



Marry linear programming to simulation modeling. Combine the perspectives of locational analysis and the general constructs of social theory. Meld time and distance as equal variables. In fact, throw in all the data variables you want. Then crunch the numbers in a manner that yields a solution page whose results are not only robust but also intuitively comprehensive, graphic and clear.

A black box would be nice, wouldn't it? It would be a general tool whose utility ranged across the myriad issues defining the OR/MS problem set. It's not a dream, however. In fact, it's halfway built. In 10 years, perhaps less, it will be mature. Its general shape is that of the current generation of geographic information systems (GIS) software. What began as a technique for automated cartography is maturing. These technologies have become a primary environment for problem solution and general modeling. Today they complement a range of more traditional OR/MS tools. Within five years, they may subsume them.

At present, there are a range of GIS programs all sharing similar attributes, each with its own strength and weakness. What they share is a means of graphically presenting (through maps and charts) data stored and manipulated in a related spreadsheet or database. Their power lies in the marriage of easily understandable graphics and increasingly powerful systems for data analysis.

GIS

In the 1960s, they used punch cards to enter the coordinates for locations so that automated plotters could draw a precise map of a specific region. The process was arduous. While fascinating to cartographers, it was of only esoteric interest to anyone else. In the following two decades the growing power of computers permitted more intense analysis with DOS-based programs for those who understood the essentials of mapping and of computer programming.

A Windows environment, better data storage systems and a new generation of point-and-click database programs have transformed GIS into a general modeling tool. Its utility lies in the ability to analyze spatially related data and then present it in a graphical form — map, histogram, chart, or as a differentiated spatial surface — that may or may not look like a standard graphical representation of the world.

"Whereas with conventional mapping (and even the earliest digital mapping) the map was the database, today the map is merely an evanescent projection of a particular view of a spatial database at a given time." [1] Maps become, in GIS, synchronous essays displaying robust solutions to complex problems. They simultaneously present a problem definition and then evidence of its solution.

"The handling and processing of enormous amounts of data, the extraction of useful information, the fast processing of information, the decision-making and subsequent implementation of decisions and control inputs are research tasks that need to be addressed." [2]

This is the promise of GIS: a single tool that permits the extraction and presentation of useful information from huge amounts of data. In coming years GIS programs may well become the engines that drive the analytics of health care planning, transportation modeling, locational analysis and political prognostics.

Indeed, this process is well underway.

As an introduction to these tools and their capabilities, a series of OR/MS-style problems ranging from the routine to the innovative are described. The intent is to provide a general survey for OR/MS readers, irrespective of their familiarity with GIS approaches.

While examples here rely primarily on two different GIS software packages — ESRI's ArcView and Caliper's Maptitude — there are an ever-growing number of GIS programs on the market today or coming to market. No special claims are made for either ArcView or Maptitude; author familiarity with both programs influenced these examples.

The Ambulance Problem

"Optimization models provide coverage in nearness or relationship to all of the elements or members of a group." [3] A classic problem is the siting of ambulances or firehouses in a manner that maximizes coverage across a region while minimizing the cost of that coverage.

The solution requires consideration of a series of constraints involving cost, number of vehicles, distance measured in time and usage. In effect, what is required is a methodology that maximizes a spatial relationship between potential supply sites and potential demand points across a variable surface whose features include population density, housing density and cost.

Illustration 1 presents a basic GIS approach to this class of problem. Published several years ago by Ritsema Van Eck, it identifies areas not covered by a Dutch ambulance service on an island in the southwest of Holland. [4] "Net driving time" from two separate ambulance depots to local homes is marked by concentric rings of 0-2.5 minutes, 2.5 to 5 minutes, 5-7.5 minutes, 7.5 to 10 minutes, etc.

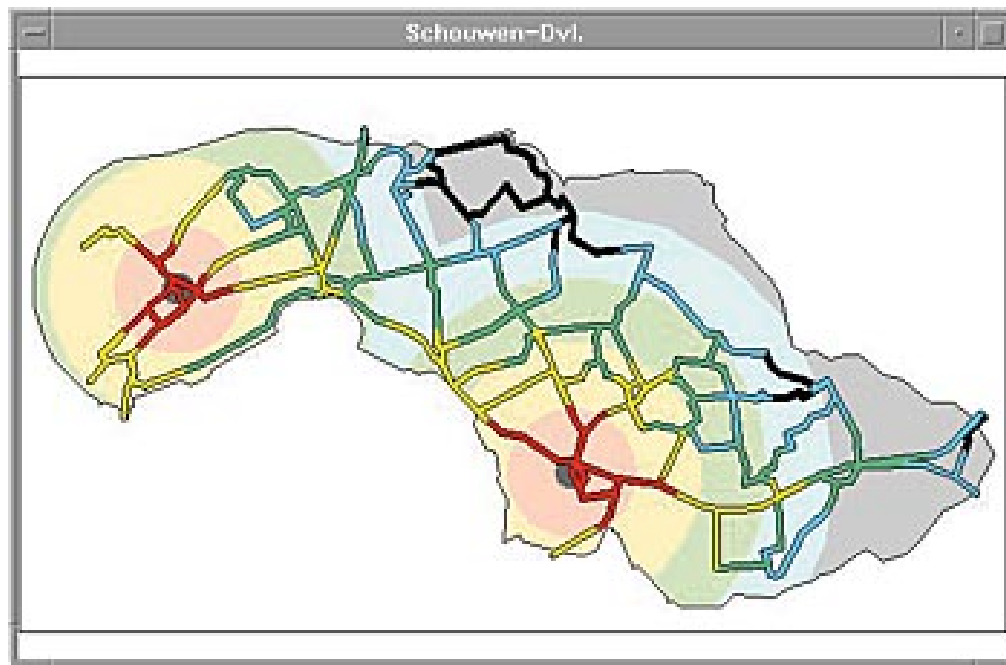


Illustration 1: A basic GIS approach to siting ambulances

Roads not adequately served — Dutch law requires ambulance service to all homes — are colored black. The solution builds from a spreadsheet in which travel time between ambulance sites and area homes is measured across the existing road network. From this point, coverage solutions easily present themselves. Introduction of a third service point is easily programmed into the set — and potential solutions can be shown in the resulting map. While this example is rather mundane by current standards, it should be noted that less than a decade ago it was considered innovative and quite revolutionary. At the same time, newer and more sophisticated approaches in this area are now coming to market. One example involves the use of flow map analysis (see, for example: <http://flowmap.geog.uu.nl/>) to provide a more detailed, raster-based surface for analysis.

Routing Problems

Routing problems have long been staple planning problems, those often taught using linear programming. Illustration 2 presents the input data and solution for a non-linear vehicle routing problem based upon travel time between warehouses and store sites. This solution was generated using TransCAD Transportation GIS software developed and distributed by Caliper Corporation (<http://www.caliper.com/>). It is included here to illustrate a GIS-based approach to a large-scale, non-linear problem type familiar to OR/MS practitioners.

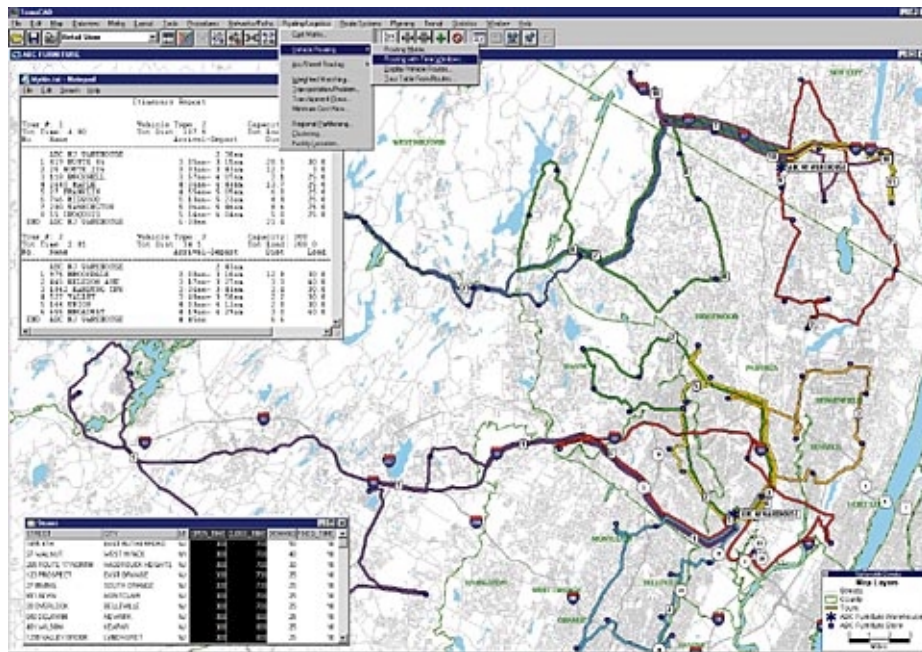


Illustration 2: Routing Problem Solution

Caliper Corporation's transportation engineer, Paul Ricotta, described the problem in this manner: A furniture company with one warehouse supplies products to 45 retail stores in its New Jersey market area. Each day, trucks deliver the goods to each of the retail stores and then return to the warehouse. Each store must be serviced within a specific time window. In addition, each store has different unloading time constraints accounted for by applying a fixed and demand-based service time.

The warehouse is similarly constrained. Because drivers work a maximum eight hours per shift, there is also a limit on the duration of each route. A further constraint fixes the capacity of each delivery truck, which is defined by the weight or volume of goods that the vehicle can carry.

The solution, displayed visually on a general road map, defines a set of routes satisfying all constraints while minimizing total travel time for the vehicle fleet. Each route can be printed graphically and in tabular form to hardcopy for use in the field. Alternate solutions might seek to minimize distance traveled.

Some of the input data, displayed in the illustration (lower left corner), are organized in a standard database format attached to the geographic component. One of the output files is also shown (upper left). This file, known as the itinerary file, outlines each route in a tabular form specifying the order and times in which the stores are serviced. There are several clear advantages to the GIS solution, Ricotta notes. First, this approach facilitates very accurate distance and travel time computation on a true street network. "Distances and travel times between points are computed from their respective latitude/longitude coordinate in real space," Ricotta says, "providing for very precise calculations."

Secondly, because the system is map-based, the level of network detail incorporated in a solution is very high. All road attributes are overt and can be included in time-distance calculations. By utilizing the database functionality inherent to this data format, the user can also specify complex road attributes. Truck exclusions on certain roads, delays at intersections, one-way streets and construction zones can all be included in the analysis and the end-product.

Thirdly, GIS provides a visual environment for the analysis. This means a graphical presentation is a reflexive component of any solution. All facets of the analysis take place in a visual and interactive environment, from preparing the input data to analyzing the solution. This relates problem formation and solution in a powerful way. That input data can be displayed both during the modeling and with the solution, offers a powerful methodology for error checking.

Finally, the visual environment provides a means of conveying highly technical information to the non-practitioner in a very straightforward and understandable manner. Users can prepare professional quality maps for employees in the field, in this case drivers, on the delivery routes.

For the client, this approach provides a graphical verification that is both intuitive and useful. An error in calculation or assumptions yields a solution that "looks wrong." If the solution is correct, maps are automatically generated for field use. In this case, maps for truckers who drive the routes were generated. And, of course, if road conditions change, editing the underlying database can easily change the solution.

The proprietary heuristic used by TransCAD is complex and robust. Other companies offer different algorithms or heuristics in their network analytic software. As advanced algorithm development continues at this and other companies, it is likely that more complex routing algorithms now reserved for top-end products will find their way into other, lower priced programs.

Redistricting

While complex, the routing problem itself is one whose solution is mundane. That is, the problem it presents is of a type that will be familiar to many readers. The next example involves analysis of a very different type of problem. Analysis of the allocation and distribution of transplantable human organs in the United States presents a monster problem in redistricting.

As a recent *OR/MS Today* article explained [5], the United States is divided for the purposes of transplant organization into 11 national districts. [6] Sixty-four organ provider organizations offer administrative and oversight services for more than 160 transplant-performing hospitals within these districts. They, in turn, utilize organs from more than 5,000 U.S. hospitals.

The efficacy of current regional boundaries has been, in this area, a critical research problem. The question is: what effect does distance have on the system's ability to harvest human organs and then transmit them to performing hospitals? Do the regions as now constituted make sense? If not, how might they be redefined?

To address this question, in another study I used Maptitude 4.0, importing from ArcView 3.1 a previously constructed database and map of 164 transplant performing hospitals. All are located in major U.S. cities for which both ArcView and Maptitude provide extensive social and economic data sets.

Next, I loaded a data set of more than 3,000 U.S. urban locations whose population is greater than 1,000. With those map layers in place, the relation between transplant performing hospitals and the urban places they supposedly serve was made explicit. A range of data based on population, income and ethnicity could be easily compared across service regions and between centers.

To understand the relation between service centers and districts, a Thiessen polygon-generating tool was used to create boundaries showing the relation between hospitals and places based on distance.

"Thiessen polygons divide a region up in a way that is totally determined by the configuration of the data points." [1] If the current regional system were designed appropriately, all urban places would be best related to hospitals in their respective regions, each to the nearest hospital.

They were not. Some were best related to hospitals in different regions. Further, multi-state regions did not reflect adequately differences in density of population across the United States. Where multiple hospitals serve a large population, very small polygons result. In other areas (in the Western United States, for example), however, polygons hundreds of square miles in area resulted.

Having identified potential problem areas with the Thiessen polygon overview, a "nearest neighbor analysis" automatically generated the distance in miles between each town and its nearest hospital. With this data, a number of redistricting solutions were presented. Choice of one over another would, of course, depend on a number of criteria inherent to the locational model chosen.

To do this work by hand would have been immensely time consuming. Distance measurement for a matrix this large would be a singularly time-consuming task. Sorting the results would have been similarly labor-intensive. With these GIS tools, results were presented both visually and in a spreadsheet in minutes. Analysis of the relations returned in both media is now well underway.

Interesting Advances

Researchers who develop their own adaptations of existing programs are authoring some of the most interesting advances in GIS. An example of this is a methodology for "what-if" analysis developed by U.S. Department of Agriculture (USDA) workers fighting the spread of Dengue Hemorrhagic Fever (DHF).

A malady common in the tropics, DHF is transmitted from person to person through the bite of several species of *Aedes* mosquitoes. Because there is no vaccine, attempting to control the spread of the illness has meant controlling the spread of the mosquitoes that carry it.

This has traditionally been a difficult and expensive procedure involving widespread spraying of expensive insecticides. The mosquitoes breed in water-holding containers, laundry drums and in trash heaps across the tropics. To be successful, mosquito control procedures must target areas whose mosquito populations are high enough to make spraying effective.

Entomologists Dana A. Focks and Jeff Carlson used ESRI's ArcView (<http://www.esri.com/>) to measure and display mosquito distribution in an area where DHF was present.

Data from a survey that inventoried water-holding containers per household, the number of mosquito pupae and the number of people per household was used to construct contour lines in a map of the area. This showed the distribution of the mosquitoes to be controlled in the area.

Contour lines show the distribution of a specific attribute across a spatial reason. Best known as a method for showing altitude on a topographic map, contour lines are widely used in GIS to show the intensity of a variety of phenomena. They can be used to illustrate disease clusters within an urban population — breast cancer, for example [7] — or of a disease carrier.

The maps generated at this stage of the analysis were used to target areas for mosquito control so the *Aedes* could be terminated in their pupae stage.

Illustration 3 shows pre- and post-treatment maps of the mosquito control sight in a neighborhood of Mayaguez, Puerto Rico. These maps were generated using ArcView and its Spatial Analyst add-on; a secondary program permits queries across multiple spatial themes.

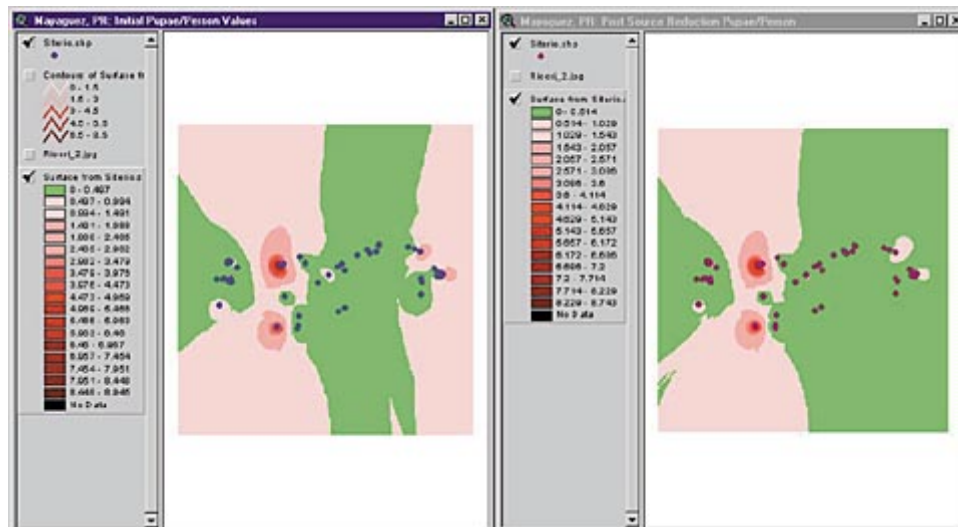


Illustration 3: Pre- and post-treatment maps of mosquito control

While most basic mapping programs are vector-based — using points, lines and polygons to construct the representative surface — this is a raster-based model. In it, data is located across a surface divided into squares in which the presence or absence of an attribute is noted. In effect, the data being studied in the delineated area becomes the surface of the resulting map, albeit one that in which each square can be immediately located in real space.

Focks and Carlson developed a script for slider bars permitting "what-if" analysis across the area. As Carlson told Arc News readers, "We can push these slider bars around to see what kind of results we'd get if, for instance, we cleaned up 25 percent of all the old tires that are sitting around collecting water and breeding mosquitoes." [8] In a recent interview, Dan Focks promised that the script for these predictive sliders, written in a simple but proprietary ArcView programming language, will be made available to other researchers in the coming months. Like many in this evolving area, unique solutions devised for specific classes of problems are shared freely with other interested persons.

The Future of GIS

At present, GIS is a specific application requiring training if it is to be used appropriately. Companies like ESRI and Caliper Corporation offer courses specific to their products. Universities are now turning out graduates trained in GIS modeling using one or another software program. As GIS matures, however, it is likely that basic techniques of application will increasingly be standardized and simplified.

As this occurs, GIS may become as commonplace as the spreadsheets and databases that store its data. Just as Insight.xla offers advanced modeling capabilities within the spreadsheet [9], future GIS programs may be sold as add-on analytic tools for programs like Microsoft Access, Excel and Borland's Paradox.

Indeed, Microsoft's promised GIS entry-level program, due out within the next six months, is likely to hasten this process. One GIS manufacturer is positioning its program to work closely with the Microsoft program. Paul Rostow at Mainfold (<http://www.manifold.net/>) believes it is in the relationship between specialized analytics and the new Microsoft program that the future of GIS can be found.

If he's correct, GIS will follow the diffusion track of two other recent technical innovations: desktop publishing and Web page authoring. Both were introduced as technical, stand-alone programs requiring specialized training. Both have become general tools in which most professionals have some knowledge. In both areas, a small number of specialists have continued to develop specific tools and approaches to advanced problems of data organization and presentation.

In the same vein, GIS approaches to standard OR/MS-style problems likely will be integrated into the field. As practitioners and clients alike become familiar with its potential, GIS will become a general tool in management science. More complex and unique problems, however, will continue to require the specialist capable of developing new approaches and new techniques. Developers will continue to find new heuristics or algorithms for specific uses.

For OR/MS specialists, this is good news. The hardest part of organizational research and management science should not be the manipulation of large data sets. It is or should be in defining the problem and knowing which of several analytic approaches to take. Whatever makes the solution clearer and easier to transmit, the better it will be for all.

And, of course, the trick still will be in choosing the scale of problem definition, the elements of its consideration and the method of approach. There is no magic in this black box. It may liberate hours, but it doesn't solve the problems analysts and researchers must carefully define, carefully address and hopefully solve.

References

1. Burrough, Peter A. and McDonnell, Rachel A., "Principles of Geographical Information systems," 1998, NY: Oxford University Press.
2. Barnhart, Cynthia, Ioannou, Petrus, Nemhauser, George and Richardson, Herbert, 1999, "NSF Workshop Maps Out Plan for Transportation," *OR/MS Today*, Vol. 26, No. 1, pp. 34.
3. Reville, Charles, 1991, "Siting Ambulances and Fire Companies: New Tools for Planners," *American Planning Association Journal*, Vol. 57, No. 4, pp. 471-484.
4. Van Eck, Ritsema J. R., 1993, "Analyse Van Transport Netwerken in GIS vor Social-Geographfisch Onderoock," *Netherlands Geographical Studies*, Vol. 164, Koninklijk Nederlands Aardrijkskundige Gennotrschap, University of Utrecht.
5. Pritsker, Allan, 1998, "Life & Death Decisions: Organ Transplantation Allocation Policy Analysis," *OR/MS Today*, Vol. 25, No. 4, pp. 22-28.
6. Koch, Tom, 1999, "The Transplant Dilemma," *OR/MS Today*, Vol. 26, No. 1, pp. 22-28.

7. Selvin, Steve, Merrill, Deane W., and Edmann, Christine, 1998, "Breast Cancer Detection: Maps of Two San Francisco Bay Area Counties," *American Journal of Public Health*, Vol. 88, No. 8, pp. 1186-1192.
8. "U.S. Military Gets ArcView GIS Reinforcements in Pest Management Battle," 1998, *ESRI Arc News*, Fall, No. 17.
9. Camm, Jeffrey D., 1999, "Software Review: Insight.Xla," *OR/MS Today*, Vol. 26, No. 1, pp. 52-55.

Tom Koch is director of Information Outreach, Ltd. He is the author of more than 10 books, including several in the field of news and information.

*E-mail to the Editorial Department of **OR/MS Today**: orms@lionhrtpub.com*

***OR/MS Today** copyright © 1999 by the Institute for Operations Research and the Management Sciences. All rights reserved.*