Statistical Hydrology

Regional Frequency Analysis

Dr. Young-Oh Kim
Seoul National University

Contents

Concept

- Substitute Space for Time (NRC, 1988)

Hann Prob 4.17

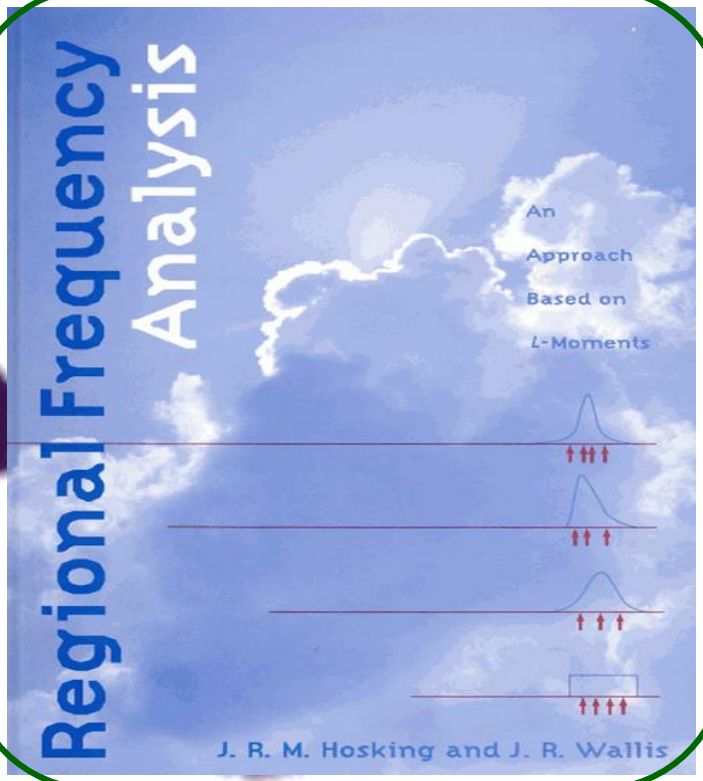
Two widely separated watersheds are selected for a study on peak discharges. If the occurrence of flood flows on the basins can be considered as independent events, what is the probability of experiencing a total of 5, 20-year events on the two watershed in a 10-years period?

At-site vs. Regional Estimates

-The at-site quantile estimates should be the best if their sample sizes are sufficiently large

- Regionalization is the inclusion in frequency analysis of data from sites other than at which quantile estimates are required

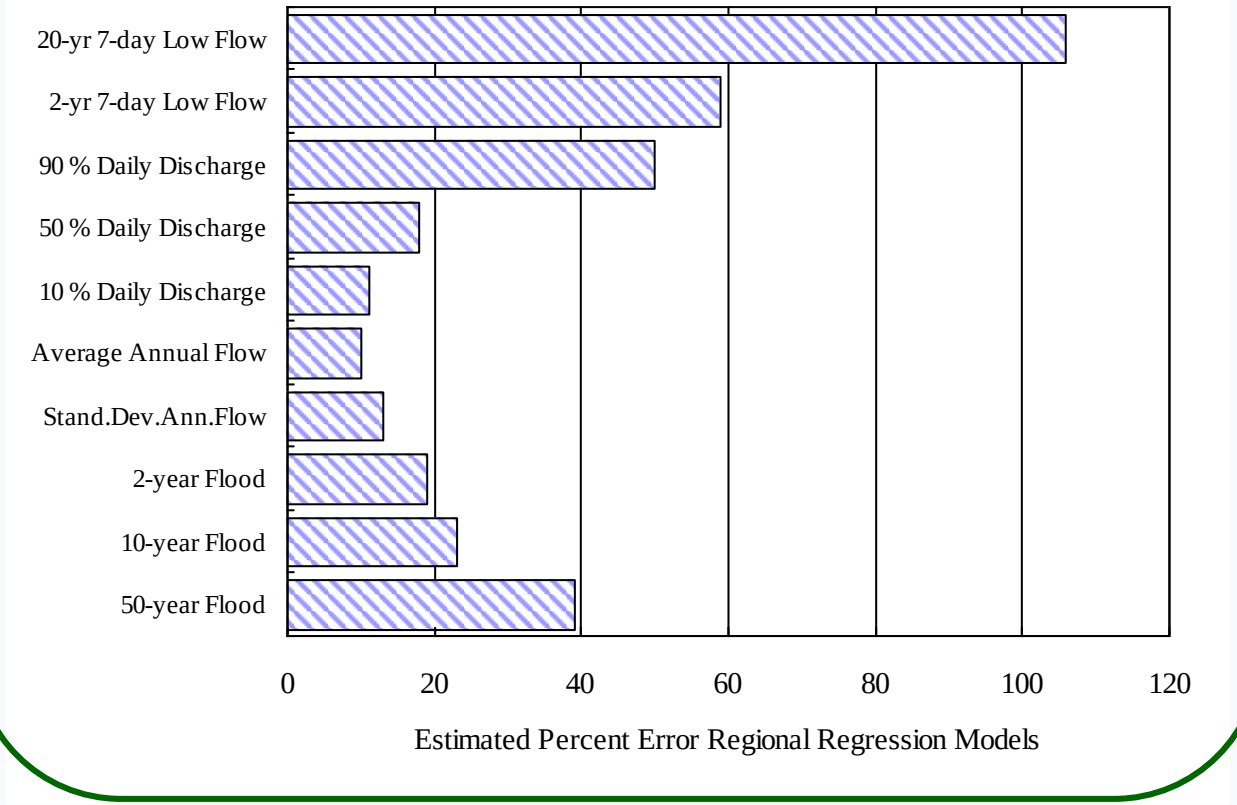
Index Flood Metho



Hosking and Wallis (1997)

- The distributions of floods at different sites in a region are the same except for a scale or index-flood parameter

Index Flood Method: Basic Concept

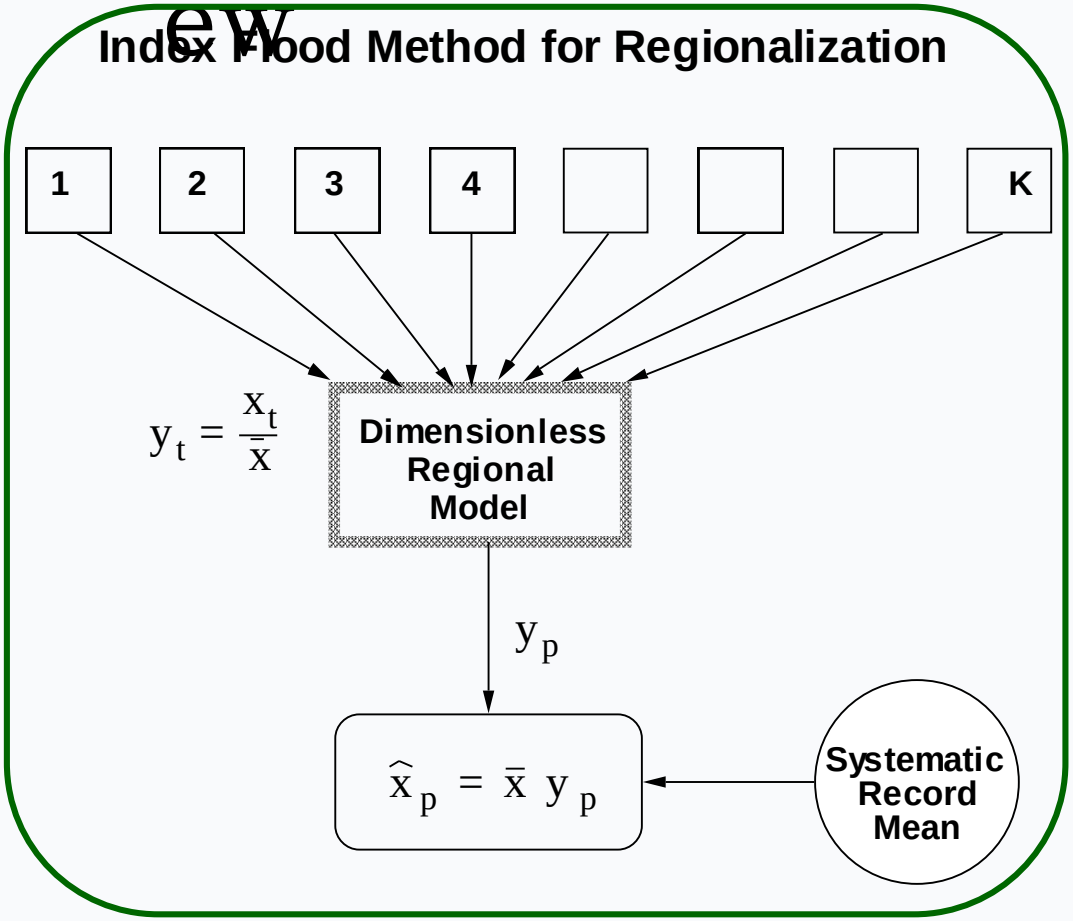


Thomas and Benson (1975)

Index Flood Method: Assumptio

- Observations at any given site are identically distributed
- Observations at any given site are serially independent
- Observations at different sites are independent
- Frequency distributions at different sites are identical apart from a scale factor
- The mathematical form of the region growth curve is correctly specified

Index Flood Method: Overview



Index Flood Method: Procedur

1.Identification of homogeneous regions

2. At each site k compute the three L-moment estimators $\hat{\lambda}_1(k)$, $\hat{\lambda}_2(k)$, $\hat{\lambda}_3(k)$ using the unbiased PWM estimators b_r.
3. To obtain a normalized frequency distribution for the region, compute the regional average of the normalized L moments of order $r = 2$ and 3:

$$\hat{\lambda}_r^R = \frac{\sum_{k=1}^{K} w_k \, [\hat{\lambda}_r(k)/\hat{\lambda}_1(k)]}{\sum_{k=1}^{K} w_k} \qquad \text{for } r = 2, 3 \qquad (18.5.1)$$

For $r = 1$, $\hat{\lambda}_1^R = 1$. Here w_k are weights; a simple choice is $w_k = n_k$, where n_k is the sample size for site k. However, weighting by the sample sizes when some sites have much longer records may give them undue influence. A better choice which limits the weight assigned to sites with longer records is

$$w_k = \frac{n_k n_R}{n_k + n_R}$$

where n_k are the sample sizes and $n_R \approx 25$; the optimal value of the weighting parameter n_R depends on the heterogeneity of a region.[138,141]

4. Using the average normalized L moments $\hat{\lambda}_1^R$, $\hat{\lambda}_2^R$, and $\hat{\lambda}_3^R$ in Eqs. (18.2.22) and (18.2.23), determine the parameters and quantiles $\hat{x}_p^R$ of the normalized regional GEV distribution.
5. The estimator of the $100p$ percentile of the flood distribution at any site k is

$$\hat{x}_p(k) = \hat{\lambda}_1^k \, \hat{x}_p^R \qquad (18.5.2)$$

where λ_1^k is the at-site sample mean for site k:

$$\hat{\lambda}_1^k = \frac{1}{n_k} \sum_{i=1}^{n_k} x_i(k)$$

Index Flood Method: Criticisms

- Observations may be correlated in time and space
- The higher moments may be heterogeneous
- Too much efforts are generally required for identification of homogeneous regions

Variants of Regionalization

Variant	Mean	Dispersion	Shape
At-site	at-site	at-site	at-site
Regional shape estimation	at-site	at-site	regional average
Index-flood	at-site	regional average	regional average
Hierarchical	at-site	regional average for subregion	regional average for full region
Fraction membership	at-site	weighted average of regional estimates	
Region of influence	at-site	weighted average of at-site estimates, for sites in site i's region of influence	
Mapping	at-site	estimated function of site characteristics	
Bulletin 17B	at-site	at-site	weighted average of at-site estimated function of site location

Regional

Generalization of Streamflow Characteristics From Drainage-Basin Characteristics

By D. M. THOMAS *and* M. A. BENSON

GEOLOGICAL SURVEY WATER-SUPPLY PAPER 1975

A study of relations for estimating streamflow characteristics from drainage-basin characteristics in four hydrologically differing regions of the conterminous United States

Regional Regressio

TABLE 9.—*Summary of Western region regression relations:*
$Y = a\ A^{b_1} S^{b_2} St^{b_4} E^{b_5} I_{24,2}^{b_6} P^{b_7} Sn^{b_8} F^{b_9} Si^{b_{10}}$

All regression coefficients are statistically significant at the 5 percent level, except those preceded by *d*

Flow Index Y	Regression Constant a*	Regression Coefficients for A	S	St	E	$I_{24,2}$	P	Sn	F	Si	-	Standard Error loge	%
P_{50}	-	-	-	-	-	-	-	-	-	-		0.485	139
	21.5	0.83	-	-	-	-	-	-	-	-		.262	64.0
	6.03	.81	-	-	-	-	-	-	0.84	-		.236	57.0
	2.31	.83	-	-	-0.48	-	-	-	1.20	-		.191	45.3
P_{25}	-	-	-	-	-	-	-	-	-	-		.476	133
	164	.83	-	-	-	-	-	-	-	-		.248	60.3
	251	.85	-	-	- .35	-	-	-	-	-		.226	54.4
	2.64	.84	-	-	- .53	-	-	-	1.12	-		.169	39.8
P_{10}	-	-	-	-	-	-	-	-	-	-		.472	132
	128	.83	-	-	-	-	-	-	-	-		.241	58.4
	200	.85	-	-	- .37	-	-	-	-	-		.215	51.5
	2.96	.84	-	-	- .54	-	-	-	1.03	-		.164	38.6
	.873	.86	-	-	- .66	-	0.50	-	.90	-		.157	36.9
P_5	-	-	-	-	-	-	-	-	-	-		.471	131
	93.9	.83	-	-	-	-	-	-	-	-		.239	57.8
	146	.85	-	-	- .37	-	-	-	-	-		.213	51.0
	3.40	.84	-	-	- .51	-	-	-	.92	-		.174	41.1
	.677	.87	-	-	- .68	-	.66	-	.75	-		.160	37.6
P_2	-	-	-	-	-	-	-	-	-	-		.481	135
	50.1	.83	-	-	-	-	-	-	-	-		.256	62.4
	67.1	.85	-	-	d- .24	-	-	-	-	-		.248	60.0
	.633	.90	-	-	- .63	-	1.33	-	-	-		.199	48.0

Regional Regression

Traditional log-linear model

$$y_i = \alpha + \beta_1 \log(area) + \beta_2 \log(slope) + \ldots + \varepsilon$$

Comments

(i) Are ε's correlated in space?

(ii) Are ε's heteroscedastic?

Bulletin 17B

Uniform approach to flood frequency analysis in United States.

LP-3 distribution + regional estimator log-space skew: addresses variability of sample skew.

Uses skew map provided in 1976:

Averaging ignored lengths of record and sample error.

Effect of cross correlation generally ignored altogether.

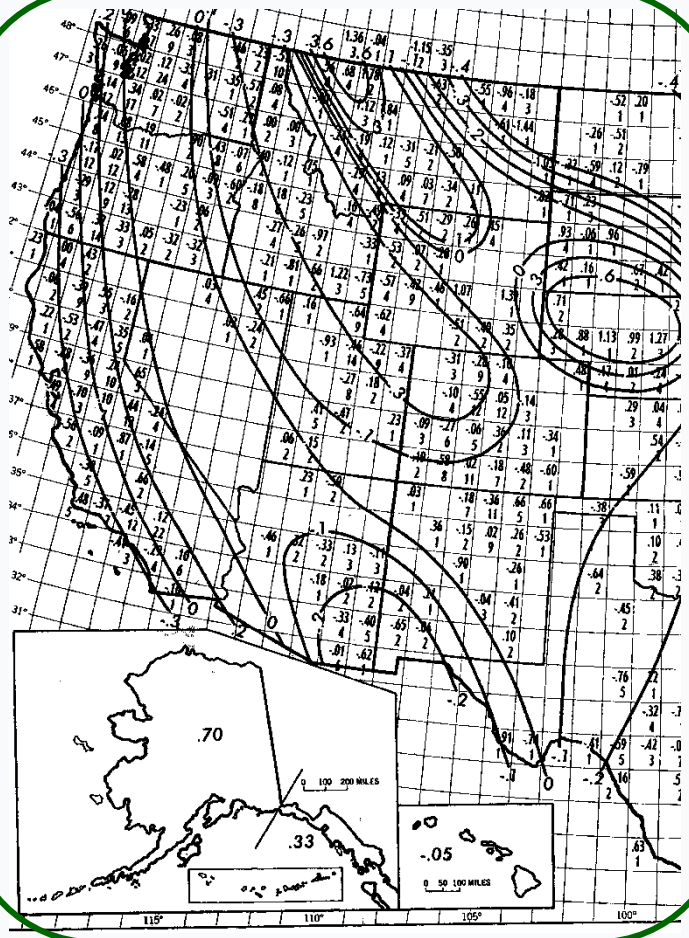
Martins & Stedinger [2002] → cross-correlation among skewness estimators as function cross-correlation of floods

Griffis, V. W. and Stedinger, J. R. (2007) "Evaluation of Flood Frequency Analysis with Bulletin 17B" J.H.E,

Griffis, V. W. and Stedinger, J. R. (2007) "Log-Pearson Type 3 Distribution and Its Application in Flood Frequency Analysis. I: Distribution Characteristics." *J.O.E,* 12(5), pp.492-500.

Bulletin 17B

Skew Map – West Coast



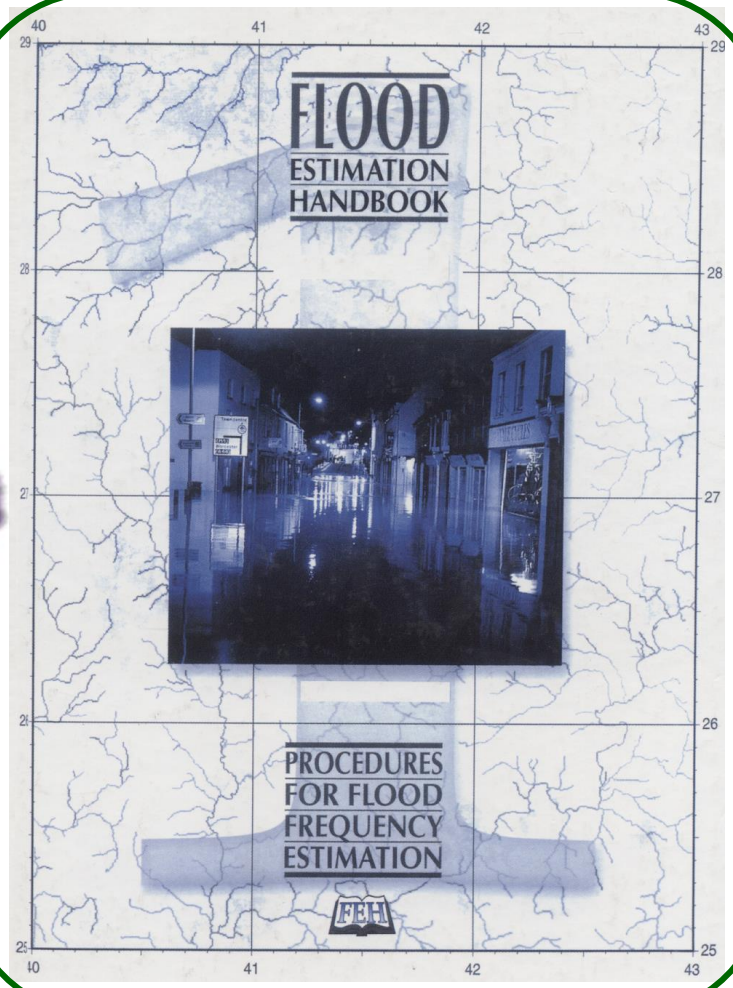
Weighted Skew

$$G_w = \frac{(G)MSE_{\bar{G}} + \bar{G}MSE_G}{MSE_{\bar{G}} + MSE_G}$$

G_w = weighted skew
G = station skew
= regional skew
= 0.302

Regional Skew map
Effective record length = 16 years

Flood Estimation Handboo



Institute of Hydrology, 1999, 5 volumes

Volume 1 contains a preliminary review, an overview of the procedures, a summary of how the methods differ from earlier guidance and an introduction to the data and software accompanying the Handbook. It includes important guidance on the choice of method to suit a particular flood estimation problem

Volume 2 procedure is built on rainfall growth factors, which relate the rainfall depth of a given return period to an index variable. Growth factors derive from the Focused Rainfall Growth Extension (FORGEX) method, which gives precedence to rainfall extremes observed locally and takes account of inter-site dependence in rainfall extremes to obtain estimates for long return periods. The index variable is the median annual maximum, which has been mapped using topographic information.

Flood Estimation Handboo

Volume 3: The statistical procedures presented in this volume use observed flood data to obtain a relationship between flood size and flood rarity (the flood frequency curve). The techniques allow questions such as 'How rare was that flood?' or 'How big is the expected 100-year flood?' to be addressed. When a catchment has a very lengthy series of flood data, the flood frequency curve can be estimated using flood peak data from the catchment alone. More commonly, there are insufficient flood data and the flood frequency curve is estimated using data pooled from a group of similar sites. Methods for selecting sensible sites for pooling, and for combining the data from the pool are presented. Volume 3 extends these methods to allow estimation of flood frequency at ungauged sites. Methods for handling special cases such as urbanised catchments and permeable catchments are also presented

Volume 4 is a comprehensive technical rewriting of the Flood Studies Report rainfall-runoff method. All the information about the method has been brought together, including relevant aspects of the basic methodology, supplementary research and recommendations, and specialist guidance on applications.

Volume 5 describes how descriptors are based on the application of a boundary (defined by a Digital Terrain Model) to digital datasets. Freed from the need to be computationally straightforward, the use of digital data allows catchments to be described with greater subtlety. Included in an appendix are 20 descriptors for 943 gauged catchments. Guidance is also given on how to use the FEH CD-ROM that holds descriptor information for all

Flood Estimation Handboo

<Recommended methods: when, T =< 27>

Length of record	Site analysis	Regional analysis	Shorthand description
< T/2 years	No	Yes	Regional analysis
T/2 ~ T years	For confirmation	Yes	Regional analysis prevails
T ~ 2T years	Yes	Yes	Joint analysis
> 2T years	Yes	For confirmation	Site analysis

<Recommended methods: when, T > 27>

Length of record	Site analysis	Regional analysis	Shorthand description
< 14 years	No	Yes	Regional analysis
14 ~ T years	For confirmation	Yes	Regional analysis prevails
T ~ 2T years	Yes	Yes	Joint analysis
> 2T years	Yes	For confirmation	Site analysis

FFA Old and New Paradig