

# 유전자알고리즘을 이용한 강우강도식 매개변수 추정에 관한 연구(II): 장·단기간 구분 방법의 제시

## Parameter Estimation of Intensity-Duration-Frequency Formula Using Genetic Algorithm(II): Separation of Short and Long Durations

신 주 영\* / 김 태 순\*\* / 김 수 영\*\*\* / 허 준 행\*\*\*\*

Shin, Juyoung / Kim, Taeson / Kim, Sooyoung / Heo, Jun-Haeng

### Abstract

In this study, the separation of short and long durations for estimation the parameters of IDF curve is suggested by using Multi-Objective Genetic Algorithm (MOGA). Objective functions are to minimize root mean squared error (RMSE) and relative RMSE between observed and computed values. The criteria for separation are two; the first one is to estimate more precisely the parameters of IDF curve and the second is to make a single IDF curve without non-continuous duration point. For this purpose 22 rainfall recording gauges operated by Korea Meteorological Administration are selected and three IDF curves that are used generally in South Korea are tested. The result shows that the IDF curve developed by Heo et al. (1999) would be the best of three tested IDF curves, and the suggested parameter estimation method using MOGA can compute more reliable parameters compared with empirical regression analysis.

**keywords** : Multi-Objective Genetic Algorithm, IDF curve, Parameter estimation, Short and long duration dividing

### 요 지

본 연구에서는 강우강도식의 매개변수를 추정하기 위해서 다목적 유전자알고리즘의 목적함수로 RMSE와 RRMSE를 적용하여 보다 객관적인 기준으로 장·단기간을 구분하는 방법을 제시하였다. 매개변수를 추정하기 위한 장·단기간을 구분하는 방법으로는 정확도를 기준으로한 방법과 그래프상의 접점을 이용하는 방법을 적용하였으며, 기상청에서 관리하는 22개 지점에 대하여 국내에서 널리 사용되고 있는 강우강도식을 이용하여 그 적용성을 살펴본다. 매개변수를 추정하는 방법에 있어서는 재현기간별로 매개변수를 구하는 방법의 정확도를 평가하였으며, 장·

\* 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 사회환경시스템공학부 BK21 연구교수

BK21 Lecturer, School of Civil and Environmental Engineering, Yonsei Univ., Seoul 120-749, Korea  
(e-mail: chaucer@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)

단기간을 구분하는 방법과 함께 재현기간별로 매개변수를 구하는 방법을 적용하였다. 국내에서 널리 사용되고 있는 세 가지 형태의 강우강도식 중에서 가장 높은 정확도를 얻을 수 있는 것은 허준행 등(1999)이 개발한 강우강도식인 것으로 나타났고, 기존의 선형회귀분석을 통해서 매개변수를 추정하는 것보다 다목적 유전자알고리즘을 적용한 결과가 더 높은 정확도를 보여주는 것으로 나타났다.

**핵심용어** : 다목적 유전자알고리즘, 강우강도식, 매개변수추정, 장·단기간 구분

## 1. 서론

본 연구는 보다 정확도가 높은 강우강도식의 매개변수를 추정하기 위해서 다목적 유전자알고리즘 기법을 적용한 연구로 여러 가지 다양한 방법을 이용하여 가장 효율적으로 매개변수를 추정하는 방법을 제시하고자 하였다. 현재 국내에서 널리 사용되는 강우강도식은 가장 기본적인 형태를 가지고 있는 Talbot, Sherman, Japanese형의 강우강도식을 활용한 통합형(이원환, 1980)과 함께 재현기간별로 매개변수 중의 일정 부분을 서로 다르게 추정하는 이원환 등(1993)이 개발한 한국 대표확률강우강도식, 그리고 재현기간과 상관없이 서로 같은 매개변수를 사용하면서도 정확도는 높은 허준행 등(1999)이 개발한 식이 사용되고 있다.

이와 같은 강우강도식의 정확도를 향상시킬 수 있는 방법의 하나로 장기간과 단기간에 따른 매개변수를 서로 다르게 추정하는 방법을 검토하였다. 확률강우강도를 전대지수에 도시해보면 자료의 양상이 변화되는 부분을 찾을 수 있는데, 이 지속기간을 중심으로 전반부를 단기간, 후반부를 장기간으로 나누어 매개변수를 추정할 경우 기존의 단기간과 장기간을 나누지 않고 매개변수를 추정하는 경우보다 높은 정확도의 매개변수를 추정할 수 있다(한국건설기술연구원, 2000). 이렇게 장·단기간에 해당하는 지속기간을 나누는데 있어서 독측으로 지속기간을 구분하는 것이 아닌 보다 객관적인 기준을 제시하고자 한다. 여기서는 주로 허준행 등(1999)이 개발한 강우강도식을 중심으로 평균제곱근오차(Root Mean Squared Error, RMSE)와 평균제곱근상대오차(Relative Root Mean Squared Error, RRMSE)를 동시에 최적화 시키는 형태의 다목적 최적화 문제의 해를 구하는 방법을 통해서 강우강도식의 매개변수를 구하는 방법을 제시하였다. 또한, 장·단기간을 구분하는 경우와 하지 않는 경우의 정확도 비교, 앞서 언급한 세 가지 강우강도식을 이용할 경우 각각의 강우강도식이 가지는 정확도 비교, 그리고 장·단기간 구분방법에 따른 정확도를 비교해서 국내에 가장 적합한 강우강도식의 매개변수 추정방법을 제시하고자 하였다.

본 연구에서는 표기상의 편리를 위하여 다음의 Eq. (1)~(3)에 표현된 강우강도식을 각각 UNI, LEE, HEO로 표기하였고, 표현상의 무리가 없는 한도내에서 Eq. (1), (2) and (3)과 혼용하였다.

$$\text{통합형 (UNI)} = \frac{a}{t^n + b} \quad t: \text{지속기간(분)} \quad (1)$$

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

$t$ : 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. 강우강도식의 매개변수 추정을 위한 다목적 유전자알고리즘의 적용

### 2.1 다목적 유전자알고리즘 적용

본 연구에서 사용된 다목적 유전자알고리즘은 NSGA-II(Non-dominated Sorting Genetic Algorithm-II, Deb et al., 2002)로 공학과 관련된 여러 분야에서 널리 사용되고 있는 다목적 유전자알고리즘 기법중의 하나이다. NSGA-II와 같은 다목적 유전자알고리즘을 효과적으로 적용하기 위해서는 반드시 사용된 목적함수들이 서로 상충되는(conflicting) 관계를 가지고 있는지를 확인해야 한다.

특히, NSGA-II와 같이 파레토지배(Pareto dominance)개념을 이용해서 최적해들 사이의 순위를 결정하는 최적화기법은 목적함수의 개수가 늘어날수록 구해지는 파레토최적해(Pareto-optimal solution)의 개수가 급격히 늘어나는 단점이 있기 때문에 최소한의 목적함수를 이용하는 것이 보다 효과적이며, 최근에는 2개 이상의 목적함수를 가지고 있는 다목적 최적화문제의 해를

구하기 위한 다양한 시도가 이루어지고 있으며(di Pierro et al., 2007), 여기서는 앞선 논문 1편에서 사용한 평균제곱근오차(Root Mean Squared Error, RMSE)와 평균제곱근상대오차(Relative RMSE, RRMSE)를 이용하여 다목적 최적화문제의 해인 파레토최적해(Pareto optimal solution)를 계산했다.

비교적 관측된 자료의 수가 많은 대구지점과 서울지점에 대하여 다목적 유전자알고리즘의 목적함수로 RMSE와 RRMSE를 적용해 본 결과 Figs. 1 and 2와 같은 파레토최적해로 이루어진 파레토프론트(Pareto-front)를 얻을 수 있었다. Figs. 1 and 2에서 가로축은 RMSE값을 나타내고 세로축은 RRMSE값을 나타내며, 두 가지 목적함수중 하나의 목적함수인 RMSE값이 적어질수록 반대로 RRMSE값은 커지며 RMSE값이 커지면 RRMSE값은 작아지는 trade-off 관계가 잘 성립함을 볼 수 있다. 대상지점인 22개 지점에 대해서도 역시 모두 비슷한 형태의 trade-off 관계가 성립함을 볼 수 있었고, 따라서 논문 1편에서 사용한 RMSE와 RRMSE는 다목적 유전자알고리즘의 목적함수로도 사용이 가능한 것으로 판단이 되어 RMSE와 RRMSE를 목적함수로 적용하여 매개변수를 추정하였다.

## 2.2 다목적 목적함수인 RMSE와 RRMSE의 특징

본 연구에서 선정한 목적함수인 RMSE와 RRMSE의 특징을 살펴보기 위하여, 논문 1편에서 대상지점으로 선택한 기상청 관측소 22개 지점에 대해서 지점빈도해석을 실시한 결과와 다목적 유전자알고리즘을 적용한 후 여기서 구한 파레토최적해중에서 하나의 목적함수값만을 최소로 만드는 최적해를 이용한 결과를 비교했다. 다목적 유전자알고리즘에서 구한 파레토최적해들은 서로 비지배(nondominated) 관계에 있는 최적해들로, 하

나의 해가 다른 해보다 더 좋다거나 나쁘다고 말할 수 없는 관계에 있다. 즉, 파레토프론트상의 최적해들은 모두 같은 순위의 우선순위를 가지고 있는 해로 흔히 1순위 최적해(rank-1 optimal solution)라고 말한다(Deb, 2001; Goldberg, 1989; Tan et al., 2005).

따라서, 다목적 유전자알고리즘의 파레토최적해는 주어진 최적화 문제를 보다 잘 이해하고 있는 사용자가 다시 해석하는 과정을 거쳐서, 같은 1순위 최적해중에서도 사용자가 원하는 최적해가 어떤 해인지를 가려내는 과정을 거치게 된다. 본 연구에서는 이와 같은 파레토최적해의 검토를 위해서 구해진 파레토프론트상의 양쪽 끝점에 위치하는 파레토최적해에 상응하는 매개변수를 이용한 결과값을 비교해 보았다. 즉, 아래의 Fig. 3은 Fig. 1에서 도시한 대구지방의 파레토최적해중에서 각각 가장 적은 RMSE값과 RRMSE값을 가지는 파레토프론트상의 양쪽 끝점을 선택하는 과정을 나타낸 것으로, 좌상단에 위치한 끝점은 RMSE값이 가장 적은 대신에 가장 큰 RRMSE값을 가지게 되며, 반대로 우하단에 위치한 끝점은 RRMSE값이 가장 적은 대신에 RMSE값이 가장 큰 결과값을 가지게 된다.

Fig. 4는 기상청 22개 지점에 대해서 다목적 유전자알고리즘을 적용한 후, 앞서 언급한 방법을 이용하여 각각 RMSE와 RRMSE를 최소로 만드는 매개변수를 선택하고 이렇게 선택된 매개변수를 적용한 HEO식을 이용해서 확률강우량을 산정한 결과값과 22개 대상지점에 대해서 FARD2002(국립방재연구소, 2002)를 이용한 지점빈도해석 결과값과의 절대편의(absolute bias)를 계산한 결과를 나타낸 것이다. 예를 들면, 지속기간 1시간에 해당하는 2개의 막대(bar)중에서 왼쪽에 위치한 가로빔을 친 막대는 전체 22개 지점에 대해서 다목적 유전자알고리즘을 수행해서 얻어진 파레토프론트상의

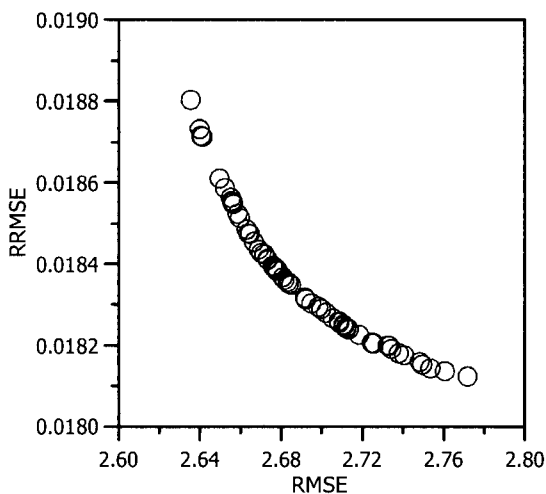


Fig. 1. Pareto Front (Daegu)

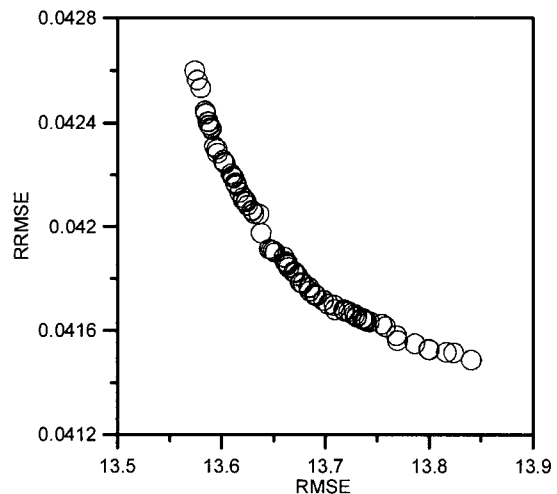


Fig. 2. Pareto Front (Seoul)

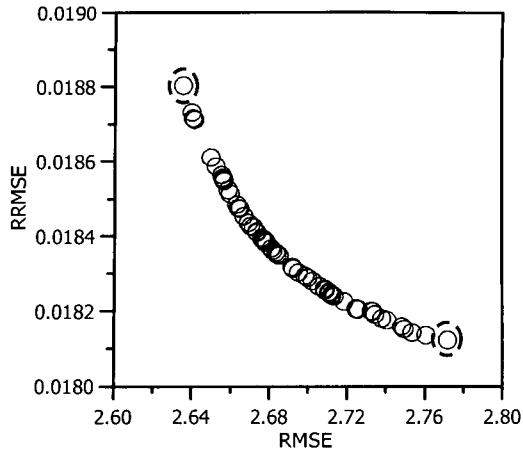


Fig. 3. Two Extreme Pareto-Optimal Solutions Which Make Only One Objective Function Minimum Value

파레토최적해들 중에서 RMSE값이 가장 적은 최적해에 상응하는 매개변수를 HEO식에 적용한 후 나온 확률강우량 값과 역시 22개 지점에 대해서 FARD2002를 적용해서 구한 지점빈도해석 결과값을 비교한 절대편의를 평균한 값을 의미한다. 같은 방법으로 오른쪽에 위치한 세로빔금으로 표시한 막대는 RRMSE값이 가장 적은 경우에 해당하는 파레토최적해를 이용했을 때의 결과를 나타낸 것이다.

Fig. 4에서 주목할 것은 단기간과 장기간에 해당하는 지속기간에서의 RMSE와 RRMSE를 사용한 경우의 절대편의 값이 서로 다른 특성을 보여준다는 것이다. 즉, 상대적으로 단기간인 지속기간 1~3시간까지의 결과를 살펴보면 RRMSE로 표시된 경우의 절대편의가 RMSE로 표시된 경우의 절대편의보다 적은 것을 볼 수 있고, 반대로 상대적으로 장기간인 지속기간 6, 9, 12, 15, 48 시간에서는 RMSE로 표시된 경우의 절대편의가 RRMSE보다 적은 것을 볼 수 있으며, 지속기간 18, 24 시간에 해당하는 경우에는 RMSE와 RRMSE가 상당히 비슷한 값을 보이는 것을 알 수 있다.

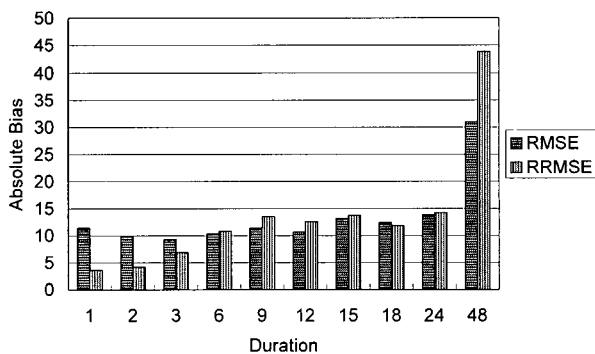


Fig. 4. Comparison of RMSE of Parameters Estimated by RMSE and RRMSE

이러한 결과가 나타나게 된 이유는 RRMSE는 참값으로 사용된 지점빈도해석의 결과값과 추정값으로 사용된 강우강도식에서 구한 확률강우량간의 상대오차를 계산하는 통계량이지만, RMSE는 전체적인 오차에 관한 절대값만을 고려하기 때문인 것으로 판단된다. 예를 들어, 확률강우량값이 50mm일 경우 5%의 오차와 확률강우량이 500mm일 경우 5%의 오차는 RRMSE의 경우 0.05로 똑같은 값이나, 실제적으로 확률강우량의 차이를 볼 경우 2.5mm와 25mm의 차이로 나타나게 된다.

이러한 이유로 RRMSE를 목적함수로 사용할 경우 작은 확률강우량 값을 가지는 단기간에서는 비교적 정확한 매개변수를 추정하나, 큰 확률강우량 값을 가지는 장기간에서는 단기간에 비하여 부정확한 매개변수를 추정하는 것이다. RRMSE와는 반대로 RMSE를 목적함수로 사용하여 매개변수를 추정할 경우 추정된 매개변수를 RMSE값이 10mm가 되도록 추정하였을 경우 전체적으로 약 10mm정도의 오차를 가지는 매개변수를 추정한 것을 의미한다. 즉 500mm에서 10mm정도의 차이는 큰 차이가 아니나, 50mm에서 10mm는 큰 차이로 볼 수 있다. 이러한 이유로 RMSE를 목적함수로 사용하였을 경우 큰 확률강우량을 갖는 장기간에서는 비교적 정확한 매개변수를 추정하고, 작은 확률강우량을 갖는 단기간에서는 장기간과 비교하여 부정확한 매개변수를 추정하는 것이다.

따라서, 결과적으로 다목적 유전자알고리즘에서 구한 파레토최적해중에서 최소의 RRMSE를 가지는 매개변수를 이용해서 구한 강우강도식은 지속기간이 단기간인 경우에 해당하는 강우량을 보다 정확하게 산정할 수 있으며, 반대로 최소의 RMSE를 가지는 파레토최적해를 적용한 경우에는 상대적으로 지속기간이 장기간인 경우에 해당하는 강우량을 보다 더 정확하게 산정할 수 있다고 할 수 있다.

### 3. 장·단기간을 동시에 고려한 강우강도식의 매개변수 추정

앞선 절에서 설명한 방법을 이용하면, 단기간과 장기간의 지속기간에 대해서 모두 높은 정확도를 보이는 강우강도식의 매개변수를 구할 수 있다. 본 연구에서는 정확도만을 높이기 위해서 RMSE와 RRMSE중에서 하나의 목적함수만을 최소화시키는 매개변수를 사용하여 장·단기간 지속기간을 나누는 경우(COMBL\_1)와 RMSE와 RRMSE를 각각 최소화시키는 매개변수를 적용한 경우의 확률강우량 그래프의 점점을 이용한 경우(COMBL\_2)로 나누어서 장·단기간 지속기간을 나누는 방법을 제시하였다.

### 3.1 정확도를 기준으로한 장단기간 구분(COMBL\_1)

정확도만을 고려하는 경우 다음의 과정을 거쳐서 장·단기간 지속기간을 구분하게 된다.

- 1) RMSE와 RRMSE 두 가지 목적함수를 사용하는 다목적 유전자알고리즘을 수행한 후에 계산된 파레토프론트 상에서, RMSE를 최소로 하는 매개변수와 RRMSE를 최소로 하는 매개변수를 선택한다.
- 2) 선택된 매개변수 중 RMSE를 최소로 하는 매개변수를 사용하여 사용자가 정한 지속기간과 재현기간별로 각각 확률강우량을 산정한다.
- 3) 선택된 매개변수 중 RRMSE를 최소로 하는 매개변수를 사용하여 사용자가 정한 지속기간과 재현기간별로 각각 확률강우량을 산정한다.
- 4) 지점빈도해석을 이용하여 확률강우량을 산정한 후, 앞의 2), 3)번 과정에서 계산된 확률강우량과의 절대편의를 지속기간과 재현기간별로 계산한다.
- 5) 각각의 지속기간과 재현기간별로 2)번과 4)번의 결과로 계산된 절대편의(RMSE를 사용한 경우)가 3)번과 4)번의 결과로 계산된 절대편의(RRMSE)보다 적다면 'RMSE'라고 표기한다. 반대의 경우에는 'RRMSE'라고 표기한다.

- 6) 지속기간별로 위의 5)번의 과정을 거쳐서 'RMSE', 'RRMSE'를 표기한 후 두 가지 중 우세한 경우를 해당지속기간의 매개변수로 선정한다.

Table 1은 위의 과정을 거쳐서 구한 대구지점의 결과값으로 지속기간 2시간에서는 RRMSE가 9개이고 RMSE가 3개인 것을 알 수 있다. 즉 지속기간 2시간에서는 RRMSE가 RMSE보다 정확한 값을 가지는 경향이 크다는 것을 의미한다. 또한, 지속기간 3시간에서는 RRMSE로 표기된 것이 6개이고 RMSE로 표기된 것이 7개로 RMSE가 RRMSE보다 더 많은 값을 가지고 있으나, 그 값의 차이가 크지 않으므로 지속기간 3시간에서는 RMSE와 RRMSE가 서로 비슷한 정확도를 가지고 있는 것을 나타낸다. 지속기간 6시간에서는 RMSE가 12개이고 RRMSE 1개로 RMSE가 RRMSE보다 정확한 값을 가지는 경향이 크다는 것을 의미한다.

Table 1에 의하면 지속기간 3시간을 기준으로 3시간보다 짧은 단기간에서는 RRMSE가 좋은 정확도를 나타내고, 지속기간 3시간보다 긴 장기간에서는 RMSE가 좋은 정확도를 나타낸다는 것을 의미한다. 즉, 지속기간 3시간을 장·단기간의 구분점으로 적용하여, 3시간보다 짧은 강우강도에서는 RRMSE가 최소가 되는 매개변수를 사용하고, 지속기간이 3시간보다 긴 강우강도에서는 RMSE가 최소가 되는 매개변수를 사용하는 것이 가장 좋은 정확도를 얻을 수 있다는 의미가 된다.

Table 1. COMBL\_1 : Short and Long Duration Separating Method

Return Period (year)	Duration(hr)				
	1	2	3	6	9
2	RRMSE	RRMSE	RRMSE	RMSE	RRMSE
3	RRMSE	RRMSE	RRMSE	RMSE	RRMSE
5	RRMSE	RRMSE	RRMSE	RMSE	RRMSE
10	RRMSE	RRMSE	RMSE	RMSE	RMSE
20	RRMSE	RRMSE	RMSE	RMSE	RMSE
30	RRMSE	RRMSE	RMSE	RMSE	RMSE
50	RRMSE	RMSE	RMSE	RMSE	RMSE
70	RRMSE	RMSE	RMSE	RMSE	RMSE
80	RRMSE	RMSE	RMSE	RMSE	RMSE
100	RRMSE	RMSE	RMSE	RMSE	RMSE
200	RRMSE	RRMSE	RRMSE	RRMSE	RMSE
300	RRMSE	RRMSE	RRMSE	RMSE	RMSE
500	RRMSE	RRMSE	RRMSE	RMSE	RMSE
Total	13	9	6	1	3
	0	4	7	12	10
	RRMSE	RRMSE	EQUAL	RMSE	RMSE

### 3.2 확률강우량 그래프상의 접점을 기준으로한 장단기간 구분(COMBL\_2)

앞서 설명한 COMBL\_1 방법으로 장·단기간을 구분하는 경우, 확률강우량의 연속성을 고려하지 않고 정확도만을 기준으로 구분하였기 때문에 장기간과 단기간 지속기간을 나누는 경계가 되는 지속기간에서는 장기간과 단기간에 대해서 구한 확률강우량이 서로 다른 경우가 발생한다. COMBL\_2는 이런 단점을 보완하기 위한 방법으로, RMSE와 RRMSE를 각각 최소로 만드는 매개변수를 적용한 강우강도식을 이용하여 확률강우량을 구한 후 이 값들을 그래프로 도시하고, 장기간과 단기간에 해당하는 그래프가 겹치는 접점을 구해서 장·단기간 지속기간의 구분점으로 선택하는 방법이다. Fig. 5는 이런 방법을 통해서 장기간과 단기간에 해당하는 지속기간을 구분한 것으로 그림상의 접점인 지속기간 18시간을 기준으로 장·단기간을 나누게 된다.

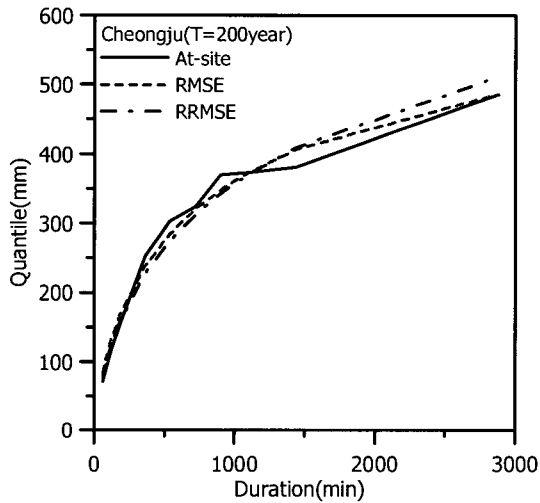


Fig. 5. COMBL\_2 : Short and Long Duration Separating Method

### 3.3 재현기간에 따른 강우강도식의 매개변수추정

본 연구에서 제시한 방법을 통해서 장·단기간의 지

속기간을 구분한 후 구한 확률강우량을 지점빈도해석으로 구한 확률강우량과 비교해 보았다. Fig. 6은 대구, 목포, 서울 그리고 부산지점에 대해서 COMBL\_1 방법을 적용하여 추정된 HEO의 매개변수로 계산된 확률강우량과 지점빈도해석을 통해 계산된 확률강우량을 나타낸 그래프이며, Fig. 7은 COMBL\_2 방법을 적용하여 추정된 HEO의 매개변수로 계산된 확률강우량과 지점빈도해석을 통해 산정된 확률강우량을 나타낸 그래프이다. 장·단기간 구분방법이 서로 다르지만 그림 상에서는 그 차이를 알아볼 수 없을 만큼 서로 비슷한 결과를 나타내는 것을 볼 수 있다.

앞에서 계산된 결과들은 재현기간을 고려하지 않고, 장·단기간만을 나누어 추정한 매개변수를 이용한 결과이다. 재현기간별로 매개변수를 다르게 산정할 경우 보다 높은 정확도를 얻을 수 있지만 매개변수의 개수가 재현기간의 수만큼 증가하게 되어 그 효용성이 떨어지게 된다. 본 연구에서는 이런 단점을 개선하기 위해서 Eqs. (2) and (3)으로 표현된 LEE 강우강도식과 HEO 강우강도식의 매개변수 중에서 최소한의 매개변수만을 재현기간별로 산정하면서도 높은 정확도를 얻는 매개변수를 선정하였다.

Eq. (1)~(3)의 강우강도식 중 UNI의 경우 재현기간별로 모든 매개변수를 추정하였고, LEE의 경우 비교적 영향이 큰  $b$ ,  $n$ 을 재현기간별로 추정하였으며, HEO의 경우 다른 매개변수들에 비해 강우강도 값에 영향을 많이 주는  $b$ ,  $n$ 을 재현기간별로 추정하였다. 이와 같은 방법으로 매개변수를 추정하면 재현기간별로 매개변수를 추정하지 않은 경우와 비교하여 더 높은 정확도를 가지는 강우강도식을 구할 수 있다. 다음의 Fig. 8은 HEO를 이용하여 COMBL\_1를 적용하고 재현기간별로 추정된 매개변수로 계산된 대전 지점의 확률강우량이고, Fig. 9는 HEO를 이용하여 COMBL\_2를 적용하고 재현기간별로 매개변수를 추정하여 계산된 대전 지점의 확률강우량이다. 그림에서 보는바와 같이 Figs. 6 and 7과 비교하여 더 높은 정확도를 보이는 것을 알 수 있다.

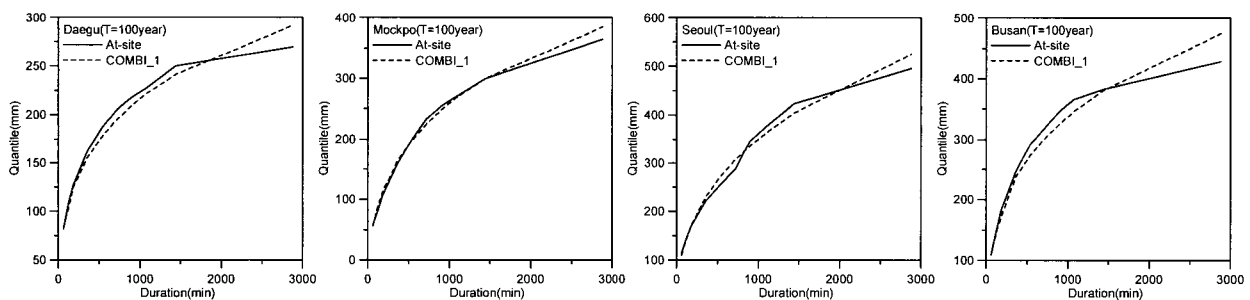


Fig. 6. Quantiles Computed by COMBL\_1 Method

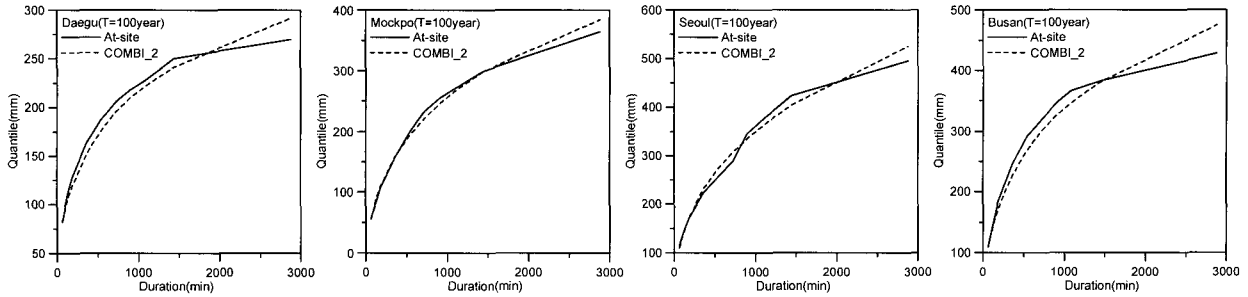


Fig. 7. Quantiles Computed by COMBI\_2 Method

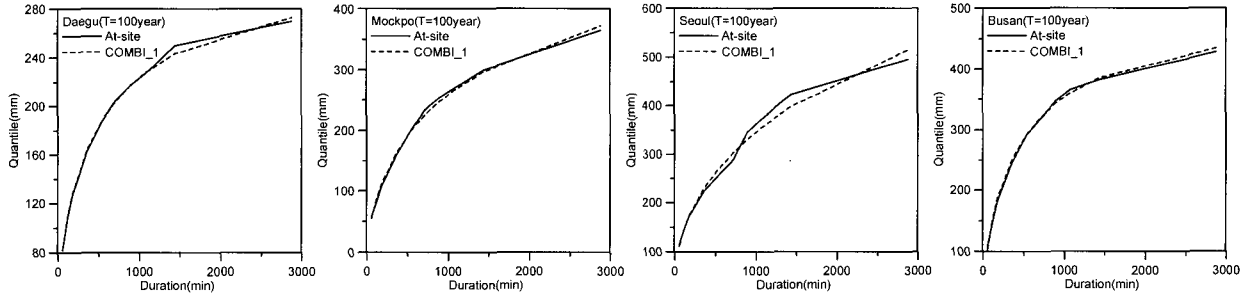


Fig. 8. Quantiles Computed by COMBI\_1 Method. The parameters of IDF curve is estimated for each return period.

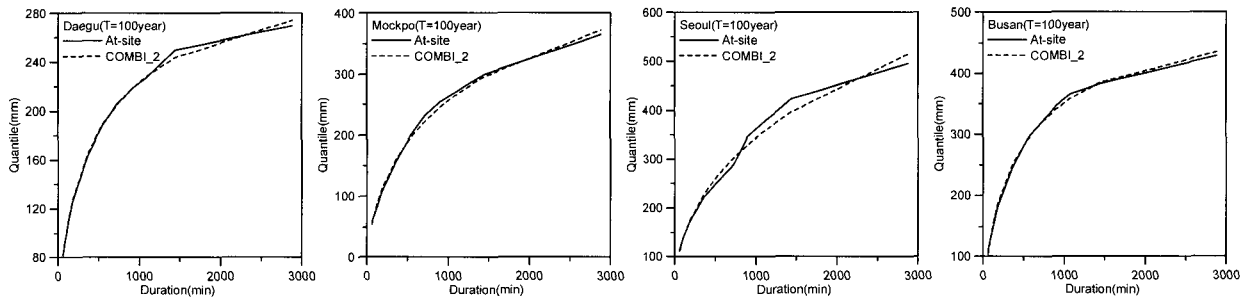


Fig. 9. Quantiles computed by COMBI\_2 method. The parameters of IDF curve is estimated for each return period.

앞 절에서 설명한 도시적인 결과와 더불어, 장·단기간을 고려하면서 동시에 재현기간별로 매개변수를 추정 한 경우와 그렇지 않은 경우에 대해서 정확도를 산정해 보았다. Table 2에서 A~D는 모두 HEO식을 이용한 결과로, A는 재현기간을 고려하지 않고 COMBI\_1을 적용하여 산정된 매개변수를 이용한 결과의 정확도이고, B는 재현기간을 고려하지 않고 COMBI\_2를 적용하여 산

정된 경우를, C는 COMBI\_1을 사용하여 재현기간별로 추정한 경우를, D는 COMBI\_2를 사용하여 재현기간별로 추정한 경우의 정확도를 각각 나타내는 것이다. E는 재현기간을 고려하지 않고 기존의 회귀방법을 통하여 추정된 매개변수의 정확도로, 한국건설기술연구원 (2000)의 매개변수자료를 사용하였다.

Table 2. RMSE, RRMSE, and COD of Each Parameter Estimation Methods

	RMSE	RRMSE	COD
A	17.3500	0.0716	0.9654
B	18.6283	0.0702	0.9600
C	8.2359	0.0359	0.9920
D	9.2683	0.0382	0.9893
E	33.3303	0.1082	0.8647

COD : Coefficient of determination

Table 2에 의하며, A, B, C 그리고 D는 E보다 적용한 모든 기준에서 좋은 정확도를 보이는 것으로 나타났다. 이는 기준에 사용되었던 회귀분석을 통한 매개변수 추정 방법보다는 RMSE와 RRMSE의 두 가지 목적함수를 만족시키는 형태의 다목적 유전자알고리즘을 사용하여 장·단기간을 나누어 매개변수를 추정하는 방법이 보다 정확하다는 것을 나타낸다. 또한, 재현기간을 고려하지 않은 A, B와 재현기간 별로 매개변수를 추정한 C, D를 비교하여 보면 재현기간별로 매개변수를 추정하는 것이, 장·단기간을 나누는 방법과 상관없이, 모두 재현기간을 고려하지 않고 매개변수를 추정하는 방법보다 정확한 매개변수를 추정하는 것으로 나타났다.

#### 4. 결과 분석

본 절에서는 강우강도식의 매개변수를, 다목적 유전자알고리즘을 이용하여 추정하는데 있어서 사용될 수 있는 다양한 방법에 대한 평가를 하기 위하여 1) 장·단기간을 나누는 경우와 나누지 않는 경우에 대한 정확도 비교, 2) UNI, LEE, 그리고 HEO 강우강도식들간의 정확도 비교, 3) 장·단기간 구분 방법으로 사용된 COMBI\_1과 COMBI\_2간의 정확도 비교를 수행하였다.

##### 4.1 장·단기간 구분을 사용한 경우와 단일 매개변수를 사용한 경우의 정확도 평가

지속기간을 장·단기간으로 구분하여 추정된 매개변수를 이용한 경우의 정확도와 장·단기간을 구분하지 않고 추정된 매개변수의 정확도를 비교하였다. Table 3은 22개 대상지점에 대하여 COMBI\_1과 COMBI\_2 그리고 장·단기간을 구분하지 않고 RMSE만을 최소화시키는 단일 목적으로 사용하여 추정된 매개변수(단일 매개변수)의 RMSE와 RRMSE 그리고 결정계수를 구해서 평균한 결과를 비교해 놓은 표이다.

Table 3을 보면 COMBI\_1은 RMSE, RRMSE, 그리고 COD 모두에 있어서 다른 두 가지 경우인 COMBI\_2, Single parameter보다 더 좋은 결과값을 보여주고 있으며, COMBI\_2는 RRMSE 한 가지 통계량에 대해서만 Single parameter보다 좋은 결과값을 보여주고 있다.

Table 3. RMSE, RRMSE, and Coefficient of Determination when Each COMBI\_1, 2, and Single Parameter Set are Used.

	RMSE	RRMSE	COD
COMBI_1	8.2359	0.0359	0.9920
COMBI_2	9.2683	0.0382	0.9893
Single parameter	8.7287	0.0498	0.9910

따라서, Table 3의 결과에 의하면 장·단기간을 나누는 방법으로는 COMBI\_1이 COMBI\_2보다 더 우수한 방법이라고 말할 수 있다.

또한 장·단기간을 나누어 추정된 매개변수의 경우 매우 짧은 지속기간과 매우 긴 지속기간에서의 확률강우량의 정확도가 높은 장점이 있다. Fig. 10은 완도지점의 재현기간 200년에 대한 지속기간 60분~180분까지 확률강우량을 나타낸 그래프이다. 그림에서 COMBI\_1은 COMBI\_1을 적용하여 추정된 매개변수를 이용한 강우강도식으로 구한 확률강우량이고, COMBI\_2는 COMBI\_2를 적용하여 추정된 매개변수의 확률강우량을, At-site는 지점빈도해석을 통해 얻은 확률강우량을, ONE PARA SET은 장·단기간을 구분하지 않고 추정한 매개변수의 확률강우량이다. 그림에서 볼 수 있는 것과 같이 장·단기간을 구분하여 추정한 매개변수를 이용한 결과(COMBI\_1, COMBI\_2)가 장·단기간을 구분하지 않고 추정한 매개변수를 이용한 경우(ONE PARA SET)보다 160분 이하의 지속기간에서 높은 정확도를 보인다.

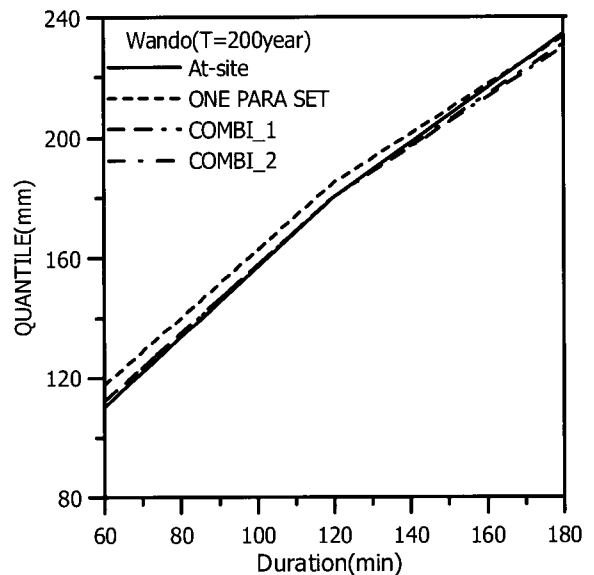


Fig. 10. Quantiles Computed by IDF Curves Whose Parameter is Estimated by Each Different Method.



## 4.2 강우강도식에 따른 장·단기간 구분 방법의 정확도 평가

Eq. (1)~(3)으로 표현된 각각의 강우강도식들을 대상으로 장·단기간을 구분하여 매개변수를 추정하여 그 정확도를 비교해 보았다. Table 4는 22개 대상지점에 대해서 Eq. (1)~(3)의 강우강도식의 매개변수를 COMBI\_1 방법을 이용하여 추정된 후 이를 이용해서 구한 확률강우량과 지점빈도해석을 이용한 확률강우량을 비교한 결과이고, Table 5는 COMBI\_2를 이용하여 같은 방법을 적용한 결과이다. 두 가지 경우 모두 HEO로 표현된 Eq. (3)이 RMSE, RRMSE, 그리고 COD에서 가장 좋은 값을 보여주는 것을 알 수 있다.

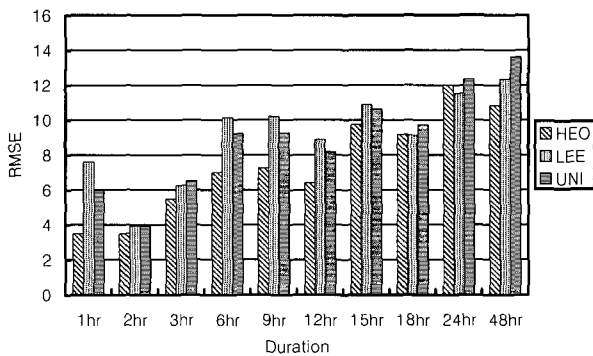


Fig. 11. RMSE computed by IDF curves such as UNI, LEE, and HEO in Eqs. (1)~(3) according to 10 durations

Tables 4, 5는 전체 지속기간에 대한 값을 평균한 것이고, 다음의 Fig. 11은 22개 대상지점에 대하여 COMBI\_1과 COMBI\_2 방법을 각각의 강우강도식에 적

용한 후, 지속기간별로 구한 RMSE값을 평균한 값을 나타내는 그래프이다. 그림에서 살펴보면 지속기간 18시간과 24시간의 경우를 제외하고는 모든 지속기간에 대해서 HEO가 다른 두 개의 강우강도식인 LEE나 UNI보다 더 좋은 정확도를 보여주고 있으며, 전반적으로 다른 두 가지 기법들보다 훨씬 적은 RMSE를 보여주고 있는 것을 알 수 있다.

## 4.3 장·단기간 구분 방법에 따른 정확도 평가

Eq. (1)~(3)의 강우강도식 중에서 비교결과 가장 정확도가 높았던 Eq. (3)을 대상으로 COMBI\_1과 COMBI\_2의 정확도를 비교해 보았다. Table 6은 각 장·단기간 구분 방법의 정확도를 나타낸 표이다.

Table 6을 살펴보면 COMBI\_1을 사용하는 것이 COMBI\_2보다 RMSE, RRMSE, 그리고 COD에 있어서 모두 더 정확한 값을 얻을 수 있는 것으로 나타났다. 그러나 1) 장·단기간을 나누는 경계가 되는 지속기간에서는 COMBI\_1의 경우 서로 다른 값을 갖는 반면에 COMBI\_2는 서로 같은 값을 갖는 차이점이 있고, 2) 각 재현기간별로 구하는 매개변수가 COMBI\_1의 경우 동일한 반면에 COMBI\_2의 경우에는 서로 달라져서 매개변수를 구하기 어렵다는 단점이 있다. 하지만, COMBI\_1으로 구한 강우량 값이 경계가 되는 지속기간에서 달라지기는 하지만 그 차이가 아주 적은 양의 값이기 때문에 전반적인 매개변수 산정의 효율성이나 정확도 면에서는 COMBI\_1 방법을 이용하여 장·단기간을 구분하는 것이 더 좋은 방법이라고 말할 수 있다.

Table 4. Comparison of Parameters' Accuracies Using COMBI\_1

IDF curve	RMSE	RRMSE	COD
HEO	8.2359	0.0359	0.9920
LEE	9.9841	0.0544	0.9874
UNI	9.9783	0.0491	0.9884

Table 5. Comparison of Parameters' Accuracies Using COMBI\_2

IDF curve	RMSE	RRMSE	COD
HEO	9.2682	0.0382	0.9893
LEE	10.9217	0.0582	0.9849
UNI	10.5048	0.0508	0.9869

Table 6. RMSE, RRMSE, and COD Computed by COMBI\_1 and 2

	RMSE	RRMSE	COD
COMBI_1	9.3994	0.0464	0.9893
COMBI_2	10.2316	0.0490	0.9870

## 5. 결 론

본 연구에서는, 강우강도식의 매개변수를 추정하는데 있어서, 다목적 유전자알고리즘의 목적함수로 RMSE와 RRMSE를 적용하여 보다 객관적인 기준으로 장·단기간을 구분하는 방법을 제시하였다. 장·단기간을 구분하는 방법으로는 정확도(COMBL\_1)와 집점(COMBL\_2)을 기준으로 하는 두 개의 구분방법을 적용해 보았으며, 22개 대상지점에 대하여 HEO, LEE 그리고 UNI형태의 강우강도식에 적용한 후 정확도를 비교해 보았다.

재현기간별로 매개변수를 추정하는 방법과 재현기간을 고려하지 않고 매개변수를 추정하는 방법의 적합도를 비교한 결과 최소한의 재현기간만을 고려한 방법을 적용하더라도 높은 정확도를 얻을 수 있음을 보였고, 적용된 HEO, LEE 그리고 UNI형태의 강우강도식 중에서 HEO 형태의 강우강도식이 가장 적용하기 적합한 강우강도식으로 나타났다. COMBL\_1과 COMBL\_2를 비교해 본 결과 COMBL\_1이 장·단기간 구분점에서 매개변수가 연속적이지 않은 단점이 있으나, COMBL\_2에 비해 높은 정확도를 보이는 것으로 나타났으며, COMBL\_2와 달리 모든 재현기간에 대하여 통일된 구분점을 사용하는 장점이 있어 COMBL\_2보다 COMBL\_1이 우리나라에 적용하기 적합한 장·단기간구분 방법으로 나타났다.

향후 RMSE와 RRMSE 외에 다른 여러 목적함수를 적용하여 좀 더 높은 정확도를 보이는 목적함수를 찾는 연구와 함께 COMBL\_1의 장·단기간 구분점에서 강우량이 연속적이지 않은 단점을 보완하는 방법에 대한 연구가 이어져야 할 것이다.

## 감사의 글

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

## 참 고 문 헌

- 국립방재연구소 (2002). **FARD2002 사용자설명서**. 연세대학교.
- 이원환 (1980). "도시 하천 및 하수도 개수계획상의 계획 강우량 설정에 관한 추계학적 해석." **대한토목학회논문집**, 대한토목학회, 제28권, 제4호, pp. 81-94.
- 이원환, 박상덕, 최성열 (1993). "한국대표확률강우강도식의 유도." **대한토목학회논문집**, 대한토목학회, 제13호, 제1호, pp. 115-120.
- 한국건설기술연구원 (2000). **강우자료 추출 및 확률강우량 산정프로그램 개발**.
- 허준행, 김경덕, 한정훈 (1999). "지속기간별 강우자료의 적정분포형 선정을 통한 확률강우강도식의 유도." **한국수자원학회논문집**, 한국수자원학회, 제32권, 제3호, pp. 247-254.
- Deb, K. (2001). *Multi-objective optimization using evolutionary algorithms*. John Wiley & Sons, Chichester.
- 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.
- di Pierro, F., Khu, S. T., and Savi, D. A. (2007). "An investigation on preference order ranking scheme for multiobjective evolutionary optimization." *Evolutionary Computation, IEEE Transactions on*, Vol. 11, No. 1, pp. 17-45.
- Goldberg, D. E. (1989). *Genetic algorithms in search, optimization & machine learning*. Addison Wesley, Reading, Massachusetts.
- Tan, K. C., Khor, E. F., and Lee, T. H. (2005). *Multiobjective evolutionary algorithms and applications*. Springer, London.

(논문번호:07-70/접수:2007.07.23/심사완료:2007.09.14)