

추계학적 신경망 모형에 의한 다목적댐의 유입량산정

○김성원*·조정석**·정재용***

1. 서론

1990년대 이후로 신경망이론은 수자원의 양적 및 질적 변량을 모형화 하는데 전통적인 방법을 대체할 수 있는 효과적인 기법으로 사용되어 오고 있다. 특히 최근에 와서는 신경망에 기반을 둔 많은 모형이 유출예측, 가뭄해석 및 저수지의 유입량의 예측에 적용되고 있다(Tokar and Johnson, 1999; Zealand et al., 1999; 김성원 외, 2001; 신현석과 박무중, 1999; Jain et al., 1999; Coulibaly, 2000a,b). 일반적으로 수공학에 있어서의 대부분의 신경망모형은 역전파 알고리즘을 이용한 다층퍼셉트론 신경망모형이다. 다목적 댐의 월별유입량을 정확하게 산정하는 것은 수자원 기술자 및 관리자들에게 주요 과업 중 하나이며, 기술적, 경제적 그리고 정치적인 범주 내에서 수자원시스템의 효과적인 운영을 가능하게 하며, 가뭄해석, 관개용수 및 발전소계획을 위하여 사용하고 있다. 다목적 댐의 월별유입량을 산정하는 기법은 일반적으로 개념적/물리적인 기법과 통계학적/추계학적 기법이 있다. 첫 번째 범주인 개념적/물리적인 개념은 유역에서 발생하는 수문학적 순환을 수학적 모형으로 표현하는 기법을 나타내고 있으며, 모형을 구성하는 여러 인자들을 통하여 미분형식의 연속 및 운동량 방정식의 해법을 이용하고 있다. 두 번째 범주인 통계학적/추계학적 기법은 단순선형, 다중선형, 비선형 회귀식, ARMA, XRAMA 기법을 의미하고 있으며, 일반적으로 다목적 댐 유입량을 산정하는 모형 구성인자의 시계열이 일치하지 않을 경우 모형의 구성 및 매개변수의 산정이 어렵다. 따라서 본 연구에서 제시하는 추계학적 신경망모형은 두 번째 범주에 속한다고 할 수 있으며, 본 연구의 목적은 낙동강유역의 안동다목적댐에서 입력변수의 자료시계열이 일치하지 않을 경우 월별유입량을 산정하는 추계학적 신경망모형을 제시하고, 훈련 및 검증을 통하여 모형의 적합성을 검토하고자 한다.

* 정희원, 동양대학교 지구환경시스템공학과, 전임강사(E-mail : swkim68@phenix.dyu.ac.kr)

** 정희원, 대구대학교 건설환경공학부, 교수(E-mail : jscho@taegu.ac.kr)

*** 정희원, 미국 콜로라도 주립대학교 토목공학과, 박사과정(E-mail : jaeyong@engr.colostate.edu)

2. 자료의 수집

본 연구에 적용할 신경망모형의 입력자료는 1977년부터 1997년까지의 안동다목적댐의 월평균유입량, 월면적강우량과 1983년부터 1990년까지의 월별 증발접시 증발량, 1983년부터 2000년까지의 안동관측소의 월평균기온 자료로 구성되어 있다. 자료의 수집은 안동다목적댐의 월평균유입량은 한국건설기술연구원(KICT)의 수문자료 D/B를 이용하였으며, 월면적강우량은 영주, 부석, 안동, 석포, 길안, 미질, 재산, 수비관측소의 일강우량을 이용하여 티센가중치에 의해 면적강우량을 산정하여 월별로 합산하였다. 또한 월 증발접시 증발량은 안동관측소의 폐쇄 및 이전 그리고 결측에 의하여 신뢰성 있는 시계열은 1983년부터 1990년까지이며 기상년보를 이용하였다. 그리고 안동관측소의 월평균기온은 기상청 웹페이지(www.kma.go.kr)에서 다운로드 받았다.

3. 수문자료의 분석

본 연구의 신경망모형의 입력자료는 월평균유입량, 월면적강우량, 월 증발접시 증발량과 월평균기온으로 구성되어 있지만, 실제 수문분석에서는 월평균유입량과 월면적강우량이 주요자료이다. 따라서 월별자료를 연단위로 합산한 연평균유입량과 연면적강우량에 대한 수문자료 분석결과는 다음 표 1과 같으며, 계열상관도(Correlogram)는 그림 1과 같다.

표 1. 수문자료의 분석

Data		Inflow(cms)	Rainfall(mm)
Statistics			
Mean		354.50	1,071.35
Standard Deviation		117.28	245.43
Skewness Coefficient		0.23	0.38
Coefficient of Variation		0.33	0.23
Maximum		575.00	1,567.30
Minimum		180.50	716.40
Correlation Coefficient	Lag-1	-0.0709	0.1574
	Lag-2	-0.0681	-0.0504

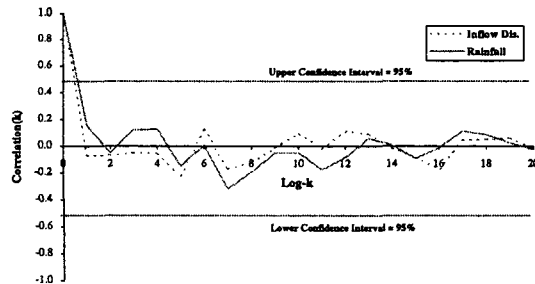


그림 1. 계열상관도

4. 입력자료의 모의발생

본 연구에서 부족한 자료 시계열을 PARMA(1,1)모형에 의하여 월평균유입량, 월면적강우량, 월 증발접시 증발량과 월평균기온 자료를 모의발생하였으며, 제시된 PARMA(1,1)모형은 다음 식 1과 같이 나타낼 수 있다.

$$y_{\nu, \tau} = \mu_{\tau} + \phi_{1, \tau} (y_{\nu, \tau-1} - \mu_{\tau-1}) + \varepsilon_{\nu, \tau} - \theta_{1, \tau} \varepsilon_{\nu, \tau-1} \quad [1]$$

여기서 ν 는 년(Year), τ 는 계절(Season)이고 $\tau = 1, 2, \dots, \omega$ 이며 계절의 수를 나타내고 있다. 일반적으로 PARMA(1,1)모형은 월유출량을 모의하는데 사용된다(Salas et al., 1980). 본 연구의 PARMA(1,1)모형은 SAMS(Salas J.D., 1998) Package를 이용하였다. 그리고 계절의 수는 12계절로 하였으며, 모의발생한 표본

(Sample)은 10개의 표본이며 각 표본마다 100년씩의 월평균유입량, 월면적강우량, 월 증발접시 증발량과 월평균기온 자료를 모의발생하였다. 모의된 자료중에서 10번째 표본의 초기에 발생되는 초기조건의 편차를 제거하기 위하여 50년간의 자료만을 이용하여 추계학적 신경망모형의 훈련자료로 사용하였다. 다음 표 2는 PARMA(1,1)모형에 의해 산정된 월평균유입량과 월면적강우량의 매개변수값이고, 근사최소자승법(Method of Approximate Least Square)에 의해 산정하였다.

표 2. PARMA(1,1)모형의 매개변수 산정

Data Month	Inflow		Rainfall	
	Phi-1	Theta-1	Phi-1	Theta-1
1	0.404	0.190	0.433	0.302
2	9.718	8.092	6.861	6.037
3	0.671	0.512	0.415	0.427
4	0.339	0.326	-0.150	-1.757
5	0.315	0.172	-0.727	-0.496
6	5.784	5.765	1.401	1.455
7	-0.113	-0.436	2.300	2.674
8	0.286	-0.041	0.792	0.698
9	-0.027	0.010	-0.456	-0.302
10	-0.339	-0.386	-0.043	0.061
11	0.482	0.497	-0.471	-0.273
12	-1.367	-1.659	-0.113	-0.184

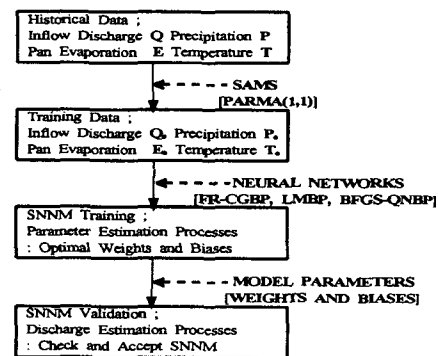


그림 2. 추계학적 신경망모형의 개발과정

5. 추계학적 신경망모형의 구축

본 연구에서 제시하는 추계학적 신경망모형에 의한 진행과정은 그림 2와 같이 나타낼 수 있고, 추계학적 신경망모형은 그림 3과 같다. 본 연구의 추계학적 신경망모형은 역전파 알고리즘을 이용한 다층퍼셉트론(MultiLayer Perceptron, MLP)형태의 기법이며, 역전파 알고리즘은 주어진 문제에 대하여 비최적화 해법을 나타내는 구배강하법이다. 본 연구에 사용된 역전파 알고리즘은 Fletcher-Reeves 공액구배 역전파(FR-CGBP)알고리즘, Levenberg-Marquardt 역전파(LMBP)알고리즘과 BFGS 의사뉴턴(BFGS-QNBP)알고리즘이며, 상세한 내용은 다음과 같다.

5.1 FR-CGBP 알고리즘

FR-CGBP 알고리즘은 연결강도와 편차에 대해 구배연산과 연속적인 보정에 있어서 일반적인 역전파 알고리즘과는 다르다. 탐색방향(Search direction) \bar{d}_k 는 각 훈련 반복횟수 k에서 계산되고, 오차함수 $E(X)$ 는 탐색방향을 따라서 최소화된다. 즉 구배하강은 오차경사의 방향으로 진행되는 것이 아니고 전 단계 오차의 연직방향으로 진행되며, 구배변화는 각 단계에서 수행된 오차최소화가 연속적인 단계로부터 어떤 장애도 없이 완전하게 보전되는 장점 때문에 이전단계의 연직으로 작용한다(Fletcher and Reeves, 1964).

5.2 LMBP 알고리즘

LMBP 알고리즘은 비선형 함수의 오차자승의 합을 최소화하는 최적기법으로 뉴턴법의 변형법이다. LMBP 알고리즘의 주요과정은 자코비안 행렬을 계산하는 것이며, 자코비안 행렬을 계산하기 위해서는 최급하강 역전파 알고리즘의 변형을 이용하여야 한다(Hagen and Menhaj, 1994).

5.3 BFGS-QNBP 알고리즘

Shanno는 공액구배 역전파(CGBP)알고리즘의 여러 가지 변형된 기법에 대한 연구를 통하여 광범위한 문제에 대하여 성능이 뛰어난 기법인 Broyden-Fletcher-Goldfarb-Shanno(BFGS) 의사뉴턴(Quasi Newton)역전파 알고리즘을 제시하였다(Shanno, 1978).

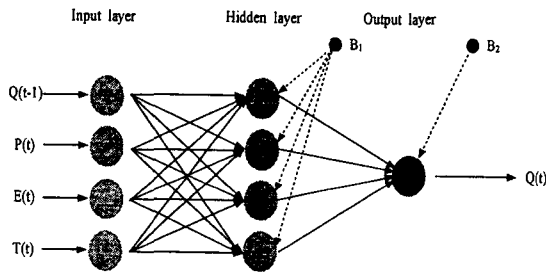


그림 3. 제시된 추계학적 신경망모형

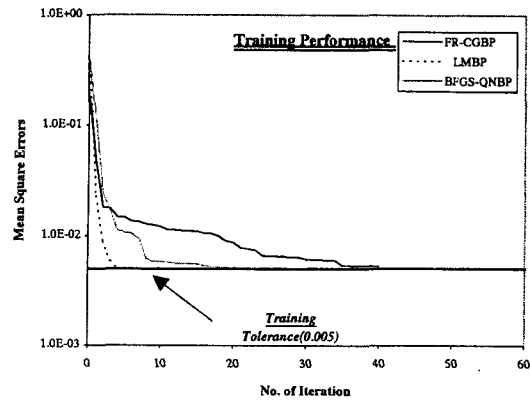


그림 4. 평균제곱오차의 수렴과정

6. 추계학적 신경망모형의 훈련

추계학적 신경망모형의 훈련을 위하여 훈련임계치(Training Tolerance)의 허용한계를 0.005로 설정하였다. 그 이유는 0.001를 허용한계로 설정할 경우 대부분의 추계학적 신경망모형의 훈련에서 평균제곱오차(MSE)는 수렴되지 않고 발산하기 때문에 경험상의 수렴한계치인 0.005를 훈련임계치로 설정하였다. 그림 4는 FR-CGBP, LMBP, BFGS-QNBP 추계학적 신경망모형의 훈련에 있어서 반복횟수에 따른 평균제곱오차의 수렴과정을 나타낸 것이고, 표 3은 훈련결과의 통계분석을 나타낸 것이다. 통계분석결과는 LMBP 추계학적 신경망모형이 PARMA(1,1)모형에 의해 모의발생된 유입량을 가장 정확하게 재현하는 것으로 나타났다. 그리고 FR-CGBP, LMBP, BFGS-QNBP 추계학적 신경망모형의 훈련에 의해 산정된 최적 연결강도와 편차를 추계학적 신경망모형의 검증에 이용하였다. 다음 표 4는 LMBP 추계학적 신경망모형의 훈련에 의해 산정된 최적연결강도와 편차를 나타낸 것이다.

표 3. 추계학적 신경망모형 훈련결과와 통계분석

Data	Algorithm	Network Configuration	Statistical Analysis			
			m(e)(cms)	s(e)(cms)	s(e)/s(y)	CC
Generating (Training)	FR-CGBP	4-4-1	-0.018	16.531	0.441	0.90
	LMBP	4-4-1	3.078	13.840	0.369	0.93
	BFGS-QNBP	4-4-1	2.074	15.972	0.426	0.91

표 4. 최적연결강도와 편차(LMBP)

Hidden Layer		Input Variables and Weights				Output Layer
		Q(t-1)	P(t)	E(t)	T(t)	
Node j	Bias B ₁	W _{j1}	W _{j2}	W _{j3}	W _{j4}	W _{ij}
1	-0.984	0.612	-8.286	1.025	4.676	-0.283
2	6.107	-0.537	-6.286	6.412	4.902	0.209
3	7.483	-1.403	-3.760	1.162	-4.793	-0.796
4	4.011	-4.030	-4.387	-2.536	-2.661	-0.299
Bias B ₂						1.079

7. 추계학적 신경망모형의 검증

추계학적 신경망모형의 검증은 훈련을 통하여 산정된 최적연결강도와 편차를 이용하여 실제 관측자료를 입력자료로 하여 월평균유입량을 산정하는 것이다. 실측자료는 각 변수에 따라 자료시계열이 일치하지 않기 때문에 공통시계열에 속하는 1983년부터 1990년까지의 월평균유입량, 월면적강우량, 월 증발점시 증발량과 월평균기온 자료를 이용하여 추계학적 신경망모형을 검증하였으며, 검증결과 LMBP 추계학적 신경망모형이 가장 우수한 결과를 나타내었다. 다음 표 5는 추계학적 신경망모형의 검증결과이며, 그림 5는 1983년부터 1990년까지의 안동댐 실측 월평균유입량과 추계학적 신경망모형의 검증결과를 비교한 것이다.

표 5. 추계학적 신경망모형의 검증에 대한 통계분석

Data	Algorithm	Network Configuration	Statistical Analysis			
			m(e)(cms)	s(e)(cms)	s(e)/s(y)	CC
Historic (Validation)	FR-CGBP	4-4-1	-0.018	16.531	0.441	0.90
	LMBP	4-4-1	3.078	13.840	0.369	0.93
	BFGS-QNBP	4-4-1	2.074	15.972	0.426	0.91

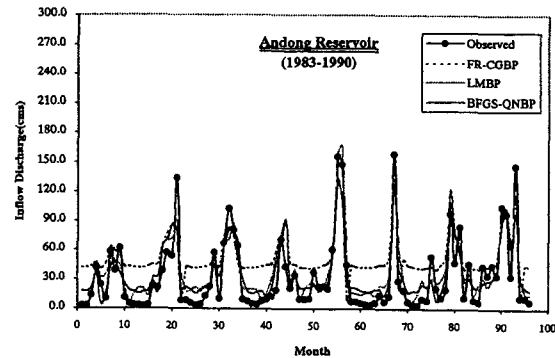


그림 5. 월평균유입량 수문곡선의 비교

8. 결론

본 연구에서는 다목적댐의 장기유입량을 산정할 수 있는 추계학적 신경망모형을 제시하였다. 사용된 자료는 안동댐 월평균유입량, 월면적강우량, 월 증발점시 증발량과 월평균기온 자료이다. 그리고 PARMA(1,1)모형을 이용하여 부족한 자료를 모의발생하였으며, 모의발생된 자료로서 추계학적 신경망모형을 훈련하여 최적연결강도와 편차를 산정하였다. 또한 추계학적 신경망모형의 훈련에 의해 산정된 최적연결강도와 편차를 이용하여 실측자료로서 추계학적 신경망모형을 검증한 결과 LMBP 추계학적 신경망모형이 가장 우수한 모형으로

판정되었다. 따라서 추계학적모형과 신경망모형을 연계하여 다목적댐의 월평균유입량의 산정과정은 합리적이고 타당하며, 기존의 물리적/개념적모형에 비교해서 결코 떨어지지 않는 우수한 모형이라 사료된다.

참고문헌

1. 건설교통부, 한국수자원공사(1991). 영천댐 도수로 도수에 따른 안동댐 및 임하댐 연계운영방안 보고서.
2. 기상청(1983-1990). 기상년보.
3. 김성원, 이순탁, 조정석(2001). “중소하천유역에서 Hybrid Neural Networks에 의한 수문학적 예측”, 한국수자원학회 논문집, 제 34권, 제 4호, pp. 303-316.
4. 신현석, 박무중(1999). “신경망을 이용한 우리나라의 시공간적 가뭄의 해석”, 한국수자원학회 논문집, 제 32권, 제 1호, pp. 15-29.
5. Fletcher, R., and Reeves, C.M.(1964). “Function minimization by conjugate gradient” *Computer Journal*, Vol. 7, pp. 149-154.
6. Jain, S.K., Das, D., and Srivastava, D.K.(1999). “Application of ANN for reservoir inflow prediction and operation” *J. Water Resour. Plng. and Mgmt.*, ASCE, Vol. 125, No. 5, pp. 263-271.
7. Coulibaly, P., Anctil, F., and Bobee, B.(2000a). “Daily reservoir inflow forecasting using artificial neural networks with stopped training approach” *J. Hydro.*, Vol. 230, pp. 244-257.
8. Coulibaly, P., Anctil, F., and Bobee, B.(2000b). “Neural network-based long-term hydropower forecasting system” *J.Comp. Aided Civ. and Infrastruct. Engrg.*, Vol. 15, No. 5, pp. 355-257.
9. Hagen, M.T., and Menhaj, M.(1994). “Training feedforward networks with the Marquardt algorithm” *IEEE Transactions on Neural Networks*, Vol. 5, No. 6. pp. 25-32
10. Salas, J.D.(1998) *SAMS ; Stochastic Analysis, Modeling, and Simulation user manual*, Colorado State University, Fort Collins, CO.
11. Salas, J.D., Delleur, J.R., Yevjevich, V., Lane, W.L.(1980). *Applied Modeling of Hydrologic Time Series, Water Resources Publications*, Littleton, CO.
12. Shanno, D.F.(1978). “Conjugate gradient methods with inexact searches” *Mathematics of Operations Research*, Vol. 3, No. 3, pp. 244-256.
13. Tokar, A.S., and Johnson, P.A.(1999). “Rainfall-runoff modeling using artificial neural networks” *J. Hydrologic Engrg.*, ASCE, Vol. 4, No. 3, pp. 232-239.
14. Zealand, C.M., Burn, D.H., and Simonovic, S.P.(1999). “Short term streamflow forecasting using artificial neural networks” *J. Hydro.*, Vol. 214, pp. 32-48.