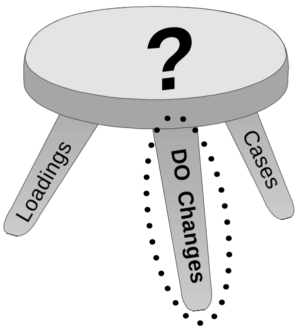
Chapter 3



An Examination of "Worst-Case" DO in Waterways Below Point Sources Before and After the CWA

Chapter 2 discussed the evolution of the BOD measurement, the impact of BOD loadings on DO levels in natural waters, and the massive amount of public and private money invested in municipal wastewater treatment to meet the mandates of the CWA. Key conclusions from the first leg of the three-legged stool approach are:

- The Nation's investment in building and upgrading POTWs significantly reduced BOD effluent loading to the Nation's waterways.
- This reduction occurred in spite of a significant increase in influent BOD loading caused by an increase in population served by POTWs.

The second leg follows up on the first leg with another question—*Has the CWA's push to reduce BOD loading resulted in improved water quality in the Nation's waterways? And, if so, to what extent?* The key phrase in the question is "to what extent?" Earlier studies by Smith et al. (1987a, 1987b) and Knopman and Smith (1993) conclude that any improvements in DO conditions in the Nation's waterways are detectable only within relatively local spatial scales downstream of wastewater discharges.

> *"Perhaps the most noteworthy finding from national-level monitoring is that heavy investment in point-source pollution control has produced no statistically discernible pattern of increases in water's dissolved oxygen content during the last 15 years [1972-87]. . . . The absence of a statistically discernible pattern of increases suggests that the extent of improvement in dissolved oxygen is limited to a small percentage of the nation's total stream miles. This is notable because the major focus of pollution control expenditures under the act [CWA] has been on more complete removal of oxygen-demanding wastes from plant effluents."*
>
> — Knopman and Smith, 1993

The purpose of the second leg of this investigation is to examine evidence that may show that the CWA's municipal wastewater treatment mandates benefited water quality on a broad scale, as well as in reaches immediately downstream from POTW discharges. The systematic, peer-reviewed approach used in this investigation includes the following steps:

- Developing before- and after-CWA data sets composed of DO summary statistics derived from monitoring stations that were screened for worst-case conditions. The purpose of the screening exercise is to mine data that inherently contain a response "signal" linking point source discharges with downstream water quality.
- Calculating a worst-case DO summary statistic for each station for each before- and after-CWA time period and then aggregating station data at sequentially larger spatial scales (reaches, catalog units, and major river basins).
- Conducting an analysis of spatial units that have before- and after-CWA worst-case DO summary statistics and then documenting the direction (improvement or degradation) and magnitude of the changes in worst-case DO concentration.
- Assessing how the point source discharge/downstream DO signal changes over progressively larger spatial scales.

Section A of this chapter provides background on the relationship between BOD loading and stream water quality and discusses the two key physical conditions (high temperature and low flow) that create "worst-case" conditions for DO. *Section B* describes the development and application of a set of screening rules to select, aggregate, and spatially assess before- and after-CWA DO data drawn from USEPA's STORET database. *Section C* presents the results of the comparison analysis of worst-case DO from before and after the CWA for reach, catalog unit, and major river basin scales. The chapter concludes with *Section D*, which provides the summary and conclusions for the second leg of this investigation.

A. Background

In both terrestrial and aquatic ecosystems, the continuous cycle of production and decomposition of organic matter is the principal process that determines the balance of organic carbon, nutrients, carbon dioxide, and DO in the biosphere. Plants (autotrophs) use solar energy, carbon dioxide, and inorganic nutrients to produce new organic matter and, in the process, produce DO by photosynthesis. Bacteria and animals (heterotrophs) use the organic matter as an energy source (food) for respiration and decomposition, and in these processes, consume DO, liberate carbon dioxide, and recycle organic matter back into the ecosystem as simpler inorganic nutrients. Water quality problems, such as depleted levels of DO, nutrient enrichment, and eutrophication (overproduction of aquatic plants), occur when the aquatic cycle of production and decomposition of organic matter becomes unbalanced from excessive amounts of anthropogenic inputs of organic carbon and inorganic nutrients from wastewater discharges and land use-influenced watershed runoff.

DO is the most meaningful and direct signal relating municipal and industrial discharges to downstream water quality responses over a wide range of temporal and spatial scales. In addition to DO's significance as a measure of aquatic ecosystem health, there are several other practical reasons for choosing DO as the signal for assessing changes in water quality, including the following:

- Historical records go as far back as the early 20th century for many major waterbodies. New York City, for example, began monitoring DO in New York Harbor in 1909 and records exist for the Upper Mississippi River beginning in 1926, the Potomac estuary in 1938, and the Willamette River in 1929 (see Wolman, 1971).
- Basic testing procedures for measuring DO have introduced few biases over the past 90 years, thereby providing the analytical consistency needed for comparing historical and modern data (Wolman, 1971).

This section provides background on sources of DO data, the relationship between BOD loading, downstream DO levels, and the two key physical conditions (high temperature and low flow) that create "worst-case" DO conditions. As will be explained, DO data collected under worst-case conditions inherently contain the sharpest signal of the point source discharge/downstream DO relationship.

Sources of DO Data

Key to this analysis is the existence of DO data with which a before- and after-CWA comparison can be made. Fortunately, systematic water pollution surveillance of many of the Nation's waterways began in 1957 in response to the 1956 Amendments to the Federal Water Pollution Control Act. Figure 3-1 is a map developed by Gunnerson (1966) displaying minimum DO concentrations throughout the United States using data collected from 1957 through 1965. It illustrates both the spatial extent of historical data and the poor DO conditions found in many of the Nation's waterways in the late 1950s and early 1960s.

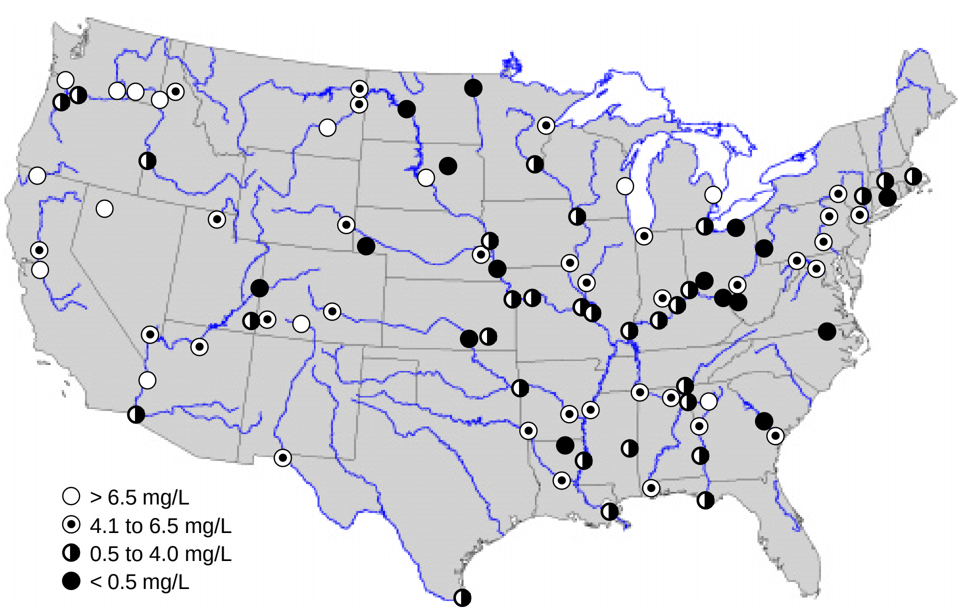


Figure 3-1

Location of sample stations and minimum DO concentrations in the contiguous 48 states from 1957 to 1965.

Source: Gunnerson, 1966.

These and more recent water quality data collected by state, federal, and local agencies are in USEPA's STORET database and available for a before- and after- CWA comparison (Gunnerson, 1966; Ackerman et al., 1970; Wolman, 1971; USEPA, 1974). Currently, the system holds over 150 million testing results from more than 735,000 sampling stations, about 4.6 million of which are DO observations recorded from 1941 to 1995 (Figure 3-2). The challenge was to figure out how to mine STORET's mountain of DO data and create before- and after-CWA data sets that inherently contain the best response "signal" linking point source discharges with downstream DO. This task is not unlike panning for gold. What was needed was a series of screens to divert away all the "rubble and debris" (noisy data), leaving a clean set of "nuggets" (signal data). Using a systematic comparison of before- and after-CWA water quality data sets, the national policy for technology- and water quality-based effluent controls can be considered a success if downstream waterways with poor water quality before the CWA can be shown to have improved significantly after the CWA.

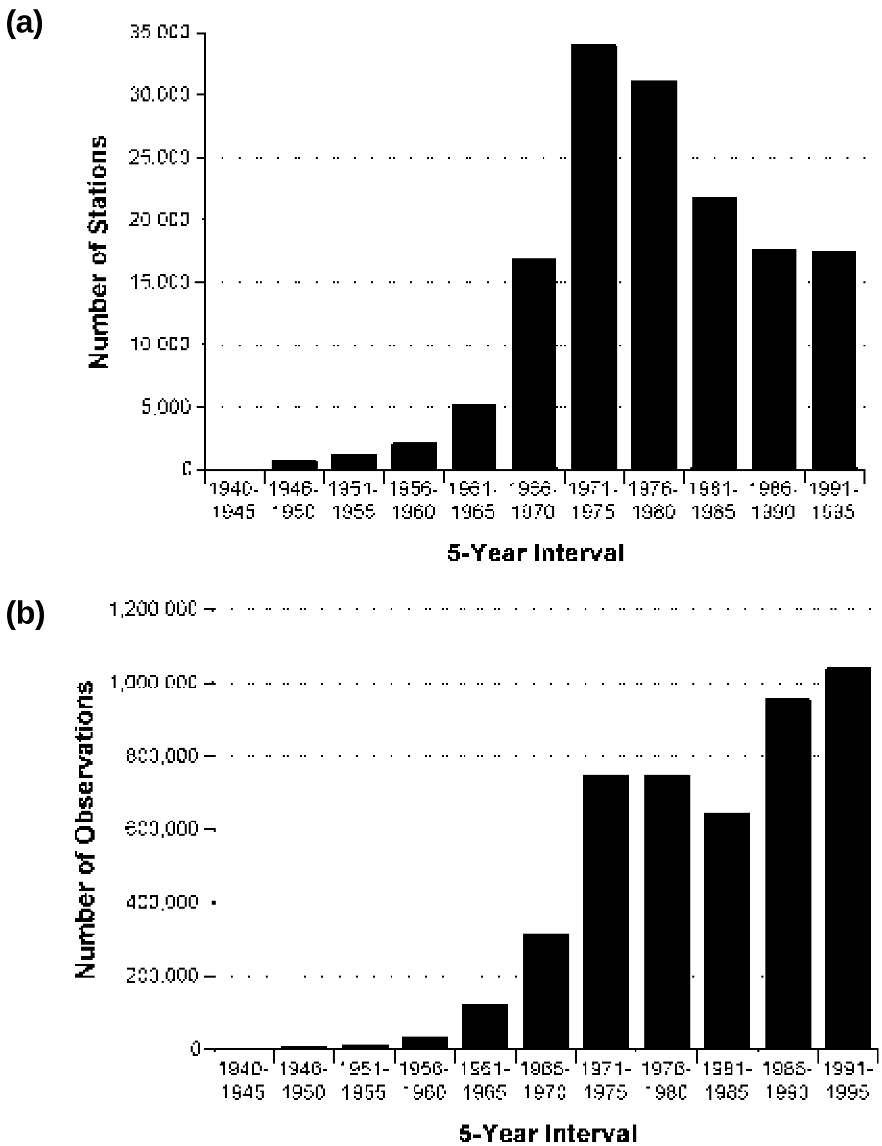


Figure 3-2

National inventory of (a) stations collecting DO data and (b) the number of DO observations made by those stations organized by 5-year intervals from 1941 through 1995.

Source: USEPA STORET.

"Worst-Case" Conditions as a Screening Tool

The first step in developing the before- and after-CWA data sets was to analyze the relationship between point source BOD loading and downstream DO levels. As the reader will see in *Section B*, the rules subsequently adopted and applied to screen out noisy data were based on eliminating physical factors that interfered with, or confounded, the point source discharge/downstream DO signal. As it turned out, the DO data that contained the strongest signal were the data collected under conditions that yielded the lowest DO levels (high water temperature and low flow). The purpose of this subsection is to explain the physical processes and spatial characteristics that make *worst-case* conditions the appropriate screening tool for developing the before- and after-CWA data sets.

Worst-Case Conditions From a Temporal Perspective

In an unpolluted stream, DO concentrations in most of the water column are typically at or near saturation. Saturation, however, varies inversely with water temperature and elevation. At typical winter water temperatures of about 10 °C, the solubility of oxygen is about 11.3 mg/L at sea level. At a higher summer temperature of 25 °C, the solubility is only about 8.2 mg/L. This high water temperature-low solubility relationship makes hot weather an especially critical period for aquatic organism survival. Higher water temperatures mean a lower reserve of oxygen is available to buffer against any additional oxygen demands made by wastewater effluent discharges.

Wastewater effluent typically has an oxygen deficit (a DO concentration below saturation). Therefore, its initial entry into a waterway causes an immediate drop in stream DO near the outfall. The effluent becomes diluted as it mixes with the stream water and moves down the channel. The BOD of the stream water thus becomes the discharge-weighted average BOD of the effluent and the stream above the discharge. The volume of streamflow (the dilution factor), therefore, is a critical variable in determining the concentration of oxygen-demanding waste. Consequently, periods of low flow in the stream channel yield the highest concentration of BOD.

The combination of unnaturally high levels of BOD inputs, high water temperature, and low stream flow creates worst-case DO levels in streams and, in turn, the most critical conditions for the survival of aquatic organisms; that is, conditions of increased oxygen demand, low oxygen solubility, and low dilution potential. Fortunately, worst-case conditions do not occur all the time. Although the BOD loading component tends to remain relatively stable over the course of a year, there are usually distinct seasonal variations in temperature and rainfall (directly related to flow). On an annual basis in the contiguous United States, the highest water temperatures and minimal flow levels usually occur from early summer to late fall. Therefore, the months of July through September are generally considered "worst-case" months for DO.

Observations of year-to-year variations in climate reveal that many areas on the earth, including the United States, experience runs of wet and dry years, a phenomenon known as *persistence*. The short time frame of historical record-keeping makes it difficult for scientists to predict exactly when these wet and dry year cycles will occur; however, more than 100 years of rainfall data have proven that they are not uncommon. Importantly, persistence tends to have a cumulative effect on stream conditions. Therefore, the worst-case scenario for DO in waterways from a temporal perspective can be further refined to include the months of July through September (worst-case months) during a run of dry years (worst-case persistence).

Defining the periods of years before and after the CWA to represent worst-case persistence was accomplished in three steps. In the first step, USGS flow data taken from approximately 5,000 gages with over 20 years of record during the period from 1951 to 1980 were classified into “dry,” “normal,” and “wet” years. Normalized ratios of summer (July to September) streamflow to long-term summer mean were computed for each gage for each year. Years with ratios less than 0.75 were considered dry; normal years had ratios from 0.75 to 1.5, and wet years were defined as having ratios greater than 1.5.

Figure 3-3 illustrates how widely mean summer flow can vary over time. The figure displays USGS gage data from the Upper Mississippi River at St. Paul, Minnesota, for the years 1960 through 1995. The scale on the left Y axis is streamflow measurements as thousands of cubic feet per second (cfs). The scale on the right Y axis is the interannual-to-long-term mean (10,658 cfs) streamflow ratio. Note that the benchmark ratio of 0.75 (which distinguishes dry from normal years) is represented by the dashed horizontal line. This graph shows that dry summers with low flow occurred in St. Paul in the years 1961, 1970, 1976, 1980, and 1987-1989. The data from this gage also show the enormous wet conditions that occurred primarily in response to the “Great Flood of 1993.” That year the mean summer flow was about 4.5 times greater than the normal mean summer flow.

For the second step, a sliding window methodology was used as an algorithm to weight and interpolate normalized streamflow ratios for multiple gages within a catalog unit. The outcome was a weighted streamflow ratio assigned to each catalog unit for each year from 1961 through 1995. Similar to the gage-scale streamflow ratio, the catalog unit-scale streamflow ratio was used to classify catalog units into dry (< 0.75), normal (0.75-1.5), and wet (> 1.5) years.

The third and final step used to define the periods of worst-case dry persistence before and after the CWA involved grouping the 35-year period from 1961 to 1995 into consecutive 5-year “time-blocks.” Then for each catalog unit, the number of years within each time-block during which the catalog unit scale streamflow ratio was below 0.75 (i.e., dry) was determined. Rather than using the seemingly obvious 5-year time-block of 1966-1970 to characterize water quality

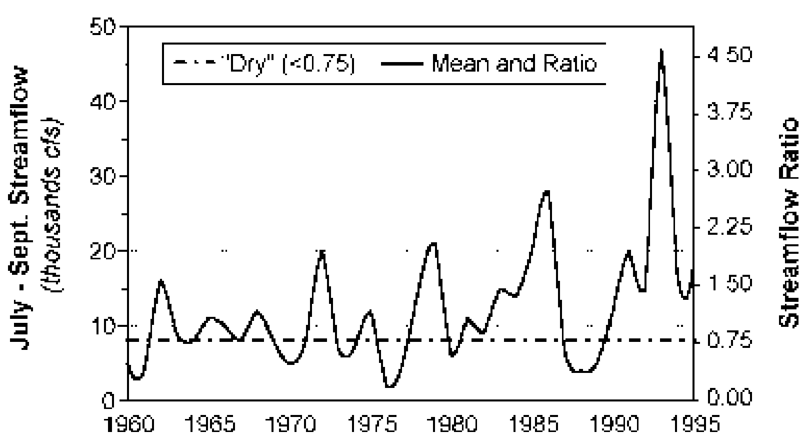


Figure 3-3

Time series of mean summer (July-September 1960-1995) streamflow and ratio of interannual to long-term (1951-1980) summer mean.

(Data from USGS Gage 05331000 on the Upper Mississippi River near Minneapolis-St. Paul, Minnesota)

conditions "before" the 1972 CWA, 1961-1965 was selected instead to represent conditions "before" the CWA while 1986-1990 was used to characterize conditions "after" the CWA.

Widespread drought conditions, a critical factor for "worst-case" water quality conditions, occurred in the Northeast, Middle Atlantic, Midwest, and Central states during both of these "before and after" 5-year time-blocks of record (i.e., 1961-1966 and 1987-1988). The widespread extent of drought conditions during the "before and after" time-blocks is shown in Figure 3-4 with maps of normalized streamflow ratios computed for each catalog unit for 1963 and 1988.

For the 5-year time-block of 1961-1965, selected to represent before-CWA conditions, 1,923 (91 percent) of the 2,111 catalog units of the 48 contiguous states were characterized by at least one year of "dry" streamflow conditions. Similarly for the 5-year time-block of 1986-1990 "after" the CWA, 1,776 (84 percent) of the 2,111 catalog units of the 48 contiguous states were characterized by at least one year of "dry" streamflow conditions. For the catalog units characterized as "dry," low flow conditions occurred for a mean period of 2.5 years during 1961-1965 and 2.7 years during 1986-1990 (Figure 3-5). Hydrologic conditions for the summers of 1963 and 1988 are shown to illustrate the similarity of the spatial extent of drought conditions within the 48 contiguous states during the before- and after-CWA time-blocks. Using this station selection approach based on summer streamflow ratios, trends identified for "before versus after" changes in DO can then be correctly attributed to changes in pollutant loadings (under comparable "dry" streamflow conditions) rather than to differences in hydrologic conditions.

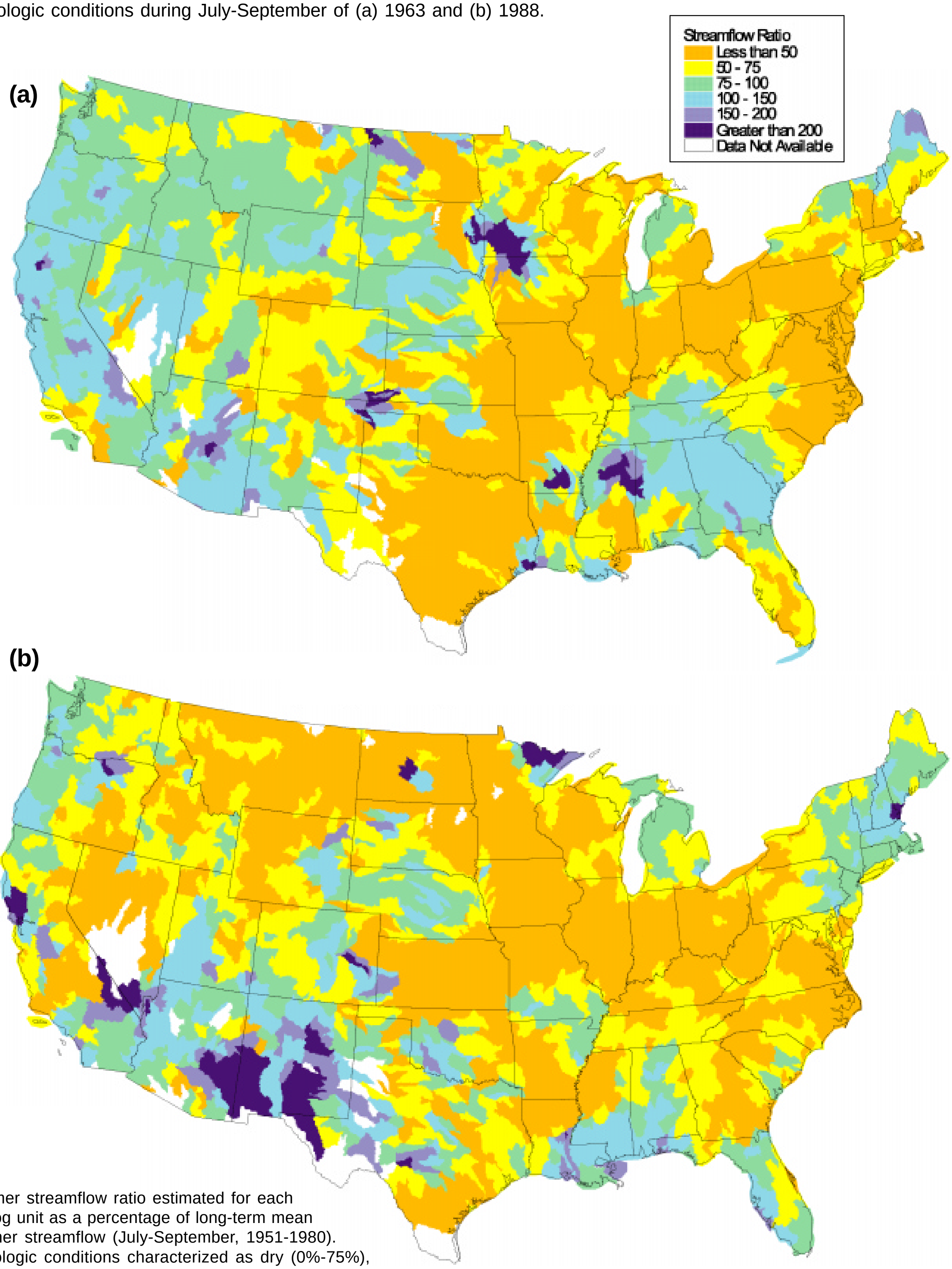
Worst-Case Conditions from a Spatial Perspective

In a clean river, upstream of any wastewater inputs, DO levels are typically near saturation. Downstream of an effluent discharge, however, measurements of DO lower than saturation exhibit a characteristic spatial pattern influenced by the loss of oxygen from degradation of organic matter and nitrification and the replenishment of oxygen transferred from the atmosphere into the river (see Thomann and Mueller, 1987; Chapra, 1997). An understanding of the spatial pattern of DO in rivers was critical for the design of the screening methodology used to detect "worst-case" conditions from a spatial perspective. Using river miles from a downstream confluence as a measure of distance along the river, Figure 3-6 illustrates spatial patterns of carbon (CBOD), nitrogen (organic-N, NH_3-N, and NO_2-N + NO_3-N), and DO in zones identified as "clean," "degradation," "active decomposition," and "recovery" that are upstream and downstream of a POTW discharge.

In streams and rivers, DO levels are maintained near saturation by the continuous transfer of atmospheric oxygen into solution in a thin surface layer of the river. The rate of transfer of atmospheric oxygen into the river (i.e., mixing of oxygen as a gas from the air into solution in the water) depends on how fast the river is running, how deep the water is, how "bubbly" the river appears to be, the water temperature, and how much oxygen is already in solution in the river. The less oxygen that is in solution in the river, the faster more oxygen can be transferred from the air into the water. In the "degradation" zone, more oxygen is being consumed by decomposition than can be replenished from the atmospheric supply of oxygen and DO levels quickly drop. In the "active decomposition" zone,

Figure 3-4

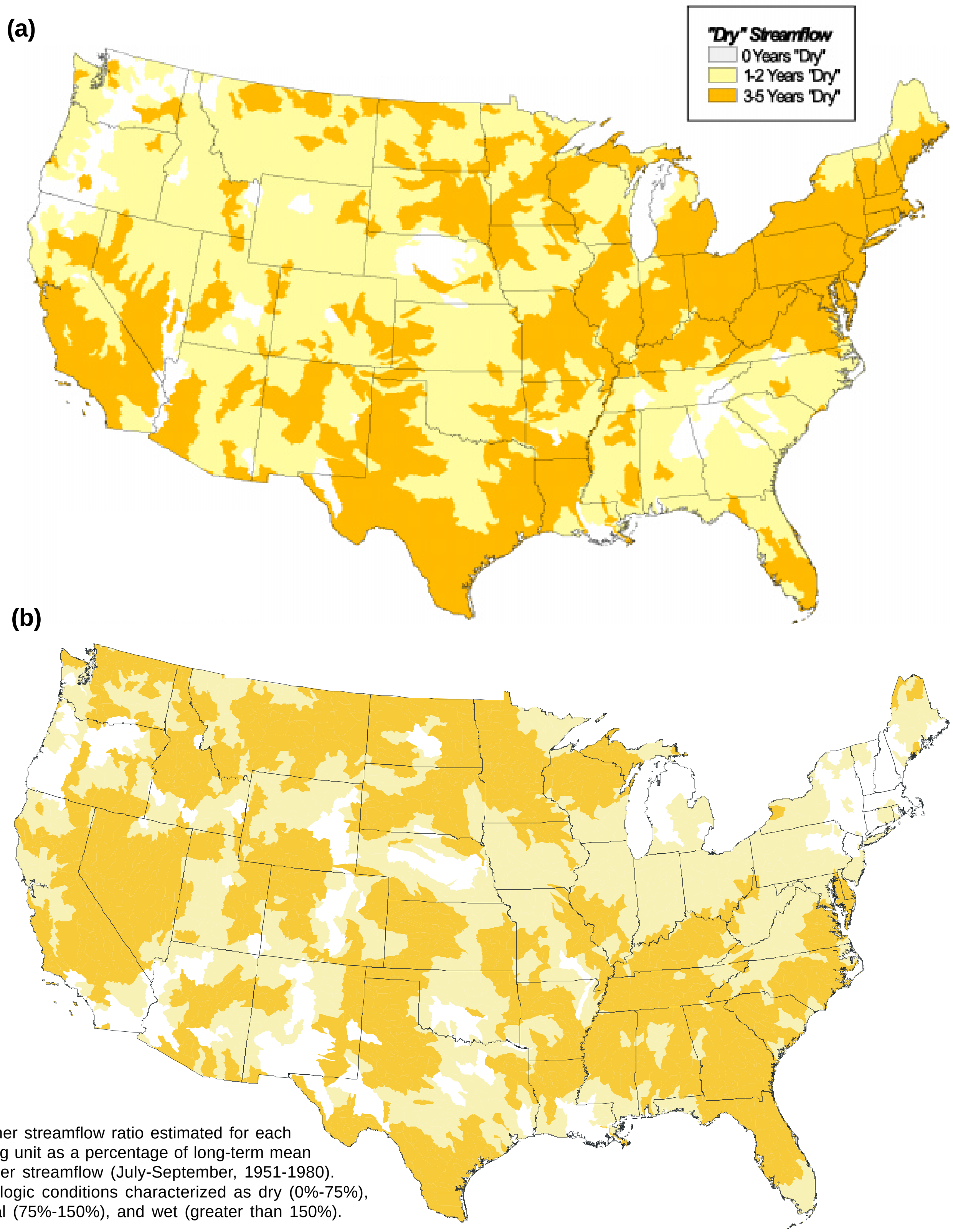
Hydrologic conditions during July-September of (a) 1963 and (b) 1988.



Summer streamflow ratio estimated for each catalog unit as a percentage of long-term mean summer streamflow (July-September, 1951-1980). Hydrologic conditions characterized as dry (0%-75%), normal (75%-150%), and wet (greater than 150%).

Figure 3-5

"Dry" hydrologic conditions during July- September of (a) 1961-1965 and (b) 1986-1990.

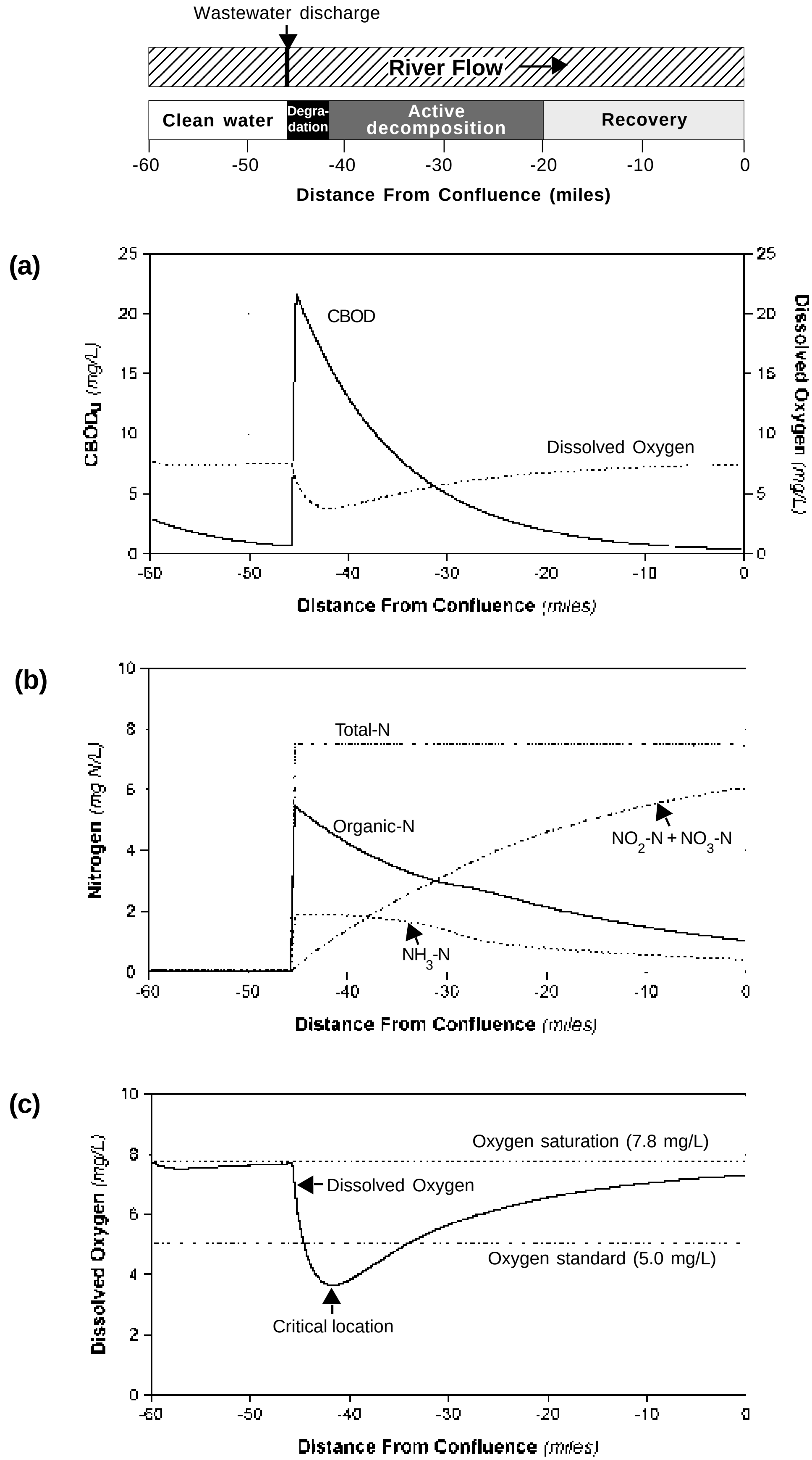


Summer streamflow ratio estimated for each catalog unit as a percentage of long-term mean summer streamflow (July-September, 1951-1980). Hydrologic conditions characterized as dry (0%-75%), normal (75%-150%), and wet (greater than 150%).

Figure 3-6

Spatial distribution of (a) organic carbon (CBOD), (b) nitrogen (organic nitrogen, ammonia, nitrite, and nitrate), and (c) DO downstream of a waste-water discharge into a river.

Source: Adapted from Chapra, 1997 and Thomann and Mueller, 1987.



more oxygen is gained by the mixing of oxygen from the air into the water than is lost by the continued decomposition of a diminishing amount of carbon (CBOD) and nitrogen (NBOD) and oxygen gradually increases. In the "recovery" zone, the rate of atmospheric replenishment of oxygen greatly exceeds the oxygen lost due to small levels of CBOD and NBOD remaining in the river and oxygen returns to the saturation level.

Immediately downstream from the POTW, the carbon concentration (CBOD) jumps from the low upstream level to a much higher flow-weighted CBOD concentration as the effluent load is diluted with the ambient upstream load (Figure 3-6(a)). Bacterial decomposition of the carbon results in a steady decrease of in-stream CBOD and a steep drop in oxygen in the "degradation" zone followed by a continued decline of CBOD with a gradual increase in oxygen in the "active decomposition" zone.

As shown in Figure 3-6(b) for the spatial patterns of nitrogen, organic nitrogen (Organic-N) and ammonia nitrogen (NH_3-N) both jump from a low upstream level following mixing of the wastewater load with the ambient upstream load. As organic nitrogen declines by hydrolysis, the nitrification process begins (if a sufficient "seed" population of nitrifying bacteria is present), ammonia is oxidized to nitrite, and nitrite is quickly oxidized to nitrate. In the figure, nitrite and nitrate are shown combined as the sum (NO_2-N + NO_3-N) of these two inorganic forms of the nitrogen cycle. As the sequential reactions of the nitrogen cycle proceed downstream, the concentration of total nitrogen (Total-N) remains unchanged to maintain the mass balance of the reactions between the organic and inorganic forms of nitrogen. In these sequential oxidation reactions of nitrification, the nitrogenous oxygen demand (NBOD) consumes oxygen faster than it can be replenished by atmospheric reaeration and oxygen drops.

The combined effect of the decomposition of carbon and nitrogen causes a characteristic critical low DO zone identified by a "sag" in the spatial distribution of oxygen (Figure 3-6(c)). Two key features of the "oxygen sag" curve are especially important for the purposes of this study:

- The magnitude of the minimum DO concentration.
- The distance downstream from a waste discharge affected by "degradation" and "active decomposition."

In designing the screening methodology to detect the "worst case" for oxygen from a spatial perspective, it is important to recognize that water quality monitoring stations located immediately downstream of wastewater inputs will most likely be within the zones of "degradation" or "active decomposition" but not necessarily at the minimum, or critical, location of the sag. For monitoring stations located considerably farther downstream from a wastewater discharge location, it is less likely that the station will be within the "degradation" or "active decomposition" zones of the river. It is more likely, rather, that the station(s) will be located in the "recovery" zone. For any stream or river, the actual locations that mark the beginning and end of these zones are highly variable. The spatial pattern of oxygen shown in Figure 3-6(c) is dependent on a number of factors, including streamflow and river velocity (travel time), depth, water temperature, the type and makeup of effluent discharged, the magnitude of the wastewater discharge load, and the degree of turbulent mixing. Rather than attempting to select stations that are located in the exact sag zone, which would undoubtedly show the sharpest downstream DO signal but in the smallest area of the waterbody, the opposite approach was taken. That is, location of the station relative to the sag zone is purposely not controlled or selected, thereby allowing representation of far larger spatial areas but at the possible sacrifice of the downstream DO signal strength.

The question originally posed in Chapter 1 is broad-based: *How have the Nation's water quality conditions changed since implementation of the 1972 CWA's mandate for secondary treatment as the minimum acceptable technology for POTWs?* The focus of the analysis is on detecting improvements in water quality conditions downstream of POTWs in the Nation as a whole, not just areas immediately below outfalls. Consequently, when the term "worst-case DO data" is used in this document, it should be taken to refer to data collected primarily during times of high water temperature and low flow conditions (i.e., "worst-case" from a temporal perspective). Spatially, no screens were developed for selecting monitoring stations located at the deepest part of the sag curve, nor even for stations in the sag curve itself. The only screening rule applied was that the water quality station had to be downstream from a point source. Thus, a station might be anywhere from within a few yards to hundreds of miles below any particular outfall. As a result, the data sets developed for the comparative before- and after-CWA analysis contain a mix of DO data from within and outside DO sag curves.

The Role of Spatial Scale in This Analysis

Recall that the objectives for this portion of the study are as follows:

- Develop before- and after-CWA data sets made up of DO summary statistics derived from monitoring stations that inherently contain a response "signal" linking point source discharges with downstream water quality.
- Calculate a DO summary statistic (10th percentile) for each station for each before- and after-CWA time period and then aggregate station data at sequentially larger spatial scales (reaches, catalog units, and major river basins).
- Conduct an analysis of all spatial units having both a before- and an after-CWA summary statistic and then document the direction and magnitude of the changes in worst-case (summer, mean 10th percentile) DO concentration.
- Assess the change in the point source discharge/downstream DO signal over progressively larger spatial scales.

The use of spatial scale is a key attribute of this analysis. Detection of positive change in signal at large (river basin) as well as small (stream reach) scales would provide evidence that the CWA's technology and water quality-based controls yielded broad as well as localized benefits (i.e., reaches both within and beyond the immediate sag curve have benefited from the CWA). If true, therefore, the second leg of the three-legged stool approach would provide further support for the claim that the CWA was a broad success.

B. Data Mining

As discussed in the previous section, the key objective in the data mining process was to screen out data collected under conditions or factors that might interfere with, or confound, the point source discharge/downstream DO signal. This section presents the six-step, peer-reviewed data mining process the study authors developed and implemented to develop the before- and after-CWA data sets to be used in the comparison analysis.

Step 1—Data Selection Rules

The data selection step incorporated three screening rules:

- DO, expressed as a concentration (mg/L), will function as the signal relating municipal and industrial discharges to downstream water quality responses.
- DO data will be extracted only from the July-September (summer season) time period.
- Only surface DO data (DO data collected within 2 meters of the water surface) will be used.

DO Concentration (mg/L) as the Water Quality Indicator

The rationale for selecting DO as the water quality indicator for this study was discussed earlier in this chapter and in Chapter 2. The only question remaining was how this parameter should be expressed in the analysis—by concentration or by percent of DO saturation. The latter measurement has some advantages because it would reduce the noise introduced by changes in temperature. However, DO expressed as mg/L concentration was ultimately selected because it is more intuitive to a broader audience. For example, USEPA has established a DO concentration of 5.0 mg/L as the minimum concentration to be achieved at all times for early life stages of warm-water biota (see Table 1-1). For this reason, this level of DO is used as a benchmark for assessing acceptable versus nonacceptable conditions. In contrast, it is somewhat more difficult to comprehend whether a DO saturation of 50, 60, or 70 percent is protective.

DO From the Time Period of July to September

Summer and early fall (July through September) is usually the best time for evaluating worst-case impacts of wastewater loading on water quality in general and DO in particular. Typically, this is when water temperatures are highest and flow is the lowest (i.e., lowest oxygen solubility and lowest dilution potential). Selecting DO data from only this time period screens out noise introduced by seasonal variations in temperature, precipitation, and flow. In addition, BOD loadings from nonpoint sources of pollution are reduced during low precipitation periods thus minimizing this contribution to DO signals.

DO from Surface Waters

In lakes, reservoirs, estuaries, coastal waters, and deep rivers, scientists typically measure DO at several depths in the water column. Often these measurements reveal significant differences between surface and bottom DO concentrations because of thermal stratification and the lack of reaeration of the bottom layer. By limiting DO data selection to the top 2 meters of a waterway, one can screen out much of the noise associated with the physical, chemical, and biological processes that occur in the lower layers and maintain some level of comparability between shallow streams and deeper waters.

Step 2—Data Aggregation Rules From a Temporal Perspective

The data aggregation from a temporal perspective step incorporated the following rules:

- 1961-1965 will serve as the time-block to represent persistent dry weather before the CWA and 1986-1990 will serve as the time-block to represent persistent dry weather after the CWA.
- To remain eligible for the before- and after-CWA comparison, DO data must come from a station residing in a catalog unit that had at least 1 year classified as dry (streamflow ratio < 0.75) out of the 5 years in each before- and after CWA time-block.

An analysis of catalog units revealed that 1,923 (91 percent) of the 2,111 catalog units in the contiguous United States experienced at least one dry summer in the 1961-1965 time-block. Further, a total of 1,475 catalog units (70 percent) experienced at least two dry summers and 886 catalog units (42 percent) experienced at least three dry summers in the before-CWA time-block. Of the catalog units that remained eligible for the comparison analysis (note that only 188 were screened out), low flow conditions remained for an average of 2.5 years. In the 1986-1990 time-block, 1,776 (84 percent) of the 2,111 catalog units in the contiguous United States experienced at least one dry summer. A total of 1,420 catalog units (64 percent) experienced at least two dry summers and 1,073 catalog units (51 percent) experienced at least three dry summers in the after-CWA time-block. Of the catalog units that remained eligible for the comparison analysis (335 were screened out), low flow conditions remained for an average of 2.7 years.

Step 3—Calculation of the Worst-Case DO Summary Statistic Rules

The calculation of the worst-case DO summary statistic step incorporated the following rules:

- For each water quality station, the 10th percentile of the DO data distribution from the before-CWA time-block (July through September, 1961-1965) and the 10th percentile of the DO data distribution from the after-CWA time-block (July through September, 1986-1990) will be used as the DO worst-case statistic for the comparison analysis.
- To remain eligible for the before- and after-CWA comparison, a station must have a minimum of eight DO measurements within each of the 5-year time-blocks.

Typically, the mean or median statistic is used to summarize a distribution of data because it describes the central tendency of the distribution. In this study, however, the emphasis is on worst-case (low) DO. Consequently, a summary statistic describing the lowest DO measurements of the data distribution was needed because these data would inherently carry a sharper point source discharge/downstream water quality signal. In other words, the objective was to characterize the worst of the DO data collected under the worst-case physical conditions (high temperature and low flow).

Because simply choosing the minimum measurement might introduce anomalous results, the 10th percentile, a more robust statistic (i.e., one that conveys information under a variety of conditions and is not overly influenced by data values at the extremes of the data distribution) was selected as the appropriate summary statistic to characterize the worst DO of a station's range of DO measurements within a time-block. An example of how one might interpret a 10th percentile value for a station is to say that 90 percent of the values collected at that station were higher than the 10th percentile value. To minimize statistical errors associated with calculating extreme percentiles, the requirement was added that a station must have a minimum of eight observations within the 5-year time-block to remain eligible for the before- and after-CWA comparison.

Step 4—Spatial Assessment Rules

The spatial assessment step incorporated one screening rule:

- Only water quality stations on portions of streams and rivers affected by point sources will be included in the before- and after-CWA comparison analysis.

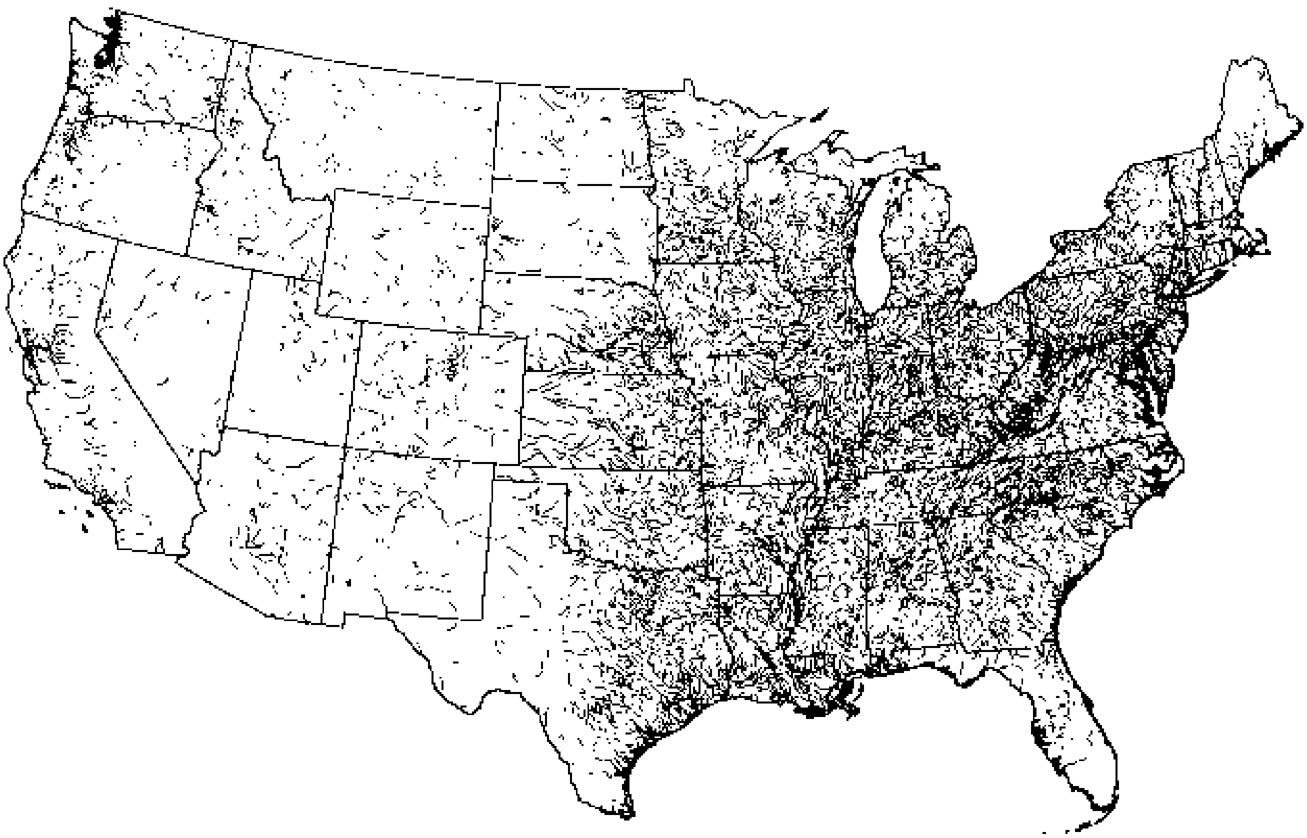
The objective was to develop before- and after-CWA data sets that contain data that inherently contain a response signal linking point source discharges with downstream water quality. Consequently, a screening rule reflecting the need to ensure that DO data came from stations located downstream, rather than upstream, from point sources was required. As noted in *Section A* of this chapter, the distance downstream was not relevant for the screening rule; the only requirement was that the station was somewhere in the downstream network.

Although the focus of this study is on effluent loading from POTWs, changes in DO are tied to industrial discharges as well. Estimates of current (ca. 1995) BOD_5 loading using the NWPCAM indicate that industrial loads are the dominant component of total point and nonpoint source loading in many catalog units associated with major urban-industrial areas (see *Section E* in Chapter 2). For this reason, and because of the fact that it is not always possible to satisfactorily distinguish between industrial and POTW outfalls because of their close proximity in many areas, this leg of the study defines "point source discharges" to include both industrial and municipal dischargers.

The upstream/downstream relationship between point source discharges and water quality monitoring stations was established using USEPA's Reach File, version 1 (RF1). RF1 is a computerized network of 64,902 river reaches in the 48 contiguous states, covering 632,552 miles of streams (see Figure 1-2). Using this system, one can traverse stream networks and establish relative positions along the river basin network of both free-flowing and tidally-influenced rivers.

Figure 3-7

Reach File version 1 stream reach network of the 48 contiguous states with point source inputs discharging to a reach.



A list of point source dischargers was developed from EPA's Permit Compliance System (PCS), Clean Water Needs Survey (CWNS), and Industrial Facilities Discharge File (IFD). Spatially integrating the dischargers with RF1 resulted in identifying 12,476 reaches that are downstream of point source dischargers (Figure 3-7) (Bondelid et al., 1999). These reaches, in turn, reside in 1,666 out of a total of 2,111 catalog units in the contiguous United States.

Example Application of the Screening Rules on DO Data From a Single Water Quality Monitoring Station

Figure 3-8 illustrates how the above screening rules were applied to monitoring station data to obtain worst-case DO data for the before- and after-CWA comparison analysis. A station located on the Upper Mississippi River at Lock and Dam No. 2 at Hastings, Minnesota, is used in this example. Figure 3-8(a) displays a time series of the entire historical record (225 observations) of raw ambient DO measurements for the station from 1957 to 1997. Note that DO concentrations fluctuate from close to zero to slightly over 15 mg/L. The apparent noise (rapid up and down movement of the DO line) is due to many factors, including seasonal streamflow-water temperature and the time scale of the graphic. Long-term interannual changes, on the other hand, might be due to persistent dry or wet weather or changes in pollutant loading from the St. Paul METRO wastewater facility.

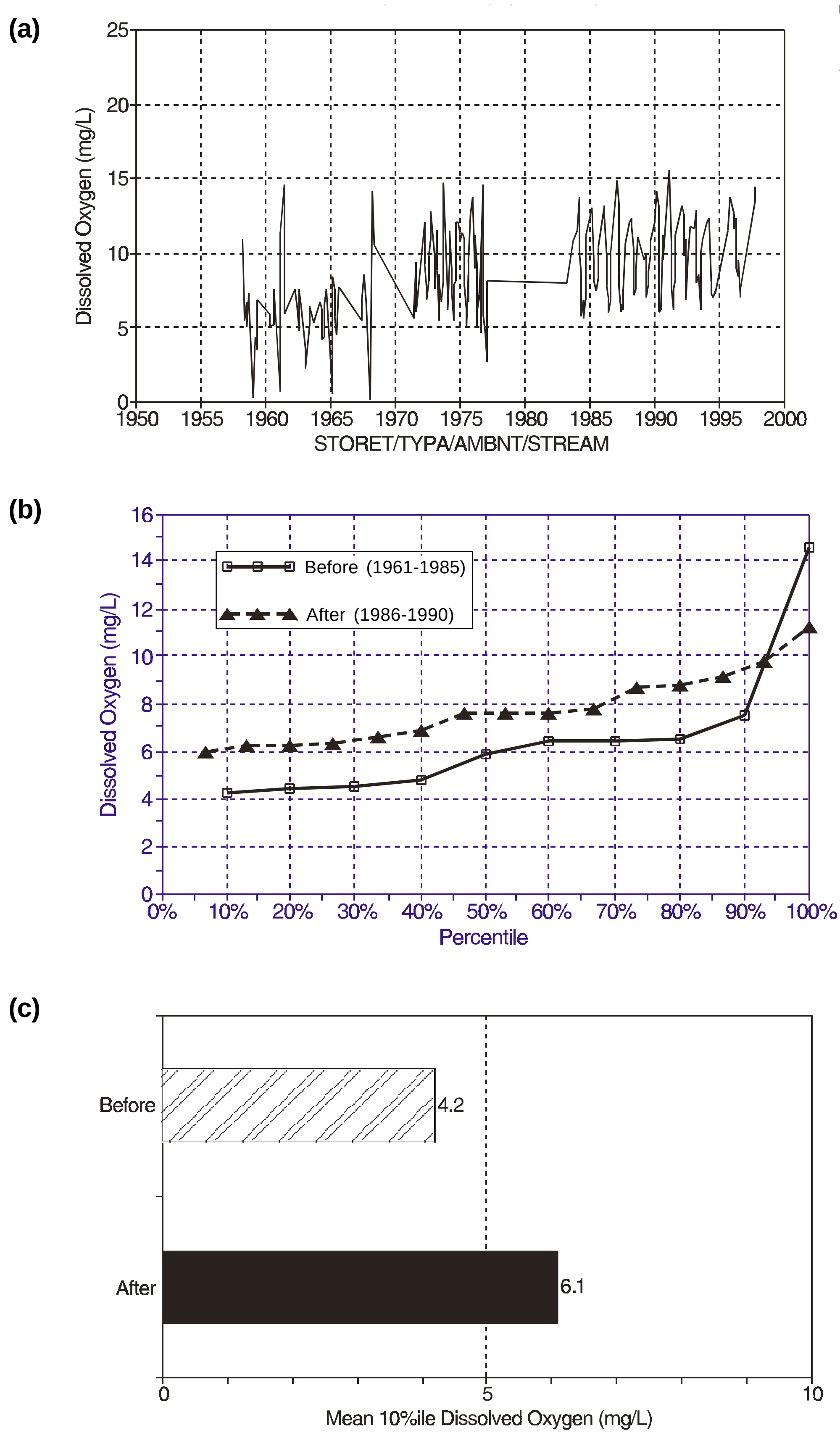


Figure 3-8

Application of the screening rules for station 21 MINN MSU-815-BB15E58 located in the Upper Mississippi River: (a) time series of historical DO observations from 1957-1997, (b) before- and after-CWA frequency distribution, and (c) 10th percentile values.

Source: USEPA STORET

In the data selection step, the study authors extracted from the raw data set surface measurements collected at the station during the summer season (52 observations). Then, in the data aggregation step, they grouped the data in 5-year time-blocks and focused in on the data from the before- and after-CWA persistent dry weather time-blocks of 1961-1965 (10 observations) and 1986-1990 (15 observations). Because (1) the catalog unit in which the station resides had at least one dry year in each of the before- and after-CWA time-blocks (streamflow ratios: 1961 [0.31]; 1964 [0.65]; 1987 [0.59]; 1988 [0.22]; and 1989 [0.40]) and (2) the number of observations for each grouping was confirmed to be greater than eight, the groupings remained eligible for the next phase.

Distributions were made for each group and the 10th percentile determined. Figure 3-8(b) displays the before- and after-CWA DO frequency distribution. Figure 3-8(c) is a bar chart comparing the 10th percentile DO values of the before- and after-CWA time-blocks. Note that the 10th percentile statistic associated with the before-CWA period is below the USEPA's minimum concentration of 5 mg/L, the level the Agency requires to be achieved at all times for early life stages of warm-water biota.

Finally, the spatial assessment phase confirmed that the monitoring station where the DO data were collected was on the Upper Mississippi River downstream of the St. Paul METRO water pollution control plant. Therefore, the station remained eligible for the comparison analysis.

Step 5—Data Aggregation Rules From a Spatial Perspective

The data mining steps described above were used to develop before- and after-CWA sets of monitoring station data. Recall that

- The before- and after-CWA data sets are collections of DO summary statistics that characterize worst-case DO at individual water quality monitoring stations across the United States for the 1961-1965 and 1986-1990 time-blocks, respectively (one DO summary statistic per station per time-block).
- The summary statistic used to characterize worst-case DO at a station is the 10th percentile value of a data distribution of actual DO measurements taken at the station during the specified time-block and recorded in STORET. For the station to be eligible for inclusion in the data set, at least eight measurements had to have been taken during the 5-year time-block.

The purpose of the data aggregation from a spatial perspective step was to assign a worst-case DO summary statistic to every eligible spatial unit defined at the reach and hydrologic unit scales for the before- and after-CWA time-blocks. This task was accomplished in two steps. First, for each data set and time-block, the mean 10th percentile value from each eligible station was computed within the spatial unit. (Since the scales are hierarchical, a station's summary statistic was effectively assigned to a reach *and* a catalog unit.) Second, the mean 10th percentile summary statistic was calculated and assigned to the spatial unit for the purpose of characterizing its worst-case DO. If a spatial unit had only one monitoring station within its borders meeting the screening criteria, the 10th percentile DO value from that station simply served as the unit's worst-case summary statistic. If, however, there were two or more stations within a spatial unit's borders, the 10th percentile values for all the eligible stations were averaged, and this mean value used to characterize worst-case DO for the unit. This averaging process reduced the correlation between stations that were located near each other. (Increased correlation reduces the effective sample size and complicates statistical comparisons. Averaging across larger spatial scales tends to reduce the correlation the most. As demonstrated later in this section, the results from the different spatial scales are generally consistent and the impact of spatial correlation is believed to be minimal.)

Step 6—Development of the Paired Data Sets (at each spatial scale)

The purpose of the sixth and final step was to prepare the before- and after-CWA data sets for the comparison analysis to be conducted at each of the three sequentially larger hydrologic scales (RF1 reach, catalog unit, and major river basin). The screening rule associated with this step was as follows

- To be eligible for the paired (i.e., before vs. after) comparison analysis, a hydrologic unit must have both a before-CWA and an after-CWA summary statistic assigned to it.

After each eligible reach and catalog unit was assigned a worst-case DO summary statistic for the appropriate before- and after-CWA time-blocks, a check was made to see which spatial units had *both* a before- and an after-CWA summary statistic. For many reaches and catalog units, factors such as the absence of dry flow conditions, station removal or changes in station location, or water quality sampling over time (see Figure 3-2) caused a summary statistic to be available for one time-block but not the other. In this case, the spatial unit was removed from the analysis.

Implementation of this final step of the data mining process yielded the following results:

- Of the 12,476 reaches identified as being downstream from point sources, 311 reaches had both before- and after-CWA worst-case DO summary statistics.
- Of the 1,666 catalog units identified as being impacted by point sources, 246 units had both before- and after-CWA worst-case DO summary statistics.
- The 311 reach-aggregated DO summary statistics were pooled by the 18 major river basins in the contiguous United States. Using the statistical requirements to conduct a paired *t*-test as a criterion, 11 of the 18 major river basins had sufficient reach-aggregated worst-case DO data to conduct the comparison analysis at the river basin level.
- The number of reaches and catalog units with *both* before- and after-CWA DO data was constrained by the limited availability of station records for the 1961-1965 before-CWA period (see Figure 3-2).

C. Comparison of Worst-Case DO in Waterways Below Point Source Discharges Before and After the CWA at Three Spatial Scales

This section presents the comparative before- and after-CWA analysis of worst-case DO data derived using the screening criteria described in *Section B* and then aggregated by spatial units defined by three scales (reach, catalog unit, and major river basin). In the discussion that follows, the term "worst-case DO" should be interpreted to mean the average 10th percentile DO statistic computed for the corresponding spatial level unless specifically noted otherwise. Also, the reader should note that a worst-case DO concentration of 5 mg/L has been adopted in this report as a general benchmark threshold for defining "desirable" versus "undesirable" levels of worst-case DO. This benchmark value was chosen primarily because USEPA has established it as the minimum concentration to be achieved at all times for early life stages of warmwater biota (see Table 1-1).

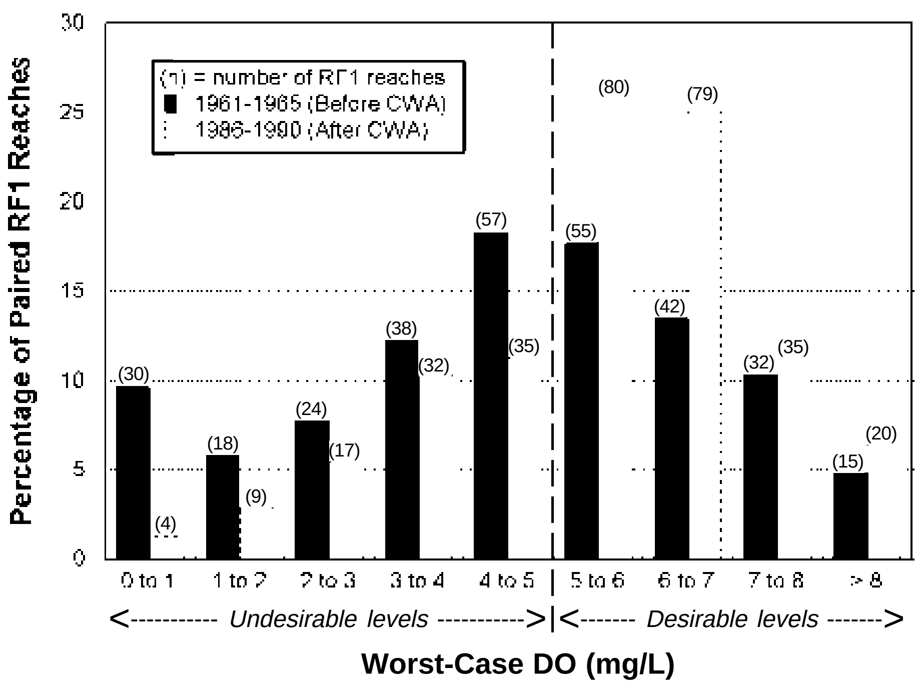
Reach Scale

A total of 311 river reaches had monitoring stations with both before- and after-CWA data and thus were eligible for comparison. Notably, these 311 evaluated reaches represent a disproportionately high amount of urban/industrial population centers, with approximately 13.7 million people represented (7.2 percent of the total population served by POTWs in 1996). Of this total, 215 reaches (69 percent) showed improvements in worst-case DO after the CWA. Figure 3-9 presents a frequency distribution of the before- and after-CWA data.

Figure 3-9

Frequency distribution comparing worst-case DO concentration of evaluated reaches before and after the CWA.

Source: USEPA STORET



Key observations from Figure 3-9 include the following:

- The percentage of evaluated RF1 reaches characterized by "very low" worst-case DO (< 2 mg/L) was reduced from 15 to 4 percent. Before the CWA, 48 reaches had very low worst-case DO. After the CWA, only 13 reaches had very low worst-case DO.
- The percentage of evaluated reaches characterized by undesirable worst-case DO (below the 5 mg/L threshold) was reduced from 54 to 31 percent. Before the CWA, 167 reaches had undesirable levels of worst-case DO. After the CWA, 97 reaches had undesirable levels of worst-case DO.
- The percentage of evaluated reaches characterized by desirable worst-case DO (above the 5 mg/L threshold) increased from 46 to 69 percent. Before the CWA, 144 reaches had desirable levels of worst-case DO. After the CWA, 214 reaches had desirable levels of worst-case DO.

By tracking individual reaches, it was revealed that 85 of the 167 reaches characterized by undesirable worst-case DO before the CWA improved to greater than 5 mg/L after the act. On the flip side, only 15 of the 144 reaches characterized by desirable worst-case DO before the CWA dropped below the 5 mg/L benchmark after the act. Thus, the net change was 70 reaches moving from undesirable levels of worst-case DO to desirable levels of worst-case DO.

Figure 3-10 is a column graph that breaks down the 85 reaches that had undesirable DO levels before the CWA and then improved past the benchmark threshold of 5 mg/L after the act according to their before-CWA worst-case DO concentration.

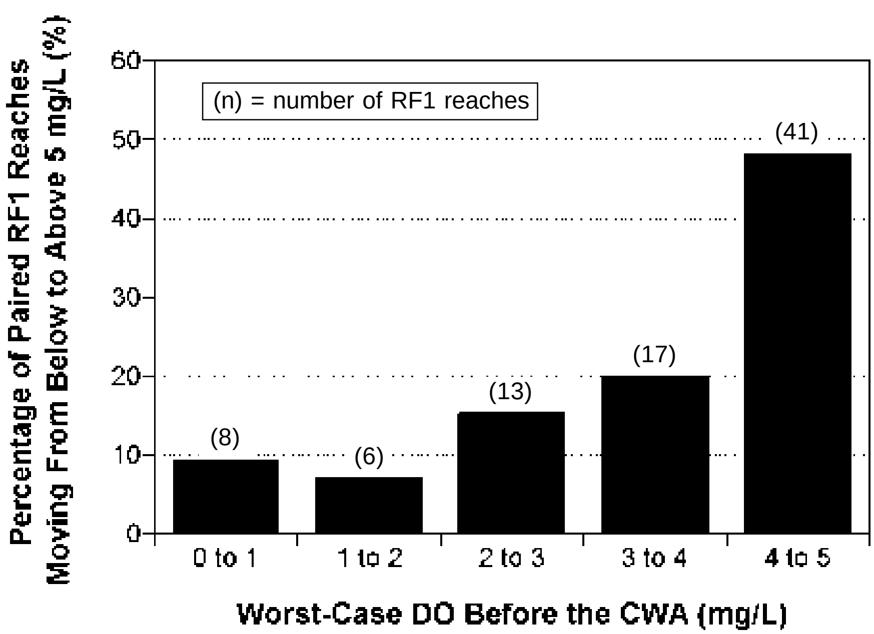


Figure 3-10

Frequency distribution of worst-case DO levels before the CWA for the 85 evaluated reaches that were < 5 mg/L before the CWA and > 5 mg/L after the CWA.

Source: USEPA STORET

Key observations from Figure 3-10 include the following:

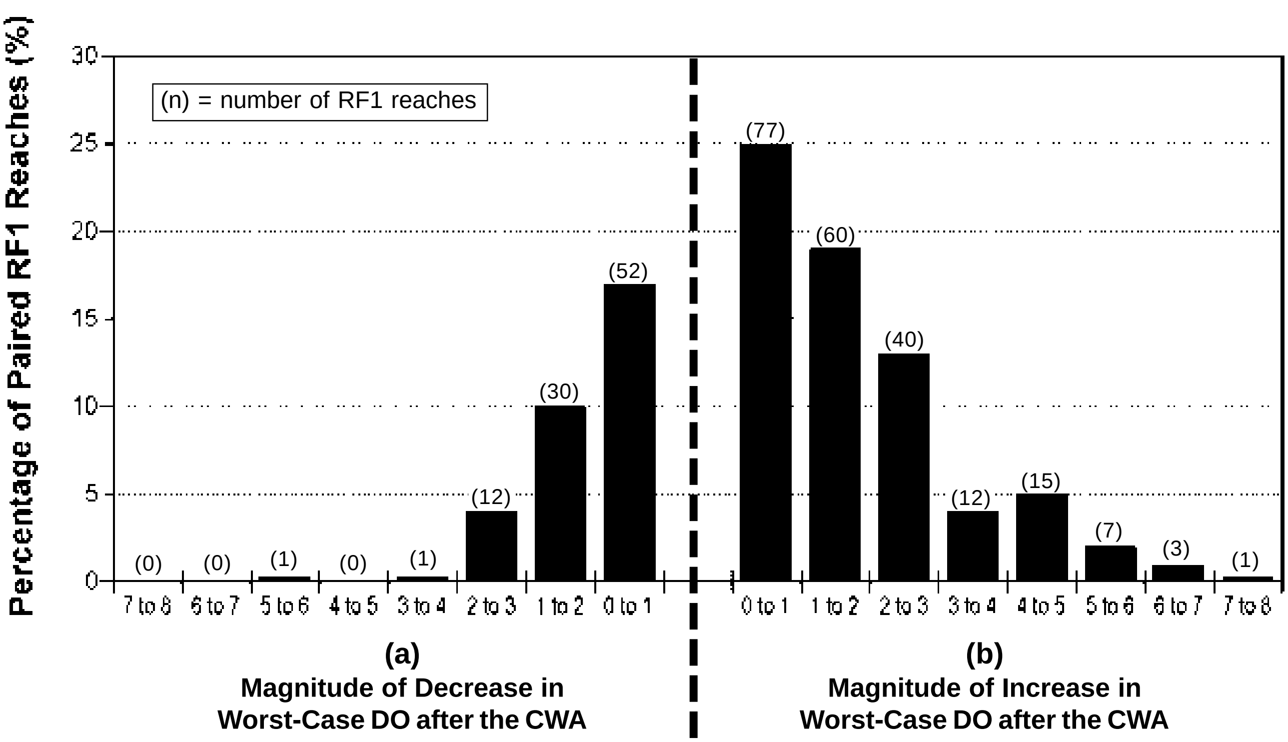
- Approximately 48 percent of the evaluated reaches (41 out of 85) that had undesirable worst-case DO levels before the CWA and then improved past the benchmark threshold of 5 mg/L after the act had before-CWA worst-case DO in the 4-5 mg/L range.
- Approximately 16 percent of the evaluated reaches (14 out of 85) that had undesirable worst-case DO levels before the CWA and then improved past the benchmark threshold of 5 mg/L after the act had very low worst-case DO (< 2 mg/L) before the CWA.

Of the 311 evaluated reaches with paired before- and after-CWA data, 215 reaches (69 percent) had increased worst-case DO and 96 (31 percent) had decreased worst-case DO after the CWA. Parts (a) and (b) of Figure 3-11 display the magnitude of degradation and improvement, respectively. Key observations from Figure 3-11 include the following:

- Approximately 36 percent of the evaluated reaches that had increases in worst-case DO (78 of the 215 improving reaches) increased by 2 mg/L or more.
- Approximately 15 percent of the evaluated reaches that had decreases in worst-case DO (14 of 96 degrading reaches) decreased by 2 mg/L or more.
- Approximately 41 percent of all evaluated reaches either stayed the same or improved or degraded by 1 mg/L or less (129 of the 311 reaches).

Figure 3-11

Frequency distribution of worst-case DO for evaluated RF1 reaches that (a) decreased in concentration (n = 96) and (b) increased in concentration (n = 215) after the CWA. *Source: USEPA STORET*



Reaches with Greatest Improvements

Table 3-1 lists the 25 river reaches with the greatest before- and after-CWA improvements in worst-case DO. Figure 3-12 presents a location map of these reaches along with a stacked column graph that shows their before- and after-CWA worst-case DO data. Key observations from Table 3-1 and Figure 3-12 include the following:

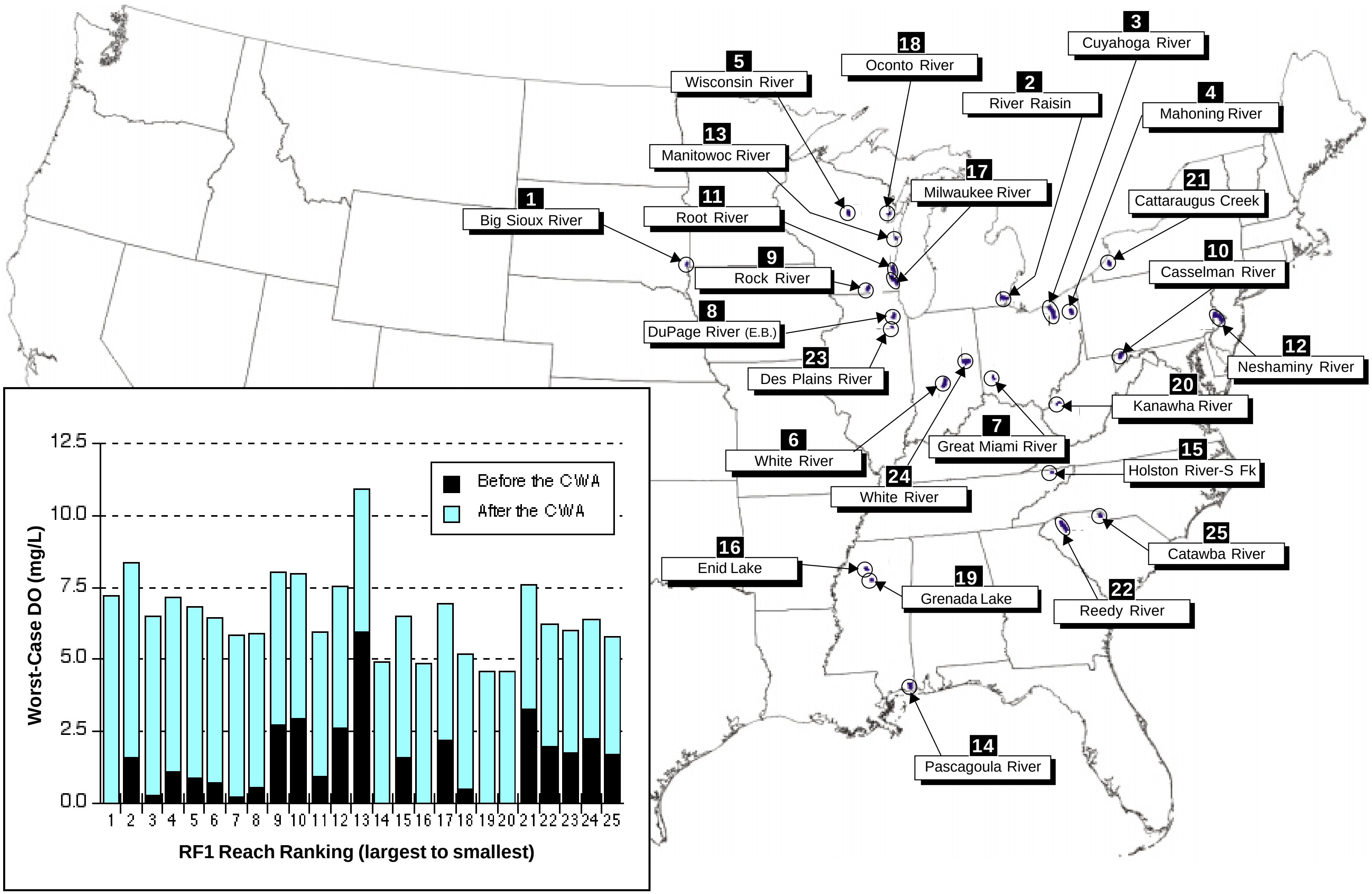
- All but one of the top 25 river reaches with the greatest before- and after-CWA improvements had before-CWA worst-case DO levels below the benchmark threshold of 5 mg/L. Five reaches had a before-CWA worst-case DO concentration of 0 mg/L.
- For 20 of the 24 reaches with before-CWA worst-case DO levels below the threshold value of 5 mg/L, after-CWA worst-case DO improved to levels greater than 5 mg/L.
- The four reaches that did not break the threshold value of 5 mg/L after the CWA all had a before-CWA worst-case DO concentration of 0 mg/L.
- Worst-case DO in the top 10 improving river reaches typically improved by about 4 to 7 mg/L (from about 0-3 mg/L in the 1961-1965 time-block to about 6-8 mg/L in the 1986-1990 time-block).

Table 3-1. Twenty-five RF1 river reaches with greatest improvements in worst-case (mean 10th percentile) DO after the CWA. *Source: USEPA STORET*

Rank	Reach ID	River Reach Name	Catalog Unit Name	Worst-case DO 1961-1965 *(mg/L)*	Worst-case DO 1986-1990 *(mg/L)*	DO Change *(mg/L)*	No. of Stations 1961-1965	No. of Stations 1986-1990
1	10170203037	Big Sioux R.	Lower Big Sioux, IA	0.0000	7.2200	7.2200	1	1
2	04100002001	River Raisin	Raisin, MI-OH	1.6000	8.3400	6.7400	2	2
3	04110002001	Cuyahoga R.	Cuyahoga, OH	0.2950	6.4967	6.2017	2	24
4	05030103007	Mahoning R.	Mahoning, OH-PA	1.0900	7.1600	6.0700	1	1
5	07070002034	Wisconsin R.	Lake Dubay, WI	0.8800	6.8400	5.9600	1	1
6	05120201004	White R.	Upper White, IN	0.6900	6.4240	5.7340	5	1
7	05080002008	Great Miami R.	Lower Great Miami, IN	0.2000	5.8600	5.6600	1	1
8	07120004018	Du Page R., E Br.	Des Plaines, IL	0.5750	5.9200	5.3450	4	3
9	07090001004	Rock R.	Upper Rock, IL	2.7600	8.0500	5.2900	1	1
10	05020006031	Casselman R.	Youghiogheny, MD	2.9600	8.0000	5.0400	1	1
11	04040002005	Root R.	Pike-Root, IL	0.9400	5.9400	5.0000	1	1
12	02040201011	Neshaminy R.	Crosswicks-Neshaminy	2.6000	7.5600	4.9600	1	1
13	04030101012	Manitowoc R.	Manitowoc-Sheboygan, WI	5.9500	10.9000	4.9500	1	1
14	03170006007	Pascagoula R.	Pascagoula, MS	0.0000	4.9200	4.9200	1	7
15	06010102004	Holston R, S Fk.	South Fork Holston,	1.6000	6.4800	4.8800	1	2
16	08030203006	Enid L.	Yocona, MS	0.0000	4.8673	4.8673	1	3
17	04040003001	Milwaukee R.	Milwaukee, WI	2.1800	6.9567	4.7767	2	3
18	04030104002	Oconto R.	Oconto, WI	0.5000	5.2000	4.7000	1	1
19	08030205018	Grenada L.	Yalobusha, MI	0.0000	4.6160	4.6160	1	4
20	05050008006	Kanawha R.	Lower Kanawha, WV	0.0000	4.5667	4.5667	2	3
21	04120102002	Cattaraugus Cr.	Cattaraugus, NY	3.3000	7.6000	4.3000	1	2
22	03050109053	Reedy R.	Saluda, SC	1.9500	6.2270	4.2770	4	10
23	07120004002	Des Plains R.	Des Plaines, IL	1.7620	6.0000	4.2380	2	1
24	05120201013	White R.	Upper White, IA	2.2267	6.3750	4.1483	3	2
25	03050103037	Catawba R.	Lower Catawba, NC	1.6780	5.8000	4.1220	5	1

Figure 3-12

Location map and distribution chart of the 25 RF1 reaches with the greatest after-CWA improvements in worst-case DO. *Source: USEPA STORET*



Catalog Unit Scale

Figures 3-13 and 3-14 are maps that display the locations and worst-case DO concentrations of catalog units potentially eligible for the paired analysis for the 1961-1965 and 1986-1990 time-blocks, respectively. The before-CWA data set contained a total of 333 catalog units. The after-CWA data set had 905 catalog units.

In the before-CWA map (Figure 3-13),

- 45 of the 333 catalog units (14 percent) have worst-case DO less than 2.5 mg/L.
- 102 of the catalog units (31 percent) have levels from 2.5 to 5 mg/L.
- 186 of the catalog units (56 percent) are characterized by worst-case DO greater than 5 mg/L.

In comparing these results with the historical data from the FWPCA surveillance network (see Figure 3-2 in Section A of this chapter), many of the catalog units characterized by poor DO conditions (DO less than 5 mg/L) in 1961-1965 correspond to the areas represented by many of the stations compiled by Gunnerson (1966) with minimum DO less than 0.5 and minimum DO between 0.5 and 4 mg/L in the 1957-1965 data set (see Figure 3-1).

In the after-CWA map (Figure 3-14),

- 49 of the 905 catalog units (5 percent) have worst-case DO less than 2.5 mg/L.
- 252 of the catalog units (28 percent) have levels from 2.5 to 5 mg/L.
- 604 of the catalog units (67 percent) are characterized by worst-case DO greater than 5 mg/L.

Undesirable levels of worst-case DO (less than 5 mg/L) are still quite prevalent after the CWA in some midwestern and southeastern watersheds, with a pattern of moderately low worst-case DO (2.5 to 5 mg/L) that appears to be characteristic of the Atlantic coastal plain from Florida to New Jersey. Higher worst-case DO (5 to 7.5 mg/L) characterizes the Piedmont region and the watersheds of the Appalachian Mountains and is likely due to cooler water temperatures. The coastal plain pattern of moderately low worst-case DO most likely reflects natural factors such as warmer summer temperatures, higher decomposition rates, and relatively long residence times within sluggish rivers and tidal waters rather than municipal or industrial point source loading within these watersheds.

Overlaying the 333 eligible catalog units in the before-CWA data set with the 905 eligible units in the after-CWA data set yielded a total of 246 intersecting catalog units that had *both* before- and after-CWA data. Notably, these 246 evaluated catalog units represent a disproportionately high amount of urban/industrial population centers, with approximately 61.6 million people represented (32.5 percent of the total population served by POTWs in 1996). Figure 3-15 presents maps that display the locations and worst-case DO concentrations of the evaluated catalog units. Figure 3-15(a) displays the catalog units that had improvement in worst-case DO after the CWA. Figure 3-15(b) displays the catalog units that had degradation in worst-case DO after the CWA. Figure 3-16 presents a frequency distribution of the before- and after-CWA data.

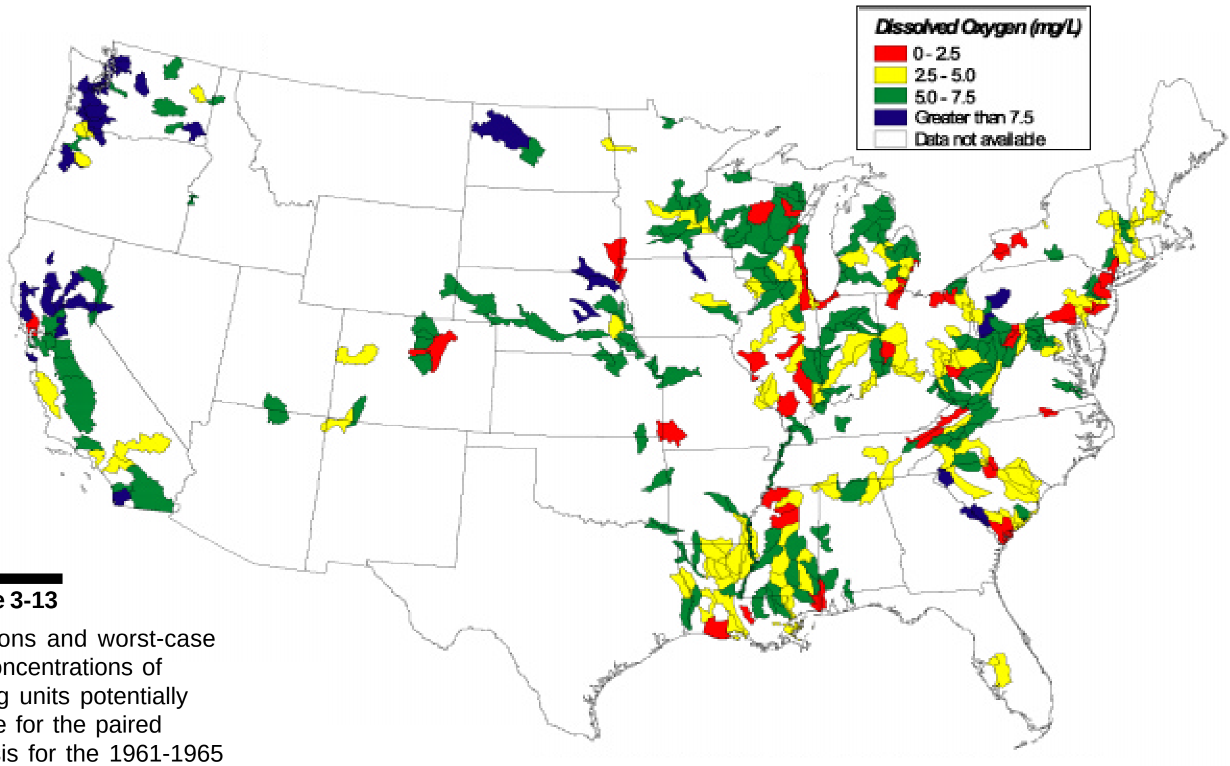


Figure 3-13

Locations and worst-case DO concentrations of catalog units potentially eligible for the paired analysis for the 1961-1965 time-block (before-CWA). N = 333 catalog units.

Source: USEPA STORET

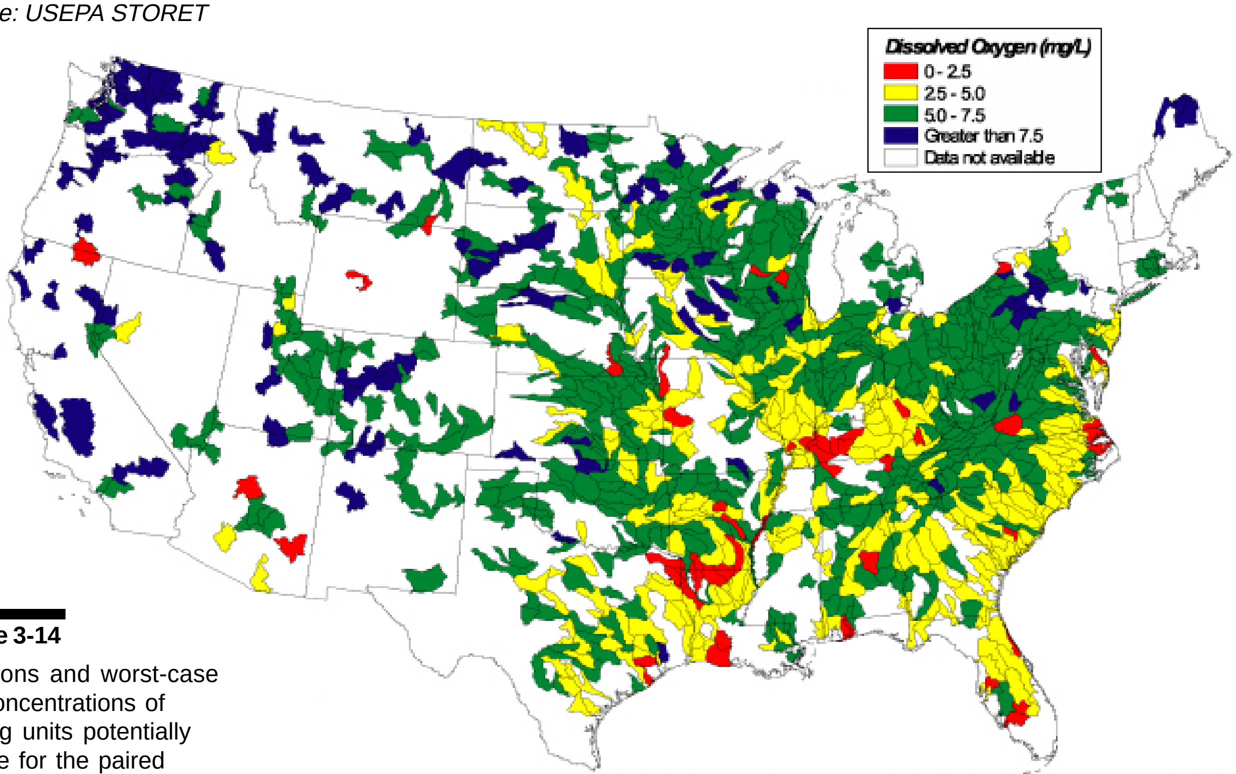


Figure 3-14

Locations and worst-case DO concentrations of catalog units potentially eligible for the paired analysis for the 1986-1990 time-block (after-CWA). N = 905 catalog units.

Source: USEPA STORET

Figure 3-15

Locations and change in worst-case DO concentrations of evaluated catalog units where (a) shows improving units (N = 167) and (b) shows degrading units (N = 79) before (1961-1965) versus after (1986-1990) the CWA.
Source: USEPA STORET

(a)

(b)

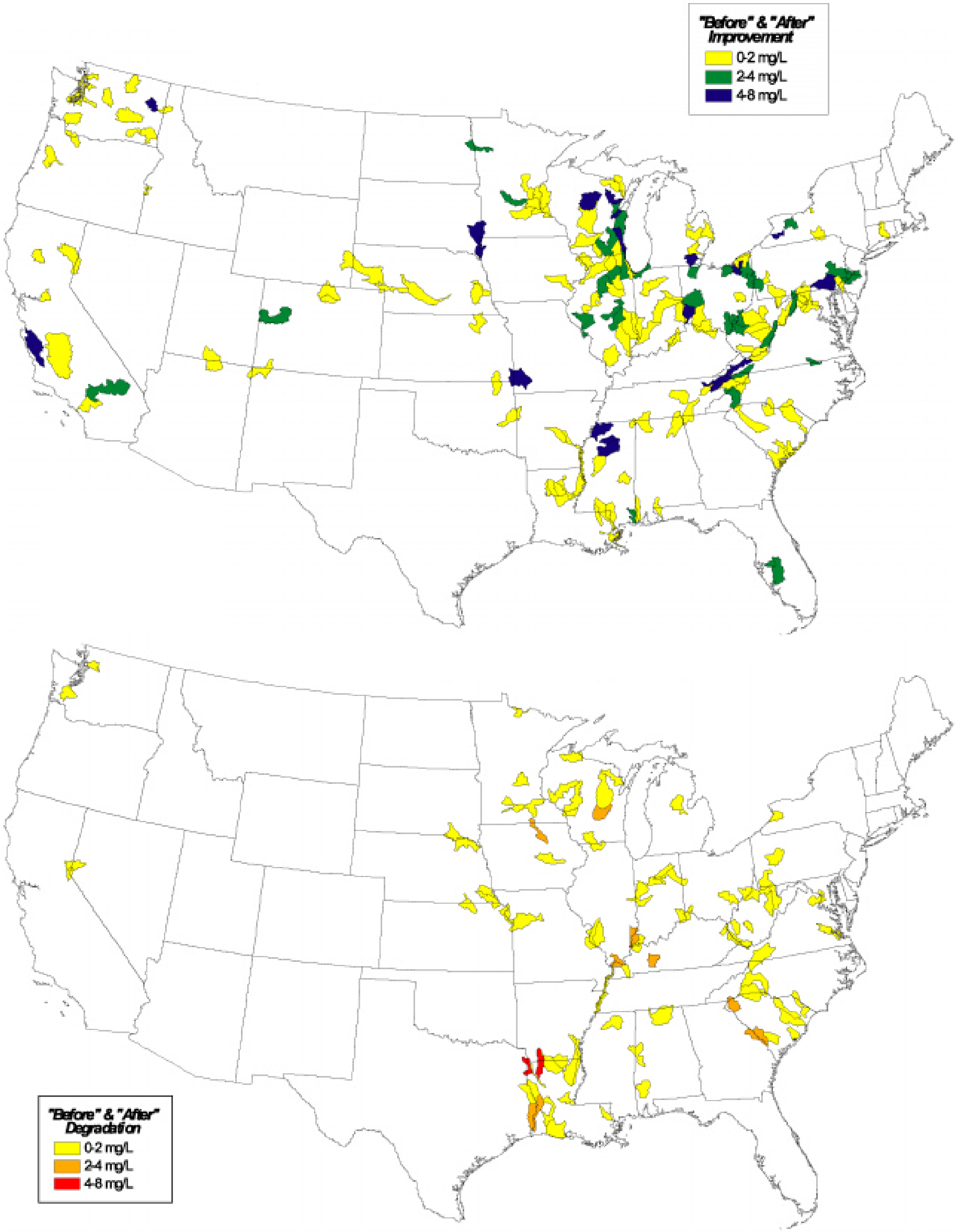
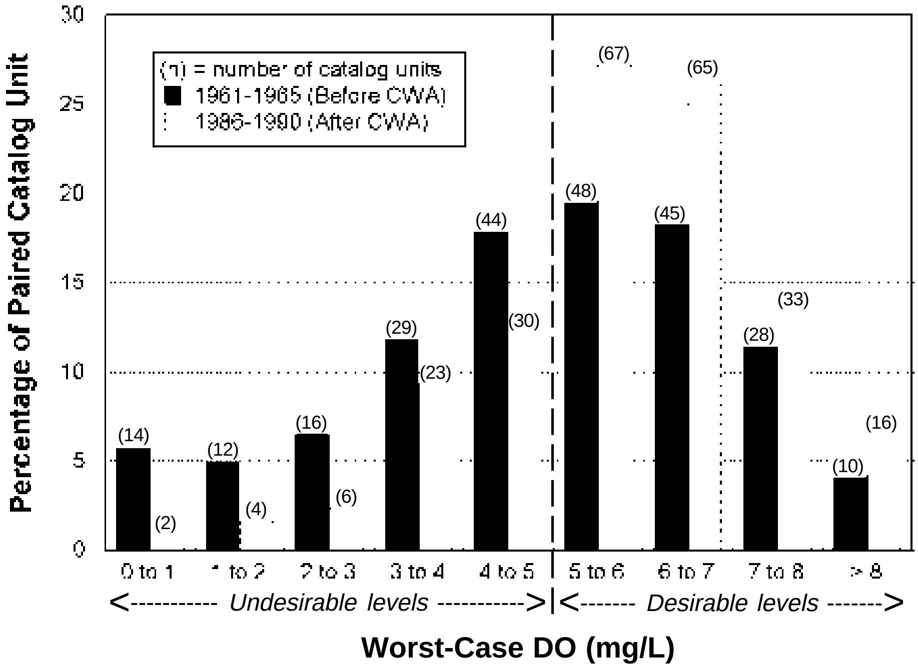


Figure 3-16

Frequency distribution comparing worst-case DO concentration of evaluated catalog units before and after the CWA. N = 246 catalog units.

Source: USEPA STORET



Key observations from Figures 3-15 and 3-16 include the following.

- 167 (68 percent) of the 246 evaluated catalog units had increases in worst-case DO after the CWA; 79 (32 percent) of the catalog units had decreases in worst-case DO after the CWA.
- The percentage of evaluated catalog units characterized by "very low" worst-case DO (< 2 mg/L) was reduced from 11 to 2 percent. Before the CWA, 26 catalog units had very low worst-case DO; after the CWA, only 6 catalog units had very low worst-case DO.
- The percentage of evaluated catalog units characterized by undesirable worst-case DO (below the 5 mg/L threshold) was reduced from 47 to 26 percent. Before the CWA, 115 catalog units had undesirable levels of worst-case DO; after the CWA, 65 catalog units had undesirable levels of worst-case DO.
- The percentage of evaluated catalog units characterized by desirable worst-case DO (above the 5 mg/L threshold) increased from 53 to 74 percent. Before the CWA, 131 catalog units had desirable levels of worst-case DO; after the CWA, 181 catalog units had desirable levels of worst-case DO.

Figure 3-17 is a column graph that describes the changes in worst-case DO that occurred after the CWA for the 246 evaluated catalog units in relation to the 5 mg/L threshold. Key observations from this figure include the following:

- 67 percent of the evaluated catalog units (166 out of 246 units) remained either above (47 percent) or below (20 percent) the 5 mg/L worst-case DO threshold.

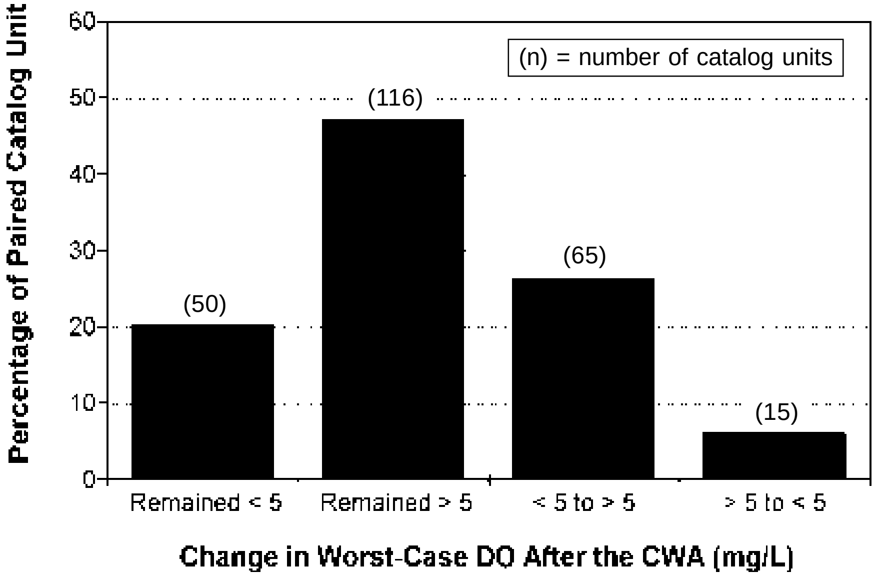


Figure 3-17

Frequency distribution of changes in worst-case DO levels after the CWA using 5 mg/L as the threshold value. N = 246 catalog units.

Source: USEPA STORET

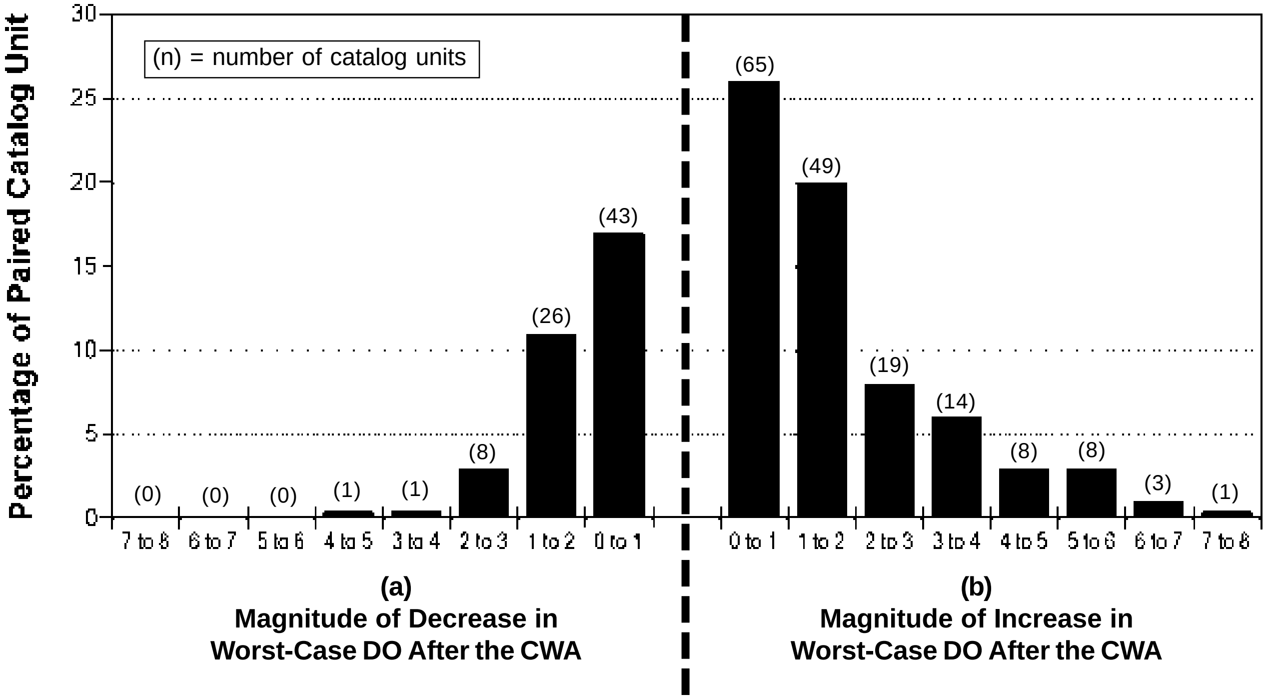
- Of the 115 catalog units that had worst-case DO concentrations below the threshold of 5 mg/L before the CWA, 57 percent (65 catalog units) increased to above the threshold after the CWA.
- Of the 131 catalog units that had worst-case DO concentrations above the benchmark threshold of 5 mg/L before the CWA, only 11 percent (15 catalog units) fell below the threshold after the CWA.

Of the 246 evaluated catalog units with paired before- and after-CWA data, 167 catalog units (68 percent) had increased worst-case DO and 79 (32 percent) had decreased worst-case DO after the CWA. Sections (a) and (b) of Figure 3-18 display the magnitude of degradation and improvement, respectively. Key observations from Figure 3-18 include the following:

- Approximately 32 percent of the evaluated catalog units that had increases in worst-case DO (53 of the 167 improving catalog units) increased by 2 mg/L or more.
- Approximately 13 percent of the evaluated catalog units that had decreases in worst-case DO (10 of 76 degrading catalog units) decreased by 2 mg/L or more.
- Approximately 44 percent of all evaluated catalog units either stayed the same or improved or degraded by 1 mg/L or less (108 of the 246 catalog units).

Figure 3-18

Frequency distribution of change in worst-case DO for evaluated catalog units that (a) decreased in concentration (n = 79) and (b) increased in concentration (n = 167) before and after the CWA. *Source: USEPA STORET*



Catalog Units with Greatest Improvements

Table 3-2 lists the 25 catalog units with the greatest before- and after-CWA improvements in worst-case DO. Figure 3-19 presents a location map of the top 10 of these units along with a stacked column graph that shows their before- and after-CWA worst-case DO concentration. Key observations from Table 3-2 and Figure 3-19 include the following:

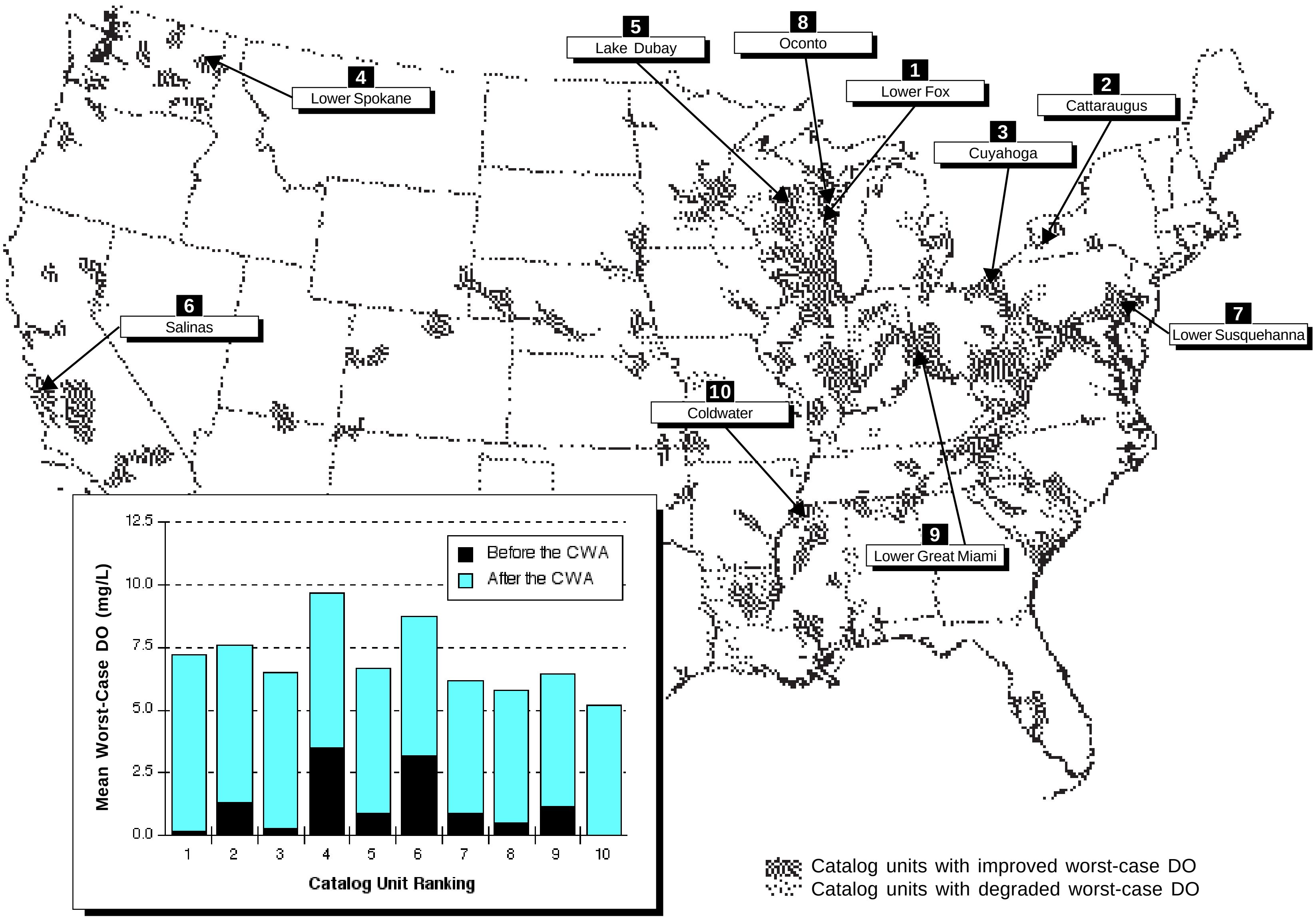
- All of the top 25 catalog units with the greatest before- and after-CWA improvements had before-CWA worst case DO levels below the benchmark threshold of 5 mg/L. Four catalog units had a before-CWA worst-case DO concentration of 0.0 mg/L.
- For 20 of the 25 catalog units, after-CWA worst-case DO improved to levels greater than 5 mg/L.
- The five catalog units that did not break the threshold value of 5 mg/L after the CWA all had concentrations of 0.6 mg/L or less in the before-CWA time-block.

Table 3-2. Twenty-five catalog units with greatest improvements in worst-case (mean 10th percentile) DO after the CWA.
Source: USEPA STORET

Rank	Reach ID	Catalog Unit Name	Worst-case DO 1961-65 *(mg/L)*	Worst-case DO 1986-90 *(mg/L)*	DO Change *(mg/L)*
1	04030204	Lower Fox, WI	0.1600	7.2050	7.0450
2	04120102	Cattaraugus, NY	1.3230	7.6000	6.2770
3	04110002	Cuyahoga, OH	0.2950	6.5008	6.2058
4	17010307	Lower Spokane, WA	3.5000	9.7000	6.2000
5	07070002	Lake Dubay, WI	0.8800	6.6833	5.8033
6	18060005	Salinas, CA	3.1800	8.7500	5.5700
7	02050306	Lower Susquehanna, MD	0.8800	6.1960	5.3160
8	04030104	Oconto, WI	0.5000	5.8000	5.3000
9	05080002	Lower Great Miami, IN	1.1850	6.4675	5.2825
10	08030204	Coldwater, MS	0.0000	5.2082	5.2082
11	10170203	Lower Big Sioux, IA	0.0000	5.1433	5.1433
12	04040002	Pike-Root, IL	0.9400	5.9400	5.0000
13	08030203	Yocona, MS	0.0000	4.8543	4.8543
14	04040003	Milwaukee, WI	2.1800	6.9567	4.7767
15	06010104	Holston, TN	0.1570	4.8686	4.7116
16	08030205	Yalobusha, MS	0.0000	4.6295	4.6295
17	06010205	Upper Clinch, TN	1.6140	6.0819	4.4679
18	02040204	Delaware Bay, NJ	0.5300	4.9100	4.3800
19	04100002	Raisin, MI/OH	4.0588	8.3400	4.2812
20	11070207	Spring, KS/MO	1.6000	5.6250	4.0250
21	04040001	Little Calumet-Galie	0.5700	4.5553	3.9853
22	18090208	Mojave, CA	4.0200	7.9767	3.9567
23	07120007	Lower Fox, IL	3.7800	7.5764	3.7964
24	07130011	Lower Illinois, IL	1.9400	5.7225	3.7825
25	04100009	Lower Maumee, OH	2.0676	5.8471	3.7795

Figure 3-19

Location map and distribution chart of the 10 catalog units with the greatest before versus after-CWA improvements in worst-case DO. *Source: USEPA STORET*



Comparison of the Change in Signal Between the Reach and Catalog Unit Scale Using the Upper White River Basin (Indiana) as an Example

Recall that the underlying objective of the second leg of the three-legged stool approach of this study was to measure the change in the response “signal” linking point source discharges with downstream water quality before and after the CWA at sequentially larger aggregations of spatial scales (reach, catalog unit, and major river basin). The theory is that if a signal change can be detected at sequentially larger scales, this would provide evidence that the CWA’s technology- and water quality-based effluent control requirements yielded broad as well as localized benefits (that is, stream reaches both within and beyond the immediate sag curve have benefited from the CWA).

The purpose of this subsection is to provide a practical comparison of reach and catalog unit signals using worst-case DO from monitoring stations in the Upper White River Basin (CU #05120201), the catalog unit in which the city of Indianapolis, Indiana, and several smaller municipalities reside.

Background

In the 1960s the citizens of the city of Indianapolis depended on primary treatment. Secondary treatment was added in the 1970s, and in 1983 the city further upgraded its POTWs to advanced wastewater treatment (AWT) to achieve compliance with water quality standards for DO. Two municipal facilities, designed to treat up to 379 cfs (245 mgd), currently discharge effluent to the White River. The base flow of the river is low; the 10-year, 7-day minimum (7Q10) flow is about 50 cfs in the channel upstream of the two POTWs. Consequently, under these low-flow conditions, Indianapolis’s wastewater effluent accounts for about 88 percent of the downstream flow.

In addition to Indianapolis, the 2,655-square-mile drainage area of the Upper White River Basin contains several smaller municipalities that also discharge municipal wastewater into the White River network. Population centers upstream from Indianapolis include Muncie, Anderson, and Noblesville. Waverly, Centerton, and Martinsville are towns located downstream of the city. Land use in the basin includes agricultural uses (65 percent) and urban-industrial uses (25 percent), with other uses accounting for the remaining 10 percent (Crawford and Wangness, 1991).

Using point and nonpoint source loading estimates of BOD_5 for contemporary conditions (16.3 metric tons/day ca. 1995) compiled for the NWPCAM (Bondelid et al., 1999), municipal loads in the basin are estimated to account for 50 percent of the total loading to basin waterways. The remaining one-half of the total BOD_5 load is contributed by major and minor industrial sources (11 percent), rural runoff (24 percent), urban runoff (13 percent), and CSOs (2 percent).

In a pre-AWT (1978-1980) and post-AWT (1983-1986) study of changes in water quality of the White River following completion of the upgrade to AWT from secondary activated sludge facilities for the city of Indianapolis, Crawford and Wangness (1991) concluded that there were statistically significant improvements in ambient levels of DO, BOD_5, and ammonia-nitrogen downstream of the upgraded municipal wastewater facilities. DO, in particular, improved by about 3 mg/L as a result of reductions in carbonaceous (BOD_5) and nitrogenous (ammonia) oxygen demands. For this study, Crawford and Wangness (1991) selected monitoring stations located about 10 and 15 miles downstream of Indianapolis's outfalls to collect data within the critical oxygen sag location of "degradation" and "active decomposition" (Waverly) and the "recovery" zone (Centerton) (see Figure 3-6).

During the before-CWA period from 1961 to 1965, streamflow conditions in the Upper White River Basin were characterized as dry, with persistent drought conditions for three consecutive summers from 1963 through 1965. During these three summers, streamflow ratios ranged from 40 to 63 percent of the long-term summer mean flow (see Figure 3-4(a) for 1963). Similarly, during the after-CWA period of 1986-1990, the Upper White River Basin was affected by the severe drought conditions of 1988 (streamflow ratio of only 34 percent of mean summer flow) that extended over large areas of the Midwest, Northeast, and upper Midwest (see Figure 3-4(b)). The hydrologic conditions of the White River are particularly critical in assessing before and after changes in DO because the municipal effluent flow of the upgraded AWT facilities (after 1983) accounted for about 88 percent of the river flow downstream of Indianapolis under low-flow conditions of the White River.

The Catalog-Level Signal

The analysis of before- and after-CWA worst-case DO data for the Upper White River catalog unit revealed that this catalog unit improved by 1.75 mg/L, from 3.80 mg/L (mean value of worst-case DO from 37 stations) before the CWA to 5.55 mg/L (mean value of worst-case DO from 14 stations) after the CWA. This level of improvement ranked it 64th out of the 246 catalog units with before and after data sets (see Appendix D). A companion examination of BOD_5 revealed that worst-case (90th percentile) loading in the catalog unit was reduced from 34.8 mg/L before the CWA (1961-1965) to 6.9 mg/L after the CWA (1986-1990).

The signal change detected provides evidence that

- The signal linking point source discharges with downstream water quality inherently resides in the before- and after-CWA worst-case DO data collected at stations throughout the Upper White River catalog unit.
- The signal is strong enough to be detected using a catalog unit scale summary statistic (mean of 10th percentile worst-case DO measurements for stations within the catalog unit).
- Improved wastewater treatment by the city of Indianapolis, as well as upgrades of wastewater treatment from small municipal facilities throughout the basin, resulted in broad water quality improvements in the Upper White River after the CWA.

The Reach-Level Signals

The POTW discharge/downstream water quality signal detected at the catalog unit scale is, in reality, a statistical aggregation of signals associated with all the monitored point source-influenced reaches in the Upper White River watershed. If one breaks the catalog unit down and examines the before- and after-CWA summary statistics for individual reaches, one would expect to find that the reaches in the "degradation" and "active decomposition" zones have more pronounced DO changes than reaches located outside those zones. An examination of reaches in the Upper White River catalog unit revealed this theory to be true. Figure 3-20 includes the locations and before- and after-CWA bar charts for each of the seven reaches in the Upper White River that have paired worst-case DO data. Figure 3-21 provides information regarding changes in worst-case (90th percentile) BOD_5 concentrations for the same reaches.

Key observations include the following:

- The reach with the greatest reduction of BOD_5 and greatest improvement in DO was the reach located immediately downstream of Indianapolis (05120201004) in the vicinity of Waverly. DO in this reach, which ranked sixth out of 311 reaches with before and after DO data nationwide (see Table 3-1), moved from 0.7 to 6.4 mg/L, an increase of 5.7 mg/L. In this same reach, the 90th percentile BOD_5 concentration declined from 58.1 mg/L to 4.3 mg/L.
- Reaches located immediately upstream of Indianapolis showed little change in before- and after-CWA DO conditions (Eagle Creek 05120201032; White River 05120201007, 05120201009; and Fall Creek 05120201006). BOD_5 , however, decreased from 20.6 to 7.0 mg/L in reach 05120201007 and from 12.4 to 3.0 mg/L in reach 05120201009. The decline in BOD_5 levels most likely reflects upgrades in municipal facilities for the small towns upstream of Indianapolis.
- Farther upstream, in the vicinity of Muncie and Anderson, greater improvements in DO were detected (along with decreasing trends in 90th percentile BOD_5 concentrations). In reach 05120201013 (Muncie), DO in the White River improved by 4.2 mg/L, from 2.2 mg/L before the CWA to 6.4 mg/L after the act. In the compilation of 311 reaches with the greatest before and after improvements in DO, this reach ranked 24th. For the reach in the vicinity of Anderson (05120201011), located downstream of Muncie, DO improved by 2.8 mg/L, from 3.4 mg/L to 6.2 mg/L. This reach ranked 44th in the nationwide ranking of stream reaches with DO improvements.
- The Lower White River catalog unit is located downstream from the Upper White River unit. Before and after station records from the most upstream reach of the basin reflect the impact of the wastewater discharges from the small towns of Centerton and Martinsville, as well as the recovery zone of the sag curve associated with the Indianapolis point source inputs. In this recovery reach of the White River (05120202031), DO improved by 1.9 mg/L, from 3.4 mg/L to 5.3 mg/L.

The aggregation of worst-case before- and after-CWA station records at the reach scale produced a variety of signals. As expected, the signal linking point source discharges with downstream water quality is most pronounced in reaches located immediately below point source discharges (in the critical portion of the sag zone). The signal became weaker farther downstream; however, in most reaches it was detectable, especially in the recovery zone of the sag curve associated with the Indianapolis discharges.

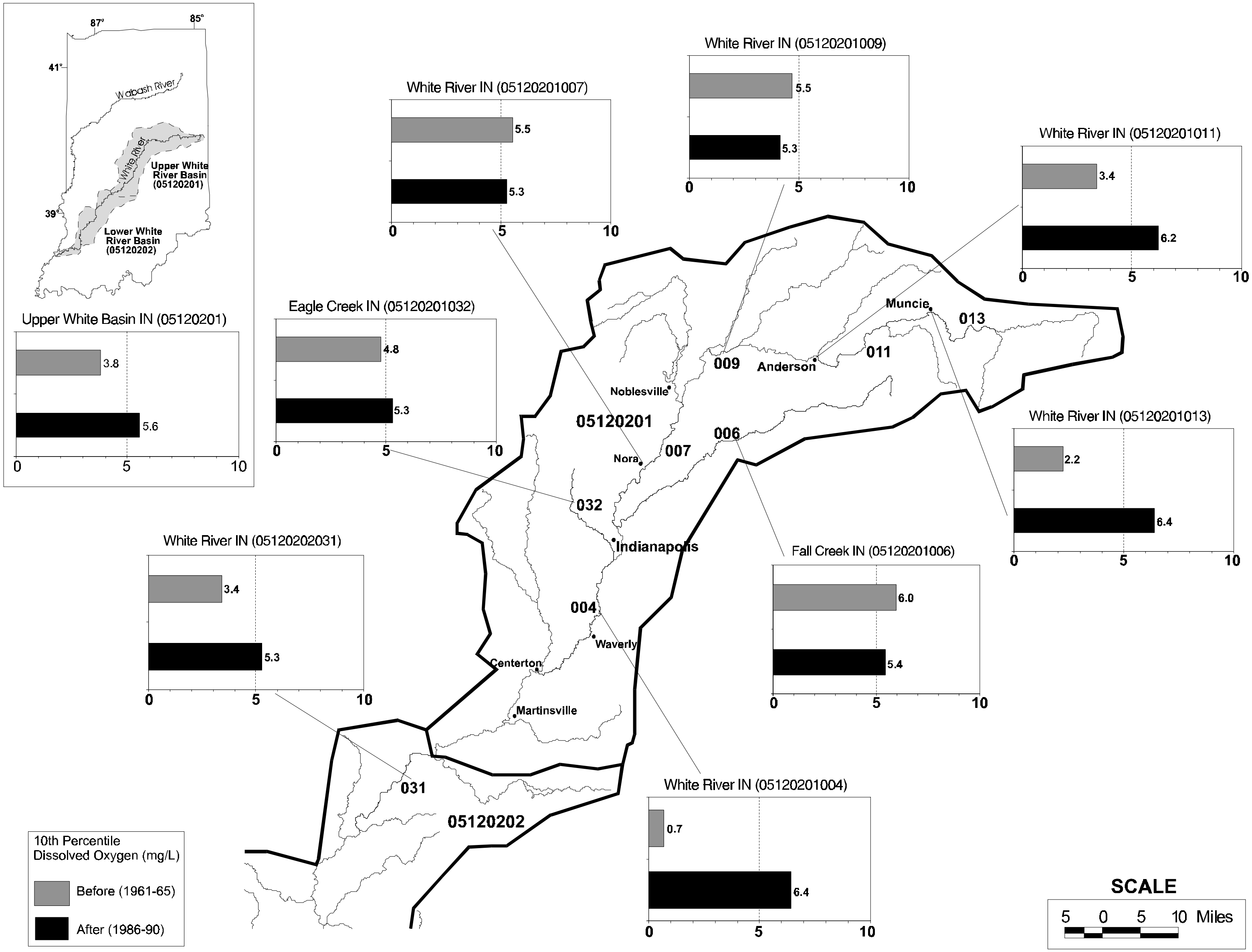


Figure 3-20. Before and after changes in worst-case DO (mg/L) for RF1 reaches of the Upper White River Basin (05120201) in Indiana.

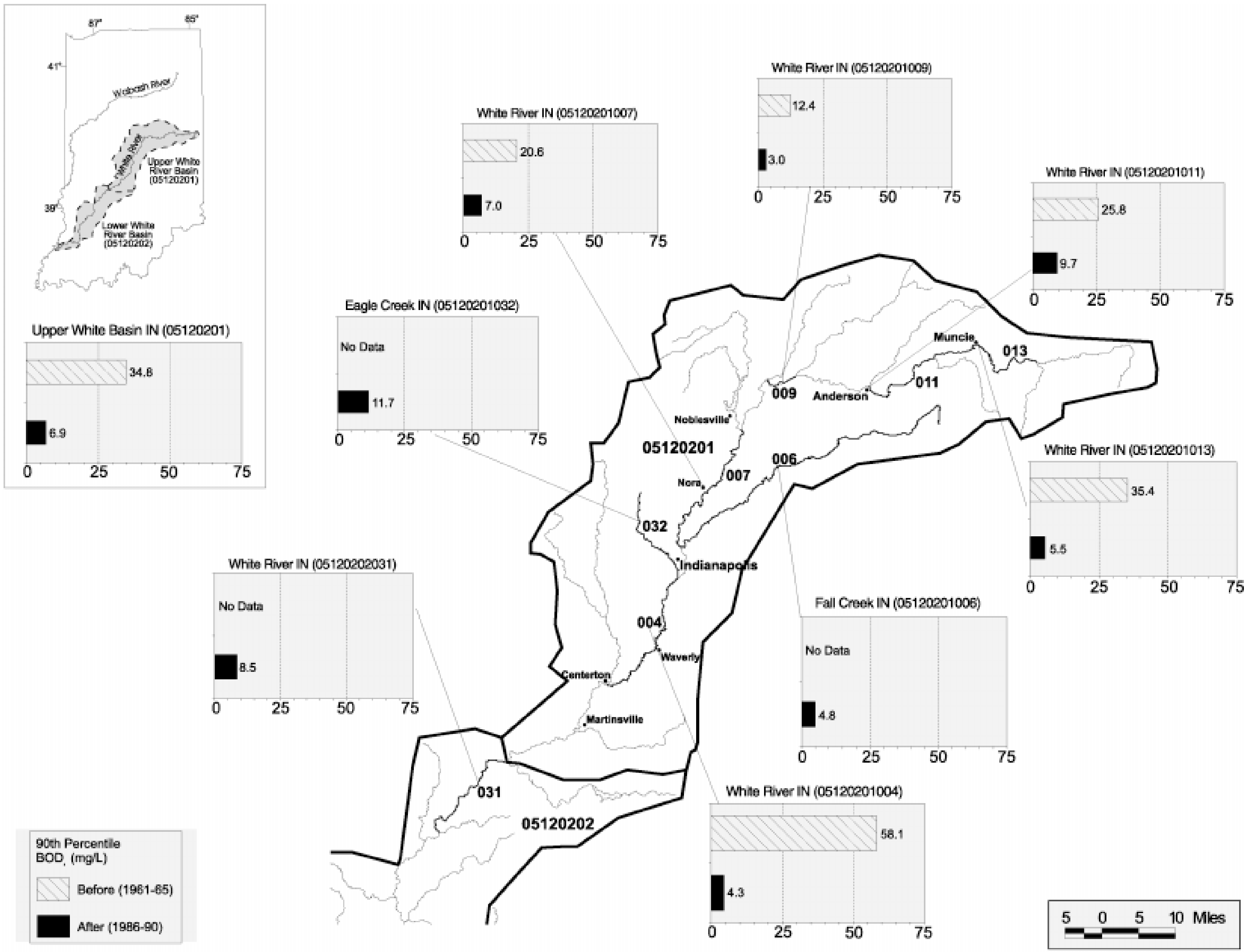


Figure 3-21. Before and after changes in 90th percentile BOD_5 (mg/L) for RF1 reaches of the Upper White River Basin (05120201) in Indiana.

Major River Basins

The stations comprising the 311 reach-aggregated worst-case DO data were pooled by the 18 major river basins of the contiguous United States for statistical analyses of the significance of changes in DO concentration before and after the CWA. These analyses were limited to the 311 evaluated reaches to improve the assurance that the data were collected from the same sample population.

Table 3-3 presents the number of observations, the results of the paired *t*-test (95 percent confidence level), and the mean of the pooled before and after worst-case DO data. The null hypothesis assumes that there is not a significant difference between the mean concentrations for the before and after periods. The means of the pooled worst-case DO data are presented as column graphs in Figure 3-22.

Table 3-3. Statistical significance of trends in mean 10th percentile (worst-case) DO by major river basin: before vs. after CWA (1961-1965 vs. 1986-1990). *Source: USEPA STORET.*

River Basin	No. of Paired Reaches	Paired *t*-test	Kolmogorov Smirnov test	Worst-case DO 1961-1965 *(mg/L)*	Worst-case DO 1986-1990 *(mg/L)*
All USA (01-18)	311	Yes	Yes	4.56	5.53
01 - New England Basin	1	*	*	4.30	6.90
02 - Middle Atlantic Basin	17	Yes	Yes	2.80	4.94
03 - South Atlantic-Gulf	61	Yes	Yes	4.10	4.73
04 - Great Lakes Basin	26	Yes	Yes	3.85	6.06
05 - Ohio River Basin	66	Yes	Yes	5.40	6.04
06 - Tennessee River Basin	19	Yes	No	4.08	5.23
07 - Upper Mississippi Basin	48	Yes	Yes	3.80	5.31
08 - Lower Mississippi Basin	25	No	No	3.79	3.94
09 - Souris-Red Rainy Basin	2	*	*	5.65	6.75
10 - Missouri River Basin	10	No	No	5.76	6.53
11 - Arkansas-Red—White Basin	7	No	No	5.36	4.60
12 - Texas-Gulf Basin	2	*	*	5.77	4.37
13 - Rio Grande Basin	0	*	*	--	--
14 - Upper Colorado River Basin	1	*	*	4.88	7.22
15 - Lower Colorado River Basin	0	*	*	--	--
16 - Great Basin	2	*	*	7.45	6.10
17 - Pacific Northwest Basin	17	Yes	No	7.61	8.21
18 - California Basin	7	Yes	Yes	5.61	7.58

Paired *t*-test: 95% confidence - 2-sided test
Kolmogorov Smirnov test: 90% confidence, 2-sided test
*insufficient data for analysis

Figure 3-22

Before vs. after trends in worst-case DO for major river basins: 1961-1965 vs. 1986-1990.

Source: USEPA STORET

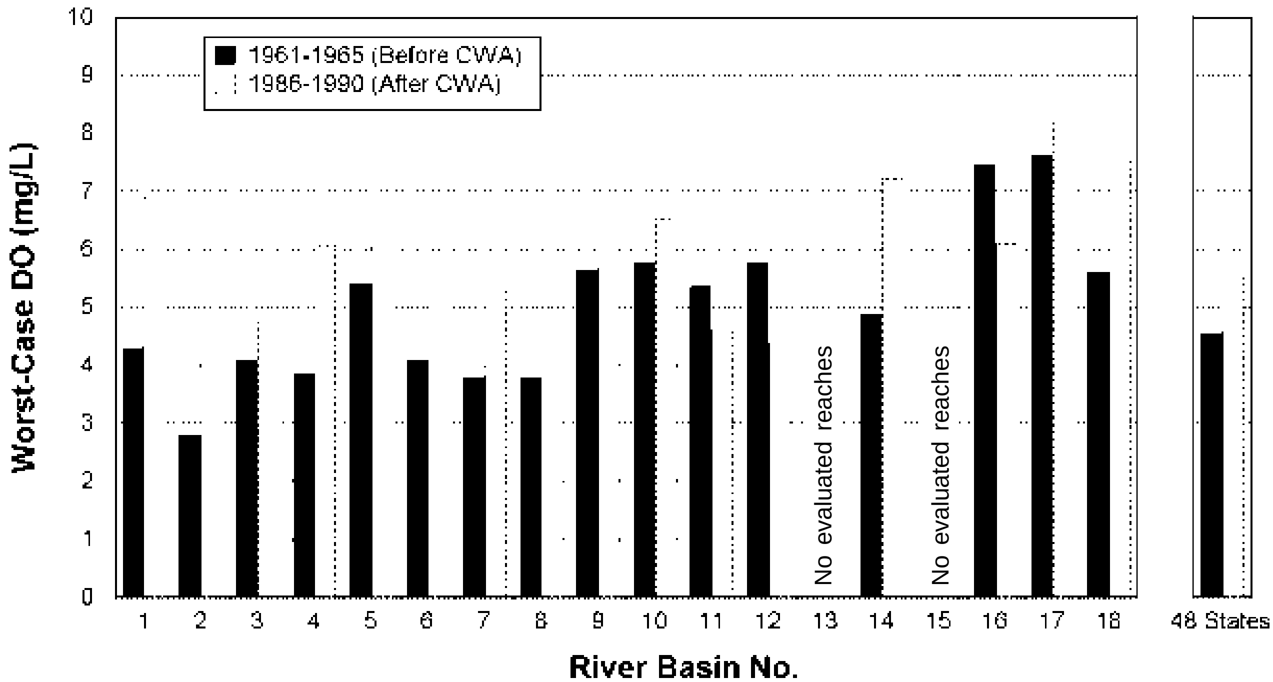


Figure 3-23 maps the results of the paired *t*-test. The darker shaded (yes) river basins indicate that there is a statistically significant difference at the 95 percent confidence level; the river basins marked with lighter shading (no) indicate that there is not a statistically significant difference between the means. Discounting the river basins, mostly in arid western states, with insufficient data for the paired *t*-test (river basins 01, 09, 12, 13, 14, 15 and 16), 8 of the 11 river basins in the Midwest, Southeast, west coast, and middle Atlantic states showed a statistically significant improvement in DO using the paired *t*-test. The visual decreases in DO in the Texas-Gulf (12), Arkansas-Red-White (11), and Great Basins (16) are not statistically significant.

Recalling that the planning and design of wastewater treatment plant upgrades are often targeted at improving worst-case (low) DO conditions, it is expected that incremental improvements for waters with higher DO conditions (e.g., approaching saturation levels of about 8 to 10 mg/L) are less likely to accrue. As a result, it was suspected that most of the gains would be for the river basins with the lowest DO concentrations before the upgrades, with fewer gains identified for basins that had not been characterized by low DO concentrations. Therefore, frequency distributions are compared in addition to the comparison of means described above.

Figure 3-23

Statistical significance of the difference between before- and after-CWA worst-case DO mean values for the 18 major river basins in the 48 contiguous states. *Source: USEPA STORET*

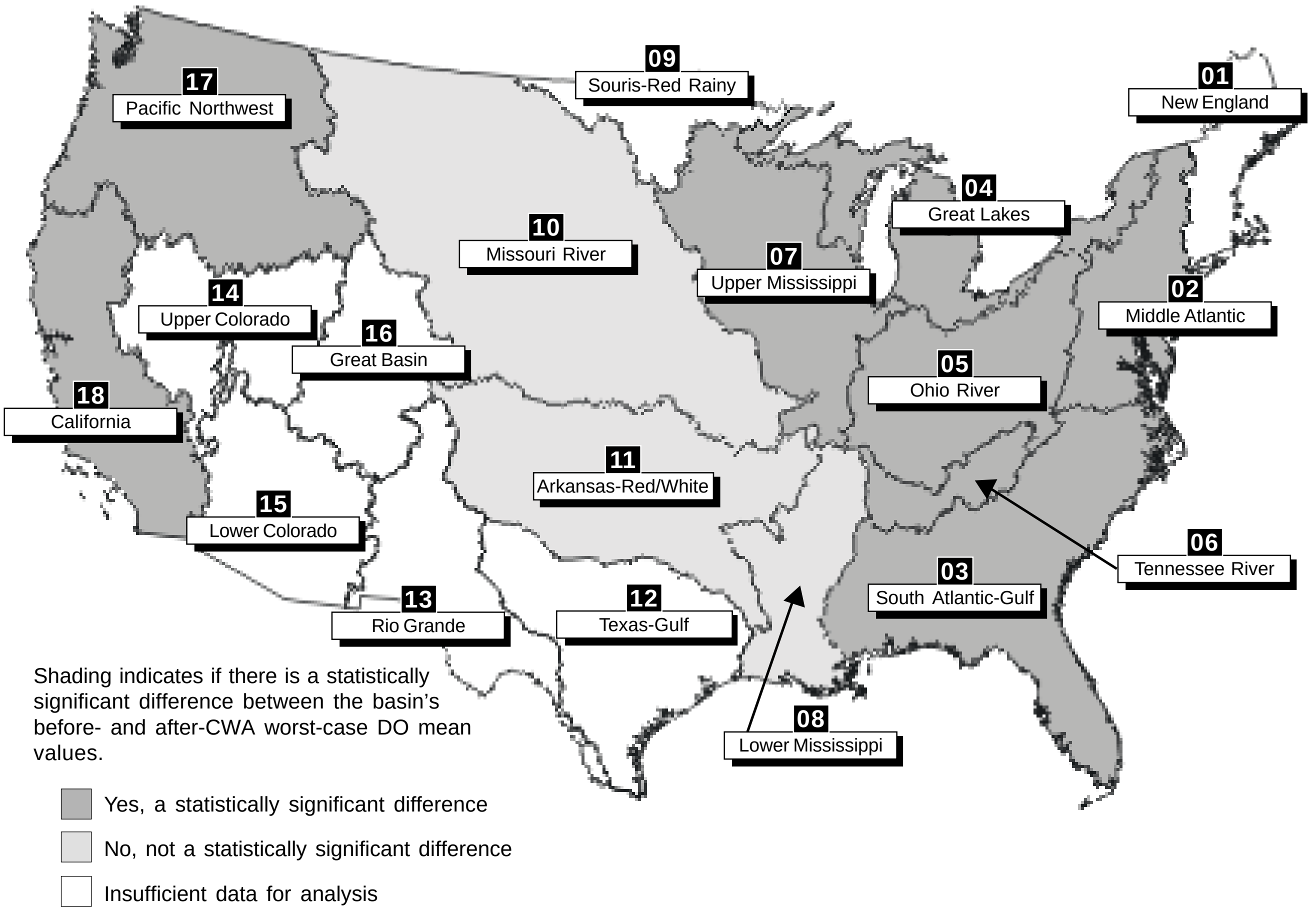
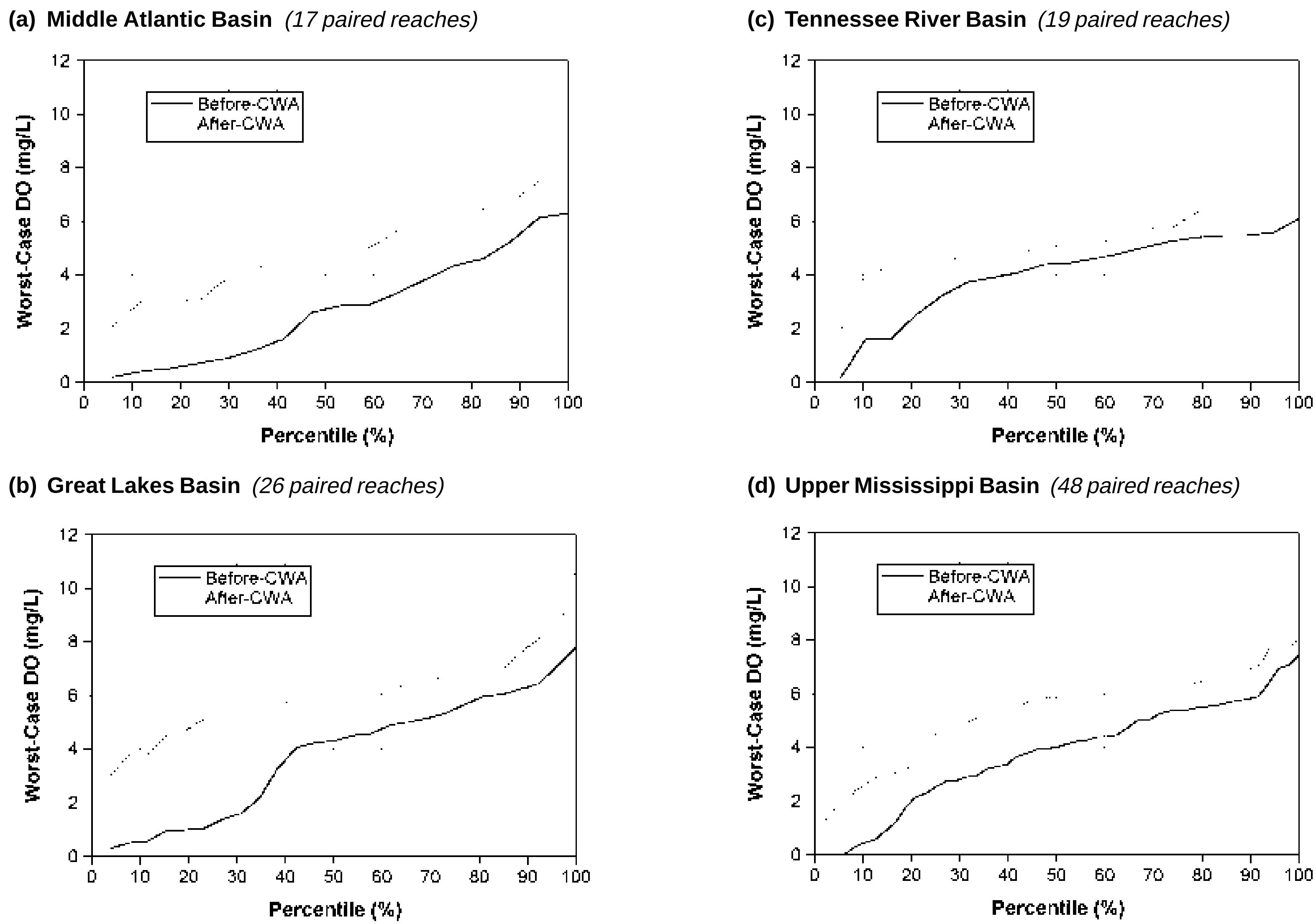


Figure 3-24 presents the before and after worst-case DO frequency distributions for the mid-Atlantic, Great Lakes, Tennessee, and Upper Mississippi major river basins. It is important to note that not only has the mean changed, but the distribution has also changed. The frequency distributions shown in the figure suggest that there have been improvements at the lower percentile levels of DO (the 10th and 20th percentiles) for these river basins. Before the CWA in the 1961-1965 time block, worst-case DO was at 1 mg/L or lower. After the act, worst-case conditions had improved to levels of about 3 to 5 mg/L.

The Kolmogorov-Smirnov test was used to statistically compare whether the before and after distributions are significantly different. The Kolmogorov-Smirnov test is a goodness of fit test that compares the empirical distributions from the two time periods. Figure 3-24, showing the empirical cumulative distribution functions of DO from the before and after periods, can be used to visualize what the Kolmogorov-Smirnov test is comparing on a statistical basis. The vertical axis presents the DO concentration corresponding to a given percentile on the horizontal axis. Referring to the mid-Atlantic basin, for example, it can be seen that about

Figure 3-24

Before- and after-CWA frequency distributions of worst-case DO aggregated by major river basin for reaches with paired before and after data sets: (a) Middle Atlantic, (b) Great Lakes, (c) Tennessee River, and (d) Upper Mississippi River basins. *Source: USEPA STORET.*



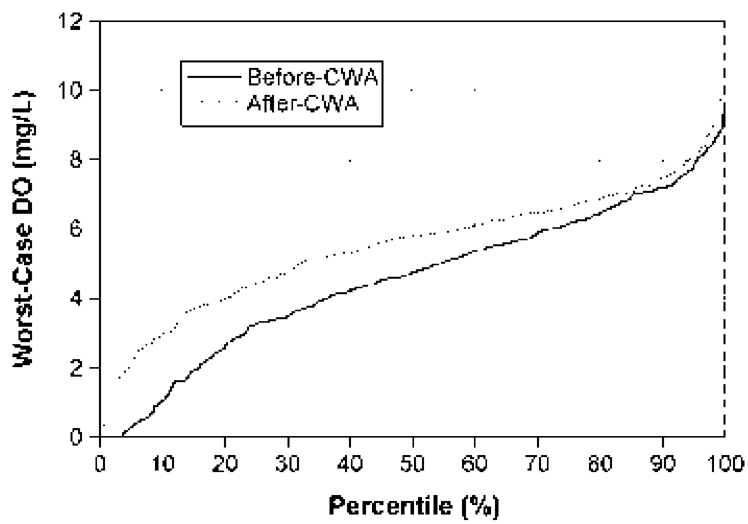
70 percent of the observations from the before period were less than 4 mg/L, whereas in the after period only 30 percent of the observations were less than 4 mg/L. The Kolmogorov-Smirnov test is a statistical comparison of the maximum distance between these curves. The results from the Kolmogorov-Smirnov test are provided in Table 3-3.

Based on the two different statistical tests, and discounting the 7 river basins with limited data, 8 of the 11 remaining river basins can be characterized by a statistically significant improvement in worst-case DO using at least one of the two tests. Mixed results (yes and no) were obtained for two basins with the Kolmogorov-Smirnov test indicating no significant improvement for the Tennessee (6) and the Pacific Northwest (17) basins, whereas the paired t-test indicated significant improvements (yes) in these basins. Overall, there is a statistically significant improvement in worst-case DO trends using both statistical tests at 6 out of 11 river basins with sufficient data. Of the five basins with at least one “nonsignificant” change, three basins (Missouri River, Arkansas Red-White, and Pacific Northwest) had a mean worst-case pooled DO level greater than 5 mg/L in the before time period and were less likely to be targeted for improved point source pollution control. It is also noteworthy that in the 25-year interval between the before- and after-CWA periods, there were no statistically significant conditions of degradation of worst-case DO for any of the major river basins. It is also noteworthy that when all 311 paired reaches are analyzed together, both tests indicate significant increases in worst-case DO (see Figure 3-25 and top row (All USA) of Table 3-3).

Figure 3-25

Before- and after-CWA frequency distributions of worst-case DO aggregated over all major river basins for the 311 reaches with paired before and after data sets.

Source: USEPA STORET.



D. Summary and Conclusions

The purpose of this chapter is to address the second leg of the three-legged stool approach for answering the question posed in Chapter 1—*How has the Nation's water quality changed since implementation of the 1972 CWA's mandate for secondary treatment as the minimum acceptable technology for POTWs?* Recall that the basic goal of the second leg was to determine the extent to which water quality improvements could be linked to the CWA's push for secondary and greater levels of treatment in the Nation's POTWs. If evidence showed that worst-case DO concentrations improved at broad, as well as localized spatial scales, the second leg of the investigation could add cumulative support for the conclusion that the CWA's mandates were successful. The following objectives were established to guide this part of the study:

- Develop before- and after-CWA data sets composed of DO summary statistics derived from monitoring stations screened for worst-case conditions.
- Develop a worst-case DO summary statistic for each station for each before- and after-CWA time period and then aggregate these data by sequentially larger spatial scales (reaches, catalog units, and major river basins).
- Conduct an analysis of the spatial units having both a before- and after-CWA summary statistic and assess the magnitude of worst-case DO change between the two time periods.
- Assess the change in the point source discharge/downstream DO signal over the progressively larger spatial scales.

Key Points of the Background Section

Section A provided background concerning the source of DO data used in this study, why worst-case conditions are an appropriate screening tool for developing the before- and after-CWA data sets, and the role spatial scale played in the second leg of this study. Key points include the following:

- The sharpest signal linking point source loading and downstream DO inherently resides in data collected in worst-case (high temperature and low flow) conditions. These worst-case conditions typically occur in the summer months (July through September) during consecutive runs of dry years (persistent drought).
- Widespread persistent drought was most pronounced in the summers in 1961-1965 (before the CWA) and 1986-1990 (after the CWA). These time-blocks were used to define the before- and after-CWA time periods for the comparison analysis.
- From a spatial perspective, worst-case critical, or minimum, DO below a point source occurs in the "degradation" or "active decomposition" zone of the oxygen sag curve. However, screening rules were not developed to select monitoring stations located within these zones because the goal of this second leg is to examine changes in the point

source discharge/downstream DO at broad scales as well as localized scales. Consequently, the only screening rule regarding location of stations eligible for the before- and after-CWA analysis is that the station must be somewhere downstream and therefore potentially influenced by a point source.

Key Points of the Data Mining Section

Section B presented the six-step data mining process used to create the before- and after-CWA data sets to be used in the comparison analysis. The screening rules associated with each step are listed below:

Step 1—Data Selection Rules

- DO, expressed as a concentration (mg/L), will function as the signal relating point source discharges to downstream water quality responses.
- DO data are extracted only from the July-September (summer season) time period.
- Only surface DO data (DO data collected within 2 meters of the water surface) are used.

Step 2—Data Aggregation Rules From a Temporal Perspective

- 1961-1965 serves as the time-block to represent persistent drought before the CWA and 1986-1990 serves as the time-block to represent persistent drought after the CWA.
- To remain eligible for the before- and after-CWA comparison, DO data must come from a station residing in a catalog unit that had at least one year classified as dry (streamflow ratio 75 percent of summer mean) out of the 5 years in each before- and after-CWA time-block.

Step 3—Calculation of the Worst-case DO Summary Statistic Rules

- For each water quality station, the 10th percentile of the DO data distribution from the before-CWA time period (July through September, 1961-1965) and the 10th percentile of the DO data distribution from the after-CWA time period (July through September, 1985-1990) are used as the station's DO worst-case statistics for the comparison analysis.
- To remain eligible for the before- and after-CWA statistical comparison, a station must have a minimum of eight DO measurements within each of the 5-year time-blocks.

Step 4—Spatial Assessment Rules

- Only water quality stations on streams and rivers affected by point sources are included in the before- and after-CWA comparison analysis.

Step 5—Data Aggregation Rules From a Spatial Perspective

- The before- and after-CWA data sets are collections of DO summary statistics that characterize worst-case DO at individual water quality monitoring stations across the United States for the 1961-1965 time-block and the 1986-1990 time-block, respectively (one DO summary statistic per station per time-block).
- For each data set and time-block, the 10th percentile value from each eligible station is aggregated within the spatial hydrologic unit. (Since the scales are hierarchical, a station's summary statistic is effectively assigned to a reach *and* a catalog unit.) A summary statistic is then calculated and assigned to the spatial unit for the purpose of characterizing its worst-case DO. If a spatial unit has only one monitoring station within its borders that meets the screening criteria, the 10th percentile DO value from that station simply serves as the unit's worst-case summary statistic. If, however, there are two or more stations within a spatial unit's borders, the 10th percentile values for all the eligible stations are averaged and this value is used to characterize worst-case DO for the unit.
- The mean 10th percentile value is computed from the eligible station's 10th percentile values for the before- and after-CWA periods.

Step 6—Development of the Paired Data Sets (at each spatial scale)

- To be eligible for the paired comparison analysis, a hydrologic unit must have both a before-CWA and an after-CWA summary statistic assigned to it.

Key Points of the Comparison Analysis Section

Section C presented the results of the comparative before- and after-CWA analysis of worst-case DO data derived using the screening criteria described in *Section B* and aggregated by spatial units defined by three scales (reach, catalog unit, and major river basin). Listed below are key observations for each spatial scale.

Reach Scale

- *Sixty-nine percent of the reaches evaluated showed improvements in worst-case DO after the CWA.* [Three hundred eleven reaches (out of a possible 12,476 downstream of point sources) survived the data screening process with comparable before- and after-CWA DO summary statistics. The number of reaches available for the paired analysis was limited by the historical data for the 1961-1965 period].
- These 311 evaluated reaches represent a disproportionately high amount of urban/industrial population centers, with approximately 13.7 million people represented (7.2 percent of the total population served by POTWs in 1996). *The top 25 improving reaches saw their worst-case DO increase by anywhere from 4.1 to 7.2 mg/L!*

- The number of evaluated reaches characterized by worst-case DO below 5 mg/L was reduced from 167 to 97 (from 54 to 31 percent).
- The number of evaluated reaches characterized by worst-case DO above 5 mg/L increased from 144 to 214 (from 46 to 69 percent).

Catalog Unit Scale

- *Sixty-eight percent of the catalog units evaluated showed improvements in worst-case DO after the CWA.* [Two hundred forty-six catalog units (out of a possible 1,666 downstream of point sources) survived the data screening process with comparable before- and after-CWA DO summary statistics].
- The number of evaluated catalog units characterized by worst-case DO below 5 mg/L was reduced from 115 to 65 (from 47 to 26 percent). The number of evaluated catalog units characterized by worst-case DO above 5 mg/L increased from 131 to 181 (from 53 to 74 percent).
- Fifty-three of the 167 improving catalog units (32 percent) improved by 2 mg/L or more while only 10 of 79 degrading catalog units (13 percent) degraded by 2 mg/L or more.
- These 246 evaluated catalog units represent a disproportionately high amount of urban/industrial population centers, with approximately 61.6 million people represented (32.5 percent of the total population served by POTWs in 1996).

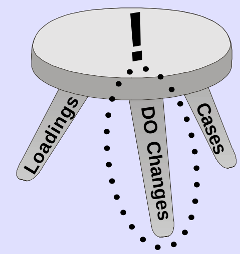
Major River Basin Scale

- A total of 11 out of 18 major river basins had sufficient reach-aggregated worst-case DO data for a before- and after-CWA comparison analysis.
- Based on two statistical tests, 8 of the 11 major river basins can be characterized as having statistically significant improvement in worst-case DO levels after the CWA. The three basins that did not statistically improve under either test also did not have statistically significant degradation.
- When all the 311 paired (i.e., before vs. after) reaches were aggregated and the statistical tests run on all 18 of the major river basins of the contiguous states as a whole, worst-case DO also showed significant improvement.

Conclusions

The statistical analyses developed for this study are not ideal. One major concern is the potential bias introduced in the ambient monitoring programs used to collect the data archived in STORET. It is believed that the analysis of data sets with data in the before and after time periods alleviates some of these concerns and that results are generally comparable for the two different statistical tests. Based on the systematic, peer-reviewed approach designed to identify and evaluate the national-scale distribution of water quality changes that have occurred since the 1960s, this study has compiled strong evidence that the technology- and water quality-based policies of the CWA for point source effluent controls have been effective in significantly improving DO. In this retrospective analysis, DO was used as the key indicator because the reduction of organic carbon and nitrogen (BOD_u) loading from municipal and industrial point sources was one of the major goals of the CWA's technology-based policy, which required industrial effluent limits and a minimum level of secondary treatment for municipal facilities. Based on ambient DO records, significant before and after improvements in many rivers and streams have been identified over national, major river basin, catalog unit, and reach-level spatial scales.

Conclusion of the second leg of the stool



There were significant after-CWA improvements in worst-case summer DO conditions in two-thirds of the hydrologic units at all three spatial scales!

The "signal" of downstream water quality responses to upstream wastewater loading and the changes in this signal since the 1960s has been successfully decoded from the "noise" of millions of archived water quality records. Given the very large spatial scale of the major river basins, it is remarkable to observe statistically significant before and after DO improvements as detected using the systematic methodology described herein. Previous evaluations of the effectiveness of the CWA (e.g., Smith et al., 1987a, 1987b, 1992; Knopman and Smith, 1993) were not able to report conclusively significant improvements in DO. In these earlier studies, however, the methodologies used were not specifically designed to separate the signal of downstream water quality response from the noise within large national databases. Using appropriate data screening rules and spatial aggregations, it has been demonstrated that improvements in water quality, as measured by improvements in worst-case DO, have been achieved since the 1960s.

The findings of this national-scale water quality assessment demonstrate three important points:

- As new monitoring data are collected, it is crucial for the success of future performance measure evaluations of pollution control policies that the data be submitted, with appropriate QA/QC safeguards, to accessible databases. If the millions of records archived in STORET had not been readily accessible, it would have been impossible to conduct this analysis to identify the signals of water quality improvements that have been achieved over the past quarter-century.
- Significant after-CWA improvements in worst-case summer DO conditions have been quantitatively documented with credible statistical techniques in this study over different levels of spatial data aggregation from the small subwatersheds of Reach File Version 1 river reaches (mean drainage area of approximately 115 mi^2) to the very large watersheds of major river basins (mean area of 434,759 mi^2).

- The data mining and statistical methodologies designed for this study can potentially be used to detect long-term trends in signals for water quality parameters other than DO (e.g., suspended solids, nutrients, toxic chemicals, pathogens) to develop new performance measures to track the effectiveness of watershed-based point source and nonpoint source controls. The key element needed to apply the data mining methodology to other water quality parameters is the careful specification of rules for data extraction that reflect a thorough understanding of the various processes that influence the spatial and temporal distributions of a water quality constituent, as well as the relevant sources of associated pollutants.

Population Affected by Reaches With Improved DO

To monetize environmental benefits derived from various environmental policy decisions, USEPA developed the NWPCAM model (Bondelid et al., 1999), which includes a link between 1990 population and RF1 reaches. As discussed in *Section E,* this model does not include all estuarine and coastal waters, and as a result, does not account for the entire US population. It is estimated that about one-third of the U.S. population is not accounted for in the model. At the same time if a person is located near two rivers, that person is counted twice since he or she can derive a benefit from environmental improvements in either river.

Recognizing this accounting procedure, the model accounts for 197.7 million people in 23,821 reaches. In the 311 reaches analyzed here (1.3 percent of reaches in the model), the model accounts for 13.7 million people (6.9 percent of the population in the model). The ratio of the percent population to percent reaches in the model demonstrates that the screening process developed for this analysis is reasonably successful in finding reaches with data near urban centers, although 57 of the 311 reaches did not have population associated with them. Of the 13.7 million people represented by the 311 reaches, 11.8 million of them (86 percent) are associated with reaches that have an increased worst-case DO from before to after the CWA. Almost one-half (45 percent) of the selected population are associated with reaches that went from worst-case DO below 5 mg/L before the CWA to greater than 5 mg/L after the CWA. Although it is unfortunate that more reaches are not considered in the current analysis (mainly because of limitations in available monitoring data for the before-CWA periods), it is helpful to consider that the corresponding 246 catalog units include 61.6 million (31.2 percent) of the 197.7 million people accounted for in the model. And three-fourths (46.5 million) of the 61.6 million people are in catalog units that had an increase in worst-case DO between the before to after time period.

Sensitivity to Using DO versus Percent Saturation

The beginning of this chapter describes the physical processes associated with atmospheric reaeration, oxygen demand, and dilution, as well as the impact of changing water temperatures and elevation. During the initial development of the screening methodology, considerable effort was spent evaluating various indicators for water quality. Ultimately, DO was selected. Another strong candidate was DO expressed as percent saturation. Use of percent saturation would effectively normalize the DO data to account for geographic differences in elevation, chlorides, and water temperature. Saturation levels of DO decrease with higher elevations, increasing chloride content, and warmer water temperatures (Chapra, 1997). Correcting for elevation would improve spatial comparisons such as those in Figures 3-13 through 3-15, and correcting for chlorides and water temperature would account for some of the unexplained variability that might exist between the before and after time periods.

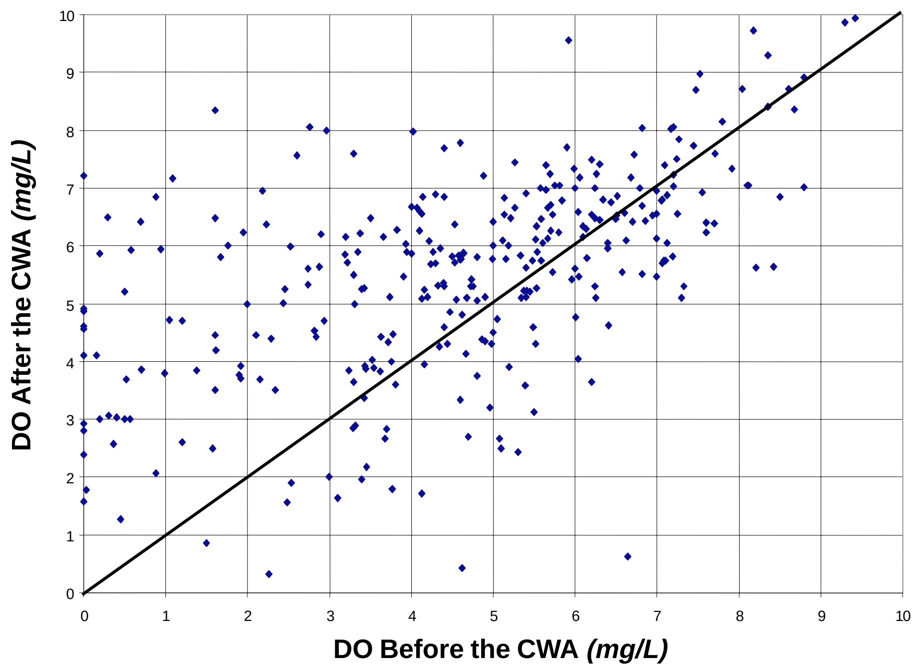
To evaluate the impact that selection of DO over DO as percent saturation might have on the analysis, two scatter plots with data aggregated to the reach level were compared. Figure 3-26(a) presents the DO after the CWA as a function of DO before the CWA. Figure 3-26(b) presents the DO (percent saturation) after the CWA as a function of DO (percent saturation) before the CWA aggregated to the reach level. The values for DO (percent saturation) were computed using the same procedure used for DO. Points above the diagonal line in either figure indicate that the DO or DO (percent saturation) increased. Although the two figures use different scales, a visual comparison suggests that there would be little difference if DO (percent saturation) were adopted over DO. Given that the public has a more intuitive understanding of DO measured as concentration, the analysis in this chapter uses DO concentration rather than percent saturation.

Figure 3-26

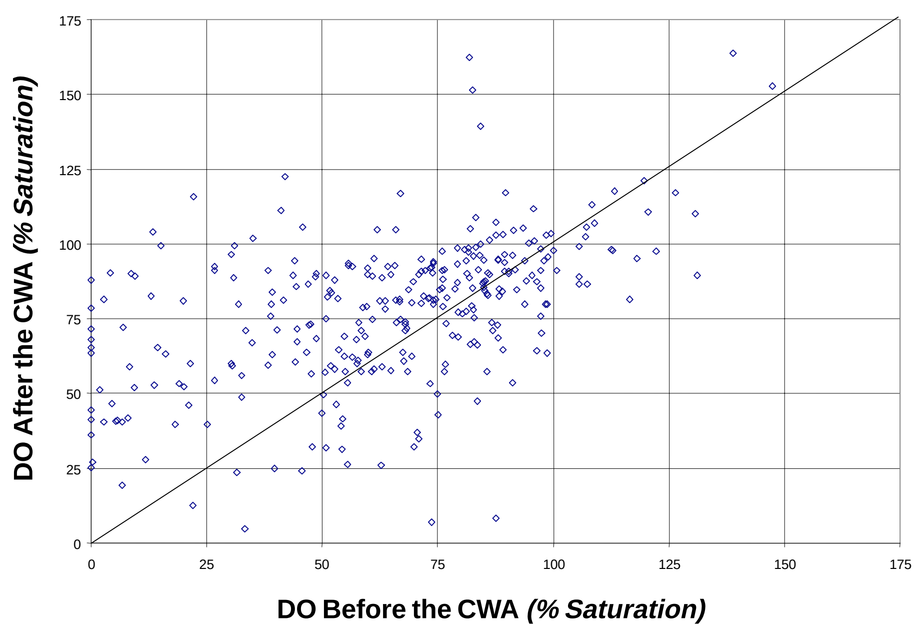
(a) Comparison of the 10th percentile DO before the CWA as a function of the 10th percentile DO after the CWA. (b) comparison of the 10th percentile DO (percent saturation) before the CWA as a function of the 10th percentile DO (percent saturation) after the CWA.

Source: USEPA STORET.

(a)



(b)



References

Ackerman, W.C., Harmeson, R.H., and Sinclair, R.A. 1970. Some long-term trends in water quality of rivers and lakes. Transactions of the American Geophysical Union, vol. 51, page 516-522.

Bondelid, T., C. Griffiths, and G. van Houten. 1999. *A national water pollution control assessment model.* Draft tech. report prepared by Research Triangle Park, NC for U.S. Environmental Protection Agency, Office of Science and Technology, Washington, DC.

Chapra, S.C. 1997. *Surface Water Quality Modeling*. McGraw Hill, Inc. New York, NY.

Crawford, C.G., and D.J. Wangness. 1991. Effects of advanced wastewater treatment on the quality of White River, Indiana. *Water Resources Bulletin* 27(5): 769-779.

Gunnerson, C.G. 1966. An atlas of water pollution surveillance in the United States, Otober 1, 1957 to September 30, 1965. Water Quality Activities, Division of Pollution Surveillance, Federal Water Pollution Control Adminstration, U.S. Department of the Interior, Cincinnati, OH.

Knopman, D.S. and R.A. Smith. 1993. Twenty years of the Clean Water Act: Has U.S. water quality improved? *Environment* 35(1):17-41.

Smith, R.A., R.B. Alexander, and M.G. Wolman. 1987a. *Analysis and interpretation of water quality trends in major U.S. rivers, 1974-81.* Water-Supply Paper 2307, U.S. Geological Survey, Reston, VA.

Smith, R.A., R.B. Alexander, and M.G. Wolman. 1987b. Water quality trends in the Nation’s rivers. *Science* 235 (27 March 1987):1607-1615.

Thomann, R.V., and J.A. Mueller. 1987. *Principles of surface water quality modeling and control.* Harper & Row, New York, NY.

USEPA. *STOrage and RETrieval Water Quality Information System.* U.S. Environmental Protection Agency, Office of Wetlands, Oceans, and Watersheds, Washington, DC.

USEPA. 1974. *National water quality inventory, 1974.* EPA-440/9-74-001. U.S. Environmental Protection Agency, Office of Water Planning and Standards, Washington, DC.

Wolman, A. 1971. The Nation’s rivers. *Science* 174:905-918.