최적강우강도식의 결정을 위한 회기법
Regression Methods for Determining Optimal Rainfall Intensity Formulas

O 안 태 진* · 류 희 정* · 박 정 응** · 김 창 완***

I. Introduction

The peak flow rate and the total runoff volume have been generally increased in a growing urban watershed due to the change of the hydrologic effects. The determination of probable rainfall intensity in an ungauged project area is crucial because the sizes of hydraulic structures are mainly affected by the design rainfall in the area. The types of the Talbot, the Sherman, and the Japanese have been generally adopted to estimate the design rainfall intensity. Lee and Park (1992) used a unified formula for Seoul district in Korea. We consider 5 types which are the Talbot, the Sherman, the Japanese, Unified-1, and Unified-2 type. Wenzel (1982) pointed out that there was no theoretical basis for these formulas and that the constants of formulas had been determined by the curve fitting procedures. The least square method has been generally used to determine the constants of the formulas.

Rousseeuw and Leory (1987) proposed the algorithm for solving the least median squares (LMS) method which does not overestimate extreme data. The value of the constants in the rainfall intensity formulas can be estimated by the least squares (LS) method, the least median squares (LMS) method, the reweighted linear squares method based on the LMS (RLS), and two linear programming models representing the constrained regression (CR) model. The purpose of this study is to estimate the reliable rainfall intensity formulas at Incheon area and to examine the applicability of the various methods.

* 안성산업대학교, 토목공학과, 교수
** 서울산업대학교, 토목공학과, 교수
*** 한국건설기술연구원, 수자원연구실, 선임연구원

2. Estimation of Probable Rainfall

In urban hydrology, time of concentration at a watershed would be within few hours. Thus, the annual maximum rainfall data from 1952 to 1996 for 11 durations such as 10, 20, 30, 40, 50, 60, 90, 120, 180, 240, and 360 minutes were provided from the Incheon meteorological station. The periods of record are 45 years. The normal, the 2-parameter lognormal, the 3-parameter lognomal, the type I extreme, the Pearson type III, the log-Pearson type III, the Gumbel-Chow, and the Iwai distribution are considered to estimate the rainfall for return periods at Incheon station. The adaptability of the parameters estimated by each distribution are reviewed. The Kolmogorov - Smirnov (K-S) test and Chi-square test (χ^2 test) are then performed to test the goodness of fit for each distribution. Table 1 shows the probable rainfall depth for each duration and the selected distribution. The Pearson type III distribution is selected for the duration of 10, 20, 30, 40, 60, 90, and 120 minutes; the Gumbel - Chow for 180 minutes; the Iwai distribution for duration of 50 and 240 minutes; the log-Pearson type III for 360 minutes.

Table 1. Probable rainfall for each duration

Return Period in years	Rainfall for Duration (mm)										
	10 min	20 min	30 min	40 min	50 min	60 min	90 min	120 min	180 min	240 min	360 min
5	17.6	27.9	36.0	42.7	47.3	53.5	66.6	77.8	92.2	102.8	118.5
10	19.9	32.5	42.1	50.3	56.3	64.2	79.6	94.1	109.5	122.4	144.0
20	21.9	36.7	47.8	57.6	65.4	74.5	91.9	109.7	126.1	141.4	170.8
30	22.9	39.1	51.0	61.7	70.9	80.4	99.0	118.6	135.6	152.5	187.2
50	24.3	41.9	54.9	66.7	77.8	87.9	107.8	129.8	147.6	166.6	208.8
80	25.4	44.6	58.5	71.3	84.4	94.7	115.7	139.9	158.5	179.6	229.8
100	25.9	45.8	60.2	73.4	87.6	97.9	119.5	144.7	163.7	185.8	240.2
150	26.9	47.9	63.2	77.3	93.5	103.7	126.3	153.4	173.0	197.3	259.7
200	27.6	49.5	65.3	80.0	97.8	107.9	131.0	159.5	179.7	205.5	274.1
selected distribu-tion	Pearson type III	Pearson type III	Pearson type III	Pearson type III	Iwai	Pearson type III	Pearson type III	Pearson type III	Gumbel-Chow	Iwai	log-Pearson type III

3. Adopted Regression Methods

The rainfall intensity herein is the average intensity over the rainfall duration. The relationship between rainfall intensity and duration is dependent on the properties of rainfall events at a rainfall station, and the following five types have been widely adopted:

Talbot type: $I=\frac{a}{t+b}$ (1)

Sherman type: $I=\frac{a}{t^n}$ (2)

Japanese type: $I=\frac{a}{t^{1/2}+b}$ (3)

Unified-1 type: $I=\frac{a}{t^n+b}$ (4)

Unified-2 type: $I=\frac{a}{(t+b)^n}$ (5)

where, I is rainfall intensity(mm/hr); t is duration(min.); a, b, and n are the constants of the formulas.

The least squares (LS) method, the least median of squares (LMS) method, the reweighted least squares based on the LMS (RLS), and the constrained regression (CR) models based on the linear programming are employed to establish formulas for describing relationship between rainfall intensity and duration of Incheon station. The least squares regression method is to determine the coefficients of a regression equation by minimizing the sum of the squared residuals and is described as following:

$$\underset{b_o,\, b_1}{\min imize} \sum_{i=1}^{n} e_i^2 \tag{6}$$

where: e_i is the residual value that is the difference between the observed value y_i^{obs} and the estimated value y_i^{comp}. The estimated value y_i^{comp} is given the following equation: $y_i^{comp}=b_o+b_1 x_i$, where b_o and b_1 are the estimated parameters; x_i is the independent variable.

The LMS method minimizes the median value of the squared residuals as in equation (7) (Rousseeuw, 1984). It is known that the LMS method is a robust method which does not overestimate extreme data.

$$\underset{b_o,\, b_1}{\min imize} \ \underset{i=1,\dots,n}{median} \ e_i^2 \tag{7}$$

The RLS method contains a weight w_i in the LS method, which is a function of the standardized LMS residuals in absolute values (Rousseeuw and Leroy, 1987). The RLS method is given in equation (8):

$$\min_{b_o,\, b_1} imize \sum_{i=1}^{n} w_i e_i^2 \tag{8}$$

where, the value of weight is a binary and is dependent on the absolute value of LMS residuals.

The linear programming (LP) is that the objective function is maximized or minimized subject to constraint equations. The CR models consider the following two criterion. The model 1 minimizes the summation of the absolute residual, whereas the model 2 minimizes the maximum absolute residual.

Model 1

$$\text{min. } z = \sum_{I=1}^{n} | e_i |$$

$$\text{s. t. } y_i^{comp} - e_i = y_i^{obs}$$

$$e_i \geq 0 \qquad i = 1, \ldots, n$$

Model 2

$$\text{min. } [\max_i | e_i |]$$

$$\text{S. T } y_i^{comp} + e_i^+ - e_i^- = y_i^{obs}$$

$$e_i^+, e_i^- \geq 0 \qquad i = 1, \ldots, n$$

$$e_i^+, e_i^- \geq 0 \qquad i = 1, \ldots, n$$

4. Application of regression methods

The classical rainfall intensity - duration formulas, the Talbot type, the Sherman type, and the Japanese type, for Incheon station are established using methods of the LS, the LMS, and RLS. The LS method was coded in the Fortran language. The LMS and the RLS method were solved using the package PROGRESS (Rousseeuw and Leroy, 1987). The root mean squared errors (RMS) are then computed to estimate the various types and methods. The Talbot type tends to yield smaller errors for longer return periods, whereas the Sherman type and the Japanese type tend to yield larger errors for longer return periods. It is known that the Japanese type is relatively reliable upon the Incheon station.

The coefficients of the Unified-1 type for the Incheon station are also determined by the LS, the LMS, the RLS, the CR models. The CR models were solved using the linear programming package LINDO. The trial and error method with the LS is employed to determine the coefficients of the Unified-1 type. Table 2 and Table 3 show the Unified-1 and Unified-2 equations for return periods and the root mean squared errors. In the Unified-1 type and the Unified-2 type, the RMS errors are significantly reduced compared to the three classical methods. It is known that all methods for the Unified-1 type or the Unified-2 type are reliable upon the estimation of rainfall intensity - duration formula.

5. summary

It is found that the LMS method and the RLS method are the superior reliability to the LS method in the classical rainfall intensity - duration types; the Talbot, the Sherman, and the Japanese. The Japanese type in the classical methods is suitable for the Incheon station. However, the Unified-1 or the Unified-2 type seem to be the most reliable formulae because a number of potential curves generated by the coefficients of the Unified-1 or the Unified-2 type are considered to determine the best coefficients of the unified type. The LS method with trial and error is relatively the best approach to the Unified-1 or Unified-2 type analysis but the other can also be an alternative method.

References

1. Lee, W. H. and Park, S. D. (1992). "A Unification of the probable rainfall Intensity formula at Seoul", J. of KSCE, Vol. 12, No. 4., pp. 135-143.
2. Kite, G. W., Frequency and Risk Analyses in Hydrology, WRP, 1980.
3. Ponce, V. M., Engineering Hydrology, Principle and Practices, Prentice Hall, New Jersey, 1989.
4. Rousseeuw, P. (1984). "Least median of squares regression," J. of Am. Stat. Assoc. Vol. 79, pp. 871-880.
5. Rousseeuw, P. and Leroy, A. (1987). "Robust Regression and Outlier Detection", Wiley, New York.
6. Wenzel, H. G. (1982). "Rainfall for Urban Stormwater Design, Urban Stormwater Hydrology edited by D. F. Kiber", Water Resource Monograph 7, American Geophysical Union, Washington, D. C..

Table 2. Probable rainfall intensity formulas and RMS errors (Unified-1 type)

( ): RMS error in *mm/hr*

Return Period in years	Rainfall Intensity Formulas				
	LS	LMS	RLS	LP-M1	LP-M3
10	$\frac{1831.7}{t^{.711}+10.26}$ (0.86)	$\frac{1818.2}{t^{.711}+10.24}$ (0.86)	$\frac{1831.7}{t^{.711}+10.26}$ (0.86)	$\frac{1821.5}{t^{.711}+10.20}$ (0.84)	$\frac{1851.8}{t^{.711}+10.39}$ (1.01)
20	$\frac{2380.9}{t^{.726}+12.88}$ (0.88)	$\frac{2325.6}{t^{.726}+12.43}$ (1.13)	$\frac{2325.6}{t^{.726}+12.47}$ (1.17)	$\frac{2409.6}{t^{.726}+13.10}$ (0.88)	$\frac{2403.8}{t^{.726}+13.01}$ (0.88)
50	$\frac{3225.8}{t^{.739}+16.61}$ (1.23)	$\frac{3030.3}{t^{.739}+15.09}$ (1.79)	$\frac{3225.8}{t^{.739}+15.84}$ (2.98)	$\frac{3184.7}{t^{.739}+16.39}$ (1.22)	$\frac{3174.6}{t^{.739}+16.33}$ (1.26)
100	$\frac{3846.2}{t^{.744}+19.04}$ (1.86)	$\frac{4000.0}{t^{.744}+19.86}$ (2.54)	$\frac{4000.0}{t^{.744}+20.16}$ (2.23)	$\frac{3759.4}{t^{.744}+18.58}$ (2.03)	$\frac{3787.9}{t^{.744}+18.74}$ (1.92)
200	$\frac{4347.8}{t^{.746}+20.69}$ (3.11)	$\frac{4761.9}{t^{.746}+23.43}$ (3.42)	$\frac{4761.9}{t^{.746}+23.38}$ (3.42)	$\frac{4273.5}{t^{.746}+19.94}$ (2.98)	$\frac{4405.3}{t^{.746}+20.98}$ (2.91)

Table 3. Probable rainfall intensity formulas and RMS errors (Unified-2 type)

( ): RMS error in *mm/hr*

Return Period in years	Rainfall Intensity Formulas		
	LS	LMS	RLS
10	$\frac{830.48}{(t+16.43)^{0.5931}}$ (0.97)	$\frac{848.30}{(t+16.43)^{0.6011}}$ (1.21)	$\frac{837.59}{(t+16.43)^{0.5961}}$ (0.99)
20	$\frac{1067.19}{(t+21.43)^{0.6085}}$ (1.05)	$\frac{1047.06}{(t+21.43)^{0.6062}}$ (1.23)	$\frac{1076.99}{(t+21.43)^{0.6115}}$ (1.09)
50	$\frac{1456.60}{(t+28.98)^{0.6185}}$ (1.28)	$\frac{1531.51}{(t+28.98)^{0.6442}}$ (1.76)	$\frac{1549.17}{(t+28.98)^{0.6448}}$ (1.50)
100	$\frac{1791.51}{(t+34.87)^{0.6429}}$ (1.74)	$\frac{1699.5}{(t+34.87)^{0.6270}}$ (2.44)	$\frac{1675.79}{(t+34.87)^{0.6241}}$ (2.42)
200	$\frac{2171.70}{(t+41.11)^{0.6549}}$ (2.51)	$\frac{2416.91}{(t+41.11)^{0.6830}}$ (3.36)	$\frac{2486.1}{(t+41.11)^{0.6852}}$ (2.81)