THE USE OF SHADE PRINT MAPS TO DISPLAY INTERPOLATED OXIDANT LEVELS

BRUCE E. TRUMBO and JOHN ORAV

TECHNICAL REPORT NO. 66
JANUARY 1983

PREPARED UNDER THE AUSPICES OF
SIAM INSTITUTE FOR MATHEMATICS AND SOCIETY

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
THE USE OF SHADE PRINT MAPS TO DISPLAY INTERPOLATED OXIDANT LEVELS

Bruce E. Trumbo
California State University, Hayward
and Stanford University

and

John Orav
Stanford University

TECHNICAL REPORT NO. 66
JANUARY 1983

STUDY ON STATISTICS AND ENVIRONMENTAL FACTORS IN HEALTH (SIMS)
PREPARED UNDER SUPPORT TO SIMS FROM
ENVIRONMENTAL PROTECTION AGENCY (EPA)
SLOAN FOUNDATION
NATIONAL SCIENCE FOUNDATION (NSF)

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
THE USE OF SHADE PRINT MAPS TO DISPLAY INTERPOLATED OXIDANT LEVELS

Bruce E. Trumbo
California State University, Hayward
and Stanford University

and

John Orav
Stanford University

Abstract

Interpolated values of ozone pollution levels are calculated for each point in a region based upon data from scattered reporting stations. The interpolated value at each point is a weighted average of station values with weights inversely proportional to the $p$th power of the distances from stations. These values are displayed for easy visual inspection and analysis on shade print maps, which represent intervals of interpolated pollutant values by varying shades of grey. A variety of such maps, produced on an Apple II-Plus computer, is used to explore the effects of: (1) varying $p$, (2) changing interval choices, (3) omitting certain stations, (4) changing individual data values, and (5) using modified distance metrics.
Key Words: Distance-weighted interpolation, shade print maps, grey-scale maps, isopleth maps, statistical maps, statistical graphics, computer graphics, microcomputers in statistics.

Acknowledgements: Some of the work for this paper was done while B. Trumbo was visiting the Statistics Department at Stanford University on sabbatical leave from California State University, Hayward. J. Orav's work was partially supported by SIMS. The authors wish to thank Professor Paul Switzer for helpful comments during the early stages of this work. Preliminary and final manuscripts were typed by Ann Cambra and Kathy Hughes (at Hayward). Their efforts are gratefully acknowledged.
1. Introduction

The monitoring of air quality in a region typically depends upon measurements taken at only a few stations. When we try to use data from scattered stations to obtain a view of pollution concentrations throughout the region, two issues arise: (1) how can we compute useful interpolated or estimated pollution values at places throughout the region and (2) how can we display these computed values effectively on a map of the region? This paper explores some aspects of the use of microcomputers in these two areas.

Some data and results from a recent SIMS technical report by Faith and Sheshinski (1979) illustrate the situation. Figure 1, reproduced from their paper, shows the log of oxidant levels averaged over 112 summer days in 1972 at 16 stations in the San Francisco Bay Area. They use spatial stochastic models, obtained by Kriging, to obtain estimated oxidant levels throughout the Bay Area. The isoplethtic map in Figure 2 shows the results obtained from one of their Kriging models.

In this report we use variations of the distance-weighted interpolation method, attributed to Shepard (1968), to obtain estimated values from the same data.
Stations

1. Petaluma
2. Napa
3. San Rafael
4. Vallejo
5. Fairfield
6. Richmond
7. San Francisco
8. Oakland
9. Pleasant Hill
10. Pittsburg
11. Burlingame
12. San Leandro
13. Redwood City
14. San Jose
15. Fremont
16. Livermore

Figure 1. Locations of air quality monitoring stations showing log average ozone levels for Summer 1972. [From Faith and Sheshinski (1979).]

Figure 2. Isopleth map. Estimated values derived from data in Figure 1 using Kriging. [From Faith and Sheshinski (1979).]

Figure 3. Shade print map. Interpolated values derived from data in Figure 1 using distance-weighted interpolation. [p=2; cut points (2.87, 3.00, 3.12, 3.25). See Sections 3 and 4.]
2. **Equipment**

A major focus of this paper is to show how shade print maps can effectively display the information interpolated from station data. Until recently this would have been a matter of only theoretical interest for most readers, but the availability of inexpensive microcomputers with nontrivial graphics capabilities now makes practical the routine production of such maps.

The computations and graphics for this report were done on an Apple II-Plus (48K) microcomputer. The 1982 price of all the equipment used is about $3300. At high resolution the Apple II gives 280 x 192 pixels (picture elements), but the maps presented here measure only 100 x 148 pixels, so a microcomputer with considerably poorer resolution would suffice. Appendix A contains a list of all equipment used and additional technical comments. Appendix B lists a sample program in Applesoft BASIC for producing maps from station data.

3. **Shade Print Maps**

From Figure 2 it is clear that isoplethic maps provide much information at a glance. The estimated
value at any one location on the map can be inferred with reasonable accuracy. For example the value at the very southern tip of the Bay is clearly about 3.06. General impressions are not always easy to formulate from such maps, however. Figure 2 shows six roughly elliptical configurations that must be either peaks or depressions in pollution values. In fact, three are peaks and three are depressions, but very careful attention must be paid to the numerical labels on the isopleths to arrive at the correct interpretation.

Figure 3 is a shade print map of the style being explored here. These maps use five patterns of dots or lines to imitate white, three shades of grey, and black. Boundaries between colored regions are isopleths whose log ozone values are called cut points.

Notice that pollution peaks (dark areas) and depressions (light areas) are easily identified on the shade print map. The maps in Figures 2 and 3 are comparable in that they are based on the same data and the four cut points that separate the five shades in Figure 3, 2.87, 3.00, 3.12, and 3.25, respectively, are also isopleths in Figure 2. But the maps differ slightly in detail because non-station values in Figure 3 are obtained by distance-weighted interpolation rather than by Kriging.
4. Distance-Weighted Interpolation

Suppose data $D(x_j, y_j)$ are collected at $n$ stations located at coordinates $(x_j, y_j)$, $j = 1, \ldots, n$. Then for $0 < p < \infty$, we define the function $D_p(x, y)$ by

$$D_p(x, y) = \sum_{j=1}^{n} w^*_j(x, y) D(x_j, y_j),$$

(1)

with

$$w^*_j(x, y) = w_j(x, y)/\sum_{i=1}^{n} w_i(x, y),$$

where

$$w_j(x, y) = [(x-x_j)^2 + (y-y_j)^2]^{-p/2}.$$  

(2)

In other words, the interpolated value $D_p$ at the point $(x, y)$ is a weighted average of the results observed at the various stations. The weight $w^*_j$ for each station is inversely proportional to the $p$th power of its Euclidean distance from the point $(x, y)$. If we interpret $1/0$ as $\infty$ in the computation of $w^*_j$, then the "interpolated" value $D_p(x_i, y_i)$ at station $i$ becomes the value $D(x_i, y_i)$ observed there. (Here is an important
difference between distance weighting and Kriging, since with Kriging the estimated and observed values at a station need not agree.)

As \( p \to 0 \), \( D_p \) becomes constant throughout the region except for possible point inconsistencies at stations. The constant value is the ordinary mean of the observed values at the \( n \) stations. As \( p \to \infty \), \( D_p \) approaches a step function. In this case each station establishes a polygonal region of influence, consisting of those points closer to itself than to any other station, over which \( D_\infty \) assumes the station value.

Figures 4(a) through 4(g) show shade print maps for various values of \( p \) from \( p = .5 \) to \( p = 30 \), each with cut points at 2.87, 3.00, 3.12, and 3.25. Figure 4(h) shows the results for \( p = 30 \) for artificial data chosen midway between cut-points and distributed to emphasize the 16 almost-polygonal regions. (Values of \( p \) greater than about 30 give intermediate computational results which exceed the allowable range of Applesoft BASIC.) In this paper we explore mainly the cases \( p = 1 \) and \( p = 2 \). Table I lists values of \( p \) and cut points for all of the shade print maps presented in this report.
TABLE I
SPECIFICATIONS OF SHADE PRINT MAPS IN FIGURES 3 THROUGH 13
BASED ON DATA OF FIGURE 1

<table>
<thead>
<tr>
<th>Figure</th>
<th>p</th>
<th>2.70</th>
<th>2.80</th>
<th>2.87</th>
<th>2.90</th>
<th>3.00</th>
<th>3.10</th>
<th>3.12</th>
<th>3.15</th>
<th>3.20</th>
<th>3.25</th>
<th>3.30</th>
<th>3.40</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>2</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*4(a)</td>
<td>0.5</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b)</td>
<td>0.75</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(c)</td>
<td>1</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(d)</td>
<td>1.5</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(e)</td>
<td>3</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(f)</td>
<td>4</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(g)</td>
<td>30</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5(a)</td>
<td>1</td>
<td>A</td>
<td></td>
<td>B</td>
<td>C</td>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b)</td>
<td>1</td>
<td>A</td>
<td></td>
<td>B</td>
<td>C</td>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(c)</td>
<td>1</td>
<td>A</td>
<td></td>
<td>B</td>
<td>C</td>
<td></td>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(d)</td>
<td>1</td>
<td>A</td>
<td></td>
<td>B</td>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6(a)</td>
<td>2</td>
<td>A</td>
<td></td>
<td>B</td>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b)</td>
<td>2</td>
<td>A</td>
<td></td>
<td>B</td>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(c)</td>
<td>2</td>
<td>A</td>
<td></td>
<td>B</td>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(d)</td>
<td>2</td>
<td>A</td>
<td></td>
<td>B</td>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>**7-13</td>
<td>2</td>
<td>A</td>
<td></td>
<td>B</td>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>D</td>
</tr>
</tbody>
</table>

A: Divides white from light grey
B: Divides light grey from medium grey
C: Divides medium grey from dark grey
D: Divides dark grey from black

*Figure 4(h) uses p = 30 with synthetic data.
**Maps in Figures 7, 8, and 11 use altered data.
Figure 12 is not a shade print map.
Figure 4. Effect of varying $p$ in distance-weighted interpolation. Data and cut points (2.87, 3.00, 3.12, 3.25) are held fixed in panels (a)-(g). In panel (h) synthetic data emphasize the almost-polygonal regions for large values of $p$. 

(c) $p = 1$  
(d) $p = 1.5$
5. Choice of Cut Points

The exact choice of values for isopleths in a map such as Figure 2 is not a crucial matter. These values must cover the appropriate range and for clarity there should be neither too many nor too few of them, but the immediate visual impact of such maps is weak and not much influenced by modifications in the choice of isopleth values. For example, redrawing Figure 2 with isopleths at 2.60, 2.75, 2.95, 3.05, 3.25 and 3.40 would not make much change in its effect.

The visual impact of shade print maps is stronger. The viewer may attach significance to the colors without much thought about what range of values each color represents. On the pollution maps shown here, for example, it is natural to associate white with "completely clean" air and black with "dangerously dirty" air. Figure 5 shows a single set of interpolated values using \( p = 1 \), represented in four different ways by different choices of cut points. Figure 6 is analogous for \( p = 2 \). (Figure 4(c) provides another example of the effect of different cut-points, for the same data, when \( p = 1 \), and Figure 3 an analogous additional example for \( p = 2 \).) These examples produce a variety of impressions but each map shows relatively clean air around the Golden Gate, relatively dirtier air.
over the eastern half of the region, and the dirtiest air over Livermore (station value 3.46).

It should also be noted that changes in cut points sometimes produce about the same effect as changes in p. Compare Figure 5(d) with 6(b), and Figure 4(e) with 6(d). Extreme choices of cut points would produce dramatically misleading effects. One could use 2.8 as the cut point between darkest grey and black to give an alarmist view of ozone pollution in the Bay Area or choose cut points with very high values to give the naive viewer the impression that ozone pollution is almost entirely absent. Powerful graphical methods must be used with care and integrity to avoid powerful abuses.

In practice a sensible procedure for the choice of cut points for pollution maps would be to relate them to familiar scales or standards. For example, the boundary between darkest grey and black might be a value recognized by State or Federal regulations as hazardous, with other cut points falling at selected percentages of this value. If shade print maps are to be used for exploratory data analysis to detect trends or unusual configurations, the set of cut points should be the same for all maps compared and chosen so that collectively the maps display a variety of the available grey-scale colors.
Figure 5. Changing the set of cut points produces a different impact even though the same interpolated values are represented. p = 1.
Figure 6. Various cut points for $p = 2$. Each panel is comparable with the corresponding one in Figure 5.
6. Sensitivity to Changes in Data

We examine briefly the effect of eliminating stations and of making changes in station values. These techniques provide crude but effective methods for validating both distance-weighted interpolation procedures and shade print maps as a means of representing interpolations.

As one would expect, the effect of eliminating centrally located stations with values not much different from their neighbors is slight. Compare, for example, Figure 7(b) (same as 3) with Figure 7(c) drawn ignoring data from stations 6 and 15. On the other hand, stations at the edges of the map with values that are extreme for their neighborhood are important to the appearance of the map. Figure 7(d) omits stations 5 and 11, for example.

Small changes in individual station values produce small effects. Large changes or changes that cross a cut point are more noticeable. Compare the maps in Figure 8 with Figure 7(b). Table II gives details of changes in station values for each map in Figures 7 and 8.
TABLE II
OMITTED AND CHANGED STATION VALUES FOR SENSITIVITY STUDY

<table>
<thead>
<tr>
<th>Fig.</th>
<th>Station Number</th>
<th>Correct Value</th>
<th>Changed Value</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>7c</td>
<td>6</td>
<td>3.02</td>
<td>omitted</td>
<td></td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>3.12</td>
<td>omitted</td>
<td></td>
</tr>
<tr>
<td>7d</td>
<td>5</td>
<td>3.33</td>
<td>omitted</td>
<td>Isolated stations with high ozone values.</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>3.26</td>
<td>omitted</td>
<td></td>
</tr>
<tr>
<td>8a</td>
<td>7</td>
<td>2.56</td>
<td>2.46</td>
<td>Lowest station value decreased.</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>3.46</td>
<td>3.56</td>
<td>Highest station value increased.</td>
</tr>
<tr>
<td>8b</td>
<td>12</td>
<td>3.31</td>
<td>3.21</td>
<td>Change crosses cut point 3.25.</td>
</tr>
<tr>
<td>8c</td>
<td>16</td>
<td>3.46</td>
<td>2.46</td>
<td>Marked effects at distant locations result from these extreme changes.</td>
</tr>
<tr>
<td>8d</td>
<td>16</td>
<td>3.46</td>
<td>2.00</td>
<td></td>
</tr>
</tbody>
</table>
(a) Station locations

(b) No changes. (Same as Figure 3.)

(c) Stations 6 and 15 omitted; $p = 2$, cut points $(2.87, 3.00, 3.12, 3.25)$.

(d) Stations 5 and 11 omitted; $p = 2$, cut points $(2.87, 3.00, 3.12, 3.25)$.

Figure 7. Effects of ignoring selected stations.
(a) Stations 7 and 16 changed.  
(b) Station 12 changed.  
(c) Station 16 changed.  
(d) Station 16 changed.

**Figure 8.** Effects of altering selected station values; p = 2, cut points (2.87, 3.00, 3.12, 3.25). See Table II for details.
7. Modified Interpolation Schemes

Distance is only an approximate measure of pollutant influence or transport. Wind flow can transport pollutants selectively in particular directions and mountain ranges can effectively isolate nearby stations from each other. If the effect of such factors is predictable and understood, it may be feasible to base the weights in equation (1) on a modified distance metric in order to make more realistic interpolated maps.

Since the Bay Area has several mountain ranges running north to south, a very crude attempt to modify distances can be based on the assumption that horizontal distances are, in effect, greater for ozone influence than vertical ones. Figure 9 shows the effect of replacing equation (2) by

\[ w_j(x,y) = \left[ (a(x-x_j))^2 + (y-y_j)^2 \right]^{-p/2}. \]

In Figure 9(a), \( a = 2 \); in Figure 9(b), \( a = 3 \). (A more realistic modification to take mountain ranges into account might be to increase the distance between points on opposite sides of a range by a certain fixed quantity.)

In order to explore a possible wind flow model, we ignore mountains and hypothesize a consistent air flow that decreases effective distances in a west-to-east direction.
and increases them in an east-to-west direction. To accomplish this we replace equation (2) by

\[ w_j(x,y) = \left[ (|x-x_j| - b(x-x_j))^2 + (y-y_j)^2 \right]^{-p/2}. \]

Figures 10(a) and 10(b) show maps with \( b = .3 \) and \( b = .5 \), respectively.

Another modification of the straightforward distance-weighting scheme can be used to take into account information about pollution levels that is somehow known, though not formally measured. In the Bay Area, for example, it might be reasonable to impose a constraint based on the known low ozone levels over the Pacific Ocean on the left side of the map. One simple implementation of this idea is to choose a line roughly parallel to and west of the coast, to assume a particular (low) constraint pollution value along that line, and to take account of this additional value in the interpolation scheme. Specifically, we increase \( n \) from 16 to 17 and we use Equation (2) as it stands for \( j = 1, \ldots, 16 \); but, for \( j = 17 \), we use

\[ w_j(x,y) = |x - \ell^{-1}(y)|^{-p}, \]
Figure 9. Exaggerated horizontal distances. See Section 7.

Figure 10. Asymmetric horizontal distances. See Section 7.
Figure 11. Constraint: a low ozone level is assumed along the line shown to the west of the map. See Section 7.

Figure 12. Boundary of the convex hull of the stations (solid line) and a recommended smaller map region (dotted line). See Section 8.

Figure 13. A version of the map in Figure 3 using larger plotting rectangles. Some detail is sacrificed for speed in plotting. See Appendix A.
where the equation of the line is \( y = \ell(x) \). Figure 11 shows a map drawn using this modification with \( p = 2 \) and a log-ozone value of 1.5 along the indicated line.

8. Dangers of Extrapolation

We have used the word interpolation so far to refer to estimated values \( D_p(x,y) \) obtained from Equation (1) at all points \((x,y)\). For points outside the convex hull of the station locations (smallest convex polygon containing all stations) the correct term is, of course, extrapolation. The rationale for using Equation (1) is weak for points far outside the hull. Figure 12 shows the convex hull for the 16 stations of this report.

Careful examination of Figures 8-11 shows that extrapolated values far from the hull are especially sensitive to variations in data or interpolation method.

The area represented on a map of interpolated values should be adjusted to avoid such points. The maps in this report would benefit from the deletion of narrow strips along the eastern and southern edges. (See dotted lines in Figure 12.)
9. Conclusion

Microcomputers can be used to produce shade print maps of interpolated pollution data. Here we have used a single set of data to show the effects of various distance-weighted interpolation schemes and the consequences of various choices of cut points in such maps.

Shade print maps are an effective means of displaying pollution data. Carefully constructed individual maps or sequences of maps should be useful for summarizing pollution data and for exploratory data analysis.

REFERENCES


Appendix A

Equipment List

The maps in this report were made using the equipment listed below (with approximate 1982 discount prices):

<table>
<thead>
<tr>
<th>Equipment</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apple II-Plus, 48K</td>
<td>$1200</td>
</tr>
<tr>
<td>Black and white CRT monitor</td>
<td>130</td>
</tr>
<tr>
<td>IDS 460G Printer</td>
<td>1300</td>
</tr>
<tr>
<td>Orange Micro &quot;Grappler&quot; Interface</td>
<td>180</td>
</tr>
<tr>
<td>Apple II disk drive</td>
<td>500</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>$3310</strong></td>
</tr>
</tbody>
</table>

Technical Notes

The disk drive is not necessary for producing individual maps but is necessary if different data sets are to be input from files or if several completed maps are to be saved for later CRT display. The printer is used to make hard copies of maps on ordinary paper. Printers costing half as much can make hard copies of graphics displays on thermal paper.

The Apple II computer has two "graphics pages" each with room for two maps of the resolution and size produced here. Thus, with minor modifications in the programs given in Appendix B, the Apple II can hold four of these maps in a form that makes instant viewing possible.

Maps with the resolution presented in Figures 3 through 12 require about 1 1/2 hours each to make, since
they consist of a 50 x 75 array of rectangles which take about 2 seconds each to compute and plot. Larger rectangles produce a coarser but still adequate map. For example, Figure 13, containing 20 x 37 rectangles, took less than 30 minutes running time. (Figure 13 is comparable with Figure 3).

Several options exist for a considerable increase in speed. A modification of our program to use Apple Integer BASIC could be written to produce results much like those we obtained here and it should run several times faster. Alternatively, the Applesoft BASIC program could be compiled by any one of several commercial compilers to run considerably faster. A similar program written in one of the compiled languages (e.g., FORTRAN or PASCAL) would also run faster.

It is probably also feasible to restructure the program so that not all rectangles need to be computed. Interpolated values change gradually (except for large p) so that if one interpolated value is far from any cut point, neighboring values do not need to be computed.

If several maps are to be made that differ only in cut points, a file of interpolated values can be made during the first run. A modified program can then read these values from file instead of computing them anew in subsequent runs.
APPENDIX B

Programs

This Appendix contains two programs in Applesoft BASIC. The first, called OZONEMAPS, is a sample program for making shade print maps of the kind shown in this report. In this precise form, it will produce sequences of maps for \( p = 2 \) using the Faith-Sheshinski data of Figure 1 with cut points chosen by the user. Thus, the program just as it stands could make the sequence of maps in Figure 6. Remarks within the program suggest easy modifications for changing the program:

(a) to draw fewer larger rectangles and thus to run faster (line 1243),

(b) to accept other sets of data (lines 1472-3 and lines 23001-2),

(c) to use other values of \( p \) (line 1720), and

(d) to adjust (slightly) the vertical or horizontal scales of the map to accommodate a CRT or plotter with some distortion (lines 1261 and 1281).

Possibilities for other more extensive changes are discussed briefly in the technical remarks below.

The second program listed, called PROTECTOR, is a brief utility program. It should be run each time the
computer is turned on before the first running of OZONEMAPS or similar programs. PROTECTOR changes the loading location of programs that follow. (See remarks in line 8 of OZONEMAPS and throughout PROTECTOR.)

Technical Remarks

These remarks will assist experienced Applesoft BASIC programmers modifying OZONEMAPS for various purposes.

If a standard set of cut points is used, lines 1510-1595 should be replaced by four statements fixing the values of E1, E2, E3, and E4.

For routine (e.g., daily or weekly) production of pollution maps it would be desirable to use disk files. First, each set of data would be entered into a file and stored on the disk. Program line 1470 would be modified to read data from disk rather than from data statements in lines 23000-23040.

Since maps plot slowly it is natural to let them run overnight. If cut points are specified within the program, it is not difficult to change the program so that up to four maps can be made without intervention. [Use a four-map loop to plot two maps on graphics page 1 (HGR) and two maps on page 2 (HGR2).]

If several regions are to be mapped, a separate program could be prepared for each. Alternatively, one modified program could read data files instead of internal
data statements to obtain map outline and station location, coordinates, etc. Note that data for the number of vertices in the map outline, for the number of stations, and for the corners of the border around the map would also need to be supplied by such files.
LISTINGS OF PROGRAMS USED TO MAKE SHADE PRINT MAPS FOR THIS REPORT

LIST

1 REM * PROGRAM NAME: OZONEMAPS
2 REM * BY B. E. TRUMBO, DEPT. OF STAT.
3 REM * CALIF. STATE UNIV., HAYWARD
4 REM * HAYWARD CA 94544
5 :
6 REM PROGRAM IN APPLESOFT BASIC
7 REM FOR APPLE II-PLUS, 48K
8 REM CAUTION: LOAD PROGRAM ABOVE HI-RES PAGE 1
9 :
10 GOTO 1100
11 :
12 REM FREQUENTLY USED SUBROUTINES PLACED FIRST FOR SPEED
13 REM OMIT ALL REMARKS FOR SLIGHT FURTHER SPEED INCREASE
14 :
150 REM SUBROUTINE--PLOTS RECTANGLES
160 FOR ROW = H TO H + HT - 1 STEP 2
170 H During G.ROW TO G + WD - 1,ROW
180 NEXT
190 RETURN
200 REM SUBROUTINE--BOOLEAN VARIABLES SELECT PRINT PATTERN
210 G1 = U < E1
220 G2 = U > = E1 AND U < E2
230 G3 = U > = E2 AND U < E3
240 G4 = U > = E3 AND U < E4
250 G5 = U > = E4
260 IF G1 THEN HCOLOR = 7: GOSUB 110: GOSUB 150
270 IF G2 THEN HCOLOR = 7: GOSUB 110: HCOLOR = 6: GOSUB 150
280 IF G3 THEN HCOLOR = 7: GOSUB 110: HCOLOR = 4: GOSUB 150
290 IF G4 THEN HCOLOR = 5: GOSUB 110: HCOLOR = 4: GOSUB 150
300 IF G5 THEN HCOLOR = 4: GOSUB 110: GOSUB 150
310 RETURN
320 :
330 :
1000 REM MAIN PROGRAM BEGINS
1010 :
1020 :
1100 NP = 72
1101 REM NUMBER OF VERTICES IN MAP OUTLINE
1110 NS = 16
1111 REM NUMBER OF MONITORING STATIONS
1120 :
1130 REM DIMENSION DECLARATIONS (USING ABOVE VARIABLES)
1140 DIM X(NP): DIM Y(NP)
1150 DIM DX(NP): DIM DY(NP)
1160 DIM WU(NS)
1170 DIM NU(NS): DIM SX(NS): DIM SY(NS): DIM S$(NS)
1180 DIM U(NS)
1190 :
1200 A% = 7: REM BACKGROUND COLOR
1210 B% = 4: REM OUTLINE COLOR
1220 C% = 4: REM BORDER COLOR
1230 D% = 4: REM STATION COLOR
1240 HT = 2: WD = 2
1241 REM HEIGHT AND WIDTH OF PLOT RECTANGLES
1242 REM FOR FASTER PLOT OF LESS DETAILED MAP
1243 REM INCREASE HT TO 4 AND WD TO 4 OR 5
1250 GW = INT (100 / WD): HW = INT (148 / HT)
1251 REM NUMBER OF PLOT COLUMNS AND ROWS
1260 A = 1: REM HORIZ. SCALE FACTOR
1261 REM SET A<1 TO SHRINK MAP WIDTH
1270 B = 75: REM LOCATION OF LEFT EDGE OF BORDER
1280 C = 1: REM VERT. SCALE FACTOR
1281 REM SET C<1 TO SHRINK MAP HEIGHT
1290 D = 5: REM LOCATION OF TOP OF BORDER
1299 :
1300 HGR : HCOLOR = A%: HFLOT 0,0: CALL 62454
1310 REM SET MIXED GRAPHICS MODE AND BACKGROUND COLOR
1320 HOME : VTAB 22
1330 PRINT " AIR QUALITY MONITORING STATIONS"
1340 PRINT " IN THE BAY AREA"
1350 FOR I = 1 TO NP: READ X(I),Y(I): NEXT
1351 REM LOADS OUTLINE ARRAY
1360 FOR I = 1 TO NP: X(I) = 2 * INT (X(I) / 2): NEXT
1370 FOR I = 1 TO NP: READ DX(I),DY(I): NEXT
1371 REM LOADS OFFSET ARRAY FOR 2ND TRACING OF OUTLINE
1380 GOSUB 3010: REM BORDER
1390 GOSUB 4010: REM OUTLINE
1400 FOR I = 1 TO NS: READ NU(I): READ SX(I): READ SY(I): READ S$(I): NEXT
1410 REM LOADS STATION INFO ARRAYS
1411 REM WHERE NU=STATION NUMBER
1412 REM SX & SY ARE STATION LOCATION COORDINATES
1413 REM S$=STATION NAME
1420 GOSUB 5010: REM PLOT AND ANNOUNCE STATIONS
1430 HOME : VTAB 22
1440 PRINT " AIR QUALITY MONITORING STATIONS"
1450 PRINT " IN THE BAY AREA";
1460 FOR ZZ = 1 TO 5000: NEXT : REM DELAY
1470 FOR I = 1 TO NS: READ U(I): NEXT
1471 REM LOADS OZONE DATA ARRAY
1472 REM IN GENERAL USE: REPLACE WITH
1473 REM DATA ENTRY OR FILE READ PROCEDURE
1490 :
1500 HOME : VTAB 22
1510 PRINT "WHAT ARE THE INTERVAL CUT POINTS?"
1520 PRINT " (ENTER 4 NUMBERS SEPARATED BY COMMAS)"
1530 INPUT E1,E2,E3,E4
1540 IF E1 > E2 OR E2 > E3 OR E3 > E4 THEN PRINT : PRINT "ENTER CUT POINTS IN INCREASING ORDER.": GOTO 1510
1550 PRINT : PRINT "CUT POINTS ARE: ";E1:"; , ";E2:"; , ";E3:"; , ";E4:"; ."
1570 PRINT "IS THIS OK? (Y OR N) ";
1580 POKE - 16368,0: GET R$: IF R$ ( "Y" AND R$ ( "N" THEN GO TO 1580
1590 IF R$ = "N" THEN PRINT : GOTO 1510
1595 PRINT
1599 :
1600 REM COMPUTE VALUES AND PLOT RECTANGLES
1610 FOR GN = 0 TO GW - 1
1620 FOR HN = 0 TO HW - 1
1630 PRINT : PRINT "COL" ; GN ; " ; ROW" ; HN ; " : VAL" ;
1640 GC = GN * WD + INT ( WD / 2 )
1650 HC = HN * HT + INT ( HT / 2 )
1660 TU = 0
1670 FOR I = 1 TO NS : WU(I) = 0 : NEXT
1671 REM WU ARE STORAGE LOCATIONS USED TO COMPUTE WEIGHTS
1680 FOR J = 1 TO NS
1690 QX = GC - SX(J) ; QY = HC - SY(J)
1700 DI = QX * QX + QY * QY
1710 REM DI IS SQUARED DISTANCE FROM RECTANGLE CENTER TO STATION J
1720 REM ADD A STEP HERE TO CHANGE TO OTHER POWERS OF DISTANCE
1730 IF DI = 0 THEN DI = 1E - 30
1740 REM PREVENTS ZERO DENOMINATORS
1750 REM IF STATION IS IN RECTANGLE
1760 REM COMPUTED VALUE WILL BE APPROX. = STATION VALUE
1770 IF U(J) = 0 THEN WU(J) = 1 / DI
1780 REM STATIONS WITH MISSING VALUES (ENTERED AS -1)
1790 REM ARE IGNORED IN WEIGHTING SCHEME
1800 TU = TU + WU(J)
1810 NEXT
1820 FOR K = 1 TO NS : WU(K) = WU(K) / TU : NEXT
1830 U = 0 : FOR I = 1 TO NS : U = U + WU(I) * U(I) : NEXT
1840 G = A * GN * WD + B : H = C * HN * HT + D : GOSUB 210
1850 REM PLOT A RECTANGLE
1860 PRINT U
1870 NEXT
1880 NEXT
1890 GOSUB 3010 : REM BORDER
1900 GOSUB 4010 : REM OUTLINE
1910 FOR AA = 1 TO 20 : PRINT CHR$(7) ; : NEXT
1920 REM ALARM TO SIGNAL COMPLETION
1930 END
1940 :
1950 :
2000 REM MAIN PROGRAM ENDS
2010 :
2020 :
2030 REM BACKGROUND PLOT SUBROUTINES
2040 :
3000 REM SUBROUTINE-DRAWS BORDER AROUND MAP
3010 HCOLOR = B %
3020 HPLLOT A * 0 + B, C * 0 + D TO A * 100 + B, C * 147 + D TO A * 100 + B, C * 147 + D TO A * 0 + B, C * 0 + D
3030 HPLLOT A * 0 + B + 1, C * 0 + D + 1 TO A * 100 + B + 1, C * 0 + D + 1 TO A * 100 + B + 1, C * 147 + D + 1 TO A * 0 + B + 1, C * 0 + D + 1
3040 RETURN
3050 :
4000 REM SUBROUTINE-DRAWS MAP OUTLINE
4010 HCOLOR = C %
4020 XF = X(1); YF = Y(1)
FOR I = 2 TO NP
XS = X(I):YS = Y(I)
HPLT A * XF + B,C * YF + D TO A * XS + B,C * YS + D
XF = XS:YF = YS
NEXT
REM 1ST TRACING ENDS, 2ND TRACING BEGINS
XF = X(I) + DX(I):YF = Y(I) + DY(I)
FOR I = 2 TO NP
XS = X(I) + DX(I):YS = Y(I) + DY(I)
HPLT A * XF + B,C * YF + D TO A * XS + B,C * YS + D
XF = XS:YF = YS
NEXT
RETURN
REM SUBROUTINE-PLOTS AND ANNOUNCES STATIONS
FOR I = 1 TO NS
HCOLOR= D%
HPLT A * SX(I) + B,C * SY(I) + D
PRINT : PRINT : PRINT
HTAB (12 - .5 * LEN (S#(I)))
PRINT "STATION ",NU(I),":" IS ",S#(I)
FOR ZZ = 1 TO 1000: NEXT : REM DELAY
NEXT
RETURN
REM DATA STATEMENTS
COORDINATES FOR MAP OUTLINE
IN 72 (X,Y) PAIRS
DATA 0,46,13,63,18,61,16,57,17,56
DATA 19,58,22,56,16,49,18,45,21,44
DATA 18,39,19,28,24,29,28,25,38,36
DATA 41,37,46,38,51,41,61,33,66,32
DATA 69,34,68,37,72,40,78,40,87,40
DATA 100,36,100,40,90,43,90,45,100,46
DATA 100,48,79,47,75,44,51,44,43,40
DATA 37,43,32,43,26,49,30,52,33,56
DATA 34,69,31,67,32,71,38,75,38,78
DATA 45,87,45,101,48,106,52,111,55,113
DATA 55,113,45,112,42,106,42,105,39,100
DATA 32,95,25,92,25,88,26,78,24,70
DATA 22,65,18,66,13,69,13,89,11,91
DATA 10,102,14,103,14,114,17,121,15,127
DATA 14,135,19,147
OFFSETS FOR SECOND TRACING OF MAP OUTLINE
DATA 1,0,0,-1,-1,0,-1,0,0,-1
DATA 0,-1,-1,0,-1,0,-1,0,-1,0
DATA -1,0,-1,-1,0,-1,0,-1,1,0
DATA 1,-1,1,1,0,-1,0,-1,0,-1
DATA 1,0,1,0,0,-1,0,-1,0,-1
DATA 0,-1,0,1,1,1,-1,0,-1
DATA 0,1,0,1,0,1,0,1,0,1
DATA 0,1,0,1,1,0,1,0,1,0
DATA 1,0,1,0,1,0,1,0,1,0
DATA 1,1,-1,1,-1,1,-1,1,-1,1
DATA -1,1,-1,1,-1,0,-1,0,-1,0
DATA 0,1,0,1,1,0,1,0,1,0
DATA 1,0,1,0,1,0,1,0,1,0
DATA 1,0,1,0
REM STATION LOCATION COORDINATES AND STATION NAMES
DATA 1,6,11,"PETALUMA"
DATA 2,40,10,"NAPA"
DATA 3,14,44,"SAN RAFAEL"
DATA 4,42,31,"VALLEJO"
DATA 5,62,20,"FAIRFIELD"
DATA 6,34,49,"RICHMOND"
DATA 7,20,70,"SAN FRANCISCO"
DATA 8,40,66,"OAKLAND"
DATA 9,58,53,"PLEASANT HILL"
DATA 10,74,49,"PITTSBURG"
DATA 11,20,92,"BURLINGAME"
DATA 12,46,79,"SAN LEANDRO"
DATA 13,38,106,"REDWOOD CITY"
DATA 14,62,130,"SAN JOSE"
DATA 15,64,104,"FREMONT"
DATA 16,74,89,"LIVERMORE"
REM OZONE DATA FOR STATIONS
REM IN GENERAL USE DATA WOULD BE ENTERED FROM KEYBOARD
REM OR READ FROM A FILE
DATA 3.21, 2.93, 2.68, 3.13
DATA 3.33, 3.02, 2.56, 2.82
DATA 3.20, 3.14, 3.26, 3.31
DATA 2.96, 3.18, 3.12, 3.46

JLIST
1 REM * PROGRAM NAME: PROTECTOR
2 :
3 REM * RUN THIS PROGRAM BEFORE OZONEMAPS
4 REM * TO PROTECT OZONEMAPS FROM DESTRUCTION
5 REM * BY ITS OWN GRAPHICS PLOT INSTRUCTIONS
8 :
9 :
100 POKE 104.96
101 REM CHANGE BEGINNING LOAD LOCATION FOR APPLESOF TO 24576
102 REM (96=6*16=$60, 24576=6*16³+$6000=1+TOP OF GRAPHICS PG 2)
103 :
200 POKE 24576,0: POKE 24577,0: POKE 24578,0
201 REM ZERO FIRST 3 BYTES OF NEW APPLESOF PROGRAM LOCATION
202 REM TO PERMIT LOADING THERE
300 END