

점강우모형 NSRPM의 매개변수 추정

Parameter Estimation of a Point Rainfall Model, Neyman–Scott Rectangular Pulses Model

금 종 호¹⁾, 안 재 현²⁾, 김 중 훈³⁾, 윤 용 남⁴⁾

1. 서론

본 연구에서는 최적화기법을 사용하여 Rodriguez-Iturbe 등(1984)과 Rodriguez-Iturbe(1986)가 언급한 점강우모형 중에서 군집특성의 고려가 가능하기 때문에 그 적용성이 우수한 것으로 알려진 Neyman–Scott Rectangular Pulses Model(NSRPM)의 매개변수를 추정하였으며, 이를 위해 강우자료의 통계적 특성값인 평균(mean), 분산(variance), 공분산(covariance) 등을 이용하는 모멘트법을 적용하였다.

일반적으로 모멘트법의 적용을 위해서는 시행착오법을 적용하는 경우가 대부분이지만, 본 연구에서는 적용의 편리함과 그 정확성이 더욱 뛰어난 것으로 알려져 있는 최적화기법을 이용하였으며, 최적화기법 중에서 비선형계획법(NonLinear Programming, NLP)과 유전자알고리즘(Genetic Algorithm, GA)으로 모형의 매개변수를 추정하였다. 이와 같은 분석을 통해 NSRPM의 매개변수 추정을 위한 최적화기법의 상대적인 적용성을 평가할 수 있었다.

Calenda와 Napolitano(1999)는 점강우모형 NSRPM의 매개변수의 적합한 범위를 제시하였다. 또한, 류재희 등(2000)은 서울지역을 대상으로 점강우모형 Rectangular Pulses Poisson Model과 Neyman–Scott Rectangular Pulses Model, Bartlett-Lewis Rectangular Pulses Model의 매개변수를 추정하여 계절 특성을 분석하였다.

2. Neyman–Scott Rectangular Pulses Model

일반적으로 유역 전반에 걸쳐 발생하는 강우를 해석할 경우에는 강우관측소가 설치된 지점에서의 측정값만을 가지고 산술평균법이나 Thiessen법 또는 등우선법 등으로 유역의 평균강우량을 산정하게 된다. 그러나, 이와 같은 경우에는 강우관측소의 분포상태, 조밀도, 유역의 지형 등에 따른 강우의 변화를 완전하게 반영할 수 없기 때문에 실제 유역에 내리는 강우를 완벽하게 재현하기는 어렵다는 한계를 가지고 있음은 주지의 사실이다.

점강우모형에서는 점강우자료의 통계적인 특성을 이용하여 강우를 보의함으로써 유역 전체에 내리는 강우를 모의할 수 있다는 장점을 가지고 있다. 이러한 점강우모형 중에서 Neyman–Scott Rectangular Rainfall Model(NSRPM)은 서로 비슷한 특성을 가지는 변량이 서로 모여서 발생하는

1) 고려대학교 대학원 토목환경공학과 석사과정

2) (주)건일엔지니어링 수자원부 차장

3) 고려대학교 토목환경공학과 부교수

4) 고려대학교 토목환경공학과 교수

군집 특성을 가진 모형으로서 점강우모형들 중에서 상대적으로 그 적용성이 우수한 것으로 알려져 있다(Rodriguez-Iturbe 등, 1984).

NSRPM은 강우의 발생을 나타내는 λ (poisson process)와 강우세포의 강도를 나타내는 μ_c (exponential distribution), 강우세포의 위치를 나타내는 β (exponential distribution), 강우세포의 지속시간을 나타내는 η_c (exponential distribution), 마지막으로 강우세포의 개수를 나타내는 random variable인 ν (poisson process) 등의 5개 매개변수를 이용하여 모의되며, 이러한 매개변수들은 실측자료에 대한 분석을 통해 추정할 수 있다.

이와 같은 5개의 매개변수를 포함한 NSRPM의 기본적인 방정식은 식(1)~(5)와 같다.

$$1\text{시간 평균} = \frac{\lambda \cdot E[\nu]}{\eta_c \cdot \mu_c} \quad (1)$$

$$1\text{시간 분산} = \frac{4\lambda \cdot E[\nu]}{\eta_c^3 \mu_c^2} (\eta_c - 1 + e^{-\eta_c}) + \frac{\lambda E^2[\nu]}{\mu_c^2 \beta (\eta_c^2 - \beta^2)} (\beta - 1 + e^{-\beta}) \quad (2)$$

$$1\text{시간 공분산} = \frac{2\lambda E[\nu]}{\eta_c^3 \mu_c^2} (1 - e^{-\eta_c})^2 + \frac{\lambda E^2[\nu]}{2\mu_c^2 \beta (\eta_c^2 - \beta^2)} (1 - e^{-\beta})^2 \quad (3)$$

$$6\text{시간 분산} = \frac{4\lambda \cdot E[\nu]}{\eta_c^3 \mu_c^2} (6\eta_c - 1 + e^{-6\eta_c}) + \frac{\lambda E^2[\nu]}{\mu_c^2 \beta (\eta_c^2 - \beta^2)} (6\beta - 1 + e^{-6\beta}) \quad (4)$$

$$12\text{시간 분산} = \frac{4\lambda \cdot E[\nu]}{\eta_c^3 \mu_c^2} (12\eta_c - 1 + e^{-12\eta_c}) + \frac{\lambda E^2[\nu]}{\mu_c^2 \beta (\eta_c^2 - \beta^2)} (12\beta - 1 + e^{-12\beta}) \quad (5)$$

위의 식 (1)~(5)와 같이 5개의 매개변수에 대해서 5개의 방정식이 존재하기 때문에 각 매개변수의 해를 구하는 것이 가능하며, 본 연구에서는 최적화 기법 중 널리 사용되는 비선형계획법(NLP)과 유전자알고리즘(GA)을 이용하여 각각의 매개변수를 추정하였다.

NSRPM 모형의 적용을 위해 추천되는 각 매개변수의 범위는 <표 1>과 같다(Calenda와 Napolitano, 1999).

<표 1> 각 매개변수의 범위(Calenda와 Napolitano, 1999)

Parameter	λ	μ_c	β	η_c	$E[\nu]$
minimum value	0.001	0.006667	0.01	0.1	2
maximum value	0.050	3.333333	0.50	5.0	100

3. 대상유역과 기초자료

본 연구에서는 소양강댐 유역을 대상 유역으로 하였다. 소양강댐 유역 내에는 강우자료의 이용이 가능한 12개의 우량관측소가 있으며, 강우자료가 공통적으로 존재하는 1987~1991년의 기간 중에서 7월의 시자료를 이용하였는데, 7월의 시자료 중에서 155일의 값을 동시간으로 결정하고 그 중 자료의 상태가 양호한 9개의 지점을 최종적으로 선택하였다. 선택된 동시간 자료의 각 지점별 1시간 평균, 분산, 1st correlation, 그리고 6시간 분산과 12시간 분산에 대한 계산 결과를 <표 2>에 나타내었다.

<표 2> 각 우량관측소에 대한 기초자료

지점		1 hr			6 hr	12 hr
관측소	지점 code	mean	variance	1st correlation	variance	variance
춘천	000101	0.70094	9.7094	0.6125	190.1237	549.1361
인제	000211	0.42798	4.0139	0.6020	75.4978	210.0020
내촌	014600	0.49946	6.7846	0.4963	107.6478	285.1715
서석	014650	0.52425	4.9164	0.5785	79.1200	216.4052
부평	015150	0.47997	5.7910	0.5894	102.5881	274.5531
현리	015350	0.46747	4.8809	0.5706	83.1040	227.2355
창촌	015380	0.58952	6.3426	0.5364	96.1198	267.1887
서화	015490	0.43091	4.7046	0.5292	72.2657	203.5338
방산	015800	0.44390	3.7006	0.6955	72.9083	200.3989

4. 최적화기법의 적용 및 결과비교

본 연구에서는 NSRPM 모형의 매개변수 추정을 위해 적용한 최적화기법으로 비선형계획법과 유전자알고리즘을 선택하였으며, 최적화기법을 적용하기 위해 식(6)과 같은 목적함수를 이용하였다. 최적화의 목적은 관측값과 추정된 값 사이의 오차를 최소화하는 것이며, 각 변수들의 크기에 따른 영향을 줄이기 위하여 분수형태로 목적함수를 구성하였다.

$$\text{Minimize } F = \left(\frac{M_{obs}}{M_{cal}} - 1 \right)^2 + \left(\frac{V_{1,obs}}{V_{1,cal}} - 1 \right)^2 + \left(\frac{C_{obs}}{C_{cal}} - 1 \right)^2 + \left(\frac{V_{6,obs}}{V_{6,cal}} - 1 \right)^2 + \left(\frac{V_{12,obs}}{V_{12,cal}} - 1 \right)^2 \quad (6)$$

여기서, M은 평균, V는 분산, C는 1차 correlation을 뜻하며 분산의 각 첨자는 계산된 분산의 시간간격을 의미한다.

4.1 비선형계획법을 사용한 매개변수 추정

본 연구에서는 비선형계획법(NonLinear Programming, NLP)의 적용을 위해 사용되는 여러 탐색방법 중에서 Davidon-Fletcher-Powell 기법을 이용하였다. 비선형계획법을 적용할 경우에는 초기값의 결정이 필수적인데, 본 연구에서는 시행착오법을 이용하여 목적함수값을 가장 작게 하는 값을 초기값으로 선택하였다. 경계값은 <표 1>과 같으며 iteration은 50,000번을 수행하였다.

이와 같은 방법을 사용하여 추정한 매개변수를 모형의 기본방정식에 대입하여 구한 강우자료의 특성값과 목적함수값은 <표 3>과 같으며, 인제관측소의 목적함수가 가장 작은 값을 가지면서 $8.9971 \times 10^{-10} \sim 1.9606 \times 10^{-2}$ 의 범위에서 목적함수가 결정되었음을 알 수 있다.

<표 3> 비선형계획법을 사용한 매개변수 최적화 결과

지점		1 hr			6 hr	12 hr	optimized value
관측소	지점 code	mean	variance	1st correlation	variance	variance	
춘천	000101	0.70093	9.7109	0.61261	190.0497	549.2706	2.6584×10^{-7}
인제	000211	0.42798	4.0138	0.60199	75.4988	210.0012	8.9971×10^{-10}
내촌	014600	0.49945	6.8675	0.49996	106.4408	285.0164	3.2809×10^{-4}
서석	014650	0.52432	4.8216	0.57118	81.4402	214.7678	1.4208×10^{-3}
부평	015150	0.47996	5.8344	0.59245	101.7186	274.8686	1.5629×10^{-4}
현리	015350	0.46747	4.8850	0.57094	83.0407	227.2207	1.6248×10^{-6}
창촌	015380	0.58653	5.9856	0.51333	81.4661	196.5225	1.9606×10^{-2}
서화	015490	0.43092	4.6881	0.52820	73.1660	201.8014	2.4117×10^{-4}
방산	015800	0.44390	3.6896	0.69367	73.2133	200.1513	3.4768×10^{-5}

4.2 유전자 알고리즘을 사용한 매개변수 추정

유전자 알고리즘(Genetic Algorithm, GA)의 경우 비선형계획법과는 달리 초기조건을 입력하지 않아도 되기 때문에 초기조건에 따른 영향이 없는 것이 특징이다. 그러나, 각 매개변수의 경계값은 비선형계획법의 경우와 동일하게 <표 1>의 값을 사용하였으며, 총 10,000회 iteration을 실시하면 모든 매개변수값이 수렴하였으므로 실행 iteration 수는 10,000회로 하였다.

유전자 알고리즘을 사용하여 추정한 매개변수를 모형의 기본방정식에 대입하여 구한 강우자료의 특성값과 목적함수값은 <표 4>와 같으며, 서석관측소의 목적함수가 가장 작은 값을 가지면서 0.0119~0.0791의 범위에서 목적함수가 결정되었음을 알 수 있다.

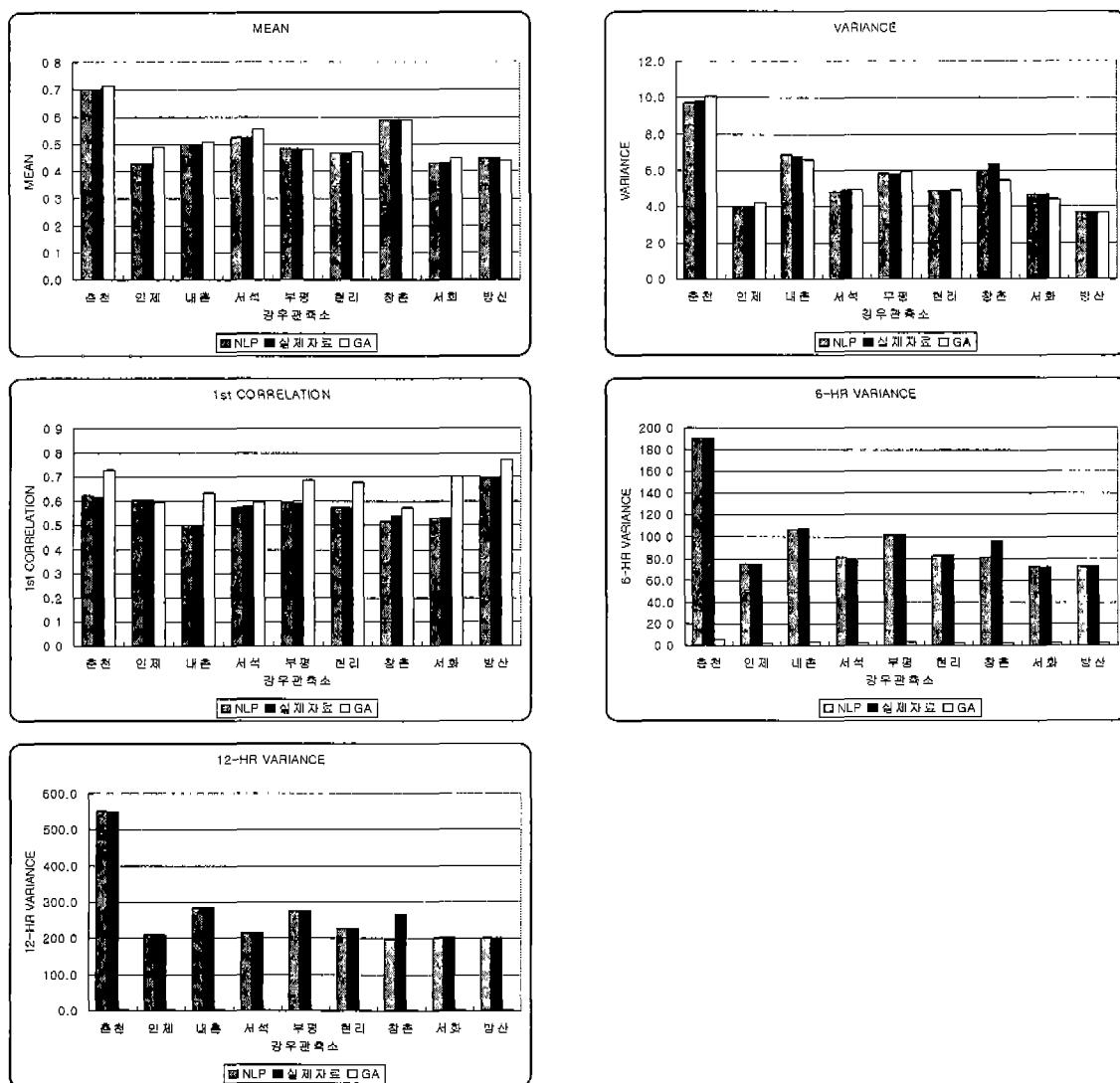
<표 4> 유전자 알고리즘을 사용한 매개변수 최적화 결과

지점		1 hr			6 hr	12 hr	optimized value
관측소	지점 code	mean	variance	1st correlation	variance	variance	
춘천	000101	0.7120	10.0816	0.7270	5.5784	3.5188	0.0362
인제	000211	0.4858	4.2320	0.5920	2.0907	1.3402	0.0251
내촌	014600	0.5080	6.6043	0.6337	3.1518	1.8943	0.0527
서석	014650	0.5535	4.9554	0.5952	2.3068	1.3943	0.0119
부평	015150	0.4774	5.9234	0.6877	2.9745	1.7912	0.0268
현리	015350	0.4705	4.9055	0.6764	2.4434	1.4723	0.0327
창촌	015380	0.5864	5.4650	0.5732	2.4507	1.4656	0.0639
서화	015490	0.4472	4.3986	0.7001	2.2569	1.3649	0.0791
방산	015800	0.4356	3.6392	0.7682	2.1300	1.3611	0.0125

5. 매개변수 추정 결과 및 비교

실제 장우자료의 통계적 특성치와 비선형계획법(NLP)을 사용하여 추정한 매개변수로 구한 통계적 특성치, 유전자알고리즘(GA)를 사용하여 추정한 매개변수로 구한 통계적 특성치를 <그림 1>에 비교하였다.

<그림 1>을 통해 알 수 있듯이 평균과 분산의 경우에는 NLP나 GA 모두 실제자료의 특성을 홀륭하게 추정하였으며, 1st correlation의 경우에는 NLP는 원 자료를 사용하여 구한 값과 비슷한 값을 보였으나 GA를 사용하여 구한 값은 실제자료보다 약간 큰 값을 공통적으로 산출하였다. 또한, 6시간 분산과 12시간 분산의 경우에는 NLP를 사용하여 매개변수를 추정한 결과는 실제자료의 특성을 그대로 나타내었으나, GA를 사용하여 매개변수를 추정한 결과 실제자료의 특성을 전혀 나타내지 못하고 실제자료를 사용한 경우보다 상당히 작은 값을 나타내고 있음을 알 수 있다.



<그림 1> 강우통계특성의 비교

6. 결론

본 연구에서는 점강우모형인 Neyman-Scott Rectangular Pulses Model(NSRPM)의 매개변수를 추정함에 있어서 최근에 널리 사용되고 있는 최적화기법을 적용하였으며, 최적화기법은 탐색성능이 우수한 비선형계획법(NLP)과 유전자 알고리즘(GA)을 이용하였다.

최적화기법을 이용하여 NSRPM의 매개변수를 추정한 결과 NLP의 경우에는 실제 강우자료를 사용하여 구한 통계적 특성치와 거의 일치하는 결과를 나타내었으나, GA의 경우에는 실제 강우자료와는 다소 상이한 결과를 나타내었다. 그러나, NLP는 상대적으로 초기값을 상황에 맞게 설정해야 하는 단점도 가지고 있는 반면, GA의 경우는 초기값에 대한 고려가 필요없는 장점도 상대적으로 가지고 있다. 또한, 본 연구는 단순하게 설정되어있는 강우자료 전체에 대해 매개변수를 추정한 것이었으므로 NLP와 GA 중에서 어느 방법이 그 적용성이 뛰어나다고 판단하기에는 무리가 따른다.

그러나, 단순하게 시행착오법을 이용한 방법에 비해서는 모형 적용을 위한 매개변수 추정에 최적화기법을 적용하는 것이 더 적합할 것으로 판단되었으며, 추후에 월별 또는 연별 변동특성 등을 찾기 위한 매개변수를 시간구간별로 나누어 추정하는 등에 대한 연구의 진행을 통해 NLP나 GA와 같은 최적화기법의 적용성에 대한 분석이 이루어진다면 더 나은 결과를 도출해낼 수 있을 것으로 사료된다.

참고문헌

- 류재희, 유철상, 윤용남(1999). 점 강우 모형의 매개변수 추정을 통한 서울지점 강우의 계절특성 분석, 1999년 학술발표회논문집, 대한토목학회, pp. 257-260.
- 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.
- GAOT, the Genetic Algorithm Optimization Toolbox, <http://www.ie.ncsu.edu/gaot/>
- Mays, L.W. and Tung, Y-K(1992). *Hydroystems Engineering and Management*, McGraw-Hill.
- Rodriguez-Iturbe, I.(1986). Scale of Fluctuation of Rainfall Models, *Water Resources Research*, Vol. 22, No. 9, pp. 15~37.
- Rodriguez-Iturbe, I., Gupta, K. and Waymire, E.(1984). Scale Considerations in the Modeling of Temporal Rainfall, *Water Resources Research*, Vol. 20, No. 11, pp. 1611~1619.