

다양한 목적 함수와 최적화 방법을 달리한 SIMHYD와 TANK 모형의 적용성 연구

The Applicability Study of SYMHYD and TANK Model Using Different Type of Objective Functions and Optimization Methods

성 윤 경* / 김 상 현** / 김 현 준*** / 김 남 원****

Sung, Yun Kyung / Kim, Sang Hyun / Kim, Hyun Jun / Kim, Nam Won

Abstract

SIMHYD and TANK model are used to predict time series of daily rainfall-runoff of Soyang Dam and Youngcheon Dam watershed. The performances of SIMHYD model with 7 parameters and TANK model with 17 parameters are compared. Three optimization methods (Genetic algorithm, Pattern search multi-start and Shuffled Complex Evolution algorithm) were applied to study-areas with 3 different types of objective functions. Efficiency of TANK model is higher than that of SIMHYD. Among different types of objective function, Nash-sutcliffe coefficient is found to be the most appropriate objective function to evaluate applicability of model.

Keywords : Rainfall-Runoff Model, SIMHYD model, TANK model, Optimization method, Objective function

요 지

일 단위 강우-유출 모형인 SIMHYD와 TANK를 소양강댐과 영천댐 유역에 적용하여 유출을 예측하였다. 7개의 매개변수를 가진 SIMHYD와 17개의 변수를 가진 TANK 모형을 국내 유역에 적용하여 모형의 적용성을 비교 평가하였다. 두 모형에 세 가지 목적함수를 달리하여 세 가지의 최적화 방법(유전자 알고리즘, Pattern Search Multi-Start, Shuffled Complex Evolution Algorithm)을 적용하여 모형과 목적함수에 따른 관측 유출량에 대한 모의 유출량의 모의 효율을 비교하였다. TANK 모형의 모의 효율이 SIMHYD 모형의 모의 효율에 비해 높게 나타났다. 목적함수를 달리할 경우는 무차원 함수인 Nash-Sutcliffe 계수를 비교하는 것이 모형의 적용성을 평가하는데 적합한 것으로 평가되었다.

핵심용어 : 강우-유출모형, TANK 모형, SIMHYD 모형, 최적화 기법, 목적함수

* 부산대학교 대학원 환경공학과 석사과정

Master student, Dept. of Environmental Engineering, Pusan National University, Pusan, 609-753
(E-mail: tj0102@hotmail.com)

** 부산대학교 공과대학 환경공학과 부교수

Associate Professor, Dept. of Environmental Engineering, Pusan National University, Pusan, 609-753
(E-mail: kimsangh@pusan.ac.kr)

*** 한국건설기술연구원 수자원 환경부 선임연구원

Senior Researcher, KICT
(E-mail: hjkim@kict.re.kr)

**** 한국건설기술연구원 수자원 환경부 수석연구원

Research Fellow, KICT
(E-mail: nwkim@kict.re.kr)

1. 서론

유역의 유출 현상을 설명하기 위해서는 그 유역의 특성에 따른 모형을 적용하는 것이 중요하다. 동일 유역에 대해서도 강우-유출 모형의 특성에 따라 모의된 유출과 모의 효율의 차이가 발생한다. 그러므로 여러 강우-유출 모형의 적용을 통하여 대상 유역에 적합한 모형을 선정하는 것이 중요하다. 모형의 적용 시에는 대상유역에 적합한 매개변수를 선정하는 것이 신뢰성 있는 유출 모의에 중요한 요소가 된다. 매개변수를 선정하기 위해 다양한 최적화 방법을 적용하여 유역의 특성과 모형에 맞는 최적화 방법을 선정하는 것이 중요하다. 탱크모형과 같은 매개변수의 수가 많은 경우의 보정은 수동 보정 보다 자동 보정이 용이하다. 매개변수의 자동 보정의 효율은 모형의 구조나 목적함수, 최적화 기법 등에 의해 결정된다(Gan 등, 1996; Gupta 등, 1999). 자동 보정된 모형의 예측 결과는 사용된 목적함수에 따라 다른 특성을 나타내며, 모형을 보정하는 목적에 따라 적절한 목적함수를 선정해야 된다(Gupta 등, 1999).

강우-유출모형의 매개변수 자동 보정에 대한 국내 연구 중 탱크모형의 매개변수 최적화에 관한 연구로는 신 등(2001)이 유전자 알고리즘과 Powell법을 이용하여 목적함수의 변화에 따른 예측 결과를 비교한 바 있으며, 강 등(2002)은 SCE-UA 법과 Annealing-Simplex 법의 전역 최적화 기법을 탱크모형의 매개변수 자동 보정에 적용하여 지역 탐색법인 Downhill Simplex법과 목적함수에 따른 적용결과를 비교한 바 있다. 또한 강 등(2003)은 토양수분 저류 구조를 갖는 4단 탱크모형에 SCE-UA 전역 최적화 방법을 이용하여 목적함수에 따라 보정 기간을 달리하여 유출을 예측하였다. 국외 연구로는 Wang(1991)이 유전자 알고리즘을 이용하여 개념적 강우 유출 모형의 보정을 하였다. Chiew 등(1993)

는 6가지의 다른 강우-유출 모형을 호주 유역에 적용하여 각 모형을 비교한 바 있으며 Duan 등(1994)은 SCE-UA법을 유역 모형의 보정에 사용한 바가 있다. Chiew 등(1994)은 MODHYDROLOG 모형을 28개의 호주 유역에 적용한 바가 있다. 그 외 연구로 Chiew 등(2002)이 호주의 331개의 유역에 SIMHYD모형의 적용성을 평가한 바가 있다.

본 연구에서는 TANK 모형과 SIMHYD 모형을 소양강댐과 영천댐 유역에 적용함으로써 두 유역의 유출 모의에 적합한 모형을 선정하고 유전자 알고리즘과 SCE-UA 법과 Pattern Search Multi-Start법을 보정에 사용하여 모형의 적용 시 가장 적합한 보정방법을 선정하였다. 또한 세 가지의 목적함수를 각 보정방법에 적용하여 매개변수 보정에 적합한 목적함수를 선택하여 높은 보정 효율을 얻고자 하였다.

II. 적용모형 및 매개변수의 최적화

2. 적용 모형

2.1 SIMHYD

SIMHYD (simplified version of the HYDROLOG model) 모형은 HYDROLOG모형을 간략화 시킨 모델로 입력 자료로는 일 강우량(mm), 일 증발산량(mm), 일 잠재 증발산량(mm)을 필요로 한다. 모형의 흐름도는 그림 1과 같다(Chiew 등, 2002). 이 모형은 유출을 침투 초과 흐름과 중간 흐름, 불투수층 유출로 구성된다. 일일 강우는 증발에 의해 매일 비워지는 차단저류를 채운다. 그러면 이를 초과한 강우는 침투능에 따라 침투하게 된다. 이때 침투된 강우는 토양 수분 함수에 따라 내부흐름 유출로 전환된다. 나머지 강우는 기저흐름의 형태로 유출된다. SIMHYD모형의 매개변수와 권장 범위를 표 1에 나타내었다.

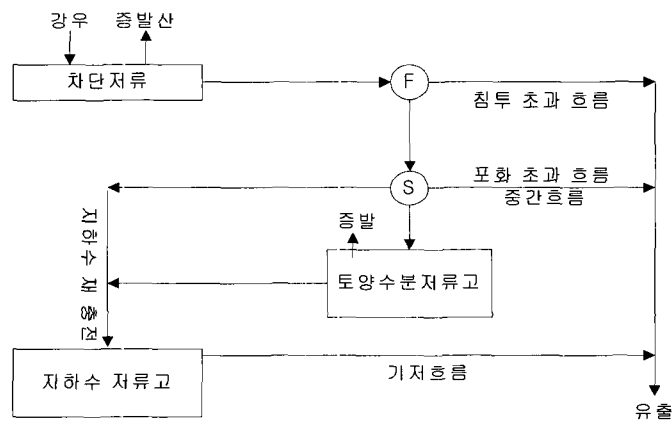


그림 1. SIMHYD 모형의 흐름도

표 1. SIMHYD 모형의 매개변수 설명

매개변수	단위	권장 범위	초기값
INSC	차단저류 (mm)	1-5	2
COEFF	최대 침투 손실 (mm)	50-300	150
SQ	침투 손실 지수	1-5	2
SMSC	토양 수분 저류능 (mm)	50-300	150
SUB	내부흐름 방정식에서의 비례 상수	0-0.5	0.2
CRAK	지하수 재충전 방정식의 비례상수	0-0.5	0.2
K	기저흐름의 선형 회귀	0.01-0.2	0.1

SIMHYD 모형에서 토양 수분 저류고의 실제 증발산량은 토양수분 습윤지수(wetness)의 선형 함수로서 추정된다(Chiew 등, 1994). 이 값은 대기로 조절되는 잠재 증발산량을 초과할 수는 없다. 실제 증발산량은 토양 수분 저류능에 대한 토양 수분 저류량과 잠재 증발산량 중 그 값이 적은 값으로 추정된다. 토양 수분의 저류능은 매개변수의 보정에 의해 결정된다. 그 관계식은 식 1에 나타내었다. SIMHYD 모형에서 사용한 잠재 증발산량을 이용한 실제 증발산량의 추정법을 나타내고 있다.

$$ET = \text{minimum of } \{10 \times SMS / SMSC, PET\} \quad (1)$$

여기서, ET 는 실제 증발산량, SMS 는 토양 수분 저류량, $SMSC$ 는 토양수분 저류능, PET 는 잠재 증발산량을 말한다.

2.2 TANK 모형

TANK 모형은 일본의 Sugawara가 1961년 처음 개

발한 모형으로 유역을 여러 개의 저류형 탱크로 가정하여 강우-유출 과정을 모형화 한 것으로 하천 유출과 관련된 개념적 수문 모형이다(Sugawara, 1995). 탱크모형에서 강우는 상부 첫번째 탱크로 유입된다. 이때 탱크의 유출공의 높이에 따라 유출이 발생한다. 유출 현상 후 저류고에 남아있던 강우는 증발산 함수에 의해 증발산이 발생하고 나머지는 두 번째 탱크로 침투된다. 첫 번째 탱크에서와 같은 매커니즘으로 마지막 탱크까지 강우가 전달되어 유출이 발생한다. 탱크모형의 입력 자료로는 일 강우량(mm), 일 유출량(mm), 일 잠재 증발산량(mm)을 필요로 한다. 모형의 흐름도는 그림2와 같다. TANK 모형에서의 첫번째 탱크에서의 유출을 표면 유출로 보고, 두 번째 탱크에서의 유출을 중간 유출이라 하고 세 번째 탱크에서의 유출을 보조기저흐름 유출이라 하고 네 번째 탱크에서의 유출을 기저흐름 유출로 가정한다. 표2는 Sugawara에 의해 권장되어지는 매개변수들의 범위와 모의에 이용된 매개변수의 범위를 나타낸 것이다.

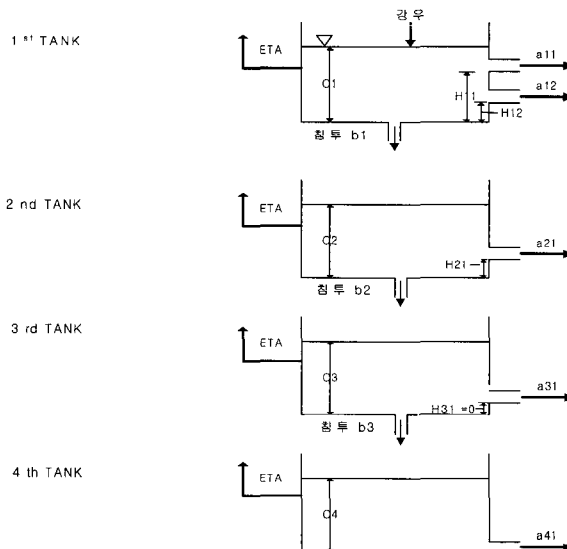


그림 2. TANK 모형의 흐름도 (배덕호 등, 2003)

표 2. Sugawara에 의해 제안된 TANK 모형 매개변수의 일반적인 범위와 이 연구의 최적화에 이용된 매개변수의 범위

매개변수	의미	Sugawara에 의해 제안된 범위	매개 변수 범위
A11	1단 상부 유출공 계수	0.1-0.5	0-0.5
A12	1단 하부 유출공 계수	0.1-0.5	0-0.3
H11	1단 상부 유출공 높이	5-10	0-300
H12	1단 하부 유출공 높이	10-60	0-300
B1	1단 침투공 계수	0.1-0.5	0.1-0.5
A2	2단 유출공 계수	0.03-0.01	0-0.1
H2	2단 유출공 높이	0-50	0-100
B2	2단 침투공 계수	0.01-0.10	0.01-0.10
A3	3단 유출공 계수	0.005-0.01	0-0.05
H3	3단 유출공 높이	0-30	0-100
B3	3단 침투공 계수	0.005-0.01	0.005-0.01
C3	3단 저류고	10-100	0-100
A4	4단 유출공 계수	0.0005-0.005	0-0.005
H4	4단 유출공 높이	0-10	0-10
C4	4단 저류고	100-500	100-1000

TANK 모형에서 토양수분의 실제 증발산량은 Beken(1979)이 제시한 방법을 이용하여 잠재 증발산량과 불포화층의 토양 수분량으로부터 일별로 추정한다.

$$Ks \cdot 1 - \exp(-\alpha S) \quad (2)$$

$$ET \leq Ks \cdot PET$$

여기서, S 는 토양 수분 저류량이고 α 는 비례상수로 최적화에 의해 보정된 상수값이다. 식 2는 TANK 모형에서 사용한 잠재 증발산량을 이용한 실제 증발산량의 추정 방법이다.

3. 최적화

다양한 모형의 산정 시 매개변수 값의 전체적인 최적조건을 찾는 것은 상당히 어려운 과제이다. 그리고 매개변수의 수가 증가 할수록 어려움은 더 커진다. TANK 모형의 매개변수는 17개, SIMHYD 모형의 변수는 7개로 매개 변수의 수가 많다. 최적화 방법으로

전역탐색 기법인 유전자 알고리즘과 SCE-UA(Shuffled Complex Evolution Algorithm)법과 지역탐색 기법인 pattern search multi-start법을 사용하였다.

3.1 유전자 알고리즘

유전자 알고리즘은 자연선택의 과정에 기초한 탐색 방법으로서 Goldberg(1989)에 의해 체계적인 방법으로 구체화되기 시작하여 최적화 문제에 적용되었고, 근래에 들어서는 함수 최적화 문제에 적용시키고자 하는 노력들이 계속하고 있는 알고리즘이다. 유전자 알고리즘은 생물학적 진화의 추상적 개념을 이용하여 연산자들이 선택, 교배, 돌연변이의 과정을 통하여 목적함수를 구성하는 변수들을 최적화하는 알고리즘이다. 유전자 알고리즘의 기본적인 처리과정은 그림3과 같다.

유전자 알고리즘은 실제의 유전자들이 가지고 있는 정보를 이진법으로 나타내는데 본 연구에서 변수들을 각 범위 내에서의 이진수를 식3과 같이 십진수로 변환하여 각 변수들의 정확도에 따라 유전자형의 길이를 결정하였다.

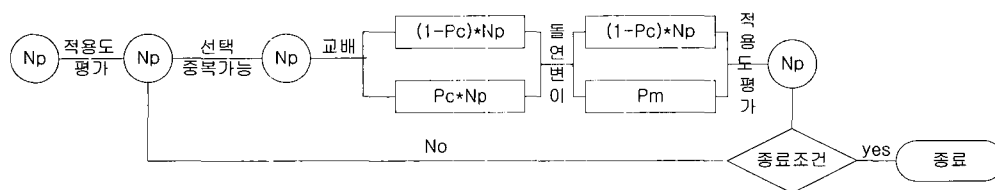


그림 3. 기본 유전자 알고리즘

$$X = P_{\min} + x \cdot \frac{P_{\max} - P_{\min}}{2^l - 1} \quad (3)$$

여기서 l 은 각 변수의 유전자형의 길이이며, P_{\min}, P_{\max} 는 각 변수의 좌우측 경계(상, 하한치), X 는 개체에서의 각 변수를 십진법으로 바꾼 값이다.

3.2 SCE-UA (Shuffled Complex Evolution Algorithm)

SCE 알고리즘은 미국 국립 기상국의 NWSRFS (National Weather Service River Forecasting System)의 보정모들 중 하나이며, 최적해를 탐색하기 위해 중요한 개념을 포함하고 있다(Duan 등, 1994). 확정론적 또는 확률론적 접근이 가능하도록 하였으며 콤플렉스의 구조적, 경쟁 진화를 하며 콤플렉스간의 혼합이 이루어진다. SCE-UA 알고리즘은 초기에 해가 존재하는 전체 매개변수 공간에서 표본을 생성하고 콤플렉스들로 분할한다. 각 콤플렉스는 Downhill Simplex 알고리즘을 사용하여 독립적으로 진화한다. 진화를 거친 콤플렉스들

은 다시 혼합되고, 새로운 콤플렉스들이 생성된다. 진화와 혼합의 과정은 수렴조건이 만족될 때까지 반복된다.

3.3 Pattern Search Multi-start 법

Pattern Search법은 직접 탐색법으로서 다중 변수 함수로 확장이 가능하고 직접 탐색법 중 효율이 비교적 높고 강력한 기능을 가졌다(이철균, 2001). Pattern Search는 동적인 방법으로 여러 차원의 탐색방향을 유지하고 생성하는 체계적인 직접 탐색법이다. 반응(목적)함수의 값이 향상을 위해 이동이 수행된다. 시험적인 이동의 성공 또는 실패는 반응함수의 변화에 대한 기본적인 지식을 제공한다. 만약 패턴의 이동이 성공적이었던 단계 크기는 최저점에 이를 때까지 점차적으로 줄어든다. 이 시점에 새로운 패턴을 정착시키는 새로운 탐색적인 이동이 일어난다(Nicolas 등, 1987). Pattern Search Multi-Start법인 경우 starting point의 수를 증가시켜 지역해에 빠질 가능성을 줄여 최적해 탐색을 가능케 하는 방법이다. 만약 반응 곡선에 과도하게 지역해가 많지 않다면 Pattern Search Multi-Start법은 보정 시간에 비해 좋은 결과를 보여줄 수 있다.

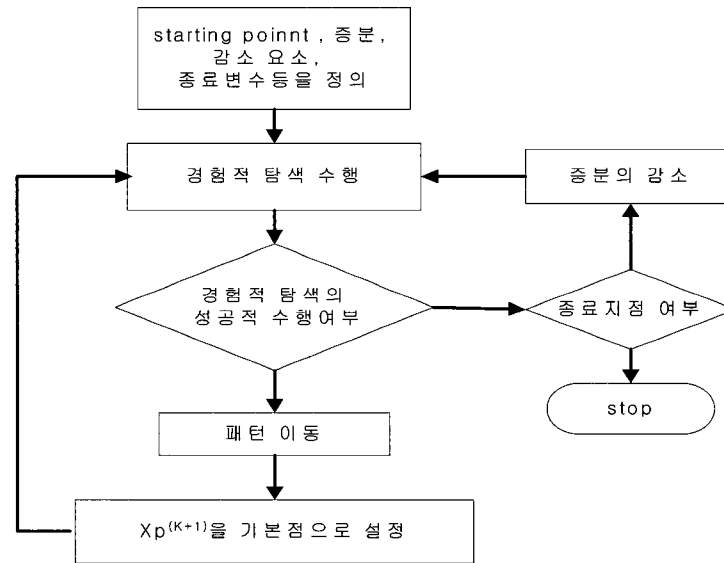


그림 4. Pattern Search 법의 흐름도

4. 목적함수

모델의 수행을 평가하기 위한 여러 가지 기준들이 있다. 그리고 선택된 기준들은 유출현상을 모의함에 있어 다양한 모델링 기법들의 차이를 설명할 수 있어야 한다. 만약 모델이 최고치의 흐름을 잘 표현할 수 있지만 적은 양의 근소한 흐름을 잘 나타낼 수 없다면 장기

간의 모의에서는 좋은 결과를 얻는 것이 가능하나 단기간의 흐름을 모의하는 데에는 부족함이 있을 수도 있다. 따라서 이 연구에서는 3가지 목적함수를 사용하였다. Nash-Sutcliffe 계수와 모의 값과 관측 값의 차의 제곱의 합과 차의 절대값의 합, 이 세가지 방법을 이용하였다. 식 4-6은 각 목적함수를 나타내고 있다. Nash-Sutcliffe 계수는 관측치와 모의치의 차이를 관측

치의 분산 정도에 대해 나타낸 값으로 관측치의 분산 정도가 커지면 상대적으로 관측치와 모의치의 차이가 적어진다. '1'에서 관측치의 분산 정도에 대한 관측치와 모의치의 차를 뺀 값으로 '1'에 가까울수록 관측치와 모의치의 차이가 적음을 나타내는 무차원 수이다.

오차의 제곱의 합은 다음과 같다.

$$OBJ1 = \sum_{i=1}^n (SIM_i - REC_i)^2 \quad (4)$$

오차의 절대값의 합은 다음과 같다.

$$OBJ2 = \sum_{i=1}^n |SIM_i - OBJ| \quad (5)$$

Nash-Sutcliffe 계수(Nash와 Sutcliffe 등, 1970)는 다음과 같다.

$$OBJ3 = 1 - \frac{\sum_{i=1}^n (EST_i - REC_i)^2}{\sum_{i=1}^n (REC_i - \overline{REC})^2} \quad (6)$$

REC 는 실측 자료 값, \overline{REC} 는 실측 자료 값의 평균, EST 는 모의된 값을 나타낸다.

III. 모형의 적용과 모의

1. 대상유역

두 가지 강우-유출 모형을 적용하기 위해 대상 유역을 선정하였다. 한강 지류인 소양강댐과 낙동강 지류인 영천댐 두 유역을 선택하였다. 그림 5와 그림 6은 대상 유역을 나타내고 있는 단위지도이다. 소양강댐은 유역면적은 약 2,703km²로서 강원도 인제군 외 4개 시·군(고성, 양구, 홍천, 춘천)에 걸쳐있다. 소양강댐은 연평균

강수량 1,168mm, 연평균 유하량은 2,129백만m³이다. 소양강 유역의 매개변수 추정에는 1986년부터 1992년까지의 6년 동안의 자료를 사용하였고 1993년부터 1996년까지의 4년 동안의 자료는 검증에 사용하였다. 강수량은 인제, 용대, 백담사, 장주대, 추대, 명계, 법영, 서화, 원동, 귀둔, 건량, 창촌, 상남, 방동, 현리, 인제, 신평, 추양 관측소의 자료를 사용하였고 증발량은 인제, 춘천 관측소의 자료를 이용하였다. 영천댐은 경북 영천시와 포함시에 걸쳐있다. 영천댐의 유역면적은 약 3901 km²으로 연평균 강수량은 1101.05 mm 이고 연평균 유출량은 582.138 mm, 연평균 잠재 증발산량은 756.245 mm 이다. 모의에는 1983년에서 1986년까지의 자료를 이용하였고 1996년부터 1998년까지의 자료로 보정을 하였다. 영천댐의 우량 자료는 적장 관측소의 값을 사용하였고 증발량은 영천, 대구, 위성의 자료를 사용하였다.

2. 모의결과

SIMHYD 모형과 TANK 모형을 소양강댐과 영천댐 유역에 적용하여 유출량을 모의하였다. 이때 각 모형의 매개변수는 최적화 방법인 유전자 알고리즘과 SCE-UA, Pattern Search Multi-Start법을 이용하여 보정한 값을 사용하였다. 실측된 유출 곡선과 모의된 유출 곡선을 그림 7과 8에 나타내었다. 그림 7은 소양강댐 유역의 유출 곡선을 비교한 것으로 우측 y축은 SIMHYD 모형을 적용한 모의 유출이고 좌측의 y축은 TANK 모형을 적용하여 나타낸 모의 결과이다. 그림 8은 영천댐 유역의 유출곡선을 비교한 것으로 그림7과 같이 나타내었다. 이때 각 모형의 매개변수를 표 3과 4에 나타내었다. 표 3은 SIMHYD 모형, 표 4는 TANK 모형의 매개변수 값으로 여러 보정 방법에 대한 평균값으로 나타내었다.

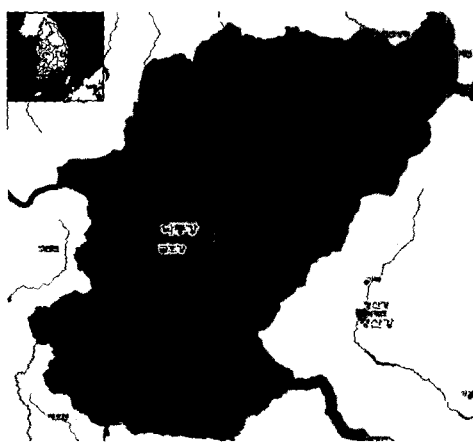


그림 5. 영천댐 유역

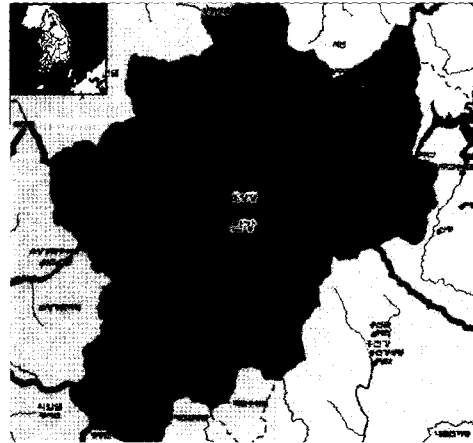


그림 6. 소양강댐 유역

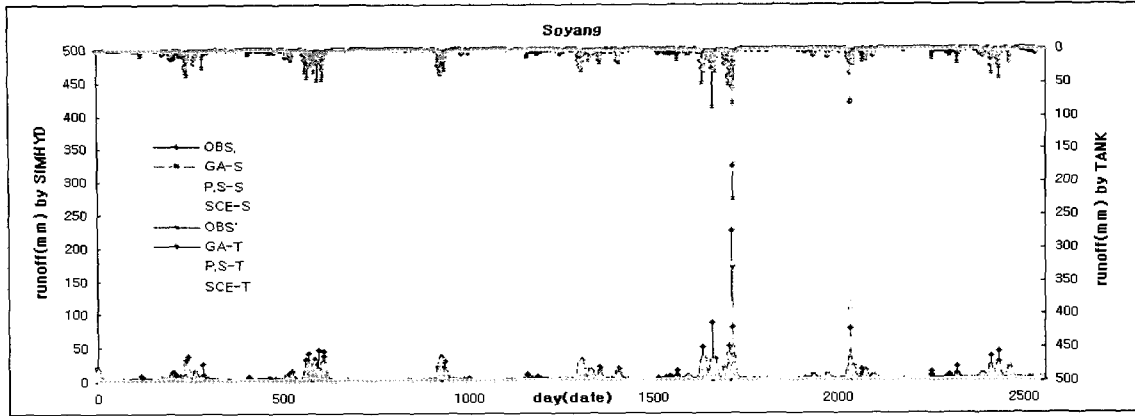


그림 7. 소양강댐 유역의 두 모형에 대한 여러 최적화 방법으로 모의한 유출과 관측유출

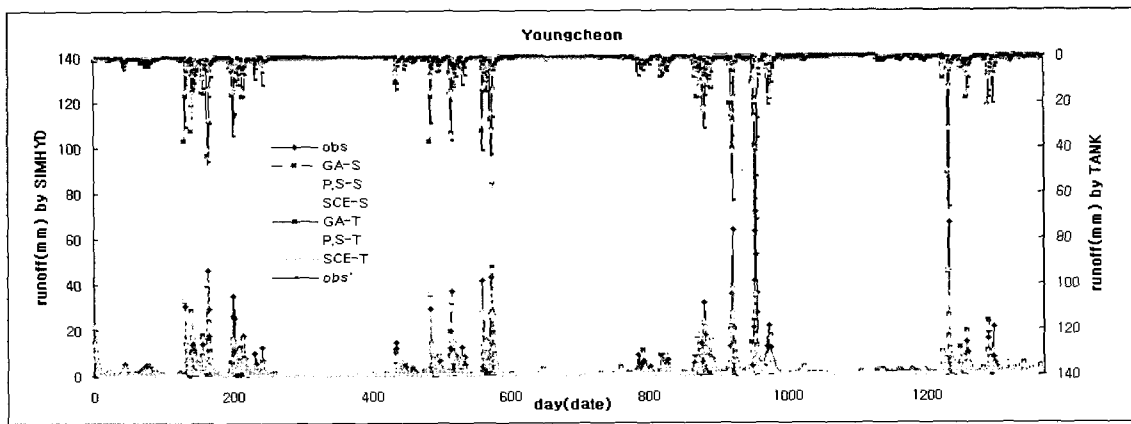


그림 8. 영천댐 유역의 두 모형에 대한 여러 최적화 방법으로 모의한 유출과 관측유출.

표 3. 보정된 SIMHYD모형의 매개변수

SIMHYD	소양강댐	영천댐
rechargeCoefficient	0.965	0.631
imperviousThreshold	0.852	3.901
infiltrationShape	1.065	0.273
rainfallInterceptionStoreCapacity	0.842	2.731
baseflowCoefficient	0.271	0.292
soilMoistureStoreCapacity	45.358	41.946
infiltrationCoefficient	310.969	179.206
interflowCoefficient	0.051	0.121
perviousFraction	0.962	0.947

표 4. 보정된 TANK모형의 매개변수

TANK	소양강댐	영천댐
H11	171.792	84.775
H12	32.322	39.573
H21	1.800	23.777
H31	9.989	21.097
a11	0.494	0.247
a12	0.330	0.310
a21	0.097	0.051
a31	0.030	0.038
a41	0.001	0.002
alpha	0.625	0.619
b1	0.114	0.019
b2	0.029	0.048
b3	0.007	0.006
C1	6.106	21.588
C2	8.301	16.186
C3	46.753	27.465
C4	895.689	345.191

TANK 모형에 비해 SIMHYD 모형은 침투치의 모의에는 효과적이거나 기저흐름을 모의하는데는 상대적으로 효율이 낮음을 알 수 있었다. 그러나 수문 곡선으로는 여러 최적화 방법에 대한 차이가 근소하여 최적화 효율

을 비교하기가 용이하지 않아 모델링의 상대적 수행을 목적함수에 따른 모의효율을 비교함으로써 분석하고자 한다. 여기서 모의효율이란 각 모의 유출량과 관측 유출량에 대한 식 6의 Nash-Sutcliffe 계수를 말한다. 표

5과 6에서는 다양한 최적화 방법에 대한 목적함수에 따른 모의 효율을 SIMHYD와 TANK모형을 비교하였다. 이때 모의 효율은 0.6이상을 양호, 0.8이상을 우수하다고 보았다. 두 유역 모두에서의 관측 유출량에 대한 모의 효율은 일별 모의를 통한 값은 0.6-0.8로 양호한 결과를 보여주고 있고 월별 모의를 통해서는 0.7-0.93정도의 모의의 효율이 높게 나타났다. 모의 효율면에 있어서 최적화 방법에 대한 큰 차이는 보이지 않으나 TANK 모형의 효율이 SIMHYD모형에 비해 높은 것을 알 수 있다.

소양강댐 유역의 모의 효율을 목적함수에 따라 살펴

보면 목적함수를 오차의 제곱의 합으로 하였을 경우 효율의 평균은 0.761로 표본 분산은 0.0017이 된다. 오차의 절대값의 합을 목적함수로 할 경우 평균은 0.715, 표본 분산은 0.0019, 마지막으로 Nash-Sutcliffe 계수를 목적함수로 하였을 경우 평균은 0.775, 표본 분산은 0.0007로 나타났다. 영천댐 유역의 경우는 오차의 제곱의 합을 목적함수로 하였을 경우 평균은 0.628, 분산은 0.0005, 오차의 절대값의 합을 목적함수로 하였을 경우 평균은 0.592, 분산은 0.0025, Nash-Sutcliffe 계수를 목적함수로 하였을 경우 평균은 0.630, 분산은 0.0004로 나타났다. 두 유역 모두의 경우 목적함수에 따른 모의 효

표 5. 소양강댐 유역에 대한 각 알고리즘별, 목적함수별 최적화 효율

소양 model	Objective function		Genetic Algorithm		SCE-UA		Pattern Search Multi	
			daily	monthly	daily	monthly	daily	monthly
TANK	calibration verification	sum of squares of errors	0.798	0.913	0.767	0.88	0.799	0.914
			0.852	0.963	0.802	0.933	0.851	0.963
	calibration verification	sum abs. {differences}	0.753	0.88	0.704	0.824	0.753	0.877
			0.863	0.962	0.822	0.914	0.862	0.96
	calibration verification	Nash-sutcliffe coef.	0.793	0.915	0.785	0.896	0.799	0.914
			0.839	0.963	0.833	0.953	0.851	0.963
SIMHYD	calibration verification	sum of squares of errors	0.734	0.905	0.693	0.853	0.776	0.863
			0.736	0.948	0.809	0.972	0.825	0.976
	calibration verification	sum abs. {differences}	0.657	0.869	0.661	0.83	0.753	0.84
			0.746	0.963	0.764	0.966	0.797	0.965
	calibration verification	Nash-sutcliffe coef.	0.772	0.886	0.727	0.868	0.776	0.861
			0.818	0.976	0.816	0.977	0.823	0.975

표 6. 두 유역에 대한 각 알고리즘별, 목적함수별 최적화 효율

영천 model	Objective function		Genetic Algorithm		SCE-UA		Pattern Search Multi	
			daily	monthly	daily	monthly	daily	monthly
TANK	calibration verification	sum of squares of errors	0.641	0.749	0.644	0.744	0.652	0.751
			0.778	0.83	0.806	0.832	0.809	0.82
	calibration verification	sum abs. {differences}	0.641	0.724	0.626	0.748	0.629	0.694
			0.828	0.865	0.743	0.867	0.801	0.909
	calibration verification	Nash-sutcliffe coef.	0.638	0.717	0.646	0.745	0.653	0.755
			0.799	0.869	0.829	0.844	0.804	0.849
SIMHYD	calibration verification	sum of squares of errors	0.619	0.725	0.588	0.723	0.621	0.73
			0.717	0.822	0.827	0.868	0.767	0.829
	calibration verification	sum abs. {differences}	0.589	0.716	0.547	0.704	0.517	0.691
			0.824	0.877	0.76	0.863	0.732	0.765
	calibration verification	Nash-sutcliffe coef.	0.621	0.724	0.597	0.723	0.622	0.731
			0.742	0.824	0.707	0.845	0.748	0.815

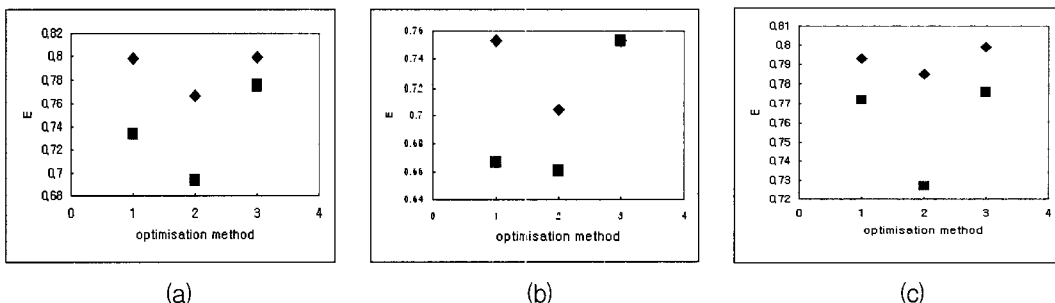


그림 9. 소양강댐 유역에 대한 SIMHYD, TANK 모형의 모의 효율 비교.

(a)는 오차의 제곱의 합을 목적함수로 한 경우의 모의 효율, (b)는 오차의 절대값의 합을 목적함수로 한 경우, (c)는 Nash-Sutcliffe 계수를 목적함수로 한 경우. X축은 최적화 방법. (1 : 유전자 알고리즘 2 : SCE-UA 3 : Pattern Search Multi-Start) (◆ : TANK, ■ : SIMHYD)

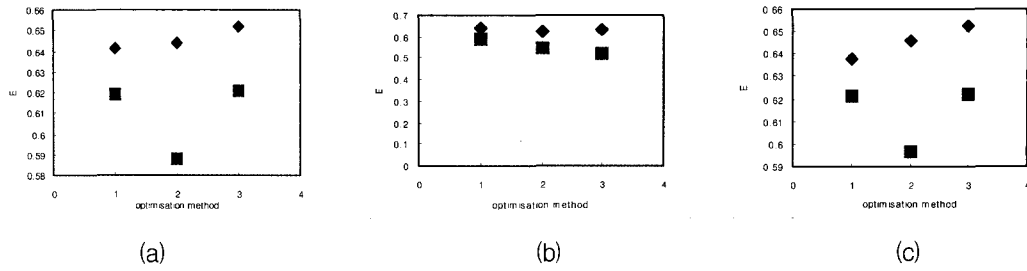


그림 10. 영천댐 유역에 대한 최적화 방법과 목적함수와 최적화 방법을 달리하여 나타낸 두 모형의 유출 모의 효율 비교(◆: TANK, ■: SIMHYD) (그림 9 참고)

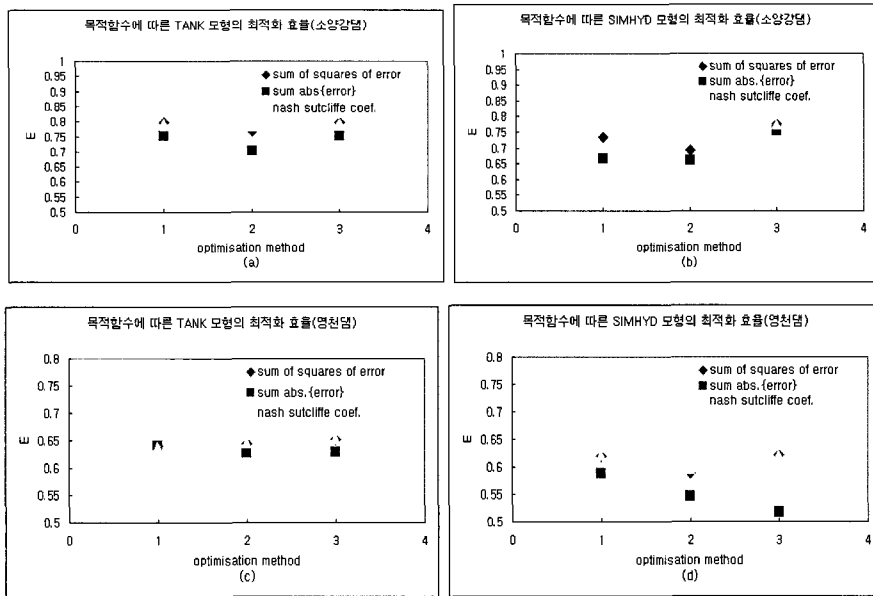


그림 11. 목적함수에 따른 모형의 최적화 효율. 각 그래프의 X축은 최적화 방법. (1: 유전자 알고리즘, 2: SCE-UA, 3: Pattern Search Multi-start)

율의 평균값은 거의 차이가 나지 않았으나 분산의 경우 오차의 절대값의 합을 목적함수로 하였을 경우 분산이 가장 크게 나타났고 Nash-Sutcliffe 계수를 목적함수로 하였을 경우 분산은 다른 경우에 비해 1/10배 정도 낮은 값을 띄고 있음을 알 수 있다. 따라서 Nash-Sutcliffe 계수는 최적화 방법에 상관없이 안정적인 값을 보여준다고 판단되었다.

그림 9와 10은 최적화 방법과 목적함수를 달리 했을 경우의 TANK모형과 SIMHYD모형의 효율을 비교한 것으로 그림 9는 소양강댐 유역에 대한 모의 효율이고 그림 10은 영천댐에 대한 모의 효율이다. 최적화 방법의 변화나 목적함수의 변화에도 불구하고 TANK 모형의 모의 수행 효율이 SIMHYD의 것보다 높음을 확인할 수 있다.

그림 11은 모의 효율로 나타낸 목적함수에 따른 모형의 수행 능력이다. X축은 최적화 방법을 나타내고 있다. 왼쪽부터 1은 유전자 알고리즘을, 2는 SCE-UA법을

을, 3은 Pattern Search Multi-Start법을 나타내고 마름모로 표시된 것은 오차의 제곱의 합을 목적함수로 한 경우의 모형의 모의 효율을, 정사각형으로 표시된 것은 오차의 절대값의 합을 목적함수로 한 경우의 모의 효율을, 삼각형으로 표시된 것은 Nash-Sutcliffe 계수를 목적함수로 한 경우의 모의 효율을 나타내고 있다.

전역 탐색기법인 유전자 알고리즘, SCE-UA법과 지역 탐색 기법인 Pattern Search Multi-Start법을 이용했지만 Pattern Search 보다 여러 지점에서 출발점을 잡는 Pattern Search Multi-Start의 최적화 효율의 개선을 가져와 지역 탐색법 임에도 불구하고 전역탐색법인 유전자 알고리즘과 최적화 효율 면에서 뒤지지 않은 양호한 결과를 얻을 수 있었다. 최적화 방법에 따른 모의 효율의 큰 차이는 없었으나 유전자 알고리즘을 보정 방법으로 사용하였을 경우 안정된 모의를 하였다. 대부분의 경우 목적함수를 Nash-Sutcliffe 계수로 한 경우의 모의 효율이 높은 것을 알 수 있다. 식 4는 유출의

최고치를 찾아내는데 효율적이고 식 5는 근소한 양의 흐름을 모의하는데 효율적이다. 유효계수 E 는 산정치와 측정치간의 일치성의 정도를 측정하는 결정 계수 R2 과 유사하지만 E는 모델이 측정치를 재현해 내는 능력을 직접적으로 측정한다는 점에서 차이가 있다. 목적 함수들의 실제 값들의 직접적인 비교보다 개선된 정보를 제공한다. 이는 목적함수의 값이 흐름부피에 따라 변화하므로 여러 다른 유역에서의 시물레이션 목적함수 값들은 각각 그 규모가 달라지기 때문이다. 뿐만 아니라 표 5과 6에서 목적함수에 따른 효율의 분산이 Nash-Sutcliffe 계수를 사용하였을 경우가 가장 낮은 것으로 나타났다. 낮은 분산은 각각 다른 최적화 방법이나 유역의 특성들의 상이함과 같은 다른 요인들에 대한 영향이 상대적으로 미소함을 의미한다. 따라서 안정적인 효율을 보여주는 Nash-Sutcliffe 계수를 사용하는 것이 유역에의 모의 효율을 나타낼 때 적합하다고 할 수 있다. 각 모형, 최적화 알고리즘, 목적함수의 선택에 따라

수행 시간의 차이가 발생한다. 유역의 강우-유출 모형을 설명하고자 할 경우 최적의 방법을 찾기 위해서는 보다 양호한 효율을 보이는 방법의 선택 못지않게 수행 시간의 고려도 중요한 부분을 차지한다. 표 7는 소양강댐 유역의 강우-유출 현상을 모의했을 때의 여러 모형과 최적화 방법, 목적함수를 달리 하였을 경우 연산에 걸린 시간을 나타내었다. 표 8은 영천댐의 경우 연산 수행 시간을 나타낸 것이다.

모형에 따른 모의 시간을 비교해볼 때 매개변수가 적은 SIMHYD모형에 비해 매개변수가 많은 TANK 모형을 적용하였을 경우가 가장 긴 시간을 소요하였다. 최적화 방법에 따른 시간은 SCE-UA법이 가장 많은 시간을 소요하였고 Pattern Search Multi-Start법이 가장 짧은 모의시간을 요구하였다. 목적함수의 선택에 따른 연산의 수행 시간은 큰 차이는 없으나 오차의 절대값의 합을 목적함수로 선택하였을 경우 수행시간이 단축되었다.

표 7. 소양강댐의 모의에 소요된 시간

소양강댐	최적화방법		Genetic Algorithm	SCE-UA	Pattern Search Multi-start
	목적함수				
TANK	오차의 제공의 합		3'40"	6'50"	1'56"
	오차의 절대값의 합		3'23"	6'35"	1'46"
	Nash-sutcliffe coef.		3'37"	6'28"	2'03"
SIMHYD	오차의 제공의 합		3'40"	5'16"	1'23"
	오차의 절대값의 합		2'46"	5'20"	1'03"
	Nash-sutcliffe coef.		2'58"	5'34"	1'12"

표 8. 영천댐의 모의에 소요된 시간

영천댐	최적화방법		Genetic Algorithm	SCE-UA	Pattern Search Multi-start
	목적함수				
TANK	오차의 제공의 합		1'57"	3'37"	1'13"
	오차의 절대값의 합		1'55"	3'17"	1'08"
	Nash-sutcliffe coef.		2'04"	3'45"	1'18"
SIMHYD	오차의 제공의 합		1'44"	3'18"	43"
	오차의 절대값의 합		1'36"	3'15"	36"
	Nash-sutcliffe coef.		1'47"	3'20"	42"

IV. 요약 및 결론

본 연구에서는 소양강댐 유역과 영천댐 유역에 TANK 모형과 SIMHYD 모형을 이용하여 유역의 유출 현상 설명에 효과적인 모형의 유출 모의능에 대해 알아 보았고 각 모형에 유전자 알고리즘과 SCE-UA법, Pattern Search Multi-Start법을 최적화 방법으로 적용하여 모형의 보정에 가장 효율적인 방법을 선정하였다. 또한 모형의 적합성을 판단하는 기준으로 세 가지 목적 함수(오차의 제공의 합, 오차의 절대값의 합, Nash-

Sutcliffe 계수)를 이용하여 목적함수에 따른 모의 효율의 변화를 알아보았다.

모의 효율만을 고려한다면 매개변수가 가장 많은 TANK 모형이 유역의 강우-유출 현상을 잘 설명한다고 할 수 있으나 연산 수행에 긴 시간이 소요되므로 가장 경제적인 방법이라고는 할 수 없다. 그러나 수행 시간의 차가 현저하지 않고 근소한 차이를 보이므로 좀더 안정되고 높은 모의 효율을 얻고 싶다면 TANK 모형을 선택하는 것이 적합하다고 판단된다.

최적화 방법 중 유전자 알고리즘법을 적용하였을 경

우 가장 높은 모의 효율과 낮은 분산을 보였다. Pattern Search Multi-Start법은 기존의 Pattern Search와는 달리 보다 여러 지점에서 출발점을 잡을 수 있고 출발점의 개수나 반복 횟수를 조절할 수 있어 최적화 효율의 개선을 가져와 지역 탐색법 임에도 불구하고 양호한 결과를 얻을 수 있었다. Pattern Search Multi-Start법은 지역 탐색법으로 전역 탐색법인 유전자 알고리즘과 SCE-UA법에 비하여 빠른 수행을 보인다. 그러나 지역 해에 빠질 확률이 있고 분산이 유전자 알고리즘에 비해 큰 값을 가지므로 안정된 모의를 수행하기 위해서는 적합하지 않다고 판단된다.

목적함수의 선정이 최적화에 결정적 역할을 하는 것은 아니지만 적절한 목적함수의 선택은 모의 효율의 향상에 기여할 수 있다. 매개변수의 보정 시 목적함수의 선택에 따른 모의 효율을 보면 오차의 제곱의 합과 Nash-Sutcliffe 계수를 목적함수로 하였을 경우 양호한 모의 효율을 보였다. 그 중 Nash-Sutcliffe 계수를 목적함수로 선택하였을 경우가 안정되면서도 가장 높은 모의 효율을 가지므로 목적함수로 적합하다고 판단된다.

감사의 글

본 연구는 21세기 프론티어 연구개발사업인 수자원의 지속적 확보기술 개발사업단의 연구비지원(과제번호 2-2-1)에 의해 수행되었습니다.

참고 문헌

강민구, 박승우, 임상준, 김현준(2002). "전역최적화 기법을 이용한 강우-유출모형의 매개변수 자동보정." **한국수자원학회 논문집**, 한국수자원학회, pp. 541-552.

신성철, 강경석, 서병하(2001). "TANK model의 매개변수 최적화에 관한 연구." **한국수자원학회 학술발표회 논문집**, 한국수자원학회, pp. 158-163.

강신욱, 이동률, 이상호(2003). "토양수분 저류구조를 가진 탱크모형의 보정에 관한 연구." **한국수자원학회 학술발표회 논문집**, 한국수자원학회, pp. 579-582.

배덕효, 정일원, 강태호, 노준우(2003). "유출성분을 고려한 Tank 모형의 매개변수 자동추정." **한국수자원학회 학술발표회 논문집**, 한국수자원학회, pp. 423-436.

이철균(2001). "Pattern Search Method와 결합된 니칭 유전 알고리즘." 동의대학교 산업기술연구지, Vol. 15, pp. 197-203.

Beken, K.(1979). "A monthly water balance model applied to two different watersheds." in *Surface and Subsurface Hydrology* 105: pp. 157-172.

Chiew, F.H.S., McMahon, T.A.(1994). "Application of the daily rainfall-runoff model MODHYDROLOG to twenty eight Australian catchments." *J.hydrol.*, 153: 383-416.

Chiew, F.H.S., Peel, M.C., and Western, A.W.(2002). "Application and testing of the simple rainfall-runoff model SIMHYD.", *Mathematical Models of Small Watershed Hydrology and Applications*. Edited by Vijay P. Singh, Donald K.Frevert. *Water Resources Publications, LLC*, pp. 335-367.

Chiew, F.H.S., Stewardson, M.J., and McMahon, T.A.(1993). "Comparison of six rainfall-runoff modeling approaches." *J.Hydrol.*, 147: 1-36.

Duan, Q., Sorooshian, S., and Gupta, V.K.(1994). "Optimal use of the SCE-UA global optimization method for calibrating watershed model." *J. Hydrol*, 158: pp. 265-284.

Gan, T.Y., and Biffu, G.F. (1996). "Automatic calibration of conceptual rainfall-runoff models: Optimization algorithms, catchment conditions, and model structure." *Water Resources Research*, 32(12), pp. 3513-3524.

Goldberg, D.E.(1989). "Genetic algorithms in search, optimization, and learning." Addison-Wesley Publishing Co., Reading, Mass.

Gupta, H.V., Sorooshian S., and Yapo P.O.(1999). "Status of automatic of calibration for hydrologic model : Comparison with multi level expert calibration." *Journal of hydrology*, 192, pp. 81-103.

Nash J., and Sutcliffe J.(1970). "River flow forecasting through conceptual models, I, A discussion of principles." *Journal of Hydrology*, 10, 282-290

Nicholas, V. Findler, Cher, LO, Ron, LO.(1987). "Pattern Search for Optimization." 29, pp. 41-50, *Mathematics and Computers in Simulation*, North-Holland.

Sugawara, M.(1995). "Tank model, in Computer models of watershed hydrology." *Water Resources Publications*, pp. 164-214, 1995.

Wang, Q.J.(1991). "The Genetic Algorithm and Its Application to Calibrating Conceptual Rainfall-Runoff Models." *Water Resources Research*, 27(9), pp. 2467-2471.

(논문번호:03-95/접수:2003.10.23/심사완료:2004.01.29)