

공간·지리적 자료의 공간자기상관성을 최소화하는 공간샘플링 기법에 관한 연구

이연수* · 이만출** · 나경범*** · 강준모****

Lee, Youn Soo*, Lee, Man Choul**, Lah, Kyung Beom***, Kang, Jun Mo****

A study on the Spatial Sampling Method to Minimize Spatial Autocorrelation of Spatial and Geographical Data

ABSTRACT

The study focused on analyzing spatial sampling by minimizing autocorrelation of spatial based on spatial and geographical data. The study concluded two different ways of minimizing autocorrelation. First, it was important to use suitable spatial sampling method to alienate spatial autocorrelation from spatial or geographical data. The shear distribution rate of public transportation in Seoul resulted in high rate of autocorrelation. However, the study showed samples eliminated autocorrelation when samples were extracted with reasonable distance(above 400m) apart. Without spatial sampling the distortion of spatial data leads to false results; therefore, spatial sampling is indispensable. Second, factors which fluctuates shear distribution of public transportation spatial sampling changed before and after spatial sampling. This was caused by incapable of controlling inherent spatial autocorrelation of the data.

Key words : Spatial sampling, Spatial autocorrelation, Urban spatial structure, Public transportation use modal share

초 록

본 연구의 목적은 공간적 또는 지리적 데이터인 도시구조특성 요소의 공간자기상관성을 최소화 시킬 수 있는 공간샘플링 기법에 대해 제시하고자 한다. 연구의 주요결과와 시사점은 다음과 같다. 첫째, 공간적 또는 지리적 자료가 지니고 있는 공간자기상관성을 제거하기 위해서는 적절한 공간샘플링 방법을 사용하여야 한다는 점이다. 서울시 전체 행정구역별 대중교통 분담률에 관한 공간자기상관성을 측정된 결과 간 의존성이 상당히 높은 것으로 분석되었다. 하지만 적절한 거리(400m)이상을 이격 시켜 공간샘플링을 실시한 후 공간자기상관성이 제거됨을 확인하였다. 공간샘플링을 실시하지 않으면, 공간적 자료의 왜곡으로 인한 잘못된 결과가 도출되며, 이를 해결하기 위해서는 공간샘플링을 하여 표본을 추출하는 과정이 필수불가결 하다는 점이다. 둘째, 공간샘플링 전후의 대중교통 분담률에 영향을 미치는 도시공간구조 특성 요인이 달라진다는 점이다. 그러나 이는 공간적 자료에 존재하는 공간자기상관성을 통제하지 못한 왜곡된 결과이다.

검색어 : 공간샘플링, 공간자기상관성, 도시공간구조, 대중교통수단분담률

* 정회원·교신저자·홍익대학교 도시계획과 박사수료 (Corresponding Author·Hongik University·zizing12@naver.com)

** 홍익대학교 도시계획과 박사수료 (Hongik University·m.c.lee@hanmail.net)

*** 홍익대학교 도시공학과 박사수료 (Hongik University·smartlah@hanmail.net)

**** 정회원·홍익대학교 도시공학과 정교수 (Hongik University·unmo@hongik.ac.kr)

Received May 2, 2014/ revised June 2, 2014/ accepted June 9, 2014

1. 서론

1.1 연구의 배경 및 목적

우리나라는 도시의 성장으로 인하여 국민들의 소득이 증가하였고 경제적 여건이 양호해 짐에 따라 인구 및 자동차의 숫자가 과거에 비하여 비약적으로 증가하였다. 이로 인해 인구과밀, 자연환경의 훼손, 교통혼잡, 에너지의 과소비, 대기오염, 소음, 탄소배출 등 도시 내에서 승용차 중심의 생활로 인한 여러 환경적인 문제에 직면한 상태이다. 이러한 승용차 중심의 도시 문제에 대응하기 위하여 친환경 녹색교통 수단으로서 그 중요성이 부각되고 있는 버스 및 지하철 등의 대중교통 이용 활성화를 위한 수많은 연구들이 진행되고 있으며 국가에서도 정책적으로 대중교통 활성화를 도모하고 있는 실정이다. 또한 수도권을 중심으로 기존의 승용차 중심의 교통체계를 지양하고 대중교통 중심의 교통체제로 전환하기 위한 노력을 기울이고 있다. 그러나 대중교통의 이용에 영향을 미치는 도시공간구조 요인들은 지리적 또는 공간적인 특성을 지니고 있기 때문에 공간적으로 일종의 상호작용을 하게 되는데 이를 공간종속성 또는 공간자기상관성이라 한다. 이는 Tobler (1970)가 주장하는 지리학의 제 1법칙인 “모든 것은 다른 모든 것과 관련되지만, 가까운 것이 멀리 있는 것보다 더 관련이 있다.”라는 법칙에서도 알 수 있다. 이러한 공간분석에 관한 논의는 1970년대부터 공간자기상관성을 고려하여 분석해야 한다고 주장되어왔다. 또한 공간분석을 위한 툴(Tool) 역시 연구자들에 의해서 사용되어졌고 발전되어왔다. 공간분석 시에 공간자기상관을 무시한 채 공간적 자료에 대한 분석을 실시한다면 결과가 왜곡되어 나타난다. 또 다른 연구자들은 공간자기상관성을 배제하고 공간분석을 실시해야 한다는 주장이 제기되었다. 후자의 연구자들의 주장은 공간자기상관이 생기는 이유는 전자의 연구자들과 같은 입장에서 동의를 하지만 공간자기상관성을 인정하기 때문에 이를 제거하거나 최소화하여 분석을 실시해야 가장 객관적인 실증분석이 가능하다는 주장이다.

본 연구에서는 주로 대중교통분담률에 영향을 미치는 도시공간구조 요인을 알아보기 위한 분석을 실시하였으며, 이 때 공간적 자료의 공간자기상관을 최소화 하기위한 공간샘플링 방법론을 탐색한다. 즉, 공간적 또는 지리적 데이터인 도시구조특성 요소의 공간자기상관성을 최소화 시킬 수 있는 공간샘플링 기법에 대해 제시하고자 한다.

1.2 연구의 범위 및 방법

본 연구의 공간적 범위는 우리나라에서 지하철 및 버스 등의 대중교통 이용수요가 가장 높은 서울시 행정구역을 대상으로 하였다. 연구는 문헌고찰을 통해 지금까지 논의되어 온 지리학적 또는 도시학적 관점에서의 공간자기상관성(Spatial Autocorrelation)에

대해서 알아보고 공간샘플링(Spatial Sampling)기법을 고찰한다. 또한 도시공간데이터(GIS data) 분석을 통한 변수의 추출, 대중교통수단 분담률 데이터를 이용하여 앞서 추출된 변수들과의 공간자기상관이 있는지의 여부를 측정한다. 공간자기상관이 있는 회귀모형과 공간자기상관성을 최대한 배제시켜 정립된 회귀모형과의 비교를 통해 공간데이터를 이용한 회귀모형의 적합성 및 차이점을 도출한다. 마지막으로 공간자기상관을 최소화할 수 있는 공간샘플링 기법을 제시하는 할 것이다.

2. 이론고찰 및 선행연구검토

2.1 공간적 상호작용에 의한 공간자기상관

공간을 연구의 대상으로 삼는 지역학(Regional Science)이나 도시경제학은 오랫동안 공간실체(Spatial Entities)들의 위치, 상호작용, 공간구조 등에 관심을 기울여 왔다. 1956년 Isard가 발표한 연구문헌인 “Location and Space Economy”에서 지리적 위치와 공간에 대한 고찰을 역설하였으며 그 이후 지역학자들은 ‘공간(Space)’의 역할을 주목하기 시작하였다. 1970년대 초 Jean Paelinck가 명명한 ‘공간계량분석론(Spatial Econometrics)’이라는 용어를 제안하면서 지역학 분야의 계량분석에서 공간의 영향에 의해 모델 설정이나 추정에 야기되는 문제들의 심각성을 인식하게 되었다(Kim, 2003, 재구성). 정치적·지리적·경제적 요소를 기반으로 하는 공간자료는 순수한 자료 자체의 정보 뿐만이 아닌 지리적 공간에 관한 정보를 함께 포함하고 있다. 이러한 공간이라는 요인을 자료에서 고려하지 못하고 있어 공간자기상관성이라는 개념이 발생하였다.

기존의 선형회귀모델에서는 공간영향을 무시한 채 분석이 실시되고 있으며 이러한 경우 사회경제현상, 인구현상, 자연현상 등 공간상에서 나타나는 특성인 의존성 및 상호작용을 통제하지 못한다(Doreian, 1981, 재구성). 또한 Berry (1971)는 어떤 공간의 특성은 단순히 그 공간의 특성에 의해서가 아니라, 다른 장소들과의 연계와 함께 일어나는 현상에 의해 영향을 받기 때문에 공간자기상관이 일어난다고 주장하였다. 이러한 주장을 뒷받침하는 Kim (2003)의 연구에서는 공간자기상관의 원인이 공간실체들의 상호작용에 있다고 주장하였다. 인간은 공간상에서 유무형의 상호작용을 끊임없이 하고 있는데 공간상의 상호작용은 정보 및 기술의 확장 등으로 나타나기도 하며, 공간상의 위계구조도 낳기도 한다고 하였다. 결과적으로 이러한 공간과정을 통해 공간 위의 한 장소에서 발생하는 현상이 다른 장소에서 현상을 결정하게 되어 공간자기상관이 발생한다고 하였다.

위의 연구들이나 이론들을 종합해 볼 때 공간자기상관성은 어떠한 현상의 지리적·공간적 패턴을 설명함에 있어 나타나는 현상에 대한 결과가 공간적 패턴에 의해 나타나는 상호작용 또는 연계특성

에 의해 발생하는 영향이라 판단된다.

공간자기상관성을 분석하기 위해 가장 보편적으로 사용되는 방법으로는 Moran's I 측도가 있으며, 이 방법은 속성값의 유사도에 관한 것으로 두 변수간의 공분산을 측정하는 기법이다. Moran's I 계수는 인접한 지역이 비슷한 특징을 가지고 있을 때 값이 커지면 양의 값을 갖으며, 서로 다른 특징을 나타내면 음의 값을 갖는다. Moran's I는 현재 나타나는 상관관계수에 가중치를 두고, 가중치는 지리적 인접성을 반영한다. I의 값이 0보다 크면 정적인 공간자기상관을 나타내고, 0보다 작은 것은 부적인 공간자기상관을 나타낸다 (Ha, 2005).

2.2 공간정보의 활용을 위한 공간샘플링(Spatial Sampling) 방법론

지리적 또는 공간적 정보를 활용하기 위해서는 공간자기상관성을 염두해두어야 한다는 것은 기존의 이론과 연구들에서 중요한 사항으로 여겨져 왔다. Haining (1993) and Irwin (2002) 등 다수의 연구와 이론에 따르면 공간자기상관성의 오차를 줄이기 위해서는 적절한 공간샘플링이 필요하다고 하였다. 따라서, 공간자기상관성을 최소화시킬 수 있는 샘플링 방법론은 대표적으로 크게 세 가지로 나뉘볼 수 있다.

첫째, 무작위 추출법(Simple Random Sampling)은 모집단에서 표본을 추출하여 이 표본에 대한 조사결과를 가지고 모집단의 성질에 관한 추정을 하는 경우, 문제가 되는 것은 추출의 방식이다. 이 방법에 의한 추정의 효율은 그리 높지 않기 때문에 실제의 조사에서는 많이 사용되지 않는다(Park, 2011). 이 방법의 경우 모집단에서 랜덤하게 샘플을 추출할 수 있지만 가까운 거리의 샘플들 간의 중복이 일어날 확률이 크기 때문에 중복성 검사를 다시 시행하여야 한다는 점에서 비효율적이다. 특히 가까운 거리의 샘플들은 정보의 '중복'이 일어날 확률이 크며, 이는 곧 가까운 거리의 정보들 사이에서 공간자기상관성이 일어날 확률이 크다는 것을 의미한다. 또한 무작위 추출법은 지역 또는 공간적인 군집이 랜덤하지 않게 분포되어진 범위 내에서는 샘플을 추출하기 힘들고 독립적인 공간 안에서 샘플을 추출해야만 한다는 점에서 샘플링의 시간적 소모가 심하다는 단점이 있다.

둘째, 층화 무작위 추출법(stratified random sampling)은 알고

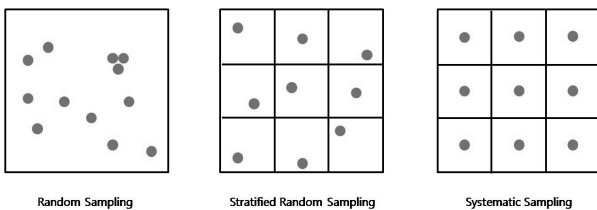


Fig. 1. Spatial Sampling Methodology

있는 정보로 모집단을 어떤 층화기준에 의하여 여러 개의 층으로 분할한 다음 각 층에서 독립적으로 일정한 수의 샘플을 임의로 추출하는 방법을 말한다(Chung, 2006). 이 방법에서는 우선 조사대상이 되는 모집단을 임의의 여러 개의 집단으로 나누고 개별 집단 하나를 '층'이라 부른다. 각 '층'이 모집단이 되어 어떤 방법으로 미리 할당된 수에 맞추어 각 층에서 표본을 추출한다. 이 방법의 장점으로는 보다 균일한 범위 내에서 샘플들을 랜덤하게 추출하기 때문에 정보의 중복성에 대한 위험은 그만큼 줄어들게 된다. 그러나 이 방법 또한 정보의 중복성을 전부 배제할 수 없다는 단점이 있다.

셋째, 계통추출법(Systematic Sampling)은 가장 간편한 임의추출의 방법이며, 조사 표본수가 많을 때는 난수표(亂數表)를 사용하는 일도 간단하지 않으므로 모집단(母集團)의 전요소(全要素)에 일련번호를 붙이고, 처음 하나의 표본을 임의추출한 다음은 일정한 간격으로 추출하는 방법이다(Doopedia, 2013). 이 샘플링 방법의 장점은 공간자료의 중복을 줄여주어 공간자기상관성을 최소화할 수 있다는 점이다. 또한 자료의 중복성 검사를 실시하지 않아도 되기 때문에 샘플링의 시간과 비용이 줄어든다. 또한 정규분포를 사용하는 모형, 가량 선형회귀분석과 같은 모형에서의 사용이 특히 유용하다. 그러나 이 샘플링 방법을 사용하기 위해서는 제한된 공간 즉, 범위 안에서 선택한 샘플이 유효하다는 것을 입증해야만 한다. 이는 앞서 서술한 Moran' I 측도법을 이용하여 유효성을 입증하는 것이 일반적으로 사용되는 방법이기도 하다.

2.3 대중교통 이용에 영향을 미치는 도시공간구조 특성 영향인자

도시공간구조는 도시를 구성하는 제반 요소들과 이들 요소들간의 상호관련성, 즉, 모든 요소들의 총체와 각 요소들 사이의 관련성을 뜻하며, 공간의 변화는 요소들의 분포로 인해 공간적 구조를 산출하는 매카니즘을 의미한다(Kim, 1996).

기존 연구들에서 사용된 도시공간구조 변수들을 Sung et al. (2006)은 도시개발밀도가 높은 역세권은 보행교통을 높이며, 승용차의 이용율이 낮다는 분석결과를 제시하였다. 이 연구에서 주목해야 할 점은 토지이용불균형지수(Mix)의 결과는 역세권내 토지이용의 혼합비율이 적절하게 배분되어 있으면 승용차에 보다 보행교통 비율이 높아지며, 또한 지하철에 비하여 도보의 이용이 높아진다는 결과를 도출하였다. Sung et al. (2008)의 또 다른 연구에서는 서울시 역세권에서의 토지이용과 도시설계 요소들이 대중교통 이용에 미치는 영향에 대하여 분석하였다. 토지이용변수로는 용도별 밀도, 다양성지수, 접근성 등의 변수를 사용하였으며, 도시설계 요소로는 가로망 및 건축물 변수를 사용하였다. 분석결과 상업 및 업무중심의 고밀개발과 토지이용의 복합화, 협소한 가로망이 밀집될수록 대중교통 이용증진에 긍정적인 영향을 미치는 것을

확인 할 수 있었다. Ewing and Cervero (2011)의 연구에서는 고밀개발, 토지이용의 복합성, 보행친화적인 가로망과 같은 TOD 계획요인을 이용하여 총통행량과 총통행거리에 미치는 영향에 대한 연구를 수행하였다. 그 결과 고밀도로 개발되고, 토지이용이 복합화 되어있으며, 보행친화적 가로망과 도시설계를 하였을 경우 통행량 또한 증가한다는 결과를 도출하였다. Cervero and Radisch (1996)의 연구에서는 개인의 소득수준 및 가족형태에 따라서 승용차 중심의 근린주교와 대중교통 중심의 근린주교의 통행수단 분담률의 비교분석한 연구를 수행하였다. Oh et al. (2009)은 대중교통 이용자와 노선수, 정류장수, 단거리 노선비율, 배차간격, 운영시간 등의 버스관련변수와 환승역수, 역사위치, 역간거리, 출입구수, 자전거면수, 주차면수를 대변하는 변수인 도시철도관련 요인들을 이용하여 주거역세권과 비주거역세권의 토지이용유형별 서울시 역세권 대중교통 이용수요의 영향인자를 실증분석 하였다. 분석결과 대중교통 이용수요에 영향력이 있는 변수들은 토지이용의 유형에 따라 서로 상이한 특성을 갖고 있음을 밝혀내었다. Sohn and Kim (2011)은 개발밀도를 측정할 수 있는 평균 용적률, 토지이용의 혼합을 측정하기 위한 변수인 개발용량, 필지면적, 주거대비 상업 및 업무용도 혼합비, 기타용도 포함여부, 보행편의성을 알아보기 위한 평균 경사도, 인구사회요인으로 분류할 수 있는 거주인구, 교육수준, 주택보유율의 변수를 이용하여 역세권의 도시공간특성을 분석하였다. 분석 결과를 살펴보면 거주인구, 개발밀도, 주거대비 상업용도 혼합비, 기타용도 포함여부(자동차관련시설, 종교시설)의 변수들이 유의미한 영향을 미치는 것으로 분석되었다. Kim et al. (2008)은 대중교통 서비스 취약계층의 공간적 분포 특성에 관하여 면적, 인구, 주택유형 등의 GIS 기반의 데이터를 사용하여 분석을 실시하였다. 분석결과 대중교통 서비스는 지역별로 차이가 크며, 농촌 및 교외 지역에서 취약계층이 많았음을 실증적으로 분석하였다. Song (1998)의 연구에서는 서울 및 수도권 지역의 통근거리를 조사하여 실제 통근거리의 49%가 직장과 주거지의 공간적 분포만으로는 설명이 되지 않는다는 결과를 분석하여 직장 근처에 거주지를 마련할 수 있는 사람들은 어떠한 이유로건 직장에서 멀리 떨어진 곳에 주거입지를 결정한다고 하였다. 이는 도시공간구조적인 측면과 더불어 정책적·제도적 측면에서 직주근접이 이루어져야 함을 나타내는 결과이다. 이 연구의 경우에는 도시공간구조와 통행행태에 관한 연구들은 도시구조특성 변수들을 기준에 사용하였던 접근법이 아닌 새로운 관점에서 접근하였다는 점에서 의의가 크다고 판단된다. 따라서 본 연구에서도 일반적인 도시구조특성 변수 이외의 자족성 측면의 변수들을 포함시키고자 한다.

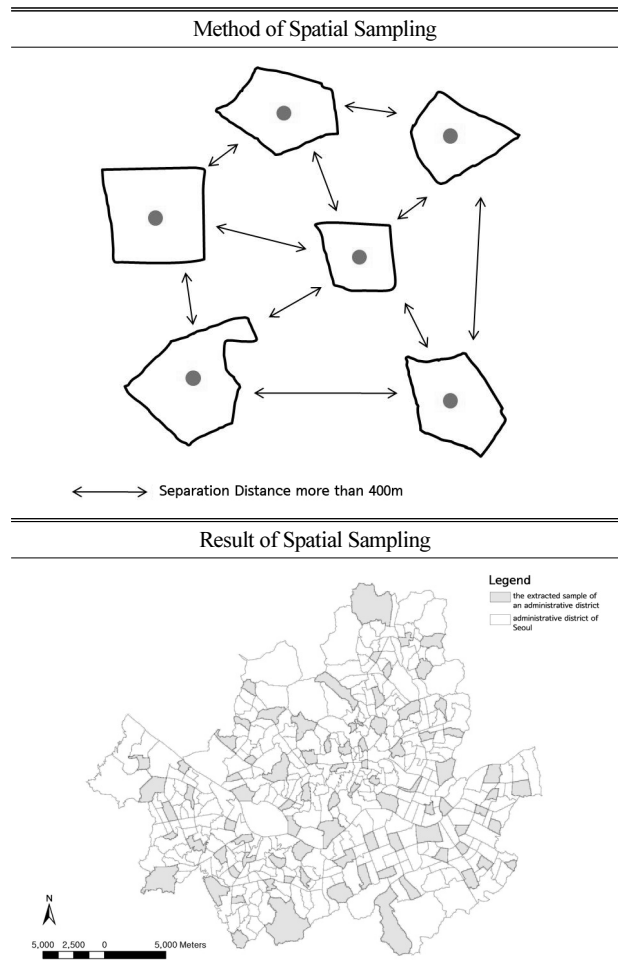
이처럼 기존 연구들에서 도시공간구조를 분석하기 위해 사용한 변수들을 살펴보면 크게 토지용도의 밀도, 토지이용의 혼합, 교통인프라, 소득수준 등의 변수들이 중요함을 확인할 수 있었다.

3. 분석의 틀 설정

3.1 공간샘플링(Spatial Sampling) 방법

본 연구에서는 서울시 행정동별 대중교통수단 분담률과 도시구조특성 변수 데이터 셋을 구축하였다. 이는 서울시 전체에 관한 데이터 셋을 구축한 것이기 때문에 지리적 공간의존성으로 자료를 왜곡하는 문제가 남아있음을 쉽게 짐작할 수 있다. 따라서 앞서 살펴본 여러 가지 공간샘플링 방법론 중에서 본 연구의 자료에 가장 적합하다고 판단되는 계통추출법을 변형한 모집단에서 일정 거리를 이격시키는 공간샘플링 방법을 사용하였다. 이 방법은 모집단의 한 지점에서 일정한 거리를 이격시켜 표본을 추출하였기 때문에 서로 인접해 있는 공간적 혹은 지리적 상관성에 의한 데이터의 왜곡을 억제할 수 있는 유용한 공간샘플링 방법이라 판단된다. 이를 위하여 ArcView 3.2프로그램의 ‘Random Sampling Generator’ 툴을 사용하여 표본 간의 최소 이격거리를 400m로 설정하여 84개의 행정동을 표본으로 추출하였다(Table 1). 본 연구에서 실시한

Table 1. Method and Result of Spatial Sampling



공간샘플링 결과 서울시 25개의 구에서 고르게 표본이 추출되었으며, 구별로 살펴보았을 때도 마찬가지로 고르게 최종적인 행정동 표본이 추출된 것을 확인할 수 있었다.)

3.2 변수의 설정

3.2.1 도시공간구조특성 변수

본 연구에서 사용한 대중교통 수단분담률에 관한 자료는 2010년 수도권장래교통수요예측 경신연구 자료를 사용하였으며, 대중교통으로 포함한 교통수단은 지하철을 포함한 철도교통수단과 버스교통수단으로 구성하였다. 가공을 거치지 않은 Low데이터는 수단교통 분담률로 구축이 되었다. 전체교통량에서 대중교통 수단분담률을 나누어 궁극적으로 사용하고자 하였던 분담률의 데이터를 구축하였다.

본 연구에서는 도시공간구조 특성 변수로 토지이용, 교통인프라, 지속성 관련 지표를 사용하였다. 토지이용 관련 변수로는 토지이용의 혼합정도를 측정하기 위한 엔트로피지수(Entropy Index)²⁾와 행정동의 주거용도의 밀집도를 측정하기 위한 주거밀도를 변수로 활용하였다. 토지이용 관련 변수로는 2010년 서울시 KLIS 데이터 상에서의 건축물과 토지의 용도에 대한 정보들을 사용하였다. 교통인프라 관련 변수로는 대표적인 대중교통인 지하철 이용의 편리성과 접근성을 측정하기 위한 지하철 진출입구수, 지하철까지의 거리를 측정하기 위한 지하철접근도, 행정동 내의 중요도로 비율을 측정할 수 있는 간선도로비율을 사용하였다. 인구사회학적 특성을 측정하기 위한 지표로는 상업 및 업무용도 인구밀도 측정을 위한 변수로서 인당 취업자수를 분석에 포함시켰으며, 미취업 인구나 대중교통 이용특성에 대한 영향을 알아보기 위한 변수로서 학생수(초중고학생수, 대학생수)를 분석에 포함하였다. 또한 해당지역의 소득수준을 측정하기 위한 대리변수(proxy)로 소득분위를 분석에 포함하였다. 지속성 관련 변수로는 직장과 주거지와의 거리를 측정하여 자동차 의존도를 측정하기 위한 직주비³⁾를 분석에 포함시켰으며, 주거지가 아닌 다른 지역으로 출퇴근 하는 통행으로 직장과 주거지의 분리정도를 파악하기 위한 교차통근율⁴⁾을 하나의 설명변

수로 설정하였다.

Table 2. Reclassification of Use

Devison	Reclassification
Residence	Apartment house, multi-family houses, multi-house, detached-house, tenement house, dormitory
Commerce	neighborhood facility, first class neighborhood facility, second class neighborhood facility, sales and business facility, sale facility
Business	public service facility, business facility

Table 3. Selected Independent Variables

Factor	Variable	Source
Land utilization	Entropy index	2010 Seoul KLIS data
	Residence density	2010 Seoul KLIS data
Traffic-infra	Number of subway station's gate	2010 Seoul KLIS data
	Access rate of subway	2010 Seoul KLIS data
	Rate of arterial road	2010 Seoul KLIS data
Characteristic of population and society	Income rank	Statistics Korea
	Number of employee per each	Metropolitan transportation demand forecast renewal study (Metropolitan Transport Association, 2010)
	First class station area passenger	yearly passenger boarding of each subway operating organization (Metropolitan Transport Association, 2010)
	Number of college student	Metropolitan transportation demand forecast renewal study (Metropolitan Transport Association, 2010)
	Number of elementary school, middle and high school	Metropolitan transportation demand forecast renewal study (Metropolitan Transport Association, 2010)
	Self-contentment	Rate of distance to the workplace
Cross commuting rate		Metropolitan transportation demand forecast renewal study (Metropolitan Transport Association, 2010)

- 1) 공간샘플링을 통하여 일정한 거리 간격으로 이격시킨 표본 행정동을 추출하였기 때문에 서울시 각 구별로 3~4개소의 행정동이 표본으로 추출되었다.
- 2) 토지이용의 혼합 요인인 주거연면적, 상업연면적, 업무 및 기타 연면적으로의 용도는 2008년 건축물 지상세 과세대상 데이터베이스의 건축물의 용도를 재분류하여 사용하였다.
- 3) 직주비 지표는 일정 공간에서 취업자수 또는 주택수에 대한 일자리수의 비율이며, 직주비가 1.25이상이면 고용초과지역, 0.75~1.25이면 직주 균형지역, 0.75이하이면 주거초과 지역으로 분류된다. 직주비 자료의 산출방법은 총 도차통근자수/총 출발 통근자수로 구해진다(Lee et al., 2012).
- 4) 교차통근율 지표는 교차통근은 주거지가 아닌 다른 지역으로 출퇴근하는 통행으로 직장과 주거지가 얼마나 많이 분리되어 있는지 파악하기

위한 것으로 산출방법은 외부출근통행량/내부출근통행량으로 구해진다(Lee et al., 2012).

3.2.2 종속변수

본 연구에서 설정한 종속변수는 2006년 수도권장래교통수요예측 경신연구를 기반으로 구축된 수도권 O-D자료 중에서 철도교통수단 통행발생량과 버스교통수단 통행발생량을 합산하고 이를 총 통행발생량으로 나누어 산출한 대중교통수단분담률 자료를 사용하였다.

3.3 현황분석

각 변수별 현황을 서울시 전체 행정동과 공간샘플링으로 추출된 표본 행정동과의 차이점은 거의 없는 것으로 나타났기 때문에 본 연구에서 사용한 공간샘플링은 적절히 이루어졌음을 짐작할 수 있다. 본 연구에서 사용된 변수들의 현황을 파악해보면 종속변수로 사용되는 대중교통수단분담률은 평균 0.56으로 나타나 사람들이 이용하는 교통수단 중 절반 이상을 대중교통이 이용되는 것을 확인할 수 있었다. 엔트로피 지수의 평균이 0.56~0.58로 토지이용의 혼합 정도는 보통수준이거나 양호하게 혼합되어 있는 것을 알 수 있다. 주거밀도가 0.65로 도시지역의 65% 수준으로 주거지가 분포되어 있는 것을 확인할 수 있었다. 교통인프라와 관련된 지하철 진출입구수는 평균 3개소 이상이 되며, 접근도가 1km 이상의 거리에 위치하고 있으며, 간선도로율은 0.29로 약 30%가량 도시의 간선도로가 설치되어 있음을 알 수 있었다. 소득분위는 4.36으로 평균적으로 서울시에는 중산층이 살고 있었으며, 인당 취업자수는 0.22~0.28로 취업률은 다소 저조하였으며, 역세권의

이용인구가 상당히 많았다. 직주비와 교차통근율과 같은 자족성 지표들은 비교적 표준편차가 크게 나타나고 있는 것으로 미루어 볼 때 서울시는 행정동별로 상반되는 형태들이 많을 것을 유추해 볼 수 있었다.

3.4 공간샘플링 전후 공간자기상관의 측정

공간자기상관이란 어떠한 값들이 공간상에 분포할 때 무작위적으로 분포하지 않고 지역적으로 인접할수록 비슷한 값을 갖거나(정적 공간자기상관) 또는 반대의 값을 갖는(부적 공간자기상관) 것을 의미한다(Lee et al., 2012). 본 연구에서는 공간통계패키지 GeoDa 1.0.1 프로그램의 공간분석 툴을 이용하여 공간샘플링 전후의 공간자기상관성을 측정하였다. 공간자기상관성을 측정하기 위한 Moran's Index 계수값을 비교해보면, 공간샘플링을 실시하지 않은 서울시 전체 행정동의 대중교통분담률의 Moran' I는 0.3387로 공간자기상관성이 상당히 높은 것으로 분석된 반면, 공간샘플링을 실시한 후의 Moran' I는 0.0625로 공간자기상관성이 거의 존재하지 않았다. 본 연구에서 제안하는 방법론을 사용하여 공간샘플링을 실시할 경우 공간자기상관성을 배제하여 분석을 실시할 수 있는 장점이 있다는 것을 실증적으로 확인할 수 있었다. 공간샘플링을 통하여 공간적 또는 지리적 데이터의 공간의존성에 대한 왜곡을 시키지 않는 범위에서 분석을 실시할 수 있음을 확인한 결과라 할 수 있다(Table 5).

공간자기상관이 발생하는 지역적 위치를 파악하기 위한 국지적

Table 4. Analysis of Present Condition

Variable	Status of total administrative district of Seoul			Status of total administrative district of Seoul after spatial sampling		
	Min	Max	Ave.	Min	Max	Ave.
Share rate of public transport	0.21	0.86	0.56	0.21	0.81	0.56
Entropy index	0.01	0.98	0.58	0.01	0.97	0.56
Residence density	0.00	1.00	0.65	0.00	1.00	0.65
Number of subway station's gate	2	28	3.58	1	20	3.19
Access rate of subway	21	23965	1172.32	21	23965	1158
Principal road rate	0.01	2.03	0.29	0.01	2.03	0.29
Income rank	2.24	9.63	4.36	2.57	8.03	4.41
Employement rate	0.04	0.96	0.28	0.04	0.88	0.22
First class station area passenger	12622	96181926	8085193	12622	53363948	8670839
Number of college student	34	33191	928.34	34	33191	1108
Number of elementary, middle and high school student	64	15003	3396.90	64	13823	3289
Rate of distance to the workplace	0.0006	0.096	0.024	0.0006	0.096	0.029
Cross commuting rate	2.34	812.30	42.84	5.16	176.41	41.20

Table 5. Comparison of Correlation Coefficient Around Sampling

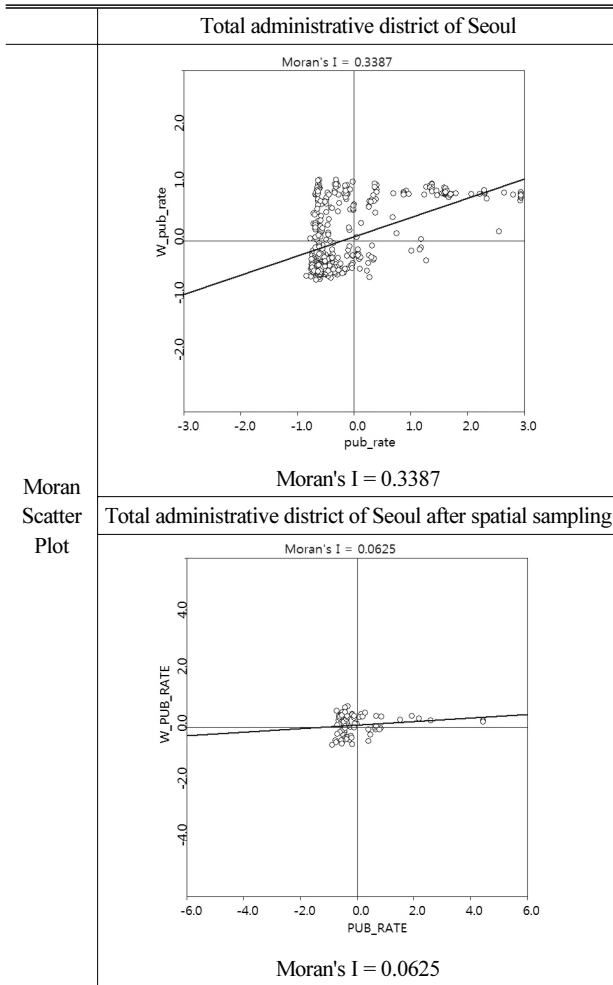
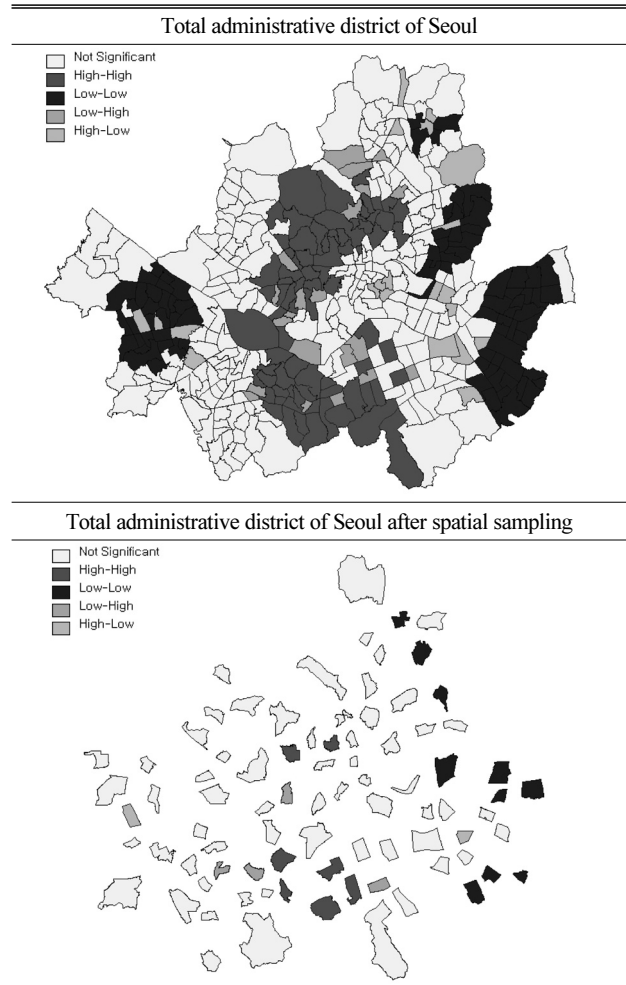


Table 6. Cluster Map



공간자기상관분석인 LISA (Local Indicator of Spatial Association)⁵⁾ 분석을 실시하였다. LISA 분석을 통해 유의수준 0.05 이내에서 공간자기상관이 존재하는 지역적 케이스(행정동)에 주변지역과 유사하게 높은 값(High-High)을 나타내는 핫스팟(Hot Spot) 지역과 주변지역과 유사하게 낮은 값(Low-Low)을 나타내는 콜드스팟(Cold Spot) 지역을 찾아주는 클러스터맵(Cluster Map)을 구축할 수 있다(Lee et al., 2012).

대중교통수단분담률의 핫스팟 지역은 종로구, 중구, 여의도, 서초구 일대의 주로 도심지역으로 나타났으며, 관악구, 동작구 등의 한강 이남지역을 중심으로 분포하고 있다. 종로구와 중구는 대중교통시의 배치가 발달된 곳이며, 여의도는 환승정류장 등의 여러 대중교통 노선들이 교차하는 지점이다. 또한 동작구와 관악구,

서초구 일대는 학원 및 학교가 밀집된 지역으로 주로 대중교통을 이용하는 학생들의 통행량이 가장 높은 지역으로 꼽힌다. 특히 이러한 지역은 대중교통 인프라가 상대적으로 발달된 지역으로, 대중교통 시설의 적절한 배치와 높은 접근성으로 인하여 핫스팟 지역으로 분류된 것으로 판단된다. 한편 대중교통분담률의 콜드스팟 지역은 강서구, 중랑구, 강동구 등의 도심지역으로 나타났으며, 이들 지역은 주로 서울시의 중심부가 아닌 외곽지역으로 상대적으로 대중교통시설이 적게 분포되어 있는 곳이다. 또한 도시외곽에 위치하고 있기 때문에 접근성 또한 좋지 않아 콜드스팟으로 분류된 것으로 판단된다. 이는 공간샘플링 전과 후 모두 비슷한 결과가 나타났으나 공간샘플링 후 강서구의 경우 콜드스팟에서 제외되었다(Table 6).

5) 지역 안에서도 공간자기상관이 위치에 따라 상이할 수 있는데, 지역 내에서 발생할 수 있는 공간자기 상관의 국지적 변이를 고려한 시각적 지표가 LISA이다(Kim, 2003).

3.5 공간샘플링 전후 회귀모형의 측정

본 연구에서는 정규분포를 이루지 못한 변수들(지하철 진출입구

수, 지하철접근도)에 한하여 로그(Log)변환을 실시하여 분석하였다. 대중교통 분담률에 관한 회귀모형을 살펴보면 공간샘플링 후의 회귀모형의 설명력(R-Square 0.372)이 공간샘플링 전의 회귀모형(R-Square 0.252) 보다 10%이상 높은 설명력을 보이고 있다. 설명변수들 사이의 다중공선성을 진단하기 위하여 분산팽창계수(VIF)를 살펴본 결과 모두 3 이하로 우려할 만한 수준의 문제는 없는 것으로 확인되었다. 회귀모형 결과 토지이용의 혼합이 적절히 이루어지고, 지하철 진출입구와 같은 대중교통시설이 많으면 대중교통의 이용이 증가한다는 일반적인 사실과 부합된 결과가 도출되었다. 또한 대학생 및 역세권 이용인구가 많을수록 대중교통이 활성화 된다는 것을 확인할 수 있었다. 그러나 초중고학생의 경우 근린주구 안에서 거주지와 학교까지의 동선이 비교적 짧기 때문에 대중교통에 부(-)의 영향을 미치는 것으로 판단된다. 자족성 지표에 관한 분석 결과는 통근통행발생량이 높은 지역에서 대중교통의 이용이 증가한다는 결과가 도출되었다. 이는 자족성이 상대적으로

높은 지역 보다 낮은 지역에 거주하는 사람들의 대중교통 이용률이 높을 것임을 짐작할 수 있으며, 거주지와 다소 먼 거리에 위치한 직장을 다니는 사람들도 대중교통의 이용이 활발한 것을 의미한다.

공간샘플링 전의 회귀모형에서는 엔트로피지수, 지하철 진출입구수, 인구, 자족도에 관련된 변수들이 유의미한 결과를 나타낸 반면, 샘플링 후 회귀모형에서는 인구 관련 변수와 자족도 관련 변수들만 유의미한 결과를 나타냈다. 공간의존성을 배제하고 분석을 실시한 결과 지역적 또는 지리적으로 연계가 중요한 변수들인 토지이용의 혼합도 및 대중교통 인프라 관련 변수들이 유의미한 변수들에서 제외된 것은 논리적으로 타당성이 있는 결과라 판단된다. 이를 통하여 알 수 있는 것은 토지이용과 대중교통 인프라는 대중교통을 활성화시키기 위하여 필수불가결한 요인들이지만, 지리적으로 연계가 이루어지지 않았을 경우에는 자칫하면 그 기능을 상실할 수 있다는 결과이기 때문에 중요한 발견이라 사료된다. 따라서 대중교통 이용율을 증가시키기 위하여 이러한 요소들의 배치가 중요할 것이며, 특히 지역적으로 연계를 이루는 토지이용 배치와 대중교통시설의 배치가 중요할 것이다. 특히 본 연구의 분석결과에 의하면 공간샘플링 전에는 토지이용의 혼합, 교통인프라 등의 변수들이 대중교통분담률에 유의미한 영향을 미치는 것으로 파악되었으나, 공간샘플링 후에는 유의미하지 않은 것으로 분석되었다. 이는 공간샘플링을 실시하지 않으면 공간자기상관성으로 인하여 결과의 왜곡이 나타나는 현상이다. 즉, 공간적 자료의 공간분석 시에는 공간의존성을 제거해주는 공간샘플링을 실시한 후 분석을 수행하여야만 결과의 왜곡이 나타나지 않는다는 것을 공간샘플링 전후의 회귀모형을 통하여 확인할 수 있었다.

Table 7. Regression Model Result

	Regression model before Spatial Sampling			Regression model after Spatial Sampling		
	β	Sig.	VIF	β	Sig.	VIF
N	419			84		
R ²	0.252			0.372		
Variable	β	Sig.	VIF	β	Sig.	VIF
Entropy index	1.118	*	1.682	0.098	-	1.966
Income rank	0.067	-	1.125	0.060	-	1.155
Number of subway station's gate ¹⁾	2.174	*	1.842	0.020	-	2.183
Access rate of subway ¹⁾	-1.328	-	1.074	-0.041	-	1.101
Principal road rate	1.452	-	1.070	0.131	-	1.076
Employement rate	0.192	-	1.157	-0.009	-	1.431
Residence density	0.593	-	1.863	0.019	-	2.139
Number of elementary, middle and high school student	-4.348	**	1.098	-0.165	-	1.136
Number of college student	4.667	**	1.058	0.171	*	1.094
First class station area passenger	3.649	**	2.007	0.398	**	2.418
Rate of distance to the workplace	-2.314	**	1.364	-0.127	-	1.518
Cross commuting rate	3.894	**	1.764	0.220	*	1.485

* Significant 90%

** Significant 95%

¹⁾ Log Transformation

4. 요약 및 결론

도시 분야에서의 공간적 또는 지리적인 성격을 지니고 있는 자료들이 대부분이다. 이러한 자료들은 필시 공간자기상관이라 불리는 공간의존성을 지니고 있다. 그러나 공간자기상관을 무시한 채 도시공간에 대한 연구를 진행하여 결과의 왜곡을 초래하는 연구들이 많았다. 본 연구에서는 공간적 자료를 바탕으로 공간의존성을 탐색한 후 일정 거리를 이격시키는 공간샘플링 기법을 이용하여 자료의 왜곡을 하지 않는 범위 내에서의 공간분석 방법론에 대하여 대중교통 분담률에 영향을 미치는 도시공간영향 요인을 추출함으로써 실증적인 분석을 하였다.

본 연구에서는 공간자기상관성을 컨트롤 및 제거해주는 공간샘플링 방법론을 사용하였으며, 이를 바탕으로 정립된 회귀모형 결과로부터 다음과 같은 시사점을 도출하였다.

첫째, 공간적 또는 지리적 자료가 지니고 있는 공간자기상관성을 제거하기 위해서는 적절한 공간샘플링 방법을 사용하여야 한다는

점이다. 서울시 전체 행정구역별 대중교통 분담률에 관한 공간자기상관성을 측정된 결과 Moran's Index 0.3387로 나타나 공간의존성이 상당히 높은 것으로 분석되었다. 하지만 적정한 거리(400m)이상을 이격 시켜 공간샘플링을 실시한 후 측정된 Moran' Index 0.0625로 나타나, 공간자기상관성이 제거됨을 확인하였다. 이를 통하여 도출되는 시사점으로는 공간샘플링을 실시하지 않으면 공간적 자료의 왜곡으로 인한 잘못된 결과가 도출되며, 이를 해결하기 위해서는 공간샘플링을 하여 표본을 추출하는 과정이 필수불가결하다는 점이다.

둘째, 공간샘플링 전후의 대중교통 분담률에 영향을 미치는 도시공간구조 특성 요인이 달라진다는 점이다. 공간샘플링 전 서울시 전체 행정구역을 대상으로 한 회귀모형의 결과를 살펴보면 토지이용의 혼합, 교통인프라, 인구 및 사회, 자족성의 대부분의 지표에서 모두 대중교통 이용에 영향을 주는 것으로 나타났다. 그러나 이는 공간적 자료에 존재하는 공간자기상관성을 통제하지 못한 왜곡된 결과이다. 따라서 공간샘플링을 실시한 후 분석된 결과에 의하면 인구 및 사회, 자족성 지표가 대중교통 이용에 영향을 미치는 것으로 나타났다. 실질적으로 인접한 공간상에 존재하는 공간자기상관성을 통제할 경우에 개인의 특성과 관련되어 있는 지표들과 대중교통 이용과 상관성이 있다는 결과이다. 개인의 활동은 인접한 공간상에서 주로 이루어진다는 점에서 공간의존성을 통제한 경우의 결과가 달라짐을 확인할 수 있는 연구 결과가 사료된다.

감사의 글

이 논문은 2013년 한국도시계획학회 학술대회에 발표한 내용을 수정·보완한 논문입니다.

References

- Cervero, R. and Radisch, C. (1996). "Mixed land uses and commuting: Evidence form the American Housing Survey." *Transportation Research Part A*, Vol. 30, No. 5, pp. 361-337.
- Chung, J. K. (2006). "A study on a decision the number of strata with stratified sampling." *Korea Institute of traditional commercial, Korea Traditional commercial academic research*, Vol. 20, No. 2, pp. 115-121 (in Korean).
- Doopedia (2013). Available at: www.doopedia.co.kr.
- Doreian, P. (1981). "Estimating linear models with spatially distributed data." *Sociological Methodology*, 12, pp. 359-388.
- Ewing, R. and Cervero, R. (2011) "Travel and the built environment: A Synthesis." *Transportation Research*, No. 1780, pp. 87-114.
- Ha, C. H. (2005). *A study on spatial structure analysis in a comurbation region using spatial autocorrelation technique*, Doctorate Thesis, Gyeongsang National University.
- Haining, R. (1993). "Spatial data analysis in the social and environmental sciences." Cambridge University Press.
- Irwin, E. (2002). "The effects of open space on residential property values." *Land Economics*, Vol. 78, No. 4, pp. 465-480.
- Isard, W. (1956). "Location and space economy." Cambridge: MIT Press.
- Kim, G. K. (2003). "Spatial autocorrelation and spatial regression take advantage of the search." *Korean Public Management Review*, The Korea Association for Policy and Evaluation, Vol. 13, No. 1, pp. 273-306 (in Korean).
- Kim, G. S. (1996). "A Study on the Theoretical spatial structure of the city." *Korean Public Management Review*, The Korea Association for Local Government & Administration Studies, Vol. 13, pp. 199-213 (in Korean).
- Kim, J. I., Kang, S. K. and Lwon, J. H. (2008). "The spatial characteristics of transit-poor in urban areas." *Journal of the Korean Association of Geographic Information Studies*, The Korean Association of Geographic Information Studies, Vol. 11, No. 2, pp. 1-12 (in Korean).
- Lee, Y. S., Jin, C. J. and Choo, S. H. (2012). "A study on spatially influencing factors ablut public transportations using spatial analysis: A case of Seoul, Korea." *Seoul Studies*, The Seoul Institute, Vol. 13, No. 4, pp. 467-472 (in Korean).
- Oh, Y. T., Kim, T. H., Park, J. J. and Rho, J. H. (2009). "An empirical analysis of influencing factors toward public transportation demand considering land use type seoul subway station Area in Seoul." *Journal of the Korean Society of Civil Engineers*, Korean Society of Civil Engineers, Vol. 29, No. 4, pp. 467-472 (in Korean).
- Park, E. T. (2011). *Economics dictionary*, Genyunsa (in Korean).
- Sohn D, W. and Kim, J. (2011). "An analysis of the relationship between the morphological characteristics of transit centers and transit riderships in seoul metropolitan region." *Journal of the Architectural Institute of Korea Planning & Design*, Architectural Institute of Korea, Vol. 27, No. 6, pp. 177-184 (in Korean).
- Song, M. R. (1998). "Urban spatial stucture and excess commuting in the Seoul Metropolitan Area." *Journal of Korea Planners Association*, Korea Planners Association, Vol. 33, No. 1, pp. 57-75 (in Korean).
- Sung, H. G., Kim, D. S. and Park, J. H. (2008). "Impacts of land use and urban design characteristics on transit ridership in the seoul rail station areas." *Journal of Korea Society of Transportation*, Korea Society of Transportation, Vol. 26, No. 4, pp. 135-147 (in Korean).
- Sung, H. G., Roh, J. H., Kim, T. H. and Park, J. H. (2006). "A study on the effects of land use on travel pattern in the rail station areas of a dense city: A Case of Seoul." *Journal of Korea Planners Association*, Korea Planners Association, Vol. 41, No. 4, pp. 59-75 (in Korean).