

ニューラルネットワークによる渇水期日流入量の推定における 効果的な教師信号の与え方の研究

○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

1. はじめに

入力と出力の因果関係が物理的・確率的に解明されていない場合でも、両者が与えられるとその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、数多くの分野で成果を上げている。著者らは、ニューラルネットワークを渇水期日流入量の推定に適用し、推定値を実際のデータと比較することでその可能性を示してきた^{1,2)}。それらの研究において、ネットワークの学習段階における教師信号は、存在するデータ全てを一括して与える方法を探ってきた。本研究では学習効率や推定精度の向上を意図して、教師信号をその性質によって分類して学習させ、その結果を従前の研究の結果と比較することで、効果的な教師信号の与え方について検討したものである。

2. ニューラルネットワークによる渇水期日流入量系列の推定

2.1 対象データ 牧尾ダムの1969～1991年の23年分の12月の日降水量を用いた。12月1～5日の日流入量・日降水量を入力し、12月6～10日の日流入量を推定、12月6～10日の日流入量・日降水量を入力し、12月11～15日の日流入量を推定・・・、というように、1年当たり5組のデータを対象として5日先までの日流入量を一括して推定した。なお、簡単に、例えば12月6～10日の日流入量を推定した場合、12月6日を1日先、12月7日を2日先、・・・、12月10日を5日先と表記している。

2.2 用いたネットワーク構造 基本的な構成である入力・中間・出力の3層からなる階層型で、流入量系列として5日先までの日流入量を一括して推定した。また、系列内の自己相関性を捉えることができるといわれているリカレント型のネットワークを用いた。学習アルゴリズムに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を用いた。入力層のユニット数は、推定期間より前5日分の日降水量・日流入量、およびフィードバック分の5個を入力するために計15個、出力層のユニット数は5日分の日流入量に対応して5個、中間層のユニット数は、入力・出力層の中間的な値として10個とした。なお、具体的に学習段階においてフィードバックさせる方法は、

Q_j ($j=t+1, t+2, \dots, t+5$): 推定期間における教師信号 (すなわち実データ)

$Q_j(L)$ ($j=t+1, t+2, \dots, t+5$): 学習段階における出力値

として、学習段階における出力層の各ユニットからの出力値と教師信号との差の絶対値である、

$$|Q_j - Q_j(L)| \quad (j=t+1, t+2, \dots, t+5) \quad (1)$$

を入力層に再入力させて入力情報として与えた。学習が終了した段階では、 $Q_j \div Q_j(L)$ となるので、(1)式の値はほぼ0となる。したがって、学習終了後のネットワークを用いた推定では、学習段階において(1)式の値を与えた入力層の入力情報としては0を与えればよく、推定の段階ではフィードバックは見かけ上消滅している。

2.3 教師信号の与え方 推定期間長が5日先までと長いために、推定期間における日流入量は、推定期間内の降雨の有無によって大きな影響を受ける。自明であるが、推定期間内に降雨があれば、日流入量は増大し、降雨がなければ日流入量はあまり変動することなく低減を続けると思われる。そこで、推定期間内の降水量によって、ネットワークを使い分けて推定を行う。つまり、全てのデータを推定期間内の降水量によって2つのグループに分けて、それぞれ別のネットワークを与えて学習させる。実際に推定を行う場合は、当然推定期間内の降水量は未知なので、降雨に関する予報などの情報により、推定に用いるネットワークを使い分けることになる。

具体的には、全115組の教師信号を、推定期間内の降水量によって順位付けし、降水量の少ない55組、降水量の多い60組、の2グループに分けた。また、115組を分割せずに一括して取り扱う場合も併せて検討している。それぞれ、教師信号の中から5組を未知と仮定して除外した残りを、ネットワークに与えて学習させ、学習済みのネットワークを用いて除外したデータを推定し、推定値と実データとの比較で推定精度を検討した。なお、学習の程度は、教師信号 (すなわち実データ) と出力値の相対誤差 ($\equiv | \text{教師信号} - \text{出力値} | \div \text{教師信号}$) の平均値により判断し、学習の程度と推定精度の関係を検討した。

3. 適用計算の結果

推定期間内の降水量が少ないグループの教師信号を用いて、学習段階の教師信号と出力値との相対誤差が30%になるまで学習させたネットワークを用いた推定値の相対誤差 ($\equiv | \text{実データ} - \text{推定値} | \div \text{実データ}$) の頻度を図-1に、また、学習の程度は等しくして、降水量が多いグループの教師信号を用いた

場合を図-2に示す。図-1をみると、1, 3, 5日先の全ての推定値について、相対誤差の頻度のピークが0.1以下のところにあり、比較的良好な精度が得られることが分かる。5日先の推定値については相対誤差が0.4~0.5になる場合もかなりあり、推定期間内の降水量が少ないとはいえ、予測期間が長くなるほど精度が悪くなる様子がうかがえる。一方、降水量が多いグループの教師信号を用いた図-2は、図-1に比べて相対誤差はかなりばらついており、1(100%)以上になる場合もいくつかある。やはり、推定期間内に降水があると流入量の変動が激しくなってしまう、今回用いたネットワークでは有効な推定ができないと思われる。

つぎに、実データと推定値の相対誤差の平均を表-1~3に示す。表-1は教師信号に降水量が少ないグループを用いた場合、表-2は降水量が多いグループを用いた場合で、表-3は比較のために、教師信号をグループ分けせずに一括して学習させた場合である。表-1によると、1~5日先の全てにおいて、学習段階の相対誤差が小さい、すなわち学習回数が多くなるほど推定値の相対誤差の平均が小さく推定精度が良くなっている。表-2との比較では、大部分の場合で表-1の値の方が小さくなっており、推定期間内の降水量が少ない方が、推定精度が高くなることが分かる。表-2で特徴的なのは、1日先の推定値を除いて、学習回数を多くしてもそれが推定精度の向上に結びついていないことである。教師信号のグループ分けをしなかった表-3と表-1との比較では、1, 2日先においては、表-3の値の方が小さく、推定精度が高いと思われるが、4, 5日先においては表-1の値の方が小さくなっており、推定期間が長い場合に教師信号を分割する効果が現れていると思われる。

4. まとめ

本研究では、教師信号を推定期間内の降水量に着目してグループ分けして学習させた場合の推定精度に与える影響を考察したが、長い推定期間においてはグループ分けの効果がみられたが、他の場合はそれほど効果的ではなかった。他の指標によるグループ分けも検討してみる必要があろう。

5. 参考文献

- 1) 鈴木正人・長尾正志・岡本祐司：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，土木学会中部支部平成6年度研究発表会講演概要集，pp.267-268,1995.
- 2) 鈴木正人・長尾正志：ニューラルネットワークによる渇水期日流入量の推定可能性，土木学会第50回年次学術講演会概要集第2部(A), pp.8-9,1995.

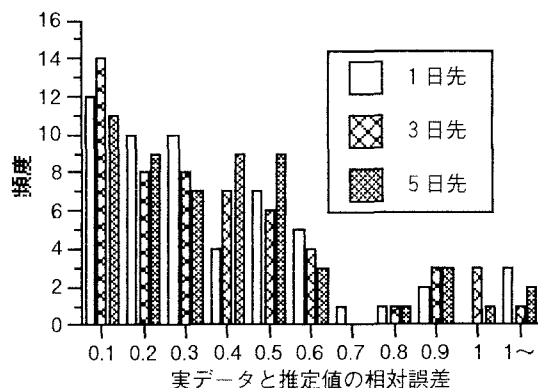


図-1 実データと推定値の相対誤差頻度 (降水小, 学習誤差30%)

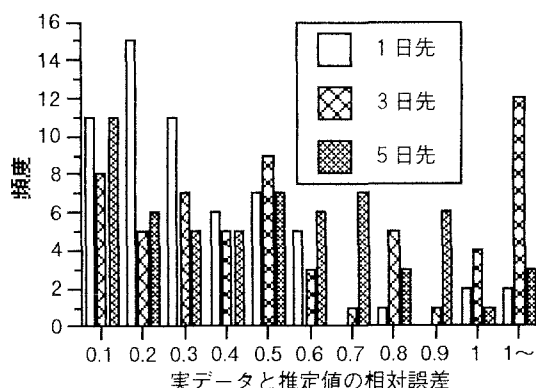


図-2 実データと推定値の相対誤差頻度 (降水多, 学習誤差30%)

表-1 実データと推定値の相対誤差の平均 (降雨小グループ)

学習段階の相対誤差	1日先	2日先	3日先	4日先	5日先
12%	0.588	0.579	0.534	0.475	0.482
36%	0.465	0.475	0.429	0.384	0.398
30%	0.342	0.378	0.351	0.322	0.344
24%	0.311	0.341	0.338	0.319	0.301

表-2 実データと推定値の相対誤差の平均 (降雨大グループ)

学習段階の相対誤差	1日先	2日先	3日先	4日先	5日先
42%	0.432	0.588	0.764	0.889	0.814
36%	0.371	0.470	0.519	0.853	0.714
30%	0.320	0.650	0.864	0.752	0.717
24%	0.405	1.358	1.713	0.953	1.190

表-3 実データと推定値の相対誤差の平均 (グループ分け無し)

学習段階の相対誤差	1日先	2日先	3日先	4日先	5日先
12%	0.375	0.347	0.371	0.655	0.459
36%	0.346	0.408	0.465	0.728	0.440
30%	0.301	0.326	0.359	0.586	0.372
24%	0.242	0.312	0.341	0.493	0.309