Neyman-Scott 구형 펄스모형의 직접적인 매개변수 추정연구

Study of Direct Parameter Estimation for Neyman-Scott Rectangular Pulse Model

정 창 삼*

Jeong, Chang-Sam

Abstract

NSRPM (Neyman-Scott Rectangular Pulse Model) is one of the common model for generating future precipitation time series in stochastic hydrology. There are 5 parameters to compose the NSRPM model for generating precipitation time series. Generally parameter estimation using moment has some problems related with increased objective functions and shows different results in accordance with random variable generating models. In this study, direct parameter estimation method was proposed to cover with disadvantages of parameter estimation using moment. To apply the direct parameter estimation, generating stochastic data variance in accordance with numbers of precipitation events of NSRPM was done. Both kinds of methods were applied at the Cheongju gauge station data. Precipitation time series were generated using 4 different random variable generator, and compared with observed time series to check the accuracies. As a results, direct method showed more stable and better results.

Keywords : NSRPM, generating future precipitation time series, direct parameter estimation

요 지

NSRPM (Neyman-Scott Rectangular Pulse Model)은 수문학분야에서 널리 쓰이고 있는 강수생성모형이다. NSRPM을 구축하기 위해서는 총 5개의 매개변수를 추정하여야 한다. 일반적으로 사용되는 모멘트를 이용하여 매개변수를 추정할 경우, 사용되는 목적함수의 증가에 따라 추정되는 매개변수의 결과가 평탄해지고 목적함수를 추가하거나 조정하기 위해서는 복잡한 수식을 다시 계산해야 하며 추정된 매개변수가 무작위변수 생성 모형에 따라 상이한 결과를 나타내는 단점이 있다. 본 연구에서는 직접적인 매개변수 추정방법을 제시하여 모멘트를 이용한 매개변수 추정의 단점을 극복하고자 하였다. 직접적인 추정방법을 적용하기 위하여 NSRPM의 강수 생성 개수에 따른 통계치 변화를 모의하여 직접적인 추정을 위한 모형을 구축하였다. 기상청 청주 지상관측소의 관측 강수 자료를 사용하여 모멘트를 이용하여 추정된 매개변수와 직접적인 방법을 이용하여 매개변수를 추정하였다. 총 4 개의 무작위변수 알고리즘을 적용하여 강수를 생성하였고, 관측 강수 시계열을 이용하여 정확도를 비교하였다. 비교 결과 직접적인 추정방법이 모멘트를 이용한 매개변수 추정방법보다 안정적이고 높은 정확도를 보이는 매개변수를 추정하는 것으로 나타났다.

핵심용어 : NSRPM, 강수생성모형, 직접적 매개변수 추정

* 교신저자, 인덕대학 토목환경설계공학과 전임강사

Corresponding Author, Assistant Prof. Dept. of Civil and Environmental Eng. Induk Univ. Seoul, Korea
(e-mail: csjeong@induk.ac.kr)

1. 서론

세계 곳곳에서 기후변화로 야기되는 이상기후에 의한 피해가 늘고 있다. 이상기후에 대비하기 위해서는 기후변화를 예측한 방재 대책과 수자원 운영 시나리오가 필요하다. 수자원 운영 시나리오를 작성하기 위해서는 장기 기후변화 모의를 수행하여야 한다. 장기 기후변화를 모의하기 위해서는 전지구 기후모형(Global Climate Model, GCM)을 이용한다. 전지구 모형을 사용하여 장기 기후변화를 예측 시 결과값은 일자료와 같은 큰 시간 스케일로 예측된다. 전지구 모형을 수자원 분야에 적용할 경우 적용목적에 따라 다운스케일할 필요가 있다. Neyman-Scott 구형펄스모형(NSRPM)은 점과정(point process)을 이용하여 강우를 생성하는 모형으로 널리 쓰이고 있다. NSRPM를 이용할 경우 강우 시계열 자료의 다운스케일을 용이하게 할 수 있다.

수공학분야에서 NSRPM은 Rodriguez-Iturbe *et al.* (1987)이 처음으로 적용하였으며, 모델의 각 분포형을 조정하여 보다 정확한 강우를 모의하는 연구가 진행되었다(Rodriguez-Iturbe *et al.*, 1988; Entekhabi *et al.*, 1989). 현재는 지속기간과 강우강도의 상관성을 고려한 NSRPM이 개발되었다(Kim and Kavvas, 2006; Evin and Favre, 2008). NSRPM의 모델의 조정외에도 모델의 매개변수를 추정하는 연구도 활발히 되고 있다. P.S.P Cowperwait *et al.* (1996)은 강수일과 무강수일의 비를 매개변수 추정을 위한 수식을 개발하였다. Calenda and Napolitano (1999)은 NSRPM으로 생성된 강우의 시간적인 변화를 이용하여 매개변수를 추정하는 연구를 하였다. Farve *et al.* (2004)은 비편의 매개변수 추정에 대한 연구를 하였다.

국내에서도 점과정을 이용한 많은 연구가 진행되어 왔다. 김재한 등(1998)은 NSRPM을 이용하여 강수 시계열을 생성하였고, 금종호 등(2001)은 NSRPM의 매개변수 추정에 대하여 연구를 하였다. 유철상 등(2002), 유철상과 김대하(2006), 김병식 등(2008)은 점과정의 응용에 대하여 연구를 하였다. Kim (2004)은 국내에 Gumbel분포를 이용하여 지속기간과 강우강도를 고려한 모형을 적용하였다. 김기욱과 유철상(2008)은 modified NSRPM을 국내에 적용하였다.

NSRPM을 구축하기 위해서는 총 5개의 매개변수를 추정하여야 한다. 일반적으로 사용되는 모멘트를 이용한 매개변수 추정방법은 수학적으로 유도된 수식을 이용하여 매개변수를 추정하기에 매개변수 추정이 용이하

고 매개변수 추정시간이 짧은 장점이 있다. 그러나 사용되는 목적함수의 증가에 따라 추정되는 매개변수의 결과가 평탄해 지고 목적함수를 추가하거나 조정하기 위해서는 복잡한 수식을 다시 계산해야 하며 추정된 매개변수가 무작위변수 생성 모형에 따라 상이한 결과를 나타내는 단점이 있다. 본 연구에서는 NSRPM의 매개변수 추정의 정확도와 사용성을 높이고자 직접적인 매개변수 추정방법을 적용하였다.

2. Neyman-Scott 구형펄스모형(NSRPM)

Neyman-Scott과정은 천문학에서 사용되었다(Neyman and Scott, 1958). Rodriguez-Iturbe (1987) 등이 기존에 수문학 분야에서 강수 생성을 위해 사용되었던 구형펄스모형(Rectangular Pulse Model)의 강수의 자기상관구조를 표현하지 못하는 단점을 극복하고자 Neyman-Scott 과정을 추가한 NSRPM을 개발하였다. 구형펄스모형은 포와송분포를 이용하여 임의의 강우를 생성하는 모형이다. 구형펄스 모형의 경우 각 강우사상이 독립적으로 나누어져 있어 실제 강우의 경우 군집화된 특성을 나타내는데 그 한계성이 있다. NSRPM은 이러한 구형펄스모형의 한계를 극복하고자 강우사상을 한 군집으로 고려하여 한 강우사상이 아닌 한 군집화된 강우를 생성한다.

2.1 NSRPM의 과정

NSRPM은 총 5개의 확률분포형으로 이루어져 있다. 우선 점과정(point process)을 이용하여 강우 군집의 시작점(origin)을 선정한다.이 때 λ 라는 매개변수가 사용되며 그 식은 Eq. (1)와 같다.

$$p = \lambda \exp(-\lambda t) \quad (1)$$

선정된 강우 군집의 시작점에 대하여 기하함수(geometric function)을 이용하여 각 강우 군집의 강우세포(cell)의 개수를 정한다. 이 때 μ_c 라는 매개변수가 사용되며 그 식은 Eq. (2)과 같다.

$$p = \frac{1}{\mu_c} \left(1 - \frac{1}{\mu_c}\right)^c \quad (2)$$

여기서 μ_c 는 강우 군집의 시작점들의 강우세포의 평균 개수를 나타낸다.

각 강우 군집의 시작점의 강우세포 개수가 결정되면 지수(Exponential)분포를 이용하여 각 강우세포의 위치를 정한다. 각 강우세포의 위치를 정할 때 기준이 되는

위치를 강우 군집의 시작점으로 보느냐 아니면 바로 앞 강우세포로 보느냐에 따라 Neyman-Scott 과정과 Bartlett-Lewis 과정으로 나뉜다. 본 연구에서는 Neyman-Scott 과정을 이용한 NSRPM모형에 대해서만 논하겠다. 강우세포의 위치를 정하는 지수(Exponential)분포는 β 를 매개변수로 갖고, 그 식은 Eq. (3)과 같다.

$$p = \exp(-\beta x) \quad (3)$$

강우세포들의 위치가 정해지면 각 강우세포들의 지속기간을 지수(Exponential)분포를 이용하여 정한다. 지속기간을 정하는 분포의 매개변수로는 η 가 사용되며, η 는 각 강우세포의 지속기간의 평균의 역수이며 강우세포의 위치를 정하는 지수분포의 식의 개형은 Eq. (4)와 같다.

$$p = \exp(-\eta x) \quad (4)$$

지속기간이 정해지면 지수 분포를 이용하여 각 강우세포들의 강우강도를 정한다. 이때는 매개변수로서 ξ 가 사용되며 각 강우세포의 강우강도의 평균의 역수이다. 강우 세포들의 강우강도를 정할 때 사용하는 식은 Eq. (5)와 같다.

$$p = \exp(-\xi x) \quad (5)$$

이렇게 정해진 강우 세포들의 강우강도들을 중첩하여 시간에 따른 강우를 생성한다(Calenda and Napolitano, 1999). NSRPM의 과정은 Fig. 1과 같다.

2.2 NSRPM의 목적함수 구성을 위한 second order property

NSRPM의 매개변수는 위의 2.1절에서 설명했듯이 λ , μ_c , β , η , ξ 총 5개가 사용된다. NSRPM의 매개변수를 추정하기 위해서는 NSRPM의 모멘트를 알아야 한다. NSRPM의 모멘트는 Rodriguez-Itrube *et al.* (1987)의 연구에 의하여 정리 되었으며 그 식은 아래와 같다. Eq. (6)은 h시간동안의 누적 강수량의 평균을 나타내고 Eq. (7)은 h시간 동안의 누적 강수량의 분산을 나타내고 Eq. (8)은 h시간 동안의 누적 강수량의 lag 1인 공분산을 나타낸다. 아래 식들은 NSRPM의 매개변수를 추정하기 위하여 사용되는 목적함수의 식들로서 2.1에서 설명된 NSRPM의 수식들을 이용하여 증

명된 수식들이다.

$$E(Y_i^{(h)}) = \frac{\lambda \mu_c h}{\eta \xi} \quad (6)$$

$$Var(Y_i^{(h)}) = \frac{\lambda}{\eta^3} (\eta h - 1 + e^{-\eta h}) \left(\frac{4\mu_c}{\xi^2} + \frac{2\mu_c(\mu_c - 1)\beta^2}{\xi^2(\beta^2 - \eta^2)} \right) - \frac{\lambda(\beta h - 1 + e^{-\beta h})2\mu_c(\mu_c - 1)}{\xi^2\beta(\beta^2 - \eta^2)} \quad (7)$$

$$Cov(Y_{i+1}^{(h)}, Y_i^{(h)}) = \quad (8)$$

$$\frac{\lambda}{\eta^3} (\eta h - 1 + e^{-\eta(k-1)h}) \left(\frac{2\mu_c}{\xi^2} + \frac{1\mu_c(\mu_c - 1)\beta^2}{\xi^2(\beta^2 - \eta^2)} \right) - \frac{\lambda(1 - e^{-\beta h})^2 e^{-\beta(k-1)h} 2\mu_c(\mu_c - 1)}{\xi^2\beta(\beta^2 - \eta^2)}$$

3. 매개변수 추정방법

본 연구에서는 NSRPM의 매개변수를 추정하기 위한 최적화 기법으로 유전자 알고리즘을 적용하였다. 모멘트를 이용하여 매개변수를 추정할 경우 통계치는 맞으나 현실에서는 일어날 수 없는 매개변수를 추정하는 경향이 있다. 특히 모멘트를 이용할 경우 무강수일과 유강수일을 고려하는 항이 목적함수로 추가되지 않을 시 추정된 매개변수가 강수일수를 정확히 모의하지 못하는 것으로 나타났다(Cowpertwait *et al.*, 1996). Cowperwait *et al.* (1996)은 무강수일과 유강수일을 나타내는 모멘트항을 만들었으나 복잡한 수식으로 인하여 널리 사용되고 있지 않다. 또한 각 분포형을 변형시키고자 할 때 매개변수로 구성된 복잡한 모멘트식을 적용하여 많은 사람들이 NSRPM를 조정하고 사용하기 어렵게 만드는 요인이 되고 있다. 또한 모멘트법을 이용하여 매개변수의 추정의 경우 seed number의 변화나 무작위변수(random number)를 생성하는 알고리즘에 따라서 생성되는 강수가 매우 상이한 결과를 나타내고 있다. 그러나 직접적인 방법을 이용하여 NSRPM의 매개변수를 추정 할 경우 모멘트를 이용한 매개변수 추정방법과 같이 복잡한 수식을 유도하지 않고 실제 생성되는 강우를 이용하여 매개변수를 추정하기에 모멘트를 이용하여 매개변수를 추정하는 방법보다 직관적으로 목적함수를 구성할 수 있다. 또한 무작위변수 생성 알고리즘을 처음부터 고정해 놓고 강우를 생성하기에 보다 무작위변수 생성 알고리즘의 맞는 매개변수를 추정할 수 있다.

3.1 유전자 알고리즘(Genetic Algorithm, GA)

유전자알고리즘은 1960년대에 Holland에 의해 창안되었다. 유전자알고리즘은 자연선택의 법칙인 자연진화 이론으로부터 나온 탐색기법으로 초기에 무작위로 초기 모집단(population)을 형성하고 이들을 부모세대(parent)로 하여 선택(selection), 교배(crossover), 돌연변이(mutation) 등의 연산과정을 거쳐 부모세대보다 진화한 새로운 자식세대(offspring)를 생성하게 된다 (Goldberg, 1989). Fig. 2는 유전자 알고리즘의 개략적인 순서를 나타낸 순서도이다.

3.2 모멘트를 이용한 매개변수 추정

NSRPM의 매개변수를 추정하는 방법으로는 많은 연구가 진행되어 왔으며, 그 중에서 모멘트를 이용하여 매개변수를 추정하는 방법이 가장 일반적으로 사용되고 있다. NSRPM의 모멘트에 대해서는 2.2절에서 설명하였다. 각 분포의 매개변수로 표현된 Eqs. (6)~(8)으로 계산된 이론적인 모멘트와 실제 관측값의 모멘트와 차이가 가장 적게하는 매개변수를 추정한다. Eq. (9)는 Cowperwait *et al.* (1996)이 제안한 식으로 가장 일반적으로 사용되는 목적함수이다. Eq. (9)에서 $\gamma(Y_{i+1}^{(h)}, Y_i^{(h)})$

Fig 1. Process of NSRPM

Fig. 2. Flowchart of geneticalgorithm

는 h시간의 누적 강수량 자료의 lag 1인 자기상관계수를 나타낸다. $\gamma(Y_{i+1}^{(h)}, Y_i^{(h)})$ 는 h시간 누적 강수량의 lag 1 상관계수를 h시간 누적 강수량의 분산으로 나누어 계산할 수 있으며 그 식은 Eq. (10)와 같다.

$$Minimize = (1 - \frac{E(\widehat{Y_i^{(h)}})}{E(Y_i^{(h)})})^2 + (1 - \frac{Var(\widehat{Y_i^{(h)}})}{Var(Y_i^{(h)})})^2 \quad (9)$$

$$+ (1 - \frac{\gamma(\widehat{Y_{i+1}^{(h)}}, \widehat{Y_i^{(h)}})}{\gamma(Y_{i+1}^{(h)}, Y_i^{(h)})})$$

$$\gamma(Y_{i+1}^{(h)}, Y_i^{(h)}) = \frac{Cov(Y_{i+1}^{(h)}, Y_i^{(h+1)})}{Var(Y_i^{(h)})} \quad (10)$$

3.3 직접적인 추정방법

3.3.1 직접적인 추정방법

직접적인 추정방법은 매개변수로 구성된 모멘트식을 이용하지 않고 NSRPM을 통하여 생성된 강수시계열 자료의 통계치와 관측된 강수량의 통계치를 비교하여 매개변수를 추정하는 방법이다. 기존에는 컴퓨터의 성능이 낮고 유전자 알고리즘과 같은 빠른 매개변수 추정 방법이 개발되지 않아 직접적인 방법이 아닌 모멘트를 이용하여 매개변수를 추정하였지만, 현재는 직접적으로

매개변수를 추정하는 것이 가능하다. 직접적인 매개변수 추정방법의 흐름은 Fig. 3과 같다.

Fig. 3. Flowchart of Direct Estimation Method for NSRPM

우선 관측된 강수량 자료의 모멘트를 구한다. 계산되는 모멘트들은 3.2절에서 설명한 모멘트들을 사용한다. 유전자 알고리즘에 NSRPM모형을 내장하도록 프로그램을 작성한다. 꼭 유전자 알고리즘일 필요는 없으나 유전자 알고리즘이 전체(global)해를 찾는 기능이 다른 최적화 방법보다 뛰어나 본 연구에서는 유전자 알고리즘을 사용하였다(참고문헌). NSRPM모형을 통해 생성되는 강수시계열 자료들의 통계치를 계산하며, 계산된 통계치들은 3.2절의 Eq. (4)를 목적함수로 구성한다.

이러한 과정을 반복하여 매개변수를 추정한다. 한번만 매개변수를 추정하지 않고 여러 번 반복하는 이유는 목적함수로 사용되는 항들의 개수가 많아 질 경우 비슷한 수준의 다른 여러 매개변수가 추정되기 때문이다. 목적함수가 비슷한 값의 정확도를 보이나 목적함수의 항이 많아 질 경우 각 매개변수로 생성되는 강수시계열의 값은 다른 양상을 보인다. 그래서 일반적으로 목적함수의 항의 개수를 5 또는 6개로 구성한다(김기욱과 유철상, 2008). 본 연구에서는 정확한 매개변수의 추정이 아닌 매개변수 추정방법들의 정확도를 비교를 목적으로 하기에 목적함수 항의 개수를 고정하였으며 P.S.P Cowperwait *et al.* (1996)의 연구에서 사용한 목적함수 구성을 이용하였다. 목적함수로는 평균, 1시간 강우자료의 분산, lag1 자기상관계수를 6시간 강우자료의 분산, lag1 자기상관계수를 24시간 강우자료의 분산, lag1 자기상관계수를 적용하였다. 구해진 여러 매개변수 중 실제 강우에 유사하다고 생각되는 매개변수를 그 지역의 NSRPM의 매개변수로 정한다.

매개변수의 범위는 Table 1과 같으며 이 매개변수의 범위는 Calenda and Napolitano (1999)의 연구에서 사용되었던 범위를 사용하였다.

3.3.2 무작위변수(random number) 생성 효과

직접적인 추정방법을 위해서는 NSRPM을 이용하여 강수시계열을 생성해야 한다. 추정된 매개변수로 생성된 강수시계열이 특정 통계치로 수렴하기 위해서는 일정 이상의 강수시계열을 생성해야 한다. 일정 이상을 생성하지 않을 경우 생성된 자료의 통계치 변화의 폭이

커 정확한 매개변수를 추정하기 어렵다. 본 연구에서는 이러한 문제를 해결하고자 NSRPM의 자료 생성 개수에 따른 모멘트값의 변화율을 모의해 보았다.

무작위변수 생성 알고리즘은 numerical recipes에 있는 ran0~ran3를 이용하였다(Press *et al.*, 1992). 총 4개의 무작위변수 생성 알고리즘에 대하여 각 알고리즘에 대하여 50개씩의 매개변수를 무작위로 생성하였다. 자료를 1,000에서 100,000까지 1,000개씩 증가시키며 모멘트들을 계산하였다. 총 200개 매개변수로 생성된 강수시계열의 모멘트들을 강우 생성개수를 1,000개씩 증가할 때의 변화율을 계산하였다. 평균, 1시간 누적자료의 분산, 1시간 누적자료의 lag 1자기 상관계수, 6시간 누적자료의 분산, 6시간 누적자료의 lag 1 자기 상관계수, 24시간 누적자료의 분산과 24시간 누적자료의 lag 1 자기상관계수를 비교 하였다.

Fig. 4는 통계치 변화율을 나타낸 그래프이다. Fig. 4를 보면 50,000개 이상부터 통계치의 변화율이 0.2를 중심으로 변하는 것을 확인 할 수 있다. 50,000개 이상부터는 생성 개수를 증가시켜도 변화율이 더 낮아 지지 않는 것을 확인 할 수 있다. NSRPM으로 생성된 강수 시계열을 50,000개 이상 생성해야 모멘트들이 신뢰할 수 있는 모멘트가 나오는 것을 확인 할 수 있다.

직접적인 추정방법을 이용하여 매개변수를 추정할 시 NSRPM의 강수시계열 생성 개수는 최소 50,000개 이상이어야 하며 본 연구에서는 모의 시간을 단축시키고자 50,000개를 생성하여 실험하였다.

Fig. 4. Gradient of Generated Precipitation Statistics

Table 1. Upper and Lower Limit for Parameter Space

Parameter	λ	μ	β	ξ	η
Minimum value	0.001	2.0	0.01	0.3	0.1
Maximum value	0.050	100.0	0.50	15.0	5.0

4. 결과 분석

4.1 대상지점

본 연구에서는 금강유역에 있는 청주지역을 대상으로 자료를 분석하였다. 사용된 자료는 청주 기상청 지상관측소의 1967~2007까지 41년의 6~9월 시강수량을 사용하였다.

4.2 정확도 비교

청주 관측소의 관측 강수 시계열을 이용하여 NSRPM의 매개변수를 추정하였다. Table 2는 각 매개변수로 생성된 강수 시계열의 통계치를 나타낸 것이다. Table 2에서 Observation은 관측 강수 시계열을 나타내고, Moment는 모멘트를 이용하여 추정된 매개변수를 나타내며, Direct는 직접적인 방법을 이용하여 추정된 매개변수를 나타낸다. ran0~ran3는 무작위변수 생성 알고리즘을 나타낸다. Moment의 경우 ran0~ran3까지 같은 매개변수를 사용한다. 모멘트를 이용하여 매개변수를 추정할 시 무작위변수를 생성할 필요가 없기 때문에 한 가지 매개변수만 추정이 된다. 추정된 한 가지 매개변수를 다양한 무작위변수 생성 알고리즘을 이용하여 강수

시계열을 생성하였다. Direct는 직접적인 방법을 이용하여 매개변수를 추정한 것이다. 직접적인 방법의 경우 매개변수를 추정 시 NSRPM이 최적화 모형 안에 내장되기 때문에 무작위변수 생성 알고리즘이 고정되어 진다. 각 무작위변수 생성 알고리즘별로 매개변수가 추정되며 총 4개의 매개변수를 추정하였다. Mean은 1시간 강수량의 평균 나타낸다. Var은 분산이며 ()안에 들어 있는 것은 누적시간을 나타낸다. AC는 자기상관계수를 ()괄호 안에 들어 있는 것은 누적시간을 나타낸다. 예를 들어 Var(6)은 6시간 누적 강수량자료의 분산을 나타낸다.

Table 2를 보면 직접적인 매개변수 추정 방법이 모멘트를 이용하여 매개변수를 추정했을 때보다 평균과 자기상관계수가 잘 맞는다. 모멘트를 이용하여 추정된 매개변수는 평균이나 자기상관계수는 잘 맞지 않으나 분산의 경우 1시간 누적강수 분산의 경우에 높은 적용성을 보이는 것을 확인할 수 있었다.

Table 3은 관측 강수 시계열의 통계치와 NSRPM을 이용하여 생성된 강수 시계열의 통계치의 상대오차를 나타낸 표이다. 평균의 경우 상대오차가 음수로 과소 추정되는 것을 알 수 있다. 1시간 누적 강수 자료의 분산의 경우 모멘트를 이용하여 추정된 매개변수는 잘 맞

Table 2. Statistics from Observed and Generated Precipitation

Type		Mean	Var(1)	AC(1)	Var(6)	AC(6)	Var(24)	AC(24)
Observation		0.2876	3.0556	0.4989	49.7648	0.4036	334.1890	0.3331
Moment	ran0	0.1373	3.9895	0.6729	88.3749	0.1523	700.2495	0.0599
	ran1	0.1399	4.6048	0.6794	100.6252	0.1630	757.1138	0.0902
	ran2	0.1210	3.6292	0.6724	75.7987	0.1588	482.5510	0.2121
	ran3	0.1015	2.7157	0.6708	54.2522	0.1290	393.2336	0.1365
Direct	ran0	0.2111	0.8528	0.4255	15.9608	0.5030	199.3156	0.4656
	ran1	0.2252	0.8809	0.4551	16.5383	0.3020	167.1981	0.1503
	ran2	0.2433	0.6988	0.4295	12.9680	0.4611	161.0822	0.4632
	ran3	0.2092	0.8844	0.4729	17.6799	0.4466	198.5383	0.3987

Table 3. Relative Error

Type		Mean	Var(1)	AC(1)	Var(6)	AC(6)	Var(24)	AC(24)
Moment	ran0	-0.52	0.31	0.35	0.78	-0.62	1.10	-0.82
	ran1	-0.51	0.51	0.36	1.02	-0.60	1.27	-0.73
	ran2	-0.58	0.19	0.35	0.52	-0.61	0.44	-0.36
	ran3	-0.65	-0.11	0.34	0.09	-0.68	0.18	-0.59
Direct	ran0	-0.27	-0.72	-0.15	-0.68	0.25	-0.40	0.40
	ran1	-0.22	-0.71	-0.09	-0.67	-0.25	-0.50	-0.55
	ran2	-0.15	-0.77	-0.14	-0.74	0.14	-0.52	0.39
	ran3	-0.27	-0.71	-0.05	-0.64	0.11	-0.41	0.20

는 것으로 나타났다. 6시간 누적 강수 자료의 분산의 경우도 모멘트를 이용하여 추정된 매개변수가 직접적인 추정방법보다 실제 강수를 잘 모의하는 것으로 나타났다. 24시간 누적 강수 자료의 분산은 직접적인 추정방법으로 추정된 매개변수가 모멘트를 이용한 매개변수보다 정확한 매개변수를 추정하였다. 분산의 경우 모멘트를 이용하여 추정한 매개변수의 경우 상대오차가 양수로 실제 강수보다 집중하여 강수가 분포되어 있다. 직접적인 추정방법으로 추정한 매개변수의 경우 상대오차가 음수로 실제 강수보다 넓게 퍼진 분포를 보였다.

Table 4는 관측 상수 시계열의 통계치와 본 연구에서 적용한 각각의 기법에 의한 결과와의 절대상대오차를 나타낸 표이다. Table 3과는 달리 각 항목들의 평균을 AVE로 표시되어 있다. Table 4의 평균(mean)의 AVE를 보면 모멘트를 이용하여 추정된 매개변수는 0.57의 절대상대오차를 보이고 직접적인 방법은 0.23의 절대 상대오차를 보인다. 분산의 경우 24시간 누적자료만이 직접적인 방법이 모멘트법보다 정확한 것으로 나타났다. 자기상관계수의 경우 모든 시간에 대해서 직접적인 방법이 더 좋은 정확도의 매개변수를 추정하는 것으로 나타났다. 직접적인 추정방법은 평균 절대상대오차가 0.4을 보이고 모멘트를 이용한 추정방법은 0.54을 보였다. 평균적인 정확도를 비교해 보았을 때 직접적인 추정방법이 모멘트를 이용한 추정방법보다 정확한 매개변수를 추정하였다.

모멘트를 이용하여 매개변수를 추정할 경우 한 종류의 최적화 알고리즘을 사용하였을 시 여러 번 반복을 하여도 같은 매개변수가 추정이 되어 한 가지 매개변수만을 추정하게 된다. 추정된 한 가지 매개변수를 이용하여 다양한 무작위변수 생성 알고리즘 적용할 경우 매우 상이한 결과가 도출된다. Table 4를 보면 모멘트를 이용하여 추정한 매개변수의 경우 각 무작위변수 생성 알고리즘의 평균 절대상대오차를 보게 되면 0.38~0.71

까지 넓은 범위에 퍼져있는 것을 확인 할 수 있다. 24시간 누적자료의 분산을 보게 되면 0.18~1.27로 어떤 무작위변수 생성 알고리즘을 사용하느냐에 따라 그 값의 변화가 큰 것을 확인 할 수 있다. 직접적인 추정방법의 경우 각 무작위변수 생성 알고리즘의 평균 절대상대오차가 0.34~0.43으로 모멘트를 이용하여 추정한 매개변수의 비하여 변화폭이 적은 것을 확인 할 수 있다. 24시간 누적자료의 분산을 보면 0.4~0.52로 모멘트를 이용하여 추정된 매개변수의 0.18~1.27보다 변화폭이 확연히 좁아지는 것을 확인 할 수 있다. 직접적인 방법으로 추정된 매개변수가 모멘트를 이용하여 추정한 매개변수보다 무작위변수 생성 알고리즘의 따른 영향이 적은 것으로 나타났다. 직접적인 추정방법은 최적화 모형 안에 무작위변수 생성 알고리즘을 내장하고 있어 알고리즘에 맞는 매개변수를 추정하여, 모멘트를 이용하여 추정하는 방법보다 안정적인 매개변수를 추정한다.

Fig. 5는 매개변수 추정된 관측 강우자료를 도시한 그래프이고 Figs. 6~13은 각 방법으로 추정된 매개변수로 생성된 강우자료를 도시한 그래프이다. Fig. 5~13 그래프의 x축은 공통적으로 시간 (일)을 나타낸다. Figs. 6~9는 모멘트를 이용하여 추정된 매개변수로 생성된 강우를 나타내고, Figs. 10~13은 직접적인 방법을 적용하여 추정된 매개변수로 생성된 강우를 나타낸다. Figs. 6~9를 보면 관측 강우와 비교하여 강우의 밀도와 분포가 성긴 것을 확인 할 수 있다. 강우의 평균이 실제 강우보다 과소추정 되었기 때문에 전체적인 강우의 총량이 적다. 평균이 적으나 분산이 보다 잘 맞는 이유는 강우량이 큰 강우사상이 많이 발생하여 분산이 잘 맞는 것으로 나타났다. Figs. 10~13을 보면 관측 강우와 비교하여 강우 밀도와 분포가 비슷하게 모의된 것으로 나타났다. 그러나 최대 강우량의 경우 비교적 잘 모의하지 못하는 것으로 나타났다. 이러한 이유는 자기상관계수를 맞추기 위하여 분산값이 비교적 부정확한 매개변수를 추정하기에 나타난 현상이라고 생각된다.

Table 4. Absolute Relative Error

Type		Mean	Var(1)	AC(1)	Var(6)	AC(6)	Var(24)	AC(24)	AVE
Moment	ran0	0.52	0.31	0.35	0.78	0.62	1.10	0.82	0.64
	ran1	0.51	0.51	0.36	1.02	0.60	1.27	0.73	0.71
	ran2	0.58	0.19	0.35	0.52	0.61	0.44	0.36	0.44
	ran3	0.65	0.11	0.34	0.09	0.68	0.18	0.59	0.38
Direct	ran0	0.27	0.72	0.15	0.68	0.25	0.40	0.40	0.41
	ran1	0.22	0.71	0.09	0.67	0.25	0.50	0.55	0.43
	ran2	0.15	0.77	0.14	0.74	0.14	0.52	0.39	0.41
	ran3	0.27	0.71	0.05	0.64	0.11	0.41	0.20	0.34
AVE	Moment	0.57	0.28	0.35	0.60	0.63	0.75	0.63	0.54
	Direct	0.23	0.73	0.11	0.68	0.19	0.46	0.38	0.40

Fig. 5 Rainfall observations

Fig. 6 Rainfall generated by ran0 using parameters estimated by moment method

Fig. 7 Rainfall generated by ran1 using parameters estimated by moment method

Fig. 8 Rainfall generated by ran2 using parameters estimated by moment method

Fig. 9 Rainfall generated by ran3 using parameters estimated by moment method

Fig. 10 Rainfall generated by ran0 using parameters estimated by direct method

Fig. 11 Rainfall generated by ran1 using parameters estimated by direct method.

Fig. 12 Rainfall generated by ran2 using parameters estimated by direct method

Fig. 13 Rainfall generated by ran3 using parameters estimated by direct method

5. 결 론

본 연구에서는 NSRPM의 매개변수 추정 시 보다 정확한 매개변수를 추정하고자 직접적인 추정방법을 적용하였다. 이를 위해 기존에 사용하였던 모멘트를 이용하여 매개변수를 추정해 보았고, 직접적인 방법으로 매개변수를 추정하였다. 직접적인 추정을 위하여 무작위변수 생성 개수의 영향에 대하여 모의하였으며, 여러 무작위 생성 알고리즘을 적용하여 그 영향을 보았다. 그 결과 아래와 같은 결과를 얻었다.

1. 직접적인 추정방법으로 추정된 매개변수는 모멘트를 이용하여 추정한 매개변수보다 자기상관계수와 평균에 대하여 정확하며 분산에 대하여서는 모멘트를 이용하여 추정한 매개변수가 더 정확한 매개변수를 추정한다. 전체적인 정확도를 비교해 보았을 때 직접적인 추정방법이 모멘트를 이용하여 추정하는 방법보다 정확한 매개변수를 추정한다.
2. 모멘트를 이용하여 추정된 매개변수의 경우 무작위변수 생성 알고리즘에 따라 같은 매개변수일지라도 다른 통계치를 가지는 강수가 생성된다. 직접적인 추정방법의 경우 최적화 모형에 무작위변수 생성 알고리즘이 내장되어 무작위 생성 알고리즘에 적절한 매개변수가 추정되어 안정적인 매개변수를 추정하는 것으로 나타났다.

직접적인 추정방법이 모멘트를 이용한 추정방법보다 정확하고 안정적인 매개변수를 추정하는 것으로 나타났다.

향후 목적함수의 항의 종류와 그 개수와 지속기간을 바꾸어 가면서 목적함수의 변화에 따른 매개변수추정의 정확도를 비교해 보는 연구가 지속되어야 한다.

감사의 글

본 연구는 21세기 프론티어연구개발사업인 수자원의

지속적 확보기술개발사업단의 연구비지원(과제번호:1-6-3)에 의해 수행되었습니다.

참 고 문 헌

- 금중호, 안재현, 김중훈, 윤용남 (2001). “점강우모형 NSRPM의 매개변수 추정.” **한국수자원학회 학술발표회 논문집**, 한국수자원학회, pp. 206-211.
- 김기욱, 유철상 (2008). “시간적 군집특성을 고려한 강우모의모형의 선정.” **한국수자원학회논문집**, 한국수자원학회. Vol. 41, No. 7, pp. 747-759.
- 김병식, 김보경, 경민수, 김형수 (2008). “기후변화가 극한강우와 I-D-F 분석에 미치는 영향 평가.” **한국수자원학회논문집**, 한국수자원학회, Vol. 41, No. 4, pp. 379-394.
- 김재한, 이정식, 이재준, 손광익 (1996). “Poisson 군집과정을 이용한 일강수계열의 모형화.” **대한토목학회 논문집**, 대한토목학회, Vol. 18, No. II-3, pp. 231-241.
- 유철상, 김대하 (2006). “구형펄스모형을 이용한 가뭄사상의 평가.” **한국수자원학회논문집**, 한국수자원학회, Vol. 39, No. 4, pp. 373-382.
- 유철상, 김남원, 정광식 (2002). “점강우모형과 강우강도-지속기간-생기빈도 해석.” **한국수자원학회논문집**, 한국수자원학회, Vol. 34, No. 6, pp. 577-586.
- Entekhabi, D., Rodriguez-Iturbe, I., and Eagleson, P. S. (1989). “Probabilistic Representation of the Temporal Rainfall by a modified Neyman-Scott Rectangular Pulse Model: Parameter Estimation and Validation.” *Water Resources Research*, Vol. 25, No. 2, pp. 295-302.
- Favre, A. C., Musy, A., and Morgenthaler, S. (2004). “Unbiased Parameter Estimation of the Neyman-Scott Model for Rainfall Simulation with

- Related Confidence Interval.” *Journal of Hydrology*, Vol. 286, pp. 168–178.
- Calenda, G. and Napolitano, F. (1999). “Parameter estimation of Neyman–Scott processes for temporal point rainfall simulation.”, *Journal of Hydrology*, Vol. 225, pp. 45–66.
- Evin, G. and Favre, A.C. (2008). “A new rainfall model based on the Neyman–Scott process using cubic copulas.” *Water Resources Research*, Vol. 44, W03433.
- Goldberg, D.E. (1989). *Genetic algorithm in search, optimization & machine learning*. Addison Wesley, Massachusetts.
- Holland J.H. (1975). *Adaptation in natural and artificial systems*. University of Michigan Press.
- Neyman, J. and Scott, E.L. (1958). “Statistical Approach to Problem of Comology.” *Journal of the Royal Statistical Society*, Vol. B20, No. 1, pp.1–43.
- Cowpertwait, P.S.P., O’Connell, P.E., Metcalfe, A.V. and Mawdsley, J.A. (1996). “Stochastic point process modelling of rainfall. I. Single-site fitting and validation.” *Journal of Hydrology*, Vol. 177, pp. 17–46.
- Rodriguez-Iturbe, I., Cox, D. R., and Isham, V. (1987). “Some Models for Rainfall Based Stochastic Point Process.” *Proceedings of the Royal Society of London*, Vol. A410, No. 1839, pp. 269–288.
- Rodriguez-Iturbe, I., Cox, D. R., and Isham, V. (1988). “A Point Process Model for Rainfall: Further Developments.” *Proceedings of the Royal Society of London*, Vol. A417, No. 1853, pp. 283–298.
- Kim, S. (2004). “On the Dependent Structure Between Rainfall Intensity and Duration.” *KSCE Journal of Civil Engineering*, KSCE, Vol. 8, No. 3, pp. 351–355.
- Kim, S. and Kavvas, M.L. (2006). “Stochastic Point Rainfall Modeling for Correlated Rain Cell Intensity and Duration.” *Journal of Hydrologic Engineering*, ASCE, Vol. 11, No. 1, pp. 29–36.
- Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P. (1992), *Numerical Recipes in C The Art of Scientific Computing Second Edition*, CAMBRIDGE UNIVERSITY PRESS, United States of America, pp. 274–286.

논문번호: 09-110	접수: 2009.09.18
수정일자: 2009.10.22/10.27	심사완료: 2009.10.27