

(株) 日本水道コンサルタント 萩原 良巳
同 上 ○ 中川 芳一

1. はじめに

地域における水環境計画を策定するに際し、計画対象地域の地域特性を十分に把握することが重要であり、とくに地域特性が鋭敏に反映されるのは計画入力である。地域における水環境計画の一構成要素となる下水道計画についてみれば、その計画入力のひとつである流達負荷量は、対象地域の地域特性を鋭敏に反映して大きな地域差が生ずるものと考えられ、この流達負荷量の把握が十分なされなければ、下水道施設の施設計画自体が不確かなものとなる。そこで、本稿では、流達負荷量の地域特性との関連に注目した構造分析法について述べる。

従来、流達負荷量の予測は、発生負荷量を原単位横上げ方式で算定し、それに一定の流達率を掛けることにより行われてきた。たしかに、流達負荷量は発生負荷量と流達率の積と考えることができ、発生負荷量、流達率はともに地域特性の影響を受け、その結果として流達負荷量に地域差が生ずるものと考えることができよう。

そこで、本稿では、2.で、負荷発生および負荷流達に影響を及ぼすと考えられる各種要因(特性値)に、統計的手法のひとつである主成分分析法(規準バリマックス法による軸の回転も含む)を適用することにより、計画対象地域の地域特性の総合的把握、地域の分類・説明等を行なう。ついで、3.で、2.の地域分析の結果を受けて地域特性を考慮した流達負荷量の構造分析を、重回帰分析法によって行なう。本稿での分析対象地域はN県のY川流域とし、分析の単位は、水質・水文データの関係上支流域とする。また、汚濁指標はBOD負荷量とし、支流最下流端における流達負荷量の実測値を用いた。なお、分析対象地域の略図を図-1に示す。

2. 主成分分析による地域の分類²⁾³⁾⁴⁾

ここでは、主成分分析により地域の性格特性の総合的把握、支流域の分類・説明を行なう。主成分分析法は、多変量解析における一手法であり、互いに相関のあるP個の特性値のもつ情報を互いに無相関なm個($m \leq P$)の総合特性値に要約するものであり、地域のように基本的な構造が不明であり、それを構成していると考えられる要因が多すぎて明確に(数量的に)その特性を記述しにくい場合、特性の集約化のための分析を行なうのに有効な手法である。

分析にあたって、流達負荷量に影響を及ぼすと考えられる種々の特性値(データ項目)のうち、経験的に重要と考えられ、入手可能であったものを支流域毎に収集した(収集したデータは昭和45年度のデータである)。収集した全てのデータ項目を用いて主成分分析を行なうことは、それらの項目間に大なり小なりの相関があることが予想されるので必ずしも必要ではなく、また、適当に項目の選択を行なった後に主成分分析を適用する方が分析結果の解釈が容易となることもある。そこで、以下の分析においては、経験的重要性、データ間の相関等を考慮して、分析に用いるデータ項目の選択を行なった。

ここでは、地域分析を流達負荷量との関係性を考慮して行なうことを目的としている。流達負荷量は支流域の面積的大小の影響を大きくうけると考えられるが、本稿では、支流域の面積的大小による流達負荷量の差異よりも、負荷発生および負荷流達の構造からみた地域の分析を行なうことを主眼とする。そこでまず、各データ項目を単位面積当りの値に変換したデータを特性値として地域

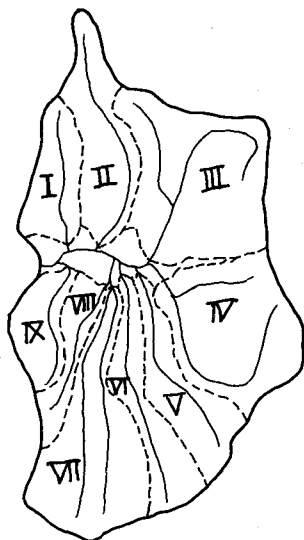


図-1 分析対象地域の略図

の分析を行なった(CASE-A)。ついで、単位面積当りの各特性値をさらに流達負荷量で割ったものを新たな特性値として地域分析を行なった(CASE-B)。すなわち、CASE-Aで、負荷発生機構からみた地域の分析を行ない、CASE-Bで、負荷発生と負荷流達の両方の構造を考慮した地域の分析を行なった。これらの分析に用いた特性値を表-1(CASE-A)、表-2(CASE-B)に示す。なお、CASE-Bの特性値は、CASE-Aの分析で得られた特性値の分類の結果を参考にし、表-1に示す特性値より選択した。以下、CASE-A、CASE-Bの分析結果について記す。

【CASE-A】 表-1に示す23項目の(単位面積当りの)特性値に主成分分析を適用した結果、表-3に示すような固有値、寄与率をもつ主成分(総合特性値)が抽出され、第4主成分(Z₄)までで全体の変動の81%を説明しており、とくに、第1、第2、第3主成分(Z₁, Z₂, Z₃)の説明力が大きい。ここで、抽出された各総合特性値の意味を解釈するの¹⁾があるが、意味の解釈がし易くなるように(標準バリマックス法による)軸の回転を行ない、回転後の各特性値の固有ベクトルにより総合特性値(軸回転後はこれを因子と呼ぶ)の意味を解釈する。

第1因子(F₁)に寄与する特性値の度合(固有ベクトル)は、X₁(人口)、X₅(し尿処理人口)、X₂₁(財政)、X₃(商業地区面積)、X₁₀(工業出荷額)、X₂(住居地区面積)、X₁₈(平地面積)等が正で大きな値を示し、逆にX₁₃(森林面積)が負で大きな値を示している。このことより、第1因子は“都市活動の強さを示すファクター”と意味付けできる。また、支流域別の因子得点より、Ⅲ、Ⅳ、Ⅰ、Ⅱ支流域では都市活動が活発であり、逆に、Ⅴ、Ⅵ支流域等では都市活動が弱いことがわかる。

第1因子の場合と同様にし、第2因子(F₂)、第3因子(F₃)の意味を解釈すると、第2因子は、支流域において工場の占める割合が大きい、オフィス街や大規模団地の占める割合が大きいかを区別する因子、すなわち、第1因子で表わされる都市活動が工業によるものかどうかを区別する因子と意味付けでき、第3因子は、農業活動の強さを示すファクターと意味付けできる。

なお、表-1とほぼ同様な(単位面積当りの)特性値を用いて、分析単位を市町村として主成分分析を行なった結果、上述の支流域単位の場合抽出された総合特性値と同様な総合特性値が抽出されたことを付記しておく。⁴⁾

【CASE-B】 表-2に示す10項目の(単位面積、単位負荷量当りの)特性値に主成分分析を適用した結果、表-3に示す固有値、寄与率をもつ主成分が抽出され、第1、第2主成分(Z₁, Z₂)で全体の変動の95%を説明している。CASE-Aの場合と同様に軸の回転を行ない、軸回転後の各特性値の固有ベクトル(表-2参照)でもって各因子の意味を解釈する。

第1因子(F₁)では、X₄(工業出荷額)、X₆(家畜頭数)、X₇(農地面積)、X₁(人口)、X₃(し尿処理人口)、X₁₀(財政)、X₂(市街地面積)が正で大きな値を示し、このことより第1因子は人間活動の強さを示し、河川水質からは“水質悪化を示すファクター”と意味付けできる。また第2因子(F₂)では、X₉(池圃数)、X₅(森林面積)、X₈(流量)が正で大きな値を示し、このことより第2因子は自然の強さを示し、河川水質からは“水質浄化を示すファクター”と意味付けできる。

このように、河川水質の立場から、第1因子は水質悪化を示すファクター、第2因子は水質浄化を示すファクターと意味付けでき、これらの因子により支流域の分類を行なうことが可能となる。図-2に各支流域の第1、

特性値(単位面積当り)	
X ₁ 人口	
X ₂ 住居地区面積	
X ₃ 商業地区面積	
X ₄ 下水処理人口	
X ₅ し尿処理人口	
X ₆ し尿未処理人口	
X ₇ 工業就業人口	
X ₈ 工場敷地面積	
X ₉ 工業排水量	
X ₁₀ 工業出荷額	
X ₁₁ 田 面 積	
X ₁₂ 畑 面 積	
X ₁₃ 森林面積	
X ₁₄ 家畜頭数	
X ₁₅ 流 量	
X ₁₆ 支 川 数	
X ₁₇ 流域山地面積	
X ₁₈ 流域平地面積	
X ₁₉ 農業用水取水量	
X ₂₀ 池 圃 数	
X ₂₁ 財 政	
X ₂₂ 流達負荷量	
X ₂₃ 流 路 長	

表-1 CASE-Aの特性値

特性値	軸回転後の固有値	F ₁	F ₂
(単位面積、単位負荷量当り)			
X ₁ 人口	0.865	0.987	
X ₂ 市街地面積	0.787	0.984	
X ₃ し尿処理人口	0.825	0.557	
X ₄ 工業出荷額	0.986	0.023	
X ₅ 森林面積	0.267	0.918	
X ₆ 家畜頭数	0.936	0.010	
X ₇ 農地面積	0.878	0.921	
X ₈ 流 量	0.399	0.905	
X ₉ 池 圃 数	-0.012	0.977	
X ₁₀ 財 政	0.800	0.535	

表-2 CASE-Bの特性値及び固有ベクトル

CASE 固有値	CASE-A					CASE-B	
	Z ₁	Z ₂	Z ₃	Z ₄	Z ₅	Z ₁	Z ₂
固有値	9.39	3.54	3.16	2.54	1.82	7.50	1.95
累積寄与率	0.41	0.56	0.70	0.81	0.89	0.75	0.95

表-3 各主成分の固有値、寄与率(CASE-A、CASE-B)

オ2因子得点による散布図を示す。図-2より、支流域は、

- a) 水質悪化も大きく、水質浄化も大きい支流域
..... II 支流域
- b) 水質悪化は大きい、水質浄化は小さい支流域
..... VI, VII, IX 支流域
- c) 水質悪化は小さい、水質浄化は大きい支流域
..... I 支流域
- d) 水質悪化も小さく、水質浄化も小さい支流域
..... III, IV, V, VIII 支流域

の4つに分類できる。

なお、各特性値の上記の総合特性値(主成分)への寄与の度合(因子負荷量)により、特性値の分類も可能であるが、その分類結果は、ここでは省略する。

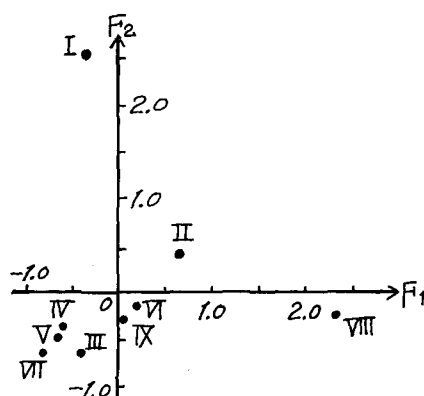


図-2 支流域散布図 (CASE-B)

3. 重回帰分析による流達負荷量の構造分析^{4) 5)}

ここでは、各支流域の流達負荷量の構造分析を行なう。流達負荷量は、前述したように発生負荷量と流達率の積と考えられ、それはともに2. で掲げた種々の特性値(データ項目)の影響を受け、地域差が生ずる。発生負荷量、流達率の構造分析を個々に行なうことも可能であるが²⁾、負荷量データは1. はじめにでも述べたように流達負荷量であり、それを発生負荷量、流達率に分離することは不可能であり、筆者らはその必要を感じない。(流達負荷量データを発生負荷量、流達率に分離することの必要性を感じないのであって、発生負荷量および流達率を個々に実測データ等をもとにして推定したり、その構造を究明することの必要性を否定するものではない。)そこで、本稿では、発生負荷量、流達率の両方を含んだ流達負荷量の構造分析を、2. の地域分析の結果を考慮して行なう。各支流域の流達負荷量を支流域の特性を示すデータ項目のいくつかより物理的に決定するモデルを設定することは、いまのところ困難である。そこで、後述するように相関分析、2. の主成分分析の結果および経験の重要性より選択した特性値(データ項目)と流達負荷量との間に重回帰分析を適用して、それらの線型一次式(重回帰モデル)を推定することにより、流達負荷量の構造分析を行なう。

重回帰分析により流達負荷量の構造式(重回帰式)を作成するにあたり、サンプル数 n と説明変数 p との間には経験的に $n-p-1 \geq 10$ なる関係が成立することが良いとされている¹⁾。ところが、水質の測定が行なわれるようになったのは近年のことであり、水質データ(流達負荷量データ)の得られる年度には限りがあるため(本稿の場合、S44年から47年までの4年度)、各支流域毎に重回帰式を作成することはサンプル数が少なくなり不可能である。そこで、2. の地域分析の結果を参考として対象支流域を2つのグループに分類し、そのグループ毎に重回帰式を作成する。すなわち、2. の流達負荷量を考慮した地域分析(CASE-B)の結果、オ1因子として水質悪化を示すファクターが抽出され、この因子の因子得点の正負により支流域を以下の2つのグループに分類する。

グループ1 (水質悪化が大きい支流域) II, VI, VII, IX 支流域

グループ2 (水質悪化が小さい支流域) I, III, IV, V, VIII 支流域

そして、これらのグループ内の支流域では、流達負荷量の構造は同じであると考え、各グループ毎に重回帰式を作成する。

つぎに、目的変数である流達負荷量を説明するのに役立つと考えられる特性値は、表-1のデータ項目のように多数考えられる。ここでそれら全ての特性値を用いた重回帰式を求めることは、それらの特性値の間には相関があり、その特性値の代表している立場が重複するため意味がなく、その一部の特性値だけを用いた重回帰式が十分間に合う。さらに、前述したようにサンプル数 n による制限を受ける。また、説明変数の数を少なくすること

は、重回帰式が簡単になり、その解釈がし易くなるという利点もある。そこで、本稿では、経験的重要な、2.の主成分分析の結果、経年データの入手のし易さ等を考慮して説明変数として用いる特性値を選出した。選出した特性値は、都市活動を代表するものとして“人口”、工業活動を代表するものとして“工業出荷額”、農業活動を代表するものとして“農地面積”、自然の状態を代表するものとして“森林面積”の4項目である。これらの4項目の特性値について昭和44年から47年までの4年度のデータを各支流域毎に収集した。この結果、サンプル数はグループ1では16、グループ2では20となる。2.でも述べたように、本稿では、地域の面積の大小による流産負荷量の差異よりも流産負荷量の構造の分析を目的とする。そのため、目的変数としては流産負荷量を各支流域の総面積で割ったものを用い、説明変数として上述の4項目の特性値を総面積で割ったものを用いて重回帰分析を行なった。表-4に採用した変数とまとめ示す。

表-4に示した変数の相関分析を行なった結果、 X_1 と他の変数との相関はグループ2の X_4 を除き低く、グループ1では X_2 と X_4 、 X_2 と X_5 、 X_4 と X_5 との相関が、グループ2では X_2 と X_3 との相関が高いことがわかった。以下、流産負荷量の構造式の選定について述べる。

まず、説明変数として X_2 、 X_3 、 X_4 、 X_5 を用いてグループ毎に重回帰分析を行なった。その結果得られた寄与率、F値、各偏帰帰係数のt値を表-5に示す。この表より、寄与率はともに約0.7であり、F値は1%水準で有意である。また、t値はグループ1の X_3 を除き10%水準で有意であり、特に、グループ1の X_4 、 X_5 、グループ2の X_2 、 X_4 は1%水準で有意である。つぎに、この分析でt値の絶対値の小さい変数(グループ1では X_2 、 X_3 、グループ2では X_5)を除去して新たに重回帰分析を行なった。その結果も表-5に示してあるが、寄与率の減少が大きい割には各偏帰帰係数のt値の有意性はあまり変化がなかった。また、他の変数を除去した場合の分析を検討したが有意な結果は得られなかった。そこで、流産負荷量構造モデルとして、説明変数として $X_2 \sim X_5$ の全てを用いた重回帰式を採用することとし、その重回帰式を以下に記す。

$$\text{グループ1: } y = -305.657 - 0.1086X_2 + 0.0546X_3 + 0.0897X_4 + 5.0170X_5 \quad (1)$$

$$\text{グループ2: } y = 4.786 + 0.0852X_2 - 0.1176X_3 + 0.0622X_4 - 2.8757X_5 \quad (2)$$

(1)、(2)式の偏帰帰係数は説明変数の測定単位によって左右されるので、偏帰帰係数相互の比較は無意味である。そこで偏帰帰係数の大きさを測定単位によって左右されないようにするため、各変数を平均0、分散1になるように基準化し、基準化された変数(y' 、 X_i' 等)ダッシュを付けて示す)に対する重回帰式を求めると、

$$\text{グループ1: } y' = -2.6248X_2' + 1.0269X_3' + 5.3461X_4' + 3.1733X_5' \quad (3)$$

$$\text{グループ2: } y' = 1.4996X_2' - 1.3630X_3' + 0.9657X_4' - 0.3826X_5' \quad (4)$$

となる。(3)、(4)式の変数の内容は(1)、(2)式と同じであり、表-4を参照)

(3)、(4)式より、基準化された目的変数 y' に与える各説明変数の影響は、グループ1では X_2' 、 X_3' 、 X_4' 、 X_5' ともかなりあるといえるが、特に X_4' 、 X_5' の影響が大きく、グループ2では X_2' と X_3' は大きい、 X_4' と X_5' は小さいことがわかる。また、グループ1では X_2' 以外は正の影響を与えているのに対し、グループ2では X_2' 、 X_4' は正の影響を、 X_3' と X_5' は負の影響を与えている。

変数	項目
目的変数 $X_1 (=y)$	流産負荷量/面積 ($kg/d/km^2$)
説明変数 X_2	人口/面積 ($人/km^2$)
X_3	工業出荷額/面積 (百万円/ km^2)
X_4	農地面積/面積 (ha/km^2)
X_5	森林面積/面積 (ha/km^2)

表-4 重回帰分析の変数

		1	2	3	4
グループ1	選択した説明変数	X_2, X_3, X_4, X_5	X_2, X_3, X_5	X_2, X_3, X_4	X_3, X_4, X_5
	寄与率	0.699	0.611	—	0.600
	F値	6.365	6.270	—	6.988
	t値	X_2	-1.899	-0.586	—
		X_3	1.789	—	-0.057
		X_4	3.302	3.862	3.380
		X_5	2.791	2.208	4.020
		切片	-305.671	-220.578	-223.466
	偏帰係数	X_2	-0.1086	-0.0111	—
		X_3	0.0546	—	-0.0066
		X_4	0.0897	0.0450	0.0413
		X_5	5.0170	3.9634	4.0020
		寄与率	0.709	—	0.633
グループ2	選択した説明変数	X_2, X_3, X_4, X_5	X_2, X_3, X_4	X_2, X_3, X_5	X_3, X_4, X_5
	寄与率	0.718	—	0.781	—
	F値	9.118	—	9.181	—
	t値	X_2	3.204	2.351	—
		X_3	-2.782	-1.732	—
		X_4	5.466	5.053	—
		X_5	-1.898	—	—
		切片	4.786	-122.884	—
	偏帰係数	X_2	0.0852	0.0539	—
		X_3	-0.1176	-0.0671	—
		X_4	0.0622	0.0625	—
		X_5	-2.8757	—	—

表-5 重回帰式の寄与率、F値、t値および偏帰帰係数

(注) F値、t値の*印は、
 * 10%水準
 ** 5%水準
 *** 1%水準
 } が統計的に有意な値であることを示す。

以上のことから、流達負荷量に対して、水質悪化の大きいグループ1の支流域では、人口、工業出荷額、農地および森林面積の全てが影響を与えているが、水質悪化の小さいグループ2の支流域では、人口、工業出荷額の与える影響が大きく、他の変数の与える影響はそれらに比べて小さいと言える。すなわち、人口で代表される都市活動、工業出荷額で代表される工業活動の流達負荷量に与える影響はグループ1、2とも同程度であるが、農地面積で代表される農業活動および森林面積で代表される自然の状態で与える影響は、水質悪化の大きいグループ1では大きい、水質悪化の小さいグループ2では小さい、(3)、(4)式の重回帰式において、流達負荷量に与える農地および森林の影響の大小が総合特性として水質悪化の大小を特色付けていると言ったことができよう。

4. おわりに

その結果を以下に要約する。

ii) 単位面積、単位負荷量当りの特性値に主成分分析を適用した(CASE-B)結果、水質悪化を示す因子と水質浄化を示す因子が抽出され、負荷発生、負荷流達構造からみた地域の分析が行なえ、これら2つの因子によって、河川水質に与えるインパクトという観点から対象流域の分類を行なうことができた。

iii) として、河川水質に与えるインパクトにより対象流域を2つのグループに分け、各グループ内の流域では流達負荷量の構造は類似であるという考えのもとで、グループ毎に重回帰分析により統計的に有意な流達負荷量の構造式(1)、(2)式を作成することができた。なお、これらの構造式は、1)で分析した地域特性と適合しており、その信頼性も明示した。

iv) 主成分分析と重回帰分析を組み合わせることにより、地域特性を考慮した要因の分類が可能となり重回帰分析における説明変数選択の情報を得るとともに、地域特性の類似性による地域の分類により各グループ毎に構造式を作成することが可能となり、サンプル数が少ない地域を対象とした場合にも重回帰分析の適用が可能となった。特に、本稿で述べたような流達負荷量構造式の作成に際して、水質データは近年のものしか入手できないことが多く、このような場合、本稿で提示した分析法は特に有効と思われる。

また、本稿の向題点、今後の課題、展望について述べる。

i) 本稿では、流達負荷量としてBOD負荷量を用いたが、水環境と人間活動との関係を記述する測度としてBOD負荷量だけでは十分の議論が残り、BOD以外の水質項目についても本稿で行なったような分析を行なう必要があるとともに、水質項目毎に本稿で用いたような統計的手法を適用して各水質項目間の関連を分析することも必要であろう²⁾。

ii) 本稿で述べたような分析は、地域特性を考慮して種々の計画を行なう場合に基礎的情報を提供するものとして大きな役割を果たすことが期待できよう。たとえば、都市活動の強さという総合特性値の大小を目安として将来の地域の下水道整備率の定むを決定すると⁴⁾、水需要構造の分析⁵⁾、将来水需要量の予測⁶⁾等が考えられる。

iii) 3.2で作成した流達負荷量構造式により、将来の流域の流達負荷量を予測する場合、将来その流域が如何なる構造に遷移していくかを判別度数法などで推定し¹⁰⁾、遷移後の構造に合致する構造式を選択し、その構造式により流達負荷量の将来予測を行なうことも考えられる。また、流達負荷量構造分析で対象流域以外の流域の影響を加味した構造モデル式を作成することも考えられる。このように構造の遷移を考慮して構造式を使いわけると予測法¹¹⁾や、流域間の影響を考慮した予測法¹²⁾により、さらに有意な計画入力の作成が可能となる。

iv) 本稿では、地域の分析に主成分分析法を、流達負荷量構造式の作成に重回帰分析法を用いたが、これらの他にも種々の統計的手法の適用も考えられ、また、システムダイナミックスにより下水道計画に限らず地域の水環境全体の構造の把握および将来の変化、発展の予測を行なうことも可能である。さらに、負荷発生、負荷流達の構造を物理的に究明し、物理的モデルの開発、検討も必要である。今後、これらの種々の手法、モデルを地域の水環境計画作成のプロセスに適用して、その構造の把握、変化過程の予測を行なうことの重要性を強調するとともに、水環境がもたらデータ¹³⁾の集積、整理の必要性を強調して本稿の結びとする。

最後に、データの収集、整理にあたり多大な御協力を得た当社・花房畜一、堀雅文の両氏および本稿の作成にあたり熱心な討議を下された当社のデータ解析プロジェクトの方々へ謝意を表します。

[参考文献]

- 1) たとえば、奥野忠一他：“多変量解析法”，日科技連，1971
- 2) 堀氏、萩原氏、菅原邦夫：“下水道計画のための自然的カット率に関する一考察”，土木学会第29回年會，1974
- 3) 萩原良巳、中川芳一：“水環境からみた地域分析”，土木学会第30回年會，1975
- 4) 萩原良巳、小泉明、中川芳一、菅原邦夫：“下水道整備計画に関するシステム論的研究Ⅱ”，土木学会第12回衛生工学研究討論会，1976
- 5) 萩原良巳、中川芳一：“下水道計画のための地域分析と流達負荷量の関連に関する一考察”，下水道協会誌・投稿集中
- 6) 萩原良巳、中川芳一：“下水道整備計画に関するシステム論的考察Ⅱ”，土木学会第12回衛生工学研究討論会，1976
- 7) 菅原邦夫：“河川における水質評価とこれの生物学的指標と理化学的指標の関連性について”，NSC研究年報Vol.4, No.2, EHC, 1976
- 8) 阿久津武彦、萩原良巳、小泉明、中川芳一、沼田篤男：“水需要予測のための地域分析”，第27回全国水道研究発表会，1976
- 9) 萩原良巳、小泉明、中川芳一：“水利用計画に関するシステム論的研究”，NSC研究年報Vol.4, No.2, EHC, 1976
- 10) 堀雅文、中川芳一、西沢常彦：“判別度数による排水水質が与えた市町村の分類に関する一考察”，土木学会第30回年會，1976
- 11) 西沢常彦、萩原良巳、小泉明：“水需要予測モデルに関する考察(2)——構造遷移を考慮した場合”，同上
- 12) 萩原良巳、小泉明、中川芳一：“水需要予測モデルに関する考察(1)——市町村間の影響を考慮した場合”，同上