

水文時系列情報のフィードバックを考慮したニューロモデルによる日流量予測

名古屋工業大学大学院 学生員○佐野 正嗣
 名古屋工業大学 正会員 長尾 正志
 名古屋工業大学 正会員 庄 建治朗

1. はじめに

流入量予測は合理的な貯水池の運用の基礎となる。有効に貯水池を運用するためにもより将来に及ぶ予測が要求される。そこで本研究では1,2日先の日流量を階層型ニューラルネットワークで予測する。時系列情報を複数同時に予測しようとするときには、出力値を入力にフィードバックさせることにより時系列データの持つ自己相関性を反映させることができるといわれている¹⁾。そこで今回は牧尾ダムの1,2日先の日流量をフィードバックのあるモデルとないモデルで予測しその効果を比較・検討した結果を報告する。

2. 対象資料および計算方法

(1) 対象資料

目的変数は牧尾ダムの翌日や翌々日の日流量とし、説明変数は従前の研究より、牧尾ダムの2日前までの日流量および同ダムの上流に位置する王滝、三浦ダムのそれぞれ2日前までの日雨量とする。学習・予測期間は1979~86年の12月14日から翌年2月11日までの1期間60日を対象とし、データの基準化には1969年から1990年までの22年間の期間を使用する。(0,1)への基準化は、日流量については岩井法により対数正規分布に従うように変換したものの非超過確率を、日雨量については雨量0の日が突出して多いので0を除いた観測値を指数分布にあてはめたものの非超過確率を使用した。

(2) 予測モデル

入力層、中間層、出力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う。データの自己相関性を反映させるために出力を入力にフィードバックさせるニューロモデルを採用する。その際、入力情報に日流量と日雨量という異質のデータを使用しているので、中間層ユニットを日流量と日雨量に対するユニットに分けることを考えた。フィードバックを日流量に対するユニットのみに行うモデル(SLモデル)と中間層ユニットを分けないモデル(FBモデル)を扱う。また比較のためにフィードバックのないモデル(SBモデル)も使用する。

(3) 計算方法

学習データ数は2期間120とし、各モデルに対し、それぞれ次の方法で計算する。中間層ユニット数を入力層ユニットと同数から1つづ減らしていき、出力層と同数の2までの各ニューロモデルを設定し、学習回数を最大5,000回とし、各100回ごとに予測値と観測値の誤差分散と結合係数を求める。この誤差分散のうち、学習が終わったと考えられる学習回数4,000回から5,000回のものを用いて各評価基準AIC, MDLを求め、各評価基準の値を平均する。AIC, MDLに小さい方から順位をつける。各ユニット数のAICとMDLの順位を加算した和が最小となる中間層ユニットを最適構造として選定する。教師信号のつぎの年の資料に対する予測値と観測値の誤差分散を求め、誤差分散の10個ごとの移動平均の最小値から、最適なモデル定数を推定し、選定された定数によって予測を行う。

3. 結果と検討

(1) ニューロモデルの構造選定

学習データに対する情報量基準の順位和により選定された構造を表1に示す。なおMIは中間層ユニット数を表す。SLモデルでのMIIは日流量に対する中間層ユニット数を表し、MIRは日雨量に対する中間層ユニット数を表す。

表1 各モデルの最適選定構造

モデル種別	79-80年	80-81年	81-82年	82-83年	83-84年
FBモデル	MI=2	MI=3	MI=2	MI=3	MI=2
SLモデル	MII=2, MIR=1	MII=1, MIR=4	MII=1, MIR=2	MII=1, MIR=1	MII=1, MIR=1
SBモデル	MI=2	MI=4	MI=4	MI=2	MI=2

(2) 予測精度の比較

最適なモデル定数での基準化した1日先、2日先の予測値と観測値の相関係数を表2に示す。なお、表で

表2 1日先、2日先の予測値と観測値の相関係数

	1日目						2日目					
	翌年			2年目			翌年			2年目		
学習データ	FBモデル	SLモデル	SBモデル	FBモデル	SLモデル	SBモデル	FBモデル	SLモデル	SBモデル	FBモデル	SLモデル	SBモデル
79-80年	0.597	0.628	0.641	0.619	0.636	0.613	0.506	0.539	0.539	0.422	0.441	0.427
80-81年	0.477	0.548	0.478	0.919	0.896	0.899	0.371	0.385	0.372	0.865	0.888	0.857
81-82年	0.876	0.890	0.889	0.042	0.177	0.052	0.844	0.883	0.833	0.192	0.141	0.127
82-83年	0.189	0.134	0.186	0.871	0.892	0.688	0.149	0.109	0.134	0.820	0.873	0.656
83-84年	0.832	0.785	0.756	0.516	0.408	0.346	0.780	0.772	0.739	0.417	0.201	0.276
平均	0.594	0.597	0.590	0.593	0.602	0.520	0.530	0.538	0.523	0.543	0.509	0.469

学習データが79-80年なら翌年とは81年を、2年目とは82年を意味する。翌年はモデル定数の決定に使用しているので既知データではあるが、未学習データなので評価対象とする。予測精度は対象年度によりかなり違う結果となった。これは表3に示す日流量データの自己相関係数の違いに影響を受けている。一般にフィードバックのあるモデルは強い自己相関のデータには精度がよく、自己相関が弱くなるとフィードバックのないモデルの方が精度がよくなるといわれる²⁾。しかし今回の結果では、表3で学習データ80-81年、82-83年のように自己相関が似た年であっても、表2のように学習データ80-81年ではフィードバックの影響がほとんどないが、学習データ82-83年ではフィードバックのあるモデルの方が精度がややよくなっている。この違いは現時点では解明されていない。フィードバックの影響をみるために、82-83年学習でのFBモデルとSBモデルによる85年の経過日数と予測流量の比較を図1に示す。また中間層ユニットの分離の差違をみるために、83-84年学習での86年の経過日数と予測流量の比較を図2に示す。図1よりSBモデルでは大きく振動しているが、FBモデルではそれがやや押さえられている。図2よりSLモデルがほぼ一定の予測をしているのに、FBモデルでは観測値の変化に対応した予測をしている。なお30日付近でFBモデルが急激に少ない予測をしているのは学習データに同じように急激に少なくなっている値があったのでその影響がでたものと思われる。

表3 基準化した日流量データの自己相関係数

	1次			3次		
	学習	翌年	2年目	学習	翌年	2年目
学習データ						
79-80年	0.870	0.605	0.609	0.755	0.528	0.451
80-81年	0.728	0.609	0.941	0.675	0.451	0.894
81-82年	0.607	0.941	0.317	0.490	0.894	0.109
82-83年	0.775	0.317	0.923	0.673	0.109	0.883
83-84年	0.629	0.923	0.729	0.502	0.883	0.361

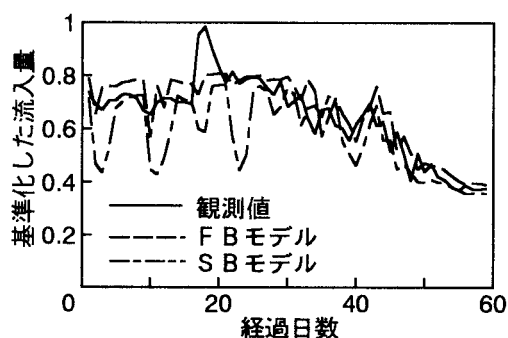


図1 82-83年学習による85年予測(1日先)

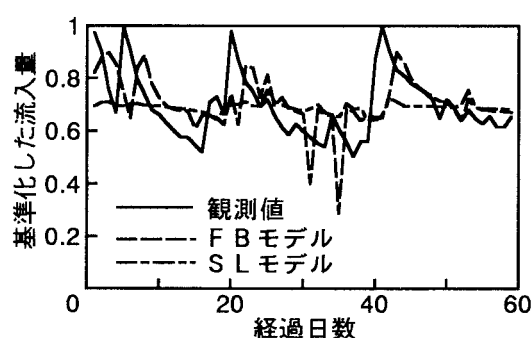


図2 83-84年学習による86年予測(2日先)

4. 結論

フィードバックの有無で大きな違いは認められなかったが、フィードバックの効果がみられる場合もあるのでフィードバックモデルの方がよいように思われる。なお、フィードバックモデルでは中間層を降雨と流量要因で分けられない方がよいようである。

参考文献 1), 2) 合原一幸編著：ニューロ・ファジィ・カオス，オーム社，pp. 62-67，1993