특징선택기법을 이용한 가뭄영향인자 분석

Identifying the Critical Drought Impact Factor with Feature Selection Approaches

김선우* · 김민기** · 정동휘*** · 유도근****

Kim, Seon Woo*, Kim, Min Ki**, Jung, Donghwi***, and Yoo, Do Guen****

Abstract

Drought generally affects a wider area for longer period of time than other natural disasters. Recently, so-called "mega-droughts" of very long durations and with more extensive impacts have been observed across the world. However, few efforts have been devoted to identifying the critical impact factors-such as climate, hydrological and socio-economic factors-for such extreme droughts. In this study, we identify the most critical factor and derive the direct/indirect relationship among nine influences including climatic, hydrological, and socio-economic factors using two feature selection methods. The two feature selection approaches we use are based on the statistical theories of principal component analysis and linear discriminant analysis. The standardized precipitation index (SPI) is used as a drought index. The study area is Boryeong city, South Korea. Drought impact analysis results obtained from the two feature selection approaches are compared and presented by polygon radar charts. The results showed that the low water level of the Boryeong dam is the most critical impact factor.

Key words : Drought Impact Factors, Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), Direct and Indirect Impact Analysis

요 지

일반적으로 가뭄은 다른 자연재해보다 광범위한 지역에 오랜 피해를 입힌다. 최근에는 10년 이상, 수십 년 동안 지속되며 전국 범위에 피해를 끼칠 수 있는 “메가가뭄”의 발생 가능성이 인지되고 있다. 하지만, 이러한 극한가뭄의 결과에 영향을 미치는 기상학적, 수문학적, 사회·경제적 영향요인에 대한 정량적 분석이 미비하였다. 본 연구는 총 9개의 기상학적(강수량, 평균기온, 평균풍속, 상대습도, 평균기압, 일조율), 수문학적(저수위, 댐방류량), 사회·경제적 인자(생활용수사용량)와 가뭄지수 간의 상호관계를 정량화하여 가뭄영향인자를 도출하고 직·간접적 영향분석을 수행하였다. 통계학적 기법 기반 특징선택 모형인 선형판별분석과 주성분분석을 사용하였으며, 가뭄의 심도는 표준강수지수(Standardized Precipitation Index, SPI)를 이용하였다. 연구 대상 지역으로 충청남도 보령시를 선정하였다. 특징선택기법별, 이용 가뭄지수 종류별 결과를 방사형 차트(Radar chart)를 이용하여 도시 및 비교하였다. 선형판별분석과 주성분분석을 각각 적용하여, 보령댐 저수위가 가뭄에 가장 큰 영향을 주는 인자임을 확인하였다.

핵심용어 : 가뭄영향인자, 주성분분석, 선형판별분석, 직·간접적 영향분석

*정회원, 계명대학교 토목공학과 학부과정(E-mail: pen100kr@naver.com)

Member, Undergraduate student, Department of Civil Engineering, Keimyung University

**계명대학교 토목공학과 학부과정(E-mail: alsrl3304@gmail.com)

Undergraduate student, Department of Civil Engineering, Keimyung University

***교신저자, 정회원, 계명대학교 토목공학과 조교수(Tel: +82-53-580-5706, Fax: +82-53-580-5706, E-mail: donghwiu@gmail.com)

Corresponding Author, Member, Assistant professor, Department of Civil Engineering, Keimyung University

****정회원, 수원대학교 건설환경에너지공학부 조교수(E-mail: dgyoo411@suwon.ac.kr)

Member, Assistant Professor, School of Civil and Environmental Engineering, The University of Suwon

1. 서 론

최근 지구온난화에 따른 기후시스템의 변화로 과거에 경험하지 못했던 기온과 강수패턴이 발생되고 있다(Hong et al., 2016). 이로 인해 기온, 강수패턴의 계절적 및 지역적 편차가 커지면서 강수의 집중, 부족으로 극한홍수 및 극한가뭄으로 인한 피해규모, 빈도가 증가하고 있다. 극한홍수의 경우, 단기적 피해 형태로 발생함에 따라 대책 수립이 상대적으로 용이한 반면에, 극한가뭄은 그 발생 예측이 어렵고, 광범위한 지역에 장시간에 걸쳐 발생한다. 따라서, 이러한 극한가뭄 대비를 위한 관련 연구나 대책 수립이 필요한 실정이다(Kim et al., 2015).

가뭄이란 강수량이나 물공급 부족으로 충분한 용수가 확보되지 못하는 현상을 말하며, 다른 자연재해에 비해 오랜 기간 지속되며 사회·경제적 피해도 야기한다(Choi et al., 2000; Lee, 2012). 국내에서 최근 기후변화로 인해 가뭄의 발생빈도, 지속시간, 피해규모가 증가하고 있는 추세이다(Choi et al., 2017). 따라서, 10년 이상, 수십 년 동안 지속되며 전국 범위에 피해를 끼칠 수 있는 “메가가뭄”의 발생 가능성을 인지하고, 이에 대한 대비 및 대응 기술 마련이 필요하다. “메가가뭄”에 대한 연구는, 기존에 발생한 일반 가뭄에 대한 원인, 전개과정, 피해 및 복구 규모를 규명 및 정립하는 것으로부터 시작될 수 있다.

현재까지 국내에서는 수행된 가뭄 관련 연구는 (1) 가뭄지수 개발(Kim et al., 2006; Lee et al., 2006), (2) 가뭄 위험도 분석(Nam et al., 2013; Kim et al., 2015), (3) 가뭄 예측(Lee et al., 2013; Park et al., 2014), (4) 가뭄 지도 개발(Chang and Park, 2004; Park et al., 2012), (5) 가뭄 시 용수배분의 우선순위 결정(Lee and Shim, 2002; Lee et al., 2004), (6) 가뭄 예·경보 시스템 구축(Yi and Yoo, 2003; Yun et al., 2018) 으로 나뉜다. 과거 가뭄 정보를 기반으로 발생 가능한 가뭄을 예측 및 그 위험도 분석에 관한 연구들이 주로 수행되었다. 이러한 연구에서 사용되는 자료의 종류에 따라 결과의 정확성, 신뢰성이 크게 달라지는 것을 확인하였다(Yoo et al., 2013). 따라서, 과거 가뭄 데이터를 활용한 가뭄 피해 예측, 파악, 위험도 분석, 모니터링 등에 관한 연구들을 수행하기에 앞서, 가뭄에 영향력이 큰 인자(가뭄 주요영향인자)를 선별하고, 정량화하기 위한 연구가 필요하다.

한편, 과거 정보를 기반으로 수행된 가뭄 연구의 경우, 기상학적 요인(Ryoo and Yoo, 2004; Park and Song, 2016), 수문학적 요인(Kang and Yoon, 2002; Kwon et al., 2006)을 기반으로 분석을 수행하였다. 최근 국민 생활 수준의 향상으로 인해 물수요량이 증가하여 생활용수 등의 용수부족으로 인한 사회·경제적 가뭄이 증가하고 있는 것으로 조사되었다(Yoo et al., 2011). 즉, 가뭄의 발생에 기상·농업·수문학적 가뭄 인자에 더해 물수요 및 공급 관련 인자가 연관되어 있는 것이다. 따라서, 가뭄 영향인자를 정량화, 평가 시 사회·

경제적 요인을 반영해야 한다.

본 연구는 총 아홉 가지의 기상학적(강수량, 평균기온, 평균 풍속, 상대습도, 평균기압, 일조율), 수문학적(저수위, 댐방류량), 사회·경제적 인자(생활용수사용량)와 가뭄지수 간의 상호관계를 정량화하여 가뭄영향인자를 도출하고 직·간접적 영향분석을 수행하였다. 분석을 위해 통계학적 기법 기반 특징선택 모형인 선형판별분석(Linear Discriminant Analysis)과 주성분분석(Principal Component Analysis)을 사용하였으며, 가뭄의 심도는 표준강수지수(Standardized Precipitation Index, SPI)를 이용하였다. 연구 대상 지역으로 충청남도 보령시를 선정하였다. 특징선택기법별, 이용 가뭄지수 종류별 결과를 방사형 차트(Radar chart)를 이용하여 도시 및 비교하였다.

2. 통계학적 기법 기반 특징선택 모형

특징선택(Feature Selection)은, 원자료 중 분류나 모형의 정확도 향상에 있어 가장 유효한 자료의 부분집합(sub-set)을 찾아내는 것을 말하며, 자료 간의 영향도 정량화에 사용된다. 본 연구에서는 가뭄에 영향을 주는 기상, 수문, 사회·경제 인자와 표준강수지수 사이의 영향도 정량화를 통해 상대적으로 가뭄지수에 영향이 큰 주요영향인자를 찾고, 인자들 간의 영향도 순위를 분석한다.

또한, 본 연구에서는 상기 인자들을 직·간접적 영향 요인으로 분류한다(Fig. 1). 직접적 영향인자는 가뭄 심도에 직접적으로 영향을 주는 인자(X1)를 말하며(Fig. 1(a)), 간접적 영향인자(X2)는 Fig. 1(b)와 같이 직접적 영향인자(X1)에 직접적으로 영향을 주지만, 가뭄지수에는 직접 영향을 주지 않는 인자를 말한다(Fig. 1(b)).

(a) Direct impact factor (b) Indirect impact factor

Fig. 1. Drought Impact Factors with Direct and Indirect Influences

이를 위해, 통계학적 기법인 선형판별분석과 주성분분석을 각각 수행하고, 기상학적, 수문학적, 사회·경제적 인자 중 어떤 인자가 가뭄과 밀접한 연관이 있는지 비교하고 평가하였다. Fig. 2는 본 연구에서 수행하는 일련의 과정을 모식도로 나타냈다. 영향도 분석 시, 가용한 모든 기상, 수문 인자를 활용할 수 없기 때문에 과거 가뭄사례 분석 및 정성적 판단에 의해 관련도가 높은 총 9개의 인자를 이용하였다. 특징선택기법에 의한 결과는 시안성이 높은 방사형 차트

Fig. 2. A Schematic of the Proposed Model of this Study

(Radar chart)를 이용하여 표출 및 비교하였다. 다음 절에서는 특징선택기법 이론에 대해 자세히 설명하였다.

2.1 선형판별분석(Linear Discriminant Analysis, LDA)

선형판별분석은 자료 간 분산과 자료 내 분산의 비율을 최대화하는 방식으로 특징선택을 수행하는 방법이다. 선형판별분석법에서는 영향인자 간의 범주를 잘 구분하기 위해 이들의 성질을 동시에 포함하는 축을 찾는 것이 중요하다. 이를 위해, 자료들을 축에 사영하였을 때, 각 범주의 평균값 간의 거리 합이 최대가 되는 축을 찾고, 범주 내의 분산이 최소가 되는 축을 찾는다. 본 연구에서는 독립변환 접근법을 이용하여 전체 분산과 자료 내 비율을 최대화하는 방식으로 특징선택을 수행하였다.

먼저, 전체 분산(S_T)과 자료 내 분산(S_W)를 계산해야 하며, 이는 Eqs. (1), (2)에 각각 나타내었다.

$$S_T = \frac{1}{N-1} \sum_{i=1}^N (x_i - m)(x_i - m)^T \quad (1)$$

$$S_W = \sum_{j=1}^M \left[\frac{1}{N_c - 1} \sum_{i=1}^{N_c} (x_i - m_j)(x_i - m_j)^T \right] \quad (2)$$

여기서, m 은 전체 자료의 평균을 나타내며, m_j ($j=1, 2, 3, \dots, M$)는 각 자료의 평균을 나타낸다. 그리고 전체 분산과 자료 내 비율을 최대화하기 위해 변환행렬 W 를 구하면 Eq. (3)과 같이 나타낼 수 있다.

$$J(W) = \left| \frac{W^T S_T W}{W^T S_W W} \right| \quad (3)$$

여기서, W 는 $S_W^{-1} S_T$ 의 고유벡터를 열벡터로 가지는 행렬로 Eq. (3)과 같은 고유치 분석을 통해 구할 수 있다. 마지막으로 분석 자료의 경우, Wilk's Lamda 방식으로 유의미성을

검증하였다. 영향인자와 표준화된 정준판별함수간의 상관계수를 나타내는 구조행렬의 값이 높은 값을 나타내면 영향력이 높고, 그렇지 않으면 낮은 영향력을 갖는다고 분석하였다.

2.2 주성분분석(Principal Component Analysis, PCA)

주성분분석은 고차원의 자료를 저차원의 자료로 환원시키는 기법이다(Boutsidis et al., 2008). 여기서, 서로 연관 가능성이 있는 고차원 공간의 표본들을 선형 연관성이 없는 저차원 공간(주성분)의 표본으로 변환하기 위해 직교 변환을 사용한다. 주성분의 차원 수는 원래 표본의 차원 수보다 작거나 같다. 주성분 분석은 자료를 한 개의 축으로 사영시켰을 때, 그 분산이 가장 커지는 축을 첫 번째 주성분, 두 번째로 커지는 축을 두 번째 주성분으로 놓이도록 새로운 좌표계로 데이터를 선형 변환한다. 주성분 분석으로부터 선택/추출된 변수는 기존 변수들의 선형결합으로 구성되며 소수의 추출 변수로써 자료를 설명할 수 있기 때문에, 고차원 형태의 자료를 차원 축소하기 위해 효과적으로 활용되고 있다(Wang et al., 2014). 이때 선택이나 추출된 변수는 원래의 자료에서 공분산 행렬로부터 얻어지는 고유값과 고유벡터를 통해 계산되며, 데이터를 설명하는 공간의 기저를 Eq. (4)로 표현한 재구성 오차(reconstruction error)를 최소화할 수 있도록 결정한다.

$$\text{Min}_{\mu, \lambda_i, V_q} \sum_{i=1}^N x_i - (\mu + V_q \lambda_i)^2 \quad (4)$$

여기서, x_i 는 관측치를 의미하고, $\mu + V_q \lambda_i$ 는 주성분분석에 의해 새롭게 구성되는 공간의 기저를 나타낼 수 있다. 예를 들어, 자료가 n 개의 변수를 갖고 있다면, 주성분분석으로부터 추출된 변수는 아래와 같이 Eq. (5)로 표현할 수 있다.

$$Y_n = \alpha_{n1} X_1 + \alpha_{n2} X_2 + \dots + \alpha_{nm} X_n \quad (5)$$

여기서, $\alpha_{i,j}, i, j = 1, 2, \dots, n$ 이며, Eq. (5)에서 Y는 추출된 변수를 나타내며, X의 선형결합의 계수값 인 α 는 기존 변수가 추출된 변수를 구성하는 데에 기여하는 정도로 표현된다. 먼저, 표준화된 각 영향인자 사이의 상관성을 파악하기 위하여 두 가지 인자 사이의 상관계수를 계산하여 상관관계 매트릭스를 작성하고, 이를 활용하여 고윳값(eigenvalues) 및 고유벡터(eigenvectors)를 추정한다. 여기서, 재정렬된 고유벡터는 주성분을 의미하며, 주성분의 분산 값을 주성분 분산누적값으로 나눈 것을 분산설명량이라고 한다. 주성분 분석에서 차원 축소를 위하여 적절한 주성분 분석의 개수를 결정하는 것은 매우 중요하다. 주성분 분석의 개수를 결정하는 방법은 다양하지만, 일반적으로 주관적인 판단에 의지하는 경우가 많다.

주로 누적기여율이 충분히 크다고 판단되는 주성분이나 주성분의 분산이 1보다 큰 주성분까지를 고려하거나, 또는 고윳값의 크기가 현저히 작아지는 주성분까지를 기준으로 한다(Jin et al., 2019). 본 연구에서는 모든 주성분 중에서 분산설명량을 가장 많이 가지고 있다고 판단되는 주성분을 선택하며, 이 중에서 큰 분산을 갖는 인자가 영향력이 높은 경우로 해석할 수 있고, 그렇지 않은 경우, 영향력이 낮다고 판단한다.

3. 대상 지역 및 자료 구축

3.1 대상 지역

본 연구에서는 상기 가뭄영향인자 도출 방법을 대한민국의 충청남도 보령시에 적용하였다. 충청남도 보령시는 최근

10년 동안 수차례의 가뭄 피해를 겪고 있으며, 보령시의 가뭄 대책 수립을 위해 여러 연구가 진행되고 있다. 특히, 보령시는 2014~2015년 극심한 가뭄을 겪으며 연강수량이 각각 1,010 mm, 785 mm 까지 감소하였다. 또한, 보령댐의 저수율이 20% 아래로 감소하면서 보령댐 하류 지역의 주민들은 용수공급 피해를 겪은 바 있다. 본 연구의 대상 지역 및 강수량 지점은 다음 Fig. 3과 같다.

3.2 연구 자료 구축

가뭄지수는 가뭄 피해의 심각도를 정량적으로 표현할 수 있는 지표이고, 현재까지 여러 가지 형태의 가뭄지수가 개발되었다. 이 중에서 표준강수지수는 가뭄은 강수량의 부족으로부터 시작된다는 것에 착안하여 개발된 가뭄지수이다(Mckee et al., 1993). 본 연구에서는 국내에서 가장 널리 이용되고 있는 표준강수지수를 분석에 사용하였다. 표준강수지수를 세 가지의 종류(SPI3, SPI6, SPI9)로 구분하여 활용하였고, 이를 기상, 수문 등의 인자에 영향을 받는 종속 인자로 고려하였다(Fig. 1의 “Drought Index”).

Fig. 2와 같이, 일차적 정성적으로 가뭄 영향인자 선정한 후 이를 구조적으로 정립할 필요가 있다. 본 연구에서는, 표준강수지수에 영향을 줄 수 있을 것으로 판단되는 총 아홉 가지 영향인자를 기상학적 요인(강수량, 평균기온, 평균풍속, 상대습도, 평균기압, 일조율), 수문학적 요인(저수위, 댐방류량), 사회·경제적 요인(생활용수사용량)으로 구분하였다(Table 1). 여기서, 기상학적 요인의 경우, 기존 점 자료의 형태에서 면적 자료의 형태로 바꿔주어야 하므로, 티센다각형법을 이용하여 구역의 평균값으로 변환하여 분석에 활용하였다(Fig. 3).

Fig. 3. Boryeong city Area Used in this Study

Table 1. Summary of Drought Index and 9 Impact Factor Data

Division	Variable	Drought Factor	Source
Drought index	D1	SPI 3	Korea Meteorological Administration (www.weather.go.kr)
	D2	SPI 6	
	D3	SPI 9	
Climate factor	C1	Monthly total precipitation	Korea Meteorological Administration (www.weather.go.kr)
	C2	Temperature	
	C3	Wind speed	
	C4	Humidity	
	C5	Atmospheric pressure	
	C6	Amount of sunshine	
Hydrological factor (Boryeong dam)	H1	Low-water level of dam	Water Resources Management Information System (www.wamis.go.kr)
	H2	Total outflow of dam	
Socio-economic factor	SE1	Residential water usage	Boryeong City Hall (https://www.brcn.go.kr)

본 연구에서 선정된 인자들의 경우, 표준강수지수(SPI3, SPI6, SPI9)와 각 인자간의 영향 정도를 정량적으로 분석하기 위해 활용되었다.

본 연구에서 활용한 모든 자료(기상학적 요인, 수문학적 요인, 사회·경제적 요인)는 월 단위 자료를 사용하였으며, 보령시에 극심한 가뭄이 발생하기 이전인 09년도 8월부터 19년 7월까지 총 10년 치 자료를 사용하였다. Table 1은 가뭄 피해의 영향인자 정량화를 위해 사용된 자료의 정보를 정리하였다.

4. 가뭄의 영향인자 정량화 분석 결과

본 연구에서 선정한 총 12가지 변수(아홉 가지의 영향인자와 세 가지의 표준강수지수) 간의 영향력을 정량화하기 위해, 기초통계 분석이 선행되어야 한다. 먼저, 표준강수지수와

각 인자간의 직·간접적 관계 파악을 위해 상관관계 분석을 수행하였다(Fig. 4). 또한, 두 특징선택기법에 의해 결과를 도출하고, 결과 해석에 앞서 그 유의성을 먼저 확인하였다. 선형판별분석법의 경우, 유의확률을 검토하여 도출 결과의 유효성을 확인하였다(P-값이 0.05 이하). 주성분분석법의 경우, 각 주성분 사이의 고유판차 차이와 분산설명량 등을 검토하였다.

Fig. 4는 원 데이터를 이용한 상관성분석 결과 획득한 상관성 매트릭스이다. 상관관계 분석 결과, SPI6과 저수위(월 최저수위)가 상대적으로 높은 상관계수 값(0.589)을 가지므로 가뭄에 직접 영향을 주는 직접영향인자로 확인되었다. 간접적인 영향인자는 SPI6과의 상관성은 낮으며, 직접영향인자인 저수위와 상관성이 높은 인자로 선정할 수 있다. 하지만 상관성 매트릭스에서 저수위와 다른 기상, 수문, 사회·경제인자 사이의 상관계수 값은 모두 0.3이하로 확인

	Atmospheric pressure	Monthly total precipitation	Humidity	Residential water usage	Amount of sunshine	Low-water level of dam	Temperature	Total outflow of dam	Wind speed	SPI6
Atmospheric pressure	1	0.562	0.715	0.283	0.438	0.152	0.917	0.272	0.667	0.095
Monthly total precipitation		1	0.550	0.145	0.610	0.077	0.580	0.547	0.280	0.142
Humidity			1	0.528	0.550	0.269	0.766	0.235	0.414	0.186
Residential water usage				1	0.054	0.081	0.496	0.051	0.234	0.217
Amount of sunshine					1	0.081	0.398	0.411	0.247	0.004
Low-water level of dam						1	0.067	0.278	0.027	0.589
Temperature							1	0.273	0.708	0.101
Total outflow of dam								1	0.210	0.357
Wind speed									1	0.051
SPI6										1

Fig. 4. Results of Correlation Coefficients with SPI 6

되어 간접영향인자를 정할 수 없었다, 그 중 가장 큰 상관계수 값 0.278이 댐방류량과 저수위 사이에서 도출되었다. 이는 기상, 수문 등의 인자와 가뭄지수 사이의 영향관계가 단순한 상관분석 - 1대 1관계로 규명할 수 없는 비선형성이 존재한다는 것은 보여준다.

선형판별분석 결과의 유효성 검증을 위해, 고유벡터를 산정한 후 Wilk's Lamda의 식을 이용하여 유의성을 확인하였는데, 모든 결과에서 P-값이 0.05 이하로 산정되었다. 주성분분석의 차원 축소 적용성을 검토한 결과, 가뭄 영향인자 정량화 분석에 주성분분석이 적용가능한 것으로 확인하였다. 특히, 가뭄의 심도에 따라 추출된 성분점수는 72.782% ~ 75.166%의 범위 값을 가졌으며, 제3 주성분까지 성분점수를 고려하는 것이 안정적으로 확인되었다.

Fig. 5는 특징선택기법(선형판별분석과 주성분분석)을 통해 산정된 영향인자별 가뭄에의 영향력을 방사형 차트로 도시한 결과이다. 여기서, 값이 0에 가까울수록, 가뭄에 대한 영향력이 적고, 1에 가까울수록 영향력이 크다. 선형판별분석의 경우, 정준판별함수 간의 상관계수를 나타내는 구조행렬의 값으로 변환하여 각 가뭄 심도에 따른 영향인자 분석을 수행하였다(Fig. 5(a)). 선형판별분석의 결과, 영향력이 가장 높은 인자는 저수위로 확인되었으며, 영향력의 정도는 0.772로 산정되었다. 특히, 표준강수지수의 기간이 증가할수록(SPI3에서 SPI9) 저수위의 가뭄에 대한 영향력은 증가하는 것을 확인할 수 있다. 반대로, 영향력의 정도가 가장 낮은 인자는 기압과 일조량으로 각각 0.047과 0.011으로 산정되었다. 월강수량의 영향력은 SPI 누적강수 고려기간이 증가하면 감소하였다.

강수량(누적강수량)으로 계산한 SPI와 월강수량의 영향도가 낮게 나온 이유는 다음과 같다. 표준강수지수(SPI3, SPI6, SPI9)는 현재 시점으로부터 3개월, 6개월, 9개월 전부터의 누적강수량을 고려하기 때문에 월강수량과는 연관도가 미비하다. 예를 들어, 7월 1일의 SPI3 값은 7월의 강수량보

다 4, 5, 6월의 강수량을 반영하고 있으며, 해당 월의 중반 이후가 되어서도 일 SPI 값은 해당 월의 강수량을 부분적으로만 고려한다. 한편, 3, 6, 9개월 동안 누적된 강수량 값이 크면, 자연적으로 유역의 댐, 저수지 등에 저류량 및 토양함수량이 늘어나게 되고 이는 저수위의 증가로 이어진다. 이러한 이유로, 표준강수지수는 댐저수위에 큰 영향을 받는 것으로 분석되었다.

주성분분석의 결과는 분산 설명량을 표준화하여 1에 가까울수록 영향력이 높도록 방사형차트에 도시하였다(Fig. 5(b)). 선형판별분석 결과와 마찬가지로, 저수위의 영향력이 가장 높게 산정되었는데, 영향도는 0.911였다. 전반적으로 SPI의 기간이 증가할수록 저수위의 영향도가 증가하였다(SPI6일 때 최대값). 선형판별분석 결과와는 달리, 댐방류량의 영향도가 두 번째로 컸다. 주성분분석에서는 월강수량, 생활용수사용량이 SPI3에 어느정도 영향이 있는 것으로 도출되었지만, 표준강수지수의 시간장경이 길어질수록 이들의 영향도는 감소하는 것으로 파악되었다. 따라서, 두 종류의 특징선택기법과 표준강수지수의 고려시간에 상관없이 저수위가 가뭄에 가장 큰 영향을 주는 것으로 확인되었다.

5. 결 론

본 연구는 과거 데이터를 기반으로, 기상학적, 수문학적, 사회·경제적인자와 가뭄지수 간의 상호관계를 정량화하기 위해 통계학적 특징선택기법(선형판별분석과 주성분분석)을 적용하였다. 먼저, 가뭄 피해와 영향도가 높다고 판단되는 기상학적 요인(강수량, 평균기온, 평균풍속, 상대습도, 평균기압, 일조량), 수문학적 요인(저수위, 댐방류량), 사회·경제적 요인(생활용수사용량) 등 총 아홉 가지 영향인자를 선정하고, 가뭄심도를 나타내는 표준강수지수(SPI3, 6, 9)와 함께 분석에 고려하였다. 충청남도 보령시를 대상지역으로 연구 자료를 구축하여 선형판별분석법과 주성분분석법을

Fig. 5. Polygon Radar Charts Presenting the Drought Impact Analysis Result of (a) LDA and (b) PCA

각각 적용하였다. 결과를 분석하여 가뭄에 직접적으로 영향을 미치는 인자를 확인하였다. 본 연구에서 다음과 같은 결론을 얻을 수 있었다.

- (1) 가뭄 피해의 직·간접 영향력을 정량화하기 위해 영향인자 간의 상관성 분석을 수행하였다. 가뭄에 직접 영향을 주는, 직접영향인자로 저수위(월 최저수위)가 선정되었고(SPI6과 상관계수 0.589), 단순한 상관성분석 - 1대1관계로는 명확한 직·간접적 영향관계를 규명할 수 없음을 확인하였다. 이는 가뭄현상과 기상, 수문 등 간의 관계에 비선형성이 존재하기 때문이다. 향후 베이저안 네트워크(확률론적 방향성 비순환 그래프 모형)의 도입이 필요하다. 베이저안 네트워크를 이용하여, 네트워크의 연결구조에 따라 각 영향인자의 영향력을 조건부 확률에 기반하여 표출가능하며, 연관성을 보다 체계적으로 도출할 수 있을 것으로 기대한다.
- (2) 선형판별분석과 주성분분석의 결과 모두 저수위의 영향력이 가장 높은 것을 확인되었다, 이는 본 연구에서 고려한 표준강수지수가 누적강수량을 기반으로 계산되는 가뭄지표이기 때문이다. 장기간의 강수량 누적 값이 크면, 자연적으로 댐 저수량 및 저수위가 증가한다. 본 연구에서 고려한 월강수량은 표준강수지수가 고려하는 시간간격과 시차가 존재하기 때문에 예상과 달리 가뭄에의 영향도가 낮게 산정되었다.

후속 연구로 (1) 생활용수사용량 외 관광객 수, 공업관련 생산량 등 보다 다양한 사회·경제적 인자를 고려한 분석을 수행할 수 있으며, (2) 다른 가뭄지수를 이용한 분석도 가능하다. 마지막으로 (3) 특성이 다른 지역에 해당 분석을 수행하여, (4) 지역별로 도출된 주요영향인자를 조합하여 지역 가뭄 시나리오를 구축, 다양한 가뭄 연구에 활용할 수 있다. 본 연구에서 제시한 분석방법을 가뭄 예측 및 전망 모형, 위험지도 개발 등 다양한 가뭄 연구분야에 활용할 수 있을 것으로 기대한다.

감사의 글

본 연구는 행정안전부 극한재난대응기반기술개발사업의 연구비 지원(2019-MOIS31-010)에 의해 수행되었습니다.

References

Boryeong City Website. Retrieved August 22, 2019, from <https://www.brcn.go.kr/>

Boutsidis, C., Mahoney, M.W., and Drineas, P. (2008). Unsupervised feature selection for principal components analysis. *Proceedings of the 14th ACM SIGKDD International*

Conference on Knowledge Discovery and Data Mining, pp. 61-69.

Chang, E., and Park, E. (2004). Mapping of drought index using satellite imagery. *Journal of the Korean Society for Geospatial Information Science*, Vol. 12, No. 4, pp. 3-12.

Choi, J.W., Heo, E.Y., and Shim, M.P. (2000). The economic impacts of water supply constraints during a drought using input-output analysis. *Journal of Korea Water Resources Association*, Vol. 33, No. 5, pp. 647-658.

Choi, J., Lee, T., Kim, S., and Shin, Y. (2017). Comparison of drought characteristics using RAI and soil moisture. *Proceedings of 2017 Annual Conference*, The Korean Society of Agricultural Engineers, p. 237.

Hong, M.H., Kim, J.H., Jung, G.J., and Jung, S.S. (2016). Rainfall threshold (ID curve) for landslide initiation and prediction considering antecedent rainfall. *Journal of the Korean Geotechnical Society*, Vol. 32, No. 4, pp. 15-27.

Jin, B., Pei, X., and Liang, H. (2019). Establishment of HPLC fingerprint, cluster analysis and principle component analysis of citri reticulatae pericarpium viride. *Medicinal Plant*, Vol. 10, No 1, pp. 69-73.

Kang, I.J., and Yoon, Y.N. (2002). A study on the hydrologic decision-making for drought management: 1. An analysis on the stochastic behavior of PDSI using markov chain. *Journal of Korea Water Resources Association*, Vol. 35, No. 5, pp. 583-595.

Kim, N.S., Kim, J.S., Jang, H.W., and Lee, J.H. (2015). Hydrologic risk analysis based on extreme drought over the Korean Peninsula under climate change. *J. Korean Soc. Hazard Mitig.*, Vol. 15, No. 4, pp. 45-52.

Kim, O., Choi, J., Jang, M., and Yoo, S. (2006). Soil moisture index development for drought assessment. *Proceedings of 2006 Annual Conference*, The Korean Society of Agricultural Engineers, pp. 464-468.

Korea Meteorological Administration (KMA) Website. Retrieved August 12, 2019, from <http://www.weather.go.kr/weather/main>

Kwon, H.J., Park, H.J., Hong, D.O., and Kim, S.J. (2006). A study on semi-distributed hydrologic drought assessment modifying SWSI. *Journal of Korea Water Resources Association*, Vol. 39, No. 8, pp. 645-658.

Lee, C.S., Choi, S.A., Shim, M.P., and Jung, K.S. (2004). Allocation of water supplied by multi-purpose dam using the estimate of weighting factors. *Journal of Korea Water Resources Association*, Vol. 37, No. 8, pp. 663-674.

Lee, D.R., Moon, J.W., Lee, D.H., and Ahn, J.H. (2006). Development of water supply capacity index to monitor droughts in a

- reservoir. *Journal of Korea Water Resources Association*, Vol. 39, No. 3, pp. 199-214.
- Lee, H.J., and Shim, M.P. (2002). Decision making for priority of water allocation during drought by analytic hierarchy process. *Journal of Korean Water Resources Association*, Vol. 35, No. 6, pp. 703-714.
- Lee, J.H., Kim, J.S., Jang, H.W., and Lee, J.C. (2013). Drought forecasting using the multi layer perceptron (MLP) artificial neural network model. *Journal of Korean Water Resources Association*, Vol. 46, No. 12, pp. 1249-1263.
- Lee, Y. (2012). Climate change and socioeconomic change effects on the four major rivers: An economic appraisal. *Journal of Environment Policy*, Vol. 11, No. 4, pp. 107-130.
- McKee, T.B., Doesken, N.J., and Kleist, J. (1993). The relationship of drought frequency and duration to time scales. *Proceedings of the 8th Conference on Applied Climatology*, Anaheim, CA, USA, pp. 179-184.
- Nam, W.H., Choi, J.Y., Jang, M.W., and Hong, E.M. (2013). Agricultural drought risk assessment using reservoir drought index. *Journal of the Korean Society of Agricultural Engineers*, Vol. 55, No. 3, pp. 41-49.
- Park, H., and Song, J. (2016). The impacts of natural disaster on local economy. *Journal of Korea Planning Association*, Vol. 51, No. 2, pp. 193-213.
- Park, J.C., Choi, K.J., and Song, S.H. (2014). Prediction of regional drought considering aspect and elevation in Jeju Island under future climate change. *Journal of Environmental Science International*, Vol. 23, No. 4, pp. 649-660.
- Park, J.Y., Yoo, J.Y., Lee, M., and Kim, T.W. (2012). Assessment of drought risk in Korea: Focused on data-based drought risk map. *Journal of the Korean Society of Civil Engineers B*, Vol. 32, No. 4B, pp. 203-211.
- Ryoo, S., and Yoo, C. (2004). A modified standardized precipitation index (MSPI) and its application. *Journal of Korean Water Resources Association*, Vol. 37, No. 7, pp. 553-567.
- Wang, Y., Jodoin, P.M., Porikli, F., Konrad, J., Benzeeth, Y., and Ishwar, P. (2014). An expanded change detection benchmark dataset. *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 393-400.
- Water Resources Management Information System Website. Retrieved July 18, 2019, from <http://www.wamis.go.kr/main>
- Yi, J., and Yoo, J. (2003). Development of drought forecasting and warning technique in basin. *Journal of the Korean Society of Civil Engineers B*, Vol. 23, No. 6B, pp. 499-506.
- Yoo, J.Y., Shin, J.Y., Kim, D., and Kim, T.W. (2013). Drought risk analysis using stochastic rainfall generation model and copula functions. *Journal of Korean Water Resources Association*, Vol. 46, No. 4, pp. 425-437.
- Yoo, J.Y., Park, J.Y., Kim, T.W., and Park, M.J. (2011). The assessment of socioeconomic droughts using a water excess deficiency index. *Journal of the Korean Society of Civil Engineers B*, Vol. 31, No. 3B, pp. 253-264.
- Yun, H.C., Gwak, Y.S., Won, C.H., and Ahn, J.C. (2018). Development of assessment techniques for drought forecasting-warning information. *Proceedings of the Korean Environmental Sciences Society Conference*, p. 115.

Received	September 26, 2019
Revised	September 27, 2019
Accepted	October 14, 2019