Annals of Operations Research, 18 (1989) 25–42

APPLICATIONS OF LOCATION MODELS

Gerard RUSHTON

Department of Geography, University of Iowa, Iowa City, IA 52242, U.S.A.

Abstract

The degree to which locational complexity and geographical complexity is represented in a location model is a critical decision that influences the quality of the application. Criteria which can be used to guide these decisions are presented and research that would better inform these decisions is described.

1. Introduction

This review of applications of location models focusses on how well analysts represent the locational complexity of the decision environment and the geographical complexity of the environment in their models and how decisions on these questions might be made so that results might be more useful in day-to-day decision-making contexts. Three kinds of location models have been developed to cope with the different kinds of questions that analysts ask about location decisions. In one literature, (spatial choice), the focus is on the locational choices made and the goal is to infer the choice rules that decision-makers used in reaching their decisions. The purpose is to develop positive theories of locational decision-making. In a second literature, (organizational decision-making), the focus is on the process of decision-making and the goal is to show who participated, what roles they play, and how did they decide to act at any particular decision-making stage. Finally, in a third literature, (location-allocation), the focus is on finding optimal locations for a decision-maker who has defined the objective criteria and wishes to find the locations that optimize the criteria.

There has been little interaction between these three literatures. Instead, each has had its separate domain of application. Yet to give effective help in most decision-contexts requires that elements of the three literatures be merged. In so far as people evaluate alternative locations and make choices, it is important that models of optimal location accurately represent these evaluations. In so far as organizations are complex decisionmaking units, it is important that models of optimal location allow the locational elements of decisions to be made in whatever contingent way is required by other, non-locational, decisions. Applications of location models are beginning to relate these literatures to one another but there are many problems to be overcome. At issue is how analysts deal with complexity. The complexity is that of an individual making a decision from among multi-attribute alternatives; of an organization making a decision in an environment where multiple goal exist; and of all of the above occurring among many combinations of locations. Separate location models exist that individually deal with each of these complexities but, not surprisingly, little progress has been made in dealing with perhaps the most common of applied location problems–the case where many individual private choices are being made, where a complex organization is involved and where many combinations of locations exist.

Papers presented at ISOLDE IV represented the judgments of researchers from many countries about the important problems in locational decision-making. They show that attempts to deal with these multiple complexities are still rare. It is no secret why this is so. In each of these three areas of locational modeling there are enough unanswered questions to keep our attention. Nevertheless, it is the lot of the applied location modeler to survive in this mine-field of unanswered questions and, usually after only a modest expenditure of resources, to say something useful about a pressing location problem where decisions must be made. In this paper I review many of the decisions that analysts make in applying a location model to the solution of a practical problem and conclude that many of these decisions are made without knowledge of the errors that are introduced. Research is needed to guide these decisions. They relate to how a given problem is represented and so will be called here "representation errors". Surprisingly little is known about the effects on the quality of solutions of the different possible decisions that can be made. Such knowledge as does exist is difficult to apply to specific cases. The paper discusses the potential for such errors to exist, what is known about each source of error, and how analysts can act to minimize the impacts of these errors on the quality of solutions reached.

2. Representing locational complexity

By locational complexity I mean the ability to identify and evaluate a large number of combinations of locations. Location-allocation models, developed in the 1960s, for the first time allowed locational complexity to be represented for geographic spaces that were realistically represented. Based on the assumption that the returns from modelling locational complexity were obviously large and significant, many application areas were soon identified. The quality of the locations selected were judged to be superior to locations selected by alternative methods.

2.1. PROBLEMS OF MEASURING LOCATIONAL EFFICIENCY

How much better are the "optimal locations" selected by a location model? Unlike 'normal' choice problems where the set of available alternatives from which choice is made usually can be identified easily, in the multiple location choice case, the composition of the choice set from which selections were made is unclear. The gain should be measured in relation to the likely objective function value if the model had not been used. In other words, the gain is a measure of what can he achieved simply by raising the quality of locational decision-making. It is not clear in many applied location studies, what assumption is being made, or standard adopted, about the location decisions that would be made if locational complexity had not been modeled. The measurement of efficiency is itself an active area of research (Charnes, Cooper and Rhodes [6], Farrell [13], Sherman [41]). We suggest that locational efficiency be defined as the performance of the actual location decisions with respect to the estimated performance of an alternative set of locations defined as a "reference set". McLafferty and Ghosh [33] define sets of locations that could have been selected by the decision-maker. Alternative assumptions about the discretion of the decision-maker to select different locations lead to the definition of alternative reference sets. It is therefore necessary to define the efficiency of any particular reference set as the performance of the best member of the set in relation to the best member from the set of all locations.

2.2. MEASURING THE PERFORMANCE OF PAST LOCATION DECISIONS

The most commonly used reference point for assessing efficiency is the ratio of the performance on a given objective function of the actual location pattern to the optimal pattern, as found by some location-allocation algorithm. There are substantial problems with this measure. What, for example, is the meaning of a report that a given location pattern is 0.94 efficient? When it is unity, the actual pattern is identical with the computed optimal pattern but it is more difficult to interpret values less than unity. In many practical circumstances the lower limit of this coefficient is substantially larger than zero and, in fact, can approach unity, even though the location decisions studied are suboptimal. This can occur, for example, when many locations are fixed and only a few location decisions are made. No matter how suboptimal these few decisions may be, relative efficiency would remain high. Even without considering constraints on location selection, the lower limit of this measure is affected by a number of characteristics of the data and problem being examined: the ratio, for example, of p (the number of service locations in the system) to m (the number of eligible or candidate sites from which p is to be selected). Where p approaches m in size, the opportunity for making a poor location selection is reduced. In such circumstances, a finding that relative locational efficiency is high is merely an artefact of the data and nothing can be imputed about the skill or purpose of the decision-maker in achieving this result. Note that the interpretation of this measure in behavioral terms, however, is based on the assumption, unrealistic in practice, that discretion in decision-making behavior allows any locations selected at some earlier time period to be abandoned without cost and for new locations to be substituted for them–also without cost.

We conclude, therefore, that the most common measure of locational efficiency has theoretical and applied problems with its use. The theoretical range in its value will vary from one problem to another, depending in an unknown way on different parameters of the problem. The behavioral possibilities for location selection implied in its use will often not exist.

2.3. THE STANDARD OF BEST PRACTICE

The argument for using "best practice" as the reference point against which to measure any benefits from modeling locational complexity is that the theoretical ideal of selecting the optimal combination of locations may be unrealistic [26]. As Farrell observed, ([13], p. 255), "it is far better to compare performances with the best actually achieved than with some unattainable ideal". Incremental location decisions can be measured against the standard of the expected productivity increases achieved in relation to the maximum that could have been achieved given the locations with the service and the places where new facilities could have been located. Any trend in these points can be interpreted as an increase or a decrease in the locational efficiency of decision-makers. Alternative reference levels can be suggested, such as the modal value, the 10th percentile or the value most recently achieved. Comparing the performance of decision-makers in one area of activity with the measure of performance derived from its own new location decisions is known elsewhere as "structural efficiency", ([13], p. 262). It measures the extent to which new location decisions keep up with the performance of those made earlier. Its use is open to the criticism that it does not reflect the extent to which the best practice in one service activity compares with the best practice in another. The moments of the distribution, if measured for several services in an area, can be used to compare the structural efficiency of different service systems. Such comparative efficiency measures could be particularly useful when different decision-making processes are followed in locating different types of services.

2.4. THE STANDARD OF ALTERNATIVE DECISION RULES

Some literature has explored the quality of alternative possible decision rules. I believe this is important because the argument for modelling locational complexity in applications of location models rests on the expected gains from use of such models. There is some reason to believe that decision-makers use simple principles to make service-location decisions, yet, neither their relative efficiency nor the circumstances that affect their performance is known.

The "principle" which implies the least directed behavior to effectively deal with locational complexity is the rule that locations are selected at random.

McLafferty and Ghosh [33] have suggested that the performance of any chosen set of locations can be measured by comparing its score on any criterion with a reference distribution of relative frequencies of scores on the criterion obtained from a random sampling of sets of locations drawn from the list of places eligible for selection by the decision-maker. If a large proportion of such sets are poorer than the set selected, it is reasonable to infer purposive behavior on the part of the decision-maker to favor the criterion in question as well as an ability to identify the patterns that perform well on the criterion. As McLafferty [32] points out, the method is particularly appropriate when the criterion of interest involves the performance of a set of locations for serving two or more sets of people whose distribution patterns are different in a study area. In such cases the optimal quality of service that either group could receive may well be different (because of their different location patterns) and quite erroneous inferences concerning the motives of decision makers have been made in analyses that have merely measured the changes in geographical accessibility to a service by different groups after new location decisions were made, (Lineberry, [28]). Larson and Stevenson [26] showed that the average travel distance to the closest service site in an optimal location pattern in an area of homogeneous demand will be 25 percent smaller than a location pattern in which facilities are randomly distributed. How random location decisions will perform depends upon the distribution of the demand to be served and the distribution of locations that can be selected. Little is known about how randomly selected location patterns typically perform in such realistic conditions.

A second principle for selecting locations is when places are selected in order of their population size. Although facilities are generally intended to serve people outside the places or communities in which they are located, decisions about their location often take place in forums where people represent their own place or community. From such a meeting may often emerge decisions favoring the larger places. It is a principle of a median point, of course, that if within a service area there exists a node with a demand greater than one-half of the total demand of the service area, it will be the median location, no matter where it is located in the local market area. If p places in a larger area, for example, contain more than half the demand of the area, the probability is high that they will be the solution to the p-median problem in the area. Other principles can be suggested.

2.5. LOCATION OF DISPENSARIES IN NIGERIA, 1979-1982

We studied the locations selected by the government of Nigeria for Dispensaries in a section of Ogun State in Nigeria, (estimated population 1,015,725), (Ayeni, Rushton and McNulty [2]). In the three years following the formation of a new civilian government in 1979, 20 dispensaries were added to the 43 that already existed in 1979. It seemed likely to us that the government would regard any question about past locational decisions as moot, but that it would be interested in finding the locations which, if added to the 43 dispensaries that existed in 1979, would provide the largest possible increase in utilization. We used a function that estimated for each location, per capita use of dispensaries in each candidate place as a function of distance from the closest dispensary, Stock [42]. Using the Teitz and Bart [44] substitution heuristic algorithm, and employing the unified linear model of Hillsman [19], we found the locations for 20 dispensaries to add to the 43 existing dispensaries in the area in 1979 which would have most increased aggregate utilization. Following McLafferty and Ghosh [33], we computed the expected utilization for fifty sets of locations, each consisting of 20 randomly selected locations from places with populations larger than 1000 that did not already have a dispensary (175 locations) and the 43 existing dispensary locations. We found that every single one of the 50 patterns would have had a larger utilization than the existing pattern of 63 dispensary locations in 1982. We found that when we chose 20 places to add to the existing 43 by selecting them at random, with probabilities proportional to their population, from the 175 places without dispensaries which had populations greater than 1000, the estimated utilization would have been 2.8 percent larger than the actual set. The third reference standard, selecting the 20 largest places without dispensaries, would have had a utilization rate 6.8 percent larger than the government's chosen locations.

We concluded, therefore, that current methods of evaluating the performance of multiple location decisions can be improved by relating measures of performance to reference points that are selected as reasonable decision-making rules that could have been applied in the problem context (Park [35]). Such rules-of-thumb can, in some circumstances therefore, work very well as an alternative to using location-allocation algorithms to select locations to provide services to a dispersed population. The gain from modeling locational complexity should be assessed in comparison with the results if one or more of these simple decision-rules had been used.

2.6. IMPORTANCE OF OPTIMIZING LOCATION SELECTION

A related question in modeling locational complexity is the significance of any given measure of difference in performance between any two patterns. This can only be answered by reference to the substance of the problem. Decision-makers will often want to estimate how many fewer facilities optimally located could have produced the same amount and quality of service or how much greater outputs could have been produced by the same number of facilities optimally located; (for a related efficiency decision see Sherman [41], p. 922). Location-allocation models were used a great deal, for example, for finding optimal locations of emergency facilities because, it was argued, the increased performance was significant-lives saved, in the case of emergency medical services; fire loss, in the case of fire equipment deployment; crime, in the case of policy patrol route optimization; ... the list could be long (see, for example, Swersey and Ignall [43]).

A common decision that the applications analyst must make is whether to sacrifice the modelling of locational complexity in order to increase the ability to deal with other kinds of complexity. Brandeau and Larson [4] discuss the difficulties of devising a realistic analytical model of an urban ambulance deployment system and after concluding that five criterion are important in ambulance location deployment, they decided to sacrifice the formal modeling of locational complexity in favor of representing the five criterion for any solution and allowing decision-makers to search for solutions by discussing the relative merits of alternative configurations. The decision-makers were being encouraged to develop their preference trade-offs for the five criteria by discussing the relative merits of alternative solutions that they generated. In this case, process complexity consisted of reconciling two primary objectives and three secondary objectives; Brandeau and Larson 1986, pp. 132–133;

"Primary performance objectives:

(i) to reduce citywide average response times to emergency calls;

(ii) to reduce citywide inequities in ambulance availability.

Secondary performance goals deemed important were:

(iii) to minimize ambulance workload imbalance;

(iv) to minimize the fraction of cells handled by backup units;

(v) to minimize the fraction of dispatches which are inter-district."

In this case the key decision the analysts made was not to use any location-allocation algorithm to search for the optimum locations for ambulance deployment. Rather, they used human intuition, supported by an evaluation routine that provided proof that the subjectively determined relations of the ambulances performed better than the current deployment pattern in terms of the five performance objectives and goals. Subjectivity in re-locating the units extended so far as to introduce new geographical data units–"dummy atoms"–into the data set to allow ambulance locations in locations that were not centroids of census tracts, but which were thought to be desirable locations for ambulances. Although no formal explanation of the subjective choice of locations appears in their published work, one can surmise that an unstated goal was to convince decision-makers that model solutions were better than the existing system and that they should make the changes indicated by the model. By involving decision-makers in the solution process and asking them to suggest re-locations for evaluation, they raised the likelihood that model results would be implemented. In this case, several existing ambulances were moved to new locations suggested by the model runs and the authors concluded that since at least one more ambulance would have been needed to achieve the same service improvements if the existing vehicles had not been re-deployed, the benefits of the model were equal to the costs of acquiring and operating an ambulance, estimated to be $150,000 per year.

3. Representing the geographical environment

Large errors can be caused by the method of representing the geography of an area in a location model. Key information are distances, travel times, or costs between places; measured of demand or need for service; the level of spatial aggregation at which data is represented; and the effect of the boundary of a study area on model results.

3.1. DISTANCE ESTIMATION

There are two issues in representing travel distances: the estimation of total (or average) distances in a system from knowledge of certain macro variables and the estimation of specific inter-point distances from secondary measures.

Kolesar [23] and others have shown that expected distances in optimally located configurations of facilities can be estimated from knowledge of the area of a region and the number of facilities. For travel times actually experienced in an area, equations calibrated from sample data on response times in operating systems have been shown to provide accurate estimates of response times for new systems (Kolesar [23], p. 190).

$$\text{maximum response time} = K/\sqrt{(\text{number of facilities})}$$

and, more generally,

$$E(D) = k\sqrt{(A/N)}$$

$$D = \alpha_i (N \exp b_i)\sqrt{A}$$

where A is the area to be served, N is the number of facilities.Kolesar shows that when needs for service occur homogeneously over area and when the number of facilities is such that the area can be divided into equal sized square blocks, then k is equal to $\sqrt{2/3} = 0.4714$. Earlier, Leamer [27] had established the same result but had cautioned against its application in areas when the assumption of homogeneous distribution of demand was not met. Even when, however, the assumptions of homogeneity of demand for service is not valid, k can be estimated by simulation of patterns of demand for service in relation to any given pattern of facilities. After a number of simulations, Kolesar found that ([27], p. 195) "the square root law holds as a good approximation under the assumption of straight-line travel even when hazards, alarms and unit locations are not homogeneous, when distances are computed with a complicated function, and when units may often be unavailable to respond." This law can be used to estimate the size of the likely response distances if any given number of facilities are deployed. It can also be used to estimate the size of the possible savings that might occur if an existing number of facilities were to be re-located. This is valuable knowledge that can be used prior to embarking on analyses to improve locational configurations of facilities, and can provide decision-makers with rough estimates on the degree to which response times can be improved by facility relocation. Although these rules are often helpful, their robustness for the range of conditions in which they can be used is not known.

Some progress has been made in making estimates of distances between places that are more accurate than the commonly assumed Euclidean or Manhattan distance metrics, (Love et al., [30], Ch. 10). Parameters of these empirical models of travel distances differ substantially over the various data sets from which the models were calibrated and Love et al. ([30], p. 262) have recently advised analysts not to generalize from these studies. Instead, calibration of models in study regions from sample spatial interaction data is still advised (Eaton et al. [12]). There is clearly scope for research that would determine the functional relationship between optimal parameter values and other measurable characteristics of regions.

3.2. SPATIAL AGGREGATION EFFECTS

The effect of employing discrete spatial structures to represent data that is distributed continuously is largely unknown, though recent research has shown that the effects on the validity of results from location-allocation models can be considerable. In most cases the decision to use a particular data structure is a matter of convenience and most analysts do not discuss the potential consequences of the data system they use or the alternatives they rejected. Optimum locations of facilities depend on data structure in three respects. (1) the actual locations identified as optimal may be suboptimal for the disaggregated data. The difference cannot be assumed to be a small local distance deviation from some unknown optimal site for each location. Instead, the true optimal location pattern may well be some quite different pattern. (2) the true objective function value for the optimal location pattern found from the aggregated data may be different from the objective function value for the location pattern found from the disaggregated data–this is the real cost to the decision-maker of accepting results from analysis of the aggregate data rather than equivalent results from the disaggregated data. It is not a "measurement error". (3) the value of the objective function for the locations identified as optimal in the aggregate data may be in error so that decisions based on it may not be valid–this is a "measurement error".

In some cases, systematic bias in the objective function value can be estimated but in many cases the direction of the bias is unknown. The problem was first studied by Gould, Nordbeck and Rystedt [17] who investigated the robustness of solutions by conducting simulations with different degrees of data aggregation and with demand estimates randomly distributed. They reported robust results for their case study and dismissed the problem as unimportant in practical work. In a more theoretical way the question has been studied by Casillas [5], Current and Schilling [9], Goodchild [15], and Hillsman and Rhoda [20]. Goodchild showed that different data aggregations of the same basic, disaggregated, data set could give optimal locational patterns that were geographically quite different even though the values for the objective function were quite close. He concluded that "aggregation tends to produce much more dramatic effects on location than on the values of the objective function". The generality of this conclusion, however, can be questioned in that the data base on which his simulations were conducted was a hypothetical distribution of equal-sized zones with uniform weights. For such a spatial distribution, the optimal locations are indeterminate in that several symmetrical location patterns will have identical values of the objective function. Casillas [5] in a series of simulations concluded that optimal location patterns remain stable across all levels of aggregation but that estimates of travel distance or cost from analyses of aggregated data had large errors. Clearly, optimum locations must be sensitive, to an important degree, at some level of spatial aggregation of data. We do not know how to identify this level in advance for any given application.

If errors in finding optimal locations and errors in measuring the objective function value arise because of aggregation of data that is continuously distributed, how can these errors be avoided? Three suggestions have been made. The first is that data be disaggregated. Unfortunately, the advantages of data disaggregation are often offset by the disadvantages of handling larger data sets and the algorithmic compromises that are often necessary as the size of data sets increase. Because the optimal degree of data aggregation is not known, errors caused by spatial data aggregation are not entirely removed by disaggregation. Indeed, other errors are introduced when estimation methods are used to predict values of data for subareas, "target zones", from larger areas. Lam [24] reviews these methods and concludes that approaches that involve the overlaying of grid cells and interpolating values to each control point often produce large errors. These methods are non-volume preserving; that is, aggregation of data from the new, smaller areas are not guaranteed to sum to the original data values for the larger areas. A second class of methods are volume preserving. One method, (Crackel [8]), uses the overlay method in which the density distribution within a larger area is assumed to be homogeneous throughout the area, the proportion of the larger area within the smaller area is measured, and the value for the smaller area is estimated from its proportions of each of the larger areas times their data value. Although volume preserving, these methods introduce error through the assumption of homogeneity of distribution within the larger area. A second method (Tobler [45]), assumes the existence of a smooth density function fit to the centroids of the larger zones; the volume preserving condition is enforced by an iterative algorithm that increments or decrements the densities within individual zones. Applications of either of these methods of areal interpolation in applications of optimizing location models are rare.

The second suggestion is that two of the three sources of error caused by spatial aggregation of data (Hillsman and Rhoda [20]) be removed prior to computing optimal location patterns. Current and Schilling [9] showed that the weighted distances from the disaggregated spatial data units that comprise each aggregated spatial unit (ASU) can be measured to the center of all other ASUs to remove Type A errors and of all disaggregated units to the center of their own ASU to remove Type B errors of aggregation. This, they argue, is a practice that should be commonly used in applied location studies. When the analyst is presented with aggregated data, it requires judgment to determine whether the errors caused by the act of spatially disaggregating the data will be smaller than the errors caused by aggregation. No general conclusions are possible since the magnitude of both types of error will depend on the degree of aggregation and the nature of the spatial distributions that have been aggregated. Typically, data sets that contain large diversity in geographical sizes of units (such as U.S. census tracts) and large variations of data within data units (such as counties) are likely to cause large errors. Disaggregating them will often be necessary to reduce error. In summary, the worst case occurs when the ratio of p/m is large, and where data units are irregular in size. Current and Schilling [9] reported that distance measures in their worst case ($p/m = 0.33$) underestimated the true travel cost by 44.6 percent. Since this error is related to p/m, it also follows that as p increases in any given analysis, the true travel cost savings which result from increasing the number of facilities, is overestimated (Current and Shilling [9], p. 107).

The reason why analysts aggregate data is to reduce computation time and data storage requirement. Current and Schilling ([9], p. 108) reported savings in CPU time used of 20 to 1 as they reduced a 681×70 data set to 30×30.

The third suggestion for dealing with spatial aggregation error in location models is more radical. It is that the problem be expressed and solved in continuous space. There is a large literature on the solution of location problems on the plane (Beaumont [3]) but in almost all cases, although facilities are located continuously in the plane, the distribution of demand is represented as a set of discrete locations. Consequently, most of the aggregation problems discussed above remain. Applications are rare where facility locations are found in continuous space and need for service is represented as a continuously defined density distribution. Leamer [27] solved the p-median problem for this case for uniformly distributed demand while Rushton [37] developed an algorithm for the case where equal-size constraints existed on the populations served by each facility–the transportation-location problem, (see Goodchild and Massam [16] and Cooper [7] for the discrete spatial demand data case of the same problem). A basis for the continuous representation of interpoint distances, times or costs, in geographic space is provided by Angel and Hyman [1]. Building on this work, Mayhew [31] proposed a method involving the drawing of tesselations of hexagons on a population density surface and then transforming them to the physical surface of the city as distorted hexagons. These "districts" then have the property of having approximately equal populations. For the problem of locating hospitals so that the maximum time distance to the closest hospital would be least, Mayhew proposed the drawing of hexagons on a travel velocity surface and the transferring of their boundaries on to a map of the city of London where they are transformed into a tessellation of distorted hexagons, each district of which has approximately equal time distance from the center to boundary. Their distribution reflected, therefore, the different travel speeds that occur in different parts of the city. Because both the methods of representing "demand" and "distance" data as continuous functions in any empirical situation have error and because it is impossible at the end to know the degree to which the problem as stated has been solved–no error theory is discussed–it seems safe to presume that Mayhew's solution method, though original and novel, is inferior to alternative methods, based on discrete spatial data, which have been carefully described and evaluated in the literature, (Daskin et al., [10]).

3.3. BOUNDARY EFFECTS

There are two ways in which boundaries enter applications of location models. The first is in affecting the quality of the solution. The second is as a measurement problem affecting the accuracy with which interaction patterns are portrayed.

Because location-allocation models can be used to compute optimal locations for irregular shaped areas with non-uniform patterns of demand, decision-makers are rarely concerned about the effect of study area boundaries on their decisions. Yet boundaries do affect the quality of solutions and in many practical cases they are arbitrary and are not recognized by users of facilities. That they result in a loss of efficiency is well known, but systematic exploration of their effect and of the gains to be made by removing arbitrary boundaries, has received little attention. Leamer [27] showed how the ideal-shaped hexagonal honeycomb of market areas found in an area of infinite extent and uniform demand becomes distorted by boundaries of study areas. He simulated the case of an industry with fixed costs of facility upkeep and transport costs proportional to distance, first in a square region, and then for the case when the region was divided into four smaller regions. He found that the loss in efficiency caused by the boundaries was less than one percent. "The rather small economic loss is surprising", he wrote, ([27], p. 241). "Dividing a square into four equal squares should be a severe test. For this reason we can conclude that the fracturing of uniform demands is not likely to cause enough distortion in optimal market areas to lead to significant economic losses". He went on to caution "that this statement cannot necessarily be extrapolated either to real locations or to non-uniform demands". Yet the evidence abounds that boundaries do cause large losses of efficiency in specific cases. The quality of solutions to covering models in particular, can be affected by the shape of a study area. Yet study areas are often rather arbitrarily defined. The authority of decision-makers often extend over some political-administrative space and though analysts have no control over this, they can suggest that

Table 1
Geographical accessibility to health service sites in Bellary district 1971-1981

Year	No. of sites	Average distance to nearest sites (km)		Percent efficiency
		existing	optimal	
1971	32	6.74	5.62	83.44
1976	47	5.19	4.27	82.29
1979	52	4.58	4.00	87.33
1981	59	4.16	3.66	88.12

Locational efficiency of primary health units selected in each period 1971-1981

Time period	No. of new sites	Average distance to nearest sites (km)		Percent inefficiency of new sites
		existing	optimal	
1971-76	15	5.19	4.52	30.2
1976-79	5	4.58	4.39	23.8
1979-81	7	4.16	3.90	38.2

$$\text{Percent inefficiency of new sites} = 1 - \frac{\text{net reduction in distance by new sites}}{\text{net reduction in distance of optimal sites}}$$

sensitivity analyses be performed to show decision-makers the effect of the boundaries on the quality of outcomes. An example will illustrate this point.

Working with colleagues from Iowa and the Indian Institute of Management, Bangalore, I have tried to compute how the "locational efficiency" of health services changed in the past ten years as the Government intervened to strengthen the health care system by opening up new primary health clinics in the rural areas of South India. After using a shortest path algorithm to compute the shortest road distances between all 600 settlements in one study area, we used a heuristic algorithm to compute locations that minimized average distance of people to their closest health clinic for each of several years during the decade, (Teitz and Bart [44], Hillsman [19]), see table 1. By comparing the actual clinic choices with the results from the algorithms, we found that the overall efficiency of the locations of the clinics increased from 83 percent efficiency in 1971 to 88 percent in 1981. We then evaluated the recent clinic location decisions and used the algorithm to determine how many person kilometers incurred by people visiting the clinics in a given year could be saved if the clinics were to be located at different sites. We used the algorithm to search for the sites where the savings would be greatest and compared these potential savings with the person distances saved in the locations actually selected. We found that recent location decisions saved a smaller and smaller percentage of the potential savings available. The five clinics which opened in the area from 1976 to 1979 realized 77 percent of the potential distance savings, whereas the 7 clinics opened between 1979 and 1981 realized only 62 percent of potential savings available. Why was this the case? Interviews in the field with the administrators of the health system and visits to some of the health sites themselves led us to suspect that the reason was that administrators were constrained in nominating sites for government approval. They were selecting sites from small regions of approximately 15,000 population each, that had earlier been identified by them in response to a government policy that "there should be one health center for every 15,000 people". We investigated the degree to which this constraint–that new health clinics be located in these regions–limited the potential for developing the most geographically accessible system. We computed (by enumeration) the best clinic location which, if added to the existing set of clinics, would most increase geographical accessibility. In one case, the eligible sites were all places that met minimum site-specific conditions. In the second case, the eligible sites were restricted to the same sites in the government's regions. We found that when we sequentially computed fifteen sites that met these conditions, there was an average loss of efficiency of 40 percent. We concluded that these regions, originally defined to aid decision-making, had now become the problem. By imposing arbitrary boundaries within which locations had to be selected, only a small percentage of the marginal improvements that could have been attained were in fact reached.

The local health administrator in this rural study area was fascinated to see the results of the computed analyses showing the efficiency of some of the places that he had played a role in selecting. He had his own mental ranking of which had been the best decisions and which the poorest. He also now knew from experience which of them were busy seeing patients and which ones were poorly patronized. The results of our analyses made sense to him. For our part we were fascinated to see the forms that he had to fill out for the government every time he nominated a new health clinic. Essentially, he was being asked to estimate the accessibility of the site to its local area–the same equation that was formalized in the objective function that was being optimized in our algorithm. With such good correspondence between model and reality, one might expect to see more such uses of location models in improving basic services in developing countries. There are, however, only a few such examples, ([29], [34], [36], [38]).

The second source of error caused by boundaries is measurement error. This occurs when the analyst selects a region for study which, although it may correspond with the space in which decisions are made, it may not be the space within which people constrain their spatial interactions. Many public facilities are located by groups who are responsible for a given territory in which some facilities serve people in neighboring territories and in which some people receive service from facilities in neighboring territories. How can this fact be incorporated in applications of location models? Hillsman ([18], p. 83 and p. 101) coded a location-allocation suite of heuristic algorithms so that analysts could define places as within or outside a study area. His programs reported statistics only for places within the study area. The analyst using such a program can let analyses proceed as if study boundaries did not exist by letting all places be feasible sites or can recognize the limited authority of the decision-maker being served by allowing only places within the study area to be selected. In either case, people inside the study area can interact with facilities outside the area and people outside the study area can be served by facilities within it. For many applications of location models this is a more realistic description of the role of a boundary. By working on each study area as if it were a self-contained island, the boundary affects the results by placing facilities in the border areas that are uncoordinated with related facilities in neighboring jurisdictions.

4. Conclusions

Applications of location models would be improved if more explicit attention were given to the benefits and costs of modelling locational complexity, facility operation complexity and the complexity of the geographical environment as coded in the application. In most cases a choice must be made between simplifying the evaluation of locational complexity by considering fewer combinations of facility locations, incorporating more detail about the process of managing and operating the system of interest in a realistic decision context, and coding and analyzing more details about the geographical characteristics of the study area. The history of applications of location models is one in which, increasingly, the treatment of locational complexity is being sacrificed in order to capture further detail of the operational process. This development is a response to user evaluations of applications of location models. As for capturing geographical complexity, the history of applications show little development: decisions to use crude distance functions and spatially aggregated data units continue to be the rule and the lack of any recognized method for evaluating the consequence of these choices on the quality of the results has encouraged analysts to simply dismiss the problem as intractable. The value trade-offs that accompany these choices are rarely made explicitly; instead, they usually follow from decisions of the analyst that are commonly made at an early point in the application and often are not rationalized. Yet, we see the disturbing spectacle of large computational efforts being incurred to find location patterns that are optimal with respect to a particular objective functions when errors that can be attributed to the operational characteristics of the system, and to the measurement of the geographical characteristics of the environment, are sufficiently large that for a large number of solutions there can be no statistically significant difference between their respective objective function values. In such cases, it is impossible to defend the use of methods that do not explicitly consider the whole range of solutions among which the true (unknown) optimal system exists.

The analyst is expected to provide guidance to the user in making these trade-offs. They should be made more deliberately and with better knowledge of the consequences. To achieve this, users need the opportunity to explore their problems in order to understand them better, and to discover the problems that would arise if they were to implement any suggested solution (Rushton [39]). They need not only to express their preferences but to form their preferences as they evaluate alternative solutions to their problem (Fischhoff and Goitein [14]). In short, they need "decision-support" (Densham and Rushton [11]). Typical locational decision problems are "ill-defined problems" (Hopkins [22]), where analysts, decision-makers and affected people need to interact to explore the boundaries of the problem,–controllable components, and possible solutions–and need to judge them in the context of the feasibility of their ultimate implementation, (Volkema [46]). The role of the analyst is to assist in the clarification and structuring of the problem; to generate preferences from the affected persons concerning criteria of value to them that may need to be sacrificed or traded-off against each other; to generate 'solutions' that meet defined criteria; and to generate other interesting alternatives for evaluation by affected people.

Although optimal location models have a role to play, it is not the traditional role of finding "the solution" to the problem. The essential task is to identify new locational arrangements of activities and associated new patterns of behavior of people in their roles as both providers and consumers of goods and services. Location models need to be developed that are accounting models that provide information on the expected behavior patterns that would prevail if location arrangements that were optimum with respect to selective space-serving criteria were to be adopted. These models should be used to generate alternatives that are "interesting" with respect to defined objectives so that people may explore their implications and provide reactions that will identify new objectives, (Holloway et al. [21]; Larson [25], Schneider [40]). The models can then, in turn, be used to generate new alternatives. In short, the challenge to the applied location modeller is to develop and test feasible decision processes rather than to find optimal decisions.

Acknowledgements

I acknowledge support from the National Science Foundation, Grant # SES-8516644 and assistance from Paul J. Densham and SooByong Park.

References

[1] S. Angel and G.M. Hyman, *Urban Fields*, (Pion, London, 1976).
[2] B. Ayeni, G. Rushton and M.L. McNulty, Improving the geographical accessibility of health care in rural areas: A Nigerian case study, Soc. Sci. Med. 10 (1987) 1083.
[3] J. Beaumont, Location-Allocation problems in a plane: A review of some models, Soc. Econ. Pl. Sci. 15 (1981) 217.

[4] M.L. Brandeau and R.C. Larson, Extending and applying the hypercube queueing model to deploy ambulances in Boston, TIMS Stud. Man. Sci. 22 (1986) 121.
[5] P. Casillas, Data aggregation and the p-median problem in continuous space, in: *Spatial Analysis and Location-Allocation Models*, eds. A. Ghosh and G. Rushton (Van Nostrand Reinhold, New York, 1987) pp. 327–344.
[6] A. Charnes, W.W. Cooper and E. Rhodes, Measuring the efficiency of decision making units, Eur. J. Oper. Res. 2 (1978) 429.
[7] L. Cooper, The transportation-location problem, Oper. Res. 20 (1972) 94.
[8] T. Crackel, The linkage of data describing overlapping geographical units - A second iteration, Historical Methods Newsletter 8 no. 3 (1975) 146.
[9] J.R. Current and D.A. Schilling, Elimination of source A and B errors in p-median location problems, Geog. Anal. 19 (1987) 95.
[10] M.S. Daskin, K. Hogan and C. ReVelle, Integration of multiple, excess, backup, and expected covering models, Env. And Pl. B., Vol. 15 (1988) 15.
[11] P. Densham and G. Rushton, Decision support systems for locational planning, in: *Behavioural Modelling Approaches in Geography* eds. H. Timmermans and R.G. Golledge (Croom Helm, London, 1987) pp. 56–90.
[12] D.J. Eaton, H.M.L. Sanchez, R.R. Lantigua and J. Morgan, Determining ambulance deployment in Santo Domingo, Dominican Republic, Jour. Oper. Res. Soc. 37 (1986) 113.
[13] M.J. Farrell, The measure of productive efficiency, J. Roy. Stat. Soc. Series A (General) Part III (1957) 253.
[14] B. Fischoff and B. Goitein, The informal use of formal models, Acad. Man. Rev. 9 (1984) 505.
[15] M.F. Goodchild, The aggregation problem in location-allocation, Geog. Anal. 11 (1979) 240.
[16] M.F. Goodchild and B.H. Massam, Some least-cost models of spatial administrative systems in Southern Ontario, Geografiska Annaler 52 (1969) 86.
[17] P. Gould, S. Nordbeck and B. Rystedt, Data sensitivity and scale experiments in locating multiple facilities, in: *Multiple Location Analysis* eds. G. Tornqvist, S. Nordbeck, B. Rystedt and P. Gould, Lund Studies in Geography Series C, General, Mathematical and Regional Geography no. 12, (Lund, Sweden, 1971) pp. 67–84.
[18] E.L. Hillsman, *Heuristic Solutions to Location-Allocation Problems: A User's Guide to ALLOC IV, V, and VI*, monograph 7 (Department of Geography, University of Iowa, Iowa City, IA).
[19] E.L. Hillsman, The p-median structure as a unified linear model for location-allocation analysis, Env. & Pl. A 16 (1984) 305.
[20] E.L. Hillsman and R. Rhoda, Errors in measuring distances from populations to service centers, Annal. Reg. Sci. 12 (1978) 74.
[21] C.A. Holloway, D.A. Wehrung and M.P. Zeitlin, An interactive procedure for the school boundary problem with declining enrollment, Oper. Res. 23 (1975) 191.
[22] L.D. Hopkins, Evaluation of methods for exploring ill-defined problems, Env. & Pl. B 11 (1984) 339.
[23] P.J. Kolesar, Travel time and travel distance models, in: *Fire Deployment Analysis* eds. W.E. Walker, J.M. Chaiken and E.J. Ignall (North-Holland, New York, 1979) pp. 157–201.
[24] N.S.-N. Lam, Spatial interpolation methods: A review, The Amer. Cart. 10 (1983) 129.
[25] R.C. Larson, Approximating the performance of urban emergency service systems, Oper. Res. 23 (1975) 845.
[26] R.C. Larson and K.A. Stevenson, On insensitivities in urban redistricting and facilty location, Oper. Res. 20 (1972) 595.
[27] E.E. Leamer, Locational equilibria, J. Reg. Sci. 8 (1968) 229.
[28] R. Lineberry, *Equality and Public Policy* (Sage, Beverly Hills, 1977).
[29] B.I. Logan, Evaluating public policy costs in rural development planning: the example of health care in Sierra Leone, Econ. Geog. 61 (1985) 144.

[30] R.F. Love, J.G. Morris and G.O. Wesolowsky, *Facilities Location Models and Methods*, (Elsevier Science Publishing Co, New York, 1988).
[31] L. Mayhew, *Urban Hospital Location* (Allen & Unwin, London, 1986).
[32] S. McLafferty, Constraints on distributional equity in the location of public services, Pol. Geog. Ort. 3 (1984) 33.
[33] S. McLafferty and A. Ghosh, Issues in measuring differential access to public services, Urban Stud. 19 (1982) 383.
[34] S.G. Nichols and R. Honey, An application of a location-allocation model to evaluate and improve geographic accessibility of vocational training centers in Jordan, in: *Proc.* Ninth Annual Applied Geography Conference (1986).
[35] S. Park, Performance of successively complex rules for locational decision-making, Ann. Oper. Res. 18 (1989) this volume.
[36] R.A. Reid, K.L. Ruffing and H.L. Smith, Managing medical supply logistics among health workers in Ecuador, Soc. Sci. Med. 22 (1986) 9.
[37] G. Rushton, Map transformations of point patterns: Central place patterns in areas of variable population density, *Papers* Regional Science Association 28 (1972) 111.
[38] G. Rushton, Use of location-allocation models for improving the geographical accessibility of rural services in developing countries, Intl. Reg. Sci. Rev. 9 (1984) 217.
[39] G. Rushton, Selecting the objective function in location-allocation analyses, in: *Spatial Analysis and Location-Allocation Models*, eds. A. Ghosh and G. Rushton (Van Nostrand Reinold, New York, 1987) pp. 345–364.
[40] J.B. Schneider, Solving urban location problems: human intuition versus the computer, J. Amer. Inst. Pl. 37 (1971) 95.
[41] H.D. Sherman, Hospital efficiency measurement and evaluation: empirical test of a new technique, Med. Care 22 (1984) 922.
[42] R. Stock, Distance and the utilization of health facilities in rural Nigeria, Soc. Sci. Med. 9 (1983) 563.
[43] A.J. Swersey and E.J. Ignall, eds., *Delivery of Urban Services*, TIMS Studies in the Management Sciences 22 (North-Holland, New York, 1986).
[44] M.B. Teitz and P. Bart, Heuristic methods for estimating the generalized vertex median of a weighted graph, Oper. Res. 16 (1968) 955.
[45] W.R. Tobler, Smooth pycnophylactic interpolation for geographic regions, J. Amer. Stat. Assn. 74 (1979) 519.
[46] R.J. Volkema, Problem formulation in planning and design, Man. Sci. 29 (1983) 639.