

Hierarchical Bayesian 기법을 통한 강우-유출모형 매개변수의 최적화 및 불확실성 분석

Parameter Optimization and Uncertainty Analysis of the
Rainfall-Runoff Model Coupled with Hierarchical Bayesian Inference
Scheme

문 영 일*, 권 현 한**

Abstracts

정교한 강우-유출 모의를 위해서는 적절한 매개변수의 추정이 필수적이며, 매개변수 추정 방법은 시행착오(trial and error)에 의한 수동보정법과 최적화방법을 사용한 자동보정법으로 구분할 수 있다. 모형의 매개변수의 수가 많은 경우 수동보정법에 의한 매개변수 추정은 매우 어렵다. 자동보정법에 사용되는 최적화방법은 Rosenbrock 알고리즘, pattern search, 캠플렉스(complex) 방법, Powell 방법 등과 같은 지역최적화 방법과 전역최적화 방법으로 나눌 수 있다. 그러나 기존 방법론들은 매개변수의 최적화를 추적하기 위한 알고리즘이 대부분이며 이들 매개변수에 관련된 불확실성을 평가하는데는 미흡한 단점이 있다. 이러한 점에서 본 연구에서는 강우-유출모형의 매개변수 추정에 있어서 불확실성을 평가할 수 있는 새로운 방법론을 검토하고자 한다. 매개변수와 관련된 불확실성을 평가하기 위한 방법은 여러 가지가 있으나 통계적으로 매우 우수한 능력을 보이는 Hierarchical Bayesian 알고리즘을 Probability-Distributed 강우-유출 모형에 적용하였다. 본 방법론은 최적화와 동시에 각 매개변수에 관련된 사후분포(posterior distribution)의 추정이 가능하므로 모형이 갖는 불확실성을 효과적으로 평가할 수 있다. 따라서, 수자원 관리에 있어서 불확실성을 고려할 수 있으므로 보다 수리수문학적 위험도를 저감할 수 있을 것으로 판단된다.

Keyword: 강우-유출모형, Bayesian 모형, 최적화, 불확실성분석

1. 서 론

국내에서는 저수지의 유입량을 예측할 수 있는 다양한 방법론과 연구가 이루어지고 있다. 유출 모형의 기본 목적은, 유출현상을 예측하기 위한 것으로 치수목적으로 시간 단위의 단기 흥수사상 모형이, 이수 목적으로는 일 단위 이상의 장기 유출 모형이 주로 사용되고 있다. 노재경(1999)은 가지야마 공식에서 이용되는 유역 특성에 관계되는 계수(f)값을 추정하기 위해서 한국수자원공사에서 운영하는 18개 지점에서 다목적댐과 용수 전용댐의 27개년 유출 자료를 가지고 f 값의 최적화 방안을 검토하였다. 김운중 등(2002)은 Mizumura(1995)가 구성한 강우-유출 모형을 유역의 물리적 특성이 반영되는 장점을 유지하면서 미래에 도래할 강우에 적용하여 유출량을 예측할 수 있는 모형으로 개선하고 그 모형의 적합성 여부를 섬진강 유역에 적용하여 검토하

* 정회원·서울시립대학교 토목공학과 교수 E-mail : yoon@uos.ac.kr

* 정회원·Columbia University Associate Research Scientist E-mail : hk2273@columbia.edu

였다. 배덕효 등(2002)은 탱크 모형의 매개변수를 추정하는데 유역의 유출특성을 모의할 수 있는 방안을 연구하였다. 이와 같이 강우-유출 모형에서 유출자료에 근거한 검정 및 보정은 필수적인 과정이며 많은 연구가 수행되고 있다. 그러나, 이러한 매개변수를 추정하는데 있어서 존재하는 불확실성을 고려하는 연구는 미약하며 특히 국내의 경우 검정 및 보정을 위한 충분한 자료를 가지고 있지 않다는 점에서 불확실성 분석은 더욱 필요하다 하겠다. 따라서 본 연구에서는 Probability-Distributed 강우-유출 모형과 Hierarchical Bayesian 기법을 결합한 강우-유출 모형을 제시하고자 하며 이를 국내 괴산댐 유역에 적용하였다.

2. Hierarchical Bayesian 기법

수리적으로 계산이 불가능하거나 복잡한 적분, 추정 등의 문제에 사용되는 Monte Carlo 기법은 최근에 수리 수문학 분야에서 위험도 및 불확실성을 평가하는 수단으로 널리 이용되고 있다 (Kwon and Moon, 2005). Monte Carlo기법은 관심이 있는 값을 확률변수의 기대값으로 표현하고 이를 모의를 통하여 추출된 동일한 분포를 따르며 서로 독립(Independent and identical Distributed: iid)인 표본들의 표본평균을 이용하여 추정하는 방법이라고 할 수 있다.

이에 반해 Markov Chain Monte Carlo 기법은 주어진 다변량 확률분포가 복잡하여 이를 따르는 iid 난수를 얻을 수 없는 경우에 사용가능한 기법으로서 iid 난수 대신 Markov Chain 난수를 추출하여 사용한다. Markov Chain을 통해 난수를 발생시킨다고 해서 정확하게 관심이 되는 확률분포를 따르지 않지만 이를 일정 시간동안 반복 후에 얻어지는 난수들은 추출을 원하는 분포에 수렴하게 된다. 따라서 Markov Chain Monte Carlo기법은 복잡한 다변량 확률분포 및 매개변수의 추정을 요하는 문제에서 주로 사용되며 또한 Bayesian 통계 기법에서 사후분포의 추론의 이용될 수 있다. 본 연구에서는 2가지 관점에서 Markov Chain Monte Carlo 기법을 이용하게 된다. 즉 종속 변수 Y 에 대해서 조건부 분포를 갖는 각 종속변수의 사후분포를 추정하게 된다.

3. Conceptual 강우-유출 모형

본 연구에서는 Moore(1985)가 제시한 Probability-Distributed 강우-유출 개념을 기본으로 하는 모형을 이용하였다. 이 모형의 일반적인 개념은 첫째로, 유역의 한 지점에서의 유출량을 정의하고 둘째로, 전체 유역의 매개변수들의 공간적 거동과 유역으로부터 통합된 유출반응을 모의하기 위해서 확률분포가 도입된다. 본 모형은 상대적으로 간단하며 강우과정 모델과 결합된 2개의 선형 저수지 모형 즉 3개의 병렬저수지를 갖는 Quick Flow와 하나의 저수지를 갖는 Low Flow로 구성된다. 따라서 관측 유출량에 대해서 5개의 매개변수를 최적화하는 것이 필요하며 즉, 최대저장용량, 토양함수량의 공간적 변동성의 정도, 2개의 저수지를 구분하는 Fraction Factor, Quick Flow와 Low Flow의 저류시간 등이 있다.

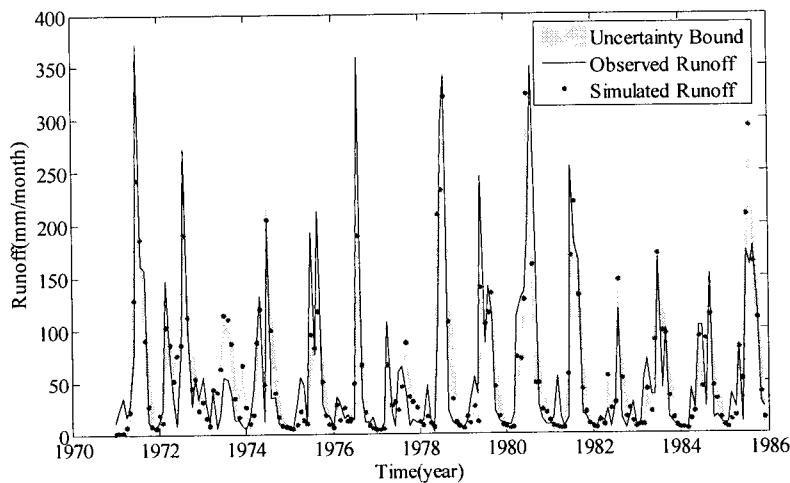
$$F(S_C) = 1 - \left(1 - \frac{S_C}{S_{C_{\max}}}\right)^{S_m} \quad (1)$$

여기서 F 는 유역의 한 지점에서의 어떤 저류량 S_C 가 발생 가능할 누가확률을, $S_{C_{\max}}$ 는 최대저장용량을 S_m 는 토양함수량의 공간적 변동성 정도를 나타내는 매개변수이다.

3. 적용 및 결론

본 연구에서는 괴산댐 유역의 월 단위 강수량, 유출량 및 증발량 자료를 대상으로 모형의 적합성을 평가하였다. 괴산댐은 1957년에 완공되어 사용 연수가 46년이나 경과되어 경제적 수명을 다하고 있으며 부분적으로 개보수가 필요한 실정이다. 괴산댐은 유역면적 671 km^2 , 댐 높이와 길이가 각각 28m, 17m 이고 유효저수용량 약 5.7백만 m^3 이며, 콘크리트 중력식 댐이다. 괴산댐 지역을 대표할 수 있는 기상관측소는 미원, 청천, 입석, 속리산, 복천암 지점의 우량관측소가 위치하고 있다. 이 중에서 30년 이상의 자료가 있는 청천 지역의 강수량 자료를 이용하여 유입량을 산정한다. 앞서 언급했듯이 강우-유출 모형의 5개 매개변수에 대해서 Hierarchical Bayesian 모형을 적용하여 매개변수를 최적화하고 이에 대한 불확실성을 평가하였다.

a)



b)

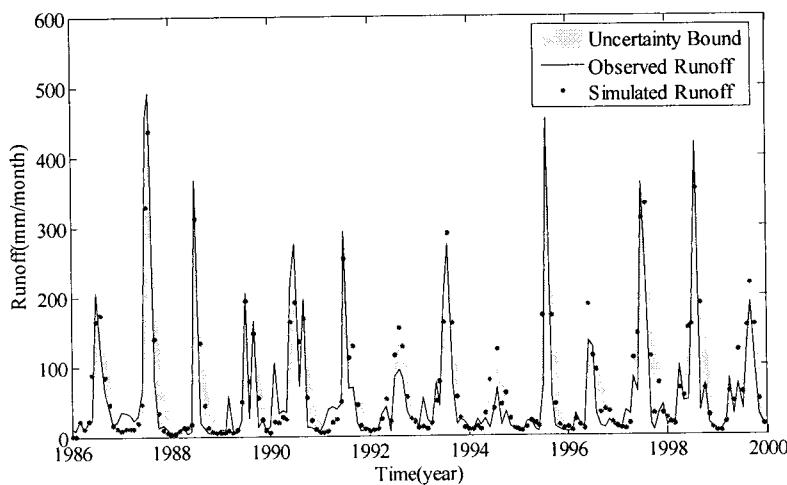


그림 1. 괴산유역 실측 유출량과 모형 유출량과의 비교 a) 검정 구간 b) 검증 구간

그림 1은 모형을 추정된 유출량과 실측 유출량과의 비교를 나타낸다. 모형의 검정과정에서는 실측자료를 이용하여 모형의 매개변수를 추정하였으며 이를 토대로 추정된 불확실성과 예측치를

그림 1-a에 나타내었다. 그림 1-b에서는 앞서 추정된 모형을 토대로 강우량과 증발량을 입력자료로 하여 유출량을 예측하였다. 그림에서 보는 바와 같이 저유량과 첨두유량에 대해서 실측치와 가깝게 모의되고 있으며 검증구간 보다 개선된 결과를 나타내 주고 있다. 검정 및 검증구간에서 모형의 통계치를 추정하여 나타내면 표 1과 같다. 검정구간에서 상관계수는 대략 0.83을 나타내고 있고 전체적으로 실측치와 유사한 거동을 보이고 있으며 상대적으로 과소 추정되고 있다. 반면 검증구간에서는 0.94의 상관관계를 나타내고 있으며 평균 및 표준편차 또한 매우 유사한 거동을 보이고 있다.

표 1. 검정 및 검증구간에서 실측치와 예측치의 통계치 비교

	R	CoE	IoA	RMSE	Mean (Observed)	Mean (Predicted)	Std (Observed)	Std (Predicted)
검정기간	0.83	0.67	0.90	43.13	56.03	47.95	75.74	63.54
검증기간	0.94	0.89	0.97	30.73	58.15	60.07	91.11	89.35

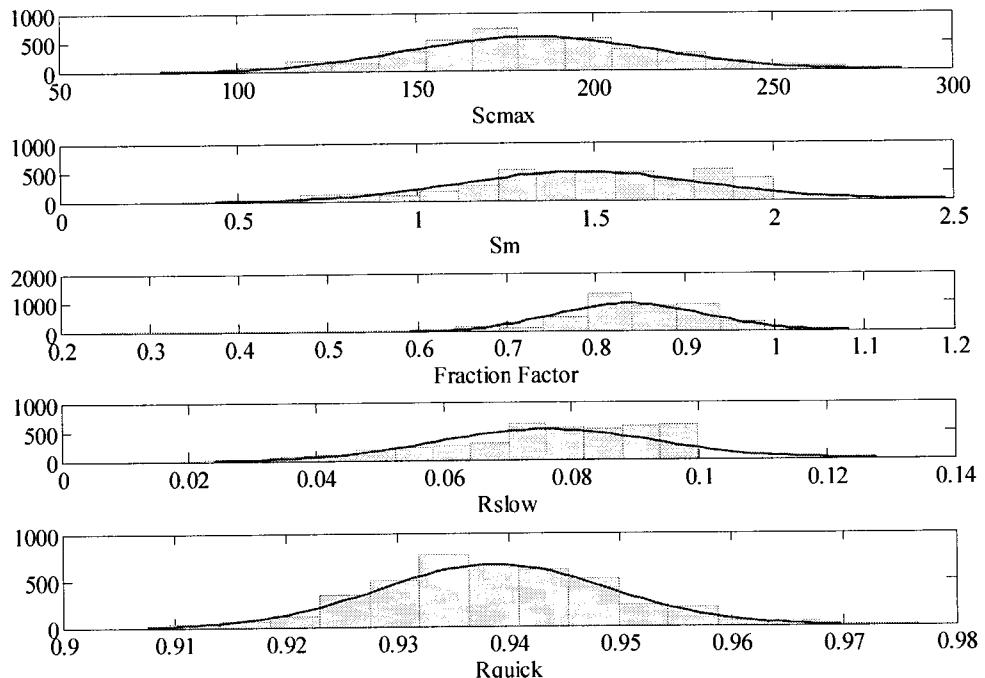


그림 2는 각 매개변수의 사후분포를 나타내며 상대적으로 작은 분산을 가지고 수렴하고 있음을 볼 수 있다. 이를 통해 추정된 모형의 적합성을 다시 한 번 확인 할 수 있으며 강우-유출관계를 모의하는데 있어서 전체적인 모형의 불확실성을 효과적으로 파악할 수 있으므로 상대적으로 수자원을 관리하는데 위험도를 감소시키고 반대로 신뢰성을 증대시키는 효과를 기대할 수 있겠다.

참 고 문 헌

1. Moore, R.J., 1985. The probability-distributed principle and runoff production at point and basin scale. *Hydrological Sciences Journal* 30(2), 273-297.
2. 김운중, 김민환, 전일권 (2002). “감수곡선을 이용한 탱크 모형과 매개변수 자동보정에 의한 유출 예측”, 대한토목학회논문집, 대한토목학회, 제22권, 제6-B호, pp. 777-784.
3. 노재경 (1999). “가지야마 월 유출량 공식의 일반화”, 한국수자원학회 학술발표회 논문집, 한국수자원학회, pp. 221-226.
- 4 배덕효, 정일원, 강태호 (2002). “유역 유출특성을 고려한 매개변수 추정에 관한 연구”, 대한 토목학회 학술발표회 논문집, 대한토목학회, pp. 38-41.