CAUTION ON REGIONAL FLOOD FREQUENCY ANALYSIS BASED ON WEIBULL MODEL

Jun-Haeng Heo[1], Dong-Jin Lee[1], and Kyung-Duk Kim[2]

1. Department of Civil Engineering, Yonsei University, Seoul, Korea
2. Korea Infrastructure Safety and Technology Cooperation, Kyunggi- Do, Korea

Abstract: Regional flood frequency analysis has been developed by employing the nearby site's information to improve a precision in estimating flood quantiles at the site of interest. In this paper, single site and regional flood frequency analyses were compared based on the 2-parameter Weibull model. For regional analysis, two approaches were employed. The first one is to use the asymptotic variances of the quantile estimators derived based on the assumption that all sites including the site of interest are independent each other. This approach may give the maximum regional gain due to the spatial independence assumption among sites The second one is Hosking's regional L-moment algorithm. These methods were applied to annual flood data. As the results, both methods generally showed the regional gain at the site of interest depending on grouping the sites as homogeneous. And asymptotic formula generally shows smaller variance than those from Hosking's algorithm. If the shape parameter of the site of interest from single site analysis is quite different from that from regional analysis then Hosking's results might be better than the asymptotic ones because the formula was derived based on the assumption that all sites have the same regional shape parameter. Furthermore, in such a case, regional analysis might be worse than single site analysis in the sense of precision of flood quantile estimation. Even though the selected sites may satisfy Hosking's criteria, regional analysis may not give a regional gain for specific and nonexceedance probabilities.

Key Words regional frequency analysis, Weibull model, Hosking's regional L-moment algorithm, regional gain, asymptotic variance

1. INTRODUCTION

Regional frequency analysis has been a standard practice in hydrology for determining flood quantile magnitudes at ungaged locations and at sites with short records (Darymple, 1960; Benson, 1962; Stedinger and Tasker, 1985; Cunnane, 1989). When the sample size is small at the site of interest, the single site analysis may result in a precision problem in estimating flood quantiles. Thus, regional flood frequency analysis has been suggested to improve such a problem. One of the popular models is regional regression, which defines the relation between flow characteristics and drainage basin characteristics. Another model for regional flood frequency analysis is the so-called index flood method. This approach to regional frequency analysis was used in the 50's and renewed in the 80's. It employs some form of a dimensionless average flood frequency curve (Darlymple, 1960). Several innovations and developments have been made on the original index flood method of regionalization (Wallis, 1980; Greis and Wood, 1981; Stedinger, 1983; Lettermaier and Potter, 1985; Hosking et al., 1985;

Hosking and Wallis, 1988; Boes et al., 1989; Heo et al., 1990), in which the annual flood data at each site is indexed by dividing site data by its mean annual flood. Then, a probability model is fitted to these indexed flood data based on the assumption of having identical distribution all over sites in a region. Once the regional parameters are estimated, the flood quantile at the site of interest is estimated by multiplying the estimated regional flood quantile by its mean annual flood. Therefore, the index flood method could combine the data at the site of interest with the data at nearby sites. Recently the regional frequency analysis have been carried out using index flood procedure with L-moments since Hosking (1985) introduced the L-moment algorithm (Rao and Hamed, 1997; Parida et al., 1998).

Consider there are k sites in a region. Then the indexed data becomes as Table 1, where n_j and μ_j are sample size and sample mean at j^{th} site, and X_{ji} is the flood data for site j and year i.

In flood frequency analysis, the most important decision is to select appropriate probability distribution for the given annual peak floods. Many probability families such as the normal, lognormal, gamma, log-Pearson type III, generalized extreme value, Gumbel, Weibull, log-Gumbel, Wakeby, and two-component extreme value

Table 1. Indexed Data of Sites in a Region

Site	Indexed data
1	X_{11}/μ_1, X_{12}/μ_1, X_{13}/μ_1, …, X_{1n_1}/μ_1
2	X_{21}/μ_2, X_{22}/μ_2, X_{23}/μ_2, …, X_{2n_2}/μ_2
.	.
.	.
k	X_{k1}/μ_k, X_{k2}/μ_k, X_{k3}/μ_k, …, X_{kn_k}/μ_k

distributions have been studied (Greis, 1983, Rossi et al., 1984; Cunnane, 1989, Bobee et al., 1993).

In this study, regional frequency analysis has been performed based on an index flood method and a two-parameter Weibull probability distribution. Even though Weibull distribution may not give a better fit to given data, it was used for the candidate distribution because analytical solution can be obtained fairly easily for regional frequency analysis. And the asymptotic regional variances of single site and regional flood quantile estimators have been estimated based on the method of maximum likelihood (ML), the method of moments (MOM), and the method of probability weighted moments (PWM). Then these formulas were applied to some actual flood data to investigate the regional gain due to regional flood frequency analysis. Finally, the results were compared with those from Hosking regional L-moment procedure.

2. REGIONAL MODELING

2.1 Asymptotic Regional Algorithm

Let us consider a two-parameter Weibull model as a parent model at each site in a region.

The cumulative distribution function (CDF) of Weibull model is given by

$$F_j = 1 - \exp[-(x/\beta_j)^{\gamma_j}],\ x \geq 0 \quad (1)$$

where, β_j is the scale parameter, γ_j is the shape parameter, j indexes the site. The rth moment about the origin and quantile at site j for the two-parameter Weibull distribution are given by

$$\mu_r(j) = \beta_j^r \Gamma(1+r/\gamma_j) \quad (2)$$

$$\xi_j(q) = \beta_j[-\log(1-q)]^{1/\gamma_j} \tag{3}$$

The index flood assumption states that $\xi_j(q)/\mu_1(j)$ does not depend on j, thus for the Weibull model

$$\xi_j(q)/\mu_1(j) = [-\log(1-q)]^{1/\gamma_j}/\Gamma(1+1/\gamma_j)$$

should not depend on j, therefore, it implies all shape parameters are identical, that is $\gamma_j = \gamma$ (Boes et al., 1989). In addition, it is assumed that the flood data $X_{j1}, X_{j2}, \ldots, X_{jn_j}$ of site j with record length n_j are independent and identically distributed and independent across the sites. From the assumption that the shape parameters of all sites in a region are identical, asymptotic variances of unbiased quantile estimators based on ML, MOM, PWM were evaluated and the gain of regional flood frequency analysis over single site analysis was quantified based on the formula (Boes et al., 1989).

2.1.1 Asymptotic Variance for ML(CRLB)

The Cramer-Rao Lower Bound (CRLB) gives the asymptotic variance of maximum likelihood estimators. If the site 1 is the site of interest for regional analysis, the CRLB for the variance of unbiased estimators of the th quantile at site 1 is given by (Boes et al., 1989)

$$CRLB[\hat{\xi}_1(q)_{ML}] = \frac{\xi_1^2(q)}{\gamma^2}\left\{\frac{1}{n_1} + \frac{\{\Gamma'(2) - \log[-\log(1-q)]\}^2}{\sum n_j \cdot C}\right\} \tag{4}$$

where $C = 1 + \Gamma''(2) - [\Gamma'(2)]^2 = \pi^2/6$, $\Gamma'(2)$ and $\Gamma''(2)$ are the first and second derivatives of the gamma function evaluated at 2, respectively.

2.1.2 Asymptotic Variance for MOM

The asymptotic variance of the estimator of the qth quantile at site 1, $\hat{\xi}_1(q)_{MOMw}$ which depends on weight-coefficient w is given by (Heo et al., 1990)

$$AVAR[\xi_1(q)_{MOMw}] = \xi_1^2(q) \times \left\{\frac{\sigma_1^2}{n_1} - \frac{2}{n}(1-w)Cov(V,U) + \frac{(1-w)^2}{n_1}\sigma_U^2 - \frac{2Cov(V,U)}{\sum n_j}\frac{S}{D} + \frac{2(1-w)\sigma_U^2}{\sum n_j} + \frac{1}{\sum n_j}\frac{S^2}{D^2}\sigma_U^2\right\} \tag{5}$$

where

$$S = w \cdot \varphi(1-1/\gamma) + (1-w) \cdot \varphi(1+2/\gamma) - \log[-\log(1-q)]$$

$$D = \varphi(1+1/\gamma) - \varphi(1+2/\gamma)$$

$$V = Z/\mu_Z, \quad U = V - Z^2/2\mu_Z$$

$$\sigma_U^2 = \frac{\Gamma(1+2/\gamma)}{\Gamma^2(1+1/\gamma)} - \frac{\Gamma(1+3/\gamma)}{\Gamma(1+1/\gamma)\Gamma(1-2/\gamma)} + \frac{\Gamma(1+4/\gamma)}{4\Gamma^2(1+2/\gamma)} - \frac{1}{4}$$

$\sigma_1^2 = \Gamma(1+2/\gamma)/\Gamma^2(1+1/\gamma) - 1$, and $\varphi(\cdot)$ is the digamma function defined by $\Gamma'(\cdot)/\Gamma(\cdot)$, and $\Gamma'(\cdot)$ is the first derivative of the gamma function.

Since $AVAR[\hat{\xi}_1(q)_{MOMw}]$ is the function of w, the minimum of $AVAR[\hat{\xi}_1(q)_{MOMw}]$ can be found by differentiating Eq. (5) with respect to w and equating to zero, which yields,

$$w_{opt} = \{\sigma_U^2 - Cov(V,U)\}/\sigma_U^2 \tag{6}$$

where $\hat{\xi}_1(q)_{MOM}$ is not an estimator, because w_{opt} is a function of γ. Set $\hat{w}_{opt} = w_{opt}(\hat{\gamma})$ and define the iterative estimator by

$$\hat{\xi}_1(q)_{MOM} = [-\log(1-q)]^{1-\hat{\gamma}_{MOM}}$$
$$[\ \hat{w}_{opt}\ \hat{\beta}_1(MOM)_1 + (1-\ \hat{w}_{opt})\cdot\ \hat{\beta}_1(MOM)_2] \quad (7)$$

then the asymptotic variance of $\hat{\xi}_1(q)_{MOM}$ is given by

$$AVAR[\ \hat{\xi}_1(q)_{MOM}] = \xi_1^2(q)\left\{\frac{1}{n_1}\sigma_V^2(1-\rho_{U,V}) + \frac{1}{\sum n_i}\frac{T^2}{D^2}\sigma_U^2\right\} \quad (8)$$

where

$T = \varphi(1+1/\gamma) - Cov(V,U)D/\sigma_U^2 - \log[-\log(1-q)]$

and $\rho_{U,V}$ is the correlation coefficient between U and V.

2.1.3 Asymptotic Variance for PWM

The asymptotic variance of the estimator of the th quantile at site 1, $\hat{\xi}_1(q)_{PWM}$ is given by (Heo et al., 1990)

$$AVAR[\ \hat{\xi}_1(q)_{PWM}] = \frac{\xi_1^2(q)}{\Gamma^2(1+G)} \times \left\{ \frac{1}{n_1}\left[-\frac{X^2}{Y} + 4H(1/2)\Gamma(1+2G) - 4\Gamma^2(1+2G)\right] + \frac{(1-X/Y+N)^2}{\sum n_i}Y \right\} \quad (9)$$

where $G = 1/\gamma$,

$$X = [\ -2^{-G} + 4H(1/2)]\Gamma(1+2G),\ -(2+2^G)\Gamma^2(1+2G)$$

$$Y = [1-2^{-G} + 4H(1/2)]\Gamma(1+2G),\ -(1+2^1+G)\Gamma^2(1+2G)$$

and $H(1/2) = {}_2F_1(2G, G; 1+G; -1/2)$ is the hypergeometric function.

For the purpose of comparison of regional analysis and single site analysis, the concept of regional gain was introduced to measure the relative analysis effect (Heo et al., 1990). The regional gain, asymptotic gain due to regionalization, is defined by the ratio of asymptotic variance evaluated from regional analysis to that from single site analysis for each scheme of the parameter estimations such as method of moments, maximum likelihood, and probability weighted moments. For example, the regional gain at site 1 for PWM can be written by Eq. (10). Note that if the value of R is larger than 1, there is no gain at site 1 from regional flood frequency analysis. In other words, single site analysis is better than regional flood frequency analysis in quantile estimating.

$$R = \frac{Avar[\ \hat{\xi}_1(q)_{PWM}]_{REGIONAL}}{Avar[\ \hat{\xi}_1(q)_{PWM}]_{SITE}} \quad (10)$$

2.2 Regional L-moment Algorithm

Hosking (1990) defined the concept of L-moments which are the linear combinations of the probability weighted moments and shown that index flood procedure based on L-moments yields robust and accurate estimates. He also suggested some statistics that could measure the discordancy of data, the homogeneity of a region, and goodness of fit, and then constructed the regional analysis procedure with 4 stages (Hosking and Wallis, 1993).

In the first stage, the sample L-moment ratios for different sites are compared. And it is checked whether the sample data has the gross errors and inconsistencies, and stationary or not. The measure of the discordancy D_i is not considered good for the regional analysis if D_i exceeds the critical value 3.

$$D_i = \frac{1}{3}(u_i - \overline{u})^T S^{-1}(u_i - \overline{u}) \quad (11)$$

where u_i is the vector of L-coefficient of variation (LCV), L-coefficient of skewness (LCS), and L-kurtosis (LCK) for a site i, S is the covariance matrix of u_i, and $\overline{u}$ is the mean of vector u_i.

For the regional frequency analysis, the region is not geophysical and it should include sites that have similar statistical characteristics. In the second stage of regional analysis, the degree of regional homogeneity is estimated using heterogeneity measure H, in which the region is accepted as homogeneous if $H < 1$, possibly heterogeneous if $1 \le H < 2$, and definitely heterogeneous if $H \ge 2$. Heterogeneity measure H is defined by

$$H = \frac{V - \mu_V}{\sigma_V} \quad (12)$$

where, V is the weighted standard deviation of the single site sample LCVs, μ_V and σ_V are the mean and standard deviation of simulated V values, respectively.

In the next stage, a regional frequency distribution is fitted to the observed data of several sites in a region by using a goodness of fit measure Z which describes the relationship between the LCS and LCK of the fitted distribution and those of the observed data. Hosking and Wallis (1993) suggested that fitted distribution be proper if Z^{DIST} would be sufficiently close to zero and that a reasonable criterion be $|Z^{DIST}| \le 1.64$, in which Z^{DIST} is given by

$$Z^{DIST} = \frac{\tau_4^{DIST} - t_4^R + B_4}{\sigma_4} \quad (13)$$

where t_4^R is the average L-kurtosis value computed from the data of a given region, B_4 is the bias of t_4^R, σ_4 is the standard deviation of t_4^R, and τ_4^{DIST} is the average L-kurtosis value computed from simulation for a fitted distribution.

Then, the regional frequency distribution obtained from goodness of fit test should be fitted to the data and scaled pertinently. Using the index flood method, Hosking and Wallis (1993) estimated the quantile magnitudes at a site for a given return period from sample mean and the regional distribution with weighted average to the data.

3. DATA APPLICATION

3.1 Data Description

Flood data were collected from gauging stations in Illinois, Indiana, and Wisconsin (Curtis, 1987). There are 394 continuous-record sites having at least 10 years but only 21 sites have more than 33 successive annual flood data. For the homogeneity problem, we trimmed the list by removing 6 sites based on the skewness, the shape parameter estimates, drainage areas, and sample size of each site in a region. For the remaining 15 sites as shown in Table 2, Weibull assumption cannot be rejected at 5% significance level by using a χ^2−goodness of fit test (Kottegoda and Rosso, 1997) and a probability plot correlation coefficients test (Filliben, 1975; Vogel and Kroll, 1989). In addition, independence assumptions for the region were checked by computing sample autocorrelation coefficients and cross-correlation coefficients and then fairly well satisfied (Salas et al., 1980). Even though the estimated shape parameters are not alike, all 15 sites are lumped into one group. In addition, for comparison and illustration, two

Table 2. General Description of Selected Sites Illinois and Wisconsin Regions

Site No.	Sample Size	Area (km^2)	Mean (cms)	Std. Dev. (cms)	Skewness	Site No.	Sample Size	Area (km^2)	Mean (cms)	Std. Dev. (cms)	Skewness.
18	60	3926	487	302	0.702	235	44	2248	83	41	1.259
45	63	631	124	68	0.452	262	43	1499	164	83	0.515
55	36	639	163	83	0.825	265	55	3240	381	216	0.624
61	43	3434	205	117	1.070	307	57	1424	169	112	1.237
63	69	1354	123	86	1.105	327	40	2807	409	255	1.169
124	35	5415	349	168	0.760	335	65	3348	288	157	0.929
144	45	932	64	28	0.376	344	45	2248	346	243	0.973
233	46	1178	265	143	0.508						

more groups are considered as

Case 1 : 45, 61, 233, 265, 335 (5 sites in a group)

Case 2 : 45, 61, 233, 265, 335, 55, 235, 262 (8 sites in a group)

Case 3 : all 15 sites in a group

For case 1, those 5 sites having values of the ML estimates of the shape parameter closely clustered were grouped. For case 2, the 3 sites with the larger ML estimates of the shape parameter were added to case 1. Considering the ML estimates of γ_s, the standard deviation of the shape parameter estimates was in the following order of case 3 > case 2 > case 1.

3.2 Comparisons of Variances

Asymptotic variances of the qth quantile estimators based on the ML, MOM, and PWM methods from the single site and the regional frequency analyses were evaluated in order to investigate the regional gains due to regional analysis. Furthermore, the comparisons of the variances between asymptotic regional variance analysis and Hosking's algorithm were performed. As for Hosking regional frequency analysis the variances of qth quantiles for the single and the regional analysis were evaluated from the relationships of relative bias (RBIAS) and relative root mean square error (RRMSE) using the simulation codes of Hosking (Hosking, 1996). The sites in a region were assumed to be independent identically distributed in the modeling setup part. For the purpose of evaluation for the effect of independence, Monte-Carlo simulation for regional analysis considering dependence of sites in a region was performed.

Table 3 shows the statistics representing homogeneity and goodness of fit following Hosking procedure. As shown in Table 3, 15 sites of case 3 cannot be considered as

Table 3. Results of Homogeneity and Goodness-of-fit Tests

Case	No. of Site	Heterogeneity Measures			Goodness-of-fit Measures				
		H_1	H_2	H_3	Weibull	GEV	Gen normal	Pearson Type 3	Gen Pareto
1	5	-1.81	-0.07	0.32	4.00	4.00	3.79	3.12	-0.69
2	8	-0.71	0.12	0.3	7.18	4.66	4.32	3.44	-0.89
3	15	2.13	0.52	-0.64	8.98	5.63	5.02	3.62	-1.88

homogeneous because of $H>2$ and the goodness of fit measure Z values for Weibull model have high values in every case.

L-moment ratio diagram has been used for determining the regional frequency distribution (Hosking, 1990; Vogel and Fennessay, 1993), and the goodness of fit tests for the two-parameter Weibull distribution was carried out using the L-moment ratio diagram, results of which are shown in Fig. 1 and Fig. 2. The two-parameter Weibull distribution is good for the regional distribution to given data,

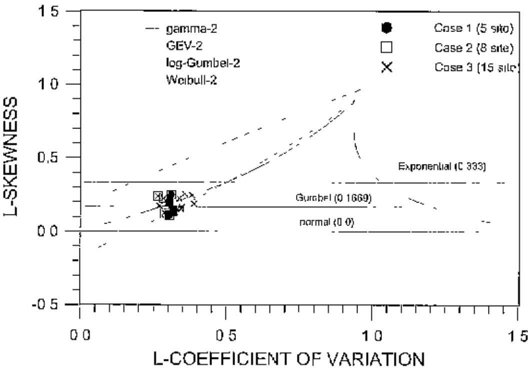

Fig. 1. L-moment Ratio Diagram (LCV vs. LCS)

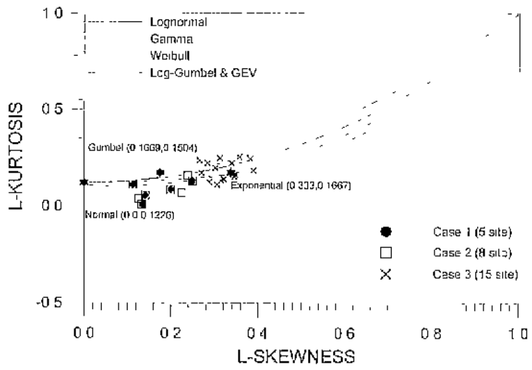

Fig. 2. L-moment Ratio Diagram (LCV vs. LCS)

Table 4. The Ratio of Variance for Various Estimation Techniques (Case 1)

Regional Shape Parameter γ = 1.93, Autocorrelation Coeff. = 0.062

Site	q	0.01	0.05	0.1	0.5	0.7826	0.9	0.95	0.99	0.999	γ_s
45	ML	0.27	0.32	0.37	0.79	1.01	0.94	0.85	0.69	0.57	1.94 (>)
	MOM	0.27	0.32	0.37	0.79	1.01	0.94	0.84	0.68	0.56	
	PWM	0.29	0.33	0.38	0.77	0.97	0.88	0.79	0.62	0.51	
	L-M (i)	0.28	0.34	0.39	0.82	1.05	0.96	0.85	0.67	0.54	
	L-M (d)	0.30	0.36	0.41	0.83	1.04	0.94	0.83	0.65	0.53	
61	ML	0.22	0.27	0.32	0.76	0.97	0.89	0.79	0.61	0.49	1.91 (<)
	MOM	0.23	0.27	0.32	0.74	0.94	0.85	0.75	0.57	0.45	
	PWM	0.21	0.26	0.31	0.76	0.99	0.90	0.79	0.62	0.49	
	L-M (i)	0.2	0.27	0.33	0.82	1.06	0.96	0.84	0.64	0.49	
	L-M (d)	0.22	0.28	0.34	0.83	1.06	0.95	0.82	0.62	0.48	
233	ML	0.21	0.27	0.32	0.78	1.02	0.95	0.84	0.67	0.54	1.95 (>)
	MOM	0.20	0.26	0.31	0.80	1.07	1.00	0.90	0.72	0.58	
	PWM	0.20	0.26	0.31	0.79	1.05	0.97	0.86	0.68	0.55	
	L-M (i)	0.20	0.27	0.33	0.84	1.11	1.01	0.89	0.69	0.54	
	L-M (d)	0.21	0.28	0.34	0.85	1.10	1.00	0.87	0.67	0.53	
265	ML	0.27	0.31	0.36	0.76	0.95	0.87	0.77	0.61	0.49	1.88 (<)
	MOM	0.27	0.31	0.35	0.75	0.95	0.87	0.77	0.6	0.48	
	PWM	0.28	0.32	0.36	0.74	0.93	0.84	0.74	0.57	0.46	
	L-M (i)	0.27	0.32	0.37	0.82	1.04	0.94	0.82	0.63	0.49	
	L-M (d)	0.29	0.34	0.39	0.84	1.04	0.93	0.80	0.61	0.48	
355	ML	0.27	0.33	0.38	0.81	1.04	0.97	0.88	0.71	0.59	1.97 (>)
	MOM	0.27	0.33	0.37	0.80	1.03	0.96	0.86	0.7	0.58	
	PWM	0.25	0.31	0.37	0.82	1.08	1.01	0.91	0.74	0.62	
	L-M (i)	0.24	0.31	0.37	0.83	1.10	1.03	0.93	0.76	0.63	
	L-M (d)	0.26	0.33	0.39	0.85	1.09	1.01	0.91	0.74	0.61	

and then considered as the representative distribution for the regional L-moment algorithm.

We defined R as the ratio of the variance from single frequency analysis to the variance from regional frequency analysis. Thus, $1-R$ represents the regional gain due to regional frequency analysis. Tables 4-6 show R values for comparisons of the variances of qth quantiles from asymptotic variance procedure based on three estimation methods and Hosking procedure considering independence L-M(i) and dependence L-M(d) characteristics.

As shown in Table 4, the estimated regional shape parameter $\gamma = 1.93$ and the estimated single site shape parameter value γ_s is shown in the last column. There are regional

Table 5. The Ratio of Variance for Various Estimation Techniques (Case 2)

Regional Parameter γ = 2.00, Autocorrelation Coeff. = 0.165

Site	q	0.01	0.05	0.1	0.5	0.7826	0.9	0.95	0.99	0.999	γ_s
	ML	0.23	0.28	0.32	0.74	0.94	0.86	0.76	0.59	0.46	
	MOM	0.23	0.27	0.32	0.74	0.94	0.86	0.75	0.58	0.45	
45	PWM	0.25	0.29	0.33	0.72	0.89	0.80	0.69	0.52	0.41	1.94 (<)
	L-M (i)	0.25	0.30	0.35	0.81	1.02	0.90	0.78	0.58	0.44	
	L-M (d)	0.32	0.37	0.42	0.84	0.99	0.86	0.74	0.55	0.42	
	ML	0.19	0.24	0.29	0.71	0.91	0.82	0.71	0.53	0.41	
	MOM	0.20	0.24	0.28	0.70	0.88	0.78	0.67	0.50	0.38	
61	PWM	0.19	0.23	0.28	0.71	0.91	0.81	0.70	0.53	0.40	1.91 (<)
	L-M (i)	0.18	0.25	0.31	0.82	1.06	0.93	0.79	0.58	0.43	
	L-M (d)	0.23	0.30	0.36	0.85	1.04	0.90	0.76	0.55	0.41	
	ML	0.18	0.24	0.29	0.73	0.95	0.86	0.76	0.58	0.45	
	MOM	0.17	0.22	0.28	0.75	1.00	0.92	0.81	0.62	0.48	
233	PWM	0.18	0.23	0.28	0.74	0.97	0.87	0.76	0.57	0.44	1.95 (<)
	L-M (i)	0.17	0.23	0.29	0.82	1.08	0.96	0.83	0.62	0.47	
	L-M (d)	0.22	0.29	0.35	0.85	1.06	0.92	0.79	0.58	0.44	
	ML	0.23	0.27	0.31	0.71	0.89	0.80	0.69	0.52	0.40	
	MOM	0.23	0.27	0.31	0.71	0.89	0.80	0.69	0.52	0.40	
265	PWM	0.25	0.28	0.32	0.69	0.86	0.75	0.65	0.48	0.37	1.88 (<)
	L-M (i)	0.25	0.30	0.35	0.84	1.06	0.92	0.79	0.57	0.42	
	L-M (d)	0.31	0.36	0.41	0.87	1.02	0.88	0.74	0.54	0.40	
	ML	0.23	0.28	0.33	0.75	0.97	0.88	0.78	0.61	0.48	
	MOM	0.23	0.27	0.32	0.75	0.96	0.88	0.77	0.60	0.47	
355	PWM	0.22	0.27	0.32	0.76	0.99	0.91	0.80	0.62	0.49	1.97 (<)
	L-M (i)	0.21	0.27	0.33	0.83	1.10	1.00	0.88	0.68	0.53	
	L-M (d)	0.27	0.34	0.39	0.86	1.07	0.96	0.84	0.64	0.51	
	ML	0.12	0.19	0.25	0.82	1.16	1.09	0.97	0.77	0.61	
	MOM	0.12	0.19	0.25	0.81	1.15	1.08	0.96	0.76	0.60	
55	PWM	0.12	0.18	0.24	0.82	1.15	1.07	0.95	0.74	0.58	2.15 (>)
	L-M (i)	0.11	0.18	0.24	0.79	1.11	1.02	0.90	0.68	0.53	
	L-M (d)	0.14	0.21	0.28	0.82	1.09	0.98	0.86	0.65	0.50	
	ML	0.13	0.20	0.26	0.85	1.20	1.14	1.03	0.82	0.66	
	MOM	0.12	0.20	0.26	0.86	1.23	1.17	1.06	0.85	0.69	
235	PWM	0.11	0.18	0.25	0.91	1.36	1.30	1.18	0.96	0.79	2.19 (>)
	L-M (i)	0.10	0.17	0.23	0.81	1.21	1.16	1.05	0.86	0.71	
	L-M (d)	0.13	0.21	0.28	0.84	1.17	1.10	0.99	0.80	0.66	
	ML	0.13	0.20	0.26	0.83	1.16	1.10	0.99	0.78	0.63	
	MOM	0.13	0.20	0.26	0.83	1.17	1.10	0.99	0.78	0.63	
262	PWM	0.14	0.20	0.26	0.81	1.12	1.03	0.92	0.71	0.57	2.16 (>)
	L-M (i)	0.13	0.20	0.26	0.80	1.12	1.03	0.91	0.70	0.55	
	L-M (d)	0.17	0.24	0.31	0.83	1.09	0.98	0.86	0.66	0.52	

gains in most cases ($R < 1.0$) and the regional gain of site of interest increases as the return period increases and decreases around specific nonexceedance probabilities. There is no regional gain for some specific nonexceedance probabilities, especially when the single shape parameter γ_s is bigger than the regional value γ ($R \geq 1.0$). The regional gain could be obtained in most cases that γ_s is smaller than the regional

Table 6. The Ratio of Variance for Various Estimation Techniques (case 3)
Regional Shape Parameter γ = 1.81, Autocorrelation Coeff. = 0.463

Site	q	0.01	0.05	0.1	0.5	0.7826	0.9	0.95	0.99	0.999	γ_s
	ML	0.11	0.18	0.24	0.81	1.15	1.08	0.97	0.76	0.60	
	MOM	0.11	0.17	0.24	0.81	1.16	1.09	0.97	0.76	0.60	
45	PWM	0.12	0.18	0.24	0.78	1.08	0.99	0.87	0.67	0.52	1.94 (>)
	L-M (i)	0.11	0.18	0.24	0.77	1.08	1.00	0.88	0.67	0.51	
	L-M (d)	0.34	0.42	0.48	0.88	1.03	0.93	0.83	0.69	0.59	
	ML	0.10	0.16	0.22	0.78	1.10	1.03	0.91	0.70	0.54	
	MOM	0.10	0.16	0.22	0.77	1.08	1.00	0.88	0.67	0.51	
61	PWM	0.09	0.16	0.22	0.78	1.10	1.02	0.90	0.68	0.52	1.91 (>)
	L-M (i)	0.09	0.15	0.21	0.77	1.09	1.00	0.87	0.65	0.49	
	L-M (d)	0.24	0.32	0.39	0.83	1.01	0.90	0.79	0.62	0.51	
	ML	0.09	0.16	0.22	0.81	1.16	1.09	0.98	0.76	0.59	
	MOM	0.08	0.15	0.22	0.83	1.24	1.18	1.06	0.83	0.66	
233	PWM	0.09	0.15	0.22	0.81	1.17	1.09	0.96	0.75	0.58	1.95 (>)
	L-M (i)	0.08	0.14	0.20	0.78	1.13	1.04	0.91	0.70	0.54	
	L-M (d)	0.24	0.33	0.40	0.87	1.05	0.94	0.83	0.66	0.55	
	ML	0.11	0.18	0.24	0.78	1.08	1.00	0.89	0.68	0.53	
	MOM	0.11	0.17	0.23	0.78	1.09	1.02	0.90	0.69	0.53	
265	PWM	0.12	0.18	0.24	0.76	1.03	0.94	0.82	0.62	0.48	1.88 (>)
	L-M (i)	0.12	0.18	0.24	0.80	1.10	1.00	0.86	0.64	0.48	
	L-M (d)	0.34	0.41	0.48	0.89	1.03	0.91	0.79	0.63	0.52	
	ML	0.11	0.18	0.24	0.82	1.18	1.11	1.00	0.79	0.63	
	MOM	0.11	0.17	0.24	0.82	1.18	1.12	1.00	0.79	0.63	
335	PWM	0.10	0.17	0.23	0.83	1.20	1.13	1.01	0.79	0.63	1.97 (>)
	L-M (i)	0.10	0.16	0.22	0.77	1.14	1.07	0.96	0.75	0.59	
	L-M (d)	0.30	0.40	0.47	0.90	1.08	1.01	0.92	0.79	0.71	
	ML	0.06	0.13	0.20	0.91	1.41	1.38	1.26	1.02	0.82	
	MOM	0.06	0.13	0.20	0.91	1.42	1.39	1.27	1.02	0.83	
55	PWM	0.06	0.13	0.19	0.9	1.39	1.34	1.21	0.97	0.77	2.15 (>>)
	L-M (i)	0.05	0.11	0.17	0.76	1.18	1.13	1.01	0.79	0.62	
	L-M (d)	0.15	0.24	0.31	0.85	1.11	1.02	0.91	0.75	0.64	
	ML	0.07	0.14	0.21	0.94	1.47	1.45	1.33	1.09	0.89	
	MOM	0.06	0.13	0.20	0.96	1.52	1.51	1.39	1.15	0.94	
235	PWM	0.14	0.25	0.33	0.86	1.14	1.10	1.03	0.91	0.85	2.19 (>)
	L-M (i)	0.14	0.25	0.33	0.86	1.14	1.10	1.03	0.91	0.85	
	L-M (d)	0.14	0.25	0.33	0.86	1.14	1.10	1.03	0.91	0.85	
	ML	0.07	0.14	0.21	0.92	1.43	1.40	1.28	1.04	0.84	
	MOM	0.07	0.14	0.21	0.92	1.44	1.42	1.30	1.05	0.86	
262	PWM	0.07	0.14	0.2	0.88	1.35	1.30	1.17	0.93	0.75	2.16 (>)
	L-M (i)	0.06	0.12	0.18	0.75	1.15	1.10	0.99	0.78	0.62	
	L-M (d)	0.18	0.28	0.35	0.86	1.10	1.02	0.92	0.77	0.67	

value γ, but the single site analysis is superior to the regional analysis around $q =$ 0.78265 as shown in Tables 5 and 6. Also, the asymptotic gain due to regionalization decreases as the difference between γ_s and γ increases. Note that the nonexceedance probability q=0.78265 is the theoretical probability that does not give the regional gain in Eq. (4), in which the second term of the numerator becomes zero for this specific q. The nonexceedance probability band is expanded as sites in a group are increased, in other words, heterogeneity is increased.

The variances of each site in a region from single site and regional analyses are shown in Fig. 3 through Fig. 8. For $q =$ 0.99, there are always regional gains for all methods. However, there are some sites which do not have regional gain for $q =$ 0.78265 depending on the estimated shape parameter.

In Fig. 8, site number 124 of case 3 has the single shape parameter greater than regional shape parameter, therefore it has the

Table 6. The Ratio of Variance for Various Estimation Techniques (case 3)
(continued) Regional Shape Parameter γ = 1.81, Autocorrelation Coeff. = 0.463

Site	q	0.01	0.05	0.1	0.5	0.7826	0.9	0.95	0.99	0.999	γ_s
	ML	0.19	0.23	0.28	0.68	0.84	0.74	0.63	0.46	0.34	
	MOM	0.16	0.21	0.26	0.69	0.89	0.79	0.67	0.49	0.36	
18	PWM	0.19	0.23	0.27	0.68	0.85	0.74	0.63	0.45	0.33	1.66 (<)
	L-M (I)	0.18	0.24	0.30	0.80	1.01	0.87	0.74	0.53	0.38	
	L-M (d)	0.52	0.53	0.57	0.89	0.96	0.82	0.71	0.53	0.42	
	ML	0.31	0.31	0.33	0.64	0.71	0.60	0.50	0.34	0.24	
	MOM	0.31	0.29	0.32	0.61	0.69	0.56	0.45	0.30	0.21	
63	PWM	0.35	0.32	0.34	0.61	0.67	0.55	0.45	0.30	0.21	1.52 (<<)
	L-M (I)	0.35	0.37	0.42	0.86	0.98	0.80	0.64	0.42	0.28	
	L-M (d)	0.97	0.79	0.76	0.95	0.94	0.76	0.62	0.43	0.31	
	ML	0.16	0.20	0.26	0.69	0.87	0.77	0.66	0.48	0.35	
	MOM	0.14	0.19	0.24	0.69	0.90	0.80	0.68	0.49	0.36	
114	PWM	0.16	0.20	0.25	0.68	0.86	0.75	0.63	0.46	0.33	1.69 (<)
	L-M (I)	0.15	0.22	0.28	0.81	1.04	0.90	0.75	0.52	0.37	
	L-M (d)	0.40	0.45	0.49	0.89	0.98	0.82	0.69	0.50	0.38	
	ML	0.06	0.12	0.20	0.97	1.55	1.55	1.42	1.17	0.96	
	MOM	0.05	0.12	0.19	0.99	1.61	1.61	1.49	1.23	1.01	
124	PWM	0.05	0.12	0.19	0.97	1.58	1.55	1.42	1.16	0.95	2.25 (>>)
	L-M (I)	0.04	0.10	0.15	0.76	1.21	1.19	1.08	0.88	0.72	
	L-M (d)	0.13	0.22	0.29	0.84	1.13	1.07	0.99	0.84	0.75	
	ML	0.20	0.24	0.28	0.67	0.81	0.70	0.60	0.43	0.31	
	MOM	0.21	0.24	0.28	0.64	0.78	0.66	0.55	0.38	0.27	
307	PWM	0.21	0.24	0.28	0.66	0.79	0.68	0.57	0.40	0.29	1.63 (<)
	L-M (I)	0.20	0.26	0.32	0.82	1.01	0.85	0.71	0.49	0.34	
	L-M (d)	0.56	0.55	0.58	0.91	0.97	0.80	0.67	0.48	0.37	
	ML	0.13	0.19	0.24	0.71	0.92	0.82	0.71	0.52	0.38	
	MOM	0.13	0.18	0.24	0.68	0.89	0.78	0.66	0.48	0.35	
327	PWM	0.13	0.19	0.24	0.69	0.90	0.79	0.68	0.49	0.36	1.74 (<)
	L-M (I)	0.12	0.19	0.26	0.80	1.04	0.91	0.76	0.54	0.39	
	L-M (d)	0.32	0.38	0.44	0.87	1.00	0.85	0.71	0.52	0.39	
	ML	0.29	0.29	0.32	0.61	0.66	0.54	0.44	0.30	0.20	
	MOM	0.24	0.25	0.28	0.60	0.70	0.57	0.46	0.30	0.20	
344	PWM	0.31	0.29	0.31	0.59	0.64	0.52	0.42	0.27	0.18	1.46 (<<)
	L-M (I)	0.32	0.36	0.42	0.88	0.97	0.77	0.61	0.39	0.25	
	L-M (d)	0.76	0.67	0.67	0.94	0.92	0.71	0.56	0.37	0.25	

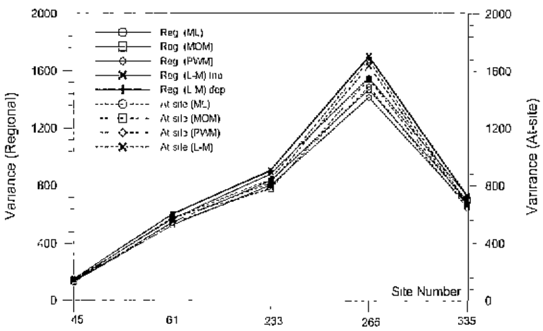

Fig. 3. Comparison of Variances for 5 Sites at $q = 0.78265$

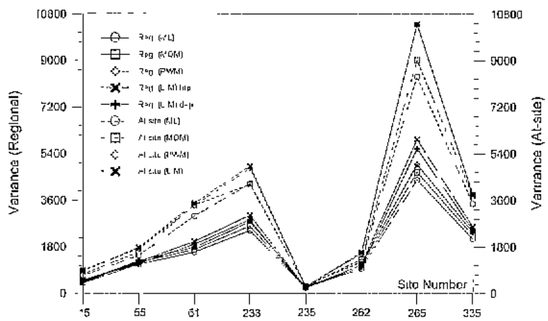

Fig. 6. Comparison of Variances for 8 Sites at $q = 0.99$

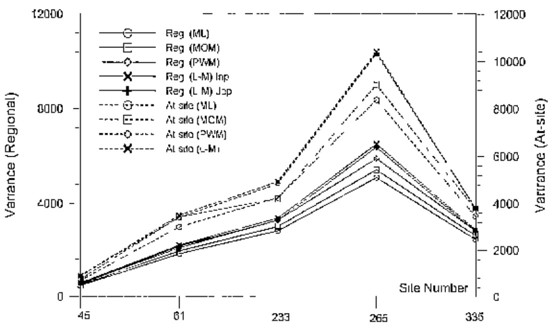

Fig. 4. Comparison of Variances for 5 Sites at $q = 0.99$

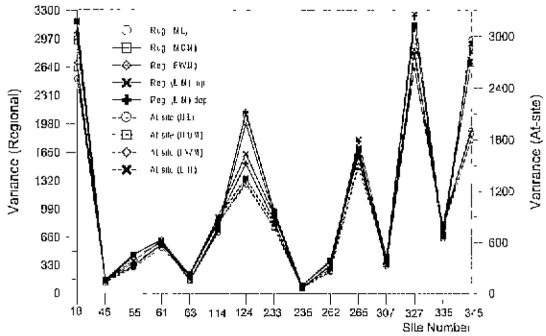

Fig. 7. Comparison of Variances for 15 Sites at $q = 0.78265$

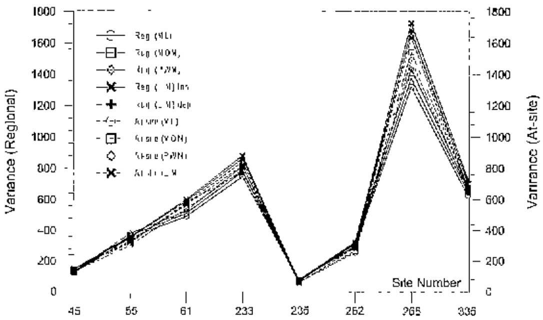

Fig. 5. Comparison of Variances for 8 Sites at $q = 0.78265$

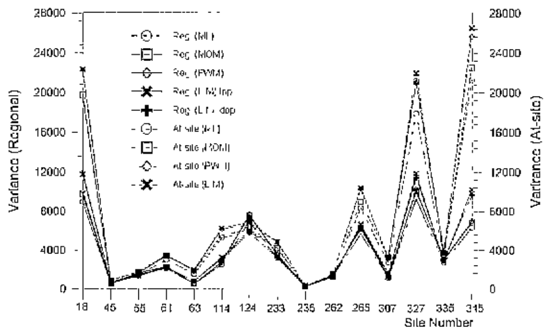

Fig. 8. Comparison of Variances for 15 Sites at $q = 0.99$

worst regional gain in a 15 site region at $q = 0.78265$ and does not give the regional even though $q = 0.99$.

4. CONCLUSIONS

The comparison between the single site and regional flood analyses have been performed in the sense of variance of quantile for a given flood data by using both asymptotic variance regional analysis and Hosking's regional L-moment algorithm. Both schemes show the similar results. Asymptotic formulas always provide the regional gains when the estimated shape parameter γ is smaller than regional shape parameter γ_s, while Hosking's algorithm does not. However, Hosking's algorithm provides smaller variances when the heterogeneity is increased and γ_s is bigger than γ for $q =$ 0.78265. For some specific sites and nonexceedance probabilities, no regional gains exist depending on the estimated single site and regional shape parameters. Furthermore, the range of no regional gain is increased as the differences between the single and regional shape parameters increase especially when $\gamma_s > \gamma$. Even though the selected sites may satisfy Hosking's criteria, regional analysis may not give a regional gain for specific nonexceedance probabilities. Thus, we have to check the estimated shape parameters from single site and regional flood frequency analysis because there might be no regional gain depending on the variability of the shape parameter for the 2-parameter Weibull model. Similar results can be expected for other probability models although further research is needed.

REFERENCES

Benson, M.A. (1962). *Evolution of methods for evaluating the occurrence of floods.* U. S. Geological Survey Water Supply Paper 1580A.

Bobee, B., Cavadias, G., Ashkar, F., Bernier, J., and Rasmussen, P. (1993). "Towards a systematic approach to comparing distributions used in flood frequency analysis." *Journal of Hydrology*, Vol. 142, No. 5, pp. 121-136.

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. 979-970.

Cunnane, C. (1989). *Statistical distributions for flood frequency analysis.* WMO Operational Hydrological Report No. 33.

Curtis, G.W. (1987). Technique for estimating flood-peak discharges and frequencies on rural streams in basins in Illinois, U.S. Geological Survey Water-Resources Investigation Report 87-991.

Darymple, T. (1960). "*Flood-frequency analyses*" U. S. Geological Survey Water Supply Paper 1543A.

Filliben, J.J. (1975). "The probability plot correlation coefficient test for normality." *Technometrics*, Vol. 17, No. 1, pp. 111-117.

Greis, N.P. (1983). "Flood frequency analysis: A review of 1979-1982." *Review of Geophysics and Space Physics*, Vol. 21, No. 3, pp. 699-705.

Gries, N.P. and Wood, E.F. (1981). "Regional flood-frequency estimation and network design." *Water Resources Research*, Vol. 17, No. 4, pp. 1167-1177.

Heo, J.H., Boes, D.C., and Salas, J.D. (1990). *Regional flood-frequency modeling and estimation.* No. 101, Water Resources Paper, Colorado State University, Fort Collins, Colorado, USA.

Hosking, J.R.M. (1990). "L-moments : Analysis and estimation of distribution using linear combinations of order statistics." *Journal of Statistical Society, Series B,* Vol. 52, No. 1, pp. 105-124.

Hosking, J.R.M. (1996). *Fortran routines for use with the method of L-moments, version 3*, IBM Research Report RC 20525, IBM Research Division, Yorktown Heights, N. Y.

Hosking, J.R.M. and Wallis, J.R. (1988). "The effect of intersite dependence on regional frequency analysis." *Water Resources Research*, Vol. 24, No. 4, pp. 588-600.

Hosking, J.R. M. and Wallis, J.R. (1993). "Some statistics useful in regional frequency analysis." *Water Resources Research*, Vol. 29, No. 2, pp. 271-281.

Hosking, J.R.M. and Wallis, J.R. (1997). *Regional frequency analysis: An approach based on L-moments.* Cambridge University Press.

Hosking, J.R.M., Wallis, J.R., and Wood, E. F. (1985). "An appraisal of the regional flood frequency procedure in the UK flood studies report." *Hydroogical Sciences Journal*, Vol. 30, No. 1, pp. 85-109.

Lettenmaier, D.P. and Potter, K.W. (1985). "Testing flood frequency estimation methods using a regional flood generation model." *Water Resources Research*, Vol. 21, No. 2, pp. 1903-1914.

Parida, B.P., Kachiroo, R.K., and Shrestha, D.B. (1998). "Regional flood frequency analysis of Mahi-Sabarmati basin (Subzone 3-a) using index flood procedure with L-moments." *Water Resources Management*, Vol. 1, No. 12, pp 1-12.

Rao, A.R. and Hamed K.H. (1997). "Regional frequency analysis of Wabash river flood data by L-moments." *Journal of Hydrologic Engineering*, Vol. 2, No. 4, pp. 169-179.

Rossi, F., Fiorentino, M., and Versace, P. (1984). "Two-component extreme value distribution for flood frequency analysis." *Water Resources Research*, Vol. 20, No. 7, pp. 847-856.

Salas, J.D., Delleur, J.W., Yevjevich, V., and Lane, W.L. (1980). *Applied modeling of hydrologic time series.* Water Resources Publications, Littleton, Colorado.

Stedinger, J.R. and Tasker, G.D. (1985). "Regional hydrological analysis 1. Ordinary, weighted and generalized least squares compared." *Water Resources Research*, Vol. 21, No. 9, pp. 1421-1432.

Stedinger, J.R. (1983). "Estimating a regional flood frequency distribution." *Water Resources Research*, Vol. 19, No. 2, pp. 435-444.

Vogel, R.M. and Kroll, C.N. (1989). "Low-flow frequency analysis using probability plot correlation coefficients." *Journal of Water Resources Planning and Management*, ASCE, Vol. 115, No. 3, pp. 338-357.

Vogel, R.M. and Fennessay, N.M. (1993). "L-moment diagrams should replace product moment diagrams." *Water Resources Research*, Vol. 29, No. 6, pp. 1745-1752.

Wallis, J.R. (1980). "Risk and uncertainties in the evaluation of flood events for the design of hydraulic structures." *Seminar on Extreme Hydrological Events · Floods and Droughts*, Centro di Cult. Sci. E. Majorana, Erice, Italy.

Jun-Haeng Heo and Dong-Jin Lee, Department of Civil Engineering., Yonsei University, Shinchon-Dong, Seodaemun-Gu, Seoul, 120-749, Korea.
(email: jhheo@yonsei.ac.kr)

Kyung-Duk Kim, Korea Infrastructure Safety and Technology Cooperation, 2311 Daehwa-Dong, Ilsan-Gu, Goyang, Kyunggi-Do, 411-410, Korea.

(Received September 2, 1999; revised October 2, 1999; accepted November 3, 1999.)