

演 題	主成分分析法とニューラルネットワークを用いた河川の上流・中流・下流を示す水質パラメータの抽出（第 2 報）－ 東京多摩川の水質データ（1997～2001）を用いて－	
発 表 者 （ 所 属 ）	○神部順子，長嶋雲兵*，青山智夫** （大東文化大，*産総研グリッド研究センター，**宮崎大工）	
連 絡 先	〒305-8561 つくば市東1-1-1 産総研先端情報計算センター tel. 0298-61-3044, Fax 0298-61-2048 e-mail: u.nagashima@aist.go.jp	
キ ー ワ ー ド	Water Contamination, Principal Component Analysis, Differential Coefficients Analysis of Input Parameter for Neural Network	
開 発 意 図 適 用 分 野 期 待 効 果 特 徴 等	東京多摩川の水質データ5年分（1997～2001）を用いて河川の上流・中流・下流を示す水質パラメータの抽出を試みた。2000年よりいくつかのパラメータの観測結果が発表されなくなったが、その妥当性も検討する。	
環 境	適 応 機 種 名	
	O S 名	
	ソ ー ス 言 語	
	周 辺 機 器	
流 通 形 態 （ 右 の い ず れ か に ○ を つ け て く だ さ い ）	・日本コンピュータ化学会の無償利用ソフトとする ・独自に頒布する ・ソフトハウス、出版社等から市販 ・ソフトの頒布は行なわない ・その他 ○・未定	具 体 的 方 法

1. はじめに：河川の上流・中流・下流を示す水質パラメータを抽出するために、東京都環境局の「公共用水域および地下水の水質測定結果」[1]に掲載の多摩川のデータを用いた。1997年から1999年の3年分（17地点の水質に関する12種類のパラメータ）について解析した結果は The Journal of Chemical Software, Vol. 8, No. 1 で報告した。今回は2000年と2001年のデータを追加し、その後の動向を把握する。

2. データと解析プログラム：追加した2000年と2001年のデータには、前回の報告に用いた12種類のパラメータのうち、Cl⁻とCONDのデータがいくつかの観測地点で記載されていない（なぜこの2つが測定項目から外れたのかは分析担当機関である国土交通省に問い合わせ中である）。また、17地点のうち、1地点（調布取水点（中流））は廃止されていた。そこで、まず、前回の報告で用いた12種類のパラメータと今回用いる10種類パラメータでの解析結果では違いが生じるかをみるために、前回の3年分のデータでの比較を行なった。その後、5年分（16地点の10種類のパラメータ）について

解析した。分析方法としては、前回と同様、主成分分析にはエクセル統計[2]を用いた。ニューラルネットワークシミュレータは、我々が開発中の Neco[3]を用いた。

3. 主成分分析の結果：1997年から1999年の3年分のデータで、 Cl^- とCONDを除いた10種類のパラメータを用いた主成分分析では、第一主成分の寄与率が71.3%、第二主成分の寄与率は9.3%となり（12種類のパラメータを用いた場合は61.0%と20.1%）、10種類のパラメータで主成分分析をかけた方が第一主成分の寄与率がより高い結果となった。もちろん Cl^- とCONDのデータを加えた12種類のパラメータで分析を行なったほうが中流と下流の分類が明確である。だが、全体的には10種類のパラメータでも充分分類が可能であった。

5年分のデータの10種類のパラメータを用いた主成分分析の結果、第一主成分の寄与率は72.6%、第二主成分の寄与率は9.2%、第三主成分の寄与率は6.9%、第四主成分の寄与率は5.7%である。第一主成分と第二主成分における16地点の主成分得点を用いたK-LプロットをFigure 1に示した。Figure 1をみると、第一主成分では、上流と中流・下流の分類が可能であり、中流と下流は分離できていない。他方、第二主成分では98年を除くと上流・中流と下流が大まかに分離されている。

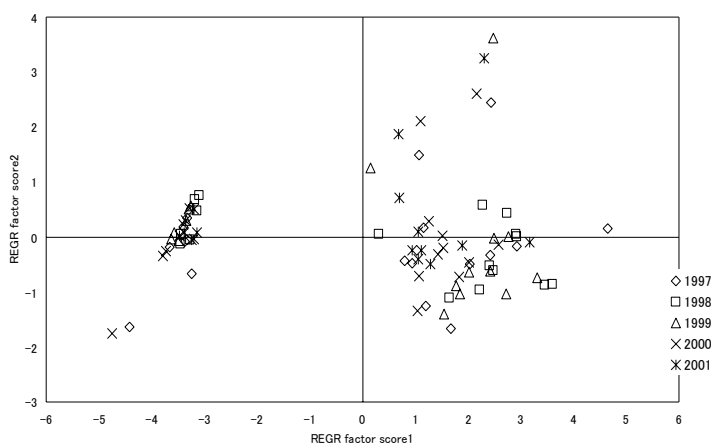


Figure1.the K-L plot of component score

Figure2 と Figure3 に 5 年分のデータに対する第一主成分と第二主成分の主成分負荷量を示した。Figure2 と Figure3 をみるとすべてのパラメータがほぼ同じ重みで寄与していることがわかる。そこで主成分分析では10種類のパラメータのうち上流・中流・下流の分類に大きな寄与をするパラメータの抽出はできないことがわかった。

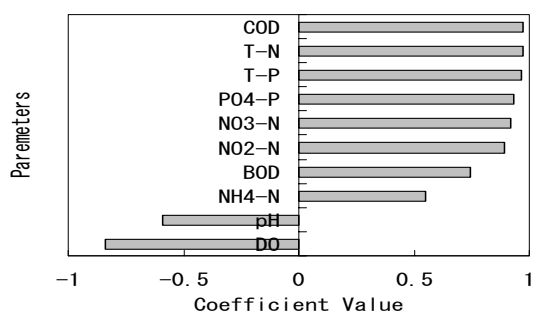


Figure2.Sorted coefficients in the first principal component

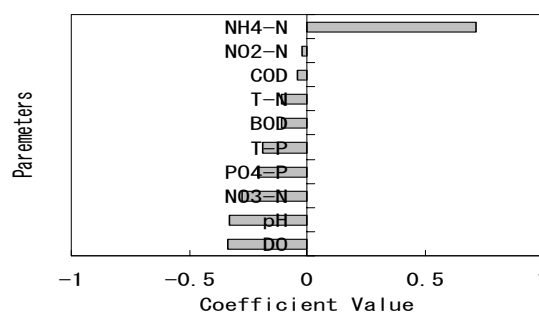


Figure3.Sorted coefficients in the second principal component

[1] 公共用水域及び地下水の水質測定結果, 東京都環境局(1997-2001 年度のデータを用いた。)

[2] 多変量解析プログラム エクセル統計 2000, <http://software.ssri.co.jp/lineup.html>

[3] ニューラルネットワークシミュレータ Neco, http://www.sccj.net/publications/cssj_jrnl.html.