

주암호 수문자료의 자기조직화 패턴분류 및 분석

Self Organized Pattern Classification and Analysis of Hydrologic Data in Juam Lake

박성천*, 진영훈**, 노경범***, 양동현****

Sung-Chun Park, Young-Hoon Jin, Kyong-Bum Roh, Dong-Hyun Yang

요 지

우리나라는 여름철에 강우가 편중되어 있고 동고서저의 산악지형으로 수자원확보가 어려운 실정이며 이는 곧 하천의 유지유량확보의 어려움과도 직결된다. 이러한 수자원확보를 위해 최근 기존 저수지 독을 높이는 사업이 전국적으로 활발히 진행되고 있으며 이는 저수지나 댐의 수체와 같은 수자원을 보다 적극적으로 활용하여 그 가치를 높임과 동시에 하천에 대한 활용도를 높이고자 하는 데 그 목적이 있다. 따라서 저수지나 댐의 저류량에 기여하는 강우량, 유입량과 같은 수문학적 자료의 심도 있는 분석이 필요하며 수문변수들이 나타내는 복잡한 패턴에 대한 연구가 이루어져야 할 것이다.

본 연구에서는 저수지나 댐의 저류량에 직접적으로 영향을 주는 수문변수들을 전체적으로 파악하기 위해 수집된 수문자료의 각각의 특성 및 자료들 사이의 복합적인 관계를 파악하였으며 이를 위하여 패턴분류 분야에서 그 적용타당성이 입증된 자기조직화 지도(Self-Organizing Map: SOM)를 이용하였다. 본 연구의 대상지점은 섬진강 유역내에 위치한 주암호를 대상지점으로 선정하였으며 패턴분석에 사용한 수문자료의 기간은 2007~2010년까지 5년간의 월평균 자료를 활용하였다.

SOM의 적용 결과, 측정수문자료에 대한 전체적인 특성을 패턴분류를 통해 분류하였으며, 각 변수에 대한 패턴별 상대성을 고려한 클러스터별 특성 및 시간적 이질성을 파악할 수 있었다. 이는 측정 자료에 대한 분석 기법개발의 일환으로 향후 수자원 확보에 대한 개발 및 정책의 기초자료로 활용될 것으로 기대된다.

핵심용어 : 주암호 , 수문자료, 패턴분류, 자기조직화 지도

* 정회원 · 동신대학교 토목공학과 교수 · E-mail : psc@dsu.ac.kr
** 정회원 · 동신대학교 토목공학과 겸임교수 · E-mail : nmdrjin@gmail.com
*** 정회원 · 목포대학교 창업지원단 센터장 · E-mail : kbyj3711@naver.com
**** 정회원 · 동신대학교 토목공학과 박사과정 · E-mail : miziri@nate.com

1. 서론

본 연구는 주암호를 관리하기 위하여 한국수자원공사에서 측정하고 있는 수문자료의 활용성을 제고하고, 독을 높이는 사업과 같은 이·차수 사업정책의 수립에 대한 기초적인 정보 제공을 위하여 자료 분석 기법을 적용하고자 하였다. 자료 분석 기법으로는 최근 패턴분류를 위해 전 세계적으로 그 적용이 활발히 이루어지고 있는 Self-Organizing Feature Map (SOFM)을 활용하였다.

먼저 연구를 위한 대상지점은 섬진강 유역 내에 위치하고 있는 주암호를 대상지점으로 선정하였으며 한국수자원공사에서 제공하는 수문자료를 이용하였다. 현재 측정되고 있는 주암호관련 수문자료들 중 댐 유지관리에 중요도가 높은 항목들을 선별하여 본 연구에 적용하였다.

2. 대상 지점 및 입력자료

본 연구의 대상지점인 주암호는 전남 남부의 화순, 보성, 순천에 걸쳐있으며 주암댐으로 인해 생겨난 인공호수이다. 주암호는 광주광역시, 나주시, 목포시, 화순군 등 전라남도 서부권에 하루 64만톤의 생활용수를 공급하고 있는 식수원으로서 유지관리에 있어 중요한 지점이다.

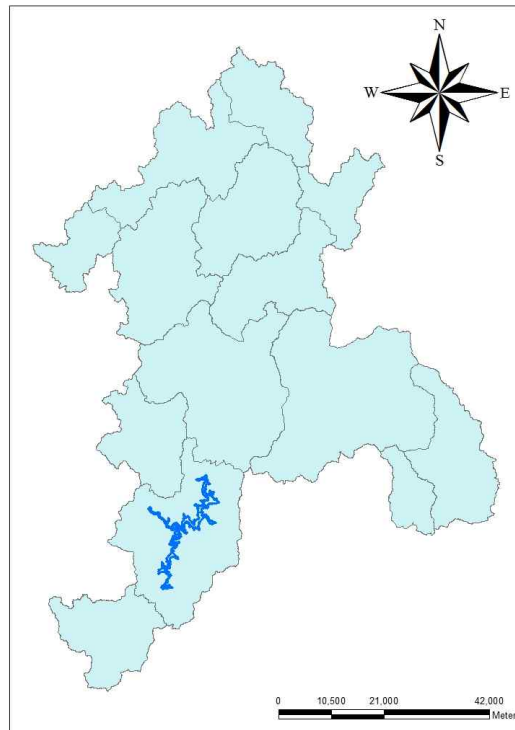


그림 1. 섬진강유역도 및 주암호 위치도

본 연구에 이용된 입력자료는 한국수자원공사에서 제공하는 주암댐의 수문자료를 수집하였으며, 댐 유지관리에 중요한 항목인 강우량, 유입량, 방류량, 저류량의 자료를 이용하였다. 입력자료의 기간은 2007~2011년의 5년간의 자료를 이용하였으며 강우량 및 저류량은 월별합계, 유입량과 유출량은 월평균으로 산정하여 적용하였다.

3. 자기조직화 특성 지도

Helsinki대학의 Kohonen(1982)은 사용자의 개입없이 스스로 입력 자료를 조직화할 수 있는 시

시스템을 소개하였으며, 이 시스템은 사용자의 도움 없이 외부의 공간에 위치하는 입력자료를 시스템이 정의한 내부의 공간으로 사상(mapping)시킨다. Kohonen은 이 시스템을 Self-Organizing Feature Map 또는 Adaptive Vector Quantizer (AVQ)라고 명명하였다. 그 이후로 다양한 종류의 SOFM이 나타나게 되었고, 많은 발전이 있었다.

SOFM은 다차원의 자료를 2차원으로 사상시켜주는 인공신경망 이론의 일종으로 패턴 분류를 위한 방법으로 최근 다양한 분야에서 활발히 적용되고 있다. 특히, SOFM은 복잡한 다차원 자료의 패턴분류와 결과에 대한 효율적인 가시화가 쉽고, 클러스터링 결과의 구조를 미리 지정해줄 수 있는 장점을 가지고 있다.

SOFM은 비감독 학습방법(unsupervised learning algorithm)으로 스스로 n차원의 입력 자료들을 클러스터링하여 2차원으로 사상시킨다. SOFM 구조는 n차원의 입력자료를 표현하는 n개의 입력노드들과 k개의 분류영역을 표현하기 위한 k개의 출력노드로 구성되어 있다. 모든 입력노드들은 모든 출력노드들과 연결되어 있고 각각의 연결강도(weight)를 가진다. 일반적으로 입력노드는 입력자료를 네트워크로 전달하며, 출력노드는 전달된 입력벡터와 입·출력 노드들을 연결하는 연결강도벡터를 이용하여 두 벡터간의 거리(distance)를 계산하는 기능을 갖는다.

이러한 과정에서 각 노드들은 학습할 수 있는 특권을 부여받기 위해 서로 경쟁하며 거리가 가장 가까운 노드가 최종적으로 승리하게 된다. 이 승자 노드만이 출력신호를 보낼 수 있는 유일한 노드가 되며, 이 노드와 이와 인접한 이웃 노드들이 제시된 입력벡터에 대하여 학습이 허용된다. 이것은 학습에 있어서 새로운 접근방식이며, 이러한 모형이 제안되기 이전에는 네트워크에 있는 모든 노드들이 반복학습과정인 훈련을 통해 연결강도를 조정하는 방법을 이용해 왔다. Kohonen 네트워크의 학습 철학은 승자만이 출력을 할 수 있는 ‘승자독점(winner take all)’ 개념이며 승자와 그의 이웃들만이 그들의 연결강도를 조정할 수 있다.

4. SOFM 패턴분류

본 연구에 사용된 SOFM의 구조를 살펴보면 자료의 총 수(60개)에 대하여 40개의 노드를 사용할 수 있으며, 주성분 분석에 의한 제 1 고유치와 제 2 고유치의 비율에 의해 가로 및 세로축에 배치될 노드의 수를 산정하였다. 결과적으로 8(세로축) × 5(가로축) 배열의 SOFM을 구축할 수 있었으며, 훈련을 통해 생성된 U-matrix를 그림 2에 나타내었다.

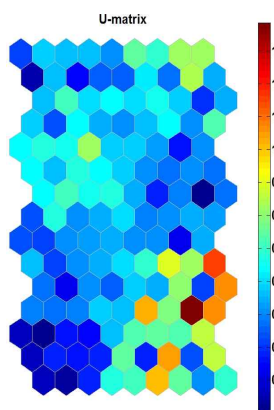


그림 2. U-matrix

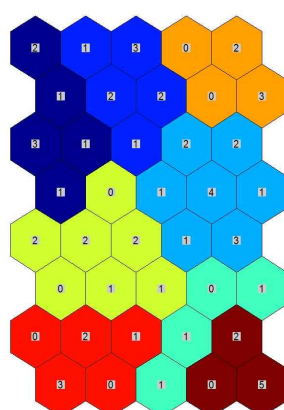


그림 3. 패턴분류 결과

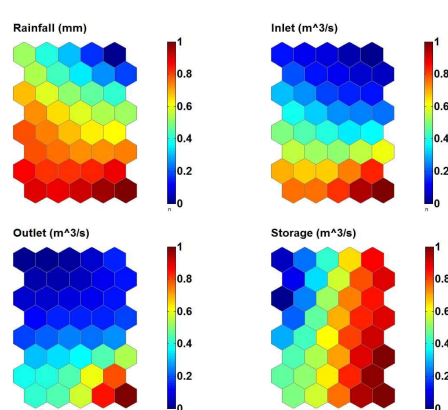


그림 4. 각 항목별 패턴분류 결과

본 연구를 위해 사용된 각 변수들에 대한 패턴분류 결과를 그림 3에 나타내었으며, 총 8개의 패턴으로 분류되었다. 각 입력변수에 대한 패턴분류 양상을 그림4의 요소평면(component plane)에 도시하였으며, 강우량과 유입량에 대한 요소평면은 유사한 변화 양상을 보이고 있다. 따라서 각 변수별 요소평면의 변화 양상에 따른 결과에 의해 강우량과 유입량은 서로 상관성을 가지고 있음을 알 수 있다.

각 클러스터별 패턴분류 특성을 나타내기 위하여 각 클러스터로 분류된 노드들의 가중치를 이용하여 그림 5와 같은 레이더 차트를 도시하였으며, 각 변수별 가중치들에 대한 1사분위~4사분위 값을 나타내었다. 분류된 노드의 가중치들을 동일변수에 대해 0과 1사이로 표준화 하여 각 클러스터들이 갖는 상대적 분류 양상을 검토하였다.

먼저, 가장 큰 강우량 자료들은 클러스터-8로 분류되었으며, 이에 따른 유입량, 저류량 및 방류량 자료 역시 큰 값을 보이고 있다. 클러스터-4와 7의 강우량 자료 역시 매우 큰 값을 보이고 있으나, 클러스터-4의 저류량 및 방류량 자료가 클러스터-7보다 큰 것으로 나타났다. 또한, 클러스터-3은 클러스터-5에 비하여 보다 작은 값의 강우량 및 유입량 자료를 포함하고 있으나, 더 큰 값의 저류량 자료를 포함하고 있는 것으로 나타났다. 클러스터-1 역시 클러스터-3과 5와 비교하여 유사한 크기의 강우량 자료를 가지고 있으나, 상당히 작은 값의 유입량, 저류량 및 방류량 자료를 포함하고 있다. 클러스터-2와 6은 공통적으로 매우 작은 값의 강우량, 유입량 및 방류량 자료를 포함하고 있으나, 클러스터-6의 저류량 자료가 상당히 큰 것으로 나타났다.

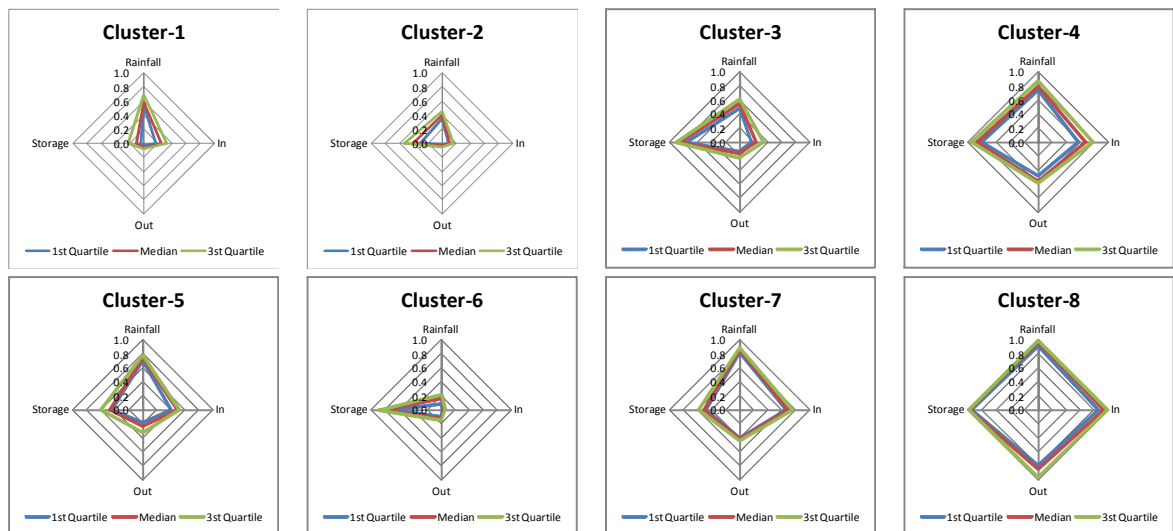


그림 5. 클러스터별 노드가중치

레이더 차트를 통한 각 패턴의 상대적 특성과 더불어 그림 6에 도시한 바와 같이 각 클러스터로 분류된 자료의 측정시기를 조사하여 각 패턴에 대한 시간적 특성을 분석하였다. 상술한 바와 같이 매우 큰 강우량 자료를 포함하고 있는 클러스터-4, 7, 8은 4월에서 10월 사이에 측정된 자료를 포함하고 있는 것으로 나타났다. 이와는 반대로 갈수기인 겨울철에 측정된 자료는 클러스터-2와 6으로 분류되었다. 클러스터-3은 갈수기 이외에도 봄과 가을에 측정된 자료들을 포함하고 있어 상대적으로 큰 값의 강우량 자료가 포함된 것으로 판단된다. 이상의 클러스터-2, 3, 6은 저류량이 다른 변수들에 비해 큰 값을 가지고 있으며, 시기적으로 갈수기 및 봄·가을에 측정된 자료를 포함하고 있으므로, 다른 클러스터들에 비해 저류량을 확보하기 위한 시기의 자료들이 다수 분류된 것으로 판단된다.

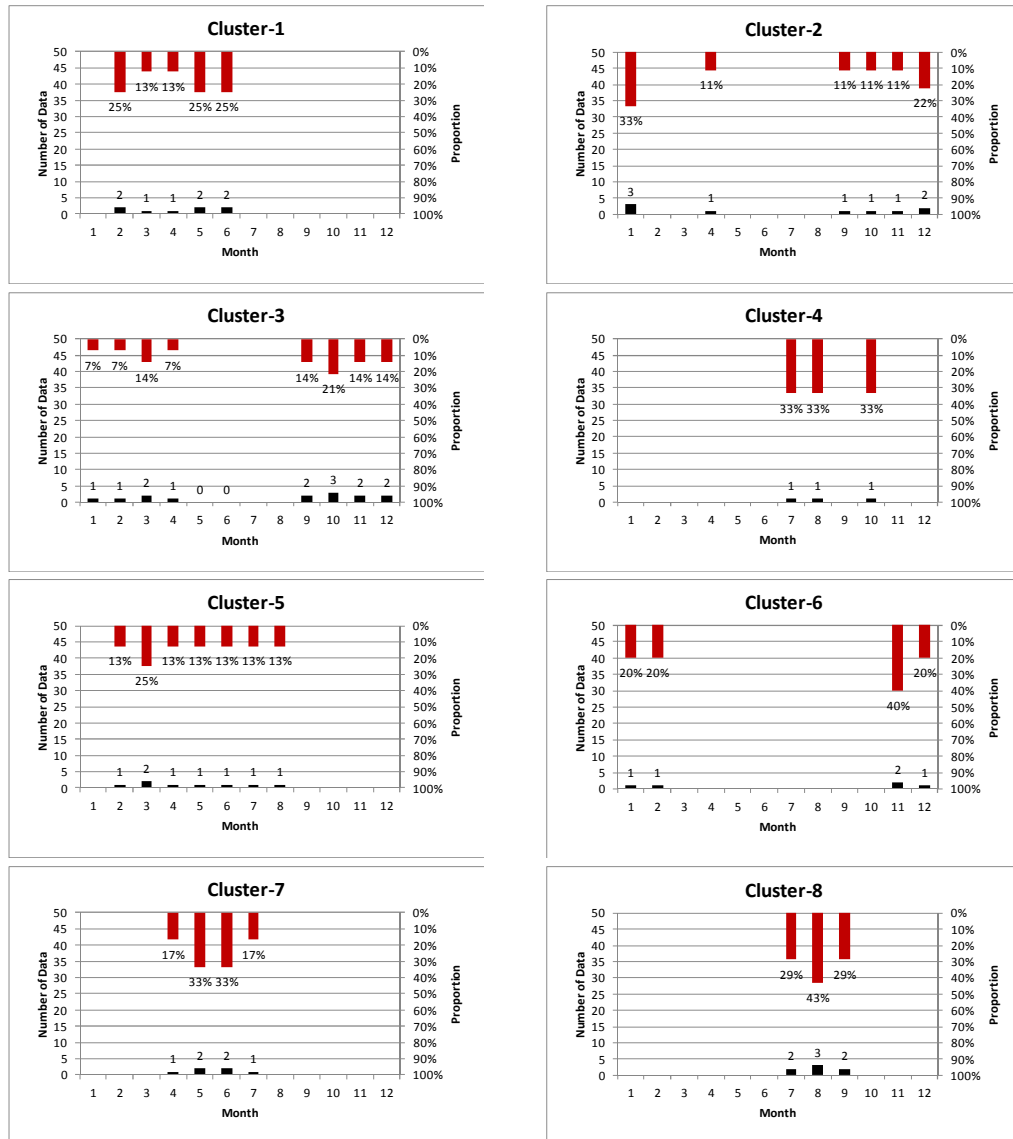


그림 6. 클러스터별 시간분포

5. 결 론

댐관리를 위해 측정되고 있는 한국수자원공사의 주암댐 수문자료의 활용성 제고를 위한 패턴분류 기법의 적용이 타당한 것으로 판단되며, 이에 따른 패턴분류 결과의 보다 심도 있는 분석을 통해 측정된 자료의 시간적 분포를 파악할 수 있었다. 이러한 결과는 향후 이·치수사업 정책의 입안 및 지속적인 추진을 위한 기초적인 정보를 제공해 줄 수 있을 것으로 기대된다.

참 고 문 헌

- Jin, Y.-H., Kawamura, A., Park, S.-C., Nakagawa, N., Amaguchi, H. and Olsson, J. (2011) Spatiotemporal classification of environmental monitoring data in the Yeongsan River basin, Korea using self-organizing map. J. Environ. Monit. 13, 2886-2894.