



## A development of hierarchical bayesian model for changing point analysis at watershed scale

Kim, Jin-Guk<sup>a</sup> · Kim, Jin-Young<sup>a</sup> · Kim, Yoon-Hee<sup>b</sup> · Kwon, Hyun-Han<sup>a\*</sup>

<sup>a</sup>Department of Civil Engineering, Chonbuk National University

<sup>b</sup>Daegu-Gyeongbuk Regional Division, Korea Water Resources Corporation

Paper number: 16-085

Received: 27 September 2016; Revised: 15 November 2016 / 19 December 2016; Accepted: 19 December 2016

### Abstract

In recent decades, extreme events have been significantly increased over the Korean Peninsula due to climate variability and climate change. The potential changes in hydrologic cycle associated with the extreme events increase uncertainty in water resources planning and designing. For these reasons, a reliable changing point analysis is generally required to better understand regime changes in hydrologic time series at watershed scale. In this study, a hierarchical changing point analysis approach that can apply in a watershed scale is developed by combining the existing changing point analysis method and hierarchical Bayesian method. The proposed model was applied to the selected stations that have annual rainfall data longer than 40 years. The results showed that the proposed model can quantitatively detect the shift in precipitation in the middle of 1990s and identify the increase in annual precipitation compared to the several decades prior to the 1990s. Finally, we explored the changes in precipitation and sea level pressure in the context of large-scale climate anomalies using reanalysis data, for a given change point. It was concluded that the identified large-scale patterns were substantially different from each other.

**Keywords:** Hierarchical bayesian model, Changing point analysis, Climate variability, Annual precipitation

## 유역단위에서의 연강수량의 변동점 분석을 위한 계층적 Bayesian 분석기법 개발

김진국<sup>a</sup> · 김진영<sup>a</sup> · 김윤희<sup>b</sup> · 권현한<sup>a\*</sup>

<sup>a</sup>전북대학교 토목공학과, <sup>b</sup>한국수자원공사 대구경북지역본부

### 요 지

최근 기후변화에 따른 기상변동성 증가로 기존 한반도의 기상패턴과 다른 이상강우 현상이 증가하고 있다. 이상강우현상에 따른 수문패턴의 변화는 수자원 계획을 수립하는데 있어 불확실성을 가중시키고 있다. 이러한 점에서 수문 시계열의 변화양상을 효과적으로 인지할 수 있으며, 유역 단위에서 일관된 변화를 평가할 수 있는 변동점 분석 개발이 필요하다. 이에 본 연구에서는 기존 변동점 분석방법에 계층적 베이지안(Hierarchical Bayesian) 기법을 연계하여 유역단위에서 계층적 변동점 분석이 가능한 모형을 개발하였다. 우리나라에 40년 이상 관측된 기상청 강수자료를 활용하여 연강수량 자료를 구축하였으며, 본 연구를 통해 개발된 모형의 적합성을 평가하였다. 분석결과, 1990년대의 강수자료의 변화 양상을 정량적으로 확인할 수 있었으며, 과거에 비해 강수의 증가 특성을 확인할 수 있었다. 최종적으로 추정된 수문자료의 변화시점 전후의 재해석자료를 이용하여 한반도 주변의 강수량과 해수면기압의 Anomaly를 분석해본 결과 변동점을 기준으로 강수량과 해수면기압의 명확한 차이를 확인하였다.

**핵심용어:** 계층적 Bayesian 모형, 변동점 분석, 기상변동성, 연강수량

\*Corresponding Author. Tel: +82-63-270-2426  
E-mail: hkwon@jbnu.ac.kr (H.-H. Kwon)

## 1. 서론

전 세계적으로 기후변화(climate change)에 대한 우려가 현실화되면서, 각 나라에서는 보다 정확한 미래 기후 예측을 위한 연구가 활발히 진행되고 있다. 이러한 기후변화는 생태계뿐만 아니라 농업, 수자원 등 다양한 분야에 직접적인 영향을 미치고 있다(Yun and Lee, 2001; Lee et al., 2009). 최근 우리나라에서는, 기후변화의 영향에 의해 기존 한반도의 기상 패턴과 다른 이상강우현상이 지속적으로 발생하고 있으며 기상변동성(climate variability) 또한 동시에 커지고 있다. 이러한 기상변동성은 수자원 계획을 수립하는데 있어 불확실성을 가중시키고 있으며, 최근 전 세계적인 가뭄에서 보듯이 궁극적으로 사회·경제적인 영향도 상당히 크다(Lee et al., 2011).

수문기상자료의 통계적 특성 변화 분석분야의 경우 현재 기후변화와 관련하여 매우 중요한 연구 분야중 하나이며, 그중에서도 수문기상자료의 변동점 분석(changing point analysis)과 관련된 연구는 수자원의 변동성을 평가하고 예측하는데 매우 중요하다 하겠다(Lee et al., 2014). 수문자료의 변동점 분석은 기본적으로 수문자료의 변동성을 파악함과 동시에 이와 연관된 대규모 기상 변동성을 이해하는데 있어서 기본적인 정보를 제공할 수 있다. 즉, 기후값의 평균이나 분산의 변화가 발생한 시점을 직접적으로 파악할 수 있을 뿐만 아니라 관측지점의 이동이나 관측기기의 변동 등으로 인한 자료의 비균질성이 나타나는 시점을 찾아내고 기후변화의 확률적 영향을 파악하는 데 활용 될 수 있다(Solow, 1987; Cox et al., 2002; Lund and Reeves, 2002; Elsner et al., 2004). 최근 많은 기후변화 연구들이 기온의 변화를 평가하기 위하여 변동점 분석기법을 이용한 연구가 진행되고 있지만, 신뢰성 있는 변동점을 도출하는데 어려움을 겪고 있다(Jung et al., 2002).

관측된 시계열 자료의 특성이 어느 시점에서 이전의 흐름과 다르게 변화하는 변동점 문제는 환경과학, 생물학, 경제학, 수문학, 기상학 등 매우 많은 분야에서 발생한다. 변동점 문제는 일반적으로 최우도(maximum likelihood) 방법 혹은 베이지안(Bayesian) 방법과 같은 모수적 방법, 비모수적 방법, 또는 의사결정론적 방법 등에 의해 수행 된다(Kim and Son, 2004).

Bayesian 변동점 분석에 있어 일반적인 이론의 개발 및 보완은 Carlin et al. (1992)에 의해 수행되었으며, 크게 단일 BCP (single Bayesian change point) 기법과 다중 BCP (multiple Bayesian change point) 기법으로 구분되어 연구가 수행되었다. 단일 BCP 기법은 주로 공액 사전분포(conjugated prior distribution)를 사용함으로써 사후확률의 계산에 있어 간단

한 면이 있으나, 단일 BCP 기법은 1개의 대상지점에 대한 1개의 변동점을 찾아내기 때문에 유역내 여러 지점에 대한 적용이 불가하다. 이와 더불어 단일 BCP 기법의 경우 먼저 주관적으로 자료의 구간을 분리한 다음 변동점을 탐색해야 하는 문제가 있어 사실상 신뢰성 있는 분석결과를 보장하기 어려운 측면이 있다. 이러한 점에서 본 연구에서는 연속된 시계열로 구성된 수문자료의 연관성을 효과적으로 고려함과 동시에 유역내 다수의 수문관측지점에 대한 변동점의 위치 및 크기를 정량적으로 평가할 수 있는 계층적 Bayesian 변동점 분석기법을 개발하였다.

Bayesian 방법과 대응하는 대표적인 방법들로는 CBS (circular binary segmentation; Olshen and Venkatraman, 2004)와 BP (break points; Bai and Perron, 2003) 기법 등이 있다. Bai and Perron의 BP 방법은 여러 조각(segments)으로 구성된 최적의 분리구간을 결정하기 위해 동적 프로그래밍 알고리즘을 이용하였다. Olshen and Venkatraman의 CBS 방법은 BS (binary segmentation; Sen and Srivastava, 1975) 방법의 개선된 형태로, BS가 단일 변화시점 검정에 기반하고 있어 큰 변화구간 사이의 작은 변화구간을 찾는 데는 어려움이 있는데, 이러한 문제를 보완한 방법이다. 이 방법은 임의의 연속 시계열에서 변화시점이 없다는 귀무가설을 검정하기 위해 우도 비율(likelihood ratio)을 검정통계량으로 사용하는 알고리즘을 사용하고, 만약 이 검정에서 귀무가설이 기각된다면 해당 연속계열은 다시 분할되고 분할된 조각(sub-segments)에서 추가로 변화가 나타나지 않을 때까지 반복된다. 그러나 BP와 달리 CBS는 주어진 변화시점 수에 대한 최적의 변화시점의 위치를 찾는 데 어려움이 있다고 알려져 있다(Barry and Hartigan, 1993).

본 연구에서는 국내 권역별 40년 이상 관측된 국내 기상청 강우관측소를 선정하여 수문자료에 대해서 개발된 모형의 적합성을 검토하였으며, 최종적으로 본 연구에서 제안한 계층적 Bayesian 변동점 분석기법을 활용하여 한반도 권역별 수문자료에 대한 정량적인 변동점의 위치(location), 크기(magnitude)를 찾아 한반도 강수의 경향성 및 변동점을 추정하고, 추정된 변화시점에 대한 미국 NOAA의 재해석자료를 이용하여 강수량과 해수면기압에 대한 관계를 분석하였다.

본 논문의 구성은 다음과 같다. 1장에서는 논문의 배경 및 목적에 대해 언급하였으며, 2장에서는 기본적인 변동점 분석 기법 및 본 연구에서 이용한 Bayesian 변동점 분석 모형, 본 연구에서 개발된 계층적 Bayesian 변동점 분석기법에 대해 설명하였다. 3장에서는 한반도 권역별 강우관측소에서 40년 이상 관측된 연강수량에 대해 계층적 Bayesian 변동점 분석

기법 적용하여 한반도 수문패턴에 대한 변동점 및 경향성 결과를 수록하였으며, 마지막으로 결론 및 토의를 4장에 수록하였다.

## 2. 연구방법

본 장에서는 기존 변동점 분석시 사용되는 기법에 대한 간단한 소개 및 본 연구에서 개발된 계층적 Bayesian 변동점 분석모델의 이론적인 배경을 수록하였다. 변동성 분석에 대한 이론적인 부분은 기존 변동성 분석과 관련된 논문들에 상세히 기술되어 있어, 본 논문에서는 변동성 분석과 관련된 가장 핵심적인 부분만을 간략히 기술하였다.

### 2.1 변동점 분석

시계열 자료의 평균에 대한 변동성을 검정하는 기법은 대표적으로 Mann-Whitney U검정(Mann and Whitney, 1947), Sign 검정(Hollander and Wolfe, 1973), T 검정, Modified T 검정(Kite, 1977) 등 통계학적 추론에 이론적 근거를 둔 기법들이 주로 사용되고 있다. 이와 같은 통계학적 추론에 이론적 바탕을 둔 변동성 분석 기반의 기법들은 자료의 통계적 특성을 분석하는데 있어 유용하지만, 통계적 검정에 있어 검정을 위해 수립한 귀무가설(null hypothesis)을 채택 또는 기각의 여부만 제시된다. 즉, 변동점 문제에 있어서는 특정 신뢰구간에 대해 동질성이 '있다 또는 없다'는 결론만을 제시하게 되므로 정량적인 변동성 분석결과를 도출하지 못하는 단점이 있다(Lee et al., 2014). 반면 Bayesian 변동점 추정 기법은 연속된 수문기상자료의 변화확률을 산정하여 변동점을 추정하는 방법이다(Barry and Hartigan, 1993). 연속적인 시계열에서의 변동점은 변화 전후 상태에 따라 결정이 되는 상대적인 개념이기 때문에 확정론적 방법으로 변동점을 결정하는 것은 어렵다. 이러한 관점에서 상대적인 변화시점의 확률을 계산하여 변동점을 결정하는 베이시안 과정은 변동점을 보다 합리적으로 추정할 수 있는 방법이다(Hwang et al., 2010).

본 연구에서 적용하고자 하는 계층적 Bayesian 변동점 분석기법은 각각의 자료가 나타내는 기존 변동점 분석 기법에 Bayesian 통계기법을 유역내 다지점에 대해서 도입함에 따라 정량적인 변동점의 위치(location), 크기(magnitude)를 산정할 수 있는 장점을 가지고 있으며, 기존의 단일지점의 적용되는 경향성 분석(trend analysis) 연구와 차별된다고 볼 수 있다.

### 2.2 계층적 Bayesian 변동점 추정기법(Hierarchical Bayesian Changing Point, HBCP)

Bayesian 이론(Bayes's theorem)이란 두 확률변수의 사전 확률(prior probability)과 사후확률(posterior probability) 사이의 관계를 나타내는 것으로 새로운 사후 근거가 제시될 때 신뢰값이 어떻게 갱신될 것인가를 말해준다. 예를 들어, 새로운 정보가 없을 때 A라는 사건이 일어날 확률을  $P(A)$ 라고 하고, B라는 새로운 정보가 주어졌을 때 A가 일어날 확률을  $P(A|B)$ 라고 하면,  $P(A)$ 는 사전확률,  $P(A|B)$ 는 사후확률을 말한다.

Bayesian을 통한 매개변수 추정기법은 기존 방법들(최우도법, 모멘트법, 확률가중모멘트법)과는 다르게 매개변수를 하나의 확률변수로 취급한다. 즉, 매개변수가 단일 값이 아닌 확률분포의 형태로 부여되며 최종적으로 매개변수의 사후분포(posterior distribution)를 추정하는데 목적을 두며, Bayes 정리를 기반으로 한다.

Bayes 정리에 대한 기본적인 개념은 다음과 같다. 주어진 확률변수  $\theta$ 와  $y$ 가 있을 때, 두 확률변수들의 결합확률분포(joint probability distribution)는 Eq. (1)과 같이 사전분포(prior distribution)  $p(\theta)$ 와 우도(likelihood)  $p(y|\theta)$ 의 곱으로 표현할 수 있다. Bayes 정리에서 확률사건  $\theta$ 와  $y$ 의 조건부 확률과 주변확률의 관계는 Eq. (2)와 같으며, Eq. (3)에서 좌변  $p(\theta|y)$ 는 사후분포를 나타내며 사전분포와 우도  $p(y|\theta)$ 의 곱으로 추정될 수 있다(Gelman et al., 2003).

$$p(\theta, y) = p(\theta)p(y|\theta) \quad (1)$$

$$p(\theta|y) = \frac{p(\theta, y)}{p(y)} = \frac{p(\theta)p(y|\theta)}{p(y)} \quad (2)$$

$$p(\theta|y) \propto p(y|\theta)p(\theta) \quad (3)$$

여기서,  $p(\theta)$ 는  $\theta$ 의 사전확률 또는 경계 확률이며, '사전'이라는 의미는 사건  $y$ 에 관한 기본적인 정보를 의미한다.  $p(\theta|y)$ 는  $y$ 가 주어졌을 때  $\theta$ 의 조건부 확률을 말하며,  $y$ 의 특정 값에 의해 결정되기 때문에 사후확률이라고도 불린다.  $p(y|\theta)$ 는  $\theta$ 가 주어졌을 때  $y$ 의 조건부 확률을 말하며,  $p(y)$ 는  $y$ 의 주변확률로서 정규화 상수의 역할을 한다.

사후평균에 대해서도 동일한 개념으로 해석할 수 있다. 분석결과에서 제시한 변화확률의 해석에 있어서 사후평균이 변화의 시점 전후로 장기적인 지속성이 없는 경우 이를 일시적 사후확률의 변화로 보고 이 시점을 변동시점으로 해석한다.

그리고 사후평균이 변화시점 전후로 장기적인 지속성이 있는 경우 이를 장기평균변화에 의한 장기적 사후확률의 변화로 보고 이 시점을 '변동점'으로 해석한다(Lee et al., 2011).

Bayesian 변동점 분석기법은 여러 분야에서 관측자료의 변동점을 추정하는데 활용되고 있다. Bayesian 방법은 일반적으로 Markov Chain Monte Carlo (MCMC)를 이용한다. MCMC 방법에서 널리 사용되는 알고리즘으로는 메트로폴리스-해스팅스(Metropolis-Hastings; MH) 알고리즘(Metropolis et al., 1953; Hastings, 1970)과 깁스 샘플링(Gibbs sampling) 방법(Geman and Geman, 1984)이 있다. 깁스 샘플링은 고차원의 경우에 사용하는 기법으로 완전 조건부 확률분포(full conditional distribution)를 이용하여 두개 이상의 확률 변수의 연속적인 표본을 생성하는 매우 효율적인 알고리즘이다.

일반적으로 기존 연구에서는 변동점 분석 시 매개변수 불확실성을 고려하지 못하는 단점이 존재한다. 이는 불확실성을 정량적으로 해석하는데 어려움이 있어 추정되는 변동점의 신뢰성 또한 확보하지 못하는 실정이다. 이러한 이유로 본 연구에서는 변동점 분석시 불확실성을 정량적으로 해석이 가능한 계층적 Bayesian 모형기반의 다지점 변동점 분석기법 모형을 개발하고자 한다.

본 연구에서 제시하고자 하는 모형과 기존 해석 절차와 다른 점은 2가지로 요약할 수 있다. 첫째, 기존 변동점 분석 절차는 단일 지점을 대상으로 변동점을 찾아내는 반면, 본 연구에서는 유역내의 각 지점별로 사전분포를 부여하여 지점별 및 지역을 대표할 수 있는 변동점을 동시에 분석이 가능한 방법이다. 이는 유사한 통계적 특성을 가진 자료집단들이 서로 정보를 교환함으로써 추정되는 매개변수의 불확실성을 저감할 수 있는 계층적 Bayesian 모형을 활용하는데 목적이 있다(Kwon et al., 2013). 둘째, 본 연구에서는 Bayesian 모형을 기반으로 매개변수의 불확실성을 정량화 할 수 있는 모형을 개발함으로써 추정되는 변동점의 신뢰성을 개선할 수 있도록 개선하였다.

본 연구에서 활용된 모형은 다음과 같다. 먼저 본 연구에서 활용된 연강수량(annual precipitation)자료에 대한 확률분포형(probability distribution)을 결정하기 위하여 BIC (Bayesian Information Criterion) 통계량을 기준으로 확률분포형을 선택하였다. BIC 통계량은 우도(likelihood)와 매개변수의 개수를 고려하여 추정되며, 우도가 크고 매개변수의 개수가 적은 경우에 우수한 모형으로 선택된다. 분포형별로 BIC를 추정하기 위한 대수우도함수는 다음과 같이 일반화 할 수 있다.

$$L = \ln \left( \prod_{i=1}^n f(x_i | \theta) \right) \quad (4)$$

여기서,  $f_k$ 는 매개변수  $\theta$ 를 가지는 확률밀도함수를 나타낸다.  $x$ 는 벡터형태의 자료를 의미하며,  $n$ 은 자료의 수를 나타낸다.

일반적으로 우도만을 가지고 최적분포형을 선택하는 경우 모분포의 자유도가 후보모델의 자유도보다 낮은 경우 잘못된 모분포를 선택할 확률이 높은 단점이 있다(Akaike, 1974). 따라서 본 연구에서는 우도, 매개변수 개수, 자료 수 등이 적절히 고려되는 BIC를 활용하였으며, 최소의 BIC값을 갖는 확률분포형을 가장 적합한 분포형으로 선택하였다. BIC 통계량에 대한 일반적인 식은 다음과 같다(Findley, 1991).

$$BIC = -2\ln(L) + k\ln(n) \quad (5)$$

여기서,  $k$ 는 확률분포형의 매개변수 수,  $n$ 은 자료의 개수를 나타낸다.

본 연구에서는 최종적으로 계산된 BIC 통계량을 바탕으로 대수정규분포(Lognormal distribution)가 자료의 확률분포형으로 적합하다고 판단하였으며, 확률밀도함수(probability density function) 및 누가확률밀도함수(cumulative density function)는 아래 Eqs. (6)~(7)과 같다.

$$f(x; \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}, x > 0 \quad (6)$$

$$F(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \int_0^x \frac{e^{-\frac{(\ln t - \mu)^2}{2\sigma^2}}}{t} dt \quad (7)$$

이때,  $\mu$ 는 위치매개변수(location parameter),  $\sigma$ 는 규모매개변수(scale parameter)를 의미한다. 따라서 연강수량 자료가 대수정규분포를 따를 때, 본 연구에서 활용된 모형은 아래 Eq. (8)과 같이 구축할 수 있다.

$$Y(\text{annual series}) = \text{Lognormal}(\mu(t, k), \sigma(k)) \quad (8)$$

본 연구에서는 연강수량의 변동성을 보다 정확하게 분석하기 위해 위치매개변수를 비정상성 모형으로 구축하였다. 즉, 평균의 분포를 변동점  $C_k$ 와 연계시킬 수 있는 다음 Eqs. (9)~(10)을 통하여 시계열로부터 평균의 변화를 효과적으로 분석할 수 있다.

$$p(\mu(t, k)) \sim \begin{cases} \alpha_{1_k}(C_k - t), & C_k > t \\ \alpha_{2_k}(t - C_k), & C_k < t \end{cases} \quad (9)$$

$$p(\sigma(k)) \sim \sigma_1^{-2} \quad (10)$$

여기서,  $t$ 는 관측된 자료의 개수이며,  $k$ 는 본 연구에서 활용된 각 유역의 기상청 강우관측소를 의미한다. 이때 위치매개변수의 경우 앞서 언급하였듯이 비정상성분석 즉 연도별로 변동성 분석이 가능하도록 Bayesian GLM (generalized linear model) 형태로 모형을 구성하였다. Bayesian GLM 모형을 구축함으로써 보다 정확한 변동점 분석이 가능할 수 있도록 하기 위함이다.

매개변수들의 사전분포( $p(\theta)$ )는 자료를 기반으로 하는 정보적 사전분포(informative priors)와 자료에 크게 상관하지 않는 비정보적 사전분포(noninformative priors)가 있으며, 비정보적 사전분포로는 정규분포, 균일분포, 지수분포가 대표적으로 활용된다(Gelman et al., 2004, Lee et al., 2010). 본 연구에서는 비정보적 사전분포를 활용하여 사후분포를 추정하였다. 앞서 언급한 바와 같이 사후분포는 사전분포와 우도함수의 곱으로 계산되며, 매개변수 추정을 위한 자료가 충분한 경우 우도함수가 비교적 정확하게 추정이 가능하기 때문에 사후분포 추정에 있어서 주도적인 역할을 하게 되며 사전분포의 중요성은 상대적으로 작아진다. 만약 사전분포에 대한 정보가 명확하지 않은 상태에서 상대적으로 좁은 범위를 갖는 사전분포를 적용하게 되면 사후분포 추정에 있어 신뢰성이 결여되는 문제점이 발생한다. 즉, 우도가 최대가 되는 지점으로 사후분포의 중심이 이동하지 못하고 멈추게 되며, 실제 매개변수의 분포와는 상이한 결과를 초래할 수 있다. 이러한 점에서 우도추정을 위한 자료가 비교적 많은 경우에 비정보적 사전분포를 적용하는 것이 왜곡된 매개변수 추정을 방지하는데 유리하다.

Eq. (11)에서  $\alpha$ 는 변동점 전후의 평균을 의미하는 회귀매개변수로서 대수정규분포의 공액분포(conjugate prior)인 정규분포를 사전분포로 결정하였다. 공액분포는 자료의 분포와 동일한 확률분포를 사전분포로 정의하는 방법으로 최종적으로 추정된 사후분포가 가정한 분포형태를 갖도록 유도하는 확률통계 개념으로서, 매개변수의 수렴 등 안정성을 확보하는데 유리한 장점이 있다(Gelman et al., 2004). Eqs. (11)~(16)에서  $N$ 은 정규분포를,  $U$ 는 균등분포를 나타낸다. 본 연구에서는 앞서 언급하였듯이 계층적 Bayesian 모형을 활용하였으며, 각 매개변수의 사전분포는 아래 Eqs. (11)~(16)과 같이 나타낼 수 있다.

$$\alpha_{1_k} \sim N(0.0, 0.0001) \quad (11a)$$

$$\alpha_{2_k} \sim N(0.0, 0.0001) \quad (11b)$$

$$\sigma_1 \sim U(0.001, 1000) \quad (12)$$

$$C_k \sim N(\mu_c, \sigma_c) \quad (13)$$

$$\mu_c \sim U(5, 40) \quad (14)$$

$$\sigma_c \sim \sigma_{1c}^{-2} \quad (15)$$

$$\sigma_{1c} \sim U(0.001, 1000) \quad (16)$$

본 연구에서 활용한 사후분포의 형태는 Bayes 정리에 의해 다음 Eq. (17)과 같이 나타낼 수 있다(Gelman et al., 2003).

$$p(\theta|y) = \frac{p(\theta, y)}{p(y)} \propto p(y|\theta) \cdot p(\theta) \quad (17)$$

여기서,  $\theta = [\mu(t, k), \sigma(k), \alpha_{1_k}, \alpha_{2_k}, \sigma_1, C_k, \mu_c, \sigma_c, \sigma_{1c}]$ 는 본 연구에서 추정되는 매개변수를 나타내며,  $p(\theta)$ 는 매개변수 사전분포를,  $p(y|\theta)$ 는 우도함수를 나타내며 Eq. (18)과 같이 표현할 수 있다.

$$p(y|\theta) = \prod_{t=1}^T \prod_{k=1}^K LN(y_t | \mu(t, k), \sigma(k)) \quad (18)$$

여기서,  $T$ 는 자료연수를 나타내며,  $K$ 는 기상청 강우관측소의 개수를,  $LN$ 은 대수정규분포를 의미한다. 위에서 정의된 우도함수와 매개변수들의 사전분포들을 Eq. (19)에 대입시킴으로써 매개변수들의 사후분포를 추정할 수 있다.

$$p(\theta|y) \propto \prod_{t=1}^T \prod_{k=1}^K LN(y_t | \mu(t, k), \sigma(k)) N(\mu(t, k) | \alpha_{1_k}(C_k - t) + \alpha_{2_k}(t - C_k), 10^{-4}) N(C_k | \mu_c, \sigma_c) U(\mu_c | 5, 40) G(\sigma_c | 0.1, 0.1) G(\sigma(k) | 0.1, 0.1) \quad (19)$$

여기서,  $G$ 는 Gamma분포를 나타낸다. Eq. (14)의 모든 매개변수에 대한 적분을 통해 직접적으로 추정하는 것은 불가능하므로, 본 연구에서는 Bayesian MCMC 기반의 해석모형을 개발하였다. Bayesian MCMC 기법은 주어진 다변량 확률분포

가 복잡하여 이를 따르는 IID (Independent and identically distributed) 난수를 얻을 수 없는 경우에 사용가능한 기법으로서 IID 난수 대신 Markov Chain 난수를 추출하여 사용한다. MCMC 기법 중 본 연구에서는 깃스샘플링을 이용하였다. 깃스샘플링은 이미 오래전부터 물리학 분야에서 복잡한 수식의 해를 찾기 위해서 사용되던 방법이였으나, Gelfand and Smith (1990)에 의하여 Bayesian 모형에 도입되어 이후 Bayesian 통계 추정시 가장 기본이 되는 이론 중 하나로서 저차원 분포로부터의 표본 생성으로 구성되기 때문에 복잡한 제한조건도 쉽게 처리할 수 있다는 장점이 있다(Geman and Geman, 1984). 깃스표본법에 대한 자세한 내용은 기존 연구문헌들을 참조할 수 있다(Kwon et al., 2008; Lee and Kwon, 2011).

Markov Chain의 Convergence를 평가하는 방법은 trace plot을 이용한 시각적 분석이나 통계학적 방법(Gelman-Rubin, Geweke, Raftery-Lewis, Heidelberg-Welch 등)이 널리 알려져 있다(Brooks and Gelman, 1988; Brooks and Roberts, 1988). 본 연구에서는 모형의 Convergence를 확증하기 위해서 3개의 Chain을 독립적으로 시행하여 Sampling이 효과적으로 혼합(mixing)되도록 하였으며, trace plot을 이용하여 Chain의 convergence 여부를 판단하였다. 최종적으로 Gelman-Rubin 통계량 검정결과를 이용하여 각 유역별 Markov Chain의 Convergence 여부를 통계적으로 판단하였다.

### 3. 연구결과

본 연구에서는 40년 이상의 강우관측 자료를 보유한 기상청 강우관측소를 선별하였으며, 한반도를 4개의 유역으로 구분하여 각각의 유역에서의 강수패턴을 분석하였다. 본 연구에서 수행한 방법은 다음과 같이 요약할 수 있다.

- 1) 본 연구에서는 변동점이 명확하게 드러나는 가상의 시나리오를 구축하여 계층적 Bayesian 변동점 분석 기법이 효과적으로 변동점을 도출하는지 확인하였다.
- 2) 계층적 Bayesian 변동점 분석 기법으로부터 각각의 유역별로 연강수량의 변동점을 도출하였으며, 수문기상자료의 변동점 분석을 하는데 있어 매개변수의 불확실성을 정량화 하였다.
- 3) 변동점 전후 기상학적 변동성이 존재하는지를 평가하기 위해서 미국 NOAA에서 제공되는 재해석자료를 이용하여 한반도 주변의 강수량과 해수면기압의 Anomaly를 분석하였다.

### 3.1 대상유역

우리나라의 국가하천은 한강, 낙동강, 금강, 영산강, 섬진강까지 5개의 권역으로 분류되지만, 섬진강의 경우 강우관측소가 상대적으로 적고 유역특성도 비슷하다고 판단하여 영산강유역과 합쳐 총 4개의 권역으로 분리하여 변동점 분석을 실시하였다. 본 연구에서는 기상청에서 관측되는 강수량 자료 중 40년 이상(1973~2013년) 관측된 60개 강우관측소의 자료를 중심으로 권역별 강수패턴 및 변동점 분석을 실시하였다. Table 1은 본 연구에서 사용된 강수지점의 관측 개시연도와 위경도 좌표를 병기하여 나타내었으며, Fig. 1은 분할된 유역 및 각 유역별 강우관측소의 위치를 도시하였다. 서울, 부산 등 주요 도시를 제외한 대부분의 강수지점은 1973년부터 강수자료가 이용가능 하였으며, 모든 강수지점에서 공통적으로 이용 가능한 1973년부터 분석을 실시하였다.

### 3.2 가상시계열을 이용한 계층적 Bayesian 변동점 분석기법 성능평가

본 연구에서는 계층적 Bayesian 변동점 분석 기법 모형의 적합성을 평가하기 위하여, MCS (Monte Carlo Simulation) 기법을 활용한 가상의 강수시나리오를 시계열 형태로 모의하여 분석에 활용하였다. MCS 기법은 관심이 있는 변량에 대해서 확률분포를 정의한 후 확률분포의 역함수로부터 확률변수를 모의하는 방법으로, 추출된 자료는 가정한 확률분포와 동일한 분포를 따르게 된다. 모의된 가상 시나리오들은 Table 1에 나타난 기상청 관측소의 강수시계열 자료로부터 추정된 대수정규분포형의 매개변수를 기준으로 구축되었다. Fig. 2의 왼쪽 그림은 40개 강수지점을 기준으로 모의된 가상시계열을 나타낸 그림으로서 모든 지점에서 1990~1995년의 범위

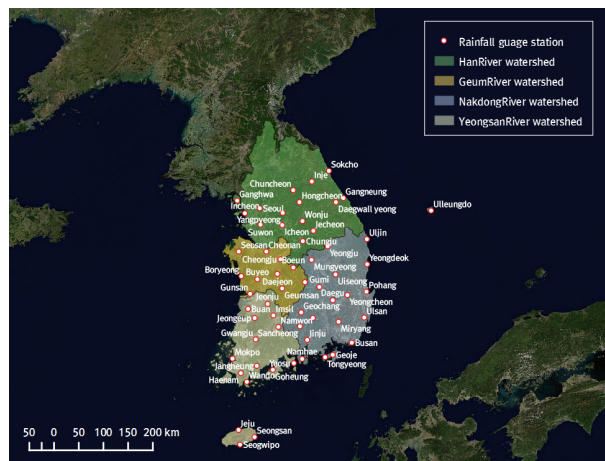


Fig. 1. A map showing 60 weather observation stations in Korea peninsula. Shaded area indicates four major watersheds

**Table 1.** Weather stations used in this study. The weather stations is operated by Korea Meteorological Administration (KMA)

Watershed	Station	Start years	Latitude (Degree)	Longitude (Degree)	Station	Start years	Latitude (Degree)	Longitude (Degree)
Han River	Chuncheon	1966	127.736	37.903	Inje	1971	128.167	38.060
	Chungju	1972	127.953	36.970	Jecheon	1972	128.194	37.159
	Daegwallyeong	1971	128.718	37.677	Seoul	1907	126.966	37.571
	Ganghwa	1972	126.446	37.707	Sokcho	1968	128.565	38.251
	Gangneung	1911	128.891	37.751	Suwon	1964	126.985	37.272
	Hongcheon	1971	127.880	37.684	Wonju	1971	127.947	37.338
	Icheon	1972	127.484	37.264	Yangpyeong	1972	127.494	37.489
	Incheon	1904	126.625	37.478	-	-	-	-
Nakdong River	Busan	1904	129.032	35.105	Pohang	1944	129.380	36.033
	Daegu	1907	128.619	35.885	Sancheong	1972	127.879	35.413
	Geochang	1972	127.909	35.667	Tongyeong	1968	128.436	34.845
	Geoje	1972	128.605	34.888	Uiseong	1973	128.689	36.356
	Gumi	1973	128.321	36.131	Uljin	1971	129.413	36.992
	Hapcheon	1973	128.170	35.565	Ulleungdo	1938	130.899	37.481
	Jinju	1969	128.040	35.164	Ulsan	1932	129.320	35.560
	Miryang	1973	128.744	35.491	Yeongcheon	1972	128.951	35.977
	Mungyeong	1973	128.149	36.627	Yeongdeok	1972	129.409	36.533
	Namhae	1972	127.926	34.817	Yeongju	1972	128.517	36.872
Yeongsan River	Buan	1972	126.717	35.730	Jeonju	1918	127.155	35.822
	Goheung	1972	127.276	34.618	Mokpo	1904	126.381	34.817
	Gwangju	1939	126.892	35.173	Namwon	1972	127.397	35.402
	Haenam	1971	126.569	34.554	Seogwipo	1961	126.565	33.246
	Imsil	1970	127.286	35.612	Seongsan	2008	126.880	33.387
	Jangheung	1972	126.919	34.689	Wando	1971	126.702	34.396
	Jeju	1923	126.530	33.514	Yeosu	1942	127.741	34.739
	Jeongeup	1970	126.866	35.563	-	-	-	-
Geum River	Boeun	1972	127.734	36.488	Chupungnyeong	1937	127.995	36.220
	Boryeong	1972	126.557	36.327	Daejeon	1969	127.372	36.372
	Buyeo	1972	126.921	36.272	Geumsan	1972	126.761	36.005
	Cheonan	1972	127.121	36.780	Gunsan	1968	127.482	36.106
	Cheongju	1967	127.441	36.639	Seosan	1968	126.494	36.777

에서 변동점의 특성을 가지도록 가상시계열을 구축하였으며 이를 대상으로 모형의 성능평가를 진행하였다.

본 연구에서 변동점 분석의 목적은 독립된 강수지점의 변동점을 분석함과 동시에 유역 단위에 평균적인 변동점을 파악하는 것이다. Fig. 2의 오른쪽 그림은 40개의 지점으로 구성된 유역의 평균적인 변동점을 평가한 결과로서 가정한 변동구간 1990~1995년에서 변동점이 정확하게 위치하고 있으며, 동시에 개별적인 강수지점의 변동성 또한 효과적으로 파악되는 것을 Fig. 3에서 확인할 수 있다.

### 3.3 관측시계열을 이용한 확률분포의 매개변수 추정 및 변동점 분석결과

본 연구에서는 변동점 분석에 있어 추정된 매개변수의 불확실성을 고려하기 위해 계층적 Bayesian 변동점 분석 기법을 적용하였다. 이를 위해 각각의 유역별 강우관측소의 변동점을 포함하여 변동점을 기준으로 분리된 구간에서 추정된 강수량의 확률분포의 매개변수들의 사후분포를 추정하고 불확실성을 평가하였다.

Bayesian 기법은 기존 최소자승법 및 최우도법과는 다르



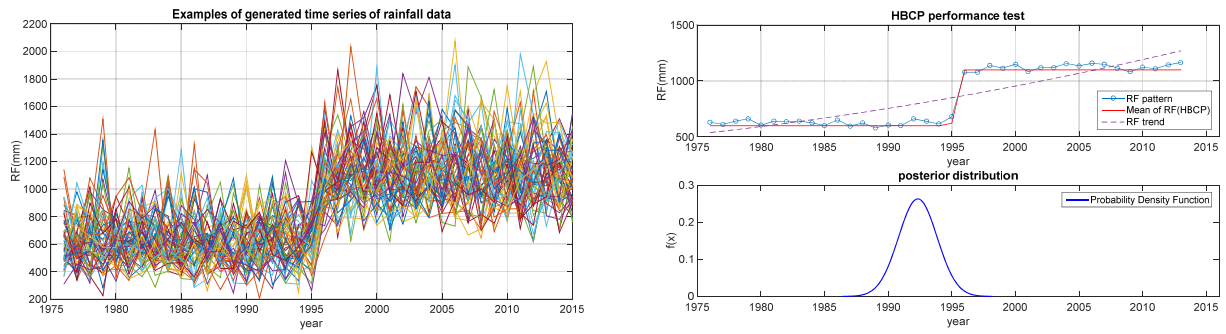


Fig. 2. Examples of generated time series of rainfall data (left) and the estimated changing point via HBCP model (right)

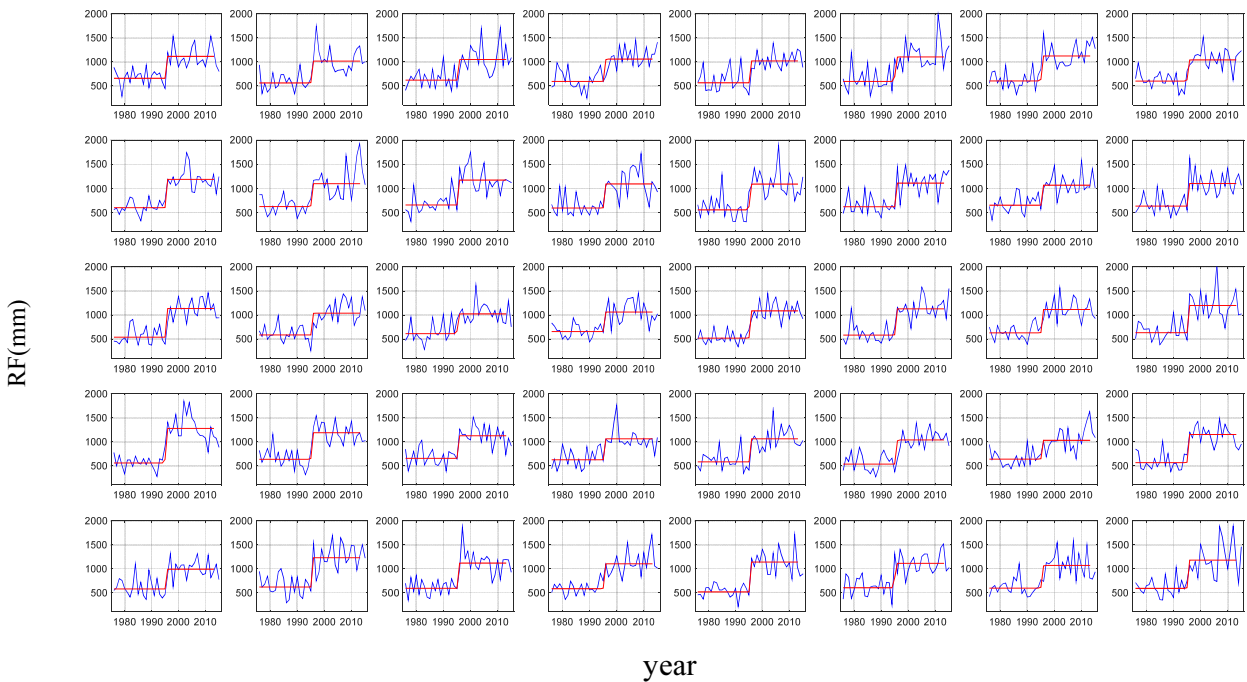


Fig. 3. The estimated changing points of individual stations using HBCP model

계 모든 매개변수에 확률분포를 부여하고 최종적으로 사후분포를 추정이 가능하기 때문에 매개변수의 불확실성을 객관적으로 정량화할 수 있다. 본 절에서는 계층적 Bayesian 변동점 분석 기법을 활용하여 실제 활용성을 검토하기 위해 기상청 연강수량 관측 자료를 대상으로 각각의 유역별 수문자료의 변동점을 분석하였다.

모든 매개변수를 직접적으로 추정하기는 현실적으로 많은 어려움이 따른다. 이에 본 연구에서는 MCMC 기법을 도입하여 각 매개변수의 사후분포를 추정하였다. 이때 매개변수의 수렴성을 확인하기 위해 Bayesian 변동점 분석 모형 내에서 3개의 Chain을 독립적으로 시행하였으며, 5,000번의 모의(iteration)를 수행하여 3,000번은 제거(burn-in) 하고, 나머지

2,000개의 Sample을 활용하여 각 매개변수 사후분포를 도출하였다.

모의된 Sampling이 이 정상적으로 혼합(Mixing)되었는지 확인하기 위하여 3개의 Markov Chain으로 부터 얻어진 2,000개의 Sample에 대한 Gelman-Rubin 검정 통계량을 확인하였다. 연쇄가 무한대로 진행될 때, 검정 통계량의 값이 1에 가까워지면 Chain에 의해 생성된 매개변수들이 동일한 분포로 수렴한다는 것을 나타낸다(Gelman et al., 2004). 모의된 사후분포에 대한 Gelman-Rubin 통계량의 값의 대부분이 1에 매우 가까움을 확인하였으며, 이는 3,000번 이후의 모의된 Sample들이 정상적으로 convergence 되었음을 나타낸다.

Table 2는 도출된 사후분포로부터 최종적으로 매개변수의



**Table 2.** Bayesian Credible interval of model parameters derived from posterior distribution

Watershed	Station	Mean	Std	2.5%	median	97.5%	Station	Mean	Std	2.5%	median	97.5%
Han River	Chuncheon	1989	3.722	1984	1989	2002	Inje	1989	3.47	1984	1989	2002
	Chungju	1989	3.424	1984	1989	2001	Jecheon	1989	3.39	1984	1989	2001
	Daegwall yeong	1989	3.646	1984	1989	2002	Seoul	1989	3.65	1985	1989	2002
	Ganghwa	1989	3.759	1985	1989	2002	Sokcho	1989	3.46	1984	1989	2001
	Gangneung	1989	3.544	1985	1989	2002	Suwon	1989	3.69	1984	1989	2002
	Hongcheon	1989	3.675	1984	1989	2002	Wonju	1989	3.72	1984	1989	2002
	Icheon	1989	3.411	1984	1989	2001	Yangpyeong	1989	3.59	1984	1989	2002
	Incheon	1989	3.395	1984	1989	2001	-	-	-	-	-	-
Nakdong River	Busan	1996	0.508	1996	1997	1998	Pohang	1996	0.46	1996	1997	1998
	Daegu	1996	0.466	1996	1997	1998	Sancheong	1996	0.49	1996	1997	1998
	Geochang	1996	0.487	1996	1997	1998	Tongyeong	1996	0.50	1996	1997	1998
	Geoje	1996	0.501	1996	1997	1998	Uiseong	1996	0.49	1996	1997	1998
	Gumi	1996	0.509	1996	1997	1998	Uljin	1996	0.50	1996	1997	1998
	Hapcheon	1996	0.502	1996	1997	1998	Ulleungdo	1996	0.51	1996	1997	1998
	Jinju	1996	0.498	1996	1997	1998	Ulsan	1996	0.50	1996	1997	1998
	Miryang	1996	0.510	1996	1997	1998	Yeongcheon	1996	0.52	1996	1997	1998
	Mungyeong	1996	0.500	1996	1997	1998	Yeongdeok	1996	0.49	1996	1997	1998
Namhae	1996	0.490	1996	1997	1998	Yeongju	1996	0.50	1996	1997	1998	
Yeongsan River	Buan	1996	0.903	1995	1997	1998	Jeonju	1996	1.05	1995	1997	1999
	Goheung	1996	1.026	1995	1997	1998	Mokpo	1996	1.03	1995	1997	1999
	Gwangju	1996	1.079	1995	1997	1998	Namwon	1996	0.96	1995	1997	1999
	Haenam	1996	1.043	1995	1997	1999	Seogwipo	1996	0.78	1995	1997	1998
	Imsil	1996	0.769	1995	1997	1998	Seongsan	1996	0.88	1995	1997	1998
	Jangheung	1996	0.777	1995	1997	1998	Wando	1996	0.99	1995	1997	1998
	Jeju	1996	0.909	1995	1997	1998	Yeosu	1996	0.99	1995	1997	1998
	Jeongeup	1996	0.966	1995	1997	1998	-	-	-	-	-	-
Geum River	Boeun	1979	5.916	1977	1978	1998	Chupung nyeong	1979	5.84	1977	1978	1997
	Boryeong	1979	5.890	1977	1978	1998	Daejeon	1979	5.88	1977	1978	1997
	Buyeo	1979	5.906	1977	1978	1998	Geumsan	1979	5.83	1977	1978	1998
	Cheonan	1979	5.868	1977	1978	1998	Gunsan	1979	5.76	1977	1978	1997
	Cheongju	1979	5.913	1977	1978	1998	Seosan	1979	5.90	1977	1978	1998

불확실성 범위를 나타내며, 매개변수별 2.5%, 50%, 97.5%의 Quantile을 추출하여 매개변수의 불확실성 구간을 정량적으로 산정하였다. 연속적으로 도출된 매개변수별 사후분포 값을 활용하여 계층적 변동점 분석 모형에서 발생하는 자료의 변동점을 정량적으로 산정하였으며, 도출된 결과는 Fig. 4와 같다.

Fig. 4에서는 시간에 따라 변동점에 대한 사후확률을 나타내었으며, 변동점의 사후확률이 가장 크게 발생하는 연도를

전후로 수문자료의 변화가 발생한 변동점으로 해석하였다. 본 연구에서 개발된 계층적 Bayesian 변동점 분석 모형을 통하여 각 유역별 1973년부터 2013년 사이에 관측된 기상청 연강수량에 대한 변동점을 Bayesian 모형 내에서 추정할 수 있었으며, 추정되는 변화시점의 구간을 정량적으로 제시할 수 있는 장점을 확인 할 수 있었다. 그림에서와 같이 유역별로 추정되는 수문자료의 변동점이 발생한 시점을 기준으로, 수문자료의 형태가 바뀌는 것을 확인할 수 있다. Fig. 4에서 파란

실선은 유역별 연강수량을 나타내며, 보라색 점선은 연강수량의 증감에 대한 경향성을 시각적으로 보여주는 회귀선을 나타낸다. 빨간선은 변동점 전후에 연강수량의 평균을 의미한다. 아래 도시된 확률밀도함수는 수문자료의 변동시점에 대한 사후분포를 나타낸 결과이다. 각각의 유역에서 도출된 사후분포를 통해 각 시점에서의 변화시점을 정량적으로 파악할 수 있다.

수문자료의 변화가 발생한 변동점에 대한 사후분포를 살펴보면, 유역별로 상이한 결과를 나타내고 있으나, 동일한 유역의 경우 대부분의 강수지점에서 유사한 변동점을 나타내는 것을 확인할 수 있다. 한강유역의 경우 1989년을 기준으로 강수량의 변동점이 나타나고 있으며, 낙동강 및 영산강유역의 경우 1994-1996년을 경계로 강수의 특성이 변동하는 것으로 평가되었다. 금강유역의 경우 타 유역과는 다르게 가장 큰 변

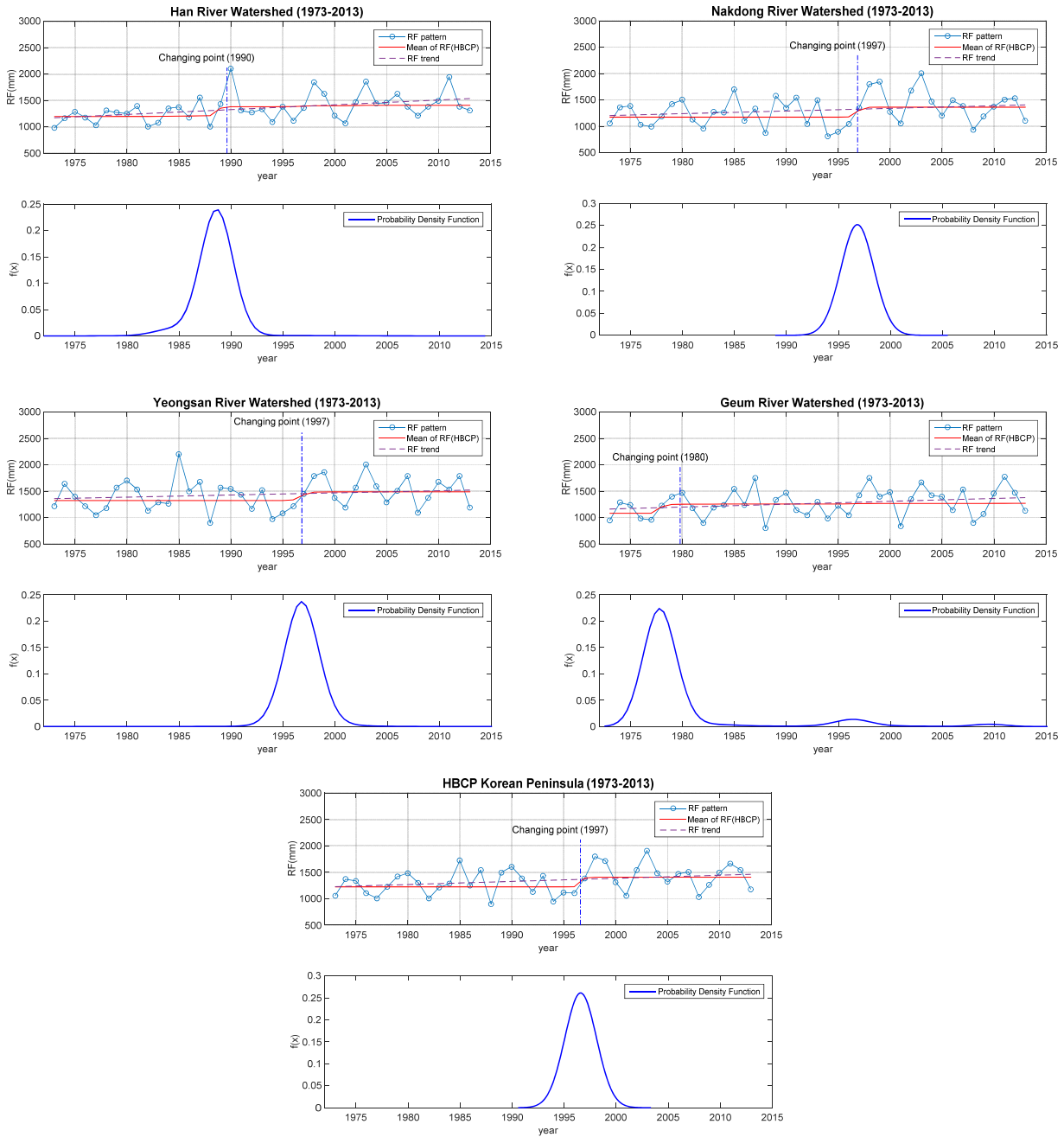


Fig. 4. The estimated changing points and their posterior distributions for four major watersheds using HBCP model

동점은 1979년에 나타나고 있다. 이와 더불어 변동점의 확률은 첫 변동점보다는 작지만 타 유역과 유사하게 1996년 정도에서 2차적인 변동점이 나타나고 있다. 전반적으로, 금강유역을 제외한 나머지 유역에서 수문자료의 변화시점은 대체적으로 1993년 전후로 나타는 것을 확인하였으며, 이 변화시점을 기준으로 연강수량의 평균이 약 200 mm 정도 상승하는 것을 확인하였다. 최종적으로 4대강유역의 모든 강수자료를 이용하여 한반도의 수문자료의 변동점을 평가한 결과 평균적으로 1997년 전후로 수문패턴의 변화가 발생하였음을 확인할 수 있었다.

### 3.4 변동점에 따른 기상학적 분석

최종적으로 본 연구에서는 우리나라에서 강수량의 변동시

점을 기준으로 해수면기압(sea level pressure, SLP)과 강수량의 공간적 분포를 Anomaly를 기준으로 평가하였다. 해수면기압과 강수자료는 미국 NOAA의 재해석(reanalysis)자료를 추출하여 활용하였으며 Anomaly는 1980~2000년의 평균을 기준으로 추정하였다.

Fig. 5에서 앞서 추정된 결과와 동일하게 양(positive)의 강수량 Anomaly를 1997년부터 2013년에서 명확하게 확인할 수 있었으며, 변동점 이전의 경우에는 반대로 음(negative) 강수량 Anomaly를 나타내고 있다. 이러한 특성은 해수면기압에서도 동일하게 나타나고 있다. 즉, 변동점 이전에는 양의 해수면기압 Anomaly를 나타내는 반면 변동점 이후에는 음의 해수면기압 Anomaly를 나타내고 있어 저기압의 분포 및 강도가 전반적으로 변동점이전보다 커졌음을 확인할 수 있다.

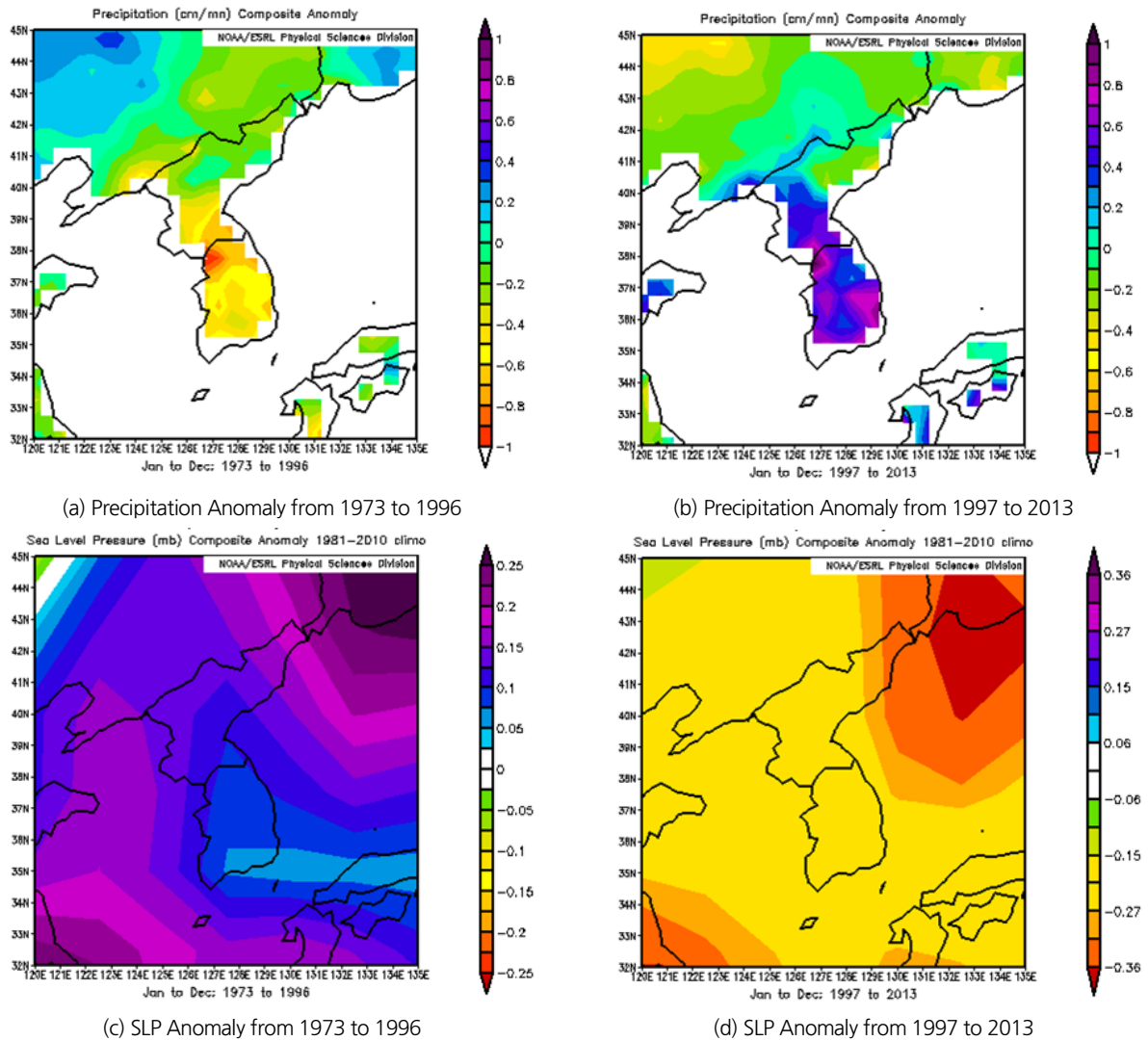


Fig. 5. The precipitation and sea level pressure anomalies for different periods determined by changing point analysis

## 4. 결론

최근 기상변동성 증가 및 기후변화로 인해 수자원 계획을 수립하는데 있어 불확실성을 가중시키고 있다. 어떠한 자연 현상도 완전히 동일하게 재현되지 않기 때문에, 시간적 순서를 고려하지 못하는 통계적 방법은 구체적인 변화시점을 명확히 파악하는데 한계가 있다. 이러한 점에서 수문 시계열의 변화양상을 효과적으로 인지하고, 유역단위에서 일관된 변화를 평가할 수 있는 변동점 분석 개발이 필요하다 하겠다. 따라서 본 연구에서는 개별 강수지점에 변동점 분석뿐만 아니라 유역단위에서 변동점 해석이 가능한 계층적 Bayesian 변동점 분석 기법을 개발하였다.

본 연구에서는 계층적 Bayesian 변동점 분석 기법을 이용하여 연강수량 자료를 대상으로 수문패턴의 변화시점을 분석하였다. 이를 위하여 비교적 장기간(1973~2013)의 관측 자료를 보유하고 있는 60개 기상관측지점에 대한 강수량 자료의 수문패턴에 대한 변동점을 추정하였으며, 유역별 연강수량의 변화시점을 확률적으로 분석하고 변동점 전후의 강수량 특성을 평가하였다. 분석결과, 상대적으로 강수패턴의 변화가 적은 금강유역을 제외한 나머지 유역에서 1990년대 초반을 기준으로 강수패턴의 변화 양상을 확인할 수 있었다. 기존 변동성 분석 기법에서도 신뢰구간을 제시해주고 있지만, 자료의 분포 특성 및 매개변수의 불확실성 등을 효과적으로 반영하지 못한다고 할 수 있다. 이러한 점에서 본 연구에서는, 수문기상 자료에 대한 변동점 분석을 하는데 있어 불확실성의 정량화와 더불어 유역별 변화시점의 분석이 가능한 계층적 Bayesian 변동점 분석 기법을 제안하였으며, 본 연구를 통해 도출된 연구 결과를 요약하면 다음과 같다.

- 1) 본 연구에서는 가상시계열 자료를 활용하여 다지점에 Bayesian 기반의 계층적 변동점 분석 모형이 적용됨에 따라 개별 지점의 변동점 뿐만 아니라 자료 전체의 전반적인 변동성을 이해하는데 유리한 특성을 확인할 수 있었다.
- 2) 변동점 분석 모형의 적용됨에 있어 발생할 수 있는 수문자료의 분포특성 및 매개변수의 불확실성 등을 정량적으로 고려할 수 있었으며, 이를 통해 추정된 변동점에 대한 신뢰성을 평가하는데 유리하였다.
- 3) 계층적 Bayesian 변동점 분석기법을 통해 각각의 특성을 나타내는 개별지점에 대한 변동점 및 유역내 다지점에서 관측된 수문시계열 자료의 통합적인 정량적인 변동점의 위치(location), 크기(magnitude)를 산정하였다.
- 4) 계층적 Bayesian 변동점 분석기법을 통해 도출된 변동점

전후 기상학적 변동성이 존재하는지를 평가하기 위해서 재해석자료를 이용하여 한반도 주변의 강수량과 해수면 기압의 Anomaly를 분석하였다. 그 결과, 분석된 한반도 수문자료에 대한 변동점을 기준으로 강수량과 해수면 기압의 명확한 차이를 확인할 수 있었다. 즉, 변동점이 기상학적 관점에서도 의미를 갖는다고 평가할 수 있다.

기후에서 변화시점을 파악하는 것은 기후변화에 대한 포괄적인 이해뿐만 아니라 극한 기상현상을 예측하고 대응하는데 있어 매우 중요하다 하겠다. 그러나 강수량과 같은 수문패턴의 변화가 구분된 유역에서 모두 같은 경향을 보이는 것은 아니며, 지역에 따라서 조금씩 다른 특성을 보이고 있다. 따라서 지형적 특성 및 관측 수문자료에 대한 지역적 차이를 고려하여 유역별 변동성의 변화를 정량적으로 분석한 본 연구 결과는 유역의 특성에 맞는 방재대책 수립 및 정책 제언 등에 대한 기초자료로 활용 가능할 것으로 사료된다.

본 연구에서는 기존 수문시계열의 변동점 분석과의 비교 검토는 이루어지지 않았다. 이러한 점에서 향후 연구에서는 기존의 변동점 분석 모형들의 장단점을 보다 정량적으로 검토하여, 보다 다양한 관점에서 변동점을 평가할 수 있는 모형으로 확장하고자 한다. 특히, 수문시계열의 변동시점과 기상학적인 변동성과의 연관성을 보다 신뢰성 있게 평가할 수 있는 기법 개발이 필요할 것으로 판단된다.

## 감사의 글

본 연구는 정부(국민안전처)의 재원으로 자연재해저감기술개발사업단의 지원을 받아 수행된 연구임 [MPSS-자연-2015-78].

## References

- Akaike, H. (1974). "A New Look at the Statistical Model Identification." *IEEE Transactions on Automatic Control*, Vol. 19, No. 6, pp. 716-723.
- Bai, J., and Perron, P. (2003). "Computation and Analysis of Multiple Structural Change Models." *Journal of Applied Econometrics*, Vol. 18, pp. 1-22.
- Barry, D., and Hartigan, J. A. (1993). "A Bayesian analysis for change point problems." *Journal of the American Statistical Association*, Vol. 88, No. 309-319, pp. 309-319.
- Brooks, S. P., and Gelman, A. (1998). "General methods for monitoring

- convergence of iterative simulations." *Journal of computational and graphical statistics*, Vol. 7, No. 4, pp. 434-455.
- Brooks, S. P., and Roberts, G. O. (1998). "Assessing convergence of Markov chain Monte Carlo algorithms." *Statistics and Computing*, Vol. 8, No. 4, pp. 319-335.
- Carlin, B. P., Gelfand, A. E., and Smith, A. F. M. (1992). "Hierarchical Bayesian analysis of change point problems." *Applied Statistics*, Vol. 41, No. 2, pp. 389-405.
- Cox, D. R., Isham, V. S., and Northrop, P. J. (2002). "Floods: some probabilistic and statistical approaches." *Philosophical Transactions of the Royal Society of London, Series A*, Vol. 360, pp. 1389-1408.
- Elsner, J. B., Niu, X., and Jagger, T. H. (2004). "Detecting shifts in hurricane rates using a Markov chain Monte Carlo approach." *Journal of Climate*, Vol. 17, pp. 2652-2666.
- Findley, D. F. (1991). "Counter examples to Parsimony and BIC." *Annals of the Institute of Statistical Mathematics*, Vol. 43, No. 3, pp. 505-514.
- Gelfand, A. E., and Smith, A. F. (1990). "Sampling-based approaches to calculating marginal densities." *Journal of the American statistical association*, Vol. 85, No. 410, pp. 398-409.
- Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2003). "Bayesian data analysis." CRC press, United States of America.
- Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2004). "Bayesian data analysis (2nd ed.)." Boca Raton: Chapman and Hall/CRC.
- Geman, S., and Geman, D. (1984). "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 6, pp. 721-741.
- Hastings, W. K. (1970). "Monte Carlo Sampling Methods using Markov Chains and their applications." *Biometrika*, Vol. 57, No. 1, pp. 97-109.
- Hollander, M., and Wolfe, D. A. (1973). "Nonparametric Statistical Methods." John Wiley & Sons, Newyork, USA.
- Hwang, S. H., Kim, J. H., Yoo, C. S., and Jung, S. W. (2010). "A probabilistic estimation of changing points of Seoul rainfall using BH Bayesian analysis." *Journal of Korea Water Resources Association*, Vol. 43, No. 7, pp. 645-655.
- Jung, H.-S., Choi, Y. E., Oh, J.-H., and Lim, G.-H. (2002). "Recent trends in temperature and precipitation over South Korea." *International Journal of Climatology*, Vol. 22, pp. 1327-1337.
- Kite, G. W. (1977). "Frequency and Risk Analysis in Hydrology." Water Resources Publication, Fort Collins, Colorado, USA.
- Lee, K. M., Jang, H. S., Kim, Y.-H., and Lee, S. H. (2011). "Change-point Analysis of Mean Temperature and Extreme Temperature in the Republic of Korea." *Journal of Korean Geographical Society*, Vol. 46, No. 5, pp. 583-596.
- Kim, K. S., and Son, Y. S. (2004). "Bayesian Change Point Analysis for a Sequence of Normal Observations: Application to the Winter Average Temperature in Seoul." *The Korean journal of applied statistics*, Vol. 17, No. 2, pp. 281-301.
- Kwon, H.-H., Casey, B., and Lall, U. (2008). "Climate Informed Flood Frequency Analysis and Prediction in Montana Using Hierarchical Bayesian Modeling." *Geophysical Research Letters*, Vol. 35, L05404.
- Kwon, H.-H., Kim, J.-Y., Kim, O.-K., and Lee, J.-J. (2013). "A Development of Regional Frequency Model Based on Hierarchical Bayesian Model." *Journal of Korea Water Resources Association*, Vol. 46, No. 1, pp. 13-24.
- Lee, J.-J., and Kwon, H.-H. (2011). "Analysis on Spatio-Temporal Pattern and Regionalization of Extreme Rainfall Data." *Journal of Korean Society of Civil Engineers*, Vol. 31, No. 1B, pp. 13-20.
- Lee, J.-J., Kwon, H.-H., and Kim, T.-W. (2010). "Concept of Trend Analysis of Hydrologic Extreme variables and Nonstationary Frequency Analysis." *Journal of Korean Society of Civil Engineers*, Vol. 30, No. 4B, pp. 389-397.
- Lee, K. M., Kwon, W. T., and Lee, S. H. (2009). "A study on plant phenological trends in South Korea." *Journal of the Korean Association of Regional Geographers*, Vol. 15, No. 3, pp. 337-350.
- Lee, S. H., Kim, S. U., Lee, Y. S., and Sung, J. H. (2014). "Intercomparison of Change Point Analysis Methods for Identification of Inhomogeneity in Rainfall Series and Applications." *Journal of Korea Water Resources Association*, Vol. 47, No. 8, pp. 671-684.
- Lund, R., and Reeves, J. (2002). "Detection of undocumented change points: a revision of the two-phase regression model." *Journal of Climate*, Vol. 15, pp. 2547-2554.
- Mann, H. B., and Whitney, D. R. (1947). "On a test of whether one of two ranmom variables is stochastically larger then the other." *Ann Math Statist.*, Vol. 18, pp. 50-60.
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). "Equations of state calculations by fast computing machines." *Journal of Chemical Physics*, Vol. 21, pp. 1087-1091.
- NOAA Earth System Research Laboratory (<http://www.esrl.noaa.gov/psd/cgi-bin/data/composites/printpage.pl>)
- Olshen, A. B., Venkatraman, E. S., Lucito, R., and Wigler, M. (2004). "Circular binary segmentation for the analysis of array-based dna copy number data." *Biostatistics*, Vol. 5, No. 4, pp. 557-572.
- Sen, A., and Srivastava, M. S. (1975). "On tests for detecting change in mean." *The Annals of Statistics*, Vol. 3, No. 1, pp. 98-108.
- Solow, A. R. (1987). "Testing for climate change: An application of two-phase regression model." *Journal of Applied Meteorology*, Vol. 26, pp. 1401-1405.
- Yun, S.-H., and Lee, J.-T. (2001). "Climate change impacts on optimum ripening periods of rice plant and its countermeasure in rice cultivation." *Korean Journal of Agricultural and Forest Meteorology*, Vol. 3, No. 1, pp. 55-70.