Proceedings of the Iowa Academy of Science

Volume 77 | Number   Article 35

1970

Preliminary Results of a Statistical Analysis of a Long-Term Limnological Study of The Coralville Reservoir

William D. Wade
University of Iowa

Donald B. McDonald
University of Iowa

Preliminary Results of a Statistical Analysis of a Long-Term Limnological Study of The Coralville Reservoir

by

WILLIAM D. WADE[1] and DONALD B. MCDONALD[2]

Abstract. The Coralville Reservoir, located on the Iowa River upstream from Iowa City, has been the subject of an intensive limnological study for the past five years. Interpretation of the physical, chemical and biological data obtained was difficult due to the large numbers of parameters measured and the frequency of sampling. The application of various statistical methods was undertaken in order to aid in analysis of the data. Preliminary correlation analyses yielded some meaningful correlations but had severe limitations. Auto-covariance and auto-correlation analysis for each parameter as well as cross-covariance and correlation tests between a number of parameters yielded useful data on the presence or absence of periodic components. Auto-covariance curves obtained from the study appear to be useful in determining optimal sampling frequency for each parameter.

The Coralville Reservoir is located on the Iowa River about five miles upstream from Iowa City, Iowa. The reservoir was constructed for flood control purposes by the U. S. Army Corps of Engineers and placed in operation in 1958. Since October 1964 the Department of Civil Engineering, University of Iowa, has been conducting a study of the limnology of the Reservoir and the Iowa River above and below the impoundment. Financial support for this project was provided by the U. S. Army Corps of Engineers, Rock Island District. Several State of Iowa agencies also cooperated by furnishing services and equipment. The primary purpose of the study was to determine the effects of a flood control reservoir on the chemical and biological characteristics of its parent river. For this purpose, samples were collected on a weekly basis at a point on the Iowa River upstream from the reservoir, at surface, mid-depth and bottom at two points in the reservoir and on the Iowa River below the Coralville Dam.

Determinations of pH, carbon dioxide, alkalinity, dissolved oxygen and temperature were made in the field at the time of collection. Turbidity, hardness, iron, phosphate, ammonia-nitrogen, nitrate-nitrogen, lignins and tannins, threshold odor, biochemical oxygen demand, coliform density and total plankton were determined in the laboratory.

ANALYSIS OF DATA

Due to the large number of parameters measured, the number

[1] Associate Professor of Electrical Engineering, University of Iowa.
[2] Associate Professor of Civil (Environmental) Engineering, University of Iowa.

of sampling sites and the frequency of sampling, it was felt that the application of statistical methods would be of value in the analysis and interpretation of the data obtained during the course of the study. Preliminary studies consisted of correlation analysis of data from the downstream river station. Correlation programs were run for the total data, for the yearly data and for the monthly data with the same months grouped together. This statistical correlation was carried out in order to determine which parameters correlated with each other at the downstream river stations. Using the statistical correlation coefficients obtained from the computer output and Table VII of Statistical Tables for Biological, Agricultural and Medical Research (Fisher, 1963), the level of sgnificance of correlation was determined if it were greater than 90%. The correlations were made for direct and inverse relationship possibilities.

Although some meaningful correlations were obtained, it became apparent that these methods had severe limitations. In order to further evaluate the relationships between various parameters and between conditions in the Iowa River above and below the reservoir, analysis for auto-covariances and auto-correlation were carried out for each parameter, in addition to cross-covariance and correlation tests between a number of parameters from both river sites.

The problems of data analysis encountered in this study are typical of those usually encountered in a large and complex system. There are enough different interrelated random parameters that simple comparison often will not reveal implicatons about cause and effect relationships. Statistical analysis can be a very useful technique for detecting relationships between parameters when random and extraneous perturbations are present. A number of questions arise in applying statistical procedures as, for example, is the process ergodic, and what is the minimum samplng rate required, and often these questions cannot be answered *a priori*.

In this research the unbiased estimations of the true averaged auto-covariances and cross-covariance of the various parameters was made using the University 360 computer. Each auto-covariance characteristic is examined to determine the approximate parameter bandwidth and whether any seasonal trends are present. With this information taken into consideration, cross-covariance curves are examined to detect possible causal connections.

At this point in the research, the analysis is more qualitative in nature than quantitative, since any simple mathematical model which might be assumed at this time could not take into account such things as the time invariance and nonlinearity of the reactions which take place in the reservoir and the non-ergodicity of the driving processes.

Results of Analysis

It appeared from the preliminary correlation studies that a relationship exists between threshold odor values and reservoir discharge. For the period October 1964 to January 1969 an inverse correlation at the 98% level of significance was observed. This indicated that the highest threshold odors occurred during periods of low flow. However, examination of the data clearly shows that several periods of elevated odors have accompanied increased runoff and high discharge rates (McDonald, *et al.,* 1969). Significant positive correlations were observed between biochemical oxygen demand and ammonia, turbidity, and tannins and lignins, as well as discharge. It is clear that there was a close association between these parameters and runoff. There were variations in the patterns of correlations in different years. In 1965 odor levels correlated with both biochemical oxygen demand and ammonia concentrations at the 99% level of significance, while in 1966 odor levels did not correlate with either parameter. Comparisons of parameters from the same month in different years yielded questionable results. The probable reason for this is that in grouping data by the same months for each year, it was assumed that corresponding months in different years had the same characteristics with respect to the parameters investigated. The results obtained indicated that this was not the case.

The results using auto-covariance analysis show that many of the parameters have strong seasonal components. Some of these, such as dissolved oxygen, were essentially annual while others, such as pH, also have strong semi-annual components. In one case (plankton) there is even a significant quarter-year component present.

On the other hand, there was a significant group of parmeters which had little or no seasonal component. These included orthophosphate, odor, and biochemical oxygen demand.

In some instances periodic components were present in the parameter below the reservoir where they were nearly or completely missing above the reservoir. Examples of these are carbon dioxide and iron.

The auto-covariance data has also been examined to determine parameter waveform bandwidths and recommended sampling frequencies (see Tabe I). For the computation of the latter values the various processes have been assumed to be a linear combination of first order auto-regressive and periodic signals.

Discussion

The use of statistical methods in the evaluation of ecological data is of value provided the investigator is aware of their limitations. Hedgpeth (1957) points out that many of the data of ecology are non-parametric and hence not amenable to parametric tests. It must be kept in mind that the ecological relationships in any body of water may be extremely complex and that many of the conditions that are affecting the behavior of the system are not or cannot be measured. The differing results obtained in different years is evidence of the profound effects of climatic conditions on the aquatic ecosystem. For this reason it is desirable to collect data over a period of several seasons if statistical analyses are to be of maximum value.

Table 1. Summary of auto-covariance data for selected parameters from the Iowa River above (A) and below (B) the Coralville Reservoir.

Parameter	Location	Periodic Components			t_o (Time Constant of Random Component)
		One Year Period	Half Year Period	Quarter Year Period	
Ortho-	A	Absent	Absent	Absent	28 days
phosphate	B	Absent	Absent	Absent	35 days
Ammonia as	A	Medium	Absent	Absent	10 days
Nitrate	B	Medium	Absent	Absent	9 days
Iron	A	Absent	Absent	Absent	14 days
	B	Medium	Absent	Absent	8 days
Carbon	A	Absent	Absent	Absent	30 days
Dioxide	B	Medium	Medium	Absent	16 days
pH	A	Medium	Absent	Absent	14 days
	B	Strong	Strong	Absent	9 days
Turbidity	A	Medium	Absent	Absent	6 days
	B	Medium	Absent	Absent	5 days
Threshold	A	Weak	Absent	Absent	8 days
Odor Number	B	Weak	Absent	Absent	21 days
Dissolved	A	Medium	Absent	Absent	11 days
Oxygen	B	Strong	Absent	Absent	11 days
Biochemical	A	Absent	Absent	Absent	10 days
Oxygen Demand	B	Absent	Absent	Absent	14 days
Temperature	A	Strong	Absent	Absent	3 days
	B	Strong	Absent	Absent	3 days
Coliform	A	Medium	Absent	Absent	6 days
Bacteria	B	Medium	Absent	Absent	6 days
Plankton	A	Medium	Absent	Absent	6 days
	B	Medium	Weak	Weak	13 days

Statistical evaluation of large quantities of data can be of value in determining causal connections. A high peak in the normalized cross-covariance curve will either indicate that one parameter strongly influences the other or that both of them are strongly influenced by a third. In either case it indicates an interrelationship which, whether direct or indirect, can often give further insight into the mechanism of the reaction. Factors which can be obtained from cross-covariance curves include reaction times which, in many cases, are also to some extent implied by the band width in the auto-covariance curves. Also the presence or absence of periodic components in the corresponding auto-and cross-covariance curves is of value in making interpretations. Thus, it is generally necessary to take into account information obtained from auto-covariance curves in the interpretation of the cross-covariance curves.

After a preliminary study an auto-covariance curve can be obtained which can be used to determine the optimal sampling frequency. The sampling frequency can then be adjusted to ensure that sufficient data is taken while keeping future sampling and analysis cost at a minimum. Sampling rates computed from the time constant of the random component for various parameters (see Table I) may be of use to other investigators in this respect.

Literature Cited

Fisher, R. A. & Yates, F., 1963. *Statistical Tables for Biological, Agricultural and Medical Research.* 6th Ed., New York, N.Y., Hafner Publishing Co.

Hedgpeth, J. W., et al., 1957. *Treatise on Marine Ecology and Paleoecology.* Geologic Society of America, 67:48.

McDonald, D. B., Fisher, N. B. & Long, K. R., 1969. J. Iowa Med. Soc.: 414-416.