

ニューラルネットワークによる日流量予測と入力情報の基準化

名古屋工業大学 フェロー会員 長尾 正志
 ○名古屋工業大学大学院 学生員 曾根 英雄
 名古屋工業大学 学生員 山田 祐次

1. はじめに

有効な貯水池運用に、流入量の正確な予測は最重要課題である。本研究室では、過去の流量時系列から簡単に非線形関係が得られるニューラルネットワークを用いて流入量予測を研究してきた。予測問題を考えるとき、データの平均的な分布特性からはずれた特異値が予測精度に影響を及ぼしたり、学習範囲内のデータに対して妥当な答えを出力するが、範囲外のデータに対しては必ずしも十分な結果が得られない場合もある。

そこで、本研究ではニューラルネットワークの入力データの基準化に際して、対数正規分布への適合、あるいは、その非超過確率の採用による予測結果の比較・検討を試みた。

2. 対象資料および計算方法

目的変数は木曽川水系牧尾ダムの翌日の日流量とし、説明変数は、牧尾ダムの2日前までの日流量および同ダム上流に位置する王滝、三浦ダムの2日前までの日雨量とする。計算には、図1に示す、入力層、中間層、出力層がそれぞれ単層からなる階層型ニューラルネットワークを使う。予測期間は、1983～88年のそれぞれ12月から翌年2月の渇水期5ヶ年とし、ニューラルネットワークの学習には予測期間の前2年分の渇水期データを用いる。

牧尾ダム渇水期の日流量は平均 $4\text{m}^3/\text{s}$ 程度に小さい値であるが、ごくまれに $230\text{m}^3/\text{s}$ もの大きな流入がある。ところで、ニューラルネットワークでは計算を行うために、入力データを0～1の間に基準化する必要がある。慣用的な基準化では、式(1)が採用されることが多い。

$$Y = \{0.9 \times (X - X_{\min}) + 0.1 \times (X_{\max} - X)\} / (X_{\max} - X_{\min}) \quad (1)$$

ただし、 X 、 $X_{\min}$ 、 $X_{\max}$ はそれぞれ観測値、観測値の最小値、最大値である。式(1)は観測値以外の日流量や日雨量にも対応できるように、0.1～0.9の間に基準化を行っている。この変換では、学習データの中に $230\text{m}^3/\text{s}$ のようなかけ離れたデータが存在すると、学習データで十分に学習させたとしても、予測精度が悪くなる。これは、渇水期の小さい流量と大きく異なる流量を勘案することにより、渇水期の小さい流量特性が再現しにくくなり、予測精度に悪影響を与えるためと思われる。そこで、高瀬による加定数0の対数正規分布への変換式(2)を用いて、原変数の特異データの変動幅を抑制するように変換する。それを、式(1)で基準化を行い、入力データとする。

$$Y = \alpha \{ \log_{10} (X / X_0) \} \quad (2)$$

ただし、 α 、 X_0 は観測値から決まる定数で、過去20年間の牧尾ダム渇水期日流入量から決定した。ところで、データ数が無限大に近ければ、0～1で基準化を行ってもよいが、現実にはデータ数が制限される。したがって、基準化の幅は現実のデータ数とのかねあいで決定されなければならない。そこで、学習範囲外のデータに対しても対応できるように、過去の渇水期日流量の非超過確率をとる。具体的には、式(2)で、対数変換された値を標準正規分布の確率変数として、式(3)により、非超過確率(累積分布)に変換する。

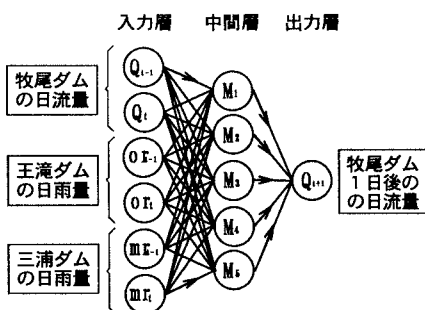


図1 計算に使用した
ニューラルネットワーク

$$Z = \int_{-\infty}^Y (1/\sqrt{2}) \exp(-t^2/2) dt \quad (3)$$

ただし、 $t = (Y - \hat{Y}_m) / \hat{\sigma}$, $\hat{Y}_m = (\sum_{i=1}^n Y_i) / n$, $\hat{\sigma}^2 = \{ \sum_{i=1}^n (Y_i - \hat{Y}_m)^2 \} / n$ である。

以下では、①従来の方法（慣用法）、②対数正規変換、③非超過確率率 という3つの基準化による予測を行い、予測値をそれぞれ原変量の大きさに戻して、予測精度の比較、検討をする。

3. 計算結果

①慣用法、②対数変換法、③非超過確率法 のそれぞれの観測値と予測値の対応として、相関係数を表1、相対誤差を表2に示す。また、表1には比較、検討のために各年の対象期間における最大流量も示す。

相関係数と相対誤差の平均値から判断すると、慣用法に比べ、対数変換法、非超過確率法の予測精度は良い結果が得られた。対数変換法と非超過確率法を比較すると、平均的には予測精度にそれほど差がない。

予測年ごとに比較すると、83-84年のように最大流量が小さい場合は、非超過確率法の予測精度は良い。しかし、最大流量が示すように、予測年に渇水特性から離れたデータが存在するほど、非超過確率法の予測精度は悪い。これは、非超過確率法を採用することにより、入力値に異常に大きな値が存在しても対応できるネットワークが構築される反面、予測値を原変量に戻す段階で、大きな予測誤差を与えた結果と考えられる。予測の一例として、83-84年の予測を図2に示す。これから、対数変換法と非超過確率法が、ほぼ観測値近傍の予測をしていることがわかる。一方、慣用法は、観測値から、かなりはずれた平均的な予測をしている。慣用法で流量の小さい変化がうまく表現できなかった結果と思われる。

4. まとめ

1) ニューラルネットワークの基準化に対数正規変換法および非超過確率法を用いることで、慣用法に比べて、予測精度が向上した。

2) 非超過確率法では、基準化の幅を検討しないで、かけ離れた値に対応するネットワークを構築しうる。しかし、予測の対象が前もって低水流量と限定された基準化では好結果であるが、そうでない場合は、むしろ慣用法を用いるべきである。

表1 観測値と予測値の相関係数

	最大流量	慣用法	対数変換	非超過確率
83-84年	9.57	0.580	0.848	0.931
84-85年	35.49	0.443	0.360	0.218
85-86年	30.54	0.670	0.656	0.523
86-87年	69.97	0.434	0.449	0.219
87-88年	8.37	0.293	0.881	0.810
平均値		0.484	0.639	0.540

表2 観測値と予測値の相対誤差

	慣用法	対数変換	非超過確率
83-84年	1.262	0.550	0.419
84-85年	1.999	2.162	2.074
85-86年	1.302	0.944	1.153
86-87年	3.539	3.174	3.568
87-88年	0.812	0.346	0.389
平均値	1.783	1.435	1.520

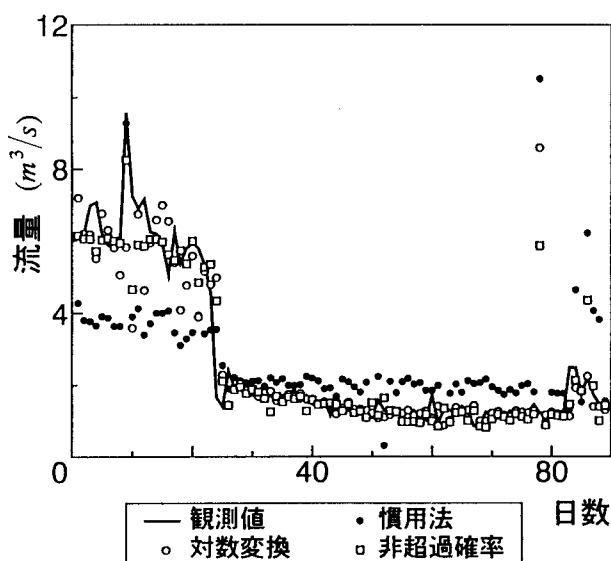


図2 83-84年の観測値と予測値の比較