

ニューラルネットワークによる渇水期流量の推定精度に関する実証的研究

岐阜工業高等専門学校 正員 鈴木正人
 名古屋工業大学 正員 長尾正志
 ○岐阜工業高等専門学校 学生員 岡本祐司

1. はじめに

入力と出力が与えられるとその対応関係を満たすように学習し、自らシステムを構築するニューラルネットワークは、入出力の関係が物理的に解明されていない場合でも構成可能なモデルとして数多くの分野で成果をあげている。水文学の分野でも短期流出量の予測への適用を始めいくつかの報告がある。本研究はニューラルネットワークの渇水期流量の推定への適用可能性の考察を目的として、現実のデータに対しどの程度の推定精度が得られるかを検討したものである。

2. 本研究で用いるニューラルネットワークおよび計算方法

研究の目的から考えて基本的な構造である、入力層・中間層・出力層の3層からなる階層型のネットワークを用い、中間層のユニット数は入力層のユニット数と等しくした。また学習アルゴリズムは教師付き学習では最も一般的なバックプロパゲーション法¹⁾を用いた。渇水期流量の推定に適したネットワーク構造や学習アルゴリズムは、それだけでも議論の対象となる問題と思われるが、渇水期流量の推定にニューラルネットワークを適用するのが初めてであり、上述のような基本的なものを用いた。

つぎに、計算方法に関して、推定精度の実証といった観点からは推定値と実データとの比較を行う必要がある。そこで、まず全てのデータのうち一部を除いたものを教師信号として与え、教師信号の入力と出力の関係を満たすように学習させる。その後、除いたデータを学習済のネットワークに入力として与えて出力させた値、すなわち推定値、と実データとの比較を行った。ここで問題になるのが「どの程度学習させれば良いか」である。ニューラルネットワークは過剰に学習させ過ぎる（過学習する）と、教師信号と類似の性質の入力値については非常に良い精度を持つ出力値が得られるが、教師信号と異質の入力値に対する出力値の精度は極端に悪くなる特性があることが知られている²⁾。そこで、本研究では学習の程度を教師信号と出力値との相対誤差により表現し、それを何段階かに変化させることにより、学習の程度と推定精度の関係を考察した。

3. 適用計算

対象データとして牧尾ダムにおける1969～1991年の12月の日流（入）量と日降水量を用いた。日流量と日降水量のそれぞれについて推定日より前1, 3, 5日分のデータを組み合わせてネットワークへ入力し、推定日流量を出力させた。データの組数は各年26組で総計598組である。データのうち各1年分を除いたものを教師信号として与えて学習させた後、除いた1年分について推定値と実データとの比較を行った。

流入量5日間、降水量5日間の計10個のデータを入力として与えた場合の、推定値と実データの相対誤差（ $= | \text{実データ} - \text{推定値} | \div \text{実データ}$ ）の出現頻度を図-1に示す。図中でerrorとは、学習段階における教師信号と出力値との相対誤差の和を意味し、それぞれ14, 12, 10, 8はデータ1個当たりの相対誤差でほぼ40, 35, 30, 25%に当たる。この値が小さいほど学習の程度が良いと表現することにしよう。ちなみに、学習回数の平均はそれぞれ、334, 2501, 3975, 10020回で、学習の程度をよくするには学習回数を飛躍的に増加させねばならない。図によると、errorが小さくなるほど、すなわち学習の程度が良いほど出現頻度のピークが小さい方に寄っており、推定精度も良くなることがわかる。

つぎに、各入力データの組み合わせにおける推定値と実データの相対誤差の平均を表-1に示す。表

で学習段階の相対誤差の和が8の列で値が記入されていない行では、学習を30万回行っても条件を満たす結果が得られなかったので計算を打ち切った。全体的にみてどの入力データの組み合わせにおいても

学習の程度が良くなるほど推定精度も良くなる。しかし、学習段階における相対誤差の和が8と10との比較においては、先の図-1に見られたほどの差はない。

これは、学習段階の相対誤差の和が8の場合には、推定値と実データの相対誤差が異常に大きなものがいくつか現れるためと思われる。入力として流量・降水量とも5日にした場合、学習段階の相対誤差の和が10の場合は推定値の相対誤差の最大値が323%であるのに比べ、8の場合の最大値は2434%もあり、これが全体の平均を引っ張っているものと思われる。

この結果だけからでは即断することはできないが、過学習に当たる現象が出てきているのかもしれない。また、入力データの組み合わせによる影響はそれほど顕著に現れてはいないが、流量5日に対しては降水量3日の組み合わせが良いようだ。

これまで推定精度を相対誤差により判断してきたが、つぎに実データにより推定値への一次回帰を行う。その場合の、相関係数および回帰式における傾きにより比較を行う。推定値と実データとの相関係数を表-2に、回帰直線の傾きを表-3に示す。表-2によると学習の程度や入力データの組み合わせによる差はあまり認められないが、表-3の回帰直線の傾きは、学習の程度が良くなるほど1に近くなる。本来データと推定値が一致すれば、回帰直線は傾きが1、切片が0になるべきものなので、その観点から見れば学習の程度が良くなるほど推定精度が高くなることを意味していると思われる。

4. まとめ

採用した構造を持つニューラルネットワークでも湧水期流量の推定がある程度可能なことが示された。今後さらに、ネットワーク構造や学習アルゴリズムを検討したり、入力に用いるデータをその特性によって分類選択するなどの工夫によって、より精度の良い推定が可能になるものと期待している。

5. 参考文献

- 1) ニューロンネットグループ・桐谷滋：入門と実習ニューロコンピュータ，技術評論社，pp.28-84，1989.
- 2) 例えば，計測自動制御学会編：ニューロ・ファジィ・AIハンドブック，オーム社，pp.1160-1167，1994.

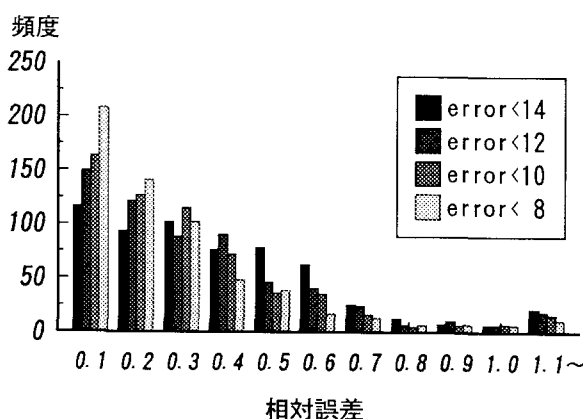


図-1 推定値相対誤差頻度（入力：流量5日 降水量5日）

表-1 推定値の相対誤差の平均（%）

入力データ		学習段階の相対誤差の和			
流量	降水量	14	12	10	8
5日	5日	35.75	31.41	28.25	27.78
5日	3日	35.27	29.34	25.59	21.25
5日	1日	35.51	34.34	25.06	28.27
3日	5日	35.28	31.11	28.71	28.61
3日	3日	36.16	30.09	26.16	22.07
3日	1日	37.97	31.14	25.13	—
1日	5日	39.05	32.36	28.47	29.26
1日	3日	40.25	32.01	24.86	—
1日	1日	45.45	37.50	27.87	—

表-2 推定値と実データ間の相関係数

入力データ		学習段階の相対誤差の和			
流量	降水量	14	12	10	8
5日	5日	0.536	0.585	0.547	0.591
5日	3日	0.524	0.613	0.538	0.636
5日	1日	0.575	0.645	0.645	0.288
3日	5日	0.585	0.597	0.582	0.613
3日	3日	0.585	0.605	0.627	0.658
3日	1日	0.594	0.635	0.672	—
1日	5日	0.606	0.602	0.623	0.597
1日	3日	0.595	0.599	0.605	—
1日	1日	0.587	0.620	0.681	—

表-3 推定値と実データ間の回帰直線の傾き

入力データ		学習段階の相対誤差の和			
流量	降水量	14	12	10	8
5日	5日	0.248	0.436	0.432	0.632
5日	3日	0.244	0.395	0.434	0.579
5日	1日	0.292	0.362	0.351	0.227
3日	5日	0.286	0.423	0.462	0.662
3日	3日	0.304	0.387	0.532	0.616
3日	1日	0.329	0.374	0.383	—
1日	5日	0.321	0.429	0.502	0.590
1日	3日	0.344	0.385	0.482	—
1日	1日	0.367	0.376	0.398	—