# Optimal Bus Stops' Allocation: A School Bus Routing Problem with Respect to Terrain Elevation

Klemen PRAH[1], Abolfazl KESHAVARZSALEH[2], Tomaž KRAMBERGER[1], Borut JEREB[1], Dejan DRAGAN[1*]

[1] University of Maribor/Faculty of Logistics, Celje, Slovenia
[2] University of Malaya (UM) and International University of Malaya-Wales (IUMW), Kuala Lampur, Malesya

[Corresponding Author indicated by an asterisk *]

*Abstract* - The paper addresses the optimal bus stops allocation in the Laško municipality. The goal is to achieve a cost reduction by proper re-designing of a mandatory pupils' transportation to their schools. The proposed heuristic optimization algorithm relies on data clustering and Monte Carlo simulation. The number of bus stops should be minimal possible that still assure a maximal service area, while keeping the minimal walking distances children have to go from their homes to the nearest bus stop. The working mechanism of the proposed algorithm is explained. The latter is driven by three-dimensional GIS data to take into account as much realistic dynamic properties of terrain as possible. The results show that the proposed algorithm achieves an optimal solution with only 37 optimal bus stops covering 94.6 % of all treated pupils despite the diversity and wideness of municipality, as well as the problematic characteristics of terrains' elevation. The calculated bus stops will represent important guidelines to their actual physical implementation.

*Key words*—Logistics, maximal covering problems, optimization, data clustering, Monte Carlo simulation, geographic information system (GIS), reduction of transportation costs, Laško, Slovenia.

## I. INTRODUCTION

The paper discourses the question of reduction of a specific type of costs, which are caused by mandatory transportation of specific categories of pupils in the municipality Laško, Slovenia. This transportation, which is enforced by government law, includes transport of pupils from their homes to the corresponding primary schools and vice versa. During the last years, such kind of transportation regularly reaches an enormously high level of costs, which can escalate even up to one million EUR per year. Since this is intolerable for such a small municipality, there exists a strong wish to decline these costs as much as possible.

The primary reason for the high level of costs is a momentary inefficient organization of transport. Namely, transportation vehicles, hired by the municipality from the outsourcers, are used for picking up and delivery of every single pupil just at his home doors. Furthermore, the transport routes are completely disorganized in the sense of unnecessarily driven additional kilometers; vehicles are frequently not fully loaded, routes are sometimes needlessly doubled or even tripled, etc. All of these problems become even worse when a rural character and a dynamic terrain are additionally taken into consideration, with home addresses of the pupils being strongly spatially dispersed.

Solving of afore-mentioned problems can be looked through a prism of resolving the so-called School Bus Routing Problem (**SBRP**) [1]. Normally, the SBRP can be solved using the following five steps [2]: 1.) data preparation; 2.) optimal bus stop selection (with a pupils' assignment to the optimal bus stops); 3.) optimal bus route generation; 4.) school bell time adjustment; 5.) optimal route scheduling.

The second step here refers to the so-called problem of optimal Bus Stops Allocation (**BSA**). In the paper, we restrict ourselves only to discuss about this sub-problem. Thus, similarly, as it was done, e.g. in the study [3], we do not cover all features of a typical SBRP problem. Conversely, we only refer to the problem of calculating the set of optimal bus stops (**OBS**) and assigning children to the nearest stop, where constraints on the distance between the bus stops and places of pupils' homes are also considered.

To solve the BSA problem as a sub-problem of SBRP problem, we have used a so-called "maximum location covering problem" (**MLCP**) approach from the field of a facility location theory [4, 5]. Solving the MLCP problem implies that the minimum fixed number of facilities must be located, which should maximize the service area [6, 7]. Thus, in our case, the number of bus stops should be minimal possible that still assure a maximal service area within a prescribed radius, while keeping the minimal possible average walking distance (**AWD**) of the pupils to the nearest stop [7].

On one side, the allocation of bus stops implies that pupils are required to walk a certain (reasonably long) distance from their homes to the nearest bus stops, which means that some comfort will be lost. However, on the other hand, the transportation vehicles would have to drive along a significantly shorter quantity of routes (distances). Hence, some compromise between two distant interests (an interest of pupils/parents versus the municipality interest) must be reached in such cases.

The algorithm for optimal BSA is based on the data clustering procedures and the Monte Carlo (**MC**) simulation method. Data clustering is based on the unique composition of hierarchical and non-hierarchical clustering that is mounted within the MC framework, while the whole mechanism is interconnected with the three-dimensional GIS data.

We believe that the paper has the following contributions: 1. There have not been many similar studies detected that would apply a three-dimensional data approach; 2. The proposed algorithm can be treated as an original not only in the field of interest but also in other similar research;  3. The results show that the suggested algorithm reaches a best solution with only 37 OBS covering 94.6 % of all addressed pupils despite the diversity and wideness of municipality, as well as the rough characteristics of terrains' elevation.


*A.   Characteristics of the previous research*

This research is a logical continuation of the previous work, presented in [7-10], where a Geographic Information system (**GIS**) [11] and its modules (*ArcLogistics* 9.3, *ArcView* 9.31, *Network Analyst*) were also applied. The following has been done in the previous research: 1.) Solving of the BSA problem to get the OBS positions; 2.) Computation of the optimal driving routes, schedules and driving  fleet considering the given schools' positions and road network characteristics, as well as the previously calculated OBS; 3.) Estimation of the driven kilometers and cost savings for the simulated optimized case if compared to the real un-optimized case; 4.) Estimation of the amount of reduced tons of $CO_2$ emissions for the optimized case, for which the simulated vehicles' km traveled (**VKT**) have been noticeably lowered (for almost 1000 km/day) in comparison with the real un-optimized kilometers.

One of the most significant deficiencies was that we had used a two-dimensional (2D) GIS, where a terrain elevation was not considered. Consequently, all the calculations were slightly distorted meaning that the results were not sufficiently optimal. So we have decided to keep moving with our research and additionally consider the third dimension by conducting three-dimensional (3D) GIS data to more accurately capture the dynamic properties of a municipality terrain. Consequently, besides some significant methodological improvements, this paper reports about replicated 3D BSA results instead of 2D results.


*B.   Bus stops allocation as an MLCP problem from a facility location theory*

Church and ReVelle first introduced the MLCP in 1974, when they applied a relaxed linear programming and branch and bound procedure [6, 12]. MLCP problem is significant for many cases in the public, as well as in the private sector ([2, 10, 13, 25]). In the public sector, authorities need to determine bases for emergency highway patrol vehicles, or they must locate fire stations, police stations, and ambulances. In all of these cases, poorly chosen locations can increase the possibility of damage or loss of life [13].

Greedy heuristic algorithms [6, 13] are also often used to solve MLCP problems, as well as other much more advanced heuristics like Genetic Algorithms [14, 15], Simulated Annealing [16, 17], Tabu Search [18, 19], Lagrangean relaxation [13, 20, 21], Lagrangean/Surrogate heuristics (Lorena, L., Pereira, M. 2002), etc.

Worwa [3] has provided an excellent formal presentation of a greedy method in the context of solving the BSA problem. Many of the MLCP approaches afore-mentioned can also be conducted for solving of the BSA problems, depending on the characteristics of the problem. Gleason was the first, who has done a pioneer work about bus stop allocation using the location theory [22]. Later, stop location problems have been considered in many other studies within the scope of facility location theory (see excellent surveys [23-25] and references therein).

*C.  Bus stops allocation as a sub-problem of the SBRP problem*

The BSA can also be considered as a sub-problem in the context of the SBRP problem if we are dealing with a design (re-design) of bus stops mostly dedicated to the pupils–commuters. Numerous authors (e.g. [1, 26]) have discussed the SBRP, and a common thread of these papers is that optimization reduces VKT and hence transportation costs.

However, surprisingly, explanation of BSA procedures and assignment of pupils to the bus stops are often omitted in the SBRP literature by assuming that the bus stops are already given, while the pupils are already assigned. Moreover, not many papers are dedicated to the explanation of the BSA procedures only, and most of them use certain types of heuristic algorithms, usually classified into two groups of methods: 1.) LAR (Location-Allocation-Routing) methods; and 2.) ARL (Allocation-Routing-Location) methods [1, 3]. Further details about different heuristic methods used in both approaches (LAR and ARL) can be found in [1].

## II. PROBLEM DEFINITION, GEOGRAPHICAL PROPERTIES OF TERRAIN, AND GIS PROCESSING

*A.  Geographical, terrain and general characteristics of the Municipality*

The Municipality can be with 189 $km^2$ and some 13.300 inhabitants classified as the medium-sized municipality in central Slovenia [27]. It is characterized by sub-alpine hills, while the flatlands spread along the Savinja River and its tributaries.

There live only 67 inhabitants per $km^2$, which represents a 65.8 % of average Slovenian population per $km^2$ [27]. Most significant settlements are located in valleys, while many smaller dispersed settlements and isolated homesteads are spread on a higher landscape territory. The Laško town is a central city with some 3.300 inhabitants and represents a nucleated urban settlement of the municipality, located in the extended part of Savinja Valley[28, 29].

Whereas valleys and ravines highly segment a landscape in the region, it is hardly passable [30]. This is also the reason why the municipality possesses one of the most diverse road networks in the country, with as much as 30 m of roads per citizen when compared with the Slovenian average of only 7 m [29]. One main road along Savinja valley, some regional roads and a large number of local roads lead through the municipality. To summarize, the municipality has a lot of settlements, a complex terrain, a current low population density, and a large number of individual road segments of the poor quality of construction [28, 29].

*B.  Problem definition and spatial distribution of the pupils and their schools*

Figure 1 depicts the locations of 562 treated pupils' home addresses (blue points) in the municipality (shadowed area). In the research, we were forced to exclude about 10% (59 pupils' addresses) of the most problematic pupils' locations due to extremely difficult and remote terrain accessibility.

For a given pupils' homes distribution, the goal is to find such (reasonably small) number of OBS from the massive group of the previously road-fragmented BSC points, which will cover the biggest possible number of pupils' addresses, while the AWD distance to the nearest stop will be minimal possible. Figure 1 also shows the positions of pupils' schools (little houses in yellow color) and their total number is 11.
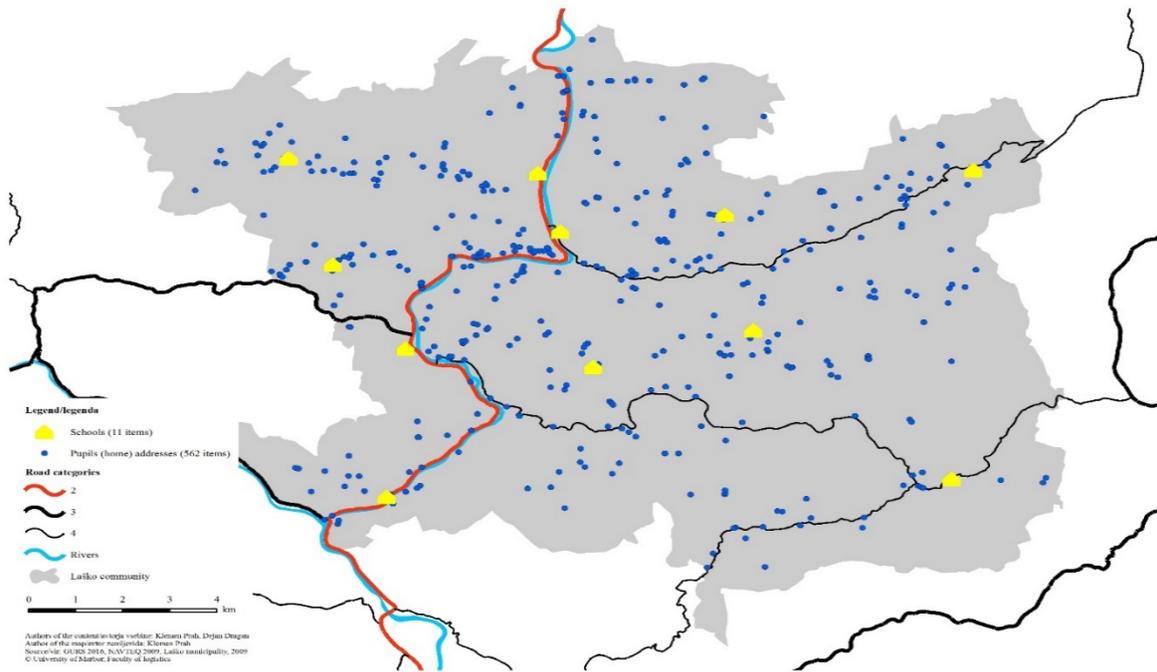
Figure 1. Positions of 562 pupils' addresses and 11 schools in the municipality of Laško. Roads of different categories split the municipality. The river Sava is illustrated at the lower left corner of the figure (blue line).

Figure 2 explains, why we have decided to treat the BSA problem in a 3D space instead of 2D space. Namely, in reality, pupils walk a distance from their homes to the potential bus stop that is not always nearly horizontal, but mostly quite inclined and therefore (sometimes significantly) longer. Thus, distortion of 2D results could have appeared in our previous studies, since the differences in height might have caused quite longer walking distances between the observed surfaces' locations than it was supposed (c.f. figure2). Specifically, taking into account a hypotenuse $d_1$ (as done in this study) is much more realistic than a straight line $d_2$ (as done in the previous studies). As will be later seen, we have had also incorporated certain randomness while calculating the (euclidian) hypotenuse $d_1$ with an intention to approach closer to the reality. This fact later essentially influences on a quality of optimality of calculated OBS since an accurately estimated walking distance is also one of the important constraints in our optimization process. Namely, we have incorporated three important types of distances as constraints, which are shown in Table 1.

Table 1. The three important types of distances involved as constraints in our optimization process.

| Type of distance | Meaning of distance |
|---|---|
| $d_{TOT}$ - **TWD** | The total walking distance of all pupils to their nearest bus stops |
| $\bar{d}_{avg}$ – **AWD** | The average walking distance that every pupil must walk to the nearest bus stop |

| $d_{max}$ - MWD | The maximal walking distance that is still acceptable for pupils' walking |
|---|---|

According to some recommendations (e.g. [31, 32] ), it is suggested that children should be motivated to walk reasonably long distances (up to 2 km) in urban areas due to several motives related to the health and sustainability reasons. However, in our circumstances, it should be taken into consideration that these recommendations might have quite been hard to implement in such a hilly non-urban terrain and simultaneously efficiently carry out solving of the addressed optimization problem. Thus, regarding the difficulties arisen from problematic and hilly (rural) terrain, we have decided that a slightly bigger distance $d_{max} = 2.5km$ might be still an acceptable maximal limit for walking.
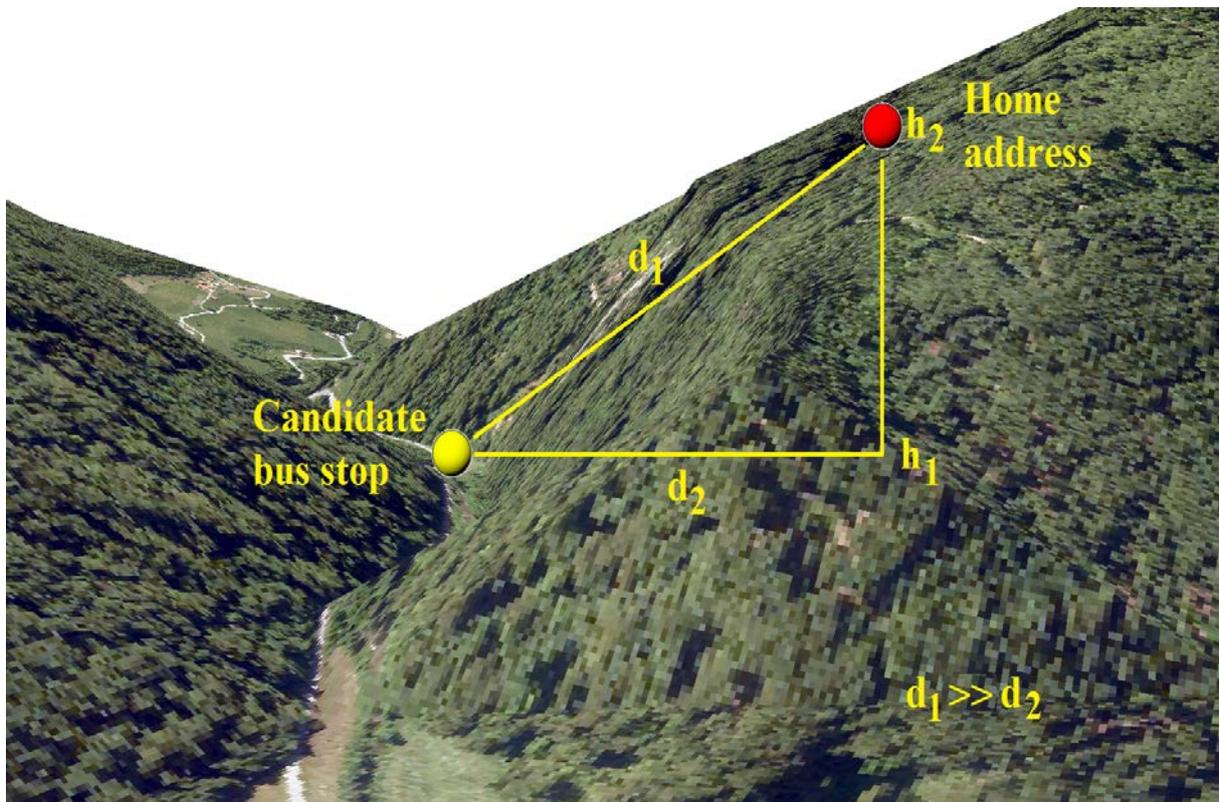


Figure 2. The dynamic complexity of the terrain: The main cause of a height difference between surface locations of home addresses and bus stops that occurs in many cases.

*C.   Processing and analyzing of adapted terrain's properties with a GIS system*

Figure 3 shows a surface of the entire municipality when its properties were adapted into a GIS system. The lowest terrain elevation in the municipality is 190.3 m, while the highest is 984.3 m. The terrain elevations between 300 and 600 m with 69.88 % of the area are those that prevail in the municipality [33].

As can be seen in figure 3, there is also an enlarged section (excerpt) of the municipality shown. Here, the fragmented road points (yellow color circles) that represent BSC candidates and some of the pupil's addresses (red color circles) can be clearly noticed. BSC points have been generated by 100 meters segmentation of every single road, and their total number collected by the GIS system is 9585. Since a such number of BSC points dispersed across the whole municipality is enormous, it should be somehow reduced. Namely, our final goal related to the financial reasons is to decrease

a total number of BSC candidates to an acceptable, reasonably low level that would still cover a maximal possible number of pupils.
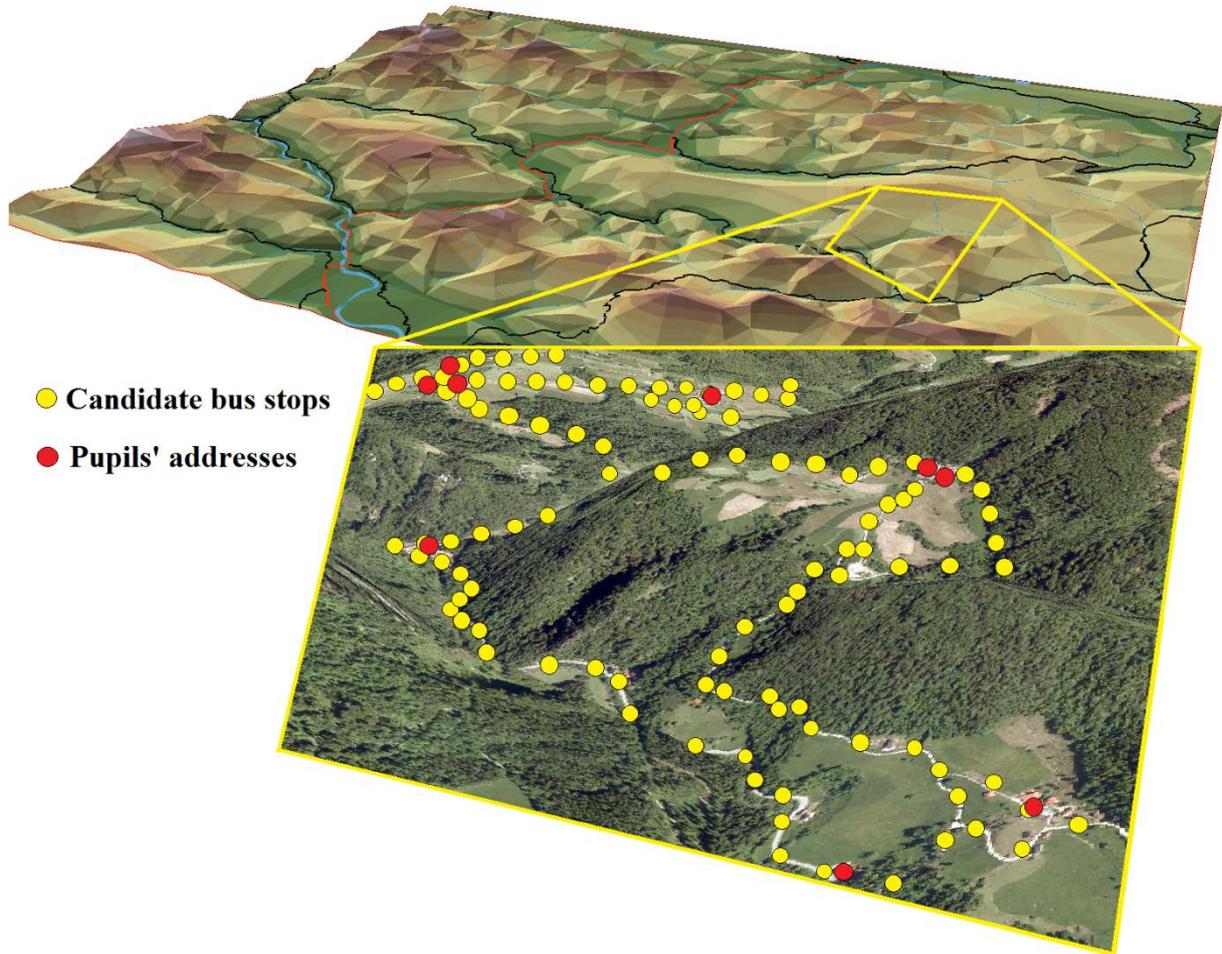


Figure 3. The terrain of the Municipality of Laško; Enlarged section (excerpt) of the municipality that illustrates BSC candidates (yellow circles), which were generated by 100 meters segmentation of every road, while the PA addresses are marked with red circles.

Similarly as in a case of collected BSC positions, also the positions of 562-59 (excluded) =503 pupil addresses (PA) were collected by means of GIS system. Some of the geographical properties of the PA points and corresponding schools are given in Table 2.

Table 2. Some of the geographical properties of the pupils' addresses (PA) and corresponding schools.

| Issue | Value |
|---|---|
| The height of the lowest PA | 201.6 m |
| The height of the highest PA | 714.6 m |
| % of all PA with a height lower than 500 m | 81.49 % |
| % of all PA with a height between 500 and 600 m | 15.12 % |

| | |
|---|---|
| % of all PA with a height higher than 600 m | 3.38 % |
| The height of the lowest school | 215.38 m |
| The height of the highest school | 570.48 m |
| Number of schools with a height lower than 300 m | Four schools (Savinja valley) |
| Number of schools with a height between 300 and 400 m | Three schools |
| Number of schools with a height between 400 and 500 m | Two schools |
| Number of schools with a height higher than 500 m | Two schools |

## III. METHODS

*A. The framework of entire research*

The heuristic optimization algorithms were used to solve the BSA problem. The algorithm for optimal BSA is based on the data clustering procedures and the Monte Carlo (**MC**) simulation method. Data clustering in our study is based on the unique composition of hierarchical and non-hierarchical clustering that is mounted within the MC framework, while the whole mechanism is interrelated with the three-dimensional GIS data.

Figure 4 shows the framework of entire research, which is intended to be carried out (based on 3D GIS data). As mentioned before, we focus on the BSA step only in this paper (see the highlighted frame in figure 4 – stage 1). Further stages of research (calculation of optimal bus routes, route schedules, and driving fleet, as well as an accurate estimation of reduced VKT and corresponding $CO_2$ emissions) are in progress and will be finished in the near future.
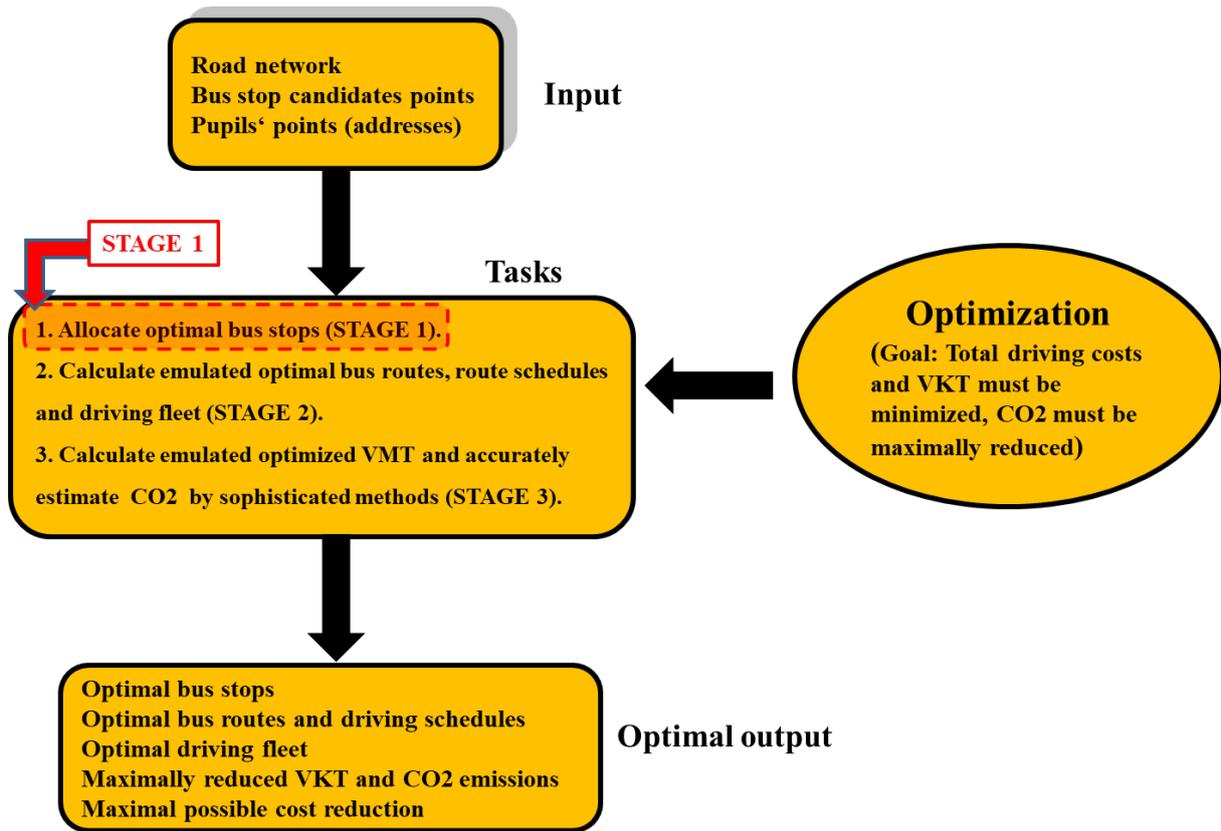
Figure 4. The framework of entire research (Stage 1- Optimal bus stops' allocation presented in this paper is highlighted within the red frame).

*B. Clustering*

Clustering is the key mechanism inside the heuristic procedure for optimal BSA. The latter is frequently used in data mining and multivariate statistical analysis for many scientific fields, from engineering, medical, life and social sciences all over to the computer science [34, 35].

Clustering is the method of grouping of a set of objects in such way that objects in the same group (cluster) are more similar to each other than to those in other groups (clusters) [36]. The classification into the connectivity-based clustering (hierarchical clustering (**HC**)) and Centroid-based (non-hierarchical) clustering (**NHC**) is usually used for the representation of two most typical families of clustering methods [36].

HC methods provide a wide-ranging hierarchy of clusters (i.e., dendrogram) that are merged at certain distances. Here, the "objects" are connected to form "clusters" based on their distance, where a cluster can be defined by the maximum distance needed to connect all of its parts. At different distances, different clusters will form, while the observation of a dendrogram can help to reveal the characteristics of any patterns present in the data. This way, it is relatively easy to choose an appropriate number of clusters.

Conversely, at NHC methods (usually *k-means* methods are used), the number of clusters must be predefined and fixed to some value *k* at first, while *k* cluster centers are formed in the second step. Finally, the objects are assigned to the nearest cluster center in such a way that the squared distances from the cluster are minimized [36]. In our approach, we have combined both types of methods, HC-based and NHC-based. Firstly, the HC-based method helped us to find a most appropriate number *k* of clusters. Afterward, the *k-means* algorithm with the previously obtained number *k* has been used to form clusters and group the closest BSC candidates around clusters' centers. Hence, by operating with the whole clusters of BSC candidates instead of each individual BSC point, we managed to simplify the entire procedure of optimal BSA. At this place it should be emphasized that we have also tried to cluster pupils instead of BSC candidates; however, in this case, the algorithm failed to give satisfactory results.

### C. Heuristic algorithm for bus stops' allocation

Figures 5 and 6 shows the main steps of the heuristic algorithm for BSA. The letters (A, B, …, J) of all blocks of both figures corresponds to each other. The BSC points are marked with a red color, while blue color highlights PA points in figure 6. As can be seen from figures 5 and 6 (block A in both figures), 9585 road data points (BSC) and 503 pupils' addresses (PA) points are collected using GIS. In the next step, an initial road data reduction is carried out (block B in both figure 5 and 6) relying on a heuristic rule that only those BSC are considered which are not too close to the nearest neighboring BSC. By this reduction, we have managed to obtain the reduced number of 2410 road points, which is a quite smaller number in comparison with the original 9585 road data points. Nevertheless, unfortunately, the reduced number of 2410 road data points is still too big in the sense of optimal BSA. Therefore, the further reduction of these points is somehow needed in order to lower the total number of optimal bus stops to the acceptable level.

For this purpose, the MC procedure is applied (see loop block C in figure 5), similarly as in our previous works (e.g. [7]). For each iteration $i_{mc}$ of the MC procedure, random three-dimensional radium $r(i_{mc}) \in N(\bar{r}, \sigma_r^2(i_{mc}))$ is applied at first (block D in figure 5), where mean $\bar{r}$ has been settled to a certain predefined constant value, while variability $\sigma_r^2(i_{mc})$ is changed by means of random generator at each iteration $i_{mc}$. Based on this radium all road data points that are more distant from any pupils' point than $r(i_{mc})$ are removed. This way, only those $N_{BSC}[r(i_{mc})]$ BSC points remain accumulated around pupils' points that are close enough to them (block E in figure 5). As it turns out, for the optimal iteration (discussed later), their total number is $N_{BSC}^*[r(i_{mc}^*)] = 1292$.

In the next step (block F in figure 5), $N_{CLU}[r(i_{mc})]$ clusters of the remaining $N_{BSC}[r(i_{mc})]$ BSC points are formed based on a clustering procedure. As it turns out, for the optimal iteration (discussed later) the total number of clusters is $N_{CLU}^*[r(i_{mc}^*)] = 49$ (clusters can be noticed in a block denoted by »F, G, H, I, J« in figure 6). As can be seen in this block, each cluster $C_i$ contains a certain number of BSC points, as well as some of the PA points.

Furthermore, the algorithm moves to block G in figure 5, where $P_{CLU}[r_{peak}(S_{N_{CLU}})]$ BSC points are randomly peaked for each cluster $C_i$. Here, $r_{peak}(S_{N_{CLU}})$ represents each cluster's »inner radium« that defines »inner circles« inside the cluster, while $S_{N_{CLU}}$ denotes the »strength« (density) of the cluster. Inner circles' centers are then fixed to the $P_{CLU}$ points meaning that these points become surrounded by inner circles. Since $r_{peak}$ depends on the density $S_{N_{CLU}}$ of the cluster, each cluster has inner circles with different inner radius $r_{peak}$.

Afterward (see block H in figure 5), each PA point in each cluster is assigned to the nearest $P_{CLU}$ point, if the mutual distance is smaller than $r_{peak}$ of the $P_{CLU}$ point's inner circle, otherwise, is marked as unassigned meaning that is not surrounded by any of inner circle. In the next step (block I in figure 5), for each cluster $C_i$, three types of distances are calculated for assigned pupils:

1. Mutual distances between $P_{CLU}$ points and corresponding PA points that are assigned to them:

$$d\left[P_{CLU}(i), PA(j)\right]_{i_{mc}} = d\left[P_{CLU}(i, i_{mc}), PA(j, i_{mc})\right] =$$

$$= d_{ij}(i_{mc}) \in N\left(d_{euclid,ij}(i_{mc}), \sigma_{d_{ij}}^2(i_{mc})\right), \tag{1}$$

$$i = 1, ..., N_{P_{CLU}}(i_{mc}), j = 1, ..., N_{PA_{ASSIGNED}}(i_{mc})$$

where implying the normally distributed randomness via $\sigma^2_{d_{ij}}(i_{mc})$ is conducted to incorporate the fact that in reality the real mutual distances are bigger than (hypothetical) euclidian distances.

2. Total distance as a sum of all mutual distances in (1):

$$d_{TOT}(N_{CLU}) = \sum_{i,j} d\left[P_{CLU}(i), PA(j)\right]_{i_{mc}} \qquad (2)$$

3. Average distance as a mean value of all mutual distances in (1):

$$\bar{d}_{avg}(N_{CLU}) = mean\left[d\left[P_{CLU}(i), PA(j)\right]_{i_{mc}}\right] \qquad (3)$$

In the final step of each MC iteration, all important results are saved for all clusters (see block J in figure 3).

A similar situation just described occurs for all other MC iterations. Namely, if next iterations of the Monte Carlo procedure are repeated, the situation in figures 5 and 6 persistently changes, since the random radium $r(i_{mc})$ is different at every further repetition. Consequently, the clusters are different at every repetition, and so are the assigned pupils' data points that correspond to $P_{CLU}(i, i_{mc})$ BSC points. Also, all distances of all clusters are persistently changed at each $i_{mc}$ repetition.
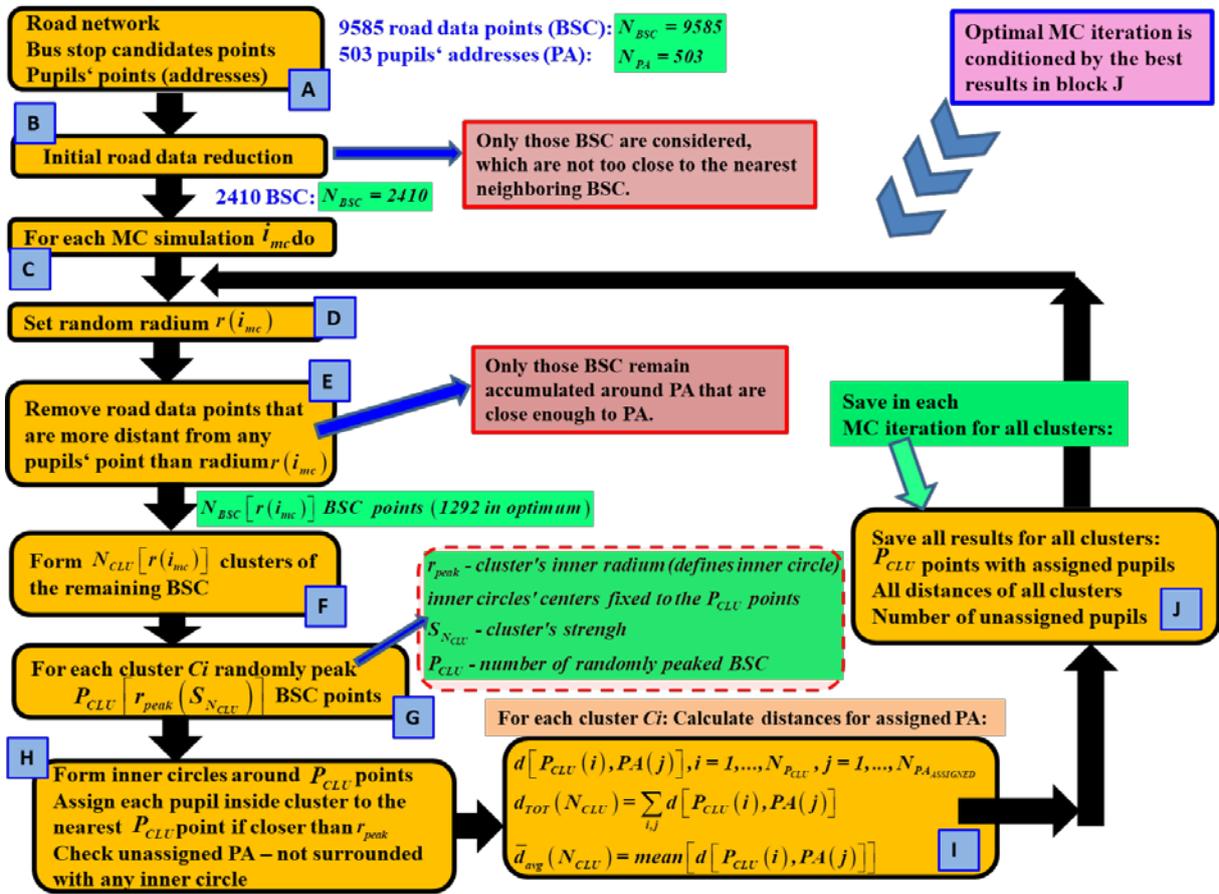


Figure 5. The main steps of the heuristic algorithm for bus stops' allocation.

Thus, when all $i_{mc} = 1, ..., N_{MC}$ iterations of the Monte Carlo procedure are finished, the following set of values of distances of all clusters $C_k(i_{mc}), k = 1, ..., N_{CLU}(i_{mc})$ will be formed:

$$d\left[P_{CLU}(i, i_{mc}), PA(j, i_{mc})\right], i = 1, ..., N_{P_{CLU}}(i_{mc}), j = 1, ..., N_{PA_{ASSIGNED}}(i_{mc})$$

$$d_{TOT}\left(N_{CLU}(i_{mc}), i_{mc}\right) = \sum_{i(i_{mc}), j(i_{mc})} d\left[P_{CLU}(i(i_{mc}), i_{mc}), PA(j(i_{mc}), i_{mc})\right] \qquad (4)$$

$$\bar{d}_{avg}\left(N_{CLU}(i_{mc}), i_{mc}\right) = mean\left[d\left[P_{CLU}(i(i_{mc}), i_{mc}), PA(j(i_{mc}), i_{mc})\right]\right]$$

$$i_{mc} = 1, ..., N_{MC}$$

Here, $d_{TOT}\left(N_{CLU}(i_{mc}), i_{mc}\right)$ represents the total distance of all mutual distances between $P_{CLU}(i, i_{mc})$ points and corresponding $PA(j, i_{mc})$ points meaning the overall distance, which should be walked by assigned pupils to the nearest bus stops identified inside each cluster $C_k(i_{mc})$. Conversely, $\bar{d}_{avg}\left(N_{CLU}(i_{mc}), i_{mc}\right)$ denotes the average distance that every individual pupil must walk to the nearest bus stop inside each cluster $C_k(i_{mc})$.
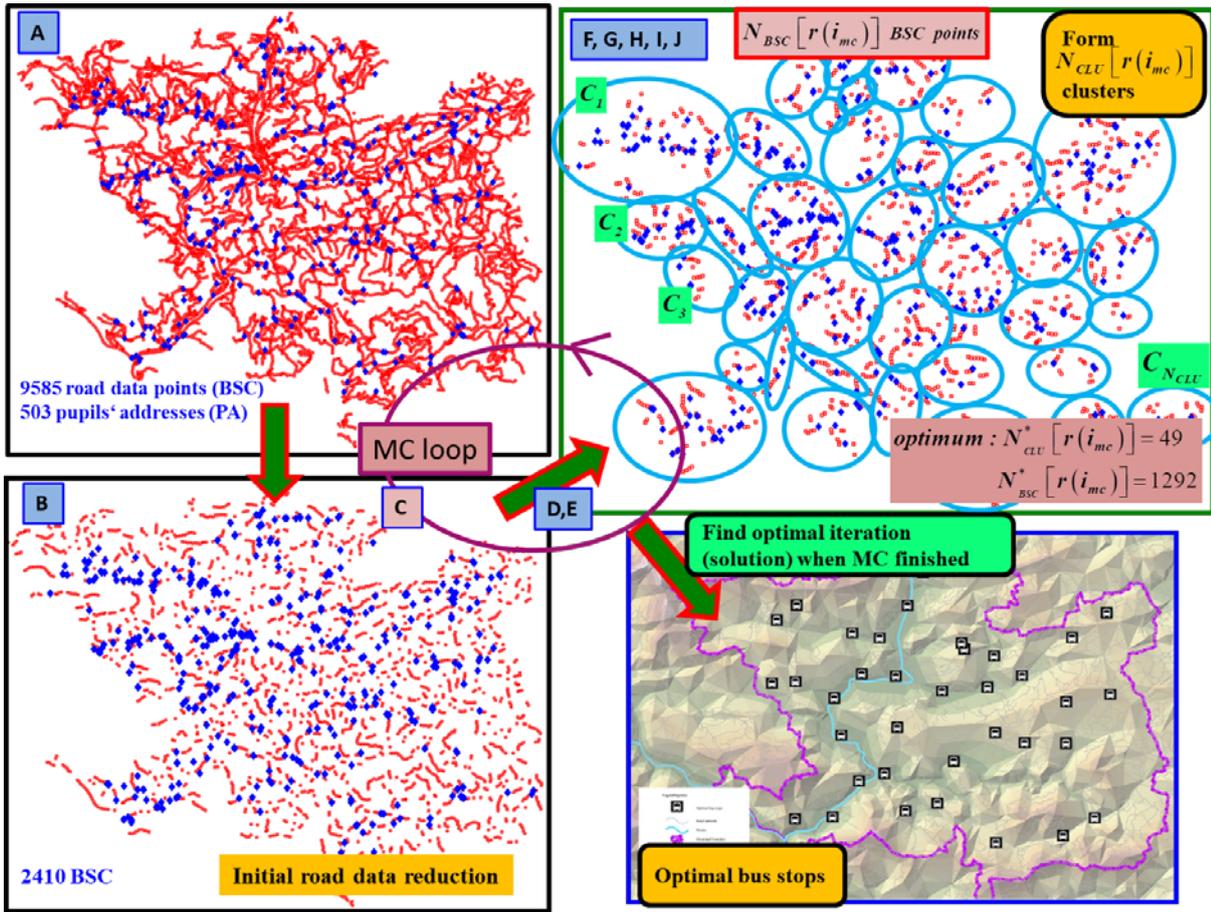


Figure 6. The illustration of BSC and PA points' data processing while executing the heuristic algorithm in figure 5: A – Given data; B – Initial road data reduction; F-J: Clustering of the remaining

$N_{BSC}\left[r\left(i_{mc}\right)\right]$ BSC points, assigning pupils to the nearest $P_{CLU}\left(i,i_{mc}\right)$ BSC points, and calculating all distances; Down-right sub-figure: Positions of optimal bus stops (OBS).

Now, our goal is to apply an optimization procedure to find such optimal MC iteration $i_{mc}^{*}$ where the sum of distances $d_{TOT}\left(N_{CLU}\left(i_{mc}^{*}\right),i_{mc}^{*}\right)$ of all clusters $C_{k}\left(i_{mc}^{*}\right)$ (total walking distance (TWD) of all assigned pupils) is minimal possible. Consequently, all the assigned pupils would have to walk the shortest possible joint total distance to their closest optimal bus stop. An additional criterion is that the average distance (AD) each individual pupil should have to walk to the nearest bus stop is minimal possible regarding the average distances $\overline{d}_{avg}\left(N_{CLU}\left(i_{mc}^{*}\right),i_{mc}^{*}\right)$ of all clusters $C_{k}\left(i_{mc}^{*}\right)$. Naturally, both criteria are trying to be reached with an additional demand that the minimal possible number of optimal bus stops (OBS), i.e., $OBS\left(i_{mc}^{*}\right)\in U_{union}\left(P_{CLU}\left(i\left(i_{mc}^{*}\right),i_{mc}^{*}\right)\right)\left\{over\ all\ clusters\ C_{k}\left(i_{mc}^{*}\right)\right\}$ collected over all clusters $C_{k}\left(i_{mc}^{*}\right)$ should be kept at the lowest possible level. The algorithm is also designed in such a way that avoids those iterations where the results would likely lead to an unacceptable number of unassigned pupils.

If the optimal bus stops are calculated as described above, they will cover as many pupils as possible, while the latter should have to walk (on average) as little as possible. Also, the TWD distance of all pupils to their nearest bus stops would be minimal possible. The distribution of obtained optimal results, i.e. OBS points is shown in the down-right sub-figure of figure 6 and will be more transparently illustrated in the sequel.

## IV. NUMERICAL RESULTS

The algorithm was developed in the MATLAB. Table 3 shows all relevant parameters used and obtained results for the case of optimal MC iteration $i_{mc}^{*}$. When the entire heuristic procedure was finished, the (X,Y,Z) coordinates of 37 OBS points were calculated (see details in table 3). The OBS were capable of covering of 503-27=476 pupils within the previously defined radium constrained by $d_{max}=2.5km$ (see section II.B), while 27 pupils should have to walk only several hundred meters more. Since the total number of observed pupils is 503, the algorithm apparently managed to achieve service to 94.6 % of all pupils by determination of calculated OBS points.

Table 3. The relevant parameters used and obtained results for the case of optimal MC iteration $i_{mc}^{*}$.

| | |
|---|---|
| Number of MC iterations to achieve appropriate results | 70000 |
| Optimal iteration $i_{mc}^{*}$ | 27085 |
| $N_{BSC}^{*}\left[r\left(i_{mc}^{*}\right)\right]$ points after removing too distant BSC points (see block E, figure 5) | 1292 |
| $N_{CLU}^{*}\left[r\left(i_{mc}^{*}\right)\right]$ clusters | 49 |
| Total walking distance (**TWD**) of all assigned pupils at $i_{mc}^{*}$ | 1054 km |

| Average walking distance (**AWD**), each individual pupil should have to walk to the nearest BS at $i_{mc}^{*}$ | 1962 m |
|---|---|
| Number of optimal bus stops $OBS\left(i_{mc}^{*}\right)$ | 37 |

The calculated results for 37 OBS positions can be represented as shown in figure 7. Their spatial distribution is relatively uniformly distributed, with the exception of most remote municipality's regions with a low density of population. Naturally, the presented OBS are based on mathematical principles only, so they must be appropriately calibrated to align hypothetical positions with the actual terrain characteristics.
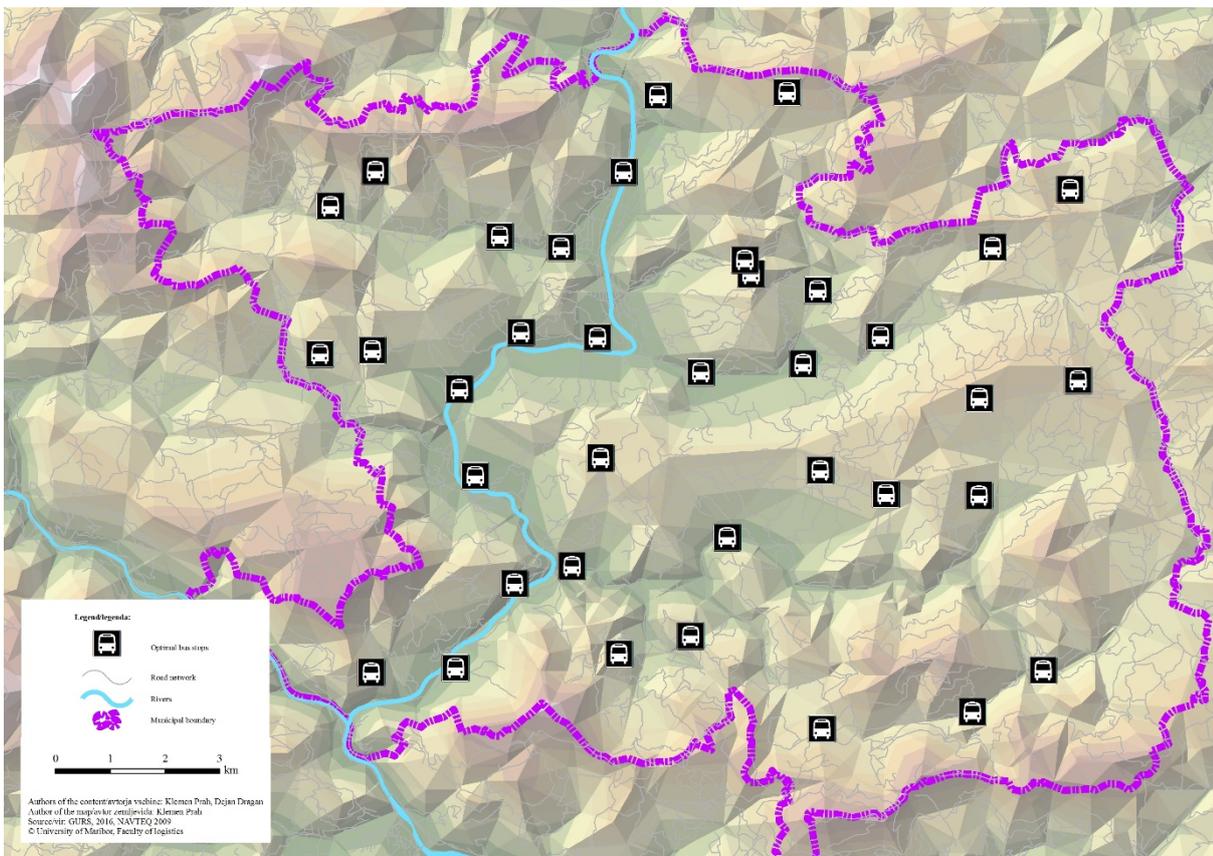


Figure 7. The spatial distribution of calculated 37 optimal bus stops positions in the Municipality of Laško.

## V. DISCUSSION

Since we managed to find an optimal solution with only 37 optimal bus stops covering 94.6 % of all treated pupils, we believe that these are quite encouraging results regarding the diversity and wideness of municipality (189 km2), as well as difficult terrain characteristics and complex terrain's elevation dynamics. Our trust in promising results is fortified also for the reason that an average distance each individual pupil should have to walk to the nearest bus stops is acceptably long (1962 m), while the total walking distance of all assigned pupils is 1054 km.

Our combined clustering-MC heuristic algorithm based on 3D GIS data has proved effective in quite difficult circumstances with a problematic dynamic rural terrain and dispersed pupils' addresses. For this reason, our approach can apparently solve even such complicated optimization problems, as the complex type of BSA problem is. This might be an important contribution of this

paper. Also, we offer research community an exciting methodology alternative to some other typical methodological approaches that are most commonly used for solving of hard optimization problems, such as MLCP problems, SBRP problems, and other location and/or vehicle routing problems. This belief might be particularly accepted for the reason that a GIS framework is also a common thread that is shared by all of such kinds of problems.

After some additional minor modifications, the optimal solution (coordinates of derived optimal bus stops) are planned to be submitted to the Laško municipality responsible personnel. By doing this, they are going to get an opportunity to trigger all necessary procedures for the purpose of actual bus stops' physical implementation.

When this implementation will be finished and correctly adapted to the actual terrain characteristics, the drivers of transportation vehicles will have to pick up the treated pupils only at the bus stops, instead of picking them at their homes individually. Since the number of all treated pupils is 503, while the number of planned bus stops is only 37, obviously the savings of transportation costs are going to be considerably significant.

## VI. CONCLUSION

The problem of cost optimization related to the mandatory pupils' transportation to their schools should be at the top of the priority list of every municipality in Slovenia. Namely, the level of such costs can escalate over any reasonable limit, which is intolerable for any municipality, particularly the small ones with a low budget. The paper addressed the optimal bus stops' allocation within the framework of a pupils' transportation cost reduction process in the Municipality of Laško.

The heuristic optimization algorithm based on data clustering and Monte Carlo simulation has been introduced for the case of three-dimensional GIS data. The results have shown that the algorithm is capable of allocating a relatively small number of optimal bus stops that can still assure a maximal service area to the majority of pupils. Furthermore, the optimal bus stops' allocation ensured that an average distance each individual pupil should have to walk to the nearest bus stop is acceptably long, while the total walking distance of all assigned pupils is within the reasonable level.

## REFERENCES

[1]     J. Park and B.-i. Kim, "The school bus routing problem : A review," *European Journal of Operational Research*, vol. 202, no. 2, pp. 311-319, 2010.

[2]     J. Desrosiers, J. Ferland, J. Rousseau, and G. Lapalme, "TRANSCOL: a multi-period school bus routing and scheduling system," *TIMS Studies in the Management Sciences*, vol. 22, 1986.

[3]     K. Worwa, "Minimization of Number of Buses in the School Bus Routing Problem," *Research in logistics and producton*, vol. 7, no. 2, pp. 127-141, 2017.

[4]     R. Farahani and M. Hekmatfar, *Facility Location: Concepts, Models, Algorithms and Case Studies.* Springer, 2009.

[5]     R. L. Francis, F. M. Jr, and J. A. White, *Facility Layout and Location: An Analytical Approach*, 2 edition ed. Englewood Cliffs, N.J: Pearson, 1991, p. 592.

[6]     R. Church and C. R. ReVelle, "The Maximal Covering Location Problem,", *Papers in Regional Science*, vol. 32, no. 1, pp. 101-118,1974.

[7]     D. Dragan, T. Kramberger, and M. Lipičnik, "Monte Carlo Simulation-based Approach to Optimal Bus Stops Allocation in the Municipality of Laško," *PROMET - Traffic&Transportation*, vol. 23, no. 4, pp. 265-278, 2012.

[8]     D. Dragan, T. Kramberger, A. Lisec, M. Intihar, and K. Prah, "Using GIS for the Optimization of Pupils Transportation: The Case of Laško Municipality," *Logistics & sustainable transport*, vol. 2, pp. 35-51, 2011.

[9]     T. Kramberger, D. Dragan, and K. Prah, "A heuristic approach to reduce carbon dioxide emissions," *Proceedings of the Institution of Civil Engineers - Transport*, vol. 167, no. 5, pp. 296-305, 2014.

[10]    D. Dragan, T. Kramberger, and K. Prah, "Transport Optimization and Estimation of Reduced CO2 Emissions," in *Sustainable Logistics and Strategic Transportation Planning*, T. Kramberger, Ed.Tomaž Kramberger, Vesna Ipavec: IGI Glogal, 2016, pp. 412-443.

[11]    J. E. Harmon and S. J. Anderson, *The Design and Implementation of Geographic Information Systems*, 1 edition ed. Hoboken, N.J: Wiley, 2003, p. 272.

[12]    V. Marianov and D. Serra, *Location Problems in the Public Sector. Facility Location: Applications and Theory*. Berlin, 2002.

[13]     M. S. Daskin, *Network and Discrete Location: Models, Algorithms, and Applications*, 2 edition ed. Hoboken, New Jersey: Wiley, 2013, p. 536.

[14]     X. Li and A. G.-O. Yeh, "Integration of genetic algorithms and GIS for optimal location search," *International Journal of Geographical Information Science*, vol. 19, no. 5, pp. 581-601, 2005.

[15]     J. H. Jaramillo, J. Bhadury, and R. Batta, "On the Use of Genetic Algorithms to Solve Location Problems," *Computers & Operations Research*, vol. 29, pp. 761-779, 2002.

[16]     A. T. Murray and R. L. Church, "Applying simulated annealing to location-planning models," (in en), *Journal of Heuristics*, vol. 2, no. 1, pp. 31-53, 1996.

[17]     J. Aerts and G. Heuvelink, "Using Simulated Annealing for Resource Allocation," *International Journal of Geographical Information Science*, vol. 16, pp. 571-587, 2002.

[18]     D. M. Jaeggi, G. T. Parks, T. Kipouros, and P. J. Clarkson, "The development of a multi-objective Tabu Search algorithm for continuous optimisation problems," *European Journal of Operational Research*, vol. 185, no. 3, pp. 1192-1212, 2008.

[19]     M. Sun, "Solving the uncapacitated facility location problem using tabu search," *Computers & Operations Research*, vol. 33, no. 9, pp. 2563-2589, 2006.

[20]     R. D. Galvao and C. Revelle, "A Lagrangean heuristic for the maximal covering location problem," *European Journal of Operational Research*, vol. 88, no 1, pp. 114-123, 1996.

[21]     V. Marianov and D. Serra, "New Trends in Public Facility Location Modeling," in *UPF Economics and Business Working Paper 755*, 2004.

[22]     J. M. Gleason, "A set covering approach to bus stop location," *Omega*, vol. 3, no. 5, pp. 605-608, 1975.

[23]     E. Carrizosa, J. Harbering, and A. Schöbel, "The Stop Location Problem with Realistic Traveling Time," in *ATMOS - 13th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems - 2013*, Schloss Dagstuhl—Leibniz-Zentrum fuer Informatik, vol. 33, pp. 80-93, 2013.

[24]     A. Schöbel, *Optimization in Public Transportation: Stop Location, Delay Management and Tariff Zone Design in a Public Transportation Network*. Springer, 2007, p. 267.

[25]     R. L. Church, "Geographical information systems and location science," *Computers & Operations Research*, vol. 29, pp. 541-562, 2002.

[26]     L. Y. O. Li and Z. Fu, "The school bus routing problem: a case study," *Journal of the Operational Research Society*, vol. 53, no. 5, pp. 552-558, 2002.

[27]     Statistični Urad. (2017, March 24). *Občina Laško*. Available: http://www.stat.si/obcine/sl/2014/Municip/Index/78

[28]     P. Horvat, *Geografija občine Laško*. Filozofska fakulteta Univerze v Ljubljani, 2006.

[29]     P. Kovač, "The Geography of the Municipality of Laško," University of Ljubljana, 2006.

[30]     D. Perko and A. M. Orožen, *Slovenija: pokrajine in ljudje*. Ljubljana, 2001.

[31]     P. Očkerl, *Priročnik za izvajanje in spremljanje poti v šolo*, Ljubljana, 2017

[32]     M. Ogrin, T. R. Planinc, M. Ilc, and A. Plevnik, Trajnostna mobilnost: priročnik za učitelje v osnovnih šolah, Ljubljana: Ministrstvo za infrastrukturo in prostor, 2013.

[33]     GURS: *Geodetska uprava Republike Slovenije: Digitalni ortofoto 050*, 2017.

[34]     R. S. King, *Cluster Analysis and Data Mining: An Introduction*, Har/Cdr edition ed. Dulles, Virginia; Boston, Massachusetts; New Delhi: Mercury Learning & Information, 2014, p. 300.

[35]     C. C. Aggarwal and C. K. Reddy, *Data Clustering: Algorithms and Applications*, 1 edition. Boca Raton: Chapman and Hall/CRC, 2013, p. 652.

[36]     B. Everitt, *Cluster analysis*. Chichester, West Sussex: Wiley, 2011.

AUTHORS

A.   **Klemen Prah,** Phd, is the Assistant Professor at the Faculty of Logistics, University of Maribor, Celje, Slovenia (e-mail: klemen.prah@um.si).

B.   **Abolfazl Keshavarzsaleh,** Msc, Msc, is the Teaching and Research Assistant at the University of Malaya (UM) and International University of Malaya-Wales (IUMW), Kuala Lampur, Malesya (e-mail: abolfazl.keshavarz.saleh@gmail.com).

C.   **Tomaž Kramberger,** Phd, is the Associate Professor at the Faculty of Logistics, University of Maribor, Celje, Slovenia (e-mail: tomaz.kramberger@um.si).

D.   **Borut Jereb,** Phd, is the Associate Professor at the Faculty of Logistics, University of Maribor, Celje, Slovenia (e-mail: borut.jereb@um.si).

E.   **Dejan Dragan,** Phd, is the Associate Professor at the Faculty of Logistics, University of Maribor, Celje, Slovenia (e-mail: dejan.dragan@um.si).