

자기조직화방법을 적용한 강우-유출과 TOC변동에 관한 패턴 분류 및 분석

Pattern Classification and Analysis of Rainfall-Runoff and TOC Variation by the application of Self Organizing Map

박성천*, 김종록**, 진영훈***, 정천리****

Jong Rok Kim, Young Hoon Jin, Cheon Lee Jeong, Sung Chun Park

요 지

본 연구는 강우-유출 및 TOC의 패턴 분류를 위하여 광주 광산 강우관측소의 강우량자료와 나주지점의 유출량 그리고 기존의 BOD 및 COD 수질농도 측정값에 비하여 적은 오차요인과 빠른 시간에 결과 값을 얻을 수 있으며 유출량과 난분해성 물질에 대한 해석이 가능하고 재현성이 탁월한 TOC자료를 사용하였다. SOM을 적용하기 위해 먼저 Map의 크기는 Garcia가 제시한 $M=5\sqrt{N}$ 을 이용하여 결정한다. 이러한 비선형적인 다변량 자료를 분석하기 위해서 Map에 의해 구분된 자료 위치를 추출하여 원자료를 재구축하고 이를 통해 원자료를 패턴별로 분류 할 수 있었다. 이러한 패턴별 분류를 통해 유출량에 따른 TOC자료를 2차원의 Map 상에 시각적으로 가시화하여 비선형적인 경향이 강한자료의 분포적 양상을 이해하는데 큰 도움이 되며, 향후 이를 통해 예측을 위한 모형화 과정에도 크게 도움을 줄 것으로 기대된다. 또한, 강우자료 또는 유출량 자료만을 이용한 단일변량의 패턴분류를 위해 SOM의 적용이 가능할 것으로 판단되며, 이는 각 변량의 본질적인 특성을 파악할 수 있을 것으로 기대된다.

핵심용어 : 인공신경망, 자기조직화, 강우-유출, TOC, 인공신경망, 패턴 분류

1. 서론

수자원의 효율적인 이용을 위해서는 홍수에 의한 침수나 각종 수질오염 사고에 의한 재난재해를 보다 적극적으로 대응하고, 하천수 수질관리 및 환경정책 입안의 기초자료로 활용을 위해서는 각종 유량과 수질자료에 대한 특성 분석이 요구되고 있으며, 비선형 경향이 강한 하천수 유출량과 총유기탄소(Total Organic Carbon: 이하 TOC)에 관한 관계를 모형화하기 위해서는 이들 상호관의 특성을 반영할 수 있는 모형개발이 절대적으로 요구된다. 이에 따른 방법론으로는 ANN모형 자체적으로 경쟁(competitive) 또는 자기조직화(self-organization)에 의하여 가중치들을 산정하여 패턴을 분류하는데 탁월한 성능을 지닌 자기조직화(Self-Organizing Map: 이하 SOM) 이론을 적용하여 강우량 자료와의 조합을 통해 하천수 유출량과 TOC변동에 따른 패턴 분리를 실시하였다. 그 결과 강우-유출 및 TOC 관계가 7개의 패턴으로 구분되었다.

2. 입력자료 구성 및 모형의 적용

본 연구에서는 강우-유출 및 수질(TOC) 분포의 비선형적 변동성을 고려한 강우-유출 및 TOC 간의 특

* 정회원·동신대학교 토목공학과 교수·E-mail : psc@dsu.ac.kr
** 정회원·동신대학교 토목공학과 석사과정·E-mail : kjr81@hanmail.net
*** 정회원·동신대학교 토목공학과 연구원·E-mail : nmdrjin@gamil.com
**** 정회원·영산강홍수통제소 조사관리과·E-mail : ch1010@moct.re.kr

성을 분석하기 위해 영산강 본류의 대표지점인 나주수위관측소와 그에 영향을 미치는 상류의 광산 강우관측소를 선정하였다.

SOM 훈련 데이터 구성은 강우-유출 및 TOC 사이의 관계를 파악하기 위해서는 긴 시간의 다수의 자료가 필요하나 본 연구에서 자료의 검토결과 강우-유출 및 TOC 관계를 파악하기 위한 최소한의 시간이 세 개의 연속자료를 나타내는 3시간으로도 그 특성 파악이 가능하다고 판단되고, 자료의 크기를 최소화하기 위하여 강우-유출 및 TOC 특성을 구분하기 위하여 3시간 자료를 SOM의 입력자료로 이용하였다.

그러므로 입력벡터 $x(t)$ 는 식 1과 같은 9차원 구조의 SOM 훈련 데이터의 구조가 된다.

$$x(t) = (P_{(t-2)}, T_{(t-2)}, Q_{(t-2)}, P_{(t-1)}, T_{(t-1)}, Q_{(t-1)}, P_{(t)}, T_{(t)}, Q_{(t)}) \dots \dots \dots \text{식 (1)}$$

여기서, P는 강우량, T는 TOC, Q는 유출량

SOM의 적용을 위해 그림 1의 SOM 적용 흐름도와 같이 강우-유출 및 TOC 자료에 대해 2003년 3월 16일~2003년 6월 10일, 2004년 2월 15일~2004년 6월 5일까지의 4774개의 자료를 선택하여 3시간의 시차를 적용한 4774×9 행렬의 SOM 입력자료를 구축하였다.

선택된 자료의 정규화 과정을 거친 후 SOM 훈련을 위해 Map 크기를 결정한다. Map 크기 결정은 Garcia 등(2003)에 의해 연구 보고된 식을 이용하였다(식 (2)).

$$M = 5\sqrt{N} \dots \dots \dots \text{식 (2)}$$

여기서, M은 Map을 구성하는 단위구조(unit)의 총 수 N은 훈련 자료의 수이다.

본 연구에서의 훈련자료는 4774쌍 이며, Map 크기는 식 (2)에 의해 산정된 345의 근사치로부터 그림 2와 같이 23×15로 결정하였다. 훈련방법으로는 일괄 훈련방법(batch training algorithm)을 이용하였고, k-means Clustering 결과 최적의 분할구역으로 구분된 결과는 그림 2와 같다. 또한 분할구역의 대표적인 강우-유출 및 TOC 간의 관계는 그림 3과 같다.

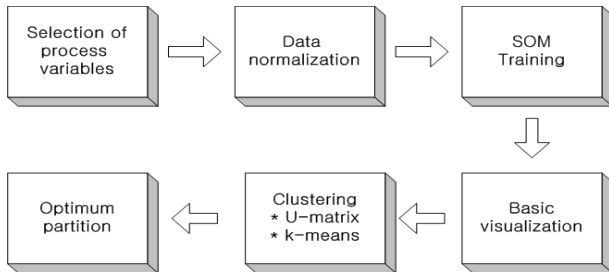


그림.1 SOM 적용 흐름도

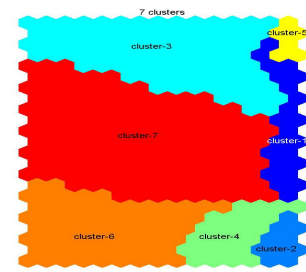


그림.2 k-means Clustering 결과

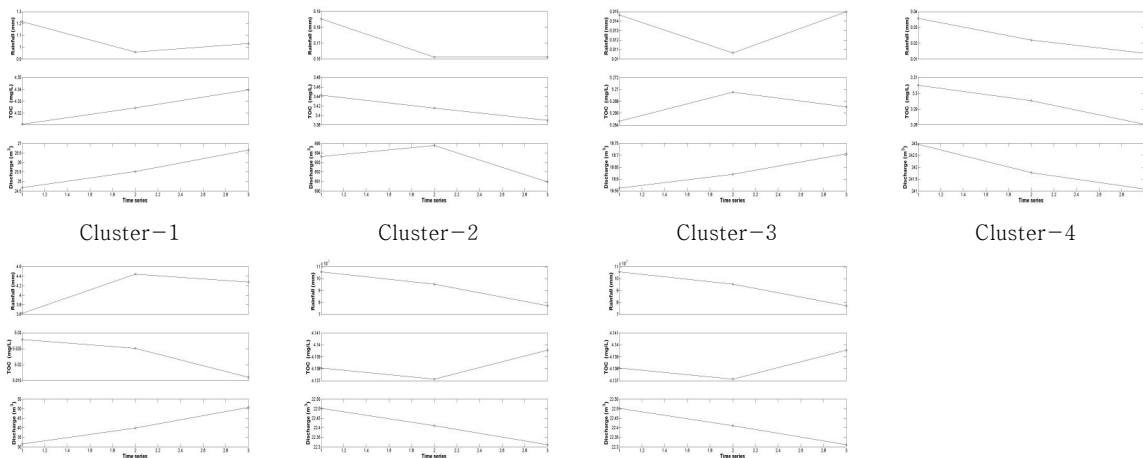


그림.3 분할구역의 대표적인 강우-유출 및 TOC관계

3. 분석 및 결과

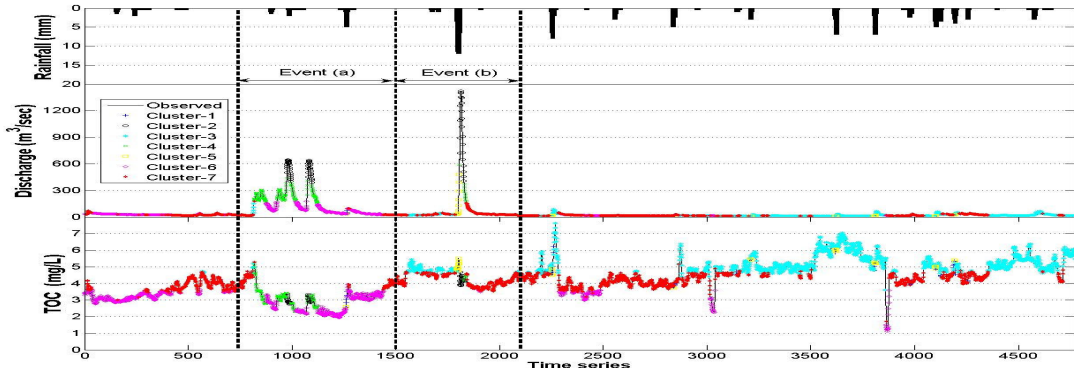


그림.4 SOM에 의해 구분된 강우-유출 및 TOC 관계 분석

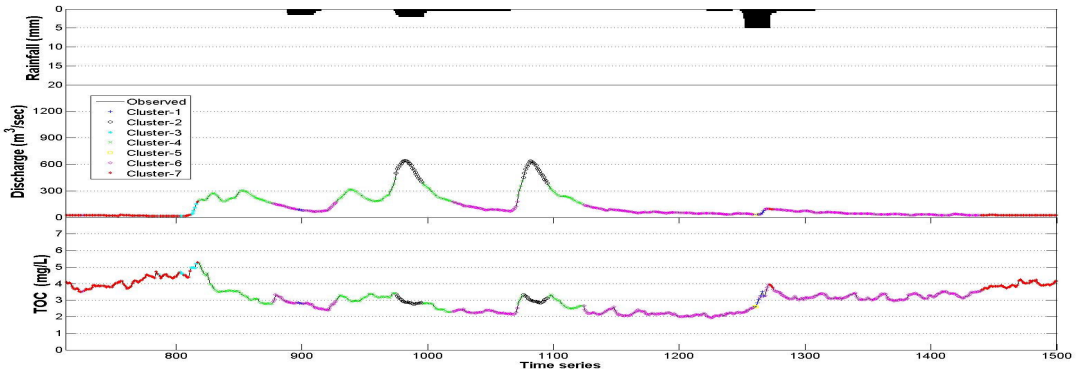


그림.5 강우-유출 및 TOC 관계 분석 (a)

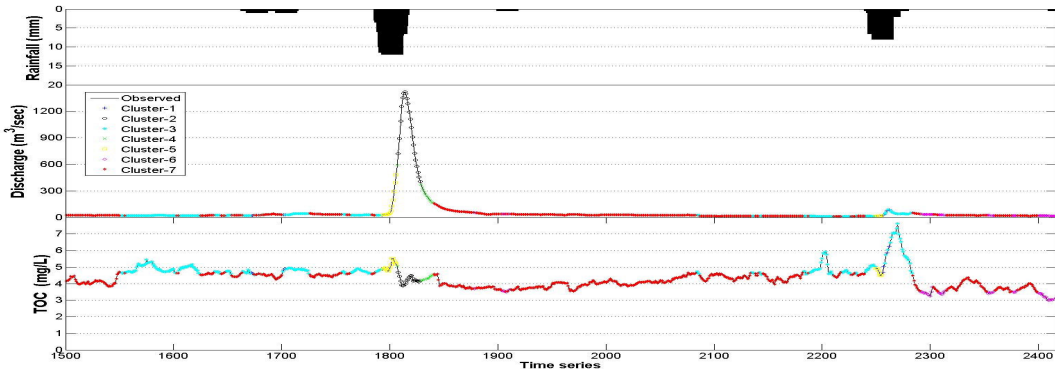


그림.6 강우-유출관계 분석 (b)

SOM에 의해 구분된 강우-유출 및 TOC 관계를 규명하기 위해 2003년 3월16일~2003년 6월 10일, 2004년 2월 15일~2004년 6월 5일 까지 나주지점의 수문곡선과 TOC농도 곡선에 분할된 구역의 입력자료를 선택하여 그림 4와 같이 도시하였다. 또한 명확한 결과의 도시를 위하여 강우발생 전후로 한 기간 (a), (b)를 선별하여 그림 5~그림 6에 도시하였다.

그 결과 Cluster-1은 강우가 시작되는 단계에 해당하며 유출량과 TOC가 점진적으로 상승하는 경향을 나타내고 있으며, Cluster-2는 강우가 종료되는 단계에서 유출량은 대규모로 발생하며 수문곡선의 침두부를 포함한 상승부와 하강부로 구분되었으며, TOC농도는 유출량의 증가로 그 값이 감소하는 경향으로 나타났다.

Cluster-3은 규모가 작은 강우가 발생하며 유출량은 점진적으로 증가하고, TOC농도는 증가하는 경향을 나타내고 있으며, Cluster-4는 강우량은 감소하고 유출량은 수문곡선에 변곡점 아래의 상승부와 하강부의 중

규모 유출량의 양상으로 구분되었고, TOC농도는 약 3.28~3.31mg/L에서 감소하는 부분으로 분할되었다.

Cluster-5는 강우량이 증가함에 따라 유출량은 급격히 증가하고, TOC농도가 비교적 5.03mg/L의 고농도에서 서서히 감소를 하는 양상으로 구분되었으며, Cluster-6은 강우가 종료되고 유출량은 감소하며, TOC농도는 2.98mg/L의 저농도에서 차차 증가하는 경향으로 구분 되었다. 마지막으로 Cluster-7은 무강우 기간에 해당되며, 유출량은 점진적으로 감소하거나 변화가 없는 반면에 TOC농도는 증가하는 경향으로 구분되었다. 이상의 결과와 같이 강한 비선형성을 가지고 있는 강우-유출 및 TOC관계가 SOM으로 강우-유출 및 TOC 패턴분류가 가능함을 확인 할 수 있었고, Cluster-2와 Cluster-4는 수문곡선의 상승부와 감수부의 변곡점을 구분지어 주는 것으로 판단되며, Cluster-5와 같이 강우가 증가함에 따라 유출량은 급격히 증가하는 수문곡선의 상승부로 구분되어지는 반면에 TOC는 유출량과 반대되는 경향인 고농도에서 차차 감소하였다. 이는 초기강우로 인한 유역내에 산재되어 있던 오염원의 집중현상으로 판단된다.

4. 결론

본 연구에서는 시·공간적 분포의 비선형적 변동성을 고려하기 위해 인공신경망의 기법의 일종인 자기조직화방법을 이용하여 강한 비선형성을 가지고 있는 강우-유출 및 TOC 관계가 SOM으로 패턴분류가 가능함을 확인 할 수 있었다. 또한 자기조직화 방법으로 구분된 분할구역(Cluster-1~Cluster-7)의 특성을 분석해보면 Cluster-1은 강우가 시작되는 단계에 해당하며 유출량과 TOC 농도가 점진적으로 상승하는 경향을 나타내고 있으며, Cluster-2는 강우가 종료되는 단계에서 유출량은 대규모로 발생하며 수문곡선의 첨두부를 포함한 상승부와 하강부로 구분되었으며, TOC농도는 유출량의 증가로 그 값이 감소하는 경향으로 나타났다.

Cluster-3은 규모가 작은 강우가 발생하며 유출량은 점진적으로 증가하고, TOC 농도는 증가하는 경향을 나타내고 있으며, Cluster-4는 강우량은 감소하고 유출량은 수문곡선에 변곡점 아래의 상승부와 하강부의 중규모 유출량의 양상으로 구분되었고, TOC농도는 약 3.28~3.31mg/L에서 감소하는 부분으로 분할되었다. Cluster-5는 강우량이 증가함에 따라 유출량은 급격히 증가하고, TOC 농도는 고농도에서 서서히 감소를 하는 경향으로 현재 명확하게 구분되어지지 않은 초기강우의 양상을 뚜렷이 재현하였다. Cluster-6은 강우가 종료되고 유출량은 감소하며, TOC농도는 2.98mg/L의 저농도에서 차차 증가하는 경향으로 구분되었다. 마지막으로 Cluster-7은 무강우 기간에 해당되며, 유출량은 점진적으로 감소하거나 변화가 없는 반면에 TOC 농도는 증가하는 경향으로 구분되었다.

이상의 결과를 토대로 치수와 이수측면에서 보다 신뢰성 있는 유출량 예측과 수질예측에 기여할 수 있을 것으로 기대되며, 초기강우 현상의 정의를 위한 새로운 접근방법을 제시할 수 있을 것으로 판단된다.

참 고 문 헌

1. 김주환(1993) 신경회로망을 이용한 하천 유출량의 수문학적 예측에 관한 연구, 박사학위논문, 인하대학교.
2. 김용구(2006) 자기조직화 이론을 적용한 강우-유출 예측모형 개발, 박사학위논문, 동신대학교.
3. 정우철(2007) SOM을 이용한 호수수 저수량과 TOC변동에 관한 패턴 분류 및 분석, 석사학위논문, 동신대학교
4. 건설교통부 영산강홍수통제소(2004) 영산강 유량측정 보고서.
5. 김용구, 진영훈, 박성천(2006) 강우-유출특성 분석을 위한 자기조직화방법의 적용, **대한토목학회 논문집**, 제26권, 제1B호, pp.61-67
5. 박성천, 진영훈, 김용구(2006) 강우-유출 예측모형 개발을 위한 자기조직화 이론의 적용, **대한토목학회 논문집**, 제26권, 제4B호, pp.389-398
5. 심순보, 김만식, 심규철(1998) 신경망 이론에 의한 다목적 저수지의 홍수유입량 예측, **한국수자원학회 논문집**, 제 31권, 1호, pp. 45-57.
5. 이관수, 박성천, 이한민, 진영훈(2000) 인공신경망의 이론의 B.P.알고리즘을 적용한 영산강의 유출량 예측에 관한 연구, **대한토목학회 논문집**, 제 20권, 제5B호, pp. 679-688.
6. Coulibaly, P., Anctil, F. and Bobee, B.(2000) Daily reservoir inflow forecasting using artificial neural networks with stopped training approach, *Journal of Hydrology*, Vol. 230, pp. 224-257.

7. Garcia, H. L. and Gonzalez, I. M.(2004) Self-organizing map and clustering for wastewater treatment monitoring, *Engineering Applications of Artificial Intelligence*, Vol. 17, pp. 215-225.
8. Hsu, K. L., Gupta, H. V., Gao, X., Soroochian, S. and Imam, B.(2002) Self-organizing linear output map(SOLO): An artificial neural network suitable for hydrologic modeling and analysis, *Water Resources Research*, Vol. 38, No. 12, pp. 38-1-38-17.
9. Luk, K. C., Ball, J. E. and Sharma, A.(2000) A Study of optimal model lag and spatial inputs to artificial neural network for rainfall forecasting, *Journal of Hydrology*, Vol. 227, pp. 56-65.
10. Kohonen, T.(1982) Self-organized formation of topologically correct feature maps, *Biological Cybernetics*, Vol. 43, pp. 59-69.