Ann Oper Res (2013) 203:389–406
DOI 10.1007/s10479-012-1174-6

Probabilistic programming models for traffic incident management operations planning

Kaan Ozbay · Cem Iyigun · Melike Baykal-Gursoy · Weihua Xiao

Published online: 4 July 2012

Abstract This paper proposes mathematical programming models with probabilistic constraints in order to address incident response and resource allocation problems for the planning of traffic incident management operations. For the incident response planning, we use the concept of quality of service during a potential incident to give the decision-maker the flexibility to determine the optimal policy in response to various possible situations. An integer programming model with probabilistic constraints is also proposed to address the incident response problem with stochastic resource requirements at the sites of incidents. For the resource allocation planning, we introduce a mathematical model to determine the number of service vehicles allocated to each depot to meet the resource requirements of the incidents by taking into account the stochastic nature of the resource requirement and incident occurrence probabilities. A detailed case study for the incident resource allocation problem is included to demonstrate the use of proposed model in a real-world context. The paper concludes with a summary of results and recommendations for future research.

Keywords Transportation · Incident management · Logistics · Quality of service · p-Efficient points · Stochastic programming · Probabilistic constraints

K. Ozbay
Department of Civil and Environment Engineering, Rutgers University, 623 Bowser Road, Piscataway, NJ 08854-8014, USA
e-mail: kaan@rci.rutgers.edu

C. Iyigun (✉)
Department of Industrial Engineering, Middle East Technical University, Ankara, Turkey
e-mail: iyigun@ie.metu.edu.tr

M. Baykal-Gursoy
Department of Industrial and Systems Engineering, Rutgers University, 96 Frelinghuysen Road, Piscataway, NJ 08854-8018, USA
e-mail: gursoy@rci.rutgers.edu

W. Xiao
Capital One, 6929 North Lakewood Avenue, Tulsa, OK 74117, USA
e-mail: xiao@rci.rutgers.edu

1 Introduction

Traffic incidents account for approximately 60 percent of the vehicle-hours lost due to congestion annually (National Conference on TIM 2002). It is now widely accepted that these congestion and congestion-related problems can be decreased by the proper use of an efficient incident management system. Since the initial investment, maintenance cost and operating cost of each response unit are considerable, there is a great need for reliable decision-support models to help evaluate and optimize the performance of such systems (Lindley 1989).

Increased attention in literature is focused on the incident response problem. The need for improved incident response models and the data available for developing such models were discussed in Ozbay and Kachroo (1999). Recognizing the highly stochastic nature of traffic and incident management operations, Pal and Sinha (2002), introduced a simulation model that could be used in designing a new freeway service patrol, as well as improving the operations of existing programs. In a paper by Ozbay and Bartin (2003), a simulation model was developed using ARENA (Kelton et al. 2001) simulation package, and was used to model and examine the effects of various incident management strategies for the incident management operations on the Washington D.C. beltway network. Repede and Bernardo (1994) used a simulation based decision support system for locating emergency medical vehicles. However, stochastic computer simulation might be time consuming especially when running multiple replications for real-size networks to obtain statistically significant results. Another problem with computer simulation is the difficulty in analyzing various components of the incident management problem in a methodological way due to the possibly quite large size of such variables when real networks are employed. In fact, Pal and Sinha (2002) pointed out the need for a systematic procedure to optimally design a freeway service patrol program using simulation. However, they and others clearly recognize the calibration, computational, and analysis related problems of simulation based approaches when a full or partial design of experiments approach is adopted.

Besides computer simulation, mathematical modeling is another frequently used approach to study incident management problems. Initial studies involve location-allocation of emergency service facilities in the context of p-median problem (see e.g., Hakimi 1964), set covering problem (see e.g., Toregas et al. 1971), maximal covering problem (see, e.g., Church and ReVelle 1974; White and Case 1974), expected maximal covering problem (Daskin 1983) (also see the references in Daskin 1987). Daskin (1987) formulated multi-objective mixed integer programming (MIP) problems to simultaneously locate, dispatch and route emergency vehicles. Zografos et al. (1993) proposed an analytical framework that could minimize the freeway incident delays through the optimum deployment of traffic flow restoration units (TFRU). According to the authors of this study, this model has been proven to be an effective tool that can model and evaluate the effects of deployment of TFRU on the overall freeway incident delays. Pal and Sinha (1997) constructed a mixed integer programming (MIP) model to determine optimal locations for response vehicles that minimizes the annual cost of response vehicles subject to a constraint on the maximum number of vehicles. It is assumed that the frequencies of incidents at potential sites in the network are given. Petty (1997) proposed a model to determine the optimal placement of tow trucks using traffic theory in combination with marginal benefit analysis.

Yin (2006) suggested a min-max bi-level programming model which assigns tow trucks to freeway service patrol beats. The objective is to minimize the maximal total travel time that may be caused by the incidents. However, it was assumed that the number of incidents is deterministic. Later Yin (2008) dealt with the randomness in the incident occurrence with a scenario-based approach and proposed a mixed-integer nonlinear programming model. Each scenario corresponds to a specific number of incidents occurring on each beat and the objective is to find a truck assignment which performs well under all scenarios. Lou et al. (2011) proposed a similar study where the coverage of the freeway segments by the beats has been considered in addition to the allocation of freeway service trucks. They presented two mixed-integer nonlinear programming models: one with deterministic travel time and number of incidents, the other one (scenario-based) with random incidents and travel times. Articles such as Yin (2006, 2008), Lou et al. (2011) and references therein discuss deployment of freeway service patrol (FSP) trucks. These models are not directly applicable (also stated in Lou et al. 2011) because FSP trucks are not dispatched in response to an incident. Instead, they are mobile emergency response units that independently roam the freeways to detect, respond to, and clear traffic incidents. But the models in this paper discuss location-allocation models for locating emergency response vehicles and determining optimal strategies for locating and dispatching those, introducing probabilistic constraints.

In a different context, Sherali et al. (2004) considered an emergency management problem aftermath of a natural disaster, a terrorist attack or an earthquake and they proposed a non-convex programming model for allocating emergency response resources to minimize the risk.

Opportunity cost-based models proposed by Sherali and Subramanian (1999) demonstrated that dispatching the closest available vehicle to the site of the current accident might not be optimal when considering the service to anticipated demands. To make this model polynomial-time solvable, the number of response vehicles required by each incident was considered to be the same, and each depot was assumed to have the same number of available vehicles. In practice, this might not be true. Due to day-to-day uncertainties, such as response vehicle breakdown, lack of drivers, etc., it is always possible to have an insufficient number of vehicles at any given day, and, the severity of accidents also varies significantly. In general, severe incidents need more response vehicles than minor incidents.

Thus, the resource demand in an incident management problem can be summarized as follows:

(1) **It is stochastic**. The available resources are stochastic and the occurrences and characteristics of incidents are stochastic.
(2) **It is a network problem**. The incidents occur randomly over the roadway network, and the location of the depots should be chosen carefully based on the topology of the network.
(3) **It is a resource allocation problem**. Resources should be allocated wisely among individual depots and between depots and patrol service to maximize the return on investment (ROI).

In this paper, we attempt to address the traffic incident management, as an operation planning problem for a given time horizon. We specifically take into account the interaction between the probabilistic occurrence rates of accidents requiring different levels of resources and the availability of adequate number of incident response units, using a stochastic programming approach. Mainly, the ideas in this paper follow the concepts in Sherali and Subramanian (1999) where they introduced "multiple-incident multiple-response (MIMR)" model seeking to find an optimal assignment of response vehicles to incidents to minimize the sum of service and opportunity cost. We develop an MIMR model by recognizing that there is a risk that some incidents might be left without service for an insufficient vehicle fleet. We propose using system-wide reliability to measure the quality of service in the network, involving multiple potential incidents and a variable number of response vehicles. A threshold confidence level is to be satisfied in servicing the incidents on the network.

We use the concept of *quality of service* and propose mathematical programming models with probabilistic constraints to model the stochastic incident response problem. Probabilistic constraints are first introduced by Charnes et al. (1958) in the storage of heating oil to meet random demands but they introduced only individual probabilistic constraints.

The joint probabilistic constrained problems have also been studied and applied for many practical problems including location and coverage problems, also adopted in this paper for emergency vehicles. The joint probabilistic constraints formulations which are more challenging both from theoretical and algorithmic points of view, were studied by Miller and Wagner (1965) and Prekopa (1970, 1973). Later, probability constraints are used in the set covering location problem under the assumption that servers could be independently unavailable and with the identical probability (Chapman and White 1974; Aly and White 1978), and in the maximal covering problem (ReVelle and Hogan 1989, see also Ball and Lin 1993) with independent but individual busy probabilities. Pal and Bose (2009) extended the reliability formulation in Ball and Lin (1993) and proposed a mixed integer programming model that locates the incident response depots and assigns response vehicles to these depots at a minimum cost. Note that in these models demand is for a single emergency service. On the other hand, in the models presented in this paper, multiple potential incidents having various demands for response vehicles are allowed, and the number of available response vehicles at each depot might be non-deterministic. The probability distributions of demand for resources on each node are allowed to be correlated.

In the case of resource allocation, some areas in the transportation network may have a higher probability of experiencing serious incidents than others, due to heavy traffic conditions or complex roadway characteristics. Thus, more resources should be allocated to those depots that are located closer to these high risk areas. Given the probability distribution of demand for resources over the network, we propose a stochastic integer programming (IP) model to determine the optimum level of resources, and the best way to allocate resources over the transportation network.

The remainder of this paper is organized as follows. In the next section, we introduce a traffic incident response problem formulation that captures the uncertainty of demand for resources and availability of resources. An IP model with probabilistic constraints and discrete random variables is proposed to address this problem with consideration given to the requirements at the sites of potential incidents. Another stochastic IP is introduced to solve the resource allocation problem. In Sect. 3, we provide a case study which is an application of the resource allocation model that deals with tow truck allocation to depot sites located on the South New Jersey road network. The solution methodology is briefly explained in Appendix. Paper concludes with a brief summary of results in Sect. 4.

2 Problem formulation

In this paper, we consider two types of incident management (IM) problems: (1) One is the incident response problem, where we determine the optimal resource assignment policy for possible incidents that might occur simultaneously; (2) The other is the resource allocation problem that will allow transportation agencies to determine the optimal location of each depot and the optimal number of vehicles assigned to the depots. We approach both of these problems as planning problems for a given time horizon.

Let us first introduce the mathematical notation used in this paper. Let $G(N, L)$ be the road network, and N and L be its node and link sets, respectively. Let D denote the set of special type of nodes, the depots, from which service vehicles are assigned, H denote the node set where potential incidents might occur. In addition, let r_i be the number of service vehicles available, thus idle, in depot $i \in D$, and N_h the number of service vehicles that might be required by a possible incident at node $h \in H$. Here, service vehicle is an abstract concept for any type of vehicles that might be needed to clear the incidents. In practice, it could be tow trucks, ambulances, police cars or fire trucks. In this paper, we consider N_h as a scalar random variable governed by the probability distribution, $P\{N_h = k\}$, that gives the probability of an incident per planning time unit requiring k response-vehicles at node h. In the case that multiple incidents occur at various nodes, the joint probability distributions, $P\{N_{h1} = k_1, N_{h2} = k_2, \ldots\}$, are assumed to be known, so that the marginal distributions could be obtained. Although N_h could be 0 even if an incident occurs, here we assume that $N_h = 0$ means no incident occurs at node h. Thus, $P\{N_h = 0\} = 1 - P_h$, where P_h denotes the probability of an incident at node h per unit planning time, and

$$P_h = \sum_{k=1}^{\infty} P\{N_h = k\}.$$

Finally, decision variables y_{ih} denote the number of service vehicles in depot i allocated to an incident that might occur at node $h \in H$. We employ the concept of quality of service to quantify the system reliability in the presence of traffic incidents.

Definition 1 Quality of Service parameter $Q_{service} \in [0, 1]$ represents the lower bound of the probability that all resources requested by all incidents will be satisfied.

It is worth mentioning that in the incident response problem, $Q_{service} = 0$ signifies that the possibility of incidents are not considered at all, while $Q_{service} = 1$ means the maximum possible resource demand by all incidents is taken into account. In the context of resource allocation problem, the quality of service makes more sense if we consider it as a measure of system reliability. Higher quality of service guarantees that more incidents can be cleared in a timely manner, thus achieving higher reliability of the transportation system. The effects of quality of service on the determination of optimal incident response policy and resource allocation strategy are discussed in the subsequent sections using real-world examples.

2.1 Incident response under demand uncertainty (P1)

This subsection focuses on how to assign response vehicles to incidents if future demand for resources is considered to be non-deterministic. In the proposed model, this probabilistic nature is reflected in the cost function as well as the constraint set. Note that, the vehicle requirement for incidents follows a discrete probability distribution (this distribution can be derived from the previous accidents by observing the frequency of number of vehicles required).

Given the quality of service level $Q_{service}$, the proposed model is formulated as follows:

(P1)

$$\text{Minimize} \quad E\{\textit{Total Response Time}\} \tag{1a}$$

$$\text{Subject to} \quad \sum_{h \in H} y_{ih} \le r_i, \quad \forall i \in D, \tag{1b}$$

$$P\left(\sum_{i \in D} y_{ih} \ge N_h, \forall h \in H\right) \ge Q_{service}, \tag{1c}$$

$$y_{ih} \ge 0, \quad \text{and integer}, \quad \forall i \in D,\ h \in H, \tag{1d}$$

where

y_{ih}: number of service vehicles dispatched from depot location i to potential incident location h,
r_i: number of service vehicles available in depot $i \in D$, and
N_h: number of service vehicles that might be required by a possible incident at node $h \in H$,
$Q_{service}$: system-wide quality service level that represents the lowest level of reliability to satisfy the potential incidents occurring anywhere in the network promptly, and $Q_{service} \in [0, 1]$.

The objective function minimizes the expected total response cost. The response cost is the vehicle delay (vehicle-hours) due to simultaneous incidents. By writing the objective function this way, the response cost is minimized while the given service quality for simultaneous incidents is guaranteed. If we do not consider the possibility of incidents, i.e., $Q_{service} = 0$, then any y_{ih} value will satisfy constraint (1c), thus minimizing the objective function forces $y_{ih} = 0$. Consequently, there is no cost in this case. Note that, one can also enforce a priority structure on the incident locations depending on the traffic volume that will be affected by the incident. This could be done by assigning weights, w_h's, between 0 and 1 to each node, where 0 corresponds to very low priority, 1 corresponds to the highest priority node. These weights are assigned with respect to the historical traffic volume data.

Constraint (1b) states that the total number of vehicles assigned from a depot should be less than or equal to the number of vehicles available at that depot. Constraint (1c) requires that the joint probability of meeting the response-vehicle requirements at each potential incident site $h \in H$, is greater than or equal to a given quality of service.

To simplify the objective function, given that an incident occurs at node h, we define t_{ih} as the expected travel time on the shortest route from depot i to node h. Any route may include multiple links. We assume here that the incident only affects node h, and not the other nodes. Thus, the expected response time is given below in terms of P_h, instead of the joint probability distribution of incidents on every node, as:

$$\text{Minimize} \quad \sum_{i \in D} \sum_{h \in H} w_h P_h t_{ih} y_{ih}. \tag{1.1a}$$

2.2 Incident response under resource uncertainty

In addition to the fact that resources required by potential incidents are not deterministic, resources available at the depots might be uncertain due to break-downs or other unexpected factors. Similarly we also employ the concept of resource reliability to quantify vehicle availability in the presence of traffic incidents.

Definition 2 Resource reliability $Q_{resource} \in [0, 1]$ represents the lowest probability that the resources at every depot location will be available to respond to all incidents.

In this case, we can formulate the problem nearly the same way as problem (P1), except that constraint (1b) is replaced by the following probabilistic constraint,

$$P\left(\sum_{h \in H} y_{ih} \le R_i, \forall i \in D\right) \ge Q_{resource}, \tag{2}$$

where R_i is a discrete random variable denoting the number of response vehicles available at depot i, (different than the deterministic capacity level, r_i in (1b)) and $Q_{resource}$ denotes the given quality of service for response vehicle availability, which is determined by the decision makers.

2.3 Resource allocation problem (P2)

The goal of a resource allocation problem in the context of traffic incident management is to determine optimal locations of each depot and the optimal fleet size of response vehicles at each depot to minimize the overall response costs, while meeting the budget constraint. Since this problem is a long-term planning problem, probability of an incident on each road for the whole planning horizon is assumed to be equal to 1, i.e., $P_h = 1$, $\forall h \in H$. One of the constraints of this problem is that the resources allocated over the network should be sufficient to adequately address all incidents. As before, assume we know the probability distribution of resource demand at node h, which is governed by the random variable N_h. Given the reliability requirement of this incident response system, the resource allocation problem can be formulated as an IP with probabilistic constraints.

(P2)

$$\text{Minimize} \quad \sum_{i \in D} \sum_{h \in H} w_h P_h t_{ih} y_{ih} \tag{3a}$$

$$\text{Subject to} \quad \sum_{h \in H} y_{ih} \le r_i, \quad \forall i \in D, \tag{3b}$$

$$P\left(\sum_{i \in D} y_{ih} \ge N_h, \forall h \in H \right) \ge Q_{service}, \tag{3c}$$

$$r_i \le M d_i, \quad \forall i \in D, \tag{3d}$$

$$c_1 \sum_{i \in D} r_i + c_2 \sum_{i \in D} d_i \le B, \tag{3e}$$

$$y_{ih} \ge 0, \quad \text{and integer}, \quad \forall i \in D,\ h \in H,$$

$$r_i \ge 0, \quad \text{and integer}; \quad d_i \text{ binary}, \quad \forall i \in D. \tag{3f}$$

We use the same notation as in the previous sections. In addition, we use M to denote a large enough number, B to denote the budget limit, and let c_1 and c_2 be the unit cost for service vehicles and the construction cost for a single depot respectively. d_i's for each node i, are binary variables defined as follows:

$$d_i = \begin{cases} 1 & \text{if node } i \text{ is a depot } (r_i > 0), \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

Constraint (3c) states the reliability of the incident response system, namely, the joint probability of satisfying the resources requested by each incident should be greater than or equal to a given value. Constraint (3d) states the logical relationship between d_i and r_i. If $d_i = 0$, then r_i must be 0; otherwise, r_i could be any positive number less than M. Constraint (3e) states that the budget constraint must be satisfied. The objective function is the sum of weighted response time.

The assumption that the depot locations are known is a realistic one since it is costly to build new depots and not a very practical approach for the short and mid-term policy decision. However, it is very important to know the optimal number of service vehicles needed in a given network.

As it is mentioned before, the main contribution of the paper is modeling explicitly the requirements of multiple response vehicles by incidents and formulating incident management problem using probabilistic constraints, as given in (1c) of problem (P1) and (3c) of problem (P2) which signify the stochastic characteristics of the requirements. In our case, the response vehicle requirements have discrete probability distributions. In Appendix, we review the methods proposed in Prekopa (1990, 1995) and Prekopa et al. (1998) to solve mathematical programming models with probabilistic constraints. This approach computes the Pareto frontiers (also known as *p-efficient* points) using a recursive enumeration algorithm (named as *p-lep* algorithm), given in Prekopa et al. (1998) for a given discrete distribution. After we obtain *p-efficient points*, probabilistic constraints (1c) in problem (P1) and (3c) in problem (P2) can be handled as in (A.2) of Appendix and these problems can be formulated as in (A.3) and further as in (A.4), shown in Appendix, which are solvable by many software packages.

The time-efficiency of this enumeration algorithm is not a major concern for a realistic incident management problem, because the number of routes considered in a typical transportation network is limited mainly to the major routes. Thus the enumeration of *p-efficient* points approach proposed by Prekopa et al. (1998) is a feasible technique. Based on this algorithm, we developed a computer program in C++, which runs very fast for the case study presented in Sect. 3 and others with several hundred nodes similar to them.

However, in some cases, for example, for multidimensional random vectors $\underline{\xi}$, the number of *p-efficient* points can be quite large and their enumeration may be difficult. It would be desirable, therefore, to avoid the complete enumeration and search for promising *p-efficient* points only. In this case, the solution method does not find all *p-efficient* points but builds them up subsequently by the use of cutting plane method. The detailed discussion is given in Prekopa et al. (1998).

3 A case study

In this section, we present an example to demonstrate the applications of the stochastic programming models proposed for incident management decision support. The example used to illustrate the solution of a realistic resource allocation problem, formulated as problem (P2) in this paper, is applied to a simplified version of South Jersey highway network where there is already an active incident management program. In this example, given the requirement of service quality, we determine the minimum resources allocated to the system in order to maximize the investment on return value.

3.1 Resource allocation problem (P2)

We consider a portion of South New Jersey road network which consists of seven major roads depicted in Fig. 1.

The distribution of the demand for the tow trucks is based on the analysis of the traffic incident data in South New Jersey.

By categorizing traffic incidents that occurred in years 2000 and 2001 in Table 1, and assuming that an average number of tow trucks required by each category is based on the severity of incidents, as listed in the first column of Table 1, we obtain the probability values which are presented in Table 2. The average travel times along each route with and without incidents are obtained by a traffic simulation software package developed at Rutgers Intelligent Transportation Systems (RITS) Laboratory, and which is based on Daganzo's cell transmission model (Daganzo 1994). The results are also shown in Table 2.

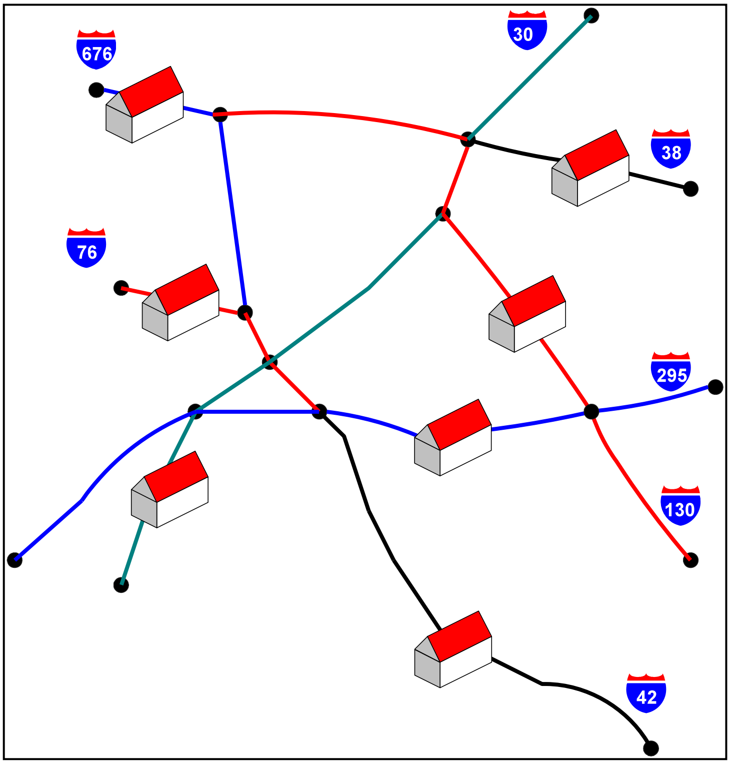

Fig. 1 Portion of the South Jersey roadway network used in this study

Table 1 Number of incidents on the major routes in South Jersey study network (2000–2001) (Values in the parenthesis are the average numbers of trucks requested by the incident in that category)

Category	US 30	NJ 38	NJ 42	I-76	US 130	I-295	I-676
HAZMAT[a] (4)	2	3	2	2	3	13	3
Vehicle-fire (3)	1	1	11	5	2	27	2
Weather-related (1)	9	3	1	1	6	3	3
Disablement-No Blocked Lanes (1)	0	2	2	3	1	6	1
Disablement-Blocked Lanes (2)	0	1	6	6	1	11	4
MVA[b]-Day Time-No Blocked Lanes (1)	21	21	22	36	30	94	8
MVA[b]-Day Time-Blocked Lanes (2)	12	14	24	25	12	81	10
MVA[b]-Night Time-No Blocked Lanes (1)	3	2	1	7	5	8	3
MVA[b]-Night Time-Blocked Lanes (2)	3	2	7	12	2	21	1
Total	51	49	76	99	62	264	35

[a]Hazardous material

[b]Multiple Vehicle Accident

The road network is then modeled as a graph shown in Fig. 2. Although in this example, each node corresponds to a route; our model allows more refined analysis by dividing the route into various sections, thus representing one route as a collection of multiple

Table 2 Average travel time and distribution of requested tow trucks on each route

Route #	Route	Average travel time (s) *without incidents*	Average travel time (s) *with incidents*	Distribution of requested tow trucks			
				1	2	3	4
1	US 30	754	6041	0.647	0.294	0.02	0.039
2	NJ 38	286	313	0.571	0.347	0.021	0.061
3	NJ 42	554	616	0.342	0.487	0.145	0.026
4	I-76	242	4301	0.484	0.443	0.052	0.021
5	US 130	1246	1666	0.677	0.243	0.032	0.048
6	I-295	773	5138	0.421	0.428	0.102	0.049
7	I-676	305	4426	0.429	0.429	0.057	0.085

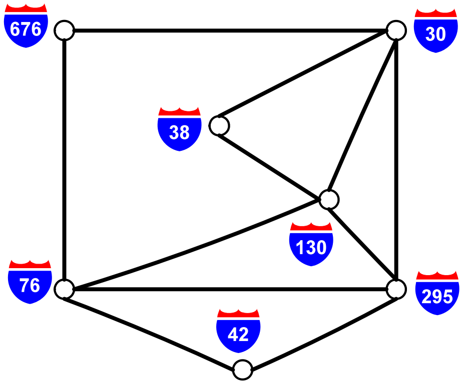

Fig. 2 Simplified graph representation of the South Jersey road network

Table 3 Average travel time between each pair of nodes (seconds)

	US 30	NJ 38	NJ 42	I-76	US 130	I-295	I-676
US 30	6041	1067	1917	5360	2420	5892	5180
NJ 38	6327	313	2203	5646	1952	6178	5466
NJ 42	7142	2168	616	4855	1765	5692	5222
I-76	6588	1614	858	4301	1908	5380	4673
US 130	7287	1559	1407	5547	1666	6384	5914
I-295	6814	1840	1389	5074	2439	5138	5441
I-676	6364	1372	1163	4606	2213	5685	4426

nodes. A link between two nodes implies there is an intersection that connects these two routes.

We assume the time for the tow truck traveling from its depot to the site of the incident on the same route is the average travel time of this route under the impact of incidents. If tow trucks that belong to another route are sent from the depot, then the total traveling time of the tow truck would be the summation of the average travel time of each node along the shortest path from the depot to the site of the incident and the average traveling time with incidents of the ending nodes. Based on these assumptions and the shortest path algorithm, the travel time between every two nodes are obtained as in Table 3.

Table 4 p-Efficient points for $Q_{service} = 0.9$

Point number	US 30	NJ 38	NJ 42	I-76	US 130	I-295	I-676
1	2	4	3	4	4	4	4
2	2	4	4	3	4	4	4
3	3	3	4	4	4	4	4
4	3	4	3	3	4	4	4
5	4	4	3	3	4	3	4
6	4	4	3	3	3	4	4
7	4	3	3	4	4	4	4
8	4	3	4	3	4	4	4
9	4	2	4	4	4	4	4
10	4	4	3	2	4	4	4
11	4	4	4	4	4	4	3
12	4	4	4	4	3	3	4
13	3	4	4	4	3	4	4
14	4	4	4	3	2	4	4
15	3	4	4	4	4	3	4

The number of tow trucks sent to an incident site is determined by the attributes of the incident. If the number of tow trucks requested by the incident exceeds the number of the available tow trucks in the nearest depot, then the rest of the trucks would be sent from the next nearest depot with idle trucks. Given the probability distribution of incidents and the average travel time at hand, we will proceed with the resource allocation analysis on South Jersey roadway network in two steps. First, we assume that tow truck depots exist along these routes, and we just need to determine the number of tow trucks assigned to each depot. Second, we determine the location of depots and the number of tow trucks allocated.

Now, assuming there is a depot constructed on each route, to maximize the return on investment, we need to determine the minimum number of tow trucks allocated to each route, while satisfying a given quality of service. Assume that there is no priority structure among the incident sites, giving all w_i's equal to 1. Let y_{ij} be the number of tow trucks assigned from the depot on route i to the site of incident on route j, and t_{ij} be the corresponding travel time. The objective function of this problem can be represented as

$$\text{Minimize} \quad \sum_{i=1}^{7}\sum_{j=1}^{7} t_{ij} y_{ij}. \tag{5}$$

The *p-efficient points* for the given probability distribution in Table 2 are calculated using p-lep algorithm presented in Prekopa et al. (1998). If the service guarantee level, $Q_{service}$ is 0.9 the p-efficient points are presented in Table 4.

Let n_{kj} be the demand for tow trucks on route j according to the kth p-efficient point, and let λ_k be the corresponding coefficients which satisfy $\sum_{k=1}^{15} \lambda_k = 1$. If we do not consider the budget limitation, then the main constraints of this optimization problem are:

$$\sum_{j=1}^{7} y_{ij} \le r_i, \quad \forall i = 1, \dots, 7, \tag{6a}$$

$$\sum_{i=1}^{7} y_{ij} \ge \sum_{k=1}^{15} \lambda_k n_{kj}, \quad \forall j = 1, \dots, 7, \tag{6b}$$

Table 5 Optimal number of tow trucks assigned to each route for various $Q_{service}$ levels

$Q_{service}$	US 30	NJ 38	NJ 42	I-76	US 130	I-295	I-676	Total
0.5	2	2	3	2	2	2	2	15
0.7	2	4	4	2	4	3	2	21
0.9	2	4	4	3	4	4	4	25

$$\sum_{k=1}^{15} \lambda_k = 1, \tag{6c}$$

$$\lambda_k \geq 0, \quad \forall k = 1, \ldots, 15,$$

$$y_{ij} \geq 0, \quad \text{and integer}, \quad \forall i = 1, \ldots, 7; \ \forall j = 1, \ldots, 7, \tag{6d}$$

$$r_i \geq 0, \quad \text{and integer}, \quad \forall i = 1, \ldots, 7.$$

Here, since the decision variable y_{ij} is an integer the problem (P2) is formulated into the relaxed form as the problem (A.4) given in Appendix. The probabilistic constraint (3c) of problem (P2) is written as a convex combination of the p-efficient points n_{kj} in constraints (6b) and (6c).

Using LINDO 6.1 (LINDO Systems Inc.) to solve this problem, we achieve the optimal resource allocation strategy. Similarly, we can compute the optimal number of tow trucks assigned to each route for various $Q_{service}$ values which is summarized in Table 5.

Note that the numbers of tow trucks we obtained in Table 5 are very conservative. In practice, the probability of having more than one incident simultaneously is quite low. Without lowering the quality of service, to avoid expensive costs due to the purchase of tow trucks and expensive maintenance cost for a large fleet, what the traffic management center (TMC) usually does is to sign service contract with private companies. For example, to achieve the ninety percent quality of service level, the TMC can probably have 4 tow trucks of its own, which is the maximum number of tow trucks a single incident might request, while the remaining tow trucks could be gathered from private contractors.

Without a surprise, to improve the quality of service more tow trucks are required. However, it is worth to note that the number of tow trucks assigned to each route does not increase simultaneously as higher quality of service is targeted. For instance, if the $Q_{service}$ value increases from 0.5 to 0.9, US 30 keeps the same number of tow trucks while NJ 38 increases its need for more tow trucks rapidly. This happens as a result of the overall travel time from US 30 being much longer than NJ 38. To cover potential incidents, assigning more tow trucks to NJ 38, instead of US 30, is a wiser choice. If the budget and operational limitations can only afford to sign contracts with 21 tow trucks, then the quality of service for the whole system cannot exceed 0.7.

Now, considering a brand-new incident management system, we need to determine the location of depots and number of tow trucks assigned to each depot. Our objective function is still to minimize the incident management cost while subject to certain budget constraints. By assuming that the annual operation cost of a single tow truck is \$10,000, the annual

Table 6 Optimal location of depots and the number of tow trucks assigned for various $Q_{service}$ levels

$Q_{service}$	US 30	NJ 38	NJ 42	I-76	US 130	I-295	I-676	Total trucks	Total depots
0.5		4	5	6				15	3
0.7		6		15				21	2
0.9		6		19				25	2

cost of a depot is \$100,000, and the total annual budget for this tow truck-depot system is \$500,000, then the constraints for this problem are given as below:

$$\begin{aligned}
&\sum_{j=1}^{7} y_{ij} \le r_i, \quad \forall i = 1, \ldots, 7, \\
&\sum_{i=1}^{7} y_{ij} \ge \sum_{k=1}^{15} \lambda_k n_{kj}, \quad \forall j = 1, \ldots, 7, \\
&\sum_{k=1}^{15} \lambda_k = 1, \\
&r_i \le M \cdot d_i, \quad \forall i = 1, \ldots, 7, \\
&\sum_{i=1}^{7} r_i + 10 \sum_{i=1}^{7} d_i \le 50, \\
&\lambda_k \ge 0, \quad \forall k = 1, \ldots, 15, \\
&y_{ij} \ge 0, \quad \text{and integer}, \quad \forall i = 1, \ldots, 7;\ \forall j = 1, \ldots, 7, \\
&r_i \ge 0, \quad \text{and integer}; \quad d_i \text{ binary}, \quad \forall i = 1, \ldots, 7.
\end{aligned} \tag{10}$$

Similarly, we use the LINDO system to solve the problem above. Since the version of LINDO we are using can only handle a maximum 50 integer variables, we relax y's as continuous variables and keep r_i as integer and d_i as binary variables. The y values we obtain in our optimal solution turned out to be integers automatically. Table 6 shows the results for various $Q_{service}$ levels.

We can see the total numbers of tow trucks we obtained in Table 6 are the same as Table 5, but much less depots are needed to reach the same level of quality of service. This result shows that if we could design an incident management system from scratch and locate depots optimally, we can even lower our costs due to the reduced number of depots required.

4 Summary and conclusion

In this paper, we proposed mathematical programming models with probabilistic constraints to account for the future service demands to determine feasible solutions for the incident response problems. The probabilistic constraints are derived from the uncertainty in the number of service vehicles requested by the potential incidents and the number of service vehicles available at the depot. Quality of service is introduced as a way to measure the impact of incident occurrence probabilities and availability of resources in planning for assignment policies. High quality of service means that incident response policy for the current set of incident occurrence probabilities will be able to respond to all incidents in a more reliable way. To put it in another way, high quality of service for the potential incidents is a conservative practice, while low quality of service means high risk. If the anticipated incidents really occur, a response policy based on a very high quality of service will have lower actual response cost compared to a lower quality of service value. We assume the distributions of resources requested by the potential incidents are known, and, from which p-efficient points could be obtained for a given quality of service requirement. Thus, the impact of quality of service is also carefully studied in this paper. The proposed formulation gives us the minimum cost response policy under the pre-selected quality of service requirements given resource and incident occurrence likelihoods. Of course, increasing the level of service requirement, will either increase the minimum response costs or will make the problem infeasible for the current constraints. Then the decision maker can select to increase the amount of available resources or reduce the quality of service requirements meaning that he/she will be willing to take higher risks for potential accidents that are likely to occur according to a probability distribution function estimated using historic data.

Considering the probability distribution of potential incidents over a transportation network, we also constructed probabilistic programming models to address the resource allocation problem for an incident management system. We consider a distributed-depot configuration, in which we need to determine the minimum number of service vehicles allocated to each depot in order to meet the reliability requirements of this system. This offers the planners the capability of considering inherent uncertainty of the system when they attempt to determine the solution that maximizes the return on investment, while remaining within certain budget constraints. In the case study of the South New Jersey road network, we compute the probability distribution of demand for tow trucks along the eight major roads by analyzing two years of incident data in this area. Based on these distributions, probabilistic programming models are constructed, and the minimum number of tow trucks allocated to each major road for various reliability requirements is obtained. For the planning of a new incident management system, our case study shows how to determine the locations of depots and the number of service vehicles assigned to the depots within the budget limits.

The mathematical programming models with probabilistic constraints discussed in this paper are solved using the concepts and algorithms presented in Prekopa et al. (1998). We also developed a computer program to enumerate all of the Pareto efficient (*p-efficient*) points for any multidimensional discrete random variable. Compared to the other models in the literature, our model has the following salient features:

(1) The concept of quality of service offers the flexibility in measuring the impact of potential incidents. This adds to the realism of the model by being able to take into account various incident occurrence probabilities.
(2) Our model can also capture the uncertainty of the resource availability and can help the decision makers to better understand the reliability of the incident management policies they adopt.
(3) More importantly, this model can accurately capture the varying probabilities of simultaneous incident occurrences. The analysis of real-world incident data shows that the incident occurrence is quite random and our model can take this uncertainty into account.

Naturally, the proposed model needs to be further improved to be implemented in real-life applications. Additional incident response data for a larger network need to be collected to better estimate the probability distributions used by the model in (P1). This data include actual characteristics of accidents, number and type of response vehicles associated with each accident and incident as well the arrival and departure time of these resources. This can be a quite involved data collection effort. However, many states have started to collect this type of performance data and it would be possible to have access relatively accurate and detailed incident response data in the near future. Aside from data required for the probabilistic aspects of the model, it can be improved by adding additional terms in the objective function beside the constraints that can capture the expected change in route travel times as a result of accidents. This will be an important enhancement to the introduce model since it will enable the model to capture the network effects of accidents in terms of delays incurred as a result of various types of accidents. Additionally, improvements in terms of adding route travel times would be to represent these travel times as distributions rather than point estimates. This will improve the realism of the model. Moreover, a time dimension that reflects the order in which resources are allocated and the impact of various time-dependent resource assignment strategies on the overall quality of service can be added. However, this type of addition is expected to further complicate the model from both theoretical and computational points of view.

In brief, the models presented in this paper can be used by the State Departments of Transportation to evaluate and improve existing Incident Management programs and to plan the future ones.

There are several opportunities for future research including: (1) model in (P1) should be improved considering the stochastic nature of the response times; (2) collect additional incident response data to develop new case studies especially with the goal of comparing the current practice with the type of decision making approach proposed in this paper that requires a very detailed computational study in addition to long-term incident response and management data collection; (3) conduct a detailed numerical study of computational aspects of these models using additional case studies; (4) a modified version of the model(s) proposed in this paper can be useful for "staffing at peak" question for transit operators under normal as well as rare conditions such as emergency evacuation problems.

Appendix

A.1 Stochastic programming review

Stochastic programming problems, similar to the ones proposed in the previous sections, are formulated in terms of a random vector $\underline{\boldsymbol{\xi}} \in \mathrm{R}^r$, $\underline{\boldsymbol{\xi}} = (\xi_1, \ldots, \xi_r)^T$, $\underline{x} \in \mathrm{R}^n$ and $r \times n$ matrix T. The symbol P denotes probability. If we require that the constraint $T\underline{x} \geq \underline{\xi}$ should hold at least with some given probability $p \in (0, 1)$, rather than for all possible realizations of the right hand side, then the following formulation is obtained:

$$\begin{aligned} \text{Minimize} \quad & \underline{c}^T \underline{x} \\ \text{Subject to} \quad & \mathrm{P}(T\underline{x} \geq \underline{\xi}) \geq p, \\ & A\underline{x} \geq \underline{b}, \\ & \underline{x} \geq \underline{0}, \end{aligned} \tag{A.1}$$

where the first constraint is a probabilistic constraint.

Because the variables in our model are integers, we concentrate on the stochastic programming with discrete random constraints. Let $\underline{\boldsymbol{\xi}}$ be an r-dimensional discrete random vector. Note that the ith element of $\underline{\boldsymbol{\xi}}$, $\underline{\xi}_i$ is a discrete random variable taking particular values z_i, and we assume the number of possible values of this discrete random variable is k_i. Before we continue with the description of the solution approach, we shall first introduce the concept of a p-efficient point, defined in Prekopa (1990) and Prekopa et al. (1998) as follows.

Definition A point $\underline{z}$ is called a p-level efficient point (*p-lep*) of a probability distribution function F, if $F(\underline{z}) \geq p$ and there is no $\underline{y} \leq \underline{z}$, $\underline{y} \neq \underline{z}$ such that $F(\underline{y}) \geq p$, where $p \in (0, 1)$.

Let the set of all p-lep's denoted by $\boldsymbol{Z}_p = \{\underline{z}^{(1)}, \underline{z}^{(2)}, \ldots, \underline{z}^{(N)}\}$, where N shows the number of p-lep's. To enumerate the p-level efficient points for a multidimensional discrete probability distribution, Prekopa et al. (1998) proposed a recursive algorithm. When we enumerate the p-lep's in $\Re^r$ (r-dimensional Euclidean space), it is assumed that an enumeration technique in $\Re^{r-1}$ is available for functions which are not necessarily probability distribution functions in the sense that the sum of all probabilities may be smaller than 1.

With the set of all *p-efficient points*, $\boldsymbol{Z}_p = \{\underline{z}^{(1)}, \underline{z}^{(2)}, \ldots, \underline{z}^{(N)}\}$, where N denotes the number of *p-efficient points*, an equivalent form of (A.1) can be formulated as follows (Prekopa et al. 1998):

$$\begin{aligned} \text{Minimize} \quad & \underline{c}^T \underline{x} \\ \text{Subject to} \quad & T\underline{x} \geq \underline{z}^{(i)}, \quad \forall \underline{z}^{(i)} \in \boldsymbol{Z}_p, \\ & A\underline{x} \geq \underline{b}, \\ & \underline{x} \geq 0. \end{aligned} \tag{A.2}$$

A straightforward way of solving (A.2) is to find all *p-efficient points* and to process all corresponding problems. For example, problem (A.2) can be solved exactly by the solution of N linear programming problems, where the ith linear programming problem has the constraint $T\underline{x} \geq \underline{z}^{(i)}$. If $\underline{x}^{(i)}$ is the optimal solution of the ith problem and $\underline{c}^T \underline{x}^{(i)} = \min_{1 \leq j \leq N} \underline{c}^T \underline{x}^{(j)}$, then $\underline{x}^{(i)}$ is the optimal solution of (A.2). For those problems with a meager number of *p-efficient points* and a small enough matrix T, this method is efficient. If the random variables in the probabilistic constraints are *r-concave* (see Dentcheva et al. 2000), we can obtain an equivalent formulation with a more convenient structure.

Many well-known one-dimensional discrete probability distributions are r-concave distributions-to name a few, Poisson distribution, geometrical distribution, and binomial distribution (Prekopa 1995). Binary random vectors and scalar integer random variables are also r-concave. If $\underline{\boldsymbol{\xi}}$ in (A.1) is an integer random vector and its probability distribution function is r-concave, then an equivalent representation of (A.2) can be obtained as (see Dentcheva et al. 2000):

$$
\begin{aligned}
&\text{Minimize} \quad && \underline{c}^T \underline{x} \\
&\text{Subject to} \quad && A\underline{x} \geq \underline{b}, \\
& && T\underline{x} \geq \underline{v}, \\
& && \underline{v} \geq \sum_{i=1}^{N} \lambda_i \underline{z}^{(i)}, \quad \underline{z}^{(i)} \in \boldsymbol{Z}_p, \\
& && \sum_{i=1}^{N} \lambda_i = 1, \\
& && \underline{x} \geq 0, \quad \underline{v} \text{ integral}, \\
& && \lambda_i \geq 0, \quad i = 1, \ldots, N.
\end{aligned} \tag{A.3}
$$

Moreover, if $T\underline{x}$ is integer, then (A.3) can be simplified further into an equivalent form:

$$
\begin{aligned}
&\text{Minimize} \quad && \underline{c}^T \underline{x} \\
&\text{Subject to} \quad && A\underline{x} \geq \underline{b}, \\
& && T\underline{x} \geq \sum_{i=1}^{N} \lambda_i \underline{z}^{(i)}, \quad \underline{z}^{(i)} \in \boldsymbol{Z}_p, \\
& && \sum_{i=1}^{N} \lambda_i = 1, \\
& && \underline{x} \geq 0, \\
& && \lambda_i \geq 0, \quad i = 1, \ldots, N.
\end{aligned} \tag{A.4}
$$

It is worth noting that if there are no conditions implying that $T\underline{x}$ is integer, then problem (A.4) cannot be used to solve problem (A.2). In fact, the objective value of (A.4) is the lower bound of the original problem by relaxing the constraints of (A.2). Since the number of service vehicles dispatched should always be an integer and the matrix T has integer entries in the resource allocation problem discussed in this paper, solving the simplified formulation (A.4) yields the same optimum solutions as to the original problem. After we obtain p-efficient points using the enumeration algorithm, problems (P1) and (P2) can be simplified into the formulation shown in (A.4), which is solvable by many software packages.

References

Aly, A. A., & White, J. A. (1978). Probabilistic formulation of the emergency services location problem. *Journal of the Operational Research Society*, *29*(12), 1167–1179.

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

Charnes, A., Cooper, W. W., & Symonds, G. H. (1958). Cost horizons and certainty equivalents: an approach to stochastic programming of heating oil by the horizon method. *Management Science*, *4*, 235–263.

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

Chapman, S. C., & White, J. A. (1974). *Probabilistic formulations of emergency service facilities location problems*. Reprint Series, IEOR Dept. Virginia Polyt. Ins. and State University.

Daganzo, C. F. (1994). The cell transmission model: a dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research. Part B*, *4*, 269–287.

Daskin, M. S. (1983). A maximum expected covering location model: formulation, properties and heuristic solution. *Transportation Science*, *17*, 48–70.
Daskin, M. S. (1987). Location, dispatching, and routing models for emergency services with stochastic travel times. In A. Ghosh & G. Rushton (Eds.), *Spatial analysis and location-allocation models* (pp. 224–265) New York: Van Nostrand-Reinhold.
Dentcheva, D., Prekopa, A., & Ruszczynski, A. (2000). Concavity and efficient points of discrete distributions in probabilistic programming. *Mathematical Programming Series A*, *89*, 55–77.
Hakimi, S. L. (1964). Optimum locations of switching centers and the absolute centers and medians of a graph. *Operations Research*, *12*, 450–459.
Kelton, W. D., Sadowski, R. P., & Sadowski, D. A. (2001). *Simulation with ARENA* (2nd ed.). New York: McGraw-Hill.
Lindley, J. A. (1989). *The urban freeway congestion problem* (Transportation Research Circular No. 344). Traffic Management for Freeway Emergencies and Special Events.
Lou, Y., Yin, Y., & Lawphongpanich, S. (2011). Freeway service patrol deployment planning for incident management and congestion mitigation. *Transportation Research. Part C*, *19*, 283–295.
Miller, L. B., & Wagner, H. (1965). Chance-constrained programming with joint constraints. *Operations Research*, *13*, 930–945.
Ozbay, K., & Bartin, B. (2003). Incident management simulation. *Simulation*, *79*(2), 69–82.
Ozbay, K., & Kachroo, P. (1999). *Incident management in intelligent transportation system*. Boston: Artech House Books.
Pal, R., & Bose, I. (2009). An optimization based approach for deployment of roadway incident response vehicles with reliability constraints. *European Journal of Operational Research*, *198*(2), 452–463.
Pal, R., & Sinha, K. (1997). A framework for locating highway incident response vehicles in urban areas. In *INFORMS national meeting*, San Diego, California, US.
Pal, R., & Sinha, K. (2002). Simulation model for evaluating and improving effectiveness of freeway service patrol programs. *Journal of Transportation Engineering*, *128*(4), 355–365.
Petty, K. F. (1997). *Incidents on the freeway: detection and management*. Doctoral dissertation, Department of Electrical Engineering and Computer Science, University of California, Berkeley.
Prekopa, A. (1970). On probabilistic constrained programming. In *Proceedings of the Princeton symposium on math. prog.* (pp. 113–138). Princeton: Princeton University Press.
Prekopa, A. (1973). Contributions to the theory of stochastic programming. *Mathematical Programming*, *4*, 202–221.
Prekopa, A. (1990). Dual method for a one-stage stochastic programming problem with random rhs obeying a discrete probability distribution. *Zeitschrift für Operations-Research*, *34*, 441–461.
Prekopa, A. (1995). *Stochastic programming*. Norwell: Kluwer Academic.
Prekopa, A., Vizvari, B., & Badics, T. (1998). Programming under probabilistic constraint with discrete random variables. In F. Giannessi & T. Rapcsák (Eds.), *New trends in mathematical programming* (pp. 235–257). Norwell: Kluwer Academic.
Proceedings of national conference on traffic incident management: a roadmap to the future, June 2002. http://onlinepubs.trb.org/onlinepubs/archive/conferences/TIM/TIMProceedings.pdf.
Repede, J. F., & Bernardo, J. J. (1994). Developing and validating a decision support system for locating emergency medical vehicles in Louisville, Kentucky. *European Journal of Operational Research*, *75*, 567–581.
ReVelle, C., & Hogan, K. (1989). The maximum availability location problem. *Transportation Science*, *23*, 192–200.
Sherali, H. D., & Subramanian, S. (1999). Opportunity cost-based models for traffic incident response problem. *Journal of Transportation Engineering*, *125*, 176–185.
Sherali, H. D., Desai, J., & Glickman, T. S. (2004). Allocating emergency response resources to minimize risk with equity considerations. *American Journal of Mathematical and Management Sciences*, *24*, 367–410.
Toregas, C., Swain, R., ReVelle, C., & Bergman, L. (1971). The location of emergency service facilities. *Operations Research*, *19*, 1363–1373.
White, J., & Case, K. (1974). On covering problems and the central facilities location problem. *Geographical Analysis*, *6*, 281–293.
Yin, Y. (2006). Optimal fleet allocation of freeway service patrols. *Networks and Spatial Economics*, *6*, 221–234.
Yin, Y. (2008). A scenario-based model for fleet allocation of freeway service patrols. *Networks and Spatial Economics*, *8*(4), 407–417.
Zografos, K. G., Nathanail, T., & Michalopoulos, P. (1993). Analytical framework for minimizing freeway-incident response time. *Journal of Transportation Engineering*, *119*, 535–549.