115641 17949

BY THE COMPTROLLER GENERAL

Report To The Congress

OF THE UNITED STATES

Better Monitoring Techniques Are Needed To Assess The Quality Of Rivers And Streams Volume I

The Environmental Protection Agency and the Geological Survey use sampling networks to assess the quality of the Nation's rivers and streams. The Council on Environmental Quality strongly endorses this approach and uses the water quality data from the networks in its annual environmental reports.

Water quality is far too complex to be monitored by these networks. Small samples, generally taken monthly, cannot account for water quality at individual sites and cannot accurately represent other locations on the same river. Nationwide reports based on data from the networks are not reliable.

Special studies of individual rivers or river segments would yield more scientifically sound and useful information on water quality. The Environmental Protection Agency and the Geological Survey should stop using their networks and shift their resources and attention to a program of well-managed special studies.

CED-81-30
APRIL 30, 1981

Request for copies of GAO reports should be sent to:

U.S. General Accounting Office
Document Handling and Information
Services Facility
P.O. Box 6015
Gaithersburg, Md. 20760

Telephone (202) 275-6241

The first five copies of individual reports are free of charge. Additional copies of bound audit reports are $3.25 each. Additional copies of unbound report (i.e., letter reports) and most other publications are $1.00 each. There will be a 25% discount on all orders for 100 or more copies mailed to a single address. Sales orders must be prepaid on a cash, check, or money order basis. Check should be made out to the "Superintendent of Documents".

COMPTROLLER GENERAL OF THE UNITED STATES
WASHINGTON D.C. 20548

B-201374

To the President of the Senate and the
Speaker of the House of Representatives

Accurate, reliable data on the actual condition of the Nation's rivers and streams are necessary for sound environmental planning and management. Existing national water quality monitoring networks operated by the Environmental Protection Agency and the U.S. Geological Survey, Department of the Interior, do not provide the type or quality of data needed. This report discusses the problems associated with the national networks and describes alternatives to them.

We are sending copies of this report to the Director, Office of Management and Budget; the Chairman, Council on Environmental Quality; the Administrator, Environmental Protection Agency; and the Secretary of the Interior.

Milton J. Socolar

Acting Comptroller General
of the United States

COMPTROLLER GENERAL'S
REPORT TO THE CONGRESS

BETTER MONITORING TECHNIQUES ARE NEEDED TO ASSESS THE QUALITY OF RIVERS AND STREAMS

D I G E S T

Sound assessments of water quality are essential to the Nation's water pollution control and cleanup programs. The Environmental Protection Agency (EPA) and the U.S. Geological Survey have established three fixed-station, fixed-interval sampling networks 1/ to assess the quality of the Nation's rivers. EPA and the Survey as well as other agencies, such as the Council on Environmental Quality (CEQ), use data from these networks to analyze whether rivers are becoming cleaner, staying the same, or getting more polluted.

Water quality assessments must be based on reliable, meaningful data if they are to be useful. However, the networks cannot provide sufficiently sound data for these assessments. Samples are taken too infrequently and stations are placed too sparsely to deal with the complex nature of water quality. Inconsistencies and errors in field and laboratory performance make network water quality data even less reliable.

WATER QUALITY IS NOT EASILY UNDERSTOOD

Water quality is complex. All biological, chemical, and physical characteristics of surface water fluctuate with time. Water quality can change dramatically during 1 day, and within a short distance. Changes can result from natural events (such as storms) or human influences (such as wastewater discharges). Complex interactions among biological, chemical, and physical processes complicate the difficulties of water quality sampling. (See pp. 1 and 2.)

SAMPLE TIMING IS CRITICAL

Sample timing is important if water quality assessments are to be meaningful and comparable. EPA and the Survey have concentrated on

---

1/ Routine sampling done periodically at specific sites.

uniform sampling frequency for their networks, limiting the frequency generally to one sample per month. Infrequent measurements cannot accurately describe water quality because rivers can change rapidly and extensively. (See pp. 18-21)

Infrequent sampling also misses important events. For example, dissolved oxygen in rivers can fluctuate dramatically within 1 day from below the minimum level of 5 milligrams per liter recommended by EPA to well above that level. (Dissolved oxygen is important because fish need oxygen to breathe.) One sample in a given day or month can describe the dissolved oxygen level only at the moment the sample was taken. The single sample cannot represent dissolved oxygen conditions for the rest of the day or month. (See pp. 19 and 20.)

Cost was a major consideration in the agencies' decisions on sampling frequencies. The agencies limited both sampling frequencies and the range of water quality analyses in order to support the number of stations they sought for the networks. (See p. 18.)

SAMPLING LOCATIONS GREATLY INFLUENCE MONITORING RESULTS

Decisions on where to take samples greatly affect individual water quality measurements and interpretations of the results. Each sampling site is affected by unique conditions, and water quality measurements taken at one site cannot normally be applied to other locations, not even those nearby in the same river. (See p. 21.)

Blanketing the Nation with enough stations to ensure that each unique stretch of waterway is monitored is not practical. However, the geographic coverage provided by the Survey's and EPA's networks is too sparse to represent water quality conditions either within individual river basins or nationwide. (See pp. 24 and 28.)

The Survey has focused primarily on riverflow, not water quality, in selecting about 500 locations for its network. Stations are located toward the downstream end of drainage basins. The Survey did not intend to detect any particular category of water quality, water use, or related land use. Instead, it generally chose to use existing flow gaging stations to minimize startup costs and to take advantage of historical flow records.

In designing its network, the Survey used an "accounting unit" concept to measure the quality of water as it leaves large drainage basins. This design concept may be appropriate for continuously measuring riverflows, but not for monthly samples of water quality. The concept fails to recognize well-known principles of stream self-purification and the highly variable nature of water quality. Water quality changes throughout drainage basins. A single station at the outlet cannot measure the conditions and changes elsewhere in the basin. (See p. 27.)

EPA encouraged States to select diverse locations for the approximately 1,000 stations in its primary network. But EPA's guidance is so broad that all the major categories of polluted waters, as well as relatively clean waters, were eligible. For example, stations could be located above and below urban or agricultural areas, at mouths of interstate streams, or at major outlets or inlets of lakes or estuaries. Apparently, EPA allowed this wide range of possibilities to gain the cooperation of States, because the States, not EPA, were to do the monitoring. (See p. 28.)

MEANINGFUL ANALYSES OF NETWORK DATA ARE UNLIKELY

The Survey, EPA, and CEQ have used network data in statistical comparisons of water quality, but these comparisons are highly questionable. The basic variability of water quality and complex causes of the variability, the limited frequency and locations of samples, and the inconsistency in field work and laboratory performance make it virtually impossible to meaningfully compare network data from month to month, season to season, and year to year. (See pp. 29 to 45.)

The Survey and EPA have known for years that network monitoring has many inherent weaknesses. But they have not studied the combined effect of all the individual difficulties on the credibility of resulting data. (See pp. 14 to 18.)

FIELD AND LABORATORY INCONSISTENCIES DEGRADE SAMPLING RESULTS

The credibility of network data is greatly diminished by inconsistencies and errors in field and laboratory performance. The Survey has noted weaknesses in field procedures of Survey and State technicians in various parts of the country. The Survey and EPA have also noted widely varying laboratory performance over time and throughout the country.

A particularly notable problem involves delays in getting samples analyzed. Because many constituents of river water are unstable, stale samples cannot produce correct or valid measurements.

These field and laboratory weaknesses add uncertainty to individual test results and to subsequent interpretations of the data. (See pp. 39 to 45.)

ALTERNATIVES?--SPECIAL WATER QUALITY STUDIES AND GREATER USE OF OTHER SIGNIFICANT INFORMATION

Special studies of water quality designed for individual river areas can overcome many weaknesses inherent in network sampling. Greater use of special studies could add immensely to our Nation's scientific knowledge of water quality changes and the reasons for the changes. (See pp. 53 and 56.)

EPA and the Survey endorse and use special studies, but they also believe that the networks should be continued for national perspective on water quality and other uses.

GAO found that national water quality assessments based on monthly network sampling were highly questionable. Better assessments of progress toward cleaner water could be achieved through well-managed special studies.

Such a program need not increase the cost of water quality monitoring. In addition to the approximately $11 million a year used for the networks, EPA and the Survey support water-quality studies through a variety of activities. With proper planning, EPA and the Survey could use the various funding sources now available to undertake a program of special studies. These studies could produce a substantial amount of good data for assessing water quality and for explaining the reasons for changes in water quality. (See p. 61.)

Other types of information related to water quality, such as changes in fish populations or reductions in discharges of contaminants, can provide insight into the Nation's efforts to reduce pollution. For example, biological monitoring is an effective method of detecting the presence of many toxic substances. EPA, States, and other agencies, such as CEQ, have already used these additional indicators in water quality reports. These sources of information could be used even more. (See p. 62.)

RECOMMENDATIONS

GAO recommends that the Secretary of the Interior and the Administrator of EPA discontinue the three national water quality networks and shift to a program of special studies of water quality.

GAO also recommends that the Administrator of EPA and the Chairman of CEQ promote the use of other available indicators of progress toward cleaner water, such as biological monitoring and reductions in waste discharges. (See p. 63.)

AGENCY COMMENTS

EPA, the Survey, CEQ, and Interior's Office of Water Research and Technology provided critical, technical, and sometimes lengthy comments on this report. Although the agencies agreed with some of GAO's concerns, they disagreed strongly with the recommendation to stop using the national water quality networks.

The agencies believe that statistical analysis will overcome the complexities of water quality and that valid assessments of the Nation's rivers can be made from infrequent (generally monthly) samples of water taken at less than 2,000 locations around the country. The agencies further believe that changes in water quality can be meaningfully analyzed without understanding the reasons for the changes.

GAO is not persuaded by the agencies' arguments and continues to believe that they should stop using the networks and devote their resources to well-managed special studies of water quality.

Chapter 5 summarizes the agencies' comments and GAO's responses. The agencies' comments are presented in their entirety together with GAO's full response to the detailed comments in a second volume to this report.

Contents

VOLUME 1

Page

CHAPTER

Page

ABBREVIATIONS

CEQ	Council on Environmental Quality
cfs	cubic feet per second
DO	dissolved oxygen
EPA	Environmental Protection Agency
GAO	General Accounting Office
mg/L	milligrams per liter
NASQAN	National Stream Quality Accounting Network
NWQSS	National Water Quality Surveillance System
STORET	Storage and Retrieval System (operated by EPA)
TDS	total dissolved solids
WATSTORE	Water Data Storage and Retrieval System (operated by Geological Survey)
ug/L	micrograms per liter

CHAPTER 1

INTRODUCTION

Human activities and natural events can dramatically change the quality of the Nation's rivers and streams. Water quality can vary substantially from hour to hour, season to season, and from place to place in the same stream. Types and concentrations of pollutants also fluctuate. The seriousness of the pollutants is tied to other chemical, physical, and biological characteristics and to the intended use of the water.

It is not easy to identify and measure the characteristics that affect the quality of the Nation's rivers and streams. In this report we evaluate three water quality monitoring networks that were intended to accomplish these tasks on a national scale.

MEASURING WATER QUALITY IS DIFFICULT

In assessing the condition of the Nation's rivers, it is essential to distinguish between the quantity of the water and quality of the water. Quantity refers to the volume of water. Quality concerns all foreign substances added to the water. This distinction is important to the issues discussed in this report.

A rough analogy to water quantity and quality is the relationship between a bus and passengers. Water quantity is a carrier, like a bus; water quality is that which is carried, like passengers. However, this analogy is far too simple to illustrate the complex makeup and behavior of water quality. Passengers board and leave a bus at designated spots; the passengers are very similar (generally all being human); they do not undergo physical change while on the bus; and they often have little interaction with other passengers.

Water quality is different on every count. Foreign substances can enter the water any time and any place; they differ vastly (e.g., mud, oil, sewage, pesticides); most foreign substances are constantly undergoing change; and they have exceptionally complex and often startling interactions. Many foreign substances, like sewage bacteria, can disappear without leaving the river. This process is called self-purification. Unstable organic matter, acted upon by the biological activity of the river, decays, using oxygen dissolved in the water. The dissolved oxygen in turn is replenished from the atmosphere and from green plants in the water. Trace metals attach to fine soils, wash into rivers by surges of water, settle to the riverbed in slow velocity reaches and pools, and become resuspended in the water when the riverflow increases again. Biological growths (e.g., plants and bacteria) attached to the river channel act as extractors and accumulators of plant nutrients in the water (e.g., phosphates, ammonia, nitrates, and certain metals). Each foreign substance added to the water undergoes its own particular change as it is carried downstream.

The process of change in water quality varies from river to river and from place to place in each river. Differences in climate (air temperature, rainfall, ice cover), riverbed structure (rocks, mud, waterfalls) and riverflow (swift or slow, rising or falling) influence the pace and form of chemical changes, decomposition, and purification. The challenge of water quality science is to determine how foreign substances enter rivers, what happens to the substances, how they affect aquatic life and man's uses of the water, and how to identify and measure the substances.

The science of hydrology (measuring quantity of water) is well developed. Largely due to the work of the Geological Survey, Department of the Interior, over the past 100 years, good topographic maps and a national network of thousands of flow-gaging stations have been established. The riverflow at many of these stations is recorded continuously by automatic monitors.

In contrast, river water quality is a relatively new science and is less developed and understood. The quality of water that passes a flow-gaging station may be altered radically along the river course. It may improve by stream self-purification and the addition of better water from tributaries, or it may drastically deteriorate from additional foreign substances, such as discharges from sewage plants, irrigation return flow, or uncontrolled runoff from city streets. Unlike continuous onsite riverflow measurements, most water quality measurements are determined by laboratory analysis of samples taken infrequently from the river.

IMPORTANCE OF WATER QUALITY INFORMATION

The Federal Water Pollution Control Act Amendments of 1972 (86 Stat. 904) obligated the United States to undertake a massive effort to restore and maintain the chemical, physical, and biological integrity of the Nation's waters. Through fiscal year 1980, $28.5 billion has been appropriated to carry out this water pollution control effort--$26.3 billion of which was for grants to municipalities to construct wastewater treatment facilities.

To meet the ambitious goals set forth in the act, it is essential to have data describing existing water quality conditions, how pollution occurs, and the effect of eliminating sources of pollution. The need for sound water quality data pervades the entire national pollution control effort from basic planning by local authorities to top-level Federal assessments of program priorities and funding.

Federal and State agencies have attempted to obtain the needed water quality data through a variety of means, including fixed-station networks and special studies. Fixed-station monitoring has been the dominant method used by water quality agencies for many years. For example, the U.S. Public Health Service began a national water pollution surveillance system in 1957 that involved fixed stations operated by all sorts of Federal, State, and local agencies.

Fixed-station networks involve routine sampling done periodically at fixed locations. Designated sampling sites are often structures, such as bridges, or convenient spots along shorelines. Sampling frequency varies considerably but monthly or quarterly sampling is common. Initially, samples from fixed-station networks were tested only for basic characteristics such as temperature and specific conductance. But gradually the range of tests expanded and now includes many physical, chemical, and biological characteristics--such as bacteria, metals, pesticides, and radioactive materials.

In contrast to the routine approach used in fixed-station monitoring, special studies are tailored to specific hydrologic and water quality conditions. Because special studies concentrate on particular problems, they vary widely in sampling frequency, number of locations, and water quality tests. However, they generally involve more intensive sampling of the affected river segments than is done through fixed-station networks.

Over the years, many criticisms have been directed at the adequacy of water quality data and monitoring programs. For example, in our report to a congressional subcommittee entitled "Water Quality Management Planning Is Not Comprehensive And May Not Be Effective For Many Years" (CED-78-167, Dec. 11, 1978), we noted that the lack of adequate data on causes and effects in water quality had plagued water quality planning for many years and apparently would continue to be a problem. In addition, the National Commission on Water Quality and the National Research Council of the National Academy of Sciences reported many inadequacies in Federal and State monitoring efforts.

PRIMARY FEDERAL AGENCIES INVOLVED IN WATER QUALITY MONITORING

The Environmental Protection Agency (EPA) and Interior's Geological Survey are the two Federal agencies most involved in water quality monitoring. EPA, because of its direct responsibilities for Federal pollution control programs, must ensure that water quality is adequately monitored. EPA relies heavily on individual States for data. The Survey monitors water quality as part of its long-standing responsibility for appraising the Nation's resources. Many other Federal agencies monitor water quality, but not on the nationwide scale of EPA and the Survey.

The Council on Environmental Quality (CEQ) does not monitor water quality but has been a major user of data produced by

EPA and the Survey. CEQ addresses national water quality status and trends in its annual reports on environmental quality.

In recent years, EPA has encouraged States to rely more on special water quality studies for decisionmaking and planning for pollution control programs and the Survey has performed more special studies. However, both agencies and CEQ continue to rely on fixed-station networks for national assessments of water quality.

OBJECTIVES, SCOPE, AND METHODOLOGY

In view of the importance of adequate, reliable data to the overall water pollution control effort, we reviewed the adequacy of fixed-station, fixed-interval water quality monitoring networks established by EPA and the Survey. We concentrated on two networks established by EPA and one established by the Survey. In this report, the term "network" will be used as an abbreviated reference to any of the three programs. The three networks are described in chapter 2.

In evaluating the best water quality data produced by network monitoring, we limited our review to rivers and streams. We did not examine monitoring needs for groundwater, lakes, reservoirs, estuaries, or oceans. Analysis of these waters is generally much more complex than for rivers and streams.

We explored a wide range of topics including network design, field sampling procedures, laboratory procedures, and data analysis. We also reviewed the legislative background of the EPA and Survey networks and technical literature 1/ pertinent to the fixed-station, fixed-interval approach to monitoring water quality.

We requested EPA and the Survey to provide examples that best demonstrated the use of fixed-station water quality information. From the examples, we selected specific stations and locations for closer review. In making our choices, we gave greatest consideration to the national assessment implications of each example and the richness, or quantity, of water quality data. One of the examples led to the case study in appendix VII. We used other examples as part of our overall evaluation, and incorporated them in chapters 3 and 4.

During the review, we were assisted by Professor Clarence J. Velz and Mr. Jerome Horowitz. Professor Velz and Mr. Horowitz assisted us in analyzing water quality data and the networks, provided advice on the technical material used in the report, and assisted in evaluating the agencies' comments on our draft

---

1/See app. I for a selected bibliography of literature reviewed.

report. Mr. Horowitz also prepared, with our assistance, the case study of the James River at Cartersville, Virginia (appendix VII), and prepared the evaluation of the agencies' comments on the case study (appendix XI).

Professor Velz is internationally recognized as a leading authority in the science of water resources and pollution control. He had an illustrious career in teaching, first as head of the Civil Engineering Department of Manhattan College, where he introduced the program of sanitary engineering involving the rational method of stream analysis and application of statistical methods. The University of Michigan invited him to revise its Department of Public Health Statistics as Chairman and to develop an interdisciplinary program in water resources management and pollution control.

Professor Velz has also served as consultant to various government agencies and private industries and has made extensive studies of numerous river systems throughout the country. Among the various committees on which he has served is the Non-Federal Advisory Committee to the Geological Survey and as initial Chairman of its Subcommittee on River Quality Assessment. For his contributions and guidance in this field, he received the Survey's John Wesley Powell Award. Professor Velz is the author of "Applied Stream Sanitation," which is used extensively as a reference book in professional circles and universities.

Mr. Horowitz is a water quality consultant in Washington, D.C. He has done studies of water quality problems for most of the Federal water agencies and has advised regional, county, and city governments on water quality matters. He has been an expert witness in water pollution litigation and at congressional hearings and has published many studies on water quality and pollution control problems throughout the United States. Mr. Horowitz is International Secretary of the Ecological Board of Southern Africa and a charter member of the Commission on Research and Development of the Association of Metropolitan Sewerage Agencies. In 1980, his outstanding achievements were honored by the Engineering News-Record.

We obtained much of the water quality data for our analyses from EPA's and the Survey's computerized information systems. (The systems are discussed on pp. 45 and 46.) We also used other sources for water quality data, such as reports of special studies by the Survey and States. We did not verify water quality data to source documents, nor did we evaluate the accuracy of laboratory or field work associated with the specific water quality data we received from the agencies' computerized systems.

Our principal contacts were with officials at both agencies' headquarters. We also visited EPA's Municipal Environmental Monitoring and Support Laboratory, Cincinnati, Ohio; the Survey's Atlanta Central Laboratory, Doraville, Georgia, and Denver Central

Laboratory, Arvada, Colorado; EPA regional offices in Seattle, Washington (Region X), and Chicago, Illinois (Region V); Survey district offices in Virginia, Washington, and Ohio; State water quality agencies in Virginia, Ohio, Washington, and Colorado; and CEQ in Washington, D.C. We also contacted other Federal, State, and local officials in different parts of the country and discussed network monitoring with several academic experts.

CHAPTER 2

AGENCIES USE FIXED-STATION NETWORKS IN MEETING FEDERAL REQUIREMENTS FOR WATER QUALITY INFORMATION

Assessment of water quality status and trends is required by Federal legislation, but the legislation does not specify how data needed for the assessments are to be obtained. EPA, the Survey, and CEQ have chosen network sampling as the principal source of data for these assessments. EPA has also required States to perform fixed-station sampling as a condition for receiving financial assistance from EPA under the Clean Water Act (33 U.S.C. 1251 *et seq*.).

WHY EPA AND THE SURVEY CHOSE FIXED-STATION NETWORKS FOR NATIONAL WATER QUALITY MONITORING

When the Survey and EPA were establishing their networks in the 1960s and early 1970s, fixed-station monitoring was widely used by Federal and State agencies. The Survey and EPA were naturally inclined to continue using this approach when they established their new national networks.

The Survey, on its own and through a large cooperative program with States, operated thousands of flow-gaging stations. Some of these sites were used for various local water quality purposes. States were also using fixed stations in their water pollution control programs. In addition, a small national network had been directed in the late 1950s and the early 1960s by the U.S. Public Health Service.

The Survey started planning for its national network in the mid-1960s. It intended to combine fixed stations and periodic special water quality studies for the monitoring sites but dropped the special studies idea because the anticipated cost of doing both types of monitoring was too high. The Survey expected the network to provide consistent water quality data that would lead to meaningful assessments of national and regional trends in water quality.

EPA consistently stressed fixed stations for its national networks. Unlike the Survey, EPA relied on others--principally States and the Survey--to do the actual monitoring work. EPA officials said it was far easier to get States to do the monitoring EPA needed through fixed stations than through an entirely different approach. While they recognized that fixed-station networks would not explain water quality conditions and changes, EPA officials hoped that the networks would provide useful water quality trend data.

EPA AND STATES ARE OBLIGATED TO MAKE MEANINGFUL ASSESSMENTS OF WATER QUALITY

The Clean Water Act is a leading source of requirements for water quality assessments. EPA and the States are required to report periodically to the Congress on the status of water pollution control efforts.

The act established reporting requirements for EPA and the States and specified goals, objectives, and timetables to guide the Nation's efforts to achieve cleaner waters. The act also requires EPA to carry out programs concerned with various categories of pollution problems, such as municipal and industrial effluents, toxic pollutants, and nonpoint sources of pollution. The act's wide-ranging provisions obligate EPA to make meaningful progress reports for congressional and public review.

Section 305(b) of the act requires each State to prepare and submit biennial reports to the EPA Administrator which must

--describe the quality of all navigable waters, taking into account seasonal, tidal, and other variations;

--analyze the extent to which all navigable waters of the State provide for protecting and propagating a balanced population of shellfish, fish, and wildlife and allow recreational activities in and on the waters;

--analyze the extent to which the above goals are being achieved and recommend any additional actions for achieving them; and

--describe the nature and extent of nonpoint sources of pollution and recommend programs to control them, including a cost estimate for implementing such programs.

The EPA Administrator, in turn, is required to transmit the reports and his analysis to the Congress.

Section 516(a) of the act requires the EPA Administrator to report annually to the Congress on measures taken toward implementing the act's objectives, including but not limited to the status of State, interstate, and local pollution control programs established under the act. The reports are to include information on the Nation's water quality. These reports are to be based on data collected from the surveillance system established under section 104(a)(5) of the act.

EPA USES FIXED STATIONS FOR ITS NATIONAL WATER QUALITY SURVEILLANCE SYSTEM

Since 1973, EPA has funded a network known as the National Water Quality Surveillance System (NWQSS), as partial response to section 104(a)(5) of the Clean Water Act. The section directs the Administrator of EPA to:

> "* * * in cooperation with the States, and their political subdivisions, and other Federal agencies establish, equip, and maintain a water quality surveillance system for the purpose of monitoring the quality of the navigable waters and ground waters and the contiguous zone and the oceans and the Administrator shall, to the extent practicable, conduct such surveillance by utilizing the resources of the National Aeronautics and Space Administration, the National Oceanic and Atmospheric Administration, the Geological Survey, and the Coast Guard, and shall report on such quality in the report required under subsection (a) of section 516." (Underscoring supplied.)

In discussing the intent of this section, the Senate Committee on Public Works stated in its report on S. 2770, which was enacted as the Federal Water Pollution Control Act Amendments of 1972 (Report Number 92-414, Oct. 28, 1971) that this surveillance system was to provide a means "to accurately inventory and determine the actual quality status of all water of the Nation." The EPA Administrator was urged to give close attention to this program to ensure high performance.

EPA started with surface freshwater and intended to expand the surveillance system to groundwater, estuaries, territorial seas, and the oceans. Because of other program priorities, however, EPA never fully developed the system. The NWQSS network of paired fixed stations was essentially all that EPA developed, and it has decreased in size. The network dropped from a peak of 188 stations in 1975 to 98 in 1978.

Stations in the network were described as "paired" because they were to be placed upstream and downstream of major land use areas, such as municipal/industrial and agricultural/rural areas. EPA planned to use these stations to analyze baselines and trends in water quality in the different types of land areas.

Major EPA uses of data from the network included:

--The first year of data was used in EPA's 1975 report to the Congress required by section 305(b) of the Clean Water Act to describe the general magnitude of pollution problems for selected water quality characteristics and the variation of water quality in different land use patterns.

--The 1975-76 EPA report to the Congress submitted under section 516(a) of the Clean Water Act summarized the same analysis.

The outlook for this limited network is uncertain. EPA has encouraged States to incorporate the stations into the network established under the Basic Water Monitoring Program discussed below. But as of April 1980, States had selected only about half of the stations. Under an interagency agreement, the Survey operated the NWQSS network for EPA. The Survey estimated the cost of operating the 98 stations in 1978 to be about $647,000. EPA established a minimum list of water quality measurements for the stations. (See app. II.) About half of the stations were sampled biweekly and the rest were sampled monthly.

EPA USES FIXED STATIONS AS PART OF ITS BASIC WATER MONITORING PROGRAM

A second EPA-sponsored network of about 1,000 fixed stations operated through States also evolved from water quality monitoring requirements specified in the Federal Water Pollution Control Act Amendments of 1972. To qualify for financial assistance from EPA under the act, the States must monitor the quality of navigable waters. Section 106(e) of the act requires that States must be providing or carrying out the

> "* * * operation of appropriate devices, methods, systems, and procedures necessary to monitor, and to compile and analyze data on * * * the quality of navigable waters * * * including biological monitoring; and provision for annually updating such data and including it in the report required under section 305 of this Act * * *."

EPA required States to use fixed stations

The States were operating thousands of fixed stations as a primary source of surface water quality data when the Federal Water Pollution Control Act was amended in 1972. EPA encouraged and required States to continue using fixed stations. The Federal regulations implementing section 106(e) of the act specified that State monitoring programs must include

"(1) compliance monitoring * * *.

(2) intensive surveys of surface waters.

(3) fixed station monitoring at representative points in surface waters." (40 C.F.R. 35.1500 et seq. App. A.)

EPA's model State monitoring program issued in 1975 more specifically stated that States were expected to rely extensively on fixed stations for their trend assessments and routine surveillance of water quality.

EPA started its Basic Water Monitoring Program network in 1977

EPA revised its model monitoring program in 1977. A major change from the previous NWQSS program was the establishment of a network of about 1,000 stations throughout the country.

EPA intended to use the new network to measure progress toward achieving water quality goals at the national level. EPA did not intend for this network to satisfy State or local needs for water quality data. According to EPA's guidelines for the monitoring program,

> "The ambient stations will be operated by the State with the data to be aggregated nationally and will be used primarily to determine national trends in water use areas (water supply, fishing/shellfishing areas, etc.), problem areas, land use areas (municipal/industrial, agricultural/rural), and in areas where future development may impact water quality and thus baseline trends are needed. These analyses are to be used in developing control and budget strategies, initiating legislation and supporting budget and grant requests at the national level." (Underscoring supplied.)

EPA suggested that many of the States' stations not selected for inclusion in the new national network be abandoned. As of April 1980, EPA and the States had designated a total of 998 stations for the national network. Many of these stations had been operated previously as part of State networks.

EPA also suggested that, in designating stations for this network, EPA regional offices and the States should consider those stations already operated as part of NWQSS and the Survey's National Stream Quality Accounting Network (NASQAN). As of April 1980, 53 NWQSS stations and 110 NASQAN stations had been designated for the new national network.

The new network became fully operational in fiscal year 1980. Based on EPA information, the annual cost to operate the network could be about $4.3 million, or $4,300 per station. The range of water quality measurements and the sampling frequencies EPA established for the stations are shown in appendix III. Generally, monthly sampling is required.

THE GEOLOGICAL SURVEY IS OBLIGATED TO PROVIDE RELIABLE, MEANINGFUL WATER QUALITY DATA

Since 1879, the Survey has had general responsibility for appraising the Nation's resources, including its water. The Survey conducts water resources investigations and research and acquires data on water supply, water quality, and water use. In total, the Survey operates about 14,000 stations for measuring surface water flow and about 7,000 stations for measuring surface water quality. A significant part of this activity is performed in cooperation with States, local agencies, and other Federal agencies who share or totally pay the costs.

NASQAN is a network of about 500 stations, which the Survey for the most part independently funds and operates. NASQAN was developed in response to the Bureau of the Budget 1/ Circular A-67, dated August 28, 1964, which made the Department of the Interior responsible for establishing a national information network for acquiring data on the quality and quantity of surface water. Except for specialized data needs, this national system was to meet the water quantity measurement requirements of all Federal agencies and the water quality measurement needs common to two or more agencies. The Secretary of the Interior designated the Survey to design and coordinate this national system.

The Survey's concept for the national system distinguishes three levels of data needs (national and regional, subregional, and local) and three functional categories. NASQAN has been designated to help satisfy national and regional data needs on streamflow and stream quality--a major component of the national information system.

The primary objectives established for NASQAN were to

--account for the quantity and quality of water moving within and from the United States,

--depict geographic variability of water quality,

--detect changes in stream quality, and

--lay the groundwork for future assessments of changes in stream quality.

1/The Bureau of the Budget was redesignated as the Office of Management and Budget by Reorganization Plan 2 of 1970, effective July 1, 1970.

NASQAN stations are generally located toward the downstream end of hydrologic accounting units 1/ in order to measure the quality and quantity of most of the water leaving the units. Sampling is done monthly for some water quality tests and quarterly for most others. (See app. IV for a listing of measurements and sampling frequencies.)

NASQAN began in 1973 with 50 stations and grew to 518 stations by fiscal year 1979. Survey personnel generally do sampling and laboratory analysis of the water. The Survey estimated the cost of operating NASQAN in fiscal year 1979 at about $5.6 million, or about $10,800 per station.

CEQ IS OBLIGATED TO REPORT ON CONDITIONS AND TRENDS IN WATER QUALITY

The National Environmental Policy Act of 1969 (42 U.S.C. 4321 et seq.) calls for the President to transmit to the Congress annually an environmental quality report. CEQ was created by the same act. It assists and advises the President in preparing the annual environmental report. Other specific requirements imposed on CEQ by the act (sec. 4344) included

> "(2) to gather timely and authoritative information concerning the conditions and trends in the quality of the environment both current and prospective, to analyze and interpret such information for the purpose of determining whether such conditions and trends are interfering, or are likely to interfere, with the achievement of the policy set forth in title I of this Act, and to compile and submit to the President studies relating to such conditions and trends; * * *."
>
> * * * * *
>
> "(7) to report at least once each year to the President on the state and condition of the environment; * * *."

CEQ relies on other agencies for the environmental data used for its assessments and annual reports. The Geological Survey and EPA have been principal sources for water quality data. For example, since 1975 CEQ has based its reports in part on data from NASQAN.

1/The siting of NASQAN stations is based on a system developed by the Water Resources Council, which divided the river drainage basins of the United States into regions and subregions. The Survey further divided these basins into 352 accounting units.

CHAPTER 3

FIXED-STATION NETWORK

MONITORING HAS MANY WEAKNESSES

EPA and the Survey face many technical, analytical, and practical difficulties in their attempts to use network data for assessing conditions and trends in water quality. Many of the difficulties stem from the agencies' decisions to use limited sampling programs to portray complex water conditions. The agencies apparently started their national networks without a clear understanding of how they were going to interpret the data and how valid their interpretations would be.

We believe that EPA and the Survey will not be able to meaningfully assess conditions and trends in water quality through their networks because:

--Water quality changes too rapidly and too often to be accurately represented by infrequent sampling.

--Water quality varies too much throughout a river basin to be represented by one or a few stations. At best, each sample represents only one specific site.

--Statistical analyses and comparisons of network water quality data are not meaningful because samples are taken from dissimilar water quality conditions.

--Inconsistencies and errors in field and laboratory work weaken the credibility of network data.

--Computerized information systems used to record and analyze network data do not have adequate information on the credibility of the data or the local conditions affecting the data.

We believe that if the networks are continued, EPA, the Survey, and others relying on the networks, such as CEQ, will be left with a mass of uncertain data of limited usefulness and questionable validity.

NETWORK MONITORING PROBLEMS ARE WIDELY KNOWN

The broad range of problems afflicting network monitoring has been recognized for many years by professionals in this field and by the agencies. The following observations illustrate the extent of concern about limited sampling programs.

Water quality experts have cited difficulties

Experts in water quality have expressed concern for years about many of the difficulties involved in fixed-station sampling. For example, a 1969 book 1/ issued by EPA's predecessor, the Federal Water Pollution Control Administration, noted that:

--Many water quality constituents, such as bacteria, dissolved oxygen, and ammonia, change rapidly, making it difficult to obtain representative measurements. For example, aquatic organisms can cause wide variations in dissolved oxygen: very high concentrations in the afternoon and very low levels near dawn.

--Streamflow is one of the primary determinants of water quality. Both the natural water quality and the effects of wastes in a stream vary as streamflow changes.

Clarence J. Velz, 2/ in his 1970 book "Applied Stream Sanitation," discussed the intricacies of sampling in streams and rivers and made several points relevant to the networks established by EPA and the Survey:

--The objective of stream sampling is to obtain a representative, reliable measure of water quality along the watercourse (a stream profile, not a snapshot of a fixed site).

--Rivers are living and dynamic, subject to hydrological and biological variations and manmade influences. Consequently, stability during sampling is very difficult to achieve and, at best, sampling results are subject to variation.

--A few well-selected locations with enough samples to define results in statistical terms provide a much more reliable picture of quality than many stations with only a few samples at each.

--Samples taken at individual stations should be representative of the main body of flow. When a river is not thoroughly mixed, it must be sampled at several positions across the river and at different depths. This situation greatly increases the number of samples needed to obtain reliable average measurements.

1/F. W. Kittrell, "A Practical Guide to Water Quality Studies of Streams," Federal Water Pollution Control Administration, 1969.

2/Professor Velz is one of the Nation's leading authorities on river water quality studies.

EPA and the Survey are aware of the problems

Both agencies have known for years that network sampling has many inherent weaknesses. Some problems, such as sample timing and location bias, have been addressed in many documents prepared by or for the two agencies. Although the agencies were aware of all categories of problems discussed in this report, we found no evidence that they had evaluated the cumulative impact of the various weaknesses on data from the networks.

Since 1972 EPA has contracted for several studies on the design of water quality monitoring networks. Several of the reports pointed out that optimal network design must be site-specific and requires knowledge of individual water quality characteristics at each site. A 1973 report 1/ prepared for EPA concluded that

"* * * even with the commitment of a substantial effort, it is apparent that these surveillance networks do not meet the overall needs of the water quality management programs. Many waste sources go undetected. Frequently, pollution problems have passed before their existence is noted by the surveillance system.

"Thus, one must conclude that water quality sampling and data analysis consume a large portion of the valuable staff time without providing the information necessary to support water quality management programs."

Several Survey reports have pointed out problems in routine monitoring of water quality. For example, Survey Circular 715-D on the Willamette River Basin stated:

"Faced with the dilemma of 'what kind of data to collect?' the tendency has been simply to 'shotgun it'--that is, to collect a lot of samples in a lot of rivers for a lot of constituents. Apparently, part of the rationale for this approach has been that once all the data are collected, compiled in books, or stored in computer files, 'users' would decide as to their needs and then obtain the data for their particular purposes. Without a clear definition of goals and information needs, programs have typically become preoccupied with long lists of 'pollution sensitive' variables. Seemingly, every year a few new variables are added to the list. Sampling frequencies for such

---

1/Robert C. Ward, "Data Acquisition Systems in Water Quality Management," Colorado State University, EPA Report No. EPA-R5-73-014, May 1973.

programs are often selected arbitrarily or based simply on convenience. Sampling sites are commonly established by political boundaries, convenience, or proximity to a particular 'polluter.'

"* * * river-quality data continue to be disappointing in terms of their interpretive utility, and criticism of data programs has mounted in recent years. As indicated in the 'Introduction,' there is a growing number of respected scientists and engineers who feel that programs based primarily on the monitoring-type approach will never prove efficient for trend or causal analysis. If this opinion is correct, and past experience indicates strongly that it is, the implications for environmental management can only be negative unless a more effective approach is developed and implemented."

In 1978, an internal Survey report 1/ summarized concerns some headquarters personnel had about evaluating NASQAN data for trends. Some of the concerns were:

"The network design constraints of NASQAN are such that the detection of trends is possible, provided the trend is sufficiently large; but, the causes of the trends will probably mostly have to be identified through ancillary studies. It is significant to note that determinations of whether or not the Nation's waters are being cleaned up will be tenuous without ancillary studies * * *.

"* * * the objectives of NASQAN are not well defined as relating to the detection of trends.

"We presently do not know which statistical tests of differences are most powerful for detection of trends in water quality. Different statistical tests may be more powerful for different constituents * * *."

"The first and most important problem identified for further study is the question of whether or not trend detection, at an acceptable level of significance, is possible within a reasonable time frame of 5 or 10 years, assuming the existing sampling frequency constraints."

1/Memorandum to Chief, Quality of Water Branch, Water Resources Division, U.S. Geological Survey, from Marvin O. Fretwell, Western Region, Water Resources Division, titled "PROGRAMS AND PLANS--Capabilities and Information Needs to be Considered in Planning for Evaluation of NASQAN Data for Trends," Nov. 30, 1978.

Officials of both agencies said that they intended to evaluate the networks after several years of operation. We do not believe they had to wait. When the networks were being established, both agencies had a substantial amount of fixed-station data available for study and the complexities of water quality analysis were widely known.

TIMING PROBLEMS DISTORT NETWORK DATA

Sample timing is important for both comparability and meaningfulness of water quality measurements. River water quality does not remain homogenous throughout the year. Rivers are different in the summer than in the winter. Within each season water quality can change dramatically and erratically as a result of changes in riverflow, weather and many other upsets to the chemical, biological, and physical systems in rivers. Infrequent sampling cannot adequately represent rapidly changing properties of water. The networks inevitably see only brief glimpses of a hodgepodge of water quality conditions. What appears to be a change in water quality could easily be a quirk of sample timing.

The networks generally collect one sample a month during normal work hours. With few exceptions, the two agencies are not attempting to isolate any particular pattern of water quality. Consequently, the measurements represent mixtures of water quality conditions and the mixtures change with time. Also, important water quality conditions can go undetected if samples are taken at about the same time during normal work days. The importance of sampling during stable, similar conditions was emphasized by Professor Velz in his 1970 book "Applied Stream Sanitation."

> "With so many factors inducing variation in water quality of a natural stream (or estuary), it is obvious that random sampling or sampling over an extended period of time is almost certain to reflect a series of distorted values of heterogeneous conditions. It should be equally obvious that sampling must be intensive over a short period to ensure stability and thus permit relating the observed water quality profile to the set of stable conditions that produced it."

Cost was a major consideration behind the agencies' decisions on sampling frequencies. The agencies limited both sampling frequencies and the range of water quality characteristics to be monitored in order to support the number of stations that they sought for the national networks.

The changeable nature of water quality over time--hourly, monthly, and yearly--is illustrated by the following two examples. Later, on page 29, we explain why this variability and the network timing problems preclude valid, meaningful analyses of network data.

Dissolved oxygen can change dramatically in a few hours

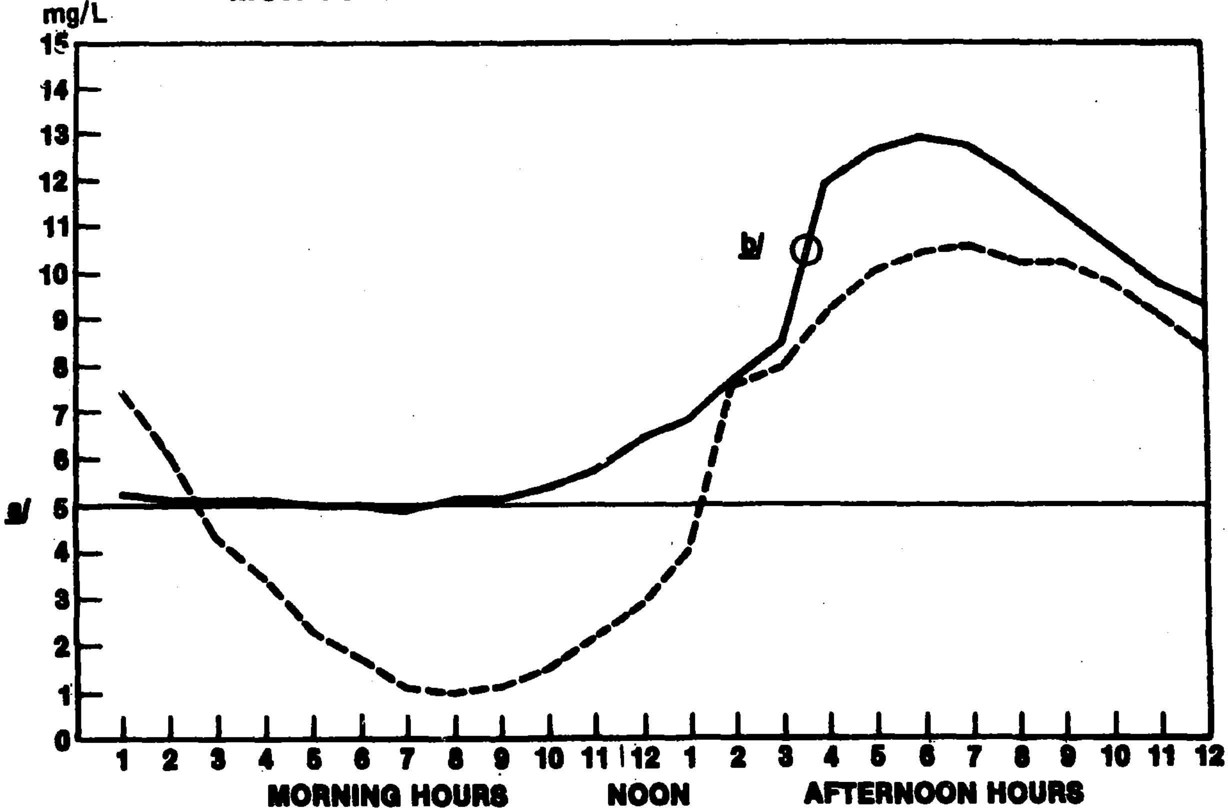
Dissolved oxygen (DO) is generally considered one of the most important indicators of pollution. DO is vital in assessing the suitability of water for fish life and is commonly used in water quality standards.

The severity of DO problems, and the extent of river length affected by them, are essential to assessments of water quality. However, DO levels in a river are not static. The levels and locations of severe depletion change in response to all sorts of hydrological, climatic, biological, and chemical influences. DO levels and their determinants are unique to each river and even to individual segments of each river. (See p. 103, app. VII, for a more detailed discussion.) Widely scattered stations and infrequent sampling cannot possibly reveal the changing nature of DO.

DO concentrations can fluctuate greatly within 1 day. One measurement taken in a day of wide fluctuation cannot record the DO activity for that day; the sample can depict the DO level only at the moment of sampling. Single samples cannot reveal whether or how much DO changes during the day. When DO changes follow regular daily patterns, sampling at a fixed time of day inevitably produces a biased picture of DO.

The graph on the following page shows how DO can fluctuate in 1 day and how misleading one measurement per day can be. DO measurements in both rivers were much lower in the morning than in the afternoon. In the networks, one measurement would represent the DO concentration for an entire sampling interval--usually 1 month. The Survey took a suite of NASQAN measurements at Milford during the afternoon of July 12, 1977. The DO reported for that visit was 10.5 milligrams per liter (mg/L). The graph reveals that on July 12, 1977, the automatic DO recorder at Milford showed that DO ranged from 4.9 mg/L to 12.9 mg/L. The recorder also showed a wide DO range on many other days that month. DO dipped below EPA's recommended minimum level of 5.0 mg/L on 15 days in July 1977. Clearly the single reading of 10.5 mg/L, which is healthy, does not accurately portray DO at Milford during July 1977. The automatic DO recorder at Miamisburg (though not a NASQAN station) also showed fluctuations and very low DO levels on many days in May 1977.

CHANGES IN DISSOLVED OXYGEN LEVELS RECORDED AT HOURLY INTERVALS BY U.S. GEOLOGICAL SURVEY AUTOMATIC MONITORING DEVICES FOR TWO RIVERS IN OHIO.



———— Little Miami River at Milford, Ohio, July 12, 1977 (Geological Survey NASQAN station no.03245500)

— — — — Great Miami River near Miamisburg, Ohio, May 20, 1977 (Geological Survey station no.03271600)

a/ 5 mg/L is EPA's recommended criterion for minimum DO.
b/ 10.5 mg/L was the DO reading listed for the July 1977 NASQAN visit to this station.

Suspended sediment concentrations are unstable

Suspended sediment in rivers comes from erosion. It is affected by rainfall and riverflow. Dirt is washed into rivers by rainfall; the river channel itself is eroded by rising riverflow. The suspended sediment may contain sand, clay, heavy metals, agricultural chemicals, toxic organics, and many other substances.

Large, intense storms and raging riverflows greatly increase erosion and concentrations of suspended sediment, but the concentrations can also fluctuate greatly with moderate changes in flow. Suspended sediment can vary between episodes of similar flow patterns. Each river basin and river segment is unique. Suspended sediment varies from location to location depending on hydraulic forces, vegetative cover; type of soil; steepness of terrain; and man's influences through agriculture, lumbering, mining, construction, dams, etc.

Suspended sediment concentrations can fluctuate even more rapidly and extensively than DO, and monthly patterns in suspended sediment are much less stable. Usually when riverflows are high, suspended sediment concentrations are high, and when riverflows are low, suspended sediment concentrations are low. However there is scatter in this relationship.

The scatter diagram on the following page shows wide ranges in suspended sediment in the Yakima River during 1974-79. For example the five measurements in January 1975-79 ranged from 4 to 490 mg/L. In May 1975, the eight samples taken that month ranged from 32 to 204 mg/L. When conditions are unstable, one sample cannot give a reliable picture.

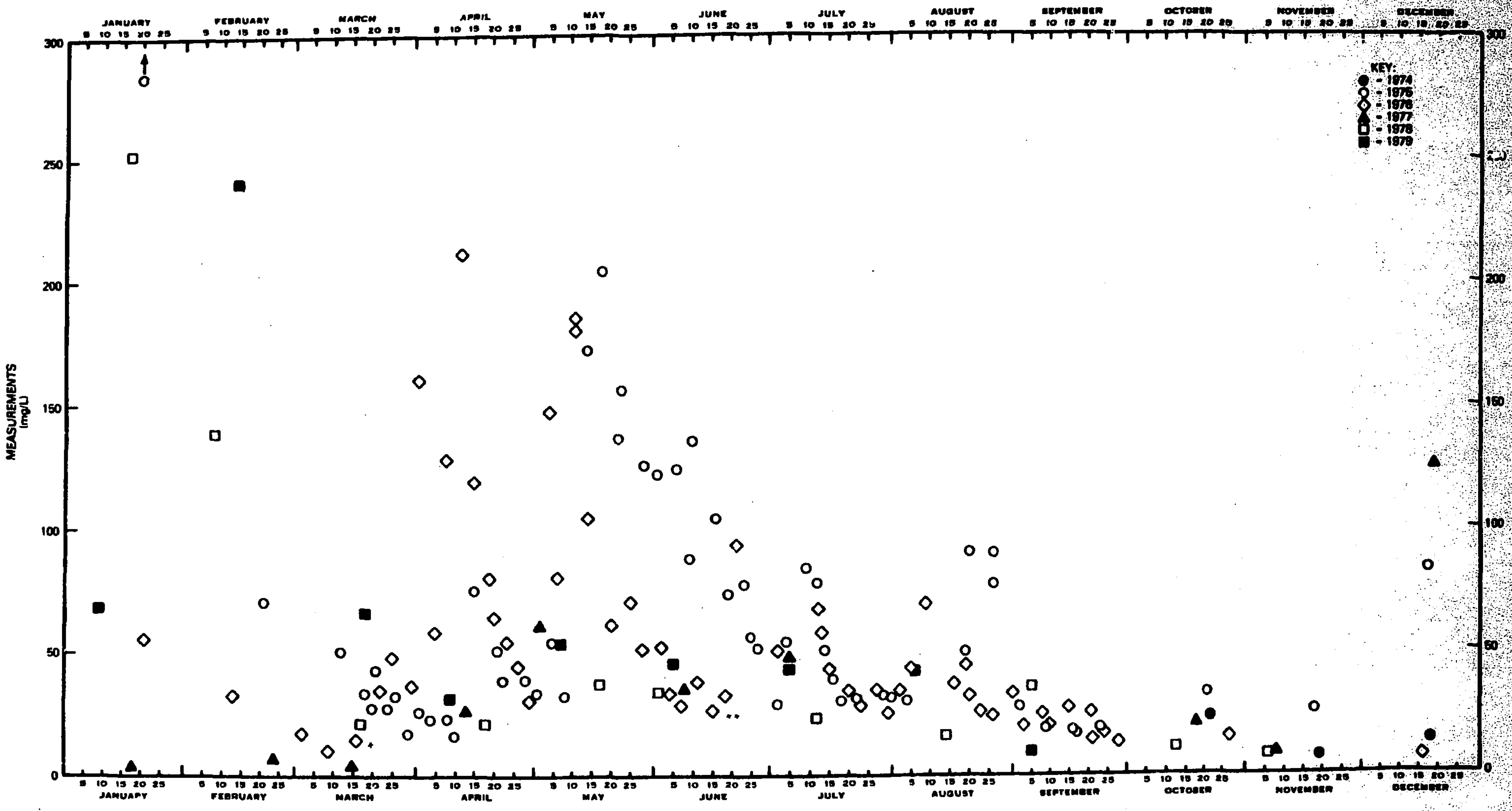
The Yakima River is more stable than many rivers in the Nation. Reservoirs help control the Yakima River, storing higher flows and releasing water later for irrigation. About one-third of the NASQAN stations in 1975 reported suspended sediment ranges wider than those at Kiona. Many stations reported measurements exceeding 1,000 mg/L. With highly variable conditions throughout the Nation, monthly network sampling cannot accurately portray concentrations or changes in suspended sediment.

NETWORK STATIONS HAVE INEVITABLE LOCATION BIAS

The placement of network stations greatly affects what they find. Water quality can vary substantially over short distances or from bank to bank at specific sites. Different chemicals, minerals, and biological matter are introduced from tributaries, irrigation return flows, sewage effluents, storm drains, surface runoff, and groundwater. A river flowing through a city can be very different from the same river flowing through farmland a few miles upstream.

The Survey and EPA have not tried to monitor any particular pattern of water quality conditions for their two main networks. The Survey tries to measure water quality at the outlets of large drainage areas. EPA seeks a great variety of water quality conditions.

SUSPENDED SEDIMENT MEASUREMENTS TAKEN AT THE NASQAN STATION ON THE YAKIMA RIVER AT KIONA, WASHINGTON, DURING 1974-1979, PLOTTED TO SHOW BOTH THE MEASUREMENT LEVELS AND DATE FOR EACH SAMPLE.



*ON MARCH 16, 1976, TWO SAMPLES MEASURING 18 mg/L WERE TAKEN.
**ON JUNE 15, 1976, TWO SAMPLES MEASURING 26 mg/L WERE TAKEN.

It is not practical to blanket the Nation with enough stations to ensure that each unique stretch of waterway is adequately monitored. In the early 1970s, EPA considered establishing a large network of fixed stations to allow tighter management of pollution control programs. EPA estimated that at least 10,000 stations would be needed and that even a network of that scale would place stations an average of about 350 miles apart. An EPA official said that the agency dropped this idea because the 1972 amendments to the Federal Water Pollution Control Act shifted the emphasis from river conditions to wastewater limitations, thereby reducing the urgency for instream monitoring.

There are no network stations on many rivers in the Nation. Few rivers have more than a small number of stations. Because water quality changes from place to place, it is impossible to generalize from one station to other places on a river or to the river as a whole. A station far away from a problem area may detect no sign of the problem.

The following examples illustrate the influence of sampling locations on water quality measurements.

<u>South Platte River quality varies greatly in Colorado</u>

The variability of water quality measurements at different points along a waterway is apparent from sampling done on the South Platte River in Colorado. The South Platte River flows east from the Rocky Mountains to its junction with the Platte River in Nebraska. The river flows through Denver shortly after leaving the mountains, then passes through about 200 miles of rangeland, farms, open areas, and small communities before entering Nebraska near Julesburg. The river and its tributaries receive municipal and industrial wastewater and urban runoff around Denver. The river is extensively used and reused for irrigation downstream of Denver.

Variability in a short distance was clearly shown in 1976-77 during a special water quality study by the State, county, and Denver.

Examples of Measurements Recorded for 2 Days During a 1976 3-Month Study of Water Quality in the South Platte River at Denver, Colorado

Sampling location	Fecal coliform bacteria (note a)		Ammonia nitrogen as N (mg/L)	
	July 21	July 23	July 21	July 23
Bowles Ave.	1,120	400	0	0
Hampden Ave.	9,000	1,700	.4	.9
Dartmouth Ave.	1,100	640	.5	.6
Evans Ave.	4,700	440	.8	1.2
Alameda Ave.	4,000	570	.8	1.0
8th Ave.	4,400	700	1.1	1.0
Colfax Ave.	15,000	1,300	1.3	1.2
Speer Blvd.	11,000	2,800	1.1	1.0
19th Ave.	100,000	1,200	.8	1.2
31st St.	32,000	700	.6	1.1
46th Ave.	3,000	1,100	.6	1.0
Franklin St.	34,000	1,300	.6	.9
York St.	28,000	1,900	.6	1.2
Highway 76	34,000	3,100	3.4	4.2
104th Ave.	150,000	4,000	4.6	4.4

a/Measured in number of colonies per 100 milliliters.

This table lists measurements in geographical sequence; Bowles Avenue is upstream of Denver, and 104th Avenue is downstream of Denver, about 24 miles from Bowles Avenue. The table shows that location greatly influences sampling results. It also reinforces our earlier discussion of timing problems (see p. 18)--many measurements taken 2 days apart were quite different.

The following table gives a broad idea of water quality variability over a longer distance in the South Platte River.

Comparison of Colorado State Monitoring Results for Fecal Coliform Bacteria at Four Locations on the South Platte River in 1976 (note a)

Station location	Median measurements	Range of measurements Smallest	Largest	Number of samples
Littleton--upstream of Denver (river mile 327)	4	less than 3	230	20
Henderson--downstream of Denver (river mile 302)	930	less than 30	b/230,000	21
Fort Morgan (river mile 189)	less than 40	9	4,300	6
Julesburg (river mile 77)	162	less than 30	430	8

a/Measurements are stated in number of colonies per 100 milliliters.

b/The 230,000 colonies/100 mL measurement is an extreme. No other measurement at Henderson exceeded 43,000 colonies/100 mL.

The river above Denver was much cleaner than it was below. Since fecal coliform bacteria die rapidly in rivers, it is unlikely that any bacteria from Denver reached Fort Morgan or Julesburg. The bacteria at these downstream stations must have come from local sources.

Water quality varies from bank to bank

Variability of water quality across a river at one site is demonstrated by the Yadkin River in North Carolina. 1/

1/Cross-sectional variability in water quality is natural and common. Water flowing from tributaries or wastewater sources will not immediately spread across the entire width of the receiving rivers. It takes time and distance for the incoming water to mix. This mixing process varies by location and changes over time even for individual locations.

During 4 hours on April 28, 1971, a continuous DO recorder operated by the Survey at Yadkin College, North Carolina, recorded a sudden plunge in DO. A visiting Survey employee repairing the recorder noted the plunge. Using a portable oxygen probe, he sampled DO from a bridge downstream of the station and found that the low measurements did not extend across the river. According to the Survey, fish were killed by the sudden plunge in oxygen. The cause of the plunge was later traced to sediment flushed from Muddy Creek by heavy rainfall. Muddy Creek is on the same side of the Yadkin River as the Survey's station. The Survey reported that, although the Yadkin station was nearly 30 miles below Muddy Creek, the sediments had not mixed with the rest of the river; instead it stayed near one bank. A Survey official told us that if the station had been located on the other side of the river, the DO plunge would not have been recorded.

Geological Survey used hydraulic accounting units in selecting NASQAN sites

The Survey established NASQAN stations near the downstream end of hydraulic accounting units (drainage areas) in an attempt to measure the quality and flow of most surface water leaving the drainage areas. (A map showing the NASQAN accounting units is in app. V.) The Survey did not intend to detect any particular category of water quality, water use, or land use. As a result, a wide variety of conditions and locations is monitored through NASQAN, such as large cities, remote rural areas, and shipping canals.

The Survey's accounting units ignore the highly variable nature of water quality and well-known principles of stream self-purification. A continuous flow gage at the outlet of a drainage area can provide a good understanding of changes in flow upstream. This location may be satisfactory for measuring riverflow, but not for measuring water quality. Foreign materials enter river water throughout its course. Each foreign substance undergoes its own particular changes along the course of the river. A water quality station at the outlet of a large drainage area cannot possibly show all that happened upstream. The station could miss much of what happened along the way, thereby giving a false or meaningless report of the wellbeing of the river. The river might actually have been very polluted in a number of reaches and yet could recover without any trace of pollution by the time the water passed the network station near the outlet. We believe this design weakness is one of NASQAN's fundamental flaws.

There are only 518 NASQAN stations in the entire Nation. Most rivers have no monitors, and few have more than one. For example, the only NASQAN station on the South Platte River between Denver and the Nebraska State line is located near the border at Julesburg, Colorado. This station cannot give a fair picture of water-quality conditions over 200 miles upstream near Denver.

EPA was vague about
where to place stations

EPA allowed States to select sites for its national network of 1,000 stations. In 1975 EPA published guidance for site selection in its document entitled "Basic Water Monitoring Program." The guidance was vague and permissive; it allowed States to include all types of polluted water and very clean water too. Stations could be placed in many ways:

"1. In a paired configuration. For example, upstream and downstream of representative land use areas (that is, municipal/industrial, agricultural/rural).

2. Single stations located in small and homogeneous subbasins. These stations may be located in specific water use areas, for example: at surface water supply intakes, within recreational areas, or within commercial fishing and shellfishing areas.

3. At locations within major rivers and significant tributaries. For example, these stations may be located:

- At the major outlets from and inputs to lakes, impoundments, estuaries or coastal areas; or

- At the mouths of major intrastate and interstate streams and significant tributaries to these streams, etc."

EPA officials considered requiring all stations to be in pairs, as they had been in NWQSS. However, the guidance was made more permissive because of resistance by State officials. States were used to fixed stations. According to EPA, the States believed that they, not EPA, should decide where the stations should be.

The same location bias problems afflicting NASQAN apply to EPA's network. Sparsely scattered stations cannot account for the wide variety of water quality in individual rivers. Pairing stations close together can be useful but can easily produce misleading information. Sampling the pairs must be well coordinated to ensure that measurements are taken from the same patch of water as it flows from the upstream station to the downstream station. Unless time of travel is accounted for, water at the downstream station may have no relation to the patch of water measured upstream, and the differences between them cannot be objectively compared.

STATISTICAL DEFICIENCIES
OF NETWORK DATA

As shown in this chapter, the credibility and usefulness of data from networks are seriously weakened by problems of sample timing, location bias, stale samples, and errors and changes in field and laboratory techniques. Even so, CEQ, the Survey, and EPA have used network data in statistical analyses of water quality. Statistical analysis does not eliminate the elementary weaknesses of network data. Instead, the elementary weaknesses become more obvious and convincing when the demanding rules of statistical analysis are applied. Basic statistical deficiencies of network data include

--an inadequate sampling plan and

--inappropriate statistical summaries.

These deficiencies are discussed below.

Inadequate sampling plan

One sample a month throughout the year at widely spaced sites cannot represent more than a tiny fraction of the complex patterns of water quality in a river. Since EPA and the Survey in general do not schedule their sampling to catch any particular pattern, the networks produce a hodgepodge of nonhomogeneous measurements. Trend analyses based on the mixed measurements are oversimplified and can be misleading when the analyses fail to account for the different water quality patterns and their fundamental determinants.

Some of the common determinants of water quality are:

--Changes in riverflow (both the size and speed of the change are important).

--Season (hot weather and cold; wet weather and dry; irrigation season and harvest season; etc.).

--Rainfall patterns (location, intensity, and duration in various parts of the river basin above the sampling site).

--Aquatic plants (including algae).

--Changes in pollution control (episodes of inadequate treatment and long-term improvements).

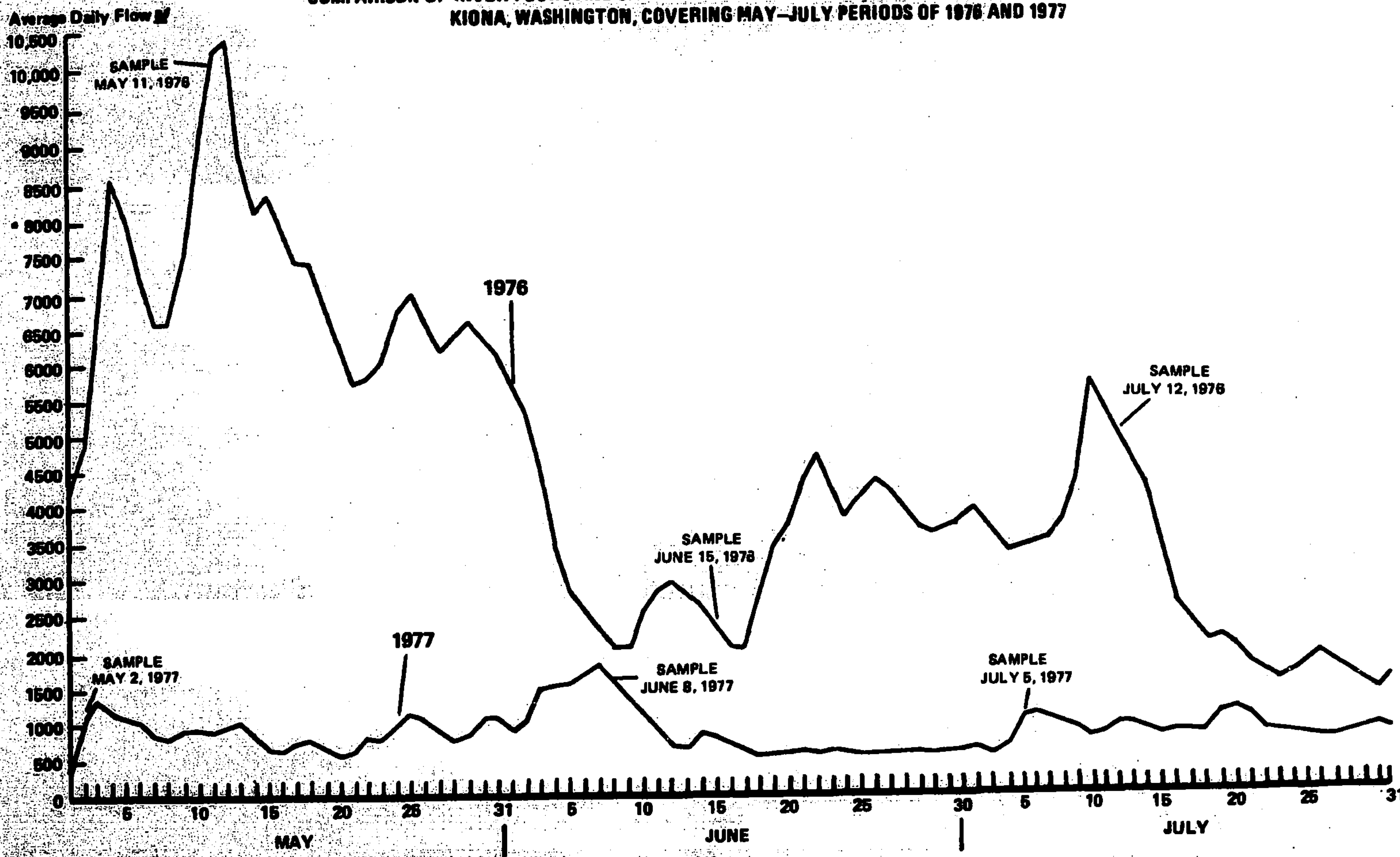
--Changes in urban growth, industrial development, and river development.

--Forest fires and other factors that affect erosion (reforestation, agricultural practices, heavy construction that disturbs land surface, new dams and reservoirs that can catch waterborne sediment before it reaches the sampling site).

There are many other determinants. Each combination of determinants produces a different pattern of water quality. It is obvious that a river in summer is different from a river in winter and that a river in flood is unlike a river in drought. Water quality will be different in each of these conditions because of the strong influence of riverflow and season. Other factors, such as changes in fertilizers, will add to the differences in water quality. The rigid sampling plans of the networks cannot flex to isolate the effects of season, riverflow, and other important determinants.

The following example illustrates the important relationship between riverflow and sample timing. The Survey takes monthly samples at a NASQAN station on the Yakima River, at Kiona, Washington. The chart on page 31 shows the riverflow records for part of 1976 and 1977, together with NASQAN sampling dates. The riverflows were much higher and less stable in 1976 than in 1977. The chart shows that the Survey took samples under all sorts of riverflow conditions: high flow and low flow, rising flow and falling flow, fairly stable conditions and extremely unstable conditions. Since water quality is greatly influenced by riverflow, it is not surprising that the samples at Kiona were quite different in 1976 and 1977. The table on page 32 shows some of the large differences for July in these 2 years.

COMPARISON OF RIVER FLOWS AND SAMPLE TIMING AT THE NASQAN STATION ON THE YAKIMA RIVER AT KIONA, WASHINGTON, COVERING MAY—JULY PERIODS OF 1976 AND 1977



a/ Stated in cubic feet per second

Selected Measurements of the Yakima River at Kiona, Washington, Recorded by the Survey for the Months of July 1976 and July 1977

Sampling date	Instantaneous riverflow (cfs) a/	Specific conductance (micromhos)	pH (standard units)	Total Kjeldahl nitrogen (mg/L)	Total hardness (mg/L)
July 12, 1976	5,000	171	7.8	0.45	63
July 5, 1977	959	365	9.2	2.9	150

a/cubic feet per second.

There are many differences between the July samples in 1976 and 1977. In 1977 the sample was saltier (higher conductance), more alkaline (higher pH), harder (higher total hardness), and it contained more unoxidized nitrogen (higher total Kjeldahl nitrogen). These differences may be largely attributed to differences in riverflow, but they may also be related to farming practices or other factors that cannot be determined from the limited data in NASQAN.

Policymakers need to know whether water quality has changed because of human actions--pollution control, soil conservation, new dams, etc. Their information needs cannot be satisfied by trend analyses that fail to sort out things that can be controlled (such as sewage and industrial waste) from those that cannot be controlled (such as rainfall). The need for this distinction is clearly illustrated in our case study of CEQ's trend analysis of the James River at Cartersville, Virginia. (See app. VII.)

Inappropriate statistical summaries

CEQ, EPA, and the Survey have published statistical summaries of network data. These summaries are largely derived from such statistics as the annual average and the standard deviation. When data have a classic bell-shaped (so-called normal) distribution, the average is an excellent measure of central tendency and the standard deviation is an excellent measure of scatter. However, when the data have an odd distribution--when there are many extreme values or when the distribution is lopsided--the average and the standard deviation are not efficient estimators of central tendency and scatter. The statistical summaries prepared by the agencies fail to specify whether the distributions are orderly, erratic, distorted, skewed, or normal. We have found that data from the networks commonly do not have a normal distribution. Consequently, the average and the standard deviation are inappropriate statistical summaries of their central tendency and their scatter.

The average and the standard deviation are strongly affected by extreme values--this is one of their inherent mathematical properties. One very high value will greatly raise the average and one very low value will greatly depress it. Although the average is the most widely known measure of central tendency, several other measures are in common use; of these, the most familiar is the median (the middle point of a series of values ranked from lowest to highest). Extreme values have no effect on the median, but they greatly affect the average.

Testing for normality

There are several ways to determine whether data are normally distributed. The following example illustrates how data can be tested for normality and how misleading yearly statistical summaries of network data can be.

The following table shows all the measurements of nitrite plus nitrate (often associated with fertilizers) recorded by the Survey at its NASQAN station on the Mississippi River at St. Paul, Minnesota, during water years 1975-77.

Measurements of Nitrite plus Nitrate (as Nitrogen) Recorded for Samples Taken at the NASQAN Station on the Mississippi River at St. Paul, Minnesota, During Water Years 1975, 1976, and 1977

	1975	1976	1977
	------(milligrams per liter)-----		
Oct.	0.3	0.1	0.1
Nov.	0.2	0.1	0.5
Dec.	0.4	0.5	0.7
Jan.	0.5	0.7	0.9
Feb.	0.5	0.8	0.8
Mar.	0.5	1.9	1.4
Apr.	2.8	0.2	0.1
May	(a)	0.0	0.3
June	5.7	0.3	8.0
July	0.8	0.1	(a)
Aug.	0.2	0.2	b/0.4
			b/0.4
Sept.	0.1	0.4	0.3

a/No samples were taken in May 1975 and July 1976.
b/Two samples were taken in Aug. 1977: one on August 1, the other on August 31.

Statistical summary

	1975	1976	1977
Average (mean)	1.09	0.44	1.16
Standard deviation	1.62	0.50	2.09
Median	0.5	0.25	0.45
Average as percent of median	218	176	258
Extremes	0.1-5.7	0.0-1.9	0.1-8.0
Scatter (as percent of values less than 1.0 mg/L)	82	92	83

Source: Individual measurements from NASQAN file. Statistical summary computed by GAO.

Two tests for normality show clearly that the data for each year were not normally distributed, which demonstrates that the measurements were drawn from different sets of water quality conditions:

1. In normal distributions the sample average and median are nearly identical. In this example, they were far from identical: the yearly averages were from 176 to 258 percent higher than the medians. Clearly, the data were lopsided in each year; the averages were strongly affected by a few high values. For example,

in water year 1975, the 11 measurements produced an average of 1.09 mg/L. However, the 11 measurements were lopsided, two were far above the average (2.8 and 5.7 mg/L) and 9 well below it.

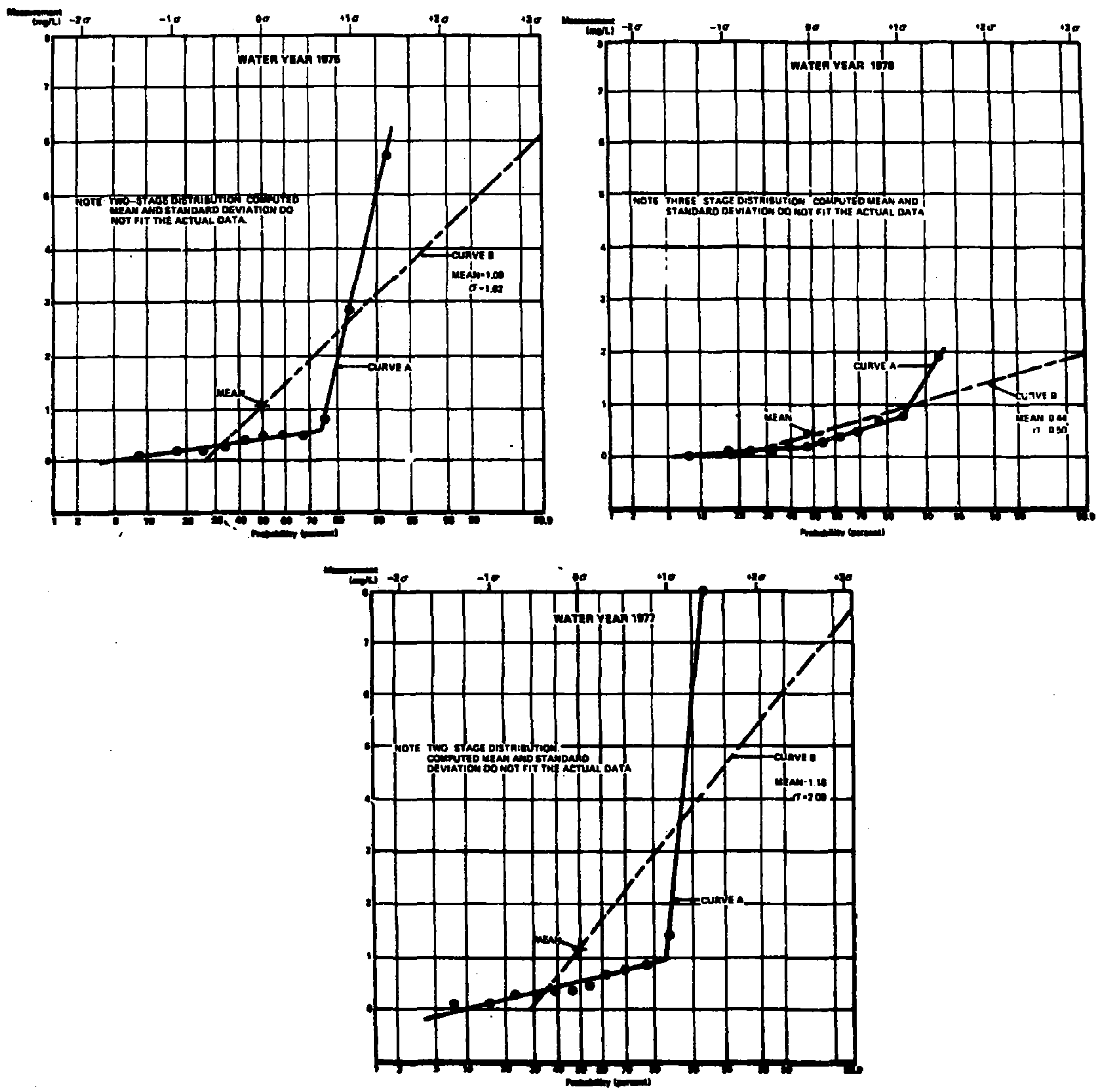
2. The charts on the following page show a graphical method of testing the St. Paul data for normality. If the data were normally distributed, all the points would lie along curve B. The charts show that the data points never followed curve B. Curve A, which is drawn from the actual measurements, shows that the data were not normally distributed in any of the 3 years. Each break point in curve A identifies a new distribution in the data. In water year 1975, there was a pronounced break point at approximately 0.6 mg/L; the data had one distribution below 0.6 mg/L and an entirely different distribution above 0.6 mg/L. In water year 1976 there were two breakpoints: one at 0.2 mg/L and the other at 0.8 mg/L; consequently there were at least three distributions in the 1976 data set. In water year 1977, there was an abrupt break at about 1.0 mg/L. Notice that the break points were different in every year; there was no stable pattern from year to year. The extreme high values were different in every year. Although the highest values each year always came between March and June, the monthly patterns were not stable.

The data show that for unknown reasons water quality fluctuated at St. Paul during each year and from year to year. These fluctuations did not follow a normal distribution, and the customary statistics (the average and the standard deviation) are of little help in describing the fluctuations. In how many months did the concentration drop because riverflows increased? (When rainfall and snowmelt swell the riverflow, concentrations of all dissolved substances--such as nitrite and nitrate--are lowered by simple dilution.) Are the fluctuations explained by agricultural practices? Did farmers actually use more fertilizer in 1977? Why was the value of 8.0 in June 1977 so much higher than any of the values in preceding years? None of these questions can be answered from the data or from the summary statistics.

<u>Taking more samples does not convert a hodgepodge of data into a normal distribution</u>

Because water quality can change so rapidly, increasing the number of samples by several hundred percent cannot be expected to change data into a normal distribution. Instead, the increase can make a nonnormal distribution more obvious. These points are clearly demonstrated in the following example.

Comparison of distribution of actual nitrite plus nitrate (as Nitrogen) measurements taken during water years 1975, 1976 and 1977 with statistically calculated distribution. NASQAN station on the Mississippi River at St. Paul, Minnesota.



KEY.

● = Actual measurements recorded by the Survey.
+ = Mean for year based on actual measurements.
σ = Standard deviation
Curve A = Graphic illustration of the distribution of the actual measurements.
Curve B = Statistically calculated distribution, based on the mean.

In water year 1978 the Survey took monthly samples of suspended sediment at Kiona (a NASQAN station on the Yakima River in Washington); 12 samples were taken that year. In water year 1976, the Survey took 63 samples--425 percent more than in 1978. The charts on page 38 show that the data were not normally distributed in either year. Increasing the number of samples by several hundred percent did not make the data more normal; instead, the extra samples made the nonnormality even more obvious.

In 1978 the suspended sediment data had at least two distributions; the break point was at 38 mg/L. (See Curve A.) In 1976 the data were a mixture of at least three distributions. The break points were at 35 and 72 mg/L. The large increase in sampling in 1976 identified yet another distribution (the break point at 72 mg/L), which did not show in the sparser 1978 data.

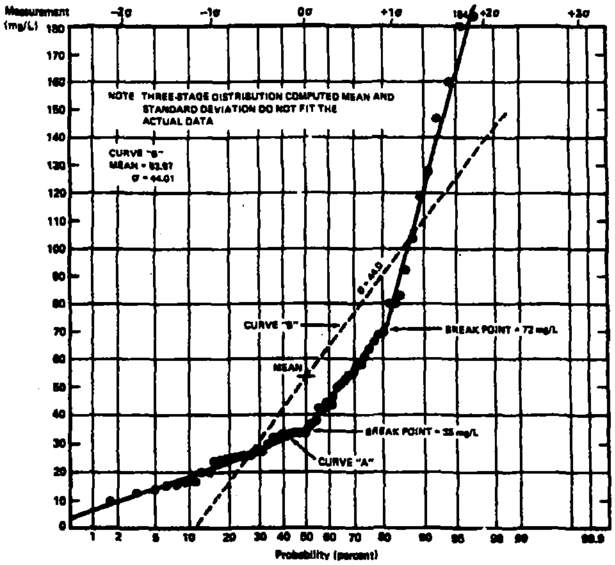
The standard statistical tests of significance cannot be correctly applied to data from the networks

Statisticians have known for decades that the standard significance tests cannot be legitimately applied to many kinds of data. The standard tests have rigorous requirements about the data to be tested. When these requirements are not met, the tests are invalid. Because data from the networks do not meet the rigorous requirements, trends derived from the data through standard statistical tests cannot be accepted as valid.

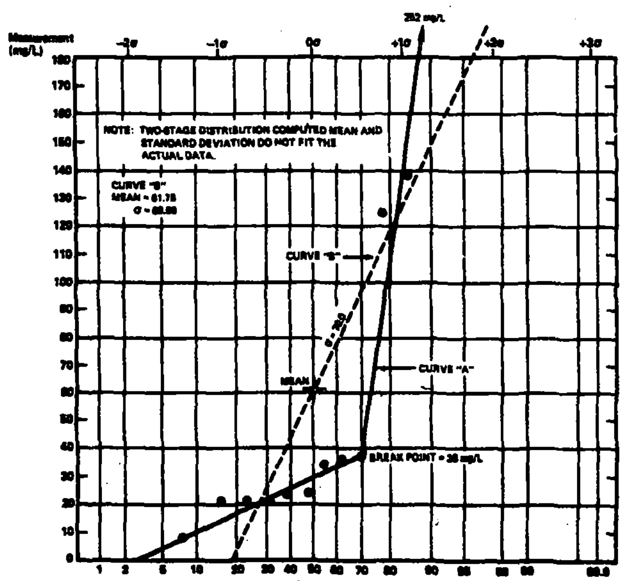
As early as 1974, scientists in the Survey avoided using standard tests to detect trends in water-quality data from networks. They recognized that the standard significance tests require data analysts to make "rigorous assumptions * * * about the underlying distributions of data used in the analysis." 1/ Among many other things, these assumptions require that the data must have a normal distribution and must come from a single, unchanging, normal distribution.

1/An Assessment of Areal and Temporal Variations in Streamflow Quality Using Selected Data from the National Stream Quality Accounting Network. U.S. Geological Survey Open-File Report 74-217, August 1974.

Comparison of distribution of actual suspended sediment measurements taken during water year 1976 with statistically calculated distribution.
NASQAN station on the Yakima River at Kiona, Washington



Comparison of distribution of actual suspended sediment measurements taken during water year 1978 with statistically calculated distribution.
NASQAN station on the Yakima River at Kiona, Washington



KEY

● = Actual measurements recorded by the Survey.

+ = Mean for year based on actual measurements.

σ = Standard deviation.

Curve A = Graphic illustration of the distribution of the actual measurements.

Curve B = Statistically calculated distribution, based on the mean.

We have already shown that water-quality data are more likely to have an odd or lopsided distribution than a normal one, and are more likely to represent a changing mixture of distributions than a single, unchanging, normal distribution. Although the Survey avoided the standard significance tests in its 1974 report, CEQ has used them in its analysis of trends in NASQAN data.

"Significance" has a special, rather stilted meaning in statistics. In statistics, significance refers to chance. When we say that differences are statistically significant at the 10-percent level, we mean that

--if we could have repeated our measurements hundreds of times;

--if there really were no changes in water quality;

--and if all the differences in the measurements were really due to measurement errors rather than to real differences in water quality;

--then we could have found differences as large as those we actually saw under these highly artificial conditions only 10 percent of the time. In other words, assuming that there really were no differences in water quality and that all the variation we saw was really measurement error, we would have been misled into believing that there were real differences 10 percent of the time.

In standard statistical tests "chance" is mathematically defined as an unchanging normal distribution. When the data do not meet the requirements of these tests, these tests cannot be correctly used.

IMPROPER PROCEDURES IN THE FIELD

It is well known that the results of water quality tests can be no better than the samples used for the tests. Improper procedures in the field add uncertainty to individual sample results and weaken subsequent interpretations of water quality status and trends. The uncertainty is increased with farflung networks because of the greater likelihood of inconsistent sampling practices.

Erroneous field procedures weaken network data in two ways. When samples have been improperly collected or handled, the test results are not valid, yet the erroneous data remain in the records. When the faults are corrected, the data becomes inconsistent, but the records are not marked to show when the inconsistency began.

Many mistakes in the field affect accuracy or representativeness. For example, water samples can become contaminated or can change concentration levels if improperly collected, preserved, or stored. Even the simplest water quality tests, such as onsite measurements of conductivity, can be done improperly.

The Survey regularly reviews its sampling programs, including observation of technicians in the field. It also has reviewed sampling programs of some of the States that participate in its cooperative projects. EPA also conducts cooperative evaluations with State agencies. In June 1979, EPA issued a directive requiring all recipients of EPA grants, including States, to follow a mandatory quality assurance program.

As shown in the following examples, the Survey often uncovers improper practices that create errors and inconsistencies in network data. 1/

--Survey reviewers found that improper procedures were used for several of the most common tests conducted in the field, such as water temperature, specific conductance, pH, and DO.

In one district, Survey reviewers observed that a technician brought samples back to the district office before doing most of the field tests. The reviewers recommended that pH always be measured in the field and that bacteriological samples be filtered and incubated in the field.

In another district, reviewers discovered that field measurements of conductivity and pH were made from the riverbank, but the rest of the samples were taken from midstream. The reviewers stated that this practice might not yield representative values.

Survey reviewers uncovered the practice of calibrating pH meters in the district office rather than at the sampling site. The reviewers noted that pH meters calibrated prematurely could be out of adjustment by the time they were used in the field.

--In several districts, Survey reviewers found fault with the way technicians were washing filters and sterilizing equipment for bacteriological samples.

--In one district, samples to be tested for nutrients were held too long before being shipped to one of the Survey's central laboratories. The Survey's reviewers found

---

1/We cite only the Survey's quality assurance reviews because the reports were readily available at the Survey's headquarters. EPA does not consolidate the results of its program, which is carried out by regional offices.

samples 9 days old in the district's refrigerator, compared with the maximum acceptable time of 7 days for storage and shipment. 1/

--In one of its reviews of State sampling procedures, the Survey found many weaknesses. For example, field personnel sent raw water samples to the State office each month where they were filtered, split, belatedly preserved, and shipped to a Survey central laboratory for analysis. According to the Survey reviewers, this procedure could result in invalid measurements because of changes in the sample between the time of collection and analysis--days or weeks later. The reviewers recommended that field filtration of samples, adherence to proper preservation procedures, and expeditious shipment of samples to the central laboratory be initiated as soon as possible.

DELAYS DURING STORAGE AND SHIPMENT

Promptness is important for many water quality tests, particularly for unstable constituents such as nutrients (nitrogen and phosphorus). Their concentrations can change substantially if tests are delayed, thereby creating invalid, inconsistent and erroneous network data. As illustrated above, the Survey has observed undesirable delays by its own districts and by States.

At our request, the Survey recorded the shipment and analysis time for all 170 NASQAN samples for nutrient analysis arriving at its two central laboratories from August 1 to 11, 1979. Of the five nitrogen and phosphorus tests involved, three may be performed up to 7 days after the water sample is collected if it has been properly preserved and stored. The test for total ammonia nitrogen should be performed less than 24 hours after the sample is taken from the river. These holding times are given in the Survey's "National Handbook of Recommended Methods for Water Data Acquisition."

The Survey's records showed that 96 percent of the 170 samples arrived at the laboratories 2 days or more after they were collected and 34 percent of the 170 samples arrived more than 7 days after collection. No samples were analyzed within 1 day--the maximum time recommended for total ammonia nitrogen. Ninety-three percent of the samples were not analyzed within 7 days. Only 44 percent were analyzed within 14 days--twice as long as the maximum holding time. Six percent of the samples

---

1/The maximum holding time of 7 days is specified in chapter 5 of the "National Handbook of Recommended Methods for Water Data Acquisition." This handbook is being prepared in increments by a Federal interagency working group led by the Survey's Office of Water Data Coordination. Chapter 5 was published in September 1977.

were held more 29 days. Because the maximum holding times were exceeded, much of the data from these samples is questionable.

The Survey's laboratory coordinator explained that workloads at the central laboratories during August and September 1979 had increased by 40 percent. He said that the results were the worst example of both the districts' collection and shipment of samples and the laboratories' analyses of them. We believe the problem may be chronic. EPA's evaluations of the Survey's two central laboratories in February 1977 and January 1978 showed that holding times were often exceeded even before the samples arrived at the laboratories. We also checked on 23 nutrient samples taken from November 1974 to January 1977 at the NASQAN station on the James River in Virginia. We found that the elapsed time between collection and testing for the 23 samples was never less than 11 days--far in excess of the 7-day limit. Ten of the samples took longer than 25 days.

INCONSISTENT LABORATORY PERFORMANCE

The credibility of network water quality data is further eroded by laboratory errors, variable performance by analysts, and inherent limitations of equipment and analytical methods. Consistent and accurate laboratory performance is important for comparing data over time and among geographic areas. Confidence in laboratory performance becomes even more important when, as done under the networks, infrequent samples are used to assess water quality. A certain degree of error is expected in all analytical methods. However, EPA and Survey tests of laboratory performance have revealed wide variances.

EPA and the Survey realize the importance of accurate and reliable laboratory measurements and have established testing programs for analytical performance. 1/ Participation by Federal, State, and private laboratories in EPA and the Survey's programs is voluntary. These testing programs show that inconsistent performance is a persistent problem. Inconsistent performance is a greater management problem for EPA's national network of 1,000 stations than for the Survey's NASQAN because many more laboratories are involved in the EPA network. The Survey relies extensively on its two central laboratories, while analyses for EPA's network are done by various State and other laboratories throughout the Nation.

1/EPA's performance evaluation program involves distilled water samples containing true value concentrations known only to EPA. Under the Survey's analytical evaluation program, participants analyze natural water samples for various constituents.

Variability among laboratories

EPA and the Survey's performance programs have shown widely varying results among laboratories concurrently analyzing similar samples. In 1977, for example, 26 of 83 laboratories participating in EPA's program deviated from the average of 0.36 mg/L total phosphorus by more than 22 percent. Participants in the Survey's 1977 test using a similar concentration of total phosphorus varied almost as much, with 6 of 34 laboratories deviating from the average by more than 23 percent.

The following table shows the performance by six laboratories that participated in two other EPA quality control tests for phosphorus.

Quality Control Tests for Total Phosphorus at Three EPA Regional Laboratories and at Three State Laboratories

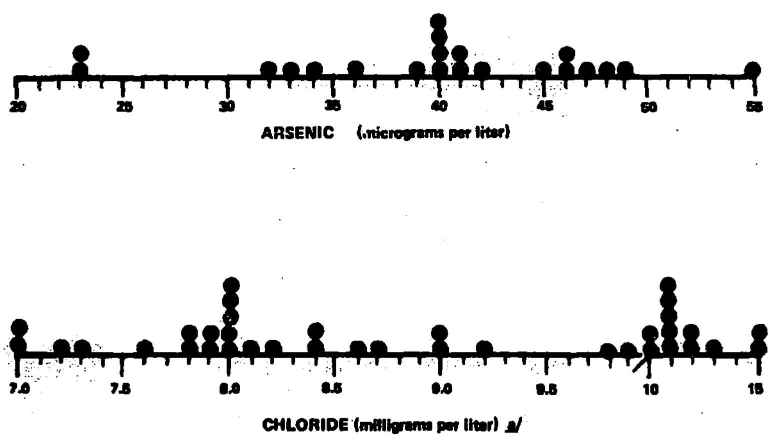
	Tests made in 1976 True value = 71 ug/L			Tests made in 1977 True value = 1.640 mg/L		
Lab #	Reported value (ug/L)	Difference from true value (ug/L)	Percent Difference	Reported value (mg/L)	Difference from true value (mg/L)	Percent Difference
1	66	-5	7	0.63	-1.01	62
2	61	-10	14	1.730	+0.09	5
3	73	+2	3	1.60	-0.04	2
4	120	+49	69	1.55	-0.09	5
5	<100	(a)	(a)	1.7	+0.06	4
6	80	9	13	1.68	+0.04	2

a/Accuracy cannot be calculated for "less than" (<100) values. Exact numbers are needed to calculate accuracy.

In 1976, four of the six laboratories gave fairly accurate answers. The accuracy of laboratory 5 cannot be calculated because of the way it reported the results. Laboratory 4 overestimated the true value by 69 percent--the largest inconsistency in the table. In 1977, five of the laboratories gave accurate answers. This time laboratory 4, which had greatly overestimated the true value in 1976, performed creditably. But laboratory 1, which had reported accurate results in 1976, underestimated the true value by 62 percent. In summary, the laboratories gave inconsistent answers and two of the laboratories gave inconsistent performances in 2 consecutive years.

The Geological Survey publishes results of its tests in tabular form and scatter diagrams. The diagrams illustrate quite clearly the inconsistent performance of the participating laboratories. As can be seen in the following example, participating laboratories reported widely scattered results for arsenic and chloride.

GEOLOGICAL SURVEY DIAGRAMS OF MEASUREMENTS
FOR ARSENIC AND CHLORIDE REPORTED BY
LABORATORIES PARTICIPATING IN ITS 1977
ANALYTICAL EVALUATION PROGRAM



a/The scale for chloride is compressed for the higher numbers beyond 9.9 milligrams per liter.

The scatter diagrams do not show the entire range of measurements because the most extreme values were excluded. In the case of arsenic, a total of 25 laboratories reported values ranging from 3 to 460 ug/L. However, the Survey excluded four extreme values outside the range of 23 to 55 ug/L from the diagram. For chloride, the actual range reported by 42 laboratories was from 3.0 to 17 mg/L, but the Survey excluded five extreme values from its scatter diagram.

Variability over time at individual laboratories

Consistent laboratory performance over time is needed in order to produce comparable measurements for trend analyses. The example of total phosphorus on page 43 shows some of the inconsistencies from year to year at individual laboratories. A laboratory may perform creditably one year and poorly the next.

EPA and the Survey do not routinely evaluate consistency over time of laboratories submitting performance test data to them. In the absence of agency evaluations, we reviewed the results for 10 laboratories that participated in EPA and Survey performance tests in 1976, 1977, and 1978.

We found that inconsistency over time was common. A laboratory testing similar concentration of individual constituents at different times would report dissimilar measurements, with one being above the expected value and the other being lower. For example, in 1976 one State laboratory reported a total zinc concentration that was 44 percent more than the 16 ug/L true value established by EPA. In a similar test in 1978, the laboratory's measurement was 25 percent less than the true value of 12 ug/L. Even if laboratories avoid being above and below the expected values in successive tests, they may consistently overestimate or underestimate the true value. For example, another State laboratory participating in EPA's program reported measurements of total dissolved solids that were 28 percent too high in 1976 and 57 percent too high in 1978.

LIMITATIONS OF THE INFORMATION SYSTEMS

Both EPA and the Survey use computerized information systems for storing and retrieving the masses of water quality data they acquire. These systems are used for water quality assessments by these and other agencies, such as CEQ, as well as consultants, contractors, and other researchers. The systems are intended to provide easy access to vast amounts of historical water quality data. However, neither information system is designed or used to record changes in local conditions, or errors and inconsistencies in field or laboratory performance that affect the credibility of the data.

STORET operations

EPA's information system, STORET (storage and retrieval system), was started in 1964 by the U.S. Public Health Service, one of EPA's predecessors. It includes a water quality file and files for other categories of information, such as municipal waste facilities, fishkills, and sewage facility construction contracts. The water quality file is used to record the monitoring results from over 200,000 (1975 EPA data) locations throughout the United States. Many different organizations can enter or retrieve water quality data, including State, local, and various Federal agencies. The Geological Survey, through a special arrangement with EPA, transfers to STORET monthly all water quality information entered into its own system.

STORET performs automatic edit checks on about 200 of the most frequently sampled water quality characteristics; however, these checks merely identify erroneous data beyond extreme upper or lower limits. For example, the STORET edit process will reject water temperatures outside the range from -2 degrees to 37 degrees centigrade. Any temperatures reported within that range will pass the STORET edit check. EPA expects participating agencies to provide detailed quality control over their data.

WATSTORE operations

The Survey established its National Water Data Storage and Retrieval System (WATSTORE) in 1971. The system includes a water quality file and several other files, including stream peak flow, daily values for some flow and quality data, and a groundwater site inventory. The water quality file in WATSTORE is also entered into STORET monthly.

Most of the data in the WATSTORE water quality file are submitted by the Survey's central laboratories through computer terminals. Other data are submitted by Survey district offices and Federal, State, or local agencies participating with the Survey in cooperative water quality monitoring projects.

WATSTORE has three mandatory computerized edit checks for data entering the water quality file. All three checks (alert system, data consistency, and chemical logic) are required for data that do not originate in the Survey's two central laboratories. Data from the central laboratories are screened by the alert and data consistency edits. The Survey relies on the central laboratories to perform the chemical logic checks before submitting data. Data not meeting the consistency check are not allowed to enter WATSTORE. Data not meeting the other edit checks are allowed into WATSTORE, but are reported separately for voluntary followup and possible corrective action by the office that submitted the data.

Computerized data lack information needed for accurate interpretation

Through the EPA and Survey computerized systems, sampling results are reduced to a collection of numbers with few indications of their quality. Neither agency qualifies the water quality data with information on local conditions affecting the individual samples, and neither agency reports the quality of field and laboratory work for each sample. Consequently, analyses of water quality conditions and trends based on STORET and WATSTORE data have to overlook these significant sources of error.

The recent study performed by the Survey on shipment and analysis time for nutrient samples (see p. 41) shows one type of error that is not identified in either computerized system. Because of shipment delays and heavy workloads in the Survey's laboratories, most of the nutrient samples were not analyzed within the prescribed 7-day period. As a consequence, it is likely that some of the concentrations of the nutrients changed before the samples were analyzed. The extent of change, which may vary for each sample, cannot be determined. These measurements were filed in WATSTORE and STORET without any notation about the delayed analyses. This situation is not unique. Delayed shipments and various errors in the field are fairly common; however, these problems are not noted in STORET or WATSTORE.

THE JAMES RIVER, VIRGINIA--A CASE STUDY OF WATER QUALITY COMPLEXITIES

Agencies attempting to interpret nationwide water quality conditions and trends from network data face imposing difficulties. As discussed in the preceding sections, many uncertainties surround network data. Many local circumstances can change water quality or affect the representativeness and comparability of periodic samples. These conditions change with time and vary from location to location. A clear understanding of the local circumstances surrounding individual samples and the rivers or streams being monitored is important if sound interpretations of water quality changes are to be made.

As part of our review of these complexities, we and our consultant studied water quality data and CEQ's interpretations of the data for one monitoring site on the James River, Virginia. The following is a brief summary of our study. (Details are in app. VII.)

CEQ used network data for the James River in a national assessment of water quality changes

In December 1978 we asked the Survey for examples that, in the Survey's judgment, best demonstrated the use of fixed-station water quality information. The Survey cited many examples, including several CEQ annual reports on environmental quality. We selected the then most recent CEQ report (December 1978) for our case study.

For its assessment, CEQ applied standard statistical analyses to 3 years of data from all NASQAN stations in operation during 1975-77. CEQ summarized waterquality changes in the following table.

Water Quality Changes (note a) at NASQAN Stations
1975-77

Water quality characteristic	Percentage of stations		
	Improved	No change	Deteriorated
Fecal coliform bacteria	7.3	88.9	3.8
Inorganic nitrogen	5.8	86.7	7.5
Organic nitrogen	3.8	83.4	12.8
Total phosphorus	4.3	83.2	12.5
Dissolved oxygen	4.5	92.9	2.6
Fecal streptococci bacteria	1.8	87.3	10.9
Dissolved solids	4.1	74.1	21.8
Dissolved zinc	9.1	86.5	4.4
Total zinc	12.8	85.8	1.4
Phytoplankton	2.0	93.6	4.4

a/Indications of change tested for statistical significance at the 90-percent level.

GAO note: Although CEQ stated that it tested indications of change for statistical significance at the 90-percent level, this statement is wrong. The indications of change were actually tested at the 10-percent level.

The James River at Cartersville, Virginia, was one of several hundred NASQAN stations CEQ used to calculate the changes reported in the table. CEQ concluded from its statistical analysis of Cartersville data that there was

--significant improvement for DO, organic nitrogen, dissolved zinc and total zinc;

--significant deterioration for dissolved solids; and

--no significant change for fecal coliform bacteria, inorganic nitrogen, total phosphorus, fecal streptococci bacteria, and phytoplankton.

We concentrated on water quality characteristics that, according to CEQ, had changed significantly.

Why we selected the James River station for review

We selected the station for detailed review because (1) it was one of the few stations that CEQ's analysis showed had statistically significant changes in water quality, (2) the changes involved both improvements and deterioration in water quality, and (3) the station was exceptionally data-rich compared with normal network stations. The NASQAN station also served as a station in EPA's NWQSS network, so samples were taken more frequently than for most of the stations involved in CEQ's study. Also, the State of Virginia sampled water quality at Cartersville and had designated the site for EPA's Basic Water Monitoring Program network. CEQ did not use the State data, but we often used them in our study.

More than statistical analysis is needed to determine if water quality is getting better or worse

CEQ concluded that dissolved solids had increased (a worsening condition), dissolved zinc and total zinc had decreased (an improving trend), and DO had improved. Relying strictly on standard statistical analysis, the Survey's data for the 3-year period do show the trends reported by CEQ. However, CEQ did not consider the important relationship between water quality and riverflow, a powerful determinant of water quality. For some constituents, such as DO, detailed information on other factors is needed to assess whether water quality is getting better or worse.

Total dissolved solids increased as riverflows decreased

Each year during 1975-77 the James River had successively lower average and minimum flows. The Survey's measurements of total dissolved solids rose as the riverflows fell. Total dissolved solids are roughly synonymous with saltiness. CEQ's statistical trend appears to be a normal consequence of the deepening drought. As the drought deepened, there was less rainfall and surface runoff to dilute the saltier groundwater.

Higher total zinc concentrations correlated with higher riverflows

Most of the measurements of total zinc at the Cartersville station were quite low. The higher readings generally occurred during high or very unstable (rapidly changing) riverflows. These dramatic riverflows were especially common during the early part of the 3-year period, when most of the higher zinc values were reported. CEQ's trend seems to have been created by dramatic riverflows and inadequate sampling at peak riverflows. The Survey took more samples during high or unstable riverflows in the earlier part of the 3-year period than in the later part. Fixed-frequency sampling made the Survey miss most of the peak riverflows after September 1975.

Dissolved zinc cannot be explained from available data

As with total zinc, dissolved zinc measurements were usually very low. Four higher values were recorded between June 1974 and June 1975, but all other measurements were at or near the detection limits. We matched the dissolved zinc measurements with riverflows and found that the patterns could not be explained as either point sources or nonpoint sources of pollution. Officials of the Survey's laboratory that produced the four higher measurements insisted that those measurements were correct. The first part of the 3-year period was weighted with four inexplicably high measurements which caused the trend reported by CEQ.

Dissolved oxygen changes cannot be explained without more information

Based on an average of about two samples per month, CEQ concluded that DO had increased (improved) during the 3-year period. In our opinion, a few samples a month from this rapidly changing river cannot show a trend in DO. DO is the result of many complex processes. As we pointed out on page 19, DO can fluctuate dramatically within 1 day. These changes vary in size in response to many biological, chemical, and physical processes.

The records for DO at Cartersville suggest that undefined phenomena dramatically change the DO supply in the river. Four of the major factors that can affect DO are: (1) physical reaeration, (2) biological reaeration, (3) temperature, and (4) decomposing wastes. (See p. 103 for explanation of these factors.) These factors need to be measured at Cartersville and sites upstream to understand DO in the river. However, the Survey and State have not sampled water quality immediately upstream of Cartersville, and there was insufficient data for the Cartersville site.

PRACTICAL LIMITATIONS TO FIXING THE NETWORKS

Rivers are not inflexible. They are variable. They change from moment to moment, from mile to mile, from their headwaters to the seas. The various organizational rigidities we have discussed make it impractical to fix the networks because an inflexible sampling program is not consistent with the variability of the Nation's waters.

The choice of sampling frequency, water quality tests, and sampling sites should all be dictated by the unique character of each river and by the purpose of the sampling effort. For example, if DO is an issue, DO should be measured around the clock, particularly when DO problems are most likely. In addition, many other water quality characteristics must be evaluated if DO patterns are to be properly understood (see app. VII), and these characteristics may need to be assessed for miles above and below the principal sampling site.

When pesticides are an issue, sampling programs should include pesticides that are widely used in that river basin and should concentrate on those times when pesticides are most likely to be a problem. However, the EPA-Survey cooperative sampling program for pesticides is very uniform and limited in the number of pesticides monitored. Under the National Pesticides Monitoring Program, the Survey collects water samples four times a year and sediment samples two times a year at about 150 NASQAN stations. EPA analyzes the samples. The list of pesticides is uniform at at all stations, although pesticide use varies greatly around the Nation. According to EPA officials, the list emphasizes pesticides that are produced in large quantities and those, like DDT, that persist in the environment long after they have been restricted or replaced.

The National Research Council criticized pesticide monitoring in a major report:

> "We found a surprising lack of integration between monitoring programs and data on the production and use of major chemicals. * * * there is little tendency for monitoring programs to be closely adjusted in space or time to take into account variations in production or use of the chemicals." 1/

In the Yakima River basin of Washington State (a major agricultural area) we found that EPA-Survey pesticide monitoring was not directed at local conditions. The agencies test for many pesticides not generally used in the Yakima basin and exclude many widely used pesticides. With misdirection and sparse sampling the monitoring program cannot produce meaningful data on pesticides in the Yakima River.

In general, the networks now sample once a month. Monthly sampling cannot show whether differences in measurements represent real changes in water quality or are random fluctuations. Taking more samples at fixed intervals is not the solution. On pages 35 to 37 we showed that increasing the sampling frequency by several hundred percent made the nonnormality of the data more obvious. Even weekly samples cannot represent more than a tiny fraction of the complex, changing patterns of water quality in a river. To make the networks produce meaningful data, monthly sampling at fixed stations would have to be converted into special studies tailored to each river.

CONCLUSIONS

Accurate, meaningful information on surface water quality is essential to the national effort to clean up our waterways. We believe that the three national networks which EPA and the Geological Survey established will not be able to provide credible data for assessing water quality on a national scale.

1/"Pest Control: An Assessment of Present and Alternative Technologies, Volume I. Contemporary Pest Control Practices and Prospects," The Report of the Executive Committee, National Research Council, 1975.

Can the networks reliably portray the quality of water at individual sampling locations? Will EPA, the Survey, CEQ, and other agencies be able to use network data to legitimately evaluate changes in water quality over time? Will the fixed stations provide useful information on the direct effect of the various Federal and State pollution control programs and projects? Can the Congress rely on the networks in its oversight of Federal pollution control programs? We believe the answer to these questions is no.

Periodic sampling at specific sites produces an array of numbers for individual water quality tests. The precise appearance of these numbers (for example, 0.075 mg/L) suggests a sense of accuracy for each sample and for averages calculated from them. However, individual measurements are very susceptible to error in each stage of monitoring--sample timing and location, fieldwork, and laboratory analysis. This potential error varies greatly among locations, among water quality tests, and among time periods. These inconsistencies in network data can create misleading or spurious trends.

In general, the networks sample once a month at widely spaced sites. This sampling approach cannot represent more than a tiny fraction of the complex, changing patterns of water quality in a river. Consequently, the networks produce a hodgepodge of measurements of widely varying water quality conditions. Trend analyses of mixed measurements are oversimplified and can be misleading because the analyses failed to account for the different water quality patterns.

In addition to the questionable accuracy and representativeness of sampling results, the networks cannot link water quality conditions with what caused those conditions. Inappropriate timing and location of samples and a lack of data on important water quality characteristics severely limit the usefulness of network monitoring for assessing changes in water quality.

We found no compelling reasons why reports on the Nation's progress toward cleaner water have to be based on repetitive fixed-station sampling. For a clear, sound understanding of the extent of pollution, reasons for it, and the effectiveness of corrective efforts, thorough, well-performed water quality studies are essential; the fixed station networks cannot provide that information.

Our recommendations are on page 63.

CHAPTER 4

ALTERNATIVES TO FIXED-STATION NETWORKS

FOR ASSESSING PROGRESS TOWARD CLEANER WATER

The primary justification for the EPA and Geological Survey networks is to provide data for assessing national water quality conditions and trends. In chapter 3 we showed why network data are a questionable basis for these assessments. More scientifically sound assessments and supporting data can be produced through special water quality studies. Also, other types of information on water quality conditions and programs can help assess progress in reducing pollution; this information includes changes in fish populations and reductions in discharges of pollutants.

Special water quality studies and other types of information are not novel. EPA, CEQ, the Geological Survey, and States have used them for years. If the agencies made more use of these kinds of information, we believe they could make better and more meaningful water quality assessments. Special studies can produce better scientific information on water quality, and the wide range of other related information can provide progress reports in terms everyone can understand.

SPECIAL WATER QUALITY STUDIES ADDRESS SPECIFIC SITUATIONS

Every river has a unique combination of hydrologic factors (e.g. streamflow, water temperature, and channel shape) and outside influences (e.g. land use and wastewater discharges) that control the magnitude and variability of water quality problems. Unlike network monitoring, special water quality studies are tailored to fit each unique river pattern. The studies can be designed to avoid the serious time and location biases of the networks and can focus on the water quality characteristics needed to understand the special problems of the river area being studied. EPA and the Survey recognize the merits of special studies and have promoted them in recent years.

Special water quality studies are distinguished by the emphasis on analysis and understanding. Properly performed studies can define the quality of water at specific locations, explain existing conditions, and show why they change. This information is essential for meaningful assessments of water quality and pollution-control programs.

Because special studies are designed for specific situations, they may vary widely in duration, number of samples, number of sampling sites, kinds of tests performed, and types of supplementary information. For example, fieldwork may last several days or several years. Some studies may require extensive information on sediments and aquatic life; other studies might not. A common trait of special studies is the close attention given to forces at work and their effects on water quality.

Characteristics of special studies

To produce sound water quality data, special studies must be well-planned and tightly managed. Data analysis must identify weaknesses or errors in the study, such as missing samples during critical water quality conditions or improper handling of samples. Some of the key features of special studies are discussed below. We are not proposing a new approach, but instead are summarizing what is already widely known by experts and the agencies themselves. For example, Survey Circular 715-D explains some of the features that distinguish special studies from routine monitoring.

Why is the study being done? This question should be the driving force throughout the study. To ensure that the right measurements are taken at the right times and places, designers of special studies must establish specific objectives. The objectives, in turn, dictate what water quality tests are to be made and where and when sampling is to be done.

The following examples illustrate how special studies avoid the time and location biases inherent in network sampling.

1. If the objective is to evaluate the effect of sewage treatment plants on water quality, some important steps in a special study would be to:

--Obtain operating information about the sewage treatment plants in the study area, including the volume of effluents, treatment processes, and important peculiarities about the system, such as unusual industrial inflows to the plant or seasonal patterns of effluent volume and content.

--Identify where effluents enter the river.

--Determine important streamflow and weather patterns that affect sampling timing and locations. Often, the most critical period for studying the effects of sewage treatment plants on water quality is during low-flow, warm periods. Sampling should be done during reasonably stable conditions after the river has reached equilibrium. Sudden changes in streamflow caused by storms, local showers, or releases from reservoirs create unstable conditions that greatly increase the complexities of sampling. Usually a short period of intensive sampling during stable conditions provides enough measurements for reliable statistical calculations. Measurements taken during stable, homogeneous conditions usually have normal distributions, thus permitting valid statistical comparisons of measurements taken in similar conditions in other years.

--Identify important hydrologic features, land use, and river activities that affect decisions on where to collect samples. Tributaries can add water of better or worse quality. Irregularities in the shape or flow of a river cause irregularities in quality measurements. Land use affects river water quality in different ways through surface runoff, ground seepage, or drainage from industrial, urban, residential, and rural areas. Some segments of rivers are more important for such uses as swimming or fishing than others.

--Determine what to sample. Water samples are essential, but sediment samples and aquatic life are important also. The type of samples needed vary with local conditions. Materials in effluents behave differently in rivers. Bacteria generally die rapidly. Suspended materials settle in quiet water. Ammonia is oxidized in swift, shallow water. Fish and other aquatic life are selectively harmed by some materials but not by others; for example, ammonia may kill fish but promote the growth of aquatic plants.

--Establish specific sampling plans, including timing, locations, and number of samples needed for valid statistical analysis. Timing is critical. Sampling should be done during stable riverflow and weather conditions, and frequent samples may be needed because of the inherent variations in measurements. The number of samples needed depends on the variability of each characteristic. Some measurements, such as fecal coliform bacteria, are inherently variable and may require many samples. Other measurements are less uncertain and may require fewer samples. The number of samples for each test is unique to each local situation. Some tests, such as DO, may require many measurements during night and daytime. Sampling locations are also critical, as we discussed above. In addition to the primary zone of study, samples should be taken upstream to (1) define the quality of water entering the primary zone and (2) provide measurements of characteristics, such as biological oxygen demand, that are needed to properly interpret related measurements in the primary zone, such as DO. (See p. 103 of app. VII.) For some characteristics, sampling times through the study area should be synchronized with river velocity to ensure that samples taken at different locations are from the same water mass.

2. If the objective is to describe pesticides contamination in an agricultural area, some important steps would be to:

--Identify the most important pesticides in the area. There are many pesticides and they are not used uniformly throughout the Nation.

--Identify when and where pesticides are applied. Pesticides are not applied uniformly throughout the year or uniformly to all agricultural land. Choice of pesticide, application schedule and application rate may vary considerably even in one county.

--Determine the important entry routes for pesticides chosen for the study, such as irrigation return canals, tributaries, aerial spraying, or groundwater.

--Identify important factors affecting when or where to sample, such as riverflow, climate reservoir releases, and river morphology.

--Determine the right mixture of sampling media, such as sediment, river water, and aquatic life. The type of pesticides being studied and local conditions all need to be considered in this step.

--Establish specific sampling plans, including timing, location, and the number of samples needed for statistical analysis.

Excellent sampling plans are of little value if they are not properly carried out. Tight management is as important for special studies as is good planning. Quality assurance can receive closer attention in special studies than in network monitoring. For example, onsite supervision can reduce mistakes in the field procedures and allow flexibility for timely adjustments and corrections. The project managers are on hand to ensure that samples are analyzed promptly and that the laboratory work is of high quality. With onsite supervision, pecularities in data can be quickly discovered and resolved. In contrast, errors in network data may not become apparent until well after samples are taken. These errors are often hard to identify or explain after the data are entered into WATSTORE or STORET. The networks have fragmented management for fieldwork, laboratory work, and interpretation of the resulting data. Special studies have unified management for all phases of the work.

Special study results can be used for national assessments

A well-managed special studies program, vigorously pursued, in a few years could cover most of the Nation's major river systems with significant river quality problems or other major interests. If such a program were supported and operated permanently, the

United States would be well on its way toward a scientifically sound basis for assessing the quality of the Nation's rivers. Through special studies, important cause-effect relationships for water quality conditions can be determined for small and large segments of rivers. The solid footing provided by initial studies would enable subsequent study teams to reach meaningful, reliable conclusions about changes in water quality and reasons for the changes. Assessments of water quality conditions and trends would be limited to the individual river areas covered by the studies. But, nationwide perspective could be produced by accumulating the assessments into periodic national reports. We believe this type and quality of information is far more useful than the information provided by network sampling.

GEOLOGICAL SURVEY HAS PERFORMED MANY SPECIAL STUDIES

The Survey has done many special studies through its own intensive river quality assessment program and its cooperative program with Federal, State, and local agencies. These programs have helped assess present water quality and offer opportunities for future assessments.

Intensive river quality assessments

Since 1973 the Survey has conducted a program of intensive river quality assessments. Overall objectives for the program are to develop techniques for studying water quality and to produce meaningful water quality assessments. Three studies have been completed and four others are underway.

The Survey established its river quality assessment program partially in response to the recommendations of its Non-Federal Advisory Committee on Water Data for Public Use. 1/ From its inception, the committee was concerned about the lack of suitable information for river basin planning and water quality management. In 1971 the committee formally recommended that the Survey should institute a program of intensive river quality assessments.

The first special study, initiated in 1973 in Oregon's Willamette River Basin, focused on problems of concern to State authorities -- DO depletion, nuisance algae, trace metals,

---

1/In response to Bureau of the Budget Circular A-67 (see p. 12), the Survey established two advisory committees, a Federal and a non-Federal. The latter, designated as the Non-Federal Advisory Committee on Water Data for Public Use, is comprised of representatives of national, State, and regional water-oriented organizations; professional and technical societies; and the academic community.

erosion, advanced waste treatment, and low-flow augmentation. The study was completed in 1976 and was reported in Survey Circulars 715A-715M. The Director of the Survey described the broad nature of the study in congressional hearings held in 1977: 1/

> "In essence, Mr. McDade, what has been involved is a study of the water quality, the chemistry, biology of the river, the sources of the contaminants, the change in water quality, including dissolved oxygen and so on, in various parts of the river, and analysis of discharge and inflow as related to seasonal changes in water quality and so on."
>
> "It has been a case of studying the anatomy of the river."

The next two studies addressed (1) problems of heat, urban and industrial wastes, flow regulation, and sediment in the upper Chattahoochee River Basin in Georgia, and (2) potential environmental problems of energy development projects (coal) on the Yampa River Basin in Colorado. Other studies underway involve the Lower Potomac River, Maryland/Virginia; Apalachicola River, Florida; Carson-Truckee Rivers, Nevada/California; and the Schuylkill River in Pennsylvania.

These studies mark what could be a shift away from fixed-station monitoring. Professor C.J. Velz, 2/ in a statement dated May 25, 1978, to the Co-Chairmen of CEQ's Task Force on Environmental Data and Monitoring, emphasized the need for this reorientation. Referring to the first two studies, Professor Velz stated:

> "It was demonstrated in the Willamette and the Chattahoochee studies that short-term, intensive, synoptic-type data, taken under stable hydrologic conditions and known pollution loadings (point and non-point sources) were indispensable in definition of cause-effect relationships and in establishing reliable bases for evaluation of planning alternatives. * * *
>
> "Drawing firm conclusions regarding river-quality problems from monitoring or surveillance data without the supporting framework of intensive river quality assessment studies, is extremely unreliable, if not

---

1/Department of the Interior and Related Agencies Appropriations for fiscal year 1978, Subcommittee of the House Committee on Appropriations, Feb. 28, 1977.

2/Professor Velz was the initial leader of the Working Group on River Quality Assessment of the Survey's Non-Federal Advisory Committee.

impossible, and in instances may actually be misleading. * * *."

He recommended that the river quality assessment program be strengthened and expanded and that periodic reassessment of rivers be done at 5- to 10-year intervals to (1) document improvement or decline in water quality and (2) evaluate the effectiveness of pollution control programs.

Subsequently, in December 1979, Professor Velz informed us that he did not regard the national networks as adequate substitutes for more thorough river studies for either national water quality assessments or pollution program management. In his opinion, one good short period of intensive sampling at a few well-selected locations to define the quality profile along the river course will provide much more reliable and useful data at less cost than years of sampling once or twice a month under all kinds of instability at widely spaced locations.

<u>Other special studies</u>

The Survey has done a variety of water resource studies in cooperation with Federal, State, and local agencies. Many of the projects relied extensively on routine fixed-station sampling, but other projects involved some of the features of special studies that we discussed on pages 54 through 56. For example:

--In cooperation with the National Park Service, the Survey studied water quality in two drainage basins of the Redwood National Park in California from September 1973 to September 1975. The purpose of the study was to determine existing water quality conditions and to identify possible effects of logging on water quality.

The Survey performed its study in three distinct phases. During several low-flow periods, samples were taken at 1- to 2-hour intervals at designated locations. Frequent samples were also taken during several storms at certain locations. In addition, some samples were taken between these periods to help determine seasonal water quality variations. The Survey established about 40 temporary monitoring sites for this study. Some of the sites were used for all three phases, while others were used for only parts of the study. Also, the range and frequency of water quality tests varied among stations.

--In a cooperative project with California, the Survey and the State Water Resources Control Board studied water quality in the Russian River during the summers of 1977 and 1978. This project, although less rigorous than the Redwood Park study, illustrates some of the extra effort that the researchers believed was needed to determine and understand river water quality.

In 1973, the State of California banned effluent discharges within the Russian River Basin during the low-flow period of May through September. Excessive bacterial contamination and nuisance growths of algae had impaired recreational use of the river.

The cooperative study was designed to determine both the extent and causes of water quality problems after the no-discharge rule was in effect. The project team concentrated on constituents that were closely related to the bacterial and algae problems. Seven of the 18 sampling sites were located in a short stretch of the river that was used extensively for recreation. To help determine the extent and sources of pollution, the team used aerial photographs taken by the Army Corps of Engineers and various forms of land-use and land-cover data. The team planned to sample upstream from where violations were noticed until the problem sources were found.

In recent years the Survey's Non-Federal Advisory Committee on Water Data for Public Use has encouraged the Survey to undertake more intensive river quality studies under its cooperative program.

EPA HAS INCREASED ITS EMPHASIS ON SPECIAL WATER QUALITY STUDIES

EPA gave special water quality studies a boost in 1977 when it requested States to use intensive surveys rather than fixed stations for pollution control program management and water quality assessments. Previously EPA expected States to rely extensively on large fixed-station networks for their assessments.

EPA, in its 1977 Basic Water Monitoring Program guidelines, stated that:

> "The intent is for the State to use the intensive survey as the primary vehicle in determining whether water quality conditions are improving or getting worse."

EPA also requested States to use intensive survey results as a basis for reports required by section 305(b) of the Clean Water Act.

Some States already were using intensive surveys to establish baseline water quality data or to evaluate changes in water quality. For example, a special 1-year water quality study was conducted on the South Platte River at Denver from July 1976 to June 1977. The primary purpose of the study was to determine the recreational suitability of the river. It was a joint effort of the Colorado Department of Health, Denver Department of Health and Hospitals, and the Tri-County Health Department. (Some of the measurements are shown on p. 25.)

During the July-September period, the project team took samples every other day at 18 sites along a stretch of 24 miles. In winter months the sampling frequency was changed to weekly. About 2,000 samples were taken during the study. The team concentrated on fecal coliform bacteria, DO, ammonia nitrogen, un-ionized ammonia, pH, temperature, and certain metals. One laboratory was used for analysis of the samples.

SHIFTING TO SPECIAL STUDIES NEED NOT INCREASE MONITORING COSTS

One reason EPA and the Survey gave for using fixed-station networks is that it is a relatively inexpensive way of getting national perspective on water quality status and trends. As shown in chapter 3, national perspectives on water quality based on network data are questionable.

The question of cost should be considered in context of overall water quality monitoring efforts of EPA, the Survey, States, and other agencies. If the network sampling programs were terminated, the estimated $11 million per year currently expended for those networks could be shifted to special water quality studies. EPA already requested States to shift some monitoring funds for this purpose under its 1977 Basic Water Monitoring Program. At that time, EPA estimated that a typical intensive water quality study could cost about $14,000, which is about three times the annual cost EPA estimated for individual network stations. While the $14,000 may be useful to States as a general guide, in actual practice, the cost of special studies can vary as substantially as the purpose and scope of individual studies. For example, a Survey official estimated that the cooperative study of the Russian River in California (discussed on p. 59) cost about $200,000. Also, the Survey estimated the cost of a typical study performed under its intensive river assessment program at about $750,000.

Other programs funded separately from the networks are also available for assessing water quality around the country. For example, the Survey's river quality assessment program and special projects, such as the Redwoods Park study discussed on page 59, have been funded separately in Survey budgets.

With proper planning, the various funding sources now available should allow EPA and the Survey to undertake a program of special studies that should produce a substantial amount of good water quality data useful for assessing current and future water quality and, more importantly, the reasons for changes in the quality.

OTHER INDICATORS OF PROGRESS TOWARD CLEANER WATER ARE AVAILABLE

Detailed chemical analyses of surface water are not the only means of assessing national progress toward cleaner water. EPA, States, the Survey, and other agencies such as CEQ, have tapped many different sources for information on water quality conditions and changes. For example, EPA and CEQ in recent years have cited improvements evidenced by increased use of waterways for swimming, fishing, and canoeing. The return of fish to previously fouled waters and changes in water color, odor, and foam are other indicators cited by EPA, CEQ, and States.

In 1977 EPA and many States included a variety of information relevant to water quality in reports required by section 305(b) of the Clean Water Act. For example, EPA in its report to the Congress summarized individual State reports of biological monitoring, improvements in fishing, and reductions in industrial waste discharges. Concerning biological monitoring, EPA stated in its report that:

> "This type of analysis is extremely useful in assessing the effects of pollution control efforts since it describes water quality in terms of the actual goals of the Act ('water quality which provides for the protection and propagation of fish, shellfish, and wi[illegible] life . . .') rather than in terms of chemical cons[illegible]tuents."

State monitoring has ranged from simple counts of fish to detailed analyses of the diversity of fish and smaller aquatic life and bioassay analyses of fish tissue for trace metals or organic chemicals. As stated previously on page 10, States are required to perform biological monitoring under the Clean Water Act. EPA has encouraged States to increase and improve their biological monitoring efforts.

EPA also has been involved in biological monitoring in two other important ways. The EPA Administrator is required under the Federal Environmental Pesticide Control Act of 1972 (7 U.S.C. 136) to establish monitoring activities for pesticides. As partial fulfillment of this requirement, EPA arranged to continue an ongoing monitoring program started in the mid-1960s by the interagency Federal Committee on Pest Control. The Fish and Wildlife Service, Department of the Interior, does the actual sampling of fish and wildlife. The primary objective of this National Pesticide Monitoring Program is to determine, on a nationwide basis, the levels and trends of selected pesticides and toxic metals. The second important involvement by EPA in biological monitoring is the research and monitoring done by or under the direction of

EPA's Environmental Research Laboratory in Duluth, Minnesota. The Laboratory has been developing techniques for analyzing the presence of toxic substances in freshwater fish and has successfully put to practice some of these techniques in field surveys.

CONCLUSIONS

Sound information of decisionmaking quality can come from special studies of water quality problems and from related information, such as fish censuses. Special studies could be planned in accordance with the relative importance of pollution problems and the pace of corrective action. Special studies can reliably document the effectiveness of cleanup programs and help decisionmakers plan future actions.

EPA is encouraging States to perform more special studies and wants them to rely extensively on these studies, rather than fixed-station monitoring, for statewide assessments of water quality. Since EPA needs the same kind and quality of information, we believe that Federal funds would be better spent if EPA also used that approach for its needs.

The Geological Survey has done several special studies in recent years. With more effort on special studies, the Survey could add to the Nation's scientific knowledge of water quality changes and the reasons for the changes. In our opinion, this added knowledge would be far more worthwhile than continuing the networks.

There is much room for increased use of other indicators of progress toward cleaner waters. For example, biological monitoring is an effective and growing technique for detecting many toxic substances.

RECOMMENDATIONS

We recommend that the Secretary of the Interior and the Administrator of EPA discontinue the three national water quality networks established by the Geological Survey and EPA and devote their resources and attention to a program of special studies of water quality. We also recommend that the Administrator of EPA and the Chairman of CEQ promote the use of other available indicators of national progress toward cleaner water.

CHAPTER 5

SUMMARY OF AGENCY COMMENTS ON THE REPORT AND GAO RESPONSES

EPA, Interior's Geological Survey and Office of Water Research and Technology, and the Council on Environmental Quality provided us with written comments on a draft of this report. Because of the length and technical nature of many of their comments, we have included the full texts of all agency comments and our detailed responses in Volume II.

EPA, the Survey, and CEQ agreed with some of our findings but disagreed strongly with our recommendation that the national networks be terminated. The Department of the Interior and the Office of Water Research and Technology also disagreed with our recommendation.

In summary, the agencies believe that valid judgments of the condition and changing quality of the Nation's rivers can be made from limited (generally monthly) samples of water taken from less than 2,000 locations around the country. The agencies believe that statistical analysis will enable them to overcome the complexities of water quality. They also believe that it is not necessary to understand why water quality conditions exist or change at the sampling locations in order to make meaningful analyses and reports.

We are not persuaded by the agencies' arguments. We remain convinced that EPA and the Survey should discontinue the existing networks and devote their resources to a program of special studies of water quality.

We believe it is misleading to characterize a large river basin or segment from monthly samples at only one location. CEQ, the Survey, and EPA have done this and indicated in their responses to our draft report that they will continue the practice. We also believe it is misleading to use annual averages to describe the quality of a river. Water quality does not remain static or homogeneous over time; it changes often, sometimes dramatically in less than an hour. Yearly averages of monthly samples cannot present a representative picture of water quality.

The seven major areas of disagreement with our report are summarized below, as well as our evaluation. The agencies' detailed comments and our responses are in Volume II of this report.

OBJECTIVES OF THE NATIONAL NETWORKS

EPA and the Survey contended that we did not evaluate their networks against the objectives established for them, but rather against objectives more appropriate for detailed water quality studies. Consequently, the agencies believe that many of our criticisms of the national networks are not valid. EPA and the Survey also believe that their networks are useful for meaningful national assessments of the status and trends in water quality. The Survey stated that NASQAN purposes (e.g. determination of correlative relationships and probability distribution) cannot be accomplished with special studies.

We did not evaluate the networks against objectives other than those established for them. We clearly understood that the networks were established to provide data for reliable and meaningful assessments of water quality, and our review focused on the ability of the networks to meet these objectives. Our report demonstrates that fundamental weaknesses in the national networks, such as time and location biases, preclude the networks from providing proper data for such assessments.

EPA agreed that the networks are not generally designed to measure the chemical, physical, and biological integrity of the Nation's waters, nor are they particularly well designed to measure human exposure to toxic pollutants. EPA further stated that network data should generally not be used to describe individual sampling sites. We agree that network data should not be used to describe individual sampling sites. However, if the networks cannot produce reliable data to assess water quality at one site on a river, it follows that such data cannot represent conditions or trends in waters that have not been measured.

We believe that the Survey should carefully consider whether NASQAN itself can accomplish its objectives and purposes. With NASQAN, we believe that the Survey started a program that cannot succeed. The same weaknesses of NASQAN that preclude reliable, meaningful national assessments of water quality seriously weaken the other NASQAN purposes.

We believe that some of the objectives and purposes of NASQAN could be met by detailed special studies. Special studies would require many samples over a short period of time. The greater amount of data from homogeneous water quality conditions would clarify relationships in the data. The special studies could provide a better foundation for future river assessments than NASQAN. Well-performed studies also can lead to sounder management decisions for individual river basins, which is something the Survey and EPA agreed their networks are not expected to do.

LIMITATIONS OF STATISTICAL ANALYSIS

The Survey, CEQ, and EPA contend that statistical tools enable them to properly interpret data from the national networks. They believe that statistical analysis can overcome the complexities of water quality.

Water quality data can be better understood through statistical analysis -- if the data are properly obtained. However, the sampling plans used in the national networks are not adequate, and statistical analysis cannot overcome deficiencies in the data. Even if the agencies could perfect their sampling plans to develop reliable yearly averages for individual water quality characteristics, the results would be limited to the specific sampling sites. The agencies would have no legitimate basis in either statistical theory or water quality principles for extending these averages to other parts of the same river basin or to other river basins.

Statistical analyses are mathematical manipulations of data. The manipulations do not improve or change the data being analyzed. Water quality measurements identify and quantify foreign materials added to the water. But the foreign materials and their concentrations are ever-changing. Different water quality conditions are induced by fluctuations in streamflow, changes in the weather, unstable biological and physical systems of the river, changes in industrial, urban, and agricultural additions to the water, and many other influences. Measurements taken once or twice a month throughout a year under changing conditions are not homogeneous, but rather heterogeneous (taken from dissimilar conditions). The large ranges in measurements found for most individual characteristics being monitored and the skewed distributions of the measurements are obvious indications of the heterogeneous conditions. There are two ways of dealing with heterogeneous conditions:

--Redesign sampling frequency and timing at each sampling site to fit the uniqueness and variability of each water quality characteristic.

--Ignore the heterogeneous factor and assume for statistical purposes that all possible water quality conditions in every month and every year are from a homogeneous population.

CEQ, the Survey, and EPA, by using annual averages to depict water quality conditions or trends, have taken the second approach -- assuming homogeneity. In our opinion, this approach is not realistic because water conditions are not homogeneous.

If instead the Survey and EPA tried to redesign the network sampling plans, they would find the new plans overwhelmingly complex. This complexity would be the opposite of what the agencies tried to accomplish by establishing orderly, infrequent sampling at the network sites. The agencies would have to specify the precision needed for each average of water quality characteristic; they would have to adjust sampling frequency and timing for each characteristic at each sampling site. This exercise would have to be done separately for each sampling site because water quality is unique to each location and it would result in massive changes in sampling frequencies and work schedules. Some water quality characteristics could require hundreds of samples each year and the sampling patterns could be different at each site. These complexities were not acceptable to the Survey or EPA when they chose the simpler network sampling plans, and we understand why. The logistical and management problems of a complex nationwide sampling program could be overwhelming.

Even if the agencies chose to use sampling plans designed for individual locations, the results would be valid only for the specific sites that had been studied. The agencies' practice of classifying water quality of a whole region from measurements taken at a single station, is, in our opinion, wrong and misleading.

Water quality conditions upstream and downstream of sampling sites are often different; substantially so in many places. Until the zone of similarity around each network station has been clearly established, sampling results cannot be generalized from the station to waters not measured. However, establishing reliable zones of similarity may not be practical.

The reality of unstable, heterogeneous conditions must also be considered when using limited network water quality data for other purposes, such as the statistical extensions claimed by the Survey (e.g. probability distributions, correlative relationships, determination of spatial transferability). Since the extent of heterogeneity cannot be determined, the resulting statistical extensions are built on assumptions and are laden with uncertain, undefined biases.

AGGREGATING RESULTS OF SPECIAL STUDIES ON A NATIONAL SCALE

EPA and the Survey agreed that special studies are particularly well suited for investigating cause and effect relationships, as well as reasons for changes in water quality. The two agencies and CEQ also stated that a program of special studies should be encouraged.

However, all three agencies stated that our draft report did not explain how the results of special studies could be aggregated on a national scale. CEQ and the Survey, argued that the networks give a better national perspective than with special studies because the networks produce more frequent information about more rivers. CEQ claimed that it could not fulfill its statutory duty to report on environmental conditions and trends if the only data available came from special studies of different watersheds in different years.

We agree that our draft report did not provide sufficient information on how special studies could provide national perspective on water quality conditions and trends. New information has been added to the report on page 56, following an expanded discussion of the characteristics of good special studies. We believe our proposal is feasible and that the resulting data would be scientifically sound and very useful to policy and decisionmakers.

We believe that CEQ can fulfill its statutory duties with special studies. The National Environmental Policy Act of 1969 (see p. 13), does not require CEQ to use networks of fixed stations or to report every year on the quality of all rivers in the country. CEQ has often used many types of information other than network data to provide a revealing national perspective on efforts to improve water quality. For example, CEQ has described national conditions and trends with information on the increased use of waterways for swimming, fishing, and boating; the return of fish to previously polluted water; and changes in water color, odor, and foam.

We are not suggesting that CEQ should rely solely on special studies for its annual reports. In fact, we recognized the value of other indicators of progress and encouraged CEQ to increase its use of these indicators.

CEQ's annual environmental reports give the impression that network sampling provides national coverage, particularly through the maps that classify all areas in the Nation according to sampling results at NASQAN stations. (A Survey NASQAN map similar to those used by CEQ is in app. V.) This impression is misleading because vast areas are improperly represented by results at single sites and many rivers are not even sampled. In addition, as we have demonstrated in this report, even the water quality measurements and averages for the sampled sites are unreliable because of various time and location biases and other weaknesses.

We do not believe CEQ can fulfill its legislative mandate ("to gather timely and authoritative information concerning the conditions and trends in the quality of the environment") through continued use of the networks. We believe that the mandate could be satisfied by soundly conceived special water quality studies.

USE OF SCIENTIFIC AND TECHNICAL LITERATURE

EPA, the Survey, and the Office of Water Research and Technology questioned whether we had adequately considered the scientific and technical literature dealing with water quality networks. Although we cited specific material from only a few sources, we reviewed an extensive amount of literature. Many studies and articles we reviewed had been prepared for EPA and its predecessors or by scientists working for the Survey. Our positions stated in the draft of the report were reached based on many sources, including technical literature. We chose not to encumber our draft report with excessively technical material, but instead, relied on more common language. Because of the agencies' concerns, we have added a selected bibliography of scientific and technical literature we reviewed. (See app. I.)

APPLICABILITY OF NETWORK SAMPLING PROBLEMS TO SPECIAL STUDIES

CEQ, EPA, and the Survey stated that many of the network weaknesses we identified, particularly quality assurance problems, also apply to special water quality studies. We agree that poor planning and management can weaken special studies, but as we stated on page 54, special studies can circumvent the time and location biases that are inherent weaknesses of the networks. Unlike the networks, special studies circumvent location bias by providing information on large stretches of rivers rather than one site. Special studies can overcome time bias by frequent sampling throughout the day and during events that rapidly alter water quality (e.g. floods). Special studies can focus on specific problems, and those problems dictate the nature and scope of the studies, including sample timing and location.

We agree that quality assurance needs close attention in any type of sampling program. Quality assurance can receive closer attention in special studies than in network monitoring because special studies have unified management onsite, whereas the networks have fragmented, scattered management.

USEFULNESS OF NETWORK DATA

EPA, CEQ, and the Survey pointed out that network data are used for a variety of purposes. EPA said that data on 2, 4, 5-T/ Silvex from a subnetwork of NASQAN were submitted as evidence in cancellation hearings under the Federal Insecticide, Fungicide and Rodenticide Act. EPA also said it plans to meet some monitoring needs of the Toxic Substances Control Act through fixed station monitoring. The Survey stated that NASQAN data are indispensable for a variety of purposes, including policy analysis, water management, and hydrological research. CEQ uses network data for its annual reports on environmental quality.

We agree that network data are being used, but we have strong reservations about the propriety of using weak data for the purposes cited by the agencies. As we have discussed in detail in chapter 3, we believe that the questionable network data, easily available through computerized information systems, have led to misleading and incorrect interpretations of water quality conditions and trends.

EPA and the Survey conceded that the networks do have drawbacks. EPA agreed that networks are not adequate for monitoring human exposure to toxic pollutants and stated it is reevaluating its monitoring program. The Survey readily acknowledged that NASQAN was never intended to be a source of information detailed enough to assess the effectiveness of pollution control measures in specific locations. We believe the networks' inability to provide reliable information on water quality raises questions about the validity of using network data for policy analysis and water management.

COST OF SPECIAL STUDIES

EPA, the Survey, and CEQ expressed concern that we had not adequately considered the cost of a program of special studies. All of the agencies stated that special studies are costly and that a program of special studies might be prohibitively expensive.

We recognize that special studies can be costly, and in the discussion in chapter 4 we gave information on the range of costs associated with different kinds of special studies. We agree that 1,500 special studies each year would cost more than maintaining the approximate 1,500 stations in the national networks.

We believe that the agencies did not fully recognize our point. Millions are currently available from various sources for water quality monitoring activities; what is needed is a well-managed national program to make the most of these expenditures. We agree that if the program were overly ambitious, the funding currently available may not be sufficient, but we cannot agree that a program of special studies would necessarily be too costly.

SELECTED BIBLIOGRAPHY OF LITERATURE REVIEWED

Beckers, Charles V., and Chamberlain, Stanley G. Design of Cost-Effective Water Quality Surveillance Systems. EPA-600/5-74-004. Washington, D.C.: U.S. Environmental Protection Agency, Office of Research and Development, 1974.

Beckers, Charles V., Chamberlain, Stanley G., and Grimsrud, G. Paul. Quantitative Methods for Preliminary Design of Water Quality Surveillance Systems. EPA-R5-72-001. Washington, D.C.: U.S. Environmental Protection Agency, Office of Research and Monitoring, 1972.

Biology of Water Pollution: a Collection of Selected Papers on Stream Pollution Waste Water and Water Treatment. Washington, D.C.: U.S. Department of the Interior, Federal Water Pollution Control Administration, 1967.

Brown, William M., III, et al. A Synoptic Approach for Analyzing Erosion as a Guide to Land-Use Planning. Circular 715-L. Reston, Va.: U.S. Geological Survey, 1979.

Briggs, John C., and Ficke, John F. Quality of Rivers of the United States, 1975 Water Year--Based on the National Stream Quality Accounting Network (NASQAN). Open-File Report 78-200. Reston, Va.: U.S. Geological Survey, 1977.

Cali, J. Paul. "Toward a National Environmental Monitoring System: Measurement is a Key." Dimensions/NBS, Dec. 1977, pp. 6-9.

Energy Resources Company, Inc. Design of Stations for Monitoring Water Quality Trends, Vol. 1. Washington, D.C.: U.S. Environmental Protection Agency, Office of Planning and Evaluation, 1976.

Enviro Control, Inc. National Assessment of Trend in Water Quality. Final Report. Washington, D.C.,: Council on Environmental Quality, 1972.

Environmental Monitoring. Hearings Before the Subcommittee on the Environment and the Atmosphere, House Committee on Science and Technology, 95th Cong., 1st Sess., Sept. 13-15, 1977.

Hawkinson, Richard O., Ficke, John F., and Saindon, Linda G. Quality of Rivers of the United States, 1974 Water Year--Based on the National Stream Quality Accounting Network (NASQAN). Open-File Report 77-151. Reston, Va.: U.S. Geological Survey, 1977.

Hidore, John J. "Annual, Seasonal, and Monthly Changes in Runoff in the United States." Water Resources Bulletin, Vol. 7, No. 3, June 1971, pp. 554-561.

Hines, Walter G., et al. Dissolved-Oxygen Regimen of the Willamette River, Oregon, under Conditions of Basinwide Secondary Treatment. Circular 715-I. Reston, Va.: U.S. Geological Survey, 1977.

Hines, Walter G., et al. Formulation and Use of Practical Models for River-Quality Assessment. Circular 715-B. Reston, Va.: U.S. Geological Survey, 1975.

Hines, Walter G., Rickert, David A., and McKenzie, Stuart W. Hydrologic Analysis and River-Quality Data Programs. Circular 715-D. Reston, Va.: U.S. Geological Survey, 1976.

Kittrell, F. W. A Practical Guide to Water Quality Studies of Streams. CWR-5. Washington, D.C.: U.S. Department of the Interior, Federal Water Pollution Control Administration, 1969.

Kittrell, F. W., and Furfari, Santo A. "Observations of Coliform Bacteria in Streams." Journal Water Pollution Control Federation, Vol 35, No. 11, Nov. 1963, pp. 1361-1385.

Kittrell, F. W., and West, A. W. "Stream Survey Procedures." Journal Water Pollution Control Federation, Vol. 39, No. 4, April 1967, pp. 627-641.

Larson, Steven P., et al. Graphic and Analytical Methods for Assessment of Stream-Water Quality--Mississippi River in the Minneapolis-St. Paul Metropolitan Area, Minnesota. Water-Resources Investigations 76-94. Reston, Va.: U.S. Geological Survey, 1976.

Lettenmaier, Dennis P. "Design Considerations for Ambient Stream Quality Monitoring." Water Resources Bulletin, Vol. 14, No. 4, Aug. 1978, pp. 884-902.

Lettenmaier, Dennis P. Design of Monitoring Systems for Detection of Trends in Stream Quality. Technical Report No. 39. Seattle, Wa.: University of Washington, Charles W. Harris Hydraulics Laboratory, 1975.

Lettenmaier, Dennis P. Detection of Trends in Stream Quality: Monitoring Network Design and Data Analysis. Technical Report No. 51. Seattle, Wa.: University of Washington, Charles W. Harris Hydraulics Laboratory, 1975.

Lettenmaier, Dennis P. "Dimensionality Problems in Water Quality Network Design." Water Resources Research, Vol. 15, No. 6, Dec. 1979, pp. 1692-1700.

Lettenmaier, Dennis P., and Burges, Stephen J. "Design of Trend Monitoring Networks." Journal of the Environmental Engineering Division, ASCE, Vol. 103, No. EE5, Oct. 1977, pp. 785-802.

Lewis, Donald H. "Optimization of State Water Quality Monitoring Systems." Compututers & Operations Research, Vol. 3, 1976, pp. 127-143.

Loftis, Jim C., and Ward, Robert C. (Unpublished). Water Quality Monitoring--Some Practical Sampling Frequency Considerations. Fort Collins, Co.: Colorado State University, Department of Agricultural and Chemical Engineering.

National Commission on Water Quality. Report to the Congress. Washington, D.C.: National Commission on Water Quality, 1976.

National Research Council. Environmental Monitoring. Vol. IV of Analytical Studies for the U.S. Environmental Protection Agency. Washington, D.C.: National Academy of Sciences, 1977.

National Water Monitoring Panel. Model State Water Monitoring Program. EPA-440/9-74-002. Washington, D.C.: U.S. Environmental Protection Agency, 1975.

Nemerow, Nelson Leonard. Scientific Stream Pollution Analysis. Washington, D.C.: Scripta Book Company, 1974.

NUS Corporation, Cyrus Wm. Rice Division. Design of Water Quality Surveillance Systems, Phase I, Systems Analysis Framework. Washington, D.C.: U.S. Department of the Interior, Federal Water Quality Administration, 1970.

Principles for Evaluating Chemicals in the Environment. Washington, D.C.: National Academy of Sciences, 1975.

Rainwater, Frank H., and Avrett, James R. "Error Inference in Systematic-Sample Statistics in Stream Quality Studies." Journal American Water Works Association, Vol. 54, No. 6, June 1962, pp. 757-768.

Rice, Jeffrey M. and Anderson, Dale E. (Unpublished). Useful Monitoring for Policy and Decision-Makers. Paper presented at American Water Resources Association Symposium, Establishment of Water-Quality Monitoring Programs, June 12-14, 1978, San Francisco, Ca.

Rickert, David A., et al. Algal Conditions and the Potential for Future Algal Problems in the Willamette River, Oregon. Circular 715-G. Reston, Va.: U.S. Geological Survey, 1977.

Rickert, David A., et al. A Synoptic Survey of Trace Metals in Botton Sediments of the Williamette River, Oregon. Circular 715-F. Reston, Va.: U.S. Geological Survey, 1977.

Rickert, David A., and Hines, Walter G. A Practical Framework for River-Quality Assessment. Circular 715-A. Reston, Va: U.S. Geological Survey, 1975.

Rickert, David A., and Hines, Walter G. "River Quality Assessment: Implications of a Prototype Project." Science, Vol. 200, No. 4346, June 9, 1978, pp. 1113-1118.

Rickert, David A., Hines, Walter G., and McKenzie, Stuart W. Methodology for River-Quality Assessment with Application to the Willamette River Basin, Oregon. Circular 715-M. Reston, Va.: U.S. Geological Survey, 1976.

Rickert, David A., Hines, Walter G., and McKenzie, Stuart W. Project Development and Data Programs for Assessing the Quality of the Willamette River, Oregon. Circular 715-C. Reston, Va.: U.S. Geological Survey, 1976.

Sanders, Thomas G., and Adrian, D. Dean. "Sampling Frequency for River Quality Monitoring." Water Resources Research, Vol. 14, No. 4, Aug. 1978, pp. 569-576.

Sanders, Thomas G., and Ward, Robert C. (Unpublished). Relating Stream Standards to Regulatory Water Quality Monitoring Practices. Paper presented at American Water Resources Association Symposium, Establishment of Water Quality Monitoring Programs, June 12-14, 1978, San Francisco, Ca.

Sharp, W. E. "Stream Order as a Measure of Sample Source Uncertainty." Water Resources Research, Vol, 6, No. 3, June 1970, pp. 919-926.

Sharp, W. E. "A Topologically Optimum Water-Sampling Plan for Rivers and Streams." Water Resources Research, Vol. 7, No. 6, Dec. 1971, pp. 1641-1646.

Shelley, Philip E. Sampling of Water and Wastewater. EPA-600/4-77-039. Cincinnati, Ohio: U.S. Environmental Protection Agency, Environmental Research Information Center, 1977.

Steele, Timothy Doak. "Beneficial Uses and Pitfalls of Historical Water-Quality Data." National Symposium on Data and Instrumentation for Water Quality Management, Madison, Wis., July 1970, pp. 346-363. Springfield, Va.: National Technical Information Service, 1970.

Steele, Timothy Doak. "The Role of Network Design in the Management and Control of Streamflow Water Quality." Bilateral U.S.-Japan Seminar in Hydrology, Honolulu, January 1971, pp. 395-427. Fort Collins, Co.: Water Resources Publications, 1971.

Steele, Timothy Doak, Gilroy, Edward J., and Hawkinson, Richard O. An Assessment of Areal and Temporal Variations in Streamflow Quality Using Selected Data from the National Stream Quality Accounting Network. Open-File Report 74-217. Reston, Va.: U.S. Geological Survey, 1974.

U.S. Environmental Protection Agency, Standing Work Group on Water Monitoring. Basic Water Monitoring Program. EPA 440/9-76-025. Washington, D.C.: U.S. Environmental Protection Agency, 1977.

Velz, Clarence J. Applied Stream Sanitation. New York: Wiley-Interscience, 1970.

Ward, Robert C. Data Acquisition Systems in Water Quality Management. EPA-R5-73-014. Washingon, D.C.: U.S. Environmental Protection Agency, Office of Research and Monitoring, 1973.

Ward, Robert C., and Nielsen, Knud Strange. Evaluating the Sampling Frequencies of Water Quality Monitoring Networks. EPA-600/7-78-169. Las Vegas, Nevada: U.S. Environmental Protection Agency, Environmental Monitoring and Support Laboratory, 1978.

Ward, Robert C. Regulatory Water Quality Monitoring: A Systems Perspective. EPA-600/7-78-228. Las Vegas, Nev.: U.S. Environmental Protection Agency, Environmental Monitoring and Support Laboratory, 1978.

Ward, Robert C. "Regulatory Water Quality Monitoring: A Systems Perspective." Water Resources Bulletin, Vol. 15, No. 2, April 1979, pp. 369-380.

Wolman, M. Gordon. "The Nation's Rivers." Science, Vol. 174, No. 4112, Nov. 1971, pp. 905-918.

MINIMUM LIST OF WATER QUALITY CHARACTERISTICS MEASURED AT ENVIRONMENTAL PROTECTION AGENCY STATIONS IN THE NATIONAL WATER QUALITY SURVEILLANCE SYSTEM

Characteristics measured a/

Field determinations
- Streamflow
- Water temperature
- Conductivity
- pH
- Transparency or turbidity

Common constituents
- Bicarbonate, carbonate, residue (total filterable), residue (total nonfilterable), calcium (total), magnesium (total), sodium (total), potassium (total), chloride (total), sulfate (total)

Major nutrients
- Total phosphorus
- Total nitrite plus nitrate
- Total Kjeldahl nitrogen
- Total ammonia nitrogen

Trace elements
- Total iron

Organic and biological
- Total organic carbon

Oxygen-related measurements
- Dissolved oxygen
- Chemical oxygen demand

a/Frequency varies, with about half of the stations sampling bi-weekly and the rest sampling monthly.

Source: EPA National Water Quality Surveillance System parameter list, Mar. 9, 1977

LIST OF WATER QUALITY CHARACTERISTICS DESIGNATED BY THE ENVIRONMENTAL PROTECTION AGENCY FOR NETWORK STATIONS UNDER ITS BASIC WATER MONITORING PROGRAM

Characteristics Measured	Sampling Frequency: Rivers and streams	Lakes and impoundments, including the Great Lakes	Estuaries and bays
Flow	monthly	--	--
Temperature	monthly	seasonally	monthly
Dissolved oxygen	monthly	seasonally	monthly
pH	monthly	seasonally	monthly
Conductivity	monthly	seasonally	--
Fecal coliform	monthly	seasonally	monthly
Total Kjeldahl nitrogen	monthly	seasonally	monthly
Nitrate plus nitrite	monthly	seasonally	monthly
Total phosphorus	monthly	seasonally	monthly
Chemical oxygen demand	monthly	--	--
Total suspended solids	monthly	seasonally	monthly
Representative fish/ shellfish tissue analysis	annually	annually	annually
Transparency, secchi disc	--	seasonally	monthly
Total organic carbon	--	--	monthly
Salinity	--	--	monthly

Source: EPA document no. 440/9-76-025, Basic Water Monitoring Program, 1977.

INITIAL LIST OF WATER QUALITY CHARACTERISTICS MONITORED AT GEOLOGICAL SURVEY STATIONS INCLUDED IN THE NATIONAL STREAM QUALITY ACCOUNTING NETWORK (NASQAN)

Characteristics measured	Frequency
Field determinations	
Discharge (flow)	continuous
Water temperature	continuous, daily or monthly a/
Specific conductance	continuous, daily or monthly a/
pH	monthly
Fecal coliform bacteria	monthly
Fecal streptococcus bacteria	monthly
Common constituents (dissolved)	
Bicarbonate, carbonate, total hardness, non-carbonate hardness, calcium, magnesium, fluoride, sodium, potassium, dissolved solids, silica, turbidity, chloride, and sulfate	monthly or quarterly b/
Major nutrients	
Total phosphorus	monthly
Total nitrite plus nitrate	monthly
Total Kjeldahl nitrogen	monthly
Trace elements (total and dissolved)	
Arsenic, cadmium, chromium, cobalt, copper, iron, lead, manganese, mercury, selenium, and zinc	quarterly
Organic and biological	
Total organic carbon	quarterly
Total phytoplankton	monthly
Three co-dominants of phytoplankton	monthly
Periphyton, biomass dry and ash weight	quarterly
Periphyton, chlorophyll a and b	quarterly
Suspended sediment	monthly

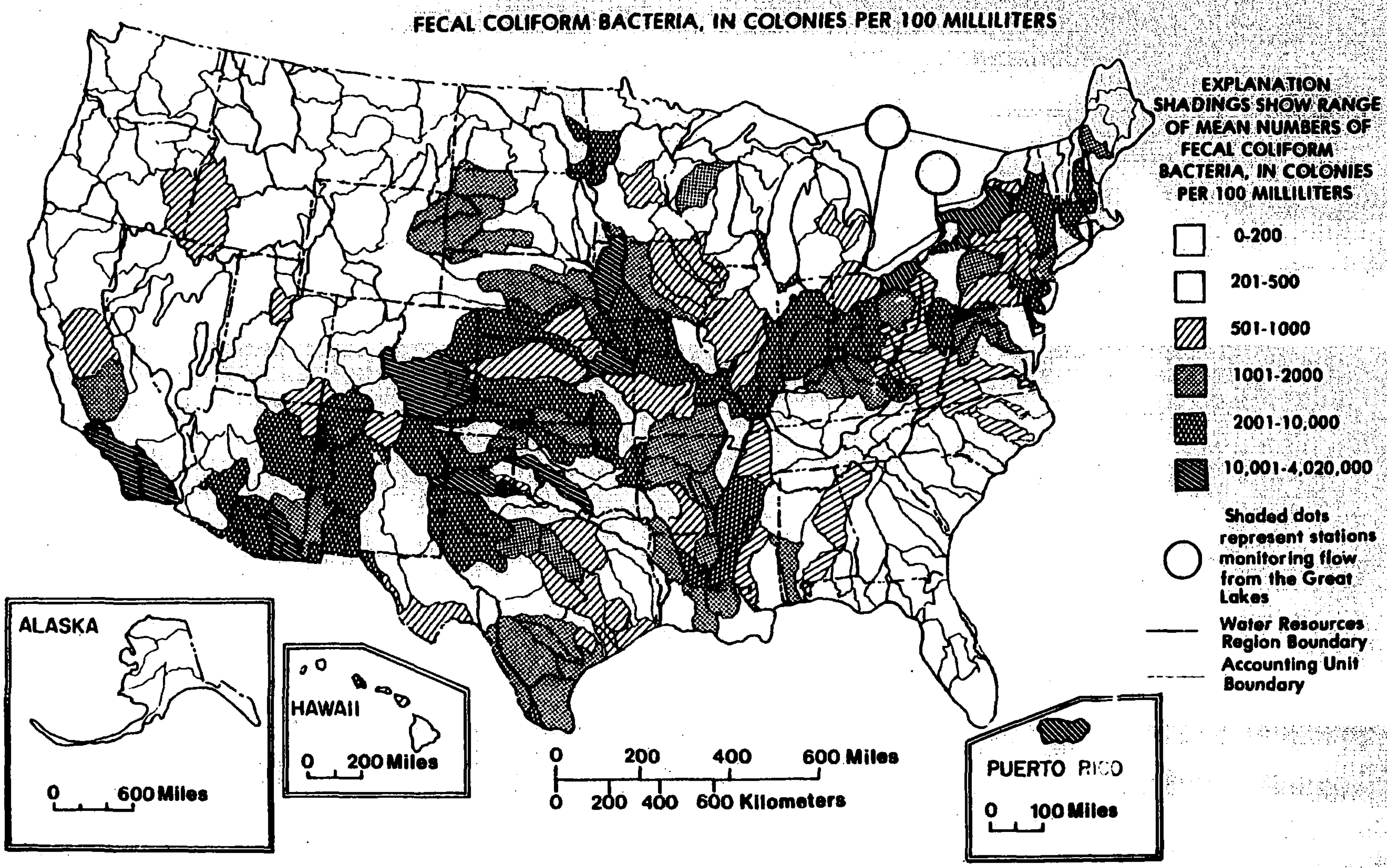
a/Continuous or daily depending on whether the station is equipped with a monitor or whether daily observations are made. Monthly measurements made at stations where a long-term record is available.

b/Quarterly or monthly, depending on whether relationships have been established between conductance and concentrations of various common constituents.

Source: U.S. Geological Survey Circular 719, The National Stream Quality Accounting Network (NASQAN).

APPENDIX VI

EXAMPLE OF A GEOLOGICAL SURVEY MAP SUMMARIZING THE RESULTS OF MONITORING AT NASQAN STATIONS--FOR FECAL COLIFORM BACTERIA MEASUREMENTS DURING OCTOBER 1974-SEPTEMBER 1975



Source: U.S. Geological Survey open file report 78-200 "Quality of Rivers of the United States, 1975 Water Year--Based on the National Stream Quality Accounting Network (NASQAN)."

EXAMPLE OF A GEOLOGICAL SURVEY TABLE SUMMARIZING THE RESULTS OF MONITORING AT A NASQAN STATION (POWDER RIVER NEAR LOCATE, MONTANA), FOR OCTOBER 1974–SEPTEMBER 1975

STATISTICAL SUMMARY OF SELECTED DISSOLVED CHEMICAL CONSTITUENTS AND REGRESSION RELATIONSHIPS OF CONSTITUENT CONCENTRATIONS TO SPECIFIC CONDUCTANCE

CONSTITUENT	CONSTITUENT (MG/L OR UNIT SHOWN) SAMPLE SIZE	MEAN	STANDARD DEVIATION	RANGE		REGRESSION SUMMARY* SAMPLE SIZE	REGRESSION COEFFICIENT, a	CONSTANT, b	CORRELATION COEFFICIENT	STANDARD ERROR OF ESTIMATE
TEMPERATURE, WATER (DEG C)	13	8.65	7.85	0.0	23.5					
SPECIFIC CONDUCTANCE (MICROMHOS)	12	1786.7	520.2	920	2370					
STREAMFLOW (CUBIC FT/SEC)	13	1687.5	2990.1	33	10800	12	-1.649	3874.917	-0.683	961.1
PH (STANDARD UNITS)	12	8.20	0.23	7.7	8.5	12	0.0001	7.9577	0.307 *	0.23
PHOSPHORUS, TOTAL	12	0.514	0.485	0.01	1.50	12	-0.00063	1.64164	-0.677	0.374
NITRITE + NITRATE, TOTAL	12	0.290	0.228	0.00	0.74	12	-0.00013	0.51434	-0.286 *	0.229
NITROGEN, KJELDAHL, TOTAL	12	1.542	1.615	0.44	5.80	12	-0.00173	4.63690	-0.558 *	1.405
PHYTOPLANKTON, TOTAL (CELLS/ML)	10	4406.0	4467.0	120	13600	10	-2.964	9475.105	-0.353 *	4433.6
SEDIMENT, SUSPENDED	12	5209.3	6763.6	31	19000					
SEDIMENT, CLAY-SILT (PERCENT)	7	70.9	28.5	18	95					
COLIFORM, FECAL (COL/100 ML)	12	194.6	370.0	3	1300					
STREPTOCOCCI, FECAL (COL/100 ML)	12	289.8	343.8	31	970					
SILICA, DISSOLVED	12	8.93	2.35	5.5	12.0	12	0.0010	7.1143	0.226 *	2.40
CALCIUM, DISSOLVED	12	110.50	28.81	64.0	150.0	12	0.0515	18.4907	0.930	11.13
MAGNESIUM, DISSOLVED	12	49.92	14.99	26.0	74.6	12	0.0266	2.3080	0.925	5.98
SODIUM, DISSOLVED	12	239.17	79.02	110.0	350.0	12	0.1357	-3.2788	0.893	37.27
POTASSIUM, DISSOLVED	12	7.09	1.48	4.3	8.7	12	0.0019	3.7237	0.665	1.16
BICARBONATE, ION	12	273.3	87.0	158	454	12	0.135	31.463	0.809	53.6
CARBONATE, ION	9	0.2	0.7	0	2	9	0.000	-0.493	0.350 *	0.7
SULFATE, DISSOLVED	12	609.17	195.75	320.0	970.0	12	0.3109	53.6589	0.626	115.66
CHLORIDE, DISSOLVED	12	107.50	51.49	31.0	220.0	12	0.0760	-28.3744	0.766	34.56
DISSOLVED SOLIDS, SUM OF CONST	12	1268.0	361.9	642	1760	12	0.672	67.417	0.915	161.3
DISSOLVED SOLIDS, RUE 180 DEG C	12	1350.8	411.1	668	1870	12	0.736	34.964	0.932	156.4
HARDNESS, TOTAL	12	481.7	132.5	270	680	12	0.238	56.727	0.934	49.8
HARDNESS, NONCARBONATE	12	258.3	77.8	140	370	12	0.125	34.938	0.836	44.8
TURBIDITY (JTU)	12	468.2	505.0	4	1200	12	-0.621	1577.597	-0.640	407.1
FLOURIDE, DISSOLVED	12	0.40	0.06	0.3	0.5	12	-0.0000	0.4330	-0.159 *	0.06

*Not significant at the 95 percent confidence level.

DURATION TABLE OF DAILY SPECIFIC CONDUCTANCE — SAMPLE SIZE = 362

DAILY SPECIFIC CONDUCTANCE IN MICROMHOS AT 25 DEG C, THAT WAS EQUALLED OR EXCEEDED FOR THE INDICATED PERCENTAGE OF TIME	1%	5%	10%	20%	30%	50%	70%	90%	95%	99%
	2870	2740	2630	2460	2280	1980	1550	898	782	708

SUMMARY OF HARMONIC ANALYSIS OF STREAM TEMPERATURE
FORM OF EQUATION: T*(D) = M + A * SIN(.0172 * D + C)

SAMPLE SIZE	HARMONIC MEAN - M (DEG C)	AMPLITUDE - A (DEG C)	PHASE ANGLE - C (RADIANS)	VARIATION EXPLAINED (%)	STANDARD ERROR OF ESTIMATE (DEG C)
362	6.13	11.17	2.89	88	2.88

SUMMARY OF MAXIMUM AND MINIMUM CONCENTRATIONS OF CONSTITUENTS SAMPLED AT A FREQUENCY OF QUARTERLY (1975 WY)

06326500 — POWDER RIVER NEAR LOCATE, MT.

CONSTITUENT	TOTAL NO. SAMPLES	TOTAL MINIMUM CONC.	TOTAL MAXIMUM CONC.	DISSOLVED NO. SAMPLES	DISSOLVED MINIMUM CONC.	DISSOLVED MAXIMUM CONC.
MINOR ELEMENTS:						
ARSENIC (AS), UG/L	4	2	60	4	0	2
CADMIUM (CD), UG/L	4	<10	10	4	0	1
CHROMIUM (CR), UG/L	4	0	100	4	0	10
COBALT (CO), UG/L	4	<50	150	4	0	1
COPPER (CU), UG/L	4	20	200	4	3	8
IRON (FE), UG/L	4	1600	170000	4	20	150
LEAD (PB), UG/L	4	<100	200	4	0	2
MANGANESE (MN), UG/L	4	30	3500	4	0	10
MERCURY (HG), UG/L	4	0.1	0.3	4	0.0	0.3
SELENIUM (SE), UG/L	4	1	5	4	1	3
ZINC (ZN), UG/L	4	60	670	4	20	40
PERIPHYTON:						
BIOMASS, DRY WT., G/SQ M	2	.79	8.39			
BIOMASS, ASH WT., G/SQ M	2	.50	6.19			
CHLOROPHYLL A, MG/SQ M	2	.0	.1			
CHLOROPHYLL B, MG/SQ M	2	.0	.0			
ORGANIC CARBON, MG/L	4	6.5	53.0			

Source: U.S. Geological Survey open-file report 78-200, Quality of Rivers in the United States, 1975 Water Year--Based on the National Stream Quality Accounting Network (NASQAN)

PROBLEMS IN INTERPRETING WATER QUALITY DATA:

A CASE STUDY, THE JAMES RIVER AT CARTERSVILLE, VIRGINIA

OCTOBER 18, 1979

REVISED TO RESPOND TO COMMENTS BY
THE U.S. GEOLOGICAL SURVEY AND THE
COUNCIL ON ENVIRONMENTAL QUALITY, FEBRUARY 28, 1980

Prepared for the

U.S. General Accounting Office

by

Jerome Horowitz
Consultant

Contents

T a b l e s

Tables (continued) Page

BACKGROUND AND PURPOSE

In December 1978 we asked the U.S. Geological Survey for examples that, in the Survey's judgment, best demonstrated the use of water quality data from fixed stations. The Survey cited many examples, including the seventh, eighth, and ninth annual reports of the Council on Environmental Quality (CEQ). These CEQ reports, according to the Survey, described conditions and trends in the quality of U.S. rivers, using data from the Survey's National Stream Quality Accounting Network (NASQAN).

We selected CEQ's ninth annual report for closer examination because CEQ, an agency in the Executive Office of the President, used network data to assess trends in water quality. In its ninth annual report, CEQ summarized water-quality changes in the following table.

"Water Quality Changes 1/
at NASQAN Stations 1975-77

Water quality characteristic	Percentage of Stations		
	Improved	No change	Deteriorated
Fecal coliform bacteria	7.3	88.9	3.8
Inorganic nitrogen	5.8	86.7	7.5
Organic nitrogen	3.8	83.4	12.8
Total phosphorus	4.3	83.2	12.5
Dissolved oxygen	4.5	92.9	2.6
Fecal streptococci bacteria	1.8	87.3	10.9
Dissolved solids	4.1	74.1	21.8
Dissolved zinc	9.1	86.5	4.4
Total zinc	12.8	85.8	1.4
Phytoplankton	2.0	93.6	4.4

1/Indications of change tested for statistical significance at the 90 percent level [sic]."

We selected the NASQAN station on the James River at Cartersville, Virginia, for detailed review because (1) CEQ concluded that there were many changes in water quality at that site, (2) the changes involved both improvements and deterioration in water quality, (3) it was exceptionally data-rich compared with normal network stations, and (4) the site served three Federal networks: NWQSS (EPA), NASQAN (Survey), and the State of Virginia's EPA-funded monitoring network.

Although CEQ did not analyze the State's data, this additional store of data made Cartersville an exceptionally good candidate for special analysis.

CEQ concluded from its statistical analysis of the Survey's Cartersville data that from October 1974 to September 1977 there was

--significant improvement for dissolved oxygen, organic nitrogen, dissolved zinc, and total zinc;

--significant deterioration for dissolved solids; and

--no significant change for fecal coliform bacteria, inorganic nitrogen, total phosphorus, fecal streptococci bacteria, and phytoplankton.

We concentrated on those water quality characteristics that, according to CEQ, had significantly changed during the 3-year period.

In conjunction with our consultant, we reviewed the Cartersville data, paying special attention to the Survey's data and to the conclusions CEQ had drawn from them. Our purpose was to cut through the standard statistical analyses to the data, and through careful examination to assess the credibility of the reported changes in water quality. The technical evaluation starting below was written by the GAO consultant, Jerome Horowitz.

SITE DESCRIPTION

Water quality samples are taken at Cartersville from State Highway 45 Bridge over the James River, about 45 miles upriver of Richmond. The river is fast-flowing and there are no large sources of wastewater nearby. The Survey alleges that water quality data for the James River at Cartersville are a valid record at the most downstream site above tidal influence where truly representative samples can be taken.

Several tributaries enter the James a few miles above Cartersville. The most important of them are the Rivanna and Willis Rivers and Byrd Creek.

Charlottesville is about 40 miles upriver (on the Rivanna). Lynchburg is about 100 miles upriver on the James proper.

The setting is rolling country, predominantly farms and forests. About 20 miles upriver at Bremo Bluff there is a large coal-fired powerplant. The State of Virginia has studied the thermal plume from this powerplant and has concluded that it does not reach Cartersville.

The river channel above Cartersville is bent and twisted; there are several rocky ledges in the channel, which create whitewater riffles. These features promote good mixing in the river.

TOTAL DISSOLVED SOLIDS: CEQ'S TREND AND THE DROUGHT

CEQ reported that total dissolved solids (TDS) significantly increased between October 1974 and September 1977. TDS is roughly synonymous with saltiness.

The Survey's TDS data unquestionably show that high values were more frequent during 1977 than 1975. But how is this apparent trend to be interpreted?

The riverflow records during 1975-77 clarify this apparent trend. Each year the river had successively lower average and minimum flows. These data are summarized in table 1. The Survey has recorded riverflows at Cartersville for about 80 years. During this long record, the average riverflow has been about 7,000 cubic feet per second (cfs).

Table 1. Yearly Average Riverflow, Yearly Minimum Riverflow, and Yearly Average TDS at Cartersville. Source: U.S. Geological Survey.

Date	Yearly Average riverflow (cfs)	Minimum daily riverflow (cfs)	Yearly Average TDS (mg/L)
October 1974-September 1975	9,076	1,840	82.6
October 1975-September 1976	6,271	865	85.6
October 1976-September 1977	5,500	555	109.5

Table 2 summarizes TDS data by month, along with monthly average and monthly minimum riverflows. All the TDS data are listed in table 2. Notice that sometimes TDS was measured only once a month; at other times it was measured several times a month. This table shows how TDS gradually rose as riverflows fell. The proportion of high TDS readings during the drought was much larger than usual.

Table 2. Total Dissolved Solids, Monthly Average Riverflow, and Monthly Minimum Riverflow at Carterville, October 1974 through September 1977. Source: U.S. Geological Survey.

Date	Monthly average riverflow	Monthly minimum riverflow	Total dissolved solids mg/L
	(cfs)	(cfs)	
1974			
October	2,261	1,840	98, 100, 62, 142
November	2,310	1,910	100, 90, 110, 100
December	7,171	2,760	100, 48, 55, 46
1975			
January	8,999	4,530	81, 40, 83
February	12,950	8,600	75, 69, 59
March	24,610	5,860	74, 49, 79, 72
April	8,456	5,360	126, 70, 69
May	12,160	6,520	86, 64, 99, 78
June	6,289	2,910	92, 87, 80, 103, 96
July	8,399	3,120	80, 189, 56, 52, 54
August	4,150	2,350	79, 85, 75, 68
September	11,270	2,490	82, 122, 89, 73, 79
October	7,355	3,870	62, 46, 122, 152
November	5,617	3,780	89, 81, 94, 112
December	4,953	3,200	82, 79, 76, 98, 76, 86
1976			
January	15,590	5,020	54, 58, 72, 72
February	8,769	5,720	62, 80
March	6,493	4,920	72, 54, 42
April	7,222	3,270	78, 68
May	4,595	2,630	76, 58
June	9,061	3,440	78, 82, 76, 69
July	2,863	1,870	124
August	1,550	1,040	184, 134
September	1,290	865	117, 113, 90
October	15,190	2,400	60, 91
November	6,033	3,480	70
December	8,441	4,470	62
1977			
January	3,579	2,670	86, 80, 179
February	3,994	2,670	176, 100
March	9,468	4,860	66, 66, 66
April	10,780	3,900	94, 81
May	3,152	2,200	92, 117
June	1,941	1,560	109, 122
July	1,118	750	148, 148
August	1,034	555	139
September	1,059	700	185, 181

The information in tables 1 and 2 shows that CEQ's "statistically significant" trend is nothing more than a normal consequence of the deepening drought. As the drought deepened there was less surface runoff to dilute the groundwater and the James River became progressively saltier. GAO presented this case study to the Survey for comment. In a letter of November 26, 1979, the water quality specialist at the Survey's northeastern region, stated:

> "The U.S. Geological Survey data for the period 1969-79 showed a statistically significant decreasing trend in dissolved solids concentration in the James River at Cartersville, Virginia." (Underscoring supplied.)

Please notice that the Survey's 10-year trend was the opposite of CEQ's 3-year trend.

TOTAL DISSOLVED SOLIDS AND SPECIFIC CONDUCTANCE: UNEXPLAINED INCONSISTENCIES

Specific conductance and TDS are closely related properties of water. In general, conductance is higher than TDS -- usually about 30 percent higher. TDS and specific conductance are among the simplest of all water quality measurements.

The Survey's Cartersville data sometimes show stable values for TDS and conductance and a stable ratio of TDS to conductance. Table 3 gives a few examples of data that seem especially credible. Notice that duplicate samples gave consistent results, that conductance was always higher than TDS, and that the ratio of TDS to conductance ran about 70 percent over a broad range of seasons and riverflows.

Table 3. Examples of the Survey's Cartersville Data, Showing Especially Stable Relationships Between Total Dissolved Solids and Specific Conductance and Consistent Results in Duplicate Samples.

Date (year/month/ day/(hour))		River-flow (cfs)	TDS (mg/L)	Specific conductance (micromhos)	TDS/ Sp.C (%)
1975/Sept/22	(10:30)	4,720	73	115	63
DUPLICATE	(11:00)		79	115	69
1976/June/1	(09:00)	16,100	78	110	71
DUPLICATE	(10:00)		82	110	75
1977/Jan/10	(09:30)	3,760	86	120	72
DUPLICATE	(09:45)		80	120	67
1977/March/7	(08:30)	6,100	66	100	66
DUPLICATE	(09:00)		66	100	66

The Survey's Cartersville data are not always of this quality. Table 4 gives examples of TDS and specific conductance that are difficult to make sense of.

Table 4. Examples of the Survey's Cartersville Data on Total Dissolved Solids and Specific Conductance That Are Difficult to Make Sense of.

Date (year/month/ day/(hour))		River-flow (cfs)	TDS (mg/L)	Specific conductance (micromhos)	TDS/ Sp.C (%)
1974/Nov/1		2,260	100	100	100
1974/Nov/15		2,530	110	110	100
1975/Jan/15		10,700	83	75	111
1975/Mar/10	(08:00)	5,740	79	110	72
DUPLICATE	(08:30)		49	110	45
1975/Mar/15		52,100	72	75	96
1975/Apr/1		16,300	126	85	148
1975/Jul/1	(08:30)	4,180	189	105	180
DUPLICATE	(21:00)		80	150	53
1975/Aug/1		3,490	79	150	53
1975/Aug/11		4,020	85	240	35
1975/Aug/15		3,300	75	125	60
1975/Oct/6		3,760	46	104	44
1975/Oct/20		30,570	152	155	98
1975/Dec/15	(09:00)	5,100	76	125	61
DUPLICATE	(09:30)		76	125	61
DUPLICATE	(noon)		98	125	78
1976/Jan/12		7,960	58	43	135
1976/Mar/8	(08:30)	5,400	54	103	52
DUPLICATE	(09:00)		72	103	70
1976/Mar/22		7,410	42	110	38
1976/Sep/20	(08:30)	1,340	90	190	47
DUPLICATE	(09:00)		113	190	59
1976/Oct/26		33,000	91	75	121
1977/Jan/25		18,700	179	140	128
1977/Feb/7		3,160	176	125	141
1977/Jul/19		945	148	260	57
1977/Aug/9		612	139	330	42
1977/Sep/6		1,080	185	330	56

There are many oddities in table 4. The ratio of TDS to conductance was extremely unstable. 1/ During the 3 years summarized in this table, the ratio varied from 35 percent to 180 percent. Duplicate samples did not always give consistent results. For example, on March 10, 1975, the duplicate sample gave consistent results for conductance but very inconsistent results for TDS (a difference of 38 percent within 30 minutes). On July 1, 1975, both TDS and conductance were inconsisent in the duplicates. Perhaps the Survey's data record has typographical errors in this pair of readings; it is difficult to believe that the river really changed so much during 1 day. TDS was greater than conductance on many occasions; surely there are errors here.

In short, these data are suspect. It is not prudent to derive sweeping conclusions about river quality from suspect data. It is not hard to identify suspect data from tables where the values are clearly impossible (e.g. TDS greater than conductance). But what assurance is there that reasonable-looking data are not inaccurate too?

RIVERFLOW PATTERNS MAY EXPLAIN CHANGES IN TOTAL ZINC

Water quality is influenced by water quantity--by hydrology. A classic example is that rivers are muddiest when riverflows are rapidly rising. Unless hydrologic influences are accounted for, analysts may draw erroneous conclusions from water quality data.

The following discussion often refers to "point sources" and "nonpoint sources" of pollution. These terms are jargon. Broadly speaking, point sources refer to distinct and discrete sources of pollution--the pollutants that come out of pipes (e.g. sewers, sewage plants, and industrial discharges). Nonpoint sources refer to loosely defined areas (e.g. runoff from forests and farmlands or seepage through geological strata).

CEQ analyzed total zinc records at Cartersville between September 1974 and October 1977 and concluded that total zinc values had significantly dropped. The following tables include data from the Survey and from the Virginia Water Control Board. CEQ did not analyze the State's data.

1/The Chief, of the Survey's Quality of Water Branch wrote us on Nov. 21, 1979, agreeing that there was "an apparent problem with specific conductance data, especially in the 1975 water year and to a lesser extent in 1976 * * *. Since 1976, the problem has apparently been resolved and the ratio of dissolved solids to specific conductance has generally been within the expected range."

The chronological display (table 5) shows that concentrations of total zinc were usually less than 40 ug/L. Most of the high values appeared between June 1974 and September 1975 but never came near violating any water quality standard. Calculated values of total daily load (pounds per day) were usually less than 1,000 pounds. Most of the high zinc loads came between June 1974 and November 1975.

Table 5. Total Zinc and Streamflow Arranged Chronologically.

Date		Total zinc concentration (micrograms per liter)	Total zinc load (thousands of lbs/day)	Instantaneous streamflow (cfs)
1974				
Mar.	12	0	0.00	6,560
June	25	40	0.77	3,564
Sept.	23	20	0.11	1,020
Dec.	23	a/70	b/2.26	5,970
1975				
Mar.	c/3	10	0.43	d/8,010
Mar.	10	a/40	b/1.24	5,740
June	2	70	5.91	15,800
June	c/23	e/10	0.16	d/2,940
Sept.	22	90	2.29	4,720
Nov.	c/14	20	1.40	d/13,000
Dec.	15	20	0.55	5,105
1976				
Mar.	8	10	0.29	5,400
May	c/14	e/10	0.14	d/2,630
June	1	30	2.60	16,100
Sept.	20	10	0.07	1,340
1977				
Jan.	10	0	0.00	3,810
Mar.	7	40	1.35	6,280
June	c/2	20	0.29	d/2,710
June	15	10	0.09	1,660
Sept.	6	0	0.00	1,060

a/This value is dissolved zinc, a part of the total zinc value. Total zinc was not tested on these days.

b/Computed from the dissolved zinc concentration. See note a.

c/Data from Virginia State Water Control Board. All other data from the Survey.

d/Instantaneous flow not available. Used mean daily flow from the Survey.

e/Total zinc could be less than 10, which is the detection limit of the instrument used on this date.

Table 6 clarifies some hydrological influences on these zinc data. Note that all the high zinc loadings were at riverflows greater than 4,500 cfs. At low riverflows, both concentrations and loads were quite low. This pattern suggests that total zinc did not behave like a point source of pollution and that suspended zinc usually accounted for most of the total zinc.

Table 6. Total Zinc Ranked by Streamflow.

Instantaneous streamflow (cfs)	Date	Total zinc concentration (micrograms/ liter)	Total zinc load (thousands of lbs/day)
16,100	June 1, 1976	30	2.60
15,800	June 2, 1975	70	5.91
a/13,000	b/Nov. 14, 1975	20	1.40
a/ 8,010	b/Mar. 3, 1975	10	0.43
6,560	Mar. 12, 1974	0	0.00
6,280	Mar. 7, 1977	40	1.35
5,970	Dec. 23, 1974	c/70	d/2.26
5,740	Mar. 10, 1975	c/40	d/1.24
5,400	Mar. 8, 1976	10	0.29
5,105	Dec. 15, 1975	20	0.55
4,720	Sept. 22, 1975	90	2.29
3,810	Jan. 10, 1977	0	0.00
3,564	June 25, 1974	40	0.77
a/ 2,940	b/June 23, 1975	e/10	0.16
a/ 2,710	b/June 2, 1977	20	0.29
a/ 2,630	b/May 14, 1976	e/10	0.14
1,660	June 15, 1977	10	0.09
1,340	Sept. 20, 1976	10	0.07
1,060	Sept. 6, 1977	0	0.00
1,020	Sept. 23, 1974	20	0.11

a/Instantaneous flow not available. Used mean daily flow from the Survey.

b/Data from Virginia State Water Control Board. All other data from the Survey.

c/This value is dissolved zinc, a part of the total zinc value. Total zinc was not tested on these days.

d/Computed from the dissolved zinc concentration. See note c.

e/Total zinc could be less than 10, which is the detection limit of the instrument used on this date.

The daily flow data suggest that total zinc behaved like a nonpoint source. Daily loads were generally high when riverflows were high and unstable, and were generally low when riverflows were low and stable.

What CEQ took to be an improving trend in water quality might be nothing more than a change in sampling riverflow patterns. Most of the exceptionally high concentrations and loads coincided with dramatic changes in riverflow. These dramatic changes were especially common between December 1974 and June 1976. As luck would have it, the Survey rarely analyzed for total zinc during dramatic changes in riverflow after about September 1975. For example, in October 1976 the streamflow suddenly jumped from about 2,500 cfs to 70,000 cfs; but the Survey took no zinc samples in October 1976. Whenever the Survey analyzed for total zinc at or near a big peak in the riverflow record, it found high values; but the Survey happened to miss most of the peaks after September 1975. The State analyzed for total zinc only once at or near a peak in the riverflow record. This trick of chance might explain the differences between the Survey's data and the State's data. But the explanation might be found elsewhere--errors in the laboratory or the computer center.

Table 5 shows that the State never reported any values greater than 20 ug/L, whereas the Survey reported several values of 70 and higher. Had CEQ used the State's data (rather than the Survey's data), it would have come to entirely different conclusions about zinc trends.

RECORDS ON DISSOLVED ZINC RESIST RATIONAL ANALYSIS

The Survey may measure dissolved zinc because it is a fish poison. Its toxicity varies with fish species, temperature and concentration of dissolved oxygen.

The chronological record (table 7) shows that dissolved zinc values were usually very low. Four exceptional values were recorded between June 1974 and June 1975. All other values were at or near the limit of detection.

Table 7. Dissolved Zinc and Streamflow Arranged Chronologically. Source: U.S. Geological Survey.

Date	Concentration (micrograms/ liter)	Load (thousands of lbs/day)	Instantaneous streamflow (cfs)
1974			
March 12	0	0.00	6,560
June 25	30	0.58	3,564
Sept. 23	10	0.05	1,020
Dec. 23	70	2.26	5,970
1975			
March 10	40	1.24	5,740
June 2	30	2.55	15,800
Sept. 22	0	0.00	4,720
Dec. 15	0	0.00	5,105
1976			
March 8	0	0.00	5,400
June 1	0	0.00	16,100
Sept. 20	10	0.07	1,340
1977			
Jan. 10	10	0.21	3,810
March 7	10	0.34	6,280
June 15	0	0.00	1,660
Sept. 6	0	0.00	1,060

This chronological record seems to show that concentrations of dissolved zinc were falling; but when the data are organized according to riverflow a different story emerges.

The dissolved zinc must come from point sources or nonpoint sources of pollution. If it comes from point sources, it should be highest when riverflows are lowest because there is less water in the river to dilute the point source discharges. Table 8 shows nothing of the kind.

Table 8. Dissolved Zinc Ranked by Streamflow.

Instantaneous streamflow (cfs)	Date	Dissolved zinc concentration (micrograms/ liter)	Dissolved zinc load (thousands of lbs/day)
16,100	June 1, 1976	0	0.00
15,800	June 2, 1975	30	2.55
6,560	Mar. 12, 1974	0	0.00
6,280	Mar. 7, 1977	10	0.34
5,970	Dec. 23, 1974	70	2.26
5,740	Mar. 10, 1975	40	1.24
5,400	Mar. 8, 1976	0	0.00
5,105	Dec. 15, 1975	0	0.00
4,720	Sept. 22, 1975	0	0.00
3,810	Jan. 10, 1977	10	0.21
3,564	June 25, 1974	30	0.58
1,660	June 15, 1977	0	0.00
1,340	Sept. 20, 1976	10	0.07
1,060	Sept. 6, 1977	0	0.00
1,020	Sept. 23, 1974	10	0.05

The table shows that both concentrations and daily loads (pounds per day) of dissolved zinc were often lowest at low riverflows. Clearly, these dissolved zinc values cannot be explained by the point source hypothesis.

Neither can the record be explained by the nonpoint source hypothesis. If the dissolved zinc were coming from nonpoint sources, daily loads should be highest when riverflows are highest. Riverflows are high when rain or melting snow washes the drainage area and collects in the river. Table 8 shows that at the highest riverflow (June 1, 1976) the daily load of dissolved zinc was zero. It was also zero at the third highest riverflow. Clearly, this data record is not consistent with the nonpoint source hypothesis.

The only clear sign of a dissolved zinc problem is the inexplicably high values between June 1974 and June 1975. We questioned the Survey's laboratory closely about changes in procedures that might explain these high values. The laboratory staff assured us that their analyses during June 1974-June 1975 were entirely credible.

If the laboratory is correct, the data record cannot be rationally explained by any hypothesis of pollution control. In short, the data record defies rational analysis. No supportable conclusions can be drawn from it.

The Chief of the Survey's Quality of Water Branch wrote us on November 21, 1979, as follows:

"The report * * * concludes that when considered with respect to riverflow the data 'resists rational analysis.' This is a puzzling conclusion since the chronological analysis suggests that dissolved zinc appeared to be a minor problem at Cartersville; there could possibly have been a more notable problem somewhere upstream for the period June 1974 to June 1975. This was the only period between March 1974 to September 1978 [sic] that dissolved zinc was found in concentrations higher than 10 ug/L regardless of flow.

"Taking the period June 1974 to June 1975 and now using the author's non-point source hypothesis that daily loads should be highest when riverflows are highest we see good evidence (see table below) that there was a discharge of zinc upstream of Cartersville around the period June 1974 to June 1975 which has subsequently left the system. There are no samples of record after June 1975 that show concentrations higher than 10 ug/L, regardless of flow.

"Date	Instantaneous Discharge (cfs)	Dissolved Zinc Concentrations (ug/L)	Dissolved Zinc Load (1000 lbs/day)
June 2, 1975	15,800	30	2.55
Dec. 23, 1974	5,970	70	2.26
March 10, 1975	5,740	40	1.24
June 25, 1974	3,564	30	0.58
Sept. 23, 1974	1,020	10	0.06"

The Survey hypothesizes that a suddenly vanishing nonpoint source rationally explains the dissolved zinc data. Notice that the flowing loads of dissolved zinc exceeded 2,500 pounds a day at high flow but were about 60 pounds a day during low flow. What happened to the nonpoint source of 2,500 pounds a day? This unspecified source, according to the Survey, was active only from June 1974 to June 1975. There was no sign of dissolved zinc at 6,560 cfs on March 12, 1974. It has since disappeared without a trace. The Survey did not explain what this unidentified area source might have been, or even where it might have been. Since it was a nonpoint source, it could not have been a factory or a sewage plant. It must have been an area permeated with dissolved zinc. The Survey did not think that its zinc data could have been wrong. Instead, we are asked to believe that an area above Cartersville had been permeated with zinc; we must further believe that this area suddenly vanished. We again conclude that the data resist rational analysis.

TEMPERATURE INSTABILITY AT CARTERSVILLE

We find that the water temperature at Cartersville is unstable (see table 9). The instability is important because temperature may affect fish in two ways: (1) sudden changes in temperature may put some fish into thermal shock and (2) temperature affects the capacity of water to hold dissolved oxygen--warm water cannot hold as much oxygen as cold water can, and fish need oxygen to breathe.

Table 9. Temperature Records at Cartersville: Data Clusters, February 1975-June 1977. Source: U.S. Geological Survey and Virginia's State Water Control Board.

				Temperatures recorded by:		
Year	Month	Day	Time	Survey technician (°C)	State employee (°C)	Survey daily observer (°C)
1975	Feb.	10	09:30	3.0		
	Feb.	10				3.5
	Feb.	11	14:20		6.7	
	Feb.	11				5.0
1975	May	1	08:30	14.5		
	May	1				14.5
	May	2	14:50		17.8	
	May	2				15.5
1975	Aug.	25	08:00	26.0		
	Aug.	25				25.0
	Aug.	26	11:15		28.9	
	Aug.	26				27.0
1975	Dec.	29	14:15		6.1	
	Dec.	29				2.5
	Dec.	30	08:30	5.0		
	Dec.	30				2.5
1977	Feb.	22	08:00	3.5		
	Feb.	22	14:00		8.0	
1977	June	1	12:00	23.0		
	June	2	13:50		2.8 [sic]	

A large coal-fired powerplant is about 20 miles upstream from Cartersville at Bremo Bluff; the State has studied the thermal plume from the powerplant and has concluded that it does not extend to Cartersville. However, several important tributaries flow into the James River just above Cartersville. These unmeasured tributaries need not have the same temperature as the James, and they may affect the temperature at Cartersville.

To make sense of the apparent temperature instability, temperatures must be read round the clock in the James River itself and in the major tributaries just above Cartersville. This work has not been done and there are no plans for doing it. It will take more than a few readings a month to define the temperature regime at Cartersville.

INCONSISTENCIES IN DISSOLVED OXYGEN

The records for DO show that undefined phenomena dramatically change the oxygen supply of the James River. The DO data cannot be interpreted without understanding these dramatic changes. Our analysis identifies common obstacles to interpreting data from fixed stations inflexibly sampled a few times a month. Two or three DO tests a month cannot explain the dramatic fluctuations throughout the DO record.

DO is important because fish breathe the oxygen dissolved in the water. Many factors affect DO; here are four of the most important ones:

--<u>Physical reaeration</u> (the entrainment of oxygen into the water, e.g. at rapids and waterfalls). Physical reaeration increases with turbulence and velocity. Water churning down a river channel is more rapidly reaerated than sluggish water.

--<u>Biological reaeration</u> (oxygen from aquatic plants). Green plants (especially algae) are net producers of oxygen in sunlight and net consumers in the dark. In waters full of algae, DO rises during the day and falls at night; DO is usually highest in the afternoon and lowest just before dawn.

--<u>Temperature</u>. The amount of oxygen that water can hold at a given temperature is a physical constant called the saturation value. For example, at 20°C the saturation value near sea level is 9.2 mg/L of DO. The amount of DO actually in the water can be compared with the saturation value by a simple percentage calculation. For example, if a water sample at 20°C holds only 6.9 mg/L of DO, it is said to be

75 percent saturated (6.9 divided by 9.2 = 75%). Cold water has a higher saturation value than warm water--cold water can hold more DO than warm water can.

--Decomposing wastes. As waterborne wastes decompose they consume oxygen. The rate at which they consume oxygen depends on the nature of the wastes, the physical geometry of the river channel, and the kinds of bacteria living in the water. Two common measurements of decomposing wastes are BOD-5 and ammoniacal nitrogen.

To make sense of DO data, all these factors must be understood.

CEQ concluded that DO significantly improved between 1975 and 1977. But riverflow fell during this period. How could DO improve during a drought? Two of the most important factors that might explain the improved DO are temperature and algae.

Discrepancies related to DO and temperature

At fairly constant riverflows (when reaeration may be assumed fairly constant), DO should rise as the water temperature falls. If it does not, we must seek the explanation in biological reaeration, decomposing wastes, or errors of measurement and reporting. Between October 1974 and September 1977 (the period of record we reviewed), DO values often failed to change as they should have according to temperature influences alone. Table 10 illustrates one set of discrepant readings.

Table 10. DO, Temperature, Percent Saturation, and Riverflow at Cartersville, 28 October - 11 November 1974.

Time	Date	Agency	Riverflow (cfs)	DO (mg/L)	Water temp. (_C)	Percent saturation	Phytoplankton Chlorophyll a (ug/L)	Phytoplankton Chlorophyll b (ug/L)
13:30	28 Oct. 74	State	1,990	10.4	16.7	107	_	
13:35	5 Nov. 74	State	1,960	9.0	17.8	94.7	_	
10:30	11 Nov. 74	Survey	1,950	7.6	10.5	67.3	14	19

Table 10 shows that DO fell as the temperature fell, even though the riverflow was virtually constant.

Theoretically, DO should have increased as the temperature fell. This theoretical difference is clearest in the "percent saturation" column; the difference between the first and third readings is nearly 40 percent. What can account for the river's having lost nearly 40 percent of its oxygen supply? On 11 November the Survey reported plenty of chlorophyll (33 ug/L) and fairly high amounts of decomposing wastes (BOD-5 of 3.5 mg/L and ammoniacal nitrogen of 90 ug/L). To interpret these discrepancies in the DO record we would need a battery of measurements for several days on chlorophyll, biological oxygen demand (BOD), and ammonia along the river and its tributaries above Cartersville. To rule out biological reaeration, we would need DO readings several times a day, including midafternoon and just before dawn. It is impossible to interpret these discrepancies from the measurements on record.

This example is not isolated. As riverflow rises and temperature falls, DO and the percent saturation should rise. Table 11 shows a counterexample.

Table 11. DO, Temperature, Percent Saturation, and Riverflow at Cartersville, 11 November - 9 December 1974.

Time	Date	Agency	Riverflow (cfs)	(mg/L) BOD-5	DO (mg/L)	Water temp. (C)	Percent saturation	Phytoplankton Chlorophyll a (ug/L)	Phytoplankton Chlorophyll b (ug/L)
10:30	11 Nov. 74	Survey	1,950	3.5	7.6	10.5	67.3	14	13
10:00	25 Nov. 74	Survey	2,450	2.6	8.6	16.0	86.0	0/12*	0/16*
09:30	9 Dec. 74	Survey	14,300	1.7	7.2	5.0	56.2	0	0

*Duplicate samples that did not give consistent results.

Between 11 and 25 November the water temperature rose and the riverflow was stable. Theoretically, DO should have decreased; but DO rose, and so did the percent saturation. By 9 December the riverflow had increased by more than 500 percent and the water temperature had plummeted. Theoretically, both DO and percent saturation should have been much higher; instead, both were much lower. On 9 December the Survey tested the river for chlorophyll and found none at all; this finding rules out biological reaeration. Tests for decomposing wastes revealed rather little (BOD-5 of 1.7 mg/L and ammoniacal nitrogen of 70 ug/L). Without full details on biological reaeration (which is not likely to have been significant at such low temperatures so late in the year) and on decomposing wastes (especially on wastes that may have been scoured up from the riverbed as the river velocity increased), the DO data cannot be interpreted.

In a letter dated November 21, 1979, the Survey's Chief, Quality of Water Branch, disagreed with the analysis of table 11:

> "Weather information and sediment concentrations are available for assisting in interpretation of the DO data. Conditions for the two November samples were clear weather, low sediment concentrations and low flow. The temperature of the sample taken on November 25 was higher and temperature-dependent photosynthesis activity could account for the higher DO value. On December 9, the flow was high from a period of storms, the weather was cloudy, and sediment concentration had increased to 131 mg/L. Lower temperature (compared to November), higher water stage with increased sediment concentration covering periphytic algae and cloudy skies reducing available sunlight would deminish [sic] photosynthetic activity and reduce the amount of oxygen added to the water. Also, runoff following storms generally has a higher BOD load than the water entering the stream during a low flow period (usually ground water). This oxygen demand reduces the amount of dissolved oxygen in the stream."

The Survey offers two explanations of the DO anomaly: (1) algal photosynthesis and (2) high BOD. Neither explanation is satisfactory; both are contradicted by the Survey's own data.

Photosynthesis by attached algae (periphyton) is speculation: the Survey has no data on periphyton or periphytonic chlorophyll at Cartersville in late 1974. The duplicate sample on November 25, 1974, shows that the Survey's analyses for phytoplanktonic chlorophyll did not give consistent results. The first analysis yielded zero for chlorophyll _a_ and _b_; the second analysis yielded 12 and 16 ug/L, respectively. These results demonstrate that the Survey could not reliably measure chlorophyll. The Survey cannot assert that chlorophyll was or was not present, or that algal photosynthesis did or did not affect the DO regime at Cartersville. Unreliable and incomplete data lead to paradox and inconclusiveness, not to firm answers.

BOD concentrations--not BOD loads--explain DO concentrations. The BOD concentrations steadily dropped between 11 November and 9 December 1975, according to the Survey's own data.

The data do not support firm conclusions on algal photosynthesis because too many phenomena were neglected (e.g. periphyton) and because the Survey's chlorophyll measurements were unreliable. All the other data (streamflow, temperature, and oxygen demand) indicate that DO should have increased on 9 December--the water was colder, swifter, and contained less BOD. It does no good to invoke algal photosynthesis when the data cannot explain important changes in the river.

We have illustrated problems in the DO record for (1) constant riverflow with falling temperature and (2) rising riverflow with falling temperature. In the next example we show DO problems with fluctuating riverflows and temperatures.

Table 12. DO, Temperature, Percent Saturation, and Riverflow at Cartersville, 3-17 May 1976.

Date	Time	Agency	River-flow (cfs)	DO (mg/L)	Water temperature (°C)	Percent saturation
3 May 76	08:30	Survey	5,960	8.2	17.0	84.5
14 May 76	13:15	State	2,630	9.3	21.1	103
17 May 76	10:30	Survey	7,130	6.9	21.0	76.7

Table 12 illustrates several inconsistencies. On 14 May the riverflow was lower and the temperature was higher than on 3 May. Theoretically, DO should have been lower: warmer water has a lower saturation value, and sluggish water has less physical reaeration. In fact, the opposite occurred: DO increased and the percent saturation jumped by nearly 20 percent.

On 17 May the riverflow was much higher and the temperature was virtually unchanged from 14 May. Theoretically, DO and percent saturation should have increased; instead, both fell. Why should the percent saturation have fallen by over 25 percent? Physical reaeration cannot be the reason. If anything, physical reaeration must have been much stronger on 17 May than on 14 May, since the riverflow had nearly tripled. The State never tests the water at Cartersville for chlorphyll or green plants, so we have no way of assessing biological reaeration. However, we can compare biological reaeration on 3 and 17 May 1976 because the Survey sometimes tests for both chlorophyll and algae.

Discrepancies related to biological reaeration

Although there are no data on algae or chlorophyll for 14 May 1976, we do have data on both items for 3 and 17 May. These data are summarized in table 13.

Table 13. Algae and Phytoplankton Chlorophyll at Cartersville, 3 and 17 May 1976.

Date	Time	Agency	Chlorophyll *a* (ug/L)	Chlorophyll *b* (ug/L)	Phytoplankton algae (cells/mL)
3 May 76	08:30	Survey	15	1	Not tested
3 May 76	09:00	Survey	Not tested	Not tested	16,000
17 May 76	10:30	Survey	43	0.0	Not tested

There was more than twice as much chlorophyll on 17 May as on 3 May. One might therefore expect that the DO on 17 May should have been higher; but a glance at table 12 shows that DO was nearly 20 percent higher on 3 May. The percent saturation was only about 8 percent lower on the 17th, which rules out any significant temperature effect. It is not sufficient to measure chlorophyll alone. Chlorophyll raises DO only when the sun is shining. Perhaps the sun was stronger on 3 May than on 17 May; we have no way of knowing from the existing records, which do not report the intensity of solar radiation. The Survey merely reported that it was cloudy on both days. Since the riverflow was unsteady, the DO fluctuations might be explained by the scouring of opaque sediments and decomposing wastes from the riverbed; but the light extinction and deoxygenation attributable to sediments are not measured by the State or the Survey. We are left with a paradox. On 17 May both the riverflow and the chlorophyll concentration were much higher than on 3 May. Nevertheless, both DO and percent saturation were lower on the 17th.

Although the Survey fairly frequently measures phytoplankton algae (the kind of algae that float freely in the water), the Survey rarely measures the kind of algae that are attached to the bed and banks of the river; attached algae are called periphyton. On the rare occasions when the Survey did test at Cartersville for periphyton, it sometimes found plenty. It makes no difference to the process of biological reaeration whether the chlorophyll is in phytoplankton or periphyton. On 14 June 1976, for example, the Survey found no phytoplanktonic chlorophyll but did find a sizable mass of periphyton and some periphyton chlorophyll. To interpret biological reaeration fairly, we need full data on chlorophyll and algae in all forms--both attached and free floating. But full data on both kinds of algae are extremely rare at Cartersville.

There is a further complication. Photosynthesis depends on more than chlorophyll. It depends on the physiological state of the algae--in a word, on their general health. There are many measures of algal physiological activity: enzyme activity, dark ammonia uptake, radiocarbon fixation. But none of these physiological tests has been used by the State or the Survey at Cartersville. Consequently, little can be said about biological reaeration at Cartersville and no firm conclusions can be drawn.

There is another complication in the record. Notice that the Survey measures phytoplanktonic algae by counting total cells per milliliter. (See table 13.) If all phytoplanktonic algae were more or less the same size, this measurement could be helpful. However, algae vary enormously in size, mass, cell respiration, photosynthetic activity, and chlorophyll content. Some algae are huge; others are minuscule. Ignoring the difference between a large algal cell and a small one (or a healthy cell and a sick one) is rather like ignoring the difference between an elephant and a flea--the size discrepancy is misleading, even though both are animals. The Survey should be encouraged to delete this misleading measurement from its list of tests.

But large or small, algae contain chlorophyll. They must contain chlorophyll, just as human beings must have hemoglobin in their blood. You just can't have algae without also having chlorophyll. Paradoxically, the Survey often reports the presence of phytoplanktonic algae but simultaneously reports the complete absence of phytoplanktonic chlorophyll. 1/ See table 14 for an extract of this impossible data set.

1/We checked all the Survey's data sheets on phytoplanktonic algae. These papers show that all types of algae were found at Cartersville. Some days the dominant algal types were tiny diatoms; on other days the dominant algae were large filamentous blue-green algae or gelatinous colonies of branched green algae. No matter what the algae were, table 14 shows that the Survey could not reliably detect chlorophyll. Clearly, there is something wrong with the chlorophyll analyses. They cannot be trusted.

Table 14. Phytoplanktonic Algae and Phytoplanktonic Chlorophyll at Cartersville. An Extract of Impossible Data, January 1975 - August 1977. Source: U.S. Geological Survey.

Date	Time	Phytoplanktonic Algae (total cells per mL)	Phytoplanktonic Chlorophyll *a* (ug/L)	Phytoplanktonic Chlorophyll *b* (ug/L)
27 Jan 75	09:30	2,700	0.0	0.0
22 Sep 75	10:30 & 11:00	490	0.0	0.0
17 Nov 75	10:00 & 10:30	220	0.0	0.0
15 Dec 75	09:00 & 09:30	160	0.0	0.0
12 Jan 76	08:30 & 09:00	170	0.0	0.0
9 Feb 76	08:30 & 09:00	300	0.0	0.0
8 Mar 76	08:30 & 09:00	1,200	0.0	0.0
5 Apr 76	08:00 & 08:30	1,600	0.0	0.0
1 Jun 76	09:00 & 10:00	4,100	0.0	0.0
23 Aug 76	08:30 & 09:00	600	0.0	0.0
20 Sep 76	08:30 & 09:00	130	0.0	0.0
2 Nov 76	08:15 & 09:00	430	0.0	0.0
10 Jan 77	09:30 & 09:45	480	0.0	0.0
19 Jul 77	13:00	710	0.0	0.0
9 Aug 77	13:00	69	0.0	0.0

On November 21, 1979, the Survey's Chief, Quality of Water Branch, wrote us that he disagreed with the analysis of table 14:

> "Most chlorophyll samples in Table 14 were collected at a different time than the cell counts and therefore should not be directly compared. However, cell counts can be used as estimates of the expected chlorophyll values by using the following table:

APPENDIX VII

Phytoplankton	Chlorophyll Content per cell (10-12g/cell) [sic]	Chlorophyll Concentration at 1000 cells/ml (ug/L)
Diatoms	0.5-5.0	0.5-5.0
Green Algae	0.1-1.0	0.1-1.0
Blue Green Algae	0.1-1.0	0.1-1.0

(Eppley and Sloane, 1966, Phys. Plantarum, V. 19, pg. 47-59.)

"The highest cell count from Table 14 (4,100 cells per ml) could be expected to produce a maximum chlorophyll value of 20 ug/L if the sample were composed entirely of diatoms containing the largest amounts of chlorophyll. Other combinations of algae would produce less than 5 ug/L. The chlorophyll samples were analyzed by method B-6501-77 (TWRI, Book 5, Chapter A4, page 209) which calls for use of acetone to extract the chlorophyll from the cells. Acetone is not particularly efficient in extracting chlorophyll; therefore the lowest value for which an estimate of precision is made is 5 ug/L. This is also the lowest detection level. Comparing the number of cells per ml given in Table 14 with the expected chlorophyll values from the table above, it is obvious that Table 14 is not a table of impossible values but a table of values below the detection limit of the method."

The Survey's arguments come to grief on several counts. The arguments based on the generalized estimates of Eppley and Sloane are contradicted by the Survey's own data at Cartersville. The Chief argues that 4,100 cells per mL would be expected to produce a <u>maximum</u> chlorophyll value of 20 ug/L. On 2 June 1975 the Survey reported 2,900 algal cells at Cartersville and gave the chlorophyll <u>a</u> concentration as 55 ug/L. The Survey did not report the genera of algae represented in the 2,900 cells. However, it <u>is</u> clear that the estimate derived from Eppley and Sloane (15 ug/L at most) is not consistent with the Survey's own data; the actual chlorophyll measurement (55 ug/L) is 267 percent higher than the highest possible estimate from Eppley and Sloane (15 ug/L).

The Chief admits that the Survey's method for chlorophyll analysis cannot reliably detect low concentrations. We recommend the method commonly used in limnology, which reliably detects concentrations well below 5 ug/L. Grinding the filter helps the extraction. The standard reference is:

Gaulterman, H. L. (1969). Method for chemical analysis of fresh water. IBP Handbook #8. Blackwell Scientific Publications, Oxford.

We agree that one sample should be used for cell counts and chlorophyll analysis. We encourage the Survey to stop running these analyses on separate samples, which is inefficient and unscientific.

Whatever the cause, the Survey's data on algae, chlorophyll, and biological reaeration explain very little.

Discrepancies between State and Survey DO data

Both the State of Virginia and the U.S. Geological Survey collect DO data at Cartersville, but their sampling programs are not coordinated. They seldom test for DO on the same day or at the same time of day. DO data from these two agencies are presented in tables 15-17. Tables 15 and 16 are organized and summarized by water year, which begins October 1. Table 17 gives the DO data in ordinary chronological order and includes other pertinent information.

The Survey's data (see table 15) are used by the U.S. Council on Environmental Quality in its annual assessments of water quality. In the supporting analyses for its ninth annual report, CEQ concluded that DO significantly improved between October 1974 and September 1977 at Cartersville. Table 15 confirms CEQ's conclusion; it shows that the annual average DO progressively increased from 8.7 mg/L (in 1975) to 9.5 mg/L (in 1976) to 10.05 mg/L (in 1977).

Nevertheless, it is odd that the DO should have improved as the James River sank deeper and deeper into drought. (See table 1, p. 89.) There are no wastewater discharges near Cartersville, and for several miles upriver the water tumbles over a rocky bed. As the riverflow declined, the river lost some of its capacity for physical reaeration; the reduced capacity for physical reaeration suggests that the DO should not have improved during the drought. Yet the Survey's data unquestionably show improvement. How can this be?

It is tempting to suggest that biological reaeration from algae more than compensated for the warmer water and the reduced physical reaeration. But the Survey's data on algae and chlorophyll are incomplete and untrustworthy (see pp. 107 to 112); it is not prudent to draw conclusions from them. Even if we used the Survey's data, we could not find much chlorophyll during 1977, when the drought was at its worst and conditions for algal growth may have been especially favorable.

Since the State also tests DO at Carterville, we analyzed the State's data. (See table 16.) We again counsel caution in drawing conclusions from the data and the annual averages. Whereas the Survey almost always took more than one DO reading a month and never missed an entire month, the State never took more than one reading a month and frequently missed several months a year. In water year 1975 it missed 1 month, in 1976 4 months. In 1977 it again missed 4 months and also reported an improbably low DO reading for August.

With this caution in mind, what do these data seem to show? If we include the improbably low DO from August 1977, the annual averages show that DO was lower in 1976 and 1977 than it was in 1975. If we exclude the improbable DO value, the averages show that DO was lower in 1976 than in either 1975 or 1977. In short, the progressive DO improvement shown by the Survey's annual averages is not confirmed by the State's averages. Had CEQ used DO data from the State rather than from the Survey, it would have reached entirely different conclusions about trends at Cartersville.

Please bear in mind that the missing months may bias the annual averages. By missing winter months (when the water is cold and DO is high), the annual average is biased low; conversely, by missing summer months (when the water is warmer and DO is lower), the annual average is biased high. If the Survey had taken its usual two DO readings a month in November and December 1976, the average for water year 1977 might have been even higher than 10.05 mg/L. Of course, the averages are affected by any kind of omission: floods, rainstorms, freezes and thaws, sudden algal blooms, and so on.

Table 17 shows another important discrepancy between the State and the Survey. Please examine the "DO" and "Percent Saturation" columns for autumn of 1974 and the first part of 1975. You will see that the Survey's values for DO and percent saturation are often much lower than the State's. For example, in October 1974 the State reported 107 percent saturation, but the Survey reported 83 percent. In November 1974 the State reported 94 percent saturation, but 6 days later the Survey reported 67 percent. It is difficult to understand why the river lost over 25 percent of its oxygen supply in 6 days. But there are more puzzling examples. On July 1, 1975, the State and the Survey took measurements only a few hours apart. Table 17 shows a difference of 26.8 percent in the DO (7.1 versus 9.0) and a 26.6 percent difference in saturation. These differences are inexplicable. It is not easy to believe that the river actually did what these data say it did. We suspect the data.

One wonders whether the Survey's DO reports (especially in November and December of 1974) are sound. Please note that there is generally better agreement in the data after 1975, especially if we disregard two glaringly improbable results in the State's data: (1) the temperature (and hence the percent saturation) on June 2, 1977, and (2) all the data for August 16, 1977. We suspect that the State mistyped some numbers. There are still some peculiarities after 1975 (e.g. May 1976), but there are fewer abnormalities in the data sequence. Of course the State failed to take readings for 5 months in 1975 and 3 months in 1976, so there are fewer occasions for discrepancy.

DO, temperature, and percent saturation are simple measures of water quality. The discrepancies between the State and the Survey cannot be attributed to sophisticated laboratory refinements. These discrepancies emphasize the frailty of conclusions drawn from these data.

Table 15. The Survey's Data on Dissolved Oxygen at Cartersville, October 1974 - September 1977.

Year	Month	Date	DO (mg/L)	Year	Month	Date	DO (mg/L)	Year	Month	Date	DO (mg/L)
1974	Oct.	7	7.5	1975	Oct.	6	8.8	1976	Oct.	4	8.4
		21	9.2			20	8.2			26	10.8
	Nov.	11	7.6		Nov.	3	10.0		Nov.	2	11.1
		25	8.6			17	10.6		Dec.	2	13.0
	Dec.	9	7.2		Dec.	2	13.0	1977	Jan.	10	14.1
		23	10.1			15	11.4			25	15.0
1975	Jan.	27	10.6			30	11.0		Feb.	7	15.2
	Feb.	10	10.4	1976	Jan.	12	14.0			22	13.4
		24	10.8			26	13.4		Mar.	7	11.3
		24	10.8*		Feb.	9	12.2			20	10.9
	Mar.	10	10.6			23	10.4		Apr.	6	10.4
		24	9.6		Mar.	8	10.6			19	9.9
	Apr.	7	10.0			22	10.2		May	4	8.8
		7	10.0*		Apr.	5	9.9			18	6.6
		21	9.4			19	8.2		June	1	8.8
	May	5	8.6		May	3	8.2			15	9.0
		19	8.4			17	6.9		July	5	7.0
	June	2	7.2		June	1	7.5			19	7.5
		17	7.4			14	8.0		Aug.	9	6.6
	July	1	7.1			28	7.3			22	7.8
		14	7.3		July	26	8.1		Sept.	6	7.6
		28	9.0		Aug.	9	7.3			21	8.0
	Aug.	11	7.1			23	7.2				
		25	6.5		Sept.	8	8.0				
	Sept.	8	7.7			20	7.6				
		22	7.5								

Water Year 1975		Water Year 1976		Water Year 1977	
DO Sum:	226.2	DO Sum:	238	DO Sum:	221.2
N:	26	N:	25	N:	22
DO Mean:	8.7	DO Mean:	9.5	DO Mean:	10.05

*Two separate DO readings were taken on these dates.

Table 16. The State's Data on Dissolved Oxygen at Cartersville, October 1974 - September 1977.

Year	Month	Date	DO (mg/L)	Year	Month	Date	DO (mg/L)	Year	Month	Date	DO (mg/L)
1974	Oct.	28	10.4	1975	Oct.	21	9.0	1976	Oct.	28	11.2
	Nov.	5	9.0		Nov.	14	9.4		Nov.	—	—
	Dec.	—	—		Dec.	29	13.2		Dec.	13	10.8
1975	Jan.	30	11.3	1976	Jan.	—	—	1977	Jan.	—	—
	Feb.	11	12.5		Feb.	—	—		Feb.	22	12.4
	Mar.	3	12.0		Mar.	3	8.8		Mar.	24	10.4
	Apr.	24	9.9		Apr.	—	—		Apr.	—	—
	May	2	9.0		May	14	9.3		May	9	9.2
	June	23	9.6		June	25	7.4		June	2	8.0
	July	1	9.0		July	2	7.3		July	20	8.8
	Aug.	26	7.2		Aug.	10	7.6		Aug.	16	1.0 (sic)
	Sept.	18	8.0		Sept.	—	—		Sept.	—	—

Water Year 1975		Water Year 1976		Water Year 1977		
DO Sum:	107.9	DO Sum:	72.0	DO Sum:	71.8	
N:	11	N:	8	N:	8	(includes August)
DO Mean:	9.81	DO Mean:	9.00	DO Mean:	8.98	
				DO Sum:	70.8	
				N:	7	(excludes August)
				DO Mean:	10.1	

Table 17. State and Survey Data on Dissolved Oxygen, Temperature, and Percent Saturation at Cartersville, October 1974 - September 1977.

Year	Month	Day	Hour	Agency	Daily Average Riverflow (cfs)	Temperature (°C)	DO (mg/L)	DO Saturation (percent)
1974	Oct.	7	12:00	GS	1,950	21.0	7.5	83.3
		21	09:30	GS	3,170	11.0	9.2	82.9
		28	13:30	VA	1,990	16.7	10.4	107.1
	Nov.	5	13:35	VA	1,960	17.8	9.0	94.7
		11	10:30	GS	1,950	10.5	7.6	67.3
		25	10:00	GS	2,430	16.0	8.6	86.0
	Dec.	9	09:30	GS	14,300	5.0	7.2	56.2
		23	09:30	GS	5,970	4.0	10.1	83.2
1975	Jan.	27	09:45	GS	24,000	6.0	10.6	84.8
		30	13:25	VA	12,100	9.4	11.3	97.3
	Feb.	10	09:30	GS	13,800	3.0	10.4	77.0
		11	14:20	VA	11,800	6.7	12.5	102.5
		24	10:00	GS	8,620	11.0	10.8	97.3
		24	10:05	GS	8,620	11.0	10.8	97.3
	Mar.	3	14:50	VA	8,010	3.3	12.0	88.9
		10	08:30	GS	5,860	6.0	10.6	84.8
		24	08:00	GS	22,600	11.0	9.6	86.5
	Apr.	7	08:30	GS	9,580	9.0	10.0	86.2
		7	08:38	GS	9,580	9.0	10.0	86.2
		21	08:30	GS	5,880	14.0	9.4	90.4
		24	13:25	VA	5,460	18.9	9.9	105.3
	May	2	14:50	VA	15,500	17.8	9.0	94.7
		5	08:30	GS	24,400	16.0	8.6	86.0
		19	08:45	GS	11,500	18.0	8.4	88.4
	June	2	08:30	GS	15,800	22.0	7.2	81.8
		17	08:00	GS	4,850	25.0	7.4	88.1
		23	15:30	VA	2,940	27.8	9.6	121.5
	July	1	08:30	GS	3,960	25.0	7.1	84.5
		1	14:15	VA	3,960	26.7	9.0	111.1
		14	08:30	GS	33,300	21.0	7.3	81.1
		28	08:30	GS	6,180	25.0	9.0	107.1
	Aug.	11	09:00	GS	4,130	25.0	7.1	84.5
		25	08:00	GS	3,500	26.0	6.5	79.3
		26	11:15	VA	3,470	28.9	7.2	92.3
	Sept.	8	11:00	GS	3,370	25.0	7.7	91.7
		18	14:00	VA	2,590	22.2	8.0	90.9
		22	10:30	GS	5,160	22.0	7.5	85.2
	Oct.	6	09:00	GS	4,440	23.0	8.8	101.1
		20	08:30	GS	28,500	15.0	8.2	80.4
		21	13:50	VA	16,200	15.6	9.0	90.0

1975 (Cont'd)								
	Nov.	3	08:30	GS	4,080	12.0	10.0	92.6
		14	13:20	VA	13,000	—	9.4	—
		17	10:00	GS	8,310	10.0	10.6	93.8
	Dec.	2	08:30	GS	4,200	7.0	13.0	106.6
		15	09:00	GS	4,000	8.0	11.4	95.8
		29	14:15	VA	10,000	6.1	13.2	105.5
		30	08:30	GS	9,410	5.0	11.0	85.9
1976	Jan.	12	08:30	GS	8,370	2.0	14.0	101.4
		26	08:30	GS	5,290	3.0	13.4	99.3
	Feb.	9	08:30	GS	8,280	4.0	12.2	93.1
		23	08:30	GS	10,600	9.0	10.4	89.7
	Mar.	3	12:30	VA	5,670	13.3	8.8	83.0
		8	08:30	GS	4,920	11.0	10.6	95.5
		22	08:15	GS	6,640	13.0	10.2	96.2
	Apr.	5	08:00	GS	12,400	12.0	9.9	91.7
		19	08:30	GS	4,350	21.0	8.2	91.1
	May	3	08:30	GS	5,960	17.0	8.2	84.5
		14	13:15	VA	2,630	21.1	9.3	103.3
		17	10:30	GS	7,130	21.0	6.9	76.7
	June	1	09:00	GS	15,200	22.0	7.5	85.2
		14	09:00	GS	3,530	22.0	8.0	90.9
		25	14:30	VA	11,700	25.0	7.4	88.1
		28	08:00	GS	7,220	26.0	7.3	89.0
	July	2	11:00	VA	4,650	25.0	7.3	86.9
		26	08:30	GS	1,870	26.5	8.1	100.0
	Aug.	9	08:15	GS	1,790	25.0	7.3	86.9
		10	11:50	VA	2,270	26.7	7.6	93.8
		23	08:30	GS	1,240	26.0	7.2	87.8
	Sept.	8	08:30	GS	1,010	22.0	8.0	90.9
		20	08:30	GS	1,440	22.0	7.6	86.4
	Oct.	4	08:00	GS	8,360	17.0	8.4	86.6
		26	13:30	GS	28,800	12.0	10.8	100.0
		28	11:45	VA	26,300	11.1	11.2	100.8
	Nov.	2	08:15	GS	12,600	9.0	11.1	95.7
	Dec.	2	08:30	GS	5,640	3.0	13.0	96.3
		13	14:10	VA	10,600	6.7	10.8	88.4
1977	Jan.	10	09:30	GS	3,800	0.0	14.1	96.6
		25	08:30	GS	3,740	1.0	15.0	105.6
	Feb.	7	08:30	GS	3,240	0.0	15.2	104.1
		22	08:00	GS	3,620	3.5	13.4	102.3
		22	14:00	VA	3,620	8.0	12.4	104.1
	Mar.	7	08:30	GS	6,780	9.5	11.3	100.0
		20	11:30	GS	8,740	11.0	10.9	98.2
		24	13:30	VA	11,200	14.0	10.4	99.9
	Apr.	6	13:30	GS	43,400	14.5	10.4	102.0
		19	12:15	GS	5,670	20.0	9.9	107.6
	May	4	12:30	GS	4,010	21.5	8.8	100.0

1977 (Cont'd)							
	9	13:30	VA	3,820	20.0	9.2	100.0
	18	12:45	GS	2,770	25.0	6.6	78.6
June	1	12:00	GS	2,420	23.0	8.8	101.1
	2	13:50	VA	2,710	2.8 (sic)	8.0	59.3 (sic)
	15	12:30	GS	1,770	23.5	9.0	105.9
July	5	12:00	GS	1,260	29.0	7.0	89.7
	19	13:00	GS	945	32.5	7.5	102.7
	20	14:15	VA	902	32.0	8.8	118.9
Aug.	9	13:00	GS	612	30.5	6.6	88.0
	16	12:50	VA	1,230	3.0 (sic)	1.0 (sic)	7.4 (sic)
	22	12:30	GS	1,110	25.5	7.8	95.1
Sept.	6	12:30	GS	1,080	29.0	7.6	97.4
	21	13:00	GS	974	27.5	8.0	101.3

CONCLUSIONS

The water quality data at Cartersville portray a complex river--a river too complex to be meaningfully described by a few samples a month. Even the most fundamental measurements, such as temperature, are filled with apparent inconsistencies, which may be traced in part to the uncoordinated sampling programs of State and Federal agencies.

Despite these complexities, CEQ published a formal report claiming statistically significant changes in water quality at Cartersville between October 1974 and September 1977.

1. Dissolved Solids and Conductance

What CEQ called a water quality trend might more fairly be described as a normal consequence of a drought. CEQ concluded that TDS significantly increased between 1974 and 1977. We agree that TDS rose during this interval, but we attribute this change to the deepening drought, which was most severe during the last year. Naturally the river got saltier as the drought deepened.

Sometimes the water quality data are not internally consistent. Duplicate samples for TDS often disagreed widely. TDS and specific conductance are two common ways to measure for the saltiness of water; these two measurements should be closely correlated. The ratio of TDS to conductance should be fairly stable, and conductance should always be greater than TDS. The data for Cartersville, however, show great instability in the ratio of TDS to specific conductance, and there were often days when TDS was reported equal to or greater than conductance. This obvious discrepancy in the data must be explained. After reviewing our analysis, the Survey agrees that the data at Cartersville (especially in 1975-76) are suspect.

2. Total Zinc

Our analysis of total zinc suggests that changing riverflow patterns and inflexible sampling schedules accounted for the trend CEQ reported. Although zinc concentrations were always at safe levels, most of the exceptionally high concentrations and loads of total zinc coincided with dramatic changes in riverflow. These dramatic changes were especially common in the first part of the data record (December 1974 through June 1976). As luck would have it, the Survey rarely analyzed for total zinc during dramatic changes in riverflow after about September 1975. However, whenever the Survey did take a sample for total zinc analysis at or near a big peak in the riverflow record, it found unusually high values. Fixed-frequency sampling caused the Survey to miss most of the peak riverflows after September 1975. The State (as it happened) analyzed for total zinc only once at or near a peak in the riverflow record. This trick of chance might explain the differences between the Survey's data and the State's data. Had CEQ used the State's data rather than the Survey's, it would have come to entirely different conclusions about zinc trends.

3. Dissolved Zinc

The data on dissolved zinc cannot be rationally explained by any hypothesis of pollution control. The data are not consistent with either a point or nonpoint hypothesis of pollution. Consequently, no supportable conclusions can be drawn from the data.

4. Temperature

The Survey and State records show unstable water temperature. The temperature regime in this section of the James River is apparently too complex to be meaningfully represented by a few samples a month at Cartersville.

5. Dissolved Oxygen

The DO data at Cartersville are not complete or consistent. The factors that might explain these inconsistencies (e.g. temperature, algae, and chlorophyll) are not reported in sufficient detail to clarify them. The algal data, in particular, often fail to make sense, and the temperature data are open to question. We do not believe that valid conclusions can be drawn from these data.

State officials told us that very detailed surveys of water quality would be necessary to answer the questions we raised about the dramatic changes in DO at Cartersville. State officials advised us that no detailed surveys have been performed on the James

River at Cartersville because no pollution problems had been identified before we presented our analysis of the data.

On close technical examination, the DO data at Cartersville do not make sense. The State and the Survey--the agencies that have collected these data--now agree that much more detailed technical work will be necessary before the DO at Cartersville can be fairly assessed. Until that careful technical work has been done, the existing water quality records should be used with extreme caution and no conclusions should be drawn from them.

☆U.S. GOVERNMENT PRINTING OFFICE: 1981—341-843:635

(087170)

AN EQUAL OPPORTUNITY EMPLOYER

UNITED STATES
GENERAL ACCOUNTING OFFICE
WASHINGTON, D.C. 20548

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE,$300

POSTAGE AND FEES PAID
U. S. GENERAL ACCOUNTING OFFICE

U.S.MAIL

SPECIAL FOURTH CLASS RATE
BOOK

BY THE COMPTROLLER GENERAL

Report To The Congress

OF THE UNITED STATES

Better Monitoring Techniques Are Needed To Assess The Quality Of Rivers And Streams: Volume II

This volume contains detailed agency comments and evaluation by GAO and its consultant.

CED-81-30
APRIL 30, 1981

Request for copies of GAO reports should be sent to:

U.S. General Accounting Office
Document Handling and Information
Services Facility
P.O. Box 6015
Gaithersburg, Md. 20760

Telephone (202) 275-6241

The first five copies of individual reports are free of charge. Additional copies of bound audit reports are $3.25 each. Additional copies of unbound report (i.e., letter reports) and most other publications are $1.00 each. There will be a 25% discount on all orders for 100 or more copies mailed to a single address. Sales orders must be prepaid on a cash, check, or money order basis. Check should be made out to the "Superintendent of Documents".

PREFACE

This volume includes the complete comments from Federal agencies together with evaluations of the agency comments by us and our consultant. Three Federal agencies were asked to review the report and provide us with comments -- the Environmental Protection Agency (EPA), the Department of the Interior, and the Council on Environmental Quality (CEQ). All three agencies commented on the report and provided us with comments from constituent groups or other interested parties. A summary of the agencies' seven major areas of concern and our responses is included in chapter 5 of volume I.

The comments are frequently long and critical of our positions. They also contain technical data and suggestions for changes to our draft report. Where appropriate, changes have been made to the final report, as discussed in our evaluations.

ENVIRONMENTAL PROTECTION AGENCY

By letter dated May 29, 1980, EPA summarized its basic concern about the analysis supporting our recommendation and provided technical comments on our draft report. In addition, EPA provided us with comments from the Wisconsin Department of Natural Resources, which reviewed the draft report at EPA's request. All of the comments provided to us by EPA are included in appendix VIII. Our evaluations are keyed to the agency's comments. Wisconsin's comments are specifically noted.

Although EPA agreed with some of the report's conclusions about the usefulness of special studies, it did not concur with our recommendation to discontinue the networks. EPA's comments generally reflect six concerns:

--Our limited use of technical and scientific literature in the report.

--Our supposed confusion over the purpose of the fixed-station networks.

--More fundamental reasons for reexamining the design of the networks than those discussed in the report.

--Our supposed limited assessment of the use of fixed-station monitoring information.

--The cost of conducting special studies.

--Continuance of the networks until a good alternative has been tested and developed.

We have responded in some detail to EPA's comments. Overall we are not persuaded by EPA's comments that our conclusions and recommendations are not valid.

DEPARTMENT OF THE INTERIOR

In a letter dated June 24, 1980, the Department of the Interior provided comments by the Geological Survey and the Office of Water Research and Technology (OWRT). The Survey's comments, include (1) an executive summary, (2) a discussion of our recommendations, (3) comments on the body of our draft report, (4) comments on the James River, Virginia, case study, and (5) several attachments and references. OWRT provided general and specific comments on our draft report. All of the comments provided by the Department are included in appendix IX.

Generally, both the Survey and OWRT disagreed with our recommendations. Our evaluations of the Survey and OWRT comments are included in appendix IX.

The Department's Assistant Secretary for Policy, Budget, and Administration stated that the Survey's general program of water resources investigations received uneven treatment and the network was evaluated against objectives other than those for which the program was designed. It was never our intention to review the entire range of the Survey's water resource investigations, as is set forth in the scope of our review on page 4, Volume I. Also, we did not evaluate the national network against improper objectives. Basically the Survey's network cannot meet its established objectives of providing accurate and meaningful data for assessments of water quality conditions and trends. Therefore, we continue to believe our conclusions and recommendations are valid.

COUNCIL ON ENVIRONMENTAL QUALITY

CEQ's response dated May 29, 1980, includes comments from Mr. John Ficke, who developed much of the information for the CEQ 1978 annual report used in conjunction with the James River, Virginia, case study we presented in appendix VII. CEQ also provided a copy of a January 11, 1980, letter to us on a previous discussion with CEQ officials on the James River case study.

All of CEQ's comments are included in appendix X as well as our evaluation of the comments. Mr. Ficke's comments on our use of information from CEQ's 1978 annual report in the James River case study have been evaluated by Mr. Jerome Horowitz, our consultant, in appendix XI.

CEQ agreed with many points made in the report, particularly concerning quality assurance and the need to encourage a strong program of special studies. CEQ also agreed that some of the funds currently used for fixed-station monitoring could be better used for special studies. But CEQ did not agree with our recommendation to discontinue the networks and expressed concern that it could not meet its legislative mandate if only data from a program of special studies were available.

We believe CEQ does not need network data to meet its legislative mandate. The mandate can be met through the results of special studies and other reliable descriptions of water quality conditions and trends. We stand by our conclusions and recommendations.

VOLUME II

ABBREVIATIONS

BOD	biological oxygen demand
CEQ	Council on Environmental Quality
DO	dissolved oxygen
EPA	Environmental Protection Agency
FIFRA	Federal Insecticide, Fungicide, and Rodenticide Act
GAO	General Accounting Office
NASQAN	National Stream Quality Accounting System
NWQSS	National Water Quality Surveillance System
OWRT	Office of Water Research and Technology
TDS	total dissolved solids
TOC	total organic carbon

APPENDIX VIII

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON, D.C. 20460

MAY 29 1980

OFFICE OF
PLANNING AND MANAGEMENT

GAO EVALUATION

COMMENT No. PAGE(S)

Mr. Henry Eschwege, Director
Community and Economic Development Division
U.S. General Accounting Office
Washington, DC 20548

Dear Mr. Eschwege:

We appreciate the interest and concern about water monitoring programs displayed in the draft of the General Accounting Office's (GAO) proposed report, "Better Monitoring Techniques are Needed for National Surface Water Quality Assessment." We have been considering many of the same points in our own review of monitoring in developing a water monitoring strategy.

The GAO draft raises several points with which we concur. For instance, we agree that intensive surveys are the best way to investigate causes of local water quality problems. Our recent program directions emphasize intensive surveys for this reason. We also concur with the need for improved analysis of data, for increased biological monitoring, and for assessing the cumulative effect of statistical errors that may build up during sample collection and analysis. In each case, our current monitoring and research programs are addressing these areas in some detail.

We have some basic concerns, however, about the analysis supporting the recommendation to discontinue the national fixed-station monitoring networks now in place.

First, we are surprised that the analysis made little or no use of the scientific and technical literature on the concepts, design, and use of fixed-station networks for analyzing water quality. The report cites only two scientific publications on monitoring, and both are primarily guides to engineers for designing in-depth investigations into local problems. These references do not address fixed-station aggregate networks. 1 9

GAO EVALUATION

COMMENT No. PAGE(S)

2

Second, we are concerned that the report confuses the purpose of national networks--to identify broad national conditions and trends--with the purpose of localized intensive surveys and scientific investigations. For example:

- A major portion of the GAO report is devoted to showing that fixed station networks cannot detect local fluctuations in water quality or explain causes for water quality problems. We are not surprised by these results, since fixed-station networks are not designed for those purposes. Although assessing local problems is an important objective for other parts of a water monitoring program, it is not the purpose of fixed-station networks.

- Networks are designed to be used in aggregate to detect broad major changes in water quality. Since they should generally not be used to draw conclusions about individual locations, the GAO report's approach of showing that one station cannot adequately analyze local problems is not particularly appropriate. The strength of a national network approach is that, as with any statistical survey, it can develop useful conclusions in the aggregate even if there are large unexplained "random" flunctions in individual observations. The GAO's analysis of an individual station is helpful in calculating the size of the "random" error, but by itself is irrelevant to deciding whether networks should be operated.

2 10-13

Third, we have more fundamental reasons than those in the GAO analysis for wanting to re-examine the design of fixed-station networks. Although the current networks can detect changes in the levels of physical and chemical pollutants, they are generally not designed to measure the chemical, physical, or biological integrity of the Nation's waters, nor are they particularly well designed to measure human exposure to toxic pollutants. EPA has already begun a re-evaluation of its monitoring program, including its fixed-station network, in light of these concerns. Neither the current networks nor the kinds of intensive surveys proposed by GAO are likely to be an adequate solution. In particular, GAO's proposal does not adequately address biological or toxic pollutant monitoring, and does not address how to standardize and aggregate results from intensive surveys on a national scale.

3 13-14

Fourth, we feel that the report is very limited in its assessment of the use of fixed-station monitoring information. While this information is used extensively to assess conditions and trends nationwide in response to the Clean Water Act, it is also used extensively as an information base for regulatory processes under other statutes such as the Federal Insecticide, Fungicide and Rodenticide Act (FIFRA). Recently, for example,

4 14

GAO EVALUATION

COMMENT No. PAGE(S)

3

data collected at a subset of the NASQAN network was used as supporting evidence in an exposure case in the 2, 4, 5-T/Silvex cancellation hearings under the authority of FIFRA. The Agency also expects to meet certain monitoring information needs of the Toxic Substance Control Act through fixed-station monitoring efforts.

Fifth, we feel that the report grossly underestimates the cost of conducting an intensive survey as part of a "well coordinated program of special studies" nationwide. The cost figure quoted in the report, $14,000 per survey, is an accurate cost figure for very simple intensive surveys conducted under the Basic Water Monitoring Program. These surveys are generally short in duration and limited to the conventional constitutents. However, if used as part of a coordinated program nationwide which must address water quality conditions beyond those represented by conventional parameters, we can expect the cost per survey to increase substantially, perhaps as much as five times the cost of a typical intensive survey currently conducted under the Basic Program. 5 15

Finally, even if our current examination of water monitoring concludes that we should de-emphasize fixed-stations, we would not want to discontinue the networks until we have developed and tested a good alternative. Even States that have switched heavily to intensive surveys and biological monitoring have found a clear need for a balanced program of both intensive surveys and fixed-stations. I am enclosing some comments along this line from the State of Wisconsin, one of the heaviest users of intensive surveys. You are also welcome to review comments we have received from other States supporting fixed-station networks. 6 15

In summary, although we agree with many of the GAO report's conclusions regarding the usefulness of intensive surveys, we do not concur with the recommendation to discontinue networks, and we suggest that GAO reconsider its analytical approach in light of some of the concerns described above. I am also enclcsing a number of more detailed technical comments on the report. I hope you find them useful.

Yours sincerely,

William Drayton, Jr.,
Assistant Administrator for
Planning and Management

Enclosures

APPENDIX VIII

Technical Comments

	GAO EVALUATION COMMENT No.	PAGE(S)
Page 3:[a] The scope of the review was limited basically to rivers and streams. This approach overlooks a major part of the Nation's waters, including lakes, estuaries, coastal zones, and wetlands. Our own review of monitoring indicates many of these waters may be extremely critical to measuring progress toward the objective of the Clean Water Act to "restore and maintain the chemical, physical, and biological integrity of the Nation's waters." While we agree these are difficult areas to monitor, we should not ignore them.	7	15-16
Page 4: The GAO's method of selecting a single station for detailed analysis gives the reader no way of knowing whether this was a typical station. As pointed out in the text of our letter, this kind of analysis is largely irrelevant to deciding whether to run networks, but it might be relevant to setting confidence bounds on conclusions drawn in the aggregate.	8	16-17
Page 9: The GAO report strongly implies that we are continuing to "encourage" States to use fixed-stations. As the report correctly states on page 52, we are actually encouraging them to reduce emphasis on fixed-stations, except for those needed for national assessments, and increase emphasis on intensive surveys.	9	17
Page 13: The GAO draft implies that networks should produce "measurements of the quality of individual rivers." As explained in the text of the letter, we do not operate fixed networks to detect and explain local problems and fluctuations, but rather to obtain a national perspective on water quality conditions and trends.	10	17
Pages 13-46: Most of the weaknesses with fixed-station monitoring that the GAO report identifies are also problems with intensive surveys. Therefore, adopting the recommendation to discontinue fixed-stations and emphasize intensive surveys will not in itself solve the problems of proper siting, timing, and quality assurance, and may in fact increase these problems.	11	17-18
Page 26: The report implies that EPA has no quality assurance program for States. This is not true. For several years we have been conducting cooperative field and laboratory quality assurance evaluations with State agencies to improve and audit their performance. Furthermore, EPA issued a directive in June 1979 to require all recipients of EPA grants (including States) to follow a mandatory quality assurance program.	12	18
Pages 35-36: Most technical experts realize that arithmetic means are inappropriate for characterizing fecal-coliform bacteria levels. Geometric means have been commonly used for decades for this reason. There is also a general recognition that non-parametric statistics and, in some cases, the use of log-normal distributions are more appropriate statistical approaches.	13	18

[a] Page numbers refer to the GAO draft report.

 APPENDIX VIII

	GAO EVALUATION COMMENT No.	PAGE(S)
Page 39:[a] The report implies that because users of the EPA data system are responsible for providing detailed quality control, their data is less reliable. We have found that this is not the case. Most State agencies are quite conscientious about checking and editing their data in EPA's data system, because they actually use this system themselves for analyzing data and helping make regulatory decisions.	14	19
Page 39: The report correctly points out the need for identifying the precision, accuracy, and other quality assurance information for the data in our information systems. We are in the process of revising the EPA system to include such features.		
Page 41: As stated in the text of our letter, the GAO analysis of a single station is inappropriate for determining the validity of a national network. National aggregate analyses are not designed to draw specific conclusions about local conditions. Furthermore; even if GAO has examples of erroneous conclusions drawn from fixed-station data, this would not be sufficient reason to conclude that the fixed networks themselves were wrong.	15	19
Page 54: The report incorrectly states that EPA provided States and areawide planning agencies nearly $500 million in fiscal years 1973-1979 to conduct planning under the Section 208 Program. The actual figure should be $330.6 million.	16	19

a/ Page numbers refer to the GAO draft report.

State of Wisconsin \ DEPARTMENT OF NATURAL RESOURCES

Anthony S. Earl
Secretary

BOX 7921
MADISON, WISCONSIN 53707

May 6, 1980

IN REPLY REFER TO: 3200

Mr. Tom Murray
Monitoring & Data Support Div.
U.S. EPA (WH-553)
401 M. Street, S.W.
Washington, D. C. 20460

GAO EVALUATION

COMMENT No. PAGE(S)

Dear Mr. Murray:

The U.S. General Accounting Office has recently prepared a draft of a proposed report titled "Better Monitoring Techniques are Needed for National Surface Water Quality Assessments." As the Wisconsin Department of Natural Resources staff person responsible for guiding surface water quality monitoring, I offer the following comments on the draft report:

The general logic of the report implies that one should either do fixed station ambient monitoring or intensive surveys. In actuality, both systems serve a useful purpose as well as automatic monitoring and remote sensing. The lesson to be derived from this report might more properly be to de-emphasize ambient monitoring and place more reliance on intensive surveys. In Wisconsin, excluding lab needs, we have the following annual resource expenditures in surface water monitoring:

Intensive Surveys	- 10 Man-Years
Ambient Fixed Station Monitoring (51 Stations)	- 1 Man-Year
Automatic Monitoring (11 Units)	- 1 Man-Year
Remote Sensing for Lake Eutrophication	- .5 Man-Year

The intensive surveys are useful in documenting cause and effect relationships from point source discharges, developing effluent limitations and managing river systems. However, it is extremely difficult to assess general water quality on a statewide basis utilizing a mixture of unrelated intensive surveys. The manpower needs for these surveys are such that they would not routinely be repeated. Also, the parametric coverage is designed for specific needs and usually do not include most of the parameters obtained at fixed monthly stations. 17 19

The fixed station ambient monitoring data was very useful in describing general water quality for the "Wisconsin 1980 Water Quality Inventory Report to Congress." 18 20
Also, some of the paired stations illustrated the effects of point source discharges. Because of the establishment of a national network station on the Sheboygan River and the subsequent fish tissue analysis, a serious PCB contamination problem was discovered. It is unlikely that an intensive survey would have uncovered such a situation. 19 20

Mr. Tom Murray -- May 6, 1980

	GAO EVALUATION COMMENT No.	PAGE(S)
There are eleven automatic monitoring stations on two major paper mill rivers in this state. Data from these stations are used in part to assist in calibration and verification of computerized mathematical models of the river. They are also used to continuously monitor the rivers to record possible waste discharges. This spring, two of the monitors documented a severe dissolved oxygen depletion for about one week. Investigation indicated that a nonpoint source event caused the situation. Severe ground frost, a rapid snowmelt and ice cover on the river combined to affect the river. The snowmelt came from an area where there is a large amount of dairy activity and winter manure spreading on fields was the main source of pollutants. In any event, this illustrates the usefulness of automatic monitors.	20	21
Because of the large numbers of lakes in this state, we have joined efforts with the University of Wisconsin and developed a LANDSAT remote sensing computerized system to classify all lakes as to the trophic status. Attempting to do this with conventional sampling would be beyond the financial resources of the Department.	21	21
In summary, each method of monitoring has its specific level of usefulness, depending on what the data will be used for. Ambient monitoring has high lab costs and low manpower requirements while the reverse is true for intensive surveys.		
Sample collection problems and quality control could be eased by good training. Such problems are associated with any method of monitoring and are probably more severe for intensive surveys. Lab quality control is also a potential problem in any system of monitoring.	22	21
Each year we assess the appropriateness of the ambient fixed station monitors. In the last year, we reduced the number of sites from 68 to 51. Additionally, the parametric coverage was reduced on the state stations. The data that is entered into STORET should have a location description. Consequently, that knowledge should be utilized when doing a large geographic analysis. It would seem extremely difficult, if not impossible to do a national water quality assessment based on only intensive surveys.	23	22
Infrequent sampling (monthly) can give a general indication of water quality. Additionally, fish tissue analysis can indicate possible organic contamination problems in the river system. If this is discovered, such as at Sheboygan, then an intensive survey is conducted to identify and illuminate the problem. Thus a part of the ambient system can be utilized as an alert to stream contamination.	24	22
As a final comment, about 80 percent of our water quality monitoring manpower is devoted to intensive surveys while only 8 percent goes into fixed station sampling. Yet, when we attempted to describe water quality statewide, we relied mostly on the fixed station data (Sec. 305b Report). The intensive survey data is published in separate reports devoted to a particular problem. There still is a value for ambient monitoring but perhaps not for as many stations as U.S. EPA and the Geological Survey have established in the past. Intensive surveys are much more valuable in managing a state program.	25	22

GAO EVALUATION

COMMENT No. PAGE(S)

Mr. Tom Murray - May 6, 1980

Enclosed for your information is Wisconsin's Sec. 305b Report[a] and an example of an intensive survey report on the Lower Fox River. These reports illustrate different data monitoring uses.

Sincerely,
Bureau of Water Quality

Jerome R. McKersie

Jerome R. McKersie, Chief
Water Quality Evaluation Section

JRM:jm
Enc.

a/ This report was not provided by EPA as part of its comments.

GAO EVALUATION OF ENVIRONMENTAL PROTECTION

AGENCY COMMENTS

1. EPA Comment

The GAO analysis made little or no use of the scientific and technical literature on the concepts, designs, and the use of fixed-station networks for analyzing water quality.

GAO Evaluation

In carrying out our analysis, we reviewed many scientific and technical studies and articles on water quality monitoring. For report presentation purposes, we could not possibly have cited all of them and therefore were selective in our use of such literature. Appendix I contains a selected bibliography of the scientific and technical literature reviewed during our work.

Based on our analysis, we believe that fundamental flaws in fixed-station networks should preclude use of the networks for analyzing water quality.

--Network monitoring does not recognize the well-known principles of stream self-purification and the rapidly changing nature of water quality. Each foreign substance added to the water undergoes its own particular changes along the course of a river. Infrequent sampling at widely separated locations cannot possibly capture these changes. Because the networks fail to account for self-purification or capture the rapid change in water quality, data from the networks cannot provide a meaningful representation of the well-being of the Nation's rivers. (See pp. 1, 2, and 18-28, vol. I.)

--Fixed-interval sampling (generally one sample a month throughout the year) produces data from widely varying conditions, and the recorded mixture of these conditions varies from year to year. Lumping the heterogeneous data into annual averages ignores common sense and elementary statistical principles. (See pp. 29-39, vol. I.)

2. EPA Comment

The report confuses the purpose of national networks -- to identify broad national conditions and trends -- with the purpose of localized intensive surveys and scientific investigations.

GAO Evaluation

We do not agree that the report confuses the purpose of national networks. EPA agrees that fixed networks cannot detect local fluctuations in water quality or explain causes for water quality problems. We do not believe that conclusions of national scope can be drawn from data that fails to make sense at the very place it was collected. If, for example, hospitals could not accurately diagnose bubonic plague, no one could be confident of plague statistics from any hospital, from any city, or for the Nation as a whole. Similarly, we believe that national water quality conditions and trends derived from inexplicable data of questionable diagnostic value are meaningless and should not be used in forming environmental policy. We believe that fundamental weaknesses in the networks' design and operation preclude the ability to identify broad national trends and conditions.

Location bias. The Nation's waters vary significantly from place to place. Some waters are exceptionally pure whereas others are muddy, salty, or laden with wastes. Even in one small area, water quality may vary significantly. Water in falls and rapids is different from water in sluggish pools and nearby marshes. Water in urban areas is different from water in farming areas. Sparsely located stations cannot account for rapidly changing properties of water and cannot reveal the wide range of water quality conditions in each river, and yet the agencies are using isolated stations to represent the quality of large river segments.

We believe EPA and the Survey have not recognized river self-purification and the highly variable nature of water quality throughout river basins in their monitoring networks, as is discussed on page 9. A river can be very polluted in many different reaches, but can recover without a trace of aftereffects by the time it reaches a sampling station. We believe that an adequate understanding of a river's water quality condition can only come from analysis of the river and the influences on it over longer stretches -- not at single sites.

The existing EPA and Survey networks are biased with respect to location. Many of the Survey's National Stream Quality Accounting Network (NASQAN) stations are at hydraulic control points, generally away from large cities and industrial complexes. NASQAN stations were generally sited to coincide with the Survey's riverflow-gaging network; they were not selected to depict the the full range of water quality within a region or even within one river. Since water quality records cannot be interpreted without reference to riverflow, it makes sense to have water quality and riverflow measured at the same site. However, this process of site selection dramatically biases the data from the

NASQAN network. It neglects important waters where most people live and work and where serious pollution problems exist.

EPA's NWQSS network was not designed like NASQAN. NWQSS stations were usually in pairs above and below cities and other pollution-prone areas. However, EPA is now deemphasizing this network and building up the Basic Water Monitoring Program, whose stations are selected primarily by the States. But EPA has not helped the States with clear guidance on station selection.

These three networks, with less than 2,000 stations spanning the entire Nation, cannot deal comprehensively with location bias and cannot ensure that the variety of waters in any river, much less in the Nation as a whole, is meaningfully accounted for.

The James River, Virginia, case study in our report demonstrates how location can distort the networks. Cartersville is the only data-rich, long lived station on the James River which is part of all three national networks. Yet Cartersville, a tiny hamlet, is upstream from the populous, industrialized, and economically most important portions of the James between Richmond and Norfolk. Also, the Cartersville location is too far downstream to reveal pollution problems or improvements in the populated areas of Charlottesville and Lynchburg.

Timing problems. Monthly network sampling causes serious timing problems. This sampling design apparently stems from the mistaken belief that water quality measurements taken throughout the year are derived from a single homogenous "population" and that the variations observed through the year can be readily handled by routine statistical methods. This assumption is incorrect.

Water quality at network stations can change dramatically from hour to hour, as is discussed in our report. (See p. 18, vol. I.) We do not believe that one, or even a few samples a month, can fairly represent these changes. For example, monthly sampling at a fixed location is comparable to measuring the temperature on the Capitol steps in Washington, D.C., on a given day of the month at the same time of the day. Although one measurement a month would establish, over a period of several years, that it is usually hotter in July than in January, the limited sampling could not give a fair picture of temperature in the Washington area. The measurements would be biased because they would always miss the cooler temperatures at night and in the suburbs. Also, in a month like March or October, when temperatures in the Washington area could easily range from subfreezing to the nineties, a single measurement would certainly give a distorted report of reality.

The response of a river to pollution is extremely complex, and this response is sensitively balanced by the biological and physical systems of self-purification that govern water quality. Homogeneity in water quality requires that the river must have time to acquire and maintain reasonable equilibrium in its sensitively poised biological and physical systems. These systems change with the seasons of the year and establish quite different "populations" of water quality. The rotation of changes through the seasons is orderly in the long run and unique for each river, but in any season the pattern can be drastically upset by day-to-day erratic change in riverflow caused by storms, reservoir releases, etc. Hence, homogenous periods in the seasonal cycle cannot be predetermined by a fixed schedule, such as monthly sampling.

To be reliable, water quality samples must be targeted at homogenous "population" groups in the seasonal cycle -- not at an agglomeration of heterogenous groups through all seasons. Furthermore, regardless of what segment of the seasonal cycle is selected, care must be taken to anticipate a reasonably stabilized period of riverflow, sufficiently removed from preceding storms and other disturbances in the flow. Since water quality actually comprises many different "populations" of water quality through the seasons, a year's accumulation of samples cannot be grouped together statistically as homogenous when they are heterogenous. The national networks do not recognize these precepts of sound sampling.

EPA and the Survey realize it is important to measure riverflow continuously because flow may change suddenly. We believe they should be equally concerned about water quality measurements, which can be just as changeable and even more susceptable to location bias and timing problems. This is one reason why we recommend special studies rather than network sampling; because sufficiently frequent samples can be taken when and where it is most important to accurately assess water quality.

Variance and significance. When data points are tightly clustered, or when they line up neatly and consistently with normal distributions, it does not take many observations to demonstrate statistical significance. But when the data are widely scattered or when they are not normally distributed, much more data is needed to demonstrate significance. As we showed in the report, water quality in the Nation's rivers and streams generally varies and infrequent sampling leads to untidy, wide-ranging data.

Some places in the Nation have relatively stable water quality. For example, the water discharged through Hoover Dam on the Colorado River is not subject to the dramatic influences of rapidly fluctuating riverflow. The dam draws from the deeper waters of the reservoir, which smooths out the fluctuating quality of the river water. Suspended materials are sedimented; water from storms and other hydrological events has time to blend into the great mass of water stored in the reservoir. Infrequent sampling (e.g. once a month) might produce useful data for assessing trends and conditions of some water quality characteristics in the immediate outflow from some dams after many years of sampling.

Most rivers and streams, however, do not have stable water quality conditions. The James River at Cartersville is more typical. Storms muddy the river, rain dilutes dissolved materials, and major impoundments do not exist to smooth out hydrological events. Consequently, sampling data are often widely scattered and more data are required to assess trends and conditions.

Infrequent sampling produces data from heterogeneous river conditions, as we demonstrated in the report with several examples. The mix of river conditions caught by monthly samples can change dramatically from year to year, giving a misleading impression that water quality is different from year to year, when it actually may not be.

Unfortunately, the networks do not consider the variance and significance of changes in water quality. All rivers are sampled at the same frequency and at the same location, which is not sufficient for characterizing highly variable water quality.

In summary, our report does not confuse the purpose of national networks with the purpose of localized surveys and investigations. Rather, our report highlights the fundamental inadequacy of sparse sampling. Broad national trends and conditions cannot be identified by the network because of timing problems, location biases, and the inability of the networks to account for rapidly changing properties of water and for stream self-purification.

3. EPA Comment

The existing networks are not designed to measure the chemical, physical, or biological integrity of the Nation's waters, nor are they particularly well designed to measure human exposure to toxic pollutants, but the kinds of intensive surveys proposed by GAO are not likely to be adequate solutions. EPA has begun a re-evaluation of its monitoring program, including its networks. GAO does not address how to standardize and aggregate results from intensive surveys on a national scale.

GAO Evaluation

We agree that the networks are not well designed to measure the chemical, physical, or biological integrity of the Nation's waters. We also agree that the existing networks are not well designed to measure human exposure to toxic pollutants.

We do not agree, however, that the use of special studies is not likely to be an adequate solution, particularly with respect to toxic pollutants. Measuring human exposure to such pollutants requires a sampling program tailored to the places where humans are likely to be exposed and relevant pollutants are likely to be present. Special studies are particularly well suited for such purposes and can include water, biological, and sediment monitoring as well as information on the source of pollutants being evaluated.

We also believe that other indicators of water quality conditions, changes, and trends, such as the return of fish to previously polluted waters, reductions in municipal and industrial discharges, and biological monitoring can be used. EPA, CEQ, and States already have used these indicators in water quality reports.

Over time, the results of individual special studies can be aggregated to show water quality conditions, changes, and trends, as discussed on page 56, volume I. Innovative uses of these studies and other indicators, can, in our opinion, produce much more useful information than networks can.

4. EPA Comment

The report is very limited in its assessment of the use of fixed-station monitoring information. Such information has been used extensively in response to the Clean Water Act and the Federal Insecticide, Fungicide and Rodenticide Act (FIFRA).

GAO Evaluation

In its comments EPA agreed that fixed-station networks are not particularly well designed to measure the chemical, physical, or biological integrity of the Nation's waters or human exposure to toxic pollutants. EPA further stated it had begun a reevaluation of its monitoring program, including its fixed-station network, in light of these concerns.

We believe it is unwise for EPA to use network data to fulfill its responsibilities under FIFRA and the Clean Water Act. Special studies, including biological monitoring, are much better vehicles for assessing water contamination by toxic substances, such as pesticides.

5. EPA Comment

The report grossly underestimates the cost of conducting an intensive survey as part of a well-designed nationwide program. Such surveys could cost as much as $70,000.

GAO Evaluation

We pointed out in the report that the costs of special studies vary considerably depending on the nature and type of each study. (See p. 61, vol. I.) We cited an example that cost $200,000.

We believe that EPA did not recognize our point -- that millions are currently available from various sources for water quality monitoring activities and what is needed is a well-managed nationwide program which taps these various sources. We agree that if the agencies were overly ambitious, the funding currently available may not be sufficient, but we cannot agree that a program of special studies would be prohibitively costly.

6. EPA Comment

EPA would not want to discontinue the national networks until it has developed and tested a good alternative.

GAO Evaluation

We agree that EPA should develop and test alternative water quality monitoring methods, but we do not agree that EPA should continue to fund national fixed-station networks which do not produce accurate, meaningful data. Our report demonstrates the serious weaknesses of networks and we cannot advocate continuing them.

7. EPA Comment

The scope of the GAO study was limited to rivers and streams and did not include lakes, estuaries, coastal zones, and wetlands. Although these other bodies of water are difficult to monitor, they should not be ignored.

GAO Evaluation

We agree that our review was limited to rivers and streams. This was clearly stated in the scope of our report. (See p. 4, vol. I.) We certainly agree that hydraulically complex waters such as lakes and estuaries are important and need to be evaluated. We restricted our review to relatively simple waters (rivers and streams) for several reasons:

--Most of the network stations are located on rivers and streams.

--Rivers are hydraulically simple in relation to estuaries, etc., it should be easier to understand water quality data from rivers.

--The Survey has a large system of flow gages on rivers and streams, but there is no comparable store of information of flow patterns in hydraulically complex water.

--Problems of interpreting water quality data from rivers and streams are more severe in hydraulically complex waters.

We do not believe that our exclusion of more hydraulically complex bodies of water detracts from the validity of the message of our report.

8. <u>EPA Comment</u>

The GAO selection of a single station for detailed analysis gives the reader no way of knowing whether the station was typical. Also, this kind of analysis is largely irrelevant to deciding whether to run networks.

<u>GAO Evaluation</u>

The station selected for our case study -- James River at Cartersville, Virginia -- is not typical because it is a data-rich station compared to most network stations. Despite its richness in data, however, neither the water quality of the James River nor the reasons for changes in chemical, physical, and biological characteristics could be determined. For example, the case study showed, among other things, that:

--Some water quality trends derived from the Survey's data were unlike those derived from the State's data at the same site.

--Several of the trends could be traced to quirks in sampling.

--Anomalies existed in the data.

--Some of the trends could be wrong because of changes in detection methods or errors in field procedures.

We do not believe the case study is irrelevant with respect to EPA's decision to use networks. Our evaluation explains that serious weaknesses in the networks preclude the use of such data for characterizing national water quality trends and conditions. (Discussed more fully in our evaluation of EPA comment 2.) The case study focused on one monitoring location to explore in depth the complexities of water quality. It was only part of our review.

We used other stations and rivers as examples throughout the report and drew extensively from technical literature.

9. EPA Comment

The report implies that EPA continues to encourage States to use fixed stations, but EPA is actually encouraging States to increase emphasis on intensive surveys.

GAO Evaluation

We agree that EPA is encouraging States to do more surveys, as is discussed on page 60, volume I. However, as noted on page 10, volume I, EPA still requires that States perform fixed-station monitoring. (40 C.F.R. 35.1500, *et seq.*, App. A.)

10. EPA Comment

The report implies that networks should produce measurements of the quality of individual rivers, but EPA operates the networks to obtain a national perspective on water quality conditions and trends.

GAO Evaluation

As discussed in our evaluation of EPA comment 2 (see p. 9), we believe that the networks do not produce accurate, credible data at an individual location and cannot account for water quality along a river. In view of these biases, distortions, and inconsistences in network data, we do not understand how such data can be used to derive national trends and conditions.

11. EPA Comment

Most of the weaknesses (proper siting, timing, and quality assurance) which GAO identifies are also problems with intensive surveys. The problems may increase if special studies are emphasized.

GAO Evaluation

Networks are inherently hampered by timing problems, location bias, and the inability to account for rapidly changing water quality; special studies can circumvent these problems. We have expanded our discussion in Chapter 4 to explain clearly why special studies can overcome these problems. (See p. 53, vol. I.) Special studies may vary widely in duration, number of sites, frequency of sampling, types of tests performed, and supplemental information obtained, depending on the water quality problems being studied. For example, several specific water quality problems were addressed in the Survey's Williamette

River assessment. (See p. 57, vol. I.) These problems dictated the nature and scope of the study, including sample timing and location.

With respect to quality assurance, we agree that this matter needs close attention in any type of monitoring. We believe, however, that quality assurance can receive closer attention in special studies than in network monitoring. As discussed on page 56 of volume I, the potential for specific quality control problems can be identified and the study team can take extra measures to mitigate or eliminate them.

12. EPA Comment

The report implies that EPA has no quality assurance program for States, but EPA issued a directive in June 1979 requiring all EPA grant recipients, including States, to follow a mandatory quality assurance program. For several years EPA has conducted cooperative field and laboratory quality assurance evaluations with State agencies to improve and audit performance.

GAO Evaluation

We did not intend to imply that EPA has no quality assurance program for States. Our review did not include an evaluation of quality assurance programs. We have revised the report to disclose EPA's 1979 mandatory quality assurance program.

13. EPA Comment

Technical experts believe that geometric means, rather than arithmetic means, are more appropriate for characterizing fecal coliform bacteria, and there is a general recognition that the use of nonparametric statistics and log-normal distributions are more appropriate statistical approaches.

GAO Evaluation

We agree. We have eliminated our discussion of arithmetic means for fecal coliform bacteria. Unfortunately, the Survey and CEQ have used arithmetic means to describe fecal coliform bacteria conditions in their published water quality reports.

14. EPA Comment

The report implies that because users of the EPA data system are responsible for providing detailed quality control, their data are less reliable. EPA has found that most State agencies are quite conscientious in checking and editing their data.

GAO Evaluation

We do not agree that the report makes such an implication. The report states (see p. 45, vol. I) that EPA expects agencies entering data through the automated system to provided detailed quality control over the data. We did not evaluate State quality control programs for data submitted to EPA.

15. EPA Comment

The GAO analysis of a single station is inappropriate for determining the validity of a national network.

GAO Evaluation

This comment is similar to EPA comment 8. Our conclusions and recommendations are not based on an analysis of a single station. We evaluated a large volume of technical literature and developed many examples for the report. Our evaluation explains that serious weaknesses in the networks should preclude the use of network data for characterizing water quality conditions and trends. The case study was prepared simply as an illustration of some of the problems inherent in interpreting data from the networks.

16. EPA Comment

The report incorrectly states that EPA provided nearly $500 million in fiscal years 1973-79 to conduct areawide planning under section 208 of the Clean Water Act. The actual figure should be $330.6 million.

GAO Evaluation

The report has been revised and we no longer refer to this program in the report.

17. EPA Comment (Wisconsin)

Special studies are useful for certain activities, but it would be extremely difficult to assess general water quality on a nationwide basis using special studies. The manpower needs for these studies are such that they would not be routinely repeated, and the coverage usually does not include most parameters obtained at fixed stations.

GAO Evaluation

We do not agree, as discussed in our evaluation of EPA comment 3.

18. EPA Comment (Wisconsin)

Fixed-station data were useful in describing water quality in the "Wisconsin 1980 Water Quality Inventory Report to the Congress," and some of the paired stations illustrated the effects of point source discharges.

GAO Evaluation

Use of fixed-station data is not required by section 305(b) of the Clean Water Act. Although Wisconsin used the data in its report to the Congress, we do not agree that the data can readily be accepted as representative of the quality of Wisconsin waters, for the reasons discussed in our evaluation of EPA comments 2 and 4.

With respect to the paired station concept, we believe that the effects of point source discharges are much better determined by movable stations and flexible sampling under special studies. Flexibility circumvents location bias and time bias and permits resources to be concentrated on places and times where problems are likely to occur.

19. EPA Comment (Wisconsin)

Because of the establishment of a national network station on the Sheboygan River and subsequent fish tissue analysis, a serious PCB problem was discovered, which would not likely have been uncovered by an intensive survey.

GAO Evaluation

The PCB problem was uncovered by fish tissue analysis, not water analysis. Since PCBs are virtually insoluble in water and accumulate in fish, monitoring for that chemical is best done through sediment and fish tissue analysis. The State followed the first fish tissue analysis with special sampling of fish sediment, municipal and industrial effluents, and river water at key locations upstream in the river. This investigation traced the PCB to an industrial waste disposal site next to the river. Water samples from the national network station on the Sheboygan River were not important to this study. Rather than supporting the continuation of national network water quality sampling, we believe this example demonstrates the value of special studies and alternative approaches, as discussed in chapter 4 of the report.

20. EPA Comment (Wisconsin)

Data from automatic monitoring stations in the State have been used extensively to monitor rivers to record possible waste discharges and to assist in the calibration and verification of mathematical models.

GAO Evaluation

Although continuous automatic monitors overcome one of the problems of fixed-stations -- time bias -- they do not overcome other problems, such as location bias. Most network stations do not use continuous automatic monitors, and therefore the problems of time bias remain. Also, only several water quality characteristics can be monitored by automatic devices. Fixed-station networks should not be needed for monitoring discharges because permits and regulations require dischargers to report to the State or EPA.

21. EPA Comment (Wisconsin)

Because of the large number of lakes, Wisconsin has developed a LANDSAT remote sensing system to classify all lakes as to trophic status. To do this with conventional sampling would be beyond its financial resources.

GAO Evaluation

As pointed out in the scope of our review on page 4, volume I, lakes were not included in our study, but this does not negate the validity of our observations about river and stream water quality monitoring.

22. EPA Comment (Wisconsin)

Sample collection problems and quality control can be eased by good training, but some problems are probably more severe for intensive surveys. Laboratory quality control is a potential problem in any system of monitoring.

GAO Evaluation

We agree that good training can ease such problems. We also agree that quality assurance needs close attention in any type of monitoring efforts. We do not agree, however, that these problems will be more severe for special studies, as is discussed in our evaluation of EPA comment 11.

23. EPA Comment (Wisconsin)

It would be difficult if not impossible to do a national water quality assessment based only on special studies.

GAO Evaluation

We do not agree that it would be impossible to make national water quality assessments based on only special studies. However we do not propose using solely special studies; other indicators exist of progress toward cleaner water. This matter is discussed in our evaluation of EPA comment 3. As is discussed in our evaluation of EPA comment 2, we do not believe that the limited number of network stations and periodic samples can portray water quality at the sampling site, upstream, downstream, or on a national basis. Special studies can circumvent the inherent weaknesses of networks.

24. EPA Comment (Wisconsin)

Infrequent sampling can give a general indication of water quality and fish tissue analysis can indicate possible organic contamination, as was done at Sheboygan. Thus a part of the ambient system can be utilized as an alert to stream contamination.

GAO Evaluation

As discussed previously (see our evaluation of EPA comments 2 and 3), we disagree that infrequent sampling can reliably indicate water quality. A tiny amount of water taken during a 1- or 2- hour visit can hardly be relied upon to reveal stream contamination during the entire month (or every 4 months, which is another common sampling frequency). As is further discussed in our evaluation of EPA comment (Wisconsin) 19, the Sheboygan situation illustrates the need for more intelligent use of biological and sediment monitoring, not network water quality monitoring.

25. EPA Comment (Wisconsin)

About 80 percent of the State's monitoring manpower is devoted to intensive surveys and only 8 percent to fixed stations. Intensive surveys are much more valuable in managing a State program. But, the State relied mostly on fixed-station data to describe water quality statewide for the 305(b) report. Ambient monitoring has a value but perhaps not for as many stations as EPA and the Survey established in the past.

GAO Evaluation

We applaud Wisconsin's efforts in using special studies rather than fixed stations for managing the water quality program. Section 305(b) of the Clean Water Act does not require the use of fixed-station data to describe statewide water quality. As we have discussed in response to several EPA comments, special studies can overcome the fundamental weaknesses of network monitoring and can provide accurate and meaningful assessments of water quality.

United States Department of the Interior

OFFICE OF THE SECRETARY
WASHINGTON, D.C. 20240

June 24, 1980

GAO EVALUATION

COMMENT No. PAGE(S)

Mr. Henry Eschwege
Director, Community and Economic
Development Division
U.S. General Accounting Office
Washington, D.C. 20548

Dear Mr. Eschwege:

On April 29, 1980, you transmitted for our review and comment the General Accounting Office's draft report entitled "Better Monitoring Techniques Are Needed For National Surface Water Quality Assessments." The U.S. Geological Survey and the Office of Water Research and Technology were asked to provide the review comments, which are enclosed.

The implications of the draft report are far reaching with regard to the Geological Survey's program of water resources investigations. Specifically, the report calls for the discontinuance of the Geological Survey's National Stream Quality Accounting Network (NASQAN). The Geological Survey is in substantial disagreement with the recommendations made in the report that pertain to its programs. The general program of water resources investigations of the Survey has received uneven treatment in the report, and the utility of its NASQAN program has been evaluated with respect to objectives other than those for which the program was designed.

The complex, but often poorly documented, nature of the criticisms contained in the report called for the extensive technical review comments. Because of the nature and substance of the comments, we strongly urge that the full text of the Geological Survey's response be included in any final GAO report.

Thank you for this opportunity to review the report.

Sincerely yours,

Larry E. Meierotto
Assistant Secretary
for Policy, Budget,
and Administration

Enclosure

APPENDIX IX

U.S. Geological Survey Response to the

Draft of a Proposed Report of the General Accounting Office

"Better Monitoring Techniques are Needed for

National Surface Water Quality Assessments"

TABLE OF CONTENTS

GAO EVALUATION

COMMENT No. PAGE(S)

EXECUTIVE SUMMARY

The draft report of the General Accounting Office entitled "Better Monitoring Techniques are Needed for National Surface Water Quality Assessment" (herein referred to as the GAO Report) discusses the use of fixed-station, fixed-interval monitoring networks to assess the quality of the Nation's rivers. The recommendations of the GAO Report call for the discontinuance of the U.S. Geological Survey's National Stream Quality Accounting Network (NASQAN) and two national programs of the Environmental Protection Agency (EPA). This response by the Geological Survey is limited to NASQAN and to the purposes incorporated into its design.

The U.S. Geological Survey initiated (in 1973) and maintains NASQAN, a fixed-site, fixed-interval monitoring network of national geographic scope, for the purpose of providing (spatially and temporally) consistent data records. Such records are indispensable for many purposes of policy analysis, water management and hydrologic research, including: (1) long term trend analysis; (2) construction of probability distributions; (3) development of water-quality standards; (4) determination of correlative relationships; and (5) determination of spatial transferability of information. Furthermore, because water quality and quantity are interrelated, care was taken in designing the network to associate each water-quality monitoring site with a stream gaging site to assist in analysis of the data, for example, in distinguishing discharge-related changes in water-quality variables from those due to other, perhaps anthropogenic, causes. "The primary objectives [of NASQAN] are (1) to account for the quantity and quality of water moving within and from the United States, (2) to depict areal variability, (3) to detect changes in stream quality, and (4) to lay the groundwork for future assessments of changes in stream quality" (J. Ficke and R. Hawkinson, Geological Survey Circular 719, 1975). However, NASQAN was never intended to be a source of information "...detailed enough to assess the effectiveness of pollution control measures on a localized basis, as prescribed by Public Law 92-500" (Ficke and Hawkinson, 1975). Nevertheless, the GAO Report has judged NASQAN by this inappropriate criterion.

1 86-88

Consistency is one of the most important characteristics of the NASQAN database, particularly in comparison to other national collections of water quality data. Not only are the same water quality variables monitored at the same locations at the same sampling frequencies nationwide over time, but methodologies for field sampling are specified and laboratory analyses are conducted only in Survey's two national laboratories. Mindful of the potential for human error in any monitoring effort, particularly in a large one of national scope, the Survey actively pursues a program of regular examination and critiquing of these activities on a district-by-district level; any deviations from the standard methodology are strictly noted in the district review reports and corrective actions are taken. Furthermore, the precision of laboratory analysis is comparable with the professionally accepted standards of precision set by the American Society for Testing and Materials (ASTM). Selected portions of the district review reports pointing out incorrect field techniques and sample handling are used in the GAO Report as evidence of poor data quality. To the contrary, the Survey contends that this searching out and correction of errors maintains the high quality of Survey data. Also, the GAO Report cites examples of Survey studies of laboratory precision at non-Geological Survey laboratories, and uses them to call the NASQAN data into question, even though analyses of NASQAN samples are performed only at Geological Survey laboratories.

2 88-89

	GAO EVALUATION COMMENT No.	PAGE(S)
The GAO Report (Chapter 1) lists the published objectives for NASQAN, but nowhere takes issue with them nor criticizes NASQAN with respect to their fulfillment. Nonetheless, the GAO Report recommends the discontinuance of the Survey's NASQAN monitoring network (as well as the two networks operated by EPA), contending that fixed-station, fixed-interval networks do not produce "reliable, meaningful surface water quality data" and should be replaced by a program of well-coordinated special studies of water quality.		
It is stated (in the GAO Report, Cover Summary) that this inability to produce "reliable, meaningful surface water quality data" arises for three reasons. The first two are both conditioned on the potential of the data to "not be representative of national water quality conditions" due to either (1) "infrequent sampling", or (2) "dissimilar monitoring locations". However, the concept of representativeness is an unspecific one (representative of median flow conditions? extreme flow? average load? extreme concentration values? etc?)--nowhere in the Report is the concept operationally defined so that allegations of the network's limitations on these grounds are unsupported. The third reason given for the inability of fixed-site, fixed-interval networks to produce "reliable, meaningful surface water quality data" is that "weaknessess in field sampling and laboratory procedures add uncertainty to interpretations of water quality conditions". However, uncertainty is inherent in any measurement since the process is inherently stochastic; rather, the concern should be with the precision of the measurements and their biases, if any--both of which are addressed by the district review and the laboratory quality assurance programs of the Survey.	3	90
The recommendations of the GAO Report call for the Survey and EPA to discontinue the three networks and to "...devote their resources to well coordinated special studies of water quality." The Survey strongly disagrees with this recommendation for the reason that the two approaches to water-quality investigation are different. The objectives of fixed-station monitoring focus primarily on description and characterization of water quality in space and time. Two major products of the NASQAN program are nationwide geographic summaries of water quality and identification of long-term trends at network sites. Because of the variable nature of hydrologic data, trend studies have necessarily awaited accumulation of sufficient data, but they are now underway and are unquestionably obtainable in conformity with accepted statistical procedures.	4	90
The NASQAN effort does not represent the only water quality program of the Survey; besides the Intensive River Quality Assessment Program, many monitoring networks are maintained and special studies are conducted through the Federal-State Cooperative Program, and additional monitoring activities of national scope are carried out often in conjunction with other Federal agencies. Special studies, such as synoptic studies and the Intensive River Quality Assessments, are of shorter duration and geographically more sparse, even though locally more intensive than NASQAN: they are the appropriate vehicle for exploring cause/effect relationships as well as reasons for changes in water quality. However, such changes are commonly perceived through programs of periodic monitoring and verified to be other than that expected from just the inherent	5	91

GAO EVALUATION COMMENT No.	PAGE(S)
6	91-92

variability of the underlying system. Thus, it is to be emphasized that the alternative offered to NASQAN, namely a coordinated series of special studies, would neither fulfill the objectives of NASQAN nor could it be financed at a level comparable to NASQAN and still maintain a national geographic scope. If NASQAN is to allow for a significant interpretive component, as in the IRQA program which GAO strongly endorsed, then the NASQAN program goals will have to be expanded and available resources increased.

The Survey's NASQAN program was created to fill a need for a database on river water quality that was national in scope and consistent in methods which would be suitable for examining conditions and trends in river water quality. The program is meeting this objective and meeting it well. The GAO Report does not question this objective nor does it provide a demonstrated alternative for fulfilling it. A case for the discontinuance of the NASQAN program has not been made.

The Geological Survey response to the GAO Report is in two parts: the first contains the body of the response which discusses in detail the recommendations made by GAO with respect to fixed-site, fixed-interval monitoring in general, as well as specifically to the NASQAN effort; the second part contains comments on the particulars of the report.

APPENDIX IX

	GAO EVALUATION	
	COMMENT No.	PAGE(S)

PART 1

INTRODUCTION

The draft report of the General Accounting Office, entitled "Better Monitoring Techniques are Needed for National Surface Water Quality Assessments" (herein, referred to as the GAO Report) recommends the discontinuance of the Survey's National Stream Quality Accounting Network (NASQAN) monitoring network as well as two networks (National Water Quality Surveillance System, Basic Water Monitoring Program) operated by the Environmental Protection Agency (EPA). The report contends that fixed-station, fixed-interval networks do not produce reliable, meaningful surface water quality data and should be replaced by a program of well coordinated special studies of water quality. It is stated that these special studies can produce water quality assessments which are more scientifically sound and more meaningful than those produced by fixed-station networks.

In this discussion we address the topic of fixed-interval, fixed-station monitoring and indicate why the suggested alternative, a collection of coordinated special studies, is not a substitute for the former. It is noteworthy that the alternative offered by the GAO Report, namely more special studies, when truly "well coordinated" nationally, approximates a fixed-station network for which "station" is synonymous with a larger areal site (than a cross section of a river) and the time interval between sampling events is on the order of years. The cost of such a series of special studies, if truly of national scope, would certainly far exceed the cost of NASQAN. 7 92

It is also to be noted that the Survey is not limited to the NASQAN network in carrying out its national responsibilities. Through the Federal-State cooperative program, it conducts many projects, including synoptic studies as well as studies of longer duration and larger scope, to examine and assess local conditions. At selected diverse locales, the Survey is carrying out a demonstration program of intensive river quality assessments as noted by the GAO, although the program has not been as extensive as planned due to budgetary and personnel constraints. Additional monitoring activities of national scope are also conducted, often in conjunction with other Federal agencies. Thus, NASQAN represents an important but not the sole water quality monitoring activity of the Survey, which is itself only a part of its overall program of water-quality investigations. 8 93

VARIABILITY IN WATER RESOURCE MEASUREMENTS

The Organic Act of 1879 charged the Geological Survey with assessing the Nation's mineral resources. Unlike other mineral resources, water has a temporal dimension in addition to its spatial definition, and quality characteristics include biological as well as physical and chemical parameters. Indeed, the term "water quality" alludes to any characteristic or set of characteristics describing the water resource, excluding only those that pertain strictly to quantity. Nonetheless, the quality and the quantity of

GAO EVALUATION

COMMENT No. PAGE(S)

a water resource are interrelated, and both continually affect and are affected by man's intentional as well as inadvertent actions.

The temporal variability of the characteristics of the water resource, in addition to a random component, may include a cyclic component (diurnal, weekly, annual, etc.) and a trending (continuously increasing or decreasing) component. Further, the analytic precision associated with any measurement of a characteristic adds to the perceived variability. Thus, because all data (whether pertaining to quality or quantity) arise from a random process, it is imperative to interpret the data from a stochastic viewpoint. The concept of "representativeness" can only be defined in this context. Any single measurement potentially could be higher or lower had it been made at some other time: with a single value, little can be inferred about the random process from which it arises. (The GAO Report gives many examples of the range of values that a record may exhibit, yet it chooses to discuss single values in isolation from the record.) Furthermore, data arising from a random process with large variance are not "difficult to compare" (GAO, Chapter 2), but rather the confidence intervals about statistics of interest are accordingly large, although they may be reduced with increased sample size.

9 93

With respect to spatial variability, both quantity and quality of surface waters may change both along the length of the river and across the cross-sectional width. However, quantity (water discharge) is defined in terms of flow through a cross-sectional area per unit time, inherently eliminating the need to separately consider variability in this dimension when measuring discharge at a site. Similarly, it is Survey policy to obtain a measurement of water quality that integrates the quality of all the water passing through the cross section, which commonly requires taking depth-integrated, discharge-weighted water samples across the cross-sectional area at the sampling site. Thus, separate consideration of cross-sectional variability is eliminated also for water quality analyses.

10 93-94

Thus, a single measurement of a characteristic of a water resource represents a single draw from a random process at a specific time, physically occurring in an instant across the entire cross-sectional area of a channel and reflecting the composite effect of whatever has occurred upstream from the site; the record at a fixed site represents the perception of this process over time isolated from spatial variability so that inferences about the true underlying process may be drawn.

A MONITORING NETWORK FOR NATIONAL ASSESSMENT

The need for a nationally consistent set of water quality information for assessing the state of the nation's river has been widely noted. (See, for example, M. Gordon Wolman, Science, 174 (905-918), 1971.) In response to this perceived need and to Bureau of the Budget Circular A-67 (1964), the Geological Survey instituted in 1973 a water quality monitoring network-NASQAN (National Stream Quality Accounting Network)--which was national in geographic scope and consistent in methodology with respect to sampling frequency as well as collection and analysis procedures.

11 94

GAO EVALUATION

COMMENT No. PAGE(S)

It should be noted that neither such a requirement nor such a solution is unique to the United States. For example, in 1974, the United Kingdom also established a monitoring network of fixed stations sampled on a regular interval to obtain data of a comparable precision so that river quality assessments could be made on a national level. (See WHO Water Quality Bulletin, 5:2, 1980.)

An additional consideration was provided for in the design of the NASQAN layout. Because it is known that discharge and water quality are interrelated and may be correlated, care was taken to associate each NASQAN station with a discharge monitoring site so that a complete discharge record as well as an instantaneous discharge measurement would be available in conjunction with the water quality record. (Further discussion of the importance of the availability of discharge information for water quality analysis can be found in the report, "National Assessment of Trends in Water Quality", submitted to the Council on Environmental Quality in 1972 by Enviro Control, Inc., NTIS report PB-210 669).

12 94

The objective of consistency is of great importance. One of the drawbacks of using data collected by a large number of different efforts using different methodologies is the inherently different levels of precision the data will exhibit--at best, an increased amount of variability is introduced in the assessment record; at worst, the data are not comparable. As noted in the GAO report, "Comparability of data over time at one location and among various locations throughout the Nation is also needed for good assessments" (Chapter 3). Although the GAO continues that it ". . . is not possible through the networks (to achieve this) . . .", it has been stated by researchers in the field that "Perhaps the greatest promise for improving performance in river-quality evaluation lies in the establishment . . . of the National Stream Quality Accounting Network by the USGS" (Edward Cleary, Journal Water Pollution Control Federation, 50:5 (831), 1978) exactly because of the consistency it provides.

NASQAN provides for a nationally uniform methodology both for field and laboratory analysis. In addition, because human error is always possible, a regular program for the examination of sampling procedures is actively pursued; deviations from the standard methodology are strictly noted and corrective procedures are taken. (Indeed, the GAO report, Chapter 3, presents several good examples of the careful critiques given by regional and headquarters personnel in cases of sampling practice deficiency.)

13 95

RECORDS FROM MONITORING NETWORKS

Monitoring networks are developed for the purpose of obtaining records which are consistent over space and time. Long (chronologically) records are imperative for many purposes of policy analysis, water management and hydrologic research. These include the following:

1. Long term trend assessment.

   Recent attention has focused primarily on long term trend assessment, almost to the point of eclipsing other uses for which long

14 95

APPENDIX IX

ALUATION

GAO EVALUATION

COMMENT No. PAGE(S)

records are indispensable. Because of the cyclic - diurnal, weekly, annual, - and random components of the temporal variability of any characteristic of the water resource, strong statistical statements concerning trends are not possible without long records. (The forementioned report by Enviro Control, Inc. provides a good example of the difficulty in the early 1970's of doing a nationally comprehensive study of apparent trends of ambient water quality: the authors used whatever data could be found in STORET in order to construct records of at least 8 years in length for each site in the study. Data were used regardless of where the sampling site was located (70 percent of the sites were in the northeastern or northwestern United States), of what the sampling frequency was (even if only at quarterly sampling intervals), or by what field or analytic methodology the data had been obtained. A further example of the use of long records in trend analysis is given by A.M.C. Edwards and J. B. Thornes, Water Resources Research, 9 (1286-1295), 1973.)

2. Construction of (empirical) probability distributions.

The distribution of a water quality characteristic of interest can be developed only if a (trend free, i.e., stationary) consistent record of some length is available. The distribution so defined is conditioned on the sampling frequency; however, incidental sample values, such as obtained in synoptic studies, can only be evaluated within the context of the statistical distribution of the characteristic of interest. (H.A.C. Montgomery and I.C. Hart, Water Pollution Control, 73 (77-101), 1974, have presented a good discussion of the utility of such information.) 14 95

3. Construction of standards.

The ability to construct standards which are neither impossible to achieve nor irrelevent is predicated on the knowledge of the range of values (and their distribution) of the characteristic of interest. For example, as pointed out by R.D. Pomeroy and G.T. Orlob (Problems of Setting Standards and of Surveillance for Water Quality Control, California State Water Quality Control Board, State of California, 1967): "A limit of 100 mg/l of hardness in a groundwater basin replenished by Colorado River water would be futile, since it is unattainable. The same limit for certain streams in the northern part of the State or in the mountains probably would not be adequate control because it would allow unnecessary downgrading of the water." (Furthermore, because there is little in the way of theory to define the bounds of the range, one may propose that standards should be developed probabilistically--for example, as discussed by M. B. Bayer, Proceedings of the 9th Canadian Symposium on Water Pollution Research, 1974.)

GAO EVALUATION COMMENT No.	PAGE(S)
14	95

4. Determination of correlative relationships.

Correlative relationships between characteristics of the water resource may be inferred from long records. (For example, see E. F. Gloyna, American Society of Civil Engineers, Journal of the Sanitary Engineering Division, 90 (127-151), 1964; and T. D. Steele and M. E. Jennings, Water Resources Research, 8, (460-477), 1972.) Such information has at least two important applications: (i) If examination indicates that similar relationships hold throughout the sampled space, such relations may be inferred to hold at ungaged sites, and where one type, but not the second type of data are available, the second may be inferred from the first. (ii) It is impossible to anticipate every "pollutant-of-the-week" and thus impossible to monitor "everything of interest", even if it were economically feasible. However, if a correlative relationship is found to exist between some water quality characteristic not previously considered and a routinely monitored one, it may be possible to infer some past history for the newly considered characteristic.

With respect to all of these points, because the methodology in NASQAN is consistent in time over the geographical space, it is possible to make comparisions among the sites and to infer what information may be transferable to non-monitored sites. It is a tenet of science that information at some level is transferable, and a task of science to discern the level of commonality. Thus, one of the goals of hydrology is to understand the similarities between rivers, even though one may conceive that in its specific details ". . . each river is an entity unto itself . . ." (David Rickert and Walter Hines, Geological Survey Circular 715-A, 1975) just as one instance of a random variable may differ from another.

This is best illustrated by example. Consider the Streeter-Phelps relationship which describes the interplay of atmospheric reaeration and the deoxygenation of polluted waters in producing the dissolved-oxygen profile along the path of flow in a stream. This relationship is generally applicable, i.e., transferable, although the specific values of the parameters of the relationship (e.g., rate of flow, reaeration rate, BOD load, etc.) must be determined for each specific situation.

It is emphasized that none of the activities described in this section--long term trend analysis, determination of probability distributions, determination of correlative relationships, standards construction, geographic transference of information--can be carried out on the basis of information obtainable from special studies, even intensive river quality assessments.

GAO EVALUATION

COMMENT No. PAGE(S)

COMPARISONS: SPECIAL STUDIES AND NASQAN

Before any comparison can be made, some semantic rigor is necessary. As used in the GAO Report, the term "special studies" refers to any activity, other than periodic monitoring at fixed stations over time, for the purpose of addressing some part cular water quality situation. A wide range of activities, varying in duration, intensity, purpose and expense is encompassed in this definition. For example, the $14,000 EPA study (cited in Chapter 4) describes a waste load assimilation study carried out for EPA by a state agency. The study entails two visits to a single site, each visit of 8 to 24 hours' duration, during which solar intensity is measured, possibly a small dye study is done, and determinations are made of diurnal DO, BOD, and nutrient levels (personal communication: Tom Murray, EPA, May 6, 1980). Through its Federal-State cooperative program, the Survey carries out studies of comparable scope and cost but they are never referred to as "intensive" as the GAO report has done.

The phrase "intensive river quality assessments" (IRQA), as used by the Geological Survey, refers to large (in terms of personnel and budget) special studies of multi-year duration, of a specific stream reach(es), generally for the purpose of addressing problems that have often been identified through information obtained from a regular periodic monitoring program. Synoptic studies are often employed in an IRQA. In an IRQA the objectives include identification of the cause of the perceived problems and exploration of possible solutions.

An intensive river quality assessment may also attempt to characterize the specifics of a particular hydrologic situation that is not well understood. As noted by GAO, such studies are well suited to explore cause/effect relations in the specific study areas, but only for effects that may be observed within the duration of the effort--at most several years. Long-term effects, for example climatic effects, are not observable in such a study. (This point was noted in the 1978 Report by the Comptroller General of the United States entitled "Water Quality Management Planning Is Not Comprehensive and May Not Be Effective for Many Years" where it was stated (on page 31) that "additional time [greater than the initial 2-year period for submitting 208 plans] may be needed to prepare plans, especially if the [208] agency needs water quality data and the data gathered during this period is not representative because of climatic or other conditions".) 15 95-96

The cost of intensive assessments, in terms of both personnel and budget requirements, is very high. (Furthermore, personnel demands are not only high but irregular, hence, additionally costly.) For example, for the period 1974 through 1976, the joint cost of the two California studies described in the GAO report exceeded one million dollars. (The two California studies--Redwoods and Russian River--were not part of the River Quality Assessment Program, but came under separate budget allocation.) For the same time period, the Willamette study cost $335,000. Furthermore, in all three studies, Survey personnel were assisted by State or other Federal agency personnel, not accounted for in the above cost figures. Adopting the time between

GAO EVALUATION

COMMENT No.	PAGE(S)
15	95-96

revisits of 10 years (the GAO report suggests every 5 to 10 years), a price in excess of $1.3 million appears extremely high to pay every 10 years to partially characterize three local situations, quite small with respect to the rest of the Nation. Furthermore, suppose a regular revisit rate of every 10 years is adopted. The ability to bridge any perceived differences in the water resource between these two time periods must still be obtained from long periodic records, such as those collected in the NASQAN program.

Table 1 summarizes the costs of the seven IRQA studies carried out or now in progress. Based on these figures, the cost of a typical IRQA study is approximately $0.75 million (1979 dollars) spent over 3 years (or $0.25 million per year).

In 1979, funding for the water-quality portion of NASQAN was $5.3 million. Since 518 stations were in the network in that year, the average annual cost per site was $10,300 approximately. This includes 12 site visits over the course of the year (not just during one hydrologic condition) and determinations of more than 60 water-quality characteristics, about 25 at a monthly frequency and the rest at a quarterly or semi-annual frequency. Although not directly applicable for waste-load assimilation work, the information available for the NASQAN site appears to provide a greater breadth of information for a general river-quality assessment than the aforementioned ($14,000) synoptic study. (Further cost information regarding NASQAN is presented in Table 2.)

If the present levels of funding for NASQAN were applied to Intensive River Quality Assessment on a 10-year restudy cycle, then it would be possible to conduct 21 such studies in any year (7 beginning in each year, running for 3 years each). The total number of sites nationally for which IRQA's could then be done at this funding level is 70 (as compared to 518 stations in NASQAN). In addition to the problem of the sparsity of coverage of such a program (70 reaches versus 518 sites), the GAO proposal does not explain how the results of the special studies would be unified to provide the kind of information sought in the NASQAN objectives.

DESIGN, REFINEMENT AND ENHANCEMENT OF DATA NETWORKS

Within the U.S. Geological Survey there are currently research efforts in the development and application of techniques for the quantitative design and refinement of hydrologic data networks. Heretofore, most such efforts, both within and outside the Survey, have been directed toward questions of water quantity (see M. Moss, Water Resources Research, 15 (1673-1676), 1979) primarily for three reasons: (1) historically, though not currently, public concern has been directed toward questions of quantity rather than quality; (2) only a few long consistent records existed for water-quality variables; and (3) water-quality problems have traditionally been cast either in taxonomic or in deterministic process terms, not amenable to statistical analysis.

Table 1. River Quality Assessments--Approximate allocated funding shown in thousands of dollars, current and (constant 1979)

FY	1973		1974		1975		1976		1977		1978		1979		1980	
Wilamette	90	(141)	157	(225)	156	(204)	22	(27)	4	(5)						
Chattahoochee					106	(139)	205	(254)	227	(226)	146	(159)				
Yampa					32	(42)	269	(334)	159	(186)	120	(131)				
Potomac											124	(135)	525	(525)	500	(475)
Apalachicola													291	(291)	300	(285)
Carson-Truckee													390	(390)	427	(406)
Schuylkill													360	(360)	350	(332)
TOTAL	90	(141)	157	(225)	294	(385)	496	(615)	390	(456)	390	(425)	1566	(1566)	1577	(1498)

Table 2. Comparison of NASQAN and River Quality Assessment Programs Since Their Inception.

Approximate allocated funding shown in thousands of dollars, current and (constant, 1979)

	NASQAN		RIVER QUALITY ASSESSMENTS	
FY	Number of Stations in Network	Funding	Funding	Number of Ongoing Assessments
1973	50	*	90 (141)	1
1974	100	486 (695)	157 (225)	1
1975	345	2224 (2913)	294 (385)	1
1976	345	2248 (2788)	496 (615)	3
1977	345	2420 (2831)	390 (456)	3
1978	445	3541 (3860)	390 (425)	3
1979	518	5295 (5295)	1566 (1566)	4
1980	518	5482 (5208)	1577 (1498)	4

* General program funds were alloted to NASQAN; not until 1974 was it a separate budget item.

	GAO EVALUATION COMMENT No.	PAGE(S)
NASQAN has been operated long enough (in 1980) to begin to offer moderately long records, consistent in space and time, as required for network design such that examination and refinement of the network could now be undertaken. The Survey is aware of many ways in which NASQAN may be enhanced to better achieve its objectives. It is within the plan for NASQAN to attend to such efforts, and this is, in fact, under way. Increased budgetary and manpower allocations would greatly expedite these efforts. Some of the possible types of enhancements to be considered are these: (1) change the frequency of sampling, (2) add new sites, (3) rotate some sites (that is, operate for a few years, cease operating for a few and then begin again), (4) delete some water-quality characteristics because of their high degree of correlation with other characteristics, (5) add some water-quality characteristics, (6) conduct special high-frequency sampling studies at NASQAN sites (to gain knowledge on variability and possible cycles in the data), (7) conduct a series of synoptic studies of river basins in concert with the sampling at the existing NASQAN fixed-station sampling. The Survey, with the aid of outside advisory groups, intends to do what the GAO has failed to do: examine the NASQAN objectives and the extent to which they are being met. The Survey is, furthermore, exploring ways that NASQAN can be enhanced to contribute more to an understanding of the causes of existing water quality.	16	96
Now that NASQAN has more than 5 years of record at more than 300 stations, there will be an increased effort to analyze the data and publish findings on such subjects as 1) the accounting of the movement of substances through the Nation's rivers, 2) the relationships between ambient water quality and the geologic, climatic and human characteristics of the river basin, 3) changes over time in water quality, 4) the relationships between water-quality characteristics, and 5) the effect of drought or long-term climatic change on water quality.	17	96
The Survey does not consider the NASQAN program to be "cast in concrete" but rather to be a program that should evolve through periodic review. The Survey is, however, committed to the concept of fixed-station, fixed-interval monitoring for the kind of national assessment function for which NASQAN is intended. NASQAN was created to fill a need, recognized widely in the late 1960's and throughout the 1970's: to have a consistent national data base for examining conditions and trends in river water quality. The GAO report does not question this need and does not provide a demonstrated alternative method of fulfilling it. The need for a nationally consistent data base continues to exist, and NASQAN should continue to fill that need.	18	97

GAO EVALUATION

COMMENT No. PAGE(S)

PART 2A

Comments on the Body of the GAO Report

[The Survey comments are listed under the headings (underlined) and paragraph numbers (underlined) used in the GAO report]

CHAPTER 1

IMPORTANCE OF WATER QUALITY INFORMATION

Paragraph 6:

	COMMENT No.	PAGE(S)
The cited reports are not specifically identified. The specific findings or recommendations should be given.	19	97

SCOPE OF REVIEW

Paragraph 4:

	COMMENT No.	PAGE(S)
The Geological Survey responded by letter January 15, 1979, with a list of 33 examples of the use of fixed station water quality information (see attachment 2 for the letter and abridged version of the list). The report examines only a part of one of these 33 examples (CEQ's trend analysis), and gives no indication of the variety of uses made of such data.	20	97-98

Paragraph 7:

	COMMENT No.	PAGE(S)
The report does not reference the work or opinions of experts within the academic community on the subject of water quality monitoring and data analysis such as: D.P. Lettenmaier (University of Washington), R.C. Ward (Colorado State University), M.G. Wolman (Johns Hopkins University).	21	98
The "...experts in the academic field..." that were contacted should be identified and their concurrence or disagreement with the recommendations of the report expressed.	22	98-99

CHAPTER 2

THE GEOLOGICAL SURVEY IS OBLIGATED TO PROVIDE RELIABLE AND MEANINGFUL WATER QUALITY DATA

After Paragraph 3:

	COMMENT No.	PAGE(S)
The objectives for NASQAN as expressed here are essentially those given by the Survey (Ficke and Hawkinson 1975). Nowhere in this report is it argued that these are inappropriate objectives. However, throughout the report NASQAN is repeatedly judged in light of different objectives (for	23	99

GAO EVALUATION

COMMENT No. PAGE(S)

example: description of variations in water quality over the length of a river, description of diel cycles, determination of causes of existing water quality or identification of waste load sources).

CHAPTER 3

FIXED-STATION NETWORK MONITORING HAS MANY WEAKNESSES

First Paragraph:

As the NASQAN station coverage became extensive enough to begin examining regional patterns, methods for describing water quality over space were developed and resulted in reports such as Briggs and Ficke (1977). More recently, the Survey has collaborated with the California Institute of Technology and Ohio State University for research on additional data interpretation techniques. Now that some 300 stations have more than five years of record, a project is underway in the Survey to develop statistically robust procedures for assessing trends in concentrations, transport, and concentrations adjusted for flow.

24 99

Paragraph 2: List of 4 points

Point 1. Referring back to the NASQAN objectives in Chapter 2: Note that providing a description of water quality at all points along a given river is not among the purposes of NASQAN. Rather, the intent is to account for the quality of the water moving into and out of the various accounting units. The approach selected for NASQAN to accomplish this objective was systematic sampling (at a fixed time interval and not influenced by the occurrence of specific types of hydrologic events). This approach was selected in accordance with methods for statistical analysis described by Rainwater and Avrett (1962) and elaborated on by Montgomery and Hart (1974).

25 99-100

The report contends that the approach taken will not produce "representative" measurements, but does not define "representative". The Survey takes the view that our measurements should characterize the unrestricted range of water-quality conditions occurring at a site. It is not the Survey's intent to charactrize only low flows, or only high flows, or only steady flows, or only one time of year, etc.

26 100

Points 2 & 3. The personnel involved in data collection receive formal training at the Survey's National Training Center and at the field offices. Their work is monitored and evaluated by District and Regional water quality specialists and by headquarters personnel. The procedures for sample collection, preservation, and shipment are standardized and documented, and the chemical analyses are carried out in Geological Survey's water quality laboratories. These efforts are intended to achieve and maintain a high level of precision and consistency nationwide. Mistakes will occur in any data collection program whether fixed station or intensive. The Geological Survey operates a quality

27 100

GAO EVALUATION

COMMENT No. PAGE(S)

assurance program to identify and resolve laboratory problems. The report makes no recommendations of ways to improve precision or lower the frequency of procedural errors.

Point 4. The Geological Survey's computerized water information system (WATSTORE) provides for remarks to be entered to indicate special circumstances (see attachment 1 for a list of such remarks). The WATSTORE system contains considerable additional "information on the variability of local conditions" such as continuous discharge, conductivity, and temperature data at most NASQAN sites. Improvements in WATSTORE are now underway to provide information on the precision of each individual analysis. At present one has to refer to tables describing precision for various types of measurements to determine the precision of any particular measurement.

28 101

Water quality experts have cited difficulties

Paragraph 1: ("Streamflow is one of the primary factors...")

It is out of a recognition of the important influence of streamflow that the Survey includes streamflow data in the NASQAN program, allowing those who interpret the data to separate out the effects of flow from the effects of other phenomena.

26 100

Paragraph 2: ("Clarence J. Velz...")

The points attributed to Velz have been taken out of context. The paraphrased statements are taken from the chapter entitled "The Stream Survey" which is prefaced with the following remark (Velz, 1970, p. 398): "In this chapter the term 'stream survey' is limited primarily to the collection of data essential to the rational method of stream analysis....There are many types of stream surveys, and the kind of data collected depends on the purpose in mind." The four points made here are made with respect to the objectives Velz is concerned with, namely, characterizing the waste-assimilation capacity of a specific river reach. This is not among the objectives of NASQAN and thus the points are not relevant.

29 101

Fourth point from Velz:

The importance of cross sectional variability of water quality is well known to the Geological Survey. The field procedure for NASQAN that is given in Quality of Water Branch Technical Memorandum No. 74.11 (which the Survey made available to GAO) is very clear on the technique to be used to collect a sample. The techniques used assure that the sample is appropriately integrated from the entire cross section. Continuous monitors for conductivity and temperature sample the flow at one location in the cross section. However, the monthly field data provide the means of determining the relationship between the measurements from the monitors and average cross sectional values.

30 102

GAO EVALUATION

COMMENT No. PAGE(S)

EPA and the Survey are aware of the problems

After Paragraph 3: (Quotation "...there is a growing number")

The quote from the Geological Survey Circular 715-D (Hines and others, 1976, p. D7) must be put in context. The statement was: "As indicated in the 'Introduction' there is a growing number of respected scientists and engineers who feel that programs based primarily on the monitoring-type approach will never prove efficient for trend or causal analysis. The sources referred to in the "Introduction" to the report by Hines and others included (1) Velz (1970) whose interests are quite different from the objectives given for NASQAN (see our discussion above); (2) Deininger (1974) who was specifically critical of programs that "are designed on an ad hoc, emergency type, basis"; and (3) Wolman (1971). The point made by Wolman in 1971 is historical in perspective; that data collection up to the present (1971) has not been conducive to the analysis of long-term trends. In no way does he argue that a data collection program conducive to such analysis is not possible. Consider the following statement from Wolman's paper (quoted in part by Hines and others, 1976):

> "Relatively few studies of the quality of the nation's rivers have been directed toward determining changes in specific parameters over long periods of time. This is perhaps not surprising because a number of disabilities interfere with a truly adequate statistical analysis of such a series. First, hydrologic records in the United States are relatively short. There are few continuous records for periods as long as 50 or 60 years. Second, techniques of observation and of analysis have changed over the years. Analytical techniques, in particular, have become more sophisticated, and routine measurements of exceedingly small quantities of contaminants are now possible which, only a few years ago, were considered impractical. Thus some comparisons reflecting changes in techniques of detection rather than real trends may be misleading. Third, changing the location or frequency of observations of water quality may distort the record. Observations of water quality are often made in the vicinity of metropolitan areas adjacent to the intakes of city water supplies. From time to time the intakes are moved to avoid sewer outfalls. While the intake may be moved upstream only a few hundred yards, the new record differs completely from the previous record, which was essentially monitoring the relation between the quality of the river and the inflow from the outfall. Fourth, adequate comparisons of specific variables related to water and to river quality require systematic correlation with hydrologic behavior. Such correlations are rarely available. Fifth, a knowledge of the 'natural background' or temporal variability of a given parameter is often essential in detecting and measuring a trend. Statistically, a trend cannot be discerned unless it is possible to discriminate between the variability of the phenomena as it might occur unaffected by the influences that one wishes to measure, in this case so-called pollution, and the variability normally associated with diurnal, annual, and significant secular climatic variations that occur in the hydrologic record over any period of time." (Wolman, 1971, page 905).

31 102-103

GAO EVALUATION

COMMENT No. PAGE(S)

This list of shortcomings of existing data describes well the shortcomings that NASQAN was designed to overcome.

The point raised by Hines and cited by the GAO report is correct in regard to certain types of causal analysis but is only a reflection of past experiences with regard to trend analysis. Hines goes on to say on page D9 "As described by Wolman (1971), many existing river-quality data have been examined for trends and cause-effect relations with disappointing results. We believe that many of these data can still be usefully interpreted, provided the data are amenable to segregation using river hydrology as the segregating tool." The NASQAN data, because they contain flow data, are specifically designed for undertaking the kind of hydrologic analysis that Hines refers to.

31 102-103

Paragraphs 3-5: ("Also, a memorandum on doubts...")

No reference is given to this memorandum. Its author is not identified and the arguments made in it are neither paraphrased nor quoted.

32 103

SAMPLE TIMING IS CRITICAL

Paragraph 3.

The quotation from Velz is not relevant to the objectives of NASQAN, as described in a previous section of the report.

33 103

Dissolved oxygen varies substantially

The Survey is well aware of the shortcomings of monthly sampling of dissolved oxygen (DO). It is one water-quality characteristic for which an empirical frequency distribution based on a series of monthly, daytime, field measurements cannot be expected to completely characterize the process. The options available to the Survey in designing the DO measurement program for NASQAN included: (1) Do not measure DO at all out of fear that the data would be misinterpreted; (2) measure DO monthly along with the regular sampling visits because measurements can be made inexpensively, given that the station would be visited anyway and the data are useful such as in providing information on the redox potential of the water which is a determinant of metal solubility; (3) station a hydrologist at the station for 24 hours a few times a year to make continuous measurements over the diel cycle; (4) set out a portable DO monitor and recording device for a 24-hour period a few times a year (this would double the travel costs for the NASQAN sampling program); (5) operate a continuous DO monitor year around at each NASQAN station (such installations require line power which is not presently available at many NASQAN sites, must be visited weekly or more frequently to insure reliable DO measurements, and would add a few thousand dollars per year to the cost of operating each station).

The Survey has selected the second of these five options. Options 3, 4 and 5 were judged too costly in terms of equipment, personnel, and travel costs, given that NASQAN sites are not necessarily in DO problem areas. It also

34 104

GAO EVALUATION

COMMENT No. PAGE(S)

appeared wasteful to fail to collect the data (option 1), given the minimal additional cost of option 2.

Paragraph after figure:

Two sentences in this paragraph are misleading: "Under the EPA and Survey national sampling program, whichever measurement was taken would be used to depict the DO concentration for the entire period represented by the sample--usually 1 month." "Under the NASQAN program the single reading of 10.5 mg/L represents the entire month of July 1977." This is not the intent of the Survey's sampling program. The entire distribution of a random variable is not to be inferred from a single value.

35 104

Many other water quality characteristics change substantially

The data presented are neither complete nor correct. An entire line of data is missing from the 1975 data, and one of the April 1975 values is incorrect. The table should read (based on our own WATSTORE retrieval)

36 104

Year of Sample	Sample Number	Dissolved Lead Concentration in µg/L April	July
1973	1	190	3
1974	1	3	6
	2	1	1
1975	1	5	15
	2	13	15
	3	6	9
1976	1	2	2
1977	1	18	4
1978	1	4	no value

Also, the report indicated a value of 82 µg/L for dissolved lead in July 1978. However, our quality assurance program identified a contamination problem with the acid ampules, supplied to the Survey by a contractor, and the suspect data were deleted from our files several months ago.

The report argues that there is no "discernible pattern" to the data. The Survey takes the view that all data exhibit some "pattern". This "pattern" can be decomposed into component parts (see Kendall, 1976). These may be described as (1) the trend-cycle component, or mean value function, (2) the seasonal component, and (3) the residual, irregular, or random component. The form of each of these parts is a part of the "pattern" of the data. The more data one collects, the better one may be able to characterize or describe this pattern. This third component may be correlated with discharge or some other variable (see Johnson and others, 1969, for example), it has a probability distribution (see Montgomery and Hart, 1974, for examples), and it has a

GAO EVALUATION COMMENT No.	PAGE(S)

serial correlation structure (see McMichael and Hunter, 1972; Lettenmaier, 1975; and D'Astous and Hipel, 1979). The characterization of this "pattern" for each constituent at each station and the comparison of these "patterns" between stations is fundamental to the objectives of NASQAN in that it lays "...the groundwork for future assessments of changes in stream quality" (Ficke and Hawkinson, 1975, p. 1.)

The report has not made clear what sort of patterns were being sought. The report appears to be suggesting that one should discontinue data collection because the authors of the report fail to discern a particular type of pattern. The Survey does not agree with this approach to data collection, and we seriously doubt that the statistical community or other Federal data service agencies would agree with this approach.

36 104

Paragraph after Table on Total Hardness

The report states that "The examples in this section illustrate the difficulty of comparing water quality." But it should be noted that the examples are all in accord with accepted statistical sampling practice, and the analysis of these data by accepted statistical techniques presents no particular difficulty. If one wishes to undertake a combined deterministic and statistical analysis of the data, there are considerable data on the "environmental influences present" to facilitate it (continuous discharge, conductivity, and temperature plus values for the other NASQAN constituents and a description of the NASQAN site).

37 105

After Paragraph 3: ("Rivers and their basins...")

The quotation from the Willamette River Quality Assessment is lacking a full citation, making it quite difficult for a reader to find the source and place it in its proper context (it is Rickert and Hines, 1975, page A-16). This quotation is a prescription for achieving the objectives of River Quality Assessment (as defined by Rickert and Hines). These objectives emphasize the evaluation of basin-development alternatives as they relate to the river-quality problems of the Willamette River basin. These objectives do not coincide, even in part, with the objectives of NASQAN as stated by Ficke and Hawkinson (1975). Thus, the criticism of NASQAN implied by this quotation has no basis.

It may be useful to consider the following analogy: An intensive survey such as the River Quality Assessment described by Rickert and Hines (1975) may be compared to the diagnostic and testing work of a physician (every patient must be viewed as an individual, and the physician should have some first-hand knowledge of the individual to do the job well). NASQAN, by contrast, may be compared to the collection and analysis of national health and mortality statistics. The statistical work is certainly no replacement for the detailed study of individuals, but it can be helpful to the physician and national policy makers. It can supply a baseline against which the physician can judge a patient's condition (e.g., just how abnormal is a pulse rate of 120 beats per minute?). It can illuminate relationships between variables (e.g., smoking and cancer) or trends (e.g., a change in respiratory disease rates since a new air pollution control device was installed on some

38 105-106

APPENDIX IX

GAO EVALUATION

COMMENT No. PAGE(S)

industrial plant) which any single physician may not notice or detect (given the small sample size). Thus it is useful: (1) to the physician who may decide to do more heart testing because of an aberrant pulse rate, or consider the patient's smoking history in deciding what tests to perform, or (2) to the policy maker who may use this evidence when weighing the costs and benefits of another air pollution mitigation action. In short, we all recognize the individuality of people (and rivers) but also recognize the need to gather statistical data on them to help us see relationships and broad patterns that the detailed work can never hope to show, and also to provide a basis for identifying aberrant conditions.

38 105-106

LOCATION OF SAMPLING STATIONS IS IMPORTANT

"EPA and the Survey" have "dissimilar" responsibilities and thus it is reasonable that they would choose "dissimilar monitoring locations." NASQAN was designed according to a well defined set of rules (see Ficke and Hawkinson, 1975), but other reasonable sets of rules are indeed possible (see, for example, Lettenmaier and Burges, 1977, or Liebetrau, 1979). The GAO has not offered any alternative approach to selecting station locations.

39 106

Water quality varies greatly throughout the length of rivers

This material on the South Platte River and on spatial variability is not relevant to NASQAN. If one has a real need to know about fecal coliform at Speer Blvd. in Denver (for example), then one must sample for it there. The development of this kind of information is not among the objectives of NASQAN.

40 106-107

Shore-to-shore variability occurs

See comments above - ("Fourth point from Velz")

30 102

Geological Survey criteria for station locations

"No specific patterns of water quality were targeted in the [NASQAN station] selection process." The intent of NASQAN was not to monitor polluted rivers, or clean rivers, or urban rivers, but to monitor the inflow and outflow of accounting units. If one were interested in the status and trends of a particular category of rivers, one could select data from NASQAN sites on such rivers for analysis.

41 107

SAMPLING ERRORS IN THE FIELD

There is no basis for the statement "...uncertainty is increased with far flung networks..." The Survey field personnel use the same methods, receive the same training, and are evaluated by common standards no matter where they are working.

The examples of problems of sampling practice "...uncovered..." in the Survey's district reviews must be viewed in the proper context. The purpose of these

42 107

APPENDIX IX

GAO EVALUATION

COMMENT No. PAGE(S)

reviews is to identify errors in procedure and prescribe the corrective action to be taken. The fact that the Survey seeks to identify errors and prescribe corrective action indicates its intent to achieve and maintain a high level of data reliability.

Delay During Storage and Shipment

The GAO report has noted a possible problem. The Survey intends to explore all of the steps in the process: sample collection, preservation, shipment, receiving, and analysis. The possibilities of changes in some or all of these steps will be considered, and their consequences for data reliability and cost will be considered.

43 108

INCONSISTENT LABORATORY PERFORMANCE

The report uses variability in analytical results among laboratories participating in the Geological Survey's Standard Reference Water Sample Program (SRWS) as a basis for casting doubt on the credibility of the Survey's water-quality data for NASQAN. However, the report's description of variability in SRWS analytical results is not pertinent to a discussion of analytical variance in NASQAN analyses because they are all performed at the Survey Central Laboratories.

A better evaluation of variability in NASQAN analytical data can be made by comparing analytical variability for standard samples analyzed within the Central Laboratories System as "blinds" (that is, samples that are not known to be "standards" by laboratory personnel) with the variability expected on the basis of comprehensive testing at similar concentrations by the American Society for Testing and Materials (ASTM Book of Standards, Part 31). Such a comparison is shown in the accompanying table for chloride determinations. Also included in the table is a comparison of mean values developed for the standard solutions through the SRWS Program with mean values developed through repeated analyses over time of the same standards within the Central Laboratories System. It can be seen that the concentrations compare favorably, and that the observed variations of the concentrations are comparable to the expected variations based on the ASTM testing program. The table also shows that the variability of the laboratory analyses has changed very little with time. Comparisons of quality-assurance data for the other determinations in question show similar results.

44 108

The GAO report makes no comparison of data reliability in fixed-station operations versus that in intensive studies. Thus, it is unclear how this argument on the reliability of the data has any bearing on the question that the report addresses: fixed-station monitoring versus special studies.

45 109

APPENDIX IX

GAO EVALUATION

COMMENT No.	PAGE(S)
44	108

Table Comparison of Analytical Results for Chloride Standard Solutions

Period of Record	Expected1/ Concentration (mg/L)	Central Lab.2/ Concentration (mg/L)	Central Lab.3/ Coefficient of Variation (± %)	ASTM4/ Coefficient of Variation (± %)
7/74-6/75	1.44 (39)	1.43 (71)	21	
1/77-12/79	1.71 (21)	1.84 (262)	17	24 *
1/75-6/76	1.84 (21)	1.57 (215)	25	
1/76-6/77	8.69 (26)	8.26 (167)	8	7 **
1/78-12/79	8.76 (25)	8.11 (152)	13	
7/74-6/75	22.9 (32)	22.5 (107)	4	
7/74-12/75	26.2 (20)	26.0 (156)	3	4 ***
6/79-12/79	27.9 (29)	27.9 (28)	4	
7/76-6/78	48.9 (24)	49.0 (88)	4	3 ****
7/74-6/76	72.2 (33)	72.7 (195)	2	
7/74-12/75	95.4 (32)	95.7 (141)	1	
1/76-12/76	122 (28)	124 (140)	4	3 *****
1/75-6/76	174 (23)	176 (228)	3	
7/74-12/75	213 (20)	213 (151)	3	
7/78-12/79	245 (33)	245 (64)	3	

1/ Consists of mean developed through Geological Survey's Standard Reference Water Sample Programs; outliers eliminated before calculation. Number in parenthesis is number of labs that participated in round-robin analysis.

2/ Mean of indicated number of determinations of the same standard solution as a "blind" sample; outliers not eliminated before calculation. Number in parenthesis is number of times standard solution was analyzed.

3/ (±1 standard deviation/mean concentration) · 100. Calculations based on number of determinations shown in parenthesis in column 2; outliers not eliminated before calculation.

4/ (±1 standard deviation/mean concentration) · 100. Calculations based on multi-laboratory determination of precision for ASTM methods; outliers eliminated before calculation.

* Coefficient of variation for ASTM method D512 at 2 mg/L
** Coefficient of variation for ASTM method D512 at 10 mg/L
*** Coefficient of variation for ASTM method D512 at 20 mg/L
**** Coefficient of variation for ASTM method D512 at 50 mg/L
***** Coefficient of variation for ASTM method D512 for greater than 50 mg/L

GAO EVALUATION COMMENT No.	PAGE(S)

<u>Comparing like to like is important</u>

If the Survey were to sample only during certain types of flow conditions or seasons, that sample would be useful for estimating characteristics of water quality for those flow conditions or seasons, but it may produce biased estimates of the overall characteristics of water quality. By sampling on a regular schedule which is unrelated to flow conditions, one gains information on the range of variation of water quality conditions. Assuming that the relevant information has been recorded (as in NASQAN), then, after several years of data collection, one may look at various sub-samples of the existing data (selected on the basis of discharge, change in discharge, season, temperature, or some other factor) and evaluate the relationships between water quality and those factors. Alternatively, one may perform analyses of the relationship between discharge and concentration (see Johnson and others, 1969) and then look for changes in this relationship over time. 46 109

The two measurements (July 12, 1976 and July 5, 1977) are of very little interest by themselves. Their usefulness arises when several more years of data have accumulated, leading to the development of an empirical probability distribution of the constituent concentrations, and of the relationship between these constituents and discharge. Then, after a number of years, one may explore the possibility that these distributions and relationships are undergoing change.

For example, the Survey has collected 24 years of monthly total hardness data for the Yakima River at Kiona. Discharge and total hardness are very closely related and concentrations may be described by the following regression equation:

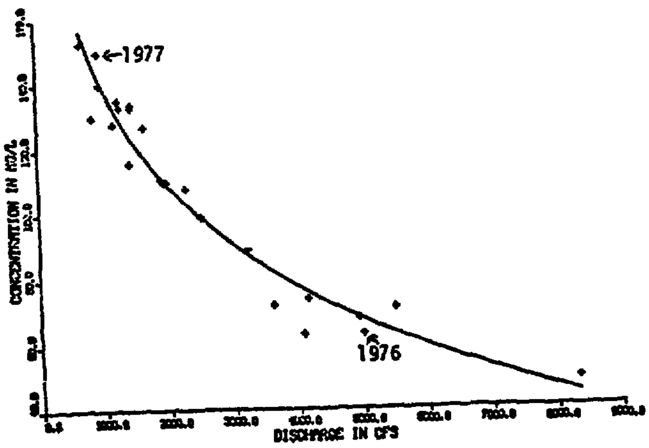
$$\hat{C} = 452.2 - 45.15 \ln Q$$

where $\hat{C}$ is the predicted concentration (mg/L) and Q is the discharge (cfs). The R^2 value for this regression is 0.94 (see figure). Given this analysis of the data, one may now pose the question of whether the 1976 or 1977 values are abnormal (that is: do the differences reflect something more than just the difference in flow). For the July 12, 1976, discharge, $\hat{C}$ is 67.6 mg/L and the measured concentration is 63 mg/L. For the July 5, 1977 discharge value, $\hat{C}$ is 142.2 mg/L and the measured concentration is 150 mg/L. The differences between the predicted and observed values are 4.6 mg/L in 1976 and -7.8 mg/L in 1977, and the standard error of estimate for an individual concentration value (from the regression) is ±7.3 mg/L. Thus, in the context of a long record of concentration and discharge measurements, one can see that neither the 1976 value nor the 1977 value is substantially different from what might be expected. In the absence of data collected from such a fixed location, fixed frequency network, it would not be possible to make this kind of interpretation, and one could only guess whether this more than doubling of total hardness in one year was an expression of some fundamental change in the basin or simply the normal consequence of the variations in discharge. 47 110

The statement "weakly supported conclusions...reached by relying on water quality measurement data without considering flow data..." is not relevant to the recommendation of the report. The Survey is in full agreement with GAO on the importance of flow data and that is why it is included in the NASQAN design and data base. 48 110-111

 APPENDIX IX

JULY TOTAL HARDNESS, YAKIMA RIVER AT KIONA, WA



Text	GAO EVALUATION COMMENT No.	PAGE(S)
<u>Mean values are not necessarily accurate indicators of water quality</u>		
The report states that "Annual means are more understandable if they are accompanied with clear information on the extent of variance among the individual measurements used to calculate the means." The Survey's latest report on NASQAN data (Briggs and Ficke, 19 7) gives sample size, standard deviation, and range, along with every mean value. Thus the data, and the Survey's reporting of it, conform to the GAO's suggestion.	49	111
For the nitrate plus nitrite data for 1975 for the Mississippi River at St. Paul, Minn., the mean is 1.088 mg/L, and the standard deviation is 1.704 mg/L (sample size 11). The report states that a "statistical test for variance revealed a standard deviation of 157 percent." This statement makes no sense, but one may take the meaning to be: the sample standard deviation was found to be 1.704 mg/L resulting in a sample coefficient of variation of 1.57. The report goes on to use these results incorrectly: stating that the likely range for the true mean extends up to 2.8 mg/L (note 2.8 = 1.088 + 1.704). The word "likely" is vague in this context but one may assume that the report is referring to the range encompassed by the sample mean plus or minus one standard error. The standard error of the mean is 1.704/ 11 = 0.514, so the "likely" range for the true mean is (0.574, 1.602). In the interest of clarity, the report should state the appropriate confidence interval implied by "likely" (68.3 percent in this case). In short, the authors have apparently assumed the standard error of the mean to be equal to the standard deviation when, in fact, it is the standard deviation divided by the square root of the sample size.	50	112
<u>Fecal coliform bacteria measures are unstable</u>		
The report indicates that a check was made on the "impact that fecal coliform variability can have on trend analysis" yet the GAO used an observation about the movement of annual means from category to category as a "trend analysis." There are many tests for trend available from the statistical literature. This ad hoc procedure used by GAO does not constitute a recognized statistical test and its sampling properties are unknown.	51	112
<u>INFORMATION SYSTEMS HAVE LIMITATIONS</u>		
WATSTORE contains parameter codes that can be used to describe conditions under which the samples were collected. Examples of these observed parameters are: algae, floating mats (severity); cloud cover (percent); detergent suds (severity); gas bubbles (severity), etc. Furthermore, the Survey is in the process of extending the capability of the WATSTORE system so that methodology and precision information can be included with the data.	52	112-113

GAO EVALUATION

Text	COMMENT No.	PAGE(S)
THE JAMES RIVER, VIRGINIA The purpose of this discussion is unclear. The report appears to be taking issue with CEQ's analysis and yet the recommendations of the report are concerned with the data collection program and not with the appropriateness of some particular use that has been made of the data. The other point that is raised in this section is that the available data (from NASQAN, NWQSS, and state monitoring activities) are not sufficient to explain the causes or identify the contributions to a particular water-quality measurement. The objectives of NASQAN are focused on description of water quality and not explanation of causes or identification of sources. Thus the fact that GAO finds "...four inexplicably high [zinc] measurements..." and finds that "Dissolved oxygen changes cannot be explained without more information" cannot be considered to be an expression of a shortcoming of NASQAN, based on the program's objectives.	53	113
PRACTICAL LIMITATIONS TO FIXING THE NETWORKS One of the objectives of NASQAN is to depict areal variability of stream quality. This demands a nationally consistent set of measurements, methods, and frequencies, rather than a program that varies from place to place. The report argues that "...if DO is an issue, DO should be measured around the clock, particularly when DO problems are most likely." The Survey agrees; if any particular constituent or characteristic of the water is a particularly important issue at some location (for example, if certain uses of the water are being impaired or major expenditures for abatement measures are being considered), then more intensive sampling, by the concerned party, may be called for. Once a specific management problem is identified, then a monitoring plan that is suited to that problem can be developed. The contribution that fixed-station, fixed-frequency data collection programs can make is not in the solving of site specific problems, but in providing background information to help guide an intensive data collection program to sample at the most appropriate times and places. In the previously cited report on the Willamette River Quality Assessment, Rickert and Hines (1975) point out the role that monitoring data plays in problem-solving studies, "Step 2 (collation and analysis of existing data) plays a vital role in data programs, even for those cases in which new data must be collected. For example, existing data were analyzed for study of DO depletion in the Willamette River. Although the records were not suitable for providing a reliable analysis of cause-effect relationships, the data did indicate (1) the general magnitude of DO depletion (up to 40 percent depletion of DO saturation), (2) the affected reaches of the river (the lower 80 miles, 129 km), (3) the yearly period of most severe DO depletion (July-August), and (4) the fact that summer flow was effectively steady state and are greatly augmented by reservoir releases. This information provided the background for a reconnaissance-level study."	54	113-114

	GAO EVALUATION COMMENT No.	PAGE(S)
Paragraph 6: ("The networks now sample once a month...") "...hundreds more samples might be needed for statistically significant data." Statistical significance (the report is assumed to be referring to significance of a test of a hypothesis concerning trend or change) can occur with virtually any amount of data (say four or more observations) Therefore, the GAO criticism is meaningless.	55	114
Paragraph 7: "...the networks [do] produce meaningful data." The report proposes "customized" studies tailored to each site. The Survey agrees with the need for such studies to enhance NASQAN. However, the objectives of NASQAN cannot be served by customized studies alone. There must be comparability, site-to-site and over time, if the objectives of NASQAN are to be met.	56	114
CONCLUSIONS		
Paragraph 4: "Comparability of data" is only a function of methods of data collection and analysis. The methods to be used in NASQAN are totally consistent nationwide and the execution of those methods is quite consistent, although not perfect. Consistency of data collection greatly facilitates the achievement of the objectives of NASQAN. What the report is recommending is to drop the present system, which considers consistency to be of paramount importance, and replace it with a collection of site specific studies that would differ in design from place to place (although the report does not define the temporal design of the proposed plan). The Survey has demonstrated in the work of Steele and others (1974) and Briggs and Ficke (1977) how the NASQAN data can be put to use to make comparisons of water quality over space and time. The report offers no examples nor does it describe how the proposed alternative program would facilitate comparisons of water quality over space and time (which are the objectives of NASQAN). As an example of the point about comparability, consider that the total hardness value for the Yakima River at a discharge of 5000 cfs is not directly comparable to one at 959 cfs (see previous discussion). But as part of a larger data set, including many flow conditions, they constitute a comparable set of values because the methods and sampling rules were consistent over time.	57	115
Paragraph 5: "...the networks do not link specific water quality conditions with what caused those conditions." Relating conditions to causes is not among the objectives of NASQAN. Once conditions have been characterized at a NASQAN site, then interested local, state, or federal agencies may mount a specific effort to track down the causes of any conditions that are impairing the uses of the water.' The purpose of NASQAN is to characterize water quality, not to explain the causes. Explanation is a logical next step after characterization.	58	115-116

	GAO EVALUATION COMMENT No.	PAGE(S)
The fact that this point arises so frequently in the report is an indication that a fundamental misunderstanding of the objectives of fixed-station sampling is at the heart of the GAO recommendation to eliminate the NASQAN network.		

CHAPTER 4

Paragraph 2:

The Survey concurs with the report on the value of intensive studies but would argue that the information is different in kind (as opposed to more or less "definitive and reliable"). The intensive studies offer a view of the local situation of water quality along a limited reach of a river but do not generally provide much information on temporal variability. The report argues that intensive studies can provide "progress reports" and yet none of the examples they give show this kind of use. It may be possible to produce useful progress reports from intensive studies but the methods have not, to our knowledge, been developed. Successive special studies using the same sites and sampling methods clearly constitute a fixed-station monitoring program. Before adopting this as the sole approach to reporting progress in water quality improvement, there should be a substantial demonstration of an acceptable methodology. 59 116-117

SPECIAL WATER QUALITY STUDIES ADDRESS SPECIFIC SITUATIONS

Paragraph 2:

At the present time it may appear that the Survey's analytical effort on water quality is concentrated in the area of intensive studies; this is largely because few long, consistent data sets have been available up to now. NASQAN has now been in existence long enough to provide a data base suitable for some types of analysis, and such work is now underway. (Note that the record is not nearly long enough yet for analysis of the effect of long-term climatic variation on water quality.) A study by Peters (Geological Survey) is examining the relationship of the transport of major ions at NASQAN stations to the geology and human population of the basin. The Survey's Water Resources Division has analyzed the relationship between discharge and concentration of dissolved solids and of total phosphorus at 308 NASQAN stations and tested for trends in concentration in transport, and in flow adjusted concentrations over the period 1972-1979. In the next few months NASQAN and other Survey network data will be examined for trends in pH, alkalinity, and sulfate in order to assess the impact of acid precipitation on the Nation's streams, and the dependence of these impacts on the general chemical character of the streams. The Survey's North Carolina District has already published reports (using NASQAN and other fixed-station data) on trends in water quality of the French Broad River (Daniel and others, 1979) and the Neuse River (Harned, 1980). Other examples of analysis of fixed-station data are given in the list that the Survey provided to GAO, attachment 2. 60 117

GAO EVALUATION

COMMENT No. PAGE(S)

Intensive river-quality assessments (IRQA)

In FY 1980 the Survey IRQA program (involving four studies) was funded at $1.6 million, as compared to $5.5 million for NASQAN. The Survey agrees with GAO as to the importance and value of the IRQA type of work. It should be noted, however, that the IRQA program is a demonstration program, designed to foster the more widespread use of that approach by state and local water management agencies and by Geological Survey District offices. It is worth commenting here on the relationship between the fixed station type of monitoring and intensive studies. Rickert and others (1976, p. C18-C19), in describing the Willamette study, observed that: "Attempts are often made to use monitoring and surveillance data for calibration and verification of applied DO models. Unfortunately, such data are usually poorly suited for these purposes." They go on to state:

> "In spite of the factors listed above, existing monitoring and surveillance data can often be useful for designing intensive studies of river-quality phenomena and for providing checks on modeling predictions. The utility of existing data for these purposes (assuming the appropriate variables were sampled) is determined by several conditions: (1) the period of record, (2) collection frequency, (3) location and number of sampling stations, and (4) the ease with which the data can be segregated and collated in relation to river hydrology. The segregation and collation are necessary to distinguish the effects of man from the natural variabilities in quality that result from temporal and spatial changes in hydrology. 61 118
>
> "Analysis of existing monitoring data for the Willamette River did provide significant insight into conditions surrounding DO depletion. The analyses indicated (1) the general magnitude of DO depletion..., (2) the affected reaches of the river..., (3) the yearly period of most severe DO depletion..., and (4) the fact that summer flow was effectively steady-state and greatly augmented by reservoir releases. This information provided the background for development of a reconnaissance-level study."

In addition Rickert (written communication, 1980), the Willamette Project Chief, stated: "The wealth of monitoring data available on the Willamette River was one reason the Willamette was selected for the prototype River-Quality Assessment. In retrospect, the existing data was instrumental to the ultimate success of the study." The point which the GAO report has completely overlooked is that these intensive studies depend on fixed-station networks for background information, for guidance in planning data collection, and for a statistical context in which to view the data collected in the intensive study.

GAO EVALUATION

COMMENT No. PAGE(S)

SHIFTING TO SPECIAL STUDIES NEED NOT INCREASE MONITORING COSTS

Paragraph 2:

Intensive water quality studies cover a wide range of levels of effort. It should be noted that the three IRQA studies thus far completed by the Survey (Willamette, Chattahoochee, and Yampa) averaged $590,000 each, and the four now underway are projected to cost an average of $890,000 each. The Redwood National Park study had a cost in excess of $840,000. The examples for which GAO has high praise are not the $14,000 type of studies that are mentioned in the report, but are the $0.25 million to $1 million.

62 118

OTHER INDICATORS OF PROGRESS TOWARD CLEANER WATER ARE AVAILABLE

The return of fish does not assure us that toxic materials are not present in concentrations sufficient to cause long-term pathological effects in fish or humans. Similarly, information on "specific pollution control actions" is not sufficient basis for evaluating progress toward cleaner water in the affected river (particularly in light of the large and poorly-known amounts of non-point source loadings in many rivers).

"Biological monitoring" can be useful, but at present only a few very specialized techniques are well developed and standardized (such as the Mussel Watch), and these pertain to only certain water uses.

63 119

CONCLUSIONS

The report finds "no compelling reasons why reports on the Nation's progress toward cleaner water have to be based on repetitive fixed-station sampling." The Survey's response to that statement is that the only practical and economical way to evaluate progress is by looking repetitively (one observation says nothing about progress) at fixed stations (if one measures in different places over time there is no comparability). The Geological Survey does not consider the present NASQAN program to be unchangeable. The Survey has and will consider changes in sampling frequency, station location, sampling methods or water quality characteristics measured. Any such changes will be considered in light of the NASQAN objectives. The GAO has not offered any suggestions on changes in the Survey's programs which would further the achievement of the objectives established for NASQAN.

64 119

RECOMMENDATIONS

The GAO has described the "special studies" it recommends but it has not explained how they could be organized to provide the kind of consistent national overview required by the Survey's mission and by the NASQAN objectives. Neither has the report provided any consideration of the costs of the "coordinated special studies" which would accomplish these objectives.

65 119-120

Part 2B

Comments on Appendix VI of the GAO Report a/

[The Survey comments are listed under the headings (underlined) and paragraph numbers (underlined) used in the GAO report]

BACKGROUND AND PURPOSE

The Council on Environmental Quality (CEQ) ninth annual report is referred to throughout Appendix VI. The Appendix includes a table of water quality changes at NASQAN stations, 1975-1977, from the CEQ report. The text from pages 96 and 98 of the CEQ report that accompany the table (identified as Table 2-1 in the CEQ report) put both the table and the use of the James River data into perspective. The text headed CHANGES from the CEQ report is:

> "Many of the pollution control facilities built during the past decade are just beginning to operate. Evaluation of their effectiveness requires good uniform data on plant performance and water quality. Fortunately, improved data networks are now providing the means for judging water quality changes, and they will continue to improve in the future.
>
> "So far, uniform water quality data exist for only 3 years, so it is premature to characterize trends definitively. But it is encouraging that bacteria levels improved through the third year.
>
> "Figure 2-5 shows fecal coliform levels from measurements at NASQAN stations during the 1975-77 water years. 'Violation rates' are the percentage of measurements in which concentrations of fecal coliform bacteria exceeded the recommended maximum for safe swimming, which many states and CEQ define as greater than 200 cells per 100 milliliters of water. (There is no legal uniform national standard; standards vary with water use and local laws and standards sometimes differ from nationally recommended criteria.) Patterns of improvement are apparent in several populous regions, particularly in the industrial urban belt south of the Great Lakes.
>
> "For other pollutants, no similar patterns of improvement are yet apparent. Levels of suspended material, nutrients, oil and grease, oxygen-demanding substances, and other materials should decline as pollution control becomes more effective. Nonpoint sources are largely responsible for some of these substances. The problems of controlling nonpoint source pollution are discussed later in this chapter.

---

a/See app. XI for GAO consultant's evaluation of agency comments on the James River case study.

> "Table 2-1 summarizes 3 years' data on stream accounting units that showed statistically significant change or lack of significant change tested at the 90 percent level, for 10 water quality characteristics. The data indicate that water quality, as measured by fecal coliform bacteria, oxygen, and zinc, has improved at 4-13 percent of the NASQAN stations and that more stations showed imprcvement than deterioration. For nitrogen, phosphorus, fecal streptococci bacteria, dissolved solids, and phytoplankton (algae), more stations showed deterioration than improvement. For all characteristics measured, most stations (74-94 percent) showed no statistically significant change."

It is important to note two points concerning the analysis of NASQAN data in the CEQ report:

(1) CEQ cautioned the reader that with only 3 years of data, it was premature to characterize trends definitively.

(2) No specific station is identified or otherwise singled out. Data from the James River were included with data from all other NASQAN stations to produce the table.

TOTAL DISSOLVED SOLIDS: CEQ'S TREND AND THE DROUGHT

Appendix VI confirms that the CEQ analyses of total dissolved solids (TDS) was correct; TDS increased during the 1975-1977 water years. Referring to CEQ's text, CEQ did not exclude weather conditions as causing change but reported conditions as they occurred. It is fortunate that the Survey has longer term data available which shows a decreasing trend in dissolved solids concentrations.

TOTAL DISSOLVED SOLIDS (TDS) AND SPECIFIC CONDUCTANCE: UNEXPLAINED INCONSISTENCIES

Specific conductance and dissolved solids concentration are closely related properties of water; however, the ratio of dissolved solids to specific conductance varies with the chemical composition of the water. The ratio has been reported as being generally between 0.54 and 0.96, with ratios above 0.75 usually coinciding with water high in sulfate or containing non-ionic materials such as organic compounds. (See J. Hem, "Study and Interpretation of the Chemical Characteristics of National Water", USGS Water Supply Paper 1473.)

The TDS determination is not one of the simplest of all water quality analyses. This determination, although straightforward, requires considerable skill by laboratory personnel. This is especially critical if the sample contains relatively low concentrations of dissolved material, which is the case for the James River at Cartersville.

The presence of organic compounds and their effect on the TDS/specific conductance ratio was apparently not addressed by GAO. It is not uncommon for the total organic carbon content in the James River at Cartersville to exceed 5 mg/L

APPENDIX IX

as C. Organic compounds often do not materially change the specific conductance of a sample, but they do, of course, contribute to the TDS. It is, therefore, possible to have a TDS/specific conductance ratio that is greater than one if the TDS are low and organic content is quite high. For the 1977 water year, the results from four samples contained both measured TDS and calculated TDS values. The range of calculated dissolved solids values were 5-12% lower than the measured TDS values. This strongly indicates that constituents other than inorganic salts are present and must be considered. Other factors, such as the hygroscopic properties of certain compounds, should not significantly affect the James River dissolved solids determinations because of the low dissolved solids concentration.

For water with low dissolved solids such as found in the James River, the laboratory method (Skougstad and others, 1979, p. 557) states a precision (expressed as a coefficient of variation) of 11 percent. Conditioned on the expected precision, dissolved solids values for duplicate samples from the James River are generally in agreement.

The report asks the question "But what assurance is there that reasonable-looking data aren't inaccurate too?" Data for the NASQAN program include three related values for each sample: field measurement of specific conductance; dissolved solids, residue on evaporation at 180°C; and dissolved solids sum of constituents. (The last value is calculated as the sum of the concentrations of the determined chemical constituents.) Close agreement between these independent measurements should provide assurance of the accuracy of the data within the stated limits of precision. As a routine procedure, the laboratory compares the value for dissolved solids at 180°C with the dissolved solids sum of constituents. If the values do not agree within the range of precision for both methods, the laboratory will rerun the sample for dissolved solids at 180°C. Also, beginning about 1976, the laboratory computer program through which the district offices receive their completed laboratory analyses has printed a warning on the analyses if the dissolved solids/field specific conductance ratio is not within acceptable limits. Clearly, there is considerable assurance that reported dissolved solids values are reasonable and accurate.

CEQ based their table and discussion on dissolved solids, not specific conductance. The facts that (1) most of the measurements on duplicate samples for dissolved solids are in agreement within stated laboratory precision, and (2) dissolved solids laboratory measurements are confirmed by calculated dissolved solids values in all NASQAN analyses, provide assurance that CEQ did work with reliable data. Dissolved solids data are <u>not</u> suspect.

Under current testing procedures, field personnel using field conductance meters are expected to produce values within ± 5 percent of the true value. The Survey was aware in 1975, and part of 1976, that the Virginia District was not consistently meeting this standard for specific conductance measurements. The District was directed to take corrective action and since 1976 the problem has been resolved.

The Survey informed GAO in April 1979 that the review of the Virginia District had identified a problem with specific conductance measurements made in that district during the period of interest. In discussions with GAO personnel, the Survey explained the importance and function of district reviews. Regional and Headquarters staff members visit district offices to ascertain, among other things, whether samples for water quality analyses are collected using prescribed techniques, whether samples are processed promptly, and whether field measurements are correctly made. Results of the reviews are transmitted to the district with specific recommendations for corrective action which must be taken to bring data collection to the required level of proficiency. GAO staff asked for and received copies of the latest reviews for California, Minnesota, Oklahoma, Utah, and Virginia. The Virginia review included specific recommendations for improvement of specific conductance measurements. Those improvements have been made.

RIVERFLOW PATTERNS MAY EXPLAIN CHANGES IN TOTAL ZINC

The use of the term "high zinc values" is misleading: the values are not high compared to nationwide occurrences or to water quality standards. The measured concentrations (90 µg/L was the highest) are so much less than the EPA public water supply criterion (5000 µg/L) as to be considered almost negligible. This is not to say that changes in values of zinc below water quality standards should be ignored. They may well indicate water-quality problems within the basin.

Appendix VI contains the following paragraph which includes several inaccuracies:

> "What CEQ took to be an improving trend in water quality might be nothing more than a change in sampling riverflow patterns. Most of the exceptionally high concentrations and loads coincided with dramatic changes in riverflow. These dramatic changes were especially common between December 1974 and June 1976. As luck would have it, the Survey rarely analyzed for total zinc during dramatic changes in riverflow after about September 1975. For example, in October 1976 the streamflow suddenly jumped from about 2,500 cfs to 70,000 cfs; but the Survey took no zinc samples in October 1976. Whenever the Survey analyzed for total zinc at or near a big peak in the riverflow record they found high values; but the Survey happened to miss most of the peaks after September 1975. The State analyzed for total zinc only once at or near a peak in the riverflow record. This trick of chance might explain the differences between the Survey's data and the State's data. But the explanation might be found elsewhere--errors in the laboratory or the computer center."

The sample collected June 1, 1976, was at a peak discharge of 16,100 cfs. The mean discharge for the previous day had been 9,060 cfs which certainly qualified the time of sampling as a time of dramatic change in streamflow. Though not as dramatic a change as from 2,500 cfs to 70,000 cfs, the June 1 streamflow was the highest discharge for the period during which zinc values were collected. It is not a "trick of chance" that certain events are missed but a matter of keeping a regular schedule. Over a period of time, many different hydrologic events will be sampled by following a regular schedule. Unless monitoring is continuous, which is impossible for most of the constituents of interest, some events will always be missed. Implicit in the decision to adopt a fixed monitoring schedule is the fact that some interesting events will not be included in the record. These are restrictions imposed by finite resources and the state of technology, as well as the objectives of the monitoring network which includes trend analysis.

RECORDS ON DISSOLVED ZINC RESIST RATIONAL ANALYSIS

The letter of November 21, 1979, from the Chief of the Survey's Quality of Water Branch to GAO puts forth a hypothesis that the data appear to support: that a non-point source of zinc existed in the basin between June 1974 and June 1975. The Survey maintains that it is plausible that there was some kind of material which contained zinc deposited in or near the stream system and it took approximately a year for the soluble zinc to flush from the system. The available data do not "prove" this hypothesis to be true, nor do they identify a specific source (they weren't intended to do so). Rather, they indicate that the elevated zinc values were episodic and provide information to develop a reasonable hypothesis concerning their origin. Data should not be dismissed as "resisting rational analysis" simply because the source producing the material in the stream is not known. Finally, if higher zinc values should be observed again in NASQAN data, the state or federal agencies may choose to look for the source of the zinc by sampling at locations higher up in the drainage network. If zinc is entering the James River from a point source, or localized non-point source, then it may be a problem in that stream reach. Once it has been established from monitoring data that zinc values are above the normal levels, and the flow conditions under which the high zinc concentrations occur have been identified, short-term special studies can then be performed to locate the zinc source.

TEMPERATURE INSTABILITY AT CARTERSVILLE

Temperature variations which occur on a daily cycle are found in every stream. The magnitude of the daily temperature variation is dependent on a number of factors including 1) surface area of the stream, because a shallow, wide stream gains heat more rapidly from the sun and air during daylight hours, and loses heat more rapidly at night than would a narrow, deep stream; 2) shape of the stream valley, because a stream in a narrow deep canyon may receive much less sun than a stream in a broad valley; 3) shading by vegetation; 4) season of the year; and 5) weather conditions. It should be noted in Table 9 of the Appendix that the Survey technician consistently visited the James River near the start of the work day. Temperatures will normally be lower at that time of day. The state employee was usually at the site in early to mid-afternoon, when the stream temperature would be approaching a maximum.

APPENDIX IX

The Appendix notes that "round the clock" readings of temperature must be made to define the temperature regime. Such readings are available for the Potomac River at Great Falls, Maryland, which is a somewhat larger river (drainage area is about 11,500 square miles as compared to 6,257 square miles for the James River at Cartersville) and is farther north. Generally, the larger the river, the less variation between maximum and minimum temperature during a single day. However, it is interesting to compare the data in Table 9 with the maximum and minimum temperature in the Potomac for the same period

		Table 9 Appendix VI			From WATSTORE	
		USGS	State	USGS daily	Potomac River at Great Falls Temperature	
Date	Time	Technician (°C)	Employee (°C)	Observer (°C)	Maximum (°C)	Minimum (°C)
Feb. 10, 1975	0930	3.0				
Feb. 10, 1975				3.5	3.0	2.0
Feb. 11, 1975	1420		6.7			
Feb. 11, 1975				5.0	3.0	2.0
May 1, 1975	0830	14.5				
May 1, 1975				14.5	13.5	13.0
May 2, 1975	1450		17.8			
May 2, 1975				15.5	15.0	13.0
Aug. 25, 1975	0800	26.0				
Aug. 25, 1975				25.0	28.5	25.5
Aug. 26, 1975	1115		28.9			
Aug. 26, 1975				27.0	30.0	27.0
Dec. 29, 1975	1415		6.1			
Dec. 29, 1975				2.5	2.5	2.0
Dec. 30, 1975	0830	5.0				
Dec. 30, 1975				2.5	2.5	2.0
Feb. 22, 1977	0800	3.5	8.0			
Feb. 22, 1977	1400					
Feb. 22, 1977					3.5	1.0
Feb. 23, 1977					5.5	1.0
June 1, 1977	1200	23.0				
June 1, 1977					25.0	23.0
June 2, 1977	1350		2.8[sic]			
June 2, 1977					25.5	24.5

The temperature measurements for the James River are within the normal range expected of diel temperature variations, keeping in mind that the James River site is about 100 miles south of the Potomac site and that the Potomac drains areas to the north in Maryland and Pennsylvania. One obvious error is the State's reported value of 2.8°C on June 2, 1977. The temperatures which are called "unstable" in Appendix VI are simply normal diurnal temperature changes.

INCONSISTENCIES IN DISSOLVED OXYGEN

Throughout the discussion of dissolved oxygen data from the James River site, it is important to keep in mind that none of the dissolved oxygen values collected by the Survey gave any indication of a problem in meeting the minimum value of 4.0 mg/L or the daily average of 5.0 mg/L dissolved oxygen set by the State standards. The lowest measured value (1974-1977) was 6.1 mg/L on July 8, 1974, at 8:15 a.m.

The author of Appendix VI appears to recognize and list the factors which must be considered in evaluating dissolved oxygen and in fact states "all these factors must be properly understood." However, in the subsequent discussion conclusions are drawn from an analysis of dissolved oxygen which ignores most of the factors. The failure of this analysis to provide explanations of the variations in DO is attributed (by the report) to inconsistency in the data rather than any failure of the analysis.

A large number of constituents and physical parameters have been measured at the James River site which can aid in interpreting dissolved oxygen values. These include time (hour of the day), temperature, biochemical oxygen demand (BOD), chemical oxygen demand (COD), total organic carbon (TOC), suspended sediment, turbidity, nitrogen and phosphorus species, chlorophyll a or b, and phytoplankton and periphyton abundance and type. All these constituents are important in determining whether measured dissolved oxygen values are reasonable or indicative of a problem.

Much of the "discrepancy" in table 10 is explained by time of day and water temperature. The two State samples were collected in the early afternoon which is, as Appendix VI points out, the period for maximum biological aeration. The Survey sample was collected in the morning when biological aeration had not reached a maximum. Water temperature also has an effect on the production of oxygen by green plants. The rate of oxygen production is less at 10.5°C, the temperature at the time the Survey made the measurement, than at 16.7° or 17.8°C when the State made its measurements.

Lower water temperatures have the effect of decreasing the rate of oxygen production by algae in the stream and increases the amount of dissolved oxygen that can be dissolved in the water. Higher temperatures have the opposite effect.

The discussion following Table 11 includes a statement from the November 21, 1979 letter from the Chief, Quality of Water Branch which says:

> "Weather information and sediment concentrations are available for assisting in interpretation of the DO data. Conditions for the two November samples were clear weather, low sediment concentrations and low flow. The temperature of the sample taken on November 24 was higher and temperature-dependent photosynthesis activity could account for the higher DO value. On December 9, the flow was high from a period of storms, the weather was cloudy, and sediment concentration had increased to 131 mg/L. Lower temperature (compared to November), higher water stage with increased sediment concentration covering periphytic algae and cloudy skies reducing available sunlight would diminish photosynthetic activity and reduce the amount of oxygen added to the water. Also, runoff following storms generally has a higher BOD load than the water entering the stream during a low flow period (usually ground water). This oxygen demand reduces the amount of dissolved oxygen in the stream."

The Survey maintains that this is valid interpretation of the data presented in Table 11. The chlorophyll a and b data will be discussed later in this reply. It is appropriate to state that, from data obtained over time at that site, there is good reason to expect to find periphyton and phytoplankton in the stream.

Table 12 data do not show inconsistencies. On two dates, May 3 and 17, 1976, both morning measurements made on cloudy days, the percent DO saturation was 84.5 and 76.7 respectively. On May 14, 1976, in the afternoon, the dissolved oxygen was at 103 percent saturation. Given the flow, the type of stream, and the time of year, these values are consistent with what one would expect.

The discussion of DO presented in Appendix VI entirely ignores the objectives of NASQAN (to describe water quality) and imposes another set of objectives (the complete understanding of the causes of each observed DO value). The fact that the GAO's consultant could not explain the observed DO values by the means of analysis he chose does not constitute a valid criticism of the NASQAN program or data.

Discrepancies Related to Biological Reaeration

The discussion following Table 13 is somewhat misleading. It is important to keep in mind that the DO concentrations on cloudy mornings when DO should not be at a maximum were 84.5 and 76.7 percent of saturation and in very close agreement. The lowest DO value measured was 6.1 mg/L, well above the state standard.

The statement, "the Survey should be encouraged to delete this misleading measurement (phytoplankton, total cells per milliliter) from its list of tests" is puzzling. Contrary to the Appendix VI contention, cell count per milliliter is a standard measurement used by biologists and is widely reported in the hydrologic literature. As has been pointed out to GAO heretofore, algae are identified and counted at the genus level. These data are available from the Survey and have been published in the Virginia data reports since 1978. Taxonomic data cannot be included in the present Survey WATSTORE system or EPA STORET data storage and retrieval system because of the hierarchical nature of taxonomic data. A separate storage and retrieval system presently contains these data. The total count of phytoplankton cells in a milliliter of water can be stored in both systems. Storage of such data in WATSTORE or STORET would alert users that taxonomic data are also available.

Following Table 14 is a quote from the November 21, 1979, letter to GAO from the Chief of the Survey's Quality of Water Branch. To put everything in proper context, the entire quote should have been included as follows:

"Most chlorophyll samples in Table 14 were collected at a different time than the cell counts and therefore should not be directly compared. However, cell counts can be used as estimates of the expected chlorophyll values by using the following table:

Phytoplankton	Chlorophyll Content per cell (10^{-12}g/cell)	Chlorophyll Concentration at 1000 cells/ml (ug/L)
Diatoms	0.5-5.0	0.5-5.0
Green Algae	0.1-1.0	0.1-1.0
Blue Green Algae	0.1-1.0	0.1-1.0

(Eppley and Sloane, Phys. Plantarum, V. 19, pg. 47-59.)

The highest cell count from Table 14 (4,100 cells per ml) could be expected to produce a maximum chlorophyll value of 20 ug/L if the sample were composed entirely of diatoms containing the largest amounts of the chlorophyll. Other combinations of algae would produce less than 5 ug/L. The chlorophyll samples were analyzed by method B-6501-77 (TWRI, Book 5, Chapter A4, page 209) which calls for use of acetone to extract the chlorophyll from the cells. Acetone is not particularly efficient in extracting chlorophyll; therefore the lowest value for which an estimate of precision is made is 5 ug/L. This is also the lowest detection level. Comparing the number of cells per ml given in Table 14 with the expected chlorophyll values from the table above, it is obvious that Table 14 is not a table of impossible values but a table of values below the detection limit of the method. Note that new laboratory methods which supercede method B-6501-77 have been developed and have even lower limits of detection. Data determined by the newer methods are placed in the data storage systems under different parameter codes."

Several points should be clarified about the chlorophyll determinations:

(1) Chlorophyll is not a constituent sampled within the NASQAN program. It was included at the James River station to complement other work.

(2) GAO was informed that a more sensitive method had been developed for chlorophyll determinations and that the old method had been superseded.

(3) Method B-6501-77, the old method, does call for grinding the filter. The complete method was provided to GAO in TWRI, Book 5, Chapter A4, page 209.

Please note that the Survey cautioned GAO about using cell counts and chlorophyll data collected at different times. The June 2, 1975, data which shows 2,900 algae cells and 55 µg/L chlorophyll a were collected at different times. Appendix VI implies that the Eppley and Sloane data are incorrect, yet it does not provide any reference to work which would indicate that it is incorrect. Obviously, there was a difference in number of cells between the two sampling times.

The Survey agrees that all samples for chemical analysis at a given site should be collected at one time. In the report, "Technical Review of Virginia District water quality activities, November 9-12, 1976," the District was told to collect all samples at the same time and to discontinue the practice of collecting related chemical and biologic constituents at different times. GAO was provided a copy of this review. Data for the 1977 water year and the subsequent period have been taken as recommended.

It should be noted that in the letter quoted above, the table heading clearly stated "(10^{-12} g/cell)"; the report mis-quotes it to be "(10-12g/cell) [sic]."

Discrepancies Between State and Survey DO Data

Appendix VI refers to Table 17 and states "the Survey's values for DO and percent saturation are often much lower than the State's." It is important to note that the Survey was consistently at the James River from early to mid-morning. The state was usually at the site in mid to late afternoon. Dissolved oxygen varies during the day as has been previously pointed out. When the State and the Survey visited the site at the same time of day, the results were much closer. Note the August 26, 1975, visit by the State and the September 8, 1975, visit by the Survey. The flow in the river was constant and both samples were collected near 11 a.m. Dissolved oxygen concentration was 7.2 and 7.7 mg/l respectively and percent saturation was 92 for both.

CONCLUSIONS

The Survey does not agree with the conclusion that the James River cannot be "meaningfully described" by the present sampling schedules. The reasons for disagreeing are given in the comments to the main part of the GAO report.

1. Dissolved Solids and Conductance

   Appendix VI affirms that CEQ did observe a statistically significant increase in dissolved solids. The text of the CEQ report does not exclude drought as a reason for a trend.

   Analyses of duplicate samples for dissolved solids are generally within the stated limits of precision for the method. No data were included in Appendix VI to show that TDS samples "often disagreed widely."

   The Survey was aware of problems with specific conductance data in 1976 and provided information to GAO showing that the Virginia District was told to take corrective action to improve their specific conductance measurements. There is no reason to suspect that TDS data are not correct and, in fact, there are considerable supporting data to show that the values are within the stated limits of precision for the method.

2. Total Zinc

   It is not a "trick of chance" or "luck" that a certain event will be missed but implicit in the decision to adopt a fixed monitoring frequency. Over a long period, many different hydrologic conditions will be sampled with a regular schedule. Unless monitoring is continuous, which is impossible for most of the constituents of interest, some events will always be missed. There are restrictions imposed by finite resources and the state of technology as well as the objectives of the monitoring network.

   Again, Appendix VI affirms that CEQ correctly showed a decreasing trend with the data they used. The text of the CEQ report qualified the use of the results by saying that with 3 years of data, "it is premature to characterize trends definitively."

3. Dissolved Zinc

   Supportable conclusions can be drawn from the dissolved zinc data. Dissolved zinc values were well below water-quality limits given in various standards and criteria. Dissolved zinc cannot be considered a problem based on the standards and criteria at the James River site. The data show that a considerable amount of zinc was in the river system from June 1974 to June 1975. The source of zinc is not known, but it suggested that if a similar pattern of dissolved zinc is observed, it might be worthwhile to conduct a more detailed investigation. Unusual transient events such as the relatively short-term elevation of zinc concentrations are more likely to be detected in a regular long-term monitoring program than during an intensive short-term sampling program.

4. Temperature
The temperature record appears to show normal diel temperature variations. Available records do not show any potential violations of State water quality standards.

5. Dissolved Oxygen

To understand the dissolved oxygen conditions at the James River site (that is, to be able to predict and model DO for the reach of river above Cartersville), one would have to collect much additional data. However, to justify the expenditure of time and money required to obtain these data, there would have to be some indication that dissolved oxygen is a problem. At the Cartersville site, none of the dissolved oxygen values collected by the Survey give any indication of a problem in meeting the minimum standard value of 4.0 mg/L or the daily average of 5.0 mg/L set by the State. The lowest value (1974-1977) was 6.1 mg/L on July 8, 1974, at 8:15 a.m. It is important that all available data be used in deciding if there is a dissolved oxygen problem at the Cartersville site.

APPENDIX IX

ATTACHMENT 1.

PARAMETER CODES FOR WATER DATA

IN ALAPHABETICAL ORDER

PHYSICAL PARAMETERS (OBSERVED)

CODE	PARAMETER
01325	ALGAE, FLOATING MATS (SEVERITY)
00032	CLOUD COVER (PERCENT)
01345	DEBRIS, FLOATING (SEVERITY)
01305	DETERGENT SUDS (SEVERITY)
01340	FISH, DEAD (SEVERITY)
01320	GARBAGE, FLOATING (SEVERITY)
01310	GAS BUBBLES (SEVERITY)
01355	ICE COVER, FLOATING OR SOLID (SEVERITY)
00022	LENGTH OF EXPOSURE (DAYS)
70971	LIGHT, ATTENUATION COEFFICIENT (ALPHA/M)
00074	LIGHT TRANSMISSION, 1 METER PATHLENGTH (PERCENT)
01330	ODOR, ATMOSPHERIC (SEVERITY)
01300	OIL-GREASE (SEVERITY)
01335	SEWAGE SOLIDS, FRESH, FLOATING (SEVERITY)
01315	SLUDGE, FLOATING (SEVERITY)
01351	STREAMFLOW (SEVERITY)
	VALUES FOR PARAMETER CODE 01351: (1=DRY, 2=LOW, 3=NORMAL, 4=FLOOD, 5=ABOVE NORMAL)
01350	TURBIDITY (SEVERITY)

APPENDIX IX

CODE PARAMETER

VALUES FOR SEVERITY: 0-NONE, 1-MILD, 2-MODERATE, 3-SERIOUS, 4-EXTREME

00041 WEATHER

VALUES FOR PARAMETER CODE 00041:

00 - CLOUDLESS
01 - PARTLY CLOUDY
02 - CLOUDY
03 - OVERCAST
10 - PRECIPITATION WITHIN SIGHT
13 - UGLY, THREATENING SKY
40 - FOG
50 - DRIZZLE
51 - SLIGHT DRIZZLE, INTERMITTENT
52 - SLIGHT DRIZZLE, CONTINUOUS
53 - MODERATE DRIZZLE, INTERMITTENT
54 - MODERATE DRIZZLE; CONTINUOUS
55 - THICK DRIZZLE, INTERMITTENT
56 - THICK DRIZZLE, CONTINUOUS
57 - DRIZZLE AND FOG
58 - SLIGHT OR MODERATE DRIZZLE AND RAIN
59 - THICK DRIZZLE AND RAIN
60 - RAIN
61 - SLIGHT RAIN, INTERMITTENT
62 - SLIGHT RAIN. CONTINUOUS
63 - MODERATE RAIN, INTERMITTENT
64 - MODERATE RAIN. CONTINUOUS
65 - HEAVY RAIN, INTERMITTENT
66 - HEAVY RAIN. CONTINUOUS
67 - RAIN AND FOG
68 - SLIGHT OR MODERATE MIXED RAIN AND SNOW
69 - HEAVY MIXED RAIN AND SNOW
70 - SNOW OR SLEET
71 - SLIGHT SNOW IN FLAKES, INTERMITTENT
72 - SLIGHT SNOW IN FLAKES, CONTINUOUS
73 - MODERATE SNOW IN FLAKES. INTERMITTENT
74 - MODERATE SNOW IN FLAKES. CONTINUOUS
75 - HEAVY SNOW IN FLAKES. INTERMITTENT
76 - HEAVY SNOW IN FLAKES. CONTINUOUS
77 - SNOW AND FOG
78 - GRANULAR SNOW (FROZEN DRIZZLE)
79 - ICE CRYSTALS
80 - SHOWER(S)

APPENDIX IX

CODE	PARAMETER
	VALUES FOR PARAMETER CODE 00041 - CONTINUED:
	81 - SLIGHT OR MODERATE RAIN SHOWER(S) 82 - HEAVY RAIN SHOWER(S) 83 - SLIGHT OR MODERATE SNOW SHOWER(S) 84 - HEAVY SNOW SHOWER(S) 85 - SLIGHT OR MODERATE RAIN AND SNOW SHOWER(S) 86 - HEAVY RAIN AND SNOW SHOWER(S) 87 - GRANULAR SNOW SHOWER(S) 88 - SLIGHT OR MODERATE HAIL OR RAIN AND HAIL SHOWER(S) 89 - HEAVY HAIL OR RAIN AND HAIL SHOWER(S) 90 - THUNDERSTORM 93 - SLIGHT THUNDERSTORM WITH RAIN OR SNOW 94 - SLIGHT THUNDERSTORM WITH HAIL 95 - MODERATE THUNDERSTORM WITH RAIN OR SNOW 96 - MODERATE THUNDERSTORM WITH HAIL 97 - HEAVY THUNDERSTORM WITH RAIN OR SNOW 99 - HEAVY THUNDERSTORM WITH HAIL
	PHYSICAL PARAMETERS (MEASURED)
00042	ALTITUDE ABOVE MEAN SEA LEVEL (TEST)
72027	AZIMUTH FROM OUTLET (DEGREES)
72028	AZIMUTH FROM SOUTHERNMOST POINT (DEGREES)
00025	BAROMETRIC PRESSURE (MM OF HG)
70969	BATTERY VOLTAGE (VOLTS)
00080	COLOR (PLATINUM COBALT UNITS)
00081	COLOR, TOTAL (PLATINUM - COBALT UNITS)
00009	CROSS-SECTION LOCATION, FEET FROM LEFT BANK LOOKING DOWNSTREAM
00001	CROSS-SECTION LOCATION, FEET FROM RIGHT BANK LOOKING UPSTREAM
00002	CROSS-SECTION LOCATION, PERCENT FROM RIGHT BANK LOOKING UPSTREAM
00003	CROSS-SECTION LOCATION, VERTICAL (FEET FROM SURFACE)

APPENDIX IX

APPENDIX IX

CODE	PARAMETER
00005	CROSS-SECTION LOCATION. VERTICAL (PERCENT OF TOTAL DEPTH)
71820	DENSITY (GM/ML AT 20 DEG C)
72019	DEPTH BELOW LAND SURFACE (WATER LEVEL) (FEET)
00003	DEPTH OF COLLECTION (FEET)
72001	DEPTH OF HOLE, TOTAL (FEET)
72025	DEPTH OF RESERVOIR (FEET)
00064	DEPTH OF STREAM, MEAN (FEET)
72008	DEPTH OF WELL, TOTAL (FEET)
72016	DEPTH TO BOTTOM OF SAMPLE INTERVAL (FEET BELOW LSD)
72003	DEPTH TO BOTTOM OF WATER-BEARING ZONE SAMPLED (FEET)
72015	DEPTH TO TOP OF SAMPLE INTERVAL (FEET BELOW LSD)
72002	DEPTH TO TOP OF WATER-BEARING ZONE SAMPLED (FEET)
00061	DISCHARGE, INSTANTANEOUS STREAM (CUBIC FEET PER SECOND)
00060	DISCHARGE, STREAM (CUBIC FEET PER SECOND)
* 99998	DISCHARGE, STREAM (MILLIONS OF GALLONS PER DAY) (00060)
72029	DISTANCE FROM OUTLET OR SOUTHERNMOST POINT (FEET)
81024	DRAINAGE AREA (SQUARE MILES)
81025	DRAINAGE AREA, CONTRIBUTING (SQUARE MILES)
72020	ELEVATION (FEET NGVD)
X 90000	ELEVATION (INCHES)

X - TEMPORARY CODE TO BE USED WITH THE DAILY VALUES FILE ONLY.

* - CODE TO BE USED AS INPUT TO THE DAILY VALUES FILE ONLY. THE INPUT DATA WILL BE CONVERTED TO THE PROPER UNITS AND STORED WITH THE CODE SHOWN IN PARENTHESES.

ATTACHMENT 2.

January 15, 1979

In Reply Refer To:
Mail Stop 412

Mr. David L. Jones
Assistant Director
Community and Economic Development Division
U.S. General Accounting Office
Washington, D.C. 20548

Dear Mr. Jones:

In response to your letter of December 26, 1978, I am pleased to provide examples of the use (and usefulness) of water quality data collected at fixed stations (see attachment).

My staff contacted several Survey field offices to assemble the examples provided here. Our responses from the field were somewhat uneven geographically because many of our senior staff who were most familiar with the uses of data we collect were on leave during the holiday period. Nevertheless we received good responses and we located many examples, especially from Georgia, California, and the Rocky Mountain states.

I am providing abstracts describing 33 examples, considerably more than the five or six you requested, because I was reluctant to try to represent the many uses of fixed-station data with so few examples. We, of course, will be pleased to provide further details for any of these examples, or for additional types of examples, according to your needs and desires.

There are two points that I hope you will keep in mind as you read over the examples contained in the attachment to this letter:

(1) Programs of the Geological Survey are extremely broad. We collect a wide variety of information for a wide variety of users, and many of our data are collected for the purpose of satisfying multiple uses.

(2) Many of our programs, especially at the local level, incorporate intensive surveys in conjunction with fixed stations. For these studies, we consider the fixed stations to be highly useful for identifying variability of water quality <u>with time</u> but to be of limited usefulness for describing variability <u>in space</u> (geographically, within the basin), while intensive surveys are best for defining variability of water

quality in space but, because they are synoptic, rather poor for defining variability with time. Thus, data from long-term fixed stations provide the most help in defining mean values, trends, or cyclic patterns in water quality, while data from intensive surveys are more useful for identifying causes and effects.

The examples given in the attachment are organized according to level of information, using the Office of Water Data Coordination's scheme for classifying data collection according to function and level. We have provided information on this type of classification previously to Mr. Peterson of GAO's Seattle office and to Mr. Edmonson and other members of your Washington staff.

I trust these brief descriptions will meet your present need for examples of the use of water-quality data from fixed stations. Please let us know if you wish further examples or additional detail on the examples described above.

Sincerely yours,

for J. S. Cragwall, Jr.
Chief Hydrologist

Attachment

 APPENDIX IX

Attachment
Prepared by Water Resources Division
January 16, 1979

EXAMPLES OF USE OF DATA FROM FIXED STATIONS

LEVEL I: NATIONAL AND REGIONAL

Annual reports of conditions and trends of U.S. rivers.--The seventh, eighth, and ninth annual reports of the President's Council on Environmental Quality (CEQ) for 1976, 77, and 78 have described conditions and trends in the quality of U.S. rivers using data from the National Stream Quality Accounting Network (NASQAN). CEQ reports have included maps shaded according to NASQAN data for bacteria, major and trace chemicals, sediment, and algae, as well as graphical and tabular presentations of data to show changes in river quality. Discussions of probable causes of conditions and trends have been included in the reports also. In addition, CEQ has used other fixed station data, such as long-term USGS data on pesticides from 60 stations in Texas, Louisiana, and Oklahoma, and data on phenols in the Ohio River basin from the Ohio River Valley Water Sanitation Commission (ORSANCO) and from STORET (EPA), to describe conditions and trends for those water-quality characteristics.

Ref: U.S. Council on Environmental Quality, 1976-78, Environmental quality, 1976, 77, 78--The seventh, eighth, and ninth annual reports of the Council on Environmental Quality: Washington, D.C., U.S. Govt. Printing Office.

Trends at NASQAN stations before 1972.--An analysis of data from approximately 80 NASQAN stations having at least 6 years of data collected before 1972 revealed significant time trends in a variety of stream temperature characteristics at 15 of the 80 stations studied. Significant trends also were found in the long-term chemical quality records at 15 of 88 stations analyzed. Some of the observed changes in temperature could be attributed to the construction and operation of reservoirs on such rivers as the Gunnison, Bighorn, Yellowstone, Colorado, Boise, and Snake Rivers. Degradation in chemical quality observed in 10 streams was attributed to mine drainage, increased irrigation, and construction of reservoirs. Improvement of water quality at 5 sites was attributed mainly to pollution abatement measures. Several instances of improvement due to abatement were noted in the Arkansas River basin. Details of the trend analyses are reported by Steele and others (1974); the work was summarized in the 1975 Annual Report of the U.S. Council on Environmental Quality (1975).

Ref: Steele, T. D., Gilroy, E. J., and Hawkinson, R. O., 1974, An assessment of areal and temporal variations in streamflow quality using selected data from the National Stream Quality Accounting Network: U.S. Geological Survey Open-File Report 74-217, 210 p.

U.S. Council on Environmental Quality, 1975, Environmental quality, 1975--Sixth annual report of the Council on Environmental Quality: Washington, D.C., U.S. Government Printing Office, 763 p.

ORSANCO network.--The Ohio River Valley Water Sanitation Commission (ORSANCO) operates a long-term program to evaluate the effects of the pollution cleanup of the river that began in 1948. They began with 11 stations in 1951; today they have more than 50 stations (Cleary, 1978). ORSANCO publishes a monthly bulletin describing water-quality conditions on the Ohio River and major tributaries. These bulletins report violations of standards and criteria, short-term (seasonal) changes, and summaries of long-term conditions and trends for selected sites along the river and its tributaries.

Ref: Cleary, Edward J., 1978, Perspective on river-quality diagnosis: Water Pollution Control Fed. Jour., v. 50, no. 5, May 1978, p. 825-832.

Analysis of salinity and other water-quality aspects of the Colorado River Basin.--Data from 17 stations in the Colorado River basin are used to summarize the severity of problems of salinity, to keep track of changes resulting from new irrigation projects and other developments, and to predict future flow depletions and salinity effects that will result from these new activities.

Ref: U.S. Dept. of the Interior, 1963-77, Quality of water, Colorado River basin: Washington, D.C., U.S. Dept. of the Interior, progress reports 1-8.

Sediment to the Oceans.--Suspended-sediment discharge data obtained from fixed stations near the mouths of 27 drainage areas during the period 1950-69 were used to estimate the sediment contributed to the oceans from the conterminous United States. Work was done by the Geological Survey as part of a UNESCO-sponsored project of the International Hydrological Decade called the "World Water Balance".

Ref: Curtis, W. F., Culbertson, J. K., and Chase, E. B., 1973, Fluvial-sediment discharge to the oceans from the conterminous United States: U.S. Geol. Survey Circ. 670, 17 p.

Dissolved solids to the oceans.--Dissolved-solids data from downstream sites in 54 river basins for the period 1966-69 were used to compute the amount of dissolved materials contributed to the oceans.

Ref: Leifeste, Donald K., 1974, Dissolved solids discharge to the oceans from the conterminous United States: U.S. Geol. Survey Cir. 695, 7 p.

Information system on water for energy programs.--Data on the quality of water for approximately 2,000 sites are being assembled for a data system describing the availabiity of water for the Nation's energy program. Using the facilities of the USGS National Water Data Exchange (NAWDEX) and STORET (EPA), data have been assembled from USGS, USEPA, Army Corps of Engineers, and Canada Inland Water Directorate sites.

Ref: Files of USGS NAWDEX office and personal communication from Mr. J. C. Sonnichsen, Hanford Engineering Development Laboratory, P.O. Box 1970, Richland, WA 99352.

Research on hypothermia in boating fatalities.--Temperature data from about 300 stations operated in 1974 by Federal, State, and local agencies were identified through the indexes of the National Water Data Exchange (NAWDEX) for use by university researchers in determining the causal involvement of hypothermia in boating fatalities. Researchers report that even though station data may not be available for the exact site and date of the boating fatalities, they have in almost every case been able to interpolate from data at nearby sites and approximately the same time.

Ref: Files of USGS NAWDEX office and personal communication from Dr. R. Michael Harnett, Clemson University College of Engineering, Clemson, SC 29631.

Water quality at Hydrologic Bench Marks.--Data collected by USGS at 57 hydrologic bench-mark stations in 37 states during the 1938-70 water years were used to define water quality in the "natural" environment. Relationships were developed between dissolved-solids concentration and discharge per unit basin area for various physical divisions of the United States. Concentrations of most major ions, trace metals, and pesticide compounds that occur at bench-mark stations are quite low when compared with concentrations in the major rivers of the country. One exception to this generalization was Bear Den Creek near Mandaree, ND, which had a dissolved-solids concentration of 3,420 mg/L.

Ref: Biesecker, J. E., and Leifeste, D. K., 1975, Water quality of hydrologic bench marks--an indicator of water quality in the natural environment: U.S. Geol. Survey Circ. 460-E, 21 p.

LEVEL II: SUBREGIONAL

Conditions and trends in the Potomac River basin.--Data covering a 10-year period from more than 100 stations in the Potomac River basin were used to assess conditions and trends. A complex pattern of change was described which included worsening of conditions in the North Branch headwaters and major tributaries, in the South Branch, and in the 20 miles of main stem above the estuary and the tributaries entering in that reach. Conditions were found to be holding steady in the main stem of the river from Cumberland, MD, to Great Falls, MD, and improving in the Shenandoah River and in the reach of the river below Washington,D.C., that extends from Blue Plains to Maryland Point.

Ref: Mason, W. T., Jr., Palmer, R. N., Sheer, D. P., and Combs, B. J., 1975, Potomac River basin water quality status and trend assessment 1962-1973: Bethesda, Md., Interstate Comm. on the Potomac River Basin, 161 p.

Temperature of streams in Georgia.--Stream temperature data collected periodically at 146 stations on streams in Georgia have been used to compute probable average temperatures for every day of the year. Results, along with streamflow and dissolved oxygen data, are used by the Georgia pollution-control agency to determine or estimate pollution loading values.

Ref: Dyar, T. R., and Stokes, W. R., 1973, Water temperature of Georgia streams: U.S. Geol. Survey report for Ga. Dept. Natural Resources, Env. Prot. Div., 317 p.

Study of mine drainage in Colorado.--An investigation to determine the effects of mine drainage was made using intensive surveys of 982 stream sites in areas of ore deposits and coal regions of Colorado. Thirteen control sites were also used, including two long-term stations from the Geological Survey's Hydrologic Bench Mark Network. The Bench Mark stations provided reference information that was used in interpreting the data from the surveys.

Ref: Wentz, D. A., 1974, Effect of mine drainage on the quality of streams in Colorado, 1971-72: Denver, Colorado Water Cons. Bd., Colorado Wat. Res. Circ. 21, 117 p., 3 pl.

Cobb, E. D., and Biesecker, J. E., 1971, The National Hydrologic Bench Mark Network: U.S. Geol. Survey Circ. 460-D, 38 p.

Effects of restoration on control of sedimentation.--Data for the Schuylkill River at Berne, PA (USGS Station 01470500), and at Philadelphia, PA (station 01474500), for the period 1948 to present have been used to demonstrate the effects of restoration work done in the period 1950-56 to control transport and deposition of sediment in the river.

Ref: Biesecker, J. E., Lescinsky, J. B., and Wood, C. R., 1968, Water resources of the Schuylkill River basin: Harrisburg, Pennsylvania Dept. Forests and Waters Bull. 3, p. 93-101.

U.S Geol. Survey, 1978, Water resources data for Pennsylvania water year 1977, v. 1 Delaware River basin: Harrisburg, Pa., U.S. Geol. Survey, p. 169-177.

Regional analysis of effects of land use.--Data from 80 stations in the Susquehanna River basin of Pennsylvania and New York for the 10-year period 1966-75 were used to assess the statistical relationship between water quality and several factors of climate, physiography, and land use. Seventeen water-quality characteristics studied represented annual mean concentrations or calculated annual yields of suspended sediment, dissolved solids, and various chemical species of nitrogen and phosphorus. Multiple-linear regressions were developed to estimate quality at specific sites or to simulate the ranges of background water quality. For example, present nitrate yields are as much as 20 times greater than simulated background yields. This condition is believed to be a result of pollution by animal wastes, application of chemical fertilizers, and increasing urbanization. Land-use variables affected by human activities and economic development had measurable impacts in all 14 of the usable regression functions.

Ref: Lystrom, D. J., Rinella, F. A., Rickert, D. A., and Zimmerman, Lisa, 1978, Regional analysis of the effects of land use on stream-water quality, methodology and application in the Susquehanna River basin, Pennsylvania and New York: U.S. Geol. Survey Water-Resources Invest. 78-12, 60 p.

Assessment of impacts of coal development.--The Geological Survey's river quality assessment of the impacts of coal development on the Yampa River basin of Colorado and Wyoming used a program of reconnaissance surveys tied to several long-term fixed stations. Data from 2 NASQAN stations and several other fixed stations provided controls for the reconnaissance data by characterizing seasonal patterns, establishing a historical base, and defining statistical relationships among hydrologic variables.

Ref: Steele, T. D., Bauer, D. P, Wentz, D. A., and Warner, J. W., 1976, An environmental assessment of impacts of coal development on the water resources of the Yampa River basin, Colorado and Wyoming--Phase I work plan: Lakewood, CO, U.S. Geol. Survey open-file report 76-36, 17 p.

Quality of the lower Mississippi River.--The water quality of the lower Mississippi River was assessed using both fixed-station data and data collected through intensive surveys. The results are being used by the Louisiana Department of Public Works and others as the basis for defining conditions and preparing water-management plans for that part of the river.

Ref: Everett, Duane E., 1971, Hydrologic and quality characteristics of the lower Mississippi River: Baton Rouge, La. Dept. Public Works Tech. Rept. 5, 48 p.

Summary of water quality in streams of Michigan.--Summaries of water-quality conditions that are prepared by the State of Michigan to meet the reporting requirements of section 305(b) of the Clean Water Act rely heavily on data from a fixed-station network that is designed to describe the status and trends of the quality of Michigan's streams. Current data are used to represent how well waters of Michigan meet criteria and standards; data for the previous 10 years are used to show trends in a number of chemical and biological factors and in a composite water-quality index.

Ref: Michigan Dept. of Nat. Res., 1977, Water quality and pollution control in Michigan: Lansing, Mich. Dept. Nat. Res., Env. Prot. Bur., 83 p.

APPENDIX IX

LEVEL III: LOCAL

Effects of water diversions in Louisiana.--In an assessment of the Atchafalaya River basin, fixed-station water-quality data were used to detemine past and potential effects of water diversions upon the aquatic biota of the stream.

Ref: Wells, F. C., and Demas, C. R., 1977, Hydrology and water quality of the Atchafalaya River basin: Baton Rouge, La. Dept. Transportation and Development, Office of Public Works, Water Res. Tech. Rept. 14, 53 p.

 APPENDIX IX

REFERENCES

Bayer, M. B., 1974, Applying probablistic water quality standards in river basin water quality optimization models: Proceedings of the 9th Canadian Symposium on Water Pollution Research, p. 25-29.

Briggs, J. C. and Ficke, J. F., 1977, Quality of rivers of the United States, 1975 water year--based on the National Stream Quality Accounting Network (NASQAN): U.S. Geological Survey Open-File Report 78-200, 436 p.

Cleary, E. J., 1978, Perspective on river-quality diagnosis: Journal Water Pollution Control Federation, v. 50, no. 5, p. 825-832.

Daniel, C. C., Wilder, H. B., and Weiner, M. S., 1979, Water quality of the French Broad River, North Carolina, an analysis of data collected at Marshall, 1958-77: U.S. Geological Survey Water Resources Investigations 79-87, 53 p.

D'Astous, F., and Hipel, K. W., 1979, Analyzing environmental time series: Journal of The Environmental Engineering Division, American Society of Civil Engineers, October 1979, p. 979-992.

Deininger, R. A., 1974, Optimization of water quality monitoring programs: Unpublished paper presented at the International Symposium on Environmental Quality Monitoring, Division of Environmental Chemistry, American Chemical Society, Los Angeles, California, April 2-4, 1974.

Edwards, A. M. C. and Thornes, J. B., 1973, Annual cycle in river water quality: A time series approach: Water Resources Research, v. 9, no. 5, p. 1286-1295

Enviro Control, 1972, National assessment of trends in water quality: U.S. Department of Commerce, National Technical Information Service, PB-210 669, 52 p.

Eppley, R. W. and Sloan, P. R., 1966, Growth rates of marine phytoplankton: Correlation with light absorption by cell chlorophyll a : Physiologia Plantarum, v. 19, p. 47-59.

Ficke, J. F., and Hawkinson, R. O., 1975, The National Stream Quality Accounting Network (NASQAN)--some questions and answers: U.S. Geological Survey Circular 719, 23 p.

Gloyna, E. F., 1964, Predictive techniques for water quality inorganics: American Society of Civil Engineers, Journal of the Sanitary Engineering Division, v. 90, p. 127-151.

Greeson, P. E., and others, editors, 1977, Methods for collection and analysis of aquatic biological and microbiological samples: Techniques of Water-Resources Investigations of the United States Geological Survey, Book 5, Chapter A4, 332 p.

Harned, D. A., 1980, Water quality of the Neuse River, North Carolina--variability, pollution loads, and long-term trends: U.S. Geological Survey Water Resources Investigations 80-36, 88 p.

Hem, J. D., 1959, Study and interpretation of the chemical characteristics of natural water: U.S. Geological Survey Water Supply Paper 1473, 269 p.

Hines, W. G., Rickert, D. A., and McKenzie, S. W., 1976, Hydrologic analysis and river-quality data programs: U.S. Geological Survey Circular 715-D, p. 7.

Johnson, N. M., and others, 1969, A working model for the variation in stream water chemistry at the Hubbard Brook Experimental Forest, New Hampshire: Water Resources Research, v. 5, no. 6, p. 1353-1363.

Kendall, M., 1976, Time-series: Charles Griffin & Co., Ltd., London, 197 p.

Lettenmaier, D. P., and Burgess, S. J., 1977, Design of trend monitoring networks: Journal of The Environmental Engineering Division, American Society of Civil Engineers, October 1977, p. 785-802.

Liebetrau, A. M., 1979, Water quality sampling: some statistical considerations: Water Resources Research, v. 15, no. 6, p. 1717-1725.

McMichael, F. C., and Hunter, J. S., 1972, Stochastic Modeling of temperature and flow in rivers: Water Resources Research, v. 8, no. 1, p. 87-98.

Montgomery, H. A. C., and Hart, I. C., 1974, The design of sampling programmes for rivers and effluents: Water Pollution Control, Metropolitan and Southern Branch, p. 77-101.

Moss, M. E., 1979, Some basic considerations in the design of hydrologic data networks: Water Resources Research, v. 15, no. 6, p. 1673-1676.

Pomeroy, R. D., and Orlob, G. T., 1967, Problems of setting standards and of surveillance for water quality control: State Water Quality Control Board, State of California, publication no. 36, 123 p.

Rainwater, F. H. and Avrett, J. R., 1962, Error inference in systematic-sample statistics in stream quality studies: American Water Works Association Journal, v. 54, p. 757-768.

Rickert, D. A., and Hines, W. G., 1975, A practical framework for river-quality assessment: U.S. Geological Survey Circular 715-A, p. 11.

Rickert, D. A., Hines, W. G., and McKenzie, S. W., 1976, Project development and data programs for assessing the quality of the Willamette River, Oregon: U.S. Geological Survey Circular 715-C.

Simpson, E. A., and Hinchcliffe, P. R., 1980, National water quality monitoring: World Health Organization (WHO) Water Quality Bulletin, v. 5, no. 2, p. 43-48.

Skougstad, Marvin W., and others, editors, 1979, Methods for determination of inorganic substances in water and fluvial sediments: Techniques of Water-Resources Investigations of the United States Geological Survey, Book 5, Chapter A1, p. 626.

Steele, T. D., Gilroy, E. J., and Hawkinson, R. O., 1974, An assessment of areal and temporal variations in streamflow quality using selected data from the National Stream Quality Accounting Network: U.S. Geological Survey Open-File Report 74-217, 210 p.

Steele, T. D., and Jennings, M. E., 1972, Regional analysis of streamflow chemical quality in Texas: Water Resources Research, v. 8, no. 2, p. 460-477.

Velz, C. J., 1970, Applied stream sanitation: Wiley-Interscience, John Wiley & Sons, 619 p.

Wolman, W. G., 1971, The nation's rivers: Science, v. 174, no. 4112, p. 905-918.

SELECTED ADDITIONAL REFERENCES

Bradley, J. V., 1968, Distribution-free statistical tests: Prentice-Hall, 388 p.

Conover, W. J., 1971, Practical nonparametric statistics: John Wiley & Sons, 462 p.

Dawdy, D. R., Kubik, H. E., and Close, E. R., 1970, Value of streamflow data for project design--a pilot study: Water Resources Research, v. 6, no. 4, p. 1045-1050.

Kendall, M., 1975, Rank correlation methods: Charles Griffin & Co., Ltd., London, 202 p.

Landwehr, J. M., 1978, Some properties of the geometric mean and its use in water quality standards: Water Resources Research, v. 14, no. 3, p. 467-473.

Larson, H. J., 1974, "Introduction to probability theory and statistical inference (2nd ed.): John Wiley & Sons, 430 p.

Lettenmaier, D. P., 1975, Design of monitoring systems for detection of trends in stream quality: University of Washington Technical Report No. 39, Department of Civil Engineering, 203 p.

Moss, M. E., and Karlinger, M. R., 1974, Surface water network design by regression-analysis simulation: Water Resources Research, v. 10, no. 3, P. 427-433.

Mosteller, F., and Tukey, J. W., 1977, Data analysis and regression: Addison-Wesley Publishing Co., 588 p.

Snedecor, G. W., and Cochran, W. G., 1967, Statistical methods, sixth edition: The Iowa State University Press, 593 p.

Ward, R. C., and Vanderholm, D. H., 1973, Cost-effectiveness methodologies for data acquisition in water quality management: Water Resources Research, v. 9, no. 3, p. 536-545.

APPENDIX IX

GAO EVALUATION

COMMENT No. PAGE(S)

Office of Water Research and Technology

REVIEW COMMENTS

GENERAL COMMENTS:

Although the document addresses a number of deficiencies in the surface water quality data-collection network systems of the Environmental Protection Agency and the U.S. Geological Survey (USGS), Water Resources Division (WRD), the majority of the allegations and the case histories are highly exaggerated and out of context. The NASQAN (National Stream Quality Accounting Network) and other surface water sampling under the guidance and responsibility of WRD are a valuable tool in providing baseline hydrological data and are used nationally by a large number of State, Federal, and municipal agencies as well as the private and industrial sectors in preparing environmental reports, decision-making process, and to further strengthen the understanding of local and regional hydrological conditions. The discontinuation of the NASQAN Stations would jeopardize national water resource efforts at a critical time in the management of the national water resource efforts.

SPECIFIC COMMENTS:

1. The national hydrological data provided by WRD are widely used. For example, various regulatory agencies utilize WRD surface water data in energy related activities in the West to prepare environmental documents and to evaluate environmental impacts resulting from such energy-related activities.

1 121

2. Page 3, section 3, paragraph 2, summarizes the scope of the General Accounting Office (GAO) document. However, the Ground Water/Surface Water interrelationships are not adequately addressed. Ground water is frequently the major surface water source. 2 121

3. Mr. Jerome Horowitz has been contracted by the GAO as a consultant. It would be most desirable to have Mr. Horowitz's credentials and qualifications addressed in this document. 3 122

4. A number of illustrations have been provided in this document to support GAO's allegations. These data can lead to multiple interpretations, and the GAO has generally only given a single interpretation. 4 122

5. The case histories and examples which are addressed in appendix VI are exaggerated. The behavior of zinc in natural waters is not only complex geochemically but also analytically. Stable background conditions for zinc are extremely difficult to achieve owing to zinc contamination. Therefore, zinc is not an ideal water quality parameter to use as an example to discredit the WRD effort. 5 122

6. The document accurately points out that WRD is responsible for the management of nearly 16,000 surface water stations including a large number of NASQAN sites. It is quite natural therefore that some discrepancy will arise. But this document only emphasizes a few isolated instances of error. A more comprehensive evaluation of the national surface water networks 6 123

APPENDIX IX

	GAO EVALUATION COMMENT No.	PAGE(S)
7. The Office of Water Research and Technology is currently analyzing the question of data-intensive models. OWRT's concern is whether or not the data collection efforts mandated by the law and the current thinking of modelers in response to the law produces a more costly data collection effort than the present one. These additional costs reduce the availability of information needed.	7	123
8. A number of analysts have considered the problem of data networks, sampling intervals, and other topics relevant to the analysis and conclusions presented in the subject report. It is difficult to tell whether this information was used inasmuch as it is not cited.	8	123-124

GAO EVALUATION OF

DEPARTMENT OF THE INTERIOR COMMENTS

1. Survey Comment

NASQAN never was intended to be a source of information detailed enough to assess the effectiveness of pollution control measures on a localized basis, yet GAO has judged NASQAN by this inappropriate criteria. The primary objectives of NASQAN are to (1) account for the quantity and quality of water moving within and from the United States, (2) depict aerial variability, (3) detect changes in stream quality, and (4) lay the groundwork for future assessments of changes in stream quality. Also, NASQAN records are indispensable for many purposes of policy analysis, water management and hydraulic research, including: (a) long-term trend analysis; (b) construction of probability distributions; (c) development of water quality standards; (d) determination of correlative relationships; and (e) determination of spatial transferability of information.

GAO Evaluation

We do not agree that we judged NASQAN based on inappropriate criteria. Because of inherent weaknesses in the NASQAN network--time and location bias, infrequent sampling, inconsistencies, and lack of consideration of variance and significance in water quality--we believe the network does not meet its established objectives. We have discussed these weaknesses in connection with EPA's comment 2, Survey comment 2, and CEQ comment 24.

Furthermore, if NASQAN was not intended to be a source of detailed information for the sampling locations and cannot provide reliable information on water quality at those locations, we do not believe that statistical manipulation of NASQAN data can portray water quality conditions over entire river basins or the entire nation.

We also do not agree that NASQAN data can or should be used for the five other purposes cited by the Survey. The same weaknesses of networks that preclude meaningful, reliable national assessments of water quality also weaken seriously the other purposes cited by the Survey. A few additional comments on the purposes cited by the Survey are appropriate.

Long-term trend analysis. The report discussed at length the use of network data for such purposes (see pp. 29-39, vol. I) and demonstrates that NASQAN data are unsuitable for this purpose because of inherent weaknesses in the networks.

Construction of probability distributions. A probability distribution assumes a stable generating function, such as a pair of dice, that will generate a stationary time series. The Survey recognized this requirement in its comments on our report. (See p. 31 of this appendix.)

> "Construction of (empirical) probability distributions. The distribution of a water quality characteristic of interest can be developed only if a (trend free, i.e., stationary) consistent record of some length is available. * * *"

But, rivers are not stable generating functions and water quality data certainly are not consistent. Rivers can be dramatically changed by suburban sprawl, new industries, shifts in pesticide use, improvements in pollution control, and shifts in weather. Water quality data are affected by these changes. Inconsistencies in network water quality data are common. Infrequent sampling produces data from dissimilar conditions. Changes and errors in field or laboratory procedures make past data inconsistent with future data. In short, historical network water quality data is a poor base for constructing probability distributions.

Development of water quality standards. We fail to see how NASQAN water quality data can have an important role in development of water quality standards. Water quality criteria published by EPA have had presumptive validity and States have had to justify, on a case-by-case basis, any deviation from these criteria. Many of the water quality properties at issue can change rapidly (such as dissolved oxygen and ammonia). NASQAN sampling cannot satisfy the technical requirements for deviations from EPA criteria. NASQAN stations are too scattered and sampling done at NASQAN stations is too infrequent to provide the data needed to properly evaluate rapidly changing properties of water.

Determination of correlative relationships. Misleading relationships can be drawn from water quality data that are from heterogenous conditions or that are unreliable for other reasons, such as questionable data produced through improper field procedures or analysis of stale samples. As we have demonstrated in the report, these weaknesses and others affect network data.

Spatial transferability of information. Use of network data for this purpose overlooks the basic concept of water quality--that each river is unique in hydrology, man's impacts, and in its water quality responses. Rapidly changing properties cannot be generalized to other locations. For example, NASQAN data from the James River at Cartersville, Virginia, do not fit conditions below or above that site since the river and influences on it are quite different at each location. Comparing NASQAN data between different rivers could be even less justified.

2. Survey Comment

Consistency is one of the most important characteristics of the NASQAN data base and the Survey quality assurance programs insure quality is maintained at a high level. GAO uses the Survey's review reports on incorrect field techniques and sample handling as evidence of poor data quality and also cites Survey studies of laboratory precision at non-Survey laboratories to call the NASQAN data into question.

GAO Evaluation

Although we agree that the Survey generally monitors for the same water quality characteristics at the same frequencies each month and has an active quality assurance program, we do not agree that NASQAN data are above suspicion or that the Survey consistently uses the same methods and procedures. For example, during the past 5 years, the Survey has changed filters and culture media used for bacterial sampling. These changes in the Survey's methods and procedures, although made to improve the bacteria data, inevitably bias the data because the new filters capture more bacteria. If at some later date improved methods for collecting bacteria samples are introduced, the new data will again be biased.

In an April 1978 article on "Microbiological Monitoring for Water-Quality Assessment" in the Journal of Food Protection, a Survey official noted that until an adequate bacteria sampler has been designed, tested, and made available, the data produced in microbiological-monitoring programs involving surface waters can be considered of questionable accuracy and can be misleading and erroneous. The article specifically stated:

> "The weakest link in the chain of events leading to production of reliable microbiological-monitoring data is a poor or inadequate sample. This results primarily from diversity of environmental conditions from which a sample must be collected. In surface waters, affinity of microbiological organisms for suspended particles necessitates that sampling procedures be designed to collect a representative sample of the

> water-sediment mixture. The key problem and challenge to microbiological monitoring is production of a sterilizable, depth-integrating sampler that will accommodate the disparity of sediment distribution as related to variations in depth and cross-section and the changes in streamflow. Until such a sampler has been designed, tested, and made readily available, the data produced in microbiological-monitoring programs involving surface waters can be considered of questionable accuracy * * * The data obtained from [current samples] * * * can be misleading and erroneous as to the true bacterial density in a body of water."

Bacterial pollution is of great concern to public health officials. CEQ has used questionable NASQAN bacterial data for trend assessments in its annual reports. But the Congress, the President, and the general public are not alerted to the changes in and inadequacies of Survey sampling procedures and methodologies which bias the data. The data remain in the NASQAN data base without any qualifications or other warning concerning their use.

We also disclosed in our report other types of inconsistent practices, such as delays in analyses of many nutrient samples (p. 41, vol. I), errors by field technicians (p. 39, vol. I), and variability in laboratory performance (p. 42, vol. I). All of these inconsistent practices make data from NASQAN inconsistent, but the inconsistencies are not revealed in the data records.

The question of consistency also applies to the water quality conditions being monitored. Without this consistency, analysis for trends in water quality is questionable. As we have explained in the report and in response to other comments by the Survey and EPA, the networks do not obtain data from homogenous water quality conditions.

3. <u>Survey Comment</u>

GAO gives three reasons for the inability of the networks to produce reliable, meaningful surface water quality data, two of which are conditioned on the potential of the data to not be representative of national water quality conditions. But nowhere does GAO operationally define the concept of representative. Also, for weaknesses in field and laboratory procedures, the concern should be for the precisions of measurements and their biases, both of which are addressed in Survey quality assurance programs.

GAO Evaluation

It does not matter how "representative" is defined. The national networks, because of inherent time and location biases, infrequent sampling, and lack of consideration of variances and significant changes in water quality, cannot produce data that are representative of median flow, extreme flow, average loading, or extreme concentration value conditions or trends in water quality.

With respect to the precision of measurements and biases, we agree that both are important and are covered in the Survey quality assurance program. Measurement precision is not of much value, however, if the number of samples is insufficient to portray water quality conditions, trends, and changes or if the measurements have time and location biases which make precise numbers misleading.

4. Survey Comment

The Survey disagrees with the recommendation to discontinue the networks and to devote their resources to well-managed special studies of water quality because the two approaches to water quality investigations are different and both are needed.

GAO Evaluation

As we have discussed previously in reply to Survey comments 1 and 3 and EPA comment 2, we do not believe the existing networks can meet their established objectives. Because of inherent weaknesses they cannot provide sound nationwide geographic summaries of water quality or identify long-term trends at network sites, within river basins, or on a national basis. We cannot agree that it makes sense to continue networks that do not produce useful data. Rather than collecting inadequate data at arbitrary sites around the Nation, we believe it is wiser to sponsor meaningful studies of water quality in fewer areas.

5. Survey Comment

NASQAN does not represent the Survey's only water quality program. The Survey carries out a variety of water quality studies and activities which are used in conjunction with NASQAN. A coordinated program of special studies would not meet the objectives of NASQAN nor could it be financed at the same level. If NASQAN is to provide a significant interpretive component as in the intensive survey program endorsed by GAO, then its program goals will need to be expanded and resources increased.

GAO Evaluation

We recognize that the Survey carries out a variety of monitoring activities, including some special studies. We do not dispute that a program of 1,500 special studies annually, about the same number as the stations in the three national networks, would be expensive. However, we do not agree that a program of special studies need be prohibitively expensive. Millions of dollars are available for various water quality monitoring efforts which could be used for special studies.

We believe the Survey has not recognized that special studies can provide greater interpretive potential than can NASQAN. The inherent weaknesses in the NASQAN program make the resulting data inadeqate for meaningful intrepretation. As the Survey has stated, NASQAN was never intended to be a source of information detailed enough to assess the effectiveness of pollution control measures.

In contrast, special studies can provide a thorough evaluation of current water quality and a solid technical basis for reliably predicting changes in river quality under altered conditions. The value of special studies for such purposes is specifically discussed in Survey Circular 715-K. The results of special studies remain valuable and useful for years and can repay their costs many times over through savings in the management of water quality programs.

We do not agree that NASQAN program goals will need to be expanded and available resources increased. As we have discussed in our evaluation of other Survey comments, we do not believe NASQAN meets its established objectives. Therefore, we cannot agree that expanding its program goals and resources would be useful or desirable. In our opinion, any additional resources should be directed to the much more useful special studies.

6. Survey Comment

NASQAN was created to provide a national data base on river quality, consistent in methods and suitable for examining conditions and trends. The program is meeting the objective. A case for discontinuing the program has not been made.

GAO Evaluation

As discussed in detail in our evaluation of other Survey comments, we do not agree that NASQAN is meeting, or can meet, its established objectives. Therefore, we believe the program should be discontinued. Also, we believe we have demonstrated that a national program of special studies can much better meet the need for reliable, accurate data, which can describe water quality conditions and trends.

7. Survey Comment

A coordinated study program of special studies would approximate a fixed-station network and the cost of a series of special studies would certainly far exceed the cost of NASQAN.

GAO Evaluation

We do not agree that a national program of special studies would approximate a fixed-station network. Unlike networks, special studies need not be distorted by time bias, location bias, and inconsistent data. Special studies also can provide reliable data on the rapidly changing properties of water, which, as we have demonstrated, networks do not. In addition, special studies provide for better understanding of water quality changes and reason for the changes throughout large river areas, in contrast to single site coverage accomplished by far-flung network sampling.

As we discussed in response to Survey comment 5, we do not disagree that a program of 1,500 special studies annually, about the same number of stations in the three national networks, would be expensive. We believe, however, that a more concentrated program of special studies, tapping the various sources of funding available for water quality monitoring efforts, can be designed.

8. Survey Comment

The Survey's efforts are not limited to the NASQAN network. The network represents an important part, but not the sole water quality monitoring activity of the Survey.

GAO Evaluation

We agree the Survey activities involve a variety of monitoring, including special studies. The report acknowledges the Survey's efforts in some of these areas.

9. Survey Comment

The temporal variability of the characteristics of the water resource, in addition to a random component, may include a cyclical component (diurnal, weekly, annual, etc.) and a trending (continuously increasing or decreasing) component. Because all data arise from a random process, it is imperative to interpret the data from a stochastic viewpoint. Data arising from a random process with large variances are not "difficult to compare" as GAO says, but rather the confidence intervals are large although they can be reduced with increased sample size.

GAO Evaluation

We fully recognize the variability of water quality characteristics in our report, and our discussion of statistical weaknesses clearly illustrates this point. (See pp. 29 to 39, vol. I.) We believe that the significant variability of data and lack of homogeneity in river conditions from which samples are drawn are overwhelming obstacles to the use of network monitoring.

The Survey offered a simple 3-factor explanation for variability in water quality measurements (random component, short-term cycles, and long-term trends). We disclose some of the serious weaknesses in the Survey's explanation in response to CEQ comment 24. However, one of the Survey's specific comments warrants additional response at this point. The Survey stated "Thus, because all data (whether pertaining to quality or quantity) arise from a random process, it is imperative to interpret the data from a stochastic viewpoint." It is not true that all data arise from a random process. Predictable patterns in water quality and quantity do exist. For example, most American rivers are warm in the summer and cold in the winter. Rivers are often muddy in flood seasons. These are not random processes, but are quite predictable.

10. Survey Comment

It is Survey policy to obtain water quality measurements that integrate the quality of all water passing through a cross-sectional area. A single measurement represents a single draw from a random process and reflects the composite effect of whatever has occurred upstream from the site.

GAO Evaluation

We cannot agree that single samples at widely scattered stations reflect the composite effect of events upstream from the station sites. As we discussed in our evaluation of EPA comment 2, water quality throughout a river basin is highly variable. Foreign substances added to water undergo changes along the course of the waterway, and the mix of the substances and the timing of their entry to the water changes frequently.

A river can be very polluted in many different reaches but can recover dramatically by the time it reaches a sampling station. Thus, a single sampling site cannot possibly reflect what has occurred upstream.

11. Survey Comment

The need for a nationally consistent set of water quality information for assessing the state of the Nation's rivers has been widely noted and NASQAN was established for this purpose. Such a requirement or solution is not unique to the United States.

GAO Evaluation

We agree that the Nation's rivers need to be assessed. We do not believe that NASQAN can produce the information needed for the assessments because of inherent timing problems, location bias, inconsistencies in data, and inability to account for self purification and rapidly changing properties of water. We have discussed these weaknesses in the report and in response to Survey comments 1 and 2 and EPA comment 2.

12. Survey Comment

In the design of NASQAN consideration was provided for the interrelationship and correlation between discharge and water quality. Care was taken to associate each NASQAN station with a discharge monitoring site so that a complete record would be available.

GAO Evaluation

We agree that water quantity (flow) information is needed to interpret water quality data. However, as we pointed out in our response to EPA comment 2, NASQAN stations are generally away from large cities and industrial complexes. The NASQAN site selection dramatically biases data from the network and neglects important waters where most people live and work.

13. Survey Comment

The objective of consistency is of great importance. It has been stated by researchers that perhaps the greatest promise for improving performance in water quality evaluation lies in NASQAN exactly because of the consistency it provides. NASQAN provides for a nationally uniform methodology for both field and laboratory analysis.

GAO Evaluation

We agree that procedural consistency is important. However, the Survey is not consistent in its methods and procedures, as pointed out in the example of inaccurate bacterial data in our evaluation of Survey comment 2. We do not deny that the Survey has an active quality assurance program. But the example also demonstrates that deviations from standardized methodology are not always strictly noted and corrective procedures taken. Even more important is consistency in the water quality conditions that are compared, whether the data are to be used to analyze conditions during a short span of time or over many years.

14. Survey Comment

NASQAN records are needed for many purposes of policy analysis, water management and hydraulic research, including (1) long-term trend assessment, (2) construction of probability distributions, (3) construction of standards, and (4) determination of correlative relationships.

GAO Evaluation

We do not agree that NASQAN data can or should be used for such purposes. This matter is discussed at length in our response to Survey comment 1 and EPA comment 2.

15. Survey Comment

The cost of special studies, in terms of both personnel and budget requirements, is high. If present levels of funding for NASQAN were applied to special studies on a 10-year restudy cycle, it would be possible to conduct 21 studies in any year or a total of 70. In addition to sparce coverage of such a program, the GAO proposal does not explain how the results of special studies would be unified to provide the kind of information sought in the NASQAN objectives.

GAO Evaluation

As we have noted in our evaluations of other Survey, EPA, and CEQ comments on this matter, we do not dispute that a program of 1,500 special studies annually, about the same as the number of stations in the three national networks, would be expensive. However, we do not agree that a program of special studies need be prohibitively expensive. What is needed is a well-managed program of special studies which taps the millions available for water quality monitoring activities.

The Survey calculated that if the present level of funding for NASQAN were applied to its Intensive River Quality Assessment program, only 70 studies could be conducted in a 10-year span. The Survey also characterized this as being sparse coverage compared to the 518 sites included in NASQAN. We believe that studies in the magnitude of the Survey's Intensive River Quality Assessment program are not at all comparable to single site coverage provided by NASQAN. For example, the Willamette River Basin study covered over 150 river miles and involved a major segment of Oregon's economic base and population. In contrast, only one station on the Willamette is used for NASQAN; about 13 miles from the downstream end of the river.

We further believe that special studies can be used to meet some of the objectives intended for NASQAN. This matter is discussed on pages 56 and 64, volume I.

16. Survey Comment

The Survey is aware of many ways in which NASQAN may be enhanced to better meet its objectives and plans to attend to such efforts. Some of the possible enhancements to be considered include (1) changing the frequency of sampling, (2) adding new sites, (3) rotating some sites, (4) deleting some water quality characteristics because of their high degree of correlation with other characteristics, (5) conducting a series of synoptic studies of river basins in concert with sampling at existing fixed stations.

The Survey, with the aid of outside advisory groups, intends to do what GAO failed to do: examine NASQAN objectives and the extent to which they are being met. Increased budgetary and manpower allocations would greatly expedite these efforts.

GAO Evaluation

We do not agree that we failed to examine NASQAN objectives and the extent to which they are being met. On the contrary, we believe that we have demonstrated that NASQAN has not and cannot meet its established objectives. This matter is discussed in our evaluation of EPA comment 2 and Survey comment 1.

17. Survey Comment

Now that NASQAN has more than 5 years of data, there will be an increased effort to analyze the data and publish findings on various subjects such as movement of substances through the Nation's rivers, changes over time in water quality, etc.

GAO Evaluation

As discussed in response to Survey comment 1 and EPA comment 2, we believe the weaknesses of the national networks preclude reliable analyses of the data.

18. Survey Comment

The Survey is committed to the concept of fixed-station, fixed-interval monitoring for the kind of national assessment function for which NASQAN is intended. NASQAN was created to fill a widely recognized need for a consistent national data base for examining conditions and trends in river water quality. GAO does not question this need and does not provide a demonstrated alternative method of fulfilling it.

GAO Evaluation

As we have previously discussed in response to the Survey's comments 2, 6, 11, and 13, we do not believe NASQAN provides consistent data. Furthermore, because of weaknesses in the existing networks we believe the Survey should question whether its commitment to NASQAN is necessary or justified. As stated in the report, we believe the Survey should discontinue its national network approach in favor of an expanded special studies program.

19. Survey Comment

The reports cited by GAO are not specifically identified. The specific findings or recommendations should be given.

GAO Evaluation

We have added a bibliography as appendix I to volume I, identifying the reports in question and other selected literature reviewed. We do not agree that our report need include detailed discussions of findings stated in other studies.

20. Survey Comment

The GAO report examined only 1 of 33 examples of use of fixed-station data that the Survey provided in January 1979, and the report gives no indication of the variety of uses made of the data.

GAO Evaluation

The January 1979 list of examples the Survey provided was in response to our requests to the Survey and EPA for five or six examples that, in their judgment, best demonstrated the use of water quality data from fixed stations. We had informed both agencies that we would not review in depth all the examples provided but would select examples to provide concrete cases for our examination of the validity of using fixed-station sampling. For our most intensive effort (see app. VII), we selected the example of CEQ use of data from NASQAN for its ninth annual report.

We reviewed aspects of other examples such as the Survey's own reports.

In the process of pinpointing examples, we screened all 33 examples provided by the Survey. We considered their recency, the variety of water quality characteristics involved, the likelihood and ease of identifying and evaluating the detailed water quality data used for the examples, and national significance of each example. We believe the examples we examined best represented what the Survey, EPA, and CEQ hope to accomplish with the national networks.

Regarding the Survey's comment about the variety of uses made of fixed-station water quality data, our report clearly stated that fixed-station monitoring had for years been the common technique used by Federal and State agencies (see pp. 3 and 7, vol. I). But past practices do not necessarily justify continued use of network monitoring. As discussed in the report, the need for better water quality data had been generally recognized for years. (See pp. 3 and 57, vol. I.) In fact, the Survey initiated its river quality assessment program in 1973 to develop techniques for studying various water quality conditions and to produce meaningful water quality assessments.

21. Survey Comment

The GAO report does not reference the work or opinions of experts within the academic community on the subject of water quality monitoring and data analysis.

GAO Evaluation

We have added a selected bibliography of literature reviewed. (See app. I.) Also, the Survey included a bibliography as part of its comments on our report. (See p. 80 of this vol.)

22. Survey Comment

The experts in the academic field who GAO contacted should be identified and their concurrence or disagreement with the recommendations of the report should be expressed.

GAO Evaluation

We do not believe it necessary to report comments from all individuals we contacted during our review. Such a procedure would be both cumbersome and prolonged. During our review we obtained the opinions of a variety of individuals familiar with fixed-station monitoring, including some in favor of the approach.

As a matter of policy, we have obtained and analyzed the comments of those agencies most directly affected by our recommendations--the Survey, EPA, and CEQ.

23. Survey Comment

Nowhere in the GAO report is it argued that the objectives for NASQAN are inappropriate. Throughout the report, NASQAN is repeatedly judged in light of different objectives.

GAO Evaluation

We do not agree. We have previously stated that NASQAN does not meet the objectives established for it. See our responses to the Survey's comments 1 and 4 and EPA comment 2.

24. Survey Comment

Now that NASQAN coverage has become extensive and data have been recorded for more than 5 years, the Survey is conducting a project to develop statistically robust procedures for assessing water quality trends.

GAO Evaluation

A substantial amount of fixed-station water quality data was available when the networks were being established. We do not believe it was necessary for the Survey or EPA to delay studying the statistical or practical complexities of network monitoring until even more data were collected. Developing statistically robust procedures for assessing trends cannot cure the time bias, location bias, and inconsistencies and errors in network data, and cannot alleviate the network's inability to account for self-purification or rapidly changing properties of water.

25. Survey Comment

The intent of NASQAN is to account for the quality of the water moving into and out of the various accounting units. Providing a description of water quality at all points along a given river is not among the purposes of NASQAN.

GAO Evaluation

The Survey contends that it can account for the quality of water through NASQAN. We do not agree. Most water quality properties respond rapidly to environmental influences; infrequent sampling cannot account for these rapid changes. We demonstrated this inadequacy of NASQAN in our report. (See pp. 18-39, vol. I.)

Although the Survey claims it does not intend to characterize water quality throughout rivers based on NASQAN, it and CEQ have used NASQAN data for that purpose. In another comment (see p. 29 of this vol.) the Survey stated that measurements taken at NASQAN stations reflect "* * * the composite effect of whatever has occurred upstream from the site * * *." We have explained why NASQAN should not be used to characterize water quality at other locations in vol. I (see pp. 21-28) and in our responses to Survey comments 1 and 10 and EPA comment 2.

26. Survey Comment

The GAO report contends that infrequent sampling at widely spaced locations will not produce representative measurements, but does not define "representative." The Survey believes that NASQAN measurements should characterize the unrestricted range of water quality conditions at a site.

GAO Evaluation

We have responded to the Survey's comment on representativeness in connection with Survey comment 3. The Survey's claim that its NASQAN monitoring characterizes the unrestricted range of water quality at individual sites is simply not valid. As we have discussed in our evaluation of EPA comment 2, infrequent sampling restricted to a span of several daytime hours each month creates an overwhelming time bias, preventing most water quality conditions from being measured.

27. Survey Comment

Mistakes in the field and laboratories will occur in any data collection program, whether fixed-station monitoring or intensive studies. The Survey has training programs and a quality assurance program to minimize the mistakes. GAO makes no recommendations of ways to lower the frequency of errors or to improve precision.

GAO Evaluation

We agree that mistakes can be made in any type of water quality effort and recognized in the report that the Survey has a quality assurance program. (Also see Survey comment 2.) However, as we have explained in the report (see p. 56, vol. I) and in response to EPA comment 11, quality assurance can receive much closer attention in special studies than in network monitoring. Errors and inconsistencies in techniques and data are likely to be identified and corrected during a special study. In the networks, however, errors and inconsistencies can persist for years before they are recognized.

28. Survey Comment

The Survey's computerized information system (WATSTORE) allows for a considerable amount of information on the variability of local conditions. WATSTORE is being improved to provide information on the precision of each analysis.

GAO Evaluation

We agree that the Survey is able to record more information in WATSTORE than it currently does. But there is much information that cannot be recorded. For example, the case study on the James River (see app. VII) disclosed many unrecorded important factors that are needed to properly interpret dissolved oxygen measurements.

As discussed in the report (see p. 46, vol. I), WATSTORE and EPA's STORET system are not designed to handle all of the information needed for accurate interpretation of water quality data. One prime example is the lack of special notations for samples tested after long delays in shipment or storage. These questionable measurements are routinely filed with other data from samples that were processed properly, creating an inconsistent and suspect record.

29. Survey Comment

The points attributed to Professor Velz have been taken out of context and are not relevant. The four points concern the waste assimilation capacity of a specific river reach. This is not among the objectives of NASQAN.

GAO Evaluation

We cannot agree. We discussed the points in question with Professor Velz before including them in our report. He agreed they were entirely appropriate. Professor Velz wrote the book, "Applied Stream Sanitation", including the chapter from which we extracted material questioned by the Survey, with the primary objective of documenting important factors essential to the rational analysis of rivers. In his and our opinion, reliable and accurate interpretation of water quality conditions and trends must recognize the complexities he describes.

30. Survey Comment

The importance of cross sectional variability of water quality is well known to the Survey. Sampling techniques used by the Survey assure that the sample is appropriately integrated from the entire cross section.

GAO Evaluation

We agree that the Survey samples at several points across rivers for some characteristics. We were not taking exception to the Survey's approach but were disclosing one more facet of varying water quality conditions in rivers.

31. Survey Comment

The material quoted from Survey Circular 715-D must be put in context. Also, the list of shortcomings of existing data noted by Wolman in 1971 describes well the shortcomings that NASQAN was designed to overcome. In addition, the authors of Circular 715-D stated in a later passage on page D9 that many existing river quality data "* * * can still be usefully interpreted, provided the data are amenable to segregation using river hydrology as the segregating tool." NASQAN data, because they contain flow data, are specifically designed for undertaking the kind of hydrologic analysis that the Circular 715-D authors refer to.

GAO Evaluation

We have expanded our quote from Survey Circular 715-D (see pp. 16 and 17, vol. I). We cited the circular to illustrate that network monitoring problems were widely known within and outside the Survey and EPA. The expanded content reinforces, rather than weakens, our point. The Survey has given the impression that NASQAN answers the problems noted by the Circular. We believe this is a wrong impression. As we have demonstrated in the report, the rigid, infrequent sampling program under NASQAN cannot adequately capture the highly complex nature of river water quality. Contrary to the impression given by the Survey, we believe the authors of Circular 715-D fully recognized this point. The abstract of the Circular set forth below, clearly reveals the subject matter of the Circular and the inherent problems with network sampling.

> "In many basins it has proven difficult to use existing river-quality data for analysis of the temporal and spatial trends and the major cause-effect relations that control critical quality conditions. Major reasons for this problem are the arbitrary nature of sampling programs that generate the quality data and a general failure to account for the background variability in quality resulting from hydrologic phenomena.
>
> "A review of prominent river-quality problems of the Nation's river shows that the timing, location, and frequency of occurrence of the problems

are largely controlled by three hydrologic characteristics--streamflow, water temperature, and channel morphology. These characteristics show marked variation from river to river and, thus, must be systematically accounted for if river-quality data are to become more useful for interpretive purposes. An approach to river-quality data programs based on hydrologic analysis and repetitive synoptic studies is proposed as an alternative to current approaches that rely heavily on routine monitoring."

32. Survey Comment

No reference is given to the memorandum. Its author is not identified and the arguments made in it are neither paraphrased nor quoted.

GAO Evaluation

We added the identification of the memorandum to our report and quoted some of the concerns expressed in the memorandum (see p. 17, vol. I).

33. Survey Comment

The quotation from Velz is not relevant to the objective of NASQAN.

GAO Evaluation

We and Professor Velz disagree. Professor Velz insists that infrequent samples are drawn from dissimilar conditions. Consequently, what appear to be changes in water quality could instead easily be a quirk of sample timing. This comment is certainly relevant to NASQAN objectives, such as detecting changes in water quality, accounting for the quality of the Nation's rivers, and depicting geographic variability of water quality.

34. Survey Comment

The Survey is well aware of the shortcomings of monthly sampling of dissolved oxygen (DO). The Survey has selected monthly sampling for DO over three other approaches because it costs less and because NASQAN sites are not necessarily in DO problem areas. According to the Survey, it also appeared wasteful to fail to measure DO since it can be done inexpensively by the field technician while collecting samples for other tests.

GAO Evaluation

Only the Survey's acknowledgement that it is well aware of shortcomings of monthly DO sampling is relevant to the cited section of our report. We included our explanation of DO daily variability primarily as educational material for the lay reader. A more complete discussion of the complexities of monitoring DO is presented in our case study. (See app. VII, vol. I.)

35. Survey Comment

The statement "Under the NASQAN program the single reading of 10.5 mg/L represents the entire month of July 1977" is misleading. The entire distribution of a random variable is not to be inferred from a single value.

GAO Evaluation

We have improved the wording of the report (see p. 19, vol. 1). Our point remains valid. One measurement of DO during a single day of wide fluctuations can be misleading and one DO measurement a month cannot adequately represent DO conditions throughout the month.

36. Survey Comment

The data presented in the example involving dissolved lead are neither complete nor correct. Also, one of the measurements has been deleted from the Survey's files because of a contamination problem. The Survey disagrees with the GAO statement that there is no discernable pattern to the data.

GAO Evaluation

The report (see pp. 19-21, vol. I) has been changed to more clearly illustrate that water quality characteristics undergo substantial changes in rivers. The Survey does not dispute this basic point. We have substituted an example involving suspended sediment in place of dissolved lead. As we noted in the report, the Survey's own data clearly show that water quality characteristics change frequently and substantially.

37. Survey Comment

The GAO report states that "The examples in this section illustrate the difficulty of comparing water quality [for individual days, months, seasons, and years]." In response, the Survey stated "But it should be noted that the examples are all in accord with accepted statistical sampling practices and the analysis of these data by accepted statistical techniques presents no particular difficulty."

GAO Evaluation

We cannot agree. The Survey is overlooking fundamental sampling and statistical principles when it contends that comparing water quality at different times is not difficult. This is discussed in detail in the report (see p. 29, vol. I) and in our evaluation of EPA comments 1 and 2; Survey comments 9, 38, and 49; and CEQ comment 24.

38. Survey Comment

The criticism of NASQAN implied by the quotation cited by GAO has no basis. The quotation is a prescription for achieving the objectives of River Quality Assessment, which do not coincide, even in part, with the objectives of NASQAN. As an analogy, an intensive study such as the River Quality Assessment may be compared to diagnostic and testing work of a physician for individual patients, while NASQAN may be compared to the collection and analysis of national health and mortality statistics.

GAO Evaluation

We disagree. We used the quotation from a Survey Circular because it clearly stated the need to understand the unique behavior of individual rivers in order to assess whether water quality in the rivers change over time. As we have stated in response to Survey comment 1, NASQAN sampling does not produce information needed to understand the behavior of individual rivers because of time bias, location bias, and the inability to account for rapidly changing properties of water or for the self-purification process involved in rivers. We believe the material quoted was pertinent, but we deleted it in revising our explanation of water quality variability.

The Survey's comparison of NASQAN with the collection and analysis of national health and mortality statistics is inappropriate. As we have demonstrated in the report (see pp. 29-39, vol. I), network sampling, including NASQAN, produces water quality data from nonhomogeneous conditions. These conditions preclude valid comparisons of data from different periods. Infrequent sampling at sparsely located stations cannot reliably measure highly complex water quality. The NASQAN sampling program was based primarily on administrative convenience rather than statistical requirements. In contrast, the national health statistics are developed from a multistage probability sample that was designed to permit a continuous sampling of the noninstitutionalized civilian population. During a year, interviews are conducted in approximately 40,000 households. Mortality statistics are based on copies of all vital records received from registration offices of all States, certain cities, and the District of Columbia.

39. Survey Comment

EPA and the Survey have dissimilar responsibilities and thus it is reasonable that they would choose dissimilar monitoring locations. GAO has not offered any alternative approach to selecting station locations.

GAO Evaluation

The Survey's comment is not relevant to the cited section of our report. We were not proposing that the two agencies monitor water quality at similar locations but instead briefly stated that "EPA and the Survey had not attempted to capture any particular pattern of water quality conditions for their two main networks." In subsequent subsections of the report (see pp. 27-28, vol. I), we discussed in more detail the criteria used by the two agencies in selecting monitoring locations.

The Survey's claim that we did not offer any alternative approach to selecting station locations is misleading. On page 50, vol. I, we explained that it was impractical to fix the networks because an inflexible program is not consistent with the variability of the Nation's waters. We pointed out that sampling sites and sampling frequency, as well as the water quality characteristics monitored, should all be dictated by the unique character of each river and the purpose of the sampling effort. We concluded that, in effect, sampling would have to be converted to special studies in order to produce meaningful data. In chapter 4 we explain some of the traits of special studies that make them superior to networks.

40. Survey Comment

The material on the South Platte River and on spatial variability is not relevant to NASQAN. If fecal coliform measurements specifically in the Denver area are important, sampling must be done there. NASQAN is not intended to develop this kind of information.

GAO Evaluation

We cannot agree that the cited passage of our report is not relevant to NASQAN. The Survey and CEQ are using NASQAN data from single downstream locations to characterize water quality throughout large river drainage areas. We very clearly demonstrated in the questioned section of our report that the location of a sampling site significantly influences water quality measurements. In the report we demonstrated that large upstream areas can have far different water quality conditions and trends than what are found at NASQAN stations. We believe the use of data obtained at NASQAN stations to describe upstream conditions is wrong.

41. Survey Comment

The intent of NASQAN was not to monitor polluted rivers, or clean rivers, or urban rivers, but to monitor the inflow and outflow of accounting units. If one were interested in the status and trends of a particular category of rivers, one could select data from NASQAN sites on such rivers for analysis.

GAO Evaluation

As stated in the report and in response to other Survey and EPA comments, we believe NASQAN cannot provide data adequate either for the accounting unit objective or for the analysis of the status and trends of various categories of rivers. The most important weaknesses of NASQAN and other networks are timing problems, location bias, and the inability to account for the rapidly changing properties of water or for self-purification. We discuss these weaknesses in the report and have elaborated on them in response to several agency comments, especially EPA comment 2 and Survey comments 1, 3, 9, and 10.

42. Survey Comment

There is no basis for the statement "* * * uncertainty is increased with far flung networks * * *." Survey field personnel use the same methods, receive the same training, and are evaluated by common standards no matter where they work.

GAO Evaluation

We disagree. The human factor is still very much involved in water quality monitoring, as has been disclosed by the Survey's own field reviews and as the Survey noted in its comment 27. The Survey's efforts to improve field work and to establish more consistency are laudable. But we continue to believe a greater likelihood exists for technical inconsistencies in widespread networks than in more tightly controlled special studies.

43. Survey Comment

The GAO report has noted a possible problem with respect to delays during storage and shipment of samples. The Survey intends to explore all of the steps in the process and will consider changes and their consequences for data reliability and cost.

GAO Evaluation

As we discussed in the report (see p. 42, vol. I), the Survey has been aware of the problem at least for several years. We do not know why the Survey had not started earlier corrective action. Because of the delays, many measurements recorded by the Survey for NASQAN are suspect.

44. Survey Comment

The description of variability of laboratory performance disclosed by the Survey's testing program is not pertinent to NASQAN because NASQAN analyses are all performed at the Survey's central laboratories. A better evaluation of NASQAN analyses can be made by comparing the variability of central laboratory analyses of standard samples with the variability expected on the basis of comprehensive testing of similar concentrations by the American Society for Testing and Materials. The Survey's central laboratories compare favorably with the expected variability over time.

GAO Evaluation

We disagree with the Survey's contention that its testing program is not pertinent to NASQAN. The two Survey central laboratories have participated in practically all of the tests. While the two laboratories generally are among the better performers, they vary over time and between themselves in the same tests. For example, in a test for ammonia (NH_3-N) in 1979, the Survey's central laboratory in Georgia reported a concentration of 0.61 mg/L while the central laboratory in Colorado reported a concentration of 1.3 mg/L.

The summary statistics cited by the Survey for analyses of chloride standard solutions are useful for general indications of laboratory performance, but they do not reveal the full range of actual measurements. Since NASQAN relies on single and separate samples once a month for most measurements, the potential error from month to month or between laboratories is quite important. In contrast, during special studies, a sufficient number of samples would be taken over a short, stable period of time to enable discovery of abnormal measurements.

45. Survey Comment

The GAO report makes no comparison of data reliability in fixed-station operations versus that in intensive studies. Thus, it is unclear how reliability of the data has any bearing on the question of fixed-station monitoring versus special studies.

GAO Evaluation

We discussed variability of laboratory performance as part of our intent to explore all major steps involved in network monitoring. Many past papers on water quality monitoring have concentrated on one or two aspects, such as statistical manipulation of data, giving less attention to other, equally important steps. We believe it is important to consider all aspects since they are so entwined; the validity and usefulness of individual measurements is affected by the composite of time bias, location bias, the complexities of water quality changes, and the quality of field and laboratory work. In our report we discuss how the special studies approach can enable the Survey or other agencies to use tighter management and quality assurance techniques than can be done through national network monitoring. (See p. 56, vol. I.)

46. Survey Comment

Sampling only during certain flow conditions or seasons would be useful for evaluating water quality during those periods. But, by sampling throughout the year on a regular schedule, which is unrelated to flow conditions, the Survey gains information on the range of variation of water quality conditions. Subsequently, the wider range of data can be evaluated for various relationships.

GAO Evaluation

We disagree that monthly network monitoring is a good technique for obtaining valid measurements of water quality throughout the year or that the data can be reliably evaluated for various relationship. We have responded at length to this matter in connection with Survey comments 1, 3, 9, and 26 and EPA comment 2.

47. Survey Comment

Measurements cited in the report for two samples taken in July of 1976 and 1977 at the NASQAN station on the Yakima River are of little interest by themselves. Their usefulness arises after several years of data have been accumulated. Then the data can be explored for various distribution and relationship possibilities. One of the characteristics GAO cited--total hardness--is very closely related to discharge (streamflow). Because 24 years of data for total hardness at the site are available, the Survey is able to calculate that the two measurements cited by GAO were about what should be expected. Without this past data it would not be possible to make this kind of interpretation.

GAO Evaluation

In the report, we had included four water quality characteristics, plus streamflow, measured in July 1976 and July 1977 at the NASQAN station to illustrate that substantially different measurements can be obtained from year to year in the same month when only one sample a month is taken. Our discussion was tied back to the flow chart for both years which revealed that the river's flow levels and patterns were quite different in the 2 years. The diagram also illustrated that the Survey's samples were taken at different flow stages each year; with some taken during dynamic changes in flow. As we stated in the report, these and many other factors could account for the differences in measurements.

We disagree with the Survey's implication that water quality data generally can be fitted into sound correlative relationships with streamflow. Although total hardness can have reasonably close correlation with streamflow because it is a summation measurement of many constituents, many characteristics have far less correlation with flow. For example, at the same NASQAN station, high concentrations of suspended sediment (another gross measurement) often were observed at remarkably different flow levels. Other characteristics such as nutrients and bacteria have even worse correlations. We have previously discussed why network data should not be used for correlative relationships in response to EPA comment 2 and Survey comment 1.

48. Survey Comment

The statement "weakly supported conclusions * * *reached by relying on water quality measurement data without considering flow data * * *" is not relevant to the recommendation of the report. The Survey is in full agreement with GAO on the importance of flow data and included it in NASQAN.

GAO Evaluation

We disagree with the Survey's contention that the statement was irrelevant to our recommendation. We recommended that the national networks, including NASQAN, be discontinued and that the Survey and EPA devote their resources and attention to a program of well-managed special studies of water quality. The statement in full, which the Survey excerpted, was

> "The case involving CEQ's use of network data which is discussed on p. 39 and in Appendix VI, demonstrates that weakly supported conclusions can be reached by relying on water quality measurement data without considering flow data and other important factors."

The Survey, in commenting on this passage, ignored the last phrase "* * * and other important factors." The case study clearly pointed out many reasons why more than monthly measurements at single sites on a river are needed to understand water quality conditions and trends. As we stated in response to EPA comment 2 and Survey comment 1 plus others, we believe that because of time bias, location bias, and the inability of network sampling to account for the rapidly changing properties of water and the self-purification process in rivers, the networks cannot adequately describe or characterize water quality. Inherently, network sampling must lead to weakly supported or vague conclusions about water quality. We have revised the passage questioned by the Survey to more clearly make this point.

49. Survey Comment

The report states that "Annual means are more understandable if they are accompanied with clear information on the extent of variance among the individual measurements used to calculate the means." The Survey's latest report on NASQAN data gives sample size, standard deviation, and range for every mean value. Thus the data, and the Survey's reporting of it, conforms to the GAO suggestion.

GAO Evaluation

We disagree. The Survey has not been revealing and dealing with the nonnormal distribution of data used for its calculations of annual means and the nonhomogeneous conditions from which samples are taken throughout each year. We have modified the report (see p. 32, vol. I) to more specifically demonstrate that annual mean values are not good indicators of water quality and that comparisons of annual averages are not valid and meaningful guides. We now clearly point out the fallacy of summarizing individual measurements taken from nonhomogeneous conditions into annual averages. We also discussed this point in response to EPA comment 2 and CEQ comment 24.

50. Survey Comment

The statistical statement in the report in the nitrite plus nitrate example is wrong.

GAO Evaluation

The Survey's comment is correct. We have changed the discussion of the example (see p. 33, vol. I) to more clearly demonstrate why statistical summaries of monthly measurements can be misleading.

51. Survey Comment

The ad hoc procedure used by GAO to test for trends in fecal coliform bacteria does not constitute a recognized statistical test and its sampling properties are unknown.

GAO Evaluation

The Survey's comment concerns a section of our report that was identified by the caption "Fecal coliform bacteria means are unstable." We have eliminated that section from the report, not because of the Survey's comment, but because of EPA's admonition that arithmetic means are inappropriate for characterizing fecal coliform bacteria (EPA comment 13 on p. 18). We agree with EPA. Unfortunately, the Survey and CEQ have used arithmetic means to describe fecal coliform bacteria conditions in published reports.

The Survey did not take exception to our basic point -- that annual averages for fecal coliform bacteria are unstable and reports based on the means are unreliable and misleading. Instead, the Survey questioned the technique we used to illustrate the point. We seriously doubt that the Survey would disagree that fecal coliform bacteria measurements are extremely variable. This is commonly known in the field of water quality and is substantiated by the Survey's own data.

52. Survey Comment

The Survey's information system (WATSTORE) contains parameter codes that can be used to describe conditions under which samples are collected. The Survey is also in the process of adding capability to WATSTORE for recording methodology and precision information.

GAO Evaluation

The Survey overlooked our basic point--that water quality data are drawn from the EPA and Survey computerized information system and used for analyses without important information on the quality of the data and the wide range of local influences that can affect individual measurements. As an example, we mentioned specifically the lack of any distinction in the information systems between measurements recorded for stale samples versus fresh samples.

We agree that WATSTORE can be used to record more information than is currently done. But, the list of additional parameters provided by the Survey covers only some of the information needed to account for and understand sampling results. In other parts of the report we discussed additional field and laboratory errors and many other temporary and long-term local influences that affect measurements recorded in the computerized information system. (Also see Survey comment 2.)

53. Survey Comment

The purpose of the discussion of water quality monitoring of the James River, Virginia, is unclear. GAO seems to be taking issue with CEQ's analysis of data and yet the report's recommendations do not address this use. Also, the points raised by GAO do not seem to address NASQAN objectives.

GAO Evaluation

We believe the purpose of this section of the report is very clear. The section questioned by the Survey is a brief summary of the case study. The Survey reviewed the case study in detail before we issued our draft report and a second time as part of our draft report. We believe the opening paragraph in the case study summary stated clearly the relevance of the study to national network monitoring. (See p. 47, vol. I.) We stated

> "A clear understanding of the local circumstances surrounding individual samples and the rivers or streams being monitored is important if sound interpretation of water quality changes are to be made."

54. Survey Comment

Fixed-station monitoring provides background information to help guide intensive data collection programs. As an example, the Survey study team which performed the Willamette River Quality Assessment used past DO data as background for a reconnaissance study.

GAO Evaluation

We agree with the basic concept that past water quality data should be considered when designing special studies. However, care needs to be taken in relying on past data generated from infrequent, widely scattered sampling because of the time and location bias and many other problems that affect the reliability and accuracy of network data. In the case the Survey used as an example, the past DO data were developed through special transverse sampling at a number of locations throughout the Willamette River. In comparison, the Survey has only one NASQAN station located near the mouth of the Willamette River. That station cannot provide useful background data for upstream stretches of the river where completely different DO conditions prevail.

55. Survey Comment

The GAO criticism is meaningless (referring to a statement that "* * * hundreds more samples might be needed for statistically significant data.").

GAO Evaluation

We have corrected the discussion (see p. 51, vol. I) to state in acceptable terms that in order to determine whether differences in measurements observed from period to period are statistically significant, it would be necessary to increase the number of samples taken. Because of differences in variation of rivers and the measured characteristics, an optimum sample size would have to be established for every site and adjusted as conditions change.

56. Survey Comment

The Survey agrees with the need for special studies tailored to NASQAN sites to enhance NASQAN sampling. However, the objectives of NASQAN cannot be served by customized studies alone. There must be comparability, site to site and over time, if the objectives of NASQAN are to be met.

GAO Evaluation

We agree that NASQAN sampling would need to be comparable over time and site to site if NASQAN objectives are to be met. However, as we discussed in the report and stated in connection with Survey comments 1 and 2, EPA comment 2, plus other comments, sampling done through the national networks, such as NASQAN, cannot produce comparable data.

57. Survey Comment

"Comparability of data" is only a function of methods of data collection and analysis. Because consistency in data collection is stressed under NASQAN, the resulting data are comparable. The GAO report does not describe how a special studies program would facilitate comparisons of water quality over space and time (which are the objectives of NASQAN).

GAO Evaluation

The Survey's comments concerned the following conclusion we stated in the draft report:

> "Comparability of data over time at one location and among various locations throughout the Nation is also needed for good assessments but is not possible through the networks due to the widely varying conditions and measurements."

That concluding statement was a summation of our demonstration in the report that infrequent, sparse sampling produces data from a hodgepodge of water quality conditions and that reliable, accurate analyses of the water quality conditions and trends cannot be developed because of the widely varying river conditions and absence of other crucial information.

We have addressed the Survey's arguments about comparability of data and use of network data for determining river water quality status and trends in response to Survey comments 1,2,3,10, and 13 and EPA comment 2. We also disclosed in the report why published reports, such as CEQ's annual reports and the one prepared by Briggs and Ficke (1977), which rely on annual averages of NASQAN data, cannot reliably portray water quality conditions. (See p. 32, vol. I.)

We believe the Survey did not grasp the meaning of the statement. In order to make the meaning of our conclusion quite clear, we changed it (see p. 52, vol. I). The change pinpoints our belief that assessments of water quality based on network data are not valid.

58. Survey Comment

Relating water quality conditions to causes is not among the objectives of NASQAN. Once conditions have been characterized at a NASQAN site, then interested local, State, or Federal agencies may track down the causes.

GAO Evaluation

We are pleased that the Survey agrees that NASQAN is not intended to identify causes of water quality conditions. As stated in the report, we believe that assessments of river water quality conditions and trends, to be meaningful to policy and decisionmakers, must include the link between water quality changes and reasons for the changes. If NASQAN is intended to supply only the measurements and not the linkage, the network will provide, at best, only part of the information needed for useful assessments. However, that part supplied by NASQAN is, in itself, of questionable accuracy and reliability. As we have demonstrated in the report and have discussed in response to several Survey comments on the report, the infrequent, sparse sampling done through NASQAN is laden with time bias and location bias and cannot account for rapid changes in water or self-purification of rivers. In our opinion, monthly sampling is woefully inadequate for characterizing conditions at NASQAN sites or for pinpointing river areas needing more detailed investigation.

The Survey contends that other agencies will perform special studies if the NASQAN data indicate a need for such studies. We believe this contention is misleading. As we discussed in the report and in response to Survey comments, many NASQAN stations are not located in river areas severely affected by human activities and thus are unlikely candidates for detailed study. Agencies concerned about water pollution control efforts presumably will give greatest priority to evaluating water quality changes in areas where important pollution control efforts are needed or have been undertaken.

59. Survey Comment

The Survey concurs with the report on the value of special studies but believes that the information from them is different than NASQAN information. Special studies generally do not provide much information on temporal (time) variability. It may be possible to produce useful progress reports from special studies but the methods have not, to the Survey's knowledge, been developed. Successive special studies using the same sites and sampling methods clearly constitute a fixed-station monitoring program.

GAO Evaluation

This Survey comment and Survey comment 64 reveal a strong desire to continue NASQAN at the expense of special studies. Special studies are discounted by the Survey as being incapable of producing useful progress reports and as being little more than expanded fixed-station sampling. We disagree with the Survey's position. In the report we explained some of the distinguishing features of special studies and discussed several examples of special studies of varying complexities and purposes.

We also believe that the Survey's statement that methods have not been developed for producing useful progress reports from special studies is overly complicating the issue. For example, followup studies can determine the results of pollution control measures by comparing river conditions before and after the corrective steps. We discuss this matter in the report (see p. 56, vol. I).

60. <u>Survey Comment</u>

NASQAN has been in existence long enough to provide a data base suitable for some types of analyses, and such work is now underway. The Survey cited several analyses, including a study of the relationship of the transport of major ions to the geology and human population of the river basin and an assessment of the impact of acid precipitation on the Nation's streams.

<u>GAO Evaluation</u>

The Survey's comment was in response to the following paragraph (see p. 53, vol. I) in our report.

> "A distinguishing feature of special water-quality studies is that emphasis is placed on <u>analysis</u> as well as data collection. Properly performed studies can reveal the quality of water at specific locations, why particular water quality conditions exist, and why they change. This information is essential for accurate and meaningful assessments of water quality and for well-founded assessments of pollution-control programs."

The Survey did not take exception to our statement, but apparently intended to claim that the same type of analyses can be done with NASQAN data. We disagree. As we have stated in the report and in response to other Survey and EPA comments, NASQAN is laden with serious time and location biases, inconsistencies, and it cannot account for the rapidly changing properties of river water or for self-purification. The examples cited by the Survey involve cause/effect relationships and complex timing and interaction of various properties of river water. We believe that meaningful, reliable analyses of those types cannot be produced from infrequent sampling at widely scattered locations (such as is done through NASQAN).

61. <u>Survey Comment</u>

The Survey agrees with GAO as to the importance and value of intensive river quality assessments. That program is a demonstration program, designed to foster the more widespread use of intensive assessments by State and local water management agencies and by the Survey's district offices. But, there is a use-a useful relationship between fixed-station data and intensive studies that the GAO report completely overlooked. Intensive studies depend on networks for background information, for guidance in planning data collection, and for a statistical context in which to view the data collected in the intensive study.

GAO Evaluation

We disagree with the Survey's insistence that networks are essential. As we have discussed in the report and in response to several Survey and EPA comments, infrequent, sparsely spaced sampling done through networks does not provide statistically reliable data. As stated in the report (see p. 63, vol. I) rather than continuing a national network that provides data of questionable validity, we believe the Survey should perform more special studies. The studies would provide far more reliable background data for future assessments than would continued NASQAN monitoring.

62. Survey Comment

Intensive water quality studies cover a wide range of levels of effort. The examples for which GAO has high praise are not the $14,000 type of studies that are mentioned in the report, but are the $0.25 million to $1 million.

GAO Evaluation

In the draft report we stated that the cost of special studies can vary substantially. We cited EPA's 1977 estimate of $14,000 for a small study and the $200,000 cost of the Survey's cooperation study of the Russian River in California, which we described on page 59 of the report. Because of the Survey's concern, we have added a comment in our report that the Survey estimates the cost of a typical study performed under its intensive river quality assessment program to be about $750,000. (See p. 61, vol. I.) Related discussions of the cost of special studies can be found under Survey comment 15 and EPA comment 5.

63. Survey Comment

The return of fish does not assure us that toxic materials are not present in concentrations sufficient to cause long-term pathological effects in fish or humans. Similarly, information on "specific pollution control actions" is not sufficient basis for evaluating progress toward cleaner water in the affected river (particularly in light of the large and poorly known amounts of nonpoint source loadings in many rivers). Biological monitoring can be useful, but at present only a few very specialized techniques are well developed and standardized (such as the Mussel Watch), and these pertain to only certain water uses.

GAO Evaluation

We do not agree with the Survey's dismissal of the use of other indicators of progress toward cleaner water. We have responded to each of the above Survey statements in connection with other Survey and EPA comments, particularly Survey comments 1, 4, 5, and 16 and EPA comment 2. Also, as the Survey noted in its comment 1, NASQAN was never intended to supply information detailed enough to assess the effectiveness of pollution control measures on a localized basis.

64. Survey Comment

The only practical and economical way to evaluate progress is by looking repetitively at fixed stations. The Survey does not consider the present NASQAN program to be unchangeable. The Survey has and will consider changes in sampling frequency, station location, sampling methods, or water quality characteristics measured. Any such changes will be considered in light of the NASQAN objectives. The GAO has not offered any suggestions on changes in the Survey's programs which would further the achievement of the objectives established for NASQAN.

GAO Evaluation

We disagree. In our draft report we pointed out many weaknesses in the networks and specifically discussed the impracticality of fixing the networks. We described why special studies are much better than network monitoring. We have discussed these points in response to several other comments by the Survey, EPA, and CEQ.

65. Survey Comment

The GAO report does not explain how special studies could be organized to provide the kind of consistent national overview required by the Survey's mission and by the NASQAN objectives. Neither has the report provided any consideration of the costs of coordinated special studies that would accomplish these objectives.

GAO Evaluation

The Survey's comment is misleading. As we have stated in response to other Survey comments, we do not believe NASQAN can satisfy the objectives that the Survey established for it. We also do not believe that the quality of data produced through NASQAN and resulting analyses meet the reputed quality of past Survey appraisals of the Nation's resources, which is the Survey's basic mission. The special studies approach, as the Survey is fully aware, offers the opportunity for the Survey and other agencies to use scientifically sound water quality monitoring techniques. As we explained in the report, unlike network monitoring, specially designed water quality studies are tailored to fit each unique river pattern and can provide essential information on why particular water quality conditions exist and why they change.

Mr. Henry Eschwege
Page 5

		GAO EVALUATION COMMENT No.	PAGE(S)
	Concentration is the important physiological parameter. An organism at the Missouri-Mississippi junction does not react to the loading, it reacts to the concentration (for example does the concentration expressed as mg/l surpass its tolerance). (As another example, drinking a half pint of sea water would be very unpleasant. If it were diluted in two gallons of distilled water and consumed gradually, it would not be unpleasant In both cases the amount of salt -- the loading -- is the same, while the concentration is very different.) Water quality violations are based on concentrations, not loading. As far as an organism is concerned, concentration is the basis of physiological, toxicological or behavioral effects. Whether concentration changes because of more or less chemical, or more or less water from flood or draught is immaterial to the organism. Both water quality relationships are important. We are in full agreement that more streamflow data are needed to accompany concentration data.	25	153
p. 35	The statistical manipulation in the first paragraph to produce a "true mean" is not valid. A 95% confidence interval would be more appropriate.	26	154
p. 36	The data in the table are interesting. What is the relation of this type of data presentation to the proposed periodic (5-10 years) intensive survey program?	27	154
p. 36-37	The comments on the fecal coliform variance are misleading. The statement concerning relationship of "sample means" to "the same category as the true mean" is not at all clear. The comparison of the mean plus or minus one standard deviation with a "quality category" does not provide useful information. Actually, there is enough information in the table to perform an analysis with a t-like statistic (t*) which shows the means not to be statistically different. If nothing else, this indicates that an analyst can perform statistical operations on variable data and draw conclusions. The fact that the means <u>appear</u> different but statistically cannot be shown to be different is the reason why eyeball analysis cannot substitute for scientific analysis.	28	154
p. 37	CEQ agrees that information systems have limitations.		
p. 39-43	Since Mr. John Ficke was the original author of the Annual Report discussion, we have asked him to respond. His remarks are attached as Appendix I. A few general comments are in order. a/	29	154

a/ The GAO consultant's evaluation of Mr. Ficke's comments on the case study is app. XI of our report. (see p. 157)

3. OWRT Comment

The GAO consultant's credentials and qualifications should be addressed in the report.

GAO Evaluation

We had included a brief statement of the consultants' qualifications in the draft report. We have expanded the statement for Mr. Horowitz and have added a similar statement for Professor Velz, who also provided us with valuable advice and assistance. (See pp. 4 and 5, vol. I.)

4. OWRT Comment

A number of illustrations have been provided in the report to support GAO's allegations. These data can lead to multiple interpretations, and the GAO has generally only given a single interpretation.

GAO Evaluation

We disagree. Again, since OWRT did not identify specific examples, we cannot respond in detail. We believe, however, that our conclusions are well-founded by the reported facts.

5. OWRT Comment

The case histories and examples which are addressed in the case study (app. VII) are exaggerated. Zinc is not an ideal water quality parameter to use as an example to discredit the Survey effort. Zinc contamination is a problem, and the behavior of zinc in natural waters is complex both geochemically and analytically.

GAO Evaluation

We do not agree that the case histories and examples are exaggerated, but agree with OWRT that monitoring zinc can be difficult. That is exactly the point that our report makes for zinc and other water quality characteristics that the Survey is attempting to monitor through NASQAN.

6. OWRT Comment

The Survey manages nearly 16,000 surface water stations, including the NASQAN sites. It is natural that some discrepancies will arise. But the report only emphasizes a few isolated instances of errors. A more comprehensive evaluation should be provided to fairly and accurately substantiate the GAO allegations.

GAO Evaluation

We disagree. Our report discloses the systemic weaknesses that pervade network-type water quality monitoring, including time bias, location bias, inconsistencies, and inability to account for the rapidly changing properties of water quality or for self-purification processes that exist in rivers. We did not rely on a few isolated instances but instead looked at data from many stations.

As we pointed out on page 12, volume I, the Survey does operate about 14,000 flow gaging stations with about 7,000 stations for various water quality purposes. We had no intention of reviewing the entire Survey water resources program. We specifically stated in the draft report that we focused our review on water quality networks. We do not believe it is necessary to evaluate the Survey's flow gaging program or the Survey's other water quality projects in order to assess the NASQAN program.

7. OWRT Comment

OWRT is concerned about the cost of data collection efforts mandated by law or done in response to law and is currently analyzing the question of data intensive models.

GAO Evaluation

We agree that cost should be considered in deciding what type of environmental monitoring effort should be done, but the quality of the resulting data should also be considered. As we have discussed in the report and in response to EPA comment 5 and Survey comment 15, a shift from network sampling to special studies could be done without increased cost and would result in better data.

8. OWRT Comment

A number of analysts have considered the problem of data networks, sampling intervals, and other topics relevant to the analysis and conclusions in the GAO report. It is difficult to tell whether this information was used inasmuch as it is not cited.

GAO Evaluation

In carrying out our analysis we reviewed many scientific and technical studies on water quality monitoring. For report presentation purposes we could not possibly have cited all of these studies, and therefore we were selective in our use of such literature. Appendix I of volume I contains a

 APPENDIX IX

bibliography of the major scientific and technical literature reviewed during our work.

EXECUTIVE OFFICE OF THE PRESIDENT
COUNCIL ON ENVIRONMENTAL QUALITY
722 JACKSON PLACE, N. W.
WASHINGTON, D. C. 20006

GAO EVALUATION

May 29, 1980

COMMENT No. PAGE(S)

Mr. Henry Eschwege, Director
Attn: Mr. Dave Jones
Community and Economic Development Division
United States General Accounting Office
Washington, D.C. 20578

Dear Mr. Eschwege:

I am responding to your April 29, 1980 request for the Chairman of CEQ to review the draft report "Better Monitoring Techniques Are Needed for National Surface Water Quality Assessment." Mr. Speth requested that I respond because of my earlier discussion with the GAO staff on this report.

General Comments

The draft report does a reasonably good job of pointing out some of the defects of existing water quality monitoring systems. Criticisms of this sort are helpful because they enable the agencies to receive feedback on the improvements they are making in response to the needs of Congress, other Federal agencies, and the user community in general. It is for this reason that earlier more substantiated criticisms, such as the two hearings before Congressman George Brown, Jr. or the 1977 National Academy study have been very useful. Some of the observations, for example, on data quality assurance made in this report are relevant to the ongoing efforts by EPA and USGS to enhance their capabilities in this area.

Unfortunately, the conclusion drawn in this report, namely that the known limitations of fixed station monitoring necessitate its replacement by so-called intensive surveys, is not substantiated by the information in the text. This is all the more so since much of the anecdotal evidence concerning fixed station monitoring defects (e.g. poor laboratory practices) would be equally applicable in a program consisting solely of intensive surveys. 1 144

CEQ's legislated mandate to report on environmental conditions and trends which it in part satisfies by its analyses in its Annual Report could not be met if the only data available were derived from a program of intensive local surveys which analyzed different watersheds each year, and only resurveyed a given body of water every several years. 2 144-145

Both fixed station and intensive survey monitoring have their strengths and limitations. CEQ would be strongly supportive of conclusions and recommendations which looked to improving water quality monitoring by taking advantage of all aspects of existing state-of-the-art monitoring in a cost effective fashion. We hope that the final version of the report will incorporate recommendations which will improve water quality data for analysis and decision making.

APPENDIX X

Mr. Henry Eschwege
Page 2

Specific Comments [a]

		GAO EVALUATION COMMENT No.	PAGE(S)
p. ii	The word "consequently" is inappropriate since the conclusion concerning representativeness does not follow from the practice of uniform sampling per month.	3	145
p. iii	The statement "The Survey has focused, primarily on quantity, not quality..." is harshly judgemental without substantiation.	4	145
p. iii	The quality assurance comments in the last paragraph are very imp.rtant. Since these problems are not unique to fixed station monitoring, the authors should indicate how they would be avoided if the recommended intensive survey program were implemented.	5	145
p. iv	The last sentence of the first paragraph is incorrect. There are many places in the text such as this, where a competent statistician should be consulted prior to preparing the final report.	6	146
p. v	In the second paragraph, the assumption is made that existing funds would permit the operation of "a systematically planned, comprehensive program of special studies." While this may very well be true, the paper should define the magnitude of such a program, indicate what constitutes "comprehensive," and provide a breakdown of costs. Given the substantive nature of the paper's recommendations, a detailed cost study of the recommendations is necessary to insure that they do not constitute an undue fiscal burden.	7	146
p. vi	The "other available indicators of national progress toward clean water" should be expanded upon. What are they? How are they obtained? Who collects them? What do they cost? How reliable are they? What is the spatial and temporal coverage? How do they relate to policy decisions?	8	146
p. 1	The third paragraph emphasizes that water quality changes downstream of a gaging station. This is also true of water quantity.		
p. 12	The analysis of CEQ's responsibilities should be modified. See Reorganization Plan Number 3 of 1970, Sec 2(5) as well as Sec 203(d)(7) of The Environmental Quality Improvement Act of 1970.	9	146
p. 13	While "EPA and the Survey face many technical, analytical, and practical difficulties...", in most cases these difficulties <u>do not</u> "...stem from the agencies' decision to use limited <u>sampling</u> programs...." For example, the case histories provided in this paper of delay in analysis of samples, incompatibility of parallel samples, etc. are not unique to fixed station monitoring. They are technical and management problems which would have to be addressed in any program, including one of special studies.	10	147

[a] Page numbers cited by CEQ refer to the GAO draft report.

APPENDIX X

Mr. Henry Eschwege
Page 3

		GAO EVALUATION COMMENT No.	PAGE(S)
p. 13	The second paragraph states a number of conclusions. The basis for these conclusions has not been presented at this point in the text. They should follow the analysis lest they appear undocumented.	11	147
p. 14	An appropriate citation to Velz' book should be provided. Quotation marks should be used where material is not paraphrased. Since Dr. Velz is used frequently as an authority in this draft, some general comments are in order: -- Many of his observations on the measurement of water are quite sound and meet with wide-spread agreement. -- As with any authority some of his opinions are not widely shared. As will be pointed out below, we feel that conventional statistical analysis, as performed in all modern scientific disciplines which measure data, permits sophisticated interpretation of data which are otherwise difficult to understand. We regard arguments by the authors of this report, in some cases based upon quotes by Dr. Velz, which claim limitations on statistical analysis capability, to be incorrect. -- Some of the quotes from Dr. Velz' book are provided out of context in this report. -- Some of Dr. Velz' comments which are applicable to the subject of his book, namely stream sanitation, are not applicable to the broader range of topics discussed in the draft GAO report.	12	147-148
p. 15	Provide citation for the 1973 report.	13	148
p. 16	Provide documentation for the sentence: "Cost was a major consideration behind the agencies' decisions on sampling frequencies."	14	148
p. 17	The statement from Dr. Velz: "...it is obvious that random sampling or sampling over an extended period of time is almost certain to reflect a series of distorted values of heterogeneous conditions" is misleading if not incorrect. Such statements as these should be reviewed by a competent statistician before the final publication is released.	15	148
p. 18	What is the source of the graph? Are these data real or fictitious?	16	149
p. 23	To be usable, the table should include the standard deviation, variance, or standard deviation of the mean.	17	149

Mr. Henry Eschwege
Page 4

		GAO EVALUATION COMMENT No.	PAGE(S)
p. 24	It is interesting to note that the Yadkin River case study would not have been detected without a fixed station already being present. The probability of a special study taking place simultaneously with the incident would be remotely small.	18	149
p. 26	The section on sampling errors is extremely important. The discussion should be extended to discuss the new EPA program on quality assurance.	19	150
	Is an intensive survey inherently different from fixed station monitoring as far as avoiding these problems?	20	150
	It is interesting to note that the fact that USGS discovered its own shortcomings is an indication that their technical audit capability works.	21	150
p. 27	Provide a citation for the statement: "Survey reviewers found fault...."	22	150
p. 28	In the last line, provide a basis for the statement: "We believe the problem may be chronic."	23	150
p. 31	The comment that "[sample timing and frequency]...undermine attempts to compare network-produced data..." is factually incorrect. The one-way classification linear model can be stated as $X_{ij} = \mu + \tau_i + \epsilon_{ij}$ As the data become less tidy, (to use a nontechnical term), the size of the term ϵ_{ij} (which reflects random variation) gets larger. Statistical theory has much to say about sampling design, sample size, and appropriateness of statistical tests which permit this fact to be taken into account to permit the analysis of data.	24	151-153
p. 34	The analysis provided on this page provides the opportunity to comment on the difference between concentration and loading in water quality analysis. Both concepts are extremely important. However, some readers may become confused by the mingling of the two ideas in the text. Loading refers to the quantity of material, as for example, so many tons of arsenic are passed into the Mississippi from the Missouri each year. This concept is important for computing mass balances and similar types of analyses. CEQ, for example, has computed the U.S. contribution of heavy metals to the marine environment in this fashion. Flow measurements and concentrations are necessary for computing loading.	25	153

Mr. Henry Eschwege
Page 5

		GAO EVALUATION COMMENT No.	PAGE(S)
	Concentration is the important physiological parameter. An organism at the Missouri-Mississippi junction does not react to the loading, it reacts to the concentration (for example does the concentration expressed as mg/l surpass its tolerance). (As another example, drinking a half pint of sea water would be very unpleasant. If it were diluted in two gallons of distilled water and consumed gradually, it would not be unpleasant In both cases the amount of salt -- the loading -- is the same, while the concentration is very different.) Water quality violations are based on concentrations, not loading. As far as an organism is concerned, concentration is the basis of physiological, toxicological or behavioral effects. Whether concentration changes because of more or less chemical, or more or less water from flood or draught is immaterial to the organism. Both water quality relationships are important. We are in full agreement that more streamflow data are needed to accompany concentration data.	25	153
p. 35	The statistical manipulation in the first paragraph to produce a "true mean" is not valid. A 95% confidence interval would be more appropriate.	26	154
p. 36	The data in the table are interesting. What is the relation of this type of data presentation to the proposed periodic (5-10 years) intensive survey program?	27	154
p. 36-37	The comments on the fecal coliform variance are misleading. The statement concerning relationship of "sample means" to "the same category as the true mean" is not at all clear. The comparison of the mean plus or minus one standard deviation with a "quality category" does not provide useful information. Actually, there is enough information in the table to perform an analysis with a t-like statistic (t*) which shows the means not to be statistically different. If nothing else, this indicates that an analyst can perform statistical operations on variable data and draw conclusions. The fact that the means <u>appear</u> different but statistically cannot be shown to be different is the reason why eyeball analysis cannot substitute for scientific analysis.	28	154
p. 37	CEQ agrees that information systems have limitations.		
p. 39-43	Since Mr. John Ficke was the original author of the Annual Report discussion, we have asked him to respond. His remarks are attached as Appendix I. A few general comments are in order. a/	29	154

a/ The GAO consultant's evaluation of Mr. Ficke's comments on the case study is app. XI of our report. (see p. 157)

Mr. Henry Eschwege
Page 6

		GAO EVALUATION COMMENT No.	PAGE(S)
	-- CEQ does not report individual observations, but rather aggregates the information into a national analysis. We did not report on the Cartersville station by itself. It appears as a single shaded polygon in a map of the United States, along with 358 other polygons. -- CEQ made a decision to report concentrations rather than loadings because this is the basis of physiological effects on organisms and because water quality criteria are expressed as concentrations. -- NASQAN maps in the Annual Report display grouped violation frequencies that damp out all but major differences in water quality.	29	154
p. 44	We agree that sampling frequency could be increased to provide increased statistical sensitivity (not validity).		
p. 45	The term "statistically significant data" is imprecise. A significance (probability) level is chosen when a null hypothesis is erected concerning the relationship among two or more sets of data.	30	154
p. 45	The statement "Although we cannot determine the precise number and extent of extra sampling..." sells the analysts' tools short. All that is needed to calculate the sample size for the confidence interval (e.g. 95%) of a mean, is the variance. The variance itself is easily calculated from the existing data (no matter how variable it may be).	31	155
p. 47	The statement "More scientifically sound assessment and supporting data can be produced through special water quality studies" is unsupported by the information in the draft GAO report. The extrapolation from a small handful of special purpose studies to a national program requires a more detailed discussion which hopefully will be provided in the final GAO report.	32	155
p. 47	The use of information such as changes in "fish population and diversity or reductions in discharge of contaminants" can be very helpful in analyzing water quality. However, these data are difficult and expensive to obtain as a high quality data set. It would be very useful for the final report to explore how such information should be gathered, and what data bases would have to be supported to make the data available. CEQ would certainly encourage the availability and use of such information.	33	155-156

Mr. Henry Eschwege
Page 7

		GAO EVALUATION COMMENT No.	PAGE(S)
p. 50	We disagree with Dr. Velz' recommendations and opinions in the middle of this page.	34	156
p. 53	The cost data provided for special studies is difficult to believe. Because this draft GAO report recommends essentially the abolition of several federal programs and creation of at least one new one, a detailed fiscal and budget analysis should be prepared.	35	156
p. 54	The question of available manpower (contractor and civil service) for the proposed program of special studies should be discussed.	36	156
p. 56	As noted above, the use of "reductions in discharges of pollutants as a water quality indicator" deserves further discussion of how it could be implemented.	37	156
App. VI	CEQ commented in an earlier letter on the GAO contractor's report. A copy is attached to this response as Appendix II. a/		

Conclusion

It is the nature of a review to be critical. We hope that these comments will be useful for the production of your final report. However, we want to emphasize that CEQ is committed to improving water quality monitoring and the quality of data it produces. We agree with many points made in the text, and feel you have raised important concerns with some of your observations, particularly as they related to quality assurance. We also agree that a strong program of special studies should be encouraged and that some of the money currently used for fixed station monitoring could be better used for special studies. However, to cast an analysis in terms of one form of monitoring or another, when each provides uniquely important information, is to create a dichotomy where none exists.

I am, of course, willing to further discuss my comments, your draft report, or any related matter with your staff at their convenience.

Sincerely,

John D. Buffington
Senior Staff Member
Environmental Data & Monitoring

cc: Gus Speth

a/ GAO Note: The earlier CEQ letter of Jan. 11, 1980 was considered in preparation of the case study.

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON, D.C 20460

May 29, 1980

(See GAO notes a and b below.)

OFFICE OF TOXIC SUBSTANCES

Dr. John D. Buffington
Senior Staff Member for
Environmental Data & Monitoring
Council on Environmental Quality
722 Jackson Place, N.W.
Washington, D.C. 20006

Dear Doug:

This is in response to your request to Jack Pickering of the U.S. Geological Survey (USGS) that I provide comments on the draft of a proposed report by the U.S. General Accounting Office (GAO) entitled, "Better Monitoring Techniques Are Needed for National Surface Water Quality Assessments." Although I have many reactions to material in the GAO draft that is distorted or just plain wrong, I will comment mostly on GAO's discussion of material used for the 1978 Annual Report (AR) of the Council on Environmental Quality.

Analysis of Changes in Water Quality

The GAO draft (pages 39-43 and 63-102) dwells at length on the analyses and data used in constructing Table 2.1 (page 96) of the 1978 AR, and in the brief paragraph of discussion on pages 96 and 98 of the AR. Table 2.1 was constructed using the following procedure:

1. Data from 357 stations of the National Stream Quality Accounting Network (NASQAN) for 10 water-quality characteristics (fecal coliform bacteria, inorganic nitrogen, organic nitrogen, total phosphorus, dissolved oxygen, fecal streptococci bacteria, dissolved solids, dissolved zinc, total zinc, and phytoplankton) for the 1975, 1976, and 1977 water years were subjected to an analysis of variance (ANOVA) to determine whether there were significant differences in the annual mean values of the 10 characteristics.

2. The computations for the ANOVA were done by a contractor who computed an "F" statistic based on the sums of squares within and between the annual data sets for each constituent at each station.

a/This letter was provided by CEQ as app. I to CEQ's response (see p. 129) to our draft report.

b/Most of Mr. Ficke's comments relate to our discussions of material used for the CEQ 1978 annual report. The material from the 1978 report was used in conjunction with the Cartersville, Va., case study in app. VII of vol. I, prepared by Mr. Jerome Horowitz, our consultant. Mr. Horowitz's evaluation of Mr. Ficke's criticisms of the case study, as well as the criticisms of the agencies, is contained in app. XI of this vol. Mr. Ficke's other comments have not been addressed specifically because they are either rhetorical or have been evaluated with respect to the agencies' comments.

3. I examined each set of data (each characteristic at each station) to determine (a) if the F statistic indicated a significant difference among the means and (b) if the differences represented a clear pattern of improvement or worsening of water quality. In comparing the F values, I used the numbers of degrees of freedom computed by the contractor, and F values to test for significance at the 90 percent level. Each characteristic at each station was labelled either "+" for improved quality, "-" for worsened quality, or "0" for no change.

4. With the help of a contractor employee, I compiled the changes and computed the percentages of the stations that had improved, not changed, or deteriorated, for each characteristic.

Criticisms by GAO

GAO staff reviewed all of the above procedures and was clearly aware of how the work was done. The GAO draft was not critical of the method of computation. They agreed that there indeed were changes in dissolved solids, zinc, and dissolved oxygen at the James River station, which was only one of the 357 stations used for the CEQ analysis. The GAO draft, however, was critical of the use of NASQAN data set for statistical analyses without regard for the basis of the changes or for possible problems in the data set. The following paragraphs respond to some of the GAO criticisms.

Validity of "trends". GAO repeatedly states that the CEQ Annual Report claims there were trends in water quality. In fact, the 1978 AR does not refer to "trends," but only to "changes." To me, a trend is a pattern that may be expected to continue, but the reported change is simply an observation of what happened. Again, the James River station for which they did detailed analyses was only one of a large set of stations used by CEQ.

Explanation of reasons for changes. GAO goes to great lengths to argue that many of the changes in water quality in the James River were caused by climatic factors. They may be correct. In fact in the 1977 AR of CEQ we pointed out that the drought conditions in many places may have affected water quality. We probably should have repeated some of that material in the 1978 AR, but space was short. A big point that GAO misses is that for many users of water (e.g., fish, municipal and industrial users) the reason for change is not important. They simply want to know what's going on. Part of CEQ's requirement under NEPA, also, is simply to report conditions and changes.

Explanations offered by GAO. Appendix VI of the GAO draft goes on at length with discussion of reasons for long-term and short-term variations in water quality, but much of the logic is faulty. It suggested that decreases in flow rate caused increases in dissolved solids concentration. But the discussion only shows coincidence, not cause and effect. They did not even test for a significant regression, much less show cause. There also are many claims that variations in concentration of zinc and dissolved oxygen or changes in temperature do not follow reasonable patterns. The text suggests, therefore, that the data are bad. What GAO really is showing, however, is that the James is a complex river and that water quality is influenced by many complicated, interrelated factors. It's not predictable -- if it were there would be no need to monitor or conduct special studies or intense surveys. In reality, however, it would be necessary to do special studies of considerable magnitude to establish the kind of predictive modeling capability that is needed to make the type of predictions that GAO would like to have, or claims to have, in faulting the actual data.

GAO's Recommendation

CEQ would not be well served if GAO's proposal to discontinue fixed station monitoring were adopted. The Council would be right back to where it was before 1975 — trying to scrounge through STORET to find adequate and reliable data to describe conditions and changes. The 1975 AR summarizes the result of one special study of trends done by the USGS, in which only 88 stations were found with suitable runs of data to support analyses of changes or trends.

CEQ's needs to describe national conditions and trends would not be met by results of intensive surveys, each designed differently to explain what is affecting water quality. The proposal of Velz to repeat surveys each 5 years would not work either; all you would have then is a network with a sampling frequency of once each 5 years.

GAO's analysis of the month-to-month, season-to-season, and year-to-year variations of quality of the James River is a good argument as to why a program of one-shot intensive surveys will not work. When would they have done the James -- 1975, 76, 77; spring, autumn, or winter? Would they have gone to the expense of adapting, verifying, and calibrating a model to explain all of the variations pointed out in Appendix VI? If they had done it, how would CEQ use the data in its Annual Report?

Strong Points in the Report

Just because the GAO draft contains conclusions that are unsupported and unacceptable to CEQ does not mean that there are not many strong points. Many are similar to points already raised by the National Commission on Water Quality, the National Academy of Sciences, the Interagency Task Force on Environmental Data and Monitoring, and several other special studies, plus staff and contractor studies for EPA, CEQ, and USGS.

Indeed, the agencies need to continue to stress quality assurance. Matters of sampling frequency and extremes need to be resolved. Variations within basins need to be studied. Biological monitoring needs to be improved. The agencies need to do more to interpret the data they collect. Coordination among agencies and at different levels of government needs to be improved. Nasty problems of time-series analyses need careful research.

Detailed Comments

The following notes pertain to particular parts of the GAO draft:

Page	Comment
i	The last paragraph also should mention seasonal patterns of change, and annual changes for natural and anthropogenic causes.
ii	USGS staff statisticians have long argued that sampling at fixed intervals provides better data for trend analyses. The report does not mention USGS policy of extra samples for extreme events.

It also does not mention the diurnal measurements of DO, conductance, and pH at NASQAN stations. Discussion of cost does not consider that cost is mostly a product of number of sites times frequency multiplied by number of measured constituents in addition to interpretation for either networks or intensive surveys.

iii GAO should support its first sentence concerning "quantity not quality." The first paragraph should explain the principle of accounting networks. In the last paragraph, GAO should support why sampling and lab work would be better for intensive surveys than for networks.

iv If network data are questionable, why would not data from intensive surveys be questionable too? GAO should acknowledge that USGS had only 88 stations for limited analyses when NASQAN started. The fourth paragraph should state how special studies will measure changes and trends.

v See comment for page ii regarding trends. GAO should show how: "The nation-wide perspective on progress toward cleaner water could be achieved through ... special studies." They should show how CEQ could use the data to satisfy NEPA. They should acknowledge USGS and EPA work on biological monitoring -- strengths and weaknesses.

vi Which available indicators are proposed? How good are they? How do they relate to criteria and standards?

2 Paragraph 4 notes variations in special studies. How should CEQ use such data to report national conditions and changes?

4 The report should repeat the long list of examples provided by USGS on the use of data from fixed stations. Concerning the last sentence on the page, it should be pointed out that this draft has not been reviewed by "experts in the academic field."

7 GAO should explain how its proposed use of intensive surveys will "determine the actual quality status of all waters of the Nation" (4th paragraph).

8 USGS never operated all of NWQSS. It was up to EPA regional offices to decide.

10 NASQAN stations always have and still do receive significant reports from states through the USGS cooperative program.

11 Discussion of cost of NASQAN (last paragraph) should point out that in 1973-74 USGS estimated $10,000 per station for operation, in 1973 dollars. USGS has compensated for inflation by greatly reducing analyses of data, research, and interpretive reporting.

12 Use of NASQAN data goes back to the 1975 CEQ Annual Report.

13 None of the discussion on this page supports how intensive surveys would produce better data to meet the needs of CEQ and other data users.

14 Diurnal measurements at NASQAN stations should be described. GAO should acknowledge other experts than Velz, whose book was written in the 1960s and published in 1970. There is a large body of literature and a lot of thinking has been done in the last dozen years.

15 GAO should acknowledge NASQAN instructions for representative sampling of a cross section. The criticism of not evaluating cumulative impact should be a recommendation. The report should cite sources of quotations and concepts. Use of monitors at many NASQAN stations should be reported.

16 The "memorandum on doubts" is cited incompletely and out of context. See earlier comment for page ii regarding trend analyses.

17 Discussion should note that DO is not on the NASQAN list (Appendix III). The whole matter of DO and NASQAN deserves discussion in the GAO report. It demonstrates the USGS awareness of potential problems with DO data.

18 GAO should give the source of the data displayed.

21 The report should discuss relative costs of the detailed visit described at the top of the page. Discussion should point out how problems of variations over time will be solved in order to make long-range trend analyses. The matter of coordination (middle paragraph) should be made into a recommendation.

24 Citation of the Yadkin River incident should point out that this was discovered during a visit to a fixed station. USGS cited this case in the list of examples provided to GAO (see my comment regarding page 4). The footnote on page 24 should acknowledge NASQAN instructions for sampling.

25 GAO should support the "not quality" statement (14 lines from bottom).

26 GAO should state why improper operations will be less of a problem for intensive surveys than for fixed stations.

28 GAO should explain why the busy-season backlog will be less for intensive surveys than for fixed stations.

31 It is not clear why laboratory performance will be better for intensive surveys than for fixed stations.

34 Data here make more of a case for repeated measurements than for one-time intensive surveys.

35 Data for St. Paul also support repeated measurements instead of one-time surveys that would describe only one set of conditions.

36 GAO needs to explain how CEQ could do better with data from a number of dissimilar intensive surveys.

39 GAO does not acknowledge new quality assurance programs of EPA. Nor do they explain how STORET and WATSTORE would do better for intensive surveys than they do for networks.

42 GAO needs to explain why an intensive survey, run maybe once during the 3-year period, would have done better at defining water quality under changing flow conditions than did a series of monthly measurements.

43 Discussion at the top of the page seems to support the CEQ analyses of improved conditions with respect to zinc. In the last paragraph, GAO should describe diurnal measurements at NASQAN stations, and the large number of automatic monitors at network stations.

44 GAO needs to describe the whole complex problem of GC-MS analysis and calibration for analyses of organic compounds. EPA now is working on redesign of the whole scheme for monitoring pesticides.

50 GAO should give limitations and cost data for the types of analyses touted by Velz. GAO implies that networks would substitute for thorough studies for management purposes. Here they should cite the Langford (OWDC) concept of function and level. The suggestion of Velz for "one good short period of intensive sampling" contradicts all the discussion of the variation in Appendix VI. It would be hard for CEQ to develop a national overview of water quality using only a series of different studies of the type done for Redwood National Park. GAO should again consider the concepts of function and level.

51 In the fourth paragraph, "researchers" are mentioned. In reality, the Redwoods study was more of a research operation than a routine survey. Again though, how does CEQ get national conditions and trends from a series of research studies?

55 CEQ has used, would like to have, and will continue to use the types of indicators given on this page. They, however, are hard to validate, hard to interpret, and hard to do well.

56 Much of the biological monitoring of pesticides is based on samples from fixed stations. GAO has failed to tell CEQ how to do good national analyses without data from fixed stations. River quality assessments of the type done by USGS on the Willamette, Chattahoochee, and Yampa just won't give CEQ what it needs. The last paragraph on page 56 should be stressed as a recommendation.

57 If the recommendation is followed, CEQ will return to days of pre-1975 and will find it extremely hard to meet requirements of NEPA.

58 The 1978 CEQ Annual Report uses "changes," not trends. See my earlier comment.

69 The CEQ only reported gross changes for more than 350 stations. It did not attempt to explain why for each one.

71 GAO does not support that the change in dissolved solids is a "normal consequence of a deepending deeping drought." It only shows that they were coincident. Neither is the 10-year trend contradictory. It is possible to have a 3-year change one direction within a 10-year trend in the opposite direction. Such phenomena are common in most environmental and economic variables.

73 Material on page 73 is not at all related to the validity of the CEQ data in the 1978 Annual Report, nor does it seem relevant to the matter of fixed stations versus intensive surveys. In fact, however, the relation between conductance and dissolved solids can vary as a function of ions and amount of silica.

75 Data repeated by GAO support the use of this station in its count of stations showing an improvement as far as zinc concentration is concerned.

77 The second sentence does not recognize the concentration of 40 micrograms per liter at a flow of 3,564 cfs. The fourth sentence ignores that "midnight dumping" or washout of holding ponds often produces heavy point-source loads at time of high flow.

78 The third paragraph is untrue. Annual averages are 68, 18, and 13, using only USGS data, and 48, 17, and 14, using both state and USGS data. CEQ's conclusion would have been the same, even if it had the state data.

80 The first paragraph is untrue. Five of 11 values of "0" or "10" are at flows greater than 5,000 cfs. The last paragraph shows that GAO assumes a very simple model, and rejects data that contradict the model. If nature were as simple as these concepts and models, we never would need monitoring or intensive studies.

83-100 Most of this discussion of DO is another case of GAO assuming very simple models for very complex situations. It is ridiculous to question the data just because they do not fit the analyst's preconceived ideas. As far as CEQ is concerned, GAO agrees (p.93) that USGS data show an improvement and that we were correct in counting it as we did for the table in the 1978 AR. I might make one note here, however. The changes were small and the F value was not far above the criterion for 90 percent confidence. This might have been one case where the statistical analysis gave a false positive, as it will do 10 percent of the time, on the average.

APPENDIX X

100-102 These conclusions merely restate some poorly supported cases made by GAO regarding effects of a drought, relations between conductance and dissolved solids, the matter of trend versus change, calling data "suspect" if they do not fit a preconceived notion, and simplifying a situation that really is very complex. On this last point, the last paragraph on page 102 is partly correct, and partly wrong. Just because the data are not understood by the investigator does not make the data wrong. I agree that more technical work would be helpful, but it is out of line to charge that CEQ should not use the data to describe national conditions and changes in water quality.

Sincerely,

John F. Ficke

John F. Ficke

cc:
R.J. Pickering, USGS

EXECUTIVE OFFICE OF THE PRESIDENT
COUNCIL ON ENVIRONMENTAL QUALITY
722 JACKSON PLACE, N. W.
WASHINGTON, D. C. 20006

(See GAO note below.)

January 11, 1980

Mr. Arthur M. Peterson
Supervisory Auditor
United States General Accounting Office
415 First Avenue North
Seattle, Washington 98109

Dear Art:

It was a pleasure to meet with you and Dr. Horowitz. As a followup to our discussion I think it is probably useful to reiterate some of our conversation.

The Council on Environmental Quality performs analyses of the conditions and trends of environmental quality each year for our Annual Report. Air and water quality are emphasized because of their all embracing importance to man and to his environment. Furthermore they have been the subject of detailed legislation.

Because we are attempting to potray the national level situation, we are severely limited by the availability of data which is both synoptic in time and national in scope. For this reason we rely extensively in the water area on the data which is stored in WATSTORE and STORET. The single data collection system which is highest in quality and is designed to have national coverage is the NASQAN system of the Department of the Interior. NASQAN can be accessed through both data systems.

CEQ, because it so extensively exercises these systems, is well aware of their shortcomings. We work closely with individual program managers within EPA and USGS as well as with OMB to improve these systems. I am enclosing a letter I sent to Dr. Marilyn Bracken of EPA in which I drew the attention of the EPA group of Deputy Assistant Administrators dealing with monitoring and data problems to some of CEQ's concerns.

Over the past year that Dr. Gevantman was on our staff on detail from NBS, he worked with both the EPA and USGS staff attempting to identify what kinds of policy questions the agencies, monitoring programs should be able to address. This initiative sprang from Dr. Gevantman's experience with the CEQ Annual Report, since these are questions which we would want to examine in future issues. I am enclosing a draft set of these questions for your information.

GAO note: This letter was received before our request for CEQ comments on our draft report. These comments were considered earlier and, if necessary, appropriate changes were made to the case study (app. VII of vol. I). Also app. XI of this vol. contains our consultant's evaluation of the agencies' criticisms of the case study. CEQ provided this letter as app. II to its response (see p. 132) to our draft report.

APPENDIX X

A variety of issues surfaced in our discussion concerning the use of NASQAN data for national level analysis. CEQ's use of these data is predicated on our confidence in the precision and accuracy of these data. The study by Dr. Horowitz attempted to cast doubt on the validity of using NASQAN data because of their "incredibility", a term he used in our discussion. The difficulties indicated by Dr. Horowitz in reconciling the data gathered by USGS with data obtained from state agencies, as well as with standard expectations of how water quality parameters should behave can be interpreted in one of two ways: 1) the data are intrinsically unreliable or 2) the data are reliable but we do not understand the relationships they portray.

Unfortunately the draft study you were kind enough to send me did not distinguish between these two problems. In fact only some of the data that were compared were in conflict. The problem of unreliable data can be dealt with by a quality assurance program such as EPA is attempting to implement in order to improve its environmental data. I consider this to be a fundamental issue. Good analysis must rest on reliable data.

I am much less concerned than is Dr. Horowitz by being forced to confront data that do not conform to my expectations as to what the expected relationships of those data should portray. Any of us who have gathered environmental data of any kind have come to realize that the universe is stochastic and not deterministic. Scientists are continually puzzled by anomalies in good data. It is the accumulation of the record of these anomalies which permits eventual explanation and deeper understanding of how the environment functions. This is why I feel that the reliability of the data is the fundamental question. If the data are good, our inability to interpret them now should not lead to the abandonment of the data collection effort, but rather to an attempt to understand why our conceptual models fail.

As to the actual use of an existing set of fixed station water quality data to draw conclusions about national trends, I would like to restate my views presented in our discussion:

1. The NASQAN system has the virtues of being national in scope, synoptic in presentation, several years in depth and is generally recognized as being reliable. To the extent that your study raises questions about reliability, I shall certainly raise this issue with USGS.

2. Station location of NASQAN is not ideal for analyzing all policy issues. Undoubtedly many local events of great importance are missed. Population centers are not as heavily weighted as perhaps might be desirable. If CEQ were locating the stations only for Annual Report analysis, I am sure we would put some stations at different locations. However, NASQAN has an internal logic and philosophy associated with station location. While its historical roots are in hydrographic data acquisition, this is insufficient reason to dismiss the system out of hand. Several USGS publications

discuss the basis of the NASQAN system and if you do not have them alredy, I am sure the Office of Water Data Coordination will provide you with them. EPA is now putting together a core set of monitoring stations which may satisfy some of our problems of the limitation of NASQAN station location.

3. The picture provided by the data retrieved from NASQAN does present a useful overview of the nation's water quality. It is imperfect for the reasons we discussed. However no monitoring system can answer all questions. Any approach to monitoring becomes imperfect when the data are used in a fashion somewhat removed from the design basis of the system (this is equally true of survey type monitoring). By the use of repetitive maps for the available data years, now five, a pattern is revealed which is informative to the reader. The picture is certainly incomplete, but nevertheless there is an emergent pattern. Because the picture is incomplete, CEQ amplifies our analysis of water quality by focusing on special problems in more narrow geographic areas. In the past for example we looked at phenols on the Ohio River. This year, Great Lakes water quality will be discussed. For these special analyses a wide variety of data are used. In this way NASQAN analysis provides the overall pattern, but more detailed analysis amplifies those topics where the NASQAN system was not designed to address them. Hence we feel that NASQAN serves as a useful base toward an improved monitoring philosophy.

4. Our discussion touched on the problem of fixed station monitoring of non-conservative water quality parameters such as D.O. (I believe USGS has indicated that sensitivity limits of standard tests answers the question of the absence of chlorophyll when algae are present. The question as to whether more sensitive test should be employed is certainly valid.) Again I would agree that fixed stations have inherent limits for identifying some water quality problems at an instantaneous measurement time. Fixed station data may be useful in identifying or analyzing local problems, but in many cases a detailed water quality analysis would be necessary. CEQ does not attempt to identify local problems with its use of NASQAN data. The pattern that emerges through the use of standard statistical techniques on such data provides a useful overview of nationwide changes that have become apparent over the period of record. Local or regional water quality problems to which the NASQAN analysis is insensitive for any reason would then be treated as case studies to the extent that their importance warrents discussion in the Annual Report.

Again let me emphasize that the 349 NASQAN accounting unit and their several years of record are not used in isolation to identify national trends. The water quality chapter of our Annual Report uses a wide variety of sources in addition to NASQAN each year.

Let me close this overly long letter with some general comments. CEQ, executive branch agencies, the congress, states and industries all have identified problems with environmental data and monitoring including water quality data. The situation now is much improved over what it was several years ago because each of us has worked from our own perspective to improve the situation. A major problem of course is that each agency generates data to satisfy its own legislated mandates. However data are then frequently used outside the context for which the data were originally generated. Activities such as those provided by OMB Circular A67 and a variety of intra-agency initiatives have resulted in a greatly improved situation. There is certainly a continuing need to see how various agency programmatic approaches such as both fixed station networks and intensive surveys can most effectively serve the nation's goal of improved environmental quality.

Whatever your final recommendations might be, I hope they reflect the complexity of the situation as well as the differences in expert professional opinion that exist among water quality analysts. The seductiveness of a simple solution to any problem usually leads to eventual additional complexity.

Sincerely,

John D. Buffington
Senior Staff Member
Environmental Data and Monitoring

cc: G. Speth, CEQ
B. Harris, CEQ
E. Strohbehn, CEQ
R. H. Langford, USGS
F. Leutner, EPA
L. H. Gevantman, NBS

GAO EVALUATION OF COUNCIL ON ENVIRONMENTAL QUALITY COMMENTS

1. CEQ Comment

The conclusion drawn that fixed-station monitoring should be replaced by intensive surveys is not substantiated, since much of the anecdotal evidence concerning fixed-station monitoring defects would be applicable to intensive surveys.

GAO Evaluation

We disagree that the conclusion is not substantiated. Our evaluation of EPA comment 11 discusses this matter at length.

2. CEQ Comment

CEQ's legislated mandate to report on environmental conditions and trends could not be met if the only data available were derived from a program of intensive local surveys.

GAO Evaluation

We fail to see the rationale of CEQ's claim that it could not meet its legislated mandate if it only had available the results of special studies. CEQ's legislated mandate, which is cited on page 13, volume I, does not specify that CEQ must use networks of fixed stations or that it must report every year on the quality of all rivers in the country. The National Environmental Policy Act does state that CEQ is "* * * to gather timely and authoritative information concerning the conditions and trends in the quality of the environment."

We are not suggesting that CEQ rely solely on special water quality studies for its annual reports. CEQ has for years used many types of information other than network water quality data that give revealing national perspective on efforts to improve water quality. In fact, we recognized in the draft of our report the value of using other indicators of progress. We recommend in the report that the Chairman of CEQ and the Administrator of EPA promote the use of these indicators.

CEQ's annual environmental reports give the impression that network sampling provides national coverage, particularly through the maps that classify all geographic areas in the Nation according to sampling results at NASQAN stations. This is misleading, however, since vast areas are improperly represented by sampling results at single sites and many rivers are not even sampled. In addition, as we have demonstrated in the report, even the water quality measurements and averages for the sampled sites are unreliable because of various time and location biases, inconsistencies, weaknesses in field and laboratory work, and statistical deficiencies in the data.

We believe CEQ cannot fulfill its legislative mandate by using network sampling results, but, the mandate could be satisfied with soundly conceived special studies.

3. CEQ Comment

The word "consequently" is inappropriate because the conclusion concerning representativeness does not follow from uniform sampling.

GAO Evaluation

The text of the report (see p. ii, vol. I) has been changed to reflect this comment.

4. CEQ Comment

The statement "The Survey has focused primarily on quantity, not quality * * *" is harshly judgmental without substantiation.

GAO Evaluation

The text of the report (see pp. ii and 27, vol. I) has been changed to better describe the criteria used by the Survey to site NASQAN stations.

5. CEQ Comment

Quality assurance is very important and the problems are not unique to fixed-station monitoring. The report should indicate how they would be avoided if the recommended special studies programs were implemented.

GAO Evaluation

We agree that quality assurance needs close attention in any type of monitoring effort. As discussed in chapter 4, however, we believe quality assurance can receive closer attention in special studies because potential problems can be specifically identified and extra measures to mitigate or eliminate them can be taken. We also discuss this matter in response to EPA comment 22.

6. CEQ Comment

The last sentence of the first paragraph on page iv of the digest (vol. I) is incorrect. A competent statistician should be consulted prior to preparing the final report.

GAO Evaluation

The text of the report (see p. iii, vol. I) has been changed.

7. CEQ Comment

The assumption is made that existing funds would permit the operation of a systematically planned, comprehensive program of special studies. Given the substantive nature of the recommendations, a detailed cost study is necessary to insure they do not constitute an undue fiscal burden.

GAO Evaluation

The text of the report (see pp. iv, 61 and 62, vol. I) states that cost need not increase. Nothing would prohibit the agencies from designing a comprehensive program within fiscal constraints. We believe it is not incumbent upon GAO to perform a detailed cost study or to state the number of studies that should be done. Such an activity is a normal function and responsibility of the agencies.

8. CEQ Comment

GAO's discussion in the digest of other available indicators of national progress should be expanded upon.

GAO Evaluation

Indicators such as biological monitoring and reductions of waste discharges are discussed in the text of the report (see p. 62, vol. I) and several types are mentioned in the digest. As noted, they have been used by CEQ and others. To describe the indicators in detail in the digest is not necessary or practical.

9. CEQ Comment

The analysis of CEQ's responsibilities should be modified.

GAO Evaluation

The text (see p. 13, vol. I) has been changed based on the information cited.

10. CEQ Comment

The technical, analytical, and practical difficulties and management problems cited would also need to be addressed in special studies.

GAO Evaluation

We agree. As we discussed in the report and in our evaluation of EPA comment 11, these problems can be overcome through special studies but not through network monitoring.

11. CEQ Comment

The second paragraph of page 14 of the text states a number of conclusions, the basis of which has not yet been presented. The conclusions should follow the analysis lest they appear undocumented.

GAO Evaluation

We do not agree. This is a matter of format and style. These matters are subsequently discussed.

12. CEQ Comment

Appropriate citations to Professor Clarence J. Velz's book should be provided, including quotations where necessary. Also, some of Professor Velz's observations are quite sound and accepted, whereas others are not widely shared, particularly some of the claimed limitations on statistical analysis capability. Some arguments used from Professor Velz's work may be taken out of context, are incorrect, or not applicable to the subject matter.

GAO Evaluation

As discussed in the scope of our review, Professor Velz provided us with substantial assistance during our work. Professor Velz' credentials are discussed in the scope of the review on page 5, volume 1. Professor Velz agrees with our use of materials from his book and other writings. We particularly cannot agree with CEQ's comment that some of Professor Velz's material we cited may not be applicable to our review. Professor Velz's book applies to stream sanitation in the broadest sense, covering the four primary sources of pollution (urban, industrial, agricultural, and natural) and explains the complexities of self-purification and interactions of a wide range of foreign substances that enter rivers and streams.

With respect to the specific comment concerning Professor Velz's claimed limitations on statistical analysis capability, it is interesting to note that some of the Geological Survey's publications have cautioned against the misuse of statistics. In Survey Circular 715-D, it is specifically stated that:

> "The application of statistics to river-quality data analysis is often based upon dubious assumption of independence of observations and randomness of sampling. However, * * * it is only within particular seasonal periods and with more phologically similar reaches of a river that river-quality phenomena can be expected to exhibit the homogeneous or stable ecological conditions that, even in the most optimistic sense, are suited to such mathematical and statistical assumptions."

13. <u>CEQ Comment</u>

Provide citation for the 1973 report.

<u>GAO Evaluation</u>

A citation has been provided in the text. (See p. 16, vol. I.)

14. <u>CEQ Comment</u>

Provide documentation for statement that cost was a major factor in the agencies' decision on sampling frequencies.

<u>GAO Evaluation</u>

The statement is explained by the next sentence in the report. In addition, the Survey clearly explained in its comment 34 that it chose monthly sampling for DO because it was the least costly option, although the Survey recognized that monthly sampling of DO was inadequate.

15. <u>CEQ Comment</u>

The statement from Professor Velz concerning distorted values of heterogeneous conditions is misleading if not incorrect. A competent statistician should review such statements.

<u>GAO Evaluation</u>

The quotation is correct as written. However, for clarity, we have added an explanation in the report that the statement concerned the importance of analyzing water quality in stable, similar conditions. (See p. 18, vol. I.)

16. CEQ Comment

Give the source of the graph on p. 20. Are these dates real or fictitious?

GAO Evaluation

The graph included in the draft report was developed by GAO for illustrative purposes only. A new graph has been provided in the text (see p. 20, vol. I) based on actual data.

17. CEQ Comment

To be usable, the table on page 26 should include the standard deviation, variance, or standard deviation of the mean.

GAO Evaluation

We disagree. The table shows data on fecal coliform bacteria at four locations on the South Platte River in Colorado. Because the data are open-ended (e.g. "less than 30") it is not correct to calculate a standard deviation or variance. In addition, EPA correctly informed us that arithmetic means are not appropriate for characterizing fecal coliform bacteria levels (see EPA comment 13). We have substituted median measurements for the arithmetic means.

18. CEQ Comment

The Yadkin River case study would not have been detected without a fixed-station already being in place. The probability of a special study taking place simultaneously with the incident would be remotely small.

GAO Evaluation

The situation was detected by a continuous monitor, not through periodic sampling at the fixed station. We do not disagree that the probability of the incident occurring during a special study may be small, but the probability of normal (generally monthly) network sampling detecting the incident is small also.

19. CEQ Comment

The section on sampling errors is important and the new EPA program on quality assurance should be discussed.

GAO Evaluation

The EPA quality assurance program is discussed on page 40, volume I.

20. CEQ Comment

Is an intensive survey inherently different from fixed-station monitoring as far as avoiding sampling and quality assurance problems?

GAO Evaluation

Yes. See our evaluation of EPA comment 11.

21. CEQ Comment

That the Survey discovered its own shortcoming is an indication that its technical audit capability works.

GAO Evaluation

We did not challenge the technical audit capability of the Survey. We only reported what had been found, which revealed many errors in field practices that result in questionable and inconsistent network data.

22. CEQ Comment

Provide a citation for the statement on page 47 that Survey reviewers found fault with several common tests conducted in the field.

GAO Evaluation

The statement is supported by the examples following it. (See p. 40, vol. I.)

23. CEQ Comment

Provide a basis for the statement on page 49 that the problems of delays during storage and shipment may be chronic.

GAO Evaluation

The statement is our opinion, which is supported by the discussion following it. (See p. 42, vol. I.)

24. CEQ Comment

The comment that sample timing and frequency undermine attempts to compare network data is factually incorrect. The one-way classification linear model can be stated as $\chi_{ij} = \mu + \tau_i + \varepsilon_{ij}$. As the data become less tidy, the size of the term ε_{ij} (random variation) gets larger. Statistical theory has much to say about sampling design, sample size, and the appropriateness of statistical tests which permit this fact to be taken into account to permit the analysis of data.

GAO Evaluation

We have clarified the sentence that CEQ did not like. Our point remains the same. The networks are not well-suited for analyses of trends in water quality because of the lack of homogeneity in river conditions and the small number of samples taken.

We believe that CEQ's method of data analysis cannot be correctly applied to annual averages from the networks. In simple terms, the equation used by CEQ says that a single measurement of water quality (χ_{ij}) is the sum of three independent factors:

--μ, the long-term constant factor;

--τ_i, the short-term (e.g. annual, seasonal, or monthly) factor; and

--ε_{ij}, the random-error factor.

The equation explicitly assumes that the three factors are additive and that there are no interactions among them; both assumptions are wrong.

CEQ used this equation to test annual averages of NASQAN data for significant differences in 3 years (water years 1975-77), and published the results in its 1978 annual report. CEQ defined μ as the 3-year average and used three values for τ_i. The first τ_i was the annual average for water year 1975; the second and third values of τ_i were the annual averages for 1976 and 1977, respectively. Everything that CEQ could not explain by this equation (whether random measurement errors or highly predictable changes resulting from river-flow or seasonal patterns) was incorrectly attributed to ε_{ij} (the random-error factor).

The long-term constant factor is meant to account for long-term predictable properties of rivers, such as rivers in the Arctic being colder than rivers in Florida. A 3-year average computed from monthly samples is not a very good estimator of μ, the long-term constant factor. Because network data often have lopsided, abnormal distributions, the 3-year average can be greatly affected by a few extreme values or a few missing values.

The short-term factor is meant to account for short-term, predictable properties of water, such as rivers being warmer in the summer and colder in the winter or muddier during a storm. CEQ's analysis defined $\tau_{\bar{i}}$, the short-term factor, as an average of measurements taken throughout each year; the analysis contained no factors shorter than 1 year. Consequently, CEQ explicitly assumed that there were no important changes in rivers during each of the 3 years. This is not true. Rivers do change dramatically during a year. They change with the seasons and with riverflows. Rivers are not homogenous and uniform during an entire year. When the underlying reality is highly variable, statistical theory requires that many samples must be taken of each distinct kind of variation in water quality. The networks do not have this flexibility.

The short-term factor can be evaluated only through intensive studies. A short-term intensive study can be designed to define water quality during a distinct type of hydrological event--a summer drought, a spring flood, a winter freeze, etc. Without sound definition of the short-term factor, the other two factors cannot be correctly estimated.

In theory, measurement error should be one of the largest components of the random error factor. For example, to estimate the random error factor in a temperature measurement, several technicians must take simultaneous temperature measurements side by side. If their readings agree closely, the random error factor is small; if they widely diverge, the random error factor is large. Network data are not derived from simultaneous measurements; consequently, CEQ has no direct evidence on the random error factor. In CEQ's method of analysis, the random error factor is used as a catch-all for all the variation that cannot be explained by the 3-factor equation, and measurement error ceases to be a dominant component of the random error factor.

CEQ's equation assumes that the long-term, short-term, and random error factors are additive and that there are no interactions among them. Both assumptions are wrong. For example, there is interaction and interdependence between the long-term trend of increased fertilizer use and the amount of nitrogen and phosphorus that washes off farmlands into rivers during Spring rains (a short-term factor). The more fertilizer there is on the fields, the more there is to wash off during rainy weather. There are many other strong relationships between long-term and short-term factors. CEQ's equation makes no provision for these relationships.

The random error factor (measurement error) is interrelated to long-term and short-term fluctuations in the concentrations of most properties of river water. Many measurements are not uniformly accurate; the size of the measurement error may be closely related to the value of the measurement. For example, in measuring extremely dilute cadmium concentrations (e.g. 1 ug/L) laboratories easily can be off by over 100 percent. At higher cadmium concentrations, measurements by laboratories should be within a few percent of the true value. When short-term seasonal factors (such as heavy rains) reduce cadmium concentrations, the relative measurement error will be increased. When other short-term seasonal factors (such as drought) increase cadmium concentrations, the relative measurement error will be decreased. As concentrations change for most properties of river water, so do the measurement errors. Consequently, the measurement error and random error factor are not independent of the other two factors.

25. CEQ Comment

The terms concentration and loading in water quality analysis are extremely important. More streamflow data is needed to accompany concentration data.

GAO Evaluation

We agree.

26. CEQ Comment

The statistical manipulation on page 35 to produce a "true mean" is not valid. A 95-percent confidence interval would be more appropriate.

GAO Evaluation

We have changed the example to disclose the lack of normal distribution in the data, which is a key test for further statistical manipulation of data. (See p. 33, vol. I.)

27. CEQ Comment

What is the relationship to the type of data presentation in the table on page 36 to the proposed periodic intensive survey program?

GAO Evaluation

The table has been eliminated from the report as part of our revision of the discussion of statistical weaknesses of network data.

28. CEQ Comment

Report comments (see pp. 36-37, vol. I) on the fecal coliform variance are misleading.

GAO Evaluation

The passage has been eliminated as part of our revision of the discussion of statistical weaknesses of network data.

29. CEQ Comment

General comments on the 1978 Annual Report:

--Individual observations were not reported, but aggregated into a national analysis.

--Concentrations rather than loadings were used because water quality criteria are expressed as concentrations.

--NASQAN maps display grouped violation frequencies which damp out all but major differences in water quality.

GAO Evaluation

CEQ's comments that the station data were aggregated into a national analysis and that NASQAN maps damp out all but major differences in water quality confirm the loose use of water quality data that we have addressed in this report. We believe we have demonstrated it is not wise to use data from network sampling, which are sparse both in time and geographic coverage, for national assessments. We discuss this in more detail in response to EPA comment 2 and Survey comments 1, 3, 9, 10, 38, and 49. Also, we have made several changes to the report to more clearly demonstrate the weaknesses of the national networks.

30. CEQ Comment

The term statistically significant data is imprecise.

GAO Evaluation

The report has been clarified. (See p. 51, vol. I.)

31. CEQ Comment

The precise number and extent of extra sampling can be calculated from the existing data.

GAO Evaluation

The report has been revised (see p. 51, vol. I) to more clearly explain that network sampling would have to be converted into special studies tailored to each river.

32. CEQ Comment

The statement that more scientifically sound assessment and supporting data can be produced through special water quality studies is unsupported by the information. A more detailed discussion of the extrapolation from a small handful of special studies to a national program is needed.

GAO Evaluation

That special studies provided much more useful data is not contested by EPA or others. In fact, Wisconsin agreed (see EPA comment 25) that such studies are much more valuable in managing a State program.

A more detailed discussion of the features of special studies and the use of studies for national assessments is provided in the report. (See pp. 54-57, vol I.) As is discussed, we believe that the use of such studies, along with other indicators of water quality, can be much more accurate and useful in describing water quality on a national basis than the existing networks.

33. CEQ Comment

The use of other indicators of water quality, such as changes in fish populations and diversity, can be very useful, but these data are very difficult and expensive to obtain. GAO should explore how such information should be gathered and the data based needed to make the data available.

GAO Evaluation

As discussed on page 62, volume I, other indicators of water quality conditions, changes, and trends, such as the return of fish to previously polluted waters, reductions in municipal and industrial discharges, and biological monitoring results, are currently available and have been used by EPA, CEQ, and States for national water quality reporting purposes. In our opinion, what is needed is not a complex system to statistically manipulate such data, but an innovative, comprehensive program to use them, along with the results of special studies, to provide a more accurate and useful picture of water quality on a national basis.

34. CEQ Comment

We disagree with Professor Velz's recommendations and opinions on page 68.

GAO Evaluation

We agree with Professor Velz's recommendation and opinions.

35. CEQ Comment

The cost data for special studies are difficult to believe. A detailed fiscal and budget analysis should be prepared.

GAO Evaluation

This comment is the same as CEQ comment 7, and our evaluation remains the same.

36. CEQ Comment

The question of available manpower for the proposed program of special studies should be discussed.

GAO Evaluation

The question of manpower is probably synonymous with the question of cost in the comment above. CEQ presents no evidence that such manpower, within fiscal constraints, would not be available.

37. CEQ Comment

The use of reductions in discharges of pollutants as a water quality indicator deserves further discussion as to how it could be implemented.

GAO Evaluation

This comment is similar to CEQ comment 32. Pages 62-63, volume I, discuss this matter.

RESPONSE OF JEROME HOROWITZ, CONSULTANT TO GAO, TO AGENCY COMMENTS ON THE JAMES RIVER CASE STUDY

The Survey and CEQ commented in detail on the James River, Virginia, case study. (See app. VII.) The Survey's comments on the case study are on pages 54 to 65 of appendix IX. The Survey's comments on the case study and Mr. Horowitz' response are below.

CEQ's comments on the case study are a letter from Mr. John F. Ficke (see pages 132 to 139, appendix X). He was the Geological Survey's NASQAN coordinator and was detailed to CEQ to assist in preparing the CEQ 1978 Annual Report. Mr. Horowitz' response to Mr. Ficke's comments begin on page 186.

We agree with Mr. Horowitz' analyses of the detailed comments from the Survey and CEQ on the case study. We do not agree with or endorse the tone of the responses.

GEOLOGICAL SURVEY COMMENTS

1. Survey Comment

The Survey quoted from CEQ's ninth annual report:

"Many of the pollution control facilities built during the past decade are just beginning to operate. Evaluation of their effectiveness requires good uniform data on plant performance and water quality. Fortunately, improved data networks are now providing the means for judging water quality changes, and they will continue to improve in the future.

"So far, uniform water quality data exist for only 3 years, so it is premature to characterize trends definitively. But it is encouraging that bacteria levels improved through the third year.

"Patterns of improvement [with respect to bacteria] are apparent in several populous regions, particularly in the industrial urban belt south of the Great Lakes.

"For other pollutants, no similar patterns of improvement are yet apparent. Levels of suspended material, nutrients, oil and grease, oxygen-demanding substances, and other materials should decline as pollution control becomes more effective. Nonpoint sources are largely responsible for some of these substances."

Consultant's Evaluation

The passage from CEQ's report is exceptionally misleading. To evaluate the effects of a treatment plant on water quality, you have to assess water quality downstream from the treatment plant. Upstream stations are useless. NASQAN stations are usually far from treatment plants. For example, Cartersville is nowhere near any treatment plant in the James River Basin. Most of the materials discharged from treatment plants are subject to rapid change. Bacteria rapidly die in streams. Suspended material settles in quiet waters. Ammonia is oxidized in swift, shallow water. In short, there is a zone of influence below treatment plant discharges. Any station outside this zone of influence cannot account for changes in water quality attributable to improvements in pollution control. Most NASQAN stations are intentionally located outside these zones of influence, and most stations are upstream of cities and their treatment plants. Consequently, NASQAN is inherently incapable of accounting for changes in water quality attributable to improved wastewater treatment.

"Improved data networks" introduce inconsistencies into the data record. Each time the networks are improved, the data are distorted by artifacts of measurement. These distortions vitiate the data for assessing real improvements.

There have been many changes in NASQAN. The Survey is on the lookout for improved methods and for deficiencies in the data. In recent years, the Survey has changed procedures for assessing bacteria and chlorophyll. As problems are identified at particular NASQAN sites (e.g. specific conductance at Cartersville), corrective actions are taken, but the questionable data remain in the record. Consequently, the data are not uniform. Data derived from new procedures are mixed with data from older procedures; questionable data are mixed with data the Survey does not question.

CEQ's patterns of improvement with respect to bacteria should be viewed with great caution. Bacteria are especially subject to rapid environmental change, but NASQAN is inherently incapable of accounting for any rapidly changing property of water. All bacterial data from the networks have been characterized by a Survey official as "of questionable accuracy * * * misleading and erroneous."

There is absolutely no reason to expect that levels of suspended material (as measured at NASQAN stations) "should decline as pollution control becomes more effective." Suspended material is another rapidly changing property of water. Most material remains suspended only so long as the water is swift and turbulent; when the water slows down, suspended materials will settle, no matter how much suspended material was put into the water. NASQAN stations are intentionally located away from waste discharges, usually upstream. NASQAN is therefore inherently incapable of accounting for improvements in suspended material due to pollution control. These same arguments apply with equal force to all properties of water commonly associated with suspended material, e.g. nutrients, oxygen-demanding substances, heavy metals, and pesticides.

Nonpoint sources are especially subject to the influence of flood and drought, and these sources are often associated with many of the rapidly changing properties of water. NASQAN is inherently incapable of accounting meaningfully for any of these properties.

For all these reasons, CEQ's reports (which are derived from NASQAN data) are distorted and misleading.

2. <u>Survey Comment</u>

"It is important to note two points concerning the analysis of NASQAN data in the CEQ report:

(1) CEQ cautioned the reader that with only 3 years of data, it was premature to characterize trends definitively.

(2) No specific station is identified or otherwise singled out. Data from the James River were included with data from all other NASQAN stations to produce the table."

<u>Consultant's Evaluation</u>

CEQ's caution to the reader is far from adequate. The reader has more to be cautious about than the 3 years of data in NASQAN, as this report shows in detail.

CEQ did not single out the James River; we did. The table and maps in CEQ's report were compiled from NASQAN data and intermediate analyses of them. The GAO report explains exactly what was done and why:

"We selected the NASQAN station on the James River at Cartersville, Virginia, for detailed review because (1)

CEQ concluded that there were many changes in water quality at that site, (2) the changes involved both improvements and deterioration in water quality, (3) it was exceptionally data-rich compared to normal national network stations, and (4) the site served three Federal networks: NWQSS (EPA), NASQAN (USGS), and the State of Virginia's EPA-funded monitoring network. Although CEQ did not analyze the State's data, this additional store of data made Cartersville an exceptionally attractive candidate for special analysis.

"* * * [We] reviewed the Cartersville data, paying special attention to the Survey's data and to the conclusions CEQ had drawn from them. Our purpose was to cut through the standard statistical analyses to the data themselves, and through careful examination to assess the credibility of the reported changes in water quality."

The President, the Congress, and the public (for whom CEQ's reports are prepared) cannot break through the quick summaries in CEQ's reports to the intermediate station-by-station analyses and the underlying reality. The case study of Cartersville shows some of the dangers in using NASQAN data for assessing changes in water quality. The closer you look, the less the data explain.

3. Survey Comment

"TOTAL DISSOLVED SOLIDS: CEQ'S TREND AND THE DROUGHT
Appendix VI confirms that the CEQ analyses of total dissolved solids (TDS) was correct; TDS increased during the 1975-1977 water years. Referring to CEQ's text, CEQ did not exclude weather conditions as causing change but reported conditions as they occurred."

Consultant's Evaluation

I agree that CEQ did not exclude weather conditions; CEQ never mentioned the weather. CEQ mentioned only new pollution control facilities and nonpoint sources as possible explanations for changes in water quality. Changes in total dissolved solids have nothing to do with nonpoint sources or new pollution control facilities; these changes can be explained only by the drought, which CEQ did not mention. It is important to show what did cause the change so that the change is correctly attributed to the weather, not to irrelevant point and nonpoint sources of pollution.

4. Survey Comment

"TOTAL DISSOLVED SOLIDS (TDS) AND SPECIFIC CONDUCTANCE: UNEXPLAINED INCONSISTENCIES Specific conductance and dissolved solids concentration are closely related properties of water; however, the ratio of dissolved solids to specific conductance varies with the chemical composition of the water. The ratio has been reported as being generally between 0.54 and 0.96, with ratios above 0.75 usually coinciding with water high in sulfate or containing non-ionic materials such as organic compounds.

"The presence of organic compounds and their effect on the TDS/specific conductance ratio was apparently not addressed by GAO. It is not uncommon for the total organic carbon content in the James River at Cartersville to exceed 5 mg/L as C. Organic compounds often do not materially change the specific conductance of a sample, but they do, of course, contribute to the TDS. It is, therefore, possible to have a TDS/specific conductance ratio that is greater than one if the TDS are low and organic content is quite high. For the 1977 water year, the results from four samples contained both measured TDS and calculated TDS values. The range of calculated dissolved solids values were [sic] 5-12 % lower than the measured TDS values. This strongly indicates that constituents other than inorganics salts are present and must be considered.

"The report asks the question 'But what assurance is there that reasonable-looking data aren't inaccurate too?' Data for the NASQAN program include three related values for each sample: field measurement of specific conductance; dissolved solids, residue on evaporation at 180° C; and dissolved solids sum of constituents. (The last value is calculated as the sum of the concentrations of the determined chemical constituents.) Close agreement between these independent measurements should provide assurance of the accuracy of the data within the stated limits of precision. As a routine procedure, the laboratory compares the value for dissolved solids at 180° C with the dissolved solids sum of constituents. If the values do not agree within the range of precision for both methods, the laboratory will rerun the sample for dissolved solids at 180° C. Also, beginning about 1976, the laboratory computer program through which the district offices receive their completed laboratory analyses has printed a warning on the analyses if the dissolved solids/field specific conductance ratio is not within acceptable limits. Clearly, there is considerable assurance that reported dissolved solids values are reasonable and accurate.

"CEQ based their table and discussion on dissolved solids, not specific conductance. The facts that (1) most of the measurements on duplicate samples for dissolved solids are in agreement within stated laboratory precision, and (2) dissolved solids laboratory measurements are confirmed by calculated dissolved solids values in all NASQAN analyses, provide assurance that CEQ did work with reliable data. Dissolved solids data are not suspect."

Consultant's Evaluation

The Survey makes several arguments in defense of its data. All of them are contradicted by the data at Cartersville. The Survey should pay more attention to site-specific data and rely less on generalized notions.

I do not contest that the ratio of TDS to conductance is generally 54 to 96 percent in a broad range of waters. But page 92 in the case study (which gives examples of TDS and conductance data where duplicate samples gave consistent results across a broad range of seasons and riverflows) shows that the ratio of TDS to conductance was generally about 70 percent in the most credible data at Cartersville. Site-specific data give a much narrower range than generalized notions derived from a broad variety of waters. The water at Cartersville does not contain high concentrations of sulfate (the James River is not the lower Colorado River) or does it contain organic carbon (the James River is not a blackwater river or a swamp).

The Survey's arguments muddy the issues with generalities that do not apply to actual conditions at Cartersville.

I did not deal with organic compounds and their effect on the ratio of TDS to conductance for two reasons:

--Total organic carbon (TOC) makes little difference to to this ratio. Please compare the last two columns in table 1 on page 164. Notice that credible ratios remain credible (e.g. December 15, 1975) and questionable ratios remain questionable (e.g. July 1, 1975 and all the data for 1977), even after adjusting TDS for TOC.

--TOC is not the correct measurement for this correction. We are interested only in dissolved compounds containing organic carbon; but the Survey does not measure dissolved organic carbon at Cartersville. TOC is inappropriate because the total includes both suspended and dissolved carbon compounds.

In short, TOC is not a relevant measurement, and the Survey has no better measurement at Cartersville. But even when we assume that all the TOC was dissolved organic carbon and correct TDS for it, the correction makes little difference to the ratio of TDS to conductance. Once again, the Survey's generalities are contradicted by real data from Cartersville.

Finally, the Survey argues that there is close agreement between TDS and the sum of dissolved constituents (calcium, sodium, chloride, etc., which are independently measured). The data in table 1 show nothing of the kind (please compare the third and fourth columns in the table). Quite often, the sum of dissolved constituents was not calculated (ND in the table), so no comparison is possible. When the sum was actually calculated, it often disagreed with the TDS value. For example, on October 15, 1974, the TDS was 62 but the sum of constituents was 115 -- a difference of 85 percent; and on April 1, 1975, the TDS was 126 but the sum of constituents was 64 -- a difference of nearly 50 percent. Once again, the Survey's generalities are contradicted by actual data. It is not true that "dissolved solids laboratory measurements are confirmed by calculated dissolved solids values in all NASQAN analyses". Table 1 shows that the calculation was often "not done."

1. Total Dissolved Solids and Related Measurements at Cartersville.

Source: WATSTORE Printout (February 28, 1979), prepared for GAO by the Survey.

Date	Time	TDS	Sum	Cond.	TOC	TDS/Cond.	(TDS-TOC)/ Cond.
1974							
Oct 15	1200	62	115	100	ND	62	---
Nov 15	2000	110	96	110	ND	100	---
Dec 23	0930	46	4*	95	2.2	48	46
1975							
Apr 1	1100	126	64	85	ND	148	---
Jul 1	0830	189	ND	105	7.4	180	173
Jul 1	0930	ND	ND	105	ND	---	---
Jul 1	2100	80	86	150	ND	53	---
Aug 11	0900	85	ND	240	12	35	30
Oct 6	0900	46	ND	104	2.6	44	42
Oct 20	0830	152	ND	155	18	98	86
Dec 15	0900	76	ND	125	3.0	61	58
Dec 15	0930	76	68	125	3.8	61	58
Dec 15	1200	98	87	125	ND	78	---
1976							
Jan 12	0830	58	ND	43	2.6	135	129
Mar 8	0830	54	ND	103	4.6	52	48
Mar 8	0900	72	62	103	15	70	55
Mar 22	0815	42	ND	110	2.0	38	36
Sep 20	0830	90	ND	190	5.0	47	45
Sep 20	0900	113	106	190	ND	59	---
1977							
Jan 25	0930	179	ND	140	3.0	128	126
Feb 7	0830	176	ND	125	2.2	141	139

*This value is surely erroneous; the dissolved calcium alone was 7.4 mg/L on this day.

Legend:

TDS is total dissolved solids in mg/L.
Sum is the sum of dissolved constituents (WATSTORE parameter 70301) in mg/L.
Cond. is specific conductance at 25 degrees Celsius, in micromhos.
TOC is total organic carbon in mg/L; note that this measurement is not the same as dissolved organic carbon--the total should be larger than the dissolved fraction of the organic carbon.

TDS/Cond. is the ratio of total dissolved solids to specific conductance, expressed as a percent.
(TDS-TOC)/Cond. is the ratio of total dissolved solids minus total organic carbon to specific conductance, expressed as a percent.
ND means "not done".
Dashes (---) indicate that the ratio cannot be calculated because one of the values needed for the calculation is missing from the data.

For all these reasons, there is no firm assurance that "CEQ did work with reliable data," and the Cartersville data that CEQ worked with provided no assurance that "dissolved solids data are not suspect."

5. Survey Comment

"Under current testing procedures, field personnel using field conductance meters are expected to produce values within [plus or minus] 5 percent of the true value. The Survey was aware in 1975, and part of 1976, that the Virginia District was not consistently meeting this standard for specific conductance measurements. The District was directed to take corrective action and since 1976 the problem has been resolved."

Consultant's Evaluation

If the problem really was resolved in 1976, it is impossible to explain the data for 1977 in table 1, which show that the ratio of TDS to conductance was still much too high. Although the Survey admits that it has reason to suspect the conductance data for 1975 and 1976, these data have not been removed from the record and are still available to CEQ and others for performing trend analyses. Insofar as the Survey has really improved that data, the records are now inconsistent, which violates one of NASQAN's stated objectives and renders the data useless for most analytical purposes.

6. Survey Comment

"The use of the term 'high zinc values' is misleading: the values are not high compared to nationwide occurrences or to water quality standards. The measured concentration (90 ug/L was the highest) [sic] are so much less than the EPA public water supply criterion (5000 ug/L) as to be considered almost negligible. This is not to say that changes in values of zinc below water quality standards should be ignored. They may well indicate water-quality problems with the basin."

Consultant's Evaluation

I agree that the zinc values never came near violating any water quality standard; those were precisely the words we used in the case study. The reader was alerted to this fact to avoid the misinterpretation the Survey has now made.

However, water quality standards are more complicated than the Survey seems to believe. EPA's "Red Book" and the water quality standards for the James River limit zinc to 0.01 of the 96-hour median lethal concentration as determined through bioassay using a sensitive resident species. This restriction was adopted to protect aquatic life, such as fish. Although the State of Virginia has not yet developed a specific zinc criterion by performing bioassays on sensitive resident species, the "Red Book" reports that rainbow trout fry experienced 54 percent mortality in 28 days when the zinc concentration was only 10 ug/L. The "Red Book" also reports that the 96-hour median lethal concentration for rainbow trout and cutthroat trout is 90 ug/L; 0.01 of this 96-hour concentration is less than 1 ug/L. Nearly all the total-zinc measurements (and about half the dissolved-zinc measurements) at Cartersville were much greater than 1 ug/L. Consequently, even the low zinc measurements at Cartersville might cause problems for aquatic life, and the State of Virginia may find (when it does the bioassays specified in the "Red Book") that total zinc should be limited to unattainably low concentrations.

Changes in zinc values at NASQAN stations are uninterpretable because total zinc is yet another of the rapidly changing properties of water. Most of the total zinc is in the form of suspended (not dissolved) material, which responds readily to changes in riverflow. When the river velocity picks up, sedimented materials are resuspended; and when the river slows down, suspended materials are sedimented. Because NASQAN's rigid once-a-month sampling scheme makes no attempt to take samples during defined riverflow patterns, the total-zinc records are uninterpretable. They certainly cannot be interpreted as signifying water quality problems within the river basin. Elsewhere in its comments on the draft report, the Survey admitted that

> "* * * the available data (from NASQAN, NWQSS, and state monitoring activities) are not sufficient to explain the causes or identify the contributions to a particular water-quality measurement. The objectives of NASQAN are focused on description of water quality and not explanation of causes or identification of sources."

7. Survey Comment

"Appendix VI contains the following paragraph which includes several inaccuracies:

'What CEQ took to be an improving trend in water quality might be nothing more than a change in sampling riverflow patterns. Most of the exceptionally high concentrations and loads coincided with dramatic changes in riverflow. These dramatic changes were especially common between December 1974 and June 1976. As luck would have it, the Survey rarely analyzed for total zinc during dramatic changes in riverflow after about September 1975. * * *'

The sample collected June 1, 1976, was at a peak discharge of 16,100 cfs. The mean discharge for the previous day had been 9,060 cfs which certainly qualified the time of sampling as a time of dramatic change in streamflow. Though not as dramatic a change as from 2,500 cfs to 70,000 cfs [during October 1976, when the Survey failed to take a sample for total zinc analysis], the June 1 streamflow was the highest discharge for the period during which zinc values were collected. It is not a 'trick of chance' that certain events are missed but a matter of keeping a regular schedule. Over a period of time, many different hydrologic events will be sampled by following a regular schedule. Unless monitoring is continuous, which is impossible for most of the constituents of interest, some events will always be missed. Implicit in the decision to adopt a fixed monitoring schedule is the fact that some interesting events will not be included in the record. These are restrictions imposed by finite resources and the state of technology, as well as the objectives of the monitoring network which includes trend analysis."

Consultant's Evaluation:

The Survey did not pay attention to the words "most" and "rarely" in the paragraph it quoted. I repeat that "Most of the exceptionally high concentrations * * * coincided with dramatic changes in riverflow" and "the Survey rarely analyzed for total zinc during dramatic changes in riverflow after about September 1975." I did not say that all the exceptionally high concentrations coincided with dramatic changes in riverflow, nor did I say that the Survey never analyzed for total zinc during dramatic changes in riverflows. The point is that the Survey missed most of the opportunities for sampling total zinc during dramatic changes in riverflow after September 1975; it is during these dramatic changes that total zinc is likely to be highest, because total zinc responds sensitively to sudden changes in riverflow. By missing these opportunities, I concluded that "What CEQ took to be an improving trend in water quality might be nothing more than a change in sampling riverflow patterns." I stand by the conclusion.

Nothing in statistical theory compels the Survey to use rigid once-a-month sampling for trend analysis. Pages 29-39 of volume I show that the Survey's sampling scheme produces a heterogeneous hodgepodge of data that cannot be legitimately analyzed. There are many kinds of sampling schemes -- stratified, clustered, nested, etc. -- that would allow the Survey to include "some interesting events" in the data record. Fixed-interval sampling is only one of many types that are widely used in scientific research; it is particularly inappropriate and inefficient for assembling useful data on water quality.

Rigid once-a-month sampling is much easier to administer than flexible sampling schemes. Its bureaucratic attractiveness, however, should not be allowed to obscure its shortcomings or to suppress the fact that other sampling plans can produce data that are meaningful, interpretable, relevant, and statistically sound.

8. Survey Comment

"RECORDS ON DISSOLVED ZINC RESIST RATIONAL ANALYSIS [The Survey] put forth a hypothesis that the data appear to support: that a non-point source of zinc existed in the basin between June 1974 and June 1975. The Survey maintains that it is plausible that there was some kind of material which contained zinc deposited in or near the stream system and that it took approximately a year for the soluble zinc to flush from the system. The available data do not 'prove' this hypothesis to be true, nor do they identify a specific source (they weren't intended to do so). Rather, they indicate that the elevated zinc values were episodic and provide information to develop a reasonable hypothesis concerning their origin. Data should not be dismissed as 'resisting rational analysis' simply because the source producing the material in the stream is not known. * * * If zinc is entering the James River from a point source, or a localized non-point source, then it may be a problem in that stream reach. Once it has been established from monitoring data that zinc values are above the normal

> levels, * * * short-term special studies can then be performed to locate the zinc source."

Consultant's Evaluation

The Survey asks us to believe a fairytale hypothesis about a suddenly vanishing nonpoint source of dissolved zinc. The Survey offers no evidence to support the hypothesis. There is evidence, however, that it must have been a very odd sort of nonpoint source, since even the Survey admits "it took approximately a year for the soluble zinc to flush from the system." Here are the concentrations and loads of dissolved zinc at Cartersville between June 1974 and June 1975:

Date	Concentration (ug/L)	Load (pounds per day)
June 25, 1974	30	580
Sept. 23, 1974	10	55
Dec. 23, 1974	70	2,260
March 10, 1975	40	1,240
June 2, 1975	30	2,550

What kind of flush could have produced increasing loads of dissolved zinc? The highest load came at the end of the supposed flush (June 2, 1975). For the entire year between June 1975 and June 1976, concentrations and loads of dissolved zinc were zero; concentrations and loads didn't taper off, as the Survey's "fairytale" hypothesis would lead us to expect. On the contrary, the concentration was highest in the middle of this magic year and the peak load came at the very end. The Survey's explanation defies rationality.

It is worth repeating from the case study a few lines about this curious hypothesis:

> "USGS hypothesizes that a suddenly vanishing nonpoint source rationally explains the dissolved zinc data. Notice that the flowing loads of dissolved zinc exceeded 2,500 pounds a day at high flow, but were about 60 pounds a day during low flow. What happened to the nonpoint source of 2,500 pounds a day? This unspecified source, according to the Survey, was active only from June 1974 to June 1975. There was no sign of dissolved zinc at 6,050 cfs on March 12, 1974. It has since disappeared without a trace. * * * Since it was a nonpoint source, it could not have been a factory or a sewage plant. It must have been an area permeated with dissolved zinc. The Survey did not think that its zinc data could have been

> wrong. Instead we are asked to believe that an area above Cartersville had been permeated with dissolved zinc; we must further believe that this area suddenly vanished. We again conclude that the data resist rational analysis."

The data clearly show that the dissolved zinc could not have come from a point source; the Survey has no basis for the notion that "the zinc is entering the James River from a point source." If there really was a problem with dissolved zinc somewhere in the James River above Cartersville, NASQAN can tell us little about it because Cartersville is the only NASQAN station on the James. Only short-term special studies can determine whether there was a problem with dissolved zinc in the upper James. Although the Survey did not conduct special studies of dissolved zinc, it maintains "that there was some kind of material which contained zinc deposited in or near the stream system." This is a fairytale, not a scientific argument from measurements and facts.

9. <u>Survey Comment</u>

> "Temperature variations which occur on a daily cycle are found in every stream. The magnitude of the daily temperature variation is dependent on a number of factors. * * * [The] Survey technician consistently visited the James River near the start of the work day. Temperatures will normally be lower at that time of day. The state employee was usually at the site in early to mid-afternoon, when the stream temperature would be approaching a maximum. [GAO] notes that 'round the clock' readings of temperature must be made to define the temperature regime. Such readings are available for the Potomac River at Great Falls, Maryland [but not anywhere on the James] * * *. The temperature measurements for the James River are within the normal range expected of diel temperature variations, keeping in mind that the James River site is about 100 miles south of the Potomac site and that the Potomac drains areas to the north in Maryland and Pennsylvania. * * * The temperatures which are called 'unstable' in [GAO's report] are simply normal diurnal temperature changes."

<u>Consultant's Evaluation</u>:

Fairytales again. Although temperature varies in every stream, the Survey has no way of knowing whether the temperature variations at Cartersville are normal or abnormal, and it has no evidence to explain the difference between its own temperature readings and those taken by the State. The

Survey has no evidence because it has never installed a continuous recording thermometer at Cartersville. Consequently, the Survey has no way of knowing about daily temperature fluctuations at Cartersville, their relative stability, or their dependence on time.

The Survey introduces the Potomac River at Great Falls (which is unlike the James at Cartersville with respect to drainage area, geography, exposure to the air, and the situation of nearby tributaries) in an attempt to show that the discrepancies among the readings taken by Survey technicians, State employees, and the Survey's daily observers are "simply normal diurnal temperature changes." Perhaps they are, and perhaps the Survey uses the Potomac to show something about the "spatial transferability of information, which is discussed elsewhere in these replies. But there is no real evidence to support this theory at Cartersville. And why? Because NASQAN takes the temperature of the river only once a month, and the Survey hasn't installed a recording thermometer. Why should anyone have to guess about temperature? It's not very hard to measure.

Several major tributaries enter the James just above Cartersville. The temperature at Cartersville is affected by temperature changes in each of these tributaries and the James itself. The Cartersville temperature responds to the flows and temperatures in each of the rivers upstream. Consequently, we must continue to insist on the case study:

> "To make sense of the apparent temperature instability [at Cartersville], temperatures must be read around the clock in the James River itself and in the major tributaries just above Cartersville. This work hasn't been done and there are no plans for doing it. It will take more than a few readings a month to define the temperature regime at Cartersville."

One problem with NASQAN's temperature measurements is time bias. It may be instructive to compare the networks with the weather measurements reported every day in the works with the weather measurements reported every day in the news. The water quality networks report on one location once a month, in general. This procedure is comparable to measuring temperature in Washington, D.C., on the Capitol steps on the first Tuesday of the month between 2:00 and 2:15 p.m. Although one measurement a month might establish -- but only after several years of data collection and analysis -- that it is usually hotter in July than in January, it could not give a fair picture of temperature in the D.C. area. The measurements would be biased because they would always miss the cooler temperatures at night and in the suburbs. It does not matter in this argument how sophisticated the thermometer is; it does not matter how many degrees the technicians have, or how much experience they have in reading thermometers, or which step they are standing on (East front or West front, top step or bottom step). The results are biased by sampling only one location only once a month, always in the afternoon. In a month like March or October, when temperatures in the D.C. area could easily range from subfreezing to the nineties, a single afternoon measurement would certainly give a distorted report of reality. The time bias is inevitable; it is built into the system and cannot be overcome by better instruments or more skillful technicians.

Appealing to temperature measurements in the Potomac is equivalent to verifying once-a-month temperature measurements on the Capitol steps by appealing to hourly temperature readings in the Philadelphia suburbs; Philadelphia is about as far north of Washington, D.C. as Great Falls is from Cartersville. What kind of measurement program would that be? What kind of science is it?

10. Survey Comment

"Throughout the discussion of dissolved oxygen data from the James River site, it is important to keep in mind that none of the dissolved oxygen values collected by the Survey gave any indication of a problem in meeting the minimum value of 4.0 mg/L or the daily average of 5.0 mg/L dissolved oxygen set by the State standards. The lowest measured value (1974-1977) was 6.1 mg/L on July 8, 1974, at 8:15 a.m."

Consultant's Evaluation

Time bias again. Elsewhere in its comments the Survey insists on the importance of biological reaeration, which depends on sunlight and temperature. Green plants and algae produce oxygen when the sun shines on them, but they breathe oxygen at night, removing it from the water. The Survey has never measured dissolved oxygen (DO) at Cartersville between midnight and dawn, when DO may be lowest because green plants could have removed large quantities of DO during the night. Although the lowest DO reading at Cartersville was 6.1 mg/L, that reading was taken in July at 8:15 a.m., when the sun had been up for hours. What might the DO have been at 3:00 a.m., before the sun rose? The Survey has no way of knowing because it has no data from predawn hours. The Survey has never measured DO around the clock at Cartersville, even though it now insists that biological reaeration greatly affects the Cartersville DO. The discrepancy between the Survey's insistence on the importance of biological reaeration and the meager data on this phenomenon at Cartersville is impossible to understand. If biological reaeration is nearly as important as the Survey now insists, predawn DO could regularly violate the State's water quality standards, but the Survey has no facts to decide one way or the other.

There are continuous DO recorders, just as there are continuous recording thermometers, and the Survey often uses them. Curiously, the Survey has never installed a continuous DO recorder at Cartersville, despite its insistence on hour-to-hour DO variations attributable to biological reaeration. Once again, there is no relation between the Survey's arguments and its data collection program.

DO is a rapidly changing property of water. The DO at Cartersville could be very high at 8:15 a.m., but very low a few hours earlier or a few miles distant. I have pointed out elsewhere that NASQAN cannot account for any rapidly changing property of water, and DO is a prime example.

11. Survey Comment

"Much of the discrepancy in table 10 [p. 104 of Volume I] is explained by time of day and water temperature. The two State samples were collected in the early afternoon which is, as Appendix VI points out, the period for maximum biological aeration. The Survey sample was collected in the morning when biological aeration had not reached a maximum. Water temperature also has an effect on the production of oxygen by green plants. The rate of oxygen production is less at 10.5°C, the temperature at the time the Survey made the measurement, than at 16.7°C or 17.8°C when the State made its measurements."

Consultant's Evaluation

The Survey openly admits the time bias in its measurements but neglects location bias. It is true that green plants may have their greatest effect on DO maxima in the afternoon, but those green plants need not be at Cartersville. The James is swift at Cartersville. Even if the river were full of green plants and algae at the Cartersville bridge, the water might not be in contact with them long enough to change the DO. Most of the algae and plants that could change the DO at Cartersville are not at Cartersville; they are somewhere upstream in the James River itself or in one of its tributaries. The Survey has no information on algae and plants upstream of Cartersville.

Although some kinds of plants and algae move with the water, most aquatic plants are rooted -- fixed to one spot -- and many aquatic algae are fixed to the bottom or sides of the river channel. The Survey has little information on attached plants and algae of any kind at Cartersville, and no information on them upstream. Without this information on upstream sites, the Survey is in the ridiculous position of claiming that the algae and plants at Cartersville explain the discrepancies between NASQAN data and the State's data on DO. This is location bias with a vengeance.

Since the Survey is fond of a fairytale hypotheses, I will concoct another one. Suppose there were lush growths of attached plants and algae 10 miles upstream of Cartersville. These growths would probably have their greatest effect on DO maxima during the afternoon. It would take time for the water to travel to Cartersville. If the river was in flood, the water would quickly travel to Cartersville; but if the river was in drought, the water would travel slowly. The effect of biological reaeration at this upstream site could not be felt at Cartersville until the water got there. The water certainly could not traverse the hypothetical 10 miles in an instant. Location bias and time bias mesh to grind up the Survey's explanations. Depending on the river velocity, the water that was oxygenated by the lush growths 10 miles upstream might reach Cartersville a few hours later (in the early evening) or much later during the night. Fairytale hypotheses get us nowhere. What is needed is data free from time bias and location bias, but NASQAN cannot produce unbiased data.

The plain fact of the matter is that DO data from NASQAN do not agree with DO data from the State, and DO trends from NASQAN data do not agree with DO trends from the State's data. The discrepancy between the State's data and NASQAN data in table 10 (p. 104 of vol. I), led me to the following conclusion:

> "What can account for the river's having lost nearly 40 percent of its oxygen supply? * * * To interpret these discrepancies in the DO record we would need a battery of measurements for several days on chlorophyll, BOD, and ammonia along the river and its tributaries above Cartersville. To rule out biological reaeration we would need DO readings several times a day, including midafternoon and just before dawn. It is impossible to interpret these discrepancies from the measurements on record."

The Survey's arguments reinforce the logic of this conclusion.

12. Survey Comment

"The discussion following Table 11 [p. 105, vol. 1] includes a statement from the November 21, 1979, letter from the [Survey's] Chief, Quality of Water Branch which says:

> 'Weather information and sediment concentrations are available for assisting in interpretation of the DO data. Conditions for the two November samples were clear weather, low sediment concentrations, and low flow. The temperature of the sample taken on November 24 was higher and temperature-dependent photosynthesis activity could account for the higher DO value. On December 9, the flow was high from a period of storms, the weather was cloudy, and sediment concentration had increased to 131 mg/L. Lower temperature (compared to November), higher water stage with increased sediment concentration covering periphytic [i.e. attached] algae and cloudy skies reducing available sunlight would diminish photosynthetic activity and reduce the amount of oxygen added to the water. Also, runoff following storms generally has a higher BOD load than the water entering the stream during a low flow period (usually ground water). This oxygen demand reduces the amount of dissolved oxygen in the stream.'

The Survey maintains that this is a valid interpretation of the data presented in Table 11. The chlorophyll a and b data will be discussed later in this reply. It is appropriate to state that, from data obtained over time at that site, there is good reason to expect to find periphyton and phytoplankton [free-floating and attached algae, respectively] in the stream."

Consultant's Evaluation

The Survey insists that its November 1979 letter presents a valid interpretation of the data. I found serious errors in that letter, and refuted the Survey's arguments in the case study as follows:

"The Survey offers two explanations of the DO anomaly [shown in table 11 of the case study]: (1) algal photosynthesis and (2) high BOD. Neither explanation is satisfactory; both are contradicted by the Survey's own data.

"Photosynthesis by attached algae (periphyton) is speculation: the Survey has no data on periphyton or periphytonic chlorophyll at Cartersville in late 1974 [the period covered by table 11]. The duplicate sample on November 25, 1974, shows that the Survey's analyses for phytoplanktonic chlorophyll did not give consistent results. The first analysis yielded zero for chlorophyll _a_ and _b_; the second analysis yielded 12 and 16 ug/L, respectively. These results demonstrate that the Survey could not reliably measure chlorophyll. The Survey cannot assert that chlorophyll was or wasn't present, or that algal photosynthesis did or didn't affect the DO regime at Cartersville. Unreliable and incomplete data lead to paradox and inconclusiveness, not to firm answers.

"BOD concentrations--not BOD loads--explain DO concentrations. The BOD concentrations steadily dropped between 11 November and 9 December 1974, according to the Survey's own data.

"The data do not support firm conclusions on algal photosynthesis because too many phenomena were neglected (e.g. periphyton) and because the Survey's chlorophyll measurements were unreliable. All the other data (streamflow, temperature, and oxygen demand) indicate that DO should have increased on 9 December--the water was colder, swifter, and contained less BOD. It does no good to invoke algal photosynthesis when the data cannot explain important changes in the river."

The Survey did not take issue with any of the technical arguments or facts in this passage. The Survey stubbornly insists that its interpretation is valid, although I showed that biological reaeration and photosynthesis are fairytales--the Survey does not have the facts to support them.

The Survey now admits that its chlorophyll measurements weren't all they could be:

"GAO was informed that a more sensitive method had been developed for chlorophyll determinations and that the old method had been superseded."

The Survey has very little evidence on algae or plants of any kind (periphyton, phytoplankton, or rooted aquatic weeds) at Cartersville, and no evidence at all for any upstream site. The Survey further admits that its method for analyzing chlorophyll has been replaced by "a more sensitive method," which is an indirect admission that the chlorophyll data I criticized will not stand up to scientific scrutiny. Despite all this, the Survey insists on the validity of its interpretation.

It is worth discussing a little further the fundamental fact that BOD concentrations -- not BOD loads, as the Survey insists -- explain DO concentrations. BOD refers to substances that react with oxygen in the water; as they react, they remove DO from solution. A load of BOD is a weight, e.g. 10 grams. Ten grams of BOD mixed into 1 gallon of water removes just as much DO as 10 grams of BOD mixed into 2 gallons of water. But there is twice as much DO in 2 gallons of water as in 1 gallon. It is the concentration of BOD -- not the load of BOD -- that explains changes in DO concentrations.

I refuted the Survey's BOD arguments in the case study and warned it that it was technically wrong to confuse concentrations and loads. I must now repeat that this confusion is a fundamental technical error. The Survey should know better.

13. Survey Comment:

> "Table 12 data (p. 107, vol. I) do not show inconsistencies. On two dates, May 3 and 17, 1976, both morning measurements [were] made on cloudy days, the percent DO saturation was 84.5 and 76.7 respectively. On May 14, 1976, in the afternoon, the dissolved oxygen was at 103 percent saturation. Given the flow, the type of stream, and the time of year, these values are consistent with what one would expect."

Consultant's Evaluation

Once again the Survey insists that the data are "consistent with what one would expect", although the data are anything but consistent. In this case we argued as follows:

> "Table 12 illustrates several inconsistencies. On 14 May the riverflow was lower and the temperature was higher than on 3 May. Theoretically, DO should have been lower: warmer water has a lower saturation value, and sluggish water has less physical reaeration. In fact, the opposite occurred: DO increased and the percent saturation jumped by nearly 20 percent.
>
> "On 17 May the riverflow was much higher and the temperature was virtually unchanged from 14 May. Theoretically, DO and percent saturation should have increased; instead, both fell. Why should the percent saturation have fallen by over 25 percent? Physical reaeration cannot be the reason. If anything, physical reaeration must have been much stronger on 17 May than on 14 May, since the riverflow had nearly tripled. The State never tests the water at Cartersville for chlorophyll or green plants, so we have no way of assessing biological reaeration."

The Survey's comment refers to cloud cover and time of day, which are relevant only to biological reaeration. Since the Survey does not have adequate data on chlorophyll, algae, and plants, and since the State has no information on any of these properties of water, it is impossible to discuss biological reaeration in a rational manner from the facts. Once again the Survey conjures up the fairytale hypothesis of biological reaeration in an empty attempt to explain serious discrepancies in the data. The Survey's comments do not deal with any of our technical arguments which are derived from the available data, not from fairytales. The data do in fact show that the James River lost over 25 percent of its oxygen resources, despite the great increase in riverflow and the constant water temperature. It will take more than a fairytale to explain this dramatic change in the river. The Survey has no relevant facts to offer (NASQAN cannot produce them) continues to spin out fairytales. Given the inherent deficiencies in NASQAN, the Survey has few hard facts to work with. It's a fairytale or nothing.

14. Survey Comment

> "The discussion of DO presented in Appendix VI entirely ignores the objectives of NASQAN (to describe water quality) and imposes another set of objectives (the complete understanding of the causes of each observed DO value). The fact that the GAO's consultant could not explain the observed DO values by the means of [the] analysis he chose does not constitute a valid criticism of the NASQAN program or data."

Consultant's Evaluation

The case study of the James River at Cartersville seems to have fulfilled its purpose rather well. It focused attention on the uselessness of the Survey's data, their inexplicable discrepancies, and their failure to agree with the State's data at the same site. In its voluminous comments, the Survey has again demonstrated its inability to account for the inconsistencies, discrepancies, and paradoxes in its own data. The Survey should not be too hard on me for being unable to explain the data; nobody could. But I did show that the Survey's Cartersville data are inconsistent, discrepant, and paradoxical -- something the Survey should have done for itself (and done something about) long ago.

Owing to time bias, location bias, the data's inability to account for any of the rapidly changing properties of water (DO is a prime example), and the deficient coverage of essential facts (such as nighttime DO and plant growths upstream from Cartersville), the Survey has had to concoct fairytale hypotheses to explain the glaring contradictions and paradoxes in the data. By showing how little NASQAN data can actually explain, the case study showed how poorly NASQAN describes water quality and the factors that influence it.

15. Survey Comment

> "The discussion following Table 13 [p. 108, vol. I] is somewhat misleading. It is important to keep in mind that the DO concentrations on cloudy mornings when DO should not be at a maximum were 84.5 and 76.7 percent of saturation and in very close agreement. The lowest DO value measured was 6.1 [sic] mg/L, well above the State standard."

Consultant s Evaluation

The lowest DO value in table 12 is 6.9 (not 6.1) mg/L. This value is much higher than the State standard of 4 mg/L. But the 6.9 value of DO was measured at 10:30 a.m., when the sun had been up for hours. The Survey does not know what the DO might have been just before sunrise. If biological reaeration at Cartersville is really important, the predawn DO might have been well below 4 mg/L. But the Survey has never measured DO at Cartersville just before dawn.

Tables 12 and 13 in the case study give data on DO and factors that might influence DO (riverflow, algae, chlorophyll, temperature, time of day) for 2 days in May 1976. Except for temperature (it was warmer on the 17th than on the 3rd of May), all the related factors suggest that DO should have been higher on the 17th; but DO was considerably lower on the 17th. Faced with this paradox in the data, the Survey now argues that biological reaeration could not have been important because of the time of day and the cloud cover. Instead, the Survey argues that temperature controlled the DO these 2 days. Please note that the Survey concocted the hypothesis of biological reaeration to explain DO readings that could not be accounted for by temperature effects. (See the Survey's comments and our replies on tables 10-12 of the case study.) When confronted with data on biological reaeration in table 13 -- data showing that biological phenomena cannot explain the DO readings -- the Survey denies that biological reaeration had any effect.

The Survey has not shown that biological reaeration was unimportant, because the NASQAN data are deficient:

--The amount of DO that green plants add to the water depends on the intensity of solar radiation. The Survey never reports on the intensity of solar radiation (measured in foot-candles). The Survey merely reports on the approximate cloud cover, which has nothing to do with the amount of solar radiation reaching aquatic plants. A clear day in January might very well pass less solar energy than a cloudy day in June. Cloud cover is an inadequate substitute for proper radiation measurements.

--The Survey has admitted that its methods for detecting chlorophyll were inadequate and have since been changed. (See p. 176.)

--The Survey has no data on algae or aquatic plants upstream of Cartersville. The Survey's biological data (limited as they are) are restricted to Cartersville; but the green plants at Cartersville cannot account for the DO effects of green plants growing upstream in the James and its tributaries. The Survey's data are incomplete and distorted by location bias.

--The Survey's chlorophyll measurements were not properly coordinated with the algal measurements. The two sets of measurements should have been conducted on one sample; instead, they were conducted on two separate samples. The Survey has recognized this procedural error and now requires proper coordination between these sets of measurements. But the data in table 13 were accumulated before the Survey recognized the error.

--Oxygen production by plants and algae depends on more than chlorophyll. It depends on the physiological state of the algae -- in a word, on their general health. I pointed out in the case study that there are many measures of algal physiological activity (enzyme activity, dark ammonia uptake, radiocarbon fixation, etc.). But neither the Survey nor the State has taken any of these physiological measurements on the algae at Cartersville.

--The DO produced by green plants is not determined by the water temperature. When algae are abundant and solar radiation is intense, water can easily be supersaturated with oxygen. The saturation value (which is determined by temperature) loses its meaning when biological reaeration is an active force. Consequently, the Survey's contention that the percent saturation values on May 3rd and May 17th were "in very close agreement" has no merit.

In sum, the Survey's data on biological reaeration are <u>inappropriate</u> (cloud cover is no substitute for solar-intensity measurements), <u>inconsistent</u> (chlorophyll analyses were changed), incomplete (the Survey has no information about plants and algae upstream of Cartersville), uncoordinated (chlorophyll measurements and algal enumerations were improperly conducted on separate samples), and inadequate (the Survey neglected all measurements relevant to the physiological robustness of aquatic plants and algae). Owing to these weaknesses in the data, the Survey cannot show whether biological reaeration did or didn't affect DO values at Cartersville, and the Survey's arguments from percent saturation values beg the question by assuming what the data cannot prove.

16. Survey Comment

"The statement, 'the Survey should be encouraged to delete this misleading measurement (phytoplankton, total cells per milliliter) from its list of tests' is puzzling. Contrary to the Appendix VI contention, cell count per milliliter is a standard measurement used by biologists and is widely reported in the hydrologic literature. As has been pointed out to GAO heretofore, algae are identified and counted at the genus level. These data are available from the Survey and have been published in the Virginia data reports since 1978."

Consultant's Evaluation

The case study explained in detail what is wrong with cell count per milliliter:

> "Algae vary enormously in size, mass, cell respiration, photosynthetic activity, and chlorophyll content. Some algae are huge; others are minuscule. Ignoring the difference between a large algal cell and a small one (or a healthy cell and a sick one) is rather like ignoring the difference between an elephant and a flea -- the size discrepancy is misleading, even though both are animals. The Survey should be encouraged to delete this misleading measurement from its list of tests."

The Survey's use of this measurement is in no way extenuated by the fact that it is "widely reported in the hydrologic literature." Careful biologists do not use this measurement because it is meaningless and explains nothing. Convention is no substitute for truth.

It is no news that the Survey identifies algae to the genus level, as pointed out in the case study (p. 109, vol. I):

> "We checked the Survey's data sheets on phytoplanktonic algae. These papers show that there were all types of

algae at Cartersville. Some days the dominant algal types were tiny diatoms; on other days the dominant algae were large filamentous blue-green algae or gelatinous colonies of branched green algae. No matter what the algae were, table 14 shows that the Survey could not reliably detect chlorophyll. Clearly, there is something wrong with the chlorophyll analyses. They cannot be trusted."

I cannot account for the Survey's failure to read this paragraph.

17. Survey Comment

"Several points should be clarified about the chlorophyll determinations:

1. Chlorophyll is not a constituent sampled within the NASQAN program. It was included at the James River station to complement other work.

2. GAO was informed that a more sensitive method had been developed for chlorophyll determinations and that the old method had been superseded.

3. Method B-6501-77, the old method, does call for grinding the filter. The complete method was provided to GAO in TWRI, Book 5, Chapter A4, page 209."

Consultant's Evaluation

It makes no difference whether chlorophyll was included in the Survey's NASQAN program or in some other program. The plain fact of the matter is that the chlorophyll measurements were obviously wrong. These erroneous data are in the Survey's WATSTORE printouts, its data files, and its reports. They are incorporated into Federal records from the fixed-station network.

I am mindful that the Survey has changed detection methods for chlorophyll. I hope that the new methods will produce better data than the old one. It is important to note that the erroneous data from the old detection method are still in the Survey's data files. Consequently the Survey's records for chlorophyll are now inconsistent (results from the new and the old methods are now mixed) and at least in part erroneous. I believe the Survey should delete all the erroneous chlorophyll data from the record.

18. Survey Comment

> "Please note that the Survey cautioned GAO about using cell counts and chlorophyll data collected at different times. The June 2, 1975, data which shows 2,900 algae cells and 55 ug/L chlorophyll *a* were collected at different times. Appendix VI implies that the Eppley and Sloane data are incorrect, yet it does not provide any reference to work which would indicate that it is incorrect. Obviously, there was a difference in number of cells between the two sampling times."

Consultant's Evaluation

It is the Survey, not GAO, that made the error of determining chlorophyll and algal cell counts in separate samples. In the case study I argue that:

> "One sample should be used for cell counts and chlorophyll analysis. We encourage the Survey to stop running these analyses on separate samples, which is inefficient and unscientific."

Although the two samples (one for chlorophyll analysis, the other for algal cell counts) are collected only a few minutes apart, they cannot be compared. The Survey has discovered the error of its ways and has instituted new procedures, in Virginia at least:

> "The Survey agrees that all samples for chemical analysis at a given site should be collected at one time. In the report, 'Technical Review of Virginia District water quality activities, November 9-12, 1976,' the District was told to collect all samples at the same time and to discontinue the practice of collecting related chemical and biologic constituents at different times. GAO was provided a copy of this review. Data for the 1977 water year and the subsequent period have been taken as recommended."

The Survey should now go through all its records and delete the erroneous data. Until these data are removed, the records will be inconsistent and will contain misleading data on water quality.

Now that the Survey has adopted scientifically adequate procedures, it will not have to depend on the gross estimates of Eppley and Sloane, which explained nothing, as I showed in the case study:

"The arguments based on the generalized estimates of Eppley and Sloane are contradicted by the Survey's own data at Cartersville. The Chief of the Survey's Quality of Water Branch argues that 4,100 cells per mL would be expected to produce a maximum chlorophyll value of 20 ug/L. On June 2, 1975, the Survey reported 2,900 algal cells at Cartersville and gave the chlorophyll a concentration as 55 ug/L. The Survey did not report the genera of algae that were represented in the 2,900 cells. However, it is clear that the estimate derived from Eppley and Sloane (15 ug/L at most) is not consistent with the Survey's own data; the actual chlorophyll measurement (55 ug/L) is 367 percent higher than the highest possible estimate from Eppley and Sloane (15 ug/L)." (See p. 111, vol. I.)

The Survey dragged in Eppley and Sloane in a futile attempt to explain the impossible chlorophyll values given in table 14 of the case study. In his letter of November 21, 1979, the Chief explained why he was using the estimates of Eppley and Sloane:

"Most chlorophyll samples in Table 14 were collected at a different time than the cell counts and therefore should not be directly compared. However, cell counts can be used as estimates of the expected chlorophyll values by using the following table [from Eppley and Sloane's 1966 article]."

Now that the Survey is properly coordinating chlorophyll analyses with cell counts, it can stop relying on the paradoxical estimates from a 1966 article and can begin working with scientifically sound facts.

19. Survey Comment

"It should be noted that in the letter quoted above [the Chief's letter to GAO, dated November 21, 1979], the table heading clearly stated '(10^{-12} g/cell)'; the report mis-quotes it to be '(10-12g/cell)[sic]'."

Consultant's Evaluation

The letter speaks for itself. The case study quoted the relevant portion of this letter exactly. I did not misquote the Chief.

20. Survey Comment

"Appendix VI refers to Table 17 and states 'the Survey's values for DO and percent saturation are often much lower than the State's.' It is important to note that

> the Survey was consistently at the James River from early to mid-morning. The State was usually at the site in mid to late afternoon. Dissolved oxygen varies during the day as has been previously pointed out. When the State and the Survey visited the site at the same time of day, the results were much closer. Note the August 26, 1975, visit by the State and the September 8, 1975, visit by the Survey. The flow in the river was constant and both samples were collected near 11 a.m. Dissolved oxygen concentration was 7.2 and 7.7 mg/l [sic] respectively and percent saturation was 92 for both."

Consultant's Evaluation

It is interesting that the Survey now claims that time bias accounts for the differences between the State's readings (generally taken in the afternoon) and the Survey's readings (generally taken in the morning). Now let me press this point to its logical conclusion. In the case study they noted that

> "The progressive DO improvement shown by the Survey's annual DO averages are not confirmed by the State's averages. Had CEQ used DO data from the State rather than from the Survey, it would have reached entirely different conclusions about trends in Cartersville."

Nothing compels the Survey to sample at Cartersville in the morning. The Survey might just as easily have sampled during the afternoon, as the State does. The Survey agrees that if they had sampled during the afternoon, its DO data would agreed with the State's. Consequently, CEQ's DO trend at Cartersville is (according to the Survey) merely an accident of a technician's work schedule.

Although the Survey claims that "dissolved oxygen varies during the day," it provides no proof for this claim because it has never measured DO at Cartersville around the clock. DO may vary from hour to hour if biological reaeration dominates the oxygen regime of the river; but we have repeatedly shown that the Survey cannot show that biological reaeration is important. The claim is pure conjecture, unsupported by fact.

The Survey alleges that the river's oxygen regime on August 26, 1975, was identical to the oxygen regime on September 8, 1975. Perhaps it was, but there are no data to support this allegation. There are no matched measurements of algae, chlorophyll, the intensity of solar radiation, and all the other phenomena controlling biological reaeration. The agreement between the State's DO reading on September 8th and the Survey's reading on August 26th could be entirely accidental. There are no data to support the Survey's argument.

21. Survey Comment

"Analyses of duplicate samples for dissolved solids are generally within the stated limits of precision for the method. No data were included in Appendix VI to show that TDS samples 'often disagreed widely.'"

Consultant's Evaluation

Table 4 of the case study does include five sets of duplicate samples:

Date	Time	TDS (mg/L)
Mar. 10, 1975	08:00	79
Mar. 10, 1975	08:30	49
July 1, 1975	08:30	189
July 1, 1975	21:00	80
Dec. 15, 1975	09:00	76
Dec. 15, 1975	09:30	76
Dec. 15, 1975	noon	98
Mar. 8, 1975	08:30	54
Mar. 8, 1975	09:00	72
Sept. 20, 1976	08:30	90
Sept. 20, 1976	09:00	113

The disagreements speak for themselves.

CEQ COMMENTS

Mr. Ficke's comments on the case study, in a letter dated May 28, 1980 (see pp. 132 to 139 in app. X), are very important because he was the Survey's NASQAN Coordinator and was detailed to CEQ to help prepare CEQ's 1978 Annual Report. The 1978 Annual Report supplied some of the material for the case study of the James River at Catersville, Virginia (see app. VII in vol. I). Mr. Ficke is currently with EPA's Office of Toxic Substances. His comments were prepared at the request of CEQ.

1. Ficke Comment

"Although I have many reactions to material in the GAO draft that is distorted or just plain wrong, I will comment mostly on GAO's discussion of material used for the 1978 Annual Report (AR) of the Council on Environmental Quality."

Consultant's Evaluation

The only part of CEQ's 1978 Annual Report discussed in the draft was one summary table. Something more should be said about the 1978 CEQ report, and this is a good place to say it.

Of the 10 trends published in CEQ's 1978 Annual Report, 2 were wrongly labeled, 1 was calculated from inaccurate data, and 2 were calculated from inconsistent data (the measurements were affected by changes in methods or equipment). In many cases the samples were stale (and therefore invalid) because too much time had elapsed between gathering and analyzing the samples. For these reasons, we question the validity of the conditions and trends identified in CEQ's report.

Incorrect labeling: CEQ incorrectly labeled both inorganic nitrogen and organic nitrogen, and the error is much more serious than a simple switching of labels. Inorganic nitrogen by definition includes ammonia, but CEQ actually analyzed nitrite plus nitrate, which excludes ammonia. Organic nitrogen by definition excludes ammonia, but CEQ actually analyzed Kjeldahl nitrogen, which includes ammonia. In short, all the trends involving nitrogen are distorted and misleading because of CEQ's mishandling of ammonia data. But ammonia is one of the most important water-quality characteristics because it is a potential fish toxin and a powerful deoxygenator of water. Ammonia is one of the major products of sewage treatment plants, a common industrial waste product, and the most widely used fertilizer in the country. The errors in CEQ's use of ammonia data are particularly puzzling because CEQ's analysis was supervised by Mr. Ficke, the NASQAN coordinator of the Geological Survey, and intimately familiar with the data CEQ was analyzing.

Inaccurate data: In its trend analyses, CEQ used inaccurate NASQAN data on phytoplankton, an important category of algae. CEQ used preliminary data, radically different from the final data now on record. For example, the NASQAN station on the Salinas River (near Spreckels, California) now shows 79,000 phytoplankton cells for water year 1975, but the preliminary data CEQ analyzed gave the 1975 count as 27,500 cells -- only a third of the final count. The final data were available when CEQ performed its trend analysis under the supervision of the Survey's NASQAN coordinator.

Inconsistent data: The Survey insists that consistency is one of the essential properties of the NASQAN data base. However, the claim that NASQAN data are consistent is false.

Methods have been changed for measuring important water characteristics. The mixing of data developed from different methods is like including apples in a count of oranges. It is unprofessional and unscientific. Nevertheless, CEQ's report did just this; it must be read with reservation and used with caution because of these grave analytical errors. In NASQAN data for water years 1975-1977, for example, methods and equipment were changed for at least two items included in CEQ's analysis:

--The filter and the growth medium used for fecal streptocci were changed.

--The filter for recovering fecal coliforms from riverwater was changed.

Although the new data are more accurate, they are not consistent with the earlier data and may result in spurious trends. A Survey biologist said that bacterial numbers should increase now that the growth medium has been improved.

In September 1976, the Survey officially changed the growth media for both kinds of bacteria and authorized the use of Millipore HC filters. Published literature shows that the HC filter is more efficient than the old HA filter in recovering fecal streptococci from rivers and streams. For these bacteria, the CEQ report states that 10.9 percent of NASQAN stations showed significantly higher numbers but only 1.8 percent showed lower numbers. In short, five times as many stations got worse as got better. It is not clear how much of this trend is real. All of the trend could be spurious -- an artifact created by improving the detection procedures.

The changes in methods to obtain NASQAN data continue today. In March 1979 the Survey adopted a single method for calibrating meters used to measure dissolved oxygen. Before this time, several different calibration procedures were used. This change may introduce a bias into the data and cause spurious trends if new and old data are mixed during statistical analysis.

Stale samples: Some of the data for assessing nitrogen trends are no good because the samples were stale when they were analyzed (see p. 42, vol. I). In its formal comments, the Survey agrees that the "GAO report has noted a possible problem," which is perhaps the least that can be said. Insofar as the samples are stale, the NASQAN data are invalid and CEQ's assessment of conditions and trends is meaningless.

2. Ficke Comment:

"The GAO draft * * * dwells at length on the analyses and data used in constructing Table 2.1 (page 96) of the 1978 AR and in the brief paragraph of discussion on pages 96 and 98 of the AR. Table 2.1 was constructed using the following procedure:

1. Data from 357 stations of the National Stream Quality Accounting Network (NASQAN) for 10 water-quality characteristics (fecal coliform bacteria, inorganic nitrogen, organic nitrogen, total phosphorus, dissolved oxygen, fecal streptococci, dissolved solids, dissolved zinc, total zinc, and phytoplankton) for the 1975, 1976, and 1977 water years were subjected to an analysis of variance (ANOVA) to determine whether there were significant differences in the annual mean values of the 10 characteristics.

2. The computations for the ANOVA were done by a contractor who computed an 'F' statistic based on the sums of squares within and between the annual data sets for each constituent at each station.

3. I examined each set of data (each characteristic at each station) to determine (a) if the F statistic indicated a significant difference among the means and (b) if the differences represented a clear pattern of improvement or worsening of water quality. In comparing the F values, I used the numbers of degrees of freedom computed by the contractor, and F values to test for significance at the 90 percent level. Each characteristic at each station was labelled either '+' for improved quality, '-' for worsened quality, or '0' for no change.

4. With the help of a contractor employee, I compiled the changes and computed the percentages of the stations that had improved, not changed, or deteriorated, for each characteristic."

Consultant's Evaluation

There are three fundamental failings in this comment:

1. The F ratios were computed without including interaction terms.

2. Because the data are not normally distributed, it is misleading to claim that the F values were "tested for significance at the 90-percent level."

3. Annual averages are misleading summary statistics where water quality data are concerned because averages by definition are far from the extreme values that endanger aquatic life and public health.

Interaction Terms. The analysis of variance makes explicit provision for interaction terms, and these provisions are standard features of many computer routines. There are strong interactions among most of the variables in water quality analysis; by ignoring these interactions, the analysis is grossly distorted.

Consider the simple case of interactions among riverflow, muddiness, and saltiness. During rainy spells, riverflow greatly increases. The large volumes of rainwater (which is very low in salt content) dilute the mineral salts in the riverwater; as the salt content decreases, the measurement error increases because it is more difficult to measure dilute salt solutions accurately. As riverflow increases, the turbulent velocity of the water increases. The raging flow muddies the water as the river tears up its channel. Rivers in flood carry enormous quantities of mud; rivers in drought carry very little. With every surge in riverflow, the mud content increases dramatically.

In short, high riverflows decrease salt content and increase mud content. This basic phenomenon, which can readily be seen, clearly illustrates the strong interactions among water quality variables. By ignoring these fundamental interactions, CEQ's analysis of variance distorts and oversimplifies the most basic properties of rivers.

Statistical Significance. All claims for statistical significance in analysis of variance are based on the assumption that the data have a normal bell-shaped distribution. In technical language, the analysis assumes that the data themselves (or their annual averages) are Gaussian. This assumption is false with respect to water quality. Water quality data are almost never Gaussian. Consider again the example of mud content. When riverflows are very high, rivers are very muddy; when riverflows are very low, rivers carry very little mud. Mud content in rivers is often very high or very low -- extreme values are common. With many extreme values, the data cannot have a bell-shaped distribution. Consequently, one of the underlying assumptions of the analysis of variance is not met and no conclusions can be drawn about the statistical significance of this analysis.

Mr. Ficke claims that the F values from the analysis of variance were tested for "significance at the 90 percent level."

This statement is a gross distortion. Because the data is not Gaussian, nothing can be said about the true statistical significance of the F values. Mr. Ficke surely means that he tested for significance at the 10-percent level -- not the 90-percent level. This fundamental error in elementary statistics shows how much this crew really knows about data analysis.

Even when the data themselves do not have a Gaussian distribution, the distribution of annual averages from these data may be Gaussian. However, CEQ used only three annual averages for each F statistic. It is preposterous to claim that three points define a Gaussian distribution.

Annual Averages. It is misleading to characterize a year of data by an annual average. The annual average oxygen content of the air we breathe is of less concern than a few minutes when the oxygen content may be zero -- which could be fatal. The same is true for fish. The annual average DO in a river is much less important than the few hours when the DO might be near zero; without DO, the fish would quickly die. It makes no difference how high DO might be later; once dead, the fish cannot come back to life.

This argument applies with equal force to all hazardous substances and conditions. Many species of fish are extremely sensitive to temperature during spawning season, but not at other times of year. Within broad limits, water temperature is irrelevant to their welfare except during spawning season. The annual average temperature is unimportant. Sudden temperature changes, especially during spawning season, are far more relevant, but annual averages tell us nothing about these sudden changes. With respect to the public health, bathing waters must not contain dangerous concentrations of bacteria when swimmers are there in the summer; it makes little difference if these waters contain bacteria during cold weather when beaches are closed.

The second goal of the Clean Water Act is to achieve "water quality which provides for the protection of fish, shellfish, and wildlife and provides for recreation in and on the water." Annual averages have no direct bearing on this goal; hazardous conditions are all that matters. Rather than doing an analysis of variance on annual averages, CEQ would be better advised to analyze violations of water-quality standards at every measurement point. Since water-quality standards are site-specific, violations will have to be tallied separately for each station.

3. Ficke Comment

"GAO staff reviewed all the above procedures and was clearly aware of how the work was done. The GAO draft was not critical of the method of computation."

Consultant's Evaluation

The case study of Cartersville did not comment on the errors and deceptions in CEQ's method of statistical analysis, and I explained why: "Our purpose was to cut through the standard statistical analyses to the data themselves, and through careful examination to assess the credibility of the reported changes in water quality." Now that Mr. Ficke has made a point of it, I have obliged him. Please refer to the preceding reply and to the discussion of the one-way classification (see p. 151), where my objections are set forth in detail.

4. Ficke Comment

> "GAO repeatedly states that the CEQ Annual Report claims that there were trends in water quality. In fact the 1978 AR does not refer to 'trends,' but only to 'changes'. To me a trend is a pattern that may be expected to continue, but the reported change is simply an observation of what happened."

Consultant's Evaluation

Mr. Ficke is welcome to his private definitions, but they are not consistent with the National Environmental Policy Act, which requires CEQ to report on trends, not changes:

> "The President shall transmit to the Congress annually * * * an Environmental Quality Report * * * which shall set forth * * * current and foreseeable trends in the quality [of the environment] * * * Each member [of the Council] shall be a person who * * * is exceptionally well qualified to analyze and interpret environmental trends * * * It shall be the duty and function of the Council * * * to gather timely and authoritative information concerning the conditions and trends in the quality of the environment both current and prospective, to analyze and interpret such information for the purpose of determining whether such conditions are interfering, or are likely to interfere, with the achievement of the policy set forth in title I of this Act, and to compile and submit to the President studies relating to such conditions and trends" (emphasis supplied).

The act clearly obliges CEQ to report on trends, not changes. Insofar as Mr. Ficke is serious about his private definitions of "trends" and "changes," he is abetting CEQ in violating the law and is suborning the Council to dereliction of duty.

5. Ficke Comment:

> "GAO goes to great lengths to argue that many of the changes in water quality in the James River were caused by climatic factors. They may be correct. In fact in the 1977 AR of CEQ we pointed out that the drought conditions in many places may have affected water quality. We probably should have repeated some of that material in 1978 AR, but space was short. A big point that GAO misses is that for many users of water (e.g., fish, municipal and industrial users) the reason for change is not important. They simply want to know what's going on. Part of CEQ's requirement under NEPA [the National Environmental Policy Act], also, is simply to report conditions and changes."

Consultant's Evaluation

Mr. Ficke has not read the law carefully. NEPA obliges CEQ to gather authoritative information concerning conditions and trends (not changes), to analyze and interpret them. Nowhere does the law say that CEQ "is simply to report conditions and changes."

CEQ's annual reports are hefty documents. It taxes credulity to read that droughts were ignored because "space was short."

In explaining these so-called "changes," CEQ did find the space to mention that new pollution control facilities and nonpoint sources should be held responsible:

> "Levels of suspended material, nutrients, oil and grease, oxygen-demanding substances, and other materials should decline as pollution control becomes more effective. Nonpoint sources are largely responsible for some of these substances."

Mr. Ficke should know that "levels of suspended material" in a river depend primarily on the turbulent velocity of the water. When riverflow is sluggish, the suspended material will settle -- no matter how much material you put into the water. Neither pollution control nor nonpoint sources explain how suspended materials behave in a river. "Climatic factors" and their hydrologic consequences do.

Unlike Mr. Ficke, I cannot say whether or how much fish want to know about "what's going on." For water quality management, however, it is essential to know why water quality is deficient so that efficient remedies can be devised. In replying to EPA and the Geological Survey, GAO explained in detail why NASQAN data fail to identify and explain "what's going on;" the principal failings of network data are time bias, spurious trends, inconsistent data, and failure to account for any of the rapidly changing properties of water.

6. Ficke Comment

> "Appendix VI of the GAO draft goes on at length with discussion of reasons for long-term and short-term variations in water quality, but much of the logic is faulty. It suggested that decreases in flow rate caused increases in dissolved solids concentration. But the discussion only shows coincidence, not cause and effect. They did not even test for a significant regression, much less show cause."

Consultant's Evaluation

The time has come to publish Horowitz' law: Rainwater is sweet because there are no large salt deposits in the sky. It is a fact that rainwater contains very little salt. Rivers in drought, however, may contain large amounts of salt because water evaporates, salt does not. There are salt deposits on earth and these salts are commonly found in groundwater or in rivers. As the proportion of rainwater in groundwater or a river increases, the salt concentration decreases. Sprey's Lemma is opposite here: Regression analysis is no substitute for having your head screwed on tight.

7. Ficke Comment

> "What GAO is really showing, however, is that the James is a complex river and that water quality is influenced by many complicated, interrelated factors. It's not predictable -- if it were there would be no need to monitor or conduct special studies or intense [sic] surveys. In reality, however, it would be necessary to do special studies of considerable magnitude to establish the kind of predictive modeling capability that is needed to make the type of predictions that GAO would like to have, or claims to have, in faulting the actual data."

Consultant's Evaluation

The James is a complex river -- far too complex to be meaningfully described or understood by a few buckets of water a month from an isolated sampling site. In replying to the Geological Survey, I demonstrated that the Survey had to resort to fairytale hypotheses in attempting to explain gross anomalies in the data -- suddenly vanishing nonpoint sources of zinc, biological reaeration without adequate data on algae, algae without chlorophyll, etc. Before water quality can be predicted, it must be understood. A few buckets of water a month from a remote sampling site cannot promote either understanding or prediction.

8. Ficke Comment

> "CEQ's needs to describe national conditions and trends would not be met by results of intensive surveys, each designed differently to explain what is affecting water quality. The proposal of Velz to repeat surveys each 5 years would not work either; all you would have then is a network with a sampling frequency of once each 5 years."

Consultant's Evaluation

Nothing in law or regulation obliges CEQ to use network data in its annual assessments of trends and conditions. But the National Environmental Policy Act explicitly directs the Council to "use authoritative information [and] * * * analyze and interpret such information for the purpose of determining whether such conditions and trends are interfering, or are likely to interfere, with the achievement of [national environmental] policy." Mr. Ficke admitted that complicated rivers are not predictable without doing "special studies of considerable magnitude." Without doing these studies, CEQ cannot fulfill its statutory obligation to determine whether such conditions and trends are interfering, or are likely to interfere with environmental policies and goals.

Mr. Ficke sees fixed networks everywhere, even in Professor Velz's suggestions for special surveys and followups. There is a big difference between a comprehensive analysis of water quality in a river (which requires sampling up and down the river, its tributaries, and its waste sources around the clock until its water quality is understood and the factors influencing it have been quantitatively evaluated) and network sampling (which involves no more than collecting a few buckets of water a month at a single, usually remote, sampling site). I cannot understand why Mr. Ficke cannot see this difference, but I suspect that most people can see it easily.

9. Ficke Comment:

> "GAO's analysis of the month-to-month, season-to-season, and year-to-year variations of quality of the James River is a good argument as to why a program of one-shot intensive surveys will not work. When would they have done the James -- 1975, 76, 77; spring, autumn, or winter? Would they have gone to the expense of adapting, verifying, and calibrating a model to explain all of the variations pointed out Appendix VI? If they had done it, how would CEQ use the data in its Annual Report?"

Consultant's Evaluation

The case study of the James River at Cartersville clearly showed that the NASQAN data from this site are a heterogeneous hodgepodge, and that the State's data from this same site frequently disagreed with NASQAN data. In commenting on this case study, the Geological Survey had to resort to fairytale hypotheses to explain gross anomalies in the data. But even if the Survey could have made sense of the data at Cartersville -- an impossible task -- it would still know nothing at all about water quality anywhere else on the James. It is impossible to describe, understand, or manage water quality in the James River from a heterogeneous hodgepodge of biased, inconsistent, distorted data from a remote sampling site. An intensive study can overcome all these difficulties when its purpose has been clearly defined and its execution is placed in competent hands.

The Survey's fairytale hypotheses amply illustrate how and when to conduct an intensive study. The Survey claims, for example, that biological reaeration may explain many of the apparent DO anomalies at Cartersville. This hypothesis can be rationally tested by comprehensively assessing the oxygen regime of the river during stable riverflows under two conditions: (1) when algae are abundant and (2) when algae are scarce. To evaluate the Survey's hypothesis of nonpoint sources of dissolved zinc, on the other hand, the river must be sampled at high flow (when nonpoint sources many be dominant).

There is nothing unmanageable about this rational approach to water quality studies. The only "trick" is to avoid collecting a heterogeneous hodgepodge of irrelevant data. The river must be studied in sufficient detail to understand its workings.

A little bit of common sense goes a long way. It has never been a secret that there is a large complex of chemical industries on the James below Richmond; it has never been a secret that the toxic pesticide Kepone was manufactured at Hopewell. It would not have taken ingenuity to study the James River for Kepone near Hopewell. If the State and the Geological Survey had studied this obvious source of toxic wastes, the tragedy of Kepone contamination of the James estuary could have been prevented.

A study of Kepone contamination near Hopewell would not have involved "the expense of adapting, verifying, and calibrating a model to explain all of the variations." It would have involved no more than looking for the obvious, then putting two and two together. Instead NASQAN data have never been collected anywhere near Hopewell and NASQAN has never collected data on Kepone. If the Survey had done a timely study on Kepone contamination near Hopewell, one of the most notorious pollution incidents of the decade might have been averted. It is probably not too farfetched to assume that CEQ might have found a way to incorporate such data into its annual reports.

Mr. Ficke asked when the James River should have been studied. Each major water quality problem defines its own period of study and its own locale. The Kepone problem should have been studied when it arose -- years ago -- but it should have been studied at Hopewell, not Cartersville. Water quality problems attendant on droughts should be studied during periods of drought, e.g. the summer of 1977. Only by studying water quality problems in their season can the investment in sampling pay off. How can CEQ develop authoritative information on the environment without thoroughly studying problems in their season? CEQ should be <u>insisting</u> on authoritative information about these seasonal variations, and should recognize that a few buckets of water a month from a single, generally remote, sampling site can never give authoritative information about these variations.

10. <u>Ficke Comment</u>

> "USGS staff statisticians have long argued that sampling at fixed intervals provides better data for trend analyses. The report does not mention USGS policy of extra samples for extreme events. It also does not mention the diurnal measurements of DO, conductance, and pH at NASQAN stations.

Consultant's Evaluation

These statistical arguments do not cut one way with respect to riverflow, the other way with respect to water quality. The Survey knows better than to measure riverflow once a month; it insists on continous riverflow measurements to ensure an adequate record. Everyone knows that water quality is greatly affected by riverflow. Why, then, does the Survey insist on continuous riverflow measurements, while pleading for the excellence of monthly water quality samples? This glaring inconsistency spotlights the contradictions in the Survey's arguments.

The "USGS policy of extra samples for extreme events" is a sometime thing. During the summer of 1977 central Virginia was in drought -- the worst in several years -- and the riverflow at Cartersville was extremely low. Yet there is no record of extra samples at the worst of the drought, nor were any extra samples taken during the subsequent floods. Cartersville is not exceptional in this respect. The Survey often fails to take samples during extreme floods because sampling sites are inaccessible or washed away. The summer of 1980 was one of the hottest and driest on record; how many extra NASQAN samples were taken during these extreme events?

At many NASQAN stations there are no "diurnal measurements of DO, conductance, and pH." At Cartersville, for example, the Survey has never measured DO between midnight and dawn, even though it suspects that biological reaeration may be important there. Cartersville is not exceptional in this regard. Mr. Ficke should know that the Survey takes very few samples at night and almost never takes samples after midnight.

11. Ficke Comment

> "CEQ only reported gross changes for more than 350 stations and did not attempt to explain why for each one."

Consultant's Evaluation

Agreed. CEQ could not have explained the changes in any event, owing to the weaknesses in the data. The case study of Cartersville and the Survey's comments on it show how little NASQAN data can really explain.

12. Ficke Comment

"GAO does not support that the change in dissolved solids is a 'normal consequence of a [sic] deepeending [sic] deeping [sic] drought.' It only shows that they were coincident. Neither is the 10-year trend contradictory. It is possible to have a 3-year change one direction within a 10-year change in the opposite direction. Such phenomena are common in most environmental and economic variables."

Consultant's Evaluation

I have dealt with this preposterous argument on page 194 of this volume. It is worth repeating that there are no large deposits in the sky, which explains why rainwater contains very little salt. Rivers full of rainwater have low salt concentrations; rivers in drought are saltier than rivers in flood. This simple fact shows that the trend in dissolved solids during the drought was "nothing more than a normal consequence of the deepening drought." Mr. Ficke's arguments are no better than his ridiculous spelling here.

I never argued that CEQ's 3-year increasing trend was contradicted by the Survey's 10-year decreasing trends. Here is what I wrote:

> "CEQ's 'statistically significant' trend is nothing more than a normal consequence of the deepening drought. As the drought deepened there was less surface runoff to dilute the groundwater and the James River became progressively saltier. GAO presented this case study to USGS for comment. In a letter of November 26, 1979, the Water Quality Specialist, Northeastern Region, USGS stated:
>
> > 'The U.S. Geological Survey data for the period 1969-79 showed a statistically significant <u>decreasing</u> trend in dissolved solids concentration in the James River at Cartersville, Virginia.' (Underscoring supplied.)
>
> Please notice that the Survey's 10-year trend was the opposite of CEQ's 3-year trend."

I cited the Survey's 10-year trend to illustrate an inherent weakness in CEQ's reports, viz. the instability of 3-year trends in NASQAN data. Because water quality is so variable, and because NASQAN data are so inadequate, the 3-year trends published in CEQ's reports will often bounce around from year to year. During the hard drought of 1980, many rivers certainly became saltier than in wetter summers. If 1981 and 1982 are wet years, CEQ will no doubt publish that dissolved solids improved over 1980. This instability (which may be attributed to tricks of chance in NASQAN sampling schedules superposed on real causes, such as floods and droughts) may explain Mr. Ficke's private definitions of "changes" and "trends."

In addition to gathering "timely and authoritative information concerning the conditions and trends in the quality of the environment," the National Environmental Policy Act obliges CEQ to "accumulate necessary data and other information for a continuing analysis of these changes or trends and an interpretation of their underlying causes." So long as CEQ relies on NASQAN data, it can never meaningfully interpret the underlying causes of trends in water quality. The Survey has frankly admitted that NASQAN data cannot -- and were never intended to -- explain causes. The Survey pointedly argued that it is "inappropriate" to judge NASQAN by such objectives as "determination of causes of existing water quality."

Unstable trends are inherent in CEQ's approach to data analysis, and NASQAN data cannot be used to determine the causes of the trends. The unstable trends make CEQ's reports misleading, and NASQAN data make it impossible for CEQ to fulfill its responsibility "for a continuing analysis of these changes or trends and an interpretation of their underlying causes."

13. Ficke Comment

> "Material on page [93] is not at all related to the validity of the CEQ data in the 1978 Annual Report, nor does it seem relevant to the matter of fixed stations versus intensive surveys. In fact, however, the relation between conductance and dissolved solids can vary as a function of ions and amount of silica."

Consultant's Evaluation

Table 4 of the case study gave examples of the Survey's Cartersville data that are difficult to make sense of. This table gives damning evidence of unreliability in the data CEQ used for its annual reports. Although the relation between conductance and dissolved solids can vary as a function of ions and amount of silica, the ratio of TDS to conductance cannot be so unstable or so extreme as it was in the data cited in table 4.

It is certainly relevant that questionable data are more readily identified in intensive surveys than in networks like NASQAN. In an intensive survey, scientists on the scene must read and use the data nearly every day. Intensive surveys force comprehensive data interpretation; when the data look doubtful, scientists on the scene can quickly discover anomalies, identify the causes, and fix the problem. In network data, on the other hand, anomalies and inconsistencies may persist for years before anyone notices them. The case study of Cartersville uncovered many examples of such problems.

14. Ficke Comment

"Data repeated by GAO support the use of this station in its count of stations showing an improvement as far as zinc concentration is concerned."

Consultant's Evaluation

The data I cited showed that concentrations of total zinc were higher before September 1977 than they were afterwards. Yes, these data show that zinc concentrations improved. I offered a rational explanation for this pattern, whereas neither Mr. Ficke nor the Survey has made sense of it.

15. Ficke Comment

Mr. Ficke stated that the discussion of total zinc "does not recognize the concentration of 40 micrograms per liter at a flow of 3,564 cfs" and * * * ignores that 'midnight dumping' or washout of holding ponds often produces heavy point-source loads at times of high flow."

Consultant's Evaluation

Mr. Ficke should read more carefully:

"Note that all the high zinc loadings were at riverflows greater than 4,500 cfs." (Emphasis supplied.)

A loading is not a concentration; if Mr. Ficke was unaware of the difference before, I trust that he understands it now.

Here is the rest of the paragraph, which contains the sentence Mr. Ficke found objectionable:

"At low riverflows, both concentrations and loads were quite low. This pattern suggests that total zinc did not behave like a point source of pollution and that suspended zinc usually accounted for most of the total zinc."

Like the Survey, Mr. Ficke has had to resort to a fairytale hypothesis--midnight dumping or the washout of holding ponds. He has offered not one scrap of evidence. I agree that such things are possible, just as I concede that once upon a time there may have been a Pied Piper of Hamlin. The overall pattern, however, suggested the conclusion I wrote:

> "The daily flow data suggest that total zinc behaved like a nonpoint source. Daily loads were generally high when riverflows were high and unstable, and were generally low when riverflows were low and stable."

Although this conclusion does not rule out fairytale hypotheses, it is consistent with the data and makes no flamboyant appeals to unbridled fancy.

16. Ficke Comment

> "Annual averages are 68, 18, and 13, using only USGS data, and 48, 17, and 14 using both state and USGS data. CEQ's conclusions would have been the same, even if it had the state data."

Consultant's Evaluation

The case study orginally read:

> "Table 5 shows that the State never reported any value greater than 20 ug/L, whereas the Survey reported several values of 70 and higher. Had CEQ used the State's data, they would have come to entirely different conclusions about zinc trends."

Mr. Ficke seems to have read this paragraph to mean that CEQ's conclusion would have been different if it had used data from both the State and the Survey. I regret the ambiguity and have removed it from the final report by inserting the phrase "rather than the Survey's data." (See p. 98 of Vol. I.) If CEQ had used the State's data rather than the Survey's, they would have concluded that there was no trend in total zinc.

17. Ficke Comment:

> "The first paragraph is untrue. Five of 11 values of '0' or '10' are at flows greater than 5,000 cfs. The last paragraph shows that GAO assumes a very simple model. If nature were as simple as these concepts and models, we never would need monitoring or intensive studies."

Consultant's Evaluation

Here are the two paragraphs Mr. Ficke objects to:

"The table shows that both concentrations and daily loads (pounds per day) of dissolved zinc were often lowest at low riverflows. Clearly, these dissolved zinc values can't be explained by the point-source hypothesis."

"If the laboratory is correct, the data record cannot be rationally explained by any hypothesis of pollution control. In short, the data record defies rational analysis. No supportable conclusions can be drawn from it."

The Survey advanced the hypothesis of a suddenly vanishing nonpoint source of dissolved zinc to explain these data. I discussed its failings on pages 168-170 supra. Mr. Ficke does not offer a counterhypothesis; he jibes that the "model" is "very simple" and that nature is not so simple. I am unaware of any model in these two paragraphs. I merely pointed out that pollution must come from point sources or nonpoint sources and that neither type of source is consistent with the data. I will happily entertain a more complicated explanation, should Mr. Ficke care to offer one. Meanwhile, I will stand by my conclusion: the data resist rational analysis.

18. Ficke Comment:

"Most of this discussion of DO is another case of GAO assuming very simple models for very complex situations. It is ridiculous to question the data just because they do not fit the analyst's preconceived ideas. As far as CEQ is concerned, GAO agrees (p. 93) that USGS data show an improvement and that we were correct in counting it as we did for the table in the 1978 AR. I might make one note here, however. The changes were small and the F value was not far above the criterion for 90 percent confidence. This might have been one case where the statistical analysis gave a false positive, as it will do 10 percent of the time, on the average."

Consultant's Evaluation

All of Mr. Ficke's claims for statistical confidence and significance are predicated on the unstated assumption that the data are Gaussian. This false impression and its misleading consequences are analyzed elsewhere in this report.

Mr. Ficke argues that it is ridiculous to question the data "just because they do not fit the analyst's preconceived ideas." I argued that there is something fishy when DO improves during a drought -- especially since the State's data at the same site showed nothing of the kind. I freely admit to preconceived ideas on both counts. Without a great deal of supporting information, it is difficult to accept CEQ's conclusion that DO really did improve during the drought. Two sets of data from the same sampling site ought not to disagree widely; but the State's data do not agree with the Survey's.

The discrepancy between the State's data and the Survey's speaks for itself. Mr. Ficke wisely refrained from offering to explain it. The Survey's explanation leads to the conclusion that CEQ's trend is merely an accident of a technician's work schedule.

Mr. Ficke also refrains from explaining why DO improved during the drought. The Survey's explanations, which centered on biological reaeration, came to grief when the data failed to support this line of argument; the data were nonexistent, untrustworthy, inconsistent, or all of the above. It is ridiculous to spend a fortune on data collection and never attempt to make sense of the records. I admit to the preconceived notion that data should be analyzed from time to time in an attempt to learn what we can from them. Data analysis means more than feeding numbers into a computer for statistically suspect analysis of variance; it means a serious attempt to relate the numbers to reality. I made this attempt in the case study and carefully set down all the arguments and the data they were based on. Mr. Ficke has done nothing of the kind. He contends that the "models" (which are more properly described as arguments) are very simple, but fails to explain why or to show how more complex explanations would be better. Surely Mr. Ficke can grace us with rational arguments if my Cartersville analysis is so poorly argued as he suggests. Until he does so, I can do no more than I have done already, namely, to argue the case from the facts.

19. Ficke Comment

"These conclusions merely restate some poorly supported cases made by GAO regarding effects of a drought, relations between conductance and dissolved solids, the matter of trend versus change, calling data 'suspect' if they do not fit a preconceived notion, and simplifying a situation that really is very complex. On this last point, the last paragraph on page 102 is partly correct, and partly wrong. Just because the data are not understood by the investigator does not make the data wrong. I agree that more technical work would be helpful, but it is out of line to charge that CEQ should not use the data to describe national conditions and changes in water quality."

Consultant's Evaluation

Again Mr. Ficke jibes that I have preconceived notions and that I oversimplified complex situations without facts or arguments to substantiate these muddy sobriquets, I will tolerate the name-calling as a minor annoyance. The conclusions summarize over 30 pages of detailed data analysis and technical argument.

Here is the paragraph Mr. Ficke characterized as partly correct and partly wrong, the case study reads:

> "On close technical examination, the DO data at Cartersville fail to make sense. The State and the Survey -- the agencies that have collected these data -- now agree that much more detailed technical work will be necessary before the DO at Cartersville can be fairly assessed. Until that careful work has been done, the existing water quality records should be used with extreme caution and no conclusions should be drawn from them."

Mr. Ficke neglected to say what he found wrong. His objection is based on the recommendation that no conclusions should be drawn from data like those. I offered nearly 20 pages of data analysis to support this paragraph. Mr. Ficke has been unable to fault any of the facts or arguments, but objects to the inescapable conclusion from these facts and arguments. Well might he, since he was largely responsible for the misleading conclusions CEQ drew from these doubtful data.

☆U.S. GOVERNMENT PRINTING OFFICE: 1981—341-843:636

(087170)