Marjorie Rush, Ph.D.
Sally Vernon, M.A.
The University of Texas Health Science Center at Houston
Houston, TX 77025

Remote Sensing and Urban Public Health

Environmental data interpreted from aerial photographs were found to be as good as census data in predicting health status in an urban community.

Introduction

Health status is the outcome of multi-factorial circumstances. One set of factors influencing a population's health is the urban environmental set. The purpose of this project was to determine the applicability of remote sensing in the form of aerial photography to urban public health problems. Spatial distributions of visual, physical, environmental characteristics in urban areas were identified from photographs, and their block groups. For each block group, selected census data, believed to be related to health, were chosen. Environmental information—aerial photographic interpretations of land uses and quality—formed another data set for each block group. The census and land-use data were independent variables in the models to predict health outcomes. Health outcome data were solicited from several local agencies and consisted of morbidity and mortality rates.

Abstract: *Measures of the environment derived from aerial photo-interpretation procedures were compared to census environmental measures for their ability to discriminate among health status areas in an urban community. Univariate and multiple step-wise linear regression analyses were performed on a sample of 37 census block groups which represented a cross-section of socio-economic levels in Houston, Texas. Six morbidity and mortality rates were the dependent variables while environmental measures from aerial photographs and from the Census formed the two independent variable sets. Environmental data gathered by remote sensing were found to be as good as census data in predicting rates of health outcomes and offered the advantage of flexibility in time and space.*

association to health problems was measured. Detailed health data, as well as census data from several environmental subareas of Houston, Texas were collected and utilized to characterize these subareas and the differences between them. Environmental characteristics were analyzed to determine if health differences among areas could be predicted through the visual expression of remote sensing data.

The analysis was carried out on a socio-economic cross-sectional sample of census

Subareas for analysis consisted of seven census tracts which were divided into 37 block groups. Socio-economic levels ranging from low to high as designated by the City of Houston Health Department were represented, with the preponderance of the block groups falling into medium-low and low socio-economic levels. The analysis was carried out on the block group level because it was felt that the sample tracts were too heterogeneous and concealed environmental differences which appeared at the block group level. Univariate and multiple stepwise linear regressions, correlation coefficients, and spatial distributions by SYMAP were utilized.

Census data were compared to land-use data in predictive ability. The census is an important data source for social and economic information, but it has several shortcomings. Most important, it is gathered only once in every ten years, and it is geographically inflexible, i.e., tract and block group boundaries are fixed. In addition, census housing data ignores neighborhood environmental conditons and provides for the most part subjective evaluation of structures. By contrast, data from aerial photos can be summarized by any desired geographical unit and flights can be flown at any time for frequent monitoring.

Previous Social Applications of Remote Sensing

Only one previous study (Mullens, 1969) attempted to correlate health indices with photographic surrogates and with socioeconomic variables. However, the ability to evaluate housing quality by using remote sensing data has been widely demonstrated, and these studies have important implications for our research since many health studies have demonstrated an association between poverty and disease or poor health status. Studies which utilized aerial photography to evaluate housing quality were reviewed prior to initiating this investigation as well as studies which showed an association between housing quality and health (Rush and Vernon, 1973).

Green (1955, 1956, and 1957) used black-and-white aerial photographs to provide data on the social structure of a city by relating it to the physical spatial structure. He found a statistical association between a number of physical and sociological variables in Birmingham, Alabama. Whereas high statistical association did exist between each of his four physical data categories and several sociological variables, any one photo-data category alone had only a limited predictive value, and so he combined various categories of social and of physical data. The four photo categories of physical structural attributes were combined by using the Guttman Scalogram method (Stouffer, 1950) to construct a "residential desirability scale." A "socio-economic status scale" based on five social data items also was constructed. A correlation of the two scales in the case of Birmingham showed that the "residential desirability scale" accounted for 78 per cent of the variation in the "socio-economic status scale."

Detailed photo interpretation of residential areas was used to delimit poverty areas and was validated by comparison with census data (Manji, 1968). With large-scale photography (about 1:10,000) poverty areas could be identified by certain indicators such as debris, junk-yards, warehouses and small businesses. It also was possible at this scale to evaluate the quality of individual housing units in each block (Mumbower and Donoghue, 1967).

Low-altitude multiband aerial photography was used to identify areas of low housing quality (Weller, 1968). Each factor in the American Public Health Association survey appraisal method was evaluated as a potential item for measurement on aerial photographs together with additional factors considered to be indicators of poor housing quality. The presence of litter, garbage, wrecked or derelict cars, and piles of lumber and rubbish in the neighborhood on both occupied and vacant lots proved to be the best single indicator of low quality housing. Other environmental features associated with low quality included a lack of landscaping together with the presence of weeds on vacant lots, the number of vacant lots, and the existence of nonresidential hazards and nuisances, primarily industrial plants and warehouses.

Bowden (1968) used color-infrared imagery at a scale of 1:60,000 to differentiate the quality of a broad range of residential neighborhoods in an area centered over downtown Los Angeles. When census tract data on income and home values were correlated with residential classifications based on aerial photographs, it was found that four broad categories of housing quality could be identified.

Horton and Marble (1970) and Moore (1970) used data gathered from a survey conducted in Los Angeles by the County Health Department in the spring of 1968 to evaluate housing quality. They used 37 structural and environmental variables which they divided into two groups—those potentially measurable by remote sensing and those not measurable by this method. Analysis showed that "for each basic housing element, the variables acting as indicators of that element tend to be highly correlated with other variables within the element" (Horton and Marble, 1970).

The authors rejected the idea of estimating overall housing quality at the parcel level based only upon remote sensor observation of environmental variables. At the block level, however, they found that the structural variables were associated with a number of environmental variables, primarily those which identify the level of upkeep of lots and the existence of land uses incompatible with residential development. Overall housing quality, then, may be estimated at the block level by using the environmental subset.

Mullens (1969) used low-altitude, large-scale, color-infrared photos to differentiate and classify types of residential areas in Los Angeles on the basis of characteristics of the physical environment. The author investigated the hypothesis that since socio-economic characteristics of large urban populations are associated with specific types of residential environments, it would be possible to associate characteristics of the physical environment with socio-economic variables within residential environments. The residential areas in this investigation were mostly low and middle income housing.

In general, Mullens found that surrogates which were good indicators for one socio-economic variable were also good indices for most of the other socio-economic variables examined. The author grouped the study areas into four categories on the basis of ranking of the areas by socio-economic variables and ranking of the areas by values of photographic surrogates. A high correlation of 0.83 existed between the ranking produced by the total of all socio-economic ranks and the ranking produced by the total of all surrogate ranks.

Mullens also obtained from local public agencies in the Los Angeles area statistics on public health, mental health, crime and delinquency. The rankings between public health as measured by morbidity occurrence rates and photographic surrogates, did not, however, have very strong correlations. The correlations between public health rankings and social rank indicators also were lower than expected. Correlations between mental health ranking and the social rank variables were similar to the correlations between the public health ranking and these social rank variables. The correlations between the mental health ranking and the photographic surrogates increased in general between 0.10 and 0.20 above public health correlations. Statistically significant but low correlations indicated that the relationship between quality of residential area and degree of criminal activity can be observed by using photography, but that photographic surrogates can only broadly separate areas of high crime from areas with lower crime rates.

Senger (1969) sought to test the validity of the relationships established by Mullens between socio-economic characteristics of the urban population and photographic surrogates from color-infrared imagery in a contrasting environmental area. This contrast was expected either to increase the significance of the Los Angeles study or to point out its limitations. The results of Senger's investigation confirmed the validity of the methodology developed by Mullens, but further research is necessary to test the criteria at different scales and in different areas (Senger, 1969; Mullens, 1969; and Davies, et al., 1972).

Davies, et al., (1972) used conventional suborbital black-and-white photography at a scale of about 1:23,000 and suborbital color-infrared photography at a scale of about 1:190,000 to examine a middle and low income residential area in Austin, Texas. The investigators found that poverty areas can be delimited from the imagery and that suitable, environmental indicators of urban blight form useful parameters in determining housing quality.

These investigations demonstrated the capability of utilizing aerial photography to evaluate residential environments. Based on the literature of both remote sensing and health studies, the following empirical generalization was formulated: Patterns of land use and residential quality and density are associated with and act as an influence upon health. This investigation is unique in its attempt to identify health status areas through the use of aerial photography.

Data and Methodology

Physical Environmental Data

Low-level color photography, taken at scales of 1:6,000 and 1:12,000 in October, 1970, was interpreted by an experienced photo interpreter. Eleven land-use categories were identified and residential land use was further broken down into quality and density levels, thus making a total of 22 designations. Density was assigned on the basis of the amount of lot frontage with less than 60′ as high density; 60′-90′, medium density; and 90′ plus, low density. Residential quality was subjectively evaluated and assigned to one of four categories by the photo interpreter: Excellent, Good, Fair, or Poor. His assessment was based on the evaluation of such things as roof condition, lot size, lot coverage, presence of swimming pools, presence of sidewalks and garages, street width and amount of vegetation. Verification of photo interpretation by ground survey was done on a sample basis and photo classifications were found to represent adequately the environmental features.

Quantification of land-use classifications into percentages of total block group area was necessary for analysis. Four ways of deriving area measurements were explored on a sample of three block groups in order to select the most efficient method for measuring land uses in the remaining block groups. The four methods included (1) image analysis slide rule, (2) planimeter, (3) 0.1-inch scaled grid overlay, and (4) systematic point counting (Hsi, 1971). It was found that estimates derived from the systematic point counting technique were as precise as those of the other three methods, and since this method was more efficient than the others, it was chosen. Advantages included the ability to measure irregularly shaped areas, to obtain independent measures of each land-use category, and to obtain rapid measurements in a less tiring way.

Some of the original 22 land-use categories were later modified by combining, separating or blocking their entry into the analytic models. Whenever an original variable was deleted, its values were always included in a combined variable which entered the computer analysis. Originally, residential density and quality were combined but these were later separated in order to determine their relative strength.

CENSUS DATA

Census data were used as a standard data set against which to compare the predictive capability of the land-use data set. Because the analysis was being done on the block group level, indices were limited to those which were reported by blocks. Several socio-economic indicators and density measures were available at this level from the 1970 Census. Quality indicators consisted of average rent, average value owner-occupied housing, percentage of rented dwellings, and computed property value. Density or overcrowding was measured by the percentage of households with greater than 1.01 persons per room, the percentage of one person households, percentage of population under 18, a ratio of total number of dwelling units to total population, and a ratio of total population to total number of dwelling units.

HEALTH DATA

Health data were collected from three agencies and consisted of a standardized age-adjusted mortality rate, morbidity rates for tuberculosis, shigella and salmonella, infectious hepatitis and meningitis, and rates of juvenile delinquency, and mental health. Data were gathered over a time span of from three to eight years with the exception of mortality data which were easily available for only one year at the block group level. Crude disease rates per 1,000 population were computed for all morbidity and behavioral data. Mortality data were standardized on the basis of sex, and three age groups by the indirect method. The difference between the crude death rate for each block group and the age-standardized death rate for that block group was used in the computer analysis. Most diseases were not well reported by private physicians, and this fact probably biased the data toward users of public health clinics.

METHODOLOGY

Multi-variate step-wise linear regression models were constructed using the BMDO2R computer program. In step-wise regression the computer program considered one variable at a time, choosing a sequence according to the highest correlation between each individual independent variable and the dependent variable. Thus, the program began with the correlation matrix. In subsequent steps the correlation became multiple, i.e., the correlation between the dependent variable and all independent variables hav-

TABLE 1. CUMULATIVE MULTIPLE CORRELATION COEFFICIENT OF FIRST TEN VARIABLES IN STEP-WISE REGRESSION MODELS.

Health Outcome	Mixed Data Model		Photo Model		Census Model	
	R	R^2	R	R^2	R	R^2
Mortality	0.84	0.70	0.65	0.42	0.70	0.50
Tuberculosis	0.89	0.79	0.67	0.45	0.81	0.65
Shigella/Salmonella	0.95	0.90	0.82	0.68	0.85	0.72
Infectious Hepatitis/Meningitis	0.77	0.60	0.71	0.50	0.49	0.24
Juvenile Delinquency	0.85	0.72	0.76	0.57	0.74	0.55
Mental Health Referrals	0.84*	0.70*	0.82*	0.66*	0.66*	0.44*

* Figures based on Step #11 because a preceding variable dropped out of the equation.

ing entered the equation. The effect of adding a variable at each step was to re-examine the variables incorporated into the model in previous steps and to alter the earlier relationship between the entered variables and the dependent variable so as to decrease their individual effect.

Any independent variable *(X)* or group of *X*'s may not be related to the dependent variable *(Y)* in a straight line. Because this was a linear model, if a group of *X*'s were related to *Y* in a curvilinear relationship, the computer tended to pick a new *X* variable to compensate for the curve and make it linear. Thus, caution must be taken in attributing explanatory power to every entering variable.

Results and Analysis

Associations of environmental data from the Census and from aerial photographs with health data were measured both together and in separate computer models so that comparisons could be made among the three. The objective was to determine if data from aerial photographs were as highly related to health outcomes as the more traditional census data. Six relationships between health and environment were explored. Patterns of land use and residential density and quality were related to

1. mortality rates
2. incidence of tuberculosis, a respiratory disease
3. - 4. incidence of shigella and salmonella, and infectious hepatitis and meningitis, communicable diseases.
5. referral rates of delinquents
6. prevalence of mental health referrals.

Table 2. Rank Order of 10 Variables in Mixed Data Model with Each Health Outcome.

	Model 1					
Photo Variables	Mort	TB	S/S	H/M	JD	MH
*R Excellent	9					10
R Medium						
R Good						
R Medium Good	10				4	
R High						
R Fair						
R Medium Fair						
R High Fair						2
R High Poor			9			
Apartments						6
Commercial					9	
Industry				6		
Unimproved				1		
Park			2	3	7	3
Green Space	8			5		5
School	3	6		4	1	
Church	4	5			3	4
Cemetery						
Streets				2		
Census Variables						
Avg. Rent		10			10	
*Avg. Value Own Occ	7					8
Total Dwelling Units		1	5			
% Blacks					8	
*% Hsdh>1.01 pop/rm		9	8		2	
*% 1 Person Hshd	5	3		10	6	
% Under Age 18			1	9		
% Over Age 62		7	10			
% Hshd Female Head						
Total People	2		4			7
% Rented Hshd	1	2	6		5	1†
*Du/Pop				8		
Pop/Du	6	4	7	7		
Computed Value Hshd		8	3			

† Variable removed in 9th step.
* R = Residential; Own Occ = Owner Occupied; Hshd = Household; Du/Pop = Total Dwelling Units ÷ Total Population.

Three multi-variate step-wise linear regression models were constructed (Table 1). The Mixed Data Model allowed free choice of either census or photographic variables to best predict each health outcome. The Photo Model used data from interpreted aerial photographs as the independent variables for each health outcome and the Census Model used only census data as independent variables.

These Models were designed to determine how much variability of health data could be explained after a limited number of independent variables entered the equation. In this case, ten steps were arbitrarily chosen. The Mixed Data Model served as a measure of the maximum explanatory potential using both sets of environmental data. Each model was compared to the Mixed Data Model to determine the ability of a single set of environmental measures in predicting health outcomes.

Multiple regression coefficients and percentages of explained variation were recorded in Table 1 after completion of ten steps in the step-wise regression equations. Table 1 indicates that with free choice of photographic and census variables 60 to 90 per cent of the variation in health outcomes was explained after only ten steps. More of the variation was accounted for by free selection from both variable sets than by either set alone.

In the separate models, the explained variation was maximized for three health outcomes by use of photographic data. The photographic variables were better predictors of infectious hepatitis/meningitis, juvenile delinquency, and mental health referrals. For two of these, the photo variables explained 26 per cent and 22 per cent more of the variations than did census variables.

Census data were better predictors for mortality, tuberculosis, and shigella/salmonella. Census variables accounted for 21 per cent more of the variability in the tuberculosis data than photo variables but were only slightly more sensitive than photo data for the other two health outcomes.

Table 2 illustrates the order of entry of variables in the Mixed Data Model. Census variables clearly entered more frequently than photo variables in predicting tuberculosis and shigella/salmonella. However, in the separate models census data explained only 4 per cent more of the variance than photo data for shigella/salmonella. The photo variables entered more frequently in predicting infectious hepatitis/meningitis and mental health referrals. Equal numbers of census and photo variables entered the equations for mortality and juvenile delinquency predictions, health outcomes with little predictive difference in the separate models.

Conclusion

Neither set of independent variables derived from the Census or photographs was clearly superior to the other for predicting the selected health outcomes. Environmental data gathered by remote sensing was found to be as good as census data in determining rates of health outcomes. In this case remote sensing was found to be an effective tool useful in identifying public health problem areas in an urban community. The applicability of this technique is being further validated in a smaller, heterogeneous urban community. Health planners may find aerial photography a faster and more current source of data than present methods such as the Census and other surveys.

Acknowledgment

This study was partially funded by the National Aeronautics and Space Administration, Johnson Spacecraft Center, Houston, Texas, under Contract NAS9-12823.

References

Bowden, L. W., *Multi-Sensor Signatures of Urban Morphology, Function and Evolution.* Technical Report No. 2, U.S.D.A. Contract No. 14-08-0001-10674, Departments of Geography, University of California at Los Angeles and Riverside, 1968.

Davies, S., Tuyahov, A., and Holz, R. K., "Use of Remote Sensing to Determine Urban Poverty Neighborhoods," *The Surveillant Science, Remote Sensing of the Environment,* Ed. Robert K. Holz, Houghton, Mifflin Co., Boston, 1972.

Green, N. E., *Aerial Photography in the Analysis of Urban Structure, Ecological and Social,* (Unpublished Ph.D. Dissertation, Department of Sociology, University of North Carolina, June, 1955).

Green, N. E., "Scale Analysis of Urban Structures," *ASR*, XXI: 8-13, February, 1956.

Green, N. E., "Aerial Photo Interpretation of the Social Structure of the City," *Photogrammetric Engineering*, 23: 89-96, March, 1957.

Horton, F. E. and Marble, D. F., "Housing Quality in Urban Areas: Data Acquisition and Classification Through the Analysis of Remote Sensor Imagery," *Second Annual Earth Resources Aircraft Program Status Review*, I: Part 15: 1-13. Houston, Texas: National Aeronautics and Space Administration, 1970.

Hsi, B. P., "Some Statistical Aspects of Point-counting Technique in Plane Sampling," *The Journal of Biomedical Systems*, 2, No. 7, 1971.

Manji, A. S., *The Uses of Conventional Aerial Photography in Urban Areas: A Review*. Technical Letter NASA-131, U.S.G.S., Washington, D.C., 1968.

Moore, E. G., "Application of Remote Sensors to the Classification of Areal Data at Different Scales: A Case Study in Housing Quality," *Remote Sensing of Environment*, 1:109-121, 1970.

Mullens, R. H., Jr., "Analysis of Urban Residential Environments Using Color Infrared Aerial Photography: An Examination of Socio-economic Variables and Physical Characteristics of Selected Areas in the Los Angeles Basin." *Interagency Report*, NASA-153, 1969.

Mumbower, L. E. and Donoghue, J., "Urban Poverty Study" *Photogrammetric Engineering* 33: 610-618, June, 1967.

Rush, M. and Vernon, S., *Urban Environmental Health and Remote Sensing*, Contract NAS9-12823, The University of Texas Health Science Center at Houston, Texas, School of Public Health, December, 1973, (unpublished).

Senger, L. W., III, *An Application of the Concepts of the L.A. Residential Environment Study to the Ontario-Upland Region of California*, Prepared by U.S.G.S. for NASA, under Contract No. W12570, Task No. 160-75-01-32-10. Work performed under U.S.G.S. Geographic Applications Program Contract No. 14-08-00001-10674, with the University of California, May, 1969.

Stouffer, S. A., et al., *Measurement and Prediction*, Princeton: Princeton University Press, 1950.

Wellar, B. S., "Generation of Housing Quality Data From Multiband Aerial Photographs," *Interagency Report*, NASA-119, U.S.G.S., Washington, D.C., 1968.

(Continued from Page 1144)

tions on geodetic disciplines in the context of present development and requirements.

Session IV: Moderator: A. J. van der Weele

H. P. Zakasnow (U.S.S.R.) (presentation by E. P. Arzanow): The course of optics for photogrammetrists (A film presentation with commentary).

J. Hothmer (G.F.R.): Negotiation of contracts.

Session V: Moderator: G. C. Tewinkel

A. Tarczy-Hornoch (Hungary): Contributions to the history of photogrammetry in Hungary.

H. L. Rogge (Netherlands) (paper read by A. J. van der Weele): International Bibliography of Photogrammetry and the Universal Decimal Classification.

H. Schmidt-Falkenberg (G.F.R.) (paper read by J. Hothmer): Photogrammetric Information and Documentation at the Applied Geodesy Institute in Frankfurt.

J. Jachimski and A. Dzierzega (Poland): Some remarks on the efficacy of the contemporary photogrammetric periodicals.

Session VI: Moderator: G. C. Tewinkel

J. Pietschner (G.D.R.): Activity Report of Working Group VI-3; "Terminology."

The technical and business sessions concluded with remarks and comments from S. G. Gamble, ISP President; Z. Sitek, Commission VI President; and the Director of the Polish Society of Photogrammetry.

—Sanjib K. Ghosh