

# 인구구조 공간분포 특성에 관한 유전자 알고리즘 적용방안 The Applicability of the Genetic Algorithm on Spatial Distribution of Demographic Characteristics

최내영\* · 이경윤\*\*  
Choei, Nae Young · Lee, Kyung Yoon

### 要 旨

본고에서 적용한 유전자 알고리즘(GenAlg)은 다윈의 유전적 진화이론에 기반한 적정해 탐색기법으로써 공간연구에 있어서도 다양한 응용연구가 선행된 바 있다. 본 연구는 최근 신도시 개발과 함께 택지개발사업 등이 활발히 진행되고 있는 화성시 동부권을 사례대상지로 GIS 툴을 이용하여 격자인구자료를 구축하고 자연환경변수와 함께 개발 촉진형 및 규제형 용도지역지구 지정현황 등 주요 인구동태 결정요인에 관한 공간정보를 구축한 다음 GenAlg 모형을 추정하고 이를 동일 설명변수를 사용한 회귀분석 결과와 비교해 봄으로써 GenAlg기법의 유용성을 정량적으로 비교분석해 보았다. 분석결과 회귀분석과 유사하게 설명변수들에 대해 비교적 일관되고 설득력 있는 계수 값들이 도출되었고 이로써 GenAlg모형은 적자생존의 유전학적 원리를 기반으로 도시현상과 관련한 제반 문제의 적정해를 찾는 유용한 학제적 탐색기법임을 확인할 수 있었다.

핵심어 : 유전자 알고리즘, 인구표면모형, 격자인구자료, 한국토지정보체계, 건축행정시스템

### Abstract

The Genetic Algorithm is one of the population surface modelling tool in the field of urban and environmental research based on the gridded population data. Taking the East-Hwasung area as the case, this study first builds a gridded population data based on the GIS databases as well as municipal population survey data. The study then constructs the attribute values of the explanatory variables by way of GIS tools. The regression model constructed with the same variables is also run as a comparative purpose at the same time. It is shown that the GenAlg output predicted as much consistent and meaningful coefficient estimates for the explanatory variables as the regression model, indicating that it is a very useful interdisciplinary research tool to find optimal solutions in urban problems.

Keywords : Genetic Algorithm, Population Surface Model, Gridded Population Data, KLIS, eAIS

## 1. 서론

인구표면모형(population surface model)은 비정형적인 인구조사구역 단위로 집계된 인구를 세부 격자에 수치적으로 공간 내삽(areal interpolation)하는 방식을 지칭하는 것으로서 특정 위치에서의 인구구조를 보다 정확히 반영하게 됨으로써 기존의 행정구역 단위 혹은 특정 조사구역단위 등으로 집계된 인구자료의 공간적 제약을 극복하는 데 있어 매우 유용한 대안 중 하나이다(Yue 등 2003, 2008). 본 연구에서 적용하는 유전자 알고리즘(Genetic Algorithm: 이하 GenAlg라 함)은

진화적 알고리즘(Evolutionary Algorithm) 관련 접근 방식 중 하나로 Holland(1975)가 처음 원천 알고리즘을 개발하였고, 이후 Koza(1990)가 개발한 유전적 프로그래밍(Genetic Programming)기법과 함께 신경망의 학습, VLSI 회로 설계, 그래프 분할 문제 등 다양한 응용분야에서 최적해 탐색기법으로 활용되어 왔다. 또한 최근에는 도시 및 환경관련 분야에서도 많은 실험적 연구가 시도된 바 있고 특히 시설 적정입지 탐색 및 앞서 언급한 인구표면모형 연구 등 도시공간 연구에서도 확장 적용되고 있다(Mutti과 Lee, 2005). 이렇듯 GenAlg기법이 공간탐색 연구에 다양하게 적용된

2010년 7월 26일 접수, 2010년 8월 23일 채택

\* 교신저자 · 정희원 · 홍익대학교 공과대학 건설도시공학부 도시공학전공 부교수(nychoei@hongik.ac.kr)

\*\* 홍익대학교 대학원 도시계획학과 석사과정(yumiyoshi11@hotmail.com)

이유는 GenAlg가 복잡한 비선형 탐색공간(complex non-linear search space)에서 경험적 자료에 기반한 확률적(stochastic) 검색방법을 통해 적정해를 자동 탐색하는 알고리즘이기 때문에 별도의 사전 모형을 설정할 필요가 없다는 실용적 장점이 있기 때문이다(Li와 Yeh, 2005). 본 연구는 최근 동탄신도시가 성공적으로 조성된 화성시 동부권 일대를 대상으로 하여 읍·면·동의 행정구역 단위로 조사된 관내 인구데이터를 격자형 인구표면모형 자료로 변형한 후 인구자료에 기초한 GenAlg모형을 적용해 봄으로써 주요 도시환경변수들의 영향력을 추정하는데 있어 GenAlg기법의 유용성을 검증해 보고자 하는 데 그 연구 목적이 있다.

2 연구의 방법 및 공간적 범위

2.1 연구의 방법

본 연구는 이를 위해 특히 대상지 일원의 a) 표고나 경사도 등 토지이용을 결정하는 자연환경 요소와 함께, b) 개발을 촉진하거나 규제하는 용도지역지구 지정현황 등 인구동태에 영향을 미칠 것으로 예상되는 다수 도시환경 변수들을 채택하고 그 변수값들을 ArcGIS v9.3을 이용하여 측정, 대입함으로써 GenAlg 기법의 설명력을 검증해 보고자 한다. 또한 객관적 비교검증을 위해 동일한 변수를 투입하여 GeoDa v0.9.8 프로그램을 이용한 공간회귀분석(Spatial Regression Analysis)을 동시 수행해 봄으로써 각각의 결과치를 정량적으로 비교검토해 보고자 한다.

2.2 연구대상지 개관

화성시는 광범위한 관내 구역을 행정편의 상 동부권, 서부권, 남부권 등 크게 세 권역으로 나누어 관리하고 있다. 본 연구는 그중 최근 수도권 내 성공적인 2기 신도시로 평가되는 동탄신도시가 조성되면서 그에 따라 인접 택지개발사업지구 조성이 활발히 진행되어 인구증가가 크게 일어나고 있는 화성시 동부권을 사례지로 선택하였다. 그림 1은 화성시 동부권의 지형 및 표고에 관한 지형공간정보와 함께 행정구역 및 인구밀도를 나타내고 있다. 그림 2는 개발진흥구역 및 택지개발예정지구 등 개발 촉진형 용도지역지구와 계획관리지역 및 보전녹지지역 등 개발 규제형 용도지역지구의 지정현황을 보이고 있다.

2.3 분석을 위한 원천자료 개관

본 연구에서는 격자형 인구자료 구축을 위해 한국토지정보체계(KLIS : Korea Land Information System)를 활용하였다. KLIS는 인구격자 분할에 필요한 필지정보뿐만 아니라 지형정보, 용도지역지구 정보 등 주요 도시계획 공간정보가 통합 구축된 종합 토지정보체계로서 본 분석에서 사례지에 대한 자연환경 및 용도지역지구 지정현황 등의 설명변수 속성값 구축을 위한 원천자료로도 활용하였다. 이와 함께 건축물정보 추출을 위해 건축행정시스템(eAIS : electronic Architectural Information System)도 사용하였는데, eAIS는 기존의 개별 건축물대장에 나타나 있는 건축연면적, 건폐율, 용적률, 층수, 건축물용도 등 공부자료를 전산화하여 구축한 비공간 속성DB이다. 특히 KLIS의 필지 지번

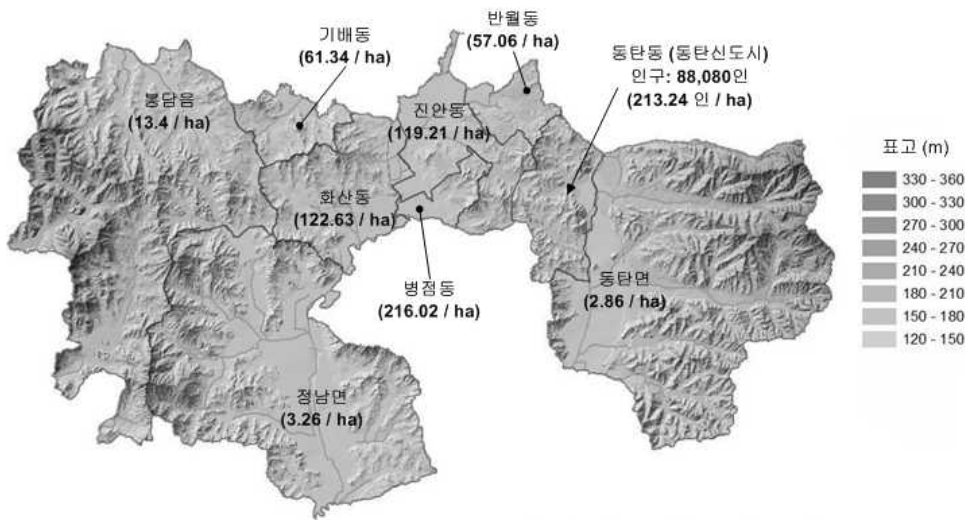


그림 1. 화성시 동부권 지형, 표고, 행정구역 및 인구밀도의 지형공간정보

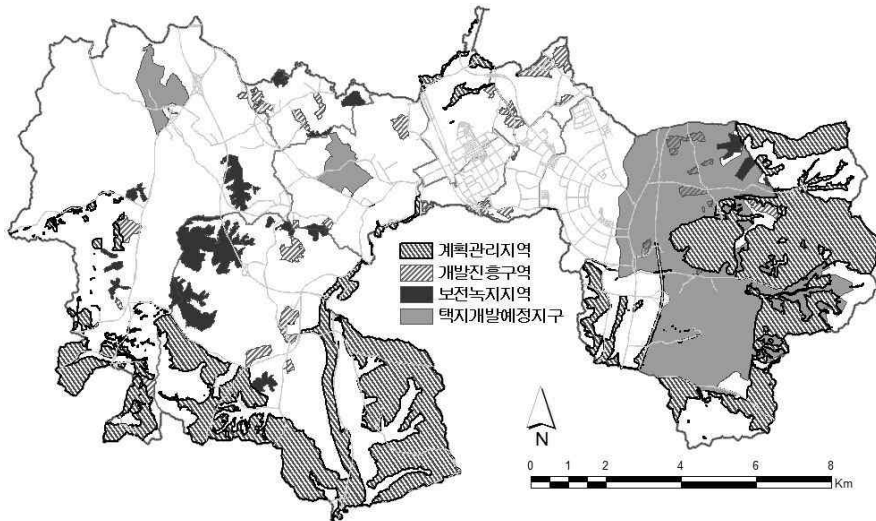


그림 2. 화성시 동부권 개발촉진형 및 개발억제형 용도지역지구 지정현황

과 eAIS의 대지 지번이 동일한 PNU(Parcel Numbering Unit) 코드체계로 작성되어 있어 두 데이터베이스를 PNU 코드를 매개로 결합(join)시킴으로써 KLIS의 토지대장 정보와 eAIS의 건축대장 정보를 개별 필지단위로 통합 연동하여 사용하였다. 또한 사례지자체인 화성시로부터 민원협조를 얻어 2008년도 관내 읍·면·동·리 단위로 자체 실사를 통해 구축한 인구조사자료를 확보하여 사용하였다.

### 3. 인구자료 구축

격자인구자료(grided population)를 구축하기 위해서는 우선 격자의 규모를 설정할 필요가 있다. 격자 크

기는 연구목적에 따라 임의로 설정하게 되는데 기존 격자공간 분석사례를 보면 인구추정의 경우 중국의 성(省 : 우리나라의 도 단위에 해당) 단위 인구추정의 경우 1×1km(Yue 등 2003)로부터 우리나라 시·군 단위 연구의 경우 200×200m(국토해양부, 2008a, 2008b)를 적용한 경우가 있다. 본 연구에서는 준 도시급 공간분석의 경우로 편의상 200×200m 규모의 격자를 선택하였으며 이 경우 화성시 동부권 전체 격자시스템은 총 4,207개 셀로 분할된다.

설정된 격자 시스템에 대해 각 셀 당 인구를 추계하는 과정은 다음과 같다. 1) 개별 동 및 리 각각의 1인당 평균 주거점유면적을 계산한다. 이는 eAIS로부터 각 행정구역별 주거건축물 연면적을 집계하여 인구조

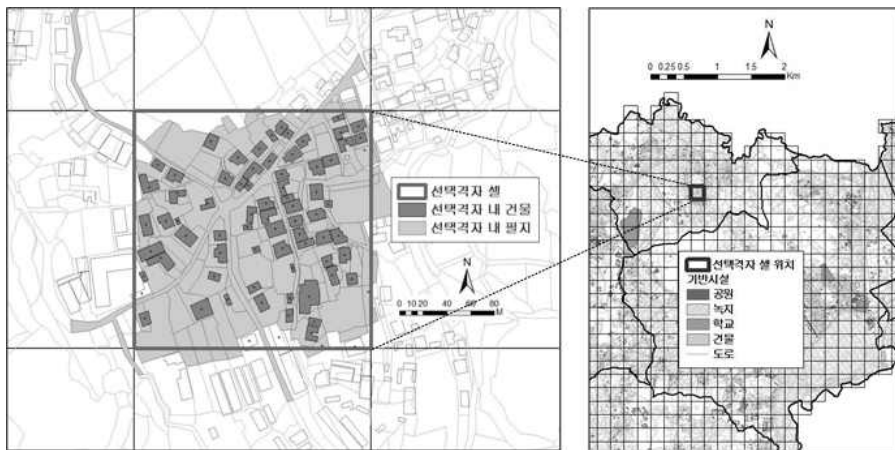


그림 3. 특정 Cell의 귀속 필지 선택결과 예시도

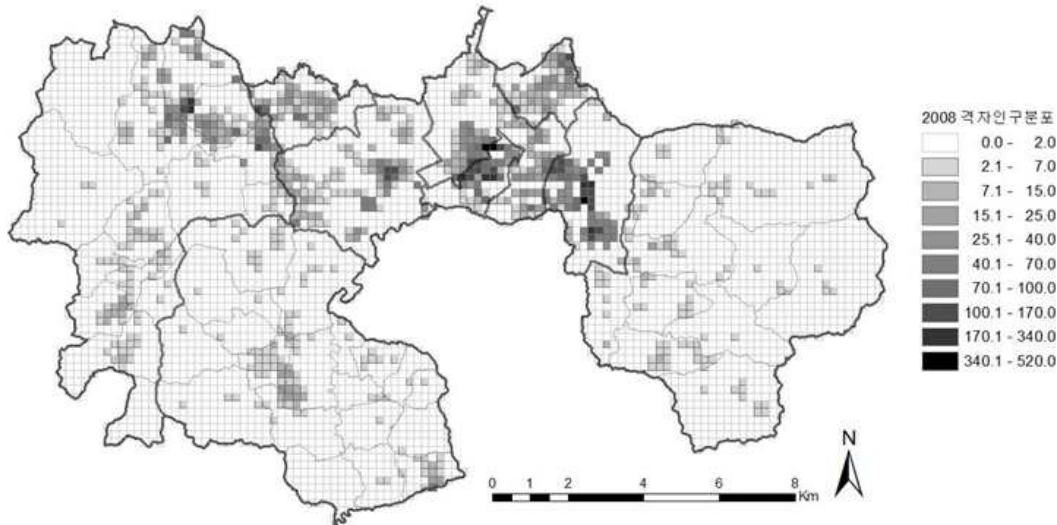


그림 4. 화성시 동부권 200X200m 격자인구분포(grided population) 추정결과

사자료의 각 해당 행정구역 인구수로 나누어 줌으로써 구한다. 2) 대상지 전체 KLIS 베이스맵을 200×200m 격자 셀로 분할한 다음 KLIS와 eAIS Join DB로부터 각 셀에 속한 주거건축물 연면적을 합산한다. 이때 필지가 격자선에 의해 분할된 경우는 필지 도형 중심점(centroid)이 속한 셀에 필지를 귀속시켜 계산한다. 그림 3은 기배동 소재 1개 셀을 확대하여 귀속 필지 및 건물이 선택(selection)된 결과를 제시하고 있다. 3) 이렇게 구한 셀별 주거건축물 연면적을 1)에서 구한 1인당 주거면적으로 나누어 해당 셀의 인구를 추계한다. 그림 4는 이와 같은 과정에 의해 화성시 동부권 전체를 200×200m 격자공간으로 분할하여 도출한 격자형 인구구조 공간분포 현황을 플로팅한 결과이다.

#### 4. GenAlg 모형 추정

##### 4.1 GenAlg 모형 추정을 위한 분석 셀 선정

그림 4의 격자분석 결과는 인구밀도가 낮은 비도시 지역에 거주인구가 적은 셀들이 많을 뿐만 아니라 중앙부 6동의 경우도 신도시와 택지개발사업지구 등에 인구가 고도 밀집한 셀들이 다수 있음을 알 수 있다. GenAlg 모형의 신뢰도를 높이기 위해서는 이와 같은 돌출값(outliers)들을 제거할 필요가 있으며 본 연구에서는 GeoDa의 백분위 맵(Percentile Map)과 상대리스크 맵(Relative Risk Map) 등 Choropleth Map 분석기법을 통해 셀들을 선별하였다(분석기법은 Anselin 등(2006) 참조). 백분위 분석에서 99% 이상의 상위 18개 고밀 셀을 제거하고 산악지역과 농림지역 및 일부

표 1. GenAlg 모형 투입 설명변수 일람

코드	범주	코드	요소
A	자연환경요소	A1	표고
		A2	경사
B	개발 촉진형 용도지역지구	B1	택지개발예정지구
		B2	개발진흥지구
C	개발 규제형 용도지역지구	C1	보전녹지
		C2	계획관리지역

도시지역 저밀 셀들을 90 Percentile까지 제거하여 유효 설명력을 가질 수 있는 셀들을 1차 선별한 후 상대리스크 분석을 통해 리스크 5.0 이상의 고 리스크 셀들을 제외시켜 보수적인 수준에서 70개 분석 셀이 최종 선택되었다.

##### 4.2 모형 설명변수 구축

본 연구의 GenAlg 모형에 투입할 자연환경 요소인 표고 및 경사도의 변수값은 GIS상에서 200×200m 셀 각각의 평균 해발고도(m) 및 평균 경사도(%)를 측정하여 구축하였다. 또한 용도지역지구에 대하여는 넓은 면적의 해당 polygon feature들을 ArcGIS 프로그램에 내장되어 있는 거리측정 도구(tool) 메뉴 중 하나인 NEAR 함수 툴을 이용하여 boundary에서 각 셀에 대해 수선의 발을 찾는 방식으로 최단 거리를 m단위로 측정하여 구축하였다. 표 1은 이와 같이 구축한 본 연구 GenAlg 모형 설명변수들의 범주와 명칭을 코드와 함께 열거한 것이다.

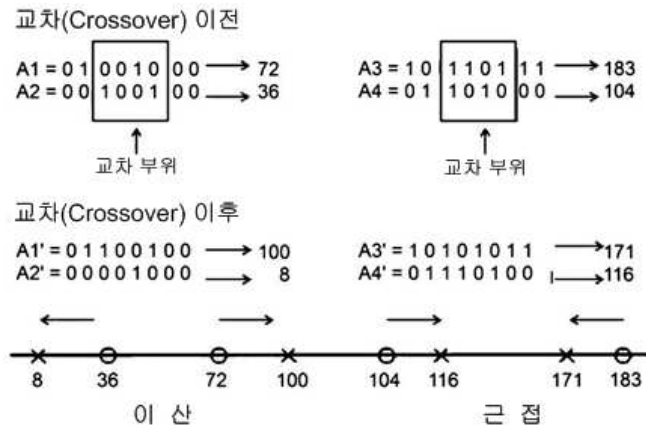


그림 5. Genetic Algorithm의 염색체 교차(Crossover) 작동기제

### 4.3 GenAlg 모형 작동기제

GenAlg의 작동기제는 자연생태계에서의 종의 진화 과정을 설명하는 다윈의 적자생존(survival of the fittest) 원리에 그 기초를 두고 있다(Balakrishnan과 Jacob, 1996). 추정코자 하는 계수들의 십진법 수치는 전산처리 과정에서 보다 긴 자릿수를 갖는 0과 1의 2진법 숫자(binary string)로 변환되는데 이때 스트링을 생명체의 유전자 염색체(chromosome) 구조와 같이 해석하여 임의의 값을 갖는 초기 개체군(initial population)들을 생성시킨 후 그림 5의 다이어그램과 같이 한 쌍씩의 선대(parents) 개체(그림에서 A1, A2 쌍과 A3, A4 쌍)로부터 2진법 스트링 일부를 교차(Crossover)시켜 후대(offspring) 개체(그림에서 A1', A2' 쌍과 A3', A4' 쌍)를 생성시킨다. 새로운 후대 개체 값들이 모형의 최대화(Max) 혹은 최소화(Min)를 달성하려는 평가함수(Evaluation Function) 값을 향상시킬 경우 보다 적합한(fitter) 개체로 채택하여 다음 세대(generation) 개체군에 생존시키고 그렇지 않을 경우 도태시킨다. 이때 그림의 A1, A2 pair의 경우는 교차 이후 후대 세대의 결과값이 십진법으로 환원하면 교차 이전값 72와 36에서 100과 8로 간격이 벌어진 이산(pull away)의 경우를, 반대로 A3, A4 pair는 이전의 183과 104에서 171과 116으로 간격이 좁아진 근접(get closer)의 예를 도식적으로 보이고 있다. GenAlg는 이와 함께 돌연변이(Mutation)도 발생시키는데 단순히 염색체 교차만으로는 때로 국지적 극대, 극소값에서 진화를 멈출 수 있기 때문에 특정 염색체 스트링 일부의 위치를 바꾸는 돌연변이(mutation)를 발생시킴으로써 전구간 최대, 최소값(global maximum or minimum) 도달 가능성을 높이는 진화적 기제로 작

용하게 된다(Su와 Chang, 2000).

본 연구에서는 GenAlg 셀 추정인구( $Y_{GenAlg}$ )를 다음과 같이 설정하였다.

$$Y_{GenAlg} = \sum_{n=1}^{N_n} \left( \theta \cdot WT_n \sum_{m=1}^{N_m} wt_m r_m \right) \quad (식 1)$$

식에서  $WT_n$ 은 A, B, C 3개 범주 각각의 가중치,  $wt_m$ 은 각 범주 내 설명변수들 각각의 가중치이며  $N_n$ 은 범주의 개수,  $N_m$ 은 각 범주 내 변수의 개수다. 특히 각 범주 또는 계수들의 상대적인 기여도를 비교하기 위해  $\sum_n WT_n = 1$ ,  $\sum_m wt_m = 1$ 로 제약조건을 설정하였다.  $r_m$ 은 설명변수 값으로서 각 변수의 최대 값에 대한 비율(ratio)로 설정하였다.  $\theta$ 는 Summation 수식을 관찰값에 근접시키는 임의의 상수이다. 따라서 GenAlg 모형을 통해 추정하는 개체(chromosome)들은  $\theta$ ,  $WT_n$  및  $wt_m$ 의 세 가지가 된다.

한편 모형 적합성을 판단하는 평가함수로는 관측된(observed) 격자인구  $Y_{Obs_i}$ 와 위 식 1에 의해 계산된  $Y_{GenAlg_i}$  간의 오차평균 제곱의 제곱근(Root Mean Square Error:  $e_{rms}$ )로 정의하고 이를 최소화(Min)하는 것으로 모형의 성능을 판단하였다(Brookes, 2001). 이때  $e_{rms}$ 는 다음 식과 같이 계산된다.

$$e_{rms} = \sqrt{\frac{1}{n} \sum_{i=1}^k (Y_{Obs_i} - Y_{GenAlg_i})^2} \quad (식 2)$$

표 2. GenAlg모형 3개 최적결과의 평가함수 요약

평가함수	GenAlg 최적해 평가함수		
	1차 결과	2차 결과	3차 결과
$e_{rms}$	6.9896605	7.0115900	7.0393161

표 3. GenAlg모형 3개 최적결과의 설명변수 가중치 요약

코드	범주	코드	요소	GenAlg 최적해 가중치					
				1차 결과		2차 결과		3차 결과	
A	자연환경요소	A1	표고	0.37255	0.13333	0.37255	0.12941	0.46667	0.14510
		A2	경사		0.87451		0.87451		0.84706
B	개발 촉진형 용도지역지구	B1	택지개발예정지구	0.11373	0.88235	0.10196	0.98039	0.11765	0.81176
		B2	개발진흥지구		0.12157		0.02353		0.19216
C	개발 규제형 용도지역지구	C1	보전녹지	0.50196	0.37255	0.50196	0.36078	0.42745	0.30980
		C2	계획관리지역		0.63922		0.62745		0.69412
가중치( $WT, wt$ ) 합계				0.98824	3.02353	0.97647	2.99608	1.01176	3.00000

5. GenAlg 모형 추정결과

표 2와 3은 다수 반복 수행한 GenAlg모형 중 3개 최적성능 결과를 요약한 것이다. 우선 표 2는 GenAlg의 평가함수(Evaluation Function)인  $e_{rms}$  값을 보이는데 약 7.0 내외에서 값이 수렴하고 있다. 참고로 회귀분석 결과에서는  $e_{rms}$ 가 6.828495가 나와 GenAlg모형이 회귀분석보다 모형오차에 있어 약간 상회하는 것으로 나타났다. 이처럼 GenAlg와 회귀분석 모두 오차가 더 낮은 수준으로 수렴치 못하는 것은 앞서 격자인구분포 개관에서도 살펴보았듯이 화성시 동부권이 고도로 도시화된 지역 주변에 저밀인구의 산악지역, 농림지역이 함께 병존하고 있는 이질적 입지특성에 기인하는 것으로 여겨진다.

표 3은 GenAlg모형의 범주와 설명변수에 대한 가중치( $WT, wt$ )를 정리한 것이다. 범주별 특징을 살펴보면 우선 3개 결과 모두에서 일관되게 A) 자연환경요소가 37~47%, C) 개발 규제형 용도지역이 42~50% 정도 이 두 범주가 압도적인 영향력을 가짐을 알 수 있고, B) 개발 촉진형 용도지역은 오직 10~11% 정도의 영향력을 보이는 것으로 나타났다. 표고와 경사는 지역 특성과 무관하게 토지이용잠재력을 결정하는 범지역적 요소임을 확인시켜주는 결과로 보인다. 용도지역지구 중에서는 현재 화성시 동부권에서 다수의 도시개발사업이 적극 진행 중이기 때문에 기왕의 개발 촉진형 용도지역보다는 역으로 규제형 용도지역 지정현황이 인구증가 억제요인으로 더 크게 작용한다는 사실을 시사하는 것으로 여겨진다.

다시 설명변수 별 영향력을 보면 A) 자연환경요소 중 경사가 84~87% 정도로 압도적인 영향력을 보인

반면 표고는 13~14% 정도로 상대적으로 영향력이 미약함을 알 수 있다. 표고는 주위와 해발고도가 달라도 평탄면을 유지할 경우 토지이용효율이 유사한 반면 경사도는 표고와 무관하게 어느 입지상황에서나 토지이용에 결정적이라는 점에서 추정치가 의미있는 것으로 판단된다. B) 개발 촉진형 용도지역 중에는 개발진흥지구가 2~19%인 반면 택지개발예정지구가 81~98%로 압도적인 영향력을 가진 것으로 나타난다. 현재 개발진흥지구는 상업, 공업용도를 망라한 도시개발사업유치를 목적으로 지정되기 때문에 거주인구 증가 유발 효과에 있어 주거전용의 택지개발사업지구가 보다 더 큰 영향력을 가지는 것으로 해석된다. C) 개발 규제형 용도지역지구는 가장 영향력이 큰 범주로 그 중 계획관리지역이 63~69%로 40~37% 수준의 보전녹지보다 큰 영향력을 보인다. 계획관리지역은 기존의 난개발 우려지역에 대한 비계획적 추가개발을 관리하기 위한 용도지역으로 역시 개발잠재력 자체는 큰 지역임을 반증하는 것으로 여겨진다. 마지막으로 표의 맨 아랫줄에는 가중치 합계가 집계되어 있다. 주어진 제약식에서  $\sum_n WT_n = 1, \sum_m wt_m = 1$  (따라서,  $\sum_{n=1}^3 \sum_{m=1}^3 wt_m = 3$ )로 주어졌으나,  $\pm 0.02$  내외의 약간의 오차가 있음을 알 수 있는데 이는 GenAlg 모형이 평가함수 최소화 과정에서 제약조건식의 값을 정확히 충족하지 못할 경우 가장 가까운 값으로 수렴시키는 특성에 기인한다.

표 4. 회귀분석모형의 설명력

$R^2$	수정 $R^2$	회귀분석 표준오차	F-통계량	확률
0.305705	0.239581	8.19602	4.62325	0.00059

표 5. 회귀분석에 의한 설명변수 통계치 요약

코드	요 소	베타	표준오차	t-통계량	확률
A1	표고	-0.0530611	0.1226791	-0.432520	0.666841
A2	경사	-0.8821053	0.4471182	-1.972868	0.052902
B1	택지개발예정지구	-0.0009065	0.0013382	-0.677415	0.500623
B2	개발진흥지구	0.0013382	0.0023771	0.562949	0.575467
C1	보전녹지	0.0016250	0.0009222	1.762214	0.082884
C2	계획관리지역	0.0030133	0.0010742	2.805057	0.006683

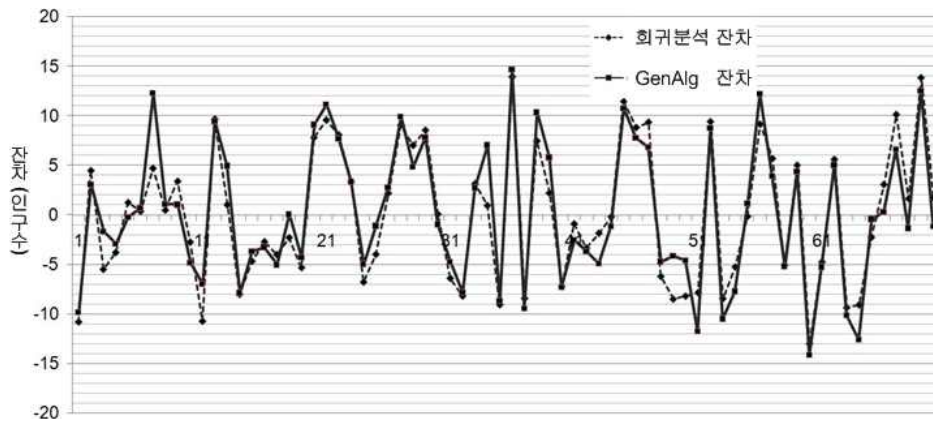


그림 6. 회귀분석과 GenAlg모형의 잔차(Residual)분포도

표 4와 5는 GenAlg모형과 동일한 변수조건으로 회귀분석을 실시한 결과를 보이고 있다. 표 4에서  $R^2$ 은 0.3057로 우수한 결과가 되지 못하는데 이는 표 5 통계량에서 보듯이 표고, 택지개발예정지구, 개발진흥지구 등 GenAlg모형에서 다소 낮은 범주 가중치와 변수 가중치를 보였던 설명변수들의 신뢰도가 낮은 데 기인하는 것을 알 수 있다. 즉 회귀분석에서 신뢰도가 낮은 변수는 GenAlg모형에서도 동시에 그 가중치가 낮게 평가된다는 점을 확인할 수 있다. 또한 회귀분석  $R^2$  값 0.3057의 제곱근은  $R$ (상관계수) 값으로서 0.5529가 되는데 GenAlg 3개 최적해의  $R$ 값도 각각 0.5560, 0.5544, 0.5556으로 측정되어 두 모형의 추정결과가 매우 근사함을 알 수 있다. 그림 6의 회귀분석과 GenAlg 모형의 잔차 분포도(residual plot)에서도 잔차의 패턴이 두 모형에서 매우 근사하게 나타나 두 모형의 설명력이 역시 매우 유사하며, 시각적으로는 GenAlg 잔차의 변동폭이 대체적으로 다소 큰 폭으로 움직임을 알 수 있는데 이는 앞서 언급했듯이 GenAlg의  $e_{rms}$  오차값이 회귀분석보다 약간 상회하고 있는 결과를 뒷받침한다.

## 6. 결 론

### 6.1 연구결과의 요약

본 연구에서는 GenAlg모형 평가함수인 모형오차 값이 회귀분석과 마찬가지로 기대보다 작지 못했다는 아쉬움이 있다. 평가함수를 더 이상 낮추지 못하는 점은 앞서 언급한 바와 같이 본 사례대상지인 화성시 동부권이 도시와 산지, 농지가 혼재하며 급속한 개발단계에서 인구동태와 도시현황이 아직 안정되지 못한 입지특성에 기인하는 것으로 여겨진다. 그러나 이 같은 한계에도 불구하고 설명변수들에 대한 분석결과는 비교적 일관되고 설득력 있는 결과로 판단되며 그 추정결과는 범용적 추정기법 중 하나인 공간회귀분석과 매우 근사한 결과를 도출할 수 있음을 알 수 있었다.

### 6.2 연구의 시사점

GenAlg모형 적용에 있어서 다양한 공간적 변수들의 구축은 GIS를 활용한 지리정보체계와의 연계가 불가분하다는 점에서 GenAlg기법과 GIS의 연계활용이 필수적이라는 점을 간과할 수 없다. GenAlg는 연산처리에 있어 생물학적 진화 기제를 기반으로 유전학적

개념을 적용하여 적정해를 찾는 방법으로서 본 연구에서와 같은 인구추정이라든가 도시현상 탐색 등 일견 유전학적 개념과 거리가 먼 최적화문제들에 있어서도 적정해를 찾을 수 있는 유용한 분석방법 중 하나임을 명확히 확인할 수 있었다. 다만 공간회귀분석기법과 비교할 때 채택한 설명변수들의 설명력을 통계적 의미에서 정량적으로 검증할 수 없다는 점과 모형 추정에 있어서도 공간회귀분석을 능가하는 우월한 분석결과를 항상 도출한다고는 볼 수 없다는 점에서 한계가 있으나, GenAlg 기법의 적용은 회귀분석에 상응하는 유용한 대안으로서 변수 선택 등에 있어 회귀분석과 병행하여 사용할 경우 매우 직관적이며 보완적인 도시분석 기법이 될 수 있을 것으로 판단된다.

### 감사의 글

이 논문은 2009학년도 홍익대학교 학술연구진흥비에 의하여 지원되었음.

### 참고문헌

1. 국토해양부, 2008a, *기반시설부담구역제도 시행 및 운영활성화를 위한 연구*.
2. 국토해양부, 2008b, *기반시설부담구역제도 운영 편람*.
3. 김형복, 홍철진, 2003, "개발에 따른 기반시설부담비용의 적정배분에 관한 연구," *국토계획*, 대한국토도시계획학회, 제 38권, 3호, pp.259-74.
4. 김형복, 전병은, 최내영, 2006, *기반시설부담금에 관한 법률 및 그 운영에 대한 특강자료*, 대한국토도시계획학회 국토도시아카데미.
5. 최내영, 2009, "인구격자 셀 버퍼공간 설정에 의한 기반시설부담구역경계 검토방안 연구," *한국지형공간정보학회지*, 한국지형공간정보학회, 제17권, 4호, pp.53-60.
6. 화성시, 2004, *화성시 기반시설부담구역 지정 및 부담계획 기준수립 연구*.
7. Anselin, L., Lozano N. and Koschinsky, J., 2006, *Rate Transformations and Smoothing*, A discussion paper of the Spatial Analysis Laboratory, University of Illinois at Urbana-Champaign.
8. Balakrishnan, P.V. and Jacob V.S., 1996, "Genetic Algorithms for Product Design", *Management Science*, Vol 42, pp.1105-1117.
9. Brookes, C.J., 2001, "A Genetic Algorithm for Designing Optimal Patch Configurations in GIS", *International Journal of Geographical Information Science*, Vol. 15, pp.539-559.
10. Holland, H.H., 1975, *Adaptation in Natural and Artificial Systems*, The University of Michigan Press, Ann Arbor.
11. Koza, J.R., 1990, *Genetic Programming: A Paradigm for Genetically Breeding Populations of Computer Programs to Solve Problems*, Stanford University Report, Report No. STAN-CS-90-1394.
12. Li, X. and Yeh, A.G., 2005, "Integration of Genetic Algorithms and GIS for Optimal Location Search", *International Journal of Geographical Information Science*, Vol. 19, pp.581-601.
13. Muttill, N. and Lee J., 2005, "Genetic Programming for Analysis and Real-time Prediction of Coastal Algal Blooms", *Ecological Modelling*, Vol. 189, pp.363-376.
14. Su, M.C. and Chang, H.T., 2000, "Application of Neural Networks Incorporated with Real-valued Genetic Algorithms in Knowledge Acquisition", *Fuzzy Sets and Systems*, Vol. 112, pp.85-97.
15. Yue, T.X., Du, Z.P. and Song, Y.J., 2008, "Spatial Models and Geographic Information Systems", *Encyclopedia of Ecology*, pp.3315-3325.
16. Yue, T.X., Wang, Y.A., Chen, S.P., Liu, J.Y., Qiu, D.S., Deng, X.Z., Liu, M.L. and Tian, Y.Z., 2003, "Numerical Simulation of Population Distribution in China", *Population and Environment*, Vol. 25, pp.141-163.