11e CONGRES INTERNATIONAL DE GENIE INDUSTRIEL – CIGI2015
Québec, Canada
26-28 octobre 2015

Le problème de redéploiement et de préaffectation des véhicules ambulanciers

VALERIE BELANGER[1], ANGEL RUIZ[2], PATRICK SORIANO[3]

[1] Centre interuniversitaire de recherche sur les réseaux d'entreprise, la logistique et le transport (CIRRELT)
Université de Montréal, C.P. 6128, succursale Centre-Ville, Montréal (Québec) H3C 3J7, Canada
valerie.belanger@cirrelt.ca

[2] Université Laval et CIRRELT
Faculté de sciences de l'administration, Université Laval, Québec (Québec), G1K 7P4, Canada
angel.ruiz@osd.ulaval.ca

[3] HEC Montréal et CIRRELT
Département de sciences de la décision, 3000, chemin de la Côte Sainte-Catherine, Montréal (Québec), H3T 2A7, Canada
patrick.soriano@hec.ca

Résumé **- Cet article présente le développement et l'analyse d'un modèle de décision pour la gestion des services préhospitaliers d'urgence (SPU). Le modèle proposé permet de résoudre le problème de redéploiement et de préaffectation des véhicules ambulanciers (PRPA) qui se distingue par l'intégration de ces trois aspects principaux. Tout d'abord, il intègre, en plus des décisions de relocalisation, l'établissement d'une liste ordonnée de préaffectation pour chaque zone de demande qui permettront de guider les décisions de répartition. De plus, il considère explicitement la capacité des véhicules à traiter des demandes urgentes. Enfin, deux familles d'objectifs liées respectivement au service à la population et aux efforts de redéploiement des véhicules sont prises en compte lors de la formulation du modèle. Les résultats obtenus montrent le fait de considérer les listes de préaffectation influence les décisions de localisation et, conséquemment, les performances espérées du système. La capacité des véhicules influence aussi les décisions, mais dans les cas étudiés, ce sont principalement les décisions de préaffectation qui sont touchées par la capacité du système. Enfin, les poids attribués aux différents objectifs ont un impact considérable sur les décisions.**

Mots clés **– Services préhospitaliers d'urgence, véhicules ambulanciers, localisation, redéploiement, préaffectation.**

1 INTRODUCTION

Les services préhospitaliers d'urgence (SPU) représentent une composante importante des systèmes de santé modernes. La mission principale des SPU consiste à répondre adéquatement aux appels de détresse en prodiguant les premiers soins aux personnes concernées et en assurant leur transport, au besoin, vers le service des urgences du centre hospitalier approprié. Les SPU remplissent donc un rôle important au sein des systèmes de santé modernes et leur capacité à répondre adéquatement aux appels peut avoir un impact crucial sur la santé des patients.

De façon générale, la gestion d'un SPU implique deux types de décisions en temps réel : la répartition des appels d'urgence et la localisation des véhicules en attente. Tout d'abord, lorsqu'un appel d'urgence est reçu, le véhicule à envoyer sur les lieux de l'incident devra être déterminé. C'est la répartition des appels. Traditionnellement, on considérera que le véhicule le plus proche est toujours sélectionné [Toro-Diaz et al., 2013]. Cette règle correspond à la règle employée de manière officielle par la plupart des SPU. En pratique, un répartiteur d'expérience pourrait toutefois choisir un véhicule différent lorsque la priorité de l'appel le permet. Dans cet esprit, certains chercheurs se sont intéressés à l'étude de différentes stratégies de répartition pour les appels de priorité moindre, mais nécessitant toujours une attention immédiate. Ainsi, des règles de répartition visant à minimiser les coûts de relocalisation [Gendreau et al., 2001] ou encore à minimiser la dégradation de la capacité de réponse du système [Andersson et al., 2007] ont été proposées, de même que des politiques de répartition basées sur la programmation dynamique [Schmid, 2012]. Schmid [2012] a d'ailleurs montré que, pour les appels de priorité moindre, mais qui requièrent toujours une intervention immédiate, le fait de considérer d'autres règles de répartition peut contribuer à améliorer le service à la population.

Entre l'affectation à deux appels de détresse, les véhicules ambulanciers sont généralement placés en attente à différents postes répartis sur le territoire à desservir. Afin de servir la population dans des délais de temps raisonnables, les véhicules disponibles pour répondre aux appels devront être localisés de manière adéquate. C'est la localisation des véhicules ambulanciers. En comparaison avec le problème de répartition des appels, le problème de localisation ou de déploiement des véhicules a été étudié de manière plus approfondie [Brotcorne et al., 2003 ; Goldberg, 2004 ; Bélanger et al., 2012 ; Başar et al., 2012]. Dans le cas statique, le plan de déploiement établi à la suite de la résolution du problème de localisation demeure inchangé, c'est-à-dire que chaque véhicule retourne au poste d'attente qui lui est attribué une fois sa mission complétée. Plus récemment, différentes études se sont aussi intéressées au problème de redéploiement *multi-période* [Repede et Bernardo, 1994 ; Rajagopalan et al., 2008 ; Başar et al., 2010 ; Schmid et Doerner, 2010 ; Saydam et al., 2013, van den Berg et Aardal, 2015] ou *dynamique* [Gendreau et al., 2001 ; Andersson et al., 2007 ; Gendreau et al., 2005 ; Nair et Miller-Hooks, 2009 ; Maxwell et al., 2009 ; Schmid, 2012 ; Naoum-Sawaja et Elhedhli, 2013 ; Mason, 2013 ; Jagtenberg et al., 2015]. Lorsque le redéploiement *multi-période* est considéré, l'horizon de planification est divisé en périodes correspondant aux variations possibles du nombre de véhicules disponibles, de la demande ou des caractéristiques du système routier. Le redéploiement des véhicules est alors permis, c'est-à-dire la modifiction du poste d'attente des véhicules disponibles, mais seulement aux changements de périodes. Le redéploiement *dynamique* considérera, quant à lui, l'état du système au moment de la prise de décision. Il pourra être résolu en temps réel ou encore *a priori* de façon à déterminer un ensemble de tables de repositionnement correspondant à différents scénarios en ce qui a trait, par exemple, au nombre de véhicules disponibles. Bien que les stratégies de redéploiement proposées permettent d'améliorer les performances globales du système, elle génère aussi des inconvénients liés, entre autres, aux déplacements plus fréquents des véhicules. La sélection d'une stratégie de redéploiement adéquate représentant un compromis intéressant entre le service à la population et les inconvénients qui y sont associés devient alors impérative.

La plupart des études effectuées jusqu'à maintenant ont considéré les décisions de localisation et de répartition de manière indépendante. En effet, la plupart des modèles proposés ne considèrent que le problème de localisation et présument une répartition des appels selon la règle du véhicule le plus proche. Bien que cette hypothèse soit réaliste en pratique, le fait d'envoyer le véhicule le plus proche à un appel de priorité moindre peut, dans certains cas, mener à une dégradation des performances globales du système, ou nécessiter un redéploiement des véhicules disponibles afin de retrouver un niveau de performance adéquat. Envoyer un véhicule différent, mais qui se situe à l'intérieur des délais acceptables, pourrait éviter de tels inconvénients. En pratique, bien que certains répartiteurs semblent le faire de façon instinctive, cela ne se traduit pas formellement dans la plupart des modèles proposés jusqu'à maintenant. L'impact des décisions de répartition sur la localisation des véhicles n'est donc généralement pas pris en compte. Au meilleur de notre connaissance, Toro-Dìaz et al. [2013] ont été parmi les premiers à intégrer explicitement la répartition et la localisation. Ils proposent alors un modèle qui intègre la localisation des véhicules ambulanciers et la répartition des appels selon une liste ordonnée pour chaque zone de demande. Néanmoins, la complexité du modèle qui utilise le modèle de l'hypercube, un modèle basé sur la théorie de files d'attente, afin de représenter le phénomène de congestion lié aux SPU le rend difficile à utiliser afin de résoudre des problèmes de taille réelle. Afin de pallier à ce problème, Toro-Dìaz et al. [2015] ont développé une métaheuristique basée sur la recherche avec tabous afin de résoudre le problème, ce qui permet l'analyse de problème de plus grande taille, soit de l'ordre d'environ 180 zones de demande et 18 ambulances. Dans les deux cas, le problème de localisation est considéré au niveau tactique/stratégique : la relocalisation et les inconvénients qui y sont liés n'y sont pas pris en compte.

Dans le présent article, nous proposons un modèle de décision pour la gestion des SPU considérant le redéploiement des véhicules, mais également une préaffectation anticipative des demandes éventuelles aux véhicules disponibles sous la forme de listes ordonnées. L'objectif de ce modèle consiste à minimiser le temps de réponse espéré tout en limitant les efforts liés au redéploiement (par exemple, le nombre de véhicules redéployés ou la distance de redéploiement). Les inconvénients liés à la relocalisation des véhicules y sont donc explicitement pris en compte. La capacité réelle des véhicules en termes de charge de travail maximale est également considérée lors de la formulation du modèle. En opposition au modèle de Toro-Dìaz et al. [2013], notre modèle considère l'incertitude liée à la disponibilité des véhicules implicitement lors de la définition du temps de réponse espéré et de la charge de travail espérée. Le modèle résultant demeure donc linéaire et pourra être utilisé afin de traiter des instances de taille réaliste.

Cet article se divise donc comme suit. Tout d'abord, le problème de redéploiement et de préaffectation des véhicules ambulanciers (PRPA), de même que ses caractéristiques principales, sont décrits et discutés. La formulation de ce problème est ensuite présentée, suivie de l'analyse du modèle grâce aux résultats obtenus lors de la phase d'expérimentation.

2 LE PROBLEME DE REDEPLOIEMENT ET DE PREAFFECTATION DES VEHICULES AMBULANCIERS

Le problème de redéploiement et de préaffectation des véhicules ambulanciers (PRPA) consiste à déterminer la localisation des véhicules disponibles pour répondre à des appels urgents de même qu'une liste de préaffectation pour chaque zone de demande. La liste de préaffectation comportera une liste ordonnée de véhicules à affecter aux demandes apparaissant dans une zone donnée. Ainsi, si un incident survient dans une zone, le premier véhicule sur la liste sera envoyé sur les lieux de l'incident si celui-ci est disponible. Dans le cas contraire, le deuxième sera envoyé, et ainsi de suite. Si aucun véhicule n'est disponible sur la liste, le véhicule disponible le plus proche sera envoyé sur les lieux de l'incident. Si aucun véhicule n'est disponible, l'appel est placé en file d'attente ou redirigé vers un autre service. La localisation des véhicules pourra engendrer ou non la relocalisation des véhicules en attente vers des postes d'attente différents de ceux qu'ils occupent au moment du

redéploiement. Conséquemment, des coûts de redéploiement pourront être encourus.

Trois caractéristiques principales distinguent le PRPA des problèmes étudiés jusqu'à présent. Tout d'abord, le PRPA considère, en plus des décisions de redéploiement, l'établissement de listes de préaffectation ordonnées pour chaque zone de demande. Ces listes de préaffectation permettront de guider les décisions de répartition, principalement pour les demandes de priorité moindre, mais nécessitant toujours une intervention immédiate. L'établissement des listes de préaffectation permettra aussi de considérer plus explicitement la capacité des véhicules, ou autrement dit, leur charge de travail. Enfin, deux familles d'objectifs liées respectivement au service à la population et aux efforts de redéploiement des véhicules sont prises en compte lors de la formulation du modèle. Le temps de réponse espéré, mesure inspirée de la définition de la couverture espérée [Daskin, 1983], sera utilisé afin de estimer le service à la population, tandis que le temps de redéploiement total permet de mesurer les efforts de relocalisation.

2.1 Formulation

Le problème de redéploiement et de préaffectation des véhicules ambulanciers (PRPA) se définit sur un graphe $G=(V,E)$ où $V=I \cup J$, $I=\{v_1,\ldots,v_n\}$ et $J=\{v_{n+1},\ldots,v_{n+m}\}$ représentent respectivement l'ensemble des zones de demande et des postes d'attente potentiels, et où $E=\{(v_i, v_j): v_i, v_j \in V\}$ est l'ensemble des arêtes reliant les sommets du graphe. Une zone de demande se définit comme un regroupement de population représenté par son centroïde et possédant une densité de population ou une demande donnée. Un site ou poste d'attente potentiel se définit plutôt comme un lieu physique où pourront être localisés un ou plusieurs véhicules et à partir duquel les véhicules se déplaceront pour rejoindre les différents appels de détresse. Ainsi, à chaque arête $(v_i, v_j) \in E$ est associé un temps de déplacement t_{ij} et à chaque sommet $v_i \in I$ est associée une densité de population ou une demande d_i. De plus, à chaque sommet $v_j \in J$ est associé un nombre maximal de véhicules p_j pouvant y être placés en attente. On définira également K, l'ensemble des véhicules disponibles, et Z, l'ensemble des positions considérées sur la liste de préaffectation, $|Z|$ représentant la taille de la liste de préaffectation. Ainsi, à chaque véhicule $k \in K$ sont associés un paramètre λ_j^k qui vaut 1 si le véhicule k est localisé en $v_j \in J$ avant la relocalisation, et 0 autrement, un paramètre θ^k qui vaut 1 si le véhicule k est considéré dans le calcul des coûts de relocalisation, et 0 autrement, et une pénalité M_{j1j2}^k associée au déplacement du véhicule k de v_{j1} en v_{j2}. On considérera également q, le taux d'occupation des véhicules, et W, la capacité ou la charge de travail maximale en termes de nombre d'interventions qu'un véhicule peut servir sur une période donnée. On présumera que tous les véhicules ont la même capacité et le même taux d'occupation. Enfin, ξ_s et ξ_r représentent les poids ou l'importance relative accordée respectivement aux objectifs liés au service à la population et aux inconvénients associés au redéploiement des véhicules lors de la prise de décision.

Le PRPA considère deux types de décisions : des décisions de localisation qui attribuent un poste d'attente à chaque véhicule disponible (c'est-à-dire en attente, nouvellement libéré, en début de quart de travail, en retour de pause) et des décisions de préaffectation, qui affectent une liste ordonnée de véhicules à chaque zone de demande. Afin de considérer adéquatement ces différentes décisions, trois groupes de variables ont été définis soit: x_{j1j2}^k, une variable binaire qui vaut 1 si le véhicule k se déplace du site $v_{j1} \in$ J vers le site $v_{j2} \in$ J à la suite du redéploiement/relocalisation, et 0 autrement (lorsque $j1=j2$, aucune relocalisation n'est requise); w_i^{zk} une variable binaire qui vaut 1 si le véhicule k est le z-ième appelé sur la liste de préaffectation de la zone de demande située en $v_i \in I$, et 0 autrement; et y_{ij}^{zk}, une variable binaire qui vaut 1 si le véhicule k, localisé en $v_j \in J$ à la suite de la relocalisation, est à la z-ième position sur la liste de préaffectation de la zone de demande située en $v_i \in I$, et 0 autrement. En considérant les paramètres et les variables tels que définis ci-haut, le PRPA se formule de la manière suivante:

$$\min \xi_s \sum_{i\epsilon I}\sum_{z\epsilon Z}\sum_{k\epsilon K}\sum_{j\epsilon J}(1-q)q^{z-1}d_i t_{ji} y_{ij}^{zk} + \xi_r \sum_{k\epsilon K}\sum_{j1\epsilon J}\sum_{j2\epsilon J}\theta^k t_{j1j2} x_{j1j2}^k$$

sous les contraintes :

(1) $\sum_{j1\in J}\sum_{j2\in J}\lambda_{kj1}x_{j1j2}^k = 1, \forall k \in K,$
(2) $\sum_{j1\in J}\sum_{j2\in J}(1-\lambda_{kj1})x_{j1j2}^k = 0, \forall k \in K,$
(3) $\sum_{k\in K}\sum_{j1\in J} x_{j1j2}^k \le p_j, \forall j2 \in J,$
(4) $\sum_{z\in Z}\sum_{i\in I}(1-q)q^{z-1}d_i w_i^{zk} \le W, \forall k \in K,$
(5) $\sum_{z\in Z} w_i^{zk} \le 1, \forall k \in K, \forall i \in I,$
(6) $\sum_{k\in K} w_i^{zk} = 1, \forall z \in Z, \forall i \in I,$
(7) $y_{ij2}^{zk} \le \sum_{j1\in J} x_{j1j2}^k, \forall z \in Z, \forall i \in I, \forall k \in K, \forall j2 \in J,$
(8) $w_i^{zk} = \sum_{j\in J} y_{ij}^{zk}, \forall i \in I, \forall z \in Z, \forall k \in K,$
(9) $x_{j1j2}^k \in \{0,1\}, \forall j1 \in J, \forall j2 \in J, \forall k \in K,$
(10) $w_i^{zk} \in \{0,1\}, \forall z \in Z, \forall i \in I, \forall k \in K,$
(11) $y_{ij2}^{zk} \in \{0,1\}, \forall z \in Z, \forall i \in I, \forall k \in K, \forall j \in J.$

La fonction objectif du PRPA présentée ci-haut consiste à minimiser les temps de réponse espérés totaux puis les temps nécessaires au redéploiement. Les contraintes (1) et (2) visent à assurer que chaque véhicule disponible soit localisé à un poste d'attente, tandis que les contraintes (3) garantissent que le nombre maximal de véhicules pouvant être localisés à un poste d'attente donné soit respecté. Les contraintes (4) assurent, quant à elles, la satisfaction de la capacité des véhicules à traiter des demandes ou, autrement dit, assurent que la charge de travail espérée d'un véhicule soit inférieure ou égale à sa capacité. Le modèle considère aussi qu'un véhicule ne pourra occuper plus d'une position sur la liste de préaffectation d'une zone de demande donnée (5) et qu'exactement un véhicule devra se retrouver à chaque position de la liste de préaffectation d'une zone de demande (6). Enfin, les contraintes (7) et (8) assurent le lien entre les variables de décision, tandis que les contraintes (9) à (11) assurent leur intégralité.

3 EXPERIMENTATION

Une série d'expérimentations a été effectuée afin d'analyser le modèle proposé pour le redéploiement et la préaffectation des véhicules ambulanciers. Plus précisément, les résultats obtenus lors de la phase d'expérimentation permettront d'analyser l'impact des décisions de préaffectation sur les décisions de relocalisation, de même que celui de considérer la capacité des véhicules sur les solutions obtenues et les performances prédites du système, et ce, pour un ensemble varié de paramètres. Deux groupes d'instances ont été générés afin de valider et d'analyser le modèle.

Le premier groupe d'instances a été généré aléatoirement à partir d'un territoire structuré comportant un nombre réduit de zones de demandes. Le territoire à desservir, comportant 24 zones de demande, a été défini sur un rectangle 16X20 km. À chaque zone de demande est associée une demande d_i qui varie selon deux cas. Dans le premier cas, *Cas U*, la demande associée à chaque zone suit une distribution uniforme de borne inférieure 1 et de borne supérieure 100. Ce cas permettra de représenter un contexte urbain avec des zones à plus forte et à plus faible densité de population. Dans le deuxième cas, *Cas R*, la demande est constante pour toutes les zones de demande, soit d_i=50, ce qui permettra de représenter un contexte rural où la demande est répartie plus uniformément sur le territoire à desservir. Chaque zone de demande comporte un poste d'attente potentiel de capacité 1. Le nombre de véhicules disponibles pour répondre aux appels variant entre 3 et 4 et leur capacité est fixée à 425 interventions. Cette valeur a été déterminée en tenant compte du profil de la demande générée. Le taux d'occupation global considéré est de 0,4. Enfin, la matrice de temps de déplacement a été définie à partir des distances euclidiennes entre deux localisations et d'une vitesse constante.

Le deuxième groupe d'instances a été généré à partir d'un cas réel. Ces instances se divisent en trois groupes correspondant à trois différentes tailles de problèmes : réduite (30 zones de demande), moyenne (149 zones de demande) et réelle (595 zones de demande). Peu importe la taille du problème, les instances ont été définies à partir du territoire ou d'un sous-ensemble du territoire des villes de Montréal et de Laval (banlieue nord de la ville de Montréal), le plus grand centre urbain de la province de Québec. Ces instances visent donc à analyser le modèle dans un contexte plus réaliste. Le territoire utilisé pour la génération des instances est représenté à la Figure 1.

Figure 1. Territoire – Instances réelles

À chaque zone de demande correspond une densité de population de même qu'une demande espérée. La demande espérée pour une zone de demande est définie comme le nombre d'interventions prévues ou contre lesquelles on souhaite se prémunir pour une période donnée. Naturellement, la demande prévue et la capacité des véhicules sont influencées par la longueur de la période choisie. Dans le cas de cette étude, la demande prévue pour une période donnée a été tirée à partir des données simulées. Ainsi, pour chaque cas de figure (réduit, moyen et réel), une liste d'appels a été générée tel que décrit dans Kergosien et al. [2015]. La liste des appels générés a ensuite été analysée plus en détail afin de déterminer le profil de demande pour chaque zone considérée. L'ensemble des postes d'attente potentiels où les véhicules pourront être placés varie, quant à lui, de 30 à 218 postes d'attente répartis sur le territoire à desservir, incluant les postes d'attente choisis arbitrairement, les centres hospitaliers et les centres opérationnels. Pour l'instance de taille réduite, chaque zone de demande inclut un poste d'attente potentiel, tandis que pour l'instance de moyenne taille, 48 zones sont sélectionnés de manière aléatoire pour devenir candidate à recevoir un véhicule en attente. Enfin, pour l'instance de taille réelle, 218 zones de demande inclut un poste d'attente potentiel. La capacité de ces localisations a été fixée à 1. La matrice comportant les temps de déplacement entre les différentes localisations sur le territoire à desservir a été extraite de la matrice comportant les temps réels entre les différents centroïdes des zones. Selon la taille du problème, le nombre de véhicules ou d'équipes de travail disponibles variera de 2 à 60 pour la période considérée, correspondant à différents scénarios allant de 7,5 à 20 zones par véhicule. Lorsque pertinent, la localisation d'un véhicule avant la relocalistion sera déterminée de trois manières : soit aléatoirement, soit à partir d'un sous-ensemble des postes d'attente optimaux ou quasi-optimaux pour un nombre de véhicules plus élevé, soit une combinaison des deux. La capacité des véhicules a été fixée à 5 interventions par période, en considérant un taux d'occupation de 0,5. Ces deux valeurs ont été déterminées à la suite de l'analyse des données générées et des résultats obtenus par simulation.

Afin de valider et d'analyser plus en détail les différentes caractéristiques du modèle, confronté à des situations variées en termes de territoire et de profil de demande, différents tests ont été menés grâce aux différentes instances générées. Les résultats

obtenus sont discutés dans la présente section. Chaque tableau de résultats présente d'abord les caractéristiques principales de l'instance et du scénario considéré, puis reporte les localisations obtenues à la suite de la résolution du modèle, le temps de réponse correspondant, le temps de calcul (Temps) et l'écart entre la solution trouvée et la solution optimale (GAP). Afin d'assurer une comparaison adéquate des scénarios considérés, le temps de réponse total (TR^T), incluant le temps de réponse total correspondant à tous les véhicules disponibles, les véhicules qui ne figurant pas sur les listes de préaffectation étant alors classés en fonction de la distance pour chaque zone de demande, et le temps nécessaire pour répondre à un appel placé en file d'attente ou redirigé vers un autre service. Dans ce cas, le temps de réponse est fixé à 360 secondes pour U et R et à 220 secondes pour R30, R149 et R595. Pour les instances aléatoires, cette valeur a été déterminée de manière arbitraire, tandis que pour l'instance pseudo-réelle, elle a été extraite des résultats des études de simulation présentés [Bélanger et al., 2014]. Tous les temps de réponse sont reportés en secondes. Le temps de réponse moyen par intervention (TR) est également donné, entre parenthèses, en secondes par intervention. Les temps de calcul et les GAP sont reportés respectivement en secondes et en pourcentage.

Tous les résultats ont été obtenus à l'aide de CPLEX 12.5 en imposant un temps limite de 28 800 secondes. Les tests ont été lancés à partir d'une machine Windows avec huit cœurs et 15 Go de RAM, installée sur un serveur équipé d'un processeur AMD Opteron 6328.

3.1 Impact des listes de préaffectation

Les résultats de l'expérimentation menée en considérant des listes de préaffectation de taille variée sont présentés ici, et ce, pour toutes les instances de taille réduite (U, R et R30). En analysant les résultats obtenus, il est possible d'évaluer dans quelle mesure le fait de considérer plusieurs véhicules au sein des listes de préaffectation a un impact sur les décisions de localisation, et par conséquent, sur les performances espérées du système.

Tableau 1. Impact des listes de préaffectation (U)

|Z|	Décisions de localisation	TR^T [1000s]	TR [s/int]	Temps [s]	GAP [%]
1	{5,8,16,19}	575,6	428	143	0,00
2	{8,11,14,17}	532,0	399	276	0,00
3	{8,11,14,16}	530,5	397	28800	5,28
4	{9,11,14,16}	530,5	397	28800	5,97

Tableau 2. Impact des listes de préaffectation (R)

|Z|	Décisions de localisation	TR^T [1000s]	TR [s/int]	Temps [s]	GAP [%]
1	{2,5,14,17}	565,9	423	2	0,00
2	{8,11,14,17}	490.8	367	793	0,00
3	{8,11,14,17}	490,8	367	28800	12,99
4	{8,11,14,17}	501,2	375	28800	18,27

Tableau 3. Impact des listes de préaffectation (R30)

|Z|	Décisions de localisation	TR^T [s]	TR [s/int]	Temps [s]	GAP [%]
1	{2,6,20,30}	1892	281	2	0,00
2	{2,4,11,20}	1822	271	1731	0,00
3	{8,11,14,16}	1822	271	28800	8,76
4	{9,11,14,16}	1822	271	28800	9,70

En observant les résultats présentés aux tableaux 1, 2 et 3, trois principaux constats peuvent être émis. Tout d'abord, il est possible de constater que, pour tous les cas étudiés, les listes de préaffectation ont un impact important sur les décisions de localisation. Ce constat est particulièrement vrai lorsque l'on inclut deux véhicules sur les listes de préaffectation plutôt qu'un seul. L'impact devient moins important lorsque trois ou quatre véhicules sont considérés : l'ajout d'un véhicule supplémentaire sur la liste de préaffectation a une contribution moins importante au sein de la fonction objectif ce qui engendre une moins grande variation en termes de localisation des véhicules. Ce constat confirme par le fait même la pertinence de la notion de double couverture proposée par plusieurs auteurs [Hogan et ReVelle, 1986 ; Gendreau et al., 1997], tout en permettant une distinction plus fine entre les différentes solutions grâce à la définition du temps de réponse espéré. Il est également possible d'observer que les performances prédites du système s'améliorent aussi lorsque l'on considère les listes de préaffectation. Selon l'instance considérée, une amélioration allant jusqu'à 56 secondes du temps de réponse par intervention peut être notée, lorsque deux véhicules sont présents sur les listes plutôt qu'un seul. Cette réduction est toutefois moins notable lorsque l'on considère trois ou quatre véhicules sur les listes de préaffectation. Enfin, les temps de calcul augmentent de manière considérable avec la taille des listes. Pour cette raison, pour l'ensemble des tests subséquents, la taille des listes de préaffectation a été fixée à 2, ce qui représente à notre avis, un compromis intéressant entre l'amélioration des décisions et le temps de calcul.

Des résultats similaires ont pu être observées pour les mêmes instances, mais en prenant en compte 3 véhicules. Pour l'instance R149, les temps de calcul deviennent considérable lorsque 2 véhicules sont inclus sur les listes de préaffectation, mais le comportement général du modèle demeure le même. Enfin, l'instance R595 n'a pu être résolue à l'intérieur du temps limite imposé. En effet, les outils en place pour la résolution du modèle ne permettent pas de résoudre l'instance R595. Nous nous limiterons donc à la discussion des résultats pour les instances de taille réduite et moyenne. Notons toutefois que l'instance de taille moyenne, soit R149, demeure de taille comparable à plusieurs instances réelles traitées dans la littérature, notamment celles considérées dans [Toro-Dìaz et al., 2015].

3.2 Impact de la capacité du système sur les décisions de localisation et de préaffectation

La capacité du système varie en fonction de deux paramètres : le nombre de véhicules disponibles et la capacité des véhicules. Différentes combinaisons des valeurs de |K| et de W ont donc été prises en compte afin de définir différents scénarios en termes de capacité du système. Ces scénarios permettront d'analyser l'impact de la capacité sur les décisions de localisation et de préaffectation, et ce, pour les trois cas de figure U, R et R30.

Dans tous les cas, W=2400 représente un système sans contrainte de capacité.

Tableau 4. Impact de la capacité du système (U)

|K|	W	Décisions de localisation	TR^T [1000s]	TR [s/int]	Temps [s]	GAP [%]
3	400	{11,14,16}	599,8	449	273	0,00
3	425	{11,14,16}	597,3	447	136	0,00
3	500	{11,14,16}	596,9	446	213	0,00
3	2400	{11,14,16}	596,9	446	277	0,00
4	400	{8,11,14,17}	532,9	399	1544	0,00
4	425	{8,11,14,17}	532,9	399	277	0,00
4	500	{8,11,14,17}	532,9	399	1148	0,00
4	2400	{8,11,14,17}	532,9	399	985	0,00

Tableau 5. Impact de la capacité du système (R)

|K|	W	Décisions de localisation	TR^T [1000s]	TR [s/int]	Temps [s]	GAP [%]
3	400	{10,14,17}	558,54	418	771	0,00
3	425	{10,14,17}	558,45	418	302	0,00
3	500	{8,11,15}	558,45	418	653	0,00
3	2400	{10,14,17}	558,45	418	6301	0,00
4	400	{8,11,14,17}	490,8	367	983	0,00
4	425	{8,11,14,17}	490,8	367	793	0,00
4	500	{8,11,14,17}	490,8	367	2146	0,00
4	2400	{8,11,14,17}	490,8	367	7018	0,00

Tableau 5. Impact de la capacité du système (R30)

|K|	W	Décisions de localisation	TR^T [s]	TR [s/int]	Temps [s]	GAP [%]
2	2	Non-réalisable				
2	3	{2,11}	2168	322	29	0,00
2	4	{2,11}	2168	322	25	0,00
2	5	{2,11}	2168	322	27	0,00
2	2400	{2,11}	2168	322	25	0,00
3	2	{2,7,20}	1972	293	336	0,00
3	3	{2,7,20}	1972	293	2074	0,00
3	4	{2,7,20}	1972	293	464	0,00
3	5	{2,7,20}	1972	293	460	0,00
3	2400	{2,7,20}	1972	293	462	0,00
4	2	{2, 4, 11, 20}	1822	271	1146	0,00
4	3	{2, 4, 11, 20}	1822	271	793	0,00
4	4	{2, 4, 11, 20}	1822	271	1731	0,00
4	5	{2, 4, 11, 20}	1822	271	1731	0,00
4	2400	{2, 4, 11, 20}	1822	271	1790	0,00

En analysant les résultats présentés aux tableaux 4, 5 et 6, il est possible de constater que, pour tous les cas étudiés, la capacité du système n'a pas d'impact réel sur les décisions de localisation lorsque l'on compare les localisations obtenues en considérant un même nombre de véhicules, mais pour différentes valeurs de la capacité. Lorsque le nombre de véhicules varie, mais que la capacité demeure la même, les décisions de localisation changent, c'est-à-dire les localisations obtenues en considérant |K| véhicules n'est pas un sous-ensemble des localisations obtenues pour |K|+1 véhicules. Les performances s'améliorent dans tous les cas lorsque le nombre de véhicules augmente. Enfin, bien que les localisations ne semblent pas influencées par la valeur de la capacité, il est possible de noter que lorsque la capacité des véhicules est limitée, les listes de préaffectation s'ajustent afin de prendre en compte la capacité réelle des véhicules à traiter des demandes urgentes, ce qui permet d'obtenir une meilleure estimation des performances du système.

3.3 Différentes versions de la fonction objectif

Les résultats présentés jusqu'à maintenant ont été déterminés en ne considérant que la minimisation du temps de réponse total (ou de manière équivalente, du temps de réponse moyen). Aucune autre mesure de la qualité du service n'a été évaluée et aucun coût de relocalisation n'a été pris en compte. De cette manière, il a été possible d'analyser et de mieux comprendre les caractéristiques du modèle, dans un même cadre de comparaison. Afin de compléter l'analyse du modèle proposé, il importe maintenant d'aller un peu plus loin dans l'analyse de la fonction objectif. C'est ce que nous ferons dans la présente sous-section.

Dans un premier temps, les objectifs en lien avec la qualité du service sont analysés ici. La minimisation du temps de réponse total permet d'assurer un bon service global, mais peut aussi mener à des zones de demande pauvrement desservies. Afin d'améliorer l'équité entre les zones de demande, il est possible de considérer, au sein de la fonction objectif, la minimisation du temps de réponse maximal, plutôt que celle du temps de réponse global. Cela aura pour effet d'améliorer le service pour la zone la moins bien servie, au détriment du temps de réponse total, mais dans quelle mesure? Différents tests ont donc été menés afin de répondre à cette question et d'évaluer les différences entre les solutions obtenues en considérant ces deux objectifs, et ce, pour U, R et R30. De plus, une mesure alternative, soit la somme du temps de réponse moyen et du temps de réponse maximal, a également été évaluée. Cette nouvelle fonction objectif permettra d'obtenir un compromis entre les performances globales du système et ses pires performances. Les tableaux 7, 8 et 9 présentent les résultats obtenus en considérant les trois fonctions objectif: temps de réponse moyen (TR_{moy}), temps de réponse maximal (TR_{max}) ou la somme des deux (TR_{m+m}). Pour chaque scénario considéré, la valeur de la fonction objectif est présentée (en gras), puis les autres mesures, calculées en fonction de la solution trouvée, sont reportées. Dans tous les cas, les coûts associés aux effort de relocalisation ne sont pas considérés, c'est-à-dire que $\xi_r = 0$.

Tableau 7. Impact de la fonction objectif (U)

Objectif	|K|	Décisions de localisation	TR_{moy} [s/int]	TR_{max} [s/int]	TR_{m+m} [s/int]
TR_{moy}	3	{11,14,16}	**319,8**	613,2	933,6
TR_{moy}	4	{8,11,14,17}	**273,6**	385,9	659,7
TR_{max}	3	{4, 13, 22}	409,2	**532,8**	942,0
TR_{max}	4	{5,8,14,23}	316,2	**385,9**	700,2
TR_{m+m}	3	{10,13,16}	328,8	547,2	**876,0**
TR_{m+m}	4	{8,11,14,17}	273,6	385,9	**659,7**

Tableau 8. Impact de la fonction objectif (R)

Objectif	|K|	Décisions de localisation	TR_{moy} [s/int]	TR_{max} [s/int]	TR_{m+m} [s/int]
TR_{moy}	3	{8,11,15}	**335,4**	576,0	911,2
TR_{moy}	4	{8,11,14,17}	**282,0**	385,9	667,9
TR_{max}	3	{3,8,20}	411,4	**532,8**	944,2
TR_{max}	4	{2,5,20,23}	340,6	**385,9**	726,5
TR_{m+m}	3	{7,10,16}	344,7	547,2	**891,9**
TR_{m+m}	4	{8,11,14,17}	282,0	385,9	**667,9**

Tableau 9. Impact de la fonction objectif (R30)

Objectif	|K|	Décisions de localisation	TR_{moy} [s/int]	TR_{max} [s/int]	TR_{m+m} [s/int]
TR_{moy}	2	{2,11}	**125,7**	261,0	386,7
TR_{moy}	3	{2,7,20}	**103,7**	288,8	345,4
TR_{moy}	4	{2,4,11,20}	**90,5**	229,0	319,5
TR_{max}	2	{7,30}	156,6	**212,3**	368,9
TR_{max}	3	{1,3,9}	146,6	**171,0**	317,6
TR_{max}	4	{7,9,28,29}	121,6	**153,0**	274,6
TR_{m+m}	2	{7,30}	152,1	212,3	**363,3**
TR_{m+m}	3	{4,10,29}	122,7	172,8	**295,5**
TR_{m+m}	4	{2,7,20,29}	96,6	159,0	**255,6**

Les résultats présentés aux tableaux 7, 8 et 9 confirment que le choix de la fonction objectif influence fortement les localisations optimales et, par le fait même, les performances globales et maximales obtenues. Naturellement, le choix de minimiser les performances moyennes amène une dégradation des pires performances, et inversement. Lorsque l'on considère la minimisation du temps de réponse moyen, le temps de réponse maximal correspondant dépasse de 0 % à 79 % le temps de réponse maximal optimal. Lorsque le temps de réponse maximal est plutôt optimisé, les performances moyennes se dégradent de 20 % à 34 % selon l'instance et le scénario considéré, ce qui demeure non négligeable. Le fait de considérer une combinaison des deux mesures permet d'obtenir des solutions intermédiaires qui se rapprochent davantage des deux objectifs: on parle alors d'une dégradation de 0 % à 21 % pour le temps de réponse moyen et de 0 % à 5 % pour le temps de réponse maximal.

Dans un deuxième temps, l'impact des coûts liés aux efforts de relocalisation sur les décisions de localisation est analysé plus en détail. Pour ce faire, différents scénarios ont été créés en considérant un ensemble de valeurs variées pour ξ_s et ξ_r. Pour chaque cas, les valeurs de ξ_s sont d'abord fixées à 1, 0,75, 0,5, 0,25 et 0, puis les valeurs de ξ_r sont fixées de sorte que $\xi_s + \xi_r = 1$. À la suite d'une expérimentation préliminaire, nous avons choisi d'inclure aussi le cas $\xi_s = 0{,}875$ et $\xi_r = 0{,}125$ afin de fournir une meilleure analyse du compromis entre le temps de réponse et le temps total de relocalisation. Cette analyse est effectuée ici pour l'instance R30, en prenant en compte différents plans de localisation initiale, c'est-à-dire avant le redéploiement. En effet, la localisation initiale des véhicules aura un impact important sur les solutions finales. Afin de représenter un éventail varié de situations, pour chaque scénario en termes de nombre de véhicules, différents plans de localisation initiale ont été pris en compte. Ces plans de localisation ont été déterminés de trois manières: en sélectionnant la localisation initiale des véhicules parmi les localisation optimales dans le cas où le nombre de véhicules disponibles est égal à |K|+1, en sélectionnant les localisations initiales de manière aléatoire, ou grâce à une combinaison des deux. Les figures 2, 3 et 4 présentent les résultats obtenus respectivement dans les cas où 2, 3 ou 4 véhicules sont disponibles pour répondre aux appels d'urgence. Sur chaque figure, le temps de réponse total est tracé en fonction du temps total de relocalisation, et ce, pour différents plans de localisation initiale. Chaque point reporté sur le graphique correspond à la solution obtenue pour pour différentes valeurs de ξ_s et ξ_r.

Figure 2. R30 – 2 véhicules

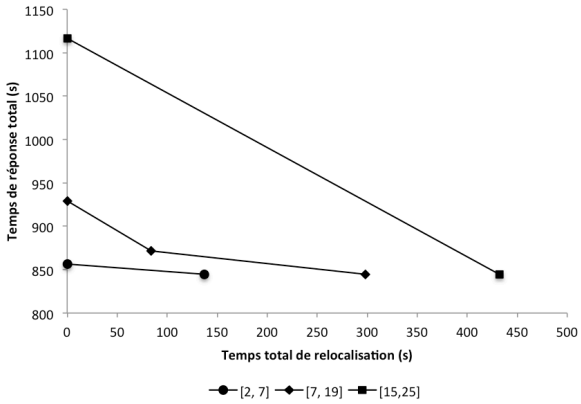


Figure 3. R30 – 3 véhicules

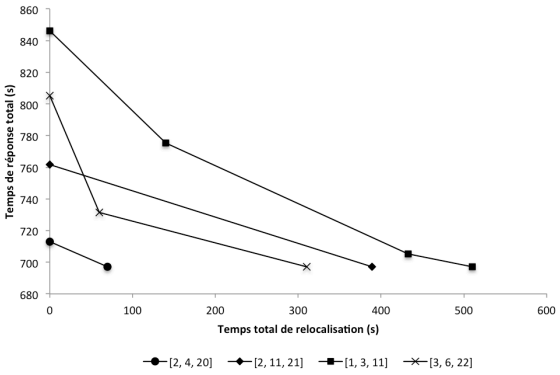
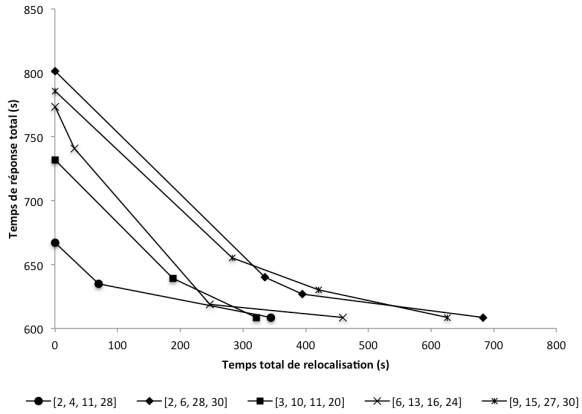


Figure 4. R30 – 4 véhicules



Les résultats obtenus pour différentes valeurs de ξ_s et ξ_r permettent d'abord de conclure que, clairement, le poids accordé à chacun des objectifs a un impact sur les décisions de localisation et de relocalisation. On a aussi observé que, pour les scénarios considérés, les efforts de relocalisation prennent rapidement le dessus sur le gain en performance au fur et à mesure que la valeur de ξ_s diminue, c'est-à-dire que rapidement, le gain en performance ne justifie plus les efforts de relocalisation. Nous avons d'ailleurs choisi d'inclure à la présentation des résultats le cas $\xi_s = 0{,}875$ et $\xi_r = 0{,}125$ de manière à fournir une meilleure appréciation du compromis entre le temps de réponse et le temps total de relocalisation. En effet, dans presque tous les cas, lorsque $\xi_s \leq 0{,}5$, les efforts de relocalisation ne justifient pas les gains en performance. La localisation initiale des véhicules avant le redéploiement a aussi un impact sur les décisions de relocalisation. Plus les localisations initiales sont proches des localisations optimales pour |K|+1, moins le gain en performance ne justifie la modification des postes d'attente des véhicules en attente. En effet, on observe que lorsque toutes les localisations initiales appartiennent à l'ensemble des localisations optimales dans le cas où |K|+1 véhicules sont disponibles, la solution courante, sans relocalisation, demeure de bonne qualité. On parle alors d'un écart de 1,4 % par rapport à la solution optimale dans le cas où 2 véhicules sont disponibles, de 2,3 % pour 3 véhicules et de 8,7 % pour 4 véhicules. Néanmoins, plus le nombre de véhicules est grand et donc, plus le système dispose d'une certaine flexibilité, plus la stratégie de « ne rien faire » est perdante. Les gains potentiels à considérer la relocalisation sont plus importants plus le nombre de véhicules est grand, et ce, peu importe les localisations initiales. D'un point de vue pratique, dans le contexte étudié, il ne devient vraiment intéressant de modifier la localisation des véhicules en attente que si le plan de localisation actuel s'écarte davantage du plan optimal pour |K|+1 véhicules. Autrement dit, à la suite de l'affectation d'un véhicule, il ne devient intéressant de revoir la localisation des véhicules restants que si leur localisation est considérablement différentes des localisations optimales pour |K|+1 véhicules.

En observant maintenant plus en détail aux figures 1, 2 et 3, il est possible de voir clairement que dès que l'on accepte de modifier un peu le plan de localisation actuel, la qualité de la solution s'améliore dans une plus grande proportion. Il est donc intéressant de modifier la localisation des véhicules (surtout si on s'écarte beaucoup du plan optimal dans le cas où |K|+1 véhicules sont disponibles) afin de déterminer une bonne solution, mais les efforts supplémentaires en termes de temps de relocalisation ne justifient pas toujours le gain en performance par la suite. Ces conclusions se vérifient aussi dans le cas R149.

4 CONCLUSION

Cet article a abordé le développement et l'analyse d'un modèle de décision pour la gestion des SPU considérant le redéploiement des véhicules, mais également une préaffectation anticipative des demandes éventuelles aux véhicules disponibles. Le PRPA tel que proposé dans ce chapitre se distingue des problèmes étudiés jusqu'à présent par trois aspects principaux. Tout d'abord, il intègre, en plus des décisions de relocalisation, l'établissement d'une liste ordonnée de préaffectation pour chaque zone de demande. Il considère aussi de manière explicite la capacité réelle ou la charge de travail des véhicules. Enfin, deux familles d'objectifs liées respectivement au service à la population et aux efforts de redéploiement des véhicules sont prises en compte lors de la formulation du modèle.

Les résultats obtenus ont permis de montrer que le fait de considérer plus d'un véhicule dans les listes de préaffectation influence les décisions de localisation et, conséquemment, les performances du système. La capacité des véhicules influence aussi les décisions, mais dans les cas étudiés ici, ce sont principalement les décisions de préaffectation qui sont touchées par la capacité du système. De cette manière, les performances prédites du système se rapprochent davantage des performances réelles du système. Enfin, la fonction objectif influence clairement les décisions et, en utilisant les mêmes mesures de la qualité du service, les poids attribués aux objectifs liés au service à la population et aux efforts de relocalisation influencent aussi les décisions. Ces constats se vérifient principalement pour les instances de taille réduite, mais aussi pour celle de taille plus réaliste (R149). Néanmoins, le modèle proposé et les outils utilisés dans leur forme actuelle ne permettent pas de résoudre toutes les instances générées dans des délais de temps raisonnables. Dans ce contexte, le développement d'une approche de résolution capable de déterminer de manière efficiente des solutions de bonne qualité au problème étudié est tout à fait justifié. Le développement d'une telle méthode sera alors abordé lors de recherches futures.

5 REMERCIEMENTS

Cette recherche a été subventionnée par le Fonds québécois de la recherche sur la nature et les technologies (FQRNT), par le biais du programme Projet de recherche en équipe, et par le Conseil de recherches en sciences naturelles et en génie du Canada (CRSNG), par le biais du Programme de subventions à la découverte, ainsi que par une Bourse d'études supérieures du Canada Alexander-Graham-Bell (BESC) décernée à la première auteure. Nous tenons ici à remercier ces deux organismes de leur appui.

6 REFERENCES

Andersson, T. Värbrand, P., (2007) Decision support tools for ambulance dispatch and relocation, *Journal of the Operational Research Society*, 58, pp. 195-201.

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, pp. 627-637.

Başar, A. Çatay, B., Ünlüyurt, T., (2012) A taxonomy for emergency service station location problem, *Operations Research Letters*, 6, pp. 1147-1160.

Bélanger, V., Kergosien, Y., Ruiz, A., Soriano, P., (2014) Empirical comparison of relocation strategies in real-time ambulance fleet management, *Document de recherche du CIRRELT*, CIRRELT-2014-73.

Bélanger, V., Ruiz, A., Soriano, P., (2012) Déploiement et redéploiement des véhicules ambulanciers dans la gestion des services préhospitaliers d'urgence, *INFOR*, 50, pp 1-30.

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

Daskin, M. S., (1983) A maximum expected location problem: Formulation, properties and heuristic solution, *Transportation Science*, 17, pp. 416-439.

Gendreau, M., Laporte, G., Semet, F., (1997) Solving an ambulance location model by tabu search, *Location Science*, 5, pp. 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, pp. 1641-1653.

Gendreau, M., Laporte, G., Semet, F., (2006) The maximal expected relocation problem for emergency vehicles, *Journal of the Operational Research Society,* 57, pp. 22-28.

Goldberg, J., (2004) Operations research model for the deployment of emergency services vehicle. *EMS Management Journal*, 1, pp. 20-39.

Hogan, K., ReVelle, C. S., (1986) Concepts and application of backup coverage, *Management Science*, 34, pp.1434-1444.

Jagtenberg, C., Bhulai, S., van der Mei, R. D., (2015) An efficient heuristic for real-time ambulance redeployment, *Operations Research for Health Care,* 4, pp. 27-35.

Kergosien, Y., Bélanger, V., Soriano, P., Ruiz, A., Gendreau, M., (2015) A generic and flexible simulation-based analysis tool for EMS management, À paraître dans *International Journal of Production Research.*

Mason, A. J., (2013) Simulation and real-time optimized relocation for improving ambulance operations. *Handbook of Healthcare Operations : Methods and Applications*, Denton (Ed), Springer : New York, pp. 289-317.

Maxwell, M. S., Restepo, M., Henderson, S. G., Topaloglu, H., (2009) Approximate dynamic programming for ambulance redeployment, *INFORMS Journal on Computing*, 22, 266-281.

Nair, R., Miller-Hooks, E., (2009) Evaluation of relocation strategies for emergency medical service vehicles. *Journal of the Transportation Research Board*, 2137, pp. 63-73.

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

Rajagopalan, H. K., Saydam, C., Xiao, J., (2008) A multiperiod set covering problem for dynamic redeployment of ambulances, *Computers and Operations Research*, 35, pp. 814-826.

Repede, J. F., Bernardo, J. J., (1994) Developing and validating a decision support system for locating emergency medical vehicles in Lousville, Kentucky, *European Journal of Operational Research,* 75, pp. 567-581.

Saydam C., Rajagopalan, H. K., Sharer, E., Lawrimore-Belanger, K., (2013) The dynamic redeployment coverage location model, *Health Systems*, 2, pp. 103-119.

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

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

Toro-Dìaz, H., Mayorga, M. E., McLay, L. A., Rajagopalan, H. K., Saydam, C., (2015) Reducing disparities in large-scale emergency medical service systems, *Journal of the Operational Research Society,* 66, pp. 1169-1181.

Toro-Dìaz, H., Mayorga, M. E. , Chanta, S., McLay, L. A., (2013) Joint location and dispatching decisions for emergency medical services, *Computers & Industrial Engineering*, 64, 917-928.

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