing the number of critical parameters. It was also found that subjects who were faster at solving the test problems tended to take large and consistent step-sizes between each design iteration, and that they are generally better at changing multiple parameters in one iteration.

5.3.5 Similarities to Simulated Annealing

A surprise finding in this study was how the heuristic optimization algorithm simulated annealing has many similarities to the problem solving process that was used by the subjects. Simulated annealing tries to find an optimal design (minimizing an objective function) by exploring a random design in the neighborhood of the current design at each iteration, and accept the new design with some probability. If the new design has a lower objective function value than the current design, then the probability of acceptance is 1, if the new design has a higher objective function, then the probability is between 0 and 0.5, and is proportional to the "temperature" of the annealing process, which decreases over time. This probability of accepting...
a worse design is analogous to the phenomenon of going away from target observed during the experiments. The difference is that in simulated annealing, the algorithm is always aware that a worse design is being accepted to avoid being stuck in a local optimum, while in the experiment some of the subjects were clearly unaware that they had selected a worse design. The similarities between simulated annealing and the subjects’s problem solving process are listed in Table 5.16.

Table 5.16: Similarities between simulated annealing and human subjects

<table>
<thead>
<tr>
<th>Search Method</th>
<th>Accepting worse design</th>
</tr>
</thead>
<tbody>
<tr>
<td>human process</td>
<td>one factor at a time,</td>
</tr>
<tr>
<td></td>
<td>deliberate/gradient desent/random</td>
</tr>
<tr>
<td>simulated annealing</td>
<td>random search in neighborhood</td>
</tr>
<tr>
<td></td>
<td>deliberate,likelihood reduce</td>
</tr>
<tr>
<td></td>
<td>over time</td>
</tr>
</tbody>
</table>

**Temperature Profile**

The probability of the simulated annealing algorithm accepting a worse design at any iteration is related to the "temperature" of the algorithm, described in the equation below:

\[
P_i = \frac{1}{1 + e^{\frac{\Delta_i}{T_i}}} \quad (5.2)
\]

where \(P_i\) is the probability of accepting a worse design at iteration \(i\), \(\Delta_i\) is the differences between the current iteration’s objective function value and previous iteration’s, and \(T_i\) is the temperature at iteration \(i\). A high temperature value correspond to a higher probability of accepting a worse design, while a low temperature means a low probability of accepting a worse design. Typically the temperature of a simulated annealing algorithm is increased following an exponential cooling schedule. A decreasing temperature profile (cooling schedule) insures that the design space is explored randomly at first, and then focus on improving the design toward the end of
To compare whether the test subjects' search patterns exhibit any characteristics of a simulated annealing algorithm, their "cooling schedules" were estimated based on their likelihood of going away from target (accepting a worse design). Each subject's progress through the design problems was equally divided into 5 segments based on the percentage of overall iterations. The ratios of iterations subjects spent going away from the target was computed for each segment, which is an indication of the likelihood that the subject going to a worse design at each iteration, and can be considered as the "temperature profile" of the subject. The "true temperature profile" of the test cannot be exactly computed since the probability of accepting worse design at each iteration cannot be computed exactly.

![Diagram](image)

**Figure 5-19:** Comparison to simulated annealing: subject's temperature profile. Estimated by the likelihood of accepting worse designs.

Figure 5-19 shows this likelihood through the problem progress. Each line represents a test subject, and the lines are color-coded according to the subject's performance ranking. The average values for the top 11 and bottom 11 ranked subjects are also plotted for ease of comparison.

Figure 5-19 shows that for faster subjects (on the red end of the color map), the likelihood of accepting a worse design drops toward zero as they progress through
the problem, whereas the slower subjects (blue end of the color map) have likelihood values that hover around 0.4 throughout the problem. This is in good agreement with the simulated annealing algorithm, the "temperature" must decrease over time for the algorithm to converge. Table 5.17 shows correlation analysis between the likelihood of accepting a worse design and the subject’s performance ranking at each progress interval. The results confirm the observation: subjects who are faster have cooling schedules similar to simulated annealing.

Table 5.17: Correlation between likelihood of accepting worse design at each progress interval and performance. Positive correlation suggest that higher performance ranking correlates to lower likelihood of accepting worse designs

<table>
<thead>
<tr>
<th>Problem Progress</th>
<th>20%</th>
<th>40%</th>
<th>60%</th>
<th>80%</th>
<th>100%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spearman’s ρ</td>
<td>0.47</td>
<td>0.29</td>
<td>0.21</td>
<td>0.67</td>
<td>0.52</td>
</tr>
<tr>
<td>p-value</td>
<td>0.03</td>
<td>0.19</td>
<td>0.34</td>
<td>0</td>
<td>0.01</td>
</tr>
</tbody>
</table>

**Neighborhood Size**

The size of the "neighborhood" in simulated annealing is analogous to the step-sizes subjects took at each iteration. Different simulated annealing algorithms may take different approaches to defining the size of the design neighborhood. In the original definition of simulated annealing, the neighborhood size is constant during the search, while in practice a variable neighborhood approach is commonly used, by setting neighborhood size proportional to the temperature at each iteration. This approach has been shown to improve the solution time and/or quality [106]

Figure 5-20 shows the average step-size change over the progress of a problem for each test subject. All subjects start with a large step-size in the beginning of problem and decrease as the problem progresses, similar to the variable neighborhood approach used in simulated annealing. The faster subjects tend to have slightly larger step-sizes consistent with the step-size analysis from the previous section, but the shape of the plots look largely the same for all subjects. Table 5.18 shows the correlation analysis between step-size at each progress interval and performance ranking. The results show that larger step-sizes in the early stages of the design process tend to
have a stronger correlation to subject's performance.

Table 5.18: Correlation between step-size at each progress interval and performance. Negative correlation suggest that higher performance ranking correlates to bigger step-size

<table>
<thead>
<tr>
<th>Problem Progress</th>
<th>20%</th>
<th>40%</th>
<th>60%</th>
<th>80%</th>
<th>100%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spearman's $\rho$</td>
<td>-0.74</td>
<td>-0.53</td>
<td>-0.50</td>
<td>-0.33</td>
<td>-0.29</td>
</tr>
<tr>
<td>p-value</td>
<td>0</td>
<td>0.01</td>
<td>0.02</td>
<td>0.13</td>
<td>0.2</td>
</tr>
</tbody>
</table>

**Solving Experimental Task Using Simulated Annealing**

A simulated annealing algorithm is setup to solve the design problems and compare the performance of the algorithm to the performance of the test subjects. The objective function of the algorithm is defined as the sum of the absolute margins of unsatisfied constraints. The algorithm was tuned to stop automatically when the objective function reaches 0, or after 600 iterations. The annealing temperature is assumed to start at 100 and decrease exponentially with a multiplication factor of 0.9 at each iteration. The neighbor function randomly varies one design parameter each time with a predefined mean step-size. A hundred simulated annealing algorithms
were run with the mean step-size varying from 0.01 to 0.4. For simplicity, only the 3x3 problem was solved using simulated annealing and presented below. The total number of iterations vs mean step-size for the simulated annealing algorithm is shown in Figure 5-21

![Figure 5-21: Simulated annealing iterations vs average step-size](image)

The results of simulated annealing searches can be observed to have two distinctive groups. One group contains the majority of simulated annealing (SA) searches, where a solution was found within the 600 iterations, and another group where the algorithm stuck on a local optimum for the entire duration and could not escape from it. A third degree polynomial is fitted to the searches that converged to a solution. The trend suggests that for very small step-sizes, more iterations are required for simulated annealing, and the algorithm is more likely to stuck in a local optimum, which is consistent with the experimental observation. The results also show that if the step-size is too high (above 0.3) there appears to be an increase in the number of iterations required, since too large of a neighborhood size makes the simulated annealing algorithm behave more like a random search [107].

Figure 5-22 compares the iterations taken each test subject to complete the 3x3 problem with the iterations taken the simulated annealing algorithm. Human test subjects tended to take fewer iterations compared to the simulated annealing algo-
Figure 5-22: Iterations vs average step-size, comparing simulated annealing to test subjects

algorithm, which is to be expected, since simulated annealing is based on a random search process. Both the simulated annealing algorithm and the test subjects’ performance confirmed that taking very small steps during the search process would lead to more iterations needed to solve the problem.

This result revealed the fundamental difference between human designer and computer based design tools. Simulated annealing is intended to use computer’s number crunching abilities to find a solution, while for the human subjects in this specific application of reverse osmosis design, the problem is well defined and there is enough knowledge about the system, that experiences and training allows designers to use fewer iterations to find the target solution. It would be interesting to further explore the similarities and differences between human and computer in the future.

Comparison of the subject’s approaches to simulated annealing reveals that the top ranked subjects exhibited characteristics in their strategies that were strikingly similar to fine-tuned simulated annealing algorithms: having relatively large search neighborhood, decreasing temperature profile and neighborhood size. In comparison, the bottom ranked subjects tended to select small search neighborhoods that limited their mobility around the design space, and maintained a constant temperature profile.
that is similar to a random walk.

5.4 Conclusions & Future Work

The work presented in this chapter was motivated by the current design process used in desalination industries, which was similar to the process of solving parameter design problems. Past studies have revealed the inefficiencies of human designers when solving parameter design problems in non-technical context. The objective of this study was to evaluate the effects of different strategies, characteristics, and technical knowledge could have on human designer’s ability to solve reverse osmosis design problems. A set of human experiments was designed and conducted with test subjects recruited based on their experiences in topics of desalination and systems engineering. Five test problems were presented to the subjects with problem scale varied from 2x2 to 10x5. The total number of design iterations taken to solve each problem was used as the performance metrics. Subjects were ranked based on a weighted sum of their performances in each test problem, and it was found that subjects with 3 or more years of experiences working or doing research in the field of desalination were consistently ranked the highest, while subjects with 1 to 2 years of experiences were on average ranked the lowest. This finding shares its underlying message with the adage: A little knowledge is a dangerous thing.

Results of the study revealed a common phenomenon of subjects consistently going away from the target design. This phenomenon was observed when the subject have an incorrect understanding of the sensitivity of two input parameters and changed the input parameters in the incorrect order/direction, therefore making the design worse in the process. This phenomenon had a strong correlation to the number of iterations needed to solve a test problem. Designers should be reminded through training and possibly through the design tool interface to take a system level view of the design constraints, and be aware of the values and sensitivities of parameters.

Making a "good" initial jump in the design space could reduce the number of critical input parameters involved in the problem, and making the problem easier
to solve. A strategy to lower the number of critical parameters is to have a good understanding of the system dynamics such that designers have a good sense of where a "good" design can be found in the design space. This knowledge about the design space can be obtained through training and experiences.

It was also found that taking large and consistent step-sizes between design iterations correlated to solving a problem in less iterations. Subjects used the one-factor-at-a-time design strategy over in 90% of design iterations, however, subjects from the fast group were able to use multiple-parameters-in-one-iteration approaches more effectively compared to subjects from the slow group.

One major conclusion from this study was the importance of understanding how coupled parameters could affect the design output simultaneously. It was observed that human had a tendency to make one-to-one associations between inputs and outputs while ignoring coupling effects, and also ignoring/forgetting the numerical values of output parameters over several iterations. This could be important to the design of design evaluation software, that presenting gradient information to the designer could be very useful to the improvement of designer's efficiency. The results from this study also reveals some limitations of current engineering training, many of the subjects were graduate students in engineering, yet they lacked a sense of system level thinking, indicated by how often they spent design iterations going away from the target design region.

A surprise finding from this work were the striking similarities found between the subjects’ problem solving strategies and the heuristic optimization algorithm simulated annealing. The top ranked subjects exhibited characteristics that were consistent with a well-tuned simulated annealing algorithm, while the bottom ranked subjects are characterized by random walk around the design space.

This study examined in depth the process of designing desalination systems from the point of view of the human designers. The study was motivated by interviews with practitioners that identified an area of real concern. It is the first human subject based experiment that involves a technical context in desalination to understand the strategies that lead to better performance in solving parameter design problems.
Future work would focus on recruiting test subjects with more diversity, especially people with multiple years working in the desalination industry. The time limit applied to the test problems were necessary to manage the duration of the experiment, but in realistic design scenarios the time constraint would usually be on the order of days and months rather than minutes, and thus future study could evaluate how the time constraint might affect a subject's abilities. Another area of future extension is to provide different supplement information to the designer to help them avoid the common pitfalls of parameter design process.
Chapter 6

Conclusions

6.1 Summary

This thesis documented a human centered approach to design research, with a focus on the desalination industry. Desalination system designers were the focus of this research. The study is motivated by the increasing world water scarcity problem and the recent increase of large-scale desalination plants built around the world, which were outlined in chapter 1.

Chapter 2 presented results from four interviews conducted with practitioners experienced in designing both municipal desalination plants and industrial water treatment systems. Two of the themes identified during the interviews – the importance of maintenance and lifecycle costs and the parameter-design approach used by designers. The findings motivated additional work in this thesis, documented in chapters 3 to 5.

Chapter 3 demonstrated a framework for integrating maintenance in the design stage by considering the effect of design parameters on component degradation. A case study of a power plant condenser was used to evaluate the effect of three maintenance strategies on design decisions: fixed maintenance interval, corrective maintenance, and predictive maintenance.

Chapters 4 and 5 describes the development, coordination, and results of a controlled experiment to study designers’ behaviors during parameter design of desalination systems. Chapter 4 described a detailed model of a complete reverse osmosis
plant, that includes a low fidelity intake & pre-treatment system model, a flexible two-pass flow structure, and a lumped-parameter solution diffusion model of the reverse osmosis membranes. Chapters 5 documented the detail design of the experiments. Five test problems were presented during the experiment with varying problem scale, and the total number of design iterations taken to solve each problem was recorded and used as performance metric. Test subjects with technical knowledge ranging from strong desalination knowledge, to some desalination knowledge, to no relevant knowledge, were recruited to participate in the experiments. Results of the study revealed that subjects with strong desalination knowledge performed the best, while subjects with some desalination knowledge performed the worst. Results also indicated that subjects had a difficult time understanding sensitivity of coupled variables, which was the major factor contributing to poor performance. There were also strong similarities between the simulated annealing algorithm and subject's problem solving approaches. Other findings include that a large initial jump toward target solution simplified the problem by reducing the number of critical parameters, taking large and consistent step-sizes between design iterations correlated to better performances.

Overall, this work made three contributions in the fields of desalination research and design research:

1. New approach to design of complex systems that combine designer behavior and system modeling and optimization

2. New framework for considering maintenance at the design stage

3. Identified positive strategies of desalination parameter design, and also potential pitfalls for future designers to avoid.

Figure 6-1 shows how this thesis research relates to the design process illustrated in Figure 2-1, and how the contributions of this thesis improves the current design process. Traditionally, maintenance has been an after-thought for system designers. This work has proposed a design approach that brings maintenance decisions to the preliminary design stage, and integrate them with the design decision making process, in order to find designs that are optimal over its lifecycle.

138
The study of designer behavior improves the detail design process by dissecting the process used by designers when solving parameter design problems and identifying the more effective strategies from the common pitfalls. The results can be used to suggest improvements to design evaluation software and also engineering education.

### 6.2 Suggestions for Future Work

There are several opportunities for future investigation in the maintenance focused approach. Predictive maintenance was assumed to be perfect in this study, however in reality there would be inevitable trade-offs for implementing predictive maintenance, such as the predictive accuracy vs cost trade-off. Future research should consider simulations with different values of prediction uncertainties, and evaluate the sensitivity of prediction accuracy and cost of predictive maintenance. It would also be interesting to conduct a few more case studies with different numbers of degradation...
components to study coupling within components, and different uncertain parameters, such as energy and resource pricing, to evaluate the sensitivity of a design to maintenance policy.

With growing economy and population, the need for desalination plants will continue to be on the rise, and water infrastructure projects of the future must consider the emergent effects of desalination plants on the region and national water supply system. The multi-disciplinary modeling process used to construct the reverse osmosis plant model could be extended to model a desalination network that include multiple desalination plants and power plants that service the water and energy demand of a local community.

Extensions to the human behavior experiments should focus on recruiting test subjects with greater diversity, to further confirm the observations from the experiment. It would be interesting to investigate how different information can be provided to the designer to improve their performance, and develop guidelines for new software design tool that will help designers navigate the complexity and coupling effects in a problem. The effects of time constraint could be further explored to find the differences between the experiment and real world design problem.

Although the results of the experimental study are specific to the reverse osmosis design problem, they suggested that human designers and computer optimization algorithms may have similarities that were previous unidentified. This finding could potentially suggest how computer based design tools should be improved in the future to amplify the strengths of humans and hide their weaknesses. The interaction and trade-off between human designers and computer design tools have not be fully understood, and warrant additional investigation. Future studies should test on a broader collection of engineering applications, as well as non-standard and ill-defined design problems. Another area of future research is to obtain a deeper, systematic understanding of the human designer’s thought process through modeling and simulation, which could inspire novel design automation technologies.

The findings of the experiment also have potential impacts in the training and testing of systems designers in universities and work places. Universities today focus
on the development of systems thinking abilities of engineering and design students. Computer design interfaces, similar to the one used in this study, are great tools to evaluate designers' system level understanding in certain engineering applications. Future implementations of this design interface could be used in universities to test the effectiveness of their engineering education, and also by employers to evaluate designer's abilities in the hiring process.
Appendix A

Designer’s Behaviors Experimental Questionnaire
Background Questionnaire

All questions optional

What is your gender?
- [ ] Female
- [ ] Male

What is your age?
- [ ] <20
- [ ] 21-30
- [ ] 31-40
- [ ] 41-50
- [ ] 51-60
- [ ] 61-70
- [ ] >71

Highest education level
- [ ] high school
- [ ] bachelor
- [ ] master
- [ ] professional degree
- [ ] PhD

Describe your background knowledge in membrane-based desalination (select all that apply)

Number of years working (or research) in the field of desalination: _______
- [ ] My work/research involves membrane-based desalination techniques
- [ ] My work/research involves desalination technologies, but not membrane-based technologies
- [ ] I took classes that involved membrane-based desalination (ie. 2.500 from MIT)
  List related classes: ___________________________________________
- [ ] No background knowledge

Additional Comments: _____________________________________________

Have you used reverse-osmosis software packages in the past? (ie. ROSA from Dow Filmtech, TorayDS from Toray Membranes, ROPRO from Koch Membrane Systems, or equivalent)

- [ ] On a regular basis
- [ ] A few times
- [ ] No

Additional Comments: _____________________________________________

Describe your background knowledge in design processes (select all that apply)

Number of years working (or research) in systems engineering related fields: _______
- [ ] I took classes on numerical optimization (ie. 16.888/ESD.77, 6.255)
  List of related classes: ___________________________________________
- [ ] I took classes on design processes (ie. 2.739, 2.744)
  List of related classes: ___________________________________________
- [ ] No background knowledge

Additional Comments: _____________________________________________
Interview Questions

How difficult were these design problems for you? Easy, moderate, or difficult?

Did you feel there was enough time for you to complete this problem?

Did you feel frustrated during any part of the experiment?

Is this design problem similar to anything you have worked on in the past?

What is your approach to achieve the end result? Do you have a sense of what the values should be?

Did your background knowledge help you with solving the problem?

Do you have any comments and suggestions for us?
Appendix B

Additional Figures
Figure B-1: Iterations vs tendency to go away from target region. x-axis is the ratio of iterations that subject is going away from target region, y-axis is total iterations. Cross marks indicate subjects who did not complete the problem.
Figure B-2: Iterations vs initial jump. x-axis is the reduction in distance-to-target in the first 4 iterations (more negative number represent larger jump), y-axis is total iterations. Cross marks indicate subjects who did not complete the problem.
Figure B-3: Iterations vs step-size kurtosis. x-axis is the step-size kurtosis, a measurement of how often an inconsistently large step is taken (larger values indicate more inconsistent step-sizes), y-axis is total iterations. Cross marks indicate subjects who did not complete the problem.
Figure B-4: Iterations vs average step-size. x-axis is the average step-size, y-axis is total iterations. Cross marks indicate subjects who did not complete the problem.
Figure B-5: Iterations vs tendency to change multiple parameters each iteration. x-axis is the ratio of iterations that 2 or more input parameters are changed, y-axis is total iterations.
Bibliography


[24] Rosa system design software.


