



## 통합모델의 강수물리과정 모수 최적화를 위한 알고리즘 비교 연구 : 마이크로 유전알고리즘과 하모니 탐색 알고리즘

An intercomparison study between optimization algorithms for parameter estimation of microphysics in Unified model : Micro-genetic algorithm and Harmony search algorithm

장지연 · 이용희<sup>†</sup> · 주상원

Jiyeon Jang, Yong Hee Lee<sup>†</sup> and Sangwon Joo

기상청 수치모델링센터 수치자료응용과

Numerical Data Application Division, Numerical Modeling Center, KMA

### 요약

기상수치예보모델의 강수물리과정은 강수 발생과 연관된 입자의 낙하속도, 부착 및 자동전환, 입자크기분포 등의 과정을 다룬다. 하지만 수치예보모델의 미세물리과정과 모수에는 상당한 불확실성이 내포되어 있다. 수치예보모델의 불확실성을 줄이기 위하여 일반적으로 모수 추정을 사용한다. 이 연구에서는 모수 추정을 위한 최적화 알고리즘으로 마이크로 유전알고리즘과 하모니탐색 알고리즘을 사용하고 우리나라에서 발생한 강수사례에 대해 통합모델의 강수물리과정에서 사용하는 모수를 최적화하였다. 두 알고리즘의 서로 다른 특성으로 인해 최적화 과정 중의 차이가 보였다. 마이크로 유전알고리즘은 440회 수행 후 약 1,033의 적합도로 수렴하였고 하모니탐색 알고리즘은 60회 수행 후 약 1,031의 적합도로 수렴하였다. 이를 통해 하모니탐색 알고리즘이 마이크로 유전알고리즘보다 더 빨리 최적의 모수를 탐색하는 것을 알 수 있었다. 따라서 계산비용이 방대한 기상수치예보모델의 최적화 문제에서 빠른 시간 내에 최적의 모수를 탐색해야 한다면 하모니 탐색 알고리즘이 더 적합하다는 것을 확인하였다.

키워드: 마이크로 유전알고리즘, 하모니탐색 알고리즘, 최적화, 강수물리과정

### Abstract

The microphysical processes of the numerical weather prediction (NWP) model cover the following : fall speed, accretion, autoconversion, droplet size distribution, etc. However, the microphysical processes and parameters have a significant degree of uncertainty. Parameter estimation was generally used to reduce errors in NWP models associated with uncertainty. In this study, the micro-genetic algorithm and harmony search algorithm were used as an optimization algorithm for estimating parameters. And we estimate parameters of microphysics for the Unified model in the case of precipitation in Korea. The differences which occurred during the optimization process were due to different characteristics of the two algorithms. The micro-genetic algorithm converged to about 1,033 after 440 times. The harmony search algorithm converged to about 1,031 after 60 times. It shows that the harmony search algorithm estimated optimal parameters more quickly than the micro-genetic algorithm. Therefore, if you need to search for the optimal parameter within a faster time in the NWP model optimization problem with large calculation cost, the harmony search algorithm is more suitable.

Key Words : Micro-Genetic Algorithm, Harmony Search Algorithm, Optimization, Microphysics

Received: Nov. 9, 2016

Revised : Feb. 13, 2017

Accepted: Feb. 14, 2017

<sup>†</sup>Corresponding authors  
gonos2004@korea.kr

이 연구는 수치모델링센터 주요사업 '기상업  
무지원기술개발연구(NIMS-2016-3100)'의 일  
환으로 수행되었다.

This is an Open-Access article distributed  
under the terms of the Creative Commons  
Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>)  
which permits unrestricted non-commercial  
use, distribution, and reproduction in any  
medium, provided the original work is  
properly cited.

## 1. 서론

기상수치예보모델의 강수물리과정은 물의 이동에 따른 상변화를 나타내는 과정으로, 기상청 통합모델(Unified Model)에서는 입자의 낙하속도, 입자의 부착 및 자동 전환과정, 입자크기분포, 강수의 증발 등의 과정을 다룬다[1]. 강수물리과정은 미세규모이기 때문에 명시적으로 나타낼 수 없어 모수화가 필요하고 대기 중의 미세물리 관측 자료가 부족하여 불확실성이 크다[2]. 기상수치예보모델의 불확실성을 줄이고 예측성을 높이기 위하여 일반적으로 최적화 기법에 의한 모수추정 방법이 많이 사용되고 있다[3,4,5].

모수 추정에는 다양한 방법이 사용되고 있지만 이 연구에서는 마이크로 유전알고리즘(Micro-genetic algorithm,  $\mu$ -GA)과 하모니 탐색 알고리즘(Harmony search algorithm, HS)을 사용하고자

한다. 두 알고리즘은 모두 메타-휴리스틱(Meta-Heuristic) 계열의 알고리즘으로 다양한 최적화 문제에서 좋은 성능을 보이고 있으며 한정된 시간 내에 주어진 문제를 최적화하기 위해 적합도 함수를 기반으로 적합도가 높은 최적 해를 탐색하게 된다. 기상수치예보모델의 모수추정에는 유전 알고리즘이 많이 사용되어 왔으며 역과정, 대류 모수화 과정[4,5]과 지면과정 모델의 모수를 최적화 한 바 있다[6,7]. 하지만 기상수치예보모델은 불연속적이고 비선형적인 문제를 다루고 있어 복잡한 계산과정을 동반하게 되므로 계산비용이 방대하여 많은 수의 개체를 사용하거나 반복수행이 어렵다. 따라서 초기 수렴속도와 탐색 범위 등을 고려한 효율적인 최적화 방법을 적용할 필요가 있다. HS는  $\mu$ -GA보다 최근에 개발된 알고리즘이며 다양한 최적화 문제를 통해 알고리즘의 이점과 효율성이 입증되었다[8]. HS는 복잡한 계산과정을 필요로 하지 않기 때문에 다양한 유형의 문제를 쉽고 효과적으로 해결하는 장점이 있고, 수학적 문제 해결에 있어 유전알고리즘과 비교하여 우수한 결과를 도출하기도 하였다[9].

따라서 이 연구에서는 통합모델의 강수 물리과정 최적화를 위해 메타-휴리스틱 계열의 알고리즘 중에서 초기 수렴성이 좋은 것으로 알려진  $\mu$ -GA와 HS를 비교하였다. 초기 수렴 속도, 모수에 대한 탐색 특징, 강수예측성능 향상에 미치는 각 모수별 민감도를 분석하여 계산비용이 방대한 기상수치예보모델의 최적화 문제에 효율적으로 사용할 수 있는 방법에 대해 고찰하고자 한다.

## 2. 실험 방법

기상수치예보모델의 일종인 통합모델의 강수물리과정을 최적화하기 위하여  $\mu$ -GA와 HS를 사용하였고, 두 알고리즘의 최적화 과정에서 나타나는 특징을 살펴보았다. 통합모델의 강수물리과정 최적화 실험 과정은 그림 1과 같다. 먼저 2가지 알고리즘에 최적화 하고자 하는 5개의 모수를 적용하고, 추정된 모수를 이용하여

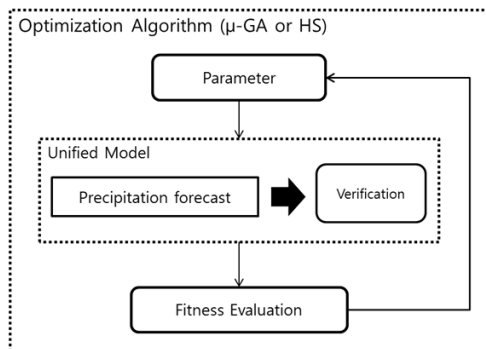


그림 1. 최적화 실험의 순서도.  
Fig. 1. Flow chart of the experiment

통합모델에서 강수량을 예측하면 강수예측속편도를 기반으로 한 적합도 평가를 통하여 최적모수를 반복적으로 탐색하는 구조로 시스템을 구축하였다.

이 절에서는 통합모델의 강수물리 모수화 과정과 최적화하고자 하는 모수, 최적화 알고리즘, 실험 사례, 적합도 함수 구성 방법에 대해 설명하고자 한다.

### 2.1 통합모델의 강수물리과정 모수화

이 연구에서는 ENDGame(Even Newer Dynamics for General Atmospheric Modeling of the Environment) 기반의 통합모델 버전 9.0을 사용하는 국지예보모델을 사용하였다. 국지예보모델의 예측 영역은 그림 2와 같이 극동아시아 지역으로 한반도 지역은 1.5 km 해상도의 고정격자 영역으로 되어 있고, 그 외의 지역은 1.5km ~ 4km 해상도의 가변격자 영역으로 되어 있다.

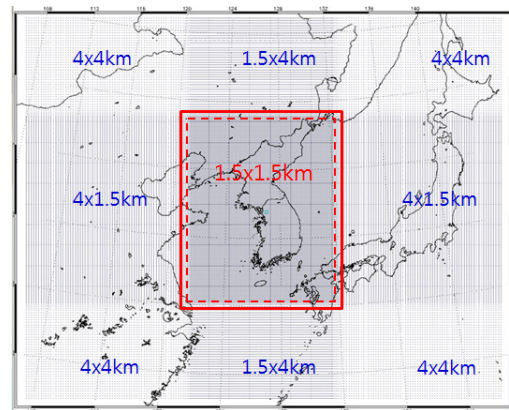


그림 2. 국지예보모델의 예측영역  
Fig. 2. Horizontal Domain of Local Prediction System.

최적화를 위해 통합모델의 강수물리과정 중에서 강우입자 크기분포, 입자의 부착 및 자동 전환 과정과 관련된 5가지 모수에 대해 최적화 실험을 수행하였다.

강우입자크기분포는 3개의 매개변수를 가지는 감마함수 형태로 식 (1)과 같이 표현된다[10]. 감마함수에서  $\alpha = 0$ 을 사용하게 되면 절편인자  $N_0$ 는 식 (2), 기울기 인자  $\lambda$ 는 식 (3)과 같이 해석 해를 가진다. 특히 통합모델에서는  $N_0 - \lambda$ 의 관계식을 식 (2)와 같이  $x_1$ 과  $x_2$ 를 사용하여 모수화하고 있다. 이 연구에서는 강우입자분포에서  $x_1$ 과  $x_2$ 를 최적화를 위한 모수로 선정하였다.

$$N(D) = N_0 D^\alpha \exp(-\lambda D) \quad (1)$$

$$N_0 = x_1 \lambda^{x_2} \quad (2)$$

$$\lambda = \left[ \frac{\pi \rho_w x_1 \Gamma(4 + \alpha)}{6 \rho_{air} q_R} \right]^{\frac{1}{4 + \alpha - x_2}} \quad (3)$$

한편, 입자의 부착과정에 사용되는 구름과 강수의 상관계수( $\rho$ , corr)도 최적화 모수로 선정하였다. 입자의 부착과정  $M_{acc}$ 은 식 (4)로 계산되고, 추가 보정과정을 거치는데 보정인자  $E$ 는 식 (5)와 같다 [11].

$$M_{acc}(q_{cl}, q_R) = 67E(f_{cl}, f_R, \rho)q_{cl}^{1.15}q_R^{1.15} \quad (4)$$

$$E(f_{cl}, f_R, \rho) = (1 + f_{cl}^2)^{-1.15/2} (1 + f_R^2)^{1.15^2/2} \times (1 + f_R^2)^{-1.15/2} (1 + f_{cl}^2)^{1.15^2/2} \times \exp(\rho 1.15^2 \sqrt{\ln(1 + f_{cl}^2) \ln(1 + f_R^2)}) \quad (5)$$

식 (4)의  $q_{cl}$ 은 구름 물 양,  $q_R$ 은 비의 양,  $f_{cl}$ 은 구름 물 양의 표준편차,  $f_R$ 은 비 양의 표준편차,  $\rho$ 는 구름과 강수의 상관계수이다.

경계층에서 지표까지 구름입자의 수능도를 점차 감소시키는 과정에 사용되는 지표입자 수능도(ndrop)와 입자 감소 고도(height)를 최적화 모수로 선택하였다. 여기서, 입자 감소 고도는 구름입자의 수능도가 감소되기 시작하는 고도를 의미한다. 이 과정은 자동 전환과 입자의 침강과정에 사용된다[12].

표 1에는 최적화를 위해 선정된 강수물리과정의 5가지 모수에 대해 기본 모수값과 탐색 범위를 요약하였다.

표 1. 최적화 알고리즘에 적용된 모수 요약  
Table 1. Summary of parameters applied to the optimal algorithm.

Parameter	Min/Default/Max
$x_1$	0.05/0.22/100
$x_2$	1/2, 2/10
Cloud-rain correlation coefficient (corr)	0.75/0.9/1.0
Surface droplet number (ndrop)	2.0e+07/7.5e+07/10.0e+07
Droplet taper height (height)	100/150/500

$x_1$ 과  $x_2$ 는 통합모델에서 현재 0.22와 2.2의 값을 각각 사용하지만 이전 버전에서는 26.2와 1.57을 사용한 바 있다[13]. 두 모수는 강수 유형에 따라 적합한 모수 값의 변동 폭이 크기 때문에 각각 0.05에서 100, 1에서 10으로 범위를 넓게 하여 모수를 탐색하도록 하였다. 구름과 강수의 상관계수(corr)는 현재 0.9의 값을 사용하고 있다. 이 값에서 구름과 강수의 상관관이 너무 낮아지지 않는 범위 내에서 범위를 정하여 0.75에서 1까지로 탐색하였다. 지표입자 수능도(ndrop)는 통합모델 내에서 모수를 이용하여 섭동을 만들 때 적용하는 범위를 사용하였다[14]. 그리고 입자 감소 고도(height)는 현재 150m의 값을 사용하고 있으며 고도를 낮추는 범위로 넓히기 위해 100m부터 시작하고 하층운의 계산에 영향을 주지 않기 위해 500m로 범위를 제한하였다.

## 2.2 최적화 유전알고리즘

### 1) 마이크로 유전알고리즘

다윈의 적자생존(survival of the fittest) 원리에 의한 자연계의 생물 유전학에 기초를 둔 유전알고리즘은 1975년 Holland에 의해 처음 도입되었다[15]. 유전알고리즘은 주어진 문제에 대한 해를 유전적 형태로 표현하여 세대를 거쳐 진화하면서 최적의 해를 탐색한다. 연산자로는 교배, 돌연변이, 재생산 등이 있으며, 좋은 유전자를 다음 세대로 유전하기 위해 염색체 선택에 있어 고도의 무작위 추출기법을 사용한다. 또한 함수의 연속성, 미분가능성 등의 제약을 받지 않아 연속 변수와 불연속 변수들이 섞여있는 문제 해결에 적합한 방법이다[15,16].

하지만 유전알고리즘은 많은 집단을 사용해야 최적의 해를 안정적으로 구할 수 있기 때문에 계산 비용이 크다는 단점이 있다. 이러한 문제를 해결하고자 Krishnakumar는 매우 작은 집단을 효율적으로 다루고 만족스러운 해가 얻어질 때까지 시작과 재시작을 반복하는 마이크로 유전알고리즘( $\mu$ -GA)을 제안하였다.  $\mu$ -GA는 세대 내의 각 개체들이 한 해에 수렴하게 되면 엘리트 개체를 보존하고 새로운 집단을 구성하여 다음 세대에서 재시작하게 된다[17].

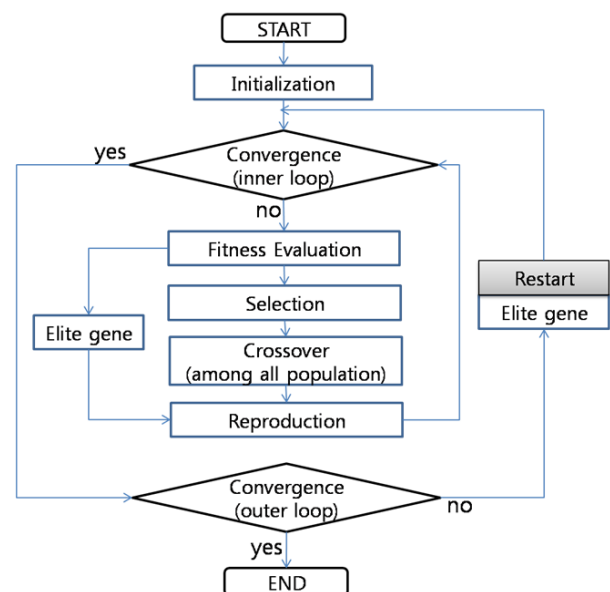


그림 3. 마이크로 유전알고리즘의 순서도[18]  
Fig. 3. Flow chart of micro-Genetic Algorithm ( $\mu$ -GA)[18]

$\mu$ -GA의 순서도는 그림 3과 같으며,  $\mu$ -GA에서는 수렴 여부를 검토하여 내부와 외부 경로 중 수행 할 경로를 정한다. 수렴 조건을 만족하면 외부 경로를 수행하고 외부경로의 종료조건을 만족하지 못하면 재시작을 수행한다. 그리고 수렴조건을 만족하지 못하면 내부 경로를 수행하게 되고, 이때 적합도 평가, 선택, 교배 과정을 수렴 조건이 만족할 때까지 반복하게 되면서 지난 세대의 최적 해인

엘리트 개체는 보존하고 나머지 해에 대해 교배과정을 통해 새로운 집단을 구성하고 다음 세대가 시작된다. 유전알고리즘과 달리 돌연변이 연산은 없지만  $\mu$ -GA는 시작과 재시작을 수행하기 때문에 돌연변이 연산 없이도 다양한 해 집단의 탐색이 가능하다.

## 2) 하모니탐색 알고리즘

HS는 2000년 Geem에 의해 처음 제안되었으며[19], 자연현상이 아닌 음악적으로 화음을 맞추는 과정을 개념화 한 알고리즘이다. 음악을 작곡할 때 가능한 다양한 음계의 조합을 시도하고, 그 화음을 기억하면서 가장 완벽한 화음을 찾게 된다. 이 과정은 공학적인 문제에 대한 최적 해를 찾는 과정과 유사한데, HS는 이 과정을 알고리즘화 한 것이다[20]. HS는 수자원 분산 네트워크의 최적 문제를 해결하기 위해 처음 제안되었으며[19], 최근에는 스토쿠 퍼즐, 댐 관리 등의 여러 최적화 분야에 적용되고 있다. HS는 그림 4와 같은 다음 과정에 따라 최적화를 수행한다.

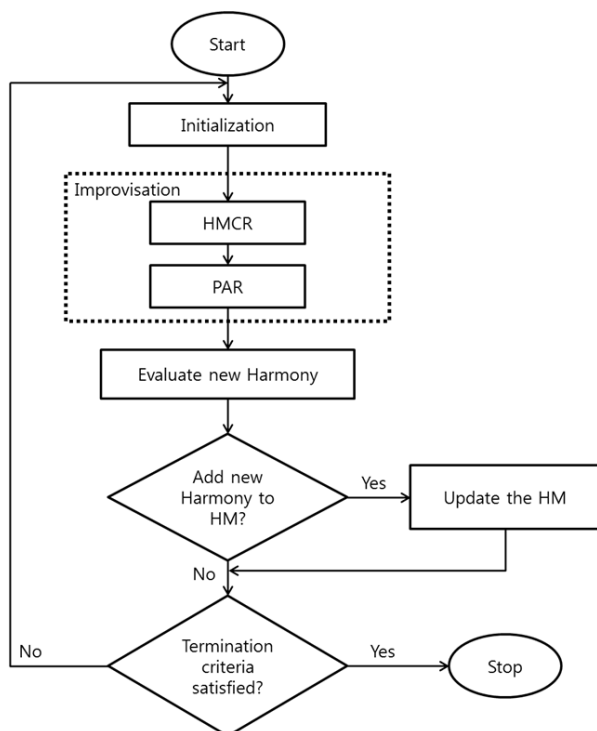


그림 4. 하모니 탐색 알고리즘의 순서도[19]  
Fig. 4. Flowchart of the HS algorithm.[19]

1. Harmony memory 초기화
2. Improvise a new harmony with Randomness  
(무작위로 새로운 하모니 생성)
3. 적합도 함수를 기반으로 높은 적합도의 경우에 대해 harmony memory 업데이트
4. 종료 조건을 만족할 때까지 2,3번 과정 반복

## 2.3 최적화 알고리즘을 이용한 실험 설계

### 1) 마이크로 유전알고리즘 설계

$\mu$ -GA를 수행할 때에는 총 세대 수, 세대별 개체 수, 모수의 개수, 돌연변이 확률, 교배 확률, 교배 방법 등에 대해 설정해야 한다. 이 연구에서  $\mu$ -GA 실험에 사용된 세대는 총 50세대이며, 각 세대 당 개체 수는 12개이다. 각 모수의 탐색범위는 표 1의 최소, 최댓값을 적용하였고 돌연변이 연산은 없도록 하였다. 교배 방법은 single point crossover와 uniform crossover 방법이 있지만 uniform crossover 방법을 사용하여 0.5의 교배율로 설정하였다. uniform crossover는 부모 사이에 고정된 혼합비율을 사용하고 0.5의 교배율은 부모 유전자의 교차점은 무작위로 선택될 수 있지만 자손은 부모의 절반씩의 유전자를 유전 받게 된다는 의미이다.

### 2) 하모니탐색 알고리즘 설계

HS에서 사전설정이 필요한 변수는 Harmony memory Size(HMS), 해 벡터의 수, Harmony Memory Considering Rate(HMCR), Pitch Adjusting Rate(PAR) 등이 있다. 새로운 하모니 벡터는 memory consideration, pitch adjustment, 무작위 선택을 기반으로 발생한다 [19]. 새로운 하모니 벡터 값은  $\mu$ -GA 실험에서와 마찬가지로 표 1의 최소, 최댓값을 적용하여 각 모수의 범위 내에서 선택되도록 하였다. HMCR은 Harmony memory에 저장된 값, 즉 이전에 사용된 값에서 하나의 값을 선택할 확률로 이 연구에서는 0.6으로 할당하였고, 선택할 값에 대해 변화를 주는 확률은 PAR으로써 0.4에서 0.85의 값을 할당하였다. 그리고  $\mu$ -GA와 동일한 횟수로 수행하기 위하여 반복 수행 횟수는 600회로 하였다.

## 2.4 사례 및 적합도 구성

우리나라에서 발생한 호우의 대표 사례를 반영하기 위하여 남부, 영동, 수도권 지역에서 발생한 3가지 강수 사례를 선정하였다. 3가지 사례는 모두 일 누적강수량 100 mm 이상의 여름철 집중호우 사례이며, 첫 번째 사례는 2014년 8월 1일 태풍 나크리(NAKRI)의 영향으로 남부지방에 일 누적강수량 약 400mm가 기록된 사례이다. 두 번째 사례는 2015년 6월 25일 온대저기압의 영향으로 전국적으로 강수가 있었으며, 특히 영동지방에 일 누적강수량 약 200 mm가 기록되었다. 세 번째 사례는 2015년 7월 24일 수도권 지역을 중심으로 일 누적강수량 약 100 mm가 기록된 사례이다. 3가지 사례 모두 집중호우가 발생한 사례이지만, 강수가 집중된 지역이 서로 다른 사례를 선정하여 다양한 강수 유형에 대해 동시에 최적화가 되도록 하였다.

3가지 강수 사례에 대해 통합모델을 이용하여 24시간 강수 예측을 수행하고, 예측한 1시간 누적 강수량에 대해 기상청 자동기상관측망(Automatic Weather Station, AWS) 관측 자료를

이용하여 검증하였다. 이 연구에서 사용한 AWS 관측자료는 사례별로 각각 699, 702, 703개 이다. 분할표를 이용하여 검증하는 강수검증 지수 중 가장 대표적인 BIAS(Bias)와 CSI(Critical Success Index)를 이용하여 적합도 함수를 구성하였다. BIAS와 CSI는 각각 식 (6)과 식 (7)처럼 계산된다. BIAS는 예보와 관측에 대한 각각의 발생 빈도를 비교하여 비율을 구한 것으로 1 미만은 과소모의, 1 초과는 과다모의를 의미한다. CSI는 무강수인 경우에 대해 적중한 경우를 제외한 지수로 1에 가까울수록 완벽한 예보를 의미한다.

$$BIAS = \frac{hits + false}{hits + misses} \quad (6)$$

$$CSI = \frac{hits}{hits + misses + false} \quad (7)$$

적합도 함수는 BIAS와 CSI를 이용하여 식 (8)과 식 (9)와 같이 구성하였다.

$$FITNESS = AVG(\sum_{i=1}^3 \sum_j FIT_{ij} / FIT_{dij}) \quad (8)$$

$$FIT_{ij} = \sum_{i=1}^3 \sum_j (\frac{1}{2(|1 - BIAS_j| + 1)} + CSI_j) \quad (9)$$

$j = 0, 1, 2, \dots, 100$

식 (9)의  $i$ 는 강수 사례,  $j$ 는 강수 임계값을 의미하고  $FIT_{dij}$ 는 통합모델의 기본 값을 사용했을 때의 값을 의미한다. 적합도를 계산할 때, 식 (8)에서 통합모델의 기본 값으로 나누어 줌에 따라 1 이상의 적합도를 가지면 기존의 통합모델 결과(CNTL)보다 개선되었다고 할 수 있다. 구성된 적합도 함수를 이용하여 3가지 강수 사례에 대해 강수량을 예측하고 적합도 평가를 하게 된다.

### 3. 실험 결과 및 분석

선정된 3가지 강수 사례에 대해 동시에 강수예측성능을 개선하도록  $\mu$ -GA와 HS를 각각 적용하여 최적화를 수행하였다. 그림 5는  $\mu$ -GA와 HS의 수행에 따른 반복 횟수별 적합도의 변화를 나타낸 그림이다.

$\mu$ -GA는 48번 반복 수행 후 최대 적합도가 1 이상으로 CNTL보다 개선된 결과를 가진 모수를 찾았다. 반면 HS는 9번째 반복 수행에서 최대 적합도가 1 이상이 되었다. 총 600번 수행 중  $\mu$ -GA는 440번째에서 최대 적합도가 되었고, HS는 60번째 수행에서 최대 적합도가 되어 HS가  $\mu$ -GA보다 초기 수렴속도가 더 빠르다는 것을 알 수 있다. 이러한 원인은 두 알고리즘의 수행 과정의 차이로  $\mu$ -GA는 한 세대 내에서 소수의 개체들에 대한 값이 할당되고 세대

내의 각 개체들에 대한 적합도 판정 후 다음 세대에 적합도 판정 결과가 반영되는 과정으로 수행되기 때문에 HS보다 수렴속도가 느린 것으로 판단된다. 그리고 HS는 탐색 초기에 빠른 속도로 높은 적합도의 값을 탐색하고 그 값에 대한 확인 절차를 지속적으로 수행하였다.

최종 수렴된 결과에서 최적 모수를 탐색하기까지의 시간을 보면 HS가  $\mu$ -GA보다 더 빨랐다.  $\mu$ -GA는 HS보다 속도는 느리나 600번 수행 되는 동안 더 높은 적합도의 모수로 11번 갱신되었다. 이는 최적 값 인근의 값을 집중적으로 탐색하는 지역 탐색의 특성으로 인한 것이며 600번 이상 수행한다면 더 높은 적합도의 모수를 찾을 가능성을 내포한다. HS는 탐색 초기에 빠른 속도로 높은 적합도의 값을 탐색하고 그 값에 대한 확인 절차를 지속적으로 수행하였다. 빠른 시간 내에 높은 적합도의 모수를 탐색하지만, 최대 적합도 탐색 이후의 과정에서는 추가적으로 더 높은 적합도를 탐색하지 못하였다.

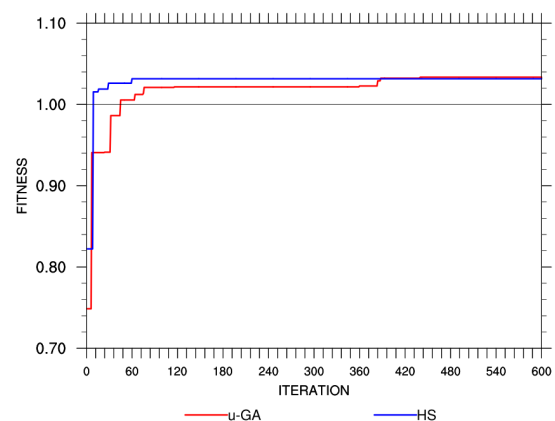


그림 5.  $\mu$ -GA와 HS의 최적화 과정의 적합도 변화.

Fig. 5. The fitness for iteration during the processes of optimization of  $\mu$ -GA and HS.

최적모수가 결정되는 과정을 살펴보기 위해 2가지 알고리즘에서 탐색한 전체모수에 대한 적합도 분포를 그림 6과 그림 7에 각각 나타내었다. 그림 6은  $\mu$ -GA가 탐색한 전체 모수에 대한 적합도 분포이고, 그림 7은 HS를 적용하였을 때의 분포이다. 각 모수간의 민감도 비교를 위해 5개의 모수( $x_1$ ,  $x_2$ , 지표입자 농도(ndrop), 구름과 강수의 상관계수(com), 입자 감소 고도(height)간의 상대 분포를 각각 나타내었다. 또한 그림에서 색은 적합도를 나타내며 빨간색으로 표시될수록 적합도가 높음을 의미한다. 그리고 ★는 통합모델에서 현재 사용하고 있는 기본 모수 값을 나타낸다.

그림 6의  $\mu$ -GA에서 탐색한 모수는 적합도가 높은 모수값 주변으로 탐색을 시작하여 점점 하나의 값으로 수렴하는 형태의 분포이고 적합도가 높은 모수 값 주변으로 모수들이 밀집되어 있는 것을 볼 수 있다. 이는 교배만 사용할 경우 염색체의 탐색범위가

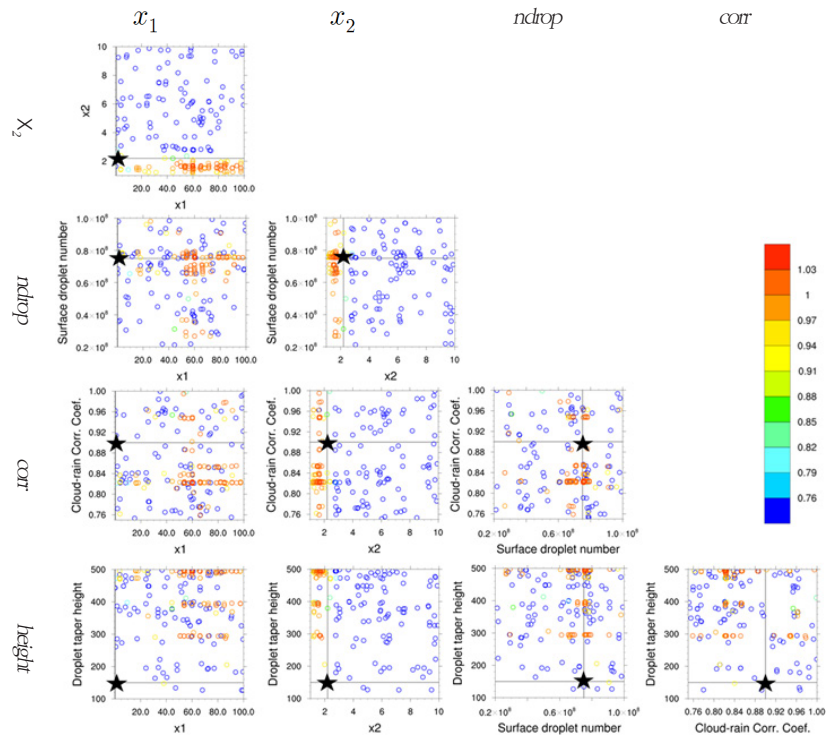


그림 6. 5가지 모수별 적합도 분포 ( $\mu$ -GA)  
Fig. 6. The distribution of fitness according to five parameters ( $\mu$ -GA)

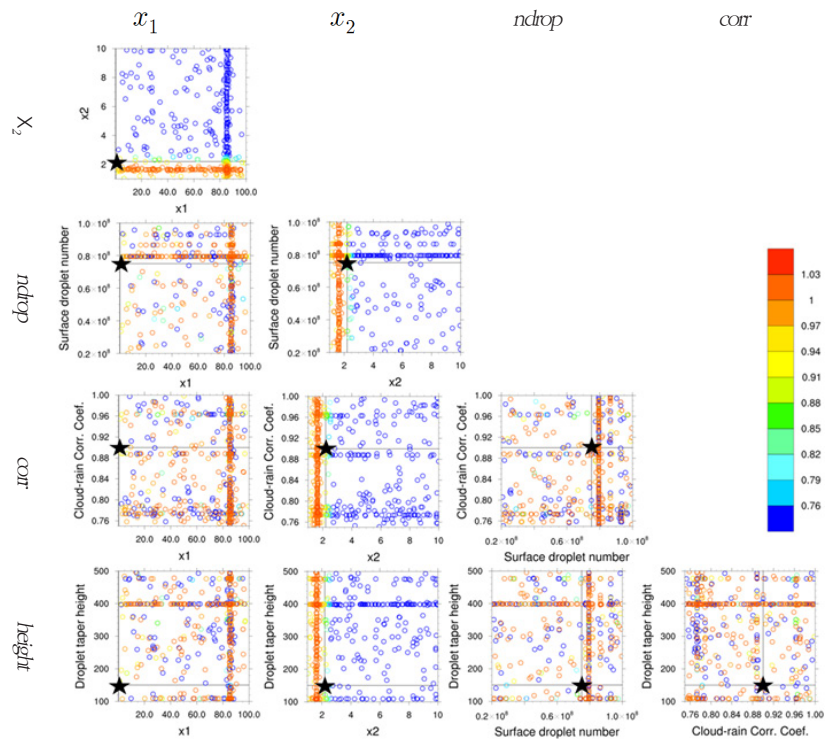


그림 7. 5가지 모수별 적합도 분포 (HS)  
Fig. 7. The distribution of fitness according to five parameters (HS)

엘리트 개체 주위로 빨리 수렴하기 때문에  $\mu$ -GA의 지역탐색 특성이 반영된 결과이다.

그림 7의 HS에서 탐색한 모수 분포를 보면 하나의 일정한 값 부근에서 지속적으로 반복 수행하는 것을 확인할 수 있다. 이는 수행과정에서 적합도가 높은 모수 값을 기억하여 그 값에 대해 조금씩 변화를 주면서 반복적으로 수행하였기 때문이다. 다섯 모수 중 corr을 제외하고 모두 하나의 값 부근에서 반복 수행하는 것을 볼 수 있다.

모수 간의 적합도를 비교해보면, 두 알고리즘에서 동일하게  $x_2$ 는 다른 모수의 값과 상관없이 약 1.7일 때 항상 적합도가 높게 나타나는 것을 볼 수 있다. 이를 통해  $x_2$ 가 강수예측성능에 가장 크게 영향을 주는 것을 알 수 있다.

$\mu$ -GA와 HS는 같은 강수 사례와 모수에 대해 탐색하였지만 탐색한 모수 분포가 달랐고 최종 탐색한 모수 값에서도 차이가 보였다. 표 2에는  $\mu$ -GA와 HS가 최종 탐색한 5가지 모수 값의 결과를 요약하였다.  $x_1$ 은  $\mu$ -GA에서 60.02, HS는 85.09로 HS가 더 높은 값으로 탐색하였다. 반면에  $x_2$ 는  $\mu$ -GA와 HS에서 모두 1.68과 1.64로 비슷한 값을 탐색하였다. corr은 각각 0.82와 0.78, ndrop은 각각  $754.8 \times 10^5$ ,  $612.6 \times 10^5$ , height는 494.34m, 398.63m로 나타났다. corr, ndrop, height는  $\mu$ -GA가 HS보다 다소 높은 값으로 탐색하였다.

그림 6과 그림 7에서 탐색한 모수의 분포와 표 2의 최종 수렴된 모수 값을 살펴보면 5가지 모수 중에서 통합모델의 강수물리과정에서 강수예측성능에 가장 민감한 모수 값은  $x_1$ 인 것을 알 수 있다.

표 2. 최적화된 모수 요약  
Table 2. Summary of optimized parameters

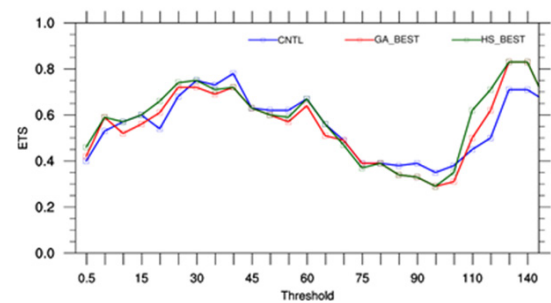
Parameter	$\mu$ -GA	HS
$x_1$	60.02	85.09
$x_2$	1.68	1.64
Cloud-rain correlation coefficient (corr)	0.82	0.78
Surface droplet number (ndrop)	$754.8 \times 10^5$	$612.6 \times 10^5$
Droplet taper height(m) (height)	494.34	398.63

두 알고리즘을 통해 최적화된 모수를 이용하여 강수 사례별로 수행한 결과를 ETS(Equitable Threat Score)로 검증하고 그림 8에 나타내었다. ETS는 식 (10)로 계산되며 이는 우연히 맞추게 되는 경우를 고려한 지수이고 -0.33에서 1까지의 값을 가진다. 1에 가까울수록 좋은 예측성을 가진 것을 의미한다.

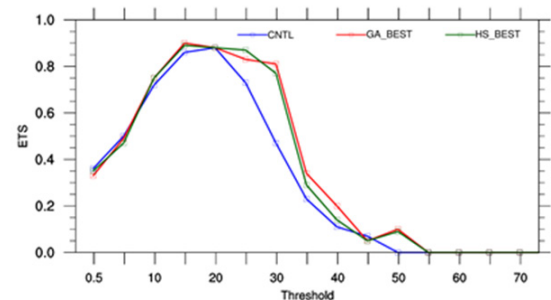
$$ETS = \frac{hits - hits_{random}}{hits + misses + false - hits_{random}} \quad (10)$$

$$hits_{random} = \frac{(hits + misses)(hits + false)}{Total}$$

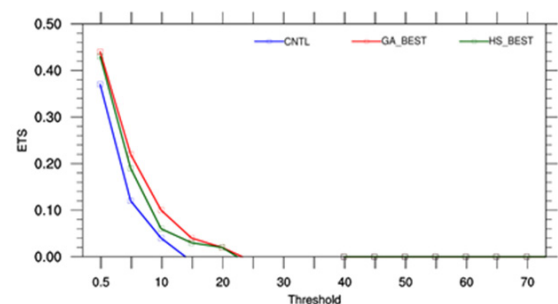
그림 8은 각각 6시간 누적강수에 대한 임계값별 검증 결과이며 (a)는 남부지방 강우 사례의 12시간 예측자료로 2014년 8월 2일 0000 UTC의 결과, (b)는 영동지방 강우 사례의 18시간 예측으로 2015년 6월 26일 0600 UTC의 결과, (c)는 수도권 강우 사례의 24시간 예측으로 2015년 7월 25일 1200 UTC의 결과이다.



(a) 2014. 8. 2. 0000 UTC



(b) 2015. 6. 26. 0600 UTC



(c) 2015. 7. 25. 1200 UTC

그림 8. 최적화된 모수에 의해 예측된 강수의 임계값별 ETS 검증 결과  
(a) 사례1, (b) 사례2, (c) 사례3. 6시간 누적강수에 대한 검증 결과이며, 파란선은 CNTL, 빨간선은  $\mu$ -GA, 초록선은 HS를 의미함.

Fig. 8. The ETSs of precipitation forecasts in terms of various thresholds (in mm). (a) case1, (b) case2, (c) case3. Lines indicate 6h accumulated precipitation forecasts for experiments CNTL(blue),  $\mu$ -GA(red), HS(green)

그림 8(a)의 남부지방 강우 사례는 약 90 mm 이상의 강수에서 두 알고리즘 모두 CNIL보다 약 ETS 0.05의 값이 감소하였지만 약 120 mm 이상의 강수에서는  $\mu$ -GA는 ETS가 약 0.1 증가하였고, HS는 약 0.15 증가하였다. 100 mm 이상의 강한 강수에서 CNIL보다  $\mu$ -GA와 HS가 예측성을 개선하는 것을 볼 수 있었다. 또한 HS의 예측성 검증 결과가  $\mu$ -GA보다 더 높은 것을 볼 수 있었다. 그림 8(b)의 영동지방 강우 사례는 20 mm 이상의 강수에서  $\mu$ -GA와 HS가 CNIL 강수 예측성을 개선하는 것을 확인할 수 있었다. 특히 약 30 mm 이상의 강수에서 두 알고리즘 모두 CNIL보다 ETS 값이 약 0.3 증가한 것을 볼 수 있다. 그림 8(c)의 수도권 강우 사례도  $\mu$ -GA와 HS가 CNIL보다 ETS 값이 약 0.05 상승하면서 강수 예측성이 개선되는 것을 확인하였다.

사례별 최적화된 결과를 임계값 별로 정리해보면 남부지방 사례에서는 110 mm 이상의 강한 강수에 대한 예측을 획기적으로 개선하였고, 영동지방의 사례는 25 mm 이상의 강한 강수에 대해서 개선하였다. 수도권 사례에서는 전체 임계값에 대해 개선하였다. 남부지방 사례에 대해서는 HS가  $\mu$ -GA 보다 높은 강수예측 성능을 보였으나, 영동지방과 수도권 사례에서는  $\mu$ -GA가 HS보다 다소 높은 강수 예측성능을 보였다.

#### 4. 요약 및 결론

이 연구에서는 통합모델 강수물리과정의 효율적인 최적화를 위해  $\mu$ -GA와 HS의 탐색 특성을 비교하였다. 우리나라 여름철에 발생한 대표적인 3가지 강수 사례를 이용하고 통합모델 강수물리과정에 사용되는 5가지 모수에 대해 최적화 실험을 수행하였다.

$\mu$ -GA를 적용한 경우 440회까지 수행 한 후 약 1.033의 적합도 값에 수렴하였고, HS를 적용한 경우 60번 반복 수행 후 약 1.031의 적합도 값에 수렴하였다. 초기 수렴성에 대한 계산 비용 측면에서 보았을 때 HS가  $\mu$ -GA보다 더 효율적인 결과를 보여주었다. 각 알고리즘별 모수 탐색 과정을 살펴보면  $\mu$ -GA는 지역탐색 특성이 반영되어 재시작을 반복하며 높은 적합도의 모수 부근으로 수렴하여 추가적으로 반복한다면 더 높은 적합도의 모수로 수렴할 가능성을 보여주었다. HS는 적합도가 높은 모수로 빠르게 수렴한 뒤 그 값에 대한 확인 작업을 반복적으로 수행하였다.

또한, 최적화된 다섯 모수 중  $\alpha$ 는 다른 모수 값에 상관없이 약 1.7에서 항상 적합도가 높다는 것을 확인하였고, 두 알고리즘에서도 상당히 유사한 값을 산출하였다. 이러한 결과는 강수 예측성에  $\alpha$ 가 가장 민감한 모수임을 보여주었다. 그리고 최적화 된 값을 적용하여 강수 예측성을 검증해 보았을 때, 두 알고리즘 모두 기존의 결과를 개선하였고 특히 강한 강수에 대해 개선되는 정도가 컸다.

이 연구를 통해 기상수치모델의 물리과정 최적화 문제에 대한 두 알고리즘의 특성을 확인하였다. HS가  $\mu$ -GA 보다 초기 수렴속도가 빠른 것을 확인하였다. 그리고  $\mu$ -GA는 수행 기간 동안 지속적으로 더 높은 적합도로 수렴되는 양상을 보여 추가적으로 더 높은 적합도의 모수로 수렴할 가능성을 보여주었다. 이러한 두 알고리즘의 탐색 특성을 이해하고 장점을 활용하여 상황에 따라 알고리즘을 적절히 선택한다면 더 적합한 최적화를 수행할 수 있을 것이다.

#### References

- [1] J. Wilkinson, R. Forbes, D. Wilson, *The large-scale precipitation parameterization scheme*, Unified model documentation paper 026, 2015.
- [2] M. Tong, and M. Xye, "Simultaneous estimation of microphysical parameters and atmospheric state with simulated radar data and ensemble square root kalman filter, Part I: Sensitivity analysis and parameter identifiability", *Monthly weather review*, vol. 136, pp. 1630-1648, 2008.
- [3] H. Moradkhani, S. Sorooshian, H. V. Gupta, P. R. Houser, "Dual state-parameter estimation of hydrological models using ensemble kalman filter", *Advances in water resources*, vol. 28, pp. 135- 147, 2005.
- [4] Y. H. Lee, S. K. Park, and D. E. Chang, "Parameter estimation using the genetic algorithm and its impact on quantitative precipitation forecast", *Annales Geophysicae*, vol. 24, pp. 3185- 3189, 2006.
- [5] X. Yu, S. K. Park, Y. H. Lee, and Y. S. Choi, "Quantitative precipitation forecast of a tropical cyclone through optimal parameter estimation in a convective parameterization", *SOLA*, vol. 9, pp. 36-39, 2013.
- [6] S. Hong, X. Yu, S. K. Park, Y. S. Choi, and B. Myoung, "Assessing optimal set of implemented physical parameterization schemes in a multi-physics land surface model using genetic algorithm", *Geoscientific Model Development*, vol. 7, pp. 2517-2529, 2014.
- [7] S. Hong, S. K. Park, and X. Yu, "Scheme-Based Optimization of land surface model using a micro-genetic algorithm: Assessment of its performance and usability for regional applications", *SOLA*, vol. 11, pp. 129-133, 2015.

- [8] X. S. Yang, "Harmony Search as Metaheuristic Algorithm", *Studies in Computational Intelligence, Springer Berlin*, vol. 191, pp. 1-14, 2009.
- [9] C. Peraza, F. Valdez, and O. castillo, "A Harmony search algorithm comparison with genetic algorithms", *Fuzzy logic augmentation of nature-inspired optimization metaheuristics*, vol. 574, pp. 105-123, 2015.
- [10] S. J. Abel, and I. A. Boutle, "An improved representation of the raindrop size distribution for single-moment microphysics schemes", *Quarterly Journal of the Royal Meteorological Society*, vol. 138, pp. 2151-2162, 2012.
- [11] I. A. Boutle, S. J. Abel, P. G. Hill and C. J. Morcrette, "Spatial variability of liquid cloud and rain : observations and microphysical effects", *Quarterly Journal of the Royal Meteorological Society*, vol. 140, pp. 583-594, 2014.
- [12] J. M. Wilkinson, A. N. F. Porson, F. J. Bornemann, M. Weeks, P. R. Field and A. P. Lock, "Improved microphysical parameterization of drizzle and fog for operational forecasting using the Met Office Unified Model", *Quarterly Journal of the Royal Meteorological Society*, vol. 139, pp. 488-500, 2013.
- [13] J. Wilkinson, R. Forbes, D. Wilson, *The large-scale precipitation parameterization scheme*, Unified model documentation paper 026, 2011.
- [14] S. Beare, W. Tennant, and C. Sanchez, *Stochastic Physics Code in the UM*, Unified model documentation paper 081, 2015.
- [15] J. Holland, *Adaptation in natural and artificial systems*, University of Michigan Press, 1975.
- [16] D. E. Goldberg, *Genetic algorithms in search, optimization and machine learning*, Addison-Wesley, 1989.
- [17] K. Krishnakumar, "Micro-genetic Algorithm for Stationary and Non-stationary Function Optimization", *Intelligent Control and Adaptive Systems*, Vol. 1196, pp. 282-296, 1989.
- [18] Y. H. Lee, J. E. Nam, S. W. Joo, "Optimization of Z-R relationship in the summer of 2014 using a micro genetic algorithm", *JKIS*, vol. 26, no. 1, pp. 1-8, 2016.
- [19] Z. W. Geem, "Optimal cost design of water distribution networks using harmony search", *Dissertation, Korea University*, 2000.
- [20] X. Wang, X. Z. Gao, K. Zenger *An introduction to harmony search optimization method*, Springer Briefs in Computation Intelligence, 2015.

## 저자 소개



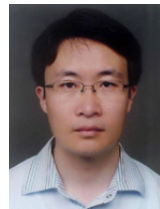
### 장지연(Jiyeon Jang)

2014년 : 부경대학교 환경대기과학과 이학석사  
2014~2016년 : 국립기상과학원 연구원  
2017년~현재 : 수치모델링센터 연구원

관심분야 : 기상수치예보, 물리과정, 최적화

Phone : +82-02-2181-0562

E-mail : jyjang415@korea.kr



### 이용희(Yong Hee Lee)

1997년 : 경북대학교 천문대기과학과 이학석사  
1998~2007년 : 기상연구소 기상연구사  
2007~2016년 : 국립기상과학원 기상연구관  
2017년~현재 : 수치모델링센터 기상연구관

관심분야 : 최적화, 기상수치예보, 딥-러닝

Phone : +82-02-2181-0542

E-mail : gonos2004@korea.kr



### 주상원(Sangwon Joo)

1990년 : 서울대학교 대기과학과 이학석사  
2002년 : 서울대학교 대기과학과 이학박사  
1992~2009년 : 기상청 위성담당관실  
수치예보과 기상연구사/연구관  
2009~2016년 : 수치예보개발과장  
수치자료응용과장  
2017년~현재 : 수치모델링센터 수치자료응용과장

관심분야 : 위성자료동화, 관측영향평가, 수치예보 해석

Phone : +82-02-2181-0541

E-mail : swjoo@korea.kr