

ニューラル・ネットワークによる渇水期日流量の予測(1)

- 階層型ネットワークの構造選定 -

名古屋工業大学 正員 長尾正志、名古屋工業大学大学院 学生員 鈴木祐一、
名古屋工業大学 佐野正嗣

1. 研究の目的と概要

(1) 目的 利水用貯水池への流入量予測は合理的な利水運用の基礎である。特に、昨年のような水不足の事態に対して、この重要性は十分に認識された課題である。その際に、雨量・流量などの入力を用いた物理的・確率的なアプローチもあるが、流域を総合した物理定数の決定などに曖昧さが介在し、結果は必ずしも満足すべきものではない。ここでは、階層型ニューラル・ネットワーク手法を用いる。特に、ネットワーク構造の選定に、最小自乗誤差や、AIC, BIC, MDLなどの情報量基準の適用を試みた。

(2) 対象資料 牧尾ダムの渇水期における翌日の日流量を目的変数、観測された日流量、日雨量を説明変数とした予測を考える。これまでの日流量時系列についての研究で、大抵2日前、精々3日前までの関連が判明しているので、説明流量としては、2日前、および3日前までを用いた。ここでは、3日前までの結果を述べる。計算には、1980-81, 1981-82, 1982-83年の12月14日から翌年2月11日までの60日間を3冬期(合計180日分)使う。また、雨量にはそれに先立つ0-6日分の影響を考える。なお、学習データとしては、先の2冬分を、検定用の未学習データとしては、後の1冬分を使う。

2. ニューラルネットワークの構成要素

入力層、中間層および出力層はそれぞれ単層とし、出力層のユニット数は1とする。

(1) モデル選定のための評価基準 情報量統計量の中心は、データへの近似誤差とモデルの自由度の調整という形態をとる。ここでは、ネットワークを構成する結合係数や閾値係数のパラメータ数を自由度として用いる。各情報量統計量として用いた評価基準は、式(1)-(3)の通りである。

$$AIC = N \cdot \log \sigma^2 + 2p \quad (1) \quad BIC = N \cdot \log \sigma^2 + p \cdot \log N \quad (2) \quad MDL = (1/2)N \cdot \log \sigma^2 + p \cdot \log N \quad (3)$$

ここで σ^2 : 最小自乗平均誤差、 p : パラメータ総数、 N : データ数

ただし、パラメータ数 p は、入力層、中間層、出力層のユニット数を、それぞれ IN , MI , OU とすると、結合係数の数 $MI(IN+OU)$ と閾値係数の数 $(MI+OU)$ の和で与えられる。

(2) 入力層のユニット数 流入量ユニット数(3)に雨量ユニット数(0, 1, ..., 6)を加えたものとする。

(3) 中間層の層数 仮定した中間層のユニット数は2, 3, ..., 7, 8である。

3. ニューラル・ネットワークの構造選定

計算手順としては、まず適当な中間層ユニット数 MI を仮定し、勘案すべき先行雨量ユニット数を決めたい。ニューラルネットの計算に入り、各学習回数に応じた自乗誤差を出す。これから情報量基準を計算する。学習回数としては、誤差の減少傾向より、1,000回から4,000回までの500回ごとの値を利用することとし、各情報量基準値はその平均を用いる。これらの結果を総合して、最良となる雨量ユニット数、および中間層ユニット数を試行的に求めていく。最終的に求められた結果を、入力層ユニット数の検討について表-1に、 $IN=7$ とした中間層ユニット数の検討については、表-2に示す。なお、表中の丸数字は各基準での順位を示す。

これから以下のような結論が得られる。

3. 1 入力層のユニット数 AIC, BICでは7ユニットがベスト、MDLでは3となる。以後7を使う。

3. 2 中間層のユニット数 AICより6が、BIC, MDLより3が選定され、以後の考察対象とする。

3. 3 ネットワークのパラメータ AIC, BIC, MDLの最小値となる100回キザミの学習回数より求めることにする。AIC, MDLにより選定されたものを、Aモデル、Mモデル、また重線形回帰モデルをMRモデルと呼ぶ。

表-1 入力層ユニット数の検討

IN	AIC	BIC	MDL
3	-761.0 ⑥	-662.0 ③	-250.5 ①
4	-780.5 ③	-662.3 ②	-235.1 ②
5	-768.8 ⑤	-631.5 ⑤	-204.1 ③
6	-779.4 ④	-623.0 ⑥	-184.3 ⑤
7	-865.5 ①	-689.9 ①	-202.2 ④
8	-524.1 ⑧	-329.4 ⑧	-6.3 ⑧
9	-855.5 ②	-641.6 ④	-146.8 ⑥
10	-740.5 ⑦	-507.4 ⑦	-64.2 ⑦

表-2 中間層ユニット数の検討

MI	AIC	BIC	MDL
2	-460.4 ⑦	-398.8 ⑦	-150.5 ⑤
3	-861.2 ②	-771.8 ①	-313.2 ①
4	-803.2 ④	-685.1 ③	-246.5 ②
5	-772.2 ⑤	-625.3 ④	-193.4 ④
6	-865.5 ①	-689.9 ②	-202.8 ③
7	-815.5 ③	-611.2 ⑤	-139.4 ⑥
9	-765.5 ⑥	-532.4 ⑥	-76.7 ⑦

表-3 学習、未学習データに対する学習回数と相関係数の関係

A モデル	適用 データ	学習回数		
		1,300	1,800	3,800
IN=7 MI=6 OU=1	学習	0.849	0.864	0.872
	未学習	0.685	0.638	0.578
M モデル	適用 データ	学習回数		
		1,500	1,900	3,000
IN=7 MI=3 OU=1	学習	0.793	0.795	0.803
	未学習	0.771	0.743	0.708

4. 結果の評価・検討

4.1 学習データに対する検討 予測値と実測値の間の相関係数で比較・評価することにしよう。

(1) 重線形回帰モデルとの比較 比較対象として、重線形回帰モデルを用いる。その結果として、学習データに対しては、相関係数の比較より、表-3右欄のように、Aモデルでは0.872(学習回数 $l=3,800$)、Mモデルでは0.803($l=3,000$)、MRモデルでは0.719となり、本モデルの優位性が認められる。

(2) AIC, MDLの比較 このように十分に学習させた場合には、AICの方がよさそうである。

4.2 未学習データに対する検討

ところで、一般に学習回数を増すと、学習データでは適合度があがるが、未学習データでは、適合がかえって悪くなるといわれている。これらの詳細な論議は別の発表によるが、ここではそれより学習回数を減らした場合に、確かに未学習データでも適合がよくなることを検証する。

(1) 重線形回帰モデルとの比較

十分な学習の場合には、MモデルではMRモデル(0.747)とほぼ同程度であるが、Aモデルではあまりよくない。少ない学習回数の相関係数の比較例を表-3に示す。これは、1,000-4,000の間で100回ごとに計算したものである。少ない回数のMモデルは、MRモデルと同等またはよい結果である。Aモデルでは、多い学習回数に対して、MRモデルよりよくなかったものが、学習回数を少なくすると次第によくなり、MRモデルに近くなる。

(2) AIC, MDLの比較 学習データに対しては、Aモデルの方がMモデルより格段とよかった。しかし、未学習データに対しては、Mモデルの方がAモデルより相関係数で10%程度大きく、MRモデルより適合性がよい。

5. 成果と問題点

5.1 成果 1) 情報量基準(AIC, MDLなど)を用いたネットワーク構造の選定が可能であることを示した。 2) 情報量基準間の比較 採用パラメータ数についていうと、AICでは多め、MDLでは少なめ、BICでは両者の中間的な特性となる。 3) 重回帰モデルとの比較では、学習データに対して、十分学習させれば重回帰モデルより良い結果が得られる。しかし、未学習データでは、MDL基準によるモデルは重回帰モデルよりよい結果が得られるが、AIC基準のモデルは重回帰モデルより良くない。 4) 総合的な評価としては、学習データに対してはAIC基準のモデルが、未学習データに対してはMDL基準モデルが適切であろう。

5.2 残された問題点 1) できるだけ広範囲に及ぶ、すなわち内挿になるような、データの学習 2) 過剰学習の回避のための学習回数の選択 3) 物理的・確率的モデルによる検討からの結果の裏付け