

## 流出負荷量の不偏推定法の一般化 A generalized unbiased load estimation method

○多田 明夫\*, 田中丸治哉\*  
○Akio TADA\* and Haruya TANAKAMARU\*

### 1. はじめに

河川を流下する物質の特定期間における総流出量は、物質の供給源側から見て流出負荷量 (load) と呼ばれる。この数値は面源汚濁負荷量価や面源原単位の算定根拠であり、流域レベルの物質収支を検討する上で重要なものである。しかしこれまでこの値を偏りなく推定する方法が存在していなかった。著者らはこの問題に対し、べき乗型 LQ 式と重点的サンプリング (Importance sampling, IS) 法を組み合わせることで、流出負荷量の不偏推定を可能とした。

しかしながらこの方法は LQ 式を決定するための事前のサンプリングと、決定された LQ 式にもとづく瞬間流出負荷推定量の大きさに比例した確率でのサンプリング— PPS (probability proportional to sample size) なサンプリング—、の二相抽出 (two-phase sampling) を必要とし、現地適用が面倒である。加えて、PPS なサンプリングの実現には一様乱数が必要とされることも現地適用にあたっての障害となる。さらに事前サンプリングと PPS サンプリングの間に LQ 式に有意な変化が生じた場合には、負荷量推定量が偏る。

上記の課題をすべて取り除くことが、当該手法の現地適用の上で重要である。このために著者らは新たに、水質サンプリング法を上記のような特定の方法に限定しない、一般化された流出負荷量の不偏推定法として IR (Importance resampling) 法を提案する。本報告ではこの手法を説明し、観測データに基づき検証する。

### 2. IS 法と IR 法

2.1 IS 法 ここで扱う流出負荷推定的前提条件は、流量データは高頻度で連続的に得られているが、水質濃度 (瞬間流出負荷量) の観測データは量的に不十分であるというものである。このような条件下では、流出負荷量を求める期間中のすべての流量観測値 ( $N$  個あるものとする) に対して、瞬間流出負荷推定量  $\hat{l}_i$  が、べき乗型 LQ 式 ( $\hat{l}_i = a_0 q_i^{a_1}$ ,  $q_i$  は瞬間流量、 $a_0 \cdot a_1$  はパラメータ) により行われる。IS 法では、この  $\hat{l}_i$  の大

きさに比例した確率で  $n_{is}$  個の水質サンプリングを行う。サンプリングは  $\hat{l}_i$  の積算軸 ( $0, \sum \hat{l}_i$ ] 上に  $n_{is}$  個の一様乱数を発生させ、それぞれ相当する時刻に行うことになる (Fig.1)。このような  $n_{is}$  個のデータから、負荷量の不偏推定量  $\hat{L}$  が次式で与えられる。

$$\hat{L} = \exp a_0 \times \left( \sum_{i=1}^N q_i^{a_1} \right) \times \left[ \frac{1}{n_{is}} \sum_{\eta=1}^{n_{is}} \exp(e_{\eta}) \right] \quad (1)$$

ここで、 $e_{\eta}$  は対数空間上での PPS なサンプリングデータに対する LQ 式の回帰残差である。(1) 式は点推定量を与えるが、右辺第三因子の信頼区間を考慮することで、 $\hat{L}$  の信頼区間を構成できる。

2.2 IR 法 IS 法で最も重要な点は、LQ 式などの瞬間流出負荷推定量の大きさに比例した確率での PPS サンプリングの実現である。これは Fig.1 にも示したように、積算負荷量軸上に等確率でサンプリング点を配置することで実現できる。逆に言えば、どのようなサンプリング法で得られたデータであっても、この軸上で一様乱数と同様に見なせるような配置をとるようなデータの取捨選択を行えば、結果的に流出負荷量の不偏推定量を得ることができるはずである。

この一様乱数と見なせるかの判断は、適合度検定 (例えば Stephens, 1986) で用いられる次式の Anderson-Darling (AD) 検定量  $A^2$  の値から判断できる。

$$A^2 = -\sum_{j=1}^n \left\{ (2j-1) \left[ \ln p_j + \ln (1-p_{n+1-j}) \right] \right\} / n - n \quad (2)$$

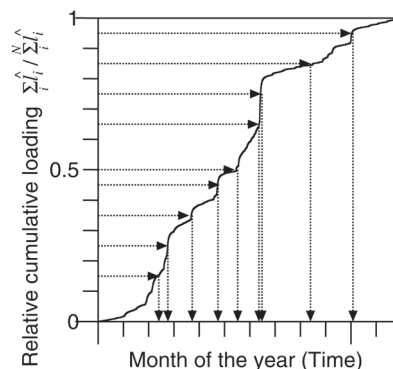


Fig.1 uniform PPS sampling points on axis of relative value of cumulative loading rate estimates

(所属) \* 神戸大学大学院農学研究科, Graduate School of Agricultural Science, Kobe University

(キーワード) 面源, 負荷量, 不偏推定, 一般化

ここで、 $n$  はサンプリング点の数、 $p_i$  はサンプリング点の相対積算負荷推定量軸 (0, 1] 上の値を昇順に並べた数列 ( $i = 1 \sim n$ ) である。相対積算負荷推定量の値は、積算負荷推定量の値を当該期間全体での合計量  $\sum_{i=1}^n A_i$  で除して得られる。なお、 $A^2$  の値は一様乱数列では小さな値をとり、そうでない場合には単調に増大する。従って、以下の手順でサンプルの取捨選択により乱数列を実現できる。まず全サンプリング点を用いて  $A^2$  の値を計算し、最も  $A^2$  の値を大きくしているサンプリング点を 1 つ除去し、再度 ( $n - 1$ ) 個のサンプルから  $A^2$  の値を評価する。この  $A^2$  の値が一様乱数と見なすことのできる閾値  $A^2_T$  よりも小さくなるまで、この繰り返しでサンプル数を減ずればよい。これにより、重点的なりサンプリングが実現できる。

この方法の利点は、AD 検定を適用するサンプリング点の数列はどのようなものでもよい、すなわちどのようなサンプリング法に基づく水質観測値に対しても適用可能である点にある。また積算負荷量軸を決定するための LQ 式は、すでに手元に得られている水質濃度と流量のデータセットから計算すればよい。このために、IS 法で起きうるような、二つのサンプリングの間での LQ 式の有意な変動にも煩わされない。

### 3. サンプリング (モニタリング) 法提案と適用結果

実際に本手法を奈良県五條市の小山林流域 (12.14ha) で観測された 10 分間隔の水質・流量データに対して適用した。対象項目は  $K^+$ 、 $Cl^-$ 、 $Na^+$ 、SS である。負荷量の計算期間は前 3 項目で 2009/5/12 ~ 2011/4/28、SS で 2012/6/7 ~ 2014/8/26 である。なお期間中の総流出高はそれぞれ 1,348、1,330、1,152、640mm であった。

IR 法の適用にあたり、効率的な現地水質モニタリング法もあわせて提案する。サンプリング点は相対積算負荷推定量軸上での一様配置により決定される。流量データが同じであれば、この軸は LQ 式の指数パラメータ  $a_1$  により決定される。この  $a_1$  の値は多くの項目で 0.8 ~ 2.2 程度の値である (例えば二瓶ら、2011)。従って、その中央値である 1.5 前後の  $a_1$  の値で軸を構成し、その軸上で一定値毎にサンプリングをすれば、効率よく様々な項目の負荷量の不偏推定ができよう。しかし  $a_1$  が大きいと大出水時に採水瓶が大量に必要となるため、現実的には  $a_1 = 1.0$ 、すなわち一定積算流量  $\Delta q$  毎の水質サンプリングを提案したい。

実際に、 $\Delta q = 50, 20, 10, 5$  mm でのサンプリングデータについて IR 法での点推定と 95% 信頼区間の構

成を行った。評価回数は、最初の  $\Delta q$  区間内に含まれる流量の観測値の数となる。この評価回数のうち、信頼区間内に負荷量の真値 (全 10 分間瞬間流出負荷量の積算値) を含む確率 (被覆確率) を Table 1 に、点推定と信頼区間の上下限の分布を Fig.2 に示した。これらの結果より、IR 法では偏りのない負荷量推定が行えること、一定積算流量毎の採水の場合被覆確率が高くなることが示された。詳細な説明は略すが、この高い被覆確率は、IR 法による信頼区間が IS 法よりも若干狭く、ばらつきも小さくなることによる。なお、サンプル数のうち一様乱数化後に採択された標本の割合は、溶存項目 ( $a_1=0.88 \sim 0.97$ ) で 90 ~ 100%、SS ( $a_1 = 1.33$ ) で 40% 強であった。一定積算流量毎では無く定期サンプリングデータ ( $a_1 = 0$  に相当) を元にするると、この採択率はもっと小さくなる。

謝辞 本研究は科学研究費補助金 (No.15K07646) の助成を受けて実施したものである。

#### 参考・引用文献

Stephens, M. A. (1986); Goodness-of-fit techniques. Series Statistics: Textbooks and Monographs, 68.

二瓶泰雄・滝岡健太郎 (2011); 出水時水質データベースに基づく L-Q 式推定法の検討、第 14 回水環境学会シンポジウム講演集、pp.235-236

Table 1 Coverage rates of 95% confidence intervals of stream loads by IR methods

$\Delta q$ (mm)	$K^+$	$Cl^-$	$Na^+$	SS
50	99.9	100	99.3	99.6
20	99.6	100	99.9	96.9
10	100	100	100	99.4
5	100	100	99.6	98.4

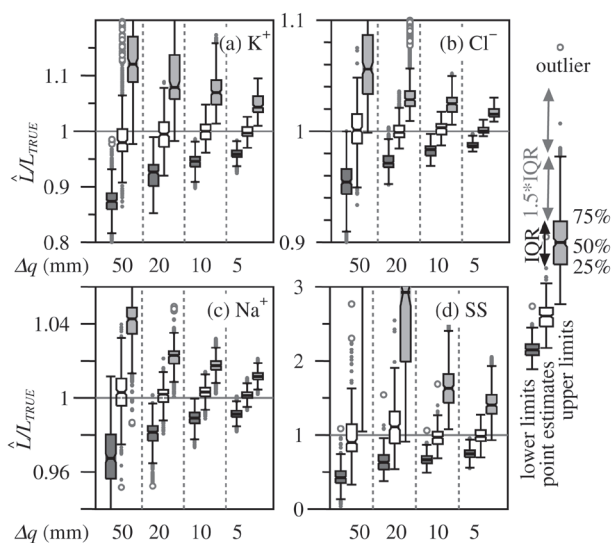


Fig.2 Point estimates and interval estimates of stream loads by IR methods