

# 유전자알고리즘을 이용한 강우강도식 매개변수 추정에 관한 연구(Ⅰ): 기존 매개변수 추정방법과의 비교

## Parameter Estimation of Intensity-Duration-Frequency Curve Using Genetic Algorithm ( I ): Comparison Study of Existing Estimation Method

김 태 순\* / 신 주 영\*\* / 김 수 영\*\*\* / 허 준 행\*\*\*\*  
Kim, Taeson / Shin, Juyoung / Kim, Sooyoung / Heo, Jun-Haeng

### Abstract

The intensity-duration-frequency (IDF) curves by Talbot, Sherman and Japanese type formulas are widely used in South Korea since the parameters are easily estimated. However, these IDF curves' accuracies are relatively worse than those of the IDF curves developed by Lee et al. (1993) and Heo et al. (1999), and different parameters for the given return periods should be computed. In this study, parameter estimation method for the IDF curve by Heo et al. (1999) is suggested using genetic algorithm (GA). Quantiles computed by at-site frequency analysis using the rainfall data of 22 rainfall gauges operated by Korea Meteorological Administration are employed to estimate the parameters of IDF curves and minimizing root mean squared error (RMSE) and relative RMSE (RRMSE) of observed and computed quantiles are used as objective functions of GA. The comparison of parameter estimation methods between the empirical regression analysis and the suggested method show that the IDF curve in which the parameters are estimated by GA using RRMSE as an objective function is superior to the IDF curves using RMSE.

**keywords** : Genetic Algorithm, IDF curve, Parameter estimation

### 요 지

현재 국내에서 사용되고 있는 Talbot, Sherman, Japanese형 강우강도식은 매개변수추정이 용이하다는 장점이 있지만, 이원환 등(1993)과 허준행 등(1999)이 개발한 강우강도식에 비하여 정확도가 떨어지며 재현기간을 고려할 수 없다는 단점이 있다. 본 연구에서는 매개변수 추정상의 어려움 때문에 널리 사용되지 않는 허준행 등(1999)이 제안

- 
- \* 연세대학교 사회환경시스템공학부 BK21연구교수  
BK21 Lecturer, School of Civil and Environmental Engineering, Yonsei Univ., Seoul 120-749, Korea  
(e-mail: chaucer@yonsei.ac.kr)
  - \*\* 연세대학교 대학원 토목공학과 석사과정  
Graduate student, School of Civil and Environmental Engineering, Yonsei Univ., Seoul 120-749, Korea  
(E-mail: ausran@yonsei.ac.kr)
  - \*\*\* 연세대학교 대학원 토목공학과 박사과정  
Ph.D candidate, School of Civil and Environmental Engineering, Yonsei Univ., Seoul 120-749, Korea  
(E-mail: sykim97@yonsei.ac.kr)
  - \*\*\*\* 연세대학교 사회환경시스템공학부 토목환경공학과 교수  
Professor, School of Civil and Environmental Engineering, Yonsei Univ., Seoul 120-749, Korea  
(E-mail: jhheo@yonsei.ac.kr)

한 강우강도식의 매개변수를 유전자알고리즘을 이용하여 추정하는 방법을 제시하였다. 기상청 22개 지점에서 관측된 강우자료의 지속기간별 년최대치자료를 구축한 후 지점빈도해석을 적용한 결과를 이용하여 강우강도식의 매개변수를 추정하였으며, 최적화기법으로 사용된 유전자알고리즘의 목적함수로는 평균제곱근오차(RMSE)와 평균제곱근상대오차(RRMSE)를 사용하였다. 회귀분석에 근거한 기존의 강우강도식과 비교한 결과, 허준행 등(1999)이 개발한 강우강도식의 매개변수를 추정하는데 있어서 RRMSE값을 최소화시키는 목적함수를 사용하는 것이 가장 정확한 결과값을 얻을 수 있는 것으로 나타났다.

**핵심용어** : 유전자알고리즘, 강우강도식, 매개변수추정

## 1. 서 론

강우강도식은 임의의 재현기간 및 지속기간의 강우강도를 나타내는 식으로, 주로 수공구조물의 설계와 수자원계획의 기준이 되는 확률강우량의 산정이 필요한 경우 실무차원에서 보다 용이하게 확률강우량을 산정하기 위하여 사용되고 있다. 국내에서는 Talbot, Sherman, Japanese 형의 강우강도식과 같이 비교적 간단한 방법으로 매개변수를 산정할 수 있는 강우강도식과 이원환 등(1993), 허준행 등(1999)이 개발한 것과 같이 보다 향상된 정확도를 얻기 위해서 다수의 매개변수를 사용하는 강우강도식이 개발되어 있다. 이 중에서 비교적 간단한 형태를 가지는 Talbot, Sherman, Japanese 형의 강우강도식은 여러 가지 상용소프트웨어를 사용하여 비교적 높은 정확도를 가지는 매개변수를 추정할 수 있기 때문에 실무에 널리 사용되고 있으나, 이원환 등(1993)과 허준행 등(1999)이 개발한 강우강도식은 상용소프트웨어를 사용하여 매개변수를 추정하는 것이 상대적으로 어렵기 때문에 실무적인 차원에서는 널리 사용되지 않는다.

강우강도식의 개발과 관련된 연구사례들을 살펴보면, 국외에서는 Bernard(1932)가 장기간의 강우강도식을 개발한 이후로, 미국 전역에 대한 확률강우량도의 개발(Hershfield, 1961), 미국 서부의 확률강우량도(Miller et al., 1973) 및 동부지역의 확률강우량도(Frederick et al., 1977) 개발 등의 연구가 진행되었고, Koutsoyiannis et al.(1998)이 강우강도식을 수학적 방식을 통해 유도하였으며, Yu and Cheng(1998)은 지역빈도해석의 결과를 이용하여 강우강우-지속기간-생기빈도 곡선을 개발하였다. 최근에는 스케일링 기법과 멀티프랙탈을 이용한 강우강도식 유도에 관한 연구가 활발하다(Bougadis and Adamowski, 2006; Langousis and Veneziano, 2007).

국내에서는 이원환 등(1993)이 한국대표확률강우강도식을 유도하였으며, 유동훈(1995)은 일반화된 확률강우

강도식을 유도하였고, 한정훈 등(1996)은 선형화기법을 이용하여 강우강도식을 유도하였다. 이후 허준행 등(1999)이 지속기간별 적정확률분포형을 선정하여 선형화기법을 이용한 강우강도식을 제안하였으며, 이재준 등(2001)은 권역별로 강우강도식을 유도하였다. 유철상 등(2001)은 점 강우모형을 이용하여 강우강우-지속기간-생기빈도를 해석하였고, 한만신 등(2006)은 강우 패턴 변화를 고려한 인천지방 확률강우강도식을 제안하였다. 최근에는 단기간에서 정확도가 높은 강우강도식의 유도에 관하여 활발한 연구가 진행되고 있다(유철상 등, 2007).

본 연구는 총 2편으로 구성된 논문의 첫 번째 논문으로, 유전자알고리즘을 이용하여 강우강도식의 매개변수를 추정하는 것이 기존에 개발된 회귀분석을 통한 매개변수 추정방법과 비교하여 어떤 장점이 있는지를 알아보기 위한 목적과 함께 매개변수 추정상의 어려움 때문에 널리 사용되고 있지 못한 허준행 등(1999)이 개발한 강우강도식의 매개변수를 유전자알고리즘을 이용해서 추정하고, 이와 같은 방법으로 구한 강우강도식을 이용한 경우의 정확도와 기존에 개발된 회귀분석을 이용한 강우강도식을 이용한 경우의 정확도를 비교해서 유전자알고리즘을 이용한 경우의 적용성을 살펴보는 데 연구의 목적이 있다.

## 2. 기본이론

### 2.1 강우강도식

우리나라에서는 Talbot, Sherman, Japanese 형과 이원환 등(1993), 허준행 등(1999)이 개발한 강우강도식이 사용되고 있으며 이 중에서 Talbot, Sherman, Japanese 형은 통합형 강우강도식(이원환, 1980)으로 나타낼 수 있다. 아래 Eq. (1)로 나타낼 수 있는 통합형 강우강도식은 a, b, n 세 개의 매개변수를 사용하여 강우강도를 산정한다. 통합형 강우강도식은 다른 강우강도식들에 비해 적은 매개변수를 사용하기 때문에 매개변수 추정

이 용이하다는 장점이 있으나, 재현기간에 대한 고려가 되어있지 않아 같은 지속기간이라고 하더라도 짧은 재현기간에 해당하는 강우강도가 긴 재현기간에 해당하는 강우강도보다 커지는 역전현상이 다른 강우강도식들에 비해 많이 나타나며, 정확도 또한 다른 강우강도식들에 비해 낮은 단점이 있다. 이원환 등(1993)이 개발한 강우강도식은 아래의 Eq. (2)와 같으며 총 a, b, c, n 네 개의 매개변수를 사용하여 강우강도를 산정한다. 허준행 등(1999)이 개발한 강우강도식은 아래의 Eq. (3)과 같으며 총 a, b, c, d, n 다섯 개의 매개변수를 사용하여 강우강도를 산정한다. Eqs. (2) and (3)은 통합형 강우강도식인 Eq. (1)보다 높은 정확도를 가지는 장점이 있으나, 다수의 매개변수를 추정해야하는 불편함이 있다.

$$\text{통합형 (UNI)} = \frac{a}{t^n + b}$$

t: 지속기간(분) (1)

$$\text{이원환식 } (\leq E) = \frac{a + b \times \log_{10} T}{c + t^n}$$

t: 지속기간(분) T: 재현기간(년) (2)

$$\text{허준행식 (HEO)} = \frac{a + b \times \ln \frac{T}{t^n}}{c + d \times \ln \frac{\sqrt{T}}{t} + \sqrt{t}}$$

t: 지속기간(분) T: 재현기간(년) (3)

## 2.2 유전자알고리즘

### 2.2.1 유전자알고리즘(Genetic Algorithm)

유전자알고리즘은 1960년대에 Holland(1975)와 그의 동료들에 의해서 창안된 기법으로 다윈이 제안한 ‘적자생존(survival of the fittest)’을 컴퓨터 기법화한 최적화 방법이다. 유전자알고리즘은 수 십 년간 발전되어 오면서 수 없이 많은 형태로 변형되었지만, 1) 해(chromosome)의 집합인 군(population)을 기본단위로 하고, 2) 교배(crossover), 돌연변이(mutation), 전치(inversion) 등의 유전자연산자(genetic operator)를 사용하며, 3) 룰렛 휠 선택법(roulette wheel selection), 토너먼트 선택법(tournament selection) 등의 선택법을 통해서 다음 세대에 더 좋은 해를 넘겨주는 방법을 통칭한다. 유전자알고리즘이 최적해를 찾아나가는 원리에 관해서는 많은 연구들이 있었지만, Holland가 제안한 이진코딩에서의 building block theory를 가장 기본적인 것으로 말할 수 있다. 예를 들어, 하나의 해(chromosome)을 구성하

는 개개의 gene은 각각 단독으로 활동하는 것이 아니라 일련의 블록(block)으로 활동하고 이런 일련의 블록들의 집합 중에서도 해를 좀 더 개선시키는 방향으로 발전시키는 블록(building block)이 생성된다는 것이 Holland의 building block theory이다. 이런 building block theory를 이해하기 위해서는 반드시 schema라는 개념의 이해가 필요하며, 여기서 말하는 schema는 해(chromosome)를 구성하는 gene을 임의의 group으로 묶을 수 있는 도구를 제공하는 것으로 Goldberg(1989)의 책에 좀더 자세히 설명되어 있다. Fig. 1은 유전자알고리즘 중에서 가장 간단한 형태인 simple genetic algorithm의 진행순서를 나타낸 순서도이다.

### 2.2.2 NSGA-II (Non-dominated Sorting Genetic Algorithm-II)

본 연구에서는 매개변수를 추정하는데 사용할 유전자알고리즘 기법으로 NSGA-II (Non-dominated Sorting Genetic Algorithm-II)를 사용하였다. NSGA-II는 최근에 가장 활발히 사용되고 있는 다목적 유전자알고리즘 기법중의 하나로써, 기존 NSGA(Srinivas and Deb, 1994)의 단점을 보완하여 엘리티즘(elitism)을 적용하였으며, 순위를 결정하기 위한 복잡도(order of complexity)를 감소시켰고, 해의 다양성(diversity)을 확보하기 위한 sharing 기법에서 사용자가 임의로 설정하는 매개변수를 없앤 군집거리방법(crowding distance)을 적용한 기법이다(Deb et al., 2002). NSGA-II는 원래 다목적 문제의 최적화를 위해서 개발되었으나, 단일목적 함수를 적용하더라도 선택법에 사용된 군집거리 계산방법이나 순위매김 방법이 단일목적 함수의 최적해를 구하는데 무리가 없고, 또한 가장 간단한 유전자알고리즘 형태인 SGA(simple genetic algorithm)보다 더 좋은 결과값을 주기 때문에 본 연구에서 사용하였다.

## 3. 분석방법

본 연구에서는 Eq. (1)~(3)의 강우강도식 중 정확도가 비교적 높다고 알려진 Eqs. (2) and (3)을 사용하였으며 지점빈도해석 결과와 유전자알고리즘을 이용하여 강우강도식의 매개변수를 추정한 후 평균제곱근오차(Root Mean Squared Error, RMSE)와 평균제곱근상대오차(Relative RMSE, RRMSE)를 이용하여 각각의 강우강도식의 적용성을 평가하였다. 또한 기존에 강우강도식의 매개변수를 추정하는 방법과의 비교를 위하여 한국건설기술연구원(2000)에서 1999년까지의 자료를 이용하여 구한 Eq. (3)의 매개변수를 사용한 결과와 비교해 보았다.

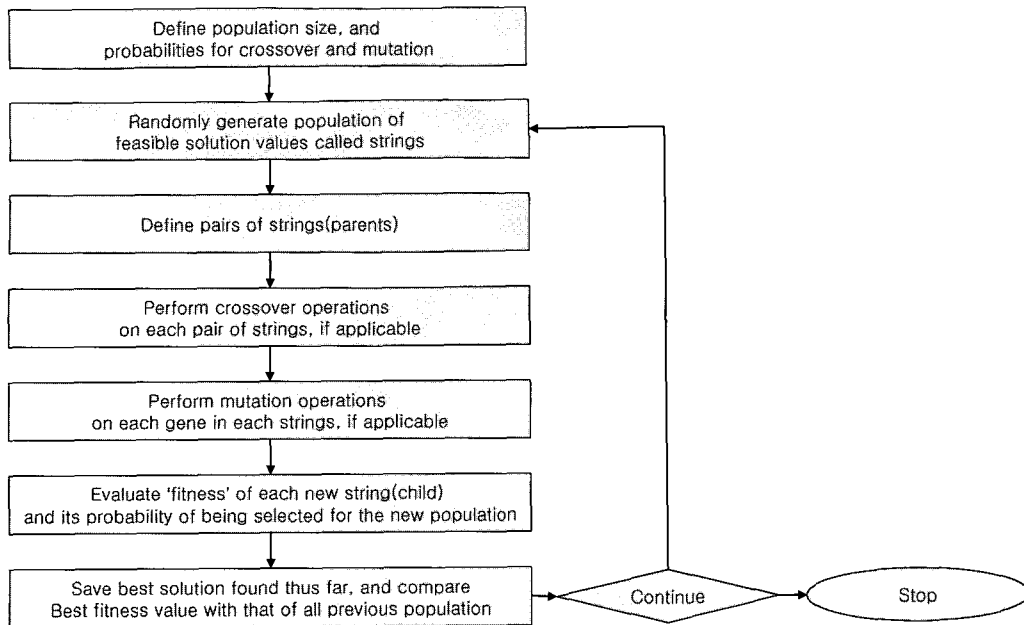


Fig. 1. Flowchart of Genetic Algorithm

### 3.1 지점빈도 해석

본 연구에서는 기상청 22개 지점의 1999년까지 강우량 자료를 이용하여 빈도해석을 실시하였고 확률가중모멘트(Probability Weighted Moments) 방법을 매개변수 추정방법으로 사용하였다. 각 지점의 적정확률분포형은 한국건설기술연구원(2000)의 연구결과를 인용하였으며, 사용된 자료는 지속기간 1, 2, 6, 9, 12, 15, 18, 24, 48시간의 강우자료를 이용하였고 재현기간 2, 3, 5, 10, 20, 30, 50, 70, 80, 100, 200, 300, 500년에 대하여 확률강우량을 산정하였다. Table 1은 본 연구의 대상 지점들의 정보를 나타낸 것이다.

### 3.2 목적함수 선정

유전자알고리즘을 사용하여 강우강도식의 매개변수를 추정하기 위해서는 적절한 목적함수의 선정이 필요하다. 본 연구에서는 평균제곱근오차(RMSE), 평균제곱근상대오차(RRMSE), 평균편차, 평균절대편의(Mean Absolute Bias), Nash & Sutcliffe Index(Nash and Sutcliffe, 1970)를 각각 사용하여 목적함수로서의 적용성을 확인해보았다. 아래 Eq. (3)~(8)에서  $n$ 은 전체 자료수,  $n_1$ 은 지속기간의 개수,  $n_2$ 은 재현기간의 개수,  $\hat{Q}_{ij}$ 는 지속기간이  $i$ 번째이고 재현기간이  $j$ 번째인 추정값,  $Q_{ij}$ 는 지속기간이  $i$ 번째이고 재현기간이  $j$ 번째인 관측값,  $\bar{Q}_j$ 는 재현기간이  $j$ 번째인 관측값의 평균을 나타낸다. Eq. (4)로 표현된 RMSE는 서로 다른 지속기간과 재현기간으로 표현된 지점빈도해석 결과와 강우강도식으로부터 구한 추정값간의 차이를 동시에 최소화시킬

수 있는 형태의 목적함수이고, Eq. (5)로 표현된 RRMSE는 관측값( $Q_{ij}$ )에 대한 상대적인 RMSE를 표현하기 때문에 작은 관측값에 대한 오차를 표현하기에 용

Table 1. Site Characteristics

Site name	Probable distribution	Record length
Sokcho	GEV	32
Chuncheon	GUM	33
Gangneung	GUM	42
Seoul	GEV	71
Incheon	GEV	47
Wonju	GEV	27
Suwon	GEV	36
Seosan	GEV	32
Cheongju	GEV	33
Daejeon	GEV	31
Chupongnyeong	GUM	45
Pohang	GEV	46
Gunsan	GUM	32
Daegu	GEV	84
Jeonju	GUM	59
Ulsan	GEV	53
Gwangju	GEV	61
Busan	GUM	57
Tongyeong	GEV	32
Mokpo	GEV	77
Yoesu	GEV	57
Wando	GUM	29

GEV: General Extreme Value Distribution  
GUM: Gumbel Distribution

이하고, Eq. (6)은 관측값( $Q_{ij}$ )과 추정값( $\hat{Q}_{ij}$ )의 평균값 사이의 차이를 최소화하기 위한 목적함수이다. 또한 Eq. (7)은 관측값과 추정값 사이의 절대적인 차이를 최소화하기 위한 목적함수이며, Eq. (8)은 주로 강우-유출 모형에서 구한 결과값의 적합도를 검토하는데 사용되는 통계량으로 결정계수(coefficient of determination)와 유사한 형태를 가지고 있다.

본 연구에서는 이와 같은 5가지의 목적함수를 단일 목적함수로 이용하여 강우강도식의 매개변수를 추정하는데 적용하였으며 그 결과 Eq. (4)의 RMSE와 Eq. (5)의 RRMSE를 사용한 매개변수 추정법을 사용하였다. 앞서 언급한 것과 같이 Eqs. (4) and (5)는 각각의 관측값과 추정값 사이의 관계를 고려하여 계산되는 통계량이기 때문에 강우강도식으로부터 구한 추정값의 정확도를 나타내는데 적절하다고 할 수 있는 반면에, Eq. (6)은 평균값에 대한 적합도만을 계산하기 때문에 사용하기에 어려움이 있다. Eqs. (7) and (8)은 식의 형태나 결과값이 Eq. (4)의 RMSE와 매우 유사하였기 때문에 보다 널리 사용되고 있는 RMSE와 상대적인 오차를 표현하는데 적합한 RRMSE를 목적함수로 사용하였다.

평균제곱근오차(RMSE) (4)

$$= \sqrt{\frac{1}{n} \sum_{j=1}^{n_2} \sum_{i=1}^{n_1} (\hat{Q}_{ij} - Q_{ij})^2}$$

평균제곱근상대오차(RRMSE) (5)

$$= \sqrt{\frac{1}{n} \sum_{j=1}^{n_2} \sum_{i=1}^{n_1} \left( \frac{\hat{Q}_{ij} - Q_{ij}}{Q_{ij}} \right)^2}$$

평균편차 =  $\left( \frac{\sum_{j=1}^{n_2} \sum_{i=1}^{n_1} Q_{ij}}{n} - \frac{\sum_{j=1}^{n_2} \sum_{i=1}^{n_1} \hat{Q}_{ij}}{n} \right)$  (6)

평균절대편의(Mean Absolute Bias) (7)

$$= \frac{1}{n} \sum_{j=1}^{n_2} \sum_{i=1}^{n_1} |Q_{ij} - \hat{Q}_{ij}|$$

Nash and Sutcliffe Index (8)

$$= \frac{\sum_{j=1}^{n_2} \left( \sum_{i=1}^{n_1} (\hat{Q}_{ij} - Q_{ij})^2 \right)}{\sum_{j=1}^{n_2} \left( \sum_{i=1}^{n_1} (Q_{ij} - \bar{Q}_j)^2 \right)}$$

### 3.3 단일목적 유전자알고리즘으로 강우강도식의 매개변수 추정

선정된 목적함수인 RMSE와 RRMSE를 이용하여 22개 대상지점에 대한 강우강도식의 매개변수를 추정하였다. 강우강도식의 매개변수는 RMSE와 RRMSE를 단일 목적으로 사용하여 NSGA-II를 적용하여 추정하였다. NSGA-II는 원래 다목적 최적화에 적합한 기법이지만 단일목적함수를 사용하더라도 최적해를 구하는데는 무리가 없으며, 또한 연속된 논문 II편의 결과와 비교하기 위하여 적용하였다. 매개변수 추정에 사용된 강우강도식은 Eqs. (2) and (3)으로써 지속기간과 재현기간을 고려해 주는 항이 강우강도식안에 존재하여 재현기간에 따른 매개변수를 산정할 필요가 없다. 그러나 Eq. (1)로 표현된 통합형 강우강도식의 경우 재현기간에 대한 항이 강우강도식 안에 존재하지 않아, 재현기간별로 매개변수를 추정해야 하며 실제로 사용되는 매개변수의 수는 재현기간 개수의 3의 배수가 된다. 따라서, 본 연구에서는 통합형을 제외한 Eqs. (2) and (3)을 이용하여 매개변수를 추정하였다.

Fig. 2는 Table 1에 나타난 22개 대상지점 중에서 5개 지점에 대한 지점빈도해석 결과와 강우강도식으로 구한 결과를 나타낸 것으로, 위에서부터 각각 부산, 여수, 울산, 인천 그리고 포항 지점에 대해서 구한 지점빈도해석결과와 강우강도식으로부터 구한 결과를 도시한

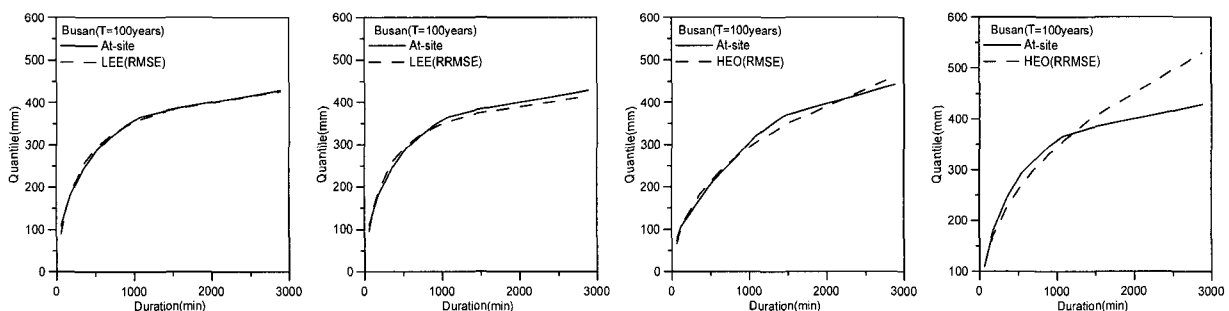


Fig. 2. Quantiles Estimated by At-Site Frequency Analysis and IDF Curve

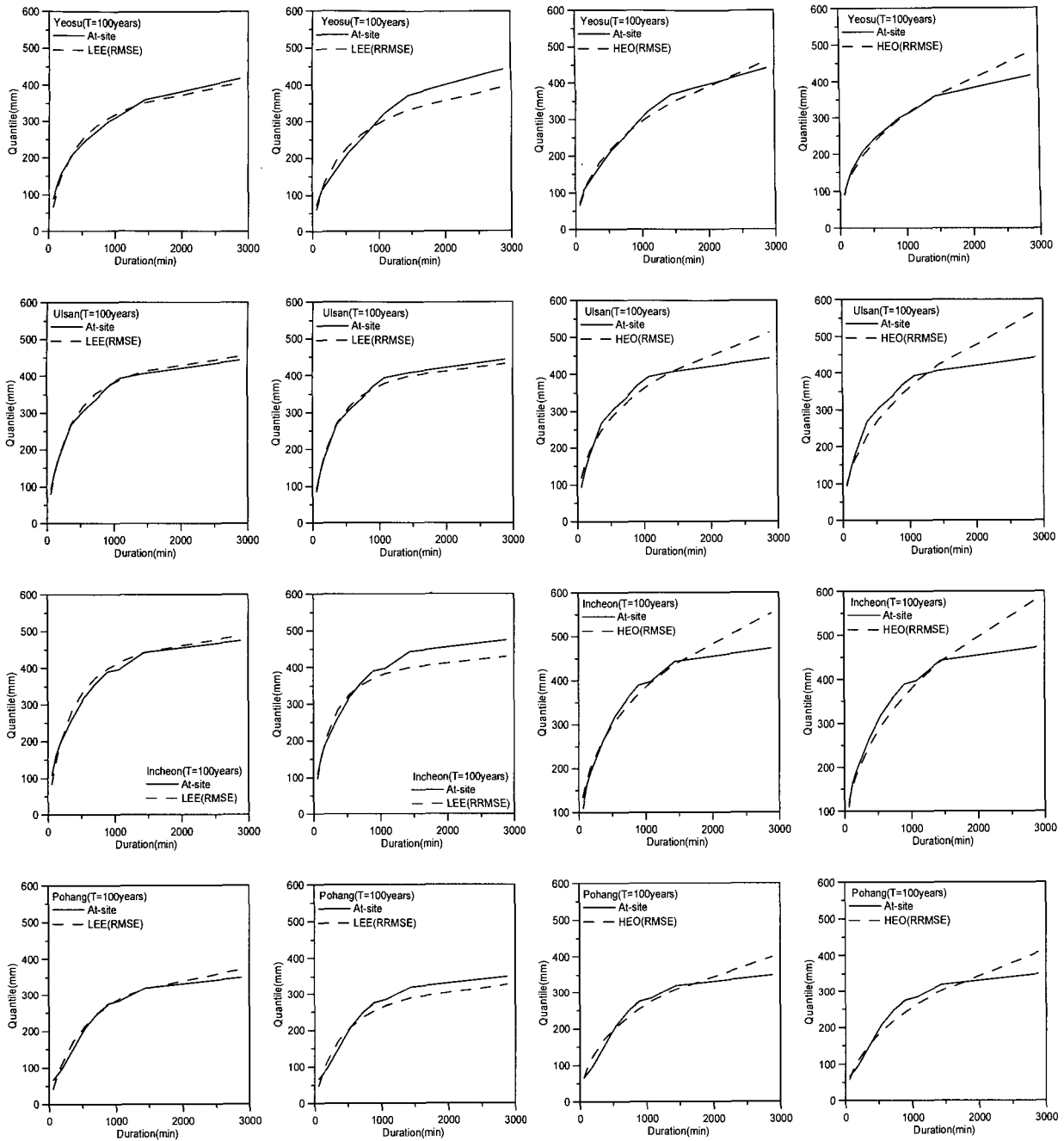


Fig. 2. Quantiles Estimated by At-Site Frequency Analysis and IDF Curve(continued)

것이다. Fig. 2에서 At-site는 지점빈도해석을 통해서 구한 확률강우량값이고, LEE(RMSE)는 Eq. (2)의 매개변수를 RMSE를 최소화시키는 목적함수를 사용하여 추정된 후 이를 통하여 구한 강우량 값이고, LEE(RRMSE)는 Eq. (2)의 매개변수를 RRMSE를 목적함수로 사용하여 추정된 값이고, HEO(RMSE)는 Eq. (3)의 매개변수를 RMSE를 목적함수로 사용하여 추정된 값이고, HEO(RRMSE)는 Eq. (3)의 매개변수를 RRMSE를 목적함수로 사용하여 추정된 값을 나타낸다.

#### 4. 유전자알고리즘을 이용한 매개변수 추정법의 정확도

기상청 22개 지점에 대하여 Eqs. (2) and (3)의 매개변수를 유전자알고리즘을 사용하여 추정하였다. 목적함수로는 RMSE와 RRMSE를 사용하였고, 단일목적 유전자 알고리즘을 이용하여 매개변수를 추정하였다.

Table 2는 추정된 매개변수로 산정된 확률강우량과 지점빈도해석을 통해 산정된 확률강우량과의 RMSE와 RRMSE 그리고 결정계수를 대상지점인 22개 지점에

Table 2. RMSE, RRMSE, and Coefficient of Determination

	RMSE	RRMSE	COD
A	18.1334	0.1132	0.9625
B	22.8404	0.0692	0.9398
C	16.5200	0.1175	0.9671
D	23.1210	0.1322	0.8997
E	33.3303	0.1082	0.8647

COD : Coefficient of determination

대해서 평균한 값을 나타낸 표이다. Table 2에서 A는 Eq. (3)을 RMSE를 목적함수로 추정된 결과의 정확도, B는 Eq (3)을 RRMSE를 목적함수로 추정된 결과의 정확도, C는 Eq (2)를 RMSE로 목적함수로 추정된 결과의 정확도, D는 Eq (2)를 RRMSE를 목적함수로 추정된 결과의 정확도, 그리고, E는 Eq (3)을 기존의 회귀방법으로 추정된 매개변수의 정확도를 나타냈다. E의 경우에 사용된 매개변수는 한국건설기술연구원(2000)의 보고서에 수록된 매개변수를 사용하였다.

Table 2에서 기존에 추정된 매개변수를 이용한 결과인 E와 유전자알고리즘을 이용하여 새롭게 추정된 매개변수를 이용한 결과인 A, B, C 그리고 D를 비교해보면, RMSE값은 기존에 추정된 매개변수를 사용한 결과인 E가 다른 4가지 경우보다 모두 낮은 정확도를 보이는 것으로 나타났고, RRMSE값을 기준으로 B를 제외한 나머지 3가지 경우에서 모두 기존에 추정된 매개변수를 이용한 결과인 E가 더 높은 정확도를 보이는 것으로 나타났다. 또한, 결정계수인 COD를 살펴보면 기존에 추정된 매개변수를 사용한 결과인 E가 다른 4가지 경우보다 모두 결정계수값이 적은 것으로 나타나서 더 낮은 정확도를 보이는 것으로 나타났다.

Eq. (4)로 표현된 RMSE는 관측값과 추정값 사이의 차이를 최소화시키는데는 효과적이지만 관측값의 절대적인 크기와 상관없이 전체적인 차이만을 최소화시키기 때문에, 상대적으로 관측값의 크기가 적은 경우에 생길 수 있는 오차가 관측값의 크기가 큰 경우에 생길 수 있는 오차보다 덜 반영되는 성질을 가지고 있다. RRMSE는 관측값과 추정값 사이의 상대적인 차이를 최소화하는 것으로 관측값의 크기에 상관없이 상대적으로 일정한 오차를 보이는 매개변수를 추정하여 지속기간에 상관없이 일정한 정확도를 보이는 장점이 있다. 특히, 강우강도식의 정확도를 나타내는데 있어서는 상대적으로 확률강우량 값이 적은 단기간의 지속기간에 대한 정확도를 보다 효과적으로 측정할 수 있는 통계량이 필요하며, 이런 측면에서는 RRMSE가 RMSE보다 기법상의 정확도를 평가하는데 있어서는 더 우월하다고 말할 수

있다.

따라서, 앞선 Table 2의 결과를 살펴보면 A~D까지의 새로 추정된 매개변수를 이용한 결과들 중에서 가장 좋은 경우는 RRMSE가 가장 적은 B를 선택할 수 있으며, B의 경우는 기존의 회귀분석에 의한 결과인 E와 비교해서는 RMSE, RRMSE, 그리고, COD 모두에 있어서 더 우수한 결과를 보여주고 있다. 결과적으로, Eq (3)을 목적함수로 RRMSE를 적용하여 매개변수를 추정하는 방법은 기존 방법보다 높은 정확도를 얻을 수 있는 매개변수를 추정하는 것으로 나타났으며 적용된 추정방법 중 가장 적용하기 적합한 것으로 판단된다.

Table 3은 22개 지점의 각 추정방법에 따른 정확도를 나타낸 표이며, Fig. 3은 대구, 목포, 서울, 광주, 그리고 전주의 각 추정방법별 확률강우량을 나타낸 그래프이다. Fig. 3은 1, 2, 3, 6, 9, 12, 15, 24, 48시간에 해당하는 총 9개의 지속기간에 대한 결과를 도시한 그림이며, Fig. 3의 오른쪽 그림에서 C의 경우가 왼쪽 그림의 B보다 더 좋은 결과인 것처럼 보이는 것은, 지속기간 24시간 미만에 해당하는 결과값을 이용해서 도시한 그래프가 상대적으로 잘 보이지 않기 때문이며, 전체 지속기간에 대한 오차의 비율을 줄이는 형태의 목적함수인 RRMSE를 사용한 경우의 결과를 Table 2에서 살펴보면 B의 경우가 다른 경우들보다 훨씬 더 좋은 RRMSE 값을 가지는 것을 알 수 있다.

## 5. 결 론

기존에 널리 사용되던 Talbot, Sherman, Japanese형의 강우강도식은 매개변수추정이 용이한 반면에 정확도가 떨어지는 단점이 있었고, 허준행 등(1999)이 개발한 강우강도식은 정확도는 높은 반면에 매개변수추정이 쉽지 않은 단점이 있었다. 본 연구에서는 허준행 등(1999)이 개발한 강우강도식을 대상으로 보다 손쉽게 매개변수를 추정하는 방법중의 하나로 유전자알고리즘을 적용하였으며 기존에 개발된 회귀분석을 통한 결과와 비교하여 적용성을 검토하였다.

Table 3. RMSE, RRMSE, and Coefficient of Determination by Each Parameter Estimation Method

Site name	A			B			C			D			E		
	RMSE	RRMSE	COD	RMSE	RRMSE	COD	RMSE	RRMSE	COD	RMSE	RRMSE	COD	RMSE	RRMSE	COD
Sokcho	11.2440	0.0466	0.9923	11.8176	0.0392	0.9915	21.8568	0.0950	0.9133	21.8568	0.1833	0.9708	24.6086	0.0862	0.9619
Chuncheon	10.8058	0.0629	0.9884	13.1205	0.0445	0.9830	12.9065	0.1346	0.9835	20.9889	0.0947	0.9564	19.5963	0.0634	0.9607
Gangneung	11.6400	0.1065	0.9889	17.3963	0.0473	0.9753	7.1245	0.0951	0.9959	12.2593	0.0667	0.9877	26.4216	0.0665	0.9411
Seoul	15.7271	0.0574	0.9780	15.9456	0.0490	0.9773	26.8189	0.1809	0.9359	40.7773	0.1128	0.8518	21.8407	0.0705	0.9567
Incheon	31.4306	0.1518	0.9123	36.5798	0.0912	0.8812	26.4298	0.1247	0.9380	35.9827	0.0975	0.8851	52.4013	0.1317	0.7524
Wonju	26.1109	0.1080	0.9528	29.3736	0.0657	0.9402	19.3888	0.1443	0.9739	30.8394	0.0980	0.9341	33.2966	0.0796	0.9218
Suwon	32.2934	0.1560	0.9452	38.0996	0.0944	0.9237	30.8136	0.1376	0.9501	42.8993	0.0935	0.9033	52.5916	0.1211	0.8524
Seosan	18.9238	0.0776	0.9766	21.6841	0.0713	0.9693	23.3522	0.1877	0.9644	38.6944	0.1166	0.9023	24.2700	0.0747	0.9609
Cheongju	22.8560	0.1327	0.9411	27.5035	0.0889	0.9147	19.9613	0.1320	0.9551	27.3990	0.1087	0.9154	43.6187	0.1569	0.7817
Daejeon	18.2219	0.0982	0.9718	23.7677	0.0721	0.9520	24.8203	0.1306	0.9476	34.6240	0.0869	0.8981	36.1317	0.1005	0.8872
Chupongnyeong	8.7899	0.0694	0.9815	10.7668	0.0472	0.9722	10.5892	0.1315	0.9731	14.8889	0.1032	0.9468	17.1891	0.0834	0.9260
Pohang	23.8830	0.1412	0.9275	26.7105	0.0996	0.9093	17.5908	0.1578	0.9606	25.3261	0.1398	0.9184	42.5452	0.3104	0.7651
Gunsan	13.1322	0.1244	0.9603	17.4379	0.0754	0.9299	7.9135	0.0592	0.9856	9.5177	0.0481	0.9791	28.1004	0.1013	0.8111
Daegu	11.7242	0.1118	0.9497	15.8424	0.0720	0.9082	10.6320	0.1102	0.9587	11.5562	0.1045	0.9512	24.5941	0.0927	0.7692
Jeonju	13.0246	0.1078	0.9662	17.6003	0.0635	0.9384	11.0811	0.0914	0.9756	14.4241	0.0754	0.9586	26.6319	0.0817	0.8540
Ulsan	29.0666	0.2169	0.9164	37.9178	0.1087	0.8577	13.0611	0.0651	0.9831	15.7667	0.0570	0.9754	23.6034	0.0844	0.9061
Gwangju	15.2078	0.1061	0.9621	18.9416	0.0640	0.9412	10.5072	0.1064	0.9819	15.5048	0.0776	0.9606	52.7857	0.1495	0.7173
Busan	20.3887	0.1673	0.9480	30.4998	0.0904	0.8836	7.5011	0.0609	0.9930	8.8695	0.0517	0.9902	51.1341	0.1256	0.6616
Tongyeong	24.2958	0.1274	0.9670	29.5510	0.0770	0.9512	28.9946	0.1416	0.9530	41.3538	0.0953	0.9045	52.5382	0.1411	0.8432
Mokpo	8.1282	0.0548	0.9907	11.0429	0.0243	0.9829	12.4680	0.1519	0.9782	20.7811	0.1043	0.9394	19.3753	0.1180	0.9462
Yoesu	10.3372	0.1013	0.9862	17.2815	0.0605	0.9615	12.6425	0.0960	0.9794	17.4072	0.0910	0.0668	17.3548	0.0581	0.9599
Wando	21.7027	0.1639	0.9712	33.6084	0.0769	0.9310	6.9854	0.0502	0.9970	6.9457	0.0328	0.9971	42.6372	0.0827	0.8859
Average	18.1334	0.1132	0.9625	22.8404	0.0692	0.9398	16.5200	0.1175	0.9671	23.1210	0.1322	0.8997	33.3303	0.1082	0.8647

COD : Coefficient of determination



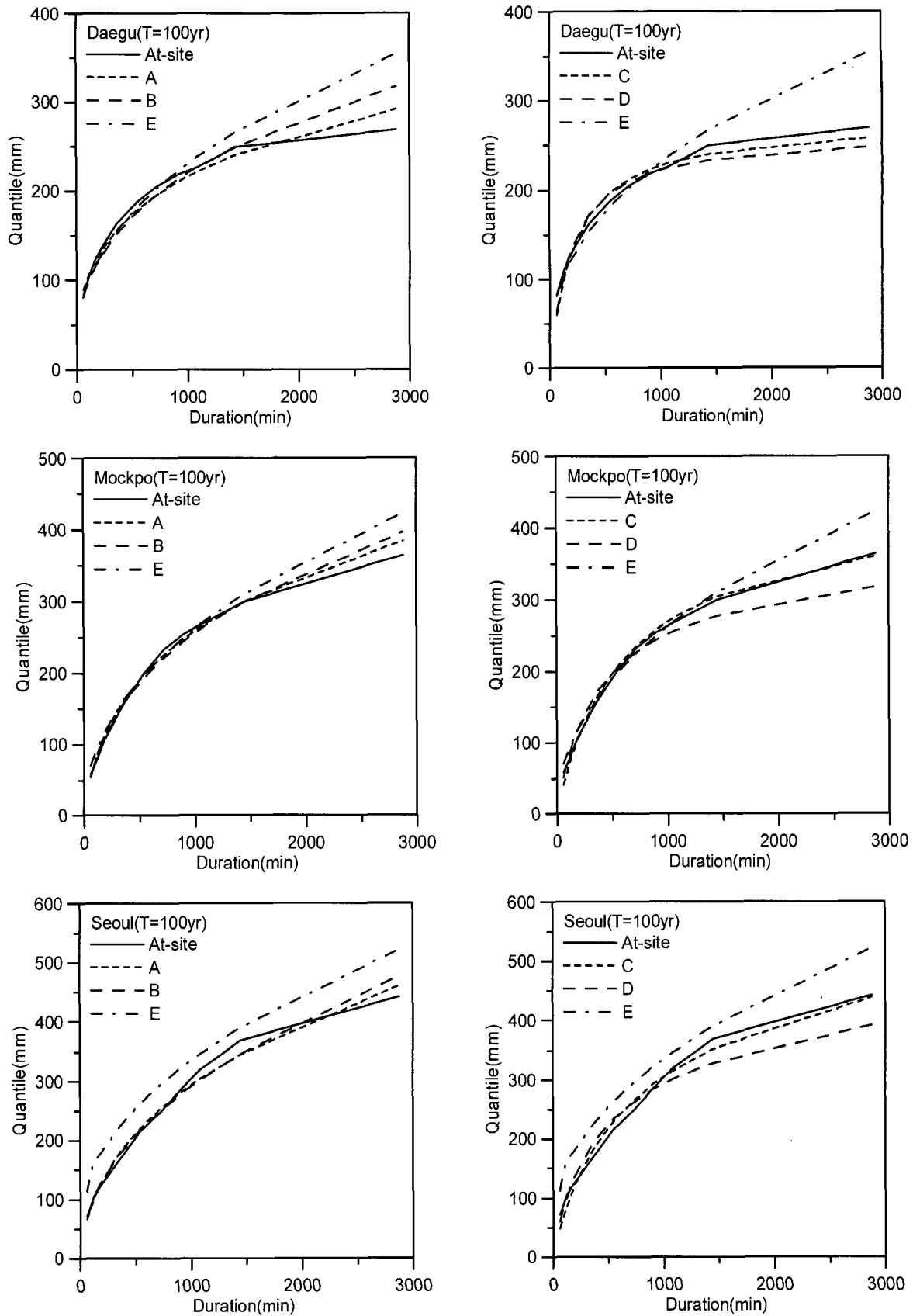


Fig. 3. Quantiles Computed by IDF Curves Whose Parameters are Estimated by A, B, C, D, and E Methods

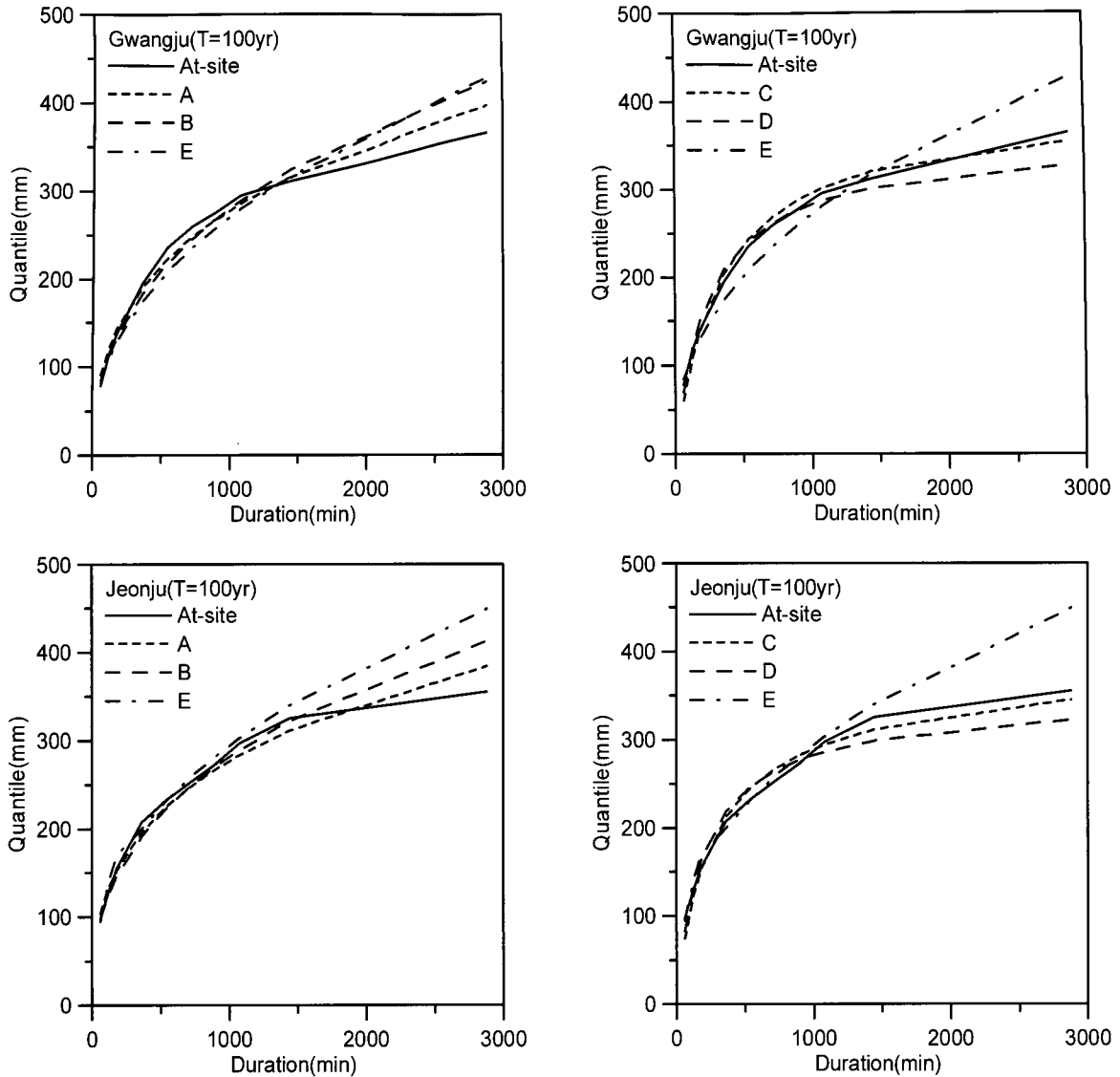


Fig. 3. Quantiles Computed by IDF Curves Whose Parameters are Estimated by A, B, C, D, and E Methods(continued)

최적화 기법인 유전자알고리즘을 효과적으로 적용하기 위해서는 주어진 최적화 문제의 수행성(performance)을 평가하는 척도인 목적함수를 적절하게 선택하는 절차가 필요하다. 본 연구에서는 유전자알고리즘의 목적함수로써 평균제곱근오차(RMSE)와 평균제곱근상대오차(RRMSE)를 사용하였으며 각각을 단일목적함수로 사용한 결과, RRMSE를 목적함수로 사용하여 추정된 매개변수를 사용한 경우의 결과가 RMSE를 사용한 경우보다 더 높은 정확도를 보이는 것으로 나타났다.

즉, RMSE를 단일 목적으로 매개변수를 추정할 경우 Eqs. (2) and (3) 모두 기존의 방법보다 RMSE값이 높은 정확도를 나타냈으나, RRMSE값은 낮은 정확도를 나타내었다. RRMSE를 목적함수로 사용할 경우 Eq. (2)의 경우 RMSE값은 기존 방법보다 정확하나

RRMSE는 낮은 정확도를 보였고 Eq. (3)의 경우 RMSE와 RRMSE 그리고 결정계수 모든 기준에서 기존방법보다 정확한 추정방법으로 나타났다. 따라서, 강우강도식의 매개변수를 추정하기 위해서는 Eq. (3)의 형태로 구성된 강우강도식을 이용하고 유전자알고리즘의 목적함수로는 RRMSE를 사용하는 것이 가장 정확도가 높은 것으로 나타났다.

### 감사의 글

본 연구는 건설교통부가 출연하고 한국건설교통기술평가원에서 위탁시행한 2003년도 건설핵심기술연구개발사업(03산학연C03-01)에 의한 도시홍수재해관리기술 연구사업단의 연구성과입니다.

## 참 고 문 헌

- 유동훈 (1995). "확률 강우강도식의 일반화." **한국수자원학회 학술발표회논문집**, 한국수자원학회, pp. 325-330.
- 유철상, 김남원, 정광식 (2001). "점 강우모형과 강우강도-지속기간-생기빈도 해석." **한국수자원학회 논문집**, 한국수자원학회, 제34권, 제6호, pp. 577-586.
- 유철상, 박창열, 김경준, 진경수 (2007). "분단위 강우강도-지속기간-재현기간 관계의 유도: 모포 마 분포의 적용." **한국수자원학회 학술대회논문집**, 한국수자원학회, pp. 168-172.
- 이원환, 박상덕, 최성열 (1993). "한국대표확률강우강도식의 유도." **대한토목학회 논문집**, 대한토목학회, 제13호, 제1호, pp. 115-120.
- 이원환 (1980). "도시 하천 및 하수도 개수계획상의 계획 강우량 설정에 관한 추계학적 해석." **대한토목학회 논문집**, 대한토목학회, 제28권 제4호, pp. 81-94.
- 이재준, 이정식, 박종영 (2001). "계획강우량 설정을 위한 권역별 확률강우강도식의 유도." **대한토목학회 논문집**, 대한토목학회, 제21권, pp. 1-10.
- 한국건설기술연구원 (2000). **강우자료 추출 및 확률강우량 산정 프로그램 개발 최종보고서**. 연세대학교.
- 한만신, 최계운, 정연중, 안경수 (2006). "최대강우 패턴 변화를 고려한 인천지방 확률강우강도식의 제안." **한국수자원학회 논문집**, 한국수자원학회, 제39권, 제6호, pp. 521-531.
- 한정훈, 김경덕, 허준행, 조원철 (1996). "선형화기법에 의한 확률강우강도식의 유도." **한국수자원학회 학술발표회 논문집**, 한국수자원학회, pp. 232-237.
- 허준행, 김경덕, 한정훈 (1999). "지속기간별 강우자료의 적정분포형 선정을 통한 확률강우강도식의 유도." **한국수자원학회 논문집**, 한국수자원학회, 제32권, 제3호, pp. 247-254.
- Bernard, M.M. (1932). "Formulas for rainfall intensities of long durations." *Trans. ASCE*, Vol. 96, pp. 592-624.
- Bougadis, J., and Adamowski, K. (2006). "Scaling model of a rainfall intensity-duration-frequency relationship." *Hydrological Process*, Vol. 20 No. 17, pp. 3747-3757.
- Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T. (2002). "A fast and elitist multiobjective genetic algorithm: NSGA-II." *Evolutionary Computation, IEEE Transactions on*, Vol. 6, No. 2, pp. 182-197.
- Frederick, R.H., Meyers V.A., and Auciello, E.P. (1977). *Five- to 60-minute Precipitation Frequency For the Eastern and Central United States*. NOAA Tech. Mere. NWS HYDRO-35, Washington, DC.
- Goldberg, D. E. (1989). *Genetic algorithms in search optimization & machine learning*. Addison Wesley, Massachusetts.
- Hershfield, D. M. (1961). *Rainfall frequency atlas of the United States for durations from 30 minutes to 24 hours and return periods from 1 to 100 years*. Technical Paper 40. US Dept of Comm., Weather Bureau, Washington, D.C.
- Holland JH. (1975). *Adaptation in natural and artificial systems*. University of Michigan Press.
- Koutsoyiannis, D., Kozonis, D., and Manetas, A. (1998). "A mathematical framework for studying rainfall intensity-duration-frequency relationships." *Journal of hydrology*, Vol. 206 No. 1, pp. 118-135.
- Langousis, A., and Veneziano, D. (2007). "Intensity-duration-frequency curves from scaling representations of rainfall." *Water Resource Research*, Vol.43 No.2, W02422.
- Miller, J. F., Frederick, R. H., and Tracey, R. J. (1973). *Precipitation-frequency atlas of the conterminous western United State*. NOAA Atlas 2. National Weather Service, Silver Spring, Maryland.
- Nash, J. E., and Sutcliffe, J. V. (1970). "River flow forecasting through conceptual models." *Journal of Hydrology*, Vol. 3 No. 3, pp. 282-290.
- Srinivas, N., and Deb, K. (1994) "Multiobjective optimization using nondominated sorting in genetic algorithms." *Evolutionary Computation*, Vol. 2, No. 3, pp. 221-248.
- Yu, P. S., and Cheng, C. J. (1998). "Incorporating uncertainty analysis into a regional IDF formula." *Hydrological Process*, Vol. 12, pp. 713-726.

(논문번호:07-69/접수:2007.07.23/심사완료:2007.09.14)

