

◆ 報 文 ◆

# 多変量解析を用いた魚種分布予測モデル

佐合純造\*

## 1. はじめに

自然環境保全の必要性が高まる中で事業や管理で生物の生息状況の調査やその影響の予測評価は重要なものとなっている。

河川においても様々な生物調査が行われている。たとえば、魚類については、ダム、護岸工事等の影響を評価するため、魚類調査が各地で行われている。また、全国的な魚類の分布状況を把握するため定期的に「河川水辺の国勢調査」が行われている。しかし、調査結果は継続的な河川管理に役立つような体系だったものになっておらず、各河川での生息状況やその変化が全体の姿として十分捉えきれていないように思われる。河川の上流から下流まで一貫して生物の生息状況を評価できる評価指標や実用的に活用できる予測手法の開発が今後の河川管理にとって急務と考えられる。

本報文では全国的かつ継続的に展開されている「河川水辺の国勢調査」を実務的に活用できるようにすることを念頭に、その魚類調査データと流域・河川の物理データを活用して多変量解析で簡易な方法によって任意の地点の魚種分布を予測する手法を提案する。これによって魚類データのない河川でも、あるいは外的環境が変化した場合の魚種分布の予測が可能となる。

## 2. 河川水辺の国勢調査とその活用策

「河川水辺の国勢調査」は平成2年度から開始されており、全国の一級河川と主な二級河川について、概ね5年に1回の頻度で「6項目の生物調査」、「河川調査」、「河川空間利用実態調査」が実施されている<sup>1)</sup>。このうち、6項目の生物調査は「魚介類（魚類及びエビ・カニ・貝類）」、「底生動物」、「植物」、「鳥類」、「両生類・爬虫類・哺乳類」、「陸上昆虫類等」である。今回対象とする「魚類調査」は本川及び主要支川で行われ、調査間隔は概略10kmである。調査は調査地点ごとに、できるだけ多くの魚類の確認が行えるように、「複

数の瀬・淵のつながりの区間」または「水面幅の5倍程度」の範囲を設定する。捕獲調査は対象年度中に2~3回行うことになっている<sup>2)</sup>。

調査のスケールは生物をどのレベルからモニタリングするかということから重要である。桜井は生物の生息場所（ハビタット）のスケールによって河川空間を階層別に表-1のように整理している<sup>3)</sup>。「河川水辺の国勢調査」のスケールは表-1から調査地点単位ではビオトープ階層、河川単位ではビオトープシステム階層に近いものと考えられる。

「河川水辺の国勢調査」のデータは河川環境をモニタリングする視点から次の2つの活用方法が考えられる。

- ① 固定地点を時系列的に追跡して、生息状況の変化を定量的にモニタリングする。しかし、現時点では河川環境の時系列的变化の定量評価については、データの蓄積が10年余程度しかなくデータのバラツキから考えてその活用は困難と考えられる。
- ② 調査地点の各生物の生息状況を横断的に比較評価する。これによって個々の調査地点のデータだけではわからない傾向や評価を求めることが可能である。「河川水辺の国勢調査」の魚類データでは1000を超える調査地点のデータが存在し

表-1 河川を生息場所からみた階層とスケール  
(文献3)より作成)

階層	内容	スケール
スーパーマイクロ ハビタット	石礫表面の生物膜や砂泥中の 間隙空間	1mm~1cm
マイクロ ハビタット	石礫や石積みの際間、倒木の 枝などがつくる空間	1cm~1m
ハビタット	早瀬・平瀬・淵・水際など広さを もつ空間	1m~100m
ビオトープ	ハビタットの有機的な集合に よって形成されている空間	100m ~10km
ビオトープ システム	川、湿地、草地、森林、農耕 地、市街地を包含する空間	10km ~1000km
ビオトープ ネットワーク	ビオトープシステムが連携して 存続する状態、生態的回廊	1000km~

Model Based on Multivariate Analysis to Estimate Proportion of River Fish Species

ており、これらデータがどこまで定量的な分析に耐えられるかどうかは不明であるが、これまで統一的なデータの分析はあまり行われておらず、検討する価値は大きいと考えられる。

本報文では②に基づいた検討を進めることにする。

### 3. 魚種分布予測モデルの考え方

#### (1) モデル

対象河川の各地点でどのような魚種がどの程度生息しているのか、その魚種分布を知ることは河川環境を評価する上で重要である。また、「河川水辺の国勢調査」の結果だけでは調査間隔が離散的であるため、河川全体の魚種分布の状況を把握しにくい。縦断方向に連続的な魚種分布が推定できれば、河川管理にも活用でき、一般住民への情報提供にも役立つ。このため、「河川水辺の国勢調査」の魚類データと流域・河川の物理データから、多変量解析を用いて簡易に扱えるモデルを作成して、河川の縦断方向に連続的に魚種分布を予測することを試みる。

魚種分布の決定要因については瀬・淵以下のスケールについては研究が進んでおり、生態学知見も多い。しかし、今回対象とするピオトープ階層（河川水辺の国勢調査のスケール）では、空間や時間スケールが異なるため、瀬・淵スケールの知見は活用できない。このため、モデルには判別分析や重回帰分析など多変量解析の手法を適用する。また計算には「JMP」（SAS社）と呼ばれるパソコン用統計ソフトを用いた。

なお、このような方法は魚類以外の生物データに対しても適用可能である。多変量解析による生物予測の例としては、「河川水辺の国勢調査」のデータではないが、鳥類や底生生物等への適用事例がある<sup>4), 5)</sup>。

#### (2) 評価指標

「河川水辺の国勢調査」の魚類データでは調査地点ごとに魚種別の捕獲個体数が示されている。これを評価指標として評価・予測できればよいが、この捕獲個体数は調査地点によって対象範囲の面積が異なるため、各調査地点を横断的に比較できる指標にはならない。しかし、各調査地点での全捕獲数（年間）に占める各魚種の捕獲数割合（以下、「魚種別生息比」という）や「魚種数」は、河川や調査地点横並びに比較評価できる実用的な指標となり得ると考えられる（厳密には統計的検証が必要であるが）。以上から「魚種別生息比」及

び「魚種数」を魚種分布の指標として用いることにする。なお、これらの値は年は同じでも時期（たとえば季節）によって変化するが、今回は年間平均値（年間の調査結果を合算したもの）を予測対象とした。

### 4. 重回帰分析

重回帰分析により流域・河川の各指標から各河川の「魚種数」と「その魚種別生息比」を推定する。重回帰式は線形式とする。重回帰式の目的変数を  $y$ 、説明変数を  $x_1, \dots, x_p$  として、式 (1) で表される。

$$y = b_1 \cdot x_1 + b_2 \cdot x_2 + \dots + b_p \cdot x_p + b_0 \quad (1)$$

ここで、 $b_0, b_1, \dots, b_p$  は回帰係数であり、最小2乗法を用いて求める。目的変数は任意地点の魚種数、各魚種の生息比とし、説明変数は河川・流域の物理データのうち、目的変数を予測するために①流域や沿川の特性を表す地質、土地利用、②流量や流量変動、③水質、水温、土砂量、④対象地点付近の人為要因、⑤堰、ダム等の縦断方向の施設の影響、⑥他の生物、等から選定する。説明変数の数はあまり多くならない範囲で重相関係数や寄与率が大きくなるように選定する必要がある。しかし、今回はまずは、全国どの河川でも適用できるモデル作成を優先することとして、容易に得られる共通する河道特性、流域データを説明変数に選定することにした。

### 5. 実河川への適用

ここでは多摩川の魚類調査データ（平成7年度）及び全国109河川の調査データ（平成5年度から11年度のうち新しいデータ）を用いて、多摩川本川河道に沿って一定間隔の連続的な魚種分布を求める。この際、データの使い方によって以下の2通りの方法で魚種分布を求めるモデル（重回帰式）を作成した。すなわち、

- ①個別河川（ここでは多摩川）のデータのみを用いて、その河川の重回帰式（多摩川単独モデル）
- ②全国河川の調査地点を用いて全国の任意な個別河川に適用できる重回帰式（全国河川汎用モデル）

①では多摩川だけのデータを用いて重回帰分析を行い、多摩川に合わせた重回帰式を作成した。適用が多摩川のみに限定されるため、対象河川には適合度の高いモデルが得られるが、他の河川への汎用性はない。

②では全国のデータを用いて、判別分析と重回帰分析を組み合わせることで魚種分布の予測を行った。モデルの適合度は前者に比べて低くなるものの、汎用性があり、魚類データがない河川や地点でも魚種分布の予測が可能となる。

以下では、①、②の手法を多摩川本川の魚種分布の予測に適用する。なお、ここでの手法は「河川水辺の国勢調査」実施河川であれば、多摩川に限らず、多くの河川で適用可能と考えられる。

(1) 多摩川の魚種分布概要

多摩川は東京湾に注ぐ幹川流路延長138km、流域面積1,240km<sup>2</sup>の一級河川である(図-1)。年降水量は1,500mm程度であり、流量は6~9月に豊富で、冬季には少なくなる。また中下流部では下水処理場からの放流水の割合が高い。全流域面積のうち山地が68%を占めるが、中下流部は東京都、神奈川県各市街地であり、流域内人口は約350万人である。平成7年度の河川水辺の国勢調査では本川に沿って河口0kmから61.5kmまでの本川の17調査地点で魚類調査が実施された。これらで確認された魚種は38種(淡水魚30種、回遊魚8種)で、外来魚6種(ブルーギル等)、絶滅危惧種5種(ギバチ等)が含まれていた。個体数では最も多く捕獲された魚類はオイカワで全体の32.4%を占めた。ついでウグイ21.3%、アユ7.5%、ウキゴリ5.5%の順であった。個体数のうち、純淡水魚が84.4%、回遊魚が15.6%であり、また、外来魚が0.2%、絶滅危惧種が1.5%である。

魚種の多様度は中流部で大きく、最大値は2.19(河口から22km地点)であった。優占して生息する魚種(生息比が最大の魚種)は下流から順に、ウキゴリ、ヌマチチブ、オイカワ、ウグイであった。また、アユが断続的に生息しており、特に河口から12.7km地点で多い。

(2) 多摩川単独モデル

まず、5. ①の単独モデルを検討する。多摩川本川では17調査地点で38種の魚種が生息していたが、このうち、捕獲数の多かった順に14魚種(図-2凡例の魚種)の魚種別生息比を予測の対象にした。調査地点によってバラツキがあるが生息量の67~96%をカバーしている。この実測結果を調査地点別に図-2に示す。この14魚種の魚種別生息比を目的変数とした重回帰式を17の調査地点のデータから作成する。

すなわち、14魚種の生息比を予測する14個の重回帰式を求めた。説明変数は4。述べたように、今回はモデル作成を優先させることとして、容易に収集可能な物理指標の中から、「調査地点の河口からの相対距離(調査地点の位置を幹川流路最上流=0、河口=1として0~1で表した値)」、「標高(m)」、「河道内緑被率(0~1)」、「河床勾配(log値)」、「水面幅(m)」の5個(p=5)とした。

重回帰式の回帰係数の記載は省略するが、重回帰式の適合度をみるため、重相関係数を求めたところ、図-7に示すように0.98~0.66の値を示し、生息割合が大きい魚種ほど適合度が良い傾向が見られた。この重回帰式を用いて、直轄区間1kmピッチの距離標地点の魚種分布を求めた結果を図-3に

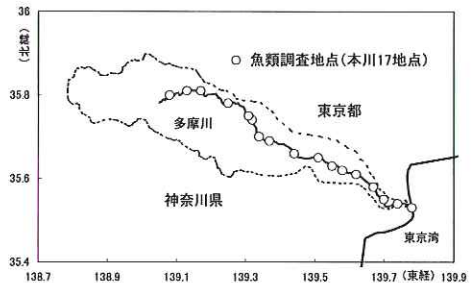


図-1 多摩川の魚類

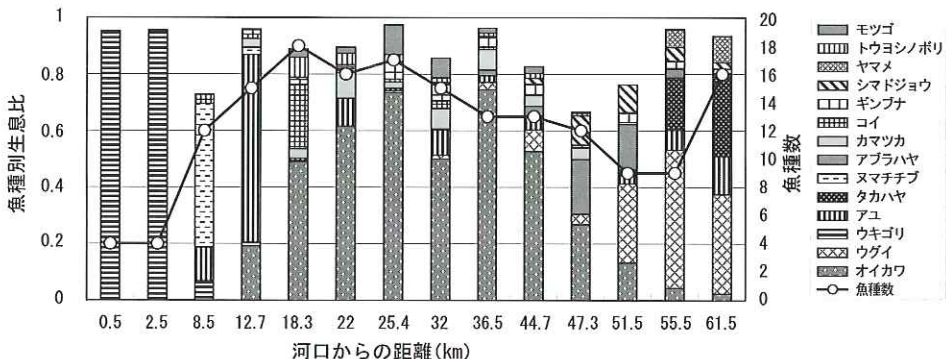


図-2 多摩川本川の実測魚種分布と魚種数 (H7データ)

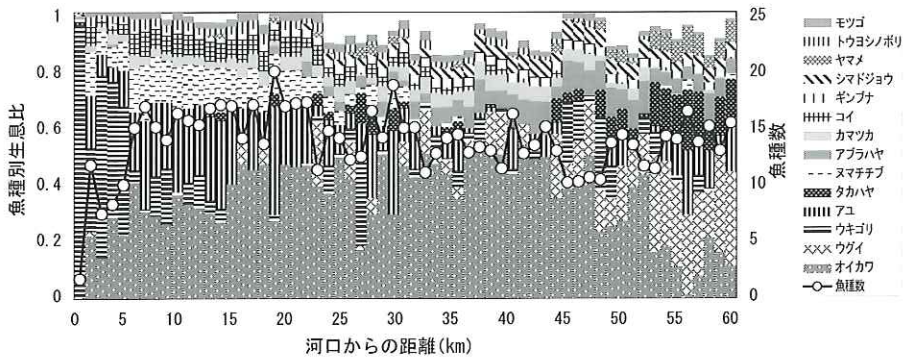


図-3 多摩川単独モデルによる予測結果 (II7データ)

示す。図-2と比較してわかるように、河川縦断的に連続的な魚種分布が表現できて調査地点の分布もほぼ再現している。本手法は調査地点と調査地点の間の分布の内挿に重回帰式を用いたものであり、地質データに例えれば地質柱状図から面的な地質断面図を作成する作業を統計的に処理したものである。当然ながら前述したように得られた重回帰式は他の河川へは適用できない。

(3) 全国河川汎用モデル

次に全国河川汎用モデルを検討する。ここでは、全国109の一級河川の本川での調査地点の魚類データ(616地点)を用いて多変量解析の手法である判別分析と重回帰分析を組み合わせて予測を行う。図-4に全国の魚類調査地点の配置状況を示す。また、図-5に予測手法のフローを示す。

全国河川のデータを用いて魚種別に生息比を予測する重回帰式を作成するが、この際、616調査地点全てのデータが統計上のサンプルとなるため、魚種によっては生息比0の調査地点も多く含まれる。このため、前述(2)の単独モデルの方法でそのまま重回帰式を作成すると、生息比0の調査地点が多い魚種では生息比0の値に引っ張られて、生息比>0の地点の予測値が小さくなり、一方で

は「生息しない」地点であっても0以上の値を予測してしまうことになる。

このため、予測値を求める前にまず、流域・河川物理データを指標にした「判別分析」によって、「生息する」(生息比>0)と「生息しない」(生息比=0)に分類分けするプロセスを追加して、魚種ごとに「生息する」、「生息しない」かを判別することにした。

判別分析は測定対象の所属グループを予測するための手法であり、ここでは魚種ごとにその魚種が生息する地点のグループ(生息比>0)、生息しない地点のグループ(生息比=0)に分類するために適用する。今回のようなグループ数が2の場合を2群の判別と言う。2群の判別を行う方法には、①マハラノビスの距離による方法、②線形判別関数による方法、③重回帰分析による方法、④ロジスティック回帰による方法、等があり、ここでは①を適用した。この方法では各魚種の生息の有無を「マハラノビスの汎距離」と呼ばれる値を用いて判断する。1つのサンプル点  $X' = (x_1, x_2, \dots, x_p)$  に対して、2つの群 ( $j=1, 2$ ) の母集団平均(重心)を  $\mu_j = (\mu_{1j}, \mu_{2j}, \dots, \mu_{pj})$ 、分散共分散行列を  $\Sigma_j$ 、その逆行列を  $\Sigma_j^{-1}$  として、下記の(2)式によって各群までのマハラノビス汎距離を計算する。

$$D_j^2 = (X - \mu_j)' \Sigma_j^{-1} (X - \mu_j) \quad (2)$$

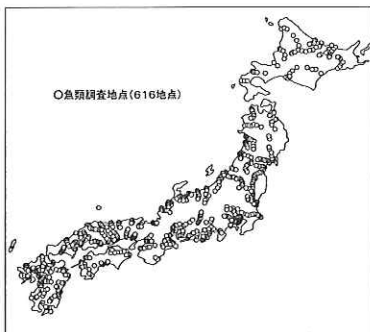


図-4 全国河川の魚類調査地点(本川)

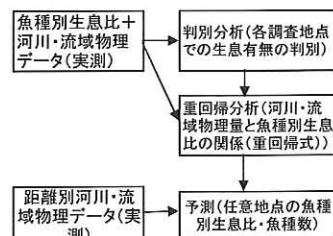


図-5 全国河川汎用モデルの計算フロー

あるサンプル点から第1群の重心へのマハラノビス距離を  $D_1^2$ 、第2群の重心へのマハラノビス距離を  $D_2^2$  と表す時、 $D_1^2 < D_2^2$  ならば第1群、 $D_1^2 > D_2^2$  ならば第2群に属すると判断する。さらに詳細な分析手法については関係文献を参照されたい<sup>6)</sup>。

判別の指標 ( $x_1, x_2, \dots, x_p$ ) は「河口からの相対距離」、「北緯」、「東経」、「標高」、「河道内緑被率」、「水面幅」の6個 ( $p=6$ ) とした。また、対象魚種は多摩川において生息比が大きい14魚種とした。

全国の河川魚類調査データ (本川616地点) に判別分析を適用して検証した結果を図-6に示す。ヤマメ、オイカワ、アブラハヤ、カマツカが比較的正しく正判別されており、14魚種平均で0.72の正判別率であった。ただし多摩川だけの判別率をみると平均の正判別率は0.64となり、全国平均に比較して低い。また、アユ、ウグイ、オイカワなどは判別率が高いが、ヌマチチブ、トウヨシノボリなどは判別率が低い。

判別分析を適用する上での課題は、「生息比=0」の調査地点は、本当は生息しているのかかわらずたまたま捕獲できなかったのか、生態学的に生息しない地点なのか不明確な点である。これを明らかにするためには地点ごとに生態調査が必要であるが、その確認は極めて困難であろう。判別分析による結果と魚の生態特性との対比や生態学的知見を取り入れた判別手法については今後とり組む必要がある。

次に、「生息比>0」の地点だけを対象にして重回帰式を適用した。この重回帰式では説明変数は「河口からの相対距離」、「標高」、「河床勾配 (log 値)」、「河道内緑被率」、「水面幅」(ここまでは多摩川単独の重回帰式と同じ)、これらに加えて「北緯」、さらに流域形態の指標として調査地点の所

属する河川の「流域面積」を加えることにした。流域形態の指標に流域面積を採用したのは、魚種数と流域面積との相関が高いことが知られているためである<sup>7)</sup>。

図-7に魚種ごとに作成した重回帰式の重回相関係数を示す。全体的に多摩川単独式の場合にくらべて適合度が高いとは言えないが、生息比の大きい魚種ほど相関係数の大きいことは多摩川単独式と同様である。以上の方法で予測した結果を図-8に示す。また、河川環境が変化した場合を想定した魚種分布の変化を予測した。すなわち、緑被率が全区間にわたって0 (植生がなくなった場合) の仮定の魚種分布の予測結果を図-9に示す。全区間にわたって魚種数が減少するとともに、魚種数も縦断的な変化が小さくなり魚種分布が単調になることがわかる。この結果はあくまで、統計的推定であり、当然ながら生態学的知見との十分な照査が必要であろう。

## 6. まとめ

河川環境の評価のために様々な調査や分析が行われているが、その多くが個別論に留まっており、一般化などの有効活用が図られていない。また、有効活用のために必要となる生物生息のモデル化は従来から工学で扱ってきた物理法則だけでは対応できないし、要素に分けた分析手法は適用しにくい。

本報文では「河川水辺の国勢調査」を活用して、全国の一級河川の魚類データに多変量解析を適用して、魚種分布を河川縦断方向に連続的に予測するモデルを作成して多摩川に適用したところ、実用的可能性のあるモデルと予測結果を得ることができた。今回の方法は多摩川以外の河川や魚類以外の生物にも適用可能と考えられる。

なお、今回は入手しやすい既存データを用いるこ

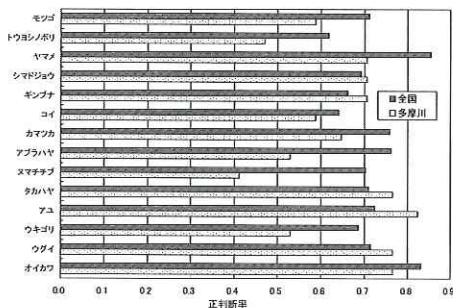


図-6 判別分析による正判別率

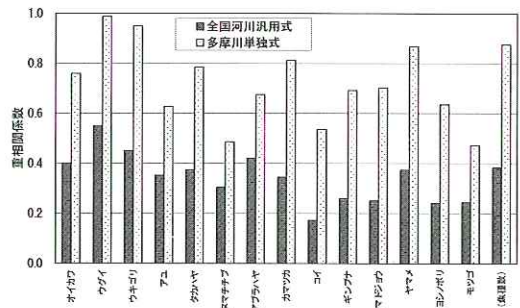


図-7 重回帰式の重回相関係数

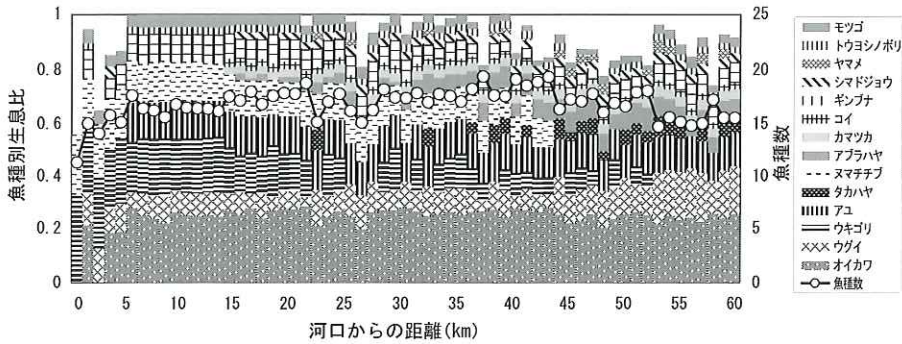


図-8 全国河川汎用モデルによる現状予測結果 (H7・多摩川本川)

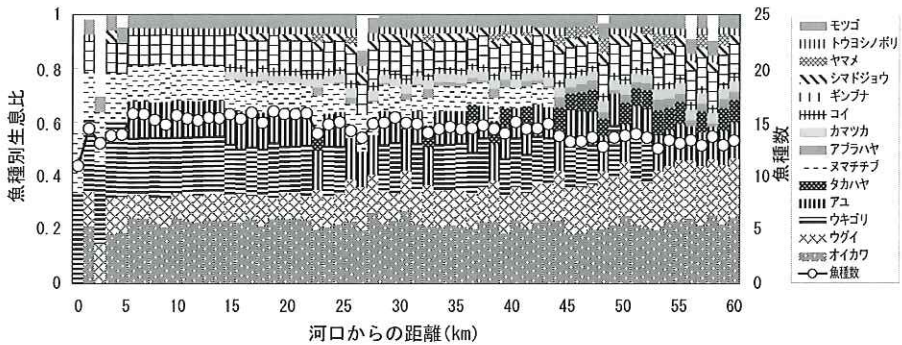


図-9 全国河川汎用モデルによる緑被率0の予測結果 (H7・多摩川本川)

とを優先しており、モデル式やパラメータ（説明変数）については十分な検討がされていない。特に、水量、水質等の重要な指標が説明変数に取り込まれていない。また、生物相互の関係、堰などの河川縦断方向の影響、支川の影響なども考慮していない。今後、本モデルの妥当性及精度アップを図るためには、これらの検討が大きな課題である。

実用性のある生物生息モデルやモニタリング手法の確立のためには、個別の生態学的知見の追求にとらわれるばかりでなく、統計分析を用いたマクロな解析や生物データと物理データをセットにした全国的なデータベースの構築が必要と考えられる。

本報文は河川水辺の国勢調査年鑑（平成2・3年度～平成11年度）のデータを用いて行ったものである。関係された方々に感謝の意を表する。

参考文献

- 1) 国土交通省河川局監修、(財)リバーフロント整備センター編集：河川水辺の国勢調査年鑑（平成2・3年度～平成11年度）、山海堂、1993～2001
- 2) 建設省河川局河川環境課監修：河川水辺の国勢調査マニュアル・河川版（生物調査編）、(財)リバーフロント整備センター、1997.4.

- 3) 桜井善雄：川づくりとすみ場の保全、信山サイテック、2003.6
- 4) 百瀬浩、藤原宣夫：広域的な生態系予測手法の開発、土木技術資料45-5、2003.5
- 5) 湯浅岳史ほか：多変量解析を用いた水生生物の出現予測、第12回環境情報科学論文集、1998
- 6) たとえば、田中豊、脇本和昌：多変量解析法、現代数学社、1983.5.
- 7) 佐合純造、永井明博：河川水辺の国勢調査結果を用いた全国河川の魚種数の特性とその評価手法、土木学会論文集No.727/VII-26、2003.2

佐合純造\*



独立行政法人土木研究所水循環研究グループ長、博（環境理工学）