

木山川水質指標のパネル分析について

崇城大学 工学部 学生員 ○ 芳坂 真一郎 正 員 田代 敬大
崇城大学 非会員 岩原 正宜 非会員 谷口 智穂

1. はじめに

一般に、野外調査の資料は様々な要因が関係する。本研究の緑川水系支川を対象とした「木山川水質調査」¹⁾の資料も地点・時間(季節)等の様々な要因の影響が混入していると推測される。ただ、この調査は採水地点 30 箇所を対象に、平成 20 年を通してほぼ毎月実施され、測定された水質指標は地点・時間がセットになったパネルデータとなっている。

本研究はパネルデータの特性を活かし、木山川の大腸菌群数・BOD 等の水質指標について、直接は観測されない地点効果・時間効果を考慮した回帰分析すなわちパネル分析を試みるものである。

2. パネル分析の概要

パネル分析は計量経済学で発達した手法で、直接には観測できない要因を明示的に検討する場合など、きわめて有効な分析手法である²⁾。

普通の回帰モデルでは

$$y_k = a_0 + a_1 x_k + e_k \quad (1)$$

において、誤差項 e_k は平均 0、分散一定等が仮定される。これに対し、パネル分析では誤差項 e_{it} に

$$e_{it} = \alpha_i + \beta_t + \varepsilon_{it} \quad (2)$$

ここに、 i : 調査地点番号 $i = 1, 2, \dots, N$

t : 調査時点番号 $t = 1, 2, \dots, T$

地点効果 α_i 、時間効果 β_t 、新たな誤差項 ε_{it} という式(2)の構造を想定する。すなわち、普通の回帰モデルでは誤差項に含まれてしまう地点効果・時間効果を、パネル分析では明示的に析出するものといえる。

ここで、効果 α_i 、 β_t と説明変数 x_{it} とに相関がある場合を固定効果モデル、無相関の場合を変量効果モデルと呼ぶが、本研究では後述のモデルについて両モデルを作成してハウスマン検定を実施したところいずれも有意水準 5%で変量効果モデルが棄却されたので、固定効果モデルについて述べる。

固定効果モデルでは、次式のパラメタを推定する。

$$y_{it} = a_0 + a_1 x_{it} + \sum_{i=1}^N \alpha_i \delta_i + \sum_{t=1}^T \beta_t \delta_t + \varepsilon_{it} \quad (3)$$

地点ダミー変数 δ_i は対象地点 j が $i=j$ の場合は 1、 $i \neq j$ の場合は 0 をとる。時間ダミー変数 δ_t も同様である。したがって時間的に一定の地点固定効果の推定値 $\hat{\alpha}_i$ は平均 0 からの差、地点的に一定の時間固定効果の推定値 $\hat{\beta}_t$ は平均 0 からの差を表わしている。

3. ATP による大腸菌群数の回帰分析

生細胞が有するアデノシン三リン酸(以下 ATP)の測定は迅速簡便であり、対象河川の大腸菌群数と ATP の関係を検討する。調査は第 4 回～第 11 回、採水点数 30、サンプル数 240 である。 y は大腸菌群数[数/ml]、 x は ATP 発光強度[RLU/100 μ l]である。

(1) 普通の回帰モデル (plain OLS)

決定係数 $R^2 = 0.267$ 調整済み決定係数 $\bar{R}^2 = 0.264$

AIC = 18.623

$$y = 2006.046 + 0.116 x \quad (4)$$

(2) 固定効果モデル

決定係数 $R^2 = 0.557$ 調整済み決定係数 $\bar{R}^2 = 0.475$

AIC = 18.419

$$y = 2288.142 + 0.064 x + \hat{\alpha}_i \delta_i + \hat{\beta}_t \delta_t \quad (5)$$

式(4)と式(5)を比較すると、決定係数 R^2 は 0.267 から 0.557 へ、自由度調整済み決定係数 $\bar{R}^2$ は 0.264 から 0.475 へと上昇している。固定効果モデルの推定値 $\hat{y}_{it}$ は地点・時点固定効果を含むため高めとなる傾向があるが、赤池情報量基準 AIC から式(5)が優れていることになる。偏回帰係数は有意水準 5%でいずれも有意である。

式(5)で興味深いのは、各地点別の地点固定効果の推定値 $\hat{\alpha}_i$ の分布である(図 1)。地点固定効果で平均 0 以上の大腸菌群数が定数項に加わる採水点は上流域の布田川 5～9 番、長田川 6・60・600 番と秋津川合流点 25 番であり、その他の採水点は負値となっている。大腸菌群数・ATP それぞれの原データにおいてもこれらの採水点では大きなピークが見られるが、式(5)の大腸菌群数・ATP の相対的關係においても大きな値を示していることになる。他の水質指標の検討においても同様の傾向がみられることから^{3) 4) ~ 8)}、

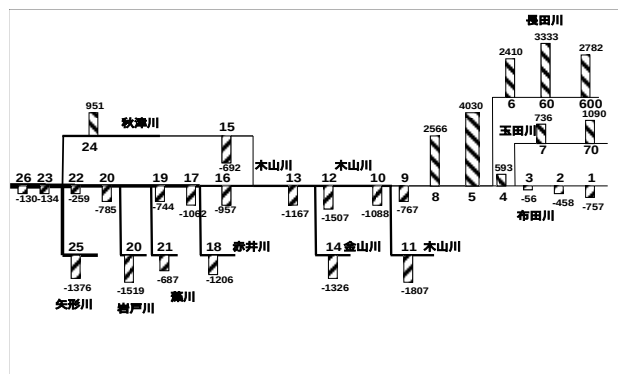


図1 式(5)の地点固定効果

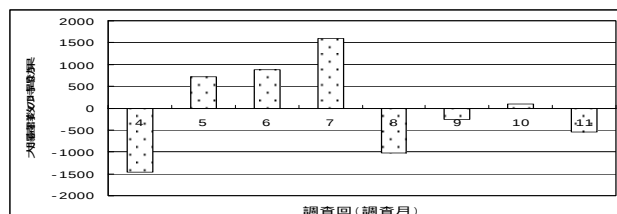


図2 式(5)の時間固定効果

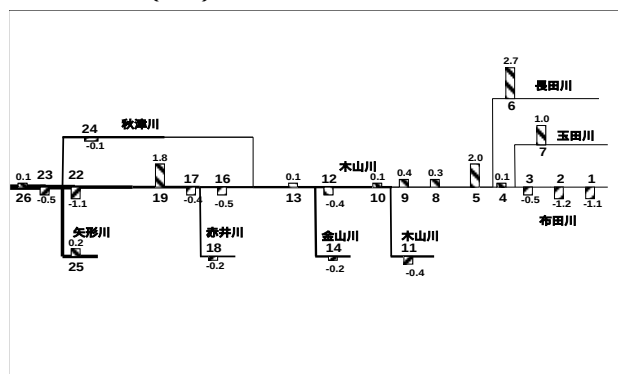


図3 式(7)の地点固定効果

これらは固有の汚濁源になっていると推測される。

一方、図2は時間固定効果の推定値 $\hat{\beta}_t$ である。平均0以上の5~7回(月)は式(5)の定数項に加わることで右辺第1項と第2項の線形回帰式は上方にシフトし、8~9回(月)では下方にシフトすることになる。これは、原データの大腸菌群数のピークが5~7回(月)頃の採水点が多いのに対し、ATPのピークは少し遅れて7~9回(月)頃の採水点が多いことに照応しているものと推測される。

4. CODによるBODの回帰分析

対象河川のBODとCODとの関係を検討する。第1回~第8回調査、採水地点数は26箇所、サンプル数176であり、 y はBOD[mg/l]、 x はCOD[mg/l]である。

(1) 普通の回帰モデル (plain OLS)

決定係数 $R^2 = 0.198$ 調整済み決定係数 $\bar{R}^2 = 0.193$

AIC = 4.666

$$y = 2.744 + 0.526x \quad (6)$$

(2) 固定効果モデル

決定係数 $R^2 = 0.534$ 調整済み決定係数 $\bar{R}^2 = 0.442$

AIC = 4.440

$$y = 2.944 + 0.385x + \hat{\alpha}_i\delta_i + \hat{\beta}_t\delta_t \quad (7)$$

一般にCODとBODの相関は高いといわれるが、本調査では多数の採水点のためBODを簡易的に測定したこともあり、式(6)の決定係数 $R^2 = 0.198$ のように低い値となった。ただこれも、式(7)のように固定効果モデルを用いると決定係数 R^2 は0.198から0.534へ、自由度調整済み決定係数 $\bar{R}^2$ は0.193から0.442へと増加する。赤池情報量基準AICからも式(7)が選択されることになる。

各地点別の地点固定効果は、図3のようになる。ここでも地点固定効果が著しいのは、上流域の布田川・長田川と秋津川合流点である。これに対し、時間固定効果は明瞭な傾向は見出せないようである。

5. おわりに

本研究では木山川の水質指標にパネル分析を試みた。固定効果モデルを適用すると、モデルの当てはまり度合いが向上するとともに、直接は観測できない採水地点固有の固定効果や時間効果を析出できた。地点固定効果からは、上流域の布田川・長田川と秋津川合流点固有の汚濁効果が明らかとなった。

【謝辞】 本研究は崇城大学工学部エコデザイン学科、ナノサイエンス学科吉田研究室・西田研究室、岩原正宜特任教授・谷口智穂研究員との共同研究「木山川水質調査」の一部である。調査にご協力いただいた関係行政機関・市民の方々・本学教職員および学生諸君に対し、記して謝する次第である。

【参考文献】 1) 澤村他「木山川水系の水質汚濁負荷について」平成20年度土木学会西部支部後援概要集。2) 松浦他「EViewsによる計量経済学入門」東洋経済新報社、2005年。3) 松井他「木山川水質の総合指標による採水点分類について」平成20年度土木学会西部支部後援概要集。4) ~8) 吉田他「化学分析法による河川水の水質汚濁の検討」、宮城他「TOC測定による河川水の水質汚濁の検討」、牧他「COD測定による河川水の水質汚濁の検討」、須崎他「硝酸イオン測定による河川水の水質汚濁の検討」、百田他「電導度測定による河川水の水質汚濁の検討」第45回化学関連支部九州合同大会講演要旨集(2008)。