

유역의 수문학적 상사성을 이용한 Nash 모형의 불확실성 평가

Assessment of Uncertainty for Applying Nash's Model Using the Hydrologic Similarity of Basins

성 기 원*

Seong, Kee-Won

Abstract

An approach determining a confidence interval of Nash's observed mean instantaneous unit hydrograph is developed. In the approach, both two parameters are treated as correlated gaussian random variables based on the theory of Box-Cox transformation and the regional similarity relation, so that linear statistical parameter estimation is possible. A parametric bootstrap method is adopted to give the confidence interval of the mean observed hydrograph. The proposed methodology is also applicable to estimate the parameters of Nash's model for un-gauged basins. An application to a watershed has shown that the proposed approach is adequate to assess the uncertainty of the Nash's hydrograph and to evaluate parameters for un-gauged basins.

Keywords : Nash's model, Box-Cox transformation, bootstrap, stochastic IUH, ungauged basinn

요 지

Nash의 관측평균순간단위도의 신뢰구간을 결정하는 기법을 개발하였다. 이 방법은 두 매개변수를 Box-Cox 변환과 유역의 상사성관계식을 이용하여 이변수정규분포의 확률변수화하고 이들의 선형 상관관계를 이용한 통계적 추정과정과 더불어 parametric bootstrap 방법을 이용한 단위도의 신뢰구간 산정 등으로 구성된다. 또한 이 방법은 미계측유역에 대한 단위도 추정에도 이용이 가능한 특징을 갖고 있다. 위천유역에 대하여 제안된 방법을 적용한 결과 제시된 방법론은 단위도의 불확실성을 평가하고 그리고 미계측 유역에 대한 매개변수 추정에 있어서 적절한 대안임을 확인할 수 있었다.

핵심용어 : Nash 모형, Box-Cox 변환, bootstrap, 추계학적 단위유량도, 미계측유역

* 건국대학교 공과대학 토목공학과 조교수

Assistant Prof., Dept. of Civil Engrg., Konkuk Univ., Seoul 143-701, Korea

(E-mail : kwseong@konkuk.ac.kr)

1. 서 론

Nash 모형은 특히 중소규모 유역의 유출을 모의하는 대표적인 개념모형(conceptual model) 가운데 한 가지로 gamma 분포함수를 기반으로 하면서 두개의 매개변수로 구성되는 모형이다. 이 Nash 모형은 구조적인 단순함에도 불구하고 모형의 유연성이 뛰어나고 매개변수의 수문학적 해석이 다양한 관점에서 가능하기 때문에 자주 이용되어왔다. 특히 최근에는 GIS 도구를 통하여 보다 다양하고 전문적인 수문지형학적 분석이 가능해짐에 따라 Nash 모형 매개변수에 대한 효율적인 추정이 이루어지고 있다. 그럼에도 불구하고 추정된 매개변수의 불확실성에 대한 해석이나 미계측 유역에 대한 적용성 문제는 여전히 해결되어야 하는 과제로 남아 있다. 특히 분석된 강우-유출 자료가 부족하여 추정된 매개변수의 분포가 적절한 대표 단위도를 도출하기에 통계학적으로 적절하지 않은 경우는 드물지 않게 발생하지만 이러한 경우에도 단순히 추정치의 평균을 이용하여 관측평균순간단위도(observed mean IUH)를 도출하는 것이 일반적이다. 이 경우 자료가 부족할수록 유도된 단위도의 신뢰성은 당연히 저하되며 결국은 부적절한 유출모의로 이어질 수 있는 개연성이 항상 존재한다. 따라서 자료가 부족할수록 또는 자료의 분산이 클수록 도출된 단위도에 대한 신뢰성을 평가할 수 있는 기법은 중요하다고 할 수 있다.

단위도의 불확실성을 완화하는 문제는 미계측 유역의 경우 보다 절실하다. 일반적으로 수문학적으로 동질성을 갖고 있다고 판단되는 미계측 유역에 대한 단위도 도출문제는 지역분석(regional analysis)이라는 방법을 통하여 주로 이루어지고 있으나 이러한 문제에 대한 접근 방법도 기존의 회귀분석방법을 보완할 수 있는 통계적인 방법을 통하여 잠재적인 불확실성을 파악하고 완화하도록 하는 것이 적절하다. 이에 대하여 성기원(1997)은 유역의 상사성이라는 수문학적 특성을 통하여 지역분석을 실시하고 미계측 유역에 대한 Nash 모형의 매개변수를 추정한 바 있으나 여전히 불확실성에 관련된 인자가 적지 않으며 개선의 여지가 있다.

반면 양질의 자료가 풍부한 단일 유역의 경우에도 모형의 불확실성은 존재한다. 이러한 불확실성은 주로 추계학적 접근을 통하여 해결하는 것이 가장 보편적인 연구 추세인데 특히, Unny(1984)는 Nash cascade를 추계학적 cascade로 확장하는 연구를 제안한 바 있으며, Sarino와 Serrano(1990)은 단일유역의 Nash

단위도를 도출하는데 있어서 매개변수를 확률변수(random variable)로 가정하여 추계학적 미분방정식(Stochastic Differential Equation, 이하 SDE)을 해결하는 방법을 제시한 바 있다. 이 방법은 Nash 단위도의 불확실성을 해석적으로 표현할 수 있다는 장점이 있으나 매개변수 가운데 규모계수(scale parameter)를 정수로 고정시켜서야하는 단점이 있으며 도출된 단위도의 분산을 효과적으로 이용하는 방안을 제시하지 않았다. 반면 Hjelmfelt와 Wang(1994) 그리고 Lin과 Wang(1996)은 Sarino와 Serrano(1990)의 모형보다 보다 일반적인 형태의 SDE를 제안하였지만 이들 역시 저류상수(storage coefficient)만을 확률변수로 취급한 모형이다. 그런데 이들의 연구에서 검증된 유역의 규모계수는 저류상수의 분산에 비하여 극도로 적은 분산값을 갖고 있어서 상대적으로 저류상수만을 확률변수로 취급하여도 별 무리가 없어 보이나 우리나라의 상당한 유역에서는 규모계수나 저류상수의 분산 규모는 큰 차이를 보이지 않으며 따라서 제안된 모형들은 적용성에 문제가 있다.

이상과 같은 문제점과 한계를 고려할 때, 자료가 부족하거나 추정 매개변수의 변동성이 큰 유역의 Nash 단위도의 불확실성을 평가하고 신뢰구간을 수립하거나 또는 계측 유역에 대한 신뢰성 있는 단위도 추정방법을 제시하기 위해서는 기존의 방법과는 다른 접근이 요구된다. 이에 본 연구에서는 Nash 모형의 두개의 매개변수를 모두 확률변수로 간주하여 상사성이 있는 유역간의 자료를 상사성 공식에 부합하도록 통합하여 적절한 변환을 통해 결합확률밀도함수(joint probability density function)를 도출한 후 parametric bootstrap 방법으로 자료를 발생시키고 이 자료를 통하여 90% 신뢰도를 갖는 단위도의 구간을 설정하는 연구를 수행하였다. 그리고 미계측 유역의 매개변수 추정을 위하여 변환된 매개변수간의 선형관계식을 도입하여 유역대표 규모계수에 대한 저류상수의 조건부확률을 계산하고 도출된 값들의 적용성을 평가하였다. 대상유역은 IHP 대표유역인 위천 소유역을 선택하였고 연구의 공정성을 위하여 윤석영과 홍일표(1994)의 자료를 통해 연구를 수행하였다.

2. 모형개발

2.1 유역의 상사성과 매개변수의 변환

Nash 모형이 나타내는 순간단위도는 gamma 분포

형 함수로서 식(1)로 나타낸다.

$$h(t) = \frac{\left(\frac{t}{k}\right)^{(n-1)} \exp\left(-\frac{t}{k}\right)}{\Gamma(n)k} \quad (1)$$

여기서 $h(t)$ 는 Nash 순간단위도이며 n 은 규모계수 그리고 k 는 저류상수로서 Nash 모형의 매개변수이다. 매개변수 산정에 있어서 n 과 k 의 독립성을 가정할 경우 m 개의 홍수사상에 대하여 추정된 매개변수는 벡터 $N=[n_1, n_2, \dots, n_m]$ 과 $K=[k_1, k_2, \dots, k_m]$ 로 나타낼 수 있다. 일반적인 경우 특정구역의 대표응답을 도출하기 위해서는 m 의 크기에 관계없이 N 과 K 의 평균값을 취하여 관측평균단위도를 작성하는데 이 경우 N 과 K 의 분산에 따른 순간단위도의 불확실성은 고려되기 어려우며 특히 분석된 사상의 수가 적은 경우 도출된 관측평균단위도의 대표성은 의심스럽다. 이를 보완하고자 본 연구에서는 소유역간에 상사성이 존재하여 매개변수 간의 관계가 상사성 경험식을 만족한다고 판단되는 경우 각 소유역간의 자료를 통합하여 확률변수 N 과 K 간의 결합확률밀도함수를 도출하는 방법을 제시하고자 한다. 성기원(1997)에 의하면 유역내 상사성이 있는 소유역 사이에는 식(2)와 같은 관계가 있음을 보인다.

$$\frac{nk}{Tc + nk} = C \quad (2)$$

여기서 Tc 는 유역의 도달시간, C 는 상수인데 따라서 식(2)는 유역의 도달시간과 Nash 매개변수의 관계는 수문학적으로 동질성이 있는 유역간에 일정한 규칙이 있다는 것을 나타낸다. 그런데 식(2)에서 n 과 k 를 분리할 경우 식(3)과 같이 나타낼 수 있다.

$$n = \frac{Tc}{k} \frac{C}{1-C} \propto \frac{Tc}{k} \quad (3)$$

식(3)은 n 과 Tc/k 사이에 선형적 관계가 있음을 보이고 있으며 수문학적으로 상사성이 있는 유역에는 일정한 기울기가 있음을 나타낸다. 따라서 n 과 Tc/k 는 서로 종속되어 있지만 실제자료에서 C 는 확률변수일 것이므로 n 과 Tc/k 를 결합확률분포를 갖는 확률변수로 가정하여도 무방하다. 그러나 n 과 Tc/k 이 속해있는 모집단의 분포를 알지 못하는 상황에서 확률변수간의 결합확률밀도 함수를 구하기 위해서는 n 과 Tc/k 의 분포는 정규분포로 변환하는 것이 유리하다. 이를 위하여

본 연구에서는 n 과 Tc/k 자료를 Box-Cox 변환(Box and Cox, 1964)하여 근사적으로 정규분포를 따르도록 하였다. 즉 벡터 $N=[n_1, n_2, \dots, n_m]$ 과 상사성 공식에 의한 변환자료 벡터 $Q=[Tc/k_1, Tc/k_2, \dots, Tc/k_m] \equiv [q_1, q_2, \dots, q_m]$ 의 원소는 각각 식(4)와 식(5)와 같은 Box-Cox 변환 공식을 통해 정규인자(normal factor)화 된다.

$$n_i(\lambda_1) = \begin{cases} \frac{(n_i^{\lambda_1} - 1)}{\lambda_1} & \text{if } \lambda_1 \neq 0 \\ \ln(n_i) & \text{if } \lambda_1 = 0 \end{cases} \quad (4)$$

$$q_i(\lambda_2) = \begin{cases} \frac{(q_i^{\lambda_2} - 1)}{\lambda_2} & \text{if } \lambda_2 \neq 0 \\ \ln(q_i) & \text{if } \lambda_2 = 0 \end{cases} \quad (5)$$

여기서 λ_1 과 λ_2 는 각각 확률변수 N 과 Q 의 변환계수(coefficient of transformation)이다. 따라서 $n_i(\lambda_1)$ 및 $q_i(\lambda_2)$ 는 각각 변환된 확률변수의 벡터 $N(\lambda_1)$ 과 $Q(\lambda_2)$ 의 i 번째 변환값을 나타낸다. 벡터 $N(\lambda_1)$ 과 $Q(\lambda_2)$ 가 정규분포를 나타내기 위해서는 정규분포를 실현할 수 있는 적절한 λ_1 과 λ_2 의 추정이 필요한데 일반적으로 통계학적으로 이용되는 방법은 정규분포로 표시되는 각각의 주변밀도함수(marginal density function) 대한 log-likelihood 함수를 최대화함으로써 λ_1 과 λ_2 를 각각 추정하는 '분리추정방법'이다. 그런데 통계학 분야에서 짝을 이루는 자료(paired data sample)간의 상관성을 바탕으로 하는 자료의 선형구조에 대한 연구가 진행되면서 Hernandez와 Johnson (1980)은 변환계수를 동시에 추정하는 방법이 주변밀도함수를 이용하여 각 변환계수를 추정하는 방법보다 통계적으로 효율적(efficient)이라 하였으며 또한 Zou와 Hall(2001)은 샘플의 수가 적을수록 각 방법에 의한 결과는 차이가 커지며 이 경우 '동시추정방법'이 우월함을 보인다. 이에 따라 본 연구에서는 변환계수를 추정하기 위하여 일반적으로 이용되는 분리추정방법 대신 동시추정방법을 적용하였는데 이 때 이용되는 joint log-likelihood function은 이변수정규분포(bivariate normal distribution)에 상응하는 함수로서 식(6)으로 나타낸다.

$$f(\lambda_1, \lambda_2 | N^*, Q^*) = \frac{m}{2} \ln(1 - \rho^2) - \frac{1}{2(1 - \rho^2)} \sum_{i=1}^m (n_i^* - 2\rho n_i^* q_i^* + q_i^{*2}) + (\lambda_1 - 1) \sum_{i=1}^m \ln n_i + (\lambda_2 - 1) \sum_{i=1}^m \ln q_i - m \ln(2\pi\sigma\tau) \quad (6)$$

여기서 $N^* = [n_1^*, n_2^*, \dots, n_m^*]$ 과 $Q^* = [q_1^*, q_2^*, \dots, q_m^*]$ 은 각각 $N(\lambda_1)$ 과 $Q(\lambda_2)$ 의 정규화된 벡터로서 짝을 이룬 변환된 샘플 $(n_i(\lambda_1), q_i(\lambda_2))$ 이 평균 (μ, ν) 와 표준편차 (σ, τ) 그리고 상관계수 ρ 로 구성된 이변수정규분포를 따른다고 할 때 N^* 과 Q^* 의 각 원소는 식(7)을 통해 계산된다.

$$n_i^* = \frac{(n_i(\lambda_1) - \mu)}{\sigma} \quad (7)$$

$$q_i^* = \frac{(q_i(\lambda_2) - \nu)}{\tau}$$

따라서 식(6)의 (n_i^*, q_i^*) 는 상관계수 ρ 를 가진 표준이변수정규분포(standard bivariate normal distribution)의 변량이다. 최적 변환계수를 얻기 위해서 식(6)을 최대화 하여 계산된 변환계수 (λ_1, λ_2) 의 추정값을 $(\hat{\lambda}_1, \hat{\lambda}_2)$ 라 할 때 이를 식(4)와 (5)에 대입함으로써 n 과 Tc/k 자료에 대한 근사적 정규분포 변환은 완성된다.

2.2 조건부확률 계산과 역변환

추정된 $(\hat{\lambda}_1, \hat{\lambda}_2)$ (이후 수식에 대한 가독성을 배려하여 $\hat{\lambda}_1$ 및 $\hat{\lambda}_2$ 는 각각 λ_1 과 λ_2 로 표기함)를 적용함으로써 확률변수 N 과 Q 가 각각 정규분포 $N(\lambda_1)$ 과 $Q(\lambda_2)$ 로 근사화 되었다면 이변수정규모형에 이용 가능한 통계학적 이론을 적용할 수 있다. 즉 n 이나 Tc/k 의 값이 주어졌을 때 이에 상응하는 Tc/k 와 n 의 변환된 조건부평균은 주어진 자료를 Box-Cox 변환하고 이를 식(8)이나 (9)에 적용함으로써 얻을 수 있다.

$$E[N(\lambda_1) | Q(\lambda_2) = \delta_1] \quad (8)$$

$$= E[N(\lambda_1)] + \rho \frac{std[N(\lambda_1)]}{std[Q(\lambda_2)]} [\delta_1 - E[Q(\lambda_2)]]$$

$$E[Q(\lambda_2) | N(\lambda_1)]$$

$$= \delta_2 = E[Q(\lambda_2)] + \rho \frac{std[Q(\lambda_2)]}{std[N(\lambda_1)]} [\delta_2 - E[N(\lambda_1)]] \quad (9)$$

식(8)과 (9)에서 $E[\]$ 는 평균을 $std[\]$ 는 표준편차를 나타내며 δ_1 과 δ_2 는 각각 변환된 임의의 Tc/k 와 n 을 의미한다. 그리고 식(8)과 (9)의 결과는 식(10)과 식(11)로 표현된 식(4)와 (5)의 역변환을 적용함으로써 최종적인 기댓값 Tc/k 와 n 으로 전환된다.

$$n = (E[N(\lambda_1) | Q(\lambda_2)] + 1)^{(1/\lambda_1)} \quad (10)$$

$$Tc/k = (E[Q(\lambda_2) | N(\lambda_1)] + 1)^{(1/\lambda_2)} \quad (11)$$

여기서 식(10)과 식(11)은 추정된 Nash 단위도의 불확실성 평가와 미계측 유역에 대한 매개변수 추정에 이용된다.

2.3 불확실성 평가

앞서 확률변수 $N(\lambda_1)$ 과 $Q(\lambda_2)$ 가 이변수정규분포를 이룬다고 하였으므로 이 분포에 대한 통계학적 성질을 이용하여 각 확률변수의 신뢰영역(confidence region)을 추정할 수 있다. 따라서 샘플의 개수가 적고 또한 분포가 파악되어있는 상태에서는 비교적 널리 알려져 있는 Efron과 Tibshirani(1993)의 non-parametric bootstrap 방법보다 parametric bootstrap 방법이 적절하며 본 연구에서는 이 방법 가운데 Monte-Carlo testing procedure(Van der Heijden 등, 1997)를 적용하였다. 이 계산을 위하여 평균 0 그리고 단위 분산을 갖는 두개의 정규변량(normal variate) Z_1 과 Z_2 를 발생시키고 계산된 상관계수 ρ 를 이용하여 식(12)를 구성한다.

$$X_1 = Z_1 \quad (12)$$

$$X_2 = \rho Z_2 + (1 - \rho^2)^{1/2} Z_1$$

여기서 X_1 과 X_2 는 각각 $n(\lambda_1)$ 과 $q(\lambda_2)$ 의 상관계수를 고려한 가중치가 되며 선형변환 관계를 고려할 때 식(13)을 통하여 불확실성을 평가할 수 있는 샘플들을 발생시킬 수 있다.

$$n(\lambda_1) = \mu + X_1 \sigma \quad (13)$$

$$q(\lambda_2) = \nu + X_2 \tau$$

여기서 평균 (μ, ν) 와 표준편차 또는 샘플을 발생하기 위한 표준오차(standard error) (σ, τ) 는 식(6)에서 설명된 값이다. 그리고 식(12)와 식(13)을 난수발생기를 이용하여 계산(1,000회)하고 역변환을 시행한 후 해

표 1. 위천 유역의 유역별/사상별 도달시간 및 Nash 모형 추정 매개변수

유역	도달시간 T_c 식(14)		사상	Nash 매개변수 추정치		T_c/k
	T_c	A		n	k	
동곡	4.29	33.8	DT861	2.73	2.13	2.01
			DT891-1	3.12	2.13	2.02
			DT892	2.74	2.71	1.59
			DT903	2.58	2.89	1.49
			DT905	2.44	2.88	1.49
			DT921-1	2.50	2.61	1.64
			DT923	2.10	3.55	1.21
			DT924	1.93	3.31	1.30
			평균	2.52	2.78	1.59
고노	4.77	109.8	DT861	2.75	2.64	1.81
			DT891-1	2.77	2.95	1.62
			DT892	3.10	2.67	1.79
			DT903	2.15	3.74	1.28
			DT904	1.87	3.47	1.38
			DT905	2.23	3.42	1.40
			DT921-1	2.56	2.61	1.83
			평균	2.49	3.07	1.58
미성	4.97	171.7	DT891-1	3.04	2.86	1.74
			DT892	3.46	2.55	1.95
			DT904	1.88	3.65	1.36
			DT905	2.21	3.57	1.39
			DT921-1	2.36	2.71	1.83
			평균	2.60	3.07	1.65
병천	5.23	300.3	DT892	3.47	3.00	1.74
			DT905	2.34	3.75	1.39
			DT921-1	2.43	2.69	1.94
			DT922	1.58	4.23	1.24
			평균	2.46	3.42	1.58
무성	5.44	472.5	DT891-1	3.34	3.68	1.48
			DT904	1.93	3.57	1.52
			DT921-1	2.47	2.70	2.02
			DT921-2	2.33	3.83	1.42
			DT922	1.67	3.97	1.37
평균	2.35	3.55	1.56			

당 유역의 T_c 를 계산하고 n 과 k 를 도출한다. 그리고 이 결과를 식(1)에 대입하면 모의된 횡수만큼의 단위도를 얻을 수 있는데, 각 시간별로 모의된 값들을 크기

순으로 정리하고 상하위 10%를 제거하면 해당 시간별 90% 신뢰구간을 계산할 수 있다.

3. 모형의 적용

3.1 대상구역 및 자료의 선정

본 연구의 수행을 위하여 위천 유역을 대상구역으로 선정하였다. 위천유역은 윤석영과 홍일표(1994), 이정식(1994), 성기원(1997) 그리고 성기원(1999) 등에 의하여 Nash 모형의 적용성과 지역분석 그리고 상사성 등을 평가받은 바 있는 연구 수행을 위한 최적의 지점이다. 본 연구를 수행하기 위한 기초자료는 강우-유출자료로부터 도출된 Nash 모형의 매개변수 n 과 k 그리고 유역의 도달시간 T_c 이다. 연구의 객관성을 확보하기 위하여 필요한 자료는 윤석영과 홍일표(1994)의 분석자료를 이용하였는데 특히 T_c 는 식(14)로 나타내어지는 제안식을 이용하였다.

$$T_c = 2.17A^{0.17} \quad (14)$$

여기서 T_c 및 유역면적 A 의 차원은 각각 hr와 km^2 이다. 계산을 위한 자료를 표 1에 수록하였다.

표 1을 살펴보면 n 과 T_c/k 는 사상별로 차이를 보여주고 있으나 유역간의 차별성은 발견할 수 없다. 따라서 상사성 관계를 나타내는 식(3)은 적용성이 있다고 판단되며 따라서 모든 유역에 대한 n 과 T_c/k 는 통합하여 단일한 모집단을 갖는다고 가정하기로 한다. 이에 따라 5개 유역 전체의 매개변수 추정값 n 과 T_c/k 를 각각 확률변수 N 과 Q 의 실현값(realization)으로 포함하고 식(6)을 적용하여 Box-Cox 변환을 위한 변환계수 (λ_1, λ_2) 를 추정하였다. 그리고 추정된 (λ_1, λ_2) 를 식(4)와 (5)에 적용하여 근사화된 정규변량을 얻을 수 있었는데 이 Box-Cox 변환에 의하여 변환된 자료의 주요 통계 특성값을 표 2에 수록하였다.

표 2를 살펴보면 $(\lambda_1, \lambda_2) = (-0.083, -1.294)$ 를

표 2. Box-Cox 변환 후 자료의 기본 통계량

확률변수	$N(\lambda_1)$	$Q(\lambda_2)$
변환계수	$\lambda_1 = -0.083$	$\lambda_2 = -1.294$
평균	0.856	0.334
표준편차	0.194	0.089
왜곡도계수	-0.173	-0.117
첨도계수	2.425	1.841
상관계수	0.704	

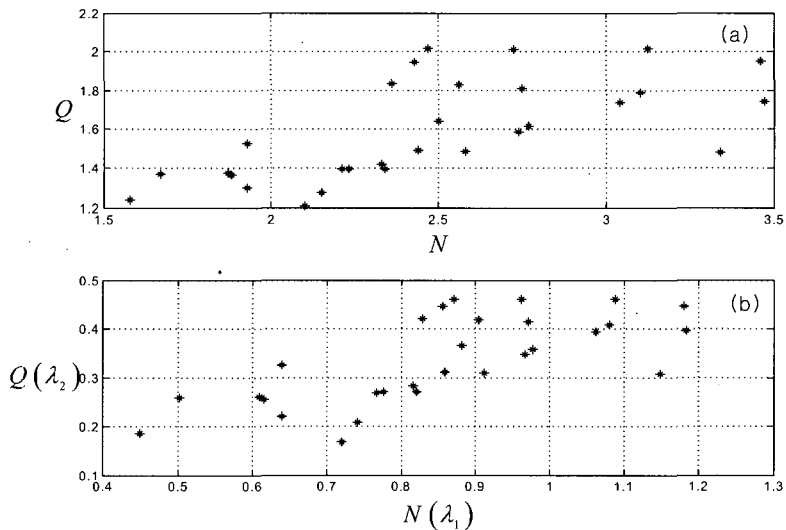


그림 1. Box-Cox 변환 전후의 n 과 T_c/k 간의 상관관계

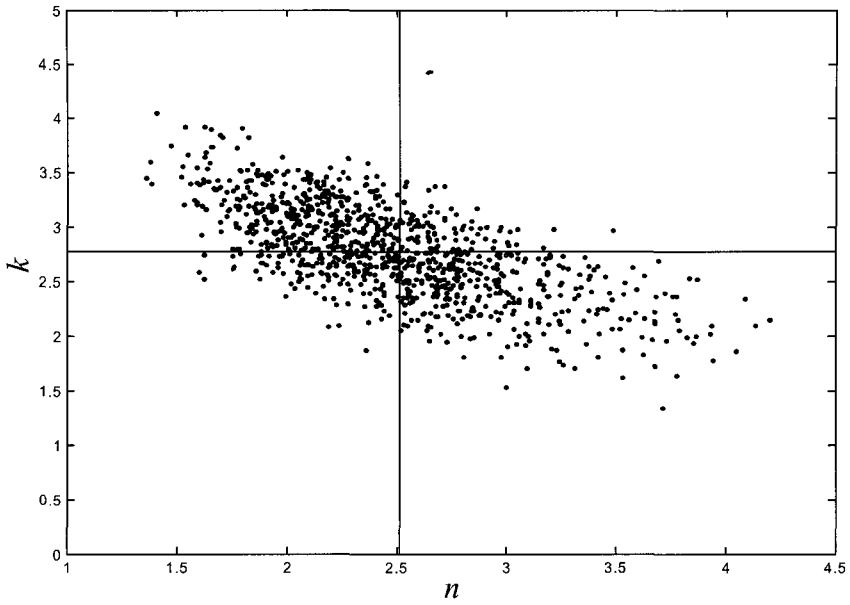


그림 2. Monte-Carlo testing procedure에 의하여 발생된 매개변수 실현값의 분포

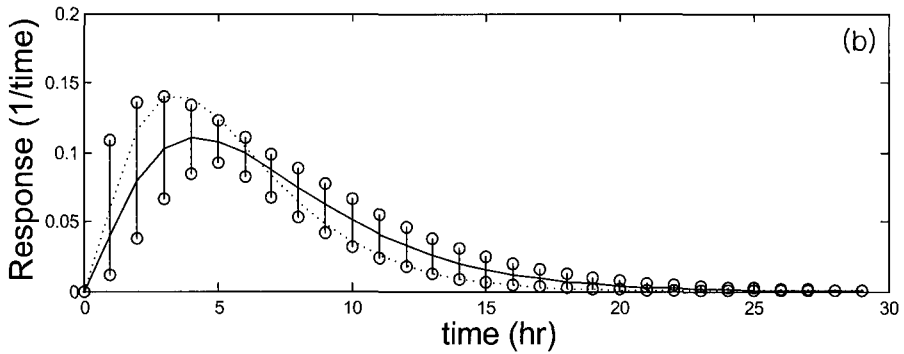
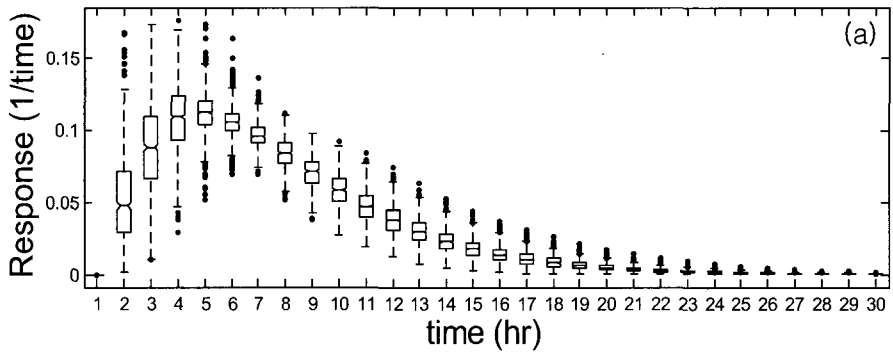


그림 3. 동곡구역의 관측평균순간단위도와 신뢰구간

적용하여 Box-Cox 변환을 실행하면 확률변수 $N(\lambda_1)$ 과 $Q(\lambda_2)$ 의 왜곡도계수(coefficient of skewness)가 실용적인 측면에서 0에 근접하였음을 발견할 수 있는데 이로부터 변환된 자료는 이변수정규분포를 근사적으로 따른다고 가정할 수 있다. 그런데 첨도계수(coefficient of kurtosis)는 $Q(\lambda_2)$ 의 경우 이상치(ideal value)인 3에 못 미치는데 이 값은 자료의 양에 깊이 의존하는 값으로 변동성이 매우 큰 값으로 알려져 있다. 따라서 이 정도의 첨도계수 값은 고려할만한 문제가 되지 않으며 오히려 자료는 적절하게 변환되었다고 판단할 수 있는 근거를 제공한다. 그리고 이 때 두 확률변수 사이의 상관관계수 ρ 는 0.704로서 비교적 뚜렷한 상관관계를 보여준다고 할 수 있는데 이를 그림 1을 통하여 확인할 수 있다.

그림 1(a)는 Box-Cox 변환 이전의 n 과 Tc/k 즉 확률변수 N 과 Q 의 상관관계를 그림 1(b)는 변환 후 $N(\lambda_1)$ 과 $Q(\lambda_2)$ 의 상관관계를 나타낸다. 그림에서 볼 수 있듯이 변환 후 상관성은 개선되었으며 자료의 분포도 집중화 되었다.

3.2 계측 유역에 대한 매개변수 추정의 불확실성 완화

동곡유역은 대상 소유역 가운데 가장 많은 8개의 매

개변수 추정자료를 갖고 있다. 그렇지만 이는 통계학적으로 충분한 양의 샘플이라 할 수 없으며 따라서 이 유역의 자료를 산술평균하여 추정된 관측평균순간단위도의 신뢰성은 그다지 높지 않다고 가정할 수도 있다. 그렇지만 매개변수의 민감도분석 이외에는 단위도의 불확실성을 평가할 수 있는 통계학적 방법은 별로 없다. 이러한 경우 식(3)의 관계를 이용하여 샘플의 수를 유역 규모로 확장하고 Box-Cox 변환을 통해 매개변수간의 상관관계를 파악하여 Monte-Carlo testing procedure를 적용하면 관측평균순간단위도의 신뢰성을 유역전체의 규모에서 간접적으로 평가할 수 있을 것이다. 이를 위하여 식(12)와 (13)을 이용하여 발생시킨 1,000개의 샘플을 역변환 하고 동곡의 Tc 를 식(14)를 통해 계산하고 적용한 후 n 과 k 의 관계를 그림 2에 도시하였다.

그림 2에서 교차실선(cross hairline)은 동곡유역의 8개 자료에 대한 매개변수 추정값의 평균을 그리고 점들은 1,000개의 매개변수 실현값을 나타낸다. 동곡의 경우 매개변수의 평균값은 실현값들의 중심에 위치하는 적절한 양상을 보여주고 있는데 점들은 평균값의 우하단부에서 약간 산포되어있는 경향을 보여준다. 이들 실현값들을 식(1)에 적용하여 1,000개의 단위도를 산출한 후 단위도 총거의 분포를 1시간 단위의 box-plot으로

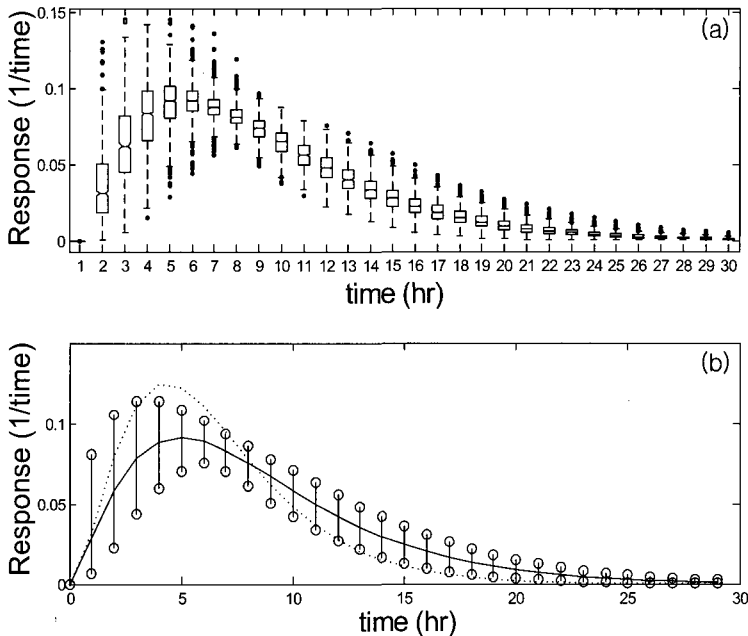


그림 4. 병천유역의 관측평균순간단위도와 신뢰구간

구현하면 그림 3을 얻을 수 있다.

그림 3(a)를 살펴보면 IQR(inter-quantile range)은 유출초기에 넓게 분포되는데 이는 유출초기의 불확실성이 상대적으로 크다는 것을 의미한다. 또한 점으로 표시된 이상치(outlier)들의 분포도 중앙값에 비하여 상향으로 비대칭인 경향을 유출초기부터 침투부 주위까지 발견할 수 있다. 이러한 증거좌표의 분포를 상하위 10%를 절삭하여 불확실성을 90% 신뢰구간으로 표시한 것이 그림 3(b)이다. 그림 3(b)에서 실선으로 표시한 단위도는 동곡유역의 관측평균순간단위도이다. 이 관측평균순간단위도의 증거는 그림 3(a)의 중앙값과 거의 일치하는데 이는 동곡의 자료의 질이 통계학적으로 분석하기에 적절하고 자료의 발생이 잘 이루어졌다는 것을 의미한다. 따라서 그림 3(b)의 신뢰구간은 동곡유역에 유효하며 이를 통하여 관측평균순간단위도의 불확실성을 평가할 수 있다. 따라서 그림을 살펴볼 때 90%의 신뢰정도로 동곡유역의 순간단위도의 최대 침투는 1시간 빠르게 발생할 수도 있는데 이 경우 단위도의 침투값은 관측평균순간단위도의 그것보다 26%를 웃돈다. 따라서 동곡유역에 대한 수공구조물 설계시 안전율을 고려한다면 Nash 모형의 매개변수의 추정치를 표 1에 나타난 평균값 대신 90% 신뢰정도에서 최대 침투값을 보인 매개변수의 조합 $(n, k) = (2.65, 2.08)$ 을 대안으로 선택할 수 있을 것이다. 이 경우에 대한 단위도를 그림 3(b)에 점선으로 나타내었다.

동곡유역과 달리 병천유역은 표 1에서 볼 수 있듯이 단지 4개의 매개변수 추정자료가 있다. 동곡유역에 실시한 분석을 병천유역에 시행한 후 단위도의 신뢰구간과 관측평균순간단위도 그리고 90% 신뢰수준에서 최대 침투값을 보이는 단위도를 그림 4에 도시하여 보았다.

그림 4(a)와 (b)를 비교하면 발생시킨 자료들에 의한 중앙값과 관측평균순간단위도는 동곡유역의 경우와 마찬가지로 거의 일치하는 경향을 보여준다. 따라서 상사성을 이용한 parametric bootstrap 방법은 타당성 있다고 판단된다. 그리고 전체 유역규모에서 평가하는 신뢰도 구간 해석 역시 적절하다. 그런데 90% 신뢰수준에서 최대 침투값을 나타내는 매개변수의 조합을 이용할 경우 도출된 단위도가 인접 시간의 신뢰구간의 상한값(upper limit)를 크게 초과한다. 이러한 현상은 선택된 매개변수가 90% 신뢰구간 외부에 단위도 증거를 발생시키기 때문으로서 적용성 측면에서 적절하다고 판단할 수 없다. 따라서 병천의 경우는 3hr에 최대 침투를 발생시킨 매개변수 보다 4 내지 5hr에서 최대 침투값을 발생시키는 매개변수의 조합 $(n, k) = (2.07, 3.07)$ 을 안전율을 고려한 설계에서 적용하는 것이 바람직하다고 사료된다. 이처럼 본 연구에서 제안된 방법론은 계측유역의 관측평균순간단위도의 적용 타당성을 유역규모에서 검증할 수 있으며 아울러 위험도를 고려한 설계 대안을 제시할 수 있기 때문에 효용성 측면에서 우월하다고 판단된다.

가지로 거의 일치하는 경향을 보여준다. 따라서 상사성을 이용한 parametric bootstrap 방법은 타당성 있다고 판단된다. 그리고 전체 유역규모에서 평가하는 신뢰도 구간 해석 역시 적절하다. 그런데 90% 신뢰수준에서 최대 침투값을 나타내는 매개변수의 조합을 이용할 경우 도출된 단위도가 인접 시간의 신뢰구간의 상한값(upper limit)를 크게 초과한다. 이러한 현상은 선택된 매개변수가 90% 신뢰구간 외부에 단위도 증거를 발생시키기 때문으로서 적용성 측면에서 적절하다고 판단할 수 없다. 따라서 병천의 경우는 3hr에 최대 침투를 발생시킨 매개변수 보다 4 내지 5hr에서 최대 침투값을 발생시키는 매개변수의 조합 $(n, k) = (2.07, 3.07)$ 을 안전율을 고려한 설계에서 적용하는 것이 바람직하다고 사료된다. 이처럼 본 연구에서 제안된 방법론은 계측유역의 관측평균순간단위도의 적용 타당성을 유역규모에서 검증할 수 있으며 아울러 위험도를 고려한 설계 대안을 제시할 수 있기 때문에 효용성 측면에서 우월하다고 판단된다.

3.3 미계측 유역에 대한 순간단위도 추정

미계측 유역에 대한 단위도를 추정하기 위해서는 지역분석 즉 유역간의 상사성이라는 정보를 활용하여야 한다. 본 연구에서는 제안된 방법을 이용하여 미계측 유역에 대한 매개변수 추정하는데 있어서 두 매개변수 n 과 k 에 대한 정보가 없다고 가정하고 다른 상사성이 있는 모든 유역의 n 을 산술평균하여 이를 미계측 유역의 n 으로 가정한 후 식(9)를 이용하여 k 를 추정하였다. 만일 n 에 대한 다른 정보나 지형학적 관계로부터 도출될 수 있는 공식이 있다면 이를 이용해도 관계없으나 동곡유역의 경우 소유역별로 n 값의 변화경향이 발견되지 않았으므로 본 연구에서는 n 의 추정값을 산술평균값으로 결정하였다. 분석을 위한 미계측 유역으로 미성과 무성유역을 선택하였는데 당연히 상사성 평가과정이나

표 3. 미계측 유역 적용을 위한 Box-Cox 변환자료의 통계량

유역	미성을 제외한 4개 유역		무성을 제외한 4개 유역	
	$N(\lambda_1)$	$Q(\lambda_2)$	$N(\lambda_1)$	$Q(\lambda_2)$
변환계수	$\lambda_1 = -0.088$	$\lambda_2 = -1.579$	$\lambda_1 = 0.027$	$\lambda_2 = -0.926$
평균	0.919	0.310	0.915	0.365
표준편차	0.221	0.080	0.204	0.108
왜곡도계수	-0.208	-0.126	-0.193	-0.200
침도계수	2.527	1.946	2.596	1.735
상관계수	0.679		0.796	

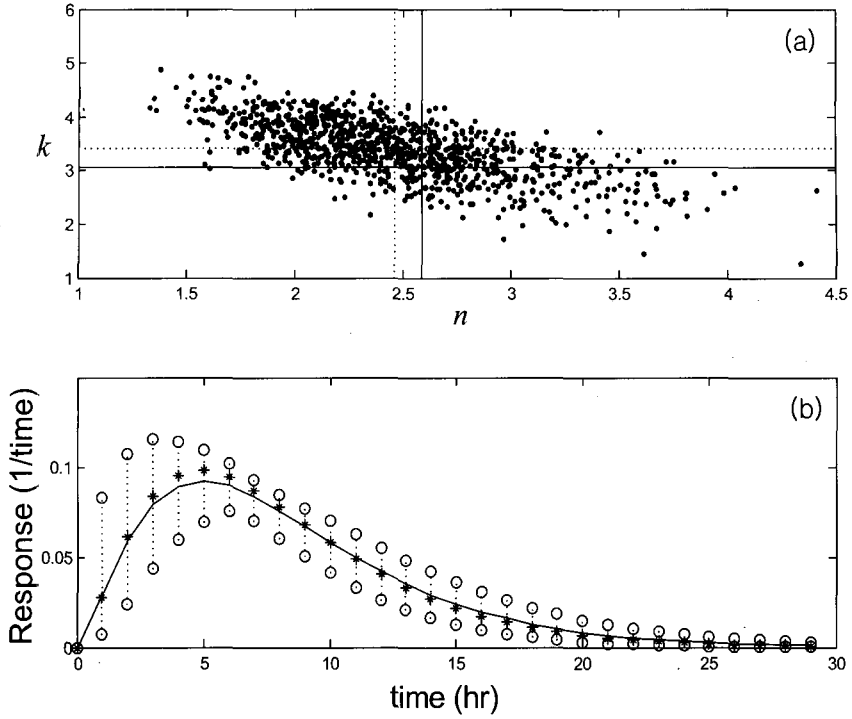


그림 5. 미성을 미계측 유역으로 가정할 경우의 추정 순간단위도와 신뢰구간

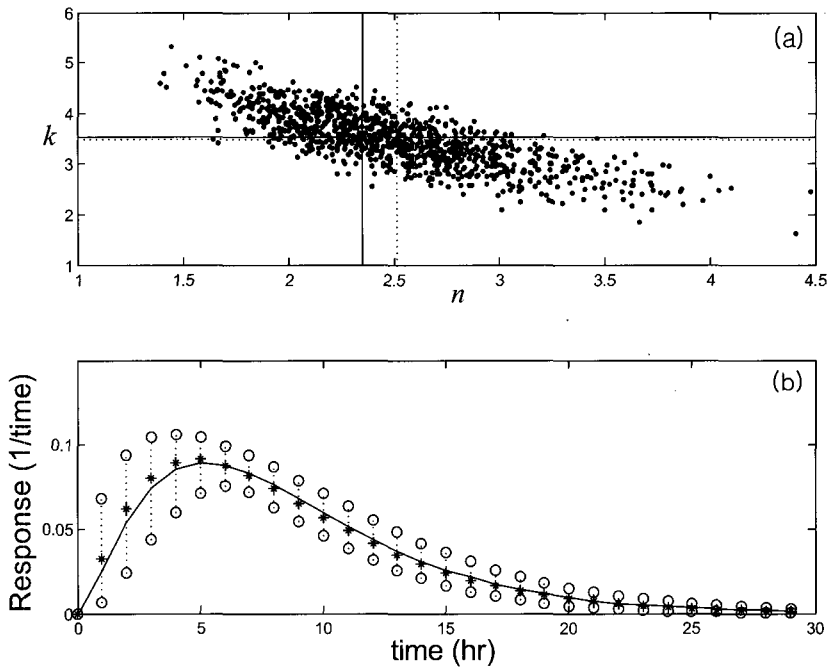


그림 6. 무성을 미계측 유역으로 가정할 경우의 추정 순간단위도와 신뢰구간

Box-Cox 변환과정에서 해당구역의 자료는 제거하였다. 이 두 구역의 자료를 각각 제거한 경우에 대한 Box-Cox 변환 결과를 표 3에 수록하였다.

표 3을 살펴보면 각각의 경우 이변수정규분포를 가정할 수 있는 수준의 통계값을 보여주고 있다. 그런데 표 3을 표 2와 비교할 경우 무성구역의 자료를 제거할 경우 상관계수가 0.704에서 0.796으로 증가하여 상관관계가 상승되는 것을 볼 수 있는데 이로부터 무성구역은 위천구역내 소구역 가운데 상사성 관계에서 가장 벗어나 있는 구역으로 판단할 수 있다. 그렇지만 미성구역은 제거될 경우 상관계수가 0.679로 감소하는데 이로 미루어 미성구역은 상사성 관계를 확고하게 하는 데 기여하는 구역으로 보인다. 미성구역의 자료를 제거함으로써 미성을 미계측 구역으로 가정하였을 때 추정되는 단위도와 매개변수의 변동성을 그림 5를 통하여 살펴본다.

신뢰구간 설정을 위하여 N 과 Q 의 실현값을 n 과 k 로 환산한 그림 5(a)는 미성구역의 상대적 중요성을 보여준다. 그림에서 교차실선은 미성구역의 관측 평균값 $(n, k) = (2.59, 3.07)$ 을 교차점선은 분포의 중심값 $(n, k) = (2.46, 3.40)$ 을 나타낸다. 앞서 언급한 바 있듯이 미성구역의 자료가 제거됨으로써 상관관계가 적지않게 저하되었는데 이는 두 교차선 사이가 멀어질수록 정도가 심한 것을 의미한다. 그렇지만 그림 5(b)를 보면 이 정도의 차이는 단위도 유도에 있어서 크게 영향을 주지 않는 것으로 보인다. 그림 5(b)에서 '*'로 표시한 단위도는 미성구역의 관측 평균값을 적용하였을 경우이며 실선은 미성을 제외한 나머지 4개 구역의 평균 n 값 그리고 이에 대한 조건부평균 k 즉 $(n, k) = (2.46, 3.38)$ 을 통해 추정된 단위도이다. 그림에서 관측평균단위도는 추정 단위도에 거의 일치하고 있으며 추정단위도의 90% 신뢰구간에 포함된다. 미성구역을 미계측구역으로 가정하는 예는 구역규모의 상사성에 대한 정보가 부족할 경우 미계측 구역에 대한 단위도 추정의 적절성을 평가하는 것으로 결과는 만족스럽다. 그렇지만 이러한 경우 이전 절에서 언급한 바 있듯이 수공구조물 설계시 90% 신뢰구간의 상한값을 선택하여 단위도를 도출하는 것도 고려할 필요가 있다. 그리고 그림 6은 또 다른 미계측 구역의 예로 가정된 무성구역에 대한 결과이다.

그림 6(a)를 그림 5(a)와 비교하였을 때 실현값들의 분포가 보다 밀집된 관계를 보여준다. 따라서 무성의 자료가 제외됨으로써 상관관계가 상승됨을 알 수 있다. 무성구역의 관측평균값은 $(n, k) = (2.35, 3.55)$ 이며 분포

의 중심은 $(n, k) = (2.51, 3.50)$ 인데 이 수치와 그림에서 볼 수 있듯이 두 교차선 사이의 간격이 그림 5(a)와 비교할 때 큰 차이를 보인다. 즉 무성구역의 자료가 제외되었을 때 두 교차선 간격의 불균형이 심하다. 이러한 현상 역시 무성 구역은 다른 구역들 간의 상사성 관계에서 약간 벗어나 있음을 보여준다. 사실 무성구역은 표 1에서 볼 수 있듯이 관측평균 n 값이 구역 가운데 가장 작지만 반면 평균 k 값은 가장 큰 특성을 갖고 있다. 그렇지만 이 정도의 불균형은 그림 6(b)를 살펴볼 때 단위도 추정에 큰 영향을 주지 않는 것으로 보인다. 즉, 무성구역을 제외한 4개 구역평균 n 값 그리고 이 값에 대한 조건부 k 값은 $(n, k) = (2.51, 3.44)$ 으로 계산되었는데 이 값을 이용하여 추정된 단위도와 관측평균단위도간에는 좋은 일치를 보여주고 있다. 이상과 같은 예를 살펴볼 때 본 연구에서 제안된 방법은 구역의 상사성을 최대한 활용하는 특징을 갖고 있으며 또한 구역규모의 신뢰구간을 제공함으로써 매개변수의 가정치에 대한 불확실성을 완화시키는 장점을 갖고 있다. 따라서 제안방법은 미계측 구역의 단위도를 추정하는데 매우 효과적이라고 판단된다. 그렇지만 타 구역에 실무적으로 적용하기 위해서는 보다 많은 자료와 다양한 사례분석을 우선 실행하여야 할 것이다.

4. 결 론

Nash 모형은 순간단위도 도출을 위하여 빈번히 이용되는 모형이지만 적용에 있어서 잠재적인 불확실성을 갖고 있다. 이러한 불확실성은 매개변수 추정방법에 의한 영향 그리고 자료의 질과 양에 따른 영향 등으로 구성된다. 그렇지만 이러한 불확실성에 대한 평가 없이 추정된 매개변수들의 산술평균으로부터 유도된 관측평균단위도를 일반적으로 이용하고 있다. 본 논문은 불확실성 인자 가운데 후자 즉 자료의 질과 양으로부터 발생하는 불확실성을 평가하는 기법을 제시하고 이에 대한 적용성을 계측구역과 미계측 구역에 대하여 분석하였다. 이러한 연구에 대한 결과로부터 다음과 같은 결론을 도출하였다.

- (1) Nash 모형의 매개변수에 대한 추정자료가 통계학적으로 충분하지 못한 경우 관측평균단위도의 대표성 및 신뢰성은 높지 않다. 이와 같은 경우 지역분석 관계식 가운데 구역의 상사성 관계식을 이용함으로써 인근의 수문학적으로 동질성이 있는 구역의 매개변수 추정자료를 분석에 이용하였다. 이 과정에서 Nash 모형의 두 매개변수

n 과 K 대신 n 과 Tc/k 는 매우 좋은 선형상관관계를 보여주었다.

- (2) 확률변수 n 과 Tc/k 의 모집단에 대한 결합확률분포를 알지 못하므로 Box-Cox 변환을 통해 이들 확률변수들이 근사적인 이변수정규분포를 따르도록 실행한 결과 자료들의 왜곡도계수 및 첨도계수는 근사적 정규분포의 수치를 보여주었다. 그런데 Box-Cox 변환 과정에서, 주어진 자료가 여전히 충분히 않은 점을 감안하여 변환계수를 추정하기 위한 방법 가운데 통계학적 효율성이 우수한 동시추정법을 이용하였다.
- (3) 확률변수 n 과 Tc/k 에 대한 결합확률분포가 결정되었으므로 각 확률변수에 대한 주변밀도함수를 이용하여 상대 확률변수에 대한 조건부평균계산과 두 확률변수간의 선형상관계수의 추정이 가능하다. 본 연구에서는 이들 통계값 및 분포특성을 이용하여 Monte-Carlo testing procedure를 적용, 자료들을 parametric bootstrap하고 이 결과 단위도 추정에 이용함으로써 단위도의 시간별 90% 신뢰구간을 추정하였는데 그 결과 단위도의 첨두까지의 불확실성이 높게 나타났다.
- (4) 계측유역에 대하여 제안방법을 적용한 결과 해당 유역의 관측평균단위도는 유역규모의 불확실성 범위의 중앙에 위치함으로써 불확실성 구간 추정이 적절하였음을 보여주었는데 이에 따라 유역규모 및 상사성을 근거로 하는 위험율을 고려한 단위도의 도출이 가능하였다.
- (5) 가상적인 미계측 유역에 대한 매개변수 및 단위도의 신뢰구간을 추정하기 위하여 n 은 계측유역의 평균값을 이용하였는데 이에 따라 k 의 조건부기대값이 계산되었으며 이를 실측자료와 비교한 결과 거의 일치된 결과를 보여주었다. 이에 따라 미계측 유역에 대한 단위도의 추정뿐만 아니라 신뢰구간의 근사적 추정이 가능함으로 불확실성을 완화할 수 있다.

이상과 같은 결론으로 볼 때 제안된 방법은 Nash 모형의 매개변수 추정의 불확실성을 경감하는데 효과적으로 기여할 수 있을 것으로 판단되며 향후 자료의 보강과 타유역에 대한 적용성을 검증한 후 실용적으로 이용되기를 기대한다.

감사의 글

이 논문은 2000년도 건국대학교 학술연구비지원에 의한 논문임

참고 문헌

- 성기원 (1997). "수문지형특성 및 시간응답특성의 상사성을 이용한 Nash 모형의 해석." **한국수자원학회 논문집**, 한국수자원학회, 제30권, 제2호, pp. 97-106.
- 성기원 (1999). "유역의 상사성을 이용한 Clark 모형의 매개변수 해석." **한국수자원학회논문집**, 한국수자원학회, 제32권, 제4호, pp. 427-435.
- 윤석영, 홍일표 (1994). **홍수유출예측을 위한 수리수문학적 인자의 결정 : Clark 및 Nash 모형을 중심으로**. 연구보고서, 건기연 94-WR-113, 한국건설기술연구원.
- 이정식 (1994). 기존 강우-유출관계의 개념적 모형 비교 연구. **국제수문개발계획(IHP)연구보고서**, 건설교통부, pp. 6.1-6.86.
- Box, G.E.P. and Cox, D.R. (1964). "An analysis of transformation." *Journal of Royal Statistical Society*, Vol. B26, pp. 211-252.
- Efron, B. and Tibshirani, R.J. (1993). *An introduction to the bootstrap*. Chapman Hall, New York.
- Hernandez, F. and Johnson, R.A. (1980). "The large-sample behavior of transformation to Normality." *Journal of American Statistical Association*, Vol. 75, pp. 855-861.
- Hjelmfelt, A. and Wang, M. (1994). "General stochastic unit hydrograph." *Journal of Irrigation and Drainage Engineering*, Vol. 120, No. 1, pp. 138-148.
- Lin, G.F. and Wang, Y.M. (1996). "General stochastic instantaneous unit hydrograph." *Journal of Hydrology*, Vol. 182, No. 1-4.
- Sarino and Serrano, S.E. (1990). "Development of the instantaneous unit hydrograph using stochastic differential equations." *Stochastic Hydrology and Hydraulics*, Springer-Verlag, Vol. 4, pp. 151-160.
- Unny, K.E. (1984). "Stochastic nature of outputs from conceptual reservoir model cascades." *Journal of Hydrology*, Vol.68, pp. 161-180.

Van der Heijden, P.G.M., Hart, H., and Dessens, J. (1997). "A parametric bootstrap procedure to perform statistical test in latent class analysis." *Application of latent trait and latent class models in the social science*, Edited by Rost, J. and Langeheine, R., Waxman Muenster, New York.

Zou, K.H. and Hall. W.J. (2001). *On estimating a transformation correlation coefficient*. SPL Technical Report #225, Harvard Medical School.

(논문번호:02-94/접수:2002.12.23/심사완료:2003.04.09)