

Grid Method 기법을 이용한 베이지안 비정상성 확률강수량 산정

Bayesian Nonstationary Probability Rainfall Estimation using the Grid Method

곽도현* / 김광섭**

Kwak, Dohyun / Kim, Gwangseob

Abstract

A Bayesian nonstationary probability rainfall estimation model using the Grid method is developed. A hierarchical Bayesian framework is consisted with prior and hyper-prior distributions associated with parameters of the Gumbel distribution which is selected for rainfall extreme data. In this study, the Grid method is adopted instead of the Metropolis Hastings algorithm for random number generation since it has advantage that it can provide a thorough sampling of parameter space. This method is good for situations where the best-fit parameter values are not easily inferred a priori, and where there is a high probability of false minima. The developed model was applied to estimated target year probability rainfall using hourly rainfall data of Seoul station from 1973 to 2012. Results demonstrated that the target year estimate using nonstationary assumption is about 5~8% larger than the estimate using stationary assumption.

Keywords : Nonstationary, Hierarchical Bayesian model, Grid method, Probability rainfall amount

요 지

본 연구에서는 Grid method를 사용하여 베이지안 비정상성 확률강수량 산정 모형을 확립하였다. 강우 극치자료의 분포로 Gumbel 분포를 채택하였으며, 분포형의 매개변수에 사전분포를 적용하고, 사전분포에 포함된 매개변수에는 초사전 분포를 적용하여 계층적 베이지안 모형을 구성하였다. Grid method는 매개변수의 발생가능 전 구간에 대하여 확률적으로 더 높은 뒷받침이 있는 하위 구간에서 난수를 직접 생성하여 집합을 구성함으로써 잘못된 결과를 도출할 수 가능성이 높은 상황에서도 보다 정확한 매개변수의 추정을 가능케 하므로 매개변수의 추정과정에서 비표준분포로 나타나는 조건부 확률밀도함수를 통한 난수의 추출은 기존에 사용해 온 Metropolis Hastings 알고리즘이 아닌 Grid method를 사용하였다. 개발된 모형은 서울의 1973년부터 2012년까지의 시강우자료를 이용하여 미래에 대한 재현기간에 따른 확률강수량을 산정하였으며, 그 결과로 기존 정상성 가정에 비해 목표연도에 따라 5%에서 8%정도의 증가율을 나타냈다.

핵심용어 : 비정상성, 계층적 베이지안 모형, Grid method, 확률 강수량

1. 서 론

종래의 강우패턴인 정상성 가정은 근래의 강우경향을 만족시키지 못하고 있다. 기후 인자들은 시간이 지남에

따라 강도와 빈도가 증가하는 비정상성 가정에 초점이 집중되고 있으며, 변화하고 있는 기후패턴을 이해하고, 대비하기 위하여 활발한 연구가 진행되고 있다. Kim et al. (2012)은 L-모멘트의 경향성 분석을 이용하여 강우의 비정상성

* 경북대학교 공과대학 건축·토목공학부 석사 (e-mail: rhkeghus7@naver.com)
Dept. of Civil Engrg., Kyungpook Ntnl. Univ., Daegu 702-701, Korea.

** 교신저자, 경북대학교 대학원 건축·토목공학부 교수 (e-mail: kimgu@knu.ac.kr, Tel: 82-53-950-5614)
Corresponding Author, Dept. of Civil Engrg., Kyungpook Ntnl. Univ., Daegu 702-701, Korea.

특성을 해석하였으며, Kim et al. (2011)은 최우도법을 사용하여 GEV 분포형의 매개변수를 추정하고 비정상성 빈도분석을 실시하였다. Jang et al. (2011)은 연 최대강우량의 회귀직선에 대한 잔차의 수문학적 빈도해석을 바탕으로 가까운 미래에 대한 확률강우량을 산정하는 방법을 제안하였으며, Lee et al. (2010)은 비정상성 강우빈도해석법을 적용하여 추정한 확률강우량에 대한 적용성 및 신뢰성을 평가하였다. 그리고 기후변화가 한반도의 미래 극한강수의 발생에 미치는 영향을 전망하기 위해 Sung et al. (2012)은 IPCC 5차 보고서(AR5)에 대비하여 새롭게 권장되는 대표농도경로(RCP)에 따라 모의된 지역기후모형의 여름철 강수량을 외부인자로 고려한 비정상성 빈도해석을 실시하였다. 이런 경향은 비단 국내뿐만 아니라 국외에서도 이미 사례를 주시하여, 관련 연구 성과를 쏟아내고 있다. Cannon (2011)은 R project의 패키지인 GEVcdn을 GEV빈도추정네트워크에 의한 비정상성 극한값을 분석함으로써, 방법론적인 초석을 닦는 연구를 진행하였다. Coulibaly and Baldwin (2005)은 Nonlinear dynamic methods를 이용한 비정상성 수문학적 시계열 자료의 예측을 연구하였으며, Villarini et al. (2010)은 장기간의 강우와 온도데이터를 기반으로 GAMLSS 모델을 이용한 비정상성 framework를 구성하였다. 그리고 Kottegoda et al. (2011)은 강우자료를 바탕으로 Gibbs procedure에 의한 비정상성 경향과 주기성에 관한 연구를 실시하였다.

또한 통계적인 부분에서는 고전적인 통계기법과 더불어 베이지안 접근을 통해 비정상성을 분석하기 위한 노력이 진행되어 왔다. Thyer and Kuczera (2003)은 장기지속 강우 자료에 대한 Hidden Markov Model (HMM)을 보정하기 위하여 베이지안 접근을 이용하였으며, Biondi and Deluca (2012)은 Calabria 지역의 실시간 홍수예측의 불확실성을 산정하기 위해 Bayesian Forecasting System (BFS)를 적용하였다. 그리고 Lee et al. (2010)은 베이지안 모형을 구성하여, Gibbs sampler에 의한 Markov Chain Monte Carlo Simulation을 통해 매개변수를 추정을 실시하였으며, POT (Peak Over Threshold) 자료로부터 계절성을 추출하고 이를 빈도해석과 연계시켜 베이지안 기법을 기반으로 하는 비정상성 빈도해석 모형의 구축에 대한 연구를 수행하였다.

베이지안 접근은 빈도주의적 접근에 비해 분석결과의 의미 해석이 용이하며, 표본 자료의 수가 적어도 사용 가능하고 Learning 과정의 표현이 가능하다는 점 등 여러 가지 장점을 가지고 최근 여러 분야에서 활용되고 있다. 베이지안 추론에서는 최종 사후분포의 조건부 밀도함수

를 통해 각 매개변수의 분포를 따르는 많은 난수를 추출하여 이들의 평균을 사용함으로써 이 통계량이 매개변수의 분포를 대표하는 값이 된다. 베이지안 기법을 통한 추론의 근본에는 난수추출이라는 불확실성을 내포한 단계가 있으며 이를 다수의 시행을 통한 통계값으로 한정하여 불확실성을 해소하고 있다. 그러나 비교적 고차원이 아닌 일반적인 베이지안 모형의 경우는 동일한 조건, 동일한 자료 하에서 안정된 추정의 결과를 나타내지만 본 연구의 비정상성 강우에 적용한 계층적 베이지안 모형의 경우는 구조상 다수의 매개변수가 발생하여 최종 사후분포가 8차원의 형태로 복잡하게 결합된다. Gibbs sampler이론에 의해 각 조건부 확률밀도함수를 차례로 순환하며 난수를 추출할 때 1cycle은 연관된 매개변수가 많아질수록 길어지며, 구조상 모든 매개변수의 난수들을 거쳐 다음 순번의 난수가 생성되기 때문에 과정 속의 불확실성은 더욱 가중된다. 이런 과정을 통해 결정된 매개변수는 매 추정마다 약간의 오차를 나타낼 수 있는 가능성을 내포하고 있다. 따라서 매 추정의 서로 다른 결과의 옳고 그름을 결정할 수 없을 뿐더러, 이를 확률 강수량 산정에 적용하기 위해서는 유일성을 확보 할 수 있는 다른 방안이 필요하다. 본 연구에서는 이와 같은 문제점을 해결하고자 표준 분포형으로 나타나지 않는 조건부 확률밀도함수의 난수 추출시 Metropolis Hastings 알고리즘 대신 Grid method를 선택하여 결과의 분산을 보정한 후 차후의 2차 보정을 통해 유일성 확보를 통한 비정상성 확률강수량 산정에 목표를 두었다.

2. 계층적 베이지안 모형

증가하는 강우의 경향성을 고려한 미래의 확률강수량 예측을 위해 계층적 베이지안 모형을 고안하였다. 연구에 사용된 자료는 서울, 강릉, 대구, 부산 4개의 각 지점의 AWS (Automatic Weather System)를 통해 획득한 1973년부터 2012년까지의 시 강우자료를 이용하였으며, 최종 모형의 결과를 각 지점마다 표와 그림으로 수록하였다. 그리고 결과 도출전의 이론 전개과정에서 이해를 돕기 위해 수록된 표는 서울지점을 대상으로 한 예시를 수록하였다. 자료는 BM (Block Maxima)방법을 기반으로 Block을 1년으로 지정하여 24시간 강우강도에 대한 총 40개의 극치자료로 구성된 집합을 형성하였다. 확률강수량의 추정을 위해서는 해당 자료들이 따르는 분포의 선정은 필수적이며, 본 연구에서 사용될 극치 강우량의 집합은 연 최대 강수량 자료의 빈도해석에 많이 이용되고 있는 Gumbel

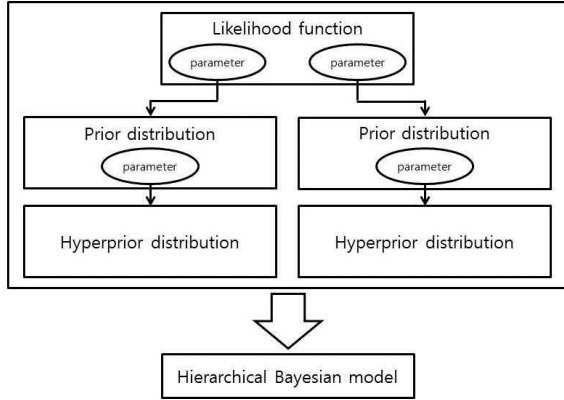


Fig. 1. Schematic Diagram of Hierarchical Bayesian Model

분포를 따른다고 가정하였다. 아래 Eq. (1)은 Gumbel 분포의 확률밀도함수를 나타낸다.

$$f(x) = \frac{1}{\beta} e^{-(z+e^{-z})}, \text{ where } z = \frac{y-\delta}{\beta} \quad (1)$$

베이저안 모형은 자료의 정보를 담고 있는 우도함수 (Likelihood function)와 사전에 미리 획득하여 신빙성 있는 정보를 보장하는 사전분포(Prior distribution)의 결합으로 형성된다. 추가적으로 사전분포가 가지는 정보인 초사전분포(Hyperprior distribution)까지 결합하여 이 모든 정보를 아울러 형성된 모형의 형태를 정보의 층(layer)이 구분된 계층적인 베이저안 모형이라 할 수 있다.

2.1 우도함수

우도함수는 베이저안 모형에서 실질적으로 획득한 자료로 형성되는 부분이며 전체적인 베이저안 모형의 틀을 잡아주는 기초부분이 된다. 비정상적인 경향을 나타내기 위하여 Gumbel 분포의 위치 모수와 축척 모수에 시간의 정보를 포함한 선형함수를 결합하여 각각 $\delta_0 + \delta_1 t_i, \beta_0 + \beta_1 t_i$ 로 변환하였다. 여기서 δ_0, β_0 는 절편모수이며 δ_1, β_1 는 기울기 모수를 나타내고 t_i 항은 시간을 나타내며 시작시점인 1973년의 t_i 는 1이 되며, 마지막인 2012년의 t_i 는 40이 된다. 이로써 우도함수는 1973년부터 2012년까지 매년의 정보를 각각 담고 있는 총 40개의 각 확률분포함수들이 하나의 묶음으로 결합된 형태로 표현할 수 있다. 아래 Eq. (2)는 결합된 우도함수를 나타낸다.

$$\prod_{i=1}^{40} \frac{1}{\beta_0 + \beta_1 t_i} \exp \left[\frac{(y_i - (\delta_0 + \delta_1 t_i))}{\beta_0 + \beta_1 t_i} + \exp \left(- \left(\frac{y_i - (\delta_0 + \delta_1 t_i)}{\beta_0 + \beta_1 t_i} \right) \right) \right] \quad (2)$$

2.2 사전분포 및 초사전분포

Eq. (2)의 형성된 우도함수에 대해 사전정보를 결합한다. 사전 분포를 적용시킬 모수는 $\delta_0, \delta_1, \beta_0, \beta_1$ 로 총 4개이며, 절편 모수를 나타내는 δ_0, β_0 는 Normal 분포를 적용하였다. 그리고 기울기 모수를 나타내는 δ_1, β_1 는 Uniform 분포를 적용하였으며, 이 사전 분포로 나타난 사전 정보들이 Eq. (2)에 결합된다. 아래의 Eqs. (3a)~(3d)는 각 사전 분포들의 형태를 나타낸다. 강우 모수의 경우 강우현상이 발생한 사건만이 주어진 정보이다. 즉, 자료에 근거하여 사후분포를 추정하기 위해서는 사전분포에 정보를 없애야 하고, 무정보적 사전분포를 적용하였다. 무정보적 사전분포의 가장 일반적인 형태가 균일(Uniform)분포이다.

$$\delta_0 | \mu_\delta, \sigma_\delta^2 \sim \mathcal{N}(\mu_\delta, \sigma_\delta^2) \quad (3a)$$

$$\delta_1 \sim \text{uniform} \quad (3b)$$

$$\beta_0 | \mu_\beta, \sigma_\beta^2 \sim \mathcal{N}(\mu_\beta, \sigma_\beta^2) \quad (3c)$$

$$\beta_1 \sim \text{uniform} \quad (3d)$$

초사전분포는 사전 분포의 매개변수인 $\mu_\delta, \sigma_\delta^2, \mu_\beta, \sigma_\beta^2$ 에 삽입된다. 여기서 μ_δ, μ_β 는 Uniform분포로 설명되었으며, $\sigma_\delta^2, \sigma_\beta^2$ 는 각각 $\pi(\sigma_\delta^2) \propto \frac{1}{\sigma_\delta^2}, \pi(\sigma_\beta^2) \propto \frac{1}{\sigma_\beta^2}$ 를 따르는 상수분포를 사용하였다. 특별히 분산의 역수를 사용하는 이유는 분산이 말하는 정보가 자료의 변동성을 보여주는 것인데 이 때, 자료의 모수에 대한 정보가 없을 때 변동성을 가장 크게 염두 해 두고 하는 것이다. 즉, 큰 분산모수를 적용하여 균일분포의 범위를 최대한 넓게 선정하고, 자료의 정보만을 최대한 사용하여 사후분포를 형성하는 것이 옳다. 특히 균일분포의 경우 모형의 민감도에 크게 영향을 미치지 않기 때문에 무정보적 사전분포 중 가장 일반적으로 사용된다.

2.3 사후분포

$$\text{posterior} \propto \text{prior} \times \text{likelihood} \rightarrow \quad (4)$$

$$\pi(\theta|y) \propto \pi(\theta) \times L(y|\theta) \times \text{hyper prior}$$

$$\text{where } L(y|\theta) = \prod_{i=1}^n f(y_i|\theta)$$

위의 Eq. (4)와 같이 Eq. (2)의 우도함수와 Eq. (3)의 사전분포, 그리고 초사전분포의 정보를 모두 결합한 형태인 사후분포는 아래 Eq. (5)와 같이 나타난다.

$$\begin{aligned} & \pi(\delta_0, \delta_1, \beta_0, \beta_1, \mu_\delta, \sigma_\delta^2, \mu_\beta, \sigma_\beta^2 | y_1, \dots, y_{40}) \\ & \propto \left[\prod_{i=1}^{40} \frac{1}{\beta_0 + \beta_1 t_i} \exp\left\{ \frac{(y_i - (\delta_0 + \delta_1 t_i))}{\beta_0 + \beta_1 t_i} \right. \right. \\ & \quad \left. \left. + \exp\left(-\frac{y_i - (\delta_0 + \delta_1 t_i)}{\beta_0 + \beta_1 t_i}\right) \right\} \right] \quad (5) \\ & \times \frac{1}{\sqrt{2\pi\sigma_\delta^2}} \exp\left\{-\frac{1}{2\sigma_\delta^2}(\delta_0 - \mu_\delta)^2\right\} \\ & \times \frac{1}{\sqrt{2\pi\sigma_\beta^2}} \exp\left\{-\frac{1}{2\sigma_\beta^2}(\beta_0 - \mu_\beta)^2\right\} \times \frac{1}{\sigma_\delta^2 \sigma_\beta^2} \end{aligned}$$

3. 매개변수 추정기법

3.1 Gibbs sampler

최종 형성된 사후분포에서 확률강수량의 추정에 직접적으로 사용되는 매개변수는 $\delta_0, \delta_1, \beta_0, \beta_1$ 이다. 그러나 이 모수들만 추출하여 추정하기란 사실상 불가능하며 전체적인 모수의 추정을 통하여 획득할 수 있다. 여기서는 마코브 연쇄 몬테칼로(Markov Chain Monte Carlo; MCMC) 방법 중 하나인 Gibbs sampler를 추정기법으로 선정하였다. Gibbs sampler란 다차원의 결합 확률분포가 복잡하여 직접 랜덤표본을 생성하기 어려운 경우 각 변수의 조건부 확률분포로부터 랜덤표본을 반복적으로 생성하면 적절한 조건 하에서 이들의 극한분포가 결합 확률밀도함수가 된다는 사실에 근거하여 난수를 생성하는 방법이다(Kim, 2011). 따라서 Eq. (5) 사후분포의 형태에서 각 매개변수들의 조건부 확률분포를 추출할 필요가 있으며, 또한 이때 개개의 조건부 확률분포로부터의 난수 발생은 어려움이 없어야 한다. Eqs. (6a)~(6h)는 해당 매개변수들의 조건부 확률분포를 나타낸다.

$$\begin{aligned} & \pi(\delta_0 | others) \propto \exp \\ & \left\{ -\sum_{i=1}^{40} \left[\frac{\delta_0 - \delta_1 t_i}{\beta_0 + \beta_1 t_i} + \exp\left(-\frac{\delta_0 - \delta_1 t_i}{\beta_0 + \beta_1 t_i}\right) \right] - \frac{1}{2\sigma_\delta^2}(\delta_0 - \mu_\delta)^2 \right\} \quad (6a) \end{aligned}$$

$$\pi(\delta_1 | others) \propto \exp\left\{-\sum_{i=1}^{40} \left[\frac{\delta_0 - \delta_1 t_i}{\beta_0 + \beta_1 t_i} + \exp\left(-\frac{\delta_0 - \delta_1 t_i}{\beta_0 + \beta_1 t_i}\right) \right]\right\} \quad (6b)$$

$$\begin{aligned} & \pi(\beta_0 | others) \propto \left[\sum_{i=1}^{40} \frac{1}{\beta_0 + \beta_1 t_i} \right] \\ & \exp\left\{-\sum_{i=1}^{40} \left[\frac{\delta_0 - \delta_1 t_i}{\beta_0 + \beta_1 t_i} + \exp\left(-\frac{\delta_0 - \delta_1 t_i}{\beta_0 + \beta_1 t_i}\right) \right] - \frac{1}{2\sigma_\beta^2}(\beta_0 - \mu_\beta)^2\right\} \quad (6c) \end{aligned}$$

3.2 Grid method

위의 조건부 확률분포 중 $\mu_\delta, \mu_\beta, \sigma_\delta^2, \sigma_\beta^2$ 는 표준분포의 형태를 띠고 있어 용이한 난수 발생이 가능한 반면, 앞의 $\delta_0, \delta_1, \beta_0, \beta_1$ 의 경우 비표준분포의 형태를 나타내고 있어 수월한 난수 추출에 지장을 준다. 이런 경우에 흔히 사용하는 방법으로 Metropolis Hastings 알고리즘이 있지만 이는 후보생성밀도함수라는 가상의 분포에서 생산한 난수들의 기각률을 통해 난수들의 적절함을 판단하는 방법으로 해당 조건부 확률분포에서 직접적인 추출이 이루어지지 않는 단점을 가지고 있다. 따라서 대안으로 보다 근본적인 부분에서 정확한 추출을 가능하게 하는 Grid method기법을 도입하였다. Grid method는 먼저 적용할 조건부 사후분포의 구간을 하위 100개의 구간으로 나눈다. 그 다음 각 구간의 중간점을 조건부 사후분포의 해당 변수에 대입하여 확률값을 산출한다. 이에 의해 각 구간은 확률 값을 가지게 되며, 1부터 100번째의 확률값이 발생한다. 그 후 $A = \sum_{k=1}^{100} a_k$ (a_k 는 각구간의 계산된 확률값)를 먼저 계산한 다음에 $b_k = \sum_{i=1}^k a_i / A, k=1, \dots, 100$ 를 계산하여 cdf 형태를 구축하

Table 1. Results of Simulation about Main Parameters

The number of trials	δ_0	δ_1	β_0	β_1
1	145.7375	0.2062	61.2391	0.0702
2	147.1871	0.2034	61.4803	0.0659
3	147.2707	0.2007	61.5319	0.0653
4	147.0824	0.2034	61.5067	0.0675
5	147.0804	0.2054	61.5581	0.0717
6	146.9354	0.2078	61.4863	0.0656
7	147.0831	0.2051	61.5652	0.0697
8	147.1722	0.2010	61.6075	0.0687
9	147.0967	0.2052	61.4493	0.0714
10	147.1722	0.2042	61.6512	0.0674

게 된다. 따라서 최종 구간인 100번째의 b_k 값은 1이 된다. 그리고 0과 1사이의 임의의 난수를 발생시킨 후 해당 난수가 포함되는 구간을 선택하고, 다시한번 그 구간 안에서 난수를 생성함으로써 최종적으로 다수의 반복된 과정을 거치면 근본적으로 조건부사후분포를 따르는 난수 표본 집단을 선정할 수 있다. 아래의 Table 1은 주관심 모수에 대한 총 10회의 독립된 시행의 추정된 매개변수를 나타낸다.

3.3 확률강우량 산정

Table 1에 나타난 것과 같이 동일한 조건과 자료를 가

지고 추정하였음에도 불구하고, 베이지안 추정의 난수추출 기반의 특성상 약간씩의 편차들이 발생한다. 모수를 산정함에 있어서 베이지안 추정은 최종적으로 발생한 해당 모수에 대한 난수의 평균을 주로 사용한다. 본 연구에서 구축한 계층적 베이지안 모형은 조건부 사후분포가 8차원의 복잡한 형태로 이루어져 있으며 Gibbs sampler를 통해 난수 추출 시 각 모수를 다음 번 추출하기 위해서는 7번의 순환을 거쳐 돌아오게 된다. 이 7번의 순환 과정에서 포함되는 랜덤한 영향이 아무리 다수의 시행을 거친다고 하여도 최종적인 모수 추정의 결과를 보면 약간씩의

Table 2. Sample Results of 30 Time Simulation

The number of trials	δ_0	δ_1	β_0	β_1
1	145.7375	0.2062	61.2391	0.0702
2	147.1871	0.2034	61.4803	0.0659
3	147.2707	0.2007	61.5319	0.0653
4	147.0824	0.2034	61.5067	0.0675
5	147.0804	0.2054	61.5581	0.0717
6	146.9354	0.2078	61.4863	0.0656
7	147.0831	0.2051	61.5652	0.0697
8	147.1722	0.2010	61.6075	0.0687
9	147.0967	0.2052	61.4493	0.0714
10	147.1722	0.2042	61.6512	0.0674
11	147.1155	0.2083	61.5166	0.0675
12	146.9841	0.2060	61.6549	0.0664
13	147.0848	0.2035	61.4627	0.0680
14	147.0248	0.2049	61.6259	0.0688
15	147.1600	0.2069	61.5268	0.0702
16	146.9937	0.2074	61.5601	0.0696
17	147.0990	0.2049	61.5614	0.0688
18	147.0504	0.2091	61.5696	0.0665
19	147.0336	0.2000	61.5864	0.0676
20	147.1798	0.2037	61.6206	0.0681
21	147.0528	0.2047	61.5891	0.0682
22	147.0069	0.2042	61.6362	0.0657
23	147.2168	0.2049	61.4987	0.0674
24	147.1783	0.2033	61.5735	0.0677
25	147.0803	0.2034	61.5244	0.0703
26	147.1359	0.2041	61.6295	0.0655
27	147.0347	0.2087	61.5688	0.0651
28	147.0914	0.2023	61.5268	0.0699
29	147.0311	0.2093	61.5874	0.0696
30	146.9682	0.2062	61.5993	0.0667
The maximum value	147.2707	0.2093	61.6549	0.0717
Round up	147.3000	0.3000	61.7000	0.1000

편차가 발생하게 되며, 이것을 확률강수량이라는 설계제원으로 사용하기에는 치명적인 부분이 된다. 같은 자료, 같은 과정을 거치면 항상 같은 결과를 주어야 하는 유일성이 결여되어 있기 때문이다. true parameter를 정확히 알 수는 없지만 설계제원으로써의 역할을 하기 위해서는 특성상 과소추정은 막아야하며 이를 위해 그 후의 수치보정 과정을 적용한다. Gumbel 분포의 구조상 추정된 매개변수가 커질수록 확률 강수량도 크게 나타난다. 이를 이용하여 고전적인 통계학에서 대표본이라 할 수 있는 30회의 시행을 거친 후, 그 중 각 매개변수들의 최대값들을 선정하여 특정 위치에서 Round up을 통한 유일성을 확보하고, 과소 추정에 대한 불안을 해소하였다. 최대값을 사용하는 이유는 설계제원이라 함은 과소추정은 막아야 제 역할을 수행할 수 있다. 구조물의 붕괴 또는 사고와 직결되

는 부분이기 때문이다. 만약 평균을 사용하게 된다면 혹여나 발생할 수 있는 평균 이상의 추정치를 커버하지 못하게 된다. 그래서 30번의 시행중 최대값을 사용하게 되었으며, 이 30번의 시행은 30번의 난수 추출 횟수가 아닌 전체 시뮬레이션을 의미한다. 그리고 round up을 하는 이유는 30번의 시뮬레이션을 1회로 가정하여 다수 시행하였을 때 최대값을 round up하면 항상 동일한 모수 추정의 결과를 획득할 수 있기 때문이다. 이는 앞선 의미와 부합하여 생각해보면 언제든 과소 추정되는 부분을 방어해 줄 수 있다. 그러나 이를 본 논문에 수록하기에는 엄청난 양의 표로 구성되어 무리라는 판단에 생략하게 되었으며, 1회 시행만의 예시를 수록한다.

위의 과정을 통해 산출된 목표연도에 따른 재현기간 별 확률강수량을 아래에 나타냈다. Gumbel 분포의 매개변수

Table 3. Probable Precipitation Amount about Return Period of Target Year at Seoul (mm)

Return period	2020	2030	2040	2070	2100
10	313.6115	316.0617	318.5120	325.8626	333.2133
50	432.8219	436.2631	439.7043	450.0277	460.3512
100	483.2188	487.0789	490.9390	502.5192	514.0995
200	533.4317	537.7092	541.9867	554.8192	567.6516

Table 4. Probable Precipitation Amount about Return Period of Target Year at Gangneung (mm)

Return period	2020	2030	2040	2070	2100
10	340.6965	343.3718	346.0471	354.0728	362.0986
50	482.5005	486.3318	490.1632	501.6572	513.1513
100	542.4488	546.7689	551.0890	564.0493	577.0096
200	602.1784	606.9855	611.7926	626.2138	640.6350

Table 5. Probable Precipitation Amount about Return Period of Target Year at Daegu (mm)

Return period	2020	2030	2040	2070	2100
10	207.7456	209.9708	212.1960	218.8716	225.5471
50	278.1026	281.1536	284.2045	293.3574	302.5103
100	307.8464	311.2464	314.6465	324.8467	335.0470
200	337.4816	341.2295	344.9774	356.2211	367.4648

Table 6. Probable Precipitation Amount about Return Period of Target Year at Busan (mm)

Return period	2020	2030	2040	2070	2100
10	294.2581	296.7083	299.1585	306.5092	313.8598
50	397.9437	401.3849	404.8260	415.1495	425.4730
100	441.7774	445.6375	449.4975	461.0778	472.6581
200	485.4511	489.7286	494.0061	506.8385	519.6710

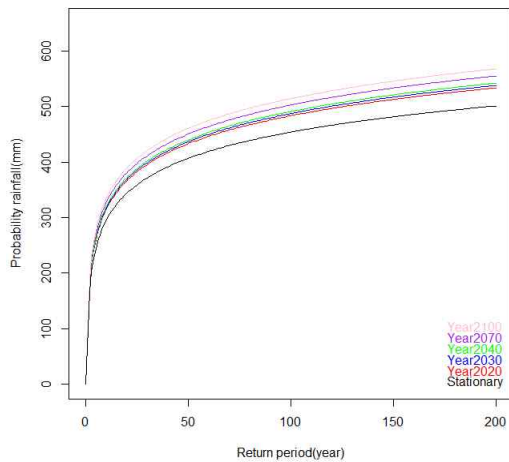


Fig. 2. Probable Precipitation Amount about Return Period of Target Year at Seoul (mm)

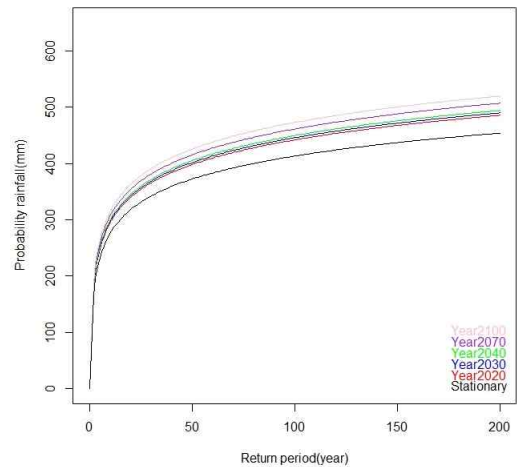


Fig. 5. Probable Precipitation Amount about Return Period of Target Year at Busan (mm)

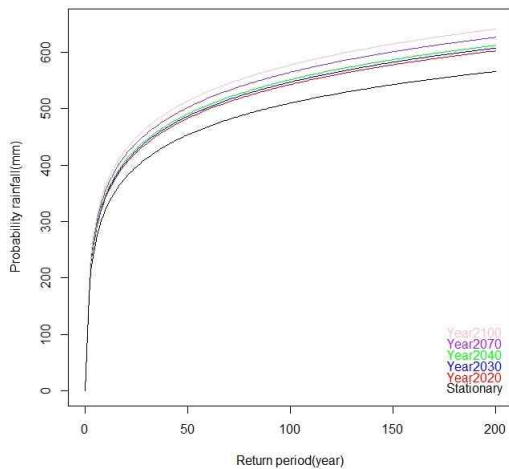


Fig. 3. Probable Precipitation Amount about Return Period of Target Year at Gangneung (mm)

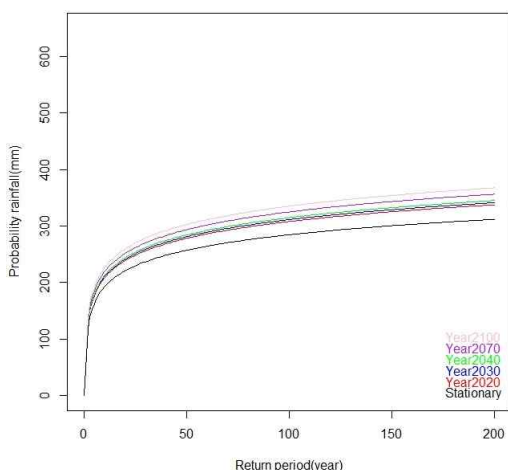


Fig. 4. Probable Precipitation Amount about Return Period of Target Year at Daegu (mm)

를 확률가중모멘트법으로 추정하여 확률 강수량을 산정한 정상성 가정에 비해 시간이 지남에 따라 점점 더 증가하는 결과를 나타냈다.

4. 결 론

본 연구는 미래에 대한 비정상성 확률강수량을 추정하기 위해 서울, 강릉, 대구, 부산 총 4개 지점의 1972년부터 2012년까지의 시강우 자료를 바탕으로 계층적 베이저안 모형을 구성하였다. 이 과정에서 기존에 주로 사용한 비표준 분포로 나타난 조건부 확률밀도함수에 대한 난수 추출 기법인 Metropolis Hastings 알고리즘을 사용하지 않고, 매개변수 전구간에 대하여 자료를 추출하므로 근본적으로 보다 정확한 매개변수 산정을 할 수 있는 Grid method를 도입하여 모형을 완성하였다. 이로써, 기존 기법이 보인 동일 시행에 대한 결과의 편차들에 대한 분산을 줄일 수 있었다. 또한 유일성 확보를 위해 2차 보정을 실시하였으며 이 일련의 과정을 통하여 정량화된 확률 강수량을 산정할 수 있었다. 모형은 서울의 경우에는 정상성 가정을 기반으로 산출된 강수량 보다 5%에서 먼 미래의 경우에는 8%가 넘는 증가율을 나타냈다. 정상성 가정을 통한 설계제원의 확보 시에는 증가하는 강수량을 감당하지 못하는 결과가 나올 수 있으며, 비정상성 강우를 충분히 고려한 결정이 필요하다.

감사의 글

이 논문은 2013년도 정부(미래창조과학부)의 재원으로

로 한국연구재단의 지원을 받아 수행된 연구임(No. 2011-0030040).

References

- Biondi, D., and De Luca, D.L. (2012). "A Bayesian approach for real-time flood forecasting." *Physics and Chemistry of the Earth*, Vol. 42-44, pp. 91-97.
- Cannon, A.J. (2011). "GEVcdn: An R package for nonstationary extreme value analysis by generalized extreme value conditional density estimation network." *Computers and Geosciences*, Vol. 37, pp. 1532-1533.
- Coulibaly, P., and Baldwin, C.K. (2005). "Nonstationary hydrological time series forecasting using nonlinear dynamic methods." *Journal of Hydrology*, Vol. 307, pp. 164-174.
- Jang, S.W., Seo, L., Kim, T.W., and Ahn, J.H. (2011). "Non-stationary Rainfall Frequency Analysis Based on Residual Analysis." *Journal of Korean Society of Civil Engineers*, Vol. 31, No. 5B, pp. 449-457.
- Kim, B.S., Lee, J.K., Kim, H.S., and Lee, J.W. (2011). "Non-stationary Frequency Analysis with Climate Variability using Conditional Generalized Extreme Value Distribution." *Journal of Korean Wetlands Society*, Vol. 13, No. 3, pp. 499-514.
- Kim, D.H. (2011). *Bayesian statistics using R and WinBUGS*, Free Academy, pp. 209-216.
- Kim, W.S., Shin, J.Y., Um, M.J., and Heo, J.H. (2012). "Analysis of Non-stationary Characteristics for Rainfall with the Trend Analysis of L-Moments." *Journal of Korean Society of Hazard Mitigation*, Vol. 12, No. 3, pp. 71-80.
- Kottegoda, N.T., Natale, L., and Raiteri, E. (2011). "Simulation of climatic series with nonstationary trend periodicities." *Journal of Hydrology*, Vol. 398, pp. 33-43.
- Lee, C.H., Ahn, J.H., and Kim, T.W. (2010). "Evaluation of Probability Rainfalls Estimated from Nonstationary Rainfall Frequency Analysis." *Journal of Korea Water Resources Association*, Vol. 43, No 2, pp. 187- 199.
- Lee, J.J., Kwon, H.H., and Hwang, K.N. (2010). "Concept of Seasonality Analysis of Hydrologic Extreme Variables and Design Rainfall Estimation Using Nonstationary Frequency Analysis." *Journal of Korea Water Resources Association*, Vol. 43, No. 8, pp. 733-745.
- 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.
- Sung, J.H., Kim, B.S., Kang, H.S., and Cho, C.H. (2012). "Non-stationary Frequency Analysis for Extreme Precipitation based on Representative Concentration Pathways (RCP) Climate Change Scenarios." *Journal of Korean Society of Hazard Mitigation*, Vol. 12, No 2, pp. 231-244.
- Thyer, M., and Kuczera, G. (2003). "A hidden Markov model for modelling long-term persistence in multi-site rainfall time series 1. Model calibration using a Bayesian approach." *Journal of Hydrology*, Vol. 275, pp. 12-26.
- Villarini, G., Smith, J.A., and Napolitano, F. (2010). "Non-stationary modeling of a long record of rainfall and temperature over Rome." *Advances in Water Resources*, Vol. 33, pp. 1256-1267.

paper number : 14-054

Received : 1 July 2014

Revised : 23 November 2014 / 25 November 2014

/ 23 December 2014

Accepted : 23 December 2014