

II-120

日流入量予測のためのニューラルネットワークへの入力情報の選定の研究

名古屋工業大学 正会員 長尾 正志
 ○名古屋工業大学大学院 学生員 田沢 孝和

1. 本研究の目的と概要

(1) 目的 利水用貯水池への流入量予測の重要性は、安定した利水供給に対する期待とともに年々高まっている。そこで、本研究はニューラルネットワークによる渇水期日流入量予測を目的として、貯水池流域における入力情報を適切に選定した実用的なネットワークの構築を行い、その有用性を実証的に示したものである。

(2) 対象資料 木曽川水系牧尾ダムの日流入量および日降水量、牧尾ダムの上流に位置する王滝ダム、三浦ダムの日降水量を説明変数とする。目的変数は、牧尾ダムの翌日の日流入量である。期間は、1980-81、1981-82、1982-83の12月14日から翌年2月11日までの1期間60日(合計180日)とする。牧尾ダムと王滝ダムおよび三浦ダムの位置関係を図1に示す。

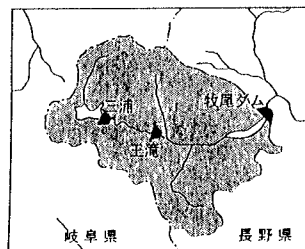


図1 牧尾ダム、王滝ダム、
三浦ダムの位置

2. 計算方法

入力層、中間層、出力層の3層からなる階層型ニューラルネットワークを用い、モーメント修正法を使ったバックプロパゲーション法により計算を行った。予測精度を判定するためには、推定値と実データとの比較を行う必要がある。そこで全データ3期間のうち、1期間を除いた残りの2期間を教師信号として与え、入力データと教師信号について最小二乗誤差学習となるように学習をさせる。その後、除外した1期間を未学習データとして日流入量を予測する。入力データは、観測値以外の日流入量および日降水量にも対応できるように、0.1から0.9の間で基準化を行った。上流地点の日降水量を説明変数に与える場合と与えない場合、それぞれにおいて長尾ら²⁾により構築されたニューラルネットワーク構造により、学習回数を最大5000回として計算した。なお、学習回数100回ごとに誤差分散、相関係数を求めた。

3. 計算と結果

(1) 上流地点の日降水量を説明変数として与えない場合

入力層は、牧尾ダムの2日前まで日流入量と同ダムの3日前までの日降水量とを説明変数とした入力層ユニット5とする。これに中間層ユニット3、出力層ユニット1を加えたネットワーク構造を用いた。

(2) 上流地点の日降水量を説明変数として与える場合

牧尾ダムの2日前までの日流入量と王滝、三浦それぞれの2日前まで日降水量とを説明変数とした入力層ユニット5、中間層ユニット5、出力層ユニット1とするネットワーク構造を用いた。最終的に牧尾ダムの日降水量を説明変数から除外するという結果となった。牧尾ダムへの降雨は流出が速く翌日の流入量に影響しないため説明変数から除外されたものと思われる。例として1980年を入力データとした場合の未学習データの誤差分散を図2に、相関係数を図3に示す。上流の日降水量を与えない場合の誤差分散の平均が0.013、上流の日降水量を与える場合は0.007となった。実測値と予測値の相関係数は上流の日降水量を与えない場合は平均で0.469であり、上流の日降水量を与える場合は0.795である。これらより上流の日降

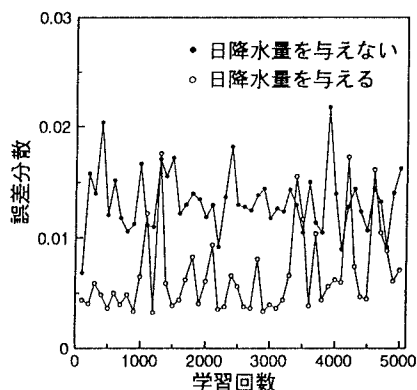


図2 1980未学習データの誤差分散

水量を説明変数に加えたほうが、飛躍的に予測精度が向上することがわかる。

未学習データでは過学習による予測精度の低下が、鈴木ら²⁾により示されている。本例でも、学習回数3400回において過学習による予測精度の低下がみられる。今回は3ケースとも1000回から2000回という少ない学習回数において最適な推定値を出力している。よって未学習データにおいては過学習を避けるために、最適な学習回数において学習を終了することが必要である。未学習データにおいては、入力データの変動の幅が大きいと予測精度が極端に低下する。これは、基準化する際に説明変数に1つでも大きな値が存在するとその値により小さな値の影響が無効となるためと思われる。そこで今後、説明変数の選択も重要であるが、選択された説明変数を基準化する際に各変数の変動状況を反映するような基準化が胆要であると思われる。学習データにおいては学習回数が多いほど予測精度が向上する。しかし学習データにおいては説明変数に上流地点の日降水量を与えない場合のほうが誤差分散と相関係数の両方において良い結果を示す場合もあった。また学習データにおいては、必要最小限の説明変数を与えたほうが予測精度が向上するということがわかった。未学習データにおいては、説明変数が多いほうが予測精度が向上した。

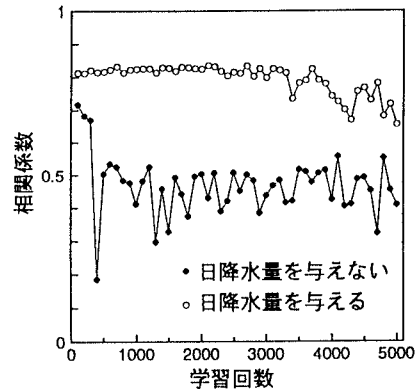


図3 1980未学習データの相関係数

5. 重線形回帰モデルとニューラルネットワークの比較

上流の日降水量を説明変数に加えた重回帰線形モデルとニューラルネットワークの最適な学習回数での誤差分散および相関係数を比較する。誤差分散は表1に、相関係数は表2に示す。

表1 重線形回帰とニューラルネットワークの誤差分散

推定年	重線形回帰		ニューラルネットワーク	
	学習	未学習	学習(回)	未学習(回)
1980	0.003	0.004	0.007(5000)	0.003(1200)
1981	0.005	0.008	0.007(4900)	0.007(4400)
1982	0.012	0.016	0.002(3600)	0.016(1300)
平均	0.007	0.009	0.005	0.016

表2 重線形回帰とニューラルネットワークの相関係数

推定年	重線形回帰		ニューラルネットワーク	
	学習	未学習	学習(回)	未学習(回)
1980	0.869	0.823	0.804(5000)	0.837(2100)
1981	0.756	0.621	0.767(4500)	0.637(3000)
1982	0.669	0.544	0.923(4800)	0.576(1500)
平均	0.765	0.627	0.831	0.683

誤差分散についていうと学習データに対して重線形回帰法が良い結果が得られた。しかし未学習データに対しては、ニューラルネットワークのほうが約6%誤差分散が小さい結果が得られた。相関係数は、ニューラルネットワークのほうが、学習データにおいて約8%、また未学習データにおいても約3%大きい結果が得られた。これらから慣用の予測手法である重線形回帰法と比べるとそれ以上または同程度の予測精度であるということがわかった。

6. 結論と問題点

結論 1) ニューラルネットワークの日流入量予測への有用性を示した。

2) 説明変数に上流の日降水量を加えることにより未学習データの予測精度が向上した。

3) 予測値と観測値の相関が最適な学習回数において重線形回帰法より優れた結果が得られた。

問題点 1) 未学習データにおいて過学習を回避する学習方法の導入

2) 説明変数とする入力データの量的変動状況を反映した基準化

参考文献 1) 長尾他:ニューラルネットワークによる渇水期日流量予測(1), 土木学会中部支部論文集, pp. 269-270, 1995

2) 鈴木他:ニューラルネットワークによる渇水期日流量予測(2), 土木学会中部支部論文集, pp. 271-272, 1995