

동질성 은닉 마코프 모형을 적용한 가뭄특성의 확률론적 평가

유지영* · 권현한** · 김태웅*** · 이승오****

Yoo, Ji-Young*, Kwon, Hyun-Han**, Kim, Tae-Woong***, Lee, Seung-Oh****

Probabilistic Assessment of Drought Characteristics based on Homogeneous Hidden Markov Model

ABSTRACT

Several studies regarding drought indices and criteria have been widely studied in the literature. If one defines the onset, severity, and end of droughts, in general, a certain threshold needs to be set to assess the drought events. However, the uncertainty associated with the threshold is a critical problem in drought analysis. To take full advantage of the inherent features in the rainfall series, a Hidden Markov Model (HMM) based probabilistic drought analysis was proposed rather than using the existing threshold based analysis. As a result, the proposed HMM based probabilistic drought analysis scheme shows better performance in terms of defining drought state and understanding underlying characteristics of the drought. In addition, the HMM based approach is capable of quantifying the uncertainties associated with the classifying meteorological drought condition in a systematic way.

Key words : Drought, Hidden markov model, Posterior probability, Transition probability

초 록

현재 국내외적으로 다양한 가뭄지수들이 개발과 더불어 가뭄평가를 위한 다양한 기준이 개발되고 있다. 가뭄지수를 이용하여 가뭄의 시작, 강도 및 종료 시점을 정의할 경우, 일반적으로 특정 기준값(threshold)에 근거한 해석이 이루어지고 있으나, 이는 실제 가뭄을 평가하는데 불확실성을 가중시키는 원인으로 작용하고 있다. 따라서 본 연구에서는 기존의 가뭄지수 기반의 가뭄판단 시 적용되어져 왔던 특정한 기준값에 근거한 해석이 아닌, 기상학적 가뭄발생의 주된 원인 중 하나인 월강수량 자료에 내재되어 있는 특징을 최대한 활용하고자 은닉 마코프 모형(Hidden Markov Model, HMM) 기반의 확률론적 가뭄해석기법을 제안하였다. 그 결과, 본 연구에서 제시한 HMM 기반의 확률론적 가뭄해석기법은 기존의 다양한 가뭄지수를 적용한 가뭄해석과 비교하여 기상학적 측면에서의 가뭄판단의 명확한 기준 제시 및 가뭄발생의 원인을 규명하는 데 있어 체계적으로 불확실성을 감안한 해석이 가능하였다.

검색어 : 가뭄, 은닉 마코프 모형, 사후확률, 전이확률

1. 서론

최근에는 지구온난화로 기인하는 기상변동성의 증가로 인해 가뭄이나 홍수 등과 같은 수문학적 재해발생 빈도가 증가하고 있다. 그러나 홍수와는 달리 가뭄은 국가적인 차원에서 지속적인 관심과 실질적인 대책마련에 대한 대비는 미비한 실정이다. 특히, 가뭄은 단순한 강수량의 부족으로 인한 물 부족에서 국한되는 것이 아니라 사회·경제적으로 심각한 피해를 연쇄적으로 지속시키는 특성이 있다.

* 정회원 · 전북대학교 토목공학과 연구교수, 공학박사 (jyoo84@gmail.com)

** 정회원 · 교신저자 · 전북대학교 토목공학과 부교수, 공학박사 (Corresponding Author · Chonbuk National University · hkwon@jbnu.ac.kr)

*** 정회원 · 한양대학교 건설환경플랜트공학과 부교수, 공학박사 (twkim72@hanyang.ac.kr)

**** 정회원 · 홍익대학교 건설도시공학부 부교수, 공학박사 (seungoh.lee@hongik.ac.kr)

Received September 5, 2013/ revised October 22, 2013/ accepted November 13, 2013

현재 국내외적으로 가뭄의 특성은 수문학적 또는 기상학적 관점에서 산정된 가뭄지수에 의해서 주로 평가되고 있으며, 일반적으로 가뭄은 평균적인 강수 및 유출 특성에 따라 건조한 정도와 지속되는 기간을 정의하여 해석된다. 그 외에도 가뭄은 지표수 및 지하수의 부족, 토양수분의 부족, 실제와 잠재증발산의 차이, 가뭄으로 인한 용수 및 전력공급의 제한 등의 다양한 원인에 따라 정의되어 진다. 이러한 가뭄은 주로 가뭄지수에 의해 상황을 판단하며, 표준강수지수(Standardized Precipitation Index, SPI), 팔머가뭄지수(Palmer Drought Severity Index, PDSI), 지표수공급지수(Surface Water Supply Index, SWSI), 유효가뭄지수(Effective Drought Index, EDI) 등과 같은 다양한 가뭄지수들이 개발되어져왔다. 이처럼 다양한 가뭄지수의 개발과 더불어 가뭄지수 별 가뭄분류의 등급도 다양해지고 있으며, 이에 따른 다양한 가뭄지수를 기반으로 가뭄을 판단할 경우, 가뭄 등급 결정 시 혼란이 발생하는 문제가 있다 (Mishra and Singh, 2010). Steinmann (2003)은 다양한 가뭄지수 별 발생빈도를 통일화하기 위한 방안을 제시한 바 있으며, 국내 연구에서도 이와 동일한 개념으로 우리나라 주요 가뭄발생 현황기록을 참고하여 가뭄심도의 분류기준으로 재작성 한 후, 다양한 가뭄등급의 통일화를 위한 연구가 수행된 바 있다(Kwon et al., 2013b).

기존 가뭄지수를 이용하여 가뭄의 시작과 끝을 정의할 경우, 일반적으로 특정 기준값(threshold)에 근거한 해석이 이루어지고 있으나, 이는 실제 가뭄을 평가하는데 불확실성을 가중시키는 원인으로 작용하고 있다. 예를 들어, 가뭄지수 중에서 기상학적 측면의 가뭄을 정의하기 위해 가장 일반적으로 적용되고 있는 표준강수지수는 계산상의 간편성과 다양한 시간 스케일을 고려할 수 있는 장점으로 인해 단기 및 장기적인 강수 부족량을 효율적으로 평가할 수 있다. 그러나 이러한 장점에도 불구하고 표준강수지수(SPI)의 대표적인 단점은 강수량의 편차가 적은 강수지점의 경우 미소한 강수량의 차이를 기반으로 가뭄의 정도를 등급으로 구분하는데 있어서 서로 다른 결과로서 나타날 수 있다는 점이다.

본 연구에서는 기상학적 측면에서 가뭄을 정의하고 수문학적 지속성 결여문제를 보완할 수 있는 방안을 수립하여 가뭄의 특성을 정량적으로 규명하고자 한다. 따라서 기존의 가뭄지수에서 적용되었던 특정한 기준값에 근거한 해석이 아닌, 강수량 시계열 자료에 내재되어 있는 특징을 활용할 수 있는 은닉 마코프 모형(Hidden Markov Model, HMM)을 도입하여 가뭄특성을 해석하고자 한다. 수문분야에서 처음으로 HMM을 적용한 연구는 미국 워싱턴 지역의 강수발생을 모델링하기 위해서 이용한 바 있다(Hughes and Guttorp, 1994). 최근 국내에서 수문분야에서는 복잡한 강수발생 패턴을 인지하고 강수량의 다양한 분포특성을 고려할 수 있도록 혼합분포를 이용한 동질성 HMM을 적용한 바 있다(Kwon et

al., 2013a). 최근 Mallya et al. (2013)은 일반적인 가뭄지수에 근거한 가뭄 해석 시 기준값 설정에 대한 불확실성을 고려한 가뭄 해석을 수행하고자 HMM 모형을 활용하였다.

이러한 HMM의 장점을 활용하여 현재까지 국내에서 수행되어진 바 없는 가뭄연구에 적용해 보고자, 본 연구에서는 강수계열의 건조 및 습윤 상태를 숨어있는 은닉상태로 가정하며 최적화 과정을 통하여 가뭄상태를 결정하는 모형을 구성하였다. 또한, 본 모형의 효율성을 객관적으로 평가하고자 기존 표준강수지수와 가뭄 해석 과정으로 비교 검토하였다. 본 연구에서는 기상청 산하의 서울관측소와 부산관측소의 월강수량을 대상으로 모형의 적합성을 평가하였다. 본 논문의 구성은 다음과 같다. 1장에서는 연구의 배경 및 목적에 대하여 서술하였으며, 2장에서는 가뭄의 정의 및 HMM에 대하여 설명하였고, 3장에서는 HMM 기반의 가뭄해석 결과에 대하여 요약하였으며, 마지막으로 결론 및 토의를 4장에 수록하였다.

2. 연구방법

2.1 가뭄의 정의

일반적으로 가뭄의 재현특성을 분석하기 위해서는 Yevjevich (1967)에 의해 제안된 연속이론(Run Theory)을 적용하여 가뭄 사상을 정의하고 가뭄의 특성변수(가뭄의 지속기간, 심도, 간격)를 대상으로 재현기간을 산정하는 방법이 효율적이라고 알려져 있다 (Dracup et al., 1980; Sen, 1980; Loaiciga and Leipnik, 1996; Mishra and Desai, 2005). 본 연구에서는 기상학적 측면의 가뭄해석을 수행하고자 3개월 누적 월강수량 시계열을 구축하였으며, 평년대비 부족량과 잉여량을 구분하고자 Eq. (1)을 이용하여 가뭄 절단수준을 결정하였다(Fig. 1 참고).

$$x_{(n,i)} = \frac{x_{(s,i)} + x_{(s+1,i)} + \dots + x_{(n-1,i)}}{n - ys} \quad (1)$$

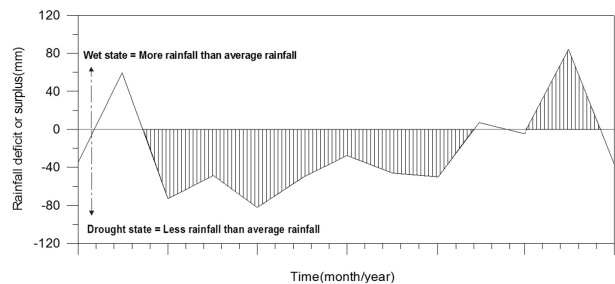


Fig. 1. Definition of Drought State Using the Run Theory for a Given Threshold Level

여기서 $x_{(n,i)}$ 는 n 년도 i 월의 가뭄절단수준을 나타내며, i 는 1, 2, 3, ..., 12 month를 나타내고, ys 는 강우자료보유의 시작년도이다.

2.2 은닉 마코프 모형 이론

본 연구에서 강수자료는 은닉상태(hidden state)의 형태를 가지는 것으로 고려하였으며 은닉상태 안에서의 천이확률은 시간에 따라 변화하지 않는 것으로 가정하였다. 즉, 동질성 HMM을 활용하였으며 시간에 따른 천이확률 및 은닉상태의 확률분포의 변동은 고려하지 않았다. HMM은 은닉상태마다 혼합분포 형태의 조건부 확률분포를 부여하고 이들 확률분포로부터 수문변량을 모의할 수 있는 방법으로 우리나라와 같이 수문변량의 변동성이 큰 경우에 적합한 방법이다(Kwon et al., 2013a). 또한, HMM은 조건부 독립가정을 기본으로 관측 자료로부터 은닉상태의 확률론적 과정(stochastic process)을 우선적으로 도출하고 이를 통하여 수문변량을 2차적으로 모의할 수 있는 이중의 확률론적 과정이라 할 수 있다. 우선 시간의 차원을 갖는 강수 시계열을 시간 T의 연속 벡터로 나타내면 Eq. (2)와 같다.

$$D_{1:T} = (D_1, \dots, D_T) \tag{2}$$

또한 시간 t에 해당하는 은닉상태(hidden state)를 나타내는 연속 계열순서(sequence)는 Eq. (3)과 같다.

$$S_{1:T} = (S_1, \dots, S_T) \tag{3}$$

HMM은 조건부 독립 가정들을 이용하여 $D_{1:T}$ 와 $S_{1:T}$ 의 결합확률(joint probability) 분포를 정의하며 두 가지 조건부 독립 가정은 다음과 같다. 첫째, 주어진 은닉상태와 k일 이전의 강수시계열로 이루어진 모형에서 관측시계열 벡터는 다른 모든 변수와 조건부적으로 독립이라고 가정한다. 둘째, 시계열의 바로 이전의 은닉상태에만 상관성이 있다고 가정하며 일반적으로 Markov Chain이라고 하면 1차 Markov Chain을 말하는 것으로 은닉상태의 확률분포는 시간적으로 오직 이전 은닉상태에 의존한다.

$$P(d_t | d_{1:t-1}, s_{1:t}) = P(d_t | s_t). \tag{4}$$

$$P(s_t | s_{1:t-1}) = \begin{cases} P(s_t | s_{t-k:t-1}) & t > k, \\ P(s_t | s_{1:t-1}) & t \leq k. \end{cases} \tag{5}$$

위의 Eq. (5)에서 추가적으로 명시된 언급이 없으면 일반적으로 $k=1$ 의 은닉상태 확률분포는 다음과 같다.

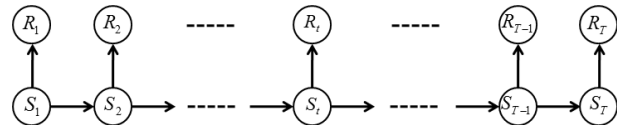


Fig. 2. Bayesian Network Representation of a Homogeneous HMM Satisfying Conditional Independence Assumptions

$$P(s_t | s_{1:t-1}) = \begin{cases} P(s_t | s_{t-1}) & t \geq 2, \\ p(s_1) & t = 1. \end{cases} \tag{6}$$

동질성 HMM은 조건부 확률 $P(s_t | s_{t-1})$ 을 추정하는데 있어서 Fig. 2와 같은 정상성 가정을 기본으로 이루어진다. 앞서 언급된 은닉상태에 조건부확률 $P(s_t | s_{t-1})$ 가 시간에 따라서 일정하다는 정상성 가정에 기본을 두고 있다. 동질성 HMM에 대해서 $\Pi = (\pi_1, \dots, \pi_K)$ 는 초기 확률벡터라 하고 $\Gamma = (\gamma_{11}, \dots, \gamma_{KK})$ 를 천이확률 행렬(transition probability matrix)이라고 정의한다. 여기서 Υ 는 각각의 은닉상태별로 결정되는 출력(emission) 확률밀도함수의 매개변수로서 은닉상태 S_t 를 기준으로 강수시계열 d_t 의 발생확률 $F_t(d) = P(d_t | S_t = i)$ 을 산정하는데 활용된다(Matsuyama, 2011). 다음 Eq. (6)은 동질성 HMM의 우도함수를 나타낸다.

$$P(d_{1:T}, S_{1:T} = s_{1:T} | \Pi, \Gamma, \Upsilon) = \left[\pi_{s_1} \prod_{t=2}^T \gamma_{s_{t-1}s_t} \right] \left[\prod_{t=1}^T F_{s_t}(d_t | d_{t-1}) \right] \tag{6}$$

본 연구에서는 강수 부족량 및 잉여량의 시계열 자료는 강수량의 월별 평균적인 특성을 기준으로 정규화된 자료인점에 착안하여 출력확률밀도함수로 Gaussian 분포를 활용하였다.

3. 적용 결과

본 연구에서는 서울관측소와 부산관측소 등 2개의 관측소의 월강수량 자료를 대상으로 모형을 적용시켰다. 분석기간은 1974년부터 2010년까지 37년간의 3개월 누적월강수량 자료를 대상으로 하였다. 분석에 앞서 월강수량은 해당 월의 평균값을 기준으로 정규화를 실시하여, 서로 다른 지점간의 상대적인 비교가 가능하도록 하였다. 모형의 적용에 앞서 최적의 은닉상태 수를 결정하기 위하여 대수우도함수(log-likelihood function)를 기준으로 서로 다른 은닉상태 수(K)를 평가하였다. 은닉상태 수(K)를 5-11까지 변화하면서 대수우도를 산출하였고 그 과정에서 K가 증가함에 따라 우도함수 또한 증가하다가 K=7을 중심으로 대수우도가 수렴하는 것으로 확인되었다. 실제, 은닉상태 수를 무한히 증가할 경우 모형의 과적합으로 인한 문제가 발생할 수 있으며, 이러한 점을

고려하여 대수우도가 수렴하는 부분에서 최적의 은닉상태 수를 결정하는 것이 사실상 효율적이라고 알려진 바 있다(Matsuyama, 2011). 본 연구에서는 HMM의 은닉상태 수를 7개로 정의하였으며, Fig. 3과 같은 은닉상태별 평년대비 강수 부족량과 잉여량을 구분할 수 있었으며, 이와 더불어 각 은닉상태별 발생 빈도를 동시 평가할 수 있었다.

서울지점과 부산지점에서는 은닉상태 1-3은 상대적으로 강수량 부족이 심한 상태를 나타내며, 은닉상태 5-7은 강수량 잉여량이 큰 특징을 나타낸다. 서울지점의 은닉상태별 빈도를 추정해보면

은닉상태 1-4는 건조한 상태로서 약 60%의 비율, 은닉상태 5-7은 습윤한 상태로서 약 40%의 비율로 추정되었다. 이 중에서 은닉상태 1은 상대적으로 가장 큰 가뭄심도를 나타내며 약 9%의 빈도를 갖고 있다. 반면, 부산지점의 은닉상태별 빈도를 추정해보면 은닉상태 1-4는 건조한 상태로서 약 65%의 비율, 은닉상태 5-7은 습윤한 상태로서 약 35%의 비율로 추정되었다. 이 중에서 은닉상태 1은 상대적으로 가장 큰 가뭄심도를 나타내며 약 18%의 빈도를 갖고 있다. 각 지점의 은닉상태별 천이확률을 추정하면 Tables 1 and 2와 같으며, 이 중에서 은닉상태 1-4의 확률은 건조한 상태간의

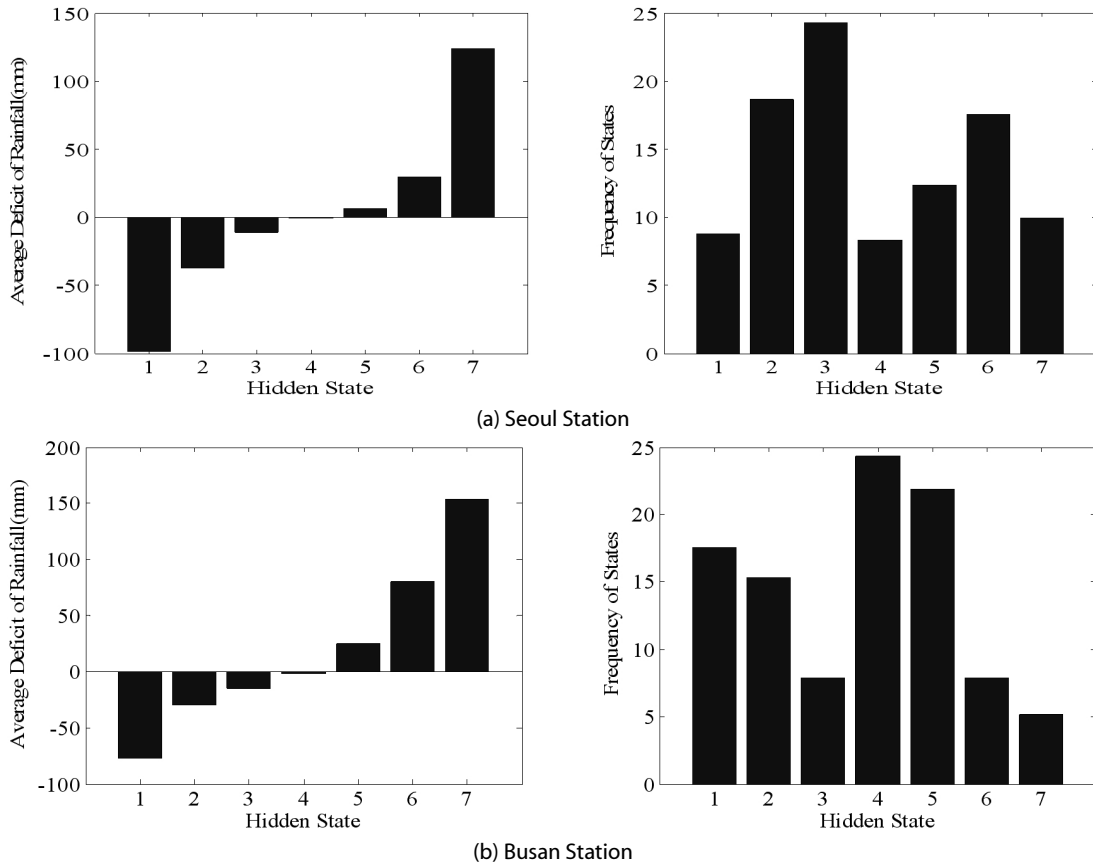


Fig. 3. Average Deficit and Surplus of Rainfall According to Hidden States and Their Frequency

Table 1. Transition Probability Matrix of Hidden States for Average Deficit of Rainfall at Seoul Station

	State-1	State-2	State-3	State-4	State-5	State-6	State-7
State-1	0.50	0.30	0.17	0.00	0.00	0.03	0.00
State-2	0.20	0.35	0.26	0.00	0.00	0.19	0.00
State-3	0.01	0.09	0.49	0.14	0.19	0.07	0.01
State-4	0.00	0.00	0.41	0.39	0.18	0.00	0.02
State-5	0.00	0.20	0.02	0.00	0.63	0.15	0.01
State-6	0.08	0.12	0.02	0.15	0.00	0.55	0.08
State-7	0.03	0.06	0.11	0.00	0.00	0.13	0.67

Table 2. Transition Probability Matrix of Hidden States for Average Deficit of Rainfall at Busan Station

	State-1	State-2	State-3	State-4	State-5	State-6	State-7
State-1	0.37	0.63	0.00	0.00	0.00	0.00	0.00
State-2	0.11	0.27	0.47	0.00	0.10	0.05	0.00
State-3	0.03	0.19	0.34	0.00	0.41	0.00	0.02
State-4	0.01	0.08	0.28	0.43	0.19	0.00	0.00
State-5	0.00	0.00	0.00	0.23	0.38	0.34	0.05
State-6	0.00	0.05	0.03	0.21	0.00	0.60	0.10
State-7	0.00	0.02	0.05	0.11	0.00	0.09	0.73

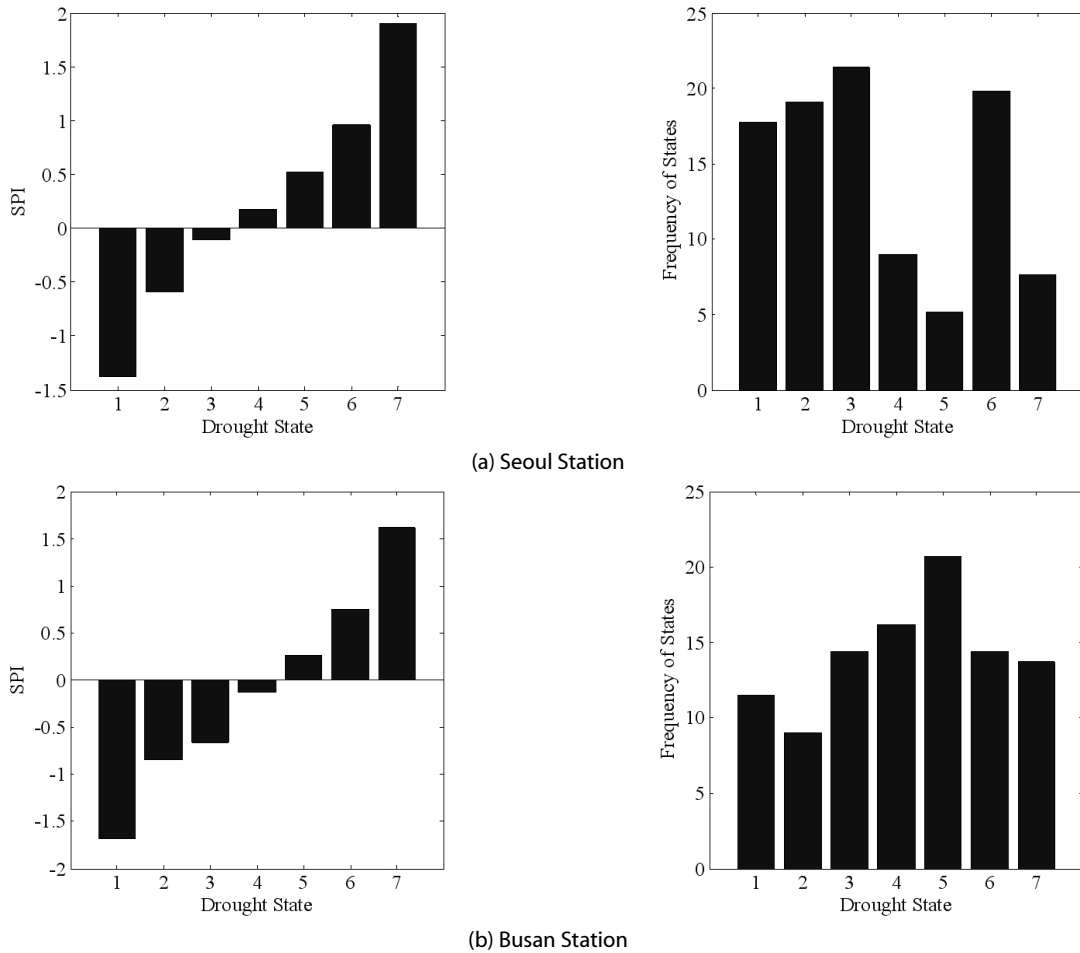


Fig. 4. SPI According to Drought State and Their Frequency

천이확률을 의미한다.

본 연구에서는 월강수량을 기반으로 분석한 은닉상태의 특성과 기존 표준강수지수와 비교하기 위해서 가뭄 등급을 7개로 동일하게 정의하여 가뭄상태를 평가하였다. 그 결과 Fig. 4와 같이 가뭄 등급별 SPI의 가뭄심도가 구분되었으며, 각 가뭄 등급별 빈도를 검토할 수 있었다. 또한, SPI를 기반으로 추정된 가뭄 등급별 천이확

률을 Tables 3 and 4에 나타내었다. Table 1의 HMM을 통해 추정된 가뭄의 천이확률을 보면 수문학적 지속성이 매우 크게 추정되는 반면 Table 3의 표준강수지수를 기반으로 추정된 가뭄의 천이확률은 수문학적 지속성이 상대적으로 크지 않다. 이는 표준강수지수의 대표적인 문제점으로 지적돼왔던 사항으로서 가뭄의 시작과 종료시점을 정량적으로 해석하는데 어려움이 있다. 즉, HMM

Table 3. Transition Probability Matrix of Drought States for SPI at Seoul Station

	State-1	State-2	State-3	State-4	State-5	State-6	State-7
State-1	0.61	0.00	0.34	0.00	0.00	0.05	0.00
State-2	0.27	0.55	0.08	0.05	0.04	0.00	0.00
State-3	0.00	0.00	0.53	0.04	0.08	0.34	0.01
State-4	0.00	0.68	0.12	0.06	0.00	0.14	0.00
State-5	0.22	0.11	0.00	0.27	0.00	0.29	0.11
State-6	0.00	0.09	0.01	0.20	0.09	0.50	0.11
State-7	0.09	0.00	0.00	0.25	0.00	0.00	0.66

Table 4. Transition Probability Matrix of Drought States for SPI at Busan Station

	State-1	State-2	State-3	State-4	State-5	State-6	State-7
State-1	0.48	0.45	0.00	0.00	0.00	0.07	0.00
State-2	0.00	0.23	0.06	0.16	0.51	0.03	0.00
State-3	0.32	0.02	0.28	0.35	0.00	0.03	0.00
State-4	0.06	0.00	0.17	0.38	0.12	0.24	0.04
State-5	0.02	0.00	0.35	0.11	0.24	0.27	0.00
State-6	0.00	0.00	0.00	0.00	0.49	0.19	0.31
State-7	0.02	0.02	0.05	0.06	0.17	0.07	0.62

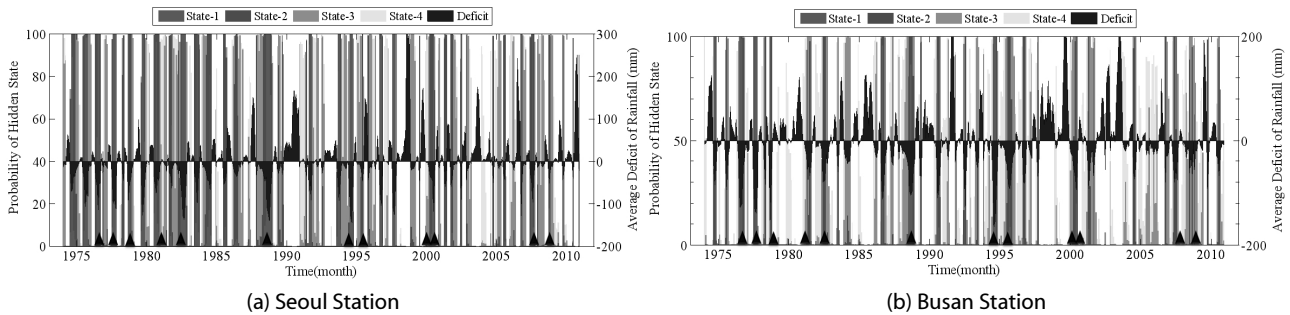


Fig. 5. Posterior Probabilities of Hidden State over Time

의 경우 극한가뭄(State 1)에서 심한가뭄(State 2)으로 천이될 확률이 0.3인데 반해 표준강수지수의 경우 0으로서 불연속적인 특성을 가지고 있다.

본 연구에서 제안하는 HMM은 Bayesian 기법으로서 매시점마다 은닉상태별 사후확률(posterior probability) 추정이 가능하며 이를 통해 불확실성을 감안한 가뭄상태를 정의할 수 있다. HMM은 특정 시기에 가뭄을 확률론적으로 해석하기보다는 은닉상태별로 확률이 부여됨으로서 가뭄의 상태를 확률론적으로 해석할 수 있다. 예를 들어, 특정시점의 은닉상태별로 확률이 주어지며 이들 확률을 모두 더하게 되면 1이 된다. 즉, 평년대비 강수량의 부족량 및 잉여량을 기준으로 서울지점과 부산지점의 가뭄 은닉상태가 정의되며 각 상태별 사후확률 계산이 가능하다. Fig. 5는 지점별 평년대

비 강수량의 부족량 및 잉여량과 가뭄의 은닉상태에 해당하는 상태 1에서 4까지의 사후확률을 시간에 따라 추정한 결과를 나타낸다. 특히, 실제 과거에 발생했던 극한가뭄 해인 1976~1978년, 1981~1982년, 1988년, 1994년~1995년, 2000~2001년, 2008~2009년을 중점적으로 검토한 결과 대부분의 가뭄상태는 은닉상태 1-2에서 사후확률이 상대적으로 크게 나타나는 것으로 확인되었다. 예를 들어, 2000년 6월의 경우에는 서울지점의 평년대비 강수부족량은 -48mm이고 은닉상태 1의 확률은 약 24%이며 은닉상태 2의 확률은 76% 이상을 나타내고 있다. 또한, 동일시점의 부산지점의 평년대비 강수부족량은 -95mm이고 은닉상태 1의 확률은 약 20%이며 은닉상태 2의 확률은 80%를 나타내고 있다. 또한, Fig. 5에서 1994년~1995년 시점의 가뭄의 상태를 보다 자세히 평가해

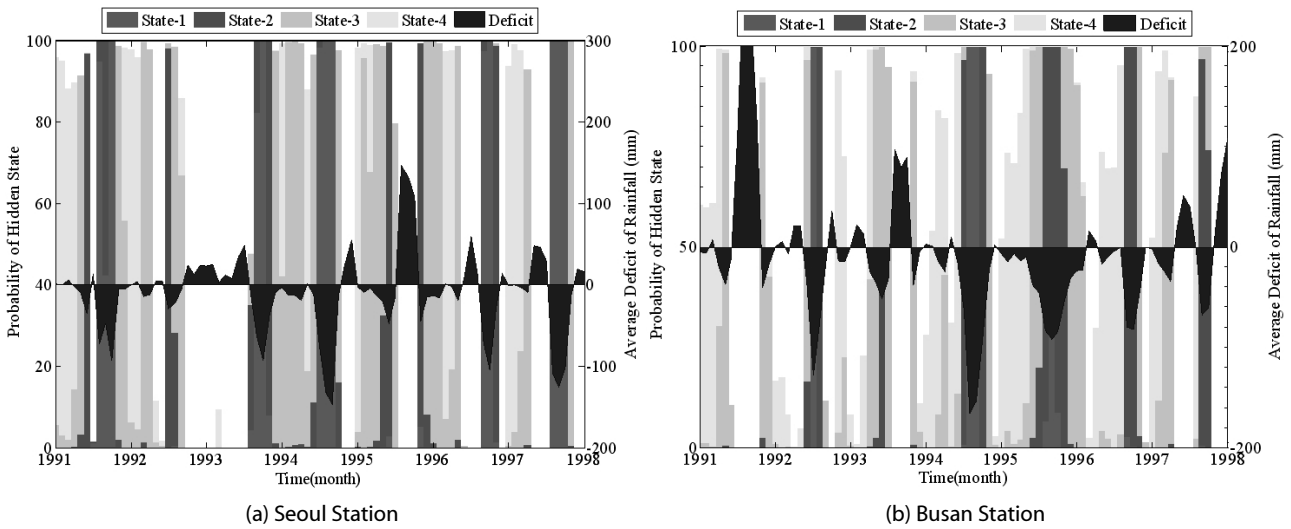


Fig. 6. Posterior Probabilities of Hidden State for a Eight-Year Block of 1991-1998 at Seoul(a) and Busan(b) Station

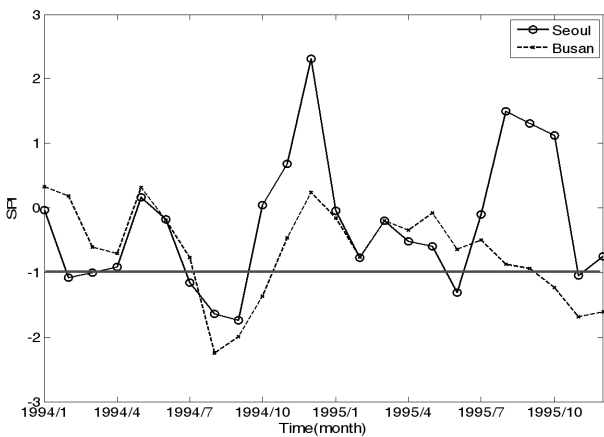


Fig. 7. SPI of Drought Period (1994-1995) for Seoul and Busan Station

보기 위해서 사후확률을 확대하여 Fig. 6에 나타내었다.

Fig. 6에서 1994년~1995년 과거가뭄을 검토해 보면, 서울지점은 1993년 상반기에는 강우의 잉여량이 발생하고 있으며 가뭄상태를 나타내는 은닉상태 1-4의 사후확률이 0으로 나타나고 있다. 1993년 하반기로 접어들면서 평년대비 3개월 누적강수량이 크게 부족하여 은닉상태 1-2의 확률이 복합적으로 크게 나타나고 있으나 1994년 상반기에는 강우의 부족량이 상대적으로 작아지면서 약한 가뭄 및 중립에 해당하는 은닉상태 3-4의 확률이 크게 증가하는 것을 확인할 수 있다. 1994년 하반기로 갈수록 다시 강우의 부족량이 크게 증가하면서 심한가뭄상태인 은닉상태 1-2의 사후확률이 100%에 근접하는 것과 같이 가뭄의 심도가 극한단계로 전이되는 것을 확인할 수 있다. 부산지점은 1994년 상반기에 다소 적은

강수부족량으로 인해 중립상태에 해당하는 은닉상태 4의 발생확률이 크게 나타났다가 1994년 중반 강수부족량이 크게 증가하면서 은닉상태 1-2의 발생확률이 급격히 증가하여 극한 가뭄단계로 전이되는 것이 확인되고 있다. 또한 1995년 부산지점의 가뭄은 서울지점에 비하여 여름철의 강수량이 적어 가뭄의 지속시간이 상당히 길어짐과 동시에 누적강수부족량이 증가하여 가뭄심도가 크게 나타나고 있다. 그러나, Fig. 7과 같이 해당 지점의 표준강수지수를 살펴보면 특정 값(-1.0)을 기준으로 1995년의 가뭄을 판단하면, 서울지점은 6월, 부산지점은 10월, 11월, 12월에만 가뭄으로 평가되고 있다.

HMM은 기본적으로 은닉상태별 방출확률밀도함수의 추정이 가능하다. 즉 평년대비 강수량의 부족량 및 잉여량은 은닉상태별로 확률밀도함수가 정의된다. Fig. 8은 서울지점과 부산지점의 가뭄상태를 나타내는 은닉상태 1-4에 대한 방출확률밀도함수의 추정결과를 나타낸다. 여기서, 좌측의 Y축은 각 은닉상태에 대한 조건부로 추정된 방출확률분포함수의 확률을 나타내며, 우측의 Y축은 전체 기간에 모든 자료의 확률밀도함수를 나타낸다. X축은 각 지점별 표준화된 강수시계열을 나타낸다. 서울지점과 부산지점의 분석기간 내 전체 자료의 확률분포는 두 지점에서 모두 정규분포에 가까운 특성을 보이고 있다. 본 연구에서는 월강수량을 기반으로 분석한 은닉상태의 방출확률밀도함수 특성과 기존 표준강수지수와 비교하기 위해서 SPI 기반의 가뭄상태별 확률분포를 검토하였다(Fig. 9 참고). 일반적으로 SPI는 특정 절단수준(-1.0)을 기준으로 가뭄이 정의되며, 서울지점의 경우에는 가뭄상태 1-3의 확률분포가 가뭄에 영향을 미치며, 부산지점의 경우에는 가뭄상태 1-4의 확률분포가 가뭄에 영향을 미치는 것으로 해석되었다. 동질성 HMM은 각

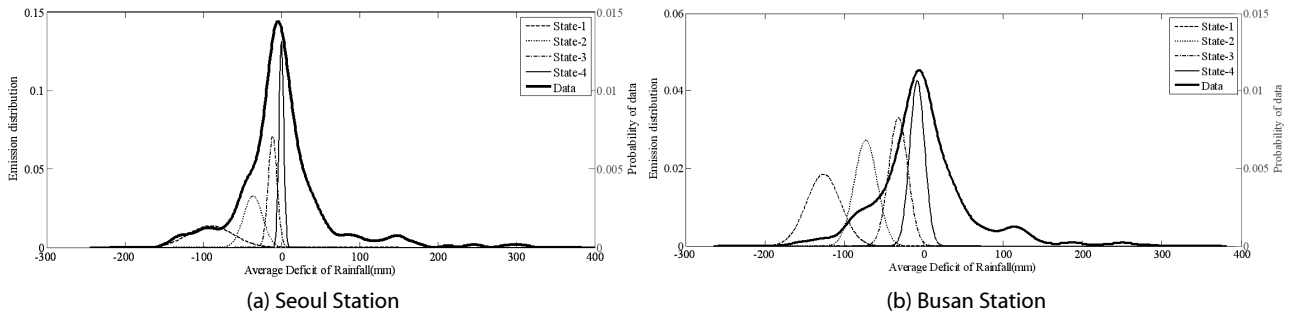


Fig. 8. Emission Distributions for all Drought States for Average Deficit of Rainfall

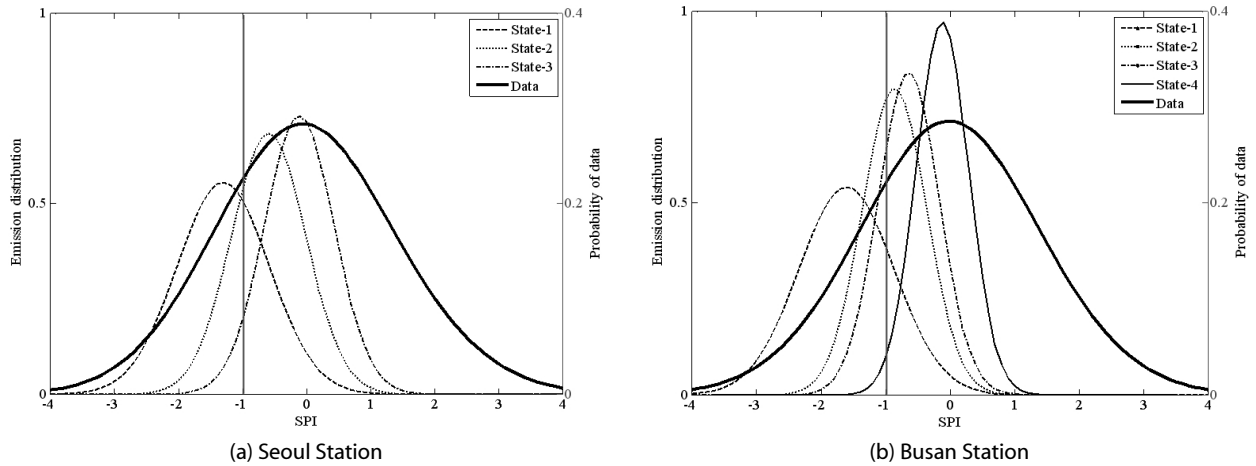


Fig. 9. Emission Distributions for all Drought States for SPI

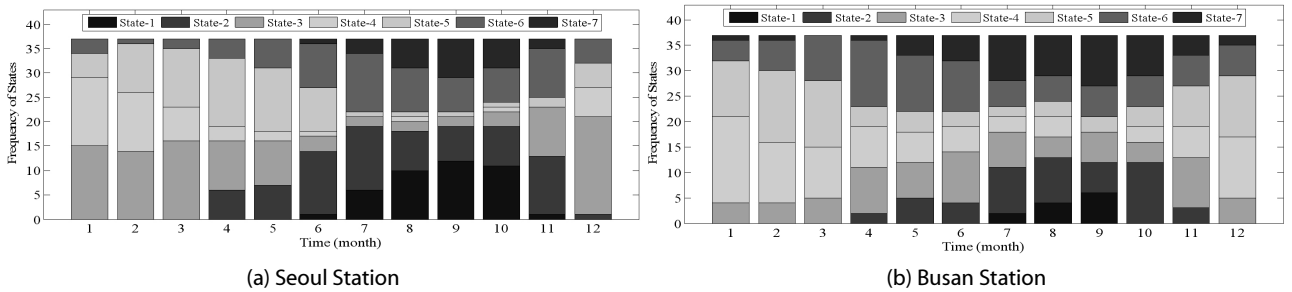


Fig. 10. Monthly Frequency Characteristic of Hidden State over Time

은닉상태별 방출확률밀도함수에 대한 통계적 특성(평균과 표준편차)을 확인할 수 있었다.

본 연구에서 정의한 가뭄은 평년대비 강수량의 부족을 의미하며, 실제 우리나라에 미치는 가뭄의 영향은 여름철과 같은 우기와 봄, 가을, 겨울과 같은 건기에 대하여 달리 해석되어야 한다. 그러므로 HMM의 각 월별 은닉상태의 발생빈도 특성을 검토하여 각 은닉상태의 월별특성을 해석하고자 한다. Fig. 10은 서울지점과 부산지점에 대한 HMM의 월별 은닉상태에 따른 발생빈도 특성을 나타낸다. 서울지점의 경우 은닉상태 1에 해당하는 가뭄은 여름철에서 늦가을

(6월-11월)까지 지속되는 6개월 동안(장기간)의 강수부족으로 인한 것으로 해석된다. 반면, 부산지점의 경우 은닉상태 1에 해당하는 가뭄은 오직 여름철에서 초가을(7월-9월)까지 지속되는 3개월 동안의 강수부족이 원인인 것으로 해석된다. 은닉상태 2에 해당하는 가뭄은 서울지점과 부산지점 모두 4월부터 겨울까지 지속된 장기간의 강수부족이 주된 원인으로 확인된다. 은닉상태 3과 은닉상태 4에 해당하는 가뭄은 대체로 1년 내내 강수부족이 원인으로 작용한다. 특히 서울지점의 은닉상태 3에 해당하는 강수부족은 겨울에서 봄 시기(12월-3월)의 가뭄빈도가 상대적으로 높게 나타나는 것으로

확인되며, 반면 부산지점은 은닉상태 4에 해당하는 강수부족은 겨울에서 봄 시기(12월-5월)의 가뭄빈도가 상대적으로 높게 나타나 는 것으로 확인된다.

4. 결론

본 연구에서는 기상학적 측면에서 가뭄을 정의하고 기존의 가뭄 지수 기반의 가뭄판단 시 적용되어져 왔던 특정한 기준값에 근거한 해석이 아닌, 강수량 자료에 내재되어 있는 특징을 최대한 활용할 수 있는 HMM 기반의 확률론적 가뭄해석기법을 제안하였다. 또한, 기상학적 측면에서 가뭄을 평가하는 데 있어, 기존 표준강수지수를 기반으로 가뭄을 평가하는데 있어서 문제점을 객관적으로 검토하였다. 그 결과, 기존 표준강수지수를 이용하여 가뭄을 판단할 경우에 가뭄의 수문학적 지속성 결여문제를 천이확률을 통해 확인할 수 있었다. 이와 더불어 특정 기준값 설정으로 인해 가뭄상태를 규정하는데 있어서 신뢰성이 결여될 개연성이 큰 것으로 나타났다. 반면 본 연구에서 제안한 HMM 기반의 확률론적 가뭄해석기법은 이러한 문제점을 상당부분 보완할 수 있었다.

월강수량을 기반으로 분석한 은닉상태별 가뭄의 천이확률특성은 기존의 표준강수지수를 적용하였을 경우와 비교하여 수문학적 지속성이 매우 크게 나타나는 것으로 확인되었다. 즉, HMM을 활용한 가뭄해석은 기존 가뭄해석 시 한계점으로 지적되었던 가뭄의 시작과 종료시점을 정량적으로 해석하는 데 있어 보다 효율적으로 가뭄상태의 지속성을 반영하여 해석할 수 있는 장점을 확인할 수 있었다. 또한, 기존의 표준강수지수를 기반으로 가뭄등급을 결정하였을 경우와 비교하여 HMM 기반의 가뭄해석은 특정지점의 각 은닉상태별 사후확률 계산이 가능하다는 장점이 있으며, 즉 각 지점별 강수발생 특성의 불확실성을 고려한 가뭄해석에 유리한 장점을 제시하였다. 더불어 HMM의 각 월별 은닉상태의 발생빈도 특성을 해석한 결과, 각 가뭄상태별로 시간적 분포 특성을 평가할 수 있으며, 각 가뭄의 은닉상태별로 강수부족현상이 발생한 주된 시점(계절)의 파악이 가능하게 되었다.

본 연구에서 제시한 HMM 기반의 확률론적 가뭄해석기법은 기존의 다양한 가뭄지수를 적용한 가뭄해석과 비교하여 가뭄판단의 명확한 기준 제시 및 가뭄발생의 원인을 규명하는 데 있어 보다 효율적인 방안이다. 이와 같이 가뭄을 정의하는 데 있어 수문학적 지속성을 재현함과 동시에 불확실성을 감안한 해석은 필수적인 것이라고 판단된다. 향후 연구에서는 기후변화에 따른 월강수량 발생패턴에 영향을 미치는 원인을 고려한 가뭄해석을 수행하고자 하며, 즉 대규모 기상변동성 현상과의 복합적인 연계성을 고려하여 수문기상학적 가뭄해석이 필요할 것으로 사료된다.

감사의 글

이 논문 또는 저서는 2010년 교육과학기술부의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(한국연구재단-2010-220-D00083)

References

- Dracup, J. A., Lee, K. S. and Paulson, E. G. Jr. (1980). "On the definition of droughts." *Water Resources Research*, Vol. 16, No. 2, pp. 297-302.
- Hughes, J. P. and Guttorp, P. (1994). "A class of stochastic models for relating synoptic atmospheric patterns to regional hydrologic phenomena." *Water Resour. Res.*, Vol. 30, No. 5, pp. 1535-1546.
- Kwon, H. H., Kim, T. J., Hwang, S. H. and Kim, T. W. (2013a). "Development of daily rainfall simulation model based on homogeneous hidden markov chain." *Journal of the Korean Society of Civil Engineers*, Vol. 33, No. 5, pp. 1861-1870 (in Korean).
- Kwon, J., Ahn, J. H. and Kim, T. W. (2013b). "Development on classification standard of drought severity." *Journal of Korean Water Resources Association*, Vol. 46, No. 2, pp. 195-204 (in Korean).
- Loaiciga, H. and Leipnik, R. (1996). "Stochastic renewal model of low-flow streamflow sequences." *Stochastic Hydrology and Hydraulics*, Vol. 10, No. 1, pp. 65-85.
- Mallya, G., Shivam Tripathi, A. M., Kirshner, S. and Rao S. Govindaraju, M. (2013). "Probabilistic assessment of drought characteristics using hidden markov model." *Journal of Hydrologic Engineering*, Vol. 18, No. 7, DOI: 10.1061/(ASCE)HE.1943-5584.0000699.
- Matsuyama, Y. (2011). "Hidden markov model estimation based on alpha-EM algorithm: Discrete and continuous alpha-HMMs." *International Joint Conference on Neural Network*, San Jose, Vol. 7, No. 5, pp. 808-816.
- Mishra, A. K. and Desai, V. R. (2005). "Drought forecasting using stochastic models." *Stochastic Environmental Research and Risk Assessment*, Vol. 19, No. 5, pp. 326-339.
- Mishra, A. K. and Singh, V. P. (2010). "A review of drought concepts." *Journal of Hydrology*, Vol. 391, No. 1-2, pp. 202-216.
- Sen, Z. (1980). "Statistical analysis of hydrologic critical droughts." *Journal of the Hydraulics Division*, Vol. 106, No. 1, pp. 99-115.
- Steinemann, A. (2003). "Drought triggers: A stochastic approach to evaluation." *Journal of the American Water Resources Association*, Vol. 39, No. 5, pp. 1217-1234.
- Yevjevich, V. (1967). "On objective approach to definitions and investigations of continental hydrologic droughts." *Hydrology Paper*, No. 23, Colorado State University, Fort Collins, pp. 4-18.