

Wakeby 분포를 이용한 강우자료의 분리효과에 관한 연구

A STUDY ON SEPARATION EFFECT OF RAINFALL DATA
USING A WAKEBY DISTRIBUTION

김 경 덕* · 허 준 행** · 조 원 철***
Kim Kyungduk · Heo Junhaeng · Cho Woncheol

1. 서 론

수공학의 2대 과업중의 하나인 자료해석에 있어서 최적 분포함수를 선정하는 것은 매우 중요하며, 동시에 실측자료의 분포특성을 정확히 나타내는 분포함수를 결정한다는 것은 매우 어려운 과제이다. 그러므로 선정된 분포함수가 과연 실측자료의 분포특성을 잘 나타내는지 면밀한 검토가 이루어져야 한다. 기존의 연구에서 많이 사용하는 분포형의 매개변수 추정은 모멘트법과 최우도법 등이 있는데, 모멘트법은 정밀도가 떨어지고, 최우도법은 우도함수를 유도하여 미분 방정식의 해를 구하기 힘든 경우가 많이 발생한다. 그리고, 기존의 연구에서는 추정된 매개변수의 적합성을 조사하지 않고 최적 분포형을 선정하여 확률 통계치를 산정하고 있는 실정이다.

본 연구에서는 우리 나라 전국 22개 지점의 7개 지속기간(30분, 60분, 2시간, 3시간, 6시간, 12시간, 24시간) 강우자료에 대해서 Wakeby 분포를 적정분포로 선정하였고, 매개변수의 추정은 Wakeby 분포의 매개변수를 추정하는데 적합하다고 알려진 확률가중 모멘트(probability weighted moments)를 사용하였다. Wakeby 분포와 비교하기 위하여 수문자료에 많이 사용하는 11가지 분포형(γ -2, γ -3, lognormal-2, lognormal-3, Gumbel, log-Gumbel-2, log-Gumbel-3, GEV, log-Pearson type III, Weibull-2, Weibull-3 분포)을 동시에 적용하여 확률가중 모멘트법으로 매개변수를 추정하고 적합성 조건을 확인하였다. 추정된 매개변수를 이용하여 균등난수 모의발생기법으로 10,000번 자료를 발생시켜 분리효과를 검토하였다. 이때, 매개변수 적합성을 고려한 경우와 그렇지 않은 경우에 대하여 각기 분리효과를 살펴보았다.

2. 본 론

2.1 Wakeby 분포와 매개변수 추정방법

Wakeby 분포는 식 (2.1) (Landwehr et al.,1978)과 같은 역함수 형태로 정의된다

$$x = m + a[1 - (1 - F)^b] - c[1 - (1 - F)^{-d}] \quad (2.1)$$

* 연세대학교 토목공학과 박사과정
** 연세대학교 토목공학과 조교수
*** 연세대학교 토목공학과 교수

여기서 F는 누가분포함수(cumulative distribution function : CDF)로서, $F = P[X \leq x]$ 로 표시되며 a, b, c, d, m은 Wakeby 분포의 매개변수이다. 매개변수 b는 분포의 왼쪽 꼬리부분을 제어하고(즉, 극소치 자료), 매개변수 d는 분포의 오른쪽 꼬리부분을 제어한다(예를 들면 극대치 자료). 확률밀도함수는 식 (2.1)을 x에 관하여 미분하여 식 (2.2)와 같다.

$$f(x) = \frac{1}{[ab(1-F)^{b-1} + cd(1-F)^{-d-1}]} \quad (2.2)$$

Wakeby 분포의 확률가중 모멘트(PWM) 추정치(Greenwood et al., 1979)는 식 (2.3)으로 주어진다

$$\begin{aligned} \hat{a} &= \frac{(\hat{b} + 1)(\hat{b} + 2)}{\hat{b}(\hat{b} + \hat{d})} \left\{ \frac{\{1\}}{2 + \hat{b}} - \frac{\{0\}}{1 + \hat{b}} - \hat{m} \right\} \\ \hat{b} &= \max[b_1, b_2] \\ \hat{c} &= \frac{(1 - \hat{d})(2 - \hat{d})}{\hat{d}(\hat{b} + \hat{d})} \left\{ \frac{\{0\}}{1 - \hat{d}} - \frac{\{1\}}{2 - \hat{d}} + \hat{m} \right\} \\ \hat{d} &= \frac{(N_1 + \hat{b}N_2)}{(N_2 + \hat{b}N_3)} \\ \hat{m} &= \frac{[\{3\} - \{2\} - \{1\} + \{0\}]}{4} \end{aligned} \quad (2.3)$$

여기서 $\{k\} = (k+1)(k+1+\hat{b})(k+1-\hat{d}) \hat{B}_k'$, $k = 0, 1, 2, 3$

$$\begin{aligned} b_1 &= \frac{(N_3C_1 - N_1C_3) + H^{1/2}}{2(N_2C_3 - N_3C_2)} \\ b_2 &= \frac{(N_3C_1 - N_1C_3) - H^{1/2}}{2(N_2C_3 - N_3C_2)} \\ H &= [(N_1C_3 - N_3C_1)^2 - 4(N_1C_2 - N_2C_1)(N_2C_3 - N_3C_2)]^{1/2} \end{aligned} \quad (2.4)$$

또한 $\hat{m} \neq 0$ 이면 (5개의 매개변수를 갖는 경우)

$$\begin{aligned} N_{4-j} &= (4)^j \hat{B}_3' - (3)^{1+j} \hat{B}_2' + 3(2)^j \hat{B}_1' - \hat{B}_0', \quad j=1,2,3 \\ C_{4-j} &= (5)^j \hat{B}_4' - 3(4)^j \hat{B}_3' + (3)^{1+j} \hat{B}_2' - (2)^j \hat{B}_1', \quad j=1,2,3 \end{aligned} \quad (2.5)$$

한편 $\hat{m} = 0$ 이면 (4개의 매개변수를 갖는 경우)

$$\begin{aligned} N_{4-j} &= -(3)^j \hat{B}_2' + (2)^{j+1} \hat{B}_1' - \hat{B}_0', \quad j=1,2,3 \\ C_{4-j} &= -(4)^j \hat{B}_3' + 2(3)^j \hat{B}_2' - (2)^j \hat{B}_1', \quad j=1,2,3 \end{aligned} \quad (2.6)$$

Wakeby 분포의 확률가중 모멘트에 의한 추정치는 식 (2.3)에서 식 (2.6)으로 부터 구할 수 있는데, 추정된 값은 매개변수 적합성 조건을 만족해야 한다. 적합성 조건은 표 2.1에 나타나 있다 (Landwehr et al., 1979a와 1979b ; Cunnane, 1989). 더욱이 $b > -1$ 과 $d < 1$ 은 Wakeby 분포의 평균 값이 존재하기 위하여 꼭 만족해야 하는 조건이다. 만약 추정된 매개변수가 적합성 조건에 맞지 않으면, $\hat{b} = \hat{b}_{max} = 50.0$ 으로 가정하고 점차적으로 \hat{b} 값을 줄여 가면서 나머지 매개변수 값을 추정한다. 이때에도 각 단계에서 적합성 조건을 만족하는가 확인해야 하며 이를 만족할 때까지 반복한다. 여기서 b가 -1.0 일 때까지 적합성 조건을 만족하지 못하면 적용한 자료에 대해 Wakeby 분포의 사용은 적합하지 못한 것이다(Landwehr et al., 1979a , 1979b).

표 2.1 매개변수 적합성 조건 (Wakeby분포)

형태	적합성 조건
I	$b + d > 0$ 또는 $b = cd = d = 0$
II	$ab = 0$ 이면 $b = 0$
III	$cd = 0$ 이면 $d = 0$
IV	$cd \geq 0$
V	$ab + cd \geq 0$

2.2 적용 및 결과 분석

본 연구에서는 전국 22개 지점의 강우자료를 사용하였으며, 지속기간은 30분, 60분, 2시간, 3시간, 6시간, 12시간, 24시간 등 7개 지속기간을 선택하였다. 표 2.2는 22개 지점과 강우기록 연수를 나타내고 있다.

표 2.2 대상 강우지점 및 자료년수

지점명	지점번호	기록년수	지점명	지점번호	기록년수
속초	90	24	포항	138	38
춘천	101	26	군산	140	24
강릉	105	34	대구	143	76
서울	108	63	전주	146	51
인천	112	40	울산	152	45
원주	114	24	광주	156	53
수원	119	28	부산	159	50
서산	129	24	충무	162	24
청주	131	25	목포	165	69
대전	133	24	여수	168	49
추풍령	135	37	완도	170	21

매개변수가 5개인 경우의 추정값은 표 2.3과 같으며, 그림 2.1과 2.2는 서울지점, 지속기간 24시간 강우자료의 경험 및 적합 확률밀도함수와 누가분포함수를 나타내고 있다. 그림에서 Wakeby 5 매개변수 분포가 실제 강우자료의 극치부분에 매우 적합하게 나타내고 있다.

표 2.3 Wakeby-5 분포의 매개변수 추정값

지점번호	지점명	자료크기	m	a	b	c	d
90	속초	24	38.8466	65.5795	18.5824	400.4721	.1409
101	춘천	22	-2.6786	89.0076	49.9000	-2761.6660	-.0236
105	강릉	34	-28.7976	112.0180	49.9000	-716.3234	-.1280
108	서울	63	51.7081	42.7420	5.5903	701.5034	.0807
112	인천	40	-162.9351	241.6904	138.9487	316.8810	.1732
114	원주	20	-25.9434	109.0423	29.5315	977.0676	.0794
119	수원	28	80.7807	6.1991	3.2360	525.8530	.1404
129	서산	24	46.0744	93.5112	1.5254	622.6848	.0519
131	충대	25	33.5242	46.3025	41.5580	-813.9635	-.0631
133	대전	23	61.3348	55.0066	1.9039	140.2972	.2296
135	추풍령	37	57.5123	58.6374	2.2002	81.0016	.2135
138	포항	38	44.1084	32.7312	8.2871	425.6466	.1003
140	군산	24	42.1335	40.9787	8.9237	206.6781	.1737
143	대구	76	38.3330	42.7439	2.8753	-764.8528	-.0533
146	전주	51	47.2786	23.8368	11.0575	2947.5790	.0208
152	울산	45	48.9139	80.1123	3.0540	64.0729	.4007
156	광주	49	51.0104	40.8488	5.3509	3395.4020	.0126
159	부산	50	18.6055	48.6925	50.2585	-622.7864	-.1619
162	충무	24	-4.2783	85.0271	49.9000	-2080.8960	-.0380
165	목포	69	42.1517	47.0874	7.0657	246.9992	.1444
168	여수	49	34.1044	50.3896	17.7616	-1493.4130	-.0466
170	완도	21	81.8320	-.5972	49.9000	9197.7110	.0117

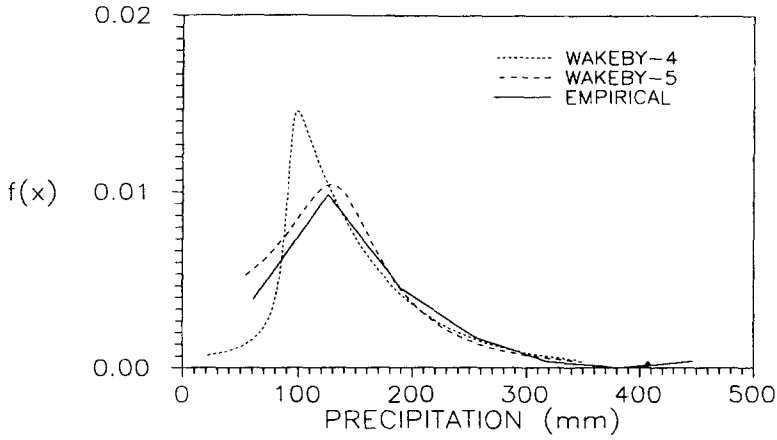


그림 2.1 경험 및 적합Wakeby 분포의 확률밀도함수(지속기간 24시간, 서울지점)

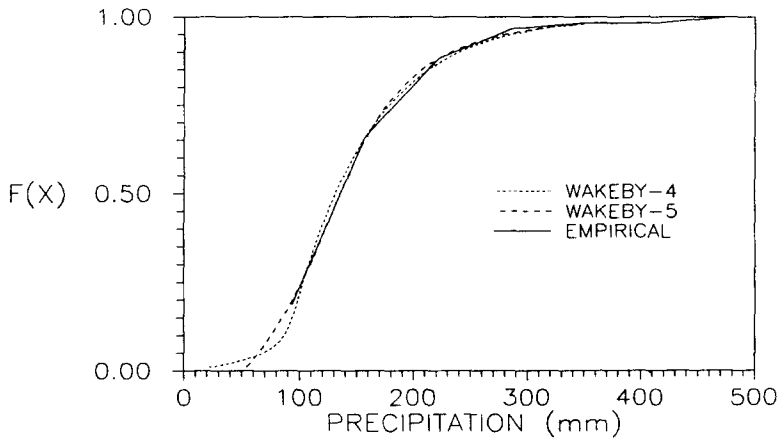


그림 2.2 경험 및 적합Wakeby 분포의 누가분포함수(지속기간 24시간, 서울지점)

적용한 11개 분포형 중에서 지속기간 24시간 강우자료에 대하여 γ -3 분포에 대한 확률가중 모멘트법에 의한 매개변수 값과, 추정된 매개변수의 적합성을 검토하여 적합성 유무를 O.K. 와 N.G.로 표시하였다. 표 2.4는 γ -3 분포의 매개변수 값과 적합성 유무를 나타내고 있다.

표 2.4 Gamma-3 분포의 매개변수 값과 적합성(지속기간 24시간, PWM)

지점번호	지점명	자료크기	Location parameter	X _{MIN}	X _{MAX}	Scale parameter	Shape parameter	Validity
90	속초	24	87.078	78.000	427.000	69.897	1.140	N.G.
101	춘천	22	69.352	68.000	289.000	49.325	1.598	N.G.
105	강릉	34	45.044	48.000	328.000	47.546	2.466	O.K.
108	서 울	63	57.386	61.000	447.000	52.990	1.739	O.K.
112	인 천	40	71.849	57.000	372.000	83.361	.858	N.G.
114	원 주	20	56.560	55.000	395.000	77.753	1.380	N.G.
119	수 원	28	89.224	86.000	460.000	118.005	.697	N.G.
129	서 산	24	-37.408	58.000	253.000	20.494	8.494	O.K.
131	충 주	25	68.501	75.000	240.000	35.582	1.646	O.K.
133	대 전	23	62.432	69.000	334.000	50.691	1.515	O.K.
135	추 풍	37	24.814	66.000	216.000	15.962	5.951	O.K.
138	포 향	38	56.707	60.000	316.000	44.225	1.448	O.K.
140	군 산	24	62.247	61.000	265.000	44.799	1.343	N.G.
143	대 구	76	3.628	42.000	248.000	19.984	5.260	O.K.
146	전 주	51	57.763	46.000	364.000	56.353	1.313	N.G.
152	울 산	45	60.181	53.000	458.000	59.658	1.541	N.G.
156	광 주	49	41.310	56.000	261.000	26.559	3.287	O.K.
159	부 산	50	44.476	37.000	439.000	54.974	1.977	N.G.
162	충 북	24	62.800	71.000	325.000	59.513	1.554	O.K.
165	목 포	69	52.373	43.000	404.000	35.151	2.069	N.G.
168	여 수	49	55.401	53.000	331.000	46.099	2.017	N.G.
170	완 도	21	84.865	68.000	460.000	116.421	.907	N.G.

2.3 강우자료의 분리효과

22개 지점, 7개 지속기간의 강우자료를 대상 분포형에 적용하여 매개변수를 추정한 후, 균등난수에 의한 모의실험을 10,000번 실시하여 분리효과를 확인하였다. 이때 X축은 왜곡도 계수의 평균, Y축은 왜곡도 계수의 표준편차로 하여 도시한다. 그림 2.3은 매개변수 적합성 조건을 만족시키지 못하는 지점을 제외하여 모의 발생시켜 분리효과를 나타낸 것이고, 그림 2.4는 매개변수의 적합성 조건과 관계없이 분리효과를 나타낸 그림이다. 그림 2.4와 같이 매개변수 적합성 조건에 관계없이 모의발생한 경우 3변수 Weibull 분포는 큰 분리효과를 나타내고 있다. 그림 2.3에서 살펴보면, GEV 분포가 가장 분리효과가 적은 것으로 나타나고 있으며, Wakeby (4, 5매개변수) 분포형을 포함한 다른 분포형도 비교적 비슷한 분리효과를 나타냄을 볼 수 있다.

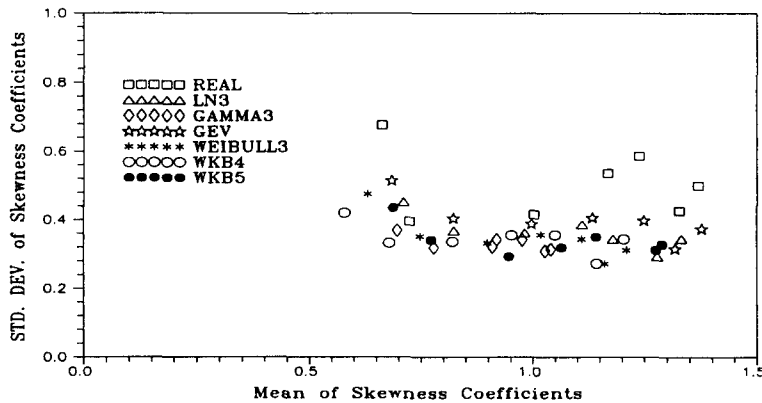


그림 2.3 강우자료의 분리효과 (I, PWM)

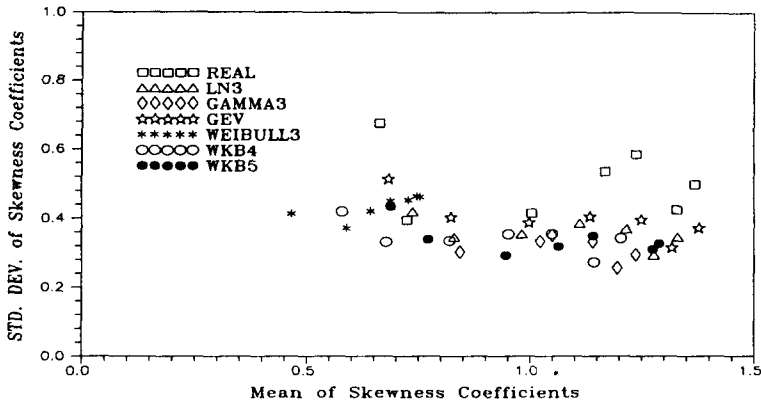


그림 2.4 강우자료의 분리효과 (II, PWM)

3. 결 론

전국 22개 지점, 7개 지속기간의 강우자료에 대한 분리효과를 검증하기 위하여 Wakeby 분포를 선정하고, 확률가중 모멘트법에 기초하여 매개변수를 추정한 후, 균등난수로 10,000번 모의발생하여 분리효과를 살펴본 결과 다음과 같은 결론을 얻을 수 있었다.

- (1) 매개변수 적합성 조건에 관계없이 모의발생하여 분리효과를 살펴보면, 3변수 Weibull 분포의 경우 실제자료에 적합하지 않으며, 매개변수 적합성 조건에 부적당한 지점을 제외하여 분리효과를 고려할 경우 실제자료와 근접한 성질을 나타냄을 알 수 있었다.
- (2) GEV 분포형이 분리효과가 가장 적게 나타났으며, Wakeby 분포도 비교적 근접하게 나타났다.

홍수 극치자료에 대하여 분리효과를 잘 나타낸다고 알려져 있는 Wakeby 분포를 우리나라의 강우자료에 적용해 본 결과 Wakeby 분포보다는 GEV 분포가 분리효과를 더 잘 나타냈으며, 분포합수를 적용할 때 매개변수 적합성 조건을 필수적으로 확인하여야 할 것으로 판단된다.

4. 참고문헌

1. 이원환, "한국 대표 확률 강우 강도식의 유도", 1993.
2. 허준행외 1인, "강우자료의 분리효과", 한국 수문학회지, 제 26권 4호, pp. 73~83, 1993.
3. Cunnane, C., "Statistical Distributions for Flood Frequency Analysis", WMO Operational Hydrology Report, No. 33, Geneva, Switzerland, 1989.
4. Greenwood, J. A., Landwehr, J. M., Matalas, N. C. and Wallis, J.R., "Probability Weighted Moments and Relation to Parameters of Several Distributions Expressible in Inverse Form", WRR, Vol. 15, No. 5, pp. 1049~1054, 1979.
5. Houghton, J. C., "Birth of a Parent: the Wakeby Distribution for Modeling Flood Flows", WRR, Vol. 14, No. 6, pp. 1105~1109, 1979.
6. Landwehr J. M., Matalas N. C. and Wallis J. R., "Some comparisons of flood statistics in real and log space", WRR, 14(3), pp. 902~920, 1978.
7. Landwehr J. M., Matalas N. C. and Wallis J. R., "Estimation of parameters and quantiles of Wakeby distributions 1. Known lower bounds", WRR, 15(6), pp. 1361~1372, 1979a.
8. Landwehr J. M., Matalas N. C. and Wallis J. R., "Estimation of parameters and quantiles of Wakeby distributions 2. Unknown lower bounds", WRR, 15(6), pp. 1373~1379, 1979b.