月流量のシミュレーションについて

大阪大学工学部　正員　室田　明
同　　　学生員○江藤剛治

1. まえがき

筆者らはここ数年にわたって"水資源計画のための月流量シミュレーションの研究"にとりくみ，i)月流量の分布はピアソンIII型分布で表わされる[1]，ii)月流量の正規化変量は単純マルコフ連鎖で表わしうる[2]，iii)低水の持続性（自己相関性）は高く，高水は偶発性が強い[2]，iv)その主たる原因は流域特性であり降雨の自己相関性の影響はそれほど大きくない[3]，などの特性を明らかにし，v)単一地点に対して，低水母集団と高水母集団を合理的に分離した流量シミュレーション・テクニックを提唱した[2]。

一方，月流量シミュレーションや作成について，つぎのような疑問が生ずる。i)誤差を含んでおり，しかも20～50年程度の標本数しかない観測資料を用いて，その10倍のオーダーの長さのシミュレーションを行なって意味があるのか，ii)もし意味があるとすれば何年程度のシミュレーションを行なうのが，工学的に最も理想的であるのか。このような重要な問題が、いまだほとんど解明されていないといって過言ではない。筆者らは情報理論を導入してこれらの問題の研究に着手しているので，本論文では，その基本的な考え方を述べる。資料は木津川支川名張川月ヶ瀬測水所1918～1965年の月流量資料を用いた。

2. 基本的考察

月流量シミュレーションの意義について情報理論的な見方をすれば，"水資源システム・シミュレーションにおいて，シミュレートされた月流量をイン・プットとして用いる場合と，観測資料を用いる場合とどちらが多くの有益な情報を得ることができるか"という問題と等値となる。ここでは一次的な考察として，i)観測流量資料（誤差は比較的小さいが標本数に限界がある。一般に10～50年程度），ii)シミュレートされた月流量資料（介入する誤差は観測流量資料より大きいが標本数はいくらでも大きくできる。），のどちらが大きな情報量を持つかを調べる。なぜなら，水資源システム・シミュレーションによって得られる情報量は，ダムの操作方式，評価関数などが同じであれば，イン・プット・システムの情報量の大小によって直接支配されると考えられるからである。

3. 月流量資料の持つ情報量

情報理論によれば，正規分布のエントロピーHは次式で表わされる。

$$H=\log(\sqrt{2\pi e}\,\sigma) \qquad (1)$$

正規化された月流量時系列$Q_1, Q_2 \cdots Q_{i-1}, Q_i, Q_{i+1}, \cdots$の第i番目の流量$Q_i$について，i)標準偏差を$\sigma_i$，ii)$Q_{i-1}$と$Q_i$の相関係数を$\rho_i$とすると，$Q_{i-1}$を与えたときの$Q_i$の条件付分布の標準偏差は，

$$\sigma_i'=(1-\rho_i^2)^{\frac{1}{2}}\sigma_i \qquad (2)$$

となるから，Q_iが正規分布に近い分布を持つとき，Q_iのエントロピーは(1)，(2)より

$$H_i=\log\{c_i\sqrt{2\pi e(1-\rho_i^2)}\,\sigma_i\} \qquad (3)$$

ここに

C_i; 正規分布とみなしたための補正係数で1に近い。

各月の平均的なエントロピーを求めるには，iを暦上の月の番号（1月，2月，……）とし，一年について平均する。

$$\overline{H}=\frac{1}{12}\sum_{i=1}^{12}\log\{C_i\sqrt{2\pi e(1-\rho_i^2)}\,\sigma_i\}$$
$$=\log\{\prod_{i=1}^{12}C_i\sqrt{2\pi e(1-\rho_i^2)}\,\sigma_i\}^{\frac{1}{12}} \qquad (4)$$

Q_iの測定上の誤差，計算上の誤差，流量シミュレーション・モデルの仮定上の誤差などを雑音とみなし，雑音の合成されたものの標準偏差をσ_{ni}とするとき，そのエントロピー$\overline{H}_n$は同様にして，

$$\overline{H}_n=\log\{\prod_{i=1}^{12}\sqrt{2\pi e}\,\sigma_{ni}\}^{\frac{1}{12}} \qquad (5)$$

月流量資料の持つ情報量は，$\overline{H}_n$が$\overline{H}$に対してそれほど大きくないとき近似的に$\overline{H}$と$\overline{H}_n$の差で表わされる。

$$\overline{H}-\overline{H}_n=\log\{\prod_{i=1}^{12}C_i\sqrt{1-\rho_i^2}\,\sigma_i/\prod_{i=1}^{12}\sigma_{ni}\}^{\frac{1}{12}} \qquad (6)$$

4．水資源システム・シミュレーションのイン・プットとしての意味の検討

求める情報量の大小は，(6)式を用いるかわりに底が1より大であることを考慮して，真数の大小で比較することにする。情報量の定義より真数は流量母集団の標本空間上の全要素の数に比例していると考えられる。これをLとすると

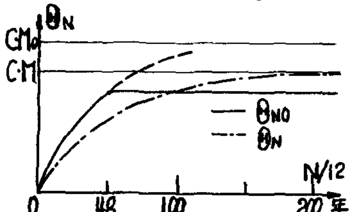

Fig.1

$$L=C'\{\prod_{i=1}^{12}C_i\sqrt{1-\rho_i^2}\,\sigma_i/\prod_{i=1}^{12}\sigma_{ni}\}^{\frac{1}{12}}$$
$$=C\cdot M \qquad (7)$$

ここに

$$C=C'\cdot\prod_{i=1}^{12}C_i \qquad (8)$$
$$M=\{\prod_{i=1}^{12}\sqrt{1-\rho_i^2}\,\sigma_i/\prod_{i=1}^{12}\sigma_{ni}\}^{\frac{1}{12}} \qquad (9)$$

Nヶ月分の流量資料とは，このL個の要素を持つ標本空間から独立に生起したN個の標本である。生起確率が等しいとすれば，L個の要素を持つ標本空間において独立な試行に対して実現するであろう要素の数Θ_Nの期待値は，L，Nがともに大であるとき

$$\Theta_N=L(1-e^{-\frac{N}{L}})$$
$$=C\cdot M(1-e^{-\frac{N}{CM}}) \qquad (10)$$

となる。一般にMは既知である。Θ_Nの大小により得られる情報量の大小が比較できる。シミュレートされた流量資料に関する量は添字なし，観測流量資料の方は添字"0"をつけると，$\Theta_N>\Theta_{N0}$のとき流量シミュレーションは水資源システム・シミュレーションのイン・プットとして意味を持つ。誤差の大きさσ_{ni}，定数Cの推定などに関する研究はまだ行なっていないので，これを適当に仮定して計算したΘ_N，Θ_{N0}の値を，サンプル数Nの関数として模式的に示したものがFig.1である。

［参考文献］

1）室田明，神田徹，「木津川における低水流量の時系列特性について」S.43関西支部年次学術講演会講演概要
2）室田，神田，江藤，「木津川水系の流量シミュレーションに関する研究」第23回土木学会年次学術講演会講演概要
3）室田，神田，江藤，「流量時系列の相関係数推定における降雨特性の導入について」第24回土木学会年次学術講演会講演集