

Model Specification and Data Aggregation for Emergency Services Facility Location

Draft Date: 6 August 2014

Key words: Facility location; aggregation errors; model specification errors; emergency medical services.

Abstract:

In this paper, we use a computational approach to explore the relative impact of aggregation errors, model choice errors, and their interaction, for an ambulance station facility location model. We compare two model choices—probabilistic and deterministic—using a year of call data from the Edmonton Emergency Medical Service. We thereby demonstrate that model choice error dominates aggregation error, which implies that aggregation error is relatively unimportant if one uses the more realistic probabilistic model. The results are robust to perturbations in the way travel distances are measured and in the spatial demand distribution.

1. Introduction

Choosing locations for facilities is an important strategic decision for many organizations. The evaluation of a set of locations can involve many quantifiable and non-quantifiable factors. Focusing on the quantifiable factors, operations researchers, geographers, and others have developed numerous mathematical models and solution algorithms for optimal facility location. Most such models use aggregated spatial data. Analysts face three critical choices when tasked with supporting decisions about where to open new facilities: (1) which model to use, (2) how much to aggregate the spatial data, and (3) how to process aggregated input data. In this paper, we use an empirical and computational approach to investigate the

impact of these choices on the quality of the location decisions prescribed by the models in the specific setting of selecting locations for ambulance stations.

Models differ in terms of their realism, solvability, and data requirements. Typically, but not always, more realistic models require more computer time to solve and more data to populate. Spatial data is aggregated because of data availability, to protect privacy, and to render models solvable. In theory, the “finest grain” data (lowest possible aggregation level) may be a single individual or a household. In practice, statistical agencies typically aggregate demographic data to spatial units with a population of at least a few hundred people, in order to protect privacy and limit the data collection effort. When data are aggregated from finer to coarser levels and when spatial units are aggregated to nodes in a network, model input data (such as population at a node and distance from one node to another) must be aggregated. It is not always clear how best to accomplish this task.

We focus on the context of the choice of locations for ambulance stations so as to maximize *coverage*, which is typically measured, in practice, as the proportion of urgent patients whose response time is below a standard: 9 minutes, for example (Fitch, 2005). We compare the use of two models that differ only in one factor: whether the response time for a call is modeled as deterministic or stochastic. We use 2008 data for urgent ambulance calls in Edmonton to calibrate the models, investigating how the choice of model and the choice of aggregation level influence solution quality. Toward that end, we compare the quality of solutions obtained with imperfect models to the quality obtained with a disaggregated model that treats response times as stochastic.

We make the following contributions in this paper: (1) We extend a framework to study aggregation error for optimal facility location to include model specification; (2) we apply the

framework in the setting of selecting ambulance station locations to maximize coverage, using data from Edmonton, Alberta, and comparing two models that differ only in the treatment of response time as stochastic or deterministic; (3) we find that choice of aggregation level and model specification interact strongly, with the more true-to-life model being insensitive to aggregation error; and (4) we show that our findings are robust to such factors as the number of facilities, the spatial distribution of demand, and the distance metric.

2. Conceptual Development

The impact of aggregation of spatial data for optimal facility location models, which recently underwent a comprehensive review by Francis et al. (2009), has been studied for about 30 years. A key issue in the study of aggregation errors is the measurement of their impact. To describe the relevant issues precisely requires some notation; we adopt and extend the notation used by Francis et al. (2009). Any discrete optimal facility location problem involves the choice of a set, X , of facility locations in order to optimize an objective function, f . The objective function is measured with respect to a set of demand points, A ; and to emphasize this, $f(X: A)$ is used to denote the objective function, without any aggregation. Following Francis et al. (2009), we also use $f(X: A)$ to denote the optimal location problem. If we replace A with a set A' of aggregated demand points, we then obtain the approximate version, $f(X: A')$, of the disaggregated problem, $f(X: A)$. Suppose that X^* optimizes $f(X: A)$ and that X' optimizes $f(X: A')$. Some authors have focused on the *location error*: the differences between the true optimal locations X^* and the locations X' that are optimal for the aggregated problem. Others have emphasized the absolute or relative *measurement error*, $|f(X: A) - f(X: A')|$ or $|f(X: A) - f(X: A')|/f(X: A)$, respectively, which can be defined for any location set X (not merely an optimal set) for which $f(X: A)$ is

nonzero. Arguably, what matters most is the *optimality error*, $f(X^*: A) - f(X': A)$. In the context of maximizing coverage, the optimality error is the *coverage loss* that results from using the locations X' that optimize the aggregated problem rather than using the locations X^* that optimize the disaggregated problem, with coverage measured, for both location sets, using the disaggregated demand points A .

We extend the Francis et al. (2009) notation to include a choice between models M and M' , where M is more realistic than M' . Our *Gold Standard* is the more realistic model with finest-grain disaggregated data, denoted $f(X: A, M)$, because it provides the most accurate coverage measurement among the models that we consider. We measure the relative coverage loss that results from using model M' and demand points A' as:

$$\text{Relative coverage loss} = \frac{f(X^*: A, M) - f(X': A, M)}{f(X^*: A, M)},$$

where X' is the location set that maximizes the less realistic model with aggregated data, $f(X: A', M')$.

Researchers in various fields have investigated ways of comparing alternative mathematical models for the same phenomenon. The study of model specification in econometrics and statistics, for example, focuses on choosing between models that predict a dependent variable as a function of a set of independent variables, where the models differ either in the independent variables included or in the functional form of the relationship (see Chapters 5 and 6 in Greene, 2012, for an overview). As another example, control theorists investigate whether a linear approximation to a nonlinear system can be used to control the system so as to achieve satisfactory long-term behavior (see Chapter 15 in Brogan, 1991, for an introduction).

In contrast, our model comparison criterion is neither how well the model predicts a dependent variable nor the long-term behavior of a process; rather it is the consequences of following the actions prescribed by alternative models. This approach is conceptually similar to the one used by Besbes et al. (2010), who tested the validity of a statistical model to predict demand, based not on the accuracy of the model’s demand predictions but on the loss in profit that would result from using the model to prescribe an order size. Besbes et al. (2010) termed this testing approach an *operations* perspective on model validation—as opposed to a *statistical* perspective. Figure 1 (inspired by Figure 1 in Besbes et al., 2010) illustrates our coverage loss error measure in a simplified setting in which the decision variable is one-dimensional and continuous, rather than multi-dimensional and discrete.

We study the coverage loss that results from using imperfect models to select ambulance station locations. We study two model imperfections: the use of spatially aggregated input data and the approximation of coverage as depending deterministically on distance.

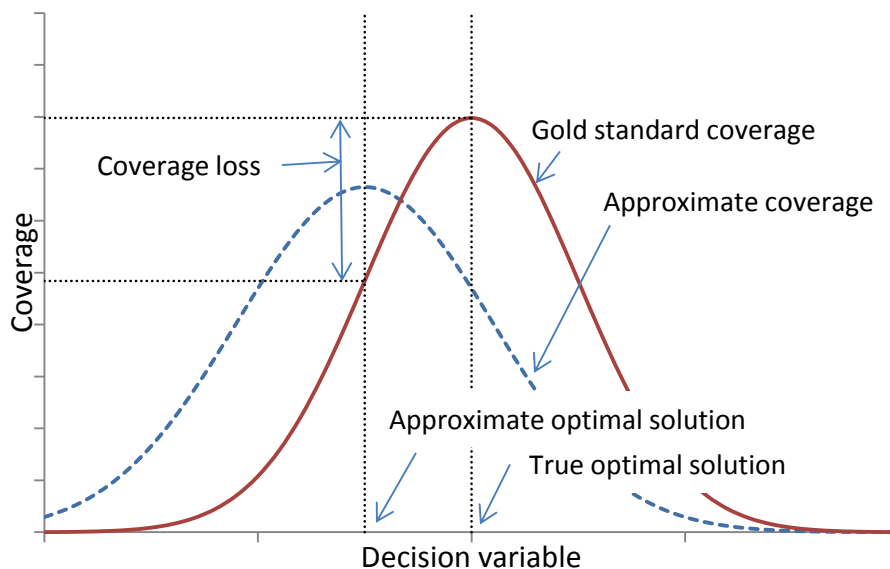


Figure 1: Illustration of coverage loss error measure.

The coverage of a geographical location in a city is not completely determined by the distance to the closest ambulance station for at least three reasons: (1) Pre-travel delay (from the time 911 is called until the responding ambulance starts driving towards the call address) is variable (Ingolfsson et al., 2008); (2) travel time is variable (Budge et al., 2010); and (3) the closest station may not have an available ambulance (Daskin, 1983), thus requiring the dispatch of an ambulance that is further away. We focus on the first two of these reasons: variability in pre-travel delay and travel time, whose sum equals the ambulance response time. We ignore the third reason (ambulance availability) because the models that we investigate are intended to support the strategic decision of selecting locations for ambulance stations, rather than the tactical decision of how to allocate ambulances to stations, which determines the availability of ambulances at individual stations. Effectively, our models assume that all stations have an available ambulance at all times, which implies that our models provide upper bounds on the coverage that can be achieved with a finite fleet of ambulances.

Figure 2 illustrates how aggregation and deterministic coverage approximation can lead to suboptimal solutions in a simple one-dimensional setting. We show station sites as triangles, demand nodes as circles, and coverage as curves—rectangular curves for deterministic coverage and bell-shaped curves for probabilistic coverage (see Appendix A in supporting information for information about how the coverage probability curves are calculated). The potential sites for Stations 1, 2, and 3 are at $x = 4, 8, \text{ and } 13$ km, for both panels. Panel a) shows four disaggregated demand nodes D1, D2, D3, and D4, at $x = 5.3, 5.9, 9.5, \text{ and } 11.1$ km, with demand magnitudes 1, 2, 1, and 2, respectively. Using deterministic coverage, Station 1 covers D1 and D2 fully, but D3 and D4 not at all. Using more true-to-life probabilistic coverage, Station 1 covers the four demand nodes with probabilities 0.76, 0.52, 0.04, and 0.01.

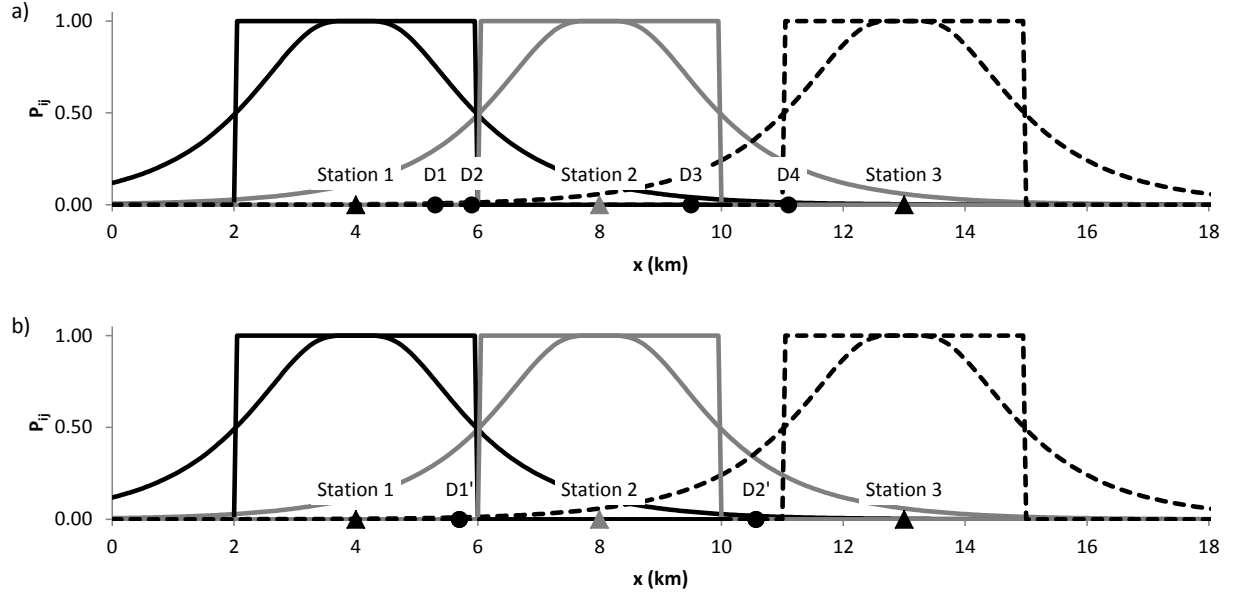


Figure 2: Illustration of probabilistic vs. deterministic coverage and its interaction with demand aggregation.

Panel b) illustrates aggregation of the four demand nodes to two demand nodes: D1 and D2 are aggregated to D1', with demand magnitude $1 + 2 = 3$, at the demand-weighted centroid of $(1 \times 5.3 + 2 \times 5.9)/(1 + 2) = 5.7$ km; and D3 and D4 are aggregated to D2', with demand magnitude 3 at 10.6 km. Using the aggregated data, Station 1 covers D1' but not D2', based on deterministic coverage, and Station 1's probabilistic coverage is 0.60 for D1' and 0.02 for D2'.

Suppose that we have a budget to open only one of Stations 1, 2, and 3. Table 1 shows the expected coverage for each station, as predicted using both deterministic and probabilistic coverage and both disaggregated and aggregated demand. According to our Gold Standard (Column 2 in Table 1: Probabilistic coverage and disaggregated demand), Station 2 maximizes expected coverage, at 39.0% of the total demand. A model that uses deterministic coverage and disaggregated demand recommends Station 1 instead, resulting in a *coverage loss* of $39.0\% - 31.2\% = 7.8\%$. If one aggregates demand, then the probabilistic coverage model continues to

prescribe Station 2, and the deterministic coverage model continues to prescribe Station 1. We also see that the accuracy of the coverage predictions (compared to the Gold Standard) worsens with aggregation for deterministic coverage, but is not greatly affected for probabilistic coverage.

Table 1: Predicted expected coverage.

| Station | Disaggregated demand | | Aggregated demand | |
|---------|-----------------------------------|---------------|-------------------|---------------|
| | Prob. coverage (Gold standard) | Det. coverage | Prob. coverage | Det. coverage |
| 1 | 31.2% | 50.0% | 30.8% | 50.0% |
| 2 | 39.0% | 16.7% | 36.4% | 0.0% |
| 3 | 20.8% | 33.3% | 18.7% | 0.0% |

In the remainder of the paper, we investigate how aggregating spatial data and approximating coverage as deterministic influence coverage loss in a realistic setting.

3. Methods, Data, and Experimental Design

In this section, we specify the two models that we compare, describe our data and the estimation of the model input parameters, and discuss the factors that vary in our experimental design.

Before defining the two optimization models that we compare, we define a set of coverage parameters P_{ij} . The deterministic model (D) and the probabilistic model (P) differ only in the way these coverage parameters are defined, as functions of the distance d_{ij} from Station j to Demand Node i . In the disaggregate P model, the coverage P_{ij} is the probability that response time $R_{ij}(d_{ij})$ for an ambulance at Station j responding to a call at Node i is less than or equal to a

coverage time threshold t_c , as illustrated in Figure 3. In the disaggregate D model, the coverage P_{ij} is 1 if the distance d_{ij} is less than or equal to a coverage distance threshold d_c and 0 otherwise. We choose the distance threshold d_c to satisfy $\Pr(R(d_c) \leq t_c) = \alpha$; that is, we set d_c to be the distance at which the α -fractile of the response time distribution equals the coverage time threshold t_c (see Figure 3). We vary the parameter α in our experimental design. Details of the computation for P_{ij} and d_c are discussed in Appendix A in the supporting information.

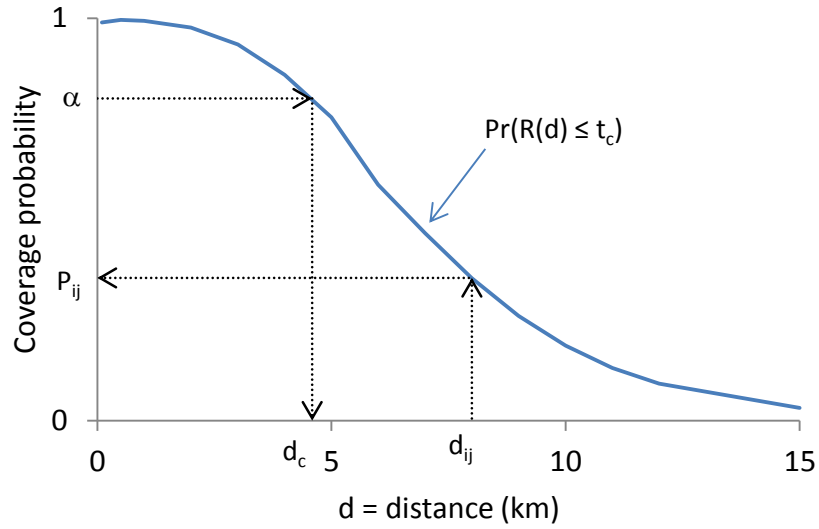


Figure 3: Illustration of the definitions of coverage probabilities P_{ij} and coverage distance threshold d_c .

To formulate the P and D models, let binary decision variable x_j be an indicator for whether potential station location $j \in J$ is one of the q locations selected and let binary decision variable y_{ij} be an indicator for whether station $j \in J$ is the one with the highest coverage P_{ij} of demand node $i \in I$ (we say that demand node i is “assigned to” station location j). The demand per time unit is D_i at node $i \in I$.

Problem Formulation:

$$\text{Maximize } \sum_{i \in I} \sum_{j \in J} D_i P_{ij} y_{ij} \quad (1)$$

subject to:

$$y_{ij} \leq x_j \quad \forall i \in I, j \in J \quad (2)$$

$$\sum_{j \in J} y_{ij} = 1 \quad \forall i \in I \quad (3)$$

$$\sum_{j \in J} x_j = q \quad (4)$$

$$x_j \in \{0,1\}, y_{ij} \in \{0,1\} \quad \forall i \in I, j \in J \quad (5)$$

The objective function (1) is the amount of demand covered. Constraints (2) and (3) together ensure that every demand node is assigned to a single “open” station. Constraint (4) sets the budget for the total number of stations to q and constraints (5) enforce binary restrictions on the decision variables.

When one changes the level of aggregation, the set of demand nodes I , the demand parameters D_i , the distances d_{ij} , and the coverage parameters P_{ij} all change. The set of potential station locations J could change, but we keep this set fixed in our experimental design. To illustrate the procedure we use to aggregate distances and coverage parameters, we assume that disaggregated Nodes 1 and 2 are aggregated into Node 3. We compute the demand, distances, and coverage parameters for Aggregate Node 3 as follows:

$$\text{Demand: } D_3 = D_1 + D_2$$

$$\text{Distance: } d_{3j} = (D_1 d_{1j} + D_2 d_{2j}) / (D_1 + D_2) \quad j \in J$$

$$\text{Coverage for P model: } P_{3j} = (D_1 P_{1j} + D_2 P_{2j}) / (D_1 + D_2) \quad j \in J$$

$$\text{Coverage for D model: } P_{3j} = 1 \text{ if } d_{3j} < d_c, 0 \text{ otherwise} \quad j \in J$$

We use Aggregation Scheme C from Daskin et al. (1989) to generate demand node sets of various sizes. Appendix B in the supporting information describes the details of this aggregation method, which is designed to distribute the aggregated nodes throughout the region

of interest, while providing a higher node density in sub-regions with a higher demand density. We use the same method to generate the set J of potential station locations.

Statistical agencies typically aggregate census data, resulting in census regions. In Canada, the smallest regions for which demographic and socioeconomic information is available are *dissemination areas* (DAs, see Statistics Canada 2012), which are designed to have a population of 400 to 700 persons. Edmonton has 1,084 DAs, which we use as our finest-grain aggregation level.

We use data on all 46,691 urgent ambulance calls made in 2008 in Edmonton, Alberta, Canada. We ignore non-urgent calls, because our interest is in modeling response to medically urgent calls rather than total system workload. The record for each call includes the call location, location of the responding ambulance at the time of dispatch, and response time. We use a coverage time threshold of $t_c = 9$ minutes.

We vary five factors in our computational experiments: Model, Aggregation, Distance Metric, Demand Distribution, and Number of Stations. We discuss the levels that we use for these experimental factors in the remainder of this section.

$M = \text{Model}$: We compare the more realistic P model to four variants of the D model: $M' = D50, D75, D80, \text{ and } D85$, corresponding to setting the coverage-distance fractile α equal to 0.50, 0.75, 0.80, and 0.85.

$A = \text{Aggregation}$: We compare the disaggregated data with 1,084 nodes to 15 aggregation levels, with 1,000, 900, 800, 700, 600, 500, 400, 300, 200, 150, 125, 100, 75, 50, and 25 aggregate nodes.

Distance Metric : Our base case distance metric is road network distance; it respects natural barriers such as the North Saskatchewan River, which splits Edmonton in two. We also

use Euclidean (straight-line) and rectilinear distance metrics, which ignore natural barriers, allowing us to test the robustness of our results to the structure of the road network.

Demand Distribution: Our base case demand parameters D_i are based on the 2008 Edmonton dataset, and they reflect a heavy concentration of demand in the city core that characterizes ambulance demand in many cities. We also used equal demand for all disaggregate demand nodes, to check the robustness of our results against demand concentration in the city core.

Number of Stations: We compared values of $q = 5, 10, 15$, and 20 as the budget for the number of stations. For comparison, there were 13 ambulance stations in Edmonton in 2008. We obtained the set J of potential station locations by using the Daskin et al. (1983) Aggregation Scheme C to select 200 of the 1,084 DAs.

We solved a total of $5 \times 16 \times 3 \times 2 \times 4 = 1,920$ instances of the station facility location optimization problem, using the Cplex (2013) solver. We were able to solve all problem instances to optimality within reasonable computation times, allowing us to focus on aggregation and model choice, without having to worry about model solvability.

The performance measure that we focus on is the relative coverage loss, which we defined as follows in Section 2:

$$\text{Relative coverage loss} = \frac{f(X^*: A, M) - f(X': A, M)}{f(X^*: A, M)},$$

where X' is the optimal solution to the approximate problem $f(X: A', M')$, with aggregation level $A' = 1,000, 900, \dots, 25$ and deterministic model $M' = D50, D75, D80$, or $D85$. The Gold Standard coverage $f(X^*: A, M)$ is the coverage obtained with the optimal solution to the probabilistic model with no aggregation. The Gold Standard coverage changes when the distance metric, the demand distribution, or the number of stations change, but the Gold Standard

coverage remains constant when the coverage distance fractile α for the D model or the aggregation level are varied.

The Gold Standard coverage is, by definition, as good or better than the coverage obtained by solving the D model or using aggregate data, which implies that $f(X^*: A, M) \geq f(X': A, M)$; therefore the relative coverage loss will never be negative. We do not claim that our Gold Standard model is perfect or the best possible model—we claim only that the P model is more realistic than the D model and that not aggregating is more realistic than aggregating.

4. Results

Figure 4 shows how relative coverage loss varies with aggregation level (on a log scale) for the five models, for a scenario with a budget for five stations, using network distances, and actual demand. This figure illustrates our two main findings: (1) If one uses the probabilistic model, then the aggregation error is negligible, even for extreme levels of aggregation and (2) all of the deterministic models result in large coverage losses that decrease inconsistently, if at all, when the level of aggregation is reduced. The evidence for the first finding is that the coverage loss for the P model is close to zero when the number of aggregated demand nodes is 50 or more. The evidence for the second finding is that the coverage loss curves for the D models are erratic and typically far above zero.

These two findings are robust across distance measures, demand distributions, and number of stations. Figure 5 provides another example of how the relative coverage loss depends on model choice and aggregation when 20 stations are selected, using Euclidean distances and uniform demand. Appendix C in the supporting information shows graphs of relative coverage loss versus aggregation for all 24 combinations of distance metrics, demand distributions, and number of stations.

To elaborate on the first finding, Figure 6 shows the maximum and average relative coverage loss (across distance metrics, demand distributions, and number of stations) for the P model. As depicted in Figure 6, even when the 1,084 demand nodes are aggregated to only 50 nodes, the coverage loss is only about 0.3% on average and less than 1% in the worst case.

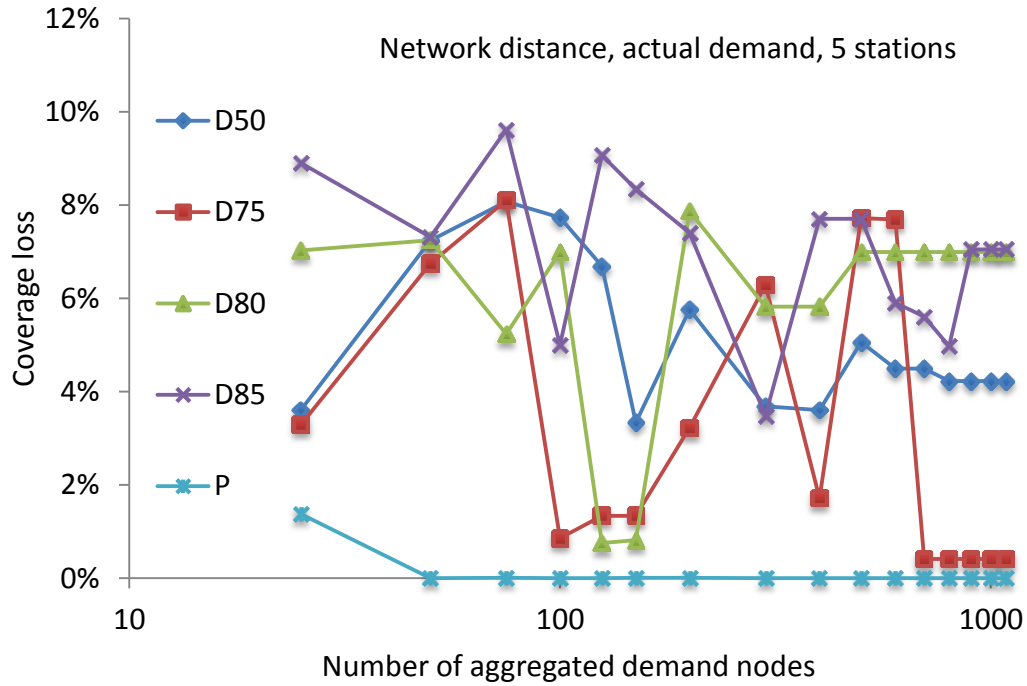


Figure 4: Relative coverage loss for network distance, actual demand, and 5 stations.

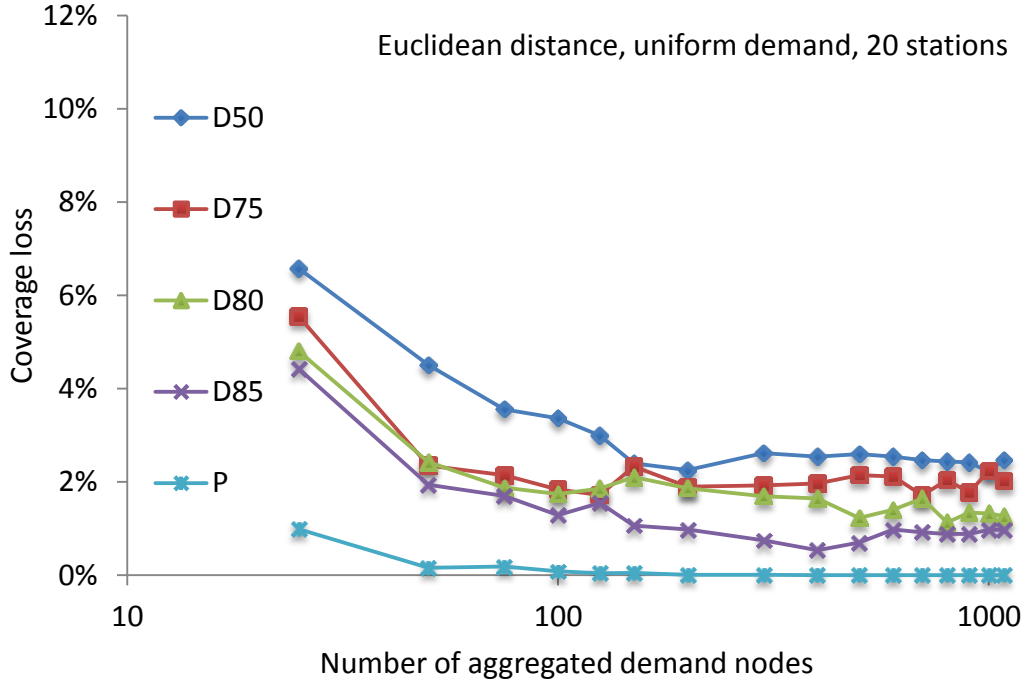


Figure 5: Relative coverage loss for Euclidean distance, uniform demand, and 20 stations.

Regarding the second finding, two questions arise: (1) whether the coverage loss for the deterministic models tends to increase with aggregation and (2) whether the coverage loss depends systematically on the coverage distance fractile. Judging from Figure 4, the answers to these questions are far from clear. Judging from Figure 5, however, it appears that coverage loss does indeed tend to increase with aggregation and that using the 85th percentile to determine the coverage distance systematically improves coverage loss (for uniform demand, Euclidean distances, and 20 stations). In the next two paragraphs, we discuss how the answer to these questions varies across distance metrics, demand distributions, and number of stations.

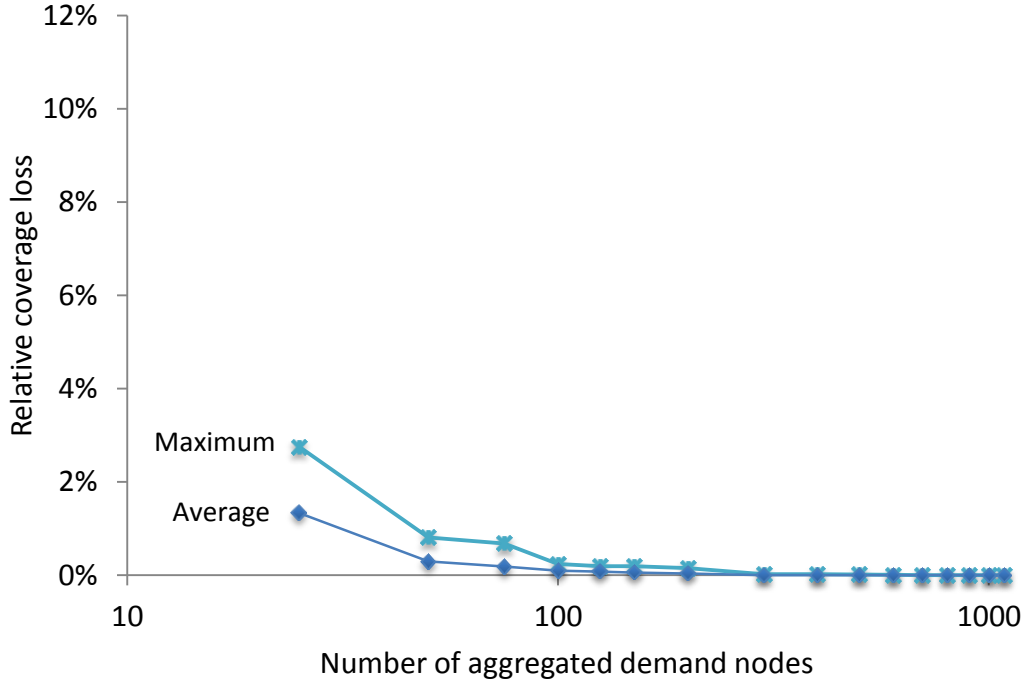


Figure 6: Maximum relative coverage loss for the P model.

One way of answering the question of whether coverage loss increases with aggregation is to use regression analysis to fit a curve to the relationship between the number of aggregate nodes, A , and the relative coverage loss. Figure 7 illustrates three regression curves for the D75 model, network distances, actual demand, and 5 stations, based on the following three regression model specifications:

$$\begin{aligned}
 \text{Linear : } \quad \text{RCL} &= \alpha + \sum_{M'} \beta_{M'} M' + \sum_{M'} \gamma_{M'} A M' + \text{error term} \\
 \text{Loglinear : } \quad \ln \text{RCL} &= \alpha + \sum_{M'} \beta_{M'} M' + \sum_{M'} \gamma_{M'} A M' + \text{error term} \\
 \text{Power law : } \quad \ln \text{RCL} &= \alpha + \sum_{M'} \beta_{M'} M' + \sum_{M'} \gamma_{M'} (\ln A)(M') + \text{error term}
 \end{aligned}$$

We fit each of these regression models to each combination of distance metric, demand distribution, and number of stations. We used dummy variables and interaction terms to allow the “intercept” and “slope” coefficients to differ among the four deterministic models $M' = \text{D50, D75, D80, and D85}$. In all three regression specifications, a negative estimated slope coefficient

γ_M indicates that the fitted relationship between the number of aggregate nodes A and the relative coverage loss, RCL, decreases, which means that coverage loss increases with aggregation.

Table 2 shows that the three model specifications lead to largely consistent results regarding the proportion of scenarios in which the fitted relationship decreases. The proportion ranges from 18 to 24 out of the 24 scenarios (75% to 100%) defined by the distance metric, demand distribution, and number of stations, so the answer to the first question—does coverage loss increase with aggregation?—is yes in most, but not all, scenarios. We reiterate, however, that there is a great deal of variation around the fitted relationships, as Figure 7 illustrates, and the practically more important finding is that when possible one should use the P model rather than the D models.

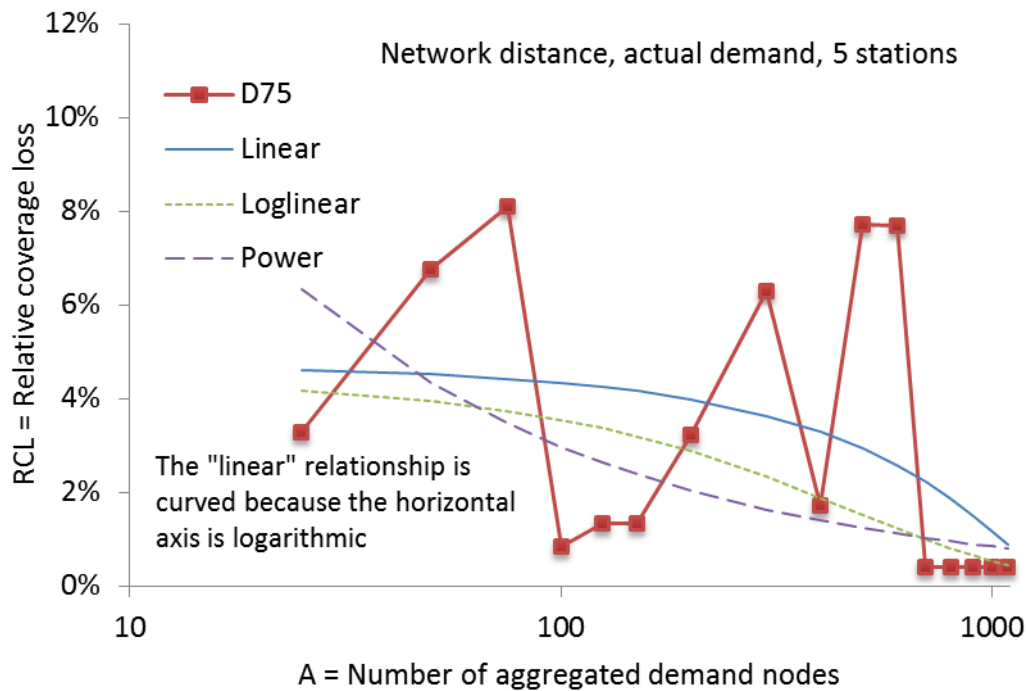


Figure 7: Estimated regression relationships between the number of aggregated demand nodes and the coverage loss.

Table 2: Decrease in coverage loss.

| | D50 | D75 | D80 | D85 |
|------------------------------------|-------|-------|-------|-------|
| Negative slope for linear model | 22/24 | 23/24 | 21/24 | 19/24 |
| Negative slope for loglinear model | 22/24 | 23/24 | 21/24 | 18/24 |
| Negative slope for power model | 23/24 | 24/24 | 21/24 | 20/24 |

The second question is whether the coverage loss depends systematically on the coverage distance fractile. We find that it does, on average, as shown in Figure 8. The lowest fractile performs better for a small number of stations, but the highest fractile performs better for a large number of stations. To understand why, first note (from Figure 3) that higher fractiles result in shorter coverage distance standards. If selecting locations for a large number of stations, a high fractile makes it more likely that the coverage areas of different stations will overlap. If the coverage areas are large enough, then all of the demand will be counted as being covered, according to the objective function of the deterministic model, and it is likely that there will be multiple alternate optimal solutions, all of which provide 100% deterministic coverage. The alternate optimal solutions do differ when measured using probabilistic coverage, however. Using a low fractile decreases station coverage areas and causes the deterministic model objective function to become a better proxy for probabilistic coverage.

If selecting locations for a small number of stations, on the other hand, the deterministic coverage areas are less likely to overlap. In this setting, using a small coverage distance fractile results in large coverage areas, which means that the optimal solution is likely to be one that spreads the stations far apart, which also tends to optimize probabilistic coverage.

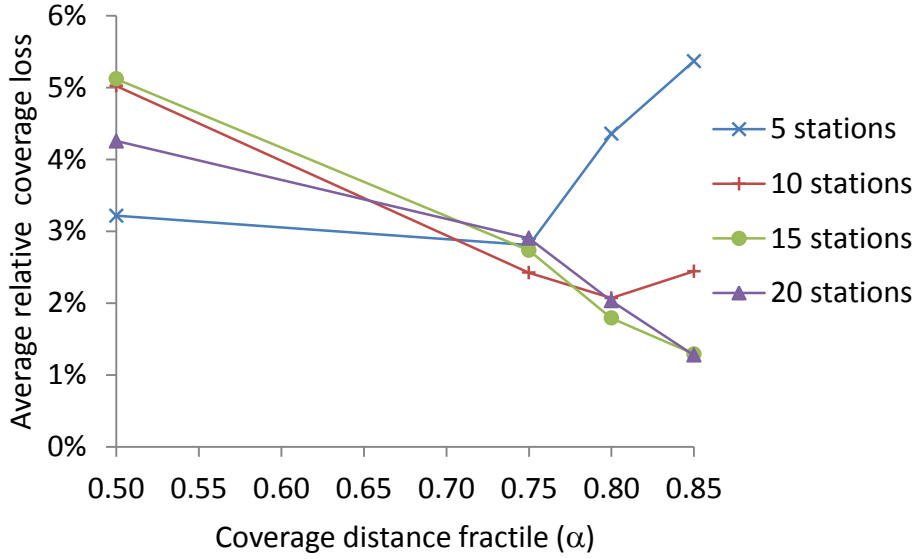


Figure 8: Average relative coverage loss for the D models, as the number of stations changes.

5. Discussion

Two natural questions about the scope of our findings arise: (1) whether the findings generalize to emergency medical service (EMS) systems in other jurisdictions and (2) whether the findings generalize to other contexts. We attempted to address the first question by varying the distance metric and the demand distribution, to test the robustness of our results to variations in geographic structure among cities; and by varying the budget for the number of stations, to test robustness to the resources available to the EMS service. Although the magnitude of the coverage loss for the deterministic models varied slightly with these factors, our two primary findings—that the probabilistic model has negligible aggregation error and that the deterministic models have large aggregation errors—held true in all scenarios that we tested (see Appendix C in the supporting information).

Although we are less confident that our findings generalize to other contexts, we note that the two models we compared are representative members of two of the most studied model

classes in optimal facility location (Daskin, 2008): covering-type models (our D model) and median-type models (our P model). Location theorists have studied aggregation errors for these model classes and concluded that solution quality for covering-type models is relatively sensitive to aggregation, compared to median-type models (Current and Schilling, 1990), but solution quality for median-type models is relatively insensitive to aggregation (Erkut and Bozkaya, 1999). Our findings are consistent with these conclusions, but it appears that no previously published studies take our perspective: comparing coverage and median models directly in a setting in which one model is clearly more realistic than the other. To understand why the P model is less sensitive to aggregation, note that it has the continuity property that if the boundaries of the regions that are aggregated into aggregate demand nodes change by a small amount, then the coverage parameters P_{ij} will change by a small amount. Intuitively, one would therefore expect that gradual changes to the aggregation level should result in only gradual changes to the optimal solution to the P model. In contrast, in the D model, small boundary changes can result in large changes in the coverage parameters (from 0 to 1 or from 1 to 0), which can cause large and difficult-to-predict changes to the optimal solution, resulting in the erratic coverage loss curves that we see in Figure 4.

One potential objection to the use of our probabilistic model is that it requires knowledge of the probability distribution of response time R_{ij} from any station j to any demand node i , whereas the deterministic model requires only distances d_{ij} , which are easily obtained with geographic information systems and online mapping systems. We begin by noting that distances are not enough even for the deterministic model—in order to arrive at a defensible coverage distance threshold d_c , one requires at least point estimates of the pre-travel delay and average travel speed, in order to estimate how far ambulances can travel within the coverage time

threshold t_c that the EMS service uses to measure its performance. We use what we believe is a relatively accurate method to model travel time distributions, based on the work of Budge et al. (2010), which requires the estimation of six parameters (see Appendix A in supporting information) from empirical travel time data. One can envision simpler methods that may achieve almost all of the benefits of the probabilistic model, without requiring extensive data collection and statistical model fitting. Perhaps the simplest method would be to use point estimates of pre-travel delay and average travel speed (information that is needed for the deterministic model) in order to obtain estimates of median response times and to assume that response times follow some convenient two-parameter positive probability distribution (such as lognormal or gamma) with a coefficient of variation that is independent of distance. The work of Budge et al. (2010) and Ingolfsson et al. (2008) suggests that coefficients of variation in the range 0.2 to 0.4 could be appropriate. Thus, use of the probabilistic model may require only one more parameter estimate, compared to the deterministic model.

Although the relationship between coverage loss and aggregation for the deterministic models and the relationship between the “best coverage distance fractile” and the number of stations are interesting, we would like to reiterate the practical consequence of our findings: Simply using the probabilistic model results in far lower coverage loss than any of the deterministic models. In all the scenarios we tested, the aggregation error associated with the probabilistic model was extremely low, if it existed at all, even for more extreme levels of aggregation than a thoughtful analyst would ever use.

6. Conclusions

In this paper, we demonstrated that the use of coverage probabilities rather than deterministic coverage thresholds reduces the deleterious effects of demand point aggregation on solution quality for ambulance station site selection optimization models. We find that for the probabilistic version of the optimization model, the effects of demand-point aggregation are minimal, even for high levels of spatial aggregation. This finding generalizes across distance metrics (existing road network distances, Euclidean distances, and rectilinear distances), spatial demand distributions (actual demand or uniform demand), the budget for the number of stations (5 to 20), and aggregation levels that range from 25 to 1,084 for a city with a population of approximately 1 million.

We used five different coverage distance thresholds for the deterministic model, based on different fractiles of the response time distribution. Although the aggregation errors for the deterministic model were erratic in general, we found that on average, the median (the 0.50-fractile) tended to work better for smaller station budgets, whereas higher fractiles (resulting in smaller station coverage areas) tended to work better for larger station budgets.

In addition to being more realistic and much less sensitive to the effects of demand point aggregation, the probabilistic model is about as computationally tractable as the deterministic one, and the additional input data required by the probabilistic model can be as little as a single parameter: a constant response-time coefficient of variation. Our findings indicate that probabilistic models are an effective method for mitigating errors associated with demand-point aggregation. The probabilistic model “softens” aggregation error, and in many cases virtually eliminates it, except for extreme levels of aggregation.

References.

- Besbes, O., Phillips, R., Zeevi, A. (2010) Testing the validity of a demand model: an operations perspective. *Manufacturing & Service Operations Management* 12, 162–183.
- Brogan, W. L. (1991) *Modern Control Theory*, third ed. Prentice Hall, Englewood Cliffs, NJ.
- Budge, S., Ingolfsson, A., Zerom, D. (2010) Empirical analysis of ambulance travel times: the case of Calgary emergency medical services. *Management Science* 56(4), 716–723.
- Current, J.R., Schilling, D.A. (1990) Analysis of errors due to demand data aggregation in the set covering and maximal covering location problems. *Geographical Analysis* 22, 116–126.
- Cplex (2013) <http://www-01.ibm.com/software/commerce/optimization/cplex-optimizer/> (last accessed 10 October 2013).
- Daskin, M., (2008) What you should know about location modeling. *Naval Research Logistics* 55, 283–294.
- Daskin, M. (1983) The maximal expected covering location model: formulation, properties and heuristic solution. *Transportation Science* 17(1), 48–70.
- Daskin, M.S., Haghani, A.E., Khanal, M., Malandraki, C. (1989) Aggregation effects in maximum covering models. *Annals of Operations Research* 18(1), 115–139.
- Erkut, E., Bozkaya, B. (1999) Analysis of aggregation errors for the p-median problem. *Computers & Operations Research* 26(10), 1075–1096.
- Fitch, J. (2005) Response times: myths, measurement & management. *JEMS: A Journal of Emergency Medical Services* 30(1), 46–56.
- Francis, R. L., Lowe, T.J, Rayco, M.B., Tamir, A. (2009) Aggregation error for location models: survey and analysis. *Annals of Operations Research* 167(1), 171–208.

Greene, W.H. (2012) *Econometric Analysis*, seventh ed. Pearson Education, Inc., Upper Saddle River, NJ.

Ingolfsson, A., Budge, S., Erkut, E. (2008) Optimal ambulance location with random delays and travel times. *Health Care Management Science* 11(3), 262–274.

Statistics Canada (2012) 2011 Census Dictionary. Catalogue no. 98-301-X2011001. Ottawa, Ontario, <http://www5.statcan.gc.ca/bsolc/olc-cel/olc-cel?lang=eng&catno=98-301-X>