RAIRO-Oper. Res. 53 (2019) 1109–1127
https://doi.org/10.1051/ro/2017084

RAIRO Operations Research
www.rairo-ro.org

THE SUPPORT UNIT LOCATION PROBLEM TO ROAD TRAFFIC SURVEYS WITH MULTI-STAGES

Marcus Vinicius Oliveira Camara[1] and Glaydston Mattos Ribeiro[1,*]

Abstract. Large countries with extensive road networks, such as Brazil, require large volumes of financial resources to perform traffic surveys. In Brazil, the biggest road traffic survey was performed in 2011 with 120 counting survey stations. This survey was divided into three stages and 83 support units provided survey teams. A support unit is a place, such as a military organization, close to the survey stations. A stage indicates that only some survey stations must be considered at a time. In large scale traffic surveys with multi-stages, we must define which support unit will serve each survey station so that travel costs for the survey teams and the costs to use the support units are minimized. We present the Support Unit Location Problem to Assist Road Traffic Survey with Multi-Stages where, given a set of available support units, each one with a coverage area, and a set of multi-stage traffic survey stations, we must select units to serve stations so that the cost is minimized. Scenarios are evaluated for a real traffic survey with 300 counting stations and four stages in Brazil. Computational experiments show that large cost reductions can be found when a mathematical model is used.

Mathematics Subject Classification. 90B06, 90C10, 90C27, 90C90.

Received March 22, 2017. Accepted October 30, 2017.

1. Introduction

Strategic planning of the transport sector has an important role in the economic development of a nation. It is responsible for the correct use of investments in infrastructure (expansion and deployment), cost reduction, and increased competitiveness and efficiency of the sector, among others. However, obtaining data to feed such mechanisms is not always an easy task, mainly due to the quality of the information obtained.

In this context, public and/or private initiatives that aim to provide data and information to governmental agencies which are used to develop a correct planning of the transport sector become relevant. In Brazil, the government uses specific mechanisms to obtain traffic data on national highways since 1975.

According to the National Department of Transport Infrastructure [14], the first traffic counting initiative in Brazil occurred through the *Programa de Contagem Sistemática de Trânsito* (Systematic Traffic Counting Program), which was carried out in the states of Rio de Janeiro, São Paulo and Minas Gerais. In 1976, with the successful implementation of this initiative, the program evolved and was replaced by the *Plano Piloto de Contagem Sistemática de Trânsito* (Pilot Plan for Systematic Traffic Counting). In 1977, the *Plano Nacional*

Keywords. Road traffic survey, traffic counting location, facility location, mathematical modeling.

[1] Federal University of Rio de Janeiro – UFRJ, Alberto Luiz Coimbra Institute for Graduate Studies and Research in Engineering – COPPE, Av. Horácio Macedo, 2030, Cidade Universitária, 21941-450 Rio de Janeiro, RJ, Brazil.

*Corresponding author: gmr.ribeiro@gmail.com

de Contagem de Trânsito (National Traffic Counting Plan) was instituted, which evolved in 1989, 1997 and 1998. However, a transport macro-planning needs information that goes beyond data from counting vehicles on highways.

According to [19], Origin-Destination (OD) matrices for demand estimation derive, usually, from large-scale traffic surveys, despite the high costs involved [23]. Thus, new national initiatives have emerged aiming to acquire information regarding OD, motivation and frequency of trips and characteristics related to the type of cargo transported. As an example, we can mention the event called *Semana Nacional de Contagem de Tráfego* (SNCT – National Traffic Counting Week) held in 2005, the *Operação Safra* (Harvest Operation) in 2006 and the *Pesquisa Nacional de Tráfego* (PNT – National Traffic Survey) in 2011, named PNT 2011 [31].

Traffic surveys become even more complex when it is considered that the quality of an estimated OD matrix for road transport depends on different factors. These factors can be mainly related to: (1) methods used for estimation; (2) correct choice of segments (links) for vehicle counting (location and number of counting points); and (3) the quantity and reliability of the data used [9, 15, 47].

The quantity and reliability of the data used suffer direct variation depending on the workforce involved, especially regarding the specific skills required to conduct traffic surveys. Since the SNCT 2005, the involvement of the *Exército Brasileiro* (EB – Brazilian Army), through its *Organizações Militares* (OM – Military Organizations), has proved to be an important contribution to the accomplishment of these surveys in Brazil. The success of EB in these operations occurs mainly because of its large distribution and operational capacity throughout the national territory, which enable the coverage of survey stations, and their experience in performing field operations.

Thus, when large-scale traffic surveys focused on transportation macro-planning are considered, at the interregional level, it is relevant to plan the assistance of the survey stations by Support Units, which serve the stations with the necessary manpower and infrastructure. With the imminence of a new National Traffic Survey, for the years 2016 and 2017 [32], in which counting and OD surveys will again be conducted by militaries, the Support Unit Location Problem to Assist Road Traffic Survey with Multi-Stages (SULPARTSMS) was developed. Due to the larger number of survey stations (300) and traffic zones (more than 5.500) and the size of the road network (more than 300 thousand kilometers), this new traffic survey was divided into four stages, *i.e.*, two stages in 2016 and two stages in 2017. The first two stages have 120 survey stations and the ones of 2017 have 180, summing up to 300 survey stations spread over the national territory.

The SULPARTSMS can be described as follows: given a set of available support units, each one with a coverage area, and a set of multi-stage traffic survey stations, we must select support units to assist traffic survey stations so that the travel costs of the survey teams and the costs associated to the units are minimized. The coverage of a support unit can be defined, for example, by a maximum distance, *i.e.*, a support unit cannot serve a survey station if the distance between them is larger than a known limit.

Figure 1a shows an example of the SULPARTSMS in which 10 survey stations need support, five in the first stage and five in the second. Five support units are available with different coverage areas represented by non-shaded and shaded areas. Figure 1b shows a solution in which two support units are selected to serve the survey stations in different stages. When a support unit is used more than once, the costs are reduced because several types of equipment have already been purchased, such as vests and signaling devices.

Thus, the main objective of this paper is to present the SULPARTSMS and propose a mathematical modeling to assist road traffic surveys considering that the total cost must be minimized. In this case, this cost is represented by travel costs and the costs to use the support units. Our model was used for the planning of the National Traffic Survey 2016–2017 in Brazil and the main results are presented in this paper, therefore a real case study. Different scenarios are shown to help decision-makers. This study is relevant because it provides a tool to support the traffic survey planning process, aiming to reduce costs and, consequently, to improve the use of resources.

As delimitations of this study, we have used the parameters and characteristics of the surveys already performed in Brazil, mainly of PNT 2011. Furthermore, estimates and simplifications were adopted for the representation of the travel costs and for the costs associated with the use of the support units. However, although

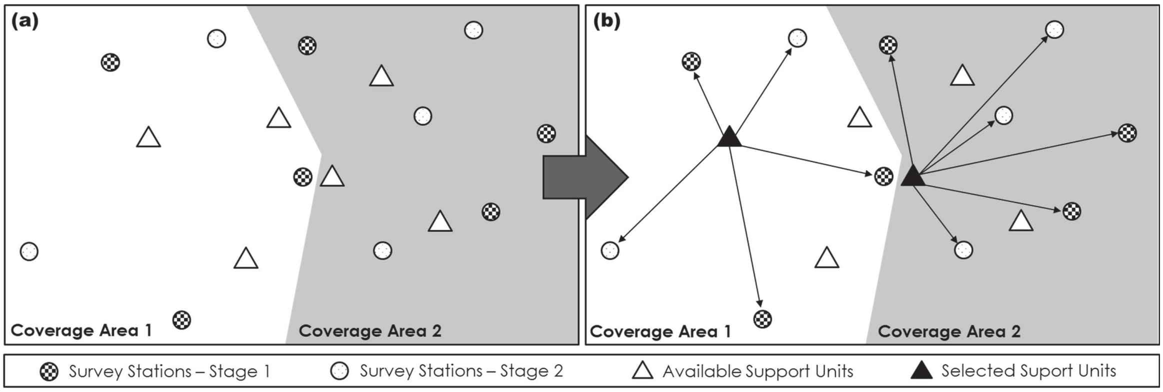

Figure 1. Fictitious problem: (a) general representation and (b) possible solution.

this paper is motivated by a problem in which the support units are OMs, our approach can be easily adapted to consider, for example, schools and institutions (public or private) instead of OMs.

The remainder of this paper is organized into five other sections. Section 2 presents a literature review about location-allocation problems and considerations about road traffic surveys. The problem addressed in this work, the assumptions and the mathematical model are presented in Section 3. Section 4 presents three scenarios, summing up to 20 computational tests for the application of the mathematical model. Section 5 describes the results and Section 6 is dedicated to final remarks.

2. Facility location and road traffic surveys: a brief review

The study of facility location has established itself as an important area of Operational Research [30]. Facility Location Problems (FLP) aim to determine the location of facilities that must interact with other elements, which may have fixed locations, such as producers, usually far from each other [13].

The theory related to facility location is extensive and influenced by the large diversity of factors involved. Based on the studies of [1,5,7,11,16,28,35,48], some important items can be listed in the process of choosing the location of facilities, such as: labor, transport and accessibility, legal issues, and implementation areas, among several others.

Different classes of problems, represented by several model types, can be found in the literature, such as p-Median, p-Coverage and p-Center models. The p-Median location problem was initially proposed in the literature by [21]. This problem consists in locating p (median) facilities so that the sum of the distances between the demand points and the nearest facility is minimized [34].

The first relevant references found in the literature about coverage location problems are from the studies of [4, 22]. However, the first mathematical formulation was proposed by [41]. For this problem, the concept of coverage must be introduced, which in this case is related to the demand points that each facility can cover. A demand point is covered by a facility if the facility can serve it. In an analogous way, the set of demand points that this facility can serve is called the coverage of this facility. Thus, for each demand point there is a subset of facilities that can serve it. The coverage area can be defined based on the distance or travel time, as described in Section 1.

The p-Center location problem consists in locating p facilities, which must be allocated to the demand points, in order to minimize the maximum distance between each demand point and the facility that serves it [39]. According to [40], the objective of the p-Center problem can also be considered to be the minimization of maximum loss in order to provide a good service by a facility.

Although the SULPARTSMS is an FLP, it does not have characteristics of just one class of problem. Even so, its main features are related to the p-Median and p-Coverage problems. The SULPARTSMS is similar to the p-Median problems in that its goal is finding facilities that serve demand points at the lowest possible cost. Besides, it considers capacities such as the capacitated p-Median problem [29, 44].

However, the SULPARTSMS also incorporates the main characteristics of the p-Coverage problems. In this case, its constraints ensure that the facilities have a set of candidate demand points to be served. Consequently, these points belong to their coverage area. Moreover, the SULPARTSMS considers a cost related to the opening of a facility (represented by use costs) in the objective function.

Despite considering some characteristics of already known problems related to FLP theory, our SULPARTSMS presents contributions not only regarding the mathematical modeling addressed, but also in the area of transportation planning and traffic engineering.

As for the transportation planning, it is hard to obtain representative mass data that can support the application of effective planning techniques and models [37, 43]. This difficulty is related to dimensions (large territorial areas, growing population and extensive road transportation network) but also, and mainly, related to the high costs involved [33] that are often prohibitive. For example, in transportation planning with a focus on prioritization of investments for the expansion of the road transportation network (new highways), improvements in pavement conditions or concession feasibility studies, it is necessary to know the annual average daily traffic on highways [18, 26, 42, 50].

In this context, several papers were presented in the literature addressing how to obtain these data, optimizing the use of resources. For problems that are similar to the SULPARTSMS, we found applications focused on traffic surveys (data collection on highways) through automated sensors/counters with the location of these counters focused on OD matrix estimates. Studies related to this topic are considered important because OD demand estimation from traffic location (traffic counters) is an important input data for transportation planning and traffic management [36].

Yang *et al.* [45] presented a mathematical model for locating a minimum amount of traffic counter stations with constraints that should guarantee coverage of all possible OD pairs (at least a fraction of travel between these OD pairs) for a given transportation network. This study was complemented by [46], with constraints that ensure coverage of all possible routes among the OD pairs considered in the network.

Considering that the costs to implement a minimum number of automatic traffic counters in large networks are high [24], new approaches have emerged aiming to maximize results with limited budget resources. For example, the maximization of captured OD pairs with a limited number of traffic counters [47, 49].

Yang and Zhou [47] have formally proposed the Traffic Counting Location Problem (TCL), which has been extensively studied (see [10, 17, 27, 38]). The TCL is also known as the Sensor Location Problem (SLP), as seen in [3, 6, 20, 51]. These papers seek to propose traffic counter locations aiming to obtain reliable data for the OD matrix estimation. The relationship between the number of traffic counters and the quality of the estimated OD matrix is detailed in [9].

As shown in the literature, most studies found in the literature are devoted to locating traffic counters, mainly focused on automatic vehicle counters. However, in the case of traffic surveys involving vehicle counting and OD interviews, it is necessary to find support units to provide logistics structure and researcher teams for the stations. Our paper presents a mathematical model to complement this gap in the literature.

3. Problem definition, assumptions and mathematical modeling

The SULPARTSMS aims to determine, from a set of possible locations, those that must be used as Support Units to assist road traffic surveys with survey teams. However, these surveys can also be performed at different time periods, here called stages, so, in addition to the locations, the mathematical model must also indicate in which stages each support unit is used.

Regarding survey stations, in the case of multi-stage traffic surveys, the mathematical modeling must also allow a set of survey stations already defined for each stage to be previously indicated. When the sets of stations, for each stage, are not defined *a priori*, the model must indicate at which stage each survey station must be attended, based on the costs to be minimized.

The main characteristics of support units are related to capacity and coverage. With regard to capacity, each unit has a maximum number of survey stations that can perform service, or simply, that can be allocated. Coverage is related to the covered area of each support unit, which may limit the set of possible stations to be served. In this sense, the coverage of a support unit refers to the set of stations that it can serve.

3.1. Assumptions for the SULPARTSMS

Taking into account the observations described above about the SULPARTSMS, the following assumptions are considered for the mathematical model development:

i. The support units selected to attend road traffic survey stations must be chosen from a set of pre-defined candidate points;
ii. The choice of a given support unit must be associated with the travel cost of the survey teams to the survey station which they are assigned;
iii. We must also consider the cost of using a Support Unit. In the case of multi-stage survey, the reuse of units at different stages may lead to cost reduction, *i.e.*, there is no need for new expenditures with the training of researchers and the purchase of accessory materials such as signaling devices, tables and chairs. This fact makes it interesting to select units to attend survey stations in different stages;
iv. The number of survey stations may indicate the need to divide the survey into multiple stages. In this case, the set of stations that must be attended at each stage may or not be previously defined. However, the number of stations that compose each stage must be pre-defined;
v. It may be interesting that the road traffic survey planning considers, for different reasons, budget constraints, a maximum number of support units at all stages or even that a maximum number of support units is employed per stage. This assumption must be fundamentally presented in the mathematical model;
vi. In addition to the capacity of each support unit, we must consider its coverage area. This area can be defined by a maximum coverage radius or even as a set of possible stations that can be served by various others reasons; and
vii. As a general rule, all survey stations must be served by a single support unit. However, in the case of impossibility of attendance due to lack of operational capacity or coverage limitations (lack of support units to attend a given region), the mathematical model must indicate the stations that have not been served.

3.2. Mathematical modeling

Let S be the set that represents all available support units for use, P be the set of all road traffic survey stations that must be served, N be the set whose elements represent the stages in which the traffic survey will be performed and Q be a set whose elements indicate the number of stages in which each of the support units can be used. For example, if a survey has four stages, then $N = \{1,2,3,4\}$ and each support unit can: not be used, be used once, twice, three times or even four times (at all stages). In this case, $Q = \{0,1,2,3,4\}$.

Based on the information above, consider that:

- U_p: represents the set of support units that are candidates to serve the survey station $p \in P$, such that, $U_p \subseteq S$. A candidate can serve a survey station if this station belongs to its coverage area;
- A_n: represents the set of survey stations that must be attended at each stage $n \in N$ of the traffic survey, such that, $A_n \subseteq P$. If all survey stations can be served in any stage, then $A_n = P \quad \forall n \in N$;
- dummy: represents a fictitious support unit of infinite capacity, used as a way to meet the requirements of Assumption (vii). Therefore, the dummy support unit must only be used in cases of lack of capacity due to labor or coverage issues. However, a penalty must be associated with the use of this dummy support unit;
- $c_{sp} \geq 0$: represents the travel cost of a survey team from a given support unit $s \in S$ to a survey station $p \in P$;

- $\gamma \geq 0$: represents the penalty (cost) for the allocation of a dummy support unit to a survey station. This penalty receives a high value, larger than a common allocation, because, as mentioned before, the allocation of a fictitious location must only be performed in cases of lack of operational capacity or coverage;
- $\theta \geq 0$: represents the penalty (cost) of using, or opening, a dummy support unit;
- $\delta_{qs} \geq 0$: represents the cost of the support unit. This cost is associated with the number of stages $q \in Q$ in which a support unit $s \in S$ is used during the whole traffic survey;
- $\mathrm{MAX}UA_n \geq 0$: represents the maximum number of support units that can be used per stage $n \in N$;
- $\mathrm{MAX} \geq 0$: represents the maximum number of different support units that can be used during all survey stages;
- $\mathrm{QTDP}_n \geq 0$: represents the amount of survey stations that must be attended in each stage $n \in N$;
- $\kappa_{ns} \geq 0$: represents the amount of survey teams that, for each survey stage $n \in N$, will be available per support unit $s \in S$ to attend the survey stations;
- $x_{nsp} \in \{0,1\}$: is a binary decision variable. If a support unit $s \in S$ is allocated to a survey station $p \in P$ in the survey stage $n \in N$, $x_{nsp} = 1$, otherwise $x_{nsp} = 0$;
- $y_{ns} \in \{0,1\}$: is a binary decision variable. If a support unit $s \in S$ is used in the survey stage $n \in N$, $y_{ns} = 1$, otherwise $y_{ns} = 0$; and
- $l_{qs} \in \{0,1\}$: is a binary decision variable. If a support unit $s \in S$ is chosen for use in $q \in Q$ survey stages, $l_{qs} = 1$, otherwise $l_{qs} = 0$.

According to the description of the problem and, at the same time, considering the established assumptions, the objective function of the SULPARTSMS is presented in (3.1).

$$\text{Minimize } z = \sum_{n\in N}\sum_{p\in P}\sum_{s\in S} c_{sp}x_{nsp} + \sum_{q\in Q}\sum_{s\in S}\delta_{qs}l_{qs} + \sum_{n\in N}\sum_{p\in P}\gamma x_{n,\mathrm{dummy},p} + \sum_{n\in N}\theta y_{n,\mathrm{dummy}} \tag{3.1}$$

The Objective Function (3.1), which must be minimized, is composed of four terms as follows:

- the first term refers to the travel costs of the survey teams from the selected support units to the allocated survey stations;
- the second term consists of the costs associated with the use of the support units at all stages, *i.e.*, it is based on the number of survey stages that each support unit is used; and
- the last two terms refer to the penalties of using the dummy support unit to serve survey stations at all stages.

Constraints (3.2) ensure that all survey stations are attended only once. For such attendance, a support unit, conventional (one whose location belongs to set S) or dummy, must be allocated to the survey station.

$$\sum_{n\in N}\sum_{s\in U_p} x_{nsp} + \sum_{n\in N} x_{n,\mathrm{dummy},p} = 1 \;\forall p \in P \tag{3.2}$$

Constraints (3.3) guarantee that, in each survey stage $n \in N$, only the survey stations planned for the referred stage will be attended. Additionally, the number of survey stations served at each stage must respect the parameter $QTDP_n$. As in constraints (3.2), the variables related to the dummy support unit are considered.

$$\sum_{p\in A_n}\sum_{s\in U_p} x_{nsp} + \sum_{p\in A_n} x_{n,\mathrm{dummy},p} = QTDP_n \;\forall n \in N \tag{3.3}$$

Constraints (3.4) ensure that, for each survey stage $n \in N$, the maximum amount of support units ($\mathrm{MAX}UA_n$) must be respected. These constraints can be used as budget limits for one or more stages.

$$\sum_{s\in S} y_{ns} \leq \mathrm{MAX}UA_n \;\forall n \in N \tag{3.4}$$

Constraint (3.5) guarantees that the maximum number of different support units must be respected during the whole road traffic survey. This limit can be defined by various reasons, such as the budget constraints.

$$\sum_{s\in S}\sum_{q\in Q} l_{qs} \leq \mathrm{MAX} \tag{3.5}$$

Constraints (3.6) guarantee that, for each survey stage $n \in N$ and for each support unit $s \in \{U_p, \text{dummy}\}$, a survey station $p \in P$ can only be served by a support unit s, if, and only if, the support unit s is selected for the stage n. This way, the allocation of survey stations to unselected support units at each stage is prohibited.

$$x_{nsp} \leq y_{ns} \ \forall n \in N, s \in \{U_p, \text{dummy}\}, p \in P \tag{3.6}$$

Constraints (3.7) are related to the operational capacity of each support unit. In each survey stage $n \in N$, the number of survey stations served by a support unit must respect the available number of survey teams (support unit capacity).

$$\sum_{p\in P\backslash s\in U_p} x_{nsp} \leq \kappa_{ns} y_{ns} \ \forall n \in N, s \in S \tag{3.7}$$

Constraints (3.8) and (3.9) guarantee, together, that the number of survey stages in which a support unit $s \in S$ is used must be equal to the quantity defined by ql_{qs}, $q \in Q$.

$$\sum_{n\in N} y_{ns} - \sum_{q\in Q} ql_{qs} = 0 \ \forall \in S \tag{3.8}$$

$$\sum_{q\in Q} l_{qs} \leq 1 \ \forall s \in S \tag{3.9}$$

Finally, constraints (3.10), (3.11) and (3.12) are related to decision variable domains.

$$x_{nsp} \in \{0,1\} \ \forall n \in N, s \in \{S, \text{dummy}\}, p \in P \tag{3.10}$$

$$y_{ns} \in \{0,1\} \ \forall n \in N, s \in \{S, \text{dummy}\} \tag{3.11}$$

$$l_{qs} \in \{0,1\} \ \forall q \in Q, s \in S \tag{3.12}$$

4. Scenarios for evaluation

In this section, three scenarios for the application of the mathematical model (3.1)–(3.12) are evaluated. The first scenario considers an application using real data from the PNT 2011. The second one considers an application based on information from a new traffic survey scheduled for 2016 and 2017 in Brazil. The third one considers an application for a hypothetical larger survey. The Military Organizations (OM) of the Brazilian Army are used as support units.

The different scenarios are presented to allow an evaluation of the results of the mathematical model for different situations, such as surveys with different numbers of survey stations and stages, different cost parameters or different amounts of survey stations per each stage. However, some common parameters are adopted for all scenarios.

We have used real data (locations and capacities) about the OMs. The capacity is based on the quantity of survey teams that each OM can provide. Thus, 221 OMs are available for traffic surveys and the capacities vary from 1 to 3 survey teams. Besides, we consider that all OMs are available for each survey stage. Thus, $\mathrm{MAX}UA_n = \mathrm{MAX} = |S|, \forall n \in N$, where $|\bullet|$ represents the total number of elements of the argument.

The estimation of the travel cost of the survey teams (military platoons) considered two trucks, as in the PNT 2011. The trucks are used to transport the militaries and equipment to the survey stations. Thus, let $dist_{sp}$ be the distance in km between a support unit $s \in S$ and a survey station $p \in P$. Considering data from

[2, 12] about the average price of fuel (diesel oil) in Brazil and the average consumption of a truck, the travel cost c_{sp} is given by (4.1):

$$c_{sp} = 2.37\text{dist}_{sp} \qquad \forall s \in S, p \in P. \tag{4.1}$$

Therefore, in order to calculate the travel cost, it is necessary to know the distances between OMs and survey stations. Thus, taking into account the georeferenced data of OM locations and survey stations, software TransCAD 5.0 [8] was used to obtain a matrix of the distance.

The cost of each OM is a value dependent on many factors of difficult estimation, such as researchers training, team management, equipment, material used, and the number of times each OM is used in different survey stages.

For this reason, for each scenario, five variations of these costs were created (five cost groups). The costs of the groups were determined according to a range of values, resulting in a more complete analysis of the impact of this parameter. Thus, the five cost groups adopted for the computational experiments are presented in Table 1, which also considers the number of stages.

Table 1. Cost groups for the use of OM in different stages.

Cost Group (US$)	Number of Stages				
	1	2	3	4	5
Cost Group 1	10	18	24	28	30
Cost Group 2	50	90	120	140	150
Cost Group 3	250	450	600	700	750
Cost Group 4	1250	2250	3000	3500	3750
Cost Group 5	6250	11250	15000	17500	18750

Cost Group 1 has the lowest values while Cost Group 5 has the highest values. The intermediate groups were obtained considering multipliers of five. For example, Cost Group 2 has values five times higher than those in Cost Group 1, Cost Group 3 has values five times higher than those in Cost Group 2, and so on until Cost Group 5. Thus, we expect to create adequate conditions to evaluate a possible trade-off between the amount of OMs used during the entire traffic survey and the total distances traveled by military troops.

The expected trade-off in this case refers to the possibility that the mathematical model indicates the use of a large number of OMs when the costs are low (expected result for Cost Group 1) and a small number of OMs when the costs are high (expected result for Cost Group 5). However, when the number of OMs is reduced, the travelled distances increase generating impacts on the travel costs.

The cost groups vary according to the number of survey stages in which each OM participates during the traffic survey and, although these costs have been constructed from estimated values, its composition complies with the principle that: if an OM is used in different stages in the same traffic survey, its cost reduces per stage.

As a penalty for allocating and using a dummy OM, large enough values were adopted to guarantee that they could only be used in cases of lack of survey teams (military platoons) to attend the survey stations or even insufficient numbers of OM to perform the traffic survey. Therefore, the value established as a penalty for the allocation of a dummy OM was $\gamma = 10^6$, and the penalty for the use of a dummy OM was $\theta = 10^6$.

It is important to highlight that the mathematical model (3.1)–(3.12) can be considered a decision support tool in the planning of traffic survey, however, for its effective application in a real situation, it is important to develop specific studies that accurately estimate the costs involved.

4.1. Scenario 1

Scenario 1 considers the application of the mathematical model for the SULPARTSMS with the parameters and characteristics of the PNT 2011 in Brazil [31]. This scenario is important to allow the comparison of the total costs related to the travel cost and the OM cost using the configuration (allocation) performed in 2011 and the total costs obtained after applying the mathematical model. With the adoption of this procedure, it is possible to evaluate the potential of reduction in the total cost if the proposed mathematical model was used during the planning of the PNT 2011.

The road traffic survey of the PNT 2011 was conducted at 120 different survey stations divided into three stages. In the first and third stages, the surveys were performed in 22 survey stations, while in the second stage all 120 stations were used.

A total number of 83 different OMs were used in all stages. In the first and third stages, 19 OMs were used and in the second survey stage all 83 OMs were used. According to data from the PNT 2011, it was possible to identify and georeference the OMs effectively used in the traffic survey. However, the relationship of all OMs available in the planning process and the choice of OMs actually adopted at that time is unknown. Therefore, besides the 221 OMs available for the traffic surveys of 2016 and 2017, we included 10 OMs that do not belong to this set but were used in 2011.

Regarding the OMs coverage areas, *i.e.*, the list of survey stations that each OM can serve, they are related with the operation areas of each OM, called *Comando Militar de Área* (CMA – Military Command Area). Each OM can only serve survey stations that belong to the coverage area of its CMA. Following this premise, this study considered information from the CMA to define the sets $U_p, \forall p \in P$.

This first scenario is composed of 10 instances (Instances 1–10), with the first five ones considering the solution implemented in the PNT 2011 but with the group costs presented in Table 1. Instances 6–10 consider the group costs presented in Table 1 and the main parameters of the PNT 2011. These instances are solved by a commercial solver as we show in Section 5.

4.2. Scenario 2

Information about the new road traffic surveys of 2016 and 2017 in Brazil indicates that four survey stages are used with 300 different survey stations throughout the national territory. With that in mind, we considered 10 instances (Instances 11–20), which also vary according to cost groups (Tab. 1), divided into two configurations described as follows:

- Configuration 1: aims to represent the new road traffic survey with the information about the real division of survey stages (amount of survey stations per stage). Stages 1, 3 and 4 with 60 survey stations and Stage 2 with 120 survey stations; and
- Configuration 2: proposes a new division of the amount of survey stations per stage. In this configuration, all survey stages have an equal amount of stations, *i.e.*, 75 survey stations per stage. This configuration aims to evaluate the potential benefits of an equal division of the survey stations through the stages.

4.3. Scenario 3

Scenario 3 has the main objective of evaluating the behavior of the mathematical model given a hypothetical road traffic survey, larger than the one of 2016 and 2017.

Thus, we expect that the mathematical model will also assist larger road traffic surveys, with a level of availability of resources similar to the current one. Thus, Scenario 3 proposes a road traffic with 500 survey stations, 500 support units to serve them, and five stages. However, in this scenario, the locations of the points (survey stations and OMs) were randomly defined using TransCAD 5.0 [8] taking into account the georeferenced database of Brazil.

As in other scenarios, the demand for a survey station is equivalent to one survey team. Regarding the availability of these teams by OMs, the maximum number of teams that each unit can provide for the traffic survey was also defined at random. Following the capacities of each OM for the real situations presented in the other scenarios, the capacities of each support unit were obtained considering the range of 1–5.

Considering as a premise the information from a previous traffic survey (Scenario 1) and from the new traffic survey (Scenario 2), we believe that the number of survey stations per stage is limited by the availability of survey teams. We defined for this scenario that the 500 survey stations are equally divided into five stages. Thus, this scenario is composed by five instances (Instances 21–25), one for each group cost presented in Table 1.

5. Computational experiments and results

In order to solve the mathematical model presented in Section 3 for the three scenarios defined in Section 4, a computer equipped with an AMD Phenom X4 1.9 GHz processor and 4 GB of RAM was used. The IBM ILOG CPLEX 12.6 [25] optimization software was used for up to 2 h on each instance. The summary of the results, based on the best solutions found by CPLEX for the 25 instances of the three proposed scenarios, is presented in Table 2.

In this table, first column shows the scenarios and the instances. The second column presents the number of traffic survey stages served. The third column indicates the number of OMs used. The fourth column shows the costs associated to the travelling of the military platoons. The fifth column shows the costs associated to the use of OMs. Finally, the last column presents the results found by CPLEX, where GAP% is the deviation between lower and upper bounds after two hours.

5.1. Results of Scenario 1

Scenario 1 allows a comparison between the decisions of the PNT planning team of 2011 regarding the allocation of military platoons performed, and such allocation made from the results of the mathematical model proposed in this paper in order to serve the 164 traffic survey stations at the time.

Although 120 different survey stations were used in 2011, 22 were repeated in Stages 1 and 3, with a total of 164 survey stations being used as input for the mathematical model. Thus, we use indicators which are extracted from the results to perform comparisons.

The indicators mentioned above refer, for example, to the amount of OMs used (per stage and throughout the traffic survey), the travelling cost of military platoons going to survey stations, the cost of using OMs, and the total value of the Objective Function (best solution value found by CPLEX), among others. In addition to the analysis of these indicators, other comparisons can be made in which importance is given to the time and computational effort required during the solution process.

With respect to the amount of OMs (regardless of the analysis being performed per survey stage or number of different OMs), the results of the mathematical model are better in all cases. This means that in Instances 6–10, a smaller number of OMs are needed to serve all 164 survey stations of the PNT 2011, as can be seen in Table 2.

In the PNT 2011 (Instances 1–5), in the first and third stages of the traffic survey, 19 OMs were used to serve the 22 survey stations, and in the second stage 83 OMs were used. It is possible to observe that the best results were found for the amount of different OMs in Instances 6–10 for Stage 2.

The results of Table 2 show that with each increase in the costs of the OMs or variation of the cost groups, a smaller number of OMs is chosen to serve the survey stations. For this reason, for Instances 6 and 7, the gains are much lower in relation to the other instances. In both Instance 6 and 7, the amount of OMs used in Stages 1 and 3 is equal to 18. For Stage 2, 82 OMs were used in Instance 6 and 78 OMs in Instance 7. For Instances 8, 9 and 10, the reduction in the amount of OMs is higher as usage costs increase. Instances 8, 9 and 10 used 16, 11 and 9 OMs, respectively, in Stages 1 and 3 while for Stage 2, 64, 44 and 41 OMs were used, respectively.

The situation presented for Scenario 1 has a very particular characteristic where the 22 stations of Stage 1 and Stage 3 coincide in location with 22 of the 120 stations of Stage 2. Thus, the results of the mathematical model show that the amount of different OMs used during the entire survey coincides with the amount of OMs used in Stage 2.

This situation occurs because the OMs selected for Stage 2 also need to serve the 22 survey stations for Stages 1 and 3, since their locations coincide. In an analogous way, it is observed that the amount of OMs in Stages 1 and 3 are the same in all cases (all instances).

TABLE 2. Summary of the results.

Scenarios / Instances			Number of survey stations served					Number of OMs used						Travel cost						OM use cost	CPLEX		
			Stage					Stage						Stage									
			1	2	3	4	5	1	2	3	4	5	$\neq^{(1)}$	1	2	3	4	5	Total		Best solution value	Time (s)	GAP%
Scenario 1	2011	1	22	120	22	-	-	19	83	19	-	-	83	6,843.6	34,152.9	6,843.6	-	-	47,840.2	1,096.0	48,936.2	0.53	0
		2	22	120	22	-	-	19	83	19	-	-	83	6,843.6	34,152.9	6,843.6	-	-	47,840.2	5,480.0	53,320.2	0.53	0
		3	22	120	22	-	-	19	83	19	-	-	83	6,843.6	34,152.9	6,843.6	-	-	47,840.2	27,400.0	75,240.2	0.53	0
		4	22	120	22	-	-	19	83	19	-	-	83	6,843.6	34,152.9	6,843.6	-	-	47,840.2	137,000.0	184,840.2	0.53	0
		5	22	120	22	-	-	19	83	19	-	-	83	6,843.6	34,152.9	6,843.6	-	-	47,840.2	685,000.0	732,840.2	0.55	0
	Mathematical Model	6	22	120	22	-	-	18	82	18	-	-	82	4,595.5	24,381.0	4,595.5	-	-	33,572.0	1,072.0	34,644.0	1.57	0
		7	22	120	22	-	-	18	78	18	-	-	78	4,595.5	24,466.2	4,595.5	-	-	33,657.3	5,160.0	38,817.3	1.59	0
		8	22	120	22	-	-	16	64	16	-	-	64	4,694.5	26,317.3	4,694.5	-	-	35,706.4	21,600.0	57,306.4	2.11	0
		9	22	120	22	-	-	11	44	11	-	-	44	7,136.4	36,545.0	7,136.4	-	-	50,817.8	74,750.0	125,567.8	6.75	0
		10	22	120	22	-	-	9	41	9	-	-	41	12,076.9	42,983.7	12,076.9	-	-	67,137.5	335,000.0	402,137.8	462.46	0
Scenario 2	Configuration 1	11	60	120	60	60	-	39	80	38	40	-	116	15,750.4	25,913.9	21,230.6	17,342.5	-	80,237.5	1,728.0	81,965.5	15.6	0
		12	60	120	60	60	-	38	71	37	39	-	108	15,787.5	29,991.8	18,408.8	16,388.0	-	80,576.2	8,050.0	88,624.2	16.57	0
		13	60	120	60	60	-	30	62	34	34	-	86	16,640.7	31,385.4	18,555.7	17,197.8	-	83,779.8	34,250.0	118,029.8	53.15	0
		14	60	120	60	60	-	24	46	23	23	-	60	20,617.7	41,858.0	23,600.4	21,215.9	-	107,292.1	122,750.0	230,042.1	1,265.00	0
		15	60	120	60	60	-	21	42	21	21	-	50	30,351.6	48,939.4	25,724.2	29,352.2	-	134,367.6	538,750.0	673,117.6	7,200.00	0.29
	Configuration 2	16	75	75	75	75	-	52	48	50	48	-	106	14,709.6	20,179.3	22,150.3	18,476.6	-	75,515.9	1,666.0	77,181.9	13.21	0
		17	75	75	75	75	-	47	49	43	46	-	96	19,053.2	17,756.6	20,454.6	18,588.7	-	75,853.3	7,740.0	83,593.3	41.31	0
		18	75	75	75	75	-	39	41	40	39	-	75	19,449.4	19,060.3	20,452.3	20,480.1	-	79,442.3	32,250.0	111,692.3	1,463.27	0
		19	75	75	75	75	-	28	28	28	29	-	40	28,572.0	27,689.5	27,025.8	24,149.6	-	107,437.0	105,750.0	213,187.0	7,200.00	1.29
		20	75	75	75	75	-	26	27	26	26	-	30	34,832.8	29,575.8	26,807.5	33,021.5	-	124,237.8	472,500.0	596,737.8	7,200.00	1.29
Scenario 3	Mathematical Model	21	100	100	100	100	100	63	66	62	67	63	310	11,317.2	13,984.7	11,658.0	12,191.1	11,664.1	60,815.4	3,184.0	63,999.4	1,903.32	0
		22	100	100	100	100	100	62	53	57	63	55	281	11,522.9	13,794.3	11,103.0	12,901.5	12,439.1	61,760.9	14,400.0	76,160.9	1,803.98	0
		23	100	100	100	100	100	43	46	43	38	48	206	14,132.5	13,777.4	14,101.9	14,567.3	14,458.8	71,038.1	53,000.0	124,038.1	7,200.00	5.59
		24	100	100	100	100	100	27	27	25	27	27	128	19,463.6	21,307.9	22,477.6	20,213.9	20,847.3	104,310.5	163,500.0	267,810.5	7,200.00	21.57
		25	100	100	100	100	100	23	22	25	23	22	58	43,570.1	45,865.9	44,690.0	41,500.4	37,816.6	213,443.1	580,000.0	793,434.1	7,200.00	31.83

Note: (1) Number of different OMs used.

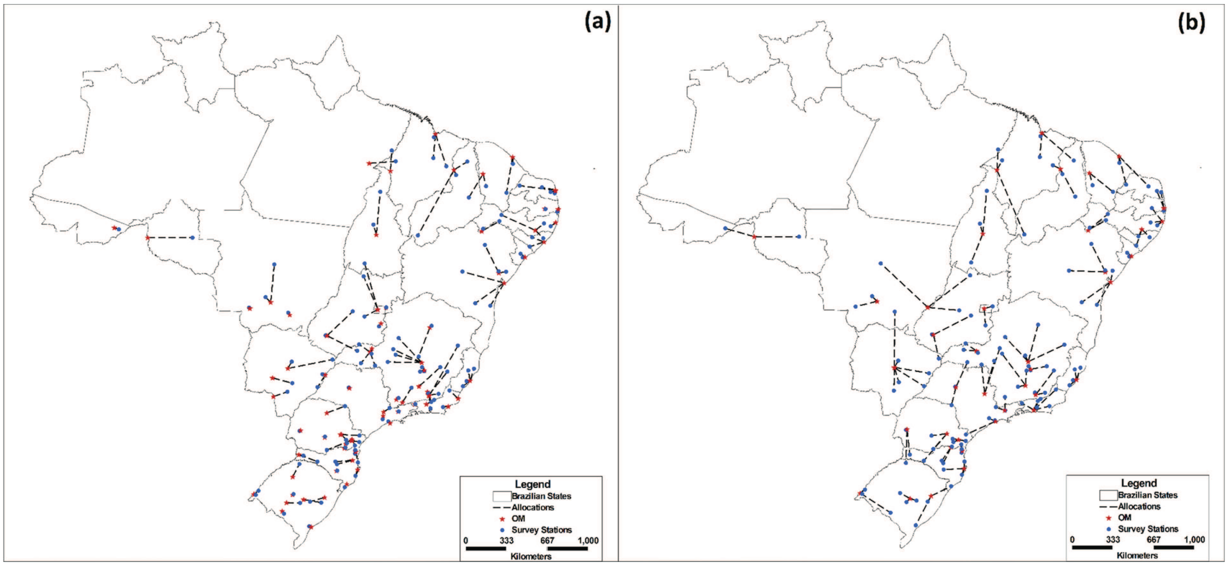

Figure 2. Comparative map between results of (a) Instance 1 and (b) Instance 10.

Thus, for the quantities of OMs used to serve the survey stations considered in Scenario 1, the best results are obtained with Group 5. In this group (Instance 10), in Stages 1 and 3, only 56.25% of the used OMs in the PNT 2011 would be enough to serve the survey stations. For Stage 2, the result is even better and only 49.40% of OMs would be sufficient.

In general, with respect to the total amount of different OMs throughout the traffic survey, as presented previously, the benefit is coincident with the results of Stage 2. Therefore, following the configurations and parameters presented for this scenario and ignoring external factors that may have influenced the choice of OMs in 2011, the survey stations could have been served with 41 OMs instead of 83.

We must also point out that the results of the mathematical model show that this amount is directly related to the variations in the cost groups established for the use of OMs. Therefore, besides the amounts of OMs, we must also analyze other indicators obtained from the variation of these costs. Figures 2a and 2b show, respectively, the solutions for Instances 1 and 10. We can see different allocations for OMs and survey stations.

When the results of Instances 6–10 are compared with the costs of the PNT 2011, we can observe that the costs related to the travelling of military troops are also smaller in three of the five comparisons. In the first and third survey stages, the reductions in travel costs were: 32.85% (Instance 6); 32.85% (Instance 7); and 31.40% (Instance 8). As for Stage 2, reductions are even more significant. At this stage, the number of survey stations is much higher compared to the others. In this case, the potential gains are: 28.61% (Instance 6); 28.36% (Instance 7); and 22.94% (Instance 8).

However, in the case of Instances 9 and 10, the high benefit in relation to the number of OMs has a negative effect on the travelling costs. In Stages 1 and 3, we have a growth of 4.28% (Instance 9) and 76.47% (Instance 10), and in Stage 2, 7.00% (Instance 9) and 25.86% (Instance 10), justified by the significant reduction in OMs used during the traffic survey.

For the total travel costs, the result follows the previously presented trends seen per stage. Again, Instances 6, 7 and 8 present better results, allowing a cost reduction of 29.82%, 29.65% and 25.36%, respectively, while Instances 9 and 10 increase these costs by 6.22% and 40.34%, respectively.

Regarding OM costs, for each Cost Group (1–5) we observe that when comparing the results from the mathematical model (Instances 6–10) and the instances representing the allocations of the 2011 PNT, respectively per Cost Group, in all cases the results of the mathematical model are better than the PNT instance results.

Therefore, in situations where the costs of the OMs are high, the mathematical model can provide good results. These results also reflect the reductions in the amount of OMs used. We can state, then, that the potential reductions in the OM use costs are: 2.19% with Cost Group 1; 5.84% with Cost Group 2; 21.17% with Cost Group 3; 45.44% with Cost Group 4; and 51.09% with Cost Group 5.

For the results of the best solution values (last column of Tab. 2), with the same comparisons made previously applied to the cost groups, again good cost reductions are possible where OMs present high use costs. However, the reductions are not increasing. In the first three cases, Instances 6, 7 and 8 compared to Instances 1, 2 and 3, the potential gains are, respectively, 29.21%, 27.20% and 23.84%. There is a decreasing behavior, while in the comparison between Instances 9 and 10 with Instances 4 and 5, the potential reductions are, respectively, 32.07% and 45.13% in the best solution found.

A possible reason for this behavior is again the variation in OM use costs. As the travelling costs for Instances 1–5 are the same, the impacts in the best solution found occur precisely when a smaller OM amount is used.

Now considering the computational complexity associated with the mathematical model, Table 2 shows that the time to find the optimal solution for Instances 1–5 was practically constant and small, as the decision variables were inserted in the model with the respective choices made in 2011. For the other instances (6–10), the computational time starts a more accentuated growth in Instance 9, with 6.75 s, and reaches 462.46 s in Instance 10.

Although the computational times are below 10 minutes, this behavior gives indications that in larger instances, as in the ones presented in Scenarios 2 and 3, the computational complexity may be related mainly to variations in cost groups which can make it difficult to obtain the optimal solution using a conventional optimization software.

5.2. Results of Scenario 2

In Scenario 2, when considering the data of the new larger traffic survey, we expect that the mathematical model can be used as a decision support tool in the choice of OMs to serve road traffic survey stations in the 2016 and 2017.

As in the evaluation of Scenario 1, the indicators used to evaluate the results of the instances are related to the number of OMs used to serve the survey stations, the travelling costs of military platoons (per stage and for the entire traffic survey), the OM use costs, the best solution value by CPLEX for the proposed mathematical model and the computational time of the solution process. In addition, special attention is given to the residual GAP% in cases where the optimal solution is not found during the defined maximum processing time.

Considering the number of OMs for Configuration 1, Table 2 shows that Stage 2 is the one that requires more OMs due to the larger number of survey stations that must be served (120). In addition, the difference in results found for the instances is related to the OM costs. In this way, we observe that, when the OM costs increase, the optimal solution indicates that a smaller number of OMs must be used.

In the worst case (Instance 11), Stages 1, 2, 3 and 4 required 39, 80, 38 and 40 OMs, respectively, with a total of 116 different OMs. However, the results of the other instances show that, with the increase in the OM costs, this amount can be reduced in about 56.90% (Instance 15). This is because Instance 15 requires 21, 42, 21 and 21 OMs in Stages 1, 2, 3 and 4, respectively. In this case, the total number of different OMs is 50.

In the case of the instances of Configuration 2, Table 2 shows that, with the equal division of survey stations between the stages, an even smaller number of different OMs is sufficient to serve the survey stations. Among the survey stages, Instances 19 and 20 stand out because, even with the increase in the OM costs, the number of OMs used is very similar in these instances. However, the difference is better observed in the amount of different OMs: 40 different OMs are required for Instance 19 and only 30 for Instance 20.

Thus, it is important to note that, in terms of the number of OMs used during the entire traffic survey, the proposed division of survey stations for the instances of Configuration 2 presents better results than the division of Configuration 1. On the other hand, it is important to highlight that the technical criteria used to perform the division of survey stations of Configuration 1 are not known.

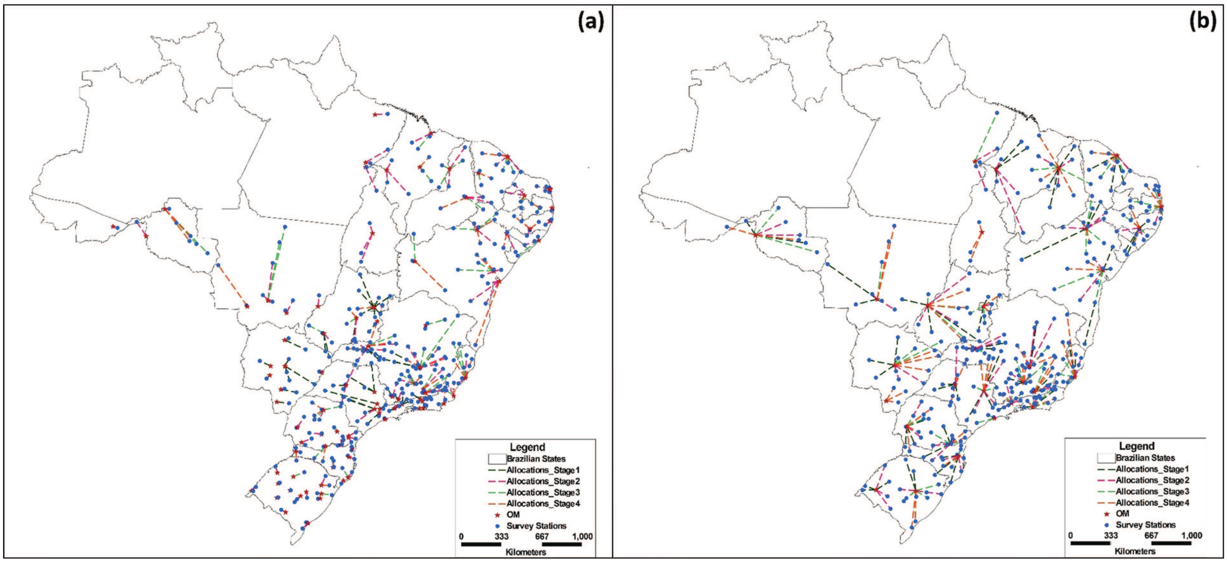

FIGURE 3. Comparative map between results of (a) Instance 11 and (b) Instance 20.

Figures 3a and 3b show, respectively, the results for Instances 11 and 20. We can see in Figure 3b that a reduced number of OMs is used, so the travelled distances are larger than in Figure 3a.

The travelling costs analysis indicates that significant reductions in the amount of OMs needed for the traffic survey, as a result of the increase in OM costs, influence negatively on the military platoons travel costs.

Considering Configuration 1 (Instances 11–15) in Table 2, we can see that the behaviors of Stages 3 and 4 stand out because, different from Stages 1 and 2, even with the increase in unit use costs and the decreased amount of OMs as a result of these costs, in those stages it was possible to achieve the lowest travel costs, even though there were higher OM unit use costs. Instances 12 and 13 presented a potential reduction of 13.29% and 12.60% for Stage 3 and 5.50% and 0.83% for Stage 4.

For the results of the instances related to Configuration 2, the travel costs are similarly distributed among the stages because they present the same number of survey stations served. As in the instances of Configuration 1, in two stages (Stages 2 and 3), the travelling costs are not increasing in all instances.

In the case of Instances 17 and 18, although the OM costs are high (Cost Groups 2 and 3) compared to Instance 16 (Cost Group 1) in Stages 2 and 3, the costs are 12.01% and 5.55% lower for Stage 2, and 7.66% and 7.67% lower for Stage 3.

When comparing the total travelling costs of the instances for Configurations 1 and 2, for instance pairs that have different cost groups (Instances 11 and 16, 12 and 17, successively up to Instances 15 and 20), the ones of Configuration 2 have better results with 75 survey stations in each stage, except the pair of Instances 14–19.

Potential gains in relation to the reduction of travelling costs are: 5.88% for Cost Group 1 (Instances 11 and 16); 5.86% for Cost Group 2 (Instances 12 and 17); 5.18% for Cost Group 3 (Instances 13 and 18); and 7.54% for Cost Group 5 (Instances 15 and 20). Even for Cost Group 4, where there was an increase in observed costs, this was only 0.14%.

As expected initially, due to the good results of the instances of Configuration 2 in terms of the number of OMs used in all comparisons between instances of Configuration 1 and 2, Instances 16, 17, 18, 19 and 20 present potential reductions in OM costs due to the equal division of survey stations between Stages 1, 2, 3 and 4.

Considering the OM costs, a significant increase in the values in function of the cost groups can be seen whereas the best results (lower costs) were those of Configuration 2. The reductions of comparative cost percentages are: 3.59% for Cost Group 1; 3.85% for Cost Group 2; 5.84% for Cost Group 3; 13.85% for Cost Group 4; and 12.30% for Cost Group 5.

As for the best solutions value, although there are benefits of the Configuration 2 compared to Configuration 1, for three cost groups (Cost Groups 1, 2 and 3), the potential gain with the division of the survey stations by 75 stations per stage is lower than 6.00%. In the comparison between Instances 11 and 16, the reduction is 5.84%, whereas between Instances 12 and 17 it is 5.68% and for Instances 13 and 18 the reduction is 5.37%. In the case of Cost Groups 4 and 5, the reductions provided by Configuration 2 are 7.33% and 11.35%, respectively.

Table 2 shows that for instances 15, 19 and 20, the optimal solution was not found within 7200.00 s . For these cases, the residual GAPs are 0.29%, 1.29% and 1.29%, respectively.

Although the residual GAP% found for the Instances 15, 19 and 20 are small (less than 2.00%), when we compare the results of Scenario 2 with those of Scenario 1, we can see that the mathematical model is sensitive to the number of survey stations and to the number of stages. Such as in Scenario 1, in the instances related to cost groups with higher OM costs, the time required to solve the instances was larger. In addition, when the number of survey stations was increased (from 164 to 300) and the number of survey stages increased from three to four, the computational times increased.

5.3. Results of Scenario 3

This scenario presents five instances, one for each cost group of Table 1. As noted in the previous scenarios, again when the OM costs increase, the results indicate that a smaller number of OMs is required to serve the survey stations. In the case of Instance 21, which presents the smallest OM costs, 63, 66, 62, 67 and 63 OMs were selected, respectively, for Stages 1, 2, 3, 4 and 5.

The results for Instance 22, which considers the Cost Group 2, are lightly better than the ones found for Instance 21. The most significant improvements are noted for Instance 23, for which the percentages of reduction in relation to Instance 21 are: 31.75% (Stage 1); 30.33% (Stage 2); 30.65% (Stage 3); 43.28% (Stage 4); and 23.81% (Stage 5).

In the case of Instance 24, the reductions are even larger compared to Instance 21, reaching 59.70% (Stage 4). However, the best results in terms of the number of OMs per stage are found in Instance 25, which requires between 22 and 25 OMs per stage to serve all the survey stations. It is important to remember that in this scenario the number of survey teams varies from 1 to 5 while in Scenarios 1 and 2 a maximum of 3 military platoons is available in each OM.

Considering the number of different OMs used throughout the traffic survey needed to serve the survey stations, Table 2 shows that between Instances 21 and 25 there is a larger variation (81.29%). In Instance 21, 310 OMs are assigned to serve the 500 survey stations, whereas only 58 OMs are needed in Instance 25.

Figures 4a and 4b show, respectively, the results for Instances 21 and 25. These figures have the same behavior of Figures 3a and 3b. When high costs are used for OMs, large travelled distances are required to serve all survey stations. These figures clearly show the trade-off mentioned before.

Considering now the travel costs per survey, we can note that Instances 24 and 25 present the highest travelling costs in the survey stages. Instance 24 has costs between 37.72% (Stage 1) and 59.39% (Stage 3) higher than those of Instance 23. In the comparison between Instances 24 and 25, the increase in costs is between 81.40% (Stage 5) and 123.85% (Stage 1).

Contrary to the trend of increasing travelling costs in the variation between Cost Groups 1 and 2 (Instance 21 for Instance 22) in Stages 2 and 3, and in the variation between Cost Groups 2 and 3 (Instance 22 for Instance 23) in Stage 2, small negative percentage variations in costs are observed. In the first case (Cost Group 1 and Cost Group 2, Stages 2 and 3), there is a reduction in travelling costs of 1.36% and 4.76%, and in the second case (Cost Group 2 and Cost Group 3, Stage 2), there is a reduction of 0.12%.

However, when the analysis is made for total travel costs throughout the traffic survey, there is a trend of increasing travelling costs as a function of the OM costs. When Instances 25 and 21 are compared, it is possible to observe that the total travelling cost is 250.97% higher. However, in the comparison between Instances 21 and 22, this increase is only 1.55%. This fact reinforces the need for an analysis that considers the trade-off.

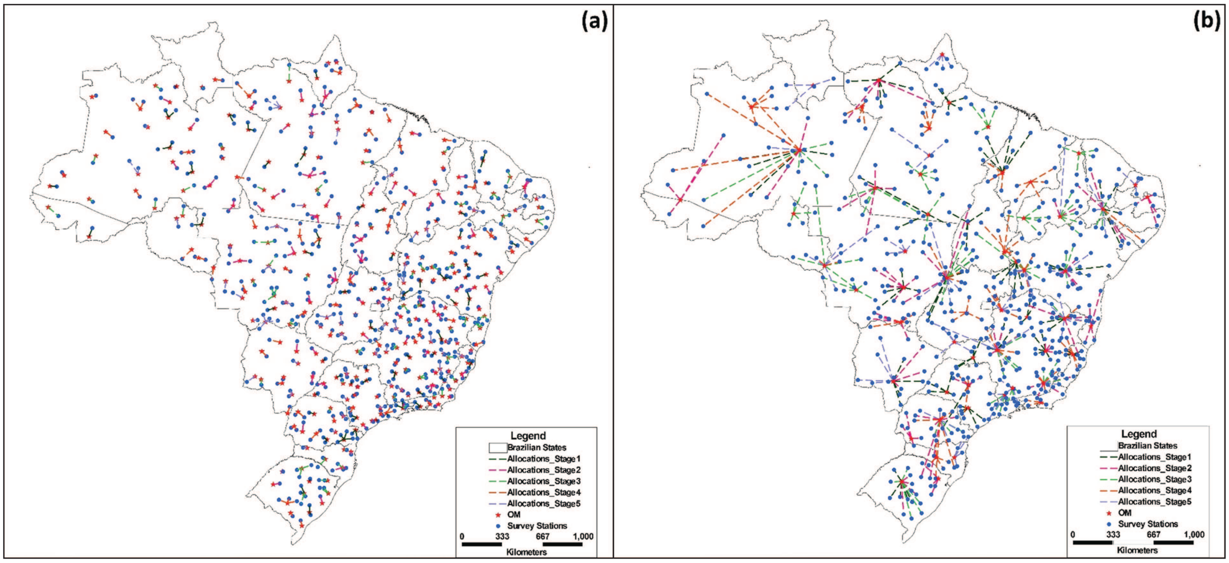

FIGURE 4. Comparative map between results of (a) Instance 21 and (b) Instance 25.

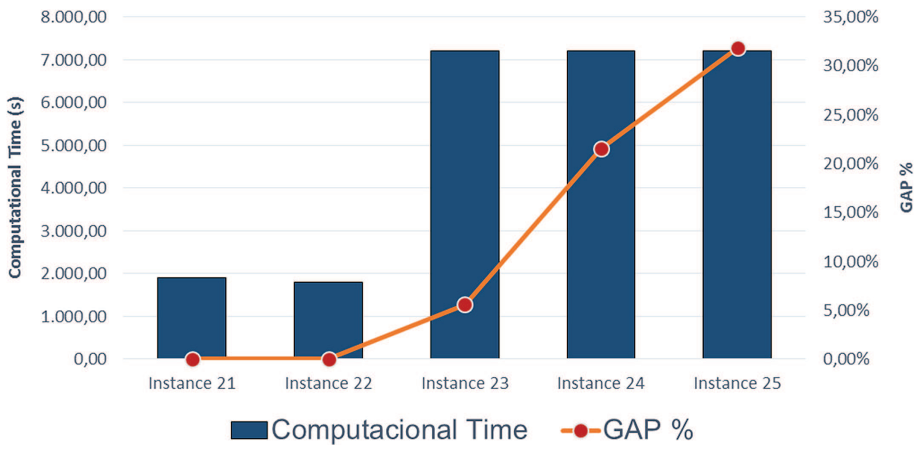

FIGURE 5. Computational time (a) and residual GAP (b) for Scenario 3 instances.

The results of this scenario based on the results of the previous ones show the trade-off between the number of OMs used, the OM costs and the travel costs during the entire traffic survey planning.

The behavior of the best solution value, see Table 2, is similar to that of OM costs. The results show that, in spite of the reduction that occurred in the number of OMs used due to the increase in OM costs, this was not enough to reduce the best solution value.

With respect to computational time and complexity of the mathematical model solution process, a significant increase can be observed. In Instances 21 and 22, while the highest time found in the other tests for Cost Groups 1 and 2 (Instances 1, 2, 6, 7, 11, 12, 16 and 17) was lower than 42 s, Scenario 3 required 1903.30 and 1804.00 s for Instances 21 and 22, respectively, as seen Figure 5.

CPLEX did not find the optimal solution for Instances 23, 24 and 25 with a maximum time of 7200.00 s. Furthermore, it is important to note that the evidence found in Scenario 2, referring to the computational complexity associated with the number of survey stations and the parameters of the model, are confirmed in this scenario.

For the instances of Scenario 3, the GAP value increases significantly from 5.59% in Instance 23 to 21.57% in Instance 24 and to 31.83% in Instance 25. Thus, this scenario indicates that the implementation of new solution techniques, such as heuristics and metaheuristics, are necessary in larger problems.

6. Final remarks

Through the theoretical discussion presented it was possible to better understand the operation of large road traffic surveys already performed in Brazil and, at the same time, to identify the need of developing new tools to assist the planning process of these surveys.

One of these identified needs is related to the locating and choice process of support units to serve survey stations. Due to their high complexity, road traffic surveys require a great mobilization of resources (financial, material and manpower) that need to be managed responsibly, especially the larger surveys.

Thus, the present study contributes to the planning of the surveys in a way that allows, through an optimization mathematical model, the efficient use of the resources related to survey teams, from the support units to the survey stations, and the use of these facilities. Therefore, the main objective of this paper was reached, *i.e.*, the Support Unit Location Problem to Assist Road Traffic Survey with Multi Stages (SULPARTSMS) was presented and mathematically modelled. The model enables, given a set of possible support units, the selection of which ones must be used to assist traffic survey stations whereas the cost is minimized.

With the three scenarios proposed for the computational experiments of the model, we could: (Scenario 1) validate the mathematical model when comparisons are made with PNT 2011; (Scenario 2) perform analyses related to the application of the mathematical model in the new traffic survey for the years 2016 and 2017 with two possible configurations of division of the survey stations between the stages; and (Scenario 3) study the behavior of the mathematical model and the results obtained for a hypothetical road traffic survey larger than the previous ones used.

In order to analyze the results, performance indicators were established that evaluate: the number of support units used per stage; the total number of different support units used during the entire traffic survey; the total costs related to the travelling of the survey teams to the survey stations; the costs incurred with the use of support units during the entire traffic survey; the total costs (Objective Function); and those related to computational performance, such as execution time and GAP% residual.

For future researches, we believe that the mathematical modeling proposed in this study can be expanded or even incorporated into the process of planning the location of survey stations. In this study, the location of the survey stations is known and used as an input. However, location processes involve complex questions in terms of network analysis (road mapping and geo-referencing, previous information about the origin and destination of journeys, and annual average daily traffic, among others) that directly influence the process of defining the best places for the survey stations. Therefore, it is expected that a mathematical modeling that simultaneously determines the location of the survey stations and the support units will allow a better planning.

Another possible approach is related to the specific studies for detailing the composition and more precise estimates for the costs related to travelling and the use of the support units, supporting, then, the correct determination of these parameters for the model. In addition, the computational experiments, especially for Scenario 3, show that the complexity involved in solving these problems justifies studies focused on specific optimization methods, such as heuristics and metaheuristics, in order to obtain good solutions with a reduced computational time.

Acknowledgements. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001, Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq (process 454569/2014-9) and Carlos Chagas Filho Research Support Foundation – FAPERJ (process 203.231/2017). Thanks are due to the referees for their valuable comments.

References

[1] R.H. Ballou Business Logistics/Supply Chain Management: Planning, Organizing, and Controlling the Supply Chain. Pearson Education India (2007).

[2] L.M.C. Bataglin and D. Alem, The location-distribution problem in the megadisaster of the Mountain region in Rio de Janeiro. *Gest. Prod.* **21** (2014) 865–881.

[3] S. Bera and K.V. Rao, Estimation of origin-destination matrix from traffic counts: the state of the art. *Euro. Transp. (Trasporti Europei)* **49** (2011) 2–23.

[4] C. Berge, Two theorems in graph theory. *Proc. Natl. Acad. Sci. USA* **43** (1957) 842–844.

[5] R. Bhatnagar, J. Jayram and Y.C. Phua, Relative importance of plant location factors: a cross national comparison between Singapore and Malaysia. *J. Business Logistics* (2003) 147–170.

[6] L. Bianco, G. Confessore and P. Reverberi, A network based model for traffic sensor location with implications on O/D matrix estimates. *Transp. Sci.* **35** (2001) 50–60.

[7] M.A.F. Caldas, J.C.C.B.S. Mello, L.A. Meza and F. Azevedo, Uso do Método de Copeland Hierárquico para Localização de Terminal Multimodal de Cargas (Use of the Copeland method for the location of the multimodal cargo terminal), in *XXIII Congresso de Pesquisa e Ensino em Transportes – ANPET*, Vitória/ES (2009).

[8] Caliper, TransCAD Transportation GIS Software – Routing and Logistics with TransCAD 5.0. Caliper Corporation, Newton, EUA (2008).

[9] A. Chen, S. Pravinvongvuth, P. Chootinan, M. Lee and W. Recker, Strategies for selecting additional traffic counts for improving OD trip table estimation. *Transportmetrica* **3** (2007) 191–211.

[10] P. Chootinan, A. Chen and H. Yang, A bi-objective traffic counting location problem for origin-destination trip table estimation. *Transportmetrica* **1** (2005) 65–80.

[11] P.T. Chuang, Combining the analytic hierarchy process and quality function deployment for location decision from a requirement perspective. *Int. J. Adv. Manuf. Technol.* **18** (2001) 842–849.

[12] CNT – Confederação Nacional do Transporte, 2015, Preços da Gasolina e do Diesel (Prices of Gasoline and Diesel). Available at: http://www.cnt.org.br/Paginas/Agencia_Noticia.aspx?n=10030&t=aumento-de-impostos-elevar%C3%A1-pre%C3%A7o-da-gasolina-e-do-diesel (2015).

[13] M.B.B. Costa, *Utilização De Modelo De Localização-Alocação Para Identificação De Zoneamento Logístico Integrado Ao Planejamento Estratégico De Transportes* (*A location-allocation model for integrated logistic zoning identification to strategic transportation planning*). Ph.D. in Dissertation in Transportation Engineering, Federal University of Rio de Janeiro. Rio de Janeiro, Brasil (2014).

[14] DNIT – Departamento Nacional de Infraestrutura de Transportes, Plano Nacional de Contagem de Tráfego (National Traffic Counting Plan). Available at: http://www.dnit.gov.br/planejamento-e-pesquisa/planejamento/contagem-de-trafego (2015).

[15] A. Ehlert, M.G.H. Bell and S. Grosso, The optimisation of traffic count locations in road networks. *Transp. Res. Part B: Methodol.* **40** (2006) 460–479.

[16] F.A. Galvão, N.D.F. Gualda and C.B. Cunha, An application of the Analytic Hierarchy Process (AHP) for locating a distribution center. In: *International Symposium on the Analytic Hierarchy Process, Bali, Indonésia* (2003).

[17] L. Gan, H. Yang and S.C. Wong, Traffic counting location and error bound in origin-destination matrix estimation problems. *J. Transp. Eng.* **131** (2005) 524–534.

[18] M. Gastaldi, G. Gecchele and R. Rossi, Estimation of annual average daily traffic from one-week traffic counts. A combined ANN-Fuzzy approach. *Transp. Res. Part C: Emerg. Technol.* **47** (2014) 86–99.

[19] Q. Ge and D. Fukuda, Updating origin–destination matrices with aggregated data of GPS traces. *Transp. Res. Part C Emerg. Technol.* **69** (2016) 291–312.

[20] M. Gentili and P.B. Mirchandani, Locating sensors on traffic networks: Models, challenges and research opportunities. *Transp. Res. Part C: Emerg. Technol.* **24** (2012) 227–255.

[21] S.L. Hakimi, Optimum locations of switching centers and the absolute centers and medians of a graph. *Oper. Res.* **12** (1964) 450–459.

[22] S.L. Hakimi, Optimum distribution of switching centers in a communication network and some related graph theoretic problems. *Oper. Res.* **13** (1965) 462–475.

[23] S.R. Hu and H.T. Liou, A generalized sensor location model for the estimation of network origin–destination matrices. *Transp. Res. Part C: Emerg. Technol.* **40** (2014) 93-110.

[24] S.R. Hu, S. Peeta and C.H. Chu, Identification of vehicle sensor locations for segmento-based network traffic applications. *Transp. Res. Part B: Methodol.* **43** (2009) 873–894.

[25] IBM, IBM ILOG CPLEX versão 12.6 (2015).

[26] M. Kayhanian, A. Singh, C. Suverkropp and S. Borroum, Impact of annual average daily traffic on highway runoff pollutant concentrations. *J. Environ. Eng.* **129** (2003) 975–990.

[27] H.J. Kim, H.I. Chung and S.Y. Chung, Selection of the optimal traffic counting locations for estimating origin-destination trip matrix. *J. East Asia Soc. Transp. Stud.* **5** (2003) 1353–1365.

[28] R.A. Lopez and N.R. Henderson, The determinants of location choices for food processing plants. *Agrobusiness* **5** (1989) 619–632.

[29] L.A.N. Lorena and E.L.F. Senne, A column generation approach to capacitated p-median problems. *Comput. Oper. Res.* **31** (2004) 863–876.

[30] M.T. Melo, S. Nickel and F. Saldanha-da-Gama, Facility location and supply chain management – a review. *Eur. J. Oper. Res.* **2** (2009) 401–412.

[31] Ministério dos Transportes, Rodovias: Contagem de tráfego nas rodovias terá maior precisão (Highways: Traffic counts on highways will have better precision). Available at: http://transportes.gov.br/component/content/article.html?id=3104 (2015).

[32] Ministério dos Transportes, DNIT, Pesquisa Nacional de Tráfego – PNT (National Traffic Survey – PNT). Brasília: Secretaria de Política Nacional de Transportes, DF (2011).

[33] B. Moya-Gómez and J.C. García-Palomares, Working with the daily variation in infrastructure performance on territorial accessibility. The cases of Madrid and Barcelona. *Eur. Transp. Res. Rev.* **7** (2015) 20.

[34] P. Rebreyend, L. Lemarchand and R. Euler, A computational comparison of different algorithms for very large p-median problems. In: Evolutionary Computation in Combinatorial Optimization. Springer International Publishing (2015) 13–24.

[35] B.C. Romero, *Análise de Localização de Plataformas Logísticas: Aplicação ao Caso do ETSP – Entreposto Terminal São Paulo da CEAGESP* (*Logistics Platform Location Analysis: Application to the Case of ETSP – Terminal Warehouse São Paulo of CEAGESP*). Dissertação de Mestrado, Engenharia de Sistemas Logísticos, Escola Politécnica da Universidade de São Paulo – POLI/USP, São Paulo/SP (2006).

[36] H. Shao, W.H.K. Lam, A. Sumalee and M.L. Hazelton, Estimation of mean and covariance of stochastic multi-class OD demands from classified traffic counts. *Transp. Res. Part C: Emerg. Technol.* **59** (2015) 92–110.

[37] J. Short and A. Kopp. Transport infrastructure: Investment and planning. Policy and research aspects. *Transp. Policy* **12** (2005) 360–367.

[38] D.J. Sun, Y. Chang and L. Zhang, An ant colony optimisation model for traffic counting location problem. *Proc. Inst. Civ. Eng. Transp.* **165** (2012) 175–185.

[39] A. Suzuki and Z. Drezner, The p-center location problem in an area. *Location Sci.* **4** (1996) 69–82.

[40] B.C. Tansel, R.L. Francis and T.J. Lowe, State of the art–location on networks: a survey. Part I: the p-center and p-median problems. *Manage. Sci.* **29** (1983) 482–497.

[41] C. Toregas, R. Swain, C. Revelle and L. Bergman, The location of emergency service facilities. *Oper. Res.* **19** (1971) 1363–1373.

[42] I. Tsapakis and W.H. Schneider IV, Determining the optimal number of seasonal adjustment factor groupings when estimating annual average daily traffic and investigating their characteristics. *Transp. Plan. Technol.* **38** (2015) 181–199.

[43] P. Waddell, Integrated land use and transportation planning and modelling: addressing challenges in research and practice. *Transp. Rev.* **31** (2011) 209–229.

[44] M. Yaghini, M. Karimi and M. Rahbar, A hybrid metaheuristic approach for the capacitated p-median problem. *Appl. Soft Comput.* **13** (2013) 3922–3930.

[45] H. Yang, Y. Lida and T. Sasaki, An analysis of the reliability of an origin–destination trip matrix estimated from traffic counts. *Transp. Res. Part B: Methodol.* **25** (1991) 351–363.

[46] H. Yang, C. Yang and L. Gan, Models and algorithms for the screen line-based traffic-counting location problems. *Comput. Oper. Res.* **33** (2006) 836–858.

[47] H. Yang and J. Zhou, Optimal traffic counting locations for origin–destination matrix estimation. *Transp. Res. Part B: Methodol.* **32** (1998) 109-126.

[48] J. Yang and H. Lee, A ahp decision model for facility location selection. *Facilities* **15** (1997) 241–254.

[49] P.K.N. Yim and W.H.K. Lam, Evaluation of count location selection methods for estimation of O–D matrices. *J. Transp. Eng.* **124** (1998) 376–383.

[50] F. Zhao and S. Chung, Contributing factors of annual average daily traffic in a Florida county: exploration with geographic information system and regression models. *Transp. Res. Rec. J. Transp. Res. Board* **1769** (2001) 113–122.

[51] X. Zhou and G.F. List, An information-theoretic sensor location model for traffic origin-destination demand estimation applications. *Transp. Sci.* **44** (2010) 254–273.