

EOF 해석과 추계학적 모형을 이용한 국내 가뭄특성의 평가

Evaluation of characteristics of the domestic drought using EOF analysis and stochastic model

유철상*, 김대하**, 김상단***, 김경준****, 김병수*****, 박창열*****
Yoo, Chulsang, Kim, Dae-ha, Kim, Sangdan,
Kim, Kyung-jun, Kim, Byung-su, Park, Chang-yeol

요 지

가뭄은 홍수와 함께 인류역사상 가장 큰 재해로 인식되어 있다. 미해양대기청의 발표에 따르면 20세기 최대 자연재해의 상위 5위 안에 4개의 가뭄이 포함되어 있다. 이러한 기록은 가뭄이 동서고금을 막론하고 국가의 흥망성쇠를 좌우할 만큼 막대한 피해를 입혀왔음을 의미한다. 그러나 가뭄의 해석은 가뭄의 정의 자체가 확실하지 않고 서서히 찾아오는 자연재해이기 때문에 그 시작과 끝을 인식하기 어렵다. 아울러 그 진행속도도 굉장히 느리며 또한 장기간에 걸쳐 지속되는 특성을 가지고 있고 시공간적으로 전파된다. 따라서 가뭄의 해석은 굉장히 까다로운 것이라 할 수 있으며 그 해석방법 또한 다양할 수 밖에 없다. 본 연구에서는 우리나라 전역 59개 지점의 표준강수지수(Standard Precipitation Index) 시계열 자료에 대한 공간적 패턴분석과 시간적인 자료확장을 시도하였다. 경험적 직교함수(Empirical Orthogonal function) 해석을 이용하여 자료의 공간적인 패턴을 확인하였고 EOF 해석에서 나타난 EOF Coefficient Time Series를 추계학적 모형에 적용하여 시간적인 자료 확장을 수행하였다. 이렇게 확장된 긴 기간의 자료를 이용하면 재현기간에 대한 평균적인 가뭄심도를 추출할 수 있으며 실제 나타난 사상의 재현기간이 어느 정도인지 평가할 수 있다. 또한 이렇게 나타난 가뭄심도를 강수부족량으로 환산하여 우리나라 대권역별 물부족량을 평가하였다.

핵심용어 : EOF 해석, 추계학적 모형, 표준강수지수

1. 서 론

가뭄은 과거와 현재에 걸쳐서 홍수와 함께 국가의 흥망성쇠를 좌우할 정도의 큰 재해로 여겨지고 있다. 미 해양대기청(NOAA)이 발표한 20세기 최대의 자연재해 중 상위 5위 안에 4개의 가뭄이 안에 포함되어 있을 정도로 가뭄은 막대한 피해를 입히는 재해라 할 수 있다. 또한 유엔은 2000년 한 해 동안 아시아 지역은 가뭄으로 인하여 약 1억 명의 기민이 발생했음을 발표한 바 있으며, 이는 한반도뿐 아니라 전 세계적으로 과거에서 최근까지도 가뭄이 지구촌에 막대한 피해를

* 정회원·고려대학교 사회환경시스템공학과 교수-E-mail : envchul@korea.ac.kr
** 정회원·고려대학교 사회환경시스템공학과 석사과정-E-mail : violt007@korea.ac.kr
*** 정회원·부경대학교 환경시스템공학부 전임강사-E-mail : skim@pknu.ac.kr
**** 정회원·고려대학교 사회환경시스템공학과 박사과정-E-mail : guitar@korea.ac.kr
***** 정회원·고려대학교 사회환경시스템공학과 석사과정-E-mail : keslo99@korea.ac.kr
***** 정회원·고려대학교 사회환경시스템공학과 석사과정-E-mail : changyeol@korea.ac.kr

발생하고 있음을 보여주는 예라 할 수 있다(경기개발연구원, 2004).

가뭄은 다른 재해와 비교했을 때 몇 가지 다른 특징이 있다. 먼저 가뭄은 아주 서서히 찾아오는 재해로 그 시작과 끝을 인식하기 매우 어렵다. 아울러 그 진행속도도 굉장히 느리며 장기간에 걸쳐 지속되는 특징을 가진다. 또한 그 진원지나 경로를 파악하기 어려우며 시공간적으로 전파된다(Wilhite and Glantz, 1985). 이러한 특징은 가뭄에 대한 예측을 어렵게 하며 따라서 이에 대한 대책 역시 마련하기 쉽지 않다. 이와 더불어 가뭄은 그 정의 자체가 공학적으로 확실하지 않기 때문에 이에 대한 혼돈으로 그 해석방법이 난무한 상황이다.

일반적으로 가뭄 해석에 사용하는 방법들은 상대적으로 긴 자료가 가용한 경우에 한한다. 자료가 긴 경우에는 관측자료 만으로도 가뭄의 특성을 파악하는 것이 가능하므로 이를 모형의 결과와도 쉽게 비교할 수 있다. 그러나 우리나라에서 대부분의 경우 가뭄을 특성화할 수 있을 정도로 충분한 길이의 강수기록은 존재하지 않는다고 할 수 있다. 홍수의 경우 최소 30-40년 정도의 기록으로도 빈도해석을 수행할 수 있으나 가뭄의 경우 수 년 동안 발생할 수도 있고 수십 년 동안 발생하지 않을 수도 있어 훨씬 긴 관측기록이 해석에 필요하게 된다(유철상 등, 2004).

본 연구에서는 자료기간이 짧은 국내 59개 지점의 3개월 표준강수지수(Standard Precipitation Index; SPI) 시계열을 경험적 직교 함수(Empirical Orthogonal Function) 해석으로 가뭄의 공간분포를 분석하였다. 또한 EOF 해석에서 나타난 시계열을 다변량 시계열모형(Multi-variate Time Series Model)에 적용하여 자료의 시간적 확장을 시도하였다. 이와 더불어 일정 재현기간에 대한 가뭄지수를 강수부족량으로 환산하여 유역별 농업가뭄 대비능력을 판단하고자 하였다.

2. 경험적 직교함수 해석(Empirical Orthogonal Function Analysis)

자연현상의 공간적 변동성은 주로 점형태로 구성된 관측망으로부터 얻어진 공간자료를 통하여 파악될 수 있다. 강수량 또한 이러한 점구조로 된 관측망의 네트워크를 통하여 공간적인 패턴을 분석해 볼 수 있다. 이 때, 어떤 관심있는 자연현상의 공간적인 변동성의 주된 특징은 경험적 직교함수(Empirical Orthogonal Function; EOF)라 불리는 기법에 의해 압축적으로 설명될 수 있다. EOF 해석은 관측망 네트워크로부터 오랜 기간 동안의 관측을 통해 얻은 시공간 정보를 분석하여 몇몇의 핵심적인 공간정보로 압축하는 기법이라 할 수 있다(Yoo and Kim, 2004).

EOF는 자료의 선형변환을 통해 비상관분포된 자료들을 구성하며 이로부터 원자료 대부분을 설명할 수 있으면서 더 작은 수의 변수들을 얻는 통계적 기법이다. 이는 식 1에 나타난 행렬의 비정칙 분해(Singular Value Decomposition, SVD)를 통해 수행될 수 있다.

$$A = U \times S \times V^T \quad (1)$$

여기서 행렬 U 의 열들을 A 의 EOF라 하고 행렬 S 의 대각요소들을 고유값이라고 한다. 행렬 S 의 고유값들은 EOF가 설명하는 원자료의 분산값이며 모든 EOF의 고유값을 더하면 원자료의 분산이 된다. 행렬 V 의 각 열은 시계열 계수로서 각 EOF의 시간적 전개를 나타낸다.

본 연구에서는 전국 59개 지점의 3SPI 시계열을 행렬 A 로 만든 후 EOF 해석을 수행하였다. 자료기간은 1973년 4월부터 2001년 12월까지로 모든 지점에 대해 동일하다. 3SPI 첫 번째부터 세 번째 EOF의 공간 분포는 그림 1에서 확인할 수 있고 분산공헌도와 EOF coefficient time series는 각각 그림 2와 그림 3에 나타내었다.

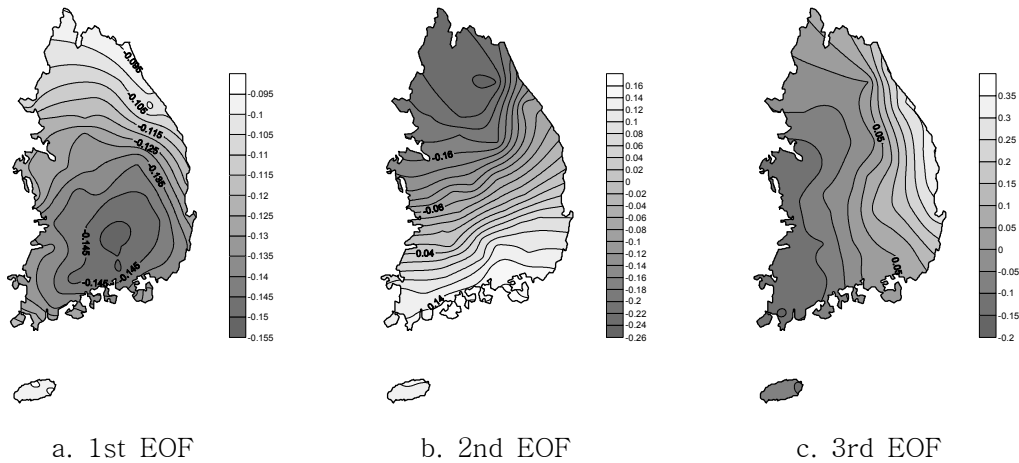


그림 1. EOF의 공간분포

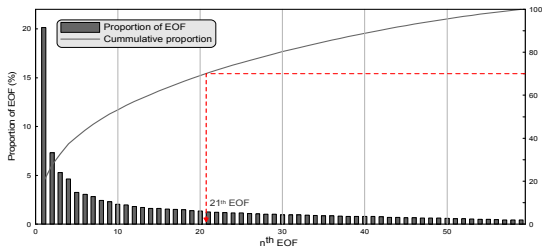


그림 2. EOF의 분산 및 누가분산

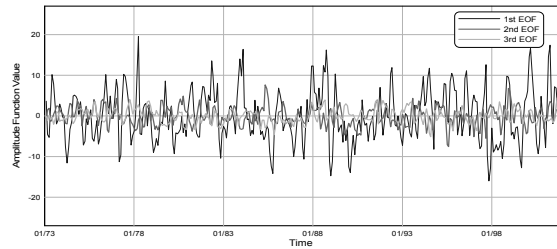


그림 3. EOF Coefficient Time Series

3. 다변량 시계열 모형 (Multivariate Time Series Model)

행렬 S 와 행렬 V 를 곱한 시계열은 EOF의 분산공헌도가 고려된 시계열로 Coefficient Time Series라 불린다. 첫 번째부터 세 번째 EOF의 Coefficient Time Series는 Fig. 3에서 확인할 수 있다. 본 연구에서는 다변량 시계열 모형중 하나인 VARMA 모형을 이용하여 시간적 자료 확장에 적용하였다. Minimum Information Criteria(MIC) 기법을 이용하여 모형을 식별하였고 그 결과는 표 1과 같다. VARMA(1,0), 즉, VAR(1)과정이 가장 작은 Information Criteria임을 알 수 있다.

표 1. VARMA 모형의 식별

Lag	MA 0	MA 1	MA 2	MA 3	MA 4	MA 5
AR 0	-30.606	-30.598	-27.451	-22.349	-15.663	-6.384
AR 1	-41.889	-37.665	-33.194	-26.322	-17.057	-3.508
AR 2	-39.719	-34.163	-27.475	-17.993	-4.436	16.142
AR 3	-35.143	-27.746	-18.161	-4.082	16.773	53.005
AR 4	-27.999	-18.018	-4.141	16.748	53.612	125.636
AR 5	-18.410	-4.214	16.913	53.013	125.467	316.542

VAR(1) 과정은 Eq. 2와 같은 식으로 표현할 수 있다.

$$y_t = \delta + \Phi_1 y_{t-1} + \epsilon_t \quad (2)$$

여기서 δ 는 상수행렬 이고 Φ_1 은 AR1 매개변수 행렬, ϵ_t 는 공분산 행렬이 Σ 인 백색잡음과정 행렬이다. 모형으로 생성된 자료와 실자료는 Fig. 4에 나타나 있으며 음영은 95% 신뢰구간을 나타낸다.

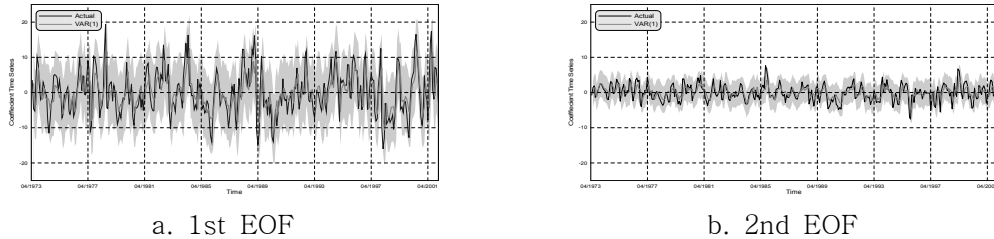


Fig 4. Actual and generated EOF Coefficient Time Series(1st-2th EOF)

4. SPI의 생성 및 재현기간별 가뭄심도의 결정

VAR(1) 모형에서 추정된 AR(1) 과정의 매개변수 행렬 Φ_1 과 백색잡음의 공분산 행렬 Σ 를 이용하면 VAR(1) 과정의 Coefficient Time Series를 임의적으로 생성할 수 있다. 또한 생성된 시계열에 원자료의 EOF 행렬 U 를 곱하면 각 지점의 3개월 SPI자료 역시 생성할 수 있다. 본 연구에서는 EOF Coefficient Time Series를 VAR(1)과정의 120,000개월 기간으로 생성하였으며 이는 10,000년 자료에 해당한다. 또한 생성된 시계열에 EOF를 곱하여 29년의 3SPI 자료를 10,000년으로 확장하였다. Fig. 5는 속초지점의 생성된 3SPI 자료를 나타낸다.

재현기간이 2년일 때 120,000개월의 자료 계열을 24개월(2년) 단위로 끊어서 최대치를 산정할 경우 5,000개의 자료계열이 생성되게 된다. 마찬가지로 재현기간이 5년일 경우 이에 대한 최대치 자료계열은 2,000개의 자료가 생기며 이러한 방법으로 재현기간을 증가시키면서 각 재현기간에 대한 최대치 자료계열을 구성하였다. 각 재현기간에 대한 자료계열은 정규분포보다는 대수정규분포 형태로 나타나는 것으로 판단되어 그 평균의 추정에 대수정규분포를 고려하였다. 그 결과는 Table 2와 같다.

표 2. 재현기간에 대한 가뭄심도

Return Period	2 years	5 years	10 years	20 years	30 years	50 years	100 years	200 years	300 years	500 years
Drought severity	-1.400	-1.813	-2.076	-2.313	-2.445	-2.602	-2.803	-2.994	-3.123	-3.266

5. 가뭄심도에 대한 강수부족량 산정과 구역별 농업가뭄 대비능력 평가

SPI는 강수량 자료를 기초로 산출되므로 SPI 가뭄지수를 이용한 심도는 강수부족량으로 환산될 수 있다. 3, 4, 5, 6월이 농업가뭄이 발생하는 주요한 달이라고 판단하였을 경우 전국 평균 3, 4, 5월 이동평균 강수량은 대수정규분포를 사용하였을 때 79.83mm, 4, 5, 6월의 경우 119.08mm이다(이 결과는 분석에 사용된 59개 지점의 평균 월강수량 자료로부터 산정되었다). 30년 빈도 가뭄이 발생하였을 경우 3SPI의 전국평균심도는 -2.445 정도라 할 수 있는데 SPI는 표준정규분포를 따르므로 이에 대한 표준정규분포의 누가확률은 0.007243이다. 아울러 누가확률 0.007243에 대한

3, 4, 5월 및 4, 5, 6월 59개 지점의 전국평균 이동평균강수량은 31.32mm, 61.64mm 이다. 따라서 3, 4, 5월 4, 5, 6월 1개월 평균 강수부족량은 각각 48.43mm, 57.43mm라 할 수 있다.

한편, 우리나라의 유역별 주요 댐 유효저수량 및 유역면적, 농경지 면적은 Table 3와 같다. 이에 댐 유효저수량을 농경지 면적으로 나누면 생활용수, 공업용수 등 농업용수를 제외한 용수사용을 모두 배제하였을 때 농경지에 공급할 수 있는 최대 공급수량 깊이의 의미를 갖는다고 할 수 있다.

표 3. 대권역별 유역면적, 댐유효저수량, 농경지면적 및 최대공급가능 수량깊이

대권역	유역면적 (km ²)	댐유효저수량(10 ⁶ m ³)	농경지면적(ha)	최대공급가능수량(mm)
한 강	23,292.83	2819.69	243,740	115.68
낙동강	23,702.02	2648.32	388,148	68.23
금 강	9,914.01	1010.10	204,565	49.38
섬진강	4,914.32	607.30	83,510	72.72
영산강	3,469.58	171.00	107,115	15.96

3, 4, 5월 및 4, 5, 6월 1개월 전국평균 강수부족량은 각각 48.43mm 및 57.43mm 이므로 Table 3에 의하면 30년 빈도의 가뭄이 발생하였을 경우 한강, 낙동강, 섬진강 유역은 어느 정도 대비할 수 있으나 금강, 영산강의 경우는 물 부족이 발생할 수 있다. 또한 같은 심도가 2개월 지속되었을 경우 대비할 수 있는 유역은 한강유역이 유일하며 한강 나머지 유역은 적절한 추가의 이수대책이 필요해 짐을 알 수 있다. 특히 영산강 유역의 경우 유역의 30% 정도가 농경지로 이루어져 있는데 반해 농업용수를 공급할 수 있는 저수용량이 상당히 부족함을 쉽게 확인할 수 있다.

6. 결론

본 연구에서는 전국 59개 지점의 3개월 SPI 자료를 EOF 해석 및 다변량 시계열 모형에 적용하여 재현기간별 최대가뭄심도를 결정하였다. EOF 해석에 의해 나타난 Coefficient Time Series를 다변량 시계열 모형에 적용하여 SPI 시계열의 시간적 확장을 수행하였다. 자료기간을 10000년으로 확장하였고 전국적인 가뭄심도를 판단하기 위해 전국 평균 지수를 이용하여 재현기간별 최대심도를 결정할 수 있었다. 이와 더불어 각 대권역의 댐 유효저수량과 농경지 면적을 이용하여 농업가뭄 대비능력을 판단하였는데 재현기간 30년 가뭄에 적절히 대비할 수 있는 이수능력을 갖춘 유역은 한강유역이 유일한 것으로 나타났다. 특히 영산강 유역은 큰 농경지 면적에 비해 저수용량이 크게 부족하여 강우량의 크기에 민감한 농업가뭄에 매우 취약할 것으로 판단되었다.

참고문헌

- 경기개발연구원 (2004). 경기북부 농업가뭄현황 감시기법 연구.
 유철상, 안재현, 류소라 (2004). "포아송과정을 이용한 가뭄의 공간분포 분석." 한국수자원학회논문집, 제37권, 제10호, pp. 813-822
 Whilhite, D.A. and Glantz, M.H. (1985). "Understanding the drought phenomenon: The Role of definition." *Water International*, Vol. 10, pp. 111-120.
 Yoo, C. and Kim, S. (2004). "EOF analysis of surface soil moisture field variability." *Advanced in Water Resources*, Vol. 27, pp. 831-842.