

A Study of the Rainfall in Korea Using Outlier Detections

Ho Geun Son[#], Jae Joon Lee⁺

Kumoh National Institute of Technology, 61 Daehak-ro, Gumi, Gyeongbuk, Korea

Abstract

The outliers in annual maximum rainfall can play an important role in the frequency analysis that estimates the probability rainfall. To ensure reliability of the frequency analysis, this study detects the outliers and conducts a qualitative analysis to investigate the causes of rainfall based on selected outliers. The outliers in annual maximum 24-hr rainfalls at 57 stations in Korea were analyzed by four different methods, such as Grubbs and Beck method, IACWD method, modified z-score method and box-plot method. The causes of the selected rainfall events as high outliers were classified into changma, typhoon, severe rain storm and frontal rainfall. The comparison of the methods shows that less outliers were found with Grubbs and Beck method than the other methods. It is found that 51.0% of the outliers occurred due to typhoon, followed by severe rain storm (35.7%), frontal storm (7.2%) and changma (6.1%). Rainfall events as outliers were mostly caused by severe rain storm in inland regions while they were caused by typhoon in coastal regions. The probability rainfall calculated without outliers from the test was compared with that before the test.

Key words: outlier detection, cause of rainfall occurrence, probability rainfall

1. 서론

매년 여름철에 발생하는 호우는 전국에 많은 피해를 발생시키고 있으며, 이상기후 현상의 영향으로 최근에는 과거보다 때 이른 폭염 및 태풍이 내습하고 있다. 발생하는 기상상황 역시 양적인 부분 및 지속기간에서 예전의 상황과는 전혀 다르게 증가하는 양상이 발생하고 있다. 과거의 호우피해가 넓은 지역에 어느 정도 예측가능하게 발생하였다면, 최근의 호우피해는

좁은 지역에 순간적인 집중호우로 인하여 예측이 어려운 상황으로 진행되고 있으며, 피해규모는 과거보다 증가하고 있다. 이와 같이 과거의 호우상황과 다르게 변화하는 현상은 체계적으로 준비되지 않은 사회 전반에 많은 피해를 발생시키고 있으며, 재해를 예측 및 저감하는 계획수립에 혼란을 발생시키고 있다. 현재 실무에서 적용하고 있는 빈도해석에서는 극치계열의 연최대치 강우자료를 활용하며, 수집된 자료에는 장비활용에 따른 기계적 오류 및 기술자가 활용하는

[#] The 1st author: Ho Geun Son, Tel. +82-54-339-5032, Fax. +82-54-339-5080, E-mail. hgson@ekr.or.kr

⁺ Corresponding author: Jae Joon Lee, Tel. +82-54-478-7616, Fax. +82-54-478-7859, E-mail. jhb365@kumoh.ac.kr

과정에서 발생하는 인적오류, 또는 이상기후에 따라 본래의 값에서 벗어나는 이상치의 값 등이 존재할 수 있다. 전술한 값들은 분석과정에서 자료 본연의 특성을 왜곡시켜 편향된 확률강우량을 산정할 수 있으며, 발생오류 중에서 기계 및 인적오류는 장비의 현대와 전산화 처리과정 등으로 해소할 수 있지만, 자료 본연에서의 이상치에 대한 처리방법은 실무에서 시행되고 있지 않은 실정이다. 또한, 이상치가 포함된 강우자료의 빈도해석을 통하여 산정된 설계홍수량을 실무에서 적용하는 경우, 이수분야에서는 고이상치로 인한 계획홍수의 과다산정이 발생하여 실제 현황과는 차이가 발생하는 경우가 나타나고 있다. 설계홍수량 산정요령(Ministry of Land, Transport and Maritime Affairs, 2012)의 경우, 이상치검정을 수행하고 결과에 따라 강우자료에서 제외하는 것으로 제시되어 있으나, 실무에서는 관측년수 부족으로 수행되지 않거나 단순 참고자료로 활용되고 있는 실정이다.

자료분석에 있어 이상치와 관련된 연구동향을 살펴보면, 이상치의 개념은 과거로부터 여러 사람들에 의해 정의되거나 적용되어 왔다. Grubbs(1969)는 표본자료에서 크게 어긋나 다른 자료들로부터 떨어져 표시되는 자료라 정의하였고, Hawkins(1980)는 기존의 데이터들과 성격이 판이하게 달라 기존의 데이터들과는 다른 분포에서 생성된 확률이 높은 데이터라 정의하였다. 또한 Johnson(1992)은 자료들로부터 분리되어 모순되는 것처럼 보이는 값이라고 하였으며, Mendenhall, *et. al.*(1993)은 분포의 중앙에서 특히 떨어져있는 값이라 정의하였다. Harmeling, *et. al.*(2006)은 실제로 발생할 확률이 매우 낮은 데이터로 이상치를 정의하기도 하였다. 이상치검정에 대하여 Thompson(1935)은 표본자료 표준편차와 관측치 편차 간의 비를 바탕으로 한 검정방법을 제안하였다. 또한 Dixon(1950-1951)은 최소 관측치와 비교하여 최대 관측치, 2순위 최대 관측치를 사용하는 방식의 검정 방법을 제안하였다. Barnett & Lewis(1994)는 Dixon이 제안한 검정 방식을 통하여 2순위 최대 관측치에 대하여 연구하였다. Grubbs(1969), Stefansky(1972)는 정규분

포를 따르는 자료에서의 이상치 검정을 제안하였으며, Grubbs & Beck(1972)은 평균과 표준편차, 그리고 유의수준에 대하여 계수 K_n 을 이용한 검정방법을 제안하였다. Pilon, *et. al.*(1985)은 K_n 을 유의수준 10%에 대한 일반식으로 개발하였으며, 미 수자원 평의회(The U.S. Water Resources Council, 1981)는 Grubbs & Beck이 제안한 검정방법에 왜곡도에 대한 검토를 추가하여 검정방법을 제안하였다. McCormick & Rao(1995)는 이 검정방법을 매우 높은 한계점을 가지는 방법이라 평하였으며, 독립변수와 예측치의 편차를 이용하여 최소평균제곱방법(Least Mean Square method, LMS)을 제시하여 기존의 방법들과 비교하였다.

강우의 이상치검정에 관한 연구로는 Choi, *et. al.*(2011)이 국내 강우의 지역빈도해석을 위하여 강우자료에 대한 자료 이상치 보정 후 지속기간별 연 최대강우량 자료를 추출하여 신뢰성 있는 자료를 구축하였다. Moon, *et. al.*(2014)은 AWS 강우자료의 품질개선을 위하여 서울지역의 강우자료를 대상으로 결측치 및 이상치 처리에 대한 적용성을 검토하여 기존 강우자료와의 비교분석을 통해 신뢰도를 검증하였다. 또한, Kim, *et. al.*(2015)이 분 단위 강우자료의 품질개선을 위하여 서울지역 강우자료를 대상으로 강우보정에 따른 지점별 비교를 수행하였으며, Jeong(2016)는 57개 지점의 국내 강우자료를 대상으로 2가지 이상치방법을 적용하여 그 결과에 따른 확률강우량을 비교·산정하였다. 또한 Lee, *et. al.*(2017)은 전국 주요 도시에 대한 이상치 강우의 발생빈도 및 변화를 분석하였다. 호우원인별 강우분석 연구로는 Song(2012)이 국내 확률홍수량의 재평가를 위하여 극한강우를 원인별로 태풍과 집중호우로 분류하여 분석하였으며, Lee, *et. al.*(2018)이 국내 강우를 대상으로 이상치를 고려한 극치강우의 발생여부를 분석하였다.

본 연구에서는 30년 이상의 기록연수를 갖는 국내 57개 지점들을 대상으로, 지속기간 24시의 연 최대치 강우자료들을 수집하여 국내 강우의 이상치 발생여부를 조사하고, 이상치 자료의 기각에 따른 확률강우량의 변화를 분석하고자 한다. 이상치 분석방법으로는

국내외에서 널리 적용되는 Grubbs and Beck 방법, IACWD 방법, 수정 z-score 방법과 Box-plot 방법을 적용하며, 확률강우량 분석은 설계홍수량 산정요령 (Ministry of Land, Transport and Maritime Affairs, 2012) 을 참고하여 Gumbel 분포형을 이용한다. 또한, 이상치가 발생하는 경우는 일반적인 기상상황이 아닌 극치 상황과 밀접한 관련이 있으므로 이상치가 발생한 강우의 원인을 장마, 태풍, 집중호우, 전선형 강우로 분류하여 국내 지점강우의 이상치 발생특성을 파악하고자 한다. 이상치 강우의 분석과 확률강우량의 변화, 지점별 이상치 강우의 발생원인 검토는 예측하기 어려운 이상기후의 영향을 고려하여 장기적으로 지점별 특화된 사상들에 대한 정도 높은 강우예측과 이수계획에 따른 수자원 확보, 치수계획에 따른 적절한 저감 계획 수립을 유도하여, 향후 국내 지역에 합리적인 수자원 활용기준을 제시할 것으로 판단된다.

II. 기본이론과 대상지점

1. 이상치 검정방법

1) Grubbs and Beck 방법

Grubbs and Beck 방법은 다양한 수문자료의 이상치 검정에 널리 쓰이고 있는 방법이며, 이 방법은 표본자료가 정규분포를 이룬다는 가정 하에 평균과 표준편차를 이용하여 상한치와 하한치를 계산하여 이상치를 판단하는 방법이다(Jung & Yoon, 2003). Grubbs and Beck 방법은 상한과 하한 임계값을 다음과 같이 정의한다.

$$X_H = \bar{x} + k_N S_x \quad (1)$$

$$X_L = \bar{x} - k_N S_x \quad (2)$$

여기서, \bar{x} 와 S_x 는 표본자료의 평균과 표준편차이며 임계값 k_N 은 표본크기와 유의수준에 따라 Grubbs and Beck에 의해 주어진다. Pilon, et. al.(1985)이 제안한 10%의 유의수준의 근사치는 식 (3)과 같다.

$$k_N = -3.62201 + 6.28446N^{1/4} - 2.49835N^{1/2} + 0.491436N^{3/4} - 0.037911N \quad (3)$$

여기서, N 은 표본의 수이며, 자료의 값이 X_H 보다 클 경우 이를 고이상치, X_L 보다 작을 경우 저이상치라고 한다.

2) IACWD 방법

IACWD(Interagency Advisory Committee on Water Data) 방법은 표본자료의 왜곡도계수를 사용하여 이상치 검정을 수행하며, 왜곡도계수가 0.4보다 크면 고이상치에 대한 검정, -0.4보다 작으면 저이상치에 대한 검정을 실시한다. 또 왜곡도계수가 -0.4~0.4 사이에 존재한다면 고이상치와 저이상치에 대한 검정을 모두 실시하는 방법이다. 고이상치와 저이상치는 식 (4)와 식 (5)와 같다.

$$y_H = \bar{y} + K_N S_Y \quad (4)$$

$$y_L = \bar{y} - K_N S_Y \quad (5)$$

여기서, y_H 와 y_L 은 각각 고이상치와 저이상치로 로 그단위이며, \bar{y} 는 평균, S_Y 는 표준편차, K_N 은 표본의 수에 따른 빈도계수이다.

3) 수정 z-score 방법

수정 z-score 방법은 표준화된 z값을 계산하는 방법으로 z값은 다음 식으로 산정된다.

$$z_i = \frac{(x_i - \bar{x})}{S_x} \quad (6)$$

여기서, \bar{x} 와 S_x 는 표본의 평균과 표준편차이며, 시계열 자료에서의 이상치는 $|z_i|$ 가 2.5를 초과하는 자료를 의미한다. 수정 z-score 방법(Iglewicz & Hoaglin, 1993)은 표본자료의 평균과 표준편차로 표준화된 z값을 이용하여 상·하 한계선을 초과하는 값을 확인하는 z-score 방법 중 향상된 중앙값 절대편차를 이용하여 이상치 여부를 판단하는 방법이다.

$$z_i = 0.675(x_i - x_{0.5}) / MAD \quad (7)$$

여기서, MAD 는 중앙값 절대편차이며 다음 식으로 산정된다.

$$MAD = \frac{1}{N} \sum_{i=1}^N |x_i - x_{0.5}| \quad (8)$$

여기서, $x_{0.5}$ 는 중앙값이며, $|z_i|$ 가 3.5를 초과하는 자료를 이상치로 간주한다.

4) Box-plot 방법

Box-plot 방법은 자료분석의 초기 단계에서 수행하는 탐색적 데이터분석에서 분포의 개괄적 특성을 파악하기 위하여 자주 사용되며, 자료의 크기보다는 순위(rank) 혹은 분위수(quantile)에 근거한다(Choi & Park, 2016). 전체 자료를 25%씩 사분위로 구분하는 방법으로 정렬된 자료계열을 중앙값, 박스, 외곽 값으로 구분한다. 정렬된 25 ~ 75% 값들을 박스로 포함하며, 사분위수의 도식화로 자료의 분포를 파악하므로서 이상치들의 위치와 자료의 비대칭 여부를 쉽게 검토할 수 있다. 박스의 길이를 사분위수범위(IQR)이라 하며, 사분위수범위는 다음 식과 같다.

$$IQR = Q_3 - Q_1 \quad (9)$$

여기서, Q_1 는 일사분위수, Q_3 는 삼사분위수이며, 상한극한값(Upper fence, UF)과 하한극한값(Lower fence, LF)은 다음과 같다.

$$UF = Q_3 + (1.5 \times IQR) \quad (10)$$

$$LF = Q_1 - (1.5 \times IQR) \quad (11)$$

2. 강우 발생원인

본 연구에서 적용한 강우 발생원인은 크게 장마, 태풍, 집중호우와 전선형 강우로 분류하였으며, 각각의 기상현상에 대한 정의는 기상청의 자료를 활용하였

다. 발생 원인으로 살펴보면 장마(changma)는 동아시아의 몬순과 연관되어 시작하는 것으로 중국, 일본 등지에서 유사하게 발생하는 비가 계속해서 많이 오는 계절현상을 뜻하며, 우리나라에서는 일반적으로 6월 하순부터 7월 중, 하순까지의 기간이 해당된다. 태풍(typhoon)은 북태평양 열대지방에서 발생하여 필리핀과 중국, 일본 및 우리나라 일대를 내습하는 열대성 저기압을 말한다. 국내에 내습하는 태풍은 주로 위도 5° ~ 25°, 수온이 약 27°C 이상인 필리핀 동부의 열대해상에서 발생하여 처음에는 북서쪽으로 이동하고 동중국해에 이르면 북동쪽으로 방향을 바꾸어 포물선을 그리면서 우리나라에는 주로 7월부터 9월에 걸쳐 내습한다. 집중호우(severe rain storm)는 짧은 시간 동안에 좁은 지역에서 많은 양의 비가 내리는 현상을 말하며, 일반적으로 1시간에 30mm이상이거나 하루에 80mm 이상의 비가 내릴 때, 또는 연강수량의 10% 정도의 비가 하루 동안에 내릴 때를 말한다. 집중호우의 지속기간은 수십분에서 수시간 정도이며, 보통 반경 약 10 ~ 20km 정도의 비교적 좁은 지역에 집중적으로 내리고, 때로는 천둥번개를 동반하기도 한다. 전선형 강우(frontal rainfall)는 물리적인 특성(온도, 수증기, 바람, 압력)이 수평면에서 대략적으로 일정한 대규모의 공기집단이라 할 수 있다. 공기기단은 북극 및 남극지역, 아열대 해양지역 그리고 건조한 아열대 육지에 분포하고 이 지역들로부터 이동하고 있으며, 만일 공기기단이 움직이는 과정 중에 주위의 공기보다 차가워지게 되면 강우가 발생하게 된다. 공기기단간의 경계를 전선이라 하며, 이는 공기의 이동에 의해 분류할 수 있다. 찬 공기가 더운 공기로 대치되면 한랭전선이 형성되고, 따뜻한 공기가 찬 공기를 밀고 오면 온난전선이 형성된다. 두 경우 모두 따뜻한 공기가 상승하며 냉각되어 강우가 발생하게 된다.

3. 확률강우량 산정방법

본 연구에서 적용한 지점별 강우분석에 따른 확률강우량 산정방법은 설계홍수량 산정요령(Ministry of

Land, Transport and Maritime Affairs, 2012)을 준수하여 확률분포형으로는 Gumbel 분포형, 모수추정방법으로는 확률가중모멘트법을 적용하였다.

1) Gumbel 분포형

Gumbel 분포형은 극치자료의 빈도해석에 많이 사용되는 분포로서 Type 1분포라고도 알려져 있다. 연최대 홍수량 또는 연최대 강우량 등의 빈도 분석에 이 분포가 사용된다. Type 1분포의 확률밀도함수는 식 (12)와 같다.

$$f(x) = \frac{1}{\alpha} \exp \left[-\frac{x-x_0}{\alpha} - \exp \left(-\frac{x-x_0}{\alpha} \right) \right] \quad (12)$$

여기서, α 는 양수인 축척 모수이며 x_0 는 위치모수로 최대 확률의 발생위치(mode)를 표시하며, 범위는 $-\infty < x < \infty$ 이다.

식 (12)를 적분하면 식 (13)과 같은 누가 분포 함수를 얻을 수 있다.

$$F(x) = \exp \left[-\exp \left(-\frac{(x-x_0)}{\alpha} \right) \right] \quad (13)$$

2) 확률가중모멘트법

모수추정방법은 확률가중모멘트법을 적용하며, 일반식은 식 (14)와 같다.

$$M_{p,r,s} = E [X^p F^r(x) \{1-F(x)\}^s] \quad (14)$$

여기서, p, q, s 는 정수이고, 일반적으로 사용하는 모집단의 확률가중모멘트는 식 (15)와 식 (16)으로 나타내며, 각각에 대한 표본자료의 불편 확률가중모멘트는 식 (17)과 식 (18)과 같이 표시할 수 있다.

$$M_{1,r,0} = E [X F^r(x)] \equiv B_r \quad (15)$$

$$M_{1,0,s} = E [X \{1-F(x)\}^s] \equiv B_s \quad (16)$$

$$\hat{B}_r = \frac{1}{N} \sum_{j=1}^N x_j \frac{(j-1)(j-2) \cdots (j-r)}{(N-1)(N-2) \cdots (N-r)}, r > 1 \quad (17)$$

$$\hat{B}_s' = \frac{1}{N} \sum_{j=1}^N x_j \frac{(N-j)! (N-s-1)!}{(N-j-s)! (N-1)!}, s \geq 1 \quad (18)$$

Table 1. Observation periods of 24hr rainfall for the 57 stations

No.	Station	Years	No.	Station	Years	No.	Station	Years
1	Sokcho	1971~2017 (47)	21	Busan	1961~2017 (57)	41	Jeongeup	1973~2017 (45)
2	Daegwanlyeong	1972~2017 (46)	22	Tongyeong	1968~2017 (50)	42	Namwon	1973~2017 (45)
3	Chuncheon	1966~2017 (52)	23	Mokpo	1971~2017 (47)	43	Jangheung	1973~2017 (45)
4	Gangneung	1961~2017 (57)	24	Yeosu	1961~2017 (57)	44	Heanam	1973~2017 (45)
5	Seoul	1961~2017 (57)	25	Wando	1973~2017 (45)	45	Goheung	1973~2017 (45)
6	Incheon	1961~2017 (57)	26	Suncheon	1973~2017 (45)	46	Yeongju	1973~2017 (45)
7	Wonju	1973~2017 (45)	27	Jinju	1969~2017 (49)	47	Mungyeong	1973~2017 (45)
8	Suwon	1964~2017 (54)	28	Ganghwa	1973~2017 (45)	48	Yeongdeok	1973~2017 (45)
9	Seosan	1968~2017 (50)	29	Yangpyeong	1973~2017 (45)	49	Uiseong	1973~2017 (45)
10	Uljin	1972~2017 (46)	30	Icheon	1973~2017 (45)	50	Gumi	1973~2017 (45)
11	Cheongju	1967~2017 (51)	31	Inje	1973~2017 (45)	51	Yeongcheon	1973~2017 (45)
12	Daejeon	1969~2017 (49)	32	Hongcheon	1973~2017 (45)	52	Geochang	1973~2017 (45)
13	Chupunglyeong	1961~2017 (57)	33	Jecheon	1973~2017 (45)	53	Hapcheon	1973~2017 (45)
14	Andong	1983~2017 (35)	34	Boeun	1973~2017 (45)	54	Milyang	1973~2017 (45)
15	Pohang	1961~2017 (57)	35	Cheonan	1973~2017 (45)	55	Sancheong	1973~2017 (45)
16	Gunsan	1968~2017 (50)	36	Bolyeong	1973~2017 (45)	56	Geoje	1973~2017 (45)
17	Daegu	1961~2017 (57)	37	Buyeo	1973~2017 (45)	57	Namhae	1973~2017 (45)
18	Jeonju	1961~2017 (57)	38	Geumsan	1973~2017 (45)	X		
19	Ulsan	1961~2017 (57)	39	Buan	1973~2017 (45)			
20	Gwangju	1961~2017 (57)	40	Imsil	1973~2017 (45)			

여기서, $\hat{B}_0 = \hat{B}'_0 = \bar{X}$ 이며, \bar{X} 는 표본자료의 평균이다. 불편 확률가중 모멘트는 선정된 분포형의 형식에 따라 편리한 식을 사용한다.

4. 대상지점

본 연구에서는 기록연수 30년 이상의 전국 57개 기상관측지점을 대상으로, 지속기간 24시간의 연최대치 강우자료를 이용하였다. 섬 지역인 제주도와 울릉도는 기록연수는 충분하나 강우특성이 내륙 및 내륙에 인접한 도서지역과 다른 양상을 띠므로 본 연구에서는 제외하였다. 적용한 강우자료는 기상청의 자료를 이용하였으며, 관측지점과 강우자료의 기록연수를 정리하면 <Table 1>과 같다.

III. 적용 및 분석

1. 이상치 검정

이상치 검정방법으로 Grubbs and Beck 방법, IACWD 방법, 수정 z-score 방법과 Box-plot 방법 등 총 4가지 방법을 대상지점에 적용하여 이상치가 발생하는 연도를 분석하였다. 각 방법에 따라 발생한 지점별 이상치 현황을 나타내면 <Table 2>와 같으며 <Table 2>를 살펴보면, 발생한 이상치 연도는 방법별 중복되는 지점의 연수를 포함하여 합계 190개 자료로 나타났다. 검정 방법별 분포현황을 살펴보면, Box-plot 방법에서 98개 연도, IACWD 방법에서 55개 연도, 수정 z-score 방법에서 22개 연도, Grubbs and Beck 방법에서 15개 연도에서 이상치가 발생하였다. Box-plot 방법과 IACWD 방법에 의한 결과는 중복되어 발생한 연도가 다수 나타났으며, 적용한 4가지 방법 모두에서 이상치가 발생한 지점도 12지점으로 나타났다.

Grubbs and Beck 방법에서는 15개 지점의 강우자료에서 이상치가 발생하였으며, 이중에서 영주, 의성, 영천 등의 3개 지점은 저이상치로 나타났다. IACWD 방법에서는 55개 지점에서 이상치가 발생하였으며, 이중에서 속초, 서울, 수원, 추풍령, 군산, 울산, 광주, 통

영, 강화, 영덕, 의성, 거제지점 등에서는 이상치가 2개년에서 나타났다. 수정 z-score 방법에서는 22개 지점에서 이상치가 발생하였으며, 이중에서 울산, 거제지점에서 이상치가 2개년에서 나타났다. Box-plot 방법에서는 가장 많은 98개 이상치가 발생하였으며, 속초, 서울, 인천, 원주, 수원, 서산, 대전, 추풍령, 포항, 군산, 울산, 광주, 통영, 목포, 완도, 강화, 홍천, 보은, 천안, 보령, 부여, 정읍, 장흥, 고흥, 영주, 문경, 영덕, 의성, 구미, 거제지점에서는 이상치가 2개년 이상으로 나타났다. 따라서 국내 지점강우를 대상으로 적용한 이상치 분석에서 이상치가 가장 많이 발생한 방법은 Box-plot 방법이며, 다음으로 IACWD 방법, 수정 z-score 방법, Grubbs and Beck 방법 순서로 이상치 발생량이 분석되었다.

2. 이상치검정 자료의 확률강우량 비교

확률강우량 산정을 위하여 본 연구에서 적용한 57개 지점의 기록연수가 서로 상이함을 고려했을 때, 분석에 적용한 4가지 이상치 검정방법 중에서 자료의 크기보다는 순위 혹은 분위수에 근거하여 분석하며 분포의 개괄적 특성을 파악하기 위하여 자주 사용되는 Box-plot 방법을 채택하였다(Choi & Park, 2016). Box-plot 방법에서 나타난 98개 연도의 자료에 대하여 해당 강우의 기각으로 자료를 보완하였으며, 각각의 보완된 자료에 의한 확률강우량을 산정하여 결과를 비교하였다. 비교방법으로는 이상치 기각전과 기각후 강우자료를 지점빈도 해석하였으며, 대상지점인 전국 57개 지점의 24시간 강우자료에 대하여 비교한 결과를 나타내면 <Table 3>과 같다.

<Table 3>을 살펴보면 이상치 기각전과 기각 후 강우자료에 대한 확률강우량을 산정하여 비교하면, 24시간 지속기간의 모든 재현기간에서 기각 전에 비하여 기각 후 강우자료에서 평균 6.1% 감소된 확률강우량을 얻을 수 있었다. 재현기간 20년에서는 확률강우량이 5.9%로 감소함을 보였고, 재현기간 30년에서는 확률강우량이 6.1%로 감소하였다. 재현기간 50년에서

Table 2. Results of four outlier detection tests for 57 stations

No.	Station	Outlier in year						No.	Station	Outlier in year						
		Grubbs and Beck		IACWD		Modified z-score				Box-plot		Grubbs and Beck		IACWD		Modified z-score
1	Sokcho		1984	2002	1984	1984	2002	30	Icheon		1990			1990		
2	Daegwanlyeong	2002		2002	2002		2002	31	Inje							
3	Chuncheon							32	Hongcheon		1990			1990	2006	
4	Gangneung	2002		2002		2002	2002	33	Jecheon							
5	Seoul		1972	2011			1972	2011	34	Boeun		1998	1998	1980	1987	
6	Incheon		1987			1972	1987	35	Cheonan		1995			1995	2002	
7	Wonju		1976			1976	2002	36	Bolyeong		1995	1995	1995	1995	2000	
8	Suwon		1972	1990		1972	1990	37	Buyeo	1987	1987	1987	1987	1995		
9	Seosan		1999			1995	1999	38	Geumsan							
10	Ulsan		1991			1991		39	Buan		2007			2007		
11	Cheongju		1995			1995		40	Imsil							
12	Daejeon		1987	1997		1987	1997	41	Jeongeup	2011	2011	2011	1998	2011		
13	Chupunglyeong		1995	2002		1995	1996	42	Namwon							
14	Andong							43	Jangheung	1981	1981	1981	1974	1981		
15	Pohang	1998		1998		1998	1997	1998	44	Heanam	1981	1981	1981	1981		
16	Gunsan		2000	2011	2000		1979	1987	45	Goheung	1981	1981	1981	1981	1998	
17	Daegu						2000	2005		46	Yeongju	2015*	2002	1994	1999	
18	Jeonju		2005			2005			47	Mungyeong				1987	2002	
19	Ulsan		1969	1991	1969	1991	1969	1991	48	Yeongdeok	1991	1991	2001	1991	1998	
20	Gwangju	1989	1989	2004	1989	1989	2004	49	Uiseong	2015*	1984	2004		1984	2004	
21	Busan		1991		1991	1991			50	Gumi				1998	2004	
22	Tongyeong	1999	1979	1999	1999		1969	1979	51	Yeongcheon	1994*					
23	Mokpo	1981	1981		1981	1981	1981	1985	52	Geochang		2004		2004		
24	Yeosu		2012			2012			53	Hapcheon						
25	Wando		1981		1981		1979	1981	54	Milyang		1979		1979		
26	Suncheon								55	Sancheong		1982		1982		
27	Jinju		2011			2011			56	Geoje		1985	1999	1985	1999	
28	Ganghwa		1997	1998		1997	1998		57	Namhae				1985	1999	
29	Yangpyeong								Total		15	55	22	98		

*: Low outlier detected in the tests

는 확률강우량이 6.4%로 감소하였으며, 재현기간 100년에는 확률강우량이 6.6%로 감소함을 보였다. 따라서 재현기간이 커질수록 확률강우량 감소량이 5.9%에서 6.6%로 점차 증가하는 양상을 보였다. 특히, 재현기간 100년의 경우에는 확률강우량의 결과를 대

지점별로 살펴보면, 부여지점이 34.1%로 가장 큰 감소를 보였으며, 다음으로 강릉지점이 33.7%이며, 가장 작은 감소를 보인 지점은 속초지점으로 8.9%의 감소율이 나타났다. 재현기간별 이상치 발생에 따른 확률강우량의 평균변화율을 나타내면 <Figure 1>과 같다.

Table 3. Comparative of rainfall quantiles in 24hr rainfall duration

No.	Station	Sortation	Return periods				No.	Station	Sortation	Return periods			
			20Year	30Year	50Year	100Year				20Year	30Year	50Year	100Year
1	Sokcho	before rejection	448.7	495.4	553.6	632.2	30	Icheon	before rejection	417.2	460.7	515.0	588.4
		after rejection	412.8	454.2	505.9	575.7			after rejection	417.2	460.7	515.0	588.4
2	Daegwanlyeong	before rejection	553.9	616.2	694.0	798.9	31	Inje	before rejection	380.2	419.5	468.6	534.7
		after rejection	425.2	467.1	519.6	590.3			after rejection	380.2	419.5	468.6	534.7
3	Chuncheon	before rejection	386.9	425.8	474.3	539.8	32	Hongcheon	before rejection	397.7	439.3	491.2	561.3
		after rejection	386.9	425.8	474.3	539.8			after rejection	397.7	439.3	491.2	561.3
4	Gangneung	before rejection	584.1	654.1	741.7	859.7	33	Jecheon	before rejection	365.6	402.1	447.8	509.4
		after rejection	406.7	448.3	500.2	570.3			after rejection	365.6	402.1	447.8	509.4
5	Seoul	before rejection	477.0	528.3	592.5	679.1	34	Boeun	before rejection	395.0	438.4	492.6	565.8
		after rejection	477.0	528.3	592.5	679.1			after rejection	347.5	383.8	429.1	490.2
6	Incheon	before rejection	411.9	456.6	512.6	588.0	35	Cheonan	before rejection	329.6	360.7	399.7	452.2
		after rejection	411.9	456.6	512.6	588.0			after rejection	329.6	360.7	399.7	452.2
7	Wonju	before rejection	397.0	437.8	488.7	557.4	36	Bolyeong	before rejection	389.9	432.7	486.1	558.1
		after rejection	397.0	437.8	488.7	557.4			after rejection	343.0	378.7	423.3	483.5
8	Suwon	before rejection	456.7	506.1	567.8	651.1	37	Buyeo	before rejection	393.7	436.2	489.2	560.8
		after rejection	456.7	506.1	567.8	651.1			after rejection	276.2	299.9	329.5	369.4
9	Seosan	before rejection	365.0	402.3	448.9	511.7	38	Geumsan	before rejection	279.9	303.8	333.6	373.8
		after rejection	365.0	402.3	448.9	511.7			after rejection	279.9	303.8	333.6	373.8
10	Uljin	before rejection	297.1	325.9	361.7	410.2	39	Buan	before rejection	298.0	325.9	360.8	407.9
		after rejection	297.1	325.9	361.7	410.2			after rejection	298.0	325.9	360.8	407.9
11	Cheongju	before rejection	291.1	318.2	352.0	397.6	40	Imsil	before rejection	291.1	318.5	352.8	399.0
		after rejection	291.1	318.2	352.0	397.6			after rejection	291.1	318.5	352.8	399.0
12	Daejeon	before rejection	364.0	401.3	447.8	510.6	41	Jeongeup	before rejection	348.3	385.0	430.9	492.7
		after rejection	364.0	401.3	447.8	510.6			after rejection	280.9	307.0	339.7	383.8
13	Chupunglyeong	before rejection	284.5	312.1	346.6	393.1	42	Namwon	before rejection	261.3	283.6	311.6	349.3
		after rejection	284.5	312.1	346.6	393.1			after rejection	261.3	283.6	311.6	349.3
14	Andong	before rejection	203.0	219.4	239.8	267.3	43	Jangheung	before rejection	468.9	519.1	581.9	666.5
		after rejection	203.0	219.4	239.8	267.3			after rejection	367.6	401.9	444.7	502.5
15	Pohang	before rejection	412.5	460.0	519.5	599.7	44	Heanam	before rejection	424.2	469.7	526.5	603.1
		after rejection	315.4	347.6	387.8	442.0			after rejection	350.7	384.7	427.3	484.8
16	Gunsan	before rejection	355.5	393.1	440.1	503.5	45	Goheung	before rejection	514.5	571.4	642.4	738.2
		after rejection	326.2	359.5	401.1	457.2			after rejection	427.3	470.7	524.9	598.1
17	Daegu	before rejection	275.2	301.4	334.1	378.3	46	Yeongju	before rejection	332.6	367.2	410.4	468.8
		after rejection	275.2	301.4	334.1	378.3			after rejection	332.6	367.2	410.4	468.8
18	Jeonju	before rejection	293.7	321.8	357.0	404.5	47	Mungyeong	before rejection	247.9	269.1	295.7	331.5
		after rejection	293.7	321.8	357.0	404.5			after rejection	247.9	269.1	295.7	331.5
19	Ulsan	before rejection	432.5	480.0	539.4	619.6	48	Yeongdeok	before rejection	297.6	327.9	365.8	416.9
		after rejection	337.2	370.2	411.5	467.1			after rejection	265.9	291.4	323.3	366.4
20	Gwangju	before rejection	347.2	382.3	426.1	485.3	49	Uiseong	before rejection	239.7	262.2	290.3	328.2
		after rejection	312.2	342.0	379.2	429.4			after rejection	239.7	262.2	290.3	328.2
21	Busan	before rejection	426.1	469.9	524.5	598.3	50	Gumi	before rejection	254.2	277.8	307.3	347.1
		after rejection	384.1	421.4	468.1	531.1			after rejection	254.2	277.8	307.3	347.1
22	Tongyeong	before rejection	380.1	418.7	466.9	532.0	51	Yeongcheon	before rejection	228.8	247.4	270.8	302.2
		after rejection	343.5	376.6	418.0	473.8			after rejection	228.8	247.4	270.8	302.2
23	Mokpo	before rejection	321.9	355.0	396.3	452.1	52	Geochang	before rejection	295.2	321.8	355.0	399.8
		after rejection	253.8	276.1	304.0	341.7			after rejection	295.2	321.8	355.0	399.8
24	Yeosu	before rejection	392.1	431.1	479.8	545.4	53	Hapcheon	before rejection	353.2	387.9	431.2	489.7
		after rejection	392.1	431.1	479.8	545.4			after rejection	353.2	387.9	431.2	489.7
25	Wando	before rejection	483.7	536.5	602.5	691.5	54	Milyang	before rejection	337.6	370.9	412.5	468.6
		after rejection	441.5	488.1	546.4	625.0			after rejection	337.6	370.9	412.5	468.6
26	Suncheon	before rejection	370.0	404.2	446.8	504.4	55	Sancheong	before rejection	466.9	515.2	575.6	657.1
		after rejection	370.0	404.2	446.8	504.4			after rejection	466.9	515.2	575.6	657.1
27	Jinju	before rejection	356.5	389.4	430.5	485.9	56	Geoje	before rejection	546.0	605.1	678.9	778.4
		after rejection	356.5	389.4	430.5	485.9			after rejection	430.8	472.5	524.8	595.2
28	Ganghwa	before rejection	481.0	531.0	593.6	677.9	57	Namhae	before rejection	489.1	536.8	596.5	677.0
		after rejection	481.0	531.0	593.6	677.9			after rejection	489.1	536.8	596.5	677.0
29	Yangpyeong	before rejection	444.4	491.1	549.4	628.1			before rejection				
		after rejection	444.4	491.1	549.4	628.1			after rejection				

Table 4. Results of outlier analysis by the cause of rainfall occurrence

No.	Station	Year	Cause of rainfall occurrence	No.	Station	Year	Cause of rainfall occurrence
1	Sokcho	1984	severe rain storm	25	Wando	2002	typhoon
		2002	typhoon			2004	typhoon
2	Daegwanlyeong	2002	typhoon	27	Jinju	2011	severe rain storm
4	Gangneung	2002	typhoon	28	Ganghwa	1997	severe rain storm
5	Seoul	1972	typhoon			1998	severe rain storm
		2011	severe rain storm	30	Icheon	1990	severe rain storm
6	Incheon	1972	typhoon			32	Hongcheon
		1987	severe rain storm	2006	severe rain storm		
7	Wonju	1976	severe rain storm	34	Boeun	1980	severe rain storm
		2002	typhoon			1987	severe rain storm
8	Suwon	1972	typhoon			1997	changma
		1990	severe rain storm			1998	frontal rainfall
9	Seosan	1995	severe rain storm	35	Cheonan	1995	typhoon
		1999	typhoon			2002	typhoon
10	Ulsan	1991	typhoon	36	Bolyeong	1995	typhoon
11	Cheongju	1995	typhoon			2000	changma
12	Daejeon	1987	changma	37	Buyeo	1987	changma
		1997	severe rain storm			1995	severe rain storm
13	Chupunglyeong	1995	severe rain storm	39	Buan	2007	changma
		1996	severe rain storm	41	Jeongeup	1998	typhoon
		2002	typhoon			2011	severe rain storm
15	Pohang	1997	severe rain storm	43	Jangheung	1974	severe rain storm
		1998	typhoon			1981	typhoon
		2005	typhoon			2002	typhoon
1979	typhoon	2004	typhoon				
16	Gunsan	1987	severe rain storm	44	Heanam	1981	typhoon
		2000	severe rain storm	45	Goheung	1981	typhoon
		2005	frontal rainfall			1998	typhoon
		2011	severe rain storm			2002	typhoon
		2012	severe rain storm	1994	severe rain storm		
18	Jeonju	2005	frontal rainfall	46	Yeongju	1999	severe rain storm
19	Ulsan	1969	frontal rainfall	47	Mungyeong	2002	typhoon
		1991	typhoon			1987	severe rain storm
		2005	typhoon	2002	typhoon		
20	Gwangju	1989	changma	48	Yeongdeok	1991	typhoon
		2004	typhoon			1998	typhoon
21	Busan	1991	typhoon			2001	frontal rainfall
22	Tongyeong	1969	frontal rainfall	49	Uiseong	1984	severe rain storm
		1979	typhoon			2004	typhoon
		1982	typhoon	50	Gumi	1998	typhoon
		1985	severe rain storm			2004	typhoon
		1999	severe rain storm			2012	severe rain storm
23	Mokpo	1981	typhoon	52	Geochang	2004	typhoon
		1985	severe rain storm	54	Milyang	1979	typhoon
24	Yeosu	2012	typhoon	55	Sancheong	1982	typhoon
25	Wando	1979	typhoon	56	Gyeoje	1985	frontal rainfall
		1981	typhoon			1991	typhoon
		1985	severe rain storm			1998	typhoon
		1998	typhoon			1999	severe rain storm

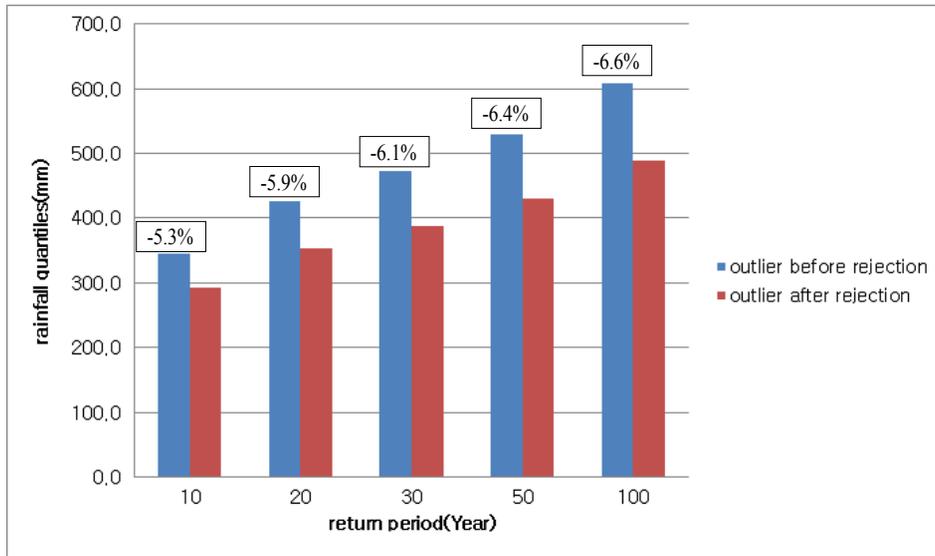


Figure 1. Comparative of average rainfall quantile for return periods

3. 이상치 자료의 강우 발생원인 분석

본 연구에서는 분석된 이상치 검정자료에서 Box-plot 방법에서 고이상치로 나타난 강우자료들을 해당 지점의 이상치 자료로 채택하였으며, 해당 이상치 강우자료의 발생 시기를 조사하여 강우 발생 원인을 분석하였다. 강우 발생원인 분석에는 해당 기간의 기상청 자료를 참조하였으며, 강우 발생원인은 크게 장마, 태풍, 집중호우와 전선형 강우로 구분하였다. 전체 지점의 이상치 자료에 해당하는 발생연도와 기상현황을 나타내면 <Table 4>와 같으며, 전체 57개 지점의 기상현황별 이상치 발생빈도를 종합하여 나타내면 <Table 5>와 같다. <Table 4>와 <Table 5>를 살펴보면, 국내에서 발생한 이상치 강우들은 98개 강우사상으로 분류할 수 있으며, 태풍의 영향으로 발생하는 비율이 51.0%이며, 다음으로 집중호우는 35.7%, 전선형 강우는 7.2%, 장마는 6.1% 순서로 나타났다. 즉, 국내에서 발생하는 극치강우는 태풍과 집중호우의 영향을 가장 많이 받고 있음을 알 수 있으며, 일반적인 기상상황인 전선형 강우와 활동기간이 정해져 있는 장마의 경우에는 영향이 미미함을 알 수 있다. 기상상황별 발생 지점 현황을 살펴보면, 속초, 강릉, 인천, 서산, 울진, 울산, 통영, 목포, 완도, 장흥, 고흥, 영덕, 거제지점들과 같이 해안에 인접한 지점들은 태풍에 의한 영향을

Table 5. The classification of outliers by the cause of rainfall occurrence

Cause of rainfall occurrence	No. of Outliers	Ratio (%)
Changma	6	6.1
Typhoon	50	51.0
Severe rain storm	35	35.7
Frontal rainfall	7	7.2
Total	98	100.0

Table 6. The classification of outliers by the occurrence year

Year	No. of Outliers	Ratio (%)
Before 1979	11	11.2
1980~1989	21	21.4
1990~1999	33	33.7
2000~2009	26	26.5
After 2010	7	7.2
Total	98	100.0

많이 받고 있으며, 그 외 내륙지점들은 집중호우의 영향에 의한 이상치 강우들이 많이 발생하였다.

증가하고 있는 기상기후에 따른 변동분석을 위하여 이상치 강우의 발생연도를 조사하면 <Table 6>과 같다. <Table 6>을 살펴보면, 1980년대에는 21.4%, 1990년대에는 33.7%, 2000년대에는 26.5%로서 과거에 비하여 증가성향으로 나타났으며, 1979년 이전과 2010년 이후에는 상대적으로 적은 11.2%와 7.2% 순서

로 발생하였다. 시기적으로 국내에 큰 영향을 준 태풍들이 1990년대와 2000년대에 다수 발생하였던 이력이 반영된 것으로 판단되며, 2010년 이후에는 이상치 강우의 발생이 상대적으로 작게 나타난 것은, 태풍 등의 많은 강우를 발생시키는 기상현상들의 발생이 예전에 비하여 감소한 현상이 반영된 것으로 판단된다.

IV. 결론

본 연구에서는 30년 이상의 기록연수를 갖는 국내 57개 지점들을 대상으로, 지속기간 24시의 연최대치 강우자료들을 수집하여 국내 강우의 이상치 발생 및 확률강우량 변화를 분석하였으며, 이상치 분석방법으로는 Grubbs and Beck 방법, IACWD 방법, 수정 z-score 방법과 Box-plot 방법, 확률강우량은 Gumbel 분포형을 이용하였다. 이상치검정의 결과에 따라 가장 많은 이상치가 발생한 Box-plot 방법을 적용하여 이상치로 나타난 연도의 자료에 대하여 해당 강우의 기각으로 자료를 보완하였으며, 보완된 자료에 따른 확률강우량의 변화를 분석하였다. 또한, 이상치가 발생한 강우의 원인을 장마, 태풍, 집중호우, 전선형 강우로 분류하여 국내 지점강우의 이상치 발생특성을 파악하였다. 본 연구의 수행으로 인해 얻어진 결과를 요약하면 다음과 같다.

1) 이상치 검정결과 Grubbs and Beck 방법에서는 15개, IACWD 방법에서는 55개, 수정 z-score 방법에서는 22개, Box-plot 방법에서는 98개 강우자료에서 이상치가 분석되었다. Box-plot 방법과 IACWD 방법에 의한 결과는 중복되어 발생한 연도가 다수 나타났으며, 적용한 4가지 방법 모두에서 이상치가 발생한 지점도 12 지점으로 나타났다.

2) 이상치 강우자료의 보완에 따른 확률강우량을 비교하면, 24시간 지속기간의 모든 재현기간에서 기각전에 비하여 기각후에서 평균 6.1% 감소하였으며, 재현기간 20년에서는 5.9%, 재현기간 30년에서는 6.1%로 감소하였다. 재현기간 50년에서는 6.4%, 재현

기간 100년에서는 6.6%로 감소하였으며, 재현기간이 커질수록 확률강우량 감소량이 5.9%에서 6.6%로 증가하는 양상을 보였다. 따라서 강우자료의 기각전·후로 확률강우량의 변화가 있음을 알 수 있으며, 확률강우량 산정시 이상치 검정결과를 적용하는 경우, 이수 및 치수 목적에 따라 설계홍수량 산정에 주의가 필요할 것으로 판단된다.

3) 적용한 이상치 검정에서 고이상치로 판정된 98개 강우사상을 이상치 강우로 선정하였으며, 선정된 강우의 발생원인 분석 결과 태풍 51.0%, 집중호우 35.7%, 전선형 강우 7.2%와 장마 6.1% 순서로 발생하였다. 대상지점들의 지형특성을 고려했을 때, 이상치 발생 강우원인들은 전반적으로 해안 및 해안과 인접 지점들은 태풍, 내륙지점에서는 집중호우에서 이상치들이 많이 발생하였다.

4) 이상치 강우들의 연도별 특성은 1980년대 21.4%, 1990년대 33.7%, 2000년대 26.5%이며, 1979년 이전과 2010년 이후에는 상대적으로 적은 11.2%와 7.2%로 나타나 예전에 비하여 태풍 등의 기상현상이 적게 발생하고 있는 국내 현황이 반영된 것으로 판단된다.

본 연구에서 적용한 전국 57개 지점 중 44개 지점에서 이상치가 나타났으며, 향후 관측지점과 기록연수들이 보완되면 지점별 특화된 사상들에 대한 이상치를 고려한 정도 높은 강우예측이 가능할 것이며, 이수 계획에 따른 수자원 확보, 치수계획에 따른 적절한 저감계획 수립을 유도하여, 향후 국내 지역에 합리적인 수자원 활용기준을 제시할 것으로 판단된다.

References

- Barnett, V. and T. Lewis. 1994. *Outliers in Statistical Data*. New York: John Wiley.
- Choi, H. S. 2016. Clustering Analysis of Particulate Matter Data Using Shrinkage Boxplot. *Journal of The Korean Data Analysis Society*. 2016(5): 2435-2443.
- Choi, M. Y., et. al. 2011. Analysis the Extreme Rainfall Data using the FORGEX Considering the Characteristics of South

- Korea Rainfall Data. Proceedings of the Korea Water Resources Association Conference. May. 2011: 27-31.
- Dixon, W. J. 1950. Analysis of Extreme Values. *Annals of Math.* 21(1): 488-506.
- Dixon, W. J. 1951. Ratios Involving Extreme Values. *Annals of Math.* 22(1): 68-78.
- Grubbs, F. E. 1969. Procedures for Detecting Outlying Observations in Samples. *Technometrics.* 11(1): 1-21.
- Grubbs, F. E. and G. Beck. 1972. Extension of Sample Sizes and Percentage Points for Significance Tests of Outlying Observations. *Technometrics.* 14(4): 847-854.
- Harmeling, S., G. Dornhege, D. Tax, F. Meinecke, and K. R. Muller. 2006. From Outliers to Prototypes: Ordering Data. *Neurocomputing.* 69(13-15): 1608-1618.
- Hawkins, D. 1980. Identification of Outliers. London: Chapman and Hall.
- Hosking, J. R. M. 1990. L-moments-Analysis and Estimation of Distribution using Linear Combinations of Order Statistics. *Journal Royal Statistical Society B.* 52(1): 105-124.
- Iglewicz, B. and David C. Hoaglin. 1993. How to Detect and Handle Outliers. *ASQC Quality Press.* c1993: 87.
- Interagency Committee on Water Data(IACWD). 1982. Guidelines for Determining Flood Flow Frequency. Bulletin 17B of Hydrology and Subcommittee. U. S. Geological Survey. Office of Water-Data Coordination.
- Jaiswal, R. K., N. K. Goel, P. Singh, and T. Thomas. 2003. L-moment based Flood Frequency Modeling. *The Institution of Engineers.* 84: 6-10.
- Johnson, R. 1992. Applied Multivariate Statistical Analysis. Prentice Hall.
- Jung, D. H. 2016. A Study on the Estimation of Probability Rainfall Considering Outliers in Korea. Master's thesis. Kumoh National Institute of Technology.
- Kim, M. S. and Y. I. Moon. 2015. A Study on Quality Control Method for Minutely Rainfall Data. *Journal of The Korean Society of Civil Engineers.* 35(2): 319-326.
- Korea Meteorological Administration. 2018. Annual Climatological Report(in Korea).
- Korea Meteorological Administration Weather Information. <http://www.weather.go.kr>
- Lee, J. S., et. al. 2017. A Study on the Occurrence of Outliers for the Rainfall in Korea. *Journal of The Korean Society of Civil Engineers.* 2017(10): 217-218.
- Lee, J. S., et. al. 2018. An Analysis of the Extreme Rainfall Considering Outliers in Korea. *Journal of The Korean Society of Civil Engineers.* 2018(10): 90-91.
- McCormick, D. L. and A. R. Rao. 1995. Outlier Detection in Indiana Flood Data, Tech. Rept. CE-EHE-95-04. School of Civil Engineering, Purdue University. W. Lafayette. 47907: 132.
- Mendenhall, W., J. E. Reimuth, and R. J. Beaver. 1993. Statistics for Management and Economics. Belmont, CA: Duxbury Press.
- Ministry of Land, Transport and Maritime Affairs. 2012. Design Flood Quantity Calculation Method.
- Moon, Y. L., et. al. 2014. A Study on Quality Improvement of Aws Rainfall Data at the Seoul Observatory. *Journal of The Korean Society of Hazard Mitigation.* 2014: 367-367.
- Pegram, G. 1997. Patching Rainfall Data Using Regression Methods. 3. Grouping, Patching and Outlier Detection. *Journal of Hydrology.* 198(1): 319-334.
- Pilon, P. I., R. Condie, and K. D. Harvey. 1985. Consolidated Frequency Analysis Package(CFA). User Manual for Version 1 - DEC PRO Series. Water Resources Branch. Inland Waters Directorate. Environment. Canada, Ottawa.
- Song, H. S. 2012. Safety Evaluation and the Application of Urban Stream by Analysis of Extreme Rainfall Considering Occurrence Cause and Climate Change. Doctor's Thesis. University of Seoul Reserved.
- Stefansky, W. 1972. Rejecting Outliers in Factorial Designs. *Technometrics.* 14: 469-479.
- Thompson, W. 1935. On a Criterion for the Rejection of Observations and the Distribution of the Ratio of Deviation to Sample Standard Deviation. *The Annals of Mathematical Statistics.* 6: 214-219.
- Tukey, J. W. 1977. Exploratory Data Analysis. Addison-wesley.
- Vogel, R. M. and Fennessey, N. M. 1993. L-Moment Diagrams Should Replace Product Moment Diagrams. *Water Resources Research.* 29(6): 1745-1752.

Korean References Translated from the English

- 국토해양부. 2012. 설계홍수량 산정요령.
- 기상청. <http://www.weather.go.kr>.
- 김민석, 문영일. 2015. 분 단위 강우자료의 품질 개선방안에 관한 연구. 대한토목학회 논문집. 35(2): 319-326.
- 문영일, 김종석, 김민석, 황성환. 2014. 서울지점의 AWS 강우자료의 품질개선에 대한 연구. 한국방재학회 학술발표대회 논문집. 2014: 367-367.
- 송현섭. 2012. 기후변화를 고려한 호우원인별 극한강우 분석에 따른 도시하천 안전성 평가 및 적용. 박사학위논문. 서울 시립대학교 토목공학과.
- 이정식, 손호근, 신창동. 2018. 이상치를 고려한 국내 극치강우 분석. 대한토목학회 학술대회. 2018(10): 90-91.
- 이정식, 신창동, 이용재. 2017. 국내 강우의 이상치 발생에 관한 연구. 대한토목학회 학술대회. 2017(10): 217-218.
- 정대회. 2016. 이상치를 고려한 국내 확률 강우량 산정에 관한 연구. 석사학위논문. 금오공과대학교 토목공학과.
- 정중호, 윤용남. 2003. 수자원설계실무.
- 최민영, 신주영, 남우성, 허준행. 2011. 국내 강우자료의 특성을 고려한 FORGEX 기법의 극한강우 분석. 한국수자원학회 학술발표회논문집. 2011(05): 25-31.
- 최호식, 박창이. 2016. 축소된 상자그림을 활용한 미세먼지 데이터의 군집분석. 한국자료분석학회지. 2016(5): 2435-2443.

Received: Sep. 3, 2019 / Revised: Sep. 20, 2019 / Accepted: Sep. 26, 2019

이상치 검정을 이용한 국내 강우의 연구

국문초록 연최대 강우자료의 이상치는 확률강우량을 산정하는 빈도해석에서 중요한 역할을 한다. 신뢰성 있는 빈도해석을 위하여 본 연구에서는 연 최대 강우자료의 통계적 검토에 의한 이상치 검정과 아울러 이상치 강우들을 선별하여 강우 발생 원인을 검토하는 정성적 분석을 수행한다. 이를 위하여 국내 57개 기상관측지점의 24시간 강우자료들을 대상으로 Grubbs and Beck 방법, IACWD 방법, 수정 z-score 방법과 Box-plot 방법을 이용하여 이상치 검정을 실시하였다. 선정된 고이상치의 강우사상에 대해 발생 원인을 장마, 태풍, 집중호우와 전선형 강우로 분류하여 분석하였다. 분석결과, 방법들의 비교를 통해 Grubbs and Beck 방법이 다른 방법에 비하여 가장 적게 이상치가 발생하였다. 강우 발생 원인 분석에서 태풍 51.0%, 집중호우 35.7%, 전선형 강우 7.2%와 장마 6.1% 순서로 나타났으며, 내륙 지점에서는 집중호우, 해안과 인접한 지역들은 태풍에 의한 이상치 들이 많이 발생하였다. 검정을 통해 발생된 이상치를 기각하여 산정한 확률강우량과 이상치 검정전의 확률강우량을 비교하였다.

주제어 : 이상치 검정, 강우발생원인, 확률강우량

Profiles **Ho Geun Son** : He received his B.S., M.S. from Kumoh National Institute of Technology, Korea in 2005. He joined the Korea Rural Community Corporation in 1996 works as a deputy director. His interesting subject and of research is Emergency Action Plan(EAP)(hgson@ekr.or.kr).

Jae Joon Lee : He received his B.A., M.A., Ph.D. from Yonsei University, Korea in 1987. He is a Professor of the Department of Civil Engineering at Kumoh National Institute of Technology, in which he has taught since 1991. His interesting subject and area of research and education is disaster prevention, Water and Wastewater, and policy implementation. He has published 100 articles in journals and written 3 books, including co-author(jhb365@kumoh.ac.kr).