강우 지역빈도해석의 적용성 연구

Study on Rainfall Regional Frequency Analysis

신홍준*, 남우성**, 허준행***, 김경덕****

Hong Joon Shin, Woo Sung Nam, Jun Haeng Heo, Kyung Duk Kim

Abstract

At-site analysis is not appropriate if the record length is shorter than target return period T. If the record length is longer than 2T years, then at-site analysis may be sufficient(Institute of Hydrology, 1999). However, in such a case, regional frequency analysis is recommended for purpose of comparison. Record lengths of annual maximum rainfall data in Korea are usually shorter than 50 years. It is therefore essential to apply regional frequency analysis for estimating rainfall quantiles of more than 100 years return period.

In this research, regional rainfall frequency analysis is performed for hourly rainfall data of South Korea. Homogeneous regions are idntified by clusgter analysis which is a standard method of statistical multivariate analysis for dividing a data set into groups. An appropriate distribution is chosen by goodness-of-fit test. GLO is found to be an appropriate distribution as a result of goodness-of-fit measure (Hosking & Wallis, 1997). Simulation experiments are performed to check the performance of frequency analysis techniques. The effects of discordant sites on quantiles are considered.

Key words: Regional Frequency Analysis, Cluster Analysis, Goodness-Of-Fit Test, Discordancy, Simulation Experiments

1. Introduction

Sufficient sample needs to fit an appropriate probability distribution and estimate parameters. In South Korea, design storm is unreliable because record length is not enough to estimate quantiles for long return periods (>100 year). Regional frequency analysis resolves this problem.

In this research, regional frequency analysis is applied to annual precipitation maxima of South Korea in order to improve reliability on rainfall quantiles. For those purposes, 1) homogeneous regions are identified and an appropriate distribution is selected, 2) differences by analysis methods, at-site frequency analysis, index flood method, and regional shape estimation method based on L-moments, are compared and analyzed, 3) the performance of regional frequency analysis is checked by Monte Carlo simulation and 4) the effects of discordant sites on

* 정회원・연세대학교 토목공학과 박사과정・E-mail : sinong@yonsei.ac.kr
** 정회원・연세대학교 토목공학과 박사과정・E-mail : nws77@yonsei.ac.kr
*** 정회원・연세대학교 사회환경시스템공학부 토목・환경공학 전공 교수・E-mail : jhheo@yonsei.ac.kr
**** 정회원·한국시설안전기술공단 댐·항만실 차장·공학박사E-mail : kkd@kistec.or.kr

quantiles are considered.

2. Data

Annual precipitation maxima are taken from 378 sites of Korea Meteorological Administration, Ministry of Construction and Transportation, and Korea Water Resources Corporation for 10 durations (1, 2, 3, 6, 9, 12, 15, 18, 24, 48 hr). Sites with record length less then 10 years are excluded from analysis.

3. Regionalization

Number of sites in a region can be chosen and optimal value can be calculated by Fuzzy cluster analysis. In present study, 378 sites are divided into 14 homogeneous regions by Fuzzy-c means technique. The following site characteristics are used in the cluster analysis.
1) Longitude (°)
2) Latitude (°)
3) Altitude (m)
4) Annual mean precipitation over 12 months (mm)
5) Annual mean precipitation over rainy season (mm)
The homogeneous regions are identified by cluster analysis and the heterogeneity measure H of each region of Hosking and Wallis (1997) are estimated. The heterogeneity test shows that 13 regions are homogeneous ($H<1$) and region NAK2 is possibly heterogeneous ($1 \leq H < 2$).

4. Identification of the Appropriate Regional Frequency Distribution

The 6 distributions are applied: Gumbel, generalized extreme value (GEV), lognormal-3 (LN3), Pearson type Ⅲ (PTⅢ), generalized Pareto (GPA), generalized logistic (GLO). Their parameters are estimated and an appropriate distribution is identified by the goodness-of-fit measure Z of Hosking and Wallis (1997). The goodness-of-fit test shows that GLO is accepted in most regions except HAN1. It is a case in the other durations.

5. Accuracy of Quantiles

Regional relative root mean square error (RRMSE) is estimated by Monte Carlo simulation based on GLO. Simulation shows that the performance of regional frequency analysis is better than that of at-site frequency analysis at the nonexceedance probability 0.995 (return period 200 year) for all regions. Especially, as nonexceedance probability of at-site frequency analysis gets larger, RRMSE increases exponentially. This indicates that at-site frequency analysis is inappropriate to quantile estimation for long return periods.
The performance of index flood method is better than that of regional shape estimation

method in 10 regions while that of regional shape estimation is better in 4 regions. Regional shape estimation performs better in case of L-CV less than 0.25 (NAK3, KUM1, YOU2, SUM) as seen in Fig. 2~Fig. 5 or range of L-CV more than 0.15 (YOU2, SUM) as seen in Fig. 4 and Fig. 5. These results agree with Stedinger and Lu (1995) and Hosking and Wallis (1997). Increase in number of sites in a region does not continuously improve the performance of regional frequency analysis. This tendency is found in case of HAN1 with many sites and TON1 with a few sites as seen in Fig. 1 and Fig. 6. This indicates that there is little gain by using regions larger than about 20 sites (Hosking and Wallis, 1997).

In conclusion index flood method performs better in case of short record length, and L-CV and L-skewness from 0.2 to 0.4 like South Korea while regional shape estimation does better in case on L-CV less than 0.25 and L-skewness more than 0.15.

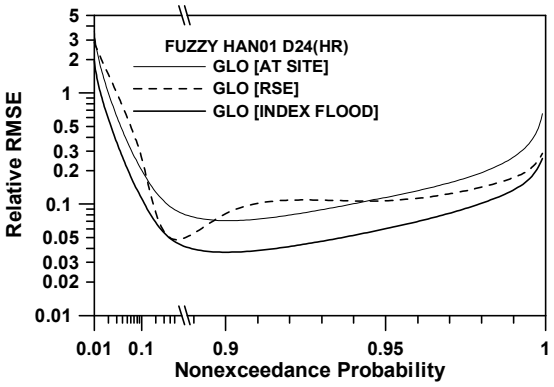

그림 1 RRMSE of HAN1

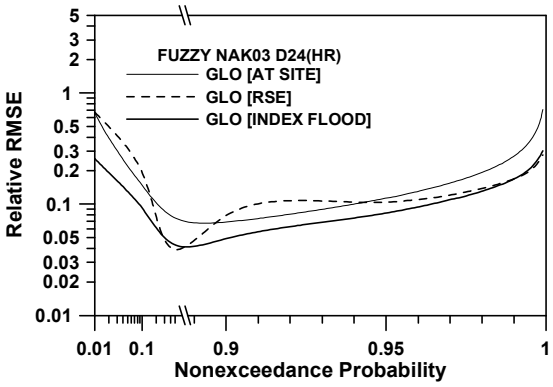

그림 2 RRMSE of NAK3

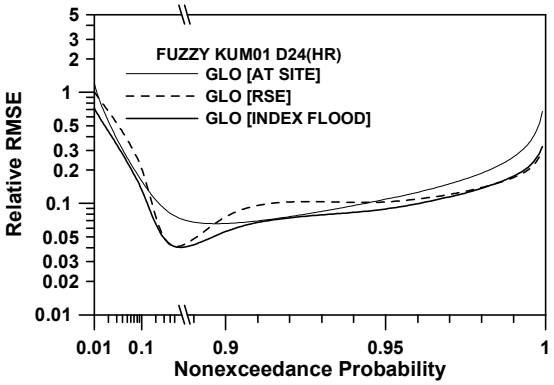

그림 3 RRMSE of KUM1

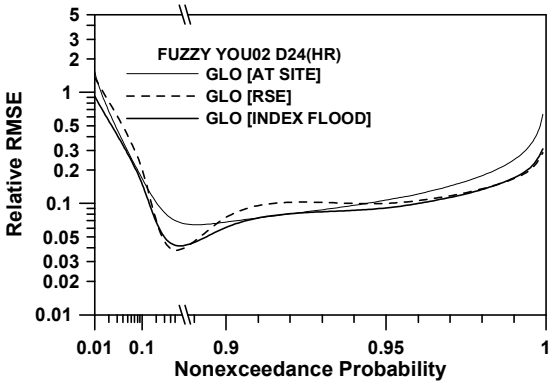

그림 4 RRMSE of YOU2

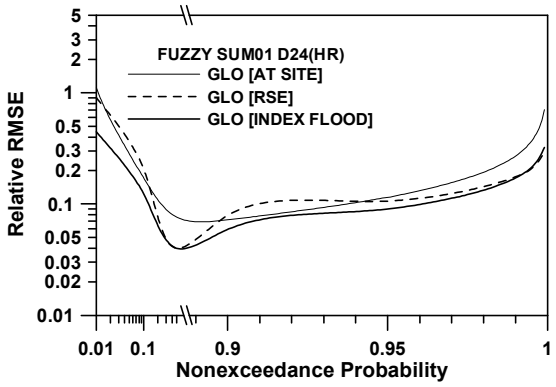

그림 5 RRMSE of SUM

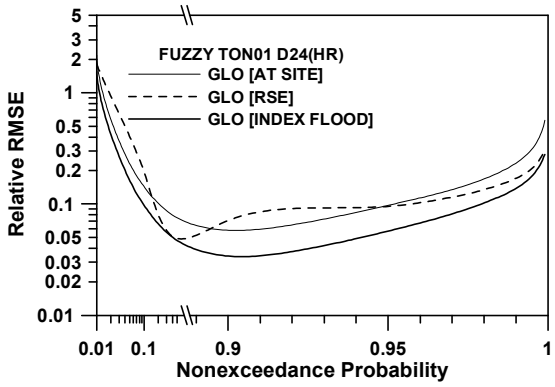

그림 6 RRMSE of TON

6. Effects of Discordant Sites

Regional frequency analysis including discordant sites is compared with that except discordant sites. GLO is the appropriate frequency distribution of South Korea in including discordant sites. In this study, regional frequency analysis is carried out including them in order to effects of discordant sites.

Discordancy test shows that scale parameter has more significant effects on quantiles than shape parameter. The result indicates L-CV has connection with quantiles. That is, quantiles increase as L-CV gets larger. It is not met with cases of duration of 48 hr in HAN2, 12 hr in HAN3, 12 hr in KUM2, 1 hr in YOU1. It might be caused by the characteristics of discordant sites. This is supported by Table 1. Sites in Table 1 are found to be discordant due to little difference between L-CV of discordant sites and regional average, and relatively large difference between L-skewness and L-kurtosis. Quantiles are affected by L-CV except this case.

It is known by Kumar and Chatterjee (2005) that quantiles including discordant sites are larger 8% more than that except them. However, in this study, it is found that quantiles depend on L-CV in this study.

Table 1. L-moment ratio of Discordant Sites

Region	Discordant Sites	L-CV	L-skewness	L-kurtosis	scale parameter	shape parameter
HAN2 (48hr)	HR017070	0.2511	0.4705	0.4259	-0.470	0.752
	Regional Average	0.2827	0.2078	0.1627	0.263	-0.208
HAN3 (12hr)	HR015150	0.1329	0.1781	0.2570	-0.178	0.106
	Regional Average	0.2458	0.2010	0.1899	0.230	-0.201
	HR015160	0.3295	0.4474	0.3639	-0.447	0.807
	Regional Average	0.2458	0.2010	0.1899	0.230	-0.201
KUM2 (12hr)	HR033510	0.1985	0.3210	0.4521	-0.321	0.389
	Regional Average	0.2452	0.2411	0.2334	0.222	-0.241
YOU1 (1hr)	HR057600	0.2161	-0.0052	0.2793	0.005	0.124
	Regional Average	0.2117	0.1876	0.1887	0.200	-0.188

8. Conclusions

At-site and regional frequency analysis by index flood method and regional shape estimation method and Monte Carlo simulation are carried out for annual precipitation maxima of South Korea. The following are concluded.

(1) Fuzzy cluster analysis divides 378 sites over 10 years of South Korea into 14 regions. All regions are found to be homogeneous by the heterogeneity measure H. GLO is identified as the appropriate frequency distribution by the goodness-of-fit measure Z.

(2) There is little difference between at-site and region frequency analysis for long record length. However, quantiles by at-site frequency analysis are not robust for short record length.

(3) Quantiles by index flood method are larger than those by regional shape estimation method in case of L-CV over 0.25 or range of L-CV over 0.15.

(4) Simulation experiments show that regional frequency analysis is more appropriate in South Korea and regional shape estimation method is more robust in case of quantile estimation for long return periods, small regional L-CV, wide range of L-CVs in a region, and long record length.

(5) L-CV has a significant effect on quantile estimation.

References

1. Hosking, J. R. M. and Wallis, J. R. (1997). *Regional Frequency Analysis*. Cambridge University Press.

2. Institute of Hydrology (1999). *Flood Estimation Handbook*, Institute of Hydrology, Wallingford, UK.
3. Kumar, R. and Chatterjee, C. (2005). "Regional Flood Frequency Analysis Using L-moments for North Brahmaputra Region of India", *Journal of Hydrologic Engineering*, Vol. 10, No. 1, pp.1-7.
4. Stedinger, J. R. and L. H. Lu. (1995). "Appraisal of regional and index flood quantile estimators", *Stochastic Hydrology and Hydraulics*, Vol. 9, pp.49-75.