

流量比例コンポジットサンプルは正しい流量荷重平均濃度を与えるか？

あるいは、面源負荷量推定でこれまで我々は何を間違えていたのか？

Can a flow proportional composite sample afford to provide the accurate flow-weighted concentration, or: what did we misunderstand in estimating nonpoint source loads?

○多田 明夫*, 田中丸治哉*

○Akio TADA* and Haruya TANAKAMARU*

1. はじめに

面源からの流出負荷量の値は、原単位の算定の基礎数値であるばかりではなく、流域レベルでの水質管理においても、さらには流域レベルの水質対策の効果の定量的評価においても重要である。このため、この値を正確に推定する必要がある、これまで国内外を通じて種々の負荷量推定法が提案されてきた（ここでいう“正確”とは、真の値に対して偏りがなく、推定値の誤差が適切に推定されている状態を指す）。しかし、未だに“正確な”流出負荷量の推定法は確立されていない。筆者らはこの問題に長年取り組んできたが、問題の所在を整理できたので、流量比例コンポジットサンプルを例に、この問題について説明したいと思う。

なおここで扱う問題は、「ある河川において、流量は時間的に十分に高い頻度で連続的に測定されているが、水質濃度の観測値は労力やコストの面から十分な頻度では観測できず、限られたデータ量である」という条件下で、例えば年間の流出負荷量を推定するというものである。流量や水質濃度値自体の測定やサンプリング誤差はないものとする。また流量比例コンポジットサンプルについては、容器保存中の変質や混合による水質変化はなく、さらに、いうまでもないことであるが、河川濃度は一定値ではないものとする。

2. 流量比例コンポジットサンプル

河川の水質試料の採取方法には、大別してグラブサンプリングとコンポジットサンプリングがある。前者は河川の流水を採水瓶などに採取する一般的な方法であり、後者は、ある基準に従って河川の流水を分取して採水瓶の中で混合し、何らかの代表値を有するような水質試料を作成する方法である。コンポジットサンプルの作成方法は、一定時間毎に一定量の河川水を採取して混合する時間コンポジットサンプルと、流量に比例して河川流水を分取し、採水瓶中で混合する流量

比例コンポジットサンプルに分類される。時間コンポジットサンプルは時間的な平均水質、流量比例コンポジットサンプルは流量荷重平均濃度を与えるので、後者に期間中の総流量を乗ずれば期間中の流出負荷量の値を得ることができる（はずである）。

それでは、流量比例コンポジットサンプルは正しい流量荷重平均濃度を与えることができるであろうか。この答えは Yes であり、また No でもある。その答えは流量比例コンポジットサンプルの作成方法に依存する。もし流量比例コンポジットサンプルを“連続的に”作成することができるのであれば、期間 $t_1 \sim t_2$ における流量コンポジットサンプルの流量荷重平均濃度 C_w は次式で得られることになり、“正確”な測定値を得ることができる。

$$C_w = \frac{\int_{t_1}^{t_2} c_i(\alpha q_i) dt}{\int_{t_1}^{t_2} \alpha q_i dt} = \frac{\int_{t_1}^{t_2} c_i q_i dt}{\int_{t_1}^{t_2} q_i dt} \quad (1)$$

ここで、 α は採水量の流量に対する比率、 c_i 、 q_i はそれぞれ時刻 i における水質濃度と瞬間流量である。期間中の総流出負荷量 L はこの流量荷重平均濃度に期間中の総流出量を乗じた次式で与えられる。これ以降の議論は、この流出負荷量の話で進めることとしたい。

$$L = C_w \int_{t_1}^{t_2} q_i dt = \int_{t_1}^{t_2} c_i q_i dt = \int_{t_1}^{t_2} l_i dt \quad (2)$$

(2) 式の c_i を河川の水質濃度、 q_i を河川の瞬間流量と見なせば、結局この問題は面源負荷量の推定と全く同じであることに気付くであろう。

3. 問題の所在

3.1 流量比例コンポジットサンプルの離散化

現実の世界では、離散化された流量比例コンポジットサンプリングがしばしば行われている。例えば、流量に比例して時々刻々採水量を変化させるのではなく、一定積算流量毎に定まった量を採水し、容器で

(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, 重点的サンプリング, 流出負荷量, 不偏推定量, 区間推定

混合するといった方法がとられる。この場合には“正確な”流出負荷量は得られない。実はこのような離散化を試みることは、コンポジット試料の濃度に偏りを導入することと同じである。流量コンポジットサンプルが正しい負荷量や流量荷重平均濃度の値を与えるのは、連続的に採水・混合が行われるときのみである。

3.2 時間コンポジットサンプルではどうか

(2) 式の最右辺を見れば、時間コンポジットサンプルでも正しい負荷量を与えるのではないかという期待を抱かせる。しかし(2)式をよく見ればわかるが、流量荷重平均濃度 C_w を正しく求めることができなければ流出負荷量の推定は不正確となる。結局、“少なくとも流量や濃度の変動を、その単位時間内では一定と見なしでも差し支えないほどの短い”時間間隔で測定しない限り、正しい負荷量の値は得られない。

4. 解答～我々は何を間違えてきたのか

それでは、離散的にしか河川の流水を採取できないのであれば、いったいどうすれば適切に流量荷重平均濃度をとらえることができるのであろうか。この答えは、半ば逆説的であるが、“コンポジット(混合)しない”ことである。また、なぜ離散化の際に値が偏ることになるのであろうか。これらを理論的に考えるために、モンテカルロ数値積分法の重点的サンプリング法を持ち出す必要がある。今瞬間的な観測負荷量の値を $l_i (=c_i q_i)$ と定義する。この値を、例えば我々のよく知ることような LQ 式などを用いて、近似値で推定することができるものとする。この推定量を \hat{l}_i と書こう。これを用いると、(2) 式は次のように変形できる。

$$L = \int_{t_1}^{t_2} \frac{l_i}{\hat{l}_i} \hat{l}_i dt \quad (3a)$$

$\hat{l}_i dt$ を $(\int \hat{l}_i dt) d\eta$ と書けば(η は 0～1 区間の変数)、(3a) 式は次式に書き直される (t_m は η に対応する時刻)。

$$L = \left(\int_{t_1}^{t_2} \hat{l}_i d\tau \right) \times \left(\int_0^1 \frac{l_m}{\hat{l}_m} d\eta \right) \quad (3b)$$

積分区間が変数の変換に伴い $t_1 \sim t_2$ から 0～1 に変更されているが、この式は(3a)式を変形しただけのものである。これを離散化した式は次式である。

$$L = \left(\sum_{\tau=t_1}^{t_2} \hat{l}_\tau \right) \times \frac{1}{n} \sum_{\eta=1}^n \frac{l_\eta}{\hat{l}_\eta} \quad (3c)$$

ここで n は離散化されたときのサンプル数である。上の式は、単に l_i を時刻に沿って積分する式とは異なり、 η 、すなわち \hat{l}_i の期待値の大きさに応じて離散化(サンプリング)することを意味している。これが重点的

サンプリングといわれるもので、 $|l_i| \propto \hat{l}_i$ であればサンプル数が少ない時でも効率的に良い近似値を与える(もし \hat{l}_i の近似が不十分であっても重点的サンプリングは不偏な推定量を与えるが、サンプル数が多数必要となり非効率となる)。一方、もし(2)式を直接離散化して時間的に等確率な乱数を用いてモンテカルロ法で計算するということになれば、非常に多くの乱数・サンプリング数を必要とすることになるであろう(ここで、 \hat{l}_i の選び方にかかわらず、データ数を増やせば 3c 式は必ず不偏推定量を与えることに注意する)。

以上より、何が問題であったのが明瞭に整理できる。まず、時間コンポジットサンプルは η の大きさに応じてサンプリングしないので“非効率であり”(偏り)、ほぼ全数に匹敵するサンプル数を取らない限り良い推定値を与えない。一方で離散的な流量層別サンプリングは η ではなく“ q_i ”の大きさに応じてリサンプリングしているので、 l_i と q_i の間に何らかの非線形性を認めうるならば、限られたサンプル数に基づく限り、必ず偏りを生じることになる。

つまり、今まで我々が正確な流出負荷量の推定に失敗していた原因は、離散化と同時に偏りを導入していることに気付かなかったことにある。「定期サンプリングの頻度を高めれば真の値に近づく」、とか、「流量に応じてサンプリングすれば、あるいは出水時にたくさんデータを取れば真の値に近づく」といった期待と、既にあった水質の定期モニタリング法を土台にすべきという先入観にとらわれて、「違う場所を掘っていた(digging in the wrong place)」のである。定期間隔の頻度をどの程度まで高めればよい推定値が得られるか、というアプローチでは正答に近づくことはできない。

5. おわりに

ここで示したように、重点的サンプリングによれば、新たなサンプリング計画の策定や不偏な流出負荷量の点推定および区間推定が可能となる。それだけではなく、定期採水データや降雨出水時の詳細調査などの、重点的サンプリング法とは異なるサンプリング法で収集されたデータに対しても、この考え方を援用することで、ほぼ不偏な流出負荷量推定が可能である。

謝辞

本研究は科学研究費補助金(No.20580263, No.23580333)の助成を受けて実施したものである。

参考・引用文献 例えば、津田孝夫(1969), モンテカルロ法とシミュレーション, 培風館