

Quantile 회귀분석을 이용한 극대강수량 자료의 경향성 분석

Trend Analysis of Extreme Precipitation Using Quantile Regression

소 병 진* / 권 현 한** / 안 정 희***

So, Byung-Jin / Kwon, Hyun-Han / An, Jung Hee

Abstract

The underestimating trend using existing ordinary regression (OR) based trend analysis has been a well-known problem. The existing OR method based on least squares approximate the conditional mean of the response variable given certain values of the time t , and the usual assumption of the OR method is normality, that is the distribution of data are not dissimilar form a normal distribution. In this regard, this study proposed a quantile regression that aims at estimating either the conditional median or other quantiles of the response variable. This study assess trend in annual daily maximum rainfall series over 64 weather stations through both in OR and QR approach. The QR method indicates that 47 stations out of 67 weather stations are a strong upward trend at 5% significance level while OR method identifies a significant trend only at 13 stations. This is mainly because the OR method is estimating the condition mean of the response variable. Unlike the OR method, the QR method allows us flexibly to detect the trends since the OR is designed to estimate conditional quantiles of the response variable. The proposed QR method can be effectively applied to estimate hydrologic trend for either non-normal data or skewed data.

Keywords : extreme rainfall, trend analysis, quantile regression, linear regression

요 지

기존 Ordinary Regression (OR) 방법을 이용한 경향성 분석은 경향성을 과소평가하는 문제점을 나타낸다. 이러한 점에서 본 연구에서는 자료의 정규분포 가정과 평균을 중심으로 경향성 평가가 이루어지는 기존 Ordinary Regression (OR) 방법을 개선한 Quantile Regression (QR) 방법을 제안하였다. 본 연구에서는 64개 강우 관측지점의 연 최대 극대강수량 자료에 대하여 QR 방법과 OR 방법에 대하여 통계적 성능을 평가하였다. QR 방법의 경향성 분석결과 47개 지점에서 5% 오차수준 내에서 t -검정을 통과한 반면 OR 방법에서는 13개 지점 만이 통계적 유의성을 가지는 것으로 나타났다. 이는 OR 방법이 자료의 평균을 중심으로 경향성을 평가하는 기법인데 반해 QR은 자료의 다양한 분위에서 경향성을 평가함으로써 극대 및 극소 부분에서의 경향성을 보다 유연하게 감지하는 이유로 판단된다. QR 방법을 통한 경향성 평가는 평균 중심의 해석문제점을 개선할 수 있으며 자료가 정규분포를 따르지 않거나 왜곡된 분포형태를 갖는 자료의 수문학적 경향성 평가에 유용하게 사용될 수 있을 것으로 판단된다.

핵심용어 : 극치강수량, 경향성분석, Quantile 회귀분석, 선형 회귀분석

* 전북대학교 대학원 토목공학과 석사과정 (e-mail: so.b.jin@jbnu.ac.kr)

Master Course, Department of Civil Engineering, Chonbuk National University, Jeonju 561-756, Korea

** 교신저자, 전북대학교 공과대학 토목공학과 조교수 (e-mail: hkwon@jbnu.ac.kr)

Corresponding Author, Assistant Professor, Department of Civil Engineering, Chonbuk National University, Jeonju 561-756, Korea

*** (주) 금강건설 과장 (e-mail: anjh2276@hanmail.net)

Master Course, Department of Civil Engineering, Chonbuk National University, Jeonju 561-756, Korea

1. 서론

수문기상학 분야에서 주요 관심사는 수문기상시스템을 특성화할 수 있는 장기간의 경향을 파악하는 것이다. 특히 기후변화 연구에서는 이러한 점을 반영하여 수문사상의 장기적인 변동성과 더불어 집중호우와 같은 극치사상들의 발생빈도를 평가하고 전망하는 것이 주된 관심사항(IPCC, 2007)이며 기상학적 분석에 선행되는 기본적인 사항이다. 대부분의 수문기상 연구에서 월, 계절, 연 등 상대적으로 작은 시간스케일에서 평균된 수문기상변량을 대상으로 장기간의 경향을 평가하는 것이 일반적이다. 그러나 최근에 극치값들을 포함하여 수문기상변량의 확률분포 전 범위를 고려한 경향성 분석에 대한 연구 필요성이 강조되고 있다(Sankarasubramanian and Lall, 2003; Weerts et al., 2010). 즉, 평균적인 범위의 경향성 분석에서 벗어나 가뭄, 홍수 등과 같이 극치자료 관점에서 경향성을 평가하는 것을 의미한다. 특히 우리나라의 경우 극치수문사상의 증가와 대규모 재해피해가 매우 큰 상관성을 가지고 있기 때문에 수문기상시스템에서 나타나는 변동성을 정확히 파악하는 과정은 매우 중요한 사항이다(권현환과 명수정, 2011).

국내외적으로 극치사상의 발생빈도 및 양적 증가에 대한 연구는 기후변화와 맞물려 여러 분야에서 활발히 진행되고 있다. 대부분의 통계학적 연구는 극치자료에 대해서 선형회귀분석, Mann-Kendall 경향성 분석 등을 적용하여 경향성의 존재 여부와 경향성의 유의성을 검증하는 것을 목적으로 하고 있다(이정주와 권현환, 2011). 그러나 일반적인 회귀분석은 최소자승법(least-squares method, LSM)에 근거하여 경향성을 파악하며 특히 독립변수(시간)와 종속변수(기상변량)의 평균값에 국한하여 경향성을 평가하는 단점이 있다. LSM 회귀분석 기법의 적용을 위해서는 관측자료의 등분산성(homoscedasticity) 가정이 만족되어야 한다. 즉 독립변수에 대한 종속변수 값들은 정규분포를 이루는데, 이러한 정규분포의 분산이 모두 일정한 것을 말하며 독립변수에 대한 종속변수가 이루는 정규분포의 분산이 동일한 것을 의미한다. 수문기상변량은 확률밀도함수에 양끝단의 확률에 따라서 수문량의 민감도가 매우 크기 때문에 등분산성 가정에 따라 상반된 결과가 추정될 가능성이 크며 이에 대한 신중한 접근한 필요하다. 특히 극치수문자료는 일반적인 수문변량과는 다른 통계학적 특성을 가지고 있어 단순히 선형회귀분석과 같은 방법으로는 극치수문량의 경향성을 효과적으로 분석하기 어렵고 이를 수자원계획의 중요한 요소인 빈도개

념과 연계시켜 해석하는데 무리가 있다고 지적되고 있다(Katz et al., 2002).

수문기상자료의 경향성을 평가하는데 있어서 기존 LSM기반의 회귀분석 방법의 또 다른 단점은 시간에 따른 수문기상자료의 분포 특성을 포괄적으로 이용할 수 없다는 점이다. 이러한 문제로 인해서 현재 우리가 사용하고 있는 기존 회귀분석 기반의 경향성 분석은 평균값에 일반화된 결과로서 수문분야에서의 관심사항인 홍수 및 가뭄과 같은 극치사상에 대한 경향성을 효과적으로 평가하는데 단점을 가지고 있다. 이러한 점에서 확률분포의 다양한 분위에서 경향성을 평가함으로써 보다 체계적으로 경향성을 평가하는 과정이 요구된다.

최근 우리나라에서도 극치사상이 빈번히 발생하고 있으며 이러한 극치사상의 경향성을 평가하는 연구들이 다수 진행되었다. 김보경과 김병식(2009)은 기후변화가 극한 강우 발생에 미치는 영향을 평가함에 있어 일 강우 자료계열 분석에 선형회귀분석을 적용하였다. 이창환 등(2010)은 비정상성 강우빈도해석법에 의한 확률강우량 평가 지점의 선정에 있어서 통계적으로 유의한 양의 경향성의 갖는 지점을 선정하기 위해서 선형회귀분석을 이용하였으며 최종적으로 비정상성 강우빈도해석법의 적용성 및 신뢰성을 평가하였다. 그러나 정성적으로 극치사상이 증가하고 있다는 예상과는 다르게 극치강우계열에 기존 회귀분석을 적용하여 경향성을 평가해보면 통계적으로 유의한 증가 또는 감소 경향을 갖는 강우관측소가 전체적으로 많지 않다. 이는 기존 회귀분석이 평균값을 기준으로 경향성을 평가한 원인으로 볼 수 있으며 극치값을 기준으로 추정된 경향성과는 상당히 다를 수 있다는 점을 보여주고 있다. 기상자료에 대해서 LSM기반 회귀분석에 기초한 경향성 분석의 또 다른 단점은 상대적으로 이상치(outlier)에 매우 민감하다는 점이다(Gaffen et al., 2000). 민감도 관점에서 기상자료의 변동성을 평가하는데 LSM기반의 회귀분석의 대안으로 비매개변수적 방법, Robust 회귀분석 방법들이 제안된 사례가 있다(Lanzante, 1996; Gaffen et al., 2000).

이러한 LSM기반 회귀분석의 단점을 보완하고 수문기상자료의 전체 범위에서 경향성을 포괄적으로 평가할 수 있는 방법론으로써 Quantile 회귀분석이 이용될 수 있다. Koenker and Bassett (1978)이 제안한 Quantile 회귀분석은 매개변수적 방법으로써 입력되는 독립변수를 이용하여 원하는 분위(percentile)에 대한 조건부 Quantile를 추정할 수 있는 방법으로 시계열에서 나타나는 다양한 범위에서의 경향성의 평가가 가능하고 계량경제학(econometrics) 분야에서

널리 이용되고 있다(Buchinsky, 1998; Chernozhukov and Hansen, 2006; Koenker, 2005). Quantile 회귀분석은 방법론적으로 우수할 뿐만 아니라 적용이 쉽고 비선형모형을 위해서도 확장이 용이한 장점이 있다. 그러나 수문기상학분야에서 Quantile 회귀분석을 적용한 사례는 매우 제한적으로 수행되었다(Bremnes, 2004; Coelho et al., 2008; Friederichs and Hense, 2007; Koenker and Schorfheide, 1994).

본 연구에서는 극치수문자료의 경향성 평가를 위해서 Quantile 회귀분석 방법을 제시하고자 하며 최종적으로 우리나라 극치강수량의 경향성을 종합적으로 검토하는데 주된 목적이 있다. 이를 위해서 장기간의 수문자료를 보유한 강수지점에 대해서 Quantile 회귀분석 방법을 주요 해석 방법으로 적용하였다. 극치강수량 및 홍수 등에 대한 경향성을 분석한 국내 관련 연구로서는 강수량과 기온 자료로부터 추출한 극한지수를 이용하여 극한 기상사상의 통계적 변화 경향을 분석한 사례(김보경 등, 2008)와 강우량의 증가 경향성을 고려한 확률강우량을 산정하기 위해 극치강수량의 이동평균과 빈도해석을 통해 증가 경향성을 확률강수량 산정에 반영한 연구가 있다(권영문 등, 2009). 이동률 등(2004)은 GCM 모의결과와 마코프 연쇄모형으로 모의된 수문자료를 바탕으로 가뭄지수 등의 변화를 모의하였고, 정대일 등(2008)은 Mann-Kendall 검정을 통해 극치강수량과 홍수량의 경향성을 판별하고, 대수정규 추세모형을 이용하여 설계홍수량을 산정하였다. 이정주 등(2010)은 극치분포를 이용하여 경향성 분석과 연계된 Bayesian 빈도해석 방법을 제안하였다. 권현한 등(2011)은 극치강수량의 경향성을 이용한 빈도해석 방법을 이용하여 미래의 확률강수량을 전망하였다.

본 논문은 극치자료의 경향성의 존재 여부를 다양한 관점에서 검증하는데 목적을 두고 있다. 본 논문에서는 첫째, 대상자료 및 Quantile 회귀분석 방법의 해석과정을 설명하였다. 둘째, 우리나라 주요 강수 지점에 연최대치 강수량을 대상으로 모형을 적용, 경향성의 존재 여부에 대한 통계학적 추론을 실시하였다. 마지막으로 해석결과를 요약하고 모형 확장 가능성 및 향후 연구방향을 제시하는 단계로 구성되었다.

2. 대상 자료 및 연구방법

2.1 대상 자료

본 연구에서는 분석을 위한 극치수문사상을 추출하기 위해 기상청에서 관리하는 64개 관측소의 일강수량 자료

를 이용하였다. 64개의 관측소의 일강수량 자료로부터 연 최대 일강수량계열(annual daily maximum rainfall series)을 산정하였다. 본 연구에서는 2011년까지 자료를 활용하여 연 최대치자료를 추출하였으며 자료연한은 짧게는 24년에서 길게는 52년의 범위를 갖는 자료계열을 구성할 수 있었다. Fig. 1은 본 연구에서 사용된 관측소명, 위치, 관측 년도를 나타낸다. 사각박스 안의 숫자는 관측 년도를 다각형 박스안의 숫자는 본 연구에서 관측소에 적용한 No.를 나타낸다.

2.2 Quantile 회귀분석

다중회귀분석, 인공신경망(neural networks)과 같은 대부분의 회귀분석 모형들은 주어진 예측인자 즉, 독립변수가 주어진다면 종속변수의 조건부 평균(conditional mean)에 해당하는 점추정치(point estimate)를 제공한다. 앞서 언급하였듯이 수문기상학분야에서는 일반적으로 LSM 회귀분석으로 추정 한다. 표준적 방법인 LSM 기반 회귀분석은 종속변수와 회귀인자의 평균 관계를 요약하는데 유용한 도구이지만, LSM 회귀분석은 종속변수와 회귀인자의 평균 관계만을 파악한다는 점에서 부분적인 정보만을 제공하게 된다. 변수 관계를 종합적으로 평가하기 위해서는 종속변수의 조건부 분포에서 여러 Quantile에 대해서 회귀인자의 영향을 파악해야 한다. 강우량을 대상으로 설명하면 강우 분포의 다양한 Quantile에서 회귀인자의 영향이 어떻게, 얼마나 달라지는지를 알아야 한다. Quantile 회귀분석은 이렇게 풍부한 정보를 얻는 가장 유용한 방법이라 할 수 있다(Koenker, 2005; Hao and Naiman, 2007). 본 절에서는 Quantile 회귀분석 방법에 대한 이론적 배경, 매개변수 추정 등에 대해서 정리하였다.

확률분포(probability distribution function, pdf)를 가지는 무작위 변량 Y 에 대해서 누가 확률은 다음과 같이 정의될 수 있다.

$$F(y) = \text{prob}(Y \leq y) \quad (1)$$

무작위 변량 Y^* 의 τ 에 해당하는 Quantile은 다음과 같은 역함수로 정의될 수 있다.

$$Q(\tau) = \infty\{y : F(y) \geq \tau\} \quad (2)$$

여기서, $0 < \tau < 1$ 이며 특히 중앙값의 경우 $\tau = 0.5$ 의 값을 갖게 된다.

무작위 변량 Y 의 무작위 표본 $\{y_1, \dots, y_n\}$ 에 대해서 표본의 중앙값은 절대편차(absolute deviation)의 합을 다음과 같이 최소화함으로써 추정이 가능하다.

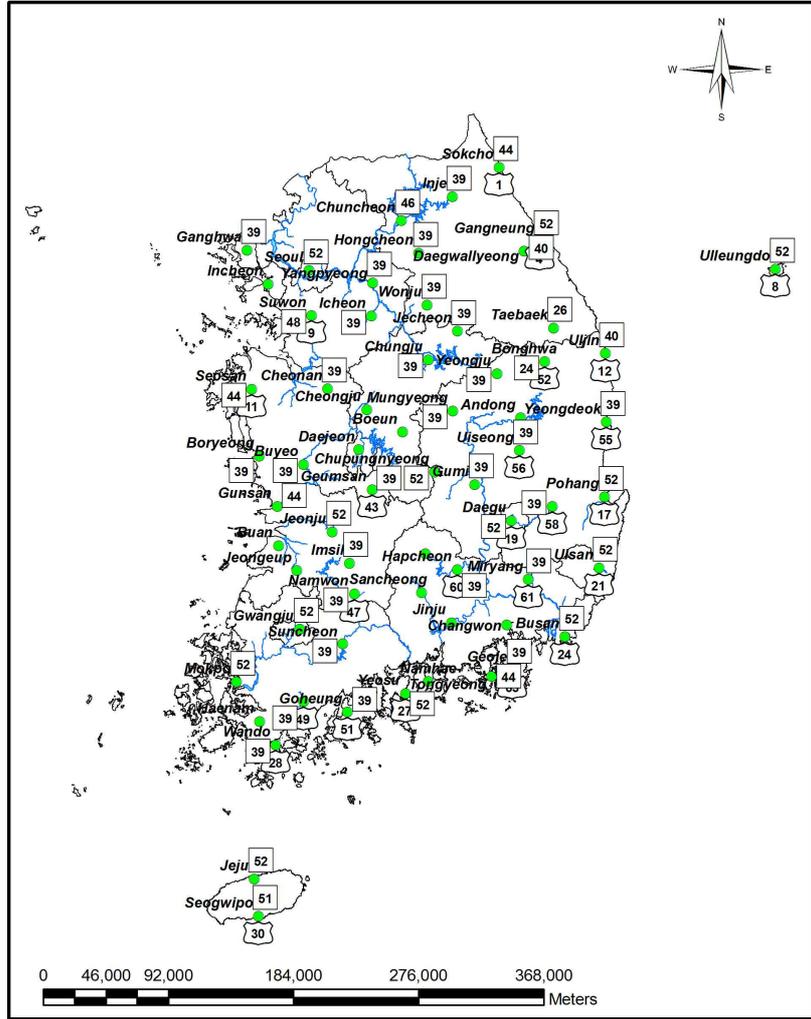


Fig. 1. Weather Stations Used in This Study

$$\operatorname{argmin}_{\xi \in R} \sum_{i=1}^n |y_i - \xi| \quad (3)$$

마찬가지로 τ 에 해당하는 표본 Quantile $\xi(\tau)$ 은 Eq. (2)의 $Q(\tau)$ 와 유사성을 가지는 값으로써 다음과 같이 최적화문제로 귀결된다.

$$\operatorname{argmin}_{\xi \in R} \sum_{i=1}^n \rho_{\tau}(y_i - \xi) \quad (4)$$

여기서, $\rho_{\tau}(z) = z(\tau - I(z < 0))$, $0 < \tau < 1$. $I(\cdot)$ 는 지시 함수(indicator function)를 나타낸다.

우선 평균에 대해서 검토해보면 잔차제곱합을 최소화 함으로써 표본평균 $\hat{\mu}$ 의 추정이 가능하며 다음 Eq. (5)와 같이 나타낼 수 있다. 또한 선형 조건부 평균 함수(linear conditional mean function) $E(Y | X = x) = x'b$ 를 도입

하여 Eq. (6)과 같이 확장이 가능하다.

$$\hat{\mu} = \operatorname{argmin}_{\mu \in R} \sum_{i=1}^n (y_i - \mu)^2 \quad (5)$$

$$\hat{b} = \operatorname{argmin}_{b \in R^p} \sum_{i=1}^n (y_i - x_i'b)^2 \quad (6)$$

선형 조건부 Quantile 함수 $Q(\tau | X = x) = x'b(\tau)$ 는 Eq. (7)과 같이 규정할 수 있으며 최적화 과정을 통해 추정이 가능하다.

$$\hat{b}(\tau) = \operatorname{argmin}_{b \in R^p} \sum_{i=1}^n \rho_{\tau}(y_i - x_i'b) \quad (7)$$

최종적으로 Quantile 회귀분석을 수행하기 위한 Eq. (8)은 최대-최소화를 위한 쌍대문제(dual problem) 형태의 일반적인 선형프로그래밍(linear programming)으로

귀결된다.

$$\min_{b \in R^p} \left[\sum_{i \in \{i: y_i \geq x'_i b\}} \tau |y_i - x'_i b| + \sum_{i \in \{i: y_i < x'_i b\}} (1 - \tau) |y_i - x'_i b| \right] \quad (8)$$

3. 적용

본 연구에서는 Quantile Regression(이하 QR)과 일반적으로 사용되는 평균에 대한 회귀분석 Ordinary Regres-

sion(이하 OR)방법을 이용하여 시간에 따른 연 최대 극대 강수량의 경향성을 평가·비교하기 위하여 종속변수 Y 를 연 최대 극대 강수량, 독립변수 X 를 시간으로 하는 선형회귀식 $Y = a + bX$ 를 적용하였다. QR 방법의 적용은 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95% 각각의 Quantile에 대하여 Eq. (8)을 이용하여 회귀계수 b 를 추정하였으며, LSM 기법을 이용한 OR 방법의 경우에도 회귀계수 b 를 추정하였다. Table 1은 회귀계수 b 에 대하여

Table 1. A Comparison of Statistical Significance Test of Regression Coefficient between QR and OR

No.	Quantile Regression		Ordinary Regression		No.	Quantile Regression		Ordinary Regression	
	t-Test	p-value	t-Test	p-value		t-Test	p-value	t-Test	p-value
1	X	0.0966	X	0.7308	*33	O.K	0.0004	O.K	0.0044
**2	O.K	0.0064	X	0.3531	*34	O.K	0.0028	O.K	0.0165
*3	O.K	0.0071	O.K	0.0344	*35	O.K	0.0004	O.K	0.0158
4	X	0.0936	X	0.1540	*36	O.K	0.0017	O.K	0.0342
*5	O.K	0.0019	O.K	0.0389	37	X	0.0903	X	0.5775
**6	O.K	0.0011	X	0.0745	*38	O.K	0.0009	O.K	0.0138
*7	O.K	0.0015	O.K	0.0496	**39	O.K	0.0497	X	0.6712
8	X	0.1128	X	0.6659	**40	O.K	0.0090	X	0.1359
9	X	0.0552	X	0.1776	41	X	0.1159	X	0.4146
**10	O.K	0.0041	X	0.1249	**42	O.K	0.0230	X	0.9626
**11	O.K	0.0455	X	0.1003	43	X	0.0657	X	0.2642
**12	O.K	0.0187	X	0.1695	44	X	0.1268	X	0.3934
**13	O.K	0.0317	X	0.1568	*45	O.K	0.0021	O.K	0.0373
**14	O.K	0.0114	X	0.1641	**46	O.K	0.0022	X	0.0606
*15	O.K	0.0017	O.K	0.0440	**47	O.K	0.0224	X	0.1472
**16	O.K	0.0115	X	0.1957	**48	O.K	0.0364	X	0.2818
**17	O.K	0.0007	X	0.2148	49	X	0.0957	X	0.6337
**18	O.K	0.0221	X	0.1244	50	X	0.1396	X	0.9042
**19	O.K	0.0123	X	0.1202	**51	O.K	0.0320	X	0.5614
**20	O.K	0.0083	X	0.0704	52	X	0.3905	X	0.9536
**21	O.K	0.0055	X	0.6406	**53	O.K	0.0016	X	0.0919
**22	O.K	0.0490	X	0.2014	*54	O.K	0.0002	O.K	0.0221
**23	O.K	0.0124	X	0.1306	**55	O.K	0.0219	X	0.5732
24	X	0.1252	X	0.7502	**56	O.K	0.0163	X	0.2945
**25	O.K	0.0109	X	0.2682	*57	O.K	0.0039	O.K	0.0156
26	X	0.0960	X	0.9102	58	X	0.1314	X	0.7104
**27	O.K	0.0153	X	0.4838	**59	O.K	0.0094	X	0.0953
**28	O.K	0.0301	X	0.5530	**60	O.K	0.0272	X	0.1578
**29	O.K	0.0163	X	0.0670	61	X	0.2138	X	0.9186
*30	O.K	0.0014	O.K	0.0426	**62	O.K	0.0112	X	0.2305
**31	O.K	0.0428	X	0.0721	63	X	0.3229	X	0.8994
**32	O.K	0.0017	X	0.4009	64	X	0.1135	X	0.3054

* QR과 OR의 결과가 모두 통계적 경향성을 보이는 경우

** QR의 결과만이 통계적 경향성을 보이는 경우

Eq. (9)를 이용한 t-검정의 p -value 값을 보여준다. Table 1에서 QR방법의 p -value는 관측소별로 적용한 10개 Quantile로 구한 p -value 중 통계적으로 가장 유의한 p -value만을 표시하였다.

$$t = \frac{b}{SE} \quad (9)$$

$$SE = \frac{s}{\sqrt{(X - \bar{X})^2}} \quad (10)$$

여기서, s 는 잔차에 대한 분산, SE 는 표준 오차, b 는 회귀 방법을 통해 추정된 회귀계수를 나타낸다. t 검정에 사용되는 자유도는 지점별 자료의 길이를 n 이라 할 때, $n-2$ 의 값이 사용된다.

Table 1에서 적용한 t-검정의 귀무가설은 “경향성이 없다”라는 $b=0$ 이며, 오차수준 5%에 대한 양측검정을 실시하였다. Table 1에서 일반적인 회귀방법인 OR 방법의 경우 64개 지점 중 13개 지점이, QR 방법의 경우 64개 지점 중 47개 지점이 귀무가설을 기각하여 통계적으로 유의한 경향성을 보여주고 있다. 표의 음영된 부분은 QR과 OR의 경우에서 통계적으로 유의한 결과가 나타난 지점을 보여주고 있다. OR의 t-검정에서 통계적으로 유의한 결과로 나타난 13개 지점이 모두 QR 방법을 적용하였을 경우의 통계적으로 유의한 결과가 나타나는 지점에 포함되어 있는 것을 확인할 수 있다. 이러한 결과는 연 최대 극대강수량의 경향성의 평가에 있어서 QR 방법을 통한 경향성 분석의 결과가 기존의 OR방법의 결과에 비하여 보

다 신뢰성 있는 결과를 보여주고 있음을 의미한다 하겠다. 즉, 추정된 회귀계수인 기울기 b 의 값이 통계적으로 유의한 결과를 보여주고 있다.

Figs. 2~6은 원주, 서귀포, 울진, 대전, 대관령의 5개 지점들의 연 최대 극대강수량에 대한 시계열과 OR과 QR로 추정된 회귀선 및 회귀계수를 보여주고 있다. Fig. 2는 Fig. 1에서 No.7에 해당되는 원주 지점에 대한 그림으로 오른쪽 그림은 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%의 Quantile을 QR 방법에 적용하여 구한 각각의 회귀계수 b 값에 대하여 t-검정 결과인 p -value를 도시한 그림이며, 점선은 경향성이 유의수준 5%에서 통계적으로 유의함을 보이는 한계선이다. 실선은 OR방법으로 구한 b 에 대한 p -value 값을 직선으로 도시함으로써, OR과 QR방법으로 구한 적합성 결과를 비교하였다. 원주 지점의 OR방법의 p -value의 값은 Table 1에서 0.0496으로 그림에서 실선과 점선이 겹쳐져 있음을 확인할 수 있으며, 50%, 60%, 70%, 80%, 90%의 Quantile을 적용한 QR 방법의 회귀계수가 통계적으로 유의한 결과를 나타내 주고 있음을 확인할 수 있다.

좌측의 그림은 연 최대 극대강수량의 시계열과 QR과 OR로 추정된 회귀분석결과를 보여주고 있다. 실선으로 표시된 직선이 OR 방법을 사용한 회귀선을 나타낸다. 통계적으로 유의한 결과값을 보여준 50%, 60%, 70%, 80%, 90%의 회귀선 중 50%와 60%, 70%의 기울기는 OR 방법의 기울기에 비하여 크게 상승하였으며, 80%, 90%의 결과는 OR의 기울기보다 약간 상승하였다. 평균적으로 QR

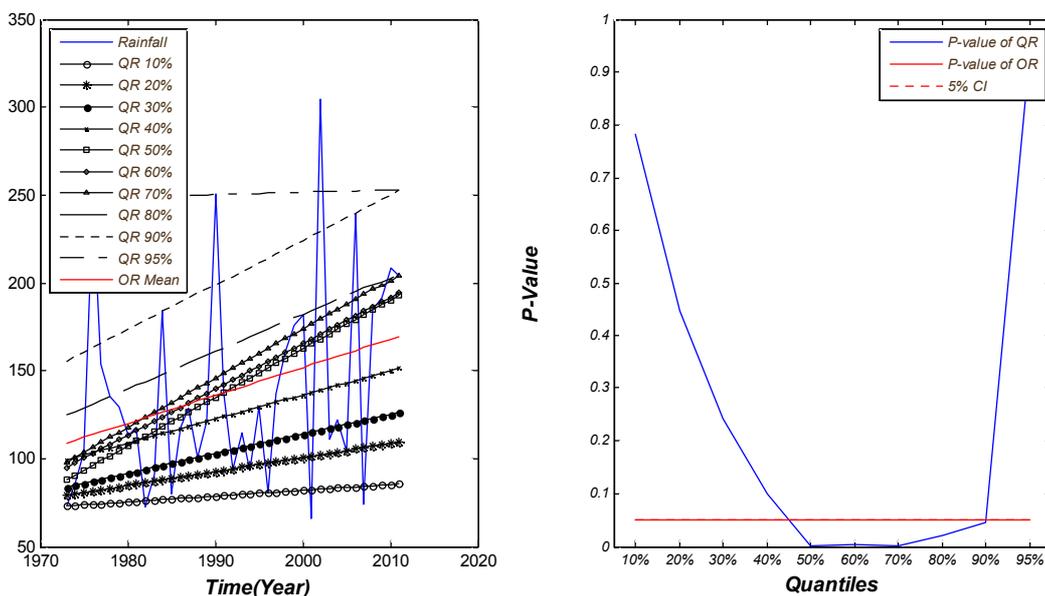


Fig. 2. Quantile Regression Results and Their p -values of Wonju Station (No.7)

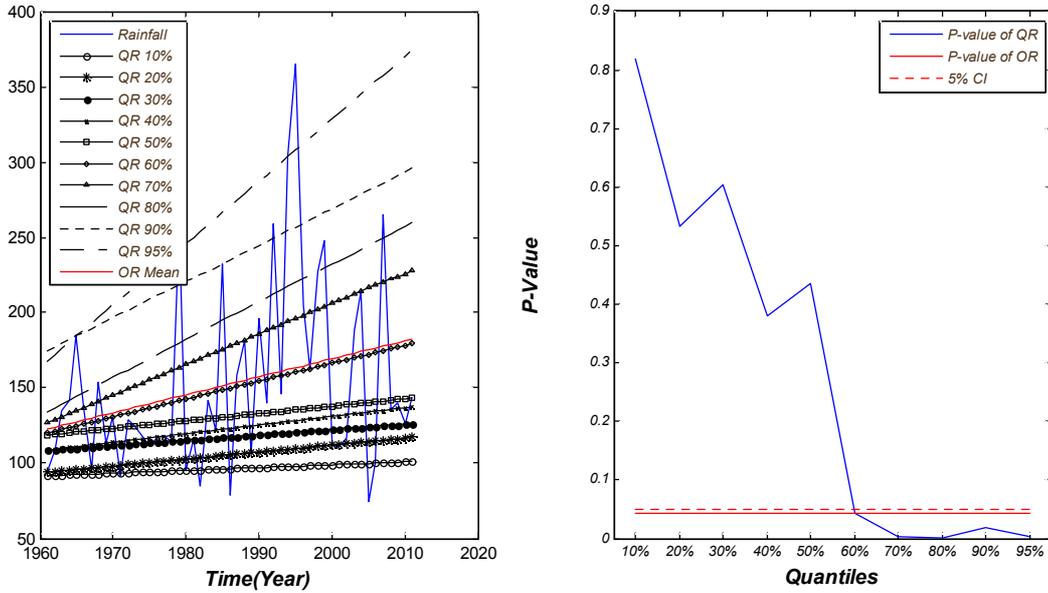


Fig. 3. Quantile Regression Results and Their p -values of Seogwipo Station (No.30)

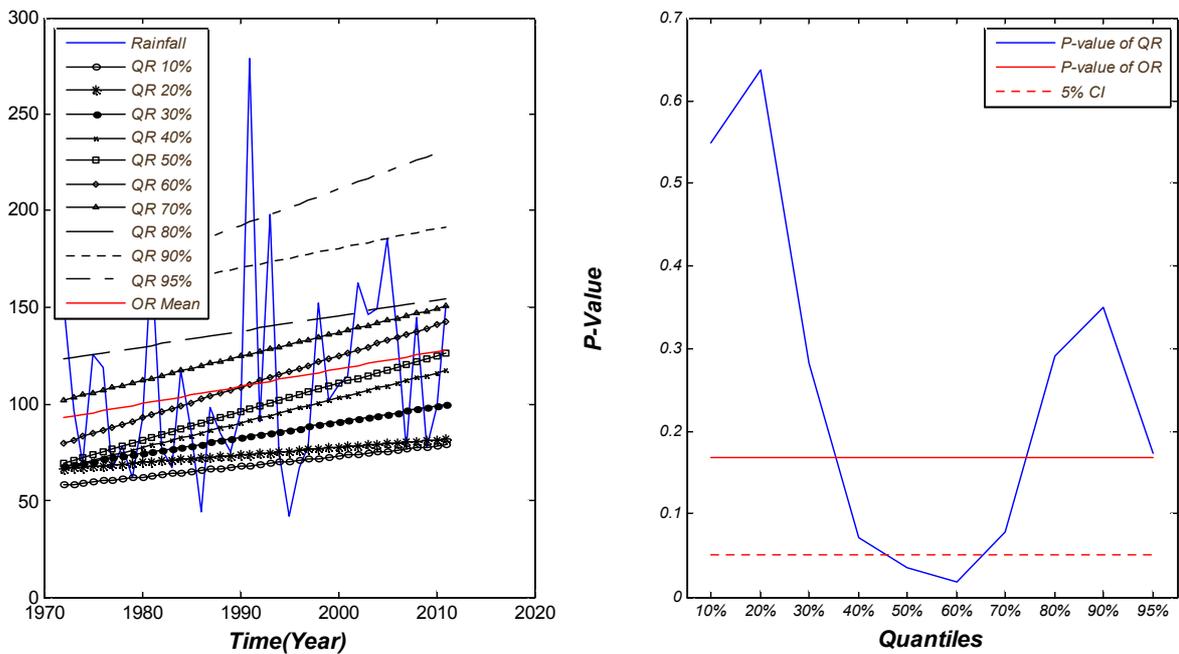


Fig. 4. Quantile Regression Results and Their p -values of Uljin Station (No.12)

방법의 결과가 OR 방법에 비하여 큰 기울기 값을 보여주고 있음을 확인할 수 있다. Fig. 3서귀포의 경우는 60%, 70%, 80%, 90%, 95%에서 통계적으로 유의한 회귀계수가 도출되었으며, QR 방법의 회귀선의 기울기가 OR 방법에 비하여 증가하고 있음을 보여주고 있다. 울진지점의 결과인 Fig. 4는 OR 방법의 p -value가 한계선 위에 위치하므로 OR 방법으로 구한 회귀계수는 통계적으로 유의하지

않다. 하지만, 50%와 60%의 Quantile에 대한 QR 회귀계수는 통계적으로 유의한 결과를 나타내주고 있으며, OR 방법의 회귀선에 비하여 기울기가 큰 회귀선이 추정되었다. Fig. 5는 대전지점의 결과를 보여주고 있으며, QR 방법에 의한 회귀선이 OR보다 경사가 급하며, 60%, 80%, 95% Quantile에 대한 QR 회귀계수는 통계적으로 유의한 결과를 주고 있음을 확인할 수 있다.

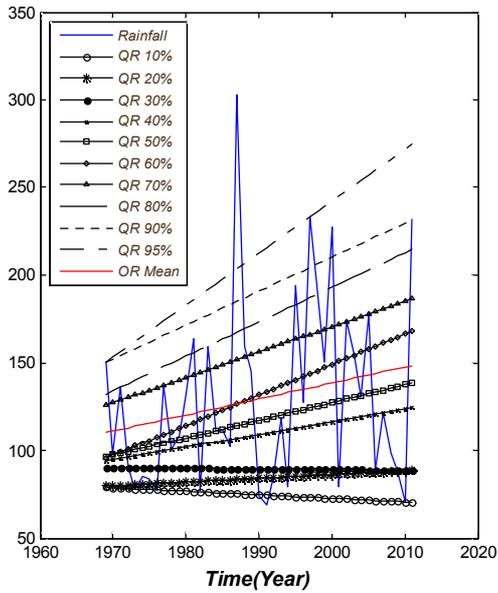


Fig. 5. Quantile Regression Results and Their p -values of Daejeon Station (No.14)

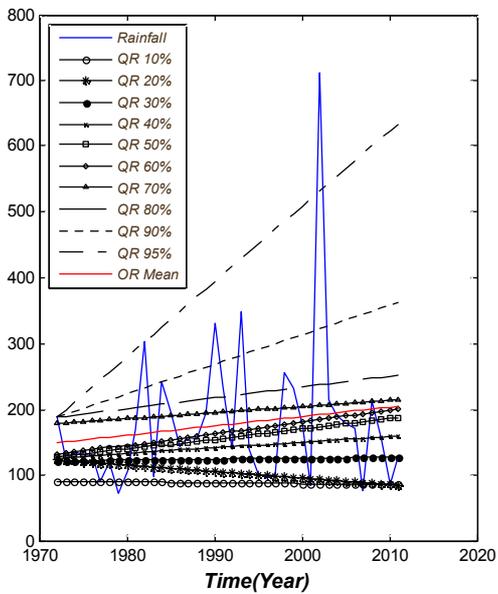


Fig. 6. Quantile Regression Results and Their p -values of Daegwallyeong Station (No.2)

Fig. 6에서 90%와 95%의 Quantile에서 통계적으로 유의한 경향성을 보이는 대관령 지역의 경우 회귀선의 기울기가 90%에서 95%로 증가함에 따라 급격하게 증가하였다. 이러한 이유는 700 mm 이상의 강우량이 포함되어 95% Quantile을 적용할 경우 이상치에 큰 가중값이 주어짐으로써 기울기가 급격하게 나타나기 때문에 95%의 값의 적용에 있어서 자료의 특성을 고려하여 적용해야 할 것으로 판명된다. 동일하게 Figs. 2~5에서도 다른 QR 회귀선에 비하여 95% 값이 급격하게 변하고 있음을 확인할 수 있

다. 이러한 95%의 결과는 전체적인 자료의 경향보다는 몇몇 이상치에 의한 결과에 크게 영향을 받으므로, 자료의 전체 경향 판단에 있어서 적절하지 않다.

Table 2는 QR 방법의 기울기가 통계적으로 유의한 결과를 보이는 47개 지점에 대하여 OR방법을 적용하여 구한 기울기 값 대비 증·감을 나타낸 표이며, 결과적으로 44개 지점의 기울기가 OR방법에 비하여 증가하며, No.25, No.28, No.42에 해당되는 3개 지점의 기울기는 음의 기울기가 증가하는 걸 볼 수 있다. Table 2에서 비교에 이용된

Table 2. A Comparison of Variation of Slope Parameter between QR and OR

No.	Station	Slope parameter		Variation	No.	Station	Slope parameter		Variation
		OR	QR				OR	QR	
2	100	1.42	4.47	Increase	31	192	1.32	1.69	Increase
*3	101	1.32	1.56	Increase	**32	201	0.90	10.22	Increase
*5	108	1.28	2.08	Increase	*33	202	2.77	3.08	Increase
6	112	1.02	1.60	Increase	*34	203	2.07	2.73	Increase
*7	114	1.59	2.57	Increase	*35	211	1.82	2.26	Increase
10	127	1.09	2.82	Increase	*36	212	1.62	2.65	Increase
11	129	0.88	1.38	Increase	*38	221	1.89	2.79	Increase
12	130	0.90	1.53	Increase	39	226	0.43	2.30	Increase
13	131	0.76	1.24	Increase	**40	232	0.89	3.73	Increase
14	133	0.91	1.85	Increase	**42	236	0.05	-5.97	Decrease(-)
*15	135	0.81	2.10	Increase	*45	244	1.30	2.97	Increase
16	136	0.74	1.55	Increase	46	245	1.64	2.77	Increase
17	138	0.85	2.24	Increase	47	247	0.81	1.50	Increase
18	140	1.09	2.67	Increase	48	256	0.78	2.24	Increase
19	143	0.62	1.30	Increase	51	262	0.73	3.70	Increase
20	146	0.73	1.21	Increase	53	272	1.10	1.91	Increase
21	152	0.30	1.82	Increase	*54	273	0.99	1.66	Increase
22	155	1.91	4.19	Increase	**55	277	0.45	5.06	Increase
23	156	0.80	1.51	Increase	**56	278	0.51	3.32	Increase
25	162	-0.73	-2.22	Decrease(-)	*57	279	1.24	1.78	Increase
**27	168	0.32	2.47	Increase	59	284	1.02	2.86	Increase
28	170	-0.71	-3.71	Decrease(-)	60	285	1.12	1.78	Increase
29	184	1.19	1.77	Increase	62	289	1.18	2.86	Increase
*30	189	1.19	2.04	Increase					

* QR과 OR의 결과가 모두 통계적 경향성을 보이는 경우
 ** 95% Quantile만이 t-검정을 통과한 경우

QR의 기울기는 지점별로 t-검정을 통과한 Quantile에 대한 기울기 값들의 평균값이 이용되었다.

Fig. 7은 QR 방법을 적용하였을 경우 통계적으로 유의한 결과를 보여준 47개 지점에 대하여 적용된 Quantile별 t-검정 통과 횟수를 도시한 도수분포표를 보여준다. 70%, 80%, 90%, 95%의 Quantile이 큰 빈도를 보여주며, 특히 80%와 95%의 Quantile이 가장 큰 빈도로 나타났다. 이러한 도수분포 형태는 자료의 분포형태가 정규분포가 아닌 Tail이 두꺼운 분포를 갖는 극치강우량의 특성이 반영된 결과로 판단할 수 있다. 즉, 기존 평균 관점에서 자료의 경향성을 판단하는 방법론의 경우 극치강우의 실질적인 증가 경향을 효과적으로 평가하지 못하고 있는 것으로 판단할 수 있다. 반면 본 연구에서 제시하는 QR 방법은 기상 변동성 및 기후변화로 인해 극치강우의 발현빈도 및 양적

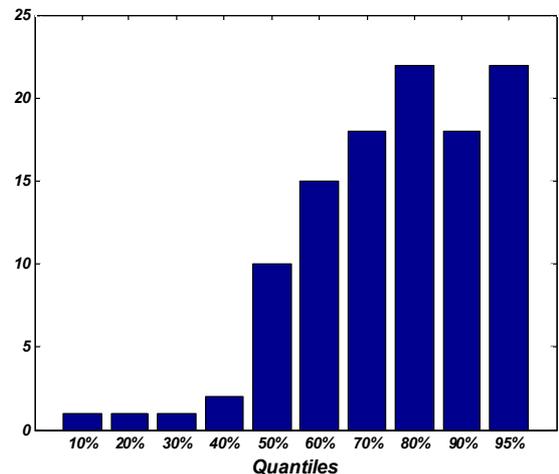


Fig. 7. Histogram Showing Acceptance Rate of t-test Given Quantiles

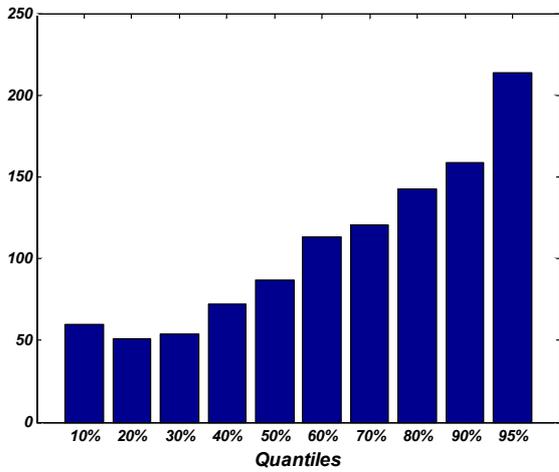
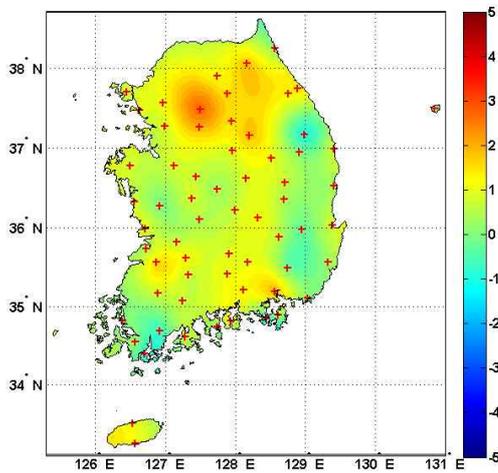


Fig. 8. Histogram Showing the Ratio of Slope Estimated from QR_{slope}/OR_{slope}

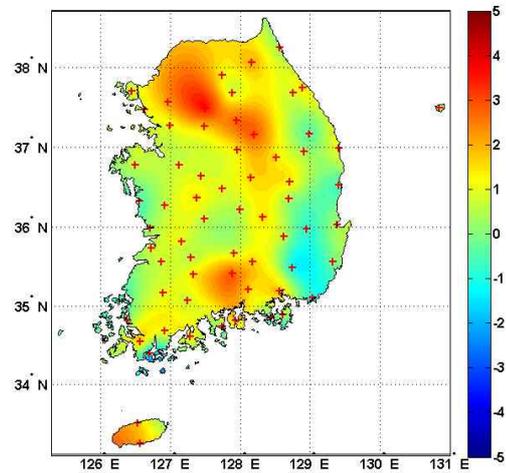
증가 경향을 보다 현실적으로 평가하고 있음을 확인할 수 있었다.

Fig. 8은 본 연구에서 이용된 64개 지점에 대한 연 최대 극대강수량에 QR 방법과 OR 방법을 적용하여 추정된 회귀선의 평균 기울기, 즉, 회귀계수 b 의 평균에 대한 비를 보여주고 있다. 즉 분위값 50% 이하에서는 QR 방법이 OR 방법 대비 상대적으로 작은 기울기를 가지고 추정되어지고 있다. 반면 60% 이상의 분위에서는 QR 방법의 기울기가 OR 방법에 비해 크게 나타났다. Fig. 7의 도수분포표에서 큰 빈도를 보여준 70%, 80%, 90%, 95%의 평균기울기의 비를 살펴보면, 70~90%의 QR값은 OR값에 비하여 1.15~1.55배의 기울기의 증가가, 95%의 QR값은 OR값에 비하여 기울기가 2.15배 증가하였다.

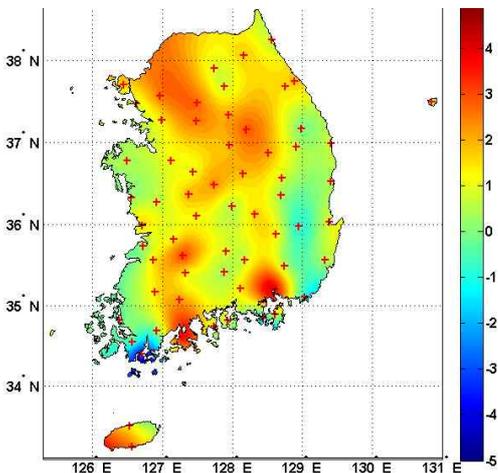
Fig. 9는 OR 방법으로 구한 회귀계수 b 와 t-검정결과



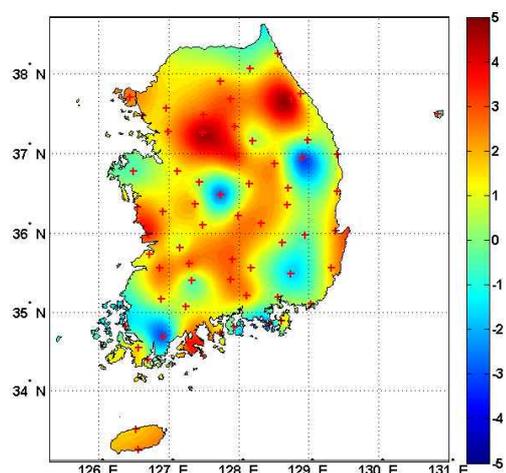
(a) Spatio-temporal distribution Regression coefficient of OR



(b) Spatio-temporal distribution Regression coefficient of 70% QR



(c) Spatio-temporal distribution Regression coefficient of 80% QR



(d) Spatio-temporal distribution Regression coefficient of 90% QR

Fig. 9. A Comparison of Spatio-temporal Distribution of Regression Coefficient between OR and QR

통계적으로 유의한 결과를 나타내는 빈도가 큰 각각의 Quantile 70%, 80%, 90%에 대하여 QR방법으로 추정된 회귀계수 b 를 64개 지점에 대하여 추정한 후 공간분포 시킨 결과를 (a), (b), (c), (d) 순으로 보여주고 있다. OR 방법의 결과 경기도와 강원도, 전라북도와 경상남도 일부의 기울기 값이 1~2의 증가 값이 나타났으며, 그 외의 부분은 모두 기울기가 0에 가까운 값이 나타나고 있다. 전체적으로 기울기의 값이 평활하게 나타나고 있다. 반면에 QR 방법을 통한 (b), (c), (d)의 경우는 OR에 의한 결과에 비하여 지역간 차이가 뚜렷하게 나타났다. QR 방법의 결과 공통적으로 경기도와 충청북도, 남해안 일부 지역의 기울기의 증가 경향이 뚜렷하며, 동해안과 전라남도 일부 지역에 연 최대 극대강수량의 감소 경향이 나타났다. 90% Quantile에 대한 공간분포 결과는 기울기의 편차가 더욱 뚜렷하게 나타나며, 충청도 일부 지역에 감소경향이 보이고 있다.

4. 결론

본 연구에서는 QR 기법을 이용한 연 최대 극대강수량의 경향성 분석을 실시하였다. 이를 위해, 64개 강우 관측 지점의 연 최대 극대강수량 자료에 대하여 QR 방법과 OR 방법으로 추정된 기울기에 대한 t-검정을 실시하여 경향성에 대한 통계적 유의성을 평가하였다. QR 방법을 통한 경향성 분석결과 64개 지점 중 47개 지점에서 5% 유의수준 내에서 t-검정을 통과한 반면 OR 방법의 경우 t-검정을 통과한 강수지점이 총 13개 지점으로 QR 방법에 비하여 약 3배의 차이를 보여주었다. 이는 OR 방법이 자료의 평균을 중심으로 경향성을 평가하는 기법인데 반해 QR 방법은 자료의 다양한 분위에서 경향성을 평가함으로써 극대 및 극소 부분에서의 경향성을 보다 유연하게 감지하는 이유로 판단된다.

QR 기법의 적용 결과 70%이상의 Quantile에서 통계적 유의성이 크게 나타나고 있음을 확인할 수 있었다. 이는 왜곡된 분포(positively skewed) 특성을 갖는 극치강수량의 특성을 반영하는 결과로 평가된다. QR 방법을 통한 적정 Quantile일의 결정에 있어서 QR 95%의 결과는 통계적으로 유의한 결과를 주지만, 경향성 분석에 있어서 이상치에 대한 민감도가 매우 큰 것으로 나타나고 있어 전체 자료의 경향성 평가에 이용하는 것은 적절치 않은 것으로 판단된다.

QR 방법의 검정결과 최적의 적합성을 보인 80% Quantile에 대한 공간분포에서 연 최대 극대강수량은 경기도, 충

청북도, 전라남도, 경상북도 일부 지역은 강한 증가경향이 나타났으며, 경상남도 지역은 약한 감소, 전라남도 일부는 큰 감소 경향이 나타나고 있음을 확인할 수 있었다. 또한, 80% QR 방법을 적용한 기울기 값은 OR 방법에 비하여 평균적으로 1.5배의 기울기 값을 보여주었다. 이러한 결과는 기존의 평균적 접근 방법으로 인하여 연 최대 극대강수량의 경향성이 과소평가될 수 있음을 보여준다. 기본적으로 평균을 중심으로 경향성을 평가해보면 자료 전체의 극값의 영향을 평활화시키는 특징을 보여주고 있어 극치 자료 계열의 경향성 검정하는데 단점이 존재한다. 반면 QR 방법을 통한 경향성 평가는 이러한 평균 중심의 해석 문제점을 개선할 수 있으며 자료가 정규분포를 따르지 않거나 왜곡된 분포형태를 갖더라도 자료의 경향성 평가에 유용하게 사용될 수 있을 것으로 판단된다. 전체적으로 극치자료에 대한 경향성 검정에 있어서 QR 방법은 OR에 의한 결과에 비하여 통계적으로 유의한 결과를 도출함으로써 경향성 분석 시 보다 신뢰성 있는 해석이 가능하였다.

추후 연구로서 보다 알맞은 적정 Quantile의 선택 방법과 비선형 Quantile Regression 방법으로 확장하여 향상된 경향성 분석 기법으로 발전시킨다면 수공학 분야에서 경향성 평가를 위한 해석도구로서 활용이 가능할 것으로 판단된다.

감사의 글

이 논문 또는 저서는 2010년 교육과학기술부의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(한국연구재단-2010-220-D00083).

참고문헌

- 권영문, 박진원, 김태웅 (2009). “강우량의 증가 경향성을 고려한 목표연도 확률강우량 산정.” **대한토목학회논문집**, 대한토목학회, 제29권, 제2B호, pp. 131-139.
- 권현한, 명수정 (2011). “Bayesian GLM 기법과 통계학적 Downscaling 기법을 활용한 미래 재해위험도 변동성 평가기법 개발.” **한국방재학회논문집**, 한국방재학회, 제11권, 제6호, pp. 207-216.
- 권현한, 소병진, 윤필용, 김태웅, 황석환 (2011). “누적평균과 이동평균을 이용한 비정상성 빈도 해석결과 비교.” **한국방재학회논문집**, 한국방재학회, 제11권, 제6호, pp. 269-280.

- 김보경, 김병식 (2009). “B2 기후변화시나리오와 극한지수를 이용한 기후변화가 극한 강우 발생에 미치는 영향분석.” **대한토목학회논문집**, 대한토목학회, 제29권, 제1B호, pp. 23-33.
- 김보경, 김병식, 김형수 (2008). “극한지수를 이용한 극한 기상사상의 변화 분석.” **대한토목학회논문집**, 대한토목학회, 제28권, 제1B호, pp. 41-53.
- 이동률, 김웅태, 유철상 (2004). “기후변화가 기상학적 가뭄과 홍수에 미치는 영향.” **한국수자원학회논문집**, 한국수자원학회, 제37권, 제4호, pp. 315-328.
- 이정주, 권현한 (2011). “극치강수량의 시공간적 특성 분석 및 지역화에 관한 연구.” **대한토목학회논문집**, 대한토목학회, 제31권, 제1B호, pp. 13-20.
- 이정주, 권현한, 김태웅 (2010). “극치수문자료의 경향성 분석 개념 및 비정상성 빈도해석.” **대한토목학회논문집**, 대한토목학회, 제30권, 제4B호, pp. 389-397.
- 이창환, 안재현, 김태웅 (2010). “비정상성 강우빈도해석법에 의한 확률강우량의 평가.” **한국수자원학회논문집**, 한국수자원학회, 제43권, 제2호, pp. 187-199.
- 정대일, 제리 스테딘저, 성장현, 김영오 (2008). “기후 변화를 고려한 홍수 위험도 평가.” **대한토목학회논문집**, 대한토목학회, 제28권, 제1B호, pp. 55-64.
- Bjørnar Bremnes, J. (2004). “Probabilistic forecasts of precipitation in terms of quantiles using NWP model output.” *Monthly Weather Review*, Vol. 132, No. 1, pp. 338-347.
- Buchinsky, M. (1998). “Recent advances in quantile regression models: a practical guideline for empirical research.” *Journal of Human Resources*, Vol. 33, No. 1, pp. 88-126.
- Chernozhukov, V., and Hansen, C. (2006). “Instrumental quantile regression inference for structural and treatment effect models.” *Journal of Econometrics*, Vol. 132, No. 2, pp. 491-525.
- Coelho, C., Ferro, C., Stephenson, D., and Steinskog, D. (2008). “Methods for exploring spatial and temporal variability of extreme events in climate data.” *Journal of Climate*, Vol. 21, No. 10, pp. 2072-2092.
- Friederichs, P., and Hense, A. (2007). “Statistical downscaling of extreme precipitation events using censored quantile regression.” *Monthly Weather Review*, Vol. 135, No. 6, pp. 2365-2378.
- Gaffen, D.J., Sargent, M.A., Habermann, R., and Lanzante, J.R. (2000). “Sensitivity of tropospheric and stratospheric temperature trends to radiosonde data quality.” *Journal of Climate*, Vol. 13, No. 10, pp. 1776-1796.
- Hao, L., and Naiman, D.Q. (2007). *Quantile regression*. Sage Publications, Inc., CA, USA.
- Katz, R.W., Parlange, M.B., and Naveau, P. (2002). “Statistics of extremes in hydrology.” *Advances in Water Resources*, Vol. 25, No. 8-12, pp. 1287-1304.
- Koenker, R. (2005) *Quantile regression*. Cambridge Univ Pr., New York.
- Koenker, R., and Bassett Jr, G. (1978). “Regression quantiles.” *Econometrica: Journal of the Econometric Society*, Vol. 46, No. 1, pp. 33-50.
- Koenker, R., and Schorfheide, F. (1994). “Quantile spline models for global temperature change.” *Climatic Change*, Vol. 28, No. 4, pp. 395-404.
- Lanzante, J.R. (1996). “Resistant, robust and non-parametric techniques for the analysis of climate data: theory and examples, including applications to historical radiosonde station data.” *International Journal of Climatology*, Vol. 16, No. 11, pp. 1197-1226.
- Parry, M., Canziani, O., Palutikof, J., Van der Linden, P., and Hanson, C. (2007). IPCC, 2007: climate change 2007: impacts, adaptation and vulnerability. *Contribution of working group II to the fourth assessment report of the intergovernmental panel on climate change*. Cambridge University Press, Cambridge.
- Sankarasubramanian, A., and Lall, U. (2003). “Flood quantiles in a changing climate: Seasonal forecasts and causal relations.” *Water Resources Research*, Vol. 39, No. 5, pp. 1134.
- Weerts, A., Winsemius, H., and Verkade, J. (2010). “Estimation of predictive hydrological uncertainty using quantile regression: examples from the national flood forecasting system (England and Wales).” *Hydrol. Earth Syst. Sci. Discuss*, Vol. 7, pp. 5547-5575.

논문번호: 12-038	접수: 2012.04.20
수정일자: 2012.05.25	심사완료: 2012.05.25