Probabilistic Envelope Constrained Multiperiod Stochastic Emergency Medical Services Location Model and Decomposition Scheme

Chun Peng
HEC Montréal & GERAD, Montréal, Canada H3T 2A7
Beijing Institute of Technology, Beijing, China 100081
chun.peng@hec.ca

Erick Delage
HEC Montréal & GERAD, Montréal, Canada H3T 2A7
erick.delage@hec.ca

Jinlin Li
Beijing Institute of Technology, Beijing, China 100081
jinlinli@bit.edu.cn

This paper considers a multiperiod Emergency Medical Services (EMS) location problem and introduces two two-stage stochastic programming formulations that account for uncertainty about emergency demand. While the first model considers both a constraint on the probability of covering the realized emergency demand and minimizing the expected cost of doing so, the second one employs probabilistic envelope constraints which allow us to control the degradation of coverage under the more severe scenarios. These models give rise to large mixed-integer programs, which can be tackled directly or using a conservative approximation scheme. For the former, we implement the Branch-and-Benders-Cut method, which improves significantly the solution time when compared to a state-of-the art Branch-and-Bound algorithm proposed in the recent literature and to using the CPLEX solver. Finally, a practical study is conducted using historical data from Northern Ireland Ambulance Service and sheds some light on optimal EMS location configuration for this region and on necessary trade-offs that must be made between emergency demand coverage and expected cost. These insights are confirmed through an out-of-sample performance analysis.

Key words: Facility location, two-stage stochastic programming, chance-constrained programming, probabilistic envelope constraint, Emergency Medical Services location, Branch-and-Benders-Cut

1. Introduction

Emergency Medical Services location plays an important role in providing adequate and high-quality medical services for the public to answer as many emergency requests as possible under dynamic and complex conditions. The locations of ambulance bases and the emergency vehicle themselves are especially important considering their effect on response time and level of coverage. Recently, EMS location has attracted substantial amount of attention in the literature through the form of ambulance location and relocation problems and healthcare facility location problems. In this paper, we study a multiperiod EMS location problem that focuses on decisions, including ambulance base locations and fleet size, and the assignment of vehicles to emergency requests and their relocation throughout the day.

Most of the literature on EMS location currently formulates the problem as a deterministic coverage location problem (see recent reviews by Aringhieri et al. (2017), Ahmadi-Javid et al. (2017), Bélanger et al. (2018)). Specifically, these models enforce that when an emergency request occurs, it can be covered by an ambulance within a certain coverage distance. However, such models disregard the fact that even within a 24-hour cycle the rate of emergency requests, travel time, and vehicles availability can vary drastically. Relocation of emergency vehicle on the network over multiperiod can therefore improve the performance of EMS by preventing areas from becoming unprotected, as well as improving the utilization of available resources. The models presented in this paper incorporate time-dependent parameters in order to identify time-dependent location strategies that make the best trade-off between the flexibility of EMS and the financial commitment.

Another key feature of the environment in which EMS operate is the pervasive presence of uncertainty. This is especially the case for factors like the amount of emergency requests at any given time, traffic conditions, operational cost, etc. Disregarding this uncertainty in a EMS location model is likely to lead to surprises regarding operational costs and might even mean that targets for demand coverage won't be met. In this regard, our proposed models will follow in the steps of Beraldi et al. (2004) and more recently Beraldi and Bruni (2009) who consider imposing a constraint (a.k.a. chance constraint as introduced in Charnes and Cooper (1959)) on the probability of covering the demand given a joint probability distribution on the amount of requests in each location. This allows EMS managers to evaluate the additional cost related to improving the reliability level of the network. This model however has the deficiency that its optimal operating strategy recommends covering none of the emergency requests for the most severe scenarios. We extend the work of Beraldi and Bruni (2009) firstly and foremost by considering probabilistic envelope constraints (see Xu et al. (2012)), instead of single chance constraint, which allow EMS managers to control how much coverage they are willing to offer as a function of how unlikely the scenario is. Other contributions are more of a numerical and empirical nature and will be discussed next.

To summarize, this paper addresses the modeling and resolution of a multiperiod EMS location problem that trades-off between expected cost and level of reliability while taking into account the decisions on ambulance base location, fleet size, assignment and relocation of vehicles in a multiperiod environment. More specifically, the contribution can be described as follows:

- We extend the static model presented in Beraldi and Bruni (2009) by considering a multiperiod environment in which vehicles can be relocated in order to better account for time-dependencies of emergency requests, operational costs, and availability of emergency vehicles, and by imposing a probabilistic envelope constraint (PEC) on the system-wide coverage level that is achieved under all realizable scenarios, instead of only the more likely ones that are addressed by chance constraint (CC) paradigm. This allows us to better characterize the trade-offs one needs to make between the cost of responding to each potential request and its likelihood of occurring.

- We develop an efficient exact algorithm for the chance-constrained version of our model that significantly outperforms the Branch-and-Bound (BB) algorithm used in Beraldi and Bruni (2009) and the CPLEX solver. The algorithm is based on a Branch-and-Benders-Cut (B&BC) scheme and is accelerated using valid inequalities and optimality cuts. This allows us to solve instances of the chance constrained stochastic program with up to 100 ambulance base locations, 150 demand locations, 6 time periods, and 200 scenarios in less than 6 minutes.
- We further propose an efficient approximation scheme for both CC and PEC stochastic programs that returns a solution guaranteed to be feasible. We observe empirically that the conservative solution of this approximation are nearly optimal for the problem instances in this study. This allows us to conclude that the PEC stochastic program can be used to derive useful insights in problems of realistic size.
- We perform a case study based on a historical dataset obtained from the Northern Ireland Ambulance Service Health & Social Care Trust (NIASHSCT) which allows us to demonstrate how our proposed stochastic programming models can be applied in a real data-driven environment. This also allows us to provide new insights on how probabilistic envelopes can be designed and on the trade-offs that exist between coverage, reliability, and expected total cost. Finally, to the best of our knowledge, we are the first to perform an out-of-sample evaluation of optimal strategies obtained for EMS location problems. This drives us to propose a new online procedure that identifies ambulance assignments for scenarios that are not observed in the in-sample dataset and confirm empirically its statistical consistency.

The rest of this paper is structured as follows. Section 2 presents a brief overview of literature. Section 3 describes the deterministic, CC and PEC version of our model. Section 4 introduces the accelerated B&BC scheme for the CC stochastic program. Section 5 presents how to adapt this B&BC scheme to the PEC stochastic program and Section 6 describes our proposed conservative approximation scheme. The numerical performance of our algorithms and case study of NIASHSCT are presented in Section 7 and 8 respectively. Finally, we give concluding remarks in Section 9.

2. Literature Review

The earliest versions of EMS ambulance location models can be considered as extensions of classical covering location models such as set covering location problem (SCLP) in Toregas et al. (1971), maximal covering location problem (MCLP) in Church and Velle (1974), double standard model (DSM) in Gendreau et al. (1997) and their variants. These early models are static and deterministic ones with decisions that model both the location of bases and fleet sizes while the objective typically aims at maximizing the coverage or minimizing the number of facilities (see Brotcorne et al. (2003)). In particular, one of the first SCLP model introduced by Toregas et al. (1971) minimizes the number of ambulances needed to cover all demand points, while ignoring other relevant aspects of these problems. Another classical ambulance location model is called MCLP and introduced by Church and Velle (1974). This model maximizes the proportion of the population that can be covered with a limited number of ambulances. Extensions of the SCLP and MCLP models have also appeared more recently in Başar et al. (2011), Schneeberger et al. (2016) and Paul et al. (2017). Recent advances for covering location problems can also be found in Marianov (2017).

In contrast to static location models, dynamic location models consider long term effects of ambulance deployment decisions. To the best of our knowledge, the first dynamic ambulance location model is proposed by Gendreau et al. (2001), where the authors account for double coverage standard (as in DSM) and penalize the frequent relocation of vehicles. More recently, Moeini et al. (2015) propose a modification to the earlier model that appears to provide better coverage for emergency request by using an adjustment parameter that accounts for its fluctuation. Similarly, Schmid and Doerner (2010) extend the work of Gendreau et al. (1997) and Gendreau et al. (2001) by accounting for capacity over a certain horizon of time. Degel et al. (2015) also formulate a multi-objective and multiperiod covering location model by incorporating time-dependent parameters and decisions. It is worth mentioning that a number of studies have also made use of exact or approximate dynamic programming (see Maxwell et al. (2010), Schmid (2012)) yet their application is limited due to the well-known curse of dimensionality. We finally refer the readers to relevant literature in the context of carsharing systems (see Lu et al. (2017), He et al. (2017, 2018)) where vehicle fleet composition and dynamics are similarly optimized although the covering of demand typically plays a less important role than in EMS.

The earliest version of stochastic programming model can be traced back to Daskin (1983) whose seminal work takes the form of the Maximal Expected Covering Location Problem (MEXCLP). In MEXCLP, the author assumes that each emergency vehicle is busy with a certain probability and that its availability is independent from other vehicles. He then formulates a model in which the expected coverage is maximized. van den Berg and Aardal (2015) later extend MEXCLP by formulating a time-dependent model with start-up and relocation costs, while Schmid and Doerner (2010) propose a multiperiod double coverage model with time-dependent travel time. At the same period, Ansari et al. (2015) account for uncertain travel times and preferences in the assignment of ambulances from different stations to demand location while Maleki et al. (2014) adapt the maximal expected coverage relocation problem in order to optimize the redeployment of ambulances.

Another important family of stochastic programming formulations appears to have been introduced for the first time by Ball and Lin (1993) in an EMS application and employs so-called "chance constraints" to control the probability that a vehicle is unable to respond to a demand call within a certain amount of time. Beraldi et al. (2004) develop a stochastic programming model with joint probabilistic constraints that aims at ensuring a reliable service level for emergency demand under a known distribution of random emergency demand while minimizing the total cost. They employ a so-called p-efficient points of the joint probability distribution to reformulate chance constraints using a MILP representation. Similarly, Zhang and Li (2015) also mitigate emergency demand uncertainty by proposing a chance constrained model that can be formulated as a second-order cone program when only the mean and covariance information are known.

A number of recent contributions have also formulated the EMS location problem in the form of a two-stage model where location and fleet size decisions are considered in the first stage while the assignment decisions are made in a second stage. Naoum-Sawaya and Elhedhli (2013) use an objective function that trades-off between relocation costs and expected coverage, using a penalty cost for unserved demand. The model in Boujemaa et al. (2017) further introduces the notion of a two-tiered system with two types of vehicles that can be employed on the network. Nickel et al. (2016) handle the trade-off between cost and coverage by imposing a lower bound on the expected coverage. An important issue associated to all of these approaches is that by using expected coverage as a measure of the risk of unmet demand, the decision model does not provide any control over the likelihood that the coverage goes below certain critical levels. This issue is naturally addressed in chance constrained formulation typically at the price of computationally inefficiency. It is worth mentioning that Noyan (2010) does propose the use of "integrated chance constraints" or second-order stochastic dominance constraints, which also circumvent this issue. One needs to know however that these risk measures are much less interpretable than chance constraints and still give rise to significant computational difficulties when considering two-stage formulations.

Our work is heavily inspired by the work of Beraldi and Bruni (2009) who consider a two-stage stochastic programming formulation that minimizes the expected cost of responding to emergency requests while imposing a chance constraint on coverage of emergency request demand. These authors design a BB algorithm to solve the resulting mixed integer linear program (MILP). Beyond replacing their model's fixed coverage cost with a linear one, we extend this work in four ways. Firstly, we extend this approach to a multiperiod setting in which ambulances can be relocated to other stations (as proposed in Schmid and Doerner (2010) and van den Berg and Aardal (2015) for covering models). Secondly, we consider probabilistic envelope constraints that allow us to also control coverage in scenarios that are considered more extreme while these are ignored in Beraldi and Bruni (2009)'s work thus causing an optimistic bias in terms of optimal expected cost. Thirdly, we propose an exact solution scheme that significantly improves the numerical performance of their BB algorithm. Finally, while Beraldi and Bruni (2009) conduct numerical experiments using test problems for the Two-Stage Capacitated Facility Location problem, we evaluate our models in a more realistic case study that uses a dataset covering historical emergency requests from Northern Ireland and draw a number of interesting practical insights.

Finally, it is worth noting that the models we propose in the following section will combine ideas from classical mathematical formulation such as the capacitated fixed-charge location problem, the p-median problem, and transportation problem and integrates them in a multiperiod setting. These types of models are in fact also commonly used in the literature that covers dynamic facility location problems where the opening, closing and relocation of facilities is optimized while accounting for uncertainty about future demand. One can for instance find many models in the recent literature: e.g. dynamic deterministic models, multiperiod stochastic models, dynamic programming based models, etc. For conciseness, we refer interested readers to the surveys by Owen and Daskin (1998), Arabani and Farahani (2012), and Nickel and da Gama (2015) for some detailed formulations. To the best of our knowledge, none of these formulations have employed probabilistic envelope constraints although we believe that they could certainly benefit from it.

3. Model Formulation

In this section, we firstly present in Section 3.1 the deterministic version of multiperiod EMS location model and related notation. Then, we incorporate emergency demand uncertainty and propose in Section 3.2 two probabilistic constrained stochastic programs.

3.1. Deterministic Model and Notation

We consider a multiperiod EMS location problem that addresses ambulance location, fleet size, allocation and relocation decisions, and where emergency requests occur randomly. Specifically, we partition the whole region into a set of zones, based on geographic division. We assume that emergency requests originate randomly from any of these zones, and consider a certain horizon divided into T periods. Given a set of potential base locations, the decision involves the identification of which subset of these locations will be operated at different times of the day and the number of ambulances assigned to these location throughout the day. This implicitly involves decisions about how ambulances are relocated throughout the day. All notation is detailed in Table 1.

Unlike most related work discussed in Section 2 under a static context, we incorporate time-dependent parameters and decisions, such as emergency demand, fleet size, service rate, related cost parameters, some of which are also in Schmid and Doerner (2010) and van den Berg and

Table 1 Notations

Sets	
$\mathcal{I}$	the set of demand zones
$\mathcal{J}$	the set of potential ambulance base locations
$\mathcal{J}_i^z$	the set of ambulance base locations from which an ambulance can serve demand in zones $i \in \mathcal{I}$
$\mathcal{J}_m^r$	the set of ambulance base locations to which an ambulance in location $m \in \mathcal{J}$ can be relocated
Parameters	
T	the number of time periods considered
f_j^t	the fixed operation cost for the base at location j at time period t
g_j^t	the marginal operation cost of hosting each ambulance at base j during time period t
c^t	the marginal transportation cost during time period t
l_{ij}	the distance between demand location i and base location j
α^t	the relocation cost for each ambulance during time period t
P_j^t	the maximum number of ambulances that can be hosted at base j during time period t
d_i^t	the number of emergency requests in demand zone i during time period t
λ_j^t	the service rate per ambulance for base location j at time period t
β^t	the minimum service level for the EMS system during time period t
Decision Variables	
x_j^t	binary variable, $x_j^t = 1$ if the base at location j will be open during time period t
y_j^t	integer variable, the number of ambulances hosted at base location j during time period t
r_{mj}^t	integer variable, the number of ambulances planned to be relocated from location m to location j (with $j \neq m$) at time period t
z_{ij}^t	integer variable, the number of emergency requests in zone i served by an ambulance at base location j during time period t

Aardal (2015). Similarly to a number of static models proposed for this problem (namely Beraldi et al. (2004), Beraldi and Bruni (2009), Noyan (2010), Naoum-Sawaya and Elhedhli (2013) and Nickel et al. (2016)), we assume the goal of the EMS manager is to minimize the total cost for operating the EMS system, which decomposes as the cost for operating bases, vehicles, the cost for hosting and maintaining them, transportation cost, and relocation cost, while ensure a minimum coverage level. This gives rise to the following multiperiod deterministic EMS location model:

$$[\text{DM}] \quad \underset{\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},\boldsymbol{r}}{\text{minimize}} \quad \sum_{t=1}^{T}\sum_{j\in\mathcal{J}} f_j^t x_j^t + \sum_{t=1}^{T}\sum_{j\in\mathcal{J}} g_j^t y_j^t + \sum_{t=1}^{T}\sum_{j\in\mathcal{J}}\sum_{i\in\mathcal{I}} c^t l_{ij} z_{ij}^t + \sum_{t=1}^{T}\sum_{j\in\mathcal{J}}\sum_{m\in\mathcal{J}} \alpha^t r_{mj}^t \tag{1a}$$

$$\text{subject to } y_j^t \leq P_j^t x_j^t \qquad \forall j \in \mathcal{J},\, t \in \{1,\ldots,T\} \tag{1b}$$

$$x_j^t \geq x_j^{t-1} \qquad \forall j \in \mathcal{J}, t \in \{2,\ldots,T\} \tag{1c}$$

$$y_j^t + \sum_{m\in\mathcal{J}_j^r} r_{mj}^t - \sum_{m\in\mathcal{J}_j^r} r_{jm}^t = y_j^{t+1} \qquad \forall j\in\mathcal{J}, t\in\{1,\ldots,T-1\} \tag{1d}$$

$$y_j^T + \sum_{m\in\mathcal{J}_j^r} r_{mj}^T - \sum_{m\in\mathcal{J}_j^r} r_{jm}^T = y_j^1 \qquad \forall j\in\mathcal{J} \tag{1e}$$

$$\sum_{i\in\mathcal{I}} z_{ij}^t \le \lambda_j^t y_j^t \qquad \forall j\in\mathcal{J},\, t\in\{1,\ldots,T\} \tag{1f}$$

$$\sum_{j\in\mathcal{J}_i^z} z_{ij}^t \le d_i^t \qquad \forall i\in\mathcal{I},\, t\in\{1,\ldots,T\} \tag{1g}$$

$$\sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}_i^z} z_{ij}^t \ge \beta^t \sum_{i\in\mathcal{I}} d_i^t \qquad \forall t\in\{1,\ldots,T\} \tag{1h}$$

$$x_j^t \in \{0,1\}, y_j^t, r_{mj}^t, z_{ij}^t \in \mathbb{N} \qquad \forall i\in\mathcal{I}, j,m\in\mathcal{J}, t\in\{1,\ldots,T\}. \tag{1i}$$

Constraint (1b) imposes the maximum number of vehicles that can be hosted at ambulance base j during time period t depending on whether the base is open or closed. Constraint (1c) imposes that an open base remains open for the rest of the day, which is a common constraint in the multiperiod facility location literature (see constraint (7) in Schilling (1980), constraint (14) in Van Roy and Erlenkotter (1982), and constraint (11.37) in Nickel and da Gama (2015)). Other possible forms of condition of operations of facilities and their associated cost are discussed in Appendix B.1. Constraints (1d) and (1e) are balance equations, which ensure that the relocation plan of emergency vehicles is feasible over the planning horizon. In particular, $\mathcal{J}_j^r$ might be such that ambulance can only be relocated to a base within a certain radius $l_{\max}^r$. It also imposes that initial and final distribution of ambulances should be the same for consistency from one day to the other (see Remark 1 below). Constraint (1f) states that the maximum number of emergency request that can be served from base j during period t depend on the number of ambulances present and on the service rate $\lambda_j^t \in \mathbb{N}$. Constraint (1g) imposes that, in each zone i, one cannot serve more emergencies than the number of requests submitted d_i^t. Once again a natural way of designing $\mathcal{J}_i^z$ consists in only including pairs of locations that are within a certain service radius $l_{\max}^z$. Finally, constraint (1h) imposes a minimum service level of β^t for the EMS system, in other words the network should be designed and managed so that at each period of the day a certain proportion $\beta^t \in [0, 1]$ of total emergency demand be served.

In this paper, some assumptions are made in order to help with the numerical resolution efforts. We assume that relocation decisions are made before any of the emergency requests are observed, and that the relocation of an ambulance does not affect its capacity to serve emergency requests. Although these assumptions constitute a limitation of our model, we believe that they do not have an important effect on the conclusion when comparing the quality of different strategic decisions. First, when the length of each period is large enough (e.g. 4 hours), the time needed for relocation within a reasonable distance becomes relatively insignificant and the number of total requests that are served in a period becomes large thus reducing the variance of the total. Regarding the nature of relocation decision, one could argue that schedules of ambulance deployment need to be defined as early as possible in order to inform drivers once of the plan for the whole day. Alternatively, one could argue that there is actually no loss in doing so (compared to an adaptive plan) in contexts where the emergency requests are independent from one period to the next.

One might also note that it is possible with the DM formulation to model the fact that some ambulances become off-duty during some periods of the day in order to reduce total cost by assigning them to an artificial node with $g_j^t = 0$. Similarly, one could model longer relocation routes that pass through similar off-duty dummy nodes which would account for the fact that the ambulance will be traveling for a full period.

REMARK 1. Note that, with Constraint (1e), model (1) effectively emulates an infinite horizon problem in which each day has exactly the same characteristics and where we wish the policy to be the same every day. In the stochastic setting, it will imply that daily emergency requests are independent and identically distributed from one day to the other. In practice, this implicit infinite horizon model might be reasonable especially when there is a strong interest in identifying a policy that is easy to operationalize by repeating it on a daily basis. Otherwise, one can also extend the model so that the horizon covers a few days of operations and implement the proposed solution using a rolling horizon approach where only the first few periods are implemented, the model is then updated with new conditions and re-optimized.

3.2. Two Probabilistic Constrained Stochastic Programming Models

An important limitation of the DM model consists in the assumption that all parameters are exactly known at the moment of designing the network. Indeed, in practice it is difficult to predict exactly where and when the emergency requests will occur. For this reason, we now propose a model that assumes that these requests occur randomly according to some distribution $\boldsymbol{d} \sim \mathbb{Q}$. In this context, it is reasonable to assume that the assignment of ambulances will be done once the demand of each period is known. In other words, each decision $z_{ij}^t : \mathbb{R}^{|\mathcal{J}|} \to \mathbb{R}$ can be adjusted to the observed emergency demand $\boldsymbol{d}^t$. Following the ideas proposed in Beraldi and Bruni (2009), this gives rise to the following two-stage chance-constrained stochastic program (CCSP):

[CCSP]

$$\underset{\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}(\cdot),\boldsymbol{r}}{\text{minimize}}\ \mathbb{E}_{\mathbb{Q}}\left[\sum_{t=1}^{T}\sum_{j\in\mathcal{J}} f_j^t x_j^t + \sum_{t=1}^{T}\sum_{j\in\mathcal{J}} g_j^t y_j^t + \sum_{t=1}^{T}\sum_{j\in\mathcal{J}}\sum_{i\in\mathcal{I}} c^t l_{ij} z_{ij}^t(\boldsymbol{d}^t) + \sum_{t=1}^{T}\sum_{j\in\mathcal{J}}\sum_{m\in\mathcal{J}} \alpha^t r_{mj}^t\right] \tag{2a}$$

subject to (1b) − (1e)

$$\mathbb{P}_{\mathbb{Q}}\left(\sum_{i\in\mathcal{I}} z_{ij}^t(\boldsymbol{d}^t) \le \lambda_j^t y_j^t\right) = 1 \qquad \forall j\in\mathcal{J},\ t\in\{1,\dots,T\} \tag{2b}$$

$$\mathbb{P}_{\mathbb{Q}}\left(\sum_{j\in\mathcal{J}_i^z} z_{ij}^t(\boldsymbol{d}^t) \le d_i^t\right) = 1 \qquad \forall i\in\mathcal{I},\ t\in\{1,\dots,T\} \tag{2c}$$

$$\mathbb{P}_{\mathbb{Q}}\left(\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_i^z} z_{ij}^t(\boldsymbol{d}^t) \ge \beta^t \sum_{i\in\mathcal{I}} d_i^t\right) \ge 1-\eta^t \qquad \forall t\in\{1,\dots,T\} \tag{2d}$$

$$\mathbb{P}_{\mathbb{Q}}\left(z_{ij}^t(\boldsymbol{d}^t)\in\mathbb{N}\right) = 1 \qquad \forall i\in\mathcal{I}, j\in\mathcal{J}, t\in\{1,\dots,T\} \tag{2e}$$

$$x_j^t\in\{0,1\}\,;y_j^t, r_{mj}^t\in\mathbb{N} \qquad \forall i\in\mathcal{I}, j,m\in\mathcal{J}, t\in\{1,\dots,T\}, \tag{2f}$$

for some $\eta^t \in [0, 1]$, where the expected total cost of operation is minimized and where constraint (2d) is a chance constraint that controls the reliability of the coverage, i.e. the required likelihood of serving at least a proportion β^t of emergency requests according to $\mathbb{Q}$ at time t. For example, when $\eta^t = 0$ then the coverage must be satisfied with probability one, and in particular it must be satisfied under any scenario that has a strictly positive likelihood of occurrence. As argued in Beraldi and Bruni (2009), this might however lead to an over-dimensioned system hence the need for the CCSP model, which can serve as a tool to evaluate different alternatives in terms of cost-reliability trade-off. Unfortunately, the above model is necessarily overly optimistic in assessing the expected total cost of managing the system given that it completely disregards how the system will respond to requests in the more extreme scenarios that have less than η^t probability of occurring at time t. This is illustrated by the following example.

Example 1. Consider a simple network with a single ambulance base, one zone that generates emergency requests, and a single period horizon. Let d_1 be drawn according to two possible scenarios: in scenario #1 $d_1 = 50$ and has 90% chance of occurring while, in scenario #2, $d_1 = 200$ with 10% chance. If one imposes a 90% chance of covering all emergency requests (i.e. $\beta = 100\%$ and $\eta = 10\%$), then, assuming that the service rate is 50, it is clear that it becomes optimal to plan for a single ambulance to be present all day. In particular the solution would take the shape $x_1 = 1$, $y_1 = 1$, and $z_{11}(50) = 50$ while $z_{11}(200) = 0$. The optimal expected cost would be $f_1 + g_1 + 45cl_{11}$ reflecting the fact that the available ambulance is left idle at the base under the more extreme scenarios while 200 emergencies are left unserved.

To correct for this deficiency of the CCSP model, we replace constraint (2d) with a PEC as popularized in Xu et al. (2012) in order to provide some control on the level of coverage achieved under all potential scenarios of emergency requests. Similarly, this gives rise to the following two-stage probabilistic envelope constrained stochastic program (PECSP):

[PECSP]

$$\underset{\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}(\cdot),\boldsymbol{r}}{\text{minimize}}\ \mathbb{E}_{\mathbb{Q}}\left[\sum_{t=1}^{T}\sum_{j\in\mathcal{J}} f_j^t x_j^t + \sum_{t=1}^{T} g_j^t y_j^t \sum_{j\in\mathcal{J}} + \sum_{t=1}^{T}\sum_{j\in\mathcal{J}}\sum_{i\in\mathcal{I}} c^t l_{ij} z_{ij}^t(\boldsymbol{d}^t) + \sum_{t=1}^{T}\sum_{j\in\mathcal{J}}\sum_{m\in\mathcal{J}} \alpha^t r_{mj}^t\right] \tag{3a}$$

$$\text{subject to } (1b)-(1e),(2b),(2c),(2e),(2f)$$

$$\mathbb{P}_{\mathbb{Q}}\left(\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_i^z} z_{ij}^t(\boldsymbol{d}^t) \geq \beta^t(\eta)\sum_{i\in\mathcal{I}} d_i^t\right) \geq 1-\eta \qquad \forall \eta\in[0,1), t\in\{1,\ldots,T\}, \tag{3b}$$

where constraint (3b) now covers all reliability levels $\eta \in [0, 1)$ using a controlled coverage envelope function $\beta^t(\eta)$. Actually, the case that $\eta = 1$ does not appear in the constraint since a chance constraint that impose 0% reliability is always redundant. Note that without loss of generality, we will assume that $\beta^t(\eta)$ is non-decreasing (i.e. the less extreme the scenario is, the more strict the targeted coverage), otherwise one can replace this function with $\beta^{t\prime}(\eta) := \sup\{\beta^t(\eta') : \eta' \leq \eta\}$ without affecting the set of feasible solutions. It is also straightforward to see that PECSP generalizes CCSP since the latter can be obtained by using the following coverage envelope function

$$\beta^t(\eta) = \begin{cases} \bar{\beta}^t & \text{if } \eta \geq \bar{\eta}^t \\ 0 & \text{otherwise.} \end{cases}$$

Looking back at Example 1, one can see how the identified deficiency can be resolved using our PECSP model. In particular, this new model allows one to describe what level of coverage is expected for the more extreme scenario. In the case that the coverage needed for all scenarios is always above 50%, e.g. $\beta(\eta) = 50\%$ for $\eta < 10\%$, then the optimal solution would recommend a larger fleet of ambulance since $200 \times 50\% > 50$. If the reliability level is more relaxed, e.g. $\beta(\eta) = 5\%$ for $\eta < 10\%$, then the same number of vehicles would be proposed yet the expected total cost would reflect the fact that a minimum proportion of requests need to be satisfied in all scenarios.

The downside of employing PECSP is the challenge that it raises from a computational perspective. Indeed, chance constrained stochastic programs are typically considered to be computationally intractable, except in rare cases with special structure. In general, chance constraints define a non-convex feasible set and can either be conservatively approximated as in Nemirovski and Shapiro (2006) or approximated using sample average approximation which gives rise to a MILP formulation (see Luedtke and Ahmed (2008), Pagnoncelli et al. (2009), Wang et al. (2011)). The challenge is even more significant when employing our PEC, which conceptually speaking imposes an infinite continuum of chance constraints. To the best of our knowledge, to this date the most efficient solution scheme employs distributionally robust versions of these constraints (see Xu et al. (2012)). Yet, while these versions might not be very interesting from a practical point of view (as it might lead to overly conservative solutions), it is also not clear how the method proposed in Xu et al. (2012) could be used for constraint (3b) given that the interaction between the envelope function and uncertainty is not additive but rather multiplicative. One should also note that constraint (3b) can be equivalently reformulated as a first-order stochastic dominance (FSD) constraint (introduced for the first time in Dentcheva and Ruszczynski (2003)) :

$$\left(\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_i^z} z_{ij}^t(\boldsymbol{d}^t)\right)/\left(\sum_{i\in\mathcal{I}} d_i^t\right) \succeq_{(1)} \Upsilon_t \qquad \forall t\in\{1,\ldots,T\},$$

where each d_i^t and Υ_t are arbitrary random variables on a joint probability space $(\bar{\Omega},\bar{\Sigma},\bar{\mathbb{Q}})$, such that $\beta^t(\mathbb{P}_{\bar{\mathbb{Q}}}(\Upsilon_t\leq y))=y$ for all $y\in(0,1]$. In the case where all decisions are static and $\bar{\mathbb{Q}}$ is discrete, which implies that the associated $\beta^t(\cdot)$ is piecewise constant, Dentcheva and Ruszczyński (2004) show that a FSD constraint can be reformulated using finitely many chance constraints, Noyan et al. (2006) further introduce a MILP formulation and propose two convex relaxations. A new MILP formulation is proposed in Luedtke (2008) which has the property of reducing to a second-order dominance constraint when the integrality constraint is relaxed. Drapkin and Schultz (2010) propose a Benders decomposition method to solve a stochastic program with a FSD constraint on the optimal value of a second-stage problem. Finally, Hu et al. (2012) provide theoretical foundations for the convergence of sample average approximations method on second-order stochastic dominance constraints which will motivate us to replace $\mathbb{Q}$ with a discrete approximation.

In sections 4 and 5, we explain how to obtain a MILP based reformulation of both CCSP and PECSP, with a general $\beta(\cdot)$ envelope, and propose an improved B&BC method that can be employed to improve numerical efficiency. A conservative approximation scheme is also proposed for PECSP (or CCSP) in order to allow the resolution of EMS location problems of realistic sizes in Section 6.

4. Solution Scheme for Scenario-based CCSP

In this section, we present a scenario-based version of CCSP, which might be obtained from applying a sample average approximation scheme, derive a MILP reformulation of the problem, and propose different variants of a B&BC method to solve the problem efficiently.

4.1. Scenario-based CCSP

In a scenario-based approach, one assumes that the random emergency requests $\boldsymbol{d}$ are drawn from a finite set of N scenarios $\{d_\omega\}_{\omega\in\Omega}$, where for simplicity $\Omega=\{1,\ldots,N\}$, hence the distribution $\mathbb{Q}$ can be characterized using a probability vector $[p_1\, p_2\, \ldots\, p_N]$ such that $\boldsymbol{p}\geq 0$ and $\sum_{\omega\in\Omega}p_\omega=1$. For each scenario $\omega\in\Omega$, we let $d_{i\omega}^t$ denote the number of emergency requests in zone i at time period t, while $z_{ij\omega}^t$ will denote the number of emergency requests in zone i at time period t served from ambulance base j. Under these conditions, one can rewrite the CCSP as:

$$\underset{\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},\boldsymbol{r}}{\text{minimize}} \sum_{t=1}^{T}\sum_{j\in\mathcal{J}} f_j^t x_j^t + \sum_{t=1}^{T}\sum_{j\in\mathcal{J}} g_j^t y_j^t + \sum_{\omega\in\Omega}\sum_{t=1}^{T}\sum_{j\in\mathcal{J}}\sum_{i\in\mathcal{I}} p_\omega c^t l_{ij} z_{ij\omega}^t + \sum_{t=1}^{T}\sum_{j\in\mathcal{J}}\sum_{m\in\mathcal{J}} \alpha^t r_{mj}^t \tag{4a}$$

subject to (1b) − (1e), (2f)

$$\sum_{i\in\mathcal{I}} z^t_{ij\omega} \le \lambda^t_j y^t_j \qquad \forall j\in\mathcal{J}, t\in\{1,\dots,T\}, \omega\in\Omega \tag{4b}$$

$$\sum_{j\in\mathcal{J}^z_i} z^t_{ij\omega} \le d^t_{i\omega} \qquad \forall i\in\mathcal{I}, t\in\{1,\dots,T\}, \omega\in\Omega \tag{4c}$$

$$\sum_{\omega\in\Omega} p_\omega \mathbf{1}\left\{\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}^z_i} z^t_{ij\omega} \ge \beta^t \sum_{i\in\mathcal{I}} d^t_{i\omega}\right\} \ge 1-\eta^t \quad \forall t\in\{1,\dots,T\} \tag{4d}$$

$$z^t_{ij\omega}\in\mathbb{N} \qquad \forall i\in\mathcal{I}, j\in\mathcal{J}, t\in\{1,\dots,T\}, \omega\in\Omega\,, \tag{4e}$$

where $\mathbf{1}\{\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}^z_i} z^t_{ij\omega} \ge \beta^t \sum_{i\in\mathcal{I}} d^t_{i\omega}\}$ is an indicator function that returns one if the required coverage is achieved at time period t under scenario ω, and otherwise returns zero.

Following an idea that dates back to Ruszczyński (2002), we next introduce a set of binary variables $\boldsymbol{\rho}\in\{0,1\}^{N\times T}$ that will be used to assess whether the coverage frequencies imposed in constraint (4d) are satisfied. This allows us to replace constraint (4d) with:

$$\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}^z_i} z^t_{ij\omega} \ge \beta^t(1-\rho^t_\omega)\sum_{i\in\mathcal{I}} d^t_{i\omega} \qquad \forall t\in\{1,\dots,T\}, \omega\in\Omega \tag{5}$$

$$\sum_{\omega\in\Omega} p_\omega \rho^t_\omega \le \eta^t \qquad \forall t\in\{1,\dots,T\}\,. \tag{6}$$

In particular, $\rho^t_\omega = 1$ indicates that the system manager plans to violate the coverage constraint under scenario ω at time t.

Proposition 1. *The CCSP problem is equivalent to the following mixed integer linear program in terms of optimal value and set of optimal solutions for* $\boldsymbol{x}$ *and* $\boldsymbol{y}$*:*

$$\text{[CCSP2]}\ \underset{\boldsymbol{x},\boldsymbol{y},\boldsymbol{r},\boldsymbol{z},\boldsymbol{\rho}}{\text{minimize}} \sum_{t=1}^T\sum_{j\in\mathcal{J}} f^t_j x^t_j + \sum_{t=1}^T\sum_{j\in\mathcal{J}} g^t_j y^t_j + \sum_{\omega\in\Omega}\sum_{t=1}^T\sum_{j\in\mathcal{J}}\sum_{i\in\mathcal{I}} p_\omega c^t l_{ij} z^t_{ij\omega} + \sum_{t=1}^T\sum_{j\in\mathcal{J}}\sum_{m\in\mathcal{J}} \alpha^t r^t_{mj} \tag{7a}$$

subject to (1b) − (1e), (4b), (4c), (6)

$$\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}^z_i} z^t_{ij\omega} \ge (1-\rho^t_\omega)\left\lceil \beta^t \sum_{i\in\mathcal{I}} d^t_{i\omega}\right\rceil \qquad \forall t\in\{1,\dots,T\}, \omega\in\Omega \tag{7b}$$

$$z^t_{ij\omega}\ge 0 \qquad \forall i\in\mathcal{I}, j\in\mathcal{J}, t\in\{1,\dots,T\},\ \omega\in\Omega \tag{7c}$$

$$r^t_{mj}\ge 0 \qquad \forall j,m\in\mathcal{J}, t\in\{1,\dots,T\} \tag{7d}$$

$$x^t_j\in\{0,1\}, y^t_j\in\mathbb{N}, \rho^t_\omega\in\{0,1\} \qquad \forall j\in\mathcal{J}, t\in\{1,\dots,T\}, \omega\in\Omega\,. \tag{7e}$$

Moreover, an optimal solution (x^*, y^*, r^*, z^*) *for CCSP can be obtained by solving a set of MILPs.*

Note that in CCSP2, we were able to relax the integrality constraint on the assignment variables $\boldsymbol{z}$ and relocation variables $\boldsymbol{r}$ without affecting the quality of the solutions obtained for $\boldsymbol{x}$ and $\boldsymbol{y}$

and optimal value. This is due to the fact that when first stage decisions $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{\rho}$ are fixed, the convex hull of the joint feasible set for $\boldsymbol{z}$ and $\boldsymbol{r}$ can be obtained simply by employing a linear relaxation (modulo the tightening of constraint (5) using a simple rounding scheme). The detailed proof can be seen in Appendix A.1.

4.2. Enhanced Branch-and-Benders-Cut Method

While stochastic programming models are generally known to be computationally challenging, a method known as the L-shaped method or Benders decomposition (BD), introduced by Benders (1962) has achieved good numerical performance in a number of applications including facility location and transportation (see for example in Martins de Sá et al. (2015), Dalal and Üster (2018), and Rahmaniani et al. (2017) for recent reviews of the topic). The main idea of BD is that the MILP that emerges in a stochastic program might decompose into a pure integer master problem (MP) and a number of smaller linear sub-problems (SPs). If so, the two sets of problems can be solved iteratively, introducing a group of additional constraints in the MP, known as Benders cuts. This is repeated until the lower bound obtained from MP reaches the upper bound that can be computed based on the solutions of the SPs.

There are a number of studies that have recently developed Branch-and-Cut (BC) methods to solve chance constrained programs (CCPs). First, the BC schemes in Song et al. (2014) and Song and Luedtke (2013) are designed for single-stage binary problems (e.g. binary bin packing and network design). More recently, Luedtke (2014) does consider mixed type of decisions yet cannot be employed in two-stage stochastic programs with costly recourse decisions, as is the case in our model. Interestingly, Liu et al. (2016) propose a BC method with two types of optimality cuts for solving two-stage CCPs of the same form as ours. While the first type resembles closely the cuts that we will propose in Section 4.2.4, some preliminary study revealed to us that no significant improvement could be obtained from the second type. Furthermore, in both case the question is still open regarding how to extend these families of cuts to the PECSP model.

In this section, we describe an implementation of the B&BC method in order to improve numerical efficiency by integrating Benders cuts directly in the BB algorithm that solves the MP. To simplify presentation we make the assumption that ambulances can be relocated and serve any location on the territory.

Assumption 1. *The graph of all feasible relocation and emergency assignments is complete, i.e.* $\mathcal{J}_i^z = \mathcal{J}$ *and* $\mathcal{J}_j^r = \mathcal{J} \setminus j$ *for all* $i \in \mathcal{I}$ *and* $j \in \mathcal{J}$.

We start in Section 4.2.1 by describing the BD scheme for CCSP2 under Assumption 1 after introducing redundant constraints that ensure that only optimality cuts will be returned by the SPs. An extension of the methods to situations where Assumption 1 does not hold is presented in Appendix B.5. We then summarize the B&BC scheme in Section 4.2.2. We finally discuss in sections 4.2.3 and 4.2.4 variants of the algorithm that involve a set of valid inequalities that can be used to tighten the MP and a way of tightening the optimality cuts returned from the SPs.

4.2.1. Benders Decomposition A traditional application of BD aims at separating the complicating integer variables $(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\rho})$ from the non-complicating variables $\boldsymbol{r}$ and $\boldsymbol{z}$ in order to accelerate the resolution of CCSP2. In particular, we consider the following reformulation of CCSP2:

$$\text{[CCSP2']} \quad \underset{\boldsymbol{x},\boldsymbol{y},\boldsymbol{\rho},\theta^r,\boldsymbol{\theta}^z}{\text{minimize}} \sum_{t=1}^{T}\sum_{j\in\mathcal{J}} f_j^t x_j^t + \sum_{t=1}^{T}\sum_{j\in\mathcal{J}} g_j^t y_j^t + \sum_{\omega\in\Omega}\sum_{t=1}^{T}\theta_{t\omega}^z + \theta^r \tag{8a}$$

$$\text{subject to (1b), (1c), (6), (7e)}$$

$$\theta^r \ge h^r(y) \tag{8b}$$

$$\theta_{t\omega}^z \ge h_{\omega,t}^z(y,\rho) \qquad \forall t\in\{1,\dots,T\}, \omega\in\Omega \tag{8c}$$

$$\sum_{j\in\mathcal{J}} y_j^t = \sum_{j\in\mathcal{J}} y_j^{t+1} \qquad \forall t\in\{1,\dots,T-1\} \tag{8d}$$

$$\sum_{j\in\mathcal{J}} \lambda_j^t y_j^t \ge \lceil \beta^t \sum_{i\in\mathcal{I}} d_{i\omega}^t \rceil (1-\rho_\omega^t) \qquad \forall t\in\{1,\dots,T\}, \omega\in\Omega \tag{8e}$$

where

$$\text{[CCSP-SP}^r\text{]} \quad h^r(\boldsymbol{y}) := \min_{\boldsymbol{r}\ge 0} \sum_{t=1}^{T}\sum_{m\in\mathcal{J}}\sum_{j\in\mathcal{J}} \alpha^t r_{mj}^t \tag{9a}$$

$$\text{subject to } y_j^t + \sum_{m\in\mathcal{J}_j^r} r_{mj}^t - \sum_{m\in\mathcal{J}_j^r} r_{jm}^t = y_j^{t+1} \qquad \forall j\in\mathcal{J}, t\in\{1,\dots,T-1\} \tag{9b}$$

$$y_j^T + \sum_{m\in\mathcal{J}_j^r} r_{mj}^T - \sum_{m\in\mathcal{J}_j^r} r_{jm}^T = y_j^1 \qquad \forall j\in\mathcal{J} \tag{9c}$$

and where

$$\text{[CCSP-SP}_{\omega,t}^z\text{]} \quad h_{\omega,t}^z(\boldsymbol{y},\boldsymbol{\rho}) := \min_{\boldsymbol{z}_\omega^t\ge 0} \sum_{j\in\mathcal{J}}\sum_{i\in\mathcal{I}} p_\omega c^t l_{ij} z_{ij\omega}^t \tag{10a}$$

$$\text{subject to } \sum_{i\in\mathcal{I}} z_{ij\omega}^t \le \lambda_j^t y_j^t \qquad \forall j\in\mathcal{J} \tag{10b}$$

$$\sum_{j\in\mathcal{J}_i^z} z_{ij\omega}^t \le d_{i\omega}^t \qquad \forall i\in\mathcal{I} \tag{10c}$$

$$\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_i^z} z_{ij\omega}^t \ge (1-\rho_\omega^t)\lceil \beta^t \sum_{i\in\mathcal{I}} d_{i\omega}^t \rceil \tag{10d}$$

Proposition 2. *CCSP2' is equivalent to CCSP2. Moreover, under the Assumption 1, given any solutions triplet* $(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\rho})$ *that satisfy* (8d) *and* (8e)*, problems CCSP-SP*^{r} *and CCSP-SP*$_{\omega,t}^z$ *are always feasible and bounded.*

In Problem CCSP2', we also include the redundant constraints (8d) and (8e) in order to ensure that $h^r(\boldsymbol{y})$ and $h^z_{\omega,t}(\boldsymbol{y},\boldsymbol{\rho})$ can be considered finite valued. A detailed proof of Proposition 2 can be found in Appendix A.2.

In order to obtain a decomposition between the optimization over $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{\rho}$, and the optimization of $\boldsymbol{r}$ and $\boldsymbol{z}$, we consider approximating $h^r(\boldsymbol{y})$ and $h^z_{\omega,t}(\boldsymbol{y},\boldsymbol{\rho})$ using a subset of their supporting hyperplanes, which is also called Benders cuts that are identified at given values of MP's solutions by solving the corresponding dual problems of CCSP-SP^r and CCSP-SP$^z_{\omega,t}$ respectively. More details of the BD approach are provided in Appendix B.2. Unfortunately, in practice this implementation can be excessively slow to converge due to the fact that each iteration involves solving a CCSP2' that takes the form of a MILP and of a size that grows with the number of iterations. For this reason, an implementation of BD in a Branch-and-Cut framework, which integrates the addition of Benders cuts inside the BB algorithm that is used to solve CCSP2', has become popular.

4.2.2. Branch-and-Benders-Cut Method Instead of solving a MP using the BB algorithm at every iteration, the B&BC method only requires traversing once the BB tree structure. This is done by introducing the new supporting hyperplanes of $h^r(\boldsymbol{y})$ and $h^z_{\omega,t}(\boldsymbol{y},\boldsymbol{\rho})$ (i.e. Benders cuts) only when reaching nodes of the BB tree for which the linear relaxation of the MILP identifies an integer solution. In CPLEX, this can be done through the *lazyconstraint* callback routine which takes the form of *CPXcutcallbackadd*. The BB tree search then resumes until the optimality is attained or a stopping criterion is met. For completeness, the detailed procedure of the proposed B&BC method is presented in Appendix B.3.

4.2.3. Strengthened Valid Inequalities In order to improve the efficiency of the B&BC method, it can be useful to identify valid inequalities that are redundant for the MILP version of CCSP2' but end up tightening its linear relaxation. Proposition 3 introduces a set of valid inequalities that are inspired from the work of Luedtke et al. (2010).

PROPOSITION 3. *For any fixed t, let the ordered scenario set $\Omega'_t := \{\omega'_k\}_{k=1}^N$ be such that for all $k<k'$ we have that $\sum_{i\in\mathcal{I}} d^t_{i\omega'_k} \le \sum_{i\in\mathcal{I}} d^t_{i\omega'_{k'}}$. Given some $\eta^t\in[0,1)$, let $\bar{k}_t \in \{1,2,\ldots,N\}$ be such that $\sum_{k=\bar{k}_t}^N p_{\omega'_k} > \eta^t$ and $\sum_{k=\bar{k}_t+1}^{N} p_{\omega'_k} \le \eta^t$, where we consider that $\sum_{k=N+1}^{N} p_{\omega'_k} = 0$. Then for all $t\in\{1,\ldots,T\}$, constraints (6) and (8e) imply that:*

$$\sum_{j\in\mathcal{J}} \lambda^t_j y^t_j \ge \lceil \beta^t \sum_{i\in\mathcal{I}} d^t_{i\omega'_{\bar{k}_t}} \rceil . \tag{11}$$

This proposition provides a set of T valid inequalities for CCSP2'. Note that for each t, the set Ω needs to be reordered appropriately and that index $\bar{k}_t$ might depend on t. The proof of Proposition 3 is provided in Appendix A.3.

4.2.4. Strengthened Optimality Cuts Looking back at the definition of $h^z_{\omega,t}(\boldsymbol{y},\boldsymbol{\rho})$, one might notice that as long as $\rho^t_\omega \in \{0, 1\}$ the optimal value of this problem remains the same when replacing constraint (10c) with

$$\sum_{j\in\mathcal{J}^z_i} z^t_{ij\omega} \le d^t_{i\omega}(1-\rho^t_\omega) \qquad\qquad \forall i\in\mathcal{I},\, t\in\{1,\dots,T\}, \omega\in\Omega\,.$$

This is due to the fact that when $\rho^t_\omega = 0$ then the two constraints are exactly the same, while if $\rho^t_\omega = 1$, then $z^t_\omega = 0$ becomes an optimal solution of CCSP-SP$^z_{\omega,t}$. It is therefore possible to consider a modified version of CCSP-SP$^z_{\omega,t}$ for which the dual sub-problem becomes:

$$[\text{CCSP-DSP}^z_{\omega,t}{}'] \;\; h^z_{\omega,t}(\bar{\boldsymbol{y}},\bar{\boldsymbol{\rho}}) = \max_{\boldsymbol{\mu}^1_{t\omega},\boldsymbol{\mu}^2_{t\omega},\mu^3_{t\omega}} \quad -\sum_{j\in\mathcal{J}}\lambda^t_j\bar{y}^t_j\mu^1_{jt\omega} - \sum_{i\in\mathcal{I}} d^t_{i\omega}(1-\bar{\rho}^t_\omega)\mu^2_{it\omega} + \mu^3_{t\omega}(1-\bar{\rho}^t_\omega)\lceil\beta^t\sum_{i\in\mathcal{I}} d^t_{i\omega}\rceil$$
$$\text{subject to} \quad -\mu^1_{jt\omega} - \mu^2_{it\omega} + \mu^3_{t\omega} \le p_\omega c^t l_{ij} \quad \forall i\in\mathcal{I}, j\in\mathcal{J}$$
$$\boldsymbol{\mu}^1_{t\omega}\ge 0, \boldsymbol{\mu}^2_{t\omega}\ge 0, \mu^3_{t\omega}\ge 0\,.$$

Hence, this reformulated sub-problem gives rise to a supporting hyperplane that will provide a tighter constraint for the MP.

5. Solution Scheme for Scenario-based PECSP

In this section, we study the numerical resolution of PECSP in a context where $\mathbb{Q}$ is modeled as a discrete distribution. We firstly formulate the scenario-based PECSP as a MILP and then adapt the B&BC algorithm.

5.1. Scenario-based PECSP

Similarly as was done in Section 4, we start by assuming that $\mathbb{Q}$ is discrete and characterized using a probability vector $\boldsymbol{p}\in\mathbb{R}^N$ such that $\boldsymbol{p}\ge 0$ and $\sum_{\omega\in\Omega} p_\omega = 1$. While it is still possible in this context to employ $z^t_{ij\omega}$ to model the conditional assignment plan of emergency requests in zone i answered by base j at time t under scenario ω and reuse much of the formulation of CCSP2, a difficulty arises in reformulating the PEC constraint (3b). Namely, under a discrete distribution $\mathbb{Q}$, it first can be reduced to:

$$\sum_{\omega\in\Omega} p_\omega \mathbf{1}\left\{\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}^z_i} z^t_{ij\omega} \ge \beta^t(\eta)\sum_{i\in\mathcal{I}} d^t_{i\omega}\right\} \ge 1-\eta \qquad \forall \eta\in[0,1),\, t\in\{1,\dots,T\}\,. \tag{12}$$

Yet, a priori it does not seem possible to introduce binary variables that would count the number of scenarios where the minimum level of coverage is not achieved since this accounting needs to be done for each reliability level η in the continuous range [0, 1). For this reason, we make a simplifying assumption about the discrete distribution $\mathbb{Q}$, which allows to identify a finite set of constraints that captures the feasible set defined through constraint (12).

Assumption 2. *The discrete distribution $\mathbb{Q}$ is a uniform distribution over all $\omega \in \Omega$. Namely, $p_\omega = 1/N$ for all $\omega \in \Omega$.*

Proposition 4. *Under Assumption 2, constraint* (12) *is equivalent to the constraint that there exists a $\boldsymbol{\rho} \in \{0, 1\}^{(N+1)\times N\times T}$ such that*

$$\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_i^z} z^t_{ij\omega} \geq \sum_{k=0}^{N-1}\beta^t_k(\rho^t_{k+1,\omega}-\rho^t_{k\omega})\sum_{i\in\mathcal{I}} d^t_{i\omega} \qquad \forall t\in\{1,\dots,T\},\omega\in\Omega \tag{13}$$

$$\sum_{\omega\in\Omega}\rho^t_{k\omega} = k \qquad \forall t\in\{1,\dots,T\},k\in\{0,\dots,N-1\} \tag{14}$$

$$\rho^t_{k\omega} \leq \rho^t_{k+1,\omega} \qquad \forall t\in\{1,\dots,T\},\omega\in\Omega,k\in\{0,\dots,N-2\} \tag{15}$$

$$\rho^t_{N\omega} = 1 \qquad \forall t\in\{1,\dots,T\},\omega\in\Omega\,, \tag{16}$$

where $\beta^t_k := \sup_{\eta<(k+1)/N}\beta^t(\eta)$.

Using Proposition 4, which proof is deferred to Appendix A.4, we are able to introduce a finite dimensional MILP reformulation for PECSP:

[PEC-SP2]

$$\underset{\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},\boldsymbol{r},\boldsymbol{\rho}}{\text{minimize}} \sum_{t=1}^{T}\sum_{j\in\mathcal{J}} f^t_j x^t_j + \sum_{j\in\mathcal{J}}\sum_{t=1}^{T} g^t_j y^t_j + (1/N)\sum_{\omega\in\Omega}\sum_{t=1}^{T}\sum_{j\in\mathcal{J}}\sum_{i\in\mathcal{I}} c^t l_{ij} z^t_{ij\omega} + \sum_{t=1}^{T}\sum_{j\in\mathcal{J}}\sum_{m\in\mathcal{J}} \alpha^t r^t_{mj} \tag{17a}$$

subject to (1b) − (1e), (4b), (4c), (7c), (7d), (14) − (16)

$$\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_i^z} z^t_{ij\omega} \geq \sum_{k=1}^{N-1}(\rho^t_{k+1,\omega}-\rho^t_{k\omega})\lceil\beta^t_k\sum_{i\in\mathcal{I}} d^t_{i\omega}\rceil \quad \forall t\in\{1,\dots,T\},\omega\in\Omega \tag{17b}$$

$$x^t_j \in\{0,1\}\,, y^t_j\in\mathbb{N}, \rho^t_{k\omega}\in\{0,1\} \qquad \forall j\in\mathcal{J},t\in\{1,\dots,T\},\omega\in\Omega,k\in\{0,\dots,N\}\,. \tag{17c}$$

The arguments that one needs to use for relaxing the integrality constraints on $\boldsymbol{z}$ and $\boldsymbol{r}$ and rounding up the values $\beta^t_k\sum_{i\in\mathcal{I}} d^t_{i\omega}$ are exactly similar as in Section 4.

Although we derive a MILP based reformulation, it still poses a great computational challenge given that the number of integer variables and number of constraints now grow at the rate of $O(N^2T)$ in PEC-SP2. Therefore, we propose an exact solution method in Section 5.2 and conservative approximation method in Section 6. We also note that the reformulation presented Proposition 4 is a special case of the MILP formulation in Luedtke (2008) for FSD constraints. In fact, for constraint (12) such a reformulation exists as long as $\mathbb{Q}$ is discrete yet might in general require $\boldsymbol{\rho}$ to be of exponential size if $\beta(\cdot)$ is strictly increasing.

5.2. Exact Branch-and-Benders-Cut Method for PECSP

We briefly summarize how to set-up the problem in order for applying the B&BC method presented in Section 4.2. In particular, we can decompose PEC-SP2' using a MP and a set of SPs:

$$\text{[PEC-SP2']} \underset{\boldsymbol{x},\boldsymbol{y},\boldsymbol{\rho},\theta^r,\boldsymbol{\theta}^z}{\text{minimize}} \sum_{t=1}^{T}\sum_{j\in\mathcal{J}} f_j^t x_j^t + \sum_{t=1}^{T}\sum_{j\in\mathcal{J}} g_j^t y_j^t + \sum_{\omega\in\Omega}\sum_{t=1}^{T}\theta_{t\omega}^z + \theta^r \tag{18a}$$

$$\text{subject to } (1b),(1c),(8d),(14)-(16),(17c)$$

$$\theta^r \geq h^r(y) \tag{18b}$$

$$\theta_{t\omega}^z \geq h_{\omega,t}^z(y) \quad \forall t\in\{1,\ldots,T\},\omega\in\Omega \tag{18c}$$

$$\sum_{j\in\mathcal{J}}\lambda_j^t y_j^t \geq \sum_{k=1}^{N-1}(\rho_{k+1,\omega}^t - \rho_{k\omega}^t)\lceil\beta_k^t\sum_{i\in\mathcal{I}} d_{i\omega}^t\rceil \;\; \forall t\in\{1,\ldots,T\},\omega\in\Omega \tag{18d}$$

$$\sum_{j\in\mathcal{J}}\lambda_j^t y_j^t \geq \kappa_t \;\; \forall t\in\{1,\ldots,T\}, \tag{18e}$$

where, for each t, we consider the ordered list $\Omega_t' := \{\omega_k'\}_{k=1}^N$ defined in Proposition 3 and let $\kappa_t := \max_{k=0,\ldots,N-1}\lceil\beta_k^t\sum_{i\in\mathcal{I}} d_{i,\omega_{N-k}'}^t\rceil$. While $h^r(\boldsymbol{y})$ is exactly as defined in (9), $h_{\omega,t}^z(\boldsymbol{y},\boldsymbol{\rho})$ needs to be redefined as PECSP-SP$_{\omega,t}^z$ with the dual variables of $\mu_{jt\omega}^1$, $\mu_{it\omega}^2$ and $\mu_{t\omega}^3$ respectively. We provide the details of the dual sub-problems associated with $h_{\omega,t}^z(\boldsymbol{y})$ and Benders cut in Appendix B.4.

$$\text{[PECSP-SP}_{\omega,t}^z\text{]} \;\; h_{\omega,t}^z(\boldsymbol{y},\boldsymbol{\rho}) := \min_{\boldsymbol{z}_w^t\geq 0}\sum_{j\in\mathcal{J}}\sum_{i\in\mathcal{I}}(1/N)c^t l_{ij} z_{ij\omega}^t \tag{19a}$$

$$\text{subject to } \;\; (10b),(10c) \tag{19b}$$

$$\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_i^z} z_{ij\omega}^t \geq \sum_{k=1}^{N-1}(\rho_{k+1,\omega}^t - \rho_{k\omega}^t)\lceil\beta_k\sum_{i\in\mathcal{I}} d_{i\omega}^t\rceil. \tag{19c}$$

The implementation of the B&BC method is exactly similar to the procedure described in Appendix B.3. Note again that imposing constraint (18d) in the MP ensures that the PECSP-SPs are always feasible and bounded under Assumption 1 when new Benders cuts are needed by the algorithm while otherwise some feasibility cuts can be generated based on the procedure described in Appendix B.5. Furthermore, constraint (18e) implements the valid inequality proposed in Section 4.2.3. Finally, it does not appear that tighter optimality cuts in Section 4.2.4 can be identified for PEC-SP2'. We wish to note that it is also possible to equivalently reformulate PECSP problem in a FSD form similar to the one studied in Drapkin and Schultz (2010) with the hope of employing the proposed cutting pane method. Unfortunately, when doing so the resulting FSD model does not satisfy the conditions imposed by the authors (e.g., continuous recourse variables, complete recourse property) which motivates the need for designing a new decomposition scheme.

6. Conservative Approximation Method

Although the exact B&BC method discussed above can be employed to solve PECSP, it might struggle when solving large scale instances. This motivates us to propose an efficient conservative approximation for PEC-SP2, which can significantly reduces the computational burden.

PROPOSITION 5. *Under Assumption 2, the probabilistic envelope constraint* (12) *is conservatively approximated by*

$$\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_i^z} z^t_{ij\omega'_{N-k}} \geq \lceil \beta^t_k \sum_{i\in\mathcal{I}} d^t_{i\omega'_{N-k}} \rceil \qquad \forall t\in\{1,\ldots,T\}, k\in\{0,\ldots,N-1\}, \tag{20}$$

where $\beta^t_k := \sup_{\eta<(k+1)/N}\beta^t(\eta)$*, and the ordered scenario set* $\Omega'_t := \{\omega'_k\}_{k=1}^N$ *is defined such that for all* $k<k'$ *we have* $\sum_{i\in\mathcal{I}} d^t_{i\omega'_k} \leq \sum_{i\in\mathcal{I}} d^t_{i\omega'_{k'}}$.

The proof can be found in Appendix A.5.

This gives rise to the following conservative approximation formulation for PECSP:

$$\text{[CAPECSP]}\ \underset{\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},\boldsymbol{r}}{\text{minimize}} \sum_{t=1}^T\sum_{j\in\mathcal{J}} f^t_j x^t_j + \sum_{t=1}^T\sum_{j\in\mathcal{J}} g^t_j y^t_j + (1/N)\sum_{\omega\in\Omega}\sum_{t=1}^T\sum_{j\in\mathcal{J}}\sum_{i\in\mathcal{I}} c^t l_{ij} z^t_{ij\omega} + \sum_{t=1}^T\sum_{j\in\mathcal{J}}\sum_{m\in\mathcal{J}} \alpha^t r^t_{mj} \tag{21a}$$

$$\text{subject to (1b)} - \text{(1e)}, \text{(4b)}, \text{(4c)}, \text{(7c)}, \text{(7d)}, \text{(20)}$$

$$x^t_j\in\{0,1\}, y^t_j\in\mathbb{N} \quad \forall j\in\mathcal{J}, t\in\{1,\ldots,T\}. \tag{21b}$$

Problem CAPECSP avoids the need to introduce extra binary variables $\rho^t_{\omega k}$ in the model. Furthermore, it can still be treated using the B&BC method by exploiting the following decomposition:

$$\text{[CAPECSP']}\quad \underset{\boldsymbol{x},\boldsymbol{y},\boldsymbol{\theta}^z,\theta^r}{\text{minimize}} \sum_{t=1}^T\sum_{j\in\mathcal{J}} f^t_j x^t_j + \sum_{t=1}^T\sum_{j\in\mathcal{J}} g^t_j y^t_j + (1/N)\sum_{\omega\in\Omega}\sum_{t=1}^T \theta^z_{t\omega} + \theta^r \tag{22a}$$

$$\text{subject to (1b)}, \text{(1c)}, \text{(8d)}, \text{(21b)}$$

$$\theta^r \geq h^r(y) \tag{22b}$$

$$\theta^z_{t\omega} \geq h^z_{\omega,t}(y) \quad \forall t\in\{1,\ldots,T\}, \omega\in\Omega \tag{22c}$$

$$\sum_{j\in\mathcal{J}} \lambda^t_j y^t_j \geq \max_{k\in\{0,\ldots,N-1\}} \lceil \beta^t_k \sum_{i\in\mathcal{I}} d^t_{i\omega'_{N-k}} \rceil \quad \forall t\in\{1,\ldots,T\}, \tag{22d}$$

where $h^r(\boldsymbol{y})$ is exactly as defined in (9), but $h^z_{\omega,t}(\boldsymbol{y})$ needs to be redefined as:

$$\text{[CAPECSP-SP}^z_{\omega,t}\text{]}\ h^z_{\omega,t}(\boldsymbol{y}) := \min_{\boldsymbol{z}^t_\omega\geq 0} \sum_{j\in\mathcal{J}}\sum_{i\in\mathcal{I}} (1/N) c^t l_{ij} z^t_{ij\omega} \tag{23a}$$

$$\text{subject to (10b)}, \text{(10c)}$$

$$\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_i^z} z^t_{ij\omega} \geq \sum_{k=0}^{N-1} \mathbf{1}\{\omega=\omega'_{N-k}\} \lceil \beta^t_k \sum_{i\in\mathcal{I}} d^t_{i\omega} \rceil. \tag{23b}$$

Note that once again that inequalities (8d) and (22d) ensure that the problems associated to $h^r(\boldsymbol{y})$ and $h^z_{\omega,t}(\boldsymbol{y})$ are both feasible and bounded under Assumption 1. For conciseness, we omit presenting the details of the dual sub-problems associated with $h^z_{\omega,t}(\boldsymbol{y})$ and Benders cut.

Regarding the CCSP model, given that it can be considered a special case of PECSP, one can implicitly establish a conservative approximation model from Proposition 5. This gives rise to what we will refer as the CACCSP model :

$$
\begin{aligned}
\text{[CACCSP]}\ \underset{\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},\boldsymbol{r}}{\text{minimize}}\ & \sum_{t=1}^{T}\sum_{j\in\mathcal{J}} f_j^t x_j^t + \sum_{t=1}^{T}\sum_{j\in\mathcal{J}} g_j^t y_j^t + (1/N)\sum_{\omega\in\Omega}\sum_{t=1}^{T}\sum_{j\in\mathcal{J}}\sum_{i\in\mathcal{I}} c^t l_{ij} z_{ij\omega}^t + \sum_{t=1}^{T}\sum_{j\in\mathcal{J}}\sum_{m\in\mathcal{J}} \alpha^t r_{mj}^t \\
\text{subject to}\ & (1b)-(1e), (4b), (4c), (7c), (7d), (8d), (21b) \\
& \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_i^z} z_{ij\omega_k'}^t \geq \lceil \beta^t \sum_{i\in\mathcal{I}} d_{i\omega_k'}^t \rceil \quad \forall k\leq \bar{k}_t,\, t\in\{1,\ldots,T\} \\
& \sum_{j\in\mathcal{J}} \lambda_j^t y_j^t \geq \lceil \beta^t \sum_{i\in\mathcal{I}} d_{i\omega_{\bar{k}_t}'}^t \rceil \quad \forall t\in\{1,\ldots,T\},
\end{aligned}
$$

where we already included the redundant constraint that ensures feasibility of the subproblems when Assumption 1 holds.

7. Numerical Experiments

To evaluate the numerical performance of our B&BC method, we employ a set of randomly generated CCSP and PECSP instances of different sizes. The parameters in each instance are generated as follows. First, each fixed operation cost f_j^t is generated independently and uniformly from [1000, 1200], each marginal operation cost g_j^t from [100, 120], each marginal transportation cost c^t from [0.5, 1], and each ambulance redeployment cost α^t from [3.5, 5]. We then generate each possible emergency request location and each ambulance base location independently and uniformly over the $[0,10]^2$ square. We also assume that the distribution is uniform over the set of scenarios, which size depends on the problem instance. At each request location and for each scenario ω, the number of requests $d_{i\omega}^t$ that can be served is drawn independently and uniformly from the set $\{1,\ldots,5\}$. The maximum number of emergency vehicles P_j^t that can be hosted at each period t and location j is uniformly drawn {5, 6, 7} while the service rate is fixed to $\lambda_j^t = 4$. For CCSP model we predefined the coverage level $\beta^t = \beta = 0.9$, while the reliability level is set to $\eta^t = \eta = 0.05$ unless stated otherwise. Finally, we fix the coverage envelope to $\beta^t(\eta) = \eta$ for the PECSP model.

All of our algorithms are implemented in C and use CPLEX 12.71. All runs were conducted on a Desktop machine with Intel(R) Xeon(R) 3.30 GHz processor and 32 GB RAM in a Windows 64-bit system. A single thread is used in all experiments. For all instances, the algorithms that employ the B&BC method are run until an optimality gap below 1% is reached. In the case of experiments where the CCSP model is solved, we set a maximum runtime of 3600 seconds for all algorithms, while we let the maximum runtime be 7200 seconds for the PECSP model.

In what follows we start in Section 7.1 by comparing the performance of our B&BC method to the BB algorithm proposed in Beraldi and Bruni (2009) and to the CPLEX solver. We then in Section 7.2 present an exhaustive computational study of the performance of employing B&BC on the exact and approximate versions of the CCSP and PECSP models when $T = 6$.

7.1. Algorithmic Comparisons

In Beraldi and Bruni (2009), the authors propose a BB algorithm to solve a problem of the similar form as CCSP yet where a single period is considered. It is therefore possible to compare the numerical efficiency of both this BB algorithm and the B&BC method for solving the CCSP model. Note however, that the B&BC cannot be directly applied to solve the model in Beraldi and Bruni (2009) because of the difference in structure of the second stage cost.

To perform this analysis we ran experiments with instances of sizes that are comparable to the sizes discussed in Beraldi and Bruni (2009). Specifically, we considered $|\mathcal{I}| \in \{40, 80, 150\}$, $|\mathcal{J}| \in \{20, 50, 100\}$, and $N \in \{10, 20, 30, 40, 100, 200\}$ while the reliability level was considered with $\eta \in \{0.05, 0.10, 0.15\}$ and $\beta = 0.9$. Our implementation of B&BC method includes valid inequalities and tighter optimality cuts presented in sections 4.2.3 and 4.2.4 and is referred as VIOC-B&BC. Each experiment involves running the algorithms on the same five instances of each class of problems. We also use the CPLEX solver to solve directly this problem.

Table 2 presents the minimum, maximum, and average CPU time (in seconds) of all methods. In the *max* column, we also report the proportion of instances that could not be solved exactly within 3600 seconds. A quick read of this table confirms that the B&BC method achieves significantly better performance than BB algorithm and the CPLEX solver. In fact, BB algorithm is unable to solve to optimality most of the problem instances that are considered within 3600 seconds while the CPLEX solver fails to solve about 30% of instances (64 out of 270). In contrast, our B&BC method allows us to solve nearly all problems (267 out 270) within an hour. In fact, we notice that the BB algorithm's performance quickly degrades as the number of scenarios considered is increased and similarly, although less drastically, as η increases.

In order to understand the poor performance of the BB algorithm, one needs to know that this algorithm starts with the search of all feasible and "non-dominated" assignments for $\boldsymbol{\rho}$ by traversing, using a customized rule, a tree where each node defines a specific assignment of binary values for $\boldsymbol{\rho}$. Once all of these candidates are identified, the algorithm iteratively solves the MILP associated to each of these candidates. While this can be efficient when η is small and the number of feasible candidates is reasonable, it quickly becomes impossible to assemble such a list for larger η's and more importantly larger N (since the number of branches at each node grows linearly in N). In contrast, our B&BC algorithm only solves a single MILP. Overall, it appears clear to us that the B&BC method is much better suited to solve problems of realistic sizes as we discuss next.

Table 2 Numerical performance comparisons with BB method in Beraldi and Bruni (2009) and the CPLEX solver. We report the average (avg), maximum (max), minimum (min) CPU time (in seconds) over each five instances.

$\lvert\mathcal{I}\rvert$	$\lvert\mathcal{J}\rvert$	η	N	BB			CPLEX			VIOC-B&BC		
				avg	max	min	avg	max	min	avg	max	min
40	20	.05	10	70	134	2	141.8	337.7	45.3	.2	.2	.2
			20	60	177	41	570	2,842.7	.7	.2	.3	.2
			30	1,483	352	189	2.5	4.4	1	.3	.4	.2
			40	3,005*	>3,600[.4]	1,672	6.6	9.4	4.1	.3	.4	.2
			100	-	-	-	91.3	143.1	49.6	.7	1.2	.4
			200	-	-	-	1,205.3*	>3,600[.2]	318	1.4	1.7	1
		.10	10	144	379	5	1.2	2.4	.4	.2	.3	.1
			20	489	1,374	115	3.8	5.1	2.9	.3	.5	.1
			30	-	-	-	6.3	8.9	2.1	.1	.2	.1
			40	-	-	-	23.1	51.4	10.8	.3	.4	.2
			100	-	-	-	469.3	852.1	70.3	.6	.9	.3
			200	-	-	-	2,455.3*	>3,600[.6]	1,655.4	1.3	2.7	.8
		.15	10	141	332	3	.4	.4	.3	.2	.2	.1
			20	1,761*	>3,600[.4]	317	5.6	10.9	3	.2	.3	.2
			30	-	-	-	13.1	18.5	9.3	.2	.3	.1
			40	-	-	-	41.3	80.3	16.9	.5	.6	.3
			100	-	-	-	2,085*	>3,600[.2]	731.2	1.1	2.9	.5
			200	-	-	-	3,581.8*	>3,600[.8]	3,581.8	1.3	1.9	.8
80	50	.05	10	106	253	4	2*	>3,600[.6]	1.8	1.2	2	.6
			20	301	1,276	29	728.3*	>3,600[.2]	6.3	1.2	1.6	.8
			30	2,791*	>3,600[.6]	1,271	17.9	30.6	11.3	2.7	4.1	1.2
			40	-	-	-	38.4	58.4	21.7	5.5	8.8	2
			100	-	-	-	780.6	1,153.1	316.2	10.7	15.5	4.6
			200	-	-	-	-	-	-	59.6	200	13.1
		.10	10	107	323	11	2.7*	>3600[.2]	1.7	.6	.7	.5
			20	1,577*	>3,600[.2]	412	18.1	33	9.6	1.2	1.4	1
			30	-	-	-	46.3	64	29	1.8	2.2	1.4
			40	-	-	-	114.7	163.3	88.1	2.7	4.1	1.9
			100	-	-	-	2,090.3*	>3600[.2]	761.7	8.4	11.9	6.4
			200	-	-	-	-	-	-	737*	>3,600[.2]	17.8
		.15	10	144	593	11	2.5	4.4	1.4	.7	.9	.4
			20	2,236*	>3,600[.2]	1,328	26.3	33.9	13.9	1.1	1.3	.9
			30	-	-	-	66.2	92.6	40.1	2.4	3.5	2
			40	-	-	-	135.9	195.4	107.5	3.2	4.4	2.5
			100	-	-	-	2,365.2*	>3600[.4]	2,124.5	7.3	14.5	3
			200	-	-	-	-	-	-	731*	>3,600[.2]	9
150	100	.05	10	779	1,900	157	6.5*	>3600[.8]	6.5	16.8	32.3	7.9
			20	1,133	3,054	307	62	89.1	41.3	27	43.7	17.8
			30	3,061*	>3,600[.8]	906	127.7	186.9	67.9	48.5	120	22.1
			40	-	-	-	256.1*	>3600[.2]	153.6	67.1	127	31.1
			100	-	-	-	3,031.9*	>3600[.6]	2,563.8	108	211	65.6
			200	-	-	-	-	-	-	277	543	156
		.10	10	927	2,830	25	12.8	53.7	10.2	8.4	12.8	3.5
			20	2,472	2,980	1,277	1,297.5	3,594.6	60	14.3	16.1	11.9
			30	-	-	-	359	447.3	233.6	25.4	34.7	19.8
			40	-	-	-	620.6	865.1	480.3	34	42.9	20
			100	-	-	-	2,449.9*	>3600[.8]	2,449.9	128	224	49.5
			200	-	-	-	-	-	-	842*	>3,600[.2]	65.6
		.15	10	1,587	2,672	109	13.1	815.8	13.1	15.6	32.2	5.4
			20	3,384*	>3,600[.6]	2,901	183.9	230.6	108.5	22.5	36.2	16.2
			30	-	-	-	481.5	672.2	289.6	47.9	77.4	32.7
			40	-	-	-	924.4	1,293.5	578.5	46.3	60.5	36.3
			100	-	-	-	-	-	-	145	203	73.9
			200	-	-	-	-	-	-	243	314	186

"-" means that none of the five instances were solved in less 3600 seconds.
"*" means the average was computed only on instances that are solved optimally in less 3600 seconds.
[·] in column of *max* means the proportion of unsolved instances in less 3600 seconds.

7.2. Numerical Performance of our B&BC algorithms

This section reports on a set of experiments performed using classes of problems of different sizes, defined by the respective sizes of $|\mathcal{I}|$, $|\mathcal{J}|$, T, and N, for both the CCSP and PECSP model. The definition of each class of problems is described in Appendix B.6. Each class of problems contains 10 randomly generated instances. Overall, the set of test instances includes 240 problems with number of ambulance base locations ranging from 40 to 150, number of emergency request locations ranging from 20 to 100, with 6 time periods, and with discrete distribution supported on a number of scenarios ranging from 50 to 200. We believe that this set of problems covers well the size of problems that should emerge in practice for small to medium sized territories.

We evaluated the CPU time needed to obtain a 1% sub-optimal solution using the following variations of the method:

- Raw-B&BC refers to applying B&BC to solve either CCSP2' and PECSP2' in the CCSP and PECSP studies respectively.
- VIOC-B&BC refers to applying B&BC to solve CCSP2' with both valid inequalities and strengthened optimality cut from sections 4.2.3 and 4.2.4.
- VI-B&BC refers to applying B&BC to solve PECSP2' with valid inequalities (18e).
- CA-B&BC refers to applying B&BC to solve CACCSP and CAPECSP in the CCSP and PECSP studies respectively.

Table 3 presents the maximum CPU time and the average time for MP and SPs required to solve the 10 instances of each class of the CCSP model (C1-C12). We also report the proportion of problems that could not be solved within 3600 seconds. The bottom row of Table 3 finally reports the average of each of these statistics over the set of twelve problem classes. The first thing that one might notice from this table is how the Raw-B&BC algorithm fails to solve most (i.e. 103 out of 120 instances) of the instances. In the case of VIOC-B&BC method, the performance is surprisingly better given that the failure rate is below 2%. One might also notice that even with the easier problems, the Raw-B&BC algorithm seems to struggle much more than the VIOC-B&BC method in terms of maximum CPU time and average time for MP and SPs. This seems to confirm that the use of the valid inequalities proposed in Section 4.2.3 and strengthened optimality cut proposed in Section 4.2.4 play a critical role in improving the numerical efficiency. It seems to be that conservative approximation does not speed up so much when comparing with the VIOC-B&BC method. But we have confirmed that CA method can quickly identify the optimal solution with a smaller tolerance (i.e. 0.1%) when comparing with the VIOC-B&BC method. Overall, one can certainly draw the conclusion that both VIOC-B&BC and CA-B&BC algorithms are valuable approaches for solving larger-scale CCSP models.

Table 3 Numerical performance for CCSP with respect to three implementations summarized in Section 7.2, in which we report the average time for MP (avmp) and SPs (avsp), maximum (max) CPU time (in seconds) and the unsolved proportion (prop) within 3600 seconds over ten instances for each data class (C1-C12), $\beta = 0.9, \eta = 0.05$.

Class	Raw-B&BC				VIOC-B&BC				CA-B&BC			
	avmp	avsp	max	prop	avmp	avsp	max	prop	avmp	avsp	max	prop
C1	-	-	-	1	0.33	2	5	.00	0.03	1	2	.00
C2	13*	14*	>3,600	.9	0.28*	3*	>3,600	.1[2.18]	0.05*	3*	>3,600	.1[1.55]
C3	61*	15*	>3,600	.9	0.44	4	6	.00	2.65	4	40	.00
C4	80*	9*	>3,600	.6	0.78	5	11	.00	0.10	6	10	.00
C5	25*	141*	>3,600	.8	0.24	12	15	.00	0.51	19	82	.00
C6	37*	114*	>3,600	.9	0.40	23	45	.00	0.12	22	41	.00
C7	-	-	-	1	17.95	56	465	.00	0.14	28	43	.00
C8	5*	72*	>3,600	.9	1.19	47	91	.00	0.30	48	98	.00
C9	9*	545*	>3,600	.6	0.23	80	163	.00	0.16	87	220	.00
C10	11*	1,415*	>3,600	.8	0.40*	144*	>3,600	.1[1.58]	0.12*	132*	>3,600	.1[1.64]
C11	-	-	-	1	0.62	212	283	.00	0.29	280	594	.00
C12	58*	1,079*	>3,600	.9	0.79	259	317	.00	0.24	300	440	.00
average	33*	378*	-	.86	2.0*	70*	-	.016	0.39*	78*	-	.016

"-" indicates that no instances were solved in less 3600 seconds.

"*" means the average was computed only on instances solved optimally in less 3600 seconds.

[·] in column of *prop* means the average optimality gap (in %) for instances beyond 3600 seconds.

Table 4 Numerical performance for PECSP with respect to four implementations summarized in Section 7.2, in which we report the average time for MP (avmp) and SPs (avsp), maximum (max) CPU time (in seconds) and the unsolved proportion (prop) over ten instances for each data class (C13-C24).

Class	Raw-B&BC				VI-B&BC				CA-B&BC			
	avmp	avsp	max	prop	avmp	avsp	max	prop	avmp	avsp	max	prop
C13	-	-	-	1[10.8]	108	4	498	.00	0.17	1	1	.00
C14	-	-	-	1[33.2]	3,667*	97*	>7,200	.1[1.78]	0.11	1	5	.00
C15	-	-	-	1[44.7]	5,689*	858*	>7,200	.4[3.11]	2.94	6	32	.00
C16	-	-	-	1[47.9]	-	-	-	1[8.90]	2.63	9	35	.00
C17	199*	1,623*	>7,200	.7[7.08]	545*	7*	>7,200	.3[1.11]	3.54	2	35	.00
C18	-	-	-	1[36.9]	4,055*	821*	>7,200	.7[2.28]	0.26	5	20	.00
C19	-	-	-	1[42.4]	3,125*	3,805*	>7,200	.7[2.83]	0.07	6	11	.00
C20	-	-	-	1[49.5]	-	-	-	1[2.28]	12.15	63	378	.00
C21	-	-	-	1[7.67]	377*	375*	>7,200	.4[1.43]	0.14*	10*	>7,200	.1[1.18]
C22	-	-	-	1[40.4]	1,268*	968*	>7,200	.4[1.20]	0.62	44	145	.00
C23	-	-	-	1[45.7]	1,026*	5,374*	>7,200	.7[3.54]	0.12	61	142	.00
C24	-	-	-	1[51.9]	-	-	-	1[6.52]	0.12	155	242	.00
average	199*	1,623*	-	.98	2,207*	1,367*	-	.54	1.91*	30*	-	.008

"-" indicates that no instance were solved in less 7200 seconds.

"*" means the average was computed only on instances that are solved optimally in less 7200 seconds.

[·] in column of *prop* means the average optimality gap (in %) for instances beyond 7200 seconds.

Table 5 The performance of conservative approximation for CCSP and PECSP when comparing with exact methods of VIOC-B&BC and VI-B&BC, in which avg gap for VIOC-B&BC (or VI-B&BC) is averaged optimality gap over ten instances for each data class, av_CA_opt is the average objectives for CA-B&BC and CA-B&BC, av_gap can be computed by $\mathbb{E}[(\textbf{CA-opt} - \textbf{LB})/\textbf{LB} \times 100\%]$, and Δ measures the average gap difference (in percentile points) between our conservative approximation and exact method.

CCSP						PECSP					
Class	VIOC-B&BC		CA-B&BC			Class	VI-B&BC		CA-B&BC		
	av_LB	av_gap%	av_CA_opt	av_gap%	Δ%		av_LB	av_gapRC	av_CA_opt	av_gap%	Δ%
C1	83,877	0.82	84,209	0.40	-0.42	C13	33,459	0.96	33,869	1.23	0.27
C2	81,786	1.89	82,184	0.50	-1.39	C14	32,016	1.02	32,447	1.35	0.33
C3	84,222	0.74	84,462	0.29	-0.45	C15	33,871	1.77	34,333	1.46	-0.31
C4	80,832	0.83	82,068	0.32	-0.51	C16	33,739	8.90	34,815	3.20	-5.70
C5	152,438	0.70	153,225	0.52	-0.18	C17	63,900	0.90	64,590	1.08	0.18
C6	150,395	0.62	151,144	0.50	-0.12	C18	64,405	1.85	65,737	2.07	0.22
C7	149,757	0.76	150,555	0.53	-0.23	C19	63,411	2.31	64,282	1.37	-0.94
C8	152,434	0.81	153,846	0.54	-0.27	C20	64,406	1.84	63,916	0.76	-1.08
C9	278,551	0.86	280,682	0.76	-0.10	C21	128,061	1.03	128,880	0.64	-0.39
C10	277,601	0.46	278,554	0.35	-0.11	C22	127,039	0.84	127,633	0.47	-0.37
C11	279,645	0.41	280,519	0.31	-0.10	C23	133,069	3.05	135,110	1.53	-1.52
C12	279,337	0.46	280,315	0.36	-0.10	C24	127,879	6.52	129,189	1.02	-5.50

Table 4 presents the numerical performance of three algorithms for PECSP models of class C13-C24 using a format that is exactly similar to Table 3 except that maximum runtime is 7200 seconds. The observations here are similar as before concerning the superior performance of VI-B&BC over Raw-B&BC, where only three instances are solved by Raw-B&BC while VI-B&BC can solve about 46% of instances. Yet, the fact that more than half of problems could not be solved to optimality within 7200 seconds still raises some concerns regarding whether VI-B&BC can reliably be used to solve PECSP model of realistic sizes. The CA-B&BC algorithm appears however much more promising in this regard. All problem instances except one were solved within 2 hours. Furthermore, over this set of 119 instances, the maximum resolution time was about 7 minutes.

The fact that the conservative models (CACCSP and CAPECSP) can be solved so efficiently raises the question of what is the loss in terms of quality. Table 5 sheds some light on this issue by presenting the average of the lower bound identified by the exact method VI-B&BC for PECSP and its remaining optimality gap for problem instances of each class (C13-C24), and VIOC-B&BC for CCSP and its remaining optimality gap for problem instances of each class (C1-C12). It also presents for both types of problems the average optimal value of the solutions produced by the conservative model, the average optimality gap (based on the obtained lower bounds). Finally, we express in the sixth and twelfth columns the difference between the average gap for conservative model and the average gap reached by exact method. Overall, regarding the CCSP model, we surprisingly see that all the values of the difference in optimality gap for CA-B&BC are negative, which indicates that the optimal solutions for conservative model are practically speaking optimal.

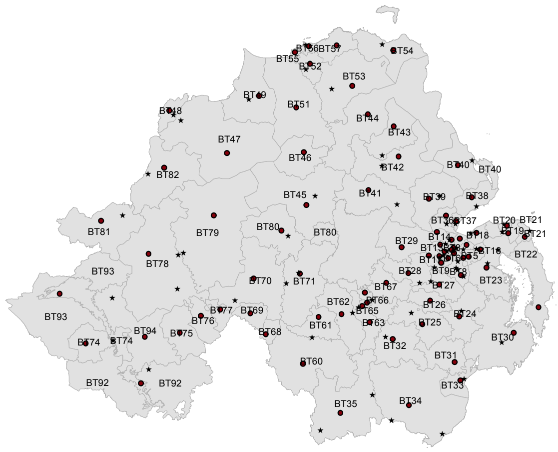


Figure 1 **ZIP code based geographical partition for Northern Ireland, solid dots represent 80 emergency demand zones and stars represent 63 potential base locations respectively ($|\mathcal{I}| = 80, |\mathcal{J}| = 63$).**

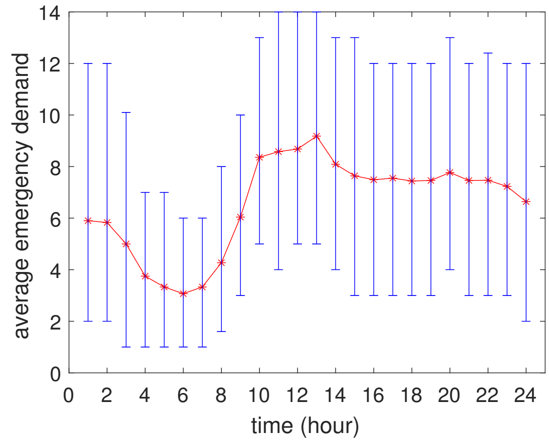


Figure 2 **Average emergency demand in 24-hour cycle, based on emergency demand data during 03/2015-12/2016. The bars represent the 5% and 95% quantiles on the hourly emergency demand respectively.**

For the PECSP model, the sizes of average optimality gaps for CA-B&BC seem to confirm that the solutions from the conservative model are of good quality. In fact, based on the fact that for 8 out of 12 problem classes the difference in optimality gap for CA-B&BC is negative and the average value is 1.98%. One might even argue that conservative solutions are of marginally better quality than the feasible solutions reached for the exact method after 2 hours of computations.

8. A Northern Ireland Ambulance Service Health and Social Care Trust Case Study

In this case study, we consider an EMS location problem in the region of Northern Ireland. We start by describing the context in which we are assuming that this EMS network is being operated together with the historical dataset that was used to perform our analysis. We then present and discuss the structure of the optimal configurations obtained from solving the CCSP and PECSP models before illustrating our findings regarding the trade-offs that can be made between expected total cost and coverage reliability. We conclude this section by performing an out-of-sample analysis that confirms that our conclusions should remain relatively valid when looking forward in time.

8.1. Context

We consider the territory to be divided into 80 emergency demand zones, each of which consists of the area where the same 4-digit ZIP code is used. We also consider 63 potential locations for ambulance bases based on existing EMS stations. Figure 1 illustrates these geographical details.

We have historical emergency demand request data gathered by the NIASHSCT over the period ranging from April 2015 until December 2016. More than 113,000 emergency requests are reported in Northern Ireland over this period. Figure 2 shows statistics about the number of requests at different hourly periods of a day during this period. We can observe that the statistics of emergency requests are highly time-dependent. For example, emergency demand stays below 5 requests per hour between 4:00am and 8:00am, while from 10:00am until 12:00am, it stays above 8 requests per hour. Moreover, we observed that this demand is unevenly distributed over the territory. In our CCSP and PECSP models, we consider that a day is divided into 6 periods of 4 hours each (i.e. 12:00am - 4:00am, 4:00am - 8:00am, ..., 8:00pm - 12:00am) and let each day of our historical dataset identify a possible scenario for the emergency demand.[1] Distances are computed using geographical distances obtained by the ArcGIS, between the center of the ZIP code zone and the location of the ambulance base. This appears reasonable given that most of the ZIP code regions are relatively small (i.e. an area of 125 km^2 on average). We let the time-dependent fixed-charge cost be $f_j^{1:T} = [200\ 300\ 350\ 450\ 500\ 500]^\intercal$, for all j, to model decreasing marginal cost as the ambulance base is run for a longer time period, given that bases that are open need to stay open until the end of the day. We further let the marginal transportation cost take the form $c = [2\ 2\ 1\ 1\ 1.5\ 1.5]^\intercal$ and let ambulance relocation costs as $\alpha = [5\ 5\ 4\ 4\ 4.5\ 4.5]^\intercal$ to model higher salaries for late night or early morning shifts. The vehicle operating and maintenance cost g_j^t is set uniformly to 40. We assume that each ambulance can serve at most $\lambda_j^t = 2$ emergency requests in each time period while the maximum capacity of each base is assumed to be $P_j^{1:T} = [2\ 1\ 2\ 3\ 2\ 1]^\intercal$ irrespective of the location j in order to simulate periods of the day where hospitals are more congested and cannot host as many ambulances. Here we consider a maximum travel distance $l_{\max}^z = 40$ km for the coverage of emergency demand and a maximum distance of $l_{\max}^r = 60$ km for relocation of vehicle. In all experiments, the distribution of demand requests consists of the empirical distribution of a set of observed daily demand realizations from the historical dataset. For the experiments conducted in sections 8.2 and 8.3, both CCSP and PECSP use 100 observations (January 2016 - mid April 2016). In Section 8.4, we will investigate the quality of solutions obtained from such empirical distributions by varying the sample size. In all three subsections, CCSP and PECSP are approximated using CACCSP and CAPECSP in order to accelerate computations and the two models are solved using CA-B&BC and CA-B&BC respectively.

For the CACCSP model, we usually consider that EMS managers want to ensure 95% of the emergency requests can be satisfies with more than 95% probability, i.e. that $\beta^t = \beta = 0.95$ while $\eta^t = \eta = 5\%$. The exception will be in Section 8.3 where the sensitivity of total expected cost to the choice of both β and η will be studied.

[1] Note that the size of each period is similar to what is used in Schmid and Doerner (2010) and van den Berg and Aardal (2015). Moreover, it is small enough to capture the time-dependent demand while avoiding the need for both a more sophisticated optimization model that captures the transient behavior of ambulances, and for large number of observations that can support the time and geographical emergency request dependencies.

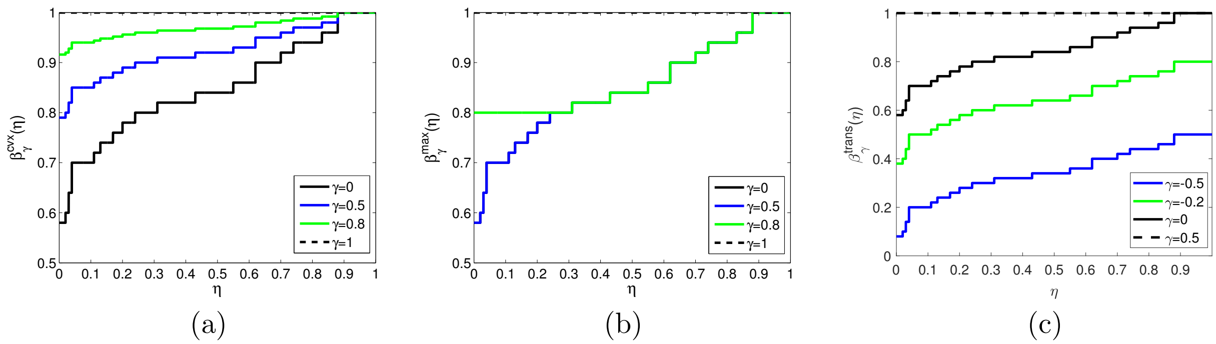


Figure 3 **Examples of probabilistic coverage envelopes obtained using the parametric families $\beta_\gamma^{cvx}(\eta)$ (in (a)), $\beta_\gamma^{max}(\eta)$ (in (b)) and $\beta_\gamma^{trans}(\eta)$ (in (c)).**

In the case of the CAPECSP model, we considered three parametric families of envelope. Given a reference first stage solution $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})$, e.g. an optimal solution to the DM formulation, we evaluate the coverage profile $\bar{\beta}(\eta)$ that is achieved by this candidate solution (we refer the reader to Appendix B.7 for more details on how to obtain $\bar{\beta}(\eta)$). Assuming that this coverage profile is considered as a benchmark for decision-makers, we assume in the first family of envelopes that $\beta_\gamma^{cvx}(\eta) := \gamma + (1-\gamma)\bar{\beta}(\eta)$, i.e. she/he is interested in finding a solution that reduces by a factor of $(1-\gamma)^{-1}$ the proportion of uncovered population for each reliability level η since $1 - \beta_\gamma^{cvx}(\eta) = (1-\gamma)(1-\bar{\beta}(\eta))$. Note that $\beta_\gamma^{cvx}(\eta)$ allows to generate a continuum of increasingly restrictive envelopes between $\bar{\beta}(\eta)$ and 1 as γ goes from 0 to 1, i.e. if $\gamma_1 \leq \gamma_2$ then $\beta_{\gamma_1}^{cvx}(\eta) \leq \beta_{\gamma_2}^{cvx}(\eta)$ for all η. Alternatively, we consider a second monotone parametric family of coverage envelopes that takes the form $\beta_\gamma^{max}(\eta) := \max\left(\gamma, \bar{\beta}(\eta)\right)$, which simply imposes that the demand coverage is at least γ over all scenarios. Finally, we consider the third family of "translated" benchmark, $\beta_\gamma^{trans}(\eta) := \min\left(1, \max(0, \bar{\beta}(\eta) + \gamma)\right)$ for any $\gamma \in [-1, 1]$, which also allows us to explore, using $\gamma \leq 0$, how much money could be saved by reducing "uniformly" the coverage profile. Examples of these parametric families of probabilistic coverage envelopes are presented in Figure 3. In the study that follows, we will employ $\beta_{0.5}^{cvx}(\eta)$ except for Section 8.3 where we study the sensitivity of the total expected cost with respect to the choice of the $\beta(\eta)$ function.

Remark 2. We note that while this case study employs real emergency request data to support its findings, the choice of cost parameters $\boldsymbol{f}$, $\boldsymbol{g}$, $\boldsymbol{c}$, and $\boldsymbol{\alpha}$ and service rate $\boldsymbol{\lambda}$ should be only considered reasonable guesses made by the authors based on values found in the EMS location literature (see for instance Noyan (2010) and Boujemaa et al. (2017)). While we are confident that the practical insights that will be drawn in our analysis should hold under more accurate choices of parameters, additional efforts would be needed to calibrate the CACCSP and CAPECSP models in a way that can provide precise guidelines on how to improve the EMS system in this region.

Table 6 **Characteristics of optimal strategies proposed by the CACCSP and CAPECSP models. The scenario ω_1 is a representative scenario for CACCSP and CAPECSP where lower coverage is achieved.**

CACCSP: Expected total cost = 52,770$						
Time periods	1	2	3	4	5	6
Total number of opened ambulance bases	11	21	21	21	21	21
Total number of occupied bases	11	21	15	14	13	21
Total number of bases providing service in ω_1	8	7	14	0	10	19
Total emergency requests in ω_1	13	10	36	46	26	37
Effective coverage level for ω_1	1.000	1.000	0.972	**0**	0.962	0.973

CAPECSP: Expected total cost = 55,326$						
Time periods	1	2	3	4	5	6
Total number of opened ambulance bases	12	22	22	22	22	22
Total number of occupied bases	12	22	15	13	15	22
Total number of bases providing service in ω_1	6	7	13	11	10	18
Total emergency requests in ω_1	13	10	36	46	26	37
Effective coverage level for ω_1	1.000	1.000	0.917	**0.804**	1.000	0.865

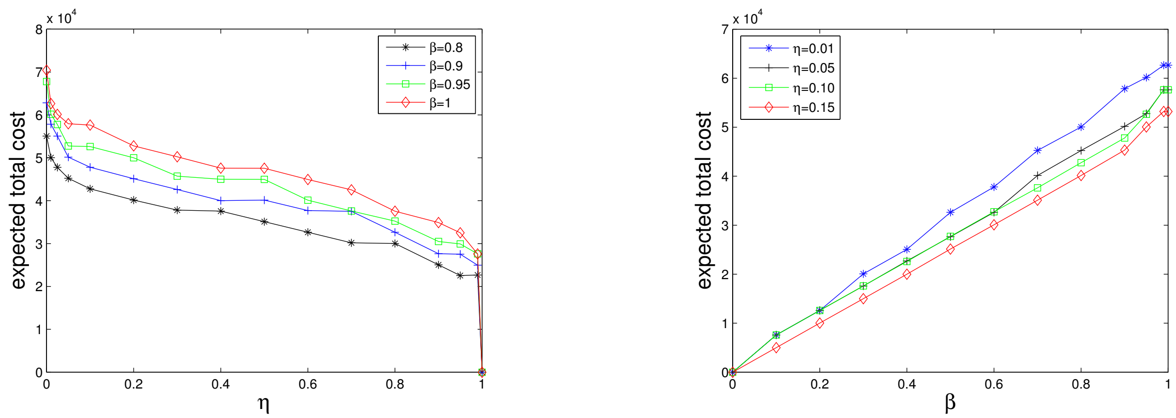
8.2. Time-dependent Optimal EMS Location Configuration

We start by presenting in Table 6 some information about the structure of the two optimal strategies obtained from solving the CACCSP and CAPECSP models. In particular, one might recognize that a large number of ambulance bases are needed at time $t = 2$ and that these bases remain open (based on our modeled requirement) although they are not all occupied in future periods. This table also presents information about the optimal activity while geographical illustrations are presented in the Appendix B.8. For the CACCSP and CAPECSP models respectively under the same scenario ω_1. In the case of CACCSP, we clearly see that scenario ω_1 was identified as an extreme scenario for time $t = 4$ where the number of emergency requests reaches 46. Consequently, the plan for this scenario and time period simply recommends keeping all ambulances idle at their bases. Scenario ω_1 is also considered a hard scenario to cover by the CAPECSP model yet we see that the use of a probabilistic envelope leads to an optimal strategy that still recommends covering a reasonable proportion of the requests, albeit achieves an 80.4% coverage for the hard scenario of $t = 4$. This is achieved at the price of incurring 2,556$ overhead in terms of expected total cost and some reduced coverage for the more optimistic scenarios. Finally, we note that the possible trade-off between expected cost and coverage at any reliability level η can easily be explored using the PECSP or CAPECSP models. This is in fact the subject of our next section.

8.3. Trade-off between Cost Management and Coverage

We next performed a sensitivity analysis to investigate what are the trade-offs between expected total cost and emergency request coverage performance. In particular, we first focus on the CACCSP model and consider the effect of changing the reliability level η and the coverage target β. The result of this analysis is presented in Figure 4 where one can already remark that a lower permitted violation probability η and a higher coverage target β necessarily leads to an increase in expected total cost. Looking more closely at Figure 4(a), one might realize that the marginal cost per percentile point (p.p.) of additional reliability becomes more expensive in the region where $\eta \leq 0.05$. Specifically, for $\beta = 0.95$ the rate is around -3,005 \$ per p.p. compared to a rate of around -262 \$ per p.p. for $\eta \in [0.05,\ 0.9]$. This is due to the fact that when $\eta \leq 0.05$, EMS system must now be prepared for the more extreme scenario, which typically occur in Winter, requiring a larger fleet size to achieve a higher reliability, in turn which incurs larger fixed cost for resources that are unused under most scenarios. The total expected cost also drastically falls to zero when η gets to 1 due to the fact that EMS system managers can then simply completely shut down its operations.

In Figure 4(b) one can observe that the effect of increasing the targeted coverage β on expected total cost is nearly linear and grows at a rate between 532 \$ per p.p. and 626 \$ per p.p. depending on the magnitude of η. It also appears that the marginal cost is higher when the permitted violated probability is smaller possibly due to the fact that the emergency requests that need to be covered in the more extreme scenario are either more numerous or more spread out on the territory. The fact that the marginal cost appears nearly constant when η is fixed is particularly interesting. Intuitively, this is due to the fact that as β is increased the EMS system manager can continue to focus his efforts on the same scenarios of emergency demand, gradually investing in more resources in order to increase his cover of these scenarios. This is unlike the case of a marginal change of η, which can force the EMS system managers to deal with new and more extreme scenarios thus creating a non-linear effect on the amount of resources needed.



(a) (b)

Figure 4 Sensitivity of optimal expected total cost to targeted coverage and reliability in the CACCSP model. (a) presents the optimal expected total cost as a function of the permitted violation probability η for different coverage level β. (b) presents the expected total cost as a function of the coverage level β for different values of permitted violation probabilities η.

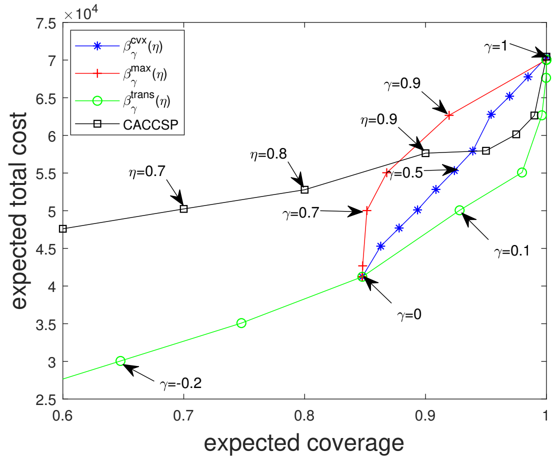


Figure 5 **Trade-off between expected total cost and expected coverage for CAPECSP model using $\beta_\gamma^{cvx}(\eta)$, $\beta_\gamma^{max}(\eta)$ and $\beta_\gamma^{trans}(\eta)$ with different values of γ and for CACCSP model with different values of η.**

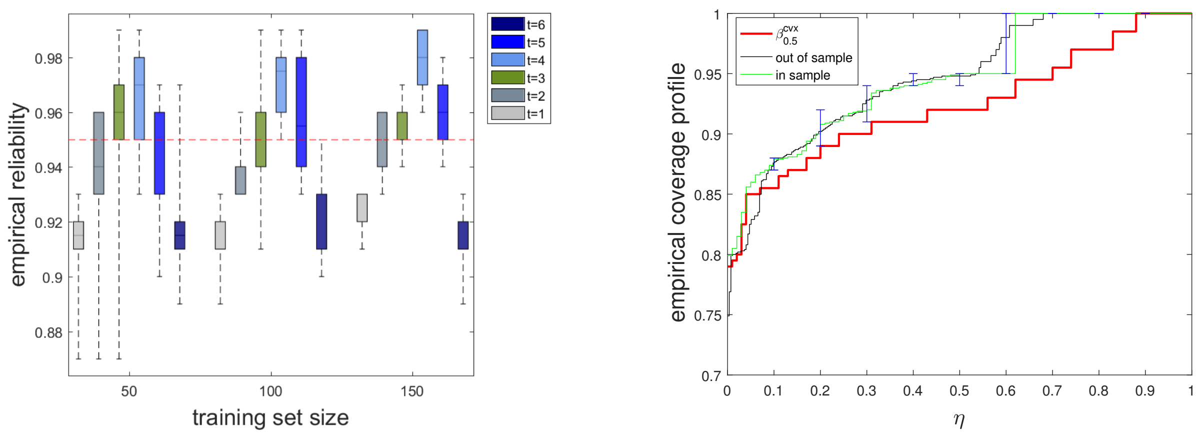
We also analysed the sensitivity of the optimal expected total cost of the CAPECSP model for different choices of probabilistic coverage envelope $\beta(\eta)$, regarding $\beta_\gamma^{cvx}(\eta)$ and $\beta_\gamma^{max}(\eta)$ with values of $\gamma \in \{0,\ 0.1,\ 0.2,\ \ldots,\ 1\}$, and $\beta_\gamma^{trans}(\eta)$ with values of $\gamma \in \{-0.2,\ -0.1,\ 0,\ 0.1,\ \ldots,\ 1\}$. Figure 5 presents the trade-off between expected coverage required by the probabilistic envelope, i.e. $\int_0^1 \beta(\eta)d\eta = \frac{1}{N}\sum_{k=0}^{N-1}\beta_k$, and the expected total cost. In words, expected coverage reflects the overall probability that an emergency request, drawn randomly from any scenario and uniformly from the requests made in that scenario, ends up being served by an ambulance in the solution that satisfies the PEC. For example, the expected coverage in the CACCSP model is $0.95 \cdot 0.95 + 0.05 \cdot 0 = 0.9025$ when $\beta = 0.95$ and $\eta = 0.05$. First, one can notice from Figure 5 that given a targeted expected coverage, the choice of probabilistic envelope has a significant impact on the investment that needs to be made. In particular, the CAPECSP model that employs $\beta_\gamma^{max}(\eta)$ identifies more expensive solutions. This is because all of the investment goes in improving coverage in the most extreme scenarios, which is more expensive (in \$ per p.p.) than in the less extreme ones since extreme scenarios have larger total demand. In particular, looking at the curve associated to $\beta_\gamma^{max}(\eta)$, we note that it costs about 8,528\$ to raise the worst-case coverage from 58% to 70%, which effectively accounts for less than 1 p.p. of expected coverage. Otherwise, we see that when using $\beta_\gamma^{cvx}(\eta)$ the marginal cost for improving expected coverage is nearly constant and close to 1,738\$ per p.p. while it is of nearly 1,072\$ with $\beta_\gamma^{trans}(\eta)$. Alternatively, one can estimate based on $\beta_\gamma^{trans}(\eta)$ that it would be possible to decrease the total expected cost by 587\$ per p.p. of expected coverage reduction. In this regard, we refer the reader to the seminal work of Keeney and Raiffa (1976) and to the extensive literature that followed for guidance on how to identify the right subjective trade-off in such a multi-objective environment. Finally, it is clear that the CAPECSP model allows us to identify strategies that strictly dominates the strategies obtained using the more restrictive CACCSP model. This is done while controlling under which type of scenario is the coverage increased or decreased.

8.4. Out-of-Sample Analysis

This section focuses on a out-of-sample analysis of the quality of solutions obtained from the CACCSP and CAPECSP models. Generally speaking this part of our study serves two purposes. The first one is to highlight the peculiarities associated to the implementation of a solution of the CACCSP/CAPECSP models given that these are approximated using scenarios. In particular, one should expect in practice that the realized emergency demand scenario is not a member of the assumed scenario list. The question therefore arises regarding how much coverage should be provided since the optimal policy is not defined for this scenario. The second objective of this study aims at verifying whether the performance of probabilistic coverage, that is estimated for an optimal EMS location using historical data can be generalized into the future.

We first shed some light on how the solution of our CACCSP or CAPECSP models can be implemented in practice. Specifically, we focus on the CACCSP model but a similar approach can be used for CAPECSP. Given an optimal solution $(\boldsymbol{x}^*, \boldsymbol{y}^*, \boldsymbol{r}^*, \boldsymbol{z}^*)$ to the CACCSP that employs an empirical distribution based on the realizations $\{\boldsymbol{d}_k\}_{k=1}^N$, one should commit to implementing the plan prescribed by $(\boldsymbol{x}^*, \boldsymbol{y}^*, \boldsymbol{r}^*)$. The peculiarity comes at any time period $\bar{t}$ where a realization $\bar{\boldsymbol{d}}^{\bar{t}}$ occurs and is most likely not part of $\{\boldsymbol{d}_k^{\bar{t}}\}_{k=1}^N$. In order to identify the right assignment to proceed with, one can solve once again the CACCSP model but with $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{r}) = (\boldsymbol{x}^*, \boldsymbol{y}^*, \boldsymbol{r}^*)$ being fixed, and with a new set of empirical realizations $\{\boldsymbol{d}'_k\}_{k=1}^N$ where each $\boldsymbol{d}_k^{t\,\prime} = \boldsymbol{d}_k^t$ except for some randomly chosen $\bar{k}$ for which $\boldsymbol{d}_{\bar{k}}^{\bar{t}\,\prime} = \bar{\boldsymbol{d}}^{\bar{t}}$. One can then implement the new optimal assignment described as $\hat{\boldsymbol{z}}^{\bar{t}}(\bar{\boldsymbol{d}}^{\bar{t}}) := \boldsymbol{z}_{\bar{k}}^{\bar{t}\,**}$. If the out-of-sample subproblems happen to be infeasible, then one can instead simply maximize the coverage. The argument that supports this procedure is that one should consider the new scenario as a scenario that is equally likely as any other anticipated scenario and to verify how this new scenario might have been treated by the CACCSP decision model. Depending on the characteristics of the observed scenario, the modified CACCSP model will decide whether it should be considered as a scenario that is too extreme to cover or not. In what follows, our out-of-sample experiments will confirm that this procedure produces a policy that has similar performance to the performance anticipated by the in-sample version of CACCSP. A more rigourous motivation for using the proposed out-of-sample policy for $\boldsymbol{z}$ is presented in Appendix B.9.

An out-of-sample experiment consists of the following. We first take a random subset of N days from the "training" period of January 2016 to December 2016. The CACCSP (or CAPECSP) model is then solved using the empirical distribution over the N observations. The optimal policy is then implemented (as described above) on every day of the "testing" period spanning from April 2015 to December 2015 (275 days) and statistics are obtained regarding the average empirical reliability and coverage. In the case of the CACCSP, we let $N \in \{50,\ 100,\ 150\}$ while for the CAPECSP model we focus on $N = 100$. Each figure presents statistics calculated over 10 different experiments.



(a) (b)

Figure 6 Out-of-sample empirical reliability for CACCSP in (a) and out-of-sample empirical coverage profile for CAPECSP in (b) in 10 out-of-sample experiments. (a) presents box plots for the reliability achieved by the solutions of CACCSP model with $\beta = 0.95$ and $\eta = 0.05$ for different training set sizes. The red line identifies the target reliability level $1 - \eta = 95\%$. (b) presents the average, minimum and maximum values for empirical coverage achieved by the solutions of CAPECSP model at time $t = 1$ and using a training set of $N = 100$ observations.

We next turn to evaluating the out-of-sample performance of the chance constraints. In this regard, Figure 6(a) presents statistics about the reliability of the coverage for the six time periods and three training set sizes. To be precise, in each experiment, we tested the implementation of the optimal in-sample CACCSP policy on 275 test scenarios and computed the probability of covering $\beta \sum_i d_{i\omega}^t$ for each period on this training set. The box plots indicate the minimum, maximum, mean, and first and third quartiles of these values over the 10 experiments. First, we can observe from this figure that the in-sample policy has a reasonably good out-of-sample performance in terms of reliability. In most experiments and for most time periods, the policy is able to achieve the 95% reliability that it was aiming for, and often reaches 100% reliability ($t = 3, 4, 5$). In the case where this target is not reached we see that most of the time, the violation is below 5%. The figure also exhibits a daily pattern. The policy is typically less reliable in the morning and the evening then during the middle of the day. This might be due to the fact that the model is more flexible in the middle of the day with large ambulance station capacities, and lower relocation and travel costs, which allows the model to more easily satisfy the requests. Finally, we believe it is possible to qualitatively state that as more training samples are used in CACCSP, the empirical reliability starts concentrating more around the 95% target.

Finally, we briefly investigate the out-of-sample performance of the PEC policy. To do so, we focus on the performance achieved for time period $t = 1$ and with a training set of size $N = 100$, although our observations are similar for other time periods. We present in Figure 6(b) statistics about the out-of-sample empirical coverage profile. This is calculated in each experiment using

$$f_{6(b)}(\eta) := \sup \left\{ \beta \;\middle|\; \frac{1}{275} \sum_{\omega'=1}^{275} \mathbf{1}\left\{ \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}_i^z} \hat{z}_{ij}^t(\boldsymbol{d}_{\omega'}^t) \geq \beta \sum_{i\in\mathcal{I}} d_{i\omega'}^t \right\} \geq 1-\eta \right\},$$

where $\{\boldsymbol{d}_{\omega'}\}_{\omega'=1}^{275}$ is the out-of-sample test set and $\hat{\boldsymbol{z}}^t(\boldsymbol{d}_{\omega'}^t)$ is the out-of-sample implementation of the PEC policy as described earlier in this section, and can be compared to the targeted probabilistic envelope $\beta_{0.5}^{cvx}(\eta)$ (red curve) used in CAPECSP. Overall, Figure 6(b) seems to indicate that out-of-sample empirical coverage is of surprisingly good quality when comparing the in-sample performance (green curve), and significantly outperforms the needed reference level (red curve).

9. Concluding Remarks

In this paper, we extend the chance constrained EMS location model presented in Beraldi and Bruni (2009) to a multiperiod setting and propose a novel PEC formulation that allows the EMS manager to control the relative level of EMS coverage achieved under every possible scenarios of emergency demand. This gives rise to two stochastic programming models (CCSP and PECSP). In order to solve instances with realistic sizes, we develop a solution scheme that is based on the B&BC method and some enhancements that are based on valid inequalities and strengthened optimality cuts. We also propose a conservative approximation model that can be solved using B&BC significantly faster. We compare the numerical performance of our solution schemes to the BB method presented in Beraldi and Bruni (2009), and the CPLEX solver. Our results demonstrate that the gain in performance is significant especially for larger-scale instances.

We also presented a case study that exploited data from NIASHSCT, where we are able to illustrate the differences between the solutions obtained from CACCSP and CAPECSP models, the use of a reference coverage profile in the design of probabilistic envelopes, and the recommended procedure for implementation of the delayed assignment decisions. Our out-of-sample analysis confirms that the policies obtained using 100 historical observations perform reasonably well on test data where characteristics of the distribution might have changed due to the passage of time.

We suspect that the models and empirical analysis presented in this paper should benefit a number of other types of applications in the public sector, e.g. fire station location, humanitarian relief location, etc. In particular, we believe that PEC constraints are among the most natural way of encoding the expectations that stakeholders have in terms of reliability when faced against uncertain operating conditions.

Acknowledgments

The authors thank the associate editor and three referees for their valuable comments which helped us improve the quality of this paper. This research was partially supported by the Natural Sciences and Engineering Research Council of Canada [Grant RGPIN-2016-05208], the Canada Research Chair program [Grant 950-230057], Groupe d'études et de recherche en analyse des décisions, and National Natural Science Foundation of China [Grants 71432002, 91746210]. The authors also thank NIASHSCT for providing the data used in case study, and thank Carlos A. Zetina for his support during the implementation of B&BC method.

References

Ahmadi-Javid, A., Seyedi, P., Syam, S. S., 2017. A survey of healthcare facility location. Computers & Operations Research 79, 223–263.

Ansari, S., McLay, L. A., Mayorga, M. E., 2015. A maximum expected covering problem for district design. Transportation Science 51 (1), 376–390.

Arabani, A. B., Farahani, R. Z., 2012. Facility location dynamics: An overview of classifications and applications. Computers & Industrial Engineering 62 (1), 408–420.

Aringhieri, R., Bruni, M. E., Khodaparasti, S., van Essen, J., 2017. Emergency medical services and beyond: Addressing new challenges through a wide literature review. Computers & Operations Research 78, 349–368.

Ball, M. O., Lin, F. L., 1993. A reliability model applied to emergency service vehicle location. Operations Research 41 (1), 18–36.

Başar, A., Çatay, B., Ünlüyurt, T., 2011. A multi-period double coverage approach for locating the emergency medical service stations in istanbul. Journal of the Operational Research Society 62 (4), 627–637.

Bélanger, V., Ruiz, A., Soriano, P., 2018. Recent optimization models and trends in location, relocation, and dispatching of emergency medical vehicles. European Journal of Operational Research.

Benders, J. F., 1962. Partitioning procedures for solving mixed-variables programming problems. Numerische Mathematik 4 (1), 238–252.

Beraldi, P., Bruni, M. E., 2009. A probabilistic model applied to emergency service vehicle location. European Journal of Operational Research 196 (1), 323–331.

Beraldi, P., Bruni, M. E., Conforti, D., 2004. Designing robust emergency medical service via stochastic programming. European Journal of Operational Research 158 (1), 183–193.

Bertsimas, D., Tsitsiklis, J., 1997. Introduction to linear optimization. Athena Scientific.

Boujemaa, R., Jebali, A., Hammami, S., Ruiz, A., Bouchriha, H., 2017. A stochastic approach for designing two-tiered emergency medical service systems. Flexible Services and Manufacturing Journal, 1–30.

Brotcorne, L., Laporte, G., Semet, F., 2003. Ambulance location and relocation models. European Journal of Operational Research 147 (3), 451–463.

Charnes, A., Cooper, W. W., 1959. Chance-constrained programming. Management Science 6 (1), 73–79.

Church, R., Velle, C. R., 1974. The maximal covering location problem. Papers in Regional Science 32 (1), 101–118.

Dalal, J., Üster, H., 2018. Combining worst case and average case considerations in an integrated emergency response network design problem. Transportation Science 152 (1), 171–188.

Daskin, M. S., 1983. A maximum expected covering location model: formulation, properties and heuristic solution. Transportation Science 17 (1), 48–70.

Degel, D., Wiesche, L., Rachuba, S., Werners, B., 2015. Time-dependent ambulance allocation considering data-driven empirically required coverage. Health Care Management Science 18 (4), 444–458.

Dentcheva, D., Ruszczynski, A., 2003. Optimization with stochastic dominance constraints. SIAM Journal on Optimization 14 (2), 548–566.

Dentcheva, D., Ruszczyński, A., 2004. Semi-infinite probabilistic optimization: first-order stochastic dominance constrain. Optimization 53 (5-6), 583–601.

Drapkin, D., Schultz, R., 2010. An algorithm for stochastic programs with first-order dominance constraints induced by linear recourse. Discrete Applied Mathematics 158 (4), 291–297.

Gendreau, M., Laporte, G., Semet, F., 1997. Solving an ambulance location model by tabu search. Location Science 5 (2), 75–88.

Gendreau, M., Laporte, G., Semet, F., 2001. A dynamic model and parallel tabu search heuristic for real-time ambulance relocation. Parallel Computing 27 (12), 1641–1653.

He, L., Hu, Z., Zhang, M., 2018. Robust repositioning for vehicle sharing. working paper.

He, L., Mak, H.-Y., Rong, Y., Shen, Z.-J. M., 2017. Service region design for urban electric vehicle sharing systems. Manufacturing & Service Operations Management 19 (2), 309–327.

Hu, J., Homem-de Mello, T., Mehrotra, S., 2012. Sample average approximation of stochastic dominance constrained programs. Mathematical Programming 133 (1-2), 171–201.

Keeney, R., Raiffa, H., 1976. Decisions with Multiple Objectives: Preferences and Value Trade-Offs. John Wiley & Sons, New York, NY, USA.

Liu, X., Küçükyavuz, S., Luedtke, J., 2016. Decomposition algorithms for two-stage chance-constrained programs. Mathematical Programming 157 (1), 219–243.

Lu, M., Chen, Z., Shen, S., 2017. Optimizing the profitability and quality of service in carshare systems under demand uncertainty. Manufacturing & Service Operations Management 20 (2), 162–180.

Luedtke, J., 2008. New formulations for optimization under stochastic dominance constraints. SIAM Journal on Optimization 19 (3), 1433–1450.

Luedtke, J., 2014. A branch-and-cut decomposition algorithm for solving chance-constrained mathematical programs with finite support. Mathematical Programming 146 (1-2), 219–244.

Luedtke, J., Ahmed, S., 2008. A sample approximation approach for optimization with probabilistic constraints. SIAM Journal on Optimization 19 (2), 674–699.

Luedtke, J., Ahmed, S., Nemhauser, G. L., 2010. An integer programming approach for linear programs with probabilistic constraints. Mathematical Programming 122 (2), 247–272.

Maleki, M., Majlesinasab, N., Sepehri, M. M., 2014. Two new models for redeployment of ambulances. Computers & Industrial Engineering 78, 271–284.

Marianov, V., 2017. Location models for emergency service applications. In: Leading Developments from INFORMS Communities. INFORMS, pp. 237–262.

Martins de Sá, E., Contreras, I., Cordeau, J.-F., Saraiva de Camargo, R., de Miranda, G., 2015. The hub line location problem. Transportation Science 49 (3), 500–518.

Maxwell, M. S., Restrepo, M., Henderson, S. G., Topaloglu, H., 2010. Approximate dynamic programming for ambulance redeployment. INFORMS Journal on Computing 22 (2), 266–281.

Moeini, M., Jemai, Z., Sahin, E., 2015. Location and relocation problems in the context of the emergency medical service systems: a case study. Central European Journal of Operations Research 23 (3), 641–658.

Naoum-Sawaya, J., Elhedhli, S., 2013. A stochastic optimization model for real-time ambulance redeployment. Computers & Operations Research 40 (8), 1972–1978.

Nemirovski, A., Shapiro, A., 2006. Convex approximations of chance constrained programs. SIAM Journal on Optimization 17 (4), 969–996.

Nickel, S., da Gama, F. S., 2015. Multi-period facility location. In: Location Science. Springer, pp. 289–310.

Nickel, S., Reuter-Oppermann, M., Saldanha-da Gama, F., 2016. Ambulance location under stochastic demand: A sampling approach. Operations Research for Health Care 8, 24–32.

Noyan, N., 2010. Alternate risk measures for emergency medical service system design. Annals of Operations Research 181 (1), 559–589.

Noyan, N., Rudolf, G., RuszczyńSki, A., 2006. Relaxations of linear programming problems with first order stochastic dominance constraints. Operations Research Letters 34 (6), 653–659.

Owen, S. H., Daskin, M. S., 1998. Strategic facility location: A review. European Journal of Operational Research 111 (3), 423–447.

Pagnoncelli, B. K., Ahmed, S., Shapiro, A., 2009. Sample average approximation method for chance constrained programming: theory and applications. Journal of Optimization Theory and Applications 142 (2), 399–416.

Paul, N. R., Lunday, B. J., Nurre, S. G., 2017. A multiobjective, maximal conditional covering location problem applied to the relocation of hierarchical emergency response facilities. Omega 66, 147–158.

Rahmaniani, R., Crainic, T. G., Gendreau, M., Rei, W., 2017. The benders decomposition algorithm: A literature review. European Journal of Operational Research 259 (3), 801–817.

Ruszczyński, A., 2002. Probabilistic programming with discrete distributions and precedence constrained knapsack polyhedra. Mathematical Programming 93 (2), 195–215.

Schilling, D. A., 1980. Dynamic location modeling for public-sector facilities: A multicriteria approach. Decision Sciences 11 (4), 714–724.

Schmid, V., 2012. Solving the dynamic ambulance relocation and dispatching problem using approximate dynamic programming. European Journal of Operational Research 219 (3), 611–621.

Schmid, V., Doerner, K. F., 2010. Ambulance location and relocation problems with time-dependent travel times. European Journal of Operational Research 207 (3), 1293–1303.

Schneeberger, K., Doerner, K. F., Kurz, A., Schilde, M., 2016. Ambulance location and relocation models in a crisis. Central European Journal of Operations Research 24 (1), 1–27.

Song, Y., Luedtke, J. R., 2013. Branch-and-cut approaches for chance-constrained formulations of reliable network design problems. Mathematical Programming Computation 5 (4), 397–432.

Song, Y., Luedtke, J. R., Küçükyavuz, S., 2014. Chance-constrained binary packing problems. INFORMS Journal on Computing 26 (4), 735–747.

Toregas, C., Swain, R., ReVelle, C., Bergman, L., 1971. The location of emergency service facilities. Operations Research 19 (6), 1363–1373.

van den Berg, P. L., Aardal, K., 2015. Time-dependent MEXCLP with start-up and relocation cost. European Journal of Operational Research 242 (2), 383–389.

Van Roy, T. J., Erlenkotter, D., 1982. A dual-based procedure for dynamic facility location. Management Science 28 (10), 1091–1105.

Wang, Q., Guan, Y., Wang, J., 2011. A chance-constrained two-stage stochastic program for unit commitment with uncertain wind power output. IEEE Transactions on Power Systems 27 (1), 206–215.

Xu, H., Caramanis, C., Mannor, S., 2012. Optimization under probabilistic envelope constraints. Operations Research 60 (3), 682–699.

Zhang, Z.-H., Li, K., 2015. A novel probabilistic formulation for locating and sizing emergency medical service stations. Annals of Operations Research 229 (1), 813–835.

Appendix A: Proof of Propositions

A.1. Proof of Proposition 1

The first step of this proof consists in demonstrating that constraint (5) is equivalent to constraint (7b) in CCSP. This can be seen from the fact that $\boldsymbol{z}$ is integer in CCSP, hence any assignment that satisfies (5) must have $\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_i^z} z_{ij}^t \geq \lceil \beta^t(1-\rho_\omega^t)\sum_{i\in\mathcal{I}} d_{i\omega}^t\rceil = (1-\rho_\omega^t)\lceil\beta^t\sum_{i\in\mathcal{I}} d_{i\omega}^t\rceil$ since ρ_ω^t is binary. The reverse is more straightforward, namely since constraint (5) is a relaxation of (7b).

The next step is to demonstrate that by relaxing the integrality constraint on $\boldsymbol{r}$ and $\boldsymbol{z}$, we still produce optimal integer solutions. Let $(\boldsymbol{x}^*, \boldsymbol{y}^*, \boldsymbol{\rho}^*, \boldsymbol{r}^*, \boldsymbol{z}^*)$ be any optimal solution of CCSP2, we will show that there exists integer valued assignments $\bar{\boldsymbol{r}}^*$ and $\bar{\boldsymbol{z}}^*$ for which $(\boldsymbol{x}^*, \boldsymbol{y}^*, \boldsymbol{\rho}^*, \bar{\boldsymbol{r}}^*, \bar{\boldsymbol{z}}^*)$ is also necessarily optimal for CCSP2. Since CCSP2 is a relaxation of CCSP, we will therefore conclude that both problems are equivalent.

We focus on the case of $\bar{\boldsymbol{z}}^*$ since the case of $\bar{\boldsymbol{r}}^*$ is similar. Specifically, for any fixed time period t and scenario ω, we know that any optimal solution to the following problem produces equivalent optimal solutions for CCSP2:

$$\begin{aligned}
\underset{\boldsymbol{z}^t_\omega \geq 0}{\text{minimize}} \quad & \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} p_\omega\, c\, l_{ij} z^t_{ij\omega} && && (24a)\\
\text{subject to} \quad & \sum_{i\in\mathcal{I}} z^t_{ij\omega} \leq \lambda^t_j y^{t\,*}_j && \forall j\in\mathcal{J} && (24b)\\
& \sum_{j\in\mathcal{J}^z_i} z^t_{ij\omega} \leq d^t_{i\omega} && \forall i\in\mathcal{I} && (24c)\\
& \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}^z_i} z^t_{ij\omega} \geq \lceil \beta^t \sum_{i\in\mathcal{I}} d^t_{i\omega}\rceil (1-\rho^{t\,*}_\omega)\,. && && (24d)
\end{aligned}$$

One can actually show that the above problem can be reformulated as the following minimum cost network flow problem (MCNFP).

$$\begin{aligned}
\underset{\boldsymbol{z}^t_\omega \geq 0, \boldsymbol{w}\geq 0, \boldsymbol{s}\geq 0, q\geq 0}{\text{minimize}} \quad & \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} p_\omega\, c\, l_{ij} z^t_{ij\omega} && \\
\text{subject to} \quad & \sum_{i\in\mathcal{I}} z^t_{ij\omega} + s_j = \lambda^t_j y^{t\,*}_j && \forall j\in\mathcal{J}\\
& w_i - \sum_{j\in\mathcal{J}^z_i} z^t_{ij\omega} = 0 && \forall i\in\mathcal{I}\\
& q - \sum_{i\in\mathcal{I}} w_i = -\lceil \beta^t \sum_{i\in\mathcal{I}} d^t_{i\omega}\rceil (1-\rho^{t\,*}_\omega) && \\
& -q - \sum_{i\in\mathcal{I}} s_i = \lceil \beta^t \sum_{i\in\mathcal{I}} d^t_{i\omega}\rceil (1-\rho^{t\,*}_\omega) - \sum_{j\in\mathcal{J}} \lambda^t_j y^{t\,*}_j && \\
& w_i \leq d_i && \forall i\in\mathcal{I}\,,
\end{aligned}$$

where $\boldsymbol{s}\in\mathbb{R}^{|\mathcal{J}|}$, $\boldsymbol{w}\in\mathbb{R}^{|\mathcal{I}|}$, and $q\in\mathbb{R}$. Based on the integrality theorem for this family of problem, it is known (see Bertsimas and Tsitsiklis (1997)) that there necessarily exists an integer valued optimal solution in contexts where all parameters in the set of constraints of the MCNFP are integers. This is necessarily the case for the instances that arise using CCSP2 since we assume that λ^t_j is an integer, and since any feasible $\boldsymbol{y}$ and $\boldsymbol{\rho}$ also are. We can therefore conclude that there is an integer assignment for $\bar{\boldsymbol{z}}^*$ that can replace the original optimal assignment while preserving the optimality of $(\boldsymbol{x}^*, \boldsymbol{y}^*, \boldsymbol{\rho}^*, \boldsymbol{r}^*, \bar{\boldsymbol{z}}^*)$ with respect to CCSP.

A similar argument can be used for the case of $\boldsymbol{r}$ where the MCNFLP more naturally takes the form of:

$$\begin{aligned}
\underset{\boldsymbol{r}\geq 0}{\text{minimize}} \quad & \sum_{t=1}^{T}\sum_{m\in\mathcal{J}}\sum_{j\in\mathcal{J}} \alpha^t r^t_{mj} && && (25a)\\
\text{subject to} \quad & \sum_{m\in\mathcal{J}^r_j} r^t_{mj} - \sum_{m\in\mathcal{J}^r_j} r^t_{jm} = y^{t+1}_j - y^t_j && \forall j\in\mathcal{J}, t\in\{1,\ldots,T-1\} && (25b)\\
& \sum_{m\in\mathcal{J}^r_j} r^T_{mj} - \sum_{m\in\mathcal{J}^r_j} r^T_{jm} = y^1_j - y^T_j && \forall j\in\mathcal{J}\,. && (25c)
\end{aligned}$$

To obtain the integer solutions for $\boldsymbol{z}^*$ and $\boldsymbol{r}^*$ once the first stage decisions $\boldsymbol{x}^*$, $\boldsymbol{y}^*$ and $\boldsymbol{\rho}^*$ are known, one can simply solve (24) for each ω and t and (25) with integer constraints on $\boldsymbol{z}$ and $\boldsymbol{r}$. This completes our proof. □

A.2. Proof of Proposition 2

This proof is divided into two parts. We first show that CCSP2 is equivalent to CCSP2'. We then confirm that solutions that satisfy constraints (8d) and (8e) give rise to versions of the CCSP-SP^r and CCSP-SP$^z_{\omega,t}$ models that are necessarily feasible and bounded.

In demonstrating the equivalence between CCSP2 and CCSP2', we start by remarking that CCSP2' follows from exploiting an epigraph reformulation of the operations $\sum_{j\in\mathcal{J}}\sum_{i\in\mathcal{I}} p_\omega c^t l_{ij} z^t_{ij\omega}$ and $\sum_{t=1}^{T}\sum_{m\in\mathcal{J}}\sum_{j\in\mathcal{J}} \alpha^t r^t_{mj}$ in the objective function, and adding constraints (8d) and (8e) to the model. We therefore only need to show that the latter two constraints are redundant in CCSP2. First, we can obtain constraint (8e) directly from the constraints (4b) and (7b):

$$\sum_{j\in\mathcal{J}} \lambda^t_j y^t_j \geq \sum_{j\in\mathcal{J}}\sum_{i\in\mathcal{I}} z^t_{ij\omega} \geq (1-\rho^t_\omega)\lceil \beta^t \sum_{i\in\mathcal{I}} d^t_{i\omega}\rceil\,,\ \forall t\in\{1,\ldots,T\}, \forall \omega\in\Omega\,.$$

Moreover, constraints (1d) and (1e) imply that

$$\sum_{j\in\mathcal{J}} y^{t+1}_j = \sum_{j\in\mathcal{J}}\left(y^t_j + \sum_{m\in\mathcal{J}^r_j} r^t_{mj} - \sum_{m\in\mathcal{J}^r_j} r^t_{jm}\right) = \sum_{j\in\mathcal{J}} y^t_j\,,\ \forall t\in\{1,\ldots,T-1\}\,.$$

We are left with showing that, under Assumption 1, both CCSP-SP^r and CCSP-SP$^z_{\omega,t}$ are always feasible and bounded when constraints (8d) and (8e) are satisfied.

In the case of CCSP-SP^r, boundedness follows from the fact that $\boldsymbol{r}\geq 0$ and that $\boldsymbol{\alpha}\geq\boldsymbol{0}$. As for the feasibility of this slave problem, this follows from the fact that $\sum_{j\in\mathcal{J}} y^t_j = \sum_{j\in\mathcal{J}} y^{t+1}_j$ for all t, and that we assumed vehicles can be transferred from any location to any other location in a single period.

In the case of CCSP-SP$^z_{\omega,t}$, boundedness again follows from the non-negativity of $\boldsymbol{z}^t_\omega$ and monotonicity of the objective function. As for feasibility, since we know that $\sum_{j\in\mathcal{J}}\lambda^t_j y^t_j \geq (1-\rho^t_\omega)\lceil\beta^t\sum_{i\in\mathcal{I}} d^t_{i\omega}\rceil$, we know that in this transportation problem there is enough offer to cover the demand. Since every source $i\in\mathcal{I}$ can serve every demand node $j\in\mathcal{J}$, we conclude that a feasible assignment can necessarily be found.

This completes our proof. □

A.3. Proof of Proposition 3

For fixed t, let the ordered scenario set $\Omega'_t := \{\omega'_k\}_{k=1}^N$ be such that for all $k<k'$ we have that $\sum_{i\in\mathcal{I}} d^t_{i\omega'_k} \leq \sum_{i\in\mathcal{I}} d^t_{i\omega'_{k'}}$. One can simply rewrite constraints (6) and (8e) as

$$\sum_{\omega'\in\Omega'_t} p_{\omega'}\rho^t_{\omega'} \leq \eta^t \qquad \forall t\in\{1,\ldots,T\}. \tag{26}$$

$$\sum_{j\in\mathcal{J}} \lambda^t_j y^t_j \geq \lceil \beta^t \sum_{i\in\mathcal{I}} d^t_{i\omega'}\rceil (1-\rho^t_{\omega'}) \qquad \forall \omega'\in\Omega'_t,\, t\in\{1,\ldots,T\} \tag{27}$$

We will prove our claim by contradiction. In order to do so, let $(\boldsymbol{y},\boldsymbol{\rho})$ satisfy constraints (27) and (26) yet violate constraint (11), namely

$$\sum_{j\in\mathcal{J}} \lambda^t_j y^t_j < \beta^t \sum_{i\in\mathcal{I}} d^t_{i\omega'_k} \qquad \forall k\geq \bar{k}_t\,,$$

given that $\boldsymbol{y}$ is integer valued. Since constraint (27) holds, then it must be that $\rho^t_{\omega_k} = 1$ for all $k \geq \bar{k}_t$. This in turn can be used to show that

$$\eta^t \geq \sum_{\omega' \in \Omega'_t} p_{\omega'} \rho^t_{\omega'} \geq \sum_{k: k \geq \bar{k}_t} p_{\omega'_k} \rho^t_{\omega'_k} = \sum_{k=\bar{k}_t}^{N} p_{\omega'_k} > \eta^t \,.$$

Since this necessarily leads to a contradiction $\eta^t > \eta^t$, we must conclude that any solution that satisfies (26) and (27) must also satisfy constraint (11). □

A.4. Proof of Proposition 4

This result is obtained by remarking that for any $t \in \{1, \ldots, T\}$ and any $k \in \{0, \ldots, N-1\}$ the constraint

$$\sum_{\omega \in \Omega} (1/N) \mathbf{1} \left\{ \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}_i^z} z^t_{ij\omega} \geq \beta^t(\eta) \sum_{i \in \mathcal{I}} d^t_{i\omega} \right\} \geq 1 - \eta \quad \forall \eta \in [k/N, (k+1)/N)$$

is equivalent to

$$\sum_{\omega \in \Omega} (1/N) \mathbf{1} \left\{ \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}_i^z} z^t_{ij\omega} \geq \beta^t(\eta) \sum_{i \in \mathcal{I}} d^t_{i\omega} \right\} \geq 1 - k/N \quad \forall \eta \in [k/N, (k+1)/N) \,.$$

Since the *sum* operator can only reach values in $\{j/N\}_{j \in \{0, \ldots, N\}}$. Hence, the constraint can be reformulated equivalently as

$$\inf_{\eta \in [k/N, (k+1)/N)} \sum_{\omega \in \Omega} (1/N) \mathbf{1} \left\{ \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}_i^z} z^t_{ij\omega} \geq \beta^t(\eta) \sum_{i \in \mathcal{I}} d^t_{i\omega} \right\} \geq 1 - k/N \,,$$

yet since $\mathbf{1} \left\{ \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}_i^z} z^t_{ij\omega} \geq \beta \sum_{i \in \mathcal{I}} d^t_{i\omega} \right\}$ is left continuous and decreasing with respect to β, it is necessary and sufficient to verify that

$$\sum_{\omega \in \Omega} (1/N) \mathbf{1} \left\{ \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}_i^z} z^t_{ij\omega} \geq \sup_{\eta \in [k/N, (k+1)/N)} \beta^t(\eta) \sum_{i \in \mathcal{I}} d^t_{i\omega} \right\} \geq 1 - k/N \,.$$

This constraint is exactly equivalent to

$$\sum_{\omega \in \Omega} (1/N) \mathbf{1} \left\{ \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}_i^z} z^t_{ij\omega} \geq \beta^t_k \sum_{i \in \mathcal{I}} d^t_{i\omega} \right\} \geq 1 - k/N \,,$$

which can then be reformulated using binary variables $\boldsymbol{\rho}$. We thus conclude that constraint (12) is equivalent to the condition that there exists $\boldsymbol{\rho} \in \{0, 1\}^{N \times N \times T}$ that satisfies

$$\sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}_i^z} z^t_{ij\omega} \geq \beta^t_k (1 - \rho^t_{k\omega}) \sum_{i \in \mathcal{I}} d^t_{i\omega} \qquad \forall t \in \{1, \ldots, T\}, \omega \in \Omega, k \in \{0, \ldots, N-1\} \tag{28}$$

$$(1/N) \sum_{\omega \in \Omega} \rho^t_{k\omega} \leq k/N \qquad \forall t \in \{1, \ldots, T\}, k \in \{0, \ldots, N-1\} \,. \tag{29}$$

The next step of this proof consists of showing that the inequality in constraint (29) can be replaced with an equality without affecting the feasibility of $\boldsymbol{z}$. In particular, for any feasible $\boldsymbol{z}$, given some $\bar{\boldsymbol{\rho}}$ that confirms the feasibility of $\boldsymbol{z}$ using (28) and (29), if $(1/N) \sum_{\omega \in \Omega} \bar{\rho}^t_{k\omega} < k/N$ for some k and some t then it must be that $(1/N) \sum_{\omega \in \Omega} \bar{\rho}^t_{k\omega} \leq (k-1)/N$ and that there is a $\bar{\omega}$ such that $\bar{\rho}^t_{k\bar{\omega}} = 0$, otherwise $(1/N) \sum_{\omega \in \Omega} \bar{\rho}^t_{k\omega} = 1 >$

$(N-1)/N \geq k/N$. One can actually show that $\hat{\boldsymbol{\rho}}$ constructed by copying $\bar{\boldsymbol{\rho}}$ at every index except for $\hat{\rho}^t_{k\bar{\omega}} := 1$ will also satisfy both constraints. Namely,

$$\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_i^z} z^t_{ij\bar{\omega}} \geq \beta^t_k(1-\bar{\rho}^t_{k\bar{\omega}})\sum_{i\in\mathcal{I}} d^t_{i\bar{\omega}} \geq \beta^t_k(1-\hat{\rho}^t_{k\bar{\omega}}+1)\sum_{i\in\mathcal{I}} d^t_{i\bar{\omega}} \geq \beta^t_k(1-\hat{\rho}^t_{k\bar{\omega}})\sum_{i\in\mathcal{I}} d^t_{i\bar{\omega}}$$
$$(1/N)\sum_{\omega\in\Omega}\hat{\rho}^t_{k\omega} = (1/N)\sum_{\omega\in\Omega}\bar{\rho}^t_{k\omega} + (1/N) \leq (k-1)/N + 1/N \leq k/N\,.$$

Based on this argument, we can conclude that there always exists a $\bar{\boldsymbol{\rho}}$ that confirms feasibility of $\boldsymbol{z}$ based on constraints (28) and (29) and also satisfies constraint (14).

The third step of this proof is to show that constraint (15) can also be imposed without affecting the feasibility of $\boldsymbol{z}$. In particular, for any feasible $\boldsymbol{z}$, given some $\bar{\rho}$ that confirms the feasibility of $\boldsymbol{z}$ using (28) and (14), if $\bar{\rho}^t_{k\omega} > \bar{\rho}^t_{k+1,\omega}$ for some $\bar{t}$, $\bar{\omega}$, and $\bar{k}$, then one can simply construct a $\hat{\boldsymbol{\rho}}$ that mimics $\bar{\boldsymbol{\rho}}$ except for $\hat{\rho}^{\bar{t}}_{\bar{k}\bar{\omega}} := \bar{\rho}^{\bar{t}}_{\bar{k}+1,\bar{\omega}} = 0$ and for $\hat{\rho}^{\bar{t}}_{\bar{k}\bar{\omega}'} := 1$, where $\bar{\omega}'$ is any scenario for which $\bar{\rho}^{\bar{t}}_{\bar{k}\bar{\omega}'} = 0 < \bar{\rho}^{\bar{t}}_{\bar{k}+1,\bar{\omega}'}$. One can verify that such an index $\bar{\omega}'$ always exists because of constraint (14). One can as a last step confirm that all constraints are satisfied for $\hat{\boldsymbol{\rho}}$:

$$\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_i^z} z^{\bar{t}}_{ij\bar{\omega}} \geq \bar{\beta}^t_{\bar{k}+1}(1-\bar{\rho}^{\bar{t}}_{\bar{k}+1,\bar{\omega}})\sum_{i\in\mathcal{I}} d^{\bar{t}}_{i\bar{\omega}} \geq \bar{\beta}^t_{\bar{k}}(1-\bar{\rho}^{\bar{t}}_{\bar{k}+1,\bar{\omega}})\sum_{i\in\mathcal{I}} d^{\bar{t}}_{i\bar{\omega}} = \bar{\beta}^t_{\bar{k}}(1-\hat{\rho}^{\bar{t}}_{\bar{k},\bar{\omega}})\sum_{i\in\mathcal{I}} d^{\bar{t}}_{i\bar{\omega}}$$
$$\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_i^z} z^{\bar{t}}_{ij\bar{\omega}'} \geq \bar{\beta}^t_{\bar{k}}(1-\bar{\rho}^{\bar{t}}_{\bar{k},\bar{\omega}'})\sum_{i\in\mathcal{I}} d^{\bar{t}}_{i\bar{\omega}'} \geq \bar{\beta}^t_{\bar{k}}(1-\hat{\rho}^{\bar{t}}_{\bar{k},\bar{\omega}'})\sum_{i\in\mathcal{I}} d^{\bar{t}}_{i\bar{\omega}'}$$
$$\sum_{\omega\in\Omega}\hat{\rho}^{\bar{t}}_{\bar{k}\omega} = \sum_{\omega\in\Omega}\bar{\rho}^{\bar{t}}_{\bar{k}\omega} - \bar{\rho}^{\bar{t}}_{\bar{k}\bar{\omega}} + \hat{\rho}^{\bar{t}}_{\bar{k}\bar{\omega}} - \bar{\rho}^{\bar{t}}_{\bar{k}\bar{\omega}'} + \hat{\rho}^{\bar{t}}_{\bar{k}\bar{\omega}'} = \sum_{\omega\in\Omega}\bar{\rho}^{\bar{t}}_{\bar{k}\omega} - 1 + 0 - 0 + 1 = k$$
$$\hat{\rho}^{\bar{t}}_{\bar{k}\bar{\omega}} = 0 \leq \hat{\rho}^{\bar{t}}_{\bar{k}+1,\bar{\omega}}$$
$$\hat{\rho}^{\bar{t}}_{\bar{k}\bar{\omega}'} = 1 \leq 1 = \hat{\rho}^{\bar{t}}_{\bar{k}+1,\bar{\omega}'}\,.$$

This allows us to conclude that there always exists a $\bar{\boldsymbol{\rho}}$ that confirms feasibility of $\boldsymbol{z}$ based on constraints (28) and (14) and also satisfies constraint (15).

The final step consists in demonstrating that for all $\boldsymbol{\rho}$ that satisfies constraint (15), constraint (28) is equivalent to constraint (13). This can be shown by arguing that for all t and all ω:

$$\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_i^z} z^t_{ij\omega} \geq \beta^t_k(1-\rho^t_{k\omega})\sum_{i\in\mathcal{I}} d^t_{i\omega}\;\forall\, k\in\{0,\ldots,N-1\}$$
$$\Leftrightarrow \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_i^z} z^t_{ij\omega} \geq \left(\max_{k\in\{0,\ldots,N-1\}}\beta^t_k(1-\rho^t_{k\omega})\right)\sum_{i\in\mathcal{I}} d^t_{i\omega}$$
$$\Leftrightarrow \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_i^z} z^t_{ij\omega} \geq (\max_{k:\rho^t_{k\omega}=0}\beta^t_k)\sum_{i\in\mathcal{I}} d^t_{i\omega}$$
$$\Leftrightarrow \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_i^z} z^t_{ij\omega} \geq \left(\beta^t_{N-1}(1-\rho^t_{N-1,\omega}) + \sum_{k\in\{0,\ldots,N-2\}}\beta^t_k(\rho^t_{k+1,\omega}-\rho^t_{k\omega})\right)\sum_{i\in\mathcal{I}} d^t_{i\omega}\,,$$

where the second equivalence follows from the fact that β^t_k is a non-decreasing sequence, and the third equivalence follows from the fact that $\boldsymbol{\rho}$ satisfies constraint (15).

This completes the proof. □

A.5. Proof of Proposition 5

First, constraint (20) can be relaxed by dropping the rounding operation in the right-hand side thus obtaining

$$\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_i^z} z^t_{ij\omega'_{N-k}} \geq \beta^t_k \sum_{i\in\mathcal{I}} d^t_{i\omega'_{N-k}} \qquad \forall t\in\{1,\ldots,T\}, k\in\{0,\ldots,N-1\}\,. \tag{30}$$

Then, based on the definition of Ω'_t, we have that for all $k<k'$, $\sum_{i\in\mathcal{I}} d^t_{i\omega'_k} \leq \sum_{i\in\mathcal{I}} d^t_{i\omega'_{k'}}$. This allows us to state that any assignment for $\boldsymbol{z}$ that satisfies constraint (30), must also satisfy

$$\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_i^z} z^t_{ij\omega'_{N-k}} \geq \beta^t_k \sum_{i\in\mathcal{I}} d^t_{i\omega'_{k'}} \qquad \forall t\in\{1,\ldots,T\}, 1\leq k'\leq N-k, k\in\{0,\ldots,N-1\}\,. \tag{31}$$

Hence, we must also have that

$$\frac{1}{N}\sum_{\omega\in\Omega}\mathbf{1}\left\{\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_i^z} z^t_{ij\omega} \geq \beta^t_k \sum_{i\in\mathcal{I}} d^t_{i\omega}\right\} \geq \sum_{k'=1}^{N-k} 1/N = 1-\frac{k}{N}. \tag{32}$$

Yet, we already showed in the proof of Proposition 4 (see Appendix A.4) that under Assumption 2, constraint (32) and constraint (12) are equivalent.

This completes our proof. □

Appendix B: Additional Supplemental Material

B.1. Variations of facility operation cost in problem (1)

It is worth observing that there are a number of structures that have been proposed in the literature for modeling the cost of opening, operating, and closing facilities such as ambulance bases and conditions on their use. In the most general version of the DM formulation, one would replace Constraint (1c) with the following ones:

$$\begin{aligned} w^t_j &\geq x^t_j - x^{t-1}_j\,,\ \forall j\in\mathcal{J}, t\in\{1,\ldots,T\} \\ v^t_j &\geq x^{t-1}_j - x^t_j\,,\ \forall j\in\mathcal{J}, t\in\{1,\ldots,T\} \\ x^T_j &= x^0_j\,,\ \forall j\in\mathcal{J}\,, \end{aligned}$$

where w^t_j and v^t_j are additional binary decision variables modeling the decisions to open or close the base location j at time period t respectively, and where x^0_j is a binary variable that is used to ensure that the plan is consistent from one day to the other and accounts for opening (and closing) decisions at the first period of the day. One could include fixed opening or closing cost in the objective function or constraints such as:

$$v^{t'}_j \leq 1 - w^t_j\,,\ \forall t'\in\{\mathrm{mod}(t,T),\mathrm{mod}(t+1,T),\mathrm{mod}(t+\Delta,T)\}\,,\ \forall t\in\{1,\ldots,T\}\,,$$

which control the minimum amount of time Δ during which a base might need to stay open. We refer interested readers to the following literature for example of formulation and ways of reformulating such constraints to obtain tighter linear relaxations, i.e. Owen and Daskin (1998), Arabani and Farahani (2012), and Nickel and da Gama (2015).

B.2. Additional Details Regarding the Benders Decomposition of CCSP

In this appendix, we describe in more the details the application of Benders decomposition to the resolution of CCSP2'. We start by describing how both $h^r(\boldsymbol{y})$ and $h^z_{\omega,t}(\bar{\boldsymbol{y}},\bar{\boldsymbol{\rho}})$ can be represented as the supremum over a set of hyperplanes. We then describe the a relaxation of CCSP2', referred as the MP, and how upper and lower bounds can be obtained for CCSP2' using optimal solutions of this problem.

Focusing first on $h^r(\boldsymbol{y})$, based on linear programming duality theory, we know that, given any fixed $\bar{\boldsymbol{y}}$ that satisfies constraints (8d), a supporting hyperplane at $\bar{\boldsymbol{y}}$ can be identified by solving the dual problem associated to CCSP-SP^r.

$$[\text{CCSP-DSP}^r]\ h^r(\bar{\boldsymbol{y}}) := \max_{\boldsymbol{\mu}^r\geq 0} \quad \sum_{t=1}^{T-1}\sum_{j\in\mathcal{J}}\mu^r_{jt}(\bar{y}^t_j-\bar{y}^{t+1}_j)+\sum_{j\in\mathcal{J}}\mu^r_{jT}(\bar{y}^T_j-\bar{y}^1_j) \tag{33a}$$

$$\text{subject to} \quad \alpha^t+\mu^r_{jt}-\mu^r_{mt}\geq 0 \qquad \forall m\in\mathcal{J}^r_j, j\in\mathcal{J}, t\in\{1,\ldots,T\}, \tag{33b}$$

where $\{\mu^r_{jt}\}_{j\in\mathcal{J},t\in\{1,\ldots,T\}}$ are the dual variables associated with constraints (9b) and (9c). Strong duality necessarily applies for CCSP-SP^r under Assumption 1, since Proposition 2 guarantees that this problem is feasible. Furthermore, since it is also bounded, the existence of an optimal solution for CCSP-DSP^r is guaranteed. A supporting hyperplane at $\bar{\boldsymbol{y}}$ therefore necessarily takes the form:

$$\sum_{t=1}^{T-1}\sum_{j\in\mathcal{J}}\bar{\mu}^{r*}_{jt}(y^t_j-y^{t+1}_j)+\sum_{j\in\mathcal{J}}\bar{\mu}^{r*}_{jT}(y^T_j-y^1_j),$$

where $\bar{\boldsymbol{\mu}}^{r*}$ is an optimal solution to CCSP-DSP^r.

Similarly, in the case of CCSP-SP$^z_{\omega,t}$, for a given solution $(\bar{\boldsymbol{y}},\bar{\boldsymbol{\rho}})$ that satisfies (8e), the dual problem takes the form:

$$[\text{CCSP-DSP}^z_{\omega,t}]\ h^z_{\omega,t}(\bar{\boldsymbol{y}},\bar{\boldsymbol{\rho}}) = \max_{\boldsymbol{\mu}^1_{t\omega},\boldsymbol{\mu}^2_{t\omega},\mu^3_{t\omega}} \ -\sum_{j\in\mathcal{J}}\lambda^t_j\bar{y}^t_j\mu^1_{jt\omega}-\sum_{i\in\mathcal{I}}d^t_{i\omega}\mu^2_{it\omega}+\mu^3_{t\omega}(1-\bar{\rho}^t_\omega)\lceil\beta^t\sum_{i\in\mathcal{I}}d^t_{i\omega}\rceil \tag{34a}$$

$$\text{subject to} \ -\mu^1_{jt\omega}-\mu^2_{it\omega}+\mu^3_{t\omega}\leq p_\omega c^t l_{ij} \quad \forall i\in\mathcal{I}, j\in\mathcal{J}^z_i \tag{34b}$$

$$\boldsymbol{\mu}^1_{t\omega}\geq 0, \boldsymbol{\mu}^2_{t\omega}\geq 0, \mu^3_{t\omega}\geq 0, \tag{34c}$$

where $\boldsymbol{\mu}^1_{t\omega}\in\mathbb{R}^{|\mathcal{J}|}$, $\boldsymbol{\mu}^2_{t\omega}\in\mathbb{R}^{|\mathcal{I}|}$, $\mu^3_{t\omega}\in\mathbb{R}$ are the dual vectors for constraints (10b)-(10d) respectively. Hence, a supporting hyperplane can be obtained in the form:

$$-\sum_{j\in\mathcal{J}}\lambda^t_j y^t_j\bar{\mu}^{1*}_{jt\omega}-\sum_{i\in\mathcal{I}}d^t_{i\omega}\bar{\mu}^{2*}_{it\omega}+\bar{\mu}^{3*}_{t\omega}(1-\rho^t_\omega)\lceil\beta^t\sum_{i\in\mathcal{I}}d^t_{i\omega}\rceil,$$

where $\bar{\boldsymbol{\mu}}^{1*}$, $\bar{\boldsymbol{\mu}}^{2*}$, and $\bar{\boldsymbol{\mu}}^{3*}$ are optimal solutions of CCSP-DSP$^z_{t,\omega}$.

Given a subset of supporting hyperplanes for CCSP-SP^r and CCSP-SP$^z_{\omega,t}$, which we will represent using $\{\bar{\boldsymbol{\mu}}^{r\tau}\}_{\tau\in G^r}$ and $\{(\bar{\boldsymbol{\mu}}^{1\tau}_{t\omega},\bar{\boldsymbol{\mu}}^{2\tau}_{t\omega},\bar{\mu}^{3\tau}_{t\omega})\}_{\tau\in G^z_{t\omega}}$ for some index sets G^r and $G^z_{t\omega}$, one can formulate a MP that returns a triplet $(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\rho})$, which is feasible in CCSP2' and which optimal value provides a lower bound for CCSP2':

$$\underset{\boldsymbol{x},\boldsymbol{y},\boldsymbol{\rho},\theta^r,\boldsymbol{\theta}^z}{\text{minimize}} \sum_{t=1}^{T}\sum_{j\in\mathcal{J}}f^t_jx^t_j+\sum_{t=1}^{T}\sum_{j\in\mathcal{J}}g^t_jy^t_j+\sum_{t=1}^{T}\sum_{\omega\in\Omega}\theta^z_{t\omega}+\theta^r \tag{35a}$$

subject to (1b), (1c), (6), (7e), (8d), (8e)

$$\theta^r\geq\sum_{j\in\mathcal{J}}\sum_{t=1,\ldots,T-1}\bar{\mu}^{r\tau}_{jt}(y^t_j-y^{t+1}_j)+\sum_{j\in\mathcal{J}}\bar{\mu}^{r\tau}_{jT}(y^T_j-y^1_j),\ \forall\tau\in G^r \tag{35b}$$

$$\theta^z_{t\omega}\geq-\sum_{j\in\mathcal{J}}\lambda^t_jy^t_j\bar{\mu}^{1\tau}_{t\omega j}-\sum_{i\in\mathcal{I}}d^t_{i\omega}\bar{\mu}^{2\tau}_{t\omega i}+\bar{\mu}^{3\tau}_{t\omega}(1-\rho^t_\omega)\lceil\beta^t\sum_{i\in\mathcal{I}}d^t_{i\omega}\rceil,\ \forall\tau\in G^z_{t\omega}, t\in\{1,\ldots,T\},\omega\in\Omega. \tag{35c}$$

Furthermore, for any feasible solution $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}, \bar{\boldsymbol{\rho}}, \bar{\theta}^r, \bar{\boldsymbol{\theta}}^z)$ of (35), one can obtain an upper bound for CCSP2' through,

$$uobj = \sum_{t=1}^{T}\sum_{j\in\mathcal{J}} f_j^t \bar{x}_j^t + \sum_{t=1}^{T}\sum_{j\in\mathcal{J}} g_j^t \bar{y}_j^t + \sum_{t=1}^{T}\sum_{\omega\in\Omega} h_{\omega,t}^z(\bar{\boldsymbol{y}}, \bar{\boldsymbol{\rho}}) + h^r(\bar{\boldsymbol{y}}). \tag{36}$$

We present the detailed procedure of the BD implementation in Algorithm 1. Note that the iteration count is denoted by v. It is well known that Algorithm 1 always terminates in a finite number of iterations (even when $stoptime = \infty$) as long as one makes sure that each $\bar{\boldsymbol{u}}^{r*}$ and $(\bar{\boldsymbol{u}}_{t\omega}^{1*}, \bar{\boldsymbol{u}}_{t\omega}^{2*}, \bar{u}_{t\omega}^{3*})$ are vertices of the feasible set described in each CCSP-DSP problem. This is due to the fact that the number of vertices of each of these feasible sets is finite so that in the worst-case, one will eventually add all of them to problem (35) and finally obtain that $lobj^v = uobj^v$.

Algorithm 1 Benders Decomposition Implementation

Input A tolerance $\epsilon \geq 0$ and maximum run time *stoptime*

Initialize $v = 0, LB = -\infty, UB = +\infty, \theta^r = \theta_{t\omega}^z = 0, G^r = G_{t\omega}^z = \emptyset$ for all t and ω.

while ($runtime \leq stoptime$ and $UB - LB > \epsilon$) **do**

  Set $v = v + 1$, solve the MP.

  Record optimal solution $(\boldsymbol{x}^v, \boldsymbol{y}^v, \boldsymbol{\rho}^v, \theta^{rv}, \boldsymbol{\theta}^{zv})$ and optimal objective $lobj^v$.

  Update $LB =: lobj^v$.

  Fix $\bar{\boldsymbol{x}} := \boldsymbol{x}^v$, $\bar{\boldsymbol{y}} := \boldsymbol{y}^v$, and $\bar{\boldsymbol{\rho}} := \boldsymbol{\rho}^v$ and solve CCSP-DSP^r and all CCSP-DSP$_{\omega,t}^z$.

  Obtain $h_{t\omega}^z(\bar{\boldsymbol{y}}, \bar{\boldsymbol{\rho}})$ and $h^r(\bar{\boldsymbol{y}})$ together with $(\bar{\boldsymbol{u}}^{r*}, \bar{\boldsymbol{u}}^{1*}, \bar{\boldsymbol{u}}^{2*}, \bar{u}^{3*})$.

  Add $\bar{\boldsymbol{\mu}}^{r*}$ to G^r, and, for all $t \in \{1, \ldots, T\}$ and $\omega \in \Omega$, add $(\bar{\boldsymbol{\mu}}_{t\omega}^1, \bar{\boldsymbol{\mu}}_{t\omega}^{2*}, \bar{\mu}_{t\omega}^{3*})$ to $G_{t\omega}^z$

  Update $UB =: \min\{UB, uobj^v\}$ where $uobj^v$ is as described in (36).

end while

return UB and corresponding optimal solution $(\boldsymbol{x}^{v^*}, \boldsymbol{y}^{v^*}, \boldsymbol{\rho}^{v^*})$ for which $uobj^{v^*} = UB$.

B.3. Implementation of the Branch-and-Benders-Cut Algorithm

Algorithm 2 describes the procedure of our implementation of the Branch-and-Benders-Cut Method.

B.4. Deriving the Benders Cuts by solving Dual Sub-Problems of PECSP-DSP$_{\omega,t}^z$

We start by noting that, the sub-problems regarding the relocation decisions in the PECSP model are the same as these in the CCSP model, so here we omit it for conciseness. For the PECSP-SP$_{\omega,t}^z$ defined in equation (19), given a pair $(\bar{\boldsymbol{y}}, \bar{\boldsymbol{\rho}})$ that is feasible according to the MP, we can apply duality theory to get an equivalent dual sub-problem:

[PECSP-DSP$_{\omega,t}^z$]

$$h_{\omega,t}^z(\bar{\boldsymbol{y}}, \bar{\boldsymbol{\rho}}) = \max_{\boldsymbol{\mu}_{t\omega}^1, \boldsymbol{\mu}_{t\omega}^2, \mu_{t\omega}^3} \; -\sum_{j\in\mathcal{J}} \lambda_j^t \bar{y}_j^t \mu_{jt\omega}^1 - \sum_{i\in\mathcal{I}} d_{i\omega}^t \mu_{it\omega}^2 + \mu_{t\omega}^3 \sum_{k=1}^{N-1} (\rho_{k+1,\omega}^t - \rho_{k\omega}^t) \lceil \beta_k^t \sum_{i\in\mathcal{I}} d_{i\omega}^t \rceil \tag{37a}$$

$$\text{subject to } -\mu_{jt\omega}^1 - \mu_{it\omega}^2 + \mu_{t\omega}^3 \leq (1/N) c^t l_{ij} \quad \forall i \in \mathcal{I}, j \in \mathcal{J}_i^z \tag{37b}$$

$$\boldsymbol{\mu}_{t\omega}^1 \geq 0,\ \boldsymbol{\mu}_{t\omega}^2 \geq 0,\ \mu_{t\omega}^3 \geq 0, \tag{37c}$$

Algorithm 2 Branch-and-Benders-Cut Implementation

Input A tolerance $\epsilon \geq 0$ and maximum run time *stoptime*
Initialize $UB = +\infty$, $LB = -\infty$, $\mathcal{N} = \{\text{root}\}$ where root is the linear relaxation of MP.
Initialize $G^r = G^z_{t\omega} = \emptyset$ for all $t \in \{1,\ldots,T\}$, and $\omega \in \Omega$.
while ($\mathcal{N}$ is nonempty and $UB - LB > \epsilon$ and $runtime \leq stoptime$) **do**
    Select a node $o' \in \mathcal{N}$ and remove it from $\mathcal{N} \leftarrow \mathcal{N}/o'$.
    Solve linear relaxation of o' with G^r and $G^z_{t\omega}$ sets to obtain optimal solution $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}, \bar{\boldsymbol{\rho}}, \bar{\theta}^r, \bar{\boldsymbol{\theta}}^z)$ and consider optimal value as *lobj* .
    if $lobj < UB$ **then**
        if $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\rho})$ is not integer valued **then**
            Update $LB := \max(LB, lobj)$.
            Branch on a non-integer variable, resulting in nodes o^* and o^{**}.
            $\mathcal{N} \leftarrow \mathcal{N} \cup \{o^*, o^{**}\}$.
        else
            Solve CCSP-DSP^r and CCSP-DSP$^z_{\omega,t}$ to obtain $h^z_{\omega,t}(\bar{\boldsymbol{y}}, \bar{\boldsymbol{\rho}})$ and $h^r(\bar{\boldsymbol{y}})$ together with $(\bar{\boldsymbol{u}}^{r*}, \bar{\boldsymbol{u}}^{1*}, \bar{\boldsymbol{u}}^{2*}, \bar{u}^{3*})$.
            Compute *uobj* as described in (36).
            if $uobj - lobj > \epsilon/2$ **then**
                Add $\bar{\boldsymbol{\mu}}^{r*}$ to G^r, and, for all $t \in \{1,\ldots,T\}$ and $\omega \in \Omega$, add $(\bar{\boldsymbol{\mu}}^1_{t\omega}, \bar{\boldsymbol{\mu}}^{2*}_{t\omega}, \bar{\mu}^{3*}_{t\omega})$ to $G^z_{t\omega}$
                $\mathcal{N} \leftarrow \mathcal{N} \cup \{o'\}$.
            end if
            if $uobj - lobj \leq \epsilon/2$ **then**
                $UB = uobj$, $(\boldsymbol{x}^*, \boldsymbol{y}^*, \boldsymbol{\rho}^*, \theta^{r*}, \boldsymbol{\theta}^{z*}) = (\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}, \bar{\boldsymbol{\rho}}, \bar{\theta}^r, \bar{\boldsymbol{\theta}}^z)$.
            end if
        end if
    end if
end while
return UB and corresponding optimal solution $(\boldsymbol{x}^*, \boldsymbol{y}^*, \boldsymbol{\rho}^*, \theta^{r*}, \boldsymbol{\theta}^{z*})$.

where $\boldsymbol{\mu}^1_{t\omega} \in \mathbb{R}^{|\mathcal{J}|}$, $\boldsymbol{\mu}^2_{t\omega} \in \mathbb{R}^{|\mathcal{I}|}$, $\mu^3_{t\omega} \in \mathbb{R}$ are the dual vectors for constraints (10b), (10c) and (19c) respectively.

Therefore, following a similar procedures as in Appendix B.2, we can obtain the corresponding Benders cuts,

$$\theta^z_{t\omega} \geq -\sum_{j\in\mathcal{J}} \lambda^t_j y^t_j \bar{\mu}^{1\tau}_{t\omega j} - \sum_{i\in\mathcal{I}} d^t_{i\omega} \bar{\mu}^{2\tau}_{t\omega i} + \bar{\mu}^{3\tau}_{t\omega} \sum_{k=1}^{N-1} (\rho^t_{k+1,\omega} - \rho^t_{k\omega}) \lceil \beta^t_k \sum_{i\in\mathcal{I}} d^t_{i\omega} \rceil, \; \forall \tau \in G^z_{t\omega}, t \in \{1,\ldots,T\}, \omega \in \Omega, \quad (38)$$

where $(\bar{\boldsymbol{\mu}}^{1\tau},\bar{\boldsymbol{\mu}}^{2\tau},\bar{\boldsymbol{\mu}}^{3\tau})$, $\tau \in G^z_{t\omega}$ is an optimal solution of PECSP-DSP$^z_{\omega,t}$

B.5. Deriving Feasibility Cuts for CCSP and PECSP

We start by noting that when Assumption 1 does not hold, the subproblems **CCSP-SP**^{r} and **CCSP-SP**$^z_{\omega,t}$ can become infeasible even though the pair $(\boldsymbol{y},\boldsymbol{p})$ is feasible in the MP. When it is the case, one needs to generate a corresponding feasibility cut and add it to the constraints of the MP until all subproblems become feasible.

For the CCSP model, the sets of feasibility cuts regarding **CCSP-SP**^{r} and **CCSP-SP**$^z_{\omega,t}$ can be defined as

$$0 \geq \sum_{j\in\mathcal{J}} \sum_{t=1,\ldots,T-1} \hat{\mu}^{r\tau}_{jt}(y^t_j - y^{t+1}_j) + \sum_{j\in\mathcal{J}} \hat{\mu}^{r\tau}_{jT}(y^T_j - y^1_j),\ \forall \tau \in H^r \tag{39}$$

and

$$0 \geq -\sum_{j\in\mathcal{J}} \lambda^t_j y^t_j \hat{\mu}^{1\tau}_{t\omega j} - \sum_{i\in\mathcal{I}} d^t_{i\omega}\hat{\mu}^{2\tau}_{t\omega i} + \hat{\mu}^{3\tau}_{t\omega}(1-\rho^t_\omega)\lceil \beta^t \sum_{i\in\mathcal{I}} d^t_{i\omega}\rceil,\ \forall \tau \in H^z_{t\omega}, t\in\{1,\ldots,T\}, \omega\in\Omega \tag{40}$$

respectively, where $\{\hat{\boldsymbol{\mu}}^{r\tau}\}_{\tau\in H^r}$ is the set of extreme rays of **CCSP-DSP**^{r} (see definition in Appendix B.2), and $\{\hat{\boldsymbol{\mu}}^{1\tau}_{t\omega},\hat{\boldsymbol{\mu}}^{2\tau}_{t\omega},\hat{\boldsymbol{\mu}}^{3\tau}_{t\omega}\}_{\tau\in H^z_{t\omega}}$ is the set of extreme ray of **CCSP-DSP**$^z_{\omega t}$ defined in Appendix B.2. Given a fixed assignment for $\boldsymbol{y}$ and $\boldsymbol{\rho}$, an extreme rays responsible for infeasibility of (39) or (40) can be obtained directly from solving subproblems **CCSP-DSP**^{r} and **CCSP-DSP**$^z_{\omega,t}$ when using a primal simplex algorithm.

In the case of the PECSP model, the set of feasibility cuts associated to **PECSP-SP**$^z_{\omega,t}$ defined in Section 5.2 can be characterized as

$$0 \geq -\sum_{j\in\mathcal{J}} \lambda^t_j y^t_j \hat{\mu}^{1\tau}_{t\omega j} - \sum_{i\in\mathcal{I}} d^t_{i\omega}\hat{\mu}^{2\tau}_{t\omega i} + \hat{\mu}^{3\tau}_{t\omega}\sum_{k=1}^{N-1}(\rho^t_{k+1,\omega} - \rho^t_{k\omega})\lceil \beta^t_k \sum_{i\in\mathcal{I}} d^t_{i\omega}\rceil,\ \forall \tau \in H^z_{t\omega}, t\in\{1,\ldots,T\}, \omega\in\Omega. \tag{41}$$

where $\{\hat{\boldsymbol{\mu}}^{1\tau}_{t\omega},\hat{\boldsymbol{\mu}}^{2\tau}_{t\omega},\hat{\boldsymbol{\mu}}^{3\tau}_{t\omega}\}_{\tau\in H^z_{t\omega}}$ is the set of extreme rays of **PECSP-DSP**$^z_{\omega,t}$ defined in Appendix B.4.

B.6. Details about the test data classes used in Section 7.2

The size of test data classes regarding the MILP based formulation is described in Table 7.

B.7. Constructing the Coverage Profile Associated to a Solution

Given a solution $(\boldsymbol{x}^*,\boldsymbol{y}^*)$, one can computed the coverage profile that this solution achieved for a specific set of equiprobable scenarios $\{\boldsymbol{d}_\omega\}_{\omega\in\Omega}$ with $|\Omega| = N$ as follows. First, for each $\omega\in\Omega$, one should solve the following linear program

$$\begin{aligned}
\bar{\beta}^*_\omega := \quad & \underset{\boldsymbol{z},\beta}{\text{maximize}} && \beta \\
& \text{subject to} && \sum_{i\in\mathcal{I}} z^t_{ij} \leq \lambda^t_j y^{t*}_j && \forall j\in\mathcal{J}, t\in\{1,\ldots,T\} \\
& && \sum_{j\in\mathcal{J}^z_i} z^t_{ij} \leq d^t_i && \forall i\in\mathcal{I}, t\in\{1,\ldots,T\} \\
& && \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}^z_i} z^t_{ij} \geq \beta\sum_{i\in\mathcal{I}} d^t_i && \forall t\in\{1,\ldots,T\} \\
& && z^t_{ij}\in\mathbb{N} && \forall i\in\mathcal{I}, j\in\mathcal{J}, t\in\{1,\ldots,T\}.
\end{aligned}$$

Letting $\bar{\beta}^*_{[k]}$ represent the k-th smallest element of the list $\{\bar{\beta}^*_\omega\}_{\omega\in\Omega}$, it is then possible to construct $\bar{\beta}(\eta) := \bar{\beta}^*_{[\lfloor \eta N+1\rfloor]}$, where $\lfloor\cdot\rfloor$ is the round down to nearest integer operation. Note that, in our case study, we used the $N = 100$ observations from January 2016 to middle of April 2016.

Table 7 **Details about the test data classes for CCSP (i.e., C1-C12) and PECSP (i.e., C13-C24) including the number of emergency demand zones ($|\mathcal{I}|$), base locations ($|\mathcal{J}|$), time periods ($T = 6$) and scenarios (N). Note that first stage variables consist of x, y, ρ and r while second stage variables consist of z.**

Class	$\lvert\mathcal{I}\rvert$	$\lvert\mathcal{J}\rvert$	N	First Stage		Second Stage	
				# of variables	# of constraints	# of variables	# of constraints
C1	40	20	50	2,940	346	240,000	18,300
C2	40	20	100	3,240	346	480,000	36,600
C3	40	20	150	3,540	346	720,000	54,900
C4	40	20	200	3,840	346	960,000	73,200
C5	80	50	50	15,900	856	1,200,000	39,300
C6	80	50	100	16,200	856	2,400,000	78,600
C7	80	50	150	16,500	856	3,600,000	117,900
C8	80	50	200	16,800	856	4,800,000	157,200
C9	150	100	50	61,500	1,706	4,500,000	75,300
C10	150	100	100	61,800	1,706	9,000,000	150,600
C11	150	100	150	62,100	1,706	13,500,000	225,900
C12	150	100	200	62,400	1,706	18,000,000	301,200
C13	20	10	20	3,120	2,570	24,000	3,720
C14	20	10	40	10,320	9,770	48,000	7,440
C15	20	10	60	22,320	21,770	72,000	11,160
C16	20	10	100	60,720	60,170	120,000	18,600
C17	40	20	20	5,040	2,470	96,000	7,320
C18	40	20	40	12,240	9,940	192,000	14,640
C19	40	20	60	24,240	21,940	288,000	21,960
C20	40	20	100	62,640	60,340	480,000	36,600
C21	80	50	20	18,000	3,250	480,000	15,720
C22	80	50	40	25,200	10,450	960,000	31,440
C23	80	50	60	37,200	22,450	1,440,000	40,800
C24	80	50	100	75,600	60,850	2,400,000	78,600

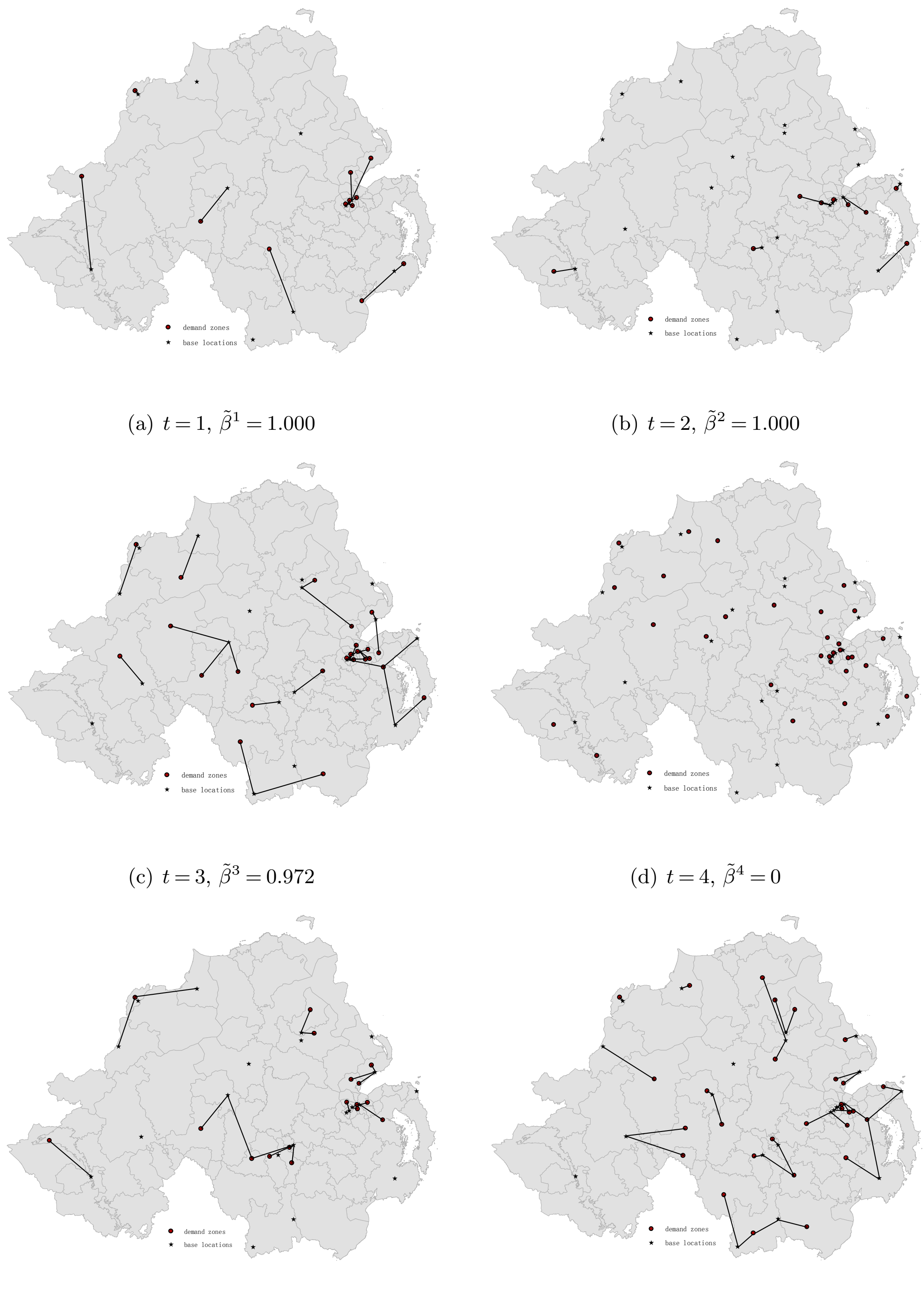
B.8. Geographical Configuration of Optimal EMS Location Strategies

Figures 7 and 8 present the geographical configuration of the time-dependent optimal EMS location strategies obtained for CACCSP and CAPECSP respectively.

B.9. Mathematical Motivation for the Procedure used in the Out-of-Sample Analysis of Section 8.4

Given some fixed t, let the set $\{\boldsymbol{d}_\omega^t\}_{\omega\in\Omega}$ of random demand vectors be identically and independently drawn from a distribution $\mathbb{Q}$. Let also $\tilde{\boldsymbol{d}}$ be the out-of-sample observation for this random vector which is also drawn independently from $\mathbb{Q}$. For the sake of simplicity, we assume that $\mathbb{Q}$ is a continuous distribution.

In what follows, we demonstrate that the procedure that is described in Section 8.4 to decide of the vehicle assignment at time t when observing the out-of-sample scenario $\tilde{\boldsymbol{d}}^t$ both in the CACCSP and CAPECSP models is consistent with the guarantees that are sought by the chance constraints and probabilistic envelope constraints of these respective models.



(a) $t=1$, $\tilde{\beta}^1=1.000$

(b) $t=2$, $\tilde{\beta}^2=1.000$

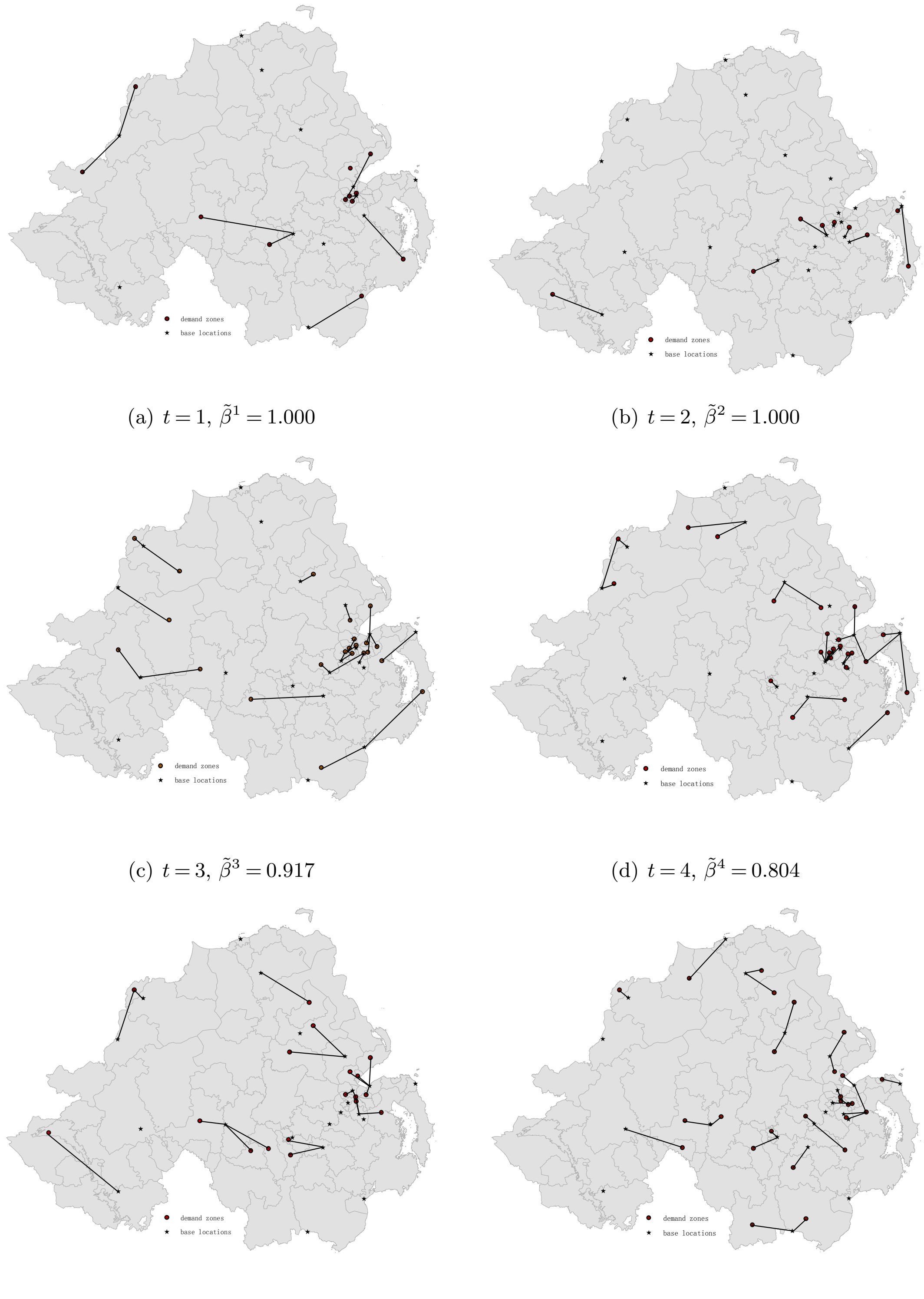
(c) $t=3$, $\tilde{\beta}^3=0.972$

(d) $t=4$, $\tilde{\beta}^4=0$

(e) $t=5$, $\tilde{\beta}^5=0.962$

(f) $t=6$, $\tilde{\beta}^6=0.973$

Figure 7 **Time-dependent optimal EMS location configuration for CACCSP under scenario ω_1 over 6 time periods ($\beta=0.95, \eta=0.05$). We also report the empirical coverage for each time period by $\tilde{\beta}^t$ under scenario ω_1. Solid dots represent the emergency demand zones, stars represents the opened bases and the black line represents the assignment respectively.**



(a) $t=1$, $\tilde{\beta}^1=1.000$ (b) $t=2$, $\tilde{\beta}^2=1.000$

(c) $t=3$, $\tilde{\beta}^3=0.917$ (d) $t=4$, $\tilde{\beta}^4=0.804$

(e) $t=5$, $\tilde{\beta}^5=1.000$ (f) $t=6$, $\tilde{\beta}^6=0.865$

Figure 8 **Time-dependent optimal EMS location configuration for CAPECSP under scenario ω_1 over 6 time periods with $\beta^{cvx}_{0.5}(\eta)$. We also report the empirical coverage for each time period by $\tilde{\beta}^t$ under scenario ω_1. Solid dots represent the emergency demand zones, stars represents the opened bases and the black line represents the assignment respectively.**

First, given an envelope function $\beta^t(\cdot) = \bar{\beta}(\cdot)$, in both the CACCSP and CAPECSP models, the coverage that is imposed on the out-of-sample scenario $\tilde{\boldsymbol{d}}^t$ can be defined as follows:

$$\tilde{\beta} = \sup_{\eta < (N - \text{Rank}(\tilde{\boldsymbol{d}})+1)/N} \bar{\beta}(\eta) = \sup_{\epsilon > 0} \bar{\beta}((N - \text{Rank}(\tilde{\boldsymbol{d}}^t) + 1)/N - \epsilon)\,,$$

where $\text{Rank}(\tilde{\boldsymbol{d}}^t)$ is the rank of $\sum_{i\in\mathcal{I}} \tilde{d}_i^t$ in the list of $\{\sum_{i\in\mathcal{I}} d_{i\omega}^t\}_{\omega\in\Omega/\omega_1}$ where we assume without loss of generality that the first scenario was the one removed randomly among the set of size $|\Omega|$. In particular, since both $\{\boldsymbol{d}_\omega^t\}_{\omega\in\Omega}$ and $\tilde{\boldsymbol{d}}^t$ are i.i.d. and $\mathbb{Q}$ is continuous, we can conclude that, given $\tilde{\boldsymbol{d}}^t$, the random variable $\text{Rank}(\tilde{\boldsymbol{d}}^t) - 1$ is distributed according to a binomial distribution $\text{Binomial}((N-1)\tilde{p}, (N-1)\tilde{p}(1-\tilde{p}))$ where $\tilde{p} = \mathbb{P}_{\mathbb{Q}}(\sum_{i\in\mathcal{I}} d_{i\omega}^t \leq \sum_{i\in\mathcal{I}} \tilde{d}_i^t)$. We are now ready to argue that in the case of the CACCSP model for all $\hat{\eta} > \eta$, there exists a $N := |\Omega|$ such that $\mathbb{P}(\tilde{\beta} \geq \bar{\beta}(\eta)) \geq 1 - \hat{\eta}$, where by $\mathbb{P}$ we denote the probability with respect to the joint distribution of $\{\boldsymbol{d}_\omega^t\}_{\omega\in\Omega}$ and $\tilde{\boldsymbol{d}}$. On the other hand, for the CAPECSP, we will show that if $\bar{\beta}(\cdot)$ is continuous, then for all $\varepsilon > 0$, there exists an $N := |\Omega|$ such that $|\tilde{\beta} - \bar{\beta}(1-\tilde{p})| \leq \varepsilon$ with high probability. This confirms that the procedure we propose as a policy to decide how much to cover once the demand is known is well motivated if we assume that the number of samples in Ω is large enough.

Looking more closely at the case of the CACCSP model, we can consider some $\Delta < \hat{\eta} - \eta$ and that N is large enough to show that:

$$\begin{aligned}
\mathbb{P}(\tilde{\beta} \geq \bar{\beta}(\eta)) &= \mathbb{P}(\sup_{\epsilon>0} \bar{\beta}((N - \text{Rank}(\tilde{\boldsymbol{d}}) + 1)/N - \epsilon) \geq \bar{\beta}(\eta)) && (42)\\
&= \mathbb{P}((N - \text{Rank}(\tilde{\boldsymbol{d}}) + 1)/N > \eta) && (43)\\
&= \mathbb{P}((N - (N-1)\tilde{p} - \xi)/N > \eta) && (44)\\
&\geq \mathbb{P}\left(\frac{N-(N-1)\tilde{p}}{N} \geq \eta + \Delta\right)\cdot \\
&\quad \mathbb{P}\left(\frac{N-(N-1)\tilde{p}-\xi}{N} > \eta \,\middle|\, \frac{N-(N-1)\tilde{p}}{N} \geq \eta + \Delta\right) && (45)\\
&= \frac{N}{N-1}(1-\eta-\Delta)\mathbb{P}\left(\frac{\xi}{N} < \Delta \,\middle|\, \frac{N-(N-1)\tilde{p}}{N} \geq \eta + \Delta\right) && (46)\\
&\geq \frac{N}{N-1}(1-\eta-\Delta)\mathbb{P}\left(|\frac{\xi}{N}| < \Delta \,\middle|\, \frac{N-(N-1)\tilde{p}}{N} \geq \eta + \Delta\right) && (47)\\
&\geq \frac{N}{N-1}(1-\eta-\Delta)\mathbb{E}\left[1 - \frac{\tilde{p}(1-\tilde{p})(N-1)}{\Delta^2 N^2} \,\middle|\, \frac{N-(N-1)\tilde{p}}{N} \geq \eta + \Delta\right] && (48)\\
&\geq \frac{N}{N-1}(1-\eta-\Delta)\left(1 - \frac{N-1}{4\Delta^2 N^2}\right) \to_{N\to\infty} (1-\eta-\Delta) && (49)\\
&> 1 - \hat{\eta} \text{ for some large enough } N\,, && (50)
\end{aligned}$$

where $\xi := \text{Rank}(\tilde{\boldsymbol{d}}) - (N-1)\tilde{p} - 1$ is the centered version of $\text{Rank}(\tilde{\boldsymbol{d}})$ such that $\mathbb{E}[\xi] = 0$. In details, we have that the first and second steps follows from the definition of $\tilde{\beta}$ while the third step follows from our definition of ξ. The fourth step follows from looking at only the events such that $\tilde{p}$ is small enough. Considering that $\tilde{p}$ is the quantile of a certain random variable, we necessarily have that $\mathbb{P}(\tilde{p} \leq p) = p$ which leads to equation (46) after simplifying the second term. We then exploit Chebyshev inequality to get expression (48) using the fact that $\mathbb{E}[\xi/N] = 0$ and $\mathbb{E}[\xi^2/N^2] = (\tilde{p}(1-\tilde{p})(N-1))/N^2$. We obtain our last expression using the fact that $\tilde{p}(1-\tilde{p})$ is maximized at $\tilde{p} = 0.5$ and that $(N/(N-1))(1-(N-1)/(4\Delta^2 N^2))$ converges to 1 as N goes to infinity.

In the case of the CAPECSP model, we can start by assuming that $\beta(\cdot)$ is K-Lipschitz continuous which in particular implies that $\tilde{\beta} = \bar{\beta}((N - \text{Rank}(\tilde{\boldsymbol{d}}) + 1)/N)$. Furthermore, we can bound with high probability the difference between the targeted coverage $\bar{\beta}(1-\tilde{p})$ and coverage $\tilde{\beta}$ obtained by our procedure as follows:

$$\begin{aligned}
|\bar{\beta}(1-\tilde{p}) - \tilde{\beta}| &= |\bar{\beta}(1-\tilde{p}) - \bar{\beta}((N - \text{Rank}(\tilde{\boldsymbol{d}}) + 1)/N)| \\
&\leq K|1 - \tilde{p} - (N - \text{Rank}(\tilde{\boldsymbol{d}}) + 1)/N| \\
&= K\left|\tilde{p} - \frac{(N-1)\tilde{p}}{N} - \frac{\xi}{N}\right| \\
&\leq K\left(\left|\tilde{p} - \frac{(N-1)\tilde{p}}{N}\right| + \left|\frac{\xi}{N}\right|\right) \leq K\left(\frac{1}{N} + \left|\frac{\xi}{N}\right|\right) \\
&\leq K(\frac{1}{N} + \frac{\sqrt{(N-1)\tilde{p}(1-\tilde{p})}}{N\sqrt{\delta}}),
\end{aligned}$$

with probability larger than $1-\delta$ based on Chebyshev inequality. Hence, since the last expression converges to 0 as N goes to infinity, it must be that for any level of confidence $1-\delta$ and accuracy $\varepsilon > 0$ there exists some N such that

$$\mathbb{P}(|\bar{\beta}(1-\tilde{p}) - \tilde{\beta}| \leq \varepsilon) \geq 1 - \delta\,.$$