

(P63) 市販流量年表データベースによる一級水系代表観測点日流量データの再構築

首都大学東京大学院	都市基盤環境学域	学生員	岡元	宏薫
首都大学東京大学院	都市基盤環境学域	正会員	河村	明
首都大学東京大学院	都市基盤環境学域	正会員	天口	英雄
首都大学東京大学院	都市基盤環境学域	正会員	中川	直子

1. はじめに

河川の洪水防御，水資源確保，環境保全などの課題に対処するには，まずその基礎的水文資料，特に河川流域の長期にわたる雨量および河川流量データの解析が必要不可欠となる．現在，日本における109の一級水系では，昭和13年（1938年）以降，国土交通省（当時の内務省土木局）により雨量観測および流量観測が実施され，これらのデータは雨量・流量年表として（社）日本河川協会より毎年刊行されている¹⁾．近年，これらの貴重なデータは同協会により平成15年分まで電子化され，DVD媒体の「雨量・流量年表データベース」²⁾として販売されている．これにより，例えば，これまで各研究者が独自に必要な日流量を流量年表より抽出し電子化していたが，この手間が省け作業能率は著しく向上したと言えよう．著者らも本データベースを活用し河川流況解析を行っていたが，その過程で明らかにおかしいと考えられるデータが存在することが分かり，冊子体の流量年表¹⁾比較したところ結構多くの相違点が見つかった．そこで著者らは全国一級水系代表観測点のみを対象とした流量年表データベースの日流量データ（以後，DVD流量年表と記す）を流量年表のそれと比較・精査することにより，DVD日流量データの入力ミスを抽出しその成因について考察し，また流量年表自体の日流量データの誤記についても検証を行っている³⁾．本研究では，全国一級水系代表観測点を対象に，DVD日流量データの入力ミスを補正し，また流量年表自体の日流量データについては明白な誤記についてのみ補正を行い日流量データの再構築を行った．次いで，再構築日流量データを月流量データに累積し，これを流量年表に記載され正しいと考えられる月流量データと比較することにより，本再構築データで補正しきれないデータについて検討を行った．

2. 対象データ

本研究では109一級水系の各幹川における代表観測点として，基本的に河川ハンドブック⁴⁾の中で基準点として挙げられている観測点を採用している．しかし，観測点は河川改修や設置場所の変更と共にその名前が変更されていることも多く，名称の異なる同一観測点を見つけ出すのは容易ではないため，本研究ではある観測点の河口からの距離と対象とする流域面積が共に5%以内の差であった場合同じ観測点とみなした⁴⁾．そして，各水系の代表観測点における日流量データをDVD流量年表より抽出した．

3. DVD日流量データの検証³⁾

DVD日流量データを精査するために流量年表のそれと比較した．まず，流量年表には日流量データのみならずそれらを月，年毎に合計した値も記載されている．ここでは，DVD日流量データを累積した年間合計値が流量年表の年間合計値と一致していればその年のDVD日流量データは全て流量年表と同じであるとみなした．年間合計値が一致しなかった場合，月，日流量データを一つ一つ見比べ，どのデータが異なっているのかを見つけ出した．この作業の結果，発見したDVD日流量データの流量年表との相違点を以下に述べる．

まず，単純な観測点名の漢字入力ミスと思われる場合が数箇所見つけた．例えば，那賀川（徳島県）において1956年のみ「古庄」観測点が「古圧」と誤記されていた．これ以外にも5水系において観測点名が年毎に不統一となっている場合が見られた．これらの入力ミスは，観測点名を入力してDVD流量年表から日流量データを抽出する場合には大きな影響がある．また，流量年表データベースの出力形式には表形式とデータ登録形式の2つがあり，前者ではデータ抽出に非常に時間がかかるため，データ登録形式にてデータを出力しテキスト(csv)形式として抽出し，抽出データを調べたところ，DVD日流量データには各数値間の区切りにカンマが使用され，csv形式(カンマ区切り)で読み込むことができる一方，千の位の後にも位取りのカンマが入力されていた．すなわち，そのままDVD日流量データをデータ登録形式にて出力してデータを抽出すると，数値間のカンマと位取りのカンマの区別がされず，例えば，1964.10 (m³/s) が1 (m³/s) と964.10 (m³/s) の2つのデータとして取り扱われることが判明した．このため，DVD日流量データをデータ登録形式として用いる場合，位取りのカンマを消去するなどなんらかの処理が必要である．

次に，対象とした109水系の代表観測点の中で，1977年以前より流量が記載されている観測点は106水系あるが，そのうち101水系で，DVD日流量データの1977年のデータには全く違う値が入っていることが判明した．その後日本

河川協会に問い合わせ、その年の正しい日流量データを入手した。

さらに、DVD日流量データを流量年表のそれと比較・精査した結果、上記以外に単純な数値入力ミスが見つかった。入力ミス種別としては、)48.20(m³/s)が482.00 (m³/s)のように小数点の位置の付け間違い、) 149.70(m³/s)が19.70(m³/s)のように抜けてしまっている、逆に1640.10(m³/s)が10640.10(m³/s)のように余計な数字が加わっている、139.21(m³/s)が139.12(m³/s)のように数字が入れ替わっている、さらに) 88.22(m³/s)が38.22(m³/s)や42.14(m³/s)が12.14(m³/s)のように、3と8や1と4のように似通った数字の読み違いが挙げられた。これらの数値入力ミス(1977年を除く)は43水系で見つかり、入力ミス種別としては)が2、)が23、)が224(合計249個)で、似通った数字の読み違いによる入力ミスが圧倒的に多く見られた。ただし、1977年を除く全データ数は1,809,683であるので、その入力ミス率としては0.014%と極めて小さい。

4. 流量年表日流量データの検証

流量年表には日流量データの数値の表のみならず、その月合計値や年合計値が記載され、さらに日流量時系列グラフも図示されている。そして、記載日流量データより算定した月合計値が記載月合計値と異なる場合が見つかり、この場合、日流量データを日流量時系列グラフと比較精査した。その結果、流量年表の日流量データにも明らかに誤記と考えられる数値が記載されていることが分かった。ここでは誤記かどうかよく分からない場合はそのままとし、明らかに誤記と考えられる場合についてのみ検証した結果、記載ミス種別としてはDVD日流量データの場合と同様に分類したところ、)が2、)が26、)が95(合計123個)となり、入力ミス率としては0.0067%である。

5. 再構築した月流量データベースの検証

上記の誤データを補正し再構築した日流量データベースの月合計値を算定し、流量年表の月合計値と比較・検証した結果を表-1に示す。残差を持つデータが約4000個ある中、-0.15~0(m³/s)の範囲で誤差を持つデータが約1500個、0~0.15(m³/s)が約2000個、つまり表示されている小数点第二位よりも小さい値を四捨五入したために生じたと想定できる範囲の残差が大半を占めていた。しかし、グラフの記載されていない古い年やグラフのオーダーが大きい河川、またグラフの記載範囲を大きく超えた値であるために明らかな誤記と判断できないが大きな残差を持つデータも存在していることが分かった。この他の残差を持たないデータは圧倒的にあり、残差があっても小さい値が大半を占めるため、再構築データベースが信頼できると言える。

6. むすび

本研究では一級水系代表観測点におけるDVD流量年表の誤データを補正するとともに、流量年表自体の明白な誤記データを補正し日流量データの再構築を行った。そして本再構築データの検証を行い補正しきれていないデータについて検討を行った結果、月合計値において-0.15~0.15(m³/s)の小さな値の残差がほとんどであり、訂正不可能な誤記データは存在するが、本研究により全国一級水系の流況解析を行うにあたり有効なデータベースを再構築できたと言える。

表-1 再構築データベースと流量年表の月合計値による比較

月流量残差	個数
~ -100	17
-100 ~ -50	17
-50 ~ -10	18
-10 ~ -1.5	59
-1.5 ~ -0.15	129
-0.15 ~ 0	1447
0 ~ 0.15	2077
0.15 ~ 1.5	105
1.5 ~ 10	63
10 ~ 50	33
50 ~ 100	18
100 ~	14
合計	3997

参考文献

- 1) 国土交通省河川局編：昭和13年度～平成15年度 雨量年表，流量年表，日本河川協会，1938-2003。
- 2) 日本河川協会：雨量・流量年表データベース，DVD版，2007。
- 3) 河村 明・岡元宏薫・天口英雄・中川直子：市販流量年表データベースによる一級水系代表観測点日流量データの検証，水文・水資源学会2009年度研究発表会要旨集 pp.90-91，2009
- 4) 日本河川協会：2007河川ハンドブック，2007。

キーワード：流量年表，データベース，一級水系，日流量データ，検証