

비동질성 Markov모형에 의한 일 강수량 모의발생

○문영일¹⁾, 서병하²⁾, 정상만³⁾, 김형수⁴⁾, 차영일⁵⁾, 박대형⁶⁾

1. 서론

지금까지 마코프모형은 일강우 모의발생에 대한 모형으로 많이 이용되고있다(서병하 등, 1999; 김주환 등, 1996; 이재준, 1987; Todorovic와 Woolhiser, 1975; Smith와 Schreiber, 1974). 즉, 무강우기간, 강우기간, 강수량을 모의발생하거나 또는 강수량을 특정 확률분포로 가정하여 강수량을 모의발생시키는 방법을 사용하였다. 그러나, 이러한 대부분 모형들은 동질성 마코프 연쇄모형으로 천이확률(강우일에서 무강우일로의 천이 또는 강우일에서 강우일로의 천이와 무강우일에서 강우일로의 천이 또는 무강우일에서 무강우일로의 천이)이 천이발생일(1~365일)에 상관없이 과거의 관측된 자료로부터 크기에 따른 상태벡터를 구성하여 이에 따른 상태별 천이 빈도수에 따라 구성하였다. 그러나, 천이확률은 계절에 따라 변하기 때문에 연중에 다르게 나타날 수 있다는 것에 문제가 있다. 즉, 강우발생의 과정은 비정상적(nonstationary)이다. 이러한 문제점을 보완한 방법을 두 가지로 크게 나누면 다음과 같다. 첫 번째가 자료를 기간(또는 계절)별로 나누어, 천이확률을 계절에 따라 각각 추정하는 방법이다(서병하 등, 1999; 김주환 등, 1996)). 이 때의 문제점은 발생과정이 계절 안에서 정상적(stationary)으로 가정하는데 있다. 두 번째 방법은 천이확률이 연중에 변한다는 것을 인정함으로써 본질적으로 비동질성의 마코프 과정을 고려하는 것으로, Fourier series를 이용하여 천이확률의 계절성을 고려하는 것이다(Feyerherm과 Bark, 1965; Woolhiser 등, 1973; Woolhiser와 Pegram, 1979). 그러나 이 때의 문제점은 천이확률이 년 중 대칭으로 변화한다고 가정하는데 있고 충분한 자료가 있어야 효과적으로 모형구성이 가능하다는 것에 있다. 본 연구에서는 비동질성 마코프 연쇄모형의 구조를 근거한 비동질성의 상태, 즉, 천이발생일에 따른 계절의 변동성을 나타낼 수 있는 마코프 연쇄모형을 위해 핵밀도함수를 사용한 Rajagopalan 등(1996)의 방법을 좀 더 확장된 비매개변수적 핵밀도함수를 이용하여 비동질성 천이확률을 구성하였다. 핵밀도함수(kernel density function)의 개략적인 개념(Silverman, 1986; Lall 등, 1993; Moon 등, 1994; Moon과 Lall, 1994)은 모든 관측치 자료를 이용하여 우리가 관심있는 날의 천이확률은, 관심 있는 날에 가깝게 발생된 천이사상에 상대적으로 큰 가중치를 주고, 관심 있는 날에서 점점 멀리 떨어질수록 점점 작은 가중치를 주어 기록치에 대한 모든 날에 대하여 천이확률을 구한다. 다음은, 난수(U[0,1])를 발생한 후, 천이확률을 이용하여 다음날에 대한 강우상태를 결정할 수 있다. 습윤일에 대한 강우량은 핵함수에 의한 확률밀도함수로부터 관측된 자료를 이용하여 구하였다.

2. 비동질성 천이확률 Markov 모형 모의과정

본 연구에서는 one-step의 비동질성 마코프 모형에 대하여 구성하였다. one-step의 2x2 천이확률 행렬을 비매개변수적 핵함수 개념을 적용하여 일별로 부드럽게 변화하는 것으로 가정하였다. 관측 강우자료로부터 핵확률밀도함수를 이용한 천이확률의 구성은 다음과 같다. 먼저 강

-
- 1) 서울시립대 토목공학과 조교수
 - 2) 인하대 토목공학과 교수
 - 3) 공주대 토목환경공학과 부교수
 - 4) 선문대 건설공학부 전임강사
 - 5) 서울시립대 토목공학과 석사과정
 - 6) 서울시립대 토목공학과 석사과정

우발생과정을 그림 1과 같이 나타내었다.

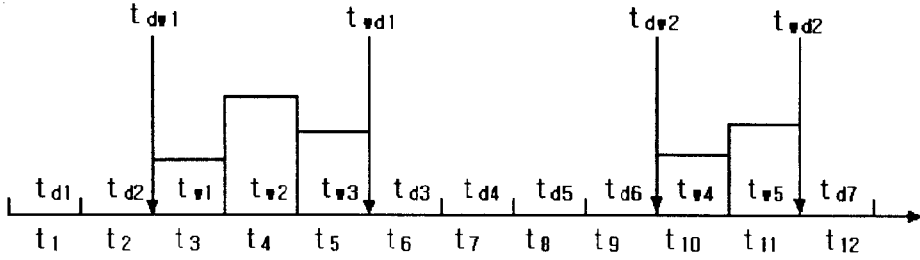


그림 1. 임의의 해에 대한 강수일

일 지수(t)는 1년 중의 1일부터 365일(윤년은 366일)을 나타낸다. 여기서, 일 강우자료로부터 네 종류의 지수로 나눌 수 있다. ① 일 지수 $t_{w1}, t_{w2}, \dots, t_{wn}$ 은 n_w 개의 강우일을 나타내고, ② 일 지수 $t_{d1}, t_{d2}, \dots, t_{dn}$ 은 n_d 개의 무강우일을 나타내며, ③ 일 지수 $t_{wd1}, t_{wd2}, \dots, t_{wd_{nwd}}$ 은 강우일에서 무강우일로 바뀌는 n_{wd} 개의 천이 발생일을 나타낸다 (즉, t_{wd1}, t_{wd2} 등은 강우일을, $t_{wd1}+1, t_{wd2}+1$ 등 무강우일을 나타낸다). ④ 일 지수 $t_{dw1}, t_{dw2}, \dots, t_{dw_{ndw}}$ 은 무강우일에서 강우일로 바뀌는 n_{dw} 개의 천이 발생일을 나타낸다 (즉, t_{dw1}, t_{dw2} 등은 무강우일을, $t_{dw1}+1, t_{dw2}+1$ 등은 강우일을 나타낸다). 여기서, 역력 t일의 천이확률 $P_{wd}(t)$ 와 $P_{dw}(t)$ 은 비매개변수적 핵밀도함수 추정치를 이용하여 자료로부터 구한다. 동질성 마코프 모형에서는 천이확률을 단순히 과거의 자료에서 천이의 수와 과거의 자료에서 강우 또는 무강우의 수에 대한 비율로써 적당히 추정되었으나, 여기서는 핵함수를 이용하여 관심 있는 사상(강우일, 무강우일, 또는 천이일)이 발생되어진 날짜에 핵함수의 중앙이 위치하도록 하여 관심의 대상이 되는 날의 부근에 일어난 사상(즉, 강우나 무강우일 또는 천이 상태)은 가중치를 더 주었고, 멀리 떨어져 일어난 사상은 가중치를 덜 주었다. 이에 대한 결과로 천이확률에 대해 핵함수 추정식은 아래와 같다. 나머지 두 개의 천이확률($P_{uw}(t)$ 와 $P_{dd}(t)$)은 $P_{wd}(t) + P_{uw}(t) = 1$ 과 $P_{dw}(t) + P_{dd}(t) = 1$ 로부터 직접 구할 수 있다.

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

$$\hat{P}_{dw}(t) = \frac{\sum_{i=1}^{n_{dw}} K\left(\frac{t-t_{dwi}}{h_{dw}}\right)}{\sum_{i=1}^{n_d} K\left(\frac{t-t_{di}}{h_{dw}}\right)} \quad (2)$$

$$P_{uw}(t) = 1 - P_{wd}(t) \quad (3)$$

$$P_{dd}(t) = 1 - P_{dw}(t) \quad (4)$$

이 때, n_{wd} 는 과거 자료에서 강우일로부터 무강우일로 바뀌는 천이의 수, n_{dw} 는 과거 자료로부터 무강우일로부터 강우일로 바뀌는 천이의 수, n_d 는 과거의 자료에서 무강우일의 수, n_w 는 과

거의 자료에서 강우일의 수이다.

상태 모의발생은 난수를 발생하여 결정하고, 만약에 현재가 강우일(W)이면 현재의 천이 확률로부터 다음 날의 강우상태를 아래 그림 2와 같이 모의 발생시킨다.

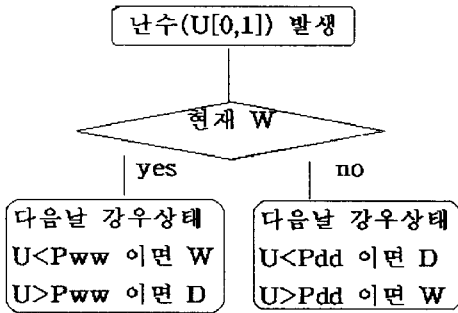


그림 2. 난수를 이용한 천이상태 결정

3. 습윤일 강우량 확률밀도함수를 이용한 강우량 모의발생

아래 그림3과 같이 관심의 대상이 되는 날(t)에 위치한 커널함수의 h_p 에 포함 되는 모든 과거자료로부터의 강우일의 강우량(즉, y_1, y_2, \dots, y_{np})을 고려하였다.

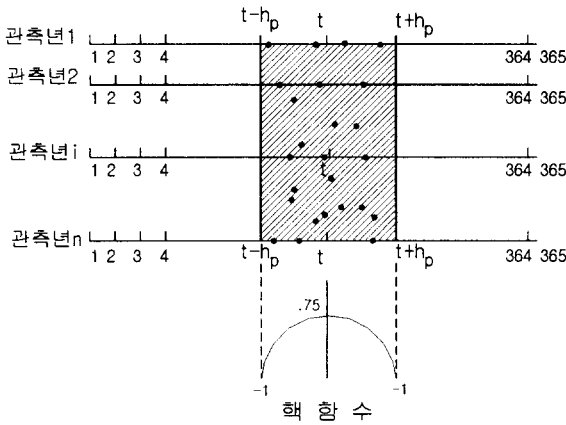


그림 3. 강우량 모의발생 과정

$$K(t) = 0.75(1-t^2), |t| \leq 1$$

관측된 강우량 np 개의 각 강우일에 해당하는 가중치 $w_{t1}, w_{t2}, \dots, w_{tnp}$ 에 의해 주어진 확률행렬에서 1과 np 사이의 난수를 발생하여 강우량 관측치 y_i 를 선택한다.

다음, 습윤일 t 에서 강우량의 식(5) 조건부 확률밀도함수를 이용하여 난수를 대응시킨다.

$$\hat{f}(y | t) = \frac{1}{y h_{LY}} \sum_{i=1}^{np} K_c \left(\frac{\log(y) - \log(y_i)}{h_{LY}} \right) \cdot K \left(\frac{t - t_i}{h_p} \right) / \sum_{i=1}^{np} K \left(\frac{t - t_i}{h_p} \right) \quad (5)$$

여기서,

$$\hat{f}(y) = \frac{1}{np} \sum_{i=1}^{np} \frac{1}{h_{LY} y} K_c \left(\frac{\log(y) - \log(y_i)}{h_{LY}} \right) \quad (6)$$

이 때 h_{LY} 는 대수화된 자료의 bandwidth. 여기서 난수 $U[0,1]$ 는 커널 함수 $K_c(\cdot)$ 에 대응하는 값이므로 모의발생된 강우량은 $y^* = \exp[\log(y) + Uh_{LY}]$ 이다.

위 과정을 50년, 80년, 100년, 200년 자료가 발생될 때까지 각 강우 관측소에 대하여 일별로 반복하였다. 고려된 강우관측소와 관측기간은 아래의 표1과 같다.

표 1. 강우관측소와 관측기간

강우관측소 위치	Data File명	관 측 기 간	기록 년수	
옥산 수위표 지점	청주	chj	1967-96	30년
	조치원	jcw	1993-95	3년
	병천	byc	1967-74, 1978-79, 1983-95	23년
	증평	jup	1973-74, 1978-79, 1983-95	17년
	진천	jch	1967-74, 1978-79, 1983-95	23년
용담댐 지점	무주	muj	1967-74, 1978-79, 1983-95	23년
	진안	jia	1967-74, 1978-79, 1983-95	23년
	계북	gyb	1993-95	3년
	장수	jas	1967-79, 1983-92, 1994-95	22년
횡성댐 지점	횡성	hos	1967-74, 1978-79, 1983-95	23년
	청일	chi	1967-74, 1978-79, 1983-95	23년

위와 같은 모의발생한 강우량을 과거 관측자료들과 통계적인 사상을 비교하기 위하여 여러 통계값과 비교하였다. 관측자료와 모의자료의 사상 비교는 (1) 강우사상의 강우량에 대한 계절별 (2) 강우사상의 습윤기간에 대한 계절별 (3) 습윤일의 강우량에 대한 계절별 (4) 무강우일의 건조기간에 대한 계절별 (5) 습윤일의 강우량에 대한 년에 대하여 통계값(①평균값 ②최대값 ③최소값 ④1사분위 값(25% quantile) ⑤중간값 ⑥3사분위 값(75% quantile) ⑦표준편차)으로 되어 있다. 이 중에서 청주관측소에 대한 결과를 표2 ~ 표4에 나타내었다.

표 2. 강우사상의 강우량에 대한 계절별 과거관측자료와 모의발생된 강우량의 통계자료(청주)

구분	25% quantile	평균 강우량 (mm)	중간값 (mm)	75% quantile	최소 강우량 (mm)	최대 강우량 (mm)	표준편차 (mm)
계절 1							
관측된강우량	0.700	7.808	2.700	9.800	0.100	95.500	12.188
모의강우량	0.815	8.086	3.170	9.621	0.051	113.542	12.929
계절 2							
관측된강우량	2.500	20.466	9.300	26.000	0.100	251.700	29.867
모의강우량	2.699	22.586	10.544	30.430	0.057	236.402	31.589
계절 3							
관측된강우량	4.200	40.062	16.300	54.500	0.100	495.000	58.159
모의강우량	5.156	40.377	19.646	55.257	0.073	399.005	54.092
계절 4							
관측된강우량	0.900	8.382	3.500	10.200	0.100	105.600	13.067
모의강우량	1.122	9.791	4.545	11.948	0.048	148.165	15.667

표 3. 강우사상의 습윤기간에 대한 계절별 과거관측자료와 모의발생된 강우량의 통계자료(청주)

구분	25% quantile	평균 습윤기간	중간값	75% quantile	최소 습윤기간	최대 습윤기간	표준편차
계절 1							
관측된강우량	1.000	1.772	1.000	2.000	1.000	7.000	1.083
모의강우량	1.000	1.688	1.000	2.000	1.000	7.571	1.097
계절 2							
관측된강우량	1.000	1.854	2.000	2.000	1.000	11.000	1.188
모의강우량	1.000	1.933	1.000	2.143	1.000	12.857	1.416
계절 3							
관측된강우량	1.000	2.277	2.000	3.000	1.000	15.000	1.729
모의강우량	1.000	2.180	1.857	3.000	1.000	14.000	1.706
계절 4							
관측된강우량	1.000	1.755	1.000	2.000	1.000	8.000	1.157
모의강우량	1.000	1.834	1.000	2.000	1.000	8.714	1.270

표 4. 습윤일의 강우량에 대한 계절별 과거관측자료와 모의발생된 강우량의 통계자료(청주)

구분	25% quantile	평균 강우량 (mm)	중간값 (mm)	75% quantile	최소 강우량 (mm)	최대 강우량 (mm)	표준편차 (mm)
계절 1							
관측된강우량	0.400	4.407	1.500	5.000	0.100	51.800	7.068
모의강우량	0.421	4.783	1.568	5.261	0.042	84.210	8.705
계절 2							
관측된강우량	1.100	11.038	4.100	13.400	0.100	156.400	17.148
모의강우량	1.091	11.691	4.102	14.372	0.047	180.896	18.782
계절 3							
관측된강우량	1.600	17.596	6.700	22.700	0.100	293.000	26.136
모의강우량	1.530	18.537	7.037	23.003	0.053	348.596	30.062
계절 4							
관측된강우량	0.500	4.776	1.800	5.900	0.100	77.400	7.849
모의강우량	0.484	5.333	1.938	6.210	0.040	118.846	9.812

4. 결과 및 결론

마코프 연쇄모형의 비동질성의 상태, 즉, 천이발생일에 따른 계절의 변동성을 나타낼 수 있는 마코프 연쇄모형을 위해 비매개변수적 핵밀도함수를 이용하여 비동질성 천이확률을 구성하였다. 추정된 비동질성 천이확률을 이용하여 비정상성(nonstationary)인 일강우량자료를 11개 강우관측소에 대하여 모의발생을 한 후 관측된 자료와 여러 통계 값을 비교하였다. 연중 천이확률이 변할 수 있다는 것을 인정한 모형의 유연성에 따라 모의된 최대강우량의 변화의 폭이 다소 크나, 나머지 통계치는 근접하게 나타났다. 또한 계절별 및 연별 모의실험의 개수를 증가할수록 모의된 통계치의 평균이 과거의 자료들의 값에 접근하였다. 따라서, 동질성 마코프 모형을 계절적인 비동질성으로 확장하여 매끄럽게 변화하는 천이확률 고려하여 모의발생된 강우량은 계절이나 연도별로 강우기간 길이, 무강우기간 길이, 그리고 강우일 강우량의 발생의 빈도구조를 유지하는 것으로 나타났다.

감사의 글

본 연구는 1998년도 한국수자원공사 학술연구비에 의하여 연구되었으며, 한국수자원공사의 지원에 대해 감사드립니다.

5. 참고문헌

- 김주환, 박찬영, 강관원. 1996. 강수계열의 상태분류에 의한 Markov 연쇄 모의발생 모형. 한국수자원학회지 제29권 제6호: 155-166.
- 이재준. 1987. 강수계열의 모의발생 모델 개발. 연세대학교 박사학위논문.
- 서병하, 강경석, 김병식. 1999. 일 강우량의 모의 발생을 통한 갈수유량 계열의 산정 및 빈도분석. 한국수자원학회지(수록예정)
- Feyerherm, A. M., and Bark, L. D. 1965. Statistical methods for persistent precipitation patterns. *Journal of Appl. Meteorology* 4: 320-328.
- Lall, U., Moon, Young-II, and Bosworth, K. 1993. Kernel flood frequency estimators: bandwidth selection and kernel choice. *Water Resources Research* 29(4):1003-1015.
- Moon, Young-II, Lall, U., and Bosworth, K. 1993. A comparison of tail probability estimators. *Journal of Hydrology* 151: 343-363.
- Moon, Young-II and Lall, U. 1994. Kernel Quantile Function Estimator for Flood Frequency Analysis. *Water Resources Research*.
- Rajagopalan, B., Lall, U. and Tarboton, D. 1996. Nonhomogeneous markov model for daily precipitation. *Journal of Hydrologic Engineering* 1(1): 33-40.
- Silverman, B. W. 1986. *Density estimation for statistics and data analysis*. Chapman and Hall, New York, N.Y.
- Smith, J. A. and Schreiber, H. A. 1974. Point processes of seasonal thunderstorm rainfall. 1. Distribution of rainfall events. *Water Resources Research* 10(3): 418-423.
- Todorovic, P. and Woolhiser, D. A. 1975. Stochastic model of n day precipitation. *Journal of Appl. Meteorology* 14(1): 17-24.
- Woolhiser, D. A. and Pegram, G. G. S. 1979. Maximum likelihood estimation of Fourier coefficients to describe seasonal variations of parameters in stochastic daily precipitation models. *Journal of Appl. Meteorology* 18: 34-42.