패턴 인식기법을 이용한 유출모형의 매개변수 최적화

정창삼*, 허준행**

Abstract

일반적으로 강우-유출모형은 lumped model과 distributed model로 크게 구분될 수 있으며, 우리나라에서는 이중 비교적 부족한 자료를 이용하여도 개략적 모의가 가능한 전자를 널리 사용하고 있다. 본 연구에서는 이러한 모형들의 매개변수를 보정하는 방법에 관해 연구하였다. 일반적으로 모형의 보정 방법에는 크게 시행오차에 의한 수동보정(manual calibration) 방법과 최적화 기법에 의한 자동보정(automatic calibration) 방법으로 나눌 수 있다. 수동보정 방법은 모형 수행결과를 수문곡선의 시각적 비교에 의해 관측치와 비교하여 모형 운영자의 주관적인 판단하에 조정하는 기법이며, 자동보정 방법은 최적화 기법을 이용8하여 특정한 산정기준(estimation criteria)을 최대 또는 최소화시켜 모형의 매개변수를 결정하는 방법이다. 이러한 최적화기법은 일반적으로 직접탐색법과 경사법으로 구분할 수 있다. 경사법은 수렴속도가 빠르지만 편미분에 의해 방향을 찾아가는 방법으로 편도함수가 필요하므로 수문모형에는 적용하기가 힘들므로 적합하지 않다. 그러나, 보다 많은 컴퓨터 수행시간을 필요로 하는 직접탐색법의 경우 수렴속도는 느리지만, 편도함수를 필요치 않으므로 수문모형의 최적화 기법으로 적합하다고 할 수 있다. 직접탐색법에는 simplex-search 법, 패턴인식(pattern-search)법, rotating-direction 법, brent 법 등이 있으며, 본 연구에서는 직접탐색법의 일종인 패턴인식(pattern -search)법을 이용하여 매개변수 최적화 과정을 모의하였다. 이러한 매개변수 보정모형을 구성한 후 이를 가장 보편적으로 사용되고 있는 유출모형인 각종 단위도법들을 결합하는 모형을 구성하였다. 또한 구성된 모형을 시범유역에 적용하여 나온 결과를 HEC-1에서 적용되고 있는 단일변량 증감법과 같은 최적화 기법을 이용한 결과와 비교・분석을 실시하였다.

본 모형을 활용하여 강우-유출 모형의 매개변수를 지속적으로 산정하고 일반화할 경우 임의의 유역의 수문기상학적 특성에 부합한 매개변수를 정량화 시킬 수 있었다.

* 연세대학교 대학원 토목공학과 박사과정

** 연세대학교 사회환경시스템공학부 토목공학과 부교수

1. 개요

HEC-1과 같은 수문모형에서 매개변수를 정하기 위한 기법으로는, 경험적인 기법과 해석적인 기법, 그리고 최적화 기법으로 나눌 수 있다. 해석적인 기법의 경우 그 해를 수학적으로 풀어낼 수 있는 경우와 그렇지 못한 경우가 있어 본 연구에서는 고려하지 않았다. 흔히 널리 사용되고 있는 경험적 방법은 시행오차에 의한 수동보정(manual calibration) 방법으로 모형 수행결과를 특정한 방법, 즉 수문곡선의 시각적 비교에 의해 관측치와 비교하여 모형 운영자의 주관적인 판단하에 매개변수를 조정하는 방법으로, 이를 위해서는 모형 및 유역의 특성을 완전히 이해함은 물론 모형보정에 대한 숙련된 경험과 감각을 갖춘 수문기술자가 필요하다. 이에 반해, 최적화기법을 이용한 매개변수 결정방법은 수학적 최적화 기법을 이용하여 목적함수를 최소 또는 최대화시켜 모형의 매개변수를 보정하는 방법이다. 이러한 기법은 경험이 풍부한 전문가가 아니더라도 가능하며, 컴퓨터의 발달로 인해 활용범위가 점차 확대되고 있는 추세이다. 이와 같은 자동보정 방법은 최적화 기법을 이용하여 특정한 산정기준(estimation criteria)을 최대 또는 최소화시켜 모형의 매개변수를 결정하는 방법이다. 자동보정 방법은 더 많은 컴퓨터 수행시간이 필요하지만 전체적인 보정과정에 소요되는 시간을 상당히 단축시킬 수 있다. 더욱이 근래의 컴퓨터의 발달로 보정에 소요되는 시간에 대한 컴퓨터 수행시간이 차지하는 비율이 적어짐으로 인해 자동보정 방법의 효용성이 증가되고 있다.

2. 최적화 기법

본 연구의 대상이 되는 매개변수 자동보정 기법의 하나인 최적화기법은 직접탐색법과 경사법으로 크게 구분할 수 있다. 경사법은 수렴속도가 빠르지만 편미분에 의해 방향을 찾아가는 방법으로 편도함수가 필요하므로 수문모형에는 적용하기가 힘들므로 적합하지 않다. 이러한 경사법(gradient based method)의 경우에는 1차 도함수(first derivative)형과 2차 도함수(second derivative)형으로 나눌 수 있다. 그러나, 보다 많은 컴퓨터 수행시간을 필요로 하는 직접탐색법의 경우 수렴속도는 느리지만, 편도함수를 필요치 않으므로 수문모형의 최적화 기법으로 적합하다고 할 수 있다. 직접탐색법에는 simplex-search 법, pattern-search 법, rotating-direction 법, brent 법 등이 있다.

이러한 직접탐색법은 매개변수에 대한 산정기준의 미분계수를 이용하지 않고 체계적인 방법에 의해 산정기준을 계산으로 방법으로, 많이 사용되는 기법들은 simplex search 법 (Nelder 와 Mead, 1965), pattern search 법 (Hooke and Jeeves, 1961), 그리고 rotating direction 법 (rosenbrock, 1960) 등이 있다.

일반적으로 비선형 최적화 문제에 대한 여러 문장에서는 1차 미분법 또는 2차 미분법이 직접탐색법보다 우수하다고 평하고 있으나 수문모형에 대해서는 그렇지가 않다. 그 이유는 여러 문헌의 기법 비교는 항상 정해가 존재하는 문제를 가지고 시험하였기 때문에 계산량만이 적은 방법이 우수한 것으로 판단하였기 때문이다. 그러나 수문 현상과 관련된 최적화 문제에서는 매개변수의 정확한 값이 종종 산출되지 않고 있기 때문에 해에 대한 안정성(Robustness), 즉 여러가지 조건하에서도 정해를 구할 수 있는 능력을 고려하여야 한다. 본

연구에서는 해의 안정성이 타 방법보다 우수하며 현재 NWS에서 NWS-RFS의 일부인 Soil Moisture Accounting 모형의 보정기법으로 사용되고 있는 Pattern Search 법을 HEC-1모형의 매개변수 자동보정을 위한 최적화 기법으로 선정하였다.

2.1 Pattern-Search 법

P-S 법은 크게 탐색 이동(Explorator Move)과 표본 이동(Pattern Move)의 두 과정으로 나누어지며 이 두 과정의 반복시행에 의해 최적화를 찾아가게 된다. 탐색 이동에서는 임의의 점에서 최적화 선정기준(Estimation Criterion)인 목적함수의 값이 개선될 수 있는 탐색경로의 방향을 결정하며 표본 이동에서는 탐색 이동에서 결정된 방향으로 이동 구간을 증가시켜가며 새로운 가능 해로 이동한다.

그림 1. 2차원 최소화 문제의 Pattern-Search 과정

2.1.1 목적함수

최적화 기법의 적용을 위해서는 최적화 산정기준인 목적함수가 정의되어야 하며 일반적으로 유출모형의 목적함수는 식(1)로 표시할 수 있다.

$$f(E) = f[Q - R(\Pi, \Theta)] \tag{1}$$

여기서 $Q = [q_1, q_2, \ldots, q_n]$ 및 R은 n시간 동안의 관측치와 모형에 의한 유출 산정치를 각각 나타내며 Π 는 모형의 입력변수, Θ는 모형의 매개변수를 나타낸다. 위식에서 오차기준 E 와 함수 f 를 여러 가지 방법으로 정의할 수 있다. 위식에서 모의발생값 R은 매개변수 Θ에 대해 양해적인 형태(explicit form)으로 나타낼 수도 있고 그렇지 않을 수도 있으나 대부분의 수문모형에서는 후자의 경우에 속한다.

본 연구에서 개발된 모형에서 사용된 목적함수는 크게 5가지로 그 대표적인 알고리즘은 실측유량과 계산유량과의 편차을 어떻게 최소화하는가에 따라 다음의 5가지 방법을 사용자가 선택적으로 사용할 수 있게 하였다.

① 일반적으로 가장 많이 사용되는 목적함수는 관측값과 모의발생값의 시간별 편차의 제곱의 합(SSR : sum of sqaure of residual)을 최소화하는 것이다. 식(2)는 관측값과 계산값의 편차의 제곱의 합의 최소화하는 목적함수를 나타내고 있다.

$$\min\ f(E) = \sum_{i=1}^{n} (q_i - r_i)^2 \tag{2}$$

여기서 q_i와 r_i는 시간 i에서 관측값과 모의발생값(계산값)을 각각 나타낸다. 위식은 유량의 제곱의 차원을 가지므로 관측 및 계산된 유량의 첨두값(peak flow)에 중점을 두고 있다.

② 두 번째 방법은 관측값과 모의발생값의 시간별 편차의 합을 최소화하는 것이다.

$$\min\ f(E) = \sum_{i=1}^{n} ABS(q_i - r_i) \quad (3)$$

여기서 q_i와 r_i는 시간 i에서 관측값과 모의발생값(계산값)을 각각 나타낸다. 위식은 식(4)는 관측값과 계산값의 편차에 대하여 동등한 비중을 주고 있다.

③ 세 번째의 방법은 관측값과 모의발생값의 시간별 편차를 관측값으로 나눈 값의 제곱의 합을 최소화하는 것이다. 식(4)는 관측값과 계산값의 편차를 관측값으로 나눈값의 제곱의 합의 최소화하는 목적함수를 나타내고 있다.

$$\min\ f(E) = \sum_{i=1}^{n} \left(\frac{q_i - r_i}{q_i}\right)^2 \quad (4)$$

여기서 q_i와 r_i는 시간 i에서 관측값과 모의발생값(계산값)을 각각 나타낸다. 위식은 편차의 크기에 따라 그 비중이 좌우되므로, 편차가 큰 경우 그 비중이 커지며 편차가 작은 경우에 그 비중이 작아지므로, 관측값에 더욱 중점을 둔 식으로 관측값에 가까이 갈수록 더욱 빠른 수렴을 하는 경향이 있다.

④ 식(5)는 관측값과 계산값에 대하여 이들의 세제곱근 값의 편차의 제곱을 한 다음 이 값을 최소화 하는 값을 목적함수로 사용하였다.

$$\min\ f(E) = \sum_{i=1}^{n} \left(q_i^{\frac{1}{3}} - r_i^{\frac{1}{3}}\right)^2 \quad (5)$$

여기서 q_i와 r_i는 시간 i에서 관측값과 모의발생값(계산값)을 각각 나타낸다. 위식은 위 ③의 목적함수에서 편차의 크기에 따라 그 비중이 좌우되므로 이를 개선하여 편차가 큰 경우에 차지하는 가중치를 줄이고 편차가 작은 경우에 차지하는 가중치를 높이는 방법을 사용하였다. 이 방법은 편차에 의한 상호간섭이 없으며, 정규분포를 이루고 있거나 아니면, Log-likelihood의 분포를 가지는 함수의 경우 매개변수의 값이 대략 이차식의 형태의 최대값이므로 대략 타원형의 contour를 가지게 된다.

⑤ 다섯 번째 식(6)은 위 ④식과 비슷한 형태를 지니고 있지만, 관측값과 계산값에 대하여 이들의 제곱근을 취하여 이들 값의 편차의 제곱의 합이 최소화 되도록 하였다.

$$\min\ f(E) = \sum_{i=1}^{n} \left(q_i^{\frac{1}{2}} - r_i^{\frac{1}{2}}\right)^2 \quad (6)$$

여기서 q_i와 r_i는 시간 i에서 관측값과 모의발생값(계산값)을 각각 나타낸다. 위식은 위의 4번째 식의 형태와 비슷하지만, 제곱근을 사용하는 것에서 차이가 난다.

2.2 모형의 구성

본 연구에서는 대표적인 단위도 분석법 가운데, Clark법, SCS법을 이용하여 유출모형을 구성하였으며, 그림 2에 나타난 개요도를 바탕으로 Fortran코드를 이용하여 계산엔진을 작성하였으며, Delphi 6.0을 이용하여 그림 3과 같이 GUI를 구성하였다.

그림 2. 모형의 개요

그림 3. 모형의 구성 및 적용

3. 모형의 적용 및 결과

본 연구를 통해 개발된 모형을 이용하여 서울시에 위치한 불광천 유역의 2001년 7월14일~7월18일 사상에 대해 적용을 실시하여 보았다. 우선 기존의 강우 및 관측 유량자료를 이용하여 HEC-1모형을 이용하여 매개변수를 최적화하여 유출 수문곡선을 작성하였으며, 본 패턴인식기법을 이용한 모형을 이용하여 매개변수를 최적화 한 후 수문곡선을 작성한 후 이를 비교하여 보았다. 모형을 비교 검토한 결과, 일반적으로 HEC-1모형에서 사용되는 최적화기법인 단일변량 증감법과 Nelder-Mead법의 단점인 첨두발생시간의 지체현상이 본 모형에서는 잘 극복됨을 알 수 있었다. 하지만, 모형의 단점인 초기값의 정의에 따라 수렴이 되지 않는 경우가 종종 발생하였으며, 목적함수에 따라 각각 상이한 사상을 최적화하는 상이한 결과가 발생하였다.

그림 4. 모형의 적용 결과 (불광천, 2001년 7월14일~7월18일 사상)

4. 결론

5가지 목적함수를 적용한 패턴인식기법을 이용하여 Lumped 유출모형에 적용할 경우 다음과 같은 결론을 얻을 수 있었다.

첫째, HEC-1 모형의 최적화 기법인 단일변량 증감법과 Nelder-Mead법을 이용하여 제곱오차, 잔차 자승 합, 절대 잔차 합, 첨두유량 백분율 오차 등을 이용한 매개변수 최적화보다 실제사상을 더 잘 모의할 수 있는 최적화된 매개변수를 얻을 수 있었다.

둘째, 개발된 모형의 목적함수 별로 각각의 사상에 대한 매개변수를 보다 정확히 모의하여 최적화할 수 있었으며, 그 패턴에 대해서는 보다 많은 사상에 대한 검정이 필요하다.

셋째, GUI모형의 구성을 통해 보다 용이하게 모형을 구성하고 모의할 수 있었다.

참고문헌

1. Beston, roger P. and Ralph F. Green, Rearch Paper No. 6, "DIFCOR - A Program to Solve Nonlinear Equation, "Tennessee Valley Authority, Knoxville, Tenn., June 1967.
2. Beston, roger P. and Ralph F. Green, "Analytically Derived Unit Graph and Runoff, "Journal of Hydraulics Division, Proceeding, ASCE, Vol. 94 No. HY6, Proc. Paper 6256, Nov. 1968.
3. Hooke, R. and T. A. Jeeves, "Direct Search Solutions of Numerical and Statistical Problems." J. Assoc. Comput. Mach., 8(2), p212-229, 1961.
4. Nelder, J.A. and R. Mead, "A Simplex Method for Functional Minimization." Comput. J., 7(4), p308-313, 1965.
5. Rosenbrock, H. H., "An Automatic Method of Finding the Greatest of Least Values of a Function." Comput. J., p175-184, 1960.

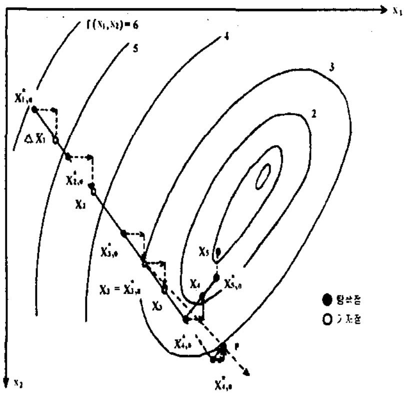

그림 1. 2차원 최소화 문제의 Pattern-Search 과정

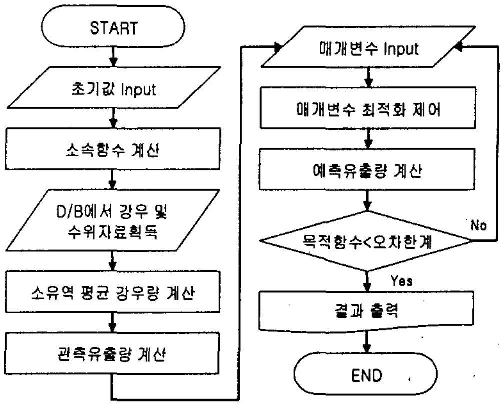

그림 2. 모형의 개8요

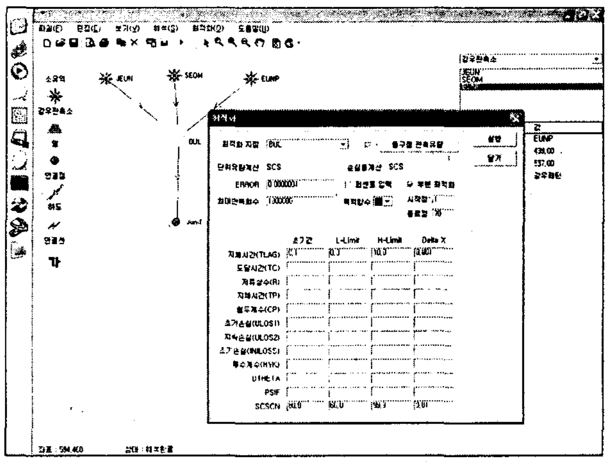

그림 3. 모형의 구성 및 적용

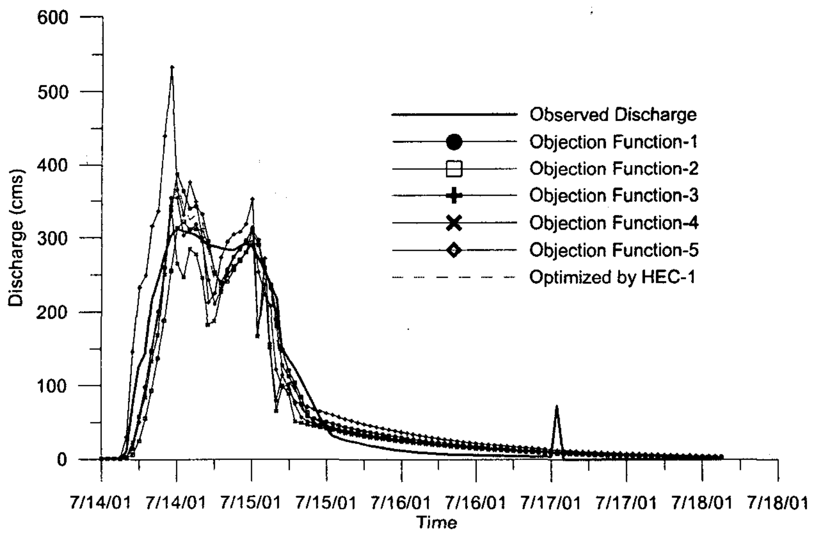

그림 4. 모형의 적용 결과 (불광천, 2001년 7월14일~7월18일 사상)