Mitigating Hard Capacity Constraints in Facility Location Modeling

by

Kayse Maass

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Industrial and Operations Engineering)
in The University of Michigan
2017

Doctoral Committee:

Professor Mark S. Daskin, Chair
Professor Hyun-Soo Ahn
Professor Xiuli Chao
Assistant Professor Siqian May Shen

Kayse Maass

leekayse@umich.edu

ORCID ID: 0000-0002-4961-4156

ACKNOWLEDGEMENTS

I am immensely grateful for the support and guidance I received from many people while undertaking this Ph.D. First and foremost, I would like to thank my advisor, Dr. Mark S. Daskin, for his continuous support and encouragement during my time at the University of Michigan. Not only did he teach me how to become a better researcher, teacher, and mentor, Mark also taught me how to become a better individual. I am especially grateful for the ways in which he challenged me to see things from a variety of perspectives, provided space for me to voice my ideas and concerns in a way that built my confidence, and shared his infectious love of puns. It is my honor to have had the opportunity to work with and learn from him.

I would also like to thank the other members of my dissertation committee: Dr. Hyun-Soo Ahn, Dr. Xiuli Chao, and Dr. Siqian Shen, for their insightful comments and suggestions of my work over the past few years. To the many other collaborators, faculty, staff, and graduate students at the University of Michigan who have enriched my life with your presence, thank you for your encouragement, laughter, and friendship.

I gratefully acknowledge the financial support from the National Science Foundation Graduate Research Fellowship Program, Rackham Merit Fellowship Program, and Department of Industrial and Operations Engineering at the University of Michigan during my tenure as a doctoral student. This support enabled me to conduct the independent research within this dissertation and disseminate my findings worldwide. I am also immensely appreciative of the time and effort Eve Gendron dedicated to help run some of the computational studies presented in this dissertation.

In addition to the mentors I have gained during my doctoral studies, many thanks are due to my predoctoral mentors who have constantly served as a source of encouragement: my second grade teacher, Mrs. Carol Wright (for finding creative ways to challenge me intellectually); my high school calculus teacher and Math League coach, Mr. Reid Froiland (for affirming my high school self in pursuing a career in STEM); my Bethel University community of support including, Prof. Patrice Conrath (for her constant eternal perspective and for showing me how operations research can help improve our world), Dr. Eric Gossett (for honing my attention to detail and for creating an environment that fostered my self-confidence), Dr. William Kinney (for encouraging me to pursue graduate school), and Drs. Paula and Steve Soneral (for their continued insight into how to pursue your dreams while maintaining a work-life balance).

I would not be who I am today without the love and support of my family. My deepest gratitude goes to my parents, Mark and Sue Lee, for always believing in me and encouraging me to be myself. I am certain that the way they integrated math and problem-solving into our daily life cultivated my love for analytic and inquisitive thought from a young age. I am also grateful for the constant encouragement from Randy and Kathy Maass, and especially appreciate of Kathy for introducing me to the field in which I am receiving my doctoral degree. Most of all, I would like to thank my loving, encouraging, and patient husband, Zac, who faithfully supports my dreams in more ways than I could have ever imagined.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

CFLP	Capacitated Fixed Charge Location Problem
i.i.d.	Independent and identically distributed
IMCLP	Inventory Modulated Capacitated Location Problem
UFLP	Uncapacitated Fixed Charge Location Problem

LIST OF SYMBOLS

Sets

C Set identifying the number of consecutive days to couple in the relaxed joint chance constraint (5.74)

D Set of days; indexed by d

$D_p \coloneqq \{d \in D: d \bmod |P| = p, d > t^*\}$ Set of days in D corresponding to cycle day $p \in P$ after the warm-up period

$\mathcal{F} \coloneqq \{(\boldsymbol{X}, \boldsymbol{Y}): \exists \boldsymbol{V}, \boldsymbol{W}, \boldsymbol{Z}$: (2.2) - (2.4), (3.6), (5.10) - (5.12), (5.15) - (5.19), and (5.22) are satisfied$\}$

$\bar{\mathcal{F}} \coloneqq \{(\boldsymbol{X}, \boldsymbol{Y}): \exists \boldsymbol{V}, \boldsymbol{W}, \boldsymbol{Z}$: (2.2) - (2.4), (3.6), (5.15) - (5.19), (5.22), and (5.80) - (5.82) are satisfied$\}$

$\hat{\mathcal{F}} \coloneqq \{(\boldsymbol{X}, \boldsymbol{Y}): \exists \boldsymbol{V}, \boldsymbol{W}, \boldsymbol{Z}$: (2.2) - (2.4), (3.6), (5.15) - (5.19), (5.22), and (5.83) - (5.85) are satisfied$\}$

I Set of demand sites; indexed by i

I_j^{ψ} For a given, facility $j \in J$ and $\psi \in \underline{\Psi}$, the set of demand sites allocated to j

$\bar{I}_j^{\psi}$ For a given, facility $j \in J$ and $\psi \in \underline{\Psi}_j$, the set of demand sites allocated to j

J Set of candidate processing facility locations; indexed by j

J^{ψ} For a given $\psi \in \underline{\Psi}$, the set of located candidate facilities that violate the backlog threshold θ

P Set of cycle-days (e.g., days of the week); indexed by p

$\wp^{\omega' j}$ Feasible region of D-M1S2(ω', j)

Q_{ij} Set of cycle days for which demand shipped from $i \in I$ would arrive at facility $j \in J$ on the weekend

R_{ij} Set of cycle days for which demand shipped from $i \in I$ would arrive at facility $j \in J$ during the weekday

$\mathcal{S}^{\psi} \coloneqq \{(i,j) \in I \times J : Y_{ij}^{\psi} = 1 \text{ and } \exists \omega' \in \Omega'; d \in \{t^* + 1, \dots, |D| + 1\} \text{ with } V_{jd}^{\omega' \psi} > \theta\}$

$\Phi^{\omega' j}$ Set of extreme points of $\wp^{\omega' j}$; indexed by ϕ

$\Phi^{\omega'}$ Set of extreme points of (5.47) - (5.52) for scenario $\omega' \in \Omega'$; indexed by ϕ

$\Phi = \cup_{\omega' \in \Omega'} \Phi^{\omega'}$

$\underline{\Phi} \subseteq \Phi$ Set of optimality cuts generated; used in solution approach 1 for the joint chance constrained problem

$\underline{\widehat{\Phi}}$ Set of optimality cuts generated; used in solution approach 2 for the joint, hybrid, and individual chance constrained problems

$\Psi^{\omega' j}$ Set of extreme rays of $\wp^{\omega' j}$; indexed by ψ

$\Psi^{\omega'}$ Set of extreme rays of (5.47) - (5.52) for scenario $\omega' \in \Omega'$; indexed by ψ

$\underline{\Psi}^{\omega'} \subseteq \Psi^{\omega'}$ Set of infeasibility cuts generated; used in solution approach 1 for the joint chance constrained problem

$\underline{\widehat{\Psi}}$ Set of infeasibility cuts generated; used in solution approach 2 for the joint chance constrained problem

$\underline{\widehat{\Psi}}_j$ Set of infeasibility cuts generated for candidate facility $j \in J$; used in solution approach 2 for the hybrid and individual chance constrained problems

Ω Set of all possible demand scenarios

$\Omega' \subseteq \Omega$ Subset of all possible demand scenarios; assumed finite

$\ddot{\Omega}$ Set of all possible capacity scenarios, assumed finite

$(\boldsymbol{\gamma}^{\omega' j\phi}, \boldsymbol{\pi}^{\omega' j\phi}, \eta^{\omega' j\phi}, \boldsymbol{\mu}^{\omega' j\phi}) \in \Phi^{\omega' j}$

$\left(\boldsymbol{\gamma}^{\omega' j\psi}, \boldsymbol{\pi}^{\omega' j\psi}, \eta^{\omega' j\psi}, \boldsymbol{\mu}^{\omega' j\psi}\right) \in \Psi^{\omega' j}$

Parameters

$a \geq 0$ Cost (in dollars) of transporting one item for one day

$b_j \geq 0$ Cost (in dollars) of holding one item in backlog for one day at facility $j \in J$

$b \geq 0$ Cost (in dollars) of holding one item in backlog for one day; not facility specific

$\hat{b}_j \geq 0$ Daily cost of purchasing one additional unit of capacity at facility $j \in J$

$\mathscr{b}_{jl}$ The incremental backlog cost associated with assigning an average of l demands to facility $j \in J$ each day rather than $l - 1$ demands

$c_{ij} = \begin{cases} 1 & \text{If } t_{ij} \leq t_{max} \\ 0 & \text{Otherwise} \end{cases}$

d^* Time horizon (in days; does not include warm-up period)

$\bar{d}_j$ Maximum number of days an item is allowed to be held in backlog at facility $j \in J$

f_j Daily fixed cost of locating at facility $j \in J$; in the cyclic-allocation model, the *average* daily fixed location cost at $j \in J$

h_i Average demand per unit of time generated at demand site $i \in I$

h_{id} Demand that is generated at demand site $i \in I$ on day $d \in D$

$\tilde{h}_{id}$ Random daily demand parameter generated at demand point $i \in I$ on day $d \in D$

h_{id}^{ω} Realization of random demand $\tilde{h}_{id}$ in scenario $\omega \in \Omega$, $\forall i \in I, d \in D$

$h_{id}^{\omega'}$ Realization of random demand $\tilde{h}_{id}$ in scenario $\omega' \in \Omega'$, $\forall i \in I, d \in D$

$\mathcal{I}$ Number of demand instances with a scenario size of $|\Omega'|$; used to determine appropriate number of scenarios to use in MIP-JCC; indexed by i

k_j Capacity of facility $j \in J$ in items processed per day

k_j^p Capacity of facility $j \in J$ in items processed per day on cycle day $p \in P$

$k_j^{max} \geq 0$ Maximum amount of additional capacity facility $j \in J$ can purchase each day

$\bar{k}_j$ Increment size of extra daily capacity available for purchase at facility $j \in J$

$\underline{k_j}$ Parameter such that $k_j - \underline{k_j}$ is the initial daily processing capacity of facility $j \in J$ before employees are trained

k_j^{rate} Rate of capacity increase due to employee learning

$\tilde{\ddot{k}}_{jd}$ Random daily capacity at facility $j \in J$ on day $d \in \{t^* + 1, \dots, |D|\}$

$\ddot{k}_{jd}^{\ddot{\omega}}$ Realization of random capacity $\tilde{\ddot{k}}_{jd}$ in scenario $\ddot{\omega} \in \ddot{\Omega}$ at facility $j \in J$ on day $d \in D$

m Maximum travel day difference allowed between a demand site and its weekday facility assignment, and the same demand site and its weekend facility assignment

M_1 Sufficiently large integer such that if $V_{jd}^{\omega'} > \theta$ for some facility $j \in J$ on some day $d \in \{t^* + 1, \dots, |D| + 1\}$ in some scenario $\omega' \in \Omega'$, then $V_{jd}^{\omega'} \leq \theta + M_1$

M_2 A sufficiently large integer such that M2S1 will not choose an allocation that results in $\sum_{\omega' \in \Omega'} p^{\omega'} Z^{\omega'} > \tau$

$\bar{M}_2$ A sufficiently large integer such that M2S1 will not choose an allocation that results in $\sum_{\omega' \in \Omega'} p^{\omega'} \bar{Z}_{jd}^{\omega'} > \tau$ for any $j \in J, d \in \{t^* + 1, \dots, |D| + 1\}$

$\hat{M}_2$ A sufficiently large integer such that M2S1 will not choose an allocation that results in $\sum_{\omega' \in \Omega'} p^{\omega'} \hat{Z}_j^{\omega'} > \tau$ for any $j \in J$

M_3 A sufficiently large integer such that if $ck_j + \theta < \sum_{i \in I} \sum_{s=0}^{c-1} h_{i,d+s-t_{ij}}^{\omega'}$ for any $c \in C, j \in J, \omega' \in \Omega', \ d \in \left\{t^* + 1, t^* + 1 + c, t^* + 1 + 2c, \ldots, \left\lfloor \frac{|D| - t^* - c}{c} \right\rfloor c + t^* + 1\right\}$ then $ck_j + \theta + M_3 \geq \sum_{i \in I} \sum_{s=0}^{c-1} h_{i,d+s-t_{ij}}^{\omega'}$ as long as $t^* + c \leq |D|$

n Maximum number of different facilities a demand site can be allocated to during the week; $\mathit{n} \in \{1, \ldots, |P|\}$

p^{ω} Probability of demand scenario $\omega \in \Omega$ occurring, with $\sum_{\omega \in \Omega} p^{\omega} = 1$

$p^{\omega'}$ Probability of demand scenario $\omega' \in \Omega'$ occurring, with $\sum_{\omega' \in \Omega'} p^{\omega'} = 1$

$\ddot{p}^{\ddot{\omega}}$ Probability of capacity scenario $\ddot{\omega} \in \ddot{\Omega}$ occurring, with $\sum_{\ddot{\omega} \in \ddot{\Omega}} \ddot{p}^{\ddot{\omega}} = 1$

$\mathfrak{p}_{\mathfrak{n}\mathfrak{m}}$ Transition probability from state $\mathfrak{n}$ to state $\mathfrak{m}$

t_{ij} Travel time (in days) between demand site $i \in I$ and candidate facility $j \in J$

$t^* = max_{i \in I, j \in J} t_{ij}$ Longest travel time between any demand site – candidate processing facility pair

t_{max} Maximum allowed travel time between a demand site and its assigned facility

$\mathcal{T}$ Interarrival time of *G/G/1* queue

v_j Initial backlog (at the beginning of day $t^* + 1$) at facility $j \in J$ if facility j is located

$\mathfrak{v}_i$ Variance of demand at demand site $i \in I$ (used in queueing formulation)

w Weight given to the population in the fixed facility location costs

z Daily base fixed location cost per facility

δ Parameter such that $1 - \delta$ represents the confidence that $\ddot{V}^{\tau' \omega'[i]}$ is a lower bound for the optimal JCC objective function value

ζ Total average daily demand generated from the demand sites

ς The number of closest facilities a demand site can be allocated to; e.g., $\varsigma = 4$ means that the only facilities a demand site $i \in I$ can be allocated to are those that are one of the four closest located facilities to $i \in I$

θ Maximum desired daily backlog level at an individual facility

$\xi(\omega) \in \mathbb{N}_0^{|I|\times|D|}$ Vector containing the demand realizations of scenario $\omega \in \Omega$; has elements $h_{id}^{\omega}, \forall i \in I, d \in D$

$\xi'(\omega') \in \mathbb{N}_0^{|I|\times|D|}$ Vector containing the demand realizations of scenario $\omega' \in \Omega'$; has elements $h_{id}^{\omega'}, \forall i \in I, d \in D$; often abbreviated ξ'

$\xi'(\omega', \dot{\imath})$ Vector containing the demand realizations of scenario $\omega' \in \Omega'$ in replication $\dot{\imath} \in \{1, 2, \dots, \mathcal{I}\}$

$\ddot{\xi}(\ddot{\omega}) \in \mathbb{N}_0^{|J|\times|D|}$ Vector containing the capacity realizations of scenario $\ddot{\omega} \in \ddot{\Omega}$; has elements $\ddot{k}_{jd}^{\ddot{\omega}}, \forall j \in J, d \in \{t^* + 1, \dots, |D|\}$

ρ_j Population corresponding to the county in which candidate facility $j \in J$ is located

τ Maximum acceptable probability of the backlog at any facility exceeding θ on any day

τ' Maximum acceptable probability of the backlog at any facility exceeding θ on any day; used in the sample approximation problem MIP-JCC in Section 5.5.1 when performing sample size calculations

Decision Variables

$\widehat{K}_{jd}$ Amount of additional unit capacity to purchase at facility $j \in J$ on day $d \in \{t^* + 1, \dots, |D|\}$

$\widehat{K}_j^p$ Amount of additional unit capacity to purchase at facility $j \in J$ for cycle day $p \in P$

$\overline{K}_{jd}$ Number of extra capacity increments purchased at facility $j \in J$ on day $d \in \{t^* + 1, \dots, |D|\}$

$E_i = \begin{cases} 1 & \text{If demand site } i \in I \text{ is multi} - \text{sourced} \\ 0 & \text{Otherwise} \end{cases}$

$G_{ij\hat{\varsigma}} = \begin{cases} 1 & \text{If facility } j \in J \text{ is the } \hat{\varsigma} \text{ closest located facility to demand site } i \in I \\ 0 & \text{Otherwise} \end{cases}$

$N_{ij} = \begin{cases} 1 & \text{If demand site } i \in I \text{ allocates any of its demand to facility } j \in J \\ 0 & \text{Otherwise} \end{cases}$

$q_{ij} = \begin{cases} 1 & \text{If demands at } i \in I \text{ are allocated to facility } j \in J \text{ on a day such that they would arrive on a weekend} \\ 0 & \text{Otherwise} \end{cases}$

$\mathbb{Q} = \begin{cases} 1 & \text{If the } \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{V}, \text{and } \boldsymbol{W} \text{ values result in a violation of the reformulated joint chance constraints } (5.10) - (5.12) \\ 0 & \text{Otherwise} \end{cases}$

$\overline{\mathbb{Q}} = \begin{cases} 1 & \text{If the } \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{V}, \text{and } \boldsymbol{W} \text{ values result in a violation of the reformulated individual chance constraints } (5.80) - (5.82) \\ 0 & \text{Otherwise} \end{cases}$

$\widehat{\mathbb{Q}} = \begin{cases} 1 & \text{If the } \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{V}, \text{and } \boldsymbol{W} \text{ values result in a violation of the reformulated hybrid chance constraints } (5.83) - (5.85) \\ 0 & \text{Otherwise} \end{cases}$

$r_{ij} = \begin{cases} 1 & \text{If demands at } i \in I \text{ are assigned to facility } j \in J \text{ on a day such that they would arrive on a weekday} \\ 0 & \text{Otherwise} \end{cases}$

$s_{ij} = \begin{cases} 1 & \text{If demand site } i \in I \text{ is allocated to faciliy } j \in J \text{ on any day of the week} \\ 0 & \text{Otherwise} \end{cases}$

$U^{\omega'} = \begin{cases} 1 & \text{Otherwise} \\ 0 & \text{If for any } c \in C \text{ the amount of arriving demand at any facility } j \in J \text{ exceeds its corresponding } ck_j + \theta \text{ value on any day in scenario } \omega' \in \Omega' \end{cases}$

V_{jd} Auxiliary decision variable representing the backlog level at facility $j \in J$ at the beginning of day $d \in \{t^* + 1, \dots, |D| + 1\}$

$\tilde{V}_{jd}$ Uncertain backlog level at facility $j \in J$ on the beginning of day $d \in \{t^* + 1, \dots, |D| + 1\}$ due to stochasticity in the demand

V_{jd}^{ω} Auxiliary decision variable representing the realized backlog level at facility $j \in J$ at the beginning of day $d \in \{t^* + 1, \dots, |D| + 1\}$ in demand scenario $\omega \in \Omega$

$V_{jd}^{\omega'}$ Auxiliary decision variable representing the realized backlog level at facility $j \in J$ at the beginning of day $d \in \{t^* + 1, \dots, |D| + 1\}$ in demand scenario $\omega' \in \Omega'$

$V_{jd}^{\omega'\psi}$ Backlog level in scenario $\omega' \in \Omega'$ at facility $j \in J$ at the beginning of day $d \in \{t^* + 1, \dots, |D| + 1\}$ in infeasible solution $\psi \in \underline{\widehat{\Psi}}$

$\bar{V}^{\omega' j}$ Decision variable in RD-M1S2(ω', j) whose optimal value is equal to the optimal objective function value of RD-M1S2(ω, j)

$\bar{V}^{\omega'}$ Decision variable in ARD-M1S2(ω') whose optimal value is equal to the optimal objective function value of ARD-M1S2(ω'); $\bar{V}^{\omega'} = \sum_{j \in J} \bar{V}^{\omega' j}$

$\tilde{\ddot{V}}_{jd}$ Uncertain backlog level at facility $j \in J$ on the beginning of day $d \in \{t^* + 1, \dots, |D| + 1\}$, due to stochasticity in the processing capacity

$\ddot{V}_{jd}^{\ddot{\omega}}$ Auxiliary decision variable representing the realized backlog level at facility $j \in J$ at the beginning of day $d \in \{t^* + 1, \dots, |D| + 1\}$ in capacity scenario $\ddot{\omega} \in \ddot{\Omega}$

$\boldsymbol{V}^{\omega' j}$ A $1 \times (|D| + 1 - t^*)$ vector whose $(d - t^*)^{\text{th}}$ element is $V_{jd}^{\omega'}$

$\boldsymbol{V}$ A $|J| \times (|D| + 1 - t^*) \times |\Omega'|$ array whose $(j, d - t^*, \omega')^{\text{th}}$ element is $V_{jd}^{\omega'}$

$\mathbb{V}$ Decision variable in RMP^n-M1S1 that represents the expected total number of items in backlog over the planning horizon

$\hat{\mathbb{V}}$ Decision variable in R-M2S1 and MIP-ICC_R-M2S1 that represents the expected total number of items in backlog over the planning horizon

W_{jd} Auxiliary decision variable representing the number of items that are processed at facility $j \in J$ on day $d \in \{t^* + 1, \dots, |D|\}$

W_{jd}^{ω} Auxiliary decision variable representing the number of items that are processed at facility $j \in J$ on day $d \in \{t^* + 1, \dots, |D|\}$ in demand scenario $\omega \in \Omega$

$W_{jd}^{\omega'}$ Auxiliary decision variable representing the number of items that are processed at facility $j \in J$ on day $d \in \{t^* + 1, \dots, |D|\}$ in demand scenario $\omega' \in \Omega'$

$\ddot{W}_{jd}^{\ddot{\omega}}$ Number of items that are processed at facility $j \in J$ on day $d \in \{t^* + 1, \dots, |D|\}$ in capacity scenario $\ddot{\omega} \in \ddot{\Omega}$

$\boldsymbol{W}^{\omega' j}$ A $1 \times (|D| - t^*)$ vector whose $(d - t^*)^{\text{th}}$ element is $W_{jd}^{\omega'}$

$\boldsymbol{W}$ A $|J| \times (|D| - t^*) \times |\Omega'|$ array whose $(j, d - t^*, \omega')^{\text{th}}$ element is $W_{jd}^{\omega'}$

$$X_j = \begin{cases} 1 & \text{If we locate at facility } j \in J \\ 0 & \text{Otherwise} \end{cases}$$

$\boldsymbol{X}$ A $1 \times |J|$ vector whose j^{th} element is X_j

Y_{ij} Fraction of demand from demand site $i \in I$ that is allocated to facility $j \in J$; if single sourcing is imposed, Y_{ij} is binary

$$Y_{ij}^{p} = \begin{cases} 1 & \text{If we allocate demands from demand site } i \in I \text{ to faciliy } j \in J \text{ on cycle day } p \in P \\ 0 & \text{Otherwise} \end{cases}$$

$Y_{ij}^{\psi} = \begin{cases} 1 & \text{If we allocate demands from demand site } i \in I \text{ to faciliy } j \in J \text{ in infeasible solution } \psi \\ 0 & \text{Otherwise} \end{cases}$

$\boldsymbol{Y}$ A $|I| \times |J|$ matrix whose $(i,j)^{\text{th}}$ entry is Y_{ij}

$Z^{\omega'} = \begin{cases} 1 & \text{If the backlog at any facility exceeds } \theta \text{ on any day in scenario } \omega' \in \Omega' \\ 0 & \text{Otherwise} \end{cases}$

$\hat{Z}_j^{\omega'} = \begin{cases} 1 & \text{If the backlog at facility } j \in J \text{ exceeds } \theta \text{ on any day in scenario } \omega' \in \Omega' \\ 0 & \text{Otherwise} \end{cases}$

$\bar{Z}_{jd}^{\omega'} = \begin{cases} 1 & \text{If the backlog at facility } j \in J \text{ exceeds } \theta \text{ on day } d \in \{t^*+1, \dots, |D|+1\} \text{ in scenario } \omega' \in \Omega' \\ 0 & \text{Otherwise} \end{cases}$

$\mathcal{Z}_{jl} = \begin{cases} 1 & \text{If the average amount of daily demand allocated to facility } j \in J \text{ is at least } l \\ 0 & \text{Otherwise} \end{cases}$

$\boldsymbol{Z}$ A $1 \times |\Omega'|$ vector whose ω'^{th} element is $Z^{\omega'}$

$\bar{\boldsymbol{Z}}_{jd}$ A $1 \times |\Omega'|$ vector whose ω'^{th} element is $\bar{Z}_{jd}^{\omega'}$

$\hat{\boldsymbol{Z}}_j$ A $1 \times |\Omega'|$ vector whose ω'^{th} element is $\hat{Z}_j^{\omega'}$

$\gamma_d^{\omega' j}$ Dual variables corresponding to a backlog balance constraint of (5.26)

$\gamma_d^{\omega' j \psi}$ Component of $\Psi^{\omega' j}$ corresponding to $\gamma_d^{\omega' j}$

$\gamma_d^{\omega' j \phi}$ Component of $\Phi^{\omega' j}$ corresponding to $\gamma_d^{\omega' j}$

$\boldsymbol{\gamma}$ A $|J| \times (|D| - t^*) \times |\Omega'|$ array whose $(j, d-t^*, \omega')^{\text{th}}$ element is $\gamma_d^{\omega' j}$

$\boldsymbol{\gamma}^{\omega'}$ A $|J| \times (|D| - t^*)$ matrix whose $(j, d-t^*)^{\text{th}}$ element is $\gamma_d^{\omega' j}$

$\boldsymbol{\gamma}^{\omega' j}$ A $(|D| - t^*)$ vector whose $(d-t^*)^{\text{th}}$ element is $\gamma_d^{\omega' j}$

$\boldsymbol{\gamma}^{\omega' \psi}$ A $|J| \times (|D| - t^*)$ matrix whose $(j, d-t^*)^{\text{th}}$ element is $\gamma_d^{\omega' j \psi}$

$\boldsymbol{\gamma}^{\omega' \phi}$ A $|J| \times (|D| - t^*)$ matrix whose $(j, d-t^*)^{\text{th}}$ element is $\gamma_d^{\omega' j \phi}$

$\boldsymbol{\gamma}^{\omega' j\psi}$ The $(|D| - t^*)$ vector component of $\Psi^{\omega' j}$ corresponding to $\boldsymbol{\gamma}^{\omega' j}$ whose $(d - t^*)^{\text{th}}$ element is $\gamma_d^{\omega' j\psi}$

$\boldsymbol{\gamma}^{\omega' j\phi}$ The $(|D| - t^*)$ vector component of $\Phi^{\omega' j}$ corresponding to $\boldsymbol{\gamma}^{\omega' j}$ whose $(d - t^*)^{\text{th}}$ element is $\gamma_d^{\omega' j\phi}$

$\mu_d^{\omega' j}$ Dual variable corresponding to a desired maximum backlog constraint of (5.29)

$\mu_d^{\omega' j\psi}$ Component of $\Psi^{\omega' j}$ corresponding to $\mu_d^{\omega' j}$

$\mu_d^{\omega' j\phi}$ Component of $\Phi^{\omega' j}$ corresponding to $\mu_d^{\omega' j}$

$\boldsymbol{\mu}$ A $|J| \times (|D| + 1 - t^*) \times |\Omega'|$ array whose $(j, d - t^*, \omega')^{\text{th}}$ element is $\mu_d^{\omega' j}$

$\boldsymbol{\mu}^{\omega'}$ A $|J| \times (|D| + 1 - t^*)$ vector whose $(j, d - t^*)^{\text{th}}$ element is $\mu_d^{\omega' j}$

$\boldsymbol{\mu}^{\omega' j}$ A $(|D| + 1 - t^*)$ vector whose $(d - t^*)^{\text{th}}$ element is $\mu_d^{\omega' j}$

$\boldsymbol{\mu}^{\omega'\psi}$ A $|J| \times (|D| + 1 - t^*)$ vector whose $(j, d - t^*)^{\text{th}}$ element is $\mu_d^{\omega' j\psi}$

$\boldsymbol{\mu}^{\omega'\phi}$ A $|J| \times (|D| + 1 - t^*)$ vector whose $(j, d - t^*)^{\text{th}}$ element is $\mu_d^{\omega' j\phi}$

$\boldsymbol{\mu}^{\omega' j\psi}$ The $(|D| + 1 - t^*)$ vector component of $\Psi^{\omega' j}$ corresponding to $\boldsymbol{\mu}^{\omega' j}$ whose $(d - t^*)^{\text{th}}$ element is $\mu_d^{\omega' j\psi}$

$\boldsymbol{\mu}^{\omega' j\phi}$ The $(|D| + 1 - t^*)$ vector component of $\Phi^{\omega' j}$ corresponding to $\boldsymbol{\mu}^{\omega' j}$ whose $(d - t^*)^{\text{th}}$ element is $\mu_d^{\omega' j\phi}$

$\eta^{\omega' j}$ Dual variable corresponding to an initial backlog constraint of (5.28)

$\eta^{\omega' j\psi}$ The component of $\Psi^{\omega' j}$ corresponding to $\eta^{\omega' j}$

$\eta^{\omega' j\phi}$ A component of $\Phi^{\omega' j}$ corresponding to $\eta^{\omega' j}$

$\boldsymbol{\eta}$ A $|J| \times |\Omega'|$ array whose $(j, \omega')^{\text{th}}$ element is $\eta^{\omega' j}$

$\boldsymbol{\eta}^{\omega'}$ A $|J|$ vector whose j^{th} element is $\eta^{\omega' j}$

$\boldsymbol{\eta}^{\omega'\psi}$ A $|J|$ vector whose j^{th} element is $\eta^{\omega' j\psi}$

$\boldsymbol{\eta}^{\omega'\phi}$ A $|J|$ vector whose j^{th} element is $\eta^{\omega' j\phi}$

$\pi_d^{\omega' j}$ Dual variable corresponding to a capacity constraint of (5.27)

$\pi_d^{\omega' j\psi}$ Component of $\Psi^{\omega' j}$ corresponding to $\pi_d^{\omega' j}$

$\pi_d^{\omega' j\phi}$ Component of $\Phi^{\omega' j}$ corresponding to $\pi_d^{\omega' j}$

$\boldsymbol{\pi}$ A $|J| \times (|D| - t^*) \times |\Omega'|$ array whose $(j, d - t^*, \omega')^{\text{th}}$ element is $\pi_d^{\omega' j}$

$\boldsymbol{\pi}^{\omega'}$ A $|J| \times (|D| - t^*)$ matrix whose $(j, d - t^*)^{\text{th}}$ element is $\pi_d^{\omega' j}$

$\boldsymbol{\pi}^{\omega' j}$ A $(|D| - t^*)$ vector whose $(d - t^*)^{\text{th}}$ element is $\pi_d^{\omega' j}$

$\boldsymbol{\pi}^{\omega' j\psi}$ The $(|D| - t^*)$ vector component of $\Psi^{\omega' j}$ corresponding to $\boldsymbol{\pi}^{\omega' j}$ whose $(d - t^*)^{\text{th}}$ element is $\pi_d^{\omega' j\psi}$

$\boldsymbol{\pi}^{\omega' j\phi}$ The $(|D| - t^*)$ vector component of $\Phi^{\omega' j}$ corresponding to $\boldsymbol{\pi}^{\omega' j}$ whose $(d - t^*)^{\text{th}}$ element is $\pi_d^{\omega' j\phi}$

$\boldsymbol{\pi}^{\omega'\psi}$ A $|J| \times (|D| - t^*)$ matrix whose $(j, d - t^*)^{\text{th}}$ element is $\pi_d^{\omega' j\psi}$

$\boldsymbol{\pi}^{\omega'\phi}$ A $|J| \times (|D| - t^*)$ matrix whose $(j, d - t^*)^{\text{th}}$ element is $\pi_d^{\omega' j\phi}$

Optimal Objective Function Values

$\bar{V}^{\omega' j}(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z})$ Optimal objective function value of M1S2(ω', j) given $\boldsymbol{X}, \boldsymbol{Y}$, and $\boldsymbol{Z}$ as input

$\bar{V}_{Dual}^{\omega' j}(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z})$ Optimal objective function value of D-M1S2(ω', j) and RD-M1S2(ω', j) given $\boldsymbol{X}, \boldsymbol{Y}$, and $\boldsymbol{Z}$ as input

$\bar{V}_{A-Dual}^{\omega'}(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z})$ Optimal objective function value of ARD-M1S2(ω') given $\boldsymbol{X}, \boldsymbol{Y}$, and $\boldsymbol{Z}$ as input

$\bar{V}_{RMP}^{n}(\boldsymbol{\gamma}, \boldsymbol{\pi}, \boldsymbol{\eta}, \boldsymbol{\mu})$ Optimal objective function value of RMP^n-M1S1 given $\boldsymbol{\gamma}, \boldsymbol{\pi}, \boldsymbol{\eta}$ and $\boldsymbol{\mu}$ as input

$\bar{V}_{Sep}^{\omega' j}(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z})$ Optimal objective function value of the separation problem corresponding to D-M1S2(ω', j) given $\boldsymbol{X}, \boldsymbol{Y}$, and $\boldsymbol{Z}$ as input

$\bar{V}_{Sep}^{\omega'}(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}) = \sum_{j \in J} \bar{V}_{Sep}^{\omega' j}(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z})$

$\hat{V}^{\omega' j}(\boldsymbol{X}, \boldsymbol{Y})$ Optimal objective function value of M2S2(ω', j) given $\boldsymbol{X}$ and $\boldsymbol{Y}$ as input

$\hat{V}_{Dual}^{\omega' j}(\boldsymbol{X}, \boldsymbol{Y})$ Optimal objective function value of D-M2S2(ω', j) given $\boldsymbol{X}$ and $\boldsymbol{Y}$ as input

$\ddot{V}$ Optimal objective function value of MIP-JCC

$\ddot{V}^{\tau' \omega' i}$ Optimal objective function value of MIP-JCC with τ' for scenario $\omega' \in \Omega'$ of replication $i \in \{1,2, \dots, \mathcal{I}\}$

$\ddot{V}^{\tau' \omega' [i]}$ Optimal objective function value of MIP-JCC with τ' for the i^{th} order statistic in scenario $\omega' \in \Omega'$

$\hat{Z}(\boldsymbol{V})$ Optimal objective function value of M2S3 given $\boldsymbol{V}$ as input

$\bar{\hat{Z}}(\boldsymbol{V})$ Optimal objective function value of MIP-ICC_M2S3(j, d) given $\boldsymbol{V}$ as input

$\hat{\hat{Z}}(\boldsymbol{V})$ Optimal objective function value of MIP-HCC_M2S3(j) given $\boldsymbol{V}$ as input

Other

$\mathbb{E}(\cdot)$ Expected value of $\cdot$

$\mathbb{E}[V_j]$ Long-term expected number of items in backlog at facility $j \in J$

$H_{j,d} = \sum_{i \in I} h_{i,t-t_{ij}} Y_{ij}$ The total demand that arrives at facility $j \in J$ on day $d \in \{t^* + 1, \dots, |D|\}$

H_j Generic total amount of demand arriving at facility $j \in J$ on any day of the planning horizon

$\mathbb{N}_0^{|I| \times |D|}$ Vector of dimension $|I| \times |D|$ which contains non-negative integer elements

n Iteration counter

$\mathbb{P}(\cdot)$ Probability of event $\cdot$ occurring

κ Total daily capacity of located processing facilities

ABSTRACT

In many real-world settings, the capacity of processing centers is flexible due to a variety of operational tools (such as overtime, outsourcing, and backlogging demand) available to managers that allow the facility to accept demands in excess of the capacity constraint for short periods of time. However, most capacitated facility location models in the literature today impose hard capacity constraints that don't capture this short term flexibility. Thus, current capacitated facility location models do not account for the operational costs associated with accepting excess daily demand, which can lead to suboptimal facility location and demand allocation decisions.

To address this discrepancy, we consider a processing distribution system in which demand generated on a daily basis by a set of demand sites is satisfied by a set of capacitated processing facilities. At each demand site, daily demands for the entirety of the planning horizon are sampled from a known demand distribution. Thus, the day to day demand fluctuations may result in some days for which the total demand arriving at a processing facility exceeds the processing capacity, even if the average daily demand arriving at the processing facility is less than the daily processing capacity. We allow each processing facility the ability to hold excess demand in backlog to be processed at a later date and assess a corresponding backlog penalty in the objective function for each day a unit of demand is backlogged.

This dissertation primarily focuses on three methods of modelling the aforementioned processing distribution system. The first model is the Inventory Modulated Capacitated Location Problem (IMCLP), which utilizes disaggregated daily demand parameters to determine the

subset of processing facilities to establish, the allocation of demand sites to processing facilities, and the magnitude of backlog at each facility on each day that minimizes location, travel, and backlogging costs. Whereas the IMCLP assumes each demand site must be allocated to exactly one processing facility, the second model relaxes this assumption and allows demand sites to be allocated to different processing facilities on various days of the week. We show that such a cyclic allocation scheme can further reduce the system costs and improve service metrics as compared to the IMCLP.

Finally, while the first two models incorporate daily fluctuations in demand over an extended time horizon, the problems remain deterministic in the sense that only one realization of demand is considered for each day of the planning horizon. As such, our final model presents a stochastic version of the IMCLP in which we assume a known demand distribution but assume the realization of daily demand is uncertain. In addition to assessing a penalty cost, we consider three types of chance constraints to restrict the amount of backlogged demand to a predetermined threshold. Using finite samples of random demand, we propose two multi-stage decomposition schemes and solve the mixed-integer programming reformulations with cutting-plane algorithms.

In summary, this dissertation mitigates hard capacity constraints commonly found in facility location models by allowing incoming demand to exceed the processing capacity for short periods of time. In each of the modelling contexts presented, we show that the location and allocation decisions obtained from our models can result in significantly reduced costs and improved service metrics when compared to models that do not account for the likelihood that demands may exceed capacity on some days.

CHAPTER 1: Introduction

Facility location plays a critical role in an organization's expenses and customer service as location decisions affect at least three key elements of a supply chain: fixed location costs, transportation costs, and the ability to provide service in a timely manner. Locating many facilities typically increases the organization's location costs but reduces transportation costs, while locating fewer facilities may reduce the location costs but increase transportation costs and drastically degrade customer service. Even if the correct number of facilities are located, poorly located facilities can negatively affect customer service and result in increased location and transportation costs. As such, facility location decisions are applicable to a broad range of areas including locating warehouses, processing plants, schools, airline hubs, hospitals, ambulances, military bases, disaster relief shelters and hazardous waste disposal sites. Location models have also been used in less traditional settings, such as database location in computer networks [Fisher and Hochbaum, 1980], the analysis of archeological sites [Bell and Church, 1985], vehicle routing [Bramel and Simchi-Levi, 1995], medical diagnosis [Reggia et al., 1983], and the alignment of candidates along a political spectrum [Ginsberg et al., 1987].

In general, facility location problems involve a set of spatially distributed customers, the location of which are known; and a set of facilities to serve customer demands. Possible questions facility location models can help answer are:

(1) How many facilities should be located?

(2) Where should the facilities be located?

(3) How large should each facility be?

(4) How should the customer demand be allocated to the located facilities?

Furthermore, the answers to these questions depend on the decision maker's objective. For example, the answers may differ depending on whether we wish to minimize cost, maximize customer service, or ensure all demand is met within a certain time frame.

The modeling foundation of this dissertation is based on the capacitated fixed charge location problem (CFLP) [Balinski, 1965]. The CFLP models an environment in which demands generated by a set of spatially-dispersed demand sites are transported to facilities for processing. While the locations of the demand sites are known in advance, only the *potential* locations for the processing facilities are known before solving the model. The premise of the CFLP is to determine how many and which of the potential processing facilities should be located, as well as how to allocate the demand to the located processing facilities so that the sum of the location and allocation costs is minimized. Furthermore, the decisions must be made in a manner that ensures that on each day, the number of demands arriving at each facility does not exceed the daily processing capacity of the facility.

1.1 The Necessity of Mitigating Hard Capacity Constraints

While capacitated location models are abundant in the literature, nearly all of the models utilize capacity constraints that are problematic for at least three key reasons. This dissertation identifies methods of addressing each of these issues and quantifies the benefit gained by utilizing improved models.

The first issue with the traditional capacity constraints is that they disregard the reality that the processing capacity of a facility is a complex function of many operational decisions and that facility managers often incorporate techniques to enable capacity flexibility. For example,

simply stating that an automobile manufacturing plant can assemble 1,000 vehicles per day does not capture the reality that the actual number of vehicles assembled depends on the number of each type of vehicle assembled and the sequence in which the vehicles are assembled. Additionally, a facility manager may utilize overtime, thus allowing the facility to process more than the specified daily capacity limit.

The second issue is that even in situations in which the processing capacity can be precisely determined, facility managers typically have operational tools that allow the facility to accept demands in excess of the stated capacity limit for short periods of time. For example, demands that exceed capacity may be stored as backlog and processed at a later date. Traditional capacity constraints do not allow for this. Instead, they employ hard capacity constraints on the number of items that can be processed each day and assume that the amount of demand that arrives at a facility for processing cannot exceed the capacity on any day (see Daskin et al. (2005) and Verter (2011) for reviews).

The third issue is that by using *average* daily demands – as is common in many operations research models in general, and location problems in particular – the traditional models fail to capture the likelihood that demand will exceed capacity on some days. For example, suppose A and B are two processing facilities, each with a capacity to process 100 units per day, and that the total amount of demand allocated to each facility follows a Poisson distribution with a mean of 95 units. Then, the daily probability of exceeding the capacity at either facility A or facility B individually is 0.282 and the probability that the demand will exceed the capacity for at least one facility on any given day is 0.485. As this simple example shows, the capacity will be exceeded on nearly half of the days, although the average demand allocated to each facility is less than the capacity.

Additionally, considering demands at an aggregate level inherently fails to provide information regarding possible temporal and spatial correlations in demands. One method of including such information into the models is to allow the data to serve as a direct model input. Saveh-Shemshaki et al. (2012) provides a seminal paper in this area that explicitly incorporates historical data directly into an extension of the traditional capacitated location model. However, as we discuss in the following section, Saveh-Shemshaki et al. (2012) allows the allocation decisions to vary each day of the planning horizon and supposes all future demands at each demand site are known in advance, which incorporates unrealistic foreknowledge into the daily allocation decisions.

A common theme linking the four aforementioned issues is that, in most facility location settings, a cost is incurred when the stated capacity is exceeded; facilities may have a cost associated with allocating additional resources to process the extra demand or loss of goodwill due to decreased customer service levels. Since traditional capacitated models neither account for capacity flexibility nor allow incoming demands to exceed capacity, they inherently underestimate the total facility, transportation, and backlog costs. Furthermore, as shown in this dissertation, they often identify the wrong facilities or even the wrong *number* of facilities to locate.

1.2 Research Contributions

To address these issues with current capacitated facility location models, we propose a new method of modeling capacity constraints. Our contributions to the facility location literature in terms of modeling, solution methodology, and managerial insights are as follows:

Modeling:

1. Capacity Flexibility: We mitigate the hard capacity constraints used in traditional facility location models by incorporating capacity flexibility. While we identify a variety of methods that can be used to incorporate capacity flexibility, we focus our modeling efforts on allowing any demands in excess of the processing capacity that arrive at a facility on a particular day to be processed on a following day. Such demands incur a penalty cost to account for the delayed processing. (That is, we focus on capacity flexibility with regard to the amount of arriving demand rather than the processing capacity; in most of the models presented in this dissertation, processing capacity is fixed, although we outline extensions that allow the processing capacity to be determined endogenously in Chapter 6.) Consistent with the CFLP, our problem is to determine the subset of the candidate processing facilities to establish and the allocation of demand sites to facilities. However, in our new model, once the facility locations and demand allocations are known, the demand stream and the daily processing capacity determine the unprocessed backlog carried over from one day to the next.

   Blood testing facilities are one example of a processing system that may benefit from a modeling framework such as the one we suggest in this dissertation. Blood samples that are drawn from patients at a clinic (i.e., the demand site) often need to be sent to an off-site regional testing facility to be analyzed [Saveh-Shemshaki et al., 2012]. After the blood sample has been analyzed, the testing facility sends a report back to the clinic indicating the results. Thus, physical demand (e.g., blood samples) is

shipped to the testing facility, but intangible results (e.g., an electronic summary) are reported back to the clinic and patient. A traditional capacitated facility location model would use hard capacity constraints to require that all of the blood samples that arrive at the testing facility each day be tested on the day they arrive. However, in practice, any additional samples that cannot be tested on the day they arrive will be held in cold storage (i.e., backlogged) and tested the next day [Adcock et al., 2012; Wong et al., 2013]. In this dissertation, we incorporate the ability to backlog demand and, furthermore, assess a penalty cost for each blood sample that is not processed on the day it arrives at the testing facility. This cost may represent electricity or inventory costs associated with using cold storage, loss-of-goodwill with the clinic due to longer wait times, etc.

While models that incorporate the option to backlog demand do exist within the current supply chain management literature, the approach we present in this dissertation differs from these contexts in two important ways: 1) The supply chain management literature assumes physical demand is shipped from processing facilities to a customer, whereas physical demands is shipped from a customer to a processing facility in the facility location context that we consider. Thus, in our context, the amount of demand available for processing at each facility each day is a function of the allocation policy as well as the travel time from the customer to the facility. 2) The supply chain management literature incorporates foreknowledge about future demand that, while appropriate for particular supply chain management contexts, is problematic for the facility location context studied in this dissertation.

In addition to utilizing backlogs as a method of incorporating daily capacity flexibility into facility location models, we identify and formulate a number of other approaches that could be considered, including utilizing overtime, outsourcing, temporary workers, and extra shifts. We also formulate ways in which endogenous capacity can be incorporated. These extensions are introduced in Chapter 6.5 and their analyses are left as an area of future work.

2. Cyclic Allocations: The first model we present assumes each demand site allocates all of its demand to a single processing facility for the entirety of the planning horizon. The second model relaxes this assumption and allows for demand to be allocated in a cyclic manner. That is, demands can be allocated to multiple processing facilities over a specific time frame (e.g., a week) but to a single facility each day. This enables the model to develop a day-of-the-week allocation scheme that considers day-to-day variations in the daily processing capacity levels of a set of candidate processing facilities and/or systematic day-to-day demand variations. We demonstrate that allowing demands at a particular site to be allocated to multiple processing facilities in such a manner can be a cost effective operational tool.

3. Daily Demand: Rather than utilizing aggregated demand parameters, we consider demands at a daily level, which allows us to explicitly incorporate the day-to-day variation in and possibly correlated nature of demands. We do this by using a daily demand dataset as a direct input into the strategic decision making model to serve as a

representation of the demand pattern. As such, our data-driven model is also a step toward incorporating the vast quantities of transactional data that are being generated on a daily basis into location models. Furthermore, variations of our data-driven model have the potential to reveal operating policies that take advantage of spatial and temporal correlations in demand that are not evident in current facility location models.

Additionally, incorporating daily demands means that the time it takes to transport demands from a demand site to a processing facility must be considered in the capacity constraints since different amounts of demand are generated each day and arrive for processing at a later date. This allows us to capture the amount of unprocessed demand from each day that will be added to the amount of demand awaiting processing on the following day.

4. Multiple Methods of Limiting Backlog: In addition to deterministic models, we present a stochastic model with uncertain demand. Rather than only assessing a backlog cost for each day an item spends backlogged, or only using chance constraints to bound the probability of having backlog exceed a predetermined threshold, we incorporate both methods into the model formulation. Furthermore, we consider three different types of chance constraints: (1) joint chance constraints that ensure the probability of any processing facility having a backlog level above the threshold on any day of the planning horizon is sufficiently small, (2) individual chance constraints to limit the amount of backlog at each facility each day, and (3) a

hybrid approach which accounts for the probability that each individual processing facility will exceed the stated maximum backlog level on any day of the planning horizon.

Solution Methodology:

The methodological contributions of this work include solving a stochastic extension of the new capacitated facility location model formulation by exploiting the model structure. This allows us to utilize a three-stage decomposition approach in which the first stage problem is precisely the traditional uncapacitated fixed charge location problem (UFLP) [Balinski, 1965]. We also present a two-stage decomposition approach and introduce a set of first-stage strengthening constraints that can be used in either the two-stage or three-stage decomposition approach. The multi-stage decomposition schemes are then transformed into mixed-integer programming reformations and solved with cutting-plane algorithms.

Managerial Insights:

By incorporating the ability to hold excess demand as backlog and process the demand at a later time, our capacitated location model better reflects actual managerial options than previous modeling formulations. We provide managerial insights into the effect of the cost associated with demands arriving at a facility that exceed the stated daily processing capacity as it pertains to location and allocation decisions, customer service levels, and overall costs. We also discuss the circumstances under which it is beneficial to utilize our data-driven, flexible capacity model as compared to a location model that employs average demands and hard capacity constraints. Insights are also given regarding the benefits of incorporating a flexible

cyclic allocation policy that can leverage systematic day-of-week patterns in demand or capacity levels that vary throughout the week.

1.3 Outline

The remainder of this dissertation is organized as follows. In Chapter 2, we review the literature on facility location models with a specific focus on capacitated models. It is here that we formally define the CFLP. We also identify current literature that considers capacity flexibility, allocation flexibility, and location models with daily demand parameters. In Chapter 3 we develop the Inventory Modulated Capacitated Location Problem (IMCLP)[1], which addresses the aforementioned issues with traditional location models. We discuss the benefits of the IMCLP in comparison to the CFLP, and present computational results from large problem instances. In Chapter 4, we extend the IMCLP by allowing for cyclic allocations, and show that such an extension can further decrease the overall cost of the system. While Chapters 3 and 4 detail deterministic models, Chapter 5 incorporates demand uncertainty. Chance-constraints are used to ensure the number of demands that are unable to be processed on the day they arrive at a processing facility is less than a user-defined threshold. We also develop and compare the performance of multiple cutting-plane algorithms used to solve the resulting stochastic formulation. Finally, we present conclusions and directions of future research in Chapter 6.

[1] The word "inventory" in the Inventory Modulated Capacitated Location Problem is synonymous with the term "backlog" that is used throughout this dissertation. Since publishing Maass et al. (2016) in which we introduced the IMCLP using the term "inventory," we have realized that using the term "backlog" in place of "inventory" promotes a clearer description of our model. Although we use the term "backlog" throughout this dissertation, we continue to refer to the model described in Maass et al. (2016) as the IMCLP.

CHAPTER 2: Literature Review

2.1 Taxonomy of Location Models

While there are numerous ways to categorize the broad spectrum of location models, Figure 1 displays a categorization, similar to that in Daskin (2008), based on the space in which the problems are modeled. *Analytic models* assume demand is distributed over a service area and facilities can be located anywhere within the service area. *Continuous models* assume that demand occurs only at discrete sites but facilities can be located anywhere within the service area. *Network models* assume that demands arise and facilities can be located only on a predetermined network consisting of nodes and arcs. In most network models, demands occur at the nodes, and facilities can be located anywhere within the network. *Discrete models* assume that demands arise on a set of nodes, and facilities are restricted to a finite set of candidate locations.

The subcategory of discrete location models can be further divided into three areas: Covering, Median, and Other. *Covering models* assume that customers are adequately served if they are within a certain distance or time of a facility location. As such, covering models are often applied to locating emergency services (such as ambulances and fire stations). The *set covering*, *max covering*, and *p-center* problems are examples of covering models. *Median models* minimize the demand-weighted average distance between a customer and a facility to which it is assigned, and are typically used in distribution planning contexts. The *p-median* and *fixed charge* problems are two types of median models. In both of these models, the facilities can

have infinite capacity (i.e., an *uncapacitated model*) or a finite capacity (i.e., a *capacitated model*). Finally, some models cannot be classified into either of these areas. For example, the *p-dispersion model* locates a given number of facilities in a manner that maximizes the minimum distance between any pair of facilities. This model is useful in locating franchise outlets and nuclear weapon silos. Other examples are *undesirable facility models* in which we seek to maximize the distance between a facility and the nearest demand node. Such models are useful when locating hazardous waste dumps, landfills, and nuclear reactors.

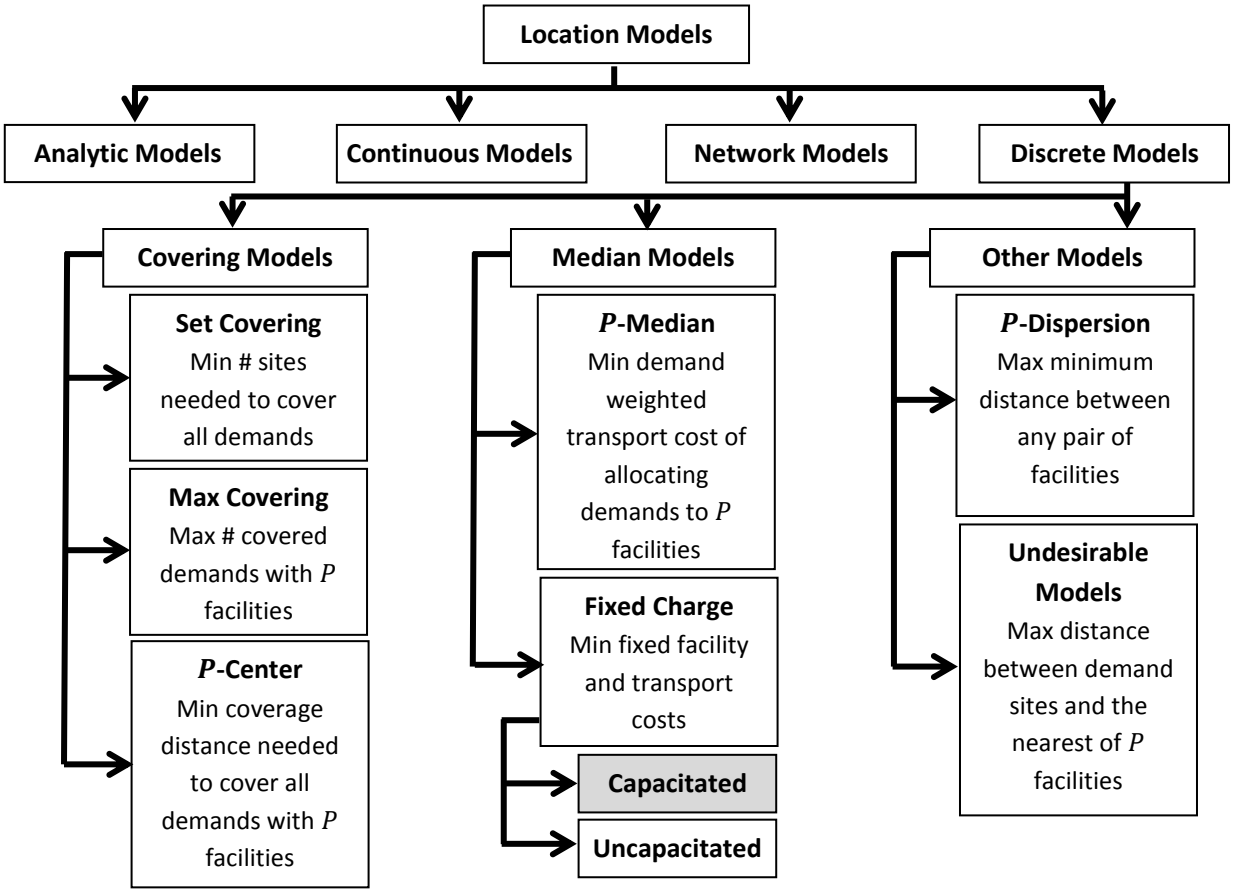

Figure 1: Taxonomy of Location Models

This dissertation focuses on models that extend the traditional capacitated fixed charge models. Specifically, we address the issue that nearly all capacitated facility location models

employ hard capacity constraints on the number of demands that can be assigned to a processing facility (see Daskin et al. (2005) and Verter (2011) for reviews) and thereby fall short in capturing the reality that the capacity is often flexible.

2.2 The Evolution of Capacitated Location Models

The traditional UFLP [Balinski, 1965] trades off the increased cost of locating additional facilities with the decreased cost of transportation as the facilities get closer to the demand sites. To formulate the model we first define the following notation:

Sets and Parameters

I Set of demand sites

J Set of candidate facility locations

h_i Average demand generated per day at demand site $i \in I$

t_{ij} Travel time between demand site $i \in I$ and facility $j \in J$

f_j Daily fixed cost of locating at facility $j \in J$

$a \geq 0$ Cost (in dollars) of transporting one item for one day

Decision Variables

$$X_j = \begin{cases} 1 & \text{If we locate at facility } j \in J \\ 0 & \text{Otherwise} \end{cases}$$

Y_{ij} Fraction of demand from demand site $i \in I$ that is allocated to facility $j \in J$

With this notation the UFLP can be formulated as follows:

Formulation

$$Min_{X,Y} \quad \sum_{j\in J} f_j X_j + a \sum_{j\in J} \sum_{i\in I} h_i t_{ij} Y_{ij} \tag{2.1}$$

$Subject\ to$

$$\sum_{j\in J} Y_{ij} = 1 \qquad \forall i \in I \tag{2.2}$$

$$Y_{ij} \leq X_j \qquad \forall i \in I; j \in J \tag{2.3}$$

$$X_j \in \{0,1\} \qquad \forall j \in J \tag{2.4}$$

$$Y_{ij} \geq 0 \qquad \forall i \in I; j \in J \tag{2.5}$$

The objective function (2.1) minimizes the sum of the fixed facility and transportation costs. Constraints (2.2) ensure that all of the demand from $i \in I$ is allocated to a processing facility. Constraints (2.3) state that demands can only be assigned to located facilities. Finally, constraints (2.4) and (2.5) are standard binary and non-negativity constraints.

It is worth noting that, given a set of located candidate facilities, an optimal solution is one in which all demand sites are assigned to the located facility that can be reached in the shortest amount of travel time. As such, the Y_{ij} variables are naturally binary-valued in an optimal solution (or can be rounded to become binary if multiple t_{ij} parameters have equal values for some $i \in I$).

However, the reality that facilities have capacities has led to the development of capacitated versions of the fixed charge location model. In a straightforward extension of the UFLP, the CFLP adds a single class of constraints composed of exogenous values that are considered for the maximum demand that can be supplied from each potential facility. The capacity constraints are of the form:

$$\sum_{i\in I} h_i Y_{ij} \leq k_j X_j \qquad \forall j \in J \tag{2.6}$$

where k_j represents the daily processing capacity of facility $j \in J$ if a facility is located at $j \in J$. We will henceforth refer to them as the traditional capacity constraints. Constraints (2.6) require that the total demand that is assigned to a facility be less than the capacity of that facility if we choose to locate at the candidate processing facility and 0 otherwise. The addition of these capacity constraints eliminates the closest assignment property that was present in the UFLP. As such, the Y_{ij} variables will not automatically take binary-values. However, we can show that at most $\sum_{j \in J} X_j - 1$ demand sites will be fractionally assigned [Daskin and Jones, 1993].

Both the UFLP and CFLP are known to be NP-hard [Cornuéjols et al., 1990]. As such, various solution algorithms have been developed for the CFLP including: branch-and-bound using linear programming relaxation [Akinc and Khumawala, 1977], Lagrangian relaxation [Nauss, 1978], and a partitioning formulation [Neebe and Rao, 1983]; Benders decomposition [Davis and Ray, 1969; Wentges, 1996]; branch-and-price [Klose and Görtz, 2007]; cross-decomposition utilizing Benders decomposition and Lagrangian relaxation [Van Roy, 1986]; and dual-based methods [Guignard and Spielberg, 1979]. Additionally, large instances of the CFLP have been solved through a variety of heuristics [Jacobsen, 1983; Delmaire et al., 1999; Ahuja et al., 2004]. Extensive research has been devoted to developing Lagrangian based heuristics for the CFLP, many of which begin by relaxing the assignment constraints [Barceló and Casanova, 1984; Pirkul, 1987; Sridharan, 1993; Holmberg et al., 1999; Rönnqvist et al., 1999]. Additionally, Klincewicz and Luss (1986) relax the capacity constraints while Beasley (1993) and Agar and Salhi (1998) relax both the assignment and capacity constraints. A review of solution techniques for the CFLP can be found in Magnanti and Wong (1990), Daskin (2013), and Verter (2011). Furthermore, despite the fact that hard capacity constraints are problematic,

they continue to appear in facility location models [Saveh-Shemshaki et al., 2012; Laporte et al., 1994; Louveaux and Peeters, 1992; Louveaux, 1986].

In addition to models in which physical items arrive at a capacitated facility for processing, many facility location models consider the context in which physical items must be shipped from a warehouse to a demand site. (For example, many retail companies that sell items online have a warehouse stocked with merchandise that is shipped to the customer upon the placement of an online order.) The former instance assumes that the capacity restrictions are placed on the amount of demand that can arrive at a processing facility, whereas the latter assumes that the warehouse has a capacity on the number of items it can pre-stock or assemble to meet the needs of the demand sites. While the capacity constraints of the models in the latter category are problematic for the same reasons as those of the former, they are widely used in the current literature. For example, Balcik and Beamon (2008) use a stochastic maximal covering model with demand uncertainty to determine the location of distribution centers that store relief supplies prior to a disaster. Capacity constraints limit the amount of relief supplies that can be pre-positioned. Additionally, Melo et al. (2006) consider the relocation of capacity when designing a supply chain network. While capacity is flexible over the long-term in the model presented by Melo et al. (2006), hard capacity constraints are enforced on a day-to-day basis.

Hard capacity constraints also appear in numerous operations research problems beyond facility location models. For example, the vehicle routing problem and its variants employ hard capacity constraints on vehicle capacities [Laporte, 1992; Toth and Vigo, 2002; Golden et al., 2008]. Additionally, the classic knapsack problem - which serves as a basis for many other operations research models including scheduling, cutting stock, and portfolio optimization

problems - utilizes hard capacity constraints on the total volume or weight of the selected items [Kellerer et al., 2010].

2.3 Methods of Mitigating Hard Capacity Constraints

Facility managers typically have many operational tools to extend capacity or to allow the facility to accept demands in excess of the capacity constraint for short periods of time. For example, a facility manager may utilize overtime or temporary workers to allow the facility to process more than the specified capacity limit. Additionally, many processing facilities use backlogging or outsourcing as a buffer against inadequate capacity.

The majority of the current facility location literature that incorporates capacity flexibility focuses on either (1) long-term flexibility or (2) penalizing the objective function without incorporating the effect that one day's unprocessed demand has on the following day. Literature that does incorporate backlogged demand focus on a supply chain management context with outgoing demands rather than incoming demands. In this chapter we will describe the differences between our models and the current literature.

2.3.1 Long-Term Capacity Flexibility

A variety of authors have acknowledged the need for capacity flexibility. However, the current literature in this area focuses on incorporating long-term capacity flexibility, rather than daily operational flexibility as we do in this dissertation. Instead of setting capacity as an exogenous input, several researchers [Luss, 1982; Klincewicz et al., 1988; Verter and Dincer, 1995] have modeled capacity as a decision variable, thereby incorporating capacity sizing decisions into the objective function. Aghezzaf (2005), Melo et al. (2006), Fleischmann et al.

(2006), Hugo and Pistikopoulos (2005), and Georgiadis and Athanasiou (2013) also consider capacity expansion decisions. While these formulations allow the capacity constraint to be flexible over a long-term planning horizon by updating the capacity level at specific time epochs, the resulting capacity levels remain inflexible in the short term. That is, between epochs, the capacity is still a hard capacity constraint. The model formulations we present relax the hard capacity constraints for both long and short term planning horizons by utilizing backlog as a buffer against inadequate capacity.

2.3.2 Penalizing Excess Demand

Many of the facility location models that recognize stochastic demands may result in demand levels that exceed the facility capacity penalize the objective function without incorporating the effect that one day's unprocessed demand has on the following day. The following papers incorporate penalties through lost profit and outsourcing costs, however, none of them consider backlogged demand.

Louveaux (1986) formulates a capacitated facility location problem with stochastic demand, production cost, and selling prices in which the goal is to maximize the expected profit by determining the optimal facility locations, facility capacities, and allocation of customers to facilities. While the model assumes that not all demand needs to be satisfied, the author does not incorporate this flexibility into the day-to-day operations of the facility. Instead, the model uses aggregated demand parameters to decide which proportion of demand at each demand site will be satisfied a priori. A penalty for demand that is not satisfied is assessed in the objective function by accounting for the profit that can be obtained by meeting demands. This method implicitly assumes that the unmet demand penalty only consists of lost profit and does not

consider any costs associated with loss of goodwill, holding backlog, etc. Louveaux and Peeters (1992) solve Louveaux's (1986) model using a dual-based heuristic while Laporte et al. (1994) develop an optimal L-shaped solution algorithm.

Albareda-Sambola et al. (2011) address the problem of assigning customers with Bernoulli distributed demand to facilities a priori. Each facility has a specified capacity on the number of customers it can serve at any given time. However, since it is uncertain whether a particular customer will place a service request (i.e., demand) during a time period, the model allows more customers to be assigned to a facility than the facility has capacity to serve at any one given time. If a facility cannot satisfy all of the service requests, it outsources some of the customers and incurs an associated cost. Thus, demands are always assumed to be satisfied in the period in which they occur, unlike the models we present in this dissertation. Additionally, Albareda-Sambola et al. (2011) does not consider daily demands or possible demand correlations.

Rather than determining the location decisions a priori, Hinojosa et al. (2014) present a two-stage model in which they consider location decisions after the uncertainty is realized. Instead, shipment distribution channels to capacitated facilities must be determined prior to realizing the demand. Once the demand (as represented by an aggregated parameter) is realized, a transportation and "location" plan is determined, where the location decisions correspond to the set-up or activation of resources necessary to process the shipment. Since distribution channels are determined a priori and facilities are capacitated, the system may have insufficient capacity to handle the realized demand. In this case, a penalty is incurred in the objective function. However, Hinojosa et al. (2014) assume any such demand is lost whereas we suppose it is

backlogged and, therefore, affects the amount of demand available for processing on the following day.

2.3.3 Backlogged Demand

Although the aforementioned facility location literature does not model unsatisfied demand as demand that is carried over into the following period, the multi-period variable-demand supply chain management (MPVDSCM) literature does capture this phenomenon through backlogged demand (see, e.g., Varthanan et al. (2012), Fahimnia et al. (2012), Torabi and Moghaddam (2012), Peidro et al. (2010)). However, there are two important distinctions between our work and the MPVDSCM literature. First, the product flow directions are reversed. In supply chain management, demand (or raw material components of demand) physically flow from the processing facility (or supplier/warehouse) to the customer, whereas in this dissertation, demand flow from the customer (e.g., clinics/hospitals) to the processing facility (e.g., blood testing lab). Thus, we account for demand that was generated on different days arriving at a processing facility on the same day due to different transportation times. For example, in our model, a processing facility may simultaneously have arriving demand that was generated one day ago at demand site A but two days ago at demand site B. Second, the MPVDSCM literature incorporates foreknowledge about the demand in each period and utilizes period-based production and distribution variables. This allows models in the MPVDSCM literature to build inventory in period t in anticipation of demand that needs to be satisfied in subsequent periods (e.g., period $t+1$). While this is not problematic for models within the supply chain management context, short-term demand allocation decisions in the facility location context should not be made based on anticipating future demand patterns and future allocation decisions. We refer to

such decisions as anticipatory decisions and discuss them further in the following section. Furthermore, our model circumvents anticipatory issues since our location and cyclic allocation decisions do not change over the time horizon.

2.4 Daily Demands and Cyclic Allocations

Most capacitated facility location models consider deterministic demands at an aggregate level by using parameter estimation techniques [Baron et al., 2008]. In contrast to the prevailing method, we consider disaggregated demands at a daily level. To the best of our knowledge, Saveh-Shemshaki et al. (2012) is the only other work that explicitly incorporates demand data directly into an extension of the traditional capacitated facility location model. However, there are some stark differences between our model and the model in Saveh-Shemshaki et al. (2012) in terms of how the daily demands influence the allocation policy, which we describe below.

While many location models enforce single-sourcing constraints, there exist extensions of the CFLP that allow demand to be allocated to multiple processing facilities. For example, it is common in the facility reliability literature to assign a percentage of demand at a site to multiple processing facilities as a means of mitigating demand fulfillment disruption if a subset of the processing facilities fail (due to power outages, stockouts, natural disasters, etc.) or to assign demands to a primary processing facility under normal working conditions but denote other processing facilities as "backup" facilities in the event of a disruption at the primary facility (see, e.g., Snyder et al. (2006)).

In models that allow demand sites to be multi-sourced, the assignments are typically either static (the same for each day) [Akinc and Khumawala, 1977; Nauss, 1978; Daskin and Jones, 1993; Ozsen et al., 2009] or dynamic (evolving as time passes) [Saveh-Shemshaki et al.,

2012; Scott, 1970; Wesolowsky and Truscott, 1975; Daskin et al., 1992; Drezner, 1995; Torres-Soto and Üster, 2011; Jena et al., 2013]. We consider two types of allocations in this dissertation. The first model we propose enforces single-sourcing constraints, while the allocation policy of the second model allows for multi-sourcing. However, the particular multi-sourcing framework we propose is different from other studies in that it is neither static nor dynamic. Instead, the allocations can vary from day-to-day in a cyclic manner. In that sense, the allocations can be considered cyclic. For example, if we let a week represent the cycle time frame, then a demand site may ship its demands to one facility *every* Sunday through Thursday and to another facility *every* Friday and Saturday. As we will show in Chapter 4, this assignment flexibility allows for additional cost savings when compared to a single allocation approach, while still providing the demand site manager with a consistent allocation policy that is easy to implement.

As previously mentioned, Saveh-Shemshaki et al. (2012) explicitly incorporates historical data directly into an extension of the CFLP and allows demand sites to be allocated to a different processing facility each day of the planning horizon. However, they assume that managers at a demand site will know the exact quantities of demand generated at every other demand site for the entirety of the (future) planning horizon and that they can determine the daily allocation decisions using this complete foreknowledge.

For example, consider a time horizon of two weeks. Suppose a facility manager can decide to send demands that are generated at site A on Mondays to processing facility X_1 or X_2, and that regardless of where the demands are sent, they will arrive for processing on Friday of the same week. The daily allocation decisions employed by Saveh-Shemshaki et al. (2012) assume that the decision of where to send the demand generated at site A can be made based on

foreknowledge of future demands and allocation decisions, as illustrated in the following example:

Week One: The decision of where to send Monday demands *in the first* week can account for the foreknowledge that on Wednesday *of the first week* demand site B will send a large amount of demand to facility X_2 that will arrive on Friday *of the first week*. Thus, if both site A and B send their demands to X_2, the incoming demand will exceed the processing capacity at X_2 and some demand will not be processed on Friday *of the first week*. However, if site A sends its Monday's demands to site X_1 it can be guaranteed that all of the demand will be processed on Friday *of the first week*. Furthermore, assigning demands from both A and B to X_2 is not feasible in Saveh-Shemshaki et al. (2012) due to the hard capacity constraints.

Week Two: Suppose that in the second week, site B still sends its demands to facility X_2 on Wednesday and that the demand shipment will arrive for processing on Friday. Since there is complete foreknowledge, demand site A knows the exact number of items facility B will send to X_2 on Wednesday and calculates that the total amount of demand that will arrive at facility X_2 on Friday will not exceed the processing capacity of X_2, even if A also sends its demands to facility X_2 on Monday for arrival on Friday. Thus, A can send its demands to X_2 in the second week.

Notice that in this example, demand site A chooses a different facility to which it will send its demand each week based on *future* information regarding other demand sites. Allowing

the allocation decisions to be made on a daily basis in this manner is problematic because, in reality, managers at a demand site will typically not know the exact daily allocation decisions (processing facility selected and quantities of demand shipped) of every other demand site for the entirety of the planning horizon.

In the second model that we develop in this dissertation, we also consider demands at a daily level and allow allocations to vary from day-to-day; however, we do not incorporate the same problematic assumption as does Saveh-Shemshaki et al. (2012) since we require that the demand allocation *policy* does not change from one time frame (e.g., a week) to another. For example, our model with a weekly cyclic allocation time frame results in a manager making the same allocation decision every Monday. That decision may be based on the historical demand patterns and on the availability of different processing facilities over the course of the week. Furthermore, the allocation decision for Monday may differ from the decision the manager would make every Tuesday or Saturday. This allows us to allocate demands in a manner that considers the effect of the day-of-the-week capacity variations and aggregate daily demand fluctuations without supposing demands are anticipatory in nature.

As far as we are aware, no facility location model currently implements such a cyclic allocation approach. The most relevant cyclic allocation literature can be found in the application areas of emergency medical vehicle relocation and nurse staffing when considering time-dependent parameters. For example, ambulances are typically repositioned to various base locations throughout the day to account for time-of-day dependent demands or travel times [Repede and Bernardo, 1994; Gendreau et al., 2001; Maxwell et al., 2010; Maleki et al., 2014]. Since the ambulances wait at their base location until they are dispatched to respond to an emergency, redeployment models can be used to aid managers in making such daily or hourly

plans to better respond to predictable demand fluctuations. Rajapopalan et al. (2008) develop a multi-period model to determine the minimum number of ambulances needed, as well as their base locations, in each time period. The model accounts for significant changes in demand patterns that occur among the time periods and ensures that coverage requirements are met. For example, such a model may determine that every day from 9:00 am – 5:00 pm a particular ambulance should be located in the city's business center, but it should be located in a residential area from 5:00 pm – 12:00 am. Repositioning in this cyclic manner allows the ambulance to be near the business center during business hours (when the residential areas are likely sparsely populated) and near residential areas during evening hours (when the business center is sparsely populated).

The redeployment model of Rajapopalan et al. (2008) however does not consider the frequency or number of repositioned ambulances; frequent repositioning may improve coverage at the cost of creating a positioning schedule that is more difficult for the deployment planners to manage. Schmid and Doerner (2010) penalize vehicle relocation in their multi-period model that seeks to determine optimal ambulance locations at various time periods such that the resulting coverage can be adequately maintained throughout the planning horizon. Rather than considering time varying demand, Schmid and Doerner (2010) incorporate time-dependent travel times. Their results indicate that neglecting time-dependent variations can lead to serious overestimation of the resulting coverage.

There also exists literature on nurse staffing that uses time-of-day or day-of-week demand data to develop shift based staffing levels. This is similar to the cyclic allocations proposed in this paper in the sense that the number of nurses assigned to work in an internal medicine, endoscopy, and intensive care unit may be 4, 4, and 7, respectively, each Monday

while 5, 4, and 6 nurses are assigned to work in the same units each Wednesday. Trivedi (1981) uses shift level demand data to determine the weekday and weekend staffing levels for full time unit nurses. Additionally, varying the input data and solving the single period newsvendor model developed in Davis et al. (2014) multiple times produces a policy based approach to determining staffing levels for different shifts.

2.5 Chapter Summary

In this chapter, we presented an overview of facility location models and identified how the models in this dissertation fit into the larger taxonomy of location models. We also reviewed the evolution of capacitated facility location models and presented the current literature related to capacity flexibility.

CHAPTER 3: Inventory Modulated Capacitated Location Problem

3.1 Motivation

As outlined in Chapter 1, the current literature on capacitated facility location models (1) disregards the reality that a facility's capacity is a function of many operational decisions, (2) does not allow facilities to accept demands in excess of their capacity constraints, and (3) utilizes aggregated demand parameters. In this chapter, we present a model that addresses these three issues by incorporating daily demands and allowing demands that arrive at a processing facility in excess of the capacity to be backlogged and processed at a later date, while incurring a penalty for the delay in processing.[1]

3.2 Model Formulation

In this chapter, we address issues with traditional capacity constraints by presenting a data-driven model in which all input parameters are deterministic and the hard capacity constraints are mitigated by utilizing a backlog of demands to be processed. We refer to this model as the IMCLP.

Although this model considers demands at a daily level, it requires the demand allocations to be fixed throughout the time horizon. Thus, both the IMCLP and the CFLP are static models in the sense that the location and demand allocation decisions are static and time-independent decisions. The CFLP captures only the average demand at each demand site while

[1] A majority of the content of this chapter has been published in Maass, et al. (2016).

the IMCLP accounts for variations around the mean demand. Therefore, the IMCLP needs to model the day-to-day processing decisions and the implied backlog accumulation that result from the confluence of daily demands and processing at the facility.

In addition to the notation previously defined, we introduce the following inputs, sets, and decision variables:

Inputs and Sets

D Set of days

h_{id} Demand that is generated at demand site $i \in I$ on day $d \in D$

$t^* = max_{i \in I, j \in J} t_{ij}$

v_j Initial backlog (at the beginning of day $t^* + 1$) at processing facility $j \in J$ if j is located

$b_j \geq 0$ Cost of holding one item in backlog for one day at facility $j \in J$

Intuitively, t^* is the longest travel time between any demand-facility pair. The days in $\{1, \dots, t^*\}$ constitute the warm-up period, and we start collecting cost metrics on day $t^* + 1$ to mitigate the model warm-up effect caused by demands not yet having arrived at a processing facility at the beginning of the model. Furthermore, we refer to $|D| - t^*$ as the number of days in the planning horizon. To ensure that the CLFP and IMCLP consider the same time horizon, we change the CFLP objective function (2.1) to

$$Min \quad (|D| - t^*)\left(\sum_{j \in J} f_j X_j + a \sum_{j \in J} \sum_{i \in I} h_i t_{ij} Y_{ij}\right). \tag{3.1}$$

Additionally, we enforce single-sourcing constraints on the CFLP so that the allocation decisions are consistent between the two models. From this point on, anytime we refer to the CFLP we are referring to (3.1), (2.2) - (2.4), (2.6) and $Y_{ij} \in \{0,1\}, \forall i \in I; j \in J$.

It is also worth mentioning that the demand parameters, h_{id}, used in the IMCLP specify the particular demand that originates at demand site i on day d. This is in contrast to the

parameters h_i in the CFLP, which represent the average daily demand generated at demand site i. Specifically, for each $i \in I$, $h_i = \frac{\sum_{d \in D} h_{id}}{|D|}$.

Finally, for ease of notation, whenever $b_{j_1} = b_{j_2}\ \forall j_1, j_2 \in J$, we drop the subscripts and let b represent the generic unit backlog cost.

Decision Variables

V_{jd} Auxiliary decision variable representing the backlog level at facility $j \in J$ at the beginning of day $d \in \{t^* + 1, \dots, |D| + 1\}$

W_{jd} Auxiliary decision variable representing the number of items that are processed at facility $j \in J$ on day $d \in \{t^* + 1, \dots, |D|\}$

The model determines the optimal long term location and allocation decisions and uses the auxiliary daily processing and backlog variables to track metrics associated with the day to day performance of the system. Thus, the location and allocation variables are static in the sense that the decision maker will decide once where to build the processing facilities and how to allocate the demand to these facilities. She will then track the daily backlog and processing levels that are a result of her location and allocation decisions. As such, the IMCLP remains a static model and can be compared to the static location and allocation decisions of the CFLP.

Formulation

With this notation, we formulate the IMCLP as follows:

$$Min_{\boldsymbol{X,Y,V,W}}\ (|D| - t^*) \sum_{j \in J} f_j X_j + a \sum_{d=t^*+1}^{|D|} \sum_{j \in J} \sum_{i \in I} h_{id} t_{ij} Y_{ij} + \sum_{d=t^*+2}^{|D|+1} \sum_{j \in J} b_j V_{jd} \qquad (3.2)$$

$Subject\ to$

$$\sum_{j \in J} Y_{ij} = 1 \qquad \forall i \in I \qquad (2.2)$$

$$Y_{ij} \leq X_j \qquad \forall i \in I; j \in J \qquad (2.3)$$

$$V_{j,d+1} - V_{jd} - \sum_{i \in I} h_{i,d-t_{ij}} Y_{ij} + W_{jd} = 0 \quad \forall j \in J; d \in \{t^* + 1, \dots, |D|\} \tag{3.3}$$

$$W_{jd} \leq k_j X_j \quad \forall j \in J; d \in \{t^* + 1, \dots, |D|\} \tag{3.4}$$

$$V_{j,t^*+1} = v_j X_j \quad \forall j \in J \tag{3.5}$$

$$X_j \in \{0,1\} \quad \forall j \in J \tag{2.4}$$

$$Y_{ij} \in \{0,1\} \quad \forall i \in I; j \in J \tag{3.6}$$

$$V_{jd} \geq 0 \quad \forall j \in J; d \in \{t^* + 1, \dots, |D| + 1\} \tag{3.7}$$

$$W_{jd} \geq 0 \quad \forall j \in J; d \in \{t^* + 1, \dots, |D|\} \tag{3.8}$$

The IMCLP's objective function (3.2) minimizes the costs associated with facility location, transportation, and backlog, while the CFLP only considers the costs associated with facility location and transportation. Recall that since we incorporated $|D| - t^*$ into the CFLP objective function (3.1), the two models calculate costs over the same time horizon.

Constraints (3.3) are flow balance constraints that update the daily backlog variables. They state that the backlog at the beginning of day $d + 1$ at processing facility $j \in J$ is equal to the amount of demand that was left unprocessed at $j \in J$ at the end of day d. This consists of the amount of demand available for processing at $j \in J$ at the beginning of day d (i.e., the amount in backlog at the beginning of day d plus the amount that arrived on day d) minus the amount that was processed on day d. Note that we assume all shipments arrive at processing facilities at the beginning of the day, before any processing occurs.

Constraints (3.4) are the capacity constraints. In contrast to the capacity constraints in traditional capacitated models, which state that the *total demand arriving* at a facility on each day must be less than the daily processing capacity at that facility, the capacity constraints of (3.4) state that, on each day, we cannot *process* more at a facility than there is capacity for processing at that facility. This form allows for facility-dependent capacity levels. Furthermore,

constraints (3.3) also ensure that we cannot process more at a facility on any given day than the sum of the amount that was left unprocessed from the previous day plus what arrives at the beginning of the day from all demand sites. Constraints (3.4) can also be extended to incorporate short-term capacity expansion decisions, such as the use of overtime (i.e., an individual working in excess of 40 hours per week). We discuss this briefly as an area of future research in Section 6.5.1.

Constraints (3.5) state that, for all $j \in J$, if facility j is located, it will have an initial backlog of v_j. Otherwise, the initial backlog level at facility j is zero. The purpose of providing the initial backlog levels v_j is to initialize the flow balance equations. Constraints (2.4) and (3.6) - (3.8) are standard binary and non-negativity constraints.

Imbedded in the IMCLP model is the assumption that demands are never processed on the day they originate (and, furthermore, that an incoming demand shipment never begins processing part way through the day). This assumption is necessary for the model to accurately calculate the amount of demand processed each day. For example, suppose that a processing facility $j \in J$ has no demand available for processing at the beginning of day $d \in D$. If the first shipment of demand that arrives for processing at that facility arrives ten minutes before closing, and the shipment contains k_j items, then the IMCLP model would believe that all k_j items can be processed on day d. However, we know that this is not acceptable. The demands arrive with only ten minutes of processing time left for the day, and therefore we would only be able to process a small fraction, if any, of the arriving demands.

While the IMCLP limits the number of demands that can't be processed on the day they arrive by assessing a penalty for backlogged demand, the formulation does not specify how long an item can be backlogged or the maximum number of demands that can be backlogged at any

given time. If desired, these types of restrictions can be added to the IMCLP formulation by adding constraints of the form:

$$V_{jd} \leq \bar{d}_j k_j \qquad \forall j \in J; d \in \{t^* + 1, \dots, |D|\} \qquad (3.9)$$

where $\bar{d}_j$ represents the maximum number of days an item at facility $j \in J$ is allowed to be held in backlog. More generally, constraints (3.9) limit the *amount* of demand that can be held as backlog, and as such $\bar{d}_j$ can take any non-negative value. We note that if $\bar{d}_j = 0$ then the IMCLP is restricted to process all items on the day they arrive and becomes a model with hard capacity constraints.

Throughout the remainder of this dissertation, we do not consider constraints (3.9) as part of the IMCLP as the penalty cost associated with holding items in backlog, b, was chosen to ensure that the backlog level did not grow over time. For example, the results for the problem instances used in this chapter never have more demand held in backlog than approximately 20% of the daily capacity. This indicates that if the demands were not processed on the first day they joined the processing queue, then they would be processed the following day. We leave analyzing the effect of incorporating constraints (3.9) for varying values of b as an area of future work.

3.3 Model Properties

In this section, we discuss some of the properties of the IMCLP. In particular, we prove that the IMCLP is NP-Hard and discuss the location, allocation, and cost differences that can arise between the IMCLP and CFLP solutions. We also show that when the allocation variables, Y_{ij}, are allowed to take on continuous values between zero and one, inclusive, the maximum number of demand sites that may be multi-sourced is different from the maximum for the CFLP.

3.3.1 NP-Hardness

The IMCLP is NP-hard as we can reduce it to the NP-hard UFLP when the parameters $b_j = 0 \ \ \forall j \in J$ and $h_{id} = h_i \ \forall i \in I, d \in \{t^* + 1, \dots, |D|\}$ [Garey and Johnson, 1979].

Theorem: The IMCLP is NP-hard since it reduces to the UFLP when $b = 0 \ \forall j \in J$ and $h_{id} = h_i$ $\forall i \in I, d \in \{t^* + 1, \dots, |D|\}$.

Proof: When $b = 0 \ \forall j \in J$ and $h_{id} = h_i \ \forall i \in I, d \in \{t^* + 1, \dots, |D|\}$, the IMCLP can be written as a UFLP with additional constraints associated with the processing ($W_{jd,} \ \forall j \in J, d \in \{t^* + 1, \dots, |D|\}$) and backlog ($V_{jd}, \ \forall j \in J, d \in \{t^* + 1, \dots, |D| + 1\}$ variables, i.e. (3.4), (3.5), (3.7), (3.8), and (3.11):

$$Min_{\boldsymbol{X,Y,V,W}} \quad (|D| - t^*)(\textstyle\sum_{j\in J} f_j X_j + a \sum_{j\in J} \sum_{i\in I} h_i t_{ij} Y_{ij}) \tag{3.10}$$

$Subject\ to$

$$\textstyle\sum_{j\in J} Y_{ij} = 1 \qquad \forall i \in I \tag{2.2}$$

$$Y_{ij} \leq X_j \qquad \forall i \in I; j \in J \tag{2.3}$$

$$X_j \in \{0,1\} \qquad \forall j \in J \tag{2.4}$$

$$Y_{ij} \in \{0,1\} \qquad \forall i \in I; j \in J \tag{3.6}$$

$$V_{j,d+1} - V_{jd} - \textstyle\sum_{i\in I} h_i Y_{ij} + W_{jd} = 0 \qquad \forall j \in J; d \in \{t^* + 1, \dots |D|\} \tag{3.11}$$

$$W_{jd} \leq k_j X_j \qquad \forall j \in J; d \in \{t^* + 1, \dots, |D|\} \tag{3.4}$$

$$V_{j,t^*+1} = v_j X_j \qquad \forall j \in J \tag{3.5}$$

$$V_{jd} \geq 0 \qquad \forall j \in J; d \in \{t^* + 1, \dots, |D| + 1\} \tag{3.7}$$

$$W_{jd} \geq 0 \qquad \forall j \in J; d \in \{t^* + 1, \dots, |D|\} \tag{3.8}$$

Note that the objective function (3.10) is the same as the objective function of the corresponding UFLP model. Let z^* and z^{**} be the optimal objective values of the IMCLP and UFLP respectively. It follows that $z^* \geq z^{**}$.

Now, note that if we find an optimal solution $(\boldsymbol{X}^*, \boldsymbol{Y}^*)$ with an objective function value of z^{**} to the UFLP ((2.2) - (2.4), (3.6), and (3.10)) and can construct an expanded feasible solution $(\boldsymbol{X}^*, \boldsymbol{Y}^*, \boldsymbol{W}^*, \boldsymbol{V}^*)$ to the IMCLP (which has a corresponding objective function value of z^{**}), then we also have $z^{**} \geq z^*$. Given the above result, this leads to $z^* = z^{**}$. Thus, $(\boldsymbol{X}^*, \boldsymbol{Y}^*, \boldsymbol{W}^*, \boldsymbol{V}^*)$ is an optimal solution to the IMCLP.

Indeed, for any optimal solution $(\boldsymbol{X}^*, \boldsymbol{Y}^*)$ to the UFLP, we are able to construct such an expanded solution $(\boldsymbol{X}^*, \boldsymbol{Y}^*, \boldsymbol{W}^*, \boldsymbol{V}^*)$ to the IMCLP by setting $\boldsymbol{W}^* = \boldsymbol{0}$ (which satisfies (3.4) for any value of $\boldsymbol{X}^* \geq \boldsymbol{0}$) and then setting the value of $\boldsymbol{V}^*$ according to (3.5) and (3.11) with $\boldsymbol{W}^* = \boldsymbol{0}$. Therefore, optimizing the specially designed instance of the IMCLP is equivalent to optimizing the UFLP, and because the UFLP is NP-hard in general, the IMCLP is also NP-hard. □

3.3.2 Benefits of the IMCLP

The following example illustrates the benefits of using the IMCLP formulation rather than the CFLP formulation, which employs hard capacity constraints and average demand values rather than daily realizations of demand. In particular, the example demonstrates that the IMCLP may not locate facilities at the same locations that the CFLP chooses. In fact, the two models may choose to locate a different number of facilities in the optimal solution. It follows that the optimal allocation decisions are different as well.

Suppose there are three demand sites, denoted by A, B, and C, and each of these sites is also a candidate facility location. Suppose further that the daily fixed cost of locating a facility is

$100 and each located facility can process 100 items per day. Demand sites A, B, and C have a daily average demand of 50, 45, and 40 units, respectively, which will be used as the average demand parameters in the CFLP and will be used to generate the daily demands in the IMCLP. The corresponding daily demand data and travel times are given in Table 1 and Table 2 below. The weight on the transportation cost is $a = 1$, the weight on the backlog cost is $b = 2$, and the initial backlog levels are set at 0. We consider the problem for a time horizon of ten days and note that the warm-up period consists of days $d \in \{1,2,3\}$ since $t^* = 3$. (Shading within the tables in this chapter indicates the warm up period.)

Table 1: Example demand

Demand Site	Day	Daily Demand Generated													Avg. Daily Demand
		1	2	3	4	5	6	7	8	9	10	11	12	13	
A		40	60	40	60	40	60	40	60	40	60	40	60	40	50
B		50	40	50	40	50	40	50	40	50	40	50	40	50	45
C		80	0	80	0	80	0	80	0	80	0	80	0	80	40
														Total Avg. Daily Demand:	135

Table 2: Example travel time (in days)

Demand Site	Candidate Facility		
	A	B	C
A	1	2	3
B	2	1	2
C	3	2	1

Using the data of Table 1 and Table 2 the optimal CFLP solution is to locate facilities at sites A and B, and to assign demands originating at A to facility A while assigning demands originating at B and C to facility B. However, the optimal IMCLP solution is to locate facilities at sites A and C, and to assign demands originating at A and B to facility A while assigning demands originating at C to facility C. The location and allocation decisions for these two solutions are depicted in Figure 2.

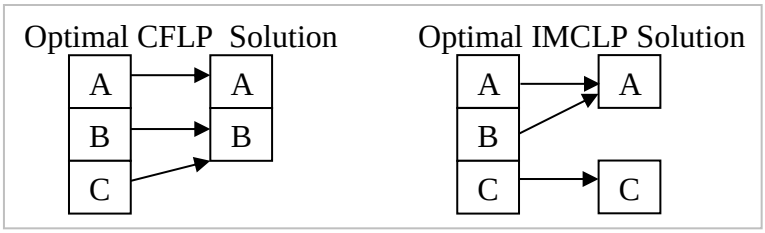

Figure 2: Optimal CFLP and IMCLP solutions for the example

If the CFLP is used to determine the optimal location and allocation decisions for a processing facility location problem where, in reality, a processing facility is able to hold excess demand in backlog and process it the following day, the total cost calculated by the CFLP will not reflect the actual cost of the system. This is because, in reality, the processing facility incurs penalty costs associated with holding excess demand in backlog and processing it the following day; the CFLP does not consider such costs. However, the IMCLP does. Thus, we compare the optimal IMCLP cost to the cost of using the optimal CFLP locations and allocations when the daily demand is realized and excess demand can be held in backlog to be processed the following day. We do this by fixing the optimal CFLP locations and allocations in the IMCLP model.

Both the optimal CFLP and IMCLP solutions choose to locate two facilities. As a result, the facility location cost term in the objective function for either solution is $2,000 ($200/day over a ten day planning horizon). Now consider the cost difference as given by the transportation and backlog terms for the optimal IMCLP solution and the IMCLP solution under the CFLP location and allocation decisions.

In the optimal IMCLP solution, demands that originate at A arrive for processing at facility A the following day (i.e., Table 2 lists the travel time as one day) and demands that originate at B arrive at facility A in two days. Facility C only processes its own demands, and these demands are available for processing the day after they originate. In the optimal CFLP solution, demands from A still arrive for processing at facility A the following day. However,

demands that originate at B and C arrive for processing at facility B in one and two days, respectively. The total amount of demand arriving each day at each facility under the two solutions is presented in Tables 3 and 4.

Tables 3 and 4 make it evident that there are some days (i.e., days 3, 5, 7, 9, 11, and 13) on which the amount of demand arriving at a facility exceeds the daily processing capacity of 100 items per facility. These excess demands will be reported as the backlog present on the following mornings (i.e., the morning of days 6, 8, 10, 12, and 14). We remind the reader that since days 1-3 fall within the warm-up period, we begin assessing the facility's operations on day four and therefore do not include the excess demand that arrives on day 3. Furthermore, due to the definition of the backlog variables representing the amount of backlog at the beginning of the day, the IMCLP captures the backlogging costs through day $|D| + 1$. Thus, although the last day of the model's time horizon is $|D| = 13$, we report the final backlog level as the about of backlog at the *beginning* of day $|D| + 1 = 14$ and note that this is the same amount of backlog that is present at the *end* of day 13.

Thus, the optimal solution to the IMCLP results in ten units of backlog at facility A at the beginning of days 6, 8, 10, 12, and 14, which results in a total backlog cost of $100 over the ten day planning horizon. However, when the CFLP locations and allocations are fixed in the IMCLP model, facility B has 20 units of backlog at the beginning of days 6, 8, 10, 12, and 14. This results in a total backlog cost of $200 over the ten day planning horizon and is double the backlog cost of the optimal IMCLP solution.

On the other hand, the optimal IMCLP solution has a total transportation cost of $1,800, which is greater than the transportation cost of $1,750 given by the CFLP solution in the IMCLP model. In the optimal IMCLP solution, the cost savings from decreasing the backlogging costs

outweigh the increased transportation costs from locating at facility C instead of B. This results in an optimal solution that is different from, and has a lower total cost than, the optimal CFLP solution once daily demands and backlogging costs are realized. Table 5 summarizes these costs.

Table 3: Daily demand received: optimal IMCLP solution

Processing Facility	Day												
	1	2	3	4	5	6	7	8	9	10	11	12	13
A	0	40	110	80	110	80	110	80	110	80	110	80	110
B	0	0	0	0	0	0	0	0	0	0	0	0	0
C	0	80	0	80	0	80	0	80	0	80	0	80	0

Table 4: Daily demand received: optimal CFLP assignments in IMCLP model

Processing Facility	Day												
	1	2	3	4	5	6	7	8	9	10	11	12	13
A	0	40	60	40	60	40	60	40	60	40	60	40	60
B	0	50	120	50	120	50	120	50	120	50	120	50	120
C	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 5: Cost comparison

	IMCLP	IMCLP with CFLP Locations and Allocations
Fixed Cost	\$2,000	\$2,000
Transportation Cost	\$1,800	\$1,750
Backlog Cost	\$100	\$200
Total Cost	\$3,900	\$3,950

While the total cost difference between the IMCLP and the IMCLP with CFLP locations and allocations is a mere \$50 for the above example, the cost savings from using the IMCLP model largely depends on the value of b. Figure 3 displays the effect of fixing the CFLP location and allocation decisions in the IMCLP as the value of b increases while a remains constant. Recall that the IMCLP objective function consists of the facility location, transportation, and backlogging costs, while the CFLP objective function only considers the first two costs. As a result, increasing b (i.e., the weight on backlog) may affect the optimal IMCLP solution but will not affect the optimal CFLP solution. Thus, we always evaluate the same CFLP solution in the IMCLP, regardless of the value of b. Figure 3 shows that as the value of b increases, the percent increase in cost from using the CFLP solution instead of the optimal

IMCLP solution increases. Hence, as b increases, it becomes more costly to approximate the optimal IMCLP solution with the CFLP solution. As a result, using the IMCLP solution will result in a lower cost than the CFLP solution when the daily demands are realized. Additionally, when $b = 11$, it becomes more costly to hold items in backlog than to locate all of the facilities and let each facility process its own demand. As such, the optimal IMCLP solution changes at $b = 11$. This solution has no backlog and, therefore, remains optimal for all values of $b \geq 11$. Therefore, using the CFLP solution instead of the IMCLP solution results in an arbitrarily more costly solution as the value of b increases.

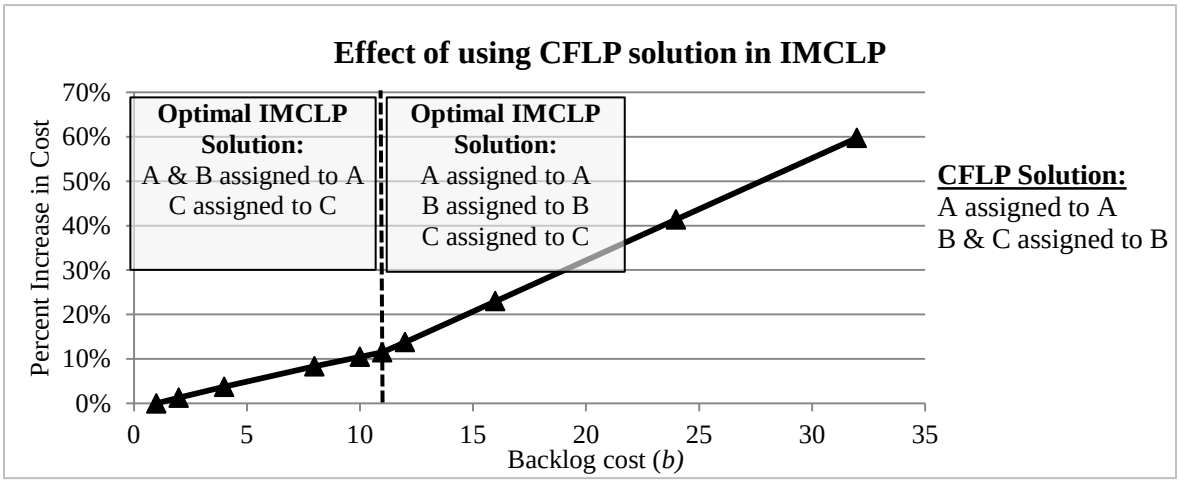

Figure 3: Effect of using the CFLP solution in the IMCLP; $a = 1$

Data-Driven CFLP:

It is worth noting that if the cost of backlog is set sufficiently high, the IMCLP model will essentially become a data-driven CFLP model (see Appendix A for formulation) in which daily demands can be used to determine the optimal model decisions, but the capacity constraints on the amount of *arriving* demand remain enforced. That is, no backlog is allowed when b is sufficiently large; the capacity constraints effectively become $\sum_{i \in I} h_{d-t_{ij}} Y_{ij} \leq k_j X_j \; \forall j \in J;\; d \in \{t^* + 1, \ldots, |D|\}$. As a result, the data-driven CFLP must locate all three facilities in the small

example, and, to minimize cost, demands must be processed at the site at which they originate. This results in a total cost of $4,350 over the 10-day time horizon as shown in Table 6.

Table 6: Cost of optimal data-driven CFLP solution

	Data-Driven CFLP
Fixed Cost	$3,000
Transportation Cost	$1,350
Backlog Cost	$0
Total Cost	$4,350

IMCLP Model Property: The optimal objective function value of the data-driven CFLP is an upper bound on the optimal IMCLP objective function value.

The increase in fixed facility cost results in the data-driven CFLP model being much more costly to implement than either the IMCLP or the IMCLP with CFLP locations and allocations when $b = 2$. In fact, since the data-driven CFLP is a more restrictive version of the IMCLP, it follows that the optimal cost of the data-driven CFLP will always serve as an upper bound on the optimal IMCLP cost. For this example, the $4,350 data-driven CFLP is 11.5% more costly than the optimal $3,900 IMCLP solution when $b = 2$. (We note that the optimal data-driven CFLP solution is precisely the IMCLP solution obtained when $b = 11$, and is indicated by a dashed line in Figure 3). Table 7 displays the amount of demand received at each facility in this situation; we note that there is no day in which over 100 demands arrive for processing.

Table 7: Daily demand received: optimal data-driven CFLP

Processing Facility	Day												
	1	2	3	4	5	6	7	8	9	10	11	12	13
A	0	40	60	40	60	40	60	40	60	40	60	40	60
B	0	50	40	50	40	50	40	50	40	50	40	50	40
C	0	80	0	80	0	80	0	80	0	80	0	80	0

3.3.3 Theoretical Maximum Number of Multi-Soured Demand Nodes

It is commonplace in logistics literature to design networks in which each demand site is served by exactly one processing facility due to the operational complexity involved in allowing the demand to be sourced by multiple processing facilities. However, one advantage of allowing demands sites to be multi-sourced is a reduction in overall logistics costs. Ozsen et al. (2009) present such cost reductions for a two-echelon capacitated facility location model that incorporates safety stock decisions (an environment that differs from the IMCLP). Their results show that the number of located facilities, multi-sourced demand sites, and the number of facilities that serve each demand site increases as the transportation weight increases. Thus, the cost reduction from allowing multi-sourcing increases as the transportation weight increases.

Additionally, while we do not consider disruptions in the network transportation structure in this model, risk diversification by multi-sourcing demand sites is favored in situations in which the network is threatened by disruptions. Mak and Shen (2012) show that supply chain networks that allow dynamic sourcing can be very robust against both disruptions and demand uncertainty. Their results indicate that single sourcing becomes less favorable as the decision maker becomes more risk adverse.

The remainder of this section investigates the theoretical implications of allowing multi-sourcing within the IMCLP. We do this by relaxing the assumption that the allocation variables are binary and instead assume that they may take on continuous values between zero and one

(i.e., $Y_{ij} \geq 0$). In Section 3.4.1 we compare the theoretical implications to empirical results generated from large data instances.

It is well known that when the assignment variables are relaxed to take on continuous values, the number of multi-sourced demand sites in the CFLP is at most the number of processing facilities that are located minus one [Bertsekas, 2003]. Daskin and Jones (1993) note that the number of demand sites is often significantly greater than the number of located processing facilities, and therefore, relatively few demand sites will be multi-sourced. Here we present a simple example to show that these results do not translate to the corresponding relaxed version of the IMCLP. Indeed, there exist situations in which the optimal solution to the relaxed version of the IMCLP requires all of the demand sites to be multi-sourced. The situation below is one such example.

<u>IMCLP Model Property:</u> The optimal solution to the IMCLP with relaxed allocation variables (i.e., $Y_{ij} \geq 0$) may require all of the demand sites to be multi-sourced. Therefore, the number of multi-sourced demand sites may greatly exceed the number of located processing facilities in the optimal solution.

Suppose we have ten demand sites labeled A through J and that demand sites A, B, C, D, and E are also candidate processing facilities. Each demand site generates 25 units of demand on two consecutive days, but does not generate any demand on the remaining eight days of a ten-day demand cycle, as described in Table 8. Let us also suppose that there is a uniform travel time so that it takes one day to transport demand from each demand site to any processing facility. Additionally, assume that the daily fixed cost of each located facility is $100 and each located

facility can process at most ten units per day. Let the transportation weight be set at $a = 1$ and the backlog weight be set to $b = 10$.

Table 8: Multi-sourced example: daily demand generated

Demand	Day																			
Site	1	2	3	4	5	6	7	8	9	10	…	92	93	94	95	96	97	98	99	100	101
A	25	0	0	0	0	0	0	0	0	25	…	0	0	0	0	0	0	0	0	25	25
B	25	25	0	0	0	0	0	0	0	0	…	25	0	0	0	0	0	0	0	0	25
C	0	25	25	0	0	0	0	0	0	0	…	25	25	0	0	0	0	0	0	0	0
D	0	0	25	25	0	0	0	0	0	0	…	0	25	25	0	0	0	0	0	0	0
E	0	0	0	25	25	0	0	0	0	0	…	0	0	25	25	0	0	0	0	0	0
F	0	0	0	0	25	25	0	0	0	0	…	0	0	0	25	25	0	0	0	0	0
G	0	0	0	0	0	25	25	0	0	0	…	0	0	0	0	25	25	0	0	0	0
H	0	0	0	0	0	0	25	25	0	0	…	0	0	0	0	0	25	25	0	0	0
I	0	0	0	0	0	0	0	25	25	0	…	0	0	0	0	0	0	25	25	0	0
J	0	0	0	0	0	0	0	0	25	25	…	0	0	0	0	0	0	0	25	25	0

The optimal solution to the IMCLP in which the assignment variables are relaxed locates all five processing facilities, instructs each of the ten demand sites to allocate 20% of their demand to each of the five located processing facilities (Table 9), and has ten units of demand arriving at each processing facility each day (Table 16). Since each facility can process ten items of demand per day, no items are backlogged. Thus, the total cost over the time horizon is $55,000 and consists of $50,000 in facility location costs and $5,000 in transportation costs.

Table 9: Multi-sourced example 2 solution: demand allocation

Demand Site	Candidate Processing Location				
	A	B	C	D	E
A	0.2	0.2	0.2	0.2	0.2
B	0.2	0.2	0.2	0.2	0.2
C	0.2	0.2	0.2	0.2	0.2
D	0.2	0.2	0.2	0.2	0.2
E	0.2	0.2	0.2	0.2	0.2
F	0.2	0.2	0.2	0.2	0.2
G	0.2	0.2	0.2	0.2	0.2
H	0.2	0.2	0.2	0.2	0.2
I	0.2	0.2	0.2	0.2	0.2
J	0.2	0.2	0.2	0.2	0.2

Table 10: Multi-sourced example 2 solution: total demand arriving at processing facility

Processing Facility	Day							
	1	2	3	4	5	6	…	101
A	0	10	10	10	10	10	…	10
B	0	10	10	10	10	10	…	10
C	0	10	10	10	10	10	…	10
D	0	10	10	10	10	10	…	10
E	0	10	10	10	10	10	…	10

Note that since ten demand sites are multi-sourced and five facilities are located, this example shows that the number of multi-sourced sites exceeds the number of facilities located minus one. Furthermore, we identify that there are no alternative optimal solutions in which fewer demand sites are multi-sourced by adding the following constraints to formulation (2.2) - (2.4) and (3.2) - (3.8), resolving the model, and observing an optimal cost that is higher than \$55,000.[1] We let N_{ij} be a binary variable that takes the value of one if demand site i allocates any of its demand to facility j (and zero otherwise), E_i be a binary variable that takes the value of one if demand site i is multi-sourced (and zero otherwise) and add the constraints:

$$N_{ij} \geq Y_{ij} \qquad \forall i \in I;\ j \in J \qquad (3.12)$$

$$\sum_{j \in J} N_{ij} \leq 1 + (|J| - 1)E_i \qquad \forall i \in I \qquad (3.13)$$

$$\sum_{i \in I} E_i \leq |I| - 1 \qquad \forall i \in I \qquad (3.14)$$

Constraints (3.12) ensure that the variable N_{ij} is one if demand site i allocates any of its demand to facility j). Then, the left hand side of constraint (3.13) counts the number of facilities to which demand site i allocates its demands. If it allocates demands to more than one facility, demand site i is multi-sourced and variable E_i must take on a value of one. Finally, constraints (3.14) ensure that at least one demand site is not multi-sourced. We note that the right hand side of constraint (3.14) represents the maximum number of multi-sourced demand sites that are allowed, and can be changed to any non-negative value the user desires. For example, if we want

[1] For the example data, the optimal cost of the solution to the IMCLP model in which we ensure at least one demand site is not multi-sourced is \$83,150.

to ensure that the number of multi-sourced sites does not exceed the number of located processing facilities minus one, the right hand side of constraint (3.14) can be replaced by $\sum_{j \in J} X_j - 1$.[1] Adding constraints (3.12) - (3.14) to formulation (2.2) - (2.4) and (3.2) - (3.8) results in a larger objective function value for this problem instance, thereby showing that there is no alternate optimum in which fewer demand nodes are multi-sourced.

3.4 Computational Results

We now provide computational results for the IMCLP. In particular, we test the sensitivity of the IMCLP solution time as the size of the problem instance increases and compare the optimal IMCLP solution to that of the CFLP. The models are coded in AMPL and the MIP problems are solved using CPLEX 12.2.0 on an Intel 3.20 GHz Xeon CPU with 8 GB of memory.

Model data are generated based on 2010 U.S. census population data from the 500 largest U.S. counties. In the discussion that follows, we refer to counties as nodes and refer to a node in terms of its population rank. (i.e., The largest county, Los Angeles, CA, has rank 1 while the 500th largest county, Boone, KY, has rank 500.) Multiple problem instances are created by varying the number of demand-generating sites and candidate processing facilities. The data instances use the largest (with respect to the population) nodes as demand-generating sites and possible candidate facility locations. For example, an instance with 100 candidate locations includes the 100 largest nodes while a data instance with 150 candidate locations includes the same nodes as the 100 candidate location dataset plus the next 50 largest nodes. From these data instances we generate the necessary model parameters as follows.

[1] For the example data, the optimal cost of the solution to the IMCLP model in which we restrict the number of multi-sourced demand nodes to be no greater than the number of located processing facilities minus 1 is $108,950

Unless otherwise noted, for each demand site $i \in I$, discrete, random daily demands are generated for a 100 day time horizon (plus the warm-up days required for initialization) from a Poisson distribution with a mean of 1/10,000 of the county population, ρ_i. To remain consistent, the average daily demand at site $i \in I$ is set to $\frac{1}{10{,}000}\rho_i$ in the CFLP instances. The second and third columns in Table 11 display the population and corresponding average daily demand parameter used to generate the daily Poisson distributed demands for the five largest demand sites, respectively. Demands were generated at this rate seven days a week at all the demand sites, and the resulting sample statistics are presented in columns four and six of Table 11. Additionally, Table 12 provides further information regarding the population range of the 500 largest counties.

Table 11: Demand statistics for 5 largest demand sites

Demand Site	Population (ρ_i)	Theoretical Average Daily Demand $\left(\frac{1}{10{,}000}\rho_i\right)$	Sample Average Daily Demand	Theoretical Standard Deviation of Demand	Sample Standard Deviation of Demand
Los Angeles, CA	9,818,605	982	984.8	31.3	29.3
Cook, IL	5,194,675	519	519.6	22.8	22.4
Harris, TX	4,092,459	409	410.1	20.2	20.4
Maricopa, AZ	3,817,117	382	380.8	19.5	19.1
San Diego, CA	3,095,313	310	309.0	17.6	16.0

Table 12: Population Range of 500 Largest US Counties

Node Rank	Max. ρ_i	Min. ρ_i	Avg. ρ_i	St Dev ρ_i
1-50	9,818,605	916,924	1,833,113	1,442,311
51-100	916,829	618,754	754,214	91,285
101-150	603,403	434,972	511,279	48,884
151-200	432,552	316,236	373,976	36,902
201-250	315,335	255,793	284,748	16,693
251-300	255,755	203,206	229,034	16,506
301-350	203,065	174,528	189,518	9,097
351-400	174,214	154,358	163,353	5,515
401-450	153,990	136,484	144,544	5,696
451-500	136,146	118,811	127,719	5,082

Distances between nodes are calculated based on the Great Circle Distance (e.g., Weisstein (2016)). We use this distance to calculate the travel time between nodes under the

assumption that it takes one day to process the outgoing order at the demand-generating site and one day to transport an item 100 miles. The travel times are integer valued. Thus, if the Great Circle Distance between two nodes is 402 miles, the travel time between the nodes is six days ($\lceil 402/100 \rceil = 5$ days for traveling plus one day for processing). Under this assumption, demands are never processed on the day they originate, in accordance with the discussion in Section 3.2. For example, suppose a processing facility is located at node A and also that the demand generated at node A is assigned to be processed at node A. All of the demands that originate at node A on day d will arrive to be processed at node A on day $d + 1$. Such a scenario resembles that of a postal service policy in which mail carriers collect mail from various houses and businesses during the day and bring the collected mail to the post office at the end of the business day. If we assume that there is no night shift at the post office, the mail must wait until the next business day to be processed. Thus, the effective travel time plus processing time is one day.

The daily facility location costs are calculated by specifying a daily base cost plus an additional cost calculated from the associated demand (population) generated at that location. Specifically, we use the formula

$$z + w\rho_j$$

to generate the daily facility location cost for each facility $j \in J$, where z denotes the daily base fixed cost and w is the population weight. For the instances that follow, we let the daily base fixed cost be $z = \$10{,}000$ and the weight be $w = \$0.0001$. For example, Los Angeles, CA, which has a population of 9,818,605, has a daily fixed cost of $\$10{,}000 + \$0.0001 * 9{,}818{,}605 = \$10{,}982$ while Du Page, IL, which has a population of 916,924, has a daily fixed cost of $\$10{,}000 + \$0.0001 * 916{,}924 = \$10{,}092$. We note that the population weight w greatly

affects which facilities are located; larger values of w result in greater disparities between the cost of locating a facility in the most populous county (i.e., Node 1- Los Angeles, CA) and in a less populous county (e.g., Node 50, Du Page, IL). We direct the interested reader to Appendix B for an analysis of the effect of the population weight on the optimal solution.

We assume that each facility, regardless of its location, has a uniform capacity of one fifth of the total average daily demand, rounded up to the nearest hundred. Thus, the capacity of each processing facility depends on which (and how many) demand nodes are considered. For example, since the 50 largest demand sites generate a daily average of 9,166 units of demand and 1,833.2 is one fifth of 9,166, each candidate processing facility has a daily processing capacity of 1,900 units. Furthermore, the capacity of each processing facility in an instance with 50 demand nodes is less than the capacity of each processing facility in, for example, a 100 demand node instance since the latter instance generates more total average demand each day. The specific capacity levels of the problem instances used in this chapter are reported in Table 13.

Table 13: Daily processing capacity per candidate facility

# Demand Nodes	Total Average Daily Demand	1/5 Total Average Daily Demand	Daily Capacity per Candidate Facility	Daily Capacity % of Total Average Daily Demand
50	9,166	1833.2	1900	20.7%
100	12,936	2587.2	2600	20.1%
150	15,491	3098.2	3100	20.0%
200	17,360	3472.0	3500	20.2%
250	18,787	3757.4	3800	20.2%
300	19,929	3985.8	4000	20.0%
350	20,878	4175.6	4200	20.1%
400	21,696	4339.2	4400	20.3%
450	22,421	4484.2	4500	20.0%
500	23,054	4610.8	4700	20.4%

If there is sufficient capacity in the system for the CFLP to accommodate all demands, then, in expectation, there will be sufficient capacity in the system for the IMCLP model to accommodate all demands. This is because the processing capacities are the same in the two

models and the daily demands in the IMCLP model are realizations of a random distribution with constant mean equal to the mean used in the CFLP. If the underlying demand process is time dependent (i.e., the mean demand at a location changes with time), then a different model must be applied [Drezner and Wesolowsky, 1991].

Conversion factors a and b are chosen to limit the number of items held in backlog from day to day to a reasonable value and will be specified in each problem instance individually. We consider a and b values to be reasonable if the backlog does not grow without bound at any processing facility when five facilities are located.

Unless otherwise noted, the following results are generated using instances constructed using the aforementioned method.

3.4.1 Empirical Number of Multi-Sourced Demand Nodes

To compare the theoretical results regarding the number of multi-sourced demand sites discussed in Section 3.3.3 with empirical results, we investigate the percent of demand sites that are multi-sourced for numerous problem instances. Each problem instance uses a time horizon of 100 days and the 50 most populous counties as candidate facility locations. The number of demand sites varies in increments of 50 to 500. The results in Table 14 indicate that although in theory all of the demand sites could be multi-sourced, no more than 8% of the demand sites were multi-sourced in the optimal solution of the data instances we considered. Moreover, all but one of the optimal solutions had at most 3% of the demand sites multi-sourced. Nevertheless, the instances with 300, 400, and 450 demand sites again confirm that the number of multi-sourced sites can exceed the number of facilities located minus one in the optimal solution.[1]

[1] Since there may be multiple alternative optima, we note that the number of multi-sourced demand nodes reported in Table 14 for each problem instance is the minimum number of multi-sourced demand nodes across all

Table 14: Number of multi-sourced demand sites; $a = 1$, $b = 3$

# Demand Nodes	# Candidate Nodes	# of Multi-Sourced Demand Nodes	% of Demand Nodes that are Multi-Sourced	# Open Facilities	Rank of Open Facilities
50	50	4	8.00%	5	2,3,6,7,13
100	50	3	3.00%	6	11,16,20,25,30,33
150	50	4	2.67%	6	11,16,25,33,35,43
200	50	5	2.50%	6	9,11,25,26,30,33
250	50	3	1.20%	6	10,11,15,16,25,35
300	50	6	2.00%	6	9,11,15,20,25,35
350	50	4	1.14%	6	2,9,11,20,25,48
400	50	9	2.25%	6	2,9,11,20,25,48
450	50	8	1.78%	6	2,9,11,20,25,48
500	50	5	1.00%	6	2,9,11,20,25,48

Using the 50 demand, 50 candidate facility locations instance as a base case, we increase the backlog weight, b. Our results indicate that the number of multi-sourced demand sites does not increase until $b > 100$. However, as b increases past 100, the number of multi-sourced demand sites significantly increases. The number of located processing facilities remains constant at five located facilities. However, in general, the backlog weight can change the number of located facilities; typically, an additional facility will be located to accommodate excess backlog if the backlog weight is sufficiently high. The results are presented in Figure 4.

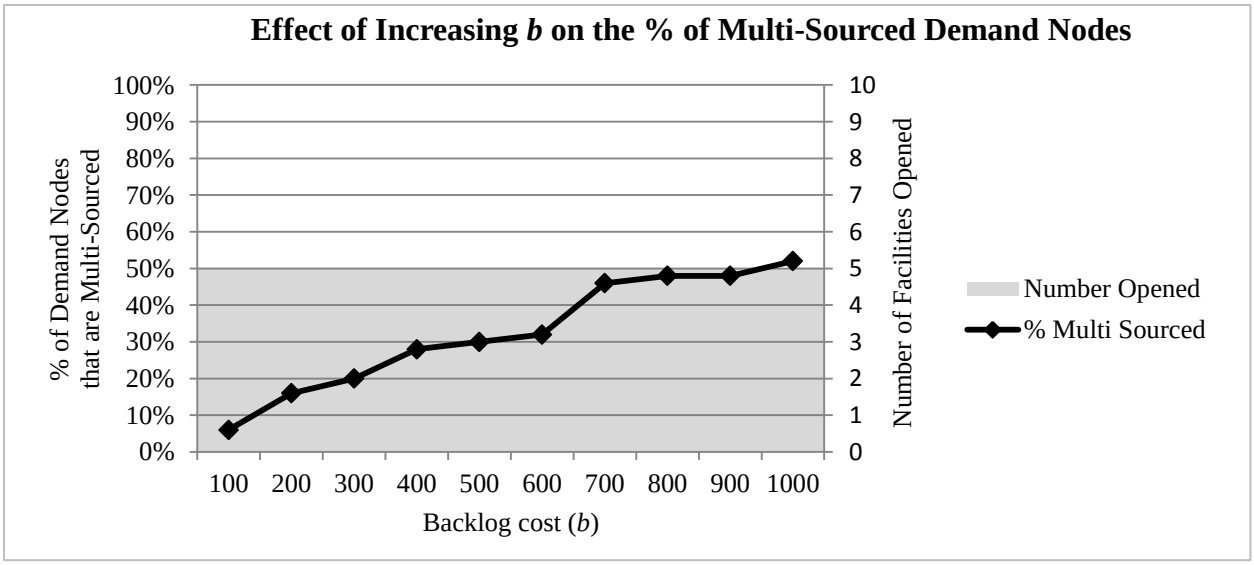

Figure 4: Effect of increasing backlog weight, 50 demand & candidate nodes, 100 days, $a = 1$

corresponding optimal solutions. This was verified by observing a higher optimal cost when we imposed a constraint that restricted the number of multi-sourced demand nodes to be less than the number reported in the fifth column of Table 14.

3.4.2 Comparison to CFLP

Next, we compare the effect of incorporating capacity into facility location models via the CFLP and the IMCLP. We find that the CFLP and IMCLP generally do not have the same optimal solutions in terms of the location and allocation decisions. In fact, the two models can choose to not only locate different processing facilities, but also locate a different *number* of processing facilities. In the following paragraph we present one such example in which the CFLP locates fewer facilities than the IMCLP. While the demand data used in the CFLP and IMCLP are generated from the same daily demand mean, the IMCLP differs from the CFLP in that if the total demand that arrives at a processing facility exceeds the processing capacity, a penalty cost is incurred because the unprocessed demand must wait in backlog to be processed at a later date. If this cost is sufficiently high, the IMCLP model will choose to locate an additional facility to mitigate the effect of the backlog penalty costs.

For the results in this section, random daily demands for the IMCLP were generated from a Poisson distribution with a mean of 1/5,000 of the node population. Since the CFLP requires as input a generic daily demand value, we set the average daily demand for the CFLP to be 1/5,000 of the demand node population to be consistent with the manner in which the IMCLP data were generated. The facility capacities, fixed location costs, and travel times between sites were generated in the manner described earlier in Section 3.4 except that a weight of $w = 0.01$ was used for the fixed location costs. In this case, the daily facility location costs of the candidate processing facilities range from \$108,186 (Node 1: Los Angeles, CA) to \$19,169 (Node 50: Du Page, IL).

We ran the IMCLP and CFLP models for the 50 demand, 50 candidate node problem instance with a 100 day time horizon. The conversion factors of $a = 1$ and $b = 100$ were used.

The traditional CFLP model located processing facilities in Fresno, CA; Pima, AZ; Travis, TX; DuPage, IL; and Westchester, NY while the IMCLP located an additional facility in Orange, FL. Note that the CFLP located five facilities while the IMCLP located six. The optimal location and allocation assignments of the two models can be seen in Figure 5 and Figure 6. The large boxes denote located facilities and the lines connect each demand site to its assigned processing facility. In both the CFLP and IMCLP solutions the demand sites were not necessarily assigned to the nearest located facility. For example, in the optimal CFLP solution, Fulton, GA is closer to DuPage, IL (589 miles, seven days) but was assigned to Travis, TX (818 miles, ten days) and in the IMCLP solution, Salt Lake, UT was assigned to Travis, TX (1068 miles, 12 days) although it is much closer to Fresno, CA (502 miles, seven days). These two demand sites are identified in Figure 5 and Figure 6 by a dotted circle.

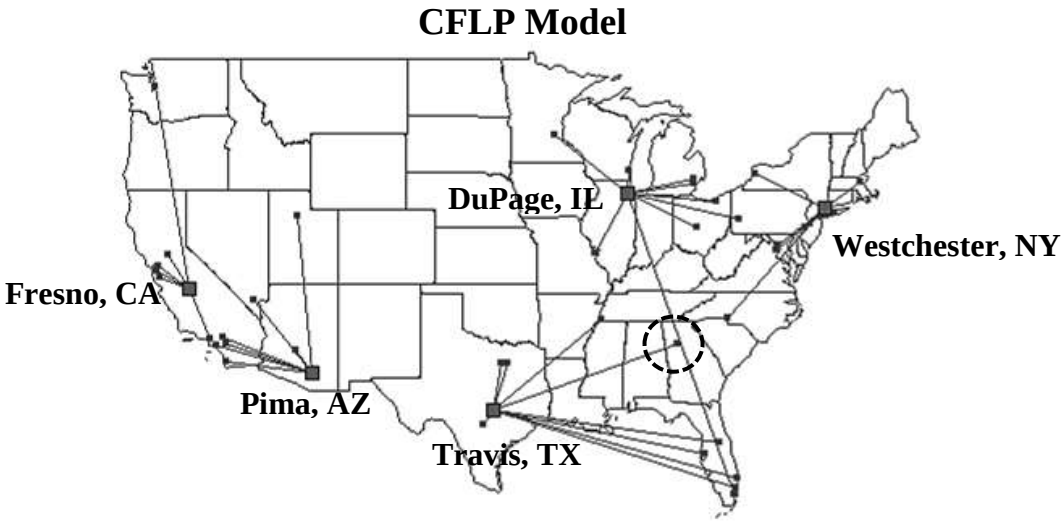

Figure 5: Optimal CFLP Solution; 50 demand & candidate nodes, $a = 1$;
mean daily demand=1/5000 node population, $w = 0.01$

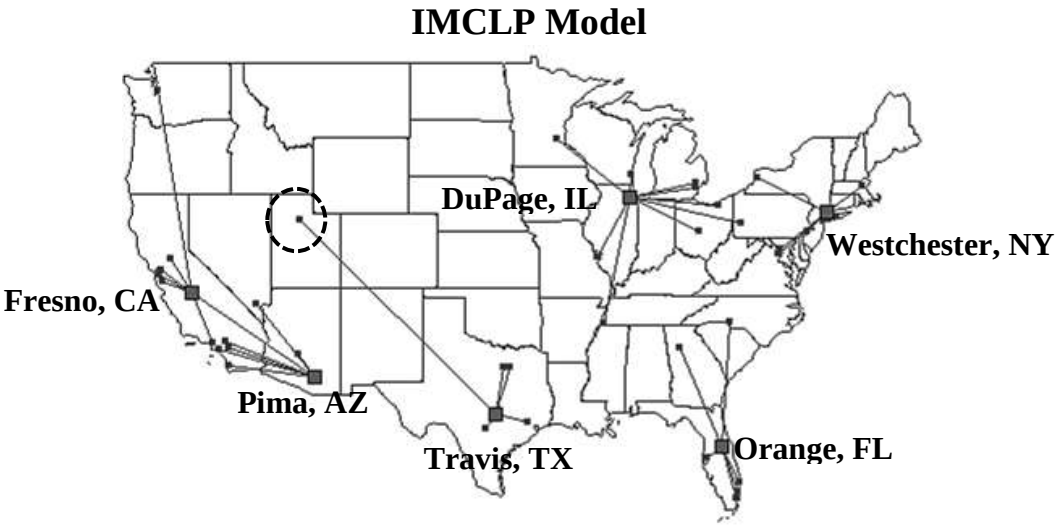

Figure 6: Optimal IMCLP solution; 50 demand & candidate nodes, $a = 1, b = 100$; mean daily demand=1/5000 node population, $w = 0.01$

As we will discuss in Section 3.4.3, preliminary results show that the CFLP solves much faster than the IMCLP. Therefore, as we did with a small example in Section 3.3., we again investigate the effect on the optimal objective function value of the IMCLP if demand allocations are given by the solution to the capacitated facility location problem as the relative cost of backlog to transportation cost varies. For this analysis we fix the transportation weight at $a = 1$ and vary the backlog weight, b. The results indicate that the solution obtained by using the CFLP allocations can provide very poor approximations to the optimal solution of the IMCLP, particularly as it becomes more costly to hold items in backlog. This is to be expected since the CFLP does not incorporate backlog in the objective function. For example, when $a = b = 1$ the optimal locations determined by the IMCLP and CFLP are the same. However, if a company experiences unit daily backlog carrying costs that are 100 times as much as the unit transportation costs, the company would see an undesirable total cost increase of 9% if the location and allocation decisions are determined via the CFLP rather than by the IMCLP. A major contributor to the cost difference between the two models is that the optimal IMCLP

solution locates six facilities whereas the corresponding CFLP solution only locates five facilities. This is in part due to the fact that the backlogging costs that would be accumulated if only five facilities were located can be mitigated by locating an additional processing facility. We also note that the optimal CFLP solution clearly does not change as the backlog weight, b, varies since backlog is not accounted for in the CFLP.

3.4.3 Effect of Problem Size

In this section we investigate the effect of problem size on solution times by increasing the number of demand and candidate nodes simultaneously, increasing the number of demand nodes only, and increasing the time horizon while keeping the number of demand and candidate nodes constant. We present the solution time in seconds or, if a provably optimal solution could not be found within one hour, we present the percent optimality gap as obtained by CPLEX 12.2.0.

Increasing the Number of Nodes

We increase the number of demand and candidate nodes simultaneously and report the solution times and processing facilities located for the CFLP, IMCLP, and IMCLP with the assignment variables relaxed from binary (i.e., Relaxed IMCLP). The solution time or optimality gap after one hour for data instances ranging from 50 to 500 demand and candidate nodes are presented in Table 15. We report a 100% optimality gap for problem instances in which no feasible solution is found within one hour. We find that the CFLP solves much quicker than both the IMCLP and Relaxed IMCLP problems. This is expected due to the additional constraints and variables in the IMCLP that track the daily activity of the processing facilities.

Table 15: Solution times (in seconds) or % optimality gap (indicated by shading) within 1 hour using a generic solver; $a = 1$, $b = 3$

# Demand & Candidate Nodes	CFLP	IMCLP	Relaxed IMCLP
50	0.52	6.35	8.11
100	192.74	1544.11	496.49
150	0.58%	3.70%	1.55%
200	0.93%	3.75%	2.26%
250	1.69%	3.54 %	2.99%
300	2.53%	5.66%	4.77%
350	2.02%	5.69%	3.13%
400	2.65%	51.26%	3.29%
450	5.14%	92.10%	100%
500	2.99%	100%	100%

Preliminary results indicate that the IMCLP model tends to locate processing facilities at nodes corresponding to larger populations (and therefore at the nodes that generate the largest demand). Thus, we reduce the set of candidate processing facilities by eliminating the candidate nodes with the lowest population to obtain an approximate solution.

We begin by using the 100 demand, 100 candidate node instance with 100 days of daily demand data, $a = 1$, and $b = 3$ to illustrate the results. We consider various candidate node sets, including the set in which all 100 of the demand nodes also serve as candidate nodes and the set with only the nodes corresponding to the ten largest populations. The tradeoff in solution time and cost is displayed in Figure 7. The cost increase is compared to the optimal cost when all 100 nodes are in the candidate node set. As expected, as the number of candidate nodes decreases, the solution time also decreases while the cost increases. We find that a similar result holds for the CFLP.

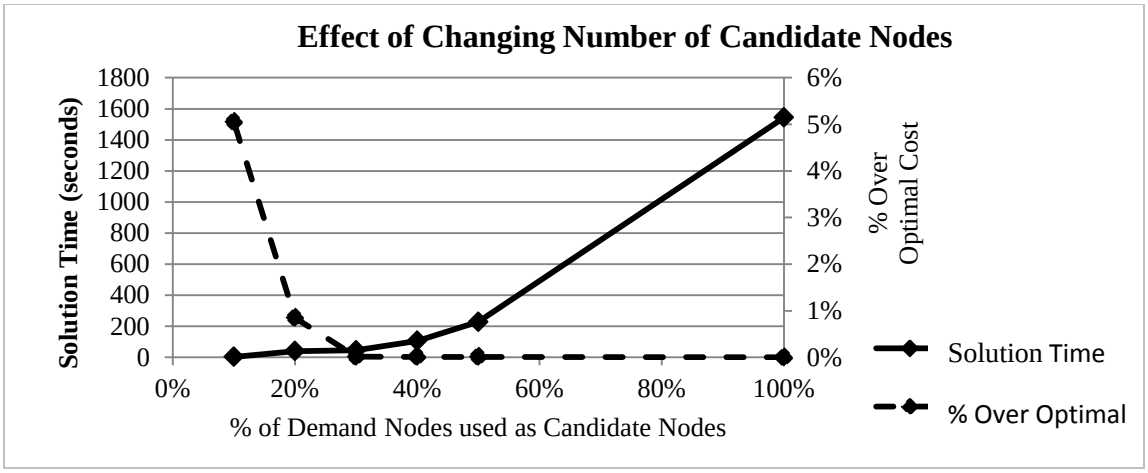

Figure 7: Effect of changing number of candidate nodes; 100 demand & candidate nodes, 100 days, $a = 1$, $b = 3$

For example, the optimal solution using all 100 nodes as candidate nodes locates processing facilities at the 11th, 16th, 20th, 25th, 30th, and 60th largest nodes (i.e., Riverside, CA; Tarrant, TX; New York, NY; Sacramento, CA; Hillsborough, FL; and Macomb, MI, respectively) and takes 1544.11 seconds to solve. However, if we restrict the set of candidate nodes to include only the 30 nodes with the largest population, the solution time is reduced to 45.29 seconds (a 97% reduction) and the cost of the solution obtained is only 0.0134% more than the optimal solution. The 30 candidate node problem locates all but one of the same processing facilities as does the problem with 100 candidate nodes; the 60th largest county (i.e., Macomb, MI) located in the 100 candidate node problem is replaced by the 15th largest county (i.e., Wayne, MI) in the 30 candidate node problem.

Since reducing the size of the candidate node set was beneficial in terms of the solution time with only a minimal increase in cost for the 100 demand node problem instance, we next vary the number of demand nodes from 50 to 500 in intervals of 50. The transportation and backlog weights are the same as those listed in Table 15. In Figure 8 we display the IMCLP optimality gap after one hour of run time when all of the demand nodes are candidate nodes, as well as when only the 50 and 100 largest demand nodes are candidate nodes. The IMCLP data

referenced in Table 15 is used for the "All Nodes" case. It is clear that reducing the size of the candidate node set to include only the 50 or 100 largest nodes significantly improves the solution progress within one hour.

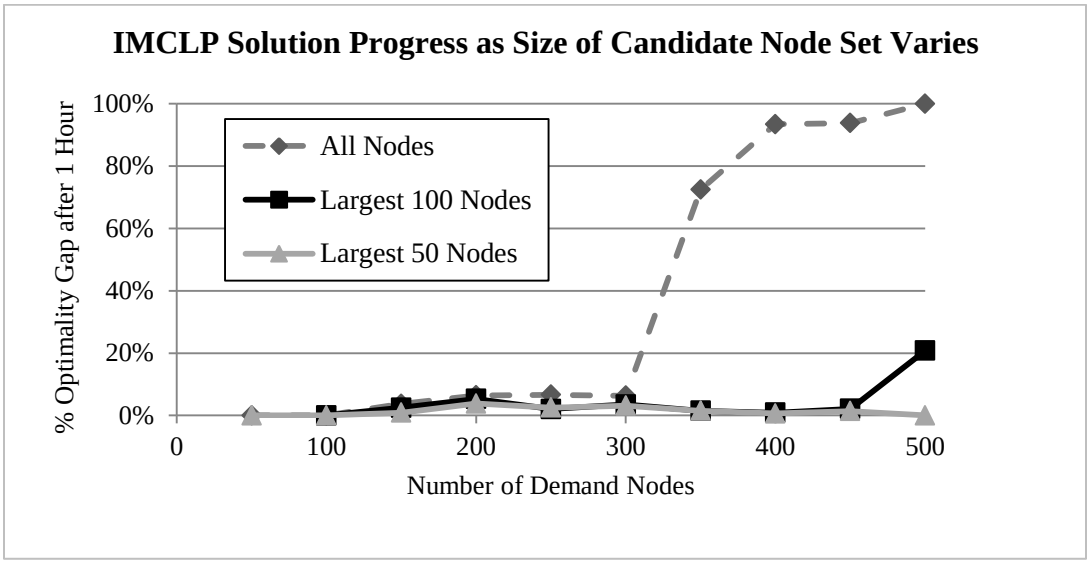

Figure 8: IMCLP solution progress as size of candidate node set varies; 100 days; $a = 1$, $b = 3$

Additionally, Figure 8 suggests that restricting the size of the candidate set to the 50 largest nodes usually does not provide a significant reduction in the solution progress when compared to that of the largest 100 node candidate node set. An exception to this is the 500 demand node instance, which solves to optimality when only the 50 largest nodes are in the candidate node set but has a 20.8% optimality gap after one hour when the 100 largest nodes are in the candidate node set.

When the same number of demand and candidate node instances are run for the CFLP we find that the resulting optimality gaps after one hour are reduced (see Figure 9). Of particular note are the instances consisting of 350-500 demand nodes when all of the demand nodes are candidate nodes as well. In the problem instance with 500 candidate nodes, the CFLP reported an optimality gap of 6.21% within one hour while the IMCLP does not find a feasible solution in

the same time frame. We note that the scale of the y-axis in Figure 9 is different than that of Figure 8.

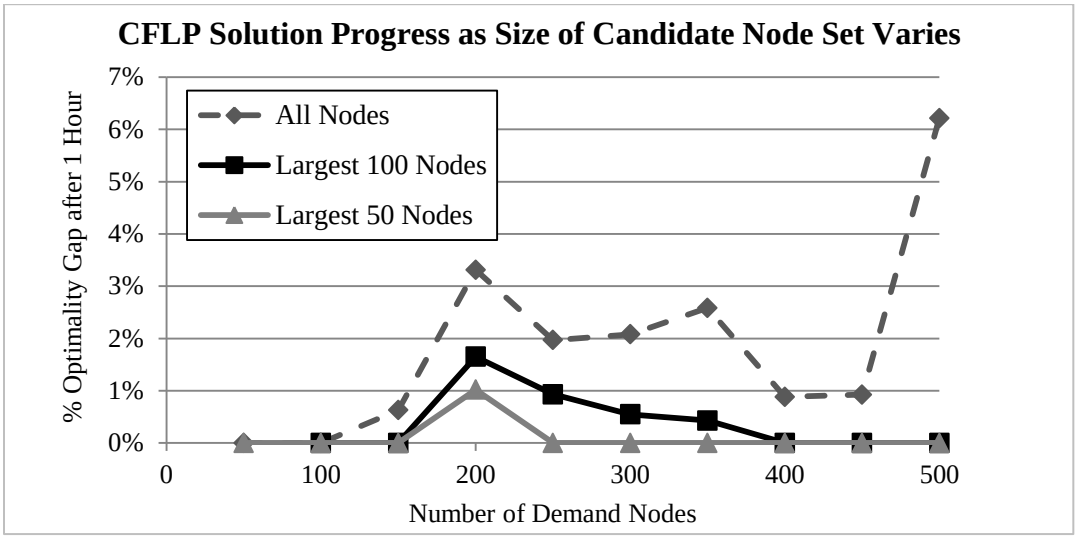

Figure 9: CFLP solution progress as size of candidate node set varies; $a = 1$

Increasing the Number of Days in the Planning Horizon

Since Table 15 indicates that the 50 demand, 50 candidate node instance produced an optimal solution relatively quickly for 100 days of daily demand data, we investigate the sensitivity of the optimal location solutions with respect to the planning time horizon. The time horizon varies from 10 to 4,000 days and the transportation and backlog weights are the same as those listed in Table 15. The daily demands in each problem instance are independent of the demands in the other problem instances. For example, the daily demands for the ten days in the ten day time horizon instance are different than the first ten days in the 4,000 day time horizon instance. Additionally, we report whether the optimal facilities that are located change as the number of days increase. Figure 10 displays the solution time for the problem. All problem instances were solved to optimality within one hour.

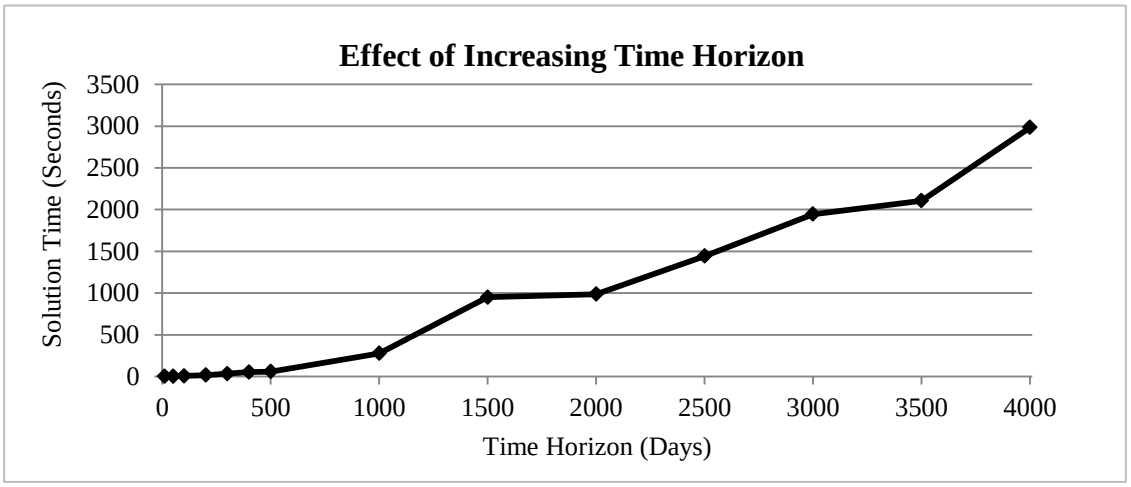

Figure 10: Effect of increasing time horizon, 50 demand & candidate nodes, $a = 1, b = 3$

We find that the optimal processing locations are relatively insensitive to the length of the time horizon. The optimal locations remain constant for time horizons of 100 to 3,500 days. For various problem instances with a time horizon of less than 100 days and the problem instance with 4,000 days a slight change in the optimal facility locations occurs. In particular, for these problem instances, it is optimal to locate a processing facility in Riverside County rather than in Orange County. However, we note that these two counties are neighboring counties in California and, as such, the increase in cost of locating the processing facility in Orange County rather than in Riverside County is only 0.0001237% in the 4,000 day instance. We also note that the solution to the IMCLP instances that locate a facility in Riverside County is precisely the optimal solution to the corresponding CFLP problem in terms of the processing facilities located.

3.5 Chapter Summary

We have argued that the hard capacity constraints that are often employed in location modeling are unrealistic since facility managers have numerous operational tools that allow a facility to accept items in excess of the stated processing capacity level. As such, we have

developed a new approach to modeling capacity constraints in a facility location and allocation model that permits storing excess demand as backlog to be processed at a later date.

We reported the difference in facility locations, demand allocations, and total cost between the CFLP and our approach. In particular, we have shown that the CFLP and IMCLP often do not locate the same facilities (and in some instances they locate a different number of facilities), and therefore also do not allocate demands to the same processing facilities. Since the CFLP does not incorporate the cost of backlogged demand, the CFLP *underestimates* the total cost, but also results in solutions with a *higher total cost* (as compared to the IMCLP) when the demand is realized and backlogging is allowed in practice. Such an underestimate could have significant implications on budget predictions.

Additionally, we compared the effect of increasing the problem size by increasing the number of demand and candidate nodes, only the number of demand nodes, and the time horizon. Computational results indicated that increasing the number of demand and candidate nodes simultaneously adds a significant burden on the IMCLP solution time. However, for the instances considered, we found that by limiting the candidate nodes to include only a subset of the largest populated demand nodes, we significantly reduced the solution time and still obtained a solution that is near optimal to the original problem. We also found that the optimal locations to the IMCLP are relatively insensitive to the length of the time horizon for the problem instances considered. Detailed analysis regarding the effect of a wider variety of model parameters on these conclusions is left as an area of future research.

A key benefit of the IMCLP model is that it incorporates a penalty cost associated with the reality that on some days the total amount of demand arriving at a processing facility may exceed the daily processing capacity. The IMCLP formulation inherently assumes that demands

do not expire; demands can be postponed indefinitely as long as the penalty cost is paid. However, we realize that in many situations this is not realistic and therefore we presented additional constraints that can be added to the IMCLP to ensure demands are processed within a specified amount of time.

Furthermore, the IMCLP can use a dataset as a direct input into the model rather than specifying deterministic demands at an aggregate level by using parameter estimation techniques. This allowed us to explicitly incorporate the variation in and the possibly correlated nature of demands. As a result, variations of this model have the potential to reveal operating policies that take advantage of spatial and temporal correlations in demand that are not evident in current facility location models. An example of such a variation is an allocation policy that varies by day of the week to take advantage of regular daily fluctuations in demand and/or variations in capacity that result from operating the processing facility for different hours each day of the week. Demand site $i \in I$ may allocate its demands to facility $j_1 \in J$ on weekends and to facility $j_2 \in J$ on weekdays, for example. Such extensions are presented in the next chapter.

CHAPTER 4: A Cyclic Allocation Model for the IMCLP

4.1 Motivation

Processing facilities can have many different operational policies, even within a single organization. For example, a portion of the facilities may operate at full capacity 24 hours per day, seven days a week, others may reduce their processing capacity on weekends, and some may only be open on weekdays. In this chapter we show that allowing demands at a particular demand-generating site to be allocated to multiple processing facilities in a cyclic manner that accounts for capacity levels that vary by day-of-the-week can be a cost effective operational tool. Furthermore, such an allocation method affords the model flexibility to utilize an allocation policy that captures systematic day-of-the-week *demand* patterns, even when all processing centers are operational at identical times.[1]

Thus, we expand the IMCLP to include a cyclic demand allocation approach. While the IMCLP restricts each demand site to allocate its demand to a single processing facility for the entire demand horizon, we now relax that assumption and allow demands to be allocated to multiple processing facilities over a specific time frame (e.g., a week), but to a single facility each day. As an example of such a system, blood is often drawn at local clinics and sent for testing at more regional centers [Saveh-Shemshaki, 2012]. If some testing centers operate five days per week and others operate seven days per week, a day-of-the-week assignment policy can reduce costs and improve the time-to-processing relative to a single assignment policy. For

[1] A majority of the content of this chapter has been published in Maass and Daskin (2017).

example, a clinic may send its blood samples to arrive at one testing facility (that is closed on weekends) on Mondays through Fridays and to another testing facility (that is open on weekends) on Saturdays and Sundays. This allows for a more efficient use of the processing system by mitigating the amount of unprocessed blood samples that are held in cold storage over the weekend.

4.2 Model Formulation

The following model formulation generates *cyclic* day-of-the-week assignments, which account for the deterministic day-to-day variation in processing capacity levels as well as deterministic cyclic patters and correlations among the demand sites.

Let the set $P \coloneqq \{0, 1, \dots, 6\}$ represent the days of the week, where element $p = 0$ represents Saturday, $p = 1$ represents Sunday, etc. In addition to the notation used in the IMCLP, we also make the following additions. Assume, without loss of generality, that the first day in D is a Sunday to match element $p = 1$ in set P. Additionally, for each $p \in P$, we can also generate the set $D_p \coloneqq \{d \in D: d \bmod |P| = p, d > t^*\}$, which contains the days in the planning horizon that correspond to the day of the week indicated by p. For example, if $D = \{1, 2, \dots, 42\}$ and $t^* = 1$, then $D_2 = \{2, 9, 16, 23, 30, 37\}$ would represent the set of Mondays in D. However, if $t^* = 2$, then we would not include day 2 in the set $D_2 = \{9, 16, 23, 30, 37\}$.

Note that we assume that the demand *rate* may vary by day-of-the week, but that it does not change from week to week. That is, while the *realizations* of the daily demand for a particular demand site will fluctuate throughout the planning horizon since they are samples from a distribution, the realizations are taken from the same distribution each Sunday. However, a demand site might have a different demand distribution for Mondays than Sundays.

In addition to the notation used to formulate the IMCLP, we introduce the following for the cyclic allocation model:

Sets and Parameters

$P \coloneqq \{0,1,\dots,6\}$ Set of cycle days (e.g., days of the week)

$D_p \coloneqq \{d \in D: d \bmod |P| = p, d > t^*\}$ Set of days in D corresponding to cycle day $p \in P$ after the warm-up period

k_j^p Capacity of facility $j \in J$ in items processed per day on cycle day $p \in P$

Decision Variables

$$Y_{ij}^p = \begin{cases} 1 & \text{If we ship demands from demand site } i \in I \text{ to faciliy } j \in J \text{ on cycle day } p \in P \\ 0 & \text{Otherwise} \end{cases}$$

With this notation, we formulate the cyclic allocation model as follows:

Formulation

$$Min_{\boldsymbol{X},\boldsymbol{Y},\boldsymbol{V},\boldsymbol{W}} \ (|D| - t^*) \sum_{j\in J} f_j X_j + a \sum_{p\in P} \sum_{d\in D_p} \sum_{j\in J} \sum_{i\in I} h_{id} t_{ij} Y_{ij}^p + b \sum_{d=t^*+2}^{|D|+1} \sum_{j\in J} V_{jd} \quad (4.1)$$

$Subject\ to$

$$\sum_{j\in J} Y_{ij}^p = 1 \qquad \forall i \in I;\ p \in P \qquad (4.2)$$

$$Y_{ij}^p \le X_j \qquad \forall i \in I; j \in J; p \in P \qquad (4.3)$$

$$V_{j,d+1} = V_{jd} + \sum_{i\in I} h_{i,d-t_{ij}} Y_{ij}^{(d-t_{ij}) \bmod |P|} - W_{jd} \qquad \forall j \in J; d \in \{t^*+1,\dots,|D|\} \qquad (4.4)$$

$$W_{jd} \le k_j^p X_j \qquad \forall j \in J; p \in P; d \in D_p \qquad (4.5)$$

$$V_{j,t^*+1} = v_j X_j \qquad \forall j \in J \qquad (3.5)$$

$$X_j \in \{0,1\} \qquad \forall j \in J \qquad (2.4)$$

$$Y_{ij}^p \in \{0,1\} \qquad \forall i \in I; j \in J; p \in P \qquad (4.6)$$

$$V_{jd} \ge 0 \qquad \forall j \in J; d \in \{t^*+1,\dots,|D|+1\} \qquad (3.7)$$

$$W_{jd} \geq 0 \qquad \forall j \in J; d \in \{t^* + 1, \dots, |D|\} \qquad (3.8)$$

As in the IMCLP, objective function (4.1) minimizes the facility location, transportation, and backlogging costs by determining the optimal facility locations and demand allocations, as well as the resulting auxiliary processing and backlog variables. This must be done in a manner such that each demand site $i \in I$ is assigned to exactly one processing facility $j \in J$ on each cycle day $p \in P$ (constraints (4.2) and (4.6)). Furthermore, constraints (4.3) state that demands can only be assigned to located facilities.

Constraints (4.2) - (4.5), (4.6) are the cyclic allocation constraints that correspond to constraints (2.2), (2.3), (3.3), (3.4), and (3.6), respectively, in the IMCLP. The major difference between this model and the IMCLP is that this model recognizes that demand sites and processing facilities may not operate in the same manner every day of the week. The cyclic formulation facilitates this by allowing the processing facilities to have different processing capacities throughout the week (i.e., we use capacity parameters k_j^p whereas the IMCLP uses a constant capacity k_j every day of the week) and by allowing demand sites to be allocated in a manner that accounts for the capacity variation and systematic day-to-day variation in the demands (i.e., we use allocation decision variables Y_{ij}^p rather than the Y_{ij} variables of the IMCLP). Such day-of-the-week allocation variables necessitate the inclusion of the sets P and D_p, both of which are not present in the IMCLP formulation.

While the daily variation in the processing capacity levels are model inputs, and therefore known a priori, we can model capacity as a decision variable with an associated cost based on the amount of capacity employed. Since the capacity level of a processing facility may dictate the respective staffing needs, such an extension considers both strategic and tactical decisions.

We discuss possible formulations of incorporating endogenous capacity flexibility in Section 6.5.1 and leave the analysis of such models as an area of future research.

4.3 Model Properties

We present a small example to illustrate that, from a cost minimization perspective, incorporating a day-of-the-week assignment policy can perform arbitrarily better than using the IMCLP model as the cost of backlogging varies.

Consider a system of five clinics that each draw an average of ten blood samples per hour that they are open. Some clinics are open nine hours each day, while others are open 8, 10, or 12 hours depending on the day of the week. Others are not open on weekends. Each clinic also serves as a candidate blood testing facility. It takes two days for demands to be shipped from one site to another for processing and one day for demands to be ready for processing at their own processing facility. We assume each located processing facility can process up to 100 demands per day and costs $200 to operate each day. Note that in this example, the processing capacity of each facility does not vary from day to day; a facility will process blood samples seven days a week if we choose to locate at that facility, but may only draw blood in the clinic on some of the days. Poisson distributed demands are generated for a time horizon of 100 days with the daily distribution mean rate listed in Table 16. The per unit per day transportation cost is fixed at $a = 2$ and we vary the per unit per day backlogging cost, b.

Table 16: Example mean demand rate pattern

	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
Site 1	80	80	100	100	100	120	120
Site 2	0	90	90	90	90	90	0
Site 3	90	10	10	10	10	10	90
Site 4	90	90	90	90	90	90	90
Site 5	90	90	90	90	90	90	90

For this example, we compare the optimal IMCLP solution to the optimal cyclic allocation model solution. The computational results indicate that for values of $b \geq 30$, the IMCLP locates all five processing facilities and assigns demands to be processed at the site at which they originate. While the cyclic allocation model also locates all five facilities, it is able to reduce the total amount of backlog in the system to approximately 10% of the total backlog accumulated in the IMCLP due to its flexible allocation policy. (As an example, the optimal cyclic assignments for the case when $b = 60$ are given in Table 17.) The difference in the total backlog cost (accumulated over the time horizon) is shown in Figure 11 and the resulting cost difference between the IMCLP and cyclic allocation model is displayed in Figure 12; clearly, using the IMCLP can become arbitrarily more costly than the cyclic allocation model as b increases.

Table 17: Optimal cyclic assignments when $b = 60$ — example instance

	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
Site 1	Site 1	Site 1	Site 1	Site 1	**Site 2**	Site 1	**Site 2**
Site 2	**Site 3**	**Site 3**	Site 2	**Site 3**	Site 2	**Site 3**	**Site 1**
Site 3	Site 3	**Site 5**	Site 3	**Site 1**	Site 3	**Site 2**	Site 3
Site 4	Site 4	Site 4	Site 4	Site 4	Site 4	Site 4	Site 4
Site 5	Site 5	Site 5	Site 5	Site 5	Site 5	Site 5	Site 5

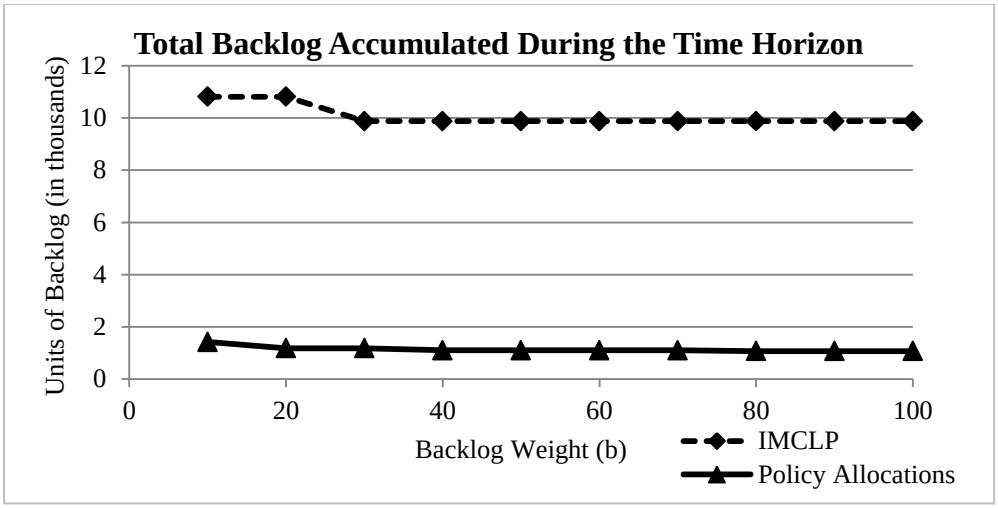

Figure 11: Total backlog—example instance

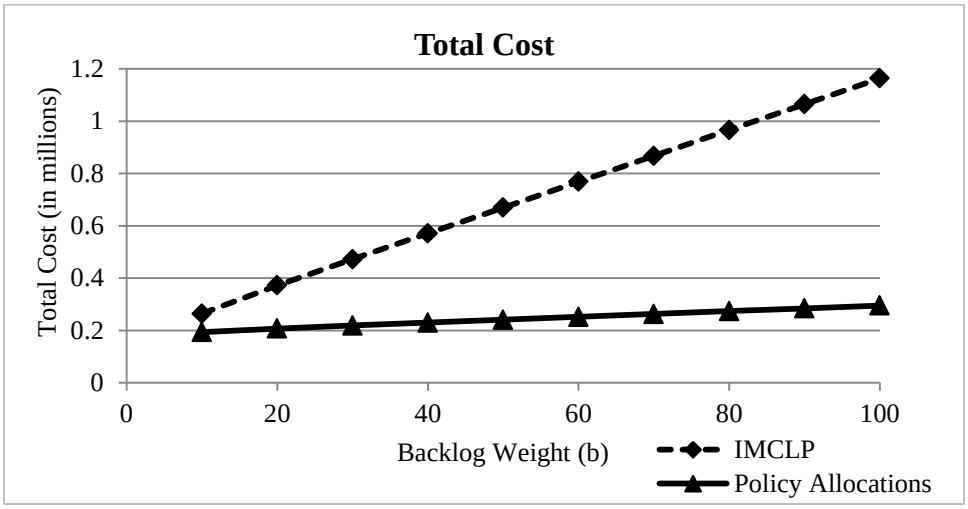

Figure 12: Total cost—example instance

4.4 Additional Constraints

Although we have shown that the cyclic allocation method has the potential for significant cost savings as compared to the IMCLP, the optimal solution that results from solving (2.4), (3.5), (3.7), (3.8) and (4.1) - (4.6) may provide an undesirable allocation decision from a managerial perspective. For example, an optimal solution may assign demands from a demand site to a different processing facility each day of the week. (In fact, in the simple 5-node example of Table 17, Site 3 ships to four of the five located facilities during the week.) This would likely

require the manager at the demand site to oversee seven individual contracts between the demand site and the processing facilities. It may be better for the manager's time to be spent overseeing only one or two contracts, even if it is more costly to do so. In this section, we describe additional constraints that may be useful for controlling the allocation policy.

4.4.1 Restricting the Number of Different Allocations

In many cases it is undesirable to allocate demands to a different processing facility each day of the week. Instead, we may wish to ensure that an individual demand site does not allocate its items to more than n different processing facilities. For any $n \in \{1, \dots, |P|\}$, we do this by defining s_{ij} as a binary variable that takes the value one if demand site $i \in I$ is allocated to facility $j \in J$ on *any* day of the week and by adding the constraints

$$Y_{ij}^{p} \leq s_{ij} \qquad \forall i \in I; j \in J; p \in P \tag{4.7}$$

$$\sum_{j \in J} s_{ij} \leq n \qquad \forall i \in I. \tag{4.8}$$

While these constraints restrict the number of unique allocations per demand site, they still allow a demand site's allocation to change every day by alternating between processing facilities. For example, if $n = 2$, then the demand site may allocate its demand to one processing facility on Sundays, Tuesdays, Thursdays, and Saturdays but allocate its demand to another processing facility on Mondays, Wednesdays, and Fridays.

IMCLP with Cyclic Allocations Model Property: Restricting n to take the value of one reduces the cyclic allocation model to the IMCLP. Since the IMCLP is NP-Hard, the cyclic allocation model with constraints (4.7) and (4.8) is also NP-Hard.

4.4.2 Weekday and Weekend Allocations

While the deterministic processing capacity of a facility may vary every day of the week, it is very common for some processing facilities to be closed on weekends while others are open seven days a week. For processing facilities that are only open Monday through Friday (i.e., weekdays), any demand that arrives at the facility on the weekend will be held in backlog and will be available for processing on Monday. Let ℓ_1 denote a facility that is only open on weekdays and let ℓ_2 represent a facility that is open seven days of the week. It is reasonable to assume that a manager at a demand site may wish to follow an allocation policy in which demands are shipped to arrive at facility ℓ_1 during the week and at facility ℓ_2 during the weekend so that the items do not have to wait in backlog at facility ℓ_1 due to the weekend closure. Since travel times are deterministic, we can ensure that such an allocation policy is adhered to in the following manner.

First, we define the sets Q_{ij} and R_{ij} to represent the set of cycle days for which demand shipped from demand site $i \in I$ would arrive at processing facility $j \in J$ on the weekend or during the weekday, respectively. Since the weekend encompasses two days (i.e., Saturday and Sunday), the set Q_{ij} contains two elements:

$$Q_{ij} \coloneqq \left\{\left(|P| - \left(t_{ij} mod |P|\right)\right) mod |P|, \left(|P| - \left(t_{ij} mod |P|\right) + 1\right) mod |P|\right\}$$

The remaining days correspond to Monday through Friday and are included in set R_{ij}:

$$R_{ij} \coloneqq \{p \in P \backslash Q_{ij}\}$$

The following examples illustrate how the Q_{ij} and R_{ij} sets change relative to the length of the travel time between a demand site $i \in I$ and a processing facility $j \in J$.

- **<u>If $t_{ij} = 1$,</u>** then demand sent from $i \in I$ on cycle days 6 (Fri.) and 0 (Sat.) will arrive at processing facility j on cycle days 0 (Sat.) and 1 (Sun.). We have

$$Q_{ij} \coloneqq \{(7 - (1\ mod\ 7))mod\ 7, (7 - (1\ mod\ 7) + 1)mod\ 7\} = \{6, 0\}$$

as desired. It follows that $R_{ij} \coloneqq \{1,2,3,4,5\}$.

- **If $t_{ij} = 3$,** then demand sent on cycle days 4 (Wed.) and 5 (Thurs.) will arrive on the weekend:

$$Q_{ij} \coloneqq \{(7 - (3\ mod\ 7))mod\ 7, (7 - (3\ mod\ 7) + 1)mod\ 7\} = \{4, 5\}$$

$$R_{ij} \coloneqq \{0,1,2,3,6\}$$

- **If $t_{ij} = 14$,** then demand sent on cycle days 0 (Sat.) and 1 (Sun.) will arrive on the weekend:

$$Q_{ij} \coloneqq \{(7 - (14\ mod\ 7))mod\ 7, (7 - (14\ mod\ 7) + 1)mod\ 7\} = \{0, 1\}$$

$$R_{ij} \coloneqq \{2,3,4,5,6\}$$

Once the sets $Q[i,j]$ and $R[i,j]$ are constructed, we can introduce binary variables

$$r_{ij} = \begin{cases} 1 & \text{If demands at } i \in I \text{ are assigned to facility } j \in J \text{ on a day such that they would arrive on a weekday} \\ 0 & \text{Otherwise} \end{cases}$$

and

$$q_{ij} = \begin{cases} 1 & \text{If demands at } i \in I \text{ are assigned to facility } j \in J \text{ on a day such that they would arrive on a weekend} \\ 0 & \text{Otherwise} \end{cases}$$

Then, the following two constraints can be implemented to enforce the weekday/weekend allocation policy:

$$\sum_{p \in R_{ij}} Y_{ij}^{p} = 5r_{ij} \qquad \forall i \in I; j \in J \qquad (4.9)$$

$$\sum_{p \in Q_{ij}} Y_{ij}^{p} \leq 2q_{ij} \qquad \forall i \in I; j \in J \qquad (4.10)$$

Together with constraints (4.2), constraint (4.9) ensures that demands at $i \in I$ are shipped to arrive at the same processing facility $j \in J$ on all five weekdays (i.e., Monday – Friday) and constraints (4.10) allow the demands to arrive at another facility on the weekends. Note that the inequality constraints of (4.10) do not *require* that the demand sites ship to multiple processing facilities; the demand site can always ship its demand to a single processing facility, even if the processing facility is closed on the weekend. Additionally, simply adding constraints (4.9) and (4.10) to the model of (2.4), (3.5), (3.7), (3.8) and (4.1) - (4.6) allows a demand site to be allocated to three processing facilities; it may be assigned to arrive at facility ℓ_1 on weekdays, facility ℓ_2 on Saturdays, and facility ℓ_3 on Sundays. If we want to ensure that a demand site does not allocate to a different facility on Saturdays than it does on Sundays, we would let $n = 2$ and include constraints (4.7) and (4.8). Alternatively, we could let $n > 2$ and add constraints

$$\sum_{j \in J} q_{ij} = 1 \qquad \forall i \in I \tag{4.11}$$

to constraint set (4.9) and (4.10) or simply change constraint (4.10) to an equality constraint.

4.4.3 Restricting the Maximum Allowed Travel Time

The optimal cyclic IMCLP solution may assign a demand site $i \in I$ to a processing facility $j_1 \in J$ that is relatively close during the week, but assign i to another processing facility $j_2 \in J$ that is very far away on weekends. To address this, we present the following three additional types of constraints that can be added to formulation (2.4), (3.5), (3.7), (3.8) and (4.1) - (4.10).

Maximal Travel Time Constraints

To ensure that no demand site is assigned to a processing facility that is unreasonably far away, we can define $t_{max} > 0$ as the maximum allowed travel time between demand sites and processing facilities, and add maximal travel time constraints of the form:

$$\sum_{j \in J} t_{ij} Y_{ij}^{p} \leq t_{max} \qquad \forall i \in I; p \in P \qquad (4.12)$$

to the model.

Alternatively, we could utilize coverage constraints [Farahani et al., 2012] by introducing the parameter c_{ij} to represent whether the travel time between a demand site and a processing facility is less than or equal to the maximal travel time allowed; the value of c_{ij} is set to one if $t_{ij} \leq t_{max}$ and 0 otherwise. Then, the following constraints state that on each cycle day, each demand site must be allocated to a processing facility that reachable within t_{max} or fewer days:

$$\sum_{j \in J} c_{ij} Y_{ij}^{p} = 1 \qquad \forall i \in I; p \in P \qquad (4.13)$$

Restricting the Time Difference between Weekday and Weekend Allocations

Another method of restricting the travel time is to add constraints that impose a limit, m, on the difference in travel time between a demand site and its weekday assignment, and the same demand site and its weekend assignment. Such constraints are of the form:

$$t_{ij_1} Y_{ij_1}^{p_1} - t_{ij_2} Y_{ij_2}^{p_2} \leq m + t^{*}(1 - Y_{ij_2}^{p_2}) \quad \forall i \in I; j_1, j_2 \in J; p_1 \in Q[i, j_1]; p_2 \in R[i, j_2]. \quad (4.14)$$

For example, when $m = 0$, demands at site $i \in I$ must take exactly the same number of days to travel from i to their weekend (i.e., Saturday, Sunday, or both) processing facility $j_1 \in J$ as it does to travel from i to their weekday processing facility $j_2 \in J$.

Closest Assignment Constraints

We consider closest assignment constraints as another method of reducing the difference in travel time between weekday and weekend assignments. Rather than restricting a demand site to be assigned to the closest located facility, we allow the demand site to be assigned to one (or two, in the case of weekday/weekend assignments with $\varkappa = 2$) of the ς closest located facilities. This is achieved by defining $G_{ij\hat{\varsigma}}$ as a binary variable that takes the value of one if facility $j \in J$ is the $\hat{\varsigma}$ closest located facility to demand site i and adding the following constraints:

$$\sum_{\hat{\varsigma}=1}^{|J|} \hat{\varsigma} G_{ij\hat{\varsigma}} + 1 \leq \sum_{\hat{\varsigma}=1}^{|J|} \hat{\varsigma} G_{il\hat{\varsigma}} + |J|(1 - X_l) \qquad \forall i \in I; j \in J; l \in \left\{\bar{J} \in J \middle| t_{ij} < t_{i\bar{J}}\right\} \quad (4.15)$$

$$\sum_{\hat{\varsigma}=1}^{|J|} \sum_{l \in J: t_{il} < t_{ij}} G_{il\hat{\varsigma}} + 1 \leq \sum_{\hat{\varsigma}=1}^{|J|} \hat{\varsigma} G_{ij\hat{\varsigma}} + |J|(1 - X_j) \quad \forall i \in I; j \in J \quad (4.16)$$

$$\sum_{\hat{\varsigma}=1}^{|J|} G_{ij\hat{\varsigma}} \leq X_j \qquad \forall i \in I; j \in J \quad (4.17)$$

$$\sum_{\hat{\varsigma}=1}^{\varsigma} G_{ij\hat{\varsigma}} \geq Y_{ij}^{p} \qquad \forall i \in I; j \in J, p \in P \quad (4.18)$$

Constraints (4.15) establish that if two facilities $j \in J$ and $l \in J$ are located, and the transportation time between facility j and demand site $i \in I$ is less than that between facility l and i, then j will be assigned a lower rank in relation to demand site i than l. Constraints (4.16) ensure that if there are $\sum_{\hat{\varsigma}=1}^{|J|} \sum_{l \in J: t_{il} < t_{ij}} G_{il\hat{\varsigma}}$ located processing facilities that are strictly closer to demand site i than processing facility j, then facility j receives a rank larger than $\sum_{\hat{\varsigma}=1}^{|J|} \sum_{l \in J: t_{il} < t_{ij}} G_{il\hat{\varsigma}}$ if j is located.

Constraints (4.17) state that facility j can be assigned at most one rank for each demand site i. If candidate facility j is not located, then facility j should not receive a rank for any demand site. We note that constraints (4.17) do not state that a rank can be given to at most one processing facility. This enables us to adequately address equidistant facilities. That is, if it takes demands at $i \in I$ two days to arrive at either processing facility $j_1 \in J$ or $j_2 \in J$, and no

processing facility is closer to $i \in I$ than j_1 or j_2, both j_1 and j_2 will be ranked as the closest processing facility to demand site i (i.e., $G_{ij_11} = G_{ij_21} = 1$). If processing facility j_3 is the next closest facility to i, then it will be counted as the third closest facility to i (i.e., $G_{ij_33} = 1$). Finally, constraints (4.18) ensure that demand site i can only be assigned to one of its ς closest located facilities.

These constraints can be used with the basic cyclic allocation model given by (2.4), (3.5), (3.7), (3.8) and (4.1) - (4.6) or in conjunction with any of the additional constraint sets listed in this section.

4.5 Computational Results

In this section, we discuss computational results for the cyclic allocation model. Our intent is to provide insights into the effect of imposing the various types of constraints introduced in the previous section rather than providing a discussion on the size of instances that can be solved using this formulation.

As in Section 3.4, discrete, random daily demands are generated for a 100 day time horizon from a Poisson distribution with a mean of 1/10,000 of the county population. However, for the instances that follow, we assumed that the five most populous counties (Los Angeles, CA, Cook, IL, Harris, TX, Maricopa, AZ, and San Diego, CA) have candidate processing facilities that, if selected, are open seven days a week over the course of a 365 day time horizon. The 45 next most populous counties have candidate processing facilities that would only be open on weekdays. Thus, the 50 most populous U.S. counties are used in the data instances in this section. (Note that while the data set represents only 1.6 percent of the 3,109 counties in the contiguous U.S., it encompasses nearly 30% of the population of these 3,109 counties.)

Additionally, on the days a processing center is open, its capacity is (approximately) 1/5 of the total average daily demand in the system; clearly, on days a processing center is closed, its capacity is zero. All backlog levels were initialized to $v_j = 0$ and integer-valued travel times between nodes were calculated via the Great Circle Distance in the manner described in Section 3.4, which incorporates the assumption that demands are never processed on the day they originate.

Since the daily facility location costs are a function of the number of days per week that the processing facility is open, we redefine f_j for the cyclic-allocation model as the *average* daily fixed location cost of facility $j \in J$. Specifically, the *average* daily facility location cost for facility $j \in J$ which is closed c days per week is given by

$$f_j = z + w\rho_j - \frac{c}{7}\left(0.2z + 0.9w\rho_j\right) \tag{4.19}$$

As in Section 3.4, z represents the daily base fixed cost, ρ_j denotes the population of candidate facility j, and w is the population weight. Formula (4.19) states that a facility incurs only 80% of the daily base fixed cost and 10% of the weighted population based cost per day it is closed as compared to a day when it is open. In the computations that follow, we assumed a daily base fixed cost of $z = \$10{,}000$ and a weight of $w = \$0.0001$. For example, the county of San Diego, CA has a population of 3,095,313 people, which makes it the 5th most populous county in the U.S. Thus, if we choose to locate a facility in San Diego, CA, it will be open seven days per week and will incur a facility location cost of $z + w\rho_j$ each day of the week. Thus, its *average* daily facility location cost will be $\$10{,}000 + \$0.0001 * 3{,}095{,}313 - \frac{0}{7}(0.2 * \$10{,}000 + 0.9 * \$0.0001 * 3{,}095{,}313) = \$10{,}310$. Orange, CA, on the other hand, has a population of 3,010,232 people and is the 6th most populous county in the U.S. This means that a facility located in Orange, CA would not be open on the weekend and will incur $z + w\rho_j$ in facility

location costs each weekday and $0.8z + 0.1w\rho_j$ in facility location costs on each day of the weekend. Thus, it's *average* daily facility location cost would be $\$10{,}000 + \$0.0001 * 3{,}010{,}232 - \frac{2}{7}(0.2 * \$10{,}000 + 0.9 * \$0.0001 * 3{,}010{,}232) = \$9{,}652.$ The average daily facility location costs for the 50 largest nodes are represented in Figure 13.

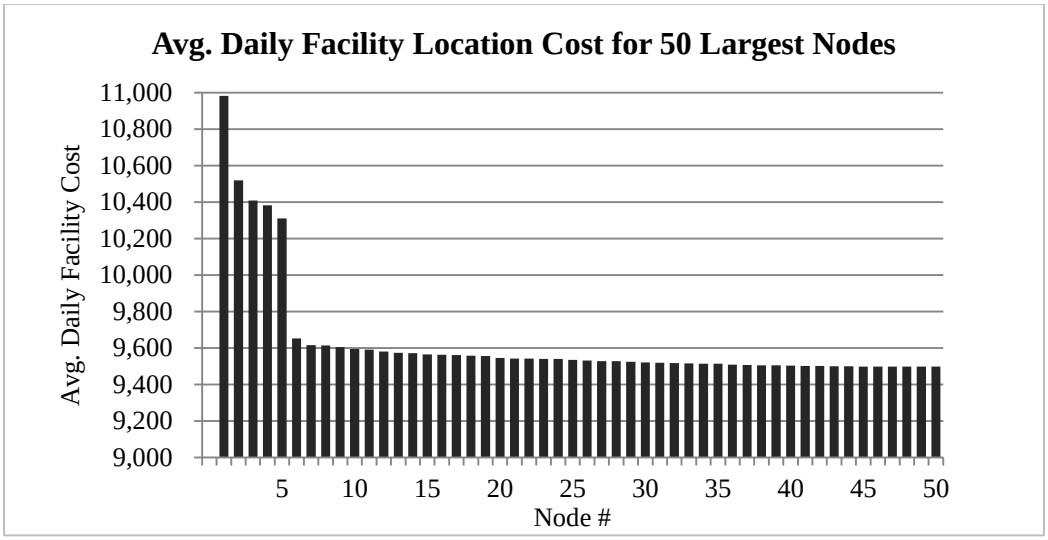

Figure 13: Average daily facility location costs for the cyclic allocation model, $z = 10{,}000$, $w = 0.0001$

Unless otherwise mentioned, costs $a = 2$ and $b = 20$ are used throughout this section and were chosen based on their ability to ensure that the number of items held in backlog from day to day does not grow without bound when six facilities are located. (Since we recognize it is often difficult to determine the backlogging cost, we discuss the implications of varying b in Section 4.5.3.)

4.5.1 Restricting the Number of Different Allocations

Table 18 presents five different cases that we studied to investigate how changing the maximum allowed number of unique allocations per demand site (i.e., n) affects the optimal

solution and the effect of imposing the weekend/weekday allocation assignments given by constraints (4.9) and (4.10).

Table 18: Effect of varying the value of n

Case	Description	Example
1	• A demand site $i \in I$ can be allocated to only 1 processing facility	• Demand from $i \in I$ is allocated to arrive at facility $j \in J$ every day of the week
2	• A demand site can be allocated to at most 2 processing facilities • Demands from $i \in I$ must be allocated to arrive at one facility Monday-Friday, but can be allocated to arrive at another facility on both Saturday and Sunday if desired	The above example as well as: • Demand from $i \in I$ is allocated to arrive at facility $j_1 \in J$ on Monday-Friday and at facility $j_2 \in J$ on Saturday and Sunday
3	• A demand site can be allocated to at most 2 processing facilities • Demands from $i \in I$ must be allocated to arrive at one facility Monday-Friday, but can be allocated to arrive at another facility on Saturday, Sunday, or both Saturday and Sunday, if desired	The above examples as well as: • Demand from $i \in I$ is allocated to arrive at facility $j_1 \in J$ on Monday-Saturday and at facility $j_2 \in J$ on Sunday • Demand from $i \in I$ is allocated to arrive at facility $j_1 \in J$ on Sunday-Friday and at facility $j_2 \in J$ on Saturday
4	• A demand site can be allocated to at most 2 processing facilities • Demands from $i \in I$ must be allocated to arrive at only one facility per day and can be assigned to arrive any day of the week	The above examples as well as: • Demand from $i \in I$ is allocated to arrive at facility $j_1 \in J$ on Sunday, Monday, Tuesday, and Friday and at facility $j_2 \in J$ on Wednesday, Thursday, and Saturday
5	• There is no restriction on the number of processing facilities to which a demand site can be allocated • Demands from $i \in I$ must be assigned to arrive at only one facility per day and can be assigned to arrive any day of the week	The above examples as well as: • Demand from $i \in I$ is allocated to arrive at facility $j_1 \in J$ on Sunday, $j_2 \in J$ on Monday, $j_3 \in J$ on Tuesday, $j_4 \in J$ on Thursday, and $j_5 \in J$ on Wednesday, Friday, and Saturday

Table 19 describes the details of each case as well as the results. The fourth column indicates whether constraint (4.10) is implemented as an inequality constraint or if it is changed to an equality constraint. Intuitively, constraints (4.9) and (4.10), as they are written, allow a demand site to assign demands to arrive at an additional facility on one, both, or neither of the weekend days. However, the equality form of (4.10) states that if a demand site assigns demands to arrive at an additional facility on the weekend, then the assignment must be made for *both* Saturday and Sunday.

Table 19: Effect of varying the value of n: computational results; $a = 2, b = 20$

Case	n	Include Constraints (4.9) and (4.10)	Constraint (4.10): Equality or Inequality	Number of Demand Sites that Allocate Demand to x Different Facilities			Total Cost	% Dec. in Cost from $n = 1$
				x =1	x =2	x =3		
1	1	No	---	50	---	---	$14,874,422	---
2	2	Yes	Equality	40	10	---	$14,347,714	3.54%
3	2	Yes	Inequality	27	23	---	$14,109,516	5.14%
4	2	No	---	23	27	---	$12,750,734	14.28%
5	7	No	---	26	20	4	$12,748,806	14.29%

Although the optimal solution locates six facilities regardless of the value of n in this problem instance, the locations of the six facilities vary slightly as n changes. In each of the instances, the optimal solution is to locate two facilities that are closed on weekends and four that are open seven days a week. Four facilities that are open for the entire week and one facility that is closed on weekends are the same regardless of the case: Los Angeles, CA; Cook, IL; Harris, TX; San Diego, CA; and Kings, NY. This represents locating facilities in the 1^{st}, 2^{nd}, 3^{rd}, 5^{th}, and 7^{th} most populous counties, respectively. The remaining facility is either Philadelphia, PA (Case 1); Palm Beach, FL (Cases 2 and 4); Hillsborough, FL (Case 3); or Broward, FL (Case 5). These are, respectively, the 21^{st}, 28^{th}, 30^{th}, and 18^{th} most populous counties.

Recall that the optimal solution to the problem with $n = 1$ is precisely the optimal solution to the corresponding IMCLP. The results show that a large cost improvement can be achieved by simply allowing each demand site to allocate its demand to two processing facilities, rather than one (Case 4 as opposed to Case 1). However, a large portion of this cost savings comes from allowing the allocations to take place any day of the week, rather than enforcing a weekend/weekday policy (Case 4 as opposed to Cases 2 and 3). Furthermore, allowing each demand site to allocate to three or more facilities provides only a marginal cost reduction (comparing Case 5 to Case 4).

4.5.2 Restricting the Maximum Allowed Travel Time

The optimal solution to (2.4), (3.5), (3.7), (3.8) and (4.1) - (4.10) may be such that some of the allocations assign a demand site $i \in I$ to a facility $j_4 \in J$ that is relatively close during the week, but assign i to another facility $j_5 \in J$ that is very far away on weekends. Thus, we utilize the constraints discussed in Section 4.4.3, to restrict the maximum allowed travel time. All of the results presented to address this situation correspond to Case 3 of the cyclic allocation model described in Table 18 and Table 19 (i.e., a weekday/weekend allocation policy given by (2.4), (3.5), (3.7), (3.8) and (4.1) - (4.10) for $n = 2$ when (4.10) is implemented as an inequality constraint).

Maximum Travel Time Constraints

When there is no restriction on the maximum travel days (i.e., $t_{max} \geq t^*$), the optimal allocation scheme is such that all demand sites are allocated to a processing facility that can be reached within 22 days; no demand-facility allocation pair requires $t^* = 29$ days of travel time in the optimal solution. Further restricting the travel time to 11 days (a 50% reduction in maximum travel days) results in only a 1.3% increase in cost and a 14% decrease in the demand-weighted average travel time. As a result, the solution that results from restricting the maximum travel time allowed to 11 days may be an attractive solution to facility managers who are concerned with the amount of time demands are in transit. However, any further restrictions in travel time come at a significant cost; restricting the maximum travel time to ten days results in a 6.3% increase in cost when compared to a maximum travel time of 11 days. The percent cost

increase due to restricting the maximum travel time can be seen graphically in Figure 14 and the details are given in Table 20.

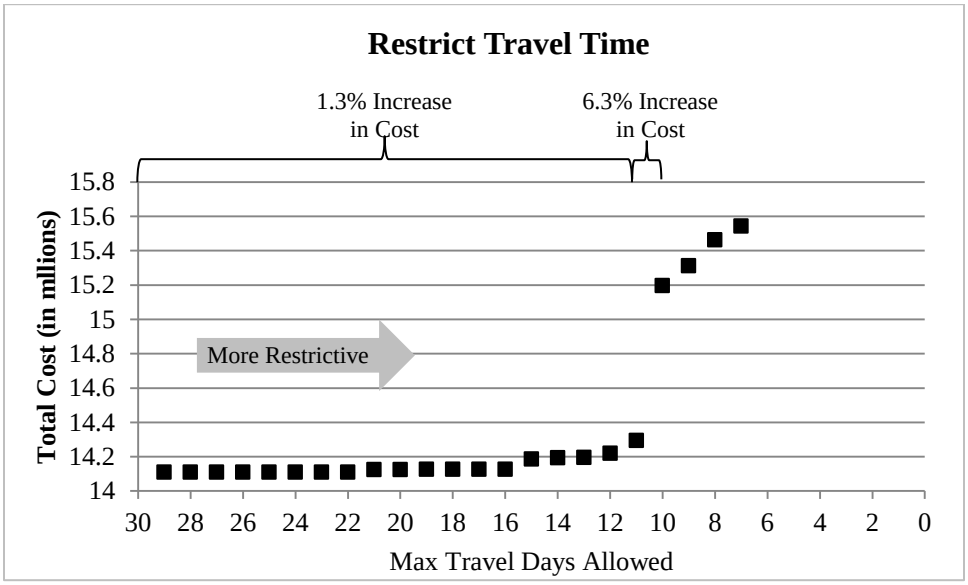

Figure 14: Cost effect of restricting travel time; $t_{\max} \in \{7,...,29\}$ days; $a = 2$, $b = 20$

Table 20: Effect of restricting travel time; $a = 2$, $b = 20$

t_{max}	# of Demand Sites Allocated to a Facility t_{max} Days Away	% Increase in Cost over Unrestricted Instance	# of Facilities Located	Cost Breakdown ($)			Total Cost ($)
				Location	Transportation	Backlog	
1	50	293.71%	50	48,176,400	1,836,932	5,537,400	55,550,732
2	24	143.61%	28	27,193,300	2,499,820	4,678,560	34,371,680
3	16	82.96%	18	17,583,400	3,222,198	5,009,860	25,815,458
4	20	48.34%	13	12,801,800	4,161,860	3,966,660	20,930,320
5	9	25.38%	10	9,948,800	4,611,892	3,129,980	17,690,672
6	7	13.63%	8	8,052,000	5,182,334	2,797,640	16,031,974
7	3	10.16%	8	8,048,600	4,999,604	2,494,720	15,542,924
8	2	9.59%	8	8,043,200	5,221,784	2,198,140	15,463,124
9	7	8.52%	8	8,048,600	5,474,212	1,789,000	15,311,812
10	1	7.70%	8	8,048,600	5,520,664	1,627,140	15,196,404
11	5	1.31%	7	7,087,000	6,084,212	1,122,700	14,293,912
12	2	0.79%	7	7,087,000	5,992,454	1,141,260	14,220,714
13	1	0.61%	7	7,087,000	6,017,134	1,091,000	14,195,134
14	3	0.60%	6	6,128,000	6,544,686	1,521,580	14,194,266
15	1	0.55%	7	7,087,600	6,027,608	1,071,860	14,187,068
16	2	0.12%	7	7,087,600	5,980,494	1,057,800	14,125,894
17-19	0	0.12%	7	7,087,600	5,980,494	1,057,800	14,125,894
20	2	0.11%	6	6,128,000	6,685,510	1,311,540	14,125,050
21	0	0.11%	6	6,128,000	6,685,510	1,311,540	14,125,050
22	1	0.00%	6	6,135,500	6,562,816	1,411,200	14,109,516
23-29	0	0.00%	6	6,135,500	6,562,816	1,411,200	14,109,516

In general, as the value of $t_{\max}$ decreases, the number of facilities located, location costs, and backlogging costs increase while the transportation costs decrease. We note that the table indicates that certain values of $t_{\max}$ (e.g., $t_{\max} \in \{13,14,15\}$) do not follow this general pattern. This is likely due to the particular demand instance used.

Additionally, we remind the reader that all demands must undergo one day of processing, which is included in the transportation cost, regardless of whether or not the demands are processed at the same location in which they originate. Thus, the transportation cost of \$1,836,932 when $t_{max} = 1$ is precisely $a = 2$ times the total number of demands generated during the planning horizon (918,466). Figure 15 displays the effect of restricting the travel time on the three cost components that comprise the objective function, as well as the optimal objective function value.

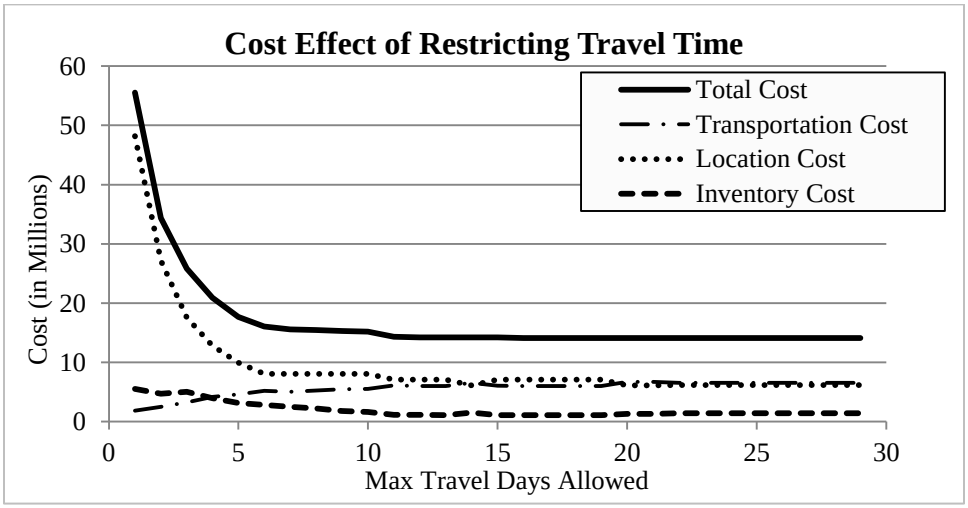

Figure 15: Cost effect of restricting travel time; $a = 2$, $b = 20$

Since the solution to the $t_{max} = 11$ constraint addition to Case 3 serves as a base case for many of the following analyses due to the reasons outlined at the beginning of this section, we display the corresponding optimal locations and arriving allocations in Figures 16-18. The optimal solution locates facilities at Los, Angeles, CA; San Diego, CA; Harris, TX; Orange, FL;

Cook, IL; Philadelphia, PA; and Queens, NY. The five most populous counties in the U.S. are identified by a circle; a processing facility located in one of these counties will be open every day of the week.

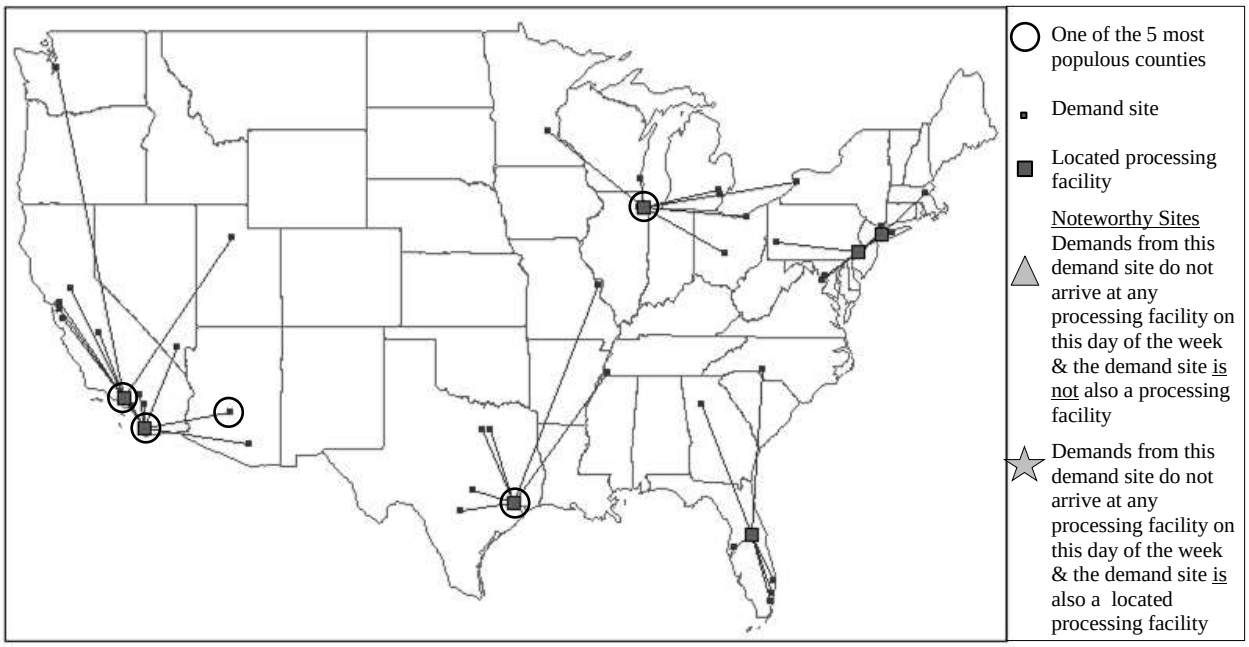

Figure 16: Optimal locations and allocations arriving on weekdays for Case 3; $a = 2$, $b = 20$, $t_{\max} = 11$

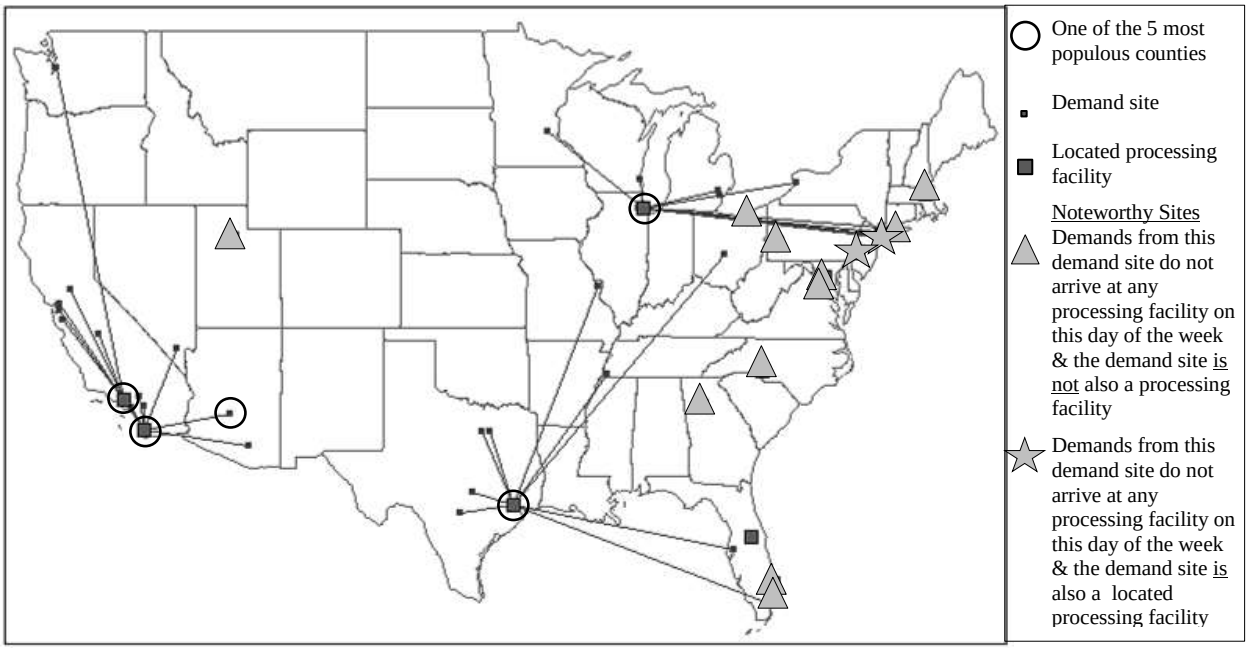

Figure 17: Optimal locations and allocations arriving on Saturdays for Case 3; $a = 2$, $b = 20$, $t_{\max} = 11$

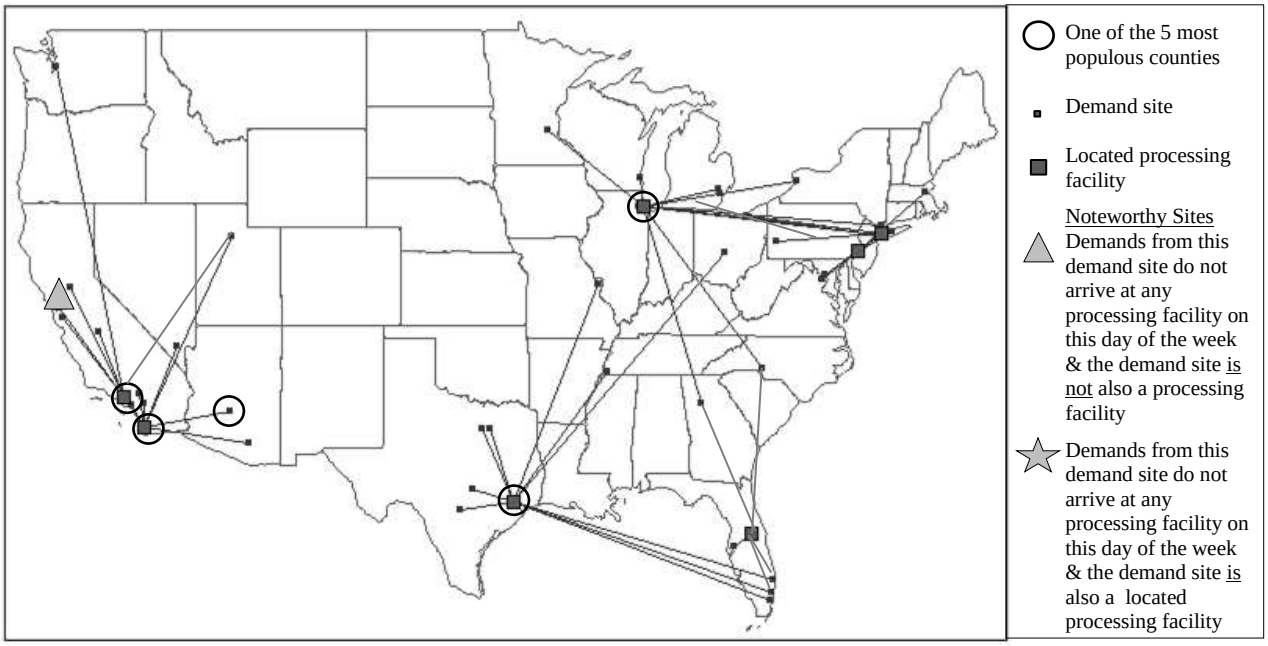

Figure 18: Optimal locations and allocations arriving on Sundays for Case 3; $a = 2$, $b = 20$, $t_{\max} = 11$

We note that the figures display where demand that is *arriving* at a processing facility on a particular day (i.e., weekday, Saturday, or Sunday) came from, rather than where the demand is being shipped on a particular day. For example, a connection from the Salt Lake, UT demand site to the San Diego, CA processing facility on Sunday's map (i.e., Figure 18) indicates that Salt Lake's demands arrive at the San Diego facility on Sundays. Since it takes eight days for the demand generated at Salt Lake to be ready for processing at San Diego, the demands must have been shipped from Salt Lake on a Saturday.

On some days, no processing facility receives demands from a particular demand-generating site. These demand sites are identified by triangles (if they are not also processing facilities) or stars (if they are processing facilities) in Figure 17 and Figure 18. For example, none of the processing facilities receive demands from Salt Lake, UT on Saturday; all of Salt Lake's demand arrives at processing facilities on either a weekday or on Sunday. This is indicated by a triangle. Additionally, none of the processing facilities receive demands from the

demand sites located in Philadelphia, PA and Queens, NY, which are both also processing facilities themselves. This is indicated by a star.

Many of the demand sites along the Atlantic coast do not allocate demands to arrive at processing facilities on Saturdays. This is because there is no nearby processing facility that is open on the weekends so it is cheaper to let the demands spend more time in transit and arrive at a facility on Sunday. This way, the demands can either be processed on Sunday (if the facility is open on weekends) or only have to incur one day (Sunday) of backlog cost rather than two (Saturday and Sunday). Demand sites and processing facilities whose demand does arrive at a processing facility on the day under consideration are denoted by a small square and a large square, respectively.

An example of the optimal cyclic allocations is depicted in Table 21 for the Salt Lake, UT demand site. Notice that demands leave the Salt Lake, UT demand site each day of the week, but only arrive at processing facilities on Sunday through Friday due to the extended travel time between Salt Lake, UT and San Diego, CA. This explains the triangle over Salt Lake, UT in Figure 17.

Table 21: Cyclic allocations for Salt Lake, UT

Day of Week Demand Leaves Salt Lake, UT	Processing Facility to which Demand is Allocated	Travel Time Between Salt Lake, UT and Processing Facility	Day of Week Demand Arrives at Processing Facility
Sunday	Los Angeles, CA	7 Days	Sunday
Monday	Los Angeles, CA	7 Days	Monday
Tuesday	Los Angeles, CA	7 Days	Tuesday
Wednesday	Los Angeles, CA	7 Days	Wednesday
Thursday	Los Angeles, CA	7 Days	Thursday
Friday	Los Angeles, CA	7 Days	Friday
Saturday	San Diego, CA	8 Days	Sunday

Restricting the Time Difference between Weekday and Weekend Allocations

While we can impose constraints (4.14) for any nonnegative integer value of m, the case of $m = 2$ is of particular interest since items that arrive at a facility that is closed on the weekends must wait in backlog for two days before they can be processed. This allows us to use the two days as (less-expensive) transportation days. Table 22 shows how varying the value of m affects the total cost, cost breakdown, and number of facilities located for the $t_{max} = 11$ restriction of Case 3. For all values of m, the model locates facilities in Cook, IL; Harris, TX; Los Angeles, CA; and San Diego, CA, which are among the five most populous counties and therefore are open seven days per week. In addition, all solutions locate a facility that is only open on weekdays in Philadelphia, PA. When $m = 0$, an additional weekday facility is located in Kings, NY; however, it is replaced with a facility in Queens, NY when $m \geq 1$. Finally, a seventh facility is located in Orange, FL when $m \geq 7$. The results show that allowing a slight increase in the travel time between weekday and weekend allocations results in a cost decrease that is nearly half the cost decrease that would be possible if the difference were unrestricted.

Table 22: Effect of restricting the time difference; $a = 2$, $b = 20$

m	Total Cost ($)	% Increase in Cost over Unrestricted Instance	# of Facilities Located	Cost Breakdown ($)		
				Location	Transportation	Backlog
0	14,773,536	3.35%	6	6,137,700	6,741,736	1,894,100
1-6	14,554,808	1.83%	6	6,135,600	6,771,708	1,647,500
7	14,373,362	0.56%	7	7,087,000	6,103,982	1,182,380
≥8	14,293,912	0.00%	7	7,087,000	6,084,212	1,122,700

Closest Assignment Constraints

In Table 23 we investigate the effect of allowing a demand site to be allocated to the closest ς facilities with respect to the difference in travel time between weekday and weekend assignments for the $t_{max} = 11$ addition to Case 3. The optimal solution in each instance had at least one assignment in which the travel time from a demand site to a processing facility was $t_{max} = 11$ and, except for the instance with $\varsigma = 1$, had a weekend assignment with an eight day longer travel time than the weekday assignment. All of the instances located seven facilities (but not necessarily the same facilities). Allowing demands to be assigned to the second closest facility results in a significant cost savings compared to the case in which all demands must be assigned to the closest located facility.

Table 23: Effect of imposing closest assignment constraints; $a = 2, b = 20$

ς	Total Cost ($)	% Increase in Cost over Unrestricted Instance	Cost Breakdown ($)			Facilities Located (Node #)
			Location	Transportation	Backlog	
1	15,950,538	11.59%	7,094,900	6,222,878	2,632,760	1, 2, 3, 5, 7, 14, 21
2	14,675,446	2.67%	7,082,200	5,723,666	1,869,580	1, 2, 3, 5, 21, 23, 30
{3,4}	14,343,048	0.34%	7,087,000	5,992,148	1,263,900	1, 2, 3, 5, 10, 21, 35
≥5	14,293,912	0.00%	7,087,000	6,084,212	1,122,700	1, 2, 3, 5, 10, 21, 35

4.5.3 Allocation Composition

To understand how many demand sites take advantage of the ability to allocate their demand to more than one processing facility, we graph the number of demand sites that adhere to each allocation method possible in Case 3 as described in Table 18, with the added $t_{max} = 11$ restriction. In particular, Figure 19 and Figure 20 display the number of demand sites that

1) only allocate demand to one processing facility;

2) allocate demand to two processing facilities, with demand arriving at a different facility on Saturday than Sunday; and
3) allocate demand to two processing facilities, with demand arriving at the same facility on Saturday and Sunday

as the cost of holding items in backlog varies.

Demand sites that allocate according to the second or third method utilize the flexibility provided by cyclic allocations. Thus, we quantify the diversity of demand allocations by giving each instance an allocation score that accounts for the number of demand sites that are assigned to each of the three possible allocation methods. Demand sites contribute a value of 1, 2, or 3, depending on whether they allocate demand according to the first, second, or third method, respectively. For example, an allocation score of 67 is given to the optimal solution when $b = 10$ since it assigns 36 demand sites to only one processing facility, 11 to two processing facilities with demand arriving at a different facility on Saturday than Sunday, and 3 to two processing facilities with demand arriving at the same facility on Saturday and Sunday (i.e., $36 * 1 + 11 * 2 + 3 * 3 = 67$). Since there are 50 demand sites, the lowest possible allocation score is 50 and corresponds to a solution that is no different from what would be obtained using the original IMCLP model (i.e., all demand sites are assigned to only one processing facility). The highest possible allocation score is 150, though the highest realized score in Figure 19 is only 72.

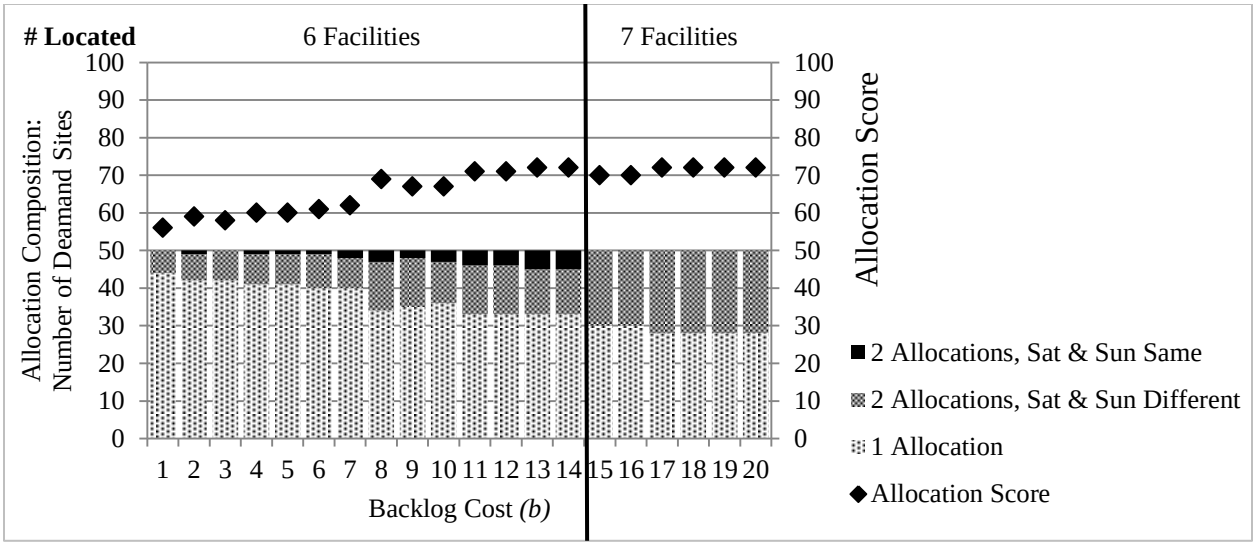

Figure 19: Types of allocations in the optimal solution as the backlog cost varies; $a = 2$

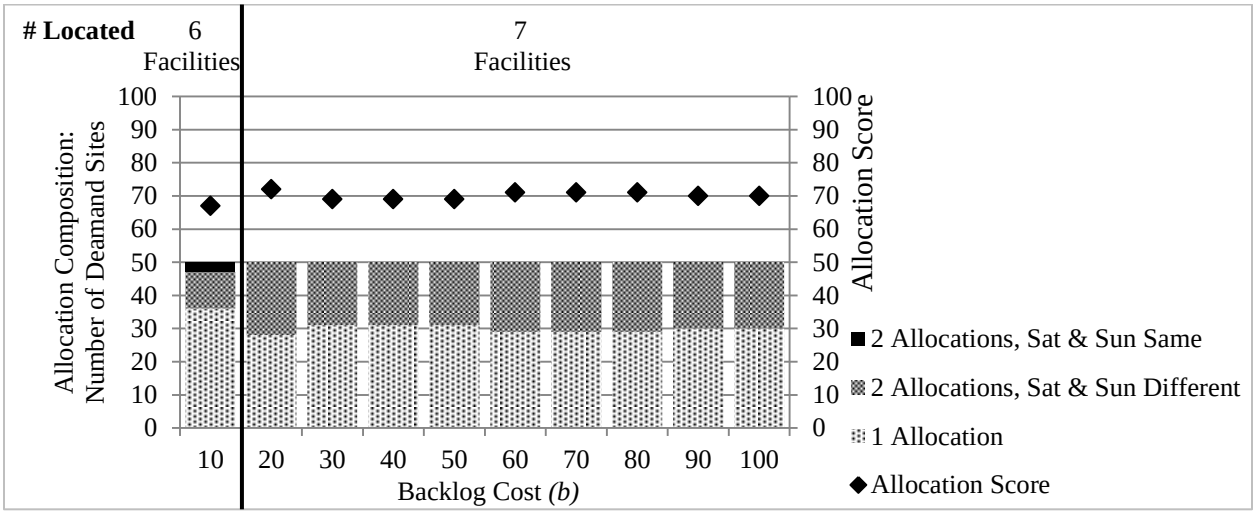

Figure 20: Types of allocations in the optimal solution as the backlog cost varies; $a = 2$

The allocation score generally increases as the backlog cost increases when six facilities are located. However, the allocation score remains relatively constant for b values that locate seven facilities in the optimal solution. Furthermore, the optimal solution to the $b \geq 15$ instances displayed in Figures 19 and 20 assigns approximately 30 demand sites to only one processing facility and 20 demand sites to two processing facilities, with demand arriving at a different processing facility on Saturday than Sunday; in this case, the allocation for six days (weekdays and one weekend day) is the same.

We further quantify the difference between the optimal cyclic solution and the optimal IMCLP solution by comparing their costs in Figure 21. The figure shows that the flexibility allowed by cyclic allocations results in increased cost savings as the cost of holding backlogged items increases. In particular, the cyclic allocations result in a 3.9% decrease in cost in our base case instance (i.e., $b = 20$) when compared to the corresponding optimal IMCLP solution. If the backlog cost were doubled (i.e., $b = 40$) we would see a 7.85% decrease in cost. At a backlog cost of $b = 100$, an 11.27% decrease in cost is achieved.

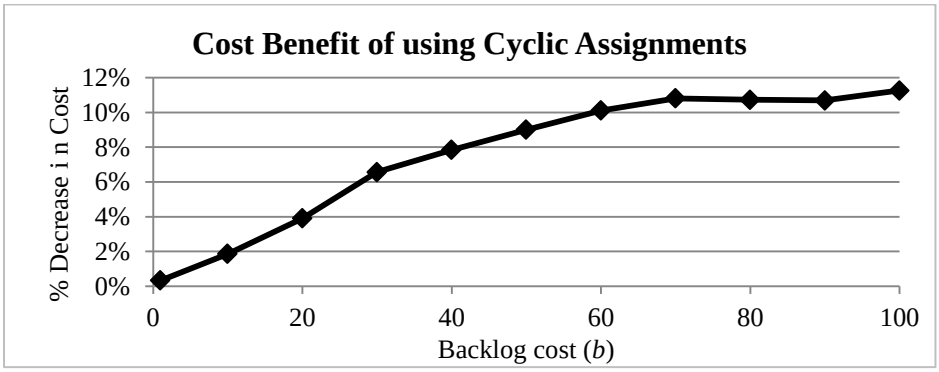

Figure 21: % decrease in cost obtained by using the cyclic allocation model as compared to the IMCLP model; $a = 2$

4.5.4 Solution Robustness

In facility location modeling it is extremely important to correctly identify which processing facilities to locate, since it is often very difficult and costly to change the locations once they are established. While it is also important to correctly identify the allocations, these can be adapted more easily than the locations. This motivates our desire to determine how robust the optimal solution is to the particular demand instance used.

To test the solution robustness, we consider ten demand instances. In each of these instances, Case 3 with the added $t_{max} = 11$ restriction is used. We first solve Instance 1 to

optimality and then compare the solution to those of the remaining instances under the following three conditions:

(1) forcing the Instance 1 allocations and locations into the remaining instances,

(2) forcing only the Instance 1 locations into the remaining instances, and

(3) solving the remaining instances to optimality without considering the Instance 1 solution.

The results indicated that each of the ten instances (under condition (3)) identified the same seven processing facilities to locate in the optimal solution. Although the optimal allocations differed slightly among the instances, these variations always resulted in less than a 0.1% increase in total cost when the Instance 1 optimal locations and allocations were imposed on the other instances (i.e., scenario (1)); the average increase in cost was 0.028%. This is particularly noteworthy since the transportation and backlogging costs are intimately linked to the allocation decisions and account for roughly 50% of the total cost. The details are displayed in Table 24.

Table 24: % Increase in Cost from using the Optimal Locations and Allocations of Instance 1 instead of the Optimal Solution for the Particular Instance; $a = 2, b = 20$

Instance	Total	Location	Transportation	Backlog
1	**0.000%**	**0.000%**	**0.000%**	**0.000%**
2	0.001%	**0.000%**	-0.025%	0.198%
3	0.060%	**0.000%**	-0.147%	1.564%
4	0.022%	**0.000%**	-0.646%	4.021%
5	0.055%	**0.000%**	-0.569%	4.008%
6	0.029%	**0.000%**	-0.543%	3.449%
7	0.008%	**0.000%**	-0.637%	3.721%
8	0.000%	**0.000%**	**0.000%**	**0.000%**
9	0.054%	**0.000%**	-0.636%	4.451%
10	0.052%	**0.000%**	-0.210%	1.870%
Average	0.028%	**0.000%**	-0.341%	2.328%

4.6 Chapter Summary

We introduced a variant of the IMCLP that allows for a cyclic allocation approach to assigning demand sites to processing facilities. This enables the model to develop a day-of-the-week allocation policy that considers deterministic fluctuations in the daily processing capacity levels of the facilities or stochastic fluctuations in the average demand by day-of-the-week. For example, our model is able to adjust the allocation policy to account for some facilities being closed on weekends. In the spirit of the IMCLP model developed in Chapter 3, the model presented in this chapter also allows disaggregated daily demand parameters to serve as inputs directly into the model and mitigates the hard capacity constraints that are typically found in capacitated facility location models by allowing unprocessed demand to be held in backlog and processed at a later date. Through computational studies, we showed that incorporating this cyclic allocation approach into the model can result in a significant cost savings as compared to models that do not incorporate this approach, such as the IMCLP. Furthermore, we note that the model can readily be restructured to deal with other cycles, such as annual cycles with demand assignments that either change seasonally or monthly in response to time-varying demand patterns.

We also presented multiple constraints that can be added to the model formulation to achieve various operational goals. These include (1) restricting the number of processing facilities to which a demand site can be assigned, (2) introducing a weekday/weekend allocation policy, and (3) bounding the difference between the travel time of a weekday assignment and a weekend assignment. Our results suggest that a significant cost reduction can be achieved by implementing the cyclic allocation model as compared to capacitated facility location models

that require a static allocation policy. Much of this benefit can be achieved even when additional constraints of the form outlined above are imposed.

CHAPTER 5: Chance Constrained IMCLP with Uncertain Demand

5.1 Motivation

In reality, a facility manager faces many sources of operational uncertainty, including uncertainty related to demand, machine processing capacity, personnel availability, purchasing/selling prices, and shipment travel times. The many sources of uncertainty make it impossible to model all of the uncertainty simultaneously due to the resulting model intractability, computational limitations, or data requirements. Thus, in this chapter, we focus solely on stochasticity that arises in demand.

We note that while the deterministic IMCLP incorporates daily fluctuations in demand over an extended time horizon, it is not explicitly formulated as a stochastic optimization problem. Thus, in this chapter, we introduce a stochastic variant of the IMCLP with uncertain demand. In addition to assessing a penalty cost associated with each day an item spends in backlog (as is done in the IMCLP and cyclic allocation model), we use three different types of chance constraints to restrict the number of demands that are backlogged to a predetermined threshold. The first approach incorporates joint chance constraints that ensure the probability of any processing facility having a backlog level above the threshold on any day of the planning horizon is sufficiently small. We also model individual chance constraints on the amount of backlog at each facility each day, as well as a hybrid approach which accounts for the probability that each individual processing facility will exceed the stated maximum backlog level on any day of the planning horizon. We then present mixed-integer programming reformulations of the

chance constraints that incorporate a finite number of scenarios from a given known demand distribution. (We refer the interested reader to Birge and Louveaux (2011) for an introduction to chance constraints.)

The resulting models are solved using two different decomposition schemes and their performance is compared to that of a generic solver. The first decomposes the problem into two stages: the long-term (i.e., location and allocation) decisions are determined in the first stage while the daily (i.e., processing and backlog) decisions are determined in the second stage. Benders decomposition is used to solve the resulting formulation. The second decomposition scheme capitalizes on the problem structure by utilizing a three-stage approach. In particular, given a feasible first-stage location and allocation solution, we can readily determine the optimal second-stage processing and backlog decisions as well as the third-stage auxiliary variables that verify whether the chance constraints are satisfied by inspection. If the chance constraints are violated, a corresponding cut is added to the first-stage problem.

5.2 Joint Chance Constrained Formulation

As mentioned, we will incorporate three types of chance constraints into the stochastic IMCLP: joint chance constraints, individual chance constraints, and a hybrid approach. The majority of this chapter will focus on the joint chance constrained model, which we formulate in this section. The individual chance constraints and the hybrid approach will be introduced in Section 5.4.

In each of the formulations, we incorporate demand stochasticity into the IMCLP using a scenario based approach in which a scenario corresponds to a realization of the daily demand for every day within the time horizon. For example, if the time horizon is one year, the demand

generated on each day of the year at each demand site is collectively referred to as a scenario. We let Ω represent the set of all possible demand scenarios, and p^{ω} represent the probability of a particular scenario $\omega \in \Omega$. It is assumed that $\sum_{\omega \in \Omega} p^{\omega} = 1$. Additionally, the parameter $\tilde{h}_{id}$ corresponds to the random daily demand generated at demand point $i \in I$ on day $d \in D$. While $\tilde{h}_{id}$ is a random parameter, h_{id}^{ω} represents a realization of the random demand $\tilde{h}_{id}$ in scenario $\omega \in \Omega$ generated at $i \in I$ on day $d \in D$ and $\xi(\omega) \in \mathbb{N}_0^{|I| \times |D|}$ is a vector containing the demand realizations $h_{id}^{\omega}, \forall i \in I, d \in D$ corresponding to scenario $\omega \in \Omega$.

Since we assume that the facility location and assignment decisions must remain constant throughout the planning horizon, variables X_j and Y_{ij}, which represent these decisions, are scenario independent in the stochastic IMCLP. For fixed location and allocation decisions, the processing and backlog variables become apparent once the scenario is observed. Thus, we let W_{jd}^{ω} and V_{jd}^{ω} be the corresponding scenario dependent variables (or equivalently, recourse variables). In summary, the decision variables for the IMCLP model with stochastic demand are:

Decision Variables

$$X_j = \begin{cases} 1 & \text{If we locate at facility } j \in J \\ 0 & \text{Otherwise} \end{cases}$$

$$Y_{ij} = \begin{cases} 1 & \text{If we assign demand site } i \in I \text{ to facility } j \in J \\ 0 & \text{Otherwise} \end{cases}$$

V_{jd}^{ω} Backlog at site $j \in J$ at the beginning of day $d \in \{t^* + 1, \dots, |D| + 1\}$ in scenario $\omega \in \Omega$

W_{jd}^{ω} Number of items processed at site $j \in J$ on day $d \in \{t^* + 1, \dots, |D|\}$ in scenario $\omega \in \Omega$

The specific aim of incorporating the joint chance constraint into the stochastic IMCLP is to minimize the total facility, expected transportation, and expected backlogging costs, while

ensuring the probability that the backlog at any facility never exceeds a predetermined level of θ is bounded from below by a given reliability level $1-\tau$. The joint chance constraint is

$$\mathbb{P}\left(\tilde{V}_{jd} \leq \theta;\ \forall j \in J, d \in \{t^*+1, \ldots, |D|+1\}\right) \geq 1-\tau \tag{5.1}$$

where $\mathbb{P}(\cdot)$ denotes the probability of event $\cdot$ occurring and $\tilde{V}_{jd}$ represents the stochastic counterpart of the random outcome V_{jd}^{ω}, $\forall \omega \in \Omega$.

5.2.1 Single Stage Formulation with Probabilistic Constraint

With this additional notation, the joint chance constrained IMCLP with stochastic demands is formulated as follows:

[JCC]:

$$Min_{\boldsymbol{X},\boldsymbol{Y},\boldsymbol{V},\boldsymbol{W}} \quad (|D|-t^*)\sum_{j\in J} f_j X_j + a\sum_{i\in I}\sum_{d=t^*+1}^{|D|}\sum_{j\in J}\mathbb{E}_{\xi}\left[\tilde{h}_{id}\right]t_{ij}Y_{ij}$$

$$+b\,\mathbb{E}_{\xi}\left[\sum_{d=t^*+2}^{|D|+1}\sum_{j\in J}\tilde{V}_{jd}\right] \tag{5.2}$$

Subject to

$$\sum_{j\in J} Y_{ij} = 1 \qquad \forall i \in I \tag{2.2}$$

$$Y_{ij} \leq X_j \qquad \forall i \in I; j \in J \tag{2.3}$$

$$\mathbb{P}\left(\tilde{V}_{jd} \leq \theta;\ \forall j \in J, d \in \{t^*+1, \ldots, |D|+1\}\right) \geq 1-\tau \tag{5.1}$$

$$V_{j,d+1}^{\omega} - V_{jd}^{\omega} + W_{jd}^{\omega} = \sum_{i\in I} h_{i,d-t_{ij}}^{\omega} Y_{ij} \qquad \forall j \in J; d \in \{t^*+1, \ldots, |D|\}; \omega \in \Omega \tag{5.3}$$

$$W_{jd}^{\omega} \leq k_j X_j \qquad \forall j \in J; d \in \{t^*+1, \ldots, |D|\}; \omega \in \Omega \tag{5.4}$$

$$V_{j,t^*+1}^{\omega} = v_j X_j \qquad \forall j \in J; \omega \in \Omega \tag{5.5}$$

$$V_{jd}^{\omega} \geq 0 \qquad \forall j \in J; d \in \{t^*+1, \ldots, |D|+1\}; \omega \in \Omega \tag{5.6}$$

$$W_{jd}^{\omega} \geq 0 \qquad \forall j \in J; d \in \{t^*+1, \ldots, |D|\}; \omega \in \Omega \tag{5.7}$$

$$X_j \in \{0,1\} \qquad \forall j \in J \tag{2.4}$$

$$Y_{ij} \in \{0,1\} \qquad \forall i \in I; j \in J \tag{3.6}$$

Where

$$\mathbb{E}_{\xi}\left[\tilde{h}_{id}\right] = \sum_{\omega \in \Omega} p^{\omega} h_{id}^{\omega} \tag{5.8}$$

and

$$\mathbb{E}_{\xi}\left[\sum_{d=t^*+2}^{|D|+1} \sum_{j \in J} \tilde{V}_{jd}\right] = \sum_{\omega \in \Omega} p^{\omega} \sum_{d=t^*+2}^{|D|+1} \sum_{j \in J} V_{jd}^{\omega}. \tag{5.9}$$

The objective function (5.2) minimizes the facility location costs, as well as the expected transportation and backlogging costs over the entire time horizon (i.e., days $t^* + 1$ to $|D|$). We note that the expected backlog at the beginning of day $|D| + 1$ is precisely the amount of backlog at the *end* of day $|D|$, and as such, including this term in the objective function allows the model to capture the backlog level at the end of the planning horizon.

Although the backlog levels will be determined to meet the desired service levels specified in the joint chance constraint (5.1) our objective function also explicitly captures the backlogging costs. This allows the model to differentiate between solutions that contain different amounts of backlog. For example, our formulation would give preference to a solution in which no items are held in backlog over a solution that meets the service level requirements but holds θ items in backlog at each facility every day, given that the two solutions have the same location and allocation cost.

As before, constraints (2.2) and (3.6) enforce single sourcing constraints on the demand sites while constraints (2.3) ensure demand sites are only assigned to located facilities. Constraint (5.1) is the joint chance constraint that ensures the probability that the backlog at any facility never exceeds a predetermined level of θ is bounded from below by a given reliability level $1 - \tau$. Constraints (5.3) are flow balance constraints to update the amount of demand held

in backlog from day to day. They also serve, along with constraints (5.4), as capacity constraints. Specifically, constraints (5.3) ensure the amount of demand processed each day does not exceed the amount available for processing while constraints (5.4) limit the maximum amount of demand processed each day to the daily processing capacities of the facilities. Constraints (5.5) state that a located facility $j \in J$ will have an initial backlog of v_j and a facility that is not located will have an initial backlog level of zero. Constraints (2.4), (3.6), (5.6), and (5.7) are standard non-negativity and binary constraints.

5.2.2 MIP Reformulation

Since formulation JCC is nonlinear due to the joint chance constraint (5.1), we present a mixed integer linear reformulation that approximates JCC by using a finite subset $\Omega' \subseteq \Omega$ of all possible demand scenarios. Furthermore, we let $\omega' \in \Omega'$ deonote elements of Ω' and define $h_{id}^{\omega'}$, $V_{jd}^{\omega'}$, $W_{jd}^{\omega'}$, $\xi(\omega')$, and $p^{\omega'}$ accordingly (with $\sum_{\omega' \in \Omega'} p^{\omega'} = 1$ and $\xi(\omega')$ often referred to as ξ' for readability. With this notation, we introduce decision variables

$$Z^{\omega'} = \begin{cases} 1 & \text{If the backlog at any facility exceeds } \theta \text{ on any day in scenario } \omega' \in \Omega' \\ 0 & \text{Otherwise} \end{cases}$$

and replace (5.1) with

$$V_{jd}^{\omega'} - M_1 Z^{\omega'} \leq \theta \qquad \forall\, j \in J;\ d \in \{t^* + 1, \dots, |D| + 1\}; \omega' \in \Omega' \quad (5.10)$$

$$\sum_{\omega' \in \Omega'} p^{\omega'} Z^{\omega'} \leq \tau \quad (5.11)$$

$$Z^{\omega'} \in \{0,1\} \qquad \forall \omega' \in \Omega' \quad (5.12)$$

Constraints (5.10) define the binary service level variables where M_1 is a sufficiently large number such that the constraint is satisfied by letting $Z^{\omega'} = 1$ when $V_{jd}^{\omega'} > \theta$ for some facility j

on some day d in some scenario $\omega' \in \Omega'$. Constraint (5.11) guarantees that the probability of violating the joint chance constraint is no more than τ. Note that we can write constraint (5.11) as

$$\sum_{\omega' \in \Omega'} Z^{\omega'} \leq \lfloor \tau |\Omega'| \rfloor \tag{5.13}$$

if each scenario $\omega' \in \Omega'$ occurs with equal probability. Finally, constraints (5.12) are standard binary constraints. This results in the following mixed integer linear reformulation of JCC:

[MIP-JCC]:

$$\ddot{V} = Min_{\boldsymbol{X},\boldsymbol{Y},\boldsymbol{V},\boldsymbol{W},\boldsymbol{Z}} \ (|D| - t^*) \sum_{j \in J} f_j X_j + a \sum_{i \in I} \sum_{d=t^*+1}^{|D|} \sum_{j \in J} \mathbb{E}_{\xi'}\left[\tilde{h}_{id}\right] t_{ij} Y_{ij}$$

$$+ b \ \mathbb{E}_{\xi'}\left[\sum_{d=t^*+2}^{|D|+1} \sum_{j \in J} \tilde{V}_{jd}\right] \tag{5.14}$$

Subject to

$$\sum_{j \in J} Y_{ij} = 1 \qquad \forall i \in I \tag{2.2}$$

$$Y_{ij} \leq X_j \qquad \forall i \in I; j \in J \tag{2.3}$$

$$V_{jd}^{\omega'} - M_1 Z^{\omega'} \leq \theta \qquad \forall j \in J; \ d \in \{t^* + 1, \dots, |D| + 1\}; \omega' \in \Omega' \tag{5.10}$$

$$\sum_{\omega' \in \Omega'} p^{\omega'} Z^{\omega'} \leq \tau \tag{5.11}$$

$$V_{j,d+1}^{\omega'} - V_{jd}^{\omega'} + W_{jd}^{\omega'} = \sum_{i \in I} h_{i,d-t_{ij}}^{\omega'} Y_{ij} \qquad \forall j \in J; d \in \{t^* + 1, \dots, |D|\}; \omega' \in \Omega' \tag{5.15}$$

$$W_{jd}^{\omega'} \leq k_j X_j \qquad \forall j \in J; d \in \{t^* + 1, \dots, |D|\}; \omega' \in \Omega' \tag{5.16}$$

$$V_{j,t^*+1}^{\omega'} = v_j X_j \qquad \forall j \in J; \omega' \in \Omega' \tag{5.17}$$

$$V_{jd}^{\omega'} \geq 0 \qquad \forall j \in J; d \in \{t^* + 1, \dots, |D| + 1\}; \omega' \in \Omega' \tag{5.18}$$

$$W_{jd}^{\omega'} \geq 0 \qquad \forall j \in J; d \in \{t^* + 1, \dots, |D|\}; \omega' \in \Omega' \tag{5.19}$$

$$X_j \in \{0,1\} \qquad \forall j \in J \tag{2.4}$$

$$Y_{ij} \in \{0,1\} \qquad \forall i \in I; j \in J \tag{3.6}$$

$$Z^{\omega'} \in \{0,1\} \qquad \forall \omega' \in \Omega' \tag{5.12}$$

Where

$$\mathbb{E}_{\tilde{\xi}'}\left[\tilde{h}_{id}\right] = \sum_{\omega' \in \Omega'} p^{\omega'} h_{id}^{\omega'} \tag{5.20}$$

and

$$\mathbb{E}_{\tilde{\xi}'}[\sum_{d=t^*+2}^{|D|+1} \sum_{j \in J} \tilde{V}_{jd}] = \sum_{\omega' \in \Omega'} p^{\omega'} \sum_{d=t^*+2}^{|D|+1} \sum_{j \in J} V_{jd}^{\omega'}. \tag{5.21}$$

5.3 Solution Methods

As is evidenced in Section 5.5.2, solving problem instances with even a small number of scenarios takes an exorbitant amount of time without employing any specialized solution techniques. Therefore, in this section, we propose two-stage and three-stage decomposition schemes to solve the mixed-integer programming reformulation of the joint chance-constrained program with cutting-plane algorithms. As a means of identifying the decomposed problems, we will use a labeling scheme of the form M#S# to identify the method and the problem stage. For example, M1S2 refers to the second stage problem of the first decomposition scheme, whereas M2S3 refers to the third stage problem of the second decomposition scheme.

5.3.1 Two-Stage Benders Decomposition Approach

The Benders decomposition [Benders, 1962] approach is one of the most effective techniques for solving large-scale linear programming and mixed-integer programming models that exhibit decomposable structures. The approach specifies two sets of variables, initial decision variables and continuous recourse variables. At each iteration, the algorithm solves a relaxed master problem at the first stage and passes a fixed first-stage optimal decision to the second stage where recourse decisions are optimal for individual scenarios. Duality theorems are then used to derive feasibility and optimality cuts, which affect the first-stage solution in

subsequent iterations. The algorithm terminates when all optimality conditions are satisfied and no additional cuts are generated. We refer the interested reader to Birge and Louveaux (2011) for a more detailed description of the decomposition technique.

Since Benders first introduced his decomposition scheme in 1962 in the context of solving a general linear program, significant research has emerged that addresses algorithmic improvements and modifications to accelerate the convergence of Benders decomposition method as applied to facility location problems. Balinski and Wolfe (1963), Davis and Ray (1969), and Geoffrion and Graves (1974) are among the earliest papers that apply Benders decomposition to a facility location problem. Magnanti and Wong (1981) address the choice of effective optimality cuts and have shown that the use of stronger cuts can reduce the number of iterations, thereby improving the convergence of the algorithm. Wentges (1996) proposes a procedure specifically for strengthening Benders cuts in the context of the capacitated facility location problem. A review by Magnanti and Wong (1990) presents an overview of decomposition methods for facility location problems. The review specifically addresses methods of accelerating Benders decomposition and model selection criterion for facility location problems.

In this section we develop a two-stage cutting plane algorithm based on the Benders decomposition scheme to solve MIP-JCC. We place the long term decisions (i.e., location, allocation) in the first stage and the daily decisions (i.e., processing, backlog) in the second stage. For simplicity, we will denote ordered sets of variables using bold face type in the following manner:[1]

[1] We note that additional similar sets are created for variables introduced later in the chapter. For readability, we adopt the convention that boldface type indicates such vectors and matrices, but we do not define every set in the text. Instead, a detailed description of each ordered set can be found in the LIST OF SYMBOLS beginning on page xvi.

$\boldsymbol{X}$: a $1 \times |J|$ vector whose j^{th} element is X_j

$\boldsymbol{Y}$: a $|I| \times |J|$ matrix whose $(i,j)^{\text{th}}$ entry is Y_{ij}

$\boldsymbol{V^{\omega' j}}$: a $1 \times (|D| + 1 - t^*)$ vector whose $(d - t^*)^{\text{th}}$ element is $V_{jd}^{\omega'}$

$\boldsymbol{V}$: a $|J| \times (|D| + 1 - t^*) \times |\Omega'|$ matrix whose $(j, d - t^*, \omega)^{\text{th}}$ element is $V_{jd}^{\omega'}$

$\boldsymbol{W^{\omega' j}}$: a $1 \times (|D| - t^*)$ matrix whose $(d - t^*)^{\text{th}}$ element is $W_{jd}^{\omega'}$

$\boldsymbol{W}$: a $|J| \times (|D| - t^*) \times |\Omega'|$ matrix whose $(j, d - t^*, \omega')^{\text{th}}$ element is $W_{jd}^{\omega'}$

$\boldsymbol{Z}$: a $1 \times |\Omega|$ vector whose ω'^{th} element is $Z^{\omega'}$

At the first stage we keep binary variables $\boldsymbol{X}$, $\boldsymbol{Y}$, and $\boldsymbol{Z}$, and constraints (2.2), (2.3), and (5.11). Without loss of generality, we add the constraint

$$\sum_{j \in J} X_j \geq 1 \tag{5.22}$$

to ensure at least one facility is located (this will become useful in the solution algorithms). Alternatively, we could add the more restrictive constraint

$$\sum_{j \in J} k_j X_j \geq \frac{\sum_{i \in I} \sum_{d=t+1}^{|D|} \mathbb{E}_{\xi'}[\tilde{h}_{id}]}{|D| - t^*} \tag{5.23}$$

in place of constraint (5.22) to ensure that the cumulative capacity of the located facilities is sufficient to handle the expected cumulative demand. We consider the relaxed master problem as:

[M1S1]:

$$Min_{\boldsymbol{X},\boldsymbol{Y},\boldsymbol{Z}} \; (|D| - t^*) \sum_{j \in J} (f_j X_j + a \sum_{i \in I} \sum_{d=t^*+1}^{|D|} \mathbb{E}_{\xi'}\left[\tilde{h}_{id}\right] t_{ij} Y_{ij} + \sum_{\omega' \in \Omega'} p^{\omega'} \bar{V}^{\omega' j}(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z})) \tag{5.24}$$

Subject to

$$\sum_{j \in J} Y_{ij} = 1 \qquad \forall i \in I \tag{2.2}$$

$$Y_{ij} - X_j \leq 0 \qquad \forall i \in I; j \in J \tag{2.3}$$

$$\sum_{j\in J} X_j \geq 1 \tag{5.22}$$

$$\sum_{\omega'\in\Omega'} p^{\omega'} Z^{\omega'} \leq \tau \tag{5.11}$$

$$X_j \in \{0,1\} \qquad \forall j \in J \tag{2.4}$$

$$Y_{ij} \in \{0,1\} \qquad \forall i \in I; j \in J \tag{3.6}$$

$$Z^{\omega} \in \{0,1\} \qquad \forall \omega \in \Omega \tag{5.12}$$

where for each scenario $\omega' \in \Omega'$ and facility $j \in J$, we optimize the following subproblem:

[M1S2(ω', j)]:

$$\bar{V}^{\omega' j}(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}) = Min_{\boldsymbol{V}^{\boldsymbol{\omega' j}}, \boldsymbol{W}^{\omega' j}} \sum_{d=t^*+2}^{|D|+1} V_{jd}^{\omega'} \tag{5.25}$$

Subject to

$$V_{j,d+1}^{\omega'} - V_{jd}^{\omega'} + W_{jd}^{\omega'} = \sum_{i\in I} h_{i,d-t_{ij}}^{\omega'} Y_{ij} \qquad \forall d \in \{t^*+1, \dots, |D|\} \tag{5.26}$$

$$W_{jd}^{\omega'} \leq k_j X_j \qquad \forall d \in \{t^*+1, \dots, |D|\} \tag{5.27}$$

$$V_{j,t^*+1}^{\omega'} = v_j X_j \tag{5.28}$$

$$V_{jd}^{\omega'} \leq \theta + M_1 Z^{\omega} \qquad \forall d \in \{t^*+1, \dots, |D|+1\} \tag{5.29}$$

$$V_{jd}^{\omega'} \geq 0 \qquad \forall d \in \{t^*+1, \dots, |D|+1\} \tag{5.30}$$

$$W_{jd}^{\omega'} \geq 0 \qquad \forall d \in \{t^*+1, \dots, |D| \tag{5.31}$$

We associate dual variables $\gamma_d^{\omega' j}, \pi_d^{\omega' j}$, $\eta^{\omega' j}$, and $\mu_d^{\omega' j}$ with constraints (5.26), (5.27), (5.28), and (5.29), respectively. For each $\omega' \in \Omega'$ and $j \in J$, the dual of M1S2(ω', j) is:

[D-M1S2(ω', j)]:

$$\bar{V}_{Dual}^{\omega' j}(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}) = Max_{\boldsymbol{\gamma}^{\boldsymbol{\omega' j}}, \boldsymbol{\pi}^{\boldsymbol{\omega' j}}, \eta^{\omega' j}, \boldsymbol{\mu}^{\boldsymbol{\omega' j}}} \sum_{d=t^*+1}^{|D|} \sum_{i\in I} h_{i,d-t_{ij}}^{\omega'} Y_{ij} \gamma_d^{\omega' j} + \sum_{d=t^*+1}^{|D|} k_j X_j \pi_d^{\omega' j}$$

$$+ v_j X_j \eta^{\omega' j} + \sum_{d=t^*+1}^{|D|+1} \left(\theta + M_1 Z^{\omega'}\right) \mu_d^{\omega' j} \tag{5.32}$$

Subject to

$$\gamma_d^{\omega' j} + \pi_d^{\omega' j} \leq 0 \qquad \forall d \in \{t^* + 1, \dots, |D|\} \tag{5.33}$$

$$\eta^{\omega' j} - \gamma_{t^*+1}^{\omega' j} + \mu_{t^*+1}^{\omega' j} \leq 0 \tag{5.34}$$

$$\gamma_{d-1}^{\omega' j} - \gamma_d^{\omega' j} + \mu_d^{\omega' j} \leq 1 \qquad \forall d \in \{t^* + 2, \dots, |D|\} \tag{5.35}$$

$$\gamma_{|D|}^{\omega' j} + \mu_{|D|+1}^{\omega' j} \leq 1 \tag{5.36}$$

$$\pi_d^{\omega' j} \leq 0 \qquad \forall d \in \{t^* + 1, \dots, |D| \tag{5.37}$$

$$\mu_d^{\omega' j} \leq 0 \qquad \forall d \in \{t^* + 1, \dots, |D| + 1\} \tag{5.38}$$

Constraints (5.33) correspond to primal variables $\boldsymbol{W}^{\omega' j}$. Constraints (5.34) correspond to the primal initial backlog variables, $V_{j,t^*+1}^{\omega'}$, constraints (5.35) correspond to primal variables $V_{jd}^{\omega'}$ for $d \in \{t^* + 2, \dots, |D|\}$, and constraints (5.36) correspond to the primal backlog variables $V_{j,|D|+1}^{\omega'}$. Constraints (5.37) and (5.38) are sign constraints on the dual variables corresponding to primal constraints (5.27) and (5.29), respectively.

We denote the feasible region of D-M1S2(ω', j) by

$$\wp^{\omega' j} \coloneqq \{\boldsymbol{\gamma}^{\boldsymbol{\omega' j}}, \boldsymbol{\pi}^{\boldsymbol{\omega' j}}, \eta^{\omega' j}, \boldsymbol{\mu}^{\boldsymbol{\omega' j}} : \text{Constraints (5.33) - (5.38) are satisfied}\}.$$

Since setting all of the dual variables equal to zero provides a feasible solution to D-M1S2(ω', j), the set $\wp^{\omega' j}$ is nonempty. The set $\wp^{\omega j}$ can be described by its extreme points $\Phi^{\omega' j} \coloneqq \left\{(\boldsymbol{\gamma}^{\omega' j}, \boldsymbol{\pi}^{\omega' j}, \eta^{\omega' j}, \boldsymbol{\mu}^{\omega' j}) | \ (\boldsymbol{\gamma}^{\omega' j}, \boldsymbol{\pi}^{\omega' j}, \eta^{\omega' j}, \boldsymbol{\mu}^{\omega' j}) \text{ is an extreme point of } \wp^{\omega' j}\right\}$ and extreme rays $\Psi^{\omega' j} \coloneqq \left\{(\boldsymbol{\gamma}^{\omega' j}, \boldsymbol{\pi}^{\omega' j}, \eta^{\omega' j}, \boldsymbol{\mu}^{\omega' j}) | \ (\boldsymbol{\gamma}^{\omega' j}, \boldsymbol{\pi}^{\omega' j}, \eta^{\omega' j}, \boldsymbol{\mu}^{\omega' j}) \text{ is an extreme ray of } \wp^{\omega' j}\right\}$. Since D-M1S2($\omega', j$) is always feasible, it is either unbounded above or has an optimal solution. If it is unbounded above, there exists an extreme ray $(\boldsymbol{\gamma}^{\omega' j\psi}, \boldsymbol{\pi}^{\omega' j\psi}, \eta^{\omega' j\psi}, \boldsymbol{\mu}^{\omega' j\psi}) \in \Psi^{\omega' j}$, such that

$$\sum_{d=t^*+1}^{|D|}\sum_{i\in I} h_{i,d-t_{ij}}^{\omega'} Y_{ij}\gamma_d^{\omega' j\psi} + \sum_{d=t^*+1}^{|D|} k_j X_j \pi_d^{\omega' j\psi} + v_j X_j \eta^{\omega' j\psi}$$

$$+ \sum_{d=t^*+1}^{|D|+1}\left(\theta + M_1 Z^{\omega'}\right)\mu_d^{\omega' j\psi} > 0 \tag{5.39}$$

in which case $\bar{V}_{Dual}^{\omega' j}(\boldsymbol{X},\boldsymbol{Y},\boldsymbol{Z}) = +\infty$. If D-M1S2($\omega', j$) has a solution, then there exists an extreme point $(\boldsymbol{\gamma}^{\omega' j\phi}, \boldsymbol{\pi}^{\omega' j\phi}, \eta^{\omega' j\phi}, \boldsymbol{\mu}^{\omega' j\phi}) \in \Phi^{\omega' j}$, such that

$$\bar{V}_{Dual}^{\omega' j}(\boldsymbol{X},\boldsymbol{Y},\boldsymbol{Z}) = \sum_{d=t^*+1}^{|D|}\sum_{i\in I} h_{i,d-t_{ij}}^{\omega'} Y_{ij}\gamma_d^{\omega' j\phi} + \sum_{d=t^*+1}^{|D|} k_j X_j \pi_d^{\omega' j\phi} + v_j X_j \eta^{\omega' j\phi}$$

$$+ \sum_{d=t^*+1}^{|D|+1}\left(\theta + M_1 Z^{\omega'}\right)\mu_d^{\omega' j\phi}. \tag{5.40}$$

Thus, we can reformulate D-M1S2(ω', j) in terms of the extreme points and extreme rays of $\wp^{\omega' j}$ as follows:

[RD-M1S2(ω', j)]:

$$\bar{V}_{Dual}^{\omega' j}(\boldsymbol{X},\boldsymbol{Y},\boldsymbol{Z}) = Min_{\bar{V}^{\omega' j}} \bar{V}^{\omega' j} \tag{5.41}$$

Subject to

$$\sum_{d=t^*+1}^{|D|}\sum_{i\in I} h_{i,d-t_{ij}}^{\omega'} Y_{ij}\gamma_d^{\omega' j\psi} + \sum_{d=t^*+1}^{|D|} k_j X_j \pi_d^{\omega' j\psi} + v_j X_j \eta^{\omega' j\psi}$$

$$+ \sum_{d=t^*+1}^{|D|+1}\left(\theta + M_1 Z^{\omega'}\right)\mu_d^{\omega' j\psi} \le 0 \qquad \forall\psi \in \Psi^{\omega' j} \tag{5.42}$$

$$\sum_{d=t^*+1}^{|D|}\sum_{i\in I} h_{i,d-t_{ij}}^{\omega'} Y_{ij}\gamma_d^{\omega' j\phi} + \sum_{d=t^*+1}^{|D|} k_j X_j \pi_d^{\omega' j\phi} + v_j X_j \eta^{\omega' j\phi}$$

$$+ \sum_{d=t^*+1}^{|D|+1}\left(\theta + M_1 Z^{\omega'}\right)\mu_d^{\omega' j\phi} \le \bar{V}^{\omega' j} \qquad \forall\phi \in \Phi^{\omega' j} \tag{5.43}$$

where constraints (5.42) and (5.43) are feasibility and optimality constraints, respectively. Furthermore, we can aggregate the optimality and feasibility cuts by candidate facility to obtain the aggregated reformulation of the D-M1S2(ω', j) problems:

[ARD-M1S2(ω')]:

$$\bar{V}_{A-Dual}^{\omega'}(\boldsymbol{X},\boldsymbol{Y},\boldsymbol{Z}) = Min_{\bar{V}^{\omega'}} \bar{V}^{\omega'} \tag{5.44}$$

Subject to

$$\sum_{j\in J}\Big(\sum_{d=t^*+1}^{|D|}\sum_{i\in I} h_{i,d-t_{ij}}^{\omega'} Y_{ij}\gamma_d^{\omega' j\psi} + \sum_{d=t^*+1}^{|D|} k_j X_j \pi_d^{\omega' j\psi} + v_j X_j \eta^{\omega' j\psi} + \sum_{d=t^*+1}^{|D|+1}\big(\theta + M_1 Z^{\omega'}\big)\mu_d^{\omega' j\psi}\Big) \le 0 \qquad \forall \psi \in \Psi^{\omega'} \tag{5.45}$$

$$\sum_{j\in J}\Big(\sum_{d=t^*+1}^{|D|}\sum_{i\in I} h_{i,d-t_{ij}}^{\omega'} Y_{ij}\gamma_d^{\omega' j\phi} + \sum_{d=t^*+1}^{|D|} k_j X_j \pi_d^{\omega' j\phi} + v_j X_j \eta^{\omega' j\phi} + \sum_{d=t^*+1}^{|D|+1}\big(\theta + M_1 Z^{\omega'}\big)\mu_d^{\omega' j\phi}\Big) \le \bar{V}^{\omega'} \qquad \forall \phi \in \Phi^{\omega'} \tag{5.46}$$

where $\Phi^{\omega'}$ and $\Psi^{\omega'}$ represent the set of dual solutions $\{\boldsymbol{\gamma}^{\omega'}, \boldsymbol{\pi}^{\omega'}, \boldsymbol{\eta}^{\omega'}, \boldsymbol{\mu}^{\omega'}\}$ that are extreme points or extreme rays, respectively, of the feasible region:

$$\gamma_d^{\omega' j} + \pi_d^{\omega' j} \le 0 \qquad \forall j \in J; d \in \{t^*+1, \dots, |D|\} \tag{5.47}$$

$$\eta^{\omega' j} - \gamma_{t^*+1}^{\omega' j} + \mu_{t^*+1}^{\omega' j} \le 0 \qquad \forall j \in J \tag{5.48}$$

$$\gamma_{d-1}^{\omega' j} - \gamma_d^{\omega' j} + \mu_d^{\omega' j} \le 1 \qquad \forall j \in J; d \in \{t^*+2, \dots, |D|\} \tag{5.49}$$

$$\gamma_{|D|}^{\omega' j} + \mu_{|D|+1}^{\omega' j} \le 1 \qquad \forall j \in J \tag{5.50}$$

$$\pi_d^{\omega' j} \le 0 \qquad \forall j \in J; d \in \{t^*+1, \dots, |D| \tag{5.51}$$

$$\mu_d^{\omega' j} \le 0 \qquad \forall j \in J; d \in \{t^*+1, \dots, |D|+1\} \tag{5.52}$$

Furthermore, we note that $\bar{V}^{\omega'} = \sum_{j\in J} \bar{V}^{\omega' j}$.

At iteration n, we can consider the relaxed problem RMP^n-M1S1 composed of only a subset of the extreme points and extreme rays by combining M1S1 and ARD-M1S2(ω'), $\forall \omega' \in \Omega'$:

[RMP^n-M1S1]:

$$\bar{V}^n_{RMP}(\boldsymbol{\gamma},\boldsymbol{\pi},\boldsymbol{\eta},\boldsymbol{\mu}) = Min_{\boldsymbol{X},\boldsymbol{Y},\boldsymbol{Z},\mathbb{V}}\ (|D|-t^*)\sum_{j\in J} f_j X_j + a\sum_{i\in I}\sum_{d=t^*+1}^{|D|}\sum_{j\in J}\sum_{\omega\in\Omega} p^{\omega'} h_{ij}^{\omega'}\, t_{ij} Y_{ij}$$

$$+b\,\mathbb{V} \tag{5.53}$$

Subject to

$$\sum_{j\in J} Y_{ij} = 1 \qquad \forall i \in I \tag{2.2}$$

$$Y_{ij} - X_j \le 0 \qquad \forall i \in I; j \in J \tag{2.3}$$

$$\sum_{j\in J} X_j \ge 1 \tag{5.22}$$

$$\sum_{\omega'\in\Omega'} p^{\omega'} Z^{\omega'} \le \tau \tag{5.11}$$

$$X_j \in \{0,1\} \qquad \forall j \in J \tag{2.4}$$

$$Y_{ij} \in \{0,1\} \qquad \forall i \in I; j \in J \tag{3.6}$$

$$Z^{\omega'} \in \{0,1\} \qquad \forall \omega' \in \Omega' \tag{5.12}$$

Initialization Cuts

$$\sum_{j\in J}\Big(\sum_{d=t^*+1}^{|D|}\sum_{i\in I} h_{i,d-t_{ij}}^{\omega'} Y_{ij}\gamma_d^{\omega' j\psi} + \sum_{d=t^*+1}^{|D|} k_j X_j \pi_d^{\omega' j\psi} + v_j X_j \eta^{\omega' j\psi}$$

$$+\sum_{d=t^*+1}^{|D|+1}\big(\theta + M_1 Z^{\omega'}\big)\mu_d^{\omega' j\psi}\Big) \le 0 \qquad \forall \omega' \in \Omega', \psi \in \underline{\Psi}^{\omega'} \subseteq \Psi^{\omega'} \tag{5.54}$$

$$\sum_{\omega'\in\Omega'} p^{\omega'}\sum_{j\in J}\Big(\sum_{d=t^*+1}^{|D|}\sum_{i\in I} h_{i,d-t_{ij}}^{\omega'} Y_{ij}\gamma_d^{\omega' j\phi} + \sum_{d=t^*+1}^{|D|} k_j X_j \pi_d^{\omega' j\phi} + v_j X_j \eta^{\omega' j\phi}$$

$$+\sum_{d=t^*+1}^{|D|+1}\big(\theta + M_1 Z^{\omega'}\big)\mu_d^{\omega' j\phi}\Big) \le \mathbb{V} \qquad \forall \phi \in \underline{\Phi} \subseteq \Phi \tag{5.55}$$

where $\Phi = \cup_{\omega'\in\Omega'} \Phi^{\omega'}$ and $\mathbb{V}$ is a decision variable that represents the total amount of backlog over the planning horizon. The sets $\underline{\Psi}^{\omega'}$ and $\underline{\Phi}$ represent the subset of feasibility and optimality cuts generated, respectively. We outline the two-stage solution method in Algorithm 1.

Algorithm 1: Two-stage cutting plane algorithm based on the Benders decomposition scheme

Step 0: Initialize $n = 1, \underline{\Psi}^{\omega'} = \emptyset, \underline{\Phi} = \emptyset$. Set the upper bound (UB) and lower bound (LB) on the optimal objective function value (5.14) appropriately (e.g., set UB equal to infinity or the objective function value of a known feasible solution, set LB equal to 0) and add the corresponding initialization cuts to RMP^n-M1S1 (e.g., $0 \leq \mathbb{V}$).

Step 1: Solve RMP^n-M1S1 and obtain an optimal solution $(\boldsymbol{X}^n, \boldsymbol{Y}^n, \boldsymbol{Z}^n, \mathbb{V}^n)$. Set the LB equal to the current value of (5.53) (i.e., $\bar{V}^n_{RMP}(\boldsymbol{\gamma}, \boldsymbol{\pi}, \boldsymbol{\eta}, \boldsymbol{\mu})$ with $(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}, \mathbb{V}) = (\boldsymbol{X}^n, \boldsymbol{Y}^n, \boldsymbol{Z}^n, \mathbb{V}^n)$ as input).

Step 2: [Feasibility Check] For each $\omega' \in \Omega'$ and $j \in J$, use $(\boldsymbol{X}^n, \boldsymbol{Y}^n, \boldsymbol{Z}^n)$ as input to solve the separation problem, which is essentially D-M1S2(ω', j) except that the right hand side of all of the constraints is replaced by the value 0. Denote the optimal objective function of each problem by $\bar{V}^{\omega' j}_{Sep}(\boldsymbol{X}^n, \boldsymbol{Y}^n, \boldsymbol{Z}^n)$. Normalize the solution[1] and let $\sum_{j \in J} \bar{V}^{\omega' j}_{Sep}(\boldsymbol{X}^n, \boldsymbol{Y}^n, \boldsymbol{Z}^n) = \bar{V}^{\omega'}_{Sep}(\boldsymbol{X}^n, \boldsymbol{Y}^n, \boldsymbol{Z}^n)$.

a) If $\bar{V}^{\omega'}_{Sep}(\boldsymbol{X}^n, \boldsymbol{Y}^n, \boldsymbol{Z}^n) \leq 0$ for all $\omega' \in \Omega'$, proceed to Step 3.

b) If $\exists \omega' \in \Omega'$ such that $\bar{V}^{\omega'}_{Sep}(\boldsymbol{X}^n, \boldsymbol{Y}^n, \boldsymbol{Z}^n) > 0$, then the optimal solution $(\boldsymbol{\gamma}^{\omega'\psi}, \boldsymbol{\pi}^{\omega'\psi}, \boldsymbol{\eta}^{\omega'\psi}, \boldsymbol{\mu}^{\omega'\psi})$ contains an extreme ray $(\boldsymbol{\gamma}^{\omega' j\psi}, \boldsymbol{\pi}^{\omega' j\psi}, \eta^{\omega' j\psi}, \boldsymbol{\mu}^{\omega' j\psi})$, which makes D-M1S2($\omega', j$) unbounded. For each such ω', add $(\boldsymbol{\gamma}^{\omega'\psi}, \boldsymbol{\pi}^{\omega'\psi}, \boldsymbol{\eta}^{\omega'\psi}, \boldsymbol{\mu}^{\omega'\psi})$ to $\underline{\Psi}^{\omega'}$. Set $n = n + 1$ and return to Step 1.

Step 3: [Optimality Check] For each $\omega' \in \Omega'$ and $j \in J$, solve D-M1S2(ω', j) with $(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}) = (\boldsymbol{X}^n, \boldsymbol{Y}^n, \boldsymbol{Z}^n)$. Denote the optimal solution by $(\boldsymbol{\gamma}^{\omega' j\phi}, \boldsymbol{\pi}^{\omega' j\phi}, \eta^{\omega' j\phi}, \boldsymbol{\mu}^{\omega' j\phi})$.

[1] For example, by dividing each element of $\boldsymbol{\gamma}^{\omega'\psi}, \boldsymbol{\pi}^{\omega'\psi}, \boldsymbol{\eta}^{\omega'\psi}, \boldsymbol{\mu}^{\omega'\psi})$ by the L2-norm of $(\boldsymbol{\gamma}^{\omega'\psi}, \boldsymbol{\pi}^{\omega'\psi}, \boldsymbol{\eta}^{\omega'\psi}, \boldsymbol{\mu}^{\omega'\psi})$.

a) If

$$\mathbb{V}^n < \sum_{\omega' \in \Omega'} p^{\omega'} \sum_{j \in J} \Big(\sum_{d=t^*+1}^{|D|} \sum_{i \in I} h_{i,d-t_{ij}}^{\omega'} Y_{ij}^n \gamma_d^{\omega' j \phi} + \sum_{d=t^*+1}^{|D|} k_j X_j^n \pi_d^{\omega' j \phi} + v_j X_j^n \eta^{\omega' j \phi} + \sum_{d=t^*+1}^{|D|+1} \big(\theta + M_1 Z^{\omega' n}\big) \mu_d^{\omega' j \phi} \Big) \quad (5.56)$$

add $(\boldsymbol{\gamma}^{\omega' \phi}, \boldsymbol{\pi}^{\omega' \phi}, \boldsymbol{\eta}^{\omega' \phi}, \boldsymbol{\mu}^{\omega' \phi})$ to $\underline{\Phi}$. Set $UB := \min\{UB, (|D| - t^*) \sum_{j \in J} f_j X_j^n + a \sum_{i \in I} \sum_{d=t^*+1}^{|D|} \mathbb{E}_{\xi'}[\widetilde{h_{id}}] \sum_{j \in J} t_{ij} Y_{ij}^n + b\, \mathbb{V}^n\}$. Then, let $n = n + 1$ and return to Step 1.

b) If the inequality (5.56) does not hold, then $(\boldsymbol{X}^n, \boldsymbol{Y}^n, \boldsymbol{Z}^n, \mathbb{V}^n)$ is the optimal solution to MIP-JCC. The algorithm terminates.

5.3.2 Three-Stage Decomposition Approach

The second decomposition solution approach removes the auxiliary $\boldsymbol{Z}$ variables that verify whether the joint chance constraint is satisfied from the first stage and places them in the third stage. Given a feasible first-stage location and allocation solution, we can readily determine the optimal second-stage processing and backlog decisions as well as the third-stage auxiliary variables that verify whether the joint chance constraint is satisfied by inspection. If the joint chance constraint is violated, we add a corresponding cut to the first-stage problem. The three stages are defined as follows:

[M2S1]:

$$Min_{\boldsymbol{X},\boldsymbol{Y}} (|D| - t^*) \sum_{j \in J} \Big(f_j X_j + a \sum_{i \in I} \sum_{d=t^*+1}^{|D|} \sum_{\omega' \in \Omega'} p^{\omega'} h_{ij}^{\omega'}\, t_{ij} Y_{ij} + b \sum_{\omega' \in \Omega'} p^{\omega'} \hat{V}^{\omega' j}(\boldsymbol{X}, \boldsymbol{Y}) \Big) \quad (5.57)$$

Subject to

$$\sum_{j \in J} Y_{ij} = 1 \qquad \forall i \in I \quad (2.2)$$

$$Y_{ij} - X_j \leq 0 \qquad \forall i \in I; j \in J \quad (2.3)$$

$$\sum_{j\in J} X_j \geq 1 \tag{5.22}$$

$$X_j \in \{0,1\} \qquad \forall j \in J \tag{2.4}$$

$$Y_{ij} \in \{0,1\} \qquad \forall i \in I; j \in J \tag{3.6}$$

For each scenario $\omega' \in \Omega'$ and facility $j \in J$, the value of $\hat{V}^{\omega' j}(\boldsymbol{X}, \boldsymbol{Y})$ in M2S1 is given by:

[M2S2 (ω', j)]:

$$\hat{V}^{\omega' j}(\boldsymbol{X}, \boldsymbol{Y}) = Min_{\boldsymbol{V}^{\omega' j}, \boldsymbol{W}^{\omega' j}} \sum_{d=t^*+2}^{|D|+1} V_{jd}^{\omega'} + \hat{Z}(\boldsymbol{V}) \tag{5.58}$$

Subject to

$$V_{j,d+1}^{\omega'} - V_{jd}^{\omega'} + W_{jd}^{\omega'} = \sum_{i\in I} h_{i,d-t_{ij}}^{\omega'} Y_{ij} \qquad \forall d \in \{t^*+1, \dots, |D|\} \tag{5.26}$$

$$W_{jd}^{\omega'} \leq k_j X_j \qquad \forall d \in \{t^*+1, \dots, |D|\} \tag{5.27}$$

$$V_{j,t^*+1}^{\omega'} = v_j X_j \tag{5.28}$$

$$V_{jd}^{\omega'} \geq 0 \qquad \forall d \in \{t^*+1, \dots, |D|+1\} \tag{5.30}$$

$$W_{jd}^{\omega'} \geq 0 \qquad \forall d \in \{t^*+1, \dots, |D|\} \tag{5.31}$$

Then, for a sufficiently large value of M_2 so that M2S1 does not choose location and allocation decisions that cause the third-stage variable, $\mathbb{Q}$, to take a positive value, we calculate $\hat{Z}(\boldsymbol{V})$ as:

[M2S3]:

$$\hat{Z}(\boldsymbol{V}) = Min_{\boldsymbol{Z}, \mathbb{Q}} \; M_2 \mathbb{Q} \tag{5.59}$$

Subject to

$$M_1 Z^{\omega'} \geq V_{jd}^{\omega'} - \theta \qquad \forall \omega' \in \Omega';\ j \in J;\ d \in \{t^*+1, \dots, |D|+1\} \tag{5.10}$$

$$\sum_{\omega'\in\Omega'} p^{\omega'} Z^{\omega'} - \mathbb{Q} \leq \tau \tag{5.60}$$

$$\mathbb{Q} \geq 0 \tag{5.61}$$

$$Z^{\omega'} \in \{0,1\} \qquad \forall \omega' \in \Omega' \tag{5.12}$$

First, we note that, for any feasible $(\boldsymbol{X}, \boldsymbol{Y})$ solution to M2S1, constraint (5.22) ensures that a feasible solution can be constructed for M2S2(ω', j). For example, if we do not process any demand at the processing facilities but instead hold all of the incoming demand in backlog (i.e., $\boldsymbol{W}^{\omega' j} = \boldsymbol{0}$), then the initial backlog levels, $V_{j,t^*+1}^{\omega'}$, are determined by constraints (5.28) and the iterative equation $V_{j,d+1}^{\omega'} = V_{jd}^{\omega'} + \sum_{i \in I} h_{i,d-t_{ij}}^{\omega'} Y_{ij} \ \forall\, d \in \{t^*+1, \dots, |D|\}$. The resulting processing and backlog values satisfy constraints (5.27), (5.30), and (5.31).

Furthermore, given $\boldsymbol{X}$ and $\boldsymbol{Y}$ values, the optimal backlog ($\boldsymbol{V}^{\omega' j}$) and daily processing variables ($\boldsymbol{W}^{\omega' j}$) of M2S2$(\omega', j)$ can easily be solved by inspection as described in Algorithm 2. To see this, note that given $\boldsymbol{X}$, constraint (5.28) automatically determines the initial backlog levels $V_{j,t^*+1}^{\omega'}$. The backlog levels for all remaining days can then be determined using the following recursive equation:

$$V_{jd}^{\omega'} = \max\{\textstyle\sum_{i \in I} h_{i,(d-1)-t_{ij}}^{\omega'} Y_{ij} - k_j X_j + V_{j,d-1}^{\omega'}, 0\} \tag{5.62}$$

Thus, for $d \in \{t^*+2, \dots, |D|+1\}$ the backlog variables can be expressed as follows:

$$V_{jd}^{\omega'} = \max\{\textstyle\sum_{i \in I} h_{i,(d-1)-t_{ij}}^{\omega'} Y_{ij} - k_j X_j + \max\{\sum_{i \in I} h_{i,(d-2)-t_{ij}}^{\omega'} Y_{ij} - k_j X_j + \max\{\dots$$

$$\dots + \max\left\{\textstyle\sum_{i \in I} h_{i,((t^*+1)-t_{ij})}^{\omega'} Y_{ij} - k_j X_j + V_{j,t^*+1}^{\omega'}, 0\right\}, 0\} \dots, 0\}, 0\}, 0\}. \tag{5.63}$$

The number of items processed each day at a given facility is the minimum of the processing capacity of the facility and the number of available items available for processing. Thus, the number of items processed at facility j on day $d \in \{t^*+1, \dots, |D|\}$ in scenario ω is:

$$W_{jd}^{\omega'} = \min\{k_j X_j, V_{jd}^{\omega'} + \textstyle\sum_{i \in I} h_{i,d-t_{ij}}^{\omega'} Y_{ij}\}.$$

Algorithm 2: Recursive algorithm for solving M2S2(ω', j)

Given $\omega' \in \Omega'$, $j \in J$, and any feasible $(\boldsymbol{X}, \boldsymbol{Y})$ solution to M2S1:

a. Set $V_{j,t^*+1}^{\omega'} = v_j X_j$

b. For $d \in \{t^* + 2, \dots, |D| + 1\}$ let $V_{jd}^{\omega'} = \max\{\sum_{i \in I} h_{i,((d-1)-t_{ij})}^{\omega'} Y_{ij} - k_j X_j + \mathrm{V}_{j,d-1}^{\omega'}, 0\}$

c. For $d \in \{t^* + 1, \dots, |D|\}$ let $W_{jd}^{\omega'} = \min\left\{k_j X_j, V_{jd}^{\omega'} + \sum_{i \in I} h_{i,d-t_{ij}}^{\omega'} Y_{ij}\right\}$

Finally, once the values of the backlog variables are determined, problem M2S3 is also easily determined by inspection. We simply check whether $V_{jd}^{\omega'} \leq \theta$ for all $j \in J, d \in \{t^* + 1, \dots, |D| + 1\}$ for each $\omega' \in \Omega'$. If the inequality is violated, we set the corresponding $Z^{\omega'} = 1$. If not, we let $Z^{\omega'} = 0$. Then, we can check whether $\sum_{\omega' \in \Omega'} p^{\omega'} Z^{\omega'} \leq \tau$. If not, we let $\mathbb{Q} = \sum_{\omega' \in \Omega'} p^{\omega'} Z^{\omega'} - \tau$ and conclude that $(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{V}, \boldsymbol{W}, \boldsymbol{Z})$ is an infeasible solution to MIP-JCC. If $\sum_{\omega' \in \Omega'} p^{\omega'} Z^{\omega'} \leq \tau$, then $\mathbb{Q} = 0$ and $(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{V}, \boldsymbol{W}, \boldsymbol{Z})$ is a feasible solution to MIP-JCC.

Thus, given a feasible $(\boldsymbol{X}, \boldsymbol{Y})$ solution to M2S1, we can readily determine the corresponding values of $\boldsymbol{V}, \boldsymbol{W}$, and $\boldsymbol{Z}$. However, we have not specified how to obtain the optimal $(\boldsymbol{X}, \boldsymbol{Y})$ values. To do so, we generate infeasibility and optimality cuts and add them to M2S1. We call the resulting formulation R-M2S1. Let $\underline{\widehat{\Psi}}$ and $\underline{\widehat{\Phi}}$ represent the set of infeasibility and optimality cuts, respectively and denote ψ as the index of set $\underline{\widehat{\Psi}}$ (i.e., $\psi \in \underline{\widehat{\Psi}}$). Furthermore, denote the allocation and backlog variables corresponding to an infeasible solution ψ by Y_{ij}^{ψ} and $V_{jd}^{\omega'\psi}$, respectively. The infeasibility cuts utilize the sets $J^{\psi} \coloneqq \left\{j \in J: \exists \omega' \in \Omega'; d \in \{t^* + 1, \dots, |D| + 1\} \text{ with } V_{jd}^{\omega'\psi} > \theta\right\}$ to represent located facilities that have violated the backlog

threshold of θ and $I_j^{\psi} \coloneqq \left\{i \in I: Y_{ij}^{\psi} = 1\right\}$ to represent demand sites allocated to processing facility $j \in J$. The problem R-M2S1 is defined as follows:

[R-M2S1]:

$$Min\ _{\boldsymbol{X},\boldsymbol{Y},\widehat{\mathbb{V}}}\ (|D| - t^*) \sum_{j\in J} f_j X_j + a \sum_{i\in I} \sum_{d=t^*+1}^{|D|} \mathbb{E}_{\xi'}\left[\tilde{h}_{id}\right] \sum_{j\in J} t_{ij} Y_{ij} + b\widehat{\mathbb{V}} \tag{5.64}$$

Subject to

$$\sum_{j\in J} Y_{ij} = 1 \qquad \forall i \in I \tag{2.2}$$

$$Y_{ij} - X_j \leq 0 \qquad \forall i \in I; j \in J \tag{2.3}$$

$$\sum_{j\in J} X_j \geq 1 \tag{5.22}$$

$$X_j \in \{0,1\} \qquad \forall j \in J \tag{2.4}$$

$$Y_{ij} \in \{0,1\} \qquad \forall i \in I; j \in J \tag{3.6}$$

Initialization cuts

$$\sum_{j\in J^{\psi}} \sum_{i\in I_j^{\psi}} Y_{ij} \leq \sum_{j\in J^{\psi}} |I_j^{\psi}| - 1 \qquad \forall \psi \in \underline{\widehat{\Psi}} \tag{5.65}$$

$$\sum_{\omega'\in\Omega'} p^{\omega'} * \left(\sum_{j\in J} \sum_{d=t^*+1}^{|D|} \sum_{i\in I} h_{i,d-t_{ij}}^{\omega'} Y_{ij} \gamma_d^{\omega' j\phi} + \sum_{j\in J} \sum_{d=t^*+1}^{|D|} k_j X_j \pi_d^{\omega' j\phi} + \sum_{j\in J} v_j X_j \eta^{\omega' j\phi}\right) \leq \widehat{\mathbb{V}} \qquad \forall \phi \in \underline{\widehat{\Phi}} \tag{5.66}$$

where $\widehat{\mathbb{V}}$ is a decision variable that represents the expected total backlog over the planning horizon. Constraints (5.65) are feasibility cuts that ensure a future optimal solution has an allocation plan that is different from any previous solution that has been deemed infeasible and constraints (5.66) are optimality cuts. Some variants of (5.65) applied to binary first-stage decisions are discussed in Laporte and Louveaux (1993) and Deng and Shen (2015). We present a proof that constraints (5.65) are valid feasibility cuts in Section 5.3.3.

Given this notation, we now present Algorithm 3, which checks whether a feasible ($\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{V}, \boldsymbol{W}$) solution to the first and second stages satisfies the mixed integer linear joint chance constraint formulation (5.10) - (5.12) and generates a feasibility cut if it is not satisfied. We can perform this check without explicitly solving problem M2S3, and therefore, we do not need the variable $\mathbb{Q}$.

Algorithm 3: Third-stage solution and feasibility check

Given a feasible solution $(\boldsymbol{X}^n, \boldsymbol{Y}^n, \boldsymbol{V}^n, \boldsymbol{W}^n)$ to the n^{th} iteration of R-M2S1 and M2S2(ω', j),

a. Initialize $Z^{\omega' n} = 0, \forall \omega' \in \Omega'$

b. For each $\omega' \in \Omega'$, check whether $V_{jd}^{\omega' n} \leq \theta \; \forall j \in J; \; d \in \{t^* + 1, \dots, |D| + 1\}$.

- If the inequality does not hold, set the corresponding $Z^{\omega' n} = 1$

c. Check whether $\sum_{\omega' \in \Omega'} p^{\omega'} Z^{\omega' n} \leq \tau$.

- If the inequality does not hold, then $(\boldsymbol{X}^n, \boldsymbol{Y}^n, \boldsymbol{V}^n, \boldsymbol{W}^n, \boldsymbol{Z}^n)$ is an infeasible solution to MIP-JCC. Cut off the solution $(\boldsymbol{X}^n, \boldsymbol{Y}^n, \boldsymbol{V}^n, \boldsymbol{W}^n, \boldsymbol{Z}^n)$ by adding the constraint

$$\sum_{j \in J^n} \sum_{i \in I_j^n} Y_{ij} \leq \sum_{j \in J^n} |I_j^n| - 1 \tag{5.67}$$

to constraint set $\underline{\Psi}$. Here $J^n := \{j \in J: \exists \omega' \in \Omega'; d \in \{t^* + 1, \dots, |D| + 1\}$ with $V_{jd}^{\omega' n} > \theta\}$ represents the located facilities that violate the backlog threshold of θ and $I_j^n := \{i \in I: Y_{ij}^n = 1\}$ represents the set of demand sites that allocate demand to processing facility $j \in J$ as obtained from the solution of R-M2S1.

- If the inequality holds, $(\boldsymbol{X}^n, \boldsymbol{Y}^n, \boldsymbol{V}^n, \boldsymbol{W}^n, \boldsymbol{Z}^n)$ is a feasible solution to MIP-JCC. As such, it produces an upper bound on (5.14). Set

$$UB = \min(UB, (|D| - t^*) \sum_{j \in J} f_j X_j^n + a \sum_{i \in I} \sum_{d=t^*+1}^{|D|} \mathbb{E}_{\xi'}[\tilde{h}_{id}] \sum_{j \in J} t_{ij} Y_{ij}^n$$

$$+ b \sum_{\omega' \in \Omega'} p^{\omega'} (\sum_{d=t^*+2}^{|D|+1} \sum_{j \in J} V_{jd}^{\omega' n})).$$

Once a feasible solution to MIP-JCC is found, we solve the dual of M2S2(ω', j) for each $\omega' \in \Omega'$ and each $j \in J$ to check the solution's optimality. The dual problem is defined as:

[D-M2S2(ω', j)]:

$$\hat{V}_{Dual}^{\omega' j}(\boldsymbol{X}, \boldsymbol{Y}) = Max_{\boldsymbol{\gamma}^{\omega' j}, \boldsymbol{\pi}^{\omega' j}, \boldsymbol{\eta}^{\omega' j}} \sum_{j \in J} \sum_{d=t^*+1}^{|D|} \sum_{i \in I} h_{i,d-t_{ij}}^{\omega'} Y_{ij} \gamma_d^{\omega' j} + \sum_{j \in J} \sum_{d=t^*+1}^{|D|} k_j X_j \pi_d^{\omega' j}$$

$$+ \sum_{j \in J} v_j X_j \eta^{\omega' j} \quad (5.68)$$

Subject to

$$\gamma_d^{\omega' j} + \pi_d^{\omega' j} \leq 0 \qquad \forall d \in \{t^* + 1, \dots, |D|\} \quad (5.33)$$

$$\eta^{\omega' j} - \gamma_{t^*+1}^{\omega' j} \leq 0 \quad (5.69)$$

$$\gamma_{d-1}^{\omega' j} - \gamma_d^{\omega' j} \leq 1 \qquad \forall d \in \{t^* + 2, \dots, |D|\} \quad (5.70)$$

$$\gamma_{|D|}^{\omega' j} \leq 1 \quad (5.71)$$

$$\pi_d^{\omega' j} \leq 0 \qquad \forall d \in \{t^* + 1, \dots, |D|\} \quad (5.38)$$

The complete algorithm for solving the three-stage problem is presented in Algorithm 4.

Algorithm 4: Three-stage cutting plane algorithm

Step 0: Initialize $n = 1, \underline{\hat{\Psi}} = \emptyset, \underline{\hat{\Phi}} = \emptyset, UB = \infty$, and $LB = -\infty$. Add an appropriate initialization cut (e.g., $\hat{\mathbb{V}} \geq 0$) to R-M2S1.

Step 1: Solve R-M2S1 and let $(\boldsymbol{X}^n, \boldsymbol{Y}^n, \hat{\mathbb{V}}^n)$ denote an optimal solution. Set the LB equal to the current value of (5.64).

Step 2: For each $\omega' \in \Omega'$ and $j \in J$, use Algorithm 2 with $(\boldsymbol{X}^n, \boldsymbol{Y}^n)$ as input to solve M2S2(ω', j). Denote the aggregated optimal solution to M2S2(ω', j) by $(\boldsymbol{V}^n, \boldsymbol{W}^n)$.

Step 3: [Feasibility Check] Using $(\boldsymbol{X}^n, \boldsymbol{Y}^n, \boldsymbol{V}^n, \boldsymbol{W}^n)$ as input, perform Algorithm 3 to determine if $(\boldsymbol{X}^n, \boldsymbol{Y}^n, \boldsymbol{V}^n, \boldsymbol{W}^n)$ is a feasible solution to MIP-JCC. If $(\boldsymbol{X}^n, \boldsymbol{Y}^n, \boldsymbol{V}^n, \boldsymbol{W}^n)$ is an infeasible solution to MIP-JCC, set $n = n + 1$ and return to Step 1. Otherwise, proceed to Step 4.

Step 4: [Optimality Check] For each $\omega' \in \Omega'$ and $j \in J$, solve D-M2S2(ω', j) with $(\boldsymbol{X}, \boldsymbol{Y}) = (\boldsymbol{X}^n, \boldsymbol{Y}^n)$. Denote the aggregated optimal solutions by $(\boldsymbol{\gamma}^{\omega'\phi}, \boldsymbol{\pi}^{\omega'\phi}, \boldsymbol{\eta}^{\omega'\phi}, \boldsymbol{\mu}^{\omega'\phi})$.

a) If $(\widehat{\mathbb{V}}^n < \sum_{\omega' \in \Omega'} p^{\omega'} \left(\sum_{j \in J} \sum_{d=t^*+1}^{|D|} \sum_{i \in I} h_{i,d-t_{ij}}^{\omega'} Y_{ij}^n \gamma_d^{\omega' j\phi} + \sum_{j \in J} \sum_{d=t^*+1}^{|D|} k_j X_j^n \pi_d^{\omega' j\phi} + \sum_{j \in J} v_j X_j^n \eta^{\omega' j\phi} \right)$ (5.72)

add $\left(\boldsymbol{\gamma}^{\omega'\phi}, \boldsymbol{\pi}^{\omega'\phi}, \boldsymbol{\eta}^{\omega'\phi}, \boldsymbol{\mu}^{\omega'\phi}\right)$ to $\underline{\widehat{\Phi}}$. Let $n = n + 1$ and return to Step 1.

b) If the inequality (5.72) does not hold, then $(\boldsymbol{X}^n, \boldsymbol{Y}^n, \boldsymbol{W^n}, \boldsymbol{V^n}, \boldsymbol{Z}^n)$ is the optimal solution to MIP-JCC. The algorithm terminates.

5.3.3 Proof of the Validity of Feasibility Cuts (5.65)

To show the validity of cut (5.65), we begin by recalling that $J^\psi := \left\{ j \in J : \exists \omega' \in \Omega';\ d \in \{t^* + 1, \dots, |D| + 1\} \text{ with } V_{jd}^{\omega'\psi} > \theta \right\}$ represents located facilities that have violated the backlog threshold of θ, and $I_j^\psi := \left\{ i \in I : Y_{ij}^\psi = 1 \right\}$ represents the demand sites allocated to processing facility $j \in J$ for $\psi \in \underline{\widehat{\Psi}}$. Additionally, we define $\mathcal{S}^\psi := \{(i,j) \in I \times J : Y_{ij}^\psi = 1 \text{ and } \exists \omega \in \Omega; d \in \{t^* + 1, \dots, |D| + 1\} \text{ with } V_{jd}^{\omega'\psi} > \theta\}$.

Let $\mathcal{F} \coloneqq \{(\boldsymbol{X}, \boldsymbol{Y})$: $\exists \boldsymbol{V}, \boldsymbol{W}, \boldsymbol{Z}$: (2.2) - (2.4), (3.6), (5.10) - (5.12), (5.15) - (5.19), and (5.22) are satisfied$\}$. For any feasible $(\boldsymbol{X}, \boldsymbol{Y})$ solution to MIP-JCC, $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathcal{F}$. Suppose another solution, $\left(\boldsymbol{X}^{\psi}, \boldsymbol{Y}^{\psi}\right)$ satisfies constraints (2.2) - (2.4), (3.6), and (5.22) but $\left(\boldsymbol{X}^{\psi}, \boldsymbol{Y}^{\psi}\right) \notin \mathcal{F}$. It follows that $(\boldsymbol{X}, \boldsymbol{Y}) \neq \left(\boldsymbol{X}^{\psi}, \boldsymbol{Y}^{\psi}\right)$.

This further implies that the allocation decisions corresponding to the two solutions are different, i.e., $\boldsymbol{Y} \neq \boldsymbol{Y}^{\psi}$. To see this, note that if $\boldsymbol{X} = \boldsymbol{X}^{\boldsymbol{\psi}}$, then it must be that $\boldsymbol{Y} \neq \boldsymbol{Y}^{\psi}$ in order for $(\boldsymbol{X}, \boldsymbol{Y}) \neq \left(\boldsymbol{X}^{\psi}, \boldsymbol{Y}^{\psi}\right)$. If instead $\boldsymbol{X} \neq \boldsymbol{X}^{\boldsymbol{\psi}}$, then $\exists \hat{\jmath} \in J$ such that $X_{\hat{\jmath}} \neq X_{\hat{\jmath}}^{\psi}$. Thus, either $X_{\hat{\jmath}} = 1$ and $X_{\hat{\jmath}}^{\psi} = 0$, or $X_{\hat{\jmath}} = 0$ and $X_{\hat{\jmath}}^{\psi} = 1$. Suppose $X_{\hat{\jmath}} = 1$ and $X_{\hat{\jmath}}^{\psi} = 0$. Then, for every $i \in I$ such that $Y_{i\hat{\jmath}} = 1$ we must have $Y_{i\hat{\jmath}}^{\psi} = 0$ since facility $\hat{\jmath}$ is not located in solution $\left(\boldsymbol{X}^{\psi}, \boldsymbol{Y}^{\psi}\right)$. Thus, $\boldsymbol{Y} \neq \boldsymbol{Y}^{\psi}$. An analogous reasoning holds for the case of $X_{\hat{\jmath}} = 0$ and $X_{\hat{\jmath}}^{\psi} = 1$.

Furthermore, the reason $\left(\boldsymbol{X}^{\psi}, \boldsymbol{Y}^{\psi}\right) \notin \mathcal{F}$ must be because too many scenario violations occur (i.e., $\sum_{\omega' \in \Omega'} p^{\omega'} Z^{\omega' \psi} > \tau$). Thus, $\exists (i', j') \in I \times J$ where j' is a facility that contributes to a scenario violation, such that $Y_{i'j'}^{\psi} = 1$ and $Y_{i'j'} = 0$. Therefore,

$$\begin{aligned}
\sum_{j \in J^{\psi}} \sum_{i \in I_j^{\psi}} Y_{ij} &= \sum_{(i,j) \in \mathcal{S}^{\psi}} Y_{ij} \\
&= Y_{i'j'} + \sum_{(i,j) \in \mathcal{S}^{\psi} \backslash (i', j')} Y_{ij} \\
&= 0 + \sum_{(i,j) \in \mathcal{S}^{\psi} \backslash (i', j')} Y_{ij} \\
&\leq |\mathcal{S}^{\psi} \backslash (i', j')| \\
&\leq \sum_{j \in J^{\psi}} \left|I_j^{\psi}\right| - 1.
\end{aligned}$$

Therefore, the inequality (5.65) holds for any feasible $(\boldsymbol{X}, \boldsymbol{Y})$ solution to MIP-JCC. □

5.3.4 Relaxed Joint Chance Constraint

Both of the decomposition approaches presented have first-stage problems that may provide location and allocation decisions that will result in infeasibility at a later stage. Thus, we propose to strengthen the first stage by adding the constraint

$$\mathbb{P}\left(\sum_{i\in I} \tilde{h}_{i,d-t_{ij}} Y_{ij} \leq\ k_j + \theta;\ \ \forall j \in J;\ d \in \{t^* + 1, \dots, |D|\}\right) \geq 1 - \tau, \qquad (5.73)$$

which is a relaxation of the joint chance constraint (5.1) built with the first-stage decisions. It states that the probability of the total amount of demand arriving at any facility on any day exceeding the processing capacity of the facility plus the maximum backlog level θ is less than τ. Clearly, if at some facility $j \in J$ on some day $d \in \{t^* + 1, \dots, |D|\}$, $\sum_{i\in I} \tilde{h}_{i,d-t_{ij}} Y_{ij} >\ k_j + \theta$ then there will be more than θ items in backlog at the beginning of day $d + 1$ since facility j can process at most k_j items per day and must hold the rest in backlog. As a result, the joint chance constraint (5.1) dominates the joint chance constraint (5.73).

The relaxed joint chance constraint (5.73) considers the amount of demand arriving at each facility on any given day, independent of the amount of demand that arrived in previous days or that will arrive in the future. For example, if $\theta = 20$ and $k_j = 100$, facility j could have 120 items of demand arrive every day and still not violate constraint (5.73). Yet, we know that 20 items will be added to the backlog queue every day, since at most 100 items can be processed each day. This means that the amount of total backlog held at facility j is a strictly increasing function and that there will be a day when we will exceed the maximum backlog level θ. Therefore, if the time horizon is long enough, we have $\mathbb{P}\left(\tilde{V}_{jd} \leq\ \theta;\ \ \forall j \in J;\ d \in \{t^* + 1, \dots, |D| + 1\}\right) = 0$ although $\mathbb{P}\left(\sum_{i\in I} \tilde{h}_{i,d-t_{ij}} Y_{ij} \leq\ k_j + \theta;\ \ \forall j \in J;\ d \in \{t^* + 1, \dots, |D|\}\right) = 1$.

We can further strengthen the formulation by expanding the relaxed joint chance constraint to couple demand arrivals for multiple days. Let C represent the set identifying the number of consecutive days to couple, then we formulate the relaxed joint chance constraint as:

$$\mathbb{P}\left(\begin{array}{c} \sum_{i\in I}\sum_{s=0}^{c-1}\tilde{h}_{i,d+s-t_{ij}}Y_{ij} \le ck_j+\theta; \\ \forall j\in J;\ c\in C;\ d\in\{t^*+1,\dots,|D|-c+1\} \end{array} \right) \ge 1-\tau \quad (5.74)$$

For example, $C=\{1,2\}$ indicates that we want to consider the amount of demand arriving at a facility on each day, as well as the amount of demand arriving within a two day span. Note that when $C=\{1\}$, constraint (5.74) is precisely (5.73). It is worth noting however that when $\{1\}\notin C$ it is possible for a scenario to generate a violation of $\sum_{i\in I}\sum_{s=0}^{c-1}\tilde{h}_{i,d+s-t_{ij}}Y_{ij}\le ck_j+\theta$ in constraint (5.74) for some $c\in C$, but not generate a violation of $\sum_{i\in I}\tilde{h}_{i,d-t_{ij}}Y_{ij}\le k_j+\theta$ in constraint (5.73). To see this, suppose $\theta=10$ and $k_j=100$ for some $j\in J$. If $C=\{2\}$ and facility j has a total of 110 arriving units of demand each day, then the one day constraint, $\sum_{i\in I}\tilde{h}_{i,d-t_{ij}}Y_{ij}\le k_j+\theta$, is satisfied (since $110\le 100+10$). However, the arriving demand over any two days totals 220 units, which causes a violation of $\sum_{i\in I}\sum_{s=0}^{c-1}\tilde{h}_{i,d+s-t_{ij}}Y_{ij}\le ck_j+\theta$ with $c=2$ (since $220>200+10$).

On the other hand, if for some scenario, $\sum_{i\in I}\sum_{s=0}^{c-1}\tilde{h}_{i,d+s-t_{ij}}Y_{ij}> ck_j+\theta$ for some number of coupled days, $c\in C$, at some facility $j\in J$ on some day $d\in\{t^*+1,\dots,|D|-c+1\}$, then there exists at least one facility j and one day d such that $V_{jd}^{\omega}>\theta$ for the corresponding ω in (5.1). As a result, the joint chance constraint (5.1) also dominates the joint chance constraint (5.74).

We can reformulate the joint chance constraint (5.74) by introducing the decision variable $U^{\omega'}$ which takes the value 0 if, for any $c\in C$, the arriving demand at any facility $j\in J$

exceeds its corresponding $ck_j + \theta$ on any day in scenario $\omega' \in \Omega'$. Letting M_3 represent a sufficiently large number, we can replace constraint (5.74) with:

$$M_3\left(1 - U^{\omega'}\right) \geq \sum_{i \in I} \sum_{s=0}^{c-1} h_{i,d+s-t_{ij}}^{\omega'} Y_{ij} - ck_j - \theta$$

$$\forall\, \omega' \in \Omega';\ j \in J;\ c \in C;\ d \in \{t^* + 1, \dots, |D| - c + 1\} \quad (5.75)$$

$$\sum_{\omega' \in \Omega'} p^{\omega'} U^{\omega'} \geq 1 - \tau \quad (5.76)$$

$$U^{\omega'} \in \{0,1\} \qquad \forall \omega' \in \Omega' \quad (5.77)$$

This formulation, however, adds an excessive number of constraints. For example, using this method on a data set with a time horizon of $|D| - t^*$ days, $|J|$ candidate facilities, and $|\Omega'|$ scenarios will add $(|D| - t^* - c + 1)|\Omega'||J|$ constraints of type (5.75) for each $c \in C$. Thus, if $C = \{2\}$, a data instance with 50 candidate facilities and ten demand scenarios results in 182,000 added constraints. If $C = \{2,3\}$, this number increases to 363,500. However, we can reduce the number of constraints by shifting the coupling so that rather than checking if every c days the arriving demand satisfies $\sum_{i \in I} \sum_{s=0}^{c-1} \tilde{h}_{i,d+s-t_{ij}} Y_{ij} \leq ck_j + \theta$ we instead check if non-overlapping couplings of c days satisfy $\sum_{i \in I} \sum_{s=0}^{c-1} \tilde{h}_{i,d+s-t_{ij}} Y_{ij} \leq ck_j + \theta$. For example, if $C = \{2\}$, instead of checking the amount of demand that arrives during days 1 and 2, 2 and 3, 3 and 4, 4 and 5, 5 and 6, etc., we would only generate constraints to check the demand arriving on days 1 and 2, 3 and 4, 5 and 6, etc. If $C = \{2,3\}$, we would generate constraints to check the amount of demand arriving on days 1 and 2, 2 and 3, 3 and 4, etc., as well as on days 1-3, 4-6, 7-9, etc. Then, such non-overlapping couplings can be achieved by specifying a constraint of type (5.75) $\forall c \in C$ such that $t^* + c \leq |D|;\ d \in \{t^* + 1, t^* + 1 + c, t^* + 1 + 2c, \dots, \left\lfloor \frac{|D| - t^* - c}{c} \right\rfloor c + t^* + 1\}$ instead of $\forall c \in C, d \in \{t^* + 1, \dots, |D| - c + 1\}$.

5.4 Additional Chance Constraints

The chance constraints that we have considered thus far are joint chance constraints; they ensure that the probability that there will be more than θ item03s in backlog on *any day* at *any facility* is no greater than τ. As such, the JCC approach does not distinguish between a scenario in which a single facility exceeds the desired backlog level θ by one unit on one day and a scenario in which multiple facilities have multiple days in which the backlog level greatly exceeds θ. This can be addressed by using individual chance constraints rather than joint chance constraints. This ensures the probability that there will be more than θ items in backlog is less than or equal to τ at *each facility* on *each day*. We can also formulate hybrid chance constraints that are less aggregate than the joint constraints but more integrated than the individual constraints; the hybrid chance constraints ensure that the probability that there will be more than θ items in backlog on *any day* is less than or equal to τ at each facility.

The individual chance constraints are formulated as

$$\mathbb{P}\left(\tilde{V}_{jd} \leq \theta\right) \geq 1 - \tau \qquad \forall j \in J;\ d \in \{t^* + 1, \dots, |D| + 1\} \tag{5.78}$$

while the hybrid chance constraints are:

$$\mathbb{P}\left(\tilde{V}_{jd} \leq \theta;\ \forall d \in \{t^* + 1, \dots, |D| + 1\}\right) \geq 1 - \tau \qquad \forall j \in J. \tag{5.79}$$

These constraints can be linearized and approximated though a finite subset of scenarios by introducing

$$\bar{Z}_{jd}^{\omega'} = \begin{cases} 1 & \text{If the backlog at facility } j \in J \text{ exceeds } \theta \text{ on day } d \in \{t^* + 1, \dots, |D|\} \text{ in} \\ & \text{scenario } \omega' \in \Omega' \\ 0 & \text{otherwise} \end{cases}$$

and

$$\hat{Z}_{j}^{\omega'} = \begin{cases} 1 & \text{If the backlog at facility } j \in J \text{ exceeds } \theta \text{ on any day in scenario } \omega' \in \Omega' \\ 0 & \text{otherwise} \end{cases}$$

We then replace the individual chance constraints (5.78) with

$$V_{jd}^{\omega'} - M_1 \bar{Z}_{jd}^{\omega'} \leq \theta \qquad \forall \omega' \in \Omega';\ j \in J;\ d \in \{t^* + 1, \dots, |D| + 1\} \quad (5.80)$$

$$\sum_{\omega' \in \Omega'} p^{\omega'} \bar{Z}_{jd}^{\omega'} \leq \tau \qquad \forall j \in J;\ d \in \{t^* + 1, \dots, |D| + 1\} \quad (5.81)$$

$$\bar{Z}_{jd}^{\omega'} \in \{0,1\} \qquad \forall \omega' \in \Omega';\ j \in J;\ d \in \{t^* + 1, \dots, |D| + 1\} \quad (5.82)$$

and the hybrid chance constraints (5.79) with

$$V_{jd}^{\omega'} - M_1 \hat{Z}_{j}^{\omega'} \leq \theta \qquad \forall \omega' \in \Omega';\ j \in J;\ d \in \{t^* + 1, \dots, |D| + 1\} \quad (5.83)$$

$$\sum_{\omega' \in \Omega'} p^{\omega'} \hat{Z}_{j}^{\omega'} \leq \tau \qquad \forall j \in J \quad (5.84)$$

$$\hat{Z}_{j}^{\omega'} \in \{0,1\} \qquad \forall \omega' \in \Omega';\ j \in J. \quad (5.85)$$

We refer to the problem consisting of (2.2) - (2.4), (3.6), (5.15) - (5.19) and (5.80) - (5.82) as the Mixed Integer Linear Programming Reformulation of the individual chance constrained Program [MIP-ICC] and refer to (2.2) - (2.4), (3.6), (5.15) - (5.19) and (5.83) - (5.85) as the Mixed Integer Linear Programming Reformulation of the hybrid chance constrained Program [MIP-HCC].

Table 25: Comparison of the types of chance constraints

Joint Chance Constraints

P(more than θ items in backlog on **any day** at **any facility**)$\leq \tau$

$$\mathbb{P}\left(\tilde{V}_{jd} \leq \theta;\ \forall j \in J,\ d \in \{t^* + 1, \dots, |D| + 1\}\right) \geq 1 - \tau$$

Hybrid Chance Constraints

P(more than θ items in backlog on **any day**)$\leq \tau$ at **each facility**

$$\mathbb{P}\left(\tilde{V}_{jd} \leq \theta;\ \forall d \in \{t^* + 1, \dots, |D| + 1\}\right) \geq 1 - \tau \qquad \forall j \in J$$

Individual Chance Constraints

P(more than θ items in backlog)$\leq \tau$ at **each facility each day**

$$\mathbb{P}\left(\tilde{V}_{jd} \leq \theta\right) \geq 1 - \tau \qquad \forall j \in J,\ d \in \{t^* + 1, \dots, |D| + 1\}$$

5.4.1 Two Stage Decomposition

Obtaining the two stage decomposition for the individual and hybrid chance constrained formulations requires only a few modifications to the two stage MIP-JCC approach. In particular, since the individual chance constraints assess the amount of backlog at each facility on each day, constraints (5.11) and (5.12) are replaced by (5.81) and (5.82) in M1S1, and constraints (5.29) are replaced by

$$V_{jd}^{\omega'} \leq \theta + M_1 \bar{Z}_{jd}^{\omega'} \qquad \forall d \in \{t^* + 1, \dots, |D| + 1\} \tag{5.86}$$

in M1S2(ω', j) to obtain the two stage MIP-ICC decomposition. Similarly, constraints (5.84), (5.85), and

$$V_{jd}^{\omega'} \leq \theta + M_1 \hat{Z}_{j}^{\omega'} \qquad \forall d \in \{t^* + 1, \dots, |D| + 1\} \tag{5.87}$$

take the place of constraints (5.11), (5.12), and (5.29) to obtain the two stage MIP-HCC decomposition. The remaining steps of the decompositions are straightforward, as they follow the method described in Section 5.3.1.

5.4.2 Three Stage Decomposition

In this section we discuss the amendments that must be made to the Three Stage Decomposition Approach when using it to solve the Individual or hybrid chance constrained formulations.

Individual Chance Constraints

The three stage MIP-ICC decomposition approach is similar to that of the MIP-JCC. In fact, the first and second stage problems are identical to M2S1 and M2S2. However, the

reformulated first stage problem for the individual chance constrained formulation, which incorporates the feasibility and optimality cuts, is slightly different from R-M2S1. As such, we define $\underline{\widehat{\Psi}_j}$ as the set of infeasibility cuts generated for facility $j \in J$ and $\bar{I}_j^{\psi} \coloneqq \{i \in I: Y_{ij}^{\psi} = 1\}$ as the set of demand sites allocated to processing facility $j \in J$ in infeasible solution $\psi \in \underline{\widehat{\Psi}_j}$. We refer to the reformulated M2S1 problem for MIP-ICC as MIP-ICC_R-M2S1:

[MIP-ICC_R-M2S1]:

$$Min_{\boldsymbol{X},\boldsymbol{Y},\widehat{\mathbb{V}}} \ (|D| - t^*) \sum_{j \in J} f_j X_j + a \sum_{i \in I} \sum_{d=t^*+1}^{|D|} \mathbb{E}_{\xi'}\left[\tilde{h}_{id}\right] \sum_{j \in J} t_{ij} Y_{ij} + b\widehat{\mathbb{V}} \tag{5.88}$$

Subject to

$$\sum_{j \in J} Y_{ij} = 1 \qquad \forall i \in I \tag{2.2}$$

$$Y_{ij} - X_j \leq 0 \qquad \forall i \in I; j \in J \tag{2.3}$$

$$\sum_{j \in J} X_j \geq 1 \tag{5.22}$$

$$X_j \in \{0,1\} \qquad \forall j \in J \tag{2.4}$$

$$Y_{ij} \in \{0,1\} \qquad \forall i \in I; j \in J \tag{3.6}$$

Initialization cuts

$$\sum_{i \in \bar{I}_j^{\psi}} Y_{ij} \leq \left|\bar{I}_j^{\psi}\right| - 1 \qquad \forall j \in J, \psi \in \underline{\widehat{\Psi}_j} \tag{5.89}$$

$$\sum_{\omega' \in \Omega'} p^{\omega'} * \left(\sum_{j \in J} \sum_{d=t^*+1}^{|D|} \sum_{i \in I} h_{i,d-t_{ij}}^{\omega'} Y_{ij} \gamma_{jd}^{\omega' \phi} + \sum_{j \in J} \sum_{d=t^*+1}^{|D|} k_j X_j \pi_{jd}^{\omega' \phi} \right.$$

$$\left. + \sum_{j \in J} v_j X_j \eta_j^{\omega' \phi} \right) \leq \widehat{\mathbb{V}} \qquad \forall \phi \in \underline{\widehat{\Phi}} \tag{5.66}$$

The feasibility cuts (5.89) ensure a future optimal solution allocates demand sites to each facility $j \in J$ in a manner that is different from any previous allocations that have violated the corresponding individual chance constraint. We present a proof that constraints (5.89) are valid feasibility cuts in Appendix C. As before, constraints (5.66) are optimality cuts.

Additionally, the third stage problem of the Individual Chanced Constrained formulation is specified for each facility, j, and each day, $d \in \{t^* + 1, \dots, |D| + 1\}$. Using the binary variable $\bar{\mathbb{Q}}$, which takes the value one if the location, allocation, processing, and backlog variables do not satisfy (5.80) - (5.82), it checks the feasibility of the first and second stage solution against the individual chance constraint:

[MIP-ICC_M2S3(j, d)]:

$$\bar{\hat{Z}}(\boldsymbol{V}) = Min_{\bar{\boldsymbol{z}}_{jd}, \bar{\mathbb{Q}}} \ \bar{M}_2 \bar{\mathbb{Q}} \tag{5.90}$$

Subject to

$$M_1 \bar{Z}_{jd}^{\omega'} \geq V_{jd}^{\omega'} - \theta \qquad \forall \omega' \in \Omega' \tag{5.91}$$

$$\sum_{\omega' \in \Omega'} p^{\omega'} \bar{Z}_{jd}^{\omega'} - \bar{\mathbb{Q}} \leq \tau \tag{5.92}$$

$$\bar{\mathbb{Q}} \geq 0 \tag{5.93}$$

$$\bar{Z}_{jd}^{\omega'} \in \{0,1\} \qquad \forall \omega' \in \Omega' \tag{5.94}$$

Similarly to the joint chance constrained algorithm, $\bar{M}_2$ is a sufficiently large integer such that M2S1 will not choose an allocation that results in $\sum_{\omega' \in \Omega'} p^{\omega'} \bar{Z}_{jd}^{\omega'} > \tau$ for any $j \in J$, $d \in \{t^* + 1, \dots, |D| + 1\}$. This feasibility check is described in Algorithm 5, and is used in Algorithm 6, which details the three-stage cutting plane algorithm for the individual chance constrained model.

Algorithm 5: Feasibility check for the MIP-ICC model

For each $j \in J$, and $d \in \{t^* + 1, \dots, |D| + 1\}$, given a feasible solution $(\boldsymbol{X}^n, \boldsymbol{Y}^n, \boldsymbol{V}^n, \boldsymbol{W}^n)$ to the n^{th} iteration of MIP-ICC_R-M2S1 and M2S2(ω', j),

a. Initialize $\bar{Z}_{jd}^{\omega' n} = 0, \forall \omega' \in \Omega'$

b. For each $\omega' \in \Omega'$ check whether $V_{jd}^{\omega' n} \leq \theta$.

- If the inequality does not hold, set the corresponding $\bar{Z}_{jd}^{\omega' n} = 1$

c. Check whether $\sum_{\omega' \in \Omega'} p^{\omega'} \bar{Z}_{jd}^{\omega' n} \leq \tau$.

- If the inequality does not hold, then $(\boldsymbol{X}^n, \boldsymbol{Y}^n, \boldsymbol{V}^n, \boldsymbol{W}^n, \boldsymbol{Z}^n)$ is an infeasible solution to MIP-ICC and the allocation to $j \in J$ contributes to the infeasibility. Cut off the current allocation to j from the solution space by adding the constraint

$$\sum_{i \in \bar{I}_j^n} Y_{ij} \leq \left|\bar{I}_j^n\right| - 1 \tag{5.95}$$

to constraint set $\underline{\widehat{\Psi}_j}$. Here $\bar{I}_j^n \coloneqq \{i \in I : Y_{ij}^n = 1\}$ represents the set of demand sites that allocate demand to processing facility $j \in J$ as obtained from the solution of MIP-ICC_R-M2S1.

Algorithm 6: Three-stage cutting plane algorithm for the MIP-ICC model

---

Step 0: Initialize $n = 1, \underline{\widehat{\Psi}_j} = \emptyset \; \forall j \in J, \underline{\widehat{\Phi}} = \emptyset, UB = \infty,$ and $LB = -\infty$. Add an appropriate initialization cut (e.g., $\widehat{\mathbb{V}} \geq 0$) to MIP-ICC_R-M2S1.

Step 1: Solve MIP-ICC_R-M2S1 and let $(\boldsymbol{X}^n, \boldsymbol{Y}^n, \widehat{\mathbb{V}}^n)$ denote an optimal solution. Set the LB equal to the current value of (5.88).

Step 2: For each $\omega' \in \Omega'$ and $j \in J$, use Algorithm 2 with $(\boldsymbol{X}^n, \boldsymbol{Y}^n)$ as input to solve M2S2(ω', j). Denote the aggregated optimal solution to M2S2(ω', j) by $(\boldsymbol{V}^n, \boldsymbol{W}^n)$.

Step 3: [Feasibility Check] Using $(\boldsymbol{X}^n, \boldsymbol{Y}^n, \boldsymbol{V}^n, \boldsymbol{W}^n)$ as input, perform Algorithm 5 to determine if $(\boldsymbol{X}^n, \boldsymbol{Y}^n, \boldsymbol{V}^n, \boldsymbol{W}^n)$ is a feasible solution to MIP-ICC.

If $\sum_{\omega' \in \Omega'} p^{\omega'} \bar{Z}_{jd}^{\omega' n} > \tau$ for any $j \in J$, $d \in \{t^* + 1, \ldots, |D| + 1\}$, then $(\boldsymbol{X}^n, \boldsymbol{Y}^n, \boldsymbol{V}^n, \boldsymbol{W}^n)$ is an infeasible solution to MIP-ICC, set $n = n + 1$ and return to Step 1. Otherwise, $(\boldsymbol{X}^n, \boldsymbol{Y}^n, \boldsymbol{V}^n, \boldsymbol{W}^n, \boldsymbol{Z}^n)$ is a feasible solution to MIP-ICC. As such, it produces an upper bound on (5.2). Set

$$UB = \min(UB, (|D| - t^*) \sum_{j \in J} f_j X_j^n + a \sum_{i \in I} \sum_{d=t^*+1}^{|D|} \mathbb{E}_{\xi'}\left[\tilde{h}_{id}\right] \sum_{j \in J} t_{ij} Y_{ij}^n$$

$$+ b \sum_{\omega' \in \Omega'} p^{\omega'} \left(\sum_{d=t^*+1}^{|D|+1} \sum_{j \in J} V_{jd}^{\omega' n}\right))$$

Step 4: [Optimality Check] For each $\omega' \in \Omega'$ and $j \in J$, solve D-M2S2(ω', j) with $(\boldsymbol{X}, \boldsymbol{Y}) = (\boldsymbol{X}^n, \boldsymbol{Y}^n)$. Denote the aggregated optimal solutions by $(\boldsymbol{\gamma}^{\omega' \phi}, \boldsymbol{\pi}^{\omega' \phi}, \boldsymbol{\eta}^{\omega' \phi}, \boldsymbol{\mu}^{\omega' \phi})$.

c) If

$$(\hat{\mathbb{V}}^n < \sum_{\omega' \in \Omega'} p^{\omega'} \left(\sum_{j \in J} \sum_{d=t^*+1}^{|D|} \sum_{i \in I} h_{i,d-t_{ij}}^{\omega'} Y_{ij}^n \gamma_d^{\omega' j \phi} + \sum_{j \in J} \sum_{d=t^*+1}^{|D|} k_j X_j^n \pi_d^{\omega' j \phi}\right.$$

$$\left. + \sum_{j \in J} v_j X_j^n \eta^{\omega' j \phi}\right) \quad (5.72)$$

add $\left(\boldsymbol{\gamma}^{\omega' \phi}, \boldsymbol{\pi}^{\omega' \phi}, \boldsymbol{\eta}^{\omega' \phi}, \boldsymbol{\mu}^{\omega' \phi}\right)$ to $\underline{\hat{\Phi}}$. Let $n = n + 1$ and return to Step 1.

d) If the inequality (5.72) does not hold, then $(\boldsymbol{X}^n, \boldsymbol{Y}^n, \boldsymbol{W}^n, \boldsymbol{V}^n, \boldsymbol{Z}^n)$ is the optimal solution to MIP-ICC. The algorithm terminates.

Hybrid Chance Constraints

As with the MIP-ICC, the first and second stage MIP-HCC problems are identical to M2S1 and M2S2 from the MIP-JCC. Furthermore, the structure of the feasibility and optimality cuts used in the reformulated M2S1 problem for MIP-HCC is identical to those of the MIP-ICC (the difference between the two models is in which particular cuts are generated). However, rather than using a single third stage problem (as in the joint chance constrained formulation) or a third stage problem for every processing facility and every day (as in the individual chance

constrained formulation), the third stage problem of the MIP-HCC formulation is specified only for each processing facility:

[MIP-HCC_M2S3(j)]:

$$\hat{\hat{Z}}(\boldsymbol{V}) = Min_{\hat{\boldsymbol{Z}}_j, \hat{\mathbb{Q}}} \ \widehat{M}_2 \hat{\mathbb{Q}} \tag{5.96}$$

Subject to

$$M_1 \hat{Z}_j^{\omega'} \geq V_{jd}^{\omega'} - \theta \qquad \forall \omega' \in \Omega';\ d \in \{t^* + 1, \dots, |D| + 1\} \tag{5.97}$$

$$\sum_{\omega' \in \Omega'} p^{\omega'} \hat{Z}_j^{\omega'} - \hat{\mathbb{Q}} \leq \tau \tag{5.98}$$

$$\hat{\mathbb{Q}} \geq 0 \tag{5.99}$$

$$\hat{Z}_j^{\omega'} \in \{0,1\} \qquad \forall \omega' \in \Omega' \tag{5.100}$$

As before, $\widehat{M}_2$ is a sufficiently large integer such that M2S1 will not choose an allocation that results in $\sum_{\omega' \in \Omega'} p^{\omega'} \hat{Z}_j^{\omega'} > \tau$ for any $j \in J$, and $\hat{\mathbb{Q}}$ is a binary variable that takes the value one in the event that a particular backlog solution $\boldsymbol{V}$ does result in $\sum_{\omega' \in \Omega'} p^{\omega'} \hat{Z}_j^{\omega'} > \tau$. Thus, the MIP-HCC can be solved using a three stage decomposition approach that is identical to Algorithm 6 with a few minor changes due to references to MIP-ICC and the details of the Feasibility Check. Algorithm 7 provides the details for the feasibility chick while Algorithm 8 provides the details for the MIP-HCC three stage decomposition.

Algorithm 7: Feasibility check for the MIP-HCC model

For each $j \in J$ given a feasible solution $(\boldsymbol{X}^n, \boldsymbol{Y}^n, \boldsymbol{V}^n, \boldsymbol{W}^n)$ to the n^{th} iteration of MIP-ICC_R-M2S1 and M2S2(ω', j),

a. Initialize $\hat{Z}_j^{\omega' n} = 0, \forall \omega' \in \Omega'$

b. For each $\omega' \in \Omega'$ check whether $V_{jd}^{\omega' n} \leq \theta, \forall d \in \{t^* + 1, \dots, |D| + 1\}$.

- If the inequality does not hold, set the corresponding $\hat{Z}_j^{\omega' n} = 1$

c. Check whether $\sum_{\omega' \in \Omega'} p^{\omega'} \hat{Z}_j^{\omega' n} \leq \tau$.

- If the inequality does not hold, then $(\boldsymbol{X}^n, \boldsymbol{Y}^n, \boldsymbol{V}^n, \boldsymbol{W}^n, \boldsymbol{Z}^n)$ is an infeasible solution to MIP-HCC and the allocation to $j \in J$ contributes to the infeasibility. Cut off the current allocation to j from the solution space by adding the constraint

$$\sum_{i \in \bar{I}_j^n} Y_{ij} \leq \left|\bar{I}_j^n\right| - 1 \tag{5.95}$$

to constraint set $\underline{\widehat{\Psi}_j}$. Here $\bar{I}_j^n \coloneqq \{i \in I : Y_{ij}^n = 1\}$ represents the set of demand sites that allocate demand to processing facility $j \in J$ as obtained from the solution of MIP-ICC_R-M2S1.

Algorithm 8: Three-stage cutting plane algorithm for the MIP-HCC model

---

Step 0: Initialize $n = 1, \underline{\widehat{\Psi}_j} = \emptyset \; \forall j \in J, \underline{\widehat{\Phi}} = \emptyset, UB = \infty,$ and $LB = -\infty$. Add an appropriate initialization cut (e.g., $\widehat{\mathbb{V}} \geq 0$) to MIP-ICC_R-M2S1.

Step 1: Solve MIP-ICC_R-M2S1and let $(\boldsymbol{X}^n, \boldsymbol{Y}^n, \widehat{\mathbb{V}}^n)$ denote an optimal solution. Set the LB equal to the current value of (5.88).

Step 2: For each $\omega' \in \Omega'$ and $j \in J$, use Algorithm 2 with $(\boldsymbol{X}^n, \boldsymbol{Y}^n)$ as input to solve M2S2(ω, j). Denote the aggregated optimal solution to M2S2(ω', j) by $(\boldsymbol{V}^n, \boldsymbol{W}^n)$**.**

Step 3: [Feasibility Check] Using $(\boldsymbol{X}^n, \boldsymbol{Y}^n, \boldsymbol{V}^n, \boldsymbol{W}^n)$ as input, perform Algorithm 7 to determine if $(\boldsymbol{X}^n, \boldsymbol{Y}^n, \boldsymbol{V}^n, \boldsymbol{W}^n)$ is a feasible solution to MIP-HCC.

If $\sum_{\omega' \in \Omega'} p^{\omega'} \hat{Z}_j^{\omega' n} > \tau$ for any $\in J$ $(\boldsymbol{X}^n, \boldsymbol{Y}^n, \boldsymbol{V}^n, \boldsymbol{W}^n)$ is an infeasible solution to MIP-HCC, set $n = n + 1$ and return to Step 1. Otherwise, $(\boldsymbol{X}^n, \boldsymbol{Y}^n, \boldsymbol{V}^n, \boldsymbol{W}^n, \boldsymbol{Z}^n)$ is a feasible solution to MIP-HCC. As such, it produces an upper bound on (5.2). Set

$$UB = \min(UB, (|D| - t^*) \sum_{j \in J} f_j X_j^n + a \sum_{i \in I} \sum_{d=t^*+1}^{|D|} \mathbb{E}_{\xi'}[\tilde{h}_{id}] \sum_{j \in J} t_{ij} Y_{ij}^n$$

$$+ b \sum_{\omega' \in \Omega'} p^{\omega'} (\sum_{d=t^*+1}^{|D|+1} \sum_{j \in J} V_{jd}^{\omega' n}))$$

Step 4: [Optimality Check] For each $\omega' \in \Omega'$ and $j \in J$, solve D-M2S2(ω', j) with $(\boldsymbol{X}, \boldsymbol{Y}) = (\boldsymbol{X}^n, \boldsymbol{Y}^n)$. Denote the aggregated optimal solutions by $(\boldsymbol{\gamma}^{\omega' \phi}, \boldsymbol{\pi}^{\omega' \phi}, \boldsymbol{\eta}^{\omega' \phi}, \boldsymbol{\mu}^{\omega' \phi})$.

a) If

$$(\widehat{\mathbb{V}}^n < \sum_{\omega' \in \Omega'} p^{\omega'} \Big(\sum_{j \in J} \sum_{d=t^*+1}^{|D|} \sum_{i \in I} h_{i,d-t_{ij}}^{\omega'} Y_{ij}^n \gamma_d^{\omega' j \phi} + \sum_{j \in J} \sum_{d=t^*+1}^{|D|} k_j X_j^n \pi_d^{\omega' j \phi}$$

$$+ \sum_{j \in J} v_j X_j^n \eta^{\omega' j \phi} \Big) \quad (5.72)$$

add $\big(\boldsymbol{\gamma}^{\omega' \phi}, \boldsymbol{\pi}^{\omega' \phi}, \boldsymbol{\eta}^{\omega' \phi}, \boldsymbol{\mu}^{\omega' \phi}\big)$ to $\underline{\widehat{\Phi}}$. Let $n = n + 1$ and return to Step 1.

b) If the inequality (5.72) does not hold, then $(\boldsymbol{X}^n, \boldsymbol{Y}^n, \boldsymbol{W}^n, \boldsymbol{V}^n, \boldsymbol{Z}^n)$ is the optimal solution to MIP-HCC. The algorithm terminates.

5.5 Computational Results

We present results relating to the optimal facility locations and computational efficiency of the solution algorithms in this chapter. While our focus is on results related to the joint chance constraints, we briefly describe the effect of imposing the individual or hybrid chance constraints rather than the joint chance constraints in Section 5.5.4.

The instances used in this chapter were generated in a manner similar to the one described in Section 3.4 with a few minor changes. Rather than using a weight of $w = \$0.0001$ (as in 3.4), we use $w = \$0.01$ in the facility location cost calculations below. Additionally, a time horizon of 365 days is employed (rather than the 100 day time horizon previously used) and

the daily processing capacity of each candidate processing facility is one fifth of the total average daily demand, rounded up to the nearest integer.

Unless otherwise noted, the computational results presented in this chapter assume that the desired maximum backlog (i.e., θ) is 5% of a processing facility's processing capacity. Since discrete, daily demand is generated for a time horizon of 365 days for ten demand scenarios, these restrictions correspond to ensuring that over the course of a ten year time horizon, we allow the amount of backlog to exceed the maximum desired level in no more than one year. It is worth noting that due to the nature of the joint chance constraint (5.1), we do not distinguish between a scenario in which one processing facility exceeds the maximum backlog level by one unit on one day of the year and a scenario in which every processing facility exceeds the maximum backlog level on most days of the year; each of these scenarios generates one violation of the joint chance constraint. Furthermore, we assume the scenarios occur with equal probability.

5.5.1 Determining the Appropriate Number of Scenarios

Stochastic models that consist of an infinite or large number of scenarios are often computationally intractable and therefore are commonly approximated by only a subset of the possible scenarios. However, this process of reducing the number of scenarios considered (referred to as scenario reduction) requires tradeoffs. On one hand, we would like the number of scenarios to be relatively modest so that the resulting model can be solved with reasonable computational effort. However, we also want to ensure that the resulting solution approximates the original problem with reasonable accuracy. In this section, we discuss our findings regarding the tradeoff between these two objectives for the joint chance constrained problem. That is, we

determine how many scenarios to consider so that the resulting MIP-JCC solution is both close to the true optimal JCC solution and solves within a reasonable amount of time.

To observe the effect of the scenario size on the objective function value, for each demand site $i \in I$ and scenario size $|\Omega'| \in \{10, 25, 50, 75, 100, 125, 150\}$, we created $\mathcal{I} = 20$ independent and identically distributed (i.i.d.) instances of daily demand realizations from a Poisson distribution with mean $\frac{1}{10{,}000}\rho_i$ (see Section 3.4). That is, we solved 20 i.i.d. instances that each had ten scenarios, 20 i.i.d. instances that each had 25 scenarios, 20 i.i.d. instances that each had 50 scenarios, etc. Parameters $a = 5,\ b = 10,\ \theta = 91$ and $\tau = 0.1$ were used, and the problem instances contained 50 demand and candidate notes.

For each value of the number of scenarios considered, Figure 22 plots the optimal objective function value for each instance (using a gray diamond) as well as two standard deviations from the mean objective function value for a particular sample number of scenarios (using a black line). We note that the vertical axis ranges from \$101,235,000 to \$101,285,000 (rather than starting at zero). We also plot the percent difference between the upper and lower bounds on two standard deviations from the mean objective function value for each scenario size (Figure 23). Our results indicate that there is a 0.039% difference between the minimum and maximum objective function values (of the 20 instances considered) when only ten scenarios are used. This percentage decreases to 0.009% when 150 scenarios are used. Thus, there is relatively little variation in the optimal objective function value even when only ten scenarios are used. Furthermore, we note that the optimal first stage decisions (i.e., the location and allocation decisions) are the same in all of the instances depicted in Figures 22 and 23. The facilities located are in DuPage, IL; Fresno, CA; Palm Beach, FL; San Bernardino, CA; Travis, TX; and Westchester, NY.

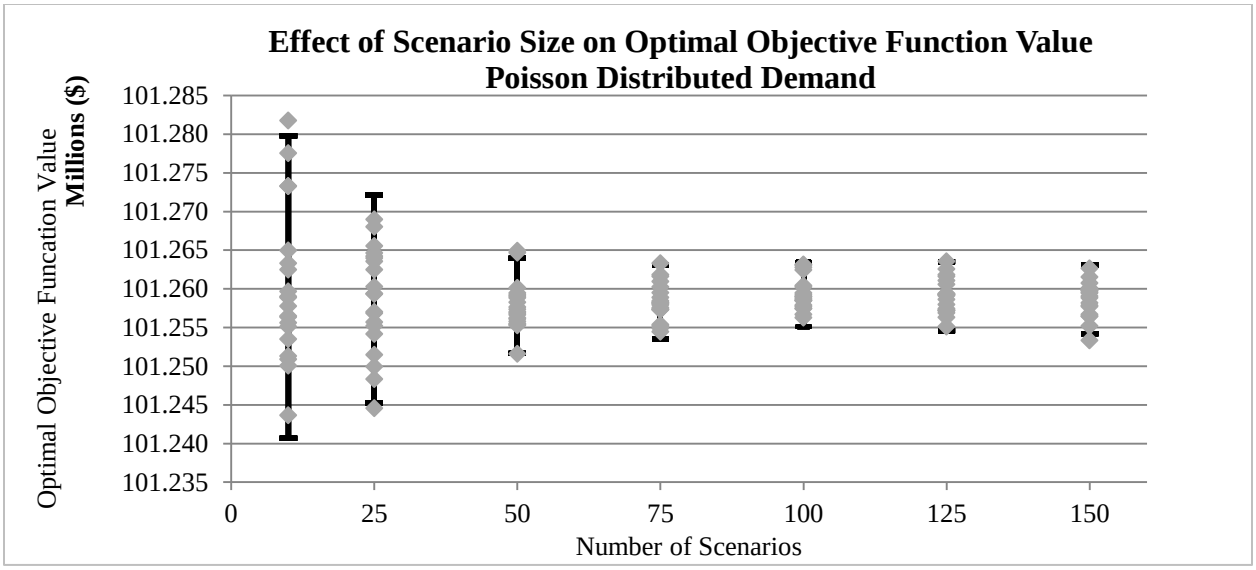

Figure 22: Variability in optimal objective function value as number of scenarios changes; 20 instances solved for each scenario size; Poisson distributed demand

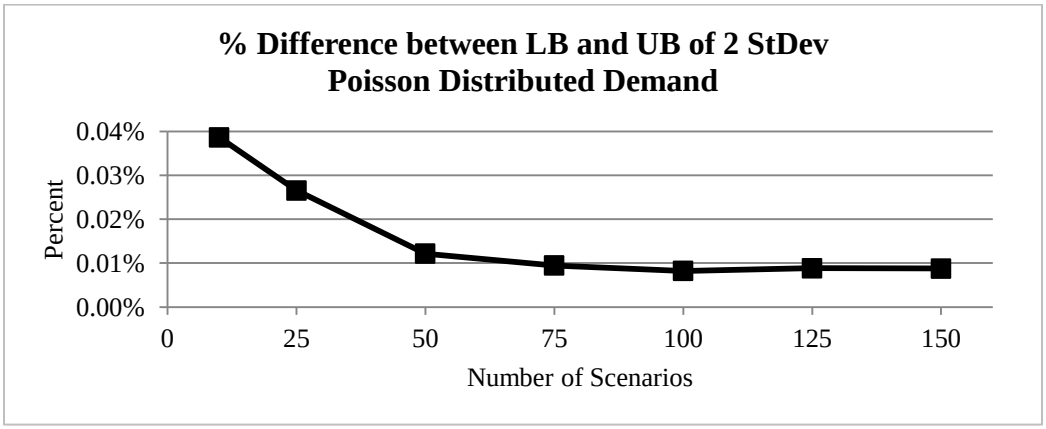

Figure 23: Percent difference between the lower and upper bounds on two standard deviations from the mean optimal objective function value; 20 instances solved for each scenario size; Poisson distributed demand

While the optimal location *and allocation* decisions are stable for all of the instances described in the previous paragraph, note that the demand at $i \in I$ was generated from a Poisson distribution with a standard deviation equal to $\sqrt{\frac{1}{10{,}000}\rho_i}$. To investigate whether the stability of the solution is an artifact of the low variability in demand, we also present results for demands that are generated from a Normal Distribution (truncated at zero) with a larger standard deviation of daily demand. While the mean of the Truncated Normal distribution for demand site $i \in I$ is

$\frac{1}{10,000}\rho_i$ (the same as the mean for the demand generated from the Poisson distribution), we set the standard deviation of the Truncated Normal Distribution to twice that of the corresponding Poisson distribution, (i.e., $2\sqrt{\frac{1}{10,000}\rho_i}$ in the Truncated Normal distribution).

As the results in Figures 24 and 25 show, the variation in the optimal objective function remains around 1.5% as the number of scenarios varies from 10 to 100 in the instances with the Truncated Normal demand distribution. Most optimal solutions locate processing facilities in DuPage, IL; Fresno CA; Orange, FL; San Bernardino, CA; Travis, TX; and Westchester NY. However, a small number of instances locate a facility in Milwaukee, WI instead of Du Page, IL or in Contra Costa, CA instead of Fresno, CA (regardless of the number of scenarios considered). Additionally, one instance of the 75 scenarios case locates seven facilities (Pima, AZ in addition to the six common facilities above).

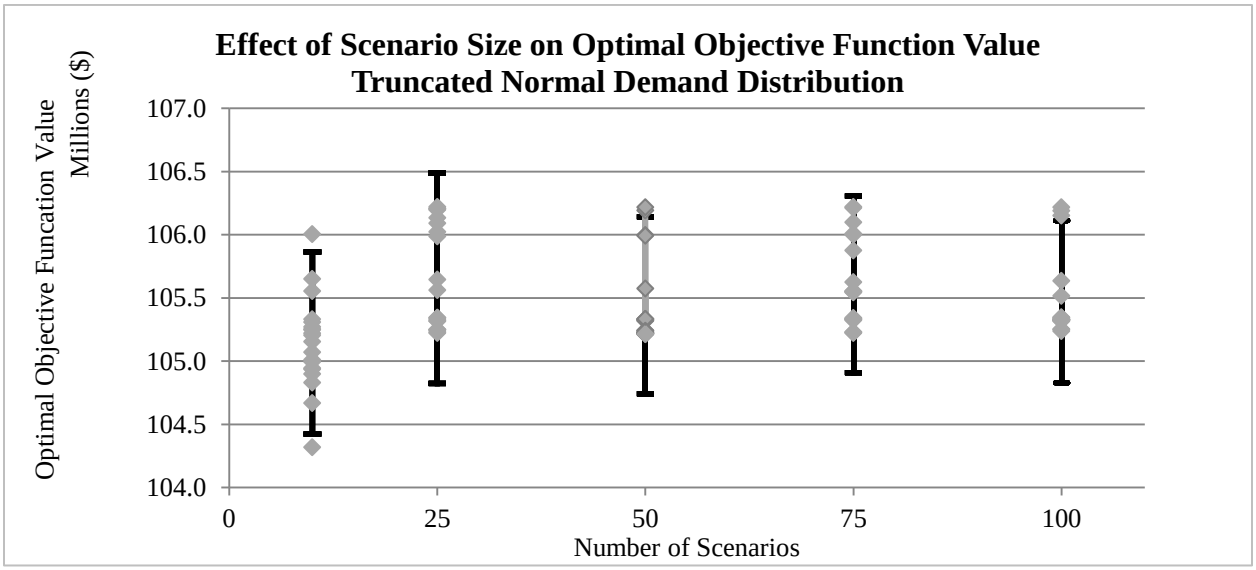

Figure 24: Variability in optimal objective function values as number of scenarios changes; 20 instances solved for each scenario size; Truncated Normal demand distribution

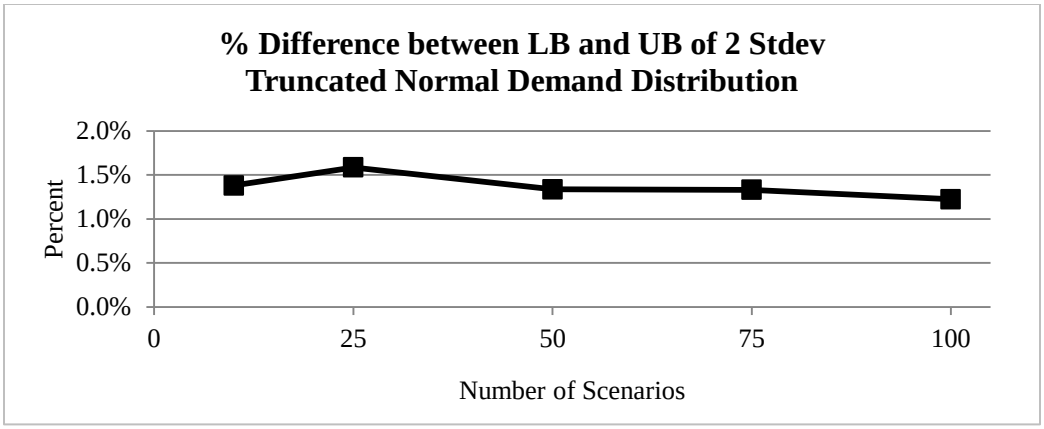

Figure 25: Percent difference between the lower bound and upper bounds on two standard deviations from the mean optimal objective function value; 20 instances solved for each scenario size; Truncated Normal demand distribution

As expected, a comparison of the stability in the optimal location and allocation decisions for the Poisson distributed demand instances and lack thereof for the instances with Truncated Normal demand distribution suggests that the number of scenarios necessary for a stable solution depends on the variability of the demand distribution. As the demand variability increases, more scenarios are needed to have high confidence that the location and allocation decisions obtained from the stochastic IMCLP solved using a subset of all possible demand scenarios, $\Omega' \subseteq \Omega$, is an optimal solution to the stochastic IMCLP with Ω.

Objective Function Lower Bound

While this analysis of the number of scenarios provides insight into the range of solutions obtained using a subset of $|\Omega'|$ scenarios in MIP-JCC, it is also important to assess the likelihood that such a solution generates a lower bound on the true JCC objective function value, as well as the likelihood that the approximated solution is a feasible solution to the JCC. For the discussions that follow, let τ refer to the acceptable exceedance probability in JCC and τ' the (perhaps different) acceptable exceedance probability used in MIP-JCC with $|\Omega'|$ sceanrios.

According to Luedtke and Ahmed (2007), if an optimal solution to JCC with τ exists and the corresponding optimal objective function value is finite, then the probability that the optimal MIP-JCC objective function value obtained with a finite subset of $|\Omega'|$ demand scenarios and an acceptable exceedance probability of τ' is a lower bound on the optimal JCC objective function value with exceedance probability τ is at least:

$$\sum_{i=0}^{\lfloor \tau' |\Omega'| \rfloor} \binom{|\Omega'|}{i} \tau^i (1-\tau)^{|\Omega'|-i}$$

Thus, a lower bound on the probability that the optimal MIP-JCC objective function value obtained with $\tau' = 0.1$ is a lower bound on the true JCC objective function value when $\tau \in \{0.01, 0.05, 0.10\}$ for a range of $|\Omega'|$ is given in Table 26. Furthermore, Luedtke and Ahmed (2007) prove that for $\tau' > \tau$, the sample approximation yields a lower bound with probability approaching one exponentially fast as $|\Omega'|$ increases.

Table 26: Lower bound on the probability that the optimal MIP-JCC objective function value obtained from a subset of $|\Omega'|$ demand scenarios with $\tau' = 0.1$ is a lower bound on the optimal JCC objective function value with τ

Number of Scenarios $\|\Omega'\|$	Probability		
	$\tau = 0.01$	$\tau = 0.05$	$\tau = 0.10$
10	0.995733799757	0.913861644101	0.736098929100
25	0.998049323110	0.872893504339	0.537094050051
50	0.999989103168	0.962223827010	0.616123007724
75	0.999999070316	0.966371998940	0.520808618577
100	0.999999993744	0.988527589933	0.583155512266
125	0.999999999460	0.989993866568	0.516016265788
150	0.999999999996	0.996396855972	0.568184357974

However, using the order statistics from $\mathcal{I}$ replications of $|\Omega'|$ independent scenarios of $\xi(\omega)$, $\omega \in \Omega$, provides a significantly more powerful lower bound on the optimal JCC objective function value. Let $\xi'(\omega', i)$ denote the vector of demand realizations for scenario $\omega' \in \Omega'$ in replication $i \in \{1, 2, \ldots, \mathcal{I}\}$. For a given τ' and for each i, solve the associated sample approximation problem MIP-JCC with $\xi'(\omega', i)\ \forall \omega' \in \Omega'$; denote the optimal objective function

value by $\ddot{V}^{\tau'\omega'i}$. Then, rearrange the values $\left\{\ddot{V}^{\tau'\omega'i}\right\}_{i=1}^{\mathcal{I}}$ to obtain the order statistics $\ddot{V}^{\tau'\omega'[i]}$ for $i = 1, \ldots, \mathcal{I}$ satisfying $\ddot{V}^{\tau'\omega'[1]} \leq \cdots \leq \ddot{V}^{\tau'\omega'[\mathcal{I}]}$. Luedtke and Ahmed (2007) prove $\ddot{V}^{\tau'\omega'[1]}$ is a lower bound for the optimal JCC objective function value with at least $1-\delta$ confidence if $\left(1 - \sum_{i=0}^{\lfloor \tau' |\Omega'| \rfloor} \binom{|\Omega'|}{i} \tau^i (1-\tau)^{|\Omega'|-i}\right)^{\mathcal{I}} \leq \delta$. Thus, the minimum objective function value taken over the $\mathcal{I} = 20$ replications for each scenario size presented in Figures 22 and 24 is a lower bound on the corresponding optimal JCC objective function value with at least $1 - \left(1 - \sum_{i=0}^{\lfloor \tau' |\Omega'| \rfloor} \binom{|\Omega'|}{i} \tau^i (1-\tau)^{|\Omega'|-i}\right)^{\mathcal{I}}$ confidence. As is evident from Table 27 (which displays the value of $\delta = \left(1 - \sum_{i=0}^{\lfloor \tau' |\Omega'| \rfloor} \binom{|\Omega'|}{i} \tau^i (1-\tau)^{|\Omega'|-i}\right)^{\mathcal{I}}$ rather than $1-\delta$), the confidence is close to one even for the $|\Omega'| = 10$ scenario case.

Table 27: The minimum MIP-JCC objective function value obtained from $\mathcal{I} = 20$ repetitions of a subset of $|\Omega'|$ demand scenarios is a lower bound on the optimal objective function value of JCC (with associated acceptable exceedance probability of τ) with confidence $1-\delta$

Number of Scenarios $\lvert\Omega'\rvert$	$\delta = \left(1 - \sum_{i=0}^{\lfloor \tau' \lvert\Omega'\rvert \rfloor} \binom{\lvert\Omega'\rvert}{i} \tau^i (1-\tau)^{\lvert\Omega'\rvert - i}\right)^{\mathcal{I}}$		
	$\tau = 0.01$	$\tau = 0.05$	$\tau = 0.10$
10	3.98864E-48	5.05745E-22	2.68425E-12
25	6.36366E-55	1.21159E-18	2.04106E-07
50	5.5719E-100	3.50235E-29	4.82883E-09
75	2.3265E-121	3.41996E-30	4.07549E-07
100	8.4192E-165	1.55989E-39	2.50889E-08
125	4.4727E-186	1.01234E-40	4.97296E-07
150	5.7608E-229	1.36029E-49	5.08121E-08

Feasibility Assessment

In addition to assessing whether the optimal MIP-JCC objective function value of a subset of $|\Omega'|$ demand scenarios is a lower bound on the true JCC objective function value, we also must consider whether a solution to the approximated problem is a feasible solution to JCC.

However, since theoretical estimates on the number of scenarios necessary to have high confidence that the sample approximation is a feasible solution to the true optimization problem are very large, the approximation problems quickly become impractical to solve. Thus, we use the optimal location and allocation decisions associated with the minimum MIP-JCC objective function value obtained from $\mathcal{J} = 20$ repetitions of $|\Omega'|$ scenarios to conduct an *a posteriori* check that counts the number of scenarios in which the chance constraints are satisfied for a problem with a much larger number of i.i.d. scenarios (Luedtke and Ahmed, 2007). If the optimal location and allocation solutions from the smaller scenario size problem satisfy the joint chance constraints in the larger sample size problem, we can have high confidence that the location and allocation decisions are feasible to JCC as well.

Recall that all of the Poisson distributed demand instances considered in Figure 22 have the same optimal location and allocation decisions, regardless of the scenario size. As a result, we find that when the optimal location and allocation decisions obtained from the data instances with Poisson distributed demand presented in Figure 22 are imposed on a data instance of 1,000 scenarios, the joint chance constraints are violated in only 18 scenarios. Since $0.018 < \tau \in \{0.05, 0.1\}$ (the values of τ used in Table 27) we can have high confidence that the location and allocation decisions are feasible to JCC with $\tau \in \{0.05, 0.1\}$. However, since the optimal location and allocation decisions vary for the Truncated Normal distributed demand instances represented in Figure 24, we assess the optimal location and allocation decisions associated with the minimum objective function value obtained over $\mathcal{J} = 20$ repetitions for each value of $|\Omega'| \in \{10, 25, 50, 75, 100\}$. We find that none of these solutions satisfy the chance constraints when imposed on a 1,000 scenario instance. This indicates a MIP-JCC sample size significantly larger than 100 scenarios is needed to simultaneously obtain a lower bound on the optimal JCC

objective function value and have high confidence in obtaining a feasible solution to JCC when demand follows a Truncated Normal distribution with mean $\frac{1}{10{,}000}\rho_i$ and standard deviation $2\sqrt{\frac{1}{10{,}000}\rho_i}$.

The remainder of the computational results presented in this chapter focuses on the Poisson distributed demand instances. Furthermore, since we found that the optimal location and allocation decisions are stable for all scenario sizes presented in Figure 22, we incorporate $|\Omega'| = 10$ scenarios to save on computation time.

5.5.2 Comparison of Solution Times

In this section, we discuss the computational efficacy of the two-stage and three-stage MIP-JCC solution algorithms and compare their performance to directly solving the problems via CPLEX. For the remainder of the results presented in this chapter, the per item per day transportation and backlog cost parameters a and b are set to 15 and 30, respectively, $|\Omega'| = 10$ scenarios are used, demand is generated from a Poisson distribution with mean $\frac{1}{10{,}000}\rho_i$ at demand site $i \in I$, and the chance constraint parameters are $\theta = 91$ and $\tau' = 0.1$.

As shown in Figure 26, the two-stage algorithm based on Benders decomposition scheme outperforms all other solution methods as it attains the optimum with 450 demand and candidate nodes within one hour of run time. On the other hand, CPLEX performs quite poorly as it is only able to solve instances with at most 150 demand and candidate nodes before running out of memory.

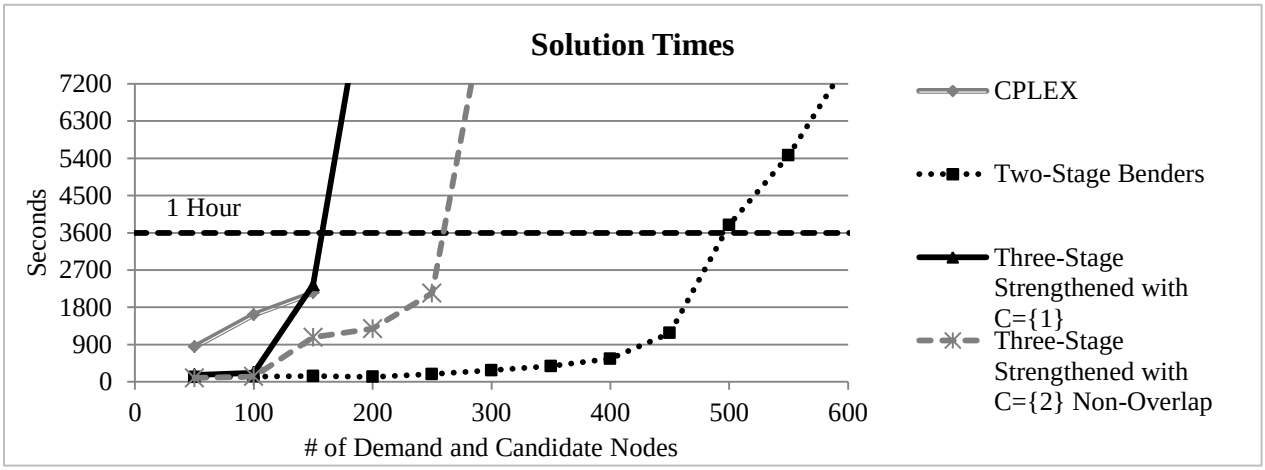

Figure 26: Comparison of solution methods; $a = 15, b = 30; \theta = 91, \tau' = 0.1$

Additionally, the unstrengthened three-stage decomposition algorithm takes hours to solve even a 50 demand and candidate node instance. However, strengthening the three-stage model with the relaxed joint chance constraints proposed in Section 5.3.4 greatly improves the solution progress. Simply adding the MIP formulation of joint chance constraint (5.73) (i.e., Three-Stage Strengthened with $C = \{1\}$) allows us to solve instances with 150 demand and candidate nodes within one hour. Furthermore, we find the three-stage decomposition algorithm with $C = \{2\}$ and non-overlapping couplings tends to be the most computationally efficient method of implementing the three-stage algorithm, although it still performs much more poorly than the two-stage decomposition approach. In fact, for 150 demand and candidate nodes, the Three-Stage Strengthened with $C = \{2\}$ Non-Overlap takes approximately eight times as long as the Two-Stage Benders approach to solve a 150 demand and candidate note instance, whereas the Three-Stage Strengthened with $C = \{1\}$ and CPLEX approaches take 16 times as long. Additionally, the strengthening constraints do not significantly affect the solution time of the two-stage decomposition approach in the instances we tested.

5.5.3 Optimal Facility Locations

Regardless of the number of candidate nodes, the optimal facility locations tend to be located near counties with the largest population since they generate the most demand. To get a sense for how important it is to locate near the most populous counties, consider that the county of Los Angeles, CA singlehandedly generates over 10% of the total average demand in the 50 demand node instance. Increasing the number of demand sites to 500 still leaves Los Angeles with 4.3% of the total average demand. In fact, in the 500 demand node case, the five most populous counties only account for 1% of the demand sites but generate 11.38% of the total average demand. The five most populous counties are displayed in Figure 27 and the optimal facility locations when there are 50 or 500 demand and candidate nodes are mapped in Figures 28 and 29, respectively. The county rank is noted in parenthesis next to the node name. For example, (1) Los Angeles, CA signifies that the demand node located in Los Angeles, CA generates more demand than any other demand node.

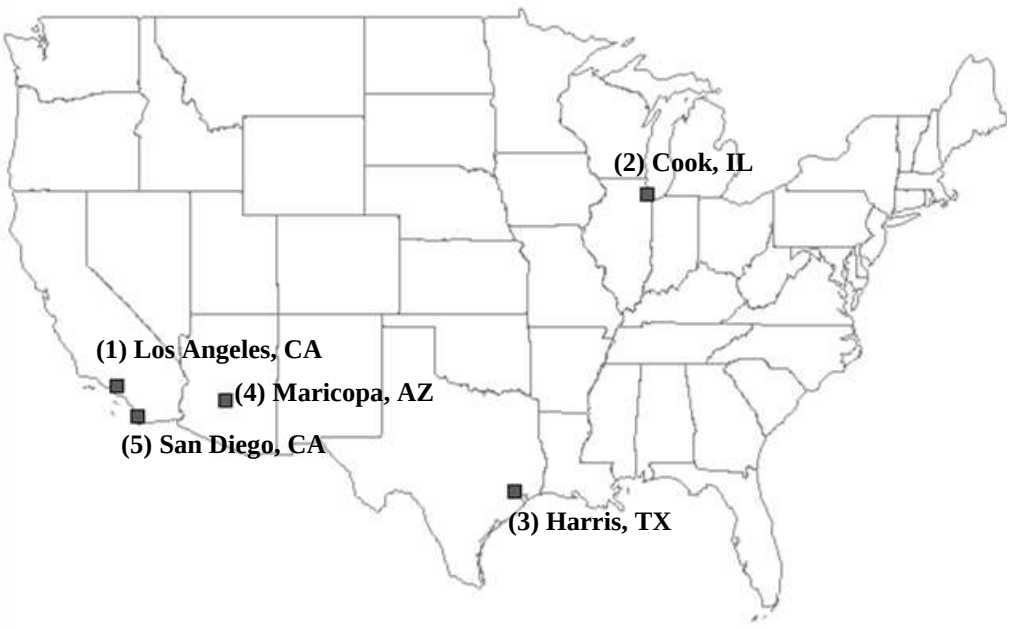

Figure 27: Five most populous counties

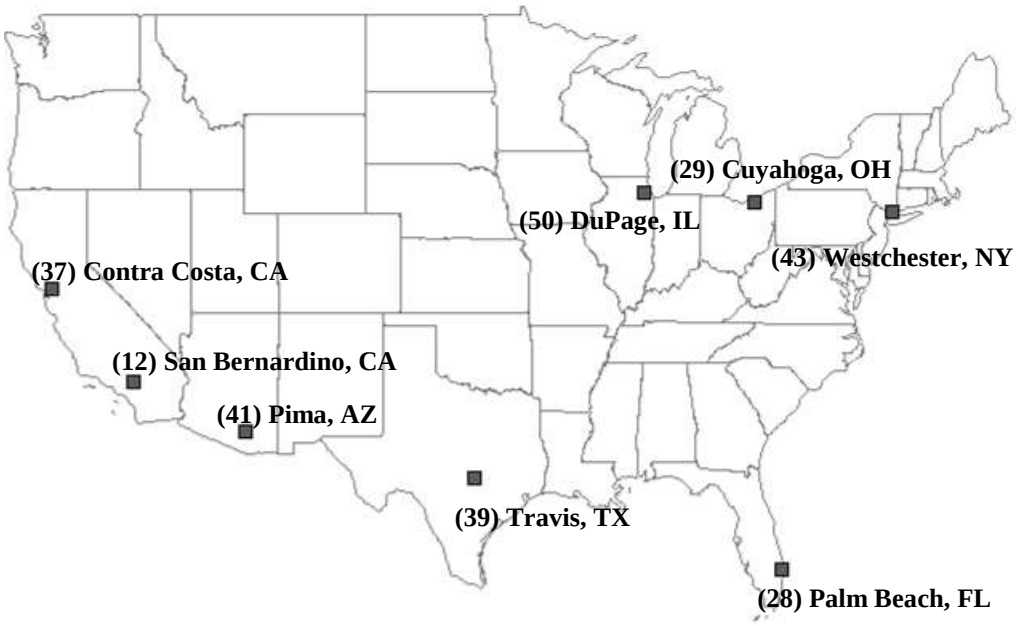

Figure 28: Optimal facility locations for the instance with 50 demand & candidate nodes; $a = 15, b = 30; \theta = 91, \tau' = 0.1$

Figure 29: Optimal facility locations for the instance with 500 demand & candidate nodes; $a = 15, b = 30; \theta = 91, \tau' = 0.1$

Recall that the data sets are generated in such a way that as the number of demand and candidate nodes increases, the newly added candidate nodes have lower fixed facility location costs. Thus, as the size of the candidate node set increases, less costly facility location options

become available near the most populous counties. Establishing a processing facility in a less populous county near a county with a large population therefore results in options with lower fixed facility location costs and minimal additional transportation costs. For data instances with 300 or more demand nodes, allowing all of the demand sites to serve as candidate facilities results in about a 13% cost reduction in total costs when compared to limiting the candidate facility set to the 50 most populous counties. Table 28 displays the percent increase in cost caused by reducing the number of candidate facilities while Table 29 illustrates that the solution time typically decreases dramatically as fewer candidate facilities are considered. (The two-stage decomposition algorithm was used to solve the instances.) Furthermore, as the number of demand nodes increases, reducing the number of candidate nodes to half of the demand nodes typically provides increasing benefit with regard to solution time as well as decreasing harm to the total cost. For example, using 150 candidate nodes in the 300 demand node instance provides a 51.71% reduction in solution time and 4.90% increase in total cost when compared to allowing all 300 nodes be candidate nodes. However, when 250 candidate nodes are used to determine the solution to the 500 demand node instance, we see an 86.74% reduction in computational time and only a 1.8% increase in total cost. Furthermore, the shaded values in Tables 28 and 29 correspond to instances in which only half of the demand sites are candidate facilities. Comparing these tradeoffs in a graphical manner, as we do in Figure 30 for the 500 demand node case, can help the user determine whether to reduce the number of candidate facilities.

Table 28: Percent total cost increases from the case in which all demand sites are candidate processing facilities; $a = 15, b = 30; \theta = 91, \tau' = 0.1$

Number Candidate Facilities	Number Demand Sites									
	50	100	150	200	250	300	350	400	450	500
50	0%	3.25%	8.42%	9.81%	11.86%	12.86%	13.02%	13.47%	13.87%	14.36%
100	-	0%	3.54%	5.50%	7.32%	8.58%	9.00%	9.28%	9.79%	10.18%
150	-	-	0%	1.69%	3.67%	4.90%	5.47%	5.62%	6.20%	6.46%
200	-	-	-	0%	1.63%	2.77%	3.13%	3.41%	4.03%	4.36%
250	-	-	-	-	0%	1.22%	1.86%	2.23%	2.86%	1.80%
300	-	-	-	-	-	0%	0.49%	0.95%	1.50%	1.80%
350	-	-	-	-	-	-	0%	0.41%	1.04%	1.30%
400	-	-	-	-	-	-	-	0%	0.68%	1.08%
450	-	-	-	-	-	-	-	-	0%	0.34%
500	-	-	-	-	-	-	-	-	-	0%

Table 29: Percent decrease in solution time compared to the case in which all demand sites are candidate processing facilities; $a = 15, b = 30; \theta = 91, \tau' = 0.1$

Number Candidate Facilities	Number Demand Sites									
	50	100	150	200	250	300	350	400	450	500
50	0%	29.92%	14.29%	54.20%	80.53%	83.27%	85.96%	89.20%	94.26%	97.91%
100	-	0%	42.14%	14.50%	62.11%	67.68%	71.63%	78.79%	89.13%	95.94%
150	-	-	0%	43.51%	37.89%	51.71%	57.02%	67.80%	83.73%	94.07%
200	-	-	-	0%	22.11%	30.42%	40.45%	56.06%	75.94%	91.59%
250	-	-	-	-	0%	12.55%	19.94%	38.64%	68.66%	86.74%
300	-	-	-	-	-	0%	11.52%	30.30%	64.04%	86.65%
350	-	-	-	-	-	-	0%	18.37%	55.99%	71.56%
400	-	-	-	-	-	-	-	0%	23.97%	68.79%
450	-	-	-	-	-	-	-	-	0%	58.24%
500	-	-	-	-	-	-	-	-	-	0%

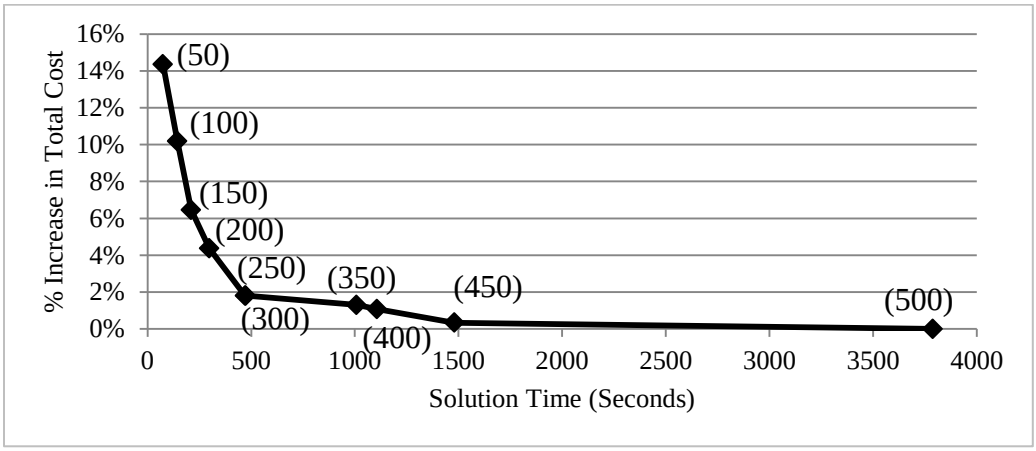

Figure 30: Tradeoffs of reducing the candidate node set for the 500 demand node instance; the number of candidate nodes used is in parenthesis; $a = 15, b = 30; \theta = 91, \tau' = 0.1$

Additionally, since we incorporate less costly facility location options as the size of the candidate node set increases, the model also decides to locate more facilities as the number of candidate nodes increases. This further helps reduce transportation and backlogging costs by establishing facilities that are near more remote counties that would otherwise incur large transportation costs. However, since the processing capacity depends on the demand sites under consideration (i.e., for a given set of demand sites, each candidate facility has a capacity of 1/5 of the total average daily demand) and the processing capacities are the same for all candidate facilities, locating more facilities decreases the total processing utilization of the system, as shown in Table 30. For example, locating five facilities in the 50 demand and candidate node instance would result in a system-wide production capacity of $1{,}834 * 5 = 9{,}170$. Since an average of 9,166 demands are generated each day, the system would have a utilization of $9{,}166 / 9{,}170 = 99.96\%$. However, the optimal solution is to locate eight facilities, which results in a system-wide capacity of $1{,}834 * 8 = 14{,}672$ and a utilization of 62.47% since the total average daily demand has not changed. Similarly, locating five facilities in the 600 demand and candidate node instance would result in a system-wide capacity of $4{,}825 * 5 = 24{,}125$ corresponding to a utilization of 99.99%. Yet, the optimal solution is to locate 20 facilities, which allows for a 96,500 processing capacity and a utilization of only 25%.

Table 30: More facilities are located as the number of demand and candidate nodes increase, which results in a decrease in capacity utilization; $a = 15, b = 30$

# Demand and Candidate Nodes	# Facilities Located	ζ =Total Avg. Daily Demand	Daily Capacity of each Located Facility	κ =Total Daily Capacity of Located Facilities	Utilization: $\frac{\zeta}{\kappa}$
50	8	9,166	1,834	14,672	62.47%
100	11	12,936	2,588	28,468	45.44%
150	12	15,491	3,099	37,188	41.66%
200	13	17,360	3,472	45,136	38.46%
250	14	18,787	3,758	52,612	35.71%
300	14	19,929	3,986	55,804	35.71%
350	14	20,878	4,176	58,464	35.71%
400	15	21,696	4,340	65,100	33.33%
450	17	22,421	4,485	76,245	29.41%
500	18	23,054	4,611	82,998	27.78%
550	20	23,618	4,724	94,480	25.00%
600	20	24,123	4,825	96,500	25.00%

Figure 31 graphically displays the increase in the optimal number of facilities to establish as the size of the candidate node set increases as well as a power regression. The regression model suggests that the number of facilities located increases by approximately the 1/3 power (R^2 value of 0.94) as the number of demand and candidate nodes increase. Since each located facility can process 1/5 of the average amount of generated demand each day, in expectation, the system has sufficient capacity to process all of the demand when at least five facilities are located. Thus, we also plot the number of *extra* facilities that are located in the optimal solution (e.g., if eight facilities are located, we consider three of them extra since only five are needed to satisfy the demand) in Figure 32. In this case, a power regression suggests that the number of extra facilities located increases to the 0.58 power (R^2 value of 0.95) as the number of demand and candidate nodes increase.

The results of the power regression models suggest that it may be possible to develop an analytic model to determine the optimal number of facilities to locate in the stochastic IMCLP. We direct the interested reader to Daskin (2010) for an example of an analytic model that determines the optimal number of facilities to locate under many simplifying assumptions (e.g.,

demand occurs uniformly over a square region and all candidate facilities have the same location cost). While analytic models typically make many assumptions about the nature of the demand and candidate facility locations, they have the benefit of being very easy to solve. In fact, many can be solved using simple calculus techniques such as the ones that Daskin (2010) uses to show that the optimal number of facilities grows as the $2/3$ power of the demand density and the unit transportation cost increases, but decreases with the $2/3$ power of the location cost. We leave the development of analytic models for the stochastic IMCLP as an area of future research.

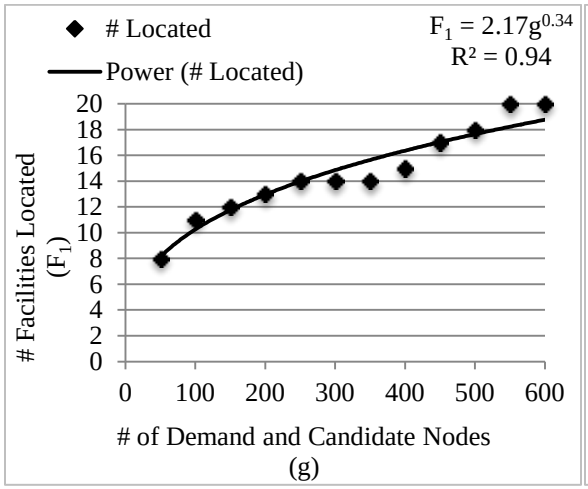

Figure 31: More facilities are located as the number of demand and candidate nodes increase.
$a = 15, b = 30, \theta = 91, \tau' = 0.1$

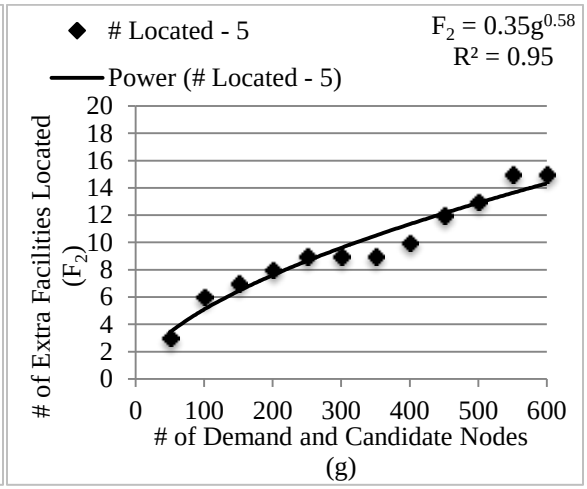

Figure 32: The optimal solution has excess system-wide capacity due to locating extra facilities.
$a = 15, b = 30, \theta = 91, \tau' = 0.1$

5.5.4 Effect of Using Individual or Hybrid Chance Constraints

The chance constrained results discussed thus far have focused on limiting the probability that any processing facility has a backlog level above the threshold on any day of the planning horizon; that is, they have focused on the joint chance constrained formulation. However, as we will see in this section, limiting the amount of items backlogged by utilizing individual or hybrid chance constraints affects the total cost as well as the frequency in which backlogged demand

exceeds the desired threshold level. Rather than suggesting which type of chance constraint should be employed, we present illustrations of the effect of using each type of chance constraint and encourage modelers to incorporate the particular type of constraints that best represents their willingness to trade off decreased costs for increased backlog levels.

Therefore, when comparing the different solutions the three types of chance constraints can provide, it is helpful to recall the differences in the way the chance constraints limit the amount of backlogged demand. The joint chance constraints state that there can be a maximum of $\tau'|\Omega'|$ scenarios with a backlog level that exceeds θ on any day at any facility. (Note that this approach does not distinguish between a scenario in which a single facility exceeds the desired backlog level θ by one unit on one day and a scenario in which multiple facilities have multiple days in which the backlog level greatly exceeds θ.) The hybrid chance constraints state that each facility can have a maximum of $\tau'|\Omega'|$ scenarios in which the backlog level at that facility exceeds θ on any day of the planning horizon, and the individual chance constraints state that for each facility-day combination, there can be a maximum of $\tau'|\Omega'|$ scenarios in which the backlog level exceeds θ at that particular facility on that particular day.

We assess the effect of using the joint, hybrid, and individual constraints on the total cost for various exceedance tolerances (τ'). Results from an instance of Poisson distributed demand with 50 scenarios and an instance of Truncated Normal distributed demand with ten scenarios are shown in Figures 33 and 34, respectively. Both instances incorporate 50 demand and candidate node. (Note that these instances correspond to those used in Section 5.5.1). The maximum desired amount of backlog is set to $\theta = 91$, which represents 5% of the daily processing capacity at each facility. For the Poisson distributed demand instances presented here, MIP-JCC and MIP-HCC produce the same optimal solution regardless of the value of τ'. However, the Truncated

Normal distributed demand instance does produce different optimal solutions to these problems for some values of τ'.

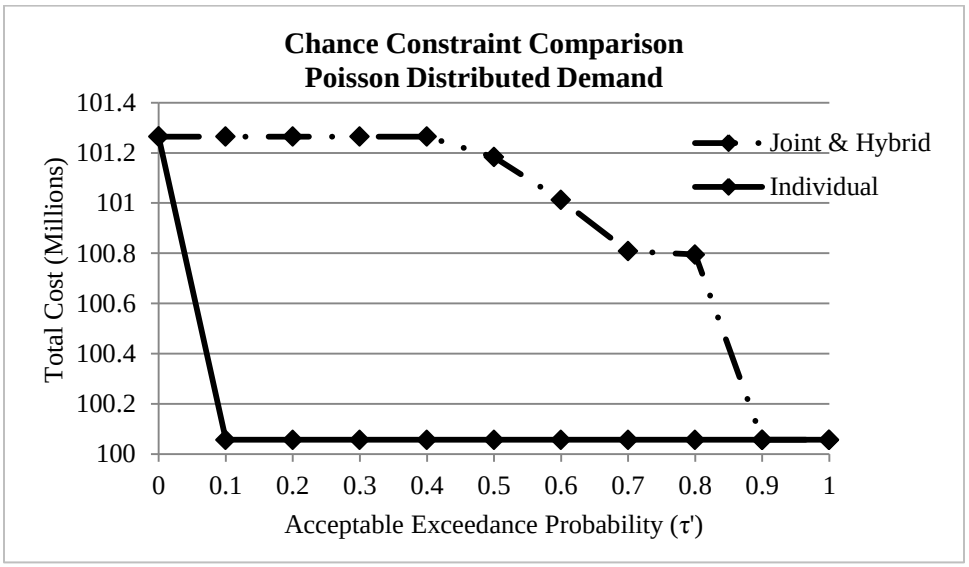

Figure 33: Comparison of the optimal cost of the MIP-JCC, MIP-HCC, and MIP-ICC chance constraints; Poisson distributed demand, 50 scenarios, $a = 5$, $b = 10$, $\theta = 91$

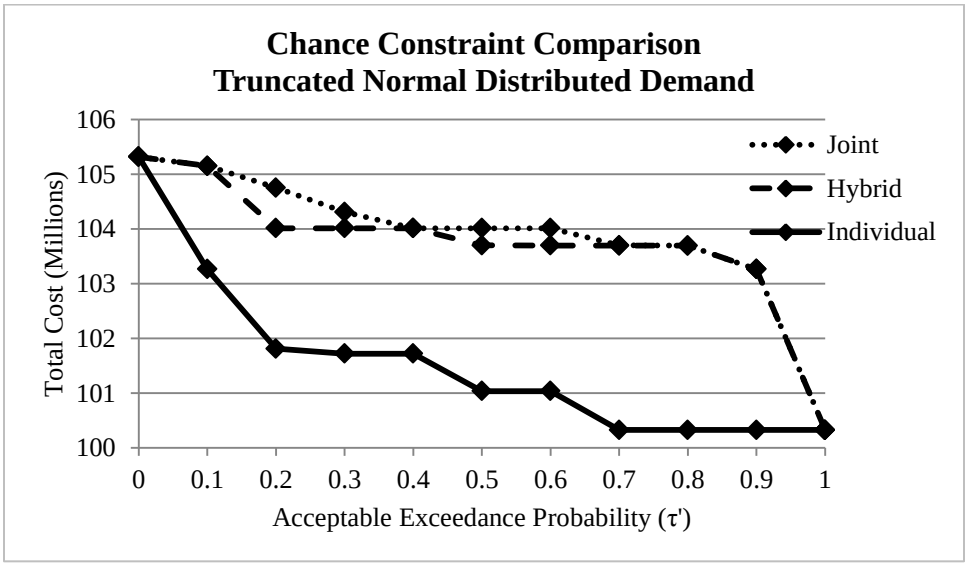

Figure 34: Comparison of the optimal cost of the MIP-JCC, MIP-HCC, and MIP-ICC chance constraints; Normally distributed demand, 10 scenarios, $a = 5$, $b = 10$, $\theta = 91$

As expected, as the value of τ' increases, the chance constraints become more relaxed thereby reducing the optimal cost. When $\tau = 0$, the chance constraints require that the amount of backlog never exceed θ. On the other hand, $\tau' = 1$ is equivalent to the stochastic formulation without the joint chance constraints (i.e., (2.2) - (2.4), (3.6), and (5.14) - (5.19)). We note that although the amount of backlog is not restricted by the chance constraints when $\tau' = 1$, it is still limited by the penalty cost in the objective function.

Additionally, for a given value of τ', the optimal MIP-JCC solution is at least as costly as the MIP-HCC solution, which in turn is at least as costly as the MIP-ICC. However, it is important to remember that while the MIP-ICC may produce a lower cost solution than the MIP-HCC or MIP-JCC, it also allows for backlog to exceed the desired maximum level θ more frequently. For example, Figure 35 shows the amount of backlogged demand at Fresno, CA in one of the Truncated Normal distributed demand scenarios when $\tau' = 0.2$. In this single scenario, the MIP-ICC solution resulted in nine days in which the backlog level exceeded $\theta = 91$ at the Fresno, CA processing facility, while the hybrid and joint chance constrained solutions only had one day of exceedance. We note that while the MIP-HCC and MIP-JCC solutions do not produce the same optimal solution when $\tau' = 0.2$ in the Truncated Normal distributed demand instance considered (as is evident in the different costs presented in Figure 34), the MIP-HCC and MIP-JCC allocate the same demand sites to Fresno, CA in both solutions. This results in the solution to both of these formulations having same amount of backlogged demand at Fresno, CA for this demand instance.

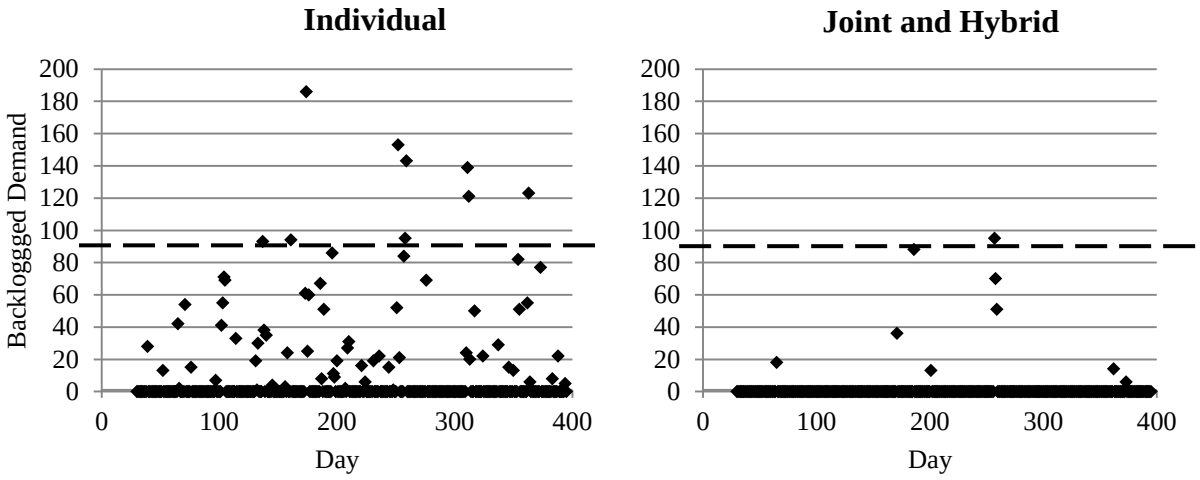

Figure 35: Example scenario of backlogged demand at Fresno, CA; Normally distributed demand, 10 scenarios, $a = 5$, $b = 10$; $\theta = 91$, $\tau' = 0.2$

5.5.5 Comparison to IMCLP Solution

Although the deterministic IMCLP allows the model to capture fluctuations in daily demand, it only considers a single realization of demand at each demand-generating site on each day. The purpose of this section is to illustrate the benefit of incorporating multiple demand scenarios and chance constraints that limit the amount of backlogged demand (i.e., the stochastic IMCLP model) compared to a single demand scenario and no restrictions (other than an associated penalty cost) on the amount of backlogged demand (i.e., the deterministic IMCLP model). As the majority of the results presented in this chapter focus on the joint chance constrained model, we specifically investigate the effect of incorporating stochasticity into the IMCLP when joint chance constraints are imposed.

Recall that in Section 5.5.1 we found that the optimal solution to MIP-JCC located the same six facilities and kept the demand allocations the same regardless of the value of $|\Omega'| \in \{10, 25, 50, 75, 100, 125, 150\}$ in the case of Poisson distributed demand with 50 demand and candidate nodes. To assess the value of considering multiple scenarios and imposing chance

constraints using the MIP-JCC verses a single scenario and no probabilistic limits on the amount of backlogged demand via the deterministic IMCLP, we use each of the 10 demand scenarios from one of the $|\Omega'| = 10$ instances corresponding to replication $i = 1$ in Figure 22 to solve a corresponding IMCLP problem. For example, we will solve one IMCLP instance that has the same demand as the *first* scenario of the $i = 1$ MIP-JCC instance with $|\Omega'| = 10$, one IMCLP instance that has the same demand as the *second* scenario of the $i = 1$ MIP-JCC instance with $|\Omega'| = 10$, etc. Then, one at a time, we force the optimal IMCLP location and allocation decisions into the $i = 1$ MIP-JCC instance with $|\Omega'| = 10$.

Through this process, we find that all ten IMCLP instances located the same six facilities as the optimal locations to the MIP-JCC instance. However, three of the 50 demand sites allocate their demand to different processing facilities in the IMCLP solutions than they do in the MIP-JCC solution. Specifically, Orange CA sends its demands to Fresno, CA in the MIP-JCC but to San Bernardino, CA in the IMCLP solutions, and Riverside, CA sends its demands to San Bernardino, CA in the MIP-JCC but to Fresno, CA in the IMCLP solutions. More notably, Salt Lake, UT allocates its demands to Travis, TX in the MIP-JCC but to Fresno, CA in the IMCLP. A spatial representation of the location and allocations of the two solutions are presented in Figure 36.

IMCLP MIP-JCC

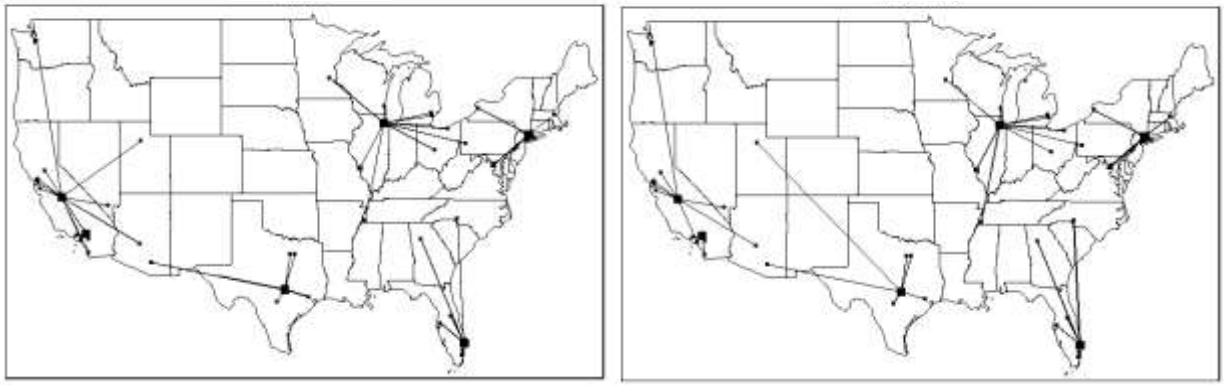

Figure 36: Comparison of the optimal IMCLP and MIP-JCC location and allocation solutions; Poisson distributed demand; $a = 5$, $b = 10$; For MIP-JCC: $\theta = 91$, $\tau' = 0.1$

Although the allocation decisions differ only slightly between the optimal solutions to the IMCLP and MIP-JCC instances, they result in significantly different amounts of backlog. As the results of Table 31 show, nine of the ten IMCLP instances have optimal solutions in which the maximum amount of backlog held at any facility on any day of the planning horizon exceeds the desired maximum backlog amount $\theta = 91$; since the joint chance constraints are not included in the IMCLP, the IMCLP therefore does not account for decision maker's preference regarding the maximum backlog amount and allows allocations that result in higher backlog levels. This means that the optimal IMCLP solution is not a feasible solution to the MIP-JCC with $\theta = 91$ and $\tau' = 0.1$ since the IMCLP solution would exceed $\theta = 91$ in 90% of scenarios.

Furthermore, Table 31 shows that the total amount of backlog (summed over all days, facilities, *and scenarios*) in the optimal MIP-JCC solution is 2,083, which is *less* than the total amount of backlog for each of the IMCLP instances that only incorporate a *single scenario* of demand. In fact, if we sum the total amount of backlogged demand over all ten of the single scenario IMCLP instances, we obtain 28,297 backlogged demands. This is *over 13 times* as many backlogged demands as compared to the optimal MIP-JCC solution.

Thus, it is not surprising that the average cost of the optimal IMCLP solutions is less than the optimal cost of the MIP-JCC solution since the IMCLP problems were not constrained by the chance constraints as was the MIP-JCC.

Table 31: Optimal IMCLP and MIP-JCC cost and backlog comparison, $a = 5$, $b = 10$; For MIP-JCC: $\theta = 91$, $\tau' = 0.1$

	Optimal OF Value	Max Backlog	Total Backlog
IMCLP Instance 1	$100,014,845	98	2,100*
IMCLP Instance 2	$100,029,145	111	2,998*
IMCLP Instance 3	$100,057,020	114	3,405*
IMCLP Instance 4	$100,086,375	123	2,725*
IMCLP Instance 5	$100,062,480	141	3,503*
IMCLP Instance 6	$100,056,630	86	2,697*
IMCLP Instance 7	$100,075,920	136	4,141*
IMCLP Instance 8	$ 99,984,675	128	2,718*
IMCLP Instance 9	$ 99,997,490	238	4,010*
IMCLP Instance 10	$100,047,325	132	3,421*
Average	$100,041,191	131	3,172*
MIP-JCC; 10 Scenarios $\theta = 91$, $\tau' = 0.1$	$101,250,878	61	2,083**
		*Over 1 scenario	**Over 10 scenarios

Furthermore, if we add a constraint to the IMCLP that limits the amount of items that can be backlogged at a facility each day to $\theta = 91$, we find that the optimal IMCLP location and allocation decisions differ from one another. Specifically, individually solving the ten IMCLP instances resulted in five different location and allocation decisions. Instances two and four returned the same location and allocation decisions, as did instances five, seven, and nine. Instances three, eight, and ten also had the same optimal location and allocation decisions, and these decisions were identical to the optimal MIP-JCC decisions.

To assess how well the optimal solutions to the IMCLP instances compare to the MIP-JCC solution in terms of the likelihood of the desired backlog level exceeding $\theta = 91$, we force each of the five different optimal solutions obtained from the ten IMCLP instances into the MIP-JCC instance with all ten scenarios and count the number of scenarios in which the IMLCP solution results in more than $\theta = 91$ backlogged demands in the MIP-JCC. Table 32 specifies

the percent of scenarios that exceed the desired maximum backlog level $\theta = 91$, as well as the maximum backlog level and the total amount of backlog across all ten scenarios. We note that although each of the IMCLP problem instances imposed a constraint to limit the amount of backlogged demand to $\theta = 91$, the maximum backlog levels reported in Table 32 often exceed $\theta = 91$. This is because the IMCLP only considers a single demand scenario while the results presented in Table 32 consider how the solution to that particular IMCLP instance with one demand scenario performs across all ten scenarios. Lastly, we remind the reader that any solutions for which θ is exceeded in over 10% of the cases are infeasible solutions to the MIP-JCC instance with $|\Omega'| = 10$, $\theta = 91$, $\tau' = 0.1$.

Table 32: Performance of optimal IMCLP instance location and allocation decisions in MIP-JCC when IMCLP solution enforces daily facility backlog limit of $\theta = 91$; $a = 5$, $b = 10$

	% of scenarios w/ backlog exceeding $\theta = 91$	Max Backlog**	Total Backlog**
Instance 1	80%	145	22,994
Instances 2 and 4	80%	149	28,781
Instances 3, 8 and 10	50%	122	21,233
Instances 5, 7, and 9 (and MIP-JCC)	0%	61	2,083
Instance 6	90%	238	31,718

***Over 10 Scenarios*

Similar to our results for when the IMCLP did not have restrictions on the amount of backlogged demand, Table 32 shows that the optimal MIP-JCC solution results in at least an *order of magnitude fewer* backlogged items than the solutions to seven of the IMCLP instances when they are forced into the MIP-JCC, *even when a limit on the amount of backlogged demand is imposed on the IMCLP*.

To aid in the understanding of Table 32, we give a detailed example for the optimal solution to instances two and four in Table 33. It shows that the maximum backlog levels are restricted to be less than $\theta = 91$ for scenarios two and four, as desired, but imposing the optimal

solution to IMCLP instances two and four on the eight remaining scenarios results in those scenarios having higher backlog levels.

Table 33: Detailed performance of optimal location and allocation decisions for IMCLP instances two and four in MIP-JCC when IMCLP solution enforces daily facility backlog limit of $\theta = 91; a = 5, b = 10$

	Max Backlog	Total Backlog
Scenario 1	95	2,189
Scenario 2	83	2,251
Scenario 3	101	2,639
Scenario 4	82	2,779
Scenario 5	120	3,361
Scenario 6	114	2,971
Scenario 7	122	3,609
Scenario 8	143	2,578
Scenario 9	149	3,935
Scearnio10	103	2,469
Over All 10 Scenarios	149	28,781

The problem instances studied in this section focused on Poisson distributed demand, which we saw provided the same optimal location and allocation decisions regardless of whether 10, 25, 50, 75, 100, 125, or 150 demand scenarios were considered in Section 5.5.1. We can therefore expect to see an even greater benefit of using the stochastic IMCLP (as compared to the deterministic IMCLP) when demand distributions with greater variability are considered.

5.6 Chapter Summary

In this chapter, we illustrated how incorporating chance constraints into the IMCLP formulation can affect the optimal solution in terms of the optimal facility locations, demand allocations, amount of items held in backlog, and overall cost. Three different types of chance constraints were presented: joint, individual, and hybrid chance constraints. In addition to assessing a penalty cost for holding items in backlog, as the deterministic formulations do, the chance constrained versions also impose probabilistic limits on the amount of demand that can be backlogged.

Through computational experiments, we showed that the stochastic, chance constrained version of the IMCLP can produce different location and allocation solutions than the IMCLP, and that even small differences in their allocations can drastically impact the amount of items held in backlog. Furthermore, if the decision maker has a limit on the desired maximum amount if items backlogged at any given time, it is important that the decision maker utilize the stochastic model; imposing the limit on a single demand scenario (i.e., using the IMCLP) often resulted in infeasible solutions to the joint chance constrained stochastic IMCLP due to the demand uncertainty.

For each of the chance constrained formulations, we proposed two solution algorithms. The first was a two-stage formulation based on the Benders decomposition scheme. The second was a three-stage decomposition that capitalized on the fact that once the location and allocation decisions are determined, the corresponding processing and backlog variables can automatically be computed. We then checked whether the joint chance constraint was satisfied. Our computational results showed that the three-stage formulation performs poorly. However, strengthening the first stage by adding a relaxed joint chance constraint markedly improved its solution progress. Both solution algorithms outperformed CPLEX in all of our instances, although we found that the two-stage decomposition scheme performed the best.

CHAPTER 6: Extensions and Conclusions

6.1 Introduction

In this final chapter, we highlight extensions of the models discussed in this dissertation and present the dissertation conclusions and contributions. Two extensions explore methods of incorporating demand backlog into capacitated facility location models without explicitly accounting for daily demands. The first such extension uses a queueing approach to model demand arriving at a processing facility while the second approach assigns an incremental cost, which represents the expected cost of backlog, to each additional unit of average demand assigned to a processing facility.

Then, in Section 6.4 we show that the framework of this dissertation can also be applied to a context in which intangible demands arrive at the facilities and result in physical demands that must be shipped from the facilities to customers. Section 6.5 discusses additional methods for capturing the effects related to incoming demand that exceed a facility's daily processing capacity. Finally, Section 6.6 summarizes the contributions of this dissertation.

6.2 Queueing Model to Approximate Average Backlog

Rather than explicitly incorporating the daily fluctuations in backlog into the model, we can approximate the average backlog at a processing facility under specific location and allocation decisions using a queueing approach. This is based on the observation that each processing facility can be viewed as a queuing system in which demands arrive (in a batch of

uncertain size due to the stochastic demand process) at the processing facility at the beginning of the day and are processed throughout the day in a batch (of up to "capacity" items) with a deterministic service time of one day. If at the end of the day there are demand that were not processed, those demand will incur one unit of "wait time" as backlogged demand. Demand that are processed on the day they arrive do not incur any waiting time.

6.2.1 Queueing Formulation

Define $H_{j,d} = \sum_{i \in I} h_{i,t-t_{ij}} Y_{ij}$ as the total demand that arrives at processing facility $j \in J$ on day $d \in \{t^* + 1, \dots, |D|\}$. Since facility $j \in J$ can process k_j units of demand each day, the number of items in backlog at the beginning of day $d + 1$, for $d \in \{t^* + 1, \dots |D|\}$, can be written as:

$$V_{j,d+1} = \max\{V_{j,d} + H_{j,d} - k_j, 0\}$$

where we have assumed that the incoming demand at facility j arrives at the beginning of the day. The preceding equation is called Lindley's equation for waiting time in a *D/G/1* queue [Lindley, 1952] with interarrival times identically equal to k_j and service time $H_{j,d}$ for customer d. Thus, if the time horizon is sufficiently long, we can study the steady state backlog cost at processing facility $j \in J$. Letting $\mathbb{E}[V_j]$ denote the long-term expected number units in backlog per day at facility $j \in J$, the expected value of the backlog term in objective function (5.2) becomes

$$b * \mathbb{E}_{\xi}\left[\sum_{d=t^*+2}^{|D|+1} \sum_{j \in J} \tilde{V}_{jd}\right] = b(|D| - t^*) \sum_{j \in J} \mathbb{E}[V_j]$$

If the total demand arriving at facility $j \in J$ on day $d \in \{t^* + 1, \dots, |D|\}$ (i.e., $\sum_{i \in I} h_{i,d-t_{ij}} Y_{ij}$) is Poisson distributed, then the total backlog at j on day d will also be Poisson distributed, but with a probability mass at 0. In general, $\mathbb{E}[V_j]$ does not have a closed form

solution, which necessitates the use of tight approximation formulas for $\mathbb{E}[V_j]$. Doing so will enable us to eliminate the processing and backlog variables from the model and will considerably reduce the problem size.

6.2.2 Approximation Formulas for $\mathbb{E}[\mathbf{V}_j]$

With a slight abuse of notation, let H_j represent a generic total amount of demand arriving at facility $j \in J$ on any day of the planning horizon. Then, if the demands at demand site $i \in I$ are i.i.d. with mean h_i and variance $\mathfrak{v}_i$ we have

$$\mathbb{E}[H_j] = \sum_{i \in I} h_i Y_{ij}$$

$$var[H_j] = \sum_{i \in I} \mathfrak{v}_i Y_{ij}$$

since $Y_{ij} = Y_{ij}^2$. With this notation, we present three approximations for the average waiting time in a *D/G/1* queue.

Upper Bound

Although Lindley's equation represents the waiting time in a *D/G/1* queue, we can approximate the waiting time with a (more general) *G/G/1* queue. Kingman's formula is one such approximation that is known to be an upper bound on the average waiting time (i.e., backlog) at facility $j \in J$ that is very close to the exact solution when the utilization factor[1] is close to 1, but does not exceed one in order to preserve stability (i.e., $\mathbb{E}[H_j]/\mathbb{E}[k_j] \approx 1$, $\mathbb{E}[H_j]/\mathbb{E}[k_j] < 1$) [Kingman, 1962a,b; Bhat, 2008]:

[1] The utilization factor, also known as the traffic intensity, represents the probability the server will be busy at any point in time. It is calculated by dividing the mean arrival rate by the mean service rate for single server queues.

$$\mathbb{E}[V_j] \leq \frac{\frac{1}{\mathbb{E}[k_j]}\left(var(k_j) + var(H_j)\right)}{2\left(1 - \frac{\mathbb{E}[H_j]}{\mathbb{E}[k_j]}\right)} \tag{6.1}$$

Since $var(k_j) = 0$, $\mathbb{E}[k_j] = k_j$, and the demands at demand site $i \in I$ are i.i.d. with mean h_i and variance $\mathfrak{v}_i$,

$$\mathbb{E}[V_j] \leq \frac{\frac{1}{\mathbb{E}[k_j]}\left(var(k_j) + var(H_j)\right)}{2\left(1 - \frac{\mathbb{E}[H_j]}{\mathbb{E}[k_j]}\right)}$$

$$= \frac{var(H_j)}{2(k_j - \mathbb{E}[H_j])}$$

$$= \frac{\sum_{i \in I} \mathfrak{v}_i Y_{ij}}{2(k_j - \sum_{i \in I} h_i Y_{ij})}$$

Lower Bound

One of the most well-known bounds for the *G/G/1* queue is attributed to Marchal (1978):

$$\mathbb{E}[V_j] \geq \frac{\left(\frac{\mathbb{E}[H_j]}{\mathbb{E}[k_j]}\right)^2 + \left(\frac{1}{\mathbb{E}[k_j]}\right)^2 var(H_j) - 2\left(\frac{\mathbb{E}[H_j]}{\mathbb{E}[k_j]}\right)}{2\left(\frac{1}{\mathbb{E}[k_j]}\right)\left(1 - \left(\frac{\mathbb{E}[H_j]}{\mathbb{E}[k_j]}\right)\right)}$$

Upon simplification, the inequality becomes

$$\mathbb{E}[V_j] \geq \frac{\left(\mathbb{E}[H_j]\right)^2 + var(H_j) - 2k_j\mathbb{E}[H_j]}{2(k_j - \mathbb{E}[H_j])}$$

which, is unfortunately a weak lower bound since it takes negative values unless $1 < \frac{1}{k_j}\sqrt{Var(H_j)}$ [Larson and Odoni, 2007]. However,

$$\frac{\frac{1}{\mathbb{E}[k_j]}\left(var(k_j)+var(H_j)\right)}{2\left(1-\frac{\mathbb{E}[H_j]}{\mathbb{E}[k_j]}\right)}-\frac{1+\frac{\mathbb{E}[H_j]}{\mathbb{E}[k_j]}}{2\frac{1}{\mathbb{E}[k_j]}} \tag{6.2}$$

is known to be a tight lower bound for the subclass of *G/G/1* queues that have an interarrival time $\mathcal{T}$ and arrival rate λ that satisfy the property

$$\mathbb{E}[\mathcal{T}-t_0|\mathcal{T}>t_0]\leq\frac{1}{\lambda}\ \text{ for all } t_0\geq 0 \tag{6.3}$$

[Marshall, 1968]. (Notice that the first term of (6.2) is precisely the upper bound (6.1).) Since clearly requirement (6.3) is satisfied for the queue we are considering in which Lindley's equation identifies interarrival times equal to k_j,

$$\begin{aligned}\mathbb{E}[V_j] &\geq \frac{\frac{1}{\mathbb{E}[k_j]}\left(var(k_j)+var(H_j)\right)}{2\left(1-\frac{\mathbb{E}[H_j]}{\mathbb{E}[k_j]}\right)}-\frac{1+\frac{\mathbb{E}[H_j]}{\mathbb{E}[k_j]}}{2\frac{1}{\mathbb{E}[k_j]}}\\ &=\frac{var(H_j)}{2(k_j-\mathbb{E}[H_j])}-\frac{k_j+\mathbb{E}[H_j]}{2}\\ &=\frac{\sum_{i\in I}\mathfrak{v}_i Y_{ij}}{2(k_j-\sum_{i\in I}h_i Y_{ij})}-\frac{k_j+\sum_{i\in I}h_i Y_{ij}}{2}\end{aligned}$$

Averaged Approximation

By averaging the upper and lower bounds, we arrive at a third approximation:

$$\mathbb{E}[V_j] \sim\frac{\sum_{i\in I}\mathfrak{v}_i Y_{ij}}{2(k_j-\sum_{i\in I}h_i Y_{ij})}-\frac{k_j+\sum_{i\in I}h_i Y_{ij}}{4}$$

6.2.3 Model Summary and Challenges

Instead of incorporating the daily backlog fluctuations directly into the model, the average backlog at a processing facility under specific location and allocation decisions can be

approximated using a queueing approach. We presented four approximations, but note that they are all nonlinear due to the allocation variables, Y_{ij}. Thus, although this method would allow us to eliminate the large number of processing and backlog variables from the model, replacing the current linear backlog term in the objective function with a nonlinear term will introduce different computational challenges. Additionally, these approximations assume that the demands are independent and identically distributed, which precludes the case in which demands are correlated.

6.3 Incremental Backlog Cost Model

An additional approach to incorporating excess demand backlog into the facility location models is to consider the incremental cost associated with assigning an additional unit of average demand to processing facility $j \in J$.

6.3.1 Linear Programming Model Formulation

Let $\mathscr{b}_{jl}$ be a parameter representing the incremental backlog cost associated with assigning an average of l demands to facility $j \in J$ each day rather than $l-1$ demands, and assume that $\mathscr{b}_{jl}$ is an increasing function of l. In addition to the standard binary location (X_j) and allocation (Y_{ij}) decisions, we also include binary variables $\mathcal{Z}_{jl}$ which take the value one if the average amount of demand allocated to facility $j \in J$ each day is at least l and will take the value 0 otherwise. Thus, $\sum_{l=1}^{\infty} \mathcal{Z}_{jl}$is the average amount of demand allocated to facility $j \in J$. Using h_i to represent the average amount of daily demand generated at demand site $i \in I$, the model can be formulated as follows:

$$Min_{\boldsymbol{X},\boldsymbol{Y},\boldsymbol{\mathcal{Z}}} \quad (|D| - t^*)\left(\sum_{j\in J} f_j X_j + a\sum_{j\in J}\sum_{i\in I} h_i t_{ij} Y_{ij} + \sum_{j\in J}\sum_{l=1}^{k_j-1} \mathscr{b}_{jl}\mathcal{Z}_{jl}\right) \tag{6.4}$$

Subject to

$$\sum_{j\in J} Y_{ij} = 1 \qquad \forall i \in I \qquad (2.2)$$

$$Y_{ij} \leq X_j \qquad \forall i \in I; j \in J \qquad (2.3)$$

$$\mathcal{Z}_{jl} \leq X_j \qquad \forall j \in J; l \in \{1, \dots, k_j - 1\} \qquad (6.5)$$

$$\sum_{i\in I} \mathbb{E}[h_i] Y_{ij} \leq \sum_{l=1}^{k_j-1} \mathcal{Z}_{jl} \qquad \forall j \in J \qquad (6.6)$$

$$X_j \in \{0,1\} \qquad \forall j \in J \qquad (2.4)$$

$$Y_{ij} \in \{0,1\} \qquad \forall i \in I; j \in J \qquad (3.6)$$

$$\mathcal{Z}_{jl} \in \{0,1\} \qquad \forall j \in J; l \in \{1, \dots, k_j - 1\} \qquad (6.7)$$

As in previous models, the objective function (6.4) minimizes the total facility location, transportation, and expected backlogging costs over the planning horizon. Constraints (6.5) ensure that the average demand at facility $j \in J$ is 0 if facility j is not located, and constraints (6.6) ensure consistency between the $\mathcal{Z}_{jl}$ variables and the total average demand assigned to facility $j \in J$. Since $\mathscr{b}_{jl}$ is an increasing function in l, constraints requiring $\mathcal{Z}_{jl} \geq \mathcal{Z}_{j,l+1}$ are not required. Finally, constraints (2.4), (3.6), and (6.7) are standard binary constraints.

6.3.2 Determining the Incremental Cost Values

One method of determining the incremental cost values, $\mathscr{b}_{jl}$, is by analyzing the discrete-time Markov Chain of the backlogged demand at a single processing facility $j \in J$ at the beginning of each day. Letting the amount of backlog present at the beginning of the day at facility $j \in J$ represent a state of the Markov Chain, the transition probability from state $\mathfrak{n}$ to state $\mathfrak{m}$ is denoted $\mathfrak{p}_{\mathfrak{n}\mathfrak{m}}$ and is given by:

$$\mathfrak{p}_{\mathfrak{n}\mathfrak{m}} = \begin{cases} P\left(H_j + \mathfrak{n} \leq k_j\right) & \mathfrak{m} \geq \mathfrak{n} - k_j,\ \mathfrak{m} = 0 \\ P\left(H_j + \mathfrak{n} = k_j + \mathfrak{m}\right) & \mathfrak{m} \geq \mathfrak{n} - k_j,\ \mathfrak{m} > 0 \\ 0 & \mathfrak{m} < \mathfrak{n} - k_j \end{cases}$$

where H_j once again represents a generic total demand arriving at facility $j \in J$ on any day.

If we assume that the number of items that arrive at the facility each day follows a Poisson distribution with mean $\mathbb{E}\left[H_j\right]$ then the transition probabilities are given by:

$$p_{nm} = \begin{cases} \sum_{z=0}^{k_j - \mathfrak{n}} \frac{e^{-\mathbb{E}\left[H_j\right]} \mathbb{E}\left[H_j\right]^z}{z!} & \mathfrak{m} \geq \mathfrak{n} - k_j,\ \mathfrak{m} = 0 \\ \frac{e^{-\mathbb{E}\left[H_j\right]} \mathbb{E}\left[H_j\right]^{(k_j + \mathfrak{m} - \mathfrak{n})}}{(k_j + \mathfrak{m} - \mathfrak{n})!} & \mathfrak{m} \geq \mathfrak{n} - k_j,\ \mathfrak{m} > 0 \\ 0 & \mathfrak{m} < \mathfrak{n} - k_j \end{cases}$$

Using these transition probabilities, we can determine the steady state probabilities for various values of the average total arriving demand, $\mathbb{E}\left[H_j\right]$, as well as the average backlog at the facility. Figure 37 displays the results for a facility that has a daily processing capacity of 100 units. It clearly shows that as the average daily demand approaches the processing capacity, the average backlog of the facility increases. This is due to the fact that as the facility operates closer to capacity, the likelihood that the total arriving demand exceeds the processing capacity increases. The difference between the average backlog value when there are $\mathbb{E}\left[H_j\right] = l - 1$ and $\mathbb{E}\left[H_j\right] = l$ items of demand arriving at facility $j \in J$ each day provides a value for the incremental backlog cost, $\mathscr{b}_{jl}$, for $l \in \{1, \ldots, k_j - 1\}$.

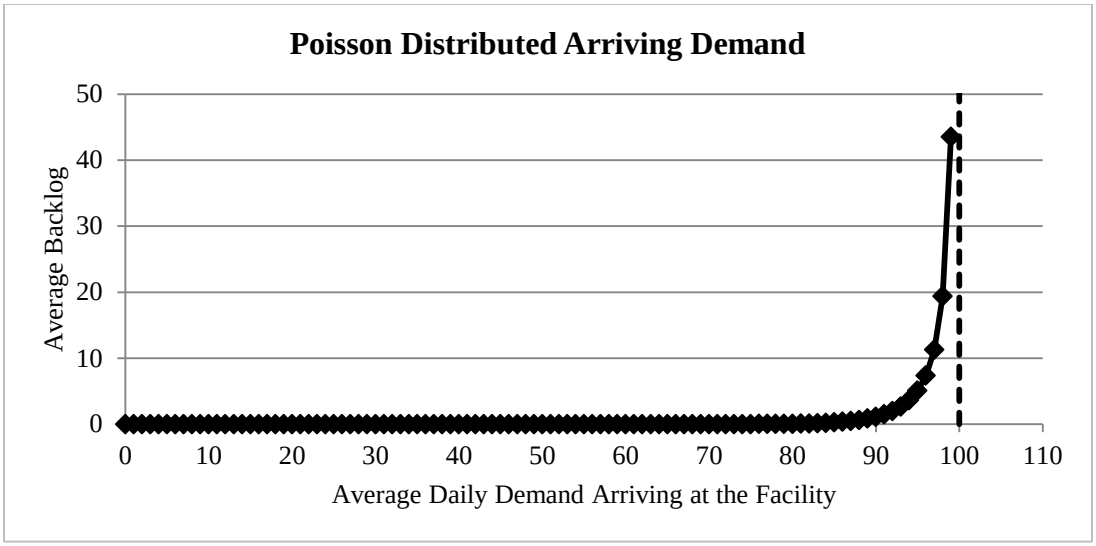

Figure 37: Average backlog at a processing facility that results from the average amount of total daily demand arriving at the facility; Poisson distributed arriving demands; $k_j = 100$

6.3.3 Model Summary and Challenges

The incremental backlog cost model presents another method of incorporating backlogging costs into a stochastic version of the capacitated facility location problem. While this approach allows us to incorporate the effect backlogged demand has on both the cost and the number of demands needing to be processed from one day to another without explicitly incorporating the backlog and processing variables, it is not without its challenges. One such challenge is in determining how large of a state space to consider. For many demand distributions, the state space is infinite, as there is a positive probability of having an arbitrarily large amount of demand arrive at a facility. However, a finite state space will likely provide a tight approximation for the desired $\mathscr{b}_{jl}$ values. Additionally, while we have only provided transition probabilities for the case when the amount of demand *arriving* at each facility follows a Poisson distribution, it is likely that the distribution of demand arising at the various demand sites is something other than i.i.d. Poisson distributions. Relaxing these assumptions on the

demand distributions pose additional analytical challenges as the decision regarding the allocation of demand sites (which may have spatially or temporally correlated demand as well as each having a different distribution of demand) directly impacts the distribution of the demand arriving at each processing facility.

6.4 Outbound Shipment Model

Thus far we have presented and discussed models that mitigate hard capacity constraints in the context of a system with physical demand (e.g., blood samples) that arrives at a processing facility (e.g., blood testing facility) and has intangible results (e.g., a report) that are reported to the customer. Incoming demands that exceed the daily processing capacity are physically held as backlog (at some cost) and are processed at a later date. However, with a slight change of notation we can modify the formulation to represent a system in which warehouses receive intangible demand orders and must ship physical goods to the demand sites. (Online retailers, including Amazon.com, are often examples of such a system.) The warehouse may have a policy that orders must be shipped within one day of the order being placed (i.e., the order needs to be en route to the demand site within one day; it does not have to arrive at the demand site within one day). If an item cannot be shipped within the specified timeframe, a penalty cost is incurred to account for expedited shipping or loss of goodwill. We refer to such an item as a "late item." The goal is to determine where to build warehouses and how to assign warehouses to demand sites so that the total location, transportation, and penalty cost is minimized, while ensuring that a specified service level is achieved.

Stochastic IMCLP Outbound Shipment Formulation

Letting J represent the set of candidate warehouse locations and I the set of demand sites, we can modify the original stochastic IMCLP notation to fit the outbound shipment context. While we present the model with the joint chance constraints, we note that substituting the hybrid or individual chance constraints for the joint chance constraint results in the corresponding outbound shipment formulation.

Modified Parameters

k_j Capacity of warehouse $j \in J$ in items shipped per day

$b \geq 0$ Penalty cost (in dollars) of shipping one item one day late

θ At any given time, the maximum desired number of late items at an individual warehouse

τ Maximum acceptable probability of having more than θ late items at any facility on any day

v_j Initial number of late items at warehouse $j \in J$ if warehouse j is located

Decision Variables

$$X_j = \begin{cases} 1 & \text{If we locate at warehouse } j \in J \\ 0 & \text{Otherwise} \end{cases}$$

$$Y_{ij} = \begin{cases} 1 & \text{If we assign warehouse } j \in J \text{ to fulfill demand from } i \in I \\ 0 & \text{Otherwise} \end{cases}$$

V_{jd}^{ω} Number of late items at warehouse $j \in J$ at the beginning of day $d \in D \cup (|D| + 1)$ in scenario $\omega \in \Omega$

W_{jd}^{ω} Number of items that are shipped from warehouse $j \in J$ on day $d \in D$ in scenario $\omega \in \Omega$

The modified model becomes:

[M-JCC]:

$$Min_{\boldsymbol{X},\boldsymbol{Y},\boldsymbol{V},\boldsymbol{W}} |D| \sum_{j\in J} f_j X_j + a \sum_{i\in I} \sum_{d=1}^{|D|} \sum_{j\in J} \mathbb{E}_{\xi}\left[\tilde{h}_{id}\right] t_{ij} Y_{ij} + b\ \mathbb{E}_{\xi}[\sum_{d=2}^{|D|+1} \sum_{j\in J} \tilde{V}_{jd}] \quad (6.8)$$

Subject to

$$\sum_{j\in J} Y_{ij} = 1 \qquad \forall i \in I \quad (2.2)$$

$$Y_{ij} - X_j \leq 0 \qquad \forall i \in I; j \in J \quad (2.3)$$

$$\mathbb{P}\left(\tilde{V}_{jd} \leq\ \theta;\ \ \forall j \in J;\ d \in D \cup (|D|+1)\}\right) \geq 1 - \tau \quad (6.9)$$

$$V_{j,d+1}^{\omega} - V_{jd}^{\omega} + W_{jd}^{\omega} = \sum_{i\in I} h_{id}^{\omega} Y_{ij} \qquad \forall j \in J; d \in D; \omega \in \Omega \quad (6.10)$$

$$W_{jd}^{\omega} \leq k_j X_j \qquad \forall j \in J; d \in D; \omega \in \Omega \quad (6.11)$$

$$V_{j,1}^{\omega} = v_j X_j \qquad \forall j \in J; \omega \in \Omega \quad (6.12)$$

$$V_{jd}^{\omega} \geq 0 \qquad \forall j \in J; d \in D \cup (|D|+1); \omega \in \Omega \quad (6.13)$$

$$W_{jd}^{\omega} \geq 0 \qquad \forall j \in J; d \in D; \omega \in \Omega \quad (6.14)$$

$$X_j \in \{0,1\} \qquad \forall j \in J \quad (2.4)$$

$$Y_{ij} \in \{0,1\} \qquad \forall i \in I; j \in J \quad (3.6)$$

The major difference between the JCC formulation and the modified M-JCC is that the new formulation does not have incoming demand that is being shipped from the demand sites to the warehouses. Instead, the warehouses ship the demand to the demand sites. Thus, we do not need to include a warm-up period in formulation M-JCC. (Note that the time horizon is now $|D|$ rather than $|D| - t^*$). Additionally, constraints (6.10) use the demand that is generated on day d to calculate the number of late items on day $d + 1$, rather than using demand that was generated $d - t_{ij}$ days ago. We use constraints (6.12) to initialize the number of late items at the beginning of the horizon. Additionally, the joint chance constraint represented by (6.9) now states that there

cannot be more than θ late items present at any warehouse on any day in more than $\tau * 100\%$ of the scenarios. If a scenario contains one year's worth of data and the model contains ten scenarios, this is equivalent to saying that we must always have less than θ late items in at least $\lfloor (1-\tau) * 10 \rfloor$ of the years.

Cyclic Allocation Outbound Shipment Formulation

A similar outbound shipment perspective can be formulated for the IMCLP with cyclic allocations. In addition to the modified notation already discussed, we introduce the redefined sets, parameters, and decision variables:

Modified Sets and Parameters

$D_p := \{d \in D : d \bmod |P| = p\}$ Set of days in D corresponding to cycle day $p \in P$

k_j^p Capacity of warehouse $j \in J$ on cycle day $p \in P$ in terms of items shipped per day

Decision Variables

$Y_{ij}^p = \begin{cases} 1 & \text{If we assign warehouse } j \in J \text{ to fulfill demand from } i \in I \text{ on cycle day } p \in P \\ 0 & \text{Otherwise} \end{cases}$

The outbound IMCLP cyclic allocation model can then be presented as:

$$Min_{\boldsymbol{X,Y,V,W}} \quad |D| \sum_{j \in J} f_j X_j + a \sum_{p \in P} \sum_{d \in D_p} \sum_{j \in J} \sum_{i \in I} h_{id} t_{ij} Y_{ij}^p + b \sum_{d=2}^{|D|+1} \sum_{j \in J} V_{jd} \tag{6.15}$$

$Subject\ to$

$$\sum_{j \in J} Y_{ij}^p = 1 \qquad \forall i \in I;\ p \in P \tag{4.2}$$

$$Y_{ij}^p \leq X_j \qquad \forall i \in I; j \in J; p \in P \tag{4.3}$$

$$V_{j,d+1} = V_{jd} + \sum_{i \in I} h_{i,d} Y_{ij}^{d \bmod |P|} - W_{jd} \qquad \forall j \in J; d \in D \tag{6.16}$$

$$W_{jd} \leq k_j^p X_j \qquad \forall j \in J; p \in P; d \in D_p \tag{4.5}$$

$$V_{j,1} = v_j X_j \qquad \forall j \in J \tag{6.17}$$

$$X_j \in \{0,1\} \qquad \forall j \in J \qquad (2.4)$$

$$Y_{ij}^{p} \in \{0,1\} \qquad \forall i \in I; j \in J; p \in P \qquad (4.6)$$

$$V_{jd} \geq 0 \qquad \forall j \in J; d \in D \cup (|D| + 1) \qquad (6.18)$$

$$W_{jd} \geq 0 \qquad \forall j \in J; d \in D \qquad (6.19)$$

6.5 Additional Methods of Incorporating Capacity Flexibility

This dissertation has focused on incorporating short-term capacity flexibility into facility location models by allowing a processing facility to accept demands in excess of its processing capacity and store the excess demand in backlog until it can be processed. While this is one way of incorporating capacity flexibility, there are numerous other avenues that are used in practice. For example, a processing facility may utilize employee overtime, hire temporary workers, outsource excess demands, or add additional physical capacity (e.g., an additional processing machine or assembly line) to capture endogenous capacity flexibility. Additionally, exogenous capacity flexibility can be captured through stochastic programming. In this section we formulate extensions to the IMCLP that incorporate both endogenous and exogenous methods of capturing capacity flexibility.

6.5.1 Endogenous Capacity Flexibility

Rather than requiring all demand in excess of the processing capacity to be held in backlog, we can formulate an extension to the IMCLP which allows extra daily capacity to be purchased on a day-to-day basis. In this model, processing facility $j \in J$ starts each day with the capacity to process k_j items. However, if the arriving demand exceeds the processing capacity, the facility manager at $j \in J$ has the option of purchasing extra capacity for that day at a unit cost

of $\hat{b}_j$ rather than incurring a facility specific backlog penalty cost b_j. In this discussion, we assume that the extra capacity can be purchased after the daily amount of arriving demand is realized. However, a stochastic extension can also be formulated to address the situation in which a facility manager must decide one day in advance (before knowing the exact amount of incoming demand) how much extra capacity to purchase. We leave this extension as an area of future work.

We let $\widehat{K}_{jd}$ be a non-negative decision variable that represents the extra capacity to purchase at facility $j \in J$ on day $d \in \{t^* + 1, \dots, |D|\}$ and $k_j^{max} \geq 0$ be a parameter representing the maximum amount of additional capacity facility $j \in J$ can purchase each day (k_j^{max} may be unbounded from above). We formulate the model as follows:

$$Min_{\boldsymbol{X},\boldsymbol{Y},\boldsymbol{V},\boldsymbol{W},\widehat{\boldsymbol{K}}}(|D| - t^*)\sum_{j\in J} f_j X_j + a\sum_{d=t^*+1}^{|D|}\sum_{j\in J}\sum_{i\in I} h_{id} t_{ij} Y_{ij} + \sum_{d=t^*+2}^{|D|+1}\sum_{j\in J} b_j V_{jd} + \sum_{d=t^*+1}^{|D|}\sum_{j\in J} \hat{b}_j \widehat{K}_{jd} \quad (6.20)$$

Subject to

$$\sum_{j\in J} Y_{ij} = 1 \qquad \forall i \in I \quad (2.2)$$

$$Y_{ij} \leq X_j \qquad \forall i \in I; j \in J \quad (2.3)$$

$$V_{j,d+1} - V_{jd} - \sum_{i\in I} h_{i,d-t_{ij}} Y_{ij} + W_{jd} = 0 \qquad \forall j \in J; d \in \{t^* + 1, \dots, |D|\} \quad (3.3)$$

$$\widehat{K}_{jd} \leq k_j^{max} X_j \qquad \forall j \in J; d \in \{t^* + 1, \dots, |D|\} \quad (6.21)$$

$$W_{jd} \leq k_j X_j + \widehat{K}_{jd} \qquad \forall j \in J; d \in \{t^* + 1, \dots, |D|\} \quad (6.22)$$

$$V_{j,t^*+1} = v_j X_j \qquad \forall j \in J \quad (3.5)$$

$$X_j \in \{0,1\} \qquad \forall j \in J \quad (2.4)$$

$$Y_{ij} \in \{0,1\} \qquad \forall i \in I; j \in J \quad (3.6)$$

$$V_{jd} \geq 0 \qquad \forall j \in J; d \in \{t^* + 1, \dots, |D| + 1\} \quad (3.7)$$

$$W_{jd} \geq 0 \qquad \forall j \in J; d \in \{t^* + 1, \dots, |D|\} \qquad (3.8)$$

$$\widehat{K}_{jd} \geq 0 \qquad \forall j \in J; d \in \{t^* + 1, \dots, |D|\} \qquad (6.23)$$

Alternatively, extra daily capacity may only be available to purchase in increments of size $\bar{k}_j$. Thus, some of the purchased extra capacity may be unused at the end of the day. To amend the above model to account for incremental daily capacity, let $\overline{K}_{jd}$ be a nonnegative integer decision variable indicating the number of increments purchased, replace $\widehat{K}_{jd}$ with $\bar{k}_j \overline{K}_{jd}$ in (6.20) - (6.22), and replace constraint (6.23) with

$$\overline{K}_{jd} \geq 0, \ integer \qquad \forall j \in J; d \in \{t^* + 1, \dots, |D|\} \qquad (6.24)$$

While the aforementioned models allow capacity to be added on a day-by-day basis, the facility manager may not have the ability to change the capacity each day of the planning horizon. Instead, the manager may prefer a cyclic method of adding capacity in which facility $j \in J$ has a base capacity to process k_j items each day and the option to increase capacity by $\widehat{K}_j^p$ (a decision variable) on cycle day $p \in P$ at a cost of $\hat{b}_j$ per unit of capacity increase. The cyclic method of adding capacity can be incorporated into the cyclic allocation model by adding the term $\sum_{p \in P} \sum_{j \in J} \hat{b}_j \widehat{K}_j^p |D_p|$ to objective function (4.1), adding $\widehat{K}_j^p$ to the right hand side of constraint (4.5), and adding constraints

$$\widehat{K}_j^p \leq k_j^{max} X_j \qquad \forall j \in J; p \in P \qquad (6.25)$$

and

$$\widehat{K}_j^p \geq 0, \ integer \qquad \forall j \in J; p \in P. \qquad (6.26)$$

6.5.2 Exogenous Capacity Fluctuations

While facility managers may have many methods of affecting the amount of capacity available each day, there also exist exogenous capacity fluctuations that managers have very little (if any) control over. Examples of such exognous fluctuations include capacity degregations due to machine breakdowns and quality control issues. Workers who assemble products or pick customer orders in a warehouse also add an element of uncertainty because each employee completes tasks at a different rate.

In this section we will present two extensions of the IMCLP that incorporate exogenous capacity fluctuations. The first is a stochastic version of the IMCLP in which the capacity is an uncertain model input. While this formulation will not incorporate uncertain demand as we do in Chapter 5, models that incorporate both uncertain demand and capacity can be developed by merging the two stochstic formulations. The second extension depicts employee training by modeling capacity that increases naturally over a period of time during the training period and eventually reaches some asymptotic behavior when the employees are fully trained.

Stochastic Daily Capacity

Let $\ddot{\Omega}$ represent the finite set of all possible capacity scenarios, and $\ddot{p}^{\ddot{\omega}}$ represent the probability of a particular scenario $\ddot{\omega} \in \ddot{\Omega}$. The parameter $\tilde{\tilde{k}}_{jd}$ corresponds to the random daily capacity available at processing facility $j \in J$ on day $d \in \{t^* + 1, \dots, |D|\}$. While $\tilde{\tilde{k}}_{jd}$ is a random parameter, $\ddot{k}_{jd}^{\ddot{\omega}}$ represents a realization of the random capacity $\tilde{\tilde{k}}_{jd}$ available at facility $j \in J$ on day $d \in \{t^* + 1, \dots, |D|\}$ in scenario $\ddot{\omega} \in \ddot{\Omega}$. We also let $\ddot{\xi}(\ddot{\omega}) \in \mathbb{N}_0^{|J|\times|D|}$ be a vector containing the capacity realizations $\ddot{k}_{jd}^{\ddot{\omega}}, \forall j \in J, d \in \{t^* + 1, \dots, |D|\}$ corresponding to scenario $\ddot{\omega} \in \ddot{\Omega}$.

The decision variables in the exogenous capacity fluctuation model include the scenario-independent location (X_j) and allocation (Y_{ij}) decisions as well as scenario-dependent backlog ($\ddot{V}_{jd}^{\ddot{\omega}}$) and processing ($\ddot{W}_{jd}^{\ddot{\omega}}$) variables. Since the backlog (and processing) variables cannot be determined until after the daily capacity is realized, we let $\tilde{\ddot{V}}_{jd}$ represent the stochastic counterpart of $\ddot{V}_{jd}^{\ddot{\omega}}$. With this notation, we can formulate the exogenous capacity fluctuation model as:

$$Min_{\boldsymbol{X},\boldsymbol{Y},\ddot{\boldsymbol{V}},\ddot{\boldsymbol{W}}}(|D|-t^*)\sum_{j\in J} f_j X_j + a\sum_{i\in I}\sum_{d=t^*+1}^{|D|}\sum_{j\in J} h_{id}t_{ij}Y_{ij} + b\,\mathbb{E}_{\tilde{\xi}}[\sum_{d=t^*+2}^{|D|+1}\sum_{j\in J}\tilde{\ddot{V}}_{jd}] \quad (6.101)$$

Subject to

$$\sum_{j\in J} Y_{ij} = 1 \qquad \forall i \in I \qquad (2.2)$$

$$Y_{ij} \le X_j \qquad \forall i \in I; j \in J \qquad (2.3)$$

$$\ddot{V}_{j,d+1}^{\ddot{\omega}} - \ddot{V}_{jd}^{\ddot{\omega}} + \ddot{W}_{jd}^{\ddot{\omega}} = \sum_{i\in I} h_{i,d-t_{ij}} Y_{ij} \qquad \forall j \in J; d \in \{t^*+1,\dots,|D|\}; \ddot{\omega} \in \ddot{\Omega} \qquad (6.27)$$

$$\ddot{W}_{jd}^{\ddot{\omega}} \le \ddot{k}_{jd}^{\ddot{\omega}} X_j \qquad \forall j \in J; d \in \{t^*+1,\dots,|D|\}; \ddot{\omega} \in \ddot{\Omega} \qquad (6.28)$$

$$\ddot{V}_{j,t^*+1}^{\ddot{\omega}} = v_j X_j \qquad \forall j \in J; \ddot{\omega} \in \ddot{\Omega} \qquad (6.29)$$

$$\ddot{V}_{jd}^{\ddot{\omega}} \ge 0 \qquad \forall j \in J; d \in \{t^*+1,\dots,|D|+1\}; \ddot{\omega} \in \ddot{\Omega} \qquad (6.30)$$

$$\ddot{W}_{jd}^{\ddot{\omega}} \ge 0 \qquad \forall j \in J; d \in \{t^*+1,\dots,|D|\}; \ddot{\omega} \in \ddot{\Omega} \qquad (6.31)$$

$$X_j \in \{0,1\} \qquad \forall j \in J \qquad (2.4)$$

$$Y_{ij} \in \{0,1\} \qquad \forall i \in I; j \in J \qquad (3.6)$$

Employee Learning

Suppose that at the beginning of employee training, the daily processing facility capacity at $j \in J$ is $k_j - \underline{k_j}$, but that it increases naturally over a period of time at a rate of k_j^{rate} and then

remains at a constant level, k_j, when employees have been fully trained on the work processes. We can model such a situation as an exponential increase in capacity and can easily implement this into the IMCLP by changing constraint (3.4) from

$$W_{jd} \leq k_j X_j \qquad \forall j \in J; d \in \{t^* + 1, \dots, |D|\} \qquad (3.4)$$

to

$$W_{jd} \leq \left(k_j - \underline{k_j} e^{-k_j^{rate}(d-(t^*-1))}\right) X_j \qquad \forall j \in J; d \in \{t^* + 1, \dots, |D|\}. \qquad (6.32)$$

6.6 Conclusions and Contributions

This dissertation argued that the hard capacity constraints that are often employed in location modeling are unrealistic since facility managers have numerous operational tools that allow a facility to accept items in excess of the stated processing capacity level. As such, we developed a new approach to modeling capacitated facility location models that takes into account the likelihood that the total amount of demand arriving at a processing facility may exceed the daily processing capacity on any day of the planning horizon. When such exceedances occur, our models permit backlogging excess demands and processing them at a later date. To ensure adequate service metrics are met, we assess a daily penalty cost for each unit of backlogged demand and, in models that consider stochastic demands, we also incorporated chance constraints to impose limits on the quantity of backlogged demand.

We reported the difference in facility locations, demand allocations, and total cost between the CFLP and our IMCLP approach. In particular, we have shown that the CFLP and IMCLP often do not locate the same facilities (and in some instances they locate a different number of facilities), and therefore also do not allocate demands to the same processing facilities. Since the CFLP does not incorporate the cost of backlog, the CFLP underestimates the

total cost. The cost benefit of using the IMCLP instead of the CFLP is particularly notable for application contexts in which it is highly desirable to process all of the demands on the day they arrive at a processing facility. Such an underestimate could have significant implications on budget predictions.

A key benefit of the IMCLP model is that it incorporates the penalty cost associated with the reality that on some days the total amount of demand arriving at a processing facility may exceed the daily processing capacity. The IMCLP formulation inherently assumes that demands do not expire; demands can be postponed indefinitely as long as the penalty cost is paid. We realize that in many situations this is not realistic and presented additional constraints that could be added to ensure demands are processed within a specified amount of time.

Our model formulations can use a dataset as a direct input into the model rather than specifying deterministic demands at an aggregate level by using parameter estimation techniques. This enabled us to develop a day-of-the-week allocation policy that considers day-to-day variations in the daily processing capacity levels of a set of candidate processing facilities and/or systematic day-to-day demand variations. For example, our model is able to adjust the allocation policy to account for some facilities being closed on weekends or for days of the week that generate more demand than other days of the week. Through computational studies, we showed that incorporating this cyclic allocation approach into the model resulted in a significant cost savings as compared to models that do not incorporate this approach, such as the IMCLP. Furthermore, we noted that the model can readily be restructured to deal with other cycles, such as annual cycles with demand assignments that either change seasonally or monthly in response to time-varying demand patterns.

We also presented multiple constraints that can be added to the model formulation to achieve various operational goals. These include (1) restricting the number of processing facilities to which a demand site can be assigned, (2) introducing a weekday/weekend allocation policy, and (3) restricting the maximum allowed travel time. Our results suggest that a significant cost benefit can be achieved by implementing the cyclic allocation model as compared to capacitated facility location models that require a static allocation policy. Much of this benefit can be achieved even when additional constraints of the form outlined above are imposed.

While the IMCLP and IMCLP with cyclic allocations consider deterministic demands, our third modeling framework incorporated demand stochasticity into the IMCLP using a scenario based stochastic optimization approach in which a scenario corresponds to a realization of the daily demand for every day within the time horizon. In addition to assessing a penalty cost associated with each day an item spends in backlog (as is done in the IMCLP), we considered three different types of chance constraints to restrict the number of demands that are backlogged to a predetermined threshold. We first incorporated joint chance constraints that ensure the probability of any processing facility having a backlog level above the threshold on any day of the planning horizon is sufficiently small. We also modelled individual chance constraints on the amount of backlog at each facility each day, as well as a hybrid approach that accounts for the number of processing facilities that exceed the stated maximum backlog level on any day of the planning horizon. We then presented mixed-integer programming reformulations of the chance constraints that incorporate a finite number of scenarios from a given known demand distribution.

Through computational experiments, we showed that the stochastic, chance constrained version of the IMCLP can produce different location and allocation solutions than the IMCLP,

and that even small differences in their allocations can drastically impact the amount of items held in backlog. Furthermore, if the decision maker has a limit on the desired maximum amount if items backlogged at any given time, it is important that the decision maker utilize the stochastic model; imposing the limit on a single demand scenario (i.e., using the IMCLP) often resulted in infeasible solutions to the joint chance constrained stochastic IMCLP due to the demand uncertainty.

The resulting models were solved using two different decomposition schemes and their performance is compared to that of a generic solver. The first approach decomposed the problem into two stages: the long-term (i.e., location and allocation) decisions were determined in the first stage while the daily (i.e., processing and backlog) decisions were determined in the second stage. Benders decomposition was used to solve the resulting formulation. The second decomposition scheme capitalized on the problem structure by utilizing a three-stage structure. In particular, given a feasible first-stage location and allocation solution, we readily determined the optimal second-stage processing and backlog decisions as well as the third-stage auxiliary variables that verify whether the joint chance constraint is satisfied by inspection. If the joint chance constraint was violated, a corresponding cut was added to the first-stage problem.

Computational results for the IMCLP, IMCLP with cyclic allocations, and the chance constrained IMCLP were presented for large-scale data sets generated from 2010 U.S. census population data by county. Using diverse data instances, we demonstrated the benefits gained from using a data-driven model with short term capacity flexibility. In each of these modelling contexts, we showed that the location and allocation decisions obtained from our models can result in significantly reduced costs and improved service metrics when compared to traditional models that do not account for the likelihood that demands may exceed capacity on some days.

Finally, we presented a number of avenues for future research including a queueing approach, incremental backlog cost model, and an outbound shipment model. Furthermore, although the focus of this dissertation was the effect of incorporating the option to backlog excess demands, managers have additional operational tools that relax the traditional inflexible capacity constraints, including overtime, adding additional production shifts, and outsourcing. Degradations in capacity are also a reality. That is, the actual capacity on any day may be less than the nominal capacity due to machine breakdowns, labor slowdowns, or power outages. As such, we briefly discussed formulations for incorporating additional types of exogenous and endogenous capacity flexibility into a data-driven capacitated facility location model.

APPENDICES

APPENDIX A: The Data-Driven CFLP

We formulate the data-driven CFLP as follows:

$$Min_{\boldsymbol{X,Y}} \quad (|D| - t^*) \sum_{j \in J} f_j X_j + a \sum_{d=t^*+1}^{|D|} \sum_{j \in J} \sum_{i \in I} h_{id} t_{ij} Y_{ij} \tag{A.1}$$

$Subject\ to$

$$\sum_{j \in J} Y_{ij} = 1 \qquad \forall i \in I \tag{2.2}$$

$$Y_{ij} \leq X_j \qquad \forall i \in I; j \in J \tag{2.3}$$

$$\sum_{i \in I} h_{i,d-t_{ij}} Y_{ij} \leq k_j X_j \qquad \forall j \in J; d \in \{t^* + 1, \dots, |D|\} \tag{A.2}$$

$$X_j \in \{0,1\} \qquad \forall j \in J \tag{2.4}$$

$$Y_{ij} \in \{0,1\} \qquad \forall i \in I; j \in J \tag{3.6}$$

The objective function (A.1) minimizes the sum of the fixed facility and transportation costs over a time horizon of $(|D| - t^*)$ days. We note that the second term of the objective function includes the daily demands, h_{id}, rather than the aggregated average demand, h_i, used in the CFLP objective function (3.1). Constraints (2.2) and (2.3), respectively, ensure that each demand site $i \in I$ is assigned to exactly one facility and that demands can only be assigned to located facilities. Constraints (A.2) are the data-driven capacity constraints that limit the amount of demand that arrives at a processing facility to be no more than the daily processing capacity of the facility. Thus, all demands are able to be processed on the day they arrive. Constraints (A.2) differ from the CFLP capacity constraints (2.6) in that constraints (A.2) are enforced on the

individual daily demands while constraints (2.6) are enforced on the average demands. Constraints (2.4) and (3.6) are standard binary and non-negativity constraints, as in the CFLP.

APPENDIX B: Effect of Population Weight

As discussed in Section 3.4, the daily facility location costs are calculated by specifying a daily base cost plus an additional cost calculated form the associated demand (population) generated at that location. Specifically, we use the formula

$$z + w\rho_j$$

to generate the daily facility location cost for each facility $j \in J$, where z denotes the daily base fixed cost, ρ_j represents the population of the candidate processing facility j, and w is the population weight.

Table 34 describes the effect of the population weight on the daily facility location cost parameters, as well as the optimal IMCLP solution for the 50 demand, 50 candidate node instance, a daily base fixed cost of $z = \$10{,}000$, Poisson distributed daily demands with a mean equal to 1/10,000 of the county population, and unit transportation and backlogging costs of $a = 1$ and $b = 3$, respectively. The first column indicates the weight attributed to the population. The second and third columns display the daily facility location cost of the least costly (Node 50: Du Page, IL) and most costly (Node 1: Los Angeles, CA) candidate processing facilities. Notice that as the population weight increases, the disparity between the minimum and maximum costs increases, as does the average daily facility location cost of the candidate facilities (column 4). The optimal solution for all five values of w locates five processing facilities. The fifth column enumerates the ranks of the five facilities that are located (recall that that lower ranks have a higher facility location cost), the sixth column reports the average of the five ranks, and the last column displays the average daily facility location cost of the facilities that are located in the

optimal solution. As expected, when the population weight, w, increases, it becomes essential to locate at less expensive candidate processing facilities (i.e., higher ranked facilities). Maps of the optimal facility locations are presented in Figure 38.

Table 34: Effect of population weight on location cost and optimal solution; Poisson distributed demand; a=1, b=3

w	Daily Facility Location Cost Parameters			Optimal Solution		
	Min (Node 50)	Max (Node 1)	Avg.	Ranks Located	Avg. Rank Located	Avg. Daily Cost of Located
$0	$10,000	$10,000	$10,000	1, 2, 3, 7, 13	5.2	$10,000
$0.0001	$10,092	$10,982	$10,183	2, 3, 6, 7, 13	6.2	$10,335
$0.001	$10,917	$19,819	$11,833	12, 39, 43, 45, 50	37.8	$11,171
$0.01	$19,169	$108,186	$28,331	39, 41, 43, 45, 50	43.6	$19,602
$0.1	$101,692	$991,861	$193,311	41, 45, 47, 49, 50	46.4	$103,345

Figure 38: The optimal facility locations listed in Table 34

APPENDIX C: Proof of the Validity of Feasibility Cuts (5.89)

Individual Chance Constraints

To show the validity of cut (5.89) we begin by recalling $\bar{I}_j^{\psi} \coloneqq \left\{i \in I : Y_{ij}^{\psi} = 1\right\}$ represents demand sites allocated to processing facility $j \in J$ in infeasible solution $\psi \in \underline{\Psi}_j$. Let $\bar{\mathcal{F}} \coloneqq \{(\boldsymbol{X}, \boldsymbol{Y}) : \exists \boldsymbol{V}, \boldsymbol{W}, \boldsymbol{Z}$: (2.2) - (2.4), (3.6), (5.15) - (5.19), (5.22), and (5.80) - (5.82) are satisfied$\}$. For any feasible $(\boldsymbol{X}, \boldsymbol{Y})$ solution to MIP-ICC, $(\boldsymbol{X}, \boldsymbol{Y}) \in \bar{\mathcal{F}}$. Suppose another solution, $\left(\boldsymbol{X}^{\psi}, \boldsymbol{Y}^{\psi}\right)$ satisfies constraints (2.2) - (2.4), (3.6), and (5.22) but $\left(\boldsymbol{X}^{\psi}, \boldsymbol{Y}^{\psi}\right) \notin \bar{\mathcal{F}}$. It follows that $(\boldsymbol{X}, \boldsymbol{Y}) \neq \left(\boldsymbol{X}^{\psi}, \boldsymbol{Y}^{\psi}\right)$.

This further implies that the allocation decisions corresponding to the two solutions are different, i.e., $\boldsymbol{Y} \neq \boldsymbol{Y}^{\psi}$. To see this, note that if $\boldsymbol{X} = \boldsymbol{X}^{\boldsymbol{\psi}}$, then it must be that $\boldsymbol{Y} \neq \boldsymbol{Y}^{\psi}$ in order for $(\boldsymbol{X}, \boldsymbol{Y}) \neq \left(\boldsymbol{X}^{\psi}, \boldsymbol{Y}^{\psi}\right)$. If instead $\boldsymbol{X} \neq \boldsymbol{X}^{\boldsymbol{\psi}}$, then $\exists \hat{\jmath} \in J$ such that $X_{\hat{\jmath}} \neq X_{\hat{\jmath}}^{\psi}$. Thus, either $X_{\hat{\jmath}} = 1$ and $X_{\hat{\jmath}}^{\psi} = 0$, or $X_{\hat{\jmath}} = 0$ and $X_{\hat{\jmath}}^{\psi} = 1$. Suppose $X_{\hat{\jmath}} = 1$ and $X_{\hat{\jmath}}^{\psi} = 0$. Then, for every $i \in I$ such that $Y_{i\hat{\jmath}} = 1$ we must have $Y_{i\hat{\jmath}}^{\psi} = 0$ since facility $\hat{\jmath}$ is not located in solution $\left(\boldsymbol{X}^{\psi}, \boldsymbol{Y}^{\psi}\right)$. Thus, $\boldsymbol{Y} \neq \boldsymbol{Y}^{\psi}$. An analogous reasoning holds for the case of $X_{\hat{\jmath}} = 0$ and $X_{\hat{\jmath}}^{\psi} = 1$.

Furthermore, the reason $\left(\boldsymbol{X}^{\psi}, \boldsymbol{Y}^{\psi}\right) \notin \bar{\mathcal{F}}$ must be because it violates the collective group of constraints (5.80) - (5.82). Since $(\boldsymbol{X}, \boldsymbol{Y})$ is a feasible solution to MIP-ICC, there must exist $i' \in I$ such that $Y_{i'\bar{\jmath}}^{\psi} = 1$ and $Y_{i'\bar{\jmath}} = 0$ for some $\bar{\jmath} \in J$ and $\bar{d} \in \{t^* + 1, \dots, |D| + 1\}$ with $\sum_{\omega' \in \Omega'} p^{\omega'} \bar{Z}_{\bar{\jmath}\bar{d}}^{\omega'\psi} > \tau$. Therefore,

$$\sum_{i\in I_{\bar{J}}^{\psi}} Y_{i\bar{J}} = Y_{i'\bar{J}} + \sum_{i\in I_{\bar{J}}^{\psi}\setminus i'} Y_{i\bar{J}}$$

$$= 0 + \sum_{i\in I_{\bar{J}}^{\psi}\setminus i'} Y_{i\bar{J}}$$

$$\leq |I_{\bar{J}}^{\psi}\setminus i'|$$

$$= \left|I_{\bar{J}}^{\psi}\right| - 1$$

Therefore, the inequality (5.89) holds for any feasible $(\boldsymbol{X}, \boldsymbol{Y})$ solution to MIP-ICC. □

Hybrid Chance Constraints

The proof that cut (5.89) is a valid feasibility cut for MIP-JCC is similar to the proof for the MIP-ICC with a few minor changes. Instead of using $\bar{\mathcal{F}}$ as the proof for the MIP-ICC does, the MIP-HCC proof utilizes $\hat{\mathcal{F}} \coloneqq \{(\boldsymbol{X}, \boldsymbol{Y}): \exists \boldsymbol{V}, \boldsymbol{W}, \boldsymbol{Z}$: (2.2) - (2.4), (3.6), (5.15) - (5.19), (5.22), and (5.83) - (5.85) are satisfied$\}$. Additionally, the reason $\left(\boldsymbol{X}^{\psi}, \boldsymbol{Y}^{\psi}\right) \notin \hat{\mathcal{F}}$ must be because violates the collective group of constraints (5.83) - (5.85). Since $(\boldsymbol{X}, \boldsymbol{Y})$ is a feasible solution to MIP-HCC, there must exist $i' \in I$ such that $Y_{i'\bar{J}}^{\psi} = 1$ and $Y_{i'\bar{J}} = 0$ for some $\bar{J} \in J$ with $\sum_{\omega'\in\Omega'} p^{\omega'} \hat{Z}_{\bar{J}}^{\omega'\psi} > \tau$.

BIBLIOGRAPHY

BIBLIOGRAPHY

Adcock, D.M., Lippi, G., & Favaloro, E.J. 2012, "Quality standards for sample processing, transportation, and storage in hemostasis testing," *Seminars in Thrombosis and Hemostasis*, vol. 38, no. 6, pp. 576-585.

Agar, M.C. & Salhi, S. 1998, "Lagrangean heuristics applied to a variety of large capacitated plant location problems," *Journal of the Operational Research Society*, vol. 49, no. 10, pp. 1072-1084.

Aghezzaf, E. 2005, "Capacity planning and warehouse location in supply chains with uncertain demands," *Journal of the Operational Research Society*, vol. 56, no. 4, pp. 453-462.

Ahuja, R.K., Orlin, J.B., Pallottino, S., Scaparra, M.P. & Scutellà, M.G. 2004, "A multi-exchange heuristic for the single-source capacitated facility location problem," *Management Science*, vol. 50, no. 6, pp. 749-760.

Akinc, U. & Khumawala, B.M. 1977, "An efficient branch and bound algorithm for the capacitated warehouse location problem," *Management Science*, vol. 23, no. 6, pp. 585-594.

Albareda-Sambola, M., Fernández E., & Saldanha da Gama, F. 2011, "The facility location problem with Bernoulli demands," *Omega*, vol. 39, pp. 335-345.

Balcik, B. & Beamon, B.M. 2008, "Facility location in humanitarian relief," *International Journal of Logistics: Research and Applications*, vol. 11, no. 2, pp. 101-121.

Balinski, M.L. 1965, "Integer programming: methods, uses, computation," *Management Science*, vol. 12, no. 3, pp. 253-313.

Balinski, M.L. & Wolfe, P. 1963, "On Benders decomposition and a plant location problem," *ARO-27, Mathematica, (Princeton, NJ, 1963)*.

Barceló, J. & Casanovas, J. 1984, "A heuristic lagrangean algorithm for the capacitated plant location problem," *European Journal of Operational Research*, vol. 15, no. 2, pp. 212-226.

Baron, O., Berman, O. & Krass, D. 2008, "Facility location with stochastic demand and constraints on waiting time," *Manufacturing & Service Operations Management*, vol. 10, no. 3, pp. 484-505.

Beasley, J.E. 1993, "Lagrangean relaxation" in *Modern Heuristic Techniques for Combinatorial Problems*, ed. C.R. Reeves, Blackwell Scientific, Oxford, pp. 243–303.

Bell, T. & Church, R. 1985, “Location-allocation modeling in archaeological settlement pattern research: some preliminary applications,” *World Archaeology*, vol. 16, pp. 354-371.

Benders, J.F. 1962, “Partitioning procedures for solving mixed-variables programming problems,” *Numerische Mathematik*, vol. 4, pp. 238-252.

Bertsekas, D. 2003, *Linear Network Optimization: Algorithms and Codes*, The MIT Press.

Bhat, U.N. 2008, “The general queue G/G/1 and approximations,” chapter 9, *An Introduction to Queueing Theory*, Springer, pp. 169-183.

Birge, J.R. & Louveaux F. 2011, *Introduction to Stochastic Programming*, 2nd edn, Springer, New York.

Bramel, J. & Simchi-Levi, D. 1995, “A location based heuristic for general routing problems,” *Operations Research*, vol. 43, pp. 649-660.

Cornuéjols, G., Nemhauser, G.L. & Wolsey, L.A. 1990, "The uncapacitated facility location problem" in *Discrete Location Theory*, eds. P.B. Mirchandani & R.L. Francis, Wiley, New York, pp. 119-171.

Daskin, M.S. 2008, “What you should know about Location Modeling,” *Naval Research Logistics*, vol 55, no. 4, pp. 283-294.

Daskin, M.S. 2010, Service Science, Wiley, New York. pp. 189-193.

Daskin, M.S. 2013, *Network and discrete location: models, algorithms, and applications,* 2nd edn, Wiley, New York.

Daskin, M.S., Hopp, W.J., & Medina, B. 1992, “Forecast horizons and dynamic facility location planning,” *Annals of Operations Research*, vol. 40, pp. 125-151..

Daskin, M.S. & Jones, P.C. 1993, “A new approach to solving applied location/allocation problems,” *Microcomputers in Civil Engineering*, vol. 8, pp. 409-421.

Daskin, M.S., Snyder, L.V. & Berger, R.T. 2005, "Facility location in supply chain design" in *Logistics Systems: Design and Optimization*, eds. A. Langevin & D. Riopel, Springer, New York, pp. 39-65.

Davis, A., Mehrotra, S., Holl, J., & Daskin, M.S. 2014, “Nurse staffing under demand uncertainty to reduce costs and enhance patient safety,” *Asia –Pacific Journal of operational Research* vol. 31, no. 1.

Davis, P.S. & Ray, T.L. 1969, "A branch-bound algorithm for capacitated facilities location problem," *Naval Research Logistics Quarterly*, vol. 16, pp. 331-344.

Delmaire, H., Diaz, J.A., Fernandez, E. & Ortega, M. 1999, "Reactive GRASP and TABU search based heuristics for the Single Source Capacitated Plant Location Problem," *INFOR*, vol. 37, no. 3, pp. 194-225.

Deng, Y. and Shen, S. 2015, "Decomposition approaches for optimizing integrated job allocation and scheduling with a joint chance constraint," online paper: http://www.optimization-online.org/DB_HTML/2014/02/4238.html.

Drezner, Z. 1995, "Dynamic facility location: the progressive *p*-median problem," *Location Science*, vol. 3, no. 1, pp. 1-7.

Drezner, Z. & Wesolowsky, G.O. 1991, "Facility location when demand is time dependent," *Naval Research Logistics*, vol .38, pp.763-777.

Fahimnia, B., Luong, L., & Marian, R. 2012, "Genetic algorithm optimisation of an integrated aggregate production-distribution plan in supply chains," *International Journal of Production* Research, vol. 50, no. 1, pp. 81-96.

Farahani, R.Z., Asgari, N., Heidari, N., Hosseininia, M., & Goh, M. 2012, "Covering problems in facility location: A review," *Computers and Industrial Engineering*, vol. 62, pp. 368-407.

Fisher, M.L. & Hochbaum, D.S 1980, "Database location in computer networks," *Journal of the ACM*, vol. 27, no. 4, pp.718-735.

Fleischmann, B., Ferber, S. & Henrich, P. 2006, "Strategic planning of BMW's global production network," *Interfaces*, vol. 36, no. 3, pp. 194-208.

Gendreau, M., Laporte, G., & Semet, F. 2001, "A dynamic model and parallel tabu search heuristic for real-time ambulance relocation," *Parallel Computing*, vol. 27 pp. 1641-1653.

Geoffrion, A. M. & Graves, G. W. 1974, "Multicommodity distribution system design by benders decomposition," *Management Science*, vol. 20, no. 5, pp. 822-844.

Georgiadis, P. & Athanasiou, E. 2013, "Flexible long-term capacity planning in closed-loop supply chains with remanufacturing," *European Journal of Operational Research*, vol. 225, no. 1, pp. 44-58.

Ginsberg, V., Pestieau, P., & Thisse, J.F. 1987, "A spatial model of party competition with electorial and ideological objectives" in *Spatial Analysis and Location-Allocation Models*, eds A. Ghosh & G. Rushton, New York: Van Nostrand Reinhold Company, pp. 101-117.

Golden, B., S. Raghavan & E. Wasil, (eds.) 2008, The Vehicle Routing Problem: Latest Advances and New Challenges, Springer, New York.

Guignard, M. & Spielberg, K. 1979, "A direct dual method for the mixed plant location problem with some side constraints," *Mathematical Programming*, vol. 17, no. 1, pp. 198-228.

Hinojosa Y., Puerto J., & Saldanha da Gama, F. 2014, "A two-stage stochastic transportation problem with fixed handling costs and a priori selection of the distribution channels," *TOP*, vol. 22, no., 3, pp. 1123-1147.

Holmberg, K., Rönnqvist, M. & Yuan, D. 1999, "An exact algorithm for the capacitated facility location problems with single sourcing," *European Journal of Operational Research,* vol. 113, no. 3, pp. 544-559.

Hugo, A. & Pistikopoulos, E.N. 2005, "Environmentally conscious long-range planning and design of supply chain networks," *Journal of Cleaner Production,* vol. 13, no. 15, pp. 1471-1491.

Jacobsen, S.K. 1983, "Heuristics for the capacitated plant location model," *European Journal of Operational Research,* vol. 12, no. 3, pp. 253-261.

Jena, S.D., Cordeau, J.F., & Gendron, B. 2013, "Dynamic facility location with generalized modular capacities," Technical Report, CIRRELT-2013-18, Université de Montrél, Montréal, Québec, Canada.

Kellerer, H. U. Pferschy & D. Pisinger, 2010, Knapsack Problems, Springer Verlag, Heidelberg, Germany.

Kingman, J.F.C. 1962a, "Some Inequalities for the GI/G/1 queue," *Biometrika*, vol. 49, pp. 315-324.

Kingman, J.F.C. 1962b, "On queues in heavy traffic," *Journal of the Royal Statistical Society: Series B*, vol. 24, pp. 383-392.

Klincewicz, J.G. & Luss, H. 1986, "A lagrangian relaxation heuristic for capacitated facility location with single-source constraints," *Journal of the Operational Research Society,* vol. 37, no. 5, pp. 495-500.

Klincewicz, J.G., Luss, H. & Yu, C. 1988, "A large-scale multilocation capacity planning model," *European Journal of Operational Research*, vol. 34, no. 2, pp. 178-190.

Klose, A. & Görtz, S. 2007, "A branch-and-price algorithm for the capacitated facility location problem," *European Journal of Operational Research,* vol. 179, no. 3, pp. 1109-1125.

Laporte, G. 1992, "The vehicle routing problem: An overview of exact and approximate algorithms," *European Journal of Operational Research*, vol. 59, no. 3, 345-358.

Laporte, G. & Louveaux, F.V. 1993, "The integer L-shaped method for stochastic integer programs with complete recourse," *Operations Research Letters*, vol. 13, no. 3, pp. 133-142.

Laporte, G., Louveaux, F. V. & van Hamme, L. 1994, "Exact solution to a location problem with stochastic demands," *Transportation Science*, vol. 28, no. 2, pp. 95-103.

Larson, R.C., & Odoni, A.R. 2007, *Urban Operations Research*, 2nd edn, Dynamic Ideas, Belmont, Massachusetts.

Lindley, D.V. 1952, "The theory of queues with a single server," *Mathematical Proceseedings of the Cambridge Philosophical* Society, vol. 48, pp. 277-289.

Louveaux, F.V. 1986, "Discrete stochastic location models," *Annals of Operations Research*, vol. 6, pp.23-34.

Louveaux, F.V. & Peeters, D. 1992, "A dual-based procedure for stochastic facility location," *Operations Research*, vol. 40, no. 3., pp. 564-573.

Luss, H. 1982, "Operations research and capacity expansion problems: a survey," *Operations research*, vol. 30, no. 5, pp. 907-947.

Maass, K.L. & Daskin, M.S. 2017, "A cyclic allocation model for the inventory-modulated capacitated location problem," *Information Systems and Operational Research*, Advance online publication.

Maass, K.L., Daskin, M.S. & Shen, S. 2016, "Mitigating hard capacity constraints with inventory in facility location modeling," *IIE Transactions*, vol. 48, no. 2, pp. 120-133.

Marshall, K.T. 1968, "Some inequalities in queueing," *Operations Research*, vol. 16, pp. 651-665.

Marchal, W.G. 1978, "Technical note- some simpler bounds on the mean queueing time," *Operations Research*, vol. 26, no. 6, pp. 1083-1088.

Magnanti, T.L. & Wong, R.T. 1981, "Accelerating Benders decomposition: algorithmic enhancement and model selection criteria," *Operations Research*, vol. 29, no. 3, pp. 464-484.

Magnanti, T.L. & Wong, R.T. 1990, "Decomposition methods for facility location problems" in *Discrete Location Theory*, eds. P.B. Mirchandani & R.L. Francis, Wiley-Interscience, pp. 209-262.

Mak, H. & Shen, Z. M. 2012, "Risk diversification and risk pooling in supply chain design," *IIE Transactions*, vol. 44, no. 8, pp. 603-621.

Maleki, M., Majilesinasab, N., & Sepehri, M.M. 2014, "Two new models for redeployment of ambulances," *Computers and Industrial Engineering*, vol. 79, pp. 271-284.

Maxwell, M.S., Restrepo, M., Henderson, S.G., & Topaloglu, H. 2010, "Approximate dynamic programming for ambulance redeployment," *INFORMS Journal on Computing*, vol. 22, no. 2, pp. 266-281.

Melo, M.T., Nickel, S. & Saldanha da Gama, F. 2006, "Dynamic multi-commodity capacitated facility location: a mathematical modeling framework for strategic supply chain planning," *Computers and Operations Research*, vol. 33, no. 1, pp. 181-208.

Nauss, R.M. 1978, "An improved algorithm for the capacitated facility location problem," *The Journal of the Operational Research Society,* vol. 29, no. 12, pp. 1195-1201.

Neebe, A.W. & Rao, M.R. 1983, "An algorithm for the fixed charge assignment of users to sources problem," *Journal of the Operational Research Society,* vol. 34, pp. 1107-1113.

Ozsen, L., Daskin, M.S. & Coullard, C.R. 2009, "Facility location modeling and inventory management with multisourcing," *Transportation Science,* vol. 43, no. 4, pp. 455-472.

Peidro, D., Mula, J., & del Mar Botella, M. 2010, "A fuzzy linear programming based approach for tactical supply chain planning in an uncertainty environment," *European Journal of Operational* Research, vol. 205, pp. 65-80.

Pirkul, H. 1987, "Efficient algorithms for the capacitated concentrator location problem," *Computers and Operations Research,* vol. 14, no. 3, pp. 197-208.

Rajapopalan, H.K., Saydam, C., & Xiao, J. 2008, "A multiperiod set covering location model for dynamic redeployment of ambulances," *Computers and Operations Research,* vol. 35, pp. 814-826.

Reggia, J., Nau, D., & Wang, P. 1983, "Diagnostic expert systems based on a set covering model," *International Journal of Man-Machine Studies,* vol. 19, pp. 437-460.

Repede, J.F. & Bernardo, J.J. 1994, "Developing and validating a decision support system for locating emergency medical vehicles in Louisville, Kentucky," *European Journal of Operational Research,* vol. 75, no. 3, pp. 567–581.

Rönnqvist, M., Tragantalerngsak, S. & Holt, J. 1999, "A repeated matching heuristic for the single-source capacitated facility location problem," *European Journal of Operational Research,* vol. 116, no. 1, pp. 51-68.

Saveh-Shemshaki, F., Shechter, S., Tang, P. & Isaac-Renton, J. 2012, "Setting sites for faster results: Optimizing locations and capacities of new tuberculosis testing laboratories," *IIE Transactions on Healthcare Systems Engineering,* vol. 2, no. 4, pp. 248-258.

Schmid, V., & Doerner, K.F. 2010, "Ambulance location and relocation problems with time-dependent travel times," *European Journal of Operational Research,* vol. 207, pp. 1293-1303.

Scott, A.J. 1970 "Dynamic location-allocation systems: Some basic planning strategies," *Environment and Planning,* vol. 3, pp. 73-82.

Snyder, L.V., Scaparra. M.P., Daskin, M.S., & Church, R.L. 2006. "Planning for disruptions in supply chain networks," chapter 9, *INFORMS Tutorials in Operations Research,* (M.P. Johnson, B. Norman B, and N. Secomandi editors). INFORMS, pp. 234-257.

Sridharan, R. 1993, "A lagrangian heuristic for the capacitated plant location problem with single source constraints," *European Journal of Operational Research,* vol. 66, no. 3, pp. 305-312.

Torabi, S.A. & Moghaddam, M. 2012, "Multi-site integrated production distribution planning with trans-shipment: a fuzzy goal programming approach," *International Journal of Production Research*, vol. 50, no. 6, pp. 1726-1748.

Torres-Soto, J.E. & Üster, H. 2011, "Dynamic-demand capacitated facility location problems with and without relocation," *International Journal of Production Research*, vol. 49, no. 13, pp. 3979-4005.

Toth, P. & Vigo, D. 2002, "Models, relaxations and exact approaches for the capacitated vehicle routing problem," *Discrete Applied Mathematics*, vol. 123, no. 1, 487-512.

Trivedi, V. M. 1981, "A mixed-integer programming model for nursing service budgeting," *Operations Research*, vol. 29, pp. 1019-1034.

Van Roy, T.J. 1986, "A cross decomposition algorithm for capacitated facility location," *Operations research*, vol. 34, no. 1, pp. 145-163.

Varthanan, P.A., Murugan, N., & Kumar, G.M. 2012, "A simulation based heuristic discrete particle swarm alogirthm for generating integrated production-distribution plan." *Applied Soft Computing*, vol. 12, pp. 3034-3050.

Verter, V. 2011, "Uncapacitated and capacitated facility location problems" in *Foundations of Location Analysis*, eds. H.A. Eiselt & V. Marianov, Springer, pp. 25-37.

Verter, V. & Dincer, M.C. 1995, "Facility location and capacity acquisition: an integrated approach," *Naval Research Logistics*, vol. 42, no. 8, pp. 1141-1160.

Weisstein, E.W. 2016, "Great Circle" From *MathWorld*- A Wolfram Web Resource. http://mathworld.wolfram.com/GreatCircle.html (Accessed on 30 Mar 2016).

Wentges, P. 1996, "Accelerating Benders decomposition for the capacitated facility location problem," *Mathematical Methods of Operations Research*, vol. 44, no. 2, pp. 267-290.

Wesolowsky G.O. & Truscott, W.G. 1975, "The multiperiod location-allocation problem with relocation of facilities," *Management Science*, vol. 22, no. 1, pp. 57-65.

Wong, D., Moturi, .S, Angkachatchai, V., Muelloer, R., DeSantis, G., van den Boom, D., & Ehrich, M. 2013, "Optimizing blood collection, transport, and storage conditions for cell free DNA increases access to prenatal testing." *Clinical Biochemistry*, vol.46, pp. 1099-1104.