

불연속 Kernel-Pareto 분포를 이용한 일강수량 모의 기법 개발

Development of Daily Rainfall Simulation Model Using Piecewise Kernel-Pareto Continuous Distribution

권현한* · 소병진**

Kwon, Hyun-Han · So, Byung Jin

Abstract

The limitations of existing Markov chain model for reproducing extreme rainfalls are a known problem, and the problems have increased the uncertainties in establishing water resources plans. Especially, it is very difficult to secure reliability of water resources structures because the design rainfall through the existing Markov chain model are significantly underestimated. In this regard, aims of this study were to develop a new daily rainfall simulation model which is able to reproduce both mean and high order moments such as variance and skewness using a piecewise Kernel-Pareto distribution. The proposed methods were applied to summer and fall season rainfall at three stations in Han river watershed in Korea. The proposed Kernel-Pareto distribution based Markov chain model has been shown to perform well at reproducing most of statistics such as mean, standard deviation and skewness while the existing Gamma distribution based Markov chain model generally fails to reproduce high order moments. It was also confirmed that the proposed model can more effectively reproduce low order moments such as mean and median as well as underlying distribution of daily rainfall series by modeling extreme rainfall separately.

Keywords : *markov chain, daily rainfall, piecewise kernel-pareto distribution, extreme rainfall*

요 지

기존 Markov Chain 모형을 통한 일강수량 모의에서 가장 큰 문제점은 극치강수량을 재현하기 어렵다는 점이다. 이러한 문제점으로 인해 수자원계획을 수립하는데 있어서 불확실성을 가중시키고 있다. 특히 일강수량 모의기법을 통해서 추정되는 빈도강수량의 과소추정으로 인해 수공구조물 설계 시에 신뢰성을 확보하는데 문제점이 있다. 이러한 점에서 본 연구에서는 기존 Markov Chain 모형에서 일강수량에 평균적인 특성과 극치특성을 동시에 재현할 수 있도록 불연속 Kernel-Pareto Distribution 기반에 일강수량모의기법을 개발하였다. 한강유역의 3개 강수지점에 대해서 기존 Markov Chain 모형과 본 연구에서 제안한 방법을 적용한 결과 여름의 일강수량 모의 시 1차모멘트인 평균과 2-3차 모멘트 모두 효과적으로 재현하지 못하는 문제점이 나타났다. 그러나 본 연구에서 제안한 불연속 Kernel-Pareto 분포형 기반 Markov Chain 모형은 여름의 일강수량 모의 시 강수계열의 평균적인 특성뿐만 아니라 표준편차 및 왜곡도의 경우에도 관측치의 통계특성을 매우 효과적으로 재현하는 것으로 나타났다. 본 연구에서 제시한 방법론은 전체적으로 기존 Markov Chain 모형에 비해 극치강수량을 재현하는데 유리한 기법으로 판단된다. 또한 극치강수량을 일반강수량으로부터 분리하여 모의함으로써 평균 및 중간값 등 낮은 차수에 모멘트 등 일강수량에 전체적인 분포특성을 더욱 효과적으로 모의할 수 장점을 확인할 수 있었다.

핵심용어 : Markov Chain, 일강수량, 불연속 Kernel-Pareto 분포, 극치강수량

1. 서 론

수자원에서 일강수량 모의기법은 다양한 목적으로 활용되고 있으며 기본적으로 수공구조물 설계 및 수자원계획을 수립하기 위한 입력 자료로서 이용된다. 수자원계획은 장기적인 목적을 가지고 수행되는 것이 일반적이며 우리가 목표로 하는 장기간의 일강수량자료의 획득이 어렵기 때문에 단기간의 일강수량자료를 장기 모의하여 이용하게 된다. 이러한 점에서 일강수량 모의발생 기법은 수자원계획의 수립에 있

어서 미래에 발생 가능한 시나리오를 제공함과 동시에 수자원이 가지는 불확실성을 정량화 하는데 매우 중요한 역할을 하고 있다. 최근에는 기후변화에 따른 수자원의 영향 평가를 위한 Downscaling 기법으로의 발전 및 적용이 국내외적으로 매우 활발하게 이루어지고 있다(김병식 등, 2003; 권현한과 김병식, 2009; Kwon *et al.*, 2010; Stehlik and Bardossy, 2002). 따라서 일강수량 모의기법의 정확성 및 적합성은 수자원계획의 전체적인 신뢰성을 결정하는데 결정적인 역할을 한다 하겠다.

*정회원 · 교신저자 · 전북대학교 토목공학과 조교수 (E-mail : hkwon@jbnu.ac.kr)

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

일강수량을 모의하는데 있어서 강수계열의 단기간의 기억(memory)을 활용한 Markov Chain 모형이 가장 일반적이다. Markov Chain 모형은 크게 두 부분으로 나뉜다. 첫째, 강수의 발생(rainfall occurrence)을 결정짓게 되며 둘째, 강수일에 대해서 강수량(rainfall amount)을 결정하는 과정으로 구분된다. 기존 Markov Chain 모형에서 강수의 발생은 이전일의 강수사상만을 고려하는 2-State Markov Chain 모형이 주로 이용되며 강수일과 무강수일 간의 또는 무강수일과 강수일간의 천이확률(transition probability)을 통해서 정의되며 모의된다. 일반적으로 단일 관측지점에 대해서 2-State Markov Chain 모형을 통해서 일강수량의 발생 특성 즉 천이확률을 효과적으로 복원할 수 있다고 알려지고 있다(김병식 등, 1999; 권현환과 김병식, 2009).

다음 단계로서 앞에서 언급했듯이 강수일에 일강수량을 부여하게 되며 확률분포(probability distribution)를 통해서 구체화된다. 일반적으로 일강수량은 주로 Gamma 분포, Exponential 분포 등에 의해서 모의되게 되며 일강수량의 통계적 특성을 효과적으로 모의한다고 알려지고 있다. 그럼에도 불구하고 기존 2-State Markov Chain 모형에서 Gamma 분포 및 Exponential 분포를 활용한 강수량 추정에서 널리 알려진 문제점은 2차 모멘트이상의 통계치를 재현하는데 어려움이 있다는 것이다. 즉, 일강수량의 평균을 재현하는데는 문제가 없지만 일강수량계열의 극치값이 관측치에 비해 현저하게 과소추정 된다는 것이다. 이러한 문제점을 해결하기 위해서 다양한 확률분포형을 적용하여 평균적인 일강수량과 극치강수량을 동시에 복원하고자 하는 연구가 진행되어왔다. 그러나 수문통계학적으로 일강수량과 극치강수량은 통계적인 분포특성이 다르기 때문에 이를 동시에 구현하는데 근본적인 어려움이 있다. 즉, 극치강수량은 Tail이 두꺼운 극치분포를 따르기 때문에 이러한 점이 고려될 수 있는 모형이 적용되어야 한다.

Pareto 분포는 일반적으로 GPD(generalized pareto distribution)분포형으로 알려져 있으며 극치수문량을 추정하는데 매우 유리한 것으로 알려지고 있다. Pareto 분포형은 주로 설계강수량 및 설계홍수량을 추정하는데 국한해서 적용된 사례가 있으며 기존 연구에서는 이에 대한 활용성 제고 방안 및 관련 매개변수 추정연구가 진행되었다. 이순혁과 맹승진(2001)은 빈도별 설계홍수량이 GEV(generalized extreme value), GLO(generalized logistic distribution) 분포형에 의한 빈도별 설계홍수량 보다 신빙도가 높은 결과가 나오는걸 보여주었다. 엄명진 등(2008)은 GPD 분포형을 활용한 빈도해석 모형에서 최적의 Threshold를 찾기 위해서 Monte Carlo Simulation을 활용하였다. Ashkar and Tatsambon (2007)은 GPD 분포를 이용한 Quantiles 추정에 있어서 ML(method of maximum likelihood), MM(method of moments), PWM(method of probability weighted moments), GPWM(method of generalized probability weighted moments) 방법들 중 알맞은 선택을 할 수 있는 조건을 제시하였다. Singh *et al.*(2003)은 2변수 GPD(generalized pareto distribution)의 추정에 대해 7가지 추정방법에 대해서 평가 비교한 결과 PWM 방법과 L-moments 방법으로 추정한 결과가 다른 매개변수 추정방법들에 비해 우수한 결과를 주고 있다고 언급하였다.

일강수량 모의에서 일강수량 통계치를 효과적으로 복원하기 위해서 특정 분포형의 가정 없이 적용이 가능한 비매개변수적 방법이 활용되어 왔다(권현환 등, 2005; 최병규 등, 2008). 그러나 기존 매개변수적 방법에 비해 통계치 복원에 유리한 측면이 있지만 여전히 극치강수량을 효과적으로 모의하는 데는 어려움이 존재하였다. 즉, 내삽(interpolation) 측면에서는 매우 유리하지만 외삽(extrapolation)에서 광역폭에 영향을 많이 받으며 극치값의 존재 여부에 따라서 외삽시에 극치강수량 복원에 어려움이 존재한다.

이러한 점에서 본 연구에서는 새로운 관점에서 일강수량의 분포 특성을 효과적으로 재현할 수 있는 방법론을 제시하고자 한다. 기존 2-State Markov Chain 모형에서 일강수량에 평균적인 특성과 극치특성을 동시에 재현할 수 있도록 불연속(piecewise) Kernel-Pareto Distribution(PKPD)을 적용한 강수량생성모형을 제안하고자 한다. PKPD 분포형은 핵밀도함수의 장점과 Pareto분포형의 장점을 동시에 고려할 수 있도록 고안된 불연속확률분포형으로서 특정 Threshold를 넘는 극치강수량에 대해서 Thick-Tail을 가지는 Pareto 분포형을 적용하고 극치강수량을 제외한 일반적인 강수량에는 핵밀도함수를 적용하게 된다. 본 연구에서는 장기간 자료를 가지고 있는 한강유역 강수량에 대해서 모형의 적합성을 검토하고자 한다. 2장에서는 새롭게 제안되는 일강수량 모의기법에 대해서 설명하고 3장에서는 한강유역 춘천, 강릉, 서울 관측소에 대해서 제안된 방법론의 적합성을 평가하였다.

2. 불연속 Kernel-Pareto 분포를 활용한 일강수량 모의기법

2.1 Markov Chain 모형

강수자료간의 계열 상관성(serial correlation)을 고려하여 미래의 상태를 현재와 과거의 상태들과의 상관관계를 이용하여 체계적으로 재추출하는 과정을 Markov 과정(Markov process)이라 하며 과거 상태의 고려를 위해서 결정되는 2차 또는 3차와 같이 한정적인 상태공간에서의 Markov 과정을 Markov Chain라 한다. 일반적으로 일강수량 자료계열은 Markov 과정을 따르는 것으로 알려지고 있으며 Markov Chain을 적용하면 일강수의 특성을 잘 묘사하는 것으로 알려져 있다(Nord, 1975; Hann, 1976; Kwon *et al.*, 2009). Markov Chain의 기본개념은 많은 연구에서 언급되었으며 간략히 요약하면 다음과 같다.

Markov Chain의 기본 개념은 수문시스템의 움직임을 예측하고자 할 때 그 계의 과거이력을 통해 미래 상태를 예측하는 것이다(Hann 등, 1976). 따라서 미래상태의 확률은 현재 또는 과거의 조건부 확률로써 다음과 같이 나타낼 수 있다.

$$P(R_{n+1}|R_n, R_{n-1}, R_{n-2}, \dots, R_1) \quad (1)$$

여기서 R_n 은 현재의 상태, R_{n-1} , R_{n-2} , R_1 은 과거의 강수 상태를 나타낸다. 미래의 강수 상태가 아주 먼 강수 상태와는 독립이며 단지 가장 근접한 현재의 강수 상태에 의해서만 영향을 받는다면 이때 Markov Chain은 아래와 같은 조건부 확률로 나타낼 수 있으며 이를 1차 Markov Chain이라 한다.

$$P(R_{n+1}|R_n) \quad (2)$$

2-State 1차 Markov Chain 모형에서 각 상태공간에 존재하는 강수사상의 수인 벡터(S)와 사상의 발생과 비 발생의 확률을 나타내는 천이확률(transition probability)은 다음과 같은 식으로 나타낼 수 있다. 즉, 2x2 천이확률 행렬이 정의된다.

$$P_{r_{ij}} = \frac{[S_{i,j}]}{\sum_{j=1} S_{i,j}} \quad (3)$$

여기서, $P_{r_{ij}}$ 는 상태 i 에서 상태 j 로 천이되는 확률을 나타내며 상태 S 에서는 반드시 $S_{i,1}, S_{i,2}$ 중 한 곳으로 천이되므로 하나의 상태에서 발생하는 천이확률의 합은 반드시 1이 된다. 만약 3-State 2차 Markov Chain 모형을 통해서 2일전까지의 강수량을 고려하여 모의한다면 천이확률은 3x3 행렬로 정의될 수 있다. Eq. (3)의 천이확률의 각 원소의 합이 일 때의 행렬을 확률행렬(probability matrix)이라 정의하며 각각의 확률행렬은 하나의 Markov 과정을 정의하며, 강수 발생여부를 판단하게 된다. 따라서 천이확률 $P_{r_{ij}}$ 를 추정하게 되면 Markov Chain의 초기 상태만을 통해 강수발생을 모의할 수 있다. 표 1은 상태-2의 1차 Markov Chain 모형에 의한 강수발생의 천이확률 행렬(transition probability matrix)을 나타낸 것이다.

표 1에서 $0 \leq a \leq 1, 0 \leq b \leq 1$ 로서, a 는 건조일 다음에 습윤일이 발생할 확률이고 b 는 습윤일 다음에 건조일이 올 확률이다. 일단 2-State 1차 Markov 모형을 통해서 발생여부가 결정되면 강수일에 대해서 강수량을 결정해야 한다. 이를 위해서 확률분포형이 적용되며 과거 일강수량 통계치를 최대한 복원할 수 있는 확률분포형을 적용하게 된다. 지금까지 일강수량 모의를 위해서는 Gamma 분포형이 일반적으로 적용되어지고 있으며 일강수량의 통계치를 복원하는데 유리한 측면이 있는 것으로 알려지고 있다. 본 연구에서 개발된 모형과 비교를 위해서 사용된 Gamma 분포형에 대해서 간략히 요약하여 나타내었다. 2변수 Gamma분포의 확률밀도함수(probability density function)는 다음과 같다.

$$f(x) = \frac{\lambda^\eta x^{\eta-1} e^{-\lambda x}}{\Gamma(\eta)} \quad (4)$$

여기서, $x, \lambda, \eta > 0$ 이고 λ 는 축척매개변수(scale parameter)를, η 는 형상매개변수(shape parameter)를 나타낸다. Gamma 함수의 Eq. (5)의 누가분포함수(cumulative density function)는 확률밀도함수를 적분함으로써 얻어질 수 있다. 누가분포함수로부터 강수량은 다음 Eq. (6)의 Quantile함수로 부터 추정이 가능하다.

$$F(x) = \int_0^x \frac{\lambda^\eta t^{\eta-1} e^{-\lambda t}}{\Gamma(\eta)} dt \quad (5)$$

Table 1. Transition probability matrix for 2-state Markov Chain Model

		t(day)	
		Dry	Wet
t-1 (day)	Dry	1-a	a
	Wet	b	1-b

$$x = F^{-1}\left(\int_0^x \frac{\lambda^\eta t^{\eta-1} e^{-\lambda t}}{\Gamma(\eta)} dt\right) \quad (6)$$

2.2 불연속 Kernel-Pareto Distribution

앞서 언급했듯이 일강수량자료는 대부분 Gamma 분포에 의해서 적합이 이루어지며 평균을 중심으로 한 일강수량의 특성을 효과적으로 모의할 수 있다. 즉, 일강수량의 특성상 0을 중심으로 대부분의 자료가 집중되어 있고 일강수량 전체적인 특성은 Gamma 분포에 의해서 모의가 가능하나 극치강수량이 혼재하는 일강수량 전체는 하나의 Gamma 분포로 적합 시키게 되면 극치강수량의 정보를 잃을 수 있는 개연성이 크다. 즉, 분포형의 양쪽 끝단에서 나타나는 극소 및 극대강수량 정보를 효과적으로 재현하기 어려운 문제점이 있다.

이러한 점에서 본 연구에서는 내삽에서 가장 우수한 재현능력을 가지는 핵밀도함수(kernel density function)방법과 극치강수량 재현에 유리한 GPD분포를 동시에 고려할 수 있는 분포형을 개발하고 이를 일강수량모의기법에 적용하고자 한다. 핵밀도함수 방법과 GPD 분포형에 대해서 간단히 살펴보면 다음과 같다.

비매개변수적 핵밀도함수 해석방법은 어떤 분포의 가정이 필요 없이 관측자료 자체에서 분포형을 유도할 수 있기 때문에 분포형 선정의 어려움을 해소할 수 있고 또한 관측자료의 적절한 분포형을 선정할 수 있으며 특히, 내삽시에 매우 우수한 능력을 발휘한다(문영일 등, 2006).

Rosenblatt(1956)는 모든 자료가 발생되어진 각각의 위치에 막대그래프의 중앙을 위치하도록 하여, 구간을 이동시킬 수 있는 이동 히스토그램을 이용한 핵밀도함수 추정법을 개발하였으며, 이는 모든 실수 x 에 대하여 Eq. (7)과 같이 정의하였다.

$$f(x) = \frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{x-X_i}{h}\right) \quad (7)$$

여기에서, X_1, X_2, \dots, X_n 은 독립적으로 동일하게 분포된 실관측치이고, $K(\cdot)$ 는 핵함수, h 는 양의 광역폭(bandwidth)이다. 본 연구에서는 광역폭을 추정하기 위해서 간단한 Rule of thumb을 이용하여 계산하였다.

$$\hat{h}_0 = 2.78 \hat{\sigma}_n^{-1/5} \quad (8)$$

그림 1은 관측된 자료위치에 핵함수를 각각 위치하여 합산한 후 확률밀도함수로 표현하였다. 그림 1에서 관측자료의

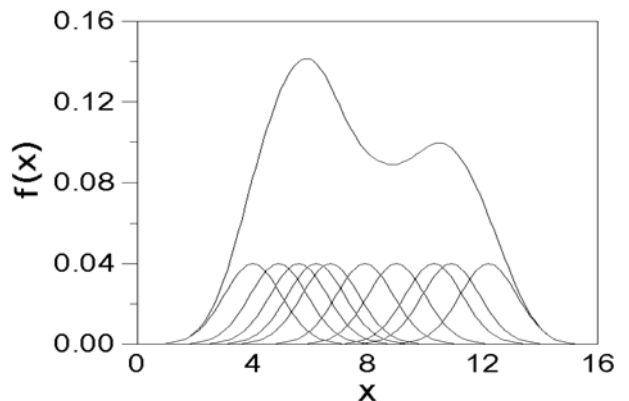


Fig. 1 Fixed kernel density function method

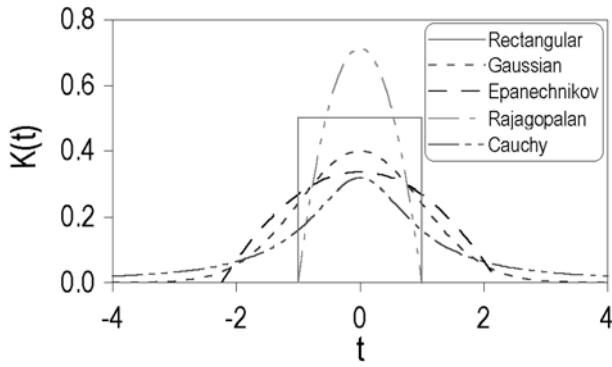


Fig. 2 Various kernel density function

개수 10, 광역폭 h 가 1인 핵함수를 사용한 핵밀도함수를 나타낸다. 위에서 언급된 핵밀도함수 추정법은 광역폭 h 의 변화 없이 일정한 값을 갖기 때문에 이를 고정 핵밀도함수 추정법 (fixed kernel density function estimator)이라 한다.

일반적인 핵함수는 중앙에서 최대치를 가지며 연속적이고 대칭인 방정식의 형태를 가진다. 확률밀도 추정시에 일반적으로 사용되는 핵밀도함수는 그림 2와 같다. 일반적으로 해석 특성에 맞는 임의의 핵함수를 사용하여도 최적의 핵함수와 비슷한 효율을 갖는다고 하였다. 그러나 사용목적에 따라 각 핵함수의 적용성을 검토할 필요가 있다. 예를 들면 연속적이고 미분 가능한 밀도함수가 필요한 경우는 경계를 갖는 핵함수보다는 경계를 갖지 않는 핵함수를 사용하는 것이 적용상에 유리하다. 본 연구에서는 극치강수량이 아닌 일반적인 강수량을 모의하기 위함으므로 일반적으로 사용되는 Gaussian 핵함수를 이용하였다.

다음으로 GPD 분포형에 대해서 살펴보면 다음과 같다. 본 연구에서 사용한 GPD분포의 확률밀도함수와 누기확률분포함수는 각각 Eq. (9)과 (10)과 같다.

$$f(x) = \frac{1}{\alpha} \left[1 - \frac{k}{\alpha} (x - \varepsilon) \right]^{-1 + \frac{1}{\alpha}} \quad (9)$$

$$F(x) = 1 - \left[1 - \frac{k}{\alpha} (x - \varepsilon) \right]^{\frac{1}{\alpha}} \quad (10)$$

여기서, k , α , ε 는 각각 Shape Parameter, Scale Parameter, Threshold를 나타낸다. 누기확률밀도함수 $F(x)$ 와 매개변수를 고려하면 극치강수량은 다음 Eq. (11)의 Quantile 함수를 통해서 추정된다.

$$x = \varepsilon + \frac{\alpha}{k} [1 - (1 - F)^k] \quad (11)$$

지금까지 설명된 두 개의 확률밀도함수 추정방법의 장점을 최대한 활용하기 위해서 일반적인 강우량에 대해서는 핵밀도함수를 적용하고자 하며 극치강수량에 대해서 GPD분포를 적용하였다. 즉, 두 개의 누기확률밀도함수를 연결시키게 되

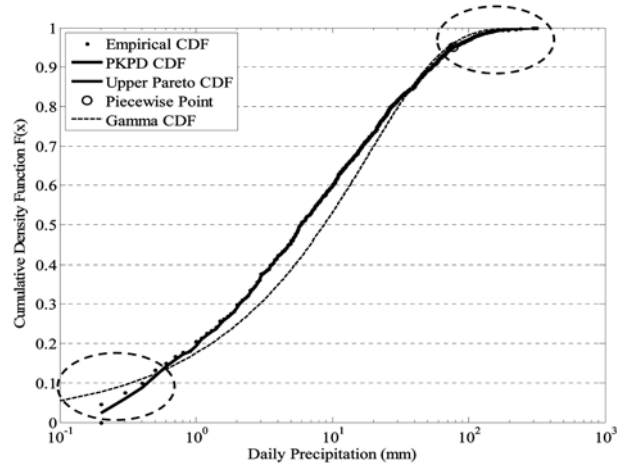


Fig. 3 An example of Piecewise Kernel-Pareto Distribution for Daily Precipitation

며 예를 들어 비초과확률 0.9이상에서는 GPD분포를 활용하고 0부터 0.9까지는 핵밀도함수를 활용하여 강수량을 모의하게 된다. 이렇게 되면 2개의 확률밀도함수가 적용되게 되며 각각의 구간에서는 독립된 확률밀도함수로부터 강수량이 모의되게 된다. 이러한 확률분포형의 형태를 불연속확률밀도함수(piecewise probability density function)이라고 하며 최근에 혼합분포와 함께 복합적인 성분을 동시에 모의하기 위해서 통계학 분야에서 많은 연구가 진행되고 있다.

본 연구에서 제안하는 PKPD 분포형의 개념을 그림 3을 통해서 나타내었다. 그림 3에서는 0.95이상의 극대강수량에 대해서는 GPD분포가 적용되었고 0부터 0.95까지는 핵밀도함수가 적용되었다. 이를 경험적 누기확률밀도함수와 일강수량 모의발생에 많이 이용되는 Gamma분포형과 비교하여 나타내었다. 그림에서 경험적인 누기확률밀도함수와 비교결과 PKPD CDF의 경우 분포형의 양 끝단에서 매우 일치되는 결과를 기존의 Gamma분포형에서는 왜곡된 모습을 보여주고 있다. 그림에서 보면 Gamma 분포형은 0.9이상에서 1에 빠르게 수렴하고 있는 반면에 PKPD 분포형은 원자료가 가지는 극치특성을 매우 잘 묘사하고 있다. 분포형에서 Tail의 거동은 극치강수량을 재현하는데 매우 민감한 요소로서 Gamma 분포형은 극치강수량을 모의하는데 무리가 있음을 그림으로부터 확인할 수 있다. 이와 더불어 0.1 이하의 누기확률부분에서도 PKPD 분포는 경험적인 누기확률밀도함수와 유사한 특성을 보이고 있으나 Gamma 분포는 꼬리가 길게 늘어져 실제 강수량이 가지는 특성이 왜곡될 가능성이 크다 하겠다.

3. 모의기법 적용

본 연구에서 제안한 방법론의 적합성은 한강유역에 서울,

Table 2. Weather stations used in this study

No.	Station Code	Station Name	Location(degree)		Observation period until 2008(year)	
			Latitude	Longitude	Daily	Hourly
1	101	Chuncheon	37.90	127.73	43	43
2	105	Gangneung	37.75	128.88	97	48
3	108	Seoul	37.57	126.95	101	48

강릉, 춘천 지점의 기상청 강수량을 대상으로 평가하였다. 3 지점 모두 30년 이상의 자료를 보유하고 있으며 특히 강릉 지점의 경우 태풍 루사 등의 영향으로 기존의 Gamma분포를 활용한 Markov 방법을 통해서서는 극치강수량을 모의하는

데 매우 어려움이 큰 것으로 알려지고 있다. 본 연구에서 사용된 강수지점에 대한 정보는 표 2와 같다. 모든 지점에서 30년 이상에 일강수량을 보유하고 있으며 본 연구에서는 일강수량을 대상으로 모형을 검토하였다. 일반적으로 Markov

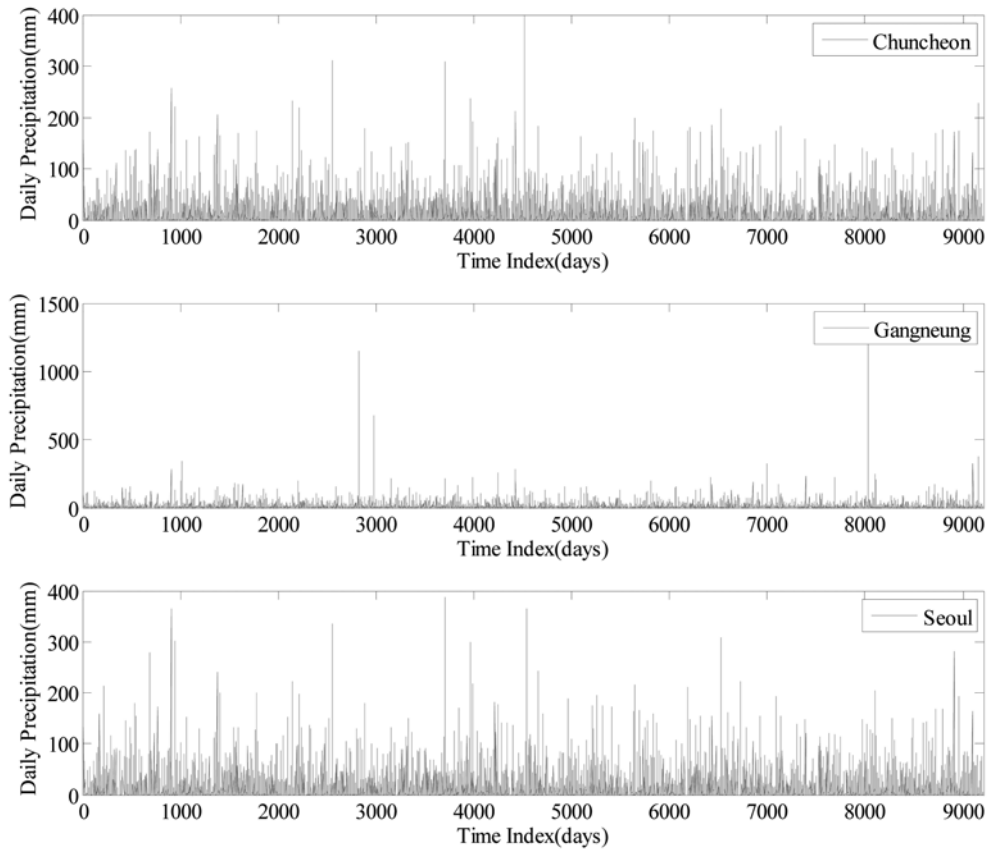


Fig. 4 Sequence of summer precipitation for three stations using PKPD based Markov Chain Model

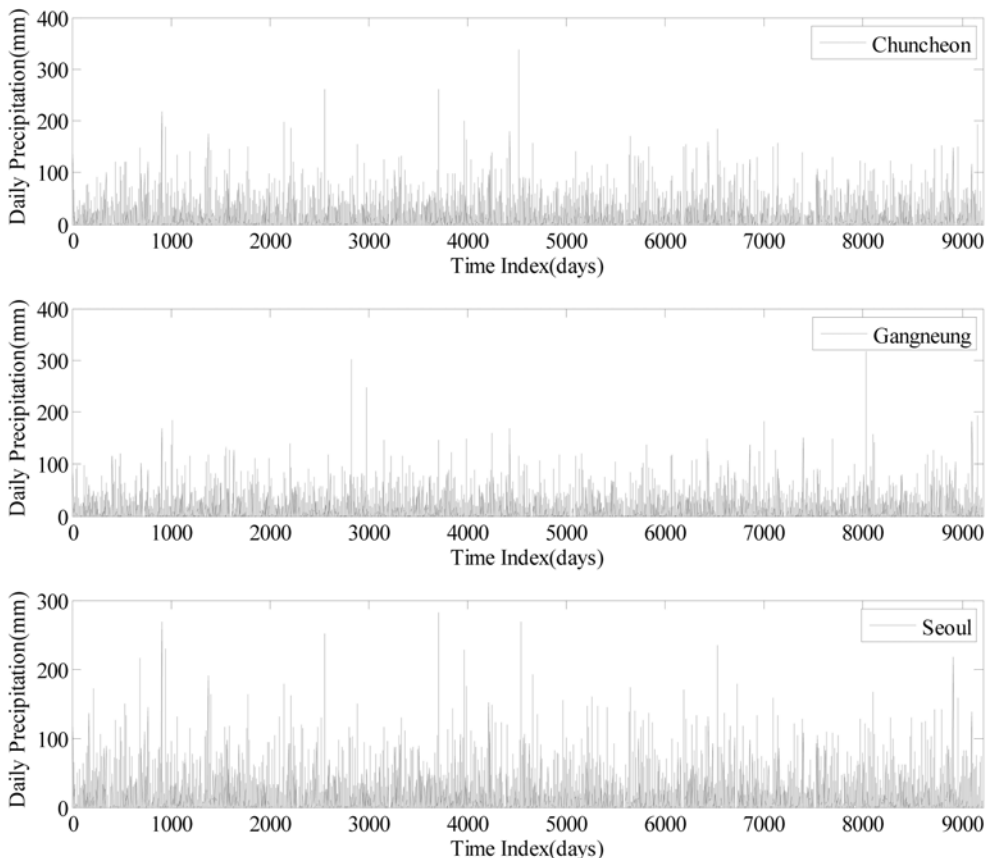


Fig. 5 Sequence of summer precipitation for three stations using Gamma Distribution based Markov Chain Model

Chain 모형은 계절별로 모형을 구축하여 모의를 실시하게 된다. 본 연구에서 제시한 방법론에 대해서 일반적인 강수 재현 능력과 여름에 극치강수량 재현 능력을 동시에 평가하기 위해서 7-9월 여름강수량과 10-12월 가을강수량을 대상으로 모의 결과를 평가하였다. 모형의 적합성은 여름강수량을 위주로 평가하여 설명한 후 가을강수량에 대해서는 결과 위주로 나타내었다.

그림 4는 한강유역 3개의 관측소에 대해서 PKPD 확률분포에 기반한 Markov Chain 모형을 이용해 여름강수량을 모의한 결과를 나타낸다. 모의기간은 7-9월 3개월간 총 92일을 100년 모의한 결과이다. 그림에서 특이할 점은 강릉관측소에 대해서 모의된 일강수량이 900 mm를 상회하는 값으로 모의되고 있다는 점이다. 이는 태풍 루사 당시 발생했던 일최대강수량을 반영한 결과로 효과적으로 극치값이 재현되고 있는 것으로 판단된다. 그림 5는 Gamma 확률분포에 기반한 Markov Chain 모형을 이용해 여름강수량을 모의한 결과이다. PKPD 확률분포에 비해서 강수량이 과소 추정되고 있음을 강수시계열로부터 확인할 수 있다.

모형의 적합성을 평가하기 위해서 가장 우선적으로 검토한 사항은 모의된 강수계열이 관측강수계열의 천이확률을 재현할 수 있느냐는 것이다. 표 3은 관측강수계열과 모의강수계열로부터 추정된 천이확률을 나타내며 모의강수계열로부터 추정된 천이확률과 관측값으로부터 추정된 천이확률과 거의 동일한 값으로 모의되는 것을 확인할 수 있다.

다음으로 모의된 강수계열로부터 4가지 통계치를 추출하여 관측강수량을 비교하였다. 모형의 적합성을 평가하기 위해서 사용되는 통계치는 중간값, 표준편차, 왜곡도, 100년 빈도 강수량이다. 그림 6은 이들 통계치를 관측값과 본 연구에서 비교 대상인 Gamma 분포형 기반 Markov Chain 모형과

Table 3. Comparison of transition probability between observation and simulation

Station	Transition Probability	Observation	Simulation
Chuncheon	p01	0.26	0.26
	p10	0.44	0.45
Gangneung	p01	0.30	0.30
	p10	0.41	0.41
Seoul	p01	0.26	0.26
	p10	0.40	0.41

비교한 그림이다. 그림에서 보는 바와 같이 본 연구에서 제안하고 있는 PKPD 기반 Markov Chain 모형이 1-3차 모멘트까지 관측치의 특성을 매우 효과적으로 재현하는 것을 확인할 수 있다. 반면 Gamma 분포 기반 Markov Chain 모형은 2차 모멘트 이상에서 강수량의 특성을 효과적으로 재현하지 못하고 있다. 그림 7에서는 극치강수량의 재현 특성을 보다 상세하게 비교하기 위해서 관측치, PKPD분포, Gamma분포의 누가핵밀도함수를 비교하였다. 분포형의 Tail 거동을 자세히 평가하기 위해서 비초과확률 0.9이상만을 도시하여 나타내었다. 그림에서 보는바와 같이 Gamma분포는 100 mm이상에서 PKPD분포에 비해 1에 빠르게 수렴하고 있으며 결국 이러한 특징이 극치강수량을 재현하는데 어려움으로 작용하게 된다. 그림 6에서 나타난 통계치들의 특성을 상세히 비교하여 표 4에 나타내었다.

표 5는 가을강수량을 모의한 결과로서 여름강수량과 동일하게 평균, 표준편차, 극치강수량, 왜곡도 등을 비교하였다. 10-12월에 100년 빈도 강수량은 수자원설계관점에 중요하지 않지만 여름계절과 모의결과와 일관성 있는 평가를 위해서

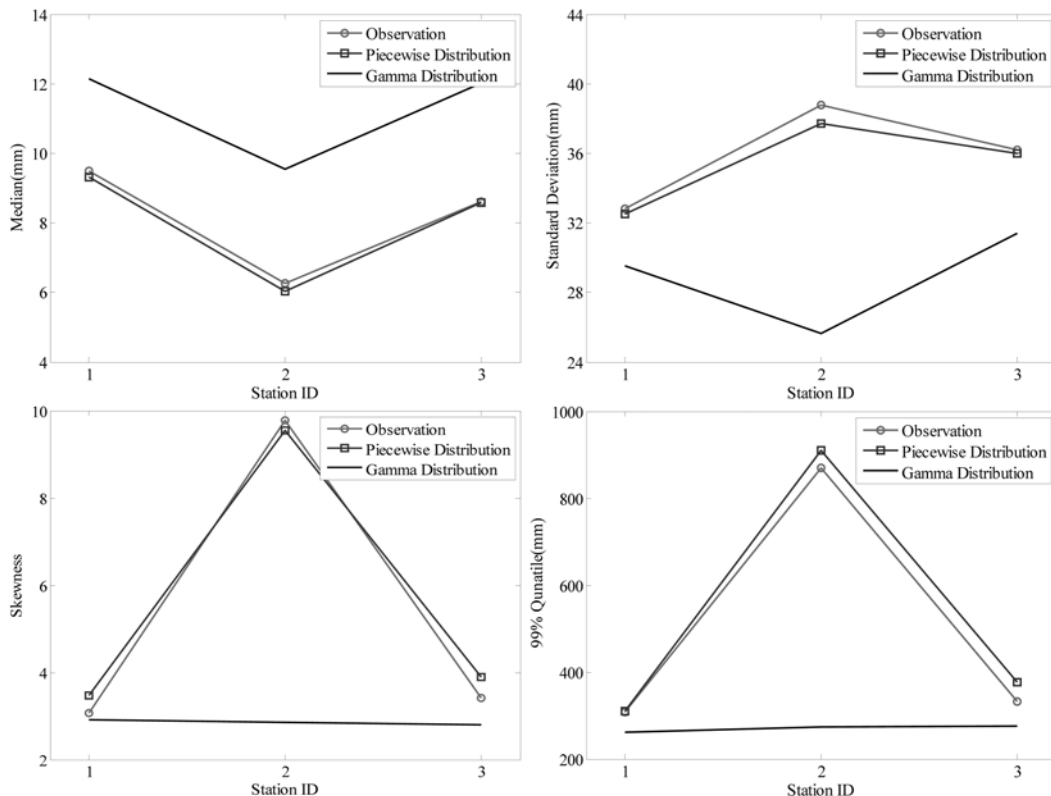


Fig. 6 Assessment of simulation performance of PKPD based Markov Chain Model (July-August-September)

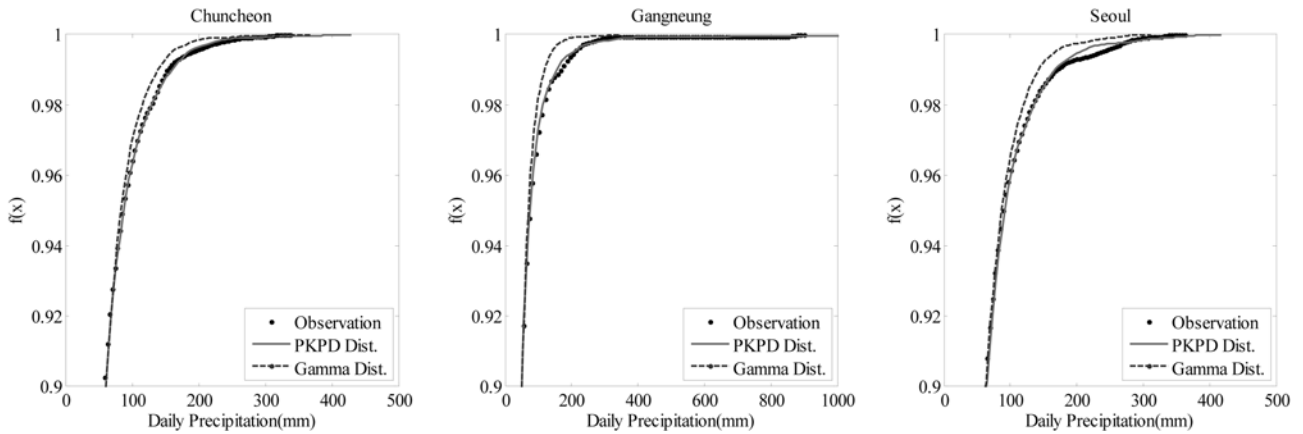


Fig. 7 Comparison of cumulative kernel density function for extreme precipitation

Table 4. Quantitative analysis of simulation performance of PKPD based Markov Chain Model (July-August-September)

Stations	Statistics	Observation	PKPD	Gamma Dist.
Chuncheon	100-Year Design Rainfall (mm)	308.5	310.3	261.4
Gangneung		870.5	909.5	274.6
Seoul		332.8	376.3	275.5
Chuncheon	Median (mm)	9.5	9.3	12.1
Gangneung		6.3	6.0	9.5
Seoul		8.6	8.6	12.0
Chuncheon	Mean (mm)	22.6	22.1	22.9
Gangneung		18.5	18.0	19.2
Seoul		22.8	22.7	23.7
Chuncheon	Standard Deviation (mm)	32.8	32.5	29.5
Gangneung		38.8	37.7	25.6
Seoul		36.2	36.0	31.4
Chuncheon	Skewness	3.1	3.5	2.9
Gangneung		9.8	9.6	2.9
Seoul		3.4	3.9	2.8

Table 5. Quantitative analysis of simulation performance of PKPD based Markov Chain Model (October-November-December)

Stations	Statistics	Observation	PKPD	Gamma Dist.
Chuncheon	100-Year Design Rainfall (mm)	70.6	65.8	45.3
Gangneung		304.0	295.0	151.0
Seoul		79.4	79.1	47.3
Chuncheon	Median (mm)	2.5	2.6	3.6
Gangneung		4.4	4.4	7.4
Seoul		2.8	3.0	3.9
Chuncheon	Mean (mm)	5.4	5.5	5.8
Gangneung		13.4	12.9	13.7
Seoul		5.8	5.9	6.2
Chuncheon	Standard Deviation (mm)	8.0	7.8	6.5
Gangneung		22.8	22.2	17.5
Seoul		8.7	8.6	6.8
Chuncheon	Skewness	3.5	3.4	2.2
Gangneung		5.1	5.3	2.8
Seoul		3.8	4.0	2.1

가을강수량의 100년빈도를 추정하여 비교하였다. 모든 면에서 기존 Gamma 기반 Markov Chain 모형에 비해 우수한 재현능력을 보여주고 있다.

4. 결 론

강수량 모의발생 기법은 수공학에서 다양한 방향으로 활용되고 있다. 특히 수공구조물 설계와 같이 불확실성을 가지는 미래 수자원을 계획하는데 있어서 필수적인 기법이다. 그러나 기존 Markov Chain 모형을 통한 일강수량 모의에서 가장 큰 문제점은 극치강수량을 재현하기 어렵다는 점이다. 이러한 문제점은 수자원설계에서 홍수 및 이수 안전도를 정량적으로 해석하는데 어려움을 주고 있으며 특히 일강수량 모의기법을 통해서 과소 추정된 빈도강수량을 이용한 수공구조물 설계 시에 신뢰성이 결여되는 문제점이 있다. 이러한 점에서 본 연구에서는 기존 Gamma 분포의 문제점을 개선할 수 있는 새로운 해석방법을 제안하였다. 즉, Markov

Chain 모형에서 일강수량에 평균적인 특성과 극치특성을 동시에 재현할 수 있도록 불연속 Kernel-Pareto Distribution 기반의 일강수량모의기법을 개발하였다. 본 연구를 통해서 얻은 결론은 다음과 같다.

1. 기존 Gamma 분포 기반 Markov Chain 모형은 일반적으로 평균적인 특성을 잘 재현하는 것으로 알려져 왔으나 여름의 일강수량 모의 시 Median 값을 효과적으로 모의하지 못하는 문제점이 나타났다. 여름의 일강수량의 평균적인 특성뿐만 아니라 2차모멘트 이상에서 관측치의 특성을 반영하지 못하고 평활화 되는 문제점이 나타났다. 그러나 가을의 일강수량에선 여름에 비해 많이 개선된 결과를 보여주고 있었다.
2. 본 연구에서 제시한 불연속 Kernel-Pareto 분포형 기반 Markov Chain 모형은 여름의 일강수량 모의 시 강수계열의 평균적인 특성뿐만 아니라 표준편차 및 왜곡도의 경우에도 관측치의 통계특성을 매우 효과적으로 재현하는 것으로 나타났다. 이를 100년 빈도 확률강수량을 통해 다시

한 번 검증한 결과 실측 강수량으로부터 추정된 값과 매우 유사한 값으로 모의되었다. 가을 일강수량의 경우에도 기존 Gamma 분포기반 Markov Chain 모형에 비해 모든 통계치에서 개선된 결과를 제시해 주었다.

본 연구에서 제시한 방법론은 전제적으로 기존 Markov Chain 모형에 비해 극치강수량을 재현하는데 유리한 기법으로 판단된다. 또한 극치강수량을 일반강수량으로부터 분리하여 모의함으로써 평균 및 중간값 등 낮은 차수에 모멘트를 더욱 정확하게 재현할 수 있는 부가적인 장점도 확인할 수 있었다.

본 연구에서 제시한 방법론은 단일지점(single site)에 대해서 적용이 가능한 모형으로 개발되었다. 그러나 추후 연구로서 동일 유역에 면적강수량을 유도하기 위해서 유역상관성(spatial coherence)을 재현할 수 있는 다지점(multi-site) 강수모의기법으로 개발이 필요할 것으로 판단된다. 이와 더불어 기후변화 연구를 위해서 기후변화 시나리오를 입력자료로 활용할 수 있는 비정상성 모형으로 확장이 필요할 것으로 판단된다.

감사의 글

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

참고문헌

권현한, 문영일, 최병규, 윤용남(2005) 신뢰성 분석을 통한 기존 댐 재개발의 적정규모 결정의 관한 연구, **한국수자원학회 논문집**, 한국수자원학회, Vol. 38, No. 2, pp. 97-110.

권현한, 김병식(2009) 비정상성 Markov Chain Model을 이용한 통계학적 Downscaling 기법 개발, **한국수자원학회 학술발표회 논문집**, 한국수자원학회, pp. 213-225.

김병식, 강경석, 서병하(1999) 일 강우량의 모의 발생을 통한 갈수유량 계열의 산정 및 빈도분석, **한국수자원학회논문집**, 한국수자원학회, Vol. 32, No. 3, pp. 265-279.

김병식, 서병하, 김남원(2003) 전이함수모형과 일기 발생모형을 이용한 유역규모 기후변화시나리오의 작성, **한국수자원학회**

논문집, 한국수자원학회, Vol. 36, No. 3, pp. 345-363.

문영일, 오태석, 김종석, 정민수(2006) 비매개변수적 Kernel Function을 이용한 지역빈도해석, **한국수자원학회 학술발표회 논문집**, 한국수자원학회, pp. 1492-1496.

엄명진, 조원철, 허준행(2008) GPD 모형 산정을 위한 부분시계열 자료의 임계값 산정방법 비교, **한국수자원학회논문집**, 한국수자원학회, Vol. 41, No. 5, pp. 527-544.

이순혁, 맹승진, 류경식, 지호근(2001) Generalized extreme value, Generalized logistic 및 Generalized pareto 분포 모형에 의한 홍수빈도분석, **한국수자원학회 학술발표회 논문집**, 한국수자원학회, pp. 140-145.

최병규, 문영일, 오태석, 박래건(2008) 비동질성 Markov 모형의 시간강수량 모의 발생을 이용한 IDF 곡선 및 홍수빈도곡선의 유도, **한국수자원학회논문집**, 한국수자원학회, Vol. 41, No. 3, pp. 251-264.

Ashkar, F. and Tatsambon, C.N. (2007) Revisiting some estimation methods for the generalized Pareto distribution, *Journal of Hydrology*, Vol. 346, pp. 136-143.

Haan, C.T., Allen, D.M., and Street, J.O. (1976) A markov chain model of daily rainfall, *Water Resour. Res.*, Vol. 12, No. 3, pp. 443-449.

Stehlik, J. and Bardossy, A. (2002) Multivariate stochastic downscaling model for generating daily precipitation series based on atmospheric circulation, *Journal of Hydrology*, Vol. 256, pp. 120-141.

Kwon, H.-H., Sivakumar, B., Moon Y.-I., and Kim B.-S. (2010) Assessment of change in design flood frequency under climate change using a multivariate downscaling model and a precipitation-runoff model, *Stochastic Environmental Research and Risk Assessment*, DOI: 10.1007/s00477-010-0422-z, in press.

Kwon, H.-H., Lall, U., and Obeysekera, J. (2009) Simulation of daily rainfall scenarios with interannual and multidecadal climate cycles for South Florida, *Stochastic Environmental Research and Risk Assessment*, Vol. 23, No. 7, pp. 879-896.

Nord, J. (1975) Some applications of markov chains, proceedings fourth conference on probability and statistics in atmospheric science, Tallahas, pp. 125-130.

Rosenblatt, M. (1956) Remarks on some nonparametric estimates of a density function. *Ann. Math. Statist.* 27, pp. 832-837.

Singh, V.P., Ahmad, M., and Sherif, M.M. (2003) A comparative evaluation of the estimators of the 2-parameter generalized Pareto distribution. *Water Engineering Research*, Vol. 4, No. 3, pp. 155-173.

(접수일: 2011.1.10/심사일: 2011.2.10/심사완료일: 2011.2.10)