



## Development of a hybrid regionalization model for estimation of hydrological model parameters for ungauged watersheds

Kim, Youngil<sup>a</sup> · Seo, Seung Beom<sup>b</sup> · Kim, Young-Oh<sup>c\*</sup>

<sup>a</sup>Institute of Health and Environment, Seoul National University

<sup>b</sup>Institute of Engineering Research, Seoul National University

<sup>c</sup>Department of Civil and Environmental Engineering, Seoul National University

Paper number: 18-026

Received: 20 March 2018; Revised: 4 June 2018; Accepted: 4 June 2018

### Abstract

There remain numerous ungauged watersheds in Korea owing to limited spatial and temporal streamflow data with which to estimate hydrological model parameters. To deal with this problem, various regionalization approaches have been proposed over the last several decades. However, the results of the regionalization models differ according to climatic conditions and regional physical characteristics, and the results of the regionalization models in previous studies are generally inconclusive. Thus, to improve the performance of the regionalization methods, this study attaches hydrological model parameters obtained using a spatial proximity model to the explanatory variables of a regional regression model and defines it as a hybrid regionalization model (hybrid model). The performance results of the hybrid model are compared with those of existing methods for 37 test watersheds in South Korea. The GR4J model parameters in the gauged watersheds are estimated using a shuffled complex evolution algorithm. The variation inflation factor is used to consider the multicollinearity of watershed characteristics, and then stepwise regression is performed to select the optimum explanatory variables for the regression model. Analysis of the results reveals that the highest modeling accuracy is achieved using the hybrid model on RMSE overall the test watersheds. Consequently, it can be concluded that the hybrid model can be used as an alternative approach for modeling ungauged watersheds.

**Keywords:** Ungauged watershed, Regional regression model, Spatial proximity model, Hybrid regionalization model, GR4J

## 미계측유역의 수문모형 매개변수 추정을 위한 하이브리드 지역화모형의 개발

김영일<sup>a</sup> · 서승범<sup>b</sup> · 김영오<sup>c\*</sup>

<sup>a</sup>서울대학교 보건대학원, <sup>b</sup>서울대학교 공학연구원, <sup>c</sup>서울대학교 공과대학 건설환경공학부

### 요 지

수문모형의 매개변수 추정에 필요한 유량 관측 자료의 수집은 시-공간적으로 제한이 있어 우리나라도 아직 상당수의 미계측유역이 존재하며, 이를 보완하고자 주변 유역의 정보를 활용하는 지역화 방법들이 연구되어 왔다. 그러나 지역적 특성이나 기후 조건에 따라 지역화 방법의 결과가 상이하 여 어느 지역에 어떠한 지역화 방법이 가장 우수하다고 판단하기 어렵다. 본 연구에서는 보편적으로 사용되는 지역화 방법인 지역회귀모형의 설명 변수에 공간근접모형으로 추정된 수문모형의 매개변수를 추가하여 회귀모형의 적합성을 향상시켰으며, 이를 하이브리드 지역화모형이라 정의하 고 기존 방법들과 비교하였다. 계측유역으로는 관측 자료가 충분한 남한의 37개 유역을 선정하였고, 수문모형은 개념적 수문모형인 GR4J를 사용 하였으며, 계측유역에 대한 수문모형의 매개변수 산정은 Shuffled complex evolution 알고리즘을 사용하였다. 유역 특성변수들 간 다중공선성 을 고려하기 위해 Variation inflation factor를 사용하였고, Stepwise regression을 통해 회귀모형의 최적 설명변수를 선택하였다. 통계 값을 통 해 모형의 적합성을 비교한 결과, 하이브리드 지역화모형에서 가장 작은 RMSE 값을 나타내었으며, 유역별 모의 값의 변동성이 줄어들어 결과의 불확실성 또한 낮아짐을 확인할 수 있었다. 따라서 하이브리드 모형이 미계측유역의 유출량 산정을 위한 하나의 대안이 될 수 있음을 확인하였다.

**핵심용어:** 미계측유역, 지역회귀모형, 공간근접모형, 하이브리드 지역화모형, GR4J

\*Corresponding Author. Tel: +82-2-880-8916

E-mail: yokim05@snu.ac.kr (Y.-O. Kim)

## 1. 서론

수문모형으로 유역의 유출량을 산정하기 위해서는 과거 관측 유량 자료를 기반으로 검보정을 진행하여야 하나, 관측 유량 자료가 부족하거나 검증되지 않은 유역의 경우 위 과정을 진행하기 어렵다(Oudin et al., 2008). 이러한 유역을 미계측유역(ungauged watersheds)이라 하며(Sivapalan et al., 2003), 미계측유역의 유출량 산정을 위해 유역을 분류하거나 그룹화를 하면서부터 지역화(Regionalization)라는 용어를 사용하기 시작하였고(Gottschalk et al., 1979), 주변 계측유역(donor catchments)의 매개변수의 추정 결과를 미계측유역의 유출량 산정에 전이하면서 지역화 용어의 정의가 확대되었다(Oudin et al., 2008). 이러한 지역화 방법은 일반적으로 두 개의 목적으로 사용된다. 하나는 계측유역의 관측유량과 유역 특성변수들 사이 관계식을 통해 미계측유역의 유출통계량을 산정하는 방법으로, 주로 다중선형회귀분석(Vogel et al., 1999; Goswami et al., 2007)이 사용되어져왔다. 이 방법은 유역 특성 자료만으로 관계식을 만들기 때문에 구조적으로 단순한 장점이 있으나, 통계량밖에 산정할 수 없다는 단점이 있다. 다른 하나는 계측유역을 대상으로 수문모형의 매개변수와 유역 특성변수들 사이 관계식을 구한 후 미계측유역의 수문모형의 매개변수를 추정하는 방법으로 주로 회귀모형(regression model), 물리 유사모형(physical similarity model), 공간근접모형(spatial proximity model) 등이 사용된다. 이 방법은 과정상에 수문모형이 사용됨으로서 불확실성이 추가되는 단점이 있으나, 추정된 수문모형의 매개변수를 통해 실시간 유량 모의가 가능한 장점이 있다. 따라서 연속적인 모의가 중요한 미계측유역 연구(Razavi and Coulibaly, 2012)에 적용하기 위한 방법으로 앞선 유출통계량 방법보다 적합한 것으로 판단하여 본 연구에서는 수문모형 매개변수 추정 방법을 선택하였다.

후자의 방법에 주로 사용되는 세 가지 모형의 장단점을 살펴보면 다음과 같다. 회귀모형은 지역화 연구에 가장 많이 사용되는 방법으로서(Young, 2006) 계측유역 수문모형의 매개변수 추정 값과 유역 특성변수들 사이 관계식을 통해 미계측유역의 유출량을 쉽게 모의할 수 있으나, 충분한 양의 자료가 필요하고 유역 특성변수들만으로 최적의 매개변수 구성을 만들기 어려우며 계측유역의 수문모형 매개변수 추정시 검보정 기간이나 입력자료의 오차에 의해 모의 결과의 변동성이 커지는 단점이 있다(Yapo et al., 1996; Perrin et al., 2007). 공간근접모형은 주변 계측유역의 유량이나 수문모형의 매개변수 정보를 가져오는 방법으로(Vandewiele et al., 1991) 추정 값을 평균하여 쉽게 사용할 수 있으나, 유역 특성변수가 전혀

고려되지 않고 추정되는 미계측유역과 주변 계측유역이 동질하다(Homogeneous)는 가정이 현실에서는 만족되기 어렵다는 단점이 있다. 물리 유사모형의 경우 유역의 물리적 특성이 비슷할 경우 수문학적 특성 또한 비슷할 것이라는 가정 하에 유역 특성변수들을 기반으로 가장 유사한 계측유역의 유출량 또는 수문모형의 매개변수 추정 값을 미계측유역에 적용하는 방법으로, 다른 방법들 보다 유역의 물리적 특성을 가장 유사하게 반영하는 장점이 있으나, 공간근접모형과 마찬가지로 기본 가정이 맞지 않을 경우 결과를 신뢰하기 어렵다.

이렇게 각 모형마다 장단점이 달라 과거 연구 사례에서는 기존 지역화 모형을 개선시키거나 여러 지역화 방법들을 비교한 후 가장 우수한 하나의 모형을 제안하여 왔다. Kim and Lee (2008)은 다중회귀모형에 bayesian 방법을 적용하여 지역 빈도분석의 불확실성을 감소시킬 수 있음을 논하였고, Moon et al. (2013)은 Hargreaves 공식의 매개변수 지역화를 위해 수정된 식을 제안하여 기존보다 개선된 기준증발산량 산정 결과를 얻었다. Kim et al. (2015)는 수문모형인 Sacramento에 Bayesian 기법을 연계하여 안정된 매개변수 추정이 가능하도록 하였다. Young (2006)은 미국의 260개 유역을 대상으로 다중회귀모형과 공간근접모형을 비교한 결과 다중회귀모형이 더 좋은 결과를 나타냄을 논하였다. Kay et al. (2006)은 물리 유사모형과 회귀모형을 수문모형인 PDM과 TATE를 사용하여 비교한 결과 PDM 모형 경우 물리 유사모형이 우수하였고 TATE 모형의 경우 회귀모형이 우수함을 보여 사용되는 수문모형에 따른 최적 지역화모형이 다를 수 있음을 논하였다. Oudin et al. (2008)은 공간근접모형, 물리 유사모형, 회귀모형 방법을 모두 비교한 결과 공간근접모형이 가장 높은 적합성을 가짐을 보였다. Jin et al. (2009)는 중국의 13개 유역에 HBV 모형을 이용하여 물리 유사모형과 평균값을 비교한 결과 오히려 평균이 더 우수한 결과를 보여줌을 논하였고, Samuel et al. (2011)은 공간근접모형, 물리 유사모형, 회귀모형을 모두 비교한 결과 공간근접모형과 물리 유사모형을 사용한 결과가 회귀모형 보다 더 좋은 값을 얻음을 보여주었다. Pugliese et al. (2016)의 경우 지구통계학적 기법(geostatistical method)과 다중회귀모형을 비교한 결과 다중회귀모형의 적합성이 높게 나왔으며 Lebecherel et al. (2016)은 공간근접모형을 기반으로 최적 적용 유역(donor catchments)의 개수를 분석하였다.

이처럼 과거 연구들에서 지역화 방법마다 상이한 적합성 결과를 보이는데, 이를 종합해보면 다음과 같다. 첫째, 유역마다 고유한 물리적, 기후적 특성을 가지고 있어 미계측유역의 유출량 산정에 사용되는 계측유역의 수가 다르다. 둘째, 사용한 지역화 방법 및 수문모형이 다르다. 셋째, 연구대상 유역

및 저자에 따라 유역 특성변수 선정 기준이 다르다. 이와 같이, 사용되는 방법과 결과가 상이하므로 현재까지도 어느 지역에 어떠한 지역화 방법이 가장 우수하다 할 수 있는 가이드라인을 제시하기 어려운 실정이다(Stoll and Weiler, 2010; Samuel *et al.*, 2011).

따라서 본 연구에서는 관측 유량 자료를 보유한 남한의 37개 계측유역을 선정하여 유역 간 물리적, 기후적 특성을 반영하고, 기존에 많이 사용되어왔던 지역회귀모형과 공간근접모형을 모두 적용하여 하나의 지역화모형으로부터 생길 수 있는 오차를 줄였다. 또한, 앞선 두 개의 지역화모형의 조합을 통해 상관관계에 따라 유역 특성변수가 자동으로 고려되면서 공간근접모형의 장점을 가져오는 하이브리드 지역화모형을 제안하고자 한다. 이러한 조합 모형은 수문학의 다른 분야에서 종종 시도되어 왔는데, multi-model ensemble (Ajami *et al.*, 2006; Duan *et al.*, 2007), combined forecasts (Jeong and Kim, 2009) 등이 그 사례 중 하나이다.

## 2. 지역화 방법

### 2.1 지역회귀모형(Regional regression model)

지역회귀모형은 주로 유역 및 기후 특성변수와 수문모형의 매개변수 사이의 관계식을 기반으로 하는데, 각 유역별로 유역 특성변수들이 다양하기 때문에 하나의 반응변수에 두 개 이상의 설명변수가 선형으로 대응되는 다중선형회귀분석이 주로 사용된다. 반면, 반응변수인 수문모형 매개변수에는 계측유역을 대상으로 검보정한 최적 매개변수 값이 자료로 사용되는데, 회귀분석 시 불확실성을 줄이기 위해 모의 정확도가 높은 유역만을 사용한다. 본 연구에 사용된 회귀식은 다음 Eq. (1)과 같다.

$$Y_p = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \varepsilon \quad (1)$$

여기서,  $Y_p$ 는 종속변수로  $p$ 번째( $=1, \dots, P$ ) 수문모형의 매개변수,  $\beta_0$ 는 절편,  $\beta_k$ 는  $k$ 번째( $=1, \dots, K$ ) 설명변수의 계수,  $X_k$ 는 계측유역의  $k$ 번째 특성변수,  $\varepsilon$ 는 오차항(error term)을 의미한다.

유역 특성변수 선정의 경우 유출량과 상관관계가 높은 변수들을 선정하기 위해 다양한 방법이 제시되어 왔는데, Merz and Blöschl (2004)은 저수지를 반영하기 위해 lake index를 사용하였으며, Parajka *et al.* (2005)은 전문가들의 조언을 통해 기상, 물리적인 변수들을 선택하였다. Oudin *et al.* (2008)

은 물리적 유사성을 고려하기 위해 기존의 지형 및 기후 변수들에 이어 건조 index를 추가하였으며, Samuel *et al.* (2011)은 특성변수로 증발산량 등을 고려하기 위해 유역의 위도, 경도를 추가하여 사용하였다. 이처럼 연구의 목적에 따라 유역 특성변수를 적용해 왔으며, 본 연구에서는 다중공선성(multicollinearity)을 고려하여 유역의 물리적 성질과 기후 특성을 모두 반영한 유역 특성변수를 기반으로 설명변수를 구성하였다. 최종 회귀식은 stepwise regression 방법 중 AIC (Akaike Information Criterion)를 활용한 backward elimination을 적용하였다. AIC는 설명변수 선정에 따른 회귀모형의 적합성을 수치로 나타내며 수가 작을수록 유의미한 결과로 볼 수 있다.

### 2.2 공간근접모형(Spatial proximity model)

공간근접모형은 서로 거리가 가까운 유역일 경우 유역 특성이나 기후, 수문 특성이 비슷할 것임을 전제로 하여 미계측 유역의 주변 계측유역으로부터 정보를 가져오는 방법이다 (Oudin *et al.*, 2008). 본 연구에서는 보편적으로 사용되는 수문모형 매개변수 평균 방법을 사용하였고, 유역 간 중심점의 거리에 따른 계측유역의 개수별 모형 결과를 분석한 결과, 가장 가까운 두 개의 유역을 사용할 때 가장 높은 모의 정확성을 보였다. 따라서 본 연구에서는 두 개의 계측유역의 평균값을 적용하였다(Eq. (2)).

$$Z_p = \frac{1}{L} \sum_{i=1}^L y_p^i \quad (2)$$

여기서,  $Z_p$ 는 공간근접모형으로 추정된 수문모형의  $p$ 번째 매개변수,  $L$ 은 사용된 주변 계측유역의 총 수,  $y_p^i$ 은 공간근접모형에 사용된  $i$ 번째( $=1, \dots, L$ ) 계측유역의  $p$ 번째 수문모형 매개변수를 의미한다.

### 2.3 하이브리드 지역화모형

#### (Hybrid regionalization model)

과거 지역화 연구에서 가장 많이 사용된 두 가지 모형은 지역회귀와 공간근접으로, 미계측유역을 대상으로 가장 높은 적합성을 보여 왔다. 본 연구에서는 지역회귀모형과 공간근접모형의 조합을 통해 모형 적합성을 개선하는 방법을 제안하며 이를 하이브리드 지역화모형(이하 하이브리드 모형)이라 정의하였다. 이 방법의 목적은 공간근접모형을 통해 나온 수문모형의 매개변수 추정 값을 지역회귀식의 설명변수에 추가하면서 기존 지역회귀모형의 적합성을 향상시키는데 있다. 공간근접모형은 유역특성변수를 전혀 고려하지 않고 수문모

형 검보정을 통해 얻어진 주변 계측구역의 수문모형 매개변수 값들을 직접 사용하므로, 구역특성변수만 사용하는 지역회귀모형과 조합하여 두 장점을 모두 가지게 된다. 공간근접모형의 매개변수 추정 값이 추가된 회귀모형은 변수들 간 다중공선성을 고려하여 설명변수를 재설정하게 되고, 반응변수와의 상관관계 분석을 통해 낮은 상관관계를 보이는 변수들은 자동으로 탈락하게 된다. 따라서 공간근접모형의 매개변수 추정 값이 설명변수로 추가되더라도 기존 구역 특성변수들보다 상관관계가 낮을 경우 회귀모형에 사용되지 않으며, 반대로 상관관계가 더 높을 경우 설명변수로 남아 회귀모형의 적합성을 향상시키게 된다. 하이브리드 모형에 사용된 식은 Eq. (3)과 같다.

$$Y_p = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \beta_{k+1} Z_{p-i} + \dots + \beta_{k+1+j} Z_{p-i+j} + \varepsilon \quad (3)$$

여기서,  $\beta_{k+1+j}$ 는 공간근접모형으로 추가된 설명변수  $Z_{p-i+j}$ 의 계수( $i=0, \dots, I, j=1, \dots, J, p>i$ )를 의미한다.

### 3. 적용 사례

#### 3.1 대상구역

본 연구에서는 수자원공사에서 제공하는 댐 상류 구역 35개와 WAMIS (Water Resources Management Information System)로부터 얻은 2개의 구역(갑천, 미호천)을 합한 총 37개 계측구역에 지역화모형을 평가하였다(Fig. 1). 관측 유량 자료의 기간은 최장 1973~2016년이었으나 구역마다 차이가 있었으므로 검보정 기간의 통일성을 위해 2000~2015년 자료만 사용하였다. 각 구역의 유출 지점은 지형학적으로 126°33'~129°16' E와 34°49'~37°56' N 사이에 위치하고 있으며, 연구에 사용된 구역의 총 넓이는 26,831.4 km<sup>2</sup>, 연평균 총 강수량은 1265.5 mm, 구역 평균 경사는 23~56%, 평균 고도는 103.8~925.9 m 이다.

#### 3.2 수문모형

지역화 연구는 주로 개념적 수문모형이 사용되어 왔는데 (Post and Jakeman, 1999; Oudin *et al.*, 2008; Jin *et al.*, 2009; Viviroli and Seibert, 2015), 본 연구에서는 비교적 간단하면서도 불확실성이 적다고 알려진(Jakeman and Homberger, 1993; Seibert, 1999; Young, 2006; Yadav *et al.*, 2007) GR4J (Perrin *et al.*, 2003)를 사용하였다. 이 모형은 4개의 매개변수

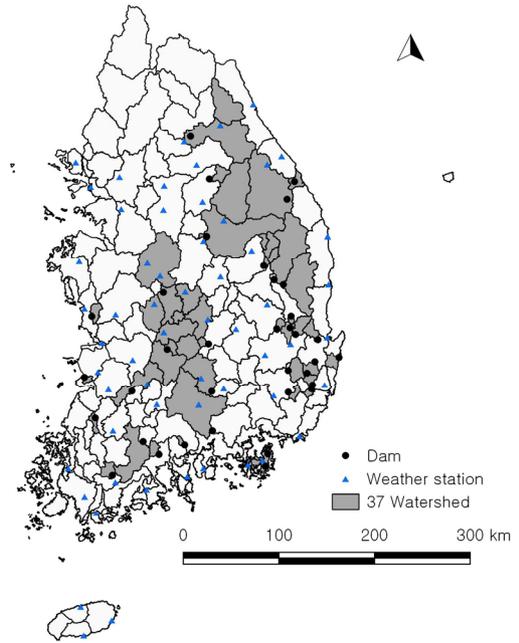


Fig. 1. Location of the 37 watersheds used in this study

Table 1. Descriptions of GR4J parameters (Perrin *et al.*, 2003)

Parameter	Description
$Y_1$	Maximum capacity of production store (mm)
$Y_2$	Groundwater exchange coefficient (mm)
$Y_3$	Maximum capacity of routing store (mm)
$Y_4$	Time peak ordinate of hydrograph unit UH1 (day)

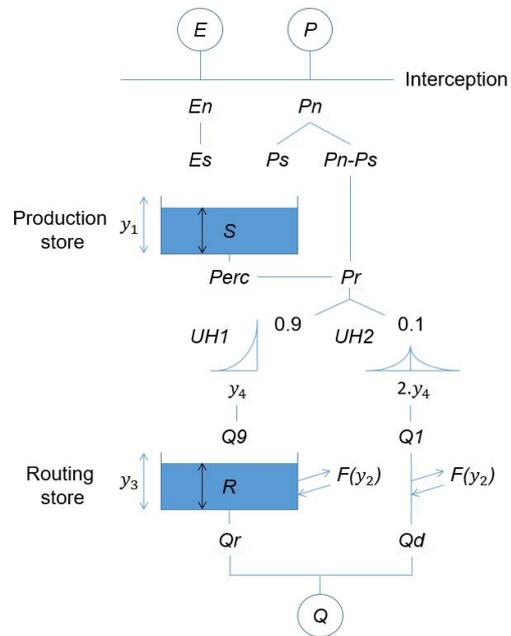


Fig. 2. Conceptual description of GR4J model (Perrin *et al.*, 2003)

를 가지고 있으며(Table 1), 두 개의 저류지(Storage) 모듈과 두 개의 단위 유량도(Unit hydrograph)를 가지고 있다(Fig. 2). 입력 자료인 강수(precipitation)와 증발산(evapotranspiration)은 첫 번째 저류지인 production store에서 침투(percolation) 등과 함께 고려된 후 유효 강수량이 된다. 두 번째 저류지인 Routing store에서는 단위 유량도를 통해 흘러들어온 유효 강수량 중 침투(infiltrate)되는 양을 고려하게 되는데, 여기서 유효 강수량의 90%는 하수(groundwater)로 들어가게 되며 나머지 10%는 지표면을 통해 유출되게 된다. 최종적으로 위 두 과정의 유량을 합친 것이 유역 유출량이 된다.

수문모형의 최적 매개변수 산정은 과거 연구 사례들을 통하여 검증된 SCE (Shuffled Complex Evolution) 알고리즘을 사용하였으며, 10년의 보정기간(2000~2009년)과 6년의 검정기간(2010~2015년)을 적용하였다. 위 과정은 오픈 소스 프로그램인 R (R statistical computing environment)(<https://www.r-project.org>)을 통해 진행되었으며, 프로그램 내 hydro-mad (Hydrological Model Assessment and Development) package (Andrews et al., 2011)를 사용하여 GR4J의 최적 매개변수를 얻었다.

### 3.3 모형 평가

계측유역을 대상으로 수문모형 매개변수 검보정 및 지역화 방법의 모의 정확성을 평가하기 위해 NSE (Nash-Sutcliffe Efficiency)와 PRMSE (Percent Root Mean Square Error), PBIAS (Percent BIAS)를 사용하였고 이는 각각 Eqs. (4)~(6)과 같다.

$$NSE = 1 - \frac{\sum_{i=1}^n (Q_i^{obs} - Q_i^{sim})^2}{\sum_{i=1}^n (Q_i^{obs} - \bar{Q}^{obs})^2} \quad (4)$$

$$PRMSE = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (Q_i^{sim} - Q_i^{obs})^2}}{\bar{Q}^{obs}} \quad (5)$$

$$PBIAS = 100 \frac{\sum_{i=1}^n (Q_i^{sim} - Q_i^{obs})}{\sum_{i=1}^n Q_i^{obs}} \quad (6)$$

여기서,  $Q_i^{obs}$ 는  $i$ 번째 관측값,  $Q_i^{sim}$ 는  $i$ 번째 모의값, 그리고  $\bar{Q}^{obs}$ 는 관측값의 평균을 말한다. NSE 값은 1에 가까울수록 높은 모의정확성을 가지며, PRMSE, PBIAS는 값이 0에 가까울수록 정확한 예측력을 의미한다.

## 4. 적용 결과

### 4.1 유역 특성변수의 선정

유역 특성변수는 WAMIS에서 제공하는 자료를 기반으로 하였고, 변수 간 다중공선성(multicollinearity)을 고려하기 위해 VIF (Variance Inflation Factor)를 사용하였는데, 이는 상관계수( $R$ )의 역수로 이루어진 Eq. (7)과 같다. 일반적으로 VIF가 10 이상의 값을 가질 경우 다중공선성이 존재한다고 볼 수 있으므로, 해당되는 변수들을 제외한 후 유역별 면적( $A$ ), 연평균 강수량( $P_r$ ), 연평균 잠재증발산량( $P_{et}$ ), 연평균 기온( $T$ ), 하천밀도( $D$ ), 평균고도( $H$ ), 최고고도( $H_{max}$ ), 토양도( $CN$ ), 침투율이 낮은 토양의 면적 비율( $c$ ) 등 총 9개의 대표 유역 특성변수를 선정하였다(Table 2).

$$VIF = \frac{1}{1 - R^2} \quad (7)$$

Table 2. Descriptors of watershed characteristics and summary statistics

Notation	Unit	Description	10th	Median	90th	Maximum VIF
$A$	km <sup>2</sup>	Area	12.1	134.1	2012.3	8.5
$P_r$	mm/yr	Annual average precipitation	1020.4	1245.0	1524.3	1.8
$P_{et}$	mm/yr	Annual average potential evapotranspiration	883.6	966.0	1023.3	1.9
$T$	°C	Annual average temperature	11.3	13.0	14.4	2.7
$D$	km/km <sup>2</sup>	Density of stream	0.9	2.2	3.6	3.0
$H$	m	Average altitude	172.7	359.6	570.2	1.5
$H_{max}$	m	Maximum altitude	589.5	1126.1	1584.2	1.6
$CN$	-	Curve number	51.0	65.0	76.0	5.0
$c$	-	Area ratio of soils with slow infiltration rates	12.6	70.6	365.0	6.4

4.2 수문모형 검보정

검보정 결과의 평균 NSE 값이 0.6 이상이면 모의 정확성이 높다고 판단하고(Young, 2006), 37개의 계측유역에 수문모형 검보정을 진행한 결과(Fig. 3), 31개 유역에서 0.6 이상의 NSE를 얻은 반면(Table 3(a)), 6개 유역(보현산댐, 감포댐, 선암댐, 장흥댐, 갑천, 미호천)에서 0.6 미만의 NSE를 얻었다 (Table 3(b) and Fig. 4). 최적화 알고리즘 수행 시 매개변수의 범위를 조정하며 여러 번 반복 수행하였음에도 일부 유역에서

낮은 모의 정확성을 가졌는데, 이는 다음 몇 가지 이유가 작용한 것으로 사료된다. 하나는 유역의 크기가 매우 작은 경우로 감포댐(3.7 km<sup>2</sup>), 선암댐(1.2 km<sup>2</sup>)이 해당되며, 이러한 유역은 관측 유량 또한 0에 가까운 값을 가져 수문모형을 통해 적절한 모의 값을 얻기 어려웠다. 다른 하나는 수위관측소 자료를 사용한 경우로 갑천, 미호천 유역이 해당되며, 수위관측소는 수위-유량 관계곡선을 통해 변환된 유량 자료를 제공하므로 관측 자료의 불확실성이 존재하여 낮은 모의 정확성을 보인 것으로 판단된다. 따라서 계측유역의 추정된 수문모형 매개변수를 사용하여 미계측유역의 수문모형 매개변수를 추정하는 지역화모형의 불확실성을 줄이기 위해, 본 연구에서는 총 37개 계측유역에 대한 검보정 결과를 바탕으로, 정확도가 높은 31개 유역을 계측유역으로 사용하여 지역화모형을 구축하고 상대적으로 정확도가 낮은 6개 유역을 미계측유역이라 가정하고 지역화모형의 적합성 평가를 진행하였다.

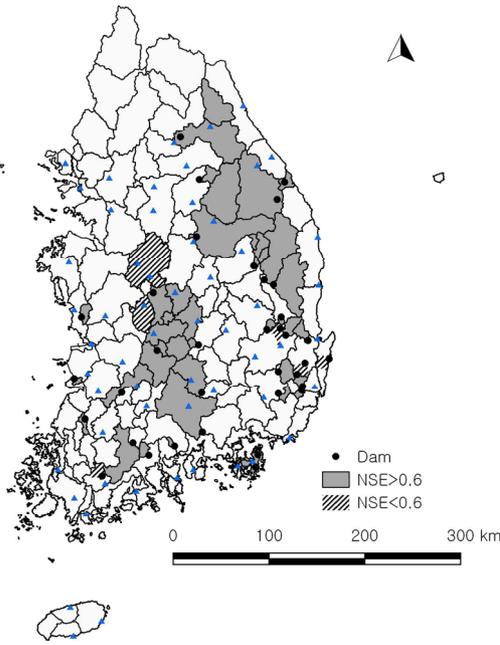


Fig. 3. The Results of calibration of GR4J model with an average NSE

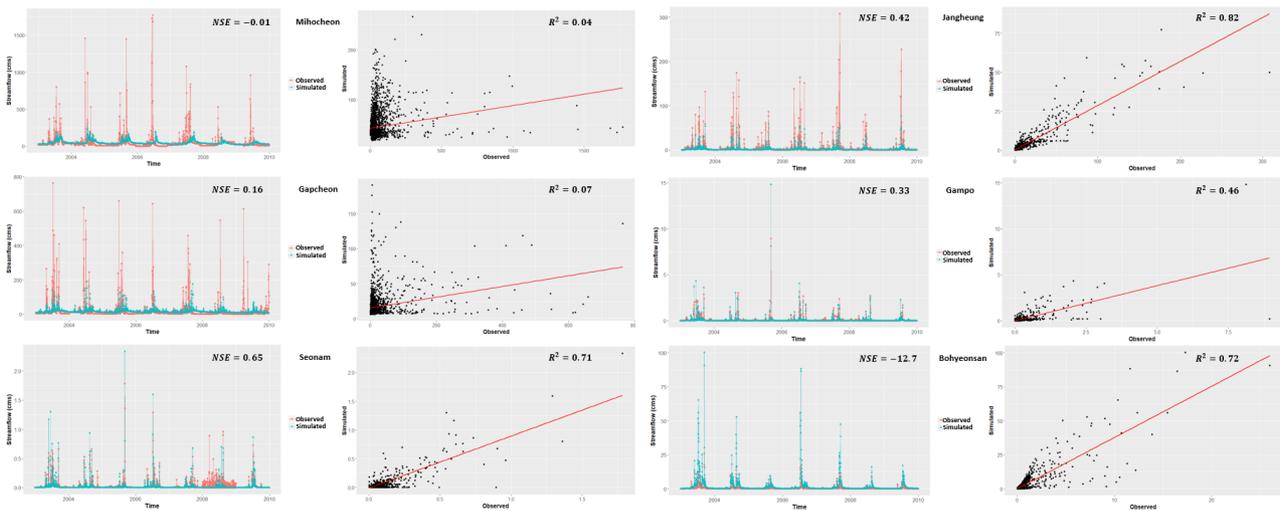
Table 3. Performance of GR4J model over the 37 watersheds

(a) 31 watersheds

NSE	Min	10th	Median	Mean	90th	Max
Calibration	0.64	0.68	0.78	0.76	0.83	0.88
Validation	0.51	0.63	0.74	0.74	0.86	0.89
Average	0.60	0.65	0.76	0.75	0.84	0.88

(b) 6 watersheds

NSE	Min	10th	Median	Mean	90th	Max
Calibration	-12.7	-6.36	0.25	-1.86	0.54	0.65
Validation	-14.8	-7.68	-0.20	-2.44	0.56	0.68
Average	-13.75	-6.98	0.06	-2.15	0.47	0.51



(a) Calibration

Fig. 4. The calibration results of 6 gauged watersheds with an average NSE of less than 0.6

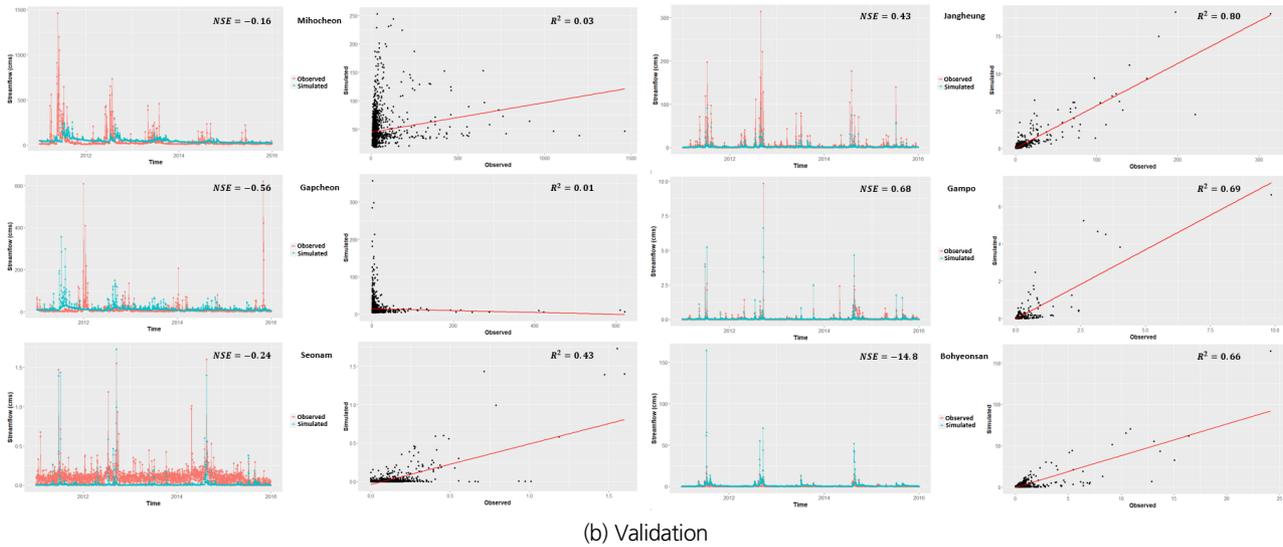


Fig. 4. The calibration results of 6 gauged watersheds with an average NSE of less than 0.6 (Continue)

4.3 지역화 방법의 모의 정확성

지역회귀모형과 하이브리드 모형 구축 결과 Table 4와 같은 회귀식을 얻었고, 조정결정계수(Adj R<sup>2</sup>)를 사용하여 수문 모형 매개변수 적합성의 향상 정도를 비교한 결과 하이브리드 모형이 최소 7%, 최대 91% 향상됨을 보였다(Table 5). 이는 기존 설명변수보다 상관성이 높은 공간근접모형의 매개변수 추정 값(Z<sub>2</sub>, Z<sub>4</sub>)이 추가되면서 전체 회귀모형의 적합성을 향상시킨 것으로 사료된다. 따라서 공간근접모형의 매개변수 추정 값은 변수 간 다중공선성 및 종속변수와의 상관관계에 따라 회귀모형에 추가되는 것을 알 수 있다.

위 결과를 기반으로 지역화모형의 적합성을 비교하면 다음과 같다. PRMSE는 지역화모형 중 하이브리드 모형과 공간근접모형이 최솟값(Min)에서 1.43으로 가장 낮았고, 중간 값

(Median)은 11.00으로 공간근접모형, 평균(Mean)과 최댓값(Max)에서는 각 16.29, 34.66으로 하이브리드 모형이 가장 우수하였다(Table 6). 검토정 결과는 최솟값과 중간 값에서 오차가 적은 반면, 평균과 최댓값에서 가장 높았는데, 이는 4개의 유역에서 우수한 모의 능력을 보였으나, 갑천 유역에서 88 이상의 높은 값을 가졌기 때문이다. 반면, 지역화모형은 모두 50 미만의 값을 가졌고, 그 중 하이브리드 모형이 가장 우수하였다. 절댓값으로 비교한 PBIAS는 최솟값을 제외하고 지역회귀모형이 가장 우수하였으나, 검토정 결과가 매우 낮은 6개 유역을 대상으로 지역화모형을 적용하였기에 전반적으로 오차가 높게 나타났다. 결과적으로, 하이브리드 모형이 기존 지역화모형과 비슷하거나 우수한 적합성을 보였으므로 미계측유역의 유출량 산정시 하나의 대안이 될 수 있을 것으로 판단된다.

Table 4. The equations of two regionalization models (regression, hybrid)

(a) Regional regression model

Parameter	Equation	R <sup>2</sup>	Adj R <sup>2</sup>
Y <sub>1</sub>	- 50.768 + 0.052A - 0.083P + 0.239P <sub>et</sub> - 6.792D + 0.076H <sub>max</sub> - 0.089c	0.67	0.58
Y <sub>2</sub>	- 5.455 - 0.005H <sub>max</sub>	0.26	0.23
Y <sub>3</sub>	19.942 + 0.039A + 0.055P - 0.073c	0.46	0.40
Y <sub>4</sub>	0.515 + 0.0001A + 0.042D + 0.001H <sub>max</sub>	0.63	0.59

(b) Hybrid model

Parameter	Equation	R <sup>2</sup>	Adj R <sup>2</sup>
Y <sub>1</sub>	668.934 + 0.023A - 0.292P - 0.114H - 8.668D - 2.749CN + 0.132H <sub>max</sub> + 1.594Z <sub>2</sub> - 1.380Z <sub>4</sub>	0.82	0.74
Y <sub>2</sub>	19.161 - 0.006P - 0.396D - 0.160CN + 0.005H <sub>max</sub> - 0.039Z <sub>4</sub>	0.53	0.44
Y <sub>3</sub>	203.343 + 0.014A - 1.916CN + 0.037H <sub>max</sub> + 0.580Z <sub>2</sub> - 0.530Z <sub>4</sub>	0.60	0.50
Y <sub>4</sub>	1.404 + 0.0001A - 0.051T - 0.0004H + 0.045D + 0.0004H <sub>max</sub> - 0.002Z <sub>2</sub>	0.71	0.63

하이브리드 모형이 일부 유역에서는 기존 지역화모형보다 낮은 적합성을 나타내었는데, 구간 별로 비교해보면 하이브리드 모형이 조합된 두 지역화모형의 영향을 받아 각 모형 사이에서 모의하는 것을 알 수 있다(Fig. 5). *PRMSE*는 최솟값과 중간값에서 공간근접모형의 영향으로 오차가 줄고, 평균(Mean)

과 최댓값에서는 지역회귀모형의 영향으로 오차가 낮아진 반면, *PBIAS*에서는 공간근접모형으로 인해 오히려 지역회귀 모형보다 높은 오차를 가졌다. 따라서 하이브리드 모형이 조합되는 과정에서 회귀모형의 적합성을 낮추는 어떤 요인이 작용한 것으로 보이며, 이는 몇 가지 근거와 함께 결론에서 논하였다.

**Table 5.** Comparison between the regression and the hybrid models for the watersheds to be ungauged using the adjusted  $R^2$

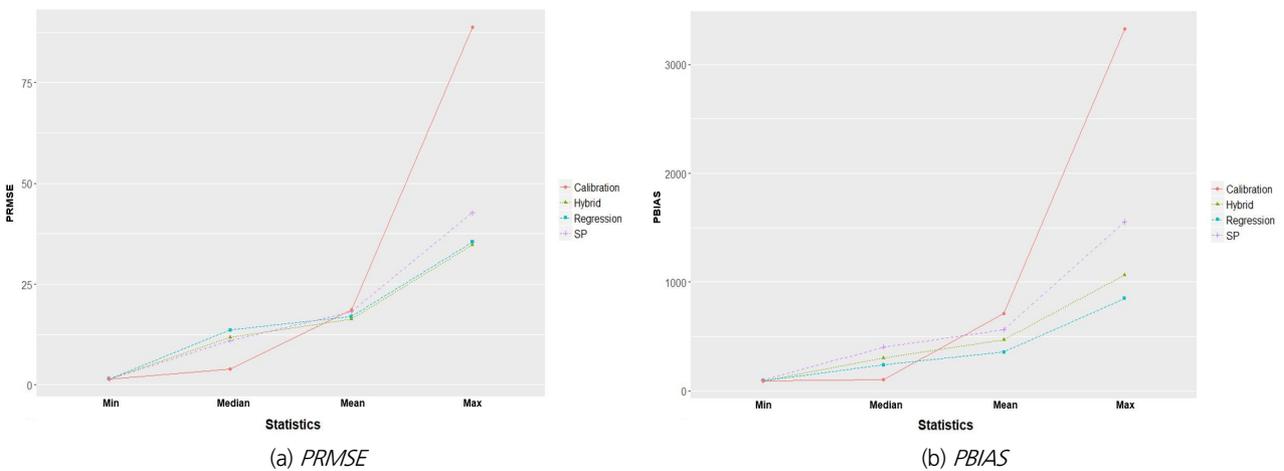
Parameter	From regression to hybrid (%)	
	Adj $R^2$	
$Y_1$	28	
$Y_2$	91	
$Y_3$	25	
$Y_4$	7	

### 5. 요약 및 결론

본 연구의 목적은 미계측유역의 유출량 산정을 위해 새로운 지역화모형인 하이브리드 모형을 개발함에 있다. 과거 지역화 연구에서 가장 널리 사용되어온 지역화 방법인 지역회귀 모형과 공간근접모형을 조합한 하이브리드 모형은 변수 간

**Table 6.** Performance of three regionalization(regression, SP, and hybrid) models for the 6 watershed to be ungauged ('calibration' indicates the results using observed data)

	Watershed	Regression			Hybrid			SP			Calibration		
		<i>PRMSE</i>	<i>PBIAS</i>	$R^2$									
Watershed	Mihocheon	22.88	375.6	0.01	19.33	509.30	0.01	17.76	707.70	0.01	11.57	557.30	0.01
	Gapcheon	35.41	850	0.01	34.43	968.50	0.01	42.68	1553.10	0.01	88.80	3327.00	0.01
	Seonam	1.43	-99.90	0.31	1.43	-99.90	0.30	1.43	-99.80	0.31	1.43	-99.80	0.42
	Jangheung	3.64	-93.70	0.14	3.65	-94.10	0.17	3.66	-95.60	0.14	3.62	-92.90	0.81
	Gampo	4.25	625	0.45	4.25	-97.80	0.55	4.24	-97.70	0.50	4.24	-96.90	0.71
	Bohyeonsan	33.86	509.30	0.67	34.66	1062.30	0.61	38.64	815.60	0.71	2.64	112.60	0.55
Statistics	Min	2.54	96.05	0.01	1.43	94.10	0.01	1.43	95.60	0.01	1.43	92.90	0.01
	Median	13.56	237.75	0.22	11.79	304.60	0.23	11.00	403.75	0.23	3.93	106.20	0.48
	Mean	16.29	357.10	0.26	16.29	471.98	0.27	18.07	561.58	0.28	18.72	714.42	0.41
	Max	35.41	357.10	0.67	34.66	1062.30	0.61	42.68	1553.10	0.71	88.80	3327.10	0.81



**Fig. 5.** Performance of three regionalization models (regression, hybrid, SP)

상관관계를 높이기 어려운 지역회귀분석의 한계와 유역의 물리적 특성 반영이 어려운 공간근접모형의 단점을 보완하여 유출량 모의 적합성을 향상시킬 수 있음을 보여주었다. 대상 유역은 관측자료가 충분히 존재하는 남한의 37개의 유역을 선정하였고, 단계별 회귀분석을 진행하여 최종 회귀식을 구성하였다. 수문모형 검보정 결과 *NSE* 값이 0.6 이상으로 높은 모의 적합성을 보인 31개 유역을 대상으로 지역화모형을 구축하였고, 나머지 6개 유역에 적용성을 평가하였다. 지역화모형 간 적합성을 비교한 결과 *PRMSE*에서 하이브리드모형이 가장 우수하였고, *PBIAS*에서는 평균적으로 지역회귀모형이 우수하였다.

결과적으로, 하이브리드 모형이 기존 지역화모형과 비슷하거나 우수한 적합성을 보여 미계측유역의 유출량 산정에 하나의 대안이 될 수 있을 것으로 사료된다. 그러나 일부 결과에서는 기존 지역화모형이 더 우수하였는데, 조합된 하이브리드 모형의 적합성이 더 낮게 나온 요인으로 다음 두 가지를 생각해 볼 수 있다. 첫째는 공간근접모형의 추정된 매개변수와 기존 유역특성변수 간 다중공선성으로 인해 종속변수와 설명력이 높았던 유역 특성변수가 제거되어 회귀모형의 설명력이 떨어진 경우이고, 둘째는 추가 또는 삭제된 설명변수가 큰지레점(high leverage points)인 경우이다.

앞선 경우는 종속변수와 상관성이 높은 유역 특성변수가 공간근접모형과 조합되면서 회귀모형에 사용되기 전 제거된 경우로, 이를 보완하기 위해서는 설명력이 높은 유역 특성변수를 우선적으로 선택하는 제한요건이 추가로 필요할 것으로 판단된다. 후자는 변수 선정과정에서 설명변수의 이상치(큰 지레점)가 추가되거나 삭제된 경우로, 적은 양의 변수로도 회귀모형의 적합성이 크게 바뀌게 될 수 있으므로, 설명변수 선정 전 자료의 이상치 적용 여부를 고려해야 할 것이다. 본 연구에서는 유역 특성변수의 설명력을 고려하지 않고 다중공선성이 높은 순서대로 변수를 제거하고, 자료의 이상치를 모두 포함하여 설명변수를 구성하였으므로 하이브리드 모형의 적합성이 영향을 받은 것으로 사료된다. 결국, 조합된 두 모형이 상호보완 되어 하이브리드 모형의 적합성을 높일 수 있지만, 적절한 설명변수의 선정이 우선시 되어야 할 것으로 판단된다.

마지막으로, 본 연구에서 미계측유역으로 사용된 곳은 수문모형 매개변수 검보정 결과가 낮은 유역이기에 지역화모형의 모의 정확성이 상대적으로 더 낮게 추정되었을 수 있다. 따라서 수문모형 매개변수 검보정 결과가 좋은 유역을 대상으로 지역화모형을 적용하여 비교하는 과정이 추가로 필요할 것으로 보인다.

## 감사의 글

본 연구는 환경부 “기후변화대응 환경기술개발사업(과제 번호: 2014001310007)”으로 지원받은 과제입니다.

## References

- Ajami, N. K., Duan, Q., Gao, X., and Sorooshian, S. (2006). “Multi-model combination techniques for analysis of hydrological simulations: application to distributed model intercomparison project results.” *Journal of Hydrometeorology*, Vol. 7, No. 4, pp. 755-768.
- Andrews, F. T., Croke, B. F., and Jakeman, A. J. (2011). “An open software environment for hydrological model assessment and development.” *Environmental Modelling and Software*, Vol. 26, No. 10, pp. 1171-1185.
- Duan, Q., Ajami, N. K., Gao, X., and Sorooshian, S. (2007). “Multi-model ensemble hydrologic prediction using Bayesian model averaging.” *Advances in Water Resources*, Vol. 30, No. 5, pp. 1371-1386.
- Goswami, M., O’connor, K. M., and Bhattarai, K. P. (2007). “Development of regionalisation procedures using a multi-model approach for flow simulation in an ungauged catchment.” *Journal of Hydrology*, Vol. 333, No. 2, pp. 517-531.
- Gottschalk, L., Jensen, J. L., Lundquist, D., Solantie, R., and Tollan, A. (1979). “Hydrologic regions in the Nordic countries.” *Hydrology Research*, Vol. 10, No. 5, pp. 273-286.
- Jakeman, A. J., and Hornberger, G. M. (1993). “How much complexity is warranted in a rainfall runoff model?” *Water Resources Research*, Vol. 29, No. 8, pp. 2637-2649.
- Jeong, D. I., and Kim, Y.-O. (2009). “Combining single-value streamflow forecasts - A review and guidelines for selecting techniques.” *Journal of Hydrology*, Vol. 377, No. 3, pp. 284-299.
- Jin, X., Xu, C. Y., Zhang, Q., and Chen, Y. D. (2009). “Regionalization study of a conceptual hydrological model in Dongjiang basin, south China.” *Quaternary International*, Vol. 208, No. 1, pp. 129-137.
- Kay, A. L., Jones, D. A., Crooks, S. M., Calver, A., and Reynard, N. S. (2006). “A comparison of three approaches to spatial generalization of rainfall-runoff models.” *Hydrological Processes*, Vol. 20, No. 18, pp. 3953-3973.
- Kim, S. U., and Lee, K. S. (2008). “Regional low flow frequency analysis using bayesian multiple regression.” *Journal of Korea Water Resources Association*, Vol. 41, No. 3, pp. 325-340.
- Kim, T. J., Jeong, G. I., Kim, K. Y., and Kwon, H. H. (2015). “A study on regionalization of parameters for sacramento continuous rainfall-runoff model using watershed characteristics.” *Journal of Korea Water Resources Association*, Vol. 48, No. 10, pp. 793-806.
- Lebecherel, L., Andréassian, V., and Perrin, C. (2016). “On evaluating

- the robustness of spatial-proximity-based regionalization methods." *Journal of Hydrology*, Vol. 539, pp. 196-203.
- Merz, R., and Blöschl, G. (2004). "Regionalisation of catchment model parameters." *Journal of Hydrology*, Vol. 287, No. 1, pp. 95-123.
- Moon, J. W., Jung, C. G., and Lee, D. R. (2013). "Parameter regionalization of Hargreaves equation based on climatological characteristics in Korea." *Journal of Korea Water Resources Association*, Vol. 46, No. 9, pp. 933-946.
- Oudin, L., Andréassian, V., Perrin, C., Michel, C., and Le Moine, N. (2008). "Spatial proximity, physical similarity, regression and ungauged catchments: A comparison of regionalization approaches based on 913 French catchments." *Water Resources Research*, Vol. 44, No. 3, W03413.
- Parajka, J., Merz, R., and Blöschl, G. (2005). "A comparison of regionalisation methods for catchment model parameters." *Hydrology and Earth System Sciences Discussions*, Vol. 9, No. 3, pp. 157-171.
- Perrin, C., Michel, C., and Andréassian, V. (2003). "Improvement of a parsimonious model for streamflow simulation." *Journal of Hydrology*, Vol. 279, No. 1, pp. 275-289.
- Perrin, C., Oudin, L., Andreassian, V., Rojas-Serna, C., Michel, C., and Mathevet, T. (2007). "Impact of limited streamflow data on the efficiency and the parameters of rainfall-runoff models." *Hydrological Sciences Journal*, Vol. 52, No. 1, pp. 131-151.
- Post, D. A., and Jakeman, A. J. (1999). "Predicting the daily streamflow of ungauged catchments in SE Australia by regionalising the parameters of a lumped conceptual rainfall-runoff model." *Ecological Modelling*, Vol. 123, No. 2, pp. 91-104.
- Pugliese, A., Farmer, W. H., Castellarin, A., Archfield, S. A., and Vogel, R. M. (2016). "Regional flow duration curves: Geostatistical techniques versus multivariate regression." *Advances in Water Resources*, Vol. 96, pp. 11-22.
- Razavi, T., and Coulibaly, P. (2012). "Streamflow prediction in ungauged basins: review of regionalization methods." *Journal of Hydrologic Engineering*, Vol. 18, No. 8, pp. 958-975.
- Samuel, J., Coulibaly, P., and Metcalfe, R. A. (2011). "Estimation of continuous streamflow in Ontario ungauged basins: Comparison of regionalization methods." *Journal of Hydrologic Engineering*, Vol. 16, No. 5, pp. 447-459.
- Seibert, J. (1999). "Regionalisation of parameters for a conceptual rainfall-runoff model." *Agricultural and Forest Meteorology*, Vol. 98-99, pp. 279-293.
- Sivapalan, M., Takeuchi, K., Franks, S. W., Gupta, V. K., Karambiri, H., Lakshmi, V., Liang, X., McDonnell, J. J., Mendiondo, E. M., O'connell, P. E., Oki, T., Pomeroy, J. W., Schertzer, D., Uhlenbrook, S., and Zehe, E. (2003). "IAHS Decade on Predictions in Ungauged Basins (PUB), 2003-2012: Shaping an exciting future for the hydrological sciences." *Hydrological Sciences Journal*, Vol. 48, No. 6, pp. 857-880.
- Stoll, S., and Weiler, M. (2010). "Explicit simulations of stream networks to guide hydrological modelling in ungauged basins." *Hydrology and Earth System Sciences*, Vol. 14, No. 8, pp. 1435-1448.
- The R Project (1997). accessed 1 April 1997, <<http://www.r-project.org>>.
- Vandewiele, G. L., Xu, C. Y., and Huybrechts, W. (1991). "Regionalisation of physically-based water balance models in Belgium. Application to ungauged catchments." *Water Resources Management*, Vol. 5, No. 3, pp. 199-208.
- Viviroli, D., and Seibert, J. (2015). "Can a regionalized model parameterisation be improved with a limited number of runoff measurements?" *Journal of Hydrology*, Vol. 529, pp. 49-61.
- Vogel, R. M., Wilson, I., and Daly, C. (1999). "Regional regression models of annual streamflow for the United States." *Journal of Irrigation and Drainage Engineering*, Vol. 125, No. 3, pp. 148-157.
- Yadav, M., Wagener, T., and Gupta, H. (2007). "Regionalization of constraints on expected watershed response behavior for improved predictions in ungauged basins." *Advances in Water Resources*, Vol. 30, No. 8, pp. 1756-1774.
- Yapo, P. O., Gupta, H. V., and Sorooshian, S. (1996). "Automatic calibration of conceptual rainfall-runoff models: sensitivity to calibration data." *Journal of Hydrology*, Vol. 181, No. 1-4, pp. 23-48.
- Young, A. R. (2006). "Stream flow simulation within UK ungauged catchments using a daily rainfall-runoff model." *Journal of Hydrology*, Vol. 320, No. 1, pp. 155-172.