

군집분석과 이변량 확률분포를 이용한 가뭄빈도해석

Drought Frequency Analysis Using Cluster Analysis and Bivariate Probability Distribution

유지영* · 김태웅** · 김상단***

Yoo, Ji Young · Kim, Tae-Woong · Kim, Sangdan

Abstract

Due to the short period of precipitation data in Korea, the uncertainty of drought analysis is inevitable from a point frequency analysis. So it is desired to introduce a regional drought frequency analysis. This study first extracted drought characteristics from 3-month and 12-month moving average rainfalls which represent short and long-term droughts, respectively. Then, the homogeneous regions were distinguished by performing a principal component analysis and cluster analysis. The Korean peninsula was classified into five regions based on drought characteristics. Finally, this study applied the bivariate frequency analysis using a kernel density function to quantify the regionalized drought characteristics. Based on the bivariate drought frequency curves, the drought severities of five regions were evaluated for durations of 2, 5, 10, and 20 months, and return periods of 5, 10, 20, 50, and 100 years. As a result, the largest severity of drought was occurred in the Lower Geum River basin, in the Youngsan River basin, and over in the southern coast of Korea.

Keywords : *drought frequency analysis, cluster analysis, regional drought characteristics*

요 지

우리나라의 경우 짧은 강우 관측 기간으로 인해 지점빈도해석에 의한 가뭄분석은 불확실성이 크다는 문제가 있어 가뭄에 대한 지역빈도해석의 적용이 필요하다. 본 연구에서는 강우의 경향이나 크기를 고려하기 위해 단기가뭄을 나타내어 줄 수 있는 3개월 이동평균강우량과 장기가뭄을 나타내어 줄 수 있는 12개월 이동평균강우량을 산정한 후, 가뭄분석을 수행하기 위한 가뭄특성변수를 추출하였다. 가뭄특성변수를 이용하여 주성분분석과 군집분석을 수행하여 가뭄의 동질성을 갖는 관측지점들을 구분하였다. 또한, 본 연구에서는 지역별 가뭄빈도해석을 위해 이변량 확률분포함수를 적용하였으며, 가뭄 특성(가뭄 지속기간과 심도)의 상호 관계를 고려하여 지역적 가뭄특성을 종합적으로 판단하였다. 또한 이변량 핵밀도함수의 적용을 통해 가뭄 발생의 분포 및 경향성을 가장 근접하게 나타내어 줄 수 있는 결합 확률밀도함수를 추정하고, 군집지역별 2개월, 5개월, 10개월, 20개월의 가뭄지속기간과 5년, 10년, 20년, 50년, 100년의 재현기간에 따른 지역적 가뭄특성을 분석하였다. 그 결과 금강하류, 영산강의 일부 권역 및 남해안 일대에서 상대적으로 큰 가뭄심도가 발생하는 것으로 나타났다.

핵심용어 : 가뭄빈도해석, 군집분석, 지역적 가뭄특성

1. 서 론

최근 지구온난화가 가속화되면서 전 세계적으로 기상재해가 급증하고 있다. 특히 강우패턴의 변화를 고려한 강수 전망 연구는 온실가스 농도 증가로 호우나 가뭄, 대설 등이 지역에 따라 서로 상반되는 변화를 가져올 가능성이 있다고 밝혔다(국립기상연구소, 2008). 또한 우리나라에 호우현상이 빈번해질 것으로 예상하였으나, 8·9월에 강수증가가 뚜렷한 반면 기후학적으로 건조한 시기인 겨울과 봄을 비롯하여 초여름까지 강수량 증가가 미미하거나 감소할 것으로 나타났다(국립기상연구소, 2009).

일반적으로 가뭄의 해석에는 그 목적에 따라 여러 가지 지표를 이용하여 가뭄을 정의하여 왔다. 그 중 강수 및 하천유량 등은 기상 및 수문학적 가뭄을 판단하기 위한 지표로 널리 사용되어 지고 있다. 특히 강수의 부족은 가뭄의 주된 요인이라 할 수 있으며, 가뭄의 정량적 평가에 효과적으로 이용될 수 있다. 즉 평균수준(혹은 절단수준)을 설정하고 가뭄의 지속기간, 심도, 발생빈도 등을 정의한 후, 이를 시계열 분석하여 가뭄의 특성을 분석하는 것이다(Yevjevich, 1967).

현재의 가뭄 해석 방법은 현실적으로 두가지의 한계를 가지고 있다. 첫째, 가뭄은 홍수와 달리 진행속도가 느리고,

*정희원 · 한양대학교 대학원 건설환경공학과 박사과정 (E-mail : 7924pooh@hanyang.ac.kr)

**정희원 · 교신저자 · 한양대학교 건설환경공학과 조교수 · 공학박사 (E-mail : twkim72@hanyang.ac.kr)

***정희원 · 부경대학교 환경시스템공학부 부교수 · 공학박사 (E-mail : skim@pknu.ac.kr)

시·공간적으로 매우 비균질하게 발생하므로(Salas *et al.*, 2005), 가뭄을 분석하고 평가하는 데 있어 공간적 특성을 고려하는 분석이 필요하다. 둘째, 가뭄은 지속기간과 심도를 주요 특성변수를 가지는 이변량 수문사상이므로(Kim *et al.*, 2003), 이를 반영한 확률 및 통계학적 해석방법의 적용이 필요하다.

가뭄의 공간적 특성과 관련한 연구는 지난 20년 동안 지속적으로 활발히 이루어지고 있다. 윤용남 등(1997)은 팔머 가뭄지수(PDSI)를 이용하여 우리나라의 대표적인 가뭄사상을 대상으로 하여 지속기간에 따른 전국 가뭄지수도를 작성함으로써 가뭄의 공간적 특성을 연구한 바 있다. 경기개발연구원(2004)과 김보경 등(2006)은 가뭄의 공간적인 분포특성을 고려하여 가뭄심도-면적-지속기간 곡선을 유도하여 가뭄 분석을 수행하였다. 경민수 등(2007)은 표준강수지수(SPI)의 지점간 유사성을 기준으로 하여 군집분석을 실시한 후, 군집별 EOF 분석과 Kriging 기법을 적용하여 가뭄심도가 가장 큰 지역을 구분하였다. 이외에도 우리나라 가뭄의 공간 분포와 관련한 다양한 연구사례는 존재하나, 지역별로 상이하게 나타나는 가뭄발생확률에 대한 가뭄지역빈도해석에 관한 연구는 활발히 진행되지 못한 실정이다.

최근 들어 수문사상의 해석 및 평가방법으로 이변량 빈도 해석 모형을 이용한 선행연구들이 활발히 이루어지고 있다. Yue *et al.*(1999), Yue(2000) 그리고 Yue and Rasmussen(2002) 등은 Gumbel 혼합모형을 이용하여 각 요소들 간의 이변량 결합 분포 및 결합 재현기간, 조건부 결합 재현기간 등과 같은 강우사상과 홍수사상의 서로 상관된 요소들 간의 결합 확률 거동을 분석하기 위한 방안을 제시하였다. 또한 Zhang and Singh(2006)은 Copula 방법을 이용하여 이변량 결합분포를 추정함으로써 기존 이변량 빈도 해석법의 한계를 완화시키는 방안에 관련한 연구를 수행하였으며, Kao and Govindaraju(2007)는 여러 가지의 극치 주변분포를 바탕으로 Copula 모형을 적용하여 동일한 강우 자료에 대해 빈도 해석을 통해 해당 지점에 가장 적합한 모형을 선정하는 방안을 제시하였다. Kim *et al.*(2003)은 이변량 핵밀도함수를 이용하여 비매개변수적인 접근을 통해 가뭄의 결합 재현기간을 추정하는 방안을 제시하였다. 연제문 등(2007)은 표준강수지수(SPI)를 이용하여 가뭄의 심도 및 지속기간을 이변수로 결정한 후 서울지역의 이차원 가뭄빈도해석 통한 가뭄 평가를 수행하였다. 특히 연제문 등(2007)은 가뭄빈도 해석에 이용될 강수자료 부족의 한계를 해결하기 위해 기록된 측우기 자료를 추가로 이용하여 분석자료 수의 길이를 늘이기 위한 노력을 수행하였다.

우리나라의 기상청 산하의 강수자료는 기록연한이 짧아 지점빈도해석을 수행하는 데 큰 불확실성을 야기하므로, 가뭄의 동질성을 갖는 지점가뭄사상들의 권역화가 선행될 필요가 있다(Hosking and Wallis, 1993). 또한, 기존의 단변량 가뭄빈도해석은 가뭄의 심도와 지속기간 각각에 따른 빈도 해석결과가 서로 상이하게 나타나는 문제를 가지고 있으므로(연제문 등, 2007), 이변량 빈도해석 모형을 가뭄빈도해석에 적용하는 것이 타당하다. 따라서, 본 연구에서는 가뭄의 단변량 지점빈도해석의 한계를 극복하기 위하여, 가뭄의 군집분석과 이변량 확률분포함수를 이용한 가뭄빈도해석을 수

행하였다. 이는 같은 지역적 동질성을 바탕으로 가뭄 특성의 상호 관계를 충분히 고려하며, 보다 일반적으로 가뭄 분석결과를 정량화 할 수 있을 것이다.

2. 가뭄특성인자 구축

2.1 가뭄사상을 통한 다변량 분석인자 구축

우리나라 기상청 산하의 강우관측 지점자료 중 30년 이상의 강우자료를 보유한 58개의 관측지점(그림 1 참고)의 월평균강우량을 이용하여, 단기가뭄사상과 장기가뭄사상을 구축하기 위해 3개월, 12개월 이동평균강우량을 산정하였다. 표준강수지수(Standardized Precipitation Index, SPI)는 보통 지속시간별 이동평균강우량을 이용하여 각 시간단위에 따라 강수 부족량을 측정하여 가뭄을 정량적으로 표현할 수 있다. 이처럼 특정 시간단위로 산정된 SPI는 각 시간단위에 따라 여러 분야에서 사용할 수 있다는 장점이 있다. 예를 들어 단기간의 시간 축척은 농업분야에서 사용되며, 비교적 장기간의 시간 축척은 수자원공급관리 등에서 사용되어지는 특성이 있다(장연규 등, 2006). 이에 본 연구에서는 가뭄을 정의하기 위해 3개월, 12개월 이동평균강우량의 강우시계열에 대한 각각의 연평균 강우량을 가뭄 절단수준으로 결정하고, 연평균 강우량보다 적은 강우가 내렸을 경우의 강우지속기간을 가뭄지속기간이라고 하며, 이때의 강우부족량을 가뭄심도로 정의하여 분석하였다. 또한 단기가뭄 및 장기가뭄에 따른 강우관측소별 가뭄발생특성을 고려할 수 있는 가뭄사상의 지속기간과 심도의 축척 매개변수(scale parameter)와 위치 매개변수(location parameter), 가뭄 지속기간과 심도의 상관계수, 가뭄의 4가지 상태전환확률을 산정하였고, 강우관측소 지점별 지형학적 특징을 고려할 수 있는 강우관측소의 위도·경도 좌표의 자료를 구축하였다. 그리하여 가뭄특성의 동질지역을 구분하는 데 이용되어질 총 11가지 다변량 가뭄인자를 구축하였다.

2.2 주성분 분석 결과

주성분 분석(Principal Components Analysis)은 여러 가지 다변량의 자료의 변동특성을 종합적으로 표현해 줄 수 있다는 장점이 있다. 본 연구에서는 가뭄분석 시 이용되어질 11가지 다변량 가뭄인자의 특성을 분석하여 주성분을 추출하였다. 주성분 분석의 방법에는 분산·공분산 행렬에 의한 방법과 상관행렬에 의한 방법 등이 있으며, 방법에 따라 주성분 분석 결과가 상이하게 나타난다. 본 연구에서는 다변량 가뭄인자의 단위가 서로 다르다는 점을 고려하여 표본상관계수행렬에 의한 방법을 선택하여 다변량 인자 간의 단위가 차이는 문제를 해결하였다. SPSS(Statistical Package for the Social Sciences)를 이용하여 표본상관계수행렬에 의한 방법의 특성인 각 변수들의 측정단위의 차이를 없애고 자료를 표준화(평균을 0, 분산을 1로)하는 분석을 실시하였다. 그 결과 기여율과 누적기여율이 산정되며 이를 해석하여 주성분의 수를 결정하였다. 일반적으로 주성분의 수를 결정하는데 있어 누적기여율이 70% 이상이면 채택이 가능하다(노형진, 2001). 표 1의 결과와 같이 제1주성분~제3주성분에서 누적기여율이 78.62%인 것으로 계산되었으며, 전체 11개의

표 1. 주성분 수 결정을 위한 기여율, 누적기여율 산정 결과

성분	기여율(%)	누적이여율(%)
1	43.126	43.126
2	20.341	63.467
3	15.153	78.620
4	8.361	86.981
5	6.637	93.618
6	2.981	96.599
7	2.704	99.304
8	0.336	99.669
9	0.331	100.000
10	0.000	100.000
11	0.000	100.000

변수에 대한 주성분의 수를 3개로 결정하였다.

주성분을 결정한 후 각 개체에 대한 주성분 점수를 계산하고 이를 이용하여 주성분 변수간 관계를 파악하였다. z_{rp} 는 r 번째 개체의 j 번째 주성분 점수이며, 아래 식 (1)은 설명변수 p 개를 이용하여 첫 번째 개체의 제1주성분부터 제3주성분까지의 점수를 산정하는 데 이용된다. 본 연구에서는 강우관측소별 가뭄의 주성분 특성분석에 목적이 있으므로 위에서의 언급한 개체란 58개의 강우관측소를 의미하고, 설명변수(p)는 11가지 다변량 가뭄인자를 의미한다.

$$z_{11} = a_{11} \left(\frac{x_{11} - \bar{x}_1}{\sqrt{s_{11}}} \right) + a_{12} \left(\frac{x_{21} - \bar{x}_2}{\sqrt{s_{22}}} \right) + \dots + a_{1p} \left(\frac{x_{p1} - \bar{x}_p}{\sqrt{s_{pp}}} \right) \quad (1)$$

$$z_{12} = a_{11} \left(\frac{x_{12} - \bar{x}_1}{\sqrt{s_{11}}} \right) + a_{12} \left(\frac{x_{22} - \bar{x}_2}{\sqrt{s_{22}}} \right) + \dots + a_{1p} \left(\frac{x_{p2} - \bar{x}_p}{\sqrt{s_{pp}}} \right)$$

$$\vdots$$

$$z_{ij} = a_{11} \left(\frac{x_{1j} - \bar{x}_1}{\sqrt{s_{11}}} \right) + a_{12} \left(\frac{x_{2j} - \bar{x}_2}{\sqrt{s_{22}}} \right) + \dots + a_{1p} \left(\frac{x_{pj} - \bar{x}_p}{\sqrt{s_{pp}}} \right)$$

여기서 a_{11}, \dots, a_{1p} 는 첫 번째 개체의 주성분 점수 산정에 이용되는 고유벡터로서 다변량(11가지 설명변수)의 상관행렬로부터 통계프로그램(SPSS)에 의해 산정되며, x_{1j}, x_{2j} 는 첫 번째 개체의 j 주성분까지의 각 설명변수 값이다. 또한, $x_1,$

x_2 는 전체 r 번째 개체의 설명변수 평균값이고, s_{11}, s_{22} 는 설명변수의 표준화 절차에서 이용되는 분산이며, 분산의 차수는 원 설명변수(p)의 개수와 같다(신재경과 문승호, 1996).

그 결과 표 2와 같이 제1주성분, 제2주성분, 제3주성분의 주성분 순위 및 득점을 계산하였으며, 여기서 제1주성분은 P_{DD}(상태전환확률), location_D(가뭄 지속기간의 위치매개변수), location_S(가뭄 심도의 위치매개변수) 등의 순서로 각각의 변수에 큰 득점이 부여되어 있음을 알 수 있다. 여기서 주성분 득점의 크기 순서로 결정되어진 주성분 순위에 따라, 제1주성분이 뜻하는 전반적 의미는 각 변수들의 합축된 의미로 포괄하여 해석할 수 있음을 알 수 있다. 그러므로 제2주성분과 제3주성분도 같은 방법으로 주성분 내 변수들의 합축된 의미로 해석할 수 있다. 또한 주성분 분석의 결과는 군집분석(가뭄동질특성 지역구분) 수행 시 기준변수(제1주성분, 제2주성분, 제3주성분)로 이용되어진다.

3. 군집분석을 통한 가뭄특성의 동질지역 구분

3.1 군집의 수 결정

우리나라는 강우관측기간이 매우 짧고 관측지점마다 강우 기록년 수의 차이가 심하게 나타나는 특징이 있기 때문에 지점별 강우자료를 이용하여 수문학적 해석을 실시할 경우 불확실성은 커지게 된다. 그러므로 현재 연구동향은 과거에 활발히 수행되어진 지점빈도해석과 관련한 연구는 위와 같은 지점빈도해석의 단점을 보완할 수 있는 지역빈도해석을 위한 연구가 활발히 진행되고 있는 실정이다. 지역빈도해석에서 가장 중요한 단계는 동질지역을 구분하는 것이며, 본 연구에서는 다변량 분석기법 중 주성분 분석을 통한 군집분석 시 이용되어질 군집기준변수(제1주성분, 제2주성분, 제3주성분)를 산정하였으며, 이를 이용하여 군집분석 기법 중 K-means 기법을 활용하여 군집분석을 수행하였다.

K-means 기법은 극단치의 영향이 적은 비계층적 군집방법 중 하나로써, K는 미리 정하는 군집의 수이다. 본 연구에서는 군집의 수를 CCC(Cubic Clustering Criterion) 값과 Pseudo-T²의 검정통계량을 동시에 고려하여 5개로 결정하였다. 여기서 CCC 값은 Searle(1983)이 제안한 방법으로 군

표 2. 주성분의 순위 및 득점

순위	제1주성분	득점	제2주성분	득점	제3주성분	득점
1	P _{DD}	0.225	P _{ww}	0.424	corr_DS	0.496
2	location_D	0.224	longitude	0.190	scale_S	0.208
3	location_S	0.181	scale_D	0.087	scale_D	0.142
4	scale_D	0.159	corr_DS	0.086	P _{Dw}	0.063
5	scale_S	0.119	P _{DD}	0.063	P _{ww}	0.027
6	latitude	0.048	latitude	0.056	longitude	-0.007
7	longitude	0.022	location_D	0.006	location_S	-0.018
8	P _{ww}	0.015	location_S	-0.062	P _{WD}	-0.027
9	P _{WD}	-0.015	P _{Dw}	-0.063	P _{DD}	-0.063
10	corr_DS	-0.049	scale_D	-0.131	location_D	-0.170
11	P _{Dw}	-0.225	P _{WD}	-0.424	latitude	-0.489

주) 기호설명: 상태전환확률(P_{DD}, P_{ww}, P_{WD}, P_{Dw}), 위치매개변수(location_D: 가뭄지속기간, location_S: 가뭄심도), 축척매개변수(scale_D: 가뭄지속기간, scale_S: 가뭄심도), 위도(latitude), 경도(longitude), 가뭄지속기간과 심도의 상관계수(corr_DS)

집의 개수와 CCC의 산점도를 그려서, CCC의 값이 3 이상 이 될 경우에 그때의 군집의 개수가 적당하다고 판단한다. 또한 Pseudo-T² 검정통계량은 두 집단간의 다변량 평균 차이를 보는 통계량으로서 군집별 이질성이 커지는 순간의 군집의 수를 나타내어 준다.

3.2 군집분석 결과

본 연구에서는 단기가뭍과 장기가뭍의 발생특성을 각각 구

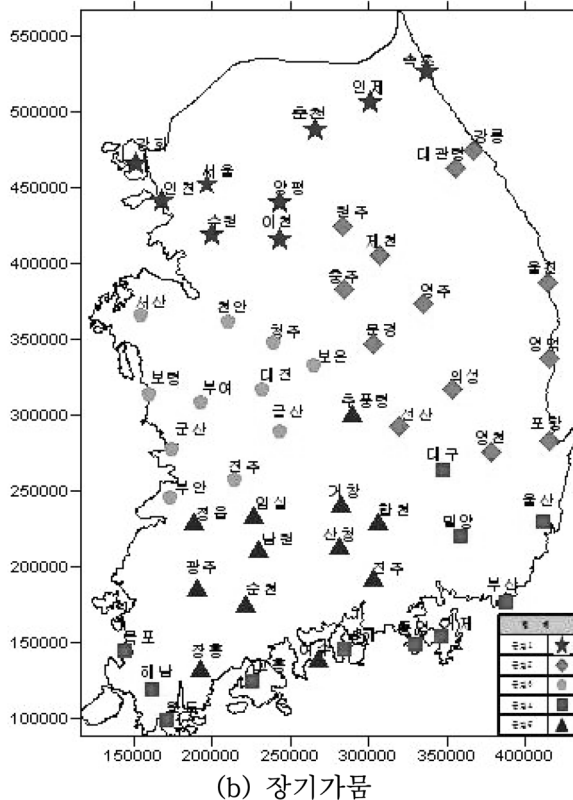
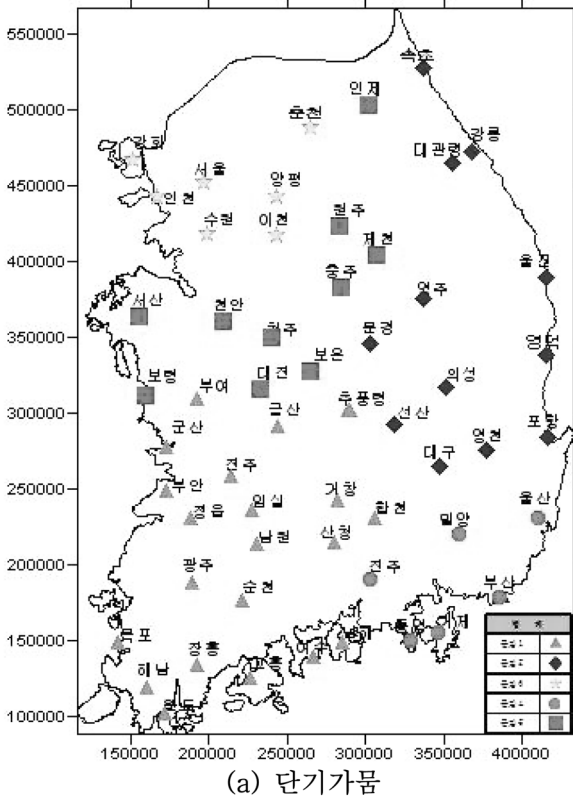


그림 1. 군집분석 결과 (a) 단기가뭍 (b) 장기가뭍

표 3. 최종 군집중심간 거리(단기 · 장기가뭍 군집분석 결과)

	군집 번호	1	2	3	4	5
단기 가뭍	1		2.120	2.613	2.108	1.885
	2	2.120		2.431	1.828	1.715
	3	2.613	2.431		3.413	2.806
	4	2.108	1.828	3.413		2.594
	5	1.885	1.715	2.806	2.594	
장기 가뭍	1		2.124	2.146	3.119	2.511
	2	2.124		1.624	2.401	1.718
	3	2.146	1.624		2.066	1.727
	4	3.119	2.401	2.066		1.721
	5	2.511	1.718	1.727	1.721	

분하여 군집분석 시 이용되어질 군집기준변수를 산정하였으며, 단기가뭍일 경우와 장기가뭍일 경우일 경우에 대한 가뭍 지역이 구분될 수 있도록 군집분석을 수행하였다. 그 결과 그림 1과 같이 각각 5개의 군집이 형성되었으며, 단기가뭍과 장기가뭍에 따른 군집분석결과가 서로 다르게 나타남을 확인할 수 있었다.

그림 1에서 춘천, 인제, 속초지점은 장기가뭍 군집결과에서는 같은 군집으로 묶여 있으나 단기가뭍 군집결과에서는 각각 서로 다른 지역으로 구분되어 지는 것을 확인할 수 있다. 이와 마찬가지로 청주, 제천, 충주, 합천, 전주, 장흥, 대구지점들의 군집결과가 달라짐을 확인할 수 있다. 표 3은 단기·장기가뭍 군집분석 결과로써 최종 군집중심간 거리를 나타내고 있으며, 표 4는 군집별 가뭍사상의 수, 지속기간과 심도의 평균과 표준편차를 나타내고 있다. 단기가뭍과 장기가뭍 분석결과를 살펴보면 군집4(남해안 일대)에서 공통적으로 평균적으로 가장 긴 가뭍 지속기간과 큰 심도가 발생함을 알 수 있고, 군집별 가뭍사상 수를 비교해 보면 단기가뭍일 경우에는 군집1(금강 권역과 영산강 권역)에서 가장 많은 가뭍사상 수가 발견됨을 알 수 있고, 장기가뭍일 경우에는 군집4(남해안 일대)에서 가장 많은 가뭍사상 수가 발견되었음을 확인할 수 있다.

4. 이변량 핵밀도 함수를 이용한 가뭍빈도 해석

4.1 결합 확률밀도함수의 추정

본 연구에서는 가뭍의 지속기간과 심도를 동시에 고려하여 가뭍의 특성을 종합적으로 판단하기 위해 이변량 가뭍빈도해석을 수행하였으며, 이변량 가뭍빈도 해석방법에는 크게 매개변수적 방법(Parametric method)과 비매개변수적 방법(Nonparametric method)로 분류할 수 있다. 이 두 방법은 대상자료의 확률적 거동이 특정 분포형을 따른다고 가정할 것인가의 유무에 따라 구분되어진다. 이변량 가뭍빈도해석을 위해 매개변수적 방법 중 대표적인 Gumbel 혼합모형의 적용성을 검토한 결과, 가뭍 지속기간과 심도의 상관관계수가 임계값 2/3 보다 커서 본 연구에서의 적용이 불가능하였다. 따라서, 본 연구에서는 비매개변수적 방법 중의 하나인 이변량 핵밀도함수를 식 (2)와 같이 적용하였다.

표 4. 군집별 가뭄지속기간 및 심도의 평균, 표준편차, 가뭄사상 수(단기 · 장기가뭄)

	군집 번호	지속기간(months)의 평균	지속기간(months)의 표준편차	심도(mm)의 평균	심도(mm)의 표준편차	가뭄사상의 수
단기 가뭄	1	3.688	3.129	109.268	144.462	1473
	2	3.716	3.171	96.047	113.930	848
	3	3.876	2.906	111.645	127.662	518
	4	4.092	3.476	145.896	178.898	489
	5	3.540	2.890	91.424	113.642	648
장기 가뭄	1	6.087	7.568	81.511	142.306	298
	2	5.840	7.479	77.996	148.513	293
	3	6.668	7.615	106.764	166.630	265
	4	8.486	9.421	141.558	223.149	354
	5	7.176	8.613	130.378	229.339	262

$$\hat{f}_{XY}(x,y) = \frac{1}{nh_x h_y} \sum_{i=1}^n \left\{ K\left(\frac{x-x_i}{h_x}\right) K\left(\frac{y-y_i}{h_y}\right) \right\} \quad (2)$$

여기서 n 은 변량의 개수, h_x 와 h_y 는 각각 x 와 y 변수의 광역폭(Bandwidth)이며, K 는 핵밀도함수(Kernel function)이다. 비매개변수적 핵밀도함수 방법에서 광역폭의 선택은 매우 중요한 문제이며, 지금까지 광역폭을 결정하는 방법은 Maximum Likelihood, Least Squares Cross Validation, Breiman Method(Adamowski, 1985; Lall *et al.*, 1993; Moon and Lall, 1994) 등이 있다. 만약 광역폭이 너무 작게 결정될 경우에는 큰 분산과 거친 추정치를 나타낼 수 있고, 반면 광역폭이 너무 크게 결정될 경우에는 큰 편차와 너무 완만한 밀도함수의 추정으로 인한 정보의 손실을 가져올 수 있다는 문제점이 있다(Kim *et al.*, 2003). 본 연구에서는 최적 광역폭의 산정을 위해 Kim *et al.*(2003)과 같이 식 (3)을 이용하였다.

$$h_{d,opt} = \left[\frac{4}{n(p+2)} \right]^{1/(p+4)} \sigma_d \quad (3)$$

여기서 σ_d 는 d 차원 분포에서의 표준편차이고, p 는 차원으로서는 이변량 빈도분석에서는 분석대상의 변량 수가 2개이기 때문에 p 값은 2로 계산하였다. 그 결과는 그림 2와 3과 같으며, 이중 그림 2는 단기가뭄 분석결과 중 가장 많은 경우관측지점의 수가 군집되어진 군집1의 결합 확률밀도함수를 추정한 결과이며, 그림 3은 장기가뭄 분석결과 중 가장 많은 경우관측지점의 수가 군집되어진 군집2의 결합 확률밀도함수를 추정한 결과를 나타내고 있다. 가뭄의 지속기간과 심도는 물리적으로 양의 범위(≥ 0)에서 정의되었지만, 결합확률분포함수를 보다 유연하고 완만하게 도시하기 위하여 그림 2와 3에서 x 축(Duration)과 y 축(Severity)의 최소값을 -15(months)와 -300(mm)으로 설정하여 도시하였다.

4.2 가뭄지속기간과 재현기간에 따른 가뭄발생 심도 추정

본 연구에서는 비매개변수적 방법을 이용한 이변량 핵밀도함수를 통하여 결합 확률분포함수를 추정하였고, 각각의 재현기간별로 가뭄지속기간에 따른 가뭄심도를 추정하였다. 그림 4는 단기가뭄의 지속기간과 심도에 따른 이변량 재현기간(Bivariate return period, BRP)을 도시하였다. 여기서 그림 4(a)는 군집분석 결과 군집1의 결과이며, 그림 5는 장

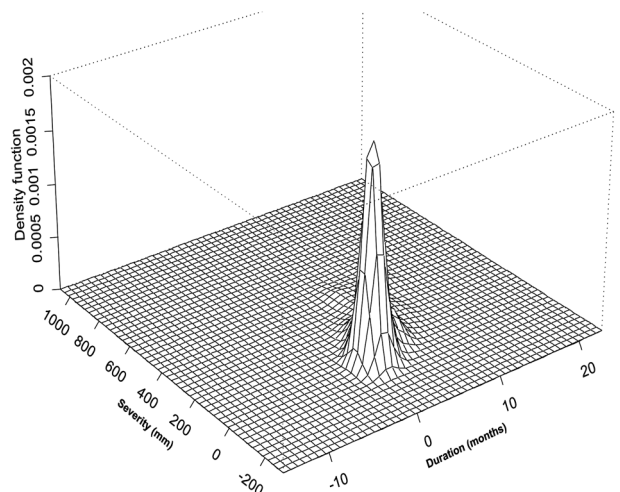


그림 2. 군집1의 결합 확률밀도함수(단기가뭄)

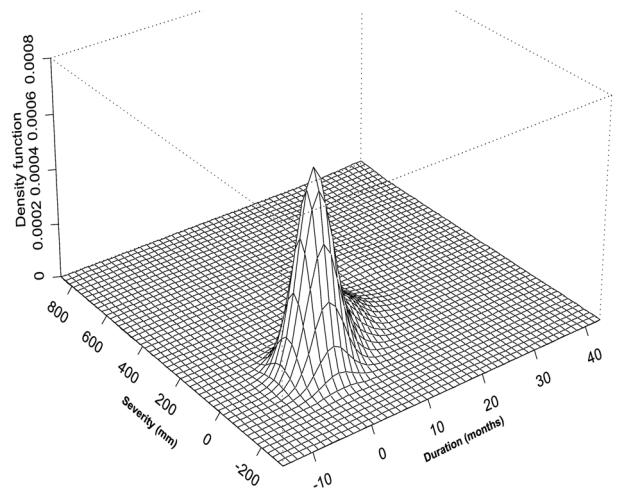
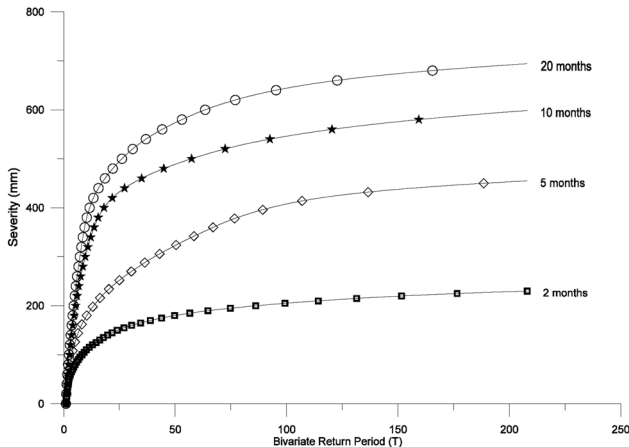
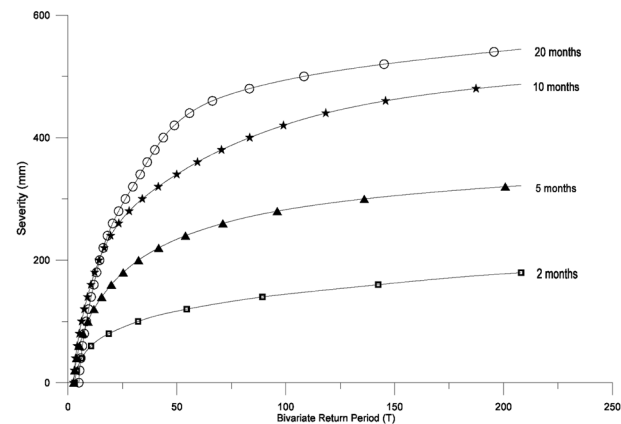


그림 3. 군집2의 결합 확률밀도함수(장기가뭄)

기가뭄의 지속기간과 심도에 따른 이변량 재현기간을 도시하였다. 그림 4(a)와 그림 4(b)를 비교해 보면, 같은 단기가뭄을 분석한 결과임에도 불구하고 군집지역별로 가뭄지속기간과 심도에 따른 재현기간이 상이하게 달라지고 있음을 확인할 수 있다. 이 결과를 재해석하면 가뭄의 발생특성은 지역별로 차이가 있으므로 가뭄빈도분석을 수행할 경우에는 선행연구로 이러한 지역적 가뭄특성을 고려해야 할 것으로 판



(a) 군집1의 BRP



(b) 군집2의 BRP

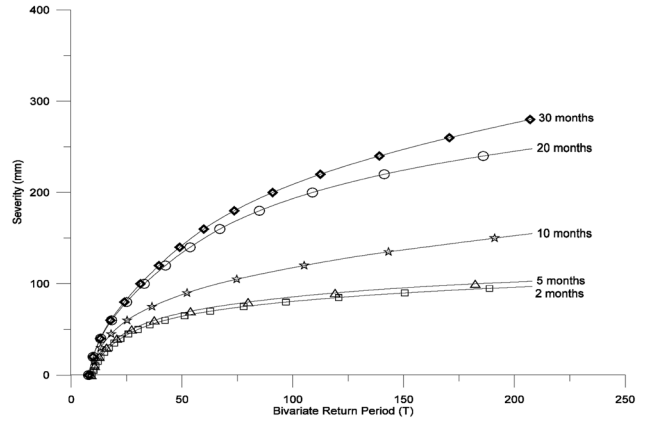
그림 4. 단기가뭍의 이변량 재현기간(Bivariate Return Period)
(a) 군집1 (b) 군집2

단된다.

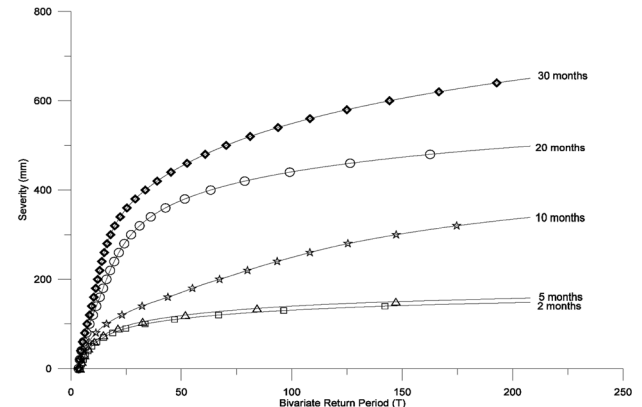
본 연구에서는 2개월, 5개월, 10개월, 20개월의 가뭄지속 기간과 5년, 10년, 20년, 50년, 100년의 재현기간에 따른 각각의 군집별 가뭄심도를 추정하였다. 그 결과 단기가뭍에 해당되는 결과는 표 5와 같고, 장기가뭍에 해당되는 결과는 표 6과 같다. 군집지역별 가뭄심도 추정결과를 살펴보면, 단기가뭍일 경우에는 금강하류 권역과 영산강 권역으로 군집화되어진 군집1에서 가뭄지속기간과 재현기간에 따른 가뭄심도가 가장 크게 나타남을 확인할 수 있는 반면, 한강 권역의 일부인 군집2에서는 상대적으로 작은 가뭄심도가 발생됨을 알 수 있다. 또한 장기가뭍일 경우에는 금강하류 권역과 영산강 권역으로 군집화되어진 군집5와 남해안 일대로 구분되는 군집4에서 상대적으로 큰 가뭄심도가 발생하였으며, 남한강 권역, 영동 권역, 낙동강북부 권역의 일부로 군집화되어진 군집2에서 보다 작은 가뭄심도가 나타났다.

5. 결 론

본 연구에서는 가뭄의 지점빈도 해석방법의 한계를 극복하기 위한 대안으로 가뭄특성이 동질한 지역의 권역화를 통해, 지역빈도해석 시 통계학적으로 분석자료 개수의 증가로 인한 분산의 감소가 커질 수 있다는 점을 착안하여 가뭄빈도



(a) 군집2의 BRP



(b) 군집5의 BRP

그림 5. 장기가뭍의 이변량 재현기간(Bivariate Return Period)
(a) 군집2 (b) 군집5

분석의 효율성과 분산결과의 정확도를 개선하기 위한 노력을 하였다. 그 후 가뭄지역빈도해석을 수행하는데 있어 가뭄의 발생특성거동을 보다 유연하게 해석하기 위해 비매개변수적인 방법을 이용한 이변량 핵밀도함수를 이용하였다.

또한 단기가뭍일 경우와 장기가뭍일 경우를 구분하기 위해 3개월 이동평균강우량과 12개월 이동평균강우량을 이용하여 가뭄분석을 수행하여 다변량 가뭄인자를 산정하고 주성분 분석을 통해 다양한 변량들로 나타내는 가뭄특성을 함축해 줄 수 있는 주성분을 결정하였다. 그 결과 가뭄특성이 동질한 지역의 구분을 위한 기준변수가 구축되었고, 이를 이용하여 군집분석을 수행한 결과 단기가뭍과 장기가뭍일 경우에 따라 몇몇 강우관측지점들의 소속군집이 변화되어짐을 확인할 수 있었다. 이는 지점별 가뭄발생특성이 가뭄의 지속기간에 따라 인근 다른 지점들과 서로 다르게 군집화 되어져 나타난 결과라고 해석할 수 있다.

본 연구는 이변량 가뭄지역빈도해석을 수행하였으며, 우선 군집분석을 수행한 결과 지역별 빈도해석을 보다 신뢰성 있게 수행할 수 있도록 가뭄자료의 길이가 길어져 지점빈도해석의 한계를 완화하였다고 판단된다. 또한 가뭄의 지속기간과 심도의 매개변수적방법을 통해 가뭄분석을 수행한 결과 적합한 확률분포형을 찾을 수 없었던 문제점을 해결하기 위해 비매개변수적 방법을 이용하여 유연한 가뭄의 거동을 분석할 수 있었다. 또한 가뭄해석시 가뭄지속기간 및 심도가

표 5. 가뭄지속기간과 재현기간에 따른 단기가뭄심도 추정결과 (mm)

재현기간	군집 번호	가뭄 지속기간(months)			
		2	5	10	20
5years	1	79.85	124.24	192.38	223.41
	2	33.59	54.79	70.19	74.84
	3	61.87	107.69	153.82	164.48
	4	51.86	86.52	117.42	128.51
	5	30.14	56.37	76.72	79.34
10years	1	109.88	177.98	305.87	373.68
	2	67.21	103.31	141.28	152.67
	3	90.52	169.62	222.68	243.72
	4	86.61	154.61	237.54	267.57
	5	58.52	106.38	153.24	160.69
20years	1	140.79	232.97	488.69	469.25
	2	97.65	147.68	218.73	250.61
	3	127.86	218.75	294.81	337.24
	4	116.37	234.27	333.35	388.63
	5	80.28	160.23	242.28	256.27
50years	1	180.39	322.92	448.69	573.72
	2	141.51	208.95	321.74	371.52
	3	180.82	283.17	329.96	457.34
	4	159.48	329.78	441.51	569.24
	5	116.42	234.38	340.61	423.81
100years	1	205.49	407.58	546.42	644.38
	2	175.39	261.86	379.51	445.67
	3	214.95	334.29	460.27	518.39
	4	195.41	387.41	534.31	659.61
	5	144.72	282.28	421.53	494.34

모두 중요한 변량으로 고려되어야 함으로 단변량 해석이 아닌 이 둘을 모두 고려할 수 있는 이변량 해석을 수행해야 하기 때문에 본 연구에서는 이변량 핵밀도 함수를 이용한 빈도해석을 수행했다. 그 결과 군집지역별 가뭄지속기간과 재현기간의 변화에 따른 가뭄심도가 지역별로 상이함을 알 수 있으며, 이는 가뭄발생의 취약지역을 발견하는데 도움이 될 수 있을 것이다.

가뭄을 정의하는데 있어 본 연구에서는 강우량만을 고려하는 기상학적 가뭄을 분석하였으나, 향후 수행되어야 하는 연구로는 가뭄평가 시 고려되어야 할 인자선정에 있어서 수문학적 가뭄을 분석할 수 있는 인자(하천유량, 저수유량)들을 추가적으로 고려하여야 할 것으로 판단된다. 또한 농업적 가뭄을 나타낼 수 있는 토양수분량과 같은 인자도 추가적으로 고려되어야 할 필요성이 있으며, 이와 같은 많은 인자를 이용하여 가뭄을 분석하는 데 효율성을 보다 높이기 위해서는 다음과 같은 방법을 제안할 수 있을 것이다. 우선 가뭄의 정의에 따라 기상학적 가뭄, 수문학적 가뭄, 농업적 가뭄으로 분류한 다음 이에 해당되는 각각의 인자들을 선정하여, 모든 인자를 이용하여 주성분 분석을 수행한다. 이에 예상할 수 있는 결과는 다변량 인자가 함축되어 표현될 수 있는 주성분을 발견하여, 이 주성분을 이용한 다양한 가뭄정의가 모

표 6. 가뭄지속기간과 재현기간에 따른 장기가뭄심도 추정결과 (mm)

재현기간	군집 번호	가뭄 지속기간(months)				
		2	5	10	20	25
10years	1	43.61	46.17	63.48	112.32	117.38
	2	14.72	16.93	23.87	30.98	32.54
	3	16.92	19.74	25.38	40.52	43.96
	4	37.81	41.85	61.12	98.81	106.55
	5	15.32	56.92	72.71	119.98	139.21
20years	1	61.15	66.15	95.67	194.62	204.74
	2	63.58	38.81	49.82	65.93	71.41
	3	34.84	38.24	49.66	96.82	118.76
	4	74.65	80.69	109.87	180.81	201.75
	5	52.77	87.74	111.75	243.98	295.36
50years	1	79.21	87.19	153.85	286.74	301.91
	2	64.85	68.59	88.61	133.92	158.71
	3	59.96	64.73	86.74	257.81	289.95
	4	109.82	118.44	168.93	285.28	329.74
	5	81.91	118.76	170.66	276.69	419.21
100years	1	91.52	102.39	199.47	342.82	357.24
	2	80.85	85.97	117.88	193.65	246.63
	3	74.96	80.56	121.52	332.76	372.18
	4	133.21	144.35	214.38	360.89	423.61
	5	130.81	139.82	249.32	440.82	497.84

두 고려된 가뭄해석도 가능할 것이라고 생각된다.

감사의 글

이 연구는 소방방재청 자연재해저감기술개발사업 [NEMA-08-NH-05] 연구비 지원으로 수행되었으며 이에 감사드립니다.

참고문헌

- 경기개발연구원(2004) 경기북부 농업가뭄현황 감시기법 연구. 기본 연구 2004-06, pp. 20-24.
- 경민수, 김상단, 김보경, 김형수(2007) 군집분석을 통한 수문학적 가뭄의 가뭄심도-가뭄면적-가뭄지속기간곡선의 작성. **대한토목학회논문집**, 대한토목학회, 제27권, 제3B호, pp. 267-276.
- 국립기상연구소(2008) 기후변화협약대응 지역기후시나리오 활용기술개발(IV). 국립기상연구소, pp. 355.
- 국립기상연구소(2009) 기후변화 이해하기 II-한반도 기후변화: 현재와 미래. 국립기상연구소, pp. 67.
- 김보경, 김상단, 이재수, 김형수(2006) 가뭄의 시공간적 분포 특성 연구: 가뭄심도-가뭄면적-가뭄지속기간 곡선의 작성. **대한토목학회논문집**, 대한토목학회, 제26권, 제1B호, pp. 69-78.
- 노형진(2001) **한글 SPSS 10.0에 의한 조사방법 및 통계분석**. 형설출판사, pp. 437-476.
- 신재경, 문승호(1996) **다변량 통계해석 입문**. 자유아카데미, pp. 110-114.
- 연제문, 변성호, 이정규, 김태웅(2007) 이차원 가뭄빈도해석을 통한 서울지역의 가뭄 평가. **한국수자원학회논문집**, 한국수자원학회, 제40권 4호, pp. 335-343.
- 윤용남, 안재현, 이동률(1997) Palmer의 방법을 이용한 가뭄의 분석. **한국수자원학회논문집**, 한국수자원학회, 제30권 4호,

- pp. 317-326.
- 장연규, 김상단, 최계윤(2006) SPI 기뻐지수의 EOF 분석을 이 용한 기뻐의 시공간적인 특성 연구. *한국수자원학회논문집*, 한국수자원학회, 제39권 8호, pp. 691-702.
- Adamowski, K. (1985) Nonparametric kernel estimation of flood frequency. *Water Resources Research*, Vol. 21, No. 11, pp. 1585-1590.
- Hosking, J.R.M. and Wallis, J.R. (1993) Some statistics useful in regional frequency analysis. *Water Resources Research*, Vol. 29, pp. 271-281.
- Kao, S. and Govindaraju, R.S. (2007) A bivariate frequency analysis of extreme rainfall with implications for design. *Journal of Geophysical Research*, Vol. 112, D13110, doi:10.1029/2007JD008522.
- Kim, T.W., Juan B. Valdes., and Yoo, C.S. (2003) Nonparametric approach for estimating return periods of droughts in arid regions. *Journal of Hydrologic Engineering*. ASCE, pp. 237-246.
- Lall, U., Moon, Y.I., and Bosworth, K. (1993) Kernel Flood frequency estimator: bandwidth selection and kernel choice. *Water Resources Research*, Vol. 29, No. 4, pp. 1003-1015.
- Moon, Y.I. and Lall, U. (1994) Kernel quantile function estimator for flood frequency analysis. *Water Resources Research*, Vol. 30, No. 11, pp. 3095-3103.
- Salas, J.D., Fu, C., Cancelliere, A., Dustin, D., Bode, D., Pineda, A., and Vincent, E. (2005) Characterizing the severity and risk of drought in the Poudre River, Colorado. *Journal of Water Resources Planning and Management*, ASCE, Vol. 131, No. 5, pp. 383-393.
- Searle, J. (1983) *Intentionality: An Essay in the Philosophy of Mind*. Cambridge University Press, Cambridge.
- Yevjevich, V.M. (1967) An objective approach to definitions and investigations of continental hydrologic droughts. *Hydrologic Paper 23*, Colorado State Univ., Fort Collins, CO.
- Yue, S. (2000) The Gumbel mixed model applied to storm frequency analysis, *Water Resource Management*, Vol. 14, pp. 377-389.
- Yue, S., Ouarda, T.B.M.J., Bobe, B., Legendre, P. and Bruneau, P. (1999) The Gumbel mixed model for flood frequency analysis. *Journal of Hydrology*. Vol. 226, No. 1-2, pp. 88-100.
- Yue, S., and Rasmussen, P. (2002) Bivariate frequency analysis: discussion of some useful concepts in hydrological application. *Hydrol Process*, Vol. 16, pp. 2881-2898.
- Zhang, L., and Singh, V.P. (2006) Bivariate flood frequency analysis using the copula method. *Journal of Hydrologic Engineering*. ASCE, Vol. 11, No. 2, pp. 150-164.

(접수일: 2010.7.23/심사일: 2010.9.7/심사완료일: 2010.9.7)