

Data Mining 기법을 이용한 자동우량과 자동수위에 의한 하류 홍수예경보 기법

Automatic Rainfall and Waterlevel Downstream Flood Warning Techniques using Data Mining Techniques

최창진*, 이정훈**, 여운기***, 지홍기****

Chang Jin Choi, Jeong Hun Lee, Un Ki Yeo, Hong Kee Jee

요 지

최근 지구 온난화에 따른 이상 기후변화로 인해 계절성 집중호우와 같은 다양한 강우패턴이 발생되고 있다. 특히 집중호우의 빈도 및 규모가 커지고 있으며 피해 또한 증가하고 있다. 이에 대한 대안으로 하도의 정비, 댐 건설, 제방의 증고와 같은 구조적인 대책과 홍수예경보, 홍수보험, 통합홍수관리와 같은 비구조적인 대책에 대한 접근이 이루어지고 있다. 그러나 미래 기후변화에 대한 예측의 한계와 구조적 대책의 물리적 한계를 감안할 때 구조적 대책에 의한 방법만으로 변화하는 기후에 대응하여 홍수재해를 완벽하게 대처하기에는 부족한 것이 사실이다. 따라서 비구조적 대책에 의한 홍수피해저감이 절실히 필요하다.

따라서 본 연구에서는 국제수문개발계획 대표유역인 낙동강유역에 위치한 위천유역을 연구대상으로 선택하였고 이러한 중소규모의 유역에서 홍수예경보의 한계를 극복하고 신뢰성을 높이기 위하여 홍수유출시에 일어나는 유역내의 복잡한 물리적인 현상을 직접 고려하지 않고 입력자료와 출력자료의 관계로부터 학습과 추론을 통해 결론을 도출해내는 신경망, 퍼지, 유전자 알고리즘과 같은 Data Mining 기법을 사용하여 자동우량과 자동수위에 의한 하류 홍수예경보시스템을 구축하기 위해 수위를 예측하였다.

핵심용어 : 신경망, 퍼지, Data mining기법, 자동우량, 자동수위

1. 서론

국지성 집중호우로 인한 빈번한 돌발홍수 발생으로 인한 소하천 및 산지하천의 피해가 증가하는 추세에서 소하천 및 산지지역의 이용도가 높아짐에 따라 국지성 집중호우 및 돌발홍수로 인한 피해, 특히 인명손실 측면에서 기존 홍수 예경보와의 연계 또는 별도의 적절한 산지하천 돌발홍수 예측모델 및 예경보 시스템의 구축이 요구되고 있는 실정이다. 따라서 자동우량경보시스템의 경보발령기준에 대한 보완 및 개선, 산지하천에 적합한 돌발홍수 예측모델 및 예경보시스템의 개발 등으로 기존 시스템을 개량 및 강화할 필요가 있다. 또한 기존의 자동우량경보는 유역 내 소량의 강우국에 의존하여 경보를 발령해 왔으나 기상변화가 심한 산지 유역을 대표하기에는 무리가 있으므로 수위국 자료를 이용한 예경보 기준을 제시할 필요가 있다.

* 정회원 · 영남대학교 공과대학 건설시스템공학과 수자원환경전공 석사과정 · E-mail : bbacci@naver.com

** 정회원 · 영남대학교 공과대학 건설시스템공학과 수자원환경전공 석사과정 · E-mail : midas0821@naver.com

*** 정회원 · 영남대학교 공과대학 건설시스템공학과 수자원환경전공 박사 · E-mail : elofy@nate.com

**** 정회원 · 영남대학교 공과대학 건설시스템공학과 교수 · E-mail : hkjee@yu.ac.kr

2. 하류 홍수예경보의 계측센서 및 개념도

하류 홍수예경보를 위해서는 집중호우를 감지하고 수위국의 감시를 통해 짧은 시간내에 홍수의 발생을 예측하여 상황에 맞는 경보를 발령할 수 있어야한다. 따라서 원활한 예경보를 위해서는 그림 1과 같이 각 계측 센서들이 유기적으로 연결되어야 하며 이들을 적재적소에 배치하는 것이 중요하다.

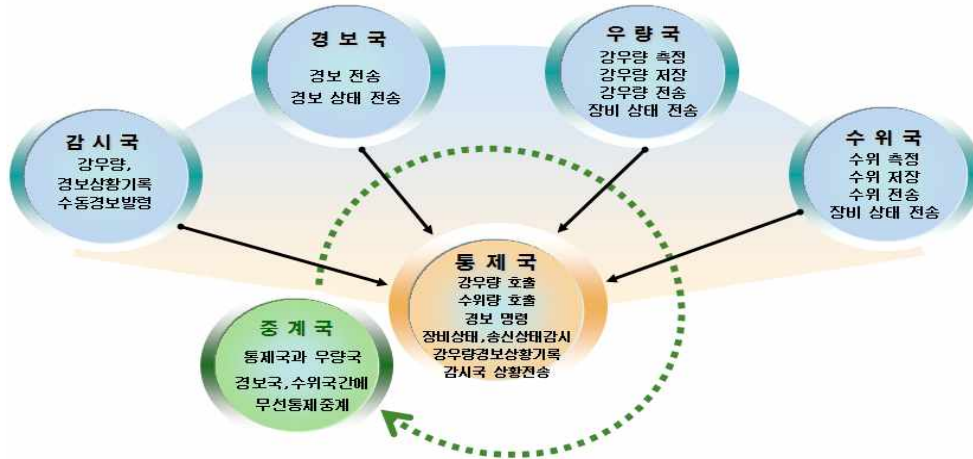


그림 1. 계측센서의 개념도

3. 데이터마이닝 기법의 적용지역

데이터마이닝 기법을 이용한 자동우량과 자동수위에 의한 하류 홍수예경보 기법의 검토를 위해 낙동강의 중류부에 위치한 감천유역을 대상으로 홍수예측을 실시하였다. 감천유역은 1998년 태풍 예니와 2002년 태풍 루사 그리고 2003년 태풍 매미 등에 의해 큰 피해를 입은 유역이며, 특히 감천 시내의 피해가 컸고 감천의 하류부에서의 침수피해도 큰 것으로 나타났다.

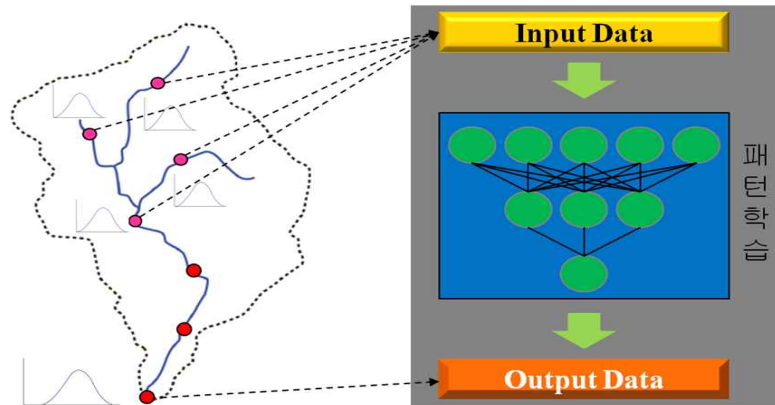


그림 2. 데이터마이닝 기법을 이용한 자동우량 자동수위에 의한 하류 홍수예경보 기법

감천 유역내에는 그림 2 및 표 1과 같이 수위관측소가 4개가 있다. 이 중 강창지점은 유역내에서 가장 먼저 관측을 시작하여 풍부한 자료를 가지고 있으나, 1987년에 폐쇄되어 분석에 사용하지 못하였으며, 동부 지점은 2010년부터 관측을 개시하여 자료의 부족으로 분석에서 제외하였다. 따라서 본 연구에서는 유역의 중앙부에 위치한 감천지점과 유역 하구부에 위치한 선산지점을 분석대상으로 하였다. 상류부에 위치한 감천 지점의 수위자료를 모형의 입력자료로 사용하였으며, 하류부에 위치한 선산지점의 수위자료를 모형의 출력 자료로 사용하여 학습과 검증을 실시하였다.

본 연구에서는 홍수예警보를 위한 수위를 예측하기 위한 모형들의 학습과 검정을 위한 자료를 선정하기 위해 현재 낙동강홍수통제소에서 홍수예警보의 발령에 사용하는 기준수위를 산정하는 방법을 사용하였다. 낙동강홍수통제소의 홍수주의보 발령기준은 ‘홍수예警보를 발령하는 지점의 수위가 계속 상승하여 주의보 경계홍수위(계획홍수량의 50%가 흐를 때의 수위)를 초과할 것이 예상되는 경우’이다. 이에 따라 모형의 출력 대상인 선산지점의 수위가 계획홍수량의 50%일 때의 수위를 초과하는 홍수사상들을 대상으로 하였다. 감천하천정비기본계획에서 제시된 선산지점의 계획홍수량은 3,000m³/sec이며, 계획홍수량의 50%는 1,500m³/sec가 된다. 이를 수위-유량 관계식을 사용하여 수위로 환산할 경우 3.7m이다. 따라서 1990년 이후의 홍수사상 중에서 이 조건에 부합하고, 관측이 누락되지 않은 자료가 양호한 호우사상들을 선정하여 분석을 실시하였다.

표 1. 감천유역의 수위관측소 현황

관측소명	관측소 코드	경도	위도	고도	비고
동부	2010660	128-11-36	36-09-46	57.972	2010년에 관측시작
강창	2008660	128-26-00	36-14-00	30.279	1987년에 관측종료
김천	2010650	128-07-58	36-06-58	61.085	
선산	2010690	128-18-40	36-13-29	32.524	

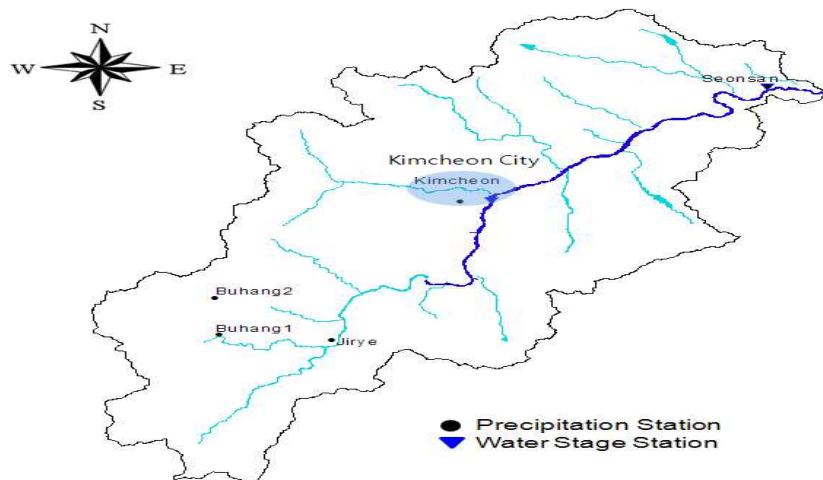


그림 3. 감천유역의 관측소 현황

4. 데이터마이닝 기법의 해석

본 연구에서는 수위예측을 위해 신경망모형(NN), 신경망과 유전자 알고리즘의 결합모형(GANN), 뉴로퍼지모형(NF), 뉴로퍼지와 유전자 알고리즘의 결합모형(GANF)을 구성하였다. 이 때 신경망 모형과 신경망과 유전자 알고리즘의 결합모형은 은닉층의 노드 수를 4개, 8개 12개로 구성함에 따라 각각 NN4, NN8, NN12, GANN4, GANN8, GANN12로 구분하였다. 이들 모형들을 비교하여 가장 정도 높은 모형을 선정하였다. 각 모형은 구조적인 차이에 따라 NN모형과 GANN모형의 구조를 같도록 하였으며, NF모형과 GANF모형의 구조를 같도록 하였다.

이 때 4모형의 학습과 검정에 사용되는 입력자료와 출력자료는 모두 같은 값을 사용하도록 구성하였다. 입력자료로는 교차상관분석을 통해 선산지점의 침투수위에 가장 높은 상관관계를 가지는 t-3, t-4, t-5, t-6의 지체시간을 가지는 김천지점의 수위를 사용하였으며, 출력자료는 선산지점의 예측시간별 수위를 사용하였다.

본 연구에서는 하류 하천의 수위를 예측하고자 하기 위해 몇 가지 인공지능기법을 사용하였다. 그 중, NN모형과 GANN모형의 경우 은닉층의 노드수를 결정할 필요가 있다. 본 연구에서는 은닉층의 노드수를 각각 입력자료의 수 n 에 비례하여 n 개, $2n$ 개 $3n$ 개인 경우를 상정하여 분석을 실시하였다.

분석에 사용된 자료는 2007년에 발생한 홍수사상을 대상으로 하였으며, 이 때 관측된 선산지점의 수위를 비교대상으로 하였다. 모형들의 예측이 실제 관측된 수문곡선을 얼마나 잘 예측하는지 살펴보기 위해 각 모형별로 선행예측시간마다 예측된 수문곡선을 실제 관측치와 비교하였다. 가장 적합한 은닉층의 노드수를 결정하기 위해 각 은닉층별로 분석된 값의 결정계수를 구하여 비교해 본 결과 가장 적합한 은닉층의 노드수는 12개인 것으로 나타났다.

또한, 앞에서 분석한 NN모형과 GANN모형과 달리 NF 모형과 GANF모형은 그 구조상 소속함수를 결정하는 것이 가장 중요하다. 일반적으로 가장 많이 사용되고 있는 소속함수들로는 중형 함수, 가우스 함수, 시그모이드 함수, 삼각형 함수, 사다리꼴 함수 등이 있다. 각 예측시간별로 적용하여 각각의 결정계수값을 비교해 본 결과 소속함수들 간에 각 지표값들의 차이는 크지 않으나 종합적으로 중형함수가 소속함수로 가장 적합한 것으로 나타났다.

앞에서 산정된 최적의 은닉층 노드수 12개를 적용한 NN12모형과 GANN12모형과 최적의 소속함수인 중형함수를 적용한 NF모형과 GANF모형을 이용하여 선산지점의 1 ~ 6시간 후의 수위를 예측한 경우 그림 4과 같은 결과를 얻을 수 있었다.

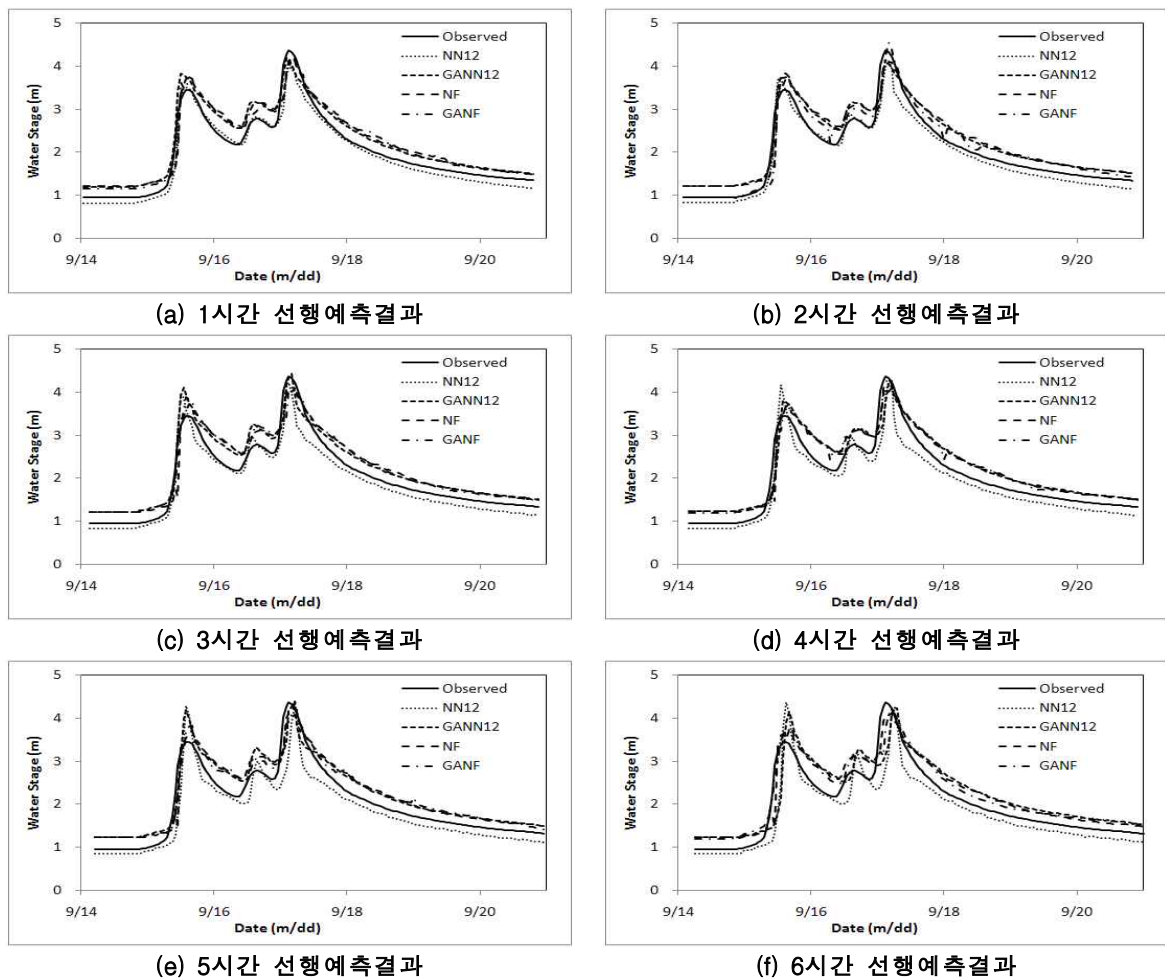


그림 4. 각 모형의 예측결과 비교

5. 결과

본 연구에서 사용된 데이터마이닝 기법들은 모두 예측력이 뛰어난 우수한 모형인 것으로 나타났으며, 어느 모형이 좋다고 말하기가 어려운 상황이다. NN12모형의 경우 예측시간이 길어짐에 따라 예측력이 떨어지는 문제가 있으며, NF모형은 예측력이나 오차의 측면에서 다른 모형에 비해 조금 떨어졌다. 유전자 알고리즘을 사용한 GANN12와 GANF 모형이 NN12모형과 NF모형에 비해 조금은 나은 결과를 보여주었다. GANN12와 GANF를 비교해 봤을 때, 다른 수치들은 비슷하게 나왔으므로 R2값이 크게 나온 GANN12모형이 예측에 감천유역에서 가장 적합한 것으로 판단된다.

감사의 글

본 연구는 소방방재청에서 주관하고 있는 자연재해저감기술개발사업단 연구개발사업의 일환으로 수행된 「지능형 돌발홍수 예경보시스템 개발(NEMA-자연-2011-48)」의 연구성과입니다.

참고 문헌

1. 지홍기, 여운기, 최현일(2009), 신경망을 이용한 단기강우예측, 대한토목학회 학술발표회
2. 지홍기, 김응석, 최현일 (2010), 미계측 유역에서의 확률강우에 대한 돌발홍수지수 산정, 한국방재학회지, 제10권 제4호, pp.81~88
3. 지홍기, 여운기, 서영민, 이승윤 (2010), 인공신경망과 유전자알고리즘의 결합모형을 이용한 수위예측에 관한 연구, 한국수자원학회논문집, 제43권 제8 호 721-731
4. 지홍기, 서영민, 여운기, 이승윤 (2010), 확률강우량의 공간분포 추정을 위한 KED 기법의 적용, 한국수자원학회논문집, 제43권 제8호, pp. 757-767
5. 지홍기, 여운기, 이순탁(2010), Neuro-Fuzzy 모형을 이용한 수위예측에 관한 연구, 대한토목학회 학술발표회
6. 지홍기, 서영민, 여운기(2011) 수위예측을 위한 데이터마이닝 기법의 비교, 대한토목학회 2011년 학술발표회
7. 지홍기, 여운기, 서영민(2011) 강우공간분포추정에서 모델 매개변수의 불확실성 해석, 대한토목학회 2011년 학술발표회
8. 지홍기, 이준호(2012) 신경망을 이용한 중소하천의 홍수예경보 기법 한국방재학회 2012년 학술발표회
9. Guidance on Flash Flood Management - Recent Experiences from Central and Eastern Europe, 2007, Associated Programme on Flood Management
10. Guidelines on Flash Flood Prevention and Mitigation, 2002, European Commission Joint Research Centre, Italy
11. Flash Flood Early Warning System Reference Guide, 2010, NOAA National Weather Service, USA