Atmospheric Environment Vol. 12, pp. 1531-1535.

0004-6981/78/0701-1531 $02.00/0

CORRELATION BETWEEN AIR POLLUTION AND SOCIO-ECONOMIC FACTORS IN LOS ANGELES COUNTY

SCOTT J. ANDERSON, BRUCE W. GARDNER, BETSY J. MOLL, GUY L. TRIBBLE, III, THOMAS F. WEBSTER, KENT R. WILSON* and MICHAEL J. ZYDA

Department of Chemistry and Energy Center, University of California, San Diego, La Jolla, CA 92093, U.S.A.

(*First received* 16 *July* 1976 *and in final form* 3 *October* 1977)

Abstract – The correlation between air pollution dosage and socio-economic group is analyzed for Los Angeles County. Census derived socio-economic data are aggregated to 31 statistical areas. Nine general socio-economic variables are used: age, time of moves, transportation, schooling, employment, income, rent, housing value and race, each subdivided into several categories, for a total of 37 categories. The air pollution dosage experienced by each of these 37 categories is computed, and the degree of correlation of dosage with each of the nine socio-economic variables is determined by analyzing the pattern of the dosages among the categories for that variable. Even though the air pollution and socio-economic variables have widely ranging spatial patterns, there is no strong correlation between the air pollution and socio-economic patterns. Weak correlations are observed in the following cases: (i) blacks breathe on the average better air as measured either in terms of oxidant level or a combined pollutant measure and (ii) by the combined measure air quality improves slightly (so slightly as to perhaps be insignificant) with education, recentness of move, income, rental level and housing value.

INTRODUCTION

While Los Angeles County is well known for its overall air pollution problem, one can ask how its effects fall upon different socio-economic groups and whether air pollution is perceived as a serious enough problem that a demographic "fractional distillation" has occurred; those who can afford it moving to areas of better air quality (and perhaps driving up housing prices) while the less affluent fill in the areas of lower air quality. One might also ask whether increased education or being young, employed, or a professional rather than blue collar worker makes one more likely to live in better air.

The narrower question of the influence of air pollution on property values and the methodology to arrive at it has been much debated (see Smith and Deyak, 1975 and references therein), with the discussion centering on the effects of non-photochemical air pollution (in contrast to the Los Angeles situation) and with different authors coming to differing, and sometimes even contradictory, conclusions.

As Gold (1970) points out, it is clear, on the basis of opinion polls, that "in the state of California today there is a general opinion, verging on consensus, that air pollution is a serious problem and, in particular, poses a health problem for all people". Is this general perception specific enough for individuals to judge area by area differences in average pollutant levels, and further do people care enough to move to areas of cleaner air?

Van Arsdol (1964, 1966) concluded on the basis of 1940, 1950 and 1960 census data and a sample of early air pollution data that "while changes in the face of smog in Los Angeles differ by contrasting populations and functions, smog itself remains as an important determinant of population change". We have analyzed 1970 Los Angeles County air pollution and census data in order to try to test this conclusion and to investigate some of the above questions in more detail. It should be recognized that our aim is largely an observational one; to make visible for Los Angeles the correlations between air pollution and socio-economic factors. The question of cause and effect is a more difficult one to demonstrate. The observed correlations (or lack of correlations) are an inducement to speculation but do not provide proof of causation.

DATA AND METHODS

The limited number of air pollution monitoring stations in Los Angeles County in 1970 restricts our air pollution spatial resolution in comparison with census socio-economic data. As shown in Fig. 1, data are used from 11 Los Angeles Air Pollution Control District monitoring stations, supplemented on the boundary of Los Angeles County by data from three non-County stations.

From this air pollution data base, we first calculate for each air quality monitoring station daily averages for every day of 1970 for total oxidant, carbon monoxide, nitrogen dioxide, sulfur dioxide and particulate matter. Missing data are filled in by linear interpolation over time. The data are then spatially interpolated for every day of 1970 to form a 50 by 60 grid of pollution values across Los Angeles County. The interpolation algorithm is essentially the distance weighted method used by the Harvard Laboratory for Computer Graphics in SYMAP (1968), modified to run much faster for

* Author to whom correspondence should be addressed.

this special application. The daily oxidant surfaces are averaged over the entire year to create an average oxidant surface for all of 1970, shown in Fig. 1.

In addition to the oxidant surface we create an air quality surface, taking into account several pollutants, in order to roughly portray overall air quality. For each day, the five pollutants listed above are combined at each grid node according to a version of the ORAQI (Babcock and Thomas, 1970; Thomas *et al.*, 1971) method of compiling an air quality index:

$$\text{ORAQI} = \left[5.7 \sum_{i=1}^{5} (C_i/S_i)\right]^{1.37}$$

in which C_i = concentration of pollutant i and S_i = Environmental Protection Agency standard for pollutant i, extrapolated to 24-h average. The standards are extrapolated from those of the U.S. Environmental Protection Agency to a common basis of 24-h averaging time. The approximation method proposed by Larsen (1968) is used, which assumes a log normal distribution of pollutant concentrations in order to relate the maxima and daily averages. The ORAQI function is applied to the five pollutant surfaces for each day, and the resultant ORAQI surfaces are averaged over the entire year to yield an air quality index surface. This rather cumbersome procedure is necessitated by the nonlinearity of the ORAQI defining equation. The resulting index is scaled such that a value of 10 corresponds to unpolluted air and a value of 100 represents the 24-h extrapolated standard for all pollutants. The total oxidant and ORAQI spatial distributions are quite similar, as the ORAQI distribution is dominated by the high oxidant levels in relation to the Federal standards which are characteristic of the Los Angeles region. The areal correlation of the two surfaces is a 0.9.

Since this study was completed, a discrepancy between the oxidant calibrations used by Los Angeles and neighboring counties has been discovered. The Los Angeles County data

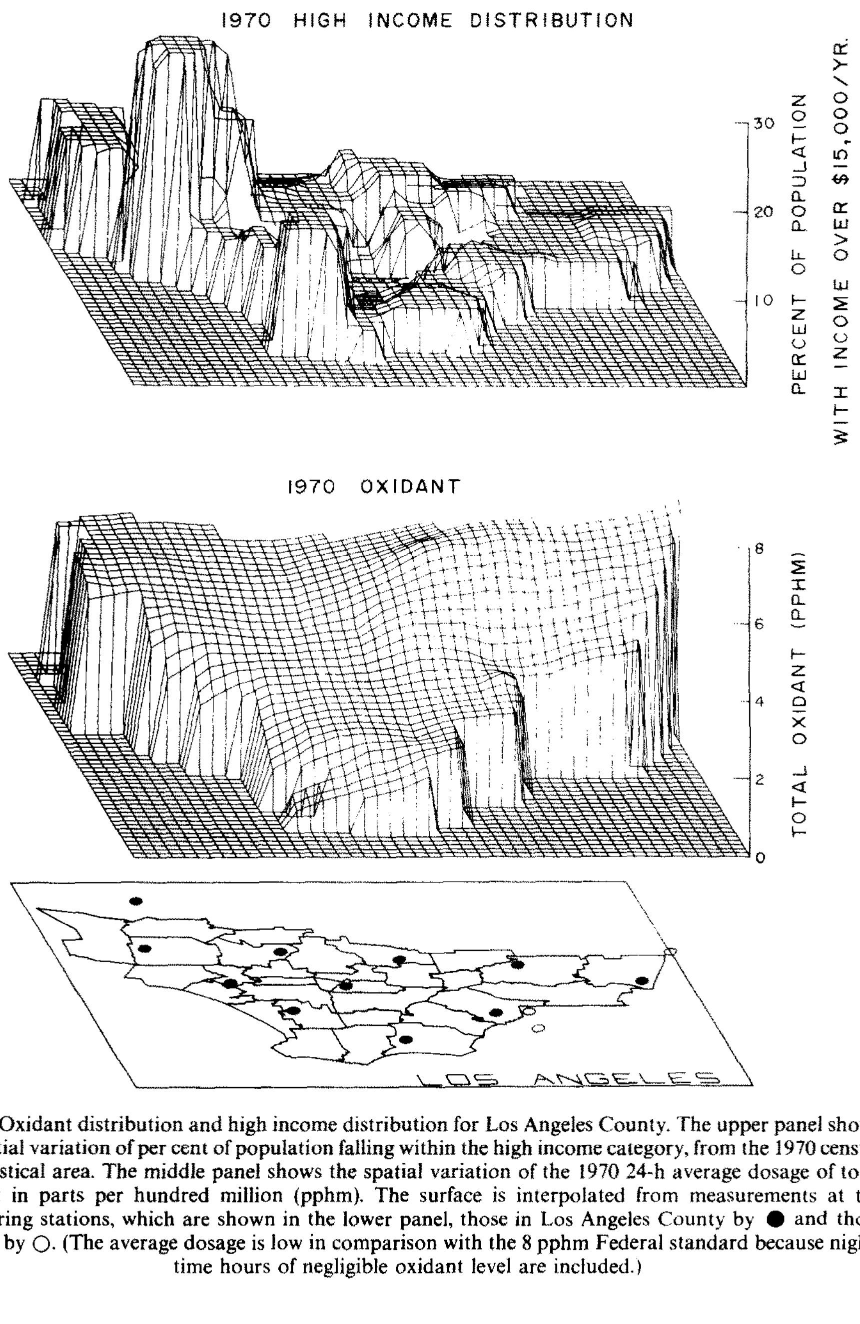

Fig. 1. Oxidant distribution and high income distribution for Los Angeles County. The upper panel shows the spatial variation of per cent of population falling within the high income category, from the 1970 census, by statistical area. The middle panel shows the spatial variation of the 1970 24-h average dosage of total oxidant in parts per hundred million (pphm). The surface is interpolated from measurements at the monitoring stations, which are shown in the lower panel, those in Los Angeles County by ● and those outside by ○. (The average dosage is low in comparison with the 8 pphm Federal standard because night-time hours of negligible oxidant level are included.)

are the more nearly correct and the oxidant data from the three stations used to provide additional information on the eastern border should be lowered by approx. a factor of 0.8 and the Los Angeles raised by a factor of 1.1 (Pitts *et al.*, 1976). The eastern border of the air quality surfaces should thus be relatively somewhat lower. It is unlikely that this change on the border of Los Angeles County would significantly alter any of the conclusions for Los Angeles presented here.

Socio-economic data base

In contrast to the air pollution data measured at only 14 points, the socio-economic data are available in the form of census returns for 1553 census tracts in Los Angeles County. Bureau of the Census fourth-count tapes of 1970 population and housing are used. Because of the spatial resolution limitations imposed by the air quality data, the socio-economic data are aggregated to 35 statistical areas as defined by the Los Angeles County Regional Planning Commission (1972), which saves subsequent computation time. Four of the statistical areas are discarded due to inadequate coverage by the air quality data base (Santa Catalina Island, San Clemente Island, the north beach areas and the San Bernardino and San Gabriel Mountains).

An example of the socio-economic data, the percentage of population in each statistical area with income over $15,000, is shown in Fig. 1. Note that this spatial distribution (as is true of the other socio-economic distributions) is quite different than the air pollution distribution of Fig. 1.

Correlations

The boundaries of the statistical areas are digitized and the oxidant and ORAQI air pollution surfaces are scaled to the map coordinates to form overlays. A polygon overlay algorithm was developed and is used to integrate the air pollution surfaces over each statistical area. Dividing the integral by the geometrical area yields an average oxidant value and an average air quality index value for each statistical area.

The degree of correlation between air pollution and socio-economic distributions is determined by calculating the average dosage of air pollution experienced by the population in each socio-economic category. These dosage levels are obtained as follows:

$$\text{DOSAGE} = N_{\text{tot}}^{-1} \sum_{i=1}^{31} \Omega_i N_i$$

where Ω_i = average total oxidant or ORAQI value for statistical area i, N_i = number of people in a particular socio-economic category in area i, and N_{tot} = total number of people in that socio-economic category in all statistical areas. The degree of correlation of air pollution with each socio-economic variable is shown graphically by the pattern of the dosages among the different categories for that variable in Figs. 2 and 3.

The correlations between oxidant air pollution and various socio-economic factors, as shown in Fig. 2, are surprisingly low. The dosages show remarkably little disparity between even the most widely differing socio-economic categories. This is in contrast to the wide variation in the dosages of total oxidant and ORAQI among the statistical areas themselves as illustrated in Fig. 1. In comparison with the magnitude of air quality variation across the County, the differences among socio-economic categories are on the whole quite small. For example, interpolated total oxidant values averaged over each statistical area vary between 0.98 and 5.0 pphm among the 31 statistical areas, while the variation among the 37 socio-economic categories ranges, with only one exception, lies between 2.95 and 3.25 pphm. While our calculational scheme and the smoothing it involves is such that the variation across socio-economic categories must always be less than the variation across statistical areas, such a marked decrease in range indicates a lack of strong correlation of the socio-economic variables with oxidant level. The one exception is a racial distinction: the air which blacks breathe is significantly better in terms of average oxidant dosage than that breathed by chicanos and whites.

Total oxidant is not the only air pollutant in Los Angeles County, so one can also study the correlation between socio-economic factors and other air quality measures. As described above, we have used the ORAQI method (Babcock and Thomas, 1970; Thomas *et al.*, 1971) in an attempt to form a single air quality index surface, which roughly takes into account the contributions of total oxidant, carbon monoxide, nitrogen dioxide, sulfur dioxide and particulate matter. Any such method which tries to combine pollutant measures is necessarily a gross approximation, and certainly no account is taken of possible synergistic effects.

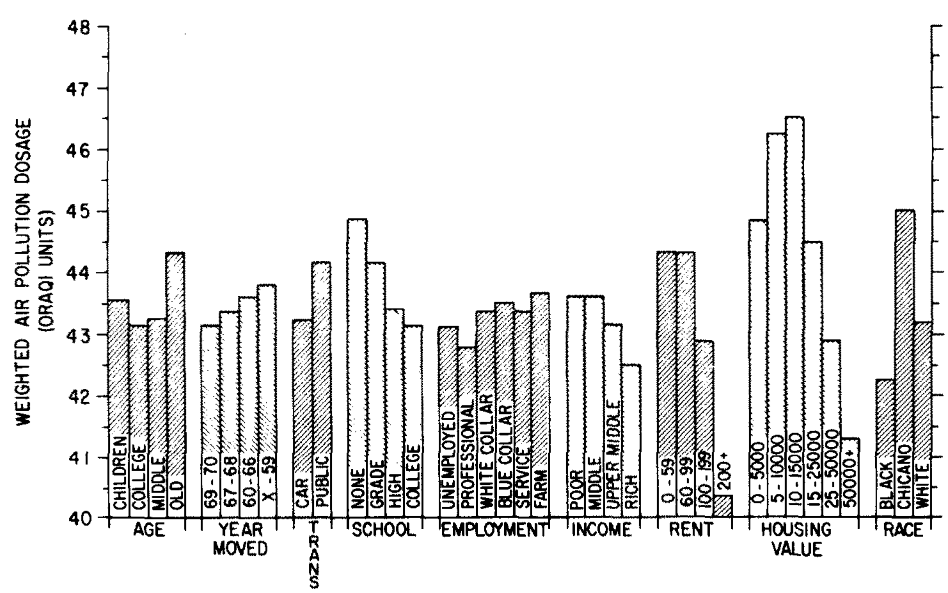

Fig. 2. Bar graph showing 1970 total oxidant dosage for 37 different socio-economic categories for 9 variables.

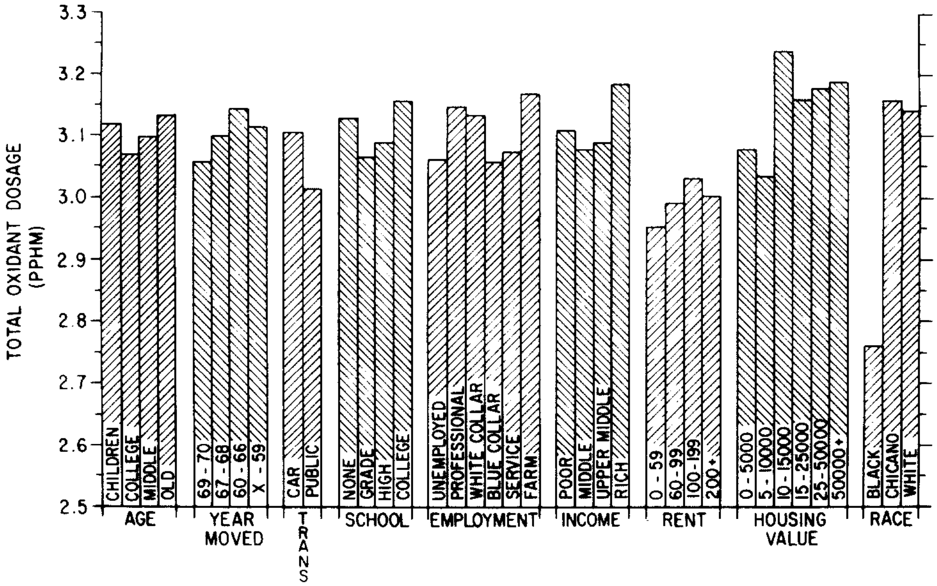

Fig. 3. Bar graph showing 1970 ORAQI air quality index dosage for 37 different socio-economic categories for 9 variables.

The ORAQI air quality index surface is almost visually indistinguishable from the total oxidant surface. Again, there is no major observed correlation of air quality with socio-economic factors.

The minor ORAQI correlations are, however, usually in the "expected" direction. The college age group, 18–29 years old, which might be expected to be the most mobile, receives a slightly lower ORAQI dosage, and those over 60 years old, presumably less mobile, the higher dosage. Those who have moved most recently have moved to better air quality than those who moved longer ago. ORAQI air quality improves steadily with educational level. Professional occupations are better on the ORAQI scale than other employment categories. The higher levels of rental and housing prices are the best in terms of ORAQI air quality. Only the dosages by racial and ethnic groups are surprising; blacks again breathing the best air on the ORAQI scale, chicanos the worst and whites intermediate.

While most of these correlations appear reasonable, we emphasize again that they are quite small, and thus many of them may be statistically insignificant.

LIMITATIONS

Three limitations to the approach we have presented should be borne in mind. First, the air quality measurements themselves are not absolutely accurate. Second, they are taken at rather widely spaced locations, so that high spatial frequencies, the localized peaks and valleys which may exist in the air pollution data, are lost. Thus very fine-grained correlations between socio-economic factors and air pollution would not be seen. Third, an individual's residential location decision is a multidimensional one, and a correlation analysis may not necessarily demonstrate a causative relationship between air pollution levels and socio-economic factors, given the many other factors such as employment location, transportation routing and zoning which also influence residential spatial patterning (Toyne, 1974).

SUMMARY AND CONCLUSION

The data presented do not lend support to the concept that the population of Los Angeles County has significantly sorted itself out geographically in response to air pollution environmental stress, such that those best able to do so have moved to areas of better air quality. It is of course possible that many people with means to do so have moved out of the air basin entirely in order to avoid air pollution. The question remains as to whether relative air quality levels are poorly perceived, or whether air quality is given a lower priority than other determinants (King, 1973) of housing location.

Acknowledgements – The research was funded as a Student Originated Study by the National Science Foundation, and we thank NSF for their help. The air quality data were provided by the Los Angeles Air Pollution Control District and the California Air Resources Board. Our thanks to L. Myrabo and K. Bond for their aid and to the Burroughs Corporation, to the National Science Foundation Division of Computer Research and to the National Institutes of Health for making the computer usage possible.

REFERENCES

Babcock L. R., Jr. and Thomas W. A. (1973) In *Proc. Third Int. Clean Air Conf.* (Edited by H. Schackmann) VDI-Verlag, Dusseldorf, D64-D66.

Gold D. (1970) Public Concern and Beliefs About Air Pollution in California: A Statewide Survey. University of California, Project Clean Air, S-11.

Harvard Laboratory for Computer Graphics (1968) *User's Reference Manual for Synagraphic Computer Mapping "SYMAP" Version V.* Harvard University, Cambridge, Massachusetts.

King A. T. (1973) *Property Taxes, Amenities and Residential Land Values.* Ballinger, Cambridge, Massachusetts.

Larsen R. I. (1968) A new mathematical model of air pollutant concentration, averaging time, and frequency. *J. Air Pollut. Control Ass.* **19**, 24–30.

Pitts J. N., Jr., Sprung J. L., Poe M., Carpelan M. C. and Lloyd A. C. (1976) Corrected south coast air basin oxidant data: some conclusions and implications. *Envir. Sci. Technol.* **10**, 794–801.

Regional Planning Commission, County of Los Angeles (1972) *Correspondence Table, Census Tracts and Statistical Areas.*

Smith V. K. and Deyak T. A. (1975) Measuring the impact of air pollution on property values. *J. region. Sci* **15**, 277–288.

Thomas W. A., Babcock L. R., Jr. and Shults W. D. (1971) Oak Ridge Air Quality Index. Publication ORNL-NSF-EP-8, Oak Ridge National Laboratory, Tennessee.

Toyne P. (1974) *Organisation, Location and Behavior.* Halsted Press (John Wiley), New York.

Van Arsdol M. D., Jr., Alexander F. and Sabagh G. (1964) Reality and the perception of environmental hazards. *J. Hlth Human Behav.* **5**, 144–153.

Van Arsdol M. D., Jr. (1966) Metropolitan growth and environmental hazards: an illustrative case. *Ekistics* **21**, 48–50.