

# 비선형 분리모형에 의한 증발접시 증발량의 해석

## Pan Evaporation Analysis using Nonlinear Disaggregation Model

김성원\* · 김정현\*\* · 박기범\*\*\*

### Abstract

The goal of this research is to apply the neural networks models for the disaggregation of the pan evaporation (PE) data, Republic of Korea. The neural networks models consist of the support vector machines neural networks model (SVM-NNM) and multilayer perceptron neural networks model (MLP-NNM), respectively. The SVM-NNM in time series modeling is relatively new and it is more problematic in comparison with classifications. In this study, The disaggregation means that the yearly PE data divides into the monthly PE data. 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 disaggregation of the nonlinear time series data. We should, furthermore, construct the credible data of the monthly PE data from the disaggregation of the yearly PE data, and can suggest the methodology for the irrigation and drainage networks system.

*Key words* : Pan Evaporation, Disaggregation, SVM-NNM, MLP-NNM.

### 1. 서론

대기와 육지는 상호 연속적인 물분자의 교환이 있는데도 불구하고 일반적으로 증발 및 증발산의 수문학적 정의는 대기로의 물이송의 순 비율에 한정되어 있다. 증발과 증발산은 제시된 지역에서 습윤효율 (Moisture efficiency)의 지표적인 변화이고, 그 크기는 하천에서 홍수유출량을 추정하는데 사용되기도 한다. 증발량은 수문학적 순환 중 하나이며, 일반적으로 간접적인 측정을 기초로 하는 질량 이송법, 에너지 수지법 및 물수지법과 같은 여러 방법으로 추정될 수 있다. 직접적인 측정기법 중 하나인 증발접시 증발량은 증발산량의 지표로 사용되고 있으며, 호수 및 저수지의 증발량을 추정하는데 가장 광범위하게 사용되고 있다 (Hargreaves, 1966; Kohler et al., 1955). 식물표면으로부터 증발산량의 추정은 물수지를 계산하고 용수수요와 공급을 추정하는데 있어서 기본적인 도구이다. 세계인구의 절반이상이 관개농업에 의한 생산에 의존하고 있으므로, 농작물 증발산량의 정량화는 매우 중요하다. 증발산량은 증발산계 (Lysimeter)를 이용하여 관측하거나 물수지 기법 혹은 기상변수를 이용하여 추정한다. 그러나 증발산계를 이용한 증발산량의 측정은 필요없이 많은 시간이 요구되는 작업이고, 정확하고 주의깊은 경험이 필요하기 때문에 항상 가능한 방법은 아니다. 따라서 기상자료를 기초로 하는 경험적인 접근법이 증발산량을 추정하는데 일반적으로 사용된다 (Allen et al., 1989; Montheith, 1965; Penman, 1948).

본 연구의 목적은 우리나라 주요 기상관측소에서 관측된 연간 소형 증발접시 증발량자료를 제시된 두가지의 신경망모형을 이용하여 월별 소형 증발접시 증발량자료로 분리하고자 한다. 따라서 연간 소형 증발접시 증발량 관측자료가 충분히 큰 기상관측소를 연구대상으로 선정하는 것이 반드시 필요하다. 또한 선정된 기상관측소의 증발접시 증발량자료를 훈련자료, 교차검증자료 및 테스트자료로 분리하여 모형을 실시하는 기법을 개발하고자 한다. 따라서 본 연구를 통하여 연간 소형 증발접시 증발량자료로부터 신뢰성있는 월별 소형 증발접시 증발량을 산정할 수 있는 방법론의 유도가 가능하다고 판단된다.

### 2. 신경망모형과 유전자 알고리즘

#### 2.1 지지벡터기구 신경망모형 (Support Vector Machine 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 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 위하

\* 정회원, 공학박사, 수자원개발기술사, 동양대학교 철도토목학과 조교수 ; E-mail : [swkim68@phenix.dyu.ac.kr](mailto:swkim68@phenix.dyu.ac.kr)

\*\* 비회원, 동양대학교 철도토목학과 4학년 ; E-mail : [idiunu@hanmail.net](mailto:idiunu@hanmail.net)

\*\*\* 정회원, 공학박사, 안동과대학 건설정보과 겸임교수 ; E-mail : [pkb5032@naver.com](mailto:pkb5032@naver.com)

여 역전과 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 최적운 영을 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. 연구범위 및 자료

본 연구에서는 제주도를 포함한 기상청의 관리하에 있는 71개의 기상관측소 중에서 장기간의 관측기록에 해당되는 90 년이 넘어가는 관측소인 인천, 부산, 목포, 강릉, 대구, 전주 및 서울관측소 중에서 자료의 결측이 거의 없는 기상관측소 인 서울 및 강릉관측소를 선정하였다. 서울관측소의 경우 1907년 10월에 기상관측을 실시한 이후 1919년, 1950-1953년 및 1991-1994년의 자료가 결측이 되어 있으나, 자료의 신뢰성이 아주 양호한 상태이다. 강릉관측소의 경우 1912년 1월에 기 상관측을 실시한 이후에 1919년, 1950-1951년 자료가 결측이 되어 있으나, 서울관측소의 경우처럼 자료의 신뢰성이 아주 양호한 상태이다.

## 4. SVM-NNM 및 MLP-NNM의 적용

### 4.1 훈련과정 (Training Performance)

본 연구의 훈련과정에서 서울관측소의 경우는 1908년부터 2007년까지 실측된 자료중에서 훈련자료는 1908년부터 1983 년까지의 결측기간을 제외한 자료를 선택하였으며, 강릉관측소의 경우는 1912년부터 1987년까지의 결측기간을 제외한 자 료를 선택하였다. 따라서 서울 및 강릉관측소에서 훈련과정에 사용된 자료는 각각 71년 및 73년의 자료시계열로 구성이 되어 있다. 그리고 SVM-NNM 와 MLP-NNM의 훈련에 따른 선행과정으로서 입력층의 입력변수인 연간 소형 증발접시 증 발량자료를 표준화하여 적용하였다. 일반적으로 입력변수들을 표준화하는 주요한 이유 중 하나는 각 입력변수들이 서로 다른 단위로서 관측된 값들을 나타내고 있기 때문에 이러한 입력변수들을 표준화하고 무차원단위에서 입력변수들을 재배 치함으로써, 각 변수들의 상사성의 효과를 제거하기 위한 것으로 판단된다 (Kim and Kim, 2008a).

### 4.2 교차검증과정 (Cross-Validation Performance)

본 연구의 교차검증과정에서 서울관측소의 경우는 1984년에서 1997년까지의 결측기간을 제외한 자료를 선택하였으며, 강릉관측소의 경우는 1988년부터 1997년까지의 결측기간을 제외한 자료를 선택하였다. 서울 및 강릉관측소에서 교차검증 과정에 사용된 자료는 각각 10년의 자료시계열로 구성이 되어 있다. 교차검증과정은 SVM-NNM와 MLP-NNM가 최소 RMSE에 도달할 때까지 모든 훈련자료를 훈련시키게 하는 것이 아니고, 각 훈련단계의 마지막에 교차검증자료로서 교차 검증하는 방법이다. 만약 과적합문제가 발생되면 테스트자료의 평균제곱오차의 수렴과정은 훈련자료가 여전히 훈련되고 있는 경우에도 감소하지 않고 증가하는 경향을 나타낼 것이다 (Bishop, 1994; Kim and Kim,2008b).

### 4.3 테스트과정 (Test Performance)

본 연구의 테스트과정에서 서울관측소의 경우는 1998년에서 2007년까지의 결측기간을 제외한 자료를 선택하였으며, 강

릉관측소의경우는 1998년부터 2007년까지의 결측기간을 제외한 자료를 선택하였다. 서울 및 강릉관측소에서 테스트과정에 사용된 자료는 각각 10년의 자료시계열로 구성되어 있다. 테스트과정은 훈련 및 교차검증과정을 통하여 선정된 매개변수를 이용하여 테스트자료를 검증하였다. 다음 그림 2의 (a)-(b)는 서울 및 강릉관측소의 테스트자료를 이용하여 SVM-NNM과 MLP-NNM의 결과치에 대한 월별 소형증발접시 증발량의 시계열을 나타낸 것이다. 또한 다음 표 1은 서울 및 강릉관측소에 대한 테스트자료에 대한 SVM-NNM과 MLP-NNM의 통계분석 결과를 나타내고 있다.

## 5. 결 론

본 연구에서는 SVM-NNM과 MLP-NNM을 이용하여 연간 소형 증발접시 증발량자료를 월별 증발접시 증발량자료로의 분리를 시도하였다. 사용된 자료는 서울 및 강릉관측소의 자료를 이용하였으며, SVM-NNM의 결과치가 MLP-NNM의 결과치보다 양호한 것으로 분석되었다. 따라서 본 연구와 병행하여 부족한 자료시계열의 경우는 추계학적 모형을 이용하여 모의된 자료를 훈련 및 교차검증자료로 이용하는 기법 및 추계학적 분리모형을 적용하여 비교분석 하는 기법등에 대한 지속적인 연구가 필요한 것으로 판단된다. 또한 본 연구를 통하여 우리나라 관개배수 시스템 구성을 위한 월별 소형 증발접시증발량 기준자료를 제공할 수 있는 방법론을 제시할 수 있을 것으로 판단된다.

## 참고문헌

1. Allen, R.G., Jensen, M.E., Wright, J.L., and Burman, R.D. (1989). "Operational estimates of reference evapotranspiration." *Agrono. J.*, Vol. 81, No. 4, pp. 650-662.
2. Bishop, C.M. (1994). "Neural networks and their applications." *Rev. Scien. Instru.* Vol. 65, pp. 1803-1832.
3. Hargreaves, G.H. (1966). "Consumptive use computations from evaporation pan data." , *In Irrigation and Drainage Special Conference*, ASCE, Las Vegas, NV, pp. 35-62.
4. Haykin, S. (2003). *Neural networks : A comprehensive foundation*, 4th Indian Reprint, Pearson Education, Singapore.
5. 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.
6. 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.
7. Kohler, M.A., Nordenson, T.J., and Fox, W.E. (1955). *Evaporation from pans on lakes*, US Department of Commerce, Weather Bureau Research Paper 38, Washington, DC.
8. Monteith, J.L. (1965). "The state and movement of water in living organism." *Proc., Evaporation and Environment, XIXth Symp.*, soc. For Exp. Biol., Swansea, Cambridge Univ. Press, NY, pp. 205-234.
9. Penman, H.L. (1948). "Natural evaporation from open water, bare soil and grass." *Proc. R. Soc. London*, 193, pp. 120-146.
10. Vapnik, V.N. (1992). "Principle of risk mimization for learning thoery" , *Adv. Neural Inf. Processing Syst.* Vol. 4, pp. 831-838.
11. Vapnik, V.N. (1995). *The Nature of statistical learning theory*, Springer Verlag, NY.
12. Principe, J.C., Euliano, N.R., and Lefebvre, W.C. (2000). *Neural and adaptive systems; fundamentals through simulation*, John Wiley & Sons, Inc., NY.

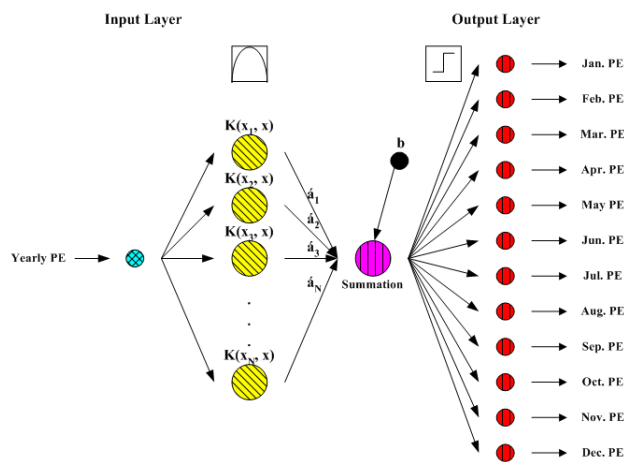


그림 1. SVM-NNM 의 구조

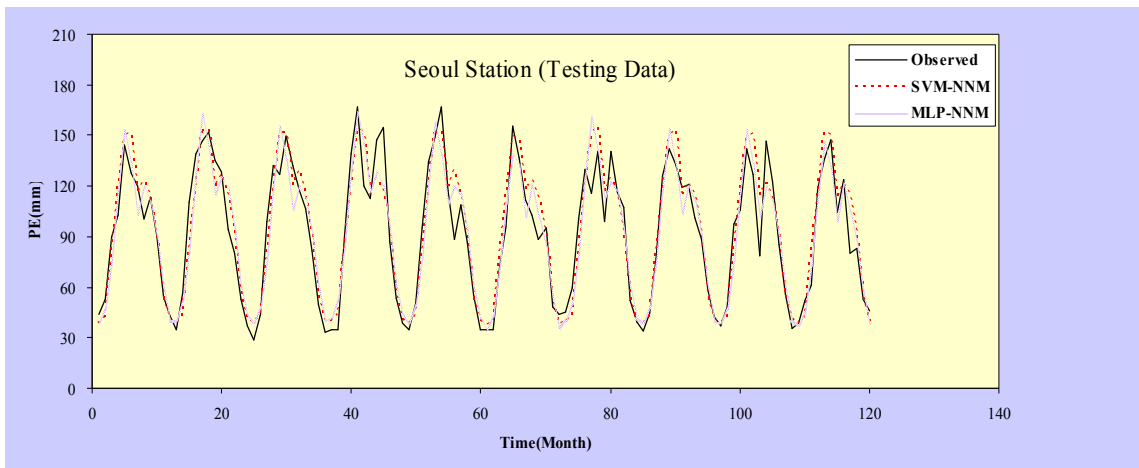


그림 2(a) 증발접시 증발량의 비교 (서울)

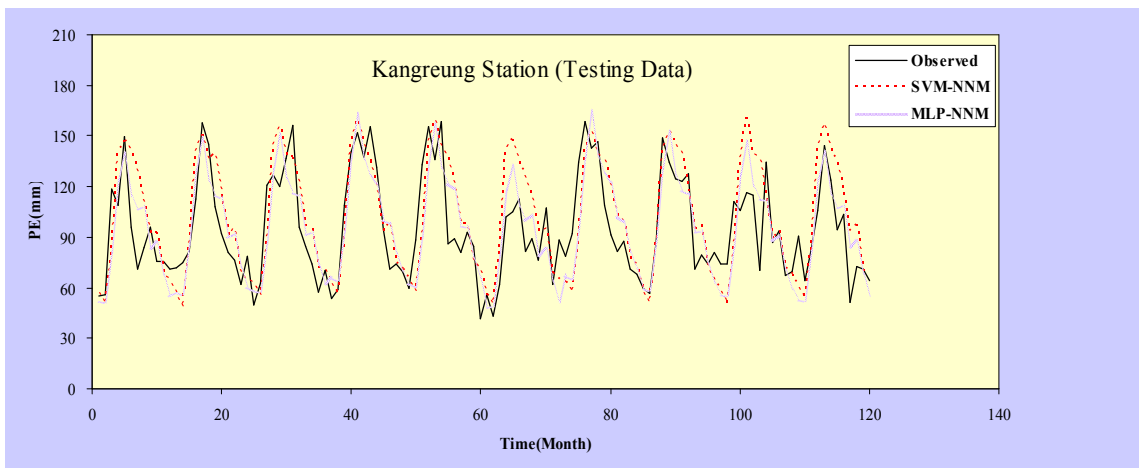


그림 2(b) 증발접시 증발량의 비교 (강릉)

표 1. 테스트자료를 이용한 SVM-NNM과 MLP-NNM 통계분석 결과 (서울 및 강릉관측소)

지점 \ 통계특성치	서울		강릉	
	SVM-NNM	MLP-NNM	SVM-NNM	MLP-NNM
CC	0.94	0.92	0.80	0.77
RMSE(mm)	13.82	15.32	18.72	22.92
E	0.88	0.85	0.619	0.43
MAE(mm)	10.35	11.39	15.23	18.16