ENERGY ANALYSIS BY MEANS OF COMPUTER GENERATED INTERACTIVE GRAPHICS

Mark Henriquez
Energy Analysis Group
Lawrence Berkeley Laboratory

Introduction

In an age where resolution of complex technical questions is characterized by the use of large data sets, informational display in the term of interactive computer graphics has become an increasingly valuable research tool. The advantages of cartographic output over alternatives such as a tabular format have increasingly become clear to our research group At the time we were involved in the early stages of a study conducted by the Energy Analysis (EAP) Group of LBL, on behalf of the Department of Energy.

Our work which is known as Technological Assessment of Solar Energy Characterizations, (TASE) is a project in which we participated with other national laboratories. It is designed to evaluate the potential of alternative energy systems to displace a given fraction of conventional oil based fuels, identify potential applications and to develop engineering specifications of specific technologies for which potential sites are identified on a county-by-county basis within Federal Region IX. This includes California, Nevada, Arizona and Hawaii.

A number of factors are examined in the execution of this task; these include county population and demographic trends, ambient air quality, land and water use patterns, to name a few. By displaying base case (existing) and scenario specific (proposed) data sets in cartographic form, the geographical distribution of many variables is available at a glance even to those who are neither familiar with nor involved in the original research.

This is not to suggest that the use of maps has totally eclipsed tabular or graphical displays. On the contrary, for any hand manipulation of data or in other cases where absolute values are described, tables are inherently superior to maps. However, by making use of both options, we have developed a more comprehensive picture than the use of either option alone would allow.

General Process Description

A collaborative effort by the Energy Analysis Program and the Computer Science and Applied Mathematics Group at the Lawrence Berkeley Laboratory has resulted in the completion of a new method for producing computer generated graphics. The software necessary for implementation has been undergoing various stages of development by CSAM over a period of years. The author, on behalf of EAP modified and perfected this technique and was the first to successfully apply it to an actual research task.

The heart of the new system is a VAX 11/780 computer operated for DOE and the Department of Labor by the Computer Science and Applied Mathematics Department at Lawrence Berkeley Laboratory. This computer has stored on line a portion of the Socio-Economic Environmental Demographic Information System, or SEEDIS, data base. Because the fraction of SEEDIS stored on the VAX is a subset of a much larger version which resides on the laboratory's separately administered BKY computer network, we were initially faced with a choice as to which system to use. The BKY system, which consists of a ODC 6400, 6600 and 7600 computers, had a cartographic/ graphics software package (13) already up and running. But because the VAX offered a less saturated user environment and lower job costs, it was felt that these advantages outweighed the developmental effort required to produce VAX color cartography. But since the VAX lacked a color hardcopy device, methods were devised to access and manipulate data on the VAX, write it on tape, and process the tape, i.e., draw the maps, elsewhere. From an applications point of view, this approach has enabled us to undertake runs consiting of up to 23 separate Federal Region by County maps at once, in a fraction of the time and cost that would have resulted from conventional methods.

At this point, it may be helpful to examine, from an applications point of view, the steps necessary to produce a color map by this new system. To begin, one enters the interactive SEEDIS monitor of the VAX to select the area and geographic level desired. The result of this step is a geocode file, which is used in subsequent steps to automatically extract the data and interface the selected information with previously created base maps which reside in the system. Possible choices for areas include one or more Federal Regions, Standard Metropolitan Statistical Areas, Census Tract, Counties or Water Quality Control regions, to name a few. Population limits on the desired area may be set at this time. In terms of data for our chosen areas, it is possible to use either packaged data already installed on line as part of the SEEDIS monitor or to insert our own original data. Examples of installed data bases developed by Ray Tessmer at BNL, the Federal Energy Regulatory Commission Electrical Generating Unit Reference File or, what we will use to illustrate the procedure, the Populations at Risk to Air Pollution (PARAP) file developed by L. Belo and the School of Public Health at the University of California at Berkeley. We may for example, select data on the concentrations of specific airborne pollutants for the counties of California, along with information of death rates due to various forms of cancer for a given segment of the population. In some cases it is desirable to determine the ratio of two variables, and straightforward arithmetic subroutines are available for this purpose.

Having assembled the desired information, we proceed to create maps within SEEDIS, using the CARTE program developed by CSAM. The graphic files thus created are saved for additional processing, and the resulting file is recorded on tape. The maps themselves are drawn by a Zeta plotter which currently is an output peripheral on the BKY system. Alternatively, the tape may be processed in such a way as to allow for cartographic output in the form of Dicomed transparencies. These transparencies are available in a variety of formats and are characterized by intense color saturation and high resolution. Samples of graphic output for the cancer mortality information described in our example are presented in this report.

Process Applications

This process has demonstrated its usefulness in applications where a number of variables interact in a complex or synergystic manner. One example is in the case of certain water quality treatment problems which are usually impacted by energy projection technologies. Previous work by the author, (Refs. 2, 14, 15) has shown one such interaction involving the degree to which a given waste stream can be treated, the cost of treatment and the size of the treatment facility. For any given treatment system, there will be a specific number of plants which may be located in a county to meet a given level of treatment. In evaluating the applicability of a given treatment technology to a county, it is helpful to know the number of separate plants that such a county can support. Competing factors in the selection of plant sites include population trends, land use patterns and ambient environmental quality. Cartographic displays are one ideal way to present this information.

The relationship of cost, degree of treatment, and size of plant may be displayed as a two-dimensional representation of the relationship between these three variables and would take the form of Figure 1. Now if we were to superimpose a three-dimensional matrix on the design envelope shown in Figure 1, it would be possible to define equi-distant points within our matrix through or near which descriptive curves for any treatment system must pass. (see Figure 2) Each point may be identified by a code specific to its location in the matrix. Each location specific code also identifies several frames of cartographic output which were constructed using data values valid for that point in the design space. By inputting two or more of the desired parameters (cost, degree of treatment, plant size), into a separate FORTAN program, points in the matrix corresponding to points in the design curve are identified and the graphical information corresponding to the number of plants per county can be easily extrapolated.

For example, consider a specific treatment technology

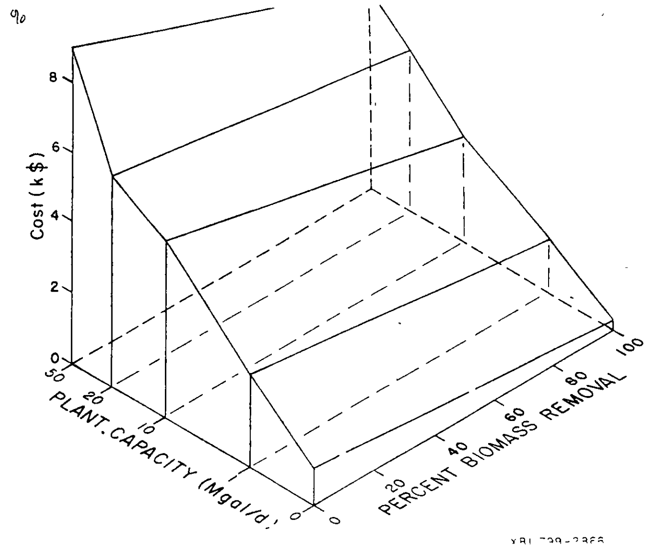


Figure 1

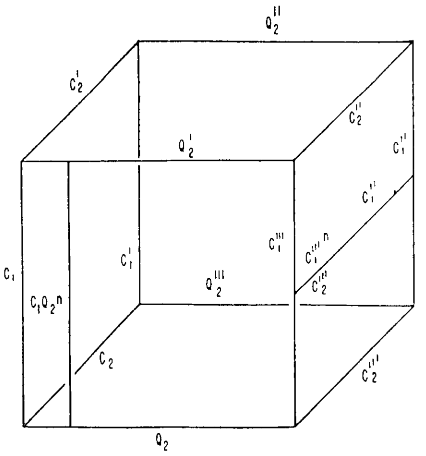
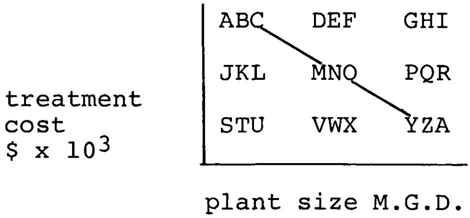


Figure 2

for which a linear relationship exists between the size of the plant and the cost of treatment over the interval of interest. Such a relationship may look like the curve described below and is shown in two dimensions for clarity. By superimposing a matrix in the form of:



on resulting design space we can identify three points, namely, ABC, MNO and YZA, which are at or adjacent to the cost/size function. Location ABC may serve as the address for one or more frames of cartographic output which show the number of treatment plants per county which may be sited if each of the plants operate within the extreme low range of plant size and the high range of treatment cost. The system may be expanded for the three dimensions.

Our new technique allows up to twenty separate frames to be stored per design point with the option of having each frame represent a multiple of the basic design space units. It is interesting to note that when the graphics data quantities are consistant with the operational characteristic equations identied for a given system, proportionality between the graphics and engineering data results. Because the maps themselves have been previously prepared, this system, which has been developed by the author for the Energy Analysis Group, has applications in instances where users are lacking sophisticated computer skills.

Methods of this type are not restricted either to water quality or energy related applications. Many fields, including aeronautical mechanical structural engineering, law enforcement, marketing and remote sensing to name a few, can benefit from interactive graphics. The reasons for the increasing popularity of interactive graphics can be largely attributed to decreasing hardware costs and the emergence of application techniques such as the one described above which seek to incorporate powerful tools into a format which is both accessible and comprehensible to the naive user. There is always the potential for a problem however, when sophisticated techniques are used by unsophisticated users.

We may take, as an illustration, our initial example of cancer mortality with respect to air pollutant concentration. Becasue both cancer and air pollution are rightly topics of public concern, there may be a tendency because of this concern to draw conclusions from existing data which may not be justified. The following example will show how the use of graphical display options may in some cases facilitate this process.

Cancer mortality data is summarized on a county base from death certificates made available from the National Center of Health Statistics. This data was manipulated to include a calculated function for the annual age adjusted rate. The result is to translate the mortality rate into the number of standard deviations above or below the mean rate for the entire U.S. This was done to equalize statistically death rates in counties of different sizes.

For example, if no deaths occur in the California counties of Alameda or Alpine the resulting zero cancer mortality rates do not accurately reflect the respective risks. This is becasue the population of Alameda is about 2 x 10^3 times as large as Alpine County. Thus zero deaths in Alameda County will be a number of standard deviations from the U.S. mean for most sites. While the number of standard deviations for a given cancer are near to zero for the entire U.S. and if even a single death resulting from this cancer is reported in Alpine county the result would be a substantial increase in death rate from the national mean - again due to the size of Alpine's population.

Other factors unrelated to statistical considerations may also mislead the researcher. For instance, cancer is characterized by long lead times between contracting the disease and succumbing to it. Given the mobility factor which characterizes human populations in this country, it is not possible at this time to determine where individuals who do succumb, initially contracted the disease. Finally, a coincidental relationship between a given pollutant and a given mortality rate does not prove a cause and effect relationship between the two.