

A modified grouping genetic algorithm to select ambulance site locations

¹*Alexis J. Comber, ²Satoshi Sasaki, ²Hiroshi Suzuki and ¹Chris Brunsdon

¹Department of Geography, University of Leicester, Leicester, LE1 7RH, UK,
Tel +44 (0)116 252 3812, Fax +44 (0)116 252 3854,

²Department of Infectious Disease Control and International Medicine, Graduate
School of Medical and Dental Sciences, Niigata University
Tel +81 25 227 2129, Fax +81 25 227 0765

* Contact author: ajc36@le.ac.uk

Abstract

This paper describes the development and application of a modified Grouping Genetic Algorithm (GGA) used to identify sets of optimal ambulance locations. The GGA was modified to consider a special case with only two groups, and the reproduction and mutation schemes were modified to operate more efficiently. It was applied to a case study locating ambulances from a fixed set of alternative locations. The sites were evaluated using data of emergency medical services (EMS) calls summarised over census areas and weighted by network distance. Census areas serviced by the same selected location defined ambulance catchments. The results indicated alternative sites for ambulances to be located, with average EMS response times improved by 1 minute 14 seconds and showed the impacts of having different numbers of ambulances in current locations and in new locations. The algorithmic developments associated with the modified GGA and the advantages of using census areas as spatial units to summarise data are discussed.

Key words: genetic algorithm, ambulance, EMS

1. Introduction

This work describes the development of a modified Grouping Genetic Algorithm (GGA) which was applied to an ambulance location problem. Site location is an important aspect of spatial planning, especially in public health domains, which have to balance current needs, expectations of future requirements and competing demands for resources. Planners require evidence to inform 'location allocation' decisions

which have traditionally been supported by GIS analyses of accessibility to facilities. As techniques have become more sophisticated (for example the evolution from deterministic analyses to ones using heuristic search algorithms), the dimensionality and scope of analyses have increased. Deterministic or brute force methods that evaluate every possible combination of solutions may be unsuitable for locational problems with thousands of demand points and a large number of subsets to be evaluated (Church, 1999). For example, in one part of the case study presented below there are over 10^{3347} sets of possible locations to evaluate. In such cases deterministic solutions may be too computationally demanding (Li and Yeh, 2005). In this context heuristic search algorithms provide useful modelling tools. Genetic Algorithms (GAs) have been suggested by a number of workers as offering a suitable search heuristic to answer site location questions (e.g. Jin and Wang 2001; Zhan et al. 2003; Ducheyne et al. 2004). Set covering models have been developed to identify ambulance site locations, including the p -median problem of locating facilities (Hodgson, 1988) and the maximal covering location problem (Church and ReVelle, 1974). P -median models seek to minimise the sum of distance (or average distance) between demand nodes and facilities (Rosing et al. 1979). They have been used to minimise distances between supply and demand points in 'location / allocation' problems (e.g. Alp et al., 2003; Snyder and Daskin, 2005). Maximal covering approaches are concerned with maximizing the coverage of a given set of facilities (e.g. Batta et al., 1989; Chiyoshi et al., 2003). Deterministic and probabilistic models of ambulance location are reviewed in Schilling et al. (1993) and Brotocorne et al. (2003) respectively.

The development and application of a modified GGA described in this paper extends other work using GAs to address the location allocation / p -median problem in a number of ways. First, the GGA is modified to consider a special case with only two groups (a discrete set of candidate locations are designated as 'used' or 'not used'), which are fixed in size. Second, it bridges the gap between theoretical models and real-world problems. Much other work in operations research literature has addressed small scale studies, based on simplified and / or simulated cases to illustrate model development. This work addresses a real logistics problem based on real locations within a regional case study including its entire road network and uses actual emergency medical services (EMS) data rather than simulated EMS demand. Third, the model allows for different numbers of facilities to be located within the solution

and their benefits evaluated. As ambulance response time to EMS cases is a critical factor in patient survival (Snyder et al., 2007; Stiell et al., 1999), the approach provides nuanced evidence to those responsible for developing spatial policy. Section 2 describes the scientific context of this work. In Section 3 details the modified GGA and its algorithmic components, and discusses the advantages of the GGA approach. The analytical methods are described in Section 4 before the results are presented (Section 5), discussed (Section 6), and some conclusions are drawn (Section 7).

2. Background

The problem addressed in this work is how to optimally allocate resources (in this case ambulance locations) based on need. In operations research other workers have considered ambulance location models using p -median and maximal covering approaches. Early work developed and applied deterministic approaches to identify ambulance locations (eg Berlin and Liebman, 1974; Daskin, 1982; Uyeno and Seeberg, 1984). Genetic Algorithms are suitable when the number of potential solutions is too large to be evaluated deterministically. Genetic algorithms are heuristic search algorithms (Goldberg, 1989), introduced by Holland (1975), that allow the survival of the fittest (Huang et al., 2004). Models that incorporate GAs are frequently referred to as ‘optimisation models’ in the literature. They simulate the process of genetic mutation, gene combination, and selection in biological evolution. The aim is to breed fitter generations, where fitness is evaluated over an objective function. Experiments in which the optimal solution is known suggest that it is likely that the optimal solution exists within the final generation of the algorithm if sufficient generations have been simulated, despite not being able to state its optimality with certainty. A set of simple scenarios illustrating this can be found in the *genalg* R package (Willighagen, 2005).

Genetic algorithms have been extensively applied to p -median, maximum covering and other spatial problems in combination with GIS-based analyses. A non-exhaustive sample of the literature is reviewed here. Alp et al. (2003) proposed a p -median location/allocation model using genetic algorithms to locate facilities in response to hypothetical demand points and to determine ‘catchments’ by allocating those points to the facilities. Li and Yeh (2005) compared GAs with simulated annealing and GIS

neighbourhood search methods in order to determine optimal locations for hospital locations based on population density and proximity to roads in Hong Kong. Vink and Schot (2000) used a GA to optimise the location of drinking water production based on environmental constraints and the transport network. Samanta and Jha (2008) used a GA linked to GIS to identify the location of a rail network and its stations. They combined network, site parameter and demographic data in a GIS to identify suitable sites evaluated on cost. Stewart et al. (2004) developed a land use planning support system using a GA with GIS. Their objective was to accommodate recent developments in land use planning including the involvement of stakeholders and a highly complex set of constraints to the decision problem in order to facilitate an interactive decision support systems. As there was no absolute 'optimal solution' to their problem they used the results of the GA analysis as the focus of interaction in order to generate user feedback on the different solutions. Genetic algorithms in combination with GIS have also been used for route planning (Huang et al., 2006; Dadkar et al., 2008), forest resource planning (Ducheyne et al., 2006) and ecological niche modelling (Levine et al., 2009). Xiao et al. (2007) provide a thorough review of algorithmic developments in the context of multi-objective spatial decision making and Li et al (2009) provide a recent extension using ant colony optimisation techniques. Other approaches specific to the ambulance site location models have incorporated hypercube structures (Larson, 1974, 1975) with genetic algorithms. Hypercube structures consider vehicle state probabilities and other related performance measures such as workloads, travel times, etc. They have been used to model emergency service provision with developments of the maximal covering location problem (e.g. Batta et al., 1989; Chiyoshi et al., 2003; Takeda et al., 2007).

There are two main themes that emerge from the literature relating to the application of GAs to solve ambulance 'location allocation' problems. First, as Brown and Sumichrast (2003) note, classical GAs can have difficulties in subsetting solutions into groups in a grouping problem. This is elaborated upon in the next section describing algorithm development. Second, classical GAs consider only individuals (locations) rather than sets or 'groups' of solutions together. It is not difficult to envisage a situation where the 'best' single location may not be included in the set of N best locations when the whole set is evaluated rather than the individual. Additionally, much of the previous research relating specifically to ambulance location problems

have considered a small number of locations: Chuang (2007) used to a GA to identify only 4 facilities; Iannoni et al (2008) combined a GA within a hypercube to consider the optimal location of five ambulances on two connected stretches of highway.

Grouping genetic algorithms are a development of the GA that consider groups or subsets of individuals rather just individuals. GGAs were proposed by Falkenauer (1998) to solve complex partitioning problems where in this study a set of potential locations are split into optimal groups and the group attributes are considered as genes. Recent work using GGAs in GIS context has applied a GGA to identify a logistics network that optimised the profitability of rice producers in Thailand (Pitaksringkarn and Taylor, 2005) and to optimally identify sets of post offices to close in order to satisfy government accessibility criteria (Comber et al., 2009). The operation of GAs and GGAs are described in more detail in the next section.

3. Algorithm Development

A number of non-genetic algorithms have been developed to address the location allocation / p -median problem – see for example Teitz and Bart (1968), Christofides and Beasley (1982), or Reese (2006) for a comprehensive overview. However, in practice these methods proved prone to finding ‘local’ optima (where the algorithm failed to converge on the *global* optimum arrangement) or to require very long running times. In the practical situations considered in this paper there are a large number of potential choices for siting, and distances are network-based rather than Euclidean. As alternative approaches, we could consider simulated annealing (e.g. Righini, 1995) or genetic approaches – which are the subject of this paper. An overview of genetic approaches is now set out.

3.1 Introduction to GAs and GGAs

The optimisation process in GA proceeds as follows. A potential set of solutions is created, sometimes called a ‘string’ (Huang et al., 2004) or more usually a ‘chromosome’. Chromosomes form individuals and are composed of genes, which in a multi-site optimisation will be sets of individual sites. The chromosome is then evaluated based on some ‘fitness’ criteria which assess its performance. Genes from

successful individuals (i.e. passing the criteria) are interchanged between individuals. This is done in a crossover which combines two chromosomes to create new chromosomes, which in turn form new individuals, in effect offspring or children. This process of creating new chromosomes from successful (i.e. highly ranked in the evaluation) ones continues usually for a predetermined number of cycles or until some convergence criteria are met. However it is possible for the optimisation to stagnate and to ensure that this does not happen mutation is an important part of the GA. Mutation occasionally adds some random new genes into the chromosome. In this way the GA 'breeds' optimal solutions by creating a more optimised (fitter) generation and hence the analogy with natural selection: crossover creates new chromosomes from successful ones, and mutation ensures diversity. So just as in natural selection, GAs ensure fitter more suitable individuals are parents to the next generation by allowing the individuals that to adapt to the fitness criteria to have offspring.

A standard genetic algorithm represents each potential solution as a fixed length 'chromosome' – in the original form this was a vector of binary elements – each of which could be thought of as a 'gene'. More generally these binary genes are replaced by integers or real numbers. However, the standard GA approach does not correspond well to the p -median problem in a grouping context. By grouping here, we mean that the chromosomes here are used to associate each individual with a group – and there are a smaller number of groups than individuals. In this case there are just two groups – the 'used' and 'not used' categories. Falkenauer (1998) proposed the *Grouping Genetic Algorithm* (GGA) and identified three major drawbacks of applying a classical GA to a grouping problem: the standard GA encoding scheme is highly redundant in relation to groups; the recombination operator may produce new population members that do not contain any of the qualities of its parents; and, other standard operators may cause too much disruption to successful population members. GGA modifies the classical genetic algorithm by developing heuristic procedures to restore group membership during the crossover operation to any members that may have been displaced by other operations. The similarities and differences between GA and GGAs are discussed in Brown and Sumichrast (2003) and in Pitaksringkarn and Taylor (2005).

The approach presented here is in the spirit of the work of Hosage and Goodchild (1986) in that it attempts to use a GGA to select an optimal subset of locations from a discrete set of points, where the subset has a fixed size. Standard GA approaches generate strings consisting of 0's and 1's which may be used to divide the locations into 'used' and 'not used' categories. However the strings generated in this way could generate a 'used' category of any size from zero (where all genes take the value 0) to the number of locations (where all genes take the value 1). Here, we only wish to consider strings with a fixed number of 1's - corresponding to the subset of locations being used having a fixed size. There is a need for strategies to avoid generating strings with either more or fewer 1's than the desired string size. One approach is to apply a penalty function to the fitness of any string that does not have the desired number of 1's. The aim here is to 'breed out' strings with the wrong number of 1's so that populations ultimately contain only strings with the desired number of 1's. However, rather than use a penalty function approach, our approach explores a search space consisting of all of the strings with the desired size number of 1's – the breeding operations guarantee that, given 'input' chromosomes with a certain number of 1's, offspring chromosomes will also have this number of 1's. Thus, rather than exploring the space of *all* binary strings, only strings with a given number of 1's are considered - and no solution in the space is infeasible. This 'shrinkage' in the search space – from strings of all combinations to those with a fixed number of 1's - leads to a considerably more efficient search than applying a penalty function approach to the full space. For example to choose 11 items out of 50, considering all possible subsets out of 50 regardless of size leads to a search space with around 10^{15} possibilities, but reducing the search space just to subsets of size 10 leaves around 10^{10} possibilities, reducing the size of the space by a factor of around 10,000.

In the Falkenauer (1998) grouping genetic algorithm this standard chromosome representation is replaced by one in which *groupings* of genes are represented, rather than values attached to each gene. Rules for combining chromosomes and mutating them have to be modified to allow for this. We also employ ideas from Alp *et al* (2003).

3.2 Modified GGA

In the original GGA paper by Falkenauer (1998), a generic approach is outlined, and it is advocated that rules for regenerating chromosomes in specific group tasks should be devised on a bespoke basis. Brown and Sumichrast (2003) found that GGA performs worse when problem-specific knowledge is not incorporated into the replacement heuristic. In the algorithm here we opt for an approach that uniquely exploits the fact that only two groups are required (points at which ambulances are sited and ones at which they are not) – and in fact this means that a potential solution can be represented by listing only one of the two groups – the other being defined implicitly as the complement of this group. In particular, we choose the smaller of the two potential groups to be the one stored as a chromosome. For example if we had to choose 10 ambulance locations out of a potential 50, each chromosome would be a set of size 10, listing the 10 points designated as an ambulance location. Alternatively, if we were choosing 40 stations, it is more efficient to consider sets of locations *not* designated as locations.

The ‘genalg’ package contributed to the R statistical programming language project (R Development Core Team, 2008) by Egon Willighagen was modified to implement the algorithm. This library provides a set of genetic algorithm tools, which were augmented with an algorithm to select a ‘best subset’ based on the optimisation of an objective function, in this case weighted person distances. A description of the ‘genalg’ package can be found at the R web site¹. Full details of this algorithm now follow.

3.3 Chromosome representation

In this algorithm, as stated above, an individual chromosome is represented as a set of k unique integers, where each integer is in the range $1 \dots N$ and where k is the number of ambulances and N is the number of candidate locations. If each generation has M chromosomes, then these can be stored in an $M \times k$ integer array.

3.4 Evaluation and fitness function

¹ <http://cran.r-project.org/web/packages/genalg/genalg.pdf>

The objective function analysed was of the form:

$$\frac{\sum_{i=C} p_i d_{ij}}{\sum_{i=C} p_i} \quad (\text{Equation 1})$$

where i indexes the census areas C , p_i is the number of emergency cases in the census area i , and d_{ij} is the distance between census area centroid i and possible ambulance location j , with j being the index of the nearest location to the census area i under a choice of k possible ambulance locations. The function computes the population weighted average distance to the nearest potential location, and in this case the function analysed the counts of actual emergency cases in each area. However it should be noted that p_i can be varied to represent different populations. For example it could be the count of over 65's, the total population or the 'at risk' population defined in some other way. Thus, the function measures access for the different subsets and the GGA optimisations seeks to minimise the value of the objective function. A similar objective function – although using Euclidean or Minkowski distances rather than network analysis, is considered in Murray and Estivill-Castro (1998).

3.5 Selection of chromosomes, crossover and mutation

Breeding of a new generation is achieved by selecting pairs of chromosomes based on their fitness, combining these to produce new offspring and possibly applying mutation. The initial selection is based on the 'roulette wheel' approach (see for example Mitchell (1998)), in which fitness is defined according to Section 3.4. When two chromosomes are selected in this way, the combination algorithm is as below.

Denote the two chromosomes as the sets S_1 and S_2 . First, find the set $S_1 \cup S_2$, that is the list of all numbers appearing in both sets. Then select a new offspring set by randomly sampling k integers from $S_1 \cup S_2$, without replacement. Repeating this operation M times will yield a new generation.

Note that if S_1 and S_2 are identical, then their offspring will also be identical, and that if they have a large number of elements in common, there is a strong chance that these will appear in the offspring. Also, a form of 'elitism' is used, where the top 10% of

the current population is automatically carried forward to the next. Finally, there is a small probability of mutation – where an element in a given subset of k ambulances is swapped for one of the $M-k$ ambulances not in the subset.

Thus, the modified algorithm has all of the elements of a standard genetic algorithm, but addresses the issue of optimal subset selection, rather than optimisation of a list of real numbers or integers.

4. Methods

4.1 Case study

The case study is to select ambulance site locations from a set of fixed locations. Niigata is a prefectural city of 804,000 people in north-western Japan. Currently there are 27 ambulance stations providing EMS. In Japan there is a one-tiered EMS system provided by the fire department. Ambulances are located in some fire stations (one in each) and in Niigata there are 35 fire stations. When an emergency call is received, the nearest available ambulance is sent to the incident. Tanigawaa and Tanaka (2006) and Lewin et al. (2005) provide overviews of EMS provision, organisation and trends in Japan. The local authorities would like to identify ambulance locations and the number of ambulances that best match EMS needs under two scenarios: current locations and possible new locations as planning is compounded by a number of factors. Predicted future population changes indicate that Japan will have an increasingly older population and the majority of EMS cases relate to older people. For example in 2002 more than 50% of emergency calls in Tokyo were for patients over the age of 50 (Tokyo Fire Department, 2002). Also, the volume of emergency calls is increasing resulting in reduced average response times. In Niigata these trends are acute. The number of ambulance emergencies cases has increased by ~50% over 10 years from 55,434 cases in 1997 to 84,730 cases in 2007 and average ambulance response times have also increased by 1.2 minutes between 1998 and 2007 (Niigata Prefecture, 2008).

4.2 Data and Preparation

The Niigata City Fire Bureau provided prehospital medical emergency records of the 21,788 cases that were attended by an ambulance in the period April to December in 2007, 21,211 of which were geocoded and selected for analysis. In Niigata there are 2076 census small areas (the smallest spatial units over which census data are reported). The road network for Niigata was provided by Increment P Corporation. Niigata currently has 35 potential locations for ambulances, which are shown with the 21,211 EMS cases and the 27 existing fire stations with ambulances in Figure 1.

Insert Figure 1 about here

A point in polygon operation was used to generate a count of EMS cases for each census area and attached as an attribute to each point. A matrix of distances between each census point and each of the 35 fire stations was created from a GIS network analysis of shortest travel time. The census areas centroids were created using the area (x, y) envelope. The use of centroids as the location for output areas is commonly used in analyses that seek to relate polygon-based objects to linear networks, and unlike in the UK for example population weighted centroid locations are not available. For some sub-areas of the census polygon, the actual distance will be over-estimated and for others it will under-estimated. In this analysis we have assumed that losses and gains balance each other out. A second distance matrix was created of distances between 1255 grid points and the census area points. This was to explore the potential of new site and to illustrate how the method could be used to explore the selection of alternative future locations upon which ambulance stations could be built. The grid points were spaced at 500m and selected to be within 50m of an existing road and shown in Figure 2. The distance matrices and the census area counts of EMS cases were used as input into the modified GGA.

Insert Figure 2 about here

4.3 Analyses

Three analyses were run on the data to evaluate locations under a number of scenarios. First, to determine whether the current location of 27 ambulances could be improved in terms of reduced EMS response times, through the selection of

alternative sets of ambulance site locations. Second, to quantify the impacts of different numbers of ambulances in terms of losses and gains in average EMS response times. Third, to select potential ambulance locations from a discrete set of new location points and to evaluate their impacts on response times. In each case up to 5 ambulances could be allocated to any one location and the GGA was run for 1500 iterations.

5. Results

5.1 Improved locations

The modified GGA was run to select 27 ambulance site locations evaluated on the network distance between each census centroids, weighted by the count of EMS cases. For each census area, the nearest site location was also calculated to indicate the catchment or service area for each ambulance location. These results are shown in Figure 3. Of the 27 locations, 23 were existing stations but 4 were not, indicating that some improvement in ambulance accessibility could be achieved with some re-allocation of resources.

Insert Figure 3 about here

Policy makers and those who make decisions about resources need to be able to quantify the benefit associated with any site re-allocation. One simple measure is to calculate the difference in times or distances between each EMS case to its nearest new ambulance site and its nearest old one. The change in response time was calculated as follows. First, the network distances from each of the 21,211 EMS case data points to the nearest of the 27 GGA modelled site locations were calculated and compared with distances to the nearest current ambulance site locations. Average distances from EMS cases to the nearest ambulance station, before and after GGA site selection, are shown in Table 1. For three locations of the 23 that were already sites with ambulances, the average travel distance to EMS cases increases by an average of 319 metres but for other 20 the distances decrease on average by 630 metres. That 23 out of the current 27 locations were selected by the GGA is an interesting result: it indicates some degree of optimality in the choice of current ambulance locations

amongst the 35 fire stations. Possible reasons include the traditional high fire risk in Japan associated with buildings constructed from timber and therefore the historical need for fire stations to be well placed. Additionally, a high proportion of EMS cases relate to the elderly and in 1986 a revision of the Firefighting Organization Acts was made that specified emergency care for people with acute illness as a major purpose of the EMS, which may have resulted in some reorganisation of local resources (Tanigawaa and Tanaka, 2006).

Insert Table 1 about here

5.2 Improved response times

The original EMS case data indicated which ambulance attended the incident. The benefit of re-allocating ambulances was also quantified in terms of improved response time, a critical factor in patient survival. A network analysis was run on the EMS data on a case by case basis, identifying whether the nearest ambulance, second nearest, third nearest, etc, responded and the response time was calculated using average road speeds and distance from the current ambulance locations. The attendance by the n^{th} nearest ambulance was then simulated with the new set of 27 GGA modelled ambulance locations using a network analysis, a second response time calculated and compared with the first. The modelled average response times between the current and modelled ambulance locations decreased from 5.35 minutes to 4.12 minutes.

5.3 Different numbers of ambulances in current and new locations

The GGA was also applied to sets of 1 to 50 locations to determine the benefits of increased or reduced numbers of ambulance site locations. This was done for two scenarios: using the 35 existing locations and using the 1255 potential new locations.

For each set a number of outputs were produced by the GGA. These included:

- the locations of the set of selected ambulance sites, allowing for up to 5 at any location;
- the nearest location for each census area (thereby providing each potential location with its catchment); and

- the person distances between the census area centroids and the nearest site location (i.e. accounting for the number of EMS cases in each area).

The latter gives an indication of the decline in distance travelled with the addition of extra ambulances.

The average distance between census area and nearest site declines as more sites are included in the set, as expected. For the 35 current sites, the benefit of additional stations levels out at around 35 – i.e. as the solution becomes spatially saturated – despite allowing for additional ambulances at each station. In contrast, the average distances to the 1255 potential locations continue to decrease as might be expected (Figure 4). The site location and its catchment as modelled by the modified GGA can be mapped for each set. Figure 5 illustrates the different locations for sets of 10, 25 and 35 ambulances modelled from the 35 current locations and the 1255 potential new locations. We can see that the overall spatial patterns are similar, as the GGA sought to minimise person distances in each case. Mapping the selected ambulance site locations provides information to support ambulance location planning and allocation of future resources.

Insert Figure 4 about here

Insert Figure 5 about here

6. Discussion

In this work EMS cases were summarised over census areas and optimal ambulance location were modelled using a modified grouping genetic algorithm. The fitness of potential solutions were evaluated using on EMS-distances between census area centroid and potential site. Of the re-allocating ambulances, 23 out of 27 were at existing sites, 4 were allocated to new sites, the average distance increased for 3 sites (by 319m) and decreased for 20 sites (by 630m). This indicates that by altering the location of just 4 out of 27 ambulances could produce considerable benefits in terms of reduced net distances. The change in response times, averaged over each case by case and based on network times factored road speeds improved by 1.23 minutes

(from 5.35 minutes to 4.12 minutes), which although not precise – the times were based on road speeds – indicate the relative gains of reallocating ambulances. Other results from this study quantified the impacts of different numbers of facilities in terms of average ‘EMS case distances’ for both current potential sites (35) and for potential new sites (1255). Ambulance response times are critical for patient survival especially for severe emergency cases, where time to pre-hospital treatment is critical (Snyder et al., 2007; Stiell et al., 1999). Specifically, several studies have noted the improved survival of patients with a high risk of mortality when the response time was less than 4 to 5 minutes (Blackwell and Kaufman, 2002; Pons et al., 2005) and have sought to improve the response times through enhancing EMS management and relocating ambulance stations (Gossage et al., 2008; Peleg and Pliskin, 2004). Additionally, ambulance response time is an important indicator of pre-hospital emergency medical services (EMS) quality (Narad and Driesbock, 1999). Although statistical models for optimal location of ambulances have been widely discussed (Brotcorne et al., 2003), practical methods for the relocation and evaluation of their benefits have been limited. Quantifying the impacts of different numbers of re-allocated facilities in this way provides evidence to support policy decisions as public health policy planners are able to evaluate the impacts of increased or reduced numbers of ambulances.

In this analysis the standard GGA was modified to do a number of things. First, it was modified to operate on a discrete set of potential points rather than generic (x, y) points as in previous work (e.g. Li and Yeh, 2005). Working with a set of discrete points is often more realistic in terms of resource planning and site location, where there may already be a set of potential sites (for example, the existing fire stations) or where sites are located on a network (such as a roads). The computational advantages of GAs are well described in the literature (Huang et al., 2004; Li and Yeh, 2005): they are able to converge on solutions much quicker than would be possible by brute force. They are an effective way to address problems in which although the search space may be specified as a set of linear constraints (as in Murray and Estivill-Castro, 1998) the objective is non-linear, and with use of network distances, difficult to consider analytically. Also, using the modified GGA as specified here makes use of a tightly-defined search space which by definition excludes infeasible solutions, and through the simplicity of the ‘breeding’ routine is relatively computationally efficient.

This overcomes the convergence problem of genetic algorithms as described by Li et al (2009, p400): “GA has problems in reaching the convergence if the number of targets is large. For example, the optimal locations of more than 15 targets can hardly be found using GA because the required length of chromosome is too long.” This work has identified sets of locations that minimise overall person network distances. Other criteria can be used as described in Comber et al. (2009) and they can be combined as described in Li and Yeh (2005). Second, in the modified GGA a site could be selected as part of the optimal set multiple times to allow for any huge spatial skew in the criteria to be accommodated. Third, the GGA was used to analyse and identify sets of $n = 1$ to $n = 50$ sites, with multiples on each site. This provides decision makers with evidence of the impacts of having different numbers of locations and allows them to determine the optimal number of sites through evaluation of the additional costs and gains (in this case decreased response times) both for existing sites and possible future sites.

Enhancing traditional GA analyses in this way has a number of advantages:

- Optimising over network distances to census areas can provide a more realistic analysis of optimal access to different goods or services. It allows disparities between different spatial and social groups to be examined and provides potential solutions. This method is generic and could readily be applied to optimise retail outlets by considering market footprints, for example.
- The approach is relatively simple. It combines a GIS network analysis of distances between different sets of points with demographic data to create a weighting that can be adjusted depending on the objectives of the study or the phenomenon under investigation.
- The approach allows planners to evaluate the current situation in terms of the number and location of resources. The result of the analysis suggests how current site location and number can be improved and quantifies the additional benefits in terms of reduced weighted access distances.

The EMS cases were summarised over census areas. The network distances between each census centroid and each potential ambulance site location were weighted by the number of EMS cases as evaluation criteria in the GGA search and select process. Goldberg et al. (1990) showed that zone structure were crucial elements in travel time

model validation. In this work census areas provided a convenient and useful spatial framework with which to develop the analysis. First, they were used to describe ambulance catchment areas indicating the expected geographical coverage for each selected site. Second, they were used to analyse average response times and to provide information about the benefits of different numbers of ambulances. Third, the use of census areas allows other socio-economic data to be integrated into the analysis. This is critical for resource planning especially in the context of changing population demographics. The number of EMS cases has been strongly correlated with specific census variables (Clark and Fitzgerald, 1999; McConnel and Wilson, 1998). Census areas allow spatially distributed models of future EMS demand to be generated, for example by linking multivariable regression analysis of EMS case data against census variables with predicted population changes.

Additional work has applied the site selection method to predicted, future EMS cases distributed over census areas to explore changes in the spatial distribution of EMS demand as well as its volume (Sasaki et al., 2010). This work selects ambulance site locations based on projected future population changes, particularly with respect to Japan's aging society. In future work, the GGA will also be further extended to select locations over the 35 existing potential sites in combination with new ones allowing for different amounts of site replacement.

7. Conclusions

This paper presents a grouping genetic algorithm approach that overcomes the limitations of traditional genetic algorithm approaches in identifying groups or subsets of solutions (in this case ambulance locations). The GGA enables sets of solutions to be identified together and in a further development, an efficiency modification was applied to the GGA to shrink the search space. The benefits and efficiency of search space 'shrinkage' is evident when compared to 'penalty function' approaches. This approach was applied to a case study to optimise ambulance locations. Ambulance response time is crucial for patient survival and computationally efficient location- allocation methods such as presented here offer the opportunity of dynamic location identification based on current need and predicted need. The method is generic, using census areas to summarise EMS case

data, a network analysis of distances to locations and a GGA implementation. The evaluation criteria can be set according to local priorities (case distances, cost of new facilities, future EMS demand) to model the location of different numbers of facilities. In this way the approach supports spatial planning by providing a more nuanced picture of service need than previous approaches.

Acknowledgements

The authors would like to thank the Japanese Society for the Promotion of Science for the short-term fellowship awarded to Alexis Comber, the Graduate School of Medical Science at Niigata University for hosting Dr Comber and the Niigata City Fire Bureau and Department of Public Health for providing data on emergency cases.

References

- Alp, O., Erkut, E. & Drezner Z., (2003). An efficient genetic algorithm for the p -median problem, *Annals of Operations Research*, 122, 21–42
- Batta, R., Dolan, J.M. & Krishnamurthy, N.N. (1989). The maximal expected covering location problem: Revisited. *Transportation Science*, 23, 277–287.
- Berlin, G. & Liebman, J. (1974). Mathematical analysis of emergency ambulance location. *Socio-Economic Planning Sciences*, 8, 323-328.
- Blackwell, T.H. & Kaufman, J.S. (2002). Response time effectiveness: comparison of response time and survival in an urban emergency medical services system. *Academic Emergency Medicine* 9, 288-295.
- Brotcorne, L., Laporte, G. & Semet, F. (2003). Ambulance location and relocation models. *European Journal of Operational Research*, 147, 451–63.
- Brown, E.C & Sumichrast, R.T. (2003). Impact of the replacement heuristic in a grouping genetic algorithm. *Computers & Operations Research* 30: 1575–1593.
- Chiyoshi, F., Galvao, R.D. & Morabito, R. (2003). A note on solutions to the maximal expected covering location problem. *Computers and Operations Research*, 30, 87–96.
- Christofides, N. & Beasley, J.E. (1982) A Tree-Search Algorithm For The p -Median Problem *European Journal Of Operational Research*,10 (2)196-204

- Chuang, C.L. (2007). A maximum expected covering model for an ambulance location problem. *Journal of the Chinese Institute of Industrial Engineers*, 24, 468-474.
- Church, R. L. & ReVelle, C. (1974). The maximal covering location problem. *Papers of the Regional Science Association*, 32, 101–118.
- Church, R.L. (1999). Location modeling and GIS. In *Geographical information systems: Volume 1*, edited by P.A. Longley, M.F. Goodchild, D.J. Maguire and D.W. Rhind (New York: John Wiley & Sons, Inc.), pp. 293–303.
- Clark, M.J & Fitzgerald, G. (1999). Older people's use of ambulance services: a population based analysis. *Journal of Accident and Emergency Medicine*, 16, 108-111.
- Comber, A.J., Brunson, C., Hardy, J. & Radburn, R., (2009). Using a GIS-based network analysis and optimisation routines to evaluate service provision: a case study of the UK Post Office. *Applied Spatial Analysis and Policy*, 2, 47-64.
- Conley, J., Gahegan, M, & Macgill, J. (2005). A genetic approach to detecting clusters in point data sets, *Geographical Analysis*, 37, 286-314
- Dadkar Y., Jones D., & Nozick L. (2008). Identifying geographically diverse routes for the transportation of hazardous materials. *Transportation Research Part E-Logistics and Transportation Review*, 44, 333-349
- Daskin, M.S. (1982). Application of an expected covering model to emergency medical service system design. *Decision Sciences*, 3, 416-439.
- Ducheyne, E., De Wulf, R.R. & De Baets, B. (2004). Single versus multiple objective genetic algorithms for solving the even-flow forest management problem. *Forest Ecology and Management*, 201, 259-273.
- Ducheyne, E., De Wulf, R.R. & De Baets, B. (2006). A spatial approach to forest-management optimization: linking GIS and multiple objective genetic algorithms, *International Journal of Geographical Information Science*, 20(8): 917-928.
- Falkenauer, E. (1998). *Genetic Algorithms and Grouping Problems*. (London: John Wiley and Sons).
- Goldberg, D.E. (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning* (Reading, MA: Addison-Wesley).

- Goldberg, J., Dietrich, R., Chen, J.M. & Mitwasi, M.G. (1990). A simulation model for evaluating a set of emergency vehicle base locations: development, validation and usage. *Socio-Economic Planning Sciences*, 24, 124-141.
- Gossage, J.A., Frith, D.P., Carrell, T.W., Damiani, M., Terris, J. & Burnand, K.G. (2008). Mobile phones, in combination with a computer locator system, improve the response times of emergency medical services in central London. *Annals of the Royal College of Surgeons England*, 90, 113-116.
- Hodgson, M.J. (1988). An hierarchical location-allocation model for primary health care delivery in a developing area. *Social Science and Medicine*, 26: 153–161.
- Holland, J. H. (1975). *Adaptation in natural and artificial systems* (Michigan: University of Michigan Press).
- Hosage, C.M. & M.F. Goodchild. (1986). Discrete Space Location–Allocation Solutions from Genetic Algorithms. *Annals of Operations Research*, 6, 35–46.
- Huang, B, Yao, L & Raguraman, K. (2006). Bi-level GA and GIS for Multi-objective TSP Route Planning. *Transportation Planning and Technology*, 29, 105-124
- Huang, B., Cheu, R.L. & Liew, Y.S. (2004). GIS and genetic algorithms for HAZMAT route planning with security considerations. *International Journal of Geographical Information Science*, 18, 769-787.
- Iannoni, A.P, Morabito, R. & Saydam, C. (2008). A hypercube queueing model embedded into a genetic algorithm for ambulance deployment on highways. *Annals of Operational Research*, 157, 207–224
- Jin, Y.Q. & Wang, Y. (2001). A genetic algorithm to simultaneously retrieve land surface roughness and soil wetness. *International Journal of Remote Sensing*, 22, 3093-3099.
- Larson, R.C. (1974). A hypercube queueing model for facility location and redistricting in urban emergency services. *Computers and Operations Research*, 1, 67–95.
- Larson, R.C. (1975). Approximating the performance of urban emergency service systems. *Operations Research*, 23, 845–68.
- Levine, R.S, Yorita, K.L., Walsh, M.C., & Reynolds, M.G. (2009). A method for statistically comparing spatial distribution maps. *International Journal of Health Geographics*, 8, 7

- Lewin, M.R., Hori, S. & Aikawa, N. (2005). Emergency medical services in Japan: an opportunity for the rational development of pre-hospital care and research. *The Journal of Emergency Medicine*, 28, 237-241.
- Li, X, He, J and Liu, X, (2009). Intelligent GIS for solving high-dimensional site selection problems using ant colony optimization techniques. *International Journal of Geographical Information Science*, 23, 399–416
- Li, X. & Yeh, A.G. (2005). Integration of genetic algorithms and GIS for optimal location search. *International Journal of Geographical Information Science*, 19, 581-601.
- McConnel, C.E. & Wilson, R.W. (1998). The demand for prehospital emergency services in an aging society. *Social Science and Medicine*, 46, 1027-1031.
- Mitchell, M. (1998). *An Introduction to Genetic Algorithms*, MIT Press, Massachusetts
- Murray, A. & Estivill-Castro, V. (1998). Cluster Discovery Techniques for Exploratory Spatial Data Analysis. *International Journal of Geographical Information Science*, 12, 431-443.
- Narad, R.A. and Driesbock, K.R. (1999). Regulation of ambulance response times in California. *Prehospital Emergency Care*. 1999, 3:131-135.
- Niigata Prefecture (2008). *Fire Fighting Annual Report*, (Niigata: Niigata Prefecture).
- Peleg, K. & Pliskin, J.S. (2004). A geographic information system simulation model of EMS: reducing ambulance response time. *American Journal of Emergency Medicine*, 22, 164-170.
- Pitaksringkarn, L. & Taylor M.A.P, (2005). Grouping Genetic Alogirhtm In GIS: A Facility Location Modelling, *Journal of the Eastern Asia Society for Transportation Studies*, 6, 2908-2920.
- Pons, P.T., Haukoos, J.S., Bludworth, W., Cribley, T., Pons, K.A. & Markovchick, V.J., (2005). Paramedic response time: does it affect patient survival? *Academic Emergency Medicine* 12, 594-600.
- R Development Core Team (2008). *R: A Language and Environment for Statistical Computing*, (Vienna: The R Foundation for Statistical Computing, <http://www.R-project.org>).
- Reese, J. (2006) Solution Methods for the p -Median Problem: An Annotated Bibliography, *Networks*, 48(3), 125-142

- Righini, G. (1995). A double annealing algorithm for discrete location/allocation problems, *European Journal of Operational Research* 86, 452–468
- Rosing, K., Hillsman, E. & Rosing-Vogelaar, H. (1979). The robustness of two common heuristics for the P-median problem. *Environment and Planning A*, 11: 373–380.
- Samanta S. & Jha M.K. (2008). Identifying Feasible Locations for Rail Transit Stations Two-Stage Analytical Model. *Transportation Research Record*, 2063, 81-88
- Sasaki, S., Comber, A.J., Suzuki, H. & Brunson, B. (2010). Using genetic algorithms to optimise current and future health planning - the example of ambulance locations. *International Journal of Health Geographics*, 9:4
- Schilling, D.A., Jayaraman, V. & Barkhi, R. (1993). A review of covering problems in facility location. *Location Science*, 1, 25–55.
- Snyder, L.V. & Daskin, M.S. (2005). Reliability Models for Facility Location: The Expected Failure Cost Case. *Transportation Science*, 39(3): 400–416.
- Snyder, D.E., White, R.D. & Jorgenson, D.B. (2007). Outcome prediction for guidance of initial resuscitation protocol: Shock first or CPR first. *Resuscitation*, 72, 45–51.
- Stewart, T.J, Janssen, R. & van Herwijnen, M. (2004). A genetic algorithm approach to multiobjective land use planning. *Computers & Operations Research*, 3, 2293 – 2313
- Stiell, I.G., Wells, G.A., Demaio, V.J., Spaite, D.W., Field, B.J., Munkley, D.P., Lyver, M.B., Luinstra, L.G. & Ward, R., (1999). Modifiable factors associated with improved cardiac arrest survival in a multicenter BLS-D system: OPALS study phase I results. *Annals of Emergency Medicine*, 33, 44–50.
- Takeda, R.A, Widmer, J.A. & Morabito, R. (2007). Analysis of ambulance decentralization in an urban medical emergency service using the hypercube queuing model, *Computers and Operations Research*, 34, 727-741.
- Tanigawa, K., & Tanaka, K. (2006). Emergency medical service systems in Japan: Past, present, and future. *Resuscitation* 69, 365-370.
- Teitz, M.B. & Bart, P. (1968). Heuristic methods for estimating the generalized vertex median of a weighted graph. *Operations Research* 16, 955-961
- Tokyo Fire Department (2002). *Current status in EMS: 2002* (Tokyo: Tokyo Fire Department).

- Uyeno, D.H. & Seeberg, C. (1984). A practical methodology for ambulance location. *Simulation*, 43, 79-87
- Vink, C., & Schot, P. (2000). Application of a genetic algorithm in a GIS-based decision support system for multi-objective optimization of drinking water production. In *4th International Conference on Integrating GIS and Environmental Modeling (GIS/EM4): Problems, Prospects and Research Needs*, Banff, Alberta, Canada, 2-8 September.
- Willighagen, E. (2005). genalg: R Based Genetic Algorithm. R package version 0.1.1. <http://cran.r-project.org/web/packages/genalg/genalg.pdf> [available 15th June 2009]
- Xiao, N., Bennett, D.A. & Armstrong M.P. (2007). Interactive evolutionary approaches to multiobjective spatial decision making: A synthetic review. *Computers, Environment and Urban Systems*, 31, 232-252.
- Zhan, H.G., Lee, Z.P., Shi, P., Chen, C.Q. & Carder, K.L. (2003). Retrieval of water optical properties for optically deep waters using genetic algorithms. *IEEE Transactions on Geoscience and Remote Sensing*, 41, 1123-1128.

List of Tables and Figures

Table 1. A comparison of the distances between actual emergency cases to current and selected ambulance site locations (n/a indicates that it was not part of the set).

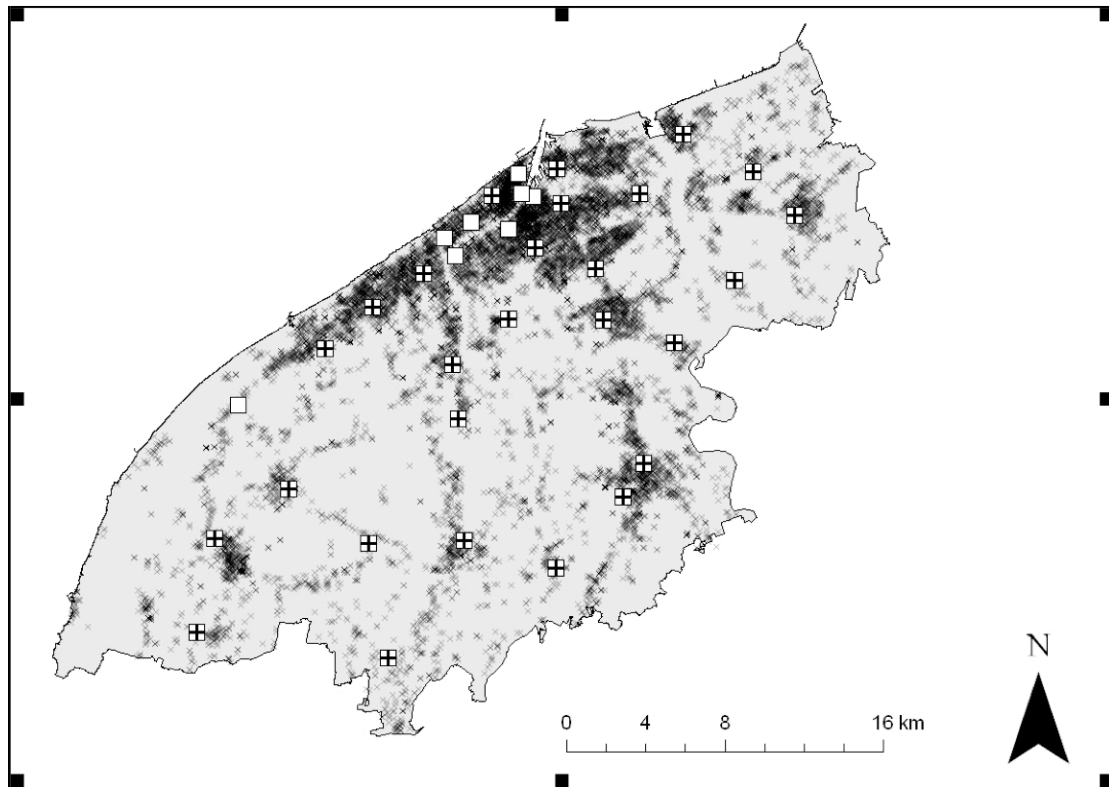


Figure 1. The study area showing fire stations without ambulances (squares), those with ambulances (squares with crosses) and the location of the 21,211 emergency cases (crosses).

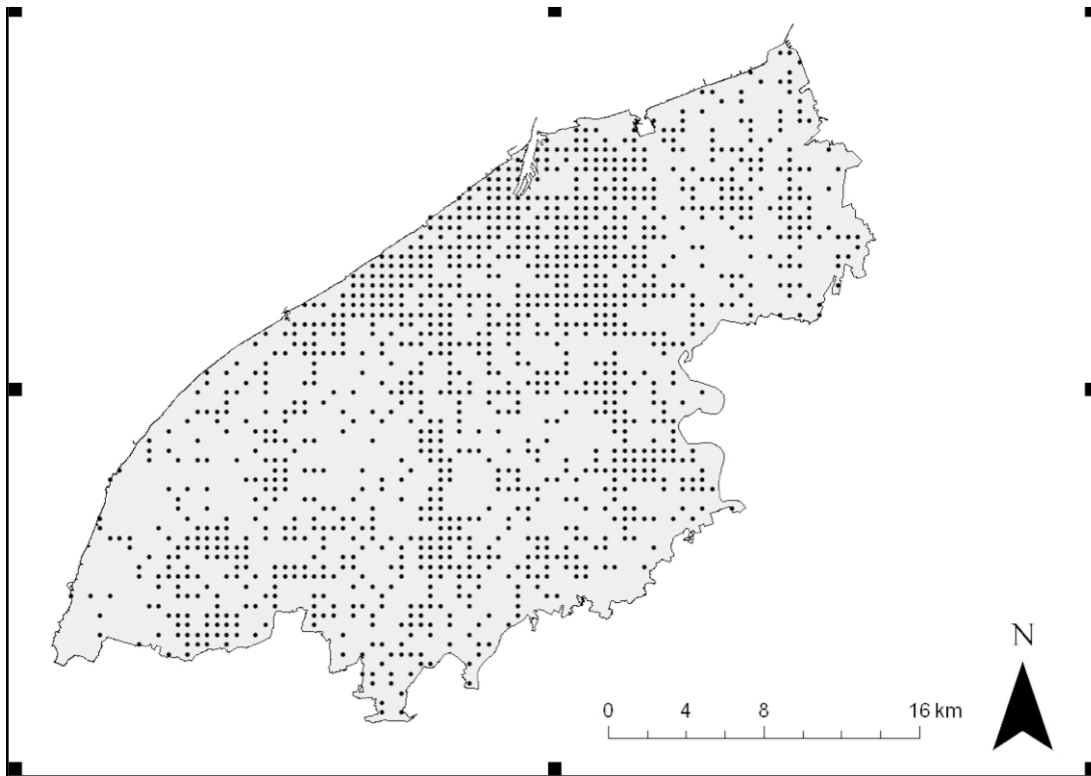


Figure 2. The study area with 1255 possible ambulance locations derived from a 500m point grid within 50m of an existing road.

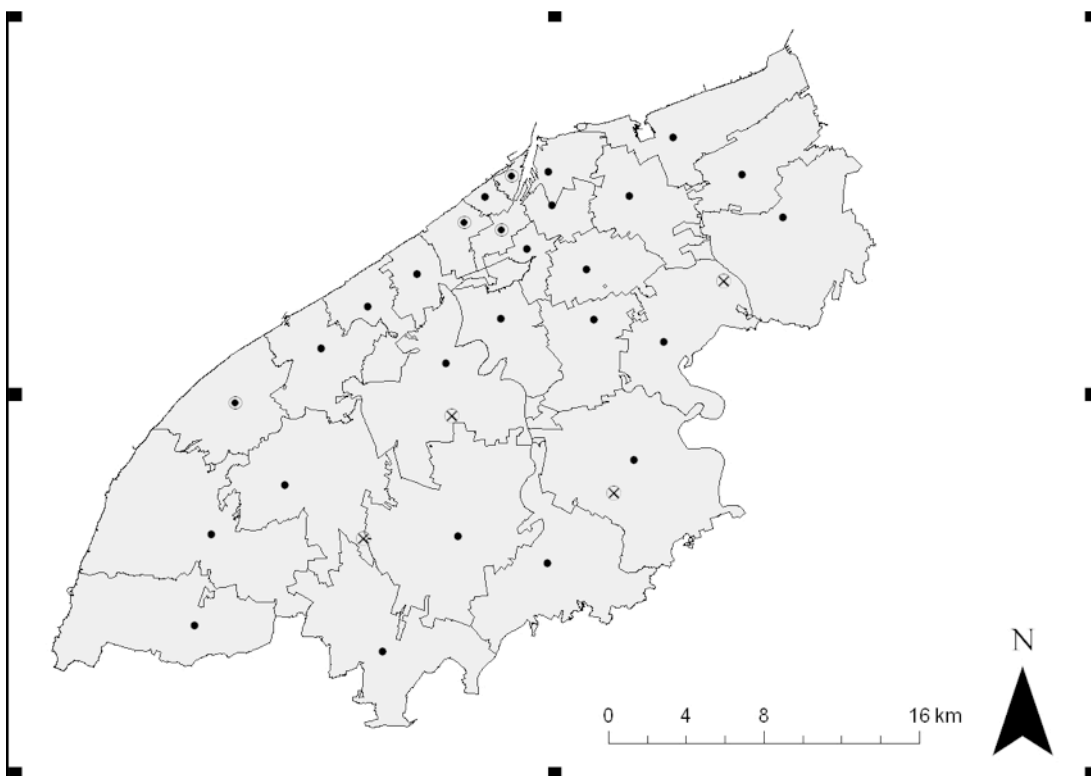


Figure 3. Selected and current ambulance site locations: selected sites that are also one of the 27 current locations (solid circle), selected sites at new locations (ringed circles) and existing sites not selected (circle with cross).

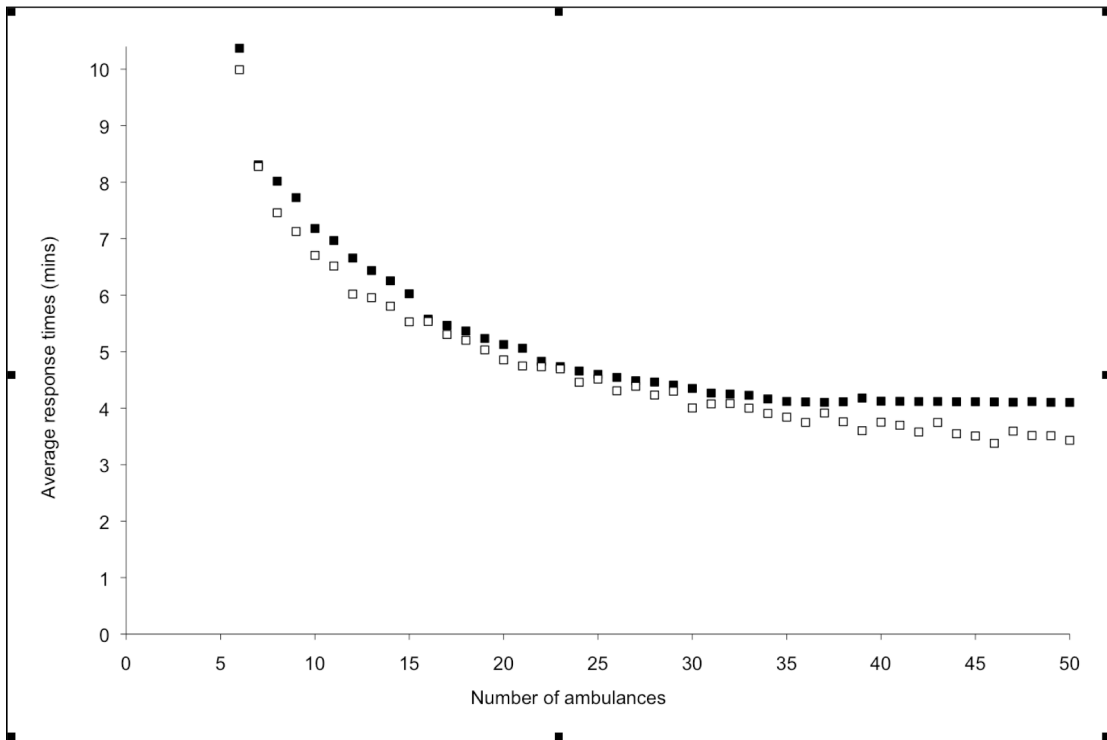


Figure 4. The change in average response times from census area centroids to their nearest ambulance as the number of potential stations increases selected from a) the current 35 potential locations (■) and b) the 1255 grid points (□).

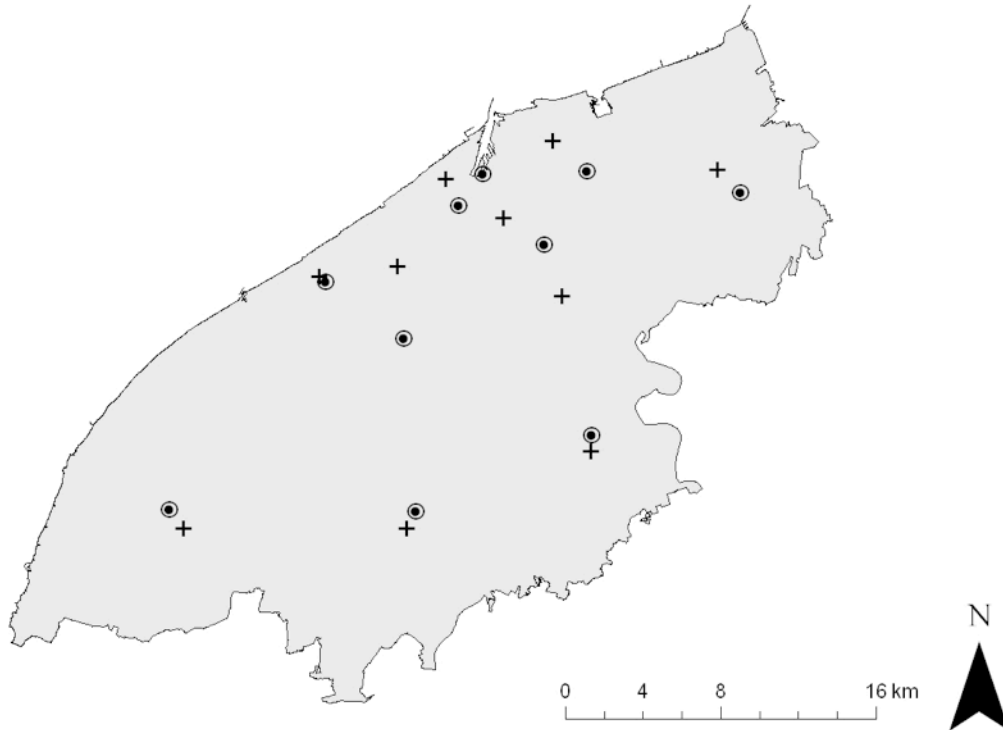




Figure 5. The distribution of selected ambulance site locations for $n =$ a) 10 ambulances, b) 20 ambulances and c) 30 ambulances, for locations based on existing sites (ringed circles) and 1255 grid points (crosses).