

# 비동질성 Markov 모형에 의한 시간강수량 모의 발생과 천이확률을 이용한 강우의 시간분포 유도

## Simulation of Hourly Precipitation using Nonhomogeneous Markov Chain Model and Derivation of Rainfall Mass Curve using Transition Probability

최 병 규\* / 문 영 일\*\* / 오 태 석\*\*\* / 박 래 건\*\*\*\*

Choi, Byung-Kyu / Moon, Young-Il / Oh, Tae Suk / Park, Rae Gun

### Abstract

The observed data of enough period need for design of hydrological works. But, most hydrological data aren't enough. Therefore in this paper, hourly precipitation generated by nonhomogeneous Markov chain model using variable Kernel density function. First, the Kernel estimator is used to estimate the transition probabilities. Second, wet hours are decided by transition probabilities and random numbers. Third, the amount of precipitation of each hours is calculated by the Kernel density function that estimated from observed data. At the results, observed precipitation data and generated precipitation data have similar statistic. Also, rainfall mass curve is derived by calculated transition probabilities for generation of hourly precipitation.

**keywords** : Kernel density function, Markov model, hourly precipitation

### 요 지

수공구조물의 설계를 위해서는 충분한 기간의 관측자료가 필요하지만, 우리나라의 수문자료는 대부분 충분한 수의 관측자료를 보유하고 있지 못하는 실정이다. 따라서 본 연구에서는 핵밀도함수를 이용한 비동질성 Markov 모형을 통해 시간강수량 자료를 모의하였다. 첫 번째로 시간강수량 자료에 변동핵밀도함수를 이용하여 천이확률을 산정하였으며, 두 번째로 난수와 천이확률을 통해 강수가 발생하는 시간을 결정하였다. 세 번째로 강수가 발생한 시간의 강수량의 크기를 핵밀도함수를 통해 추정하였다. 분석결과에서 모의된 시간강수량은 관측시간강수량과 비슷한 통계적 특성을 보이고 있는 것으로 나타났다. 또한, 시간강수량의 모의발생을 위하여 산정한 천이확률을 이용해 강수의 무차원시간분포곡선을 유도하였다.

**핵심용어** : 핵밀도함수, 마코프 모형, 시간강수량

\* 서울시립대학교 공과대학 토목공학과 박사과정

The University of Seoul, Siripdae-gil 13(Jeonngong-dong 90), Dongdaemun-gu, Seoul 130-743 Korea  
(e-mail: bkchoi@samaneng.com)

\*\* 교신저자 · 서울시립대학교 공과대학 토목공학과 교수

The University of Seoul, Siripdae-gil 13(Jeonngong-dong 90), Dongdaemun-gu, Seoul 130-743 Korea  
(e-mail: ymoon@uos.ac.kr)

\*\*\* 서울시립대학교 공과대학 토목공학과 박사과정

The University of Seoul, Siripdae-gil 13(Jeonngong-dong 90), Dongdaemun-gu, Seoul 130-743 Korea  
(e-mail: waterboy@uos.ac.kr)

\*\*\*\* (주) 삼안 수력부 대리

Saman Corporation, 1-20, Beyolyang-dong, Gwacheon-si, Gyeonggi-do 427-707 Korea  
(e-mail: rkpark@samaneng.com)

## 1. 서 론

안전한 수자원의 확보 및 홍수 방어와 같은 치수 계획을 수립하기 위해서는 수문관측자료의 확보가 필수적이라 할 수 있다. 특히, 수재해를 방어하기 위해서는 긴 기간 동안에 관측된 자료를 활용하여 대상 구조물에 적절한 설계 기준을 설정하여야 한다. 주요한 국가하천의 경우에 설계빈도를 200년으로 설정하고 있으나 우리나라의 관측된 유량자료는 가장 긴 경우에도 50년 내외의 관측 자료를 보유하고 있는 실정이다. 따라서 부족한 관측 수문 자료의 확장을 위해 많은 노력을 기울여 오고 있다.

따라서 대표적인 수문자료 중에 하나인 강수량 자료에서 관측 기간이 짧은 자료를 확장하기 위한 많은 추계학적 모형들이 개발되어 왔으며, 현재까지 개발된 모형으로는 Markov 연쇄모형(Gabriel and Neumann, 1962), 이산형자기회귀이동평균 모형(Chang *et al.*, 1984), Point 과정 모형(Todorovic and Woolhiser, 1969)이 있다.

수문 자료의 일관측자료 모의를 위하여 Bruhn *et al.*(1980)과 Richardson (1981)은 일별 기상자료를 대상으로 모의발생 모형을 개발하였다. 이원환과 이재준(1985)은 2변량 모형을 이용하여 건조기간과 습윤기간의 단일지점에 대하여 일강수량을 모의하였고, 이재준(1987)과 허준행 등(1997)은 선형 또는 다중회귀식을 사용하여 관측소간의 공간적인 가중치를 주어 각 지점의 일강수량을 모의하였다. 한국수자원공사(1999)에서는 Markov 연쇄이론을 이용해 강수량을 모의하여 강수량을 확충한 후 이를 강우-유출모형에 이용하기도 하였다. 차영일(2004)은 다변량 핵밀도함수를 이용한 비동질성 Markov 모형을 이용하여 일강수량을 모의하여 평가하였다. 강경석(2000)은 자료의 유역 내 각 지점의 공간적 특성을 고려하여 강수량을 모의 발생하였으며, 이창훈과 김승(1995)은 주성분 분석과 회귀분석을 이용하여 계절적인 특성이 고려된 월강수량 자료로부터 결측 구간을 보완하여 년 월평균 강수량을 추정하였지만 이는 일강수량 자료의 확충을 하기 위한 것이 아닌 월강수량 자료를 보완하고 이를 토대로 년 및 월평균 강수량 추정을 위한 것이라 할 수 있다.

수문 자료의 시간관측 자료의 모의를 위해 강태호 등(1999)은 장래 입력강수량을 시간강수량에 의한 시계열 모형으로 구축한 바 있으며 이재준과 이정식(2002)은 시간강수계열의 군집성을 고려한 Point 과정의 추계학적 모형인 Neyman-Scott형의 군집 과정을 통하여 강수사상의 군집성을 고려할 수 있는 강수발생과정을

모형화한 연구를 수행하였다.

따라서 본 연구에서는 핵밀도 함수를 이용한 비동질성 Markov 연쇄이론을 통해 시간강수량을 모의하였다. 시간강수량의 모의를 위하여 적용한 비매개변수적 방법은 Lettenmaier and Burges(1977)가 자료의 모멘트를 만족시키고 단일 첨두를 가지는 밀도함수에 맞도록 Cubic Spline Fit를 확률밀도함수에 이용함으로써 소개되기 시작하였으며, Yakowitz(1985)와 Adamowski(1985)이 각각 독자적으로 핵밀도 함수와 관련된 연구를 통해 수문학에 적용되었다. 이 후, Adamowski and Labatiuk(1987), Bardsley(1988, 1989), Lall *et al.*(1993), Moon and Lall(1994), Moon and Lall(1996) 등에 의하여 수문 분야와 연관성을 갖으며 발전되어 왔다. 이 중에서 극치 수문 자료를 이용한 빈도해석 등에 활용하기 위한 변동핵밀도함수(variable kernel density function)가 확률홍수량과 확률갈수량 및 확률강수량 산정에 활용되었다(Lall *et al.*, 1993; Moon and Lall, 1994; Moon and Lall, 1996; 차영일 등, 2006).

따라서 본 연구에서는 핵밀도함수를 활용한 비동질성 Markov 모형을 통해 우리나라에서 관측된 시간강수량의 모의를 수행하여 평가하였다. 또한, 시간강수량의 모의를 위해 핵밀도함수에 의해 구성된 천이확률을 이용하여 강수의 시간분포 모형을 구축하여 평가하였다.

## 2. 핵밀도 함수와 비동질성 Markov 연쇄 모형의 이론

본 연구에서는 핵밀도함수를 이용한 비동질성 Markov 모형을 통해 시간강수량 자료의 모의를 수행하였다. 관측된 시간강수량의 발생시기와 강수량에 핵밀도함수를 적용하여 천이확률을 구성하고 비동질성 Markov 연쇄과정을 통해 시간 강수량을 모의하였다.

### 2.1 핵밀도 함수의 이론

비동질성 Markov 모형을 이용해 시간강수량을 모의하기 위해서는 관측 자료로부터 천이확률을 구성하여야 한다. 본 연구에서는 천이확률의 구성을 위해 Rajagopalan 핵함수를 이용하였으며, 강수량의 결정을 위해 Epanechnikov 핵함수를 적용하였다. 또한, 광역폭의 선택을 위해 Scott(1992)가 제안한 LSCV(least squares cross validation) 기법을 이용하였다.

#### 2.1.1 핵밀도 함수

Breiman *et al.*(1977)은 고정 핵밀도함수 추정법의 특성과 자료의 지역적인 밀도를 고려하는 K-th Nearest Neighbor방법을 결합한 변동 핵밀도함수 추정법을 제

안하였다. 즉,  $K(x)$ 를 핵함수라 하고  $K$ 를 양의 정수로 놓고,  $d_{j,K}$ 를 한 개의 자료  $x_j$ 에서 그 나머지 자료  $(n-1)$ 개 중에서  $k$ 번째로 가까운 지점에 있는 자료까지의 거리라고 하면, 변동 핵밀도함수 추정법은 다음 Eq. (1)과 같이 정의 할 수 있다.

$$f(x) = \frac{1}{nh} \int \frac{1}{d_{j,k}} K\left(\frac{x-X_j}{hd_{j,k}}\right) dx \quad (1)$$

여기서  $hd_{j,k}$ 는 양의 변동 광역폭이다. 이러한 핵함수(kernel function)는 다음과 같은 Eqs. (2)~(4)의 가정을 만족하여야 한다.

$$\text{면적} : \int K(t)dt = 1 \quad (2)$$

$$\text{평균} : \int t K(t) dx = 0 \quad (3)$$

$$\text{분산} : \int t^2 K(t)dt = \text{constant} \quad (4)$$

여기서,  $t = (x - X_i)/h$ 이고, 이 때  $x$ 는 임의의 실수이며,  $X_i$ 는 실 관측된 자료이다. 일반적인 핵함수는 Eqs. (2)~(4)와 같은 특성을 가지며 대개  $t=0$ 에서 최대치를 가지며 연속적이고 대칭인 방정식의 형태를 가진다. 여러 핵함수의 종류는 Table 1과 같으며, 각 핵함수가 갖는 모양은 Fig. 1과 같다.

### 2.1.2 광역폭의 선택

비매개변수적 핵밀도 함수법에서 광역폭( $h$ )의 선택은 매우 중요한 문제로 다루어져 왔다. 적절한 광역폭을 선택하는 방법은 Rule of Thumb, LSCV(least

squares cross validation), BCV(biased cross validation), Cross Validated MISE(mean integrated squared error), Maximum Likelihood, 또는 Adamowski Criterion, Plug-In 기법, Smoothed Bootstrap 기법(Adamowski, 1985; Lall et al., 1993; Moon and Lall, 1994)등이 존재하며, 이로부터 최적의  $h$ 를 구할 수 있다.

본 연구에서는 비동질성 Markov 모형의 적용을 위해 관측자료로부터 천이확률을 결정하였다. 천이확률을 결정하기 위한 광역폭은 24시간에서 2160시간(90일) 사이에서 선택되어지도록 하였으며, LSCV 기법에 의해 242시간이 최적의 천이확률을 구성하기 위한 광역폭으로 선택되었다. LSCV 기법은 Rudemo(1982)와 Bowman(1984)에 의해 제안된 방법으로 다음 Eq. (5)와 같이 정의되는 함수  $R$ 을 최소로 하는 광역폭  $h$ 가 최적의 광역폭이라 정의하였다.

$$R(\hat{f}) = \int \hat{f}^2 - 2 \int \hat{f}f \quad (5)$$

여기서,  $f$ 는 임의의 확률밀도함수를 의미하며, 이에 대한 추정치를  $\hat{f}$  라고 한다. LSCV의 기본원리는 자료로부터 주어진  $h$ 의 구간에 대하여  $R(\hat{f})$ 를 최소화시키는  $\hat{f}$ 를 구하는 것이다. 여기서  $\int \hat{f}f$  항은  $\hat{f}$ 로부터 구할 수 있다.

$$\hat{f}_i(x) = \frac{1}{(n-1)h} \sum_{j \neq i} K\left(\frac{x-X_j}{h}\right) \quad (6)$$

위 Eq. (6)과  $E[\hat{f}] = \int \hat{f}f$ 의 개념을 이용하면 Eq. (5)는 Eq. (7)로 표현된다.

Table 1. Basic Equations of Kernel Function

Name	kernel function, $K(t)$
Rectangular	$\frac{1}{2}$ for $ t  < 1$ , 0 if otherwise
Gaussian	$\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right)$
Epanechnikov	$\frac{3}{4} (1 - \frac{1}{5}t^2) / \sqrt{5}$ for $ t  < \sqrt{5}$
Rajagopalan	$\frac{3h}{(1-4h^2)} (1-t^2)$ for $ t  \leq 1$
Cauchy	$\frac{1}{\pi(1+t^2)}$

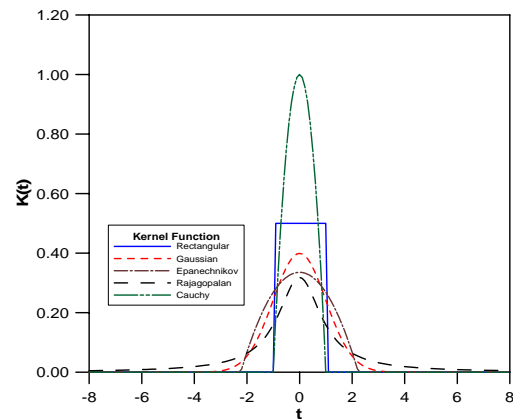


Fig. 1. Shapes of Kernel Function

$$M_0(h) = \int \hat{f}^2 - 2n^{-1} \sum_i \hat{f}_{-i}(X_i) \quad (7)$$

LSCV의 개념은 광역폭  $h$ 에 걸친  $M_0$ 를 최소화하는 것으로,  $M_0$ 의 값은 전적으로 관측 자료에 의존하게 된다.

## 2.2 비동질성 Markov 모형의 이론

### 2.2.1 비동질성 Markov 모형

비동질성 Markov 모형의 구성을 위해 One-Step의  $2 \times 2$  천이확률 행렬을 핵밀도함수를 적용하여 산정하였다. 1년 중에 1시간부터 8,760시간까지의 시간강수량 자료를 다음의 4가지 종류로 구분하였다. 첫 번째,  $t_{w_1}, t_{w_2}, \dots, t_{w_{n_w}}$ 는  $n_w$ 개의 습윤시간사상, 두 번째, 시간지수  $t_{d_1}, t_{d_2}, \dots, t_{d_{n_d}}$ 는  $n_d$ 개의 건조시간 사상, 세 번째,  $t_{wd_1}, t_{wd_2}, \dots, t_{wd_{n_{wd}}}$ 는 습윤시간사상에서 건조시간사상으로 바뀌는  $n_{wd}$ 개의 천이발생시간, 마지막으로  $t_{dw_1}, t_{dw_2}, \dots, t_{dw_{n_{dw}}}$ 는 건조시간사상에서 습윤시간사상으로 바뀌는  $n_{dw}$ 개의 천이발생시간으로 구분하였다. 여기서, 1년 중에 가장 마지막 시간인 8,760번째 시간은 그 다음연도의 1번째 시간과 연속되는 것으로 천이확률을 구성하게 된다.

다음 Fig. 2에서  $t$ 시간의 천이확률  $P_{wd}(t)$ 와  $P_{dw}(t)$ 는 비매개변수적 핵밀도함수 추정치를 이용하여 자료로

부터 구한다.  $P_{wd}(t)$ 는  $t_{wd_1}, t_{wd_2}$  등의 습윤시간사상에서  $t_{wd_1} + 1, t_{wd_2} + 1$  등의 건조시간사상으로 천이되는 것을 뜻하며,  $P_{dw}(t)$ 는  $t_{dw_1}, t_{dw_2}$  등의 건조시간사상에서  $t_{dw_1} + 1, t_{dw_2} + 1$  등의 습윤시간사상으로 천이되는 것을 의미한다. 따라서 각각의 연도별로 천이확률을 산정하여 최종적으로 이를 중첩시켜 강수의 발생상태를 결정하는 천이확률을 구성하게 된다. 각각의 습윤상태에서 건조상태로의 천이확률과 건조상태에서 습윤상태로의 천이확률은 다음의 Eqs. (8)~(11)을 통해 산정하게 된다.

$$\hat{P}_{dw}(t) = \frac{\sum_{i=1}^{n_{dw}} K\left(\frac{t - t_{dwi}}{h_{dw}}\right)}{\sum_{i=1}^{n_d} K\left(\frac{t - t_{di}}{h_{dw}}\right)} \quad (8)$$

$$\hat{P}_{wd}(t) = \frac{\sum_{i=1}^{n_{wd}} K\left(\frac{t - t_{wdi}}{h_{wd}}\right)}{\sum_{i=1}^{n_w} K\left(\frac{t - t_{wi}}{h_{wd}}\right)} \quad (9)$$

$$P_{dw}(t) + P_{ww}(t) = 1 \quad (10)$$

$$P_{dw}(t) + P_{dd}(t) = 1 \quad (11)$$

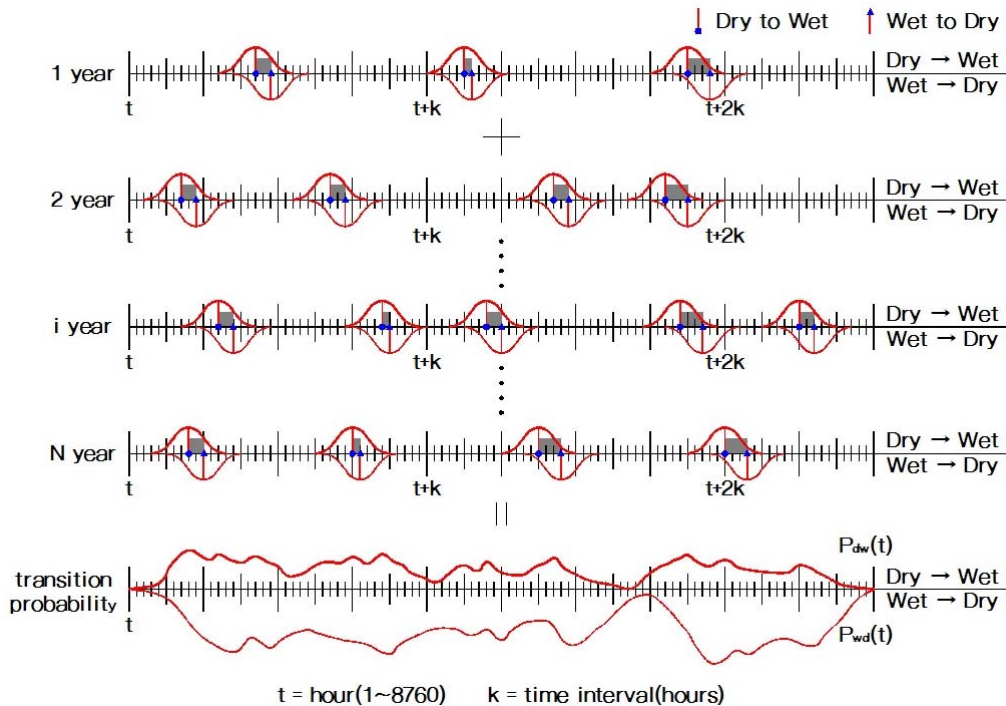


Fig. 2. The Construction Method of Transition Probabilities for Each Duration

여기서,  $h_{wd}$ 와  $h_{dw}$ 는 각각 습윤시간에서 건조시간으로, 건조시간에서 습윤시간으로 천이될 때의 광역폭이고,  $t_{wi}, t_{di}, t_{wdi}, t_{dwi}$ 는 각각  $i$ 번째 습윤시간, 건조시간, 습윤시간에서 건조시간, 건조시간에서 습윤시간으로 천이되는 시간을 의미하며,  $P$ 는 추정된 확률값을 의미한다. 또한,  $nwd$ 는 관측자료에서 습윤시간에서 건조시간으로 천이 되는 사상수,  $ndw$ 는 건조시간에서 습윤시간으로 천이 되는 사상수,  $nd$ 는 관측자료에서 건조시간개수,  $nw$ 는 습윤시간개수,  $K(\cdot)$ 는 핵함수,  $h(\cdot)$ 는 광역폭,  $t$ 는 구하고자 하는 시간이며,  $t()$ 는 위에서 설명한 것과 마찬가지로 ()안의 것이 발생한 시간이다. 이와 같이, 천이확률이 구성되면 각 시간별로 난수를 발생시켜 현재상태에 따른 조건부 천이확률로부터 다음 시간의 건조 또는 습윤상태를 결정하게 된다.

### 2.2.2 시간강수량의 크기 결정

시간강수량의 모의과정에서 습윤시간이 결정되면 다음으로 발생하는 강수량의 크기를 산정하여야 한다. 시간강수량의 크기를 결정하기 위해서 Fig. 3과 같이 첫 번째로 강수가 발생한 시간의 광역폭( $h_p$ )을 결정하여,

선정된 광역폭에서 관측된 시간강수량 자료들의 핵밀도 함수로부터 강수량의 크기를 모의발생시키게 된다.

따라서 건조시간을 제외한 습윤시간에서는 앞에서 설명한 LSCV기법(Rajagopalan et al., 1996)을 통해 광역폭을 선정하게 된다. 선정된 광역폭에 포함되어 있는 강수사상에 핵함수를 통해 확률밀도함수를 추정하고 추정된 밀도함수에 난수를 대응시켜 강수량의 크기를 결정하게 된다. 즉, 습윤시간  $t$ 에서 강수량의 Eq. (12)의 조건부확률밀도함수에 난수를 대응시킨다.

$$\hat{f}(y | t) = \frac{\frac{1}{y h_{LY}} \sum_{i=1}^{np} K_c \left( \frac{\log(y) - \log(y_i)}{h_{LY}} \right) \cdot K \left( \frac{t - t_i}{h_p} \right)}{\sum_{i=1}^{np} K \left( \frac{t - t_i}{h_p} \right)} \quad (12)$$

여기서,  $\hat{f}(y) = \frac{1}{np} \sum_{i=1}^{np} \frac{1}{h_{LY} y} K_c \left( \frac{\log(y) - \log(y_i)}{h_{LY}} \right)$ 가 된다. 또한, 난수  $U[0, 1]$ 는 핵함수  $K_c(\cdot)$ 에 대응하는 값이므로 모의발생된 강수량은  $y^* = \exp[\log(y_j) + U h_{LY}]$

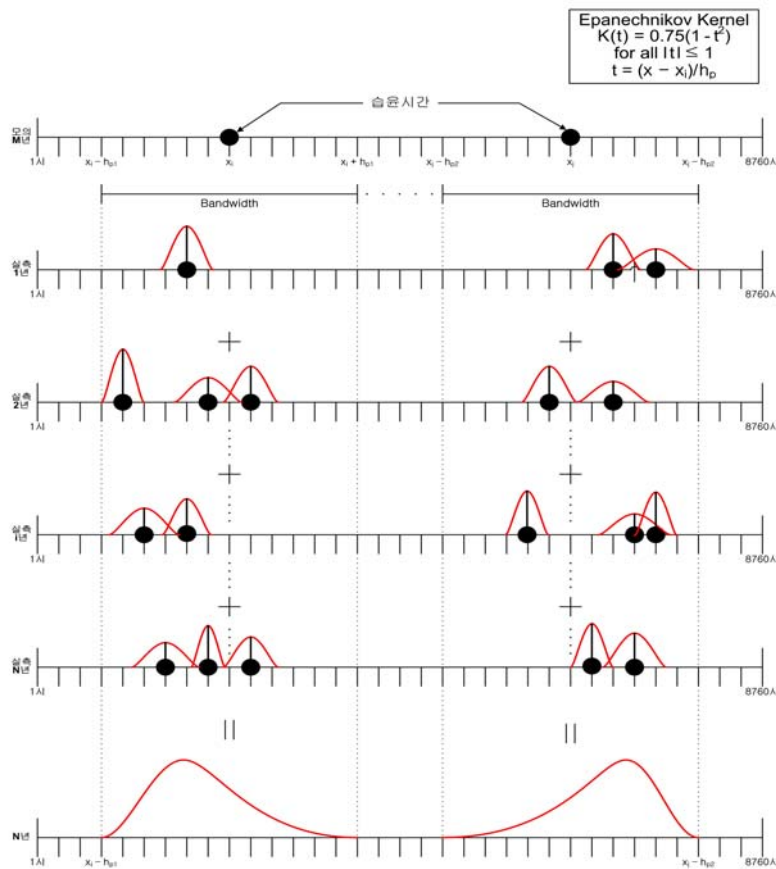


Fig. 3. Determination of Generated Hourly Precipitation using Epanechnikov Kernel Function

이다. 이 때  $y_i$  는 교란을 목적으로 선택한 과거 관측치의 강수량이며,  $h_{L,T}$ 는 대수화된 자료의 광역폭(Bandwidth)이다.

### 3. 비동질성 Markov 모형을 이용한 시간강수량의 모의발생

비동질성 Markov 모형을 구성하고 평가하기 위해서 기상청 관할 추풍령 지점에서 관측된 시간강수량 자료를 활용하였다. 기상청 관할 추풍령 관측소는 북위 36°13', 동경 128°00', 해발높이 242.5m에 위치하고 있다. 1961년부터 시간강수량의 관측을 시작하여 2006년까지 총 46개년의 시간강수량 관측자료가 존재하고 있다. 그러나 1961년부터 1980년까지는 우기에 해당하는 5월부터 10월까지만 시간강수량을 관측하고 11월부터 익년 4월까지의 시간강수량을 관측하지 아니하였다. 그러나 비동질성 Markov 모형의 구성을 통해 시간강수량을 모의하기 위해서는 전기간에 걸친 관측자료가 필요하므로 전년에 걸쳐 시간강수량이 관측된 1981년부터 2006년까지의 26년 동안 관측된 시간강수량 자료를 이용하여 모의를 수행하였다.

비동질성 Markov 모형이 갖는 동질성 Markov 모형(Nord, 1975)과의 차이점은 다음과 같다. 첫 번째로 천이확률의 구성에 있어서 동질성 Markov 모형은 임의의 시간에서 강수가 발생하였을 경우에 다음시간의 강수상태를 결정하는 천이확률은 조건부확률과정을 통해 산정하게 되는데, 천이확률이 조건부확률을 산정하는 동일기간별로 같은 값을 갖게 된다. 이에 반해 비동질성 Markov 모형에서는 천이확률의 결정하는데 있어 핵밀도 함수를 이용함으로써 각 시간 단위별로 천이확률이 다르게 계산된다. 따라서 실제로 발생한 강수사상과 가장 유사한 천이확률을 구하게 되는 것이다. 그러므로 비동질성 Markov 모형은 실제 강수가 발생한 시간별로 천이확률을 계산하여 모의를 수행하므로 실측강수자료

와 모의강수자료가 동질성 Markov 모형보다 더 근사하다고 할 수 있다.

두 번째로 습윤사상에 발생하는 강수량의 결정은 동질성 Markov 모형은 관측된 강수량 중에서 무작위로 추출하는 방법을 이용한다. 따라서 모의된 결과가 관측값에 비해 과소 추정될 가능성이 크며 극한값의 추정이 어렵다. 반면에 핵밀도 함수를 이용한 비동질성 Markov 모형은 습윤 사상이 결정되면 습윤 사상이 발생한 주변에 변동핵밀도 함수를 이용하여 강수량을 결정함으로써 극한 강수 사상의 추정이 가능한 장점이 있다.

#### 3.1 관측자료를 이용한 시간별 천이확률 구축

비동질성 Markov 모형을 이용하여 추풍령 지점에서 관측된 1981년부터 2006년까지 26년동안 관측된 시간강수량자료를 입력자료로 구축하여 시간강수량을 관측 기간과 동일한 26년간의 모의를 수행하였다.

비동질성 Markov 모형을 통해 시간강수량을 모의하기 위해서는 관측자료로부터 건조시간에서 습윤시간 또는 습윤시간에서 건조시간으로의 천이확률을 계산하여야 한다. 본 연구에서는 Rajagopalan 핵밀도함수를 이용하여 천이확률을 계산하였다. 다음 Fig. 4는 건조시간에서 습윤시간으로의 천이확률이며, Fig. 5는 습윤시간에서 건조시간으로의 천이확률을 도시한 결과로써 각각의 시간에 따라 계산된 천이확률을 나타낸 것이다.

Fig. 4와 Fig. 5에서 우리나라의 겨울에 해당하는 시간에는 건조시간에서 습윤시간으로 천이되더라도 그 다음시간에는 다시 건조시간으로 천이될 가능성이 큰 것으로 나타났으며, 여름인 우기에 해당하는 시간에는 반대의 경향을 보이는 것으로 나타났다. 또한, Fig. 4와 Fig. 5는 발생한 천이확률은 각 연도와 시간별로 다른 값을 갖게 되며, 추후에 시간강수량의 모의발생을 위해서는 각연도별로 계산된 천이확률을 Fig. 6 및 Fig. 7과 같이 중첩시켜 강수량을 발생여부를 결정하게 된다.

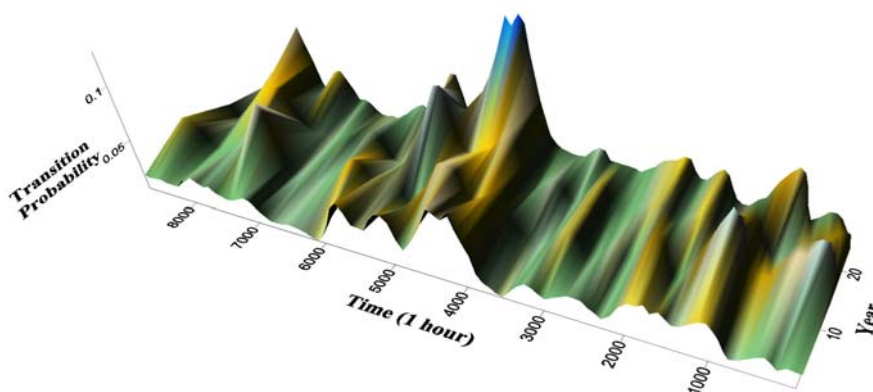


Fig. 4. Transition Probabilities Form Dry Condition to Wet Condition



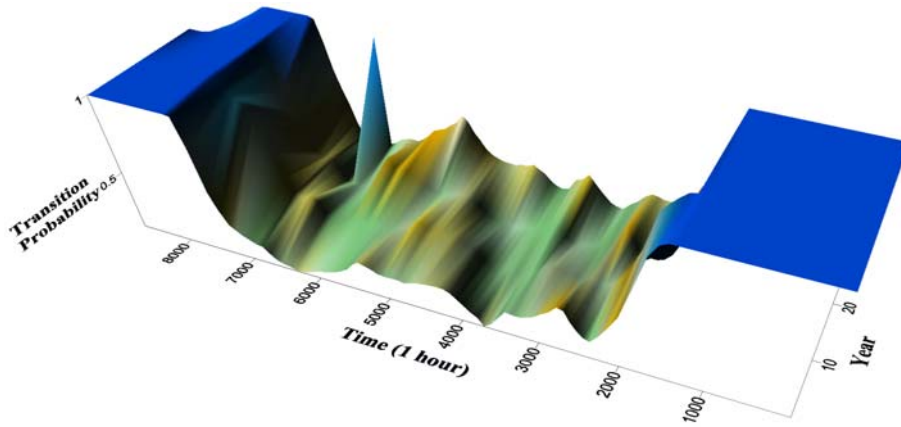


Fig. 5. Transition Probabilities Form wet Condition to Dry Condition

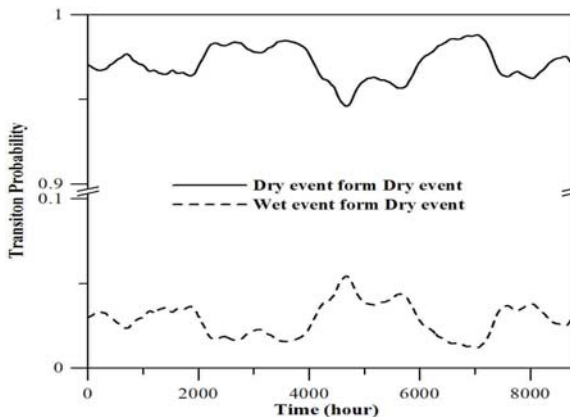


Fig. 6. Transition Probabilities of Superposition from Dry Condition

Fig. 6은 이전 시간에 강우가 발생하지 않은 건조사상일 때, 바로 다음 시간에 강우의 발생 여부를 판단하는 천이확률을 시간별로 중첩시켜 산정한 결과이다. 즉, 건조시간으로부터 습윤시간이나 건조시간으로 천이되는 확률을 도시한 결과로써, 이전시간이 건조시간일 때 다음 시간에 강우가 발생할 수 있는 확률(Fig. 6에서 Wet event form Dry event)로 연중 큰 차이는 보이지 않으나 여름에 해당하는 부분의 천이확률이 약간 더 크음을 알 수 있다. Fig. 7은 습윤시간으로부터 다음 시간에 강우가 발생하거나 발생하지 않는 경우를 도시한 결과이다. 우기에 해당하는 구간에서는 연속되는 강수가 발생할 수 있는 천이확률이 크게 산정됨을 알 수 있었으며, 건기에 해당하는 겨울철에는 이전 시간에 강수가 발생하였더라도 그 다음 시간에는 강수가 발생하지 않는 확률이 매우 크음을 알 수 있었다.

### 3.2 관측자료를 이용한 모의시간강수량 자료의 비교·검증

앞에서 산정된 천이확률을 이용하여 습윤시간을 결정하였으며, Epanechnikov 핵함수와 LSCV기법을 통

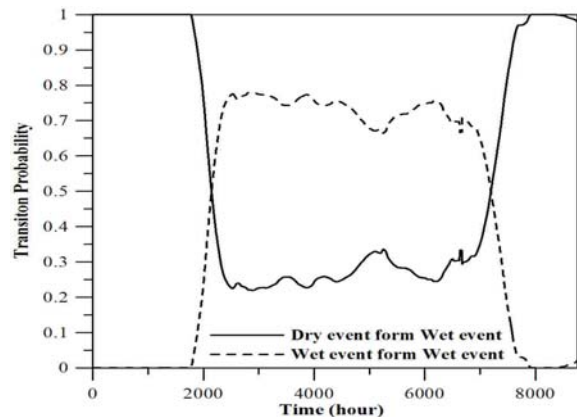


Fig. 7. Transition Probabilities of Superposition from Wet Condition

한 광역폭을 결정하여 강수량의 크기를 결정하였다. 즉, 1981년부터 2006년까지의 26년 동안 관측된 시간강수량 자료를 통해 천이확률을 구성하여 시간강수량의 모의를 수행하였다.

시간강수량의 관측자료와 모의자료의 통계적 특성을 비교하였다. 첫 번째는 관측강수량과 모의강수량에서 연최대치 1시간강수량을 추출하여 비교하였으며, 두 번째로 연강수량으로 합산하여 그 특성을 비교하였다. 세 번째로 관측자료와 모의자료의 전기간에 걸쳐 발생한 시간강수량의 통계치를 산정하여 비교하였으며, 네 번째로 습윤시간에 발생한 강수량의 특성을 비교해 주었다. 마지막으로 T-test와 F-test를 통해 시간최대강수량과 연강수량의 평균과 분산이 동일한 모집단에서 추출된 자료인지에 대한 검증을 수행하였다.

Table 2에서 보는 바와 같이, 1시간단위로 관측된 시간강수량의 모의 결과에서 관측치와 모의치가 크게 다르지 않은 것으로 나타났다. 다만, 표준편차가 약간 감소하는 경향을 갖는 것으로 보이나 큰 차이는 나타나지 않고 있다. 다음은 관측자료와 모의자료의 전기간과 습윤시간의 통계치를 나타낸 것이다.

**Table 2. Comparison of Annual Maximum Precipitation and Yearly Precipitation between Observed and Generated Data**

Classification	annual maximum precipitation		Yearly precipitation	
	Observed data	Generated data	Observed data	Generated data
Average	42.6	45.9	1,379.3	1,423.3
St. Dev.	11.7	11.1	324.7	243.5
Min.	21.5	28.3	828.7	866.7
25%	35.0	38.5	1,185.4	1,291.0
Median	42.3	44.8	1,348.8	1,430.8
75%	50.5	52.3	1,626.6	1,536.0
Max.	65.1	71.7	2,070.0	1,922.3

**Table 3. Comparison of Statistics for Rainfall of All Events between Observed and Generated Data**

Classification	Observed data	Generated data
Number of data	227,760	227,760
Sum of precipitation	35,862.8	36,970.4
Average	0.157	0.162
St. Dev.	1.241	1.252
Min.	0.000	0.000
25%	0.000	0.000
Median	0.000	0.000
75%	0.000	0.000
Max.	65.100	71.688

**Table 4. Comparison of Statistics for Rainfall of Wet Events between Observed and Generated Data**

Classification	Observed data	Generated data
Number of data at wet time	15,420	15,347
Sum of precipitation	35,862.8	36,970.4
Average	2.326	2.409
St. Dev.	4.195	4.215
Min.	0.100	0.100
25%	0.300	0.340
Median	0.900	0.932
75%	2.500	2.728
Max.	65.100	71.688

Table 3에서 관측자료와 모의자료의 통계적 특성은 거의 비슷한 값의 평균과 표준편차를 갖는 것으로 나타났다. 전기간의 자료를 대상으로 기본적인 통계량을 산정한 결과에서 각 분위값들이 영(zero)를 갖는 것으로 산정되었다. 이는 강수가 발생한 시간보다 발생하지 않는 시간이 훨씬 많이 존재하기 때문이다.

따라서 Table 4에서 보는바와 같이, 습윤(강수가 발생한)시간자료만을 대상으로 통계량을 산정하여 비교하였다. Table 4에서 자료의 수는 습윤시간 중에서 강수량이 0.1mm 이상인 자료의 개수로 관측자료에서는 전

시간 대비 6.770%에 해당하는 기간 동안에 강수가 발생한 것으로 나타났으며, 모의자료에서는 전시간 대비 6.738%의 값을 갖는 것으로 나타났다.

Table 1부터 Table 4까지의 비교를 통해 관측자료와 모의자료를 비교분석한 결과에서 핵밀도함수를 이용한 비동질성 Markov 모형은 비교적 모의가 잘 이루어진 것으로 나타났다.

다음의 Table 5는 T-test와 F-test를 통해 연최대시간강수량과 연강수량이 관측자료와 모의자료가 같은 모집단에서 추출된 결과인지에 대하여 유의수준 5%로 검



Table 5. The Results of T-test and F-test between Observed and Generated Data

T-test	Classification	Degree of freedom		t value	Pr >  t
	Time rainfall	50		-1.07	0.2915
	Annual rainfall	50		-0.55	0.5833
F-test	Classification	Degree of freedom	Degree of freedom	F value	Pr >  F
	Time rainfall	25	25	1.10	0.8209
	Annual rainfall	25	25	1.78	0.1571

정을 수행하였다.

Table 5에서 나타낸 바와 같이, 시간강수량의 F-test의 결과에서 자유도 25, 199의 F-분포에서 F값인 1.10보다 클 확률이 0.8209로 유의수준 0.05보다 크므로 “귀무가설( $H_0$ ): 분산이 동일하다”는 가설을 기각할 수 없게 된다. 따라서 관측자료와 모의자료의 분산이 동일하며, t-test 결과 중의 시간강수량에서도 t값의 절대값 0.49보다 t-분포에서 클 확률이 0.2915이므로 두 표본집단의 평균이 같은 모집단에서 추출된 결과임을 의미하게 된다.

그러므로 비동질성 Markov 모형에 의한 시간강수량의 모의결과는 관측자료와 통계적 일관성을 유지하면서 짧은 관측기간을 가진 자료를 확장하는데 이용할 수 있는 것으로 판단된다. 즉, 연최대치시간강수량이나 연강수량과 같이 관측기간이 부족한 지점을 대상으로 자료를 확장할 수 있다.

그러나 비동질성 Markov 모형을 이용한 시간강수량의 모의는 Markov 모형의 특성인 조건부확률과정에 난수를 대응시키는 무작위과정에 의해 다음 상태가 결정되므로 관측치에서 시간강수사상이 연속적으로 발생하였을 때에 이에 대한 반영이 모의자료에서 약간 부족한 것으로 나타났다. 이 부분의 보완을 위한 추가적인 연구가 필요한 것으로 사료된다.

#### 4. 천이확률을 이용한 강수의 시간분포 유도

추풍령 지점의 시간강수량 모의는 변동핵밀도함수를 이용한 비동질성 Markov 모형을 적용하였으며, 시간강수량의 모의를 위한 각 시간별 천이확률은 관측자료 분석을 통해 산정하였다. Fig. 5에서 나타난 강수사상에서 무강수사상으로 천이되는 확률을 1에서 빼주게 되면 Fig. 8과 같이 강수사상에서 강수사상으로의 천이확률을 계산할 수 있게 된다.

본 연구에서는 강수사상에서 강수사상으로의 천이확률을 계산하여 각 연도별로 지속시간 24시간에서 가장 큰 천이확률을 갖는 구간을 추출하였다. 모의에 입력된 시간강수량 자료는 총 26년이므로 26개의 연최대 천이확률계열을 구성하였다.

천이확률로 표현되는 연최대계열을 이용하여 강수의 무차원누가시간분포곡선을 유도하였다. 유도방법은 각각의 24시간 천이확률 중에서 천이확률이 가장 큰 구간과 지속시간을 3등분하여 가장 큰 값을 갖는 구간이 일치하는 사상별로 구분하였다. 따라서 Fig. 9는 초기(advanced type)·중기(centered type)·말기(delayed type)에 가장 큰 천이확률을 갖는 경우를 도시한 결과이다.

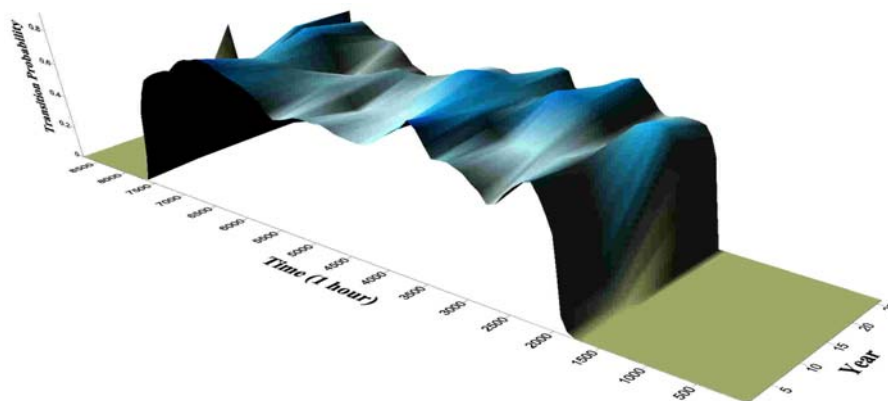


Fig. 8. Transition Probabilities form Wet Condition to Wet Condition

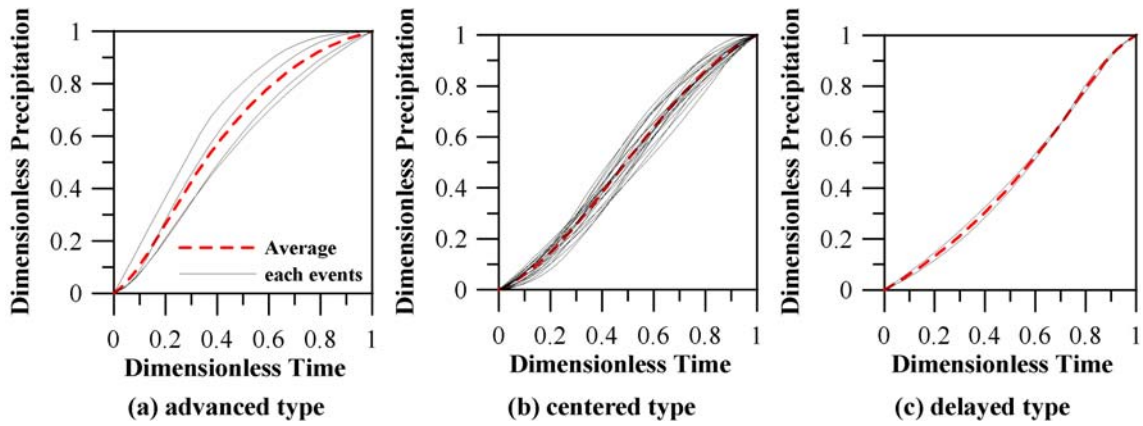


Fig. 9. The Derivation of Rainfall Mass Curve using Transition Probability

Fig. 9에서 강수초기(advanced type)에 가장 큰 천이 확률이 발생한 경우는 4개, 중기(centered type)는 20개이며 말기(delayed type)에는 2개 사상이 포함되는 것으로 나타났다. 또한, 각각의 전·중·후의 천이확률을 평균한 값도 함께 도시하였다.

천이확률을 이용하여 계산한 강수의 시간분포곡선을 한국확률강우량도(건설교통부, 2000)에서 제시하고 있는 추풍령 지점의 Huff 분포와 비교하여 Fig. 10에 도시하였다.

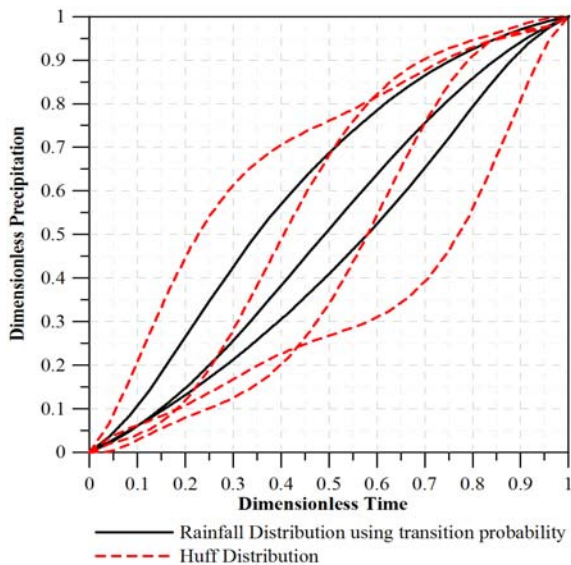


Fig. 10. Comparison to Huff-distributions

본 연구에서 유도한 시간분포 곡선과 Huff의 강우시간분포 곡선을 Fig. 10에서 비교하였다. Huff 분포에 비해 본 연구에서 유도한 시간분포곡선이 보다 완만하게 증가하며, 초기분포(advanced type)는 Huff 분포 중에서 1분위와 2분위 분포에 근사하고, 중기분포(centered type) 및 말기분포(delayed type)는 Huff 분포 중에서

2, 3분위 분포와 근사한 것으로 도시되었다.

강수의 시간분포를 유도하는데 있어, 천이확률을 이용한 시간분포의 이점은 다음과 같다. 즉, 실측강수의 시간분포는 변동이 매우 크다. 이는 강수가 발생하고 발생하지 않는 사상이 연속해서 측정되기 때문이다. 그러나 천이확률을 이용하는 경우에는 핵밀도 함수를 이용해 강수의 발생 빈도와 크기를 고려하여 천이확률을 산정한 자료를 이용하므로, 확률과정을 통한 강수의 시간분포를 묘사할 수 있는 장점이 있다.

## 5. 결론

본 연구에서는 핵밀도함수를 이용한 비동질성 Markov 모형을 통해 추풍령 지점에서 관측된 1981년부터 2006년까지의 26년 자료를 활용하여 시간강수량을 모의하였다. 모의시간강수량자료와 관측시간강수량자료의 통계치 비교와 T-test와 F-test를 통해 동일한 모집단에서 추출된 자료인지에 대한 검증을 실시하였으며, 천이확률을 이용한 강수의 시간분포에 적용을 시도하였다.

- 1) 핵밀도 함수를 이용한 비동질성 Markov 모형을 통해 산정된 천이확률은 추풍령 지점의 강수특성을 잘 반영해주는 것으로 나타났다. 즉, 비동질성 Markov 모형은 각각의 단위시간마다의 천이확률을 산정하여 습윤시간과 건조시간을 모의하므로 실측사상에 유사한 모의자료를 발생시킬 수 있다.
- 2) 비동질성 Markov 모형에 의한 시간강수량의 모의 결과를 관측강수량 자료의 통계적 특성치와 비교하였다. 비교·분석 결과에서 모의된 자료는 실측자료와 매우 유사한 강수자료를 강수관측자료가 부족한 지점의 자료를 연장할 수 있는 좋은 대안이 될 수 있다.

3) 산정된 천이확률을 이용하여 유도된 강수분포는 실측강수사상을 반영한 변동핵밀도함수를 통해 산정되므로, 유도된 강수의 무차원시간분포를 설계강수량에 적용할 수 있다.

핵밀도함수를 이용한 비동질성 Markov 모형은 관측 자료로부터 천이확률을 구성하고 핵함수를 이용하여 강수량의 크기를 결정하므로 모의시간을 증가시키면 강수량의 최대값이 증가될 수 있는 강수사상의 불확실성도 고려할 수 있다. 그러나 시간강수량의 습윤상태와 건조상태가 발생한 난수에 의해서 결정되므로 시간강수량의 지속성을 반영하는 데는 다소 부족한 것으로 나타나 이에 대한 추후 연구가 필요한 것으로 나타났다.

### 감사의 글

본 연구의 일부는 소방방재청 자연재해저감기술개발사업(과제명 : 사면붕괴 예측 및 대응기술 개발) 연구비 지원으로 수행되었으며 이에 감사드립니다.

### 참고 문헌

건설교통부 (2000). **한국확률강우량도**.

강경석 (2000). **다지점 일 강우모형에 의한 일 유출량의 모의발생**. 박사학위논문, 인하대학교.

강태호, 이재준, 이종태 (1999). "도시하천에서의 강우, 유출 및 수질 예측(I)-시간강우계열 모형 개발 및 적용." **대한토목학회 논문집**, 대한토목학회, 제19권, 제II-1호, pp. 23-34.

이원환, 이재준 (1985). "시간적 확률구조를 고려한 일 강수량의 모의발생에 관한 연구." **연세대학교 산업기술연구소 논문집**, 제17집, 제1권, pp. 37-46.

이재준 (1987). 강수계열의 모의발생모델 개발. 박사학위논문, 연세대학교.

이재준, 이정식 (2002). "시간강수계열의 강수발생과정에 대한 추계학적모형." **한국수자원학회논문집**, 한국수자원학회, 제35권, 제1호, pp. 109-124.

이창훈, 김승 (1995). "회귀분석에 의한 한국의 연·월 평균 강수량의 추정." **대한토목학회 논문집**, 대한토목학회, 제15권, 제5호, pp.1255-1266.

차영일 (2004). **다변량 핵밀도함수를 이용한 일강수량 모의발생에 관한 연구**. 박사학위논문, 서울시립대학교.

차영일, 김범순, 문영일 (2006). "극치값 추정에 적합한 비매개변수적 핵함수 개발." **한국수자원학회 논문집**, 한국수자원학회, 제39권, 제6호, pp. 187-194.

허준행, 김정덕, 조원철 (1997). "다변량 추계학적 모형을 이용한 낙동강 유역의 가뭄해석에 관한 연구." **한국수자원학회 논문집**, 한국수자원학회, 제30권, 2호, pp. 155-163.

한국수자원공사 (1999). **확률강수량 산정 및 이용방법 연구**.

Adamowski, K. (1985). "Nonparametric kernel estimation of flood frequency." *Water Resources Research*, Vol. 21, No. 11, pp. 1585-1590.

Adamowski, K., and Labatiuk, C. (1987). "Estimation of flood frequencies by a nonparametric density procedure." *Hydrologic Frequency Modeling*, pp.97-106.

Bardsley, W.E. (1988). "Toward a general procedure for analysis of extreme random events in the earth sciences." *Math. Geol.*, Vol. 20, No. 5, pp. 513-528.

Bardsley, W.E. (1989). "Using historical data in nonparametric flood estimation." *Journal of Hydrology*, Vol. 108, pp. 249-255.

Bowman, A. W. (1984). "An alternative method of cross-validation for the smoothing of density estimates", *Biometrika*, Vol. 71, pp.353-360.

Breiman, L., W. Meisel, and Purcell, E. (1977). "Variable kernel estimates of multivariate densities." *Technometrics*, Vol. 19, No. 2, pp. 135-144.

Bruhn, J. A., Fry, W. E., and Fick, G. W. (1980). "Simulation of daily weather data using theoretical probability distributions." *Jour. Appl. Mete.*, Vol. 19, No. 9, pp. 1029-1036.

Chang, J.J., Kavvas, M.L., and Delleur, J.W. (1984). "Daily precipitation modeling by discrete autoregressive moving average processes" *Water Resources Research*, Vol. 20, No.5, pp. 556-580.

Gabriel, K. R., and Neumann, J. (1957). "Distribution of weather cycles by length." *Quart Jour. Royal Meteor. Soc.*, Vol. 83, pp. 375-380.

Lall, U., Moon, Y. I., and Bosworth, K. (1993). "Kernel flood frequency estimators: bandwidth selection and kernel choice." *Water Resources Research*, Vol. 29, No. 4, pp. 1003-1015.

Lettenmaier, D.P., and Burges, S. J., (1977). "An operational approach to preserving skew in hydrologic models of long-term persistence."

- Water Resources Research*, Vol. 13, No. 2, pp. 281-290.
- Moon, Y. I., and Lall, U. (1994). "Kernel quantile function estimator for flood frequency analysis." *Water Resources Research*, Vol. 30, No. 11, pp. 3095-3103.
- Moon Y.I., and Lall, U. (1996). "Atmospheric flow indices and interannual Great Salt Lake variability." *Journal of Hydrologic Engineering*, Vol. 1, No. 2, pp 55-61.
- Nord, J. (1975). "Some applications of markov chains", *Proceedings Fourth Conference on Probability and Statistics in Atmospheric Science*, Tallahas, pp.125-130.
- Rajagopalan, B., Lall, U. and Tarboton, D. (1996). "Nonhomogeneous markov model for daily precipitation." *Journal of Hydrologic Engineering*, Vol. 1, No. 1, pp.33-40.
- Richardson, C. W. (1981). "Stochastic simulation of daily precipitation, temperature, and solar radiation." *Water Resource Research*, Vol. 17, No. 1, pp. 182-190.
- Rudemo, M. (1982). "Empirical choice of histograms and kernel density estimators", *Scand. J. Statist.*, Vol. 9, pp.65-78.
- Scott, D. W. (1992). *Multivariate density estimation: theory, practice and visualization*. New York, Wiley.
- Todorovic, P. and Woolhiser, D. A. (1975). "A stochastic model of n-day precipitation." *Journal of Applied Meteorology*, Vol. 14, pp. 17-24.
- Yakowitz, S. (1985). "Markov flow models and the flood warning problem." *Water Resources Research*, Vol. 21, No. 1, pp. 81-88.

(논문번호:07-143/접수:2007.12.11/심사완료:2008.01.24)