

확률 강우량의 변동성 분석
An Analysis of the variability of rainfall quantile estimates

정성인*, 유철상**, 윤용남***1)
Sung In Jung, Chul Sang Yoo, Yong Nam Yoon

.....
Abstract

Due to the problems of global warming, the frequency of meteorological extremes such as droughts, floods and the annual rainfall amount are suddenly increasing. Even though the increase of greenhouse gases, for example, is thought to be the main factor for global warming, its impact on global climate has not yet been revealed clearly in rather quantitative manners. Therefore, the objective of this study is to inquire the change of precipitation condition due to climate change by global warming. In brief, this study want to see its assumption if rainfall quantile estimates are really changing.

In order to analyze the temporal change, the rainfall quantile estimates at the Seoul rain gauge stations are estimated for the 21-year data period being moved from 1908 to 2002 with 1-year lag. The main objective of this study is to analyze the variability of rainfall quantile estimates using four methods.

Next, The changes in confidence interval of rainfall quantile are evaluated by increasing the data period. It has been found that confidence interval of rainfall quantile estimates is reduced as the data period increases. When the hydraulic structures are to be designed, it is important to select the data size and to re-estimate the flood prevention capacity in existing river systems.

Key words : Global Warming, Rainfall quantile estimate, Confidence interval,
.....
.....

지구온난화로 인해 강우의 량 뿐 아니라 가뭄과 홍수와 같은 극치 강우사상의 빈도가 증가하고 있다. 하지만 지구 온난화가 얼마나 많은 강수를 발생시켰는지는 명확하지가 않다. 따라서 실제로 지구온난화로 인해 강우사상이 변하고 있는가를 분석하는 것이 이 논문에 목적이다. 서울 강우 관측소의 관측자료를 이용해 1908~2002년까지 21개 자료부터 한 개씩 추가해 가는 방법으로 빈도분석을 하여 확률강우량에 대한 변동성을 분석하였다. 다음으로 자료기간의 증가에 따른 확률강우량의 신뢰구간을 평가하였는데 확률강우량의 신뢰구간은 자료의 기간이 증가할수록 감소하는 것을 알 수 있었다.

핵심용어 : 지구온난화, 확률강우량, 신뢰구간

1. 서론

최근 들어 우리나라뿐 아니라 전 세계 곳곳에서 가뭄, 홍수, 한파 등 이상 기상이 빈번히 발생하고 있다.

* 정회원·동부엔지니어링 수자원부
** 정회원·고려대학교 사회 환경시스템 공학과 교수
***정회원·고려대학교 사회 환경시스템 공학과 교수

이로 인해 계획 홍수량을 초과하는 극치 강우량이 빈번하게 발생하여 기존 홍수 방어 시설물의 치수 안전도를 저하시키고 있다고 한다. 이런 현상의 발생원인이 지구온난화로 인한 기후변화라는 논쟁이 제기되고 있다. 하지만 아직까지 역학적-물리적 이론에 근거한 명확한 설명은 충분하지 않다(차은정, 2001).

지구 온난화로 인한 증발량이 얼마나 자주, 얼마나 많은 강수를 발생시켰는지는 명확하지가 않으며 어느 지역에 강수를 증가시켰는지에 대해서도 명확하지가 않다. 따라서 본 연구에서는 지구온난화로 인한 기후변화가 강우사상에 실제로 얼마나 많은 영향을 주었는가를 알아보고자 한다.

강우사상이 변하면 수공구조물의 수문학적 설계기준인 확률강우량이 변하게 된다. 이상기후로 인하여 강우량이 증가하였다면 확률강우량도 변하게 될 것이고 그렇게 된다면 기존에 설계된 수공구조물의 치수 안전도는 저하되게 된다. 그렇기 때문에 수공학 분야에서는 강우사상의 변화와 함께 확률강우량의 변화를 파악하는 것이 중요하다고 할 수 있겠다. 따라서 본 연구에서는 확률강우량의 변화추이를 알아보고 최근에 관측되는 강우에 의해 확률강우량이 유의한 변동성을 갖고 있는지에 대해서 분석해 보고자 한다.

2. 서울지역 강우관측 자료와 확률강우량의 정상성 판별

2.1 콕스-스튜어트의 추세검정

서울지역 강우관측 자료와 확률강우량이 시간의 경과에 따라 관측되는 값들이 어떤 형태의 추세를 보이는지를 알아보기 위해 콕스-스튜어트의 추세검정을 실시하였다.

<표 1> 서울지역 관측 자료 및 확률강우량에 대한 콕스-스튜어트 추세검정 결과

자료 계열	자료형태	분포형	빈도	(-)부호의 개수	(+)부호의 개수	동점인 경우	Z	비고
일 최대 자료	서울지역 강우관측 자료	-	-	20	25	0	-0.745	추세 없음
	확률강우량	EV1	10	30	5	0	4.226	하향 추세
			20	27	8	0	3.212	하향 추세
			50	19	15	1	0.507	추세 없음
			100	25	10	0	2.535	하향 추세

표 1은 콕스-스튜어트 추세검정을 이용하여 서울지역 강우관측 자료와 확률강우량에 대하여 유의수준 5%에서 추세검정을 한 결과이다.

유의수준 α 가 0.05이므로 대표본근사이론에 의해, 계산된 Z값이 -1.96보다 작으면 상향추세가 있는 것이고 +1.96보다 크면 하향 추세가 있는 것이다. Z값이 -1.96과 +1.96사이에 있으면 상향적 추세나 하향적 추세가 없다고 할 수 있다. 표 1을 보면 Z값이 -0.745이기 때문에 서울지역 강우관측 자료에 대해서는 추세가 없다는 결론을 얻을 수 있다. 하지만 확률강우량의 추세검정에서는 50년 빈도를 제외하고는 하향추세가 있는 것을 확인 할 수 있었다.

2.2 Dickey-Fuller의 단위근 검정

자료의 정상성을 판별하기 위하여 가장 많이 쓰이는 방법은 단위근검정이다. 시계열이 단위근을 가지고 있는지를 검정하는 방법에는 여러 가지가 있는데 여기에서는 Dickey-Fuller 단위근 검정방법을 이용하였다. 서울지역 강우관측 자료와 확률강우량이 정상성을 가지고 있는 가를 판별하기 위하여 통계 프로그램인 SAS(Statistical Analysis System)를 이용하여 Dickey-Fuller(DF)의 단위근 검정을 하였다. 서울지역 강우관측자료와 확률강우량 대한 단위근 검정 결과는 아래 <표 2>와 같다.

<표 2> 서울지역 관측 자료 및 확률강우량에 대한 Dickey-Fuller 단위근 검정

자료 계열	자료형태	분포형	빈도	유의 확률(p-value)	유의 수준	비 고
일 최대 자료	서울지역 강우관측 자료	일 최대 자료	-	0.0000954542	0.05	정상시계열
	확률강우량	EV1	10	0.7140129189	0.05	비정상시계열
			20	0.6263155675	0.05	비정상시계열
			50	0.1211159544	0.05	비정상시계열
			100	0.4871571047	0.05	비정상시계열

Dickey-Fuller의 단위근 검정결과 서울지역 강우 관측 자료에 대해서는 유의 수준 5%에서 정상시계열이 임의 확인되었지만 확률강우량에 대해서는 비정상 시계열임이 확인되었다. 이와 같이 확률강우량에서 추세가 발견되고 비정상 시계열이라는 결과가 나온 것은 확률강우량의 몇몇 경우에서 추세가 나타난 것은 자료기간이 강우자료를 확률강우량 자료로 변환시킬 때 자료기간이 짧아진 것으로 인한 것이라 생각되며 비 정상시계열로 판정 된 것은 확률강우량 추정과정에서 앞·뒤 값의 상관성이 매우 커졌기 때문인 것으로 판단된다.

그렇기 때문에 확률강우량이 추세가 나타나고 비정상 시계열로 판별되었지만 비정상 시계열로 판단하는 것은 다소 무리가 있다. 따라서 본 연구에서는 4장에서 제시하는 3가지 방법으로 확률강우량의 변동성을 분석함으로써 확률강우량이 정상성을 판별하고자 한다.

3. 확률강우량의 변동성 분석

3.1 자료기간 중 최대 값의 재현기간 평가

확률강우량의 변동성에 대해서 알아보기 위해 자료기간 중 최대 값의 재현기간을 추정하여 보았다. 정상적인 시계열 자료에서는 n개의 자료 중에 가장 큰 값($m=1$)의 재현기간은 $n+1$ 이라는 것을 Weibull의 확률도 시공식을 통해서 알 수 있으며 이는 자료 기간 중 최대 값의 재현기간은 자료의 총수인 n과 비슷한 값을 갖는 다는 것을 의미한다.

관측된 서울 강우량 자료 중 년 강수량자료에 대해서 처음 21개의(1908~1928년)자료에서부터 일개년씩의 자료를 추가해 가면서 분석을 한 후 빈도분석에 사용된 자료 중 최대 강우관측 값에 대한 재현기간을 추정하였다.

자료기간 중 최대 값의 재현기간 추정에 사용된 강우자료가 정상적인 시계열 자료라면 이를 이용하여 산정된 확률강우량도 정상적이기 때문에 n개의 자료 중 최대 값에 대한 재현기간은 자료의 총수, n과 비슷할 것이라는 것이다. 따라서 정상적인 시계열 자료라면 최대 값의 재현기간과 자료의 총수가 비슷한 값을 갖게 되므로 아래 그림 3.1에서 검은 실선으로 그려진 $y=x$ 그래프 근처에 추정된 값들이 존재하게 될 것이다.

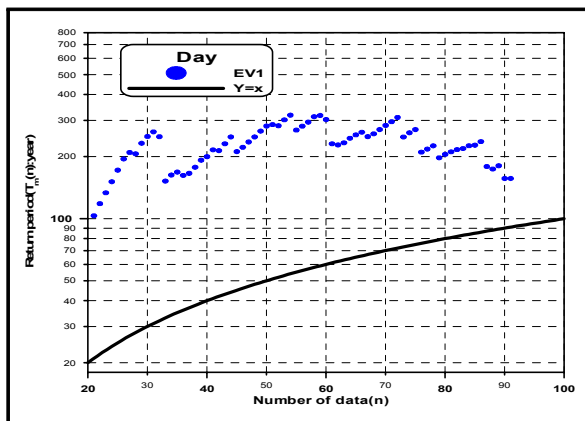


그림 3.1 자료기간중 최대값의 재현기간 평가

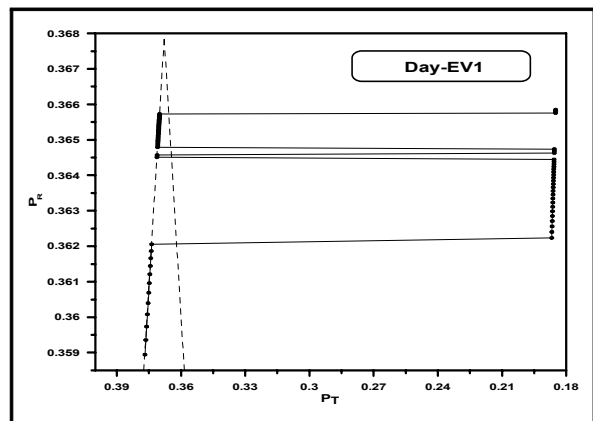


그림 3.2 기록년수 초과 강우 발생빈도

일 최대 자료를 이용하여 최대 값에 대한 재현기간을 추정한 결과 처음에는 점들이 $y=x$ 그래프와 상당히 이격되어 있는 것을 볼 수 있지만 자료의 수가 증가함에 따라 근접해 가는 것을 확인 할 볼 수 있다. 자료의 개수가 증가함에 따라 자료기간 중 최대 값에 대한 재현기간은 $y=x$ 그래프에 점점 더 접근해 가는 것을 볼 수 있었는데 이는 원 자료 계열과 확률강우량이 자료의 개수가 증가함에 따라 안정성을 갖는다는 것을 의미하는 것이다. 따라서 최근 들어 지구온난화로 인해 극단적인 강우 사상이 많이 발생하여 자료계열이 정상성을 갖고있지 않는다는 판단은 하기 어렵었다. 또한 위 그림 3.1에서와 같이 자료의 수가 증가함에 따라 안정성을 갖는 경향으로 간다는 것은 지금 우리가 가지고 있는 강우관측 기록이 아직은 부족하다는 것을 반증하는 결과이기도 하다.

3.2 기록년수를 초과하는 재현기간을 갖는 강우의 발생빈도 평가

기록년수를 초과하는 재현기간을 갖는 강우란 n개의 자료를 가지고 빈도분석을 하였을 때 n년보다 큰 재현기간을 갖는 사상을 말한다. 즉 기록년수 n에 해당하는 확률강우량보다 큰 사상을 말한다. 기록년수를 초

과하는 재현기간을 갖는 강우의 발생빈도 평가 방법은 우선 n개의 자료를 가지고 빈도분석을 하여 n년 빈도의 확률강우량을 산정한다. 확률강우량의 산정에 사용된 n개의 자료 중에 n년의 재현기간에 해당하는 확률강우량 보다 큰 강우 관측치가 몇 개가 기록되었는가를 조사한다. 여기에서도 3.1절에서와 같이 자료의 개수에 따른 기록년수를 초과하는 강우사상의 발생횟수를 조사하여 그 발생확률을 계산하였다. 그 결과는 그림 3.2와 같다.

정상적인 시계열 자료라면 n개의 자료를 가지고 빈도분석을 했을 때 n년 보다 큰 사상의 출현은 한번으로 예측 할 수 있다. n년 동안에 T-년 사상의 발생예상 횟수는 T/n이므로 n년 동안에 n년 초과사상의 발생예상 횟수는 n/n 즉, 1번이라고 예상할 수 있기 때문이다. 그림 3.2에서 정상적인 강우가 발생하였을 경우 점선 위에 점이 위치하게 되고 기록년수를 초과하는 강우사상이 2회 이상 발생할 경우는 그래프 오른쪽에 점이 존재하게 된다. 그림 3.2를 보면 기록년수를 초과하는 강우 사상의 발생횟수가 처음에는 한번으로 정상적인 경향을 보이지만 어느 시점에서는 자료의 개수가 추가되어도 계속해서 발생횟수가 2회 인 것을 볼 수 있다. 연속해서 발생횟수가 2회가 나타나는 것은 극단적인 사상이 계속해서 나타나서라기 보다는 이전에 발생된 극단적인 사상이 자료가 추가되어도 계속해서 기록년수를 초과하는 강우사상의 발생횟수에 영향을 주는 것이라 할 수 있다. 자료의 수가 더욱 많아지면 다시 정상적으로 되돌아오는 것을 볼 수 있다. 이번 절에서도 자료기간이 증가함에 따라 기록년수를 초과하는 강우사상의 발생횟수가 1 또는 0으로 수렴해 가는 것으로 보아 자료의 수가 많아지는 최근에 확률강우량이 더욱 더 정상적인 자료계열이 되어 가는 것을 확인 할 수 있었다.

3.3 최대 관측치의 재현기간에 대한 신규자료의 영향 평가

3.1절에서는 관측기간 중 최대 값에 대한 재현기간을 추정해 보았다. 여기에서는 1개의 자료를 추가하여 n+1개의 자료로 빈도분석을 한 후 n+1번째의 자료를 제외시킨 후 n개의 자료 중에 최대 값을 찾아 그 값에 대한 재현기간을 추정하여 3.1절의 결과와 비교해 보았다.

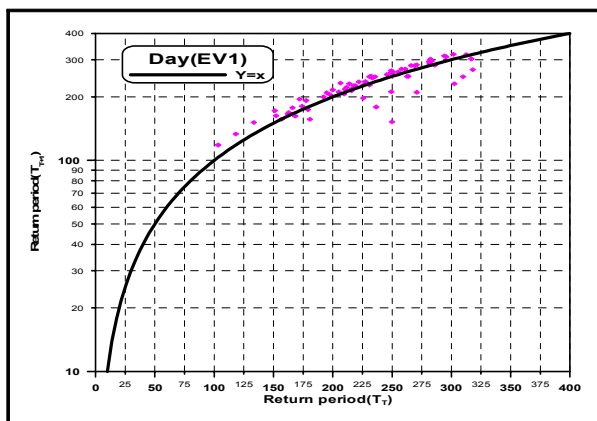


그림 3.3 최대 관측치의 재현기간에 대한 신규 관측치의 영향 평가

이런 재현기간을 추정하여 3.1절의 결과와 비교해 보았다. 이번 절에서 알아보고자 하는 것은 신규자료가 최대 관측치의 재현기간에 대한 영향을 알아보는 것이다. 그림 3.3에서 신규자료 즉, n+1번째 값이 이전의 자료들에 비해 큰 변화가 없을 경우 점들이 실선 가까운 곳에 존재하게 될 것이다. 하지만 n+1번째 값이 크게 상승하였을 경우에는 점은 실선 아래에 위치하게 된다. n+1번째 값이 다른 값에 비해 크면 클수록 점들은 오른쪽 아래로 이동하게 된다. 따라서 오른쪽 아래에 많은 점들이 존재하면 이상기후로 인해 극단적인 사상이 많이 발생하였음을 의미한다.

그림 3.3을 보면 대부분의 경우 실선 가까운 곳에 점들이 위치하고 있는 것을 볼 수 있으며 실선 오른쪽 아래로 이격되어 있는 점들이 만치 않은 것으로 확인되었다. 최대 관측치의 재현기간에 대한 신규자료의 영향 평가에서도 확률강우량의 변동성에 큰 영향을 주는 극단적인 강우사상이 많이 관측되지 않았음을 확인할 수 있었다.

신규자료가 추가됨으로 인해 최대 관측치의 재현기간이 증가하는지 감소하는지에 대한 정량적인 검토를 해 보았다. 신규자료가 추가됨으로 인해 최대 관측치의 재현기간이 증·감을 주는 강우 관측기록이 없다고 말할 수 있으려면 이들 점들에 대해서 회귀분석을 한다면 그 기울기가 1이 되어야 할 것이다. 따라서 기울기를 1이라고 할 수 있는가에 대한 가설검정이 필요하다. 회귀직선에 대한 기울기의 신뢰구간을 유의 수준 5%로 추정한 결과는 표 3.1과 같다.

<표 3.1> 회귀 직선의 기울기에 대한 신뢰구간 검정

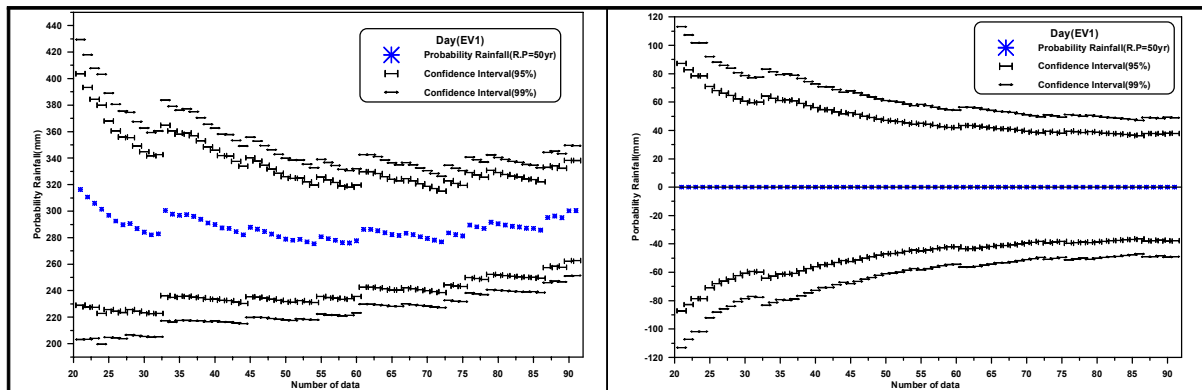
자료 계열	분포형	표준오차	신뢰구간범위 표준오차×1.96	하한 값	기울기(a)	상한 값
일 최대	EV1	0.059144	0.1159	0.8807	0.9966	1.1125

표 3.1을 보면 기울기의 신뢰구간이 1포함하는 것으로 확인됐다. 따라서 5% 유의수준에서 기울기, a는 1이라고 할 수 있다. 최대 관측 치의 재현기간에 대한 신규자료의 영향 평가에서는 신규자료의 추가로 인한 확률강우량의 유의한 변동성은 없다고 할 수 있다.

4. 확률강우량의 신뢰구간

이번 장에서는 지금까지 빈도분석을 해왔던 방법과 마찬가지로 자료의 개수를 21개에서부터 하나씩 증가시키면서 신뢰구간의 변화와 확률강우량의 변화 추이를 살펴보았다. 확률강우량의 신뢰구간의 변화를 살펴보기 위한 빈도는 50년 빈도를 사용하였다. 그 결과는 아래와 같다.

확률강우량의 신뢰구간을 살펴보는 이유는 확률강우량의 변동과 변동정도를 추정하는 것이 확률강우량의 정상성을 이해하는데 도움이 되기 때문이다. 같은 신뢰도를 갖으면서 신뢰구간이 좁아진다는 것은 확률적으로 그 만큼 정확한 예측이 가능하다는 것이며 이는 예측에 대한 불확실성이 작다는 것과 같은 의미이다. 정상적인 강우가 관측되었다고 하면 확률강우량의 신뢰구간은 표본의 크기가 클수록 즉, 관측자료의 수가 많을수록 좁아진다. 하지만 극단적인 사상으로 인해 자료가 불안정해지면 확률강우량의 예측에 대한 불확실성이 커지게 되므로 신뢰구간은 커지게 된다. 왼쪽 그림은 자료의 수가 증가함에 따라 확률강우량과 신뢰구간의 변화 추이를 그린 것이고 오른쪽 그림은 확률강우량 값을 0으로 맞춘 후에 신뢰구간의 폭이 자료의 수가 증가함에 따라 어떻게 변하는지를 그린 것이다.



<그림 4.1> 확률강우량과 신뢰구간 변화

강우의 관측은 일정하지 않으므로 관측되는 강우자료는 증가 혹은 감소 할 수 있다. 관측되는 강우가 계속해서 증가 혹은 감소하게 되면 확률강우량 또한 증가 또는 감소하게 되어 예전에 산정된 신뢰구간을 벗어나게 될 것이다. 위 그림을 보면 최근에 산정된 확률강우량 값들이 과거에 계산된 신뢰구간을 크게 벗어나지 않고 있음을 확인할 수 있다. 따라서 현재 일어나고 있는 확률강우량의 변화는 아주 정상적인 범위 안에서 이루어지고 있다는 판단을 내릴 수 있다. 한편 오른쪽 그림을 보면 자료의 수가 증가함에 따라 신뢰구간의 폭이 대체로 감소하고 있는 것을 볼 수 있다. 자료의 수가 증가함에 따라 신뢰구간의 폭이 좁아진다는 것은 그만큼 확률강우량에 대한 불확실성이 감소한다는 것이고 이는 확률강우량이 안정적이라는 것이다.

5. 결론

본 연구에서 3가지 방법으로 확률강우량의 변동성을 분석해본 결과 서울지역 확률강우량이 비정상적이라는 판단을 내리기 어려웠다. 더욱이 자료의 수가 증가하면 확률강우량은 더욱더 안정성을 갖게된다는 것을 알 수 있었다. 따라서 확률강우량이 비정상 시계열이기보다는 정상시계열 가깝다는 결론을 얻을 수 있었다. 특히 확률강우량의 신뢰구간의 추정에서도 자료의 수가 증가할수록 불 확실성이 줄어드는 것을 확인할 수

있었다. 또한 확률강우량의 신뢰구간을 추정에서도 확률강우량이 이전에 설계된 수리구조물의 안정성을 위협할 만한 큰 변동이 없었다는 것을 알 수 있었고 자료가 증가하면서 신뢰구간의 폭도 대체로 감소하고 있는 것을 볼 수 있었다. 이렇게 신뢰구간 폭이 감소한다는 것은 확률강우량에 대한 불확실성이 줄어드는 것이고 이는 그 만큼 확률강우량이 안정성을 갖는다는 것을 의미한다.