

# Chapter 18

## Aggregation in Location



Richard L. Francis and Timothy J. Lowe

**Abstract** Location problems occurring in urban or regional settings may involve many tens of thousands of “demand points,” usually individual residences. In modeling such problems it is common to aggregate demand points to obtain tractable models. We discuss aggregation approaches to a large class of location models, consider various aggregation error measures, and identify some effective measures. In particular, we focus on an upper bounding methodology for the error associated with aggregation. The chapter includes an example application.

### 18.1 Introduction

Many location problems involve locating services (called *servers*) with respect to customers of some sort (called *demand points*, and abbreviated as DPs). Usually there is travel between servers and DPs, so that travel distances, or (more generally) travel costs, are of interest. Location models represent these travel costs, and solutions to the models can provide locations of the servers of (nearly) minimal cost. For books on location models and modeling, see Daskin (2013), Drezner (1995), Drezner and Hamacher (2002), Eiselt and Marianov (2011), Francis et al. (1992), Handler and Mirchandani (1979), Love et al. (1988), Mirchandani and Francis (1990), and Nickel and Puerto (2005).

A common difficulty with modeling location problems that occur in urban or regional areas is that the number of DPs may be quite large, since each private residence might be a DP. In this case it may be impossible, and also unnecessary, to include every DP in the corresponding model. Further, the corresponding problems

---

R. L. Francis

Industrial and Systems Engineering Department, University of Florida, Gainesville, FL, USA

e-mail: [Francis@ise.ufl.edu](mailto:Francis@ise.ufl.edu)

T. J. Lowe (✉)

Management Science Department, Tippie College of Business, University of Iowa, Iowa City, IA, USA

e-mail: [timothy-lowe@uiowa.edu](mailto:timothy-lowe@uiowa.edu)

may be NP-hard to optimize (Kariv and Hakimi 1979). For problems as diverse as those including the location of branch banks (Chelst et al. 1998), tax offices (Domich et al. 1991), network traffic flow (Sheffi 1985), and vehicle exhaust emission inspection stations (Francis and Lowe 1992) a popular aggregation approach is used: to suppose every DP in each postal code area or zone of the larger urban area is at the centroid of the postal code area or zone, and to compute distances accordingly. The result is a smaller model to deal with, but one with an intrinsic error. If the modeler wishes to aggregate to have a small number of *aggregate demand points* (abbreviated as ADPs), and also desires a small error, then aggregation becomes a nontrivial matter.

It is tempting to ask the following question: How many ADPs are enough? There are no general answers to this question. This is because there are important tradeoffs in doing aggregation. Aggregation often decreases: (1) data collection cost, (2) modeling cost, (3) computing cost, (4) confidentiality concerns and (5) data statistical uncertainty. The first four items seem self-explanatory; item (5) occurs because aggregation leads to pooled data, which provides larger samples and thus smaller sample standard deviations. The price paid for aggregation is increased model error: instead of working with the actual location model we work with some approximate location model. How to trade off the benefits and costs of aggregation is still an open question. The question is open in part because there is no general agreement on how to measure the aggregation error, and also because there is no accepted way to attach a cost to model aggregation error. To the best of our knowledge, professional judgment is generally used to do the tradeoffs. Francis et al. (2009) provide a survey of various demand point aggregation error measures and an extensive literature discussion. In fact, much of the early material in this chapter, and Table 18.4, is from that paper.

One can categorize location models as *strategic*, *tactical*, or *operational* in scope. As pointed out by Bender et al. (2001), planar distances are often used for strategic-level location models, and network distances for tactical-level location models. Such models are often converted to equivalent mixed integer programming (MIP) models for solution purposes, using some finite dominating set principle to reduce the set of possible locations of interest to a finite set (Hooker et al. 1991). Thus, results to follow for these planar and network models also apply to their MIP representations, including those for the  $p$ -median,  $p$ -center, and covering location models. These models are the subject matter of Chaps. 2, 3, and 5 respectively. Operational-level location models are not too common (mobile servers are one example), but for such models no aggregation may be best. Note that the scope of the location model may well indicate the degree of aggregation; a more detailed scope requires a more detailed aggregation.

## 18.2 Terminology and Examples

We suppose that servers and DPs are all either points in the plane, or on some travel network. In either case, there is some well-defined set of server points and DPs, say  $\Omega$ , and a distance  $d(x,y)$  between any two points  $x, y$  in  $\Omega$ . If  $\Omega$  is a travel network (assumed undirected) then  $d(x,y)$  is usually the length of a shortest path between  $x$  and  $y$ . For planar problems when  $\Omega = \mathbb{R}^2$ , with  $x = (\chi_1, \chi_2)$ ,  $y = (\psi_1, \psi_2)$ ,  $d(x,y)$  is often the  $\ell_p$ -distance:  $\|x - y\|_p = [|\chi_1 - \psi_1|^p + |\chi_2 - \psi_2|^p]^{1/p}$ , with  $p \geq 1$ . Taking  $p = 1$  or  $2$  gives the well-known rectilinear or Euclidean distance, respectively. The limiting case of the  $\ell_p$ -distance as  $p$  goes to infinity, denoted by  $\|x - y\|_\infty$ , is given by  $\|x - y\|_\infty = \max\{|\chi_1 - \psi_1|, |\chi_2 - \psi_2|\}$ , and is called the Tchebyshev distance. The Tchebyshev distance in  $\mathbb{R}^2$  is sometimes analytically convenient because it is known (Francis et al. 1992) to be equivalent to the planar rectilinear distance under a 45-degree rotation of the axes. We define the diameter of  $\Omega$  by  $diam(\Omega) = \sup\{d(x,y) : x, y \in \Omega\}$ , with the understanding that possibly  $diam(\Omega) = +\infty$ . More generally,  $\Omega$  can be a metric space (Goldberg 1976) with metric  $d$ , but no loss of insight occurs by considering the network and planar cases for  $\Omega$ .

Suppose we have  $n$  DPs,  $v_j \in \Omega, j = 1, \dots, n$ . Denote the list (or vector) of DPs by  $V = (v_1, \dots, v_n)$ . When we aggregate, we replace each DP  $v_j$  by some ADP  $v'_j$  in  $\Omega$ , obtaining an ADP list  $V' = (v'_1, \dots, v'_n)$ . While the DPs are usually distinct, the ADPs are not, since otherwise there is no computational advantage to the aggregation. When we wish to model  $q$  distinct ADPs, we let  $\Gamma$  denote the set of  $q$  distinct ADPs, say  $\Gamma = \{\gamma_1, \dots, \gamma_q\}$ . We use the former (latter) ADP notation when the correspondence between DPs and ADPs is (is not) of interest. Usually we have  $q \ll n$ .

For any positive integer  $p$ , let  $S = \{s_k, \dots, s_p\}$  denote any  $p$ -server, the set of locations of the  $p$  servers,  $S \subset \Omega$ . (This symbol  $p$  is a different symbol from the one defining the  $\ell_p$ -distance.) Denote the location model with the given original DPs by  $f(S:V)$ , and the one with the aggregate DPs by  $f(S:V')$ . The notation  $f(S:V)$  and  $f(S:V')$  captures a key idea that *an aggregation is a replacement of  $V$  by  $V'$ , with the entries of  $V'$  not all distinct*.

For the large class of location models with similar or indistinguishable servers, with only the closest one to each DP assumed to serve the DP, for any  $p$ -server  $S \subset \Omega$  and DP  $v \in \Omega$  we denote by  $D(S,v) \equiv \min\{d(s_k,v) : k = 1, \dots, p\}$  the distance between  $v$  and a closest element in  $S$ . We then define the closest-distance vectors  $D(S,V) \equiv (D(S,v_j) : j = 1, \dots, n)$ ,  $D(S,V') \equiv (D(S,v'_j) : j = 1, \dots, n) \in \mathbb{R}_+^n$ . Suppose  $g$  is some ‘‘costing’’ function with domain  $\mathbb{R}_+^n$  attaching a cost to  $D(S,V)$  and  $D(S,V')$ . This gives original and approximating location models  $f(S:V) \equiv g(D(S,V))$  and  $f(S:V') \equiv g(D(S,V'))$ , respectively. Important and perhaps best-known instances of  $g$  are the  $p$ -median and  $p$ -center costing functions,  $g(U) = w_1 u_1 + \dots + w_n u_n$ , and  $g(U) = \max\{w_1 u_1, \dots, w_n u_n\}$  respectively; the  $w_j$  are positive constants, often called weights, and may be proportional to the number of trips between servers and DPs. Thus  $f(S:V)$  is either the  $p$ -median model,  $w_1 D(S,v_1) + \dots + w_n D(S,v_n)$ ,

or the  $p$ -center model,  $\max\{w_1 D(S, v_1), \dots, w_n D(S, v_n)\}$ . These models originate from Hakimi (1965) (each is called unweighted if all  $w_j = 1: j = 1, \dots, n$ ). They are perhaps the two best-known models in location theory. The covering model, a model related to the center model, will be described later in this chapter. Subsequently, we refer to the  $p$ -center,  $p$ -median, and covering location model as PCM, PMM, and CLM respectively. These models are NP-hard to minimize (Kariv and Hakimi 1979; Megiddo and Supowit 1984).

Consider several aggregation examples which serve to illustrate our notation and basic aggregation ideas. Let  $J = \{1, \dots, n\}$  denote the set of all DP indices. We suppose, for these examples, that the DPs will be aggregated into two postal code area centroids. Let  $J_i$  denote the subset of indices of the DPs in postal area  $i = 1, 2$ . Let  $\gamma_i$  denote the centroid of postal area  $i = 1, 2$ . Clearly, the  $J_i$  form a partition of  $J$ . To aggregate the DPs in the postal code areas into the centroids we replace each  $v_j$  with  $j \in J_i$ , by  $\gamma_i$  for  $i = 1, 2$ . Thus  $v_j' = \gamma_i$  for  $j \in J_i$  and  $i = 1, 2$ . Hence  $V'$  is now the  $n$ -vector of ADPs, and  $\Gamma = \{\gamma_1, \gamma_2\}$  is the ADP set.

**Example Aggregation 1, PMM**

$$f(S : V) = \Sigma \{w_j D(S, v_j) : j \in J\}.$$

Let  $\omega_1 = \Sigma\{w_j: j \in J_1\}$ ,  $\omega_2 = \Sigma\{w_j: j \in J_2\}$ . We then have  $f(S:V') = \Sigma \{w_j D(S, v_j') : j \in J\} = \Sigma\{w_j D(S, \gamma_1) : j \in J_1\} + \Sigma\{w_j D(S, \gamma_2) : j \in J_2\} = \omega_1 D(S, \gamma_1) + \omega_2 D(S, \gamma_2)$ .

This example illustrates how aggregation error can occur. If only  $p$ -servers are of interest (with  $p \geq 2$ ), then taking  $S$  to be  $\{\gamma_1, \gamma_2\}$  minimizes  $f(S:V')$  with minimal value of 0, giving a useless underestimation of  $\min\{f(S:V):S\}$ .

If there is only one server,  $S = \{s\}$ , and the  $\ell_p$ -distance is used, then it is known that this 1-median under-approximation is valid for all  $s$ . Letting  $\omega = \Sigma\{w_j: j \in J\}$ , and  $\gamma = \Sigma\{(w_j/\omega) v_j: j \in J\}$  be the centroid of the DPs, so that  $f(s:V') = \omega \|s - \gamma\|_p$ , it is known (Francis and White 1974) that  $f(s:V) \geq f(s:V')$  for all  $s$ . This is an important reason why underestimation can occur for PMM aggregation when few centroid ADPs are used. It is also known that for  $\ell_p$  distances (Plastria 2001) the difference  $f(s:V) - f(s:V')$  goes to zero as  $s$  gets farther from  $\gamma$  along an infinite ray with one end point at  $\gamma$ . There are good theoretical reasons due to self-canceling error (Plastria 2000, 2001; Francis et al. 2003) for using centroids as ADPs for the PMM, but none that we know of for the PCM and CLM. Indeed, better choices than centroids are available for the latter two models.

**Example Aggregation 2, PCM**

$$f(S : V) = \max \{w_j D(S, v_j) : j \in J\}.$$

Let  $w_1^+ = \max\{w_j: j \in J_1\}$ ,  $w_2^+ = \max\{w_j: j \in J_2\}$ . We then have  $f(S:V') = \max\{w_j D(S, v_j') : j \in J\} = \max\{\max\{w_j D(S, v_j') : j \in J_1\}, \max\{w_j D(S, v_j') : j \in J_2\}\} = \max\{\max\{w_j D(S, \gamma_1) : j \in J_1\}, \max\{w_j D(S, \gamma_2) : j \in J_2\}\} = \max\{w_1^+ D(S, \gamma_1), w_2^+ D(S, \gamma_2)\}$ . Again, if only  $p$ -servers ( $p \geq 2$ ) are of interest, then taking

$S$  to be  $\{\gamma_1, \gamma_2\}$  minimizes  $f(S:V')$  with minimal value of 0, giving an underestimate of  $f(S:V)$ .

**Example Aggregation 3, CLM** Minimize  $|S|$  subject to  $D(S, v_j) \leq r_j, j \in J, S \subset \Omega$ , where  $r_j$  is a “covering radius” associated with  $v_j$ . All but two covering constraints for the aggregated model are redundant. Define  $\rho_1 = \min\{r_j: j \in J_1\}$ ,  $\rho_2 = \min\{r_j: j \in J_2\}$ . Thus, the aggregated model has constraints  $D(S, \gamma_1) \leq \rho_1, D(S, \gamma_2) \leq \rho_2, S \subset \Omega$ . This means it takes at most two servers to solve the aggregated model. CLMs and PCMs are known to be closely related (Kolen and Tamir 1990). We shall see that aggregation results developed for one model often also apply to the other.

These examples of models illustrate two equivalent approaches for representing  $n$  DPs with an aggregation of  $q$  ADPs. Either we have a partition of the DP index set  $J$  into  $q$  sets  $J_1, \dots, J_q$  with one ADP per set, or for each  $v_j$  there is a replacing ADP  $v_j'$ , with each  $v_j'$  in the set  $\Gamma$  of  $q$  distinct ADPs. In either case, three aggregation decisions (Francis et al. 1999) must be made: (1) the number of ADPs, (2) the location of ADPs, (3) the replacement rule: for each  $v_j$ , what is  $v_j'$ ? The (reasonable) replacement rule often used is to replace each DP by a closest ADP. Further, for the aggregation to be computationally useful we require the number of ADPs,  $q$ , to be less (usually much less) than the number of DPs,  $n$ ; also it is reasonable to have  $p \ll q$ . The authors note that versions of these three aggregation decisions occur in location modeling. Hence results in location theory help in doing DP aggregation, so DP aggregation is a sort of “second-order” location problem to solve prior to solving the original or “first-order” problem.

These three examples may suggest that as more ADPs are used the aggregation error decreases – ideally, if we could use  $q = n$  ADPs, we have no aggregation error at all. In fact there are classes of location models where the law of diminishing returns applies: aggregation error decreases at a decreasing rate as  $q$  increases (Francis et al. 2004b). Thus a very small value of  $q$  may cause a very high aggregation error, while a large value of  $q$  might give little less error than an appreciably smaller value of  $q$ .

### 18.3 Case Study

This section is based on the work by Dekle et al. (2005), where supplemental information may be found. We refer to the authors of this study as the “team”.

FEMA is an acronym for Federal Emergency Management Agency, a national U.S. agency that deals with disasters such as fires, floods, hurricanes, tornadoes, and terrorist attacks. This work stems from a FEMA request to all counties in Florida to identify possible locations for disaster recovery centers (DRCs). FEMA describes a DRC as “a facility established in or nearby the community affected by the disaster, where people can meet face-to-face with representatives from Federal, State, local and volunteer agencies to obtain assistance.” For the county this study deals with, Alachua County, FEMA required the identification of at least three DRCs, which

could be called upon at very short notice for use in a local disaster. Alachua County had a population of about 219,000 at the time of the study. The east-west and north-south dimensions of the county are about 32 and 30 miles (51.5, 48.3 km) respectively; the land area is about 874 square miles (2266 km<sup>2</sup>).

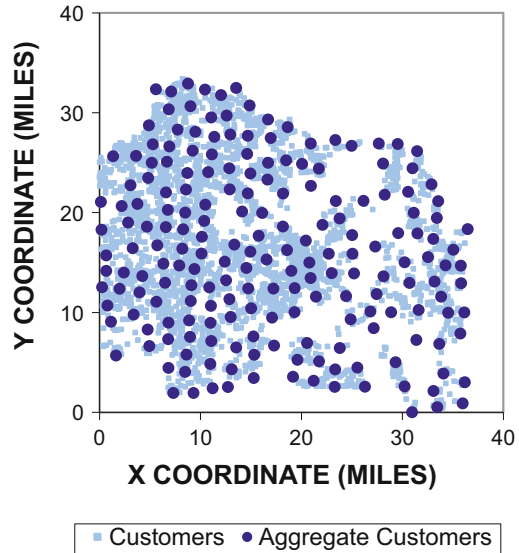
FEMA provided seven DRC requirements/evaluation criteria. The County accepted all of these requirements, but added four more, including that the proposed DRC locations should be buildings allowing reasonable travel distances to them by potential users. This criterion was the most challenging to satisfy, and led to the principal objective of the study. The team spent a substantial effort discussing with their Alachua County sponsor possible principal objectives for the study; eventually they agreed upon the following idealized one: minimize the total number of DRCs needed, subject to each county resident being within a specified distance  $r$  (called the radius) of a closest DRC. Thus if  $B(s,r)$  denotes the set of all points in the plane whose distance from a given point  $s$  is at most  $r$ , a requirement meaning that each county resident location must be in at least one  $B(s,r)$  for some DRC location  $s$ ; that is, each resident point in the county must be “covered” by at least one  $B(s,r)$  for some DRC location  $s$ . Hence the location requirement specifies a “covering” problem (see Chap. 5). It was the belief of the team (eventually confirmed) that if they could solve this idealized problem meeting the location requirement, then they could find nearby locations that would meet all the other requirements.

A natural and important question was how to measure distances between points. Ideally, shortest path distances on the existing road network would have been used, but these were unavailable due to the very limited study budget. Since the county had a largely rectilinear/right-angle road network, the team, with the agreement of its sponsor, settled on the use of rectilinear distances: for any planar points  $s = (s_1, s_2)$ ,  $t = (t_1, t_2)$ ,  $d(s,t) = |s_1 - t_1| + |s_2 - t_2|$  defines the rectilinear distance between  $s$  and  $t$ .

We refer to resident locations as “demand points”, abbreviated as DPs. For any real aggregation location problem, obtaining and dealing with DP data will probably be a major part of the problem-solving effort. Interaction with the county property appraiser’s office elicited the information that principal DP data sources could be obtained from GIS data available in a library, and from the county property appraiser’s office. The county DP data was arranged by “parcels” of land. There were about 6600 parcels, and for each parcel the following information was known:  $x$  and  $y$  coordinates of the parcel center, the total heated square footage of the parcel buildings, and whether parcel buildings were residential or commercial. The parcel locations were used as residential location/DPs, and as possible DRC sites. As many as 3900 of the parcels seemed possibly usable for DRC sites, as they had public or commercial buildings whose total usable footage exceeded 2000 ft<sup>2</sup>. Figure 18.1 shows a plot of all the DPs, as well as the aggregated DPs (yet to be discussed).

Covering models are discussed in Chap. 5; they provide a way to compute, for a specified covering radius  $r$ , a minimal set of locations, say  $S = \{s_1, \dots, s_k\}$ , so that each DP is contained in at least one  $B(s_i, r)$ . To formulate the covering problem using all the available data as an integer program model would give a constraint matrix with about 6600 rows and 3900 columns. The size of this model was beyond the resources

**Fig. 18.1** Plot of demand points and aggregate demand points



available to the team to deal with. The covering algorithm readily available to the team was one in Excel, which could deal with at most 200 variables/columns. The need to somehow aggregate the DP data and the potential site data thus became quite evident.

In a later section we discuss a useful error bound for covering location problems,

$$\max \{d(v_j, v_j') : j = 1, \dots, n\},$$

where  $v_j$  is the location of DP  $j$ , and  $v_j'$  is the ADP that replaces  $v_j$ ; the  $v_j$  are distinct while the  $v_j'$  are not. Choosing the  $v_j'$  to keep this error bound small keeps the covering error small. Note, if there are  $n$  distinct  $v_j'$ , that  $\max\{d(v_j, v_j') : j = 1, \dots, n\}$  may be viewed as the objective function of an  $n$ -center problem with DPs  $v_j$  and facility locations defined by the  $v_j'$ . This observation indicated that it would be reasonable to modify some  $p$ -center algorithm to locate the ADPs. As discussed in Dekle et al. (2005), the team used a variation of a Dyer and Frieze (1985) pick-the-farthest (PTF) algorithm to pick the ADPs. Possible center locations were also similarly aggregated. Figure 18.1 illustrates that the algorithm chooses well-dispersed locations. A number of runs of the PTF algorithm were made, and finally solutions were chosen that reduced the number of DPs from 6600 to 198 and the number of potential DRC sites from 3900 to 162.

The team's version of the Dyer-Frieze algorithm works as follows. First, an arbitrary DP is chosen as an ADP. Next, a DP whose closest-distance to an ADP is farthest is then chosen to be an ADP. Continuing, at each iteration a DP is chosen as an ADP whose closest-distance to the collection of ADPs is farthest. This process continues until the closest-distance of every remaining DP to the collection of ADPs

**Table 18.1** (a) Shows how some DRC performance measures changed with various  $r$  values for the idealized stage 1; (b) Shows similar results for the actual stage 2 results

	a			b		
	Idealized			Actual		
Travel limit $r$ (miles)	10	15	20	10	15	20
Maximum travel distance (miles)	10.9	15.1	20.3	14.0	25.8	26.94
Average travel distance (miles)	4.9	9.1	7.6	4.86	6.76	7.36
% Parcels covered	99.78	99.96	99.92	97.7	89.8	97.4
Average covering violation (miles)	0.184	0.84	0.184	1.05	2.80	2.55

is no more than a “control parameter”  $b$ . This parameter may be adjusted to provide a computationally manageable number of ADPs. Dyer and Frieze give a low-order implementation of this approach.

Subsequently, the covering location model is solved using the ADPs as DPs; the model formulation guarantees that each ADP will be within the radius  $r$  of at least one center. However, original DPs not chosen as ADPs may possibly not be within such a radius  $r$ ; supposing that  $v$  is any such unchosen DP, the algorithm guarantees that some ADP, say  $v'$ , satisfies  $d(v', v) \leq b$ . Thus for any center  $s$  that covers  $v'$ ,  $d(s, v) \leq d(s, v') + d(v', v) \leq r + b$ . Hence if  $b$  can be kept small then the uncovered DPs will be nearly covered, as was true in this application (see Table 18.1).

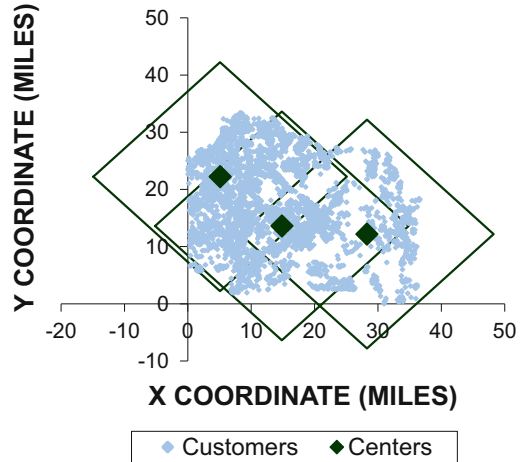
Note that  $\max\{d(v_i, v'_i) : \text{all unchosen DPs } v_j\} \leq b$  when the algorithm terminates, so keeping  $b$  small guarantees a small aggregation error. Aggregation error is discussed in the next section.

Once the DPs were aggregated and the potential DRC sites were also similarly aggregated, the covering location problem could be solved. We call the covering location problem the idealized problem, while we call the one that considers all 11 criteria the actual problem. The team solved the idealized problem first, and then sought good solutions to the actual problem that were “close” to those of the idealized problem. This approach greatly simplified the problem and worked acceptably.

Because of initial uncertainty about an appropriate value of  $r$ , the greatest distance any resident should need to travel to a closest DRC, it was decided to treat  $r$  as a parameter of the study, try various  $r$  values, and then evaluate the resultant solutions. The team eventually chose  $r$  values of 10, 15 and 20 miles (16.1, 24.2 and 32.3 km respectively). By solving the idealized covering model with these three  $r$  values, solutions were found requiring 8, 4 and 3 DRCs respectively; see Fig. 18.2 for the case of three DRCs; note Fig. 18.2 illustrates three  $B(s, 20)$  regions. The team then proceeded to solve the actual problem by finding potential DRC locations near the idealized solutions which would meet the other evaluation requirements. To aid in this effort, they and the sponsor developed a score card, much like a grade card, on which they could score each potential location considered; most of the buildings considered were schools, churches, recreation centers, or government buildings. Table 18.1a illustrates some DRC performance measures for the solutions to the idealized problem. Discrepancies between Table 18.1a performance measures and



**Fig. 18.2** The set of points within 20 miles of three Disaster Recovery Centers (DRCs)



the three different radius measures are due to aggregation effects, and can be seen to be quite small. Table 18.1b shows performance measures for the actual problem. There are some bigger discrepancies than in Table 18.1a, but these locations scored well on all the other criteria. Also it was recognized that the proper choice of a radius value  $r$  was somewhat subjective.

A number of modeling insights were gained in the course of this study, including the following. (1) Sponsors may not have a principal objective. (2) The choice of a model may be somewhat subjective. (3) Getting and working with all the data can be most of the work in an aggregation/location study. (4) Data aggregation can be essential and helpful. (5) The covering location model solutions were easy to explain to the sponsor, in part due to the figures. (6) The well-dispersed locations of the covering model also had political and geographic redundancy advantages.

The three-location solution to the actual location model for  $r = 20$ , which covered 97.4% of the parcels, was accepted by the sponsor. The following is a quote from a letter the sponsor provided to the team.

The Florida Division of Emergency Management has requested that all county emergency management offices provide at least three sites preidentified as potential DRCs. With completion of this project, Alachua County is now able to comply with this request . . .

Overall, this was an outstanding project which has provided the Office of Emergency Management with tangible results. When DRCs must be opened in the future, it will be based upon careful research and problem solving rather than guesses on which locations would be best.

In closing, we remark that this approach easily generalizes to covering problems using network distances, given adequate network data. The approach worked well, and controls the covering error. We recommend its use for aggregating covering location problems, as well as unweighted  $p$ -center problems.

### 18.4 Aggregation Error Measures

While there can be other types of error in location models, the one we focus on is *demand point aggregation error*, which result from replacing DPs by ADPs. Thus, instead of actual distances we obtain approximating ones. The use of these approximating ADPs creates error. It is thus important for the location modeler who does the aggregation to be aware of the aggregation error being created. The modeler who does DP aggregation intentionally introduces error into the model. The use of ADPs is the *cause* of the aggregation error, but there are error *effects*—including inaccurate values of the objective function and of server locations, due to using the approximating distances. It is important to consider both cause and effects in order to get the whole picture. There are a number of ways to measure error effects; further, the magnitude of aggregation effects can depend on the model structure—for the same aggregation, some models can have more error than others. What is clear, in any case, is that the way to minimize DP aggregation error is to not aggregate DPs—certainly this is what we recommend when it is feasible. *The ideal way to aggregate DP data is to not aggregate it.*

If DP data must be aggregated, then we need to consider aggregation error measures. We list and summarize ten such measures in Table 18.2. Some of these error measures are discussed in NaimiSadigh and Fallah (2009), as well as in Irawan and Salhi (2015a). All of the error measures in Table 18.2 have an ideal value of zero. One simple way to measure aggregation error is to consider *ADP-DP distances*. If these distance values are all zero then ADPs and DPs are identical, so there is no error. Later we establish a relationship between ADP-DP distances and other error measures, including the *distance difference error*. For the PMM,

**Table 18.2** Various demand point aggregation error measures for a location model  $f(S:V)$ . Ideal error measures have value zero for all  $j$  and all  $S$

No.	Error name	Error definition
1	ADP-DP distances	$d(v_j', v_j), j \in J$
2	Distance difference error	$D(S, v_j') - D(S, v_j), j \in J, \text{ all } S$
3	DP error, PMM	$e_j(S) = w_j [D(S, v_j') - D(S, v_j)], j \in J, \text{ all } S$
4	Total DP error, PMM	$e(S) = \sum \{e_j(S): j \in J\}, \text{ all } S$
5	ABC error for PMM: $J_1, \dots, J_q$ is a partition of $J = \{1, \dots, n\}; \omega_i \equiv \sum \{w_j: j \in J_i\}$ for $i = 1, \dots, q$	$abc_i(S) = \omega_i D(S, \gamma_i) - \sum \{w_j D(S, v_j): j \in J_i\}, \text{ all } S$
6	Absolute error, any location model	$ae(S) =  f(S:V') - f(S:V) , \text{ all } S$
7	Relative error, for all $S$ with $f(S:V) > 0$	$rel(S) = ae(S)/f(S:V), \text{ all } S$
8	Maximum absolute error	$mae(f', f) = \max \{ae(S): S \subset \Omega,  S  = p\}$
9	Error bound $eb$	A number $eb$ with $ae(S) \leq eb$ for all $S$
	Ratio error bounds (when $f(S:V), f(S:V') > 0$ )	$ f(S:V')/f(S:V) - 1  \leq eb/f(S:V),$ $ f(S:V)/f(S:V') - 1  \leq eb/f(S:V')$ for all $S$
10	Location error	a measure, $diff(S', S_*)$ , of the “difference” between $p$ -servers $S'$ and $S_*$

this distance difference error leads to an error we call *DP error*. Like the difference error, the DP error can be negative or positive. Still considering the PMM, note that the *total DP error*  $e(S)$  in Table 18.2 satisfies  $e(S) = f(S:V') - f(S:V)$ , the difference between the aggregated PMM and the original model. Even though no DP error is zero, the total DP error can be zero or nearly zero, since negative errors can cancel out positive errors—this is the concept of *self-canceling error*. Unfortunately self-canceling error only applies to models with an additive cost structure.

Next, consider *ABC errors* for the PMM, due to Hillsman and Rhoda (1978), pioneering aggregation researchers. Note that ABC errors are sums of the DP errors which are organized by the ADPs. Suppose we represent an aggregation by a partition of  $J = \{1, \dots, n\}$ , say  $J_1, \dots, J_q$ , such that for  $i = 1, \dots, q$ , every DP  $v_j$  with  $j \in J_i$  is aggregated into the ADP  $\gamma_i$ ; that is,  $v_j' = \gamma_i$  for  $j \in J_i$ . Thus  $\sum \{w_j D(S, v_j): j \in J_i\}$  is replaced in the aggregate model by  $\sum \{w_j D(S, \gamma_i): j \in J_i\} = \omega_i D(S, \gamma_i)$ , where  $\omega_i \equiv \sum \{w_j: j \in J_i\}$ . In the parlance of Hillsman and Rhoda, the ABC error illustrates their *Source A* error, which they define actually as  $\omega_i D(S, \gamma_i)$ . Using  $\omega_i D(S, \gamma_i)$  instead of  $\sum \{w_j D(S, v_j): j \in J_i\}$  is a source of error. The special case of Source A error when  $\gamma_i \in S$ , so that  $\omega_i D(S, \gamma_i) = 0$ , is their *Source B* error. If  $\omega_i D(S, \gamma_i) = 0$ , then it is useless as an estimate of  $\sum \{w_j D(S, v_j): j \in J_i\}$ . The *Source C* error is a related sort of allocation error involving closest-distance definitions. Suppose  $s_k \in S$  is closest to  $\gamma_i$ ; we might then assume that every  $v_j \in J_i$  will be closest to  $s_k$ . However, in reality, some  $v_j \in J_i$  may be closer to another element of  $S$  than  $s_k$ . In effect, we would allocate some DPs to a wrong server location that is not closest to them. Note  $abc_i(S) = \sum \{e_j(S): j \in J_i\}$  for all  $i$ , so total ABC error is  $e(S) = f(S:V') - f(S:V)$ . ABC error can be negative or positive, again resulting in possible self-cancellation effects. Hillsman and Rhoda recognize and discuss both total error and error self-cancellation.

Now consider any location model  $f(S:V)$  with  $p$ -server  $S$  and its approximation  $f(S:V')$ . A difficulty with error measures that can be negative or positive is that a smaller error (e.g.,  $-3000$ ) can be worse than a bigger error (e.g.,  $+3$ ). We can avoid this difficulty by using the (nonnegative) *absolute error*,  $ae(S) \equiv |e(S)| = |f(S:V') - f(S:V)|$  defined for all  $S$ . This measure is familiar from the calculus for measuring how well one function approximates another. Related to  $ae(S)$  is the idea of an *error bound*: a number  $eb$  for which  $ae(S) \leq eb$  for all  $S$ . An equivalent way to define an error bound, using  $f'$  and  $f$  to denote the functions  $f(S:V')$  and  $f(S:V)$  respectively, is based on the *maximum absolute error*,  $mae(f', f)$ , a number which may very well be quite difficult to compute. Any error bound is then an upper bound on the maximum absolute error. Good error bounds may be much easier to compute than the maximum absolute error. A *relative error* can be defined when  $f(S:V)$  is always positive:  $rel(S) \equiv ae(S)/f(S:V)$ , perhaps converted to percent. Depending on the model structure,  $ae(S)$  may be large but  $rel(S)$  may still be small due to the magnitude of  $f(S:V)$ . Relative error is not affected by the measurement scale chosen, whereas the preceding error measures are.

Assuming  $f(S:V) > 0$  and  $f(S:V') > 0$  for all  $S \subset \Omega$ , the relative error idea gives other equivalent ways of expressing the error bound, for all  $S \subset \Omega$ :

$$\left| \frac{f(S:V')}{f(S:V)} - 1 \right| \leq \frac{eb}{f(S:V)} \iff \left| \frac{f(S:V)}{f(S:V')} - 1 \right| \leq \frac{eb}{f(S:V')}.$$

If the model  $f(S:V)$  is on a national scale, but aggregation is done on a city/town scale (e.g.,  $eb = 10$  miles,  $f(S:V) = 500$  miles), we could have relatively small ratios  $eb/f(S:V)$  and  $eb/f(S:V')$ , in which case the model ratios would be nearly one and we would have a good aggregation. By contrast, if the model is on a city/town scale and the aggregation is also on a city/town scale, we might have a poor aggregation. *We need the aggregation scale to be substantially smaller than the model scale in order to have a good aggregation.* This is one reason that aggregation may be of more interest for problems of city/town/regional scope than those of national or international scope.

There is another way to view the use of an aggregation error bound. The error bound allows us to draw conclusions about a family of original models, instead of just one. If the actual location model is  $F(S:V)$  instead of  $f(S:V)$ , but the error bound applies to both, that is  $|f(S:V') - F(S:V)| \leq eb$  and  $|f(S:V') - f(S:V)| \leq eb$  for all  $S$ , then whatever conclusions we draw about the function  $f$  using the error bound inequality also apply to the function  $F$ . While we lose accuracy when we aggregate, we gain the ability to draw approximate conclusions about a family of original functions. As a general example of the function  $F$ , suppose instead of the DP set  $\{v_j: j \in J\}$  we have a different DP set, say  $\{b_j: j \in J\}$ , defining  $F$ , while all other model data is the same as for  $f(S:V)$ . If each DP  $b_j$  is aggregated into  $v'_j$ , then each of the functions  $F$  and  $f$  will be aggregated into the same approximating model, denoted by  $f'$ . Further, if also  $d(v_j, v'_j) = d(b_j, v'_j)$  for  $j \in J$ , then the methods we present later would provide both  $F$  and  $f'$ , and  $f$  and  $f'$ , with the same error bound. The data for  $F$  and  $f$  differ, but are sufficiently similar that the aggregation does not detect the differences.

Denote (globally) minimizing solutions to any original and approximating location models  $f(S:V)$  and  $f(S:V')$  by  $S_*$  and  $S'$  respectively. While we usually cannot expect to find  $S_*$  if we must aggregate DPs, we can still obtain some information about  $S_*$  if we know an error bound  $eb$  and  $S'$ . Geoffrion (1977) proves that if  $|f(S':V') - f(S_*:V)| \leq eb$ , then  $|f(S':V) - f(S_*:V)| \leq 2eb$ . Supposing  $f(S':V) > 0$ , we thus have  $|1 - f(S_*:V)/f(S':V)| \leq 2eb/f(S':V)$ . Hence, if  $2eb$  is small relative to  $f(S':V)$ , we may reasonably accept  $S'$  as a good substitute for  $S_*$ . We assume henceforth that we can compute  $S'$  but not  $S_*$ . Note that if we wish to use  $S'$  to approximate  $S_*$ , then it makes no sense to allow  $p \geq q$ , for then we can place a new facility at every ADP and may achieve a minimal approximating function value of  $f(S':V) = 0$ . Certainly it is desirable to have  $p \ll q$ .

**Table 18.3** Various types of optimality errors for any location model  $f(S:V)$ . Ideal error measures are zero

No.	Error name	Error definition
1	Total error at $S'$	$e(S') = f(S':V') - f(S':V)$
2	Opportunity cost error	$f(S_*:V) - f(S':V')$
3	Optimality error	$f(S_*:V) - f(S':V)$

Various authors, cited in Francis et al. (2009), have proposed different types of optimality errors which we list in Table 18.3. The first error can be computed, and indicates how well the approximating function estimates the original function at  $S'$ . For large models, the second two errors cannot be computed without knowing  $S_*$ . They can be computed for smaller models where  $S_*$  can be found without the need to aggregate, or for larger models if one *assumes* the algorithm used to solve the original problem provides  $S_*$ . Unless one can be certain that  $S_*$  is known, or that some properties of  $S_*$  are known, the latter two measures do not seem useful.

Although it is reasonable to use some measure of the difference between  $f(S:V')$  and  $f(S:V)$  to represent aggregation error, doing so results in what may well be called the *paradox of aggregation* (Francis and Lowe 1992). Often our principal reason to aggregate is because we cannot afford, computationally, to make many function evaluations of  $f(S:V)$ . We want to aggregate to make the error small; however, algorithms to do this typically require numerous function evaluations of  $f(S:V)$  and thus cannot be used for this purpose. Usually it is practical, however, to compute error measures for at least a few  $S$ , and we certainly recommend doing so whenever possible. For example, given we know  $V$  and  $V'$ , we can use a sampling approach to compute a random sample of size  $K$  of  $p$ -servers, say  $S_1, \dots, S_K$ , compute  $f(S_k:V')$  and  $f(S_k:V)$  for each sample element  $S_k$ , and then compute a sample error estimate of any error measure of interest.

*Location error* (Casillas 1987; Daskin et al. 1989) involves some comparison of the  $p$ -server locations  $S_*$  and  $S'$ . There are several difficulties with using this concept. First, if we really knew  $S_*$  we would not need to do the aggregation. Second, when  $|S_*| \geq 2$ , there appears to be no accepted way to define the difference between  $S_*$  and  $S'$ . Third (assuming we do know  $S_*$ ), the function  $f(S:V)$ , particularly if it is the PMM function, may well be relatively flat in the neighborhood of  $S_*$ , as pointed out by Erkut and Bozkaya (1999). This means we could have some  $S'$  with  $f(S':V)$  only a little larger than  $f(S_*:V)$ , but  $S'$  is “far” from  $S_*$ . Fourth,  $S'$  and  $S_*$  may not be unique global minima. Why are comparisons made between  $S'$  and  $S_*$ ? We speculate they are made in part due to unstated subjective evaluation criteria, or known but unstated supplementary evaluation criteria. As another possible example of the use of location error, we might solve the approximating model with three different levels of aggregation (numbers of ADPs), obtaining three corresponding optimal  $p$ -servers say  $S'$ ,  $S''$  and  $S'''$ . In this case, differences between successive pairs of these  $p$ -servers might be of interest; we might want to know how stable the optimal server locations are as we change the level of aggregation (Murray and Gottsegen 1997).

Subjective or unstated aggregation error criteria may well be important, but are not well-defined. Thus, two analysts can study the same DP aggregation and not agree on whether it is good or not. Further, if a subjective evaluation derives from some visual representation of DPs and ADPs, such an analysis may single out some relatively simple visual error feature that is inappropriate for the actual model structure. For example, a visual analysis could not evaluate the (computationally intense) absolute error for the PMM. Some generally accepted way to measure location error is desirable.

How should we measure the location error  $\text{diff}(S, Y)$ , the “difference” between any two  $p$ -servers  $S$  and  $Y$ ? The answer is not simple, because the numbering of the elements of  $S$  and of  $Y$  is arbitrary, and we must find a way to match up corresponding elements. Further,  $S$  and  $Y$  are not vectors, but sets. We propose the use of a method discussed by Francis and Lowe (1992). For motivation, consider the case where for each element  $s_k$  of  $S$  there is only one “nearby” element of  $Y$ , say  $y_k$ . In this case we might use either  $\max\{d(s_k, y_k): k = 1, \dots, p\}$  or  $\sum\{d(s_k, y_k): k = 1, \dots, p\}$  as  $\text{diff}(S, Y)$ . More generally, define the  $p \times p$  matrix  $C = (c_{ij})$  with  $c_{ij} = d(x_i, y_j)$ . Define an assignment (permutation matrix) to be any  $0/1$   $p \times p$  matrix  $Z = (z_{ij})$  having a single nonzero entry of one in each row, and a single nonzero entry of one in each column, and let  $P$  denote the set of all such  $p!$  assignments (permutation matrices). Define the objective function value  $v(Z)$  for every assignment  $Z \in P$  by  $v(Z) \equiv \max\{c_{ij} z_{ij}: Z \in P\}$ , so that  $v(Z)$  is the largest entry in  $C$  for which the corresponding entry in  $Z$  is one. Define  $\Delta(S, Y) = \min\{v(Z): Z \in P\}$ , so that  $\Delta(S, Y)$  is the minimal objective function value of the min-max assignment problem with cost matrix  $C$ . We propose using  $\Delta(S, Y)$  for  $\text{diff}(S, Y)$ . There are several good reasons for using  $\Delta(S, Y)$ . One reason is that it has all the properties of a distance (see Goldberg 1976): *symmetry*:  $\Delta(S, Y) = \Delta(Y, S)$ ; *nonnegativity*:  $\Delta(S, Y) \geq 0$  and  $\Delta(S, Y) = 0 \iff S = Y$ ; *triangle inequality*:  $\Delta(S, Y) \leq \Delta(S, Z) + \Delta(Z, Y)$  for any  $p$ -servers  $S, Y$  and  $Z$ . Another reason, further explored in Francis et al. (2009), is that it is related to the absolute error. (We could also use the optimal value of the conventional min-sum assignment model for  $\text{diff}(S, Y)$ . This optimal value also has all the properties of a distance, but we know of no useful relationship between it and absolute error.) We call the distance  $\Delta$  the *min-max distance*. Note, for any two  $p$ -servers  $S, Y \subset \Omega$ ,  $\Delta(S, Y) \leq \text{diam}(\Omega)$ . Further, when  $p = 1$  the min-max distance is just the usual distance,  $d(x_1, y_1)$ .

Both min-max and min-sum assignment models are well-studied and are efficiently solvable in low polynomial order for any set of real coefficients (Ahuja et al. 1993). In the assignment models we study, the coefficients typically correspond to distances between points in some geometric spaces, e.g., planar Euclidean or rectilinear cases. For these geometric models significantly more efficient algorithms have become available (Agarwal et al. 1999; Agarwal and Varadarajan 1999; Efrat et al. 2001, and Varadarajan 1998).

There are a number of relationships between the error measures of Table 18.2. These relationships, some of which may not be obvious, are a subject of discussion in Francis et al. (2009), where there are also numerical examples of many of the error measures. It also seems worth pointing out that error measures 2 through 7 of

Table 18.2 are local error measures, since they depend on  $S$ . By contrast, measures 1, 8 and 9 may be considered global error measures.

There is no general agreement on which aggregation error measure is best. Until the research community agrees on one or more error measures, progress in comparing various aggregation approaches, and in building a cumulative body of knowledge, will necessarily be limited. The lack of agreement on error measures also limits progress in trading off aggregation advantages and disadvantages. Further, because comparisons of various aggregation algorithm results should all be based on the same error measures, there is currently little point in developing a data base of DPs that can be used by the profession to test their aggregation methods. We personally recommend the uses of relative error based on absolute error and/or error bounds, together with ADP-DP distances. The bound in the inequality  $|f(S_*:V)/f(S':V)| \leq 2 eb/f(S':V)$  seems particularly promising.

An alternative to using some low computational order approach to aggregate the original demand point set, and then solving the resulting aggregated location model to optimality, is to use some low computational order metaheuristic approach (Pardalos and Resende 2002; Reeves 1993; Resende and de Sousa 2004) to approximately minimize the original, unaggregated location model. The first approach gives bounds on optimality to the original model. Remove space The second approach introduces an additional source of error, since a heuristic is used, but may possibly result in a better solution. Promising examples of this second approach include the work of Avella et al. (2012), Irawan and Salhi (2015b), and Irawan et al. (2014). Given the current state of the art, which approach is best is not known. Indeed, “best” may not even be well-defined, since there is no generally accepted measure of aggregation error.

## 18.5 Error Bounds

We have argued that an upper bound on the absolute error is among the best representations and measures of the error associated with an aggregation. We have used the symbol  $eb$  to represent this upper bound so that with  $f(S,V)$  a general location model,  $|f(S:V') - f(S:V)| \leq eb$ .

Consider now obtaining error bounds for the PMM and PCM, say  $eb_{pmm}$  and  $eb_{pcm}$ , with these two models defined in Examples 1 and 2 respectively. Both error bounds are direct consequences of the triangle inequality for shortest distances, which holds for all  $j \in J$  and all  $S \subset \Omega$ :

$$-d(v_j', v_j) \leq D(S, v_j') - D(S, v_j) \leq d(v_j', v_j) \iff |D(S, v_j') - D(S, v_j)| \leq d(v_j', v_j). \quad (18.1)$$

The  $p$ -median and the  $p$ -center models have the following error bounds respectively:

$$eb_{pmm} = \Sigma \{w_j d(v_j', v_j) : j \in J\}, eb_{pcm} = \max \{w_j d(v_j', v_j) : j \in J\}.$$

The error bounds themselves can be viewed as location models; if  $v_j'$  is the closest ADP to  $v_j$  (which is reasonable), then we have

$$eb_{pmm} = \Sigma \{w_j D(\Gamma, v_j) : j \in J\}, eb_{pcm} = \max \{w_j D(\Gamma, v_j) : j \in J\}.$$

Since it is of interest to have small error bounds when doing aggregation, we can view each of the latter two error bound expressions as a location model, and use heuristic location minimization algorithms to compute  $\Gamma$ . Thus, doing aggregation may be viewed as solving a location problem.

We remark for PMM, if  $S$  is restricted to being in a finite set of possible sites, and there are fixed site costs but the sites are not aggregated, then the site fixed costs can be added to the objective function without affecting the error bound.

Francis et al. (2009) give an extensive discussion of the use of the above error bounds for aggregation. The conditions for the PMM error bound to be tight are much stronger than for the PCM error bound to be tight, and this is reflected by better computational experience for the PCM than the PMM. However, computational experience does show that the PMM error bound is well correlated with sample absolute error measures, and that it makes sense to locate ADPs so as to keep the PMM error bound small.

Another location problem of interest is the covering location model, defined by Example 3. Since  $D(S, v_j) \leq r_j$  is equivalent to  $D(S, v_j)/r_j \leq 1$ , from (18.1) we obtain.

$$| D(S, v_j')/r_j - D(S, v_j)/r_j | \leq d(v_j', v_j)/r_j, \text{ for all } j \in J \text{ and all } S \subset \Omega. \tag{18.2}$$

Thus we obtain  $n$  error bounds, one for each original constraint. Clearly, it makes sense to aggregate so as to keep these error bounds small.

Let us now build on (18.2), the basic error bound idea for constraints. Generally, we have location constraints of the form  $f_j(S) \leq r_j, j \in J, S \subset \Omega$ . Suppose each function  $f_j(S)$  is replaced by some approximating function, say  $f_j'(S)$ , resulting in some constraints that are not distinct for the aggregated model of  $f_j'(S) \leq r_j, j \in J, S \subset \Omega$ . If we now define functions  $f(S)$  and  $f'(S)$  by  $f(S) \equiv \max\{(1/r_j) f_j(S) : j \in J\}, f'(S) \equiv \max\{(1/r_j) f_j'(S) : j \in J\}$ , then the constraints for the two models are equivalent to  $f(S) \leq 1$  and  $f'(S) \leq 1$  respectively. Hence, we can view  $f'(S)$  as an aggregated version of the function  $f(S)$ , and apply whatever function error measures are of interest. It is known (Francis et al. 2004a) for example, that if  $f_j'(S)$  and  $f_j(S)$  have error bound  $b_j (= d(v_j', v_j)/r_j$  for the CLM) for  $j \in J$ , then  $f(S)$  and  $f'(S)$  have the (unitless) error bound  $eb = \max\{b_j : j \in J\}$ . For the CLM, the resulting error bound is identical in form to that for the PCM; hence aggregation methods providing small PCM error bounds also can provide small CLM error bounds, and vice-versa.



When  $f(S)$  and  $f'(S)$  are any original and aggregated functions with some error bound  $eb$ , it follows directly that  $f'(S) \leq 1 - eb \Rightarrow f(S) \leq 1$ ;  $f(S) \leq 1 \Rightarrow f'(S) \leq 1 + eb$ . Thus, the constraint  $f'(S) \leq 1 - eb$  gives a restriction of the original constraint, while  $f'(S) \leq 1 + eb$  gives a relaxation. Each can be easier to deal with than the original constraint and may be used to compute lower and upper bounds on the optimal objective function value of the original model. Supposing  $eb \ll 1$  (which is clearly desirable), feasibility conclusions about one model thus allow us to draw feasibility or “near-feasibility” conclusions about the other model.

Following Francis et al. (2004c), Table 16.4 illustrates the use of error bounds as discussed to obtain a relaxation and restriction of the aggregated CLM as well as a relaxation and restriction of the original model.

Francis et al. (2004c) used the approach of Table 18.4. They solved to optimality a CLM with almost 70,000 original CLM constraints by solving several aggregated CLMs each with less than 1000 covering constraints. Their computational experience was usually that the minimal objective function value of the original model was underestimated when solving the approximating model without enough ADPs, which is consistent with the discussion in Sect. 18.2. The case study of Sect. 18.3 uses some of these aggregation ideas.

The error bound  $\max\{w_j d(v_j', v_j) : j \in J\}$  for the PCM and CLM for some choice of the  $w_j$  including  $w_j = 1/r_j$  is quite robust. It applies to an obnoxious facility location model (Francis et al. 2000; Erkut and Neuman 1989) and, when doubled, to a  $p$ -center hub location model (Gavriliouk 2003; Ernst et al. 2002a, b). Although most of the error bound results were developed for the case of discrete demand, Francis and Lowe (2014) study the relationship between error bounds for the discrete and the continuous demand cases.

**Table 18.4** Relaxation and restriction of both the original and aggregated covering location models assuming all  $\delta_j < r_j$

Constructing Aggregated CLM		
1	Definitions	$\gamma_1, \dots, \gamma_q$ : the $q$ distinct ADPs $\delta_j \equiv d(v_j', v_j), j \in J; \delta_j < r_j, j \in J$ $\beta_i \equiv \min\{r_j - \delta_j : v_i' = v_j\}, i = 1, \dots, q$ $\rho_i \equiv \min\{r_j + \delta_j : v_i' = v_j\}, i = 1, \dots, q$
2	Original covering constraints	$D(S, v_j) \leq r_j, j \in J, \text{ all } S$
3	Aggregate constraints	$D(S, v_j') \leq r_j, j \in J, \text{ all } S$
4	Restrictions of both original and aggregate constraints	$D(S, v_j') \leq r_j - \delta_j, j \in J, \text{ all } S \iff$ $D(S, \gamma_i) \leq \beta_i, i = 1, \dots, q, \text{ all } S$
5	Relaxations of both original and aggregate constraints	$D(S, v_j') \leq r_j + \delta_j, j \in J, \text{ all } S \iff$ $D(S, \gamma_i) \leq \rho_i, i = 1, \dots, q, \text{ all } S$

## 18.6 Conclusions

For location problems with many thousands of demand points, aggregation is often essential. This chapter has dealt with the topic of demand point aggregation for location models. We have pointed out that demand point aggregation causes error and presented some possible ways of measuring this error. Our focus has been on the concept of an error bound, an upper bound on the maximum absolute error due to aggregation. Error bounds are given for three key location models: the  $p$ -median model (PMM), the  $p$ -center model (PCM) and covering location model (CLM). We have shown that minimizing the error bounds for (PMM) or (PCM) results in a location problem. This is a concept that we have called “the paradox of aggregation”. We have also presented an application of the covering location model to a real public sector location problem in the state of Florida, and have demonstrated error bound analysis for this problem.

Difficulties in computing actual errors lead to the concept of an error bound, and this error bound can be used as a surrogate for the maximum absolute error. In fact, error bounds can be computed for many other location models since many of these models share properties with (PMM), (PCM), or (CLM). In addition, error bound analysis can be extended to more general costing functions  $g$  if  $f(S) = g(D(S, V))$  and the costing function  $g$  is subadditive and nondecreasing (SAND) (see Francis et al. 2000, 2009).

Based on our work on demand point aggregation for location modeling, we offer the following observations:

1. The work of Hillsman and Rhoda (1978) is widely recognized and influential; in particular, self-canceling error is a helpful concept for models with additive structure;
2. There is little average-case analysis of aggregation error;
3. Much more research on aggregation for the median problem has been done than for center, covering and other models;
4. Progress is definitely being made in understanding aggregation error;
5. Aggregation error bounds can be useful, particularly for center and covering models;
6. Aggregation error measures used vary greatly, and there is no agreement on how to measure error; hence it is pointless to ask which aggregation algorithm is best, since “best” is not defined.

## References

- Agarwal PK, Efrat A, Sharir M (1999) Vertical decomposition of shallow levels in 3-dimensional arrangements and its applications. *SIAM J Comput* 29:912–953
- Agarwal PK, Varadarajan KR (1999) Approximation algorithms for bipartite and nonbipartite matchings in the plane. In: 10th ACM-SIAM symposium on discrete algorithms (SODA). ACM, New York, pp 805–814

- Ahuja RK, Magnanti TL, Orlin JB (1993) Exercise 12.23. In: *Network flows: theory, algorithms, and applications*. Prentice Hall, Englewood Cliffs, p 505
- Avella P, Boccia M, Salerno S, Vasilyev I (2012) An aggregation heuristic for large scale p-median problem. *Comput Oper Res* 39:1625–1632
- Bender T, Hennes H, Kalcsics J, Melo T, Nickel S (2001) Location software and interface with GIS and supply chain management. In: Drezner Z, Hamacher H (eds) *Facility location: applications and theory*. Springer, Berlin
- Casillas PA (1987) Data aggregation and the p-median problem in continuous space. In: Ghosh A, Rushton G (eds) *Spatial analysis and location-allocation models*. Van Nostrand Reinhold Publishers, New York, pp 227–244
- Chelst KR, Schultz JP, Sanghvi N (1998) Issues and decision aids for designing branch networks. *J Retail Bank X*:5–17
- Daskin MS (2013) *Network and discrete location: models, algorithms, and applications*, 2nd edn. Wiley, Hoboken
- Daskin MS, Haghani AE, Khanal M, Malandraki C (1989) Aggregation effects in maximum covering models. *Ann Oper Res* 18:115–139
- Dekle J, Lavieri M, Martin E, Emir-Farinas H, Francis RL (2005) A Florida county locates disaster recovery centers. *Interfaces* 35:133–139
- Domich PD, Hoffman KL, Jackson RHF, McClain MA (1991) Locating tax facilities: a graphics-based microcomputer optimization model. *Manag Sci* 37:960–979
- Drezner Z (ed) (1995) *Facility location: a survey of applications and methods*. Springer, Berlin
- Drezner Z, Hamacher HW (eds) (2002) *Facility location: theory and algorithms*. Springer, Berlin
- Dyer M, Frieze A (1985) A simple heuristic for the p-center problem. *Oper Res Lett* 3:285–288
- Efrat A, Itai A, Katz MJ (2001) Geometry helps in bottleneck matching and related problems. *Algorithmica* 31:1–28
- Eiselt HA, Marianov V (2011) *Foundations of location analysis*. Springer, New York
- Erkut E, Bozkaya B (1999) Analysis of aggregation errors for the p-median problem. *Comput Oper Res* 26:1075–1096
- Erkut E, Neuman S (1989) Analytical models for locating undesirable facilities. *Eur J Oper Res* 40:275–291
- Ernst A, Hamacher HW, Jiang HW, Krishnamorthy M, Woeginger G (2002a) Uncapacitated single and multiple allocation p-hub center problems. Report. CSIRO, Melbourne
- Ernst A, Hamacher HW, Jiang HW, Krishnamorthy M, Woeginger G (2002b) Heuristic algorithms for the uncapacitated hub center single allocation problem. Report. CSIRO, Melbourne
- Francis RL, Lowe TJ (1992) On worst-case aggregation analysis for network location problems. *Ann Oper Res* 40:229–246
- Francis RL, Lowe TJ (2014) Comparative error bound theory for three location models: continuous demand vs discrete demand. *TOP* 22:144–169
- Francis RL, Lowe TJ, Rayco MB, Tamir A (2003) Exploiting self-canceling demand point aggregation error for some planar rectilinear median problems. *Nav Res Log* 50:614–637
- Francis RL, Lowe TJ, Rushton G, Rayco MB (1999) A synthesis of aggregation methods for multi-facility location problems: strategies for containing error. *Geogr Anal* 31:67–87
- Francis RL, Lowe TJ, Tamir A (2000) On aggregation error bounds for a class of location models. *Oper Res* 48:294–307
- Francis RL, Lowe TJ, Tamir A (2004a) Demand point aggregation analysis for a class of constrained location models: a penalty function approach. *IIE Trans* 36:601–609
- Francis RL, Lowe TJ, Tamir A, Emir-Farinas H (2004b) Aggregation decomposition and aggregation guidelines for a class of minimax and covering location models. *Geogr Anal* 36:332–349
- Francis RL, Lowe TJ, Tamir A, Emir-Farinas H (2004c) A framework for demand point and solution space aggregation analysis for location models. *Eur J Oper Res* 159:574–585
- Francis RL, McGinnis LF, White JA (1992) *Facility layout and location: an analytical approach*, 2nd edn. Prentice-Hall, Englewood Cliffs
- Francis RL, Rayco MB, Lowe TJ, Tamir A (2009) Aggregation error for location models: survey and analysis. *Ann Oper Res* 167:171–208

- Francis RL, White JA (1974) Facility layout and location: an analytical approach. Prentice-Hall, Englewood Cliffs, p 324. problem 7.25
- Gavriliouk EO (2003) Aggregation in hub location models. M Sc thesis, Department of Mathematics, Clemson University, Clemson, SC
- Geoffrion A (1977) Objective function approximations in mathematical programming. *Math Prog* 13:23–37
- Goldberg R (1976) *Methods of real analysis*, 2nd edn. Wiley, New York
- Hakimi SL (1965) Optimum location of switching centers and the absolute centers and medians of a graph. *Oper Res* 12:450–459
- Handler GY, Mirchandani PB (1979) *Location on networks: theory and algorithms*. MIT Press, Cambridge, NJ
- Hillsman EL, Rhoda R (1978) Errors in measuring distances from populations to service centers. *Ann Reg Sci* 12:74–88
- Hooker JN, Garfinkel RS, Chen CK (1991) Finite dominating sets for network location problems. *Oper Res* 39:100–118
- Irawan C, Salhi S (2015a) Aggregation and non aggregation techniques for large facility location problems - a survey. *Yugosl J Oper Res* 25:313–341
- Irawan C, Salhi S (2015b) Solving large p-median problems by a multistage hybrid approach using demand points aggregation and variable neighborhood search. *J Glob Optim* 63:537–554
- Irawan C, Salhi S, Scaparra M (2014) An adaptive multiphase approach for large unconditional and conditional p-median problems. *Eur J Oper Res* 237:590–605
- Kariv O, Hakimi SL (1979) An algorithmic approach to network location problems: part 1, the p-centers; part 2, the p-medians. *SIAM J Appl Math* 37:513–560
- Kolen A, Tamir A (1990) Covering problems. In: Mirchandani PB, Francis RL (eds) *Discrete location theory*. Wiley, New York, pp 263–304
- Love R, Morris J, Wesolowsky G (1988) *Facility location: models and methods*. North-Holland Publishers, Amsterdam
- Megiddo N, Supowit KJ (1984) On the complexity of some common geometric location problems. *SIAM J Comput* 13:182–196
- Mirchandani PB, Francis RL (eds) (1990) *Discrete location theory*. Wiley, New York
- Murray AT, Gottsegen JM (1997) The influence of data aggregation on the stability of p-median location model solutions. *Geogr Anal* 29:200–213
- NaimiSadigh A, Fallah H (2009) Demand point aggregation analysis for location models. Chap. 22. In: Farahani R, Hekmatfar M (eds) *Contributions to management science*. Physica-Verlag, Heidelberg
- Nickel S, Puerto J (2005) *Location theory: a unified approach*. Springer, Berlin
- Pardalos PM, Resende M (eds) (2002) *Handbook of applied optimization*. Oxford University Press, Oxford
- Plastria F (2000) New error bounds in continuous minisum location for aggregation at the gravity centre. *Stud Locat Anal* 14:101–119
- Plastria F (2001) On the choice of aggregation points for continuous p-median problems: a case for the gravity center. *TOP* 9:217–242
- Reeves C (1993) *Modern heuristic techniques for combinatorial problems*. Blackwell Scientific Press, Oxford
- Resende MGC, de Sousa JP (eds) (2004) *Metaheuristics: computer decision-making*. Kluwer Academic Press, Boston
- Sheffi Y (1985) *Urban transportation networks: equilibrium analysis with mathematical programming models*. Prentice-Hall, Englewood Cliffs, pp 14–16
- Varadarajan KR (1998) A divide and conquer algorithm for min-cost perfect matching in the plane. In: *Proceedings 38th annual IEEE symposium on foundations of computer sciences*. IEEE, Los Alamitos, pp 320–331