

무작위변량을 이용한 강우빈도분석시 내외삽오차에 관한 연구

A Study on Error of Frequency Rainfall Estimates Using Random Variate

최한규* 엄기옥**
Choi, Han Kyu Eom, Ki Ok

Abstract

In the study rainfall frequency analysis attempted the many specific property data record duration it is difference from occur to error-term and probability distribution of concern manifest. error-term analysis of method are fact sample data using method, in other hand it is not appear to be fault that sample data of number to be small random variates.

Therefore, day-rainfall data to randomness consider of this study sample data to the Monte Carlo method by randomize after data recode duration of form was choice method which compared an assumed maternal distribution from splitting frequency analysis consequence.

In the conclusion, frequency analysis of chuncheon region rainfall appeared samll RMSE to the Gamma II distribution.

In the rainfall frequency analysis estimate RMSE using random variates great transform, RMSE is appear that return period increasing little by little RMSE increased and data number increasing to RMSE decreasing.

키워드: 평균제곱오차

Keywords: RMSE(Route Mean Square Error)

1.서론

수공구조물의 수문학적인 설계나 평가 그리고 수자원계획시에 아주 중요한 것은 홍수량 자료라고 할 수 있는데, 실제로 대부분의 유역은 미계측 유역이기 때문에 홍수량자료를 직접 이용할 수 없으며 자료의 수가 소량이기 때문에 그에 따른 홍수량 산정이나 강우량 산정등 어려움이 있으며, 또한 소량

의 자료치를 이용하여 그에 따른 지역의 강우량을 이용한 홍수량추정에 있어서도 한계가 있음을 지적해 온 것을 알 수 있다. 따라서 강우-유출시스템의 거동특성이 유사하다는 가정 하에 강우자료를 통계 분석하여 설계수문량으로 이용하고 있다.

현재 국내 수자원 실무에서는 미계측 유역의 홍수량을 추정하기 위하여 강우자료를 빈도 분석하여 이를 강우-유출모형의 입력자료로 이용하고 있다. 강우자료의 기록기간은 대체로 30년 정도이며 분석 목적에 따라서 재현기간 2년, 5년, 10년 등의 내삽하는 경우와 50년, 100년, 200년 등 자료기간 밖으로 외삽하는 경우가 있다. 그러나 실제로 빈도분석

* 강원대학교 토목공학과 교수, 공학박사

** 강원대학교 토목공학과 석사과정

한 강우량이 어느 정도의 오차를 가지고 있는지, 또는 어느 정도까지 외삽할 수 있는지 등의 분석결과와 거동특성이 충분히 알려진 바 없기 때문에 자료에 적합한 분포가 모분포라 가정하여 임의적으로 빈도분석하여 강우유출모형의 입력자료로 이용하고 있으며, 이 결과의 평가 또한 아주 다양한 의견이 개진되고 있는 것이 현실이라고 할 수 있다.

본 연구에서는 이러한 현실을 감안하여 강우빈도 분석시 발생하는 여러 특성치중 자료 기록기간에 의한 차이로 발생하는 오차정도와 이때 수반되는 확률분포와의 관계를 밝히고자 시도하였다. 이들 자료기록기간에 따른 오차정도를 분석하는 기법에는 실제 샘플자료를 이용 방법김남원, 최한규[3] 등이 있을 수 있으나, 표본자료의 수가 적음으로 인하여 무작위 변량특성을 충분히 나타낼 수 없는 단점이 있다.

따라서 본 연구에서는 일 강우량 자료의 무작위성을 고려하여 Monte Carlo 방법에 의해 표본자료를 충분히 모의한 후에 자료기록기간의 형태로 스플릿 하여 빈도분석하여 얻어지는 결과치와 가정된 모분포치와 비교하는 방법을 채택하였다. 따라서 오차에 대한 거동 특성을 규명함으로써 빈도해석시 강우자료를 판단하는 기준으로 활용하는 동시에 제한기간에 따른 오차의 정도에 따라 가장 합리적인 확률분포를 제시하여 계획 설계업무에 기여하고자 한다.

홍수빈도에 관한 여러분석 형태에 관한 찬반론자들이 과거 수십년동안 논쟁을 하였다. 특정방법의 지지자들은 그것이 몇몇 다른 방식들보다 우월하며 몇 개의 이론적 설명과 사례연구들에 의해 증명되었다.

그러나 그들의 이론적 설명은 여전히 의심스러운 가설들을 포함하고 있었다.

도수분포의 모든 분석적 형태의 선택에 있어서 최우선 고려사항의 분포는 측정된 데이터에 적합하게 만드는 것이다.

Benson[17]은 미국 전역에 위치한 10개의 지류에서의 홍수빈도분석에 대한 여섯가지 방법을 연구하였다.

이들 자료에 대한 기록은 40년부터 97년까지 55년간의 평균에 걸쳐 있다.

그 여섯 방식은 Gamma 분포, Gumbel분포, 데이터의 대수를 이용한 Gumbel분포, Log-Normal 평균분포 Log Person Type III 그리고 하천의 방법으로 이루어져 있다.

홍수빈도분석은 그 중요성 때문에 금세기 초부터 수문통계학 분야에서 관심의 대상이 되어 왔다. 분석대상으로는 연최대치 계열(Annual Maximum Series)과 부분기간 계열(Patial Duration Series)이 주로 사용되어 왔다. 분포형으로는 극치분포(Extreme Type I 또는 Extreme Type II), 대수정

규분포(Lognormal II), log pearson II, Gamma 분포 등이 주로 사용되고, 각각 분포형의 매개변수들은 주로 모멘트방법으로 추정되어 왔다.

해외에서의 연구는 수문자료계열의 빈도분석을 위한 전통적인 방법에서는 확률분포의 매개변수 결정을 위해 간편하고 용이한 모멘트법을 사용해 왔으나, Hann[12], Hosking[20]과 등은 전통적인방법으로 결정된 확률분포형으로 모의된 자료계열간에는 큰 편차가 항상 존재할 뿐만 아니라 빈도해석의 유연성에 큰 문제가 있음을 지적한 바 있다. 홍수빈도분석을 위한 주요연구를 보면 우선 국외적으로는 Greenwood, Landwehr, Matalas 및 Wallis는 매개변수 유도를 위한 확률가중모멘트를 정의하고 홍수빈도분석에 사용되는 Gumbel분포형을 비롯한 5개 확률분포형의 확률가중 모멘트식을 제시함으로써 L-모멘트법에 의한 5개 분포형의 매개변수 유도를 가능하게 하였다. Landwehr 와 Matalas 및 Wallis는 Gumbel의 확률분포에 있어서 확률가중모멘트에 근거한 매개변수를 추정하고 이를 모멘트법과 최우법에 의해서 구해진 매개변수와 추정된 홍수량의 비교분석을 수행하였으며 이어서 이들은 Wakeby분포에 있어서 매개변수를 확률가중모멘트법에 의해서 구하는 방법을 발표하였다. 또한 Hosking과 Wallis 및 Wood는 확률가중모멘트법을 이용하여 GEV확률분포의 매개변수와 확률홍수량의 추정치를 구하였고 이 추정치들에 대한 편의는 GEV분포의 극치에 대한 말미부분에서의 홍수량 추정시만을 제외하고는 매우 작게 나타났다.

Pilon과 Adamosk는 L-모멘트법을 이용하여 지역 홍수빈도분석을 수행한 결과 캐나다 Nova Scotia지방의 홍수자료는 GEV 확률분포형에 적합하고 L-모멘트를 이용한 GEV 분포모형에 의한 지역홍수빈도분석 결과가 일치점 GEV 분포분석 결과 보다 신빙성이 높은 것으로 발표하였다. 또한 Vogel, McMahon 및 Chiew는 호주 61개 지점의 홍수자료를 대상으로 L-모멘트도를 사용한 홍수빈도분석 결과 GEV 및 Wakeby 확률분포형이 적절한 모형으로 판정되었다고 발표하였다.

국내에서는 김남원, 최한규[3]는 일 강우량 자료인 7개지점의 70년 이상의 자료를 가지고 강우빈도분석시 내.외삽 오차에 관한 연구를 발표하였다.

윤용남, 박무종[2]은 L-Moment법을 이용하여 월 강우량 자료의 지역가뭄빈도해석을 하여 5개 지점의 주요가뭄기간 동안의 개개 소유역의 가뭄지속기간별 최소가뭄우량을 산정한 후 이를 가뭄우량빈도 분석결과와 비교연구하였다. 이와같이 홍수자료의 부족과 홍수빈도분석에 대한 인식 부족 때문에 홍수빈도분석은 실무에서 거의 수행되고 있지 않은 것이 현실이다. 그러므로 우리나라에서 주로 쓰이거나 또는 표준적인 홍수 분포형을 제시하기는 곤란하며, 미계측지역의 자료치의 수가 많지 않을 때 무

작위성을 고려하여 Monte Carlo 방법에 의해 표본 자료를 충분히 모의한 후에 적정분포형에 따른 재현기간별 빈도오차에 대한 연구는 아직 수행하지 않은 것으로 알려졌다.

본 연구에서는 강우빈도시에 자료의 특성으로 인하여 발생하는 오차의 특성을 정량적으로 밝히고자 시도하였다.

공구조물의 수문학적인 설계나 평가 그리고 수자원 계획시에 아주 중요한 것은 홍수량 자료라고 할 수 있는데, 실제로 대부분의 유역은 미계측 유역이기 때문에 홍수량자료를 직접 이용할 수 없으며 따라서 강우-유출시스템의 거동특성이 유사하다는 가정하에 강우자료를 통계분석하여 설계수문량으로 이용하고 있다.

현재 국내 수자원 실무에서는 미계측 유역의 홍수량을 추정하기 위하여 강우자료를 빈도분석하여 이를 강우-유출모형의 입력자료로 이용하고 있다. 강우자료의 기록기간은 대체로 30년 정도이며 분석 목적에 따라서 재현기간 2년, 5년, 10년 등의 내삽하는 경우와 50년, 100년, 200년 등 자료기간 밖으로 외삽하는 경우가 있다. 그러나 실제로 빈도분석한 강우량이 어느 정도의 오차를 가지고 있는지, 또는 어느 정도까지 외삽할 수 있는지 등의 분석결과와 거동특성이 충분히 알려진 바 없기 때문에 자료에 적합한 분포가 모분포라 가정하여 임의적으로 빈도분석하여 강우유출모형의 입력자료로 이용하고 있으며, 이 결과의 평가 또한 아주 다양한 의견이 개진되고 있는 것이 현실이라고 할 수 있다.

본 연구에서는 이러한 현실을 감안하여 강우빈도분석시 발생하는 여러 특성치중 자료 기록기간에 의한 차이로 발생하는 오차정도와 이때 수반되는 확률분포와의 관계를 밝히고자 시도하였다. 이들 자료기록기간에 따른 오차정도를 분석하는 기법에는 실제 샘플자료를 이용 방법 김남원, 최한규[3] 등이 있을 수 있으나, 표본자료의 수가 적음으로 인하여 무작위 변량특성을 충분히 나타낼 수 없는 단점이 있다.

따라서 본 연구에서는 일 강우량 자료 무작위성을 고려하여 Monte Carlo 방법에 의해 표본자료를 충분히 모의한 후에 자료기록기간의 형태로 split 하여 빈도분석하여 얻어지는 결과치와 가정된 모분포치와 비교하는 방법을 채택하였다.

2 분석방법

강우자료의 모 분포를 적절히 가정하여야 모의기법에 의한 표본을 추출할 수 있다. 여기에서 분석하고자 하는 것은 무작위 강우계열을 이용하여 빈도분석할 경우 발생하는 오차정도를 파악하기 위한 것이다. 따라서 이러한 무작위 계열의 모의발생을 위해서는 가상의 통계량을 설정하는 경우와 실제 통

계량을 설정하는 경우로 나눌 수 있다. 본 연구에서는 모의발생을 위하여 후자의 경우인 실제 통계량 사용법을 이용하기로 하였다. 이를 위해 기상청 산하 춘천 측후소의 일 강우량자료를 이용하였고, 이 강우량자료의 통계량을 모분포의 특성치로 가정하기로 하였다.

춘천측후소에서 관측된 자료는 1913년부터 1996년까지 84년간이나 이중 일 최대강우기록은 59년간만 보존되어 있었다. 따라서 59년의 통계치를 이용하여 모분포의 특성을 가정하였다. 김남원과 최한규[3]에서 밝힌바와 같이 자료기록수가 적을 때 고차변수의 확률분포보다 2변수의 확률분포형이 뛰어났다. 따라서 계산의 복잡성을 피하고 김남원, 최한규[3]의 결과를 평가하기 위하여 본 연구의 모분포의 후보 및 검증대상의 샘플분포의 후보로 Log-Normal II 분포, Extreme Type I 분포, Gamma II를 선정하였다. Kolmogorov-Smirnov 검정을 실시한 결과 가장 적합한 분포로 Extreme Type I 분포가 가장 적합한 것으로 나타났다. 따라서 춘천지방의 자료특성을 갖춘 Extreme Type I 분포를 모분포의 빈도분석과 또한 무작위 발생의 분포형으로 설정하였다. 자료의 무작위 계열검정결과는 표 1과 같다.

표 1. 무작위성 검토결과

유의 주준	Anderson Test			Tunning point test		
	r(1)	검정 구간	결과	U_c	U_{1-a}	결과
0.025	0.006	[-.31, 27]	○	2.195	2.25	○

Monte Carlo방법에 의한 무작위 변량 10000개를 발생하였으며, 모의 발생한 10000개를 각각 자료수 10개, 20개, 30개, 40개, 50개, 60개씩 순서대로 자료계열을 분리한 후에 표본분포로 선정한 Log-Normal II분포, Extreme Type I 분포, Gamma II분포로 빈도분석을 실시하였고 특정빈도년의 빈도우량과 모분포의 Extreme Type I 방법에 의한 춘천지방의 빈도분석결과치와 비교하였다. 표본 분포와 모분포의 빈도분석결과를 Root Mean Squar Error(RMSE)를 기준으로 비교하였으며 식(1)과 같다.

$$RMSE(\%) = \left[\frac{\sum (Q_{Ti} - Q_T)^2 / M}{M} \right]^{1/2} \times 100 / M \quad (1)$$

여기서, Q_{Ti} 는 가정된 모분포의 재현기간별 강우량이며, Q_T 는 표본 자료에 의한 재현기간별 강우량이다. 또한 M은 표본빈도분석의 총회수이다. 본 연구의 분석을 위해 선택한 재현기간은 2년, 5년, 10년, 20년, 50년, 100년이며 자료의 수는 20~

60개 이다. 참고로 매개변수 추정방법은 업계에서 가장 많이 이용하고 있는 Moment방법을 사용하였다.

3. 결과 분석

3.1 Extreme Type I 을 모분포로 하였을 경우

3.1.1 Extreme Type I 을 표본분포로 하였을 경우

Extreme Type I 을 모분포로 가정하여 Extreme Type I 을 표본분포로 채택하여 RMSE를 산정한 결과는 표3.1 와 같다.

표 3.1 표본분포 Extreme Type I 에 따른 RMSE(%)

재현기간	자료개수에 따른 RMSE(%)				
	20	30	40	50	60
2	10.79	8.82	8.05	6.79	6.35
5	12.17	9.76	9.21	7.85	7.07
10	13.45	10.87	10.14	8.70	7.77
20	14.48	11.77	10.88	9.37	8.33
50	15.53	12.07	11.64	10.06	8.92
100	16.16	13.25	12.09	10.46	9.27

내삽인 2년, 5년, 10년, 20년, 50년인 경우에 2년에서는 자료수가 20개 일 때 10.79%, 60개 일 때 6.35%의 RMSE가 나타났으며 50년에서는 자료의 수가 20개 일때는 15.53%, 60개 일때는 8.92%이었으며, 자료의 개수가 60개일 때 다른 분포보다 RMSE가 현격히 작게 나타났음을 알 수 있다.

외삽일 경우 100년에서는 자료수가 20개 일때는 16.16%, 60개 일때는 9.27%로 나타났으며, Extreme Type I 분포의 RMSE는 자료의 개수가 증가함에 따라 RMSE는 감소하는 것으로 나타났다. Extreme Type I 을 모분포한 경우 Log-Normal II 분포와 Gamma II 분포 보다 전체적으로 작은 RMSE가 나왔다.

3.1.2 Log-Normal II 을 표본분포로 하였을 경우

Extreme Type I 을 모분포로 가정한 후 Log-Normal II 을 표본분포 채택하여 산정한 결과는 표 3.2와 같다.

표 3.2 표본분포 Log-Normal II 에 따른 RMSE(%)

재현기간	자료개수에 따른 RMSE(%)				
	20	30	40	50	60
2	15.11	14.06	13.56	12.99	12.73
5	12.22	10.55	10.07	9.10	8.89
10	13.66	11.39	10.43	9.17	8.78
20	17.28	14.55	12.94	11.56	10.94
50	23.76	20.38	17.84	16.37	15.45
100	28.16	25.42	22.09	20.53	19.58

내삽인 2년, 5년, 10년, 20년, 50년인 경우 2년에서는 자료수가 20개 일 때 RMSE가 15.11%, 60개 일 때 12.73%의 RMSE가 나타났으며, 재현기간이 5년과 10년이 RMSE보다 3%정도 높게 나타났으며 이는 난수발생시 우연오차의 발생으로 생각된다. 5년일 경우 자료개수 20개 일때는 12.55%이며 60개 일때는 8.89%로 나타났다. 50년 일 경우 자료의 수가 20개 일때는 23.76%, 60개 일때는 15.45%의 RMSE가 나타났다. 자료수의 증가함에 따라 재현기간에 따른 RMSE가 감소하는 것으로 나타났으나 자료의 수에 따른 RMSE는 최소 8.89%에서 최대 23.76%로 거동 폭은 작은 것으로 나타났다.

외삽일 경우 100년에서는 자료의 개수에 관계없이 RMSE가 크게 나타나는 것을 알 수 있다. 자료수가 20개 일때는 28.16%, 60개 일때는 19.38%로 나타났다. 또한 외삽일 때 다른 분포보다 RMSE가 크게 나온결과 표본분포로 적합하지 않는 것을 알 수 있다.

3.1.3 Gamma II 을 표본분포로 하였을 경우

Extreme Type I 을 모분포로 가정한 후 Gamma II 을 표본분포로 채택하여 산정한 결과는 표 3.3과 같다.

표 3.3 표본분포 Gamma II 에 따른 RMSE(%)

재현기간	자료개수에 따른 RMSE(%)				
	20	30	40	50	60
2	12.20	11.36	7.42	5.83	5.81
5	12.44	11.97	7.53	6.25	6.92
10	13.81	13.58	9.15	8.20	8.47
20	15.04	14.98	10.52	9.84	9.905
50	16.36	16.46	11.97	11.62	11.61
100	17.20	17.23	12.88	12.78	12.79

내삽인 2년과 5년에서는 자료의 수가 20개일 때 RMSE가 각각 12.20%와 12.44%였으며, 자료의 개수가 50개 일때는 RMSE가 5.83%과 6.25%이었다. 또한 60개 일때는 RMSE가 각각 5.81%와 6.92%이

며 이는 다른 확률표본분포보다 내삽에서 가장 작은 RMSE를 나타내며 Extreme Type I를 모분포로 할 경우 Extreme Type I를 표본분포로 하는 것이 가장 적절한 것으로 판단된다. 외삽인 100년 일 때의 RMSE는 자료의 개수 20개 일 때 17.20%로 나타났고 자료의 개수가 60개 일 때는 12.79%의 RMSE가 나타났다.

그림 3.1 재현기간에 따른 RMSE(%)분석 (Extreme Type I)

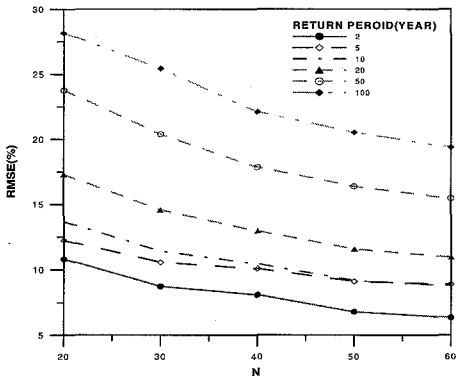
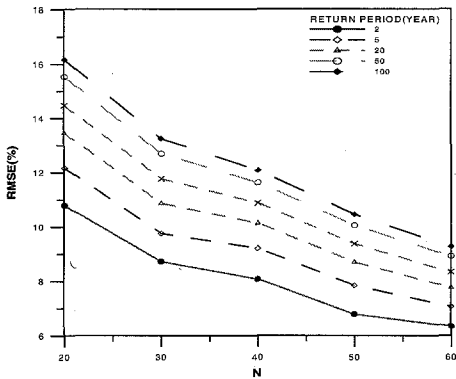


그림 3.2 재현기간에 따른 RMSE(%)분석 (Log-Normal II)

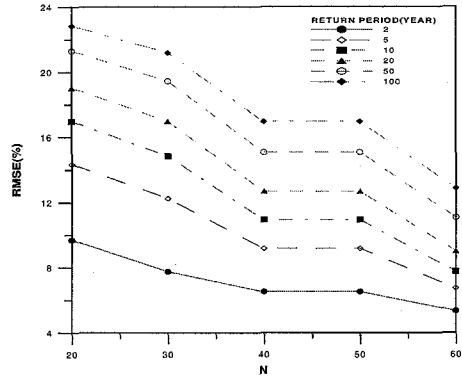


그림 3.2 재현기간에 따른 RMSE(%)분석 (Gamma II)

3.2 Log-Normal II을 모분포로 하였을 경우

3.2.1 Extreme Type I을 표본분포로 하였을 경우

Log-Normal II를 모분포로 가정하여 Extreme Type를 표본분포로 채택하여 RMSE를 산정한 결과는 표 3.4 와 같다

표 3.4 표본분포 Extreme Type I에 따른 RMSE(%)

재현기간	자료개수에 따른 RMSE(%)				
	20	30	40	50	60
2	10.35	8.72	7.32	6.85	6.39
5	13.17	11.19	9.62	8.74	8.20
10	15.00	12.75	11.14	10.05	9.46
20	16.46	14.02	12.40	11.91	10.56
50	18.03	15.44	13.83	12.53	11.89
100	19.07	16.42	14.83	13.50	12.87

내삽인 경우 2년, 5년, 10년, 20년, 50년인 경우 2년에서는 자료수와 상관없이 가장 작은 RMSE가 나타났다. 자료수의 증가함에 따라 재현기간에 따른 RMSE가 감소하는 것으로 나타났으나 자료의 수에 따른 RMSE는 최소 6.39%에서 최대 18.03%로 거동 폭은 작은 것으로 나타났다. 외삽일 경우 100년에서는 자료수가 20개 일 때는 19.07%, 60개 일 때는 12.87%로 나타났으며, Extreme Type I 분포와 Lognormal II 분포 그리고 Gamma II 분포를 표본분포 한 결과 Extreme Type I 분포로 할 경우가 가장 RMSE가 적게 나타났다.

3.2.2 Log-Normal II을 표본분포로 하였을 경우
Lognormal II를 모분포로 가정하여 Lognormal II를 표본분포로 채택하여 RMSE를 산정한

결과는 표 3.5와 같다.

표 3.5 표본분포 Log-Normal II에 따른 RMSE(%)

재현기간	자료개수에 따른 RMSE(%)				
	20	30	40	50	60
2	11.42	10.32	9.55	9.34	8.97
5	12.66	11.33	10.27	9.87	9.65
10	14.07	12.34	11.06	10.48	10.26
20	15.57	13.41	11.92	11.18	10.93
50	17.53	14.81	13.07	12.11	11.80
100	19.98	15.85	13.93	12.82	12.46

내삽인 경우 2년, 5년, 10년, 20년, 50년이며 재현기간이 2년일 경우에는 자료수가 20개 일때는 11.42% 이 RMSE가 나타났고, 60개 일때는 8.97%이며 재현기간이 증가할수록 오차의 거동이 증가하는 것을 알 수 있으며 자료의 개수의 증가함에 따라 RMSE는 감소하여 재현기간이 50년 일때는 자료개수가 20개일 때 17.53%이며, 60개 일때는 11.80%의 RMSE의 거동이 일어났다.

외삽인 경우의 100년에서는 자료의 개수가 19.88%~12.46%의 RMSE로 나타났다.

3.2.3 Gamma II를 표본분포로 하였을 경우

Log-normal II를 모분포로 가정하여 Log-normal II를 표본분포로 채택하여 RMSE를 산정한 결과는 표 4.6와 같다.

표 4.6 표본분포 Gamma II에 따른 RMSE(%)

재현기간	자료개수에 따른 RMSE(%)				
	20	30	40	50	60
2	9.69	7.74	6.51	6.10	5.23
5	14.34	12.24	9.16	7.14	6.72
10	16.96	14.84	10.34	8.10	7.73
20	19.02	16.97	12.68	9.51	8.98
50	21.30	19.44	15.07	11.89	11.16
100	22.84	21.18	16.93	13.91	12.87

내삽인 경우 2년, 5년, 10년, 20년, 50년이며 재현기간이 2년일 경우에는 자료수가 20개 일때는 9.69%의 RMSE가 나타났으며, 재현기간별로 가장 작은 RMSE를 나타내고 있다. 60개 일때는 5.23%이며 재현기간이 증가할수록 오차의 거동이 증가하는 것을 알 수 있으며 자료의 개수의 증가함에 따라 RMSE는 감소하여 재현기간이 50년 일때는 자료개수가 20개일 때 21.03%이며, 60개 일때는 11.16%의 RMSE의 거동이 일어났다.

외삽인 경우의 100년에서는 자료의 개수가 22.84%~12.87%의 RMSE로 나타났다. 자료의 개수에 상관없이 Extreme Type I분포와, Lognormal II분포의 외삽보다 큰 RMSE가 나타나 이 결과로 외삽에서는 표본분포에 적합지 않은 것으로 나타났다.

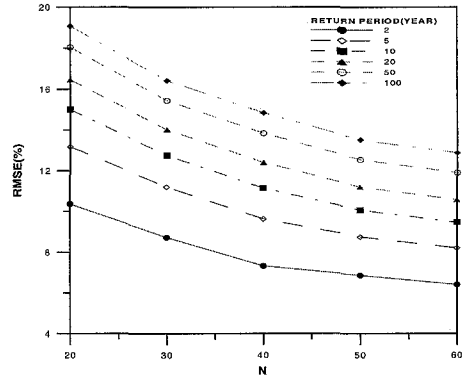


그림 3.4 재현기간에 따른 RMSE(%)분석 (Extreme Type I)

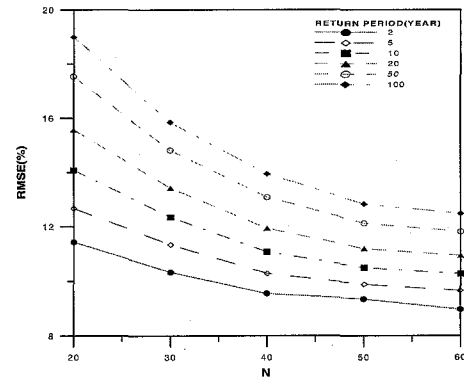


그림 3.5 재현기간에 따른 RMSE(%)분석 (Log-Normal II)

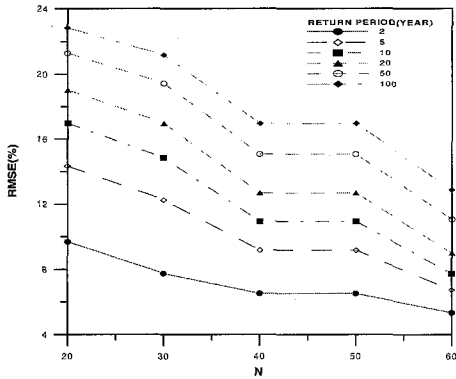


그림 3.6 재현기간에 따른 RMSE(%)분석 (Gamma II)

3.3 Gamma II 을 모분포로 하였을 경우

3.3.1 Extreme Type I 을 표본분포로 하였을 경우

Gamma II 를 모분포로 가정하여 Extreme Type I 을 표본분포로 채택하여 RMSE를 산정한 결과는 표 3.7와 같다.

표 3.7 표본분포 Extreme Type I 에 따른 RMSE(%)

재현기간	자료개수에 따른 RMSE(%)				
	20	30	40	50	60
2	8.24	6.95	6.12	5.69	5.24
5	18.94	16.15	14.53	13.59	12.06
10	23.42	20.05	18.08	16.86	14.97
20	26.56	22.77	20.56	19.14	16.99
50	29.66	25.47	23.03	21.41	19.04
100	30.05	27.12	24.55	22.82	20.33

내삽인 2년, 5년, 10년, 20년, 50년인 경우 2년에서는 자료수가 20개 일 때 RMSE가 8.24%, 60개 일 때 5.24%의 RMSE가 나타났으며, 5년 일 경우 자료의 수가 20개 일때는 18.94%, 60개 일때는 12.06%의 RMSE가 나타났으며, 재현기간이 2년에서 5년으로 증가할수록 오차의 범위가 큰 변화의 폭으로 증가됨을 알 수 있다. 자료수의 증가함에 따라 재현기간에 따른 RMSE가 감소하는 것으로 나타났으나 자료의 수에 따른 RMSE는 최소 5.24%에서 최대 29.66%로 거동 폭은 큰 것으로 나타났다. 외삽일 경우 100년에서는 자료수가 20개 일때는 30.05%, 60개 일때는 20.33%로 나타났다. 또한 외삽에서는 자료의 개수에 상관없이 가장 큰 RMSE를 나타내고 있다.

3.3.2 Log-Normal II 을 표본분포로 하였을 경우

Gamma II 를 모분포로 가정하여 Log-normal II 분포를 표본분포로 채택하여 RMSE를 산정한 결과는 표 3.7과 같다.

표 3.8 표본분포 Log-Normal II 에 따른 RMSE(%)

재현기간	자료개수에 따른 RMSE(%)				
	20	30	40	50	60
2	8.47	7.86	7.55	7.38	7.15
5	18.76	17.81	17.23	16.98	16.66
10	22.62	21.36	20.58	20.24	19.83
20	25.23	23.65	22.67	22.24	21.73
50	27.72	25.72	24.45	23.90	23.25
100	29.24	26.88	25.38	24.73	23.98

외삽인 경우 재현기간이 2년에는 자료수가 20개일 때 RMSE 8.47%이며, 60개일때는 7.15%의 RMSE로 나타났다.

이 Log-Normal II 분포를 표본분포로 하여 비교한 것이 다른 분포에서 비교한결과 보다 내삽에서 오차가 크게 나오는 것을 볼수 있으며 재현기간이 2년일 때 RMSE가 5년일 때 10%이상의 RMSE가 증가하는 것을 나타내었다.

외삽에서는 자료개수가 20개일 때 30.33%이고 60개일 때는 25.61% RMSE의 거동이 나타났다.

또한 Extreme Type분포와 같이 내삽에서는 재현기간이 2년에서 5년까지의 RMSE의 오차가 다른 재현기간보다 크게 증가하는 것을 알 수가 있다.

3.3.3 Gamma II 를 표본분포로 하였을 경우

Gamma II 를 모분포로 가정하여 Gamma II 를 표본분포로 채택하였으며 모분포와 표본분포가 같은 확률분포로 되어있다. 분석한결과와 표3.9와 같으며 외삽인 경우 재현기간이 2년에는 자료수가 20개일 때 RMSE 2.21%이며, 60개일때는 0.9%의 RMSE로 나타났다.

표 3.9 표본분포 Gamma II 에 따른 RMSE(%)

재현기간	자료개수에 따른 RMSE(%)				
	20	30	40	50	60
2	2.21	1.39	1.23	0.99	0.99
5	2.58	1.85	1.66	1.27	1.07
10	2.66	1.97	1.75	1.37	1.17
20	2.67	2.01	1.77	1.39	1.20
50	2.65	2.02	1.77	1.39	1.21
100	2.63	2.00	1.75	1.38	1.20

50년에는 자료의 개수가 20개 일 때 2.65%이며 자료의 개수가 증가할수록 자료의 개수가 60개 일 때 1.21%까지 감소하는 거동을 볼수가 있다. 외삽인 경우에도 재현기간이 100년인 경우 2.63%에서 1.20%까지의 RMSE의 거동이 보이며 Gamma II를 모분포로 한 경우가 Log-Normal II와 Extreme Type I를 모분포로 한 결과보다 현저히 작은 RMSE의 거동을 보이고 있으면 이는 춘천측 후소의 관측된 값이 빈도분석을 한결과 Gamma II분포가 적합한 것을 판명됨을 알 수가 있다.

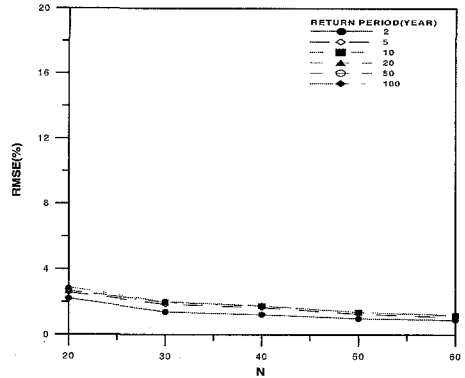


그림 3.9 재현기간에 따른 RMSE(%)분석 (Gamma II)

4. 결론

본 연구를 통해 무작위 변량을 이용한 빈도분석시 내외삽오차를 비교분석한 결과 다음과 같은 결론을 얻었다.

1. 춘천지방의 가우자료의 빈도분석은 Gamma II 분포가 작은 RMSE로 적합한 것으로 나타났다.
2. 강우빈도분석시 무작위 변량을 사용한 내외삽오차는 자료수가 증가할수록 RMSE가 감소하며, 재현기간이 증가할수록 RMSE가 증가하는 것으로 나타났다. 재현기간과 자료수 따라서 크게 변화하며 RMSE는 실제 설계시 사용되는 재현기간별 강우량이 실제강우량보다 어느 정도의 크기로 과대 또는 과소 평가될 수 있는지 보여준다.
3. Extreme Type I의 모분포로 가정하여 산정한 결과 Log-Normal II 분포가 외삽에서 자료수가 20개일 때 29.16%와 60개일 때 19.39%로 다른 분포보다 높은 RMSE의 거동이 나타났다. Extreme Type I 표본분포가 자료개수에 따른 RMSE가 가장 낮은 것으로 나타났으며, Log-Normal II 분포와 Gamma II 분포 보다 전체적으로 작은 RMSE가 나왔다.
4. Log-Normal II의 모분포로 산정한 결과가 재현기간이 2년인 경우 Gamma II 분포가 9.69%~5.22%로 가장 적합하며 Extreme Type I 분포도 비슷한 거동을 보이고 있다. 외삽일 경우 Gamma II 분포가 22.89%~12.97%로 크게 나왔다.

본 연구에서는 2변수 확률분포에 국한되었고 모분포도 Extreme Type I의 한가지 방법만을 선정하

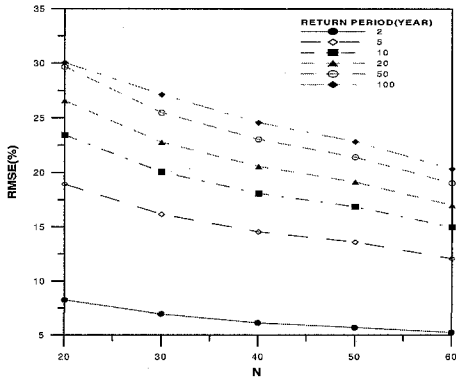


그림 3.7 재현기간에 따른 RMSE(%)분석 (Extreme Type I)

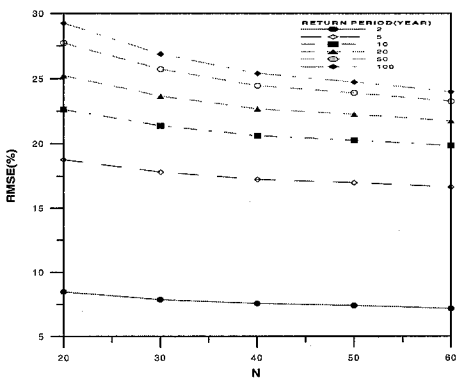


그림 3.8 재현기간에 따른 RMSE(%)분석 (Log-normal II)

여 분석하였으므로 향후 연구과제로 여러 종류의 모분포 선정에 따른 3변수 확률분포에 대한 연구가 이루어져야 할 것으로 생각된다. 또한 매개변수 추정에 있어서 전통적인 방법인 Moment방법 한가지만을 선정하였고 그 밖의 방법인 L-Moment법, 최우도법, 확률가중모멘트등의 방법을 이용한 연구가 필요하다.

참 고 문 헌

- [1] 김남원, 최한규, “홍수 빈도분석상 일 유량자료 이용과 유용성”, 수자원학회 학술발표논문집 '96 pp250~255
- [2] 윤용남, 박무중, “L-Moment법을 이용한 월 강우량 자료의 빈도분석”, 수자원학회 학술발표 논문집 '96 pp259~266
- [3] 김남원, 최한규, “강우빈도해석시 내.외삽 오차에 관한 고찰”, 대한토목학회 학술발표 논문집 (II) pp29~32
- [4] 류경식, Weibull-3 및 GEV분포모형의 L-모멘트법에 의한 홍수빈도비교분석, 충북대학교 대학원, 1990
- [5] 이순혁, 홍성표, 박명근, Low Outliers 를 고려한 홍수빈도분석에 관한 연구, 한국농공학회지, Vol. 30, No. 4, pp. 62 ~ 70, 1988.
- [6]. 박창남, 확률홍수량 산정을 위한 지역빈도분석, 한양대학교 산업대학원, 1991
- [7] 윤태훈, “지역화한 홍수빈도분석”, 대한토목학회지 제 21 권 제 3 호, 1973
- [8]. 김형문, “서부경남지역 최대일강우량 빈도분석에 관하여”, 경상대학교 대학원, 1988
- [9] 윤용남, 공업수문학, 청문각, 1999
- [10] 최석범, 수자원공학, 창원, 1999
- [11] 허준행, 김경덕, 조원철 “연최대 강우자료의 빈도해석을 통한 지역적 고찰에 관한 연구”, 수자원학회논문집 '96 pp199~204
- [12] Haan, Charles Thomas (1979) Statistical method in hydrology, Iowa state University
- [13]. Elderton, W.P. 1963. Frequence Curve and Correlation. Fourth, Harren
- [14]. Salas, J. D., Smith, R. A. Tabios, G. and Heo, Jun-Haeng, Statistical Computer Techniques in Hydrology and Water Resources, Department of Civil Engineering Colorado State University, 1993.
- [15] Yevjevich, V., Probability and Statistics in Hydrology, W. R. P., Colorado, pp.118 ~ 167, 1972.
- [16] Dalrymple, T., "Flood Frequence Analyses", USGS Water Supply Paper, 1960
- [17] Benson, M. A., "Characteristic of Frequence Curve Based on a Thoretical 1000-Years Report", USGS Water Supply Papper, 1960
- [18] Vogel, R. M. and N.M. Fennessey, L-Moment Diagrams should replace Product Moment Diagrams, Water Resource Research, Vol.29, No.6, pp. 1745 ~ 1752, 1993.
- [19] Chow, V. T. "A general formulas for hydrologic frequence analysys", Transaction, American Geophysical Union, Vol, 32, pp 23 1~237.
- [20] Hosking, J.R.M. "Fortran routines for use with the method of L-moment" Reserch Report RC-17097, IBM T. J. Watson Reserch Center, Yorktown Height, New York