

하천유역에서 기후변화에 따른 이상호우시의 최적 수문예측시스템

The Optimal Hydrologic Forecasting System for Abnormal Storm due to Climate Change in the River Basin

김성원* · 김형수**

Abstract

In this study, the new methodology such as support vector machines neural networks model (SVM-NNM) using the statistical learning theory is introduced to forecast flood stage in Nakdong river, Republic of Korea. The SVM-NNM in hydrologic time series forecasting is relatively new, and it is more problematic in comparison with classification. And, the multilayer perceptron neural networks model (MLP-NNM) is introduced as the reference neural networks model to compare the performance of SVM-NNM. And, for the performances of the neural networks models, they are composed of training, cross validation, and testing data, respectively. From this research, we evaluate the impact of the SVM-NNM and the MLP-NNM for the forecasting of the hydrologic time series in Nakdong river. Furthermore, we can suggest the new methodology to forecast the flood stage and construct the optimal forecasting system in Nakdong river, Republic of Korea.

Key words : Optimal forecasting system, Abnormal storm, SVM-NNM, MLP-NNM

1. 서론

한반도 대운하 건설정책은 한강과 낙동강 540km를 연결하며, 경부운하를 중심으로 하여 호남 대운하, 금강 운하 등 3개축으로 추진하며 한반도를 하나의 물길로 이어가면서 생활경제권을 형성하는 정책이다. 이 정책은 현 이명박 대통령의 핵심 공약으로 가장 주목받는 정책 중 하나이며, 운하기간과 공사비산정, 수질오염개선과 환경파괴 및 국내여건에서의 적합성여부 등으로 주요 논란의 핵심이 되고 있다. 특히 기술적인 측면에서는 운하건설시 치수, 이수 및 환경적인 측면에서 많은 논란이 되고 있다. 낙동강에서의 치수적인 측면을 고려하면 낙동강 대운하 건설이전에 기존의 자료를 이용하여 최적 홍수예측시스템의 구성이 낙동강 대운하 건설에 필요한 신뢰성 있는 수문자료를 제공하는 것이 가능할 것이다. 따라서 하천 홍수량 및 홍수위의 정확한 예측은 수자원의 계획과 운영에 있어서 주요한 입력정보를 형성한다. 특히 주어진 수문년에서의 연속적인 하천홍수위의 측정을 기준으로 이루어진다면 다음 두 가지의 측면에서 도움이 될 것이다. 첫 번째는 홍수시의 임박한 홍수의 예경보에 대한 실시간 정보를 제공할 수 있으며, 두 번째는 홍수시에 저수지의 방류량을 제어하는데 도움이 될 것이다 (Kim and Kim, 2008a).

본 연구에서는 낙동강유역의 주요 수위표지점인 진동지점에서 루사 (Rusa) 혹은 매미 (Maemi)와 같은 이상호우의 발생시에 최적 홍수위예측시스템을 구성하기 위하여 지지벡터기구 신경망모형 (Support Vector Machines Neural Networks Model, SVM-NNM) 과 다층퍼셉트론 신경망모형 (MultiLayer Perceptron Neural Networks Model, MLP-NNM)을 제시하였으며, 입력층 노드로는 고령교 및 적포교 지점의 하천홍수위 자료를 이용하였다. 또한 제시한 모형의 운영기법에 대한 훈련, 교차검증 및 테스트과정을 통하여 최적 매개변수의 선정 및 모형의 적합성을 검토하였으며, 본 연구결과를 통하여 최적 홍수위 예측시스템 구축에 이용할 수 있을 것으로 판단된다.

2. 신경망모형과 유전자 알고리즘

2.1 지지벡터기구 신경망모형 (Support Vector Machines Neural Networks Model, SVM-NNM)

지지벡터기구 신경망모형 (SVM-NNM)은 다음 두 가지 개념에 의하여 운영되는 새로운 개념의 분석기법이다. 첫 번째는 고차원공간으로의 자료의 전환이며, 이것은 복잡한 결정 평면들을 가진 복잡한 문제를 선형 판별식의 함수로 사용할 수 있는 단순한 문제로 전이 할 수 있다. 두 번째는 SVM-NNM의 경우 분류에 대한 좀 더 많은 정보를 제공할 수 있기 때문에 결정 평면 근처에 있는 단지 그들 입력을 이용하고, 훈련의 개념에 의하여 동기를 가지게 된다. SVM-NNM 적용에 있어서 첫 번째의 단계는 고차원공간으로 전이하는 것이다. 일반적으로 이것은 각 자료에 가우시안을 설정하는 반경

* 정희원-공학박사수자원개발기술사, 동양대학교 철도토목학과 조교수 ; E-mail : swkim68@phenix.dyu.ac.kr

** 정희원-공학박사인하대학교 환경토목공학부 부교수 ; E-mail : sookim@inha.ac.kr

기초함수 (Radial Basis Function, RBF)를 사용하여 이루어진다. 그리하여 특성공간 (Feature Space)은 자료의 수만큼 증가하게 된다. 그러나 일반적으로 반경기초함수는 마지막 결과치를 계산하기 위하여 가우시안의 선형조합을 훈련시키기 위하여 역전파 훈련알고리즘 (BackPropagation Training Algorithm)을 이용한다. 그러나 본 연구에서 SVM-NNM은 훈련과정을 위하여 대규모의 극한 분류 (Margin Classifier)의 개념을 사용한다. 이것은 입력공간으로 부터의 분류의 용량을 분리하고 동시에 양호한 일반화를 제공한다. 이것은 분류를 위한 이상적인 조합이다. 훈련알고리즘은 반경기초함수에 확장된 아다트론 알고리즘 (Adatron Algorithm)을 기초로 하고 있다. 아다트론 알고리즘은 핵함수 (Kernel Function)에 의하여 내부공간에서 패턴의 내적을 교체함으로써 반경기초함수로서 쉽게 확장되어 질 수 있다 (Vapnik, 1992, 1995; Principe et al., 2000). 본 연구에서는 단계간격 (Step Size)은 0.001, 최대 반복횟수는 10,000회로 설정하였다. 다음 그림 1은 본 연구에 적용된 SVM-NNM 구조를 나타낸 것이다.

2.2 다층퍼셉트론 신경망모형 (Multilayer Perceptron Neural Networks Model, MLP-NNM)

다층 퍼셉트론 신경망모형 (MLP-NNM)은 SVM-NNM의 수행능력을 비교하기 위한 신경망모형으로 선택되었다. 일반적으로 MLP-NNM은 입력층, 출력층 및 그 사이에 1개 혹은 그 이상의 은닉층으로 구성되어 있으며, 각 층에서의 노드는 그 다음 층의 노드와 상호 긴밀하게 연결되어 있다. 또한 본 연구에서 MLP-NNM을 적용하는 과정은 훈련, 교차검증 및 테스트과정으로 구성되어 있다. 일반적으로 훈련 및 교차검증과정을 통하여 MLP-NNM의 최적 매개변수를 결정하며, 결정된 매개변수를 이용하여 MLP-NNM의 테스트과정을 수행한다 (Haykin 2003). 본 연구에서는 훈련알고리즘을 레벤버그-마쿼트 (Levenberg-Marquardt) 알고리즘을 이용하였으며, 최대 반복횟수는 10,000회로 설정하였다.

2.3 유전자 알고리즘 (Genetic Algorithm, GA)

본 연구에서 유전자 알고리즘 (GA)은 첫 번째는 훈련자료 및 교차검증자료를 이용하여 SVM-NNM 및 MLP-NNM 을 훈련시키는 것이며, 두 번째는 첫 번째 과정에서 형성된 SVM-NNM 및 MLP-NNM을 이용하여 테스트자료에 대한 최적의 성능을 위하여 전 범위를 검토하는 것이다. GA는 각 입력층 노드에 대한 평활인자 승수 (Multiplier)를 찾고 있으며, 비록 훈련과정에 대하여 탐색기법으로 GA를 사용하지 않은 경우보다 시간은 좀 더 걸리지만 테스트자료에 대하여 가장 양호하게 운영될 수 있는 SVM-NNM 및 MLP-NNM을 구성할 수 있다. 본 연구에서는 GA의 적용을 위하여 세대수 (Generation)는 100, 염색체수 (Chromosome) 50 및 최대 반복횟수는 10,000회로 설정하였다.

3. 신경망모형의 구성 및 적용

3.1 분석대상지점 및 수문자료의 선정

본 연구에서의 신경망모형을 적용하기 위한 주요 수위표지점은 낙동강유역의 분류에 위치한 진동 수위표지점으로 선정하였으며, 입력자료로는 고령교 및 적포교 수위표지점의 홍수위자료를 이용하였다. 따라서 각 수위표지점의 홍수위사상은 기왕의 자료중에서 자료의 신뢰성이 높은 4개의 시간 홍수위자료를 선택하였다. 또한 홍수위사상 1-2까지를 신경망모형의 훈련과정, 홍수위사상 3을 신경망모형의 교차검증과정 및 홍수위사상 4를 신경망모형의 테스트과정에 사용하였다. 본 연구에 사용된 자료는 국가 수자원관리종합정보 시스템 (Water management information system, WAMIS)의 홈페이지 (www.wamis.go.kr)의 무료 공개자료를 이용하였다. 다음 표 1은 본 연구에 사용된 홍수위자료를 나타내고 있다.

3.2 훈련과정 (Training Performance)

본 연구의 훈련과정에서는 SVM-NNM 와 MLP-NNM의 훈련에 따른 선행과정으로서 입력층의 입력변수인 고령교와 적포교의 홍수위자료를 표준화하여 적용하였다. 일반적으로 입력변수들을 표준화하는 주요한 이유 중 하나는 각 입력변수들이 서로 다른 단위로서 관측된 값들을 나타내고 있기 때문에 이러한 입력변수들을 표준화하고 무차원단위에서 입력변수들을 재배치함으로써, 각 변수들의 상사성의 효과를 제거하기 위한 것으로 판단된다. 따라서 본 연구에서 훈련과정에 사용된 자료 수는 480개로 구성하였다.

3.3 교차검증과정 (Cross-Validation Performance)

본 연구의 교차검증과정은 SVM-NNM와 MLP-NNM가 최소 RMSE에 도달할 때까지 모든 훈련자료를 훈련시키게 하는 것이 아니고, 각 훈련단계의 마지막에 교차검증자료로서 교차검증하는 방법이다. 만약 과적합문제가 발생되면 테스트자료의 평균제곱오차의 수렴과정은 훈련자료가 여전히 훈련되고 있는 경우에도 감소하지 않고 증가하는 경향을 나타낼 것이다 (Bishop, 1994; Kim and Kim, 2008b). 따라서 본 연구에서 교차검증과정에 사용된 자료 수는 240개로 구성하였다.

3.4 테스트과정 (Test Performance)

본 연구의 테스트과정은 훈련 및 교차검증과정을 통하여 산정된 매개변수를 이용하여 테스트자료를 검증하였다. 따라서 본 연구에서 테스트과정에 사용된 자료 수는 240개로 구성하였다. 다음 그림 2는 낙동강 진동 수위표지점에서 홍수위 예측의 테스트결과에 대한 비교를 나타낸 것이다. 또한 다음 표 2는 낙동강 진동 수위표지점에서 테스트자료에 대한

SVM-NNM과 MLP-NNM의 통계분석 결과를 나타내고 있다. 그림 2와 표 2를 통하여 SVM-NNM의 결과치가 MLP-NNM의 결과치보다 미소하게 양호한 것으로 나타났다.

4. 결 론

본 연구에서는 낙동강 하천유역에서의 홍수위의 예측과정을 통한 최적 홍수위예측시스템 구성을 위하여 통계학적 학습 이론을 기초로 한 신경망모형인 SVM-NNM 및 보편적인 신경망모형인 MLP-NNM을 이용하였다. 따라서 낙동강 진동 수위표지점에서 홍수위예측을 위하여 고령교 및 적포교 수위표지점의 홍수위자료를 입력자료로 선정하였다. 그 결과 SVM-NNM의 결과치가 MLP-NNM의 결과치보다 미소하게 양호한 것으로 분석되었다. 그러나 본 연구에서는 홍수위예측에 대한 시간적인 분포에 해당하는 선행시간의 고려가 전혀 없었으며, 입력자료의 홍수위지점도 2개 지점만 선택하였으며 홍수위 사상도 선택된 몇 개만 이용하였다. 따라서 언급한 사항에 대하여 모두 고려할 경우 본 연구에서 분석된 결과와 다른 결과를 유도할 수도 있을 것이다. 하지만 본 연구에서 제시된 SVM-NNM은 홍수위예측의 새로운 방법론으로 제시할 수 있을 것으로 판단되며, 이를 기초로 하여 최적 수문예측시스템을 구성할 수 있을 것으로 판단된다. 또한 낙동강에서의 이러한 최적 홍수위예측시스템은 궁극적으로 낙동강과 한강을 연결하는 경부운하의 구조적인 설계 및 계획홍수위의 선정에 있어서 주요한 기초 자료를 제공할 수 있으리라 판단되며, 경부운하의 신뢰성을 확보하는데 도움이 될 것이라 판단된다.

참고문헌

1. Bishop, C.M. (1994). "Neural networks and their applications." *Rev. Scien. Instru.* Vol. 65, pp. 1803-1832.
2. Haykin, S. (2003). *Neural networks : A comprehensive foundation*, 4th Indian Reprint, Pearson Education, Singapore.
3. Kim, S., and Kim, H.S. (2008a). "Uncertainty reduction of the flood stage forecasting using neural networks model." *J. Americ. Water Resor. Associ.*, Vol. 44, No. 1, pp. 148-165.
4. Kim, S., and Kim, H.S. (2008b). "Neural networks and genetic algorithm approach for nonlinear evaporation and evapotranspiration modeling." *J. Hydro.*, Vol.351, pp. 299-317.
5. Vapnik, V.N. (1992). "Principle of risk minimization for learning theory" , *Adv. Neural Inf. Processing Syst.* Vol. 4, pp. 831-838.
6. Vapnik, V.N. (1995). *The Nature of statistical learning theory*, Springer Verlag, NY.
7. Principe, J.C., Euliano, N.R., and Lefebvre, W.C. (2000). *Neural and adaptive systems; fundamentals through simulation*, John Wiley & Sons, Inc., NY.

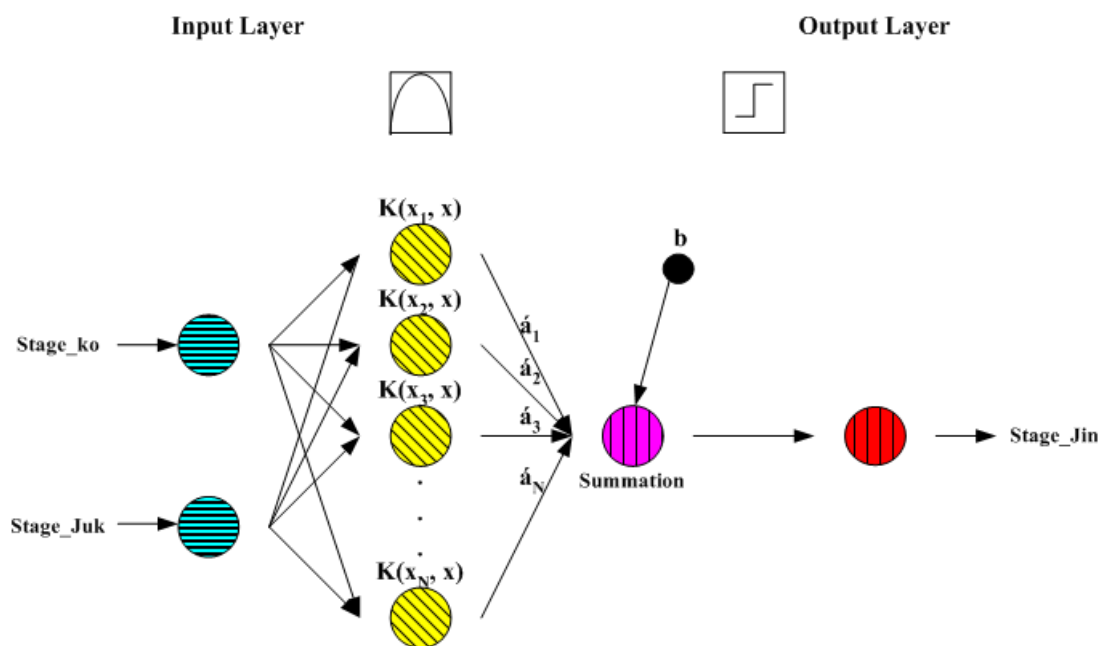


그림 1. SVM-NNM 의 구조

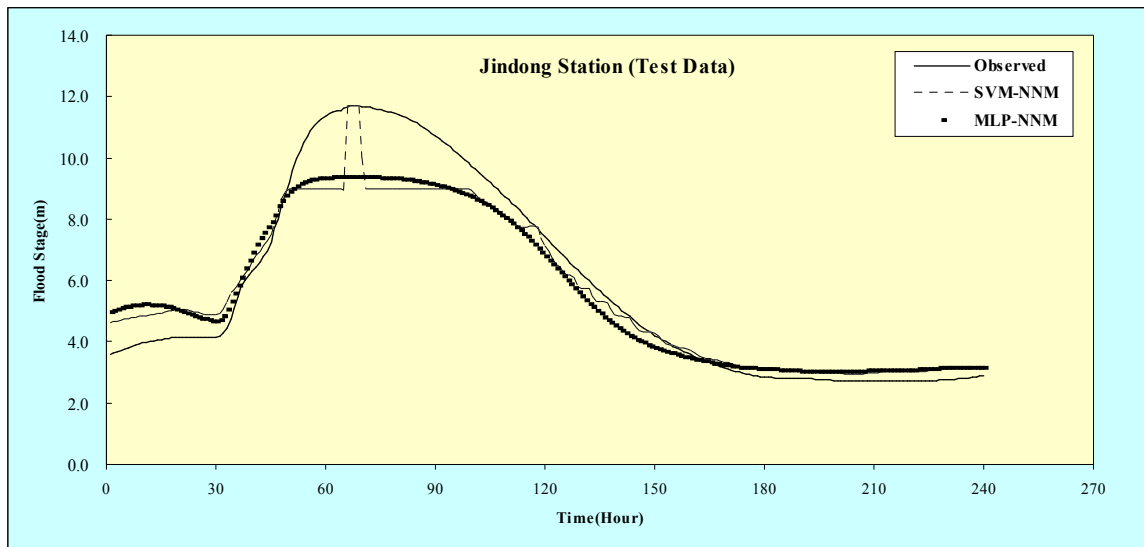


그림 2 홍수위의 비교 (진동지점, 검증자료)

표 1. 본 연구에 사용된 홍수위자료

지점	일시	총 시간 (Hr)	시간최대수위 (m)	비고
고령교	1991.08.22 - 08.31	240	8.35	
	1993.07.11 - 07.20	240	8.21	
	1993.08.07 - 08.16	240	9.05	
	2003.09.11 - 09.20	240	12.07	태풍 매미
적포교	1991.08.22 - 08.31	240	8.34	
	1993.07.11 - 07.20	240	7.90	
	1993.08.07 - 08.16	240	8.97	
	2003.09.11 - 09.20	240	11.82	태풍 매미
진 동	1991.08.22 - 08.31	240	8.78	
	1993.07.11 - 07.20	240	8.01	
	1993.08.07 - 08.16	240	9.01	
	2003.09.11 - 09.20	240	11.70	태풍 매미

표 2. 테스트자료를 이용한 SVM-NNM과 MLP-NNM 통계분석 결과 (진동지점)

Station	Statistical Index	Models	
		SVM-NNM	MLP-NNM
Jindong	CC	0.977	0.972
	RMSE(m)	0.991	1.007
	E	0.904	0.900
	MAE(m)	0.037	0.786