

유전자 알고리즘을 이용한 수위-유량 관계곡선의 보정

○노경범¹⁾, 이한민²⁾, 박성천³⁾, 이관수⁴⁾

1. 서론

어떤 수위관측단면에서의 하천수위와 그에 상응하는 하천유량을 동시에 측정하여 상당한 기간동안의 자료를 수집하면 수위와 유량간의 관계를 표시하는 검정곡선을 얻을 수 있으며 이를 수위-유량관계곡선(stage-discharge relation curve) 혹은 Rating Curve라고 한다. 이러한 수위-유량곡선의 형식으로는 영유량(零流量) 수위를 고려한 $Q = a(H + b)^n$ 지수식이 있다. 본 연구에서는 영유량 수위를 고려한 지수식을 유전자 알고리즘을 이용하여 보정하였다.

본 연구에서 실시한 유전자 알고리즘(Genetic Algorithm; GA)은 생물진화(선택도태나 돌연변이)의 원리로부터 착안된 알고리즘으로서, 확률적 탐색이나 학습 및 최적화를 위한 기법의 한가지이다.

2. 이론적 배경

2.1 유전자 알고리즘

유전자 알고리즘(Genetic Algorithm; GA)은 생물의 진화 메커니즘을 모방한 탐색 알고리즘이다. 이 유전자 알고리즘의 기본 연산자는 선택, 교배, 돌연변이의 연산자가 있으며, 알고리즘의 도식도는 그림 1과 같다.

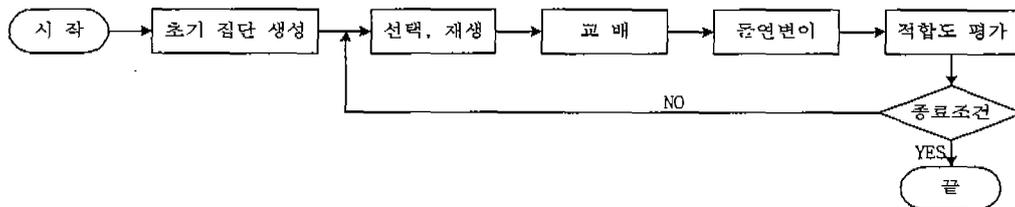


그림 1. 유전자 알고리즘 도식도

유전자 알고리즘에서 해의 후보는 유전자형(genotype)으로서 염색체(chromosome)에 일차원적으로 표현되며, 각 세대는 개체(individual)의 집합이다. 각 세대에 있어서 개체의 수는 집단의 크기(population size)라고 부른다. 이러한 유전자 알고리즘의 처리 순서는 첫 번째 단계에서 초기집단을 생성하고 두 번째 단계에서 각 개체에 대해 적합도 평가, 선택, 교차 및 돌연변이 과정을 종료조건이 만족될 때까지 반복한다. 본 연구에서는 각 개체의 염색체를 실수 값으로 표현하여 실시하였다.

- 1) 전남대학교 토목공학과 공학석사
- 2) 전남대학교 공과대학 토목공학과 박사과정
- 3) 동신대학교 토목환경공학부 환경공학과 부교수
- 4) 전남대학교 건설지구환경공학부 토목공학과 교수

2.1.1 선택(selection)

모집단은 매 세대마다 일정수의 개체를 유지하고 매 세대에서 각 개체의 적합도(fitness)를 평가하여 이에 따라 다음 세대에 생존할 개체들을 확률적으로 선택한다. 본 연구에서는 우선 초기 모집단을 무작위로 발생시켰다. 이렇게 초기에 발생된 개체들은 목적함수(식 (1))에 의해 목적함수 값이 할당되어지고, 할당된 목적함수 값에 의해 각 개체는 적합도 값을 갖는다.

$$ObjF = \sqrt{(Q_{obs} - Q_{sim})^2 / N} \quad (1)$$

여기서, Q_{obs} :실측값, Q_{sim} :모의값, N :자료의 수

본 연구에서 적용한 적합도 함수는 목적함수의 역수를 사용하였으며 식 (2)와 같다.

$$fitness = 1 / ObjF \quad (2)$$

적합도 값을 할당받은 개체는 각각 선택확률에 의해 선택과정을 거치후 교배과정을 실행한다.

2.1.2 교배(crossover)

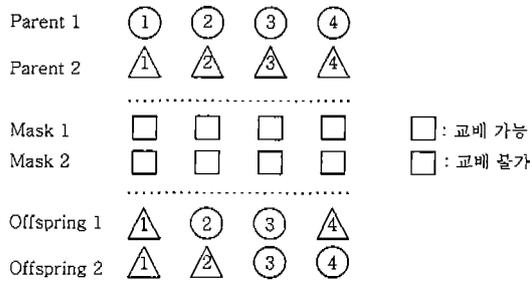


그림 2. Discrete Recombination

교차는 두 부모의 염색체를 조합하여 자식의 염색체를 만드는 조작이다. 본 연구에서 실시한 교배 방법은 개체들간에 변수의 값을 교환하는 불연속 재조합(discrete recombination)법을 이용하였다.

그림 2에서 보는 바와 같이 각각의 부모는 마스크(mask)를 통하여 자손을 생성하게 된다. 이렇게 생성된 자손은 다음으로 돌연변이 과정을 거친다.

2.1.3 돌연변이(mutation)

돌연변이는 유전자를 일정한 확률로 변화시키는 조작이다. 염색체가 1차원 문자열로 표현되어 있는 경우 그 중에서 의미있는 패턴이 발생한다. 이와 같은 패턴을 스키마타(schemata)라고 부르며 만약 돌연변이율이 너무 크면 스키마타가 전부 파괴되어 임의탐색(random search)으로 변할 확률이 높다. 돌연변이가 없는 경우에는 초기 유전자 조합이외의 공간을 탐색할 수 없으며, 결국 찾고자 하는 해의 질에도 한계가 드러난다. 본 연구에서는 돌연변이율을 0.1로 하여 수행하였다.

이 세 가지 유전 연산과정을 수행한 후 목적함수를 평가하며 만일, 종료조건에 목적함수 값이 도달했을 경우 수행이 종료되지만, 그렇지 않을 경우 선택, 교배, 돌연변이의 유전 연산과정을 반복하게 된다.

2.2 기존 수위-유량 관계곡선식 개발 방법

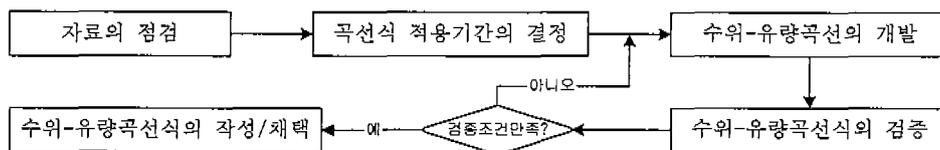


그림 3. 수위 유량곡선 작성도

관측 수위(H) 및 유량(Q)과의 관계를 나타낸 곡선이 수위-유량곡선(H-Q)이며, 이 수위-유량 곡선을 사용하여 연속적으로 측정되는 수위에서 유량을 산출한다.

수위-유량곡선을 작성할 때의 순서는 그림 3과 같으며, 곡선식 형식으로는 영유량 수위를 고려한 지수식을 사용하며 식 (3)과 같다.

$$Q = a(H + b)^n \quad (3)$$

여기서, Q는 유량(m^3/sec), a, b 및 n은 상수이며 b는 영유량 수위

n은 하도 단면 형상 등에 관계하며 직사각형 수로인 경우 3/2, 포물선 형태인 경우 2, 삼각형이나 반원형인 경우 5/2의 값을 가진다. 또한 비교적 폭이 넓은 하천인 경우 1.3~1.8 범위를 가지며 비교적 깊고 좁은 단면인 경우 보통 2보다 크고 종종 3을 초과할 수도 있다. 자연하천에서 단면이 포물선인 경우가 많으며, 이때 n 값은 약 2가 되므로 곡선식형은 2차식이 된다.

수위-유량곡선식의 영유량 수위인 b값을 결정하는 방법은 시행착오법, 직접법, 도해법 및 컴퓨터 프로그램을 이용한 방법 등이 있지만, 모든 추정치는 현장조사를 통해 확정해야 한다.

시행착오법은 시행착오에 의해 수위(H)에 적용하기 위한 최적의 b값은 수위와 유량의 관계가 직선으로 나타난 경우이다.

회귀 분석법(Regression analysis)은 수위-유량곡선식을 계산함에 있어서, 가정된 곡선식 결정의 이론적 근거와 신뢰성을 확인하기 위하여 회귀분석 및 상관분석을 실시하여야 한다. 주어진 자료군을 이용하여 표준개산오차, 상관계수 및 분산분석표를 작성하여 회귀성 검정 및 표준개산오차로부터 적합성 검정을 실시하여야 한다.

2.3 유전자 알고리즘에 의한 수위-유량 곡선의 보정

본 연구에서는 수위-유량곡선의 지수식 형태인 식 (3)에서 구하고자 하는 매개변수 a, b 및 n을 유전자 알고리즘에 적용하기 위해 염색체로 표현하였다. 이렇게 염색체로 표현된 각 매개변수는 유전자 알고리즘에 적용하여 최적화된 매개변수 값을 얻었다.

본 연구에서 실시한 유전자 알고리즘의 수행조건은 표 1과 같으며, 하나의 개체집단을 생성하여 잠재 해의 최적화를 수행하는 SGA(simple genetic algorithm)과 달리 다중 부 개체 집단을 생성하여 수행하는 방법인 MPGA(multiple populations genetic algorithm)를 이용하여 수위-유량관계 곡선식의 보정을 실시하였다.

표 1. 유전자 알고리즘 수행 조건

GGAP	P_c	P_m	MAXGEN	INSR	SUBPOP	MIGR	MIGGEN	NIND
0.8	1	0.1	800	0.9	6	0.2	20	25

여기서, GGAP:세대차이, P_c :교배율, P_m :돌연변이율, MAXGEN:최대 세대수, INSR:재삽입율,

SUSPOP:부개체집단수, MIGR:이주율, MIGGEN:이주 주기, NIND:개체의 수

유전자 알고리즘의 적합도 평가를 위한 목적함수로는 식 (1)을 이용하였으며, 목적함수 값에 의해 각 개체에 대한 적합도 값을 할당하였다.

3. 대상지점 및 자료

본 연구에서 실시한 연구 대상지점 보성강 수계의 검백지점, 영산강 수계의 오례지점과 섬진강 수계의 오수지점을 연구 대상지점으로 선정하였고, 이용된 자료는 1997~1999년에 실측된 자료를 이용하여 기 개발된 곡선식에 대한 보정을 실시하였다.

4. 분석결과

본 연구에서 실시한 유전자 알고리즘에 의한 수위-유량 관계곡선식의 보정 결과를 표 2~표 4 까지 비교하였다.

표 2는 검백지점에서 '97~'99년에 기존 방법에 의해 개발된 수위-유량관계곡선식과 유전자 알고리즘에 의해 구해진 값을 나타내고 있으며, 그림 4는 '98년 자료에 대해 기존식과 유전자 알고리즘에 의해 개발된 식에 의해 구해진 값을 나타내고 있다. 그 결과 전체적으로 유전자 알고리즘에 의해 산출된 값이 우수한 결과를 나타냈으며, 그림 4에 도시된 '98년 값을 비교해보면 기존 식에 의한 값은 수위 5.46(m)의 유출량에서 약 135(CMS)정도 차이를 보였으나 유전자 알고리즘에 의해 산출된 값은 약 56(CMS)정도의 차이를 나타냈다. 또한, 유전자 알고리즘에 의해 개발된 상관계수 0.995는 기존 개발식의 0.988보다 우수한 값을 나타냈다.

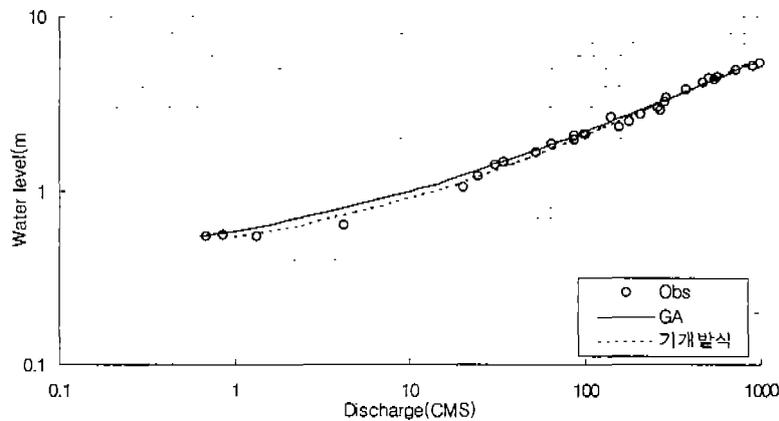


그림 4. 수위-유량 관계 곡선(검백, 98년)

표 2. 수위-유량 관계곡선식 비교(검백)

분석 자료	수위범위	수위-유량곡선식	표준개산 오차	상관계수	비 고
'97	0.48 ≤ h ≤ 4.41	$Q = 32.25153(h - 0.31677)^{2.21569}$	49.07	0.988	기 개발식
		$Q = 61.97619(h - 0.4799)^{1.596143}$	15.19	0.995	G.A
'98	0.55 ≤ h ≤ 5.46	$Q = 36.39897(h - 0.39321)^{1.94404}$	37.58	0.988	기 개발식
		$Q = 24.67975(h - 0.27771)^{2.20848}$	27.58	0.995	G.A
'99	0.50 ≤ h ≤ 4.00	$Q = 19.576(h - 0.260)^{2.563}$	36.99	0.973	기 개발식
		$Q = 50.50424(h - 0.4999)^{1.86744}$	27.72	0.983	G.A

오레지점에 대한 기 개발 수위-유량 관계곡선식과 유전자 알고리즘에 의해 개발된 식을 표 3에 나타냈으며, 그림 5는 '98년 측정된 자료에 대해 기 개발식과 유전자 알고리즘에 의해 개발된 식을 도시했다. 그림 5에서 보는 바와 같이 오레지점에 대한 '98년 수위-유량관계 곡선은 기 개발식과 비슷한 값을 나타내고는 있지만 표준개산오차와 상관계수를 비교해보면 유전자 알고리즘에 의한 값이 0.990으로 0.981보다 우수한 결과를 보였다.

오레지점에 대한 수위-유량곡선의 보정 결과 전체적으로 GA에 의해 계산된 값이 우수한 값을 보였으며, 특히 '97년에 자료에 대한 표준개산오차 및 상관계수 값이 크게 향상되었다.

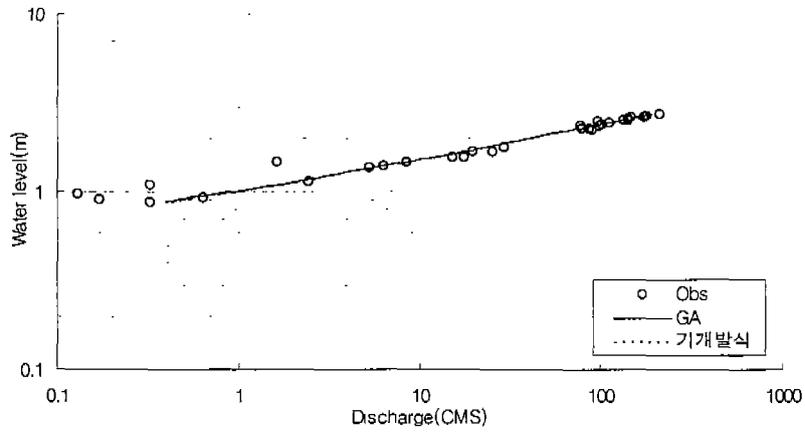


그림 5. 수위-유량 관계 곡선(오례, 98년)

표 3. 수위-유량 관계곡선식 비교(오례)

분석 자료	수위범위	수위-유량곡선식	표준개산 오차	상관계수	비 고
'97	0.90 ≤ H ≤ 3.47 (H=h+1)	$Q = 22.93580(H - 0.97907)^{2.69940}$	49.07	0.966	기 개발식
		$Q = 42.25775(H - 0.1199)^{1.593196}$	5.50	0.993	G.A
'98	0.88 ≤ H ≤ 2.73 (H=h+1)	$Q = 4.06545(H - 0.28624)^{4.34815}$	10.24	0.981	기 개발식
		$Q = 4.10411(H - 0.26642)^{4.26599}$	9.17	0.990	G.A
'99	0.18 ≤ h ≤ 1.54	$Q = 42.145(h - 0.011)^{1.9594}$	4.38	0.993	기개발식
		$Q = 21.51569(h + 0.230672)^{2.798092}$	3.12	0.996	G.A

오수지점에 대한 기 개발된 수위-유량관계곡선과 유전자 알고리즘에 의해 구해진 곡선을 그림 6에 도시하였으며, 표 4에서 '97~'99년까지 기존에 개발되어 이용되고 있는 곡선식과 유전자 알고리즘에 의해 개발된 식에 대한 검증을 실시하였다.

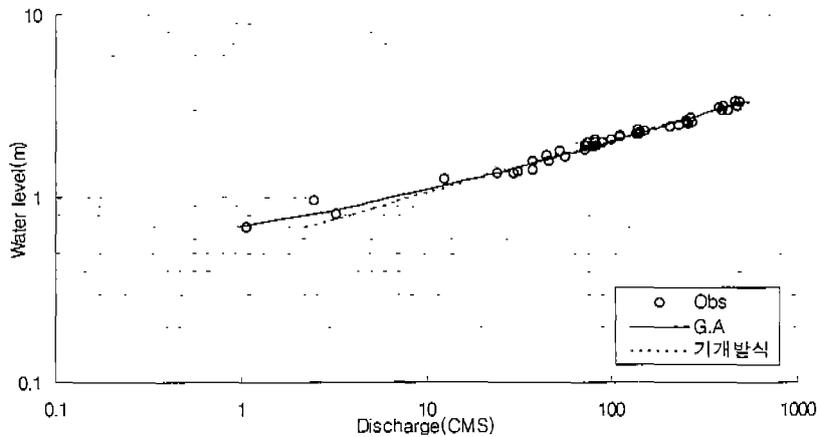


그림 6. 수위-유량 관계 곡선(오수, 99년)

표 4. 수위-유량 관계곡선식 비교(오수)

분석 자료	수위범위	수위-유량곡선식	표준개산 오차	상관계수	비 고
'97	0.94 ≤ H ≤ 3.94	$Q = 21.10532(H - 0.44026)^{3.06685}$	40.27	0.988	기 개발식
	(H=h+1)	$Q = 70.40872(H - 0.84408)^{2.239334}$	25.73	0.995	G.A
'98	0.87 ≤ H ≤ 4.13	$Q = 11.57331(H - 0.23325)^{3.22133}$	30.11	0.987	기 개발식
	(H=h+1)	$Q = 38.50348(H - 0.66612)^{2.53820}$	23.37	0.996	G.A
'99	0.86 ≤ H ≤ 3.37	$Q = 8.512(H - 0.015)^{3.449}$	27.99	0.981	기 개발식
	(H=h+1)	$Q = 26.310758(H - 0.393478)^{2.739925}$	21.25	0.989	G.A

그 결과 유전자 알고리즘에 의해 개발된 식이 기존 식에 비해 표준개산오차와 상관계수에서 전체적으로 우수한 값을 나타냈다. 이는 수위-유량곡선의 각 변수가 유전자 알고리즘에 의해 최적화된 변수 값을 찾은 것으로 판단된다.

5. 결 론

본 연구에서 실시한 유전자 알고리즘에 의한 수위-유량곡선의 보정 결과 기존 방법으로 개발된 수위-유량 관계곡선보다 모든 지점에서 우수한 결과를 보였다. 이것은 기존의 지수식 수위-유량 관계곡선 개발 방법에서 충분히 최적화되지 않은 변수들의 값이 유전자 알고리즘을 적용했을 경우 최적화된 변수값을 구한 것으로 판단된다. 또한, 수위관측소에서 기록된 수위를 유전자 알고리즘에 의해 최적화된 수위-유량관계식을 이용 유량으로 환산할 경우 부정확한 수위-유량관계식에 의해 발생할 수 있는 실제 유량에 대한 오차를 최소화 할 수 있으며 신뢰성있는 유량자료를 산출 할 것으로 판단된다.

6. 참고문헌

1. 기타노 히로아키,(1993), GA의 기초이론, 공학응용 및 인공생명 유전자 알고리즘, 대청 컴퓨터 월드
2. 김여근(1999) 등, 메타휴리스틱, 영지문화사
3. 노경범, "유전자 알고리즘과 신경망의 결합에 의한 유출량 예측", 전남대학교 공학석사학위논문, 2001
4. 한국수자원공사, "수자원기초조사편람"
5. 한국수자원공사, "97, 98 주암댐 유량측정보고서", 1997,1998
6. 한국수자원공사, "주암댐 하천 유량측정 등 수문기초조사 보고서", 1999
7. Chipperfield, A. et al., "Genetic Algorithm Toolbox for use with 5. Matlab. Version1.2", Department of Automatic Control and Systems Engineering, University of Sheffield, 1994
8. Michalewicz(1994), Genetic Algorithms+Data Structures=Evloution Programs, Springer-Verlag