

## 年間河川負荷量の正確な信頼区間の提案 Precise confidence interval for annual river loads

○多田 明夫\*・田中丸治哉\*

○Akio TADA\* and Haruya TANAKAMARU\*

1. はじめに 年単位などの長期間の河川負荷量の正確な推定は、農業、工業、生活排水などの人間活動の、水環境への影響評価の基礎数値である。著者は Horvitz-Thompson (HT) 推定量に基づく不偏推定法と信頼区間の構成法を提案したが (Tada and Tanakamaru, 2022), この方法には、月1度(年12個)の定期調査データに基づいた場合、懸濁物質関連の水質項目(SS, TPなど)で95%信頼区間の幅が非常に広くなり、かつ真値を信頼区間が含む確率である被覆確率の値が90%に満たないという課題があった。特に真値の数百倍にもなりうる信頼区間幅は提案法の実用性を大きく損なう。我が国の公共水域(河川)の水質観測頻度が月1度であることと年間数十程度の水質調査回数が多いことを考えると、そのようなデータからより正確な信頼区間を構成することの重要性は大きい。本発表では、より正確な信頼区間の構成法として、不偏推定量であるHT推定量とLQ式を組み合わせた bias-corrected regression estimator(BCRE)を提案し、実際のデータへの適用結果について紹介する。

2. 方法 BCREは回帰式とHT推定量を組み合わせた regression estimator (Särndal et al., 1992) の一種であり、HT推定量の推定精度を向上させるものである。河川負荷量の計算期間において、 $n$ 個の瞬間負荷量 $l_i$ (水質試料採取時の濃度×流量)のデータ( $i=1\sim n$ )が確率 $p$ で得られているものとする。このサンプリング確率 $p$ は実際に標本を収集していない時間も含めた母集団の $N$ 個の単位時間すべてに対して定義され、 $j$ 番目の要素(単位時間)に対して $p_j$ と表記する( $j=1\sim N$ )。また日流量などの補助変量の情報からLQ式により日負荷量の推定量 $\hat{l}_j$ が得られるものとする。このとき、期間中の総流出負荷量のBCRE推定量 $L_B$ は(1)および(2)式で与えられる。式中で $L_{HT}$ ,  $L_M$ ,  $L_{MHT}$ はそれぞれ年負荷量のHT推定量, LQ式による年負荷量の推定量,  $L_M$ のHT推定量であり、 $\hat{\gamma}$ は $L_M$ の偏りの推定量

$$L_B = L_{HT} + \hat{\gamma}(L_M - L_{MHT}) = (L_{HT}/L_{MHT})L_M \quad (1)$$

( $=L_{MHT}/L_{HT}$ )である。LQ式には

Loadestモデル(Runkelら, 2004)のmodel 1, 2, 4から最適なモデルを選択利用した。

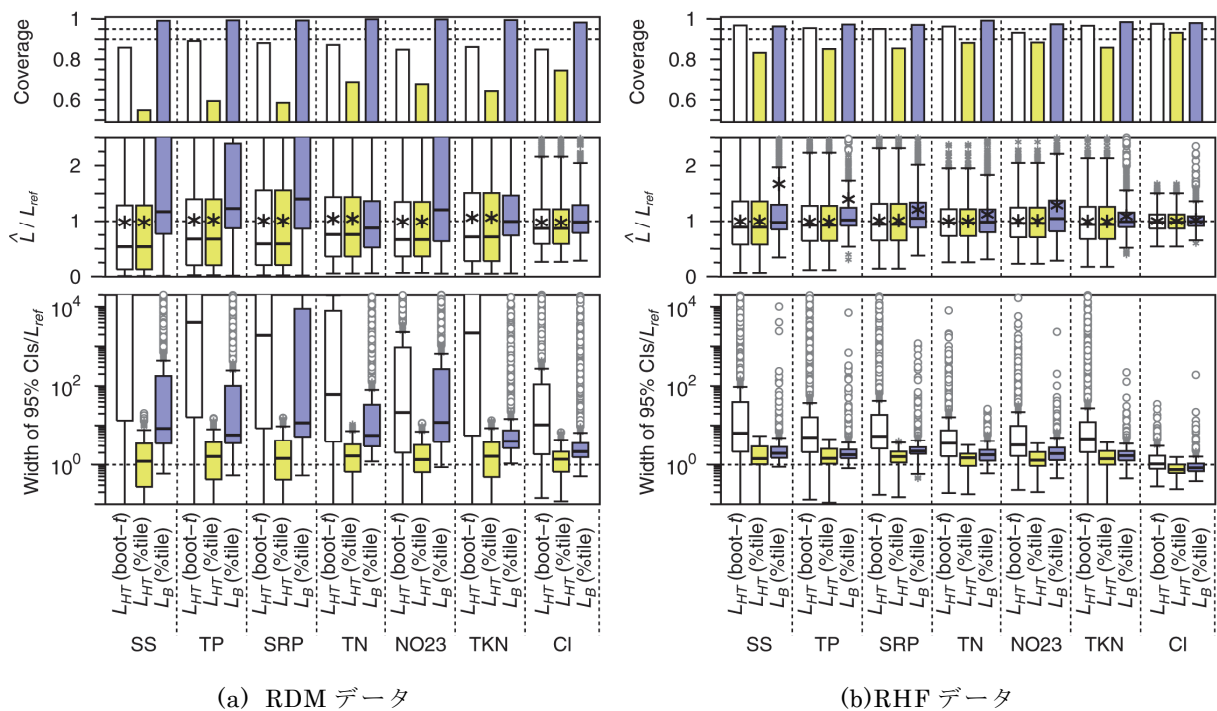
$$L_{HT} = \sum_{i=1}^n \frac{l_i}{p_i}, \quad L_{MHT} = \sum_{i=1}^n \frac{\hat{l}_i}{p_i} \quad (2)$$

実際に米国の複数流域での日単位の流量データ(USGS, 2016)と水質データ(Heidelberg University, 2019)を用い、25の流域・年のデータセットに対して、月1回の定期調査相当の(月内の採水日をランダムに選ぶ)方法で収集したデータ(RDMデータ)と、この年12個のデータに高流量時(年間日流量の上位20パーセント)からランダムに8標本要素を追加した年20個のデータ(RHFデータ)に基づき、SS, TP, SRP (Soluble reactive phosphorus), TN, NO<sub>2+3</sub> (亜硝酸塩+硝酸塩), TKN (総ケルダ

(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) 河川, 流出負荷量, 信頼区間, LQ式, regression estimator

ール態窒素), CI の年負荷量を推定した。信頼区間の構成は bootstrap パーセンタイル法に従った。推定量の分布は, 標本を母集団から Monte Carlo 法でリサンプリング (繰り返し回数 2,000 回) して計算した。

3. 結果と考察 Fig.1 に Rock Creek 流域 (面積 90km<sup>2</sup> で 72%が農地) の年負荷量の推定量と中央 95%信頼区間の被覆確率を示した。比較のために, bootstrap-*t* 法と bootstrap パーセンタイル法による HT 推定の結果 ( $L_{HT}$ ) も示した。この図からも明らかなように, HT 推定量はすべての項目に対して不偏推定量を与える一方で, いくつかの項目で被覆確率が信頼水準の 95%を下回っている。BCRE ( $L_B$ ) はそのような項目でも 90%以上の被覆確率を与え, かつ信頼区間の幅も  $L_{HT}$  から大きく縮小している。また, BCRE は不偏推定量ではなく, 中央値がより不偏 (真値に近い) 推定量である。さらに高流量時のデータの追加は不確かさの削減に大きく貢献することも明らかである。



(a) RDM データ (b) RHF データ  
Fig.1 年単位負荷量の推定結果 (Rock creek 2015 データ,  $L_{ref}$  は真値)

4. おわりに HT 推定量は unbiased but imprecise という問題があったが, BCRE は less-biased but precise である。本手法を利用すれば, 多くの国で蓄積されている低頻度の長期間水質モニタリングデータからより有益な情報を引き出すことが可能となろう。

参考・引用文献

Heidelberg University. (2019). Heidelberg University National Center for Water Quality Research Tributary data download [data file]. Retrieved from <https://ncwqr.org/monitoring/data>

Runkel, R.L., Crawford, C.G., Cohn, T.A., 2004. Load Estimator (LOADEST): A FORTRAN program for estimating constituent loads in streams and rivers. U. S. Geol. Surv. Tech. and Methods, Book 4, chap. A5, 69 pp., U. S. Geol. Surv., Denver, Colo.

Särndal, C.E., Swensson, B., Wretman, J., 1992. Model assisted survey sampling. Springer Series in Statistics. Springer-Verlag, New York.

Tada, A., Tanakamaru, H., 2022. Unbiased estimates and confidence intervals of riverine loads for low-frequency water quality monitoring strategies. Water Resour. Res., 58(5). DOI:10.1029/2022wr031941

U.S. Geological Survey. (2016). National water information system data available on the World Wide Web (USGS water data for the nation) [Data file]. Retrieved from <http://waterdata.usgs.gov/nwis/>, <https://doi.org/10.5066/F7P55KJN>