131802

United States General Accounting Office

GAO

Report to the Administrator
Environmental Protection Agency

December 1986

WATER QUALITY

An Evaluation Method for the Construction Grants Program— Methodology

131802

GAO/PEMD-87-4A

Volume 1

GAO

United States
General Accounting Office
Washington, D.C. 20548

Program Evaluation and
Methodology Division

B-221558

December 17, 1986

The Honorable Lee M. Thomas
Administrator, Environmental Protection
Agency

Dear Mr. Thomas:

Within the past 2 years, several important studies have been published that address present water quality in the United States, how water quality has changed, what pollution sources exist, and whether the Construction Grants Program has contributed to any water-quality improvement that may have occurred. In particular, GAO's recent report entitled The Nation's Water: Key Unanswered Questions About the Quality of Rivers and Streams (GAO/PEMD-86-6) reviewed these studies and reported on the gaps in what is known about the effect of the Construction Grants Program on the quality of the nation's water. The present report demonstrates a way of closing some of these gaps by using data and methods of analysis available to the Environmental Protection Agency. This is the first volume of the report. It contains recommendations to you in chapter 3.

As you know, 31 U.S.C. 720 requires the head of a federal agency to submit a written statement on actions taken on our recommendations to the House Committee on Government Operations and the Senate Committee on Governmental Affairs not later than 60 days after the date of the report and to the Committee on Appropriations in the House and in the Senate with the agency's first request for appropriations made more than 60 days after the date of the report.

We are sending copies of this report to appropriate House and Senate committees, members of the Congress from the states mentioned in the report, and the director of the Office of Management and Budget. We will also make copies available to interested organizations, as appropriate, and to others upon request.

Sincerely yours,

Eleanor Chelimsky
Director

Executive Summary

Purpose

Since the Clean Water Act was enacted in 1972, the federal government has spent more than $39 billion to assist in constructing and upgrading municipal sewage-treatment plants under the Construction Grants Program. The U.S. Environmental Protection Agency (EPA) has not yet performed an adequate evaluation of the program's effectiveness in cleaning up the nation's waters, although extensive water-quality data have been collected. GAO undertook this research to develop guidelines for evaluating treatment plant upgrades and tested it in four case studies, using only available data and software available to EPA.

Background

The majority of the Construction Grants Program funds are spent on increasing the capacity of existing municipal sewage-treatment plants and improving their efficiency in removing specific pollutants from the wastewater they discharge into rivers and streams. Past evaluations of the program's activities are deficient inasmuch as they examined only the changes in plant efficiency that resulted from upgrades; some evaluations failed to demonstrate rigorously the connection between changes in plant discharge and changes in stream water quality.

Results in Brief

An adequate evaluation of the Construction Grants Program should be based on stream water quality and should address four issues: the changes in plant discharge that result from an upgrade, the history of water quality in the receiving stream, the relationship between these two measures over time, and the possibility of alternative explanations for stream water-quality changes. Using existing EPA data and software and applying common statistical techniques, GAO evaluated the effectiveness of four treatment plant upgrades funded by the program.

Stream data for adequately assessing the effect of all treatment plant upgrades do not yet exist, but GAO believes that many upgrades could be assessed with the data that do exist. Assessing them would provide a more realistic estimate of the program's effectiveness than is now available.

GAO's Analysis

GAO believes that its method is appropriate because the method successfully answered four essential evaluation questions with extant data and available software.

Analysis of Plant Effluent

1. Did the upgrade of the sewage-treatment plant decrease the amount of pollutants the plant discharged?

A plant whose treatment level has been raised clearly has not improved water quality if there is no decrease in the concentration of pollutants it discharges. GAO's method uses data related to pollutants discharged from treatment plants, including such measures as suspended solids, biochemical oxygen demand, and ammonia. Before-and-after pollutant measures are compared in common statistical tests. Analyzing discharge reports that sewage-treatment plants are required to submit to EPA, GAO found statistically significant postupgrade decreases in the pollutants discharged from each plant that it examined. (See pages 20-22 and 32.)

Stream Water-Quality Changes

2. Did water quality improve downstream from the treatment plant?

An upgrade that has not improved the quality of water downstream from the plant has not succeeded in its ultimate purpose. GAO analyzed the history of the water-quality indicators affected by treatment plant discharges as recorded at water-quality monitoring stations. The analytic procedures GAO applied to these data to improve the sensitivity of its analysis included adjusting the readings for variations in stream flow, temperature, and upstream pollution levels and performing separate analyses of low-flow observations. GAO found statistically significant improvements in downstream water quality in three of the four cases. These changes were more evident under low-flow conditions. (See pages 22-28 and 32.)

Relationship Between Plant and Stream Indicators

3. Were changes in the plant's effluent related to changes in stream water-quality indicators?

Changes in stream water quality must be shown to be associated with changes in a treatment plant's effluent before they can be attributed to the upgrade. GAO correlated the available stream readings with the discharge data in the four cases and found that, for the most part, changes in plant discharge were moderately reflected in stream water-quality observations. GAO concluded that this procedure tends to understate the connection, because the discharge undergoes biochemical changes after entering the stream, and because the two sets of records use different data-recording methods. (See pages 29-30 and 33.)

Alternative Explanations

4. Can other reasonable explanations of water-quality conditions be excluded?

A correlation between changes in a plant's discharge levels and stream indicators does not mean that the plant's upgrade is the sole determinant of a change in water quality downstream. GAO examined the discharge records of other sources of pollution in the vicinity of the plants and estimated the relative importance of each source. GAO was able to distinguish the effect of the upgrade from the effects of other major pollution sources. (See pages 30-31 and 33.)

Recommendations

GAO recommends that the administrator of EPA perform additional evaluations of treatment plant upgrades that use available data and methods similar to those developed by GAO. The purpose of these evaluations should be to assess, insofar as possible, the effects of the Construction Grants Program on stream water quality.

Agency Comments

U.S. Environmental Protection Agency

EPA generally agreed with GAO's methodology and its application to the four case studies but said that it is applicable to only a small fraction of plant upgrades. GAO believes that the methodology could be used for more upgrades than EPA estimated but agrees that the particular data sources and statistical methods GAO used in the case studies are not universally applicable and would have to be supplemented in many analyses. However, GAO believes that this consideration does not affect the need to examine the empirical evidence on stream water quality in assessing the effects of the Construction Grants Program, particularly where relevant data are already available to EPA. (See pages 49-53.)

U.S. Geological Survey

The U.S. Geological Survey generally agreed with GAO's concern that evaluations of upgrades be based on empirical evidence and offered some specific technical comments on the methodology.

GAO made changes to the report, as appropriate, basing them on the agencies' comments, which appear in appendixes I and II.

Contents

Figures

Abbreviations

ASIWPCA	Association of State and Interstate Water Pollution Control Administrators
DO	Dissolved oxygen
EPA	U.S. Environmental Protection Agency
GAO	U.S. General Accounting Office
GICS	Grants information control system
NASQAN	National stream quality accounting network
NPDES	National pollutant discharge elimination system
STORET	STOrage and RETrieval

Introduction

Background

The literature reporting on the condition of the nation's water quality in the late 1960's and early 1970's is replete with accounts of water pollution. For example, in 1970, the Federal Water Quality Administration, responsible for implementing federal water pollution laws, reported that

> "Almost any day, in the waters near any large population center in the United States and, increasingly, in the countryside, we can see the signs of water pollution . . . . Use of our waters to receive and carry away wastes has seriously damaged our ability to enjoy other water uses, such as swimming and boating, sport and commercial fishing." (U.S. Federal, 1970, p. 7)[1]

More recently, the Association of State and Interstate Water Pollution Control Administrators (ASIWPCA) recalled that in the early 1970's, "reports were all too frequent of fish disappearing from rivers and streams, lakes choked with algae, and beaches posted against swimming or shellfishing" (ASIWPCA, 1984, p. 2).

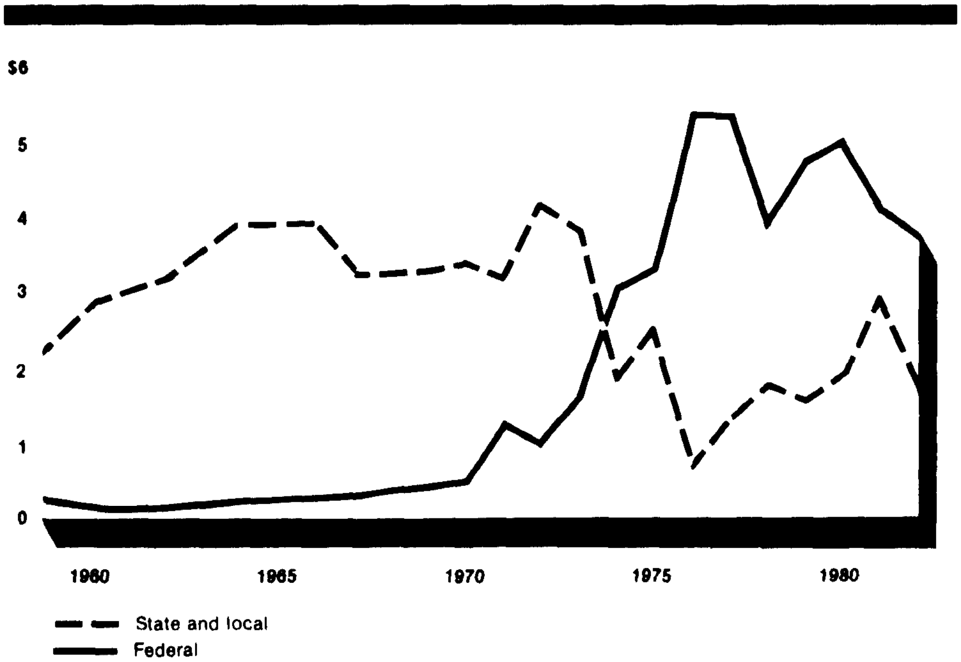
From the reaction of citizens to pollution problems, in the form of increased awareness and ecological concern, major environmental legislation was enacted in the early 1970's. Foremost in respect to water quality were the 1972 amendments to the Federal Water Pollution Control Act (Public Law 92-500). In 1977, this legislation became known as the Clean Water Act. Section 101 of the Clean Water Act of 1972 stated that in order to restore and maintain the physical, chemical, and biological integrity of the nation's water, it is national policy to provide federal financial assistance to construct publicly owned waste-treatment plants, among other things. Title II, section 201, authorized the Construction Grants Program, which requires and assists in the development and implementation of waste-treatment management plans and practices, in order to achieve the goals of the Clean Water Act. Title II also authorized the administrator of the U.S. Environmental Protection Agency (EPA) to make grants to any municipality, intermunicipality, state, or interstate agency for the construction of publicly owned treatment works.

The 1972 act significantly changed the funding, particularly federal funding, provided for publicly owned wastewater-treatment projects in the United States. Under this act, any federal grant for treatment works may fund up to 75 percent of the cost of construction. Prior to the passage of the Clean Water Act, most expenditures for wastewater treatment were from state and local sources. Since the passage and

[1]Interlinear citations are spelled out in the bibliography.

implementation of this act, federal funding for wastewater facilities has dominated. (See figure 1.1.)

Figure 1.1: Federal, State, and Local Funding for the Construction of Wastewater-Treatment Plants in 1959-82



[a]1982 constant dollars in billions. The state and local data are in doubt.

Source: EPA, Study of the Future Federal Role in Municipal Wastewater Treatment (Washington, D.C.: 1984), p. 3-2.

The federal government has spent approximately $39 billion to assist municipalities to construct wastewater-treatment facilities. The bulk of this, or $37 billion, has been spent since the implementation of Public Law 92-500—that is, in the last 13 years. This makes the Construction Grants Program one of the nation's largest nonmilitary public construction programs. Changes in federal authorization for municipal assistance are shown in table 1.1.

Table 1.1: Federal Assistance to Municipal Wastewater Treatment in 1948-84

Public law	Annual authorization	Type of funding	Maximum federal share
80-845 (1948)	$22.5 million	Loans	Lesser of 33-1/3% or $250,000
84-660 (1956)	$50 million	Grants	30% or $250,000
87-88 (1961)	$80 million FY 1962 $90 million FY 1963 $100 million FY 1964-67	Grants	30% or $600,000
89-234 (1965)	$100 million	Grants	Greater of 30% or $1.2 million
92-500 (1972)	$18 billion FY 1973-75	Grants	75%
95-217 (1977)	$1 billion FY 1977 $4.5 billion FY 1978 $5 billion each FY 1979-82	Grants	75% or 85% for "innovative or alternative processes"
97-35 (1981)	$2.548 billion FY 1981	Grants	75% or 85% for "innovative or alternative processes"
97-117 (1981)	$2.4 billion each FY 1982-85	Grants	55% starting in FY 1985

Authority to obligate funds for the Construction Grants Program under the Clean Water Act expired at the end of fiscal year 1985, and the disagreement was significant within the Congress, and between the Congress and the administration, concerning what level of funding should be reauthorized. For example, at the October 1983 meeting of EPA's management advisory group for the Construction Grants Program, the executive director of the Association of Metropolitan Sewage Agencies said that the current grant authorization level of $2.4 billion annually was not enough to enable communities to achieve the federally mandated clean-water goals within the specified time (Feinthel, August 3, 1984, p. 18). But at an April 7, 1983, ASIWPCA seminar on financing sewage treatment, officials of the Water Pollution Control Federation argued that the Congress never intended the EPA grants program to continue indefinitely and proposed that the Construction Grants Program be phased out over 10 years (Feinthel, p. 6).

Currently, EPA is attempting to identify feasible alternative financing schemes to the federally funded Construction Grants Program. This effort is intended to gradually remove the federal government from the financing of wastewater-treatment plants. The president's fiscal year 1986 budget proposed phasing out the program by 1990 and disallowing new construction after fiscal year 1985.

Bills that would authorize the funding of the Construction Grants Program into the 1990's have been passed in both houses of the Congress.

In July 1985, the House authorized annual expenditures of $2.4 billion through 1990 (in H.R. 8) for the construction of sewage-treatment plants and an additional $12 billion in grants in fiscal year 1986-94 to enable states to set up revolving loan funds to finance them. With S. 1128, the Senate had already approved amendments to the Clean Water Act in June, in which $2.4 billion would be authorized annually for fiscal years 1986-88 and another $1.4 billion for fiscal years 1989-90; this bill would also authorize a total of $8.4 billion in fiscal years 1989-94 for a revolving loan fund for the states.

EPA has expressed misgivings about the level of funding called for in these bills. Shortly before the House voted on a proposed reauthorization of the Construction Grants Program in 1984, the administrator of EPA was reported to have written to the House minority leader, expressing serious reservations about the House bill, especially the multibillion dollar increases in authorizations for wastewater treatment. The letter suggested that any bill with higher construction grants authorizations could draw a presidential veto ("Bipartisan Support . . . ," June 22, 1984, p. 292). In a letter read to the House during floor debate on H.R. 8 in 1985, the EPA administrator criticized the funding levels as excessive, particularly in view of the serious deficit problems in the national budget.

Why is there major disagreement on the appropriate funding level for the Construction Grants Program? One reason may be the lack of adequate information on the program's effectiveness. As we pointed out in The Nation's Water: Key Unanswered Questions About the Quality of Rivers and Streams (GAO/PEMD-86-6, September 1986), no adequate national evaluation of the program has been conducted.

Objective, Scope, and Methodology Overview

Objective

The objective of this evaluation was to determine whether available data could be used to draw conclusions about the effect of expenditures under the Construction Grants Program. Specifically, our aim was to answer the following question: Is it possible to develop and demonstrate a method for evaluating the effect of upgrading a sewage-treatment plant on the health of the body of water into which the plant's effluents

<u>are discharged?</u> To answer this question affirmatively, we needed to develop a methodology capable of responding to the following four questions.

1. <u>Did upgrading the sewage-treatment plant decrease the amount of pollutants that it discharged?</u> This is the sine qua non of a case-specific evaluation. If an improvement in a plant's discharge history has not occurred after an upgrade, further investigation of the upgrade's effectiveness would be pointless. It may seem self-evident that upgrading the level of treatment applied to waste must decrease the quantity of pollution discharged into a receiving stream, but an upgrade may not have this effect in all cases. For example, a simultaneous expansion in the amount of influent waste may mean that the pollutant discharge does not decrease substantially, or it may even increase. In a case like this, it would be critical to examine the change in the plant's efficiency in removing pollutants from the influent, not simply the pollutant loading in its effluent.

2. <u>Did water quality improve downstream from the sewage-treatment plant?</u> The primary goal of upgrading a sewage-treatment plant is to improve, or at least maintain, the health of the receiving stream. An upgrade that does not reach this goal is ineffective. We noted in <u>The Nation's Water: Key Unanswered Questions About the Quality of Rivers and Streams</u> (cited above) that most of the previous assessments of upgrades funded by the Construction Grants Program have not adequately examined the effect of upgrades on water quality in streams below the treatment plants.

3. <u>Is there a relationship between changes in a plant's effluent and changes in stream water-quality indicators?</u> Changes can occur in streams for many reasons. The simple fact that a stream's water quality improved after an upgrade does not imply a causal connection. To assert a causal connection, one must identify some positive evidence of association between the decrease in pollutant discharge from the upgraded treatment plant and water-quality improvement where downstream observations are made.

4. <u>Can other reasonable explanations of a stream's water quality be excluded?</u> Any attempt to assess the stream effect of an upgrade must take into account changes in other pollution sources. An upgrade in a treatment plant may be only one part of a larger effort to improve regional water quality, and its effect can be confused with the results of actions intended to decrease pollutant discharge from other municipal or

industrial point sources or even with changes in activities producing nonpoint-source pollution. Conversely, the stream effect of an upgrade may be masked by coincidental increases in pollutant discharge from other sources. It may be impossible in many cases to rule out absolutely the possibility of alternative explanations of water-quality changes, but some attempt must be made to identify and assess the likely effect of changes in the profile of discharge from other significant sources of pollution in the vicinity of the upgraded treatment plant.

These four questions address the one overall question of whether an upgrade funded by the Construction Grants Program improves stream water quality. They are integrally linked to one another in the logic of our presentation and in our analysis. All the possible sets of responses to the four questions can theoretically assume 16 different patterns, which can be reduced to the four basic patterns presented in table 1.2.

Table 1.2: Alternative Response Patterns for Our Evaluation Questions[a]

Pattern of response		Did upgrade improve water quality?	Comment
1.	Q1: No Q2: Yes or no Q3: Yes or no Q4: Yes or no	No	Effluent failed to improve
2.	Q1: Yes Q2: Yes Q3: Yes Q4: Yes	Yes	Improvements in stream water quality are attributable to sewage-treatment plant because other factors can be excluded
3.	Q1: Yes Q2: No Q3: No Q4: Yes	Unknown	Effluent improved but not stream water quality, because of other factors; upgrade may have maintained water quality, which would otherwise have been degraded
4.	Q1: Yes Q2: Yes or no Q3: Yes or no Q4: No	Unknown	Cause cannot be established between upgrade and stream water quality

[a]The questions shown in the first column are
Q1: Did the sewage-treatment plant upgrade decrease discharged pollutants?
Q2: Did downstream water quality improve?
Q3: Were changes in effluent related to changes in stream water-quality indicators?
Q4: Can alternative explanations be excluded?

In the first pattern shown in the table, an upgrade resulted in no significant decline in the concentration of pollutant discharge from a sewage-treatment plant, answering the first question in the negative. The remaining questions become academic: the upgrade was patently ineffective. In the second pattern, both the effluent quality and the stream water quality improved, and other possible explanations of the stream improvement have been excluded. Therefore, a causal relationship may be inferred.

Patterns 3 and 4 in table 1.2 represent gray areas in which causal attribution cannot be made with certainty. Pattern 3 could occur in a situation in which an upgrade was effective in maintaining stream water quality at a constant level despite substantial increases in pollution from other sources. (Confirmation would probably require traditional water-quality modeling procedures to hypothesize what the stream water-quality would have become without the upgrade.) Alternatively, pattern 3 could appear if there were severe measurement error in the stream water-quality data or if the downstream monitoring station were located inappropriately for measuring the influence of changes in the treatment-plant's effluent.

Pattern 4 represents a situation in which, despite clear improvement in a treatment-plant's effluent, other pollution sources exert significant influence on a stream's health and cannot be excluded as an explanation of observed variations in stream water quality. In this case, even if a significant improvement has occurred in stream water quality, it cannot be definitively attributed to the upgrade and any resultant improvement in effluent quality.

For each of these four questions, we applied traditional statistical techniques to the relevant data. Although we reviewed many data sources, all the data we report in this study came from the STORET data base, except for the discharge monitoring reports, most of which we collected at EPA's region III headquarters. STORET is EPA's computerized data base and contains the results of tens of millions of samples taken from more than 200,000 unique collection points. EPA has also begun to computerize the data from discharge monitoring reports, which will make all the data bases we used for this report accessible from one computer system.

Scope: Restrictions on Methodological Applications

Our intent was to develop and apply methods of evaluating the effect of upgrading a sewage-treatment plant on the health of the body of water into which the sewage-treatment plant discharges. We set five important limitations on our methodological research. First, we excluded analysis of influences on oceans, estuaries, and lakes, each of which presents special problems, and we restricted our analyses to rivers and streams.[2]

Second, we did not attempt to examine the stream effect of new construction; we examined only sewage-treatment plant upgrades funded by the Construction Grants Program. We included in the category of upgrades any change to a plant intended either to increase its capacity or to improve its efficiency in removing pollutants from its waste discharge. The program's funds can be used to upgrade existing plants and to construct new plants. We limited the application of our methods to case studies of upgrades because these represent the bulk of the program's expenditures, both past and projected. Most of the funding under the Construction Grants Program has gone into upgrades. In EPA's 1984 needs survey report, it estimated that more than three quarters of plant-related expenditures to be needed between 1984 and 2000 will be spent on changes to existing plants.

We believe, however, that our methodology could be applied to new construction, although an interpretation of the results might be problematic in some cases. For example, for a plant newly constructed to serve a population previously dependent on septic tanks for waste disposal, we would anticipate some degradation in downstream water quality, because of the additional pollutant load from the plant that, prior to its construction, had been dispersed into the ground and entered the groundwater table.

Third, we restricted our analysis to locations where relevant data were already adequate and available. This restriction reduced the number of treatment-plant changes we could investigate and, to a large extent, dictated the type of data we used in our study. We and others have recommended that water-quality studies not confine their definitions of water quality to traditional chemical and biological parameters (like dissolved oxygen and fecal coliform bacteria), but in most cases, only these measurements have been collected long enough and with sufficient frequency to allow meaningful statistical analysis. Furthermore, these parameters are the criteria by which EPA regulates the quality of discharge it allows from sewage-treatment plants. For these reasons, we employed traditional water-quality criteria.

[2]We have used the terms "river," "stream," and "water body" interchangeably in this report, except where the context requires otherwise.

Fourth, we restricted our evaluation to wastewater-treatment plant upgrades that received funding under Public Law 92-500. This did not severely restrict our case selection, because most upgrades since 1972 received Public Law 92-500 funds, and it allowed us to focus on the principal federal wastewater-treatment program.

Fifth, we confined our analysis to treatment plant upgrades in Pennsylvania. This allowed us to conserve resources and concentrate our identification and collection of existing data. We selected four diverse cases involving, among others, large and small communities; rural and urban settings; dominant agricultural, industrial, and municipal pollution; point and nonpoint pollution sources; and treatment-plant upgrades of different complexity. Within these constraints, we selected from a list of 10 projects under the Construction Grants Program 4 sites where, in the opinion of our expert consultants, we would be most likely to detect the influence of the upgrade on stream water quality from extant downstream monitoring station data. The sewage-treatment plant upgrades we examined are outlined in table 1.3.

Table 1.3: The Sewage-Treatment Plants in Our Four Case Studies

Sewage-treatment plant	Flow (mgd)[a]	Cost (million)[b]	Date[c]	Receiving stream
Allentown	29.9	$14.9	9/79	Lehigh River
Hamburg	0.5	1.0	7/76	Schuylkill River
Lansdale	2.4	16.1	1/81	Neshaminy Creek
Tamaqua	1.1	2.2	6/77	Little Schuylkill River

[a]Post-upgrade average.

[b]Grants information control system data.

[c]Effective date, or the date on which the effect of an upgrade becomes observable. This definition is explained in chapter 2.

We chose the case study method because existing data were available with which to test the feasibility of our new approach to measuring the effect of treatment-plant upgrades on the quality of the water into which the treatment plants discharge their effluent. We chose a multiple case study design because we were testing a new approach; therefore, much of our developmental work was exploratory and subject to iterative changes. Case studies were an excellent vehicle for this endeavor because they allow, with limited resources, an analysis of data availability and data accuracy and a test of whether a methodology can be developed that is sufficient to address cause-and-effect questions.

Data and Analytic Methods

We organized each case study around the four questions enumerated above. In this section, we briefly describe the data we used to address each question and the statistical techniques we applied to these data. We discuss the data and techniques in greater detail in chapter 2.

Amount of Pollutants Discharged

Data Used. For our source of information on effluent changes, we used the discharge monitoring reports submitted by the plants. Each sewage-treatment plant is required to submit to EPA and the state a summary record of its monthly pollutant discharge. Each report includes the average daily amount of each pollutant for which a plant's permit sets limits. A typical report includes the amount of wastewater discharged, or flow; total suspended solids; biochemical oxygen demand, or BOD; and fecal coliform bacteria. If there are particular problems in a river, the plant may also be required to report its average daily discharge of ammonia, phosphorus, or some other effluent constituent. We extracted from these reports all available monthly averages of flow, total suspended solids, BOD_5, fecal coliform bacteria, ammonia, and phosphorus for the period from at least 2 years prior to the upgrade to the date of the latest available data.

Analytic Method. We computed the mean of these monthly averages for the pre-upgrade period and for the entire postupgrade period. We tested these two means for a statistically significant difference for each type of pollutant. We also calculated the annual mean of each effluent constituent for each year of record. We have included these statistics in the second volume of this report.

Downstream Water Quality

Data Used. The data base we used to determine stream water quality for each case study consisted of observations of standard water-quality parameters made at fixed monitoring stations in the vicinity of the sewage-treatment plant. A monitoring station is less than 10 miles downstream from each treatment plant in the four case studies. For one case, monitoring-station data from an upstream location were available and served as baseline data.

The observations on water quality we extracted from the monitoring-station records included measures of the stream concentration of dissolved oxygen, BOD_5, ammonia, nitrite, and nitrate. These measurements were typically taken once a month from the mid-1970's to the present and less frequently prior to the mid-1970's. All the data we used for this question are in STORET.

Analytic Method. We applied some tranformations to these data to allow meaningful analysis. In particular, we adjusted the data to compensate for their representation of chemical concentrations observed at different flow levels in the rivers. We then compared the average pre-upgrade and postupgrade readings and tested for statistically significant differences between the two readings. We made these comparisons for averages of "all observations" and, separately, for averages of "low-flow observations."

Relationship Between Effluent and Stream Water-Quality Indicators

Data Used. The data base we used to answer this question consisted of the data for the two preceding questions: monthly average discharge, taken from the plants' discharge monitoring reports, and the STORET records of stream conditions for the same period.

Analytic Method. We computed correlation coefficients for the relationship between the average plant effluent levels for each month (total suspended solids, BOD_5, fecal coliform bacteria, and, if available, ammonia) and the observations of stream concentrations of dissolved oxygen, BOD_5, ammonia, nitrite, and nitrate during the corresponding months.

Other Reasonable Explanations

Data Used. We identified all the significant point-source dischargers of pollution that had received permits to discharge into the stream above each monitoring station whose data we used in assessing the effect of an upgrade. We collected the discharge monitoring reports of these dischargers and extracted the same information we extracted for the upgraded treatment plants. We added the monthly averages in these reports to the data base for the preceding question.

Analytic Method. The methods we used to investigate alternative explanations of stream changes depended on the complexity of each case. In some instances, a simple examination of the relative volume of discharge from point sources other than the treatment plant clearly indicated that the plant was the dominant point source of pollution. In these instances, there was no need for further analysis. In less clear cut instances, we used multiple regression analysis to compare the relative influence of the rival point-source explanations for downstream change.

Methodology

In this chapter, we provide the details of the method we applied to our case studies in order to respond to each of the four questions we defined in chapter 1 as the components of an evaluation of the effect of sewage-treatment plant upgrades. Slightly paraphrased, the four questions are outlined as the four principal headings below.

Did the Upgrade Decrease Discharged Pollutants?

Data Sources: Discharge Monitoring Reports

Copies of discharge monitoring reports are maintained at three locations: at the sewage-treatment plants, which are required to maintain these reports for at least 3 years; at the regional office of Pennsylvania's Department of Environmental Resources, and at EPA's regional office. We obtained most of our records from EPA's region III, because of its central location and because it maintains a comprehensive set of reports. Some of the reports are in hard copy; others are on microfiche. Recently, some of the reports have been entered into a computerized data base called the "permit compliance system."

From the discharge monitoring reports, we extracted the readings for average flow and the monthly average effluent loadings of suspended solids, BOD_5, fecal coliform bacteria, and ammonia, where it was available. These averages are based on samples of the concentration (usually expressed as milligrams per liter, or mgl) of these constituents taken with varying frequency during the month. For some plants, multiple samples are taken 1 day a week, and a "composite sample," or the average of the readings from all samples taken that day, is computed. For other plants, one "grab sample" is taken to represent a week's reading. These concentrations are converted into loadings by multiplying by flow levels for the day and are then averaged, in order to provide daily average effluent loadings for the month.

Effective Date of Upgrade

Determining the exact date on which one should expect to see the effect of an upgrade on a plant's discharge history is difficult, at best. It would entail a level of data collection and expert investigation beyond that necessary for our purposes. Rather than employ an a priori method to determine an exact date, we used three independent sources of information: the data from the sewage-treatment plant, the discharge monitoring reports, and EPA's grants information control system (GICS).

GICS maintains, in a computerized data base, various types of information about Construction Grants Program projects, including funding levels and milestones. We used the GICS "works in operation" date to screen for projects for which a sufficiently long time-series of effluent and water-quality observations could be expected for both before and after the upgrades. We also verified this date by interviewing sewage-treatment plant personnel.

The ultimate criterion for determining an operational date for an upgrade, however, is the discharge monitoring report. Every upgrade that improves the efficiency of waste treatment should be marked by a clearly observable decrease in the concentration of effluent loadings from the treatment plant. The date we used to demarcate the pre-upgrade and postupgrade periods was derived from an inspection of effluent records from around the time suggested by GICS and the treatment plant personnel.

Because of the nature of some upgrades, we also developed a three-level definition of upgrade phases and applied it to the Allentown case study, the largest treatment plant in our sample. Upgrades frequently involve the modification, replacement, or addition of several component subsystems, and these components are brought on-line serially when they have been completed. Even after all construction has been completed, a proper balance between components is achieved in a period of adjustment and stabilization, when system dysfunctions are resolved and plant operators acquire expertise in the most efficient operation of the new equipment. This period may last for a year or more.

For this reason, an analysis that defined an effective upgrade date as the time GICS or plant personnel indicate that the upgrade became fully operational, or as the point of sharpest decline in effluent loadings record in the discharge monitoring reports, could underestimate the effect of the upgrade by assigning a portion of the reduction in pollution achieved by the upgrade to the pre-upgrade period. A more realistic estimate of an upgrade's effect might be obtained by excluding from analysis a period of 12 to 24 months prior to the completion date as a transitional period and comparing postupgrade data with the pretransition data.

We used the latter method in analyzing the upgrade at Allentown. For the Allentown plant, our examination of the discharge monitoring reports suggested that effluent pollutants declined gradually, but unevenly, for some 20 months before the completion of the upgrade. However, we failed to find any substantial difference in analytic power between the two definitions of the pre-upgrade period, mostly because the Allentown plant incompletely reported its discharge monitoring data during the construction.

Since the exclusion of a transitional period diminishes the power of statistical tests and the ability to detect differences between the two periods by reducing the number of usable pre-upgrade observations, we do not recommend that this approach be used routinely in evaluations of the effect of the Construction Grants Program. We believe it should be considered only where discharge monitoring reports indicate an extended transitional period of instability. Our case study experience leads us to believe that it is preferable in most cases to base the effective date of an upgrade on an inspection of discharge monitoring reports supplemented by information supplied by GICS and plant personnel.

Analytic Methods

We computed the mean level of each of the effluent parameters we extracted from the discharge monitoring reports for postupgrade and, where appropriate, pretransition as well as pooled pre-upgrade and transition observations. We computed a t statistic to determine the statistical significance of the difference between the mean of the postupgrade levels and the mean of the pre-upgrade or pretransition level. In the case study chapters, we report the results of these tests and present the annual mean for each effluent constituent.

Did Downstream Water Quality Improve?

Data Sources: Fixed Monitoring Stations

Since our intent was to examine the effect of sewage-treatment plant upgrades on stream water quality from existing data, our selection of cases was limited to plants located above the water-quality sampling locations that possessed long-term records, and our water-quality criteria were restricted to the parameters that had been sampled regularly at those locations.

We relied on the EPA-developed software to identify monitoring stations situated downstream from plants whose upgrades had been completed within a period of time that allowed for sufficient pre-upgrade and post-upgrade data. Specific treatment plants and other dischargers can be identified by NPDES permit number in the EPA computer programs, and other dischargers and water-quality monitoring locations whose readings are contained in STORET can be searched upstream and downstream.

From the fixed monitoring station data on STORET, we extracted the records of the water-quality constituents that had been regularly monitored and that were most likely to have been affected by changes in treatment plant effluent loadings. The constituents included dissolved oxygen, ammonia, nitrite, and nitrate and, in some places, BOD_5 and phosphorus.

Data Transformations

In addition to extracting the observations of stream concentrations of the different pollutant indicators, we extracted flow and water temperature readings from STORET, in order to adjust the raw water-quality data. We converted dissolved oxygen readings into a more meaningful form that was based on a stream's oxygen saturation. We also calculated a flow-adjusted form of each water-quality observation that we based on a model of the relationship between the concentration of the constituent and stream flow.

Dissolved Oxygen Deficit

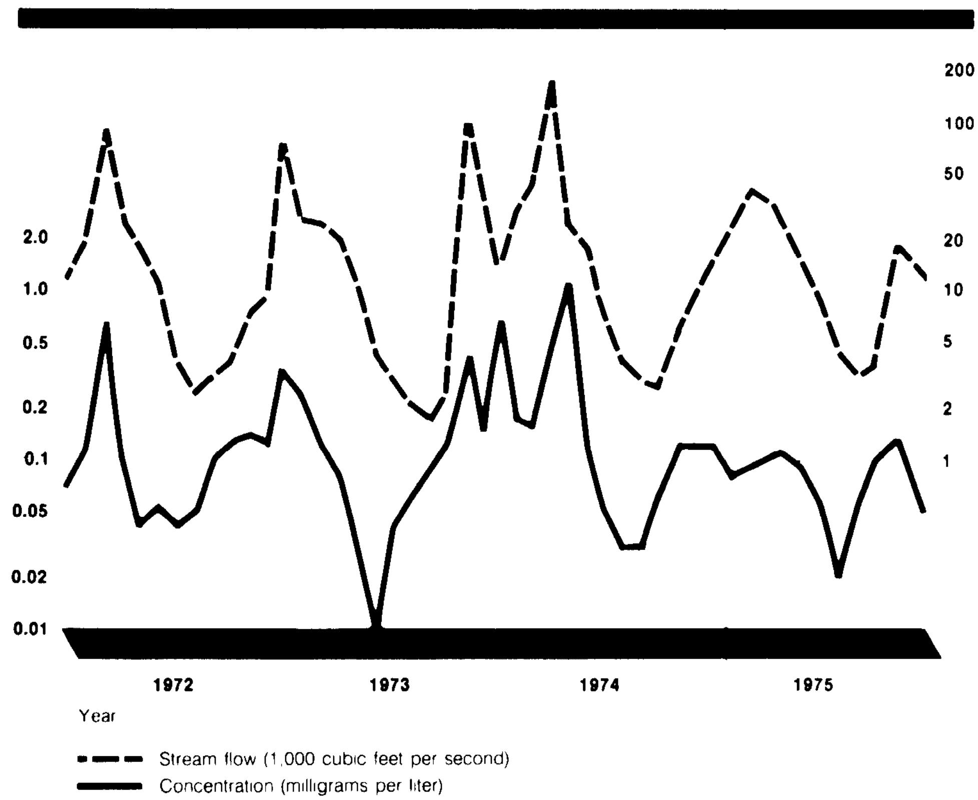
The concentration of dissolved oxygen in water is affected by the ability of water to retain oxygen in solution. The primary determinant of this ability is water temperature. For this reason, adjustments are commonly made to dissolved oxygen measurements in order to reflect the effect of temperature (and sometimes of other, less critical influences), and a statistic is derived that expresses the extent to which the observed level of dissolved oxygen represents the maximum concentration of oxygen in the water, given the water temperature. This procedure can be readily performed on STORET data, since nearly all STORET observations record water temperature.

We applied a common algorithm to calculate dissolved oxygen (DO) saturation potential from water temperature. For this report, we converted DO saturation to DO deficit: the difference between the actual reading and saturation. Expressed mathematically, $D = 1 - (O/S_t)$, where D = DO deficit, or O = observed DO level and S_t = DO saturation level at a given temperature. Therefore, when we discuss dissolved oxygen, we are discussing the unsaturated condition of the water body in question. The larger the deficit, the poorer the water quality.

Flow-Adjusted Concentrations

The effect of a given quantity of a pollutant on a body of water is obviously a function of the amount of water containing the pollutant. A kilogram of BOD_5 will have a stronger effect on a small creek than on a large river. This relationship poses two parallel problems for any analysis that attempts to relate a time series of data on water quality in a stream with a treatment plant's discharge history: (1) concentrations of pollutants in a stream measured under different flow conditions are not directly comparable and (2) the effect of pollutant discharge into a stream may not be clearly discernible except when flow is low (typically, during the summer).

Figure 2.1: Stream Flow and Total Phosphorus Concentration in the Klamath River, California, in 1972-75

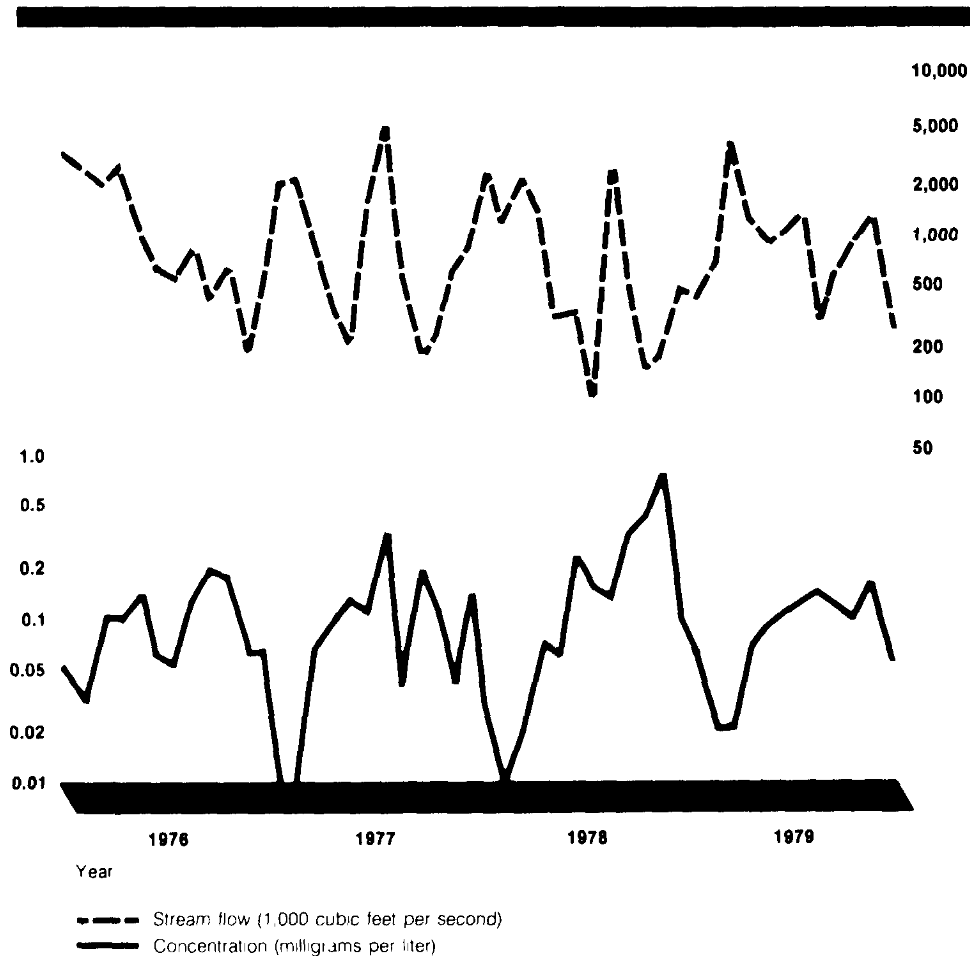


Source: R.A. Smith, R.M. Hirsch, and J.R. Slack, A Study of Trends in Total Phosphorus Measurements at NASQAN Stations, Water-supply paper 2190 (Reston, Va.: U.S. Geological Survey, 1982), p. 7.

These phenomena tend to distort or conceal measures of water quality, so we took two different steps to adjust for them. First, we standardized water-quality data for flow. Second, we examined the relationship between effluent and stream quality within different flow strata. The second step did not involve data transformation but did involve examining selected subsets of the water-quality data. We discuss the second step below, in the section entitled "flow stratification."

A simple way to standardize concentrations of pollutants recorded at different flow levels would be to multiply them by their respective flows

Figure 2.2: Stream Flow and Total Phosphorus Concentration in the Black River, South Carolina, in 1976-79



Source: R.A. Smith, R.M. Hirsch, and J.R. Slack, A Study of Trends in Total Phosphorus Measurements at NASQAN Stations, Water-supply paper 2190 (Reston, Va.: U.S. Geological Survey, 1982), p. 7.

to determine pollutant loads. Research performed at the U.S. Geological Survey, however, has demonstrated that the relationship between flow and concentration is frequently not linear and can take several different forms. Examples are shown on pages 24 and 25 in figures 2.1 and 2.2, which plot phosphorus concentrations over time along with flow measures in two rivers and demonstrate the different forms that the relationship between flow and the concentration of a common chemical pollutant can assume.

The concentration of phosphorus in the Klamath River, as depicted in figure 2.1, varies in nearly exact proportion to river flow. This strong, direct correlation suggests that the presence of phosphorus in the river stems mostly from nonpoint sources such as agricultural runoff. As rainfall increases, river flow increases, and so does the amount of phosphorus washed into the river or stirred up from the river banks and bottom.

A different phenomenon occurs in the Black River, as depicted in figure 2.2. Here, phosphorus concentration is almost a mirror image of river flow. This suggests a relatively constant source of phosphorus that is not affected by precipitation. As the river flow increases, the stream phosphorus is diluted.

The U.S. Geological Survey has developed procedures that fit 14 different models to the relationship between flow and concentration. These procedures are purely pragmatic. They attempt not to explain the relationship between flow and concentration at a monitoring site but only to describe it and statistically remove as much of the flow-associated variation from concentration readings as possible. (More details on these models will be found in the second volume of this report.)

Using these procedures with minor modifications, we derived a set of flow-adjusted concentrations to parallel the STORET readings. The flow-adjusted concentrations are the residuals from the regression model that best fits the data—that is, the differences between the readings predicted by the model and the observed values. They represent the stream concentrations purged of the detectible effects of variations in flow.

Analytic Methods

The methods we used to detect changes in water quality downstream from upgraded sewage-treatment plants included selecting specific subsets of monitoring station data for separate analyses (flow stratification), comparing observations with baseline conditions (data differencing), and determining with statistical techniques whether a change in water quality downstream from a plant was coincident with its upgrade.

Flow Stratification

To respond to the problem that the effect of pollutants is more discernible (and more environmentally significant) during periods of low flow, we stratified water-quality observations into flow strata. We developed a "low flow" data base by examining only observations made at flow levels within the bottom quartile of flow readings. We found that low-flow measurements of pollutants typically averaged higher than their overall means. For example, the dissolved oxygen deficit downstream from the Tamaqua plant averaged 8 percent before the upgrade; the low-flow average was 19 percent for the same period.

We experimented with an alternative low-flow stratum consisting of observations in the lower half of flow readings. We abandoned this alternative, since it added little explanatory power to that already offered by examining the lowest quartile. The results of the procedure are presented in our discussion of the Allentown case study.

We call the data base for all observations, whatever the flow level, "full flow." Because the relationship of flow to concentration can be expected to be different at lower flow levels, we developed separate flow adjustments for full-flow and low-flow data sets.

Another common procedure is to examine observations made during summer months, which are usually the periods of lowest flow. This has the advantage of narrowing the range of variations in temperature, another important influence on the assimilative capacity of rivers. However, we chose to employ explicit flow stratification for its more precise control of stream volume, the most critical determinant of a pollutant's concentration. (We discussed our adjustments for the effect of temperature variations on stream dissolved oxygen in the section above on dissolved oxygen deficit.)

Data Differencing

Attempts to relate stream water quality to individual dischargers can be confounded by the influence of other sources of degradation or improvement in water quality. These influences may be either point sources, such as industrial or municipal discharges, or nonpoint sources, such as road salt or agricultural runoff containing fertilizers and pesticides. While it is probably impossible to control for all sources, it is possible to exert partial control, which increases the likelihood of detecting an effect specific to a particular discharger.

Where adequate data are available from monitoring stations both above and below a treatment plant, only pollution sources that lie between the stations need be considered for alternative explanations of changes in water quality. The background influence of other pollution upstream can be eliminated by subtracting the readings at the upstream station from those at the downstream station. We call the results "differenced data." A positive result from differencing indicates a concentration of the pollutant in question that is greater downstream than upstream.

In a perfect differencing situation, an upstream station would be located immediately upstream from a treatment plant and the other station would be a relatively short distance downstream. This situation would minimize the length of the uncontrolled stretch of river and the need to collect additional data from point sources. Unfortunately, our experience suggests that it is very rare.

In our Allentown case study, we were able to use differenced data because of the availability of upstream water-quality observations. We were able to apply data from three monitoring stations, one above and two below the plant. In this case, we used two sets of differenced data, one for each pair of upstream and downstream monitoring stations.

We calculated mean levels of dissolved oxygen deficit, BOD_5 (where available), ammonia, nitrite, nitrate, and fecal coliform bacteria from the STORET data set associated with our target plants for the same period for which we calculated each plant's effluent means. We repeated these calculations for differenced data, where they were available, and for the low-flow condition described above. We performed the same calculations on the corresponding flow-adjusted data. We tested the statistical significance of the change from the before-upgrade to the after-upgrade periods by computing the t statistic.

Were Changes in Effluent Related to Changes in Stream Water-Quality Indicators?

Data Sources: Combined Data

To answer this question, we used a combination of the data bases we used for the two preceding evaluation questions: discharge monitoring reports and stream water-quality observations. These two data bases are not completely parallel. Both consist of a time series of monthly estimates of pollutant levels, but the discharge data represent the amount of specific pollutants discharged into a river while the water-quality observations represent the stream concentration of the same or related pollutants.

Their reliability also differs. The discharge data report the daily average of effluent loadings estimated from weekly, daily, or more frequent observations. Most of the stream readings represent a single observation and are clearly less reliable indicators of a month's water quality than an average of more frequent observations would be.

Analytic Methods

We determined the degree of association between these two time series by calculating the Pearson product moment correlation coefficient for the two sets of monthly readings. We compared the water-quality readings for every month with the effluent averages reported in those months on the discharge monitoring reports. We correlated the latter data with the full set of stream observations and with the low-flow subsets discussed above, with unadjusted and flow-adjusted observations, and with differenced data, where appropriate.

For practical and theoretical reasons, we did not limit our correlation analysis simply to one-to-one relationships (for example, the relationship between effluent and stream BOD_5 or between effluent and stream ammonia). Practically, gaps in both sets of data limited our ability to perform such analysis. For example, BOD_5 readings were available, because required for all discharge monitoring reports, but stream BOD_5 readings were much less common. Stream ammonia concentrations, in contrast, had been routinely recorded each month for the STORET data bases but had not been required in the monitoring reports for two cases and had been recorded only after the upgrades for the two other cases.

More importantly, there are substantive reasons for investigating the association of one effluent constituent with multiple measures of stream pollution (and vice versa). These have to do with the dynamics of a stream's assimilation of pollutants. For example, the ammonia discharged from a treatment plant is oxidized through bacterial action in the stream first into nitrite and then into nitrate. This nitrification process depletes the stream oxygen. Thus, the ammonia discharged into a stream affects the level of not merely the stream ammonia but the nitrite, nitrate, dissolved oxygen, and BOD_5 as well. The relationships vary in space and time. Effluent ammonia is likely to be highly correlated with stream observations of ammonia made close to the discharge point but may be more highly correlated with nitrate than with ammonia concentration further downstream.

Can Alternative Explanations Be Excluded?

The controls on pollution sources other than a treatment plant that are imposed by the differencing procedures and flow adjustments can be viewed as the first step toward ruling out the possibility that what appears to be an improvement in water quality caused by an upgrade is actually caused by a decrease in pollutant discharge from other sources. Differencing excludes the influence of the point and nonpoint sources above the upstream monitoring station, and flow adjustment diminishes the influence of flow-related sources, particularly nonpoint sources. It is possible, however, to use existing data to assess more directly the likelihood that observed water-quality changes stem from sources other than the treatment plant in question.

Data Sources

The software package that we used to identify monitoring stations in the vicinity of upgraded sewage-treatment plants can also be used to identify other point sources discharging into the rivers. We collected discharge monitoring data for the major point sources located above the criterion monitoring station whose influence could not be excluded by differencing. These rival point sources included industrial and municipal dischargers. We extracted from the discharge monitoring reports the same information (where available) that we extracted for the target plants: flow, BOD_5, suspended solids, fecal coliform bacteria, and ammonia.

The number of additional point sources we considered varied from case to case. For example, at Tamaqua, only one other point source, another municipal sewage-treatment plant, was included in the analysis. At Allentown, despite the application of differencing techniques, we had to examine five other point sources that influence downstream water quality.

Analytic Method

The complexity of our analyses depended on the complexity of each situation. For all cases, we tested for changes in effluent with the procedure we used for treatment plant effluent: we calculated the mean values before and after an upgrade and computed t tests to determine statistically significant differences between means. In some instances, it was evident that other dischargers could not satisfactorily account for stream changes, because their discharges had remained at a constant level through the period of analysis or were so small in comparison to the plant's discharge as to be trivial additions to the stream's pollution.

In other instances, where rival point sources had approximately the same magnitude as the upgraded plant and a discharge history that paralleled the plant's, we applied multiple regression analysis to assist in estimating each source's independent effect on downstream water quality. We used the effluent BOD_5 and suspended solids loading from the plant to predict water-quality readings. While these may not reflect the most parsimonious models for the relationship between effluent and stream water quality, they make it possible to distinguish the relative importance of different pollution sources.

In chapter 3, we draw together the findings and the lessons we learned from our case studies, present our conclusions on the success of each aspect of our methodology, and estimate the feasibility of extending the methodology to a much larger sample. We have included the detailed results of the case studies in the second volume of this report.

Summary, Conclusions, Recommendations, and Agency Comments and Our Response

The objective of our evaluation was to determine the feasibility of basing an assessment of the effect of Construction Grants Program upgrades on currently available data. We developed a method for evaluating the effect of upgrading sewage-treatment plants on the quality of the body of water into which a treatment plant discharges its effluents. In chapter 2, we introduced several concepts and procedures for analyzing the types of data that were available to us for measuring the effect of upgrading a treatment plant on downstream water quality. In this chapter, we summarize our experience with these procedures and discuss their relative success and failure when applied to four different upgrades. Before proceeding to these methodological considerations, however, we summarize the principal substantive findings that emerged from our application of the method to four case studies. These findings are case specific; that is, they cannot be generalized to other upgrades. We present them as examples of the results that can be obtained by means of these evaluation techniques when relevant data are available.

Summary of Findings From Case Studies

We examined the effect of treatment plant upgrades in four eastern Pennsylvania municipalities: Allentown, Hamburg, Lansdale, and Tamaqua. We present the details of these case studies in the second volume of this report. Here, we outline the principal findings, organizing them around our four evaluation questions.

Did the Upgrade Decrease Discharged Pollutants?

At all four sites, the average amount of BOD_5, suspended solids, and fecal coliform bacteria discharged after an upgrade was substantially lower than before the upgrade. The amounts decreased despite significant increases in the volume of wastewater discharged from all the plants except Tamaqua. At Hamburg, the initial decline in pollutant discharge was not consistently maintained after the upgrade but surged, approaching some pre-upgrade levels for an extended period that began approximately 3 years after the upgrade.

Did Downstream Water Quality Improve?

Water quality in the receiving streams below the Allentown and Tamaqua treatment plants improved significantly after their upgrades. Marginally significant improvements were also found at the monitoring station downstream from the Hamburg plant, the smallest plant we studied. The improvements in effluent quality resulting from the Lansdale upgrade were not reflected in an improvement in water quality downstream.

Were Changes in Effluent Related to Changes in Stream Water-Quality Indicators?

We found a statistically significant association between discharge from the treatment plants and downstream water quality in every case but Lansdale. The association was stronger at low flow, or when stream observations were adjusted for flow. At Lansdale, we found no statistically significant relationship between plant effluent and stream conditions.

Can Alternative Explanations Be Excluded?

The streams to which the Allentown and Lansdale plants discharge effluents receive effluents from several other point sources, both municipal and industrial. Comparing the effluent history of the major competing point sources, we determined that the influence of the Allentown treatment plant on water quality at two locations downstream was significant and distinguishable from that of other dischargers that contributed significantly to the stream's water quality. The failure of the Lansdale upgrade to improve water quality at the downstream monitoring station stemmed, at least in part, from increased levels of pollutant discharge from two other nearby municipal waste-treatment plants during the postupgrade period. Tamaqua and Hamburg, in contrast, were the principal point sources affecting water quality in their receiving streams; the improvements noted in these streams were not the result of a reduction in point-source pollution from other sources.

These case-specific findings resulted from applying a set of analytic techniques to extant data relevant to four upgrades. In the next sections, we discuss the techniques we used to analyze the data and report methodological conclusions that emerged from these case studies. We hope they can serve as guidelines in designing future evaluations of Construction Grants Program upgrades. Below, we address the fundamental issue of data availability and reliability.

The Data: Availability and Reliability

Effluent Data

The effluent data in the discharge monitoring reports were generally available and sufficiently reliable for our case studies. Copies of reports were available for holders of NPDES permits at the plants, state field offices, and EPA's regional office, although their comprehensiveness varied from case to case. Permit-holders are required to keep records for only 3 years, but in several cases we were able to obtain much older reports from them. The state field office that was responsible for monitoring three of our cases was able to provide discharge monitoring reports for approximately the previous 3-year period. EPA's regional office was the richest source of information; in most cases, its records dated back to 1978 in hard copy and to the original permit dates, usually 1974 or 1975, on microfiche.

These data are being computerized by EPA. Region III has computerized records from 1982 onward. We used the hard copy of the records EPA provided. Obtaining direct access did not seem an appropriate allocation of our resources, because we required effluent records from 1976 or even earlier for many of the permit-holders. However, studies examining the effect of more recent construction could use the computerized data base exclusively.

We did not attempt to verify the accuracy of these data directly. The data were clearly subject to error from various sources such as poorly calibrated instrumentation, inadequate laboratory procedures, and clerical mistakes. At Tamaqua, for example, we discovered that flow measurements were inaccurate, to an undefinable degree, because of a malfunctioning flow meter.

However, we had limited the scope of our evaluation to existing data, in order to discover whether the various types of data employed in this study could be integrated in a way that would yield reasonable explanations of the stream effect of treatment plant changes. We believe we were successful in this undertaking. At Tamaqua, despite the acknowledged inaccuracies in flow—and, therefore, in pollutant load estimates—it was possible to demonstrate a relationship between effluent patterns and downstream water quality. We assume that this relationship would have been even clearer with more accurate effluent data.

Grant Information Control System

Our use of the GICS data was fairly limited. We obtained our original list of the universe of completed Public Law 92-500 projects in EPA's region III from the GICS. We also used the GICS dates for the start of construction, works in operation, and project completion as preliminary screens to identify projects whose timing appeared to allow for adequate before-and-after records.

Our experience with the GICS suggested that it was generally adequate for this limited purpose. However, where "works in operation" dates were missing, were identical with project completion dates, or indicated an unlikely length of time (that is, more than 2 years between the start of construction and the start of operations), the GICS information should be supplemented with information taken either from direct contact with the treatment plant or from the discharge monitoring reports. In any event, the best date for the effectiveness of a plant's upgrade is the date of a clear change in the time series of plant effluent records. For two of our cases, this date was within a month of the GICS date; for a third case, the GICS provided no date; for the fourth case, the date was 4 years before the GICS date.

We demonstrated another approach to the timing question in our Allentown case study. Here, we excluded a 2-year period of records as being typical of neither the pre-upgrade period nor a full upgrade. This approach seems more appropriate for large-scale upgrades, in which new processes are phased in as they are completed. In the Allentown analysis, our estimate of the upgrade's effect may have been enhanced somewhat, but our overall findings were not substantially different from a parallel analysis that included data from the 2-year transition.

Water-Quality Data: STORET

In reviewing possible locations and data sources for our case studies, we examined STORET and other municipal and private data sources. None of the non-STORET data bases were adequate, because their records were incomplete. All the stream data we used for our case studies were contained in STORET and had been derived from long-term records collected at fixed monitoring stations.

With the exception of some stream flow data, all our stream data were collected at stations maintained by the Pennsylvania Department of Environmental Resources. These stations generally contained continuous monthly records beginning in the mid-1970's and had earlier quarterly data. It had not been our intention to limit our water-quality data sources to one agency and one data bank, but it turned out that the only adequate data sources were these monitoring stations.

The use of the water-quality data in STORET has the distinct advantage of easy availability and use. The STORET system can be accessed through remote computer terminals. The software allows an immediate interface between STORET data bases and SAS, a common statistical software package.

STORET's greatest virtue is its vast wealth of water-quality data. Until recently, this was its major weakness. In 1976, the number of sampling locations contained in STORET had reached 200,000. In 1979, EPA estimated that STORET contained 57 million observations. Searching a data base as massive as this for stations that satisfy the requirements of timing, detail, and location needed in order to identify long-term trends in water quality affected by specific changes in point sources could be inordinately time-consuming and expensive. Moreover, most of the records in STORET are too limited for establishing long-term water-quality trends.

In the 1970's, STORET developed software that allowed the selection and retrieval of data by state, county, hydrologic unit, and user-defined "windows" of latitude and longitude. The software was helpful but limited in utility. The development of the "river reach file" in 1981 and software linking this river indexing system with water-quality monitoring locations, the U.S. Geological Survey's "flow-gaging" stations, and dischargers under the NPDES program have greatly facilitated the search process. Interactive computer programs that are now available allow the researcher to specify a particular point source by NPDES number and receive immediately a list of all other point sources as well as water-quality monitoring stations, gauging stations, and public water supplies located on the river reach and a graphic representation of their location. The user may then search up or down the stream for the same information on adjacent reaches. Another software package allows the investigation of this information in varying levels of detail. The interactive software makes a large-scale attempt to match point sources of pollution and appropriate downstream data practical, perhaps for the first time, although it has deficiencies that are to be expected in a system still under development.

We encountered some difficulties in using this search sequence. First, the programs are not accompanied by "user-friendly" documentation, presumably because their primary, if not exclusive, users are the EPA staff involved in their development. Second, their identification of point sources and monitoring stations is not infallible. We found that one sewage-treatment plant had been mislocated by 150 river miles. The Tamaqua monitoring station we relied on for defining downstream conditions was not identifiable through the search program; we found it only through a conversation with a staff member in the Pennsylvania Department of Environmental Resources who was familiar with the area. Another monitoring station had been moved several miles downstream from the original location recorded in STORET, without any record of the move, which had occurred at approximately the same time as an upgrade to an upstream treatment plant. Without an external verification of the monitoring station's location, our trend analysis of the data might well have resulted in our attributing apparent water-quality improvements to the upgrade that were in reality the effects of moving the observation point further downstream from the plant. A similar undocumented move may have corrupted a data base upstream from the Allentown plant.

The location of point sources and monitoring stations relative to one another was not always evident from the STORET programs, particularly for the complex networks of small streams in our Lansdale case study. At Lansdale, three monitoring stations, each with extensive, long-term records, appeared to be appropriate for estimating the upgrade's effect, but after consulting with local officials, we discovered that only one station was in a position that allowed us to anticipate useful results from analyzing its data. The second was located on a small tributary, and the third appeared to be too far downstream to detect upgrade-related changes.

For Allentown and some other cases, a river mile index system had been applied to river reaches. For a small number of identified instances, the system provides exact mileage from the downstream end of a river reach for each point source and monitoring location. For the data that are available, this greatly facilitates the exact identification of points of interest in the reach.

In previous reports, we have pointed out the difficulties of using data from fixed monitoring stations to characterize stream water quality.[1] To a large extent, our criticisms centered on the question of how representative such data are on various levels. First, we expressed doubt that the collective data from the national networks can be considered truly representative of the nation's water quality. It is unlikely, for example, that the quality of the 1.8 million miles of rivers and streams in the United States can be adequately defined by some 300 NASQAN stations. Second, we pointed out that characterizing the conditions in a river from samples taken at only one location in the river can be misleading. For example, NASQAN's only source of water-quality data for more than 200 miles of the South Platte River is one station. Water-quality conditions

[1]See, for example, Water Quality Management Planning Is Not Comprehensive and May Not Be Effective for Many Years, CED-78-167, December 11, 1978, and Better Monitoring Techniques Are Needed to Assess the Quality of Rivers and Streams, CED-81-30, April 30, 1981.

are so subject to spatial variations that a sample taken at one location may bear little relationship to a sample taken a few miles upstream or downstream in the same river.

In designing this study, we took into consideration our earlier reservations about the inadequacy of fixed monitoring station data for providing representative samples. We chose a case study approach because we recognized that a nonstatistical sample of sewage-treatment plant upgrades matched with downstream water-quality data could not be confidently considered representative of upgrades under the Construction Grants Program across the nation. We recognized also that the readings at the locations we used for our case studies may not be typical. We attempted to detect the effects of changes upstream in relatively close proximity to the changes in point-source pollution resulting from sewage-treatment plant upgrades in the belief that if no effect could be discerned at our monitoring locations, then it was highly unlikely that samples taken further downstream would show changes attributable to the upgrades.

In the reports cited above, we have also questioned how representative fixed monitoring station data are of conditions at the stations. Readings taken once a month are subject to large variations because of chance and typical conditions such as sudden rainstorms and brief surges in discharge from a point source. Every measurement of a concentration is strongly affected by stream flow. Dissolved oxygen levels and oxidation vary from day to day, and even within one day, depending on cloud cover, temperature, flow, and other factors.

We attempted to minimize the effects of external factors. We developed flow-adjustment models for each location and each constituent at each flow stratum. We converted dissolved oxygen readings into dissolved oxygen saturation estimates that took temperature variations into account. Where possible, we also controlled for temporal variation by using upstream stations as baselines. (Monthly observations at adjacent stations maintained by the same agency are typically taken within a few hours of one another, which reduces differences from short-term variations.)

Even with these error-reduction strategies, some error remained in the water-quality data, and this difficulty was compounded when these data were correlated with discharge monitoring reports of monthly effluent averages. We have no reason to believe that the error is nonrandom, but this "noise" in the data will make it difficult to demonstrate the hypothesized associations between effluent levels and downstream water quality. Our findings, in most instances, are consonant with this assertion, since significant correlations tended to be few and moderate, but they did exist.

Other error-reduction procedures were possible, some automatically implemented by or available from STORET. STORET data on the water-quality parameters we examined are routinely screened for such impossible or highly implausible readings as water temperature less than −2 degrees Celsius, pH greater than 12, and negative concentrations. In addition, STORET flags some data that should be qualified. A common example is a reading that represents the threshold of detection for a particular laboratory procedure or device, so that the true value of a concentration may be not zero but some value between zero and the detection threshold. We excluded all such data from our analysis.

We could have taken other steps to further reduce variance. For example, we could have excluded all values more than two or three standard deviations from the site mean for each constituent. We chose not to take such steps in order to perform a more realistic test of our essential hypothesis—namely, that the effect of the Construction Grants Program on water quality can be determined from available data.

Analytic Procedures

In this section, we summarize how we used the four case studies to construct the methodology we present in the report.

Flow Adjustment

Our experience with the U.S. Geological Survey's flow-adjustment procedures was mixed. We attempted to develop flow-adjustment models for each of the five parameters (at Lansdale we added a sixth, phosphorus) that we examined at each of six monitoring stations. We tested the 14 suggested models for each combination of location and parameter at different flow strata. We used the criterion of statistical significance ($p < 1$) that the Geological Survey used.

We were able to develop statistically significant flow-adjustment models for 49 of the 77 final combinations (64 percent) of parameter, location, and stratum that we used in our analysis.[2] As would be expected, statistically significant flow-adjustment models were least likely to be found where the smallest number of observations were available—that is, when we analyzed only low-flow observations. Eighty-four percent of the modeling attempts were successful for observations made under all flow conditions, but only 48 percent were successful for the lowest-quartile observations.

The flow data we used for adjusting concentrations were collected at the same location and time as the other observations of criterion parameters for four of the six monitoring stations we used in our evaluation. Because flow data were insufficient or totally lacking at the two other monitoring stations, we applied flow data collected at flow-gauging stations some miles upstream or downstream from the water-quality observations made at these two stations. Our success at finding significant flow-adjustment models at these stations was similar to our success with flow data collected at the criterion monitoring stations.

The relative effectiveness of adjusting stream data for flow was demonstrated in our analysis. Where flow adjustment was possible, it generally tended to sharpen differences between before-and-after stream data sets and to increase the positive correlation between effluent and stream data.

Despite the apparent utility of flow-adjusted data, we do not believe that they should be the sole measure of stream changes. On many occasions, adequate flow-adjustment models cannot be developed from the 14 models suggested by the Geological Survey and unadjusted data must be used. More importantly, the procedure can be considered a "black box." Basing the selection of a model solely on a single statistical criterion is intuitively unsatisfactory and can end in misleading results. For example, it is probable that what appears to be the influence of stream flow on stream concentrations is in many cases the effect of an exogenous third variable correlated with both flow and concentration. Of course, this is also the procedure's strength. It is the only readily available way of controlling, although indirectly and imperfectly, for the

[2] A statistically significant model is a least-squares regression equation, in which a function of flow accounts for a portion of the variation in stream concentration of a given parameter at confidence level greater than 9 in 10 (p. < 10). We borrowed this criterion, and the set of models, from the Geological Survey. The Geological Survey has reported a 71-percent success rate in adjusting phosphorus data for flow from 289 NASQAN stations, but its results are not directly comparable with ours for three reasons: (1) the Geological Survey did not attempt to develop separate models for low-flow subsets of observations; (2) it used only 11 of the 14 models, subsequently adding 3 more; and (3) the data sets it used generally cover shorter periods than ours.

effects of nonpoint-source pollution on stream water quality without substantially increasing the costs of data collection and analysis.

For these reasons, all the results we report in the second volume are expressed in both unadjusted and flow-adjusted data. The decision of whether to add flow adjustment, with its attendant complexities, to future water-quality analyses must be made individually from considerations of time and expense. On balance, the advantages of the procedure appear to outweigh its small costs and risks of misinterpretation.

Data Differencing

The differencing procedure has been suggested by water-quality researchers as a straightforward method of providing an upstream baseline against which to judge water-quality changes.[3] We were able to apply the procedure to the Allentown case study by identifying monitoring stations. The Lansdale case may be considered "source differenced," since the sewage-treatment plant is near enough to the headwaters of its receiving stream to obviate the need for upstream baseline data.

In the Allentown case, using differenced data shed some light on the analysis but also introduced complications. Some water-quality parameters were significantly degraded at the baseline location in the period being examined. During the same period, downstream conditions improved. Therefore, an analysis of merely differenced data over time would have overstated the downstream effect of the Allentown upgrade.

Given this experience, we concluded that data differencing must be used with caution and that it is not appropriate where baseline conditions undergo significant changes over time. The difficulty of applying differencing to sewage-treatment plant upgrade evaluations lies in the nonconservative nature of the water-quality parameters affected by upgrades—dissolved oxygen, BOD_5, ammonia, nitrite, and nitrate, among others. These stream criteria are transformed by natural processes when they are carried downstream, so that any one-to-one comparison between downstream and upstream trends can be misleading. For example, an upstream increase in ammonia concentrations does not imply a comparable rise in downstream levels, but it may be

[3]See Frankel, 1979, p. 159. Some form of differencing was also implicit in the design of the national water quality surveillance system, a national network of paired monitoring stations that EPA abandoned in 1981.

reflected in increased nitrite or nitrate concentrations downstream. Without detailed hydrologic modeling of the expected effect of upstream changes on downstream observations, it could be impossible to separate the effect of natural changes from the effect of other interventions, such as upgrades, into the dynamics of the stream.

In summary, where upstream data exist, they should be used in the interpretation of downstream conditions, but they should be used cautiously. Caution becomes increasingly important as the distance between upstream and downstream monitoring stations increases. It is unlikely that analysts will find a perfect differencing situation: an upgraded sewage-treatment plant located immediately below one monitoring station with a second station a few miles downstream. For this reason, formal differencing will only rarely have practical utility. However, other identifiable upstream influences on water quality should be considered in evaluations by including relevant discharge data from upstream point sources in the analysis. (We discuss this further at the end of this section.)

Flow Stratification

For Allentown, we were able to analyze three overlapping stream data sets: the set of all readings (excluding STORET's flagged data), the set of readings coincident with the lower half of the flow readings, and the set of observations made at the lowest 25 percent of flow conditions. We concluded (as we discussed in chapter 2) that an analysis of a lower-half stratum would not contribute significantly to the findings, and we changed our stratification to two strata: all readings and low-flow, or the lowest 25 percent of, observations. We continued to develop flow-adjustment models for these two strata.

In nearly all our analyses, an examination of the low-flow observations resulted in clearer evidence of change and stronger correlations with point-source discharge. However, with the large reduction in the number of available data points (since, by definition, a low-flow analysis uses only one quarter of all readings), the confidence levels of our findings were also reduced. For this reason, unless a very large number of observations is available (perhaps 200 or more), it appears prudent to use both strata in an analysis.

In our preliminary analyses, we used summer readings as surrogates for low-flow conditions. However, we found that low flow was not limited to the summer months for our case studies. Therefore, we believe that explicit flow stratification is a more effective tool, particularly in view of the data sorting and selection capabilities that are available for computer-assisted analyses. (The use of only summer observations has the distinct advantage of limiting the range of variations from temperature and daylight, but we believe that flow is a more important determinant of stream concentrations.)

The Relationship Between Effluent and Stream Water Quality

We believe it is unwarranted to assert that an improvement in stream water quality that occurs at approximately the same time as an upstream upgrade has been caused by the upgrade. For this reason, we tested the correlation of discharge monitoring reports and stream water quality over time to see if the changes in a treatment plant's effluent were reflected in the readings at our downstream monitoring locations. They were for three of our four cases. We recognize that correlation does not imply causality, but we believe that a significant positive correlation between the discharge monitoring reports and downstream readings makes a causal leap somewhat less risky. At the least, the absence of a demonstrable relationship would make causal attribution highly suspect.

The strength of these correlations almost always understates reality, because of the nature of the two data sets being compared. The value of a correlation coefficient is commonly interpreted as the square root of the amount of common variance between two sets of observations. In the correlation between the discharge monitoring reports and stream data, this estimate should be considered close to the lower boundary of the true relationship.

For example, we found that 23 percent of the variance of downstream ammonia readings at low flow in the Little Schuylkill River could be explained by the level of BOD_5 in the effluent from the Tamaqua treatment plant. This does not mean that 77 percent of stream ammonia is caused by point or nonpoint pollution from sources other than the plant. Some portion of the unexplained 77 percent is presumably a function of individual stream readings' reflecting conditions not completely representative of the months for which they were surrogates. The influence of the treatment plant's effluent on water quality was most likely stronger than what is indicated by the correlation coefficient.

Alternative Explanations of Water-Quality Trends

A causal link between treatment plants and water quality is not conclusively demonstrated in any particular instance by a significant correlation between discharge monitoring reports and stream readings. Since a complete analysis must consider the possibility that changes in other point sources can better explain changes in a river, we examined the reports of competing point sources in our case studies. It was possible to exclude the possibility of the influence of some of them, simply because of their small amounts of effluent or their relatively long distance from a monitoring location. We excluded minor dischargers (for example, a trailer park near Lansdale) and dischargers more than one river reach above the discharge point.

We also examined records for changes that paralleled those at a sewage-treatment plant and might as easily explain water-quality changes. At Allentown, for example, we discovered that another plant had been upgraded at approximately the same time as the plant we studied, requiring a more detailed analysis of their combined effect on downstream water quality. In the Lansdale case, we demonstrated that pollutant loading from other sources might be totaled and contrasted in influence with that of the Lansdale plant. For the major competing point sources, we employed multiple regression techniques in an attempt to identify the influence of different point sources relative to one another.

Summary and Comment

In our four case studies of sewage-treatment plant upgrades, we have demonstrated methods of identifying and accessing available data, transforming them for analysis, dating and quantifying an upgrade, and relating the upgrade to water quality and other influences on the health of a stream. We have not demonstrated all possible approaches to analyzing the available data on the effect of the Construction Grants Program, and our approach is not necessarily the best use of the data.

For example, some analysts might prefer to apply classic Box-Jenkins time-series analysis to these data. We considered doing this, but we concluded that the demands of this analysis would be too rigorous for the available time series, which were irregularly spaced and sometimes rather short and contained extensive gaps. Nevertheless, techniques that could impute missing values and other approaches to improving the quality of the data could make these time series amenable to Box-Jenkins analysis.

Other analysts might argue that given the nonnormal distribution of many water-quality variables, only nonparametric techniques would be appropriate. In deference to this position, we applied nonparametric tests for trends to the Allentown water-quality data. The conclusions were essentially the same as those of our more traditional approach.

This confirmed our belief that the classical approaches were robust and accurate enough to suit our needs and the idiosyncratic nature of the data bases we examined.

We believe that the data that have been and are being collected can be used in assessing the effect of the Construction Grants Program and other interventions on the water quality they are intended to improve. Given current limits on evaluation funds, we believe that rather than abandoning an evaluation effort or setting in motion impractical or duplicative data collection efforts, it makes sense to use the data that are already available.

The Feasibility of Expanding the Use of Our Method

Whether the method we applied to the four case studies can be used on a larger scale depends on two essential considerations: the generalizability of the approach and its cost.

Generalizability

We can make only a gross estimate of the extent to which our method would be applicable to more cases than four in eastern Pennsylvania. While all four were in close geographic proximity, they varied between urban and rural settings, large and small treatment plants, and upgrades from secondary to advanced treatment. A critical consideration is the extent to which it is reasonable to anticipate detecting an upgrade's effect at a downstream monitoring station.

The factors that would make this reasonable include the amount of wasteflow from the plant and its constituents, the size of the receiving stream, the decrease in pollutant discharge after the upgrade, the distance between the discharge point and the monitoring station, the extent to which the effluent is mixed with the stream before reaching the monitoring station, and the stream's aeration and nitrification rates. Complex models could take these factors into account, but confidence in their predictions would require additional data. To screen out cases not amenable to simpler evaluation, rules of thumb on effluent amount, stream flow, and the distance between plants and monitoring stations might be useful and appear to be suggested by the four cases.

Three of the four sewage-treatment plants contributed substantially to stream flow, for example. Allentown's and Lansdale's effluent constituted 3.4 percent of the stream flow at the downstream observation point, and Tamaqua's effluent accounted for nearly 5 percent. During 1 month of the period we examined, nearly one quarter of the stream flow measured below Tamaqua was effluent from the plant. See table 3.1.

Table 3.1: Summary of the Relationship Between Wasteflow and Stream Flow at Downstream Monitoring Stations in Our Four Case Studies

Sewage-treatment plant	Distance (miles)	Mean wasteflow (mgd)	Mean stream flow (cfs)		Ratio of wasteflow to stream flow[a]		
			Full	Low[b]	Full	Low	Maximum
Allentown[c]							
WQN124	5.5	29.5	2,308	650	3.4%	7.5%	8.6%
WQN123	17.0	29.5	2,850	864	2.6	5.3	20.8
Hamburg	1.8	0.4	709	176	0.2	0.3	6.9
Lansdale	10.0	2.2	279	34	3.4	8.1	15.4
Tamaqua	2.5	1.2	77	16	4.8	11.7	24.2

[a]Converted to common units; based only on months for which both wastewater flow and stream flow data were available.

[b]Lowest 25 percent.

[c]WQN = water-quality monitoring station.

The Hamburg plant, in contrast, contributed only a trickle to the Little Schuylkill River. However, a small but statistically significant decrease in stream pollution after the Hamburg upgrade was related to its effluent. This may be because, at the observation point, the Little Schuylkill River is small and quite clean, the plant's effectiveness improved dramatically (cutting its BOD_5 discharge, for example, to one eighth of its pre-upgrade average), the monitoring station is quite close to the plant's discharge points, and no other significant point sources confound the relationship between the plant and monitoring station records.

In contrast to Hamburg, Lansdale wastewater accounts for more than 8 percent of stream flow during low flow. Nevertheless, the effect of the Lansdale upgrade, which resulted in a two-thirds reduction in BOD_5 effluent, could not be detected downstream, perhaps because of an offsetting increase in pollutant discharge from nearby plants after the Lansdale upgrade.

Therefore, rules of thumb for predicting an effect on a stream are not easily derived from our cases. More cases should be examined, and more sensitive screening criteria should be developed. We suggest that cases worth investigating are those where long-term data are available from a point less than 2 miles downstream from an upgrade or where monitoring station data are available within 10 miles of a plant that contributes more than 2 percent to the stream flow.

Another consideration is whether data similar to Pennsylvania's are available in other states. Our discussions with administrators of natural and environmental resources in several states indicated that they are, and so did our review of the STORET data for several states.

Finally, we believe that our approach could be applied to a large number of Construction Grants Program upgrades, even though the stream monitoring station system cannot provide a sample of water quality representative of all rivers in the United States. An increase in the number of case studies to which our approach is applied would lead toward greater confidence in the results as useful estimates of the national situation.

Case Analysis Costs

Computer and staff considerations can provide very gross estimates of the costs of conducting case studies. Extrapolating from our experience is of only limited help, because much of our effort was spent in developing our approaches and because of the differences in the complexity of different case studies. For example, the Allentown case study was much more complex than the Tamaqua case and, therefore, had higher computer costs and took longer.

The computer cost of the Lansdale data calculations was less than $20. This included applying the flow-adjustment models to the raw data, calculating and conducting statistical tests on flow-adjusted and stratified data, and conducting regression analyses to test for alternative causes of improved water quality. This cost does not include the preliminary computer costs of identifying and retrieving the relevant data bases and would vary considerably, depending on the computer facilities used for analysis. As for the cost of time, it took one analyst about 25 to 30 hours to provide the necessary computer instructions and analyze the output data for the Lansdale case study.

It could be argued that EPA could perform these tasks more quickly than we did. After only four case studies, our analysts are still relatively inexperienced at using the STORET data base and computer software routines. Since the data base is stored at EPA and the software is developed there, EPA analysts have more experience with them and, therefore, should be able to conduct the computer operations more quickly. The initial analysis might take longer but should proceed smoothly after a few cases.

Conclusions

Our findings illustrate the problems associated with achieving the goals of the Clean Water Act to "restore and maintain the chemical, physical, and biological integrity of the Nation's waters." However, we can make no reliable estimate from our sample of the extent to which the Construction Grants Program has improved water quality in the United States, nor can we assess the relative frequency with which the successes and failures we found in our sample would be found by examining all Construction Grants Program projects. Nevertheless, we believe that our methodology can be applied to a large number of additional cases.

Applying our methodology to more cases would help illuminate future funding decisions and clarify the competing demands for funds. We believe that a much greater use of EPA's various data bases and software packages would increase EPA's ability to assess the effect of Construction Grants Program expenditures on stream water quality, at least case by case. The cost of implementing this methodology appears to be minimal. Most of the software has been developed, the STORET data are readily available, and the discharge monitoring data are being computerized.

The limits to the applicability of our methodology have not been tested, but they could be defined by extending it to additional cases. Each additional case would bring EPA one step closer to an understanding of the effect that the Construction Grants Program has on the nation's water quality. If the method proves applicable to a sufficient number of cases, a national estimate of the program's effectiveness could be developed.

The benefits of fully integrating the components of our methodology could extend well beyond the accumulation of evidence on the effects of sewage-treatment plants on stream water quality. The ready accessibility of an integrated network of data could improve the combined effectiveness of the various separate but complementary offices within EPA, each with its own responsibility for helping achieve the goals of the Clean Water Act. Moreover, as the base of case-specific evaluations increases, commonalities among the cases might emerge and contribute to the development of a decisionmaking framework for future funding priorities.

Recommendations

It is likely that billions of dollars will be spent on the Construction Grants Program in the next 5 years. We recommend that EPA perform additional evaluations of treatment plant upgrades that use available data and methods similar to the ones we developed. These evaluations should be intended to determine the feasibility of performing a broadly based and methodologically sound evaluation of the Construction Grants Program that makes optimal use of the data already in EPA's possession and that identifies and remedies the gaps in its information systems. Further, we recommend that EPA improve the reliability and usability of its water-quality data base by ensuring the internal consistency of its data collection practices, updating its data on the geographical locations of plants and stations to reflect changes in them, and expanding its use of river mile indicators for monitoring stations and point sources.

Agency Comments and Our Response

The U.S. Environmental Protection Agency and the U.S. Geological Survey, through the U.S. Department of the Interior, provided general and specific comments on a draft of this report. Their letters are printed in appendixes I and II. We made specific changes in the report in response to some of their comments, which are discussed below.

U.S. Environmental Protection Agency

EPA found our case studies "excellent" and the method "fairly rigorous." EPA's reservations about methodological rigor appear to have been based on the concern that appropriate water-quality data are not available for evaluating the effectiveness of the majority of the Construction Grants Program projects. We are substantially in agreement with this qualification on the universal applicability of our methodology, and we addressed this consideration in the draft report (especially in chapter 3). We do not believe that this constraint lessens the advisability of examining the stream effects of the $40 billion program.

We believe a clarification should be made in the term "GAO's method." Our approach to the case studies was founded on three propositions:

1. Any assessment of the program's effect on the cleanliness of the nation's waters should be based on historical water-quality data.

2. Attributing improvements in water quality to a project under the program requires the verification of four conditions: (a) the project resulted in a decrease in the pollutants being discharged from the plant,

(b) water quality improved after the upgrade, (c) the plant's discharge history is associated with water-quality patterns, and (d) other reasonable explanations of water-quality changes have been excluded.[4]

3. Specific statistical procedures are appropriate for testing the four preceding conditions.

As we discussed in the report, the statistical procedures to be used would depend on data exigencies, hydrological considerations, assumptions about statistical distributions, and other factors. However, the first two propositions are essential and should be applied as rigorously as possible within the constraints of the available data resources.

EPA commented that our four case studies do not represent most Construction Grants Program projects, inasmuch as STORET water-quality data are adequate for only a small fraction of cases. EPA supported this assertion, commenting that in a recent study it had found only 700 instances in which a STORET monitoring station and a sewage-treatment plant were located on the same river reach. Knowing that water-quality data do not exist for all plant upgrades, we believe that examining the effect of only a fraction of the 700 plants would constitute a significant addition to information about the program's effects and would allow a much more precise definition of the limitations of an evaluation restricted to the use of extant data.

In addition, we suggest that EPA's estimate of appropriate data bases is overly pessimistic, since it bases "appropriateness" on the number of monitoring stations identified by STORET as being on the same river reach as treatment plants. If we had used this criterion, we would have found only two of the five downstream monitoring stations on whose data we based our case studies. The two monitoring stations downstream from Allentown were located on the next reach downstream from the sewage-treatment plant, and a third station, below Tamaqua, was not identifiable from STORET software.

EPA expressed concern that our methodology "may not work for waterways with complex hydrology and water quality problems." We agree that some situations may well prove too complex for a method similar to ours. For example, we explicitly excluded lakes, estuaries, and oceans from consideration (as we indicated in chapter 1). In principle, however,

[4]The special case in which an upgrade designed to prevent pollution increases does not have an immediate effect on stream water quality is discussed below.

streams with regulated flows could be examined if the stream flow data that were available allowed procedures, such as those we demonstrate, to compensate for flow variations in receiving streams.

Additionally, we concede that an extreme multiplicity of dischargers could overwhelm a causal analysis while pointing out that two of our case studies dealt with relatively complex situations involving multiple dischargers. At Allentown, we considered many rival dischargers and excluded the majority as not offering alternative explanations for historical water-quality changes; we were able to draw substantive conclusions about the relative influences of three major dischargers to the receiving stream. At Lansdale, we demonstrated that the absence of a clear effect of the upgrade on water quality at our data collection point probably stemmed from offsetting increases in discharge from several other sewage-treatment plants.

EPA pointed out that only one quarter of the program's funding has been spent on net pollutant reduction over the treatment life of the upgraded facilities and expressed concern that an evaluation based on water quality would not indicate the full effects of the program. We acknowledge that it may prove very difficult, if not impossible, to measure the effect on the water quality in a stream from improvements, funded under the program, that do not directly involve treatment plant construction or upgrades. However, the criterion of the program's effectiveness is ultimately the extent to which water quality has been improved—at least, beyond the level of what quality would be if there were no program.

EPA also expressed concern about our statement that an upgrade that has not decreased pollutant discharge can have no beneficial effect on water quality. EPA pointed out that expanding a plant's capacity may not result in a reduction in absolute levels of pollutant discharge but may prevent long-term degradation of water quality. We agree with this qualification and have clarified the statement in question. Our intent was to point out that a sewage-treatment plant upgrade that does not improve treatment efficiency is by definition ineffective. It is possible that an upgrade that only expanded a plant's capacity would have no effect on the concentration of pollutants in the wastewater but would have a long-term effect by maintaining the effectiveness of treatment that would otherwise be degraded by influent increases. However, an increase in capacity would in itself normally decrease pollutant concentrations by allowing greater detention periods and, therefore, greater control over biodegradation.

EPA expressed concern that some state or local environmental agencies might misinterpret the report as suggesting that fixed-station fixed-interval sampling is the only way to evaluate pollution control programs. We welcome this comment and join with EPA in emphasizing that other data collection methods can be used. Our use of data from such stations was dictated by the objective of exploring the usefulness of available data as well as by our judgment that collecting primary water-quality data would not be an appropriate use of our resources. Intensive before-and-after surveys or other data collection designs could well be more cost effective in measuring the effect of an intervention, provided that our four evaluation questions were adequately addressed. However, until more data have become available from such studies, relevant fixed-station data can be a valuable information source.

EPA commented that water-quality criteria other than the traditional ones, which we used, may be necessary for evaluating the condition of a stream. We agree that water quality cannot be comprehensively described by monitoring a small number of pollutants. We conceded this point in chapter 1 and have discussed it more extensively in The Nation's Water: Key Unanswered Questions About the Quality of Rivers and Streams (GAO/PEMD-86-6). The water-quality criteria we used for our case studies were not only the criteria for which data were most readily available but were also the criteria that were most relevant to the interventions we examined.

Sewage-treatment plant upgrades are generally intended not to limit the discharge of toxic pollutants but only to remove the pollutants found in sewage from human waste. In the two case studies of upgrades to advanced treatment, in which the upgrades were designed to alleviate problems from excessive levels of specific chemicals (ammonia and phosphorus) in the receiving streams, we performed additional analyses on the relationship between plant effluent changes from the plants and stream changes in these water-quality parameters. We strongly agree, however, that an analysis of the few parameters (such as dissolved oxygen, BOD_5, pH, and ammonia) that have been monitored most frequently is inadequate for characterizing the health of a body of water, which may be impaired, for example, by toxic substances for whose removal sewage-treatment plants are not typically designed.

EPA noted that the use of monthly averages to characterize a plant's effluent may mask short-term fluctuations. We agree, having discussed this in chapters 2 and 3. We recognize that EPA, in its regulatory responsibility, must also consider monthly extremes of pollutant discharge. We believe, however, that the monthly average statistics in the discharge monitoring reports provide the best available indicator of a plant's typical performance.

EPA suggested that we amend our recommendation that it extend the application of our methodology. We believe that our methodology is flexible enough to allow for the use of data sources and analytic procedures other than those we employed. We continue to believe that an adequate evaluation of the Construction Grants Program must encompass the four questions we addressed in this report. We believe further that an evaluation of the program should begin with the analysis of data already in hand.

U.S. Geological Survey

General Comments

The U.S. Geological Survey generally agreed with our concern that the evaluation of the effects of the Construction Grants Program should be based on empirical evidence of stream water quality. It suggested, however, that we have seriously understated the effort required for evaluating individual sewage-treatment plant upgrades, stating that the complexity of evaluations requires an expenditure of substantial resources that is well justified in view of the nation's already sizable investment in the program.

Our objective was to define the logical requirements of an adequate evaluation of projects under the program, to assess whether extant data bases could be used for an evaluation, and to demonstrate their use. We did not attempt to estimate the number of upgrades whose effectiveness could be investigated with available data; we believe this task is EPA's. We agree with the Geological Survey that the assessment of the effect of an upgrade cannot be made rationally in the absence of familiarity with the hydrology of the area and local sources and patterns of water pollution. For this reason, we interviewed state and local officials who were familiar with the areas we studied, and we reviewed and verified EPA's records of other relevant sources of pollution. In one case, we even verified on site the location of a dam in relation to a monitoring station we intended to include in our analysis.

The Geological Survey has misinterpreted our estimate that 25 to 30 hours were required to perform data analyses for our Lansdale case study. As we indicate in chapter 3, we had previously spent much more time identifying appropriate case study sites, reviewing the hydrology of the region, and collecting and formatting the necessary data. However, we did not consider the effort we spent in data collection, our most time-consuming phase, relevant to estimating the effort required for evaluating planned or recently completed upgrades. Our case studies required collecting discharge data from as early as 1974. We extracted these data from hard copies of reports at EPA, state offices, and the sewage-treatment plants. Current discharge data are now being computerized by EPA and can be accessed much more easily.

For this reason, we believed that we could realistically estimate only the time required for data analysis, which occurred after all data had been collected and prepared for computer-assisted analysis. We explicitly limited our estimate to the cost of time to this latter phase of the evaluation. More comprehensive estimates of the level of effort should be performed by the agency or individuals designing an evaluation and should take into consideration the ready accessibility of the necessary data, the expertise of the evaluation staff, the complexity of the cases, and, most importantly, the level of detail required to achieve the purposes of the evaluation.

The Geological Survey recommended the use of additional documentation, particularly "environmental impact statements" and other planning documents required for grant awards. Such data could be of great value to evaluators. They would allow, as the Geological Survey recommends, the comparison of design criteria for an upgrade with actual results in effluent quality. While we acknowledge that such comparisons would be valid and valuable, we hesitate to recommend their universal application. We agree that this would require "perhaps an order of magnitude more effort," inasmuch as the documentation could be expected to vary greatly in detail and format and is not centrally located, unlike the data bases we employed. Nevertheless, the approach the Geological Survey recommended could well be appropriate to a local evaluation of a limited number of upgrades or to a larger effort to determine the adequacy of pre-upgrade estimates of effects on water quality on the basis of a representative sample of cases.

The Geological Survey distinguished a "before and after" study, as it characterized our report, from a "with and without" study. It is unclear whether the agency intended to draw a qualitative distinction between the two. We suggest that our "before and after" methodology is equivalent to a "with and without" approach to temporally sequential conditions. If the comment means that simple before-and-after comparisons are subject to internal validity threats, such as the possible changes in background conditions the agency enumerated, we fully concur.

It was for this reason that we made background interviews, concerning monitoring testing locations, and attempted to control for other influences such as flow variations and changes in other point-source discharges. As we mention in chapter 3, we were forced to qualify one case study and abandon another because of changes in monitoring locations. We urge evaluators to use some form of screening for validity threats and strongly suggest additional data collection in the presence of apparent anomalies.

However, if the Geological Survey is recommending that evaluations of the program be based on a "with and without" methodology that relies on hydrologic modeling techniques, we disagree. Hydrologic modeling could take one of two forms. The program's effect could be described in terms of changes in the proportion of a stream pollutant attributable to plant effluent—that is, how much difference a decrease in effluent makes when diluted in the stream—without adjustments for stream biochemical transformations. This might yield useful information but would fail to describe the effect of an upgrade on the dynamic conditions of the stream. Alternatively, one could estimate changes in water quality associated with an observed decrease in a particular discharge component. This would certainly be valid, but its demands in costs and expertise would be beyond the resources of most program evaluators.

Similar considerations accompany our response to the Geological Survey's suggestion that we discuss the need for, and design of, water-quality sampling programs. Our use of data from fixed water-quality monitoring stations should not be taken as an implicit recommendation to expand the present monitoring network. We point out in chapter 3, and the Geological Survey implied in its comments, that these stations cannot provide the best evaluation criteria if they are not properly situated and maintained with the explicit purpose of measuring the effect of sewage-treatment plants. We made no attempt to examine the needs and associated costs of a system that could adequately describe the effect of the Construction Grants Program. We are aware of the significant work toward improving the design of the monitoring network over the past decade. The Geological Survey might profitably examine, in the light of this research, the costs and benefits of modifying the networks it currently maintains.

Specific Technical Comments

This report was designed to highlight the need for evaluations based on water-quality changes from sewage-treatment plant upgrades, to identify the logical requirements of an adequate evaluation, and to demonstrate how some relatively simple statistical methods could be applied to extant data to fulfill these requirements. We acknowledge that other statistical procedures might be preferable in some analyses. We welcome suggestions for additional and alternative procedures. We state in the report that the choice of alternatives must be made by the individual evaluator after considering an evaluation's intended audience, available resources, and the objectives. The following response to the specific suggestions of the Geological Survey should be read with this in mind.

1. Did the upgrade of the sewage-treatment plant decrease the amount of pollutants it discharged? We agree with the Geological Survey that the information on ammonia discharge is relevant mostly to the Allentown and Lansdale case studies, since these two upgrades were specifically designed to reduce ammonia discharge. For this reason, we did not analyze changes in ammonia discharge in the Hamburg and Tamaqua cases. We did, however, examine changes in ammonia in their receiving streams in relation to BOD_5 decreases in plant effluent, because of the overlap between these two biochemical measures.

We disagree that the assumption of independence in the t tests may result in overstatements of the statistical significance of before-and-after changes. We assume that adjusting statistically for seasonality, autocorrelation, and other forms of statistical dependence would have raised the significance levels. In any event, the magnitude of the changes in all four cases renders discussion of statistical significance largely academic.

We agree that cyclical patterns in discharge records may have introduced some systematic bias into the data, but we believe it is insignificant in view of the magnitude of pollution decreases observed and the time series data we used, which were sufficiently extensive to compensate for patterns of annual or shorter duration.

2. Has water quality changed downstream from the sewage-treatment plant? We discuss in this report the question of whether any sample or set of samples can be considered truly representative of a stream or a

river system (see particularly chapter 3). As the Geological Survey pointed out, the location of a sampling point can have important effects on the degree to which it reflects upstream changes. Because we chose to use existing data bases for our case studies, we had no control over the location and timing of sampling or over analytic procedures in the laboratory, except to exclude cases that appeared to be clearly unsuitable. Homogeneity may be assumed in the data collection procedures for our cases, since the data were all collected by the same agency in approximately the same period, with very similar frequency, from stations relatively close to one another.

Nevertheless, it is unlikely that the position of the stations coincided with the dissolved oxygen sag points for the effluent discharged from the sewage-treatment plants we studied. This coincidence would occur only where the sampling location had been chosen for measuring the maximum dissolved oxygen effect of the upgrade. Even then, other sampling points in the river would be required for appropriate measures of the maximum effect on other water-quality indicators such as ammonia and phosphorus. For some cases, extensive before-and-after data collection may be justified, but a large-scale evaluation might depend on data collection designs that are less sophisticated.

The Geological Survey called our attention to the fact that an evaluation that examines only stream water-quality improvements from an upgrade will fail to describe the potential benefit to the groundwater from a plant expansion intended to treat waste previously treated only by septic tank disposal. We agree that this evaluation constraint requires particular attention in an assessment of the effect of an expansion only. None of our cases fits this category.

The Geological Survey commented that our deletion of STORET "remarked," or flagged, data from our analyses could obscure before-and-after differences by systematically excluding concentrations below analytical-detection limits. We accept the caution, but since less than 1 percent of the data available from our monitoring stations was "remarked," we believe that the effect of this procedure is minimal and served generally to improve the reliability of the analysis. STORET "remarks" are used to indicate potential problems with the reported values. These problems may stem from any of 20 different threats to reliability, including "estimated value, not accurate" and "actual value is known to be greater than value given." In excluding these observations, we acted on the advice of the EPA officials who were most familiar with STORET.

Other options were available. For example, we could have selectively retained subsets of "remarked" data and included them in nonparametric tests such as the Wilcoxon rank-sum test. We decided not to do so for analytic simplicity. The consequent loss, if any, in statistical power was more than offset by reliability improvements. We encourage further investigation into the practical effects of different treatments of "remarked" data.

We agree that seasonality was not explicitly accounted for in our analyses. An explicit adjustment for seasonality would not substantially alter the results of our analysis, for several reasons. First, our flow-adjustment procedures compensated for much of the season-related variation. Second, the longitudinal data we used were extensive enough to smooth out variations from one season to another. Third, adopting the Geological Survey's nonparametric tests, which are based on a comparison of readings from the same month in different years (for example, January 1976 compared to January 1984 or February 1976 compared to February 1984), would not have added sufficient power to the analyses to justify abandoning intuitively more simple procedures. Finally, as we note in chapter 3, we supplemented the analysis of one of our cases with an analysis using the Geological Survey's procedures. The results were very similar, and we recommend for consideration the research methods developed by the Geological Survey that we have referenced in our bibliography.

3. Were changes in the plant effluent related to changes in stream water-quality indicators? In connection with this evaluation question, the Geological Survey suggested several other topics for discussion. These included the relationship between current water quality and stream standards or previously predicted water quality and the cost and benefit of observed stream changes. These worthwhile subjects were outside the objective of our study.

The Geological Survey proposed that we address the question, "What would concentrations have been without the upgrade?" We discussed this subject above, in our treatment of "with and without" studies. In the Lansdale case, where we found no change in water quality despite a substantial decrease in pollutant load after the upgrade, we addressed this anomaly by documenting other changes in discharge to the stream.

The Geological Survey suggested using a more general measure of association between effluent and stream constituents than the linear correlation coefficient we used. We analyzed these data with both Spearman's rho and Kendall's tau but found no improvement in their explanatory power.

4. Can other explanations of the stream's water-quality conditions be excluded? The Geological Survey inquired why we did not provide multiple regression procedures in the Lansdale case study and asserted that this is the only case in which other point sources appear to be important. We disagree with this assertion, since the influence of the Bethlehem Steel Corporation and the Bethlehem sewage-treatment plant were demonstrated in the Allentown case study. In the case studies in which we used regression approaches, they were employed as examples of possible alternatives for determining whether upstream changes that appeared to correspond with sewage-treatment plant upgrades might be attributable to other point-source changes. In the Lansdale case, this question was largely academic, since no effect from the upgrade was found downstream. However, because the data from other point sources were available to us, we attempted to determine whether they provided an explanation for the lack of water-quality improvement. A comparison of the discharge patterns from other upstream treatment plants revealed that the total pollutant load from all sources increased in the postupgrade period, despite the upgrade at Lansdale. We also applied regression techniques to the Lansdale case study, but the results were inconclusive because of a combination of factors, including the multicollinearity of the effluent measures and the lack of flow data from the monitoring station.

The Geological Survey requested further information on the regression procedures we used in the case studies and on our reasons for choosing the models reflected in the regression statistics presented in tabular form in the statistical appendixes to the second volume of this report. We have added greater detail to our discussion of this approach in chapter 2 of the present volume, in order to clarify the purpose of these tables. It should be evident from the tables that not all models are significant and not all predictors are significant within the significant models. Rather than present the most parsimonious models with a different set of predictors for different station-parameter combinations, we chose to provide as complete a set of statistics as data availability would permit and, thus, to allow the reader to examine the statistical fit and relative influence of predictors of the individual models.

Comments From the U.S. Environmental Protection Agency

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

MAR 19 1986

OFFICE OF
POLICY, PLANNING AND EVALUATION

Mr. J. Dexter Peach
Director
Resources, Community, and Economic
Development Division
U.S. General Accounting Office
Washington, D.C. 20548

Dear Mr. Peach:

On February 10, 1986, the General Accounting Office (GAO) issued a draft report to the Environmental Protection Agency (EPA) for review and comment. The report is entitled "Water Quality: A Proposed Methodology For The Construction Grants Program". According to Public Law 96-226, EPA reviewed the report and provides the statement that follows concerning the draft report.

In the draft report, GAO presented a fairly rigorous method for assessing water quality improvements associated with individual municipal wastewater treatment projects. GAO selected four wastewater treatment facilities for review and completed excellent case studies, using currently available information to assess water quality improvements on a case specific basis.

Feasibility of Using GAO's Methodology

However, the four Pennsylvania case studies are not representative of most Construction Grant program (CGP) projects. The GAO methodology is designed to make use of the storage and retrieval (STORET) stations with long periods of records on surface water quality just downstream of the discharge. The Office of Water (OW) conducted a study recently of the 3500 treatment plants completed under the Clean Water Act (CWA) and found that less than 20 percent of the plants have a STORET station on the same reach. There are even fewer useful downstream STORET stations because many either lack long periods of records or are located outside the area of maximum water quality impact. Extending the GAO method to a large number of cases, as recommended, may therefore prove to be difficult.

- 2 -

Additionally, while the methodology may apply to relatively simple cases with appropriate STORET data, it may not work for waterways with complex hydrology and water quality problems. For instance, the methodology may not work well for estuaries, waterways with multiple dischargers, or streams with regulated flows.

See page 2.

The Executive Summary of the report states that "An adequate evaluation of the Construction Grants Program should be based on stream water quality...." EPA agrees that stream water quality is an important factor in evaluating CGP projects, but it is not the only factor to consider in making comprehensive national conclusions on the effectiveness of the CGP. Over one half of the CWA dollars invested in completed facilities have gone to treatment facilities, but only about one half of those CWA dollars have funded upgrades resulting in net pollutant reduction over the design life of the facilities. (See enclosure). The other CGP funding has covered facilities for public health improvements, reserve capacity for population growth, rehabilitation of wastewater infrastructure, and construction of interceptors in lieu of treatment. Such facilities serve to maintain surface water quality despite growth or protect groundwater against contamination.

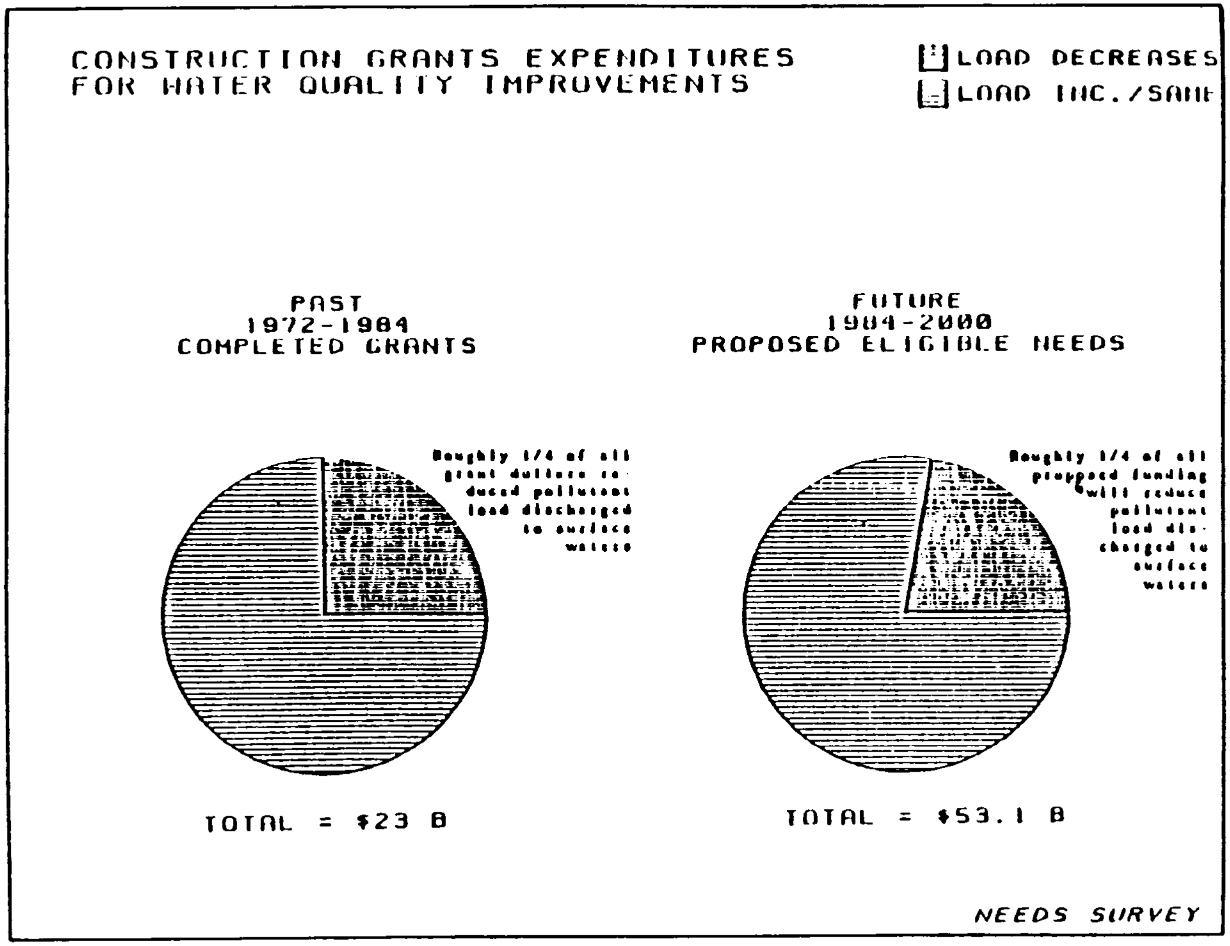
See diagram on page 62.

Although expansion facilities which do not decrease the pollutant load may not immediately improve water quality, they may prevent water quality from worsening. For instance, if a facility continued to receive increasing flows without expanding its capacity to adequately treat the flow, the future water quality would be much worse than that with the expansion. Thus, the report's Executive Summary statement, "if pollutants discharged from the plant are not decreased, clearly the upgrade can have no beneficial effect on water quality" is not necessarily correct.

See pages 3 and 51.

Other Technical Considerations

While the methodology appears to be successful in certain cases where routine fixed station monitoring has taken place over a number of years, this should not be used as the principal reason to establish new monitoring stations without a careful analysis. While EPA realizes the report does not recommend sampling programs, EPA is concerned that Federal, state, or local water pollution control managers who wish to evaluate effects of wastewater treatment plants may conclude from this report that because this sampling and analytical methodology works in these selected cases, it is the only approach which will work. This is not necessarily true. Depending on the situation, there may be more efficient ways to design sampling and analysis to detect instream effects than to conduct fixed-station, fixed-interval sampling to detect changes and trends. For instance, carefully designed intensive surveys before and after a pollution control event may be able to demonstrate water quality relationships more quickly and more efficiently.


CONSTRUCTION GRANTS EXPENDITURES
FOR WATER QUALITY IMPROVEMENTS
LOAD DECREASES
LOAD INC./SAME
PAST
1972-1984
COMPLETED GRANTS
FUTURE
1984-2000
PROPOSED ELIGIBLE NEEDS
Roughly 1/4 of all grant dollars reduced pollutant load discharged to surface waters
Roughly 1/4 of all proposed funding will reduce pollutant load discharged to surface waters
TOTAL = $23 B
TOTAL = $53.1 B
NEEDS SURVEY

- 3 -

Improvements in a small number of routinely monitored pollutants do not necessarily imply that all water quality objectives have been achieved. For instance, the presence of toxic pollutants in water, fish, or sediments may need to be addressed. Also, biological field surveys, bioassays, and human health risk assessment may be needed to assess overall water quality conditions.

Now page 17.

The study used monthly averages to evaluate changes in effluent quality. This approach may mask daily or other short-term fluctuations in effluent quality (Page 1-15).

In summary, GAO's methodology can be used to evaluate water quality impacts from upgraded facilities where (a) there are adequate stream monitoring data, and (b) the hydrologic and water quality situation is relatively simple with long periods of record. Such facilities, however, do not represent most facilities funded under the Construction Grants program. Other methods, perhaps in combination with GAO's methodology, could be used to evaluate the effectiveness of the Construction Grants program. The GAO recommendation to adopt and extend its methodology through legislative mandate and special funding or through Agency initiative has merit, but should be amended to afford flexibility in the methodologies used to evaluate water quality and other impacts of wastewater treatment works.

EPA appreciates the opportunity to comment on the draft report.

Sincerely,

Milton Russell

Milton Russell
Assistant Administrator
for Policy, Planning and Evaluation

Enclosure

Comments From the U.S. Department of the Interior

United States Department of the Interior

OFFICE OF THE SECRETARY
WASHINGTON, D.C. 20240

MAR 2 4 1986

Mr. J. Dexter Peach
Director, Resources, Community
and Economic Development Division
U.S. General Accounting Office
Washington, D.C. 20548

Dear Mr. Peach:

This is in response to your February 10 letter to Secretary Donald Paul Hodel transmitting for review and comment the General Accounting Office (GAO) draft report entitled "Water Quality: A Proposed Evaluation Method for the Constuction Grants Program." The draft report addresses the need to evaluate changes in water quality due to facilities funded by the Construction Grants Program (CGP) under the Clean Water Act. We agree that more information is needed to evaluate the effectiveness of the CGP. The enclosed comments were prepared by the U.S. Geological Survey and they pertain to the level of effort required, the kinds of information needed, and the methods suggested for conducting such an evaluation.

We are concerned that the proposed methods may not adequately document the effectiveness of Sewage Treatment Plant (STP) upgrades and that the estimate of the effort involved in documenting the effectiveness may be too low. We suggest that the GAO consider broadening its recommendations to include new sampling programs and development of methodologies for estimating the effectiveness of STP upgrades. GAO may wish to recommend targeting a subset of existing facilities for sampling and study to provide feedback that would be useful in improving preconstruction analyses regarding STP upgrade effectiveness.

Thank you very much for the opportunity to review and comment on this draft report.

Sincerely,

(SGD) JOSEP[illegible]

Acting Assistant Secretary for
Water and Science

Enclosure

U.S. Geological Survey Comments on

U.S. General Accounting Office Draft Report

"Water Quality: A Proposed Evaluation Method for the Construction Grants Program"

The General Accounting Office (GAO) draft report discusses the need for improved evaluation of the effectiveness of sewage treatment plant upgrades funded by the Construction Grants Program of the Clean Water Act. The draft report describes the questions that such an evaluation program should address, gives some examples of types of analyses that can be used to answer these questions, and suggests changes in the Clean Water Act to require, and fund, such analyses.

The U.S. Geological Survey (USGS) agrees with the general concern for documentation of changes in water quality and the association of these changes with specific actions such as sewage treatment plant (STP) upgrades. However, the GAO's evaluation is in the nature of "before and after" questions, whereas the goal of effectiveness studies should be to answer "with and without" questions. "Before and after" analyses may not address the effectiveness question because the upgrading may have occurred nearly simultaneously with other changes: changes in data collection or analytical techniques, changes in flows to the upgraded treatment plants or to other plants, or changes in land use or hydrologic conditions (e.g., new dams or diversions) which may affect water quality.

"With and without" analyses commonly require more understanding of the hydrology and waste water systems than is involved in the "before and after" types of statistical analyses applied in the draft report. This suggests that the GAO's estimate of 25 to 30 hours to evaluate the effectiveness of a given

2

plant may be a significant underestimate. The data base and software requirements which the draft discusses are important, but an understanding of the hydrologic setting and knowledge of the several types of changes mentioned above are also necessary. This knowledge is unlikely to be acquired in less than several weeks of site visits, review of existing scientific documents, initial data analyses, technical review, and report preparation. Even such an analysis, requiring perhaps an order of magnitude more effort than suggested by GAO, does not seem excessive when considering the size of the investments in the treatment plant upgrades.

The draft report does not utilize the extensive body of analyses and reports, such as Environmental Assessments or Impact Statements (EA/EIS), which must exist in the planning of a treatment plant upgrade. These should provide both pre-plant data and predictions of the water-quality consequences of the planned upgrade. They would allow evaluation of whether plant effluents are meeting design criteria, in addition to whether any significant decrease in concentration, regardless of the size of that decrease, occurs. These reports are a vital part of any post-construction analysis. The routine comparison of actual results to anticipated results of STP upgrades could be formalized into a feedback mechanism which would improve the accuracy of the analytical techniques used in the planning of future STP upgrades.

The draft report also does not address the suitability of available data for making "with and without" comparisons. The most accurate statistical and hydrologic methods are of no use if the location, timing, parameter coverage, and quality assurance of the data collection before and after the upgrade are not carefully designed. For example, if sampling does not occur at locations where dissolved oxygen concentrations are low and at times (season and time of

3

day) when they are low, then the "before and after" comparisons may miss or grossly understate the effectiveness of the plant upgrade. We suggest, therefore, that the GAO should consider broadening its recommendations to include consideration of the need for, and design of, sampling programs as well as the need for data analysis.

The following comments or suggestions deal with the specific methods of analysis suggested or used in the GAO draft report, and are organized by the four questions used for that evaluation.

1. Did the upgrade of the STP decrease the amount of pollutants discharged?

   - Two of the plants (Tamaqua and Hamburg) were only upgraded to secondary treatment, which is not designed to remove ammonia nitrogen. We question the purpose of statistical tests for decreases in ammonia discharges in these cases.

   - The t-test on estimated daily loads was computed based on an assumption of independence. This assumption is incorrect due to the way that daily loads are estimated (using a single weekly measurement of concentration) and because daily loads generally exhibit weekly and annual cycles. Thus, the signficance of the test results may be greatly overstated.

   - The t-test does not consider seasonality, which could be important if combined sewers are present.

2. Have changes occurred in water quality downstream from the STP?

   - The location of the downstream station(s) is crucial to proper evaluation of impact. Their location must be known in relation to the dissolved oxygen (DO) sag. Unlike conservative constituents, the most appropriate location for measuring STP impacts is not typically the closest point to the STP. Hydraulic considerations may also be important--the presence of a dam downstream, for example, often creates low DO conditions at the impoundment, and higher DO as waters flow over the dam. Knowledge of the locations of sampling stations in relation to these factors is crucial, and should be available in the EA/EIS documents.

- Other benefits may accrue because of the upgrade in addition to in-stream improvements below the plant. Protection of ground-water supplies from pollution by septic tank effluent, for example, may be an important benefit envisioned in the planning process. Such critical benefits would not be considered by GAO's present methodology.

- Were the methods and times of collection appropriate for evaluating impact? Surface grab samples, for example, would produce inadequate measures of DO deficit. Point samples (not depth or width integrated) might either miss, or exaggerate, any impact of the upgrade, depending on location and mixing conditions. Such data collection design and quality-assurance questions are not dealt with in the GAO methodology.

- The GAO methodology recommends deletion of all data with concentrations below analytical detection limits prior to statistical analysis. This results in an upward bias of mean concentrations, possibly obscuring any existing "before versus after" differences. Methods which are capable of incorporating these data into the analysis (Wilcoxon rank-sum test) should be used rather than a t-test.

- Seasonality is not accounted for in the GAO methodology. Such seasonality could occur in the flow-concentration relationship, as a result of seasonal nonpoint-agricultural inputs of nutrients from upstream, or seasonal changes in waste loadings. Thus a t-test, even on flow-adjusted concentrations, may not adequately discern "before versus after" differences. The seasonal rank-sum test should be preferred, as seasonality can be accounted for, if present.

3. Does a relationship exist between changes in the plant effluent and changes in instream water-quality indicators?

- If the answer to this question is yes, it is still not clear that the change is: a) adequate in terms of an instream standard or criteria, b) in accordance with the changes predicted during design, or c) beneficial in relation to the cost incurred.

- If the answer is no, significant changes caused by the upgrade may have been masked by other changes in the basin. This was reported by GAO at the Lansdale STP. We believe the appropriate question is "What would concentrations have been without the upgrade?"

- The measure of association used, the correlation coefficient, is a measure of linear association. However, the relationship may not necessarily be a linear one. A more general measure of association (Spearman's rho or Kendall's tau) is more appropriate.

- Previous comments on station location, and the lack of sampling design, also apply here.

6

4. Can other reasonable explanations of stream water-quality conditions be excluded?

- We find the explanation of the multiple regression procedure to be incomplete, making evaluation or reproduction of the technique difficult--example equations are not provided. It is unclear which models are being compared to determine significance of added factors.

- In the one situation where other sources appear important (Lansdale), why was multiple regression not done?

- Tables presented in volume two do not present the significance of added upstream discharges. They instead present coefficients and r^2 for certain models. If explanations were given of why these models are presented, and why others are not, the methodology could be more closely followed by future analysts.

- Therefore, we do not think that success in the stated objective of using multiple regression to define "relative influence" is documented.

In summary, the USGS is generally supportive of the GAO's recommendation but expresses concerns that the level of effort at data collection and data analysis are significantly underestimated, considering the complexity of the problem and the size of the investment.

Bibliography

ADAMS, B. J., and R. S. Gemmell. Water Quality Evaluation of Regionalized Wastewater Systems. Research report 7. Urbana, Ill.: Water Resources Center, University of Illinois, 1973.

ANDERSON, R. L. "Statistics and the Environment: III." Amstat News, November 1983, pp. 2-3.

ASIWPCA (Association of State and Interstate Water Pollution Control Administrators). The States' Evaluation of Progress, 1972-1982. Washington, D.C.: 1984.

BENJES, H. H., Jr. Handbook of Biological Wastewater Treatment. New York: Garland STPM Press, 1979.

"Bipartisan Support Given Water Act Bill in House Debate Despite President's Opposition." Environment Reporter, 15:8 (June 22, 1984), 292-93.

BOX, G. P., and G. M. Jenkins. Time Series Analysis: Forecasting and Control. San Francisco: Holden-Day, 1976.

BRIGGS, J. C., and J. F. Ficke. Quality of Rivers of the United States, 1975 Water Year: Based on the National Stream Quality Accounting Network. Open-file report 78-200. Reston, Va.: U.S. Geological Survey, May 1978.

CAMP, T. R., and R. L. Meserve. Water and Its Impurities. Stroudsburg, Pa.: Dowden, Hutchinson and Ross, 1974.

CBO (U.S. Congressional Budget Office). Public Works Infrastructure: Policy Consideration for the 1980's. Washington, D.C.: April 1983.

CENTER, L. W., and L. G. Hill. Handbook of Variables for Environmental Impact Assessment. Ann Arbor, Mich.: Ann Arbor Science Publishers, Inc., 1979.

CHAMPNEY, L. "A Case Study in the Implementation of the Federal Water Pollution Control Act Amendments." Water Resources Bulletin, 15 (December 1979), 6.

"Chemical Rivers: Cleaner Today, But Clean Enough?" Chemical Week, September 29, 1982, pp. 60-64.

CONOVER, W. J. Practical Nonparametric Statistics. New York: John Wiley and Sons, 1971.

CONSERVATION FOUNDATION. State of the Environment: An Assessment at Mid-Decade. Washington, D.C.: 1984.

COSTLE, Douglas M. "EPA's Construction Grants Program: Perspectives on Costs." Pollution Engineering, September 1978, pp. 40-43.

CRAWFORD, C. G., J. R. Slack, and R. M. Hirsch. Nonparametric Tests for Trends in Water-Quality Data Using the Statistical Analysis System. Open-file report 83-660. Reston, Va.: U.S. Geological Survey, 1983.

D'ASTOUS, F., and K. W. Hipel. "Analyzing Environmental Time Series." Journal of the Environmental Engineering Division: Proceedings of the American Society of Civil Engineers, 105 (October 1979), 979-92.

DELAWARE Department of Natural Resources. Delaware Water Quality Inventory. 3 vols. Dover: 1984.

——. 1982 Delaware Water Quality Inventory. 3 vols. Dover: 1982.

DISTRICT OF COLUMBIA. 1982 Water Quality Assessment and Pollution Control Status Report. Washington, D.C.: 1982.

ELLIS, Franklin W., et al. "Integrating Water Quality and Construction Grants Management." Journal of the Water Pollution Control Federation, September 1984, pp. 1022-29.

EPA (U.S. Environmental Protection Agency). Activities of the Grants Assistance Programs. Washington, D.C.: December 1984.

——. National Water Quality Inventory: 1982 Report to Congress. Washington, D.C.: 1984.

—— and U.S. Fish and Wildlife Service. 1982 National Fisheries Survey. 3 vols. Washington, D.C.: 1984.

FEINTHEL, Roger. "Outlook for Financing Construction of Sewage Treatment Facilities." Environment Reporter, 15:14 (August 3, 1984), 1-24.

FICKE, J. F., and R. O. Hawkinson. "The National Stream Quality Accounting Network (NASQAN): Some Questions and Answers." Circular 719. Reston, Va.: U.S. Geological Survey, 1975.

FISCHER, A. J. "Sanitary Sewage Treatment." In H. W. Gehm and J. I. Bregman (eds.), Handbook of Water Resources and Pollution Control, pp. 481-520. New York: Van Nostrand Reinhold, 1975.

FRANKEL, M. The Social Audit Pollution Handbook: How to Assess Environmental and Workplace Pollution. London: Macmillan, 1979.

FULLER, F. C., Jr., and C. P. Tsokos. "Time Series Analysis of Water Pollution Data." Biometrics, 27 (December 1971), 1017-34.

GAO (U.S. General Accounting Office). Better Data Collection and Planning Is Needed to Justify Advanced Waste Treatment Construction, CED-77-12. Washington, D.C.: 1976.

———. Better Monitoring Techniques Are Needed to Assess the Quality of Rivers and Streams, CED-81-30. Washington, D.C.: 1981.

———. Costly Wastewater Treatment Plants Fail to Perform as Expected, CED-81-9. Washington, D.C.: 1980.

———. Examination into the Effectiveness of the Construction Grant Program for Abating, Controlling, and Preventing Water Pollution, B-166506. Washington, D.C.: 1969.

———. Wastewater Discharges Are Not Complying with EPA Pollution Control Permits, RCED-84-53. Washington, D.C.: 1983.

GIANESSI, L. P., H. M. Peskin, and G. K. Young. "Analysis of National Water Pollution Control Policies I: A National Network Model." Water Resources Research, 17:4 (August 1981), 796-801.

GREESON, P. E., C. J. Velz, and D. A. Rickert. "River Quality Assessments." Water Resources Bulletin, 13:3 (June 1977), 445-53.

HAWKINSON, R. O., J. F. Ficke, and L. G. Saindon. Quality of Rivers of the United States, 1974 Water Year: Based on the National Stream Quality Accounting Network. Open-file report 77-151. Reston, Va.: U.S. Geological Survey, 1977.

HIPEL, K. W., D. P. Lettenmaier, and A. I. McLeod. "Assessment of Environmental Impacts. Part One. Intervention Analysis." Environmental Management, 2:6 (1978), 529-35.

HIRSCH, R. M., J. R. Slack, and R. A. Smith. "Techniques of Trend Analysis for Monthly Water Quality Data." Water Resources Research, 18:1 (February 1982), 107-21.

JOHNSON, N. M., et al. "A Working Model for the Variation in Stream Water Chemistry at the Hubbard Brook Experimental Forest, New Hampshire." Water Resources Research, 5:6 (December 1969), 1353-63.

KOCH, R. W., T. G. Sanders, and H. F. Morel-Seytoux. Detection of Water Quality Changes through Optimal Tests and Reliability of Tests. Completion report 104. Fort Collins, Colo.: Colorado Water Resources Research Institute, Colorado State University, September 1980.

——. "Regional Detection of Change in Water Quality Variables." Water Resources Bulletin, 18:5 (October 1982), 815-21.

LEO, W. M., R. V. Thomann, and T. W. Gallagher. Before and After Case Studies: Comparisons of Water Quality Following Municipal Treatment Plant Improvements. Washington, D.C.: U.S. Environmental Protection Agency, May 1984.

LETTENMAIER, D. P. "Design Considerations for Ambient Stream Quality Monitoring." Water Resources Bulletin, 14:4 (August 1978), 884-902.

——. Design of Monitoring Systems for Detection of Trends in Stream Quality. Technical report 39. Seattle, Wash.: Charles W. Harris Hydraulics Laboratory, University of Washington, August, 1975a.

——. Detection of Trends in Stream Quality: Monitoring Network Design and Data Analysis. Technical report 51. Seattle, Wash.: Charles W. Harris Hydraulics Laboratory, University of Washington, 1975b.

——. "Detection of Trends in Water Quality Data from Records with Dependent Observations." Water Resources Research, 12:5 (1976), 1037-46.

——. "Dimensionality Problems in Water Quality Network Designs." Water Resources Research, 15:6 (December 1979), 1692-1700.

———. "Intervention Analysis with Missing Data." Water Resources Research, 6:1 (February 1980), 159-71.

———, and S. J. Burges. "Design of Trend Monitoring Networks." Journal of the Environmental Engineering Division: American Society of Civil Engineers, 103 (October 1977), 785-802.

———. "Use of State Estimation Techniques in Water Resource System Modeling." Water Resources Bulletin, 12:1 (February 1976), 83-99.

LETTENMAIER, D. P., K. W. Hipel, and A. I. McLeod. "Assessment of Environmental Impacts. Part Two. Data Collection." Environmental Management, 2:6 (1978), 537-54.

LEWIS, D. H. "Optimization of State Water Quality Monitoring Systems." Computers and Operations Research, 3 (1976), 127-43.

LIEBETRAU, A. M. "Water Quality Sampling: Some Statistical Considerations." Water Resources Research, 15:6 (December 1979), 1717-25.

LOFTIS, J. C., and R. C. Ward. "Sampling Frequency Selection for Regulatory Water Quality Monitoring." Water Resources Bulletin, 16:3 (June 1980), 501-07.

LONGEST, Henry L. "The Construction Grants Program: Getting It All Together for the 80's." Water and Sewage Works, March 1980, pp. 36 and 52.

MCCLEARY, R., and R. A. Hay. Applied Time Series Analysis for the Social Sciences. Beverly Hills, Calif.: Sage, 1980.

MARYLAND Department of Health and Mental Hygiene. Maryland Water Quality Inventory. Baltimore: 1982.

———. Maryland Water Quality Inventory. Baltimore: 1984.

MATHUR, S. P. "Thermal Discharges." In H. W. Gehm and J. I. Bregman (eds.), Handbook of Water Resources and Pollution Control, pp. 719-79. New York: Van Nostrand Reinhold, 1975.

MELVIN, N. W., and R. H. Gardiner. The BOD_5/DO Ratio: A New Analytical Tool for Water Quality Evaluation. Washington, D.C.: U.S. Environmental Protection Agency, 1973.

NATIONAL COMMISSION on Water Quality. Report to the Congress. Washington, D.C.: March 18, 1976a.

———. Water Quality Analysis/Environmental Impact Assessment. Report PS-252-167. Springfield, Va.: National Technical Information Service, 1976b.

NATIONAL RESEARCH Council. Environmental Monitoring. Vol. 4. Analytical Studies for the U.S. Environmental Protection Agency. Washington, D.C.: National Academy of Sciences, 1977.

PENNSYLVANIA Department of Environmental Resources. 1984 Water Quality Inventory. Harrisburg, Pa.: 1984.

———. 1982 Water Quality Inventory. Harrisburg, Pa.: 1982.

PESKIN, H. M., et al. A Framework for Analyzing National Water Pollution Control Policy. Washington, D.C.: Resources for the Future, June 12, 1979.

PICKERING, R. J., and J. F. Ficke. "Design of Nationwide Water-Quality Monitoring Networks." Journal of American Water Works Association, 68:2 (1976), 82-85.

RASIN, V. J., Jr., and K. M. Brooks. Potomac River Basin Water. Technical publication 82-1. Rockville, Md.: Interstate Commission on the Potomac River Basin, December 1982.

SANDERS, T. G., and D. D. Adrian. "Sampling Frequency for River Quality Monitoring." Water Resources Research, 14:4 (1978), 569-76.

———, and B. B. Berger. Designing a River Basin Sampling System. Completion report 62. Amherst, Mass.: Water Resources Research Center, March 1986.

SANDERS, T. G., and R. C. Ward. "Relating Stream Standards to Regulatory Water Quality Monitoring Practices." Proceedings of the American Water Resources Association, Symposium on Establishment of Water Quality Monitoring Programs. Minneapolis: American Water Resources Association, 1979.

SANDERS, T. G., et al. Design of Networks for Monitoring Water Quality. Littleton, Colo.: Water Resources Publications, 1983.

SENECA, J. J., and M. K. Taussig. Environmental Economics. Englewood Cliffs, N.J.: Prentice-Hall, 1979.

SMITH, R. A., and R. B. Alexander. A Statistical Summary of Data from the USGS's National Water Quality Networks. Open-file report 83-533. Reston, Va.: U.S. Geological Survey, 1983.

———. A Study of Trends in Dissolved Oxygen and Fecal Coliform Bacteria at NASQAN Stations. Open-file report 82-1019. Reston, Va.: U.S. Geological Survey, December 1982.

SMITH, R. A., R. M. Hirsch, and J. R. Slack. A Study of Trends in Total Phosphorus Measurements at NASQAN Stations. Water-supply paper 2190. Reston, Va.: U.S. Geological Survey, 1982.

STEELE, T. D. "Data and Information Aspects of Modeling." Proceedings of the IUFRO Workshop on Water and Nutrient Simulation Models, pp. 21-28. Birmensdorf, Switzerland: Swiss Federal Institute of Forestry Research, 1981.

———. "Water Quality." In M. G. Anderson and T. B. Burt (eds.), Hydrological Forecasting. New York: John Wiley and Sons, forthcoming.

———, E. J. Gilroy, and R. O. Hawkinson. 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, August 1974.

STEIN, Murray. "Enforcement Measures in Water Quality Control." In J. W. Pratt (ed.), Statistical and Mathematical Aspects of Pollution Problems, pp. 85-91. New York: Marcel Dekker, 1974.

STIFF, M. J. River Pollution Control. Chichester, England: Ellis Horwood Ltd., 1980.

STRAW, R. B., et al. "Using Routine Monitoring Data to Identify Effects and Their Causes." In G. A. Forehead (ed.), Applications of Time Series Analysis to Evaluation. San Francisco: Jossey-Bass, 1982.

THOMANN, R. V. "Some Statistical Analyses of Water Quality in the Delaware River." In J. W. Pratt (ed.), Statistical and Mathematical Aspects of Pollution Problems, pp. 297-311. New York: Marcel Dekker, 1974.

THOMANN, R. V. "Time-Series Analysis of Water-Quality Data." Journal of the Sanitary Engineering Division: Proceedings of the American Society of Civil Engineers, 93 (1967), 1-23.

TIAO, G. C. "Use of Statistical Methods in the Analysis of Environmental Data." American Statistician, 37:4 (November 1983), 459-70.

U.S. COUNCIL on Environmental Quality. Environmental Quality 1983. Washington, D.C.: 1984.

U.S. FEDERAL Water Quality Administration. Clean Water for the 1970's. Washington, D.C.: 1970.

U.S. GEOLOGICAL Survey. National Water Summary 1983—Hydrologic Events and Issues. Reston, Va.: 1984.

VAUGHN, W. J., and C. S. Russell. "Monitoring Point Sources of Pollution." American Statistician, 37:4 (November 1983), 476-86.

VELZ, C. J. Applied Stream Sanitation. New York: Wiley-Interscience, 1970.

———. "Stream Analysis—Forecasting Waste Assimilation Capacity." In H. W. Gehm and J. I. Bregman (eds.), Handbook of Water Resources and Pollution Control, pp. 216-61. New York: Van Nostrand Reinhold, 1975.

VIRGINIA Water Control Board. Water Quality Inventory. Richmond: 1982.

———. Water Quality Inventory. 2 vols. Richmond: 1984.

WAKEFORD, A. C., and G. Knowles. "Enhancement of Water Quality Monitoring Systems by Incorporating Water-Quality Modeling and Time-Series Analysis." Water Pollution Control, 77:4 (1978), pp. 471-91.

WEST VIRGINIA Department of Natural Resources. West Virginia Water Quality Status Assessment 1981-1983. Charleston, W. Va.: n.d.

———. West Virginia Water Quality Status Assessment 1979-1981. Charleston, W. Va.: n.d.

WOLMAN, M. Gordon. "The Nation's Rivers." Science, November 26, 1971, pp. 905-18.

Glossary

Advanced Treatment	Wastewater treatment beyond the secondary, or biological, stage that includes removal of nutrients such as phosphorus and nitrogen and a high percentage of suspended solids.
Biochemical Oxygen Demand, or BOD	A measure of the oxygen consumed in the biological processes that break down organic matter in water. Therefore, it indicates the quantity of organic waste; large quantities of organic waste "demand" large amounts of dissolved oxygen for decomposition, posing a strain on the ecosystem. BOD_5 is a 5-day measure of biochemical oxygen demand.
Combined Sewers	A system that carries both sewage and storm-water runoff. In dry weather, all flow goes to the wastewater-treatment plant. During a storm, only part of the flow is intercepted, because of overloading; the remaining mixture of sewage and storm water overflows, untreated, into the receiving stream.
Dissolved Oxygen	A measure of the concentration of oxygen dissolved within a body of water, often used as a measure of the water's health.
Dissolved Oxygen Deficit	The difference between 1 and the percentage of dissolved oxygen saturation.
Dissolved Oxygen Sag Point	The location downstream from a point source of pollution where the pollutant discharge has its maximum effect on the stream's dissolved oxygen.
Dissolved Oxygen Saturation	The ratio, expressed as a percentage, of observed dissolved oxygen to the maximum amount of oxygen soluble under observed conditions, especially temperature.
Effluent	The discharge from an industrial or municipal wastewater-treatment plant into water such as a river or stream.
Effluent Load	A measure of the quantity of pollution being discharged from a point source into a body of water.

Fecal Coliform Bacteria	A group of organisms common to the intestinal tracts of humans and animals; their presence in water indicates pollution and potentially dangerous contamination.
Flow	The passage of a volume of liquid in a unit of time. As wasteflow, it is commonly measured in millions of gallons per day (mgd); as stream flow, in cubic feet per second (cfs).
Flow-Adjusted Concentration	The concentration of a stream water-quality indicator after mathematical adjustment to compensate for variations in stream flow.
Influent	Flow inward to an industrial or municipal wastewater-treatment plant.
NASQAN	National stream quality accounting network, more than 300 monitoring stations around the nation at which many water-quality characteristics are measured at regular intervals.
Nitrification	The biochemical process in which ammonia is oxidized to nitrate compounds. Some treatment plant upgrades are classified as advanced nitrification treatment, with the goal of reducing high ammonia levels in the water.
Nonpoint-Source Pollution	Diffused pollution resulting from water runoff from urban areas, construction sites, agricultural and silvicultural operations, and the like.
NPDES	National pollutant discharge elimination system, a permit program that imposes discharge limitations on point sources, basing them on national performance standards for new sources or on water-quality standards.
pH	A chemical measure of acidity and alkalinity; in water, the lower the pH is, the more acid is the water. A pH measure of 7 is neutral.

Point-Source Pollution	Pollution discharged through a pipe or other discrete source from municipal wastewater-treatment plants, factories, confined animal feed-lots, or combined sewers.
Primary Treatment	The first stage in the treatment of sewage that uses screens and settling tanks to remove material that settles or floats.
River Reach	A segment of a river or stream of specific length. Most reaches extend between the points of confluence with other streams.
Secondary Treatment	The second stage in wastewater-treatment systems in which bacteria consume the organic content of wastes in trickling filters or activated sludge.
Sewage-Treatment Plant	A series of tanks, screens, and other processes by which pollutants are removed from domestic sewage.
STORET	A computerized data base utility that EPA maintains for the STOrage and RETrieval of parametric data on the quality of the waterways within and contiguous to the United States.
Stream Flow	See Flow.
Tertiary Treatment	See Advanced treatment.
Wasteflow	See Flow.
Water-Quality Criterion	A scientific requirement on which may be based a decision or judgment concerning the ability of water quality to support a designated use.
Water-Quality Standard	A government regulation mandating enforceable limits on water quality.

WQN A prefix identifying water-quality monitoring stations maintained by the Pennsylvania Department of Natural Resources.

Requests for copies of GAO reports should be sent to:

U.S. General Accounting Office
Post Office Box 6015
Gaithersburg, Maryland 20877

Telephone 202-275-6241

The first five copies of each report are free. Additional copies are $2.00 each.

There is a 25% discount on orders for 100 or more copies mailed to a single address.

Orders must be prepaid by cash or by check or money order made out to the Superintendent of Documents.

United States
General Accounting Office
Washington, D.C. 20548