

신경망을 이용한 강우 결측지점의 보완과 강우예측

안상진* · 최병만** · ○연인성*** · 김기석****

1. 서론

강우는 계절라성이라고 불릴만큼 극부적인 성격이 강해지면서 재산피해 뿐만아니라 인명피해까지 유발되고 있으며 예측이 또한 어렵다. 수자원의 불균형이 심한 우리나라의 경우 장기간 실측된 수문기상 기초자료가 부족한 실정이어서 수공구조물 설계를 위해 정확한 강우량 자료의 취득이 선행되어야 한다.

특히 호우로 피해를 입는 지역은 수해등으로 인하여 강우 결측이 생기며 차후 강우 예측에도 많은 어려움이 있을 것으로 판단된다. 본 연구에서는 기존의 강우 결측보완 방법들과 신경망 모형을 비교하여 결측보완에 대한 신경망 모형의 타당성을 검토하고 근래의 자료들을 대상으로 결측지점에서의 자료의 보완과 실시간 강우예측을 동시에 수행하고자 한다.

2. 기본 이론

2.1 신경망 이론

신경망은 뉴런들이 거미줄처럼 복잡하게 연결되어 상호작용을 하는 생물학적 신경망의 반응 전달 과정을 수학적으로 모형화한 것이다. 연산능력을 가진 노드(node)들이 학습에 의해 작동하도록 구성된다. 연산과정은 입력과 연결강도 벡터들간의 곱의 형태로 표현되는 결합함수, 결합함수로부터 도출된 결과를 가지고 반응 여부를 결정하는 활성화함수, 그리고 전달을 위한 출력의 과정으로 형성되며 그림 1과 같다.

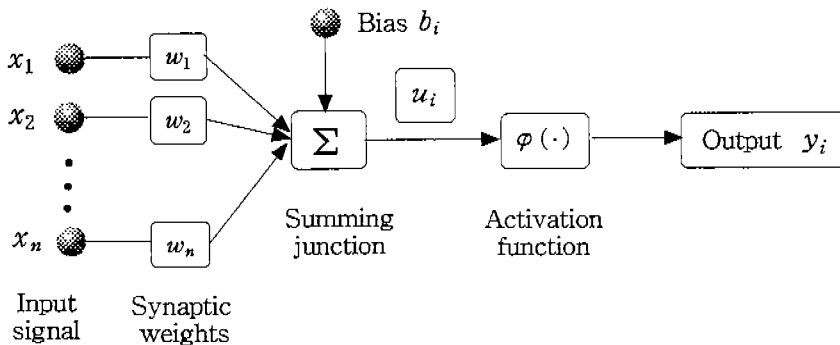


그림 1. 모형의 비선형 연산과정

* 정회원, 충북대학교 공과대학 교수
** 정회원, 충북대학교 대학원 토목공학과 박사과정, 수자원공사 조사기획처
*** 정회원, 충북대학교 대학원 토목공학과 박사과정
**** 정회원, 충북대학교 대학원 토목공학과 석사과정

입력 X 와 연결강도 W 는 식(1), 식(2)와 같이 벡터형태로 표시되며 가중합 s_i 는 식(3), 바이어스 b_i 를 포함한 가중합 u_i 는 식(4), 그리고 출력값 y_i 는 식(5)와 같다.

$$X = [x_1 \ x_2 \ \dots \ x_n] \quad (1)$$

$$W = [w_1 \ w_2 \ \dots \ w_n] \quad (2)$$

$$s_i = \sum_{p=1}^n w_p x_p \\ = x_1 w_1 + x_2 w_2 + \dots + x_n w_n \quad (3)$$

$$= XW^T$$

$$u_i = s_i + b_i \quad (4)$$

$$y_i = \varphi(s_i + b_i) = \varphi(u_i) \quad (5)$$

여기서 w_p 는 단일 출력시 입력층 뉴런 p 의 출력값에 대한 연결강도, x_p 는 입력층 뉴런, b_i 는 형성되어 있는 연결강도에 영향을 갖는 편견을 나타내며 φ 는 활성화 함수로서 비선형 연속함수이다.

2.2 결측보완 방법

RDS 방법은 결측지점으로부터 주위 관측점까지의 거리에 따른 가중인자를 결정하여 보완하는 방법으로 평면좌표계에서 결측지점의 좌표를 (x_0, y_0) , 주위의 관측점 좌표를 (x_t, y_t) 라 할 때, 두 지점간의 제곱거리 D_t^2 는 식(6)과 같고 t 관측점의 가중인자 W_t 는 식(7)과 같으며 t 관측점에서의 강우량 P_t 와 가중인자의 곱에 의한 결측지점의 보완 강우량은 식(8)과 같다. 또한 표 1에 기타결측보완 방법들에 의한 보완강우량 산정방법을 나타내었다.

$$D_t^2 = (x_t - x_0)^2 + (y_t - y_0)^2 \quad (6)$$

$$W_t = \frac{\frac{1}{D_t^2}}{\sum_{t=1}^n \frac{1}{D_t^2}} \quad (7), \quad R = \sum_{t=1}^n W_t P_t \quad (8)$$

표 1. 기타 결측보완 방법들의 보완강우량 산정 방법

구 분	Thiessen 가중법	정상년강우량법 (Normal Ratio Method)	산술평균법 (Arithmetic Mean Method)
보완강우량(R) 산정	$R = \frac{\sum_{t=1}^n (A_t P_t)}{\sum_{t=1}^n A_t}$	$R = \frac{N_x}{n} \sum_{t=1}^n \left(\frac{P_t}{N_t} \right)$	$R = \frac{1}{n} \sum_{t=1}^n P_t$
비 고	t 관측 지점의 다각형 산정면적 A_t 와 강우량 P_t 로 산정	결측지점에서 정상년평균 강우량 N_x 와 t 관측지점의 정상년 평균 강우량 N_t 로 산정	n 개의 주위 관측소 강우량을 산술평균으로 산정

3. 모형의 구성

3.1 강우 결측보완 신경망 모형

강우 관측소 12개 지점 중 결측 지점을 제외한 11개 지점의 동일시간 시우량 자료를 입력층로 사용하여 결측지점의 동시간 시우량 자료를 출력하는 단일 출력층으로 구성하였다.

3.2 강우 예측 신경망 모형

입력층의 자료로서 동일 호우사상에 대한 동일 시간(t)까지의 12개 강우관측소 시우량 자료를 사용하였으며, 목표값으로 결측지점의 1시간($t+1$), 2시간($t+2$), 3시간($t+3$)후의 강우량을 설정하여 1, 2, 3시간 예측 결과가 동시에 출력되도록 구성하였다.

3.3 두 모형의 연계

강우예측을 위해 결측보완 신경망 모형의 출력치가 강우 예측 모형의 입력층에 자동으로 내삽되어 이용되도록 구성하였으며, 그림 2는 신경망 모형의 연계 모식도를 나타내고 있다.

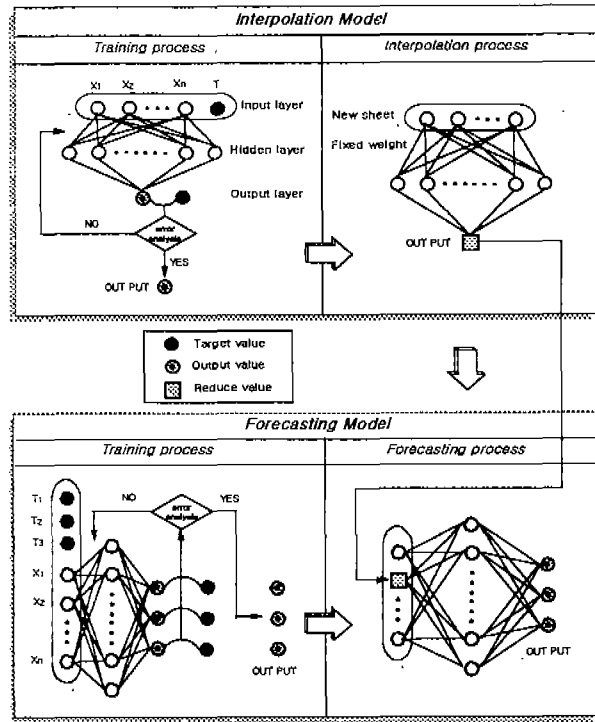


그림 2. 신경망모형의 연계모식도

4. 적용

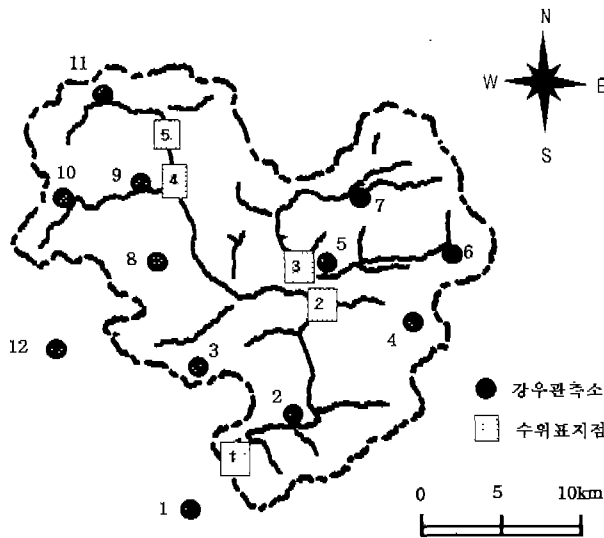


그림 3. 보청천 유역도

그림 3에 도시한 보청천 유역의 강우관측소 No.5(관기)와 주변의 11개 강우관측소, No.1(묘금), No.2(청산), No.3(능월), No.4(중늬), No.6(평운), No.7(삼가), No.8(송정), No.9(삼산), No.10(동정), No.11(이원), No.12(안내)의 12개 지점의 강우기록 자료를 이용하였으며, 1983년 6월부터 1998년 9월까지의 홍수기 호우사상을 중심으로 총 강우량 20mm 이상의 44개의 강우사상을 선정하였다.

선정자료는 시우량 자료로서 3시간 이상의 무강우 상태는 별개의 강우사상으로 간주하였다.

4.1 강우결측보완

보청천 유역의 No.5(관기) 지점을 결측지점으로 가정하고, 은닉층과 학습회수를 달리하여 83년부터 94년까지의 30개 강우사상을 학습여 12개의 결측보완 신경망 모형을 구성하였다. 학습하지 않은 95년 이후의 11개 강우사상을 대상으로 각 모형에 의해 보완된 결측자료를 분석하여 표 2에 나타내었다.

표 2에서 보는 바와같이 가장 좋은 결과를 보인 INN-IV모형을 최적의 결측보완 신경망 모형으로 선정하였으며 학습하지 않은 95년 7월 11일 호우사상에 대해 기존의 강우 결측보완 방법들 (RDS : 거리역산법, Thiessen : Thiessen 가중법, NRM : 정상년 강우량법, AMM : 산술평균법)과 비교하여 표 3에 나타내었으며 결과를 그림 4와 그림 5에 도시하였다.

표 2. 강우 결측보완 신경망 모형들의 보완결과 통계분석

INN Model	Configuration of Network		Error		
	Hidden Layer	Training iterations	CC	NRMSE	NMAXE
INN-I	11 (N)	100	0.7627	0.5338	0.5815
INN-II		500	0.7452	0.5503	0.7343
INN-III		1000	0.7311	0.5714	0.7786
INN-IV	22 (2N)	100	0.7781	0.5196	0.5698
INN-V		500	0.7485	0.5519	0.6129
INN-VI		1000	0.6946	0.5955	0.6937
INN-VII	33 (3N)	100	0.7548	0.5415	0.5962
INN-VIII		500	0.7722	0.5240	0.5900
INN-IX		1000	0.7559	0.5479	0.6335
INN-X	44 (4N)	100	0.7429	0.5578	0.7025
INN-XI		500	0.7194	0.5744	0.5601
INN-XII		1000	0.7337	0.5634	0.7851

표 3. 각 방법에 의한 결측보완 결과 비교

Rainfal Events	Item	NMAXE	NRMSE	CC	PEP	0 th Moment Error
950711	INN-IV	0.2065	0.2441	0.9577	-10.6468	0.0734
	RDS	0.5735	0.5781	0.8576	-57.354	-0.2879
	Thiessen	0.6194	0.6147	0.8355	-61.9396	-0.3119
	NRM	0.6593	0.6495	0.8199	-65.9253	-0.3825
	AMM	0.6234	0.6217	0.8146	-62.338	-0.3228
960704	INN-IV	0.2500	0.2478	0.9337	-10.1573	0.0322
	RDS	0.3011	0.3138	0.8856	-22.0687	0.0228
	Thiessen	0.3576	0.4071	0.8025	-23.9258	-0.0634
	NRM	0.3781	0.4290	0.7786	-25.6226	-0.0795
	AMM	0.4091	0.4189	0.7903	-21.8182	-0.0125

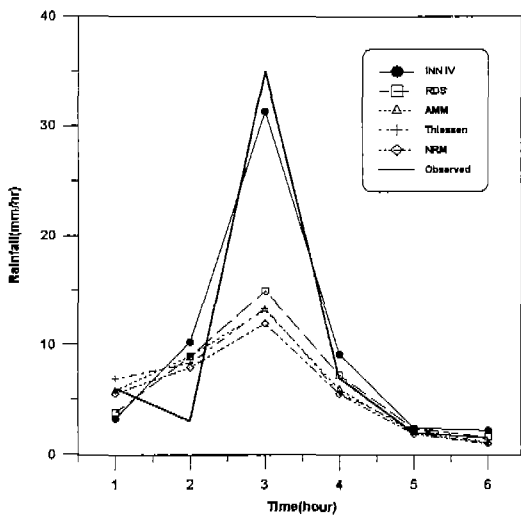


그림 4. 강우 결측보완 결과(1995.07.11)

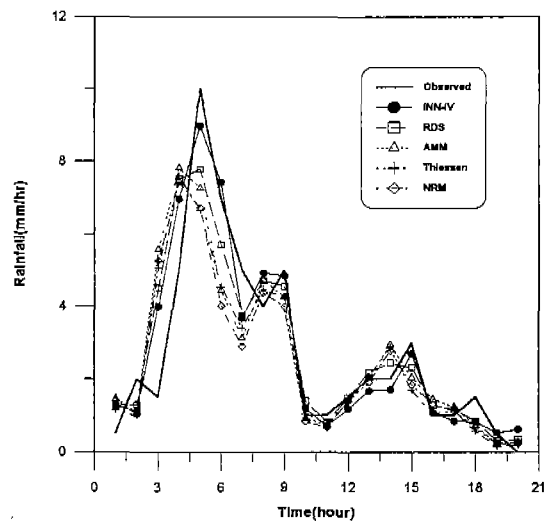


그림 5. 강우 결측보완 결과(1996.07.04)

4.2 강우예측

입력층과 은닉층은 최적 상태로 고정하고 1983년부터 1994년까지의 30개 강우사상을 대상으로 학습하였으며 학습회수에 따라 4가지 종류의 모형으로 구분하였다. 표 4는 1995년에서 1996년까지 No.5(관기)지점의 9개 강우사상에 대해 각 모형에 의해 선행예측 되어진 1~3시간 점강우량을 분석한 결과이다.

표 4. 학습회수에 따른 강우예측 결과 분석

FNN Configuration of Network				Forecasting Time	Error		
Input Layer	Hidden Layer	Output Layer	Training iterations		CC	NRMSE	NMAXE
12(N)	24(2N)	3	100 (FNN I)	1hr	0.5389	0.6942	0.8256
				2hr	0.3912	0.7559	0.8343
				3hr	0.3559	0.7701	0.8860
				Mean	0.4287	0.7401	0.8486
			300 (FNN II)	1hr	0.4423	0.7344	0.9051
				2hr	0.4244	0.7475	0.7945
				3hr	0.2043	0.8053	0.9611
				Mean	0.3570	0.7624	0.8869
			500 (FNN III)	1hr	0.6571	0.6546	0.8038
				2hr	0.3667	0.7638	0.8681
				3hr	0.3227	0.7732	0.9271
				Mean	0.4488	0.7305	0.8663
			1000 (FNN IV)	1hr	0.4930	0.7254	0.8161
				2hr	0.3467	0.7693	0.9296
				3hr	0.1907	0.8104	0.9681
				Mean	0.3435	0.7684	0.9046

표 4에서 알 수 있듯이 9개 사상을 대상으로 예측한 결과, FNN (Forecasting Neural Network) III 모형이 가장 좋은 통계결과를 보여 최적 모형으로 선정하였으며 1시간, 2시간, 3시간 순으로 좋은 예측 결과를 나타내었다.

선정된 강우 예측보완 신경망 모형인 INN IV 모형과 강우예측을 위해 선정된 FNN III 모형을 결합하여 결측지점으로 가정한 No.5 지점의 1시간~3시간까지 선행 예측을 하였다. 1997년 7월 1일과 1997년 8월 3일 그리고 1998년 8월 11일 발생한 집중호우를 대상자료로 이용하였으며, No.5 지점의 1997년 7월 1일 강우사상에 대한 예측은 2, 3시간 예측이 오히려 실측치에 근사했으며 1997년 8월 3일, 그리고 1998년 8월 11일 강우사상에 대해서는 1시간 예측이 근사한 결과를 보여주고 있다. 보완된 강우량 자료를 이용하여 시간예측을 한 결과를 그림 6 ~ 그림 9에 도시하였다.

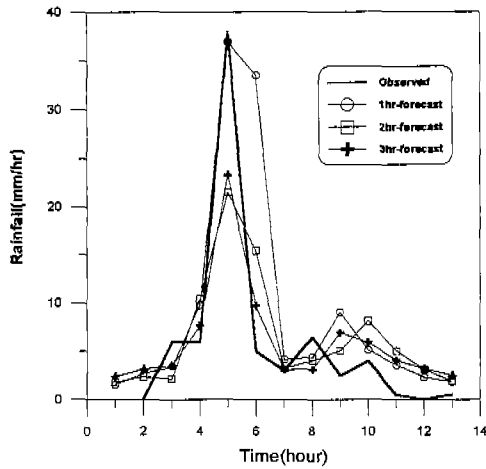


그림 6. 강우예측 시우량 곡선(1997.07.01)

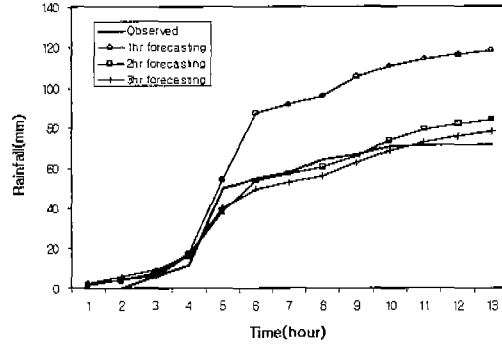


그림 7. 강우예측 누가우량 곡선(1997.07.01)

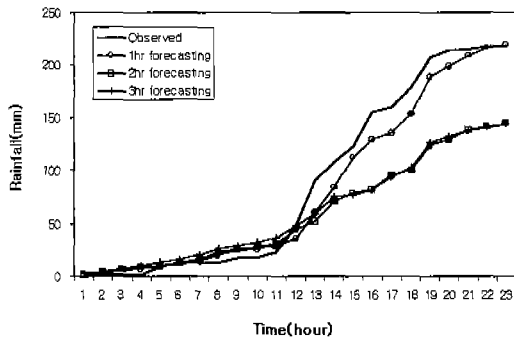


그림 8. 강우예측 누가우량 곡선(1997.08.03)

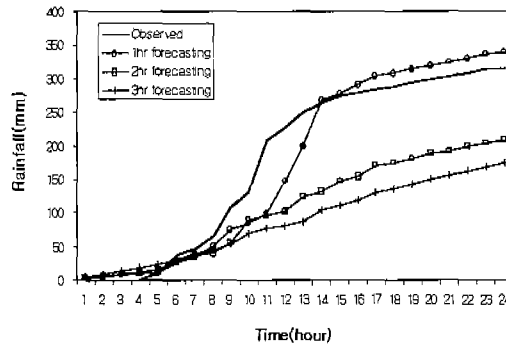


그림 9. 강우예측 누가우량 곡선(1998.08.11)

5. 결론

본 연구에서는 역전파 알고리즘을 이용한 다층 신경망 모형으로 강우 자료의 결측을 보완하고 보완된 자료를 이용하여 강우 결측지점의 시우량을 3시간까지 선형예측하도록 수행하여 다음과 같은 결론을 얻었다.

신경망 모형으로 강우 결측자료를 보완한 결과 오차가 적고 침투치에 근접하였으며 총강우량의 크기도 잘 산정하는 것으로 나타났으며 대체로 신경망법, RDS 방법, Thiessen 방법, 산술평균법, 정상년강우량법 순으로 실측치에 근사함을 보였다. 강우 결측보완과 강우 예측에 있어서 입력층의 뉴런이 N개일 경우 은닉층의 뉴런 수는 두 배인 2N개의 신경망 구조가 가장 좋은 결과를 보여주었으며 보완된 자료로 신경망 모형을 연계한 강우 예측은 강우의 진행 과정을 잘 파악하는 것을 알 수 있었으나 1시간 예측을 제외하곤 침투 강우량을 과소추정하는 경향을 보였다.

1. French, M.N., Krajewski, W.F., AND Cuykendall, R.R.(1992). "Rainfall forecasting in space and time using a neural network." J. of Hydrology, Vol. 137, pp. 1~31.
2. 오남선, "신경망이론과 Kalman Filter 강우예측모형의 비교연구", 서울대학교 박사학위논문, 1995.9.