관개배수 네트워크 시스템 구축을 위한 시계열자료의 모형화

Modeling of Time Series for Irrigation and Drainage Networks System

김성원*

Abstract

The goal of this research is to apply the neural networks model for the disaggregation of the pan evaporation (PE) data, Republic of Korea. The neural networks model consists of recurrent neural networks model (RNNM). The disaggregation means that the yearly PE data divides into the monthly PE data. And, for the performances of the neural networks model, it is composed of training and test performances, respectively. The training and test performances consist of the historic, the generated, and the mixed data, respectively. From this research, we evaluate the impact of RNNM for the disaggregation of the nonlinear time series data. We should, furthermore, construct the credible data of the monthly PE from the disaggregation of the yearly PE data, and can suggest the methodology for the irrigation and drainage networks system.

Key words : RNNM, Irrigation and Drainage Networks System, Hydrologic Time Series, Disaggregation Model

1. 서 론

증발접시의 이용은 도구적인 제한과 관측오차, 유수의 교란, 조류 혹은 다른 동물들에 의한 유수의 이용뿐만 아니라, 증발량 관측의 정도를 감소시킬 수 있는 여러 가지 유지문제와 같은 실제적인 문제에 의하여 영향을 받는다 (Jensen, et al., 1990). 증발접시가 이용될 수 없는 곳에서의 증발접시 증발량의 추정에 대한 실제적인 의미는 수문학자, 농학자 및 기상학자들에게 고려할 만한 중요성이 있는 것이다 (Kisi, 2006). 증발과정은 자연현상에서 매우 심한 비선형관계이다 (Kisi, 2006; Eslamian et al., 2008). 그러므로 이러한 필요성은 진행과정에 있어서 내재적인 비선형성을 나타낼 수 있는 모형을 통하여 접근 할 수 있다.

본 연구의 목적은 관개배수 네트워크의 구축을 위한 수문학적 시계열자료의 모형화를 개발 및 적용하는 것이다. 특히 연 증발접시 증발량의 시간적인 분해를 위하여 순환신경망모형 (Recurrent neural networks model, RNNM)을 개발 및 적용하는데 있다. 그리고 RNNM의 수행평가를 위하여 훈련 및 테스트과정으로 구성되었다. RNNM의 훈련과정을 위하여 실측, 모의 및 혼합자료와 같은 세 가지 형태의 자료가 사용되었으며, 테스트과정을 위해서는 실측자료만 이용되었다. 본 연구에서는 비선형 수문학적 시계열자료의 시간적 분해를 위해서 RNNM의 적용성을 평가하였다. 게다가 연 증발접시 증발량 자료의 시간적 분해로부터 신뢰성있는 월 증발접시 증발량자료를 구축할 수 있을 것이다.

2 순환신경망모형 (Recurrent Neural Networks Model, RNNM)

순환신경망모형 (RNNM)은 은닉층 및 출력층과 같은 임의의 한 층의 노드에서의 결과치가 이전 층의 노드로 피드백이 이루어지며, 피드백을 받는 층의 노드의 결과치는 시간 t 에서의 이전 층의 입력자료와 전 시간 (t-△t)에 존재한 피드백 한 층의 노드 자체의 결과치의 함수로 나타낼 수 있다. 여기서 △t는 연산의 한번 과정에 대한 시간을 나타내고 있다. 이러한 순환신경망모형은 모형의 결과치가 현재와 이전의 입력자료에 의존하기 때문에 단기간 기억장치 (Short-term memory)와 같은 유사한 특징을 나타내고 있다 (Giles et al., 1997; Li et al., 1989; Ryeu and Chung, 1996). RNNM은 시간적 및 공간적인 패턴을 인식하고 상기시킴으로서 전방향신경망모형보다 잠재적으로 강력하나, 운영에 있어서 전방향신경망모형보다 많이 복잡한 단점이 있다. 또한 전방향신경망모형은 출력값이 단지 구성된 입력층 노드의 함수로 구성되나 RNNM은 출력값이 시간의 함수로 구성되어 있으며, 주어진 입력자료와 초기 출력값에 대하여 안정적인 결과치로서 수렴하게 된다. 그리고 이러한 안정적인 결과치를 유도하기 위하여 전방향신경망모형과는 상이하게 많은 은닉층 노드수를 요구하고 있다 (Demuth and Beale, 2000). 본 연구에서 RNNM의 은닉층 노드수는 12개, 은닉층의 전이함수는 signoid, 학습규칙은 Quickprop 역전파 훈련알고리즘, 단계크기 (Step size)는 0.01 및 모멘텀은 0.50으로 설정하였으며, 출력층의 경우 출력층 노드수는 12개, 출력층의 전이함수는 linearsignoid, 학습규칙은 Quickprop 역전파 훈련알고리즘, 단계크기 (Step size)는 0.01 및 모멘텀은 0.50으로 설정하였다. 또한 최대반복회수는 10,000회, 오차임계치는 0.001로 설정하였다. 다음 Figure 1은 본 연구에 적용된 RNNM 구조를 나타낸 것이다.

3. 연구범위 및 자료

* 정회원, 공학박사, 수자원개발기술사, 동양대학교 철도토목학과 부교수 ; E-mail : swkim1968@dyu.ac.kr

본 연구에서는 제주도를 포함한 기상청의 관리하에 있는 71개의 기상관측소 중에서 서울, 강릉, 인천, 부산, 제주 및 목포관측소를 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 기상관측을 실시한 이후에 1919년, 1950-1951년 자료가 결측이 되어 있으나, 서울관측소의 경우처럼 자료의 신뢰성이 아주 양호한 상태이다. 인천관측소는 1949년 1월에 기상관측을 실시한 이후에 1950-1951년 자료가 결측이 되어 있으나, 자료의 신뢰성은 양호하다. 그리고 부산은 1965년 1월부터, 제주 및 목포는 1980년 1월부터 기상관측을 실시하였고 결측자료는 없는 상태이다.

4. 추계학적 모형

주기성 자기회귀이동평균 모형은 주기성 자기회귀 (Periodic Autoregression, PAR)모형에 주기성의 이동평균 매개변수를 포함하여 확장하며 일반적으로 PARMA (p,q)로 나타낸다. 저차수의 PARMA 모형은 주기적인 수문시계열의 모형화에 사용된다. 예를들면 PARMA (1,1) 모형은 다음 식(1)과 같이 나타낼 수 있다.

$$y_{v,\tau} = \mu_{\tau} + \Phi_{1,\tau}(y_{v,\tau-1} - \mu_{\tau-1}) + \varepsilon_{v,\tau} - \Theta_{1,\tau}\varepsilon_{v,\tau-1} \quad (1)$$

여기서 v 는 년(year), τ는 계절(Season)이고 $\tau = 1, 2, \cdots\cdots, \omega$이며 계절수를 나타내고 있다. 식 (1)과 같은 모형은 주로 월유출량계열에 적용된다 (Salas et al., 1980). 본 연구에서는 신경망모형의 훈련자료를 모의발생 하기위하여 PARMA (1,1) 모형을 이용하였다. 그리고 모의발생한 표본은 2개의 표본으로 하였으며, 500년에 해당되는 월별 증발량을 모의발생시켰다. 모의된 자료 중에서 첫 번째 표본자료는 선택하지 않았으며, 두 번째 표본자료를 선택하였다. 또한 두 번째 표본자료에서 초기에 발생되는 편차 (Bias)를 제거하기 위하여 초기의 50년 자료는 제거하였으며, 각 PARMA (1,1) 모형의 매개변수는 근사 최소자승법 (Method of approximate least square)을 이용하여 산정하였다. 따라서 제시된 RNNM의 입력층 변수의 모의자료의 경우는 모의발생 시킨 월별 소형 증발량 증발량자료를 합한 결과치에 해당되는 연간 소형 증발접시 증발량자료이고, 출력층 변수는 모의발생시킨 월별 소형 증발접시 증발량자료이다.

5. RNNM의 적용

5.1 훈련과정 (Training Performance)

5.1.1 실측자료의 이용

본 연구의 훈련과정에서 훈련자료는 서울관측소의 경우는 1908년부터 2002년까지의 결측기간을 제외한 86년의 자료를 선택하였으며, 강릉관측소의 경우는 1912년부터 2002년까지의 결측기간을 제외한 88년의 자료를 선택하였다. 인천관측소의 경우는 1949년부터 2002년까지의 결측기간을 제외한 52년의 자료를 선택하였다. 부산관측소의 경우는 1965년부터 2002년까지의 38년의 자료를 선택하였다. 그리고 제주 및 목포관측소는 1980년부터 2002년까지의 23년의 자료를 선택하였다. 그리고 RNNM의 훈련에 따른 선행과정으로서 입력층의 입력변수인 연간 소형 증발접시 증발량자료를 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 (Kim and Kim, 2008a, b).

5.1.2 모의자료의 이용

연구의 모의자료의 이용을 위한 훈련과정에서 훈련자료는 “5.1.1 실측자료의 이용” 에서 선정한 월 증발접시 증발량자료를 모의자료의 구축을 위하여 이용하였다. 따라서 RNNM의 훈련과정을 위한 입력층자료는 본 연구에서 선정한 추계학적 모형인 PARMA (1,1)을 이용하여 500년의 자료를 모의발생시켰다. 모의발생된 월 증발접시 증발량자료 중에서 초기에 발생하는 편차 (Bias)를 제거하기 위하여 초기의 50년 자료는 제거하였으며, 나머지 450년의 월 증발접시 증발량자료를 이용하였다. 따라서 입력층자료에 해당하는 연 증발접시 증발량은 각 월 증발접시 증발량자료를 합하여 선정하였다. 그러므로 RNNM의 훈련과정을 위한 입력층자료는 450개의 모의된 연 증발접시 증발량자로로 구성되어 있다.

5.1.3 혼합자료의 이용

본 연구의 혼합자료의 이용을 위한 훈련과정에서 훈련자료는 “5.1.1 실측자료의 이용” 에서 선정한 실측된 월 증발접시 증발량자료와 “5.1.2 모의자료의 이용” 에서 선정한 모의된 월 증발접시 증발량자료를 합한 혼합자료로 구성하였다. 따라서 RNNM의 훈련과정을 위한 입력층 자료수는 각각 서울관측소는 536개, 강릉관측소는 538개, 인천관측소는 502개, 부산관측소는 488개, 제주 및 목포관측소는 473개의 혼합된 연 증발접시 증발량자료로 구성되어 있다.

5.2 테스트과정 (Test Performance)

본 연구의 테스트과정에서 테스트자료는 서울, 강릉, 인천, 부산, 제주 및 목포관측소에서 동일한 시계열인 2003년부터 2007년까지의 5년간의 자료를 선택하였다. 테스트과정은 훈련과정을 통하여 산정된 매개변수를 이용하여 테스트자료를 검증하였다. 다음 Figure 2(a)-(f)는 서울, 강릉, 인천, 부산, 제주 및 목포관측소의 실측자료를 이용한 훈련과정의 수행결과 산정된 매개변수 및 테스트자료를 이용하여 RNNM의 테스트과정의 수행결과에 대한 월 증발접시 증발량자료의 시계열의

비교를 나타낸 것이다. 또한 다음 Table 1은 실측자료를 이용한 훈련과정의 수행결과 산정된 매개변수를 이용하여 RNNM의 테스트과정의 수행결과를 나타내고 있다.

6. 결 론

본 연구에서는 RNNM을 이용하여 연간 소형 증발접시 증발량자료를 월별 증발접시 증발량자료로의 분리를 시도하였으며, 사용된 자료는 서울, 강릉, 인천, 부산, 제주, 및 목포관측소의 자료를 이용하였다. RNNM의 훈련과정의 수행결과 산정된 각 6종류의 매개변수 군을 이용하여 테스트과정을 수행하였다. 본 연구를 통하여 분석에 이용된 자료의 수는 RNNM의 수행능력에 영향을 끼치는 것으로 판단된다. 또한 본 연구를 통하여 RNNM을 이용하여 연 증발접시 증발량자료에 대한 시간적 분해모형의 구축에 대한 적용성을 검토하였으며, 그에 따른 잠재성을 분석하였다. 그러나 시간적 분해모형의 보편적인 해석과 결론을 유도하기 위해서는 예를 들면 강우 및 유출량과 같은 더욱 다양한 시계열자료를 이용한 검증과정이 필요하며, 많은 종류의 신경망모형을 이용한 지속적인 연구가 필요한 것으로 판단된다.

참고문헌

1. Demuth, H. and Beale, M. (2000). *Neural network toolbox : for use with MATLAB user's guide*, The MathWorks Inc.
2. Eslamian, S.S., Gohari, S.A., Biabanaki, M., and Malekian, R. (2008). "Estimation of monthly pan evaporation using artificial neural networks and support vector machine." *J. Appl. Sci.,* Vol. 8, No. 19, pp. 3497-3502.
3. Giles, C.L., Lawrence, S., and Tsoi, A.C. (1997). "Rule inference for financial prediction using resurrent neural networks." *Proc. of 1997 IEEE/IAFE Conference on Computational Intelligence for Financial Engrg.*, IEEE Press, Piscataway, NJ, pp. 253-259.
4. Jensen, M.E., Burman, R.D., and Allen, R.G. (1990). *Evapotranspiration and irrigation water requirements*, ASCE Manual and Report on Engineering Practice No. 70, ASCE, NY. USA.
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. Kisi, O. (2006). "Daily pan evaporation modeling using a neuro-fuzzy computing technique." *J. Hydro,*, Vol. 329, pp. 636-646.
8. Li, J., Michel, A.N., and Porod, W. (1989). "Analysis and synthesis of a class of neural networks : linear system operating on a closed hypercube." *IEEE Trans. on Circuits and Systems*, Vol. 36, No. 11, pp. 1405-1422.
9. Ryue, J.K. and Chung, H.S. (1996). "Chaotic recurrent neural networks and their application to speech recognition." *J. Neurocomputing*, Vol. 13, pp. 281-294.
10. Salas, J.D., Delleur, J.R., Yevjevich, V., and Lane, W.L. (1980) *Applied modeling of hydrologic time series,* Water Resor. Pub., Littleton, CO.

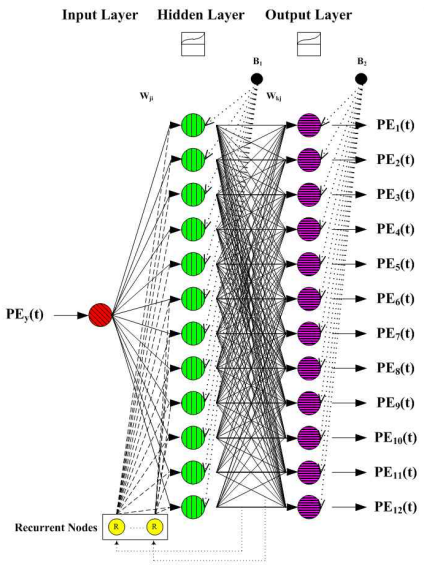
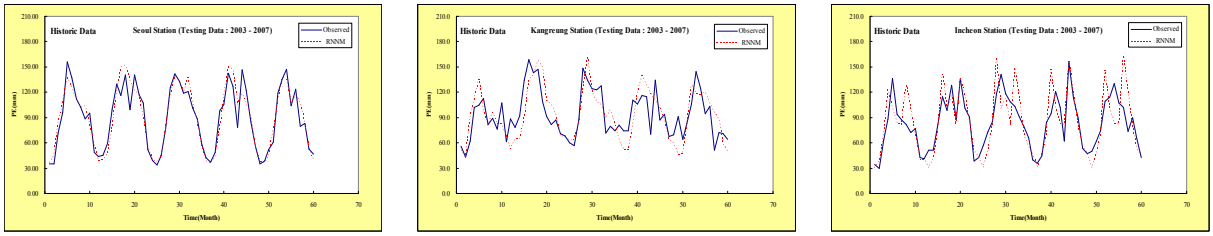


Figure 1. RNNM structure



(a) Seoul (b) Kangreung (c) Incheon

(d) Busan (e) Jeju (f) Mokpo

Figure 2. Comparison of the monthly PE for the testing performance (Historic data)

Table 1. Statistical analysis of the monthly PE for the testing performance (Historic data)

Station	Statistical Index	SVM-NNM
Seoul	CC	0.940
	RMSE	12.978
	E	0.878
	AARE	0.0131
Kangreung	CC	0.740
	RMSE	21.074
	E	0.418
	AARE	0.0329
Incheon	CC	0.792
	RMSE	22.847
	E	0.452
	AARE	0.0003
Busan	CC	0.818
	RMSE	17.122
	E	0.570
	AARE	0.0037
Jeju	CC	0.788
	RMSE	24.650
	E	0.579
	AARE	0.0258
Mokpo	CC	0.808
	RMSE	21.659
	E	0.643
	AARE	0.0618