韓國水資源學會論文集
第41卷 第1號・2008年 1月
pp. 49~63

DOI: 10.3741/JKWRA.2008.41.1.049

Bayesian MCMC를 이용한 저수량 점 빈도분석: Ⅱ. 적용과 비교분석

At-site Low Flow Frequency Analysis Using Bayesian MCMC: Ⅱ. Application and Comparative Studies

김 상 욱* / 이 길 성**
Kim, Sang Ug / Lee, Kil Seong

Abstract

The Bayesian MCMC(Bayesian Markov Chain Monte Carlo) and the MLE(Maximum Likelihood Estimation) methods using a quadratic approximation are applied to perform the at-site low flow frequency analysis at the 4 stage stations (Nakdong, Waegwan, Goryeonggyo, and Jindong). Using the results of two types of the estimation method, the frequency curves including uncertainty are plotted. Eight case studies using the synthetic flow data with a sample size of 100, generated from 2-parmeter Weibull distribution are performed to compare with the results of analysis using the MLE and the Bayesian MCMC. The Bayesian MCMC and the MLE are applied to 36 years of gauged data to validate the efficiency of the developed scheme. These examples illustrate the advantages of the Bayesian MCMC and the limitations of the MLE based on a quadratic approximation. From the point of view of uncertainty analysis, the Bayesian MCMC is more effective than the MLE using a quadratic approximation when the sample size is small. In particular, the Bayesian MCMC is a more attractive method than MLE based on a quadratic approximation because the sample size of low flow at the site of interest is mostly not enough to perform the low flow frequency analysis.

keywords : Uncertainty, Bayesian MCMC, MLE using a quadratic approximation, 2-paramter Weibull distribution, synthetic flow data, gauged data, frequency curve

요 지

본 연구에서는 Bayesian MCMC 방법과 2차 근사식을 이용한 최우추정(Maximum Likelihood Estimation, MLE) 방법 방법을 이용하여 낙동강 유역의 본류지점인 낙동, 왜관, 고령교, 진동지점에 대한 점 빈도분석을 수행하고 그 결과로써 불확실성을 포함한 빈도곡선을 작성하였다. 통계적 실험을 통한 두 가지 추정방법의 분석을 위하여 먼저 자료의 길이가 100인 8개의 합성 유량자료 셋을 생성하여 비교 연구를 수행하였으며, 이를 자료길이 36인 실측 유량자료의 추정결과와 비교하였다.

Bayesian MCMC 방법에 의한 평균값과 2차 근사식을 이용한 취우추정방법에 의한 모드에서의 2모수 Weibull 분포의 모수 추정값은 비슷한 결과를 보였으나, 불확실성을 나타내는 하한값과 상한값의 차이는 Bayesian MCMC 방

* 서울대학교 BK21 안전하고 지속가능한 사회기반건설 사업단 박사 후 연구원
Post-Doctor, Seoul National University BK21 SIR Group, Seoul National University, Seoul, 151-744, Korea
(e-mail: plethor1@snu.ac.kr)
** 서울대학교 공과대학 건설・환경공학부 교수
Professor, Dept. of Civil and Environmental Engineering, Seoul National University, Seoul, 151-744, Korea
(e-mail: kilselee@snu.ac.kr)

법이 2차 근사식을 이용한 취우추정방법보다 불확실성을 감소시켜 나타내는 것을 알 수 있었다. 또한 실측 유량자료를 이용한 결과, 2차 근사식을 이용한 취우추정방법의 경우 자료의 길이가 감소됨에 따라 불확실성의 범위가 합성 유량자료를 사용한 경우에 비해 상대적으로 증가되지만, Bayesian MCMC 방법의 경우에는 자료의 길이에 대한 영향이 거의 없다는 결론을 얻을 수 있었다. 그러므로 저수량 빈도분석을 수행하기 위해 충분한 자료를 확보할 수 없는 국내의 상황을 감안할 때, 위와 같은 결론으로부터 Bayesian MCMC 방법이 불확실성을 표현하는데 있어서 2차 근사식을 이용한 최우추정방법에 비해 합리적일 수 있다는 결론을 얻을 수 있었다.

핵심용어 : 불확실성, Bayesian MCMC, 2차 근사를 이용한 최우추정방법, 2모수 Weibull 확률분포, 합성 유량자료, 실측 유량자료, 빈도곡선

1. 서 론

WMO(1974)는 저수(low flow)를 '가뭄이 지속되는 경우 하천에 흐르는 유량'으로 정의한 바 있다. 그러나 저수는 가뭄이 지속되는 경우가 아닌 경우에도 인간의 활동을 위해 필요한 생활·공업·농업용수의 취수나 지하수량 증가에 의해 발생될 수 있으므로, 위의 개념으로 '저수'를 정의하기에는 부족한 면이 많다. Smakhtin (2001)은 '어떤 하천의 주변적인 요인이나 인간의 활동 그리고 가뭄의 지속적으로 인해 어떤 지점의 유량이 확보될 수 없는 흐름의 상황'으로 저수를 정의한 바 있으며, 이는 Dracup et al.(1980)이 제안한 '수문학적 가뭄(hydrological droughts)'의 개념을 일부분 차용한 것이라 할 수 있다.

정량적으로 저수량을 분석하기 위해서 여러 가지 지표들이 사용될 수 있는데, 대표적인 것들이 MAR(Mean Annual Runoff), MDF(Mean Daily Flow), MF(Median Flow), AMF(Absolute Minimum Flow) 등이다. 또한 위와 같은 개략적인 지표들 외에도 유황곡선의 작성이나 빈도분석의 결과를 이용해서도 여러 가지 저수량을 분석할 수 있는 결과를 얻을 수 있다. 현재까지 저수분석(low flow analysis)에서 가장 많이 사용되는 값으로 미국에서는 빈도분석을 이용한 7일 지속기간 10년 빈도유량(7Q10)이고 국내에서는 355위 유량의 10년 빈도에 해당하는 유량(기준갈수량)이라 할 수 있다. 두 값을 산정하기 위해서는 빈도분석을 수행해야 하는데, 빈도분석은 주어진 자료를 나타낼 수 있는 확률분포함수의 선정, 모수의 추정, 추정된 모수를 이용한 재현기간별 분위수(quantile) 산정의 절차를 통하여 통상적으로 산정된다. 즉 홍수량이나 저수량의 발생빈도에 해당되는 유량을 통계적으로 분석함으로써 필요한 재현기간에 해당되는 유량을 얻을 수 있는 방법이라 할 수 있다.

저수량 빈도분석에 자주 사용되어진 바 있는 확률분포는 Weibull, Gumbel, Pearson Type Ⅲ 그리고 Log-normal 분포로써 많은 선행 연구자들의 연구에서 찾아 볼 수 있으며, 이와 관련된 연구는 주로 주어진 자료가 어떤 분포에 가장 잘 적합 되는지를 검토하는 데 초점이 맞추어져 있다고 할 수 있다(Matalas and Benson, 1968; Burges et al., 1975; Boughton, 1976; Rao, 1980; Boes et al., 1989; Nathan and McMahon, 1990).

확률분포함수가 결정되면 확률분포함수에 포함되어져 있는 모수가 자료를 올바르게 나타내도록 추정되어져야 하는데, 이를 위해 모수적 추정방법으로는 적률법(method of moment), 확률가중적률법(probability weighted moment method), L-moment 방법, 최우추정(maximum likelihood estimation)방법 등이 주로 사용되며 다양한 비모수적 방법도 적용될 수 있다. 모수의 추정이 완료되면 주어진 재현기간에 대한 분위수를 추정하게 되며, 추정된 분위수에 대한 신뢰구간을 산정하고 재현기간에 따른 빈도곡선을 작성함으로써 원하는 재현기간에 따른 저수량 또는 홍수량을 구하여 사용할 수 있게 된다.

그러나 국내 대부분의 주요 중장기 계획은 빈도분석의 확정적인(deterministic) 값만을 이용하여 수립되고 있으며, 빈도분석결과의 불확실성을 반영한 확률적인(probabilistic) 값이 이용되는 계획은 찾아보기 힘들다. 즉 대부분의 연구는 확률분포의 선정과 그에 따른 모수의 추정에만 초점이 맞추어져 있고 추정된 모수의 불확실성에 따르는 빈도유량을 산정하는 연구는 많지 않은데, 이는 불확실성에 대한 인식 부족과 함께 불확실성에 대한 계산방법이 현실을 제대로 반영하지 못함으로써 빈도분석 결과의 정확성에 대한 신뢰도가 낮은 것에 기인한다. 특히 본 연구의 Ⅰ편의 연구동향에서 언급한 바와 같이 빈도유량의 불확실성에 대한 연구는 홍수량을 대상으로 하는 경우에는 진행된 사례가 있으나 저수량 빈도분석에 있어서 불확실성을 고려한 연구는 찾아보기 힘들다. 그러므로 본 연구에서는 불확실성을 계산

하는데 있어서 기존에 사용된 정상성(normality), 선형성(linearity) 등의 가정조건을 사용하지 않는 Bayesian 방법론을 사용하여 저수량 빈도분석을 수행함으로써 기존 방법과 Bayesian 방법을 비교평가하고 개선점을 제시하고자 하였다.

본 연구는 I 편에서 제시된 이론적 배경과 Bayesian MCMC 방법을 적용하기 위해 제시된 여러 가지 요소들을 사용하여 실제적으로 낙동강 유역에 대한 저수량 점 빈도분석을 수행하기 위한 빈도분석의 과정을 순서대로 진행하였다. 특히 저수량 빈도분석을 위한 자료의 선정과 한계성을 먼저 분석하였으며 합성 유량자료를 생성하여 통계적 실험을 수행함으로써 Bayesian MCMC 방법과 그에 대한 비교 대안으로 2차 근사식을 이용한 최우추정방법을 비교 분석하였다. 또한 최종적으로 얻어진 두 가지 추정방법을 이용한 모수 추정값을 이용하여 불확실성을 포함한 빈도곡선을 재현기간별로 산정하여 빈도곡선을 작성하였다.

2. 대상유역의 선정 및 자료의 선정

본 절에서는 저수량 점 빈도분석을 수행하기 위한 대상유역과 실측 유량자료의 저수특성 및 한계점, 그리고 Bayesian MCMC 방법과 이차근사식을 이용한 최우추정방법의 비교분석을 위한 통계실험을 수행하기 위한 모의자료의 발생 과정을 서술하였다.

2.1 대상유역 및 실측 유량자료의 선정

본 연구의 수행을 위하여 선정된 유역은 낙동강 유역(Fig. 1)으로서 유역면적은 23,702 km^2 으로 수자원단위지도상에서 중권역 22개와 표준유역 191개로 구성되어져 있다. 본 연구는 저수빈도분석에서 나타나는 불확실성의 감소에 초점이 맞추어진 연구로서 향후 적용연구로 불확실성을 포함한 빈도유량을 사용한 수질모형의 구축, 오염총량제와 관련되어 국내에서 사용하고 있는 저수량(Q_{275})의 불확실성 분석 그리고 취수허가 시 필요한 기준유량의 결정 등의 연구가 진행될 수 있으므로 국내에서 몇 번의 수질사고의 경험이 있고, 오염총량제의 적용이 가장 많이 진행되어 있는 낙동강유역을 대상유역으로 선정하였다.

본 연구에서 사용된 낙동강 유역의 유량자료는 수자원단위지도를 기준으로 하여 그림 1과 같이 본류상의 4개의 수위관측지점인 낙동, 왜관, 고령교, 진동지점의 실측 유량자료와 10개의 중권역의 유역 유출자료를 이용하였다. 단, 8번과 10번 소유역은 이후 타 연구에서 진행될 지역빈도분석 연구와의 일관성 유지를 위하여 두 개의 중권역을 하나의 소유역으로 구분하여 사용하였다. 점 빈도분석을 수행하기 위해서는 일반적으로 자료의 개수가 30개 이상이 되어야 함을 추천하고 있는데(Rao and Hamed, 2000), 본 연구에서 사용되어지는 각 유역의 말단에 위치되어 있는 대부분의 수위관측소의 경우에는 자료의 기간이 약 20년 미만으로 점 빈도분석이 직접 수행되기에는 부족한 면이 있으므로 각 소유역별 저수량 빈도분석을 위해서는 점 빈도분석 외에 지역빈도분석을 수행하는 것이 타당할 수 있다. 그러므로 본 연구에서는 본류상에 위치한 낙동, 왜관, 고령교, 진동지점만을 대상으로 하여 1966년부터 2001년까지의 36개년에 해당하는 일 유량자료를 수집하고 점 빈도분석을 수행하였다.

저수량 빈도분석을 수행하고자 하는 경우 가장 먼저 산정해야 하는 것은 자연유량(natural flow)이다. 자연유량이란, 유역 내 인위적인 유량이 조절이 없는 자연상태의 유역으로부터 발생되는 유역의 고유한 유량으로써 실측 유량자료를 자연유량으로 환산하기 위해서는 유역 내 존재하는 댐으로 인한 조절량(유입량, 방류량, 수면 증발량 등), 실제 취수량과 같은 인위적 조절 유량의 일자료가 필요하다. 그러나 국내 축적된 자료를 이용하여 임의 지점에서 자연유량을 산정하는 것은 필요자료의 부족으로 인하여 올바른 자연유량이 산정되기 매우 어려운 실정이다.

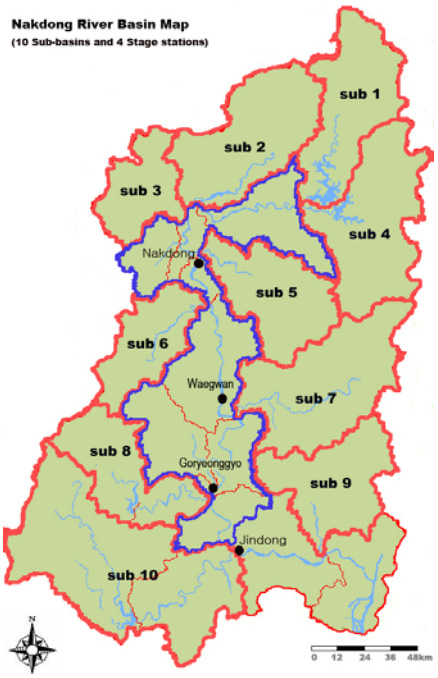


Fig. 1. Nakdong River Basin Map

이와 같은 자료의 부족 및 부정확성에 대한 문제는 저수량 빈도분석과 그 결과 값을 이용한 유지유량 또는

환경유량 등의 설정에 매우 큰 영향을 미칠 수 있으므로 시급하게 해결되어져야 하는 문제라고 할 수 있다. 일반적으로 취수 허가량은 실제 취수량보다 큰 값을 가지는 것으로 추정되고 있으며, 특히 농업용수의 경우 계절별로 허가량과 실제 취수량과의 차이가 매우 크므로 자연유량 산정 시에는 필히 실제 취수량을 사용할 필요가 있다. 현재 건교부에서는 이를 위하여 '하천유수사용 실적관리시스템'을 지속적으로 운영하고 있으며, 조만간 이 시스템을 이용하여 얻어진 실제 취수량을 자연유량의 산정에 이용할 수 있을 것으로 판단된다.

위와 같은 자료의 한계성을 바탕으로 본 연구에서는 실측된 수위로부터 환산된 유량자료를 자연유량으로 탈조절화(deregulation)하기 위하여 각 관측소간의 취수허가량 자료만을 사용하였고, 취수 허가량 자료만을 사용하여 환산된 유량을 4개 수위관측지점에서의 자연유량으로 가정하여 빈도분석을 수행하였다. 향후 보다 합리적인 자연유량의 산정결과가 얻어질 수 있다면, 본 연구에서 사용된 자료를 대체하여 수행함으로써 합리적인 자연유량의 사용이 빈도분석의 최종결과에 얼마나 큰 영향을 미치는지도 살펴볼 수 있으리라 판단된다.

또한 본 연구의 수행을 위해서는 본 연구의 Ⅰ편에서 수행한 바와 같이 Bayesian 방법에 큰 영향을 미치는 자료에 기반한 사전분포를 구축하기 위한 부가적인 유량자료가 필요한데, 이러한 부가적인 유량자료를 위해서 Fig. 1의 낙동강 유역의 10개 중권역의 출구 유출량을 사용하였다. 이 유량 역시 실측 유량자료를 이용하여 자연유량으로 환산되어져야 하지만, 지류지점의 수위관측소는 자연유량으로의 환산에 있어서 같은 어려움이 있고 수위기록의 기간이 짧아 사전분포를 구축하는 데 있어서 비합리적인 값을 산정하도록 할 수 있다. 이러한 문제를 해결하기 위해서 건교부와 한국수자원공사(2006)는 낙동강 유역의 유역조사사업을 수행하면서 낙동강 유역 각 중권역의 36년간 자연유량 자료를 PRMS(Precipitation-Runoff Modeling System)모형을 이용하여 모의한 바 있다. 모형을 이용한 자연유량 자료는 강우자료와 유역특성을 이용하여 원하는 기간만큼의 유량자료를 수문학적 모형을 통하여 모의할 수 있다는 장점이 있어 자연유량을 산정하기 힘든 국내 실정에 대한 대안으로서 사용될 수 있는 방법이라 할 수 있다. 그러나 수문학적 모형이 가지는 모형의 매개변수는 유역이 자연유량을 모의할 수 있도록 보정될 필요가 있으므로, 이를 위해 건교부와 한국수자원공사(2006)는 유역내에 댐이 없고 취수시설이 가장 적은 안동댐유역, 임하댐 유역, 합천댐 유역과 부족한 대상유역을 보충하기 위하여 한강유역의 도암댐 유역과 괴산댐 유역을 대상으로 PRMS 모형의 매개변수를 자연유량이 잘 모의되도록 보정한 후, 보정된 매개변수를 낙동강 유역으로 전이하여 사용함으로써 낙동강 유역의 중권역별 자연유량을 모의한 바 있다. 본 연구에서는 자료에 기반한 사전분포의 구축을 위하여 중권역 자연유량의 자료를 별도로 모의하지 않고, 위와 같은 과정을 통해 모의된 낙동강 유역의 중권역별 자연유량 자료를 이용하여 연구를 수행하였다. 단 자료의 선정에 있어서 최근까지의 자료를 사용하면 보다 최근의 저수유출특성을 분석할 수 있지만, 본 연구에서 사용한 10개 소유역에서의 자연유량자료는 본 연구를 통하여 모의된 자료가 아니라 위에서 언급한 바와 같이 건교부와 한국수자원공사(2006)에서 PRMS모형을 사용하여 모의한 1966년부터 2001년까지의 36년간의 자료이므로 지점 빈도분석도 이와 같은 기간의 자료를 사용하여 모의하였다. 그러므로 향후에는 최근까지의 자료를 이용하여 10개 소유역에서의 모의자료를 생성하여 사용함으로써 최근의 저수 유출특성을 반영한 빈도분석 결과를 얻을 수 있으리라 판단된다.

2.2 통계실험을 위한 합성 유량자료의 발생

본 연구의 최종 목표는 Bayesian MCMC 방법의 결과와 2차 근사식을 이용한 취우추정방법으로 빈도분석을 수행하고 그 결과를 불확실성 측면에서 비교한 후, 두 방법의 비교분석 결과와 적용가능성을 알아보는 것이다. 이와 같은 통계적 모형간의 비교를 위해서는 실측 유량자료를 이용한 적용을 수행하기 이전에 참값을 알고 있는 합성 유량(synthetic flow)자료를 이용하여 어떤 방법이 참값에 가까운 결과를 모의하는 지, 불확실성 측면에서 어떤 결과를 나타내는지를 알아보기 위한 통계적 실험을 수행할 필요가 있다.

본 연구에서는 2모수 Weibull 분포를 사용하고, 그 모수의 값으로써 Table 1과 같은 8개의 값을 사용하여 자료의 길이가 100인 7일 지속기간 최소유량(7Q)을 발생시킨 후, 각각의 자료 셋을 이용하여 Bayesian MCMC 방법과 2차 근사식을 이용한 취우추정방법을 이용하여 다시 자료의 모수의 2.5%, 평균, 97.5%에 해당되는 모수를 추정하였다. 단, Table 1에서 각각의 모수들은 낙동강 유역의 14개 지점에 대한 36년간 유량자료를 최우추정방법으로 추정한 후, 상한과 하한을 정하고 모수별 임의적인 구간을 선정하여 결정하였으며 7번째 자료 셋의 모수인 2.8371와 34.4203은 진동지점의 36년간 유량자료로부터 얻어진 36개의 7Q자료를 최우추정방법으로 추정하여 얻은 추정치로써 두 가지 추정방법을 이용하여 진동지점에서 얻어진 모수의 추정치를

Table 1. Weibull parameters for generation of 8 synthetic data sets

Data set number	Shape parameter, α	Scale parameter, β
1	1.8000	10.0000
2	2.0000	15.0000
3	2.2000	20.0000
4	2.4000	25.0000
5	2.6000	30.0000
6	2.8000	35.0000
7	2.8371	34.4203
8	3.0000	40.0000

Table 2. Statistical properties of each data (at Jindong station)

Types of data	Mean	Standard deviation	Coefficient of variation	Coefficient of skewness	Coefficient of kurtosis
Synthetic data (100 years) with α = 2.8371, β = 34.4203	32.6943	13.3770	0.4092	0.3235	2.5272
Gauged data (36 years)	32.8461	12.8604	0.3915	0.3788	2.8204

얼마나 합리적으로 추정할 수 있는지를 알아보기 위하여 사용되었다. 또한 7번의 모수값들은 통계적으로 발생시킨 100개의 합성 유량자료와 실측 유량자료로부터 산정된 36개의 실측 유량자료와의 통계치를 비교함으로써 합성 유량자료가 올바르게 발생되었음을 입증하는 자료로도 사용되었다(Table 2). Table 2의 평균, 표준편차, 변동계수(coefficient of variation), 왜곡도(coefficient of skewness), 첨도(coefficient of kurtosis)를 비교해보았을 때, 합성 유량자료와 실측 유량자료의 통계적 특성치가 거의 유사한 것을 알 수 있으며, 이로부터 합성 유량자료를 이용하여 저수량 빈도분석을 수행하기 위한 Bayesian MCMC 방법과 2차 근사식을 이용한 취우추정방법으로 통계적 모의 실험을 하는 것이 무리가 없다는 것을 알 수 있었다. 단, 여기서 합성 유량자료의 통계치는 자료길이 100인 100개의 앙상블을 먼저 생성하고 이들의 통계치를 산술평균하여 나타낸 것이다.

2.3 저수(低水)특성 및 빈도분석 대상유량의 결정

진동지점에서 얻어진 실측 유량자료를 자연유량으로 환산한 후, 저수 특성을 나타낼 수 있는 여러 가지 지표를 산정하였다. 산정된 지표는 1966년부터 2001년까지의 36년간의 일자료를 이용한 MAR(Mean Annual Runoff), MF(Median Flow), AMF(Absolute Minimum Flow)로서 MAR은 년 평균유량, MF는 365개의 일 유량의 중간값, AMF는 365개의 일 유량 중 가장 작은 유량을 뜻한다. MAR은 홍수량이 시간적으로 밀집되어 있는 경우에는 이 값의 영향을 받아 저수분석을 위해서는 사용되지 못하는 것이 일반적이므로, 주로 저수 분석대상은 MF와 AMF 사이의 값이 된다. Fig. 2는 36년간 진동지점에서의 MAR, MF, AMF를 표시한 그림으로써 저수유량의 주요 관심이 되는 MF는 52.64~206.26 cms에서 변동되고, AMF는 8.85~52.32 cms에서 변동

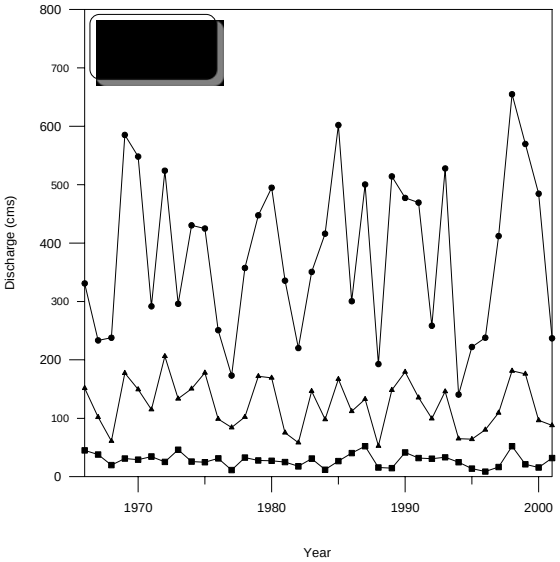


Fig. 2. Representative Low flow Indices at Jindong Station

되는 것을 알 수 있는데, 우리나라의 경우에는 특히 홍수량이 시기적으로 많이 집중되어 있고 하상경사가 커서 홍수 시 유량이 급속하게 감소하는 경향이 있으므로 중앙값인 MF보다는 AMF를 사용하는 것이 저수량 분석에 효과적일 수 있다.

그러나 위와 같은 지표들은 저수량의 변동과 같이 정성적인 부분만을 평가하는 데 있어서 주로 사용되어지며, 빈도 유량과 같은 정량적인 분석은 국내에서는 갈수량인 Q_{355}, 미국에서는 7일 지속기간 또는 10일 지속기간 최소유량이 사용된다. 위에서 산정된 MAR, MF, AMF외에 저수량 분석을 위해 국내에서 주로 사용되어지는 갈수량(Q_{355})과 7일 지속기간 최소유량(7Q) 및 10일 지속기간 최소유량(10Q)을 진동지점에서 산정하여 그림3에 나타내었다. 단, Fig. 3은 년도별로 산정된 지표들을 순위별로 정리한 그림이다.

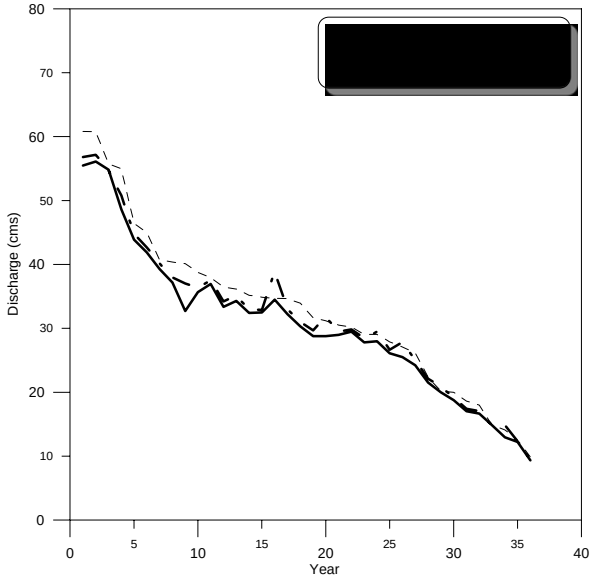


Fig. 3. Annual Minimum Flows for Different Duratins at Jindong Station

순위별로 표시된 Fig. 3의 결과를 보면, 1개년을 제외한 모든 년도에서 7Q가 10Q 보다 작은 값으로 산정되었으며, 마찬가지로 1개년을 제외한 모든 년도에서 Q_{355}보다 작은 값으로 산정 된 것을 알 수 있다. 조탁근 등(2007)은 국내 7개 댐 유역의 자료를 이용하여 Q_{355}와 7Q를 비교한 결과 위와 유사한 결과를 산정한 바 있으며, 이로부터 우리나라가 갈수기 유량을 미국보다 크게 평가하고 있어 이수관리 측면에서 상대적으로 보수적인 경향을 가지고 있음을 제안한 바 있다. 본 연구에서 산정된 결과를 이용해도 이수관리 측면에서 같은 결론을 내릴 수 있으며, 우리나라가 미국보다 이수관리 측면에서 더욱 여유 있는 유지유량이나 환경유량과 같은 하천관리유량이 산정되어 사용되고 있어 수자원관리 측면에서 불리한 면이 있음을 제시할 수 있었다.

빈도분석을 위해서 사용하고자 하는 자료가 결정되면 빈도분석을 수행하기 이전에 채택된 자료의 독립성과 정상성에 대한 사전 통계검정을 해야 한다(Rao and Hamed, 2000). 이를 위해서 본 연구에서는 Wald-Wolfowitz(1943)이 제안한 W-W 시험을 수행하였다. 이 시험은 주어진 자료, $x_1, x_2, \ldots x_n$가 있을 때, 다음의 표준정규변량 z를 산정하여 이 값을 임의의 유의수준에 해당되는 검증치와 비교함으로써 자료의 정상성과 독립성을 검정한다.

$$z = \frac{(R - \overline{R})}{[Var(R)]^{1/2}} \tag{1}$$

$$R = \sum_{i=1}^{n-1} x_i x_{i+1} + x_1 x_n \tag{2}$$

$$\overline{R} = \frac{(s_1^2 - s_2)}{n-1} \tag{3}$$

$$Var(R) = \frac{s_2^2 - s_4}{n-1} - \overline{R^2} + \frac{(s_1^4 - 4s_1^2 s_2 + 4s_1 s_3 + s_2^2 - 2s_4)}{(n-1)(n-2)} \tag{4}$$

여기서 R, $\overline{R}$, $Var(R)$은 각각 통계량과 이에 대한 평균 및 분산을 의미하며 $s_r = nm_r'$이고 m_r'은 r번째 모멘트를 나타낸다.

위와 같은 식들을 이용하여 자료 길이가 36인 10개의 소유역별 유출량과 4개의 수위관측소에서의 유량을 검정한 결과를 나타내면 다음 Table 3과 같다. Table 3에서 자료길이 36인 14개 지점의 유량 자료를 이용한 자료의 독립성과 정상성에 대한 검정은 유의수준 5.0 %에서 모두 만족스러운 결과를 얻을 수 있었다. 그러므로 본 연구에서는 산정된 여러 가지 저수유량의 지표들 중에서 가장 작은 값으로 평가할 수 있는 7일 지속기간 최소유량(7Q)만을 대상으로 하여 빈도분석을 수행하였으며, 재현기간은 1년부터 100년까지 반복 계산하여 최종적으로 7Q에 대한 빈도곡선을 작성하였다.

3. 2차 근사를 이용한 취우추정방법과 Bayesian MCMC 방법의 수행

본 연구의 Ⅰ편에서는 Bayesian MCMC 방법과 이를 수행하기 위한 Metropolis-Hastings 알고리즘의 이론적 배경과 Bayesian MCMC 방법을 수행하는데 있어서 중

Table 3. Results of the Preliminary Test at Significance Level 5.0 %(=1.96)

Name	Type	Number of data	Test statistic(z)	Test result
Sub-basin 1	Simulated	36	0.868	Accepted
Sub-basin 2	Simulated	36	0.498	Accepted
Sub-basin 3	Simulated	36	0.679	Accepted
Sub-basin 4	Simulated	36	0.432	Accepted
Sub-basin 5	Simulated	36	0.801	Accepted
Nakdong	Gauged	36	0.832	Accepted
Sub-basin 6	Simulated	36	0.698	Accepted
Waekwhan	Gauged	36	0.871	Accepted
Sub-basin 7	Simulated	36	0.521	Accepted
Goryunggyo	Gauged	36	0.698	Accepted
Sub-basin 8	Simulated	36	0.854	Accepted
Sub-basin 9	Simulated	36	0.956	Accepted
Sub-basin 10	Simulated	36	0.707	Accepted
Jindong	Gauged	36	0.656	Accepted

요한 자료에 기반한 사전분포와 제안분포를 제안하였다. 또한 Bayesian MCMC 방법의 비교대안으로 2차 근사식을 이용한 취우추정방법에 대한 이론적 배경과 가정조건으로 인해 발생하는 적용상의 단점들을 언급한 바 있다.

3.1 2차 근사를 이용한 취우추정방법의 수행

본 연구의 Ⅰ편에서 언급한 바와 같이 최우추정(Maximum Likelihood Estimation, MLE)방법은 우도함수를 최대화하는 모수를 결정하는 최적화 문제이다. 이를 위해서는 먼저 우도함수의 형태 및 특성을 파악하여 최우추정에 필요한 알고리즘을 검토할 필요가 있다. 선정된 확률밀도함수는 본 연구의 Ⅰ편에서 사용되었던 2모수 Weibull 분포(Eq. (5))이다. 또한 이에 대한 우도함수를 다시 나타내면 다음 식과 같고, 이를 각각의 모수의 범위 내에서 그림으로 나타내면 Fig. 4와 같다. 아래 식에서 α는 형상모수(shape parameter)이고 β는 축척모수(scale parameter)이다.

$$f(x) = \frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1}\exp\left[-\left(\frac{x}{\beta}\right)^{\alpha}\right] \tag{5}$$

$$L(\mathbf{x}\,|\alpha,\beta) = \left(\frac{\alpha}{\beta}\right)^{n}\prod_{i=1}^{n}\left(\frac{x_i}{\beta}\right)^{\alpha-1}\exp\left[-\sum_{i=1}^{n}\left(\frac{x_i}{\beta}\right)^{\alpha}\right] \tag{6}$$

Fig. 4로부터 주어진 모수 α, β의 범위 내에서 우도함수가 단일 모드를 가지는 것을 알 수 있으며, 이로부터 전역 최적화알고리즘을 사용하지 않고 국부 최적화

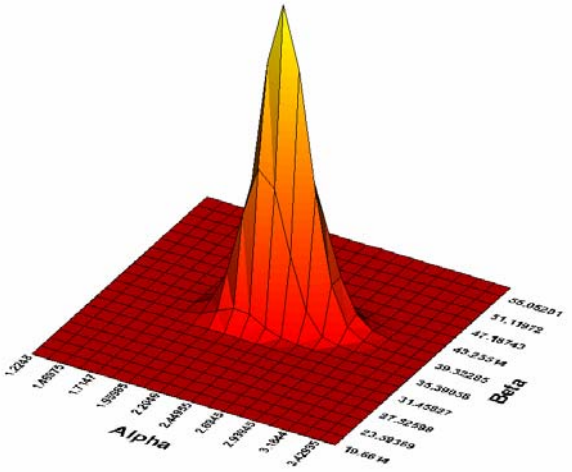


Fig. 4. Surface Plot of Likelihood Function with Shape Parameter = 2.8371 and Scale Parameter = 34.4203

알고리즘을 사용하여도 우도함수를 최대화하는 모수를 탐색할 수 있음을 알 수 있다. 그러므로 본 연구에서는 우도함수를 선형화하기 위하여 양변에 로그를 취한 후, 형상모수와 축척모수에 대한 편미분방정식의 값이 0이 된다는 두 개의 편미분 방정식을 구성한 후, 이를 반복법을 이용하여 모수의 최적값을 추정하였다. 그러나 이러한 도해적인 방법을 통해 최적화 알고리즘을 경사법만으로 결정하여 사용하는 것은 이후 추정결과에서 모수들의 변화에 따라 불연속적인 모수들로 추정될 수 있는 우려가 있으므로, 최우추정방법을 이용한 모수 추정 시 전역최적화 알고리즘을 사용함으로써 보다 나은 최

우추정결과를 얻을 수 있을 것으로 판단된다.

본 연구의 주된 관심사는 추정된 모수뿐만 아니라, 모수의 추정과정에서 발생되는 불확실성을 산정하는 것이므로, 이를 나타내기 위하여 연구의 Ⅰ편에서 제시한 2차 근사식을 이용한 신뢰구간의 범위에 대한 근사식을 이용하였다. 이를 다시 나타내면 다음과 같다.

$$\Pr[\hat{\boldsymbol{\theta}} - \frac{z_{(1-\alpha/2)}}{\sqrt{I}} < \boldsymbol{\theta} \le \hat{\boldsymbol{\theta}} + \frac{z_{(1-\alpha/2)}}{\sqrt{I}}] \approx 1-\alpha \qquad (7)$$

3.2 Bayesian MCMC 방법의 수행

본 연구의 Ⅱ편에서는 제안된 알고리즘과 분포들을 이용하고 합성 유량자료와 실측 유량자료를 대상으로 Bayesian MCMC 방법을 이용한 점 빈도분석을 낙동강 유역의 본류에 위치한 낙동, 왜관, 고령교, 진동지점에 대하여 수행하였으며 이를 위해 제안된 관련 식을 요약하면 다음과 같다. 또한 Metropolis-Hastings 알고리즘은 Matlab을 이용하여 코딩하였으며, 수행시간은 Pentium Ⅳ-3.2G를 이용하여 자료의 길이가 100인 합성 유량자료의 경우 약 28분 정도가 소요되었고, 자료의 길이가 36인 실측 유량자료인 경우에는 17분 정도가 소요되었다. 단 논문의 Ⅰ에서 검토된 군집분석의 결과를 이용하여 낙동강유역을 하나의 동질한 유역으로 보고, 진동지점에서 추정된 λ, a, b의 추정 값이 낙동, 왜관, 고령교의 사전분포에도 동일하다고 가정함으로써 연구를 진행하였으며, 향후에는 각 지점별로 다른 사전분포나 다른 추정치를 사용하여 연구를 수행함으로써 보다 정확한 결과를 얻을 수 있으리라 판단된다.

① 빈도분석을 위한 Bayesian 방법:

$$\pi(\alpha,\ \beta\,|\mathbf{x}) = \frac{L(\mathbf{x}|\alpha,\ \beta)\pi(\alpha,\ \beta)}{\int_0^\infty \int_0^\infty L(\mathbf{x}\,|\alpha,\ \beta)\pi(\alpha,\ \beta)d\alpha d\beta} \qquad (8)$$

② 우도함수:

$$L(\mathbf{x}\,|\alpha,\beta) = \left(\frac{\alpha}{\beta}\right)^n \prod_{i=1}^{n}\left(\frac{x_i}{\beta}\right)^{\alpha-1} \exp\left[-\sum_{i=1}^{n}\left(\frac{x_i}{\beta}\right)^{\alpha}\right] \qquad (9)$$

③ 자료에 기반한 사전분포:

$$\pi(\alpha,\beta) = \frac{b}{a}\left(\frac{\beta}{a}\right)^{b-1} \exp\left[-\left(\left(\frac{\beta}{a}\right)^b + \lambda\alpha\right)\right] \qquad (10)$$

(여기서, λ, a, b는 각각 추정된 0.98, 27.15, 1.92)

④ 제안분포: $q(\boldsymbol{\theta}_{j+1}|\boldsymbol{\theta}_j) = q_1(|\boldsymbol{\theta}_{j+1} - \boldsymbol{\theta}_j|)$

(여기서, q_1은 이변량 정규분포) (11)

4. 합성 유량자료를 이용한 결과의 분석

위에서 구축된 모형과 통계적 실험을 위하여 생성된 8개의 합성 유량자료 셋으로부터 얻어진 자료길이 100인 8개 셋의 7Q유량을 이용하여 Metropolis-Hastings 알고리즘을 100,000번 반복하여 사후분포로부터 모수를 샘플링하였다. 단, 알고리즘의 안정성을 위하여 최초 샘플링된 1,000개는 무시하고 99,000개의 추정치만을 이용하여 통계적 특성치를 산정하였다. 또한 Bayesian MCMC 방법의 비교 대안으로 최우추정법을 통해 얻어진 모수의 추정값과 이에 대한 신뢰구간을 산정하였다.

Table 4와 5는 두 가지 방법의 수행으로 얻어진 모수의 추정값을 나타낸 것으로써, Table 4는 형상모수의 결과이고 Table 5는 축척모수의 결과를 나타낸 것이다. Table 4와 Table 5의 값들을 비교해보면, Bayesian MCMC 방법의 평균값과 취우추정방법의 모드에서의 값의 차이는 형상모수의 경우 0.03~0.30이고 축척모수의 경우는 0.05~2.00으로 큰 차이를 보이지 않는다. 즉, 확정적인 값만을 필요로 하는 경우에는 취우추정방법보다 적용 절차가 복잡한 Bayesian MCMC 방법을 사용하는 것이 큰 장점을 가지지 못함을 알 수 있다. 만약 우도함수가 다중모드를 가지거나 모수의 차원이 커지는 경우에도 이와 같은 현상은 동일하게 일어날 것이므로 국부최적화 알고리즘 대신 전역최적화 알고리즘을 사용하면 취우추정방법을 이용하여 추정된 모수의 최적값을 확정적인 개념에서 추정하여 사용할 수 있을 것이다.

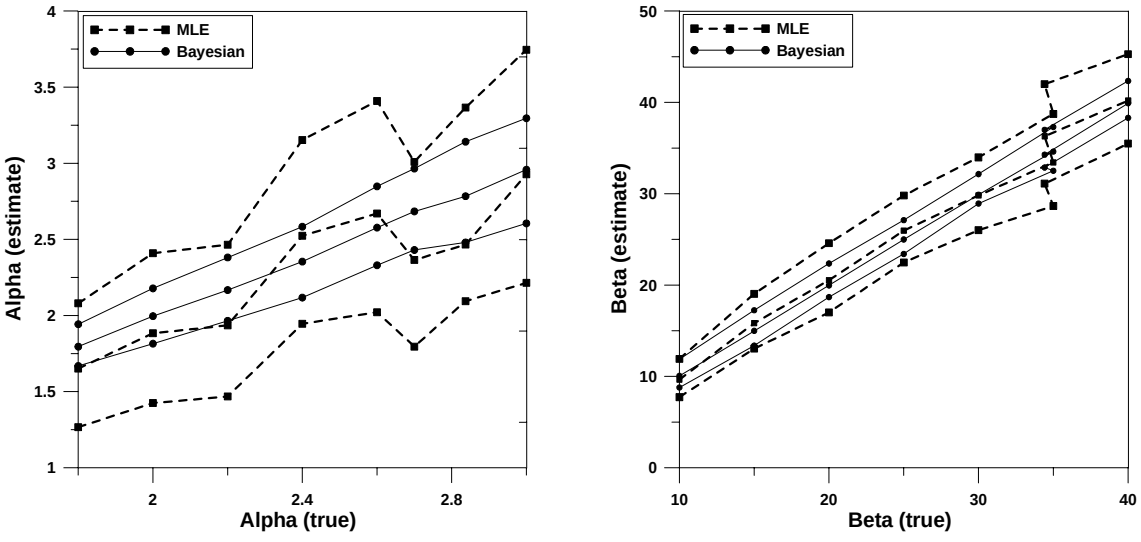
그러나 확률적인 개념이 필요한 경우, 즉 불확실성을 고려한 모수의 추정은 두 가지 방법에 의한 결과로부터 Bayesian MCMC 방법이 우수하다는 결론을 얻을 수 있다. Table 4와 5의 2.5 %와 97.5 % 사이의 추정치의 차이를 살펴보면 Bayesian MCMC 방법의 경우 형상모수가 0.3~0.7의 범위를 가지는 반면 2차 근사식을 이용한 취우추정방법은 0.8~1.5사이의 범위를 나타내는 것을 알 수 있으며, 축척모수의 경우도 Bayesian MCMC 방법의 경우 3.1~4.8의 범위를 가지며 2차 근사식을 이용한 취우추정방법의 경우 4.2~11.0의 범위를 가지는 것을 알 수 있다. 즉, Bayesian MCMC 방법의 경우 2차 근사식을 이용한 취우추정방법보다 감소된 불확실성의 범위를 나타내는 것을 알 수 있고, 이로부터 불확실성의 측면에서는 Bayesian MCMC 방법이 2차 근사식을 이용한 취우추정방법보다 우수한 방법이라는 결론을 내릴 수 있었다. 이와 같은 추정방법에 따른 모수의 산정값을 그림으로 나타내면 Fig. 5와 같고 이로부터 Bayesian MCMC 방법을 사용하는 경우 모수의 범위가 감소되어져 나타나는 것을 알 수 있다.

Table 4. Comparison of Shape Parameter with 8 Synthetic Data Sets

True shape parameter (α)	MLE				Bayesian MCMC				Difference (b) and (e)
	2.50% (a)	Mode (b)	97.50% (c)	Difference (c) and (a)	2.50% (d)	Mean (e)	97.50% (f)	Difference (f) and (d)	
Set 1 : 1.8000	1.2663	1.6509	2.0797	0.8134	1.6684	1.7954	1.9427	0.2743	0.1445
Set 2 : 2.0000	1.4250	1.8834	2.4089	0.9839	1.8146	1.9956	2.1780	0.3634	0.1122
Set 3 : 2.2000	1.4682	1.9349	2.4648	0.9966	1.9650	2.1670	2.3811	0.4161	0.2321
Set 4 : 2.4000	1.9455	2.5235	3.1526	1.2071	2.1176	2.3537	2.5830	0.4654	0.1698
Set 5 : 2.6000	2.0216	2.6695	3.4089	1.3873	2.3298	2.5771	2.8478	0.5180	0.0924
Set 6 : 2.7000	1.7960	2.3647	3.0088	1.2128	2.4303	2.6837	2.9645	0.5342	0.3190
Set 7 : 2.8371	2.0941	2.4654	3.3653	1.2712	2.4798	2.7832	3.1411	0.6613	0.3178
Set 8 : 3.0000	2.2143	2.9270	3.7453	1.5310	2.6048	2.9578	3.2955	0.6907	0.0308

Table 5. Comparison of Scale Parameter with 8 Synthetic Data Setss

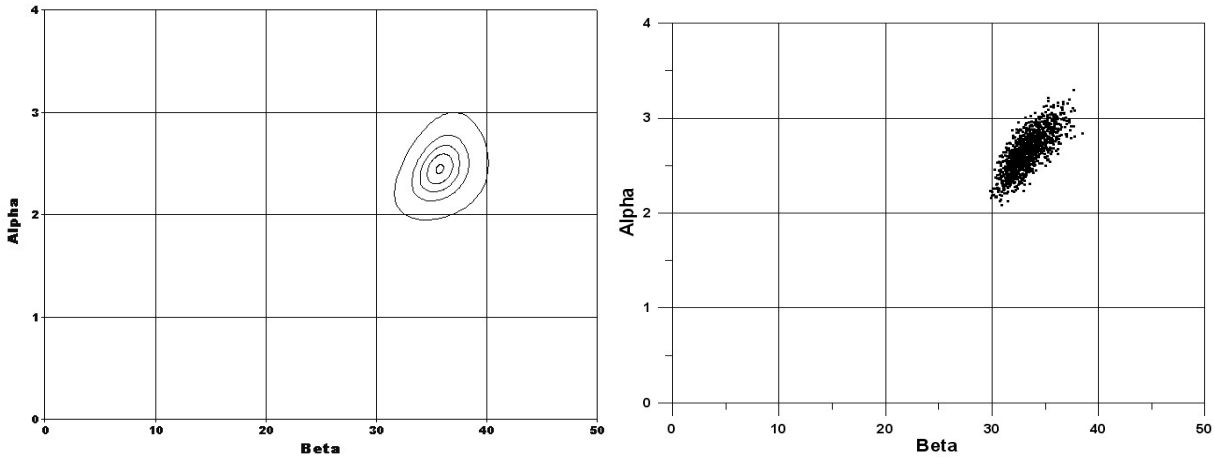
True scale parameter (β)	MLE				Bayesian MCMC				Difference (b) and (e)
	2.50% (a)	Mode (b)	97.50% (c)	Difference (c) and (a)	2.50% (d)	Mean (e)	97.50% (f)	Difference (f) and (d)	
Set 1 : 10.0000	7.7280	9.6401	11.9111	4.1831	8.7791	10.0618	11.8916	3.1125	0.4217
Set 2 : 15.0000	13.0307	15.8242	19.0417	6.0110	13.3686	14.9855	17.2377	3.8691	0.8387
Set 3 : 20.0000	17.0018	20.5283	24.5813	7.5795	18.6785	19.9589	22.3572	3.6787	0.5694
Set 4 : 25.0000	22.4696	25.9496	29.8029	7.3333	23.3969	24.9838	27.0996	3.7027	0.9658
Set 5 : 30.0000	26.0055	29.8188	33.9782	7.9727	28.9208	29.8720	32.1437	3.2229	0.0532
Set 6 : 35.0000	28.6512	33.4262	38.7364	10.0852	32.5087	34.6072	37.2976	4.7889	1.1810
Set 7 : 34.4203	31.1011	36.2876	41.9878	10.8867	32.8653	34.2711	36.9877	4.1224	2.0165
Set 8 : 40.0000	35.4751	40.1901	45.2724	9.7973	38.2996	39.9030	42.3400	4.0404	0.2871



(a) Shape Parameter (b) Shape Parameter

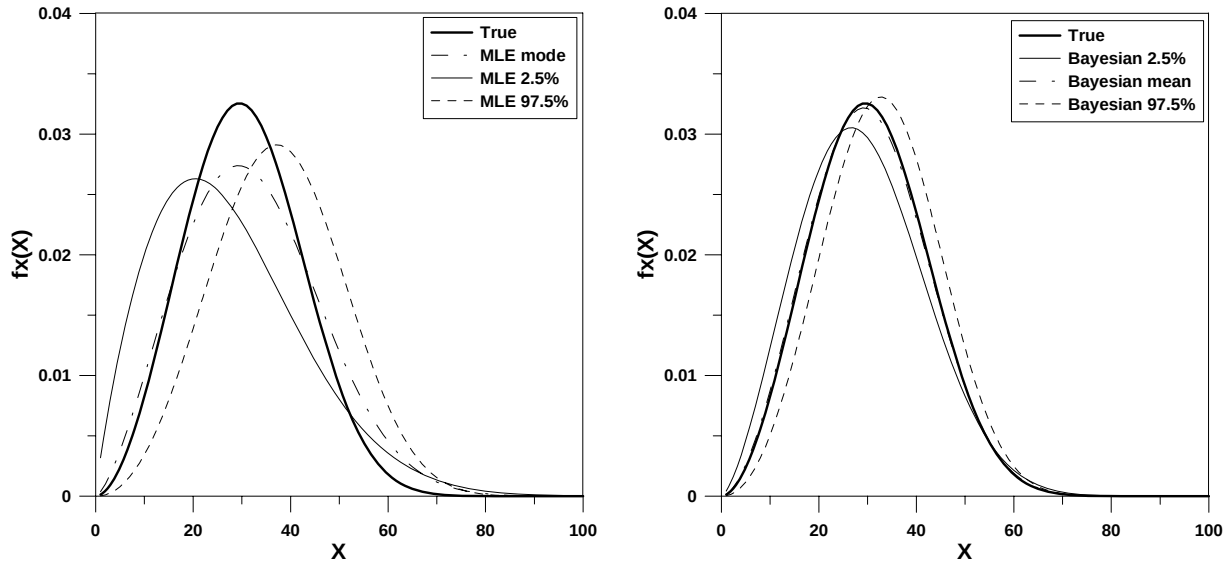
Fig. 5. True Parameter v.s Estimated Parameter

또한 Fig. 6은 각각 합성 유량자료 7번에 대한 Bayesian MCMC 방법과 2차 근사식을 이용한 취우추정방법을 이용하여 얻어진 사후분포의 샘플링 공간을 나타낸 것이며, Fig. 6에서도 위의 결론과 마찬가지로 Bayesian MCMC 방법이 2차 근사식을 사용한 취우추정방법보다 불확실성 측면에서 우월하다는 결론을 보다



(a) MLE (b) Bayesian MCMC

Fig. 6. Contour Plots of MLE and Bayesian MCMC(Case: set 7)



(a) MLE (b) Bayesian MCMC

Fig. 7. Probability Density Function with MLE and Bayesian MCMC

확실히 알 수 있었다.

그러나 불확실성 분석의 최종적인 목표는 모수의 불확실성에 대한 분석보다는 모수의 불확실성에 따른 확률밀도함수의 불확실성에 있다고 할 수 있다. 그러므로 본 연구에서는 합성 유량자료 셋 7번의 추정결과에 따른 확률밀도함수의 불확실성을 Fig. 7에 나타내었다. Fig. 7은 추정된 모수에 따른 2모수 Weibull 분포의 변동을 나타냄으로써 사용하고자 하는 확률밀도함수의 불확실성을 나타내고 있다고 할 수 있으며, Bayesian MCMC 방법이 2차 근사식을 이용한 취우추정방법보다 신뢰구간에 따른 변동이 적으므로 불확실성 측면에서 우수하다는 것을 알 수 있다.

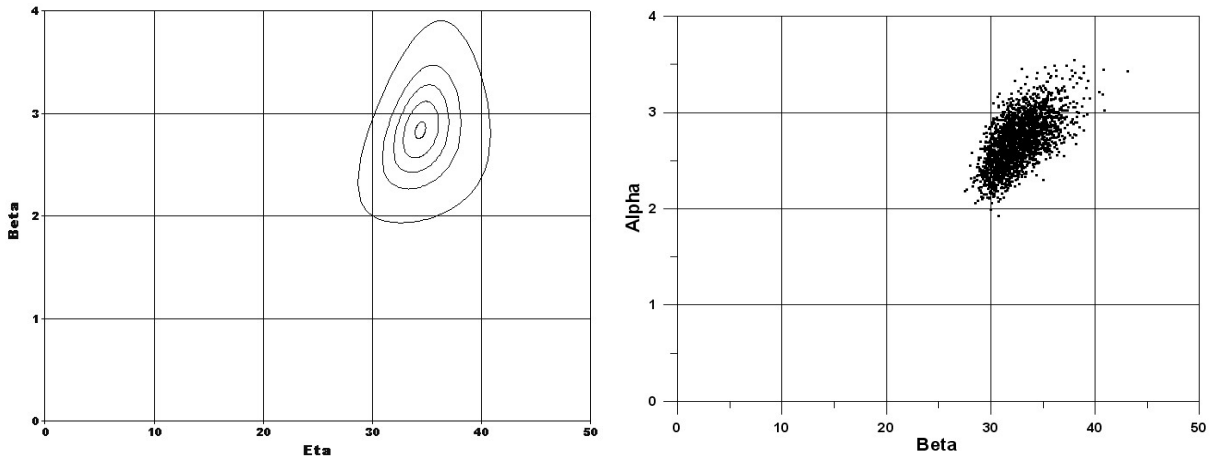
5. 실측 유량자료를 이용한 결과의 분석

본 절에서는 4개 지점의 실측유량을 이용하여 4절에서 수행한 것과 같은 과정을 다시 수행하였다. Table 6은 진동지점의 36년간의 자료를 이용하여 추정한 형상모수와 축척도모수의 결과를 나타낸 것이며, 위의 결론과 같이 Bayesian MCMC 방법이 2차 근사식을 이용한 취우추정방법보다 불확실성 측면에서 우수한 결과를 얻을 수 있으며, 또한 평균값과 모드값에 있어서는 큰 차이를 보이지 않는 것을 알 수 있다.

또한 각각의 방법에 따른 샘플링 공간을 나타내면 Fig. 8과 같고, 진동지점의 자료를 이용한 추정값을 이용하여 2모수 Weibull 확률밀도함수를 나타내면 Fig. 9

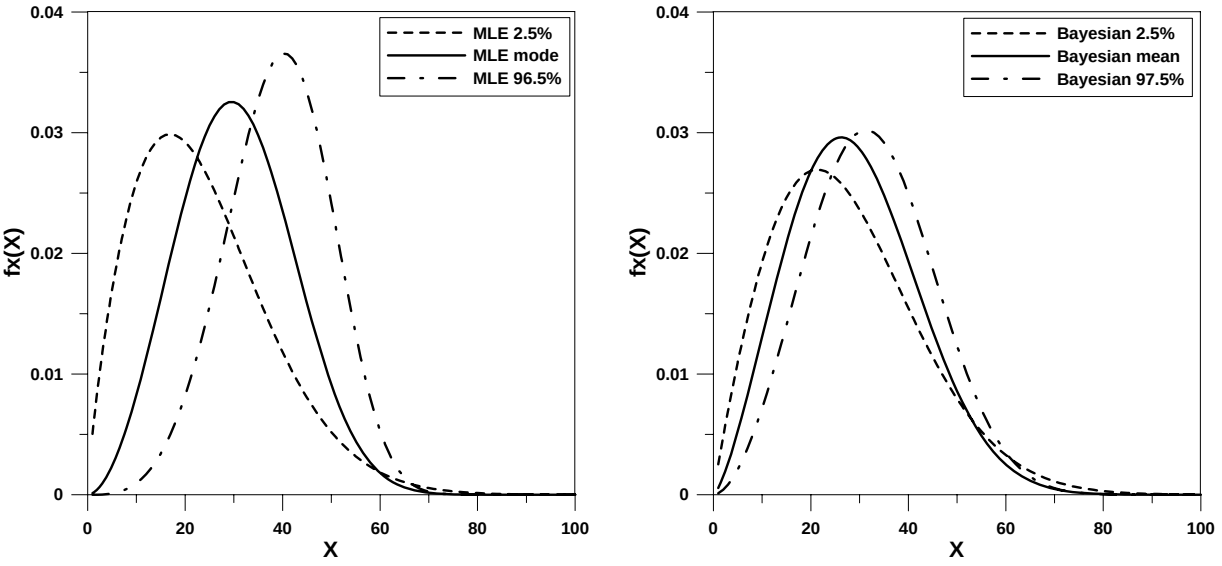
Table 6. Results of Parameter Estimation with 36 Gauged Data(at Jindong Station)

Weibull parameters	MLE			Bayesian MCMC		
	2.5 %	Mode	97.5 %	2.5 %	Mean	97.5 %
Shape (α)	1.7823	2.8371	4.1511	1.9311	2.3897	2.7996
Scale (β)	26.8116	34.4203	43.1168	31.1398	32.9181	36.6753



(a) MLE (b) Bayesian MCMC

Fig. 8. Contour Plot of MLE and Bayesian MCMC (at Jindong)



(a) MLE (b) Bayesian MCMC

Fig. 9. Probability Density Function with Gauged Data (at Jindong)

와 같다. Bickel and Doksum(1977)은 2차 근사식을 이용한 신뢰구간식을 제안하면서 자료의 길이가 감소되는 경우 신뢰구간에 의해 추정된 하한값과 상한값의 정확성이 감소되어 불확실성 측면에서 과대추정될 수 있다는 결론을 제시한 바 있다. Fig. 8a를 보면, 4절의 2차 근사식을 이용한 취우추정방법에 의한 결과(Fig. 6a)보다 더욱 넓은 샘플링공간을 보여주는 것을 알 수 있으며, Fig. 9a에서는 확률밀도함수의 변동이 Fig. 7a보다 큰 것을 알 수 있다. 이에 비해 Bayesian MCMC 방법의 결과인 Fig. 8b와 Fig. 9b는 자료의 길이가 감소되어도 2차 근사식을 이용한 취우추정방법보다 불확실성이 크게 증가되지 않는 것을 알 수 있다. 이로부터 Bickel

and Doksum(1977)이 제시한 결과와 같은 결과를 얻을 수 있었으며, 특히 국내에서 저수량 빈도분석을 수행하고자 하는 경우에는 대부분의 지점에서 자료의 길이가 짧은 한계성이 존재하므로 2차 근사식을 이용한 취우추정방법보다 Bayesian MCMC 방법을 사용하는 것이 저수량 빈도분석에서 존재하는 불확실성을 표현하는 데 있어서 합리적일 수 있다는 결론을 얻을 수 있었다.

6. 2모수 Weibull 분포의 적정성 검토와 빈도곡선의 작성

본 절에서는 낙동, 왜관, 고령교 및 진동지점에서의 불확실성을 포함한 빈도곡선을 작성하기 이전에 5절에서 추정된 진동지점에서의 36년 자료를 이용한 Bayesian MCMC 방법의 평균 추정값을 이용하여 2모수 Weibull 분포의 실측 유량자료에 대한 적정성을 검토하였다. 이를 검토하기 위하여 추정된 각각의 모수와 36개의 진동지점의 7Q유량을 Weibull 확률지에 나타낸 결과 Fig. 10과 같은 결론을 얻을 수 있었으며, 같은 모수를 이용하여 Weibull 확률분포와 36개의 비교한 결과 Fig. 11과 같은 결과를 얻을 수 있었다. 두 가지 그림으로 36개의 실측 유량자료와 추정된 모수를 이용한 결과가 적정하다는 결론을 내릴 수 있었으며, 이로부터 진동지점에서 2모수 Weibull 분포를 사용하는 것이 타당하다는 결론을 얻을 수 있었다.

위에서 제시한 실측 유량자료에 대한 Bayesian MCMC 방법과 2차 근사를 이용한 취우추정방법을 낙동, 왜관, 고령교 지점에도 같은 과정으로 적용하였다. 단, 사전분포는 진동지점을 대상으로 하는 자료에 기반한 분포만을 타 지점에도 사용하였으며, 각 지점에서의 자료에 기반한 사전분포를 구축하여 점 빈도분석을 Bayesian MCMC 방법으로 수행하는 과정은 추가적인 연구를 통하여 수행하고 그 결과를 비교함으로써 또 다른 의미 있는 결론을 얻을 수 있을 것으로 판단된다. 최종적으로 두 가지 추정방법에 의해 얻어진 모수의 2.5 %, 평균값(또는 모드값), 97.5 %에 해당되는 모수를 이용하여 낙동, 왜관, 고령교, 진동지점에서의 1년부터 100년의 재현기간에 해당되는 분위수를 추정하였다. 그 결과를 10년, 50년, 100년 빈도에 대해서만 요약하여 나타내면 다음 Table 7과 같다.

Table 7에서 추정된 빈도유량을 이용하여 빈도곡선을 작성하기 위해서는 36개의 실측 유량자료의 위치를 결정하기 위한 확률도시식이 필요하다. 본 연구에서는 주로 저수량의 확률도시에 사용되어진 다음의 Gringorten(1963) 식을 이용하여 실측 유량자료의 위치를 산정하였다.

$$\xi(i) = \frac{i - 0.44}{n + 0.12} \tag{12}$$

여기서 $i = 1,2,...,n$이고, $n = 36$이다.

Figs. 12~15에는 각각 낙동, 왜관, 고령교, 진동지점에 대하여 최종적으로 구축된 불확실성을 포함한 빈도곡선을 나타내며 모든 지점에서 실선으로 나타낸 그림이 Bayesian MCMC 방법이고 점선으로 나타낸 그림이 2차 근사식을 이용한 취우추정방법에 의한 결과로서 각각 아래로부터 2.5 %, 평균, 97.5 %를 나타낸다. 4개의 모든 지점에서 평균값에 대한 빈도곡선의 차이는 두 방법에 있어서 큰 차이를 나타내지 않지만, 불확실성 측면에서는 Bayesian MCMC 방법이 불확실성을 감소시켜 나타낼 수 있음을 확인할 수 있었다.

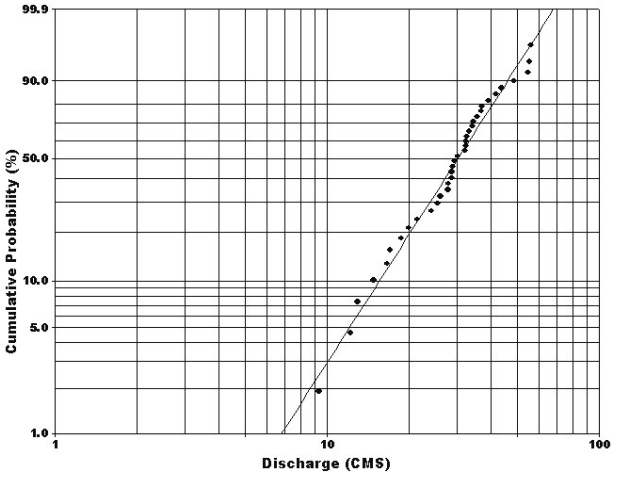


Fig. 10. Weibull Probability Plot with Gauged Data (at Jindong)

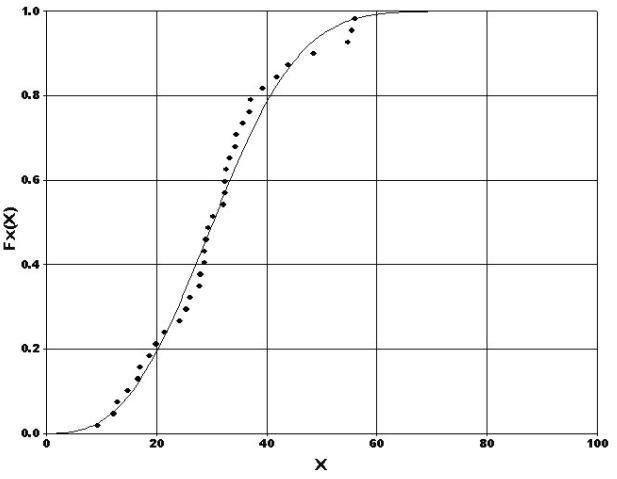


Fig. 11. Cumulative Weibull Distribution Functionand Gauge Data (at Jindong)

Table 7. Results of Quantile Estimation with Gauged Data

Station	Return Period (Year)	MLE			Bayesian MCMC		
		2.5%	Mode	97.5%	2.5%	Mean	97.5%
Nakdong	10	3.30	6.65	10.53	5.75	7.12	8.18
	50	1.36	3.87	7.35	3.15	4.17	5.08
	100	0.92	3.03	6.23	2.41	3.28	4.09
Waegwan	10	3.56	7.32	11.78	5.16	8.09	10.70
	50	1.37	4.03	7.83	2.25	4.61	7.00
	100	0.92	3.13	6.58	1.58	3.62	5.84
Goryeonggyo	10	4.25	8.87	14.40	5.50	8.22	10.61
	50	1.57	4.72	9.29	2.36	4.38	6.92
	100	1.02	3.57	7.61	1.62	3.31	4.56
Jindong	10	7.59	15.57	25.07	9.71	12.87	16.42
	50	3.00	8.70	16.84	4.13	6.43	9.10
	100	2.03	6.80	14.24	2.83	4.76	7.09

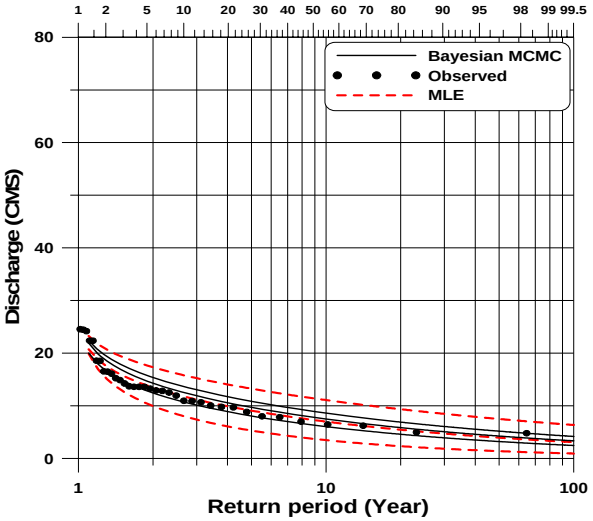


Fig. 12. Low Flow Frequency Curve (at Nakdong Station)

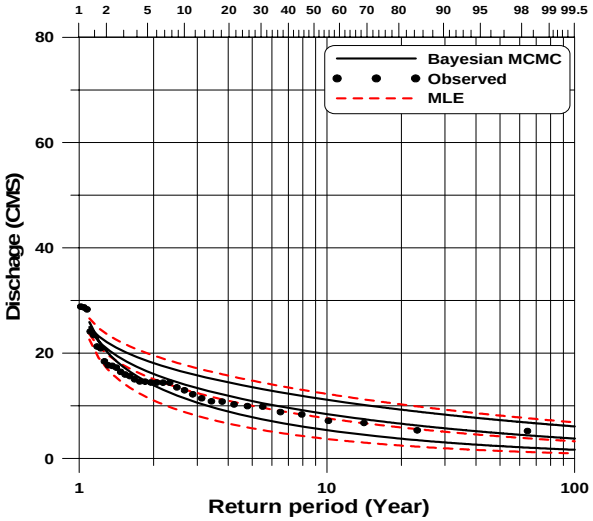


Fig. 13. Low Flow Frequency Curve (at Waegwan Station)

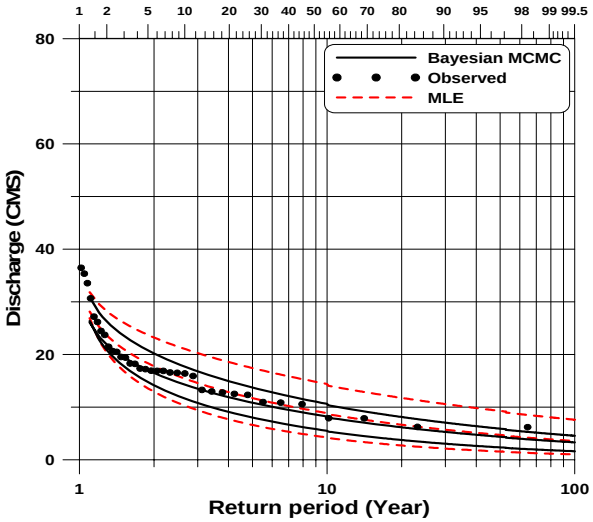


Fig. 14. Low Flow Frequency Curve (at Goryeonggyo Station)

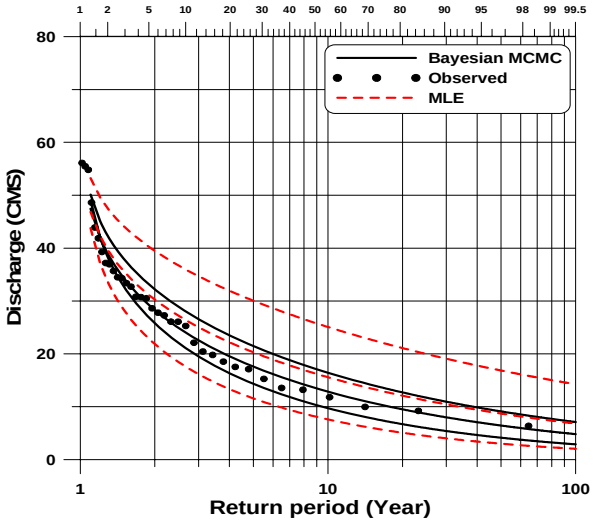


Fig. 15. Low Flow Frequency Curve (at Jindong Station)

7. 결론 및 향후 과제

본 연구는 Bayesian 방법을 이용하여 저수량 빈도분석을 수행하기 위한 연구로써 I 편에서 수행된 이론적 배경과 Bayesian MCMC 방법과 2차 근사식을 이용한 취우추정방법을 이용하기 위해 구축된 자료에 기반한 사전분포, 제안분포, 신뢰구간의 산정식 등을 이용하여 낙동강 유역의 본류지점인 낙동, 왜관, 고령교, 진동지점에 대한 점 빈도분석을 수행하고 그 결과로써 불확실성을 포함한 빈도곡선을 작성하였다.

두 가지 추정방법의 분석을 위하여 먼저 자료의 길이가 100인 8개의 합성 유량자료 셋을 각각 다른 8개의 2모수 Weibull 분포의 모수를 이용하여 생성하고, 생성된 자료를 다시 이용하여 두 가지 추정방법에 의해 모수를 추정하는 통계적 실험을 수행하였다. 그 결과 Bayesian MCMC 방법에 의한 평균값과 2차 근사식을 이용한 취우추정방법에 의한 모드에서의 추정값은 비슷한 결과를 보였으나 불확실성을 나타내는 하한값과 상한값의 차이는 Bayesian MCMC 방법이 2차 근사식을 이용한 취우추정방법보다 불확실성을 감소시켜 나타낼 수 있음을 알 수 있었다. 또한 진동지점의 36년간의 실측 유량자료를 이용하여 산정된 자료길이 36인 실측 유량자료를 이용하여 같은 방법을 수행한 결과, 평균값과 모드에서의 추정값은 큰 차이를 보이지 않는 유사한 결과를 얻을 수 있었다. 그러나 불확실성 측면에서는 2차 근사식을 이용한 취우추정방법의 경우 자료의 길이가 감소됨에 따라 불확실성의 범위가 합성 유량자료를 사용한 경우에 비해 상대적으로 증가되지만 Bayesian MCMC 방법의 경우에는 자료의 길이에 대한 영향이 거의 없다는 결론을 얻을 수 있었다. 그러므로 저수량 빈도분석을 수행하기 위해 충분한 자료를 확보할 수 없는 국내의 상황을 감안할 때 위와 같은 결론으로부터 Bayesian MCMC 방법이 불확실성을 표현하는데 있어서 2차 근사식을 이용한 취우추정방법에 비해 합리적일 수 있다는 결론을 얻을 수 있었다.

두 가지 방법을 낙동, 왜관, 고령교 지점에 동일하게 적용하여 최종적으로 진동을 포함한 4개 지점에 대한 불확실성을 포함한 빈도곡선을 작성할 수 있었으며, 이를 이용하면 저수관리 측면에서 불확실성을 고려한 관리를 수행할 수 있을 것으로 판단되며 이를 위해서는 보다 다양한 지점에 개발된 방법의 적용 및 평가가 필요하다고 할 수 있을 것이다. 특히 실무적으로 위에서 얻은 불확실성을 포함한 빈도유량을 사용함으로써 유지유량의 산정, 오염총량제에서 사용되는 저수량의 산정 그리고 수질모형을 이용한 장기수질예측 등의 연구에 있어서 불확실성을 고려할 수 있을 것으로 판단된다.

또한 본 연구에서는 실측 유량자료를 허가취수량만을 이용하여 자연유량으로 환산하여 사용하였는데, 이에 대한 개선으로서 자연유량을 산정함에 있어서 실제 취수량 자료 및 기타 자료의 확보 및 검증에 대한 보다 활발한 연구가 진행되어야 할 것으로 판단되며, 향후에는 입력자료로서 7Q10 뿐만 아니라 국내에서 사용되고 있는 기준갈수량 등을 사용하여 빈도유량과 불확실성을 산정함으로써 개발된 알고리즘의 국내 적용성을 검토해야할 필요가 있다고 판단된다.

감사의 글

본 연구는 21세기 프런티어 연구개발 사업인 수자원의 지속적 확보기술개발 사업단(과제번호 1-7-3)의 서울대학교 공학연구소를 통한 연구비 지원(30 %)과 서울대학교 BK21 안전하고 지속가능한 사회기반건설사업단의 연구비 지원(70 %)에 의해 수행되었습니다. 연구비 지원에 심심한 감사의 뜻을 표합니다.

참 고 문 헌

건교부, 한국수자원공사 (2006). **낙동강 유역조사사업 보고서**

조탁근, 김영오, 이길성 (2007). "국내 미계측유역의 갈수량 산정 개선." **한국수자원학회 논문집**, 한국수자원학회, 제40권, 제2호, pp. 113-124.

Bickel, P.J., and Doksum, K.A. (1977). *Mathematical Statistics: Basic Ideas and Selected Topics*. Holden-Day, Inc., San Francisco, CA.

Boes, D.C., Heo, J-H.,Salas, and J.D. (1989). "Regional flood quantile estimation for a Weibull model." *Water Resources Research,* Vol. 25, No. 5, pp. 979-990.

Boughton, W.C. (1976). "Flood estimation from short records." *Journal of Hydraulic Division*, pp. 241-253.

Burges, S.J., Lattenmaier, D.P., and Bates, C.L. (1975). "Properties of the three parameter log-normal probability distribution." *Water Resources Research*, Vol. 11, No. 2, pp. 229-235.

Dracup, J.A., Lee, K.S., and Paulson, E.G. (1980). "On the definition of droughts." *Water Resources Research*, Vol. 16, No. 2, pp. 297-302.

Gringorten, I.I. (1963). "A plotting rule for extreme probability paper." *Journal of Geophysics Research*, Vol. 68, No. 3, pp. 813-814.

Matalas, N.C., and Benson, M.A. (1968). "Note on the standard error of the coefficient of skewness." *Water Resources Research*, Vol. 4, No. 1, pp. 204-205.

Nathan, R.J., and McMahon, T.A. (1990). "Practical aspects of low flow frequency analysis." *Water Resources Research*, Vol. 26, No. 9, pp. 2135-2141.

Rao, A.R., and Hamed, K.H. (2000). *Flood Frequency Analysis*. CRC Press, Boca Raton, FL.

Rao, D.V. (1980). "Log Pearson type 3 distribution: Evaluation." *Journal of Hydraulic Division*, pp. 853-872.

Smakhtin, V.U.(2001). "Low flow hydrology: A review." *Journal of Hydrology*, Vol. 240, pp. 147-186.

Wald, A., and Wolfowitz, J. (1943). "An exact test for randomness in the nonparametric case based on serial correlation." *Annals of Mathematical Statistics*, Vol. 14, pp. 378-388.

WMO. (1974). *International Glossary of Hydrology*. WMO, Geneva.

(논문번호:07-107/접수:2007.10.10/심사완료:2007.12.21)