

## 확률강우량의 공간분포추정에 있어서 Bayesian 기법을 이용한 공간통계모델의 매개변수 불확실성 해석

서영민 · 박기범<sup>1)\*</sup> · 김성원<sup>2)</sup>

영남대학교 건설시스템공학과, <sup>1)</sup>안동과학대학 건설정보과, <sup>2)</sup>동양대학교 철도토목학과  
(2011년 6월 21일 접수; 2011년 9월 5일 수정; 2011년 12월 7일 채택)

### Uncertainty Analysis of Parameters of Spatial Statistical Model Using Bayesian Method for Estimating Spatial Distribution of Probability Rainfall

Young-Min Seo, Ki-Bum Park<sup>1)\*</sup>, Sungwon Kim<sup>2)</sup>

Department of Civil Engineering, Yeungnam University, Kyeongsan 721-749, Korea

<sup>1)</sup>Department of Construction Information Andong Science College, Andong 760-820, Korea

<sup>2)</sup>Department of Railroad Civil Engineering, Dongyang University, Kyeongbuk 750-711, Korea

(Manuscript received 21 June, 2011; revised 5 September, 2011; accepted 7 December, 2011)

#### Abstract

This study applied the Bayesian method for the quantification of the parameter uncertainty of spatial linear mixed model in the estimation of the spatial distribution of probability rainfall. In the application of Bayesian method, the prior sensitivity analysis was implemented by using the priors normally selected in the existing studies which applied the Bayesian method for the purpose of assessing the influence which the selection of the priors of model parameters had on posteriors. As a result, the posteriors of parameters were differently estimated which priors were selected, and then in the case of the prior combination, F-S-E, the sizes of uncertainty intervals were minimum and the modes, means and medians of the posteriors were similar to the estimates using the existing classical methods. From the comparative analysis between Bayesian and plug-in spatial predictions, we could find that the uncertainty of plug-in prediction could be slightly underestimated than that of Bayesian prediction.

**Key Words** : Probability rainfall, Spatial distribution, Parameter uncertainty, Spatial linear mixed model, Bayesian inference

#### 1. 서론

신뢰성 있는 강우공간분포의 추정은 국지성 호우의 확인 및 분석, 면적평균강우량 추정의 정확도 향상, 강우-유출모델, 특히 분포형 수문모델에서의 모의결

과에 대한 신뢰도 향상에 큰 영향을 미친다. 공간자료에 대한 공간분포추정을 위한 기법으로는 Thiessen 다각형, IDW(inverse distance weighted) 및 등우선법 등과 같은 확정론적 기법과 크리깅 기법으로 대표되는 공간통계학적 기법(spatial statistical method) 또는 지구통계학적 기법(geostatistical method)으로 구분할 수 있다. 기존의 확정론적 기법들의 경우 단순히 면적비 또는 거리를 추정인자로 사용하기 때문에 공간자료가 가지는 공간변동의 통계학적 특성을 적절히

\*Corresponding author : Ki-Bum Park, Department of Construction Information Andong Science College, Andong 760-820, Korea  
Phone: +82-53-321-1517  
E-mail: pkb5032@naver.com

반영하지 못하는 단점을 가지지만 이러한 기법들의 대안으로 사용되는 공간통계학적 기법은 기존의 확정론적 기법과는 달리 자료의 공간상관성을 고려하기 때문에 최근 공간자료에 대한 공간분포추정과 관련하여 여러 분야에서 많은 연구들이 진행되고 있는 추세이다.

대표적인 공간통계학적 기법으로는 기지의 모델 매개변수를 가정한 선형예측기법인 단순크리깅(simple kriging), 기지의 공분산 매개변수를 가정하고 평균을 필터링하는 선형예측기법인 보통크리깅(ordinary kriging), 추세항의 모델링에 따라 보통크리깅을 확장한 일반크리깅(universal kriging), 회귀크리깅(regression kriging), KED(kriging with external draft)가 있으며, 비가우시안 과정(non-Gaussian process)에 대한 비선형 예측기법인 DK(disjunctive kriging), 비매개변수적 예측기법인 지시크리깅(indicator kriging) 등이 있다. 그리고 가우시안 공간과정하에서 공간모델요소에 대한 완전한 매개변수적 설정을 가정하여 관측자료에 대한 조건부 예측기대치를 평가할 수 있는 모델기반접근법(model-based approach)(Diggle 등, 1998)이 있으며, 이러한 모델기반접근법에 기초한 공간선형혼합모델(spatial linear mixed model)은 공간모델요소를 계층적으로 구성하여 Bayesian 기법을 이용한 모델 매개변수의 불확실성 평가를 가능케 한다.

일반적으로 공간통계학적 기법은 관측된 공간자료로부터 공간상관구조의 모델링 및 매개변수 추정이 이루어지고 이를 바탕으로 미계측지점에 대한 공간예측을 수행하게 된다. 여기서, 기존의 공간통계학 분야에서 널리 적용되어온 모델 매개변수의 추정기법에는 보통최소제곱법(ordinary least squares, OLS), 가중최소제곱법(weighted least squares, WLS), 최우도법(maximum likelihood, ML), REML(restricted maximum likelihood) 등이 있으며, 이러한 기법들을 이용하여 추정된 매개변수들은 공간예측에서 플러그인 추정치(plug-in estimate)로 적용되기 때문에 매개변수에 대한 불확실성을 설명할 수 없는 단점을 가진다.

반면에 Bayesian 추정은 매개변수를 확률변수로 처리함으로써 공간예측에서 매개변수의 불확실성을 설명할 수 있기 때문에 강우공간분포 추정에 있어서 모델 매개변수의 불확실성을 정량화할 수 있으며, 이를 바탕으로 강우공간분포, 강우-유출모의 및 홍수위

험도 추정에 있어서의 불확실성을 정량화하여 보다 정교하고 신뢰성있는 추정결과를 제공하고 홍수위험 관리에 있어서의 의사결정지원을 위한 유용한 분석도구 개발에 기여할 수 있는 장점을 가진다. 이와 관련하여 최근 외국에서는 공간예측에 있어서 모델 매개변수의 불확실성을 정량화하기 위하여 Bayesian 기법의 적용이 증가하고 있는 추세이다.

공간통계학적 모델의 매개변수 추정에 있어서 Bayesian 기법을 적용한 기존의 외국 연구들을 살펴보면, Handcock과 Stein(1993)은 평균 매개변수 및 척도 매개변수에 대하여 Jeffrey 사전분포를 적용하였으며, Cressie(1993)은 공간통계학에서의 Bayesian 접근방법을 요약하였다. De Oliveira 등(1997)은 비가우시안 랜덤필드(non-Gaussian random field)를 고려하였으며, 예측에서 매개변수의 불확실성을 설명하기 위하여 BTG(Bayesian transform Gaussian) 모델을 도입하였다. Ecker와 Gelfand(1997)은 Bessel 함수들의 유한혼합을 이용하여 Bayesian 베리오그램 적합절차를 제안하였으며, Le와 Zidek(1992)는 매개변수적 상관함수 대신 공분산 행렬에 대한 역 Wishart 사전분포를 채택한 Bayesian 기법을 제안하였다. Berger 등(2001)은 공간 가우시안 과정의 매개변수에 대한 사전분포로서 무정보적 사전분포를 채택한 Bayesian 기법을 적용하였으며, Diggle과 Ribeiro Jr(2002)는 모델 매개변수 추정에서의 불확실성을 설명하기 위하여 Bayesian 및 모델기법접근법을 적용하였다.

외국의 경우 공간예측에 있어서 모델 매개변수의 불확실성 추정에 대한 활발한 연구가 이루어지고 있는 반면 국내의 경우 수자원 분야와 관련하여 서 등(2010b)과 같이 공간변동성을 고려한 공간예측에서의 불확실성 추정과 같은 유사연구는 있으나 공간예측모델에 대한 매개변수의 불확실성 추정과 관련된 연구는 거의 찾아보기 힘든 실정이다.

따라서 본 연구에서는 확률강우량의 공간분포추정에 있어서 모델 매개변수의 불확실성을 정량화하고 홍수위험도 추정에서 불확실성 인자 추정 및 중요도 분석을 위한 사전연구 수행을 위하여 공간선형혼합모델과 Bayesian 기법을 적용하였다. Bayesian 기법의 적용에서 관련 기존 연구들과는 달리 모델 매개변수에 대한 사전분포 선택에 따른 영향을 평가하기 위하

여 Bayesian 기법을 적용한 기존 연구들에서 일반적으로 채택된 바 있는 사전분포들을 대상으로 총 24개의 사전분포 조합을 구성하여 사전분포 민감도 분석을 실시하였다. 그 결과로부터 모델 매개변수에 대한 적정사전분포를 선정하고 결과된 매개변수의 사후분포와 기존의 매개변수 추정기법을 적용한 매개변수 추정치를 비교하였으며, 마지막으로 Bayesian 공간예측치와 플러그인 공간예측치의 비교를 통해 모델 매개변수 추정에 있어서 Bayesian 기법의 적용성을 평가하였다.

## 2. Bayesian 기법을 이용한 모델 매개변수 추정

### 2.1. 공간선형혼합모델

어떤 영역  $D$  내에서 일변량 공간자료  $(x_i, y_i), i = 1, \dots, n$ 는 공간위치  $x_i$ 와 그 위치에서 확률변수  $Y$ 의 측정치  $y_i$ 로 구성되며, 이러한 공간자료는 공간추계과정(spatial stochastic process)에 대한 표본위치에서의 실현치(realization)으로 간주된다.

기존의 공간통계학에서 기본이 되는 조건부 모델 설정에서 확률변수  $Y$ 는 Eq. (1)과 같이 공간과정  $S(x)$ 가 주어졌을 때 정규분포를 따르고 조건부 독립으로 가정된다.

$$[Y(x_i)|S] \sim N(D(x_i)\beta + S(x_i), \tau^2 I), \quad i = 1, \dots, n \quad (1)$$

여기서,  $[\cdot]$ 는 확률변수의 분포를 의미하며,  $D(x_i)\beta$ 는 모델의 평균부 또는 공간추세,  $D(x_i)$ 는  $n \times p$  공변량 행렬(covariate matrix),  $p$ 는 공변량의 개수,  $\beta$ 는 평균 매개변수(mean parameter),  $\tau^2$ 는 조건부 분산 또는 측정오차분산으로서 덩어리 효과(nugget effect)를 나타낸다. 공간과정  $\{S(x); x \in D \subset \mathbb{R}^2\}$ 는 영평균, 분산  $\sigma^2$  및 상관함수  $\rho(h; \phi)$ 를 가지는 정상 가우시안 과정(stationary Gaussian process)이며, 여기서  $\phi$ 는 상관 매개변수(correlation parameter),  $h$ 는 두 위치간의 벡터거리이다.

Eq. (1)과 같이 정의된 공간모델은 계층적 구성(hierarchical scheme)을 이루는 공간선형혼합모델

(spatial linear mixed model)로 Eqs. (2)~(4)와 같이 다시 나타낼 수 있으며, 공간과정  $S(x)$ 는  $\sigma_k^2$ 에 의해 척도화된 잠재과정  $T_k(x)$ 의 합으로 분해될 수 있다 (Ribeiro Jr과 Diggle, 1999).

$$\begin{aligned} \text{Level 1: } Y(x) &= D(x)\beta + S(x) + \varepsilon(x) \quad (2) \\ &= D(x)\beta + \sum_{k=1}^K \sigma_k T_k(x) + \varepsilon(x) \end{aligned}$$

$$\text{Level 2: } T_k(x) \sim N(0, R_k(\phi_k)), \varepsilon(x) \sim N(0, \tau^2 I) \quad (3)$$

$$\text{Level 3: } [\beta, \sigma^2, \phi, \tau^2] \sim pr(\cdot) \quad (4)$$

여기서, 표준화된 잠재정상공간과정  $T_k$ 는 영평균, 분산 1, 상관행렬  $R_k(\phi_k)$ 를 가지며,  $R_k(\phi_k)$ 의 요소는 상관 매개변수  $\phi_k$ 를 가지는 상관함수  $\rho_k(h; \phi_k)$ 에 의해 주어진다. 잠재과정  $T_1, \dots, T_K$ 는 상호독립이고 공간과정  $S$ 는 척도화된 잠재과정의 합인  $S(x) = \sum_{k=1}^K \sigma_k T_k(x)$ 로 나타낼 수 있다. 만약 그 과정이 등방성(isotropic)일 경우  $h$ 는 스칼라값으로서 두 위치간의 유클리드 거리(Euclidean distance)로 축소되며,  $\sigma_k^2$ 는 척도 매개변수(scale parameter)이고  $\varepsilon(x)$ 는 표본위치  $x$ 에서의 오차(잡음)벡터로서 영평균과 분산  $\tau^2$ 를 가지는 독립공간과정, 즉 공간백색잡음(spatial white noise)이다. 그리고 세 번째 레벨은 Bayesian 추정에서 모델 매개변수에 대한 사전분포  $pr(\cdot)$ 로 구성된다.

### 2.2. 모델 매개변수에 대한 Bayesian 추정

공간추세  $\mu(x)$ 에 대하여 선형모델을 가정한 가우시안 모델을 고려할 경우 공간추세에 있어서 다항추세평면(polynomial trend surface) 또는 공간참조된 공변량을 포함할 수 있다. 따라서  $\mu(x) = D\beta$ 에 대하여 Eqs. (2)~(4)로 주어진 공간선형혼합모델에 의한  $Y$ 의 분포는 Eq. (5)와 같이 나타낼 수 있으며, 대수우도함수는 Eq. (6)과 같이 나타낼 수 있다(Diggle과 Ribeiro Jr, 2007).

$$[Y] \sim N(D\beta, \sigma^2 R(\phi) + \tau^2 I) \quad (5)$$

$$L(\beta, \tau^2, \sigma^2, \phi) = -0.5\{n \log(2\pi) + \log\{(\sigma^2 R(\phi) + \tau^2 I)\} + (y - D\beta)^T (\sigma^2 R(\phi) + \tau^2 I)^{-1} (y - D\beta)\} \quad (6)$$

선형회귀형태의 공간추세를 포함하는 Eq. (5)와 같은 가우시안 모델에 대하여 덩어리 효과(nugget effect)를 무시, 즉  $\tau^2 = 0$ 일 경우 특정 사전분포를 설정함에 따라  $\beta$ 와  $\sigma^2$ 에 대한 사후분포를 해석적으로 유도할 수 있으며, 이러한 결과를 이용하여 모든 매개변수들이 미지값일 경우에 대한 사후분포 및 예측분포를 유도할 수 있다.

기지의 고정된  $\phi$ 에 대하여  $(\beta, \sigma^2)$ 에 대한 공액사전분포로서 정규분포와 척도화 역카이제곱분포의 곱으로 나타낼 수 있는 정규척도화 역카이제곱(normal-scaled-inverse- $\chi^2$ ) 분포를 고려할 수 있다. 따라서 각각 Eqs. (7) and (8)과 같이 정규분포와 척도화 역카이제곱분포를 가지는  $\beta$ 와  $\sigma^2$ 에 대한 사전분포를 설정할 수 있으며, 표기의 편의를 위하여  $(\beta, \sigma^2)$ 에 대한 사전분포는 Eq. (9)와 같이 나타낼 수 있다(Diggle과 Ribeiro Jr, 2007).

$$[\beta | \sigma^2, \phi] \sim N(m_b, \sigma^2 V_b) \quad (7)$$

$$[\sigma^2 | \phi] \sim \chi_{S_I}^2(n_\sigma, S_\sigma^2) \quad (8)$$

$$[\beta, \sigma^2 | \phi] \sim N\chi_{S_I}^2(m_b, V_b, n_\sigma, S_\sigma^2) \quad (9)$$

여기서,  $m_b$ 와  $V_b$ 는 각각  $\beta$ 에 대한 평균과 공분산 행렬이며,  $n_\sigma$ 와  $S_\sigma^2$ 는 각각  $\sigma^2$ 에 대한 척도화 역카이제곱분포에서의 자유도와 척도 매개변수이고  $\chi_{S_I}^2(n_\sigma, S_\sigma^2)$ 는 Eq. (10)과 같은 확률밀도함수를 가진다.

$$\pi(z) \propto z^{-(n_\sigma/2+1)} \exp(-n_\sigma S_\sigma^2 / (2z)), \quad z > 0 \quad (10)$$

베이즈 정리(Bayes' Theorem)를 이용하여 사전분포 Eq. (9)와 우도함수 Eq. (6)을 결합하면 Eq. (11)과 같은  $(\beta, \sigma^2)$ 에 대한 사후분포를 얻을 수 있다(Diggle과 Ribeiro Jr, 2007).

$$[\beta, \sigma^2 | y, \phi] \sim N\chi_{S_I}^2(\tilde{\beta}, V_{\tilde{\beta}}, n_\sigma + n, S^2) \quad (11)$$

여기서,  $\tilde{\beta} = V_{\tilde{\beta}}(V_b^{-1}m_b + D'R^{-1}y)$ ,  $V_{\tilde{\beta}} = (V_b^{-1} + D'R^{-1}D)^{-1}$ 이고  $S^2$ 는 Eq. (12)와 같다.

$$S^2 = \frac{n_\sigma S_\sigma^2 + m_b' V_b^{-1} m_b + y' R^{-1} y - \tilde{\beta}' V_{\tilde{\beta}}^{-1} \tilde{\beta}}{n_\sigma + n} \quad (12)$$

한편 더욱 현실적인 매개변수 추정으로서 모든 모델 매개변수에 대한 불확실성을 고려한 경우를 생각할 수 있다. 마찬가지로  $\tau^2 = 0$ 이고 단일 상관 매개변수  $\phi$ 를 고려할 경우 모델 매개변수에 대한 사전분포는  $[\beta, \sigma^2, \phi] = [\beta, \sigma^2 | \phi][\phi]$ 라고 간주할 수 있다. 즉, 모델 매개변수에 대한 사전분포는 Eq. (9)와  $\phi$ 에 대한 사전분포  $[\phi]$ 의 곱이라 할 수 있으며, 이때 사후분포는 Eq. (13)과 같이 나타낼 수 있다. 여기서,  $[\beta, \sigma^2 | y, \phi]$ 는 Eq. (11)과 같고  $[\phi | y]$ 는 Eq. (14)와 같다(Diggle과 Ribeiro Jr, 2007).

$$[\beta, \sigma^2, \phi | y] = [\beta, \sigma^2 | y, \phi][\phi | y] \quad (13)$$

$$p(\phi | y) \propto \pi(\phi) |V_{\tilde{\beta}}|^{1/2} |R|^{-1/2} (S^2)^{-(n+n_\sigma)/2} \quad (14)$$

여기서,  $\pi(\phi)$ 는  $\phi$ 의 사전분포이고  $(\beta, \sigma^2, \phi)$ 의 사전분포로서  $\pi(\beta, \sigma^2, \phi) \propto \pi(\phi)/\sigma^2$ 를 고려할 경우에는 Eq. (14)에서  $n_\sigma = -p$ 가 된다. 원칙적으로  $\phi$ 의 사전분포는 연속지지집합(continuous support set)을 가져야 하지만 실제 적용에서는 일반적으로 이산사전분포를 적용하며,  $\phi$ 의 사전분포로서 균등형, 지수형, 역수형 및 제곱역수형이 적용될 수 있다(Ribeiro Jr과 Diggle, 1999).

### 3. 적용 분석

#### 3.1. 적용구역 및 분석개요

본 연구는 확률강우량의 공간분포추정에 있어서 공간통계학적 모델의 매개변수에 대한 불확실성을 추정하기 위하여 공간선형혼합모델 및 Bayesian 기법을

적용하였다. 이를 위하여 국제수문개발계획(IHP)에 의해 선정 및 운영된 바 있는 위천대표유역을 포함하고 있어 다수의 양질의 수문자료를 보유하고 있으며, 유역의 대부분이 산지로 이루어져 있고 도시화율이 낮은 위천유역을 분석유역으로 선정하였다.

입력자료인 대상유역내 지점확률강우량은 위천유역의 18개 관측지점에 대한 확률강우량 분석결과(확률분포형: Gumbel, 100년 빈도, 지속시간 24시간)(서 등, 2010a)를 사용하였다. 본 연구의 경우 공간선형혼합모델의 매개변수의 불확실성 해석에 초점을 두고 있기 때문에 공간상관성에 대한 베리오그램 모델에서 등(2010a, 2010b)에서 채택한 구형모델(spherical model)을 동일하게 적용하였다.

모델 매개변수 추정, 공간예측 등을 위하여 통계계산 및 그래픽을 목적으로 GNU 프로젝트로 개발된 프로그래밍 언어인 R(R development Core Team, 2011)을 이용하였으며, 본 연구의 대상 유역도와 입력자료인 확률강우량에 대한 버블차트는 각각 Figs 1 and 2와 같다.

본 연구에서는 공간선형혼합모델의 매개변수, 즉 평균 매개변수, 척도 매개변수 및 상관 매개변수에 대한 불확실성을 추정하기 위하여 Bayesian 기법을 적용하였으며, 그 결과를 비교하기 위하여 기존의 공간통계학에서 널리 적용되고 있는 모델 매개변수 추정 기법인 보통최소제곱법, 가중최소제곱법, 최우도법 및 REML을 적용하였다. 그리고 Bayesian 기법을 적용함에 있어서 사전분포 선택에 따른 민감도 평가 및

최종 사전분포 선택을 위하여 기존 연구들(Ribeiro Jr 과 Diggle, 1999; Diggle과 Ribeiro Jr, 2002; Diggle과 Ribeiro Jr, 2007)에서 적용된 바 있는 사전분포들을 대상으로 Bayesian 추정을 실시하였으며, 여기서 평균 매개변수의 경우 정규분포 및 균등분포, 척도 매개변수의 경우 척도화역카이제곱분포, 균등분포 및 역수분포, 상관 매개변수의 경우 균등분포, 지수분포, 역수분포 및 제곱역수분포를 적용하였다. 본 연구의 분석절차에 대한 개요는 Fig. 3과 같다.

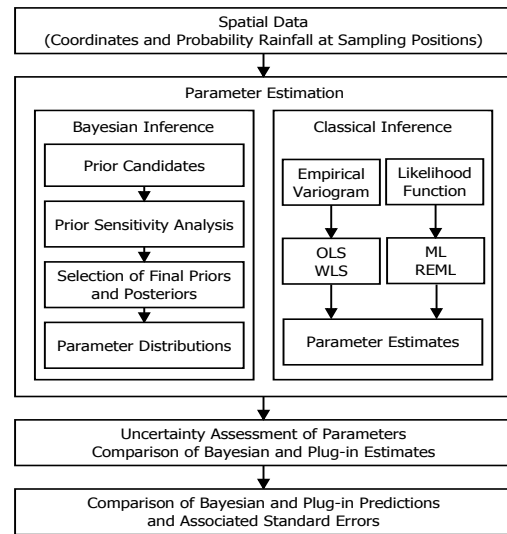


Fig. 3. Flow Chart for Bayesian and Classical Inference in Parameter Estimation and Spatial Prediction.



Fig. 1. Study Area and Raingauge Location.

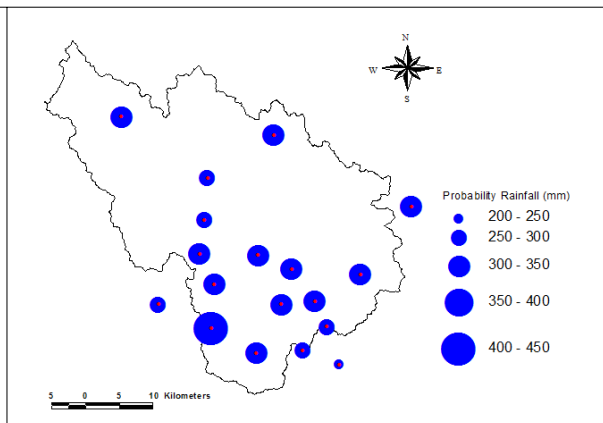


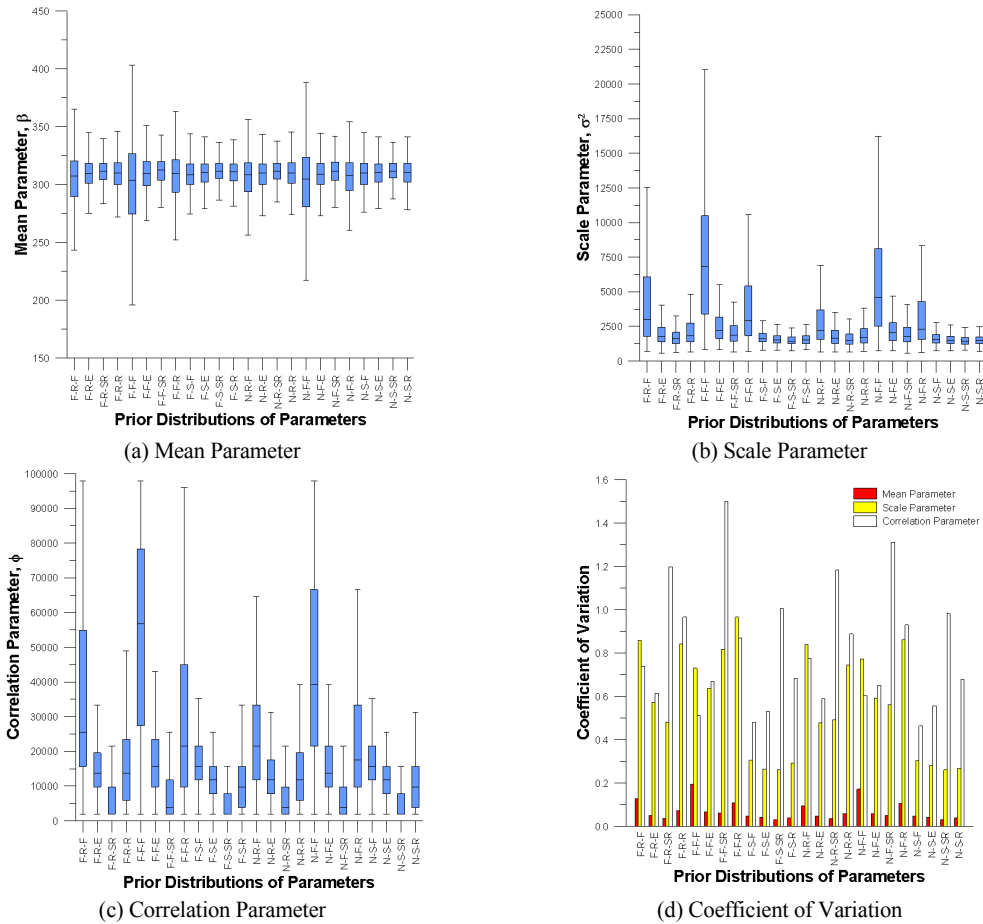
Fig. 2. Bubble Plot of Probability Rainfall. (Duration 24hr, Return Period 100yr)

한편 공간추세 및 공간상관성에 대한 모델링을 단순화시키기 위하여 보조변수와 덩어리 효과는 고려하지 않았으며, 사후분포로부터 표본을 모의하기 위해서 다음과 같은 절차를 적용하였다. 먼저 지지집합을 이산화하고 Eq. (14)를 이용하여  $p(\phi|y)$ 를 계산하였으며, 다음으로  $[\phi|y]$ 로부터  $\phi$ 의 값을 모의한 후  $[\beta, \sigma^2|y, \phi]$ 에 모의값을 추가하고 다시 이 분포에 대한 모의를 수행하였다. 그리고 표본발생수 만큼 1,000회의 모의를 반복하여 최종적으로 모델 매개변수들에 대한 결합사후분포로부터  $(\beta, \sigma^2, \phi)$ 에 대한 표본을 발생시켰다.

3.2. 사전분포 민감도 분석

본 연구에서는 확률강우량의 공간분포추정에 있어서 공간선형혼합모델에 대한 매개변수의 불확실성을 정량화하기 위하여 Bayesian 기법을 적용하였다. Bayesian 기법은 매개변수에 대한 사전분포를 자료에 기반한 우도함수와 베이즈 정리를 이용하여 갱신시켜 매개변수에 대한 사후분포를 유도하는 기법으로서 일반적으로 자료의 길이가 충분하지 않을 경우 갱신과정에서 사전분포의 영향을 압도할 수 없기 때문에 추정결과가 사전분포의 선택에 따라 상당한 영향을 받을 수 있다.

따라서 본 연구에서는 Bayesian 기법의 적용에 있



※ F: flat, R: reciprocal, E: exponential, SR: squared reciprocal, S: scaled-inverse- $\chi^2$ , N: normal

Fig. 4. Prior Sensitivity Analysis for Model Parameters.

어서 모델 매개변수들에 대한 사전분포 선택에 따른 사후분포 추정결과의 영향을 평가하고 이러한 결과로부터 최종 사전분포를 선택하기 위하여 각 매개변수 별로 고려된 사전분포들의 모든 조합을 토대로 하여 Bayesian 추정을 실시하였다. 이러한 분석방법을 사전분포 민감도 분석이라고 하며, 본 연구에서는 총 24 개의 사전분포조합을 고려하였다.

Fig. 4은 모델 매개변수의 사전분포 선택조합에 따른 각 매개변수별 사후분포를 박스그림으로 요약하여 나타낸 것으로서 사전분포의 선택에 따라 매개변수에 대한 불확실성 추정범위가 상이하게 나타남을 알 수 있다.

Fig. 3-(a)로부터 척도 매개변수 및 상관 매개변수에 대한 사전분포로서 각각 역수분포와 균등분포 또는 모두 균등분포로 선택할 경우 평균 매개변수에 대한 불확실성 범위가 상대적으로 크게 추정되었으며, 사전분포 선택조합에서 척도 매개변수의 사전분포로서 척도화역카이제곱분포를 선택한 경우가 다른 경우에 비해 평균 매개변수에 대한 불확실성 범위가 작은 것으로 분석되었다. 평균 매개변수에 대한 최빈값, 평균값 및 중앙값의 경우 기존 추정기법들의 결과(Table 1 참조)와 대체로 비슷하게 추정되었다.

Fig. 3-(b)로부터 척도 매개변수에 대한 사전분포로서 척도화역카이제곱분포를 선택한 경우 다른 경우에 비해 척도 매개변수의 불확실성 범위가 작게 추정됨을 알 수 있다. 그리고 Fig. 3-(c)로부터 상관 매개변수의 사전분포로서 균등사전분포를 선택한 경우 대체로 기존의 매개변수 추정기법들에 비해 최빈값, 평균값, 중앙값 및 불확실성 구간의 크기가 크게 추정되었으며, 제곱역수형 사전분포의 경우 우측왜곡의 정도가 심하고 다른 사전분포조합 및 기존 기법들에 비해 최빈값, 평균값 및 중앙값이 작게 추정되었다.

Fig. 3-(d)는 각 매개변수에 대한 사전분포 선택조

합에 따른 변동계수를 비교한 것으로서 다른 매개변수들에 비해 평균 매개변수의 변동계수가 훨씬 작게 나타났으며, 대체로 상관 매개변수의 변동계수가 가장 크게 나타남을 알 수 있다. 그리고 사전분포 선택조합 중에서 척도 매개변수에 대하여 척도화역카이제곱 분포, 상관 매개변수에 대하여 균등분포, 지수분포, 및 역수분포를 선택한 경우 대체로 다른 조합들에 비해 변동계수가 낮게 추정되는 것으로 분석되었다.

따라서 사전분포 선택조합에 따른 매개변수 추정치의 민감도 분석(Fig. 4) 및 기존 추정기법과의 비교 (Table 1)로부터 최종적으로 사전분포 F-S-E를 선택하여 이후의 분석에 고려하였다. 여기서, 사전분포의 선택은 Fig. 4로부터 불확실성 구간 및 변동계수의 크기가 작고 사후분포의 최빈값, 평균값 및 중앙값이 기존 기법들의 추정치(Table 1)와 크기가 비슷한 사전분포 조합을 기준으로 선택되었다.

Table 1은 기존의 모델 매개변수 추정기법, 즉 보통 최소제곱법, 가중최소제곱법, 최우도법 및 REML을 적용하여 추정된 모델 매개변수값과 Bayesian 기법을 적용하여 추정된 모델 매개변수값을 비교한 것으로서 Bayesian 기법을 이용한 매개변수 추정치의 경우 최종 선택된 사전분포 F-S-E를 고려한 매개변수의 사후분포에 대한 평균값 및 중앙값을 나타낸다. Table 1로부터 매개변수의 사후분포에 대한 최빈값, 평균값 및 중앙값은 기존 매개변수 추정기법을 이용한 추정결과와 대체로 비슷한 결과를 보여줌을 알 수 있다.

### 3.3. 모델 매개변수에 대한 불확실성 정량화 및 공간예측 비교

본 연구에서는 모델 매개변수에 대한 불확실성을 정량화하기 위하여 3.2절에서 최종선택된 사전분포 F-S-E에 대하여 각 매개변수의 사후분포를 추정하고 요약통계치를 정리하였다. 전절에서 살펴본 바와 같

**Table 1.** Comparison of Parameter Estimates from Bayesian and Classical Methods

MODEL PARAMETERS	CLASSICAL INFERENCE				BAYESIAN INFERENCE		
	OLS	WLS	ML	REML	MODE	MEDIAN	MEAN
Mean Parameter	312.2	312.2	311.1	310.1	310.4	310.4	309.8
Scale Parameter	1,387.3	1,560.0	1,385.8	1,531.3	1,411.7	1,530.0	1,598.0
Correlation Parameter	11,405.8	13,012.6	11,719.2	12,933.5	11,100.0	11,750.0	12,010.0

이 모델 매개변수에 대한 사전분포 선택에 따라 매개변수 추정치 및 불확실성 범위가 상이하며, 이러한 사전분포 민감도 분석을 이용한 최적 사전분포 선택을 통해 매개변수 추정치가 기존 방법에 비해 편이되지 않고 불확실성 범위가 상대적으로 작은 사전분포를 설정함으로써 보다 신뢰성 있는 매개변수 추정 및 불확실성 추정이 이루어질 수 있다.

각 매개변수에 대하여 추정된 사후분포에 대한 요약통계치 및 확률분포는 각각 Table 2 및 Fig. 5와 같다. Table 2는 모델 매개변수의 사후분포에 대한 요약통계치를 나타낸 것으로서 사후분포에 대한 통계적 특성치와 불확실성 범위를 정리하였으며, 불확실성 범위의 경우 범위(range), 사분위수 범위(interquartile range), 신용구간(credible interval)으로 나타내었다. Table 2로부터 각 매개변수들의 불확실성 범위를 살펴보면 평균 매개변수의 경우 범위가 [247.1, 349.5]로서 크기가 102.4, 사분위수 범위가 [302.1, 317.9]로서 크기가 15.8이고 척도 매개변수의 경우 범위가 [797.9, 4,028.0]으로서 크기가 3,230.1, 사분위수 범위가 [1,296.0, 1,844.0]으로서 크기가 548.0으로 분석되었으며, 상관 매개변수의 경우 범위가 [1,959.0, 35,260.0]으로서 크기가 33,301.0, 사분위수 범위가 [7,836.0, 15,670.0]으로서 크기가 7,834.0으로 분석되

었다. 그리고 각 매개변수에 대한 변동계수를 살펴보면 평균, 척도 및 상관 매개변수가 각각 0.047, 0.266, 0.530으로서 산포도의 정도가 평균, 척도 및 상관 매개변수의 순으로 낮음을 알 수 있다.

Fig. 5는 모델 매개변수에 대한 사후분포를 확률밀도함수로 나타낸 것으로서 평균 매개변수의 경우 대체로 확률밀도함수의 형상이 대칭형이며, 척도 및 상관 매개변수의 경우 확률밀도함수의 형상이 우측으로 왜곡된 비대칭형임을 알 수 있다. 각 매개변수의 사후분포에 대한 최빈값을 Bayesian 추정치의 대표값으로 할 경우 Bayesian 추정치와 기존 매개변수 추정기법으로부터 추정된 매개변수 추정치를 비교해보면 평균 매개변수의 경우 ML 추정치 및 REML 추정치와 Bayesian 추정치는 비슷한 결과를 나타내고 있으나 OLS 추정치 및 WLS 추정치와 큰 차이를 보이지 않으며, 척도 매개변수의 경우 ML 추정치 및 OLS 추정치와 Bayesian 추정치가 비슷한 결과를 나타내고 상관 매개변수의 경우 ML 추정치 및 OLS 추정치가 Bayesian 추정치와 비슷한 결과를 나타냄을 알 수 있다. 따라서 기존의 모델 매개변수 추정기법 중에서 최우도법을 적용한 ML 추정치가 Bayesian 추정치와 가장 비슷한 결과를 나타내는 것으로 분석되었다.

Table 2. Summary Statistics for Posteriors of Model Parameters

PRIOR	SUMMARY STATISTICS OF POSTERIOR	MEAN PARAMETER	SCALE PARAMETER	CORRELATION PARAMETER	
F-S-E	Minimum	247.1	797.9	1,959.0	
	1st Quartile	302.1	1,296.0	7,836.0	
	Median	310.4	1,530.0	11,750.0	
	Mean	309.8	1,598.0	12,010.0	
	3rd Quartile	317.9	1,844.0	15,670.0	
	Maximum	349.5	4,028.0	35,260.0	
	Standard Deviation	12.9	424.6	6,363.0	
	Coefficient of Variation	0.047	0.266	0.530	
	Interquartile Range	15.8	548.0	7,834.0	
	Range	102.4	3,230.1	33,301.0	
	Credible Interval (95%)	Lower	282.9	928.3	2,691.5
		Upper	335.9	2,576.8	34,327.9



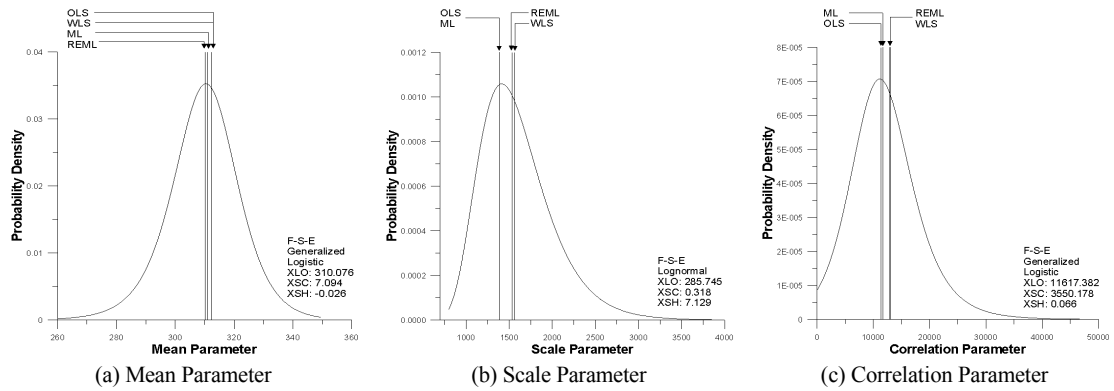


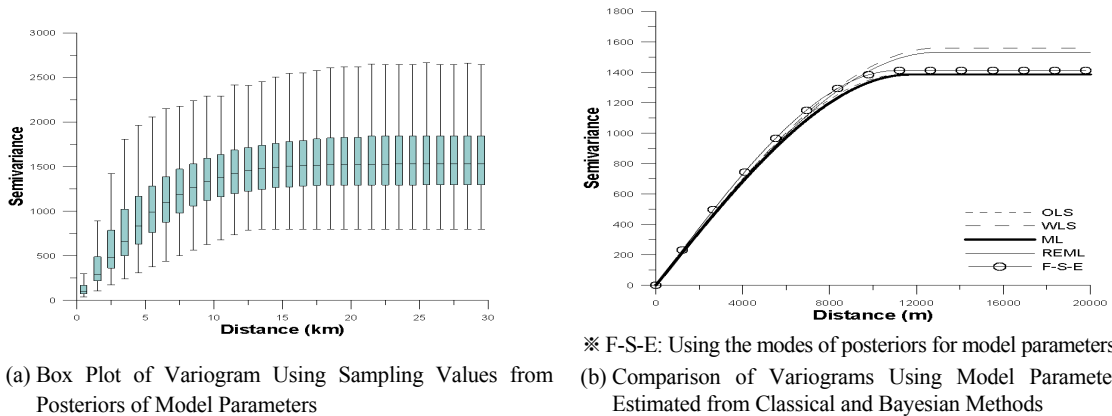
Fig. 5. Posteriors of Model Parameters.

Fig. 6은 모델 매개변수에 대한 사후분포로부터 베리오그램의 불확실성 구간을 나타낸 상자그림(Fig. 6-(a))과 기존 및 Bayesian 매개변수 추정치에 대한 베리오그램(Fig. 6-(b))을 비교한 것이다. Fig. 6-(b)에서 Bayesian 추정치로부터 구축된 베리오그램은 ML 및 OLS 추정치로부터 구축된 베리오그램과 비슷한 결과를 나타냄을 알 수 있다.

Fig. 7은 Bayesian 및 플러그인 공간예측에 대한 공간예측치 및 표준오차를 비교한 것으로서 Fig. 7-(a)로부터 기존 매개변수 추정기법을 이용한 매개변수 추정치에 대한 플러그인 예측치와 Bayesian 예측치의 비교결과 대각선에 대한 값의 밀집도가 가장 높은 ML 추정치를 이용한 플러그인 예측치와 Bayesian 예측치가 가장 비슷한 결과를 나타내는 것으로 분석되었다. 따

라서 ML 추정치를 이용한 플러그인 예측치는 Bayesian 예측치의 대안으로 사용될 수 있음을 알 수 있다.

Fig. 7-(b)로부터 플러그인 예측치에 대한 표준오차는 Bayesian 예측치에 대한 표준오차보다 다소 과소 추정되는 것으로 분석되었으며, 표준오차가 커질수록 플러그인 예측치와 Bayesian 예측치에 대한 표준오차의 차이가 대체적으로 다소 커지는 것으로 나타났다. 일반적으로 공간예측에서 관측밀도가 낮은 지역의 경우 표준오차가 크게 나타나기 때문에 이러한 차이는 관측지점이 밀집되지 않은 지역에서 두드러지게 나타나며, 이러한 결과로부터 플러그인 예측치는 Bayesian 예측치에 비해 공간예측에 대한 불확실성이 다소 과소평가될 수 있으며, 특히 관측밀도가 낮은 지역의 경우 이러한 경향은 더욱 크게 나타남을 알 수 있다.



(a) Box Plot of Variogram Using Sampling Values from Posteriors of Model Parameters

(b) Comparison of Variograms Using Model Parameters Estimated from Classical and Bayesian Methods

Fig. 6. Variograms from Classical and Bayesian Methods.

※ F-S-E: Using the modes of posteriors for model parameters

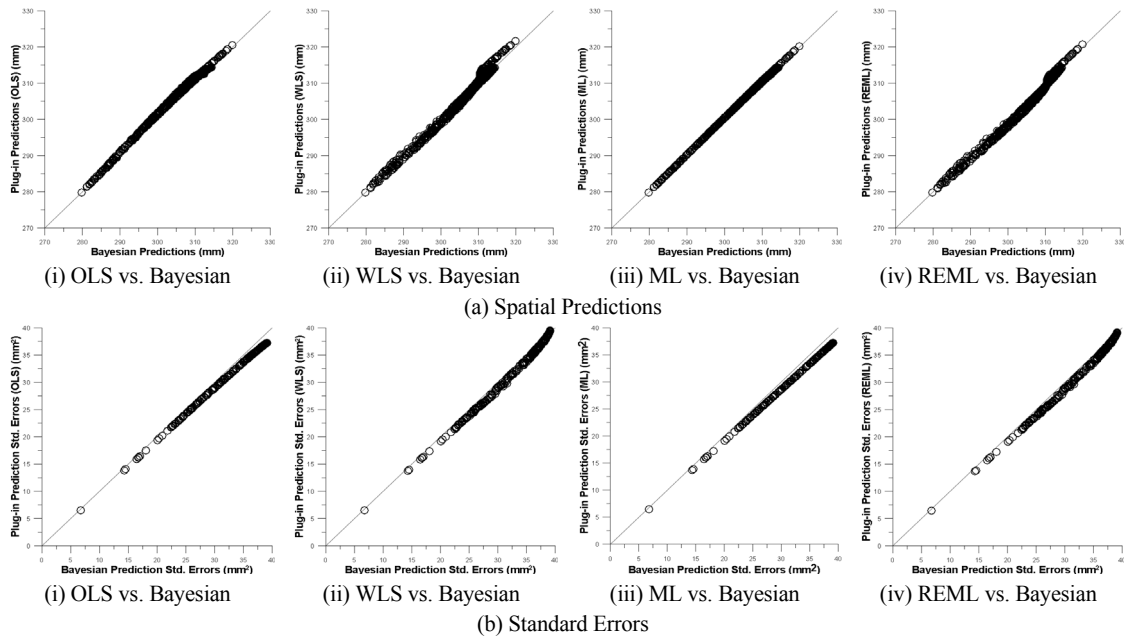


Fig. 7. Comparison of Plug-in and Bayesian Predictions, and Associated Standard Errors.

4. 결론

본 연구는 확률강우량의 공간분포추정에 있어서 모델 매개변수에 대한 불확실성 정량화하기 위하여 Bayesian 기법을 적용하였다. 위천유역을 대상으로 한 적용분석에서 모델 매개변수에 대한 사전분포 설정 및 사전분포에 따른 영향을 평가하기 위하여 Bayesian 기법을 적용한 기존 연구들에서 일반적으로 채택된 바 있는 사전분포들을 대상으로 사전분포 민감도 분석을 실시하였다.

사전분포 민감도 분석결과로부터 매개변수의 불확실성 추정에 적용할 최적 사전분포를 선정하고 결과된 매개변수의 사후분포와 기존의 매개변수 추정기법을 적용한 매개변수 추정치를 비교하였다. 그리고 모델 매개변수에 대한 Bayesian 추정치와 기존 기법들로부터 결과된 추정치들로부터 베리오그램 비교, 공간예측치 및 표준오차에 대한 비교분석을 실시하였으며, 이러한 일련의 비교분석들로부터 얻은 결과를 정리하면 다음과 같다.

- 1) 모델 매개변수에 대한 사전분포 선택에 따른 민감도 분석 결과 Bayesian 기법을 이용하여 추정

된 매개변수의 불확실성 범위는 사전분포의 선택에 따라 상이한 결과가 도출됨을 확인할 수 있었으며, 최종 사전분포조합으로서 불확실성 범위가 가장 작고 최빈값, 평균값 및 중앙값이 기존 매개변수 추정기법의 추정결과와 비슷한 F-S-E(평균 매개변수: 균등분포, 척도 매개변수: 척도화역카이제곱분포, 상관 매개변수: 지수분포)가 선택되었다.

- 2) 모델 매개변수에 대한 Bayesian 추정치로부터 구축된 베리오그램은 OLS 추정치 및 ML 추정치에 대한 베리오그램과 비슷한 결과를 보이는 것으로 분석되었다.
- 3) Bayesian 및 플러그인 공간예측에 대한 공간예측치 및 표준오차의 비교로부터 Bayesian 예측치는 ML 추정치를 이용한 플러그인 예측치와 가장 비슷한 공간예측결과를 보이는 것으로 분석되었으며, 플러그인 예측치에 대한 표준오차는 Bayesian 예측치에 대한 표준오차보다 다소 과소추정되는 것으로 분석되었다. 이러한 결과로부터 플러그인 예측치는 Bayesian 예측치에 비해 공간예측에 대한 불확실성이 다소 과소평

가될 수 있음을 확인할 수 있었다.

- 4) 본 연구에서 적용한 Bayesian 모델 매개변수 추정기법은 기존 기법과 거의 비슷한 추정결과를 나타내고 있으나 모델 매개변수에 대한 사후분포를 추정함으로써 매개변수에 대한 불확실성을 정량화할 수 있기 때문에 공간선형혼합모델의 매개변수 추정 및 공간예측에서의 신뢰도를 향상시키는 데 있어서 그 적용성을 확인할 수 있었다.

본 연구에서는 기존 연구들과는 달리 사전분포 민감도 분석을 통해 최적 사전분포를 선택함으로써 사전분포 선택에 따른 불확실성을 최소화하였다. 외국의 경우와 같이 유역내 강우관측지점이 많을 경우 사전분포의 선택에 따른 영향은 크지 않지만 국내와 같이 유역내 강우관측지점이 적을 경우는 Bayesian 기법의 적용에 있어서 사전분포의 영향을 입력자료가 압도할 수 없기 때문에 본 연구에서와 같이 사전분포의 민감도를 평가하는 것은 중요한 과정이라 할 수 있다.

한편 확률강우량의 공간분포 추정에 있어서 모델 매개변수에 대한 불확실성 평가는 홍수 위험도 해석에서 홍수량의 불확실성에 가장 큰 영향을 미치는 면적평균강우량의 불확실성에 대한 영향인자를 규명하는데 중요한 역할을 할 수 있다. 그러나 본 연구에서는 불확실성 요소로서 모델 매개변수에 대한 불확실성만 고려하고 있기 때문에 향후 추가연구를 통해 강우자료오차, 강우공간변동성 등과 같은 추가적인 불확실성 요소에 대한 고려로 확대되어야 할 것이다.

본 연구에서 적용한 공간통계기법 및 Bayesian 기법은 입력자료의 수가 많을수록 결과의 신뢰성이 보장되기 때문에 관측지점의 수를 어느정도 확보할 수 있는 중규모 이상의 유역으로 적용이 제한된다. 그리고 Bayesian 기법의 적용에 있어서 본 연구의 경우 매개변수에 대한 사전분포의 설정은 기존 연구들에서 적용하고 있는 사전분포들, 즉 부적절 사전분포 또는 공액사전분포에 한정되어 있기 때문에 민감도 분석에 적용될 사전분포들의 선택에는 불가피하게 분석자의 주관에 개입될 수밖에 없는 한계를 가진다. 또한 선택된 사전분포에 따라 모델 매개변수의 불확실성이 과대 또는 과소추정될 수도 있기 때문에 향후 연구에서는 정보적 사전분포 구축을 통한 객관적인 사전분포

의 선택에 대한 추가연구의 필요성이 제기될 수 있을 것이다. 이와 관련하여 정보적 사전분포의 구축 및 적용을 위하여 bootstrap 및 Bayesian MCMC를 결합한 형태의 방법론에 대한 연구가 제안될 수 있을 것이다.

### 참 고 문 헌

서영민, 여운기, 이승윤, 지흥기, 2010a, 확률강우량의 공간분포 추정을 위한 KED 기법의 적용, 한국수자원학회논문집, 43(8), 757-767.

서영민, 여운기, 이승윤, 지흥기, 2010b, 확률강우량의 공간분포에 대한 불확실성 해석 - CEM과 SGS 기법의 비교, 한국수자원학회논문집, 43(11), 933-944.

Berger, J., De Oliveira, V., Sanso, B., 2001, Objective Bayesian analysis of spatially correlated data, *Journal of the American Statistical Association*, 96, 1361-1374.

Cressie, N., 1993, *Statistics for spatial data*, Wiley, New York.

De Oliveira, V., Kedem, B., Short, D., 1997, Bayesian prediction of transformed Gaussian random fields, *Journal of the American Statistical Association*, 92(440), 1422-1433.

Diggle, P. J., Ribeiro Jr, P. J., 2002, Bayesian inference in Gaussian model-based geo-statistics, *Geographical and Environmental Modelling*, 6(2), 129-146.

Diggle, P. J., Ribeiro Jr, P. J., 2007, *Model-based geostatistics*, Springer, New York.

Diggle, P. J., Tawn, J., Moyeed, R., 1998, Model based geostatistics (with discussion), *Journal of the Royal Statistical Society, Series C*, 47(3), 299-350.

Ecker, M., Gelfand, A., 1997, Bayesian variogram modeling for an isotropic spatial process, *Journal of Agricultural, Biological, and Environmental Statistics*, 2(4), 347-369.

Handcock, M., Stein, M., 1993, A Bayesian analysis of kriging, *Technometrics*, 35(4), 403-410.

Le, N., Zidek, J., 1992, Interpolation with uncertain spatial covariance: a Bayesian alternative to kriging, *Journal of Multivariate Analysis*, 43(375), 351-374.

R Development Core Team, 2011, R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.

Ribeiro Jr, P. J., Diggle, P. J., 1999, Bayesian inference in Gaussian model-based geostatistics, Technical Report ST-99-08, Department of Mathematics and Statistics, Lancaster University, Lancaster, UK.