

홍수위 예측을 위한 유전자 알고리즘과 신경망 이론의 결합

이한민¹⁾, 이연길²⁾, 박성천³⁾, 이관수⁴⁾

1. 서론

강우-유출관계 예측을 위한 모형화는 수문자료의 비선형성과 시공간적 다양성에 의한 많은 오차를 내포하고 있다. 특히, 개념적 모형은 실제 상황을 모형화하는 데 많은 어려움이 있을 뿐만 아니라 모형을 이용하기 위해 많은 노력과 시간이 투자되어야 한다. 그러나, 최근 신경망이나 카오스, 퍼지이론 등의 인공지능 모형이 수문학에 적용되면서 이러한 문제점을 해소하고자 하는 노력이 계속되어 왔다. 본 연구에서는 인간의 신경계를 수학적으로 모형화한 ANN(역전파 학습 알고리즘을 이용한 인공신경망) 방법과 여기에 중의 진화를 모방한 GA(유전자 알고리즘)와의 결합을 이용한 GANN 방법에 의해 하천에서의 실시간 홍수위를 예측하였다. 본 연구에서는 특히 목표지점의 수위를 예측하기 위해 기존의 연구들이 목표지점만의 강우자료와 수위자료를 입력자료로 취하였던 것에 비해, 영산강 수계 나주 수위관측소의 홍수위 예측을 목표로 상류 선암, 남평 그리고, 마죽지점과 유역내 평균강우를 입력자료로 이용하는 한편, 섬진강 유역 송정 수위관측소의 홍수위를 예측하는 데에는 강우자료를 제외한 상류 구례, 곡성, 즉곡 및 남원 수위관측소의 수위자료만을 입력자료로 활용하였다. 연구에 사용된 자료는 최고수위와 자료기간을 고려하여 훈련과정과 보정과정으로 분리하여 모의 실험을 실시하였으며, 훈련과정과 보정과정에 사용되지 않은 3개 사상에 대해 검증 실험을 실시하였다.

검증결과에 대한 통계분석을 통하여 선정된 나주 지점과 송정 지점에 대한 각각의 최적 모형으로 실시간 홍수위 예측을 실시한 결과, 나주 지점과 송정 지점 모두 GANN 모형이 ANN 모형보다 다소 높은 예측력을 나타내었다. 따라서 본 연구의 결과를 바탕으로 신경망과 유전자 알고리즘 기법을 적용한 실시간 홍수예보 프로그램 및 시스템 구축에 관한 지속적인 연구가 기대된다.

2. 대상 유역선정

영산강 유역은 섬진강 유역과 함께 호남지역에 위치하는 대표적인 하천유역으로서, 섬진강과 분수령을 이루는 노령산맥의 용추봉(EL. 570m)에서 발원하여 광주를 거쳐 목포에서 서해안으로 흘러 들어간다. 영산강 유역은 북위 34°40' ~ 35°29' 과 동경 126°27' ~ 127°05' 사이에 위치하고 있으며, 유역면적은 3,429km², 동서와 남북간의 최장거리는 각각 61.3km 및 89.7km이며, 유역의 평균 폭은 26.5km인 우리 나라 5대강에 속하는 하천이다. 영산강의 발원지인 용추봉에서 하구지점까지의 유로 연장은 129.5km이다. 수계 내에는 건설교통부 산하 17개소의 수위 관측소 및 14개소의 강우 관측소가 설치되어 있어 자료의 습득이 용이한 편이다.

섬진강 유역은 우리 나라 남부지역 중서부에 위치하고 있으며, 총 유역면적은 4,896.5km²이고, 본류의 유로연장은 212.3km이다. 유역의 최남단은 북위 34°30' 26", 최북단은 북위 35°50' 00"이며, 최동단은 동경 127°53' 05", 최서단은 동경 126°51' 50"로 본유역의 동쪽에는 낙동강 유역, 서쪽에는 영산강 유역과 동진강 유역, 북쪽에는 금강 유역과 만경강 유역이 접하고 있다. 총 유역면적 4,896.5km² 중 전라남도가 47%, 전라북도가 44%, 경상남도가 9%를 차지하고 있으며, 해당 행정구역은 3도 4시 11군 8읍 89면에 달하고 있다

1) 광주광역시 지하철 건설본부
2) 전남대학교 토목공학과 박사수료
3) 동신대학교 토목·환경공학부 부교수
4) 전남대학교 토목공학과 교수

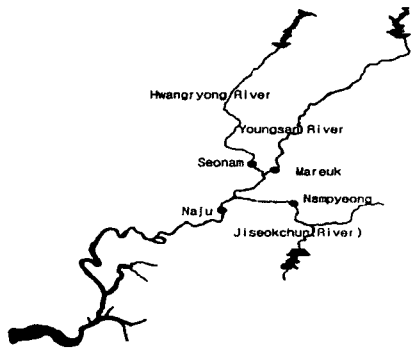


Fig. 2.1 Youngsan river basin



Fig. 2.2 Seomjin river basin

3. 모형의 적용

3.1 나주 지점의 예측모형 구성

나주 수위관측소 지점은 영산강 제1지류 중에서도 규모 면에서 1, 2위에 해당하는 황룡강과 지식천 합류점 하류에 위치하고 있어 영산강의 중하류부를 대표한다고 할 수 있다. 또한 나주 수위관측소 지점은 영산강 홍수통제소의 홍수예경보 지점으로서 그 중요도가 매우 높은 지점이다. 따라서 본 연구에서는 나주 지점에서의 실시간 홍수위 예측 모형을 구성하고자 하였다. 우선 지점별 특성에 맞는 최적모형을 구성하고자 입력층 노드와 은닉층 노드의 크기를 각각 아래와 같이 구성하였다. 즉, 입력층 노드 수(I)를 n이라면 은닉층 노드 수(J)는 $n \sim 2n$ 까지 변화시켰다.

Table 3.1 Number of processing element for flood forecasting models in Naju

Type	Model	Input layer(I)	Hidden layer(J)
I	1	7	7~14
	2	7	7~14
II	3	17	17~34
	4	17	17~34
	5	19	19~38
III	6	14	14~28
	7	16	16~32

Table 3.2 Number of processing element for flood forecasting models in Songjeong

Model	Input layer (I)	Hidden layer (J)
8	16	16~32
9	17	17~34
10	18	18~36
11	19	19~38

3.2 송정 지점의 예측모형 구성

송정 수위관측소 지점은 섬진강 제1지류이면서 동복댐, 주암댐, 주암조절지댐, 보성강댐 등 모두 4개의 댐이 위치하는 보성강의 합류점 아래에 위치하며, 하류부를 대표하는 하동 수위관측소와 함께 섬진강의 주요 홍수예경보 지점이다. 특히 섬진강은 우리나라의 5대강 중에서도 가장 경사가 급하고 하폭이 좁아 홍수위 예측이 어려운 지점으로 알려져 있다. 본 연구에서는 송정 지점에서의 실시간 홍수위 예측 모형을 구성하고자 하였다.

우선 지점별 특성에 맞는 최적모형을 구성하고자 입력층, 은닉층의 크기를 각각 아래와 같이 나주 지점에 이어 Model 8~11로 표현하였다. 즉, 입력층 노드 수(I)를 n이라면 은닉층 노드 수(J)는 $n \sim 2n$ 까지 변화시켰다. 따라서, Model 11의 경우 입력층 노드의 수가 19개이므로 은닉층 노드의 수를 입력층의 2배까지 설정하여 모의를 수행하므로 은닉층의 수를 최대 20개까지 설정하였다.

4. 실시간 홍수위 예측

4.1 예측방법

전절에서 선정된 지점별(나주, 송정), 예측방법별(ANN, GANN) 최적모형에 의한 실시간 홍수위 예측은 다음과 같이 두가지 방법으로 비교하여 실시하여 보았다. 첫째, 기존의 연구들이 제시해왔던 방법으로 입력층 노드수를 설정한 후 그에 따른 은닉층 노드수의 변화에 따른 예측결과를 비교한 후 결과가 가장 좋은 은닉층 수를 나타내는 case를 최적모형으로 채택하는 방법으로서 최적case의 weight와 bias를 이용하여 실시간 홍수위를 예측하는 방법이며, 한 지점에서 단 1개의 예측치가 존재한다. 둘째, 모형의 구성과정에서 통계적 기준에 의해 가장 좋은 결과를 나타내는 model을 선정한 후 model의 모든 은닉층 case에 대해 예측을 실시하는 방법으로서 여러 개의 예측값이 산출된다. 즉, 입력층 노드수가 n이라면 은닉층 수는 n~2n까지 변화하므로 결국 총 n+1개의 예측값이 존재하게 된다. 따라서, 본 연구에서는 n+1개의 예측치를 다시 통계처리하여 최종적으로 1개의 예측치를 도출하였다. 지점별, 예측방법별 실시간 예측모형을 Table 4.1에 정리하였다.

Table 4.1 Summary of real time forecasting model

Station	Forecasting Method	Model No.	I. L	H. L	Remarks
Naju	ANN	6	14	14~28	
	GANN	6	14	14~28	
Song-jeong	ANN	8	16	16~32	
	GANN	8	16	16~32	

Table 4.2 Summary of real time forecasting 2002 flood

Station	Method	Model	Peak W.L (m)		Error of		R ²
			Obs.	Sim.	Peak W.L (cm)	Ratio (%)	
Naju	ANN	6	7.26	7.1430	11.70	1.61	0.9923
	GANN	6	7.26	7.3573	9.73	1.34	0.9905
Song-jeong	ANN	8	9.19	9.5320	34.20	3.72	0.9875
	GANN	8	9.19	9.4659	27.59	3.00	0.9877
			10.22	10.4731	25.31	2.48	

4.2 예측 결과

2002년 8월 나주 지점과 송정 지점에서 발생한 홍수에 대해 예측을 실시한 결과 Table 4.2와 같이 나타났다. 송정 지점의 경우 단일 호우지만 침수가 2회 발생하였기에 나누어 분석하였다.

위 표에서 알 수 있듯이 나주 지점과 송정 지점에서 모두 매우 근소한 차이지만 GANN의 결과가 ANN의 결과보다 예측력이 좋은 걸로 나타나고 있다. 비록 그 차이가 크진 않지만 앞서도 언급되었듯이 초기치 부여 방법에서 ANN 방법은 단 하나의 값으로 출발하지만 GANN은 여러 개의 값을 부여함으로써 global minimum의 탐색확률이 더 높다는 면에서 GANN이 보다 효율적으로 판단된다.

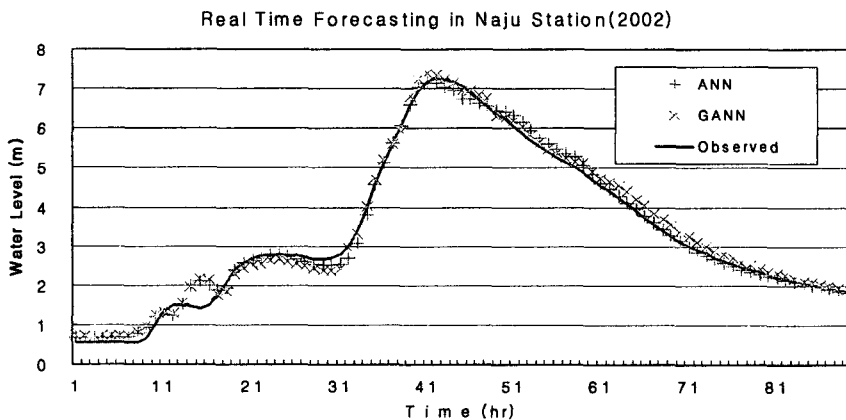


Fig. 4.1 Result of simulated W.L against observed W.L by ANN & GANN in Naju

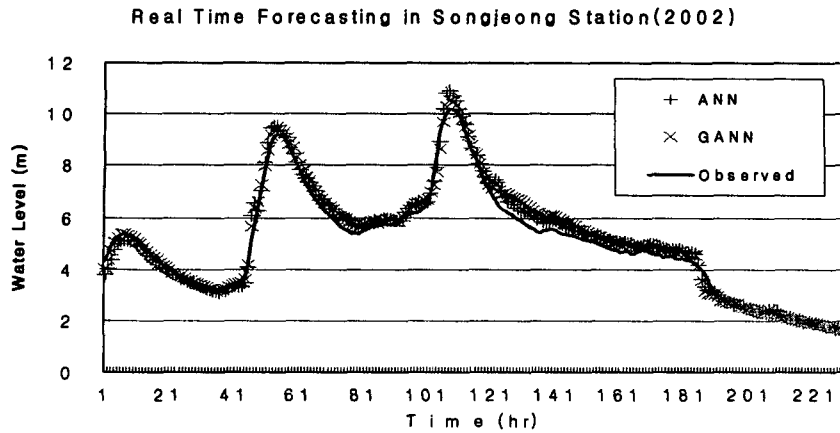


Fig. 4.2 Result of simulated W.L against observed W.L by ANN & GANN in Songjeong

5. 결론

1. 신경망 이론의 경우 난수발생을 통해 단 한 개의 초기치를 발생시키므로 local minimum에 빠질 위험이 큰 반면에, 유전자 알고리즘은 여러 개의 초기치를 발생시킨 후 다음 세대로의 진화과정을 통해 global minimum에 가장 근접한 개략적인 해를 신경망 과정의 입력자료로 전달함으로써 local minimum에 빠질 위험을 감소시키기 때문에 GANN 방법이 ANN 방법보다 정확한 예측이 가능하였다.
2. 나주 수위관측소에서의 수위예측 결과, 철두에서의 오차가 ANN Model 6이 11.70cm(1.61%), GANN Model 6이 9.73cm (1.34%)를 나타내었으며, 송정 수위관측소에서는 ANN Model 8이 각각 철두에서 34.20cm (3.72%), 49.10cm (4.80%), GANN Model 8이 각각 27.59cm(3.00%), 25.31cm(2.48%)의 오차를 나타냄으로써, 나주 및 송정 수위관측소 모두 GANN의 결과가 ANN의 결과보다 우수하게 나타났다.
3. 향후 예측의 정확도를 높이기 위해 신경망의 초기치 설정 문제와 유전자 알고리즘의 주요변수 설정방법, 그리고 신경망 이론의 단점인 훈련되지 않은 수위의 예측능력을 개발함으로써 이상호우 또는 국지성 호우의 경우에도 예측이 가능한 모형의 개발에 대한 연구가 계속되어야 할 것이다.

참고문헌

1. 영산강 홍수통제소, 1998. 영산강홍수예경보, 건설교통부.
2. 진영훈, 2000. "하천의 유출량 예측을 위한 인공신경망의 적용", 전남대학교 석사학위 논문
3. Alper, Kutay O., Bagdassarian Carey K. Coupled genetic algorithm/kohonen neural network(GANN) for projection of three-dimensional protein structures onto the plane, The college of William and Mary
4. Babovic, V., 1998. A Data Mining Approach to Time Series Modelling and Forecasting, Hydroinformatic'98, 847-856.
5. Brath, A., Burlando, P., Rosso, R., 1988. Sensitivity analysis of realtime flood forecasting to on-line rainfall predictions. In: Siccardi, F., Bras, R.L. (Eds.), Selected papers from the Workshop on Natural Disasters in European-Mediterranean Countries, Perugia, Italy, 469-488.
6. Campolo, M., Andreussi, P., Soldati, A., 1999. River flood forecasting with a neural network model. Water Resour. Res. 35 (4), 1191-1197.
7. Cook, D.F., Ragsdale, C.T., Major, R.L., 2000. Combining a neural network with a genetic algorithm for process parameter optimization, Engineering Applications of Artificial Intelligence 13, 391-396