

OPTIMAL SENSOR PLACEMENT PROBLEMS UNDER UNCERTAINTY:
MODELS AND APPLICATIONS

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Todd Zhen

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

December 2019

Purdue University

West Lafayette, Indiana

**THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF DISSERTATION APPROVAL**

Dr. Carl D. Laird, Co-Chair

School of Chemical Engineering

Dr. Gintaras V. Reklaitis, Co-Chair

School of Chemical Engineering

Dr. Joseph F. Pekny

School of Chemical Engineering

Dr. Mario Ventresca

School of Industrial Engineering

Approved by:

Dr. John A. Morgan

Head of the School Graduate Program

To my friends and family, for their support. And to myself, for persevering.

ACKNOWLEDGMENTS

I would like to give my most sincere gratitude to my advisor, Dr. Carl Laird, for the endless support he has provided me. I came into this PhD program directly from my undergrad, was young and reckless, and had no background in optimization. And yet he accepted me with open arms and has never once doubted my abilities, even though I was constantly plagued by self-doubt. He has always gone out of his way to teach me every chance he can, to be a source of encouragement, and a good friend. I am proud to be a student of Dr. Laird. I could not have asked for a better mentor to guide me through this journey.

I would also like to thank my committee members, Dr. Gintaras Reklaitis, Dr. Joseph Pekny, and Dr. Mario Ventresca for their insightful comments and constructive feedback on my work. I am grateful to my co-authors and collaborators at the National Renewable Energy Laboratories (Bri-Mathias Hodge, Tarek Elgindy, S M Shafiqul Alam, Jianhua Zhang), Sandia National Laboratories (Katherine Klise), and Kenexis Consulting (Sean Cunningham, Edward Marszal) for their contributions and support.

I am especially indebted to my fellow group members, Arpan Seth, Yankai Cao, Jianfeng Liu, Michael Bynum, Jose Santiago Rodriguez, and Daniel Laky, and graduate school colleagues, Taufik Ridha and Rick Ray, for being the best friends and coworkers I could have asked for and for making every day a blast. I wouldn't be here without you guys. You have taught me so much and I will always cherish the memories and bonds that we've forged. I wish each of you the absolute best in your futures.

Finally, I would like to thank my friends and family at home who have supported me throughout this journey. Thank you to my friend Henry, for always listening

and providing wisdom; To my friend Simon, for being there to share a laugh; To my friend Christine, for always reminding me of the joys of life; And to my parents, for supporting me every step of the way.

This work was authored in part by Alliance for Sustainable Energy, LLC, the manager and operator of the National Renewable Energy Laboratory for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by the U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Solar Energy Technologies Office. The views expressed in this work do not necessarily represent the views of the DOE or the U.S. Government. We would like to thank the Purdue Process Safety and Assurance Center for their support in this work. We would like to thank the Graduate Assistance in Areas of National Need (GAANN) Fellowship for funding of this work.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energys National Nuclear Security Administration under contract DE-NA0003525.

Disclaimer: This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

Part of this dissertation is reprinted from “Optimizing Placement and Expansion of Resources in Smart Grid Communications Networks” by Zhen, T., Elgindy, T., Alam, S., Hodge, B. M., and Laird, C.D., *IET Smart Grid*, July 2019.

Part of this dissertation is reprinted from “Optimal Placement of Flame Detectors in Petrochemical Facilities” by Zhen, T., Klise, K., Cunningham, S., Marszal, E., and Laird, C.D., accepted in *Process Safety and Environmental Protection*, August 2019.

TABLE OF CONTENTS

	Page
LIST OF TABLES	ix
LIST OF FIGURES	x
LIST OF PUBLICATIONS AND PRESENTATIONS	xiii
ABSTRACT	xv
1 INTRODUCTION	1
1.1 Basic Models and Classification	1
1.1.1 Set Covering Problem	2
1.1.2 Maximal Coverage Location Problem	3
1.1.3 Uncapacitated/Capacitated Facility Location Problem	5
1.1.4 p -Median and p -Center Problems	6
1.2 Placement Under Uncertainty	9
1.2.1 Stochastic and Robust Optimization Placement Models	11
1.3 Applications of Optimal Sensor Placement Methods	12
1.4 Dissertation Outline	15
2 OPTIMAL PLACEMENT OF DATA CONCENTRATORS	17
2.1 Background	19
2.2 Introducing the Smart Grid communication network	22
2.2.1 Home Area Network (HAN)	23
2.2.2 Neighborhood Area Network (NAN)	24
2.2.3 Wide-Area Network (WAN)	25
2.2.4 Defining the network graph	26
2.2.5 QoS Factors	27
2.2.6 Path Loss Propagation Model for Communication Radius	29
2.3 Smart Grid Expansion Problem (SGEP)	31
2.3.1 Redundancy	37
2.4 Case Study	38
2.5 Numerical Results	42
2.5.1 System Resilience	51
2.5.2 Multi-Period Expansion	55
2.6 Conclusions and Future Work	57
3 OPTIMAL FLAME DETECTOR PLACEMENT WITH NON-UNIFORM UNAVAILABILITIES	59

	Page	
3.1	Background	60
3.2	Problem Formulation	66
3.2.1	Basic Flame Detector Formulation (Ignoring Detection Failure)	66
3.2.2	Probabilistic Coverage Formulation (P-FDP)	69
3.2.3	Iterative Solution Algorithm	73
3.3	Numerical Results	75
3.3.1	Results for Case A: Tank Storage Facility	81
3.3.2	Results for Case B: Wellhead Drilling Facility	90
3.3.3	Results for Case C: Offshore Hydrocarbon Separations Module	94
3.3.4	Computational Scalability	98
3.4	Conclusions	100
4	PARAMETER ESTIMATION OF INFECTIOUS DISEASE TRANSMISSION PARAMETERS AND OPTIMAL SELECTION OF DISEASE DETECTION SITES	102
4.1	Background	105
4.2	City-to-Metapopulation Spatial Coupling Parameter Estimation	109
4.2.1	Simulation	109
4.2.2	City-to-Metapopulation Formulation	112
4.2.3	Estimation Results	114
4.3	Simultaneous City-to-Metapopulation with Power Law Model	119
4.4	City-to-City Spatial Coupling Parameter Estimation	121
4.5	Simultaneous City-to-City Spatial Coupling Parameter Estimation with Seasonal Transmission	124
4.6	Optimal Selection of Disease Detection Sites	129
4.6.1	Problem Formulation	131
4.6.2	Selection Results	132
4.7	Conclusions	136
5	SUMMARY	138
5.1	Optimal Placement of Data Concentrators	138
5.2	Optimal Placement of Flame Detectors	140
5.3	Optimal Selection of Infectious Disease Detection Sites	142
5.4	Recommendations for Future Work	143
	REFERENCES	146
A	Derivation of the Stochastic Simulation Model	154
A.1	Nomenclature	154
A.2	Derivation for Binomial Distribution	155
A.3	Derivation for Negative Binomial Distribution	156
A.4	Derivation for Hazard model	157
A.5	Seasonal Transmission Parameters	158

VITA 159

LIST OF TABLES

Table	Page
2.1 Numerical Values for SUI Model Parameters	30
2.2 Reference Test Case A Topology	41
2.3 Parameter Values for SGEP-A/-MM/-R	42
2.4 Numerical Results for SGEP-A, SGEP-MM, and SGEP-R solutions	43
2.5 Numerical Results for SGEP-A, SGEP-MM, and SGEP-R with redundancy $C=2$	56
3.1 Notation for Problem FDP	69
3.2 Notation for P-FDP	71
3.3 Numerical Performance of P-FDP formulation comparing Itera- tive Algorithm and BARON for Case A	87
3.4 Numerical Performance of Iterative Algorithm for Case A with $p_{l,e} \in [0.5, 0.99]$	90
3.5 Numerical Performance of P-FDP formulation comparing Itera- tive Algorithm and BARON for Case B	92
3.6 Numerical Performance of Iterative Algorithm for Case B with $p_{l,e} \in [0.5, 0.99]$	94
3.7 Numerical Performance of P-FDP formulation comparing Itera- tive Algorithm and BARON for Case C	96
3.8 Numerical Performance of Iterative Algorithm for Case C with $p_{l,e} \in [0.5, 0.99]$	98
3.9 Solution Times for Scalability Analysis with average times over 10 realizations for $p_{l,e} \in [0.5, 0.99]$	99
4.1 Numerical Timing Results	121
4.2 Estimates for α parameters in gravity model (954-city dataset, distance boundary of 60 miles)	124
A.1.1 Description of model variables and parameters	154

LIST OF FIGURES

Figure	Page
2.1 Network Diagram	26
2.2 RTCA network diagram.	41
2.3 Distribution of connections per data concentrator for SGEP-A solutions with budget limit (a) $K=8$, (b) $K=25$, and (c) $K=50$	45
2.4 Distribution of connections per data concentrator for SGEP-R solutions with budget limit (a) $K=8$, (b) $K=25$, and (c) $K=50$	46
2.5 Distribution of connections per data concentrator for SGEP-MM solutions with budget limit (a) $K=8$, (b) $K=25$, and (c) $K=50$	47
2.6 SGCN solution for 8 newly-installed data concentrators with (a) SGEP-A, (b) SGEP-MM, and (c) SGEP-R formulations, respectively. Black links denote the underlying power line feeder. Red links denote the connections from smart meters to data concentrators, and from data concentrators to the edge router.	50
2.7 SGCN solution with (a) SGEP-A, (b) SGEP-MM, and (c) SGEP-R formulations for 20 newly-installed data concentrators (redundancy $C=2$)	55
2.8 Comparison of SGEP-MM/-R vs SGEP-MM/-R with Multi-Period Expansion	57
3.1 Flame detector layout example	67
3.2 Linear under-estimators	75
3.3 Overhead view of Case A	77
3.4 Isometric view of Case B facility with 24 candidate flame detector locations	79
3.5 Isometric view of Case C facility with 81 candidate flame detector locations	80
3.6 Optimal placement results for FDP formulation for Case A with budget $k=10$	83
3.7 Optimal placement results for P-FDP formulation for Case A with budget $k=10$ and $p_{l,e}=0.7$	83

Figure	Page
3.8 Expected coverage of FDP assuming perfect detection vs. FDP/P-FDP under imperfect conditions for Case A with (a) FDP assuming perfect detection ($p_{l,e}=1.0$) vs. imperfect conditions ($p_{l,e}=0.7$) and (b) Expected coverage of FDP vs. P-FDP under imperfect conditions ($p_{l,e}=0.7$)	84
3.9 Percentage improvement of P-FDP solutions against FDP solutions for Case A for probabilities of successful detection $p_{l,e}=0.5, 0.7, \text{ and } 0.9$	85
3.10 Percentage improvement of P-FDP solutions against FDP solutions for Case A for probabilities of successful detection $p_{l,e} \in [0.5, 0.99]$	89
3.11 Percentage improvement of P-FDP solutions against FDP solutions for Case B for probabilities of successful detection $p_{l,e}=0.5, 0.7, \text{ and } 0.9$	91
3.12 Percentage improvement of P-FDP solutions against FDP solutions for Case B for probabilities of successful detection $p_{l,e} \in [0.5, 0.99]$	93
3.13 Percentage improvement of P-FDP solutions against FDP solutions for Case C for probabilities of successful detection $p_{l,e}=0.5, 0.7, \text{ and } 0.9$	95
3.14 Percentage improvement of P-FDP solutions against FDP solutions for Case C for probabilities of successful detection $p_{l,e} \in [0.5, 0.99]$	97
3.15 Normalized solution time versus grid size with average times over 10 realizations for $p_{l,e} \in [0.5, 0.99]$	100
4.1 SIR compartment-based model	109
4.2 Comparison between simulated data and case data for Cambridge, UK (population 91,300)	112
4.3 Distribution of estimated spatial coupling parameters on 100 simulations for $c_j=10000$ (Example: Birkenhead, population 141,550)	115
4.4 Spatial coupling estimation for the 954 urban communities in England and Wales with reporting fraction $r=0.55$ and initial susceptible fraction $s_0=0.04$	116
4.5 Sensitivity analysis on reporting fraction r	118
4.6 Sensitivity analysis on initial susceptible fraction s_0	119
4.7 Comparison of linear regressions between estimated parameters (in blue) and findings in (Bjørnstad and Grenfell, 2008) (red line) for reporting fraction $r=0.55$ and initial susceptible fraction $s_0=0.04$	121

Figure	Page
4.8 Comparison of Seasonal Transmission Parameter Estimates for normalized β over time (biweeks) between values estimated in (Xia et al., 2004) and (Finkenstädt et al., 2002) with our results for the 60-city and 954-city datasets	126
4.9 Comparison of outbreak frequency between case data for Cambridge (est. population 105,950) and simulated outbreaks (10 simulations) using spatial and seasonal parameter estimates	128
4.10 Comparison of outbreak frequency between case data for Teignmouth (est. population 10,540) and simulated outbreaks (10 simulations) using spatial and seasonal parameter estimates	129
4.11 Optimal Selection Scheme for TNC impact metric with detection threshold level of 12 cases and budget $K=6$	134
4.12 Objective with TNC impact metric versus detection threshold level for levels from 10 to 100 cases	135
4.13 $\log_{10}(\text{Objective})$ with TNC impact metric versus budget limit from $K=1$ to $K=60$	136
A.5.1 Seasonal Transmission Parameters over 1 year	158

LIST OF PUBLICATIONS AND PRESENTATIONS

Publications

- Zhen, T., Elgindy, T., Alam, S., Hodge, B. M., and Laird, C.D., “Optimizing Placement and Expansion of Resources in Smart Grid Communications Networks”, *IET Smart Grid*, published July 2019. <https://doi.org/10.1049/iet-stg.2019.0006>
- Zhen, T., Klise, K., Cunningham, S., Marszal, E., and Laird, C.D., “Optimal Placement of Flame Detectors in Petrochemical Facilities”, accepted in *Process Safety and Environmental Protection*, August 2019.
- Zhang, J., Hasandka, A., Wei, J., Alam, S., Zhen, T., Elgindy, T., Florita, A., Mohagheghi, S., and Hodge, B. M., “A Hybrid Communications System Architecture Design Framework for Distributed Energy Resources Coordination”, draft completed.
- Zhen, T., Cummings, D.A., and Laird, C.D., “A Nonlinear Programming Framework for Estimating Spatial Coupling and Seasonal Transmission Parameters in Disease Transmission”, draft completed.

Conference Presentations

- Zhen, T., and Laird, C.D., “A Mixed-Integer Programming Framework for the Optimal Placement of Flame Detectors in Chemical Processing Facilities”, FO-CAPD 2019, Copper, CO, July, 2019.
- Zhen, T., Klise, K., and Laird, C.D., “Optimal Placement of Flame Detectors in Petrochemical Facilities”, AIChE Spring Meeting & 15th Global Congress on Process Safety, New Orleans, LA, March, 2019.

- Zhen, T., and Laird, C.D., “A Nonlinear Programming Framework for Estimating Spatial Coupling and Seasonal Transmission Parameters in Disease Transmission”, AICHE Annual Meeting, Pittsburgh, PA, October, 2018
- Zhen, T., Klise, K., Nicholson, B., and Laird, C.D., “A Mixed-Integer Programming Framework for Combined Placement of Fire and Gas Detectors in Chemical Processing Facilities”, AICHE Annual Meeting, Pittsburgh, PA, October, 2018.
- Zhen, T., Elgindy, T., Alam, S M., Hodge, B. M., and Laird, C.D., “Optimal Resource Expansion and Placement in Smart Grid Communications Networks”, AICHE Annual Meeting, Minneapolis, MN, November 2017.
- Zhen, T., Watson, J. P., and Laird, C.D., “A Nonlinear Optimization Approach for Estimating Spatial Transmission Parameters in Infectious Disease Spread”, ICS 2017: INFORMS Computing Society Conference, Austin, TX, January, 2017.
- Zhen, T., and Laird, C.D., “A Nonlinear Programming Framework for Estimating Spatial Coupling in Disease Transmission”, AICHE Annual Meeting, San Francisco, CA, November, 2016.
- Zhen, T., and Laird, C.D., “A Nonlinear Optimization Approach to the Estimation of Spatial Transmission Parameters in Infectious Disease Spread”, AIChE Annual Midwest Conference, Chicago, IL, March, 2016.

ABSTRACT

Zhen, Todd PhD, Purdue University, December 2019. Optimal Sensor Placement Problems Under Uncertainty: Models and Applications. Major Professor: Carl D. Laird.

The problem of optimally placing sensors can often be formulated as a facility location problem. In the literature of operations research, facility location problems are mathematical optimization problems where one or more facilities must be placed in relation to a given number of demand points or customers. Within the context of sensor placement, for example, this translates to placing wireless communication nodes that connect to a set of users or placing smoke detectors to adequately cover a region for safety assurances. However, while the classical facility location problem has been extensively studied, its direct applicability to and effectiveness for the optimal sensor placement problem can be diminished when real-world uncertainties are considered. In addition, the physics of the underlying systems in optimal sensor placement problems can directly impact the effectiveness of facility location formulations. Extensions to existing location formulations that are tailored for the system of interest are necessary to ensure optimal sensor network design.

This dissertation focuses on developing and applying problem-specific optimal sensor placement methods under uncertainty in sensor performance. With the classical discrete facility location problems as a basis, our models are formulated as mixed-integer linear and nonlinear programs that, depending on the specific application, can also be in the form of a stochastic program, a robust optimization framework, or require probability distributions for uncertain parameters. We consider optimal placement problems from three different areas, particularly the optimal placement of data concentrators in Smart Grid communications networks, the optimal placement of

flame detectors within petrochemical facilities, and the optimal selection of infectious disease detection sites across a nation. For each application, we carefully consider the underlying physics of the system and the uncertainties and then develop extensions of previous sensor placement formulations that effectively handle these qualities. In addition, depending on the degree of nonlinear complexity of the problem, specific relaxations and iterative solution strategies are developed to improve the ability to find tractable solutions. All proposed models are implemented in Pyomo, a Python-based optimization modeling language, and solved with state-of-the-art optimization solvers, including IPOPT, Gurobi, and BARON for nonlinear, mixed-integer, and mixed-integer nonlinear programs, respectively. Numerical results show that our tailored formulations for the problems of interest are effective in handling uncertainties and provide valuable sensor placement design frameworks for their respective industries. Furthermore, our extensions for placement of sensors under probabilistic failure are appropriately general for application in other areas.

1. INTRODUCTION¹

Sensor placement strategies have been extensively studied in areas ranging from structural health monitoring to leak detection in municipal water networks to wireless sensor networks. The strategies can generally be divided into two types: (i) exact methods based in mathematical programming, or (ii) heuristic-based methods including proprietary algorithms and placement rules. In this dissertation, we will focus on reviewing methods from the former, and present frameworks by which sensor placement is handled for various applications.

We begin by discussing basic models for sensor placement, including the Set-Covering Problem and Maximal Covering Location Problem, as well as expanded approaches such as the p -Median, p -Center, and Capacitated Facility Location Problems in Section 1.1. We describe methods for consideration of uncertainty in facility location problems in Section 1.2. The contributions of the dissertation are then described in Section 1.3, followed by an outline in Section 1.4.

1.1 Basic Models and Classification

Optimal sensor placement formulations, within the context of mathematical programming, can be traced back to the classical facility location problem. The facility location problem consists of determining the “best” location for one or several facilities or equipments to serve a set of demand points. The facility location problem has been studied in the fields of operations research, location science, and optimization, among others. We refer the reader to the book by Laporte et al. (2015) which

¹Part of this chapter is reprinted from “Optimizing Placement and Expansion of Resources in Smart Grid Communications Networks” by Zhen, T., Elgindy, T., Alam, S., Hodge, B. M., and Laird, C.D., *IET Smart Grid*, July 2019.

provides comprehensive details on the facility location problem and its corresponding formulations, as well as on various applications and extensions across research domains.

The optimal sensor placement problem, like the facility location problem, includes a large family of model formulations that vary based on the physics of the underlying system. For example, a detection sensor (as in a smoke alarm or contamination detector) may only need to send signal in one direction and, in the case of binary detection, require little to no information post-processing. However, wireless sensors, within the context of wireless sensor networks, are required not only to detect and serve demand points, but may also relay, process, and/or fuse information between other sensors. An optimal sensor placement formulation that serves well for a problem within one context may not be applicable for one in a different context. In general, however, optimal sensor placement models can be described by the class of facility location problem formulations, and extensions can be made depending on the underlying physics of the system. In the following sections, we describe the relevant problem classes within the field of location science, and their applicability to optimal sensor placement problems.

1.1.1 Set Covering Problem

In general, two types of covering problems are (i) the Set-Covering Problem (SCP), where full service to demand points is required, and (ii) the Maximal Covering Location Problem (MCLP), where coverage is maximized. The first mention of covering problems in literature can be traced back to (Hakimi, 1965) where the problem of where to distribute a number of policemen to patrol a highway was investigated. However, the problem was first mathematically formulated by (Toregas et al., 1971) within the context of facility location. The classical set covering problem involves minimizing the total cost of placing a number of facilities such that every demand

point or customer is covered. The corresponding mathematical formulation is defined as follows.

$$\underset{y}{\text{minimize}} \quad \sum_{i \in I} y_i \quad (1.1a)$$

$$\text{subject to} \quad \sum_{i \in I_n} y_i \geq 1 \quad \forall n \in N \quad (1.1b)$$

$$y_i \in \{0, 1\} \quad \forall i \in I \quad (1.1c)$$

where I is the set of candidate locations i where facilities can be placed, N is the set of all demand points n that must be covered, I_n is the set of candidate locations that can cover demand point n , and y_i is a binary variable indicating whether a facility is placed at location i . The objective function minimizes the total number of facilities that are placed. The first constraint ensures that all demand points are covered by one or more facilities. The second constraint denotes y_i as a binary variable.

In the context of sensor placement, each facility in this case would instead be a sensor with coverage or servicing capabilities. For example, a smoke detector with a fixed radius of coverage, a visual camera with a field-of-view, or a wireless sensor module with a communication radius. The SCP can also easily be expanded to specify a level of coverage, where a cost is associated with covering each element. In addition, the coverage constraint can be expanded to require each element or demand area (in the case of facility location) be served by two or more sensors based on a minimum percent of covered area. Farahani et al. (2012) provides an exhaustive updated survey of literature on covering problems in facility location and their various extensions. Wang (2011) details a number of coverage models for sensor networks, including a review of coverage metrics and models for point, area, and barrier coverage problems.

1.1.2 Maximal Coverage Location Problem

The maximal coverage problem (MCLP) was first introduced by (Church and Revelle, 1972) and describes maximizing the coverage of target demand subject to

distance constraints. The MCLP is a natural extension of the SCP where the coverage constraint is relaxed and coverage is maximized subject to a fixed budget of facilities.

Extending the SCP formulation and notation from the previous section, the MCLP formulation is defined as follows.

$$\begin{array}{ll} \underset{x,y}{\text{maximize}} & \sum_{n \in N} x_n \end{array} \quad (1.2a)$$

$$\text{subject to} \quad \sum_{i \in I_n} y_i \geq x_n \quad \forall n \in N \quad (1.2b)$$

$$\sum_{i \in I} y_i = P \quad (1.2c)$$

$$y_i \in \{0, 1\} \quad \forall i \in I \quad (1.2d)$$

$$0 \leq x_n \leq 1 \quad \forall n \in N \quad (1.2e)$$

where x_n is a variable indicating whether demand point n is covered by the network of facilities (or sensors) and P is a parameter denoting a limit on the number of facilities that can be placed. Conventionally, the variable x_n is denoted as a binary variable. But in the case of this specific formulation, the binary constraint can be relaxed because the nature of the problem will converge x_n to its bounds, 0 or 1. The objective function maximizes the number of covered or satisfied demand points x_n . The coverage constraint in Equation 1.2b describes the relationship between y_i and x_n , where a demand point n is satisfied only if a corresponding facility i that can cover the demand point is placed. The budget constraint in Equation 1.2c limits the number of placed facilities i .

The MCLP has been widely used for sensor placement problems, especially in the fields of wireless sensor networks and visual surveillance. Bagula et al. (2015) expanded the MCLP for placement of wireless parking sensors within garages to detect presence of vehicles. Sterle et al. (2016) proposed a unified stepwise approach for 2D and 3D coverage problems for omni-directional and unidirectional wireless sensor networks based off the MCLP. Murray et al. (2007) applied the MCLP for surveillance monitoring with visual sensors of a university campus.

1.1.3 Uncapacitated/Capacitated Facility Location Problem

Traditionally, the defining characteristic in Facility Location Problems is to locate a number of facilities across a set of potential locations in order to serve user demand. The placement decisions are generally assumed to be strategic while the demand allocation decisions are considered flexible or operational. In the design of a wired communications network, for example, both placement and allocation decisions would be considered strategic, while in a wireless communications network, the allocation decisions may be operational as changes in allocation do not incur cost. We refer the reader to (Fernández and Landete, 2015a) for an in-depth review of fixed-charge facility location problems. Daskin et al. (2005) reviews the capacitated facility location problem and extensions in supply chain network design, including transportation costs, stochastic scenario-based formulations in literature, and reliability models. Examples and an overview of the fixed charge facility location problems, including capacitated and uncapacitated flow, can be found in (Bertsekas, 1998; Laporte et al., 2015).

The Capacitated Facility Location Problem (CFLP) enforces a constraint on the number of goods or services that each facility can handle (i.e. limit on sensor readings, message size). The Uncapacitated Facility Location Problem (UFLP) is the general form without this limit on capacity. In general, the CFLP is defined below, following the notation convention defined for the previous SCP and MCLP formulations.

$$\begin{array}{ll} \underset{x,y}{\text{minimize}} & \sum_{i \in I} f_i y_i + \sum_{i \in I} \sum_{n \in N} c_{in} x_{in} \end{array} \quad (1.3a)$$

$$\begin{array}{ll} \text{subject to} & \sum_{i \in I_n} x_{in} = 1 \quad \forall n \in N \end{array} \quad (1.3b)$$

$$\sum_{n \in N} g_n x_{in} \leq q_i y_i \quad \forall i \in I \quad (1.3c)$$

$$y_i, x_{in} \in \{0, 1\} \quad \forall i \in I, n \in N \quad (1.3d)$$

where x_{in} is a binary variable indicating whether demand at n is served by facility i , y_i is a binary variable indicating whether a facility is placed at location i , f_i is the

cost of placing facility i , c_{in} is the cost of serving demand point n , g_n is the demand at n , and q_i is the capacity of facility i . The objective function minimizes the total cost of placing a facility as well as the total cost of allocating resources to the demand points. Constraints 1.3b ensure that all demands are served. Constraints 1.3c sets a capacity limit on the demands that can be served by a facility i .

In comparison to the MCLP, the CFLP does not impose limits on the number of facilities that are placed. Rather, the total number of placements are minimized according to the cost objective. In addition, the capacity constraint in Equation 1.3c performs similarly to the coverage constraint in Equation 1.2b in that only if a facility is placed at location i (i.e. $y_i=1$) can demand allocations to facility i exist. However, the capacity constraint also imposes limits on the number of potential allocations or links that can be established between facility i and demand point n .

In the context of sensor placement, the CFLP has been predominantly used for problems in designing wireless sensor networks, where facilities are wireless sensors that have bandwidth-capped communication to service nodes (Qiu et al., 2004; Rowaihy et al., 2007; Mo et al., 2011).

1.1.4 p -Median and p -Center Problems

Presumably independently of each other and concurrently with the development of the Capacitated and Uncapacitated Facility Location Problems, Hakimi (1964) published the seminal paper on the problem of locating facilities such that the sum or the maximum of the distances from all nodes of the network to the facilities is minimized. These two problems have since been known as the p -median and p -center problems, respectively.

Specifically, the p -median problem (PMP) is that of placing p facilities to minimize the demand weighted-average distance between the closest selected facilities and demand nodes (Laporte et al., 2015). A thoroughly annotated bibliography of p -median-type problems can be found in (ReVelle et al., 2008). The p -median problem draws close parallels with the capacitated facility location problems and the maxi-

mal coverage location problems, as we will show in the following formulation, and has historically lied at the heart of many practical location problems. Within sensor placement, the p -median formulation has been applied extensively to problems in optimal gas detector placement (Legg et al., 2012, 2013; Benavides-Serrano et al., 2014; Gomes et al., 2014) and water contamination detector placement (Berry et al., 2005; Hart and Murray, 2010; Krause et al., 2008; Casillas et al., 2013; Krause and Guestrin, 2007). A number of papers have also expanded the p -median problem to consider scenario-based coverage, and a variety of other metrics that are extended from SCP or MCLP formulations (Farahani et al., 2012; Laporte et al., 2015; Karatas et al., 2016).

The general PMP formulation is defined below, following the notation defined for the previously described formulations.

$$\begin{array}{ll} \underset{x,y}{\text{minimize}} & \sum_{i \in I} \sum_{n \in N} w_n c_{in} x_{in} \end{array} \quad (1.4a)$$

$$\begin{array}{ll} \text{subject to} & \sum_{i \in I_n} x_{in} = 1 \quad \forall n \in N \end{array} \quad (1.4b)$$

$$\sum_{i \in I} y_i = P \quad (1.4c)$$

$$x_{in} \leq y_i \quad \forall i \in I, n \in N \quad (1.4d)$$

$$y_i \in \{0, 1\} \quad \forall i \in I \quad (1.4e)$$

$$x_{in} \geq 0 \quad \forall i \in I, n \in N \quad (1.4f)$$

where w_n is the weight of serving demand node n , c_{in} is the cost of serving demand n with facility i , and P is the limit on placed facilities. The objective function minimizes the demand-weighted total cost. Constraints 1.4b ensure that all demand nodes are served, and is equivalent to Constraints 1.3b in the CFLP formulation. Constraints 1.4c define a budget for the number of placed facilities, and is equivalent to Constraints 1.2c in the MCLP formulation. This can also be relaxed to an inequality. Constraints 1.4d ensure that a demand node is served only if a corresponding facility is placed, and is similar to the coverage constraints found in Constraints 1.2b in

the MCLP formulation. Constraints 1.4e require that the placement variables y_i be binary and Constraints 1.4f enforce that the assignment variables x_{in} be non-negative. Interestingly, in comparison to the assignment variables in CFLP, the assignment variables x_{in} here can be relaxed to non-negative values as long as the total demand at node n can be assigned to the nearest placed facility.

The PMP formulation, like the CFLP, shares a minimum objective function, but differ in their solution behavior. Whereas the objective in the PMP formulation is comprised of a single summation for the total or average distance to demand nodes, the CFLP objective includes multiple cost objectives that can lead to competing behaviors, especially if it is formulated as a profit maximization instead. Therefore, the PMP is, in some cases, considered more stable than the CFLP, although both have been studied and used extensively over the past 60 years.

Similar to the PMP, the p -center problem (PCP) is that of placing p facilities to minimize the maximum distance between a demand point and a demand-servicing facility. While the total or average distance minimization in PMP tends to favor clusters, the PCP yields a more spatially-dispersed demand-facility assignment, and has been useful for location problems that need to account for adequate spatial dispersion, such as location of ambulances and police stations. Within sensor placement, this translates to the placement of contamination detectors in water distribution networks that minimize maximum detection times (Berger-Wolf et al., 2005; Watson et al., 2008, 2009; Krause and Guestrin, 2009), or in the design of wireless sensor networks that minimize maximum network hop distance (Youssef and Younis, 2010). Calik et al. (2015) describes in detail the p -center problem, solution methods, variants, and extensions.

The general PCP formulation is an extension of the PMP formulation, and is defined below following the notation previously used for the aforementioned formulations.

$$\underset{x,y}{\text{minimize}} \quad z \quad (1.5a)$$

$$\text{subject to} \quad \sum_{i \in I} \sum_{n \in N} c_{in} x_{in} \leq z \quad \forall n \in N \quad (1.5b)$$

$$\sum_{i \in I_n} x_{in} = 1 \quad \forall n \in N \quad (1.5c)$$

$$\sum_{i \in I} y_i = P \quad (1.5d)$$

$$x_{in} \leq y_i \quad \forall i \in I, n \in N \quad (1.5e)$$

$$y_i \in \{0, 1\} \quad \forall i \in I \quad (1.5f)$$

$$x_{in} \geq 0 \quad \forall i \in I, n \in N \quad (1.5g)$$

where z is an auxiliary variable for the maximum of the distances between the demand points and their facilities. Together with the objective, Constraints 1.5b ensure that the objective value is no less than the maximum of the distances between the demand points and corresponding facilities. The rest of the constraints are equivalent to those of the PMP formulation.

In the next section, we describe how the p -median and p -center formulations can be utilized in designing sensor networks under uncertainty.

1.2 Placement Under Uncertainty

Snyder (2006) provides an excellent and extensive survey of approaches to optimization under uncertainty within the context of facility location. The reader is referred to (Snyder, 2006) for additional details and papers of interest, as only a general overview of the approaches and reasoning are covered in this section.

There is a large body of literature concerning facility location problems with possibly congested facilities, which attempt to capture the uncertain event that a customer may need service from a facility that is occupied with another customer.

Such models are most commonly utilized in siting ambulances or other emergency services. They can attempt to guarantee adequate service either by (i) requiring redundant coverage or (ii) explicitly considering queueing models. Closely related is the branch of literature that considers models in which facilities may be unable to provide service due to facility distributions or link failures.

Snyder (2006) divides decision-making into three categories: (i) certainty, (ii) risk, and (iii) uncertainty. In certainty situations, all parameters are deterministic or known, whereas uncertainty and risk situations involve randomness. In risk situations, uncertain parameters can be characterized by known probability distributions. In uncertainty situations, parameters are uncertain and probability distributions are unknown. Problems concerning risk fall under stochastic optimization, while those concerning uncertainty fall under robust optimization.

The goals of both these optimization classes of problems are to find a solution that will perform well considering the possible realizations of the random parameters, where performance is described by a metric that varies with the application. For example, stochastic optimization problems can generally focus on optimizing an expected value of an objective function, while robust optimization problems will attempt to handle the worst-case performance of the system. In addition, these random parameters can either be described by continuous probability distributions or characterized by a discrete set of scenarios. The scenario approach has two main disadvantages: (i) identification of the scenarios is itself a daunting task, and (ii) a small number of scenarios is favored for practicality and computational handling, but limits the state space of the evaluation. Nevertheless, the scenario approach can be more tractable and has the benefit of allowing statistical dependency of parameters across time and space, which can be difficult to achieve with continuous probability distributions.

In sensor placement, there is also another form of uncertainty arising from the performance of the sensor. This comprises uncertainties such as sensor outages and undetected events, which can be critical in the effectiveness of the sensor network.

While scenarios can also be used in this type of problem, such as specifying outage scenarios, the more common case is to use probability distributions to quantify the uncertainty. Therefore, both probability- and scenario-based approaches can be suitable depending on the goal of the sensor placement problem and nature of the uncertainty involved in the design.

1.2.1 Stochastic and Robust Optimization Placement Models

The stochastic programming (SP) problem consists of first and second stage decision variables. Generally, locations are first-stage variables to be set now, and demand assignments (in the case of facility location problems) are set after the uncertainty has been resolved in the second-stage or recourse decisions.

The most common objective in SP is to optimize the mean (or expected value) outcome of the system, for instance, by minimizing the expected cost. This is typically done using a scenario-based PMP or CFLP formulation, whereby the expected cost, distance, or demand, weighted across scenarios, is minimized. Snyder (2006) survey a large number of papers within this context, including those involving dynamic location, mean-variance models, and chance-constraints.

In contrast to the SP problem, robust optimization (RO) problems typically focus on minimax cost (as in PCP-based formulations), and are used when no probability information is known about uncertain parameters. The primary motivation for this problem is that it does not require the modeler to have prior knowledge of scenario probabilities.

The minimax problem is one that minimizes the maximum cost across all scenarios or demand-weighted costs. Since this type of problem prioritizes handling of the worst-case scenario, it can provide poor solutions for the scenarios that do not have max cost. Thus, minimax objectives are useful for optimizing those applications in which the system must function well in worst-case scenarios.

The literature for this type of minimax problem within facility location is not as large as that of the expected value problem in SP, due to its limited practical

use-cases and solution difficulty. Snyder (2006) describes papers that show general tricks and heuristic algorithms that can be applied for minimax problems, such as setting all uncertain parameters to their upper bounds and solving as a deterministic problem. Exact methods for minimax problems, however, are still in their infancy. Recently, Legg et al. (2013) proposed an optimal gas detector placement formulation under uncertainty that considered Conditional-Value-at-Risk within a stochastic programming framework, and generated placements with better worst-case behavior and fewer scenarios with high detection times when compared with those that minimized mean detection times.

1.3 Applications of Optimal Sensor Placement Methods

In this dissertation, the focus will be on applying and extending a combination of these classical facility location formulations for the design of sensor networks under uncertainty for three different domains: placement of data concentrators within Smart Grid communications networks, placement of flame detectors within petrochemical facilities, and location of infectious disease detection sites within a country. With each of these problems, the classical facility location formulations cannot solve them sufficiently, and model extensions must be developed to handle their unique natures.

Optimal Placement of Data Concentrators for the Expansion of the Smart Grid Communications Network

Evolving power systems with increasing renewables penetration, along with the development of the smart grid, calls for improved communication networks to support these distributed generation sources. Automatic and optimal placement of communication resources within the Advanced Metering Infrastructure are critical to providing a high-performing, reliable, and resilient power system.

The Smart Grid communications network is characterized by modules that collect power usage data from consumers, which then relays the information wirelessly to data concentrators that process the data before sending it to the central utility station.

The problem of placing data concentrators is formulated as a CFLP-based model that aims to maximize the residual network capacity rather than minimize cost.

Three network design formulations based on CFLP and p -center approaches are proposed to minimize network congestion by optimizing residual buffer capacity through placement of data concentrators and network routing. Results on a case study show that the proposed models improve network connectivity and robustness, and increase average residual buffer capacity. Maximizing average residual capacity alone, however, results in both oversaturated and underutilized nodes, while maximizing either minimum residual capacity or total reciprocal residual capacity can yield much-improved network load allocation. Consideration of connection redundancy improves network reliability further by ensuring quality-of-service in the event of an outage. Analysis of multi-period network expansion shows that the models do not deviate significantly from optimal when used in a rolling horizon planning context (within 5% deviation), and are effective for utility planners to use for Smart Grid expansion.

Optimal Placement of Flame Detectors in Petrochemical Facilities Considering Non-Uniform Unavailabilities

Flame detectors provide an important layer of protection for personnel in petrochemical plants, but effective placement can be challenging. A mixed-integer non-linear programming formulation based on the Maximal Expected Coverage Location Problem is proposed for optimal placement of flame detectors while considering non-uniform probabilities of detection failure. We show that this approach allows for the placement of fire detectors using a fixed sensor budget and outperforms models that do not account for imperfect detection. We develop a linear relaxation to the formulation and an efficient solution algorithm that achieves global optimality with reasonable computational effort. We integrate this problem formulation into the Python package, Chama, and demonstrate the effectiveness of this formulation on a small test case and on two real-world case studies using the fire and gas mapping software, Kenexis Effigy.

Parameter Estimation for Infectious Disease and Optimal Selection of Cities for Early Detection of Disease Spread

Highly infectious pathogens can cause localized outbreaks that are followed by regional extinction as susceptible hosts are exhausted. Spatial transmission of these pathogens can once again spark new outbreaks. Study of spatial transmission parameters is essential to predicting and mitigating these outbreaks, as well as implementing strategies for allocating resources for early detection of outbreaks.

We first address the issue of estimating disease transmission parameters by proposing a flexible, scalable modeling framework. Spatial coupling parameters are a measure of the level of mixing of individuals between regional subpopulations. Seasonal transmission parameters describe the impact of seasonal variation due to effects like degree of interaction. This work first focuses on estimating both spatial coupling and seasonal transmission parameters with a nonlinear programming approach to generate an effective model for simulating the spread of disease across a network of subpopulations.

A framework is presented for efficient estimation of parameters in city-to-city spatial transmission models by inferring transport information from localized disease case data using a statistical, hazard-based SIR model. . A statistical hazard-based approach focusing on disease fade-out periods provides the basis for the estimation. Then, the model is extended to simultaneously estimate the city-to-city spatial transmission parameters, as well as seasonal transmission parameters. The proposed approach for this large-scale estimation is readily scalable to larger problem sizes and substantially reduces solution times. The estimation is demonstrated using records of measles outbreaks in 954 cities across England and Wales between 1944 and 1964.

Using the parameter estimates, a stochastic simulator is constructed to model spatio-temporal disease dynamics and generate disease transmission scenarios, we identify optimal locations to place infectious disease detection centers using a stochastic programming, p -median-based formulation. We compare the results between the

minimization of two impact metrics - (i) time until detection and (ii) expected cumulative number of infected individuals before detection.

Throughout this thesis, we demonstrate that facility location formulations are remarkably general and applicable in a variety of sensor placement contexts. Furthermore, as we show, these formulations can be extended to handle physics and constraints arising in new applications while retaining reasonable computational performance, even in a mixed-integer nonlinear programming context. This opens the door for exact methods by extending formulations instead of relying on heuristic simulation-based approaches (e.g. genetic algorithms) that do not provide optimality guarantees.

1.4 Dissertation Outline

The goal of this dissertation is to develop optimal sensor placement models that consider uncertainty for applications across three separate domains. In Section 1.1, we discussed the classical models for sensor placement problems, including the Set Covering Problem, Maximal Coverage Location Problem, Capacitated Facility Location Problem, and p -Median and p -Center Problems. We then showed methods for extending optimal placement formulations to consider uncertainty in Section 1.2.

In Chapter 2, we focus on the optimal placement of data concentrators for the expansion of the Smart Grid communications network. We propose and compare the effectiveness of three mixed-integer linear and nonlinear formulations that optimize placement of data concentrators based on extensions to the Capacitated Facility Location and p -center problems. The proposed formulations attempt to maximize residual capacity in the network in order to reduce chances of network congestion and ensure network reliability through three different ways: (i) maximizing total residual capacity, (ii) maximizing minimum residual capacity, and (iii) minimizing sum of the reciprocal residual capacity. Objective (i) follows the structure of a maximum CFLP-based formulation (maximizing residual capacity rather than minimizing cost), while Objective (ii) is a robust p -center maximin formulation. Objective (iii) represents a

hybrid of the previous two objectives in solution behavior, timing, and effectiveness. We use the proposed formulations to place data concentrators according to a fixed budget as well as impose redundancy requirements to capture the uncertain event that an established data concentrator goes out of service. Chapter 2 concludes with results on multi-period expansion to show the effectiveness of the models for incrementally improving network capacity in a rolling horizon context.

In Chapter 3, we extend past work on the placement of gas detectors within petrochemical facilities to consider the placement of flame detectors with non-uniform unavailabilities. A mixed-integer nonlinear program is proposed based on the Maximal Expected Coverage Location Problem in which the expected coverage of a region-of-interest is maximized, assuming the flame detectors have the possibility of failing to capture combustion events. An iterative solution strategy and corresponding relaxation is also proposed to improve solution tractability and timing for the formulation. The results from this formulation are compared with those from the classical Maximal Coverage Location Problem, where coverage is maximized assuming flame detectors have perfect detection capabilities.

In Chapter 4, the focus is on developing the framework to be able to select optimal locations for infectious disease detection sites. To do so, extensive work is done in building an effective model that accurately represents infectious disease propagation across subpopulations. Using a nonlinear programming approach, spatial transmission parameters are estimated that provide key information for constructing city-to-city disease transmission models. A stochastic simulator is developed to generate scenarios of infectious disease epidemics across a nation-wide scale, specifically for measles. The scenarios represent uncertainties in the centers of origin for epidemics, and are entered into a stochastic p -median formulation where the optimal locations for infectious disease detection sites are selected by minimizing time until detection or total expected cases.

Finally, Chapter 5 concludes this dissertation with a summary and proposed directions for future work.

2. OPTIMAL PLACEMENT OF DATA CONCENTRATORS¹

The adoption of new renewable standards and accelerated cost reductions are driving sharp growth in renewable energy technologies. In particular, the number of distributed solar photovoltaic (PV) installations is growing rapidly. However, as distributed solar power becomes an increasing fraction of total energy generation, the electric power grid must ensure continued reliability and cost-effectiveness. The development of the smart grid, as a natural evolution of the electric power grid, seeks to incorporate new technologies in order to support these distributed generation sources.

New grid devices, such as phasor measurement units (PMUs) and smart meters, will be built to provide the existing Supervisory Control and Data Acquisition (SCADA) systems with vital power flow information, allowing utility operators to monitor and control power flow and maintain operational efficiency. The Advanced Metering Infrastructure (AMI) is designed to measure, collect, process, and relay energy consumption data from these smart meters to utility operators. The AMI is typically comprised of a multi-tiered network of various communication devices and technologies. The AMI is the most fundamental part of the Smart Grid, and can pose a challenging problem to solve in terms of node placement and routing design (Saputro et al., 2012).

While technology selection and hardware capabilities are important to the effectiveness of the Smart Grid, the placement of hardware for AMI applications is equally vital. Manual reconfiguration of network deployments are often impractical since devices and infrastructure are fixed. Multi-tiered AMI networks require sensors and gateways deployed at specific locations to effectively operate. Routing assign-

¹Part of this chapter is reprinted from “Optimizing Placement and Expansion of Resources in Smart Grid Communications Networks” by Zhen, T., Elgindy, T., Alam, S., Hodge, B. M., and Laird, C.D., *IET Smart Grid*, July 2019.

ments between nodes can have significant effects on network quality and performance. Bottlenecks and islanding can occur in the network due to highly congested nodes and poor routing, leading to difficulties serving customers and ineffective system control. Careful planning of node placement and routing design can significantly alleviate these Quality-of-Service (QoS) issues, while at the same time improving performance. Automatic placement and optimization of network topology can be used to provide a cost-effective, high-performing, reliable, and secure communication network in support of the expanding smart grid.

In this work, we develop a procedure to optimally place communication hardware and route communication within AMI networks in a way that best maximizes QoS performance for utility customers, provides redundant network connectivity in the event of a security threat or outage, and provides for effective expansion of the network.

To do so, we first propose three mixed-integer linear programming (MILP) approaches for optimal placement of data concentrators to maximize the network residual buffer capacity while ensuring network connectedness between smart meters and electrical control centers. We apply the models to a case study with empirical parameters and standards. We analyze the network residual buffer capacities of the results, comparing distributions of connections and placement topologies. We then extend the formulations to improve network resilience by imposing connection redundancy, and conclude by testing formulation effectiveness in planning for Smart Grid communication network expansion.

This chapter is organized as follows. In Section 2.1, we summarize and highlight related work. In Section 2.2 we describe the Smart Grid communication network and its supporting technologies, as well as define assumptions about the network that we will use in the optimization formulation. Section 2.3 outlines the mathematical formulations for the smart grid expansion problems. Section 2.4 describes the applied case study in detail. Results of the optimization are presented in Section 2.5.

2.1 Background

In this work, we are focused on the problem of expanding the distribution grid communications network as part of a growing, integrated Smart Grid. For consistency and conciseness, we will simply refer to this problem as the Smart Grid expansion problem.

The Smart Grid expansion problem is twofold; it consists of (i) a facility location problem and (ii) a routing assignment problem. Facility location and routing assignment problems have been studied extensively for applications in supply chain, emergency, pipeline, and telecommunications networks (Daskin et al., 2005; Fernández and Landete, 2015b). We refer to topology (or network) design as the process of placing nodes within a network, and to routing design as the process of link assignment and bandwidth allocation between nodes.

There have been a number of previous studies on the use of mathematical programming approaches for designing wired/wireless mesh networks with regard to QoS, cost, and/or traffic allocation. Kojima et al. (2018) investigated aggregation point placement in a wired SCADA network and proposed a MILP framework involving communication scheduling. Qiu et al. (2004) presented a MILP formulation based on the capacitated facility location problem to optimally place internet transit access points (ITAPs) in a wireless mesh network, while also handling QoS requirements like throughput and fault tolerance. Amaldi et al. (2008) presented a MILP formulation for constructing Mesh Access Points (MAPs) in the form of a multi-commodity network flow problem with consideration to network traffic allocation. Jahromi and Rad (2013) designed a communication topology for meshed node placement and assignments of wired connections within the power communication network, and formulated a multi-objective, nonlinear problem considering cost due to network packet loss, delay, and budget. The previous work on wired networks and mesh networking cannot be directly applied to our context because mesh routing requires different specifica-

tions and characteristics than Smart Grid AMI for distributed power systems, which are typically multi-layered systems with various interacting technologies.

In addition, many related works reported using approximation algorithms and heuristic-based solvers to solve their respective location/routing problems (Amaldi et al., 2008; Qiu et al., 2004; Jahromi and Rad, 2013; Rastgoo and Sattari-Naeini, 2016; Tavasoli et al., 2016; Klinkert, 2018; Hassan et al., 2018), citing high time-complexities when solving NP-hard problems with exact methods. We instead argue that the problems, while complex in nature, can be handled effectively with exact solution methods (i.e. Branch-and-Cut as implemented in state-of-the-art commercial solvers), and the benefits outweigh the costs in doing so. Approximation algorithms are typically fast and specifically tailored to the application, but provide minimal optimality guarantees. On the other hand, exact methods, by nature of the methods themselves, provide measures of optimality. Given that the Smart Grid expansion problem is a planning and design problem, fast solution times may be less of a concern than the advantages of optimality guarantees provided by exact methods.

There is existing work that investigates the optimal placement problem of data concentrators within a smart grid context. Huang and Wang (2015) evaluate a MILP framework for constructing a three-layer Smart Grid AMI that focuses on ensuring network demand while minimizing total cost, and designed an approximation algorithm handle the problem. Klinkert (2018) present a MILP framework for repeater and data collector placement based on fixed-charge network flow within the context of a smart grid communications network. They also consider a multi-tier AMI involving meters, repeaters, data collectors, and a central data center, analogous to the AMI presented in this work. Their MILP formulation also focuses on ensuring network feasibility for a minimum cost placement. Tavasoli et al. (2016) considered the optimal placement of data aggregation points (i.e. data concentrators) in a hybrid wireless and wired communication network by minimizing installation costs of the access points, fiber-optic backbone, and Worldwide Interoperability for Microwave Access (WIMAX) communication costs. Kong (2019) studied the inter-network re-

liability problem concerning data aggregator placement, specifically the problem of where to place data aggregators to ensure communication and power robustness. They proposed a minimum cost objective subject to QoS requirements and network robustness assurances. Each of these related works (Huang and Wang, 2015; Klinkert, 2018; Tavasoli et al., 2016; Kong, 2019) presented models that minimized cost while ensuring network demand is fulfilled. In this work, however, we argue that while optimizing for cost alone is an effective short-term solution for cost-limiting scenarios, it is insufficient to guarantee network performance for the long-term. A more rigorous objective, namely one that prioritizes QoS and network capacity, is required to best ensure network longevity and future performance, especially for planning a long-standing Smart Grid network.

Gourdin and Klopfenstein (2006) has compared different QoS objectives for communication routing design. They condense the resource management solutions for handling QoS factors (i.e. bandwidth, delay, jitter, and packet loss rate) down to two essential criteria: avoiding congestion and limiting path length. They define congestion as the difference between arc capacity and arc flow, and denote this relationship as residual capacity. They also describe three objective functions that can handle such congestion: (i) maximize the minimum arc residual capacity for all arcs in the network, (ii) maximize the total amount of available network bandwidth (also equivalent to maximizing the total/average residual capacity across the network), and (iii) minimize the sum of the reciprocal arc residual capacities. They analyzed the effectiveness of each of these objectives on a wired communication network. While their methods focus on optimal routing of arc flow (splittable) with consideration to minimal communication path length (i.e. number of hops required from source to destination), they did not consider node placement (i.e. network design).

Our application focuses on the distribution system communications network with two-hop communication links and intermediate data concentrator nodes. Due to the nature of data concentration, we do not consider splittable routes. Our approach prioritizes node placement and introduces binary variables, yielding a mixed-integer

linear programming design problem. We focus on optimal handling of aperiodic data traffic by optimizing the residual buffer capacity of data concentrators as an analog to optimizing QoS factors like bandwidth, throughput, packet loss, and delay requirements in the network. We expand upon the objectives proposed by (Gourdin and Klopfenstein, 2006) within a topology design framework, and extend them by incorporation of communication redundancy to improve network reliability. In addition, our models prove effective for multi-period expansion of the communications network. To the best of our knowledge, no prior work has been done on solving the problem of data concentrator placement while prioritizing QoS performance within an AMI network.

In summary, the placement of data concentrators to *maximize network residual buffer capacity under the impacts of budget constraints, network connectivity requirements, distance limitations due to path loss, and redundant communication links to enhance robustness for an expanding Smart Grid communications network* is a unique challenge that we aim to address in this chapter.

2.2 Introducing the Smart Grid communication network

We first describe the Smart Grid communication network (SGCN) in detail and then propose an optimization framework to handle the expansion of the network to ensure QoS guarantees and network connectedness. Specifically, we focus on the advanced metering infrastructure (AMI), which is the data communication architecture between the smart meter and the meter data management system (MDMS). The AMI is used to transfer real-time information from meters, including fault, outage, and usage, to the utility control center, where utility operators will use the information to control power flow and maintain operational efficiency. The AMI typically involves a hierarchical network architecture utilizing a variety of communications technologies (Niyato and Wang, 2012). We will focus on the AMI when referring to SGCN in this chapter.

The general framework of the SGCN consists of a multi-tier system: (i) a Home Area Network (HAN) consisting of smart appliances, Internet-of-Things (IoT) devices, and distributed energy sources communicating to a nearby smart meter (SM), (ii) a Neighborhood Area Network (NAN) consisting of data concentrators (DC) that receive the information from nearby smart meters, process and relay the information onto (iii) the edge router (ER) or a control center of the energy provider within the Wide Area Network (WAN) (Niyato and Wang, 2012; Andreadou et al., 2018; Klinkert, 2018). Edge routers contain secure fiber optic connections to monitoring stations on a SCADA network within the transmission system. A number of communication technologies for the SGCN can be used, and can be divided into wired and wireless types (Kuzlu et al., 2014; Andreadou et al., 2018). For the purposes of this chapter, we will focus on wireless communications technologies.

2.2.1 Home Area Network (HAN)

The HAN is localized to the customer domain and encompasses communication within the typical residential home, including smart appliances, solar panels, light controls, and various other sensors and actuators. Home devices will send power readings, usage, and outages over the HAN to the smart meter outside the home for AMI application. A number of wireless technologies can be used within the HAN, including Bluetooth, ZigBee, and IPv6 over Low-Power Wireless Personal Area Networks (6LoWPAN) (Saputro et al., 2012).

Bluetooth, Institute of Electrical and Electronics Engineers (IEEE) 802.15.1 standard, is a low power, short range radio frequency standard operating on the 2.4 GHz band (Parikh et al., 2010). However, Bluetooth communication distances are short (between 1m - 100m), are highly influenced by IEEE 802.11 networks, and offer weak security compared to similar network standards. 6LoWPAN is a low-power wireless personal area network (WPAN) built on the IPv6 routing protocol, and offers interoperability and QoS guarantees within a HAN. Disadvantages include network security and poor service discovery (i.e. automatically locating other nodes and higher layer

services) (Saputro et al., 2012; Niyato and Wang, 2012). ZigBee (built on the IEEE 802.15.4 standard) was developed by the ZigBee Alliance and is recommended as the common choice for Home Area Network (HAN) communication. ZigBee devices and routing protocols allow for meshed networking within the HAN, and provides QoS guarantees for information relaying across devices. Advantages include low cost, low power consumption, minimal data rates, and wide-spread usage in existing smart home devices. Additional information on ZigBee communications can be found in (Saputro et al., 2012).

2.2.2 Neighborhood Area Network (NAN)

The NAN connects smart meters to data concentrators for AMI applications, and can involve a series of relays and mesh networking involving technologies such as IEEE 802.11s Radio Frequency (RF) Mesh, IEEE 802.11 2.4 GHz Wireless Local Area Network (WLAN), WIMAX, Power Line Communication (PLC), or cellular technologies. Compared to the HAN and WAN, communications technologies can vary significantly within the NAN, and often there is no smart grid-specific standard within this domain. For our purposes, we focus chiefly on wireless communications and will not discuss PLC technologies, although more information can be found in (Saputro et al., 2012).

Mesh networking can present a cost-effective and viable solution for improved network resilience (Xu et al., 2018). However, mesh networking can have significant signal interference and fading effects that make the quality of these connections unreliable, presenting crucial security problems. Methods to address these security issues in mesh routing have been investigated (Saputro et al., 2012). WIMAX, based on the IEEE 802.16 standard, also provides promise with high throughput, low latency, high security standards, and traffic management tools (Patel et al., 2011). However, WIMAX capabilities are typically relegated to specific architectures, and would require constructing a utility-proprietary network with high capital costs.

IEEE 802.11-based 2.4 GHz wireless LAN (WLAN) provides a robust, high speed, point-to-point communication. Its ubiquity allows for easy installation and low costs. Moreover, WLAN technologies are undergoing continuous improvement of data rates, service, and coverage. High reliability can be achieved by proper routing design and system design techniques (Parikh et al., 2010). WLAN is considered an excellent candidate for smart grid systems (Niyato and Wang, 2012), specifically in remote metering and monitoring applications, and has been recognized by the National Institute of Standards and Technology (NIST) via the International Electrotechnical Commission (IEC) 61850 standard for application in smart grid environment (Parikh et al., 2010).

While cellular technologies have been studied for smart meter transmissions, they are generally preferred for links between data concentrators and edge routers (Andreadou et al., 2018).

2.2.3 Wide-Area Network (WAN)

The WAN provides connections between the smart grid and the core utility network. Within the WAN, communication between data concentrators and edge routers can include technologies utilized in the NAN as well (e.g. cellular technologies, WIMAX, RF Mesh or broadband fiber optic connections for wired options (Niyato and Wang, 2012; Saputro et al., 2012)).

Cellular and Long-Term Evolution (LTE) technologies (3G/4G on 824-894 MHz/1900MHz spectrum) are on licensed frequency bands and provide benefits of low interference, high reception rate and information security, extensive data coverage, and no maintenance costs (since operation and network is maintained by carriers) (Parikh et al., 2010). Since utility control centers require high levels of reliability, LTE connections are suitable for WAN communication (Patel et al., 2011). Additionally, the existing infrastructure for cellular communications allows for rapid deployment of Smart Grid communication hardware.

2.2.4 Defining the network graph

The network description addressed in this chapter is summarized in Figure 2.1. The SGCN of interest in this chapter considers wireless communications between PV

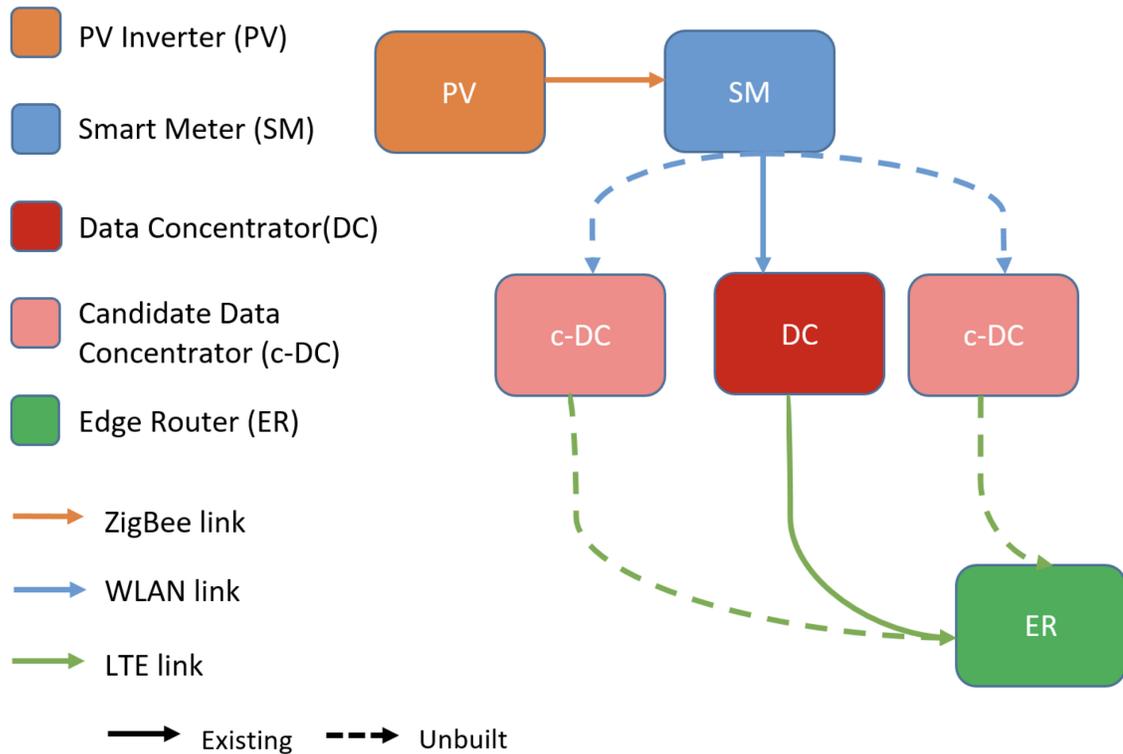


Figure 2.1.: Network Diagram

inverters, smart meters, data concentrators, and edge routers. In addition, we assume PV inverters transmit to smart meters via Zigbee, smart meters transmit to data concentrators via 2.4 GHz WLAN, and data concentrators relay to edge routers via LTE technology. While PV inverters are typically installed close to smart meters, data concentrators and edge routers are installed separately from the PV inverter-smart meter HAN. We do not consider mesh networking or multi-hop communications in this work. Since the network communication is wireless, connections can be established if transmitters and receivers are within communication range of one another, but the assignment of connections will be decided by the optimization formulation described

in Section 2.3. We use the terms “link”, “arc”, and “connection” interchangeably to denote the connections in the network. Specific parameters for the network are provided in Section 2.4. For additional information on Smart Grid technologies and practices, QoS requirements for Smart Grid integration to SCADA systems, and background on advanced metering infrastructure, we refer the reader to (Budka et al., 2014).

2.2.5 QoS Factors

There are several critical QoS factors that affect network performance: packet loss, path loss, effective throughput, network criticality, network availability, available bandwidth, latency, and connection outage probability (Rastgoo and Sattari-Naeini, 2016). For our purposes in providing a time-invariant, deterministic routing assignment and optimal placement of communication nodes, we will consider only packet loss, path loss, effective throughput, available bandwidth, and latency as QoS factors of interest in this work. We consider the time-varying and stochastic QoS parameters (network availability, criticality, and outage probabilities) as future work.

Packet loss is the fraction of transmitted packets lost in the network and commonly occur due to queuing congestion and buffer overflows. Path loss depends on factors such as antenna height, frequency, and link distance. Effective throughput, the amount of data successfully passed through a link at a given time, is highly dependent on the processing within each network router to determine the transmission rate in order to forward the packet with a minimum number of hops (Rastgoo and Sattari-Naeini, 2016). Since packet loss and effective throughput are highly dependent on the effectiveness of underlying network technologies and routing protocols (Qiu et al., 2004), we choose to focus on *path loss*, *available bandwidth*, and *latency* as the principal QoS factors when placing network hardware and assigning network routes.

Latency, or network delay, is defined as the time needed for a bit of data to travel from one node to another. Latency can be divided into four parts: (i) processing delay

- the time taken for a node to process the packet header, (ii) queuing delay - the time a packet spends in routing queues before being processed, (iii) transmission delay - the time taken to push the packet's bits onto the link, and (iv) propagation delay - the time for a signal to reach its destination. The propagation delay is dependent on the distance and geography between communication nodes, while the majority of latency (processing, queuing, and transmission delays) is dependent on the communication hardware and wireless routing protocols (Nagai et al., 2013).

In the Smart Grid Expansion Problem, propagation delay is the relevant QoS metric of interest within latency, while processing, queuing, and transmission delays are handled by the hardware and wireless technologies (Park et al. (2019) has recently developed a low-latency congestion control algorithm for wireless LTE networks). Lopez-Aguilera et al. (2010) has shown that the propagation delay for outdoor IEEE 802.11 wireless networks is dependent on the slot time and translates up to 6 km communication range for IEEE 802.11b without performance degradation. For larger distances, the propagation delay becomes greater than the slot time and performance is degraded.

For our proposed approaches, we ensure path loss and propagation delay in the network are within QoS requirements by defining the effective communication range of wireless technologies to within limits that ensure minimal performance degradation (refer to Section 2.2.6). Once the communication ranges are established, the optimization can then consider optimal placements for wireless hardware and routing decisions by maximizing available network bandwidth.

Following the convention of (Gourdin and Klopfenstein, 2006), we refer to available network bandwidth as residual capacity, specifically on the capacities of buffers at each data concentrator. Each data concentrator has a buffer size that can handle incoming messages. Exceeding this buffer limit negatively impacts the packet loss rate, and increases communication delay (Gourdin and Klopfenstein, 2006). Therefore, routing assignments are made by maximizing residual capacities across the network.

2.2.6 Path Loss Propagation Model for Communication Radius

The maximum allowable distance between nodes (communication radius) can be predicted using a minimum threshold of received power required for signal reception and a model for signal propagation. The Stanford University Interim (SUI) propagation model (Abhayawardhana et al., 2005; Hari and Smith, 2003) is chosen to represent the 2.4 GHz WLAN link between smart meters and data concentrators in our case study.

The SUI models are specified for three types of terrain (A, B, and C), where Type A is associated with hilly terrain and moderate to heavy foliage with maximum path loss, Type B is associated with either mostly flat terrain with moderate to heavy tree densities or hilly terrains with light tree densities, and Type C is associated with flat terrain and light tree densities with minimum path loss. More information and associated parameters on the SUI models can be found in (Abhayawardhana et al., 2005).

Given that the case study we will be considering in Section 2.4 is representative of a typical suburban area, we have elected to use the terrain Type B specification, although the proposed approach is general and can be used with Type A and Type C terrains. The path loss equation with correction factors and other parameters are defined below,

$$P = X_g + 10\gamma \log_{10} \left(\frac{d}{d_0} \right) + X_f + X_h + s \quad (2.1a)$$

$$X_g = 20 \log_{10} \left(\frac{4\pi d_0}{\lambda} \right) \quad (2.1b)$$

$$\gamma = \alpha - \beta h_b + \frac{\omega}{h_b} \quad (2.1c)$$

$$X_f = 6.0 \log_{10} \left(\frac{f}{2000} \right) \quad (2.1d)$$

$$X_h = -10.8 \log_{10} \left(\frac{h_r}{2000} \right) \quad (2.1e)$$

where P is the path loss (defined as the ratio of the transmitted to received power), X_g is the free space or line-of-sight (LOS) gain, d is the distance between the transmit-

ter (smart meter) and the receiver (data concentrator) antennas, d_0 is the reference distance ($d > d_0$), s is a shadow fading factor, h_b is the base station height above ground, γ is the path loss exponent, X_f and X_h are correction factors for the operating frequency and the antenna height, respectively, f is the operating frequency of the transmission signal, λ is the wavelength from the communication frequency, and h_r is the receiver antenna height above ground. The terrain constants α , β , and ω used in our study are given in Table 2.1, along with ranges and values for the other parameters in the path loss equation.

Table 2.1.: Numerical Values for SUI Model Parameters

Model Parameter	Value
d_0 (m)	100
s (dB)	8.2 - 10.6
h_b (m)	10 - 80
f (MHz)	2400
h_r (m)	6
α	4.0
β (m^{-1})	0.0065
ω (m)	17.1
λ (m)	0.125

Bounding Equation 2.1a for maximum path loss and expanding P_{max} , we have

$$P_{max} = P^{trans} - P_{min}^{rec} \quad (2.2)$$

where P_{max} is the maximum path loss incurred when the received power is at its minimum, P^{trans} is the transmitted power (in dBm) and P_{min}^{rec} is the minimum received power (in dBm). From Equations 2.1a-2.1e, we obtain a relationship for the communication distance in terms of P_{max} ,

$$d \leq d_0 10^{\frac{P_T}{10\gamma}} \quad (2.3)$$

where $P_T = P_{max} - X_g - X_f - X_h - s$. Smart meter transceivers can operate at peak transmitted power up to 1 W (30 dBm), but typically averages low transmitted power since the transmission intervals are very short and infrequent (GSMA, 2015). We therefore assume average transmitted powers of $P^{trans} = 0.2$ W (23.0 dBm) for our test case (Rupasinghe and Guvenc, 2014). Minimal received signal power for wireless networks (IEEE 802.11 variants) are typically $P_{min}^{rec} = -100$ dBm. Combining these power specifications leads to $P_{max} = 123$ dB. Therefore, the smart meter communication radius, in accordance with the SUI propagation model parameters in Table 2.1, is $d \leq 0.93$ km with $s = 9.0$ and $h_b = 10$, in order to remain within minimal received signal power range for all receivers. In addition, since this communication range for 2.4 GHz WLAN is less than the reported distance in order to ensure minimal propagation delay (i.e. 6 km according to (Lopez-Aguilera et al., 2010)), setting this as the distance limit between smart meters and data concentrators will ensure acceptable QoS in the optimal design.

2.3 Smart Grid Expansion Problem (SGEP)

A larger number of smart meters will eventually be constructed as the Smart Grid grows, distributed generation becomes more widespread, and an increasing number of IoT devices become commonplace. The Smart Grid expansion problem focuses on placement of additional data concentrators (by selecting from a set of candidates denoted in Figure 2.1 as c-DC) and assignment of communication links between smart meters and data concentrators in order to handle the load on the network, while minimizing congestion and improving network connectivity and robustness. We consider the residual buffer capacity of the network as the key metric to define congestion levels in the network.

In this section, we show three formulations that focus on improving network residual buffer capacity within a mixed-integer linear program to form the Smart Grid Expansion Problem (SGEP).

The Smart Grid Expansion Problem - Average (SGEP-A) focuses on maximizing average residual buffer capacity and is defined by the following mixed-integer linear program,

$$\begin{array}{ll} \text{maximize} & \sum_{l \in L} r_l \\ & x, y \end{array} \quad (2.4a)$$

$$\text{subject to} \quad \sum_{l \in L} y_l \leq K \quad (2.4b)$$

$$x_{a,l} \leq y_l \quad \forall l \in L, a \in \mathcal{A}_l \quad (2.4c)$$

$$\sum_{l \in \mathcal{L}_a} x_{a,l} = 1 \quad \forall a \in A \quad (2.4d)$$

$$r_l = b_l y_l - \sum_{a \in \mathcal{A}_l} x_{a,l} f_a \quad \forall l \in L \quad (2.4e)$$

$$y_l \leq \sum_{a \in \mathcal{A}_l} x_{a,l} \quad \forall l \in L \quad (2.4f)$$

$$y_l, x_{a,l} \in \{0, 1\} \quad \forall l \in L, a \in \mathcal{A}_l \quad (2.4g)$$

$$r_l \geq 0 \quad \forall l \in L \quad (2.4h)$$

where A is the set of all smart meters that must be serviced, L is the set of data concentrators (existing and candidate), \mathcal{L}_a is the set of candidate data concentrators within communication range of smart meter a , as determined by the path loss model from Equation 2.3. Here, \mathcal{A}_l is the set of all smart meters that are within the communication range of candidate data concentrator l (i.e. the distance between smart meter a and data concentrator l is less than or equal to d defined in Equation 2.3). The binary variable y_l indicates if a data concentrator at location l is selected for installation, binary variable $x_{a,l}$ indicates if smart meter a is assigned a connection to data concentrator l , and r_l is a variable defined as the residual buffer capacity for data concentrator l . Parameters f_a denotes the flow of data or amount of data to be processed daily (depending on units of buffer capacity) from smart meter a , K defines

the limit on total allowable number of data concentrators in the network (includes existing data concentrators), and b_l is the buffer capacity on data concentrator l .

Problem Formulation 2.4a-2.4h seeks to select the set of data concentrators that should be built (y_l) along with the connection routing that maximizes the residual capacity of the constructed network. Constraints 2.4b limit the total number of newly installed data concentrators to be within budget K . Constraints 2.4c enforce that a connection between smart meter a and data concentrator l can occur only if data concentrator l is installed and is within communication range of a . Constraints 2.4e describe the residual capacity, defined as the difference between the allowable bandwidth and the total throughput across each data concentrator. This definition corresponds to residual bandwidth in (Gourdin and Klopfenstein, 2006). Constraints 2.4f enforce that a data concentrator will not be installed unless at least one smart meter is assigned. Note that if $y_l=0$ (the data concentrator l is not selected), then no connections are possible for that concentrator (i.e. $x_{a,l}=0 \forall a \in A$ and $r_l=0$ from Constraints 2.4e). In addition, y_l is fixed to 1 for all existing data concentrators. Constraints 2.4g define the binary variables in the optimization, while Constraints 2.4h bound the residual capacity to be non-negative.

While the average residual buffer capacity seems like a reasonable metric, it does nothing to ensure bandwidth is balanced. Consider a simple case of two data concentrators, each with bandwidth b . A design that results in each data concentrator at 50% residual bandwidth has the same average as a design where one data concentrator is at 10% and the other at 90%. Therefore, this formulation would conceivably yield saturated nodes due to the nature of the averaging metric. Therefore, we also reformulate SGEP-A to (i) maximize the minimum residual buffer capacity and (ii) to minimize the sum of the reciprocal residual buffer capacities. These alternative objectives present options for penalizing oversaturated nodes in the network, while also handling node placement. In the case of reciprocal residual buffer capacities, the problem becomes nonlinear, and off-the-shelf MILP solvers are no longer directly applicable.

The reformulation to maximize the minimum residual buffer capacity, called the Smart Grid Expansion Problem - Maximin (SGEP-MM), is as follows

$$\underset{x,y,z}{\text{maximize}} \quad z \quad (2.5a)$$

$$\text{subject to} \quad r_l \geq z \quad \forall l \in L \quad (2.5b)$$

$$\sum_{l \in L} y_l \leq K \quad (2.5c)$$

$$x_{a,l} \leq y_l \quad \forall l \in L, a \in \mathcal{A}_l \quad (2.5d)$$

$$\sum_{l \in \mathcal{L}_a} x_{a,l} = 1 \quad \forall a \in A \quad (2.5e)$$

$$r_l = b_l - \sum_{a \in \mathcal{A}_l} x_{a,l} f_a \quad \forall l \in L \quad (2.5f)$$

$$y_l \leq \sum_{a \in \mathcal{A}_l} x_{a,l} \quad \forall l \in L \quad (2.5g)$$

$$y_l, x_{a,l} \in \{0, 1\} \quad \forall l \in L, a \in \mathcal{A}_l \quad (2.5h)$$

$$r_l \geq 0 \quad \forall l \in L \quad (2.5i)$$

Objective 2.5a now maximizes an auxiliary variable z , defined as the minimum residual capacity over all connections in the network, enforced through Constraints 2.5b. The objective function in SGEP-A was focused on maximizing the sum of the residual buffer capacity. For SGEP-A, if a data concentrator was not selected, r_l must be forced to zero for that specific data concentrator. In the SGEP-MM formulation, we are maximizing the worst-case buffer capacity. This worst-case should be identified for only the selected data concentrators. Therefore, Equation 2.5f is written to ensure that unselected data concentrators do not impact this worst case objective. That is, numerically, the residual capacity r_l for unselected data concentrators are set equal to the total bandwidth to ensure that they are not impacting the objective function as written. All other constraints remain the same as those of SGEP-A.

As opposed to the SGEP-A formulation, which optimizes for maximum average residual capacity over the entire network, the proposed formulation SGEP-MM improves reliability by assigning connections to data concentrators in an attempt to

maximize the worst-case, or *minimum* residual capacity. By ensuring that the minimum residual buffer capacity is maximized, utility planners can be confident about minimum QoS guarantees the system will be able to provide.

The reformulation to minimize the sum of the reciprocal residual buffer capacities, called the Smart Grid Expansion Problem - Reciprocal (SGEP-R), is as follows

$$\underset{x,y}{\text{minimize}} \quad \sum_{l \in L} \frac{1}{r_l} \quad (2.6a)$$

$$\text{subject to} \quad \sum_{l \in L} y_l \leq K \quad (2.6b)$$

$$x_{a,l} \leq y_l \quad \forall l \in L, a \in \mathcal{A}_l \quad (2.6c)$$

$$\sum_{l \in \mathcal{L}_a} x_{a,l} = 1 \quad \forall a \in A \quad (2.6d)$$

$$r_l = b_l - \sum_{a \in \mathcal{A}_l} x_{a,l} f_a \quad \forall l \in L \quad (2.6e)$$

$$y_l \leq \sum_{a \in \mathcal{A}_l} x_{a,l} \quad \forall l \in L \quad (2.6f)$$

$$y_l, x_{a,l} \in \{0, 1\} \quad \forall l \in L, a \in \mathcal{A}_l \quad (2.6g)$$

$$r_l \geq 0 \quad \forall l \in L \quad (2.6h)$$

Objective 2.6a now minimizes the sum of the reciprocal residual buffer capacities in the network. Constraints 2.6e have been modified in the same manner as SGEP-MM. SGEP-R follows the same assumption that the residual capacities of unselected data concentrators are equal to its bandwidth rather than zero in order for the reciprocal objective to perform as expected. All other constraints remain the same as those of SGEP-A.

While SGEP-A maximizes the total residual capacity of the network, it fails to consider the distribution of links to each node. In other words, the SGEP-A objective is indiscriminant of a network with oversaturated nodes and concentrated links versus a network with evenly distributed links. The SGEP-R objective, however, does seek an even distribution of links in the network while also maximizing the total residual capacity. The reciprocal objective function used in SGEP-R has been referred to as

a tractable penalty model, falling between the extreme approaches used in SGEP-A and SGEP-MM in terms of the number of connections routed to data concentrators (Gourdin and Klopfenstein, 2006).

Due to the reciprocal objective, SGEP-R becomes a mixed-integer nonlinear programming problem, and can be challenging to solve. Fortunately, the nonlinearity is convex, and this problem can be approximated to arbitrary accuracy by applying linear under-estimators to approximate the nonlinear function. First, linear under-estimators can be applied to the nonlinear function using a Taylor Series expansion as follows

$$\frac{1}{r_l} \simeq \frac{1}{r_{l,m}^*} - \frac{1}{r_{l,m}^{*2}}(r_l - r_{l,m}^*) \quad \forall l \in L, m \in M_l \quad (2.7)$$

where $r_{l,m}^*$ denotes the points along the domain of r_l (denoted by $m \in M_l$) to apply the under-estimators. These points can be generated uniformly or algorithmically to improve convergence times. In our cases, we have chosen to generate them uniformly. In addition, because the nonlinear function is convex with respect to the minimization

objective, the constraint can be converted to an inequality. The result is the following relaxed SGEP-R formulation,

$$\underset{x,y,v}{\text{minimize}} \quad \sum_{l \in L} v_l \quad (2.8a)$$

$$\text{subject to} \quad v_l \geq \frac{1}{r_{l,m}^*} - \frac{1}{r_{l,m}^{*2}}(r_l - r_{l,m}^*) \quad \forall l \in L, m \in M_l \quad (2.8b)$$

$$\sum_{l \in L} y_l \leq K \quad (2.8c)$$

$$x_{a,l} \leq y_l \quad \forall l \in L, a \in \mathcal{A}_l \quad (2.8d)$$

$$\sum_{l \in \mathcal{L}_a} x_{a,l} = 1 \quad \forall a \in A \quad (2.8e)$$

$$r_l = b_l - \sum_{a \in \mathcal{A}_l} x_{a,l} f_a \quad \forall l \in L \quad (2.8f)$$

$$y_l \leq \sum_{a \in \mathcal{A}_l} x_{a,l} \quad \forall l \in L \quad (2.8g)$$

$$y_l, x_{a,l} \in \{0, 1\} \quad \forall l \in L, a \in \mathcal{A}_l \quad (2.8h)$$

$$r_l \geq 0 \quad \forall l \in L \quad (2.8i)$$

where v_l is an auxiliary variable for the reciprocal residual buffer capacity, which is approximated by Constraint 2.8b. This relaxation allows the problem to be solved as a mixed-integer linear program. For the rest of the chapter, SGEP-R will refer to the formulation comprising Eq. 8a-8i.

2.3.1 Redundancy

In addition, due to the flexibility of the MILP framework, constraints in all three formulations can easily be modified to consider connection redundancy. To do so, we can modify constraints 2.4d, 2.5e, and 2.8e and apply a redundancy threshold to the right-hand side as follows

$$\sum_{l \in \mathcal{L}_a} x_{a,l} = C \quad \forall a \in A \quad (2.9)$$

Constraint 2.9 provides redundancy in the event that a data concentrator is unavailable for a connected smart meter, either because of component failure or unexpected congestion, ensuring that each smart meter can connect to at least C data concentrators. For example, if $C=2$, then this formulation ensures that each smart meter can connect to one of two data concentrators (e.g. primary and secondary) while ensuring that the data concentrator can feasibly handle the combined bandwidth from all potential connections identified.

2.4 Case Study

To show the application of the SGEP formulations for a SGCN, and to evaluate the capabilities of the modeling method in answering Smart Grid communication system problems, a case study designing the topology of a semi-urban SGCN was conducted.

The Reference Test Case A (RTCA), shown in Figure 2.2, was adapted from Pacific Northwest National Laboratory (PNNL) taxonomy feeder R2-25.00-1 (Schneider and Yousu Chen David Chassin Robert Pratt Dave Engel Sandra Thompson ii, 2008) and represents a moderately-populated suburban area. It was placed arbitrarily in an area with good solar resource data to provide representative scale for the network, but has no relation to actual locations depicted. The characteristics of the network are listed in Table 2.2.

The RTCA network spans 10.0 km from East to West and 8.0 km from North to South.

A grid layout of candidate locations for data concentrators was chosen where new nodes are allowed to be placed at any vertex on the grid. The density of the grid can be assigned as a parameter to the optimization. In this case, a 44x44 matrix of candidate locations was used, totaling 1936 candidate locations with approximately 200 m between each candidate node.

The communication range of the smart meters via 2.4 GHz WLAN is dependent on the transmit power of the transmitters (smart meters), the receive power at receivers

(data concentrators), terrain, elevation, and line-of-sight. Sources report widely ranging values for physical ranges of 802.11 signals, from 25 m for 802.11a/b/g (Kuzlu et al., 2014) to 80-450 m for 802.11ah (for smart cities in urban settings) (Bellekens et al., 2018) to 1.0 km for 802.11ac (Kuzlu et al., 2014). Instead, we will utilize a path loss propagation model to predict the communication range of WLAN, described in Section 2.2.6, particularly for a suburban topology that reflects the RTCA test case. Nevertheless, the approach taken in this research can be easily extended to urban environments that utilize closer range protocols (as in the IEEE 802.11ah standard for smart cities). The communication range is vital to construct the sets \mathcal{L}_a and \mathcal{A}_l , which determine what smart meter-data concentrator pairs are communicable.

Given the size of the system under consideration, the LTE communication range allows data concentrators to communicate with the edge router with minimal loss regardless of their proximity to each other. In addition, each PV inverter is tied to a smart meter in the network, but additional smart meters exist for which no PV inverters are linked. These additional smart meters represent residents who have not yet integrated solar generation into their household, but will in the near future.

Therefore, the Smart Grid expansion problem is formulated to place additional data concentrators in order to ensure network connectivity for these emerging sources of distributed generation. A multi-period expansion will be performed using the proposed formulations to evaluate the effectiveness of the methods for network expansion over time. The optimization is focused on the placement of the data concentrators and the assigned connections between the smart meters and the data concentrators.

Table 2.3 summarizes the parameter values used in the optimization. To facilitate the case study, the buffer capacities and data transfer sizes are kept constant across all data concentrators and smart meters in the network, but the proposed formulations are amenable to varying values for these parameters.

Andreadou et al. (2018) studied smart meter traffic and data concentrator processing for AMI systems, particularly the effects on the network for smart meter message sizes from 4 bytes to 12,000 bytes and transmission intervals from 4 to 24 times per

day. They showed that for an urban network of 101 meters to 1 data concentrator, 12,000 bytes transmitted 24 times per day from each meter required about 11 Mbits of data to be processed daily at each data concentrator. For our case, the amount of data or size of packets to be processed daily from each smart meter was set to be at this maximum (11 Mbits) in order to simulate a scenario where a large amount of daily traffic occurs at each data concentrator in the network. This also ensures that the resulting network will be able to handle increases in data rates from future demand. Because the approach relegates the handling of packet transmission and reception rates to the hardware technologies, the placement problem needs only to consider best locations to place the hardware and the specific routes in order to ensure network capacity can handle high-traffic events.

An LTE range of 50.0 km was chosen to represent a moderately-populated suburban area for RTCA from (Kuzlu et al., 2014). For more rural environments, evolved UMTS Terrestrial Radio Access Network (eUTRAN) can be used, which is an air interface for LTE cellular network that supports coverage up to 100 km with acceptable performance (Rémy and Letamendia, 2014). For denser urban environments, LTE coverage ranges are typically only reported up to 20 km, according to the COST 231-Hata model (Hamid, 2013). The proposed approach is general and can be extended to urban or rural environments with appropriate choice of LTE coverage.

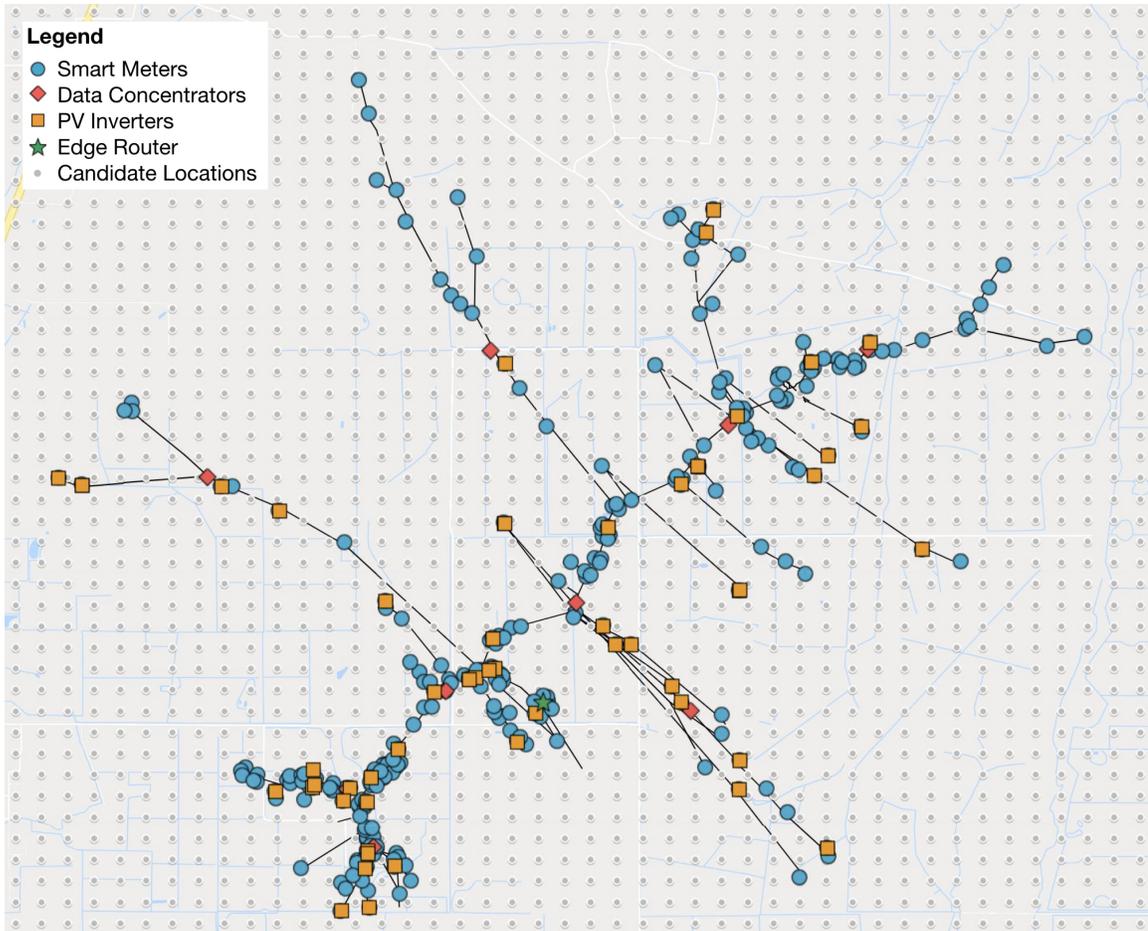


Figure 2.2.: RTCA network diagram.

Table 2.2.: Reference Test Case A Topology

Node Type	Total Count
PV Inverters	57
Smart meters	275
Data concentrators	10
Edge router	1
Candidate locations for DCs	1,936

Table 2.3.: Parameter Values for SGEP-A/-MM/-R

Model Parameters	Value
b_l	640 Mbits (Andreadou et al., 2018)
f_a	11 Mbits (Andreadou et al., 2018)
WLAN Range	0.93 km [refer to Section 2.2.6]
LTE Range	50.0 km (Rémy and Letamendia, 2014; Kuzlu et al., 2014)

2.5 Numerical Results

In the first set of numerical results, we consider all three formulations (SGEP-A, SGEP-MM, and SGEP-R) with a single period, no redundant connections. The optimization problems were constructed using Pyomo (Hart et al., 2011a, 2017) and solved with Gurobi Optimizer. Gurobi Optimizer is a commercial optimization solver primarily used for linear, quadratic and mixed-integer linear or quadratic programming problems (Gurobi Optimization, 2018). Gurobi Optimizer is well-established in the operations research literature as a state-of-the-art solver for mixed-integer problems of this class. The results were obtained on a computer running a 2.4 GHz Intel Core i5, dual-core processor with 8 GB RAM. Table 2.4 lists the computational results for SGEP-A, SGEP-MM, and SGEP-R with different budgets of data concentrators (budget K) based on the RTCA network. A total of 10 uniformly-distributed linear under-estimators were used for the reciprocal function for each data concentrator in Constraint 2.8b.

Given a budget limit, fixed assignments from Constraints 2.4c, and the assumption of constant parameter values in this case study, the average (or total) residual buffer capacity of the network can be pre-computed assuming the budget allows a fully connected network. As such, SGEP-A (Formulation 2.4) reduces to a purely feasibility formulation. The SGEP-A formulation leads to both oversaturated connections and

Table 2.4.: Numerical Results for SGEP-A, SGEP-MM, and SGEP-R solutions

Budget K	Model	Objective value	Residual Buffer Capacity	
			max (%)	min (%)
8	SGEP-A	73.74	98.28	0.31
8	SGEP-MM	55.31	94.84	55.31
8	SGEP-R	1329.30	93.12	53.59
10	SGEP-A	76.37	98.28	0.31
10	SGEP-MM	67.34	93.12	67.34
10	SGEP-R	1328.56	93.12	63.91
25	SGEP-A	86.50	98.28	0.31
25	SGEP-MM	84.53	96.56	84.53
25	SGEP-R	1327.42	93.12	84.53
50	SGEP-A	92.12	98.28	0.31
50	SGEP-MM	89.69	98.28	89.69
50	SGEP-R	1327.07	94.84	84.53

underutilized data concentrators, as evidenced by the unchanged maximum and minimum residual buffer capacities of 98.29% and 0.31%, respectively. The SGEP-MM formulation, on the other hand, seeks to maximize worst-case residual buffer capacity, yielding values of 55.31%, 67.34%, 84.53%, and 89.69% for budget limits of 8, 10, 25, and 50 data concentrators, respectively. The SGEP-R formulation seeks to minimize the sum of the reciprocal of the residual buffer capacity and performs similarly to SGEP-MM, yielding minimum values of 53.59%, 63.91%, 84.53%, and 84.53% for budget limits of 8, 10, 25, and 50 data concentrators, respectively.

The histogram comparisons in Figures 2.3 - 2.5 show clearly the distribution of connections per data concentrator in the solutions provided by SGEP-A, SGEP-MM, and SGEP-R for an increasing budget size. Since SGEP-A only considers feasibility,

the distributions shown in Figure 2.3 are not surprising – a large number of connections to a few data concentrators along with many underutilized data concentrators with only one or two connections. On the other hand, formulations SGEP-MM and SGEP-R result in more balanced distributions of data flow with the same average residual buffer capacity over the network.

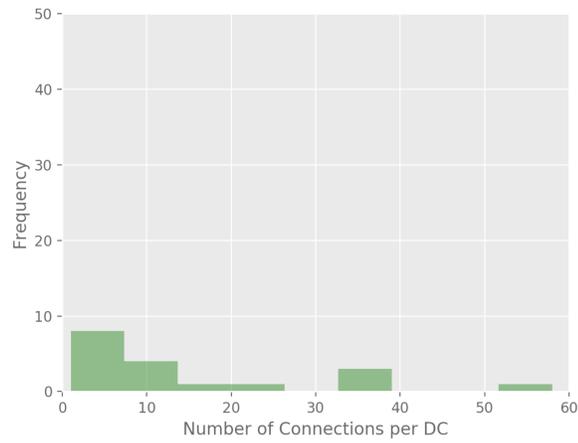
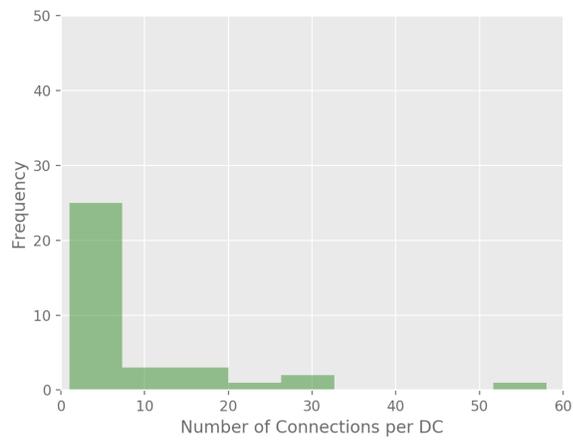
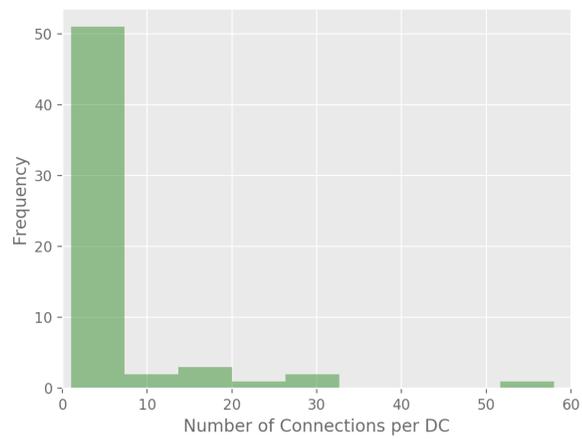
(a) SGEPA with $K=8$ (b) SGEPA with $K=25$ (c) SGEPA with $K=50$

Figure 2.3.: Distribution of connections per data concentrator for SGEPA solutions with budget limit (a) $K=8$, (b) $K=25$, and (c) $K=50$.

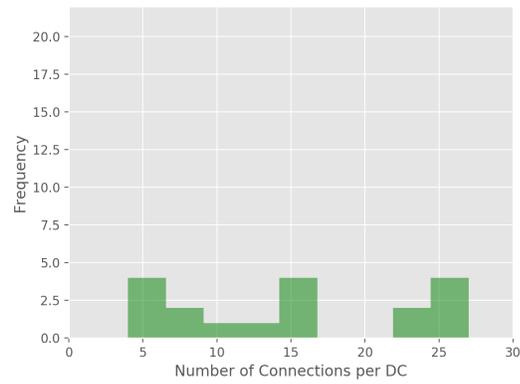
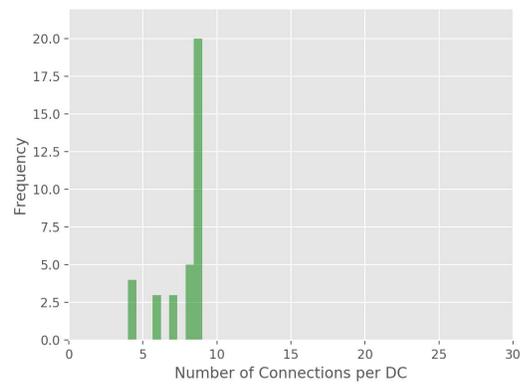
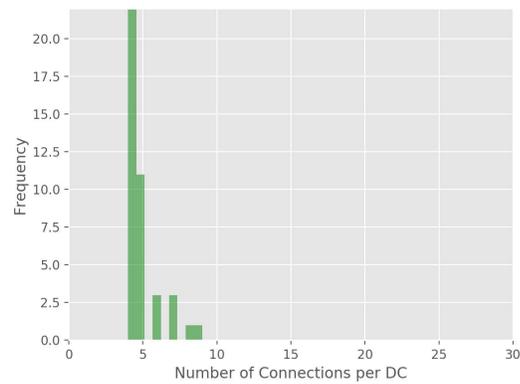
(a) SGEPR with $K=8$ (b) SGEPR with $K=25$ (c) SGEPR with $K=50$

Figure 2.4.: Distribution of connections per data concentrator for SGEPR solutions with budget limit (a) $K=8$, (b) $K=25$, and (c) $K=50$.

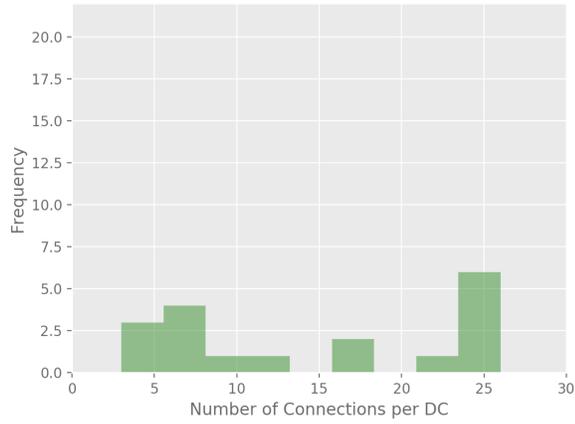
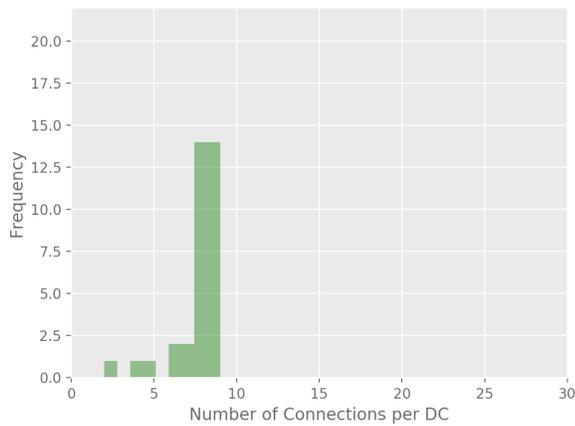
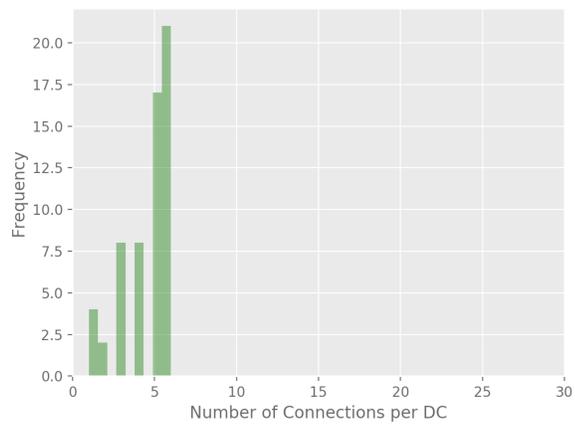
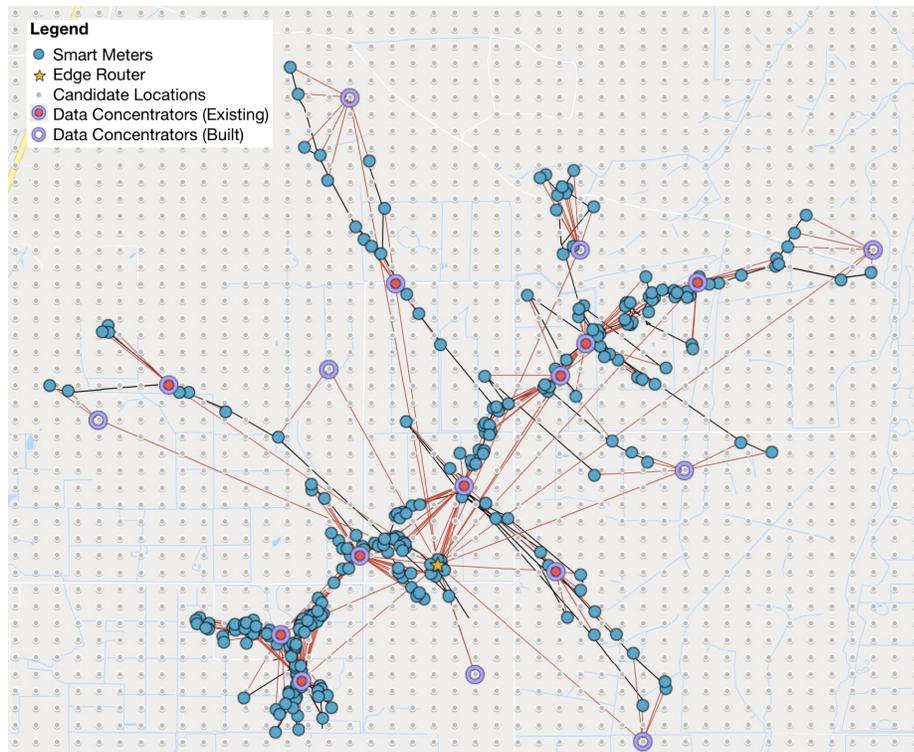
(a) SGEP-MM with $K=8$ (b) SGEP-MM with $K=25$ (c) SGEP-MM with $K=50$

Figure 2.5.: Distribution of connections per data concentrator for SGEP-MM solutions with budget limit (a) $K=8$, (b) $K=25$, and (c) $K=50$.

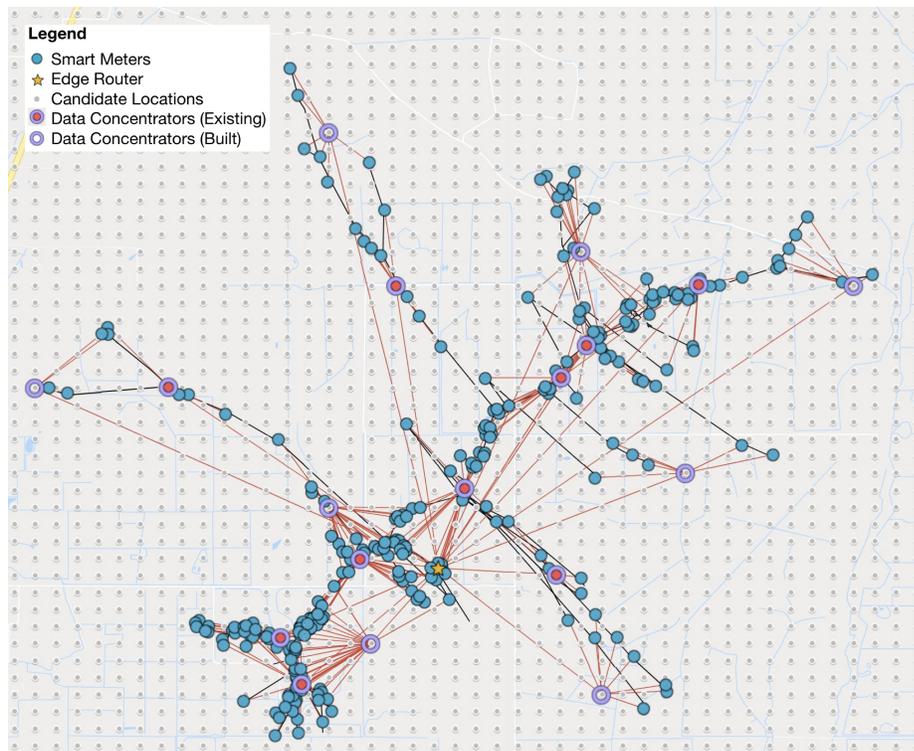
The reciprocal objective in SGEP-R penalizes designs with data concentrators that have low residual buffer capacity. As Figure 2.4 shows, SGEP-R attempts to equalize the number of connections per data concentrator across the network to reduce bottlenecks. This behavior effectively reduces the maximum number of connections per data concentrator by consolidating the distribution. As the budget limit for data concentrators is increased, this consolidation becomes more pronounced, yielding a progressively tighter distribution. By distributing connections effectively, the network robustness and QoS are improved in the case of node outages.

SGEP-MM behaves similarly to SGEP-R. SGEP-MM maximizes the minimum residual buffer capacity of the network, which, in test cases where the amount of data to be processed daily f_a for each smart meter and the buffer capacities b_i of each data concentrator are equal, is equivalent to minimizing the maximum number of connections per data concentrator. Rather than reducing the maximum number of connections as a result of consolidating in the case of SGEP-R, SGEP-MM directly minimizes the maximum value. By doing so, SGEP-MM not only distributes connections similar to SGEP-R to reduce congestion bottlenecks, but also provides the guarantee that the minimum residual buffer capacity is maximized. One disadvantage of SGEP-MM is that, by prioritizing the minimization of the high end of the distribution, SGEP-MM neglects the low end, as shown in Figure 2.5c. This can lead to a few underutilized nodes. Compared with Figure 2.4c, SGEP-R provides better management of the low end of the distribution, but does not provide the guarantees of SGEP-MM.

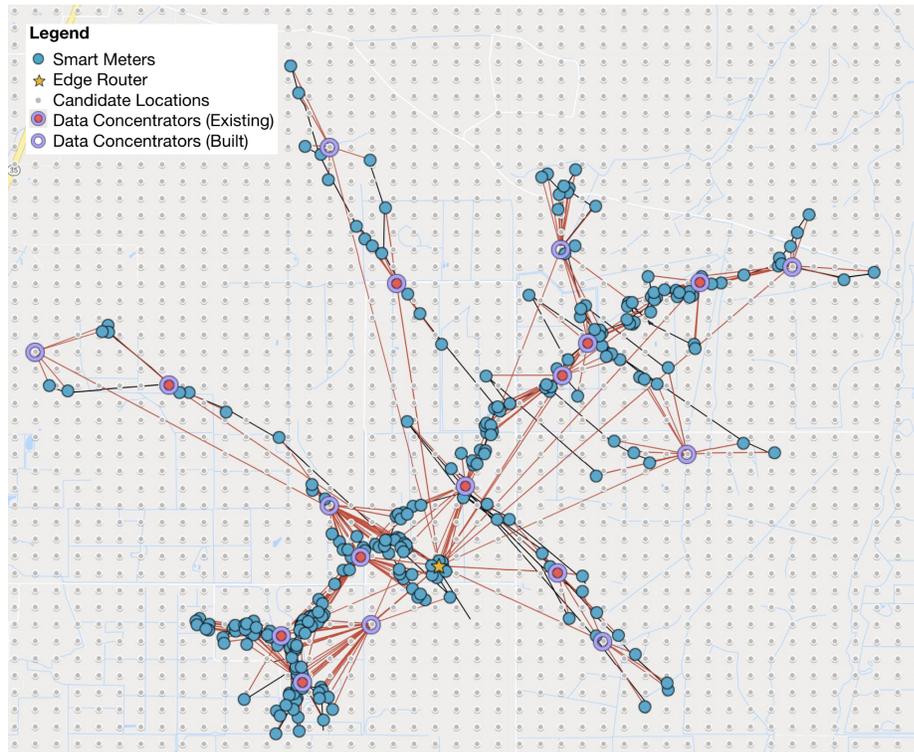
Figure 2.6 shows the resulting placement of data concentrators following the SGEP-A (Figure 2.6a), SGEP-MM (Figure 2.6b), and SGEP-R (Figure 2.6c) formulations for budget $K=8$ data concentrators. This limit is also the fewest number required in order for the problem to remain feasible (i.e. fully connected). As observed, the placements can be distant from the feeder network. Since we have assumed a radius of communicability for each smart meter rather than distance-dependent performance, this result is not surprising.



(a) SGEp-A



(b) SGEp-MM



(c) SGEPR

Figure 2.6.: SGCN solution for 8 newly-installed data concentrators with (a) SGEPA, (b) SGEPM, and (c) SGEPR formulations, respectively. Black links denote the underlying power line feeder. Red links denote the connections from smart meters to data concentrators, and from data concentrators to the edge router.

Both SGEPM and SGEPR assign connections to more effectively utilize each data concentrator. As such, not only do they handle network demand more efficiently by spreading the load across data concentrators, but the resulting placement locations are also closer to the feeder network as a consequence of increasing the number of connections to data concentrators. By staying within vicinity of the feeder network, the solutions of SGEPM and SGEPR become more practical for utility operators and contractors, since the infrastructure upon which the data concentrators can be

installed may already be in place. SGEP-R shows a slightly better result than SGEP-MM in terms of adhering to the feeder network, although the difference is minimal.

The SGEP-R formulation was also tested on a feeder-centric candidate topology, where the data concentrators were restricted to be placed only along existing utility poles. The findings from this test showed that the placements from the gridded topology were very similar to those from the feeder topology. The restricted feeder topology results in a similar number of connections per data concentrator. Minor discrepancies in location of placement between the two were due to the fact that the optimization model considered any connecting node within communication radius to be of equal value. In other words, as long as a smart meter and corresponding data concentrator were within communication range, there was no added value in locating the data concentrator closer to the smart meters. This is evident in the results of the gridded case, where placements of data concentrators can be seen in locations that extend outwards in directions away from the utility lines, in intermediate distances between its connected smart meters.

2.5.1 System Resilience

To address system resilience, we use the modified formulations that enforce redundancy. This formulation enforces multiple possible connections for each smart meter. With backup connections, we improve the resiliency and robustness of the network in the face of possible security threats, outages, or node failures. In the case of a node outage, the formulation ensures that a nearby data concentrator will be able to continue processing the information originally handled by the damaged unit since routing assignments and placement strategies have accounted for the additional buffer capacity necessary.

Table 2.5 lists numerical results for SGEP-A, SGEP-MM, and SGEP-R with redundancy threshold $C=2$, where $K=20$ is the minimum number of additional data concentrators required for a feasible, fully-connected network. Figure 2.7 shows the resulting placement of data concentrators following the SGEP-A (Figure 2.7a), SGEP-

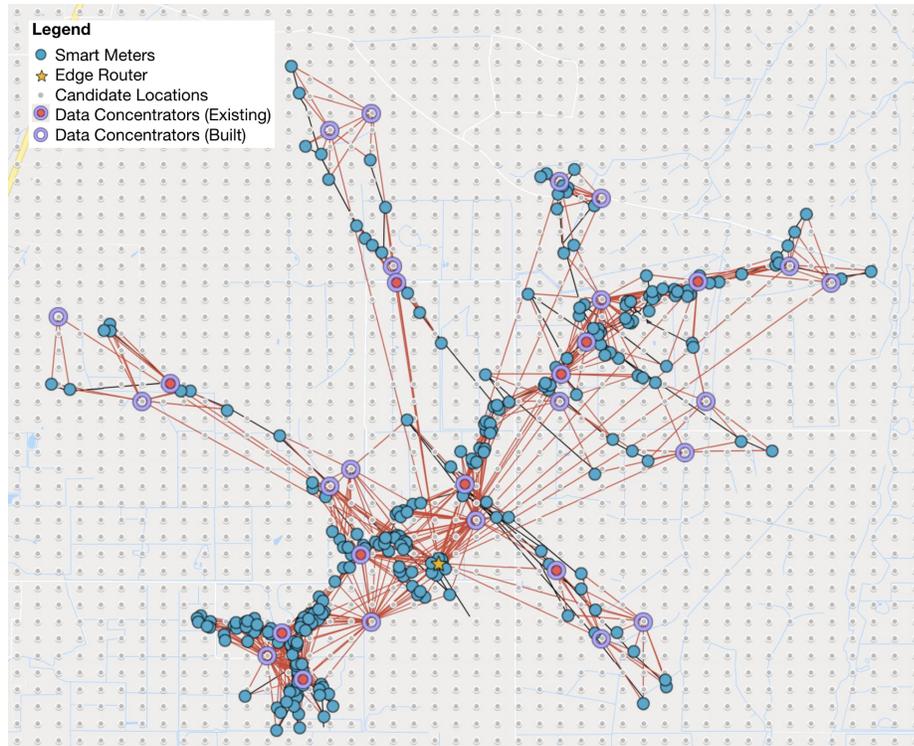
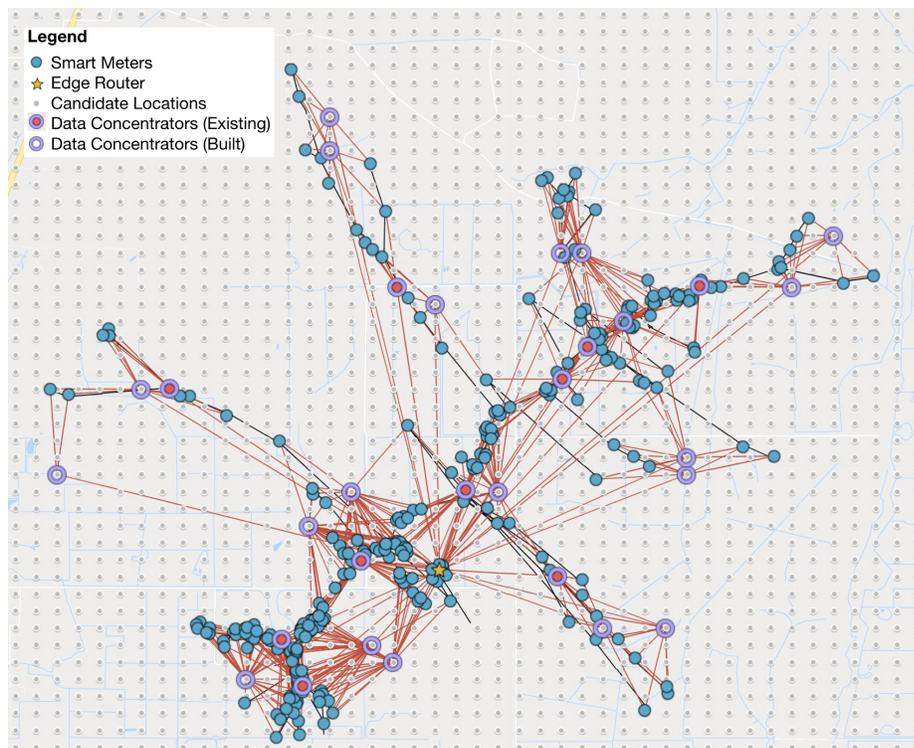
MM (Figure 2.7b), and SGEP-R (Figure 2.7c) formulations for budget $K=20$ data concentrators and redundancy threshold of $C=2$.

Since SGEP-A focuses on feasibility only, data concentrators are overutilized with very low residual buffer capacity (0.31%), along with underutilized data concentrators yielding high residual buffer capacities of 98.28%. On the other hand, SGEP-MM produces solutions with higher minimum residual buffer capacities (43.28%, 62.19%, 70.78%, and 82.81% for $K=20, 25, 30,$ and $50,$ respectively) while maintaining high maximum residual buffer capacities as well (93.12%, 93.12%, 94.84%, and 96.56% for $K=20, 25, 30,$ and $50,$ respectively). SGEP-R performs close to SGEP-MM, producing minimum residual buffer capacities of 41.56%, 62.19%, 63.91%, and 81.09% for $K=20, 25, 30,$ and $50,$ respectively and maximum residual buffer capacities of 93.12%, 91.41%, 93.12%, and 94.84% for $K=20, 25, 30,$ and $50,$ respectively.

In comparison with the numerical results from the $C=1$ solutions, the redundancy solutions for $C=2$ yield lower residual buffer capacities overall. This is not surprising, since consideration of redundancy requires the potential of additional throughput to data concentrators, leading to lower overall residual buffer capacities for the same number of nodes (minimum residual capacities of 84.53% for SGEP-MM with $C=1$ versus 62.19% for SGEP-MM with $C=2$ for $K=25,$ for example). However, as the number of data concentrators increases, this difference decreases (minimum residual capacities of 89.69% for SGEP-MM with $C=1$ versus 82.81% for SGEP-MM with $C=2$ for $K=50,$ for example). This effect is due to the fact that the redundancy threshold, in our case, does not scale with the number of data concentrators, but with the number of smart meters. Since the number of smart meters in the network is fixed, increasing the number of allowable data concentrators will always serve to alleviate congestion by increasing total buffer capacity.

While this phenomenon holds for SGEP-MM and SGEP-R, the same cannot be said of SGEP-A. SGEP-A could still assign the maximum number of connections to a few data concentrators, regardless of the budget limit, while sparsely assigning the minimum number of connections to the rest of the data concentrators.

Analyzing Figure 2.7 shows all models placing data concentrators in locations that can be very close to others. This result is due to the fact that, in the event that a data concentrator fails, the smart meters in the area must be able to reconnect to an alternative data concentrator nearby. An important trade-off occurs between the desire to prevent against node failure by applying redundancy versus effective utilization of data concentrators. Nevertheless, SGEP-MM and SGEP-R handle this trade-off well by more evenly distributing route assignments among data concentrators despite placing ones that can be in close proximity to each other. Future formulations should consider addressing the assumption of independence of failure and spatial distribution of data concentrators.

(a) SGEPA with $C=2$ (b) SGEPM with $C=2$

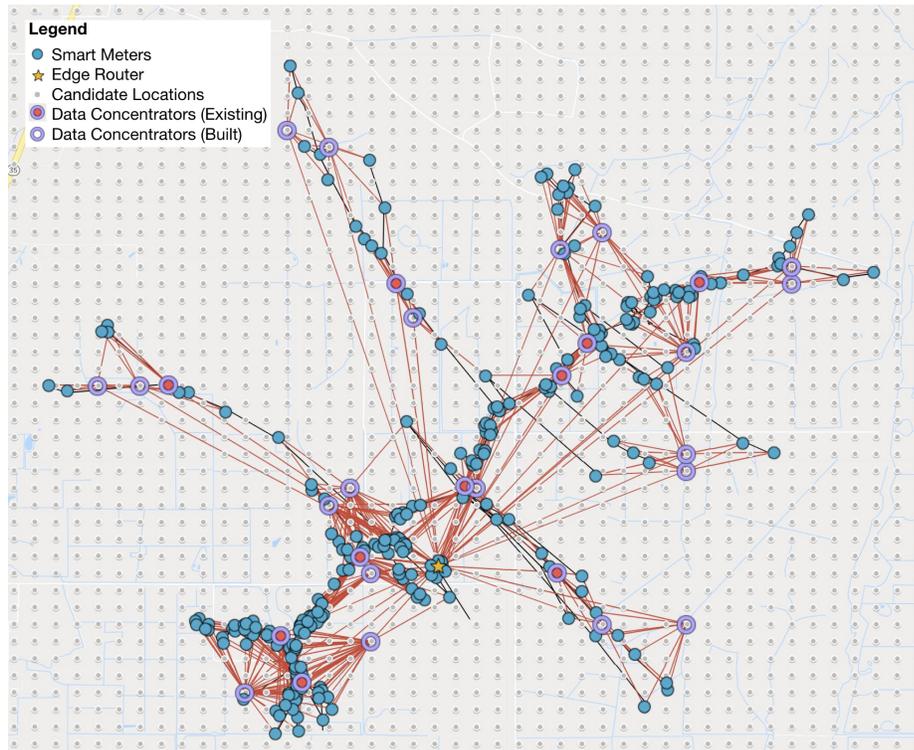
(c) SGEP-R with $C=2$

Figure 2.7.: SGCN solution with (a) SGEP-A, (b) SGEP-MM, and (c) SGEP-R formulations for 20 newly-installed data concentrators (redundancy $C=2$)

2.5.2 Multi-Period Expansion

As this study focuses on the expansion problem, it may be reasonable to assume that utility planners will need to resolve the expansion problem multiple times over the course of decades as the Smart Grid grows and distributed generation spreads. Rather than solving a large-scale multi-period problem, we verify the effectiveness of a sequence of single-period solves in a rolling horizon. In terms of the optimization, we model this rolling horizon approach by progressively solving the expansion problem with an increasing budget while requiring that nodes from each prior placement be considered built.

Table 2.5.: Numerical Results for SGEP-A, SGEP-MM, and SGEP-R with redundancy $C=2$

Budget K	Model	Objective value	Residual Buffer Capacity	
			max (%)	min (%)
20	SGEP-A	68.49	98.28	0.31
20	SGEP-MM	43.28	93.12	43.28
20	SGEP-R	1339.31	93.12	41.56
25	SGEP-A	72.99	98.28	0.31
25	SGEP-MM	62.19	93.12	62.19
25	SGEP-R	1335.84	91.41	62.19
30	SGEP-A	76.37	98.28	0.31
30	SGEP-MM	70.78	94.84	70.78
30	SGEP-R	1334.75	93.12	63.91
50	SGEP-A	84.24	98.28	0.31
50	SGEP-MM	82.81	96.56	82.81
50	SGEP-R	1333.10	94.84	81.09

Figure 2.8 shows the system performance for a series of single-period optimizations (with 5 data concentrators added in each period). This is compared with the best possible solution obtained if the placements were re-optimized for the new number of data concentrators. From this comparison, we see that while placing a full set of data concentrators at once (with the basic SGEP-MM formulation) yields a higher minimum residual capacity, the difference compared to the rolling horizon solution is minimal, with an average gap of 4%. This gap is dependent on the number of nodes utilized for each progressive addition. The minimum residual buffer capacities of the rolling horizon solutions and re-optimized solutions for SGEP-R are nearly identical, and lie close to those of the rolling horizon solutions of SGEP-MM.

Investigation of SGEP-A was omitted here since the minimum residual buffer capacities of SGEP-A solutions remain unchanged, falling automatically to their extremes (minimum of 0.31%).

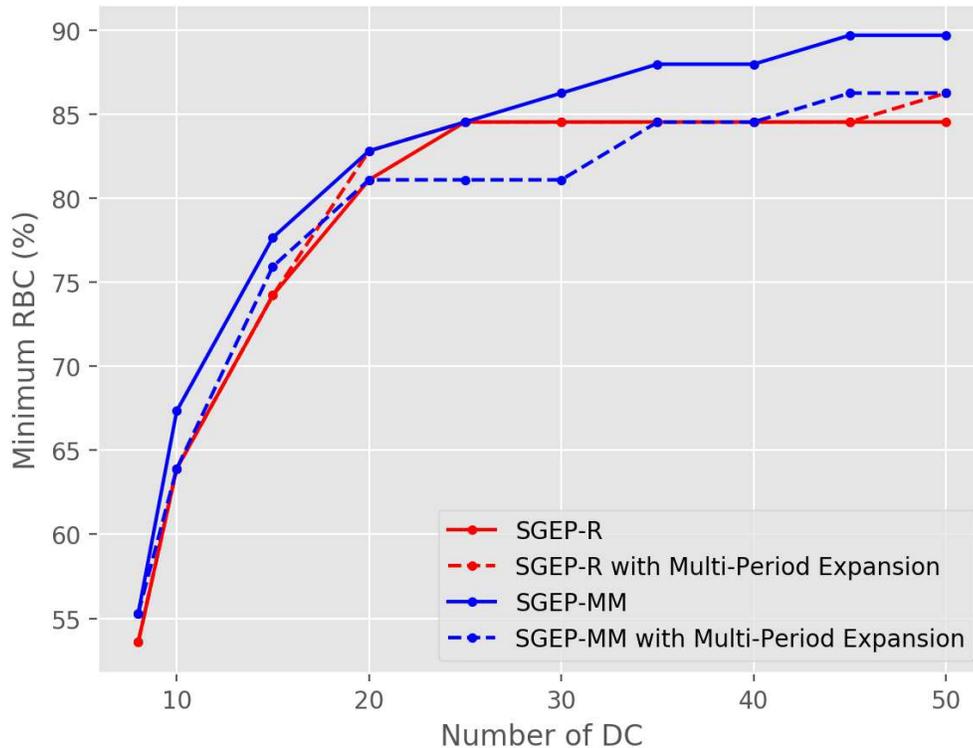


Figure 2.8.: Comparison of SGEP-MM/-R vs SGEP-MM/-R with Multi-Period Expansion

2.6 Conclusions and Future Work

Our key conclusions in this work are as follows. First, our proposed formulations worked to design optimal placement topologies of data concentrators in Smart Grid communications networks. Second, the SGEP-A model can be useful for targeting (i.e. minimum number of data concentrators for feasibility), but is poor for routing assignment, since the total (or average) residual buffer capacity can be computed

directly. Third, the SGEP-MM and SGEP-R models provide better solutions in terms of routing assignment, and leads to better placement within vicinity of the feeder network, a practical consideration useful for ensuring that chosen placements can be built. In addition, SGEP-MM provides minimum residual buffer capacity guarantees that are useful for establishing and ensuring QoS standards. Fourth, these formulations can address resilience concerns and are effective for rolling horizon network expansion.

Future plans of study include evaluation of the proposed approaches on a variety of urban environments and network sizes, as well as showing the scalability of the approaches across a larger span of problems. In addition, investigation of the optimal spatial distribution when placing nodes is important, since clustered placement of data concentrators can reduce their effective utilization despite provisioning for node outages. Also, multi-period placement approaches with stochastic growth models for the expansion of the Smart Grid network over time can address concerns such as worst-case scenarios and placement under uncertainty. Finally, investigation of statistical metrics such as Conditional Value-at-Risk (CVaR) can be useful when considering alternative objectives for optimal placement prioritizing risk mitigation.

3. OPTIMAL FLAME DETECTOR PLACEMENT WITH NON-UNIFORM UNAVAILABILITIES

In petrochemical facilities, flame detectors provide an important layer of protection for personnel. Flame detectors, as opposed to smoke or heat detectors, are optical sensors that utilize information from a visual field capture to detect the presence of flames. Flame detectors can utilize a number of different visual sensing technologies, including ultraviolet (UV), infrared (IR), combination UV/IR, and visible light. For effective and reliable detection, flame detectors require a visual path to the flame, free of obstructions, referred to as line-of-sight (LOS). Petrochemical facilities, however, are typically characterized by complex physical geometries arising from the large number of obstructions such as valves, pipes, tanks, and reactors in the plant. In addition, a significant number of uncertainties, including visual field decay, unknown obstructions, sensor outages, and cost limitations, make the problem of where to optimally place flame detectors extremely challenging. Traditional approaches for flame detector placement are ad-hoc, rule-of-thumb methods that are not typically quantitative in nature, and do not account for these uncertainties.

Numerical optimization techniques provide a way to overcome these challenges. Camm et al. (2002) proposed the Maximal Expected Coverage Problem (MECP) as a way of optimizing expected coverage over a region of space, applied to the selection of nature reserve sites to maximize the expected number of species. In this work, we translate this formulation to the problem of optimal placement of flame detectors. To the best of our knowledge, no prior work has been done on solving the problem of optimally locating optical flame detectors within process facilities by maximizing expected coverage in the face of detection failure.

In this chapter, we present a mixed-integer nonlinear programming formulation for determining the optimal plant-specific placement of optical flame detectors by maximizing the expected coverage of 3D space while accounting for path-dependent probability of detection failure. A linear relaxation of the original formulation is developed along with an iterative method to achieve global convergence to the original nonlinear formulation. Under real-world conditions, we show that this probabilistic formulation, with consideration of uncertainties, outperforms the max coverage formulation that neglects potential for detection failure. We utilize the state-of-the-art fire and gas mapping software developed by Kenexis Consulting Corporation, Kenexis Effigy (Kenexis, 2019), to evaluate the effectiveness of the formulations on real-world facility geometries. In addition, we integrate our developed models into the Python package, Chama (Klise et al., 2017), for general usage.

The chapter is organized as follows. Section 3.1 provides background on industrial fire detection and provides motivation for this work. This is followed by Section 3.2 in which the mixed-integer linear programming formulation for optimal placement of flame detectors assuming perfect detection is presented, along with the probabilistic extension in the form of a mixed-integer nonlinear programming problem, and a corresponding convex relaxation and iterative solution algorithm. Section 3.3 compares the perfect detection formulation and probabilistic extension with analysis on one small-scale test case and two real-world facility geometries. The formulations are evaluated under real-world conditions, first compared across fixed probability levels of detection failure, and then across randomized probabilities of detection failure.

3.1 Background

The optimal flame detector placement problem investigated in this chapter is related to the optimal camera placement problem. The optimal camera placement problem considers the best locations to place a number of visual camera sensors (fixed or motion-tracking) to ensure coverage over a region of interest, and is derived from the Art Gallery problem. The Art Gallery Problem (AGP) is the problem of

determining the minimum number of guards required to cover the interior of an n -wall art gallery room (??, ORo). This problem, originally an exercise in geometry and an application of the classic set covering problem, has been used as the foundation for much work in the fields of camera and directional sensor placement.

Olague and Mohr (1998) presented a solution to the problem of positioning camera sensors in 3D space to minimize the uncertainty in viewing a set of observable points P . They defined the coverage based on relative accuracy of obtaining 3D measurements and attempted to minimize the measurement error using a genetic algorithm. Recently, Erdem and Sclaroff (2004) presented the problem of optimal camera placement utilizing realistic camera capabilities and obtained a solution with integer linear programming (ILP) techniques while handling arbitrary polygonal shapes. Zhao and Cheung (2009) defined a grid-based solution to optimal camera placement with visual tagging based on a set of visibility conditions. Hanoun et al. (2016) most recently investigated placement of a camera network for target coverage within manufacturing facilities with visibility and quality-of-resolution constraints. Much of the models concerning camera placement assumes data fusion capabilities. For fire safety systems, however, optical flame detectors are attempting to process information to alert personnel quickly in the event of an emergency, and would not require the two-way communication functionality that would be typical for camera sensor networks. In addition, while the camera placement problem is well-studied, they do not consider uncertainties such as detection failure, which can impact reliability in a fire safety system design.

The Maximal Expected Covering Location Problem (MEXCLP) was originally introduced by (Daskin, 1982, 1983). The work addressed the problem of locating a number of facilities to maximize the expected demand covered by an available emergency medical service facility. Camm et al. (2002) then proposed the Maximal Expected Coverage Problem (MECP) and the Linearized Maximal Expected Coverage Problem (LMECP), an extension and a linear approximation of the MEXCLP, respectively, within the context of placing nature reserve sites in order to maximize

expected coverage of species habitats. The formulation in (Camm et al., 2002) improved upon the MEXCLP by assuming location-dependent probabilities, rather than assuming a system-wide probability.

The mixed-integer nonlinear programming formulation we use to maximize expected coverage for the placement of flame detectors is similar to the MECP extension proposed by Camm et al. (2002). We translate it to instead consider probabilities of detection failure for each optical flame detector in the system to be independent of each other, and allow for individual specifications according to distance from detector to target. Rather than perform a linear approximation of the nonlinear probabilistic model as in (Camm et al., 2002), our approach performs a convex relaxation of the model using linear under-estimators via a Taylor-Series approximation. In addition, we develop an iterative method of adding linear under-estimators that tightens our relaxed formulation to reach the global optimum of the original nonlinear model. We show that, for varied probabilities of detection, our iterative approach is computationally efficient, and competitive with commercial global MINLP solvers.

Fire detection sensors can be divided into three detection categories: smoke/gas, heat, and flame – each of which are detected via sensing techniques such as chemical- or optical-based methods. It is also important to note that not all fire detectors are suited to detect all fires. Chemical-based detectors are predominantly used to detect the presence of smoke that is created from smoldering fires (e.g. wood and coal), or various harmful gases (e.g. carbon monoxide (CO)). These chemical-based detectors typically employ catalysts or radioactive ions that interact with harmful compounds in the air to alert personnel of gases, and are point detectors that must be placed in the path of the evolving gas cloud to be effective. Fonollosa et al. (2018) provides a comprehensive review of chemical sensors and corresponding algorithms for fire detection.

While chemical-based detectors are effective for smoke and gases, they are ineffective for detecting presence of heat or flames. Optical-based detectors, on the other hand, have smoke detection capabilities as well as heat and flame-detection capabil-

ities, depending on the technology and suitability to fire. Optical-based detectors are typically comprised of UV, IR, visual, or a combination of these technologies to detect presence of fires. For example, video-based systems, utilizing specialized algorithms, can recognize and detect flames. Similarly, optical smoke detection can also be done with specialized algorithms and visual sensors. Comprehensive systems can combine IR array technology with video imaging to locate flames within the field of view in addition to detection capabilities. Bogue (2013) provides an in-depth review of current sensor technologies for fire detection.

One major advantage of optical-based detectors for flame detection is their ability to provide detection coverage over large physical areas due to their camera-like field-of-view. Placed intelligently, significantly fewer optical-based flame detectors would be required to ensure coverage over a target area than would a network of chemical-based point detectors. Second, optical-based flame detectors can, in many cases, provide much faster detection times than other types of flame detection. Yuan et al. (2018) compared the response times of various fire detection sensors (e.g. CO, smoke, and visual flame detectors) placed in different locations around a central flame. They found that visual flame detectors produced the shortest detection times regardless of location in their test scenario, whereas CO and smoke detectors provided less reliable detection and required significantly more time (in some cases, 10x longer) to detect the same fire.

A number of studies have been done surrounding algorithms that can better identify flames from image data from optical flame detectors. Milke et al. (2003) investigated multiple sensor technologies and data sources to identify the best variables to consider in order to discriminate flames from nuisance sources. Several studies have used support vector machines to classify flames from visual sensor data (Wang et al., 2006; Ko et al., 2009). Thuillard (2002) investigated and designed a new flame detector using fuzzy-wavelet algorithms. Wu et al. (2019) recently present a convolutional neural network-based visual fire detection method to handle the intelligent and automatic identification of flames as part of the on-board software within visual

flame detector systems. In this work, we focus instead on optimally placing optical flame detectors, and assume that, should a flame scenario occur in the field of view of an optical flame detector, the detector will be able to successfully detect the flame with a certain probability.

Previous formulations addressed rigorous mathematical programming approaches for optimal placement of gas detectors. Legg et al. (2012) initially proposed a stochastic programming framework to handle placement of gas detectors using dispersion simulations generated using computational fluid dynamics (CFD) models. This work was then expanded in (Legg et al., 2013) to handle worst-case scenarios by considering Conditional-Value-at-Risk (CVaR). Subsequently, Benavides-Serrano et al. (Benavides-Serrano, 2014; Benavides-Serrano et al., 2014) extended the stochastic programming framework to consider unavailabilities and voting to handle detector failure and false alarms. Recently, Rad et al. (2017) investigated the optimal placement problem for gas detectors using a risk-based approach, formulated as a maximum weighted coverage problem, and proposed a greedy algorithm.

While the formulations described above focused on point-based or line-of-sight gas detectors, a different set of assumptions are required for placing optical flame detectors. The first-to-detect concept employed in (Legg et al., 2012, 2013; Benavides-Serrano et al., 2014; Benavides-Serrano, 2014) (i.e. minimizing detection time by placement of gas detectors near or close to likely gas dispersions) is no longer applicable for optical-based detector systems. Instead, the design of a flame detection system must consider whether a flame event in a region-of-interest (ROI) is visible and detectable by the detection network as a whole.

The placement of optical flame detectors in the petrochemical industry is typically based upon rules-of-thumb that rely on identifying equipment of interest and convenient mounting points for detectors. Several sets of internationally and regionally-recognized standards and guidelines for fire detection and alarm systems can be found in NFPA 72 (Association, 2019), ISO 7240-1:2014 (Nunes, 1999), and EN 54 (Nunes, 1999). However, these guidelines predominantly provide qualitative methods

for placement of point detectors or alarms (i.e. gas, smoke, and/or heat detectors), and only briefly mention optical flame detector placement. Indeed, little work has been done on the problem of optimally placing optical flame detectors within petrochemical facilities. Obi (2014) evaluated the placement of flame detectors between using 2D or 3D detector mapping simulation tools, and showed that 3D mapping techniques provided better information to the user to improve flame detector system designs. Yang et al. (2012) used color classification and a recognition algorithm to identify the placement of visual fire detectors. They assigned hazard ratings to key equipment, corresponding to a color mask, and used an exhaustive approach to scan all possible locations along a wall for each detector in order to determine best placement according to the highest coverage values.

There is a need for a systematic approach to optical flame detector placement that addresses the following gaps:

- The existing practices are based on rule-of-thumb approaches that require significant human input. While 3D detector mapping software has become more widespread, it still relies on human interaction through manual placement of individual detectors within the software, while providing no guarantee of optimality. A rigorous optimization-based approach to flame detector placement has the potential to provide quantifiably improved solutions while allowing design engineers to focus on improving input data and exploring alternative solutions through increased automation.
- Within the petrochemical industry, it is imperative that alarm systems are reliable, otherwise there can be catastrophic consequences. Therefore, the design of any safety system must rigorously consider uncertainties. In the context of optical flame detectors, these uncertainties can take the form of detector failure, wiring issues, and partial or unknown obstructions, for example. An effective detector placement approach should be able to consider and account for these uncertainties.

The mathematical programming formulation presented in this chapter addresses the above concerns in the form of a mixed-integer nonlinear programming formulation that considers uncertainty in detection, and yet is solvable to global optimality with reasonable computational effort.

3.2 Problem Formulation

3.2.1 Basic Flame Detector Formulation (Ignoring Detection Failure)

Figure 3.1 is used to describe the terminology that is used in this chapter. Figure 3.1 shows an overhead view of a simple flame detector placement within a facility with two large obstructions, A and B. The optical flame detector, denoted in red with the number 1, has coverage (or vision) of the area in green. Each block of open space that is covered by the flame detector is denoted as an “entity”. Within each entity lies a possibility of a flame event. The aim of the flame detector placement problem is to maximize coverage of the unobstructed entities in the facility in order to minimize the chance that a flame goes undetected.

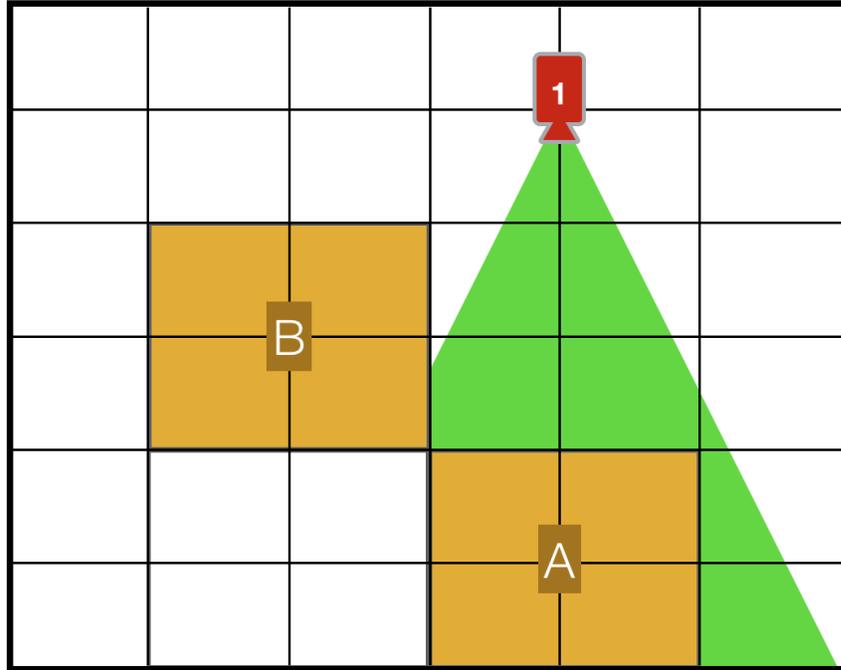


Figure 3.1.: Flame detector layout example

The initial approach for placement of optical flame detectors that does not consider detection failure is in the form of a set covering problem, where detectors are placed in order to maximize total coverage of a set of unobstructed entities or regions of interest with the assumption that the detectors provide perfect coverage (or detection).

The flame detection problem (FDP) is formulated as a mixed-integer linear program given by

$$\underset{x,y}{\text{maximize}} \quad \sum_{e \in E} y_e w_e \quad (3.1a)$$

$$\text{subject to} \quad \sum_{l \in L} x_l \leq k \quad (3.1b)$$

$$\sum_{l \in L_e} x_l \geq y_e \quad \forall e \in E \quad (3.1c)$$

$$x_l, y_e \in \{0, 1\} \quad \forall l \in L, \forall e \in E \quad (3.1d)$$

where E denotes the set of entities e that are unobstructed (i.e. regions of interest), L denotes the set of all candidate locations l for detectors, L_e denotes the set of

candidate detector locations that can observe entity e , y_e is an indicator variable that has a value of 1 if entity e is observed by at least one detector, w_e is a weighted parameter for each entity e that allows priority of areas to be considered, x_l is a binary variable denoting whether a detector is built at location l , and k is a parameter denoting the detector budget limit. The notation used in the FDP formulation is summarized in Table 3.1.

The objective function 3.1a maximizes the total number of entities $e \in E$ observed by the placed detectors. Constraints 3.1b limits the number of detectors to be placed to be no greater than budget limit k . Constraints 3.1c ensures that each entity e can only be observed if a detector is placed within observable range. Constraints 3.1d specifies the binary decision variables x_l and y_e .

The coverage set L_e is specified through a data pre-processing step. First, candidate locations for detectors are specified in the topology of interest. For each candidate location, a ray-tracing analysis is performed to obtain the observable region captured by each candidate detector location. This information is organized into a set containing all entities e observed by each candidate detector location l . This set is then inverted to obtain L_e , the set of all candidate detector locations that have observable vision of entity e , where $L_e \subseteq L$.

Table 3.1.: Notation for Problem FDP

Symbol	Definition
L	Set of all potential detector locations
E	Set of all entities e to be observed
L_e	Set of detector locations that can observe e
x_l	Binary variable indicating 1 if detector is built at location l
y_e	Binary variable indicating 1 if entity e is observed
w_e	Entity weight
k	Maximum number of detectors allowed

3.2.2 Probabilistic Coverage Formulation (P-FDP)

The FDP formulation, while suitable for a preliminary analysis, fails to suffice for real-world process safety conditions because it does not consider uncertainties in detection. It assumes that detectors, once placed, provide perfect detection across its field-of-view. In reality, however, detection is rarely perfect, and any number of uncertain factors can affect the detection rate.

We therefore consider the case where detection is imperfect and associate a probability of detection for each detector-entity pair, and place detectors in such a way

to maximize the expected coverage. The resulting formulation is a mixed-integer nonlinear program (P-FDP) and is given by

$$\underset{x, \sigma}{\text{maximize}} \quad \sum_{e \in E} \sigma_e w_e \quad (3.2a)$$

$$\text{subject to} \quad \sum_{l \in L} x_l \leq k \quad (3.2b)$$

$$\sigma_e = 1 - 1\left[\prod_{l \in L_e} (1 - p_{l,e} x_l)\right] \quad \forall e \in E \quad (3.2c)$$

$$x_l \in \{0, 1\} \quad \forall l \in L \quad (3.2d)$$

$$0 \leq \sigma_e \leq 1 \quad \forall e \in E \quad (3.2e)$$

where σ_e is the expected coverage of entity e (i.e. probability of an event at entity e being detected), $p_{l,e}$ is the probability of a detector at location l successfully detecting an event at entity e , and w_e is the weight of entity e . Table 3.2 summarizes the notation in this formulation.

Constraint 3.2c describes the probability that a flame event at entity e is detected by the selected detectors. The second term in the right-hand side of Constraint 3.2c is the probability that an event at entity e is not detected. This probability is, of course, equal to 1 for all locations that do not observe entity e , and $(1 - p_{l,e} x_l)$ for locations that could observe entity e (recall that x_l is a binary variable that indicates whether or not a detector at location l is selected). Therefore, Constraint 3.2c is formulated as the complement of the probability that a flame event at entity e is not detected by the network.

In addition to accounting for uncertainties in detection, another major advantage of the P-FDP formulation is in its handling of partial obstructions in lower resolution geometries. The granularity of the discretized facility geometry determines the number of nonlinear constraints 3.2c and σ_e variables, which can drastically affect the problem size. Therefore, while it is advantageous to use a larger granularity (i.e. lower resolution) to reduce problem size, complexity, and solution times, a lower resolution of the discretized geometry leads to less information about obstructions.

The detection probability $p_{l,e}$ is impacted by all factors that lead to a chance of failed detection by the network of flame detectors. This includes, but is not limited to, factors such as defective detection software, detector maintenance downtime, and wiring issues. Since $p_{l,e}$ describes the detector-entity pairing, it is also useful for treatment of partially obscured entities. This probability term allows this uncertainty in partial obscurity to be considered when placing flame detectors.

Table 3.2.: Notation for P-FDP

Symbol	Definition
L	Set of all potential detector locations
E	Set of all entities e to be observed
L_e	Set of detector locations that can observe entity e
x_l	Binary variable indicating 1 if detector is built at location l
σ_e	Probability of entity e being observed
w_e	Entity weight
$p_{l,e}$	Probability of detector at location l successfully observing entity e
k	Maximum number of detectors allowed

3.2.2.1 Relaxed Form (RP-FDP)

The expected coverage model in Constraint 3.2c makes the P-FDP formulation highly nonlinear and difficult to solve. We propose a convex relaxation of the P-FDP formulation using linear under-estimators that can be solved as a mixed-integer linear program. We also introduce an iterative algorithm that solves the original P-FDP problem to global optimality using the relaxed form.

We begin by defining the nonlinear product term in Constraint 3.2c as γ_e , the probability that an event at entity e is not detected, where

$$\sigma_e = 1 - \gamma_e \quad \forall e \in E \quad (3.3)$$

and

$$\gamma_e = \prod_{l \in L_e} (1 - p_{l,e} x_l) \quad \forall e \in E \quad (3.4)$$

Next, a log-transform is performed on Equation 3.4, resulting in

$$\ln(\gamma_e) = \sum_{l \in L_e} \ln(1 - p_{l,e} x_l) \quad \forall e \in E \quad (3.5)$$

We then redefine $\ln(\gamma_e)$ as $\bar{\gamma}_e$. In addition, since x_l is a binary variable, it can be factored out of the logarithmic expression while maintaining an equivalent result, as follows.

$$\bar{\gamma}_e = \sum_{l \in L_e} x_l \ln(1 - p_{l,e}) \quad \forall e \in E \quad (3.6)$$

This transformation yields a linearized form of the original nonlinear product. The original product γ_e is reconstructed from $\exp(\bar{\gamma}_e)$. Due to the nature of the objective, this expression can be relaxed as an inequality producing a convex constraint, as follows.

$$\gamma_e \geq \exp(\bar{\gamma}_e) \quad (3.7)$$

A set of linear under-estimators can then be generated using a Taylor-Series Approximation that bounds the right-hand side and provides a tight relaxation to Equation 3.7, as follows,

$$\gamma_e \geq \exp(\bar{\gamma}_{e,m}^*) (\bar{\gamma}_e - \bar{\gamma}_{e,m}^* + 1) \quad \forall e \in E, m \in M_e \quad (3.8)$$

where $\bar{\gamma}_{e,m}^*$ are points along the domain of $\bar{\gamma}_e$ where linear under-estimators are generated and M_e is the set of all points in the domain of $\bar{\gamma}_e$.

This leads to the following model, referred to as the Relaxed Probabilistic Coverage Flame Detector Problem (RP-FDP):

$$\begin{aligned} & \underset{x, \sigma}{\text{maximize}} && \sum_{e \in E} \sigma_e w_e && (3.9a) \end{aligned}$$

$$\begin{aligned} & \text{subject to} && \sum_{l \in L} x_l \leq k && (3.9b) \end{aligned}$$

$$\sigma_e = 1 - \gamma_e \quad \forall e \in E \quad (3.9c)$$

$$\bar{\gamma}_e = \sum_{l \in L_e} x_l \ln(1 - p_{l,e}) \quad \forall e \in E \quad (3.9d)$$

$$\gamma_e \geq \exp(\bar{\gamma}_{e,m}^*) (\bar{\gamma}_e - \bar{\gamma}_{e,m}^* + 1) \quad \forall e \in E, m \in M_e \quad (3.9e)$$

$$x_l \in \{0, 1\} \quad \forall l \in L \quad (3.9f)$$

$$0 \leq \sigma_e \leq 1 \quad \forall e \in E \quad (3.9g)$$

As the number of linear under-estimators increases, the model provides a more accurate and tighter relaxation of the original problem. Therefore, an iterative algorithm is developed to solve the RP-FDP efficiently by intelligently placing additional under-estimators at each iteration and comparing this lower bound with the best candidate solution to achieve global optimality.

3.2.3 Iterative Solution Algorithm

An initial set of equidistant linear under-estimators are generated across the domain of $\bar{\gamma}_e$ for each entity e . The RP-FDP is the master problem in this case and is solved as a mixed-integer linear program. A relative optimality gap between the relaxed RP-FDP and the original P-FDP subproblem is calculated by (i) substituting the solution from RP-FDP into the P-FDP problem, (ii) calculating the objective value of P-FDP, and (iii) determining the percentage difference between the RP-FDP and P-FDP objectives. Then, an additional linear under-estimator is added using the current RP-FDP solution and the problem is re-solved. This iterative sequence continues until the optimality gap between the RP-FDP and P-FDP problems falls below a given tolerance or until a maximum iteration count has been achieved.

To summarize, the steps of the iterative algorithm are as follows:

Step 0 Initialization A set of 3-5 equidistant linear under-estimators across the domain of $\bar{\gamma}_e$ are added to initialize the relaxed problem.

Step 1 Solving the Upper Bounding Master Problem The mixed-integer master problem (RP-FDP) is formulated and solved to global optimality. If this problem converges, update the current upper bound and proceed to *Step 2*. Otherwise, the algorithm terminates and the current upper bound is the best value of the objective function.

Step 2 Solving the Nonlinear Subproblem With the integer solution from *Step 1*, the nonlinear subproblem P-FDP is fully determined and becomes a forward simulation. The optimality gap is then calculated between the RP-FDP and P-FDP solutions. If the optimality gap is below the tolerance, the algorithm terminates and the global solution is obtained. Otherwise, proceed to *Step 3*.

Step 3 Refining the Master Problem The current MILP master problem (RP-FDP) is refined by adding linear under-estimators to eliminate the current integer solution. This is done by adding an under-estimator at the previous value of $\bar{\gamma}_e$ for all $e \in E$. Proceed to *Step 1*.

Figure 3.2 depicts a visual representation of the iterative strategy using linear under-estimators. As more under-estimators are added, the set of constraints generates a tighter bound and converges to the original nonlinear function.

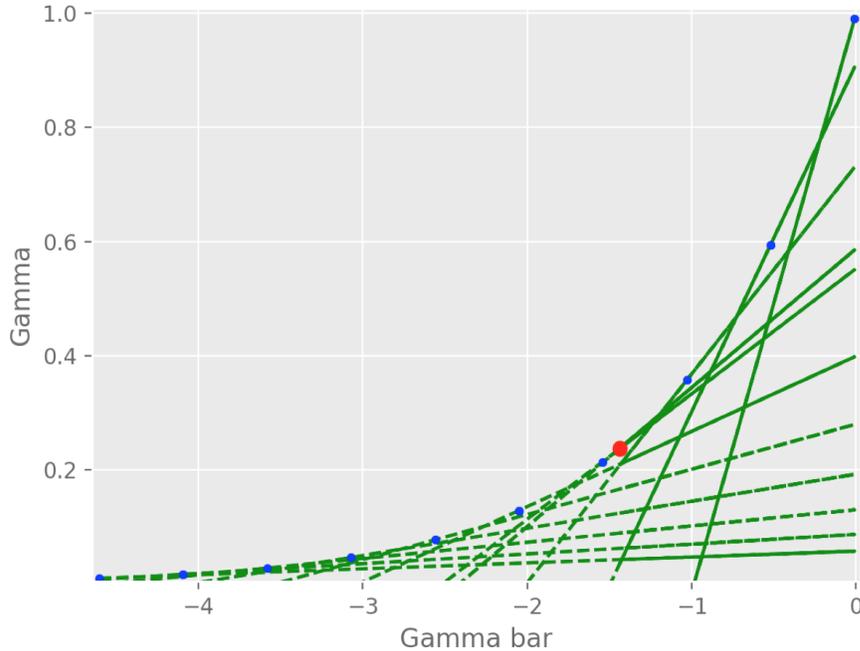


Figure 3.2.: Linear under-estimators

3.3 Numerical Results

The optimization problems presented in this chapter were written in Python using the Chama (Klise et al., 2017) and Pyomo packages (Hart et al., 2011a, 2017).

Chama is an open-source Python package that provides mixed-integer linear programming formulations to determine optimal sensor locations that maximize monitoring effectiveness. Chama is a general-purpose optimization software tool and is applicable to a wide range of sensor placement optimization problems. Currently, Chama is being actively used to design sensor networks for airborne and water pollutant monitoring. The FDP model in this chapter was solved using the Coverage formulation within Chama. In addition, the P-FDP and RP-FDP formulations developed in this chapter have been added to the Chama library for general usage.

The Python Optimization Modeling Objects (Pyomo) package provides a complete optimization modeling platform that allows for full scripting capabilities in a Python environment. Pyomo was used to write the proposed optimization models (P-

FDP and RP-FDP) in this chapter, including all variables, constraints, and objective functions.

The P-FDP formulation was solved using BARON 17.8.9 (Tawarmalani and Sahinidis, 2005; Sahinidis, 2017) while the global solution algorithm using the RP-FDP formulation was solved using Gurobi Optimizer (Gurobi Optimization, 2018). BARON is an optimization solver that solves non-convex optimization problems to global optimality and is particularly suited for mixed-integer nonlinear programs. Gurobi Optimizer is a commercial optimization solver that is well-established in the operations research field as a state-of-the-art solver for mixed-integer problems. All analyses were performed on a Intel Xeon E5-2697v2 12-Core 2.70 GHz machine with 250 GB RAM.

A preliminary test case, denoted as Case A, was created to show the effectiveness of the formulations on a small-scale facility. Case A, shown in Figure 3.3, represents a warehousing facility with 10 storage tanks and 3 wall structures, and spans 100 units by 100 units in area and 2 units in height. Each storage tank and wall structure measures 1 unit in height. A total of 50 candidate detector locations were considered for Case A, located on the edges and throughout the facility. Uni-directional detectors are placed at the sites on the edges of the facility, while each candidate location within the interior of the facility allows the placement of four uni-directional detectors, one oriented in each cardinal direction (i.e. at 0° , 90° , 180° , and 270° with respect to the horizontal axis). Each detector has a 90° field-of-view. Discretizing the space into 1 cu unit entities yields 8,300 unobstructed blocks.

The coverage set for Case A was generated using a ray-casting technique. A set of rays extend from each candidate detector location. A ray extends until it reaches an obstruction (e.g. storage tank or wall structure) or the boundary of the facility. Each unobstructed unit entity that the ray passed through is added to the set of entities that a candidate detector location can observe. A total of 300 discretization points about a 90° field-of-view are used to initialize the ray-casting from each candidate detector location. A step size of 0.1 units is used in casting each ray.

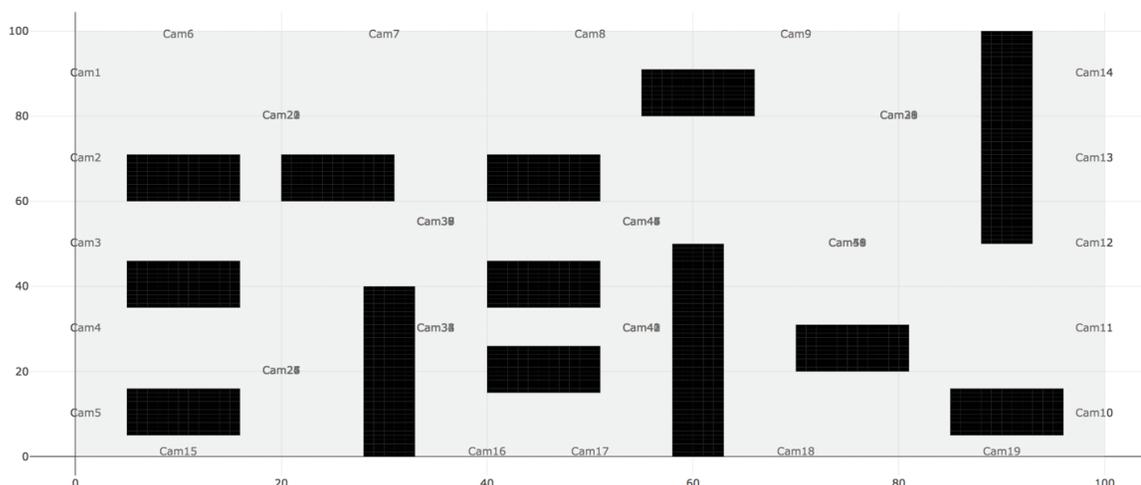


Figure 3.3.: Overhead view of Case A

The proposed formulations were also tested on real-world facilities, denoted in this chapter as Cases B and C, which were generated by Kenexis Consulting Corporation using the fire and gas mapping software, Kenexis Effigy (Kenexis, 2019). Kenexis Effigy Fire and Gas Mapping Software is a state-of-the-art fire and gas mapping tool that performs both geographic and scenario-based coverage calculations. Kenexis Effigy is built on the techniques from the ISA 84.00.07 Technical Report (ISA-TR84.00.07, 2010) that standardize the process for the design of fire and gas systems and was developed by industry experts to achieve a more quantitative approach to process safety management. For both Cases B and C, Kenexis Effigy also provided the coverage sets identifying the physical space captured within the field-of-view of each candidate detector location.

The first real-world case study, Case B, is a wellhead drilling facility, typical for offshore oil platforms. The facility geometry, shown in Figure 3.4, includes 16 wellheads, stairs, storage cabinet, and support structures, and spans 30m by 20m in area and 10m in height. A total of 24 candidate detector locations are considered for Case B, located on the surrounding support columns. Each flame detector was specified as a multi-spectrum infrared (MSIR) detector for hydrocarbon hazard identification

and a horizontal field-of-view of 90° . Discretizing the space into 1 cu ft entities yields 162,722 unobstructed blocks, 57% of which are seen when all 24 detectors are allowed. The reported coverage percentages in the proceeding analyses are normalized with respect to this maximal coverage percentage.

Case C represents a typical offshore platform hydrocarbon separations module and is a representative case of a complex process geometry. The facility geometry, shown in Figure 3.5, includes a host of process equipment (tanks, supports, pipes, etc.) spanning across 3 platform levels, measuring 73ft by 73 ft in area and approximately 35 ft in height. A total of 81 candidate detector locations were considered for Case C, located at appropriate mounting points throughout the facility. Each flame detector was specified as a multi-spectrum infrared (MSIR) detector for hydrocarbon hazard identification and a horizontal field-of-view of 90° . Discretizing the space into 1 cu ft entities yields 129,061 unobstructed blocks, 81% of which are seen when all 81 detectors are allowed. The reported coverage percentages in the proceeding analyses are normalized with respect to this maximal coverage percentage.

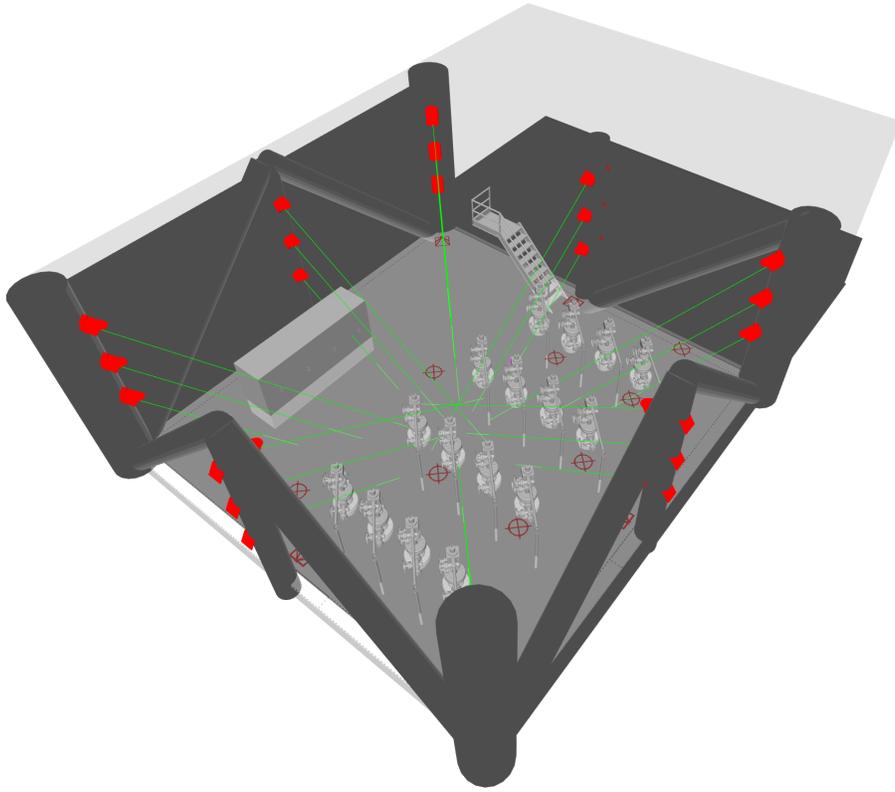


Figure 3.4.: Isometric view of Case B facility with 24 candidate flame detector locations

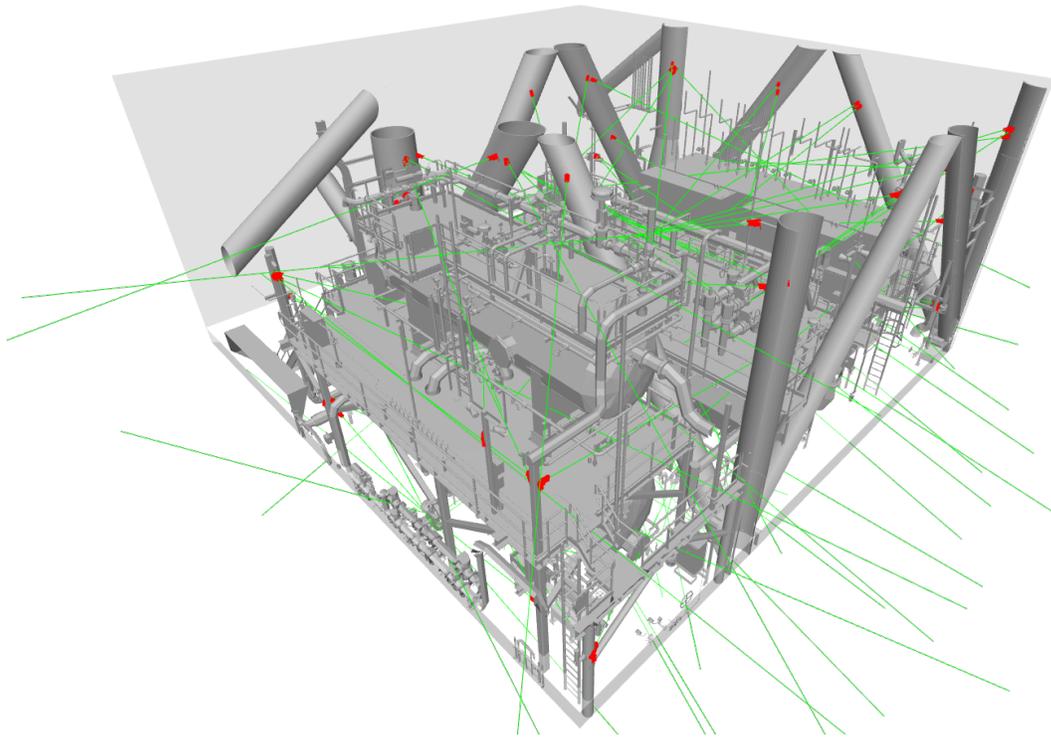


Figure 3.5.: Isometric view of Case C facility with 81 candidate flame detector locations

The FDP and P-FDP formulations were applied to all test cases. Recall that the FDP formulation places detectors assuming perfect detection, while the P-FDP formulation considers detection failure by modeling detection as a set of probabilities. The RP-FDP formulation is the relaxed form of P-FDP and is used as the master problem in the proposed global solution algorithm to solve P-FDP more efficiently. In the following results, the P-FDP model is solved using both BARON (when possible) and the developed algorithm. The solutions from both methods are the same, but solution times vary.

The relative optimality gap of the global iterative algorithm is set to 0.1% and the outer iteration limit is set to 100. For comparison, the relative optimality gap for

BARON is also set to 0.1% with a CPU time limit of 3 hours (10800 seconds). For evaluation purposes, the iterative algorithm was not given a CPU time limit.

Two separate analyses are performed concerning the set of detection probabilities. The first analysis considers the case where the detection probabilities $p_{l,e}$ are fixed to constant values of 50%, 70%, and 90% for all unit entities. In other words, the probability of successful detection of a flame event within any block by the set of selected detectors would be 50% for $p_{l,e}=0.5$, for example.

The second analysis considers the case where the detection probabilities $p_{l,e}$ are varied relative to each location l and entity e . Specifically, the detection probabilities $p_{l,e}$ were randomly sampled from a uniform distribution between the values 0.5 and 0.99. In this analysis, the probability of successful detection of a flame event within any block by the set of selected detectors could be any value within the interval $[0.5,0.99]$. The real facility analysis, these probabilities could be determined considering a number of factors, including the amount of partial obstruction for each detector-entity pair.

3.3.1 Results for Case A: Tank Storage Facility

The placement results of the FDP and P-FDP formulations for Case A are shown in Figures 3.6 and 3.7, respectively. Figures 3.6 and 3.7 show how the resulting placements of the FDP and P-FDP formulations differ. Each point depicts coverage from the selected detectors according to the corresponding optimization formulation. The P-FDP formulation attempts to maximize expected coverage. In so doing, coverage redundancy is automatically accounted for. On the other hand, the FDP formulation seeks to maximize total coverage assuming each detector has perfect detection capabilities. This difference can be seen in the placement results, where the FDP formulation results in fewer overlapping coverage fields than that of the P-FDP formulation.

As expected, if we compare the optimal expected coverage assuming perfect detection with that obtained when considering probability of detector failure, the for-

mulation considering detector failure is lower. To this end, we compare the optimal placement using FDP when assuming perfect detectors (detection failure probability of 0%) and when subjected to a detection failure probability of 30%. These results are shown in Figure 3.8a. Once detection failure is considered, the expected coverage of the resulting placement drops significantly, almost 15% at points.

This is not a realistic assessment, however, since the FDP formulation ignores the reality of detector failure. To analyze the impact of ignoring detector failure, in the next analysis, we place the detectors using the same two approaches (FDP assuming perfect detection, and P-FDP assuming 30% failure probability), but then we evaluate these placements using the more realistic case including failure probability of 30%. These results are shown in Figure 3.8b. As the number of detectors that are placed increases, so does the discrepancy between the FDP and P-FDP solutions under real-world conditions. The P-FDP solutions show consistent improvements in coverage compared with the FDP solution, reaching up to 10%. Figures 3.8a and 3.8b also show the diminishing returns in placing more detectors past a certain threshold. The expected coverage increases 60% when placing 1 to 10 detectors. However, placing more than 10 detectors for Case A yields no more than 15% additional expected coverage, all the way up to 40 placed detectors.

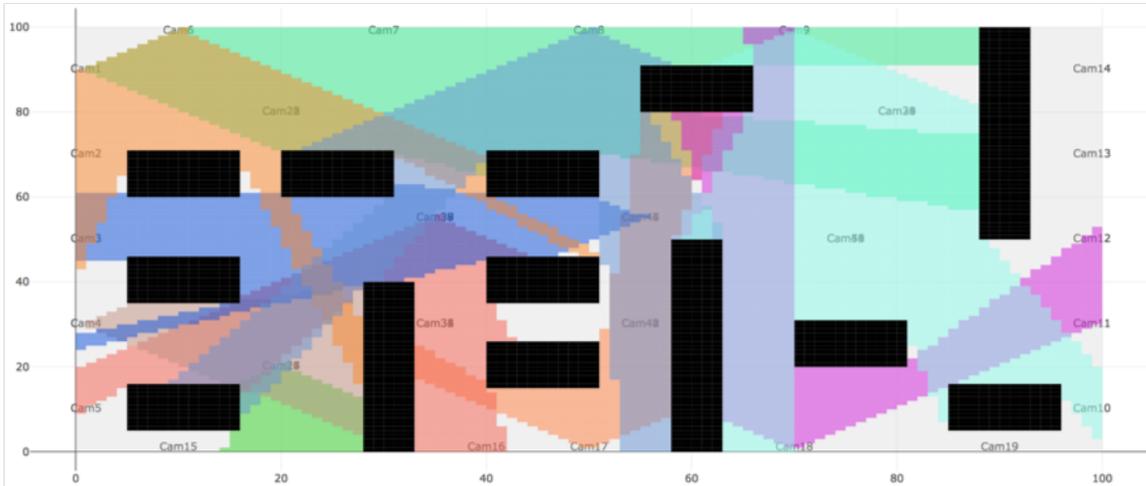


Figure 3.6.: Optimal placement results for FDP formulation for Case A with budget $k=10$

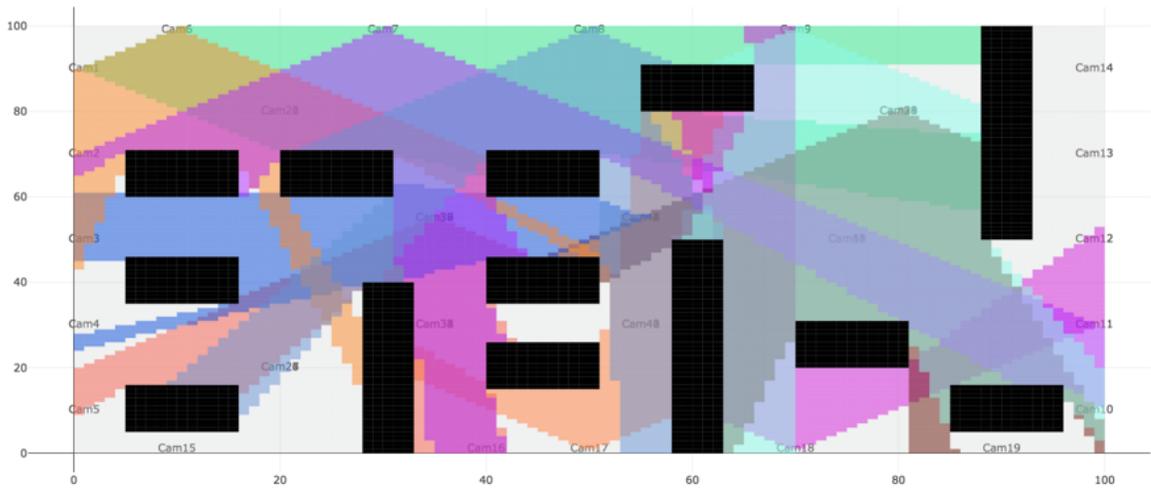
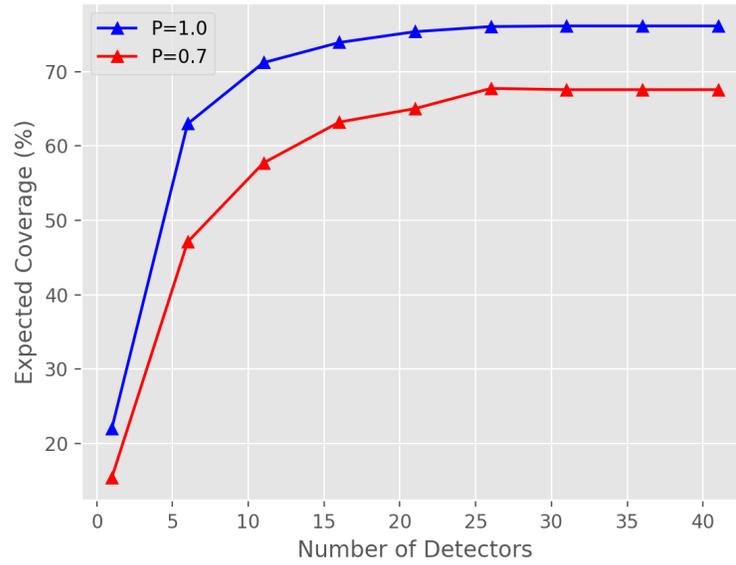
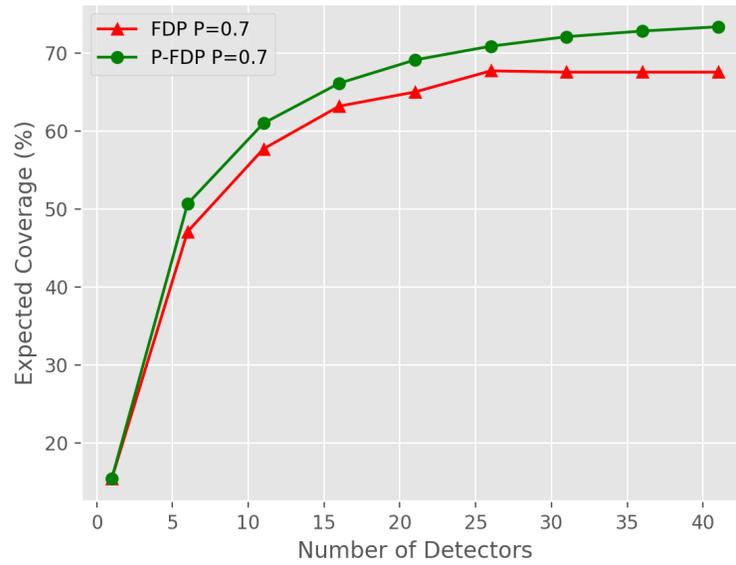


Figure 3.7.: Optimal placement results for P-FDP formulation for Case A with budget $k=10$ and $p_{l,e}=0.7$



(a)



(b)

Figure 3.8.: Expected coverage of FDP assuming perfect detection vs. FDP/P-FDP under imperfect conditions for Case A with (a) FDP assuming perfect detection ($p_{l,e}=1.0$) vs. imperfect conditions ($p_{l,e}=0.7$) and (b) Expected coverage of FDP vs. P-FDP under imperfect conditions ($p_{l,e}=0.7$)

In addition, the percent improvement of the P-FDP solution over the FDP solution is more pronounced when fewer detectors are considered. Figure 3.9 shows the percent improvement of the P-FDP solution across three detection probability levels $p_{l,e}=0.5, 0.7,$ and 0.9 for Case A. For high detection probability (50%) we see increasing improvement in the P-FDP solution over the FDP solution, reaching 20% improvement. A detection probability of 70% shows a moderate 5 to 8% improvement over the FDP solution. A detection probability of 90% shows the smallest improvement of around 2% over the FDP solution. Nevertheless, this value still represents a non-trivial improvement when designing a reliable fire safety system in a petrochemical facility.

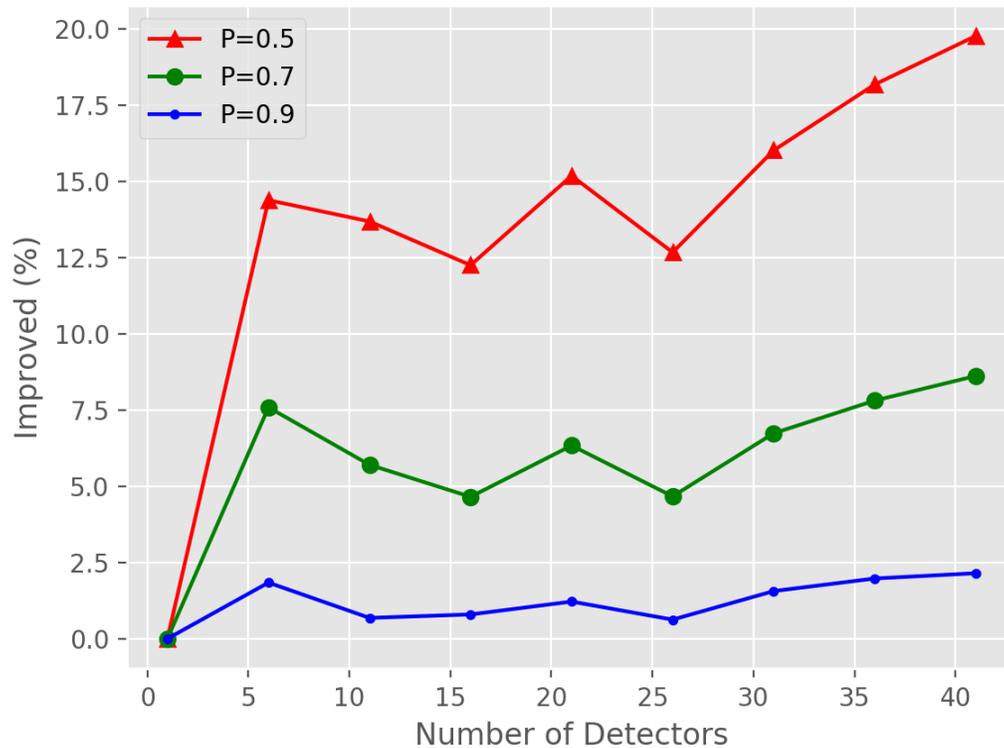


Figure 3.9.: Percentage improvement of P-FDP solutions against FDP solutions for Case A for probabilities of successful detection $p_{l,e}=0.5, 0.7,$ and 0.9 .

The FDP formulation solved consistently within 1.35 seconds for Case A, but it is a considerably smaller model compared to the P-FDP formulation and does not consider probabilities. Table 3.3 lists the numerical results for the P-FDP formulation for Case A across three probability levels $p_{l,e}=0.5, 0.7,$ and 0.9 . Listed are the detector budget k , the optimality gap percentage, the iteration count for the global solution algorithm, the timing of the solution algorithm in seconds, and the timing for BARON in seconds.

From Table 3.3 we find that the proposed global solution algorithm using linear under-estimators is well suited to solve the probabilistic flame detector placement problem. Our iterative algorithm converges to 0.1% optimality gap in 6 or fewer iterations. Generally, as the detection failure rate increases ($p_{l,e}$ decreases), our algorithm takes longer to converge over BARON. In this case, higher failure rates introduces more probabilistic variation between placements and requires additional under-estimators to attain the same degree of tightness of the relaxation when compared to problems considering lower detection failure rates. When the detection failure rate is at its lowest, the iterative algorithm performs fastest and converges in one iteration. At this point, the P-FDP formulation approaches the timing of the FDP formulation, since the probability of successful detection is reaching the perfect detection assumption.

The timing of the iterative algorithm is comparable with that of BARON at high detection probability levels ($p_{l,e}=0.9$) and takes about 5 times as long when $p_{l,e}=0.5$. In some cases, namely when few detectors, the performance of the algorithm fares better than BARON. When few detectors can be placed, there may be clear placements that provide the best coverage. In the same vein, when the number of detectors that can be placed approaches the total number of available locations, the choice of the best coverage that can be gained from the remaining locations may also be clear. However, when the number of detectors that can be placed and the number of available locations are both large, the optimization requires significantly more time to evaluate the available combinations. There may exist a large number

of integer solutions that give similar objective values. This increase in computational time occurs for both the iterative algorithm and BARON.

Table 3.3.: Numerical Performance of P-FDP formulation comparing Iterative Algorithm and BARON for Case A

Budget k	$p_{l,e}=0.5$				$p_{l,e}=0.7$				$p_{l,e}=0.9$			
	Gap (%)	Iter	Alg	BARON	Gap (%)	Iter	Alg	BARON	Gap (%)	Iter	Alg	BARON
1	0.00	4	41	93	0.00	4	41	42	0.02	1	7	71
6	0.00	4	82	52	0.00	2	38	58	0.02	1	20	40
11	0.06	4	115	57	0.00	5	166	38	0.02	1	34	32
16	0.06	3	49	30	0.00	3	93	64	0.03	1	16	8
21	0.02	4	65	30	0.00	3	96	8	0.03	1	14	9
26	0.03	6	121	8	0.10	3	62	9	0.03	1	26	9
31	0.04	3	45	8	0.10	2	23	8	0.03	1	19	10
36	0.00	3	45	8	0.08	2	23	8	0.03	1	12	10
41	0.05	3	46	8	0.04	2	30	8	0.02	1	12	10

In the previous case, it was assumed that the probability of detection failure was the same for all detector-entity pairs. However, the probabilities depend on several factors, including partial obstruction, and we also consider the case where these probabilities are not all the same. We call this the non-uniform case. The percentage improvement of the P-FDP solution against the FDP solution was also compared for the non-uniform case. Figure 3.10 shows the comparison for Case A for detection probability $p_{l,e}$ drawn from a uniform distribution between 0.5 and 0.99. In this case, we find that the improvement follows a similar pattern to the case for uniform probabilities. The percentage improvement in this case follows similarly to the solution for a detection probability level of 0.7. This makes sense, since the mean of the uniform distribution between 0.5 and 0.99 is 0.745, which is approximately where the profile would lie if the analysis were performed using uniform probabilities.

While a uniform distribution was used as the probability sampling in our analysis, the problem formulation is general for any probability sampling distribution. In future work, the investigation of a probability model that follows an exponential decay as a function of distance from the optical flame detector and consider partial obstruction would be valuable towards improving estimates of this uncertainty.

Table 3.4 lists the numerical timing results of the iterative algorithm on Case A for the non-uniform case. The iterative algorithm converges to 0.1% gap in 6 or fewer iterations. The timing of solving the variable probability scenario is greater on average than the timing for any of the previous cases due to the increase in problem complexity with the non-uniform probabilities.

It is important to note that for the non-uniform, the commercial global optimizer BARON fails to generate a solution within the CPU time limit (3 hours) defined for this case study. While BARON was able to solve the problem formulations using uniform probability levels, for the case of flame detector placement problems with non-uniform probabilities, a tailored approach was required for efficient solution. Support for non-uniform probabilities more accurately reflects uncertainty in real-world scenarios as opposed to uniform probability levels.

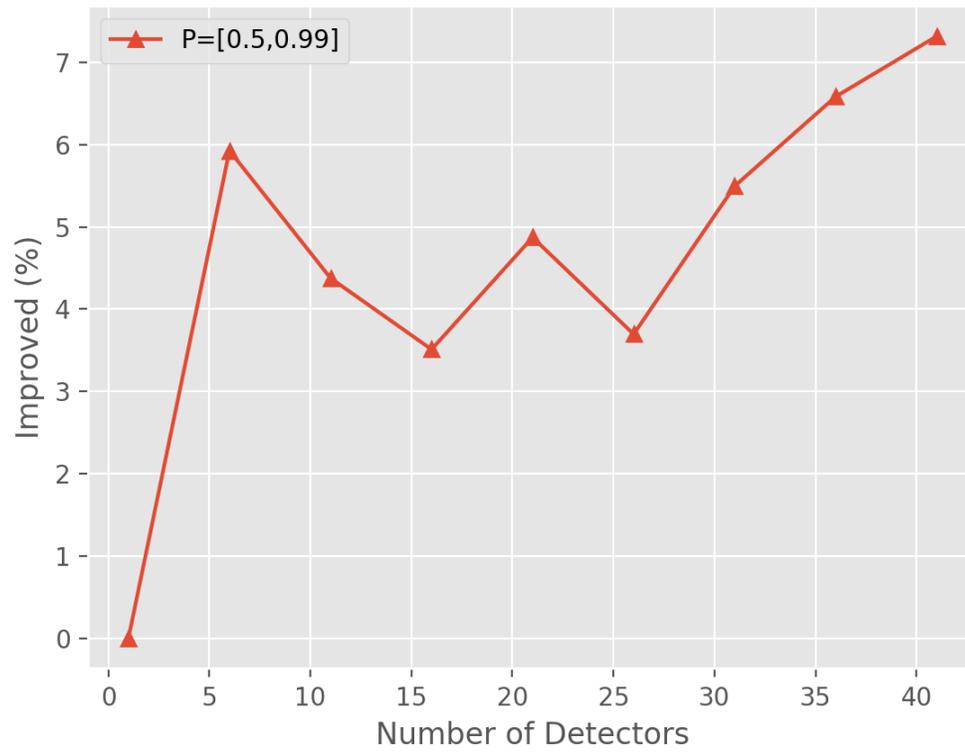


Figure 3.10.: Percentage improvement of P-FDP solutions against FDP solutions for Case A for probabilities of successful detection $p_{l,e} \in [0.5, 0.99]$

Table 3.4.: Numerical Performance of Iterative Algorithm for Case A with $p_{l,e} \in [0.5, 0.99]$

Budget k	Gap (%)	Iter	Time (s)
1	0.00	4	41
6	0.00	2	36
11	0.00	4	115
16	0.01	6	202
21	0.02	4	90
26	0.03	3	249
31	0.04	3	95
36	0.04	3	48
41	0.06	2	37

3.3.2 Results for Case B: Wellhead Drilling Facility

Figure 3.11 shows the percent improvement of the P-FDP solution across detection probability levels $p_{l,e}=0.5, 0.7, \text{ and } 0.9$ for Case B. In comparison to the results from Case A, those of Case B show markedly smaller improvements in using P-FDP over FDP. For a detection probability of 50%, the improvement reaches up to 2.5% while the same probability level for Case A reaches 20% improvement. The discrepancy lies in the facility geometry of the respective test cases, as well as the candidate detector locations. In Case A, the facility geometry was 2-dimensional, while the geometry of Case B is 3-dimensional. This increase in dimension allows detectors placed at higher vantage points in Case B to cover a much wider area than those in Case A. In addition, the geometry and candidate locations in Case B are highly symmetrical, which leads to very similar coverage areas with different placement configurations. The combination of these factors in Case B results in similar coverage capabilities using the FDP formulation, resulting in minor improvements when using

the P-FDP formulation instead. Nevertheless, the improvement in using the P-FDP formulation is non-trivial and remains a better placement method compared with the FDP formulation. Note that this improvement is more pronounced with the more complex geometry considered in Case C.

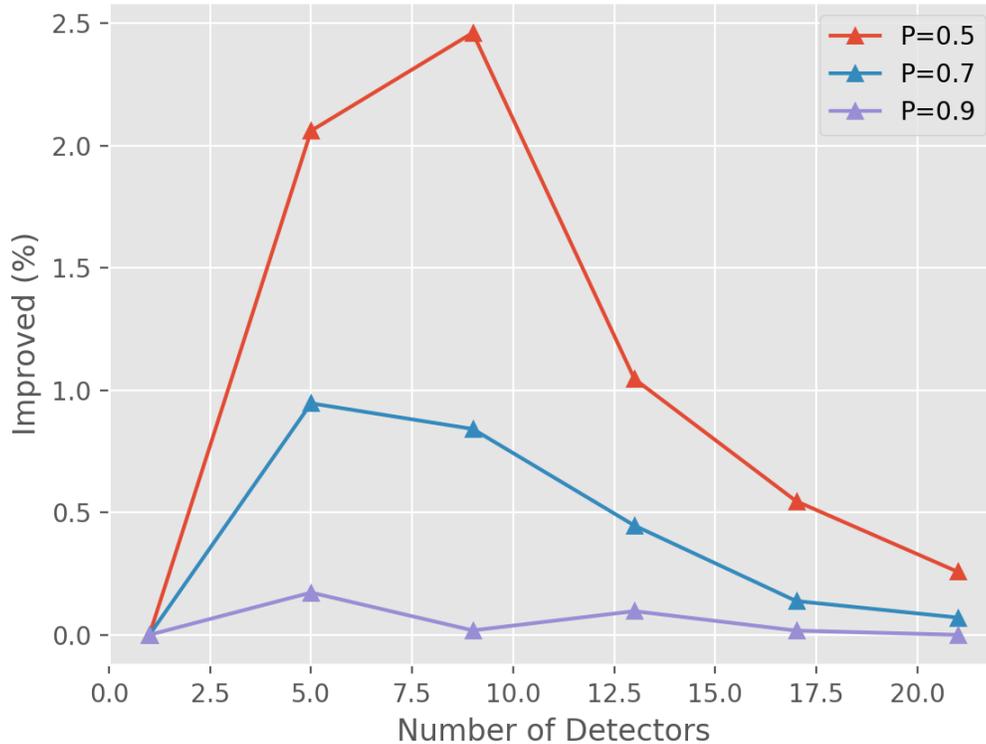


Figure 3.11.: Percentage improvement of P-FDP solutions against FDP solutions for Case B for probabilities of successful detection $p_{l,e}=0.5, 0.7,$ and 0.9 .

Table 3.5 lists the numerical timing results for the P-FDP formulation for Case B across three probability levels $p_{l,e}=0.5, 0.7,$ and 0.9 . The iterative algorithm converges to 0.1% optimality gap in 5 or fewer iterations. For higher detection failure levels ($p_{l,e}$ decreases), the problem becomes more challenging to solve and requires additional iterations. For low detection failure levels ($p_{l,e}=0.9$), the iterative algorithm converges in one iteration and requires the shortest computational time. Because Case

B represents a larger geometry, it requires significantly more time to solve compared to Case A. The iterative algorithm was executed without a fixed CPU time limit for evaluation purposes. It should be noted that the greatest timing coincided with the budget limits and probability levels that generated the greatest percentage improvement in the P-FDP solution compared with the FDP formulation. At one point, the algorithm required over 3 hours to evaluate budget limit $k=5$ with probability level $p_{l,e}=0.7$.

Note also that for budget limit $k=5$, and partially for $k=9$ and $k=13$, the commercial global solver BARON failed to solve the problem within the CPU time limit (3 hours). However, it is worthwhile to mention that we only considered the default options for BARON. Tuning option values may improve the performance of BARON in solving these problems. In addition, BARON may have been able to generate a solution for the missing values if given additional CPU time. Comparatively, the iterative algorithm was able to converge to 0.1% gap for nearly all of the evaluations in Case B in less time than BARON required.

Table 3.5.: Numerical Performance of P-FDP formulation comparing Iterative Algorithm and BARON for Case B

Budget k	$p_{l,e}=0.5$				$p_{l,e}=0.7$				$p_{l,e}=0.9$			
	Gap (%)	Iter	Alg	Time (s) BARON	Gap (%)	Iter	Alg	Time (s) BARON	Gap (%)	Iter	Alg	Time (s) BARON
1	0.00	5	3015	1121	0.00	4	1219	1117	0.01	1	298	1227
5	0.00	4	8056	-	0.00	4	-	-	0.03	1	1784	-
9	0.05	3	1709	-	0.00	4	9175	-	0.02	1	593	1118
13	0.04	4	3825	-	0.00	3	1675	1438	0.01	1	224	834
17	0.00	3	994	1502	0.06	2	501	1354	0.00	1	175	845
21	0.00	2	381	1294	0.00	2	383	1046	0.00	1	134	728

Figure 3.12 shows the comparison for Case B considering non-uniform detection probabilities. Here, again, $p_{l,e}$ is drawn from a uniform distribution between 0.5 and

0.99. In this case, we find that the performance improvement follows a similar pattern to the case for a fixed probability level of 0.7.

Table 3.6 lists the numerical results of the iterative algorithm on Case B for detection probability $p_{l,e}$ drawn from a uniform distribution between 0.5 and 0.99. The algorithm was able to converge to 0.1% gap in 4 or fewer iterations. The timing for the non-uniform case fares similar to that of the uniform case for the lower detection probability levels. Again, for the non-uniform case, the commercial global optimizer BARON was unable to generate a solution within the CPU time limit (3 hours) for any budget.

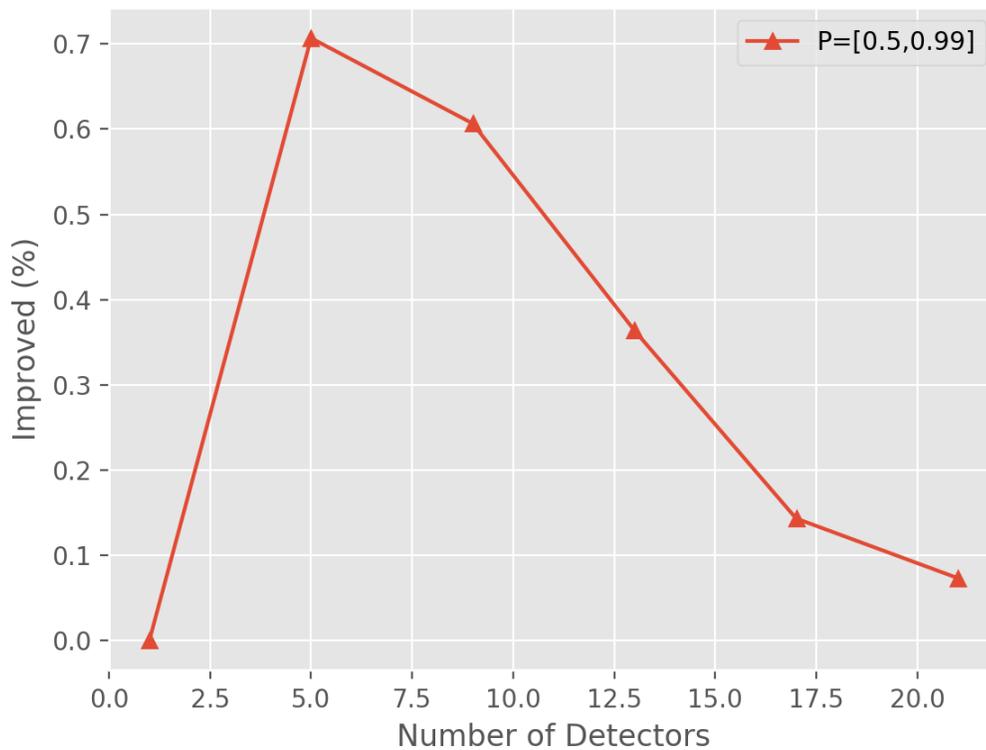


Figure 3.12.: Percentage improvement of P-FDP solutions against FDP solutions for Case B for probabilities of successful detection $p_{l,e} \in [0.5, 0.99]$

Table 3.6.: Numerical Performance of Iterative Algorithm for Case B with $p_{l,e} \in [0.5, 0.99]$

Budget	Gap (%)	Iter	Time (s)
1	0.00	4	2951
5	0.00	3	8831
9	0.00	3	5702
13	0.04	2	1029
17	0.03	2	494
21	0.00	2	474

3.3.3 Results for Case C: Offshore Hydrocarbon Separations Module

Figure 3.13 shows the percent improvement of the P-FDP solution across detection probability levels $p_{l,e}=0.5, 0.7, \text{ and } 0.9$ for Case C. The results for Case C show larger improvements in using P-FDP over FDP when compared with the results for Case B. For a detection probability of 50%, the improvement in Case C reaches up to 10%. Again, the discrepancy lies in the facility geometries. Case C is a highly non-symmetrical facility geometry which leads to marked improvements in the placements from P-FDP versus FDP. The results for Case C also show a significant drop-off in improvement percentage after placing about 5 detectors. Similarly, a drop-off can be seen in the Case B results after placing 10 detectors. This drop-off is due to the diminishing returns when placing additional detectors, previously encountered in the trade-off curve in Figure 3.8. As more detectors are placed, the coverage areas become more similar for both FDP and P-FDP solutions. Interestingly, this phenomenon is not encountered in the Case A results, possibly due to the fewer number of solutions that capture similar spaces and the larger present factor of obstructions in a 2-dimensional geometry.

The results in Case C also show an eventual decay of improvement as the number of detectors placed approaches the total set of possible candidate placements. In this case, the placement results for both the FDP and P-FDP solutions converge to the same selected detectors, leading to zero improvement in using P-FDP over FDP for detector budgets that approach the maximum number of considered locations.

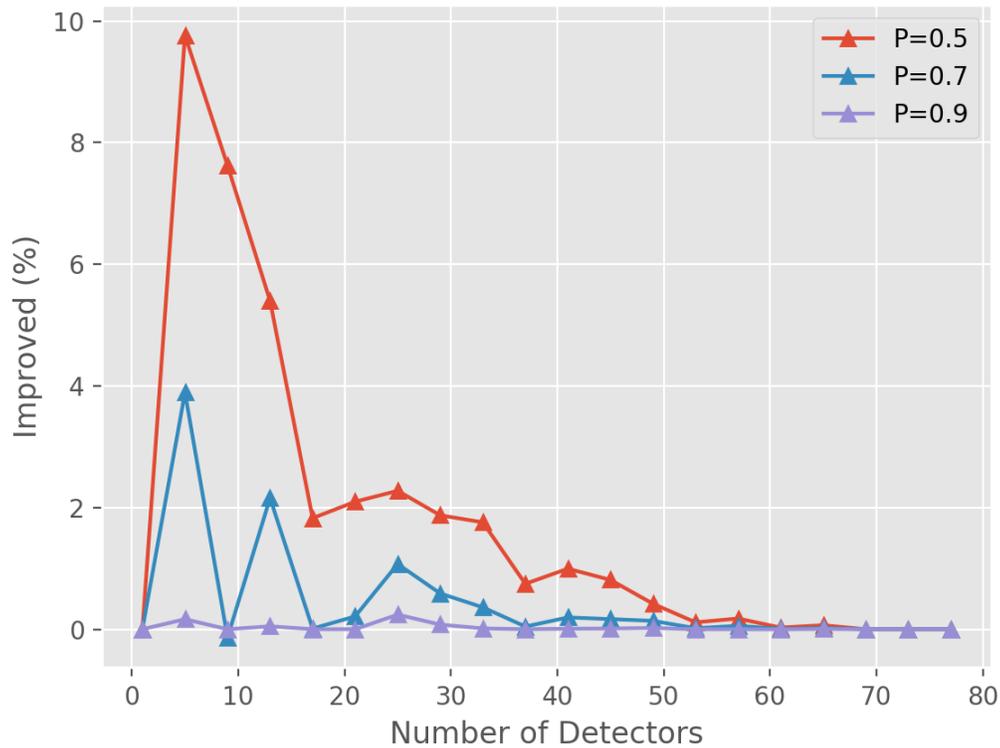


Figure 3.13.: Percentage improvement of P-FDP solutions against FDP solutions for Case C for probabilities of successful detection $p_{l,e}=0.5, 0.7,$ and 0.9 .

Table 3.7 lists the numerical timing results for the P-FDP formulation for Case C across three probability levels $p_{l,e}=0.5, 0.7,$ and 0.9 . The iterative algorithm converges to 0.1% optimality gap within 5 iterations. As detection failure levels increase ($p_{l,e}$ decreases), the problem becomes more challenging to solve and requires additional iterations. For the highest detection rates $p_{l,e}=0.9$, the algorithm converges in one

iteration and generally required less time than BARON. The algorithm required the longest evaluation times for intermediate budget levels $k=11, 21,$ and $31,$ especially as the detection probability decreased. Regardless, when detector budget limits are small or nearing the total available locations, the algorithm performs faster than BARON and represents the lowest solution times of the series of evaluations.

Table 3.7.: Numerical Performance of P-FDP formulation comparing Iterative Algorithm and BARON for Case C

Budget k	$p_{l,e}=0.5$				$p_{l,e}=0.7$				$p_{l,e}=0.9$			
	Gap (%)	Iter	Alg	BARON	Gap (%)	Iter	Alg	BARON	Gap (%)	Iter	Alg	BARON
1	0.00	3	384	442	0.00	3	384	471	0.01	1	88	621
11	0.00	5	7874	948	0.00	4	8704	889	0.02	1	1681	853
21	0.00	5	4027	879	0.03	3	2517	868	0.02	1	609	763
31	0.00	4	1180	727	0.00	3	1358	711	0.02	1	302	580
41	0.08	2	297	515	0.07	2	350	510	0.02	1	141	492
51	0.04	2	327	495	0.02	2	447	476	0.01	1	103	479
61	0.00	2	378	484	0.00	2	238	476	0.01	1	117	461
71	0.00	2	233	430	0.00	2	236	477	0.01	1	92	475

Figure 3.14 shows the comparison for Case C with non-uniform detection probabilities. As before, $p_{l,e}$ is drawn from a uniform distribution between 0.5 and 0.99. In this case, we find that the improvement follows a pattern similar to the case for a fixed probability level of 0.7.

Table 3.8 lists the numerical timing results of the iterative algorithm on Case C for detection probability $p_{l,e}$ drawn from a uniform distribution between 0.5 and 0.99. Note that the point of greatest improvement in expected coverage ($k=11$) from Figure 3.14 required the longest solution time (over 4 hours). As the detector budget increases and the number of available locations decreases, there is a significant reduction in solution time. In general, the timings for the non-uniform case match

closely with those of the uniform. Again, for the non-uniform case, the commercial global optimizer BARON fails to generate a solution within the CPU time limit (3 hours).

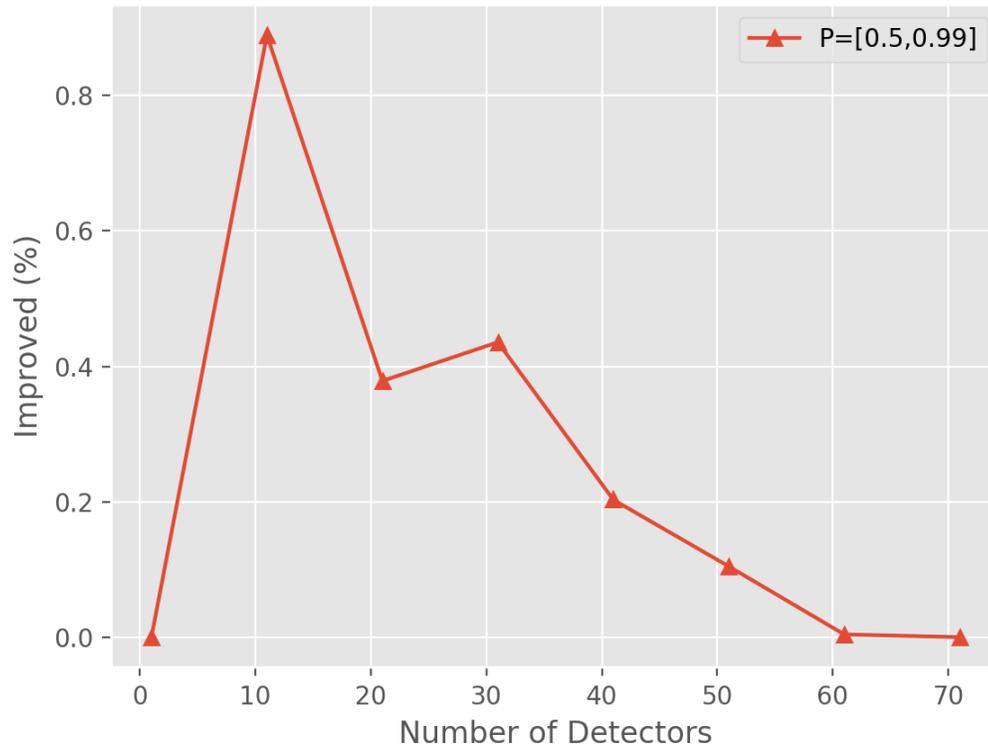


Figure 3.14.: Percentage improvement of P-FDP solutions against FDP solutions for Case C for probabilities of successful detection $p_{l,e} \in [0.5, 0.99]$

Table 3.8.: Numerical Performance of Iterative Algorithm for Case C with $p_{l,e} \in [0.5, 0.99]$

Budget	Gap (%)	Iter	Time (s)
1	0.00	3	431
11	0.00	5	-
21	0.02	3	3061
31	0.05	3	1914
41	0.06	2	383
51	0.02	2	376
61	0.00	2	351
71	0.00	2	262

3.3.4 Computational Scalability

The computational scalability of the P-FDP formulation and iterative solution algorithm was tested across a set of topologies with increasing entity size. First, a set of two-dimensional topology instances (in this case, 10 instances) were generated, where each topology instance had a different set of obstacles that were randomly placed within the respective topology. The set of instances were generated across a series of grid densities: 100, 2,500, and 10,000 entities (corresponding to 10 m, 50 m, 100 m square grid sizes). Then, using a set of 20 candidate detector locations that were placed in evenly-spaced locations along the boundaries of the topologies (consistent across all topology instances) and a budget of 5 candidate detectors, the P-FDP formulation was applied to each instance and grid size, assuming non-uniform detection probabilities $p_{l,e}$.

Table 3.9 lists the timing results of the scalability analysis for the set of 10 instances across three grid sizes, where solution times correspond to the total time required for the solver (in this case, Gurobi Optimizer) to converge to less than 0.1%

optimality gap using the iterative solution algorithm. Each timing entry in Table 3.9 represents an average over 10 stochastic realizations of the non-uniform detection probability $p_{l,e}$ sampled from a uniform distribution between 0.5 and 0.99 for the corresponding grid size and topology instance. Table 3.9 shows clearly that the solution time increases with respect to the grid size. The timing results vary across instances, with some instances performing significantly better than others. Figure 3.15 shows the normalized timings for each instance across the three grid sizes, where solution times are normalized to the time required to solve the 100 m² grid size. The figure shows that while, in general, the solution times for the instances scale with an increase in grid size, a few instances (specifically, instances 6 and 10) scaled better than the others.

Table 3.9.: Solution Times for Scalability Analysis with average times over 10 realizations for $p_{l,e} \in [0.5, 0.99]$

Instance	Grid Density		
	100 m ²	2,500 m ²	10,000 m ²
1	0.15	10.71	71.04
2	0.29	6.11	14.32
3	0.10	7.03	21.35
4	0.23	6.84	33.63
5	0.26	7.98	38.48
6	0.07	8.09	10.92
7	0.14	7.07	20.01
8	0.26	16.94	82.92
9	0.11	17.55	89.86
10	0.06	6.92	8.76

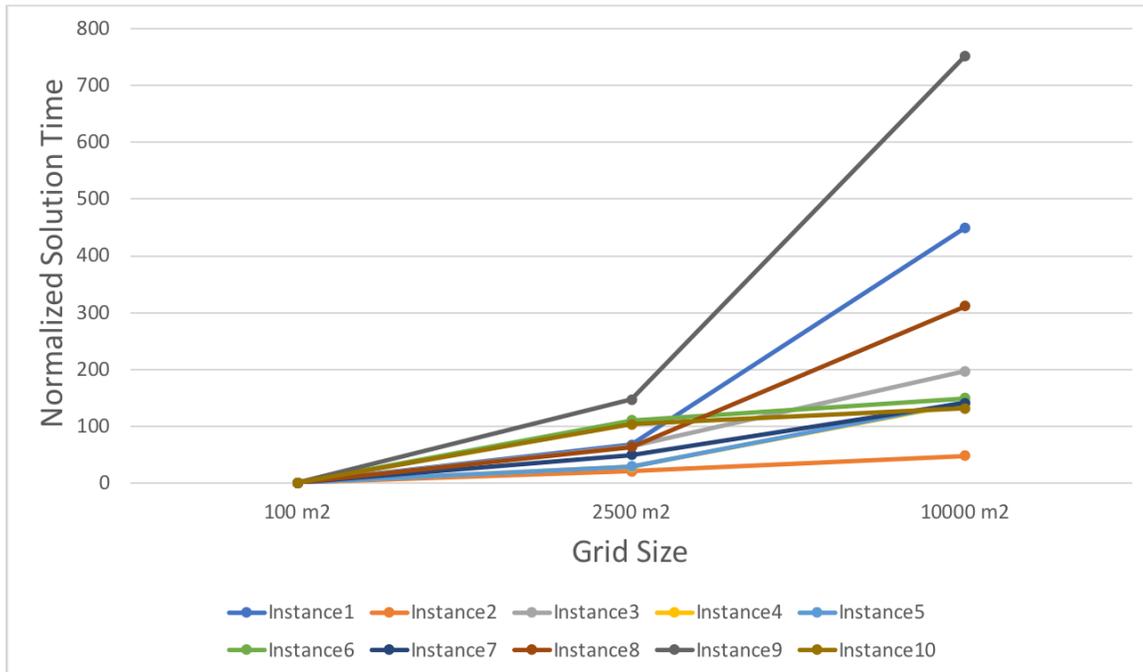


Figure 3.15.: Normalized solution time versus grid size with average times over 10 realizations for $p_{l,e} \in [0.5, 0.99]$

The solution times were also analyzed in relation to the objective function values for each solution. Checking the correlation between the solution times and objective function values yielded a Spearman's correlation coefficient of $r = -0.7$ with a p -value of 9.7×10^{-16} , indicating a strong negative correlation between the solution time and objective value. This relationship implies that problems with low expected coverage due to the location of obstructions take longer to solve than those with a similar number of obstructions with a comparatively higher expected coverage.

3.4 Conclusions

Our key conclusions in this work are as follows. First, our proposed P-FDP formulation results in consistent improvements in expected coverage compared with the FDP formulation when subjected to real-world conditions. The improvement of the P-FDP model over the FDP model reaches up to 20% for some cases when assuming

high detection failure probabilities, while the improvement decreases to 1-2% when assuming 90% success rate in detection. Recall that the probability of detection failure includes consideration of partial obstruction and camera range in addition to failure or unavailability of the detector itself. Second, the P-FDP model addresses redundancy in coverage, which leads to enhanced reliability in situations where particular detectors are in service or in need of maintenance. Third, the P-FDP formulation is agnostic to the specific process and underlying detection technology. Fourth, the P-FDP formulation when using non-uniform probabilities is effectively solved with our global solution algorithm that is tailored specifically for the flame detector placement problem and solves efficiently with reasonable computational effort.

Future avenues of study include evaluating the behavior of the P-FDP formulation with detection probabilities modeled as exponential decay over distance with respect to the field-of-view specifications of the candidate detectors, as well as investigation into a decomposition method for handling large-scale facility geometries (or high resolution discretizations). In addition, a combined optimization formulation for flame and gas detector placement with shared budget constraints would provide a holistic approach to process safety design.

4. PARAMETER ESTIMATION OF INFECTIOUS DISEASE TRANSMISSION PARAMETERS AND OPTIMAL SELECTION OF DISEASE DETECTION SITES

The work in this chapter focuses on first estimating both spatial coupling and seasonal transmission parameters with a series of nonlinear programming approaches for a network of subpopulations. Then, the estimated parameters are used to form the simulation basis for a stochastic programming framework based on the classical p -median formulation to determine the optimal locations for placing detection sites in order to quickly respond to infectious disease epidemics across a nationwide scale.

Despite large successes in the control of infectious diseases by improvements in drugs, nutrition, vaccines, and lifestyles, transmissible pathogens still pose an enormous threat to human mortality. Mortality rates attributed to infections are projected to stay at the current levels of 13 to 15 million deaths per year until at least 2030 (Heesterbeek et al., 2015). Long-standing infectious agents and newly emerging ones still pose a continuing challenge — for example, severe acute respiratory syndrome (SARS) and Middle Eastern respiratory syndrome (MERS) coronaviruses; measles and whooping cough; and the recent major Ebola outbreak in 2014-2015 (Heesterbeek et al., 2015).

The development of effective control methods for infectious disease spread involve multiple layers of complexity, including an increased understanding of the spatial transmission patterns in relation to susceptibility, infectiousness, and recovery rates, as well as the duration of pathogen immunity. In light of this complexity, computational tools and mathematical models have become essential to garner information for developing the evidence base for decision-making, and for understanding epidemiological patterns throughout the population. In order for these reliable models to

be constructed, however, the parameters within the models must be accurately estimated, which can be quite difficult, and can require efficient algorithms and solution procedures.

The seasonal transmission parameter measures the level of inter-population mixing on a temporal scale. Previous works have shown that the seasonal transmission parameter is correlated with school-term holidays (Word et al., 2012). The spatial transmission parameter (also called the spatial coupling parameter) captures the level of movement between subpopulations and provides important insights for understanding dynamics of diseases within and between communities. Spatial coupling between communities directly correlates with spatial synchrony, which is a key epidemiological phenomenon describing the likelihood of disease persistence within a metapopulation (Keeling and Rohani, 2002). In the absence of spatial synchrony, disease reintroduction events prevent localized extinctions from occurring and becoming permanent, leading to an increased probability of disease persistence. By quantifying the spatial coupling parameters for a metapopulation, we can better understand disease persistence within this network, as well as implement more informed control measures.

Currently, a gap exists for a model that uses disease fade-out information to better estimate city-to-city spatial coupling for many cities (instead of city-to-metapopulation). We attempt to fill this gap while also incorporating an efficient solution method for the parameter estimation.

The estimations in this work are demonstrated using records of measles outbreaks in 954 cities across England and Wales between 1944 and 1964. The nonlinear programming approaches for this large-scale estimation accurately reproduce longstanding parameter estimates, are readily scalable to larger problem sizes, and substantially reduce solution times compared to Markov-Chain Monte-Carlo methods. We present our work on efficient estimation of spatial coupling parameters over a series of four parameter estimation formulations.

First, we discuss the spatial coupling parameter in the form of a city-to-metapopulation relationship, which relates the city of interest (i.e. host city) with the larger

surrounding metapopulation (i.e. aggregate of cities within the same country/region, excluding the host city). We show a stochastic simulation by which the city-to-metapopulation spatial coupling parameter can be used to simulate transmission dynamics. Then we present a nonlinear programming framework to estimate the parameter for each city individually. A statistical, hazard-based Susceptible-Infectious-Recovered (SIR) model focusing on disease fade-out periods provides the basis for the estimation.

Second, we extend the prior formulation by applying a power law model to derive the relationship between the city-to-metapopulation spatial coupling parameter and the population size of the host city. We do so by solving the city-to-metapopulation parameter estimation problem over all cities in the metapopulation simultaneously, subject to the power law model.

Third, using the city-to-metapopulation spatial coupling parameter and corresponding estimation formulation as a basis, we propose a model extension to estimate the city-to-city spatial coupling parameter, which decouples cities from the metapopulation and describes the level of transmission between cities in a pairwise relationship. In this extension, we apply a gravity model to relate the city-to-city spatial coupling parameter as a function of the sizes and distances between the pairwise cities. The resulting nonlinear programming formulation solves for the city-to-city spatial coupling parameters over all cities in the metapopulation simultaneously, subject to the gravity model. We also extend the stochastic simulator proposed earlier to handle city-to-city spatial dynamics.

Fourth, we utilize the flexible nature of our parameter estimation framework to simultaneously estimate the seasonal transmission parameter in addition to estimating the gravity model-based, city-to-city spatial coupling parameter.

Finally, the stochastic simulator, in conjunction with the previously estimated city-to-city spatial transmission parameters found in PCC-GS, is then used to generate a family of epidemic scenarios for measles. This set of scenarios serves to quantify the uncertainty in epidemic origin and infection transmission across subpopulations,

and enters into a stochastic programming framework based on the classical p -median problem to identify optimal locations for the placement of infectious disease detection sites. The optimal placement framework has fast solution times and serves to help policymakers and healthcare professionals alike in determining epidemic control strategies on a countrywide scale.

The sections in this chapter are divided as follows. Section 4.1 provides a background on infectious disease modeling and parameter estimation, epidemiological models for transmission networks, and efficient solution algorithms. Section 4.2 focuses on the estimation and simulation of the city-to-metapopulation spatial coupling parameter, including a stochastic simulation used to simulate data and help validate the estimation in Section 4.2.1, an estimation framework for the city-to-metapopulation spatial coupling parameter in Section 4.2.2, and an extension to simultaneously estimate the city-to-metapopulation spatial coupling parameters for all cities in a region according to a power law model in Section 4.3. Section 4.4 presents an extension to estimate city-to-city transmission parameters using a gravity model basis and is further developed in Section 4.5 to include estimation of seasonal transmission parameters. Section 4.6 describes the development of a scenario-based stochastic programming method for the optimal selection of disease detection sites using the estimated transmission parameters to generate epidemic scenarios. Section 4.7 concludes the chapter with a discussion and summary of results and findings.

4.1 Background

The difficulty in modeling infection processes arises from the complex interactions between individuals (i.e. length of time of interaction, physical contact with individual or contaminated object, etc.) and the distribution of infection characteristics at various spatiotemporal scales. Additionally, the contact patterns between individuals are highly stochastic and are often described by one of two fundamental classes of models: agent-based models and compartment-based models (Heesterbeek et al., 2015). Agent-based models focus on network-level interactions that involve large

numbers of distinct individuals with unique sets of traits such as age, sex, location, immunity status. These types of models are often used to simulate epidemic growth, rather than endemic cycles (epidemic refers to outbreaks, whereas endemic refers to the cyclical recurrence of these outbreaks). They provide excellent frameworks for representing stochastic contact patterns, but the complexity of the networks and the high level of detail make them computationally impractical for parameter estimation on the endemic level.

On the other hand, compartment-based modeling categorizes individuals into their stage of infection, e.g. susceptible, infectious, exposed, etc. Since individuals are described by states and aggregated into compartments with same average characteristics, compartment-based models are more computationally tractable in estimation frameworks than agent-based models. For example, the susceptible-infectious-recovered (SIR) model is commonly used to represent the stages of measles (Hethcote, 2000). Unvaccinated newborns are susceptible to measles until they become infected, after which they may recover and gain life-long immunity to the disease. By discretizing the compartment traits to match the progressive stages of the specific disease, a compartment-based model can effectively model large-scale disease spread. Because of the computational efficiency and applicability of solution algorithms, we choose to use compartment-based modeling when formulating our problem.

An in-depth analysis on deriving spatial transmission information from the coupling parameter was first performed by Keeling and Rohani (2002). To study the interchange and exact mixing patterns of individuals between populations, Keeling and Rohani (2002) defined a two-city SIR model using spatial coupling as a key parameter to represent transmission. Incorporating the spatial coupling into both a phenomenological model (SIR with coupling parameter in incidence rate) and a mechanistic model (SIR with constraints for rates at which individuals leave and return to home cities, and ratio of time spent in the temporary to the permanent population), Keeling and Rohani (2002) showed that equivalent coupling levels produced identical movement rates.

Whereas Keeling and Rohani (2002) fixed the coupling parameter to study its effects, Xia et al. (2004) expanded the scope to estimate city-to-city spatial coupling for many cities using a gravity model. Xia et al. (2004) estimated the spatial coupling parameters using data from epidemic growth and minimizing one-step-ahead prediction error.

Given that spatial coupling is correlated to disease persistence, Bjørnstad and Grenfell (2008) hypothesized that more accurate estimates for the coupling parameters could be drawn from fade-out data (periods of zero new infections) rather than from epidemic growth trajectories. Bjørnstad and Grenfell (2008) defined a model based on hazards (the probability of a new infection occurring after a period of fade-out) and estimated the city-to-metapopulation spatial coupling for many cities.

While small-scale models may give important clues for disease spread, large-scale models are necessary for understanding infectious disease spread across metapopulations (e.g. multiple cities, regions, or countries). Previously, key parameters could be chosen to fit models using case data from recurring epidemics, as well as simulating a possible discrete set from qualitative techniques. However, due to the lack of comprehensive data (e.g. infectious and recovery rates are not observed for all individuals recorded and high under-reporting rates), many researchers use statistical methods to infer and reconstruct probable parameters for disease models.

To model stochasticity and reconstruct the numerous parameters in large-scale models, Cauchemez et al. (2011) presented a stochastic continuous time model using a statistical approach to analyze time-series epidemic data. To estimate the parameters, they used the Metropolis-Hastings Markov-Chain Monte-Carlo (MCMC) approach to sample from distributions to identify the parameters. MCMC approaches are notoriously computationally expensive, and the simulation performed by Cauchemez et al. (2011) required 20 hours per run for 20 total cities.

Recently, Jandarov et al. (2014) investigated the estimation of gravity model parameters for spatial transmission, building directly off the work of (Bjørnstad et al., 2002; Grenfell et al., 2002; Finkenstädt et al., 2002; Xia et al., 2004). Jandarov

et al. (2014) used an MCMC approach as well to estimate the spatial transmission parameters, making the assertion that likelihood-based inference methods are computationally intractable, and resorted to assuming a Gaussian process to emulate the gravity model. They showed computation times of 10 hours for the Gaussian emulation approach when compared to an MCMC approach, which required 3 days of computation time.

In light of the significant resource requirements to fulfill large-scale MCMC statistical sampling methods, research has been done using large-scale optimization methods to solve these difficult parameter estimation problems. Word et al. (2012) demonstrated an optimization approach to estimating transmission parameters with a nonlinear, continuous time model. Using an SIR model for measles as a framework, Word et al. (2012) proposed a nonlinear programming formulation involving the seasonal transmission parameter varying with time. Using existing measles case data from various sources, Word et al. (2012) showed estimation of the parameters matching closely to existing findings performed independently by others in the field. In addition, by using an interior-point solution method for the parameter estimation, a problem with over 16,000 variables and 15,000 constraints required only 190 seconds in CPU time to converge.

In this research, we adopt the SIR model for measles and simulate endemic disease cycles to show that incorporation of the spatial coupling parameter does indeed capture the essence of space-time disease dynamics. We then build upon the current models by first reproducing the hazard model proposed by (Bjørnstad and Grenfell, 2008) using the efficient solution approach proposed by (Word et al., 2012). We then extend both methods with a gravity model to decouple transport between the cities and estimate intercity disease transmission. Finally, the estimated parameters are used to form the simulation basis for a stochastic programming framework, based on the classical p -median formulation, to determine the optimal locations for disease detection sites in order to detect and mitigate infectious disease epidemics on a nationwide scale.

4.2 City-to-Metapopulation Spatial Coupling Parameter Estimation

We begin by focusing our scope to measles, an established, prototypical disease in quantitative epidemiology and population dynamics. Since measles is easily diagnosed and was historically subject to mandatory notification, medical personnel have retained excellent data records for measles outbreaks (Anderson and May, 1992). Additionally, measles-immune individuals have perfect cross-immunity to all variant strains. Data from measles outbreaks in England and Wales from 1944 to 1964 were used for the proposed models. These data contain sets for both a 60-cities subset and the full 954-cities set in England and Wales, with new reported cases every biweek for the 20 year period and before massive vaccination interventions. Due to misdiagnoses and failure to seek medical attention, an under-reporting bias of about 40-60% was characterized in these data (Bjørnstad et al., 2002). Given the cyclic nature of measles, the data holds key insights into regional persistence and spatial immigration due to the recurrent fade-out periods and epidemic periods.

4.2.1 Simulation

We first created a discrete-time, stochastic simulation of multi-city disease dynamics using spatial coupling parameters to show that real data can be modeled by spatial transmission parameters, and to help validate the estimation procedure. The SIR model represents the progressive stages of measles. A newborn falls into the “susceptible” (S) category before becoming “infected” (I), and finally “recovering” (R) into life-long immunity. This relationship is shown in Figure 4.1.

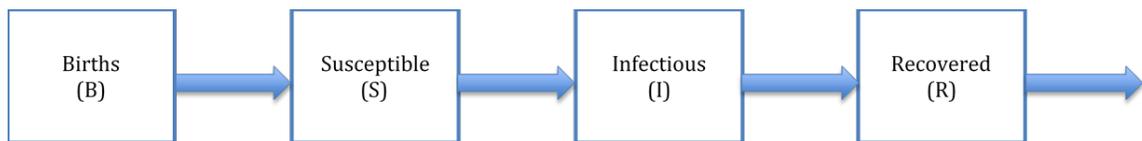


Figure 4.1.: SIR compartment-based model

The mean proportion of susceptibles was initialized at 4% for each city (although it can be as large as 8%) and the total population for each city was kept constant because the population variance over the estimation period for the England and Wales dataset was minimal (Bjørnstad et al., 2002). The stochastic SIR simulation is as follows.

$$S_{t+1,j} = S_{t,j} + B_{t,j} - I_{t+1,j} \quad \forall t \in T, j \in J \quad (4.1a)$$

$$I_{t+1,j} \sim \text{NegBin}(\lambda_{t,j}, I_{t,j} + l_{t,j}) \quad \forall t \in T, j \in J \quad (4.1b)$$

$$\lambda_{t,j} = \frac{\beta_{u,j}}{N_j} (I_{t,j} + l_{t,j})^\alpha S_{t,j} \quad \forall t \in T, j \in J, u = 1, \dots, 26 \quad (4.1c)$$

$$l_{t,j} \sim \text{Bin}(1, 1 - \exp(-c_j x_{t,j} \bar{y}_{t,j})) \quad \forall t \in T, j \in J \quad (4.1d)$$

$$x_{t,j} = \frac{S_{t,j}}{N_j} \quad \forall t \in T, j \in J \quad (4.1e)$$

$$\bar{y}_{t,j} = \frac{\sum_{k \neq j} I_{t,k}}{\sum_{k \neq j} N_k} \quad \forall t \in T, j \in J \quad (4.1f)$$

where $S_{t,j}$, $B_{t,j}$, and $I_{t,j}$ denote the number of susceptible individuals, newborns, and infected individuals in city j at time t , respectively; N_j is the total population of city j ; $\lambda_{t,j}$ describes the number of expected infected individuals; $l_{t,j}$ is the number of immigrant infections into city j at time t ; α is a correction parameter for discretization of the continuous infection process; c_j is the measure of spatial coupling of city j to all other cities; $x_{t,j}$ is the fraction of local susceptible individuals and $\bar{y}_{t,j}$ the fraction of non-local infected individuals. Equations 4.1 are indexed over all bi-weeks $t \in T$ and cities $j \in J$, with u denoting the indices of bi-weeks per year for the seasonal transmission rate $\beta_{u,j}$.

Equation 4.1a describes a balance over the susceptible individuals in the local region (newborns $B_{t,j}$ enter the susceptible population and infected individuals $I_{t,j}$ leave the susceptible population). The susceptible balance ignores mortality, since fatality from measles and child mortality rates have been historically very low in developed nations.

Equation 4.1b defines the stochastic epidemic growth, which is randomly sampled from a negative binomial distribution $\text{NegBin}(a, b)$ with expectation $a = \lambda_{t,j}$ and

clumping parameter $b=I_{t,j} + l_{t,j}$, previously described in (Bjørnstad and Grenfell, 2008).

Equation 4.1c describes the expected infections in a region and introduces stochasticity into the simulation. The seasonal transmission rate $\beta_{u,j}$ is a parameter that represents seasonal dynamics (i.e. more infections occur during the school year as opposed to the holiday season), and has been defined for the data set by Xia et al. (2004), as well as Bjørnstad and Grenfell (2008). A value of $\alpha=0.97$ was used in Equation 4.1c, which serves to correct for discretizing an inherently continuous process and account for nonlinearity in the growth of infections (Grenfell et al., 2002; Xia et al., 2004; Glass et al., 2003). The immigrant infections $l_{t,j}$ are randomly sampled for each city j and time t from a binomial distribution, $\text{Bin}(n, p)$, with $n=1$ trials and $p=(1 - \exp(-c_j x_{t,j} y_{t,j}^-))$ probability of success, where success is defined as contact between a local susceptible individual and a non-local infected individual. Equation 4.1d denotes this relationship and results in either 0 or 1 immigrant infections at every time step. This relationship was previously described in (Bjørnstad and Grenfell, 2008). Equations 4.1e and 4.1f describe the expressions for computing the fraction of local susceptible individuals and the fraction of non-local infected individuals, respectively.

The spatial coupling parameter c_j describes the disease transmission rate to city j in relation to the larger metapopulation (or pool of cities) (Bjørnstad and Grenfell, 2008; Finkenstädt et al., 2002; Grenfell et al., 2002; Finkenstädt and Grenfell, 2000). The spatial coupling parameter is critical to the understanding of inter-city relationships. This parameter will be the focus of the parameter estimation process. For the simulation, spatial coupling values from Bjørnstad and Grenfell (2008) are used.

Figure 4.2 compares the results of a single stochastic simulation against the real case data for the city of Cambridge, UK over the 20 year period. The seasonal patterns are consistent between the two data sets, and the magnitudes of the epidemic peaks are comparable. Most importantly, the simulation reproduces the dynamics of fade-out periods and re-introductions of infections into the city.

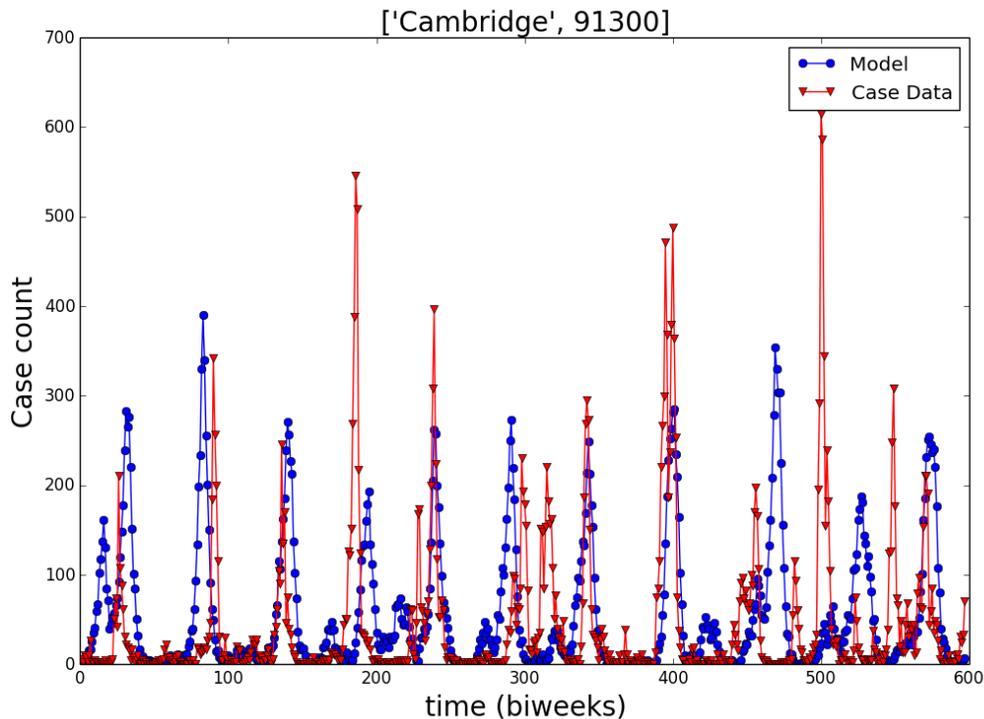


Figure 4.2.: Comparison between simulated data and case data for Cambridge, UK (population 91,300)

Note that no estimation has been performed in this portion. Values for the spatial coupling were calculated using the linear function $\log(c_j)=0.69 + 0.98 \log(N_j)$ described by Bjørnstad and Grenfell (2008). The initial susceptible fraction S_0 was assumed to be 4% and the reporting fraction r was 55%. The seasonal transmission rates for biweeks 1 through 26 are outlined in A.5.

4.2.2 City-to-Metapopulation Formulation

In this section, we discuss the development of an estimator for spatial transmission of measles among populations in the United Kingdom. The spatial coupling parameter c_j in the model formulation represents the level of mixing within and movement between one city and the metapopulation, as well as information about the importance

of the city in terms of disease spread potential (Finkenstädt and Grenfell, 2000; Xia et al., 2004; Bjørnstad and Grenfell, 2008).

The problem of estimating the city-to-metapopulation spatial coupling parameter can be formulated as a nonlinear programming problem on an independent, city by city basis as follows.

$$\max_{c_j} \quad \sum_{t \in H} \ln(1 - \bar{h}_{t,j}) + \sum_{t \in H^*} \ln(\bar{h}_{t,j}) \quad (4.2a)$$

$$\text{subject to} \quad \bar{h}(t, j) = \frac{\frac{\beta_{u,j}}{N_j} S_{t,j}}{1 + \frac{\beta_{u,j}}{N_j} S_{t,j}} (1 - \exp(-c_j x_{t,j} \bar{y}_{t,j})) \quad (4.2b)$$

$$\forall t \in T, u = 1, \dots, 26$$

$$\text{where} \quad H = \{t | I_{t-1,j} = 0, I_{t,j} = 0 \quad \forall t \in T\} \quad (4.2c)$$

$$H^* = \{t | I_{t-1,j} = 0, I_{t,j} > 0 \quad \forall t \in T\} \quad (4.2d)$$

where the variable for spatial coupling c_j is the only variable being estimated. Parameters N_j , $S_{t,j}$, $x_{t,j}$, $I_{t,j}$, and $\bar{y}_{t,j}$ are pre-computed from the data set, and the seasonal transmission parameters $\beta_{u,j}$ are initially defined using values from (Bjørnstad and Grenfell, 2008).

Constraint 4.2b defines the discrete-time hazard $\bar{h}_{t,j}$, which represents the probability that a disease fade-out period will end and a new epidemic will result. The hazard equation was first formulated by Bjørnstad and Grenfell (2008) based on the idea that more information could be gathered about spatial disease transmission in times of disease fade-out than in times of epidemic growth. This hazard is governed by the joint probability of spatial contact between a susceptible individual and a non-local infectious individual, along with the probability that an epidemic will result from such a contact. The derivation of this expression can be found in A.4.

The estimation is done by maximizing the log-likelihood of the fade-out data over the variable c_j , represented by the objective function in Equation 4.2a, for individual cities j within a metapopulation, one city at a time. The objective operates over a summation of the hazard for all periods in which the preceding biweek had no

new infected cases. Then, if the infected cases I at biweek t is 0, then the natural logarithm of the fade-out probability $(1 - \bar{h}_{t,j})$ is taken. Otherwise, if the infected cases I at biweek t is greater than 0, then the natural logarithm of the hazard probability $\bar{h}_{t,j}$ is taken. The sets H and H^* denote this relationship, and is shown in Sets 4.2c and 4.2d. Conceptually, this objective function models the behavior of fade-out periods, in which fade-out is the result of the probability that neither spatial contact nor epidemics occur (i.e. $1 - \bar{h}_{t,j}$), and the end of fade-out is the probability of both spatial contact and epidemics occurring (i.e. $\bar{h}_{t,j}$). To minimize instances of cities with little to no fade-out periods, the estimation was performed on cities with populations less than 250,000 people. Note that while the cities j of interest included only cities with populations fewer than 250,000, all cities were included in the calculation of the fraction of non-local infectious individuals $\bar{y}_{t,j}$.

4.2.3 Estimation Results

This problem was modeled using Pyomo (Hart et al., 2012, 2011b) and solved using IPOPT (Wachter and Biegler, 2006). Pyomo is an open-source optimization modeling language based on Python. IPOPT is an open-source, nonlinear interior-point solver. All estimations were performed on a Intel Xeon E5-2697v2 12-Core 2.70 GHz processor.

In order to validate the estimator, estimations were performed on 100 stochastic simulations of case data, generated from an initialized value for the spatial coupling parameter for each city. Figure 4.3 shows a sample distribution of the estimated coupling parameters for the city of Birkenhead (with an average population of 141,550 from 1944-1964) over 100 stochastic simulations.

The mean of the set of distributions do not deviate more than 4.5% of the initialized (true) coupling value for all cities in the data set. General observations show slightly higher deviations from the true value in cities with smaller populations than in those with larger populations.

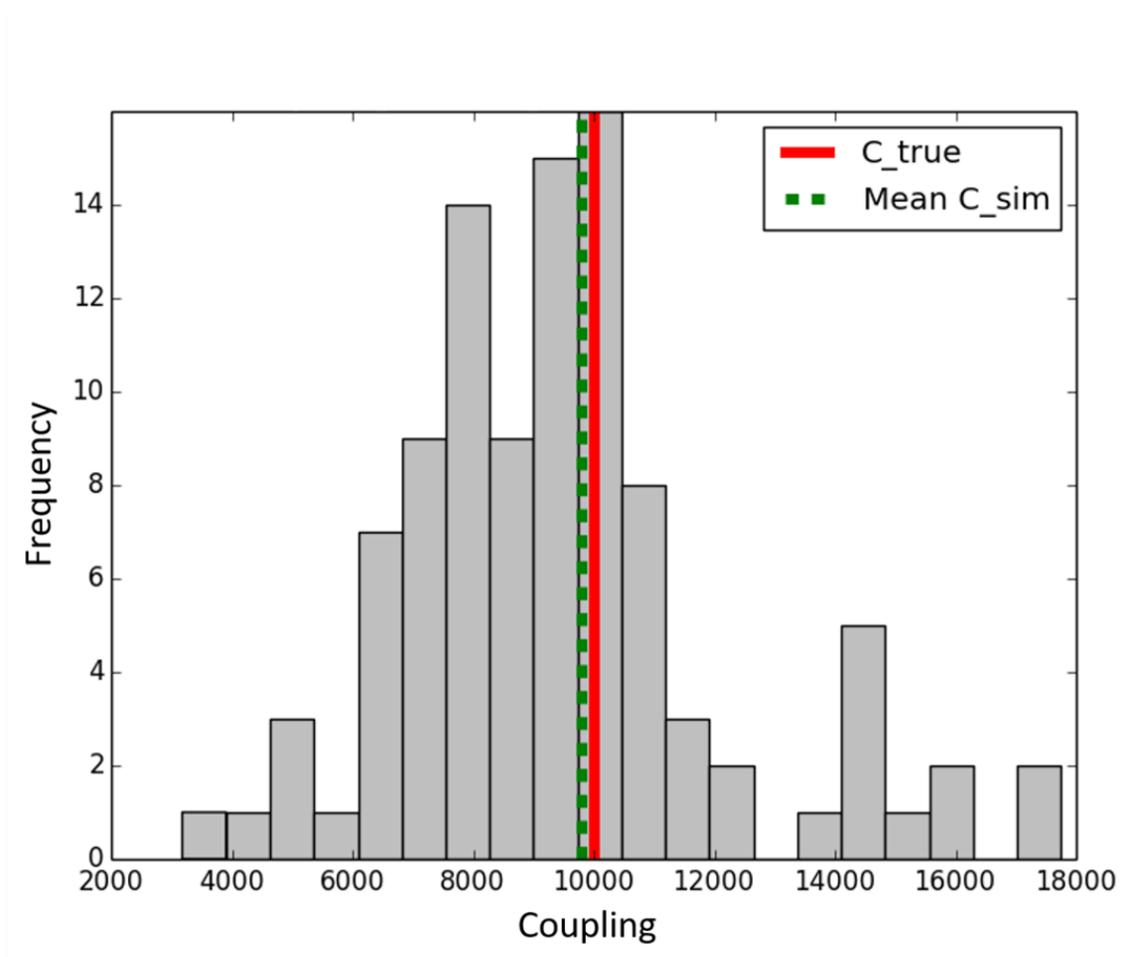


Figure 4.3.: Distribution of estimated spatial coupling parameters on 100 simulations for $c_j=10000$ (Example: Birkenhead, population 141,550)

Figure 4.4 shows the estimated values for the spatial coupling parameters in each of the 954 cities and towns from real case data, plotted against their respective population sizes. The color gradient ranges from cities with most fade-out periods (dark blue) to cities with least fade-out periods (white). Intuitively, the larger the city, the fewer fade-out periods in disease data since a larger population provides more individuals through which disease can spread. The estimated spatial coupling values for the cities match the values contained in findings reported in (Bjørnstad and Grenfell, 2008). The figure shows a strong log-linear correlation between the spatial coupling parameter and the population size, which reasons that spatial coupling is a

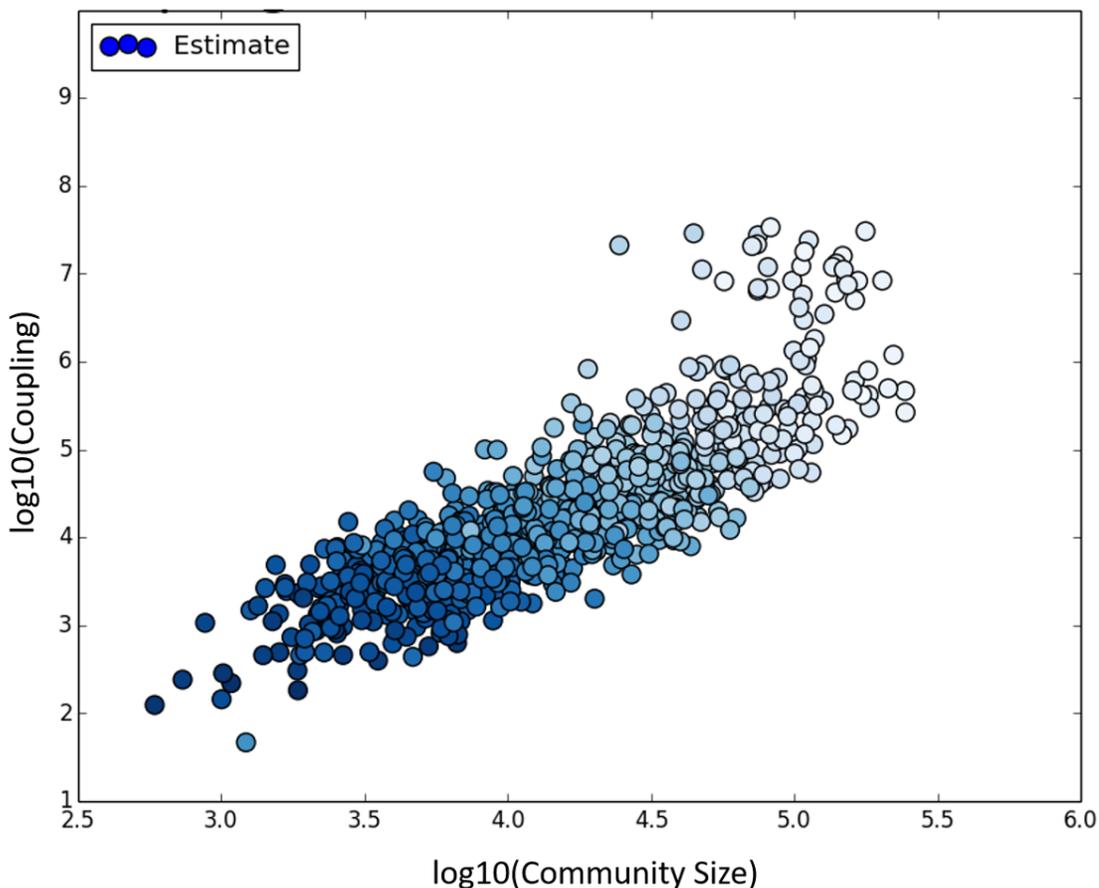


Figure 4.4.: Spatial coupling estimation for the 954 urban communities in England and Wales with reporting fraction $r=0.55$ and initial susceptible fraction $s_0=0.04$

direct function of the population and larger cities have a greater impact on spatial transmission within a metapopulation. Another observation from the figure is that the variance in the estimation is larger at high populations. Larger cities have fewer fade-out periods, and thus provide the estimator with less information. This is not surprising, since the objective function is a summation over the periods of fade-out. If little fade-out occurs for a given city, there is less information for the estimation.

Several key assumptions were made in this problem formulation, including the seasonal transmission rate $\beta_{u,j}$, the reporting fraction r , and the initial susceptible

fraction s_0 . The seasonal transmission rate $\beta_{u,j}$ was defined as a fixed set of values, derived from the England and Wales data set reported in previous works (Bjørnstad and Grenfell, 2008; Xia et al., 2004). The reporting fraction r , which affected the total reconstructed case data, was defined as 55% (from range of 40-60% (Bjørnstad and Grenfell, 2008)). The initial susceptible fraction s_0 was defined as 4% (from range of 4-8% (Bjørnstad and Grenfell, 2008)). These values were selected based on susceptible reconstruction results and other literature estimates (Word et al., 2012). A sensitivity analysis on the estimation was performed over the reporting fraction and initial susceptible fraction to understand the impact of these parameters and assumptions. Figures 4.5 and 4.6 show the sensitivity analyses on the two parameters. The scatter points denote parameter estimates from the PCM formulation in Equation 4.2, while the line corresponds to the linear regression ($c_j=1.99N_j^{0.98}$) from estimates performed in (Bjørnstad and Grenfell, 2008).

In Figure 4.5, a higher reporting fraction leads to an increase in the coupling values for cities. Conceptually, a higher reporting fraction produces fewer reconstructed infections in the real case data. This increase in reporting fraction also leads to a closer match with reported values from (Bjørnstad and Grenfell, 2008) (shown in the red line). It may be that the results reported in (Bjørnstad and Grenfell, 2008) were obtained using a greater reporting fraction than what was assumed here.

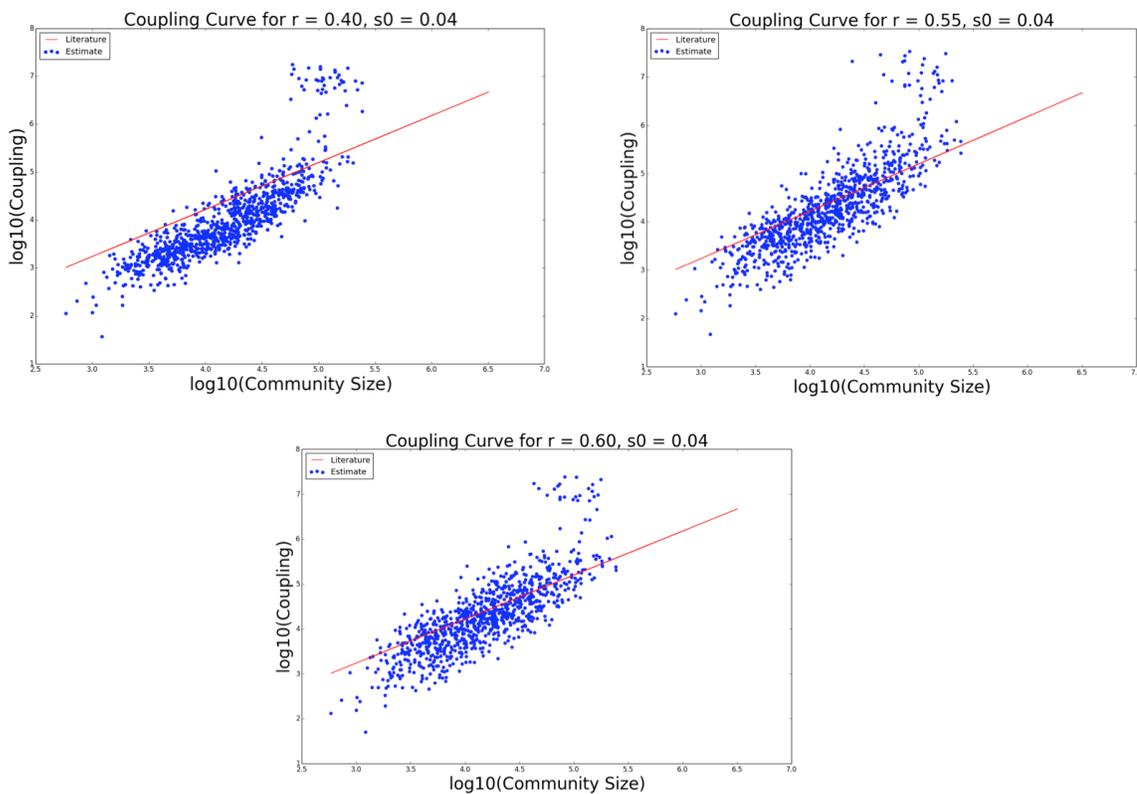


Figure 4.5.: Sensitivity analysis on reporting fraction r

Figure 4.6 shows a decrease in coupling values for all cities with an increase in the initial susceptible fraction. Conceptually, a greater initial susceptible fraction, along with a fixed number of infections, results in a lowered spatial coupling parameter because the infections have a lowered effect on the susceptible population.

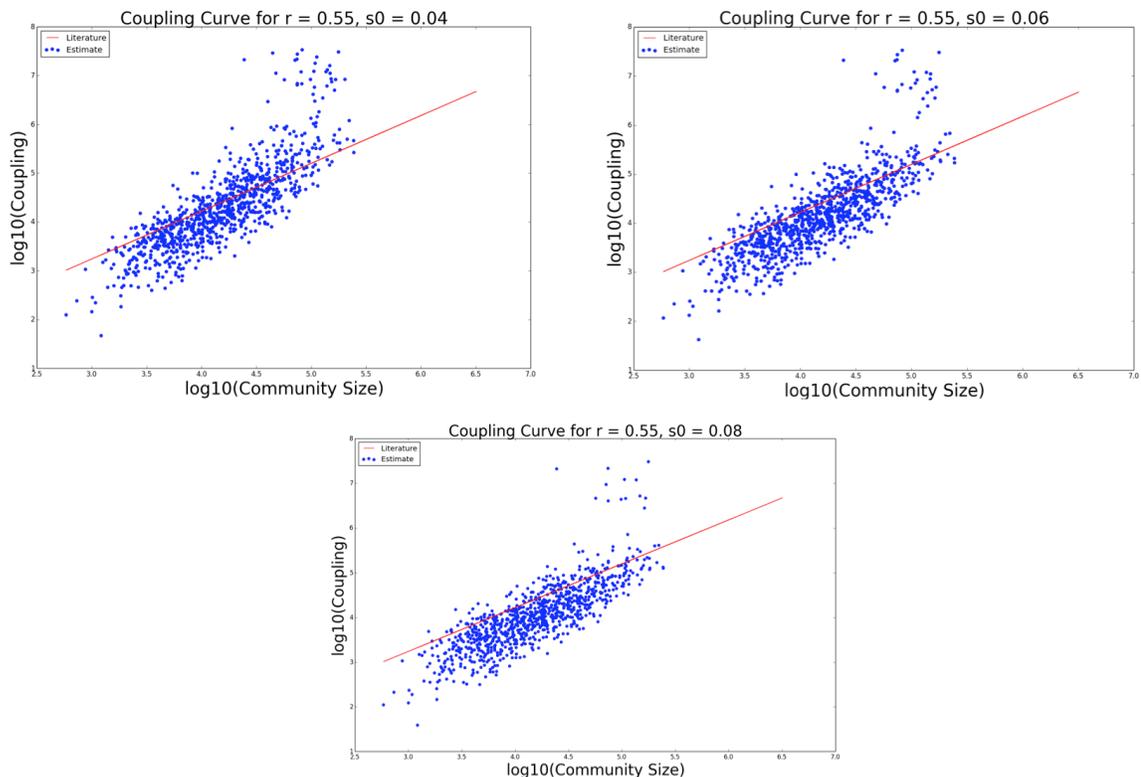


Figure 4.6.: Sensitivity analysis on initial susceptible fraction s_0

4.3 Simultaneous City-to-Metapopulation with Power Law Model

To understand the log-linear relationship between the coupling parameter and the population size, a power-law constraint is added to the nonlinear programming formulation, shown in Equation 4.3.

$$c_j = \theta N_j^\gamma \quad (4.3)$$

This constraint, after a logarithmic transformation, yields a linear form and will identify a consistent relationship between the city-to-metapopulation spatial coupling parameter c_j and the population size N_j of city j . Until now, the estimation procedure (including solution methods used by Xia et al. (2004) and Bjørnstad and Grenfell (2008), as well as the nonlinear programming formulation by Word et al. (2012)) had been done on an independent, city by city basis. Now we simultaneously estimate

θ, γ over all cities. This now significantly increases the problem size and the computational challenges. Nevertheless, we show that the nonlinear programming approach remains computationally efficient and the problem remains tractable. The modified formulation to consider simultaneous estimation of the city-to-metapopulation spatial coupling using the power-law model is shown below.

$$\max_{\theta, \gamma} \quad \sum_{j \in J} \left(\sum_{t \in H} \ln(1 - \bar{h}_{t,j}) + \sum_{t \in H^*} \ln(\bar{h}_{t,j}) \right) \quad (4.4a)$$

$$\text{subject to} \quad \bar{h}(t, j) = \frac{\frac{\beta_{u,j}}{N_j} S_{t,j}}{1 + \frac{\beta_{u,j}}{N_j} S_{t,j}} (1 - \exp(-c_j x_{t,j} \bar{y}_{t,j})) \quad (4.4b)$$

$$\forall t \in T, j \in J, u = 1, \dots, 26$$

$$c_j = \theta N_j^\gamma \quad \forall j \in J \quad (4.4c)$$

$$\text{where} \quad H = \{t | I_{t-1,j} = 0, I_{t,j} = 0 \quad \forall t \in T\} \quad (4.4d)$$

$$H^* = \{t | I_{t-1,j} = 0, I_{t,j} > 0 \quad \forall t \in T\} \quad (4.4e)$$

The objective in Equation 4.4a now operates as a summation over all cities $j \in J$. The hazard probability and corresponding fadeout sets H and H^* remain the same from Equations 4.2b, 4.2c, and 4.2d.

The results with the simultaneous estimation formulation yielded $c_j = 0.91 N_j^{0.99}$ for an $\theta = 0.91$ and $\gamma = 0.99$, which shows close agreement with the regression performed in (Bjørnstad and Grenfell, 2008) (i.e. $c_j = 1.99 N_j^{0.98}$). This comparison is shown in Figure 4.7, where the red line shows the regression reported in (Bjørnstad and Grenfell, 2008) and the blue points are the results from our simultaneous estimation.

The numerical timing results for the simultaneous estimation approach are summarized in Table 4.1. Given over 300,000 constraints and 300,000 variables for the estimation using the 954-city dataset, the solver converged in under three minutes. With the computational efficiency of this problem formulation, scale-up to larger

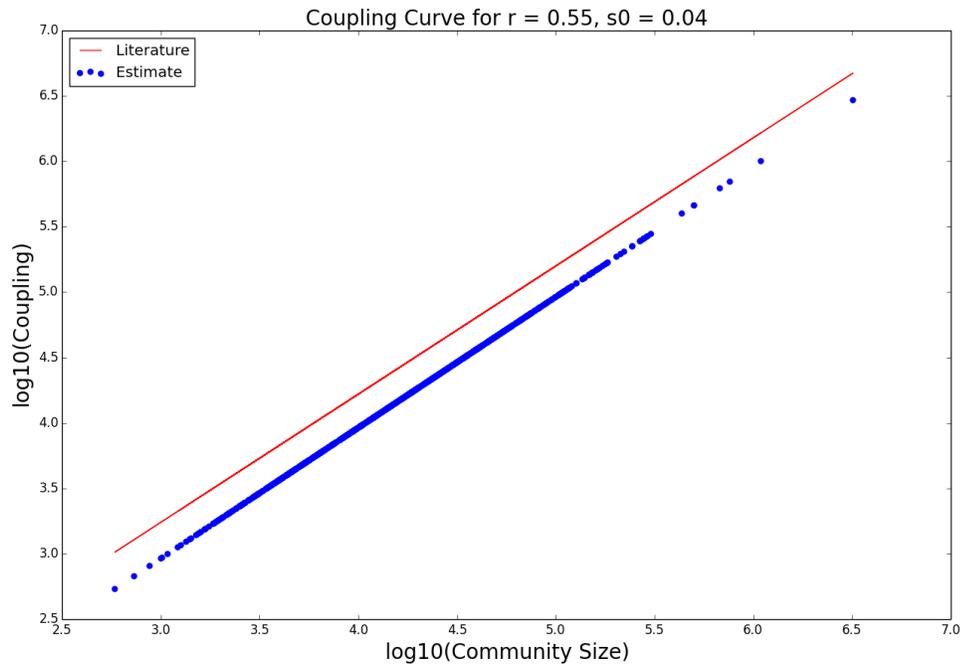


Figure 4.7.: Comparison of linear regressions between estimated parameters (in blue) and findings in (Bjørnstad and Grenfell, 2008) (red line) for reporting fraction $r=0.55$ and initial susceptible fraction $s_0=0.04$

data sets and additional constraints is very viable, which will be explored in the next sections.

Table 4.1.: Numerical Timing Results

Data set	Iterations	Solution Time (s)	Constraints	Variables
60 cities	12	2.21	6530	6532
954 cities	74	126.10	310934	310936

4.4 City-to-City Spatial Coupling Parameter Estimation

In this section, we transition to estimating spatial transmission between cities on a city-to-city, pairwise basis. While the previous formulation provided a model for c_j ,

the physical significance of the city-to-metapopulation spatial coupling parameter c_j lies only in its relationship to the metapopulation. In order to understand the spatial transmission between respective cities in a disaggregated manner, however, we must further specify the spatial coupling parameter as $c_{j,k}$, where j denotes the local city of interest and k denotes a nearby, non-local donor city (note that “donor” refers to the uni-directional transmission from city k to city j). Redefining the spatial coupling changes the problem formulation as well. In this case, Equation 4.2b is now replaced by Equation 4.5. Since the coupling parameter is now dependent on the interactions between individual cities, it enters into the summation that previously yielded the non-local infected fraction $y_{t,j}^-$.

$$\bar{h}(t, j) = \frac{\frac{\beta_{u,j}}{N_j} S_{t,j}}{1 + \frac{\beta_{u,j}}{N_j} S_{t,j}} \left(1 - \exp\left(-x_{t,j} \sum_{k \neq j} \frac{I_{k,t} c_{j,k}}{N_k}\right) \right) \quad (4.5)$$

A gravity model is adopted for this new spatial coupling parameter, introducing an additional set of constraints on $c_{j,k}$, shown by Equation 4.6.

$$c_{j,k} = \frac{\exp(\alpha_1) N_j^{\alpha_2} N_k^{\alpha_3}}{d_{j,k}^{\alpha_4}} \quad (4.6)$$

The gravity model has been used extensively to relate magnitude of travel with population sizes in transportation models (Xia et al., 2004). The term *gravity* comes from the analogous form of the gravity equation between two large, physical bodies. The model has parameters N_j , the population at city j ; N_k , the population at city k ; and $d_{j,k}$, the distance between cities j and k . The exponents in the model, α_1 , α_2 , α_3 , and α_4 , determine the relationship between the parameters in the model. Variables α_2 , α_3 , and α_4 are assumed to be positive so that the transmission between the cities increases with the populations of the cities and decreases with the distance between the cities, while α_1 is assumed to be negative to reflect the appropriate relationship

in the exponential term. The modified formulation for the simultaneous estimation of the pairwise spatial coupling $c_{j,k}$ using a gravity model (PCC-G) is as follows.

$$\begin{aligned}
& \max_{\alpha_i} && \sum_{j \in J^*} \left(\sum_{t \in H} \ln(1 - \bar{h}_{t,j}) + \sum_{t \in H^*} \ln(\bar{h}_{t,j}) \right) \\
\text{subject to} &&& c_{j,k} = \frac{\exp(\alpha_1) N_j^{\alpha_2} N_k^{\alpha_3}}{d_{j,k}^{\alpha_4}} \\
&&& \bar{h}_{t,j} = \frac{\frac{\beta_u}{N_j} S_{t,j}}{1 + \frac{\beta_u}{N_j} S_{t,j}} \left(1 - \exp\left(-x_{t,j} \sum_{k \neq j} \frac{I_{k,t} c_{j,k}}{N_k}\right) \right) \\
&&& \forall t \in T, j \in J \quad \text{and} \quad u = 1, \dots, 26
\end{aligned} \tag{4.7}$$

Correspondingly, an extension to the stochastic SIR simulation previously defined in Equations 4.1a - 4.1f is also developed using the gravity model to reflect city-to-city transmission, as follows.

$$S_{t+1,j} = S_{t,j} + B_{t,j} - I_{t+1,j} \quad \forall t \in T, j \in J \tag{4.8a}$$

$$I_{t+1,j} \sim \text{NegBin}(\lambda_{t,j}, I_{t,j} + l_{t,j}) \quad \forall t \in T, j \in J \tag{4.8b}$$

$$\lambda_{t,j} = \frac{\beta_{u,j}}{N_j} (I_{t,j} + l_{t,j})^\alpha S_{t,j} \quad \forall t \in T, j \in J, u = 1, \dots, 26 \tag{4.8c}$$

$$l_{t,j} \sim \text{Bin}\left(1, \left(1 - \exp\left(-x_{t,j} \sum_{k \neq j} \frac{I_{k,t} c_{j,k}}{N_k}\right)\right)\right) \quad \forall t \in T, j \in J \tag{4.8d}$$

$$c_{j,k} = \frac{\exp(\alpha_1) N_j^{\alpha_2} N_k^{\alpha_3}}{d_{j,k}^{\alpha_4}} \tag{4.8e}$$

$$x_{t,j} = \frac{S_{t,j}}{N_j} \quad \forall t \in T, j \in J \tag{4.8f}$$

Equations 4.1a, 4.1b, 4.1c, and 4.1e stated previously are maintained in Equations 4.8a, 4.8b, 4.8c, and 4.8f, respectively. Equation 4.1d is expanded here in the Equation 4.8d to consider the re-derived expression representing the probability of contact between a nonlocal infected and a local susceptible individual (expressed in Equation 4.5). The gravity model is included in the simulation to model the city-to-city transmission parameter $c_{j,k}$ in Equation 4.8e.

Given the distance between cities, the size of the country of interest, and the assumption that individuals typically do not have daily commutes over large distances,

the model only assumes the pairwise existence of spatial coupling for cities within a certain radius of one another. That is, city-to-city pairings were not considered for cities that fell outside the distance boundary parameter. For our tests we used a distance boundary of 60 miles.

The estimation results for the gravity model for all cities in the 954-city dataset are $\alpha_1 = -2.175$, $\alpha_2 = 1.127$, $\alpha_3 = 0.054$, and $\alpha_4 = 1.056$, listed in Table 4.2.

Table 4.2.: Estimates for α parameters in gravity model (954-city dataset, distance boundary of 60 miles)

Parameter	Value
α_1	-2.175
α_2	1.127
α_3	0.054
α_4	1.056

4.5 Simultaneous City-to-City Spatial Coupling Parameter Estimation with Seasonal Transmission

In this section, we extend the formulation to simultaneously estimate the seasonal transmission parameters, β_u , in addition to the parameters in the gravity model. We also compare these estimates with existing findings performed by others in the field. We then compare epidemic predictions from our model estimates with those performed by others and with actual reported cases.

The seasonal transmission parameters had been estimated previously by (Word et al., 2012; Xia et al., 2004; Finkenstädt et al., 2002; Bjørnstad et al., 2002; Grenfell et al., 2002), which are defined in Appendix A.5 and were used for the estimation done thus far. However, we now allow this parameter to be estimated and introduce a least squares term into the objective to minimize the difference between number of

predicted and recorded infections. The extended formulation is shown in Formulation 4.9.

$$\max_{\alpha_i, \beta_i} \sum_{j \in J^*} \left(\sum_{t \in H} \ln(1 - \bar{h}_{t,j}) + \sum_{t \in H^*} \ln(\bar{h}_{t,j}) \right) - \rho \sum_{j \in J} \sum_{t \in T} (I_{t,j} - \hat{I}_{t,j})^2 \quad (4.9a)$$

$$\text{subject to } c_{j,k} = \frac{\exp(\alpha_1) N_j^{\alpha_2} N_k^{\alpha_3}}{d_{j,k}^{\alpha_4}} \quad (4.9b)$$

$$\bar{h}_{t,j} = \frac{\frac{\beta_u}{N_j} S_{t,j}}{1 + \frac{\beta_u}{N_j} S_{t,j}} \left(1 - \exp\left(-x_{t,j} \sum_{k \neq j} \frac{I_{k,t} c_{j,k}}{N_k}\right) \right) \quad (4.9c)$$

$$\hat{I}_{t,j} = \frac{\beta_u S_{t,j}}{N_j} I_{t-1,j}^\gamma \quad (4.9d)$$

$$\forall t \in T, j \in J \quad \text{and} \quad u = 1, \dots, 26$$

Constraint 4.9d is a one-step-ahead method to predict the number of new infected cases I , where \hat{I} is the predicted value. The parameter ρ is used to weight the multiple objectives and is set to 1 for the subsequent analyses.

The estimation results for the seasonal transmission parameter β (normalized) are shown in Figure 4.8. Figure 4.8 compares our estimation results with existing estimates performed in (Xia et al., 2004) and (Finkenstädt et al., 2002). Our estimates match seasonality patterns well with the existing estimates; namely, peaks around the 4th, 12th, and 21st biweeks and troughs around the 9th, 18th, and 25th biweeks. In addition, the estimation reproduced the relative extrema in values and slopes. The α_i parameters did not change significantly from previous estimates (shown in Table 4.2).

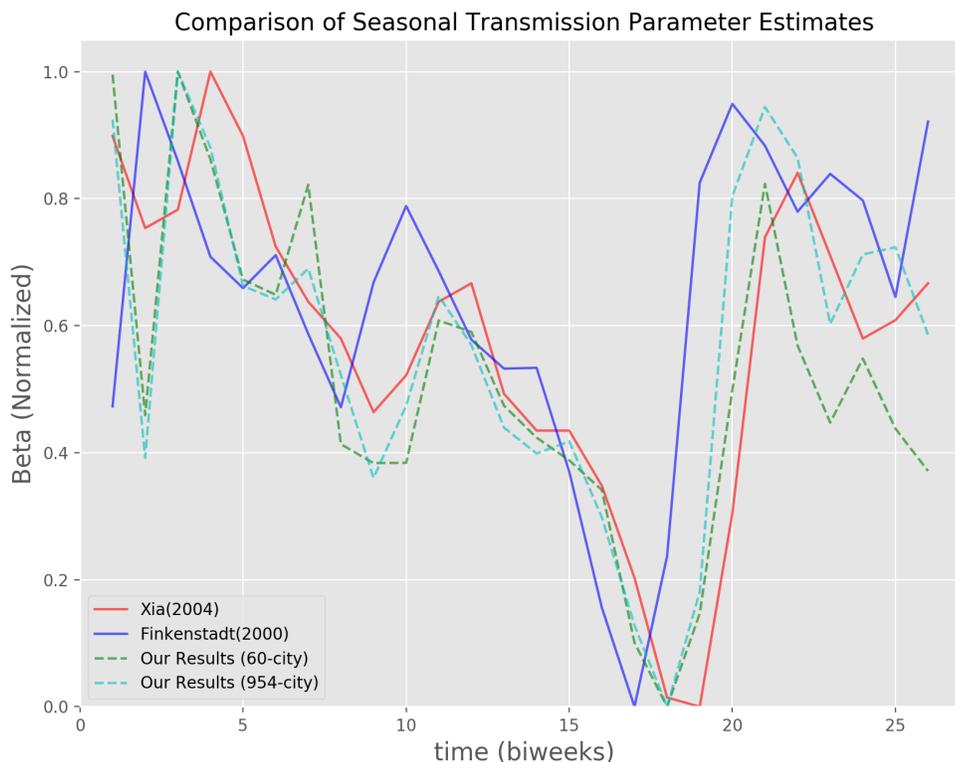


Figure 4.8.: Comparison of Seasonal Transmission Parameter Estimates for normalized β over time (biweeks) between values estimated in (Xia et al., 2004) and (Finkenstädt et al., 2002) with our results for the 60-city and 954-city datasets

Finally, we can combine the parameter estimates up to this point (α_i from Table 4.2 and $\beta_{u,j}$) with the pairwise city simulator in Equations 4.8 to simulate time-varying infected case data. In Figures 4.9 and 4.10, we compare the frequency of outbreaks in 10 simulations with those from the original measles case data for England and Wales from 1944-1964 for the cities of Cambridge (est. population 105,950) and Teignmouth (est. population 10,540), respectively. The outbreak size is reported in terms of the infected fraction, which is the number of cases divided by the population size of the city. For the larger city of Cambridge, we see remarkable agreement in the mean and shape of the distributions. The scale of the fade-out periods with

respect to the epidemic periods are also close, showing that the parameter estimation formulation and simulator can indeed capture and recreate fade-out dynamics that are in agreement with the data. The similarities in the long tail as cases increase show that the proposed methods can also simultaneously match the scales of epidemics at the relative infrequency of their occurrence. This strong agreement in distributions is consistent in the findings for cities with populations larger than about 75,000 people. As city sizes decrease, the simulations are less able to recreate outbreak frequencies that match the original case data, as shown in Figure 4.10. While the mean of the simulated distribution of outbreaks matches well with that of the original case data, the simulation is no longer able to approximate the scale and frequency of epidemics as well as it can for larger cities. As population sizes decrease, the number of fade-out periods overwhelm the number of periods with infected cases, which ultimately leads to a lower calculated probability of hazard or reintroduction of the disease. As a result, the simulation generates more fade-out periods and fewer outbreaks. The sporadic and sparse nature of endemic cycles in small cities (populations fewer than 35,000 to 55,000) makes simulations of infections in these cities notoriously difficult. Nevertheless, the relatively close approximation of the distribution of outbreaks that the proposed methods are able to produce even for small cities is encouraging and provides credence to use the methods in decision-making and control strategies.

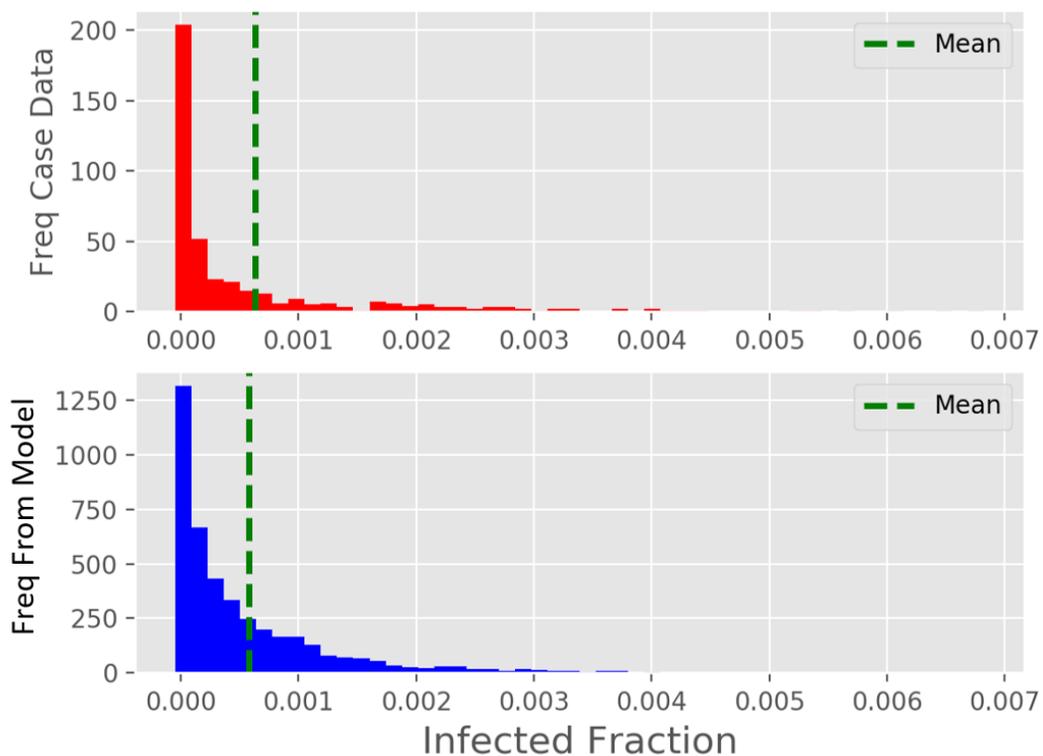


Figure 4.9.: Comparison of outbreak frequency between case data for Cambridge (est. population 105,950) and simulated outbreaks (10 simulations) using spatial and seasonal parameter estimates

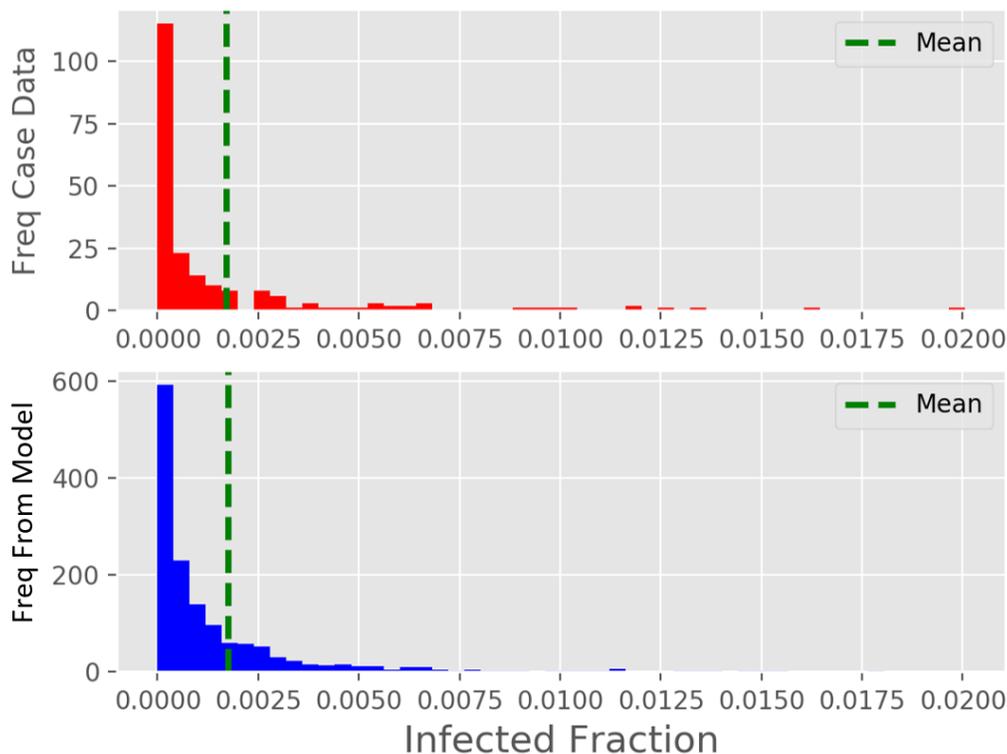


Figure 4.10.: Comparison of outbreak frequency between case data for Teignmouth (est. population 10,540) and simulated outbreaks (10 simulations) using spatial and seasonal parameter estimates

With confidence in the estimation of spatial coupling parameters for infectious disease transmission, we can use the insights from the work done so far to propose an optimal selection method for the location of disease detection sites across a metapopulation.

4.6 Optimal Selection of Disease Detection Sites

In this section, we build off of the classical p -median approach that has been used extensively in the fields of operations research and location science Laporte et al. (2015) to identify optimal locations for disease detection sites while accounting for the uncertainty in sources of epidemics.

Healthcare resources, including but not limited to infectious disease teams, appropriate medical supplies, and support personnel, can often be in short supply when an infectious disease outbreak occurs. The fast spread of infectious diseases puts additional pressure on epidemic response teams and stretches healthcare resources further.

While one solution would be to increase the total amount of available resources across all possible healthcare locations, this tactic is typically cost prohibitive and may not be reasonable given the vast number of treatments required to handle the known roster of infectious diseases. Another solution, which will be discussed in this section, is the optimal selection of infectious disease detection sites. A disease detection site can be defined as the placement or allocation of additional healthcare resources in order to maximize the chances of early detection of a specific (or related family of) infectious disease(s). This solution aims to provide a more reasonable allocation of resources while also aiming to detect infectious disease epidemics early to implement subsequent control measures.

For the work described here, we have chosen to evaluate the selection of disease detection sites for the targeted detection of measles. Note that while the optimal selection method is designed to be agnostic to the specific disease, the scenario generation method discussed here is disease-specific and should be modified according to the corresponding spatial and seasonal transmission patterns of the targeted disease.

4.6.1 Problem Formulation

The impact formulation, based on the p -median facility location problem, for the optimal selection of disease detection sites is described as follows.

$$\underset{x,s}{\text{minimize}} \quad \sum_{a \in A} \alpha_a \sum_{i \in L_a} d_{ai} x_{ai} \quad (4.10a)$$

$$\text{subject to} \quad \sum_{i \in L_a} x_{ai} = 1 \quad \forall a \in A \quad (4.10b)$$

$$\sum_{i \in L} s_i \leq K \quad (4.10c)$$

$$x_{ai} \leq s_i \quad \forall a \in A, i \in L_a \quad (4.10d)$$

$$s_i \in \{0, 1\} \quad \forall i \in L \quad (4.10e)$$

$$0 \leq x_{ai} \leq 1 \quad \forall a \in A, i \in L_a \quad (4.10f)$$

where A is the set of all epidemic scenarios a , L is the set of all candidate locations i that can be selected, L_a is the set of all locations that are capable of detecting epidemic scenario a , α_a is the probability that scenario a occurs, d_{ai} is the impact coefficient which describes the measure of impact that will be incurred if scenario a is first detected at location i , x_{ai} is an indicator variable that denotes 1 if location i is selected and is also the first to detect epidemic scenario a , s_i is a binary variable indicating whether location i is selected, and K is a parameter for the limit on the number of locations that can be selected.

The use of scenarios to quantify uncertainty in the objective function yields a stochastic program that minimizes the total expected impact-weighted cost. Constraints 4.10b ensure that all scenarios are detected by the set of selected detection sites. Constraints 4.10c define a budget limit for the number of selected detection sites. Constraints 4.10d ensure that an epidemic scenario is detected only if a corresponding detection site is selected. Constraints 4.10e require that the placement variables s_i be binary and Constraints 4.10f enforce that the assignment variables x_{ai} be continuous between 0 and 1.

The epidemic scenarios for the impact formulation that comprise set A are generated using the stochastic SIR simulation defined in Equation 4.8. For each scenario, the simulation is first initialized with 10 infected measles cases in a single city. Then, the infection spread is simulated for 104 biweeks (4 years) across the set of cities while the number of cases in each city at each time period are recorded. For the subsequent analyses, a total of 60 scenarios are conducted, one for each city in the data set. The transmission radius $d_{j,k}$ was bounded to cities within 60 miles.

The impact coefficient d_{ai} can represent a number of different measures of impact. For our cases, we define the impact metric as the total number of infected cases across the metapopulation before an outbreak has been detected, denote as TNC (Total Number of Cases). This metric also requires specification of a threshold level for detection. For the subsequent analyses, we have chosen a threshold level of 12 infected cases. This parameter can have a large impact on the optimization results. An analysis is shown in the next section that describes the sensitivity of this threshold parameter. The impact coefficients d_{ai} are precomputed using the set of epidemic scenarios A .

The set L_a , defined as the set of all sites that can detect epidemic scenario a , is also precomputed using the set of epidemic scenarios A . A dummy site is included within each location set L_a to handle the case when no existing candidate site is capable of detecting epidemic scenario a . The impact coefficient d_{ai} corresponding to a dummy site that is selected is set to a large value in order to penalize undetected scenarios. For our results, we set the undetected impact coefficient to the maximum impact coefficient corresponding to the scenario in the set of scenarios A . In other words, the impact of not detecting scenario a is equivalent to the maximum number of cases that correspond to scenario a at a given threshold level.

4.6.2 Selection Results

The simulation was written in Python while the optimal detection site formulation was built using the Impact Formulation method contained in the Python package

Chama (Klise et al., 2017), and solved using Gurobi Optimizer using default settings. Chama is an open-source Python package that provides mixed-integer linear programming formulations to determine optimal sensor locations that maximize monitoring effectiveness. Chama is a general-purpose optimization software tool and is applicable to a wide range of sensor placement optimization problems, including designing sensor networks for airborne and water pollutant monitoring. Both the simulation, optimization, and subsequent analyses were performed on a 2.4 GHz Intel Core i5 machine with 8 GB RAM. Solution times for the optimal detection site formulation were consistently below 0.10 seconds with an optimality gap of 0.00%.

Figure 4.11 shows the optimal selection of disease detection sites for the TNC impact metric, along with a threshold detection level of 12 cases and a budget of 6 detection sites for the 60-city UK dataset. The locations and approximate population sizes (circa 1944-1964) of the selected cities are London (3.2 Million), Bristol (436,000), Southampton (191,000), South Shields (108,000), Exeter (76,000), and Macclesfield (36,000). Also of interest is that the optimization selected sites in centrally-located areas, rather than ones with the largest populations.

Note that the distance bound of 60 miles that was applied to the simulation can greatly affect the selection of sites because a tighter boundary can yield islanded clusters of cities where epidemics will be localized. Since the optimization formulation (specifically Constraint 4.10b) requires that each scenario be captured by the set of detection sites (otherwise a heavy penalty will be incurred), a detection site will nearly always be selected within islanded clusters of cities.

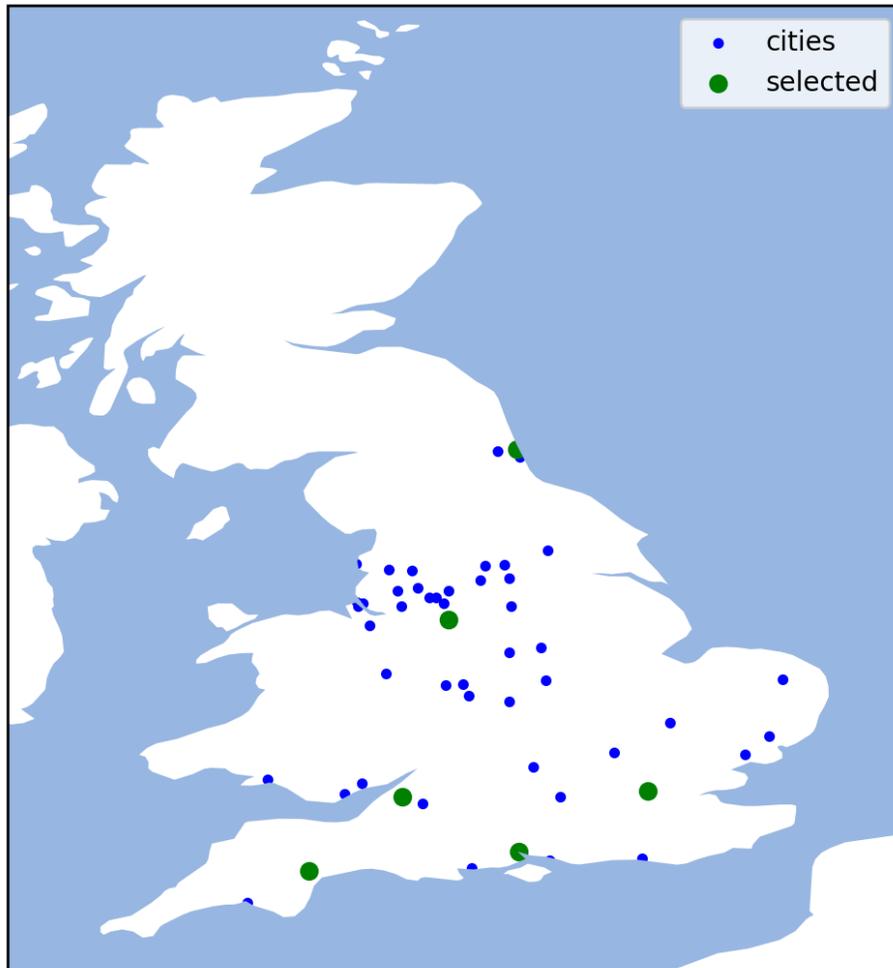


Figure 4.11.: Optimal Selection Scheme for TNC impact metric with detection threshold level of 12 cases and budget $K=6$

Figure 4.12 illustrates the sensitivity of the objective (with the TNC impact metric) with respect to the detection threshold level. The analysis was performed with detection threshold levels from 10 to 100 cases. The figure shows a gradual increase in the total number of expected cases with an increase in the detection threshold. When fewer resources are able to be allocated to improve detection thresholds, more outbreaks will go undetected, leading to an increase in the total number of people who will become exposed to and infected by the disease.

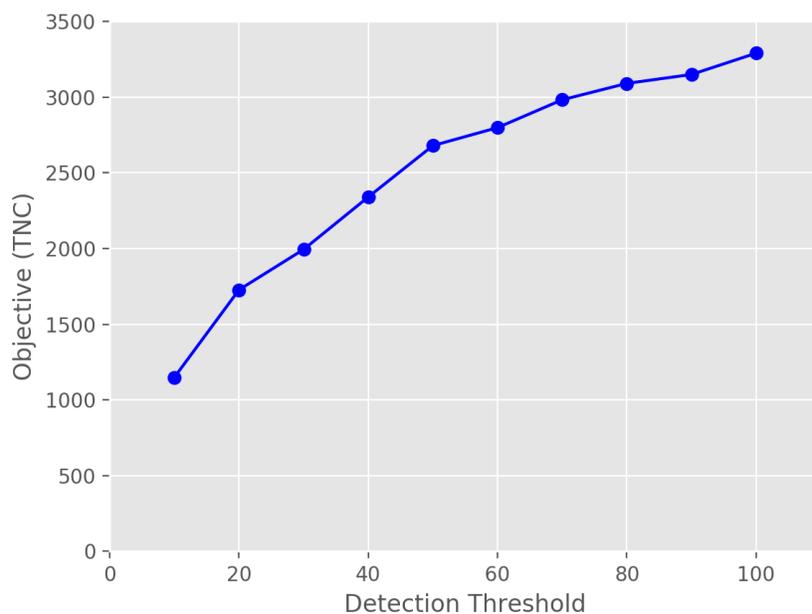


Figure 4.12.: Objective with TNC impact metric versus detection threshold level for levels from 10 to 100 cases

Figure 4.13 shows the impact of the budget K on the objective with TNC impact metrics at a threshold level of 12 cases. The objectives are plotted on a logarithmic scale to more clearly represent the stark decrease that occurs. The difference in the total expected number of cases between placing one detection site and 10 detection sites is nearly 2 orders of magnitude difference, from over 18,000 cases to just under 500 cases.

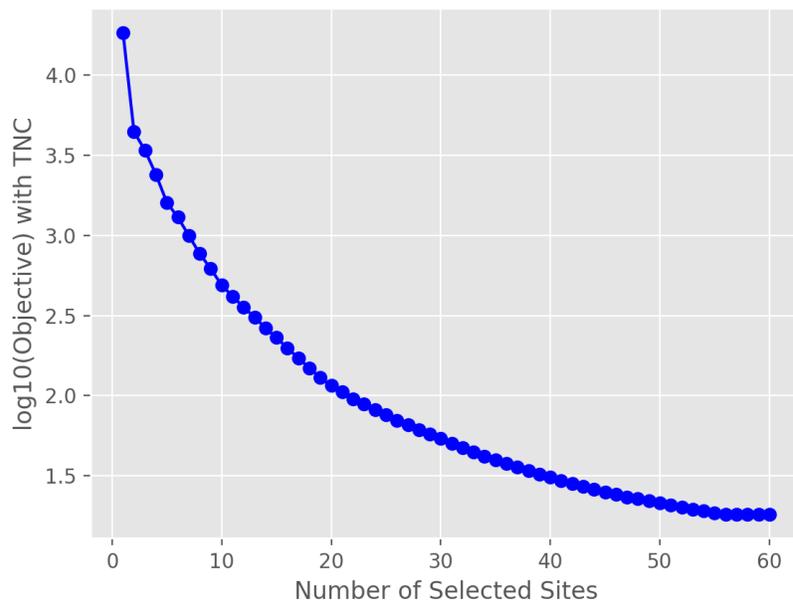


Figure 4.13.: $\log_{10}(\text{Objective})$ with TNC impact metric versus budget limit from $K=1$ to $K=60$

The trade-off effect shown in this figure means that a few optimally selected sites (in this case, just 10 detection sites out of 60 possible candidate sites) can have a dramatic impact on the possibility of early detection and mitigation of infectious disease epidemics. In a resource-constrained region with limited access to healthcare facilities and emergency personnel, the optimal selection of sites to locate these finite resources can be crucial for preventing larger outbreaks.

4.7 Conclusions

From this work, we have presented a nonlinear programming framework for estimation of city-to-city spatial and seasonal transmission rates by inferring transport information from localized disease case data, as well as a rigorous selection methodology for the location of disease detection sites using a scenario-based stochastic program. We have shown that nonlinear programming is an effective tool for esti-

mating spatio-temporal disease models from case data using a hazards-based model. We use this to estimate spatial coupling parameters for three relationships: (i) individual city-to-metapopulation, (ii) simultaneous city-to-metapopulation with power law model, and (iii) city-to-city transport with gravity model. We then extended the city-to-city gravity formulation to estimate seasonal disease transmission parameters in addition to transport parameters. With the estimated parameters, we were able to generate simulations of measles cycles on a metapopulation scale that reproduce fade-out characteristics seen in the data. We then constructed simulations as scenarios for the basis of a scenario-based stochastic programming formulation that identified optimal locations for disease detection sites. The methods presented in this work are readily scalable to larger problem sizes, have fast solution times, and are flexible to handle increasing parameter uncertainty. The optimal site selection method uses the set of simulations to handle uncertainty in the source and spread of epidemics on a countrywide scale, and leads to a solution that performs well for healthcare policymakers and emergency medical teams alike in selecting locations that minimize detection time and total number of infected cases.

5. SUMMARY

While the optimal sensor placement problem has been well studied given the extensive number of fields where it applies, it remains a difficult problem to solve and can require significant investment in model formulation depending on the application. Starting from the main contributions present in literature for facility location problems (e.g. Maximal Coverage Location Problem, Capacitated Facility Location Problem, and p -median/ p -center based problems), our work has been devoted to extending, testing, and improving approaches for the placement of sensors across three different application areas: optimal placement of data concentrators for the expansion of the Smart Grid communications network, optimal placement of flame detectors within petrochemical facilities, and optimal selection of infectious disease detection sites across a large population. In this chapter, we summarize the major contributions of each work and provide suggestions for future avenues of research.

5.1 Optimal Placement of Data Concentrators

In Chapter 2, we presented work on the optimal placement of data concentrators for the expansion of the Smart Grid communications network to serve the distribution power system based on extensions to the Capacitated Facility Location and p -center problems. A central issue in the expansion of the Smart Grid is the placement of additional data concentrators to service a growing network of smart meters while maintaining a high level of reliability and Quality-of-Service. While this specific problem has not been previously addressed, related works involving design of Automated Metering Infrastructure for the Smart Grid have focused on development of approximation algorithms and heuristics to solve cost minimization problems. In addition, few studies in this field have addressed the location problem and routing

design problems simultaneously within a single formulation, while minimizing chance of congestion and ensuring QoS guarantees.

The focus in this work is on optimal handling of aperiodic data traffic by optimizing the residual buffer capacity of data concentrators as an analog to optimizing QoS factors like bandwidth, throughput, packet loss, and delay requirements in the network. The objectives proposed by (Gourdin and Klopfenstein, 2006) are extended through a topology design framework, along with incorporation of communication redundancy to improve network reliability. In addition, our models prove effective for multi-period expansion of the communications network. To the best of our knowledge, no prior work has been done on solving the problem of data concentrator placement while prioritizing QoS performance within an Automated Metering Infrastructure network.

The placement of data concentrators to maximize network residual buffer capacity under the impacts of budget constraints, network connectivity requirements, distance limitations due to path loss, and redundant communication links to enhance robustness for an expanding Smart Grid communications network is a unique challenge that we address in this work. Three network design formulations based on mixed-integer linear programming approaches were proposed to handle Smart Grid expansion with QoS guarantees: (i) SGEP-A – maximizing the average residual buffer capacity, (ii) SGEP-MM – maximizing the minimum residual buffer capacity, and (iii) SGEP-R – minimizing the total reciprocal residual buffer capacity. Objective (i) follows the structure of a maximum CFLP-based formulation, while Objective (ii) is a robust p -center maximin formulation. Objective (iii) represents a hybrid of the previous two objectives in solution behavior, timing, and effectiveness.

We applied the design formulations to a case study with empirical parameters and solved them to optimality using Gurobi. The proposed formulations generated feasible and effective design topologies for data concentrators in our test network. The maximum CFLP-based SGEP-A can be useful for feasibility analysis, but is poor for routing assignment, since the total residual buffer capacity does not capture the

benefit of more evenly distributed capacity. Compared with SGEP-A, the p -center-based SGEP-MM and hybrid reciprocal-based SGEP-R models both provide better solutions in terms of routing assignment and lead to better placement within vicinity of the feeder network. Between the SGEP-MM and SGEP-R models, SGEP-MM provides minimum residual buffer capacity guarantees that are useful for establishing and ensuring QoS standards, but can require high solution times to achieve optimality. The SGEP-R formulation, while inherently nonlinear, can be relaxed to a mixed-integer linear program using linear underestimators and solved to optimality faster than that of SGEP-MM while returning similar placements. The hybrid reciprocal-based formulation of SGEP-R also has the advantage of ensuring equal distributions of connections per data concentrator so the network is balanced rather than over- or under-saturated as in the case of SGEP-A. Our proposed formulations build off of the CFLP and p -center formulations to address outage uncertainties through incorporation of redundancy, and are effective for rolling horizon network expansion to address future network demands. This application has demonstrated that facility location formulations can provide a reasonable basis for optimal placement, even when significant extensions are necessary.

5.2 Optimal Placement of Flame Detectors

Chapter 3 addresses the optimal flame detector placement problem in petrochemical facilities with an extension of the maximal expected facility location problem to consider non-uniform unavailabilities in detection failure. In the petrochemical industry, the chief concern is the implementation of safe and reliable systems to ensure the well-being of operation personnel. Flame and gas detection systems have historically been designed using rule-of-thumb, semi-quantitative methods.

In this research, we have proposed a rigorous, quantitative approach to the problem of optimal flame detector placement while considering uncertainties in detection, such as detector outages and partial obstructions, that can lead to failure to detect combustion events. Our approach, based on mixed-integer nonlinear programming,

generates a flame detection network design of optical flame detectors that maximizes expected geographic coverage. The mixed-integer nonlinear programming formulation we use to maximize expected coverage for the placement of flame detectors is similar to the Maximal Expected Coverage Problem extension proposed by Camm et al. (2002), which was applied to the selection of nature reserve sites to maximize the expected number of species. We translate it to consider instead placement of optical flame detectors while assuming independent probabilities of detection failure, and allow for individual detector specifications as a function of distance from detector to target. Rather than perform a linear approximation of the nonlinear probabilistic model as in (Camm et al., 2002), our approach performs a convex relaxation of the model using linear under-estimators via a Taylor-Series approximation. To the best of our knowledge, no prior work has been done on solving the problem of optimally locating optical flame detectors within process facilities by maximizing expected coverage in the face of detection failure.

Our P-FDP formulation (based on the Maximal Expected Coverage Problem) assumes that complications may occur in detection systems, such as partial obscurities, detector failure, wiring issues, etc. The proposed formulation includes a parameter to capture a probability of detection failure. This probability can be set for each detector-entity pair. The formulation becomes a mixed-integer nonlinear program that is solved to global optimality using a tailored iterative algorithm that relies on refining a relaxed master problem with linear under-estimators. We test our approach on three separate test cases: (i) Case A – a basic 2D warehouse geometry, (ii) Case B – a real-world oil wellhead facility, and (iii) Case C – a real-world offshore platform hydrocarbon separations module. We utilize the state-of-the-art fire and gas mapping software developed by Kenexis Consulting Corporation, Kenexis Effigy (Kenexis, 2019), to generate real-world facility geometries. In addition, we integrate our developed models into the Python package, Chama (Klise et al., 2017), for general usage. We compare our proposed P-FDP formulation to the corresponding FDP formulation which assumes perfect detection (i.e. the classical Maximal Coverage Location

Problem) and evaluate the performance of the developed solution algorithm against BARON across both uniform and non-uniform probabilities.

We showed that our proposed P-FDP formulation results in consistent improvements in expected coverage compared with the FDP formulation when subjected to detector failure conditions. The improvement of the P-FDP model over the FDP model reaches up to 20% for some cases when assuming high detection failure probabilities, while the improvement decreases to 1-2% when assuming 90% success rate in detection. The probability of detection failure that we have incorporated into the model considers partial obstruction and detector range in addition to failure or unavailability of the detector itself. The P-FDP model also addresses redundancy in coverage to reduce the chance of undetected events, leading to improved reliability in situations where certain detectors may be out-of-service. Our formulation is agnostic to the specific chemical process, facility layout, and underlying detection technologies, and requires only coverage information from the set of candidate detector locations to perform the optimization. This formulation is general and useful for other coverage problems with similar uncertainty characteristics. Furthermore, when considering non-uniform probabilities, our global solution algorithm solves efficiently with reasonable computational effort and is competitive with off-the-shelf global solvers like BARON. The iterative solution algorithm and model extensions proposed here can be readily applied to sensor placement problems in other research areas, particularly ones that aim to place sensors to maximize observability of regions while considering detector failure conditions.

5.3 Optimal Selection of Infectious Disease Detection Sites

Chapter 4 describes work in the field of epidemiology, particularly in the development of a parameter estimation framework for construction of a city-to-city spatial transmission model and in the development of an optimal placement framework for the selection of infectious disease detection sites on a nationwide scale based on the classical p -median problem.

We first present a nonlinear programming framework for estimation of city-to-city spatial and seasonal transmission rates by inferring transport information from localized disease case data. We have shown that nonlinear programming is an effective approach to estimate spatio-temporal disease dynamics with a statistical hazards-based model. We have considered three models of spatial coupling: (i) individual city-to-metapopulation, (ii) combined city-to-metapopulation with power law model, and (iii) city-to-city transport with gravity model. We then extended the city-to-city gravity model to also estimate seasonal disease transmission parameters simultaneously with the transport parameter. The methods presented in this work are readily scalable to larger problem sizes, have faster solution times than existing parameter estimation methods, and are flexible to handle increasing parameter uncertainty.

Using our estimated model, we then developed a stochastic programming approach for the optimal selection of infectious disease detection sites based off of the classical p -median approach for facility location. Our approach handles uncertainties in epidemic origins and infection transmission by minimizing the impact from a set of epidemic scenarios generated using the results from parameter estimation work. We consider the total expected number of infected individuals before detection as the impact metric. Our optimal location framework results in fast solution times and provides useful information for policymakers and health care professionals alike in designing nationwide disease mitigation programs and allocation of emergency health-care resources. This work illustrates the effectiveness of nonlinear programming for estimation of these challenging models with computational performance that greatly exceeds that of sampling-based methods.

5.4 Recommendations for Future Work

For the work on optimal placement of data concentrators in Chapter 2, future avenues of study include evaluation of the proposed approaches on applicability to different environment types and on scalability to various network sizes. Further investigation of the reciprocal-based placement method can lead to insights for general

facility location problems, as the method has shown distinct advantages when compared with the established CFLP and p -center-based formulations. Investigating the optimal spatial distribution of node placement is also important, since clustered placement of data concentrators can reduce their effective utilization despite provisioning for node outages. One possible alternative model that could remedy this would be a p -median based formulation, since the distance-centrality property of the p -median problem can be used to prioritize effective node utilization while limiting clusters of placed nodes. Multi-period placement approaches using stochastic growth models for the expansion of the Smart Grid network over time can also address concerns such as worst-case scenarios and placement under uncertainty.

Concerning optimal flame detector placement discussed in Chapter 3, future work would include investigating the behavior of the maximal expected facility location-based P-FDP formulation by modeling detection probabilities as a function of distance-to-detector according to the field-of-view specifications of the candidate detectors. In addition, the development of a decomposition method would be vital to handle large-scale facility geometries (or high resolution discretizations) within reasonable computational time. A combined optimization formulation for flame and gas detector placement with shared budget or cost constraints would also provide a holistic approach to process safety design.

For the future of parameter estimation of infectious disease spread discussed in Chapter 4, the hope is to further refine model formulations to better capture and portray spatial transmission of disease spread. This includes testing a wider array of formulations concerning the spatial coupling parameter (e.g. a radiation-based model as a complement to the gravity model). In addition, investigation of the extent of our methods by applying it to new types of datasets and infectious diseases would be valuable.

Future work on optimal selection of infectious disease detection sites includes investigation of dynamic placement models such as time-dependent deployment of medical teams and resources to serve communities in need that take into account local

epidemic trajectories as well as spatial transmission. In addition, to complement the stochastic programming frameworks proposed in this dissertation, a robust optimization framework to handle uncertainties in disease spread by providing guarantees for worse-case scenarios such as optimal quarantine zones would be practical for policymakers in controlling epidemics. To integrate with the parameter estimation work, development of an information-based optimization formulation that selects the locations that best provide information for estimation spatial transmission parameters would be valuable, such as those based on Fisher Information Matrices or corresponding methods utilized in the structural health monitoring fields that localize sensors so as to maximize estimation of structural failure modes.

REFERENCES

REFERENCES

- Abhayawardhana, V., I. Wassell, D. Crosby, M. Sellars, and M. Brown (2005). Comparison of Empirical Propagation Path Loss Models for Fixed Wireless Access Systems. *2005 IEEE 61st Vehicular Technology Conference 1(c)*, 73–77.
- Amaldi, E., A. Capone, M. Cesana, I. Filippini, and F. Malucelli (2008). Optimization models and methods for planning wireless mesh networks. *Computer Networks 52(11)*, 2159–2171.
- Anderson, R. and R. May (1992). *Infectious Diseases of Humans: Dynamics and Control*. Dynamics and Control. OUP Oxford.
- Andreadou, N., E. Kotsakis, and M. Masera (2018). Smart meter traffic in a real LV distribution network. *Energies 11(5)*.
- Association, N. F. P. (2019). NFPA 72: National Fire Alarm and Signaling Code 2013 Edition. Technical report.
- Bagula, A., L. Castelli, and M. Zennaro (2015). On the design of smart parking networks in the smart cities: An optimal sensor placement model. *Sensors 15(7)*, 15443–15467.
- Bellekens, B., L. Tian, P. Boer, M. Weyn, and J. Famaey (2018). Outdoor IEEE 802.11ah Range Characterization Using Validated Propagation Models. *2017 IEEE Global Communications Conference, GLOBECOM 2017 - Proceedings 2018-Janua*, 1–6.
- Benavides-Serrano, A. J. (2014). *Mathematical Programming Formulations for the Optimal Placement of Imperfect Detectors with Applications to Flammable Gas Detection and Mitigation Systems*. Ph. D. thesis, Texas A&M University.
- Benavides-Serrano, A. J., S. W. Legg, R. Vázquez-Román, M. S. Mannan, C. D. Laird, R. Vaquez-Roma, M. S. Mannan, C. D. Laird, and M. O. Kay (2014). A stochastic programming approach for the optimal placement of gas detectors: Unavailability and voting strategies. *Industrial and Engineering Chemistry Research 53(13)*, 5355–5365.
- Berger-Wolf, T. Y., W. E. Hart, and J. Saia (2005). Discrete sensor placement problems in distribution networks. *Mathematical and Computer Modelling 42(13)*, 1385–1396.
- Berry, J. W., L. Fleischer, W. E. Hart, C. A. Phillips, and J.-P. Watson (2005). Sensor Placement in Municipal Water Networks. *Journal of Water Resources Planning and Management 131(3)*, 237–243.

- Bertsekas, D. (1998). *Network Optimization: Continuous and Discrete Models*, Volume C.
- Bjørnstad, O. N., B. F. Finkenstädt, and B. T. Grenfell (2002). Dynamics of measles epidemics: Estimating scaling of transmission rates using a Time series SIR model. *Ecological Monographs* 72(2), 169–184.
- Bjørnstad, O. N. and B. T. Grenfell (2008). Hazards, spatial transmission and timing of outbreaks in epidemic metapopulations. *Environmental and Ecological Statistics* 15, 265–277.
- Bogue, R. (2013). Sensors for fire detection. *Sensor Review* 33(2), 99–103.
- Budka, K. C., J. G. Deshpande, and M. Thottan (2014). *Communication Networks for Smart Grids Making Smart Grid Real* (1 ed.). London: Springer.
- Calik, H., M. Labbé, and H. Yaman (2015). p-center problems. In *Location science*, pp. 79–92. Springer.
- Camm, J. D., S. K. Norman, S. Polasky, and A. R. Solow (2002). Nature Reserve Site Selection to Maximize Expected Species Covered. *Operations Research* 50(6), 946–955.
- Casillas, M., V. Puig, L. Garza-Castañón, and A. Rosich (2013). Optimal Sensor Placement for Leak Location in Water Distribution Networks Using Genetic Algorithms. *Sensors* 13(11), 14984–15005.
- Cauchemez, S., A. Bhattarai, T. L. Marchbanks, R. P. Fagan, S. Ostroff, N. M. Ferguson, and D. Swerdlow (2011, February). Role of social networks in shaping disease transmission during a community outbreak of 2009 H1N1 pandemic influenza. *Proceedings of the National Academy of Sciences of the United States of America* 108(7), 2825–30.
- Church, R. and C. Reville (1972). The Maximal Covering Location Problem. *Science* 32(6), 101–118.
- Daskin, M. S. (1982). Application of an expected covering model to emergency medical service system design. *Decision Sciences* 13(3), 416–439.
- Daskin, M. S. (1983). A Maximum Expected Covering Location Model: Formulation, Properties and Heuristic Solution. Technical Report 1.
- Daskin, M. S., L. V. Snyder, and R. T. Berger (2005). Facility location in supply chain design. In *Logistics systems: Design and optimization*, pp. 39–65. Springer.
- Erdem, U. M. and S. Sclaroff (2004). Optimal placement of cameras in floorplans to satisfy task requirements and cost constraints. *Proceedings of the International Workshop on Omnidirectional Vision, Camera Networks and Non-classical Cameras (OMNIVIS)* (May), 30–41.
- Farahani, R. Z., N. Asgari, N. Heidari, M. Hosseininia, and M. Goh (2012). Covering problems in facility location: A review. *Computers and Industrial Engineering* 62(1), 368–407.
- Fernández, E. and M. Landete (2015a). Fixed-charge facility location problems. In *Location science*, pp. 47–77. Springer.

- Fernández, E. and M. Landete (2015b). *Fixed-Charge Facility Location Problems*, pp. 47–77. Cham: Springer International Publishing.
- Finkenstädt, B. F., O. N. Bjørnstad, and B. T. Grenfell (2002). A stochastic model for extinction and recurrence of epidemics: estimation and inference for measles outbreaks. *Biostatistics (Oxford, England)* 3, 493–510.
- Finkenstädt, B. F. and B. T. Grenfell (2000). Time series modelling of childhood diseases: a dynamical systems approach. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 49(2), 187–205.
- Fonollosa, J., A. Solórzano, and S. Marco (2018). Chemical sensor systems and associated algorithms for fire detection: A review. *Sensors (Switzerland)* 18(2).
- Glass, K., Y. Xia, and B. T. Grenfell (2003). Interpreting time-series analyses for continuous-time biological models - Measles as a case study. *Journal of Theoretical Biology* 223(1), 19–25.
- Gomes, E. G., R. D. A. Medronho, J. Victor, and B. Alves (2014). Gas Detector Placement in Petroleum Process Unit Using Computational Fluid Dynamics. *International Journal of Modeling and Simulation for the Petroleum Industry* 8(1), 17–24.
- Gourdin, E. and O. Klopfenstein (2006). Comparison of different QoS-oriented objectives for multicommodity flow routing optimization. *Proceedings of the International Conference on Telecommunications (ICT 2006)* (1), 1–4.
- Grenfell, B. T., O. N. Bjørnstad, and B. F. Finkenstädt (2002). Dynamics of measles epidemics: Scaling noise, determinism, and predictability with the TSIR model. *Ecological Monographs* 72(2), 185–202.
- GSMA (2015). Smart meters: Compliance with radio frequency exposure standards. Technical report, London.
- Gurobi Optimization, L. (2018). Gurobi optimizer reference manual. <http://www.gurobi.com>. Accessed: 2018-11-30.
- Hakimi, S. L. (1964). Optimum locations of switching centers and the absolute centers and medians of a graph. *Operations research* 12(3), 450–459.
- Hakimi, S. L. (1965). Optimum distribution of switching centers in a communication network and some related graph theoretic problems. *Operations research* 13(3), 462–475.
- Hamid, M. (2013). Path Loss Models for LTE and LTE-A Relay Stations. *Universal Journal of Communications and Network* 1(4), 119–126.
- Hanoun, S., A. Bhatti, D. Creighton, S. Nahavandi, P. Crothers, and C. G. Esparza (2016). Target coverage in camera networks for manufacturing workplaces. *Journal of Intelligent Manufacturing* 27(6), 1221–1235.
- Hari, K. V. S. and M. S. Smith (2003). *Channel Models for Fixed Wireless Applications Background*.
- Hart, W. E., C. Laird, J.-P. Watson, and D. L. Woodruff (2012). *Pyomo—optimization modeling in python*, Volume 67. Springer Science & Business Media.

Hart, W. E., C. D. Laird, J.-P. Watson, D. L. Woodruff, G. A. Hackebeil, B. L. Nicholson, and J. D. Siirola (2017). *Pyomo—optimization modeling in python* (Second ed.), Volume 67. Springer Science & Business Media.

Hart, W. E. and R. Murray (2010). Review of Sensor Placement Strategies for Contamination Warning Systems in Drinking Water Distribution Systems. *Journal of Water Resources Planning and Management* 136(6), 611–619.

Hart, W. E., J.-P. Watson, and D. L. Woodruff (2011a). Pyomo: modeling and solving mathematical programs in python. *Mathematical Programming Computation* 3(3), 219–260.

Hart, W. E., J.-P. Watson, and D. L. Woodruff (2011b). Pyomo: modeling and solving mathematical programs in python. *Mathematical Programming Computation* 3(3), 219–260.

Hassan, A., Y. Zhao, L. Pu, G. Wang, H. Sun, and R. M. Winter (2018). Evaluation of clustering algorithms for DAP placement in wireless smart meter network. *Proceedings of 2017 9th International Conference On Modelling, Identification and Control, ICMIC 2017 2018-March*(Icmic), 1085–1090.

Heesterbeek, H., R. M. Anderson, V. Andreasen, S. Bansal, D. De Angelis, C. Dye, K. T. D. Eames, W. J. Edmunds, S. D. W. Frost, S. Funk, T. D. Hollingsworth, T. House, V. Isham, P. Klepac, J. Lessler, J. O. Lloyd-Smith, C. J. E. Metcalf, D. Mollison, L. Pellis, J. R. C. Pulliam, M. G. Roberts, and C. Viboud (2015, March). Modeling infectious disease dynamics in the complex landscape of global health. *Science* 347(6227), aaa4339–aaa4339.

Hethcote, H. W. (2000). The mathematics of infectious diseases. *SIAM review* 42(4), 599–653.

Huang, X. and S. Wang (2015). Aggregation Points Planning in Smart Grid Communication System. *IEEE Communications Letters* 19(8), 1315–1318.

ISA-TR84.00.07 (2010). *Guidance on the Evaluation of Fire, Combustible Gas and Toxic Gas System Effectiveness*. Number January 2010.

Jahromi, A. E. and Z. B. Rad (2013). Optimal topological design of power communication networks using genetic algorithm. *Scientia Iranica* 20(3), 945–956.

Jandarov, R., M. Haran, O. N. Bjornstad, and B. Grenfell (2014). Emulating a gravity model to infer the spatiotemporal dynamics of an infectious disease arXiv : 1110.6451v1 [stat.ME] 28 Oct 2011. *Journal of the Royal Statistical Society* 63(3), 423–444.

Karatas, M., N. Razi, and H. Tozan (2016). A comparison of p-median and maximal coverage location models with Q-coverage requirement. *Procedia Engineering* 149(June), 169–176.

Keeling, M. J. and P. Rohani (2002). Estimating spatial coupling in epidemiological systems: a mechanistic approach. *Ecology Letters* 5(1), 20–29.

Kenexis, C. C. (2019). Effigy. <https://www.kenexis.com/software/effigy/>. Accessed: 2019-01-30.

- Klinkert, A. (2018). *Designing Smart-grid Telecommunications Systems via Interval Flow Network Optimization*. Ph. D. thesis, Southern Methodist University.
- Klise, K. A., B. Nicholson, and C. D. Laird (2017). Sensor placement optimization using chama. Technical report, Sandia National Lab.(SNL-NM), Albuquerque, NM (United States).
- Ko, B. C., K. H. Cheong, and J. Y. Nam (2009). Fire detection based on vision sensor and support vector machines. *Fire Safety Journal* 44(3), 322–329.
- Kojima, H., T. Tsuchiya, and Y. Fujisaki (2018). The aggregation point placement problem for power distribution systems. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences E101A*(7), 1074–1082.
- Kong, P. Y. (2019). Cost Efficient Data Aggregation Point Placement with Interdependent Communication and Power Networks in Smart Grid. *IEEE Transactions on Smart Grid* 10(1), 74–83.
- Krause, A. and C. Guestrin (2007). Near-optimal Observation Selection using Submodular Functions. *Association for the Advancement of Artificial Intelligence*, 1650–1654.
- Krause, A. and C. Guestrin (2009). Robust sensor placement for detecting adversarial contaminations in water distribution systems. In *World Environmental and Water Resources Congress 2009: Great Rivers*, pp. 1–10.
- Krause, A., J. Leskovec, C. Guestrin, J. M. VanBriesen, and C. Faloutsos (2008). Efficient sensor placement optimization for securing large water distribution networks. *Journal of Water ...* 134(6), 516–526.
- Kuzlu, M., M. Pipattanasomporn, and S. Rahman (2014). Communication Network Requirements for Major Smart Grid Applications in HAN, NAN and WAN. *Computer Networks* 67, 74–88.
- Laporte, G., S. Nickel, and F. S. da Gama (2015). *Location Science*, Volume 528. Springer.
- Legg, S. W., A. J. Benavides-Serrano, J. D. Siirola, J. P. Watson, S. G. Davis, A. Bratteteig, and C. D. Laird (2012). A stochastic programming approach for gas detector placement using CFD-based dispersion simulations. *Computers and Chemical Engineering* 47, 194–201.
- Legg, S. W., C. Wang, A. J. Benavides-Serrano, and C. D. Laird (2013). Optimal gas detector placement under uncertainty considering Conditional-Value-at-Risk. *Journal of Loss Prevention in the Process Industries* 26(3), 410–417.
- Lopez-Aguilera, E., J. Casademont, and J. Cotrina (2010). Propagation delay influence in IEEE 802.11 outdoor networks. *Wireless Networks* 16(4), 1123–1142.
- Milke, J. A., M. E. Hulcher, C. L. Worrell, D. T. Gottuk, and F. W. Williams (2003). Investigation of multi-sensor algorithms for fire detection. *Fire Technology* 39(4), 363–382.
- Mo, Y., R. Ambrosino, and B. Sinopoli (2011). Sensor selection strategies for state estimation in energy constrained wireless sensor networks. *Automatica* 47(7), 1330–1338.

- Murray, A. T., K. Kim, J. W. Davis, R. Machiraju, and R. Parent (2007). Coverage optimization to support security monitoring. *Computers, Environment and Urban Systems* 31(2), 133–147.
- Nagai, Y., L. Zhang, T. Okamawari, and T. Fujii (2013). Delay performance analysis of LTE in various traffic patterns and radio propagation environments. *IEEE Vehicular Technology Conference*, 1–5.
- Niyato, D. and P. Wang (2012). Cooperative transmission for meter data collection in smart grid. *IEEE Communications Magazine* 50(4), 90–97.
- Nunes, R. W. (1999, Feb). Safety alarms and European standards. (CERN-ST-99-035), 6 p.
- Obi, E. C. (2014). *Optimization of Flame and Gas Detectors*. Ph. D. thesis, University of Stavanger, Norway.
- Olague, G. and R. Mohr (1998). Optimal 3d sensors placement to obtain accurate 3d points positions. *Proceedings of the Fourteenth International Conference on Pattern Recognition* 1, 16–20.
- Parikh, P. P., M. G. Kanabar, and T. S. Sidhu (2010). Opportunities and challenges of wireless communication technologies for smart grid applications. *IEEE PES General Meeting (Cc)*, 1–7.
- Park, S., J. Lee, J. Kim, J. Lee, and K. Lee (2019). ExLL : An Extremely Low-latency Congestion Control for Mobile Cellular Networks. *CoNEXT*, 307–319.
- Patel, A., J. Aparicio, N. Tas, M. Loiacono, and J. Rosca (2011, oct). Assessing communications technology options for smart grid applications. In *2011 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pp. 126–131. IEEE.
- Qiu, L., R. Chandra, K. Jain, and M. Mahdian (2004). Optimizing the placement of integration points in multi-hop wireless networks. *Proceedings of ICNP*.
- Rad, A., D. Rashtchian, and N. Badri (2017). A risk-based methodology for optimum placement of flammable gas detectors within open process plants. *Process Safety and Environmental Protection* 105, 175–183.
- Rastgoo, R. and V. Sattari-Naeini (2016). Tuning Parameters of the QoS-Aware Routing Protocol for Smart Grids Using Genetic Algorithm. *Applied Artificial Intelligence* 30(1), 52–76.
- Rémy, J. and C. Letamendia (2014). *LTE Standards*. ISTE. Wiley.
- ReVelle, C. S., H. A. Eiselt, and M. S. Daskin (2008). A bibliography for some fundamental problem categories in discrete location science. *European Journal of Operational Research* 184(3), 817–848.
- Rowaihy, H., S. Eswaran, M. Johnson, D. Verma, A. Bar-Noy, T. Brown, and T. La Porta (2007). A Survey of Sensor Selection Schemes in Wireless Sensor Networks. In *Proceedings of SPIE*, Volume 6562, pp. 11 pages.

- Rupasinghe, N. and I. Guvenc (2014). Licensed-assisted access for WiFi-LTE coexistence in the unlicensed spectrum. *2014 IEEE Globecom Workshops, GC Wkshps 2014* (March), 894–899.
- Sahinidis, N. V. (2017). *BARON 17.8.9: Global Optimization of Mixed-Integer Nonlinear Programs*, User’s Manual.
- Saputro, N., K. Akkaya, and S. Uludag (2012). A survey of routing protocols for smart grid communications. *Computer Networks* 56, 2742–2771.
- Schneider, K. P. and P. P. Yousu Chen David Chassin Robert Pratt Dave Engel Sandra Thompson ii (2008). Modern Grid Initiative Distribution Taxonomy Final Report. Technical report, Pacific Northwest National Laboratory.
- Snyder, L. V. (2006). Facility location under uncertainty: A review. *IIE Transactions (Institute of Industrial Engineers)* 38(7), 547–564.
- Sterle, C., A. Sforza, A. E. Amideo, and C. Piccolo (2016). A unified solving approach for two and three dimensional coverage problems in sensor networks. *Optimization Letters* 10(5), 1101–1123.
- Tavasoli, M., M. H. Yaghmaee, and A. H. Mohajezadeh (2016). Optimal placement of data aggregators in smart grid on hybrid wireless and wired communication. *2016 4th IEEE International Conference on Smart Energy Grid Engineering, SEGE 2016*, 332–336.
- Tawarmalani, M. and N. V. Sahinidis (2005). A polyhedral branch-and-cut approach to global optimization. *Mathematical Programming* 103, 225–249.
- Thuillard, M. (2002). A new flame detector using the latest research on flames and fuzzy-wavelet algorithms. *Fire Safety Journal* 37(4), 371–380.
- Toregas, C., R. Swain, C. ReVelle, and L. Bergman (1971). The location of emergency service facilities. *Operations research* 19(6), 1363–1373.
- Wachter, A. and L. T. Biegler (2006). On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical Programming* 106(1), 25–57.
- Wang, B. (2011). Coverage problems in sensor networks: A survey. *ACM Computing Surveys (CSUR)* 43(4), 32.
- Wang, H., S. Zheng, C. Chen, W. Yang, L. Wu, X. Cheng, M. Fei, and C. Hu (2006). Discriminating fire detection via support vector machines. *Lecture Notes in Control and Information Sciences* 344, 176–181.
- Watson, J.-P., W. E. Hart, and R. Murray (2008). Formulation and optimization of robust sensor placement problems for contaminant warning systems. In *Water Distribution Systems Analysis Symposium 2006*, pp. 1–13.
- Watson, J.-P., R. Murray, and W. E. Hart (2009). Formulation and optimization of robust sensor placement problems for drinking water contamination warning systems. *Journal of Infrastructure Systems* 15(4), 330–339.

- Word, D. P., D. a. T. Cummings, D. S. Burke, S. Iamsirithaworn, and C. D. Laird (2012). A nonlinear programming approach for estimation of transmission parameters in childhood infectious disease using a continuous time model. *Journal of the Royal Society, Interface / the Royal Society* 9, 1983–97.
- Wu, H., D. Wu, and J. Zhao (2019). An intelligent fire detection approach through cameras based on computer vision methods. *Process Safety and Environmental Protection* 127, 245–256.
- Xia, Y., O. N. Bjørnstad, and B. T. Grenfell (2004). Measles metapopulation dynamics: a gravity model for epidemiological coupling and dynamics. *The American naturalist* 164(2), 267–281.
- Xu, S., Y. Qian, and R. Qingyang Hu (2018). Reliable and resilient access network design for advanced metering infrastructures in smart grid. *IET Smart Grid* 1(1), 24–30.
- Yang, T. P., V. S. Asirvadam, and N. B. Saad (2012, jun). Optimal placement of fire detector using dual-view. In *2012 4th International Conference on Intelligent and Advanced Systems (ICIAS2012)*, pp. 599–603. IEEE.
- Youssef, W. and M. Younis (2010). Optimized asset planning for minimizing latency in wireless sensor networks. *Wireless Networks* 16(1), 65–78.
- Yuan, L., R. A. Thomas, J. H. Rowland, and L. Zhou (2018). Early fire detection for underground diesel fuel storage areas. *Process Safety and Environmental Protection* 119, 69–74.
- Zhao, J. and S. C. S. Cheung (2009). Optimal visual sensor planning. *Proceedings - IEEE International Symposium on Circuits and Systems* (June), 165–168.

APPENDICES

A. DERIVATION OF THE STOCHASTIC SIMULATION MODEL

A.1 Nomenclature

Table A.1.1.: Description of model variables and parameters

Variable	Description	Units
t	Time index	biweek
T	Set of all time intervals	–
j	City index	–
J	Set of all cities	–
$S_{t,j}$	Number of susceptible individuals	people
$B_{t,j}$	Number of births	people
$N_{t,j}$	Population	people
$I_{t,j}$	Number of infected individuals	people
$l_{t,j}$	Number of immigrant infections	people
$\beta_{u,j}$	Seasonal Transmission parameter	1/biweek
u	Seasonal index	–
α	Nonlinear correction parameter	–
c_j	Spatial coupling parameter	people
$x_{t,j}$	Fraction of local susceptible individuals	–
$y_{t,j}$	Fraction of non-local infected individuals	–
\hbar	Hazard probability	–

A.2 Derivation for Binomial Distribution

Assume that the contact between local susceptible individuals and non-local infectious individuals follows a Poisson process, with a rate of contact

$$k = Crx\bar{y}$$

where C is a constant coefficient; r is the rate of people moving in to the local region; \bar{y} is the probability that a non-local individual is infectious; and x is the probability that a local individual is susceptible (Note: the spatial coupling parameter c_j is the product of C , r , and the time T). Also assume that infected cases occur independently and are identically distributed.

Since the probability of infected cases can be modeled by a Poisson process, the interval of time between cases can be described by an Exponential distribution. The probability that no contact occurs between non-local infected individuals and local susceptible individuals is equivalent to the probability of the interval of time between cases being greater than an interval x (where x is the fade-out period and l_t are the number of contacts). Thus,

$$P(X > x) = P(l_t = 0) = \frac{e^{-kT}(kT)^0}{0!} = e^{-kT}$$

and the probability that there is at least one contact (or the probability that the fade-out period is less than interval x) is

$$P(X \leq x) = P(l_t > 0) = 1 - e^{-kT}$$

For relatively weak city-to-metapopulation coupling (Bjørnstad and Grenfell, 2008), we can assume the number of contacts l_t is binary and draws from the cumulative distribution function

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(u)du$$

Therefore, the number of contacts l_t can be represented by a random sample over the Binomial distribution

$$\text{Bin}(n, p) = \text{Bin}(1, 1 - e^{-c_j x \bar{y}})$$

where n is the number of trials and p is the probability of having at least one spatial contact.

A.3 Derivation for Negative Binomial Distribution

Assume the per capita growth rate of infected individuals can be represented by the geometric distribution. The geometric distribution describes the probability of “ x trials made before a success”. In this case, assume the number of incidences within time period t , I_{t+1} , represents x trials and an infected individual represents a success. The expectation of I_{t+1} for a geometric distribution is

$$E(I_{t+1}) = \frac{1-p}{p}$$

where p is the probability of an individual becoming infected.

The negative binomial distribution describes the probability of “ x trials made before s successes”, where s is fixed. Comparing these two distributions, we can see that the negative binomial distribution is a particular case of the geometric distribution, namely, when $s=1$. Therefore, for a number of initial infections $s=I_t + l_t$, the expectation of I_{t+1} is

$$E(I_{t+1}) = \frac{s(1-p)}{p} = \frac{(I_t + l_t)(1-p)}{p}$$

where p is the probability of an individual becoming infected.

Recall that the estimated number of infections can be modeled by the standard incidence rate. In our case,

$$E(I_{t+1}) = \lambda_t = \frac{\beta_u}{N}(I_t + l_t)^\alpha S_t$$

Using the previous relationship, we have

$$\frac{(I_t + l_t)(1-p)}{p} = \frac{\beta_u}{N}(I_t + l_t)^\alpha S_t$$

and

$$p = \frac{1}{1 + \frac{\beta_u S_t}{N}(I_t + l_t)^{\alpha-1}}$$

Therefore, the stochastic growth of an epidemic can be represented by a negative binomial process, $\text{NegBin}(n, p)$, with $n = I_{t,j} + l_{t,j}$ initial infections ($n > 0$) and

$$p = \frac{1}{1 + \frac{\beta_{u,j} S_{t,j}}{N_j} (I_{t,j} + l_{t,j})^{\alpha-1}}$$

probability of an individual becoming infected.

Note that this formulation is analogous to that used by Bjornstad et al. (Bjornstad and Grenfell, 2008), wherein the negative binomial process $\text{NegBin}(a, b)$ has parameters: expectation $a = \lambda_t$, and clumping parameter $b = I_t + l_t$.

A.4 Derivation for Hazard model

The probability that spatial contact will occur is $(1 - \exp(-c_j x_{t,j} \bar{y}_t))$, which can be derived from the exponential distribution (see A.2). Note that this probability expression implies that the larger the spatial coupling parameter of the city, the more likely spatial contact will occur between a local susceptible individual and a non-local infectious individual.

The probability that no new infection will occur if one non-local infection l_t is introduced at the beginning of time t is

$$P(I_{t+1} = 0 | I_t + l_t = 1) = p = \frac{1}{1 + \frac{\beta_u S_t}{N} (I_t + l_t)^{\alpha-1}}$$

The probability that a local epidemic occurs from a non-local infection is

$$P(I_{n+1} > 0 | I_t + l_t = 1) = 1 - p = \frac{\frac{\beta_u}{N} S_t}{1 + \frac{\beta_u}{N} S_t}$$

Therefore, the probability of a hazard, or the joint probability of spatial contact between a susceptible individual and a non-local infectious individual, along with the probability that an epidemic will result from such a contact is

$$\bar{h} = \frac{\frac{\beta_{u,j}}{N_j} S_{t,j} (1 - \exp(-c_j x_{t,j} \bar{y}_t))}{1 + \frac{\beta_{u,j}}{N_j} S_{t,j}}$$

A.5 Seasonal Transmission Parameters

The seasonal transmission rates for bi-weeks 1 through 26 were defined as $\beta_{u,j} = 30 \times (1.24, 1.14, 1.16, 1.31, 1.24, 1.12, 1.06, 1.02, 0.94, 0.98, 1.06, 1.08, 0.96, 0.92, 0.92, 0.86, 0.76, 0.63, 0.62, 0.83, 1.13, 1.20, 1.11, 1.02, 1.04, 1.08)$.

Figure A.5.1 below shows the parameters plotted over a one-year horizon. These transmission rates were the result of estimations performed by Bjørnstad et al. and Grenfell et al. (Bjørnstad and Grenfell, 2008; Xia et al., 2004).

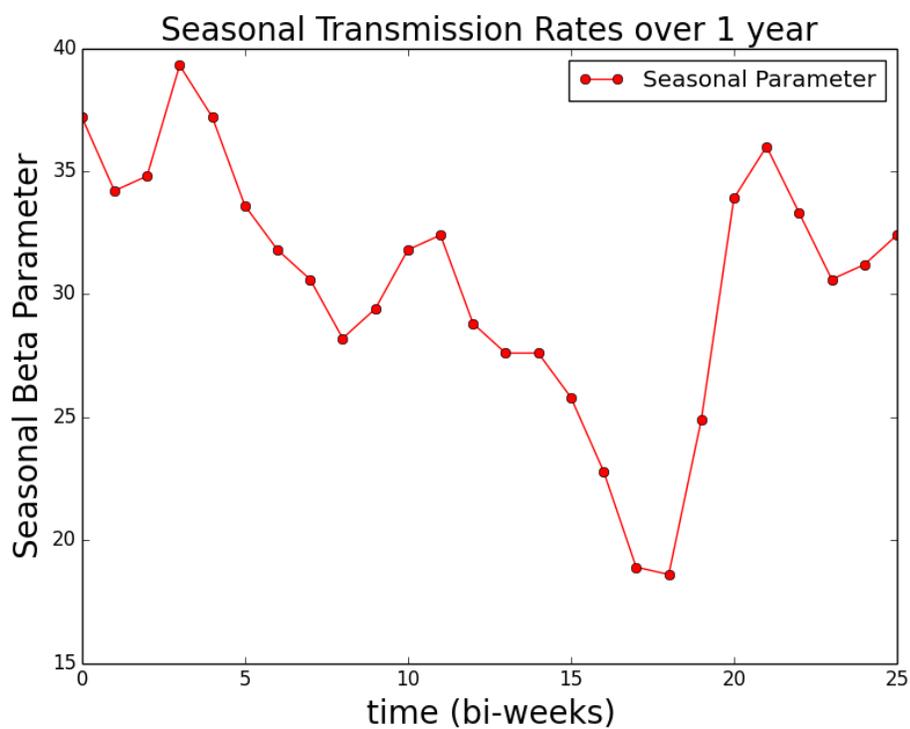


Figure A.5.1.: Seasonal Transmission Parameters over 1 year

VITA

VITA

Todd Zhen was born in Manhattan, NY, USA and grew up in Atlanta, GA, USA. He received his Bachelor of Science in Chemical and Biomolecular Engineering from Georgia Institute of Technology, Atlanta, GA. In August 2014, he started graduate school at Purdue University, West Lafayette, IN. In January 2015, he joined the Carl Laird Research Group, where his research focused on the development and application of mathematical optimization techniques for optimal sensor placement problems.

During his graduate program, Todd also interned at the National Renewable Energy Laboratories and Evonik Industries, where he gained valuable skills in the research and development of solutions for engineering problems.