Hierarchical Bayesian 방법을 이용한 수위-유량 관계 곡선 유도 및 불확실성 분석

Derivation and Uncertainty Analysis of Rating Curve Using Hierarchical Bayesian Model

권 현 한*, 문 영 일**, 최 병 규***, 김 민 석****

Hyun Han Kwon, Young Il Moon, Byung-Kyu Choi, Seok Min Kim

요 지

정확한 유량산정은 수자원 연구에서 기초가 되는 자료를 생산한다는 관점에서 홍수 및 가뭄관리에서 매우 중요한 부분이라 할 수 있다. 국내에서 유량측정의 정확성을 높이고자 진보된 계측기의 개발 및 분석 방법에 관한 연구가 꾸준히 진행되고 있다. 일반적으로 유량을 추정하기 위해서 특정단면에서의 수위를 측정하여 이를 수위-유량 관계곡선을 통해서 유량으로 환산하게 된다. 즉 수위-유량 관계를 측정한 후 이를 회귀분석 방법으로 내삽 및 외삽을 실시하여 유량을 추정하게 된다. 그러나 수위-유량 관계곡선에서 저수위와 고수위를 하나의 곡선식으로 하게 되는 경우 정도가 낮아지게 되므로 많은 경우에 있어서 저수위, 고수위를 각각의 곡선으로 구하여 사용하고 있다. 이러한 경우 정량적으로 변곡점을 구하기보다는 경험적으로 저수위와 고수위를 구분하고 있으며, 수위-유량 관계를 회귀식에 의해서 추정하게 되므로 이에 대한 불확실성 또한 정량화할 필요가 있다. 이러한 관점에서 본 연구에서는 불확실성 분석과 함께, 저수위-고수위를 정량적으로 구분할 수 있는 Hierarchical Bayesian 방법을 도입하여 수위-유량곡선식을 유도하고자 한다.

핵심용어 : 수위-유량 관계곡선, Hierarchical Bayesian Model, Uncertainty

1. 서 론

수위-유량 관계곡선은 매년 한강을 비롯한 중요하천의 홍수통제소, 댐 관리를 목적으로 개발하고 있는 수자원공사, 농업 용수등 기타 수자원 이용을 목적으로 하는 농어촌 진흥공사 등이 개발하고 있으며 동일 지점이라 하더라도 개발 주체에 따라 조금씩 차이가 있다. 최근 들어 수위-유량 곡선에 대한 신뢰도의 문제에 대한 연구가 이루어지고 있으나, 기존 개발방식에서 크게 변화하지는 못하고 있다.

수위-유량 관계곡선이 수자원분야의 가장 기초적이며 중요한 자료임을 비춰볼 때 이에 대한 최근의 연구는 그 중요도에 비하여 적었다고 할 수 있으며, 다만 기초자료인 유량자료의 신뢰성을 분석하거나, 하천의 수리학적 특성의 반영 등을 목적으로 연구가 진행되고 있다.

앞에서도 언급한 것과 같이 1994년 건설교통부가 수위-유량 관계곡선에 대하여 정리하고, 개선안을 제시, 중요 하천에 대한 수위-유량 곡선식을 정한바있다. 이러한 연구나 개발은 기존의 곡선식 개발의 보완이나 신뢰도 향상에 목적을 두는 것으로 수위-유량곡선식 개발 방법에 대한 고찰

* 정회원·한국건설기술연구원 수자원연구실 선임연구원·공학박사Email: hkwon@kict.re.kr

** 정회원·서울시립대학교 토목공학과 교수·Email: ymoon@uos.ac.kr

*** 정회원·(주) 삼안건설기술공사 수력부 부사장·공학박사Email: bkchoi@samaneng.com

**** 정회원·서울시립대학교 토목공학과 석사과정 sharpkms@nate.com

이나 연구는 적다고 할 수 있다.

또한 수위-유량 관계곡선에 대한 연구는 주로 수리학적 특성의 반영, 즉 유수에 대한 하천의 하도 통제기능과 단면 통제기능 등에 대한 연구가 있으며, 이러한 연구 역시 기존 곡선식의 보완이나 신뢰도 향상에 목적을 두고 있는 것이다.

2. 수위-유량 곡선식

지수형과 포물선형 중에서 상대적 우수성은 없으며 분석의 편의성에 따라 선택하여 쓰게 된다. 그러나 영유량의 처리과정 가지는 번잡성을 해결할 수 있다면 수리적인 의미를 가지는 지수형에 의한 측정보다 포물선형이 상대적으로 보편화 되어있다.

현재 사용 중인 수위-유량 곡선식의 형태는 크게 USGS, USBR, WMO (World Meteorological Organization)에서 작성된 기준에 기본을 두고있으며, WMO에서 발간된 Homes Manual에 언급된 수위-유량 곡선식의 일반적 형태는 아래의 2가지 형태가 주류를 이룬다.

$$Q = a(H - H_0)^b \tag{1}$$

$$Q = aH^2 + bH + c \tag{2}$$

여기에서 a ,b ,c 및 H_0 는 매개변수 이다. 식(1)의 경우 하천 수위와 유량의 물리적 개념을 표현해주는 식으로서 위에 언급된 일반화된 유량곡선식의 형태를 반영해 주고 있다. 그러나 실제로는 매개변수 b 에 대한 선정이 어렵고, 갈수유량의 처리가 어려운 단점이 있다. 반면에 식(2)의 경우 영유량의 처리 또는 갈수량의 처리에 적합한 표현으로서, 갈수량의 처리가 손쉬워지는 장점이 있다. 또한 매개변수에 대한 선정이 간편하여 사실상 운용이 쉬워진다.

3. 마코프 체인 Monte Carlo 모의

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, 2006). Monte Carlo기법은 관심이 있는 값을 확률변수의 기댓값으로 포현하고 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 마코프 체인 Monte Carlo 기법은 주어진 다변량 확률분포가 복잡하여 이를 따르는 iid 난수를 얻을 수 없는 경우에 사용가능한 기법으로서 iid 난수 대신 마코프 체인 난수를 추출하여 사용한다. 마코프 체인을 통해 난수를 발생시킨다고 해서 정확하게 관심이 되는 확률분포를 따르지 않지만 이를 일정 시간동안 반복 후에 얻어지는 난수들은 추출을 원하는 분포에 수렴하게 된다. 따라서 마코프 체인 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 마코프 체인 Monte Carlo 기법을 이용하게 된다. 즉 종속변수 Y 에 대해서 조건부 분포를 갖는 각 종속변수의 사후분포를 추정하게 된다.

마코프 체인 Monte Carlo기법의 대표적인 방법으로 깁스표본법(Gibs Sampling)이 있으며 이를 본 연구에서 이용하였다. 깁스표본법은 원하는 다변량 확률분포에서 iid 표본을 추출하는 것이 복잡하거나 난해한 경우 이용 가능한 방법으로서 3개의 변수를 갖는 다변량 확률분포를 이용하여 설명하면 다음과 같다. 3개의 변수를 갖는 다변량 확률밀도함수를 $f(x,y,z)$라고 하자. 깁스표본법

은 확률밀도함수로부터 직접 표본을 추출할 수는 없으나 각각의 변수들의 대해서 다른 두 변수들이 주어졌을 때의 조건부 분포가 알려져 있고 이로부터의 표본추출이 가능한 경우에 사용할 수 있다. 알고리즘을 간단히 정리하면 다음과 같다.

[1] 세변수에 대한 초기 값$(x^{(0)}, y^{(0)}, z^{(0)})$을 부여한다.

[2] i번째 난수 벡터 $(x^{(i)}, y^{(i)}, z^{(i)})$가 주어졌을 때 $i+1$번째 난수 벡터를 다음과 같은 조건부 분포에서 추출한다.

(1) $x^{i+1} \sim f(x|y^{(i)}, z^{(i)})$

(2) $y^{i+1} \sim f(y|x^{(i+1)}, z^{(i)})$

(3) $z^{i+1} \sim f(z|x^{(i+1)}, y^{(i+1)})$

[3] 위의 과정을 충분히 반복한 후 초기의 일정부분 난수를 제거한 이후의 난수들을 이용한다.

위의 알고리즘에서 보듯이 깁스표본법은 조건부 분포들에서 조건으로 주어지는 변수들의 값은 정확하게 바로 직전의 단계에서 주어진 값들이 사용되게 되며 따라서 조건부 분포에서 추출된 난수들이 안정 상태에 도달하는 것이 주어진 다변량 확률분포를 정확히 따르는 난수가 되는 척도가 되며 깁스표본법을 구현하는데 가장 중요한 부분이 된다.

4. 수위 유량 곡선 유도

본 연구에서는 2000년-2006년 까지 우이천 장월교 지점의 수위 유량자료를 이용하였다. 총 117개의 사상을 대상으로 분석을 실시하였으며 앞서 제시한 식 (1)과 식(2)를 이용하였다. Figure 1a는 식 (1)을 이용하여 단일 곡선을 추정한 결과이며 회색으로 된 구간은 모형으로부터 추정된 불확실성 구간을 나타난다. Figure 1b는 식 (2)를 이용하였으며 저수위 고수위를 나누는 기준 값을 하나의 변수로 가정하여 Bayesian 분석으로부터 추정하여 Rating Curve를 추정하였다.

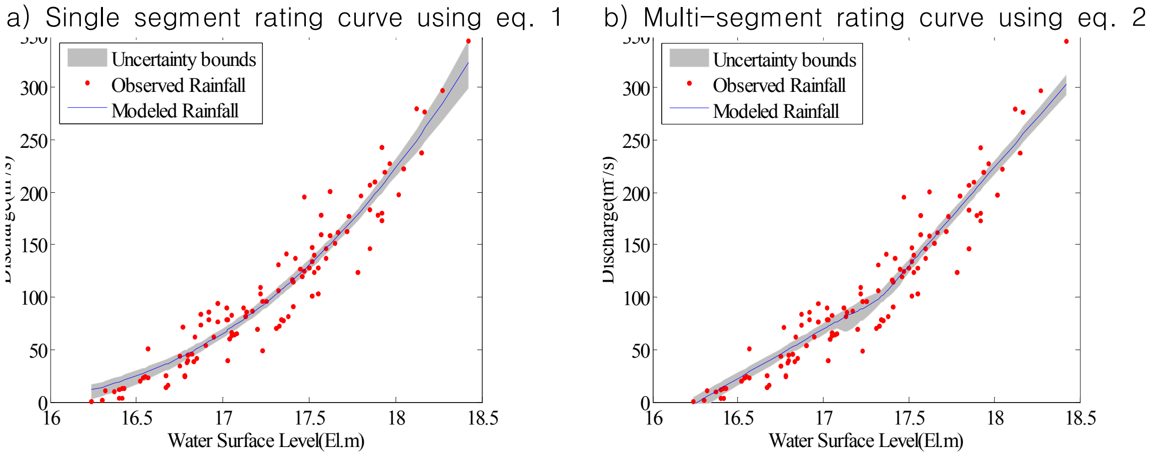

Figure 1. Rating curve derivation using Hierarchical Bayesian Model

각 모형별 매개변수들에 대한 불확실성 구간 및 중간값을 Table 1에 나타내었다. 모형 2에서 고수위와 저수위를 나누는 기준 값으로서 17.3 El.m이 추정되었으며 추정된 분산을 고려하면 통계적으로 유의한 값으로 추정되었다.

Table 1. Credible interval of the parameters for two models

node	mean	sd	2.50%	median	97.50%
alpha[1]	27.0	11.6	10.9	29.5	48.2
alpha[2]	15.5	0.2	15.1	15.6	15.9
alpha[3]	2.4	0.3	2.0	2.3	2.9
node	mean	sd	2.50%	median	97.50%
Inflection Point	17.3	0.1	17.1	17.3	17.5
alpha[1]	−1842.0	98.9	−1977.0	−1856.0	−1621.0
alpha[2]	40.7	8.3	26.1	43.6	50.2
alpha[3]	4.1	0.2	3.8	4.1	4.6
alpha[4]	−1149.0	127.7	−1328.0	−1174.0	−952.2
alpha[5]	47.7	11.0	27.8	53.4	60.9
alpha[6]	1.4	0.3	1.0	1.4	1.9

Table 2. Fitting results for the models

	Model 1	Model 2
R	0.96	0.96
R^2	0.91	0.91
N-S	0.98	0.98
RMSE	20.78	20.72

5. 결 론

수위-유량 관계곡선에서 저수위와 고수위를 하나의 곡선식으로 하게 되는 경우 정도가 낮아지게 되므로 많은 경우에 있어서 저수위, 고수위를 각각의 곡선으로 구하여 사용하고 있다. 이러한 경우 정량적으로 변곡점을 구하기보다는 경험적으로 저수위와 고수위를 구분하고 있으며, 수위-유량 관계를 회귀식에 의해서 추정하게 되므로 이에 대한 불확실성 또한 정량화할 필요가 있다. 이러한 관점에서 본 연구에서는 불확실성 분석과 함께, 저수위-고수위를 정량적으로 구분할 수 있는 Hierarchical Bayesian 방법을 개발하였다.

본 방법론을 우이천 유역 장월교 지점 수위-유량 자료를 대상으로 검토한 결과 기존방법과 유사한 결과를 얻을 뿐만 아니라, 정량적으로 저수위, 고수위를 구분할 수 있었다. 이와 함께 각 매개변수들에 대한 불확실성을 파악할 수 있으므로 계측자료의 불확실성과 모형의 불확실성을 효과적으로 구분이 가능하다.

감사의 글

본 연구는 한국건설기술연구원 『Seed Money 사업』 에 의해 지원되었습니다.

참고문헌

Kwon, H-H and Y.-I. Moon, Improvement of Overtopping Risk Evaluations Using Probabilistic Concepts for Existing Dams, Stochastic Environmental Research and Risk Assessment, Vol. 20(4), 2006.