

유전자 알고리즘을 이용한 강우-유출모형의 매개변수 추정에 관한 연구

A Study on Parameters Estimation of Rainfall-Runoff Model
by Using the Genetic Algorithms

○박봉진*, 차형선**, 김주환***

1. 서론

현재 국내 하천의 홍수예측을 위한 홍수예경보시스템과 다목적댐의 홍수조절을 위하여 사용되고 있는 수문학적 모형으로는 저류함수모형이 주로 사용되고 있다. 저류함수모형의 매개변수 추정방법은 홍수수문자료 등의 부족으로 인하여 일본의 利根川공식과 木村공식을 주로 사용하고 있으나, 우리나라의 유역특성을 충분히 반영하지 못하는 등 많은 문제점을 가지고 있다. 따라서 그동안 이러한 문제점을 해결하고 우리나라 유역특성에 적합한 매개변수를 산정하는 데 많은 노력을 기울여 왔지만 홍수 수문자료의 부족과 이 모형이 가지는 자체의 한계성 등 여러가지 문제점으로 인하여 아직까지 만족할 만한 성과를 얻을 수 없었다. 우리나라에서는 유출모형의 매개변수를 자동보정기법으로 추정하기 위하여 Standardized Davidon-Fletcher-Powell(남궁달, 1985), Partten Search(1991), 최적화기법(심순보, 1992), 퍼지기법(이정규 등, 1996) 등의 방법을 적용하는 다양한 연구가 진행되고 있다. 외국에서는 최근에 들어 유출모형의 자동보정을 위하여 유전자 알고리즘(Shie. Yui Liang 외 2인 1995 ; Wang Q.J. , 1991)의 적용에 관한 연구논문들이 발표되고 있다. 본 연구에서는 국내하천의 홍수예경보시스템의 운영과 다목적댐의 홍수조절관리를 위하여 사용되고 있는 수문학적 모형인 저류함수법의 매개변수 산정을 위하여 자연계의 적응 및 도태에 의한 Darwinian의 진화과정을 수학적으로 체계화시킨 유전자 알고리즘(Genetic algorithmss)을 적용하여, 대청댐 유역의 과거 홍수사상을 검토하고 최적 매개변수를 산정하여 제시 하고자 하였다.

2. 유전자 알고리즘

유전자 알고리즘은 Darwinian의 적자생존에 바탕을 둔 가장 많이 알려진 진화이론의 알고리즘이다. 이는 생명체의 자연도태을 기계적인 학습영역에 적용한 탐색 알고리즘의 일종으로 Holland에 의한 1975년에 최초로 제안되었다. 유전자 알고리즘은 무작위로 초기모집단(Init Population)을 형성하고 이들을 부모세대로 하여 선택(Selection), 교차(Crossover), 돌연변이(Mutation) 등의 연산과정을 거쳐 부모세대 보다 진화한 새로운 세대를 생성하게 된다.

2.1 모집단

유전자 알고리즘은 초기모집단을 무작위로 선택된 염색체들을 형성함으로써 시작된다. 일반적으로 유전자는 "1"과 "0"의 2개의 문자로 구성된 2진수(Binary Code)를 변환하여 생성하게 되는데 염색체는 이들 "1"과 "0"으로 구성된 유전자들이 집합이다. Goldberg(1985)는 24비트를 가지는 염색체에 대하여 51개의 모집단 크기를 제안한 바 있다.

* 한국수자원공사 특수사업본부 굴포천 건설사무소 과장

** 한국수자원공사 낙동강사업본부 과장

*** 한국수자원공사 수자원연구소 선임연구원

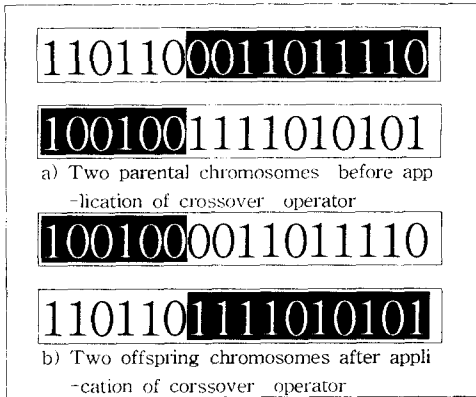


그림 1. 교차의 모식도

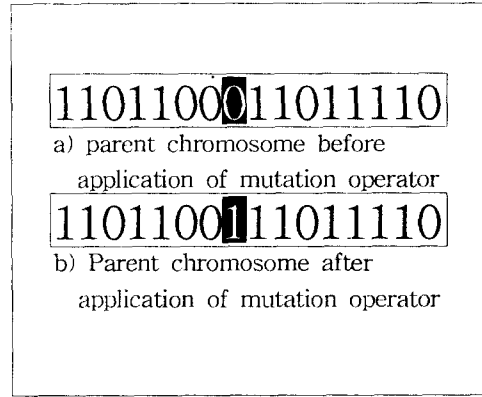


그림 2. 돌연변이의 모식도

2.2 재생(Reproduction)

2.2.1 선택

선택은 Charls Darwin의 적자생존의 철학을 적용하는 연산자의 하나이다. 주어진 환경에 가장 적합하고, 가장 강력한 유전자가 더 많은 생존의 기회를 갖게 되며 다음 세대를 구성하기 위하여 선택되게 된다. 새로운 세대는 구세대의 염색체들 중에서 적합도에 따라 비례선택하여 새로운 세대를 형성하게 된다.

2.2.2 교차

교차는 한쌍의 염색체가 유전형질을 서로 교환하는 유전학 기법의 하나이다. 선택과정으로부터 얻어진 모집단에 있는 부모중에서 한 쌍이 무작위로 추출되어 교차라는 과정을 거치게된다. 선택된 염색체가 재결합 하기 위한 교차위치는 무작위로 선택되며, 교차의 시행후 두 가지 새로운 자손을 형성하게된다. 교차위치는 균등분포를 따른다. 교차에 의해 생성된 자손은 그들의 부모를 대신하여 그림 1에서와 같이 새로운 세대를 형성하게 된다.

2.2.3 돌연변이

돌연변이는 아주 작은 확률로 염색체 내의 비트 하나가 변환하여 발생하는 염색체의 우연적이고 임의적인 변화이다(그림 2 참조). 1987년 Goldberg는 돌연변이의 발생확률을 0.01로 제안하였다(Goldberg, 1989).

3. 유출모형의 적용

3.1 저류함수법

홍수유출계산을 위한 저류함수법(Kimura, 1961)은 유역과 하도에서의 유출량과 저류량의 관계를 표시하는 저류함수를 홍수과의 운동방정식에 대입하여 홍수과의 연속방정식을 계산함으로써 홍수유출량을 결정하는 방법이다. 저류함수법은 홍수유출을 Manning 공식으로 표시할 수 있는 지표면유출이라 가정하여, 유역과 하도구간의 저류량 S를 유출량의 Q의 지수함수로 나타내었다.

$$S_l = K Q_l^p \quad (1)$$

식 (1)은 홍수과의 운동방정식으로써 S_l 은 유역의 저류량, Q_l 은 유역의 유출량, K, p는 유역 및 하도의 저류상수이다.

유역에 대한 연속방정식은 다음의 물수지방정식으로 나타낼수 있다.

$$\frac{1}{3.6} f r_{ave} A - Q_t = \frac{dS_t}{dt} \quad (2)$$

여기서 f 는 유역유입계수, r_{ave} 는 시간당 유역 평균강우량, A 는 유역면적(km²), $Q_t(t) = Q(t + T_t)$ 으로서 유역의 지체시간 T_t 을 고려한 유역으로부터의 직접유출량(m³/sec), S_t 은 유역저류량(m³/sec)이다. 하도에 대한 연속방정식은 다음과 같이 표시할 수 있다.

$$\sum_{j=1}^n f_j I_j - Q_t = \frac{dS_j}{dt} \quad (3)$$

여기서 f_j 는 하도의 평균유입계수, I_j 는 유역 및 지류역에서 하도로 유입하는 유입량, 또는 하도구간의 상류단으로의 유입량이다. $Q_t(t) = Q(t + T_t)$ 은 지체시간 T_t 을 고려한 하도구간 하류단의 유출량(m³/sec)이다. S_j 는 하도구간에 저류되는 저류량(m³/sec)이다.

4. 매개변수의 산정

저류함수법에 의한 홍수유출계산에 사용되는 매개변수는 유역유출과 하도유출의 두가지 형태로 구분할 수 있다. 유역의 유출을 결정하기 위한저류함수의 상수에 해당하는 저류상수 K , P 및 지체시간(T_t)과 유출상태가 포화점에 이르기까지의 1차유출율(f_1), 포화점까지의 누가우량인 포화우량(R_{sa})과, 하도구간의 유출현상을 해석하기위한 하도구간의 저류상수 K , P 및 지체시간(T_t)을 결정하여야 한다.

4.1 유전자 알고리즘의 적용

본 연구에서 저류함수의 매개변수를 보정하기 위한 유전자 알고리즘의 적용을 위하여 우선 저류함수의 매개변수를 일본 利根川유역의 경험식을 이용하여 각 소유역별로 산정된 과거의 매개변수값(다목적댐 홍수유출해석 및 홍수기 저수지 운영프로그램 설명서, 1993)으로 고정시켜 놓고 이 값에 특정 비율벡터 $\Lambda = (\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_n)$ 를 곱하여 식 (4)와 같이 조정함으로써 이루어지도록 하였다.

$$\theta c = \theta e \cdot \Lambda^T \quad (4)$$

여기서 θe 는 경험공식으로 산정된 매개변수 $K, P, T_t, f_1, R_{sa}, K', P', T_t'$ 이며, θc 는 매개변수의 보정치이다. 유출모형의 매개변수를 보정하기 위한 유전자 알고리즘의 적합도의 판단기준은 식 (5)와 같이 관측유량과 모의유량의 시간별 오차의 제곱의 합(Sum of Square of Residual : SSR)을 목적함수로 하였다.

$$\text{Minimize } f(E) = \sum_{i=1}^n (Q_o - Q_s)^2 \quad (5)$$

여기서 Q_o 와 Q_s 는 각각 관측유량과 모의유량을 나타낸다.

본 연구에 적용한 유전자 알고리즘의 연산자 값은 한세대의 모집단 수를 45개, 교차될 확률을 0.6, 돌연변이의 발생 확률은 0.01으로 결정하였으며, 진화를 위한 발생 세대수는 각각 10세대, 20세대, 50세대, 100세대 등으로 모의운영하여 수렴하는 정도를 검토한 결과, 이 중에서 SSR값이 90% 이상 개선되어 수렴상태가 적합하다고 판단되는 50세대로 선정하였다. 저류함수모형에서 보정해야 할 매개변수는 유역 및 하도의 매개변수를 포함하여 모두 8개이므로 각 매개변수에 해당하는 유전자의 길이는 4비트로 결정하였다. 길이가 4비트인

유전자 하나가 생성할 수 있는 비율벡터는 16가지 이므로 각 유전자의 형성에 필요한 비율벡터의 결정값(Decision value)을 표 1과 같이 결정하였다.

표 1. 유전자형성에 필요한 비율벡터의 결정값

10진수	2진수	특성값
0	0000	0.4
1	0001	0.5
2	0010	0.6
3	0011	0.7
4	0100	0.8
5	0101	0.9
6	0110	1.0
7	0111	1.1
8	1000	1.2
9	1001	1.3
10	1010	1.4
11	1011	1.5
12	1100	1.6
13	1101	1.7
14	1110	1.8
15	1111	1.9

5. 적용결과

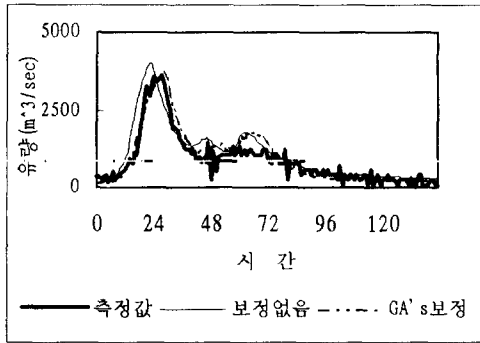
본 연구에서 선정한 대상홍수자료는 대청댐의 홍수기록 중 12개 사상이다. 여기서 선정된 홍수사상중에서 8개의 홍수사상은 유전자 알고리즘을 적용하여 매개변수를 보정하고, 4개의 홍수사상은 매개변수 보정후 매개변수의 개선을 검증에 위하여 사용하였다. 각 홍수사상은 기존의 경험공식으로 산정된 매개변수값을 이용하여 모의운명을 실시하였으며, 유전자 알고리즘에 의한 매개변수 보정을 50세대씩 3회씩 실시하여 비교·검토하였다.

분석결과 그림 3에서 보는 바와 같이 유전자 알고리즘을 적용하였을 때의 모의운영 결과는 매개변수의 보정전과 비교하여 첨두홍수량 및 첨두홍수량 도달시간 등 모든면에서 향상되었다.

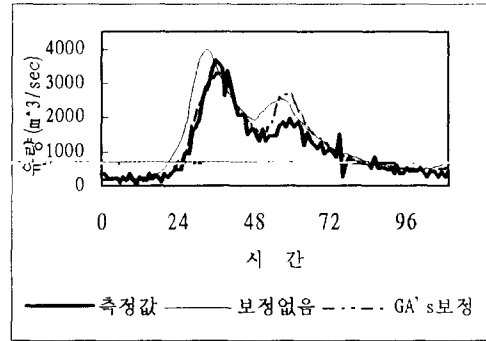
저류함수법의 매개변수들의 민감도를 분석하면 일반적으로 K와 P의 민감도가 가장 크게 나타나는 것으로 분석된다(PC-그래픽을 이용한 다목적댐 홍수관리 종합 컴퓨터 시스템 개발연구, 1991). 따라서 이들 매개변수들의 변화폭이 가장 클 것으로 예상하였으나, 산정된 매개변수를 분석한 결과 R_{50} 의 최대값이 1.9 최소값이 0.3으로 가장 큰 변화폭을 보였으며, 다음으로 변화폭이 큰 것은 f_1 로 분석되었다. 이들 매개변수는 주로 포화 강우량과 1차 유출류에 관계된 매개변수들이다. 따라서 이정규와 이창해(1996) 연구에서도 제시한 바와 같이 저류함수법은 포화우량을 기준으로 유출률이 획일적으로 변화한다고 보는 점 등의 문제점을 우선적으로 해결하여 할 과제로 분석되었다. 또한 유전자알고리즘으로 8개의 홍수사상을 분석하여 저류함수모형의 매개변수를 산정하고, 그 매개변수의 평균값을 적용하여 매개변수 산정에 사용하지 않은 홍수사상 4개를 적용하여 분석한 결과, 일부 홍수사상은 상당히 향상되었으나 일부 홍수사상은 전혀 개선되지 않았다. 이는 저류함수모형은 홍수사상별로 각기 다른 매개변수를 갖고 있으며 그 변화폭이 크기 때문이며 8개의 매개변수로는 대청댐 유역의 각 소유역별 특성을 충분히 반영하지 못함을 알 수 있었다.

표 2. 유전자 알고리즘에 의한 저류함수모형의 매개변수 특성

항 목	K	P	T_1	f_1	R_{so}	K'	P'	T_1'
최 대	1.8	1.8	1.7	1.8	1.9	1.8	1.7	1.8
최 소	0.4	0.5	0.3	0.3	0.3	0.2	0.4	0.8
평 균	1.2	0.9	0.7	1.0	1.4	1.1	1.1	1.4
표준편차	0.336	0.342	0.447	0.505	0.490	0.455	0.366	0.280

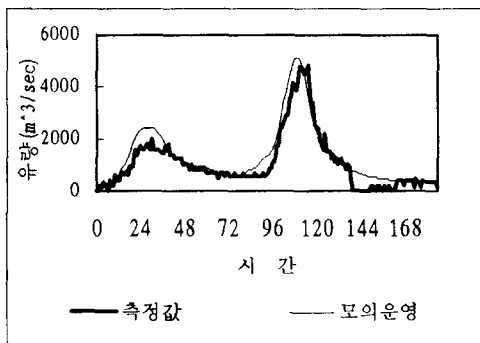


a) 1983. 7.27 홍수사상 분석결과

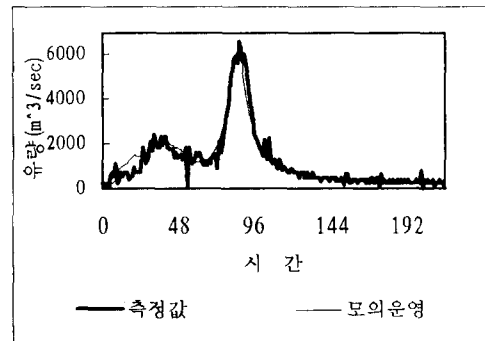


b) 1985. 7. 6 홍수사상 분석결과

그림 3. 유전자알고리즘을 적용한 유출모형 모의운영결과



a) 1985.9.15 모의운영결과



b) 1987.7.28 모의운영결과

그림 4. 유전자알고리즘에 의한 매개변수 보정후 모의운영결과

6. 결론

본 연구에서는 국내하천의 홍수에정보시스템의 운영과 다목적댐의 홍수조절관리를 위하여 사용되고 있는 수문학적 모형인 저류함수법의 매개변수 산정을 위하여 지금까지 사용된 경험공식을 바탕으로 자연계의 진화 및 도태이론에 근거한 유전자 알고리즘을 도입·적용하였다. 유전자 알고리즘을 이용한 저류함수모형의 매개변수 추정을 위하여 대청댐 유역의 과거 홍수사상을 검토하고 최적 매개변수를 산정하여 제시하였으며 그 결과를 요약하면 다음과 같다.

1. 유전자 알고리즘을 적용한 대청댐의 모의운영결과에서는 매개변수의 보정전과 비교하여 침투홍수량 및 침투홍수량 도달시간 등 모든 면에서 향상되어 저류함수모형의 최적 매개변수를 추정 하기 위한 유전자 알고리즘의 적용성을 입증할 수 있는 것으로 분석되었다.
2. 유전자 알고리즘에 의해 추정된 매개변수의 특성을 분석한 결과 R_{so} 가 최대값이 1.9 최소값이 0.3으로 가장 큰 변화폭을 보였으며, 다음으로 변화폭이 큰 것은 f_1 로 나타나 강우에 관련된

매개 변수의 보정이 가장 중요한 것으로 식되었다.

3. 저류함수모형의 매개변수 추정단계에서 이용되지 않은 4개의 홍수사상에 대하여 매개변수 특성값의 평균값을 적용하여 분석한 과에서는 경우에 따라 일부 홍수사상은 정확도가 상당히 향상되었으나 일부 홍수사상은 전혀 개선되지 않음을 보여 주었는데 저류함수모형은 홍수사상별로 각기 다른 매개변수를 가지고 있고 그 변화폭도 커서 8개의 매개변수로는 대청댐 유역의 각 소유역의 특성을 충분히 반영하지 못하는 것으로 판단되며 이에 대한 모형의 검토가 필요하다.

저류함수모형은 다목적댐의 홍수조절 예측모형으로 우리나라에서 중요한 역할을 하고 있으나 국내의 홍수수문자료의 부족을 고려하여 이제는 그 동안 축적된 홍수수문자료를 바탕으로 저류함수법의 한계를 극복하고 우리나라의 홍수특성에 맞는 자체적인 모형의 개발 및 개선이 필요할 때라고 생각된다.

참 고 문 헌

- 남궁달 (1985). "저류함수법에 의한 강우-유출모형의 변수추정." 한국수문학회지, 제18권, 제2호, pp. 175-184.
- 다목적댐 홍수유출해석 및 홍수기 저수지 운영프로그램 설명서 (1993). 연구보고서. 한국수자원공사.
- 심순보 (1992). "최적화 기법에 의한 저류함수 유출모형의 자동보정", 대한토목학회논문집, 제12권, 제3호, pp. 127-137.
- 이정규, 이창해 (1996). "저류함수법의 시변성 매개변수 조정에 퍼지이론 도입에 관한 연구." 한국수자원학회지, 제29호, 제4권, pp. 149-159.
- PC-그래픽을 이용한 다목적댐 홍수관리 종합 컴퓨터 시스템 개발연구 (1991). 연구보고서, 한국수자원공사.
- David E. Goldberg (1989). "Genetic Algorithms in Search, Optimization, and Machine Learning. Addison-wesley Publishing Company, Inc..
- Kimura, Toshimitsu (1961). "저류함수법에 의한 홍수유출 추적법." 일본건설성 토목연구소.
- Shie. Yui Liong, Weng Tat Chab, Jaya ShreeRam (1995), "Peak-Flow Forecast with Genetic Algorithm and SWMM." Journal of Hydraulic Engineering, August, pp. 613-617.
- Wang Q.J. (1991). "The Genetic Algorithm and Its Application to Calibrating Conceptual Rainfall-Runoff Models." Water Resources Research, Vol. 27, No 9, pp 2467-2482.