회귀분석 방법을 이용한 서울지점 극치 수문량 산정

Estimates Extreme Hydrologic Event at Seoul Using Regression Analyses

김광섭* · 이기춘**

Kim, Gwangseob · Lee, Gichun

Abstract

In this study, we estimated probable precipitation amounts at Seoul for the target year(2020, 2030, 2040 year) using the 1 and 24 hour annual maximum precipitation data from 1973 through 2009. The Gumbel distribution was chosen and the probability weighted moment method was used to estimate model parameters. The behavior of the mean of extreme precipitation data, scale parameter, and location parameter amount using the linear regression and the logistic regression methods. The probable precipitation amount for the target years using the linear regression methods is much higher than that using a stationary frequency analysis. The probable precipitation amount using the logistic regression showed stable increase but probable precipitation amounts was almost same for different target years since the behavior of logistic curve converges before 2040 year. In estimating probable precipitation, the logistic regression is able to reflect the increase behavior of hydrologic extreme reasonably while probable precipitation amounts for target years using a linear regression method have continuous increase.

Key words : Probable precipitation amount, Linear regression, Logistic regression

요 지

본 연구에서는 서울지점의 목표연도별(2020, 2030, 2040년) 재현기간에 따른 확률강수량을 산정하기 위하여 1973년부터 2009년까지 지속시간 1시간 및 24시간에 대한 연 최대 강수량 자료를 사용하였다. Gumbel 분포를 이용해 빈도해석을 실시하였으며, 확률가중모멘트법을 사용하여 분포형의 위치매개변수와 척도매개변수를 산정하였다. 서울지점의 누적평균 극치강수량에 대하여 선형회귀분석을 실시하는 방법과 한강유역에 속하는 14개 지점들의 누적평균 극치강수량에 대하여 Logistic 회귀분석을 실시하는 방법을 이용해 각 방법에 따른 확률강수량을 정상성임을 가정하여 산정된 재현기간에 따른 확률강수량과 비교하였다. 선형회귀분석에 비해 Logistic 회귀분석의 결정계수가 더욱 높은 것으로 나타났고, 확률강수량의 경우, 선형회귀분석을 통한 확률강수량은 정상성에 비하여 매우 높게 나타남을 보였으나, Logistic 회귀분석에서 최대편차를 이용해 산정된 확률강수량의 경우 선형회귀분석에 비해 비교적 안정적인 증가를 보였다. 확률강수량 산정에 있어 정상성 가정에 기초한 전형적인 빈도해석과 비교하였을 때, Logistic 회귀분석은 극치의 선형 증가 거동을 반영하면서도 선형회귀분석을 사용한 목표연도별 확률강수량 산정시 나타나는 지속적 선형 증가의 문제를 보완할 수 있을 것으로 판단된다.

핵심용어 : 확률강수량, 선형회귀분석, Logistic 회귀분석

1. 서 론

지구 온난화로 인해 발생된 기후변화의 영향으로 수공구조물 및 홍수방어시스템을 통한 홍수방어능력을 초과하는 강우를 발생 시키는 등 홍수피해를 증가시키고 있다. 2001년부터 2010년까지 10년 동안 평균 68명의 인명피해와 1조 7,044억 원의 재산피해가 발생하였으며, 2010년의 피해발생을 월일별로 구분하였을 때, 피해액을 기준으로 호우에 의해 1,808억 원으로 약 42%의 피해를 입었다(소방방재청, 2011). 또한 우리나라뿐만 아니라 세계적으로 홍수와 가뭄 같은 이상기후현

상의 빈도와 강도가 증가하는 지역이 많이 보고되고 있다. 파키스탄에서 1929년 이래 최악의 집중호우로 인해 2,000만 명의 이재민이 발생하였으며, 홍수로 국토의 20%가 침수하는 등 150억 달러의 피해가 발생하였다. 중국에서는 6월에 100년만의 폭우로 381명이 사망하고 838억 위안(약 15조 원)의 피해가 발생하였으며, 8월의 홍수로 인해 산사태로 1400명 이상이 사망하는 등 세계적으로 이상기후로 인한 인명 피해 및 경제적 피해가 증가할 뿐만 아니라 자연재해 발생과 피해 규모가 최근 60년 동안 증가하는 추세로 나타났다(기상청, 2010).

*정회원 · 경북대학교 건축토목공학부 부교수(E-mail : kings@knu.ac.kr)

**정회원 · 경북대학교 건축토목공학부 석사과정(교신저자)

일반적으로 수자원관리를 위한 시스템을 계획하고 운영함에 있어 강수 극치자료의 정상성 가정에 기초하지만, 최근 들어 각국들은 강수 극치자료의 변화는 이러한 정상성 가정을 수자원계획과 운영에 있어 더 이상 기본가정으로 받아들이기 힘들다고 인식하고 있다. 비정상성 시계열의 빈도해석에 있어 확률분포의 매개변수가 시간에 따라 변화하기 때문에 기존의 정상성 가정에 따른 빈도해석과는 다른 새로운 접근법이 요구된다. 또한 초과확률이나 관련된 불확실성 추정에도 있어서도 비정상성을 고려할 수 있는 새로운 접근방법이 요구되어지며, 따라서 관측 자료가 독립적이고 정상성을 띠는 기본가정을 벗어나는 요소들은 빈도해석 결과의 타당성에 영향을 미치는 결과를 초래하게 된다. 최근 수문학적 시계열 자료를 모형화하고 특정 재현기간의 극치사상을 추정하기 위해 확률분포를 관측 자료에 적합 시킬 때, 시계열자료의 비정상성 측면을 고려한 여러 가지 기법들이 검토되어지고 있다.

수자원의 계획 및 설계에 있어 비정상성을 적용하고자 하는 연구들 중 국외 연구로, Renard et al.(2006)은 확률모형들과 지수분포형, Generalized Pareto(GP) 분포형, Gumbel 분포형, Generalized Extreme Value(GEV) 분포형 등 극치분포형을 고려하고 Bayesian 기법을 적용하여, 사전분포를 특성화하고 사후분포는 확률모형을 정량화함으로써 보다 현실적인 빈도해석 모형을 도출할 수 있다고 하였다. Ouarda and Adlouni(2008)은 시계열 자료에 대한 경향성을 Bayesian 방법을 통해 매개변수를 산정하였으며, GEV 분포와 GP 분포를 적용하였다. Sugahara et al.(2009)은 Brazil Sao Paulo의 1933-2005년까지 일강수량 극치 빈도해석을 위해 Peaks-Over-Threshold(POT)와 GP 분포형을 적용하였으며, 비모수적 추세검정을 위해 Mann-Kendall 검정을 수행하였다. 국내에서 수행된 비정상성 빈도해석 관련 연구는 권영문과 김태웅(2009)은 서울 강우관측소의 46개년(1961-2006) 시 강우자료를 Gumbel 혼합모형에 적용해 빈도해석을 수행하였으며, 이변량 강우빈도해석을 통해 결합누적분포함수를 선정한 후, 결합 재현기간, 조건부 재현기간을 산정하였다. 이와 같은 이변량 강우빈도해석은 다양한 호우특성들의 확률적 거동에 대한 예측정보를 제공함으로써 수공구조물의 계획 및 설계, 위험도 평가 등의 문제해결에 유용하게 사용될 수 있다고 제안하였다. 이정주 등(2010)은 서울 지방의 24시간 임의지속시간의 강수량을 대상으로 하여 POT 자료를 이용해 계절성을 추출하였고, 이를 빈도해석과 연계시켜 Bayesian 기법을 기반으로 하는 비정상성 빈도해석 모형을 구축하였다. 장선우 등(2011)은 연최대강우량의 회귀직선에 대한 잔차의 수문학적 빈도해석을 바탕으로 하여, 가까운 미래로 설정된 목표연도의 확률강수량을 산정하는 방법을 제안하였다.

본 연구는 서울지점에 대해 선형회귀분석과 Logistic 회귀분석을 이용하여 비정상성 빈도해석을 수행하고 목표연도별 확률강수량을 산정하였다. 서울지점의 누적평균 극치강수량과 연도와의 선형회귀분석을 수행하는 방법과 한강유역의 속하는 지점들의 누적평균 극치강수량을 전체 평균하여 산정된 강수량과 연도와의 Logistic 회귀분석을 수행하는 방법을 이

용하여, 두 가지 방법을 이용한 비정상성 빈도해석의 결과들을 비교하였다.

2. 연구방법

2.1 분석 자료

서울지점의 목표연도(2020, 2030, 2040년)별 비정상성 빈도해석을 수행하기 위하여 한강유역에 속하는 14개 기상관측지점의 1973년부터 2009년까지 시강수량 자료를 구축하였다. 수집된 관측 자료를 통해 지속시간 1시간과 24시간에 대한 지속강우를 계산하였으며, 각 지속시간별 강수량의 연최고치를 산정하여 본 연구 자료로 사용하였다. 또한 지속시간별로 초기 20년을 기준으로 1년씩 추가하여 누적강수량을 구한 후 평균하여, 각 기간별 연 최대 누적평균 강수량을 산정하였다.

그림 1. 한강유역 14개 지점

2.2 확률분포함수

본 연구에서는 Gumbel 분포를 사용하여 서울지점의 비정상성 빈도해석을 실시하였다. Gumbel 분포는 일반화된 극치분포(Generalized Extreme Value Distribution, GEV)에서 형상매개변수(shape parameter) 일 때를 나타내며, Type-I 극치분포(Extreme Value Type-I, EV1)라 부르기도 한다. 이 분포는 극치자료 중에서도 연 최대 홍수량 및 강수량 자료의 빈도분석에 많이 사용되고 있으며, 확률밀도함수는 다음과 같다.

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

여기서, α 는 축척매개변수(scale parameter)이며, x_0 는 위치매개변수(location parameter)이다. 매개변수의 추정법은 자료에 이상치가 있는 경우에도 왜곡특성이 크게 나타나지 않고 표본의 크기가 작아도 영향을 덜 받으며, 상대적으로 안정적인 장점을 지니고 있어, 최근에 많이 사용되고 있는 확률가중도멘트법을 사용하였다. Gumbel 분포의 위치매개변수와 축척매

개변수를 구하는 식은 다음과 같다.

$$\alpha = \frac{l_2}{\ln 2}, \quad x_0 = l_1 - 0.5772157\alpha \quad (2)$$

여기서, l_1 과 l_2 는 표본자료의 1차 및 2차 L-모멘트이다.

또한 Gumbel 분포의 누적분포함수는 다음과 같다.

$$F(x) = \exp\left[-\exp\left(-\frac{x-x_0}{\alpha}\right)\right] \quad (3)$$

목표연도 재현기간별 확률강수량 산정에 사용된 Gumbel 분포의 누적분포함수의 역함수는 다음과 같다.

$$x_T = x_0 - \alpha \ln[-\ln F(x)] \quad (4)$$

2.3 회귀분석

본 연구에서는 목표연도별 누적평균 극치강수량을 구하기 위하여 선형회귀분석과 Logistic 회귀분석을 이용하여 산정하였으며, 산정된 누적평균 극치강수량을 통해 선형회귀분석을 이용하여 매개변수를 추정하였다. 일반적인 회귀분석 모형에서 독립변수가 $-\infty$ 에서 $+\infty$ 의 범위에 있다면 종속변수도 마찬가지로 $-\infty$ 에서 $+\infty$ 의 범위를 가지게 되는 것에 반해, Logistic 회귀분석은 연속형을 가지는 독립변수에 따른 종속변수의 관계를 0과 1 사이에 확률로 제시한다. 또한 각 독립변수가 종속변수에 미치는 영향의 정도를 보다 정량적으로 표현할 수 있으며, 이러한 Logistic 함수는 다음과 같다(권현환과 문영일, 2007).

$$p(x) = \frac{e^{\delta_0 + \delta_1 x_1 + \dots + \delta_i x_i}}{1 + e^{\delta_0 + \delta_1 x_1 + \dots + \delta_i x_i}}, \quad i = 1, 2, \dots, N \quad (5)$$

여기서 δ 는 회귀계수를 나타내며 x 는 독립변수를, N 은 종속변수를 나타내며, Logistic 함수를 아래와 같이 변형시킴으로써 일반적인 회귀모형과 같은 형태가 나타난다.

$$g(x) = \ln\left[\frac{p(x)}{1-p(x)}\right] = \delta_0 + \delta_1 x_1 + \dots + \delta_i x_i, \quad i = 1, 2, \dots, N \quad (6)$$

3. 자료개요 및 분석결과

3.1 경향성 및 정상성 검정

서울지점의 1973년부터 2009년까지 시 강수 자료를 이용해 지속시간 1시간과 24시간에 대한 연 최대 강수량 자료를 구축하고 경향성과 정상성 검정을 실시하였다. 본 연구에서는 Mann-Kendall, Wald-Wolfowitz, Run test를 이용해 지속시

간 1시간 및 24시간의 연 최대 강수량 자료의 경향성 검정을 실시하였고, 정상성 검정을 위해 Augmented Dickey-Fuller (ADF), Phillips Perron(PP), Kwiatkowski-Phillips-Schmidt-Shin(KPSS) test를 실시하였다. 다음 표 1은 각 검정별 p-value와 경향성의 유무, 정상성과 비정상성 판별 결과를 나타낸다. 지속시간 1시간의 연 최대 강수량의 경우, 경향성 검정들 중 2가지 방법이 경향성이 있는 것으로 나타났으며, 정상성 검정의 경우에는 하나의 검정에서만 비정상성이 나타남을 보였다. 24시간 연 최대 강수량의 경향성 및 정상성 검정에서는 ADF test 만이 비정상성으로 나타남을 보였다.

3.2 자료구축

서울지점의 1973년부터 2009년까지 지속시간 1시간과 24시간에 대한 연최고치 강수량을 구축하였고, 초기 20년을 기준으로 하여 1년씩 강수 자료를 추가하여 나타나는 변화는 최근 변화양상을 가장 잘 표현할 수 있는 것으로 알려져 있기에, 본 연구에서 초기 20년 기준으로 1년씩 강수 자료를 누적시켜 기간으로 평균함으로써, 각 기간별 누적평균 극치강수량 자료를 구축하였다(안재현 등, 2000). 또한 지속시간별 연 최대 강수량을 통하여 Gumbel 분포의 매개변수인 위치매개변수와 축척매개변수를 확률가중모멘트법을 이용하여 산정하였다(표 2). 지속시간 1시간 및 24시간 연 최대 강수량 자료 모두 기간이 증가함에 따라 누적평균 극치강수량이 대체적으로 증가하는 것으로 나타났으며, 위치매개변수 및 축척매개변수 또한 증가하는 경향을 보였다.

표 3은 한강유역에 해당하는 14개 지점의 누적평균 극치강수량 자료를 모두 합산한 후 평균하여 한강유역에 대한 누적평균 극치강수량을 산정하였다. 또한 한강유역과 서울지점의 각 기간에 대한 누적평균 극치강수량 자료를 통하여 편차를 구해, 각 편차에 대한 평균값, 최댓값을 산정하였다. 지속시간 1시간의 경우 한강유역 전체의 평균이 기간이 늘어남에 따라 일반적으로 증가하는 것으로 나타났으며, 서울지점과의 편차도 증가하는 것으로 나타났다. 지속시간 24시간의 연 최대 강수량에서도 한강의 누적평균이 기간이 늘어남에 따라 증가하는 것으로 나타났으나 서울지점과의 편차의 경우, 경향을 보이지 않았다.

3.3 목표연도별 확률강수량 산정

본 연구에서는 서울지점의 목표연도 2020년, 2030년, 2040년에 대한 확률강수량을 산정하기 위해, 목표연도별 누적평균 강수량을 산정하였고, 평균 극치강수량과 위치매개변수 및 축척매개변수의 상관성을 이용하여 목표연도에 해당하는 매개변수를 추정하였다. 목표연도에 해당하는 누적평균 극치강수

표 1. 지속시간 연 최대 강수량별 경향성 및 정상성 검정

	경향성 검정						정상성 검정					
	Kendall		Wald		Run		ADF		PP		KPSS	
1시간	0.069	X	0.024	O	0.030	O	0.216	O	0.010	X	0.050	X
24시간	0.108	X	0.567	X	0.864	X	0.380	O	0.010	X	0.100	X

※ <경향성 검정> O : 경향성 있음, X : 경향성 없음 <정상성 검정> O : 비정상성, X : 정상성

표 2. 지속시간 연최고치 강수량별 누적평균 극치강수량 및 확률분포함수의 매개변수

자료기간	지속시간 1시간에 대한 연 최대 강수량			지속시간 24시간에 대한 연 최대 강수량		
	누적평균	위치매개변수	축척매개변수	누적평균	위치매개변수	축척매개변수
1973-1992	33.27	11.13	38.27	160.12	128.77	54.31
1973-1993	34.01	11.40	38.10	156.97	125.86	53.89
1973-1994	33.95	11.13	37.44	154.13	123.36	53.30
1973-1995	34.24	10.87	38.66	156.76	125.55	54.07
1973-1996	34.19	10.62	38.47	157.25	126.59	53.11
1973-1997	34.17	10.38	38.49	156.09	126.27	51.66
1973-1998	34.77	10.84	38.63	163.99	129.89	59.08
1973-1999	35.25	10.89	38.29	167.61	132.49	60.85
1973-2000	35.45	10.67	40.23	166.01	131.66	59.52
1973-2001	36.16	12.34	40.44	170.98	134.71	62.83
1973-2002	36.53	12.23	39.78	174.82	137.46	64.71
1973-2003	37.05	12.50	39.63	175.08	138.23	63.84
1973-2004	37.02	12.26	39.36	173.75	137.61	62.61
1973-2005	37.26	12.09	39.21	172.48	137.04	61.42
1973-2006	37.34	11.87	39.20	175.88	139.46	63.10
1973-2007	36.83	11.93	39.05	173.27	136.73	63.31
1973-2008	36.63	11.81	39.25	172.85	136.89	62.30
1973-2009	36.62	11.61	39.67	173.31	137.70	61.70

표 3. 지속시간 연 최대 강수량별 누적평균 극치강수량 및 확률분포함수의 매개변수

자료기간	지속시간 1시간에 대한 연최고치 강수량					지속시간 24시간에 대한 연최고치 강수량				
	한강 평균	서울 평균	편차	위치 매개변수	축척 매개변수	한강 평균	서울 평균	편차	위치 매개변수	축척 매개변수
1973-1992	35.04	39.70	4.66	11.13	38.27	157.31	160.12	2.81	128.77	54.31
1973-1993	35.08	40.59	5.51	11.40	38.10	156.84	156.97	0.13	125.86	53.89
1973-1994	35.00	40.37	5.37	11.13	37.44	156.60	154.13	-2.48	123.36	53.30
1973-1995	35.24	40.51	5.27	10.87	38.66	157.65	156.76	-0.90	125.55	54.07
1973-1996	35.20	40.33	5.13	10.62	38.47	157.21	157.25	0.04	126.59	53.11
1973-1997	35.48	40.16	4.68	10.38	38.49	157.64	156.09	-1.55	126.27	51.66
1973-1998	36.16	41.03	4.87	10.84	38.63	160.63	163.99	3.36	129.89	59.08
1973-1999	36.21	41.54	5.33	10.89	38.29	163.21	167.61	4.40	132.49	60.85
1973-2000	36.39	41.61	5.22	10.67	40.23	162.86	166.01	3.15	131.66	59.52
1973-2001	36.65	43.28	6.63	12.34	40.44	162.60	170.98	8.39	134.71	62.83
1973-2002	36.91	43.59	6.68	12.23	39.78	168.42	174.82	6.40	137.46	64.71
1973-2003	37.12	44.26	7.14	12.50	39.63	168.94	175.08	6.14	138.23	63.84
1973-2004	37.12	44.09	6.97	12.26	39.36	168.77	173.75	4.98	137.61	62.61
1973-2005	37.20	44.24	7.04	12.09	39.21	168.33	172.48	4.16	137.04	61.42
1973-2006	37.37	44.19	6.82	11.87	39.20	170.52	175.88	5.36	139.46	63.10
1973-2007	37.20	43.71	6.51	11.93	39.05	169.01	173.27	4.27	136.73	63.31
1973-2008	37.27	43.44	6.17	11.81	39.25	169.25	172.85	3.60	136.89	62.30
1973-2009	37.29	43.32	6.03	11.61	39.67	170.34	173.31	2.97	137.70	61.70

량을 산정에는 선형회귀분석과 Logistic 회귀분석을 이용하였고, 회귀분석 결과를 이용하여 확률분포함수의 매개변수를 추정하고 목표연도별 확률강수량을 산정하여 두 가지 방법에 대한 결과를 비교·분석하였다.

3.3.1 목표연도별 누적평균 극치강수량 산정

그림 2는 서울지점의 지속시간별 누적평균 강수량과 선형

회귀분석을 실시한 결과와 서울이 포함된 한강유역 14개 지점의 누적평균 강수량 자료들을 평균한 강수량에 대해 Logistic 회귀분석을 실시한 결과를 나타내고 있다. 서울지점의 누적평균 강수량과 선형회귀분석을 실시한 결과, 지속시간 1시간과 24시간 모두 증가하는 것으로 나타났다. 한강유역의 누적평균 강수량의 경우, 기간이 지남에 따라 초기 1992년부터 1996년의 평균 강수량에 비해 증가는 하지만 그 이후에

(a)서울 1시간

(b)서울 24시간

(c)한강 1시간

(d)한강 24시간

그림 2. 서울과 한강유역의 누적평균 극치강수량 자료에 대한 선형회귀분석 및 Logistic 회귀분석

표 4. 각 회귀분석에 대한 결정계수

분 류	지속시간 1시간에 대한 연 최대 강수량		지속시간 24시간에 대한 연 최대 강수량	
지 점	서울지점 선형회귀	한강유역 Logistic	한강유역 선형회귀	한강유역 Logistic
결정계수	0.7966	0.9862	0.7842	0.9625

는 점점 수렴하는 형태를 나타내고 있어, 유역의 누적평균 강수량의 거동을 비정상성으로 보기에 한계가 있음을 보였다.

한강유역과 서울지점의 누적평균 강수량에 대하여 선형회귀분석과 Logistic 회귀분석을 실시한 결과를 이용해 결정계수를 산정한 결과, 표 4와 같이 나타났다. 지속시간 1시간 및 24시간의 연 최대 강수량 자료를 이용하여 목표연도 평균 강수량을 산정하는 모형의 적합성을 판별한 결과, 선형회귀분석에 대한 결정계수가 각각 0.7966과 0.7842로 높은 값을 나타냈으나, Logistic 회귀분석이 0.9470과 0.9396으로 더 높은 값을 나타냈다.

각 회귀분석을 통해 목표연도에 따른 평균 강수량을 산정하였다. 선형회귀분석을 통한 목표연도별 평균 강수량은 기간이 증가함에 따라 강수량 또한 계속 증가하는 변동을 보였으나, Logistic 회귀분석의 경우, 유역의 거동이 증가하는 형태를 보이다가 2003년 이후로는 수렴하게 되어 목표연도별 평균 강수량이 동일한 값을 가졌다.

3.3.2 Gumbel 분포함수의 매개변수 추정

그림 3은 서울지점의 누적평균 강수량과 Gumbel 분포의 매개변수 간의 관계를 나타내고 있다. 선형회귀분석과 Logistic 회귀분석을 이용해 산정된 목표연도별 누적평균 강

수량을 이용하여 다음 식과 같이 나타낼 수 있으며, 산정된 매개변수는 표 5와 같다. 표 5에서 최대편차 및 평균편차는 한강유역의 누적평균 강수량과 서울지점의 누적평균 강수량의 편차 중 평균 및 최대편차를 이용해, 한강유역의 Logistic 회귀분석을 통해 서울지점의 강수량에 대한 위치 및 축척매개변수를 각각 나타낸 것이다. 강수량의 변동에 밀접한 연관을 지닌 매개변수이기에 선형회귀분석에 의한 매개변수들은 모두 증가하는 것으로 나타났으나, Logistic 회귀분석에 의한 누적평균 강수량을 이용하여 매개변수를 추정하였을 경우에는 각 목표연도별 강수량이 같기 때문에 매개변수도 마찬가지로 같은 값을 나타내었다.

$$\bar{x}_m = a \times R_m + b \quad (7)$$

$$\alpha_m = c \times R_m + d \quad (8)$$

여기서, $\bar{x}_m$ 은 모든 목표연도(m)의 위치매개변수, α_m 는 목표연도(m)의 축척매개변수를 나타내며, R_m 은 목표연도의 평균 강수량, a와 b는 누적평균 극치강수량과 위치매개변수 간 선형회귀식의 매개변수, c와 d는 누적평균 극치강수량과 축척매개변수 간 선형회귀식의 매개변수이다.

(a) 지속시간 1시간 위치매개변수

(b) 지속시간 1시간 축척매개변수

(c) 지속시간 24시간 위치매개변수

(d) 지속시간 24시간 축척매개변수

그림 3. 각 지속시간별 누적평균 극치강수량과 Gumbel 분포의 매개변수 관계

표 5. 서울지점의 각 회귀분석에 따른 목표연도별 매개변수 추정치

지속시간 1시간에 대한 목표연도별 누적평균 강수량					
매개변수	선형회귀분석			Logistic 회귀분석	
	2020년	2030년	2040년	최대편차	평균편차
위치 매개변수	40.03	42.30	44.58	37.36	36.36
축척 매개변수	13.39	14.39	15.35	12.24	11.80
지속시간 1시간에 대한 목표연도별 누적평균 강수량					
매개변수	선형회귀분석			Logistic 회귀분석	
	2020년	2030년	2040년	최대편차	평균편차
위치 매개변수	149.88	158.76	167.63	140.40	136.78
축척 매개변수	73.29	80.52	87.75	65.58	62.63

3.3.3 목표연도별 재현기간에 따른 확률강수량 산정

서울지점의 지속시간 1시간과 24시간에 대한 연 최대 강수량 자료를 이용해 누적평균 극치강수량 및 위치, 축척매개변수를 산정하였고, 다른 방법으로서 한강유역에 속하는 14개 지점에 대한 지속시간별 연 최대 강수량 자료를 이용해 한강유역에 대한 누적평균 극치강수량 자료를 산정하였다. 산정된 누적평균 극치강수량 자료를 통해 선형회귀분석 및 Logistic 회귀분석을 이용하여 목표연도에 따른 서울지점의 평균 강수량을 산정하여 각 회귀분석에 따른 목표연도별 위치매개변수와 축척매개변수를 구축하였다. 구축된 목표연도별 위치 및 축척매개변수를 통해 서울 지점의 확률강수량을 산정한 결과를 그림 4와 표 6과 같이 나타냈다. 선형회귀분석을 이용해 산정된 확률강수량의 경우, 목표연도가 증가함에 따라 확률강수량이 증가하는 것으로 나타났다. 하지만 장기간의 확률강수량을 산정할 경우, 정상성임을 가정한 확률강수량에 비해 매

우 높아질 것으로 나타났다. 이는 목표연도별 평균 강수량이 선형으로 증가함으로써 매우 높은 평균 강수량이 나타나, 목표연도별 위치 및 축척매개변수가 높은 값을 가지게 되어 확률강수량이 매우 높게 나기 때문이다. Logistic 회귀분석을 이용해 최대편차를 사용할 경우 산정된 확률강수량은 정상성임을 가정하여 산정된 확률강수량에 비해 비교적 안정적인 증가를 보였으며, 평균편차를 사용하였을 경우 정상성과 거의 비슷한 결과를 보였다. 지속시간 1시간 연 최대 강수량을 이용해 정상성임을 가정하여 산정된 확률강수량 대비 증가율은 선형회귀분석의 경우, 2020년에는 최대 최대편차를 이용하였을 때 최대 약 13.09%의 증가를 보였고 2030년에는 최대 20.71%, 2040년에는 약 28.33%의 증가율을 보였다. Logistic 회귀분석을 이용하여 산정된 확률강수량의 증가율은 최대편차를 이용하였을 때 최대 4.13%의 증가율을 보였으며, 평균편차를 이용하였을 때에는 최대 0.78%의 증가를 보였다. 지

(a)

(b)

그림 4. 지속시간별 연 최대 강수량을 이용한 재현기간에 따른 확률강수량

표 6. 서울지점의 각 회귀분석에 따른 목표연도별 확률강수량 및 증가율

지속시간 1시간 연 최대 강수량 자료를 이용한 비정상성 빈도해석								
재현기간(Year)		5	10	20	30	50	100	200
정상성 확률강수량(mm)		54.00	62.70	71.10	75.90	81.90	90.00	98.10
선형회귀분석을 이용한 확률강수량 (mm)	2020	60.11	70.16	79.80	85.35	92.28	101.63	110.94
	증가율(%)	11.32	11.90	12.24	12.45	12.67	12.92	13.09
	2030	63.86	74.65	84.99	90.94	98.38	108.42	118.42
	증가율(%)	18.26	19.05	19.54	19.82	20.13	20.46	20.71
	2040	67.61	79.13	90.18	96.54	104.49	115.21	125.89
	증가율(%)	25.20	26.20	26.84	27.19	27.58	28.01	28.33
Logistic 회귀분석을 이용한 확률강수량 (mm)	최대편차	55.71	64.89	73.70	78.76	85.10	93.64	102.15
	증가율(%)	3.16	3.49	3.65	3.77	3.91	4.05	4.13
	평균편차	54.06	62.92	71.41	76.30	82.41	90.65	98.87
	증가율(%)	0.11	0.35	0.44	0.53	0.63	0.73	0.78
지속시간 24시간 연 최대 강수량 자료를 이용한 비정상성 빈도해석								
재현기간(Year)		5	10	20	30	50	100	200
정상성 확률강수량(mm)		230.20	276.50	321.00	346.50	378.50	421.50	464.50
선형회귀분석을 이용한 확률강수량 (mm)	2020	259.82	314.82	367.58	397.93	435.87	487.05	538.04
	증가율(%)	12.87	13.86	14.51	14.84	15.16	15.55	15.83
	2030	279.54	339.96	397.92	431.27	472.95	529.17	585.19
	증가율(%)	21.43	22.95	23.96	24.46	24.95	25.54	25.98
	2040	299.25	365.10	428.27	464.60	510.03	571.30	632.34
	증가율(%)	30.00	32.04	33.42	34.08	34.75	35.54	36.13
Logistic 회귀분석을 이용한 확률강수량 (mm)	최대편차	238.77	287.98	335.19	362.34	396.29	442.08	487.70
	증가율(%)	3.72	4.15	4.42	4.57	4.70	4.88	4.99
	평균편차	230.73	277.73	322.82	348.75	381.17	424.91	468.48
	증가율(%)	0.23	0.45	0.57	0.65	0.71	0.81	0.86

속시간 24시간 연 최대 강수량의 경우에도 선형회귀분석을 이용한 확률강수량의 경우 높은 증가율을 보였으나, Logistic 회귀분석을 이용해 산정한 확률강수량은 선형회귀분석을 이용한 확률강수량에 비해 적은 증가율을 보였다.

4. 결 론

본 연구는 서울지점의 목표연도별(2020, 2030, 2040년) 재현기간에 따른 확률강수량을 산정하기 위하여, 서울뿐만 아니

라 한강유역에 속하는 전 지점들의 지속시간 1시간과 24시간에 대한 연 최대 강수량 자료를 구축하였다. 구축된 연 최대 강수량 자료에 대해 20년을 기준으로 1년씩 강수 자료를 누적시켜 기간으로 평균함으로써, 각 기간별 누적평균 강수량 자료와 위치 및 축척매개변수를 산정하였다. 서울지점의 누적평균 강수량과 연도간의 선형회귀분석을 실시한 방법과 한강유역 14개 지점의 누적평균 극치 강수량을 평균한 값에 대한 Logistic 회귀분석을 실시한 방법을 이용해 확률강수량을 산정하여 두 방법을 비교하였다.

서울지점의 누적평균 강수량에 대한 선형회귀분석에 비해 한강유역의 누적평균 강수량에 대한 Logistic 회귀분석을 실시하였을 때의 상관계수가 더욱 높게 나타남을 보였다. 선형 회귀분석은 기간이 늘어남에 따라 단기간의 경우 큰 증가를 보이지 않으나, 계속 증가하기 때문에 장기간에 대한 평균 강수량의 산정에 있어 매우 높은 수치가 나타나, 이로 인해 위치 및 축척매개변수 또한 매우 높은 값을 보였다. 따라서 목표연도별 재현기간에 따른 확률강수량을 산정하였을 때, 정상성임을 가정하여 산정된 확률강수량에 비해 단기간의 경우에도 약 20~30% 정도의 증가율을 보였으며, 장기간의 경우에는 그 이상의 증가율을 보일 것이기 때문에 선형회귀분석에 따른 비정상성 빈도해석에 있어 한계가 있음을 보였다. 한강유역의 누적평균 강수량의 경우 기간이 늘어남에 따라 증가하는 형태를 나타내다가 수렴하는 형태로 나타나면서 목표연도별 평균 강수량이 수렴한 값을 가지기 때문에 정상성임을 가정한 평균 강수량과 동일한 형태를 보였다. Logistic 회귀분석의 경우, 최대편차를 이용하여 산정된 재현기간에 따른 확률강수량의 경우에는 선형회귀분석을 통한 확률강수량에 비해 매우 안정적인 증가를 보였으며, 평균편차를 이용하였을 때는 정상성임을 가정하여 산정된 확률강수량과 비슷한 결과를 나타냈다. 따라서 확률강수량 산정에 있어 정상성 가정에 기초한 전형적인 빈도해석과 비교하였을 때, Logistic 회귀분석은 극치의 선형 증가거동을 반영하면서도 선형회귀 분석을 적용한 확률강수량의 변화가 지속적 선형증가의 문제점을 보완할 수 있는 새로운 방법이라 판단된다.

감사의 글

본 연구는 한국건설교통기술평가원 건설기술혁신사업의 '기후변화에 의한 수문 영향분석과 전망(09-기술혁신 C01)' 과제에 의해 지원되었습니다.

- 권영문, 김태웅 (2009) 이변량 강우 빈도해석을 이용한 서울지역 I-D-F 곡선 유도, **대한토목학회논문집**, 대한토목학회, 제29권, 제2B호, pp. 155-162.
- 권현한, 문영일 (2007) 기상정보 및 태풍특성을 고려한 계절 강수량의 확률론적 모형 구축, **대한토목학회논문집**, 대한토목학회, 제27권, 제1B호, pp. 45-52.
- 기상청 (2010) **이상기후 특별보고서**, 기상청.
- 소방방재청 (2011) **재해연보 2010**, 소방방재청.
- 안재현, 김태웅, 유철상, 윤용남 (2000) 자료기간 증가에 따른 확률강수량의 거동특성 분석, **한국수자원학회논문집**, 한국수자원학회, 제33권, 제5호, pp. 569-580.
- 이정주, 권현한, 황규남 (2010) 극치수문자료의 계절성 분석 개념 및 비정상성 빈도해석을 이용한 확률강수량 해석, **한국수자원학회논문집**, 한국수자원학회, 제43권, 제8호, pp. 733-745.
- 장선우, 서린, 김태웅, 안재현 (2011) 잔차시계열 분석을 통한 비정상성 강우빈도해석, **대한토목학회논문집**, 대한토목학회, 제31권, 제5B호, pp. 449-457.
- Ourada, T.M.B.J. and El Adlouni, S. (2008) Bayesian inference of non-stationary flood frequency models, *World Environmental and Water Resources Congress 2008*, Honolulu.
- Renard, B., Lang, M., and Bois, P. (2006) Statistical analysis of extreme events in a non-stationary context via a Bayesian framework: case study with peak-over-threshold data, *Stoch Environ Res Risk Assess*, Vol. 21, pp. 97-112.
- Sugahara, S., Rocha, R.P.da., and Silverira, R. (2009) Non-stationary frequency analysis of extreme daily rainfall in Sao Paulo, Brazil, *International Journal of Climatology*, Vol. 29, pp. 1339-1349.

© 논문접수일 : 2012년 04월 19일
 © 심사의뢰일 : 2012년 04월 20일
 © 심사완료일 : 2012년 04월 30일