

물수지 분석을 위한 불확실성 정량화 방안

Quantifying Uncertainty for the Water Balance Analysis

이승욱*, 김영오**, 이동률***

Seung Uk Lee, Young-Oh Kim, Dong Ryul Lee

요 지

수자원장기종합계획에서는 물의 과부족 또는 가용한 물을 정량적으로 평가하기 위해 물수지 분석을 실시한다. 물수지 분석은 미래 예측되는 용수수요량과 공급가능량을 비교하는 단순한 과정이지만, 분석 과정에 포함되어 있는 자료와 모형의 불확실성으로 인하여 물수지 분석을 실시한 각종 보고서마다 서로 다른 결과를 보여주고 있어 국민의 신뢰를 얻지 못한 실정이다. 본 연구에서는 Monte Carlo simulation 기법 중 Latin Hypercube Sampling에 기반한 확률적 모사로 물수지의 불확실성을 표현하고, 이를 기존에서 제시한 단일 물부족량과 비교하여 불확실성의 범위와 특성을 분석하였다. 또한 민감도 분석을 수행함으로써 입력변수들 간의 상대적 중요도를 산정하여 수자원계획 수립시 투자 우선순위를 제시하였다.

핵심용어: 물수지 분석, 불확실성, Monte Carlo Simulation, Latin Hypercube Sampling, 민감도 분석

1. 서론

기존의 물수지 분석방법은 기본 가정이 많아 제한적이고 각 변수들의 변동 가능성을 무시함으로써 정해진 흐름도의 결과변수인 물부족량의 단일 값을 제시하고 있다. 이러한 방법은 문제를 단순화시켜 이해가 쉽도록 하며 결과를 나타내는데 드는 노력을 줄여주지만, 입력변수들의 자연적 변동성 또는 정확하게 알려지지 않는 불확실성의 존재로 발생할 수 있는 다양한 결과에 대한 예측이 불가능하다. 일반적으로 예측기간이 길어지면 불확실성들은 더욱 증폭되기 때문에 10년 앞을 내다보는 수자원장기종합계획은 불확실할 수밖에 없으므로 이에 대한 고려가 필수적이다. 최근 들어 김남원 등(2001)은 물수지 분석에서의 불확실성을 인위적인 형태와 자연적인 형태 그리고 시스템의 형태로 나누어 낙동강 유역에 대하여 분석하였고, Marjolein 등(2002)은 다양한 기후변화의 영향을 고려한 TARGETS 모형을 이용하여 네덜란드 남부에 대해 물수지 분석을 하였다.

본 연구에서는 수자원장기종합계획(건설교통부, 2000)의 물수지 분석에서 적용된 자료와 기본적 이론을 바탕으로 물수지 분석모형을 다시 구축하였고 이를 '수정 수장기'라 칭하였다. 이를 바탕으로 금강수계 301 소유역에 대해 물수지 분석의 대한 불확실성과 민감도 분석을 시행하였다.

2. 수정 수장기 물수지 분석

수정 수장기 물수지 분석모형의 기본 가정은 수자원장기종합계획(건설교통부, 2000)에서 제시한 가정들을 그대로 따랐다. 그러나 기존의 반순별 분석을 월별로 전환하여 분석하였고, 목표저수량 대신 20년 빈도의 갈수 유입량을 고려하는 적정확보수위와 최소확보수위를 고려하였다. 기존의 물수지 분석에서 고려한 생공용 수전용댐 이용량, 농업용댐 이용량, 유역별 용수의 절감량 및 중수도 이용량은 고려하지 않았다. 본 연구에서

* 서울대학교 지구환경시스템공학부 석사과정 • E-mail : lsu0102@snu.ac.kr

** 서울대학교 지구환경시스템공학부 조교수 • E-mail : yokim05@snu.ac.kr

*** 한국건설기술연구원 수자원환경연구부 수석연구원 • E-mail : dryi@kict.re.kr

적용한 금강수계는 전체 16개 소유역으로 나뉘져 있으며, 여기서 본류는 308, 310, 313, 314, 316 소유역으로 저수지 운영율의 영향을 받는 지역이고, 나머지 소유역은 지류지역으로 저수지 운영율의 영향을 받지 않는다. 저수지 운영에서 댐의 월방류량 결정은 최소확보수위와 적정확보수위를 기준으로, 현재수위가 상시만수위보다 높을 경우에는 최대발전방류와 여수로 방류를 하도록 하며, 현재수위가 상시만수위와 적정확보수위 사이에 있을 경우에는 최대발전방류를, 적정확보수위와 최소확보수위 사이에 있을 경우에는 기본계획방류를, 최소확보수위와 사수위 사이에 있을 경우에는 우선순위 (1.유지용수 2.생공용수 3.농업용수)에 따라 차등 공급하도록 하였다.

3. Latin Hypercube Sampling

일반적으로 불확실성 해석을 하기 위해 광범위하게 사용되는 접근법인 Monte Carlo simulation은 오랫동안 성공적으로 사용되어 왔다. 이는 불확실한 입력변수의 표본추출과 모형수행을 반복하여 실행함으로써 결정론적 문제를 확률적 문제로 전환하는 모의기법이다. 본 연구에서는 표본추출기법으로 전체 집합 S에서 n개의 점이 골고루 추출되도록 고안된 Latin Hypercube Sampling(LHS)을 선택하였다. 그림 1에서 보는 바와 같이 각각의 입력변수 x_k 를 n개의 등확률로 나눈 후 여기에서 하나씩 표본을 추출하는 방법이다. 입력변수의 자료가 충분히 크거나 많을 때는 random sampling을 사용하는 것이 더 효과적이지만 자료가 크지 않고 입력변수가 많을 때는 LHS가 훨씬 효과적인 결과를 나타낸다고 알려져 있다(Iman, 1981).

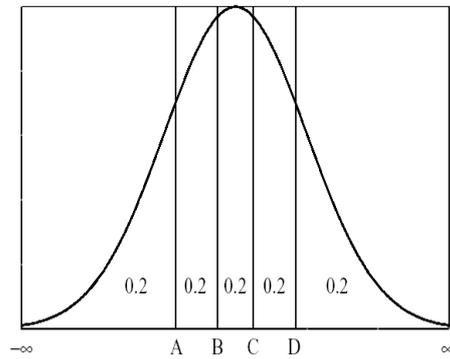


그림 1. Latin Hypercube Sampling의 등확률 분할 사례

4. 입력변수의 선정

본 연구에서는 수많은 물수지 분석 입력자료 중 물수지에 가장 큰 영향을 주는 자연유량, 생활용수, 공업용수, 농업용수, 회귀율을 대표 입력변수로 선정하였다. 이들에 대한 적절한 분포를 설정하여 LHS를 통해 나타나는 물부족량의 불확실성을 평가하고자 한다. 각 입력변수에 대한 분포설정은 다음과 같다.

자연유량은 수자원장기종합계획(건설교통부, 2000)에서 탱크모형으로 산정한 자연유량 32개년(1966년 10월~1998년9월) 시나리오가 존재하므로 자료의 다양한 특성이 반영되도록 시나리오를 그대로 사용하였다. 목표연도 2011년에 대하여 생공용수의 분포 설정은 선형회귀의 이론을 사용하였다. 먼저 과거의 생공용수 수요량 자료(1966년~1998년)를 근거로 하여 선형회귀식을 선정한다. 선형회귀의 기본 가정으로 오차(ϵ_t)의 항은 정규분포를 따른다. 즉 $\epsilon_t \sim N(0, \sigma_{\epsilon}^2)$ 이며 이를 이용하면 목표연도 2011년의 생공용수는 정규분포, 즉 $N(\mu, \sigma_{\epsilon}^2)$ 을 따른다고 할 수 있다. 여기서 μ 은 회귀식에서의 2011년에 해당하는 수요량 추정값이 된다. 농업용수의 분포 설정은 생공용수와는 달리 과거자료의 변동성이 매우 커서 선형회귀분석에 어려움이 있었다. 같은 유역이라 하더라도 자료의 특성으로 인해 지역에서 재배되는 농작물의 성장 속도, 경작의 변화율 등에 의해서 월별 변화가 매우 크고 연간 변화도 무작위로 변하기 때문에 판단된다. 따라서 본 연구에서는 이러한 농업용수의 분포를 설정하기 위하여 maximum entropy theory를 적용하였다. 이는 불확실성에 대한 양을 나타내는 entropy를 이용하여 주어진 제약 조건에 따라 entropy를 최대화시킴으로써 자료의 분

표 1. Maximum Entropy Theory

Given constraints	Assign probability distribution
min, max	Uniform
mean, standard deviation	Normal
mean, max, mode	Triangle
mean, max, min, standard deviation	Beta

포를 설정할 수 있는 방법이다. 표 1에서 보는 바와 같이 가지고 있는 정보의 제약 조건들에 따라서 maximum entropy theory를 사용하여 자료들의 분포를 설정할 수 있다. 따라서 농업용수의 경우에 과거 자료 즉, 32개년(1968년~1998년)의 농업용수의 수요량에 대해서 최대값을 max로, 최소값을 min으로 하고, 수자원장기종합계획(건설교통부, 2000)에서 추정된 2011년 농업용수 수요 추정량을 mode로 두어 삼각 분포로 설정하였다. 현재 우리나라의 물수지 분석에서 사용되는 생공용수, 농업용수들의 회귀율은 각기 65%, 35%로 사용되고 있다. 이러한 회귀율에 대해서도 앞서 언급한 entropy theory를 적용하여 그 분포를 설정하고 범위를 규명하였다. 먼저 생공용수의 회귀율은 Linsley 등(1979)에 의하면 일반적으로 60~70%에 해당한다고 제시하고 있어 이를 최대, 최소 값으로 두고 수자원장기종합계획(건설교통부, 2000)에서 사용된 65%를 mode로 두어 삼각분포로 설정하였다. 농업용수 회귀율의 경우 농업용수 수요량 조사 종합보고서(농림부, 1999) 등 기존 연구사례를 바탕으로 그 범위를 5~65%로 설정하고 수자원장기종합계획(건설교통부, 2000)에서 사용한 35%를 mode로 정하여 이 또한 삼각분포로 설정하였다.

5. 불확실성 분석 및 민감도 분석

목표연도 2011년에 해당하는 소유역별 6개 입력변수의 분포를 설정한 후 32개년 자연유량에 대해 설정된 분포로부터 4개씩 표본을 추출하여 LHS 분석을 시행하였다. 따라서 32개년에 대해 $4096(=4^6)$ 의 년자료를 모의 발생하였다. 생활용수, 공업용수, 농업용수의 분포로부터 나온 표본은 년별 용수 수요량을 나타내므로 이를 다시 월별로 변환하여 분석하여야 한다. 각 용수별 월별 보정계수는 수자원장기종합계획(건설교통부, 2000)에서 제시한 값들을 사용하였으나, 공업용수는 년중 균일하게 사용된다고 보았다. 불확실성 분석으로, 1개년 자연유량 시나리오에 대해 수정 수장기의 단일값과 LHS 분석을 통해 나타난 결과값의 평균을 비교하였고, 또한 각 시나리오에 대해 나타난 결과의 95% 신뢰구간은 물론 수정 수장기의 단일값과도 비교하였다. 또한 전체 32개년에 대한 merged CDF를 구하여 수정 수장기의 단일값(1/30년 가뭄발생)의 위치를 규명하고 이를 분석하였다.

입력변수의 불확실성에 대한 결과값의 민감도를 분석함으로써 결과값에 가장 큰 영향을 미치는 입력변수를 확인하고 그 영향의 정도를 수량화할 수 있다. Monte Carlo simulation의 한 부분으로 수행되는 민감도 해석은 종종 회귀분석에 기본을 두고 있다(Iman, 1981). 본 연구에서는 LHS 분석의 결과로 나타난 물부족량에 대해 이에 따른 입력변수들의 민감도를 표준화회귀계수와 부분상관계수를 적용하여 분석하였으며, 각 입력변수들의 정량화된 불확실성과 민감도의 관계를 규명함으로써 물수지 분석에서의 이들의 중요 우선순위를 제시하였다. 표준화회귀계수와 부분상관계수는 Imam(1981)이 처음 제시한 민감도 분석방법으로, 표준화회귀계수는 모든 입력변수에 대한 상대적인 민감도를 나타내어 주고, 부분상관계수는 다른 입력변수들을 배제한 상태에서 한 요소에 대한 결과값의 민감도를 나타내어주는 지표이다.

6. 결과 분석

금강수계 301 소유역에 대한 불확실성 분석결과와 민감도 분석결과를 표 3과 그림 2~3에 제시하였다. 그림 2에서 보는 바와 같이 전체 32개년별 자연유량 시나리오에 대해 수정 수장기 물부족량과 1개년에 대한 LHS 분석 평균 물부족량 비교시, 수정 수장기의 값이 LHS 분석 평균값보다 큰 경우는 13개년, 같은 결과인 경우는 17개년, 수정 수장기의 물부족량의 값이 적은 경우는 2개년으로 나타났다. 또한 32개년에 대해 대체적으로 수정 수장기의 물부족량이 LHS 분석한 결과의 95% 신뢰구간에 포함되었으며, 수정 수장기에서의 30년 1회 가뭄에 해당하는 물부족량의 값은 merged CDF의 93%에 해당하였다. 그러나 다른 소유역에 대해서 불확실성 분석 결과, 95% 신뢰구간에 대한 단일값의 위치는 각기 다른 상반된 결과를 보였다. 301 소유역에 대해서 단일값은 95% 신뢰구간의 중앙값에 근사하게 도시되었지만, 309 소유역의 경우는 95% 하한선, 316 소유역의 경우는 95% 상한선에 가까이 나타났다. 이는 수정 수장기에서의 단일값의 추정이 과소 및 과대 추정되었음을 의미한다.

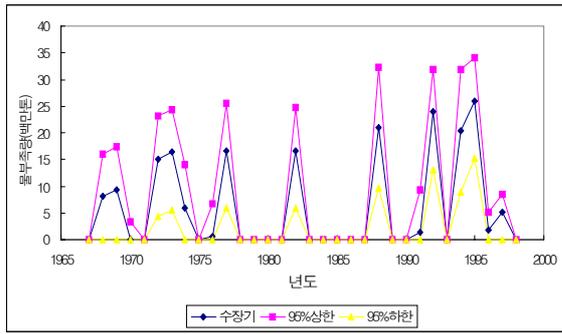


그림 2. 95% 신뢰구간 비교

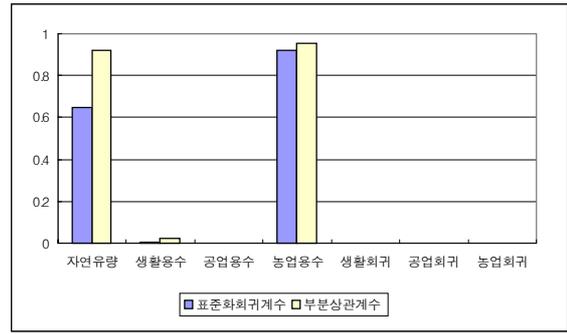


그림 3. 입력변수별 민감도 비교

표 2. 민감도 분석 결과

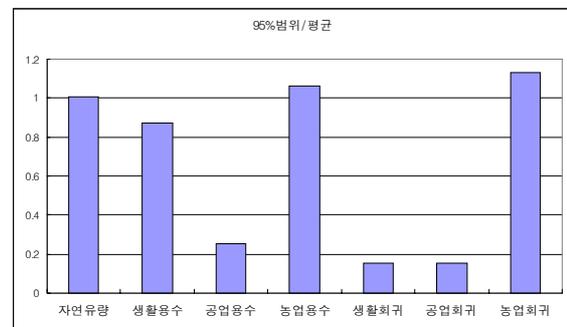
전 체	자연유량	생활용수	공업용수	농업용수	생활회귀	공업회귀	농업회귀
표준화 회귀계수	-0.649	-0.007	0.000	0.922	0.000	0.000	0.000
부분 상관계수	-0.919	-0.025	0.000	0.952	0.000	0.000	0.000
전체 순위	2	3	4	1	4	4	4

그림 3과 표 2에서처럼, 민감도 분석의 결과는 농업용수가 가장 큰 민감도를 나타냈으나 회귀율에 대한 민감도는 0의 값을 나타냈다. 이는 다른 입력변수들에 비해 회귀율은 물수지 분석에서 결과 물부족량에 거의 영향을 끼치지 않음을 의미한다. 또한 자연유량과 생활용수는 물부족량에 대해 음의 상관관계를 보였고, 나머지 입력변수는 양의 상관관계를 나타냈다. 표준화회귀계수와 부분상관계수는 동일 입력변수에 대해 다른 값을 보이는 것은 그들의 절대값에 근거한 입력변수의 중요도 순서는 같은 결과를 보였다.

마지막으로 입력변수들의 불확실성과 민감도 사이의 관계를 알아보기 위해, 우선 입력변수 불확실성의 상대적 크기 비교가 필요하였다. 본 연구에서는 그림 4~5와 같이 입력변수의 변동계수와 95% 신뢰구간을 평균으로 나눈 값을 비교대상으로 선정하여 분석하였다. 그 결과, 변동계수는 공업용수가 가장 크게 나타났고, 평균을 고려한 95% 신뢰구간은 농업용수 회귀율이 가장 크게 나타났다. 다음으로 변동계수와 평균을 고려한 95% 신뢰구간의 비교에서 나타난 불확실성 순위를 그림 5에서와 같이 민감도 순위와 비교하였다. 즉 순위번호가 작을수록 민감도와 불확실성이 큰 경우를 말한다. 결과로써 농업용수 회귀율의 경우 불확실성이 가장 크나, 민감도는 가장 작게 나타났고, 농업용수의 경우는 가장 큰 민감도를 나타내는 반면에 불확실성의 크기는 전체 7개 입력변수 중에서 우선순위가 세 번째로 나타났다.



(a) 변동계수



(b) 95%범위/평균

그림 4. 불확실성의 정량화

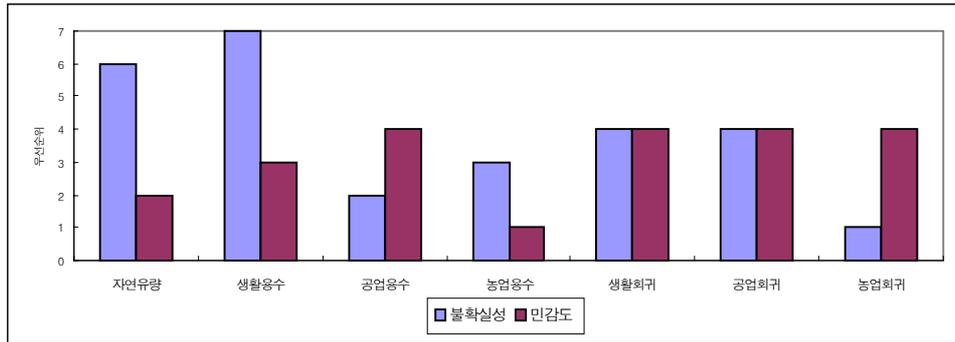


그림 5. 불확실성과 민감도 관계

7. 결론

금강수계 301 소유역에 대해서 불확실성 분석결과, 수정 수장기에서 계산한 물부족 단일값은 95% 신뢰 구간의 중앙값에 근사하게 도시되었고, 각 입력변수들에 대한 민감도 분석에서는 농업용수와 자연유량의 민감도가 가장 크게 나타났으며, 회귀율의 민감도는 0에 가까운 값을 나타냈다. 또한 입력변수들의 불확실성과 민감도를 비교해 본 결과, 농업회귀율의 경우 입력변수들 사이에서 가장 큰 불확실성을 보였으나 결과값에 미치는 영향, 즉 민감도는 거의 없는 것으로 나타났다. 이는 물수지 분석시 가장 변동 가능성이 크고 측정이 어려운 입력변수임에는 분명하나 결과에는 아무런 영향을 미치지 못하고 있음을 의미한다.

본 연구에서 제안한 불확실성 분석기법으로 입력변수의 불확실성으로 인해 야기되는 물부족량의 불확실성 정도를 추정할 수 있었고, 장래 수자원장기종합 계획시 추정되는 물부족량에 대한 불확실성을 정량화하는 방법을 제시하였다. 또한 불확실성과 민감도의 관계 규명으로 입력변수들이 가지는 불확실성과 민감도에 대한 중요도 순위와 이들의 관계를 나타냄으로써 장래 수자원 정책 수립시 투자 우선순위를 정하는 기반을 제공하였다.

8. 감사의 글

본 연구는 21세기 프론티어연구개발사업인 수자원의 지속적 확보기술개발사업단의 연구비지원(과제번호 1-5-1)에 의해 수행되었습니다.

9. 참고 문헌

1. 건설교통부 (2000). **수자원장기종합계획**, 건설교통부, 한국수자원공사.
2. 김남원, 원유승 (2001). "물수지 분석의 불확실성 및 범위", **2001년 한국수자원학회 학술발표회 논문집, 한국수자원학회**, pp.607-612.
3. 농림부 (1999). **농촌용수 수요량조사 종합보고서**, 농림부, 농어촌진흥공사.
4. Iman R. L.(1981). "An Approach to Sensitivity Analysis of Computer Models." *Journal of Quality Technology*, Vol. 13, pp.174-183.
5. Linsley, Ray K., Franzini, and Joseph B. (1979). *Water Resources Engineering*. 3rd ed., McGraw-Hill, New York, NY.
6. Marojolein B. A., Van Asselt, and Jan Rotmans. (2002). "Uncertainty in Integrated Assessment Modelling." *Climatic Change*, Vol. 54, pp. 75-105.