水質環境モニタリングデータの処理，蓄積と利用の現状*

—— 公共用水域の場合 ——

松 本 幸 雄

1. はじめに

全国の公共用水域における水質汚濁の常時監視は，水質汚濁法15条により，都道府県知事により行われることとされており，昭和46年以降，都道府県ごとに毎年度作成される測定計画に基づいて実施されている。

公共用水域測定データは指定水域ごとに環境基準との比較において水質を評価するために用いられ，その結果は毎年12月に前年度の水質測定結果として環境庁水質保全局より公表されている。この集計の過程で得られた結果の一部を集約したものが水質年鑑¹⁾である。また，水質の全国的かつ経年的な調査としては，昭和33年より建設省河川局により行われている一級河川を対象（約1,000地点）とした調査²⁾があり，一部の地点は公共用水域水質測定地点と重複する。これらの調査により得られた測定データおよび関連資料は，わが国の水環境を記述する極めて貴重な情報である。本稿では，公共用水域水質測定データを例にとり水質環境モニタリングデータの全国レベルにおける処理，蓄積，利用の現状を紹介する。

2. 国立公害研究所における情報業務

かつてわが国で発生した深刻な公害への対策の一環として，環境質のモニタリングを中心に，環境公害の諸分野において，多数の機関で大量の測定または調査データが作成されていた。しかし，データの多面的有効利用の体制は必ずしも十分とはいえず，このため，国立公害研究所が発足した昭和49年頃には，環境データの多面的利用ニーズに効率的に応ずるためのシステ

マチックな整理，蓄積，処理の体制の必要性が各方面より指摘されていた。

このような背景のもとに国立公害研究所では，昭和52年度より，「環境データベース」作業に着手した。おおまかにいえば，数値情報，文献情報，および情報源情報（ある主題に関する情報を所有する人，組織に関する情報）にわけて情報の流れを全国的な立場から整理し，必要なものについてはデータベース化を図り，情報の効率的な利用を図ろうとするものであった。

数値情報として，環境公害関連情報の種類の多さを考えるとすべてを取り上げることは事実上不可能であり，かつ，すでに多くの機関でデータの作成から統計書や磁気テープ等の作成まで行われているものまで組織化の対象に含めることは非現実的でもあるので，昭和52年度当時にはまだ体系的な統計としての扱いの確立していなかった環境質データ（環境常時監視データ）を中心にした。これは基本的に現在迄続いている^{3),4)}。

データの国レベルでの流れを組織化するためには，単に与えられた測定データを電子計算機システムに入力してデータベースをつくるだけでは不十分である。地方自治体における測定データの精度管理の実態の把握に始まって，都道府県での処理，国への報告様式，報告されたデータのチェック，利用体制まで含んだプロセス全体を配慮する必要がある。水質環境のデータファイルはこのような一連の作業のなかで生じる果実である。

3. 水質環境モニタリングデータの処理

昭和53年度まで，公共用水域水質測定結果は，原データと同時にその集計結果も，自治体より環境庁に報告されていた。環境庁水質保全局においては，この自治体の集計結果をもとに全国的な水質状況を評価し，水質年鑑を編集していた。一方，この時期の国立公害研究所の水質環境データファイルは，原データの記載されている調査表（個表と呼ばれた）から，必要な項目（主にBOD，COD）を選んでパンチ入力をおこない，必要なチェックプログラムと編集プログラムにより作成した。従って，環境庁水質保全局の水質評価作業は国立公害研究所の水質ファイルとに直接の関係はなく行われ，評価のための各種集計と水質ファイルとの間

Yukio Matsumoto

昭和43年 東京大学理学部物理学科卒業

45年 同大学院修士課程修了

51年 国立公害研究所

59年 同環境情報部情報調査室長

〔趣 味〕 ジョギング，ハイキング

* Processing, Storage and Use of Water Quality Monitoring Data, 昭和62年3月12日受理

に不整合性が生ずる余地が残されていた。

昭和54年度より報告システムが変わった。自治体から環境庁へ報告された原データは、国立公害研究所にわたり、まず磁気ファイルに収録され、それをもとに全国の水質評価が行われるようになった。これにより水質モニタリングデータの一元的管理が可能になり、全国公共用水域の水質評価に関してはすべて同一の水質ファイルにより行われる体制となった。

公共水域水質データの流れについて、測定計画に基づいてある年度の測定が終了した後の段階から簡単に紹介する。

(1)公共用水域水質測定調査：年度の始めにその前年度のすべての水質測定地点を対象に、測定地点の位置情報や水質評価に必要な項目(表1)を調査する。また、新規に測定を開始した地点および当てはめ水域の変更のあった測定地点については測定地点、当てはめ水域の境界を記入した5万分の1地形図も同時に提出を求める。

(2)水質マスタファイルの作成：(1)の調査結果に、絶対番号(4.1節参照)など国立公害研究所で独自の情報を加えてファイルを作成する。これを「水質マスタファイル」と呼んでいる(4.2節参照。その性格からいえば「測定地点情報ファイル」とでも呼ぶのが適切かもしれない)。

(3)水質データの収録：国立公害研究所は、都道府県より環境庁が報告を受けた水質測定結果の原データを磁気ファイルに収録する(水質データファイル(一次))。報告は、磁気テープによる県が33、環境庁により定められた様式の個表によっている県が14である(昭和60年度測定)。

個表による報告はキー入力する。磁気テープによる報告の場合、自治体独自の仕様により報告する自治体も多く、国立公害研究所の仕様(国公研仕様と呼んでいる)に変換・編集する。

収録の過程では、ニューメリックチェックや測定項目ごとに設けた上下限チェック等にかけることにより、誤りの可能性のある数値を摘出し、必要に応じて、環境庁や都道府県の担当者に確認する。

(4)一次集計とチェック：でき上がった水質データファイル(一次)と水質マスタファイルを用いて集計プログラムにより、各種集計表を作成する(これを一次集計とよぶ)。集計表の種類は30種以上あるが、大きく分けると単県集計と全国集計とあり、単県集計については環境庁経由で自治体の担当者に送付され、チェックを受けた後回収される。全国集計についても環境庁の担当官による総合的な観点からのチェックを受ける。特に、BOD、CODの高濃度水域や有害物質検出上位地点等は、公表後の影響も大きいので特に誤りのないことを確認する。

(5)水質データファイルの確定：回収したチェック済みの一次集計表に基づいて水質データファイルを修正する。この過程によって、自治体からの報告段階に含まれていた単純ミス(個表記入ミスや磁気テープ作成上の誤りなど)およびデータ収録の際に生じた誤りが除去され、原データファイルとしての水質データファイル(二次)は確定する。

(6)二次集計：確定した水質データファイルを用いて再度(二次)集計を行う。

(7)水質評価結果の公表：環境庁は、集計結果に基づき全国の公共用水域の水質を評価し、毎年12月に「昭和・・年度公共水域水質測定結果について」として公表する。また、この集計結果を中心に都道府県が公表している結果も付け加えて「全国公共用水域水質年鑑」¹⁾(表2)として出版している。

4. データの蓄積

前節に述べた流れに従って収録された水質関連のデータは環境データベースの磁気ファイルの1つとし

表1 水質測定地点調査の内容(概略)

測定地点の名称・コード等	(1)地点統一番号(環境基準を考慮した地点コード。4.1節参照) (2)水域名称、測定地点名称(漢字、カタカナ) (3)測定地点の新設、中止、再開等に関する情報
測定地点の位置情報	(1)地点の属する河川・湖沼の名称(本流から三次支流まで) 海域の場合は灘、湾、小湾、地先名 (2)地点の存在する場所の5万分の1地形図の図幅名称とコード (3)地点の緯度経度
水質の評価に必要な情報	(1)環境基準の指定年月日 (2)指定類型、達成期間、指定区分 (3)県際水域に属するか否か (4)測定機関、測定法に関する情報

表2 水質年鑑の掲載項目とその内容

測定項目	水質年鑑掲載内容						
	検体数	未達成数	最小値	最大値	平均値	測定日数	日平均値
健康項目 カドミウム, シアン, 有機リン, 鉛, 六価クロム, ヒ素, アルキル水銀, PCB 総水銀	○ ○	○ ○		○ ○		○	
生活環境項目 pH DO, SS, 大腸菌群数, n-ヘキサン抽出物 BOD, COD	○ ○ ○	○ ○ ○	○ ○ ○	○ ○ ○	○		◎ ^{a)}
その他項目(栄養塩類等)	○	○	○	○	○		
全窒素・全リン(告示法による)	○	○	○	○	○		◎ ^{b)}

a) BOD・CODは、日平均値について、検体数、未達成数、未達成率、最小値、最大値、平均値、中央値、75%値を扱っており、また月平均値についても扱っている。

b) 全窒素・全リンは、日平均値について、検体数、未達成数、最小値、最大値、平均値を扱っている。

て蓄積される。以下に、蓄積の概要について述べる。

4.1 水質地点コードの現状

データを計算機で扱う場合には色々なコードが用いられる。中でも、各種データファイルは年度単位に作られているため、年度間のファイルを接続する場合、地点を同定するためのコードが特に重要になる。建設省や地方自治体においては独自のコードが用いられているが、公共用水域について最初の全国的な地点コードは水質年鑑において用いられた「年鑑コード」(表3)である。昭和54年度のシステム変更に伴い昭和52年度以降の測定結果には、「統一地点番号」(表4)が正式コードとして用いられることになった。

統一地点番号は、環境監視点の体系化を主目的とし、当てはめ水域の有無、環境基準点か補助点の区別をコード化したものである。しかし、欠点としては、類型未指定水域が新規に指定されたり、補助点が基準点に変わったとすると、それ以後はコードがまったく変わり、そのことがコードからは判別できない点である。

この欠点を補うために、現在は経年的に不変な地点に固有な番号が「絶対番号」(省略)として併用されている。このコードは水域当てはめや補助点基準点の異動があっても永久に変えない。

4.2 データファイルの種類と内容

3節で示した水質データの流れの中で生ずるファイルを中心に、国立公害研究所では、次の4通りのファイルを作成している⁴⁾。

(1)水質マスタファイル：公共用水域の約8,000以上の水質測定地点(このうち、指定水域内の地点は約6,800。表5参照)について、3節(1)(2)で述べた調査内容(表1)

表3 水質年鑑コード

X ₁ X ₂ X ₃ X ₄ X ₅ X ₆ X ₇	
X ₁ X ₂	都道府県コード(総務庁の標準地域コードによる) 0 1 ~ 4 7
X ₃	水域分類コード(R, L, Sのいずれかとする) R: 河川, L: 湖沼, S: 海域
X ₄ X ₅ X ₆ X ₇	地点番号0 0 1 ~ 9 9 9の番号を付ける。
X ₇	予備

を中心に、付加情報(地点の継続性を確認のための絶対番号など)とともに作成したファイルである。

このファイルはデータの検索や環境基準適否の判定の基礎となる情報を含んでいる。

(2)水質データファイル：公共用水域水質測定の実データを収録したものである。主な収録項目とデータ量を表5に示す。

(3)流量マスタファイル：水質データの解析には汚染質の濃度だけでなく汚濁負荷量も評価することが必要である。水質データファイルにも流量の記録はあるが、解析の際の補助情報として建設省作成の流量年表をもとに流量測定点と公共用水域水質測定点との関連を示す情報などを付加してファイルを作成する。このファイルは、表6に示すように、流量測定点に関する属性情報を収録したものである。

(4)流量データファイル：上述の流量年表の測定点ごとに流量の原データを収録している(表6参照)。

4.3 水質測定地点情報と地図

水質測定地点の位置情報をもとに5万分の一地形図台帳を毎年更新する。これをもとに20万分の一の地図

表 4 (a) 地点統一番号

X ₁ X ₂ X ₃ X ₄ X ₅ X ₆ X ₇						
X ₁ X ₂	都道府県コード 0 1 ~ 4 7					
X ₃ X ₄ X ₅	水域コード (表 4 (b))					
X ₆ X ₇	地点コード (表 4 (b))					

表 4 (b) 水域および地点コード

水 系	水域の類型指定	地点の性格	使用可能なコード	
			水域コード	地点コード
河 川	有	基準地点 補助地点	001~199 001~199	01~49 51~99
	無	—	201~399	01~99
湖 沼	有	基準地点 補助地点	501~599 501~599	01~49 51~99
	無	—	401~499	01~99
海 域	有	基準地点 補助地点	601~699 601~699	01~49 51~99
	無	—	701~999	01~99

帳⁵⁾を刊行している。

また、5 万分の一、20 万分の一の地形図と同一縮尺で測定点の位置と地点統一番号を XY プロッタに出力するシステムを開発した。

5. データの利用

5.1 データファイルの利用体制の現状

環境数値データベースは、広く環境研究や環境行政の基礎資料として用いられることを願って整備されているものであり、利用の促進を図ることが望ましい。しかし、現状では、いくつかの限界がある。

まず、オンライン利用については（検討中の部分はあるが）サービス体制にはない。1つの理由は、環境数値データをディスクに常駐することの非効率性のためである。また、外部利用者に対して国立公害研究所

表 5 水質データファイルの収録内容

データの年度		46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
地点数 ^{a)}																
河 川		1,077	1,612	2,296	2,382	3,263	3,634	3,795	4,009	3,978	4,073	4,129	4,165	4,232	4,169	4,159
湖 沼		55	115	150	169	245	275	280	354	330	350	366	371	378	369	407
海 域		323	647	1,104	1,119	1,732	2,005	2,091	2,187	2,053	2,124	2,125	2,178	2,201	2,207	2,207
計		1,455	2,374	3,350	3,670	5,240	5,914	6,166	6,550	6,361	6,541	6,620	6,714	6,811	6,745	6,773
総調査地点数						5,883	5,914	7,168	8,196	7,951	8,142	8,310	8,340	8,387	8,317	8,425
データ量(キロバイト)		24,270	39,840	53,540	56,650	82,430	86,950	96,870	123,180	122,602	123,777	125,056	124,469	121,242	119,444	118,234
主 な 入 力 項 目	流 量						○	○	○	○	○	○	○	○	○	○
	pH								○	○	○	○	○	○	○	○
	BOD ^{b)}	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
	COD ^{c)}	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
	SS							○	○	○	○	○	○	○	○	○
	DO						○	○	○	○	○	○	○	○	○	○
	大腸菌群数								○	○	○	○	○	○	○	○
	油分等								○	○	○	○	○	○	○	○
	全窒素													○	○	○
	全リン													○	○	○
	カドミウム								○	○	○	○	○	○	○	○
	シアノ								○	○	○	○	○	○	○	○
	有機リン								○	○	○	○	○	○	○	○
	鉛								○	○	○	○	○	○	○	○
健 康 項 目	クロム(6価)								○	○	○	○	○	○	○	○
	ヒ素								○	○	○	○	○	○	○	○
	総水銀								○	○	○	○	○	○	○	○
	アルキル水銀								○	○	○	○	○	○	○	○
	PCB								○	○	○	○	○	○	○	○
そ の 他	チッ素化合物						○	○	○	○	○	○	○	○	○	○
	リン化合物						○	○	○	○	○	○	○	○	○	○

a) 地点数は生活環境項目の類型指定水域のみ、総調査地点数とは区別。b) 河川について測定。c) 湖沼、海域について測定。

表6 流量ファイルの内容

流量マスタ ファイル	建設省流量観測所コード, 水質・流量測定 地点一致および不一致表示, 流域名称, 支 流名称, 観測所名称, 流域面積, 河口から の距離, 測定開始年
流量データ ファイル	測定開始年以降の最大・豊水・平水・低水・ 渇水・最小, 年平均, 年総量値, 日流量デー タ, 月合計, 月平均値, 年間合計, 年間平 均値

の電子計算機システムが, 所外からのオンラインアクセスに対応できる体制にないことも他の理由である。

オフラインの利用は, 「磁気テープ貸し出し規定」によって行われており, その対象は, (大まかにいえば) 大学を含む研究機関と国または地方公共団体とされている。利用を希望される方は国立公害研究所環境情報部情報調査室へ連絡をとりたい。

5.2 解析用のソフトウェア

通常の業務用データベースではデータファイルを扱うのにソフトウェアとして, (1)検索機能を含むデータベース管理システムが用いられる。また, ファクトデータベースである環境数値データには, (2)データ解析システムが重要となる。

(1)については, 過去に独自の管理システムを開発したが, 研究用にデータを利用するには電算機資源の効率の点から適当でなく現在は常駐していない。環境数値データの場合, 単純な形式で原データを保存し, 目的に応じた加工を行うためのソフトウェアを準備するという基本的考えで対処している。ただし, 水質年鑑(表2)レベルの要約データファイルに関しては行政面から検索機能の需要が考えられるので現在オンライン利用のための体制の検討を進めているところである。

(2)については, 環境データベースのなかで一般的に提供できるのは統計的手法に関するソフトウェアである。しかし, 従来の統計の方法を適用する場合には注意が必要である。環境数値データは, 異常値や欠測値の混入が避けられないことや, データに仮定するモデルをあらかじめ決めることが困難なことが多い。従って, 統計解析にあたっては, (a)異常値の処理, 分布形の判定, 欠測処理等多くの前処理が必要なことや, (b)手法の適用後にモデルと残差の整合性を吟味する等, 多くのプロセスを段階的, 対話的にすすめる必要が生じる。当研究所では数値データ利用ソフトウェアの1

つとして, 計算機と対話しながらデータの特徴を調べするためのプログラム⁶⁾(CDA Conversational Data Analysis Program 検索機能は持たない)を開発し, 所内利用者には公開している。(所外の問い合わせに対しても可能な範囲で対応している。)

5.3 さらに高度な利用

データは環境研究や環境行政に一層広く利用していただくことが望ましい。もちろん, 水質データは行政データとして作成されているものであり, この側面から管理上検討されるべき点があると思われるが, 単に技術的側面に限っても, データ要約の方法, 水質以外の関連データ(発生原データ, 経済社会データ等)との接続を容易にする方法, オンラインによるサービスの技術的可能性等, 検討することは多い。国立公害研究所では研究レベルでこれらの検討を進めている。

一方, 環境庁では行政上のデータ活用の観点から水質環境情報高度利用システムの開発を61年度より進めており, 利用技術の開発の面からも期待は大きい。

6. おわりに

以上述べたように, 水質関連のデータは毎年, 国立公害研究所の環境データベースのデータファイルとして蓄積される体制がようやく整ったところである。これらのデータはわが国の水質環境状態を記録するものとして将来にわたって貴重な国民的財産となるものと考えられる。今後, 一層, 有効利用のための体制が進展することを期待する。

おわりにあたって, モニタリング(常時監視)にご尽力されている環境庁の担当官ならびに地方自治体の担当者の方々に敬意を払うとともに, 今後の援助を賜わるよう心からお願いする次第である。

文 献

- 1) 環境庁水質保全局監修, 全国公共用水域水質年鑑, 芙蓉情報センター。
- 2) 建設省河川局編, 水質年表, 関東建設弘済会。
- 3) 藤原正弘(1985) 環境データの組織化と環境問題への適用に関する研究, 京都大学工学部衛生工学科学学位論文。
- 4) 廣崎昭太(1982) 環境情報のデータベース, 環境情報科学, 11, (2), 25-32。
- 5) 国立公害研究所(1982) 全国公共用水域水質測定地点図, 国立公害研究所研究資料第23号。
- 6) 廣崎昭太, 新藤純子, 芳賀敏郎(1986) 環境データ解析のための対話型解析プログラム, 国立公害研究所資料第30号。