

SCE-UA 및 GA 알고리즘을 이용한 NWS-PC 모형의 매개변수 추정

Estimation of NWS-PC Model Parameters using SCE-UA and Genetic Algorithm

○강경석¹⁾, 강신욱²⁾, 이동률³⁾

1. 서론

일반적으로 유역유출 모형의 목적은 실제와 가장 근사하게 유출현상을 재현하는 것으로서, 이를 위해서는 합리적인 모형구축이 제일 중요하지만 일관성 있는 최적의 모형 매개변수를 결정하는 것도 중요하다. 그러나 개념적 모형들의 강우-유출을 지배하는 매개변수들은 추정하기가 어려울 뿐만 아니라 추정과정에서 실제 현상과는 상이한 결과를 출력하는 경우가 발생하게 된다. 이러한 문제점을 해결하기 위한 연구로 과거 40년 동안 최적화기법과 자동화기법 등이 개발되어 왔다.

이러한 연구들과 더불어 지난 40년간의 과학기술은 인간의 지적인 능력을 모방하는 시스템을 개발하기 위한 연구에 집중되어 신경망 이론 및 유전자 알고리즘 등을 개발하여 여러 분야에서 성공적으로 응용되고 있으며, 최근 수자원분야에도 그 적용성에 대한 연구가 활발하게 적용되고 있다.

따라서 본 연구에서는 비선형 최적화 문제에서 효과적인 탐색 방법으로 알려져 있는 유전자 알고리즘과 개념적 강우-유출 모형의 전역해를 추정하는데 합리적이라 알려진 SCE-UA 방법을 이용하여 토양의 물 이동 경로가 비교적 실제 구조에 가깝도록 설계된 토양함수 모형(Sacramento Soil Moisture Accounting Model, SAC-SMA)과 HEC-1의 운동파(Kinematic wave) 하도추적 모형을 결합하여 만든 미 기상국의 NWS-PC 모형의 매개변수 추정을 위한 활용성을 검토하고, 유출량 산정의 정확성 향상 및 예측에 효율적으로 적용할 수 있도록 하였다.

2. NWS-PC 모형의 개요

미국 국립 기상국(National Weather Service)의 수문예측 사업 그룹은 강수, 눈의 축적과 용설, 토양 함수상태의 계산, 유출의 하도추적, 매개변수 최적화 등의 유출 예측체계를 컴퓨터 프로그램화하여 NWSRFS(National Weather Service River Forecast System)을 개발하였다. 이 NWSRFS의 최종 모형은 프로그램의 규모가 방대하여 Main-frame 전산기만을 사용해야 수행 될 수 있어 NWSRFS의 축소 모형인 NWS-PC(Tabios III et. al, 1986)모형을 개발하였다.

1) : 한국건설기술연구원 수자원환경연구부 선임연구원(hydrokks@kict.re.kr)

2) : 한국건설기술연구원 수자원환경연구부 연구원(sukang@kict.re.kr)

3) : 한국건설기술연구원 수자원환경연구부 수자원연구그룹장(dryi@kict.re.kr)

NWS-PC모형은 NWSRFS의 SAC-SMA모형과 HEC-1의 운동파(kinematic wave)추적 프로그램을 조합하여 구성되었으며 추적방법의 대안으로 단위도와 Muskingum 방법을 결합하여 적용할 수 있도록 개발되었다. NWS-PC 모형은 15개의 부프로그램으로 구성되어 있으며 모형의 흐름도는 그림 1~2와 같다.

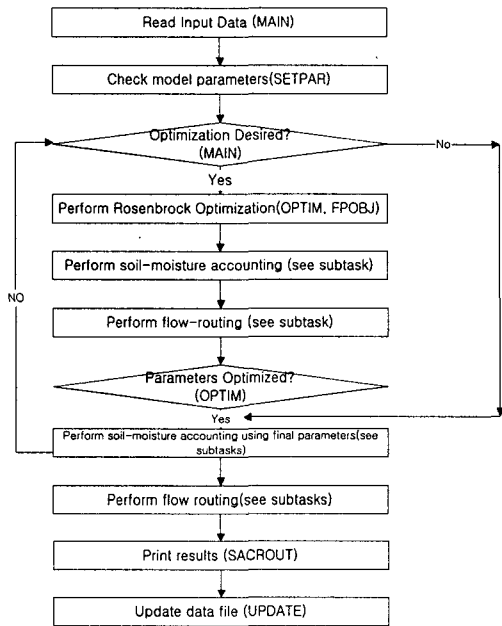


그림 1. NWS-PC 모형 흐름도

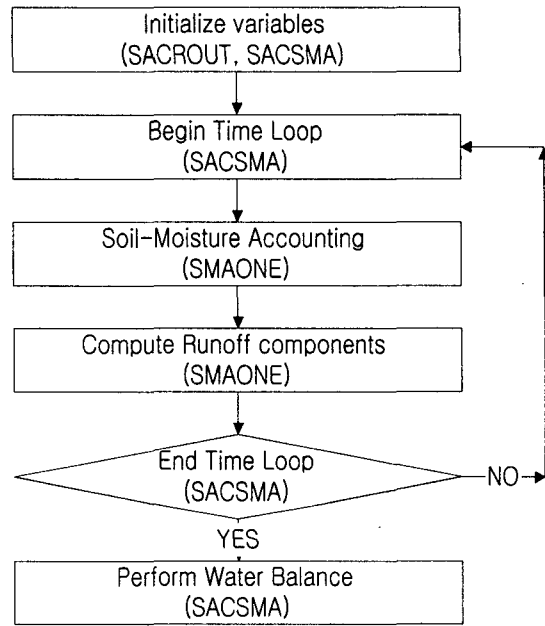


그림 2. SAC-SMA 모형 흐름도

3. 전역 최적화방법

NWS-PC 모형에서 SAC-SMA 모형의 매개변수 추정 방법은 시행착오법과 자동보정방법을 사용할 수 있으며, 자동보정시에는 Rosenbrock 제약 최적화 기법(Kuester and Mize, 1973)을 이용한다. 그러나 SAC-SMA 모형과 같이 다수의 매개변수를 가진 유출모형의 경우에 목적함수를 최적화시키는 매개변수 군은 무수히 많으므로 지역최적화 기법을 이용할 경우에 보다 최적조건에 근접하여 수렴한 해를 찾기는 어렵다. 이러한 문제를 해결하기 위하여 전역최적화 기법 중 개념적 강우-유출모형의 전역해를 구하기 위하여 많이 사용되는 SCE-UA(shuffled complex algorithm - University of Arizona ; Duan 등, 1992) 및 유전자(Genetic) 알고리즘을 SAC-SMA 모형의 매개변수 추정방법으로 이용하였다.

3.1 SCE-UA 알고리즘

SCE-UA는 자연진화의 과정을 최적화 탐색과정에 도입한 것으로, s개의 선정된 점을 이용하여 하나의 모집단을 구성하며, 이 모집단을 몇 개의 집합체(complex)로 다시 분할하여 각각 다른 방향으로 탐색 진행하며 진화한다. 계속적인 진화의 과정에서 집합체는 재분할되며, 각 콤플렉스는

simplex 알고리즘을 사용하여 독립적으로 진화된다. 이 과정에서 각 집합체에서 얻은 탐색공간에 대한 정보를 공유한다.

SCE-UA는 이러한 경쟁적 진화(competitive evolution)와 혼합(shuffling)을 이용한 집합체의 재생산을 통하여 모집단에 포함된 정보가 다음 세대로 전달되도록 하며, 이는 진화의 과정에서 탐색 공간에 대한 정보가 퇴보되는 것을 막아준다. 이러한 특징은 SCE-UA가 광범위한 문제의 영역에서 전체 최적해를 효과적으로 탐색할 수 있도록 한다. 진화와 혼합의 과정은 수렴조건이 만족될 때까지 반복된다. 자세한 과정은 Duan(1991)의 연구를 참고할 수 있다.

3.2 유전자 알고리즘(GA)

유전자 알고리즘(Genetic Algorithm ; GA)은 John Holland(1975)가 그의 저서 "Adaption in Natural and Artificial Systems"에 처음으로 소개된 기법으로, 자연선택과 유전 메커니즘에 입각해 통계적으로 탐색하는 진화적 과정의 알고리즘이며, 자연생태계의 진화과정에서 관찰된 "적자생존의 원리"를 컴퓨터 알고리즘과 결합시켜 정립된 확률적 탐색 최적화 알고리즘이다.

유전자 알고리즘은 그 개념과 이론이 단순하고, 해의 탐색성능이 우수하여 여러 분야의 최적화 또는 의사결정문제에 다양하게 적용되고 있다. 특히 복잡한 해 공간의 탐색성능이 우수하여 변수와 제약이 많은 문제를 푸는데 적합한 기법이고, 또한 모형에 대한 유연성이 높아 제약식이나 목적함수의 변경이 용이하다는 장점이 있다.

유전자 알고리즘이 다른 전통적인 최적화 방법들에 비해서 가지는 특징은 다음과 같다.

- (1) 단일해가 아닌 해의 집단을 동시에 진화시켜 나가는 방식이다. 유전자 알고리즘에서는 점(point)이 아닌 군(population)에 기반한 방법이다.
- (2) 유전자 알고리즘에서는 확정적(deterministic) 규칙이 아닌 확률적 천이 규칙을 따른다.
- (3) 목적함수의 무제약성으로, 대상함수에 대하여 연속성이나 미분가능성 등의 제약성을 요구하지 않는다.
- (4) 국부해(local optimum)에 수렴하는 문제를 효과적으로 억제한다. 하지만 최적해 근처에 도달하여 최종 목표값으로 수렴하기 위해서는 많은 시간을 소모하기도 한다.

4. 매개변수 추정 및 검증

본 장에서는 NWS-PC 모형의 운영에 필요한 기본 입력자료를 작성하고, 모형 매개변수의 추정 및 검증을 실시하여 최적 매개변수를 결정하고, 이를 이용하여 모형의 적용 가능성을 검토하고자 대청댐 상류의 용담댐 지점을 대상구역으로 선정하여 NWS-PC 모형의 매개변수를 추정하였다.

4.1 모형의 입력자료 구성

NWS-PC 모형을 이용하여 유역의 유출량을 모의하기 위해서 요구되는 자료는 강수량과 증발산량 및 유출량 등의 수문자료와 하도추적을 위한 유역특성자료가 필요하다. 따라서 본 연구에서는 수문자료의 경우 '수자원장기종합계획(건설교통부, 2000)'과 유역특성자료의 경우 '수자원계획의 최적화 연구(건설교통부, 2000)'에서 사용된 용담댐 유역의 자료를 이용하였다.

4.2 매개변수 추정 및 검증

용담댐 유역에 대해 유전자 알고리즘과 SCE-UA 방법을 이용하여 SAC-SMA 모형의 16개 토양함수상태 매개변수 중 SIDE를 제외한 15개 매개변수를 추정하였다. 목적함수는 오차제곱합을 최소화하는 것을 사용하였다. 그리고 두 알고리즘을 비교하기 위한 비교 지표로 평균편차의 비율(PBIAS)와 Nash-Sutcliffe 통계량(NS)과 관측 및 모의 유출량의 통계적 특성을 비교 검토하였다.

비교지표인 PBIAS는 '0'에 가까울수록 모의치와 관측치가 근접하는 것이며, NS는 '1'에 가까울수록 관측치와 모의치의 오차가 적다는 것을 나타낸다.

$$PBIAS = \frac{\sum_{i=1}^N (q_i^{obs} - q_i^{sim})}{\sum_{i=1}^N q_i^{obs}} \times 100 \quad (1)$$

$$NS = 1 - \frac{\sum_{i=1}^N (q_i^{sim} - q_i^{obs})^2}{\sum_{i=1}^N (q_i^{obs} - q^{mean})^2} \quad (2)$$

여기서, N 은 자료의 수, q_i^{obs} 는 관측유량, q_i^{sim} 은 모의유량, q^{mean} 은 관측유량의 평균값이다.

표 1. GA 및 SCE-UA 알고리즘에 의해 추정된 매개변수

매개변수	GA	SCE-UA	매개변수	GA	SCE-UA
UZTWM	12.954	36.076	LZTWM	56.989	31.921
UZFWM	2.498	2.920	LZFSM	587.262	996.226
UZK	0.525	0.710	LZFPM	609.651	808.812
PCTIM	0.087	0.100	LZSK	0.021	0.037
ADIMP	0.362	0.415	LZPK	0.002	0.002
RIVA	0.074	0.125	PFREE	0.364	0.018
ZPERC	69.760	129.168	RSERV	0.265	0.208
REXP	2.843	3.435	SIDE	0.0	0.0

용담댐유역의 1970년 1월 1일 ~ 1973년 12월 31일의 자료를 이용하여 유전자 알고리즘과 SCE-UA 알고리즘에 의해 추정된 매개변수는 표 1과 같다. 표에서 살펴보면 대부분의 매개변수가 적정 범위내에 존재하고 있을 뿐만 아니라 유사한 값을 보이고 있으나 포화상에서 건조상으로 변할 경우 증가되는 침투량에 관계되는 계수인 ZPERC는 유전자 알고리즘과 SCE-UA 알고리즘이 상당한 차이를 보이고 있다. 이러한 이유는 ZPERC를 산정하기 위해 요구되는 LZTWM, LZFSM, LZFPM, LZSK, LZPK 등 5개의 매개변수 중 하나라도 잘못 추정되면 ZPERC는 상당한 오차를 발생하게 된다. 따라서 GA와 SCE-UA에 의해 산정된 상기 5개 매개변수가 서로 달라 발생한 것으로 판단된다.

상기 표 1의 매개변수에 의해 각 방법별로 모의된 모의 유출량의 통계적 특성과 비교지표는 표 2와 같고, 이를 살펴보면 최대치의 경우 두 방법 모두 관측치에 비해 과소하게 모의하였으나 기타 통계적 특성은 관측유출량과 유사한 결과를 제시하고 있다. 뿐만 아니라 비교 지표를 살펴보면 NS의 경우 SCE-UA에 의한 결과가, PBIAS의 경우는 유전자 알고리즘에 의한 결과가 우수한 것으로 나타났으나 두 경우 모두 큰 차이를 보이지는 않았다. 그림 3과 4는 추정된 매개변수에 의해 모의된 용담댐의 유출수문곡선이다.

표 2. 용담댐유역 유출량의 통계적특성
(매개변수 추정시)

(단위 : m³/sec)

항 목	관 측	계 산	
		GA	SCE-UA
평 균	24.55	20.58	20.59
최 소	2.17	3.02	2.47
최 대	1,149.56	540.94	539.66
표준편차	66.80	43.41	43.07
1사분위	5.42	5.30	5.30
2사분위	9.76	7.60	8.38
3사분위	18.44	14.02	15.48
PBIAS(%)	-	7.31	7.59
NS	-	0.70	0.67

표 3. 용담댐유역 유출량의 통계적특성
(매개변수 검증시)

(단위 : m³/sec)

항 목	관 측	계 산	
		GA	SCE-UA
평 균	21.44	21.70	21.51
최 소	2.17	2.93	2.67
최 대	896.87	649.63	662.29
표준편차	52.11	47.42	47.82
1사분위	5.42	5.33	5.10
2사분위	8.68	7.34	7.52
3사분위	16.27	14.50	15.12
PBIAS(%)	-	-1.19	-0.30
NS	-	0.58	0.58

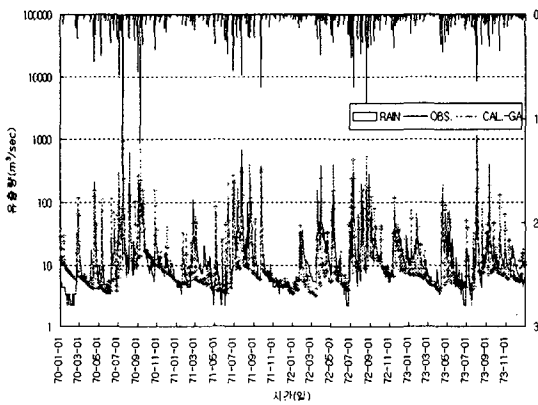


그림 3. 추정 매개변수에 의한 용담댐 유역 유출수문곡선(1970.1.1 ~ 1973.12.31, GA)

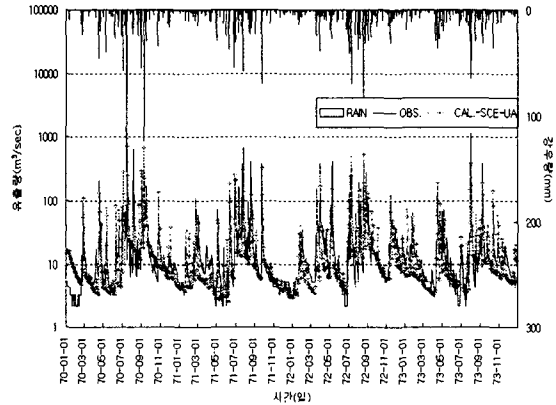


그림 4. 추정 매개변수에 의한 용담댐 유역 유출수문곡선(1970.1.1 ~ 1973.12.31, SCE-UA)

한편 유전자 및 SCE-UA 알고리즘에 의해 추정된 매개변수의 검증을 위하여 용담댐 유역의 1974년 1월 1일 ~ 1976년 12월 31일의 유출수문곡선을 각 방법별로 추정된 표 1의 매개변수를 이용하여 모의한 결과 모의 유출량의 통계적 특성은 표 3과 같고 이를 도시하면 그림 5 및 그림 6과 같다. 모의 유출량의 통계적 특성을 살펴보면 최대치의 경우 관측치에 비해 GA는 약 28%, SCE-UA는 약 26% 과소하게 모의하였으나, 기타 통계적 특성은 관측유출량과 유사한 결과를 보여주고 있다. 또한 각 방법별 비교 지표를 살펴보면 유전자 알고리즘에 의한 경우 PBIAS는 -1.19%, NS는 0.58, SCE-UA 방법의 경우는 PBIAS가 -0.30%, NS는 0.58로 나타나 PBIAS는 SCE-UA에 의한 결과가 우수하게 나타났으며 NS의 경우 두 가지 방법 모두 동일한 값을 나타내었다.

뿐만 아니라 관측 및 모의 유출수문곡선을 살펴보면 전반적으로 강우 입력에 따라 유역 유출 반응이 유사하게 거동하고 있어 본 연구에서 유전자 알고리즘과 SCE-UA 알고리즘에 의해 추정된 매개변수들은 상식적으로 물리적인 범위내에 있음을 알 수 있다.

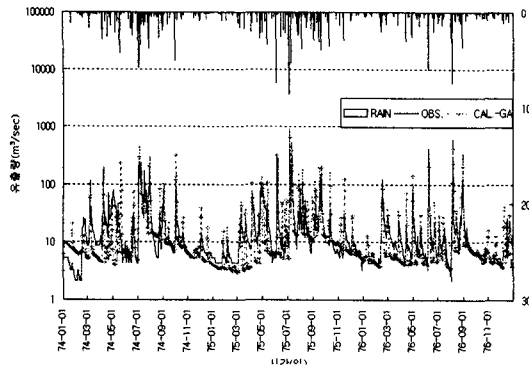


그림 5. 추정 매개변수에 의한 용담댐 유역 유출수문곡선(1974.1.1 ~ 1976.12.31, GA)

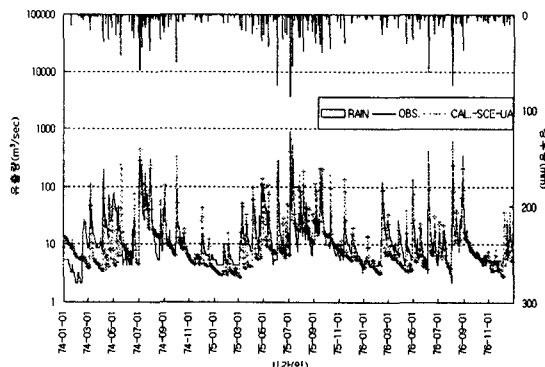


그림 6. 추정 매개변수에 의한 용담댐 유역 유출수문곡선(1974.1.1 ~ 1976.12.31, SCE-UA)

5. 결론

본 연구에서는 NWS-PC 모형의 매개변수를 추정하기 위하여, 대상유역의 1970년 1월 1일 ~ 1973년 12월 31일의 유출량 자료를 이용하여 유전자 알고리즘과 SCE-UA 알고리즘에 의한 SAC-SMA 모형의 최적 매개변수를 추정하였으며, 각 방법에 의해 추정된 매개변수를 검증하기 위하여 1974년 1월 1일 ~ 1976년 12월 31일의 모의 및 관측유량 계열의 통계적 특성인 평균, 표준편차, 최대, 최소, 1사분위값, 2사분위값 및 3사분위값과 PBIAS와 NS를 비교 검토하였다. 그 결과 유전자 알고리즘과 SCE-UA 알고리즘에 의한 모의유량 계열이 최대유량을 제외한 분석항목에서 관측 유량계열과 근사한 통계적 특성치를 재현하여 다수의 매개변수를 가지는 개념적 강우-유출 모형의 매개변수 최적화를 위한 방법으로 활용가능성이 높은 것을 알 수 있었다.

6. 참고문헌

- 강경석, 서병하 (1998). "SAC-SMA 모형을 이용한 일 유출량 산정." '98년 한국수자원학회 학술발표회 논문집, 한국수자원학회, pp. 146 ~ 152.
- 강경석, 서병하 (2001). "유전자 알고리즘을 이용한 SAC-SMA 모형의 매개변수 추정." '2001년 한국수자원학회 학술발표회 논문집, 한국수자원학회, pp. 341 ~ 346.
- 건설교통부 (2000). 수자원계획의 최적화연구(IV)(기후변화에 따른 수자원계획의 영향평가).
- 건설교통부 (2001). 수자원장기종합계획.
- Duan, Q. (1991). *A Global Optimization Strategy for Efficient and Effective Calibration of Hydrologic Models*, Ph.D. dissertation, Dept. of Hydrol. and Water Resour., Univ. of Arizona, Tucson.
- Duan, Q., Sorooshian, S. and Gupta, V.K. (1992). "Effective and efficient global optimization for conceptual rainfall-runoff models." *Water Resources Research*, Vol. 28(4), pp. 1015 ~ 1031.
- Goldberg, D.E.(1989). *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley Publishing Company. Inc..