

# Generalized Logistic 분포형을 이용한 지역빈도해석의 불확실성 추정

## Uncertainty Assessment of Regional Frequency Analysis for Generalized Logistic Distribution

신홍준\* · 남우성\*\* · 정영훈\*\*\* · 허준행\*\*\*\*

Shin, Hongjoon · Nam, Woosung · Jung, Younghun · Heo, Jun-Haeng

### Abstract

Confidence intervals of growth curves are calculated to assess the uncertainty of index flood method as a regional frequency analysis. The asymptotic variance of quantile estimator for the generalized logistic distribution is introduced to evaluate confidence intervals. In addition, the variances of at-site frequency estimator and regional frequency estimator are used to evaluate an efficiency index. The efficiency indexes for 14 homogeneous regions based on 378 stations show that index flood method estimators are more efficient than at-site frequency estimators. It is shown that the number of sites in a region needs to be limited for regional gain.

**Keywords :** *efficiency index, asymptotic variance, index flood method, at-site frequency analysis, generalized logistic distribution*

### 요 지

본 연구에서는 홍수지수법의 불확실성을 평가하기 위해 우리나라 강우자료의 지역빈도해석에 적합한 것으로 제안된 generalized logistic 분포형의 quantile에 대한 점근 분산식을 이용하여 성장곡선에 대한 신뢰구간을 산정하였다. 또한 지점 빈도해석과 지역빈도해석에 의한 quantile의 분산을 이용하여 빈도해석의 효율성 지표(efficiency index)를 계산하였다. 우리나라 378개 강우 관측 지점을 바탕으로 구분한 14개 동질 지역에 대해 효율성 지표를 계산한 결과 홍수지수법이 지점빈도 해석보다 불확실성이 더 작은 quantile을 추정하는 것으로 나타났다. 한 지역에 포함되는 지점 개수가 과다하지 않도록 조정하는 것이 지역빈도해석의 효율성 측면에서 나온 것으로 나타났다.

**핵심용어 :** 효율성 지표, 점근 분산식, 홍수지수법, 지점빈도해석, Generalized Logistic 분포형

### 1. 서 론

빈도해석을 통한 확률수문량은 상대적으로 짧은 표본자료로부터 추정되기 때문에 발생빈도에 따른 불확실성의 증가에 의해 많은 영향을 받는다. 일반적으로 홍수빈도해석은 30~60년의 연최대수문계열의 자료를 이용하기 때문에 작은 초과확률, 예를 들어 0.01 보다 작은 확률에 대해 신뢰할만한 추정량을 얻기는 힘들다. Benson(1962)은 빈도해석을 통한 신뢰할만한 추정량은 자료개수의 2배보다 작은 재현기간에 대해서만 얻을 수 있다고 주장하였으며, Hosking 등(1984)은 신뢰할만한 수준의 추정량은 비초과확률 또는 재현기간이  $1-1/n$ 보다 작을 경우에만 얻을 수 있다고 주장하였다. 지역빈도해석과정은 “공간”을 이용하여 “시간”을 늘

림으로써 표본크기를 증가하여 이러한 어려움들을 극복하고자 하는 방법으로 이것은 적절한 홍수지수(index flood)를 사용하여 지역내의 연최대수문계열을 일반화(normalizing or indexing)하여 이루어지며, 이러한 지역빈도해석은 설계목적에 필요한 재현기간에 대한 일반화된 표본 수문량의 신뢰할 만한 확률수문량을 제공한다. 하지만 홍수지수법(index flood method)의 적용은 특정지점에서의 홍수지수 추정으로 인한 추가적인 불확실성을 반드시 고려해야 하며, 일반화된 수문추정량과 그와 연관된 홍수지수의 불확실성을 함께 고려해야 한다. 홍수지수법은 극치사상의 통계적 확률수문량 추정에 적절한 과정을 제공하며 많은 나라에서 적용되어 왔으나, 빈도해석을 통한 수문추정량의 상대적인 불확실성에 대한 조사는 아직 많이 이루어지지 않고 있다.

\*정회원 · 연세대학교 대학원 토목공학과 박사과정 (E-mail : sinong@yonsei.ac.kr)

\*\*연세대학교 대학원 토목공학과 박사과정 (E-mail : nws77@yonsei.ac.kr)

\*\*\*정회원 · 연세대학교 대학원 토목공학과 박사과정 (E-mail : yhjung2000@yonsei.ac.kr)

\*\*\*\*정회원 · 교신저자 · 연세대학교 사회환경시스템공학부 토목환경공학과 교수 (E-mail : jhheo@yonsei.ac.kr)

Boes 등(1989)은 이변수 Weibull 모형을 가정하여 확률 수문량의 불편의 추정량에 대한 분산을 계산하기 위해 CRLB(Cramer-Rao Lower Bound)의 식을 유도하였으며, 이 식을 통해 지점빈도해석과 지역빈도해석을 비교하였다. Lu와 Stedinger(1992)는 확률가중모멘트와 지점빈도해석에 의한 추정량을 표본의 평균으로 나누는 일반화된 자료를 이용하여 GEV 분포형의 근사적 분산을 구하였다. Heo 등(2001)은 이변수 Weibull 분포형을 고려하여 모멘트법, 최우도법, 확률가중모멘트법에 의한 확률수문량의 근사적 분산을 유도하였으며 이를 이용하여 근사적 분산의 적용성을 검토하였다. De Michele와 Rosso(2001)는 확률가중모멘트를 이용하여 GEV 분포형에 일반화된 홍수량 자료를 적용하여 계산되는 확률홍수량의 분산을 검토하였으며, 이를 홍수지수법에 의한 추정량과 조합하여 홍수지수법을 이용한 지역빈도해석의 분산을 평가하는 불확실성 모형을 개발하였으며, 이를 이용하여 지역빈도해석과 지점빈도해석간의 선택기준을 제시하는 효율성지수를 소개하였다.

본 논문은 지역빈도해석의 불확실성 해석을 위하여 지역빈도해석에서 사용되고 있는 홍수지수법에 의한 추정량의 분산을 구하였으며, 최근 허준행 등(2007)에 의해 우리나라 지역빈도해석에서 적합하다고 알려진 generalized logistic 분포 및 홍수지수법에 의한 추정량을 이용하여 이를 신흥준 등(2007a, 2007b)의 generalized logistic 분포의 수문추정량(quantile)에 대한 점근 분산식에 적용하였다. 또한 각 방법별 추정량의 분산을 이용하여 지점빈도해석과 지역빈도해석 선택에 있어서 구분 지표를 구하고자 하였다.

## 2. 수문추정량의 분산

### 2.1 Generalized Logistic(GL) 분포형

GL 분포형은 이변수 logistic 분포형의 일반화된 모형이며 kappa 분포형의 특수한 모형이라고 할 수 있다. Logistic 분포형으로부터의 일반화 과정은 기존 문헌들에서 정의된 것들과는 다른 것으로, Ahmad 등(1988)의 log-logistic 분포형을 재매개변수화(reparameterization)한 형태라고 할 수 있다. 또한 GL 분포형의 명칭은 generalized Pareto 분포나 generalized extreme value 분포와 같이 분포형의 특징을 반영하기 위해 붙여진 이름으로 GL 분포형의 누가분포함수와 확률밀도함수는 다음과 같다.

$$F(x) = \left[ 1 + \left\{ 1 - \frac{\beta}{\alpha}(x-x_0) \right\}^{1/\beta} \right]^{-1} \quad (1)$$

$$f(x) = \frac{1}{\alpha} \left\{ 1 - \frac{\beta}{\alpha}(x-x_0) \right\}^{1/\beta-1} \left[ 1 + \left\{ 1 - \frac{\beta}{\alpha}(x-x_0) \right\}^{1/\beta} \right]^{-2} \quad (2)$$

여기서  $x_0$ 는 위치매개변수,  $\alpha$ 는 규모매개변수,  $\beta$ 는 형상매개변수이다.

GL 분포형의 모멘트법, 최우도법, 확률가중모멘트법을 이용한 수문추정량에 대한 점근 분산식은 신흥준 등(2007a)에 의해 유도된 바 있으며, 그 중 확률가중모멘트법에 의한 수문추정량의 점근 분산식은 다음과 같다.

$$\begin{aligned} Var(\hat{x}_T) &= Var(\hat{x}_0) + Var(\hat{\alpha}) \left( \frac{1-(T-1)^{-\beta}}{\hat{\beta}} \right)^2 \\ &+ Var(\hat{\beta}) \left( \frac{\hat{\alpha} \log(T-1)}{\hat{\beta}} \frac{1}{(T-1)^{\hat{\beta}}} - \frac{\hat{\alpha}}{\hat{\beta}^2} \left( 1 - \frac{1}{(T-1)^{\hat{\beta}}} \right) \right)^2 \\ &+ 2Cov(\hat{x}_0, \hat{\alpha}) \left( \frac{1-(T-1)^{-\beta}}{\hat{\beta}} \right) + 2Cov(\hat{x}_0, \hat{\beta}) \\ &\left( \frac{\hat{\alpha} \log(T-1)}{\hat{\beta}} \frac{1}{(T-1)^{\hat{\beta}}} - \frac{\hat{\alpha}}{\hat{\beta}^2} \left( 1 - \frac{1}{(T-1)^{\hat{\beta}}} \right) \right) + 2Cov(\hat{\alpha}, \hat{\beta}) \left( \frac{1-(T-1)^{-\beta}}{\hat{\beta}} \right) \\ &\left( \frac{\hat{\alpha} \log(T-1)}{\hat{\beta}^2} \frac{1}{(T-1)} - \frac{\hat{\alpha}}{\hat{\beta}^2} \left( 1 - \frac{1}{(T-1)^{\hat{\beta}}} \right) \right) = \frac{\alpha^2}{n} v(\beta, T) \end{aligned} \quad (3)$$

여기서,

$$Var(\hat{x}_0) = \frac{1}{n} [g_{00}^2 V_{00} + g_{01}^2 V_{11} + g_{02}^2 V_{22} + 2g_{00}g_{01} V_{01} + 2g_{00}g_{02} V_{02} + 2g_{01}g_{02} V_{12}] \quad (4.a)$$

$$var(\hat{\alpha}) = \frac{1}{n} [g_{10}^2 V_{00} + g_{11}^2 V_{11} + g_{12}^2 V_{22} + 2g_{10}g_{11} V_{01} + 2g_{10}g_{12} V_{02} + 2g_{11}g_{12} V_{12}] \quad (4.b)$$

$$var(\hat{\beta}) = \frac{1}{n} [g_{20}^2 V_{00} + g_{21}^2 V_{11} + g_{22}^2 V_{22} + 2g_{20}g_{21} V_{01} + 2g_{20}g_{22} V_{02} + 2g_{21}g_{22} V_{12}] \quad (4.c)$$

$$\begin{aligned} Cov(\hat{x}_0, \hat{\alpha}) &= \frac{1}{n} [g_{10}g_{00} V_{00} + g_{11}g_{01} V_{11} + g_{12}g_{02} V_{22} \\ &+ (g_{10}g_{01} + g_{11}g_{00}) V_{01} + (g_{10}g_{02} + g_{12}g_{00}) V_{02} \\ &+ (g_{11}g_{02} + g_{12}g_{01}) V_{12}] \end{aligned} \quad (4.d)$$

$$\begin{aligned} Cov(\hat{x}_0, \hat{\beta}) &= \frac{1}{n} [g_{20}g_{00} V_{00} + g_{21}g_{01} V_{11} + g_{22}g_{02} V_{22} \\ &+ (g_{20}g_{01} + g_{21}g_{00}) V_{01} + (g_{20}g_{02} + g_{22}g_{00}) V_{02} \\ &+ (g_{21}g_{02} + g_{22}g_{01}) V_{12}] \end{aligned} \quad (4.e)$$

$$\begin{aligned} Cov(\hat{\alpha}, \hat{\beta}) &= \frac{1}{n} [g_{20}g_{10} V_{00} + g_{21}g_{11} V_{11} + g_{22}g_{12} V_{22} \\ &+ (g_{20}g_{11} + g_{21}g_{10}) V_{01} + (g_{20}g_{12} + g_{22}g_{10}) V_{02} \\ &+ (g_{21}g_{12} + g_{22}g_{11}) V_{12}] \end{aligned} \quad (4.f)$$

또한  $g_{ij}$ 와  $V_{ij}$ 는 신흥준 등(2007a)의 (B.4)~(B.8)에 나타나 있다.

### 2.2 홍수지수법(index flood method)

지역빈도해석은 수문학적으로 동질하다고 판정된 지역 내의 이용 가능한 모든 자료를 이용하여 상대적으로 큰 재현 기간에 대해 수문추정량을 구할 때 사용되는 방법이다. 이 방법을 적용하기 위해 선행되어야 하는 것은 수문학적으로 동질한 지역 내에 많은 수의 관측자료가 사용 가능해야 한다는 것이다. 그 후 매개변수적 확률분포는 연평균홍수량 같은 일반화된 특정 홍수지수에 적용시키게 된다. 홍수지수법은 일반적으로 지역빈도해석에서 사용되는 방법으로 지점의 표본자료를 동질한 지역내의 수문량인 홍수지수(지점 평균 수문량)로 나눔으로써 구해지며 이는 지역내의 모든 지점에서 동일한 성장곡선(growth curve)이라 불리는 분포형을 따라가게 된다. 따라서 특정 지점에서 확률수문량은 성장곡선

의 추정량에 지점의 홍수지수를 곱함으로써 추정하게 된다. 이러한 과정은 아래 Eq. (5)로 표현될 수 있다.

$$q_T = q_{index} \cdot x_T \quad (5)$$

여기서  $q_T$ 는 특정 지점에서 재현기간 T년의 수문량,  $x_T$ 는 지역내의 재현기간 T년의 일반화된 수문량,  $q_{index}$ 는 홍수지수로 일반적으로 연평균수문량이다.

만약 일반화된 연평균 수문자료가 GL 분포형에 적합하다고 가정하면 재현기간 T년의 확률수문량은 다음과 같다.

$$\hat{x}_T = \hat{x}_0 + \frac{\hat{\alpha}}{\beta} [1 - (T-1)^{-\beta}] \quad (6)$$

Eq. (6)을 Eq. (5)에 대입하면 아래와 같다.

$$q_T = q_{index} \cdot \left[ \hat{x}_0 + \frac{\hat{\alpha}}{\beta} [1 - (T-1)^{-\beta}] \right] \quad (7)$$

따라서 확률수문추정량의 변동성은  $\hat{q}_T$ 가 평균  $E[\hat{q}_{index}]$ ,  $E[\hat{x}_T]$ 와 분산  $Var[\hat{q}_{index}]$ ,  $Var[\hat{x}_T]$ 를 가지는 두 변수  $\hat{q}_{index}$ 와  $\hat{x}_T$ 의 곱으로 표현되는 변수임을 가정함으로써 추정될 수 있다(Stedinger 등, 1992). 하지만 위의 가정이 성립하려면  $\hat{q}_{index}$ 과  $\hat{x}_T$ 끼리 서로 독립적이어야 하는데 실제로는 서로 독립적이라고 할 수가 없다. 따라서 본 논문에서는 실제 적용 가능한 불확실성 모형을 위해  $\hat{q}_{index}$ 과  $\hat{x}_T$ 는 서로 독립적이라고 가정하였고 그로 인해 아래와 같은 관계가 성립하게 된다.

$$E[\hat{q}_T] = E[\hat{q}_{index}]E[\hat{x}_T] \quad (8)$$

$$\begin{aligned} Var(\hat{q}_T) &= E[(\hat{q}_{index}\hat{x}_T)^2] - E^2[\hat{q}_{index}\hat{x}_T] \\ &= E[\hat{q}_{index}^2]E[\hat{x}_T^2] - E^2[\hat{q}_{index}]E^2[\hat{x}_T] \\ &= E[\hat{q}_{index}^2]E[\hat{x}_T^2] - E^2[\hat{q}_{index}]E[\hat{x}_T^2] \\ &\quad + E^2[\hat{q}_{index}]E[\hat{x}_T^2] - E^2[\hat{q}_{index}]E^2[\hat{x}_T] \\ &= Var(\hat{q}_{index})E[\hat{x}_T^2] + E^2[\hat{q}_{index}]Var(\hat{x}_T) \end{aligned}$$

$$\begin{aligned} &= Var(\hat{q}_{index}) \left\{ E[\hat{x}_T^2] - E^2[\hat{x}_T] + E^2[\hat{x}_T] \right\} + E^2[\hat{q}_{index}]Var(\hat{x}_T) \\ &= Var(\hat{q}_{index}) \left\{ E[\hat{x}_T^2] - E^2[\hat{x}_T] \right\} + Var(\hat{q}_{index})E^2[\hat{x}_T] \\ &\quad + E^2[\hat{q}_{index}]Var(\hat{x}_T) \\ &= Var(\hat{q}_{index})Var(\hat{x}_T) + Var(\hat{q}_{index})E^2[\hat{x}_T] + E^2[\hat{q}_{index}]Var(\hat{x}_T) \quad (9) \end{aligned}$$

따라서 특정 지점에서 자료개수 n의 연최대수문계열이 존재한다면 관측값의 표본평균으로  $q_{index}$ 를 표시할 수 있으며 그로 인해  $\hat{q}_{index}$ 의 표본 분산은 다음과 같다.

$$\begin{aligned} Var(\hat{q}_{index}) &= Var(\hat{\mu}_0) \left( \frac{1}{n'} \sum_{i=1}^{n'} Q_i \right) = \frac{1}{n'^2} Var \left( \sum_{i=1}^{n'} Q_i \right) \\ &= \frac{1}{n'^2} \sum_{i=1}^{n'} Var(Q_i) = \frac{1}{n'^2} n' Var(Q) = \frac{1}{n'} Var(Q) \quad (10) \end{aligned}$$

### 3. 적 용

지역빈도해석은 동질한 수문학적 특성을 갖는 수문 관측 지점들을 한데 묶은 동질 지역에 적용하게 된다. 본 연구에서는 허준행 등(2007)이 구분한 동질 지역을 바탕으로 총 378개의 지점의 연최대 강우자료를 이용하여 지속시간 24시간에 대한 지역빈도해석을 수행하였다. 허준행 등(2007)은 강우 관측 지점의 위도, 경도, 고도, 연평균강수량, 7-9월의 평균강우량을 바탕으로 군집해석 기법인 fuzzy-c means 기법을 활용하여 우리나라 378개 강우 관측 지점을 14개의 지역으로 구분하였다. 구분된 14개 지역과 각 지역에 포함된 강우 관측 지점의 개수, 자료수, 매개변수 및 각 지역의 동질성 여부는 Table 1과 같다. 동질성 여부는 Hosking과 Wallis(1997)가 제안한 이질성 척도 (heterogeneity measure, H)에 근거한 것으로  $H < 1$ 이면, 동질한 (homogeneous) 지역으로,  $1 \leq H < 2$ 이면, 비교적 동질한 (possibly heterogeneous) 지역으로,  $H \geq 2$ 이면, 이질한 (heterogeneous) 지

Table 1. Results of Regional Frequency Analysis (허준행 등, 2007)

Region	Sub Region	Number of Site	Number of Data	Location Parameter	Scale Parameter	Shape Parameter	Homogeneity
Hangang	HAN01	55	1209	0.899	0.246	-0.235	Hom. <sup>a</sup>
	HAN02	25	422	0.920	0.242	-0.192	Hom.
	HAN03	42	756	0.887	0.249	-0.254	Hom.
Kumgang	KUM01	12	255	0.911	0.209	-0.241	Hom.
	KUM02	18	345	0.898	0.223	-0.257	Hom.
	KUM03	20	273	0.951	0.204	-0.141	Hom.
Nakdonggang	NAK01	37	587	0.920	0.201	-0.227	Hom.
	NAK02	55	989	0.928	0.204	-0.205	Poss. Hom. <sup>b</sup>
	NAK03	26	423	0.928	0.193	-0.214	Hom.
Namhaean	NAM01	14	436	0.898	0.220	-0.260	Hom.
Sumjingang	SUM01	10	161	0.928	0.210	-0.199	Hom.
Tonghaean	TON01	9	298	0.893	0.220	-0.272	Hom.
Youngsangang	YOU01	28	427	0.968	0.180	-0.107	Hom.
	YOU02	27	598	0.932	0.200	-0.199	Hom.

<sup>a</sup>Hom : Homogeneous

<sup>b</sup>Poss. Hom. : Possibly Homogeneous

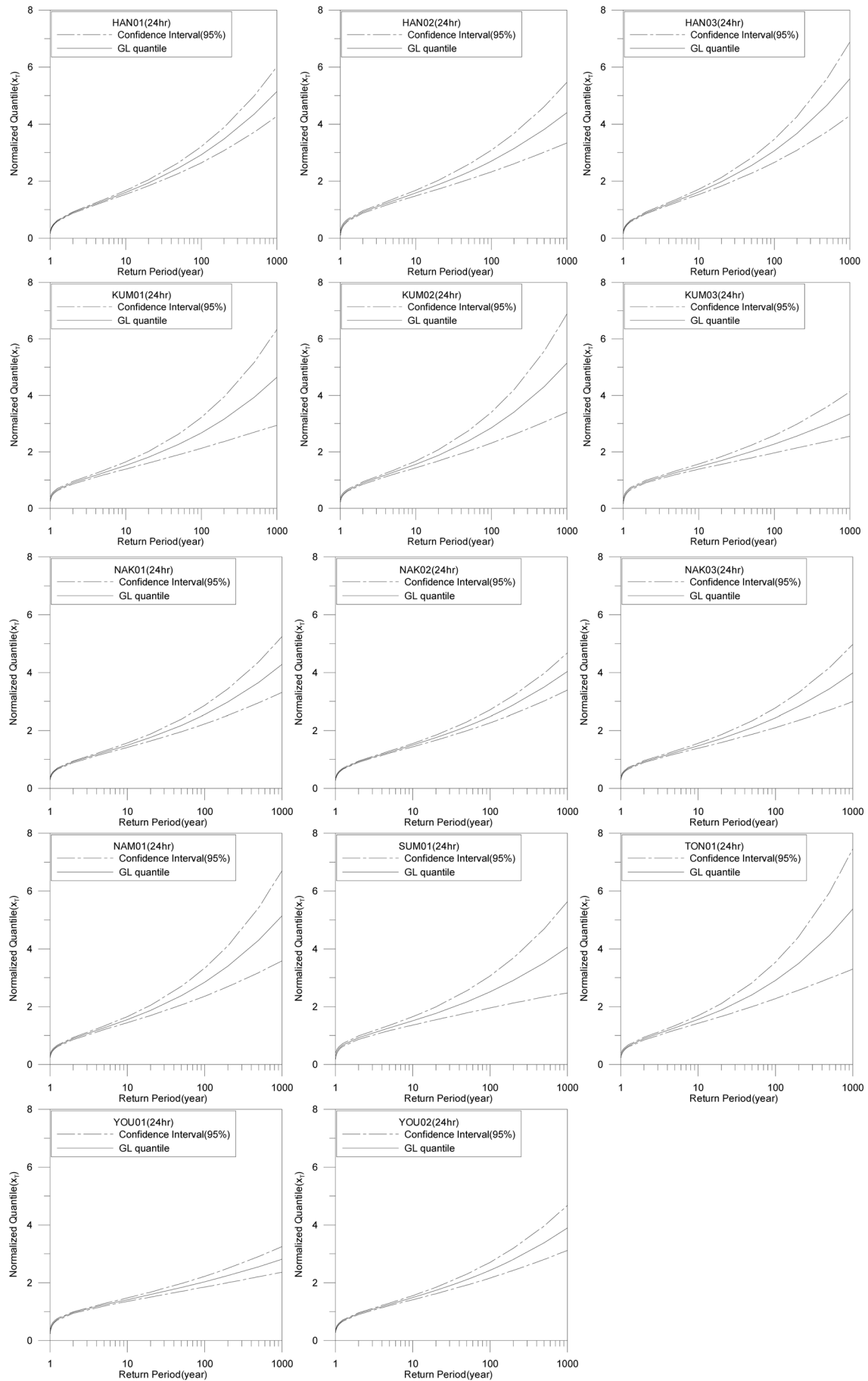


Fig. 1 Growth Curves and Confidence Intervals for 14 Regions

역으로 판단한다. Table 1의 지역과 Hosking과 Wallis(1997)의 적합성 척도(goodness-of-fit measure)를 기준으로 선정된 적정 확률분포형은 generalized logistic 분포였기 때문에 본 연구에

서는 GL 분포를 바탕으로 분석을 수행하였다.

다음의 Fig. 1은 일반화된 강우자료(normalized data)와 홍수지수법 및 Eq. (3)의 점근 분산식을 이용하여 quantile

과 신뢰구간을 나타낸 그림이다. 그림에서 금강 유역 (KUM01, KUM02), 남해안(NAM01), 섬진강(SUM01), 동해안(TON01)의 경우 다른 지역에 비해 상대적으로 신뢰구

간이 넓은데, 이는 이 지역들이 타 지역에 비해 지역내의 지점개수 및 자료개수가 상대적으로 적기 때문인 것으로 판단된다.

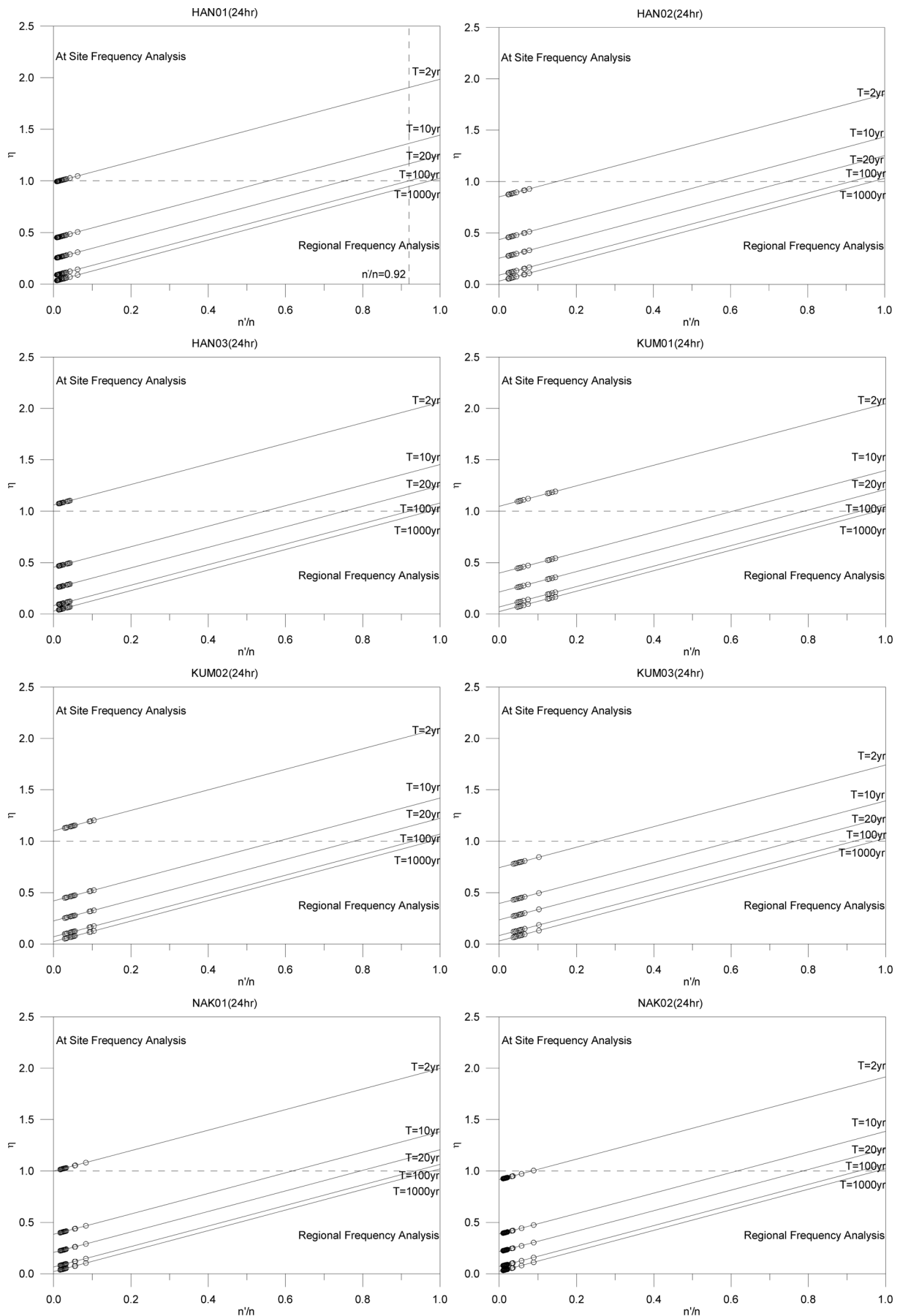


Fig. 2 Regional Relative Efficiency Index between Regional Frequency Analysis and At-Site Frequency Analysis

#### 4. 지점빈도해석과 지역빈도해석의 비교

위의 결과들을 바탕으로 지점빈도해석(at-site frequency analysis, AFA)과 지역빈도해석(regional frequency analysis, RFA)간의 효율성을 비교하기 위해 Eq. (11)과 같은 두 방법간의 상대적인 효율성 지표를 계산하였다.

$$\eta = \frac{\text{Var}(\hat{q}_T)_{RFA}}{\text{Var}(\hat{q}_T)_{AFA}} \quad (11)$$

즉,  $\eta$ 가 1보다 큰 값을 가지게 되면 그 지점에서는 지점 빈도해석이 지역빈도해석보다 더 유리하며 1보다 작은 값을 가지게 되면 지역빈도해석이 더 유리한 결과를 나타냄을 의미한다. 따라서 Eq. (9)와 Eq. (3)을 Eq. (11)에 대입하고 간단히 나타내면 다음과 같다.

$$\begin{aligned} \eta &= \frac{\text{Var}(\hat{q}_T)_{RFA}}{\text{Var}(\hat{q}_T)_{AFA}} \\ &= \frac{E^2[\hat{q}_{index}]_{RFA} \text{Var}(\hat{x}_T)_{RFA} + E^2[\hat{x}_T]_{RFA} \text{Var}(\hat{q}_{index})_{RFA} + \text{Var}(\hat{x}_T)_{RFA} \text{Var}(\hat{q}_{index})_{RFA}}{\text{Var}(\hat{x}_T)_{AFA}} \\ &= \frac{\hat{\mu}_Q^2 \frac{\alpha^2}{n} v(\beta, T) \left[ \hat{x}_0 + \frac{\hat{\alpha}}{\hat{\beta}} \left\{ 1 - (T-1)^{-\hat{\beta}} \right\} \right]^2 \frac{\alpha^2}{n' \beta^2} (g_2 - g_1^2) + \frac{\alpha^2}{n} v(\beta, T) \frac{\alpha^2}{n' \beta^2} (g_2 - g_1^2)}{\hat{\mu}_Q^2 \frac{\alpha^2}{n} v(\beta, T) + \frac{\alpha^2}{n' v(\beta, T)} + \frac{\alpha^2}{n' v(\beta, T)}} \\ &= \frac{n'}{n} + \frac{\left[ \hat{x}_0 + \frac{\hat{\alpha}}{\hat{\beta}} \left\{ 1 - (T-1)^{-\hat{\beta}} \right\} \right]^2 (g_2 - g_1^2)}{\beta^2 v(\beta, T)} + \frac{\alpha^2 (g_2 - g_1^2)}{n \beta^2} \quad (12) \end{aligned}$$

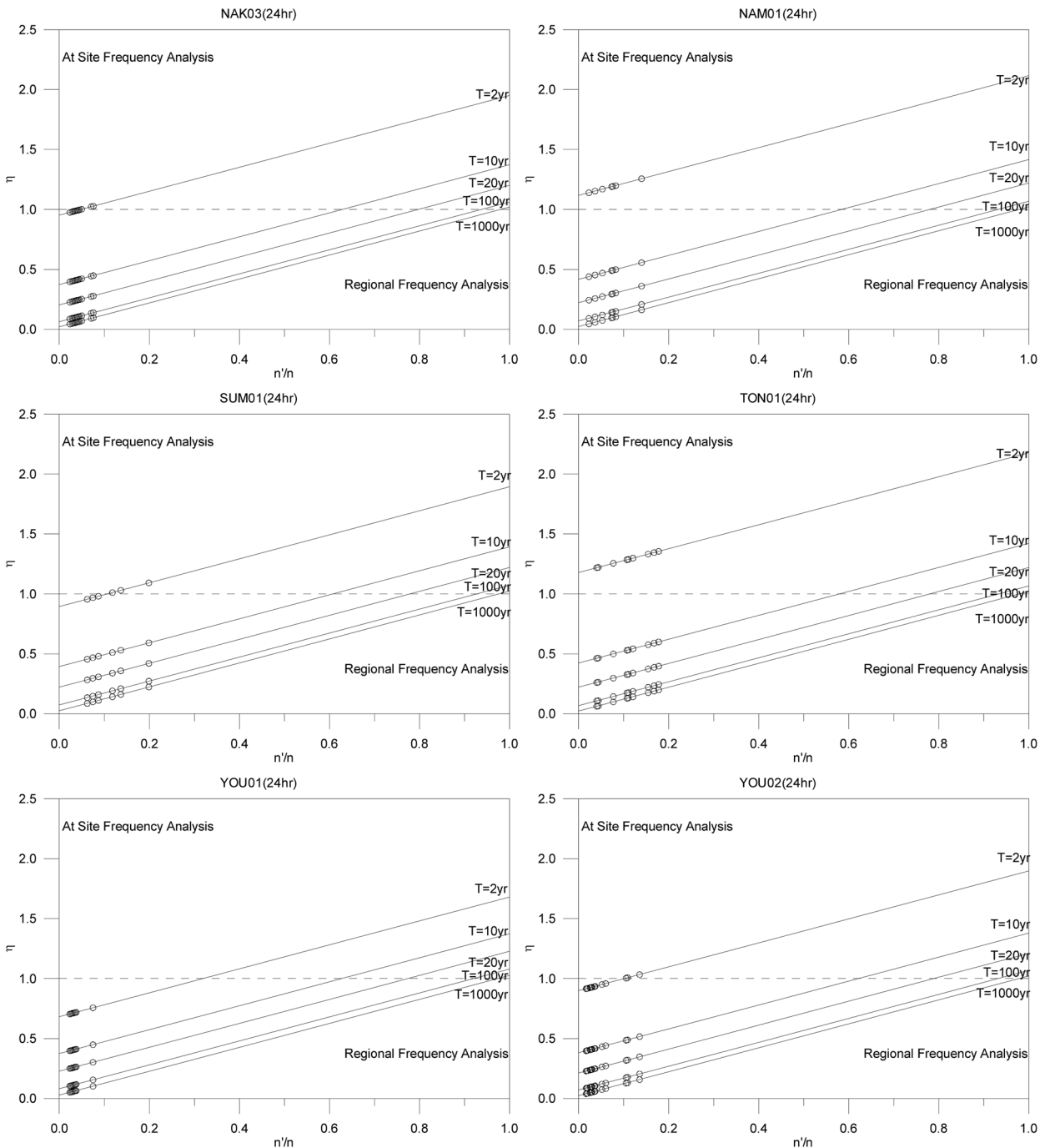


Fig. 2 Regional Relative Efficiency Index between Regional Frequency Analysis and At-Site Frequency Analysis

여기서,  $g_r = \Gamma(1+r\beta)$ ,  $v(\beta, T)$ 는 식 (3)에서  $\alpha^2/n$ 를 제외한 나머지 항,  $n$ 은 지역전체의 자료개수,  $n'$ 은 특정지점의 자료개수이다. 또한  $\alpha', \beta'$ 는 지점빈도해석의 매개변수이고  $\alpha, \beta$ 는 지역빈도해석의 매개변수이며 일반화된 표본 자료를 이용하였기 때문에  $\alpha' = \alpha\mu_Q$ ,  $\beta' = \beta$ 의 관계가 성립한다.

다음의 Fig. 2는 Eq. (12)의 지역빈도해석과 지점빈도해석의 quantile에 대한 분산을 이용하여 각각의 지역에 적용해 본 결과이며, 각 지역내의 모든 지점을 점으로 표시하였다. 예를 들어 HAN01 지역의 전체자료개수( $n$ )가 100개이고 특정지점의 자료개수( $n'$ )가 30개라고 하면 자료개수의 비율( $n'/n$ )은 0.3이고, 재현기간 100년의 quantile 추정을 위해서는 효율성 지표가 1일 때의 자료개수의 비율 0.92, 즉 전체 지역자료개수의 92% 정도인 92개 이상의 자료개수를 특정지점에서 보유해야 지점빈도해석을 수행하는 것이 지역빈도해석을 수행하는 것보다 더 정확하다는 것으로 여기서는 지역빈도해석이 더 유리하다는 것을 의미한다.

적용 결과 일반적으로 재현기간 2년을 제외하고는 대부분의 지역 및 재현기간에 대하여 지점빈도해석보다 지역빈도해석을 수행하는 것이 quantile에 대한 분산이 작게 나오며 이를 통해 지점빈도해석보다 지역빈도해석의 결과가 불확실성이 더 작은 quantile을 추정할 수 있게 해주는 것으로 나타났다. 다만 그림에서 볼 수 있듯이 대부분의 지역에서 지점의 자료개수와 지역전체의 자료개수의 비율이 0~0.2 사이로 지점의 자료개수가 지역전체의 자료개수에 비해 상당히 적은 수를 나타내고 있으며, 이는 지역빈도해석의 적용시 필요한 자료개수보다 훨씬 많은 자료가 지역 내에 존재한다고 할 수 있다. 따라서 현재 상태에서 지역빈도해석을 적용할 경우 필요이상의 자료개수로 인하여 늘어난 자료개수만큼의 지역빈도해석을 통한 이득은 크게 기대할 수 없을 것으로 판단되며 이는 효율성 측면에서 지역을 좀 더 세분화하는 것이 더 좋을 것이라 판단된다.

## 5. 결 론

본 논문에서는 최근 지역빈도해석에서 많이 사용되고 있는 generalized logistic 분포형과 확률가중모멘트법을 이용한 quantile에 대한 점근 분산식을 이용하여 지점빈도해석의 가장 큰 단점인 자료부족의 극복을 위한 지역빈도해석의 불확실성 추정을 수행하였다. 이를 위해 지역빈도해석에서 가장 많이 사용되는 홍수지수법을 이용하여 quantile을 추정하였으며, 여기에 분포형 자체의 quantile에 대한 점근 분산식을 적용하여 홍수지수법으로 추정되는 quantile의 발생범위를 나타내었다. 또한 연구 수행을 위해 지역내의 모든 지점들간에는 서로 독립이라는 가정을 하였으며, 지점빈도해석과 지역빈도해석에 의한 quantile의 분산을 이용하여 효율성지표 (efficiency index)를 제시하였다. 이 지표를 우리나라 전체 14개 지역에 적용해 본 결과 현재 구분되어 있는 지역을 기준으로 하였을 경우 일반적으로 재현기간 2년을 제외하고는 대부분의 지역 및 재현기간에 대하여 지점빈도해석보다 지역빈도해석을 수행하는 것이 quantile에 대한 변동성이 더 작게 나오며 이를 통해 지점빈도해석보다 지역빈도해석의 결과가 불확실성이 더 작은 quantile을 추정할 수 있게 해주는

것으로 나타났다. 하지만 현재 상태 그대로 지역을 구분하여 지역빈도해석을 적용할 경우 필요이상의 자료개수로 인하여 늘어난 자료개수만큼의 이득은 기대할 수 없을 것으로 판단되며 이는 효율성 측면에서 지역을 좀 더 세분화하는 것이 더 좋을 것이라 판단된다. 또한 본 연구에서는 지역내의 지점간 상관성은 고려하지 않았으나 향후 지역빈도해석의 불확실성 추정에 영향을 미칠 수 있는 지역내의 지점간 상관성을 고려한 연구가 지속적으로 수행되어야 할 것으로 판단된다.

## 감사의 글

본 연구는 국토해양부가 출연하고 한국건설교통기술평가원에서 위탁시행한 2003년도 건설기술혁신사업(03산학연 C01-01)에 의한 도시홍수재해관리기술연구사업단의 연구성과입니다.

## 참고문헌

- 신홍준, 김수영, 허준행(2007a) Generalized Logistic 분포형의 신뢰구간 추정을 위한 점근 분산 유도 : I. 점근 분산식의 유도. **대한토목학회논문집**, 대한토목학회, 제27권, 제3B호, pp. 331-338.
- 신홍준, 김수영, 허준행(2007b) Generalized Logistic 분포형의 신뢰구간 추정을 위한 점근 분산 유도 : II. 모의실험 및 적용. **대한토목학회논문집**, 대한토목학회, 제27권, 제3B호, pp. 339-344.
- 허준행, 이영식, 신홍준, 김경덕(2007) 우리나라 강우자료의 지역 빈도해석 적용성 연구(I) : 확률강우량 산정. **대한토목학회논문집**, 대한토목학회, 제27권, 제2B호, pp. 101-111.
- Ahmad, M.I., Sinclair, C.D., and Werritty, A. (1988) Log-logistic flood frequency analysis with historical information, *Journal of Hydrology*, Vol. 98, pp. 205-224.
- Benson, M.A. (1962) Evolution of methods for evaluating the occurrence of floods. *USGS Water Supply Paper*, 1580A.
- Boes, D.C., Heo, J.H., and Salas, J.D. (1989) Regional flood quantile estimation for a Weibull model. *Water Resources Research*, Vol. 25, No. 5, pp. 979-990.
- De Michele, C. and Rosso, R. (2001) Uncertainty assessment of regionalized flood frequency estimates. *Journal of Hydrology Engineering*, Vol. 6, No. 6, pp. 453-459.
- Heo, J.H., Boes, D.C., and Salas, J.D. (2001a) Regional flood frequency analysis based on a Weibull model : part 1. Estimation and asymptotic variances. *Journal of Hydrology*, Vol. 242, No. 3-4, pp. 157-170.
- Heo, J.H., Salas, J.D., and Boes, D.C. (2001b) Regional flood frequency analysis based on a Weibull model : part 2. Simulations and applications. *Journal of Hydrology*, Vol. 242, No. 3-4, pp. 171-182.
- Hosking, J.R.M., Wallis, J.R., and Wood, E.F. (1984) Estimation of the generalized extreme value distribution by the method of probability weighted moments. *Technometrics*, Vol. 27, No. 3, pp. 251-261.
- Hosking, J.R.M. and Wallis, J.R. (1997) *Regional frequency analysis: An approach based on L-moments*, Cambridge University Press.
- Lu, L.H. and Stedinger, J.R. (1992) Variance of two- and three-parameter GEV/PWM quantile estimators and a regional homogeneity test. *Journal of Hydrology*, Vol. 138, pp. 247-267.
- Stedinger, J.R., Vogel, R.M., and Foufoula-Georgiou, E. (1992) Frequency analysis of extreme events, *Handbook of Hydrology*, Ch18., McGraw-Hill, New York.

(접수일: 2008.8.21/심사일: 2008.9.17/심사완료일: 2008.9.17)