



기본 | 22-24

인공지능 기법을 적용한 소지역단위 장래인구 예측 방법론 개발 기초 연구

A Study on Small Area Population Forecast Model with Artificial Intelligence

이보경, 장요한, 정예진, 홍사흠, 김동근

기본 22-24

인공지능 기법을 적용한 소지역단위 장래인구 예측 방법론 개발 기초 연구

A Study on Small Area Population Forecast Model with Artificial Intelligence

이보경, 장요한, 정예진, 홍사흠, 김동근

■ 저자

이보경, 장요한, 정예진, 홍사흠, 김동근

■ 연구진

이보경 국토연구원 부연구위원(연구책임)

장요한 국토연구원 국토데이터랩 팀장

정예진 국토연구원 연구원

홍사흠 국토연구원 국토계획평가센터장

김동근 국토연구원 연구위원

■ 연구심의위원

김태환 국토연구원 부원장

김대종 국토연구원 공간정보사회연구본부장

이영주 국토연구원 연구위원

민성희 국토연구원 연구위원

임시영 국토연구원 부연구위원

김오석 고려대학교 교수

안윤상 국토교통부 사무관

주요 내용 및 정책제안

FINDINGS & SUGGESTIONS



주요 내용

- 1 인구감소시대에 국가와 지자체를 지속가능하게 운영하기 위해서는 데이터기반의 과학적인 증거를 바탕으로 적절한 계획과 정책 수립이 필요
- 2 인공지능 기법을 적용하여 '격자단위 장래인구 예측 AI 모델(Gridded Population Forecast with AI, 이후 GPAI)'를 개발하고 향후 국토 정책 및 계획에서의 활용방안을 제언
- 3 장래인구 예측 결과 2020년 평균 인구밀도를 100으로 보았을 때 2050년에 91수준으로 감소하며, 한정된 인구 밀집 지역과 대부분의 인구 희박지역으로 양극화될 것으로 전망

학술적 기여 및 정책제안

- 1 1km 격자단위 장래인구 예측을 위해 공간 빅데이터를 활용하고 이종의 두 모델(코호트요인법-인공지능 기법)의 결합 아키텍처를 개발하여, 향후 후속 연구의 방향성을 제시
 - 장래인구 예측에 다양한 출처의 공간 빅데이터의 융합·활용 방법을 제시하고, 격자단위 데이터의 시·공간 관계성을 반영한 인공지능 기법 모델을 개발
 - 인구학적 특성과 공간적 특성을 함께 고려할 수 있도록 코호트요인법 모델과 인공지능 기법 모델을 앙상블(ensemble)로 결합한 아키텍처 설계 방식을 최초로 제안
 - 딥러닝 알고리즘을 활용하여 인구분포에 큰 영향을 미치는 건축물 수, 주거용 연면적 등의 공간적 특성 변화를 전망하여 장래인구 예측에 활용
- 2 GPAI는 (1) 유연한 정책·계획의 공간 구역 설정, (2) 미래예견적 계획수립을 위한 과학적 의사결정 지원 및 (3) 실효성있는 지자체 계획 수립의 근거 등으로 활용 가능
- 3 향후 소지역단위 장래인구예측 모델이 신뢰성있게 실무에서 활용되기 위해서는 후속 연구를 수행하고 활용 확산을 주도할 수 있는 권한과 역량을 갖춘 조직과 제도 마련이 필요



1. 연구 개요

□ 연구 배경과 필요성

- 인구감소시대를 맞이함에 따라 지방소멸의 위기가 본격적으로 대두되고 있으며, 국토 내 공간별로 겪고 있는 소멸의 위기 수준은 천차만별
 - 수도권 일부 지역은 인구증가와 밀집에 따른 인프라 부족 등의 문제를 겪고 있지만 지방 소도시는 생산인구의 급격한 감소로 지역경제 침체를 걱정
 - 수도권과 비수도권 간의 인구 격차뿐 아니라 같은 지자체 내에서의 서로 다른 공간 간의 인구 격차 역시 국토의 균형발전을 저해하는 요인
 - 그러나 국가와 지자체는 ‘소멸의 위협에 직접 노출된 장소는 어디인지?’, ‘실제 이 공간에 적절한 인구감소 관련 지원이 이루어지고 있는지?’, ‘지금까지 투입된 예산이 인구감소 대응에 효과가 있었는지?’ 등 공간의 관점에서 관련 이슈에 답하지 못하고 있는 실정
- 지속가능한 국토·지역·도시 운영을 위해서는 데이터기반의 과학적인 증거를 바탕으로 적절한 계획과 정책 수립이 필요
 - 윤석열 정부의 지역균형발전 국정과제에서는 인구감소에 따른 소멸 위기를 겪고 있는 지역에 대해 효율적인 자원 배분 및 인프라 경영·관리 등 적절한 정책과 계획을 수립·시행할 수 있는 역량을 요구(제20대 대통령직인수위원회, 2022. 4. 27)
 - 2022년 말 공표된 ‘인구구조 변화와 대응방안’에서는 데이터와 과학에 기반하여 정책 효과성을 검증하고, 공간 빅데이터를 활용한 권역 설정을 통한 인구정책 추진의 필요성을 강조(저출산고령사회위원회·관계부처, 2022. 12. 28)
 - 고해상도 소지역단위 장래인구를 활용하면 앞으로 어느 공간이 인구감소시대에 가장 취약할지 전망하고, 이들을 정밀하게 타겟팅한 장소기반·증거기반 정책을 수립할 수 있을 것으로 기대
- 공공과 민간부문에서 공간 빅데이터 생산과 공개가 증가하고 있고, 분석 컴퓨팅 환경이 발전함에 소지역단위 장래인구 예측이 가능해진 상황

□ 연구 목적과 내용

- 이 연구의 목적은 인공지능 기법을 적용하여 소지역단위 장래인구 예측모델을 개발하고 국토·지역·도시 정책 및 계획에서 활용할 수 있는 방안을 모색하는 것
 - 첫째, 소지역단위 장래인구 예측 방법 개발 관련 국내·외 연구동향과 사례를 조사하고 시사점을 도출하여 개발 방향을 정립
 - 둘째, 1km 격자단위¹⁾ 장래인구 예측 방법인 ‘격자단위 장래인구 예측 AI 모델 (Gridded Population Forecast with AI, 이후 GPAI)’을 개발하였으며, GPAI 아키텍처를 인구학적 특성과 공간적 특성을 모두 고려할 수 있는 인공지능 기반 모델과 코호트인법 모델을 결합하는 방식으로 설계
 - 셋째, GPAI를 전국에 적용하여 2050년까지 인구분포 변화를 전망하고, 인구 증가지역, 인구 감소지역, 무거주화 위험지역 등 국토공간을 유형화하여 비교·분석
 - 넷째, GPAI가 실무에서 신뢰성있게 활용되기 위해 필요한 방법론 고도화 방안, 활용 예시 및 활용 확산 전략 등을 제언

□ 기대효과

- [정책] 소지역단위 장래인구 예측 방법론과 적용 결과는 인구감소시대에 공간의 맞춤형 계획수립을 위한 기초자료로 활용될 것으로 기대
 - 소지역, 특히 격자단위 장래인구 예측 결과는 생활권부터 광역권까지 공간적 정합성을 가지고 인구변화와 관련된 정보를 제공할 수 있음
 - 정책적 관심에 따라 자유롭게 구역을 설정하여 제한된 자원을 가장 효과적으로 배치하고 운영하는 의사결정의 기초자료로 활용 가능
- [학술] 공간적 특성을 반영한 인공지능 기반 모델과 인구학적 특성에 따른 코호트인법 기반 모델을 앙상블(ensemble)²⁾ 보팅(voting)³⁾ 방법으로 결합하는 접근 방식을 제안

1) 소지역의 공간범위는 활용 목적과 시대에 따라 다양하게 정의될 수 있음. 이 연구에서는 선행연구 검토 결과와 우리나라 공간 데이터 현황과 정책 및 계획에서의 활용을 고려하여 1km격자단위를 소지역으로 채택

2) 앙상블(Ensemble)은 데이터에 기반한 알고리즘의 과대 또는 과소적합을 방지하기 위해 다양한 알고리즘이 조화를 이루어 최종값을 도출하는 데이터마이닝 방법임

3) 보팅(voting)은 앙상블 중 하나로, 예측 결과값이 명목형 변수인 경우 하드보팅, 소프트보팅, 가중치 투표 등으로 하나의 값을 선택하고 연속형인 경우(장래인구 규모) 각 알고리즘을 통해 도출된 결과의 평균을 활용하여 두 모델을 결합

2. 연구 동향 및 사례 조사

□ 소지역단위 장래인구 예측 방법 관련 선행연구 현황

- 국내·외 관련 선행연구 동향을 면밀히 파악하여 각 방법론의 장점과 한계를 검토
 - 국내외 소지역단위 인구예측 관련 선행연구를 방법론 관점에서 기초모델, 코호트 생존율 활용 모델, 공간통계모델, 마이크로시뮬레이션, 도시성장모델, 인공지능 기법으로 분류하여 검토(<표 1> 참조)

표 1 | 소지역단위 장래인구 예측 방법 유형 및 특징

분류	방법론 및 모델	내용
기초모델	<ul style="list-style-type: none"> • 경향외삽법: Linear/Exponential model, Holt-Winters, Autoregressive Integrated Moving Average 등 • 비율법: share of population 등 	<ul style="list-style-type: none"> • 최소한의 데이터를 활용하여 직관적으로 소지역인구 현황을 예측 • 주로 연령별 인구에 비해 총인구 규모를 예측하는데 활용 • 비율법의 경우 매개변수(주택 등)를 활용하여 지역 인구를 소지역으로 배분 • 적용이 간편하고 이해가 쉽다는 장점
코호트 생존율 활용 모델	<ul style="list-style-type: none"> • 코호트요인법 • Hamilton-Perry model 	<ul style="list-style-type: none"> • 코호트요인법은 인구변동요인의 복잡성과 연계성을 고려하여 인구균형방정식을 만들어 장래인구를 추계 • Hamilton-Perry model은 코호트별 인구변화 비율이 일정하게 지속된다는 가정에 따라 두 시점의 연령별 인구 규모의 변화율을 활용
공간통계 모델	<ul style="list-style-type: none"> • Spatial-temporal regression • Geographically-weighted regression 	<ul style="list-style-type: none"> • 공간데이터의 특성을 반영한 통계모델 • 인구예측 결과와 여기에 영향을 주는 요인들 간의 관계를 유추할 수 있어 의사결정자가 공간 정책 또는 계획을 수립하는데 필요한 정보를 제공
마이크로 시뮬레이션	<ul style="list-style-type: none"> • 아일랜드 SMILE • 호주 Spatial MSM 	<ul style="list-style-type: none"> • 개인의 행동을 예측하고 집계하여 지역의 인구규모를 전망하는 방법으로 '소지역'이라는 공간단위를 설정할 필요가 없음
도시성장모델	<ul style="list-style-type: none"> • Cellular Automata(CA) • 유전적 알고리즘(Genetic Algorithm, GA) 	<ul style="list-style-type: none"> • 토지이용과 교통을 고려한 소지역단위 인구예측 모델 • 최근 사회 경제 및 기후변화에 따른 인구 분포의 장기적 변화를 전망하는데 널리 활용
인공지능 기법	<ul style="list-style-type: none"> • 딥러닝: ANN(Artificial Neural Network), LSTM(Long Short-Term Memory) 등 • 머신러닝: XGboost, Random Forest 등 	<ul style="list-style-type: none"> • 머신러닝 등 인공지능 기법이 인구 전망에 적용된 것은 최근 2~3년 전부터 진행되기 시작함 • 인구분포와 밀접하게 관련된 요인변수를 머신러닝을 통해 예측하고 인구를 할당하는 연구도 등장

주: 본 연구 표 2-2의 일부 내용을 발췌한 것임

자료: 저자 작성

□ 격자단위 인구통계 관련 사례 조사

- 미국, EU, 일본 및 우리나라에서 격자단위 인구통계 생산과 관련된 많은 연구가 진행되었으며, 다양한 실무에 활용되고 있음
 - 국내·외 격자단위 인구통계는 대체로 1km 격자를 기준으로 생산·공표
 - 격자단위 인구통계 생산 방법의 경우 우리나라는 개인 주소정보를 바탕으로 지오코딩하여 격자별 실제 인구를 집계하는 반면, 해외에서는 행정구역 단위 센서스 등록인구를 임의의 가정에 따라 격자에 배분하는 방식
- 일본 국토교통성의 국토정책국은 2015년도 국세조사 정보를 활용하여 2050년까지 5년단위의 격자단위 장래인구 예측 방법과 결과를 공개
 - 격자단위 장래인구 예측 결과와 함께 자원 배분 및 인프라 계획에 어떻게 적용할 수 있는지와 관련된 후속 연구 및 도구 개발을 지속

3. 소지역단위 장래인구 예측 방법론 개발

□ 개발 방향

- 연구 동향 및 사례 조사를 통해 개발 방향을 도출하고 GPAI 아키텍처를 설계
 - 첫째, 시간이 지나도 변하지 않는 표준화되고 규격화된 소지역인 1km 격자단위로 장래인구를 예측할 수 있는 방법론을 개발
 - 둘째, Tobler가 제안한 셀룰러 지오그래피(cellular geography) 개념을 바탕으로 격자 간의 이웃효과를 고려할 수 있도록 알고리즘을 설계
 - 셋째, 인구학적 특성(인구규모, 사망률, 출생률, 순이동률 등)과 공간적 특성(건축물 밀도, 주거용 건축물 현황, 토지 경사도 등)의 변화를 복합적으로 고려하는 방법론을 고안
 - 넷째, 시·공간 위계가 복잡하게 연계된 대용량의 공간 빅데이터를 학습할 수 있도록 인공지능 기법을 적용
 - 마지막으로, 널리 활용되고 있는 코호트요인법과 인공지능 기법 기반 모델을 결합하여, 인구학적 특성과 공간적 특성을 모두 고려할 수 있는 장래인구 예측 방법론을 개발

□ GPAI 아키텍처 설계

- 전통적 장래인구 추계모형인 코호트요인법과 건축물 수, 주거용 연면적, 경사도 등 공간적 특성을 반영한 인공지능 기법을 결합한 새로운 방법론을 제안(〈표 2〉 참조)
 - 모델1은 인구학적 특성과 공간적 특성을 고려한 인공지능 기법 기반 모델이며, 모델2는 인구학적 특성(인구구조) 시나리오 기반 코호트요인법 기반 모델
 - 앙상블 보팅(ensemble voting)방법 통해 두 모델의 결과를 결합·장래인구를 예측

표 2 | GPAI 아키텍처 설계

구분	모델1	모델2
주요 가정	과거의 인구변화 추세를 학습하고, 장래의 공간적 특성 변화 예측을 통해 격자별 장래인구를 전망 • 과거의 해당 격자 및 주변 격자의 인구 변화가 장래의 인구 규모에 영향 • 과거의 해당 격자 및 주변 격자의 공간적 특성이 장래의 공간적 특성에 영향	격자단위 인구학적 구조(성별 및 연령)와 인구학적 변동요인에 따른 자연적 증감을 통해 격자별 장래인구를 예측 • 현재 격자단위의 성별 연령별 인구구조에 인구학적 변동요인을 적용하여 자연적 증감에 따른 장래인구변화 예측 ※ 인구이동에 따른 인구의 사회적 증감은 시도단위에서 고려
활용 데이터	인구학적 특성(시도 및 격자 단위)과 공간적 특성(격자단위)을 변수로 활용 • 2050년 시도단위 추계 사망률, 출생률, 순이동률 등 인구변동요인(통계청, 2022)과 격자단위 인구 변화(2000~2020)를 활용 • 격자단위 건축물수, 주택수, 주거용 연면적, 경사도 등 공간적특성(2000~2020)를 활용	인구학적 특성(시도단위)과 인구 구조적 특성(격자단위)을 변수로 활용 • 2050년 시도단위 추계 사망률 및 출생률과 같은 인구변동요인(통계청, 2022) • 2020년 격자단위 성별/연령별 인구구조(행정안전부 주민등록 통계)를 기준인구로 설정
적용 방법론	3개 단계(인구학적-공간적 특성 학습, 장래의 공간적 특성 변화 예측, 장래의 인구학적-공간적 특성 기반의 격자단위 인구 예측)으로 구성 • 1단계: 인구학적 특성, 공간적 특성 및 격자단위 인구분포와의 관계를 학습을 위해 이웃효과를 고려한 랜덤포레스트(Random Forest) 수행 • 2단계: CNN-LSTM(Convolutional Neural Network-Long short Term Memory) 알고리즘을 적용하여 공간적 특성 예측 • 3단계: 장래의 인구학적 변동요인과 2단계 공간적 특성 예측 결과를 1단계에서 학습한 랜덤포레스트에 적용하여 격자단위 장래인구를 예측	이보경(2019)에서 제안한 격자단위 장래인구 추계를 위한 코호트요인법을 활용 • 1단계: 격자단위 기준인구 정의 • 2단계: 격자별 코호트 요인법적용 • 3단계: 격자별 장래인구 가중치 산출 • 4단계: 시도단위 장래인구 추계를 격자별 가중치를 기준으로 배분하여 통계청과 정확성을 갖춘 격자단위 장래인구 예측값 도출
최종 결과	1km 격자단위 장래인구 예측(2025년 ~ 2050년)	

자료: 저자 작성

4. GPAI 학습 및 장래인구 예측 결과

□ 전국 및 광역권 장래인구 변화 전망

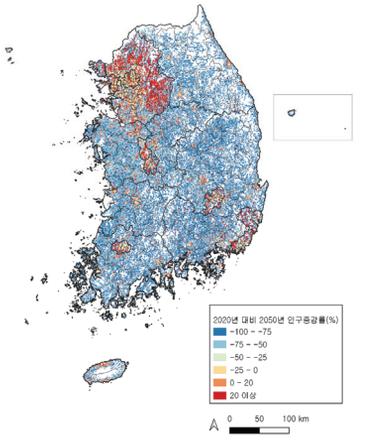
- GPAI를 적용하여 2050년까지 장래인구를 예측한 결과, 2020년 인구밀도를 100으로 보았을 때 2050년의 인구밀도는 91수준으로 감소하고, 인구거주 지역 중 장래 무거주화 위험 및 인구 감소지역이 87%에 달할 것으로 전망(<표 3> 참조)

표 3 | 전국 2020년 ~ 2050년 인구밀도와 인구증가, 인구감소 및 무거주화 위험 지역 비중 변화

구분	2020년	2025년	2030년	2035년	2040년	2045년	2050년
평균 인구밀도 (전 국토 기준)	483명*/km ²	479명/km ²	477명/km ²	473명/km ²	467명/km ²	456명/km ²	441명/km ²
평균 인구밀도 (인구거주지역 기준)	724명/km ²	696명/km ²	694명/km ²	695명/km ²	691명/km ²	683명/km ²	672명/km ²
무거주화 위험 지역 비중 (2020년 기준)	-	1.6%	1.7%	1.9%	2.3%	2.9%	3.9%
인구감소지역 비중 (2020년 기준)	-	58.2%	74.9%	79.8%	81.4%	83.2%	83.3%
인구증가지역 비중 (2020년 기준)	-	40.2%	23.4%	18.3%	16.3%	13.9%	12.8%

주: *2020년 전 국토 기준 인구밀도의 경우 SGIS 격자통계는 개인정보보호를 위해 5인이하를 0처리한 경우가 있어, 해석의 왜곡을 방지하기 위해 비마스킹된 국토지리정보원 2020년 10월 기준 1km 격자단위 인구통계를 활용
자료: 저자 작성

그림 1 | 2020년 대비 2050년 인구증감률 전망



자료: 저자 작성

- 2050년에는 현재(2020년)에 비해 수도권, 충남 북부-세종-대전 및 일부 광역시 외곽지역 공간에서는 인구가 증가하지만, 대부분의 국토공간에서 인구감소 현상을 겪을 것으로 전망(<그림 1>)
 - 수도권과 비수도권의 인구비중은 2020년 약 50.2%에서 2050년 약 53.0%로 격차가 커질 것
 - 2020년 대비 2050년 수도권 인구 규모는 96.5% 비수도권의 경우 86.5%로 모두 감소하며, 비수도권에서의 인구감소 폭이 더 클 것으로 나타남
 - 2050년에 대전·세종·충청권의 총인구는 증가하는 것으로 전망되었으나, 격자단위 장래인구 예측 결과 해당 지역에서도 공간 간의 인구양극화가 심화될 것

□ 장래인구 변화에 따른 공간 유형과 특성

- 격자가 위치한 지자체의 특성(수도권, 광역시, 도내 시 지자체와 도내 군 지자체)과 인구 증감 패턴(인구감소, 무거주화 위험, 인구증가)을 종합하여 각각의 격자를 유형화하고, 유형별 정주여건을 비교(<표 4> 참조)
 - 수도권 → 광역시 → 도내 시 지자체 → 도내 군 지자체 순으로 인구밀도가 낮아질 것으로 전망
 - 대표적인 정주여건인 생활인프라(초등학교와 약국) 접근성을 유형별로 비교한 결과, 인구증가 공간이 대체로 인구감소 및 무거주화 위험 공간에 비해 생활인프라 이용 편의성이 양호한 것으로 나타남
 - 도내 시·군 지자체에서 인구가 증가할 것으로 예상되는 공간(격자)의 경우 수도권보다도 생활인프라 이용편의성이 양호한 것으로 나타나, 총인구가 감소할 것으로 예상되는 군 지역이라고 하더라도 정주여건이 좋으면 인구가 증가할 수 있음을 시사

표 4 | 인구감소 유형별 정주여건

유형	유형별 비중	2050년 평균 인구밀도	평균 초등학교 접근성 (2020년 기준)	평균 약국 접근성 (2020년 기준)
수도권	100%	2,590명/km ²	4.2km	3.2km
무거주화 위험 공간	1.9%	-	7.2km	5.3km
인구감소 공간	57.5%	2,766명/km ²	4.4km	3.4km
인구증가 공간	40.5%	2,464명/km ²	3.7km	3.0km
광역시	100%	2,471명/km ²	4.0km	3.2km
무거주화 위험 공간	3.2%	-	6.7km	5.3km
인구감소 공간	60.5%	2,703명/km ²	3.9km	3.1km
인구증가 공간	36.3%	2,300명/km ²	4.1km	3.1km
도내 시 지자체(경기 제외)	100%	444명/km ²	6.5km	4.3km
무거주화 위험 공간	3.7%	-	9.7km	6.5km
인구감소 공간	87.1%	293명/km ²	6.7km	4.4km
인구증가 공간	9.2%	2,054명/km ²	3.2km	2.5km
도내 군 지자체(경기 제외)	100%	100명/km ²	8.0km	5.0km
무거주화 위험 공간	4.7%	-	10.6km	6.7km
인구감소 공간	90.6%	30명/km ²	8.1km	5.0km
인구증가 공간	4.7%	1,558명/km ²	4.1km	2.6km

자료: 생활인프라 접근성은 2020년 국토지리정보원 국토조사 결과를 활용하여 저자 작성

5. 정책 제언 및 결론

□ GPAI의 독창성 및 학술적 기여

- 첫째, 다양한 소지역의 공간단위 중 1km 격자단위로 장래인구를 예측하기 위해 격자단위 공간데이터의 특성을 반영할 수 있는 방법론을 개발
 - 다양한 출처의 원천데이터로부터 격자단위의 공간 데이터를 가공·생산하여, 공간적 특성에 대한 새로운 접근고 다양한 데이터를 장래인구 예측에 활용할 수 있음을 제시
 - 1km격자 공간단위 데이터의 시·공간 관계성을 정의하고 인공지능 기법 모델에 적용
- 둘째, 인구학적 특성과 공간적 특성을 함께 고려할 수 있도록 코호트요인법 모델과 새로운 인공지능 기법 모델을 앙상블로 결합한 아키텍처 설계 방식을 최초로 제시
- 셋째, 랜덤포레스트 알고리즘으로 정밀한 격자단위 인구-인구학적 특성-공간적 특공간 관계를 밝히고, 딥러닝 알고리즘을 통해 인구분포에 큰 영향을 미치는 건축물 수, 주거용 연면적 등의 공간적 특성 변화를 예측하여 인공지능 기법을 다중 결합
- 마지막으로, GPAI 학습을 위해 활용한 데이터의 한계와 우리나라 인구변화 추세 및 기후변화와 같은 사회·환경 현상 측면에서 보완해야하는 사항 등을 검토하여 방법론 고도화 방안과 후속 연구로의 확장 가능성을 시사(〈표 5〉 참조)

표 5 | GPAI 고도화 방안

구분	고도화 방안	
데이터 측면	인공지능 기법을 학습할 수 있는 시계열 데이터가 비교적 시점이 짧고 변동성이 큼	→ 지역에 특화된 국지적 GPAI방법론을 개발하여 최대한 오차를 줄일 수 있도록 함
	인구학적 변동요인, 특히 인구이동에 대한 정보가 행정구역단위로만 공개되고 있음	→ 격자단위의 인구이동관련 공간통계를 생산·공표할 수 있도록 관련 연구와 제도 마련
	토지이용 변화를 예측할 수 있는 공간적 특성 변수를 충분히 활용하지 못함	→ 토지이용과 관련된 인문환경과 자연환경적 요소를 고려하여 격자단위의 토지이용 변화 및 거주가능지 분포를 예측하는 모델을 개발
사회·환경 현상 측면	지금까지 우리나라는 인구가 증가해왔으나, 현재를 기점으로 완전히 다른 인구감소 추세를 경험하게 될 것임	→ 2008년(1억2,808만 명)을 정점으로 13년째 인구가 감소하고 있는 일본의 변화 추세를 함께 학습하는 방법론 개발
	택지개발에 따른 대규모 주택공급과 같은 중앙정부와 지자체 개입이 방법론에서 고려되지 못함	→ 각종 중앙정부와 지자체의 주택공급에 따른 인구이동 효과를 반영할 수 있도록 방법론을 고도화
	기후변화에 따른 인구거주지역 변화 및 기후변화에 대응하기 위한 토지이용 변화와 같은 현상을 방법론에 반영하지 못함	→ 기후변화에 따른 거주지 변화 시나리오를 고려하여 소지역단위의 장래인구 분포를 예측할 수 있는 방법론 개발

자료: 저자 작성

□ 정책·계획 부문 활용 및 확산 방안

- GPAI 적용 결과와 모델은 (1) 유연한 정책·계획의 공간 구역 설정, (2) 미래 예견적 계획수립을 위한 과학적 의사결정 지원 및 (3) 인구변화 대응을 위한 실효성있는 지자체 계획 수립의 근거 등에 활용 가능
 - 첫째, 1km 격자라는 공간범위의 특성에 따라 예측 공간 범위와 위치가 표준화되어 있기 때문에 통시적이면서도 유연하게 구역을 지정하여 정책 및 계획수립과 집행이 가능
 - 둘째, 미래예견적 정책 및 계획을 수립할 수 있도록 장래의 인구변화 정보를 활용하여 시뮬레이션을 통해 과학적 의사결정을 내리는 데 도움
 - 셋째, 시·군·구보다 작은 소지역단위 인구정보를 통해 급격한 인구변화에 큰 영향을 받는 장소를 타겟팅할 수 있어, 지자체에서 실효성있는 도시·군기본계획 등 공간계획 수립이 가능
- 앞으로 격자를 포함한 소지역단위 장래인구예측 모델이 신뢰성있게 실무에서 활용되기 위해서는 방법론 고도화를 위한 후속 연구 수행과 함께 활용 확산을 위한 인센티브 부여 방법, 전문 조직 및 제도 마련 방안 등에 대한 논의가 필요
 - 미국 캘리포니아의 경우 재무국(DOF, State of California Department of Finance)은 지자체로부터 상향식 의견 수렴을 통한 인구예측을 수행하여 주 전체의 계획과 예산의 기준 자료로 활용
 - 깃허브(GitHub) 등의 오픈소스 플랫폼을 통해 GPAI 모델 결과를 공개하여 이해관계자가 모델 고도화에 참여할 수 있는 오픈이노베이션 전략 마련 필요

차례

CONTENTS

주요 내용 및 정책제안	iii
요 약	v

제1장 서론

1. 연구 배경과 목적	3
2. 연구 범위와 방법	6
3. 선행연구 검토와 차별성	11
4. 기대효과	14

제2장 소지역단위 장래인구 예측 연구 및 사례 조사

1. 소지역단위 장래인구 예측 방법론 연구 동향	19
2. 격자단위 인구통계 관련 사례 조사	36
3. 시사점	44

제3장 소지역단위 장래인구 예측 방법론 개발

1. 개발 방향	51
2. 격자단위 장래인구 예측 AI 모델(Gridded Population Forecast with AI, GPAI) 아키텍처 설계 · 56	
3. GPAI 학습과 장래인구 예측을 위한 공간 데이터	69

제4장 GPAI 학습 및 장래인구 예측 결과

1. GPAI 인공지능 기법(모델1) 학습 결과	83
2. GPAI 적용 우리나라 장래인구 예측	89
3. GPAI 고도화 방안	106
4. 소결: 방법론의 독창성 및 학술적 기여	113

제5장 정책 제언 및 결론

1. 정책 및 계획 부문 활용방안	117
2. 활용 확산을 위한 정책 제언	123
3. 결론	129

참고문헌	133
SUMMARY	143
부 록	146

표차례

LIST OF TABLES

〈표 1-1〉 선행연구 현황 및 차별성	12
〈표 2-1〉 전 세계 격자단위 인구데이터 현황	37
〈표 2-2〉 소지역단위 장래인구 예측 방법 유형 및 특징	44
〈표 3-1〉 격자단위 장래인구 예측 AI 모델(Gridded Population Forecast with AI, GPAI) 아키텍처 설계	56
〈표 3-2〉 2000~2020년 격자단위 인구 기초통계량	69
〈표 3-3〉 인구학적 특성의 데이터 출처	73
〈표 3-4〉 시도별 인구추세(과거와 미래, 2020년~2050년)	76
〈표 3-5〉 공간적 특성의 데이터 출처	77
〈표 4-1〉 1단계의 변수 중요도 및 변수 내용(상위 15개)	84
〈표 4-2〉 외생변수별 정산기준	86
〈표 4-3〉 전국 2020년~2050년 인구밀도와 인구증가, 인구감소 및 무거주화 위험 지역 비중 변화	92
〈표 4-4〉 광역권 2020년~2050년 인구밀도와 인구증가, 인구감소 및 무거주화 위험 지역 비중 변화	94
〈표 4-5〉 비수도권 인구감소 지정지역(행안부)과 인구감소 비(非)지정지역 2050년 인구 현황 비교	98
〈표 4-6〉 인구감소 유형별 정주여건	100
〈표 4-7〉 시도 광역시 공표 장래인구추계와 GPAI 예측 비교	104
〈표 4-8〉 택지지구 그룹별 유입거리별 인구 비중 누적 비율의 평균	111
〈표 5-1〉 GPAI 개발 관련 데이터 및 알고리즘 공개 방안 예시	128
〈표 5-2〉 GPAI 고도화 방안	131

그림차례

LIST OF FIGURES

〈그림 1-1〉 연구 흐름	10
〈그림 2-1〉 출산율의 공간적 이질성과 장래인구 전망에 따른 격자단위 고령자 비율 변화 ..	23
〈그림 2-2〉 시군구단위 장래인구추계와 폭염일수 간의 융합 활용	24
〈그림 2-3〉 200m 격자단위 장래인구 예측 방법론 및 활용 변수	27
〈그림 2-4〉 UIAF 모델링 구성	30
〈그림 2-5〉 소지역단위 인구추계를 위한 방법론 다이어그램	30
〈그림 2-6〉 격자단위 인구분포 예측 시뮬레이션 방법	31
〈그림 2-7〉 장래인구예측을 위한 LSTM 아키텍처	34
〈그림 2-8〉 인공지능 기법을 활용한 인구/건물 분포 예측 프로세스	35
〈그림 2-9〉 공간의 정밀성에 따른 정확성과 활용성 관계(대구광역시 및 인접 지역의 총인구 예시)	39
〈그림 2-10〉 DMVST-Net의 아키텍처 구성	41
〈그림 2-11〉 장래인구·세대예측 도구 킷 매뉴얼: 대상지역 장래인구 및 세대 예측 ..	43
〈그림 3-1〉 소지역단위로 살펴본 인구감소지역(강원도 양양군)의 30대 인구변화 사례 ..	52
〈그림 3-2〉 2020년 시군구단위 인구 분포와 소지역(1km 격자)단위 인구 분포 비교 ..	53
〈그림 3-3〉 Tobler의 셀룰러 지오그래피(Cellular Geography) 모델	54
〈그림 3-4〉 시간에 따른 공간구분별 인구 데이터 구축 현황	58
〈그림 3-5〉 시공간적 분포를 고려한 모형의 격자 데이터 배분 방법 도식화	59
〈그림 3-6〉 예측모형의 입력과 출력변수 설정 도식화	60
〈그림 3-7〉 점진적 온라인 학습 방법 적용 예시	62
〈그림 3-8〉 GPAI 다중 모델 결합 모형 아키텍처	63
〈그림 3-9〉 GPAI 모델1 개념 순서도	63
〈그림 3-10〉 CNN-LSTM 모형 개념도	65
〈그림 3-11〉 GPAI 모델2 개념 순서도	67
〈그림 3-12〉 모델1과 모델2 앙상블 결합 기반의 GPAI 아키텍처	68
〈그림 3-13〉 2000년 및 2020년 전국 인구분포(1km 격자)	70

차례

CONTENTS

〈그림 3-14〉 2000년 대비 2020년 인구 증감률(1km 격자)	71
〈그림 3-15〉 경기 남부와 충남 북부 일대 2000년과 2020년 전국 인구분포(1km 격자) ·	72
〈그림 3-16〉 시도별 인구학적 변동요인 추세(2000년~2050년)	75
〈그림 3-17〉 토지이용(건물) 현황과 인구밀도 간의 상관관계	79
〈그림 3-18〉 경상도에 따른 인구분포 예시	80
〈그림 4-1〉 2020년 격자별 실제 인구와 랜덤포레스트 예측 결과 비교	85
〈그림 4-2〉 주요 외생변수(공동주택 수, 단독주택 수, 주거용 연면적, 건축물 수)의 2020년 참값과 2025년 예측값 관계	87
〈그림 4-3〉 주요 외생변수의 참값(2000~2020)과 모델1 2단계 예측값의 관계	88
〈그림 4-4〉 GPAI를 활용한 전국 1km 격자단위 장래인구 예측 결과(2020년~2050년)	90
〈그림 4-5〉 GPAI를 활용한 전국 1km 격자단위 장래인구 변화 전망(2020년 대비 2050년 인구증감률)	91
〈그림 4-6〉 2020년도와 2050년도 부산광역시, 울산광역시 및 경상남도 일대 인구분포 변화 ·	95
〈그림 4-7〉 2020년도와 2050년도 충청남도 및 세종특별자치시 일대 인구분포 변화 ·	96
〈그림 4-8〉 비수도권 인구감소 지정지역과 인구감소 비(非)지정지역의 인구변화 추세 ·	97
〈그림 4-9〉 2050년 장래인구 예측결과의 아웃라이어 현황	102
〈그림 4-10〉 2000년 대비 2020년, 2020년 대비 2050년 인구변화 증가율 아웃라이어 분포 ·	103
〈그림 4-11〉 행정구역 변화(2000년 대비 2021년) 및 공주-연가-세종-청주-청원 일대 행정구역 변화	107
〈그림 4-12〉 과거-현재-미래의 건축물수 및 총인구수 변화 추세	108
〈그림 4-13〉 세종특별자치시-공주시-청주시 일대 인구 변화(2000년~2020년)	109
〈그림 4-14〉 경기도 남양주시 택지개발 계획 구역(지구자정-붉은색, 실세계획-푸른색, 부분준공-노란색)	110
〈그림 4-15〉 그린피스가 예측한 해수면 상승에 따른 2030년 우리나라 침수 영향	112
〈그림 5-1〉 중층적 공간구조와 격자단위 공간 데이터 활용의 예시	118
〈그림 5-2〉 대구광역시 및 인접지역의 1km 격자단위 인구 분포와 경계 변화	119
〈그림 5-3〉 격자단위 장래인구예측을 활용한 부산 동구 작은도서관 폐점결정 시뮬레이션 예시	121
〈그림 5-4〉 연구성과(데이터 및 소스코드) 공개 예시	127



CHAPTER 1

서론

1. 연구 배경과 목적	3
2. 연구 범위와 방법	6
3. 선행연구 검토와 차별성	11
4. 기대효과	14

01 서론

소지역단위 장래인구 예측 방법은 인구감소시대에 적절한 계획 수립의 기초자료로 유용하게 활용될 수 있다. 이 연구의 목적은 인구학적 특성과 공간적 특성을 고려하여 인공지능 기반의 1km 격자단위 장래인구 예측 방법론을 개발하고 이의 활용방안을 제시하는 데 있다. 기존의 행정구역 단위의 장래인구 전망 방법론의 한계를 검토하여 상세한 격자단위 장래인구 예측 방법론의 차별성을 도출하였고, 방법론 개발과 그 결과를 활용함으로 예상되는 정책적·학술적 기대효과를 제시하였다.

1. 연구 배경과 목적

1) 연구 배경과 필요성

우리나라가 인구감소시대를 맞이함에 따라 지방소멸의 위기가 본격적으로 대두되고 있으며, 국민 대부분(88%)이 인구감소 위기를 체감하고 있다(인구보건복지협회, 2022). 그런데 국토 내 공간별로 겪고 있는 소멸의 위기 수준은 천차만별이다. 예를 들어, 수도권 일부 지역은 인구증가와 밀집에 따른 인프라 부족 등의 문제를 겪고 있지만, 지방 소도시는 생산인구의 급격한 감소로 지역경제 침체를 걱정하고 있다. 비단 수도권과 비수도권의 차이뿐만 아니라, 같은 지자체 내에서도 공간에 따라 인구증감의 격차가 큰 상황이다. 즉, 우리나라는 총인구 규모의 감소와 인구증감의 공간적 빈익빈 부익부(이보경, 2019) 현상을 겪고 있다.

그러나 국가와 지자체는 공간 맞춤형 전략보다는 시·군·구와 같은 행정구역 틀 안에서 지역을 진단하고 지원책을 마련해왔다¹⁾. 따라서 국가와 지자체에서는 ‘소멸의 위험에 직접 노출된 장소는 어디인지?’, ‘실제 이 공간에 적절한 인구감소 관련 지원이

1) 국토교통부·행정안전부 성장촉진지역, 행정안전부 인구감소지역 등의 경우 지역을 선정하고 지원하는 공간단위는 시·군·구이고 그 기준은 장래의 인구변화가 아니라 과거의 변화 및 현상을 나타내는 회고적 지표임

이루어지고 있는지?’, ‘지금까지 투입된 예산이 인구감소에 효과가 있었는지?’ 등 공간의 관점에서 발생하고 있는 인구감소의 이슈에 답하지 못하고 있다.

인구감소시대에 대응하고 적응해나가기 위해서는 데이터기반의 과학적인 증거를 바탕으로 적절한 계획과 정책을 수립해야한다. 윤석열 정부의 지역균형발전 국정과제에서는 인구감소에 따른 소멸 위기를 겪고 있는 지역에 대해 효율적인 자원 배분 및 인프라 경영·관리 등 적절한 정책 및 계획을 수립하고 시행할 수 있는 역량 강화를 요구하고 있다(제20대 대통령직인수위원회, 2022. 4. 27). 지자체의 기획과 경영역량을 높이기 위해서는 이들이 활용할 수 있는 과학적 근거자료와 계획수립의 도구가 필요한데, ‘상세한 공간단위(소지역)’의 ‘장래의 인구’ 정보는 미래예견적 계획 수립을 통해 지역의 지속가능성을 도모하는 데 중요한 역할을 한다. 같은 맥락에서 2022년도 말에 저출산고령사회위원회·관계부처에서 발표한 ‘인구구조 변화와 대응방안’에서도 데이터와 과학에 기반하여 정책의 효과성을 검증할 수 있도록 인구정책을 추진하도록 가이드하고 있다. 고해상도의 소지역단위 장래인구를 활용하면 앞으로 어느 공간이 인구감소시대에 가장 취약할지 전망하고, 이들을 정밀하게 타겟팅하여 장소기반·증거기반의 정책을 수립할 수 있을 것으로 기대된다.

과거에는 데이터 및 분석 환경이 뒷받침되지 않아 고해상도의 장래인구를 예측하는데 어려움이 있었다. 이제는 장래인구 예측에 활용될 수 있는 공간 빅데이터가 많아졌고, 분석 컴퓨팅 환경이 발전함에 소지역단위 인구예측 방법론 개발이 가능해졌다(이보경, 2019). 데이터 관점에서는 장래인구를 전망하기 위해서 소지역단위로 기준인구 설정 및 인구변화 추세를 파악해야 한다. 최근 국토지리정보원의 국토조사와 통계청의 통계지리정보서비스(SGIS)²⁾에서 100m, 1km 등의 격자단위 인구통계를 제공함에 따라 관련 자료를 확보할 수 있게 되었다. 그리고 대용량의 상세한 공간단위 정보를 학습하여 예측할 수 있는 인공지능 기법과 컴퓨팅 환경이 발전하였다. 잘 학습된 정보를 바탕으로 작동하는 인공지능 모형은 수만 건에 달하는 복잡한 계산식을 단 수 초 만에 계산하는 등 시간과 공간의 물리적인 한계를 극복할 수 있다.

2) 통계지리정보서비스(SGIS)(<https://sgis.kostat.go.kr/>, 최종접속일: 2022.11.22.)

2) 연구 목적

이 연구의 목적은 인공지능 기법을 적용하여 소지역단위 장래인구 예측모델을 개발하는 것이다. 나아가 향후 국토·지역·도시 정책 및 계획에서 소지역단위 장래인구 예측 결과를 활용할 수 있는 방안을 모색하고자 한다. 연구는 다음과 같은 순서로 진행되었다. 첫째, 소지역단위 장래인구 예측 관련 국내·외 연구동향과 사례를 조사하고 방법론 개발방향을 정립하기 위한 시사점을 도출하였다. 둘째, 다양한 소지역 공간범위 중 1km 격자단위 장래인구 예측 방법론인 ‘격자단위 장래인구 예측 AI 모델(Gridded Population Forecast with AI, 이후 GPAI)’을 개발하였다. 소지역은 활용 목적과 시대에 따라 다양한 공간범위로 정의될 수 있다. 사례 조사 결과와 정책적 활용성을 고려하여 이 연구에서는 1km격자를 소지역으로 채택하였다. 그리고 인구학적 특성과 공간적 특성 모두 고려할 수 있는 방법론을 개발하기 위해 인공지능 기반 모델과 코호트요인법³⁾ 기반 모델을 결합하는 아키텍처로 GPAI를 구성하였다.

셋째, GPAI를 전국에 적용하여 2050년까지 인구분포 변화를 전망하였다. 예측 결과를 통해 어느 공간의 인구가 감소할지, 또는 인구가 증가하는 지역은 어디인지 그리고 지자체별로 그 현황은 어떻게 다른지 비교하였다. 이를 통해 격자단위 장래인구를 활용하여 정책적 관심지역인 인구 증가지역, 인구 감소지역, 무거주화 위험지역 등을 도출하고 공간 맞춤형 계획을 수립할 수 있는 가능성을 시사하였다.

넷째, GPAI가 실무에서 신뢰성있게 활용되기 위해 필요한 방법론 고도화 방안, 활용 예시 및 활용 확산 전략 등을 제안하였다. 우선 GPAI 적용 결과를 면밀하게 검토하여 방법론의 고도화 방안과 후속 연구 수행 방향을 도출하였다. 그리고 적절한 자원 배분, 인프라 계획 및 국토공간계획 수립에 등 관련 정책 및 계획 수립에 활용할 수 있는 구체적인 예시를 제시하였다. 또한 오픈이노베이션을 통한 지속적인 GPAI의 고도화를 위해 깃허브(GitHub)와 같은 플랫폼 활용방안을 제안하였다.

3) 코호트 요인법은 “기준인구에 출생아 수와 국제순이동은 더하고, 사망자 수는 빼는 인구균형방정식을 적용하여 다음 해 인구를 반복적으로 산출해 나가는 인구전망 방법이다.”(김경수·김상미, 2021: 47)

2. 연구 범위와 방법

1) 주요 개념과 연구 범위

(1) 주요 개념과 내용적 범위

① 장래인구 전망의 개념과 우리나라 현황

장래인구 전망은 ‘추계(projection)’와 ‘예측(forecast, prediction)’으로 구분된다. “추계(projection)는 인구변동과 관련된 일련의 가정에 기초한 장래인구의 규모와 구조에 관한 진술”이며, “일련의 가정들이 충족될 때 실현되는 장래인구에 대한 조건적 전망”이다(Smith et al, 2001, Wilson·Rees, 2005, 우해봉, 2009: 74 재인용). 반면, “예측(forecast)은 현재의 과학적 지식에 기초하여 연구자가 제시하는 실현 개연성이 가장 높은 장래인구의 규모와 구조에 관한 진술”로 정의할 수 있다. “예측은 항상 추계에 해당하지만 그 역은 성립하지 않”는다(우해봉, 2009: 74). 대부분의 사용자는 장래인구추계와 예측 결과를 구분하지 않고, 통계청에서 공표하는 장래인구추계를 장래인구 전망치로 생각하고 활용하고 있다.

이 연구의 ‘장래인구 예측 모델’은 주어진 조건과 가정에 따라 실현되는 장래인구를 전망한다는 면에서 추계의 성격을 갖추고 있으나, 인공지능 기법을 활용하여 정확도를 고려한다는 면에서 협의의 예측 모델로 볼 수 있다.

② 소지역

인구와 관련된 연구에서 ‘소지역’에 대한 보편적 정의는 없으나, 목적에 따라 데이터 확보가 가능한 지역 분류에서 가장 세밀한 공간단위를 의미한다(Smith·Morrison, 2005: 2). 인구 관련 연구에서 고려된 해외의 소지역단위는 미국의 센서스 트랙(census tracts), 호주의 Statistical Areas Level2, 영국의 워드(ward)와 집계구(census super output areas), 일본의 도시/농촌 블록(city/rural block) 등이 있다(강영욱·장세진, 2008; 이영주·이보경·오창화, 2020).

우리나라의 경우 시간이 지나면서 조사 방법, 데이터 처리 관련 기술과 분석 환경의 개선, 제도 마련, 역량 강화 등에 의해 ‘소지역’에 대한 정의가 시·군·구, 읍·면·동, 집계구, 격자 등 점점 좁은 공간 범위로 조밀해지고 있다. 최근 행정구역단위 통계의 경우 경계의 시계열 변화로 인해 정보의 누락 및 일관성 확보가 힘들고 다양한 공간정보와의 융합이 어려운 문제점이 대두되고 있다. 이에 따라 국토지리정보원 국토조사, 통계청 SGIS 등 공공부문과 이동통신사와 같은 민간부문에서 격자단위 통계 생산과 활용이 증가하는 추세이다. 특히 제3기 인구정책 TF에서 “통계청이 생산한 인구구조 변화와 관련된 통계를 격자(格子) 단위 공간정보와 결합하여 정밀한 융·복합 데이터 생산”할 필요성을 제기하였다(관계부처 합동, 2021. 1. 27. : 12).

이와 같은 배경에서 이 연구에서는 1km⁴) 격자단위의 장래인구 예측 모델을 개발하였다. 제2장에서 격자를 포함한 다양한 공간단위의 소지역에 대한 관련 연구 동향과 사례 조사를 시행하였고, 제3장부터는 소지역을 1km 격자에 한정하여 방법론을 개발하고 정책적 활용 방안을 도출하였다.

③ 예측 대상

예측 대상 장래인구는 1km 격자 내 거주하는 총인구수⁵⁾다. 통계청 공표 장래인구 추계에서는 성별/연령별 인구수를 제시하고 있으나, 주어진 연구 기간과 정보를 고려하여 이 연구에서는 총인구수에 초점을 맞추어 GPAI를 개발하였다.

4) 공간의 정밀성을 확보하면서도 오차를 최소화하여 모델의 효용성을 극대화할 수 있도록, 읍·면·동, 시·군·구 등 행정구역보다 작으면서 국토공간계획에 활용할 수 있는 공간단위인 소지역을 1km 격자로 정의하였으며, 자세한 설명은 제2장 참조

5) 인구밀도(명/km²)

(2) 시간적 범위

이 연구에서는 통계청 공표 시도단위 장래인구추계와 정합성을 갖기 위해 시간적 범위를 2025년부터 5년 주기로 2050년까지(2025년, 2030년, 2035년, 2040년, 2045년, 2050년)로 설정하였다. 통계청은 5년 주기로 전국단위 향후 50년의 장래인구추계와 시도단위의 향후 30년의 장래인구추계를 작성하여 공표하고 있다. 가장 최신의 통계청 공표 장래인구추계는 2022년 5월에 공개한 2050년까지 시도단위 장래인구추계⁶⁾이며, 이 연구에서도 통계청 장래인구추계와 정합성을 갖추기 위해 2050년을 시간적 범위로 설정하였다. 그리고 국토종합계획, 광역도시계획, 도시·군기본계획 등 국토공간계획은 20년 주기로 수립하고 있다. 이때 연도의 끝자리가 0 또는 5년이며 5년마다 계획의 타당성을 전반적으로 재검토하여 정비하고 여건 변화에 따라 일부 조정한다. 따라서 이 연구에서도 2025년부터 2050년까지 5년단위로 장래인구를 예측하였다.

(3) 공간적 범위

이 연구의 공간적 범위는 전 국토를 대상으로 한다. 전 국토를 1km 격자 크기로 구획한 소지역단위를 인구예측의 공간단위로 설정하였다. 과거로부터 현재까지 소지역(1km 격자)단위로 변해온 전 국토의 인구분포 데이터와 공간적 특성을 학습하여 GPAI를 개발하였다. 그리고 이 방법을 전 국토에 적용하여 2050년까지 소지역단위 인구분포를 전망하였다. 이를 통해 전국 약 10만 개 소지역(1km 격자)의 2000년부터 2050년까지 인구변화를 분석하였다.

6) 2019년도 3월에 공표한 장래인구추계는 정책수요에 따른 특별추계이며, 2021년 12월 정기 장래인구추계(2020~2070년)와 2022년 5월 장래인구추계(시도편):2020~2050년을 발표

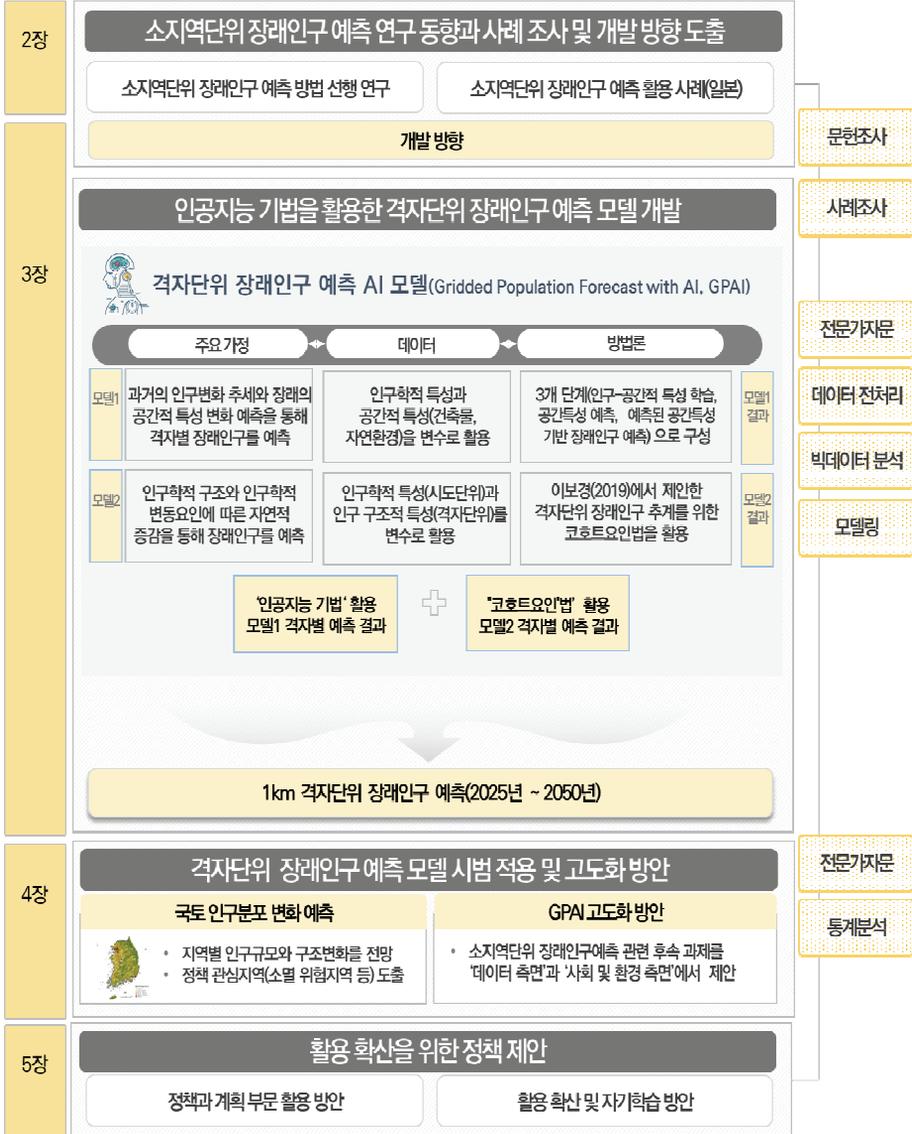
2) 연구 방법과 흐름

소지역단위 장래인구 예측 방법론 개발을 위한 연구 방법과 흐름은 다음과 같다(그림 1-1) 참조). 먼저, 문헌 및 사례 조사를 통해 국내외에서 진행된 선행연구와 일본의 격자단위 장래인구 예측 및 활용 사례를 검토하였다(제2장). 문헌 및 사례 조사 결과를 바탕으로 전문가 자문을 실시하고 우리나라에서 구득할 수 있는 관련 데이터를 검토하여 소지역단위 장래인구예측 모델의 개발 의의와 방향을 설정하였다. 이를 바탕으로 소지역단위 장래인구 예측을 위한 데이터와 방법론의 아키텍처를 구성하였다. 아키텍처는 인공지능 기반 모델(모델1)과 전통적인 코호트요인법 기반 모델(모델2)로 이루어졌다. 모델1은 과거부터 현재까지 인구수뿐 아니라 인구분포와 밀접하게 관련된 공간적 특성(건물, 주택, 물리적 환경 등)을 예측하여 이를 외생변수로 장래인구를 예측한다. 그리고 모델2는 인구감소 추세를 반영할 수 있는 인구학적 변동요인을 세밀하게 시뮬레이션하여 장래인구를 예측한다. 이 두 개의 모델을 앙상블⁷⁾로 결합하여 최종 소지역단위 장래인구를 예측하였다. 이 과정에서 다양한 출처의 공간 빅데이터를 수집하여 전처리 과정을 거쳐 1km 격자단위의 데이터로 가공·융합하고 학습에 활용하였다(제3장).

학습된 GPAI 모델을 전국의 1km 격자에 적용하여 2050년까지의 인구변화를 탐색하고 거주 공간이 어떻게 변할지 전망하였다. 인구변화에 따라 지역을 유형화하여 유형별 주거환경(주요 인프라 접근성)을 통계분석을 통해 비교하고 정책적 관심 사항을 도출하였다. 특히 모델을 고도화시키기 위해 GPAI 초기 모델의 한계를 정리하고 이를 극복하기 위한 향후 과제를 제시하였다(제4장). 마지막으로 연구의 결과를 전문가 세미나와 학회를 통해 공유하고 전문가와의 토론을 수행하였다. 이 과정에서 향후 정책과 계획에서 어떻게 활용할 수 있을지, 이 모델이 계속해서 발전하고 누구나 쉽게 활용할 수 있을지에 대한 전략을 논의할 수 있었다. 마지막 장에서는 정책적 활용방안, 활용 확산 전략 및 시사점을 도출하였다(제5장).

7) 앙상블(Ensemble)은 데이터에 기반한 알고리즘의 과대 또는 과소적합을 방지하기 위해 다양한 알고리즘이 조합을 이루어 최종값을 도출하는 데이터마이닝 방법임

그림 1-1 | 연구 흐름



자료: 저자 작성

3. 선행연구 검토와 차별성

1) 선행연구 현황

통계청에서 공표하는 장래인구추계단위인 전국 및 시·도단위보다 작은 공간단위로 장래인구를 예측하는 방법론을 제안한 연구는 <표 1-1>과 같다. 우리나라 소지역단위 장래인구 전망 관련 연구는 대부분 권역 또는 시·군·구 공간범위를 대상으로 코호트요인법 기반으로 진행됐다. 민성희 외(2016)는 시·군·구 간 인구이동 매트릭스에 마르코프 연쇄모형을 적용하여 다지역 인구이동을 고려한 코호트요인법을 제안하였다. 읍·면·동 단위의 인구예측 방법론을 제안한 연구도 있는데, 최현정·최석환·홍성조(2019)는 Mankiw-Weil 모형을 활용하여 면적별 주택 수 기반의 읍·면·동 단위 인구예측 방법론을 개발하였다.

소지역단위 장래인구에 대한 수요가 증가함에 따라 국내외에서 격자단위로 인구를 예측하는 방법에 관한 연구가 최근 진행되기 시작했다. 임은선 외(2017)는 격자단위 주택용 건물비중을 활용하여 2040년까지 우리나라 장래인구를 500m 격자단위로 예측하였다. 격자별 인구의 연령 구조와 건물의 현황을 가지고 시나리오 기반으로 인구 소멸지역, 유지지역, 증가지역을 구분하였다. 장래인구추계에 전통적으로 활용된 코호트요인법을 격자단위에 적용한 선행연구도 있는데, Breidenbach·Kaeding·Schaffner(2019)는 난민 등 해외에서의 인구 유입을 과거추세를 바탕으로 예측하고 이를 코호트요인법에 적용하여 독일의 장래인구를 1km 격자단위로 예측하였다. 이보경(2019)은 500m 격자단위 장래인구 전망을 위해 코호트요인법을 활용하여 통계청 공표 시도단위 장래인구전망을 재할당하는 방법론을 제안하였다.

표 1-1 | 선행연구 현황 및 차별성

구분	선행연구와의 차별성		
	연구목적	연구방법	주요 연구내용
주요 선행연구	1 제목: 민성희 외(2016). 국토계획 수립지원을 위한 인구분석 방법 연구 목적: 소지역 인구분석 방법을 설정	<ul style="list-style-type: none"> • 문헌 고찰 및 사례 검토 • 통계 분석 및 사례 적용 • 전문가 의견 수렴 	<ul style="list-style-type: none"> • 국내외 인구추계 방법론 검토 • 소지역단위(시군구) 인구추계 방법 개발 및 적용 • 국토계획을 위한 인구추계방법 활용 방안을 모색
	2 제목: 최현정·최석환·홍성조(2019). 읍면동 단위 장래인구 추정모형 개발에 관한 연구 목적: 도시군기본계획 계획인구를 읍면동단위로 배분할 수 있는 방법론을 개발	<ul style="list-style-type: none"> • 이론적 배경 검토 • 분석모형 구축 및 적용 	<ul style="list-style-type: none"> • 도시계획과 부동산 정보를 활용하여 주택과 인구 변화간의 관계를 도출 • 주택과 인구 변화간의 관계를 바탕으로 장래의 주택개발에 따른 인구 변화를 읍면동 단위로 예측
	3 제목: 임은선 외(2017). 빅데이터 시대의 국토정책 추진방향 목적: 빅데이터 시대를 맞이하여 실증기반의 국토정책을 추진하기 위해 빅데이터를 활용하는 방안을 모색	<ul style="list-style-type: none"> • 계획수립 시 데이터 활용 여건과 실태조사 및 관계기관 의견 수렴 • GIS, 통계분석 등 • 해외 사례 검토 • 빅데이터기반 국토정책 추진을 위한 전략 도출 	<ul style="list-style-type: none"> • KCB 빅데이터를 활용하여 혁신도시 여건 변화를 모니터링 • 500m 격자 지표로 추정된 장래 시점의 인구지표와 인구감소에 따른 생활기반시설 폐점 시나리오를 적용 • 국토정책 혁신 전략과 추진과제 도출
	4 제목: 이보경(2019). 2040년 장래인구 분포 전망 연구 목적: 통계청의 장래인구 추계와 정합성을 가지는 소지역단위(500m 격자) 장래인구 분포를 전망	<ul style="list-style-type: none"> • 소지역단위 장래인구 전망 방법론 검토 • 코호트요인법을 활용한 장래인구 분포 가치치 도출 및 적용 	<ul style="list-style-type: none"> • 시도 단위 장래인구추계 결과를 소지역에 할당하여 2040 미래 국토인구분포 도출 • 2040 미래 국토의 격자단위 인구분포를 2019 국토의 격자단위 인구분포와 비교하여 지역별 인구증감 수준 및 인구 소멸 위험지역을 전망
	5 제목: Breidenbach·Kaeding·Schaffner(2019). Population projection fo Germany 2015-2050 on grid level. Jahrbucher fur Nationalokonomie und Statistik 239 목적: 독일 장래인구를 1km 격자단위로 전망	<ul style="list-style-type: none"> • 소지역단위 장래인구 전망 방법론 검토 • 해외유입 인구에 대한 전망 방법론 검토 및 개발 • 코호트요인법을 활용한 장래인구 전망 	<ul style="list-style-type: none"> • 코호트요인법을 활용하여 독일의 장래인구를 1km 격자단위로 전망 • 난민 유입 등 해외에서 인구 유입을 과거 추세를 바탕으로 예측하고 현재 인구 규모 비중으로 할당하여 방법론을 개발
본 연구	제목: 인공지능 기법을 적용한 소지역단위 장래인구 예측 방법론 개발 기초 연구 목적: 우리나라 인구데이터에 적용할 수 있는 장래인구 전망 방법론 개발	<ul style="list-style-type: none"> • 최신의 소지역단위 장래인구 전망 방법론을 검토 • 관계부처와의 협업을 통해 우리나라 인구데이터 현황과 이를 활용한 전망 방법 개발을 시도 	<ul style="list-style-type: none"> • 문헌조사 및 전문가 자문을 통해 우리나라 인구데이터에 활용할 수 있는 전망 방법론 개발 • 소지역단위 장래인구 전망을 통해 지역의 변화를 유형화하고 정책적 시사점을 도출

자료: 저자 작성

2) 본 연구의 차별성

행정구역단위 장래인구 전망과 관련된 선행연구의 한계는 다음과 같다. 첫째, 시·도 및 시·군·구는 지자체 내 공간계획을 수립하기에 범위가 너무 넓다. 공간계획을 수립하기 위해서는 인프라 및 시설 주변의 수요를 추정해야 하는 경우가 많은데, 이와 같은 장소기반 의사결정을 위해서는 소지역단위 인구분포 정보가 필요하다. 물론 읍·면·동 단위의 경우 충분히 지역의 특성을 반영한 소지역단위로 볼 수 있으나 2050년까지 계획을 수립하는 데 있어 행정구역의 변화가 빈번하여 과거 데이터를 축적·학습하는데 한계가 있다. 둘째, 사망률, 출생률, 인구순이동 등의 시계열 추세를 반영하여 급변하는 지역의 현황을 다양한 시나리오로 접근하는 것이 불가능하다. 코호트요인법에서는 공간의 특성을 고려하지 않기 때문에 행정구역이라는 큰 범주에서 총량만 전망할 수 있고, 토지이용 변화에 따른 인구의 공간적 분포를 예측할 수는 없다는 한계가 있다. 전통적인 코호트요인법 기반의 행정구역 단위 장래인구추계의 한계로, 소지역단위 인구예측 분야는 공공과 민간 분야에서 수요가 매우 높지만 그 필요성에 비해 극히 제한적으로 연구됐다.

이 연구의 주요 차별점은 다음과 같다. 첫째, 기존 연구에서 행정구역단위 인구전망 방법을 제안하였다면 본 연구는 1km 격자단위로 인구를 전망한다. 임은선 외(2017)과 이보경(2019)도 격자단위의 장래인구 예측 방법론을 개발하였으나, 토지이용의 변화와 같은 공간적 특성을 충분히 반영하지 못했다는 한계가 있다. 이를 극복하기 위해 시·공간의 복잡하고 대용량의 과거 데이터를 학습할 수 있는 인공지능 기법을 적용하였다. 또한 기존 장래인구 예측 방법론이 과거 데이터 기반으로 추정된 파라미터를 제시하는 것에 그쳤다면, 본 연구에서는 예측모델의 활용을 위한 정책적 제언과 활용 확산 방안을 제시하는데 차별성이 있다.

4. 기대효과

소지역, 특히 격자단위 장래인구 예측 방법론과 적용 결과는 인구감소시대에 적절한 계획 및 정책을 수립하기 위한 기초자료로 활용될 수 있다. 2021년도 제3기 범부처 인구정책 TF에 따르면, ‘축소사회에 대응’과 ‘지역소멸에 선제 대응’을 위해서는 미래의 바람직한 발전 방향을 설정하는 지침 성격 계획과 일상생활이 영위되는 공간에 대한 장소기반 전략이 적합성과 연계성을 가져야 한다(관계부처합동, 2021). 즉, 생활권과 같은 작은 공간부터 광역권까지 관통할 수 있는 데이터가 필요한데, 1km 격자단위의 장래인구 예측 결과가 기초자료로 유용하게 활용될 수 있다.

격자단위 장래인구 정보를 활용하면 미래 공간구조를 유연하게 설정할 수 있다. 인구증감의 공간별 빈익빈 부익부에 따라 현재 행정구역 단위에 의존한 의사결정이 미래 변화 여건을 충분히 반영하지 못할 가능성이 크다. 저출산고령사회위원회·관계부처에서 공표한 ‘인구구조 변화와 대응방안’(2022. 12. 28.)에서도 공간 빅데이터를 활용하여 행정구역이 아닌 광역권 또는 생활권 단위의 인구정책을 추진하도록 가이드하고 있다. 격자단위 장래인구 정보를 활용하면 정책적 관심에 따라 자유롭게 구역을 설정하고 의사결정에 참조할 수 있을 것이다. 그리고 격자단위 장래인구 분포를 통해 어느 지역이 고립되거나 인구가 집중되어 혼잡비용이 발생하는 등의 문제를 사전에 예측할 수 있다. 따라서 관련된 이슈가 발생하지 않도록 미래예견적인 대응 방안을 마련할 수 있을 것이다. 특히 제한된 자원을 가장 효과적으로 배치하고 운영하는데 근거자료로 활용될 수 있을 것이다. 소지역단위 장래인구예측 자료를 활용하여 장래 변화를 시뮬레이션하고 다양한 인프라 또는 사업 시행 여부를 과학적으로 결정할 수 있을 것으로 기대된다.

학술적 측면에서 이 연구의 기대효과는 인공지능 기법을 적용하여 격자단위 장래인구 예측의 가능성을 시사함으로 관련된 후속 연구 수행의 기틀을 마련한 데 있다. 우선, 공간적 특성을 반영한 인공지능 기반 모델과 인구학적 특성에 따른 코호트요인법 기반 모델을 앙상블(ensemble)⁸⁾을 통해 결합하는 접근방식을 제안하였다. 앞으로 전

통적인 코호트요인법에 다양한 방법론을 결합하여 각 방법론이 가지는 장점을 극대화하고 한계를 보완하는 형태의 새로운 접근방식의 연구가 수행될 수 있다.

또한 이 연구에서는 행정구역에 구애받지 않는 격자단위로 공간 데이터를 가공·융합하고 장래인구 예측에 활용할 수 있는 구체적인 예시를 제시하였다. 그리고 격자단위 공간통계의 특징인 이웃효과를 인공지능 기법에서 고려하여 학습하고 격자단위 인구를 예측할 수 있도록 아키텍처를 설계하였다. 이를 통해 복잡한 시·공간 위계를 가진 대용량 데이터를 인공지능 기법으로 학습하는 방법을 개발하여 관련 방법론의 고도화에 기여하였으며, 향후 이 아키텍처를 기반으로 모델을 진화시킬 수 있을 것으로 기대된다.

이 연구에서는 방법론 개발에 행정안전부, 국토교통부, 농림축산식품부, 국토지리정보원, 통계청 등 모두에게 공개된 데이터와 R의 인공지능 기법 오픈소스를 활용하였다. 따라서 누구나 같은 출처의 데이터와 알고리즘 라이브러리를 활용하여 연구 결과를 재현할 수 있을 것으로 기대된다. 특히 다양한 데이터 전처리 방법과 알고리즘 결합 방법을 깃허브와 같은 오픈소스 플랫폼을 통해 공개한다면, 향후 관련 연구를 수행하거나 플랫폼을 구축해야하는 전문가(공공 부문 지원 민간 사업자, 스타트업, 연구자)가 좀 더 쉽게 인공지능 기법을 활용한 격자단위 장래인구 예측 모델을 발전시켜 적용할 수 있을 것이다.

8) 앙상블(Ensemble)은 단일 모델을 활용하였을 때 보다 더 나은 예측 성능을 위해 다양한 알고리즘을 결합하는 방법



CHAPTER 2

소지역단위 장래인구 예측 연구 및 사례 조사

1. 소지역단위 장래인구 예측 방법론 연구 동향 19
2. 격자단위 인구통계 관련 사례 조사 36
3. 시사점 44

02 소지역단위 장래인구 예측 연구 및 사례 조사

본 장에서는 소지역단위 장래인구예측 방법론을 개발하기 위해 관련된 선행된 방법론 개발 연구를 검토하여 각 방법론의 장점과 한계를 정리하였다. 그리고 전 세계 격자단위 인구 통계 현황, 격자단위 공간빅데이터 분석 방법론 동향 및 일본의 격자단위 장래인구 예측 방법론 개발과 활용 현황 사례를 조사하였다. 이와 같은 선행 연구 및 사례 조사를 통해 장래인구예측 방법론 개발을 위한 시사점을 도출하였다.

1. 소지역단위 장래인구 예측 방법론 연구 동향

최근 활용 가능한 데이터와 방법론의 발전으로 소지역단위 인구에 관한 관심이 증가하고 있으며, 특히 인공지능 기법을 적용한 장래인구 예측 방법론이 제안되기 시작하였다. 국내외 소지역단위 인구예측 관련 선행연구를 방법론 관점에서 기초모델, 코호트요인법, 공간통계모델, 마이크로시뮬레이션, 도시성장모델, 인공지능 기법으로 분류하여 검토하였다.

1) 장래인구 전망 기초모델

소지역단위 인구예측 방법론 중 가장 간단한 모델인 기초모델에는 경향외삽법(extrapolative model)과 비율법(ratio method)이 있다. 경향외삽법은 이론보다 과거의 추세를 기반으로 한다. 최소한의 데이터만을 요구하고 계산이 단순하며, 직관적으로 이해된다는 장점이 있지만, 연령구조 및 성별 인구를 예측하는 데 한계가 있다. 비율법은 권역(또는 소지역이 이루고 있는 광역의 공간단위)과 소지역의 관계를 바탕으로 광역단위의 추계인구를 소지역단위로 할당하는 방법이다.

Rayer(2008)이 미국의 2,482개 카운티 인구를 예측하는데 다양한 경향외삽법과 비율법을 적용하여 비교를 수행하였다. 그 결과 정확도가 높은 방법은 양의 인구성장률을 고려한 비율법으로 나타났다. 하지만 예측 기간이 30년 이상으로 증가할수록 오차가 매우 커졌으며, 선형/지수 모델과 같은 추세법에서는 지역의 인구가 음수로 추정되는 등 실제와 다른 큰 오차를 보이기도 하였다.

기초모델 중 매개변수를 활용하는 방법으로 주택정보 활용 모델(housing-unit model)이 있다. Baker et al. (2008)은 신규 택지개발 계획과 주택 재고에 로지스틱모형을 적용하여 주택 수를 예측하고, 이를 근거로 미국 뉴멕시코의 census 트랙 공간단위의 장래인구를 전망하였다. Hauer·Evans·Alexander(2015)는 미국 조지아주의 sub-county 단위 인구를 예측하기 위해 선형/지수모델을 활용하여 70년간의 주택 수를 예측하였다. 이 모델은 선형/지수모델로 직접 인구를 예측할 때 소지역 내 인구가 음수로 추정되는 오류를 방지할 수 있으며, 소지역단위로 예측된 인구를 통해 지구단위 계획을 조정하고 시뮬레이션할 수 있다는 장점이 있다.

국내에서는 건축물 용도별 연면적 및 토지이용 현황을 활용하여 소지역단위 인구를 추정하는 모형연구가 극히 제한적으로 진행되고 있다. 주거용 건물 분포를 활용하여 장래의 인구를 예측하기 위한 연구(임은선 외, 2017; 최현정·최석환·홍성조, 2019)와 읍·면·동 또는 시·군·구 인구를 더 작은 공간단위의 분포로 이해하기 위해 도시환경변수를 사용한 연구(이창효·이승일, 2006; 최내영, 2010; 이석준 외, 2014) 등이 있다.

이창효·이승일(2006)은 집계 인구 관련 자료를 활용하여 미시적 공간단위인 대지(건축물)로 배분하기 위한 가상 인구분포 추정 모델을 제안하였다. 인구분포 추정을 위하여 건축물대장, 토지대장 등의 속성자료를 분석하였으며, 건축물 속성 기반 대지를 미시적 공간 단위로 설정하였다. 이석준 외(2014)는 모든 인구가 주거용 건축물에 거주한다고 했을 때, 서울시 총인구를 세밀한 공간 단위로 배분하는데 건축물대장 정보를 활용한 대시매트릭매핑¹⁾방법을 활용하였다. 또한, 2010년도 센서스 인구의 집계

1) 대시매트릭매핑은 공간 보간법(Areal interpolation)을 통해 데이터를 분해(Data de-aggregation)하는 것으로, 하나의 권역 또는 상위의 공간단위에서 좀 더 작은 단위의 하위 공간단위로 데이터를 전환하는 과정임.

구단위 분포를 추정하기 위해 건축물대상상 주거용 건물의 연면적과 층수 정보를 활용하였으며, 25개 자치구에 대한 예측 회귀분석의 R^2 가 평균 0.771로 도출되는 등 매우 설명력이 높은 것으로 나타났다.

최현정·최석환·홍성조(2019)는 읍·면·동단위로 인구를 예측할 수 있는 방법론을 개발하고자 하였으며, 수원시를 대상으로 적용 가능성을 검토하였다. 읍·면·동단위 표본 데이터를 활용하여 면적별 주택수와 연령별 인구수의 관계를 밝히고, 주택-인구모형을 통해 대상지역인 수원시의 사회적 인구 증가분을 산정하였다. 장래 사회적 인구 증가분은 경기도의 도시·군기본계획 지침에 따른 외부유입률을 활용하였다. 이와 같은 과정을 거쳐 예측된 인구를 주택-인구모형을 통해 다시 동단위로 배분하여 장래인구를 전망하였다.

임은선 외(2017)는 건축물 데이터를 활용하여 500m 격자단위로 인구자료를 배분하고 장래인구를 추정하였다. 이 연구는 주거용 건물 비율 값을 가중치로 사용해 2040년까지의 인구추정 값을 격자별로 분할하는 방식을 사용하였다. 2차적으로 읍면동별 초고령자(85세 이상) 비율이 상위 68% 이상에 해당하고, 500m 격자별 인구가 적은 지역은 2040년에 새로운 이주자가 없을 것이라는 가정하에 시나리오를 적용하여 인구를 추정하였다.

이와 같은 기초모델은 적용이 간편하고 이해가 쉽다는 장점이 있으나 다음과 같은 한계도 존재한다. 첫째, 지역이 작아질수록 예상하지 못한 요인으로 인해 인구규모와 인구구성의 변화 추세가 쉽게 바뀔 수 있다. 광역단위 인구를 예측할 때는 예측 대상 인구규모가 크기 때문에 지금까지의 거시적인 추세가 적용될 가능성이 높다. 그러나 공간단위가 조밀해질수록 인구규모에 영향을 주는 요인이 다양해지고 개개인의 의사결정에 큰 영향을 받기 때문에 과거의 추세에서 벗어날 확률이 크다. 둘째, 기초모델은 인구의 성별 연령별 구성과 정교한 토지이용을 고려하지 않기 때문에 장래의 인구규모가 왜 그렇게 변하는지에 대한 설명에 부적합하다는 한계가 있다. 따라서 의사결정자 입장에서는 어떤 계획/정책을 수립하고 시행했을 때 그 효과가 어떻게 지역에 반영될지 설명하지 못한다면 해당 방법론의 활용 범위는 한정적일 수밖에 없다.

2) 코호트 생존율 활용 모델: 코호트요인법 및 Hamilton-Perry모델

코호트요인법은 20세기 이후 가장 보편적으로 활용되는 인구추계 방법론으로 알려져 있다(Preston·Heuveline·Guillot, 2001: 119-120, 우해봉 외, 2016: 33 재인용). 또한, 앞서 소개한 기초모델의 한계를 극복하고자 출생, 사망, 이동이라는 인구변동요인의 복잡성과 각각의 요인들의 연계성을 고려하여 인구균형방정식을 만들어 장래인구를 추계하는 방법이다. 코호트요인법은 인구변동요인별로 그 수준을 시나리오로 설정하여 다양한 범위의 장래인구추계 제시가 가능하다. 과거에는 단순히 기준인구에 출생자 수, 사망자 수, 전입/전출자 수를 고려한 회계방식에서 머물렀으나, 최근 확률적 접근법도 적용되어 고도화되고 있다.

[코호트 요인법 인구균형방정식 예시]

$$Pop_t = Pop_{t-l} + Birth_{t-l,t} - Death_{t-l,t} + Migration_{t-l,t} \quad (1)$$

Pop_t : t 년도 인구

$Birth_{t-l,t}$: $t-l$ 년도에서 t 년도까지 출생아 수

$Death_{t-l,t}$: $t-l$ 년도에서 t 년도까지 사망자 수

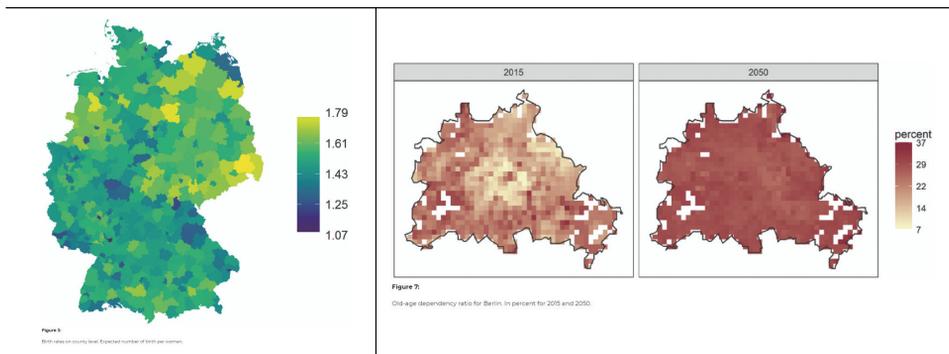
$Migration_{t-l,t}$: $t-l$ 년도에서 t 년도까지 순이동자 수

소지역단위 장래인구추계에는 코호트의 생존율을 고려하여 코호트요인법을 변형한 Hamilton-Perry 모델(Hamilton·Perry, 1962)이 활용되고 있다. Hamilton-Perry 모델은 기존에 축적된 센서스 상의 인구변동요인 정보를 활용하지 않고, 코호트별 인구 변화 비율이 일정하게 지속된다는 가정에 따라 두 시점의 연령별 인구규모의 변화율(cohort change ratio, 이후 CCR)을 활용하여 장래인구를 예측한다. Hamilton-Perry 모델은 소지역단위의 시계열 데이터가 충분하지 못할 때, 코호트요인법의 장점을 유지하면서 간단하게 적용할 수 있다는 장점이 있다. 하지만, 따라서 인구를 선형적으로 예측한다는데 코호트요인법의 차이가 급격한 추세 변화에 지나치게 높거나 낮은 인구 규모를 예측하는 단점을 지니고 있어, 상한치와 하한치 그리고 조금 더 큰 광역권의 총인구 규모와 같은 컨트롤 변수가 필요하다.

Swanson·Schlottmann·Schmidt(2010)은 CCR을 적용하여 센서스 트랙(census tracts)단위로 장래의 인구 변화를 예측하고, 이후 소지역단위의 인구규모가 너무 크거나 작지 않도록 상한선과 하한선 기준을 마련하여 소지역단위 장래인구추계를 조정하였다. Baker·Swanson·Tayman(2021)는 미국의 65,221개 센서스 트랙을 대상으로 Hamilton-Perry 모델의 방법의 정확도를 평가하였다. 1990년과 2000년 두 시점의 인구변화를 활용하여 2010년의 인구를 예측하고, 2010년의 센서스 결과를 활용해 그 정확도를 사후 검증하였다. 오차율은 평균절대비오차(Mean Absolute Percent Error)와 평균 백분율 오차(Mean Algebraic Percentage Error)를 활용하였다.

Breidenbach·Kaeding·Schaffner(2019)은 코호트 요인법 개념을 활용하여 1km 격자단위로 독일의 인구분포를 2015년부터 2050년까지 예측하였다. 이 연구에서 1km 격자단위로 인구예측을 시도한 이유는 첫째, 다양한 공간단위로 집계 가능한 격자체계의 유연성, 둘째, 시계열 변화에 영향을 받지 않는 지속성, 셋째, 데이터 처리 및 저장이 용이한 효율성, 넷째, zip 코드나 행정구역에 비해 지역을 더 정밀하게 탐색할 수 있다는 정밀성 때문이다. 또한, 이 연구는 격자별 기존의 인구규모와 구성, 전국 수준의 사망률, 해당 지역의 출산율에 코호트요인법을 적용하였다. 이때 사망률은 모든 지역에 동일하게 적용하고, 출산율은 공간적 이질성이 크기 때문에 격자가 속한 지역에 따라 다르게 적용하였으며, 국내외 인구이동은 제외하였다(〈그림 2-1〉 참조).

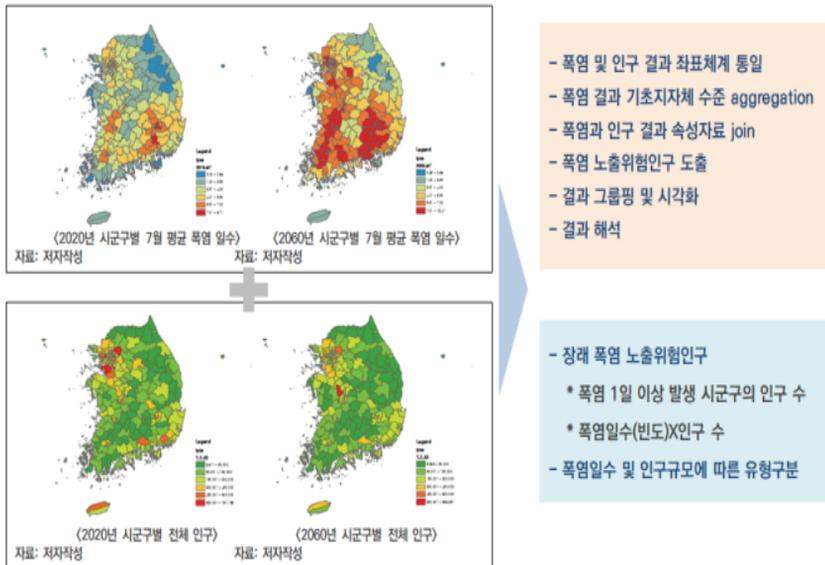
그림 2-1 | 출산율의 공간적 이질성과 장래인구 전망에 따른 격자단위 고령자 비율 변화



자료: Breidenbach·Kaeding·Schaffner(2019), 그림3·그림7, p739·p742

한국환경연구원의 심창섭·김오석(2019)은 통계청의 17개 시도별 인구추계가 지역의 미시적인 특성을 반영하지 못한다는 한계를 지적하였다. 이 연구는 폭염의 영향을 받는 인구를 추계하기 위해 전국을 37개 권역으로 나누어 2060년까지의 인구 변화를 시뮬레이션하고, 시·군·구단위로 배분하였다. 권역별로 코호트요인법을 적·용하여 인구를 예측하고, 시·군·구별 2018년도 성별/연령별 인구규모를 고려하여 권역별 인구추계 결과를 할당하였다. 이후 시군구별로 예측된 폭염 일수를 시군구별 장래인구추계 결과와 융합하여 장래 폭염 노출 위험 인구수를 산출하였다(<그림 2-2> 참조).

그림 2-2 | 시군구단위 장래인구추계와 폭염일수 간의 융합 활용



자료: 심창섭·김오석(2019), 그림3-15, p36

Kim·Kim(2020)은 우리나라가 인구감소·초고령화 시대를 맞이함에 따라 2020년과 2050년 사이에 초고령화 현상을 가장 먼저 또는 가장 늦게 경험할 것으로 예상되는 지역이 어디인지 37개의 권역에 대한 전망을 실시하였다. 권역별로 연령별, 성별 인구를 추계하기 위해 자기회귀누적이동평균 및 랜덤워크모형과 같은 외삽법을 활용한 인

구균형방정식을 제안하였다. 그리고 2000년부터 2018년까지의 출생률, 사망률 및 국내 순이동률과 같은 인구학적 변동요인을 17개 시도보다 작은 37개 권역으로 구성하여 2020년부터 2050년까지 권역별 인구규모, 성별, 연령별 인구를 추계하였다.

이보경(2019)은 코호트요인법을 활용하여 전국의 500m 격자단위 장래인구를 전망하였다. 분석 결과, 2040년까지 인구가 감소하는 격자는 159,917개로 현재 인구 거주 지역(197,856개 격자)의 80% 이상에서 인구가 감소할 것으로 예상하였으며, 2040년 전국 인구 5만 명 미만 시·군·구는 61개로 예상되었다. 이는 2019년 52개보다 9개 증가한 결과이다. 본 연구는 인구학적 요인을 고려하고 코호트요인법을 변형하여 격자단위로 장래인구를 전망한 연구로서 의의가 있으나, 미시공간 단위일수록 공간의 특성이 인구분포에 영향을 미치는 점을 반영하지 못하였다.

이처럼 코호트 생존율 기반의 인구전망법은 인구학자들에게 인정받은 유일한 전망법이지만, 소지역단위로 예측할 경우 데이터의 변동이 커지기 때문에 세분화된 지역에 적용하기 어려운 한계가 있다. 국내에서 전통적인 코호트요인법이 적용된 최소의 공간단위는 시·군·구이다. 인구변화의 큰 영향을 주는 인구이동 패턴을 읍·면·동단위로 계산하는데 큰 부담이 있기 때문이다. 또한, 대부분의 코호트요인법(Hamilton-Perry 모델과 같이 변형된 코호트 요인법 포함)은 소지역단위의 인구와 가구(세대 수)가 분포하는데 영향을 미치는 사회경제적, 인프라, 물리환경 등 공간의 특성을 적절히 고려하지 못한다는 한계가 있다. 이는 예측이 필요한 지리적 단위가 작아질수록 코호트요인법에서 다루는 인구학적 요인을 넘어서는 요인(예: 거주가 가능한 토지의 규모, 주거와 관련된 서비스 및 공공과 민간의 개발계획 등)이 인구분포에 영향을 미친다고 볼 수 있으며(Murdock et al., 1991), 지금까지 코호트요인법과 결합하여 활용할 수 있는 소지역단위의 공간 특성 데이터가 부재하였기 때문이다.

3) 공간통계모델

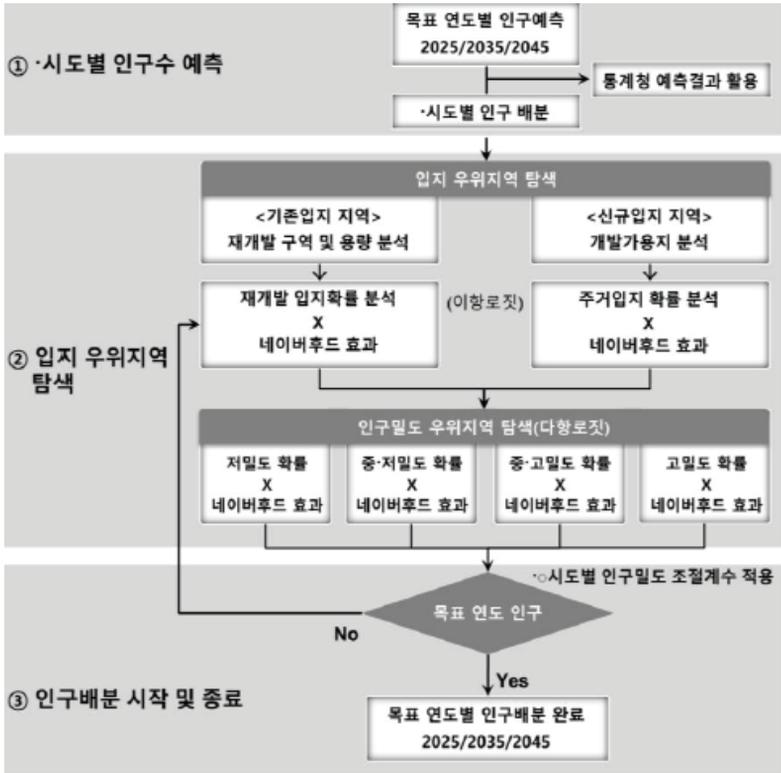
전통적인 장래인구추계 방법을 소지역단위에 적용하는데 가장 큰 장애물이었던 이유는 공간의 특성을 반영하지 못한다는 데 있다. 이를 극복하기 위해 공간통계모델을 활용한 소지역단위 장래인구예측 방법론이 연구되었다. 공간통계모델에서는 기초모델과 코호트요인법에서 간과하고 있었던 공간의 특성과 인접 공간 간의 상호관계가 인구 변화에 미치는 영향을 고려할 수 있다(Chi·Zhou·Voss, 2011). 일반적인 회귀분석 모델과 달리 공간통계모델에는 이웃효과(neighborhood effects) 및 공간자기상관성을 반영할 수 있기 때문이다(Anselin 1988, Lesage·Pace 2009). 공간통계모델은 인구예측 결과와 결과에 영향을 주는 요인들 간의 관계를 유추할 수 있어 의사결정자가 공간 정책 또는 계획을 수립하는데 정보를 제공할 수 있다는 장점이 있다.

Chi·Voss(2011)은 이웃들의 특성과 그들과의 거리를 고려한 시공간 회귀 인구예측 모형(spatio-temporal regression forecasting approach)을 제안하였다. 시공간 회귀 인구예측모형의 주요변수로 인구 성장 추세와 사회, 경제, 환경 특성을 사용하였다. 이 연구에서 제안한 모델이 공간의 특성과 공간 간의 상관관계를 고려하여 현상을 잘 반영한다고 볼 수 있으나, 미국 위스콘신의 MCD(Minor Civil Divisions)의 1990년부터 2000년까지의 인구를 예측한 결과, 단순한 선형회귀 모형에 비해 오차가 작게 나타나진 않았다. 물론, 지역의 특성에 따라 인구예측 시 시공간 회귀모형이 선형회귀 모형에 비해 더 뛰어난 예측력을 보이는 경우가 있었다. Chi·Zhou·Voss(2011)이 다양한 회귀모형을 밀워키의 센서스 트랙에 적용하여 예측한 결과, 도심 지역에 대해서는 시공간 회귀 모형이 추세외삽법에 비해 더 뛰어난 예측력을 보여주었다.

김현중·여관현(2019)은 셀 기반의 로지스틱 회귀모형(cell-based logistic regression model)을 활용하여 우리나라의 인구분포 변화를 미시적인 수준에서 예측할 수 있는 모형을 구축하고자 하였다(<그림 2-3> 참조). 이 모형을 활용하여 주거 입지 확률을 예측하였으며, 200m 격자단위로 2045년까지의 한국 장래인구를 예측하였다. 인구 및 경제특성, 근린 특성, 교통 특성, 자연 특성을 종속변수로 설정하고, 종속변수는 주거

입지/재개발구역 및 주거밀도로 설정하여 인구가 분포하는 입지 확률을 예측하였다.

그림 2-3 | 200m 격자단위 장래인구 예측 방법론 및 활용 변수



자료: 김현중, 국토연구원 개최 소지역단위 인구 전망 연구 동향 세미나(2022.04.15) 발표자료 활용

공간통계모델은 공간 특성과 공간 간의 상호관계를 반영할 수 있다는 장점이 있다. 그러나 선행연구에서 공통적으로 도심 지역을 제외하고 기존 연구에서 보여준 예측력이 단순한 회귀에 비해 뛰어나지 않다는 한계가 있다고 밝히고 있다. 따라서, 공간의 특성을 고려하되 예측력을 높일 수 있는 방법론 개발이 필요하다.

4) 마이크로 시뮬레이션

앞서 소개한 기초모델, 코호트요인법 기반 모델, 공간통계 모델이 지역의 인구 규모를 예측하는 방식이었다면, 마이크로 시뮬레이션은 개인의 행동을 예측하고 이를 집계하여 지역의 인구 규모를 전망하는 방식이라는 점에서 차이가 있다. 또한, 개인 단위의 예측으로 ‘소지역’이라는 공간 단위를 설정할 필요가 없다는 특징이 있다.

Ballas·Clarke·Wiemers(2005)는 SMILE(Simulation Model for the Irish Local Economy)을 통해 아일랜드의 소지역단위 인구를 단기 예측하였다. SMILE은 아일랜드 비도시지역의 정책 효과 변화와 경제 발전을 시뮬레이션하기 위해 개발된 다이나믹 공간 마이크 시뮬레이션 모델이며, 인구예측 부분이 핵심 역할을 한다. SMILE은 연령, 성별, 지역에 따른 사망확률, 연령, 결혼 여부 및 지역에 따른 출산확률을 고려하여 1991~1996년 그리고 1996~2002년의 소지역(sub-county 단위)별 인구변화를 시뮬레이션하였다. 인구이동의 경우 연령, 성별과 카운티단위로 랜덤하게 추출하여 이동확률을 적용하였으며, 국제 이동은 데이터의 한계에 따라 모델에서 고려하지 않았다.

호주는 인구와 가구 특성을 소지역(Statistical Local Areas)단위로 예측할 수 있는 SpatialMSM(Spatial Microsimulation Model)을 개발하였다(Harding·Vidyattama·Tanton, 2011). 소지역단위의 개인과 가구의 특성을 예측하는 것은 정책 수립에 중요한 정보이다. 예를 들어, 한정된 자원(초등학교, 복지시설 등)을 배분한다거나 또는 취약인구 현황을 전망하여 예방 정책을 수립하는 등으로 활용할 수 있다. SpatialMSM은 캔버라 대학의 호주 경제 및 사회정책 연구소인 NATSEM(National Centre for Social and Economic Modelling)이 개발한 공간 마이크로 시뮬레이션 모델로 소지역단위 가구와 개인의 특성을 파악하기 위한 도구이다. SpatialMSM/08C(Cassells et al, 2010)은 소득 및 가구 설문조사(Survey of Income and Housing)를 활용하여 호주 1,422개 SLA(Statistical Local Areas)단위의 개인 및 가구 특성을 재현데이터로 구성하고, 2027년까지의 전망치를 제공한다. 이는 주로 단기 예측에 활용되며, 장기 예측의 정확도에 한계가 있다(Wilson et al., 2022: 885).

5) 도시성장모델(urban growth model) 기반 시뮬레이션

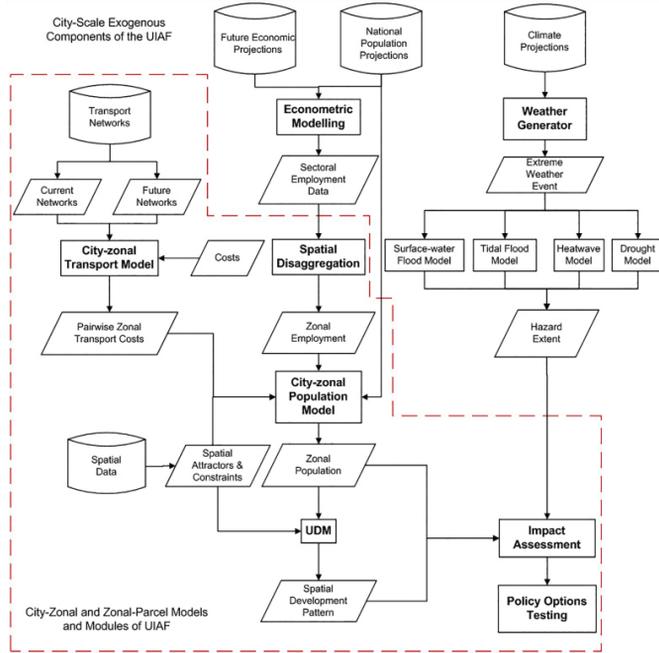
도시성장모델의 장점은 통계모델보다 인구변화와 관련 있는 이론적 배경을 잘 반영하고 사회경제, 물리적 환경 및 개발계획을 소지역단위로 고려할 수 있다는 것이다. 인구분포는 토지이용 및 교통인프라의 변화와 관련이 높다. 이와 관련하여 Triantakonstantis·Mountrakis(2012)는 토지이용과 교통을 고려한 모델을 검토하여 토지이용이 인구이동 등 지역의 인구 변화와 어떠한 관계가 있는지 살펴보았다. Triantakonstantis·Mountrakis(2012)가 도시성장모델 메타분석을 실시한 결과, 가장 많이 활용된 모델링 방법은 셀룰러 오토마타(Cellular Automata, 이후 CA)²⁾ 방법으로 나타났다. Ulan·Neumann이 1940년대에 CA모델을 처음 소개한 이후, John Conway가 1970년대에 'Game of Life' 모형을 제안하면서 CA모델은 1980년대부터 본격적으로 도시의 성장을 시뮬레이션하는데 많이 활용되고 있다.

Ford et al. (2019)는 미래 개발계획 시나리오에 근거하여 공간 간 상호관계(spatial interaction) 인구 모델과 CA를 활용한 토지개발 모델을 결합하여 미래 도시화 지도를 전망하였다. 또한, 기후변화에 따른 사회경제의 변화 과정을 시뮬레이션하기 위해 다양한 공간단위 통합모델인 UIAF(Urban Interated Assessment Framework)를 개발하였다. 이 연구에서는 국가 또는 지역 권역에서 예측한 인구 및 고용자 추계를 소지역의 사회경제 및 토지이용 변화 전망을 바탕으로 배분한 특징이 있다(<그림 2-4> 참조).

McKee et al. (2015)은 토지이용, 대도시와의 거리, 현재의 인구 등을 고려하여 30 arc-second(약 1km) 격자단위 미국 인구를 2050년까지 예측하였다. 전국 수준의 장래인구 규모에서 격자단위로 바로 할당한 것이 아니라 코호트요인법을 적용하여 카운티 단위로 인구를 예측하고, 각 카운티에서 할당된 격자로 인구를 배분하는 방식으로 진행하였다. <그림 2-5>와 같이 인구를 배분하는데 활용한 변수는 물리적 환경과 토지 피복, 경사도, 인구, 교통 인프라 및 대도시와의 거리등 토지이용을 통해 개발 압력을 고려하여 카운티 단위 장래인구 추계를 격자단위로 할당하였다.

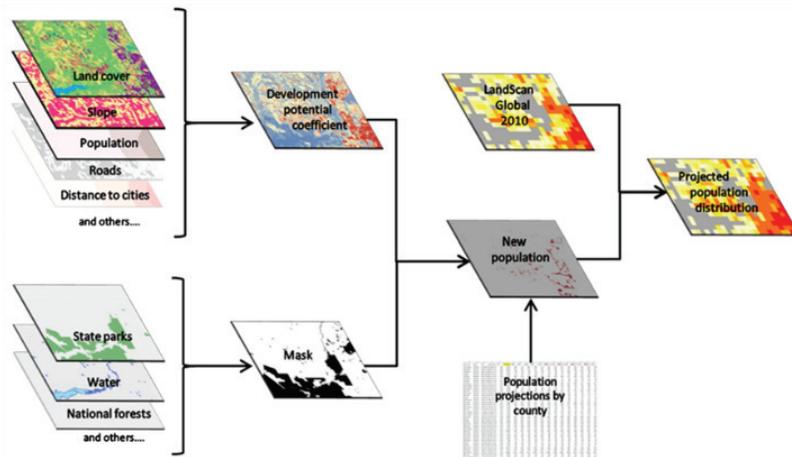
2) CA는 이차원의 격자별 상태와 인접한 셀과의 상호작용을 통해 전이규칙에 따라 변화를 시뮬레이션하는 방법론임

그림 2-4 | UIAF 모델링 구성



자료: Ford et al.(2019), 그림1, p230

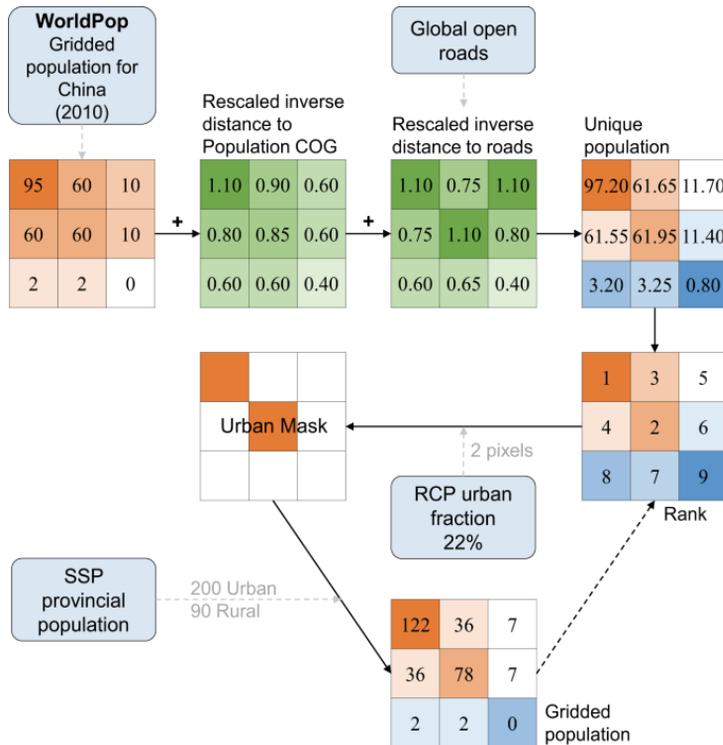
그림 2-5 | 소지역단위 인구추계를 위한 방법론 다이어그램



자료: McKee et al.(2015), 그림1, p1346

Chen et al. (2020a)은 SSPs(Shared Socioeconomic Pathways)에 따라 2010년부터 2100년까지 중국의 인구를 연령별, 성별, 교육 수준별로 추정하였다(그림 2-6) 참조). 이 연구는 31개 광역단위로 장래인구를 전망하고, 30 arc-seconds 격자로 배분하였다. 격자단위 인구밀도(현재), 인구중심도, 도로까지의 거리(역순)을 고려하여 인구분포의 가중치를 계산하였으며, 기준인구(unique population)를 도출하고 이들의 순위(rank)를 매겼다. RCP(Representative Concentration Pathway) 시나리오에 따른 도심지역(urban) 면적 배분에 따라 Unique population의 순위 순서대로 도심과 비도심을 격자별로 지정(urban mask)하였으며, 지역별 SSPs 시나리오별 예측된 장래인구 전망을 도심-비도심 구분과 기준인구의 비율에 따라 격자에 할당하였다.

그림 2-6 | 격자단위 인구분포 예측 시뮬레이션 방법



자료: Chen et al.(2020a), 그림2, p7

최내영(2010)은 유전적 알고리즘(Genetic Algorithm, 이후 GA) 방법론을 활용하여 도시환경변수를 활용해 격자단위 인구를 추정하는 연구를 수행하였다. 화성시 동부권 일대를 대상으로 읍·면·동의 행정구역 단위로 조사된 관내 인구데이터를 격자형 인구표면모형 자료로 변형한 후 인구자료에 기초한 GA모형을 적용하였다. 본 연구에 활용된 도시환경 변수는 표고, 경사도, 용도지역지구 지정현황 등이며, 이후 객관적 비교검증을 위해 동일한 변수로 공간회귀분석을 실시하여 결과값을 정량적으로 비교하였다. 여러 변수 중 경사도와 같은 자연환경과 용도지역지구가 인구분포에 가장 큰 영향을 미치는 것으로 나타났다.

국가교통DB(2018. 01)은 전국의 여객 통행을 추계의 바탕이 되는 읍·면·동 단위의 장래인구를 전망하기 위해 별도의 모형을 구축하였다. 장래 교통 수요 예측을 위해 읍, 면, 동(3, 489개) 5세 별(17개 구간) 인구추계모형을 개발하여 2020, 2025, 2030, 2035, 2040, 2045년 5세 단위 인구 예측값을 생성하였다. 이 연구의 특징은 택지개발과 같은 토지이용을 고려하였다는 데 있으며, 토지이용계획에 따라 미개발지역에서 시행되는 사업의 계획인구는 100% 반영하고, 기존 시가지에서 시행되는 사업은 준공 후 추가로 유입되는 인구만 고려하여 인구학적 요인으로 예측한 인구에 가산하였다. 토지이용계획 상 계획인구 규모에 따라 입주 시기를 조정하였는데, 예를 들어 10만 명 이상의 계획인구 규모를 가진 토지이용계획에는 준공연도에 전체의 30%, 준공+1년에 40%, 준공+2년 15%, 준공+3년 10%, 준공+4년에 5%가 유입된다고 가정하였다. 그런데 계획인구가 다른 지역 또는 같은 지역 내 인구이동의 제로섬계입임에도 불구하고 이를 어떻게 처리하였는지, 그리고 대규모 택지개발의 경우 전체 택지의 준공 시점과 택지 내 실제 인구 유입의 시점에 차이가 있음에도 불구하고 준공일자를 기준으로 유입인구의 흐름을 할당하였다는 것에 한계가 있다.

McKee et al. (2015)와 Chen et al. (2020a)의 연구와 같이 권역별로 인구를 예측하고 소지역단위의 토지이용 변화를 시뮬레이션하여 인구를 배분하는 것은 다양한 공간 단위에서 인구 규모의 정합성을 맞출 수 있고 격자별 예측값의 오차를 줄일 수 있다는 장점이 있다. 하지만, 도시성장 모델도 기존의 통계모델과 마찬가지로 소지역단위의

데이터가 필요하다는 점에서 범용적으로 활용되는데 한계가 있다. 또한, 대부분 모델에서 인구규모 전망 시 인구학적인 요인과 자연적 증감요인을 고려하지 않는 것은 인구학자(또는 관련 연구 흐름)에게 잘 받아들여지지 못하는 부분이다.

6) 인공지능 기법

머신러닝 등 인공지능 기법이 인구전망에 적용된 것은 앞서 소개한 다섯 가지 유형의 모델에 비해 비교적 최근이다. 또한, 머신러닝모델은 다른 분야의 예측 부문에 많이 활용되고 있음에도 불구하고 소지역단위 인구전망에 적용된 선행연구 최근 2~3년 전부터 진행되었으며(Riiman et al., 2019; Grossman et al., 2022, Grossman·Wilson·Temple, 2022), 관련 연구가 극히 드문 상황(Grossman et al., 2022: 2)이다. 지금까지 진행된 머신러닝 기법 적용 연구는 대부분 새로운 머신러닝 기법이 기존의 코호트 요인법 또는 단순 추세법과 같은 모델에 비해 유의미한지 비교하는 내용이다.

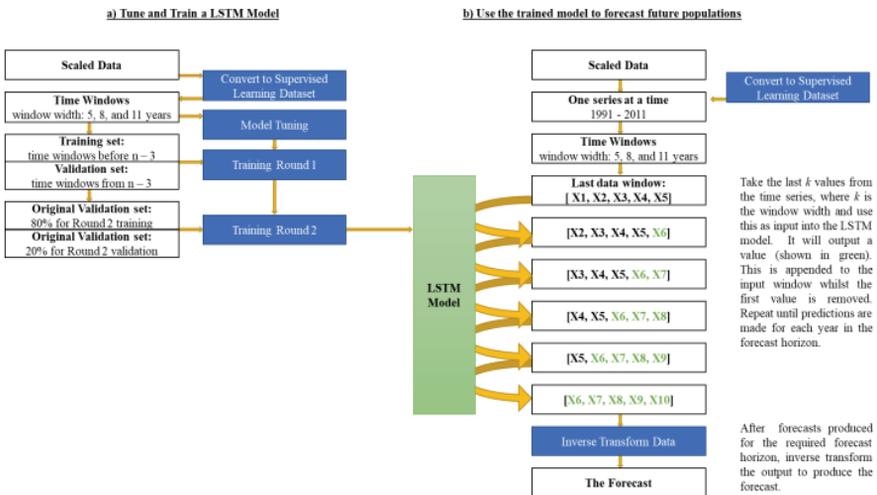
Riiman et al. (2019)은 ANN(artificial neural network) LSTM(Long Short-Term Memory) 모델을 활용하여 미국 알라바마 주 카운티의 10년 간 인구 변화를 예측하고, 코호트요인법으로 예측한 결과와 비교하였다. 이 연구는 연간 인구를 예측하는 것과 10년 단위로 조사되는 센서스 인구를 예측하는 모델을 각각 제안하였으며, 각 모델은 모든 알라바마 카운티의 데이터를 학습한 것과 각 카운티별로 구분하여 학습한 두 개의 유형으로 나누어 개발하였다. 가장 좋은 예측력을 보인 것은 각 카운티별로 ANN LSTM모델을 학습한 경우이며, 특히 연간 인구보다 10년 단위의 센서스 인구를 예측했을 때 오차율을 나타내는 MAPE(Mean Absolute Percent Error)가 5%로 나타나, 코호트요인법에 비해 더 유용한 모델임을 확인하였다.

Grossman et al. (2022)은 머신러닝 기법을 적용하여 인구를 예측할 때 단일 모델보다 여러 모델을 병합하여 활용하면 더욱 정확도가 높아질 것이라고 주장하였다. 또한, 앙상블 기법을 활용하여 6개 모델(CSP:, LIN/EXP, MEX, VG, THETA, LGBM)의

값을 평균낸 MEAN-ALL 모델 또는 6개의 모델 중 최대값과 최소값을 제외하고 평균 낸 TRIMMED-ALL 모델 등과 같은 통합모델의 정확도를 비교하였다.

Grossman·Wilson·Temple(2022)은 호주의 장래인구를 SA2(Statistical Area Level2) 단위로 예측하는데 LSTM모델을 활용하였다(〈그림 2-7〉 참조). 또한, 딥러닝모델의 최적 하이퍼파라미터를 지정할 수 있도록 지원하는 Keras Tuner를 사용하여 딥러닝 레이어 수와 시계열 예측 구간을 설정하였다. SA2단위의 실제 인구 벤치마킹 모델(선형/지수 모델) 결과와 LSTM 모델 예측 결과를 비교한 결과 단기 예측(5년 이내)일수록 LSTM 기반 모델의 예측도가 높은 것으로 나타났다. 그리고 LSTM은 인구 규모가 5,000명 이상인 지역을 예측하는데 정확도가 더 높았다. 즉, 인구가 희박하고 외딴 곳에 떨어진 지역의 경우 전통적인 기초모델(선형/지수 모델)의 정확도가 높은 반면, 주요 도심의 경우 LSTM의 예측력이 좋은 것으로 나타났다. 즉, Riiman et al. (2019)와 Grossman·Wilson·Temple(2022)을 통해 공통적으로 머신러닝 기반 예측모델의 정확도가 지역의 특성(인구 규모 또는 도심/비도심 등)에 따라 달라지는 것을 확인하였다.

그림 2-7 | 장래인구예측을 위한 LSTM 아키텍처

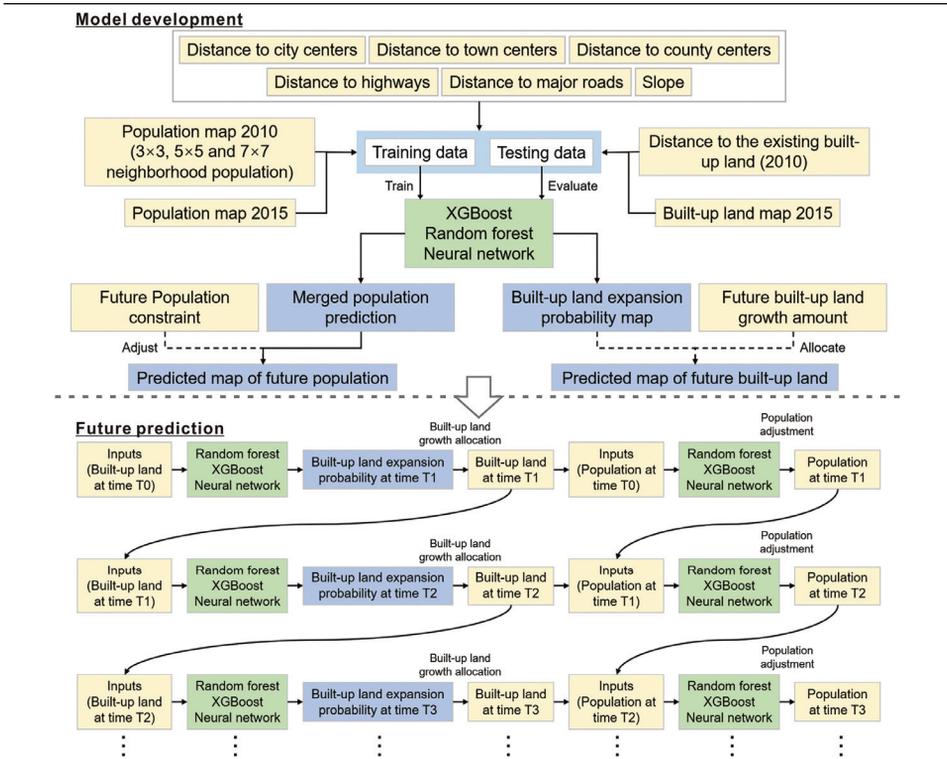


자료: Grossman·Wilson·Temple(2022), 그림1, p10

LSTM과 같이 과거의 인구규모 변화 추세를 학습하여 장래의 인구규모 변화를 예측하는 시계열 분석 방법론이 아니라, 인구분포와 밀접하게 관련된 외생변수를 머신러닝을 통해 예측하고 인구를 할당하는 연구도 최근 등장하였다. 예를 들어, Weber(2020)은 머신러닝 기법을 적용하여 2005년~2009년 사이의 데이터 학습을 통해 2011년~2015년까지의 독일 municipal 수준의 인구순이동을 예측하였다.

Chen et al. (2020b)는 이 연구에서 XGboost, Random forest, Neural network 등 인공지능 기법을 활용하여 장래의 건축물과 인구분포를 예측하였다(그림 2-8) 참조). 인구분포와 밀접하게 관련된 토지이용, 건물 분포 등을 인공지능 기법으로 예측할 수 있다면 자연스럽게 해당 지역의 인구 규모도 전망할 수 있다.

그림 2-8 | 인공지능 기법을 활용한 인구/건물 분포 예측 프로세스



자료: Chen et al.(2020b), 그림1, p5

2. 격자단위 인구통계 관련 사례 조사

1) 국내·외 격자단위 인구통계 현황

격자단위 인구통계는 전 세계에서 연구·활용되고 있다. 해외에서는 대부분 국가 또는 지역단위의 센서스 등록인구를 임의의 가정에 따라 격자단위로 배분하는 방식으로 인구분포 데이터를 구축하고 있다. 반면, 국내는 주민등록시스템과 센서스 기반 인구총조사 정보를 활용하여 주소 정보를 바탕으로 건물에 지오코딩하여 격자로 집계하는 방법으로 데이터를 구축하고 있다(<표 2-1> 참조).

우선 해외에서 가장 대표적으로 활용되고 있는 격자단위 인구통계는 GPW(Gridded Population of the World)이다. 컬럼비아대학 부설 국제지구과학정보네트워크 센터(CIESIN)에서 전 세계의 인구분포를 격자단위로 구축한 최초의 데이터이다. GPW는 국가단위의 센서스 인구를 Tobler의 피크노필랙틱(smooth pycnophylactic)³⁾ 보간법을 사용하여 격자별로 인구를 할당하는 방법으로 만들어졌다. 그리고 GRUMP(Global Rural-Urban Mapping)는 GPW 데이터를 기초로 활용한 격자단위 인구통계이다. 격자단위로 공간의 특성인 중심지까지의 거리 및 주거용 건물의 위치, 그리고 심야 시간대의 빛 분포와 같은 위성 정보 등을 종합하여 인구를 추정하는 방법으로 만들어졌다. LandScan은 ORNL(Oak Ridge National Laboratory) 30 arc-second(약 1km) 격자에 인구를 추정한 결과이다. US Census에서 집계된 행정구역단위 자료를 활용하여 도로망도, 경사도, 토지피복, 야간 조명 등을 가중치로 인구를 배분하였다. 미국뿐 아니라 유럽연합의 JRC(Joint Research Centre)는 전 세계의 건물 분포 레이어와 인구 밀도 분포 레이어를 250m와 1km 격자단위의 GHSL(Global Human Settlement Layer)로 공표하고 있다. 인구를 할당하는 매개 변수로 건물의 분포를 활용하였다. GHSL은 전 세계의 도시화 정도를 측정하고 분류하는 DEGURBA(Degree of Urbanisation)의 기초자료이다.

3) 피크노필랙틱 보간법(Pycnophylactic interpolation)은 “단위지역 내 총인구를 유지하면서 주변지역의 인구를 고려하여 인구분포를 보간하는 방법으로 대시메트릭 지도제작법에 의한 인구분포를 연속적이고 부드럽게 하여 좀더 현실적인 인구분포도를 작성할 수” 있음(Kim·Choi, 2011, p.67)

표 2-1 | 전 세계 격자단위 인구데이터 현황

	격자단위 인구	출처	방법론	공간단위	연도
해외	GPWv4.11 (Gridded Population of the World)	CIESIN (Center for International Earth Science Information Network), Columbia University	피크노필랙틱 보간법(Pycnophylactic interpolation)을 활용하여 센서스 등록인구를 격자단위로 할당	1km 격자	2000, 2005, 2010, 2015, 2020
	GRUMPv1 (Global Rural Urban Mapping Project)	CIESIN,; International Food Policy Research Institute, The World Bank, Centro Internacional de Agricultural Tropical	GPW와 동일하게 피크노필랙틱 보간법을 활용하지만, 격자가 도시(Urban)지역에 해당하는 여부에 따라 가중치를 조정	1km 격자	1990, 1995, 2000
	LandScan (LandScan Global Population Database)	ORNL (Oak Ridge National Laboratory)	격자별 도로망도, 경사도, 토지피복, 야간 조명 등을 바탕으로 가중치로 산출하여 행정구역 단위 센서스 인구를 배분	30 arc-seconds 격자 (약 1km)	매년 (2000~2017)
	GHS-POP (Global Human Settlement Layer -Population)	유럽위원회 JRC(Joint Research Centre)	건물의 밀도를 고려하여 데시메트릭 매핑을 통해 인구를 배분	250m, 1km	1975, 1990, 2000, 2015
국내	국토조사	국토지리정보원, 행정안전부	주민등록인구통계의 주소정보를 활용하여 거주지에 지오코딩하고, 격자단위로 집계	100m, 250m, 500m, 1km, 10km 등	2014년 이후 1년에 2회 갱신
	소지역 공간통계자료	통계청	센서스 기반 인구총조사 정보를 정보보호기법을 적용하여 격자단위 통계로 공개	100m, 500m, 1km, 10km 등	2022년 최초 공개 및 2000년도부터 제공

자료: Calka-Bielecka(2019: 3) 표1, Leyk et al.(2019: 13~14) 표2, 국토지리정보원(map.ngii.go.kr, 최종접속일: 2022.11.05.), 통계청 보도자료(2022.07.01.)을 종합 검토하여 저자 재작성

국내의 격자단위 인구통계는 주민등록인구 시스템 및 인구총조사의 주소정보를 활용하여 실제 인구 거주지를 지오코딩하여 격자단위로 집계하는 형태로 격자단위 인구통계를 생산한다는데 해외 관련 통계와 차별점이 있다. 우리나라에서 공신력있게 격자단위 인구통계를 생산·공개하는 기관은 국토지리정보원과 통계청이다. 국토지리정보원은 국토조사를 통해 2014년부터 매년 2회(6월, 10월) 행정안전부 주민등록시스템에 등록된 인구를 토대로 격자단위(100m, 250m, 500m, 1km 등) 인구통계를 공표하고 있다. 통계청은 통계지리정보서비스(SGIS)⁴⁾를 통해 인구총조사 결과를 토대로 2000년부터 5년 단위, 2015년 이후 1년 단위의 격자단위(1km, 10km 등) 인구통계를 공개하고 있다.

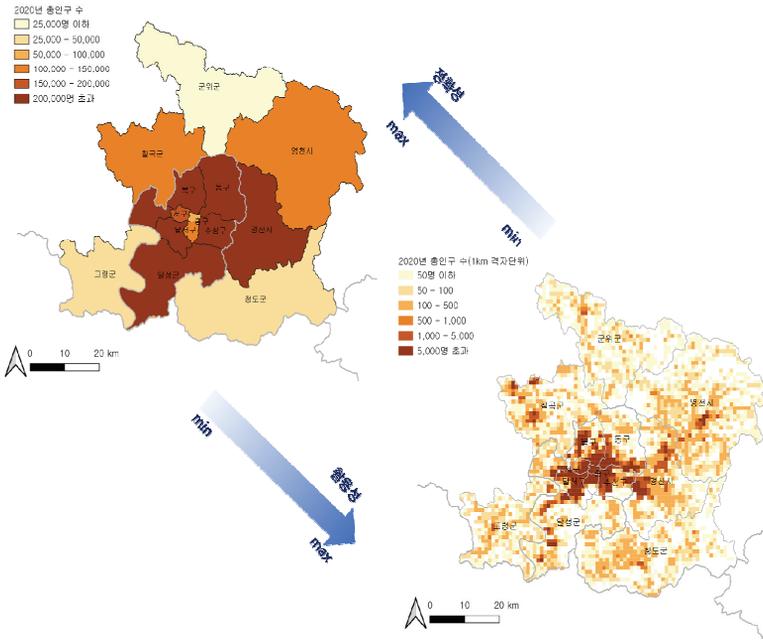
국토지리정보원과 통계청의 격자단위 인구통계의 차이점은 (1) 전자는 주민등록인구를 활용하는 것과 후자는 인구총조사 결과를 활용하는 것, (2) 전자는 개인정보 보호를 위해 5인 이하는 N/A로 제공하지만 후자는 실제 인구 값에 일정 구간의 허수를 조합하여 값을 제공한다는 것, (3) 전자는 100m, 250m, 500m, 1km, 10km 등 격자단위 통계를 제공하지만 후자는 100m, 1km, 10km 등인 점이다. 국토지리정보원의 국토조사가 다양한 격자단위 통계를 제공하고 개인정보보호를 위해 처리한 방식이 비교적 단순하여 값의 왜곡이 적다는 점 등에서 장점이 있다. 그리고 통계청 SGIS의 경우 2000년도부터 격자통계를 제공하여 시계열 데이터 확보가 용이하고 우리나라 장래 인구추계가 인구총조사 정보를 활용한다는 점에서 장점이 있다.

인구부문의 격자단위 공간통계의 경우 전 세계적으로 1km 격자가 범용적으로 활용되고 있다(<표 2-1> 참조). <그림 2-9>의 두 지도는 대구광역시와 인접 경북 시·군·구의 인구분포와 1km 격자단위의 인구분포를 비교한 것이다. 100m 격자와 같이 소지역단위의 장래인구를 예측할 경우 공간의 정밀성을 확보하여 결과의 활용성은 증가한다. 그러나 예측해야하는 대상 격자 수가 증가하면 예측 결과의 총 오차도 기하급수적으로 증가하여 예측 정확도가 감소하는 한계가 있다. 예를 들어, 좌측 상단의 대구광역시 및 인접 지역의 장래인구를 시·군·구단위로 예측할 때 예측 대상이 총 14개지만 우

4) 통계지리정보서비스(SGIS)(<https://sgis.kostat.go.kr/>, 최종접속일: 2022.11.22.)

측 하단의 1km 격자단위의 경우 예측 대상은 3,046개이며, 100m 격자단위로 예측할 경우 예측 대상은 304,600개가 된다. 이 연구에서는 공간의 정밀성을 확보하면서도 오차를 최소화하여 모델의 효용성을 극대화할 수 있도록, 읍·면·동, 시·군·구 등 행정 구역보다 작으면서 국토공간계획에 활용할 수 있는 공간단위인 소지역을 1km 격자로 정의하였다. 그리고 장래인구 예측에 시계열 데이터 확보가 중요하고 인구총조사의 기준인구가 통계청 공표 장래인구추계의 기초자료로 활용되고 있으므로⁵⁾ 이 연구에서는 통계청 SGIS에서 제공하는 격자단위 인구통계를 활용하였다.

그림 2-9 | 공간의 정밀성에 따른 정확성과 활용성 관계(대구광역시 및 인접 지역의 총인구 예시)



자료: 저자 작성

5) 우리나라에는 인구규모와 장래인구 전망에 대해 명문화된 법률이 없으나, 「통계법」제18조10의 규정에 따라 통계청장의 승인 아래 작성하는 승인통계를 공식 통계로 받아들이고 있다. 통계청은 5년단위로 장래 50년 후 인구규모를 연단위 성별/연령별/시나리오별로 작성하여 공표하고 있는데 이는 코호트요인법에 근거하고 있다. 우리나라 통계청에서는 코호트요인법의 인구균형방정식의 변수 설정을 위해 기준연도 7월 1일 시점의 3개월 이상 거주하는 내국인과 외국인을 대상으로 인구총조사 정보를 활용하여 기준인구로 정의하고, 인구변동요인인 출생, 사망, 국제이동은 해당연도 7월에서 다음연도 6월까지의 발생 수를 참조한다.

2) 격자단위 공간 빅데이터 관련 선행연구 현황

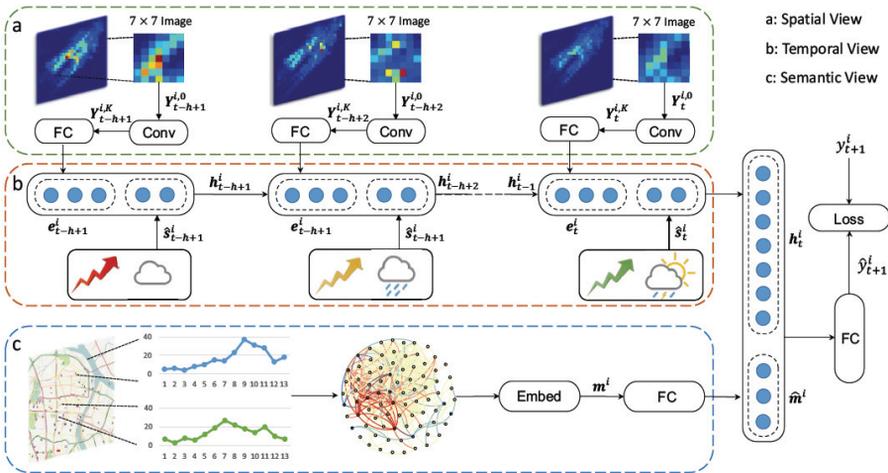
소지역의 공간단위는 국가, 활용 목적 및 적용 방법론에 따라 행정구역, 통계단위 및 격자 등으로 다양하게 정의될 수 있다. 앞서 격자를 포함한 소지역단위 장래인구 예측 방법론을 검토하였는데, 이 절에서는 딥러닝 등 인공지능 기법의 관점에서 격자 단위 공간 빅데이터가 어떻게 활용되고 있으며 관련 방법론의 연구 동향을 살펴보았다.

격자단위 공간 빅데이터에 인공지능 기법을 적용한 선행 연구는 주로 사람과 재화의 즉각적인 이동이 확인되는 교통량, 유동인구 등 중심으로 이루어져 왔다. 격자단위 공간 빅데이터에 딥러닝을 적용한 선행 연구 관련 연구 분야는 (1) 이종 모형의 결합(Yao et al., 2018, Yao et al., 2019, Lin et al., 2020), (2) 딥러닝 모형의 고도화를 위한 시공간 데이터의 변환(Yao et al., 2018, Yao et al., 2019, Lin et al., 2020, Song·Han·Stettler, 2020, Du et al., 2020), (3) 입력 데이터 및 외생변수의 세분화(Liang et al., 2019), (4) 모형 아키텍처의 가중치 변화(Lin et al., 2020, Grossman et al., 2022, Grossman·Wilson·Temple, 2022) 등으로 분류할 수 있다. 즉, 지금까지 소개된 격자단위 공간 빅데이터에 적용된 딥러닝 기법은 기계가 인지하기 쉽고 공간의 특성이 잘 훈련될 수 있도록 입력 데이터와 아키텍처 및 모형 구성을 설계 하는데 주안점을 두었다고 볼 수 있다.

격자단위 공간 빅데이터에 딥러닝을 적용한 대표적인 연구로 Yao et al. (2018)이 있다. 미국 Uber, 중국 Didi Chuxing과 같은 승차공유서비스수요를 소지역단위(0.7km 격자단위)로 예측하기 위해 CNN의 아키텍처와 LSTM의 아키텍처를 결합하여 딥러닝 모델을 개발하였다. 격자단위 공간데이터의 특성과 거리에 따른 이웃효과를 고려하기 위해 분석 대상 격자와 이웃의 격자(7x7, 자신 포함 총 49개 격자)의 값을 함께 고려할 수 있는 CNN과 시계열의 자기상관성을 고려할 수 있는 LSTM의 구조를 결합하였다. 시계열(0.5 시간 단위)별로 구축된 CNN 구조에서 1차원 변환 값을 LSTM 입력층으로 활용하였다. 그리고 최근 4시간 값을 과거의 추세 입력값으로 함께 활용하였고, 평균 수요, 위·경도 값, 날씨 정보, 휴일 여부 등을 함께 활용하였다.

Yao 외(2018)은 이 연구에서 구조적으로는 거리 가중치를 고려하여 지점 간의 (유클리디안 직선거리 계산법) 거리가 가까울수록 효용가치가 상승하는 지수함수 형태를 산출하여 종합모형에 추가하였다. <그림 2-10>의 a 파트는 공간의 관점에서 격자단위 공간데이터의 이웃효과를 고려하여 합성곱 모델을 구성하는 방법이다. 그리고 b는 이 합성곱을 시계열 관점에서 LSTM 모델에 결합하는 방법이다. 마지막으로 c는 격자단위 데이터는 아니지만 승차공유서비스에 영향을 미칠 수 있는 해당 지역과 주변 지역의 영향요인을 거리 함수를 고려하여 예측 모델에 적용한 것이다.

그림 2-10 | DMVST-Net의 아키텍처 구성



자료: Yao 외. (2018). 그림1, p.2591

3) 일본의 격자단위 장래인구 예측 개발 및 활용 사례

일본 국토교통성의 국토정책국은 2015년도 국세조사 정보를 코호트요인법에 적용하여 2050년까지 5년 단위의 500m 및 1km 격자단위 장래인구 예측 방법과 결과를 공표하였다. 이 결과는 2014년 일본 국토의 그랜드 디자인 2050과 제2차 국토형성계획의 기초자료로 활용되었다. 격자단위 현재 및 장래의 인구분포와 고령화 수준은 공간에서 직관적으로 지역의 현황을 파악하고, 지역별로 주민생활을 유지·향상하고 소득과 고용을 증진하여 지역을 지속가능하게 하기 위한 전략을 마련하는 데 큰 도움을 준다⁶⁾. 일본 국토교통성의 격자단위 장래인구를 인구감소·저출산·고령화, 경제성장력 약화, 재정력 저하 등의 장래 변화에 적응하기 위한 자원배분에 활용되고 있다.

일본 정부는 격자단위 장래인구 예측 방법을 개발하고 이 결과의 활용 방법에 대해 매년 지속적으로 연구를 이어가고 있다. 2016년에는 생활 관련 서비스 이용 가능성 시뮬레이션 분석 연구, 2017년에는 인프라 정비 방안 및 대도시권 고령화 분석, 2018년에는 앞서 제시한 추계 방법과 활용방안을 정교화하는 작업 등을 순차적으로 시행하였다⁷⁾. 그리고 격자단위 장래인구 정보는 빅데이터 기반 지역경제분석시스템(RESAS, Regional Economy Society Analyzing System)를 통해 웹으로 모두에게 공개되고 있다⁸⁾.

일본 국토기술정책종합연구소는 2017년부터 기존에 시구정촌 단위 장래인구예측 결과의 활용성을 높이기 위해 조밀한 마을단위로 5세별·성별 인구를 예측할 수 있는 ‘장래인구·세대예측 도구’를 개발하였다. 2018년에는 마을단위 장래인구와 세대를 격자단위로 배분하고, 이를 그래프와 지도로 표현할 수 있는 기능을 구성하여 공개하였다. 이 도구는 누구나 쉽게 활용할 수 있도록 엑셀 기반의 도구로 개발되었으며, 전체 프로세스와 조작 방법을 <그림 2-11>과 같이 제공하여 실무에서의 활용도를 높이고 있다.

6) 일본국토교통성(国土交通省), <https://www.mlit.go.jp/common/001135250.pdf>, p1~p2를 저자가 번역하였음. 최종 접속: 2022년 11월 21일

7) 일본국토교통성(国土交通省), https://www.mlit.go.jp/kokudoseisaku/kokudoseisaku_tk3_000086.html, 웹사이트 내용을 저자가 번역하였음. 최종 접속: 2022년 11월 21일

8) RESAS(<https://resas.go.jp/>)는 지역의 현황을 파악하고, 미래를 예측하여 자발적이고 효율적인 정책 및 계획을 수립할 수 있도록 지원하는 웹 기반 시각화 및 분석 도구임(이영주·임은선·차미숙, 2015: 53).

3. 시사점

소지역단위 장래인구 예측은 국내에서는 생소하고 제한적으로 연구되어온 분야이다. 따라서, 지금까지 국내외에서 진행된 관련 연구의 동향을 면밀히 파악하여 우리나라에서 구득할 수 있는 관련 데이터의 종류, 공간 단위 및 시계열 범위 등을 고려하여 방법론의 개발방향을 설정할 필요가 있다. 소지역단위 인구예측을 방법론에 따라 6개 분야로(기초모델, 코호트 생존율 활용 모델, 공간통계 모델, 마이크로 시뮬레이션, 도시성장모델, 인공지능기법) 구분하여 검토하였으며, 각 방법론의 활용 모델, 내용, 한계 및 선행연구에서 주로 다룬 소지역의 공간단위를 요약한 내용은 <표 2-2>와 같다.

표 2-2 | 소지역단위 장래인구 예측 방법 유형 및 특징

분류	특징		선행연구
기초모델	모델	<ul style="list-style-type: none"> 경향외삽법: Linear/Exponential model, Holt-Winters, Autoregressive Integrated Moving Average 등 비율법: share of population, share of population growth model 등 	<ul style="list-style-type: none"> Rayer(2008) Baker et al. (2008) Hauer et al. (2015) 이창호·이승일 (2006) 최내영(2010) 이석준 외(2014) 임은선 외(2017) 최현정·최석환·홍성조(2019)
	내용	<ul style="list-style-type: none"> 최소한의 데이터를 활용하여 직관적으로 소지역인구 현황을 예측 주로 연령별 인구에 비해 총인구 규모를 예측하는데 활용 비율법의 경우 매개변수(주택 등)를 활용하여 지역 인구를 소지역으로 배분 적용이 간편하고 이해가 쉽다는 장점 	
	한계	<ul style="list-style-type: none"> 공간단위가 조밀해질수록 인구규모에 영향을 주는 요인이 다양해지고 개개인의 의사결정에 큰 영향을 받기 때문에 과거의 추세에서 벗어날 확률이 높음 인구의 성별 연령별 구성과 도시/토지 이용을 고려하지 않기 때문에 장래의 인구규모가 왜 그렇게 변하는지에 대한 설명에 한계 	
코호트 생존율 활용 모델	모델	<ul style="list-style-type: none"> 코호트요인법 Hamilton-Perry model 	<ul style="list-style-type: none"> Swanson·Schlot tmann·Schmidt (2010) Baker·Swanson·Tayman(2021) Breidenbach·Ka eding·Schaffner (2019) 심창섭·김오석 (2019) 이보경(2019)
	내용	<ul style="list-style-type: none"> 코호트요인법은 출생, 사망, 이동이라는 인구변동요인의 복잡성과 연계성을 고려하여 인구균형방정식을 만들어 장래인구를 추계 Hamilton-Perry model은 코호트별 인구변화 비율이 일정하게 지속된다는 가정에 따라 두 시점의 연령별 인구규모의 변화율을 활용 	
	한계	<ul style="list-style-type: none"> 급격한 추세 변화에 지나치게 높거나 낮은 인구규모를 예측할 수 있다는 단점이 있어 광역권의 총 인구 규모와 같은 컨트롤 변수필요 소지역단위의 인구와 가구가 분포하는데 영향을 미치는 사회경제적, 인프라, 물리환경 등 공간의 특성을 고려하지 못함 	

분류	특징		선행연구
공간통계 모델	모델	<ul style="list-style-type: none"> • Spatial-temporal regression • Geographically-weighted regression 	<ul style="list-style-type: none"> • Chi·Voss(2011) • 김현중·여관현 (2019)
	내용	<ul style="list-style-type: none"> • 소지역 내 다양한 특징과 이웃효과와 같은 공간데이터의 특성을 반영한 통계모델 • 인구예측 결과와 여기에 영향을 주는 요인들 간의 관계를 유추할 수 있어 의사결정자가 공간 정책 또는 계획을 수립하는데 필요한 정보를 제공 	
	한계	<ul style="list-style-type: none"> • 도심지역을 제외하고 기존 연구에서 보여준 예측력이 단순한 회귀에 비해 뛰어나지 않다는 한계 	
마이크로 시뮬레이션	모델	<ul style="list-style-type: none"> • 아일랜드 SMILE(Simulation Model for the Irish Local Economy) • 호주 Spatial MSM(Spatial Microsimulation Model) 	<ul style="list-style-type: none"> • Ballas·Clarke·Wiemers (2005) • Harding·Vidyattama·Tanton (2011) • Cassells et al. (2010)
	내용	<ul style="list-style-type: none"> • 개인의 행동을 예측하고 이를 집계하여 지역의 인구규모를 전망하는 방법으로 '소지역'이라는 공간단위를 설정할 필요가 없다는 특징 	
	한계	<ul style="list-style-type: none"> • 주로 단기 예측에 활용되며, 장기 예측의 정확도의 한계가 있음 	
도시성장 모델	모델	<ul style="list-style-type: none"> • Cellular Automata(CA) • 유전적 알고리즘(Genetic Algorithm, GA) 	<ul style="list-style-type: none"> • Triantakostas·Mountrakis(2012) • Ford et al. (2019) • McKee et al. (2015) • Chen et al. (2020a) • 최내영(2010) • 국가교통DB (2018.01)
	내용	<ul style="list-style-type: none"> • 토지이용과 교통을 고려한 소지역단위 인구예측 모델 • 최근 사회 경제 및 기후변화에 따른 인구 분포의 장기적 변화를 전망하는데 널리 활용 	
	한계	<ul style="list-style-type: none"> • 소지역단위의 데이터가 필요하다는 점에서 범용적으로 활용되는데 한계 • 대부분의 모델에서 인구규모를 전망하는데 인구학적인 요인과 자연적 증감요인을 고려하지 않는데서 인구학자(또는 관련 연구 흐름)에게 잘 받아들여지지 않음 	
인공지능 기법	모델	<ul style="list-style-type: none"> • 딥러닝: ANN(Artificial Neural Network), LSTM(Long Short-Term Memory) 등 • 머신러닝: XGboost, Random Forest 등 	<ul style="list-style-type: none"> • Riiman et al. (2019) • Grossman et al. (2022) • Grossman·Wilson·Temple(2022) • Weber(2020) • Lloyd·Soricchetta·Tatem(2017) • Chen et al. (2020b)
	내용	<ul style="list-style-type: none"> • 머신러닝 등 인공지능 기법이 인구 전망에 적용된 것은 최근 2~3년 전부터 진행되기 시작함 • 인구분포와 밀접하게 관련된 요인변수를 머신러닝을 통해 예측하고 인구를 할당하는 연구도 등장 	
	한계	<ul style="list-style-type: none"> • 머신러닝 기법을 적용한 연구 대부분 5년 또는 10년과 같은 단기 예측에 활용되는데, 이는 시계열 데이터 확보에 한계로 충분히 모델을 학습시킬 수 없기 때문임 	
소지역 공간단위	<ul style="list-style-type: none"> • 미국 카운티, 센서스 트랙, MCD(Minor Civil Division) • 호주 SLA, SA2 • 한국 시군구, 읍면동, 집계구 • 격자(100m, 200m, 500m, 1km) 		

자료: 저자 작성

선행연구 검토를 통해 본 연구의 모델개발에 시사점을 준 기존 소지역단위 장래인구 예측 방법의 한계를 종합하면 다음과 같다. 첫째, 지역이 작아질수록 예상하지 못한 사건으로 인구규모 및 구성의 추세가 쉽게 바뀔 수 있다. 광역단위 인구를 예측할 때는 예측 대상 인구규모가 크기 때문에 지금까지의 거시적인 추세가 적용될 가능성이 크지만, 공간단위가 조밀해질수록 인구 규모에 영향을 주는 요인이 다양해지고 개개인의 의사결정에 큰 영향을 받기 때문에 과거의 추세에서 벗어날 확률이 높다. 따라서 기존의 기초모델 또는 코호트요인법과 같은 방법론을 격자단위 장래인구예측에 단독으로 활용하는데 한계가 있다.

둘째, 코호트요인법(Hamilton-Perry 모델과 같이 변형된 코호트요인법 포함)은 소 지역단위의 인구와 가구(세대 수)가 분포하는 데 영향을 미치는 사회경제적, 인프라, 물리환경 등 공간의 특성을 적절하게 고려하지 못한다. 이와 같은 한계를 극복하고자 도시성장모델 기반 시뮬레이션 모형은 토지이용, 교통인프라 등을 고려한 CA(Cellular Automata) 등의 방법론을 적용하여 장래인구의 분포를 예측하였으나, 기후변화 관련 연구 흐름 또는 도시계획의 관점에서 시나리오에 기반한 모델로서 예측 정확도에 대해서는 큰 관심이 없는 것으로 나타났다.

즉, 전통적인 인구추계모형에서 다루는 출산율, 사망률, 인구이동을 다루면서 토지 이용 등 공간의 특성을 소지역단위로 반영하여 인구의 분포를 예측할 수 있는 복합적인 모델이 필요하다. 이를 위해서는 공간과 시간상으로 매우 복잡하고 대용량의 데이터를 처리할 필요가 있는데, 이를 위해서는 머신러닝, 딥러닝 등 인공지능 기법을 적용하는 것이 적절하다. 그러나 인공지능 기법이 장래인구예측 분야에는 다른 학문 분야에 비해 매우 보수적으로 인구예측에 적용되어 온 것이 사실이다. 최근 2~3년 사이에 진행된 관련 연구는 간단한 딥러닝 모델이 기존의 선형/지수 모델에 비해 얼마나 좋은가를 비교하는 데 그쳤다. 그리고 예측 자체를 위해 모델을 학습하는 연구는 거의 없는 실정 이므로 이와 관련된 연구 수행은 정책적·학술적으로 모두 의의가 있다.

인공지능 기법을 적용한 소지역단위 장래인구 예측 방법론을 개발하는데 사례조사 결과에서 참조할 수 있는 시사점은 다음과 같다. 첫째, 인공지능 기법 기반의 인구예측

모델은 5년 또는 10년과 같이 비교적 단기간의 인구예측 모델을 개발하였으나, 우리나라에서 공간계획에 활용하기 위해서는 적어도 20년 이상의 중장기적 시점의 인구를 예측할 수 있어야 한다. 선행연구가 단기예측에 집중된 가장 큰 이유는 시계열 데이터 확보에 한계로 충분히 모델을 학습시킬 수 없기 때문이다. 우리 연구에서도 시나리오 기반이 아니라 데이터 기반으로 2050년까지 약 30년의 인구를 예측하는데 LSTM과 같이 소지역단위 인구의 과거 시계열 변화를 모두 가지고 있어야 하는 모델의 경우 활용에 제약이 있다. 그러나 Grossman·Wilson·Temple (2022)이 주장한 것과 같이 인구분포에 영향을 미치는 요인을 예측하는 모델과 이로 인해 달라지는 인구 분포를 예측하는 모델을 병합하여 활용한다면 이와 같은 한계를 극복하고 다양한 인구 분포 영향요인을 고려한 예측력 높은 방법론 개발을 기대할 수 있다. 또한 아키텍처 측면에서 예측 기간이 길어질수록 인공지능 모델 적용 결과에 변동성이 커질 수 있어 이를 단독으로 활용하는 것보다 장기 장래인구 추계에 안정적인 코호트요인법을 결합할 필요가 있다. 둘째, 소지역, 특히 격자단위 시계열 데이터를 학습할 경우 시간과 공간의 이웃효과를 고려할 수 있어야 한다. Tobler(1979)는 셀룰러 지오그래피(Cellular Geography)를 개념을 바탕으로 격자단위 데이터의 시간적·공간적 위계 관계를 모델링에 반영해야함을 강조하였다. 특히 지역이 조밀해질수록 공간적 이웃효과가 강하게 나타날 수 있는데, 이를 고려한 모델을 개발해야 한다. 지금까지 널리 활용된 코호트요인법의 장점과 인공지능 기법을 공간 빅데이터에 적용한 모델의 장점을 반영할 수 있는 복합적인 방법론을 개발한다면, 이론적으로 의사결정자들을 모두 이해시키면서 예측력 높은 모형을 개발할 수 있을 것으로 기대된다.

국내·외의 격자단위 인구통계 관련 사례조사 결과 국내외에서 공통적으로 1km 격자단위의 인구통계를 활용하고 있음을 알 수 있다. 공간의 정밀성에 따른 활용성, 오차 및 사례 조사 결과에 따른 본 연구에서도 소지역을 1km 격자단위로 설정하여 연구를 진행하였다. 그리고 국내에서 사용할 수 있는 격자단위 인구통계는 국토지리정보원과 통계청에서 각각 주민등록인구와 인구총조사를 바탕으로 생산하고 있다. 장래인구 예측에 시계열 데이터 확보가 중요하고 인구총조사의 기준인구가 통계청 공표 장래인구추

계의 기초자료로 활용됨을 고려하여, 이 연구에서도 통계청 SGIS에서 제공하는 격자 단위 인구통계를 활용하였다.

격자단위의 장래인구 예측 결과를 정책 및 계획에서 활용하기 위한 일본 사례에서 얻을 수 있는 시사점을 다음과 같이 정리할 수 있다. 첫째, 소지역단위 장래인구 예측 모델과 함께 활용할 수 있는 사례 연구와 시뮬레이션 도구가 함께 수반되어야 한다. 국토교통성에서는 단순히 격자단위 장래인구 예측 결과를 제공하는 것뿐 아니라, 자원 배분 및 인프라 계획 등에 어떻게 적용할 수 있는지 후속 연구를 지속적으로 수행하였다. 이 과정에서 예측 모형의 정교화 관련 연구도 함께 진행되었다. 둘째, 실무자가 활용하기 쉽도록 엑셀 기반의 도구를 개발하고 수요에 맞는 매뉴얼을 개발할 필요가 있다. GIS분석 등 데이터 분석에 익숙하지 않은 실무자가 쉽게 지역의 현황을 파악하고, 이를 보고서 또는 계획에 삽입하기 위해서는 엑셀을 통해 계산은 물론 그래프와 지도까지 한번에 시각화할 수 있도록 도구를 개발할 필요가 있다. 셋째, 예측 결과를 온라인으로 공개하여 다양한 이해관계자가 정보를 얻고 의견을 공유할 수 있도록 해야 한다. 장래인구 예측 결과와 이를 바탕으로 시뮬레이션할 수 있는 도구를 국토연구원 인터랙티브 리포트⁹⁾와 같은 웹서비스로 공개할 수 있다.

9) 국토연구원 인터랙티브 리포트(interactive.krihs.re.kr/interactive, 최종접속일: 2022.11.20.)



CHAPTER 3

소지역단위 장래인구 예측 방법론 개발

1. 개발 방향 51
2. 격자단위 장래인구 예측 AI 모델(Gridded Population Forecast with AI, GPAI) 아키텍처 설계 56
3. GPAI 학습과 장래인구 예측을 위한 공간 데이터 69

03 소지역단위 장래인구 예측 방법론 개발

본 장에서는 선행연구 및 사례 조사의 시사점을 바탕으로 소지역단위 장래인구 예측 방법론 개발 방향을 정립하였다. 개발 방향에 따라 ‘격자단위 장래인구 예측 AI 모델(Gridded Population Forecast with AI, GPAI)’ 아키텍처와 각 모델의 알고리즘을 개발하였다. 그리고 GPAI 학습을 위해 다양한 출처의 인구학적 특성과 공간적 특성 데이터를 검토하고, 학습용 데이터 구축을 위해 수행한 공간 데이터 수집·전처리·가공 과정을 설명하였다.

1. 개발 방향

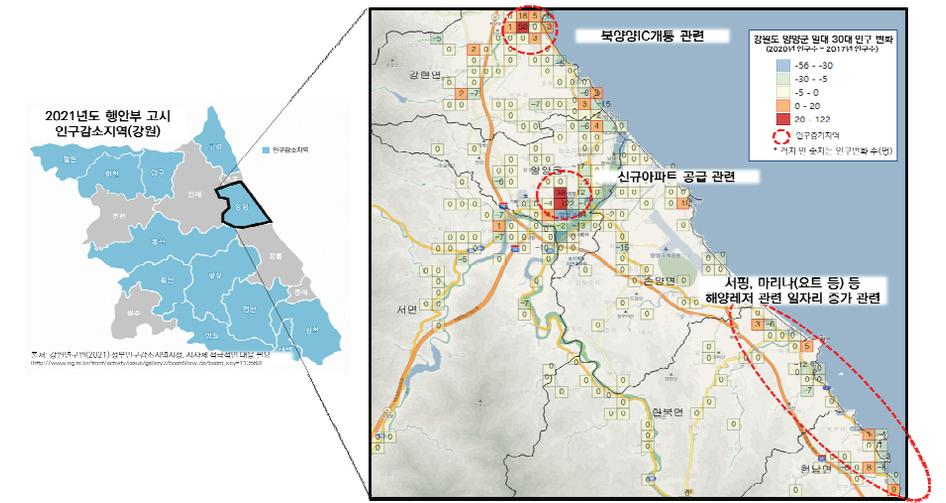
가속화되는 인구감소 추세에 따라 우리나라 중앙정부는 소멸의 위험이 큰 지자체를 선정하여 예산을 지원하고 있다. 예를 들어, 국토교통부·행정안전부는 균형발전특별법에 따라 연평균 인구변화율, 소득수준, 재정 상황 및 생활인프라 접근성 등을 평가하여 70개 성장촉진지역을 지정('19.09.10. 고시)¹⁰⁾하여 지역개발 사업을 지원하고 있다. 그리고 행정안전부는 평균인구증감률, 인구밀도, 고령화비율, 조출생률, 재정자립도 등 인구와 재정을 평가하여 89곳의 인구감소지역을 지정('21.10.19. 고시)하여 10년간 기금을 지원하고 있다. 예시로 든 대표적인 인구감소지원 정책을 포함하여 대부분의 지역 지원 정책은 시·군·구와 같은 행정구역 단위로 기획 및 집행되고 있다. 따라서 지자체에서는 우선적으로 대응이 필요한 공간이 어디인지 파악하거나, 국가에서는 지자체에 투입된 지원이 실제 인구 소멸의 위험을 겪고 있는 공간에 영향을 미쳤는지, 어떤 역할을 했는지 등에 대해 평가하고 다음 계획에 환류하는 데 한계가 있다.

10) 성장촉진지역은 균특법 제2조제6호에 따라 “생활환경이 열악하고 개발수준이 저조하여 사회적·경제적 성장을 촉진하기 위해 필요한 기반시설 구축 등에 특별한 배려가 필요한 지역”을 의미하지만, 해당 지자체 선정에 인구 변화율이 큰 영향을 미치며, 70개 성장촉진지역이 행정안전부 지정 인구감소지역과 일치

* 2021년도 국토교통부 성장촉진지역 대상으로 인구를 유입할 수 있는 농산어촌 주거플랫폼 조성 사업을 공모하는 등 소멸위험에 대응할 수 있는 계획과 사업 발굴에 초점

시·군·구라는 공간단위에서는 지역이 성장촉진지역 또는 인구감소지역인지 아닌지로 이분화된다. 그런데 조금 더 정밀하게 지역을 내부를 살펴보면 한 시군구 내에서도 인구가 증가하는 공간, 유지하는 공간, 감소하는 공간 등 다양한 인구변화 패턴을 확인할 수 있다. 예를 들어, 2021년도 행정안전부 인구감소지역으로 지정된 양양군의 경우 북양양C개통, 신규아파트 공급, 레저인구의 증가 등으로 관련된 장소의 30대 인구가 증가하는 패턴(2017년 대비 2020년)을 보인다(<그림 3-1> 참조). 강원도 양양군은 인구감소지역이지만, 내부의 현황은 공간별로 매우 상이하다.

그림 3-1 | 소지역단위로 살펴본 인구감소지역(강원도 양양군)의 30대 인구변화 사례



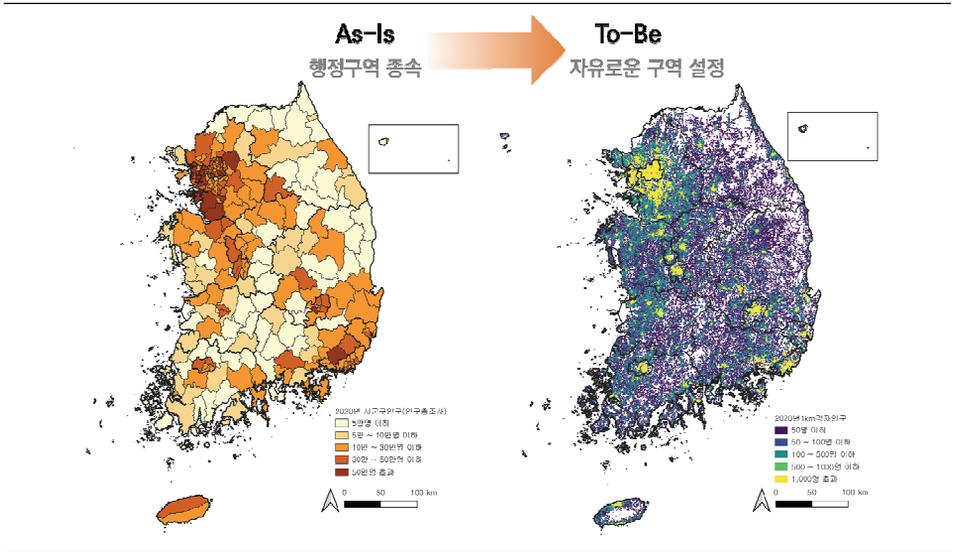
주: 주황색과 붉은색으로 표시된 부분이 30대 인구가 2017년 대비 증가한 지역임
 자료: 저자 작성

최근, 시·군·구보다 더 작은 소지역단위에서의 장래인구 예측의 활용 수요가 증가하고 있다. 2022년 말 공표된 ‘인구구조 변화와 대응방안’에서는 데이터와 과학에 기반하여 정책 효과성을 검증하고, 정밀한 공간단위의 공간 빅데이터를 활용하여 인구정책을 추진해야 함을 강조하였다(저출산고령사회위원회·관계부처, 2022. 12. 28.). 특히 최근 도시·군기본계획은 정부 중심에서, 지자체 단위로, 더 나아가 생활권 단위로 세밀하게 장소기반의 계획을 마련하는 방향으로 전환 중이다. 소지역단위 공간정

보는 N분 도시의 개념 또는 커뮤니티계획 등과 같은 생활권계획에서는 매우 중요한 역할을 한다.

이와 같은 배경에서, 인구감소시대의 데이터 기반 국토정책·계획 수립의 지원을 위한 소지역단위 장래인구 예측 방법론 개발의 방향은 다음과 같다. 첫째, 이 연구에서는 다양한 소지역 공간단위 중 1km 격자단위를 소지역으로 정의하고, 전국의 1km 격자 단위 장래인구를 예측할 수 있도록 방법론을 설계하였다. 다양한 소지역 공간범위 중 격자는 시계열 변화에 따라 구역이 변하지 않아 과거부터 장래까지 국토의 변화를 모니터링하는데 큰 장점이 있다(〈그림 3-2〉 참조). 최근 인구 감소에 따라 행정구역 재편의 수요가 증가함에 따라 행정구역에 유연한 공간단위의 데이터 수요가 증가하고 있다. 그리고 제2장에서 살펴본 것과 같이 1km 격자는 시·군·구 또는 읍·면·동보다 고해상도이면서 더 작은 격자에 비해 데이터 저장과 처리에 소요되는 자원과 및 분석결과 오차를 최소화할 수 있다는 장점이 있다.

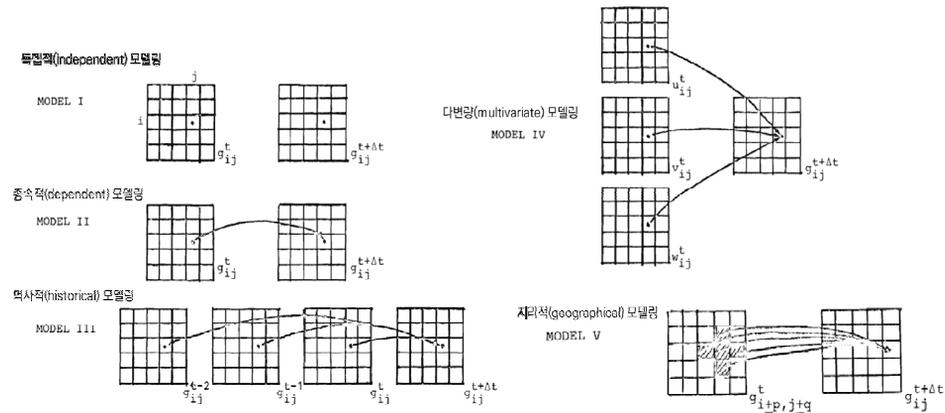
그림 3-2 | 2020년 시군구단위 인구 분포와 소지역(1km 격자)단위 인구 분포 비교



자료: 인구총조사(통계청) 활용 저자 작성(https://kostat.go.kr/portal/korea/kor_ip/6/3/1/1/index.board, 최종 접속일: 2022.11.03.)

두 번째 개발 방향은, 예측 대상 격자에 주변 격자가 영향을 미치는 이웃효과를 고려할 수 있는 방법론을 설계하는 것이다. 기존 시·도 또는 시·군·구 단위의 인구전망 연구에서는 지리적 근접성에 따른 상호작용을 고려할 필요가 없었으나, 분석대상과 활용 데이터가 조밀해짐에 따라 모델은 이웃 공간(격자)과의 상호작용을 반영해야 한다. Tobler(1979)가 제시한 셀룰러 지오그래피(Cellular Geography)의 5가지 모델의 틀에서(〈그림 3-3〉 참조), 이 연구에서 제안하는 방법론은 같은 공간(격자)의 직전 상태에 종속적이고(model 2), 같은 공간의 변화 추세를 반영하는 역사적 모델이며(model 3), 해당 공간 내 다양한 인구변동요인의 영향을 받으며(model 4), 인접한 이웃공간의 영향을 받는 지리적 모델(model 5)를 고려할 수 있도록 개발하였다.

그림 3-3 | Tobler의 셀룰러 지오그래피(Cellular Geography) 모델



자료: Tobler(1979) Figure 2. p4 번역

세 번째 개발 방향은 인구학적 특성(인구학적 변동요인: 인구규모, 사망률, 출생률)과 공간적 특성(토지이용 및 물리적 환경: 건축물 밀도, 주거용 건축물 현황, 토지 경사도 등)의 변화를 복합적으로 고려할 수 있도록 모델을 설계하는 것이다. 장래인구추계에 전통적이고 보편적으로 적용되는 코호트요인법을 적용하는데 시·군·구 이하 단위인 소지역으로 공간단위가 정밀해질 경우 방법론 대상인 성별·연령별 인구 규모가 작아짐에 따라

변동성이 커 그만큼 오차가 증가할 수 있다¹¹⁾. 따라서 행정구역보다 작은 격자단위의 인구예측을 위해서는 광역권 내에서 인구분포에 영향을 미치는 다양한 요인을 고려해야 한다. 선행연구 검토를 통해 주거용 건물 등 토지이용이 소지역단위 인구분포와 큰 관련이 있음을 확인하였다.

네 번째 개발 방향은 시·공간 위계가 복잡하게 연결되어있는 대용량의 공간빅데이터를 학습할 수 있도록 인공지능 기법을 적용하는 것이다. 코호트요인법이 시나리오 기반으로 인구학적 변동을 잘 반영할 수 있고 보편적으로 활용되어왔다는 장점이 있으나, 단순한 시계열의 인구변동요인만 고려한다는 데 한계가 있다. 따라서 이 연구에서는 인구분포에 영향을 주는 인구학적 특성과 공간적 특성을 과거부터 현재까지 데이터로 학습하여 오차를 가장 줄이는 접근방식을 가진 인공지능 기법을 방법론에 활용하고자 한다.

마지막으로, 인공지능 기법을 적용한 인구예측 결과와 전통적인 코호트요인법을 활용한 시나리오 기반의 장래인구 예측 결과를 결합하여 최종 1km 격자단위 장래인구 예측값을 제시함으로써 각 모델의 한계를 보완할 수 있도록 하였다. 인공지능 기법은 과거의 추세를 학습하여 예측력 높은 장래의 패턴을 예측하는데 장점이 있으나, 사망과 출생과 같이 시나리오에 따른 인구증감을 반영하지 못한다는 한계가 있다. 따라서 이를 보완할 수 있도록 이보경(2019)에서 제안한 코호트요인법을 활용하여 격자단위 장래인구를 예측하고 인공지능 기법 적용 결과와 앙상블을 통해 결합하였다. 앙상블은 데이터를 학습하여 패턴을 예측하였을 때 데이터의 변동에 따라 예측 결과의 변동성이 큰 경우 이를 최소화하여 안정성을 확보하기 위한 방법이다. 이 연구의 결과값이 연속형 변수(인구수)임을 감안하여 앙상블 중 보팅(voting)¹²⁾ 방법론을 적용하여 각 모델의 결과의 평균을 최종 모델의 결과로 제시하였다.

11) 김기환, 국토연구원 개최 소지역단위 인구 전망 연구 동향 세미나(2022.04.15.) 발표자료

12) 보팅(voting)은 서로 다른 알고리즘의 결과를 결합하는 방법으로, 예측 결과값이 명목형 변수인 경우 하드보팅, 소프트보팅, 가중치 투표 등으로 하나의 값을 선택하고 이 연구와 같이 연속형인 경우(장래인구 규모) 각 알고리즘을 통해 도출된 결과의 평균을 활용하여 두 모델을 결합

2. 격자단위 장래인구 예측 AI 모델(Gridded Population Forecast with AI, GPAI) 아키텍처 설계

이 연구에서는 소지역단위 장래인구예측 방법론으로 ‘격자단위 장래인구 예측 AI 모델(Gridded Population Forecast with AI, 이후 GPAI)’을 개발하였다. GPAI는 출산율, 사망률 등 인구학적 변동요인과 토지이용 등 공간적 특성을 모두 장래인구 예측에 반영한 방법론이다(<표 3-1> 참조).

표 3-1 | 격자단위 장래인구 예측 AI 모델(Gridded Population Forecast with AI, GPAI) 아키텍처 설계

구분	모델1		모델2
주요 가정	<p>과거의 인구변화 추세를 학습하고, 장래의 공간적 특성 변화 예측을 통해 격자별장래인구를 전망</p> <ul style="list-style-type: none"> • 과거의 해당 격자 및 주변 격자의 인구 변화가 장래의 인구 규모에 영향 • 과거의 해당 격자 및 주변 격자의 공간적 특성이 장래의 공간적 특성에 영향 	앙상블 보팅 (Ensemble Voting) 결합	<p>격자단위 인구학적 구조(성별 및 연령)와 인구학적 변동요인에 따른 자연적 증감을 통해 격자별장래인구를 예측</p> <ul style="list-style-type: none"> • 현재 격자단위의 성별 연령별 인구구조에 인구학적 변동요인을 적용하여 자연적 증감에 따른 장래인구변화 예측 ※ 인구가동에 따른 인구의 사회적 증감은 시도단위에서 고려
활용 데이터	<p>인구학적 특성(시도 및 격자 단위)과 공간적 특성(격자단위)을 변수로 활용</p> <ul style="list-style-type: none"> • 2050년 시도단위 추계 사망률, 출생률, 순이동률 등 인구변동요인(통계청, 2022)과 격자단위 인구 변화(2000~2020)를 활용 • 격자단위 건축물수, 주택수, 주거용 연면적, 경사도 등 공간적특성(2000~2020)를 활용 		<p>인구학적 특성(시도단위)과 인구 구조적 특성(격자단위)을 변수로 활용</p> <ul style="list-style-type: none"> • 2050년 시도단위 추계 사망률 및 출생률과 같은 인구변동요인(통계청, 2022) • 2020년 격자단위 성별/연령별 인구구조(행정안전부 주민등록 통계)를 기준인구로 설정
적용 방법론	<p>3개 단계(인구학적-공간적 특성 학습, 장래의 공간적 특성 변화 예측, 장래의 인구학적-공간적 특성 기반의 격자단위 인구 예측)으로 구성</p> <ul style="list-style-type: none"> • 1단계: 인구학적 특성 공간적 특성 및 격자단위 인구분포와의 관계를 학습을 위해 이웃효과를 고려한 랜덤포레스트(Random Forest) 수행 • 2단계: CNN-LSTM(Convolutional Neural Network-Long short Term Memory) 알고리즘을 적용하여 공간적 특성 예측 • 3단계: 장래의 인구학적 변동요인과 2단계 공간적 특성 예측 결과를 1단계에서 학습한 랜덤포레스트에 적용하여 격자단위 장래인구를 예측 		<p>이보경(2019)에서 제안한 격자단위 장래인구 추계를 위한 코호트 요인법을 활용</p> <ul style="list-style-type: none"> • 1단계: 격자단위 기준인구 정의 • 2단계: 격자별 코호트 요인법적용 • 3단계: 격자별 장래인구 가중치 산출 • 4단계: 시도단위 장래인구 추계를 격자별 가중치를 기준으로 배분하여 통계청과 정합성을 갖춘 격자단위 장래인구 예측값 도출
최종 결과	1km 격자단위 장래인구 예측(2025년 ~ 2050년)		

자료: 저자 작성

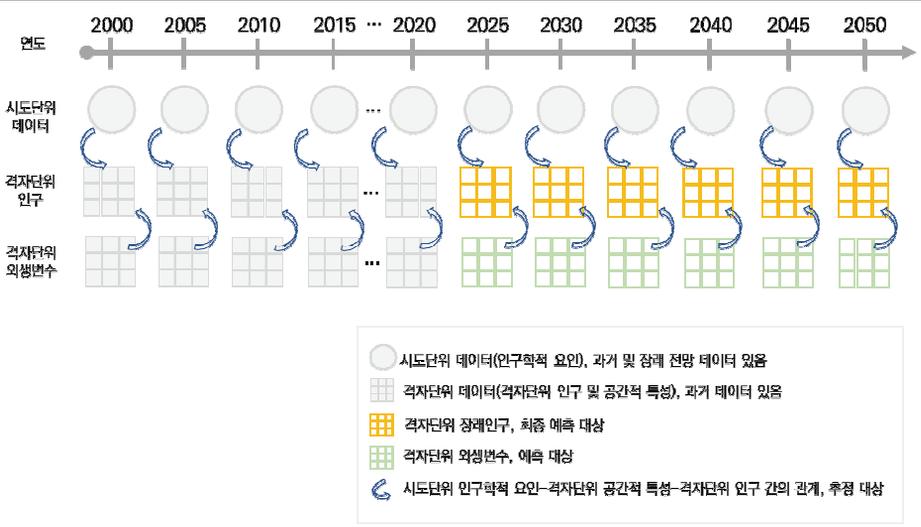
모델1은 인구학적 특성과 공간적 특성을 고려한 인공지능 기법 기반 모델이며, 모델 2는 인구학적 특성에 따른 변화에 초점을 맞춘 코호트요인법 기반 모델(이보경, 2019)이다. 두 모델의 격자단위 장래인구 예측 결과를 앙상블 중 보팅(voting) 방법을 통해 결합하여 2025년부터 5년단위로 2050년까지 1km 격자단위 장래인구 예측값을 제시한다. 각 모델에 대한 상세한 설명은 다음과 같다.

1) [모델1] 장래의 인구학적 특성과 공간적 특성 변화를 고려한 인공지능 기법

인공지능 기법을 통해 충분한 사전정보를 학습하면 관찰하고자 하는 대상의 숨겨진 패턴을 예측할 수 있다. 2000년부터 2020년까지 격자단위 인구 데이터를 학습하여 2050년까지 격자단위 인구를 예측하는 인공지능 기법을 개발하였다. 통계청에서 공표한 2000년부터 2020년까지의 인구통계를 전체 학습데이터로 구축하고 향후 예측 대상이 되는 2050년까지의 시도별 인구 추계정보를 기준으로 격자단위 인구를 예측하는 배분적 접근방법을 제안하였다.

지난 20년간의 인구정보뿐 아니라 인구분포에 영향을 미치는 공간적 특성 정보를 외생변수로 활용하여 풍부한 학습 데이터를 구축하고자 하였다. <그림 3-4>의 가로축 실선은 2000년부터 2050년까지 시간의 흐름을 의미한다. 그리고 회색으로 표시된 각 데이터는 실제 또는 이미 전망된 값이 있음을 의미하고, 다른 색을 모델링을 통해 도출해야 하는 대상을 나타낸다. 예를 들어, 2000년부터 2020년까지 시·도단위 실제 인구와 2050년까지 장래인구추계는 통계청에서 공표하고 있어 해당 데이터를 활용할 수 있다. 2000년부터 2020년까지 격자단위 인구의 경우 통계청 SGIS에서 공개하고 있다. 그리고 인구분포에 영향을 주는 공간적 특성을 나타내는 외생변수도 2000년부터 2020년까지는 데이터 가공 및 융합을 통해 1km 격자단위 데이터로 구축할 수 있다. 그런데 2020년 이후 격자단위 장래인구(노란색), 격자단위 외생변수(초록색) 및 인구학적요인-격자단위 공간적 특성-격자단위 인구 간의 관계(파란색)은 데이터 학습을 통해 예측 또는 규명해야 할 대상이다.

그림 3-4 | 시간에 따른 공간구분별 인구 데이터 구축 현황



자료: 저자 작성

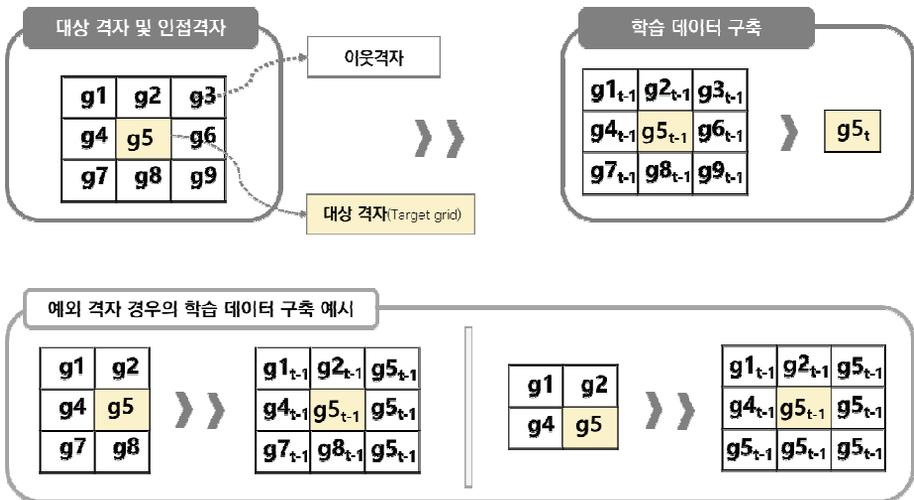
격자단위 장래인구예측을 위해 활용할 수 있는 정보를 파악하였고, 예측 대상을 정의하였다. 다양한 공간 위계를 가진 지난 20년간의 시계열 데이터를 학습하여 앞으로 약 30년간의 장래의 패턴을 예측하기 위해서는 (1) 시공간 데이터의 특성 및 위계 정보를 활용하고 (2) 사전정보 부족에 따른 과소식별로 인한 오차를 최소화할 수 있는 모형이 필요하다. 따라서 시공간 위계 정보 활용을 위해 (1-①) 격자단위 데이터의 이웃효과와 (1-②) 시계열 및 공간 간의 논리적 위상을 모형에 반영하였다. 그리고 과소식별 오차 최소화를 위해 (2-①) 점진적 온라인 학습 방법과 (2-②) 다중 모형결합방식을 적용하였다.

(1) 시공간 데이터의 특성 및 위계 정보 활용 모형

① 격자단위 정보의 시공간적 이웃효과 반영

Tobler의 셀룰러 지오그래피 모델 개념을 바탕으로 대상 격자 정보와 함께 이웃 격자의 정보까지 고려할 수 있도록 학습데이터를 구축하였다(〈그림 3-5〉 참조). 시간적(historical) 경향성과 공간적(spatial) 상관성을 모두 반영하기 위해 격자로 취합되는 목표 시점(t)의 목표 격자(target grid)의 값을 대상 격자와 이웃 격자의 이전 시점(t-1) 정보를 통해 예측하여 격자 간의 시공간적 관계를 학습하도록 하였다. 전국의 모든 격자를 대상으로 시공간적 이웃효과를 반영할 수 있도록 학습데이터를 구축하였으며, 국토 경계에 위치하거나 섬 지역 등 이웃 격자 정보 확보가 어려운 격자는 대상 격자 정보를 반복하여 일관된 학습데이터 구성을 유지하는 방식으로 설계하였다.

그림 3-5 | 시공간적 분포를 고려한 모형의 격자 데이터 배분 방법 도식화

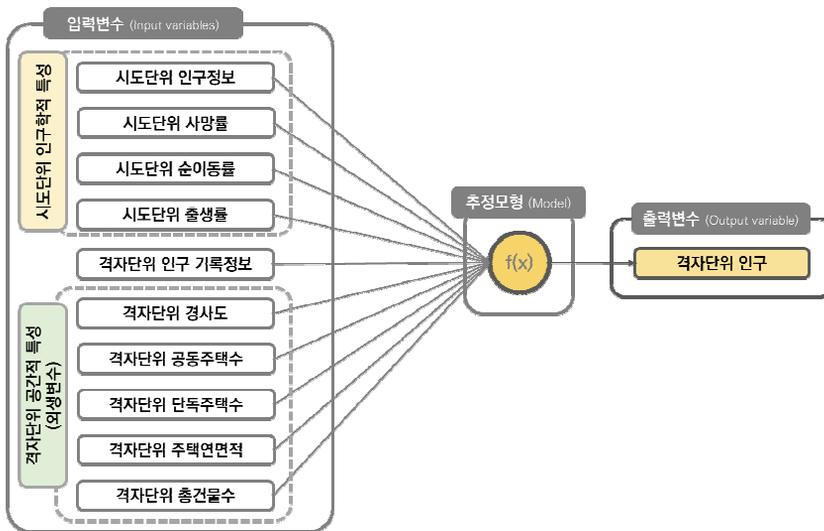


자료: 저자 작성

② 학습 대상 변수의 시계열 및 공간 간의 논리적 위상 고려

학습 대상 데이터의 공간단위는 시·도단위와 격자단위가 혼합되어 있다. 장래인구 추계 및 인구변동 요인의 경우 2000년부터 2050년까지 시·도단위 데이터를 활용하였다. 2000년도부터 2020년도까지 공간적 특성(경사도, 공동주택수, 단독주택수, 주택연면적, 총 건물 수)의 경우 다양한 출처로부터 데이터를 수집하여 격자단위로 가공하였다¹³⁾. 과거 인구의 경우 2000년부터 2020년까지 격자단위 데이터를 구득할 수 있다. 따라서 장래인구 예측을 위해서는 시계열과 시도-격자 간의 관계 등 데이터의 복합적인 관계에 대한 논리적인 위상을 고려해야한다. 입력변수로부터 출력변수를 학습하여 격자단위 인구를 추정하는 학습구성은 <그림 3-6>과 같다. 본 연구를 위해 활용할 수 있는 각 입력변수를 각 시공간 단위에 맞게 구성한 뒤, 추정모형을 거쳐, 출력변수로 그 패턴을 학습하도록 구성하였다.

그림 3-6 | 예측모형의 입력과 출력변수 설정 도식화



자료: 저자 작성

13) 활용 변수에 대한 상세한 설명은 본 장의 3절 참조

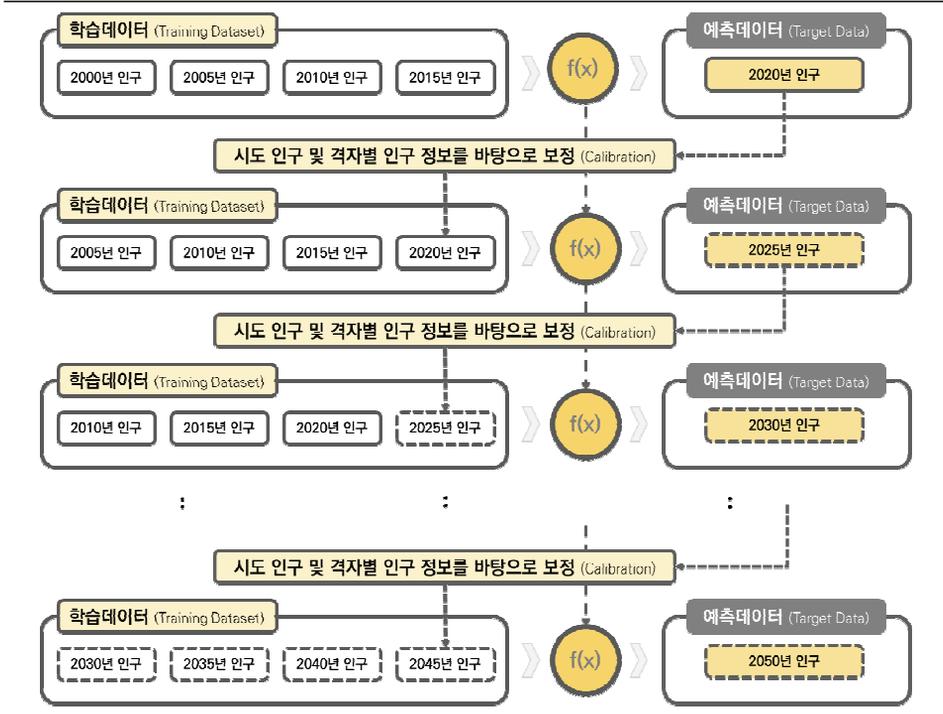
(2) 과소식별에 따른 오차를 최소화하기 위한 모형

① 점진적 온라인 학습

인공지능 기법을 개발할 때 학습 대상이 되는 사전정보의 시계열이 충분히 확보되지 못한 경우 부정확한 결과를 초래하는 과소식별(under-estimation) 문제를 일으킬 수 있다. 머신러닝 모형에서 활발하게 활용하고 있는 점진적 온라인 학습(incremental online learning)을 적용하여 과소식별을 극복하고자 하였다. 점진적 온라인 학습 방법은 최근 재귀순환신경망(RNN, Recurrent Neural network) 및 LSTM 딥러닝 모형 등에서 주로 사용되는 학습 방법으로, 시계열 데이터를 전수로 학습하는 경우 발생하는 계산시간 부담과 오차를 줄이기 위해 고안된 방법론이다. 점진적 온라인 학습은 훈련데이터를 시계열 순서대로 일정 단위 분할 한 뒤 분할된 데이터를 순서대로 일부 중첩하여 직전 데이터의 학습 결과를 바탕으로 다음 데이터의 학습과 예측을 하는 방식으로, 점진적으로 그 오차를 보정해가며 예측도를 높이는 것이 특징이다.

이 연구에서는 4개 시점의 데이터를 바탕으로 장래 1개 시점의 데이터를 추정하는 방식으로 점진적 온라인 학습을 구현하였다(〈그림 3-7〉 참조). 훈련과 검증에 충분한 양의 학습데이터가 아닌 점을 고려하여, 2000년부터 2015년까지의 4개 시점(2000, 2005, 2010, 2015)의 데이터를 초기 학습대로 구성한 뒤, 2020년도 실제 값을 잘 예측할 수 있도록 모형을 지도학습(Supervised learning)하고 그렇게 학습된 아키텍처를 유지하였다. 다음 학습에서는 직전 단계에서 생성된 모형 아키텍처를 통해 2005년부터 2020년까지의 데이터를 학습하여 2025년 격자 인구를 예측한다. 이와 같은 학습방식을 동일하게 2050년까지 격자단위 장래인구 예측에 적용하였다.

그림 3-7 | 점진적 온라인 학습 방법 적용 예시

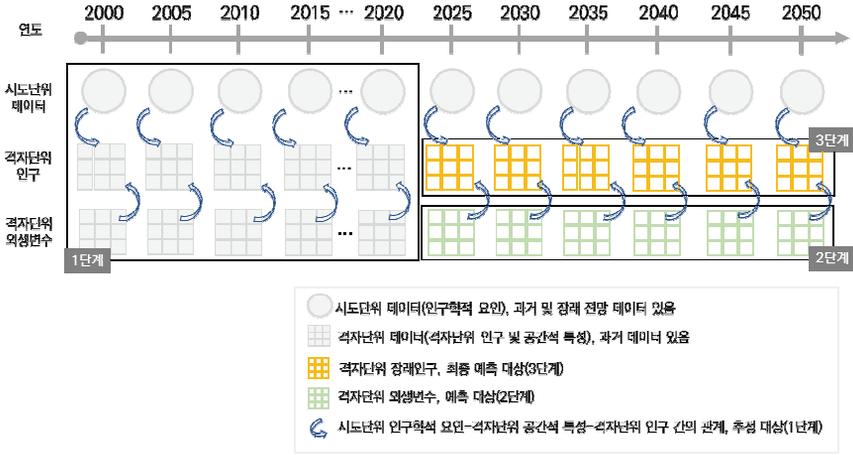


자료: 저자 작성

② 다중 모형결합방식

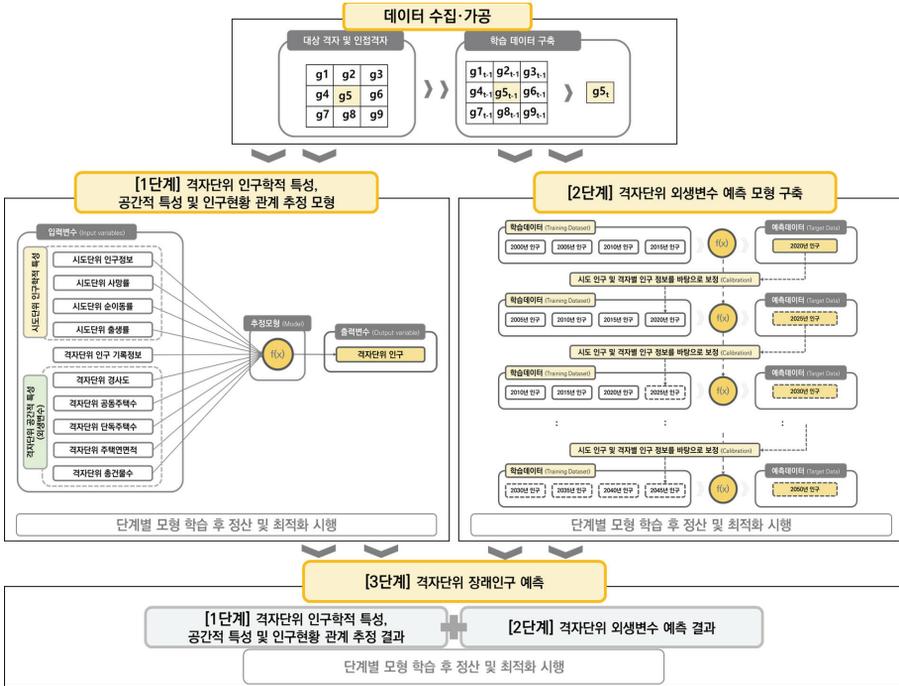
과소식별에 기인하는 오차를 최소화하기 위한 두 번째 방법으로 시공간 논리적 구성을 최대한 활용하기 위해 다중 모형결합방식으로 모델1을 설계하였다. 다중 모형결합 방식은 Yao 외(2018, 2019), Liang 외(2019), Song·Han·Stettler(2020) 등 최근 관련 연구에서 활발하게 시도되고 있는 방법이다. 이 연구에서도 GPAI의 모델1을 3개 단계로 구성하였다(〈그림 3-8〉 참조). 모델1의 [1단계]는 시공간 특성을 과거로부터 학습하고, [2단계]는 장래 외생변수를 추정하며, [3단계]에서는 앞의 두 단계의 결과를 활용하여 격자단위 장래인구를 예측하는 과정이다. 다양한 모형 및 알고리즘이 복합적으로 작용하는 만큼, 최적화 방법과 학습 결과에 대한 정산 단계를 거쳐 모형의 학습 결과 오차를 최소화하도록 설계하였다(〈그림 3-9〉 참조).

그림 3-8 | GPAI 다중 모델 결합 모형 아키텍처



자료: 저자 작성

그림 3-9 | GPAI 모델1 개념 순서도



자료: 저자 작성

(3) 단계별 적용 알고리즘

모델1의 첫 번째 단계에서는 랜덤포레스트(random forest)¹⁴⁾ 알고리즘을 활용하여 과거 연도의 시도단위 및 격자단위 인구, 인구학적 특성(출생률, 사망률, 순이동률 등), 공간적 특성(건물, 주거용 연면적 등) 간의 관계와 시간에 따른 변화 추세를 학습하였다. 이전 시점으로부터 추출할 수 있는 인구 및 외생변수 간의 시공간 상관관계를 바탕으로 현시점 격자단위 인구의 분포를 학습하고 그 오차를 최소화하는 방식으로 학습을 수행하였다.

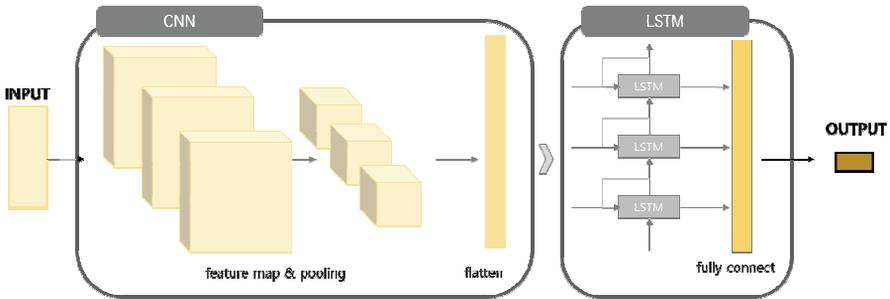
지도학습(supervised learning)에 속하는 랜덤포레스트는 주어진 각 데이터로부터 연고자 하는 목표값 간의 관계를 간단한 분류모형(CART, classification regression tree)을 이용하여 상관관계를 도출하고, 이를 계층적 관계로 산출한 뒤 다수의 상관관계를 종합하여 앙상블의 형태로 평균치를 치환하는 접근방법이다. 모델1의 1단계에서 랜덤포레스트 알고리즘을 활용한 가장 큰 이유는 랜덤포레스트 학습과정에서 미세한 오차를 적용하여 모형 적용 결과에 영향을 주는 정도를 가늠할 수 있는 검증단계가 포함되어 입력변수의 중요도와 민감도 등을 함께 분석해볼 수 있기 때문이다. 즉, 어떤 변수가 장래인구 분포에 큰 영향을 미치는지 해석할 수 있다는 장점이 있다.

두 번째 단계에서는 CNN-LSTM(Convolutional Neural Network-Long short Term Memory) 알고리즘을 활용하여 2020년 이후 정보가 없는 격자단위 공간적 특성을 예측하였다. 이 단계에서는 학습한 공간적 특성 대상은 공동주택 수, 단독주택 수, 주거용 건물연면적 및 건축물 수와 같은 외생변수이다. CNN-LSTM 알고리즘에 2000년부터 2020년까지 데이터를 적용하여 학습하고, 각 변수의 2050년까지의 격자단위 값을 예측하였다.

14) 랜덤포레스트라는 이름은 독립변수와 종속변수 간의 각 관계를 파악하는 단계가 나무의 가지가 뿔어 나가는 방식과 흡사하고, 이후 앙상블의 단계에서는 다수의 나무가 숲을 이루는 방식과 흡사하다 하여 붙여진 이름임. 랜덤포레스트는 소수의 계층적 관계도로부터 발생할 수 있는 과소식별의 오차를 다수의 샘플링을 통해 보완한 모형으로, 시계열분석뿐만 아니라 분류분석 등 그 다양하게 활용되고 있다. 구체적인 랜덤포레스트에 대한 설명은 <부록 1>과 같음.

CNN-LSTM 알고리즘은 최근 딥러닝 알고리즘의 발달과 함께, 딥러닝 알고리즘 간 이종 결합을 통해 시도되고 있는 모형이다. 합성곱(convolution) 모형에서 오는 데이터 특성 변환(feature map) 및 데이터 차원 축소의 강점과 LSTM 모형에서 오는 점진적 온라인 학습의 시계열 강점을 종합하여 융복합한 것이 특징이다(〈그림 3-10〉 참조). CNN-LSTM 알고리즘은 방대한 양의 데이터를 사람이 확인하기 힘든 미세한 데이터의 특성까지 포괄하는 섬세함뿐만 아니라 본 과업과 같이 제한된 시계열 데이터가 구축될 때 소량의 데이터로부터 가능한 데이터 특성을 다수 도출할 수 있다는 장점이 있다. 최근 다양한 연구에서 활용되고 있는 CNN-LSTM은 입력 데이터의 특성을 추출하는 단계(feature map)를 거쳐 차원을 변환(pooling)하고, 시계열 특성을 추출하는 단계(LSTM)를 통해 점진적 온라인 학습방식으로 데이터를 학습하였다.

그림 3-10 | CNN-LSTM 모형 개념도



자료: 저자 작성

세 번째 단계는 [1단계]의 랜덤포레스트를 통해 학습한 인구학적 특성-공간적 특성-격자단위 인구 간의 관계에 [2단계]에서 CNN-LSTM으로 예측한 공간적 특성인 외생변수를 적용하여 2050년까지 격자단위 장래인구를 예측하는 과정이다. 따라서 [2단계]의 예측값을 [1단계]의 랜덤포레스트 알고리즘에 적용하여 결과값을 도출하기 때문에 별도의 인공지능 기법 학습을 수행하지는 않는다. 3단계에서는 GPAI 모델에서 예측한 격자단위 장래인구 값이 통계청에서 공표한 장래인구 추계와 정합성을 갖출 수

있도록 배분의 기준을 마련하였다. 그리고 진행 과정에서는 (식 3-1)에서부터 (식 3-5)에서 제시된 목적함수와 조건제약식 등을 반복적으로 적용하여 장래 연도의 격자별 인구추계의 신뢰도를 유지하고자 노력하였다.

목적함수(Objective function)

$$Min \sum_{i=1}^I (\text{모형의 격자별 인구추정}_i - \text{격자별 인구추정 보정값}_i) \quad (\text{식 3-1})$$

제약 조건식(Subject to, s.t.)

$$\text{격자}_{ik} \text{의 인구} \geq 0 \quad (\text{식 3-2})$$

$$\text{격자}_{ik} \text{의 인구} = 0, \text{ 만약, 격자}_{ik} \text{의 주택수} = 0 \quad (\text{식 3-3})$$

$$\text{통계청 추계 } t \text{시점 시도}_k \text{ 장래인구} = \sum_i^{I_k} \text{격자}_{ik} \text{의 } t \text{시점 장래인구} \quad (\text{식 3-4})$$

$$\forall i \in I, k \in K \quad (\text{식 3-5})$$

(식 3-1)은, 본 연구에서 개발한 목적함수로, 격자_{*i*}의 모형 추정 인구와 정산 단계를 통해 보정되는 격자_{*i*}의 인구 간의 차이를 최소화하는 것을 의미한다. (식 3-2)부터 (식 3-5)는 본 목적함수의 제약 조건식이다. (식 3-2)은 시도_{*k*}에 속한 격자_{*ik*}의 인구는 음수가 될 수 없다는 조건을 의미한다. (식 3-3)은 시도_{*k*}에 속한 격자_{*ik*}의 인구추계 결과는 추계 해당 연도의 격자_{*ik*}의 주택 수 등 거주 건축물이 없는 경우 거주 인구도 0명으로 예측한다는 의미이다. (식 3-4)는 통계청에서 공표한 2050년까지의 시도단위 장래인구 추계의 합은 해당하는 시도에 위치한 모든 격자_{*ik*}의 같은 연도 장래인구 예측값의 합과 일치한다는 의미이다. 이를 통해 GPAI 모델로 예측한 결과와 통계청에서 공표한 장래인구추계가 시도단위에서 정합성을 갖게 된다. (식 3-5)는 해당 목적함수와 제약식이 모든 시도와 격자에 동일하게 적용됨을 의미한다.

2) [모델2] 인구학적 변동요인을 고려한 코호트요인법

모델2는 이보경(2019)가 코호트요인법을 활용하여 격자단위 장래인구추계를 시물레이션한 방법론을 활용하였다(〈그림 3-11〉 참조). 첫 번째 단계에서는 국토지리정보원의 500m 격자단위 성별·연령별 인구를 기준인구로 설정하였다. 그리고 두 번째 단계에서는 격자가 속한 지역의 출산율과 사망률을 적용하여 5년 단위로 격자별 코호트요인법에 따른 인구 변화를 전망하였다. 지역별·성별·연령별 출산율과 사망률을 격자단위 인구에 적용함으로써 격자별 자연적 인구 증감을 예측할 수 있다.

그리고 3단계에서는 5년마다 자연적 인구증감이 반영된 격자단위 인구를 활용하여, 격자가 속한 시·도단위 총 장래인구에서 해당하는 격자의 장래인구의 비중을 가중치로 산출하였다. 이미 격자단위의 장래인구를 예측하였으나, 이는 자연적 증감에 의한 변화만 고려한 것으로 시·도 간의 인구이동에 따른 사회적 증감을 인구변화에 고려하기 위해 가중치를 산출한 것이다. 4단계에서는 이 가중치를 통해 통계청에서 공표한 시·도단위 장래인구추계를 할당하였다. 이를 통해 통계청에서 공표한 장래인구추계의 시·도단위에서 정합성을 가지는 격자단위 장래인구 예측값을 구할 수 있다. 마지막 단계에서는 모델과의 격자크기를 1km로 맞추기 위해 500m 격자단위로 예측한 장래인구 추계를 위치에 따라 1km로 집계하였다.

그림 3-11 | GPAI 모델2 개념 순서도



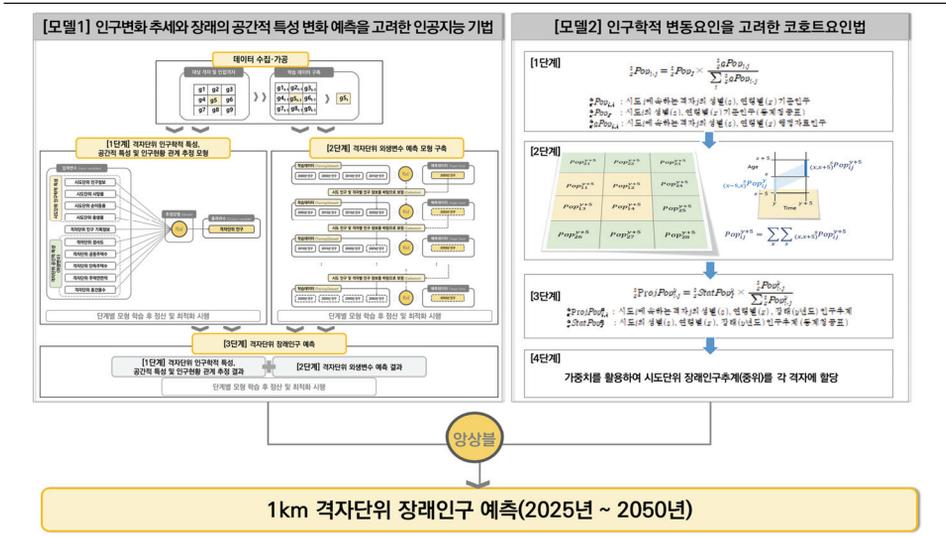
자료: 이보경(2019) 연구결과 활용 저자 재작성

3) 앙상블을 통한 모델1과 모델2 결합과 장래인구 예측

모델1과 모델2에서 도출된 격자별 장래인구 예측값을 앙상블의 보팅으로 결합하였다 (<그림 3-12> 참조). 앙상블은 데이터에 기반한 알고리즘의 과대 또는 과소적합을 방지하기 위해 다양한 알고리즘이 조화를 이루어 최종값을 도출하는 데이터마이닝 방법이다. 앙상블에는 보팅(voting), 배깅(bagging), 부스팅(boosting) 등으로 이루어져있다. 보팅은 서로 다른 알고리즘을 결합하는 것, 배깅은 같은 유형의 알고리즘을 활용하되 학습 데이터 샘플링을 부스트래핑을 통해 다양하게 추출하는 것 그리고 부스팅은 다양한 알고리즘을 순차적으로 학습하여 직전의 알고리즘 결과를 보완해나가는 방식이다.

이 연구에서는 모델1의 결과와 모델2의 서로 다른 알고리즘을 결합하기에 앙상블 중 보팅 방법을 사용하였다. 그리고 결과값이 격자별 총 인구수인 연속형 변수이기 때문에 두 값을 평균값을 시도단위의 장래인구 추계를 할당하는 기준으로 활용하였다. 앙상블 기법을 적용하여 두 모델에서 활용한 알고리즘의 장점을 모두 반영하면서 한계를 보완할 수 있는 복합적 GPAI를 설계할 수 있었다.

그림 3-12 | 모델1과 모델2 앙상블 결합 기반의 GPAI 아키텍처



자료: 저자 작성

3. GPAI 학습과 장래인구 예측을 위한 공간 데이터

1) 격자로 본 우리나라 인구분포

2050년까지 1km 격자단위 장래인구 예측을 위해 통계청 SGIS의 인구총조사 센서스 기반 1km 격자단위 인구 데이터를 활용하였다. 통계청은 2022년 격자단위 인구통계를 공표하면서 2000년부터 2020년까지 격자단위 인구정보를 공개하였다(통계청, 2022. 07. 01. 보도자료). 이 연구에서 분석대상으로 설정한 격자는 전국을 포함하는 107,442개의 1km 격자이다. 해당 격자들의 2000년, 2010년, 2020년 격자단위 인구 현황을 비교한 결과(〈표 3-2〉 참조), 1km 격자 평균 인구 수(인구밀도)는 2000년 428명, 2010년 448명, 2020년 470명으로 증가하는 추세이다¹⁵⁾. 격자단위 인구 분포의 표준편차는 약 2,500명으로 우측 꼬리가 긴 형태의 분포를 보인다(히스토그램 왜도와 첨도 참조). 따라서 격자단위 인구밀도의 중앙값은 20~23명 수준으로 나타났다.

표 3-2 | 2000~2020년 격자단위 인구 기초통계량

구분	2000	2010	2020
평균(인구밀도)	428.01	447.70	469.50
중앙값	23	20	23
표준편차	2,580.34	2,586.23	2,488.09
최대값	60,853	45,934	43,846
히스토그램 왜도	9.86	9.02	8.24
히스토그램 첨도	117.15	97.94	83.18

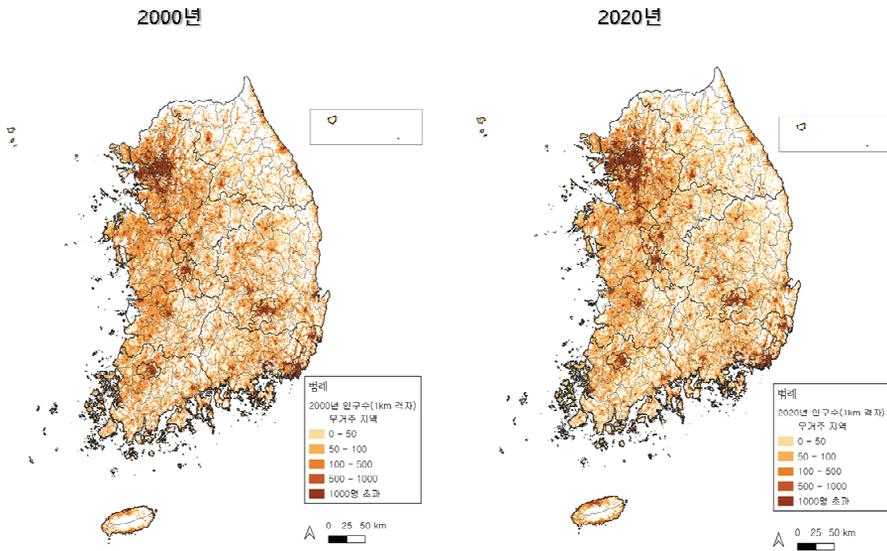
자료: 통계청 격자단위 인구총조사 자료를 활용하여 저자 작성

2000년과 2020년의 1km 격자단위 인구분포는 〈그림 3-13〉과 같다. 수도권 및 특·

15) 통계청 공표 격자단위 인구통계는 비밀보호기법을 적용한 결과로 실제값이 0~4 사이인 경우 0을, 1~5 사이인 경우 5를, 6 이상인 경우 M-2~M+2의 값으로 변환되었음. 따라서 기초통계량의 수치가 센서스 기반 인구총조사의 참값과 상이할 수 있음.

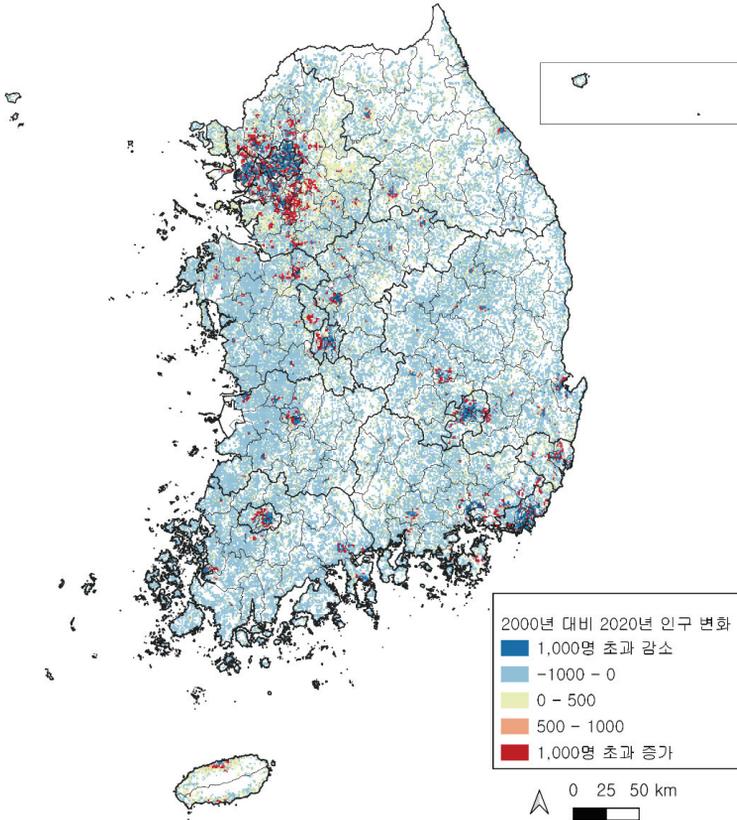
광역시에 1km 격자 내 1,000명 이상 거주하는 인구 지역이 밀집되어있다. 격자단위 인구통계를 활용하면 행정구역의 변경과 무관하게 공간상의 인구 시계열 변화를 살펴볼 수 있다. 지난 20년간의 인구변화를 살펴보면(<그림 3-14> 참조), 특·광역시 원도심 인구는 감소하고, 주변 지역(예: 경기도 화성시, 오산시, 충청남도 천안시, 경상남도 양산시, 경상북도 경산시 등)의 인구가 크게 증가하는 패턴이 나타난다. 즉, 지난 20년간 도심의 영역이 특·광역시의 인근 교외로 확장되었다고 볼 수 있다.

그림 3-13 | 2000년 및 2020년 전국 인구분포(1km 격자)



자료: 저자 작성

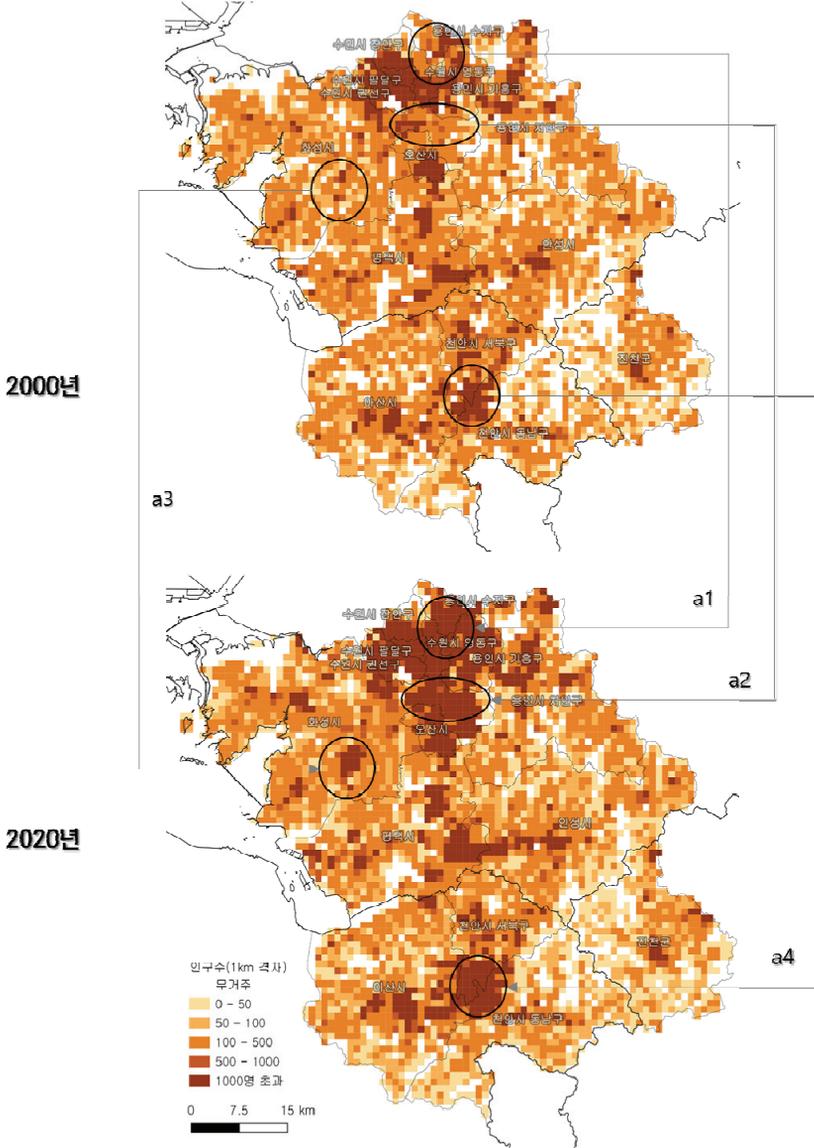
그림 3-14 | 2000년 대비 2020년 인구 증감률(1km 격자)



자료: 저자 작성

인구 증가 패턴이 눈에 띄는 지역인 경기도 남부와 충청남도 북부 지역의 2000년과 2020년 격자단위 인구수를 확대하면 <그림 3-15>와 같다. ‘a-1’은 수원시 영통구의 광고 일대와 용인시 수지구 성북동, 상현동 일대, ‘a-2’는 화성시 병점, 동탄 일대와 오산시 세마동, 신장동 일대이다. ‘a-3’은 화성시 향남읍, ‘a-4’는 천안시 서북구 남측(백석동, 쌍용동, 불당동 등)과 동남구 서측(봉명동, 일봉동 등)이다. 지난 2000년부터 2020년까지 격자단위 인구 현황과 변화 추세를 학습하여 2050년까지 격자단위 장래인구를 예측하고자 한다.

그림 3-15 | 경기 남부와 충남 북부 일대 2000년과 2020년 전국 인구분포(1km 격자)



주: 'a-1'은 수원시 영통구의 광교 일대와 용인시 수지구 성북동, 상현동 일대 / 'a-2'는 화성시 병점, 동탄 일대와 오산시 세마동, 신장동 / 'a-3'은 화성시 향남읍 일대/ 'a-4'는 천안시 서북구 남측(백석동, 쌍용동, 불당동 등)과 동남구 서측(봉명동, 일봉동 등) 일대

자료: 저자 작성

2) 장래인구 분포 영향요인

(1) 인구학적 특성

격자단위 인구분포를 학습하고 장래의 변화를 예측하기 위해 과거로부터 장래까지 인구학적 특성(사망, 출생, 인구가동) 정보가 필요하다. 분석대상 공간단위인 1km 격자단위 로 시계열 데이터가 축적되어 있다면 이를 활용하여 인구학적 변동요인 각각을 예측할 수 있다. 그러나, 관련 데이터가 시·도단위로 구축·공표되고 있다, 따라서 이 연구에서는 Breidenbach·Kaeding·Schaffner(2019)이 독일의 1km 격자단위 인구 예측 시 활용한 것과 같이 시도 내에서의 인구학적 변동요인은 균일하다는 가정을 하고 시도 단위 통계를 그 안에 속한 격자단위에 그대로 반영하였다. 격자단위 장래인구 예측 방법론에 활용한 인구학적 특성 관련 변수의 정의와 데이터 출처는 <표 3-3>과 같다.

표 3-3 | 인구학적 특성의 데이터 출처

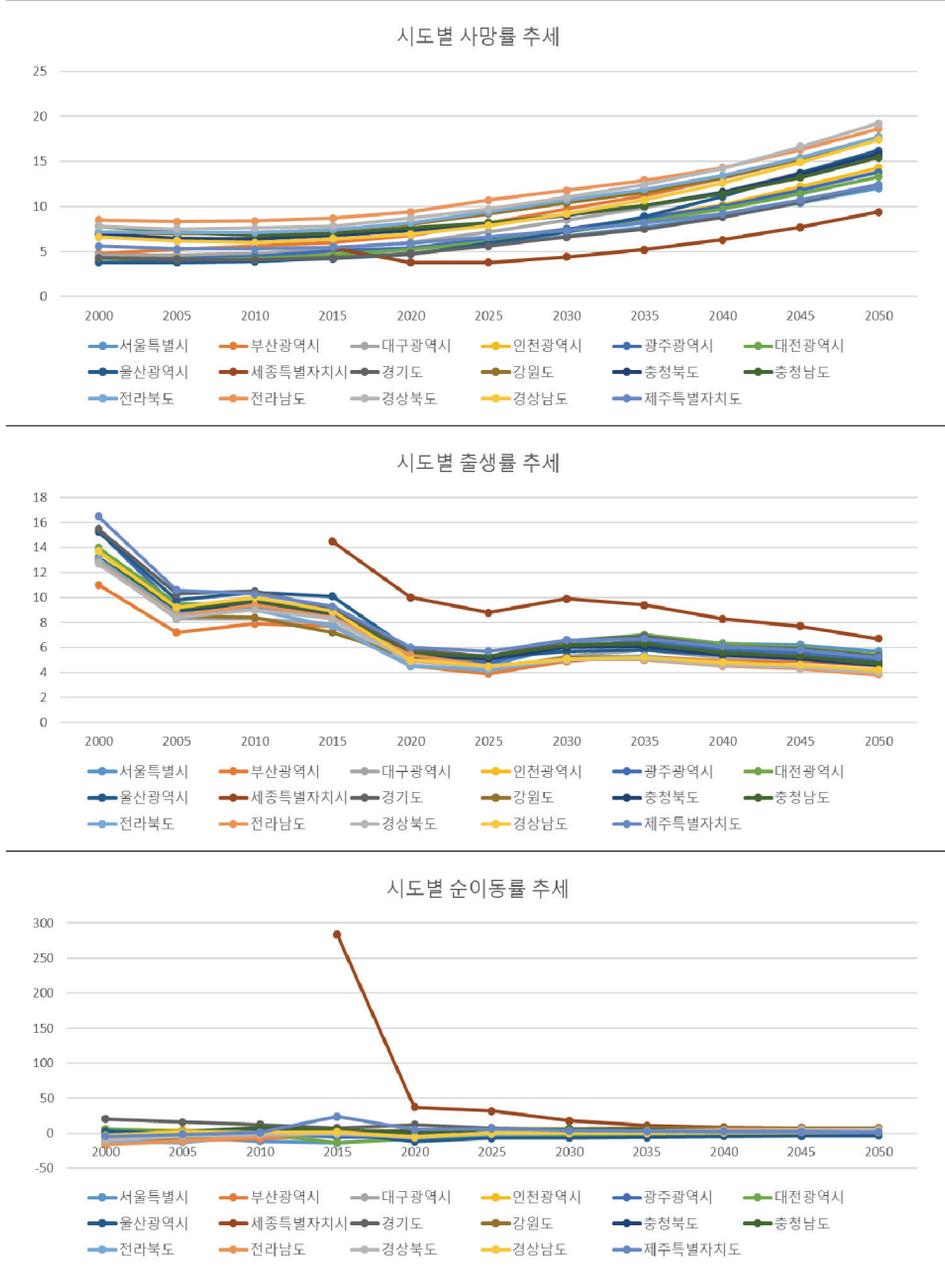
구분	내용	시점	출처	공간 단위
사망률	인구천명당 사망자수	2000~2020	인구동향조사 - 시도/인구동태건수	시도
		2021~2050	장래인구추계 - 장래인구변동요인/시도	
출생률	인구천명당 출생아수	2000~2020	인구동향조사 - 시도/인구동태건수	
		2021~2050	장래인구추계 - 장래인구변동요인/시도	
순이동률	인구천명당 이동자 수	2000~2021	국내인구가동통계	
		2022~2050	장래인구추계 - 장래인구변동요인/시도	
인구 수	시도단위 총 인구수(과거) 및 장래인구추계	2000~2020	통계청 인구총조사	
		2025~2050(5년)	장래인구추계	

자료: 통계청 장래인구추계, 인구동향조사, 인구총조사 및 국내인구가동통계 통계설명자료(<https://kosis.kr/>, 최종 접속 일: 2022.06.23.)를 바탕으로 저자 작성

2000년부터 최근까지는 실제 인구동태 및 인구이동 통계를 활용하여 데이터를 구축하였고, 이후 2050년까지는 통계청 장래인구추계를 위해 공표된 변동요인 추세를 활용하였다. 사망률과 출생률의 경우 2020년을 기준으로 비교하였을 때, 계산한 값과 통계청에서 장래인구추계를 위해 활용한 값이 모두 일치하는 것을 확인하였다. 따라서 2000년에서 2020년까지는 과거의 데이터를 실제 데이터를, 2021년부터 2050년까지는 추계 데이터를 적용한다. 인구이동의 경우 실제값과 추계에 활용한 값 간에 차이가 있는데, 이는 COVID-19에 따른 인구이동의 최근 급격한 변화 추세를 추계에서 반영하지 않았기 때문이다. 따라서 2000년부터 2020년까지는 실제 인구이동자 수를 활용하여 순이동률 데이터를 구축하였고, 2025년부터 2050년까지는 통계청 공표 장래인구추계에서 공표한 값을 활용하였다.

<그림 3-16>은 2020년부터 2050년까지 인구학적 변동요인의 시도별 추세이다. 세종특별자치시의 경우 2014년 출범이후 데이터가 축적되기 시작하였다. 따라서 이 연구에서는 전국의 평균값을 세종특별자치시에 위치한 격자들의 과거 인구학적 변동요인 수치로 적용하였다. 시도별 사망률 추세를 살펴보면 초고령화 사회에 진입함에 따라 2050년까지 모든 시도에서 꾸준히 증가하는 것으로 전망되었다. 시도별 출생률 추세는 2000년~2005년, 2015년~2020년 급격한 출생률 감소를 겪었고 장래인구추계에서는 2025년까지 출생률이 감소하는 추세를 유지하지만 이후에는 작은 폭에서 증가와 감소를 이어갈 것으로 전망하였다. 시도별 순이동률의 경우 세종특별자치시 출범에 따라 2015년도의 순이동률 비율이 해당 지역에 매우 높은 것으로 나타났으며, 세종특별자치시로의 인구유입이 점차 감소하지만 2030년까지는 유지되고 이후에는 다른 시도와 유사한 수준으로 수렴할 것으로 전망되었다. 순이동률은 사망률 및 출생률과 다르게 시도 간의 차이가 큰 것으로 나타났는데, 예를 들어 2020년 기준 경기도의 순이동률은 12.5%인데 반해 울산광역시의 경우 -11.9%로 나타났다. 통계청 공표 장래인구추계에서는 순이동률의 지역간 편차가 실제보다 더 작은 폭에서 차이가 날 것으로 공표하여(2025년 최대-경기도 7.2%, 최소-울산광역시 -7.1%), 향후 분석 결과 해석에 이를 염두에 둘 필요가 있다.

그림 3-16 | 시도별 인구학적 변동요인 추세(2000년 ~ 2050년)



자료: 저자 작성

격자단위 장래인구 수를 학습하고 예측할 때, 통계청 공표 시·도단위 장래인구추계와 적합성을 갖추기 위해 보정단계를 거쳤다. 2000년부터 2020년까지 인구총조사에 따른 실제 인구와 2020년 이후 2050년까지 장래인구추계 변화 추세는 <표 3-4>와 같다. 2000년도에 가장 인구가 많은 곳은 서울이었으나, 2050년에는 경기의 인구가 1,435만 명으로 가장 인구가 많은 광역지자체가 될 것으로 전망된다. 인천의 경우 2000년에 인구수로 7위였으나 2050년에는 수도권 인구 집중 추세에 따라 인구수 3위 광역지자체가 된다. 2000년 대비 2050년 인구변화를 전망해보면 경기, 인천, 충남, 충북, 제주 및 세종은 인구가 증가할 것으로 나타났다. 다만 인구증가 속도는 세종을 제외하고 모든 지자체에서 2020년 이후 감소할 것으로 나타났다. 이와 같은 시계열 변화 추세를 격자가 어떤 시·도에 위치해있는지에 따라 반영하여 장래인구 예측에 활용하였다.

표 3-4 | 시도별 인구추세(과거와 미래, 2020년 ~ 2050년)

시도	2000년 인구수 (만명)	2020년 인구수 (만명)	2050년 장래인구 추계(만명)	2000년 인구순위	2050년 인구순위	2000년 대비 2020년 인구 변화(%)	2020년 대비 2050년 인구 변화(%)	2000년 대비 2050년 인구 변화(%)
경기도	898	1,351	1,435	2	1	50.4	6.2	59.8
서울특별시	990	959	792	1	2	-3.1	-17.4	-20
인천광역시	248	295	281	7	3	19	-4.5	13.7
경상남도	298	333	277	4	4	11.9	-16.9	-7
부산광역시	366	335	251	3	5	-8.6	-25	-31.4
경상북도	272	264	226	5	6	-2.9	-14.7	-17.2
충청남도	185	218	219	10	7	18	0.7	18.7
대구광역시	248	241	181	6	8	-2.8	-25	-27.2
충청북도	147	163	160	12	9	11.3	-1.7	9.4
전라남도	200	179	152	8	10	-10.4	-14.9	-23.7
전라북도	189	180	149	9	11	-4.6	-17.2	-21
강원도	149	152	148	11	12	2.3	-2.6	-0.4
대전광역시	137	149	125	13	13	8.8	-16.2	-8.9
광주광역시	135	148	121	14	14	9.2	-17.9	-10.3
울산광역시	101	114	84	15	15	11.9	-25.6	-16.8
제주	51	67	70	16	16	30.7	4.1	36.1
세종	-	35	63	17	17	-	78.2	-

주: 세종은 2014년 출범에 따라 2000년 인구 등의 정보가 누락되었음
 자료: 저자 작성

(2) 공간적 특성

격자단위 장래인구 예측 관련 선행연구에서는 건축물 수, 도심/비도심 여부 등 토지 이용을 인구 배분의 주요 변수로 활용하였다. Chen et al. (2020a)는 XGboost, Random forest, Neural network 등 인공지능 기법을 활용하여 장래의 건축물과 이에 따른 인구분포를 예측하였으며, Ford et al. (2019)은 미래 개발계획 시나리오에 근거하여 공간간의 상호관계 인구 모델과 CA를 활용한 토지개발 모델을 결합하여 미래 도시화를 전망하였다. Lloyd·Sorichetta·Tatem(2017)은 인구분포에 영향을 미치는 요인 중 건축물 분포뿐 아니라 토지 피복 및 경사도와 같은 물리적 환경을 고려하였다.

이와 같은 선행연구를 바탕으로, 이 연구에서는 공간적 특성으로 토지이용과 물리적 환경을 고려하였다. 우선, 토지이용은 주거와 밀접하게 관련된 총 건축물 수, 공동주택 수, 단독주택 수, 주거용 건물의 연면적 합을 국토교통부 세움터 정보를 활용하여 2000년부터 2020년까지 데이터로 구축하였다. 물리적 환경은 휴토람에서 공표한 토지경사도를 활용하였다. 상세한 변수의 내용과 출처는 <표 3-5>와 같다.

표 3-5 | 공간적 특성의 데이터 출처

구분	내용	집계 단위	시점	원천데이터 출처	
토지 이용	건축물 수	격자 내 총 건축물 수	1km 격자 POI 건물데이터를 격자단위로 집계	2000년 ~ 2020년	국토교통부 세움터
	공동주택 수	격자 내 총 공동주택 수, 주용도코드가 공동주택으로 할당된 건물의 수를 집계			
	단독주택 수	격자 내 총 단독주택 수, 주용도코드가 단독주택으로 할당된 건물의 수를 집계			
	주거용건물 연면적	격자 내 주거용 건물의 연면적 합, 주용도코드가 공동주택/단독주택인 건물의 연면적 합			
물리적 환경	토양경사도	경사도 조사는 현장에서 경사계를 통해서 측정, 실제 거리 100m에서 꼭지점까지의 기울기 즉, 높이를 측정하여 결정	1km 격자 토양특성을 격자단위로 배분	90년대 측정* *휴토람 문의 결과	휴토람

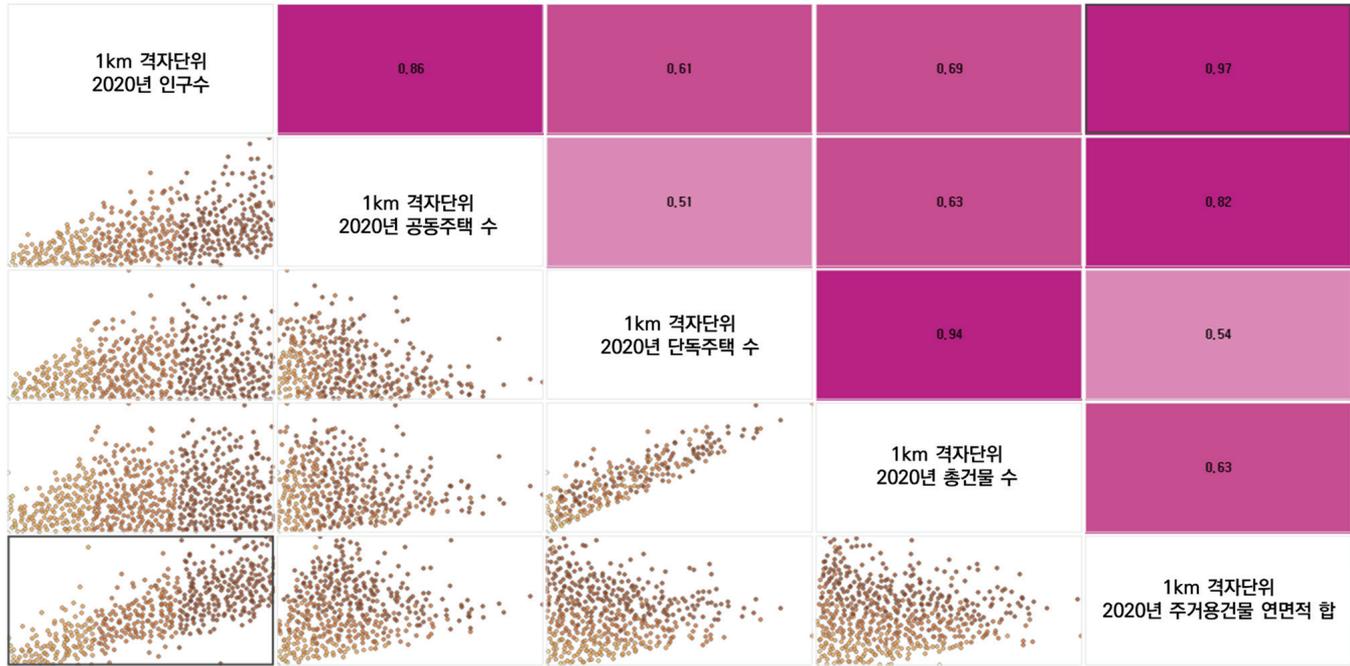
자료: 저자 작성

격자단위 공간적 특성을 나타내는 외생변수의 시계열 데이터를 구축하기 위해 국토교통부 건축행정시스템 세움터 자료를 활용하였다. 세움터는 지자체에서 건축물 승인을 받기 위해 등록한 모든 정보를 보유하고 있다. 세움터의 표제부 상 주소정보를 활용하여 전국의 건축물을 POIs(Point of Interests)로 변환하였다. 그리고 각 건축물의 주용도, 연면적, 사용승인연도 정보를 활용하여 1km 격자별 건축물 수, 공동주택 수, 단독주택 수, 주거용건물 연면적 데이터를 2000년부터 2020년까지 구축하였다. <그림 3-17>은 2020년 기준으로 1km 격자단위의 인구수, 공동주택수, 단독주택수, 총건물 수, 주거용건물 연면적 합의 상관관계를 도식화한 것이다. 인구수와 가장 상관관계가 높은 변수는 주거용건물 연면적 합(0.97)으로 나타났다¹⁶⁾. 격자 내 인구수는 공동주택수-총건물수-단독주택 수 순서로 상관관계가 높은 것으로 나타났다. 모든 변수에서 상관관계가 0.6 이상으로 유의하게 나타났으며, 외생변수로 활용하기에 적합한 것으로 판별된다.

2000년부터 2020년까지 격자단위 공간적 특성 데이터를 활용하여 모델1의 1단계에서 같은 시기에 시·도별 인구, 격자단위 인구분포, 인구학적 변동요인 및 공간적 특성 간의 관계를 학습하였다. 그리고 공간적 특성 데이터의 시계열 변화와 이웃효과를 고려하여 모델1의 2단계에서 CNN-LSTM 기법을 적용하여 2050년까지 공간적 특성인 외생 변수의 변화를 예측하였다. 1단계와 2단계의 학습 결과를 활용하여 모델1의 3단계에서는 2050년까지 인구학적 특성과 공간적 특성을 고려한 격자단위의 장래인구를 예측하였다.

16) 이때 주거용건물 연면적의 경우 세움터 데이터의 오류로 일부 격자의 연면적 합이 실제에 비해 너무 크게 기록되어 있어 분포를 고려하여 100만㎡ 이상의 값을 가지는 289개 격자는 아웃라이어 처리를 하였음.

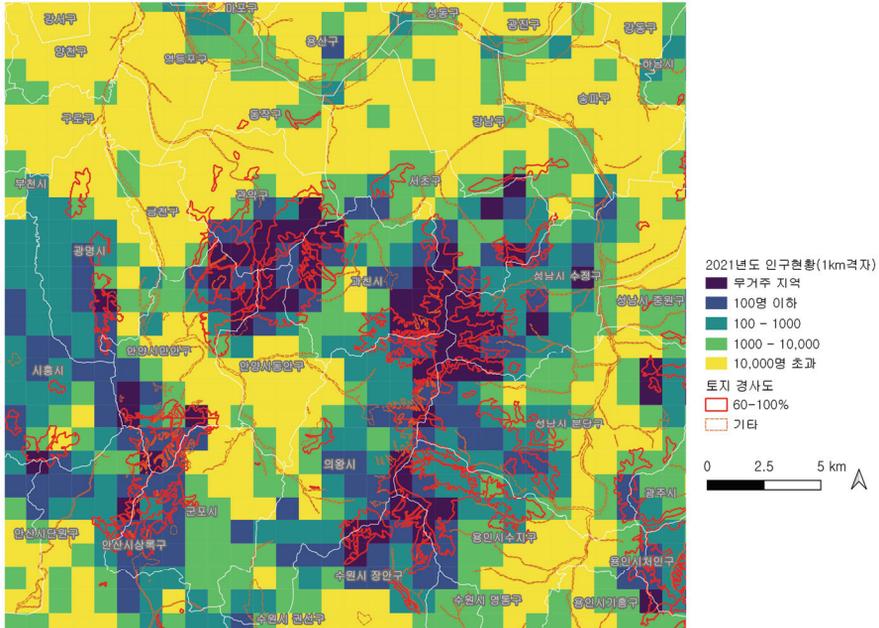
그림 3-17 | 토지이용(건물) 현황과 인구밀도 간의 상관관계



주: 주거용건물 연면적 합은 전체 분석대상 중 아웃라이어로 고려되는 연면적 총합 100만㎡ 초과 격자(289개)를 제외하여 상관계수를 계산
 자료: 저자 작성

토양경사도는 토양특성에 따라 매우 세분화되어 측정되어, 경사도 카테고리별로 격자에서 차지하는 비중을 계산할 수 있다는 장점이 있다. 경사도는 토양특성으로 구분된 구획별로 2%이하, 2~7%, 7~15%, 15~30%, 30~60%, 60% 초과 등 6개 카테고리별로 구분되어 있으며, 이를 1km 격자와 공간결합하여 각 토양특성의 경사도를 해당 토양특성이 격자에서 차지하는 비중만큼 부여하는 전처리 작업을 수행하였다. 1km 격자 내 각 경사도의 카테고리별 면적 비중을 산출하였으며, 6개 변수로 활용할 수 있다. 다만, 토양경사도 시계열 변화 데이터를 얻을 수 없어, 경사도가 크게 변하지 않는다는 가정으로 90년대 정보를 활용한다. 이와 같은 정보의 활용 적합성을 검토하기 위해 최근 격자단위 인구와 경사도를 융합하여 살펴본 결과(〈그림 3-18〉 참조), 경사도가 60% 초과인 지역의 경우 대부분 아무도 거주하고 있지 않은 무거주지역으로 분류되어 변수로 활용하는데 의의가 있음을 확인하였다.

그림 3-18 | 경사도에 따른 인구분포 예시



자료: 저자 작성



CHAPTER 4

GPAI 학습 및 장래인구 예측 결과

1. GPAI 인공지능 기법(모델1) 학습 결과 83
2. GPAI 적용 우리나라 장래인구 예측 89
3. GPAI 고도화 방안 106
4. 소결: 방법론의 독창성 및 학술적 기여 113

04 GPAI 학습 및 장래인구 예측 결과

본 장에서는 격자단위 장래인구 예측 AI 모델(Gridded Population Forecast with AI, GPAI)에서 모델의 인공지능 기법 학습 결과를 살펴보았다. 그리고 학습된 GPAI를 전국에 적용하여 2050년까지 우리나라 인구분포 예측 결과를 제시하였다. 전국 인구분포와 함께 7개 광역권과 인구 감소 지정지역 등 정책 관심지역에 대해 인구변화 추세를 소지역인 1km 격자단위로 분석하고 인구가 증가하거나 또는 무거주화 위험에 놓인 지역을 탐색하였다. 예측 결과 검토를 통해 GPAI 개선에 필요한 사항들을 검토하였고, 이로부터 고도화방안을 도출하였다.

1. GPAI 인공지능 기법(모델1) 학습 결과

격자단위 장래인구 예측 AI 모델(Gridded Population Forecast with AI, GPAI)는 인구학적 특성과 공간적 특성을 고려한 인공지능 기법 기반 모델1과 인구학적 특성에 따른 변화에 초점을 맞춘 코호트요인법 기반 모델2를 결합한 방법론이다. 인공지능 기법을 활용한 모델1의 경우 과거의 시계열 데이터를 활용하여 알고리즘을 학습하는 과정이 중요하며, 각 과정의 학습 결과에 따라 예측값의 신뢰도가 결정된다. 모델1에서 인공지능 기법은 1단계와 2단계에 적용되었다. 1단계에서는 인구학적특성-공간적 특성-격자단위 인구분포 간의 관계를 도출하기 위해 랜덤포레스트(random forest) 알고리즘을 학습하였다. 2단계에서는 공간적 특성을 외생변수로서 2020년 이후 각 변수값을 예측하는 CNN-LSTM(Convolutional Neural Network-Long-Short Term Memory) 알고리즘을 학습하였다.

1단계 랜덤포레스트 학습 과정에서 주요 하이퍼파라미터(hyper-parameter)인 가지 생성분기점 값과 나무집계 최적의 수를 결정하기 위해 점진적 최적화 방법인 언덕등반 기법(hill climbing) 알고리즘을 적용하였다. 그리고 모험학습 과정에서 검증 비율은 머신러닝 방법론에서 널리 사용되고 있는 황금비율(1.618)을 이용하여 그 비율을 중복

샘플링 경우를 포함하여 랜덤하게 산정하여 실행하였다. 모형학습에 사용된 입력데이터의 구축은 시간과 공간 간의 이웃효과를 포함하기 위하여 격자별로 인접 격자를 모두 포함하여 구축하였고, 최종 총 300가지의 입력변수를 구성하였다¹⁾.

모형의 정확도는 머신러닝 분야에서 통상적으로 활용되는 효과적도 중 하나인 평균 제곱근오차(Mean Squared Error, MSE)로 검증하였다. 1단계의 랜덤포레스트 학습 결과, 최적의 가지 생성분기점은 293, 나무집계 최적의 수는 946으로, MSE 0.000008232771에서 최적 수렴하였다. 학습에 활용된 다양한 변수 중 모형 MSE기준 가장 높은 민감도를 나타내는 변수는 해당연도 목표 격자의 주거용 건물 연면적이 약 19.1%로 가장 큰 영향을 미치는 것으로 나타났다. 그리고 목표 격자의 직전년도 인구 변수가 5.7% 비중으로 두 번째로 높게 확인되었고, 목표 격자의 직전 연도의 주거용 건물 연면적 등이 뒤를 이었다(〈표 4-1〉 참조). 가중치가 큰 변수 대부분 목표 격자 자신의 값의 영향을 가장 많이 받았으며, 이웃 격자의 영향을 받는 특성은 공동주택 수로 나타났다. 즉, 목표 격자와 인접한 이웃 격자들의 공동주택 밀도가 목표 격자의 장래인구 분포에 영향을 미친다고 해석할 수 있다.

표 4-1 | 1단계의 변수 중요도 및 변수 내용(상위 15개)

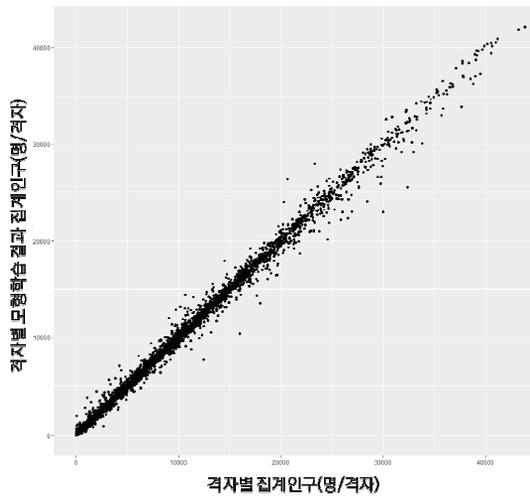
순위	변수명	중요도	변수 내용	구분
1	ARA_gTg_2020	19.1%	목표 격자의 해당연도(t) 주거용건물 연면적	외생변수
2	HisPop_gTg_2015	5.7%	목표 격자의 직전 연도(t-1) 인구 정보	격자인구
3	ARA_gTg_2015	3.0%	목표 격자의 직전 연도(t-1) 주거용건물 연면적	외생변수
4	HisPop_gTg_2010	2.9%	목표 격자의 전 연도(t-2) 인구 정보	격자인구
5	ARA_gTg_2010	2.6%	목표 격자의 전 연도(t-2) 주거용건물 연면적	외생변수
6	ARA_gTg_2005	2.2%	목표 격자의 전 연도(t-3) 주거용건물 연면적	
7	APT_gTg_2020	2.1%	목표 격자의 해당연도(t) 공동주택 수	
8	APT_gCT_2000	2.0%	목표 격자와 인접한 격자의 전 연도(t-4) 공동주택 수	외생변수
9	APT_gTg_2010	1.6%	목표 격자의 전 연도(t-2) 공동주택 수	
10	HisPop_gTg_2005	1.5%	목표 격자의 전 연도(t-3) 인구 정보	격자인구
11	Grd_gLM_ud60	1.5%	목표 격자와 인접한 격자의 경사도 30% ~ 60% 비율	외생변수
12	HisPop_gTg_2000	1.4%	목표 격자의 전 연도(t-4) 인구 정보	격자인구
13	APT_gTg_2005	1.3%	목표 격자의 전 연도(t-3) 공동주택 수	외생변수
14	BUD_gTg_2005	1.3%	목표 격자의 전 연도(t-3) 건축물 수	
15	APT_gCT_2005	1.2%	목표 격자와 인접한 격자의 전 연도(t-3) 공동주택 수	

자료: 저자 작성

1) 입력변수의 변수명과 내용은 〈부록 2〉를 참조

2020년 격자별 인구 참값과 1단계 랜덤포레스트 알고리즘 학습을 통해 예측된 2020년 격자별 인구를 비교한 결과는 <그림 4-1>과 같다. 산포도의 X축은 2020년 기준 격자별 집계인구 참 값을, 산포도의 Y축은 학습 결과로 예측된 집계인구(명/격자)를 의미한다. 산포도에서 확인할 수 있듯이, 모형의 결과는 대각선에서 수렴하는 형태 ($R^2=0.998$)로 모형의 학습이 적절히 이루어진 것으로 나타났다.

그림 4-1 | 2020년 격자별 실제 인구와 랜덤포레스트 예측 결과 비교



자료: 저자 작성

2단계 CNN-LSTM 알고리즘 학습 과정에서는 모형의 오차 가능성을 최소화하기 위해 외생변수(공동주택수, 단독주택수, 주택연면적, 총건물수 등)에 대한 개별 모형을 적용하였다. 시계열 데이터 학습을 위해 학습 단위(2개년도)별로 합성곱(Convolution)을 적용하여 데이터 특성을 최대화한 뒤, LSTM모형으로 짝을 지어 대략적인 반복 횟수는 100회에, 목표 오차율은 0.0003로 일괄 적용한 뒤 학습하는 방식으로 수행하였다. 4개의 외생변수에 대해서 107,442개 격자 단위당 개별 CNN-LSTM 모형을 총 429,768개 구축하여 장래 연도 외생변수 예측에 활용하였다.

부족한 학습데이터로 발생할 수 있는 예측 격자의 오차 위험을 최소화하기 위해, 외생변수별로 추가 정산 단계를 거치도록 하여 비교적 안정적인 예측을 수행할 수 있도록 유도하였다. 추가 정산 단계는 총 두 단계로 구성하였는데, 정산의 기준은 가장 최근의 참값이라 할 수 있는 2020년 격자별 인구와 각 외생변수 값을 기준으로 각 격자의 경험적인 정보를 활용하여 진행하였다. 먼저 첫 번째 정산에서는 예측된 격자별 정보를 바탕으로 <표 4-2>에서 기술된 조건에 따라 격자별 예측값을 치환한다. 이후 전국 격자의 2020년 각 외생변수 값과 예측값 간의 회귀관계를 도출한 뒤, 절편 및 기울기를 적용하여 두 번째 정산을 수행하였다.

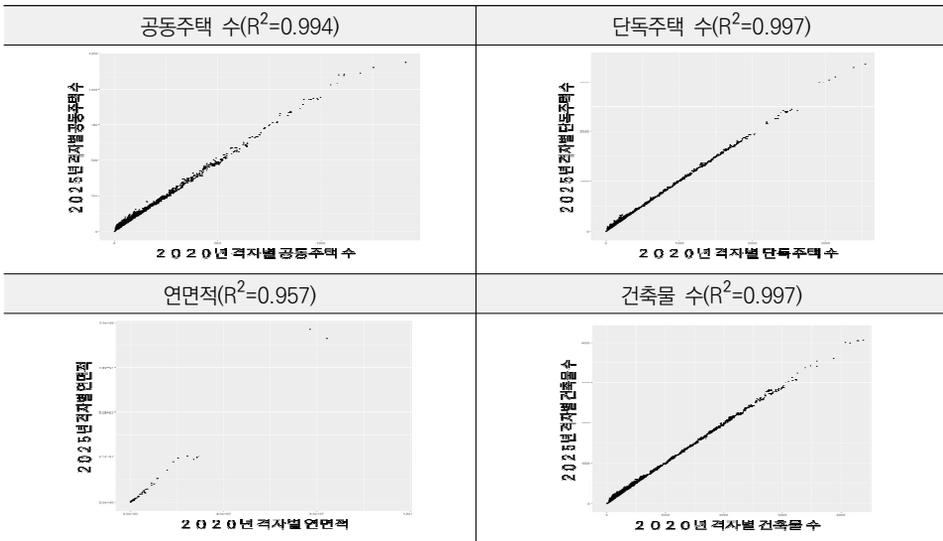
표 4-2 | 외생변수별 정산기준

외생변수 구분	정산기준
건축물 수	<ul style="list-style-type: none"> • 2020년 기준 격자 내 건물 수가 40개 이하이거나 2020년 기준 격자 인구가 10명 이하면, 예측 연도에도 2020년도 건축물 수가 변하지 않는다고 간주하고 해당 격자의 2020년 건축물 수를 활용 • 2020년 기준 격자 내 건물 수가 40개 이상이고 2020년 기준 격자 인구가 10명 초과인 경우, 해당 격자의 예측된 값을 활용
공동주택 수	<ul style="list-style-type: none"> • 2020년 기준 격자 내 공동주택 수가 4개 이하이거나 2020년 기준 격자 인구가 10명 이하면, 예측 연도에도 2020년도 공동주택 수가 변하지 않는다고 간주하고 해당 격자의 2020년 공동주택 수를 활용 • 2020년 기준 격자 내 공동주택 수가 4개 이상이고 2020년 기준 격자 인구가 10명 초과인 경우, 해당 격자의 예측된 값을 활용
단독주택 수	<ul style="list-style-type: none"> • 2020년 기준 격자 내 단독주택 수가 20개 이하이거나 2020년 기준 격자 인구가 10명 이하면, 예측 연도에도 2020년도 단독주택 수가 변하지 않는다고 간주하고 해당 격자의 2020년 단독주택 수를 활용 • 2020년 기준 격자 내 단독주택 수가 20개 이상이고 2020년 기준 격자 인구가 10명 초과인 경우, 해당 격자의 예측된 값을 활용
주거용건물 연면적	<ul style="list-style-type: none"> • 2020년 기준 격자 내 연면적이 2000 이하이거나 2020년 기준 격자 인구가 10명 이하면, 예측 연도에도 2020년도 연면적이 변하지 않는다고 간주하고 해당 격자의 2020년 연면적이 값을 활용 • 2020년 기준 격자 내 연면적이 2000 이상이고 2020년 기준 격자 인구가 10명 초과인 경우, 해당 격자의 예측된 값을 활용

자료: 저자 작성

CNN-LSTM 알고리즘 학습에 따른 4개 외생변수(공동주택 수, 단독주택 수, 주거용 연면적, 건축물 수)의 2025년 예측값과 2020년 참값을 비교한 결과, 모두 R^2 값 기준 높은 상관도를 보인 것을 확인할 수 있다(그림 4-2 참조). 또한 외생변수의 2000년부터의 추세와 2단계 결과로 예측된 2025년부터 2050년까지 예측값의 추세는 <그림 4-3>과 같다. 일관된 비교를 위해 모든 그래프의 X축은 모두 2020년으로 통일하여 도식화하였는데, 참값인 2000년부터 상대적으로 분산된 추이가 2015년까지 2020년에 근접할수록 대각선으로 수렴하는 추이가 확인되었다. 한편, 예측값인 2025년부터 2050년으로 진행할수록 상대적으로 높은 수치가 포함된 격자의 수가 2020년 대비 감소하는 추이²⁾가 확인되었고, 장래 연도에 근접할수록 Y축에 가까워지는 감소 추이가 확인되었다. 이러한 추이는 모형의 예측과 정산의 반복과정에서 인구밀집도가 매우 높거나 낮은 격자의 2020년 인구정보가 CNN-LSTM 예측에서 예측된 값의 상한값을 조절하는 등 정산의 단계에서 해당 현상의 주된 원인으로 작용했을 것으로 해석할 수 있다.

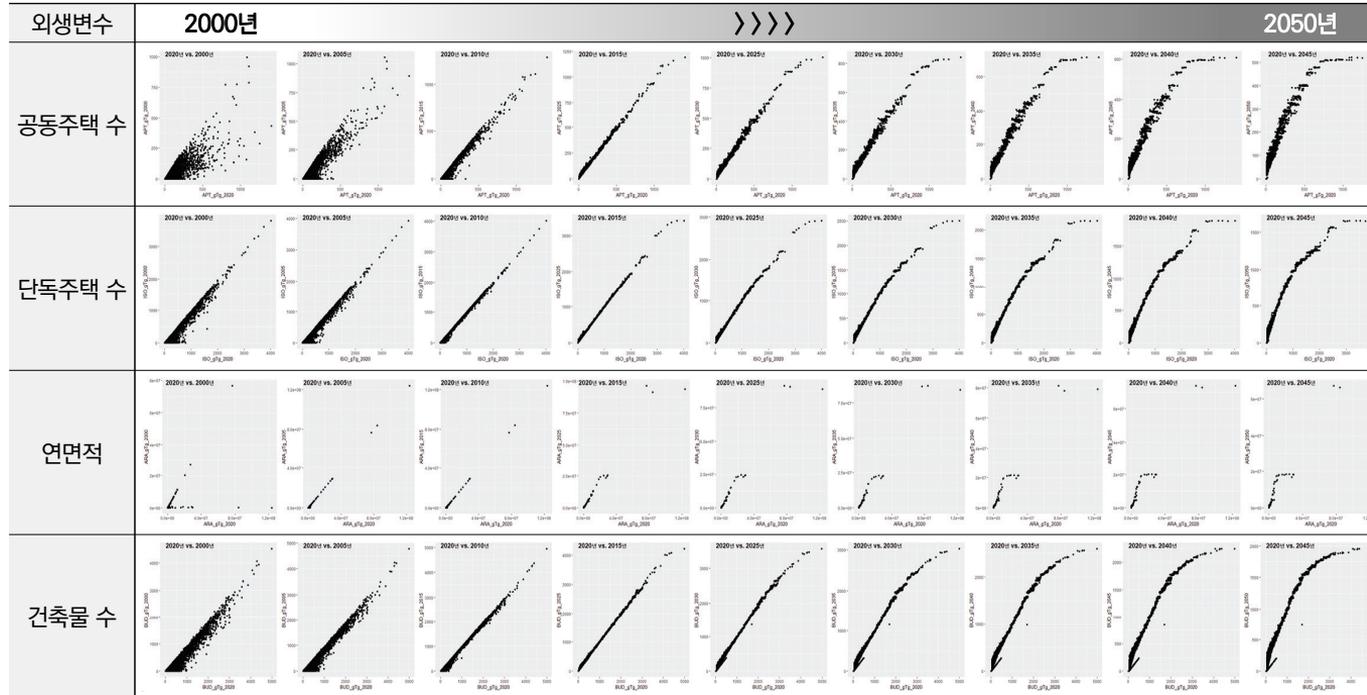
그림 4-2 | 주요 외생변수(공동주택 수, 단독주택 수, 주거용 연면적, 건축물 수)의 2020년 참값과 2025년 예측값 관계



자료: 저자 작성

2) 그래프에서 대각선이 오른쪽으로 휘는 추세

그림 4-3 | 주요 외생변수의 침강(2000~2020)과 모델1 2단계 예측값의 관계



주: 모든 그래프의 X축은 2020년임. 2020년을 기준으로 2000년부터 2015년까지의 격자별 외생변수와 2050년까지 예측된 격자별 외생변수 비교 (왼쪽에서 오른쪽으로 갈수록 2050년에 가까움)

자료: 저자 작성

2. GPAI 적용 우리나라 장래인구 예측

1) 전국 및 광역권 장래인구 변화 전망

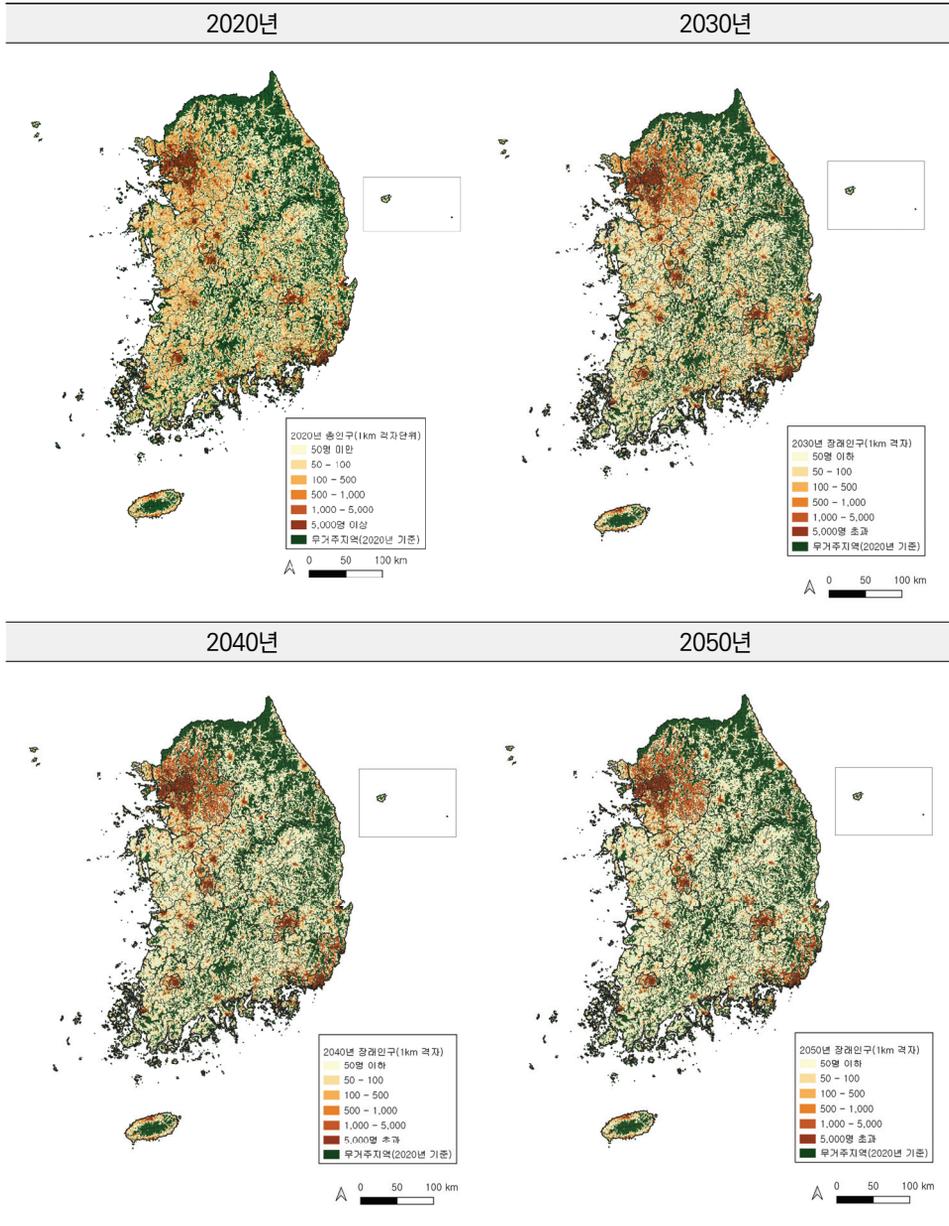
GPAI를 적용하여 2025년부터 2050년까지 우리나라 격자단위 장래인구를 5년 단위로 예측하였다³⁾. 2020년 대비 2030년, 2030년 대비 2040년, 2040년 대비 2050년 등 10년 단위의 인구변화는 격자별로 상관관계가 매우 높지만⁴⁾, 시간이 지날수록 변동 폭이 커지는 것을 확인하였다. 2020년의 인구분포부터 2050년까지 10년 단위 전국 인구분포 전망은 <그림 4-4>와 같다.

2020년도 인구분포와 2050년 장래인구 분포를 비교하였을 때 일부 수도권과 광역지자체를 제외하고 인구밀도가 전체적으로 감소할 것으로 나타났다. 인구감소시대를 맞이함에 따라 2050년도에는 인구밀도는 매우 집중된 지역 일부와 대부분의 인구희박지역으로 양극화될 것으로 전망된다. 2020년 기준으로 30년 후 2050년 격자단위 인구증감율은 <그림 4-5>와 같다. 2050년에는 현재(2020년)에 비해 수도권, 충남 북부-세종-대전 및 일부 광역시 외곽지역 공간에서는 인구가 증가하지만, 대부분의 국토에서 인구감소 현상을 겪을 것으로 전망된다.

3) 5년 단위 장래인구 분포는 <부록 4> 참조

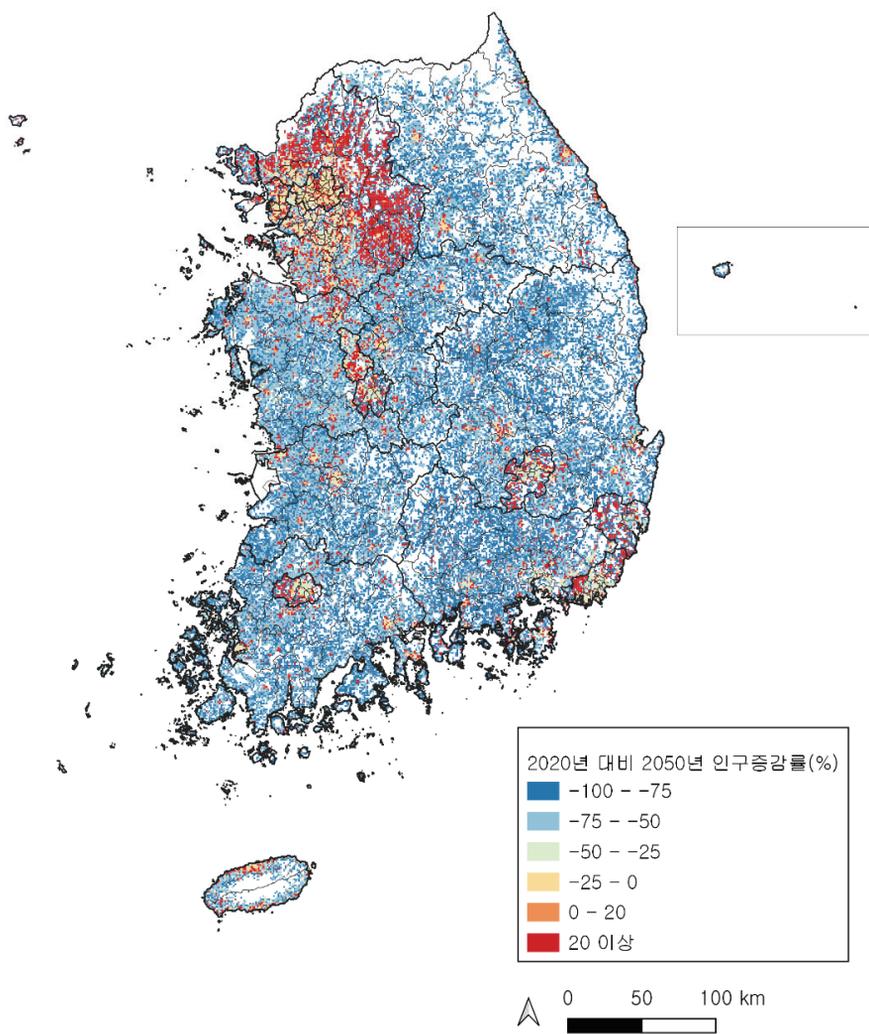
4) 2020년, 2030년, 2040년 및 2050년 인구 예측의 상관관계는 <부록 3> 참조

그림 4-4 | GPAI를 활용한 전국 1km 격자단위 장래인구 예측 결과(2020년~2050년)



자료: 저자 작성

그림 4-5 | GPAI를 활용한 전국 1km 격자단위 장래인구 변화 전망(2020년 대비 2050년 인구증감률)



자료: 저자 작성

2050년까지 예측한 장래인구를 활용하여 인구밀도 및 인구감소와 무거주화 위험지역의 비중 변화를 <표 4-3>과 같이 살펴보았다. 전 국토를 기준으로 평균 인구밀도는 2050년까지 감소할 것으로 전망되었다. 예를 들어, 2020년 인구밀도를 100으로 보았을 때 2050년의 인구밀도는 91수준일 것으로 나타났다. 전국의 평균 수치인 인구밀도로 인구감소 패턴을 파악하기 어려우므로 2020년에는 인구가 거주하였으나 장래에 아무도 살지 않을 것으로 예상된 ‘무거주화 위험 지역’, 2020년 대비 인구가 감소할 것으로 예상된 ‘인구감소 지역’ 그리고 인구가 증가할 것으로 예상된 ‘인구증가 지역’의 비중을 살펴보았다. 그 결과 2050년 무거주화 및 인구감소지역 비중이 전체 인구 거주지역의 87.2%를 차지할 것으로 나타났다.

표 4-3 | 전국 2020년 ~ 2050년 인구밀도와 인구증가, 인구감소 및 무거주화 위험 지역 비중 변화

구분	2020년	2025년	2030년	2035년	2040년	2045년	2050년
평균 인구밀도 (전 국토 기준)	483명*/km ²	479명/km ²	477명/km ²	473명/km ²	467명/km ²	456명/km ²	441명/km ²
평균 인구밀도 (인구거주지역 기준)	724명/km ²	696명/km ²	694명/km ²	695명/km ²	691명/km ²	683명/km ²	672명/km ²
무거주화 위험 지역 비중 (2020년 기준)	-	1.6%	1.7%	1.9%	2.3%	2.9%	3.9%
인구감소지역 비중 (2020년 기준)	-	58.2%	74.9%	79.8%	81.4%	83.2%	83.3%
인구증가지역 비중 (2020년 기준)	-	40.2%	23.4%	18.3%	16.3%	13.9%	12.8%

주: *2020년 전 국토 기준 인구밀도의 경우 SGIS 격자통계는 개인정보보호를 위해 5인이하를 0처리한 경우가 있어, 해석의 왜곡을 방지하기 위해 비마스킹된 국토지리정보원 2020년 10월 기준 1km 격자단위 인구통계를 활용
 자료: 저자 작성

전국을 7개 광역권(수도권, 강원권, 대전·세종·충청권, 대구·경북권, 부산·울산·경남권, 광주·전남·전북권, 제주도)로 구분하여 2050년까지 인구변화를 살펴보았다(〈표 4-4〉 참조). 수도권과 비수도권의 인구비중은 2020년 약 50.2%에서 2050년 약 53.0%로 그 차이가 커질 것으로 나타났다. 그리고 2020년 인구를 100%로 보았을 때 2050년 수도권 인구 규모는 96.5%, 비수도권의 경우 86.5%로 모두 감소하지만 비수도권의 감소 폭이 더 클 것으로 나타났다. 구체적으로 2020년 대비 2050년 인구 증감 패턴에 따라 격자를 무거주화 위험, 인구감소, 인구증가 지역으로 유형화하여 살펴보면(〈표 4-4〉 참조), 수도권의 경우 2020년 대비 2050년 인구가 증가하는 공간이 약 40.5%이지만 비수도권의 경우 8.3%에 불과하다.

비수도권 내 광역권에 따라 인구변화 추세가 다르게 나타났는데, 대전·세종·충청권과 제주권의 경우에는 인구가 증가하지만 부산·울산·경남권의 경우 2050년 인구는 2020년 인구의 약 77.9%의 수준일 것으로 나타났다. 대전·세종·충청권과 부산·울산·경남권의 2020년 대비 2050년 인구분포의 변화는 〈그림 4-6〉 및 〈그림 4-7〉와 같다. 전체 인구규모 관점에서는 부산·울산·경남권 인구는 크게 감소하고 대전·세종·충청권은 증가할 것으로 전망되었으나, 격자단위의 고해상도 인구분포 변화에 따르면 두 지역에서 공통적으로 인구밀도의 양극화가 심화될 것으로 나타났다.

비수도권 내 광역권별 인구증감 패턴에 따른 유형의 비중을 살펴보면, 부산·울산·경남권의 경우 인구증가지역의 비중이 13.1%로 인구가 증가할 것으로 나타난 대전·세종·충청권의 인구증가지역 비중(9.6%)에 비해 높는데 이는 행정구역단위의 인구추세가 우리가 실제 생활하고 있는 공간범위인 소지역단위에서는 완전히 다른 패턴을 보일 수 있다는 것을 반증한다. 따라서 행정구역단위의 인구규모 변화뿐 아니라 고해상도의 격자단위 장래인구추계를 활용하여 인구밀도가 유지 또는 증가하는 공간과 인구가 크게 감소하거나 무거주화될 공간을 파악하여 장소기반의 공간계획을 수립해야 한다.

표 4-4 | 광역권 2020년 ~ 2050년 인구밀도와 인구증가, 인구감소 및 무거주화 위험 지역 비중 변화

구분	수도권	비수도권							
		강원권	대전·세종·충청권	대구·경북권	부산·울산·경남권	광주·전남·전북권	제주		
2020년 총 인구수	25,982천명	25,752천명	1,533천명	5,526천명	5,053천명	7,866천명	5,099천명	672천명	
2050년 장래 인구 수	25,085천명	22,272천명	1,481천명	5,673천명	4,063천명	6,125천명	4,229천명	698천명	
2020년 대비 2050년 인구 비중	96.5%	86.5%	96.6%	102.7%	80.4%	77.9%	82.9%	103.9%	
평균 인구밀도 (인구거주 지역 기준)	2020년*	2,566명/km ²	427명/km ²	199명/km ²	412명/km ²	399명/km ²	853명/km ²	323명/km ²	555명/km ²
	2050년	2,590명/km ²	371명/km ²	193명/km ²	424명/km ²	321명/km ²	664명/km ²	268명/km ²	577명/km ²
2050년 무거주화 위험지역 비중	1.9%	4.2%	6.6%	3.0%	4.8%	4.3%	3.7%	3.2%	
2050년 인구감소지역 비중	57.5%	87.5%	88.1%	87.4%	88.4%	82.7%	90%	80.2%	
2050년 인구증가지역 비중	40.5%	8.3%	5.3%	9.6%	6.8%	13.1%	6.3%	16.6%	

주1: *2020년 전 국토 기준 인구밀도의 경우 SGIS 격자통계는 개인정보보호를 위해 5인 이하를 0처리한 경우가 있어, 해석의 왜곡을 방지하기 위해 비마스킹된 국토지리정보원 2020년 10월 기준 1km 격자단위 인구통계를 활용

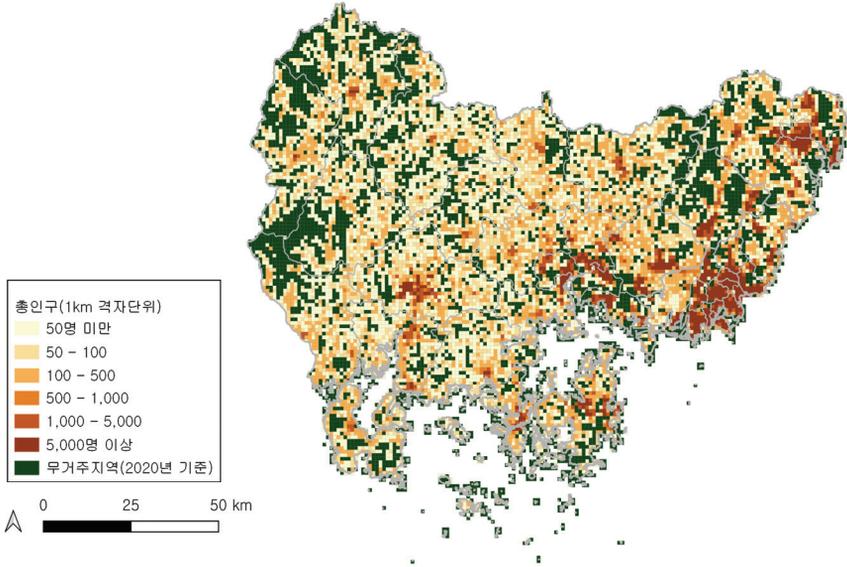
주2: 격자의 위치(중심점)를 기준으로 시도단위 인구를 합산한 것으로, 통계청 공표 총조사 인구 및 장래인구추계와 근소한 차이가 발생할 수 있음

주3: 무거주화 위험 지역 비중, 인구감소지역 비중, 인구증가지역 비중은 2020년도 기준 2050년 인구 변화에 따라 격자(소 지역)의 유형을 구분하고, 2020년 인구 거주 격자 중 해당하는 유형의 격자 수의 비중을 계산한 것임

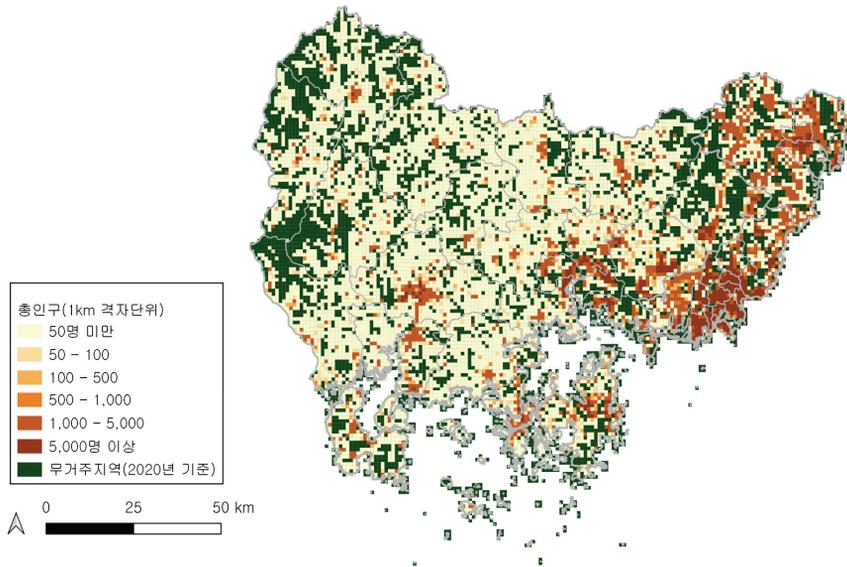
자료: 저자 작성

그림 4-6 | 2020년도와 2050년도 부산광역시, 울산광역시 및 경상남도 일대 인구분포 변화

2020년



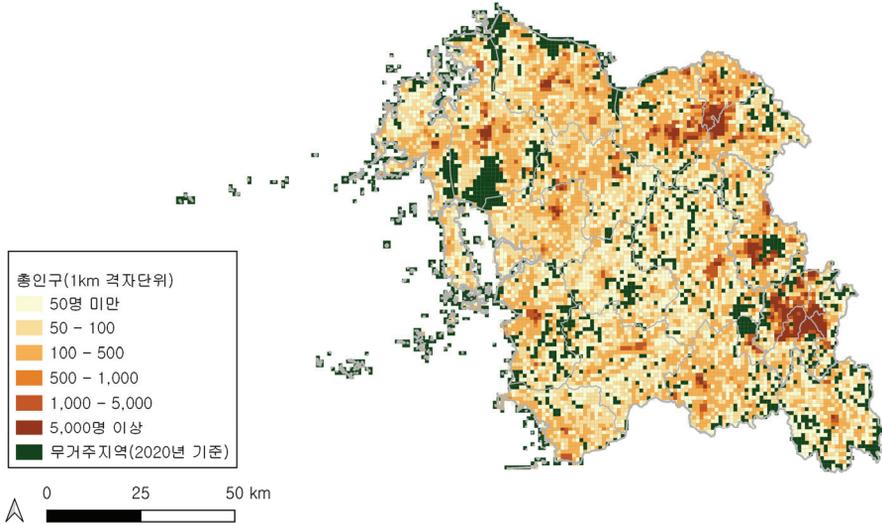
2050년



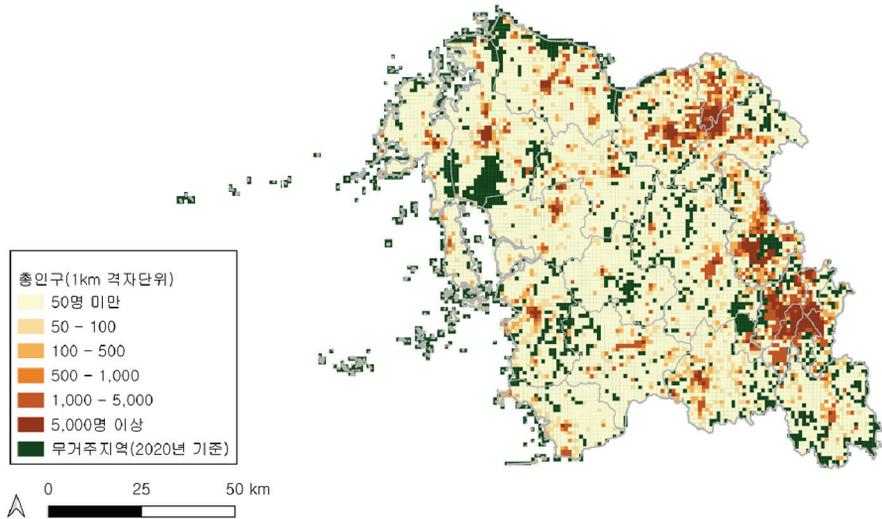
자료: 저자 작성

그림 4-7 | 2020년도와 2050년도 충청남도 및 세종특별자치시 일대 인구분포 변화

2020년



2050년



자료: 저자 작성

수도권과 비수도권의 장래인구 전망에 따르면(〈표 4-4〉 참조), 비수도권의 인구감소지역 비중이 수도권에 비해 약 30%p 가량 클 것으로 나타났다. 이와 같이 소멸의 위험이 높은 비수도권 지역 중 2021년에 행정안전부에서 지정하여 큰 정책적 관심을 받는 인구감소 지정지역과 인구감소 비(非)지정지역의 장래인구 변화 패턴을 비교하였다. 1km 격자단위 장래인구추계를 비수도권에 위치한 인구감소 지정지역(85개 시·군·구)과 인구감소 비지정지역(88개 시·군·구)으로 구분하여 2050년까지 인구분포의 변화를 살펴본 결과는 〈그림 4-8〉 및 〈표 4-5〉와 같다.

비수도권 인구감소 지정지역의 지난 20년간 인구변화를 살펴보면, 2020년 인구 규모는 2000년에 비해 약 81%로 인구감소 비지정지역이 같은 시기 103%로 인구가 증가한 것에 비해 크게 감소한 것으로 나타났다(〈그림 4-8〉 참조). 그런데 2000년 기준으로 50년 후 인구변화를 살펴보면, 2000년에 비해 2050년 인구감소 지정지역의 인구는 약 74% 수준이고 인구감소 비지정지역의 인구는 약 88% 수준으로 감소할 것으로 전망되었다. 즉, 지금까지 인구변화로 인구감소 비지정지역이더라도 앞으로 30년간은 인구감소를 경험할 것이며, 그 감소 폭은 도리어 인구감소 지정지역보다 더 클 것으로 나타났다(〈표 4-3〉 2020년 대비 2050년 인구 비중에 따르면 인구감소 지정지역은 91%, 인구감소 비지정지역은 85%).

그림 4-8 | 비수도권 인구감소 지정지역과 인구감소 비(非)지정지역의 인구변화 추세



자료: 2000년, 2020년 총 인구는 통계청 인구총조사의 시·군·구별 인구통계를 활용하여 저자 작성

2020년의 비수도권 인구감소 지정지역의 평균 인구밀도는 129명/km²으로 인구감소 비지정지역 평균 인구밀도 924명/km²에 비해 매우 낮은 것으로 나타났다. 이와 같은 차이는 2050년에도 유지되는데, 평균 인구밀도가 각각 114명/km²과 795명/km²이 될 것으로 예측되었다. 다만, 인구감소 지정지역에 비해 비지정지역의 인구밀도가 더 큰 폭으로 감소할 것으로 나타났다. 2020년 대비 2050년 인구증감 패턴에 따라 격자를 무거주화 위험, 인구감소, 인구증가 지역으로 유형화하고 인구감소 지정지역과 인구감소 비지정지역에서 각 유형이 차지하는 비중을 비교하였다. 인구감소 지정지역의 경우 95.1%, 인구감소 비지정지역의 경우 85.9% 지역에서 인구가 감소하거나 또는 더 이상 인구가 거주하지 않는 무거주화지역이 될 것으로 전망되었다. 지난 20년간(2000년~2020년) 비수도권의 인구감소 지정지역이 인구감소 비지정지역에 비해 급격한 인구감소를 경험하였으나 앞으로 30년간은 이와 같은 구분 없이 모든 지역에서 인구밀도가 크게 감소하고, 무거주화 및 인구감소지역에 대부분의 공간이 노출될 것으로 나타났다.

표 4-5 | 비수도권 인구감소 지정지역(행안부)과 인구감소 비(非)지정지역 2050년 인구 현황 비교

구분		인구감소 지정지역 (비수도권 85개 시군구)	인구감소 비(非)지정지역 (비수도권 88개 시군구)
2000년 총 인구수*		5,770천 명	20,521천 명
2020년 총 인구수*		4,683천 명	21,101천 명
2050년 장래 인구 수		4,257천 명	18,008천 명
2000년 대비 2020년 인구 비중(%)		81.2%	102.8%
2000년 대비 2050년 인구 비중(%)		73.9%	87.8%
2020년 대비 2050년 인구 비중(%)		91.0%	85.3%
평균 인구밀도 (인구거주지역 기준)	2020년**	129명/km ²	924명/km ²
	2050년	114명/km ²	795명/km ²
2020년 인구거주 격자 수		37,285개	22,660개
2050년 무거주화 위험 지역 비중		4.7%	3.4%
2050년 인구감소지역 비중		90.4%	82.5%
2050년 인구증가지역 비중		4.7%	14.1%

주1: *2000년, 2020년 총 인구수는 통계청 인구총조사의 시·군·구별 인구 통계를 활용

주2: **2020년 전 국토 기준 인구밀도의 경우 SGIS 격자통계는 개인정보보호를 위해 5인이하를 0처리한 경우가 있어, 해석의 왜곡을 방지하기 위해 비마스킹된 국토지리정보원 2020년 10월 기준 1km 격자단위 인구통계를 활용

주3: 무거주화 위험 지역 비중, 인구감소지역 비중, 인구증가지역 비중은 비수도권 인구감소지역 85개 시군구 전체 또는 비수도권 비인구감소지역 88개 시군구 전체 인구 거주 격자 중 각 유형에 해당하는(2020년도 기준 2050년 인구 변화) 격자 수 비중을 계산한 것임

자료: 저자 작성

2) 장래인구 변화에 따른 공간 유형과 특성

인구감소시대에 적절한 국토공간계획 수립을 위해 인구가 증가하거나 감소하는 패턴에 따라 구분된 공간 유형별로 어떤 특성이 있는지 파악하는 것이 매우 중요하다. 따라서 격자가 위치한 지자체의 특성(수도권, 광역시, 도내 시 지자체와 도내 군 지자체)과 인구증감 패턴(인구감소, 무거주화 위험, 인구증가)을 종합하여, 각 격자의 유형을 구분하였다. 그리고 각 유형별로 현재(2020년 기준) 정주여건을 생활인프라 접근성으로 비교한 결과 <표 4-6>과 같이 나타났다. 정주여건을 파악하는 다양한 변수 중 생활에 필수적인 공공시설인 초등학교와 민간시설인 약국까지의 이용 편의성⁵⁾이 각 유형별로 어떻게 차이가 나는지 살펴보았다.

지역유형별 무거주화 위험공간, 인구감소 공간, 인구증가 공간의 비중을 살펴보면, 도내 군 지자체의 무거주화 위험 공간이 4.7%로 수도권 1.9%에 비해 매우 높은 것으로 나타났다. 그리고 인구감소 공간의 비중의 경우 수도권은 전체 수도권 인구거주 격자 중 약 57.5%에서 인구가 감소할 것으로 예상되나, 광역시는 60.5%, 도내 시 지자체는 87.1%, 그리고 도내 군 지자체는 90.6%로 나타났다. 이와 같은 맥락에서 인구가 증가할 것으로 예상되는 공간의 비중의 경우 수도권과 도내 군 지자체의 양극화는 매우 크게 나타나는데, 수도권의 경우 약 40.5% 지역에서 인구가 증가하지만 도내 군 지자체의 경우 단 4.7% 지역에서만 인구가 증가할 것으로 나타났다. 이는 앞서 <표 4-2>를 통해 광역권별 인구변화를 검토한 결과와 일치한다.

5) 국토지리정보원에서 공표한 국토지표의 격자단위 생활인프라 접근성(500m 기준)을 활용하여 1km 격자에 속하는 500m 격자의 초등학교와 약국까지의 접근거리 평균을 이용 편의성으로 정의

표 4-6 | 인구감소 유형별 정주여건

유형	유형별 비중	2050년 평균 인구밀도	평균 초등학교 접근성 (2020년 기준)	평균 약국 접근성 (2020년 기준)	
수도권	100%	2,590명/km ²	4.2km	3.2km	
무거주화 위험 공간	1.9%	-	7.2km	5.3km	
	인구감소 공간	57.5%	2,766명/km ²	4.4km	3.4km
	인구증가 공간	40.5%	2,464명/km ²	3.7km	3.0km
광역시	100%	2,471명/km ²	4.0km	3.2km	
무거주화 위험 공간	3.2%	-	6.7km	5.3km	
	인구감소 공간	60.5%	2,703명/km ²	3.9km	3.1km
	인구증가 공간	36.3%	2,300명/km ²	4.1km	3.1km
도내 시 지자체(경기 제외)	100%	444명/km ²	6.5km	4.3km	
무거주화 위험 공간	3.7%	-	9.7km	6.5km	
	인구감소 공간	87.1%	293명/km ²	6.7km	4.4km
	인구증가 공간	9.2%	2,054명/km ²	3.2km	2.5km
도내 군 지자체(경기 제외)	100%	100명/km ²	8.0km	5.0km	
무거주화 위험 공간	4.7%	-	10.6km	6.7km	
	인구감소 공간	90.6%	30명/km ²	8.1km	5.0km
	인구증가 공간	4.7%	1,558명/km ²	4.1km	2.6km

자료: 생활인프라 접근성은 2020년 국토지리정보원 국토조사(국토지리정보원, 2021) 결과를 활용하여 저자 작성

2050년 장래인구로 추정된 인구밀도를 유형별로 살펴보면, 수도권 → 광역시 → 도내 시 지자체 → 도내 군 지자체 순으로 인구밀도가 낮아질 것으로 나타났다. 수도권과 광역시의 경우 인구감소 공간이 인구증가 공간에 비해 인구밀도가 높은 것으로 나타났는데, 이는 도단위 지역의 인구감소 공간이 인구증가 공간에 비해 현저하게 인구밀도가 낮은 것과 정반대의 패턴을 보여준다. 이는 수도권과 광역시의 인구거주지역 확장에 따라 상대적으로 인구밀도가 낮은 공간에서 인구가 증가할 것으로 전망되었기 때문이다. 반면, 도내 시 지자체와 도내 군 지자체의 경우 인구밀도가 낮은 곳이 소멸의 위험에 크게 노출된 것으로 볼 수 있다.

생활인프라(초등학교와 약국) 접근성 관점에서 유형별로 살펴보면, 인구증가 공간이 대체로⁶⁾ 인구감소 및 무거주화 위험 공간에 비해 생활인프라까지 접근성이 짧아 이용이 편리한 것으로 나타났다. 지역별로 살펴보면 수도권에 비해 도내 시·군 지자체(경기도 제외)의 생활인프라 이용 편의성이 취약한데, 도내 시·군 지자체 중 인구증가 공간 유형의 경우 도리어 수도권에 비해 이용환경이 양호한 것으로 나타났다. 예를 들어, 약국까지의 접근성을 비교해 보면 수도권 평균 접근성은 3.2km로, 도내 시 지자체(4.3km)와 도내 군 지자체(6.7km)를 비교했을 때 수도권의 약국 이용 편의성이 좋은 것으로 나타났다.

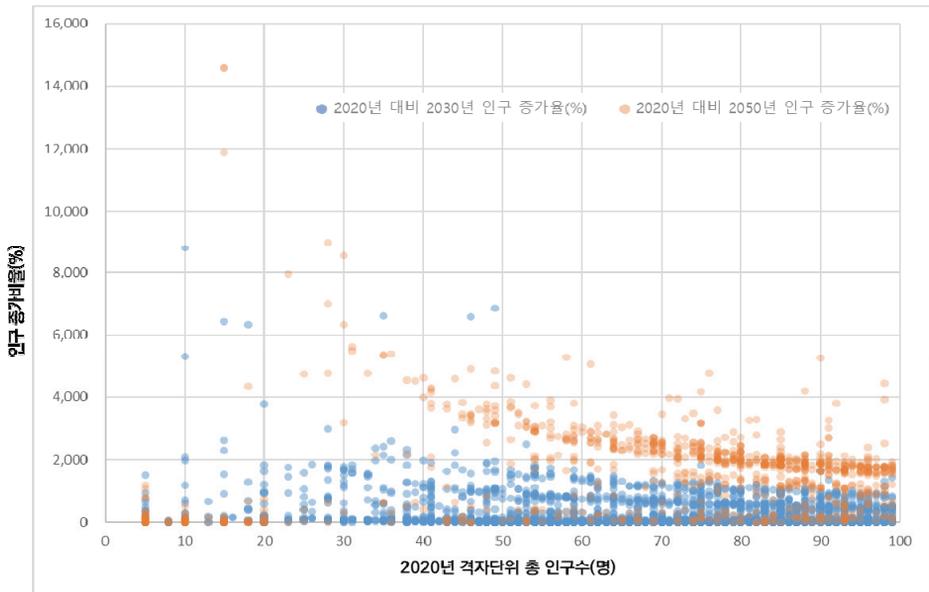
대체로 수도권과 광역시 전반에서 생활인프라 이용이 편리할 것으로 인식되고 있다. 그런데 분석 결과에 따르면, 도내 시·군 지자체에서 인구가 증가할 것으로 예상되는 공간(격자)의 경우 수도권과 비교했을 때 인프라 환경이 더 양호한 것으로 나타났다. 예를 들어, 수도권의 인구증가 공간의 평균 약국 접근성은 3.0km인데 도내 시·군 지자체내 인구증가 공간의 평균 약국 접근성은 각각 2.5km와 2.6km로 나타났다. 이와 같은 결과는 장래에 총인구가 크게 감소할 것으로 추계된 군 지역이라고 하더라도, 생활인프라 이용 편의성으로 나타난 정주여건이 좋은 공간이면 인구가 증가할 가능성이 있음을 시사한다.

6) 광역시 인구증가 공간의 경우 인구감소 공간에 비해 약 0.2km 가량 초등학교 접근성이 먼 것으로 나타남. 이는 광역시의 주변 지역 개발로 인해 현재 시점의 초등학교 인프라가 먼 지역으로 인구거주지역이 확장되어 해당 지역이 인구증가 공간 유형으로 분류되기 때문으로 해석할 수 있음.

3) 장래인구 예측 결과의 아웃라이어 검토

이 연구는 격자단위 인구학적 특성과 공간적 특성을 반영하기 위해 인공지능 기법과 코호트요인법을 활용한 장래인구예측 모델 개발을 처음 시도하여, 그 가능성을 보여줬다는 데 큰 의의가 있다. 다만, 이 연구에서 제시한 GPAI 초기모델을 적용해 도출된 5년단위 장래인구 예측 결과를 바로 실무에 사용하기에는 여러 가지 이슈가 발생할 가능성이 있다. 예를 들어, <그림 4-9>는 2020년 격자별 인구수가 200명 이하일 때 2030년과 2050년에 예측된 인구의 증가 비율의 아웃라이어(outlier, 이상치)를 비교한 것이다. 인구증가율의 이상치를 2020년과 비교하였을 때 장기예측일수록 이상치가 큰 격자가 잦은 빈도로 나타났다. 이는 선행연구 중 Grossman·Wilson·Temple (2022)이 LSTM을 활용하여 5년 이내 단기예측이 중장기 예측보다 정확도가 더 높게 나타난 것과 같은 맥락이라고 볼 수 있다.

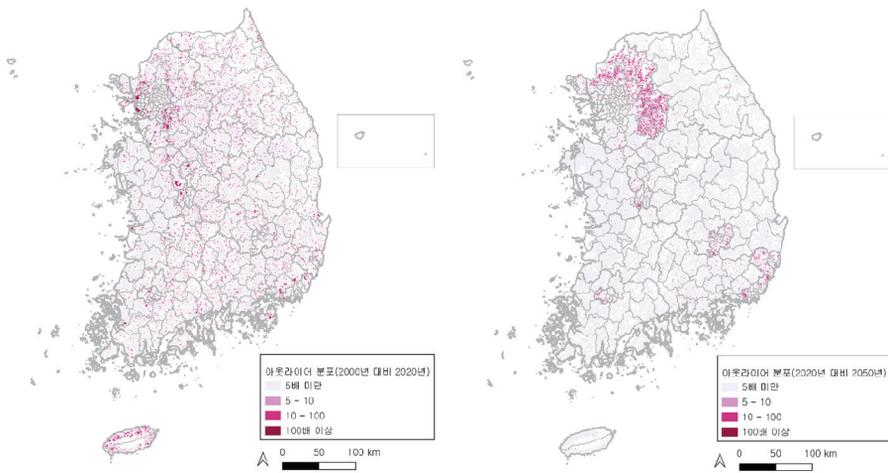
그림 4-9 | 2050년 장래인구 예측결과의 아웃라이어 현황



자료: 저자 작성

2000년부터 2020년까지 20년간의 격자단위 인구밀도의 변화 분포를 살펴보면, 실제 인구변화의 패턴에서도 아웃라이어를 발견할 수 있었다(〈그림 4-10〉 좌측). 2050년 예측결과 아웃라이어로 탐지된 격자의 위치와 비교해보면(〈그림 4-10〉 우측), 이들 지역은 대부분 인구밀도가 상대적으로 낮은 지역 또는 최근 개발 추세에 따라 갑자기 인구가 증가한 지역으로 볼 수 있다. 이와 같은 지역에 따른 이상치 및 오차 발생은 Riiman et al. (2019)와 Grossman·Wilson·Temple(2022)이 공통적으로 지적한 것과 같이 모델의 정확도가 인구 규모 또는 도심/비도심에 따라 달라지는 것을 잘 보여준다.

그림 4-10 | 2000년 대비 2020년, 2020년 대비 2050년 인구변화 증가율 아웃라이어 분포



자료: 저자 작성

〈그림 4-10〉의 우측에서 상대적으로 아웃라이어가 적은 전라남도와 아웃라이어가 눈에 띄게 많은 경기도에 대해 광역지자체에서 공표한 장래인구추계와 GPAI의 결과를 비교하였다(〈표 4-7〉 참조). 각 광역지자체는 2019년도 말에 통계청에서 배포한 코호트요인 기반의 KOSTAT-SPP 도구를 활용하여 2017년~2037년 시·군·구별 장래인구를 추계·공표하였다⁷⁾. 해당 추계의 기준인구 시점은 2017년으로 2022년에 공표한 시

도단위 통계청 장래인구추계를 따르는 GPAI 결과와 일부 차이가 있다. 이를 감안하더라도 경기도의 경우 두 장래인구 간에 큰 차이를 보이는 지역이 나타났다.

예를 들어, 경기도 가평군, 양평군, 포천시, 여주시와 같이 경기도 외곽지역에 대해 GPAI가 장래인구를 더 크게 예측한다. 이는 광역 지자체가 활용한 KOSTAT-SPP는 코호트요인법에 따라 인구변동요인만 고려하였으나, GPAI는 인구변동요인과 함께 토지이용의 변화 패턴을 함께 고려하여 격자단위 장래인구를 예측하기 때문이다. 경기도 외곽지역에 대해 장래인구가 더 많이 예측된 것은 토지이용 변화에 따라 최근 해당 지역의 인구거주지역이 확장되는 패턴을 학습하였기 때문이다. 즉, 인구거주 공간의 확장과 함께 경기도 외곽지역의 특성인 전원주택 또는 세컨드하우스의 개발 등의 토지개발 등 공간의 특성 변화 추세에 따라 GPAI 모델에서는 경기도 외곽지역의 인구가 크게 증가할 것으로 예측한 것이다. 이와 같은 GPAI의 특성을 고려하여 지역 내 인구증가 예측 결과를 해석하고 계획 및 정책 수립에 활용할 필요가 있다.

표 4-7 | 시도 광역시 공표 장래인구추계와 GPAI 예측 비교

시도	시군구	시도 광역시 공표 장래인구추계		GPAI 예측		차이 (B/Ax100)
		2030년	2035년(A)	2030년	2035년(B)	
전라남도	목포시	208,040	200,539	216,020	206,339	102.9
	여주시	249,130	242,191	279,210	275,446	113.7
	순천시	262,351	258,710	251,310	248,454	96
	나주시	130,015	133,607	129,177	128,481	96.2
	광양시	144,231	142,208	150,977	148,560	104.5
	담양군	44,002	44,720	44,820	45,075	100.8
	곡성군	26,656	26,941	25,847	24,523	91
	구례군	25,010	25,660	26,077	27,087	105.6
	고흥군	57,441	57,329	48,617	46,980	81.9
	보성군	36,016	36,023	32,264	30,594	84.9
	화순군	55,934	55,688	53,747	52,242	93.8
	장흥군	34,084	34,391	31,435	31,106	90.4
	강진군	29,937	29,731	32,677	32,164	108.2
	해남군	59,921	58,848	53,754	51,358	87.3
영암군	50,204	49,211	53,021	52,493	106.7	

7) 경기도(2019.12) 2017년 기준 경기도 시군별 장래인구특별추계 지침서

시도	시군구	시도 광역시 공표 장래인구추계		GPAI 예측		차이 (B/Ax100)
		2030년	2035년(A)	2030년	2035년(B)	
경기도	무안군	76,817	75,534	73,861	72,950	96.6
	함평군	29,187	29,518	20,628	19,291	65.4
	영광군	47,121	46,681	41,701	42,712	91.5
	장성군	42,312	43,045	40,608	38,594	89.7
	완도군	45,428	44,782	41,592	40,691	90.9
	진도군	27,073	26,811	26,067	25,636	95.6
	신안군	34,136	34,256	25,101	22,092	64.5
경기도	수원시	1,280,120	1,295,409	1,054,181	977,581	75.5
	성남시	914,452	908,420	864,203	804,078	88.5
	의정부시	464,634	466,033	428,492	398,774	85.6
	안양시	526,624	513,300	501,713	461,345	89.9
	부천시	803,509	792,263	687,254	633,789	80
	광명시	290,384	281,302	305,112	278,540	99
	평택시	587,527	601,882	562,834	533,638	88.7
	동두천시	95,121	96,028	115,614	121,244	126.3
	안산시	638,732	616,778	614,581	573,382	93
	고양시	1,082,941	1,097,556	1,009,019	951,333	86.7
	과천시	45,358	42,824	74,206	70,042	163.6
	구리시	209,600	212,108	185,340	174,224	82.1
	남양주시	753,576	768,758	821,464	843,510	109.7
	오산시	252,550	257,545	220,657	206,019	80
	시흥시	554,913	553,590	468,945	436,084	78.8
	군포시	261,824	256,917	250,214	231,396	90.1
	의왕시	152,188	148,682	138,734	129,672	87.2
	하남시	379,594	401,594	231,079	219,261	54.6
	용인시	1,156,652	1,170,131	1,043,270	1,017,361	86.9
	파주시	515,582	528,704	593,668	667,476	126.2
	이천시	234,111	237,041	408,525	513,732	216.7
	안성시	208,657	211,903	228,947	239,660	113.1
	김포시	548,686	566,183	454,296	441,848	78
	화성시	1,124,629	1,175,683	848,680	792,848	67.4
	광주시	452,106	468,760	491,104	549,031	117.1
	양주시	246,292	248,746	344,773	404,683	162.7
	포천시	153,220	153,642	324,501	439,525	286.1
	여주시	117,980	121,288	296,002	433,521	357.4
	연천군	43,396	44,017	102,486	141,734	322
	가평군	65,630	67,822	247,406	352,636	519.9
양평군	133,057	140,181	503,794	663,448	473.3	

자료: 광역지자체 공표 장래인구추계(경기도: https://kosis.kr/statHtml/statHtml.do?orgId=210&tblId=DT_210008_2017001&conn_path=I2 / 전라남도: https://kosis.kr/statHtml/statHtml.do?orgId=215&tblId=DT_215007_001&conn_path=I2, 최종접속: 2022.11.18.)와 본 연구 결과를 활용하여 저자 작성

3. GPAI 고도화 방안

이 연구에서 개발한 GPAI를 전국에 적용하여 2050년까지 장래인구를 예측한 결과 인구 변화 패턴을 포함하여 아웃라이어의 분포를 확인할 수 있었다. GPAI 아키텍처 설계와 적용 결과 과정에서 향후 GPAI가 고도화되기 위해 필요한 사항을 데이터 측면과 사회·환경 현상 측면에서 도출할 수 있었다. 각 측면에서 현재 단계의 GPAI의 한계와 함께 후속 과제를 통해 방법론을 고도화하는 방안을 제시하였다.

1) 데이터 측면의 고도화 방안 및 향후 과제

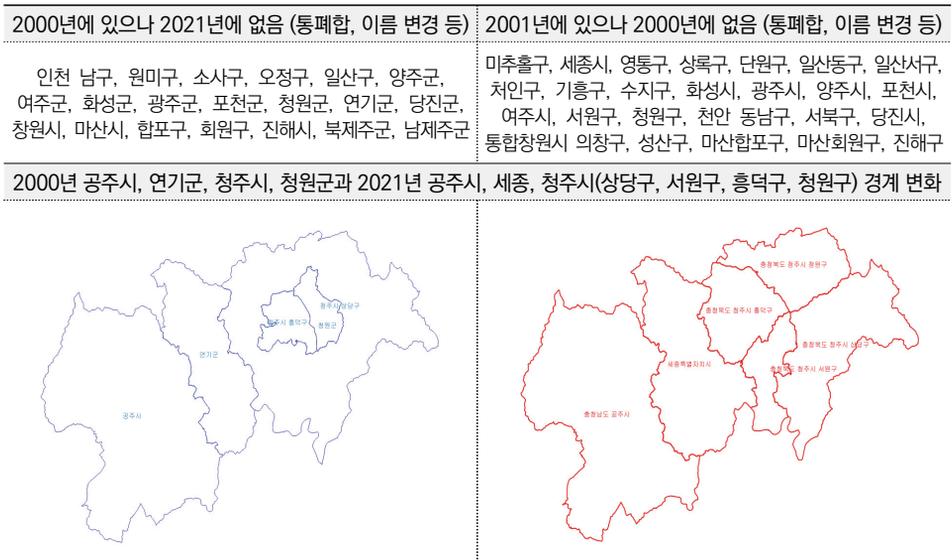
첫 번째 데이터 측면의 고도화 방안은 국지적 GPAI 방법론의 개발이다. 앞서 <그림 4-10>에서 확인한 것과 같이 아웃라이어로 도출된 지역 대부분이 인구밀도가 상대적으로 낮은 지역 또는 최근 개발 추세에 따라 갑자기 인구가 증가한 지역으로 나타났다. Chi·Zhou·Voss(2011)이 다양한 회귀모형을 밀워키의 센서스 트랙에 적용하였을 때 도심과 비도심 여부에 따라 예측력의 차이가 있음을 확인하였다. 따라서 이와 같은 한계를 극복하고 방법론을 고도화하기 위해서는 향후 지역에 특화된 국지적 모델을 개발할 필요가 있다.

두 번째 데이터 측면에서 고도화 방안은 사망률, 출산율 및 인구이동과 같은 인구학적 특성 관련 데이터를 격자단위로 구축하는 것이다. 사망률과 출산율의 경우 시도별 데이터를 격자의 위치에 따라 적용하였는데, 특히 인구이동의 경우 같은 지역 내 인구이동의 비중이 높는데, 시·군·구 또는 시·도단위의 장래예측에서는 지자체 간의 인구이동만 고려하면 되지만 격자의 경우 같은 지자체 내 격자간의 이동 데이터가 필요하다. 향후 격자단위의 인구이동 데이터를 구축하는데, 격자의 위치에 따라 읍·면·동 또는 시·군·구의 순이동자 수를 공간 보간을 통해 할당할 수 있다. 물론 이 과정에서 어떤 기준으로 공간 보간을 수행할지, 그 결과가 타당한지에 대한 연구가 진행되어야 할 것이다.

나아가 정확한 격자단위 인구이동 데이터를 얻기 위해서는 행정안전부와 통계청에서

주소 기준 읍·면·동 또는 시·군·구 단위의 인구이동 데이터를 구축하는 것과 같이 국가 관심지점정보 기준을 차용하고 있는 격자단위로 통계를 생성할 수 있겠다. 현재 인구 이동과 관련하여 연구자가 가장 어려움을 겪고 있는 분야 중 하나가 행정구역의 빈번한 변화로 인해 과거와 현재까지의 인구이동의 변화 추세를 파악하는데 어려움이 있다는 것이다(〈그림 4-11〉 참조). 향후 격자단위의 순이동자 수 및 전출-전입 빅데이터를 생산함으로써 인구이동 및 장래인구 예측과 관련된 연구 분야를 발전시킬 수 있을 것으로 기대된다.

그림 4-11 | 행정구역 변화(2000년 대비 2021년) 및 공주-연기-세종-청주-청원 일대 행정구역 변화



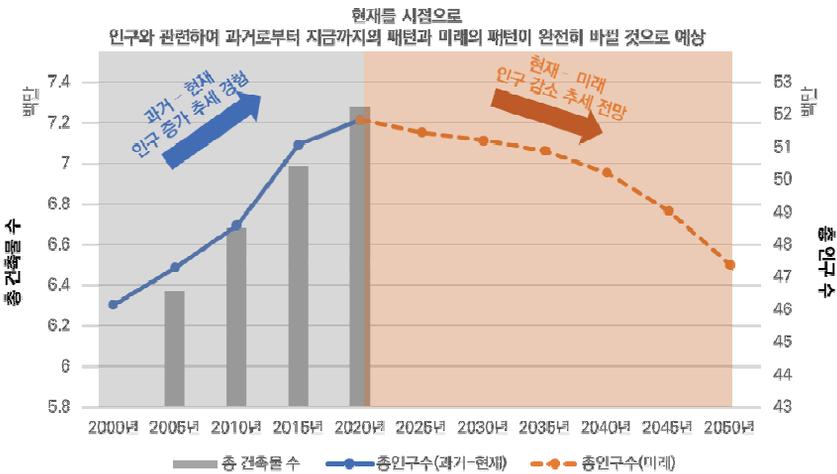
자료: 저자 작성

세 번째 데이터 측면의 고도화 방안은 공간적 특성을 나타내는 변수를 다양하게 활용하는 것이다. 장래의 토지이용과 관련하여 토지 피복, 국토환경성평가 등급, 용도지역지구 규제, 그린벨트 등급, 도로까지 접근성 및 하천 등에 따라 건축물의 확산 패턴이 달라질 수 있는데, 이 연구에서는 과거로부터 현재까지의 건축물 분포의 변화에만 초점을 맞추었다. 토지이용과 관련된 인문환경과 자연환경적 요소를 고려하여 격자단위의 토지이용 변화 및 거주가능지 분포를 예측하는 모델을 후속 연구로 수행할 수 있다.

2) 사회·환경 현상 측면의 고도화 방안 및 향후 과제

지금까지 우리나라는 지속적으로 인구가 증가해왔으며, 토지개발 등에 따라 건축물 수도 증가해왔다. 이와 같은 과거부터 현재까지의 인구 및 국토개발의 특성은 향후 인구감소시대를 맞이하며 크게 변화할 것으로 예상된다(〈그림 4-12〉 참조). 그런데 인공지능 모델, 특히 LSTM 방법의 경우 과거로부터의 변화 추세를 학습하여 장래의 현상을 예측하는 시계열 모델이기 때문에, 이와 같은 과거로부터의 변화가 예측 결과의 큰 오차를 가져올 수 있다. 만약 지금까지 인구가 증가한 것과 같은 추세로 2050년까지 인구가 증가하는 것을 전제로 한다면, 과거의 데이터를 학습했을 때 시계열 추세가 잘 반영될 수 있지만 증가가 감소로 변환되는 현재 시점에서는 과거의 시계열 추세가 도리어 오차를 크게하는 요인이 될 수 있는 것이다. 이와 같은 특수한 우리나라의 인구 감소 사회 현상 측면을 모델에 반영하기 위해, 2008년(1억 2,808만 명)을 정점으로 13년째 인구가 감소하고 있는 일본의 변화 추세를 함께 학습하는 방법론 개발 등 향후 국제 공동연구를 수행할 필요가 있다.

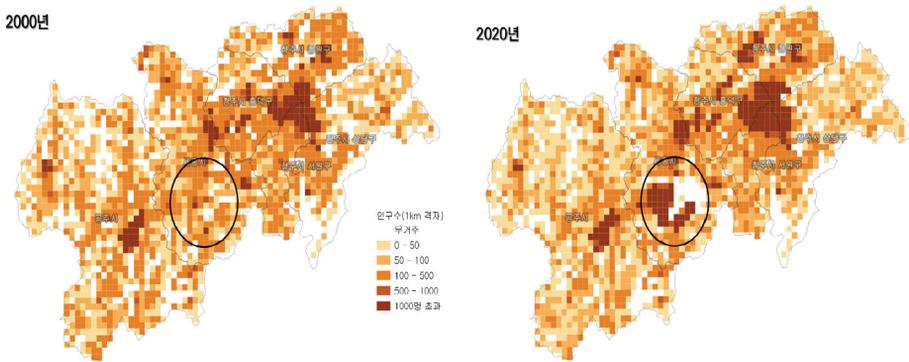
그림 4-12 | 과거-현재-미래의 건축물수 및 총인구수 변화 추세



자료: 총건축물수(국토교통부 건축물통계), 총인구수(과거-현재: 통계청 인구총조사, 미래: 통계청 장래인구추계) 활용하여 저자 작성

두 번째 사회 현상의 측면에서 고도화 방안은 정책/계획에 따른 택지개발에 따른 대규모 주택공급과 같은 개입을 방법론에 반영하는 것이다. <그림 4-13>은 2000년부터 2020년까지 세종특별자치시-공주시-청주시 일대의 인구변화를 1km 격자단위로 나타낸 것이다. 그림의 원으로 강조한 부분은 세종특별자치시의 행정중심복합도시 개발에 따라, 급격히 인구가 증가한 곳이다. 이 택지지구는 사업 대상지는 물론 공주, 충주를 포함한 주변 지역의 인구분포에 큰 영향을 미쳤다.

그림 4-13 | 세종특별자치시-공주시-청주시 일대 인구 변화(2000년 ~ 2020년)

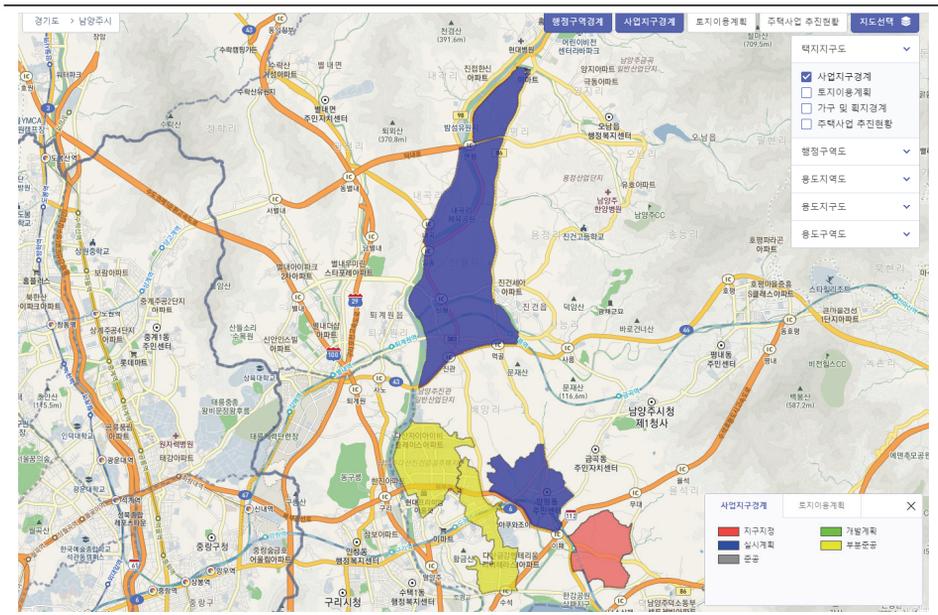


자료: 저자 작성

격자단위 장래인구 전망이 지자체에서 활용되기 위해서는 현재의 제도적 테두리 내에서 가급적 지자체의 의지를 최대한 확보하는 차원에서 인구가 전망되어야 하며, 이를 위해서는 중앙정부 및 지자체의 장래 주택공급 계획이 반영되어야 한다. 인공지능기법을 비롯한 모델 기반의 장래인구 예측치는 중앙정부나 지자체의 주택공급 등의 개발 개입이 없을 때 자연스러운 인구변동 및 토지이용 변화에 따른 인구분포 변화 기준(baseline)으로 볼 수 있다. 향후 이 결과를 토대로 각종 중앙정부와 지자체의 주택공급에 따른 인구이동 효과를 반영하는 방법으로 고도화 연구를 수행할 수 있다.

예를 들어, <그림 4-14>와 같이 경기도 남양주시 ‘남양주 양정역세권 도시개발구역’ (지구지정), ‘남양주 다산지금 공공주택지구’ (부분준공), ‘남양주 다산진건 공공주택지구’ (부분준공), ‘남양주 진접2 공공택지지구’ (실시계획), ‘남양주왕숙 공공주택지구’ (실시계획), ‘남양주왕숙2 공공주택지구’ (실시계획)가 2030년 전까지 주택공급계획에 따라 인구분포의 기준(baseline)이 어떻게 바뀔지 시뮬레이션이 가능하다.

그림 4-14 | 경기도 남양주시 택지개발 계획 구역(지구지정-붉은색, 실시계획-푸른색, 부분준공-노란색)



자료: 택지정보시스템(<https://jigu.go.kr/index.do>, 최종접속일: 2022.11.05.)

이보경 외(2021)에서는 주택공급에 따른 인구이동 패턴을 사전에 파악하여 같은 지자체 내에서의 인구이동 그리고 다른 지자체에서의 인구이동을 구분하여, 전자의 경우 관련 전출지에서의 인구감소(또는 공가연쇄작용)와 후지에서는 지역간 인구이동의 제로섬게임 등을 함께 고려할 필요성을 강조하였다. 중앙정부와 지자체의 대규모 주택공급이 인구분포 변화에 미치는 영향을 반영하기 위해 인구분포의 기준(baseline)에 주택공급 예정 구역(택지지구)을 공간을 기준으로 융합하고, 공급 예정 주택 규모와 실제

인구의 이동 경로 정보를 활용하여 주택공급에 따른 파급효과를 전망해볼 수 있다. 구체적으로 이보경 외(2021)에서 제시한 택지지구 유형별 유입인구의 거리분포를 활용하면, 주택공급에 따라 거리별로 어떻게 인구가 이동하여 인구분포에 영향을 미칠지 시뮬레이션할 수 있다(〈표 4-8〉 참조). 이와 같은 후속 과제 수행을 통해 주택공급과 같은 개입이 없을 때 인구분포를 기준(baseline)으로 정의하고, 정책 개입 시 인구분포의 변화를 비교하여 택지개발의 효과를 시뮬레이션하고 적정입지를 선정하는데 의사결정 자료로 활용할 수 있을 것이다.

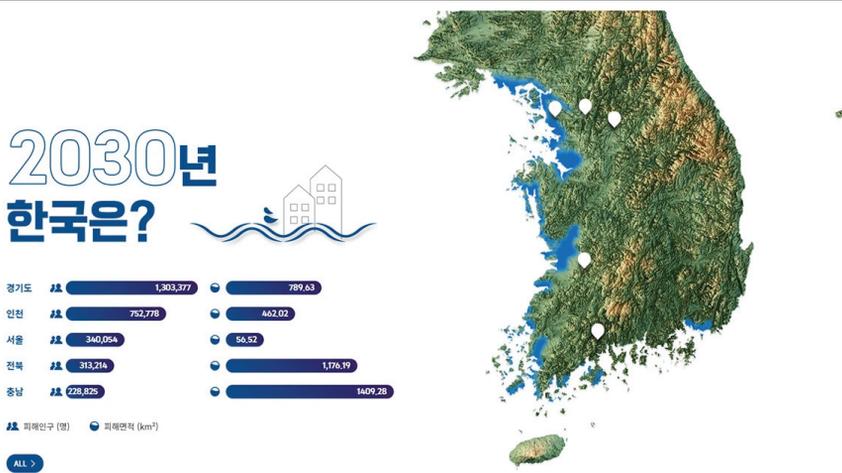
표 4-8 | 택지지구 그룹별 유입거리별 인구 비중 누적 비율의 평균

유입거리	유입거리별 인구 비중 누적 평균(%)		
	수도권 택지지구(110개)	비수도권 혁신/행복도시(10개)	비수도권 일반 택지지구(158개)
4km	34.16	23.83	42.62
8km	55.30	44.68	62.78
12km	67.42	54.92	73.16
16km	74.36	61.66	78.36
20km	79.20	65.49	81.40
24km	82.71	68.24	83.10
28km	85.33	69.38	84.12
32km	87.21	70.16	84.99
36km	88.48	70.93	85.58
40km	89.61	71.75	86.04
50km	91.66	73.29	87.18
60km	93.07	74.42	88.32
70km	94.26	75.79	89.64
80km	94.98	77.42	90.88
90km	95.39	79.23	92.04
100km	95.72	81.54	92.92

자료: 이보경 외(2021), 표4-13, p76

세 번째 고도화 방안은 기후변화에 따른 인구거주지역 변화와 기후변화에 대응하기 위한 탄소중립 정책 등으로 인한 토지이용 패턴의 변화와 같은 현상을 방법론에 반영하는 것이다. <그림 4-15>와 같이 그린피스에서 미국 기후변화 연구기관 클라이밋 센트럴의 연구 결과를 바탕으로 장래 우리나라 국토의 침수 현황을 시뮬레이션하였다. 그 결과 해수면 상승 및 태풍 등 복합적 영향으로 2030년에는 국토의 5% 면적이 침수되고 약 332만 명이 피해를 겪을 것으로 예상되었다⁸⁾. 거주 가능한 국토가 변하면 격자단위 장래인구 예측의 결과 역시 큰 영향을 받는다. 따라서 기후변화에 따른 거주지 변화 시나리오를 고려하여 격자단위의 장래인구 분포를 예측할 수 있는 방법론 개발 연구가 향후 수행되어야 할 것이다. 예를 들어, 제2장에서 검토한 Chen et al. (2020a)의 연구와 같이 기후변화에 따른 정책적 완화 및 대응조치 시나리오인 SSPs에 따라 토지이용을 시뮬레이션하고, 이에 따른 인구 거주지의 변화를 격자단위 인구분포 모델링의 주요 변수로 활용하는 방법론을 개발하는 등의 연구로 발전할 수 있다.

그림 4-15 | 그린피스가 예측한 해수면 상승에 따른 2030년 우리나라 침수 영향



자료: 그린피스 웹사이트(<http://www.climate.or.kr/>, 최종 접속일: 2022.11.05)

8) 그린피스 웹사이트(<https://www.greenpeace.org/korea/update/14990/blog-ce-flood-web-page/>, 최종 접속일: 2022.11.05.)

4. 소결: 방법론의 독창성 및 학술적 기여

GPAI 아키텍처를 설계하고, 적용하여 그 결과를 검토하고, 향후 고도화 방안을 제시하는 일련의 과정에서, 이 연구가 가지는 학술적 독창성과 이론에 대한 기여점은 크게 네 가지로 꼽을 수 있다. 첫째, 다양한 소지역의 공간단위 중 1km 격자단위로 장래 인구를 예측하는 방법론을 개발한 것이다. 둘째, 전통적 코호트요인법 모델과 새로운 인공지능 기법 모델을 결합한 새로운 방법론을 제안하였다는 것이다. 셋째, 랜덤포레스트와 CNN-LSTM 등 다양한 인공지능 기법을 활용하여 정밀한 공간단위의 인구-인구학적 특성-공간적 특성의 관계를 밝히고, 딥러닝을 활용하여 처음으로 인구분포에 큰 영향을 미치는 건축물 수, 주거용 연면적 등과 같은 공간적 특성의 변화를 예측하였다는 것이다. 마지막으로 GPAI를 적용한 결과를 면밀하게 검토하여 고도화 방안을 제시하면서 향후 후속 연구의 진행 방향을 설정하였다는 것이다.

소지역이라는 공간적 범위는 시대, 활용 목적 및 국가 통계 구축 현황 등에 따라 다양하게 정의될 수 있다. 기존 선행연구 대부분이 소지역이라고 하더라도 행정구역 또는 통계 집계구 단위로 인구를 전망해왔다. 제3기 인구정책 TF에서 “통계청이 생산한 인구구조 변화와 관련된 통계를 격자(格子) 단위 공간정보와 결합하여 정밀한 용·복합 데이터 생산”(관계부처 합동, 2021. 1. 27. : 12)의 중요성을 강조하는 등 격자단위 장래인구 전망에 대한 수요가 증가함에 따라 이 연구에서는 1km 격자단위의 장래인구 예측 방법론을 개발하였다. 따라서 최대한 격자가 가지는 공간적 특성을 모델에 반영하기 위해 다양한 출처의 원천데이터로부터 격자단위의 공간 데이터를 가공·생산하였다. 이를 통해 전통적인 인구학적 변동요인뿐 아니라 공간적 특성에 대한 새롭고 다양한 데이터를 활용할 수 있는 가능성을 시사하였다. 그리고 Tobler의 셀룰러지오그래피 개념을 바탕으로 1km격자 공간단위 데이터의 시·공간 관계성을 정의하고 인공지능 기법에 적용하였다는 데 기존 방법론과 차별성이 있다.

둘째, 이 연구에서는 GPAI를 개발하기 위해 인구학적 특성에 따라 시나리오 기반으로 인구의 변화를 추계하는 코호트요인법 모델과 과거 인구학적·공간적 특성을 학습하

여 정밀한 공간단위의 인구분포를 예측하는 인공지능 모델을 결합하였다. 전통적으로 사망률, 출생률, 순이동률 등 인구학적 변동요인을 고려한 코호트요인법이 널리 활용됐다. 그러나 전국 또는 시·도단위 장래인구 추계에 활용된 코호트요인법을 정밀한 공간단위에 적용하는 데 한계가 있으므로, 인구 거주와 밀접하게 관련된 토지이용 현황을 반영할 필요가 있다. 제2장의 사례조사에서 검토한 것과 같이 기존 선행연구에서 코호트요인법, 공간통계모델, 인공지능 기법 등을 각각 활용하여 장래인구를 예측해왔다. 이 연구에서는 각 방법론의 한계를 보완할 수 있도록 코호트요인법 기반 모델과 공간통계 특성을 반영한 인공지능 기법 기반 모델을 결합하는 아이디어를 최초로 제시하였다는데 방법론의 독창성을 가진다.

셋째, GPAI는 정밀한 공간단위의 인구분포에 직접적으로 영향을 미치는 건축물 수, 주거용 연면적 등의 공간적 특성의 변화를 예측하는 인공지능 기법을 포함하고 있다. 장래인구를 예측하기 위해 인공지능 기법을 적용하는 것은 매우 보수적으로 진행되어 최근 2~3년 사이에 소수의 연구가 시행되었다. 이들 대부분 간단한 딥러닝 모델이 기존의 선형/지수 모델에 비해 얼마나 좋은가를 비교하고 그 활용의 가능성을 제안하는 것에 그쳤다. 예측 대상 공간단위가 작아질수록 그 지역에 사람이 거주할 수 있는 토지이용 등의 공간적 특성이 큰 영향을 미치는 데, GPAI에서는 이 요인을 인공지능 기법을 적용하는 접근방식을 처음으로 제안하였다.

마지막으로 모델의 고도화 방안과 향후 과제를 제시함으로써 소지역단위 장래인구 예측 방법론 관련 이론의 새로운 연구 방향을 개척하였다는 데 학술적으로 기여하는 바가 크다. GPAI 적용 결과를 통해 살펴본 것과 같이 지금 버전의 GPAI를 바로 정책이나 계획에 적용하는데 여러 가지 이슈가 발생할 가능성이 있다. 이해관계자가 받아들일 수 있고 또한 실무에서 쉽게 활용할 수 있도록 GPAI 방법론 및 활용 도구 등의 관점에서 다양한 후속 연구가 요구된다. 따라서 GPAI 학습의 기초가 되는 데이터 측면과 최근 크게 변화하고 있는 사회·환경 현상 측면에서 GPAI 고도화 방안과 구체적인 후속 과제를 도출하였다.



CHAPTER 5

정책 제언 및 결론

- 1. 정책 및 계획 부문 활용방안 117
- 2. 활용 확산을 위한 정책 제언 123
- 3. 결론 129

05 정책 제언 및 결론

이 연구에서는 소지역 중 격자단위 장래인구 예측 AI 모델(Gridded Population Forecast with AI, GPFI)을 정책과 계획에서 활용하기 위한 방안을 모색하였다. 그리고 GPFI가 실무에서 신뢰성 있게 활용되기 위해 필요한 사항을 제언하였다. 마지막으로 GPFI아키텍처 설계 과정, 적용 결과, 고도화 방안 및 정책 활용방안 등을 요약하여 연구의 결론으로 제시하였다.

1. 정책 및 계획 부문 활용방안

이 연구에서 제안한 격자단위 장래인구 예측 AI 모델(Gridded Population Forecast with AI, GPFI) 적용 결과와 모델은 활용 측면에서 다음과 같은 세 가지 장점이 있다. 첫째, 1km 격자라는 공간범위의 특성에 따라 예측 공간 범위와 위치가 표준화되어 있기 때문에 통시적이면서도 유연하게 구역을 지정할 수 있다는 것이다. 둘째, 미래예견적 정책 및 계획을 수립할 수 있도록 장래의 인구변화 정보를 활용하여 시뮬레이션을 통해 과학적 의사결정을 내리는 데 도움을 줄 수 있다. 셋째, 시·군·구보다 작은 소지역단위 인구정보를 통해 급격한 인구변화에 큰 영향을 받는 장소를 타겟팅할 수 있어, 지자체에서 관련 계획 수립 시 근거자료로 활용할 수 있다. 이와 같은 장점을 바탕으로 정책 및 계획 부문에서의 구체적인 활용 방안은 다음과 같다.

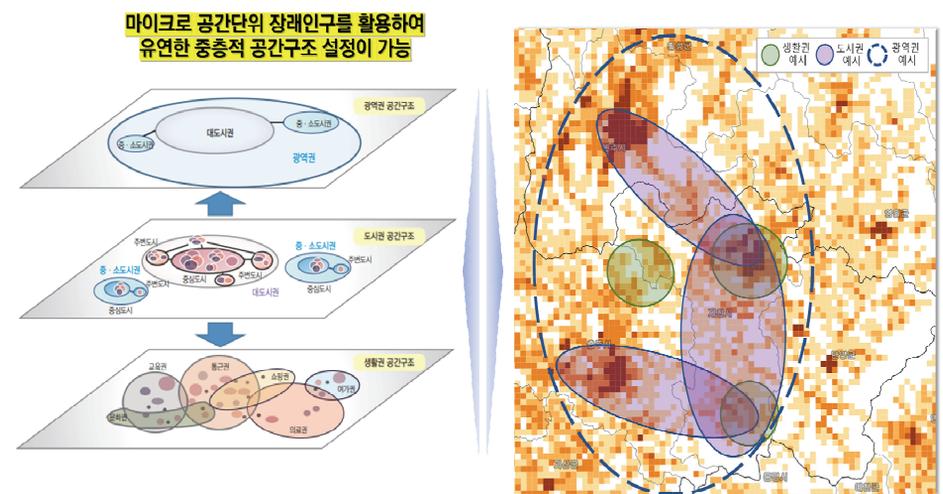
1) 유연한 정책·계획의 공간 구역 설정

격자단위 장래인구 정보를 활용하면 생활권부터 광역권까지 이슈에 따라 유연하게 공간 구역을 설정하여 정책 집행 및 평가를 할 수 있다. 제5차 국토종합계획(2020~2040)에서는 발전전략 중 하나로 인구감소에 대응하여 유연한 도시개발·관리와 집약적인 도

시공간구조 재편의 필요성을 강조하였다(국토교통부, 2019: 28). 그리고 2022년도 말 정부는 '인구구조 변화와 대응방안' (저출산고령사회위원회·관계부처 합동, 2022. 12. 28.) 을 통해 국토·지역·도시 정책의 집행 단위가 행정구역 기준이기 때문에 실제 인구 규모와 경제활동과 괴리되어 있다는 점을 지적하였다. 그리고 이를 극복하기 위해 공간 빅데이터를 활용하여 행정구역에 구애받지 않은 광역 경제·생활권 등의 정책구역을 설정하고 육성하는 방안을 제시하였다.

GPAI의 1km 격자단위 장래인구 정보를 활용하면 장래인구 변화를 고려한 다양한 정책집행 단위를 설정할 수 있다. 국토 균형발전 관점에서도 기능적 측면에서 광역권-도시권-생활권으로 이루어진 중층적 공간구조 정의의 필요성이 강조되어 왔다(민성희 외., 2020). 기존의 행정구역에 구애받지 않고 기능적 측면으로 구성된 공간구조를 정의하고, 기초단위부터 광역 연계까지 정합성을 갖춘 정보로 계획을 수립하는데 1km 격자단위 장래인구 정보를 활용할 수 있다(<그림 5-1> 참조).

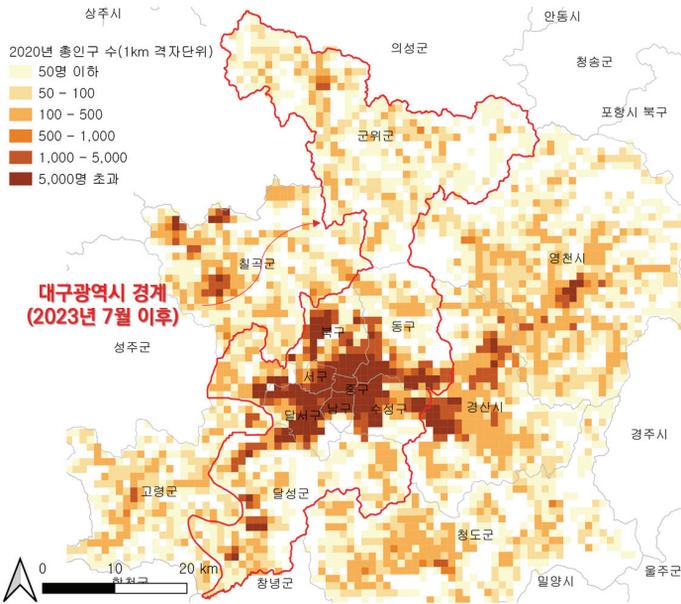
그림 5-1 | 중층적 공간구조와 격자단위 공간 데이터 활용의 예시



자료: 민성희 외(2020) 그림2, p4를 활용하여 저자 재구성

물론, 행정구역은 정책을 시행하고 집행하는 단위이기 때문에 인구감소에 적응하기 위한 축소사회를 위해서는 유연한 정책관심지역의 설정 뿐 아니라 행정구역의 재설계가 필요하다. 이미 중앙정부와 광역지자체에서도 인구감소시대에 적응하기 위해 행정구역 개편을 실행하고 있다. 인천광역시는 2026년까지 중구와 동구를 제물포구로 통합하고, 중구와 영종도를 영종구로 분리 독립할 계획을 하고 있다. 그리고 2023년 7월 1일자로 경상북도 군위군이 대구광역시에 편입될 계획이다(<그림 5-2> 참조)¹⁾. 인구감소시대 지자체의 지속가능한 경영을 위해서는 비단 인천광역시와 대구광역시-군위군뿐 아니라 모든 지자체가 조정의 대상이 될 수 있다. 중앙정부와 지자체는 국토·지역의 관점에서 장래의 인구분포와 함께 통근·통학과 같은 경제활동의 공간 간 연계성을 진단·전망하여 행정구역 조정의 과학적 근거로 활용할 수 있을 것이다.

그림 5-2 | 대구광역시 및 인접지역의 1km 격자단위 인구 분포와 경계 변화



자료: 저자 작성

1) 2022년 12월 8일 '경상북도와 대구광역시 간 관할구역 변경에 관한 법률안' 통과

2) 미래예견적 계획수립을 위한 과학적 의사결정 지원

인구감소 시대에 적응하기 위해서는 단순히 그 상황을 받아들이는 것이 아니라 인구가 살아가는데 필요한 근원적인 플랫폼을 재구조화하는 것이 필요하며(민보경, 2021), 이 과정에서 미래예견적 의사결정이 요구된다(이보경 외, 2020). 특히 인구감소에 따라 소멸의 위기를 맞이하고 있는 지자체에서는 장래 인구변화 정보를 바탕으로 효과적이고 지속가능한 자원 배분과 인프라 운영 등 계획을 수립해야 한다. 예를 들어, 인구감소에 따라 세수의 감소 등에 따라 지자체의 도로, 상수도 등은 운영 및 관리에 직접적인 영향을 받는다. 조만석 외(2020)는 인구 규모별로 총괄원가의 시계열 추세를 제시하며, 인구밀도가 낮은 군지역의 경우 원가가 지속적으로 증가하고 있어 효율성 및 지속가능성의 이슈를 제기하였다. 지방상수도의 경우 기초지자체단위의 직영 기업으로 운영되고 있어 지속적인 인구 감소는 기초지자체의 예산 부족과 직접적으로 연계되어, 향후 지역의 지속가능성에 큰 타격을 미칠 수 있다. 이와 같은 상황에서 수도사업의 경영 효율성 강화를 위해 소지역단위 장래인구 예측 정보를 활용하면 미래 변화에 적응할 수 있는 상수도 공급 인프라의 효율적 재배치를 검토할 수 있을 것이다.

이채정(2021)은 지역별 주요 사회서비스의 분포와 격차를 정밀한 공간단위로 분석하여 인구구조 변화에 적응할 수 있는 효율적인 서비스 전달 체계 구축 및 운영 방안 마련의 필요성을 강조하였다. 이때 시군구 또는 읍면동 단위 접근이 아닌, 좀더 정밀하고 행정구역에 구애받지 않은 생활공간 단위로 서비스 현황과 수요의 변화를 예측하는데 격자단위 장래인구 예측을 활용할 수 있다. 예를 들어, 이보경 외(2020)은 이보경(2019)의 500m 격자단위 장래인구 데이터를 활용하여 도서관과 어린이집의 시설별 장래 배후수요 변화를 시뮬레이션하고, 각 시설의 통폐합 및 신규시설 신설에 따른 지역의 생활인프라 서비스의 포용성 변화를 제시하였다. 전국의 어린이집, 작은도서관 및 국공립도서관을 대상으로 시설 주변(서비스권역)에 거주하는 장래인구 변화에 따라 향후 해당 시설을 사용할 잠재력있는 배후수요가 변한다고 가정하였다.

국토연구원에서는 이보경 외(2020) 연구 결과를 바탕으로 누구나 장래의 인구 변화를 고려하여 시설계획을 세울 수 있도록 인터랙티브 리포트를 개발하여 공개하였다(〈그림 5-3〉 참조). 이보경 외(2020)의 연구결과와 국토연구원 인터랙티브 리포트를 활용하여 지자체는 지속가능한 생활SOC 운영을 위해 인구가 감소하는 지역에서는 소외인구를 최소화하면서 시설을 효율적으로 운영할 수 있는 입지를 선정할 수 있고, 인구가 증가하는 지역에서는 인구가 가장 많이 증가할 곳을 대상으로 신규시설 입지를 설정할 수 있다. 이와 같이 GPAI의 장래인구 예측값을 활용하면, 장래의 인구변화를 고려하여 수요 변화를 예상하고 시설의 배치 등과 같은 자원 배분의 의사결정을 좀 더 과학적으로 내릴 수 있을 것으로 기대된다.

그림 5-3 | 격자단위 장래인구예측을 활용한 부산 동구 작은도서관 폐점결정 시뮬레이션 예시

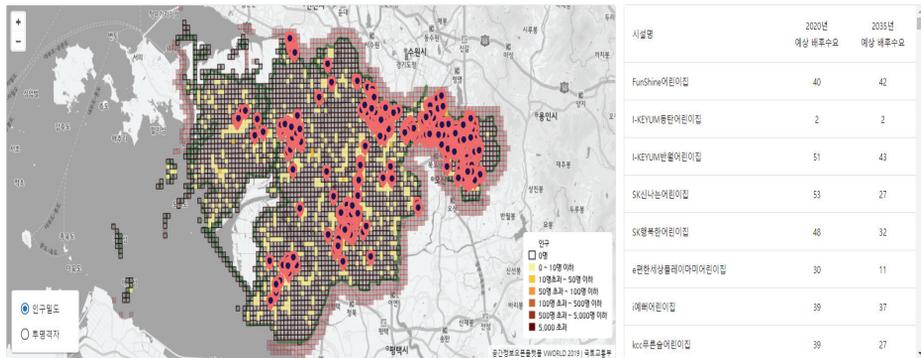
2040년까지 5년마다 미래 생활인프라 시설의 배후수요가 어떻게 변할지, 관심 지역과 시설(도서관과 어린이집)을 선택해서 지도로 확인해보세요.

* 배후수요는 시설부터 서비스권역 가장 내에 거주하는 인구 규모를 의미합니다.



경기도 화성시 어린이집 내 어린이집 현황과 배후수요 현황

경기도 화성시 내 어린이집 현황 (886개소)



자료: 이보경 외(2020)를 바탕으로 구축된 국토연구원 인터랙티브 리포트 캡처

(<https://interactive.krihs.re.kr/interactive/dnspopsimNew/index.html>, 최종 접속일자: 2022.12.30.)

3) 인구변화 대응을 위한 실효성있는 지자체 계획수립

지자체는 도시·군기본계획을 통해 20년 후 미래상을 정립하고 이를 위한 공간구조 설정 등의 공간계획을 수립한다. 이 과정에서 계획인구는 시가화예정 용지를 비롯한 각종 기반 시설의 용량을 결정하는데 근거로 활용되는 핵심적인 사항이다. 지금까지 지자체에서는 과부족에 대한 우려 및 지자체 사업에 대한 수요를 고려하여 최대한의 계획인구를 설정하여 계획을 수립해왔다. 물론, 과부족에 대한 우려는 충분히 인정되지만, 지자체의 계획인구 과대 추정 및 활용은 여러 가지 우려 및 부작용이 있는 것이 현실이다. 이와 같은 상황에서 지자체에서 격자단위 장래인구 정보를 활용하여 계획을 수립한다면, 인구변화에 대응할 수 있는 적절한 계획을 수립할 수 있을 것이다.

국가는 지자체의 계획을 검토하는 과정에서 정밀한 공간단위의 인구 기준을 활용하여 사회적 증가인구의 근거가 되는 개발사업 등의 입지에 따라 보다 효율적, 객관적으로 계획의 실효성을 확인할 수 있다. 현재 기본계획은 기본적으로 도에서 승인하며 도 승인 이전에 국토계획평가(중앙정부)를 수행하도록 하는데, 도 승인 과정 및 국토계획평가에서 지자체가 설정된 계획인구는 계획수립지침을 토대로 철저히 검증한다. 그러나 이미 확정된 개발사업의 경우에도 인구유입률 등의 측면에서 여전히 모호하게 계획인구가 설정되는 경우가 많고, 이를 검증하기 위한 객관적인 수단도 부재한 실정이다. 격자단위 장래인구를 활용하면 이와 같은 중앙정부 및 광역지자체의 시군단위 계획의 효율성과 지속가능성을 담보하기 위한 역할을 강화할 수 있을 것이다.

지자체 내부의 균형발전에 대한 수요가 전 국가적 균형발전에 대한 필요성만큼이나 절실하게 인식되고 있다. 격자단위 장래인구 예측 결과를 활용하면, 실제 지역 내부 인구과밀 혹은 과소 지역을 세부적으로 진단하고 향후 가능한 다양한 이슈에 대응할 수 있는 도시·군기본계획 수립이 가능해질 것으로 기대된다. 생활권 단위 인구 배분은 기반 시설의 배분과 직접적으로 관련 있다고 할 수 있기 때문에 지자체 차원에서는 계획 수립 이후 추후 실질적인 시설 건설이나 도입과 관련된 예산 마련 및 사업 우선 순위 선정에도 이러한 격자단위 인구 예측모형이 효과적으로 활용될 수 있다.

2. 활용 확산을 위한 정책 제언

1) 실무에서의 활용을 위한 정책 제언

격자와 같은 소지역단위 장래인구 예측 모형이 완벽하게 미래의 다양한 여건 변화를 반영하여 절대적으로 정확한 인구 규모를 추정할 수 없기 때문에 지역의 상황이나 맥락과는 상이한 결과가 나타날 가능성도 존재한다. 또한 직접 계획을 수립하는 지자체 입장에서는 계획 수립의 권한과 의지라는 지자체 고유의 기능을 제한하는데 역할을 할 수도 있다. 이와 같은 이슈는 근본적으로 다양한 지자체의 상황과 장래의 변화 시나리오를 모델에서 충분히 고려하지 못하여 지자체의 의사결정자가 수용하는 데 한계가 있기 때문이다. 따라서 단기적으로는 계획수립권한을 인정하는 현재의 제도적 테두리 내에서 가급적 지자체의 의지를 최대한 확보하는 차원에서 인구를 검토하되, 보조적 지표나 세부 조정을 위한 과학적인 증거로서 소지역단위 장래인구 예측 결과를 활용할 필요가 있다.

중장기적으로는 국가 및 통계청에서 공표하는 하향식 장래인구추계와 정합성을 가지면서 지자체의 다양한 상황과 의지를 상향식으로 수렴할 수 있는 소지역단위 장래인구 예측 과정이 마련되어야 할 것이다. 현재는 하향식으로 국가(통계청)에서 전국 또는 시도단위의 장래인구 추계를 공표하고 이를 지자체에서 받아들이고 있다. 그런데 이 과정에서 지자체의 개발 시나리오가 포함되고 있지 않다 보니, 실제 장래인구추계와 지자체의 다양한 정책 및 계획에 따른 인구변화가 일치되지 않는 경우가 많다. 특히 예측 대상 공간단위가 작아질수록 국가의 대규모 택지개발 또는 지자체의 토지이용 계획 등에 따라 인구분포가 달라질 가능성이 크다.

이를 위해서는 무엇보다 국가, 광역 및 기초 지자체의 관련 정책과 계획을 통합하여 소지역단위 장래인구 예측 과정을 이끌어갈 수 있는 권한과 역량을 갖춘 조직이 필요하다. 예를 들어, 미국 캘리포니아의 경우 재무국(DOF, State of California Department of Finance)은 상향식 의견 수렴을 통한 인구예측을 수행하여 주 전체의 계획과 예산의 기준 자료로 활용하였다(최성연, 2013: 130). 캘리포니아 재무국은 광역 및 카운티

단위의 인구 예측력을 높이고 지역의 특성을 반영하기 위해 카운티, COG(Council of Governments), MPO(Metropolitan Planning Organization) 등의 계획가 또는 인구학자의 의견을 바탕으로 지역의 특수한 상황을 예측치에 반영하는 과정을 거쳤다.

관련 권한과 역량을 갖춘 조직은 중앙부처, 통계청 또는 국책연구기관 등 공공기관에 설립할 수 있다. 예를 들어, 국토기본법 제25조의 2의 국토모니터링²⁾ 제도가 상향식 소지역단위 장래인구 예측과 관련 조직 운영의 법적 근거로 활용될 수 있다. 국토모니터링 제도는 국토의 변화상과 국토계획 및 국토정책에 대한 추진상황을 주기적 또는 수시로 점검할 수 있도록 기틀을 마련하였다. 이와 같은 제도와 역량을 바탕으로 국토모니터링 추진 조직이 소지역단위 장래인구 예측 모델을 지속적으로 개발하고 활용확산하는 중추적인 역할을 할 수 있을 것이다.

소지역단위 장래인구 예측을 담당하는 조직이 앞으로 수행해야 할 과제를 다음과 같다. 우선, 데이터 분석을 통해 각종 정책과 계획의 개입이 소지역단위 인구변화에 미치는 영향을 규명해야 한다. 그리고 인구변화에 영향을 미치는 국가 또는 지자체의 각종 정책 및 계획과 전문가 의견을 수렴한다. 셋째, 인구변화에 영향을 미치는 정책 및 계획이 시행되지 않았을 때를 포함하여 다양한 시나리오를 도출하고, 앞서 규명한 정책 및 계획의 개입과 인구변화 관계를 바탕으로 시나리오별 소지역단위 장래인구 예측 방법론을 개발한다³⁾. 넷째, 상향식 장래인구 예측 결과가 국가 및 통계청에서 공표하는 하향식 장래인구추계와 정합성을 가지고 활용될 수 있도록 조율한다. 다섯째, 지자체의 실무 담당자가 예측 결과를 쉽게 활용할 수 있도록 일본의 ‘장래인구·세대예측 도구 킷 매뉴얼’(<그림 2-11> 참조) 및 RESAS와 같은 도구를 개발하고 교육한다.

2) 제25조의2(국토모니터링의 추진 등) ① 국토교통부장관은 국토의 변화상과 국토계획 및 국토정책에 대한 추진상황을 주기적 또는 수시로 점검(이하 “국토모니터링”이라 한다)할 수 있다. ② 중앙행정기관의 장 및 지방자치단체의 장은 국토계획 및 국토정책을 수립할 때, 국토모니터링 결과를 반영하도록 노력하여야 한다. ③ 국토교통부장관은 체계적이고 효율적인 국토계획의 수립과 국토정책의 추진을 위하여 국토모니터링체계를 구축·운영할 수 있다. ④ 국토교통부장관은 국토모니터링체계를 구축·운영하기 위하여 필요한 경우 관계 기관에 자료제공을 요청할 수 있다. 이 경우 이를 요청받은 관계 기관은 정당한 사유가 없으면 이에 따라야 한다.

3) 국토교통과학기술진흥원(KAIA) 발주하고 국토연구원 국토모니터링연구센터에서 수행하는 미래국토 전망 및 의사결정 지원 시스템 기술 개발 기획'과제의 기술 개발 대상으로 이 연구에서 제안한 GPAI 고도화 방안이 포함될 수 있도록 협의 중

2) 연구성과 공개를 통한 활용 확산 전략

컴퓨터 기술과 ICT환경의 발달로 개인 단위에서 정보를 습득하고 활용해볼 기회는 날이 갈수록 늘어가고 있다. 예를 들어, 유튜브와 같은 동영상 공유 플랫폼에서 유튜브 환경에서 사용자 참여형 콘텐츠(User Created Contents, UCC) 등에 대한 제작과 공유, 확산이 보편화 되었다. 이와 같은 흐름은 비단 콘텐츠 소비에서만 그치는 것이 아니라, 정보의 습득과 공유, 활용에서도 발견할 수 있다. 2007년 공개된 깃허브(GitHub)는 대표적인 온라인 소스코드 보관 및 공유의 개발자 커뮤니티이다. 깃(Git)은 오픈소스로 개발된 파일을 관리하고 여러 사용자가 함께 작업하고 이력을 공유할 수 있는 시스템을 의미하며, 이 깃을 사용하는 프로젝트를 지원하는 온라인 웹 호스팅 서비스가 깃허브인 것이다(윤진혁·김영진, 2018). 이 깃허브에 참여하는 개인, 민간, 대학 및 연구소가 폭발적으로 증가⁴⁾함에 따라, 프로그래밍 기반의 다양한 분석모델의 오픈소스 R&D⁵⁾의 새로운 장이 열렸다.

오픈소스 R&D는 특히 인공지능 분야에서 깃허브와 같은 플랫폼을 통해 급격히 성장하고 있는데, 구글과 같은 강력한 산업 주체 및 연구소에서 연구성과를 공개함에 따라 생태계가 조성된 것이다. 인공지능 분야의 오픈소스 생태계는 사용자들의 인공지능의 활용 진입 장벽을 완화하였고, 다양한 외부 연구자가 연구성과를 검토하여 모델을 검증하고 발전 방향을 제시하며 후속 연구에 활용되는 등 큰 역할을 하고 있다. 특히 이 분야의 연구자들은 연구성과를 학술지에 투고하는 것뿐 아니라 아카이브(ArXiv)에 공개하여 성과를 검증는 추세이다. GitXiv는 아카이브에 공개한 논문과 깃허브에 공개한 소스코드를 모두 제공하는 플랫폼으로 연구자들의 성과를 평가하고, 개방된 연구성과가 후속 연구에 활용될 수 있도록 선순환 환경을 형성하고 있다. 예를 들어, 개와 고양이를 구별하는 인공지능 모델이 공개되고, 이를 다른 데이터에 학습하여 암세포를 진

4) 2022년 6월 기준, 공공 및 민간 참여기업을 포함하여 전 세계 8천 3백만 개발자(팀)와 2억 개의 커뮤니티(repositories)이 깃허브에 참여(<https://en.wikipedia.org/wiki/GitHub>, 2022년 12월 26일 검색)

5) 오픈소스 R&D는 오픈이노베이션의 방법으로 기업 및 연구소의 내부 자원을 공개하여 외부자원의 참여를 통해 제품을 개선 및 신제품을 개발하는 것을 의미(과학기술&ICT 정책·기술동향 140호, 2019)

단하는 모델로 발전하는 것과 같이 계속해서 모델을 고도화시킬 수 있는 바탕이 되는 것이다(과학기술&ICT 정책·기술동향 140호, 2019).

이 연구를 통해 개발한 인공지능 기반의 인구예측 기초모델인 GPAI의 경우 민간뿐만 아니라 중앙정부 및 지자체 정책 자료로도 그 활용도가 매우 높은 데 반해, 인공지능 모델의 아키텍처와 소스코드 같은 지식 공유·확산할 수 있는 환경은 조성되지 않은 상황이다. 다양한 원인 중 하나로, 국토연구원을 포함한 국책연구기관의 인공지능 분야 성과물은 대부분 개념도, 적용 결과 및 활용방안 등을 담은 보고서를 제작하는 데 초점이 맞추어져 있으며, 실제 인공지능 관련 연구의 가장 중요한 핵심인 모델링을 위한 소스코드에 대한 검증, 평가 및 재사용을 통한 확산에 대한 관심이 저조하기 때문으로 볼 수 있다. 데이터 기반의 과학적 의사결정의 중요성이 증가함에 따라 기존의 정성적 연구 결과뿐 아니라 정량적 분석에 따른 결과와 그 분석 방법론 역시 연구의 주요 성과물로서 공유되고, 향후 연구에 활용될 필요가 있다.

GPAI 개발에 활용된 데이터, 아키텍처 및 소스코드를 공개하여, 인구예측에 인공지능 기법을 활용하는데 진입장벽을 낮추고 모델의 고도화를 위한 오픈 이노베이션이 필요하다. 성과의 공유와 확산, 집단지성의 효과를 극대화하기 위해서는 주제와 목적에 따라 적절한 공유와 확산 도구를 활용할 필요가 있다(김대중 외, 2014). **이영주 외(2019)**에서 제시한 국토연구원 인터랙티브 리포트⁶⁾가 정보 등의 공유와 확산을 위해 최적화되어 있는데 반해, 인공지능 기법 등과 관련하여 고도화된 코드 등 정보의 지속 가능한 공유와 확산을 위해서는 깃허브, GitXiv와 같은 다른 형태의 플랫폼이 요구되는 바이다.

정량적 분석 및 인공지능 등 모델링을 개발하는 연구성과를 공개하는 방법으로 <그림 5-4>와 같이 데이터 및 소스코드를 온라인 플랫폼에 공개하고 링크를 포함한 정보를 보고서에 함께 수록하는 방안이 있다. Olen·Lehsten(2022)는 2100년까지 전 세계의 격자단위 인구 예측하기 위한 방법론, 활용 데이터 및 소스코드를 공개한 예시이다. 해당 사례의 공개 내용과 방법에 따라 본 연구의 결과를 공개하는 예시는 <표 5-1>과

6) 국토연구원 인터랙티브 리포트 (<http://interactive.krihs.re.kr/interactive>, 2022년 2월 26일 검색)

같다. 예측 결과 데이터와 소스코드의 접근방법의 경우 깃허브에 공개가 가능하나, 연구보고서 외 데이터 및 소스코드 성과의 공개 정책이 마련되지 않은 상태이다.

깃허브, GitXiv 또는 인터랙티브 리포트와 같은 국토연구원 자체 지식공유 플랫폼이 마련되면, 이 연구의 성과물인 데이터와 소스코드를 업로드하고 이에 대한 평가 의견과 활용 확산이 가능할 것이다. 예를 들어, 국토연구원의 인터랙티브 리포트와 같은 쌍방향 소통도로 공개되면서 연구성과가 단순히 일방향적인 보고서에서 그치는 것이 아니라 다양한 참여자들의 연구 결과를 탐색하고 정책의 인사이트를 얻는 데 활용되고 있다. 이처럼 인공지능 모델링을 포함한 정량분석에 활용된 데이터와 알고리즘을 공유하고 확산할 수 있는 라이선스 기준과 플랫폼을 마련해야 한다. 또한 내부적으로 연구자들이 활용 데이터와 코드를 공개하는 데 인센티브 부여 방식 등의 제도적으로 활성화 방안을 마련한다면, 연구성과의 품질 향상 및 활용 확산을 통한 향후 과제 수행 등의 선순환 생태계가 조성될 것으로 기대된다.

그림 5-4 | 연구성과(데이터 및 소스코드) 공개 예시

Subject	Global and Planetary Change
Specific subject area	Population projection
Type of data	Gridded dataset
How data were acquired	Simulated
Data format	Annual files of population density at 1 km resolution in Geotif format.
Parameters for data collection	The input data were selected to contain the necessary data to estimate the future population in relation to IPCC CMIP6 scenarios.
Description of data collection	Input data to model future population
Data source location	SSP population - https://tntcat.iiasa.ac.at/5spDb/ RCP urban fraction - https://luh.umd.edu/ WorldPop - Unconstrained global mosaics - https://www.worldpop.org/geodata/summary?id=24767 Roads - https://sedac.ciesin.columbia.edu/data/set/groads-global-roads-open-access-v1 Water bodies - https://sedac.ciesin.columbia.edu/data/set/grump-v1-national-admin-boundaries/data-download (created with national boundaries data) Country borders - http://thematicmapping.org/ Urban extent - https://www.researchgate.net/publication/339873537_MGUP_annual_global_2001_2018 .
Data accessibility [1]	Repository name: DataGURU Data identification number: https://doi.org/10.18161/global_popcount.201610
Code accessibility [2]	Direct URL to data: https://dataguru.lu.se/app#worldpop Repository name: GitHub Direct URL to code: https://github.com/niklasbokeolen/world_population

자료: Olen-Lehsten(2022), p.2, Specifications Table

표 5-1 | GPAI 개발 관련 데이터 및 알고리즘 공개 방안 예시

분류	내용
주제	2050년까지 대한민국 장래인구 예측
분야	인구추계 및 예측
데이터 유형	격자단위 공간통계
데이터 생성 방법	인공지능 및 시나리오 기반 분석 결과
데이터 포맷	2025년부터 2050년까지 5년단위의 전국 1km 격자단위 인구예측 결과
변수 선정 기준	인구변화에 영향을 주는 인구학적 변동요인 및 공간적 특성 요인을 선정
원천데이터 및 접근방법	<ul style="list-style-type: none"> • 사망률 출처: 통계청, 「인구동향조사」, 2021, 2022.11.06, 사망자수, 조사망률, 기대수명 링크: https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=INH_1B8000F_02&conn_path=l2, 최종접속일: 2022.12.1. • 출생률 출처: 통계청, 「인구동향조사」, 2021, 2022.11.06, 출생아수, 합계출산율, 자연증가 등 링크: https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=INH_1B8000F_01&conn_path=l2, 최종접속일: 2022.12.1. • 순이동률 출처: 통계청, 「국내인구이동통계」, 2022.09, 2022.11.06, 시군구별 이동자수 링크: https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1B26001_A01&conn_path=l2, 최종접속일: 2022.12.1. • 격자단위 인구 수 정의: 통계청 지리정보서비스(SGIS) 링크: https://sgis.kostat.go.kr/view/pss/openDataIntrncn, 최종접속일: 2022.12.1. • 장래인구변동요인 정의: 통계청, 「장래인구추계」, 2050, 2022.11.06, 장래 인구변동요인(출생, 사망, 인구이동)/시도 링크: https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1BPB004&conn_path=l2, 최종접속일: 2022.12.1. • 건축물 현황 출처: 국토교통부 건축행정시스템 세움터 표제부 링크: https://cloud.eais.go.kr/, 최종접속일: 2022.12.1. • 토양경사도 출처: 국립농업과학원, 흙토람 링크: http://soil.rda.go.kr/soil/index.jsp, 최종접속일: 2022.12.1.
결과 데이터 접근방법	향후 공개
소스코드 접근방법	향후 공개

자료: 저자 작성

3. 결론

인구정책은 인구에 급격한 변화가 나타나면 이 자체를 대상으로 삼아 현상을 완화(mitigation)하려는 정책과, 그 변화에 따라 어떻게 사회가 변할 것인지 전망하여 이에 적응(adaptation)하는 정책으로 구분할 수 있다(김한나 외 2020: 117). 우리나라는 급격한 인구감소와 고령화 사회에 대응하기 위해 2005년 저출산·고령사회기본법을 제정하고 2006년부터 제1차 저출산·고령사회기본계획(2006~2010)을 수립·추진해왔다. 지금까지(제1차~3차 기본계획- 2006~2020) 기본계획은 인구변화를 ‘완화’하기 위해 출산율을 높이는 데 초점을 맞춰왔다. 이와 같은 십수 년의 노력과 예산 투입에도 불구하고 우리나라 인구는 급변하고 있다. 따라서 최근 수립된 인구구조 변화와 대응 방안, 제4차 저출산·고령사회기본계획(2021~2025) 및 제3기 범부처 인구정책 TF는 인구 문제와 관련하여 ‘완화’ 기조에서 ‘적응’ 기조로 변화의 필요성을 공감하고 정책의 방향을 바꾸었다.

인구감소 시대에 ‘적응’하기 위해서는 계획(planning)이 수반되어야 한다(김한나 외 2020: 118). 그리고 계획수립에 가장 기초자료는 장래의 인구이다. 국토·지역·도시 관점에서도 인구감소 시대에 ‘적응’하기 위한 적절한 계획수립은 국토공간의 인구 분포 변화 전망에 기초한다. 국가 관점에서는 통계청에서 추계한 장래인구를 활용하여 다양한 정책과 계획의 시나리오를 구상하고 있으나, 국토공간계획에서 장래인구는 지자체 개발 근거를 마련하기 위한 수치일 뿐 계획의 과학적 근거로 작동하고 있지 않다.

과학적인 장래인구 정보에 기반한 계획수립의 가장 큰 걸림돌은 관련 정보의 부재이다. 이 연구에서는 2050년까지의 장래인구 변화를 격자단위로 예측할 수 있는 방법론, GPAI(Gridded Population Forecast with AI)를 개발하였다. GPAI의 가장 큰 특징 두가지는 다음과 같다. 첫째, 시계열 변화에 영향을 받지 않은 1km 격자단위의 장래인구 예측 방법론을 설계하였고, 이 과정에서 발생하는 격자망 내 타겟격자와 주변 격자간의 이웃효과를 고려하였다. 둘째, 복잡하고 대용량의 공간 빅데이터를 학습하여 장래인구 예측의 정확도를 높일 수 있는 인공지능 기법을 방법론과 인구구조의 특성을

잘 반영하는 코호트요인법 모델을 결합하였다. 인구학적 변동요인(인구규모, 사망률, 출생률)과 토지이용(건축물 밀도, 주거용 건축물 현황, 토지 경사도 등)의 변화를 복합적으로 고려할 수 있도록 두 모델을 각각 개발하여 최종 단계에서 앙상블을 통해 격자단위 장래인구 예측값을 제시하였다.

GPAI를 2000년부터 2020년까지 데이터에 학습하여 2050년까지 인구를 예측한 결과 전 국토를 기준으로 평균인구밀도는 2050년까지 감소할 것으로 전망되었다. 전국의 평균 수치인 인구밀도로 인구감소 패턴을 파악하기 어렵기 때문에 인구감소 및 무거주화 위험 공간의 비중을 살펴보았다. 2050년에는 2020년의 인구거주지역 기준으로 대부분의 공간(87%)이 인구 감소 또는 무거주화 위험에 노출될 것으로 전망된다. 전국토의 관점에서 2000년도와 비교하여 50년간의 인구 분포의 변화를 예측해보면, 인구밀도는 매우 집중된 지역 일부와 대부분의 인구희박지역으로 양극화될 것으로 나타났다. 2050년까지 인구변화 예측 결과와 격자의 입지(수도권, 광역시, 도내 시 지자체와 도내 군 지자체)를 고려하여 공간 유형을 구분하여 정주여건을 비교하였다. 이때 생활에 필수적인 초등학교와 약국까지의 이용 편의성(접근성)을 유형별로 비교하였는데, 인구증가 공간이 대체로 인구감소 및 무거주화 위험 공간에 비해 생활인프라까지 접근성이 짧아 이용이 편리한 것으로 나타났다. 특히 인구가 크게 감소할 것으로 예상되는 군 지역이라고 하더라도 정주여건이 월등히 양호한 공간이면, 인구가 증가할 것으로 나타났다.

이 연구에서 제시한 GPAI 초기모델을 바로 정책이나 계획에 적용하는데 여러 가지 이슈가 발생할 가능성이 있다. 인공지능 기법의 경우 데이터의 시계열 변동에 큰 영향을 받기 때문에 지금까지 인구증가 추세를 학습한 결과로 인구감소시대의 인구분포를 전망하는 데 한계가 있기 때문이다. 이 연구를 바탕으로 향후 관련된 연구가 지속적으로 수행될 수 있도록 GPAI 초기모델의 고도화 방안을 데이터 측면과 사회·환경 현상 측면에서 검토하고, 필요한 향후 과제를 제시하였다는 데 큰 의의가 있다(〈표 5-2〉 참조).

표 5-2 | GPAI 고도화 방안

구분	고도화 방안		
데이터 측면	인공지능 기법을 학습할 수 있는 시계열 데이터가 비교적 시점이 짧고 변동성이 큼	→	지역에 특화된 국지적 GPAI방법론을 개발하여 최대한 오차를 줄일 수 있도록 함
	인구학적 변동요인, 특히 인구이동에 대한 정보가 행정구역단위로만 공개되고 있음	→	격자단위의 인구이동관련 공간통계를 생산·공표할 수 있도록 관련 연구와 제도 마련
	토지이용 변화를 예측할 수 있는 공간적 특성 변수를 충분히 활용하지 못함	→	토지이용과 관련된 인문환경과 자연환경적 요소를 고려하여 격자단위의 토지이용 변화 및 거주가능지 분포를 예측하는 모델을 개발
사회·환경 현상 측면	지금까지 우리나라는 인구가 증가해왔으나, 현재를 기점으로 완전히 다른 인구감소 추세를 경험하게 될 것임	→	2008년(1억2,808만 명)을 정점으로 13년째 인구가 감소하고 있는 일본의 변화 추세를 함께 학습하는 방법론 개발
	택지개발에 따른 대규모 주택공급과 같은 중앙정부와 지자체 개입이 방법론에서 고려되지 못함	→	각종 중앙정부와 지자체의 주택공급에 따른 인구이동 효과를 반영할 수 있도록 방법론을 고도화
	기후변화에 따른 인구거주지역 변화 및 기후변화에 대응하기 위한 토지이용 변화와 같은 현상을 방법론에 반영하지 못함	→	기후변화에 따른 거주지 변화 시나리오를 고려하여 소지역단위의 장래인구 분포를 예측할 수 있는 방법론 개발

자료: 저자 작성

과학적인 근거를 바탕으로 인구감소시대에 적용할 수 있는 국토공간계획수립을 수립하는 데 가장 큰 걸림돌은 관련 정보의 부재인데, 이 연구에서는 격자단위 장래인구 예측 모델과 적용 결과가 기초자료로 활용될 가능성을 확인하였다. 다만, 여전히 모델을 실무에 바로 적용하는데 다양한 이슈가 있다. 따라서 이를 극복하고 실무에서 유용하게 활용되기 위해서는 앞서 일본 사례와 깃허브 등을 검토하여 제언한 활용 확산 전략과 후속 과제가 실행될 수 있도록 제도, 수행 조직 및 예산의 뒷받침이 필요하다. 향후 소지역단위 장래인구 예측 방법론이 의사결정에 활용할 수 있도록 고도화하고, 그 결과를 이해관계자들이 신뢰하고, 실무자가 쉽게 활용할 수 있도록 환경을 조성하기 위해서는 모든 부처의 요구사항을 조사하고 활용을 확산할 수 있는 역량과 권위가 있는 조직에서 이를 담당하여 추진해야 할 것이다.

참고문헌

REFERENCE



【 인용문헌 】

- 강영욱·장세진. 2008. 소지역 통계공표구역 설정지침 보완연구. 한국지도학회지, 8(2), 47-57.
- 경기도. 2019. 2017년 기준 경기도 시군별 장래인구특별추계 지침서
- 과학기술&ICT 정책·기술 동향 제140호. 2019. 과학기술정보통신부·한국과학기술평가원.
- 관계부처합동. 2021. 1. 27. 제3기 인구정책 TF 주요과제 및 추진 계획.
- 국가교통DB. 2018. 01. 2017년 제2권 전국 여객 O/D 전수화 및 장래수요예측. 한국교통연구원
- 국가균형발전 특별법 시행령. 2022. 대통령령 제32733호(2022. 6. 28., 타법개정)
- 국가균형발전 특별법. 2023. 대통령령 법률 제18877호(2022. 6. 10., 타법개정)
- 국토교통부. 2019. 제5차 국토종합계획(2020-2040). 국토교통부.
- 국토기본법. 2022. 법률 제18829호(2022. 2. 3., 일부개정)
- 국토지리정보원. 2021. 국토진단 및 모니터링을 위한 2021 국토조사
- 김경수·김상미. 2021. NABO 내국인 인구 시범추계: 2020~2040. 경제현안분석, 제101호.
국회예산정책처.
- 김기환. 국토연구원 개최 소지역단위 인구 전망 연구 동향 세미나(2022. 04. 15.) 발표자료
- 김대종·황명화·윤서연·서태성. 2014. 과학적 국토계획 수립을 위한 공간지식플랫폼 구축
및 활용방안 연구(I). 국토연구원.
- 김한나·고우림·임예진·정명구·조영태. 2020. 지방인구정책의 문제점과 대안. 한국인구학회,
43(4), 115-138.
- 김현중. 국토연구원 개최 소지역단위 인구 전망 연구 동향 세미나(2022. 04. 15.) 발표자료
- 김현중·여관현. 2019. 한국의 미래 인구분포 변화에 대한 예측. GRI 연구논총, 21(1), 301-328.
- 민보경. 2021. 인구충격에 대응하는 지역의 미래 전략 : 완화와 적응. 국회미래전략 insight, 34호.
- 민성희·변필성·김선희·차은혜·이철호·안용진. 2016. 국토계획 수립지원을 위한 인구분석

-
- 방법 연구. 국토연구원.
- 민성희·이순자·홍사흠·조정희·유현아. 2020. 국토 균형발전을 위한 중축적 권역 중심의 공간 전략. 국토정책Brief. 국토연구원.
- 손재선·장요한·오창화·임은선·황명화·이경주. 2021. 감염병 방역정책 지원을 위한 시공간 시뮬레이션 모형 개발 연구. 국토연구원.
- 심창섭·김오섭. 2019. 저출산·고령화를 고려한 폭염 노출위험인구 전망 및 지역별 대응전략. 한국환경연구원.
- 우해봉. 2009. 우리나라 인구추계의 정확성과 시사점. 조사연구. 10(2), 71-96.
- 우해봉·양지윤·조성호·안형석. 2016. 인구추계 방법론의 현황과 평가. 한국보건사회연구원.
- 윤진혁·김영진. 2018. GitHub 로 보는 오픈 이노베이션의 미래. 한국과학기술정보연구원.
- 이보경. 2019. 2040년 장래인구 분포 전망 연구. 국토연구원.
- 이보경·이영주·홍사흠·박미선·오창화. 2021. 주택공급에 따른 인구이동 분석모형 개발 및 활용방안 연구. 국토연구원.
- 이보경·황명화·박정은·이후빈·장요한·임룡혁·박진호·박서현·정도채·이민파. 2020. 미래예견적 국정관리지원: 생활인프라 수요-공급 예측 부문 모델링 연구. 세종: 경제인문사회연구원.
- 이석준·이상욱·홍보영·엄홍민·신휴석·김경민. 2014. 대시메트릭 매핑 기법을 이용한 서울시 건축물별 주거인구밀도의 재현. Spatial Information Research, 22(3), 89-99.
- 이영주·이보경·오창화. 2020. 국민활동 데이터를 이용한 개방형 지리인구통계 분석도구 개발 및 활용방안. 국토연구원.
- 이영주·이보경·오창화·이한진·김승범. 2019. 국토모니터링을 위한 인터랙티브 리포트 구축 연구. 국토연구원.
- 이영주·임은선·차미숙. 2015. 지역밀착형 정책지원을 위한 공간정보 융복합 활용방안 연구. 국토연구원.
- 이창효·이승일. 2006. 건축물 단위의 인구분포 추정. 국토계획, 41(6), 37-50.
- 이채정. 2021. 저출생·고령사회 심화에 따른 사회서비스 전달체계 개선 방향 검토: 아동

및 노인 대상 주요 사회서비스 시설의 분포 분석을 중심으로, 국가미래전략 Insight, 33호, 국회미래연구원.

인구보건복지협회. 2022. 「지방소멸 체감도 및 우선지원책」 설문조사, [http://www.ppfk.or.kr/sub/data/report_material.asp? mode=view&bid=9&s_type=&s_keywo rd=&s_cate=&idx=48729&page=1](http://www.ppfk.or.kr/sub/data/report_material.asp?mode=view&bid=9&s_type=&s_keywo rd=&s_cate=&idx=48729&page=1)

임은선 외. 2017. 빅데이터 시대의 국토정책 추진방향. 국토연구원.

저출산고령사회위원회·관계부처 합동, 2022. 12. 28. 인구구조 변화와 대응방안. [https://www.betterfuture.go.kr/front/notificationSpace/pressReleaseDetail.do;jsessionid=08E68063C144B0DFF8CC09538A677733.node10? articleId=184](https://www.betterfuture.go.kr/front/notificationSpace/pressReleaseDetail.do;jsessionid=08E68063C144B0DFF8CC09538A677733.node10?articleId=184)

제20대 대통령직인수위원회. 2022. 4. 27. 지역균형발전 비전 대국민 발표.

조만석·박태선·한우석·이상은·안종욱·서정섭·신정우. 2020. 인구감소·지방분권시대에 대응한 지방상수도 정책 개선방안 연구. 국토연구원.

최내영. 2010. 도시환경변수를 이용한 격자 인구추정에 있어서의 유전적 알고리즘기법 활용 연구. 한국지리정보학회지, 13(3), 119-130.

최성연 2013. chapter 경제불황과 지역인구의 예측, '한국과 미국의 광역화 추세와 지방정부간 협력: 대구경북과 남캘리포니아를 중심으로'. 대구경북연구원.

최현정·최석환·홍성조. 2019. 읍면동 단위 장래인구 추정모형 개발에 관한 연구. 부동산분석, 5(3), 67-87.

통계법. 2020. 법률 제17339호(2020. 6. 9. 타법개정)

통계청 보도자료. 2022. 07. 01. 소지역별 인구·사업체 등 공간통계 융합자료 서비스 확대, [https://kostat.go.kr/portal/korea/kor_nw/1/1/index.board? bmode=read&aSeq=419091](https://kostat.go.kr/portal/korea/kor_nw/1/1/index.board?bmode=read&aSeq=419091)

Anselin, L. 1988. A test for spatial autocorrelation in seemingly unrelated regressions. *Economics Letters*, 28(4), 335-341.

Baker, J., Ruan, X., Alcantara, A., Jones, T., Watkins, K., McDaniel, M., Frey,

-
- M., Crouse, N., Rajbhandari, R., Morehouse, J., Sanchez, J., Inglis, M., Baros, S., Penman, S., Morrison, S., Budge, T., & Stallcup, W. 2008. Density-dependence in urban housing unit growth: An evaluation of the Pearl-Reed model for predicting housing unit stock at the census tract level. *Journal of Economic and Social Measurement*, 33(2-3), 155-163.
- Baker, J., Swanson, D., Tayman, J. 2021. The accuracy of Hamilton-Perry population projections for census tracts in the United States. *Population Research and Policy Review*, 40(6), 1341-1354.
- Ballas, D., Clarke, G. P., Wiemers, E. 2005. Building a dynamic spatial microsimulation model for Ireland. *Population, Space and Place*, 11(3), 157-172.
- Breidenbach, P., Kaeding, M., Schaffner, S. 2019. Population projection for Germany 2015-2050 on grid level (RWI-GEO-GRID-POP-Forecast). *Jahrbücher für Nationalökonomie und Statistik*, 239(4), 733-745.
- Breiman, L. 2001. "Random Forests." *Machine Learning*, 45, 5-32.
- Calka, B., Bielecka, E. 2019. Reliability analysis of LandScan gridded population data. The case study of Poland. *ISPRS International Journal of Geo-Information*, 8(5), 222.
- Cassells, R., Harding, A., Miranti, R., Tanton, R., McNamara, J. 2010. Spatial microsimulation: Preparation of sample survey and census data for SpatialMSM/08 and SpatialMSM/09. Technical Paper No 36, Canberra: National Centre for Social and Economic Modelling (NATSEM) (April).
- Chen, Y., Guo, F., Wang, J., Cai, W., Wang, C., Wang, K. 2020a. Provincial and gridded population projection for China under shared socioeconomic pathways from 2010 to 2100. *Scientific Data*, 7(1), 1-13.
- Chen, Y., Li, X., Huang, K., Luo, M., Gao, M. 2020b. High-resolution gridded population projections for China under the shared socioeconomic pathways.

-
- Earth's Future, 8(6), e2020EF001491.
- Chi, G., Voss, P. R. 2011. Small-area population forecasting: Borrowing strength across space and time. *Population, Space and Place*, 17(5), 505-520.
- Chi, G., Zhou, X., Voss, P. R. 2011. Small-area population forecasting in an urban setting: A spatial regression approach. *Journal of Population Research*, 28(2-3), 185-201.
- Du, B., Hu, X., Sun, L., Liu, J., Qiao, Y., Lv, W. 2020. Traffic demand prediction based on dynamic transition convolutional neural network. *IEEE Transactions on Intelligent Transportation Systems*, 22(2), 1237-1247.
- Ford, A., Barr, S., Dawson, R., Virgo, J., Batty, M., Hall, J. 2019. A multi-scale urban integrated assessment framework for climate change studies: A flooding application. *Computers, Environment and Urban Systems*, 75, 229-243.
- Grossman, I., Bandara, K., Wilson, T., Kirley, M. 2022. Can machine learning improve small area population forecasts? A forecast combination approach. *Computers, Environment and Urban Systems*, 95, 101806.
- Grossman, I., Wilson, T., Temple, J. 2022. Forecasting small area populations with Long Short-Term Memory Networks.
- Hamilton, C. H., Perry, J. 1962. A short method for projecting population by age from one decennial census to another. *Social Forces*, 41(2), 163-170.
- Harding, A., Vidyattama, Y., Tanton, R. 2011. Demographic change and the needs-based planning of government services: Projecting small area populations using spatial microsimulation. *Journal of Population Research*, 28(2), 203-224.
- Hauer, M. E., Evans, J. M., Alexander, C. R. 2015. Sea-level rise and sub-county population projections in coastal Georgia. *Population and Environment*, 37(1), 44-62.
- Kim, H., Choi, J. 2011. A hybrid dasymetric mapping for population density surface

-
- using remote sensing data. *Journal of the Korean Geographical Society*, 46(1), 67–80.
- Kim, K. W., Kim, O. S. 2020. Super aging in South Korea unstoppable but mitigatable: a sub-national scale population projection for best policy planning. *Spatial demography*, 8(2), 155–173.
- LeSage, J., Pace, R. K. 2009. *Introduction to spatial econometrics*. Chapman and Hall/CRC.
- Leyk, S., Gaughan, A. E., Adamo, S. B., de Sherbinin, A., Balk, D., Freire, S., ... Pesaresi, M. 2019. The spatial allocation of population: a review of large-scale gridded population data products and their fitness for use. *Earth System Science Data*, 11(3), 1385–1409.
- Liang, X., Wang, G., Min, M. R., Qi, Y., Han, Z. 2019 May. A deep spatio-temporal fuzzy neural network for passenger demand prediction. In *Proceedings of the 2019 SIAM international conference on data mining* (pp. 100–108). Society for Industrial and Applied Mathematics.
- Lin, H., Bai, R., Jia, W., Yang, X., & You, Y. 2020 August. Preserving dynamic attention for long-term spatial-temporal prediction. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 36–46).
- Lloyd, C. T., Sorichetta, A., Tatem, A. J. 2017. High resolution global gridded data for use in population studies. *Scientific data*, 4(1), 1–17.
- McKee, J. J., Rose, A. N., Bright, E. A., Huynh, T., Bhaduri, B. L. 2015. Locally adaptive, spatially explicit projection of US population for 2030 and 2050. *Proceedings of the National Academy of Sciences*, 112(5), 1344–1349.
- Murdock, S. H., Hamm, R. R., Voss, P. R., Fannin, D., & Pecotte, B. 1991. Evaluating small-area population projections. *Journal of the American Planning*

-
- Association, 57(4), 432–443.
- Olen, N. B., Lehsten, V. 2022. High-resolution global population projections dataset developed with CMIP6 RCP and SSP scenarios for year 2010–2100. *Data in Brief*, 40, 107804.
- Preston, S. H., Heuveline, P., Guillot, M. 2001. *Demography: Measuring and Modeling Population Processes*. MA: Blackwell Publishing.
- Rayer, S. 2008. Population forecast errors: A primer for planners, *Journal of Planning Education and Research*, 27, 417–430.
- Riiman, V., Wilson, A., Milewicz, R., Pirkelbauer, P. 2019. Comparing artificial neural network and cohort-component models for population forecasts. *Population Review*, 58(2).
- Smith, S. K., & Morrison, P. A. 2005. Small-area and business demography. *Handbook of population*, 761–785.
- Song, J., Han, K., Stettler, M. E. 2020. Deep-MAPS: Machine-Learning-Based Mobile Air Pollution Sensing. *IEEE Internet of Things Journal*, 8(9), 7649–7660.
- Swanson, D. A., Schlottmann, A., Schmidt, B. 2010. Forecasting the population of census tracts by age and sex: An example of the Hamilton-Perry method in action. *Population Research and Policy Review*, 29(1), 47–63.
- Tobler, W. R. 1979. Cellular geography. In *Philosophy in geography* (pp. 379–386). Springer, Dordrecht.
- Triantakonstantis, D., Mountrakis, G. 2012. Urban growth prediction: a review of computational models and human perceptions.
- Weber, H. 2020. How well can the migration component of regional population change be predicted. A machine learning approach applied to German municipalities. *Comparative Population Studies*.
- Wilson, T., Grossman, I., Alexander, M., Rees, P., & Temple, J. 2022. *Methods*

for small area population forecasts: State-of-the-art and research needs. *Population Research and Policy Review*, 41(3), 865–898.

Yao, H., Liu, Y., Wei, Y., Tang, X., & Li, Z. 2019, May. Learning from multiple cities: A meta-learning approach for spatial-temporal prediction. In *The World Wide Web Conference* (pp. 2181–2191).

Yao, H., Wu, F., Ke, J., Tang, X., Jia, Y., Lu, S., ... Li, Z. 2018, April. Deep multi-view spatial-temporal network for taxi demand prediction. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 32, No. 1).

【 홈페이지 】

- 건축물 현황. 국토교통부 건축행정시스템 세움터 표제부(<https://cloud.eais.go.kr/>, 2022. 12. 1. 접속)
- 격자단위 인구수. 통계청 지리정보서비스(SGIS)(<https://sgis.kostat.go.kr/view/pss/openDataIntrcn>, 2022. 12. 1. 접속)
- 광역지자체 공표 장래인구추계(경기도: https://kosis.kr/statHtml/statHtml.do?orgId=210&tblId=DT_210008_2017001&conn_path=I2, 2022. 11. 18. 접속)
- 광역지자체 공표 장래인구추계(전라남도: https://kosis.kr/statHtml/statHtml.do?orgId=215&tblId=DT_215007_001&conn_path=I2, 2022. 11. 18. 접속)
- 국토연구원 인터랙티브 리포트 (<http://interactive.krihs.re.kr/interactive>, 2022. 2. 26. 접속)
- 국토지리정보원(map.ngii.go.kr, 최종접속일: 2022. 11. 05. 접속)
- 그린피스 웹사이트(<http://www.climate.or.kr/>, 2022. 11. 05. 접속)
- 그린피스 웹사이트(<https://www.greenpeace.org/korea/update/14990/blog-ce-flood-web-page/>, 2022. 11. 05. 접속)
- 사망률. 통계청, 「인구동향조사」, 2021, 2022. 11. 06, 사망자수, 조사망률, 기대수명 (https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=INH_1B8000F_02&conn_path=I2, 2022. 12. 1. 접속)
- 순이동률. 통계청, 「국내인구이동통계」, 2022. 09, 2022. 11. 06, 시군구별 이동자수 (https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1B26001_A01&conn_path=I2, 2022. 12. 1. 접속)
- 이보경 외(2020)를 바탕으로 구축된 국토연구원 인터랙티브 리포트 캡처(<https://interactive.krihs.re.kr/interactive/dnspsimNew/index.html>, 2022. 12. 30. 접속)
- 인구총조사(통계청) (https://kostat.go.kr/portal/korea/kor_ip/6/3/1/1/index.board, 2022. 11. 03. 접속)
- 일본국토교통성(国土交通省) (<https://www.mlit.go.jp/common/001135250.pdf>, 2022. 11. 21. 접속)

일본국토교통성(国土交通省)https://www.mlit.go.jp/kokudoseisaku/kokudoseisaku_tk3_000086.html(f, 2022. 11. 21. 접속)

장래인구변동요인. 통계청, 「장래인구추계」, 2050, 2022. 11. 06, 장래 인구변동요인 (출생, 사망, 인구이동)/시도([https://kosis.kr/statHtml/statHtml.do? orgId=101&tblId=DT_1BFPB004&conn_path=I2](https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1BFPB004&conn_path=I2), 2022. 12. 1. 접속)

출생률. 통계청, 「인구동향조사」, 2021, 2022. 11. 06, 출생아수, 합계출산율, 자연증가 등 ([https://kosis.kr/statHtml/statHtml.do? orgId=101&tblId=INH_1B8000F_01&conn_path=I2](https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=INH_1B8000F_01&conn_path=I2), 2022. 12. 1. 접속)

택지정보시스템(<https://jigu.go.kr/index.do>, 2022. 11. 05. 접속)

토양경사도, 국립농업과학원, 흙도람(<http://soil.rda.go.kr/soil/index.jsp>, 2022. 12. 1. 접속)

통계지리정보서비스(SGIS) (<https://sgis.kostat.go.kr/>, 2022. 11. 22. 접속)

통계청 장래인구추계, 인구동향조사, 인구총조사 및 국내인구이동통계 통계설명자료 (<https://kosis.kr/>, 2022. 06. 23. 접속)

GitHub(<https://en.wikipedia.org/wiki/GitHub>, 2022. 12. 26. 검색)

NILIM 2020(<http://www.nilim.go.jp/lab/bcg/siryou/2020report/ar2020hp.pdf>, 2022. 11. 21. 접속)

SUMMARY



A Study on Small Area Population Forecast Model with Artificial Intelligence

Bo Kyeong Lee, Yo Han Jang, Yae Jin Jung, Sa Heum Hong, Dong Geun Kim

Key words: Population Prediction, Artificial Intelligence, Small Area, Population-Declining Era

South Korea is currently in an era with a declining population and faces the heightened risk of extirpation. While most citizens (88%) are already aware of the risk, the level of risk varies widely in different areas of the national territory. Most policymakers and planners have started working on developing strategies to help solve this problem; however, they struggle with the existing frameworks of the nation or administrative district, which limits the development of place-based strategies that could be substantiated. To develop such practical policies in the era of the declining population and implement them, a finer unit of a small area is necessary rather than a bigger unit of an administrative district. Furthermore, the data regarding potential changes in the future population should be collected rather than the retrospective data covering the past and the present.

The purpose of this study is to develop a population prediction model for small-area units through the use of the artificial intelligence (AI) technique and to find ways to forecast the population in small areas at the national, local, and

urban levels of policy-making and planning. First, previous studies on developing a way to predict the future population in small-area units, which is an unfamiliar concept in South Korea, were reviewed to identify the advantages and limitations of each methodology and set the direction of development. Second, we develop a 1km grid level population prediction prediction method, called Gridded Population Forecast with AI (GPAI) by applying the AI technique and cohort component method to comprehensive temporal-spatial data containing the demographic and spatial characteristics. Third, the GPAI model is applied to the entire nation to forecast the changes in population distribution up to 2050 and to categorize and comparatively analyze the population variation in areas of depopulation and those with a high risk of being unpopulated. Lastly, a way to advance the GPAI toward a high level of prediction accuracy through continuous self-supervised learning and to simultaneously use the model on the political frontier is proposed.

The GPAI architecture is a comprehensive model used to predict the population distribution reflecting both demographic and spatial characteristics at 1km grid level. These characteristics include demographic variables such as birth rate, death rate, and population migration and spatial characteristics such as land use, conventionally analyzed in population forecast models. The GPAI consists two models; Model I is an AI-based model incorporating demographic and spatial characteristics, and Model II is a model based on the cohort component method with a focus on the changes induced by demographic characteristics. The final results of model prediction on future populations in the grids are combined through the Bagging of Ensemble method to present the predictive values on future populations.

Considering the South Korean population density in 2020 as 100, the future population predicted using the GPAI showed a reduction in population density

to 91 by 2050, and the percentage of areas with a high risk of being unpopulated and areas of depopulation was predicted to be as high as 87%. Compared to in the present (2020), in 2050, the capital area, the northern parts of Chungcheongnam-do, Sejong-si, and Daejeon-si, and the satellite area of certain metropolitan cities shows increased population, whereas depopulation would be exhibited by most parts of the national territory. Each grid is categorized by integrating the characteristics of the local government in the grid (that of the capital and metropolitan areas and cities and provinces) and the patterns of population variation (depopulation, risk of being unpopulated, and increased population). In addition, the settlement conditions at present (2020) were compared in terms of infrastructure accessibility regarding livelihood. In the capital and metropolitan areas, the population density turns out higher in areas of depopulation than in areas of increased population, displaying a pattern opposite to that of provinces with markedly low population density in areas of depopulation than in areas of increased population. Furthermore, in grids predicted to show increased population by the local governments of cities and provinces, compared to the capital area, the infrastructure condition is more favorable. This indicates that even in a province predicted to show a decline in total population at administrative unit, the population could still increase through favorable settlement conditions at small area units.

The biggest obstacle in establishing national territorial plans using scientific evidence toward the adaptation in the era of the declining population is the lack of necessary data, and the output of the prediction model in this study on future populations in small-area units could serve as basic data. To promote follow-up studies and the development of strategies toward the wider use of the model proposed in this study, relevant systems, implementing organizations and budgets should be taken into serious consideration.



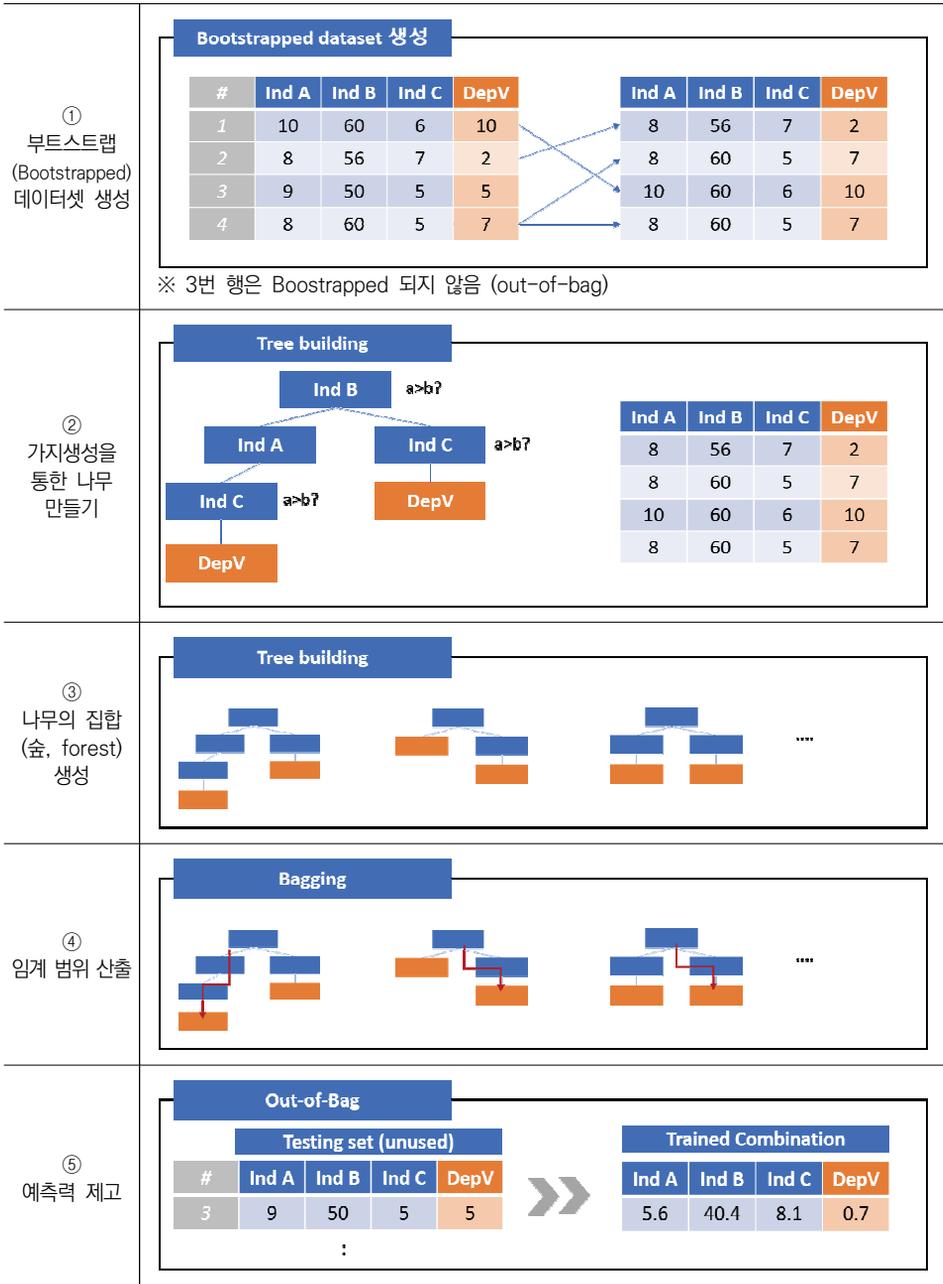
▣ 부록 1. 랜덤포레스트(random forest) 모형

- 랜덤포레스트(random forest) 알고리즘은 2001년 Breiman에 의해 처음 제안된 이래, 높은 예측성과 정확성 등 강점 때문에 공학, 통계학, 인문학 등 다양한 학문 분야에 널리 활용 중인 알고리즘 중 하나임
 - 단일 모형 기반의 예측에서 발생할 수 있는 예측 결과 편향성과 오차 위험성을 다수 모형을 종합하여 평균값으로 치환하는 형태인 앙상블(ensemble) 접근 방법을 통해 극복하여 예측력 제고뿐만 아니라 오차율 최소화에도 효과적으로 알려져 있음

- 랜덤포레스트의 높은 예측력은 주어진 데이터로부터 무작위 추출(random sampling) 후, 독립변수와 종속변수 간의 관계를 단순하고 반복적으로 접근하여 그 관계를 파악하는 방식에서 두드러짐
 - 방대한 양의 데이터를 한꺼번에 분석하고 파악할 때 요구되는 분석 및 계산 부담을, 랜덤하게 선택되는 데이터 간의 관계를 반복 학습하여 각 독립변수에 따른 종속변수의 특성을 추출하고, 더 나아가서는 독립변수의 증감 및 변형에 따른 종속변수의 영향력(importance)과 민감도까지 도출할 수 있는 특징이 있음
 - 때문에, 지도학습(supervised learning) 군에서 빈번하게 발생하는 오차 편향성에 강하고(robust), 데이터 모수를 활용하는 오프라인학습(offline learning, 또는 batch learning) 군에서 취약한 대규모 데이터 처리에서 오는 부담을 완화하는데 최적의 모형 중 하나로 분류되기도 함

-
- 랜덤포레스트의 계산순서는 다음과 같음(<그림 부1-1> 참조)
 - 먼저, 무작위 추출(중복 추출포함)을 통한 독립적 샘플데이터(bootstrapped data)를 특정 비율에 따라 모수로부터 추출함. 해당 데이터는 모형의 개발과 훈련에 필요한 데이터로, 이 단계에서 추출되지 않은 나머지 데이터(out-of-bag)는 훈련 완료 후, 검증(validation)의 단계에 활용됨
 - 두 번째 단계에서는, 샘플데이터의 일련의 독립변수를 무작위로 선택하고 종속 변수와의 관계를 파악함
 - 세 번째 단계에서는 앞선 단계에서 파악된 무작위 선택을 중복하여 각 독립변수와 종속변수 간의 관계를 일반화하는 단계인데, 이 두 단계가 마치 가지가 뻗어 나가는 단계와 유사하다 하여 가지와 나무 생성(Tree building) 단계라고도 불림
 - 도출된 임계값을 바탕으로 각 독립변수와 종속변수 간의 관계가 파악되면, 각 독립변수에 다양한 범위의 오차율을 적용하여 임계범위를 산출하고, 이 단계에서 파악된 임계범위로부터 각 변수의 중요도를 추정함.
 - 마지막 단계에서는, 검증데이터를 이용하여 모형의 완성도를 높임

그림 부1-1 | 랜덤포레스트 모형의 개념적인 전개



출처: Breiman 2001의 내용을 손재선 외. (2021) p.118 그림 4-3에서 재인용

▣ 부록 2. GPAI 모델 1 입력 변수

표 부2-1 | GPAI 입력변수 및 주요 내용

NO	변수 그룹	모델의 변수명	설명
1	목표 격자 ID	grid_ID	목표 격자 ID
2	목표 격자가 속한 시도단위의 인구	POPSido_2000	목표 격자가 속한 시도단위의 2000년 집계 인구
3		POPSido_2005	목표 격자가 속한 시도단위의 2005년 집계 인구
4		POPSido_2010	목표 격자가 속한 시도단위의 2010년 집계 인구
5		POPSido_2015	목표 격자가 속한 시도단위의 2015년 집계 인구
6		POPSido_2020	목표 격자가 속한 시도단위의 2020년 집계 인구
7	목표 격자가 속한 시도단위의 사망률	DeathSido_2000	목표 격자가 속한 시도단위의 2000년 집계/추계 사망률
8		DeathSido_2005	목표 격자가 속한 시도단위의 2005년 집계/추계 사망률
9		DeathSido_2010	목표 격자가 속한 시도단위의 2010년 집계/추계 사망률
10		DeathSido_2015	목표 격자가 속한 시도단위의 2015년 집계/추계 사망률
11		DeathSido_2020	목표 격자가 속한 시도단위의 2020년 집계/추계 사망률
12	목표 격자가 속한 시도단위의 순이동률	MoveSido_2000	목표 격자가 속한 시도단위의 2000년 집계/추계 순이동률
13		MoveSido_2005	목표 격자가 속한 시도단위의 2005년 집계/추계 순이동률
14		MoveSido_2010	목표 격자가 속한 시도단위의 2010년 집계/추계 순이동률
15		MoveSido_2015	목표 격자가 속한 시도단위의 2015년 집계/추계 순이동률
16		MoveSido_2020	목표 격자가 속한 시도단위의 2020년 집계/추계 순이동률
17	목표 격자가 속한 시도단위의 출생률	BirthSido_2000	목표 격자가 속한 시도단위의 2000년 집계/추계 출생률
18		BirthSido_2005	목표 격자가 속한 시도단위의 2005년 집계/추계 출생률
19		BirthSido_2010	목표 격자가 속한 시도단위의 2010년 집계/추계 출생률
20		BirthSido_2015	목표 격자가 속한 시도단위의 2015년 집계/추계 출생률
21		BirthSido_2020	목표 격자가 속한 시도단위의 2020년 집계/추계 출생률
22	목표 격자와 인접한 격자의 과거시점 인구정보	HisPop_gLT_2000	목표 격자와 인접한 격자(왼쪽위)의 2000년 인구 정보
23		HisPop_gLT_2005	목표 격자와 인접한 격자(왼쪽위)의 2005년 인구 정보
24		HisPop_gLT_2010	목표 격자와 인접한 격자(왼쪽위)의 2010년 인구 정보
25		HisPop_gLT_2015	목표 격자와 인접한 격자(왼쪽위)의 2015년 인구 정보
26		HisPop_gCT_2000	목표 격자와 인접한 격자(중간위)의 2000년 인구 정보
27		HisPop_gCT_2005	목표 격자와 인접한 격자(중간위)의 2005년 인구 정보

NO	변수 그룹	모델의 변수명	설명
28		HisPop_gCT_2010	목표 격자와 인접한 격자(중간위)의 2010년 인구 정보
29		HisPop_gCT_2015	목표 격자와 인접한 격자(중간위)의 2015년 인구 정보
30		HisPop_gRT_2000	목표 격자와 인접한 격자(오른쪽위)의 2000년 인구 정보
31		HisPop_gRT_2005	목표 격자와 인접한 격자(오른쪽위)의 2005년 인구 정보
32		HisPop_gRT_2010	목표 격자와 인접한 격자(오른쪽위)의 2010년 인구 정보
33		HisPop_gRT_2015	목표 격자와 인접한 격자(오른쪽위)의 2015년 인구 정보
34		HisPop_gLM_2000	목표 격자와 인접한 격자(왼쪽중간)의 2000년 인구 정보
35		HisPop_gLM_2005	목표 격자와 인접한 격자(왼쪽중간)의 2005년 인구 정보
36		HisPop_gLM_2010	목표 격자와 인접한 격자(왼쪽중간)의 2010년 인구 정보
37		HisPop_gLM_2015	목표 격자와 인접한 격자(왼쪽중간)의 2015년 인구 정보
38		HisPop_gTg_2000	목표 격자(가운데중간)의 2000년 인구 정보
39		HisPop_gTg_2005	목표 격자(가운데중간)의 2005년 인구 정보
40		HisPop_gTg_2010	목표 격자(가운데중간)의 2010년 인구 정보
41		HisPop_gTg_2015	목표 격자(가운데중간)의 2015년 인구 정보
42		HisPop_gRM_2000	목표 격자와 인접한 격자(오른쪽중간)의 2000년 인구 정보
43		HisPop_gRM_2005	목표 격자와 인접한 격자(오른쪽중간)의 2005년 인구 정보
44		HisPop_gRM_2010	목표 격자와 인접한 격자(오른쪽중간)의 2010년 인구 정보
45		HisPop_gRM_2015	목표 격자와 인접한 격자(오른쪽중간)의 2015년 인구 정보
46		HisPop_gLB_2000	목표 격자와 인접한 격자(왼쪽아래)의 2000년 인구 정보
47		HisPop_gLB_2005	목표 격자와 인접한 격자(왼쪽아래)의 2005년 인구 정보
48		HisPop_gLB_2010	목표 격자와 인접한 격자(왼쪽아래)의 2010년 인구 정보
49		HisPop_gLB_2015	목표 격자와 인접한 격자(왼쪽아래)의 2015년 인구 정보
50		HisPop_gCB_2000	목표 격자와 인접한 격자(중간아래)의 2000년 인구 정보
51		HisPop_gCB_2005	목표 격자와 인접한 격자(중간아래)의 2005년 인구 정보
52		HisPop_gCB_2010	목표 격자와 인접한 격자(중간아래)의 2010년 인구 정보
53		HisPop_gCB_2015	목표 격자와 인접한 격자(중간아래)의 2015년 인구 정보
54		HisPop_gRB_2000	목표 격자와 인접한 격자(오른쪽아래)의 2000년 인구 정보
55		HisPop_gRB_2005	목표 격자와 인접한 격자(오른쪽아래)의 2005년 인구 정보
56		HisPop_gRB_2010	목표 격자와 인접한 격자(오른쪽아래)의 2010년 인구 정보
57		HisPop_gRB_2015	목표 격자와 인접한 격자(오른쪽아래)의 2015년 인구 정보
58	목표 격자와	Grd_gLC_ud2	목표 격자와 인접한 격자(왼쪽위)의 경사도 2% 미만 비율

NO	변수 그룹	모델의 변수명	설명
59	목표 격자와 인접한 격자의 경사도 정보	Grd_gCT_ud2	목표 격자와 인접한 격자(왼쪽중간)의 경사도 2% 미만 비율
60		Grd_gRT_ud2	목표 격자와 인접한 격자(오른쪽위)의 경사도 2% 미만 비율
61		Grd_gLM_ud2	목표 격자와 인접한 격자(왼쪽중간)의 경사도 2% 미만 비율
62		Grd_gTg_ud2	목표 격자의 경사도 2% 미만 비율
63		Grd_gRM_ud2	목표 격자와 인접한 격자(오른쪽중간)의 경사도 2% 미만 비율
64		Grd_gLB_ud2	목표 격자와 인접한 격자(왼쪽아래)의 경사도 2% 미만 비율
65		Grd_gCB_ud2	목표 격자와 인접한 격자(중간아래)의 경사도 2% 미만 비율
66		Grd_gRB_ud2	목표 격자와 인접한 격자(왼쪽아래)의 경사도 2% 미만 비율
67		Grd_gLC_ud7	목표 격자와 인접한 격자(왼쪽위)의 경사도 2% ~ 7% 비율
68		Grd_gCT_ud7	목표 격자와 인접한 격자(왼쪽중간)의 경사도 2% ~ 7% 비율
69		Grd_gRT_ud7	목표 격자와 인접한 격자(오른쪽위)의 경사도 2% ~ 7% 비율
70		Grd_gLM_ud7	목표 격자와 인접한 격자(왼쪽중간)의 경사도 2% ~ 7% 비율
71		Grd_gTg_ud7	목표 격자의 경사도 2% ~ 7% 미만 비율
72		Grd_gRM_ud7	목표 격자와 인접한 격자(오른쪽중간)의 경사도 2% ~ 7% 비율
73		Grd_gLB_ud7	목표 격자와 인접한 격자(왼쪽아래)의 경사도 2% ~ 7% 비율
74		Grd_gCB_ud7	목표 격자와 인접한 격자(중간아래)의 경사도 2% ~ 7% 비율
75		Grd_gRB_ud7	목표 격자와 인접한 격자(왼쪽아래)의 경사도 2% ~ 7% 비율
76		Grd_gLC_ud15	목표 격자와 인접한 격자(왼쪽위)의 경사도 7% ~ 15% 비율
77		Grd_gCT_ud15	목표 격자와 인접한 격자(왼쪽중간)의 경사도 7% ~ 15% 비율
78		Grd_gRT_ud15	목표 격자와 인접한 격자(오른쪽위)의 경사도 7% ~ 15% 비율
79		Grd_gLM_ud15	목표 격자와 인접한 격자(왼쪽중간)의 경사도 7% ~ 15% 비율
80		Grd_gTg_ud15	목표 격자의 경사도 7% ~ 15% 비율
81		Grd_gRM_ud15	목표 격자와 인접한 격자(오른쪽중간)의 경사도 7% ~ 15% 비율
82		Grd_gLB_ud15	목표 격자와 인접한 격자(왼쪽아래)의 경사도 7% ~ 15% 비율
83		Grd_gCB_ud15	목표 격자와 인접한 격자(중간아래)의 경사도 7% ~ 15% 비율
84		Grd_gRB_ud15	목표 격자와 인접한 격자(왼쪽아래)의 경사도 7% ~ 15% 비율
85		Grd_gLC_ud30	목표 격자와 인접한 격자(왼쪽위)의 경사도 15% ~ 30% 비율
86		Grd_gCT_ud30	목표 격자와 인접한 격자(왼쪽중간)의 경사도 15% ~ 30% 비율
87		Grd_gRT_ud30	목표 격자와 인접한 격자(오른쪽위)의 경사도 15% ~ 30% 비율
88		Grd_gLM_ud30	목표 격자와 인접한 격자(왼쪽중간)의 경사도 15% ~ 30% 비율
89		Grd_gTg_ud30	목표 격자의 경사도 15% ~ 30% 비율

NO	변수 그룹	모델의 변수명	설명
90		Grd_gRM_ud30	목표 격자와 인접한 격자(오른쪽중간)의 경사도 15% ~ 30% 비율
91		Grd_gLB_ud30	목표 격자와 인접한 격자(왼쪽아래)의 경사도 15% ~ 30% 비율
92		Grd_gCB_ud30	목표 격자와 인접한 격자(중간아래)의 경사도 15% ~ 30% 비율
93		Grd_gRB_ud30	목표 격자와 인접한 격자(왼쪽아래)의 경사도 15% ~ 30% 비율
94		Grd_gLC_ud60	목표 격자와 인접한 격자(왼쪽위)의 경사도 30% ~ 60% 비율
95		Grd_gCT_ud60	목표 격자와 인접한 격자(왼쪽중간)의 경사도 30% ~ 60% 비율
96		Grd_gRT_ud60	목표 격자와 인접한 격자(오른쪽위)의 경사도 30% ~ 60% 비율
97		Grd_gLM_ud60	목표 격자와 인접한 격자(왼쪽중간)의 경사도 30% ~ 60% 비율
98		Grd_gTg_ud60	목표 격자의 경사도 30% ~ 60% 비율
99		Grd_gRM_ud60	목표 격자와 인접한 격자(오른쪽중간)의 경사도 30% ~ 60% 비율
100		Grd_gLB_ud60	목표 격자와 인접한 격자(왼쪽아래)의 경사도 30% ~ 60% 비율
101		Grd_gCB_ud60	목표 격자와 인접한 격자(중간아래)의 경사도 30% ~ 60% 비율
102		Grd_gRB_ud60	목표 격자와 인접한 격자(왼쪽아래)의 경사도 30% ~ 60% 비율
103		Grd_gLC_ud100	목표 격자와 인접한 격자(왼쪽위)의 경사도 60% ~ 100% 비율
104		Grd_gCT_ud100	목표 격자와 인접한 격자(왼쪽중간)의 경사도 60% ~ 100% 비율
105		Grd_gRT_ud100	목표 격자와 인접한 격자(오른쪽위)의 경사도 60% ~ 100% 비율
106		Grd_gLM_ud100	목표 격자와 인접한 격자(왼쪽중간)의 경사도 60% ~ 100% 비율
107		Grd_gTg_ud100	목표 격자의 경사도 60% ~ 100% 비율
108		Grd_gRM_ud100	목표 격자와 인접한 격자(오른쪽중간)의 경사도 60% ~ 100% 비율
109		Grd_gLB_ud100	목표 격자와 인접한 격자(왼쪽아래)의 경사도 60% ~ 100% 비율
110		Grd_gCB_ud100	목표 격자와 인접한 격자(중간아래)의 경사도 60% ~ 100% 비율
111		Grd_gRB_ud100	목표 격자와 인접한 격자(왼쪽아래)의 경사도 60% ~ 100% 비율
112		Grd_gLC_other	목표 격자와 인접한 격자(왼쪽위)의 기타 경사도 비율
113		Grd_gCT_other	목표 격자와 인접한 격자(왼쪽중간)의 기타 경사도 비율
114		Grd_gRT_other	목표 격자와 인접한 격자(오른쪽위)의 기타 경사도 비율
115		Grd_gLM_other	목표 격자와 인접한 격자(왼쪽중간)의 기타 경사도 비율
116		Grd_gTg_other	목표 격자의 기타 경사도 비율
117		Grd_gRM_other	목표 격자와 인접한 격자(오른쪽중간)의 기타 경사도 비율
118		Grd_gLB_other	목표 격자와 인접한 격자(왼쪽아래)의 기타 경사도 비율
119		Grd_gCB_other	목표 격자와 인접한 격자(중간아래)의 기타 경사도 비율
120		Grd_gRB_other	목표 격자와 인접한 격자(왼쪽아래)의 기타 경사도 비율

NO	변수 그룹	모델의 변수명	설명
121	목표 격자와 목표 격자와 인접한 격자의 과거부터 현재까지의 공동주택 수	APT_gLT_2000	목표 격자와 인접한 격자(왼쪽위)의 2000년 공동주택 수
122		APT_gLT_2005	목표 격자와 인접한 격자(왼쪽위)의 2005년 공동주택 수
123		APT_gLT_2010	목표 격자와 인접한 격자(왼쪽위)의 2010년 공동주택 수
124		APT_gLT_2015	목표 격자와 인접한 격자(왼쪽위)의 2015년 공동주택 수
125		APT_gLT_2020	목표 격자와 인접한 격자(왼쪽위)의 2020년 공동주택 수
126		APT_gCT_2000	목표 격자와 인접한 격자(중간위)의 2000년 공동주택 수
127		APT_gCT_2005	목표 격자와 인접한 격자(중간위)의 2005년 공동주택 수
128		APT_gCT_2010	목표 격자와 인접한 격자(중간위)의 2010년 공동주택 수
129		APT_gCT_2015	목표 격자와 인접한 격자(중간위)의 2015년 공동주택 수
130		APT_gCT_2020	목표 격자와 인접한 격자(중간위)의 2020년 공동주택 수
131		APT_gRT_2000	목표 격자와 인접한 격자(오른쪽위)의 2000년 공동주택 수
132		APT_gRT_2005	목표 격자와 인접한 격자(오른쪽위)의 2005년 공동주택 수
133		APT_gRT_2010	목표 격자와 인접한 격자(오른쪽위)의 2010년 공동주택 수
134		APT_gRT_2015	목표 격자와 인접한 격자(오른쪽위)의 2015년 공동주택 수
135		APT_gRT_2020	목표 격자와 인접한 격자(오른쪽위)의 2020년 공동주택 수
136		APT_gLM_2000	목표 격자와 인접한 격자(왼쪽중간)의 2000년 공동주택 수
137		APT_gLM_2005	목표 격자와 인접한 격자(왼쪽중간)의 2005년 공동주택 수
138		APT_gLM_2010	목표 격자와 인접한 격자(왼쪽중간)의 2010년 공동주택 수
139		APT_gLM_2015	목표 격자와 인접한 격자(왼쪽중간)의 2015년 공동주택 수
140		APT_gLM_2020	목표 격자와 인접한 격자(왼쪽중간)의 2020년 공동주택 수
141		APT_gTg_2000	목표 격자(가운데중간)의 2000년 공동주택 수
142		APT_gTg_2005	목표 격자(가운데중간)의 2005년 공동주택 수
143		APT_gTg_2010	목표 격자(가운데중간)의 2010년 공동주택 수
144		APT_gTg_2015	목표 격자(가운데중간)의 2015년 공동주택 수
145		APT_gTg_2020	목표 격자(가운데중간)의 2020년 공동주택 수
146		APT_gRM_2000	목표 격자와 인접한 격자(오른쪽중간)의 2000년 공동주택 수
147		APT_gRM_2005	목표 격자와 인접한 격자(오른쪽중간)의 2005년 공동주택 수
148		APT_gRM_2010	목표 격자와 인접한 격자(오른쪽중간)의 2010년 공동주택 수
149		APT_gRM_2015	목표 격자와 인접한 격자(오른쪽중간)의 2015년 공동주택 수
150		APT_gRM_2020	목표 격자와 인접한 격자(오른쪽중간)의 2020년 공동주택 수
151		APT_gLB_2000	목표 격자와 인접한 격자(왼쪽아래)의 2000년 공동주택 수

NO	변수 그룹	모델의 변수명	설명
152		APT_gLB_2005	목표 격자와 인접한 격자(왼쪽아래)의 2005년 공동주택 수
153		APT_gLB_2010	목표 격자와 인접한 격자(왼쪽아래)의 2010년 공동주택 수
154		APT_gLB_2015	목표 격자와 인접한 격자(왼쪽아래)의 2015년 공동주택 수
155		APT_gLB_2020	목표 격자와 인접한 격자(왼쪽아래)의 2020년 공동주택 수
156		APT_gCB_2000	목표 격자와 인접한 격자(중간아래)의 2000년 공동주택 수
157		APT_gCB_2005	목표 격자와 인접한 격자(중간아래)의 2005년 공동주택 수
158		APT_gCB_2010	목표 격자와 인접한 격자(중간아래)의 2010년 공동주택 수
159		APT_gCB_2015	목표 격자와 인접한 격자(중간아래)의 2015년 공동주택 수
160		APT_gCB_2020	목표 격자와 인접한 격자(중간아래)의 2020년 공동주택 수
161		APT_gRB_2000	목표 격자와 인접한 격자(오른쪽아래)의 2000년 공동주택 수
162		APT_gRB_2005	목표 격자와 인접한 격자(오른쪽아래)의 2005년 공동주택 수
163		APT_gRB_2010	목표 격자와 인접한 격자(오른쪽아래)의 2010년 공동주택 수
164		APT_gRB_2015	목표 격자와 인접한 격자(오른쪽아래)의 2015년 공동주택 수
165		APT_gRB_2020	목표 격자와 인접한 격자(오른쪽아래)의 2020년 공동주택 수
166		목표 격자와 목표 격자와 인접한 격자의 과거부터 현재까지의 단독주택 수	ISO_gLT_2000
167	ISO_gLT_2005		목표 격자와 인접한 격자(왼쪽위)의 2005년 단독주택 수
168	ISO_gLT_2010		목표 격자와 인접한 격자(왼쪽위)의 2010년 단독주택 수
169	ISO_gLT_2015		목표 격자와 인접한 격자(왼쪽위)의 2015년 단독주택 수
170	ISO_gLT_2020		목표 격자와 인접한 격자(왼쪽위)의 2020년 단독주택 수
171	ISO_gCT_2000		목표 격자와 인접한 격자(중간위)의 2000년 단독주택 수
172	ISO_gCT_2005		목표 격자와 인접한 격자(중간위)의 2005년 단독주택 수
173	ISO_gCT_2010		목표 격자와 인접한 격자(중간위)의 2010년 단독주택 수
174	ISO_gCT_2015		목표 격자와 인접한 격자(중간위)의 2015년 단독주택 수
175	ISO_gCT_2020		목표 격자와 인접한 격자(중간위)의 2020년 단독주택 수
176	ISO_gRT_2000		목표 격자와 인접한 격자(오른쪽위)의 2000년 단독주택 수
177	ISO_gRT_2005		목표 격자와 인접한 격자(오른쪽위)의 2005년 단독주택 수
178	ISO_gRT_2010		목표 격자와 인접한 격자(오른쪽위)의 2010년 단독주택 수
179	ISO_gRT_2015		목표 격자와 인접한 격자(오른쪽위)의 2015년 단독주택 수
180	ISO_gRT_2020		목표 격자와 인접한 격자(오른쪽위)의 2020년 단독주택 수
181	ISO_gLM_2000		목표 격자와 인접한 격자(왼쪽중간)의 2000년 단독주택 수
182	ISO_gLM_2005		목표 격자와 인접한 격자(왼쪽중간)의 2005년 단독주택 수

NO	변수 그룹	모델의 변수명	설명
183		ISO_gLM_2010	목표 격자와 인접한 격자(왼쪽중간)의 2010년 단독주택 수
184		ISO_gLM_2015	목표 격자와 인접한 격자(왼쪽중간)의 2015년 단독주택 수
185		ISO_gLM_2020	목표 격자와 인접한 격자(왼쪽중간)의 2020년 단독주택 수
186		ISO_gTg_2000	목표 격자(가운데중간)의 2000년 단독주택 수
187		ISO_gTg_2005	목표 격자(가운데중간)의 2005년 단독주택 수
188		ISO_gTg_2010	목표 격자(가운데중간)의 2010년 단독주택 수
189		ISO_gTg_2015	목표 격자(가운데중간)의 2015년 단독주택 수
190		ISO_gTg_2020	목표 격자(가운데중간)의 2020년 단독주택 수
191		ISO_gRM_2000	목표 격자와 인접한 격자(오른쪽중간)의 2000년 단독주택 수
192		ISO_gRM_2005	목표 격자와 인접한 격자(오른쪽중간)의 2005년 단독주택 수
193		ISO_gRM_2010	목표 격자와 인접한 격자(오른쪽중간)의 2010년 단독주택 수
194		ISO_gRM_2015	목표 격자와 인접한 격자(오른쪽중간)의 2015년 단독주택 수
195		ISO_gRM_2020	목표 격자와 인접한 격자(오른쪽중간)의 2020년 단독주택 수
196		ISO_gLB_2000	목표 격자와 인접한 격자(왼쪽아래)의 2000년 단독주택 수
197		ISO_gLB_2005	목표 격자와 인접한 격자(왼쪽아래)의 2005년 단독주택 수
198		ISO_gLB_2010	목표 격자와 인접한 격자(왼쪽아래)의 2010년 단독주택 수
199		ISO_gLB_2015	목표 격자와 인접한 격자(왼쪽아래)의 2015년 단독주택 수
200		ISO_gLB_2020	목표 격자와 인접한 격자(왼쪽아래)의 2020년 단독주택 수
201		ISO_gCB_2000	목표 격자와 인접한 격자(중간아래)의 2000년 단독주택 수
202		ISO_gCB_2005	목표 격자와 인접한 격자(중간아래)의 2005년 단독주택 수
203		ISO_gCB_2010	목표 격자와 인접한 격자(중간아래)의 2010년 단독주택 수
204		ISO_gCB_2015	목표 격자와 인접한 격자(중간아래)의 2015년 단독주택 수
205		ISO_gCB_2020	목표 격자와 인접한 격자(중간아래)의 2020년 단독주택 수
206		ISO_gRB_2000	목표 격자와 인접한 격자(오른쪽아래)의 2000년 단독주택 수
207		ISO_gRB_2005	목표 격자와 인접한 격자(오른쪽아래)의 2005년 단독주택 수
208		ISO_gRB_2010	목표 격자와 인접한 격자(오른쪽아래)의 2010년 단독주택 수
209		ISO_gRB_2015	목표 격자와 인접한 격자(오른쪽아래)의 2015년 단독주택 수
210		ISO_gRB_2020	목표 격자와 인접한 격자(오른쪽아래)의 2020년 단독주택 수
211	목표 격자와 목표 격자와 인접한 격자의 과거부터	ARA_gLT_2000	목표 격자와 인접한 격자(왼쪽위)의 2000년 주거용건물 연면적
212		ARA_gLT_2005	목표 격자와 인접한 격자(왼쪽위)의 2005년 주거용건물 연면적
213		ARA_gLT_2010	목표 격자와 인접한 격자(왼쪽위)의 2010년 주거용건물 연면적

NO	변수 그룹	모델의 변수명	설명
214	현재까지의 주거용건물 연면적	ARA_gLT_2015	목표 격자와 인접한 격자(왼쪽위)의 2015년 주거용건물 연면적
215		ARA_gLT_2020	목표 격자와 인접한 격자(왼쪽위)의 2020년 주거용건물 연면적
216		ARA_gCT_2000	목표 격자와 인접한 격자(중간위)의 2000년 주거용건물 연면적
217		ARA_gCT_2005	목표 격자와 인접한 격자(중간위)의 2005년 주거용건물 연면적
218		ARA_gCT_2010	목표 격자와 인접한 격자(중간위)의 2010년 주거용건물 연면적
219		ARA_gCT_2015	목표 격자와 인접한 격자(중간위)의 2015년 주거용건물 연면적
220		ARA_gCT_2020	목표 격자와 인접한 격자(중간위)의 2020년 주거용건물 연면적
221		ARA_gRT_2000	목표 격자와 인접한 격자(오른쪽위)의 2000년 주거용건물 연면적
222		ARA_gRT_2005	목표 격자와 인접한 격자(오른쪽위)의 2005년 주거용건물 연면적
223		ARA_gRT_2010	목표 격자와 인접한 격자(오른쪽위)의 2010년 주거용건물 연면적
224		ARA_gRT_2015	목표 격자와 인접한 격자(오른쪽위)의 2015년 주거용건물 연면적
225		ARA_gRT_2020	목표 격자와 인접한 격자(오른쪽위)의 2020년 주거용건물 연면적
226		ARA_gLM_2000	목표 격자와 인접한 격자(왼쪽중간)의 2000년 주거용건물 연면적
227		ARA_gLM_2005	목표 격자와 인접한 격자(왼쪽중간)의 2005년 주거용건물 연면적
228		ARA_gLM_2010	목표 격자와 인접한 격자(왼쪽중간)의 2010년 주거용건물 연면적
229		ARA_gLM_2015	목표 격자와 인접한 격자(왼쪽중간)의 2015년 주거용건물 연면적
230		ARA_gLM_2020	목표 격자와 인접한 격자(왼쪽중간)의 2020년 주거용건물 연면적
231		ARA_gTg_2000	목표 격자(가운데중간)의 2000년 주거용건물 연면적
232		ARA_gTg_2005	목표 격자(가운데중간)의 2005년 주거용건물 연면적
233		ARA_gTg_2010	목표 격자(가운데중간)의 2010년 주거용건물 연면적
234		ARA_gTg_2015	목표 격자(가운데중간)의 2015년 주거용건물 연면적
235		ARA_gTg_2020	목표 격자(가운데중간)의 2020년 주거용건물 연면적
236		ARA_gRM_2000	목표 격자와 인접한 격자(오른쪽중간)의 2000년 주거용건물 연면적
237		ARA_gRM_2005	목표 격자와 인접한 격자(오른쪽중간)의 2005년 주거용건물 연면적
238	ARA_gRM_2010	목표 격자와 인접한 격자(오른쪽중간)의 2010년 주거용건물 연면적	
239	ARA_gRM_2015	목표 격자와 인접한 격자(오른쪽중간)의 2015년 주거용건물 연면적	
240	ARA_gRM_2020	목표 격자와 인접한 격자(오른쪽중간)의 2020년 주거용건물 연면적	
241	ARA_gLB_2000	목표 격자와 인접한 격자(왼쪽아래)의 2000년 주거용건물 연면적	
242	ARA_gLB_2005	목표 격자와 인접한 격자(왼쪽아래)의 2005년 주거용건물 연면적	
243	ARA_gLB_2010	목표 격자와 인접한 격자(왼쪽아래)의 2010년 주거용건물 연면적	
244	ARA_gLB_2015	목표 격자와 인접한 격자(왼쪽아래)의 2015년 주거용건물 연면적	

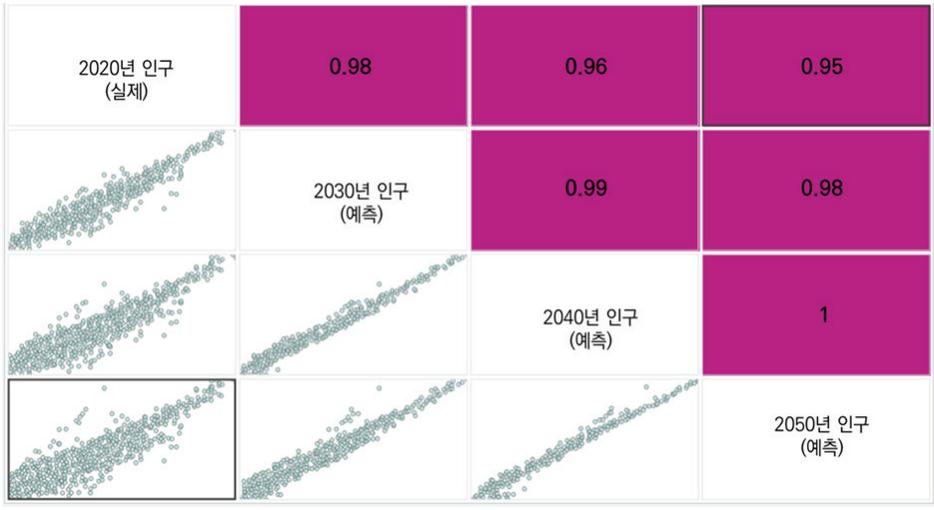
NO	변수 그룹	모델의 변수명	설명	
245		ARA_gLB_2020	목표 격자와 인접한 격자(왼쪽아래)의 2020년 주거용건물 연면적	
246		ARA_gCB_2000	목표 격자와 인접한 격자(중간아래)의 2000년 주거용건물 연면적	
247		ARA_gCB_2005	목표 격자와 인접한 격자(중간아래)의 2005년 주거용건물 연면적	
248		ARA_gCB_2010	목표 격자와 인접한 격자(중간아래)의 2010년 주거용건물 연면적	
249		ARA_gCB_2015	목표 격자와 인접한 격자(중간아래)의 2015년 주거용건물 연면적	
250		ARA_gCB_2020	목표 격자와 인접한 격자(중간아래)의 2020년 주거용건물 연면적	
251		ARA_gRB_2000	목표 격자와 인접한 격자(오른쪽아래)의 2000년 주거용건물 연면적	
252		ARA_gRB_2005	목표 격자와 인접한 격자(오른쪽아래)의 2005년 주거용건물 연면적	
253		ARA_gRB_2010	목표 격자와 인접한 격자(오른쪽아래)의 2010년 주거용건물 연면적	
254		ARA_gRB_2015	목표 격자와 인접한 격자(오른쪽아래)의 2015년 주거용건물 연면적	
255		ARA_gRB_2020	목표 격자와 인접한 격자(오른쪽아래)의 2020년 주거용건물 연면적	
256		목표 격자와 목표 격자와 인접한 격자의 과거부터 현재까지의 건축물 수	BUD_gLT_2000	목표 격자와 인접한 격자(왼쪽위)의 2000년 건축물 수
257			BUD_gLT_2005	목표 격자와 인접한 격자(왼쪽위)의 2005년 건축물 수
258			BUD_gLT_2010	목표 격자와 인접한 격자(왼쪽위)의 2010년 건축물 수
259			BUD_gLT_2015	목표 격자와 인접한 격자(왼쪽위)의 2015년 건축물 수
260	BUD_gLT_2020		목표 격자와 인접한 격자(왼쪽위)의 2020년 건축물 수	
261	BUD_gCT_2000		목표 격자와 인접한 격자(중간위)의 2000년 건축물 수	
262	BUD_gCT_2005		목표 격자와 인접한 격자(중간위)의 2005년 건축물 수	
263	BUD_gCT_2010		목표 격자와 인접한 격자(중간위)의 2010년 건축물 수	
264	BUD_gCT_2015		목표 격자와 인접한 격자(중간위)의 2015년 건축물 수	
265	BUD_gCT_2020		목표 격자와 인접한 격자(중간위)의 2020년 건축물 수	
266	BUD_gRT_2000		목표 격자와 인접한 격자(오른쪽위)의 2000년 건축물 수	
267	BUD_gRT_2005		목표 격자와 인접한 격자(오른쪽위)의 2005년 건축물 수	
268	BUD_gRT_2010		목표 격자와 인접한 격자(오른쪽위)의 2010년 건축물 수	
269	BUD_gRT_2015		목표 격자와 인접한 격자(오른쪽위)의 2015년 건축물 수	
270	BUD_gRT_2020		목표 격자와 인접한 격자(오른쪽위)의 2020년 건축물 수	
271	BUD_gLM_2000		목표 격자와 인접한 격자(왼쪽중간)의 2000년 건축물 수	
272	BUD_gLM_2005		목표 격자와 인접한 격자(왼쪽중간)의 2005년 건축물 수	
273	BUD_gLM_2010		목표 격자와 인접한 격자(왼쪽중간)의 2010년 건축물 수	
274	BUD_gLM_2015		목표 격자와 인접한 격자(왼쪽중간)의 2015년 건축물 수	
275	BUD_gLM_2020		목표 격자와 인접한 격자(왼쪽중간)의 2020년 건축물 수	

NO	변수 그룹	모델의 변수명	설명
276		BUD_gTg_2000	목표 격자(가운데중간)의 2000년 건축물 수
277		BUD_gTg_2005	목표 격자(가운데중간)의 2005년 건축물 수
278		BUD_gTg_2010	목표 격자(가운데중간)의 2010년 건축물 수
279		BUD_gTg_2015	목표 격자(가운데중간)의 2015년 건축물 수
280		BUD_gTg_2020	목표 격자(가운데중간)의 2020년 건축물 수
281		BUD_gRM_2000	목표 격자와 인접한 격자(오른쪽중간)의 2000년 건축물 수
282		BUD_gRM_2005	목표 격자와 인접한 격자(오른쪽중간)의 2005년 건축물 수
283		BUD_gRM_2010	목표 격자와 인접한 격자(오른쪽중간)의 2010년 건축물 수
284		BUD_gRM_2015	목표 격자와 인접한 격자(오른쪽중간)의 2015년 건축물 수
285		BUD_gRM_2020	목표 격자와 인접한 격자(오른쪽중간)의 2020년 건축물 수
286		BUD_gLB_2000	목표 격자와 인접한 격자(왼쪽아래)의 2000년 건축물 수
287		BUD_gLB_2005	목표 격자와 인접한 격자(왼쪽아래)의 2005년 건축물 수
288		BUD_gLB_2010	목표 격자와 인접한 격자(왼쪽아래)의 2010년 건축물 수
289		BUD_gLB_2015	목표 격자와 인접한 격자(왼쪽아래)의 2015년 건축물 수
290		BUD_gLB_2020	목표 격자와 인접한 격자(왼쪽아래)의 2020년 건축물 수
291		BUD_gCB_2000	목표 격자와 인접한 격자(중간아래)의 2000년 건축물 수
292		BUD_gCB_2005	목표 격자와 인접한 격자(중간아래)의 2005년 건축물 수
293		BUD_gCB_2010	목표 격자와 인접한 격자(중간아래)의 2010년 건축물 수
294		BUD_gCB_2015	목표 격자와 인접한 격자(중간아래)의 2015년 건축물 수
295		BUD_gCB_2020	목표 격자와 인접한 격자(중간아래)의 2020년 건축물 수
296		BUD_gRB_2000	목표 격자와 인접한 격자(오른쪽아래)의 2000년 건축물 수
297		BUD_gRB_2005	목표 격자와 인접한 격자(오른쪽아래)의 2005년 건축물 수
298		BUD_gRB_2010	목표 격자와 인접한 격자(오른쪽아래)의 2010년 건축물 수
299		BUD_gRB_2015	목표 격자와 인접한 격자(오른쪽아래)의 2015년 건축물 수
300		BUD_gRB_2020	목표 격자와 인접한 격자(오른쪽아래)의 2020년 건축물 수
301	목표 격자의 인구	POP_gTg_2020	목표 격자의 2020년 인구 (목표값, Target)

자료: 저자 작성

부록 3. 2020, 2030, 2040 및 2050년 인구 예측 상관관계

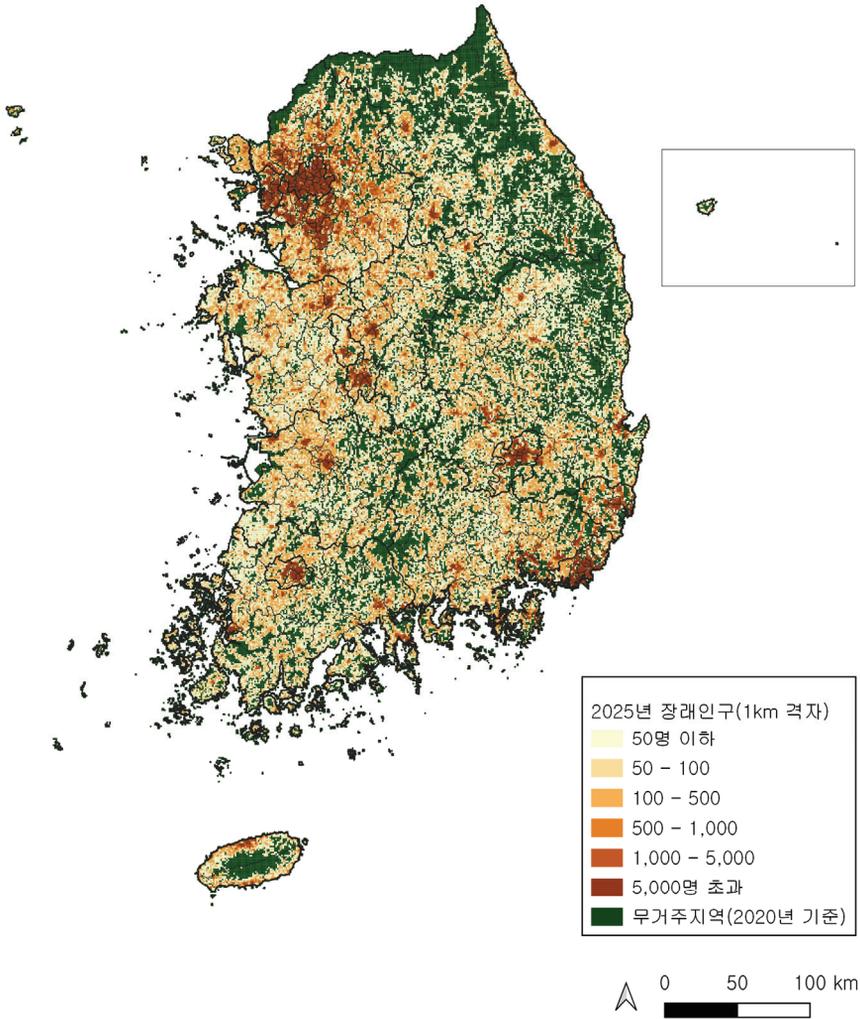
그림 부3-1 | 2020, 2030, 2040 및 2050년 인구 예측 상관관계



자료: 저자 작성

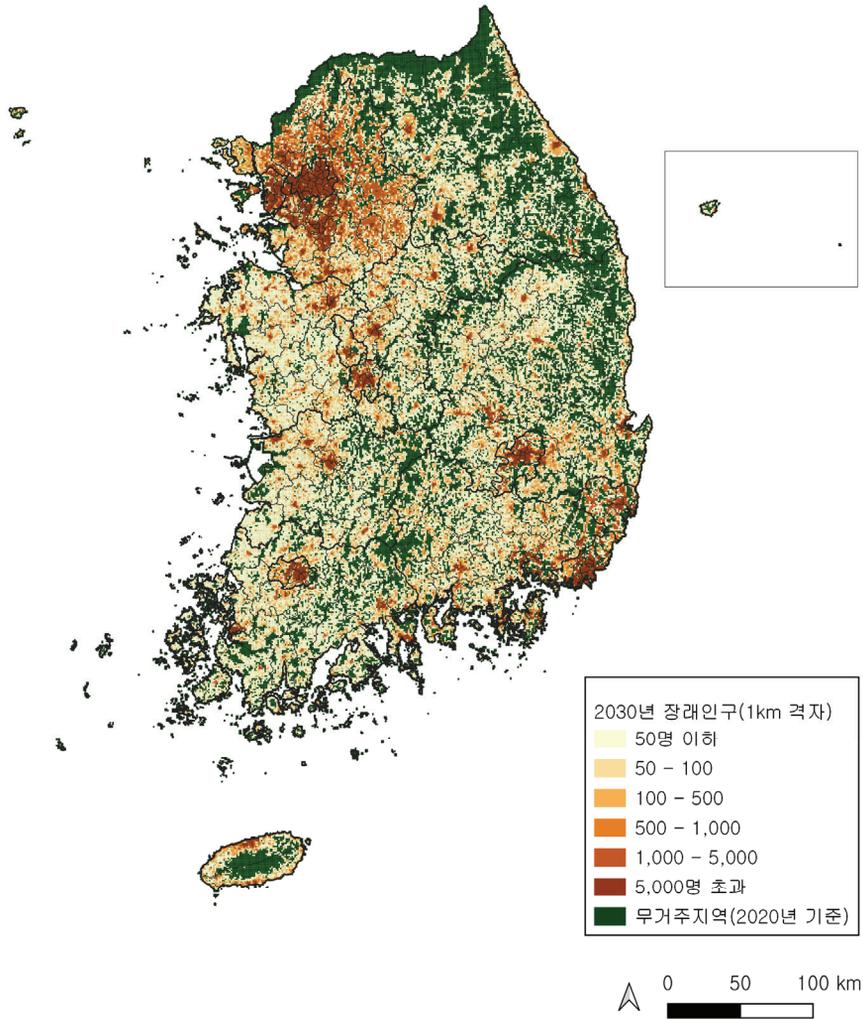
부록 4. 2025년 ~ 2050년 장래인구 예측 결과 전국 인구분포

그림 부4-1 | 2025년 장래인구 예측 결과 인구 분포



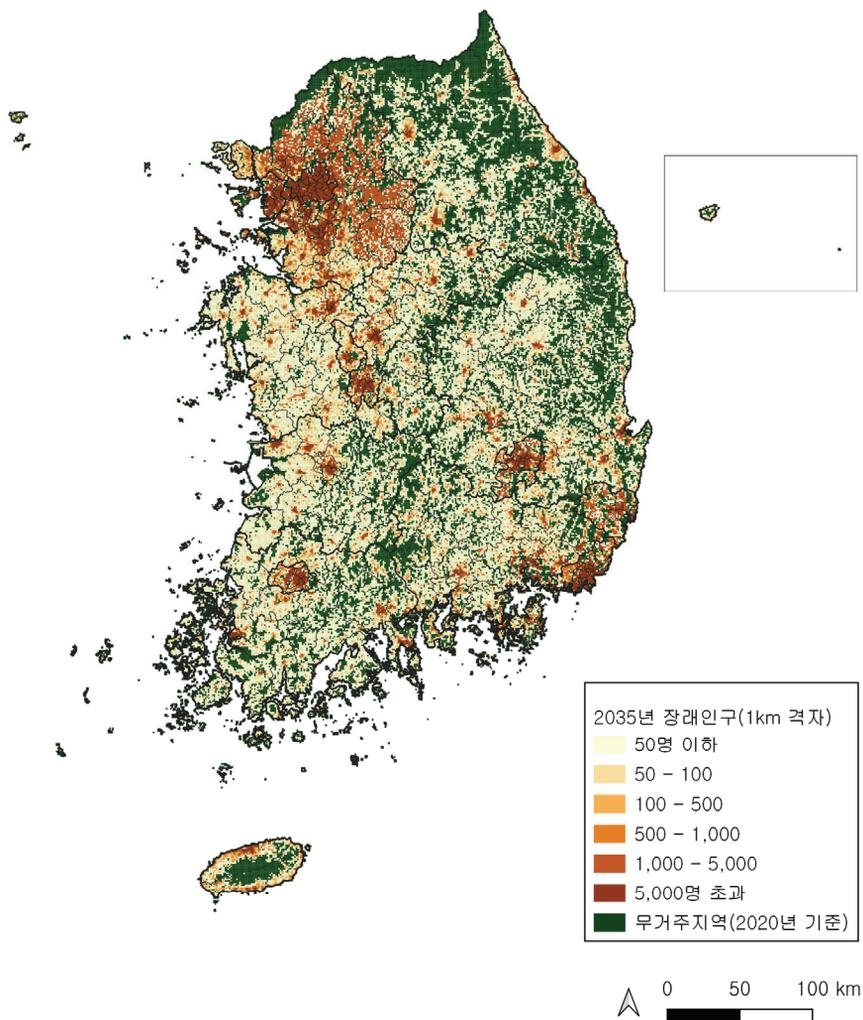
자료: 저자 작성

그림 부4-2 | 2030년 장래인구 예측 결과 인구 분포



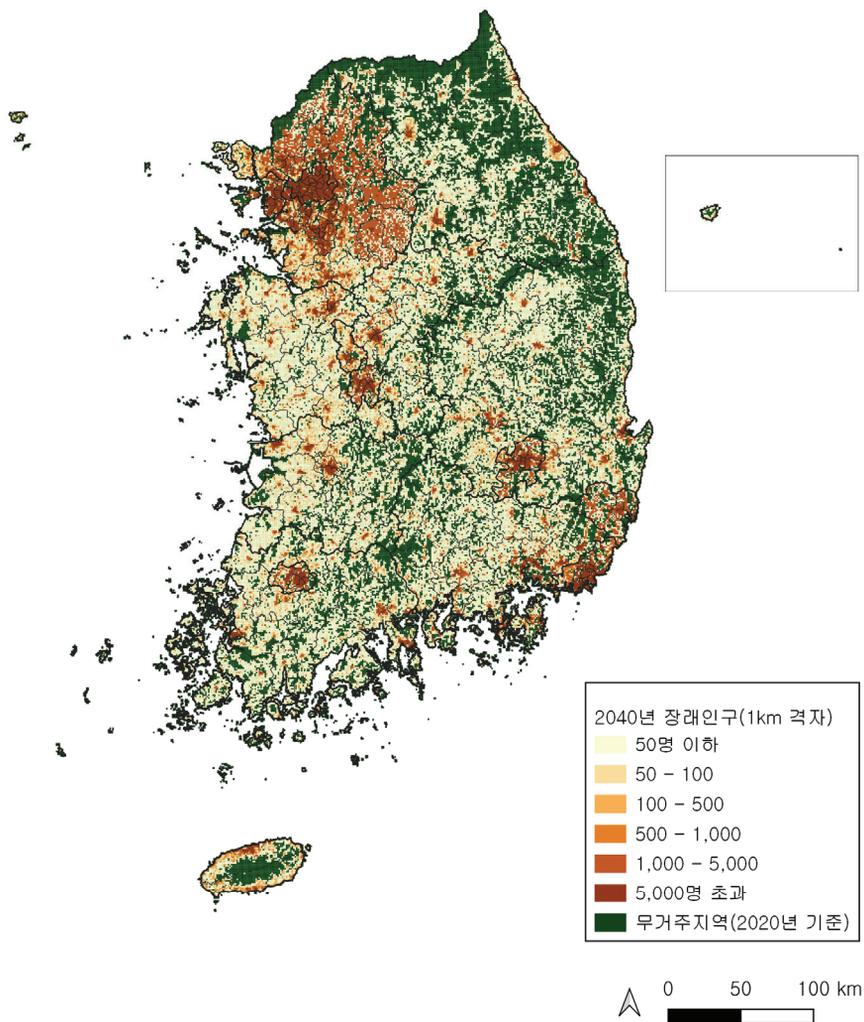
자료: 저자 작성

그림 부4-3 | 2035년 장래인구 예측 결과 인구 분포



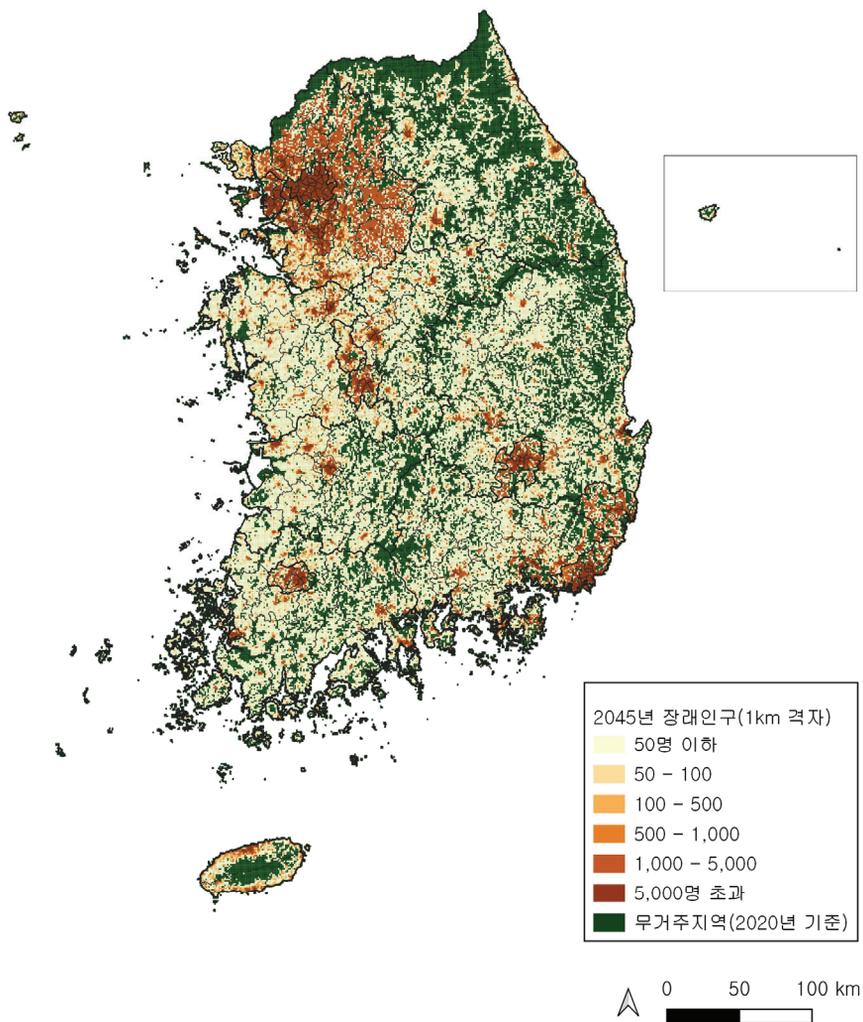
자료: 저자 작성

그림 부4-4 | 2040년 장래인구 예측 결과 인구 분포



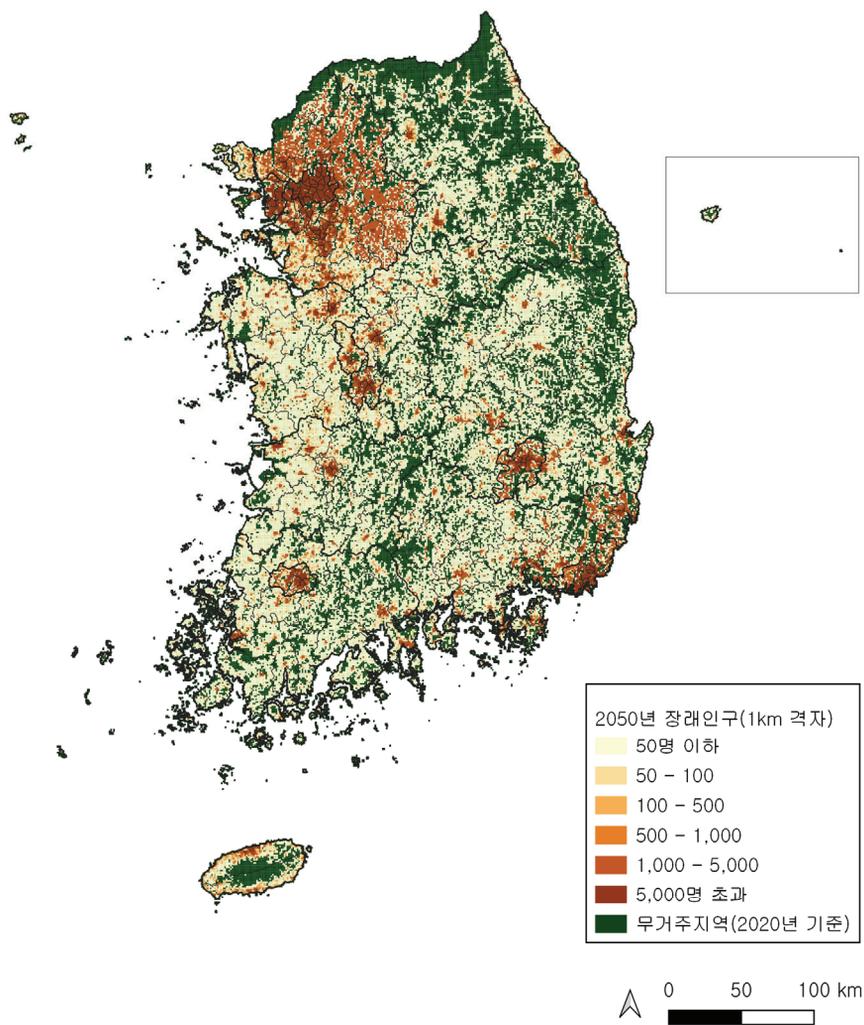
자료: 저자 작성

그림 부4-5 | 2045년 장래인구 예측 결과 인구 분포



자료: 저자 작성

그림 부4-6 | 2050년 장래인구 예측 결과 인구 분포



자료: 저자 작성

기본 22-24

인공지능 기법을 적용한 소지역단위 장래인구 예측 방법론 개발 기초 연구

저 자 이보경, 장요한, 정예진, 홍사흠, 김동근

발 행 인 강현수

발 행 처 국토연구원

출판등록 제2017-9호

발 행 2022년 12월 31일

주 소 세종특별자치시 국책연구원로 5

전 화 044-960-0114

팩 스 044-211-4760

가 격 7,000원

I S B N 979-11-5898-808-1

홈페이지 <http://www.krihs.re.kr>

© 2022, 국토연구원

이 연구보고서를 인용하실 때는 다음과 같은 사항을 기재해주시시오.

이보경, 장요한, 정예진, 홍사흠, 김동근. 2022. 인공지능 기법을 적용한 소지역단위 장래인구 예측 방법론 개발 기초 연구. 세종: 국토연구원.

이 연구보고서의 내용은 국토연구원의 자체 연구물로서 정부의 정책이나 견해와는 상관없습니다.

이 연구보고서는 한국출판인협회에서 제공한 KoPub 서체와 대한인쇄문화협회가 제공한 바른바탕체 등이 적용되어 있습니다.

인공지능 기법을 적용한 소지역단위 장래인구 예측 방법론 개발 기초 연구

A Study on Small Area Population Forecast Model with Artificial Intelligence



제1장 서론

제2장 소지역단위 장래인구 예측 연구 및 사례 조사

제3장 소지역단위 장래인구 예측 방법론 개발

제4장 GPAI 학습 및 장래인구 예측 결과

제5장 정책 제언 및 결론



KRIHS 국토연구원

(30147) 세종특별자치시 국책연구원로 5 (반곡동)
TEL (044) 960-0114 FAX (044) 211-4760

