



## Estimation of design floods for ungauged watersheds using a scaling-based regionalization approach

Kim, Jin-Guk<sup>a</sup> · Kim, Jin-Young<sup>b</sup> · Choi, Hong-Geun<sup>a</sup> · Kwon, Hyun-Han<sup>a\*</sup>

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

<sup>b</sup>ISAN Corporation

Paper number: 18-031

Received: 4 May 2018; Revised: 17 July 2018 / 24 July 2018; Accepted: 24 July 2018

### Abstract

Estimation of design floods is typically required for hydrologic design purpose. Design floods are routinely estimated for water resources planning, safety and risk of the existing water-related structures. However, the hydrologic data, especially streamflow data for the design purposes in South Korea are still very limited, and additionally the length of streamflow data is relatively short compared to the rainfall data. Therefore, this study collected a large number design flood data and watershed characteristics (e.g. area, slope and altitude) from the national river database. We further explored to formulate a scaling approach for the estimation of design flood, which is a function of the watershed characteristics. Then, this study adopted a Hierarchical Bayesian model for evaluating both parameters and their uncertainties in the regionalization approach, which models the hydrologic response of ungauged basins using regression relationships between watershed structure and model. The proposed modeling framework was validated through ungauged watersheds. The proposed approach have better performance in terms of correlation coefficient than the existing approach which is solely based on area as a predictor. Moreover, the proposed approach can provide uncertainty associated with the model parameters to better characterize design floods at ungauged watersheds.

**Keywords:** Scaling, Bayesian, Design floods, Ungauged watershed, Regionalization

## 스케일링 기법 기반의 지역화를 통한 미계측 유역의 설계 홍수량 산정

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

<sup>a</sup>전북대학교 토목공학과, <sup>b</sup>(주)이산 수자원부

### 요 지

설계홍수량 산정은 일반적으로 수자원설계 목적을 위해 요구되며 수자원 관련 계획, 안정성 그리고 수공구조물의 위험도를 평가하기 위해 추정된다. 그러나 설계목적에 위한 국내의 유량자료는 매우 제한적이며, 강우자료와 비교해 봤을 때 상대적으로 관측년수가 상당히 적은 실정이다. 이러한 점에서 본 연구에서는 기 수립된 하천의 재현기간에 따른 설계홍수량 및 유역특성인자(면적, 유역경사)로부터, 설계홍수량을 멱함수 형태로 지역화하여 미계측 유역에서 설계홍수량 산정이 가능한 모형을 개발하였다. 제안된 설계홍수량 지역화 모형의 매개변수 산정과 불확실성을 정량적으로 평가하기 위해 계층적 Bayesian 모형을 활용하였으며, 최종적으로 교차검증 관점에서 모형의 적합성을 검증하였다. 모형 적용 결과, 기존 면적기반의 홍수량 산정식에 비해 약 0.3 이상 높은 상관성을 가지며 홍수량을 추정하는 결과를 확인하였다. 본 연구를 통해 제안된 모형은 검증과 성과 도출된 결과를 통해 유역특성에 따른 재현기간별 설계홍수량을 효과적으로 재현하는데 유리할 뿐만 아니라, 동시에 모형의 매개변수 및 결과에 대한 불확실성 정보를 제공함으로써 미계측 유역의 홍수량을 평가하는 기초자료로서 활용 가능할 것으로 판단된다.

**핵심용어:** Scaling, Bayesian, 설계홍수량, 미계측 유역, 지역화

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

## 1. 서론

최근 기상변동성 증가에 따라 국지성 집중호우로부터 동반하는 돌발홍수에 대한 빈도가 증가하는 경향을 보이고 있는 등 예측하지 못한 기상재난으로 인해 피해 규모가 대형과 되어가는 추세이다(Kim *et al.*, 2010; Kim *et al.*, 2012). 가뭄·홍수 등 수재해 예방 측면에서, 정확한 홍수량 추정 은 가장 핵심적인 사항이라 할 수 있다. 한국조사연보(MLIT, 2016)에 따르면, 국토교통부, 한국수자원공사, 환경부에서 우리나라에 약 487여 개소의 유량관측소를 설치하여 주요지점에 대한 홍수량을 모니터링하고 있다. 그러나 우리나라의 경우, 주요지점을 제외한 유역에서는 양질의 홍수량 자료 취득이 어려우며, 매년 수위-유량 관계 곡선식 등을 개발하여 홍수량을 추정하고 있는 실정이다. 관측 홍수량자료가 가용한 경우, 일반적으로 설계홍수량 산정은 홍수빈도해석을 적용하는 것이 가장 적절한 방법으로 알려져 있다(Stedinger *et al.*, 1993). 그러나 미계측 유역에 대한 자연유출량 자료는 취득이 어려울 뿐만 아니라, 강수량 자료에 비해 상대적으로 자료연한이 짧아 빈도해석 시 불확실성이 크게 발생한다는 단점이 있다(Chae *et al.*, 2018).

기존의 미계측 유역 홍수량 추정과 관련된 연구에서는 확률강우량의 확률분포 선정과 강우-유출모형의 매개변수 추정에 관한 연구가 주로 수행되었으며, 유역 특성을 반영한 미계측 유역의 홍수량 추정 연구는 다소 미진하였다(Kwon *et al.*, 2004). 그러나 최근 미계측 유역의 정확한 확률홍수량 산정을 위한 극치홍수사상 재현, 지형인자를 활용한 유역의 홍수량 추정 및 매개변수에 대한 불확실성 분석 연구 등 미계측 유역의 확률홍수량 예측을 위한 국내외 연구가 활발히 진행 중에 있다.

미계측 유역의 홍수량 추정 방법에는 지역홍수빈도분석(regional flood frequency analysis), 확률강우량과 강우-유출관계를 이용하는 방법, 강우-유출모의에 의한 동역학적 분석방법 등이 있으며(Kim *et al.*, 2016b), 계측 유역에 대한 지역홍수빈도분석 결과를 활용하여 미계측 유역의 홍수량을 간접적으로 추론하는 방법이 일반적이다(Kim and Won, 2004).

지역홍수빈도분석법은 유역단위에서 특정 빈도로 발생할 수 있는 홍수량의 크기를 판단하는데 목적이 있으며, 지역화를 통해 유역의 수문량에 대한 동질성을 평가하고 이를 기반으로 수문자료를 통합적으로 고려한 빈도분석 절차가 수반된다. 홍수량의 지역홍수빈도분석 방법으로는, Darymple (1960)이 제안한 지수홍수법, 모멘트추정법, 전이법 등이 있으며, 유역면적과의 경험식을 이용하는 방법, 합리식을 이용하는 방법, 유역의 지형학적, 수문기상학적 요소를 이용한 다

중회귀분석 등의 방법을 통하여 미계측 유역의 확률홍수량을 산정할 수 있다(Brath *et al.*, 2001).

미계측 유역의 홍수량 산정 및 매개변수 불확실성의 평가와 관련된 연구는 다음과 같다. Chae *et al.* (2018)은 설계홍수량 산정방법 중 홍수빈도 해석법과 설계 강우-유출 관계 분석법의 차이가 얼마만큼 발생하는지 확인을 통해, 설계 강우-유출 관계 분석법으로 산정된 값을 보정하여 확률홍수량을 산정하는 방법을 제시하였다. Lee *et al.* (2017)은 미계측 유역의 설계홍수량을 산정하는 방안으로 유역의 면적비를 적용하는 방안을 제안하였으며, Kim *et al.* (2016a)은 지방하천의 유역 특성과 유출량과의 관계를 파악하고, 설계홍수량을 추정하는 회귀모형을 구성하여 유역특성과 설계홍수량과의 관계를 통계학적으로 분석하였다. Kwon *et al.* (2013b)은 극치강수량 모의 후, Bayesian HEC-1 모형과 연계하여 강우-유출 관계의 불확실성을 고려한 홍수빈도곡선을 개발하였다. Yi and Lee (2013)는 fuzzy 회귀분석기법을 이용하여 미계측 산지하천유역의 설계홍수량 산정 모형을 개발하였으며, Choi and Lee (2006)는 지형의 기후학적 단위유량도로부터 미계측 소유역의 돌발홍수 기준우량을 제시하기 위한 한계유출량을 산정하였다.

외국에서는 Tofiq and Guven (2014)가 기후변화에 따른 홍수량 변화를 평가하기 위해 선형 유전자 프로그래밍(linear genetic programming, LGP) 및 전지구기후모형(global circulation model, GCM)을 연계하여 미래의 설계홍수량을 전망하였다. Ohnishi *et al.* (2004)는 Creager *et al.* (1945)이 제시한 유역면적과 가능최대홍수량(probable maximum flood, PMF) 사이의 비선형성을 수식화하여 제공하는 비홍수량산정기법의 일종인 Creager 공식을 이용하여 유역면적과 침투유량의 상관관계를 검증하였으며, Rogger *et al.* (2012)는 호주에서 홍수빈도분석을 통해 산정된 설계홍수량과 강우-유출모형을 통해 산정된 설계홍수량의 차이를 비교하였다. Jakob and Jordan (2001)은 과거에 기록된 홍수사례를 분석하여 설계홍수량 산정 시 발생하는 문제점을 제기하였으며, 지형학적 인자를 활용하여 홍수빈도분석 연구를 수행한 바 있다.

이상의 국내외 연구에서 제시된 연구방법들은 공통적으로 홍수량 추정을 위해서 강우-유출 모형의 매개변수를 관측 강우-유량 자료를 통해 검보정하거나 지형자료를 이용한 경험식을 이용하여 추정한다. 그러나 우리나라 중소규모 유역의 경우 대부분 미계측 유역인 점을 감안했을 때, 비홍수량을 이용한 경험적인 방법이 일반적인 방법이라 할 수 있다. 이러한 점에서 본 연구에서는 계층적 Bayesian GLM (generalized linear model) 기법을 이용한 경험적 설계홍수량 산정방법을

제시하고자 한다.

본 연구에서 제안하는 모형과 기존에 활용되는 미계측 유역 홍수량 추정 모형의 다른 점은 세 가지로 요약할 수 있다. 첫째, 기존의 미계측 유역 홍수량 추정 시 단일 유역을 대상으로 홍수량을 추정하는 반면, 본 연구에서는 권역별 다수의 설계홍수량자료를 활용하였으며, 이를 토대로 우리나라 4대 권역 설계홍수량에 대하여 Bayesian GLM 기법을 통해 지역화를 수행하였다. 둘째, 지역화된 모형으로부터 권역을 대표하는 회귀분석에 근거한 홍수량 산정식을 제시함과 동시에, 권역 내 미계측 유역의 홍수량 추정이 가능하다는 장점이 있다. 셋째, 유사한 통계적 특성을 가진 자료집단들이 서로 정보를 교환하고 매개변수의 불확실성을 저감할 수 있는 계층적 Bayesian 모형을 활용하여 매개변수의 불확실성 구간을 추정할 수 있는 방안을 제시함으로써 미계측 유역 홍수량 추정 시 정량적인 매개변수의 불확실성 분석이 가능하다는 장점이 있다.

최종적으로 본 연구에서 제시된 방법론에 대한 적합성을 평가하기 위해 기수립 설계홍수량 자료를 통해 모형을 검증하였으며, 개발된 모형을 대상으로 설계홍수량이 수립된 유역을 미계측 유역으로 가정하여 교차검증 관점에서 모형의 검증을 수행하였다.

본 논문의 구성은 다음과 같다. 1장에서는 본 연구와 관련된 기존 연구 내용에 대해 서술하였으며, 2장에서는 일반적인 홍수량 추정 방법 및 본 연구에서 제안한 기법에 대해 설명하였다. 3장에서는 개발된 미계측 홍수량 산정모형의 검증 및 재현기간별 홍수량 분석을 수행한 결과를 제시하였으며, 4장에서는 결론과 향후 연구 방향을 기술하였다.

## 2. 연구방법

본 연구에서는 하천기본계획에서 산정된 재현기간에 따른 홍수량에 대한 지역화를 수행하였다. 이러한 점에서 본 절에서는 미계측 유역에서 활용되는 홍수량 추정 이론과 본 연구에서 개발한 계층적 Bayesian 기반 미계측 유역의 홍수량 지역화 방법을 요약하였다.

### 2.1 강우-유출 방법

강우-유출 방법은 유역특성에 따른 강우-유출 관계를 수립하고 이를 이용하여 특정 강우량에 대한 유출량을 구하는 방법이며, 입력되는 확률강우량의 빈도와 추정되는 홍수량의 빈도가 같다는 기본가정 하에 분석이 이루어진다. 대부분의 강우-유출 모형은 총강우량 중 유효우량에 따른 직접유출량

과 기저유량을 합하여 산정하며, 일반적으로 하천기본계획 수립 시 활용되는 강우-유출 방법에 의한 설계홍수량 산정 절차는 다음과 같다.

- (1) 유역 상하류의 홍수량의 변화를 파악할 수 있는 정도의 구간 설정 및 유역 면적 분할, 지류 합류점 및 주요 구조물 지점 등을 전반적으로 고려하여 홍수량 산정지점 선정
- (2) 강우량자료의 빈도분석을 통한 강우지속기간별 재현기간별 확률강우량 산정 및 강우강도식 유도
- (3) 확률강우량을 Huff 방법 등을 통해 시간분포 후 우량주상도 작성
- (4) SCS 방법의 유출곡선지수(CN) 등을 이용하여 손실우량을 분리 후 유효우량주상도 작성
- (5) 강우지속기간별 재현기간별 유효우량주상도를 단위에 적용하여 직접유출수문곡선 및 기저유량을 고려한 소유역별 홍수수문곡선 산정
- (6) 홍수수문곡선을 활용하여 저수지추적, 홍수수문곡선 합성 등을 통하여 전체 유역의 홍수수문곡선 작성
- (7) 첨두홍수량이 최대가 되는 임계지속기간을 선정하여 재현기간별 홍수량 산정
- (8) 재현기간별 산정지점별 첨두홍수량에서 채택된 계획빈도 첨두홍수량을 설계홍수량으로 결정

### 2.2 비홍수량을 활용한 빈도별 비홍수량 추정 모형

미계측 유역에서 정확한 홍수량을 추정하기 위해서는 수문학적으로 동질한 지역의 다지점 정보(홍수량 자료)를 활용한 지역화 과정이 필요하다(Kim et al., 2016b). 이에 본 연구에서는 권역별 기수립된 하천정비기본계획 보고서로부터 강우-유출방법을 통해 산정된 유역면적 1,000 km<sup>2</sup> 이하의 재현기간(return period)에 따른 확률홍수량과 해당유역의 비홍수량(유역면적, 유역경사) 자료를 구축하였다.

유역경사가 유역의 유출에 미치는 영향을 살펴보면, 유역경사가 클수록 지표면 유출의 속도가 빨라져 유역의 집중시간이 짧아지며 동시에 첨두유량은 커지게 된다. 뿐만 아니라, 유역경사가 급할수록 사면의 식생밀도는 조밀하게 되고 토양은 침식이 용이하므로 침투능이 저하되어 유출을 가속시키는 결과를 초래한다(Fleming, 1977). 유역면적은 유출량과 직접적인 관계가 있으며 강우-유출관계를 표시하는 경험공식의 주요 변수로 사용되고 있다. 홍수량과 유역면적 간의 경험식은 멱함수 법칙(power law)으로 설명될 수 있으며, 일반적인 멱함수 형태의 미계측 유역의 홍수량 산정식은 Eq. (1)과 같이 나타낼 수 있다(Yoon, 2007; Kim et al., 2016b).

$$Q = cA^m \tag{1}$$

여기서,  $Q$ 는 홍수량( $m^3/s$ )을 의미하며,  $A$ 는 유역면적( $km^2$ ),  $c$ 와  $m$ 은 회귀분석에 의해 결정되는 회귀상수를 나타낸다. 멱함수 형태의 미계측 유역 홍수량 지역화 수행 시 홍수량은 유역

면적의 멱승에 비례하므로 홍수량과 유역면적 사이의 상관성을 검토하는 과정이 요구된다(Fill and Steiner, 1998; Malamud and Turcotte, 2006; Kim *et al.*, 2016b). 본 연구에서는 권역별 강우-유출 방법을 통해 산정된 설계홍수량에 대한 지역화 분석을 수행하기에 앞서, 멱함수 형태의 유역면적과 재현기간

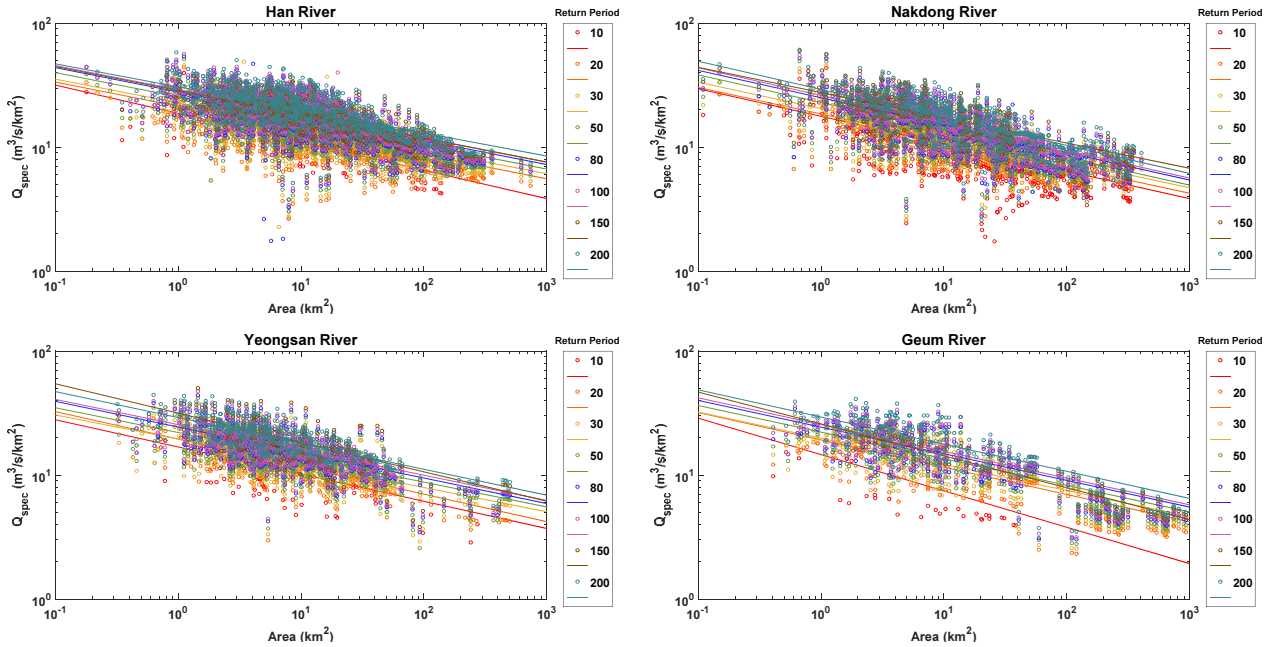


Fig. 1. Correlation analysis of specific discharge - area

Table 1. Estimated a set of parameters of the existing approach for different return periods and watersheds

Watershed	Return period	Parameter		R	Return period	Parameter		R
		$c$	$m$			$c$	$m$	
Han river	10	18.7136	-0.2292	0.740	80	27.9256	-0.1955	0.660
	20	21.3733	-0.1951	0.647	100	29.0791	-0.1947	0.664
	30	22.7824	-0.1906	0.628	150	28.6774	-0.1928	0.694
	50	25.5468	-0.1948	0.661	200	30.6750	-0.1863	0.659
Nakdong river	10	17.7659	-0.2217	0.679	80	25.0423	-0.2215	0.706
	20	18.5538	-0.2135	0.710	100	26.1482	-0.2224	0.715
	30	20.4838	-0.2133	0.665	150	27.5672	-0.2037	0.674
	50	22.9359	-0.2225	0.691	200	29.2981	-0.2263	0.732
Yeongsan river	10	16.8895	-0.2196	0.638	80	24.4614	-0.2061	0.663
	20	19.5077	-0.2221	0.703	100	25.5238	-0.2039	0.654
	30	19.5121	-0.1954	0.613	150	31.6135	-0.2368	0.715
	50	22.0606	-0.2002	0.638	200	29.1514	-0.2090	0.716
Geum river	10	14.6580	-0.2936	0.656	80	24.4148	-0.2153	0.769
	20	19.2347	-0.2203	0.793	100	25.5564	-0.2160	0.777
	30	19.8591	-0.2111	0.753	150	25.8658	-0.2575	0.941
	50	22.2466	-0.2149	0.766	200	29.4292	-0.2191	0.792

에 따른 설계홍수량의 상관성을 확인하였다. 권역별 비홍수량과 유역면적 사이의 관계는 Fig. 1에 도시하였으며(여기서,  $Q_{spec}$ 는 비홍수량(specific discharge) ( $m^3/s/km^2$ )을 의미), 역함수 형태의 선형상관분석 결과는 Table 1과 같다.

Table 1에 나타냈듯이, 면적과 비홍수량의 상관계수(correlation coefficient)  $R$ 은 평균적으로 0.65 정도를 갖는 것으로 파악되었다. 단일 설명변수로 면적을 사용했다는 측면에서 비교적 높은 상관계수라 할 수 있지만 회귀선을 중심으로 산포도가 매우 큰 것을 확인할 수 있으며, 면적 이외에 다른 지형학적 인자를 고려한 회귀식 개발이 필요할 것으로 판단된다. 이러한 점에서 본 연구에서는 미계측 유역의 홍수량 추정을 위한 회귀식을 개발하는데 있어 유역면적과, 유역경사, 재현기간 등을 고려하였으며, 최종적으로 모형의 매개변수에 대한 불확실성에 대하여 정량적으로 평가하기 위해서 계층적 Bayesian 모형 기반의 미계측 유역 홍수량 산정 모형을 개발하였다.

본 연구에서 제안하는 계층적 Bayesian 모형은 매개변수(parameter)를 단계적으로 모형화하여 각각의 매개변수에 사전분포(prior distribution)를 정해주면, 계층모형의 모든 매개변수에 대한 사후분포(posterior distribution)를 계산하는 것이 가능하다는 장점이 있다. 즉 유사한 통계적 특성을 가지는 자료집단들이 정보를 서로 교환함으로써 추정되는 매개변수의 불확실성을 저감하도록 고안된 계층적 Bayesian 모형의 장점을 이용하는데 목적이 있다(Kwon et al., 2013b). 이에 본 연구에서는, Bayesian 모형을 기반으로 매개변수의 불확실성을 정량화할 수 있도록 모형을 개발함으로써 추정되는 미계측 유역의 홍수량에 대한 신뢰성을 개선할 수 있도록 하였다.

### 2.3 Bayesian GLM 기법을 활용한 미계측 유역 홍수량 지역화

최근 국내외에서는 수문학적 빈도분석 시 불확실성 평가 및 해석결과에 대한 신뢰성 평가가 일반적으로 요구되고 있으며, 다양한 관점에서 불확실성을 의사결정에 활용하고자 노력하고 있다(Kim et al., 2017). 본 연구에서는 권역별로 기 수립된 빈도별 홍수량 자료와 비홍수량을 활용하여 홍수량의 공간적 지역화를 수행하였다. Bayesian GLM 기법을 도입하여 역함수 형태의 미계측 유역 홍수량 산정모형을 구축하였으며, 해석 결과에 대한 불확실성 구간을 제시하였다. Bayesian 기법에서 추정되는 불확실성 구간은 신뢰구간(credible interval)이라는 용어로 사용되며, 기존의 신뢰구간(confidence interval)과 유사한 개념이라 할 수 있다.

Bayesian 이론에 대한 기본개념은 다음과 같다. Bayesian 이론(Bayes's theorem)이란 두 확률변수의 조건부확률을 나

타내는 정리이다. A라는 사건이 발생하였을 때의 확률( $P(A)$ )을 기준으로 B라는 새로운 자료에 의해서 조건부확률이 갱신되어 사건이 일어날 확률( $P(A|B)$ ) 사이의 관계를 나타낼 수 있다. 여기서,  $P(A)$ 는 사전확률(prior probability),  $P(A|B)$ 는 사후확률(posterior probability)을 말한다(Gelman et al., 2004). Bayesian 이론을 통한 매개변수 추정기법은 기존 방법들(최우도법, 모멘트법, 확률가중모멘트법)과는 다르게 매개변수를 하나의 확률변수로 취급하며, 매개변수가 단일 값이 아닌 확률분포의 형태로 부여되어 최종적으로 매개변수의 사후분포를 추정하는데 목적을 둔다(Kwon et al., 2013a). Bayesian 방법은 MCMC (Markov Chain Monte Carlo)를 이용하며, MCMC 방법에서 널리 사용되는 알고리즘으로는 메트로폴리스-해스팅스(Metropolis-Hastings, MH) 알고리즘과 깁스 샘플링(Gibbs sampling) 방법 등이 있다.

Bayesian 이론은 Eqs. (2)~(4)로 나타낼 수 있다. 주어진 확률변수  $\theta$ 와  $y$ 가 있을 때, 두 확률변수들의 결합확률분포는 Eq. (2)와 같이 사전분포  $p(\theta)$ 와 우도  $p(y|\theta)$ 의 곱으로 표현할 수 있다. Bayes 정리에서 매개변수  $\theta$ 와 변량  $y$ 의 조건부확률과 주변확률의 관계는 Eq. (3)과 같으며, Eq. (4)에서 좌변  $p(\theta|y)$ 는 사후분포를 나타내며 사전분포와 우도  $p(y|\theta)$ 의 곱으로 추정될 수 있다(Gelman et al., 2003).

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

$$p(\theta|y) = \frac{p(\theta, y)}{p(y)} = \frac{p(\theta)p(y|\theta)}{p(y)} \quad (3)$$

$$p(\theta|y) \propto p(y|\theta)p(\theta) \quad (4)$$

여기서,  $p(\theta)$ 는  $\theta$ 의 사전확률 또는 경계확률이다. ‘사전’이라는 의미는 사건  $y$ 에 관한 기본적인 정보를 의미하며,  $p(\theta|y)$ 는  $y$ 가 주어졌을 때  $\theta$ 의 조건부 확률을 나타낸다.  $p(\theta|y)$ 는 특정한 값  $y$ 에 의해 추정되기 때문에 사후확률이라 부른다.  $p(y|\theta)$ 는  $\theta$ 가 주어졌을 때의  $y$ 에 대한 조건부 확률을 나타내며,  $p(y)$ 는  $y$ 의 주변확률로 정규화 상수의 역할을 한다.

동일 빈도의 호우규모로 인한 유역으로부터의 침투홍수량을 유역면적으로 나누어 단위유역 면적당 유량, 즉 비홍수량으로 표시해보면, 비홍수량은 유역면적이 클수록 작은 값을 나타낸다. 이는 유역면적이 클수록 홍수도달시간이 길어질 뿐만 아니라 저류로 인한 지체효과가 더 크기 때문에 나타나는 현상으로 이해할 수 있으며, 유역의 수문응답특성에 대한 분석 시 유리하다고 알려져 있다(Yoon, 2007). 본 연구에서는

유역들에 대한 설계홍수량과 유역면적의 관계가 멱함수 법칙을 따른다고 가정하였으며, 유역의 지역화를 통해 미계측 유역의 홍수량을 평가하였다. 본 연구에서 제안한 홍수량 지역화 모형은 미계측 유역 홍수량 산정식인 Eq. (1)의 홍수량( $Q$ ) 함수를 Eq. (5)와 같은 비홍수량( $Q_{spec}$ )에 대한 멱함수 형태로 구축하였다.

$$Q_{spec} = \alpha A^\beta \quad (5)$$

여기서,  $\alpha, \beta$ 는 본 연구에서 비홍수량을 활용한 회귀분석에 의해 결정되는 회귀상수를 나타낸다. 유출량과 유출 특성 변수들의 대수 값은 선형관계를 갖는 것으로 알려져 있다. 이에 본 연구에서는 대수정규분포(lognormal distribution)를 Bayesian GLM 모형의 확률분포형으로 선정하였다(Riggs, 1973). 대수정규분포에 대한 확률밀도함수(probability density function) 및 누가확률밀도함수(cumulative density function)는 Eqs. (6) and (7)과 같다.

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

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

여기서,  $\mu$ 는 위치매개변수(location parameter),  $\sigma$ 는 규모매개변수(scale parameter)를 나타낸다. 홍수량 자료가 대수정규분포를 따른다고 가정하였을 때, 본 연구에서 활용된 GLM 모형은 Eq. (8)과 같이 나타낼 수 있다.

$$Y(\text{specific discharge}) = LN(\mu(t), \sigma^2) \quad (8)$$

여기서,  $LN$ 은 대수정규분포를,  $t$ 는 유역별 홍수량 자료의 개수를 나타내며 Bayesian GLM 모형에 사용되는 모든 매개변수는 각각의 확률분포를 가지게 된다. 즉, Eq. (8)의  $\mu(t)$ 는 지형인자를 예측인자로 활용되는 회귀식 형태의 연결함수(link function)를 의미한다. 연결함수의 회귀매개변수와  $\sigma$ 는 각각의 사전분포로 정의된다. 앞서 언급하였듯이, 유출량자료는 대수정규분포로 가정하였으며 연결함수로서 Eq. (5)의 유량과 비홍수량과의 멱함수를 활용하여 정리하면 Eq. (9)와 같은 선형회귀(linear regression) 형태로 나타낼 수 있다.

$$p(\mu(t)) \sim \log(\alpha(t)) + \beta \cdot \log(\text{Area}(t)) \quad (9)$$

여기서,  $\beta$ 는 면적에 대한 회귀계수를 나타낸다. Eqs. (8) and (9)의 매개변수들에 대한 사전분포는 Eqs. (10) and (11)과 같이 정의될 수 있으며,  $\Gamma$ 는 감마분포(gamma distribution)를,  $N$ 은 정규분포(normal distribution)를 나타낸다.

$$p(\sigma^2) \sim \Gamma(1, 1) \quad (10)$$

$$p(\beta) \sim N(0, 1) \quad (11)$$

본 연구에서 기존 연구와 다른 점은 면적뿐만 아니라 유역경사와 설계홍수량의 설계빈도를 통합적으로 해석할 수 있는 공식을 제시하는 것이라 할 수 있다. 이러한 점에서 유역면적에 해당하는 평균유량( $\beta \log(\text{Area})$ )을 기준으로 유역경사와 설계빈도를 고려하기 위하여 절편에 해당하는  $\alpha(t)$ 를 유역경사와 설계빈도의 연결함수 형태로 가정하였다. 즉, 경사가 급해지면 동일 유역면적에 해당하는 평균유량( $\beta \log(\text{Area})$ )에 가장효과를 주게 되며, 동일하게 설계빈도가 커지게 되면 평균유량을 기준으로 선형적으로 증가된 유량으로 추정되도록 하였다. 이러한 관계를 식으로 나타내면 Eq. (12)와 같은 선형회귀 형태로 나타낼 수 있다.

$$p(\alpha(t)) = \exp(-2(\theta_1 + \theta_2 (\log(\text{Slope}(t))) + \theta_3 (E(t)))) \quad (12)$$

$$\theta_1 \sim N(0.0, 0.0001) \quad (13)$$

$$\theta_2 \sim N(0.0, 0.0001) \quad (14)$$

$$\theta_3 \sim N(0.0, 0.0001) \quad (15)$$

Eqs. (13)~(15)에서  $\theta_1, \theta_2, \theta_3$ 은 홍수량의 평균을 추정하기 위한 회귀매개변수로서 사전분포를 대수정규분포의 공액분포(conjugate prior)인 정규분포로 선정하였다. 공액분포는 자료의 분포와 동일한 확률분포를 사전분포로 정의하는 방법으로 최종적으로 추정된 사후분포가 가정한 분포형태를 갖도록 유도하는 확률통계 개념이며, 매개변수의 수렴 등 안정성을 확보하는데 유리한 장점이 있어 회귀식의 매개변수는 일반적으로 정규분포로 가정한다(Gelman *et al.*, 2004). Eqs. (13)~(15)의 정규분포를 이용한 사전분포에서 0.0001은 정밀도(precision)를 의미하며 분산의 역수를 의미한다. 즉, 매우 폭이 넓은 사전분포를 사용하였으며 이는 추정해야 할 매개변수에 비해 상대적으로 많은 자료가 이용되는 경우 적절하다고 알려지고 있다(So *et al.*, 2017). 본 연구에서 제안하는 홍수

량 산정식은 미계측 유역에서 홍수량에 대한 재현기간을 고려하여 추정하였다. 이에 Eq. (12)에 홍수량의 초과확률(exceedance probability) 인자를 추가하여 홍수량 회귀모형을 구축하였으며,  $E(t)$ 는 초과확률 인자로 재현기간의 역수이다. 본 연구에서 활용한 사후분포의 형태는 Bayesian 이론에 의해 Eq. (16)과 같이 나타낼 수 있다(Gelman *et al.*, 2003).

$$p(\theta|y) = \frac{p(\theta, y)}{p(y)} \propto p(y|\theta) \cdot p(\theta) \quad (16)$$

여기서,  $\theta = [\mu(\sigma, \beta, \theta_1, \theta_2, \theta_3)]$ 는 본 연구에서 추정되는 매개변수를 나타내며,  $p(\theta)$ 는 매개변수 사전분포를,  $p(y|\theta)$ 는 우도함수를 나타내며 Eq. (17)과 같이 표현할 수 있다.

$$p(y|\theta) = \prod_{t=1}^T LN(y|\mu(t), \sigma^2) \quad (17)$$

여기서, T는 자료의 개수를 의미한다. 위에서 정의된 우도함수와 매개변수들의 사전분포들을 Eq. (18)과 같이 대입시킴으로써 매개변수들의 사후분포를 추정할 수 있다.

$$\begin{aligned} p(\theta|y) &\propto \prod_{t=1}^T LN(y|\mu(t), \sigma^2) \\ &\times N\left(\mu(t) \mid \log\left(\exp(-2(\theta_1 + \theta_2(\log(\text{Slope}(t))) + \theta_3(E(t))))\right) + \beta + \log(\text{Area})\right) \quad (18) \\ &\times N(\theta_1 \mid 0, 0.0001) \times N(\theta_2 \mid 0, 0.0001) \\ &\times N(\theta_3 \mid 0, 0.0001) \times \Gamma(\sigma^2 \mid 1, 1) \times N(\beta \mid 0, 1) \end{aligned}$$

Eq. (18)의 모든 매개변수에 대하여 직접적인 적분을 통해 추정하는 것은 불가능하므로, 본 연구에서는 Bayesian MCMC 기법을 활용하였다. Bayesian MCMC 기법은 주어진 다변량 확률분포가 복잡하여 이를 따르는 IID (independent and identically distributed) 난수를 얻을 수 없는 경우에 사용가능한 기법으로써 IID 난수 대신 Markov Chain 난수를 추출하여 사용한다(Kwon *et al.*, 2013a). MCMC 기법 중 본 연구에서는 깃스샘플링을 이용하였다. 깃스샘플링은 이미 오래전부터 물리학 분야에서 복잡한 수식의 해를 찾기 위해서 사용되던 방법이었으나, Gelfand and Smith (1990)에 의하여 Bayesian 모형에 도입되어 이후 Bayesian 통계 추정 시 가장 기본이 되는 이론 중 하나이다. 깃스샘플링은 저차원 분포로부터의 표본 생성으로 구성되기 때문에 복잡한 제한조건도 쉽게 처리할 수 있다는 장점이 있다(Geman and Geman, 1984). 깃스표본

법에 대한 자세한 내용은 기존 연구문헌들을 참조할 수 있다 (Kwon *et al.*, 2008; Lee and Kwon, 2011).

Markov Chain의 수렴(convergence)여부를 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, 1998). 본 연구에서는 모형의 수렴을 확증하기 위하여 3개의 Chain을 독립적으로 시행하여 Sampling이 효과적으로 혼합(mixing) 되도록 하였으며, Trace plot을 이용하여 Chain의 수렴여부를 판단하였다. 최종적으로 Gelman-Rubin 통계량 검정결과를 이용하여 각 회귀모형의 매개변수에 대한 Markov Chain의 수렴여부를 판단하였다.

### 3. 연구결과

#### 3.1 대상유역

우리나라는 지형적으로 동쪽은 유역의 경사가 가파르고 대체적으로 하천의 길이가 짧으며, 서쪽으로는 경사가 완만하여 대하천이 발달하기 좋은 여건을 가진다(Ministry of Environment, 2006). 우리나라의 국가하천은 한강, 낙동강, 금강, 영산강, 섬진강을 기준으로 하여 5개의 대권역으로 구분할 수 있다. 그러나 본 연구에서는 영산강, 섬진강권역은 지형학적으로 인접해 있으며, 권역의 면적이 작아 권역의 유역 특성도 유사하다는 가정 하에 하나의 권역으로 간주하였다. 최종적으로 우리나라를 4개의 권역으로 분리하여 설계홍수량의 지역화 분석을 수행하였으며, 제주도 권역은 제외하였다. 우리나라를 4개의 대권역으로 분할한 권역의 형태는 Fig. 2에 도시하였다.

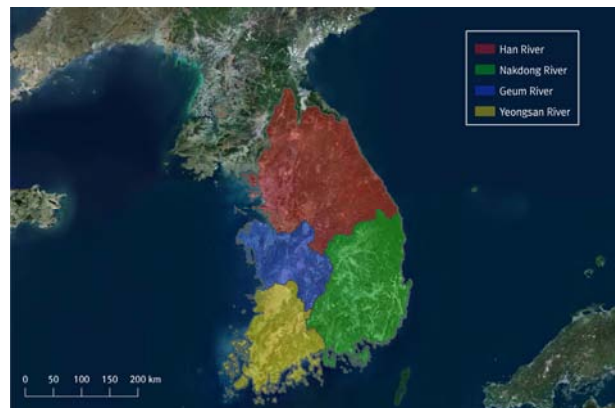


Fig. 2. A map showing Korea peninsula. Shaded area indicates four major watersheds

### 3.2 Bayesian 기법을 활용한 홍수량 지역화 분석결과

#### 3.2.1 매개변수 추정 및 불확실성 분석결과

본 연구에서는 권역별 재현기간에 따른 설계홍수량과 지형인자들을 대상으로 지역화 분석을 수행하기에 앞서, 선형 회귀 분석을 수행하였다. 그 결과는 Fig. 1에서 확인할 수 있으며, 멱함수 형태의 선형회귀 분석 시 유역면적에 따른 빈도별 홍수량의 추세선이 재현기간에 따라 선형 형태를 이루고 있는 것을 확인하였다. 그러나 유역의 특성이 완전히 고려되지 않아 특정빈도에서 추정된 홍수량의 추세선이 역전되는 현상이 나타나며, 매개변수의 불확실성이 고려되지 않는 등 이에 대한 고려가 필요하다. 이러한 점에서 본 연구에서는 모형의 매개변수와 추정되는 결과에 대한 불확실성을 고려하기 위해 Bayesian GLM 기법을 활용하여 미계측 유역에 적용 가능한 홍수량 지역화 모형을 구축하였다. 권역별 홍수량 지역화 분석 시 유역면적, 유역경사 등 비홍수량 뿐만 아니라, 특정 홍수 사상이 발생할 수 있는 확률까지 종합적으로 고려하기 위하여 재현기간을 모형의 독립변수로 활용하였다.

Bayesian 기법은 모든 매개변수에 다양한 확률분포를 부여하고 최종적으로 사후분포를 추정이 가능하다는 점에서 정규분포를 사용하는 최소자승법과 같은 선형 모형과 차별성이 있으며, 매개변수의 불확실성을 객관적으로 정량화 할 수 있다는 장점이 있다(Gelman *et al.*, 2003). 본 연구에서는 매개변수를 추정하기 위하여 MCMC 기법 중 깃스샘플링 기법을 활용하였으며, 10,000번의 모의를 수행하고 최종 2,000개의 표본(sample)을 활용하여 각 매개변수의 사후분포를 도출하였다. 모의된 사후분포 추정 시 수치적으로 안정성을 확인하기 위하여 독립된 3개의 Chain으로 부터 얻어진 2,000개의 표본에 대한 Gelman-Rubin 검정 통계량을 계산하여 매개변수의 수렴여부를 확인하였다. 연쇄가 무한대로 진행될 때, 검정 통계량의 값이 1에 가까워지면 Chain에 의해 생성된 매개변수들이 동일한 분포로 수렴한다는 것을 나타낸다(Gelman *et al.*, 2004). 최종적으로 모의된 사후분포에 대한 Gelman-Rubin 통계량 결과의 대부분이 1에 매우 가깝다는 것을 확인하였으며, 이는 모의된 표본들이 정상적으로 수렴되었다는 것을 나타낸다. Table 2는 도출된 사후분포에 대한 매개변수의 불확실성 구간을 나타내며, 매개변수별 2.5%, 50%, 97.5%의 분위(quantile)를 추출하여 매개변수의 불확실성 구간을 정량적으로 산정하였으며, Fig. 3에 매개변수의 사후분포를 도시하였다.

Fig. 3에서 파란 선은 매개변수의 사후분포를, 빨간 선은 매개변수의 중앙값(median)을 나타낸다. 기존 선형회귀방정식은 미계측 유역의 홍수량을 예측하는데 있어 자료 자체의 분

Table 2. Credible interval of model parameters derived from posterior distributions

Watershed	Parameter	Mean	2.5%	50%	97.5%
Han river	$\beta$	-0.2129	-0.2176	-0.2129	-0.2083
	$\sigma$	22.2099	18.2000	21.9600	27.5600
	$\sigma_\alpha$	22.3430	18.2650	22.1000	27.8700
	$\theta_1$	-0.6145	-0.6170	-0.6145	-0.6118
	$\theta_2$	0.0008	0.0001	0.0008	0.0015
	$\theta_3$	1.2122	1.1550	1.2120	1.2690
Nakdong river	$\beta$	-0.2314	-0.2375	-0.2314	-0.2254
	$\sigma$	17.7185	14.1350	17.5050	22.8500
	$\sigma_\alpha$	18.2584	14.4350	17.9200	23.6800
	$\theta_1$	-0.5968	-0.6001	-0.5969	-0.5934
	$\theta_2$	-0.0023	-0.0031	-0.0023	-0.0016
	$\theta_3$	0.8258	0.7796	0.8263	0.8749
Yeongsan river	$\beta$	-0.2270	-0.2345	-0.2269	-0.2195
	$\sigma$	23.9028	19.2900	23.6200	29.9650
	$\sigma_\alpha$	25.1599	20.2650	24.8300	32.1250
	$\theta_1$	-0.6015	-0.6059	-0.6014	-0.5971
	$\theta_2$	-0.0015	-0.0032	-0.0015	0.0001
	$\theta_3$	1.1698	1.0890	1.1700	1.2490
Geum river	$\beta$	-0.2692	-0.2764	-0.2693	-0.2618
	$\sigma$	20.8768	17.0150	20.6400	26.2300
	$\sigma_\alpha$	22.2374	17.7700	22.0000	27.9300
	$\theta_1$	-0.6156	-0.6204	-0.6156	-0.6111
	$\theta_2$	-0.0031	-0.0041	-0.0031	-0.0021
	$\theta_3$	1.3456	1.2220	1.3440	1.4720

포특성을 정규분포에 한정하는 문제점이 있다. 이러한 문제점은 본 연구에서 제시하는 Bayesian GLM 방법을 활용하여 개선될 수 있다. 즉 자료의 분포특성을 정규분포가 아닌 원자료가 가지는 분포를 직접 고려할 수 있다. 본 연구에서는 홍수량 자료에 적합성을 가지는 대수정규분포를 대표 분포로 결정하였으며, 분포의 특성을 결정짓는 평균값은 유역면적, 유역경사 및 재현빈도의 함수로 정의된다. Fig. 4는 본 연구에서 제안하는 Bayesian GLM 모형을 활용한 권역별 멱함수 형태의 홍수량 지역화 분석결과를 도시하였다. Fig. 4의 지역화 결과를 살펴보면, 빈도에 따라서 도시한 빈도별 비홍수량 추세선이 Fig. 1의 결과와 다르게 권역별로 모두 역전하지 않고 정상적으로 추정되는 것을 확인할 수 있다. Fig. 4에 미계측 유역의 홍수량 추정을 홍수량 지역화 결과를 Table 3에 권역별 멱함수 형태로 홍수량 산정식을 제시하였다. 여기서,  $Q_{spec}$ 는 미계측 유역의 비홍수량( $m^3/s/km^2$ ),  $A$ 는 유역면적( $km^2$ ),  $S$ 는 유역경사(%),  $T$ 는 재현기간(year)을 나타낸다.



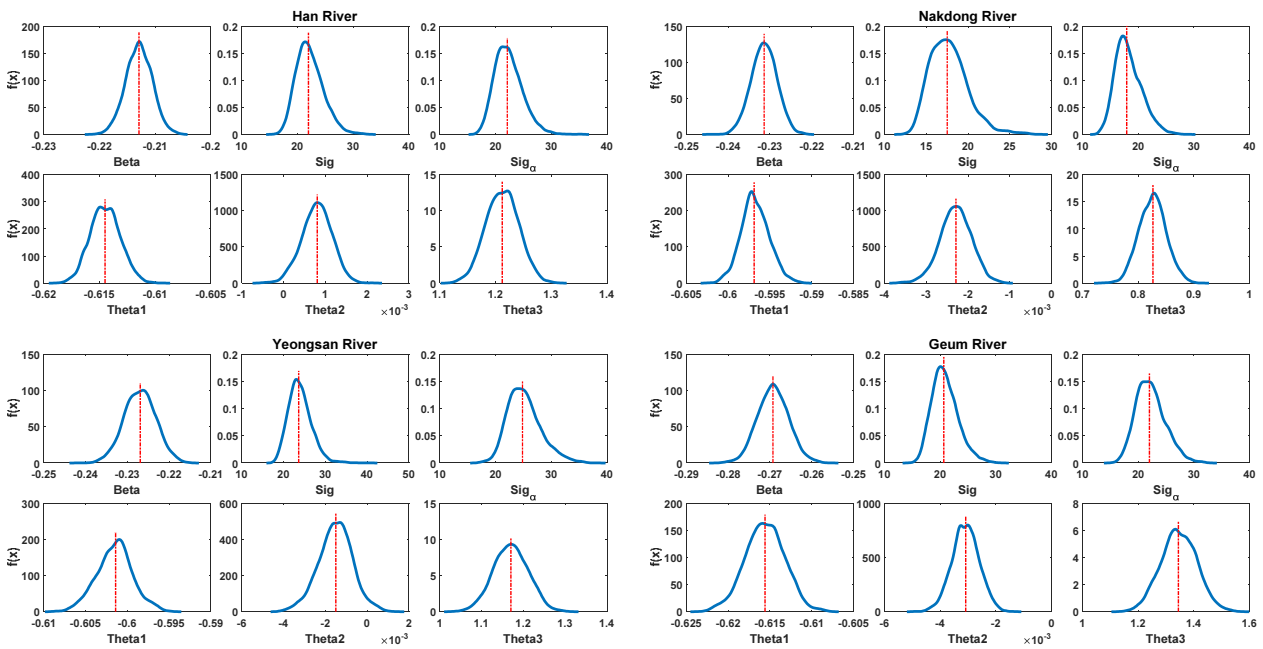


Fig. 3. Posterior distributions of model parameters at four major watersheds

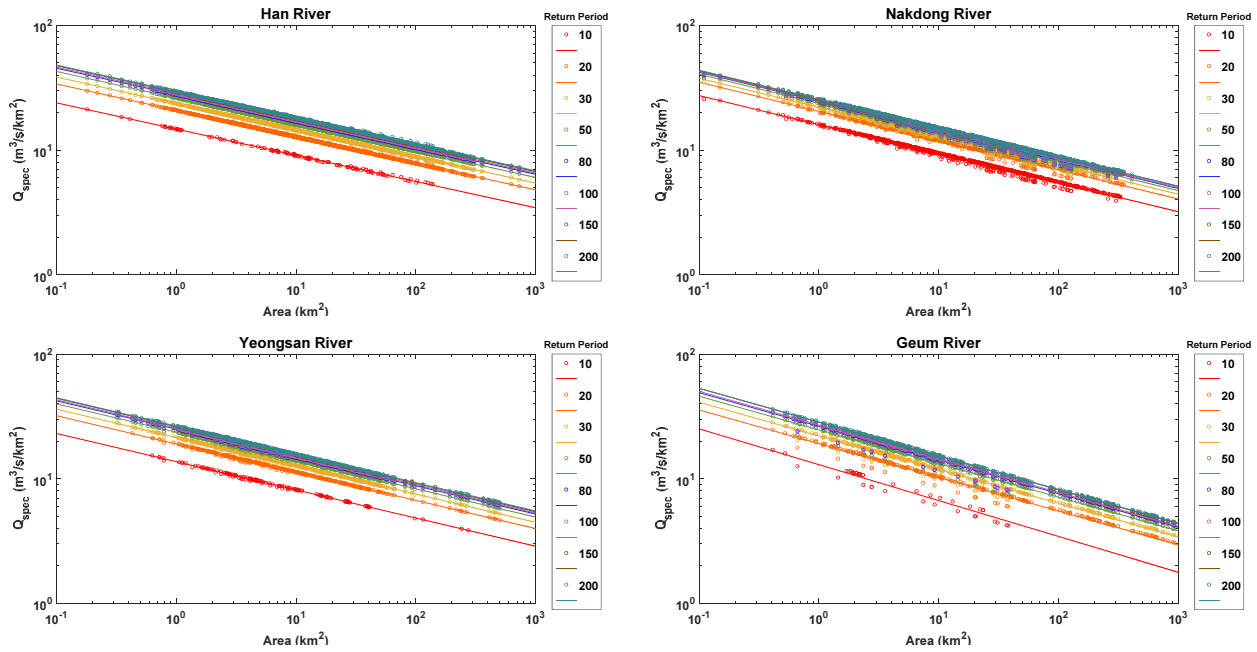


Fig. 4. Regionalization results by Bayesian GLM analysis at four major watersheds

Table 3. Regionalized design rainfall equations to estimate specific discharge for the ungauged watersheds

Watershed	Equation
Han river	$Q_{spec} = \exp(\log(\exp(\exp(-2(-0.6145 - 0.0008\log(S) + 1.2122(1 - (1/T)))))) - 0.2129\log(A))$
Nakdong river	$Q_{spec} = \exp(\log(\exp(\exp(-2(-0.5968 - 0.0023\log(S) + 0.8258(1 - (1/T)))))) - 0.2314\log(A))$
Yeongsan river	$Q_{spec} = \exp(\log(\exp(\exp(-2(-0.6015 - 0.0015\log(S) + 1.1698(1 - (1/T)))))) - 0.2270\log(A))$
Geum river	$Q_{spec} = \exp(\log(\exp(\exp(-2(-0.6156 - 0.0031\log(S) + 1.3456(1 - (1/T)))))) - 0.2692\log(A))$

### 3.3 홍수량 지역화에 따른 권역별 미계측홍수량 추정 결과

본 연구에서 제안한 권역별 설계홍수량 지역화 결과와 기 수립된 설계홍수량과의 비교를 통해 미계측 유역에서의 적용성을 검토하였다. 그 결과 모형을 통해 추정되는 홍수량이 기

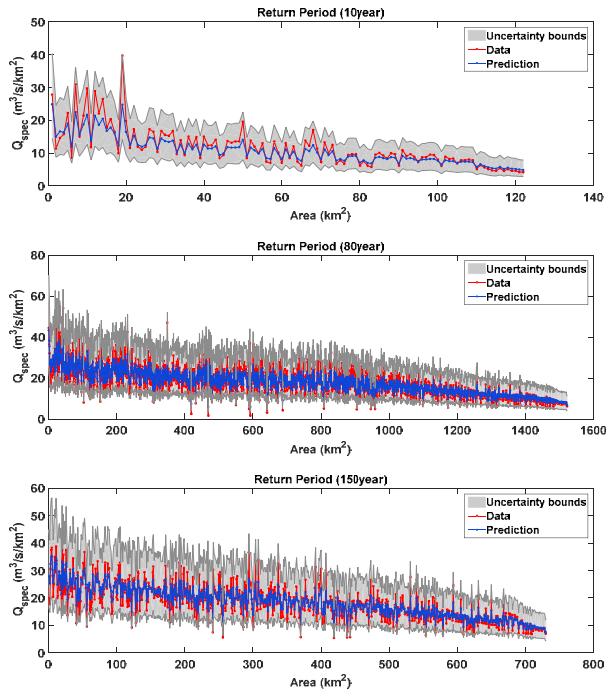


Fig. 5. A comparison of specific discharge between data and simulation for different return periods (10, 80, 150 year) at Han river watershed

존 자료의 통계적 특성을 효과적으로 재현하는 것을 확인하였다. 대표적으로 한강유역에 적용한 결과를 Figs. 5~6에 도시하였다. Fig. 5를 살펴보면, 본 연구에서 제안한 미계측 유역 홍수량 추정 모형이 기수립된 설계홍수량에 대하여 약 0.9의 높은 상관성을 가지며 재현기간에 따른 설계홍수량을 재현하는 것을 확인할 수 있다. Fig. 6의 고유량 부분에서 추정되는 홍수량은 시각적으로는 과소 추정되는 것으로 나타나지만, 본 연구에서 제시한 불확실성 구간 안에 들어오는 것을 확인하였다. Bayesian 모형의 매개변수는 자료의 개수에 따른 영향, 즉, 표본오차(sampling error)에 매우 큰 영향을 준다. 이러한 점에서 고유량 부분의 홍수량 추정에 있어 극단치 자료의 부족으로 나타나는 문제로서 추정된 매개변수의 신뢰성이 그만큼 결여된다고 볼 수 있다. 본 연구에서 제안하는 홍수량 지역화 모형은 중·소규모 유역의 면적, 경사 정보와 설계홍수량의 재현기간 등을 활용하여 미계측유역의 홍수량 추정 시 활용성을 가지는데 목적이 있다. 고유량의 홍수량 추정 시 과소 추정 되는 현상을 해결하기 위해서는, 앞서 언급한 유역특성 인자 이외의 홍수량과 연관된 변수에 대한 고려와 고유량 부분의 충분한 자료 구축을 통해 해결 가능할 것으로 사료된다.

권역별 비홍수량의 선형회귀분석을 통한 홍수량 산정 결과와 지역화를 통해 산정된 미계측 유역 홍수량에 대한 상관분석을 수행하였다. 그 결과, 미계측 유역의 지역화를 통해 추정되는 홍수량이 모형 구축 시 사용한 자료와 약 0.9 이상의 높은 상관성을 나타낸 반면 선형회귀분석은 0.6 정도의 상관성을 갖는 것으로 평가되었다. 권역별 추정된 결과에 대한 비교를 Table 4에 제시하였다.

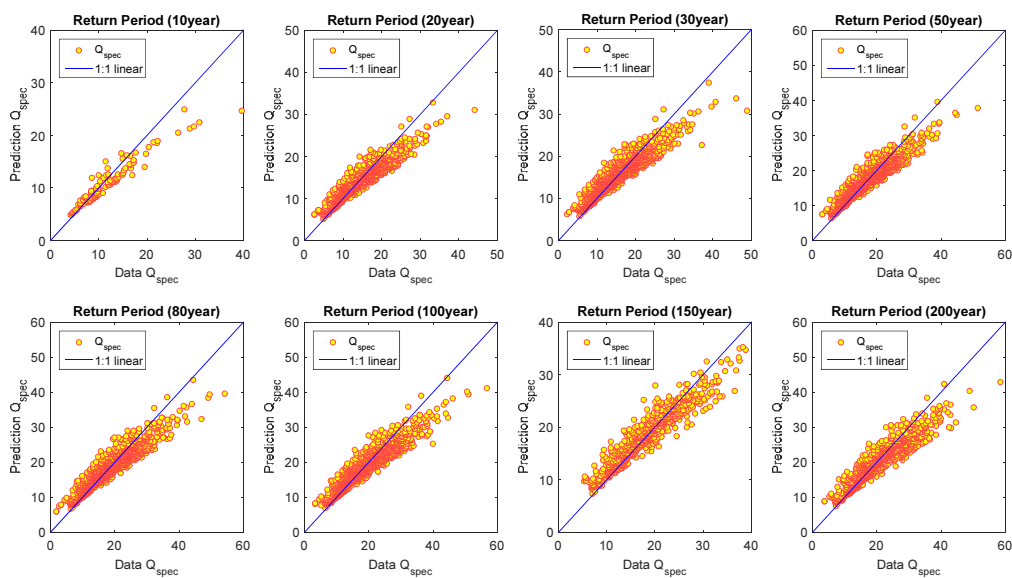


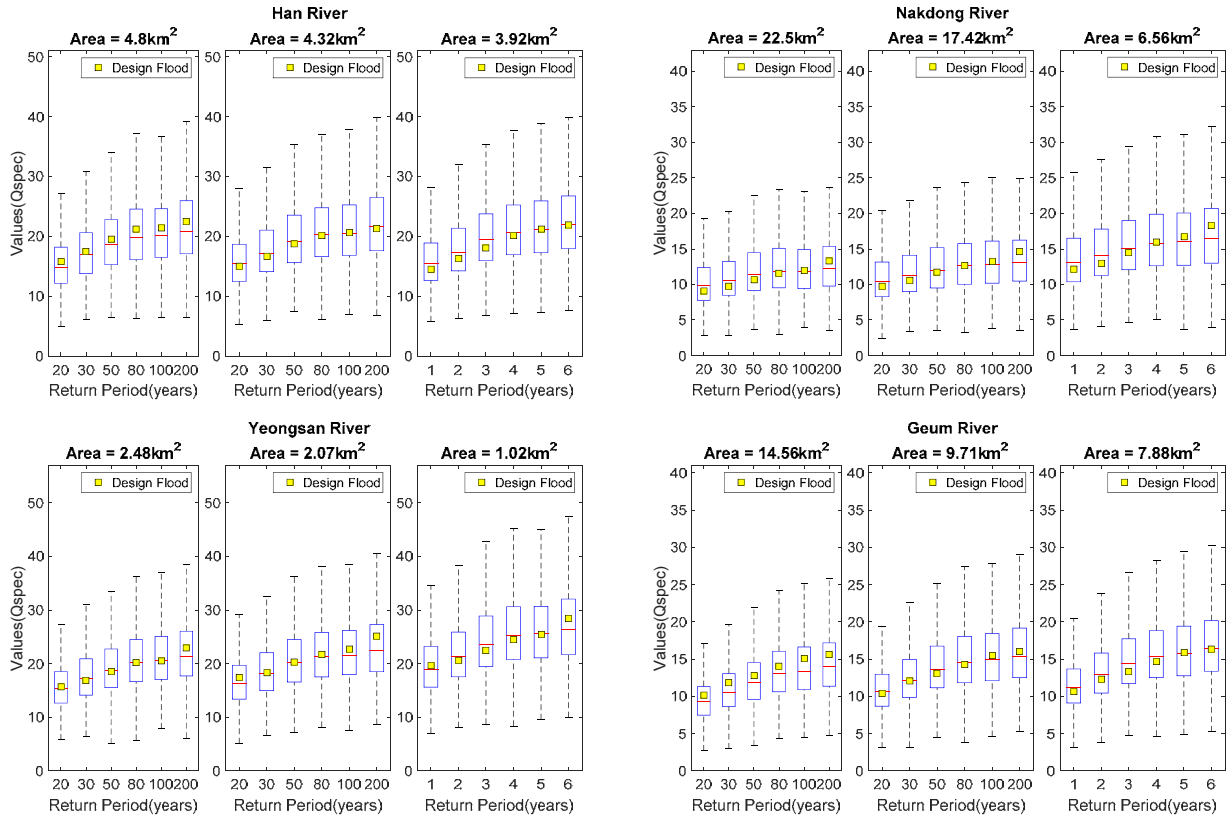
Fig. 6. A scatter plot of specific discharge for different return periods between data and simulation (Han river watershed)

**Table 4.** Comparison of model performance between the proposed model and existing approach with coefficient of correlation

Watershed	Return period	R		Return period	R	
		Linear	Bayesian GLM		Linear	Bayesian GLM
Han river	10	0.740	0.957	80	0.660	0.944
	20	0.647	0.943	100	0.664	0.943
	30	0.628	0.939	150	0.694	0.946
	50	0.661	0.944	200	0.659	0.940
Nakdong river	10	0.679	0.944	80	0.706	0.948
	20	0.710	0.945	100	0.715	0.949
	30	0.665	0.942	150	0.674	0.939
	50	0.691	0.946	200	0.732	0.952
Yeongsan river	10	0.638	0.941	80	0.663	0.934
	20	0.703	0.947	100	0.654	0.932
	30	0.613	0.926	150	0.715	0.948
	50	0.638	0.929	200	0.716	0.943
Geum river	10	0.656	0.961	80	0.769	0.944
	20	0.793	0.953	100	0.777	0.945
	30	0.753	0.941	150	0.941	0.985
	50	0.766	0.944	200	0.792	0.946

최종적으로 본 연구에서는 권역별 홍수량 지역화 분석 시 활용하지 않은 유역의 일부를 미계측 유역으로 가정하여 교차 검증(cross validation) 관점에서 모형의 적합성을 검증하였다. 각 권역별 3개의 소규모 유역을 선정하여 본 연구에서 제안한 모형을 통해 추정된 홍수량을 Boxplot을 활용하여 시각적으로 검토한 결과, 전체적으로 모의된 홍수량이 기수립된 설계홍수량 자료와 높은 상관성을 가지는 것을 확인할 수 있으며 그 결과를 Fig. 7에 도시하였다. 미계측 유역 산정 모형의 결과와 일부 유역에서 수립된 홍수량과의 적합성을 통계적 평가 수단으로 상관계수(R), 일치계수(Index of agreement, IoA), 평균제곱오차(root mean square error, RMSE) 등을 활용하여 통계적 지표로서 모형의 적합성을 평가하였다. 그 결과, 모든 권역에서 평균적으로 상관계수, 일치계수 등 통계적 지표에서 0.9 이상 높은 상관성을 나타내는 것을 확인할 수 있으며, Table 5에 권역별 통계적 분석 결과를 나타내었다.

본 연구에서 통계적 평가 수단 사용된 지표 중 상관계수 R은 모형의 효율성을 나타내는 무차원 기준으로 1에 근접할수록 모의능력이 우수함을 나타낸다. 평균 제곱근 오차(Root Mean Square Error, RMSE)는 실제 값과 모의된 값의 차이를 알려 주는데 사용되는 척도이며 0에 가까울수록 모형이 우수함을



**Fig. 7.** Cross-validated analysis of the proposed approach for four major watersheds

**Table 5.** A comparison through statistical assessment between design flood and predicted discharge

Watershed	Han river	Nakdong river	Yeongsan river	Geum river
R	0.953	0.955	0.953	0.912
IoA	0.971	0.963	0.971	0.950
RMSE	1.045	0.835	1.045	0.865
Median (Obs.)	21.206	12.765	21.206	13.563
Median (Pre.)	20.830	12.818	20.830	13.301
SD (Obs.)	3.322	2.518	3.322	2.025
SD (Pre.)	3.108	1.986	3.108	2.021
Cv (Obs.)	0.157	0.197	0.157	0.149
Cv (Pre.)	0.149	0.155	0.149	0.152

나타낸다. 일치계수(Index of Agreement, IoA)는 평균과 분산사이의 차이를 추정할 때 상관계수의 근거한 통계치의 민감성을 보완하고자 Willmott (1981)에 의해 개발되었으며, 0부터 1의 범위를 가지며 1에 가까울수록 모형의 예측능력이 우수함을 뜻한다. 앞서 언급하였듯이, 전체적인 검정결과를 통계적으로 분석하여 비교해보면, 미계측 유역 홍수량 지역화 모형으로부터 추정된 홍수량은 기수립된 설계홍수량과 높은 상관성을 보이고 있으며, 상관계수, 일치계수, 평균제곱근오차 등의 통계적 지표로부터 우수한 모의능력을 나타내는 것을 확인할 수 있다. 이밖에도 모형의 편의 및 기본 통계적 특성치를 모의하는데 있어 본 연구에서 제안한 미계측 유역 홍수량 산정모형이 상당히 유사한 결과를 보이고 있다. 이는 본 연구에서 고려된 비홍수량과 신뢰성 있는 매개변수 추정으로 인한 결과로 판단된다. 이와 더불어 해석결과에 대한 불확실성 정보를 정량적으로 제공한다는 점에서 기존 방법론과 차별성을 갖는다 할 수 있다.

#### 4. 결론

본 연구에서는 유역면적과 유역경사 등 유역의 전반적인 특성을 나타내는 자료를 활용하여 권역별 미계측 유역에서 홍수량 추정이 가능한 모형을 제시하였다. 각 권역별 기수립된 설계홍수량 자료의 지역화를 통해 미계측 유역 홍수량 산정모형의 매개변수를 추정하였으며, Bayesian 기법을 도입하여 모형의 매개변수와 산정되는 결과에 대한 불확실성을 정량적으로 평가하였다. 최종적으로 개발된 모형을 이용하여 하천 기본계획 수립 시 산정된 설계홍수량과 비교 분석을 수행하였다. 본 연구는 전체적인 내용은 요약하면 다음과 같다.

- 1) 본 연구에서는 유역면적 1,000 km<sup>2</sup> 이하의 기수립된 설계 홍수량 자료를 활용하여 재현기간에 따른 유역면적-비홍수량 사이의 상관관계를 멱함수 형태의 회귀식을 이용하여 자료 전체의 전반적인 특성을 파악하였다.
- 2) 권역별 설계홍수량 자료의 지역화 분석을 수행하는데 있어 Bayesian GLM 기법의 도입으로부터 비홍수량(유역면적, 유역경사)와 재현기간을 복합적으로 고려한 해석이 가능하였으며, 모형에서 추정되는 사후분포에 대한 불확실성을 정량적으로 평가하였다.
- 3) 권역별 홍수량 지역화 해석 모형을 통해 추정된 결과를 다양한 통계적 지표를 통해 검증한 결과 기존 선형회귀분석 기법에 비해 우수한 모형적합 능력을 나타내었다. 이와 더불어 추정된 매개변수 모두 불확실성 구간이 상대적으로 크지 않았으며, 사용된 독립변수에 대한 적합성 또한 확인할 수 있었다.
- 4) 유역의 일부를 미계측 유역으로 가정하여 교차검증 관점에서 제안된 모형의 적합성을 평가하였다. 본 연구에서 제안된 모형의 결과가 기존의 설계홍수량 값과 0.9 이상의 상관성을 나타내는 것을 확인할 수 있었으며, 다양한 통계적 지표를 통해서도 모형의 적합성을 확인하였다.

본 연구에서 제안된 미계측 유역 홍수량 산정모형은 매개변수에서 발생하는 불확실성의 정량화를 통해 해석결과에 대한 신뢰성을 제고하였으며, 권역별 자료를 더욱 확충한다면 우리나라 미계측 유역의 홍수량을 추정하는 도구로서 활용성이 높을 것으로 기대된다.

## 감사의 글

본 연구는 국토교통부/국토교통과학기술진흥원의 지원으로 수행되었음(과제번호 18AWMP-B083066-05). 저자들은 전북대학교 방재연구센터에 소속되어 연구를 수행하였습니다.

## References

- Brath, A., Castellarin, A., Franchini, M., and Galeati, G. (2001). "Estimating the index flood using indirect methods." *Hydrological Sciences Journal*, Vol. 46, No. 3, pp. 399-418.
- 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.
- Chae, B. S., Choi, S. J., Ahn, J. H., and Kim, T. W. (2018). "Estimation of flood quantile in ungauged watersheds for flood damage analysis based on flood index of natural flow." *Journal of the Korean Society of Civil Engineers*, Vol. 38, No. 1, pp. 175-182 (in Korean).
- Choi, H., and Lee, S.-J. (2006). "Verification about threshold discharge computation using GIUH on ungauged small basin." *Journal of Korea Spatial Information Society*, Vol. 14, No. 1, pp. 15-27.
- Creager, W. P., Justin, J. D. W., and Hinds, J. (1945). *Engineering for dams*. Vol. 1, General Design, John Wiley, New York.
- Dalrymple, T. (1960). *Flood-frequency analyses, manual of hydrology: Part 3 (No. 1543-A)*. USGPO.
- Fill, H. D., and Stedinger, J. R. (1998). "Using regional regression within index flood procedures and an empirical Bayesian estimator." *Journal of Hydrology*, Vol. 210, No. 1, pp. 128-145.
- Fleming, G. (1972). *Computer simulation techniques in hydrology*. In *Computer Simulation Techniques in Hydrology*, Elsevier.
- 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. PAMI-6, No. 6 pp. 721-741.
- Jakob, M., and Jordan, P. (2001). "Design flood estimates in mountain streams the need for a geomorphic approach." *Canadian Journal of Civil Engineering*, Vol. 28, No. 3, pp. 425-439.
- Kim, E. S., Choi, H. I., and Jee, H. K. (2010). "Estimation of the flash flood index by the probable rainfall data for ungauged catchments." *Journal of the Korean Society of Hazard Mitigation*, Vol. 10, No. 4, pp. 81-88.
- Kim, E. S., Jang, H. Y., Kim, S. D., Park, M. J., and Choi, H. I. (2012). "Analysis of rainfall-flood severity relations in small ungauged catchments." *Journal of the Korean Society of Hazard Mitigation*, Vol. 12, No. 2, pp. 199-208.
- Kim, J.-Y., Kim, J.-G., Cho, Y.-H., and Kwon, H.-H. (2017). "A development of Bayesian Copula model for a bivariate drought frequency analysis." *Journal of Korea Water Resources Association*, Vol. 50, No. 11, pp. 745-758.
- Kim, M.-H., Ceon, I.-K., and Chung, M. (2016b). "The regression model's independent variables importance decision for the design flood estimation." *Journal of Korean Society of Environmental Technology*, Vol. 17, No. 4, pp. 378-391.
- Kim, N. W., and Won, Y. S. (2004). "Estimates of regional flood frequency in Korea." *Journal of Korea Water Resources Association*, Vol. 37, No. 12, pp. 1019-1032.
- Kim, N. W., Lee, J. E., Lee, J., and Jung, Y. (2016a). "Regional frequency analysis using spatial data extension method: II. Flood frequency inference for ungauged watersheds." *Journal of Korea Water Resources Association*, Vol. 49, No. 5, pp. 451-458.
- Kwon, H. H., Brown, C., and Lall, U. (2008). "Climate informed flood frequency analysis and prediction in Montana using hierarchical Bayesian modeling." *Geophysical Research Letters*, Vol. 35, No. 5, L05404.
- Kwon, H.-H., Kim, J.-G., and Park, S.-H. (2013b). "Derivation of flood frequency curve with uncertainty of rainfall and rainfall-runoff model." *Journal of Korea Water Resource Association*, Vol. 46, No. 1, pp. 59-71.
- Kwon, H.-H., Kim, J.-Y., Kim, O.-K., and Lee, J.-J. (2013a). "A development of regional frequency model based on hierarchical Bayesian model." *Journal of Korea Water Resources Association*, Vol. 46, No. 1, pp. 13-24.
- Kwon, H.-H., Park, D.-H., and Moon, Y. I. (2004). "Derivation of flood frequency curve using uncertainty analysis of single event rainfall-runoff model-uncertainty analysis of rainfall-runoff model." *Journal of The Korean Society of Civil Engineers*, Vol. 24, No. 3B, pp. 229-239 (in Korean).
- Lee, J., Park J., Song, Y., Jun H., and Lee, J. (2017). "Estimation of design flood discharge by areal ratio for ungauged basin." *Journal of Wetlands Research*, Vol. 19, No. 3, pp. 335-344.
- 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 (in Korean).
- Malamud, B. D., and Turcotte, D. L. (2006). "The applicability of power-law frequency statistics to floods." *Journal of Hydrology*, Vol. 322, No. 1-4, pp. 168-180.

- Ministry of Environment (2006). *Hydraulic/hydrological investigation and development of analytic method in Yeongsan and Seomjin river basins*.
- Ministry of Land, Infrastructure and Transport (2016). *Korea annual hydrological report*.
- Onishi, R., Kato, T., Imaizumi, M., and Tanji, H. (2004). "A study to determine an equation of enveloping curve for regional flood peaks by maximum rainfall data." *Paddy and Water Environment*, Vol. 2, No. 2, pp. 83-90.
- Riggs, H. C. (1973). *Regional analysis of stream flow characteristics*. US Government Printing Office.
- Rogger, M., Kohl, B., Pirkl, H., Viglione, A., Komma, J., Kimbauer, R., Merz, R., and Blöschl, G. (2012). "Runoff models and flood frequency statistics for design flood estimation in Austria - Do they tell a consistent story?" *Journal of Hydrology*, Vol. 456-457, pp. 30-43.
- So, B.-J., Kim, J.-Y., Kwon, H.-H., and Lima, C. H. (2017). "Stochastic extreme downscaling model for an assessment of changes in rainfall intensity-duration-frequency curves over South Korea using multiple regional climate models." *Journal of Hydrology*, Vol. 553, pp. 321-337.
- Stedinger, J. R., Vogel, R. M., and Foufoula-Georgiou, E. (1993). "Frequency analysis of extreme events." in *Handbook of Hydrology*.
- Tofiq, F. A., and Guven, A. (2014). "Prediction of design flood discharge by statistical downscaling and general circulation models." *Journal of Hydrology*, Vol. 517, pp. 1145-1153.
- Willmott, C. J. (1981). "On the validation of models." *Physical Geography*, Vol. 2, No. 2, pp. 184-194.
- Yi, J., and Lee, T. (2013). "Development of design flood estimation system for mountainous river basin using fuzzy regression method." *Journal of Korean Society of Hazard Mitigation*, Vol. 13, No. 6, pp. 343-350.
- Yoon, Y. N. (2007). *Hydrology-Basics and application*. Cheong-MoonGak.