SPATIAL HETEROGENEITY OF WATER QUALITY PARAMETERS

by

S.R. Esterby

NWRI CONTRIBUTION #86-70

National Water Research Institute
Canada Centre for Inland Waters
867 Lakeshore Road, P.O. Box 5050
Burlington, Ontario L7R 4A6

Hétérogénéité spatiale des paramètres de la qualité de l'eau

S.R. Esterby
Division de l'écologie aquatique
Institut national de recherche sur les eaux

L'élément d'espace, présentant généralement des aspects structurel et stochastique, joue un rôle important quant à la variabilité des mesures de la qualité de l'eau dans plusieurs genres d'études diverses. Il est tout aussi nécessaire de détecter l'hétérogénéité spatiale que d'estimer cet élément. Les objectifs d'une analyse de l'hétérogénéité spatiale peuvent se diviser en deux types généraux : 1) la caractérisation de l'hétérogénéité soit par regroupement des zones, soit par l'expression d'une variable de la qualité de l'eau qui est fonction de la zone et 2) l'essai et l'estimation de l'hétérogénéité pour des zones bien définies. Dans le premier cas, la caractérisation est obtenue au moyen de méthodes graphiques et de méthodes de regroupement et d'analyse du comportement des surfaces, suivies d'une estimation à l'intérieur des divisions spatiales ou d'une estimation des fonctions. Dans le deuxième cas, il faut utiliser des méthodes comme l'analyse de la variance et des méthodes non paramétriques analogues qui donnent une idée de la zone. Le document traitera d'abord de la nature des ensembles de données et des objectifs de l'analyse, puis des méthodes statistiques. Il renfermera des exemples de l'analyse statistique des données sur la qualité de l'eau.

RÉSUMÉ ADMINISTRATIF

En analysant les données sur la qualité de l'eau, on s'est penché davantage sur les méthodes statistiques pour déterminer les tendances temporelles plutôt que l'hétérogénéité spatiale. Cette dernière est tout aussi importante, même lorsqu'on se préoccupe surtout d'une certaine tendance temporelle, car il ne faut pas confondre les facteurs autres que le temps à l'effet temporel. Le présent document décrit les méthodes qui conviennent pour analyser les données sur la qualité de l'eau, y compris celles qu'on utilise déjà et celles qu'on aurait intérêt à utiliser.

EXECUTIVE SUMMARY

In the analysis of water quality data, more attention has been given to statistical methods for the determination of time trends than for the determination of spatial heterogeneity. The latter is equally important, even when a time trend is of primary concern, since factors other than time must not be mixed up with the time effect. The present paper describes methods suitable for the analysis of water quality data, including methods already in use and methods which could be used to advantage.

SPATIAL HETEROGENEITY OF WATER QUALITY PARAMETERS

S.R. ESTERBY
National Water Research Institute, Canada Centre for Inland Waters, Burlington, Ontario, Canada

1. INTRODUCTION

In water quality monitoring, it is difficult to think of a situation in which a decision does not have to be made about where to collect samples. Thus, at least potentially, all water quality measurements have a spatial component. In general, knowledge of how water quality parameters vary over the region of interest is necessary to the understanding of a system. It may be of primary interest, as, for example, in the study of the transport of a pollutant, or of secondary interest in that it may be necessary to remove the spatial component in order to detect changes over time. Dependent upon our state of knowledge, the analysis of water quality data for the spatial component can be placed under one of the following objectives: 1) characterization of heterogeneity, or 2) testing for and estimation of a well-defined spatial component. In this paper, only the first objective will be considered.

The nature of the data and the reasons for the analysis will determine the methods used for characterizing heterogeneity. Three broad classes of procedures are available: 1) grouping methods, 2) spatial autocorrelation methods and 3) methods which involve fitting a function for the relationship between a water quality parameter and location, with or without the assumption of independence.

In the present paper, a general overview of methods suitable for water quality studies has been attempted. The topic is clearly too extensive to review here. Although the discussion has been limited to spatial variation, no suggestion that analysis for spatial and temporal variability are entirely separable is intended. Haugh (1984) suggests directions such

analyses might take for regularly spaced sample points based on extensions of the Box and Jenkins approach to time series analysis.

It will be useful to consider an example in which spatial heterogeneity is important before discussing methodology. Such an example is the issue (Barica, 1982) of anoxic conditions in Lake Erie (for a brief review see Kwiatkowski, 1984). An objective of the 1978 Great Lakes Water Quality Agreement between Canada and the United States was the restoration of year round aerobic conditions in the bottom waters of the Central Basin of Lake Erie. The historic records of hypolimnetic dissolved oxygen have been examined for the existence of a time trend by several authors (Dobson and Gilbertson, 1971; Charlton, 1979; and Rosa and Burns, 1981). Data selection was considered essential by all three sets of authors (Anderson et al., 1984). Dobson and Gilbertson (1971) used dissolved oxygen concentrations from samples for which the temperature was within 3°C of the minimum. Charlton (1979) used only near-bottom values at stations that had a depth over 15 m, were stratified and showed no evidence of incursion of Eastern Basin water. Burns and Rosa (1981) attempted to establish a representative, homogeneous area by calculating a depletion rate distribution map. Some of these criteria are clearly spatial considerations. Others are reasons why spatial variation would exist. This raises an additional point. Spatial heterogeneity can sometimes be associated with a supplementary measurement such as temperature. Inclusion of this additional variable in the model may make simplifying assumptions, such as independence of errors, tenable and permit the use of conventional statistical methods.

2. DATA SETS AND NOTATION

Notation is given here (Table 1) for a data set collected over both space and time and containing a number of water quality parameters and descriptors of the location or water mass, which will be called supplementary observations in this paper. The primary use for the notation will be to show which dimensions of such a general data set can be or are normally handled by a particular method of analysis. The matrix of measurements on p water quality parameters at I stations at

TABLE 1
Description of a general water quality data set.

Description of Rows			Subscripts of Y Matrix				Subscripts of X Matrix[a]			
Station	Depth	Row	Water Quality Parameter				Supplementary Observation			
			1	2	...	p	1	2	...	q
1	1	1	1,1	1,2	...	1,p	[b]			
	2	2	2,1	2,2	...	2,p				
	.	.	.	.		.				
	.	.	.	.		.				
	.	.	.	.		.				
	n_1	n_1	n_1,1	n_1,2	...	n_1,p				
2	1	n_1+1	n_1+1,1	n_1+1,2	...	n_1+1,p				
	2	n_1+2	n_1+2,1	n_1+2,2	...	n_1+2,p				
	.	.	.	.		.				
	.	.	.	.		.				
	.	.	.	.		.				
	n_2	n_1+n_2	n_1+n_2,1	n_1+n_2,2	...	n_1+n_2,p				
	.	.	.	.		.				
	.	.	.	.		.				
	.	.	.	.		.				
I	1	$n-n_I+1$	$n-n_I+1$,1	$n-n_I+1$,2	...	$n-n_I+1$,p				
	2	$n-n_I+2$	$n-n_I+2$,1	$n-n_I+2$,2	...	$n-n_I+2$,p				
	.	.	.	.		.				
	.	.	.	.		.				
	.	.	.	.		.				
	n_I	n	n,1	n,2		n,p				
Vector notation, length=n			$\underline{y}_1$	$\underline{y}_2$	...	$\underline{y}_p$	$\underline{x}_1$	$\underline{x}_2$	...	$\underline{x}_q$
Vector notation, length=I[c]			$\underline{y}_{1\,1}$	$\underline{y}_{2\,1}$	...	$\underline{y}_{p\,1}$	$\underline{x}_{1\,1}$	$\underline{x}_{2\,1}$	...	$\underline{x}_{q\,1}$
Vector notation, length=I[d] derived variable			$\underline{z}_1$	$\underline{z}_2$	...	$\underline{z}_p$	$\underline{u}_1$	$\underline{u}_2$	...	$\underline{u}_q$

[a] Subscripts of matrix X take the same form as those of matrix Y.

[b] Assuming the coordinates of the stations appear in columns 1 and 2 of the X matrix, the elements of a column denoted by unbroken vertical lines are equal since they correspond to the coordinate for the same station.

[c] The second subscript indicates that the vector includes only the elements corresponding to the first depth for each station.

[d] An example of a derived variable is some function of the values of the original variable for all depths sampled at a particular station.

time t is denoted by $Y_t = \{y_{ij}\}_t$ and the matrix of q corresponding supplementary observations by $X_t = \{x_{ik}\}_t$. The number of values for each parameter is n since samples are assumed to be collected at n_i depths at the ith station and

$$n = \sum_{i=1}^{I} n_i .$$

The subscript t will be dropped to simplify the notation, but it is implicit. For ease of description, it will be assumed that the first two columns of X contain the coordinates of the station, and the third column the depth of the measurement.

3. THE IMPORTANCE OF CHARACTERIZING SPATIAL HETEROGENEITY

The idea, that, if sampling is being conducted over time and space, the variability in space should be accounted for in the design of the sampling program and the analysis of the data, is so fundamental that it needs no elaboration. Two of the objectives of statistical design and analysis which are achievable in environmental field studies are the increase in precision by removing as much as possible from the error term and the elimination of bias. An ignored spatial component could enter either as large variability or as bias. Although the examples given below are for other components, the comments are equally applicable to spatial variability.

In the first example the day to day variability was overlooked at the design stage. In studying spatial heterogeneity of phytoplankton, Platt et al. (1970) drew the conclusion that the between-station variance rose as sampling area increased to a density of ten stations per mile, then remained relatively constant. By identifying the points in their Figure 1 by date and replotting in the original variance units (Figure 1), one sees that estimates of between-station variance which were obtained on the same day are approximately equal. The impression is obtained that the date of sampling is important in sorting out the reasons for different between-station variances. However, since variances based on markedly different distances between stations were obtained for one day only, the day-to-day variation in the between-station variances cannot be assessed.

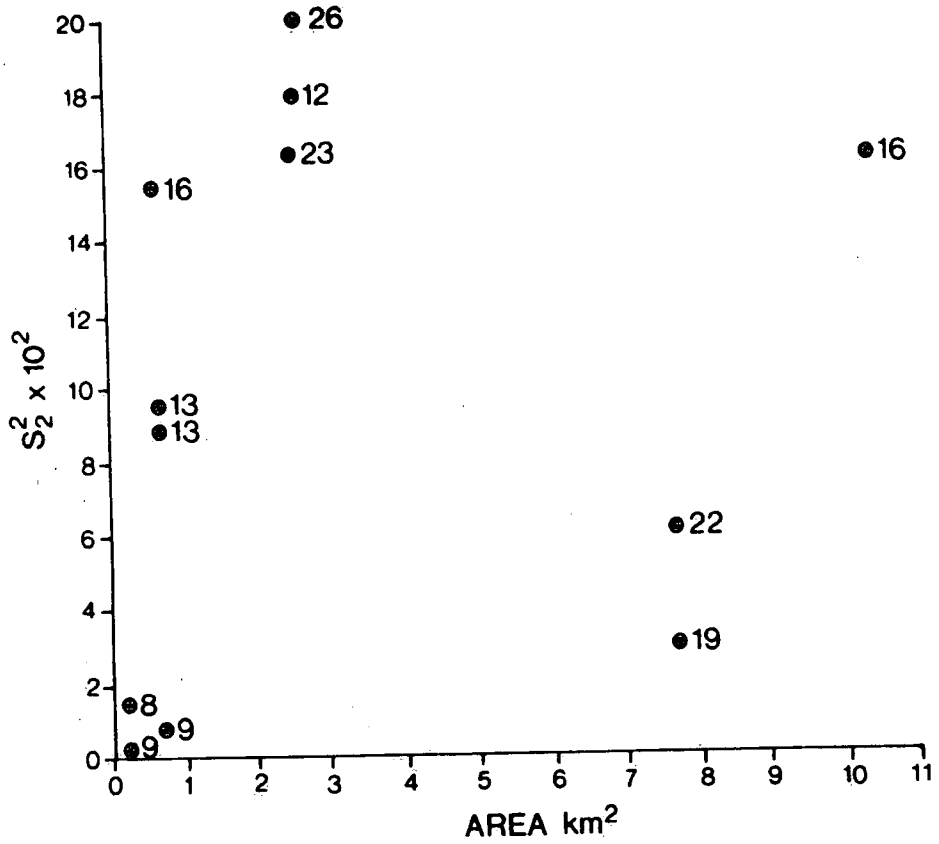

Fig. 1 Plot of between-station variance and sample area as reported by Platt et al. (1970). Numbers give the day of sampling.

An example of adjusting for a supplementary observation, by the inclusion of an additional term in a model of a time trend, is given by El-Shaarawi (1984) in further work related to the issue of Lake Erie anoxia. Water level was treated as a covariate in the analysis of oxygen depletion rates for the existence of a time trend. This provides a statistical solution to the problem of correcting for hypolimnion thickness which had been identified earlier by Charlton (1979) and Rosa and Burns (1981).

4. GROUPING PROCEDURES

In the present context, the common element to the procedures in this broad class is the division of a set of points in space, usually sampling stations, into two or more groups such that members in the same group are more similar to each other than to members in other groups with respect to one or more

water quality parameters. Included are geometrical and clustering methods. The geometrical methods consist of plotting points in low-dimensional Euclidean space so that points which are similar to one another occur close together. Clustering procedures use a mathematical criterion to partition the set of points into homogeneous groups. The two types of methods are complementary since the clustering procedures provide objectivity and the geometrical procedures display natural groupings (Gordon, 1981). To discover the structure in a data set, more than one procedure will often be needed.

These procedures are widely used in ecology and recognition of their importance is to be found in the book by Pielou (1984), which is devoted entirely to this topic. The geometric method, factor analysis, is most frequently used method in geology (Jöreskog et al., 1976). Books concerned more with the methods, than with applications in a particular discipline, include Anderberg (1973), Hartigan (1975), Gordon (1981) and Lebart et al. (1984).

With a few exceptions, these procedures divide the data set into groups strictly on the basis of the values of the parameters and no information about location is used in this categorization. The examination of the spatial distribution of the members of the groups is done after the categorization is complete, usually by plotting the groups on a map or other diagram which represents the location in space. Constrained clustering methods, which require members of a cluster to be spatially contiguous, are available for data on which a linear ordering has been imposed (e.g. transect or depth profile data) but other spatial arrangements are much more complicated. Constrained clustering methods have been applied to pollen percentages for sediment cores (Gordon, 1981).

Grouping procedures are generally applied to the values of one or more water quality parameters or some combination of water quality parameters and supplementary observations at one sampling depth, or to derived values corresponding to a specified part of the water column. For example, the matrices $Y = (\underline{y}_{11}\ \underline{y}_{21}\ \cdots\ \underline{y}_{p1})$ or $Z = (\underline{z}_1\ \underline{z}_2\ \cdots\ \underline{z}_p)$ might be used (Table 1). A simple example of a derived variable is the mean of the concentration for all depths sampled at a station.

4.1 Examples of grouping based on water quality data

Clustering methods have been developed for the zonation of a lake into regions, homogeneous with respect to the level of one or more water quality parameters. A clustering method, which defines a homogeneous set of concentrations as one which can be fitted by a Poisson distribution, was applied to surface coliform concentrations at sampling stations on Lake Erie, cruise by cruise (El-Shaarawi et al., 1981; and Esterby and El-Shaarawi, 1984). A spatial pattern which changed with season was found to be qualitatively consistent from year to year, whereas apparent biases from year to year made comparison of concentrations impossible. This is an example where a grouping procedure was a natural choice due to the discontinuous nature of the spatial regions of similar concentration. A second procedure, suitable for a larger class of water quality parameters, is based on a linear additive model for the station and cruise components of data collected in one year (El-Shaarawi and Shah, 1978). Allowance is made for non-orthogonality and the need to transform the data. Provided the station component is significantly different from zero, a criterion based on the change in the residual sum of squares is used to group stations. The procedure has been used mostly for a single water quality parameter (.e.g El-Shaarawi and Kwiatkowski, 1977) but the multivariate extension was also given by El-Shaarawi and Shah.

Existing data on water quality parameters related to acid precipitation for sets of lakes in regions of Eastern Canada have been analyzed by grouping methods (El-Shaarawi et al., 1985; and Haemmerli and Bobée, this volume) to determine subsets of similar lakes which can subsequently be used in the design of future data collections. El-Shaarawi et al. applied both a graphical method, in which the lakes were plotted on the first three principal components, and a k-means non-hierarchical nearest-centroid clustering method to determine groups of lakes. Plots of the clusters in space, the percentage of variation explained by a given number of clusters, and descriptions of the variables within clusters, by means of summary statistics, histograms and Q-Q plots, were used to select the number of clusters and to characterize the clusters.

Two examples of the use of clustering methods which are considerably different from the above examples are the following. The first example is another analysis of dissolved oxygen concentrations in Lake Erie, the issue which was discussed in the introduction. Anderson et al. (1984) used a clustering method on three-dimensional spatial data, consisting of the temperature and the dissolved oxygen concentration of surface and bottom samples at each station, to divide these points into groups which were assumed to be either hypolimnetic or not hypolimnetic. El-Shaarawi et al. (1985) grouped stations on the Niagara River by applying a complete linkage clustering method to the standardized Euclidean distances between stations calculated separately from the ranks of chlorinated organic substances in water and in suspended sediments. Ranks were used as a means of dealing with values below the detection limit and, for each phase, the stations were ranked separately for each substance.

5. SPATIAL AUTOCORRELATION METHODS

If the value of a variable depends upon the values of the same variable at neighbouring locations, then spatial autocorrelation exists. The problem of determining whether autocorrelation exists is more difficult for spatial data than for a time series since dependence may extend in all directions for the spatial data but only into the past for the time series. The monograph by Cliff and Ord (1973) provides a thorough treatment of the spatial autocorrelation measures which were introduced earlier by Moran and Geary. Although only examples from geography were used by Cliff and Ord, the methods are applicable to any fixed set of points in space.

Let $\underline{y}' = [y_1, y_2, \ldots, y_n]$ be the vector of values of a variable at m points in space, for example, $\underline{y}$ could be $\underline{y}_1$, $\underline{y}_{11}$ or $\underline{z}_1$ corresponding to the first water quality parameter of matrix Y (Table 1). Then the general forms of the Moran and Geary spatial autocorrelation coefficients are

$$I = \left[m \sum_{\substack{i=1 \\ i \neq j}}^{m} \sum_{j=1}^{m} w_{ij} (y_i - \bar{y}) (y_j - \bar{y})\right] / W \sum_{i=1}^{m} (y_i - \bar{y})^2 \qquad (1)$$

and

$$c = \left(\frac{m-1}{2W}\right)\left[\sum_{\substack{i=1\\ i\neq j}}^{m}\sum_{j=1}^{m} w_{ij}\,(y_i - y_j)^2\right] \Big/ \sum_{i=1}^{m}(y_i - \bar{y})^2 \qquad (2)$$

respectively, where

$$\bar{y} = \left(\sum_{i=1}^{m} y_i\right)/m \qquad \text{and} \qquad W = \sum_{\substack{i=1\\ i\neq j}}^{m}\sum_{j=1}^{m} w_{ij} \ .$$

The spatial information enters only through the matrix of weights $\{w_{ij}\}$. The test for spatial autocorrelation is in the form of the test of the hypothesis of no spatial autocorrelation, that is a random distribution in space, against an alternative as specified by the matrix $\{w_{ij}\}$.

Jumars et al. (1977) first applied this method in the field of ecology. These authors recognized that the method uses the spatial information in the data, which is ignored in tests such as Fisher's index of dispersion, but is less restrictive than spectral analysis which requires intensive and regularly-spaced samples. The matrix $\{w_{ij}\}$ requires the definition of which pairs are connected, in some sense, and either a measure of adjacency or of distance between pairs of connected localities (Sokal, 1979). When no a priori distance is available, a spatial correlogram, obtained using the unweighted autocorrelation coefficient, can be used to suggest the form. A more recent application and a list of other ecological applications are given by Mackas (1984).

One of the applications of the above method is in the estimation of patch diameters. Spectral analysis has been used in ecology for this purpose. Two examples are Platt et al. (1970) and Lekan and Wilson (1978). In both cases, densely and regularly sampled transects with gradients of temperature and salinity were used and the objectives included inferences about patch size or length scale of phytoplankton. Platt et al. acknowledged that this method would underestimate patch size since the transect may not pass through the widest part of the patches.

As far as the author can determine, the methods based on the Moran and Geary spatial autocorrelation coefficients have not been applied to water quality parameters other than chlorophyl

(Mackas, 1984), which is the same indicator of phytoplankton biomass used in the spectral analyses by Platt et al. (1978) and Lekan and Wilson (1978). The form of $\{w_{ij}\}$ needs to be considered for each new application and for water quality parameters, the literature on ecological applications is more relevant than that on geographical applications.

6. THE PARAMETER EXPRESSED AS A FUNCTION OF POSITION

Methods of the previous section provide indirect inferences about the response surface. In this section, the methods are based upon an explicit relationship between a variable and its position. However, there is considerable overlap between the two sections since spatial autocorrelation is also considered here.

As in the previous section, procedures are univariate. Consider again $\underline{y}' = [y_1, y_2, \ldots, y_m]$, the vector of values of a variable at m points in space which could be $\underline{y}_1$, $\underline{y}_{11}$ or $\underline{z}_1$ (Table 1). The vector $\underline{y}'$ gives values of either a water quality parameter or a variable derived from a water quality parameter, but the term parameter will be used only for parameters of a model in this section. Let $\underline{x}_i = (x_{i1}, x_{i2})$ or $\underline{x}_i = (x_{i1}, x_{i2}, x_{i3})$ provide the location of the ith variate value in two- or three-dimensional space for $i = 1,2,\ldots,m$. Then the value of the variable at the ith location, expressed as the sum of deterministic and random terms, is

$$y_i = f(\underline{x}_i) + U(X) \qquad (3)$$

where X is a matrix of m rows given by the row vectors $\underline{x}_i$. If the value of the random term at location i does not depend upon the value at nearby locations,

$$y_i = f(\underline{x}_i) + e_i \qquad (4)$$

where the e_i's are uncorrelated. For f linear in the parameters, ordinary least squares can be used to fit the model in (4). In the more general form, (3), some assumption about stationarity is required. If the dispersion matrix is assumed

known, generalized least squares can be applied (Rao, 1965, p. 180).

Hand-drawn contour maps are often used to display spatial variability. To avoid the subjectivity of this procedure, geologists (among others) have adopted several procedures, based on well-known statistical methods which assume the variable can be expressed as a function of position. The procedure which consists of 1) fitting a polynomial regression model of the variable on the spatial coordinates, 2) using the regression model to estimate the mean value of the variable at grid points and 3) constructing contours of constant value using some numerical interpolation procedure is called trend surface analysis (Davis, 1973). A procedure, which accounts for spatial autocorrelation, known as kriging, was developed by French geomathematicians led by Matheron. A useful reference on this method is the paper by Watson (1971), in which he relates the technique to the English language statistical literature and thus, in effect, defines the new terminology using standard statistical terms. In kriging, the observations are assumed correlated, and the correlation takes the form of a known parametric function called the variogram. Contours are constructed by predicting the value of the variable at unobserved points. Watson (1971) shows that this kriging estimator is the best-linear unbiased estimator (BLUE) for prediction of y_p at an unobserved point p, and that it consists of the sum of the BLUE for the mean value, $E(y_p)$, at the unobserved point p and a term which accounts for the dependence on nearby values. Note that this distinction between estimating the mean value at an unobserved point and the value of an individual observation is analogous to what is done in regression (Draper and Smith, 1981, pp. 28-31). Equation (3) covers kriging under the "intrinsic hypothesis", i.e. stationarity of the mean, and universal kriging, i.e. the case with drift, (Delhomme, 1978) since $f(\underline{x}_i)$ reduces to a constant in the first case.

Delhomme (1978) states that an advantage to kriging is that it provides an estimate of precision which most techniques for contouring, including least squares, do not. As can be seen from the analogy with regression which was given above, this is not so. The method of ordinary least squares provides an

estimate of the precision, however, the method of trend surface analysis, as defined above, does not use this capability. El-Shaarawi and Esterby (1981) have shown how this can be done to construct contours of constant value, for either the mean or an individual observation, and to attach confidence bands to these contours. The method was applied to surface temperature data from Lake Erie.

Whatever the assumption about the form of the autocorrelation, the difficulty is in estimation of the dispersion matrix. This requires estimation of the variogram in kriging (Delhomme, 1978) which is usually done by ad hoc procedures. Iterative procedures for regression with correlated errors are given by Cliff and Ord (1973) and Cook and Pocock (1983). Maximum likelihood methods of estimation are considered by Cook and Pocock (1983) and Mardia and Marshall (1984).

Methods of testing for spatial autocorrelation in regression residuals are given by Cliff and Ord (1973) and have been applied in paleoecology (Howe et al., 1984). Cliff and Ord (1973) stress the fact that detection of autocorrelation in the residuals may be due to one of the following: 1) an inadequate form for the relationship between dependent and independent variables, such as using a linear model when curvature is present, 2) omission of one or more regressors and 3) the need for autocorrelation structure in the model. Clearly, in 1) and 2), means of removing the autocorrelation from the residuals exist, which are simpler than the methods incorporating spatial autocorrelation.

Data collected in space need not exhibit spatial autocorrelation for various reasons. It may not be detectable because distances between the points in space are larger than the distance within which dependence occurs. Cook and Pocock (1983) discuss aggregation to remove correlation. Analogously, both spacing and the use of means or medians over seasons have been suggested as methods of reducing serial correlation in the analysis of water quality data for temporal trends (van Belle and Hughes, 1984). The consequences of using ordinary least squares, when errors are correlated, are inefficient estimators of the regression parameters and a downwards biased estimator of the variance with the latter resulting in an overestimate of the significance of the regression (Cliff and Ord, 1973).

Methods in this section also provide the capability of expressing the variabile as the sum of spatial components, temporal components and other explanatory variables such as temperature. An example of this, which encompasses many of the points discussed in this section, is the complicated model used by Eynon and Switzer (1983) to construct contour maps of rainfall pH.

7. DISCUSSION

The many dimensions of water quality data sets make analysis difficult. Data is often collected to meet objectives related to monitoring the change in water quality conditions which are necessarily too general to be of help in reducing the dimensions of the problem. Thus, cluster analysis and related methods, which do not use the spatial location but can be used to examine the structure of multivariate data, are complementary to the univariate methods which do use the spatial location. The analyst can expect to use the classes of methods discussed here in an iterative fashion, coupled with scientific understanding of the system, to arrive at a characterization of spatial structure.

Of the methods discussed, only the grouping procedures are strictly for the purpose of discovering structure in the data. The other methods, even in the characterization stage, are used for testing hypotheses and estimation.

REFERENCES

Anderberg, M.R., 1973. Cluster Analysis for Applications. Academic Press, New York, 359 pp.

Anderson, J.E., El-Shaarawi, A.H., Esterby, S.R. and Unny, T.E., 1984. Dissolved oxygen concentrations in Lake Erie (U.S.A.-Canada), 1. Study of spatial and temporal variability using cluster and regression analysis. J. Hydrol., 72: 209-229.

Barica, J., 1982. Lake Erie depletion controversy. J. Great Lakes Res., 8(4): 719-722.

Charlton, M.N., 1979. Hypolimnetic oxygen depletion in central Lake Erie: Has there been any change? Inland Waters Dir., Sci. Publ. Ser. No. 110, 25 pp.

Cliff, A.D. and Ord, J.K., 1973. Spatial Autocorrelation. Methuen, New York, 178 pp.

Cook, D.G. and Pocock, S.J., 1983. Multiple regression in geograhical mortality studies, with allowance for spatially correlated errors. Biometrics, 39: 361-371.

Davis, J.C., 1973. Statistics and Data Analysis in Geology. Wiley, New York.

Delhomme, J.P., 1978. Kriging in the hydrosciences. Adv. Water Resources, 1: 251-266.

Dobson, H.H. and Gilbertson, M., 1971. Oxygen depletion in the hypolimnion of the Central Basin of Lake Erie, 1929-1970. Proc. 14th Conf. Great Lakes Res., Int. Assoc. Great Lakes Res., pp. 743-748.

Draper, N.R. and Smith, H., 1981. Applied Regression Analysis. Wiley, New York, 709 pp.

El-Shaarawi, A.H. and Kwiatkowski, R.E., 1977. A model to describe the inherent spatial and temporal variability of parameters in Lake Ontario, 1974. J. Great Lakes Res., 3: 177-183.

El-Shaarawi, A.H. and Shah, K.R., 1978. Statistical procedures for classification of a lake. Scientific Series No. 86. Inland Waters Directorate, National Water Research Institute, Canada Centre for Inland Waters, Burlington, Ontario.

El-Shaarawi, A.H., Esterby, S.R. and Dutka, B.J., 1981. Bacterial density in water determined by Poisson or negative binomial distributions. Appl. Environ. Microbiol. 41: 107-116.

El-Shaarawi, A.H. and Esterby, S.R., 1981. Analyse de régression de la variation spatial d'une caractéristique limnologique. Eau du Québec, 14: 222-228.

El-Shaarawi, A.H., 1984. Dissolved oxygen concentrations in Lake Erie (U.S.A.-Canada), 2. A statistical model for dissolved oxygen in the Central Basin of Lake Erie. J. Hydrol., 72: 231-243.

El-Shaarawi, A.H., Esterby, S.R., Warry, N.D. and Kuntz, K.W., 1985. Evidence of contaminant loading to Lake Ontario from the Niagara River. Can. J. Fish. Aquat. Sci., 42: 1278-1289.

El-Shaarawi, A.H., Esterby, S.R., Clair, T. and Lemieux, R., 1985. Spatial variability of acidifiation-related water quality parameters for lakes in Eastern Canada. NWRI Contribution #85- . Environment Canada, Burlington, Ont.

Esterby, S.R. and El-Shaarawi, A.H., 1984. Coliform concentrations in Lake Erie - 1966 to 1970. Hydrobiologia, 111, 133-146.

Eynon, B.P. and Switzer, P., 1983. The variability of rainfall acidity. Canad. J. Statist., 11: 11-24.

Gordon, A.D., 1981. Classification. Chapman and Hall, London, 193 pp.

Hartigan, J.A., 1975. Clustering Algorithms. Wiley, New York, 351 pp.

Haugh, L.D. 1984. An introductory overview of some recent approaches to modelling spatial time series. In: O.D. Anderson (Editor), Time Series Analysis: Theory and Practice 5. North Holland, Amsterdam, pp. 287-301.

Howe, S. and Webb, T., III, 1984. Calibrating pollen data in climatic terms: improving the methods. Quaternary Science Reviews, 2: 17-51.

Jöreskog, K.G., Klovan, J.E. and Reyment, R.A., 1976. Geological Factor Analysis. Elsevier, Amsterdam, 178 pp.

Jumars, P.A., Thistle, D. and Jones, M.L., 1977. Detecting two-dimensional spatial structure in biological data. Oecologia, 28: 109-123.

Kwiatkowski, R.E., 1984. Lake Erie review. In: A.H. El-Shaarawi (Editor), Statistical Assessment of the Great Lakes Surveillance Program, 1966-1981, Lake Erie. Sci. Ser. No. 136, Inland Waters Directorate, Environment Canada, Burlington, Ontario, pp. 3-26.

Lebart, L., Morineau, A. and Warwick, K.M., 1984. Multivariate Descriptive Statistical Analysis. Wiley, New York, 231 pp.

Lekan, J.F. and Wilson, R.E., 1978. Spatial variability of phytoplankton biomass in the surface waters of Long Island. Estuarine and Coastal Marine Science, 6: 239-251.

Mackas, D.L., 1984. Spatial autocorrelation of plankton community composition in a continental shelf ecosystem. Limnol. Oceanogr., 29: 451-471.

Mardia, K.V. and Marshall, R.J., 1984. Maximum likelihood estimation of models for residual covariance in spatial regression. Biometrika, 71: 135-146.

Pielou, E.C., 1984. The Interpretation of Ecological Data. Wiley, New York, 263 pp.

Platt, T., Dickie, L.M. and Trites, R.W., 1970. Spatial heterogeneity of phytoplankton in a near-shore environment. J. Fish Res. Board Can., 27: 1453-1473.

Rao, C.R., 1965. Linear Statistical Inference and its Applications. Wiley, New York, 522 pp.

Rosa, F. and Burns, N.M., 1981. Oxygen depletion rates in the hypolimnion of central and eastern Lake Erie. A new approach indicates change. Presented at Workshop on Central Basin Oxygen Depletion, Dec. 2-3, 1981, Canada Centre for Inland Waters, Burlington, Ontario.

Sokal, R.R., 1979. Ecological parameters inferred from spatial correlograms. In: G.P. Patil and M. Rosenzweig (Editors), Contemporary Quantitative Ecology and Related Ecometrics. International Co-operative Publishing House, Fairland, Maryland, pp. 167-196.

van Belle, G. and Hughes, J.P., 1984. Nonparametric tests for trend in water quality. Water Resour. Res., 20: 127-136.

Watson, G.S., 1971. Trend Surface Analysis. Jour. Math. Geology, 3: 215-226.