

SARIMAX 모형을 이용한 공공자전거 수요추정과 평가: 서울시의 COVID-19 영향을 중심으로

Estimation of Shared Bicycle Demand Using the SARIMAX Model: Focusing on the COVID-19 Impact of Seoul

홍 정 열 * · 한 은 룡** · 최 창 호*** · 이 민 서**** · 박 등 주*****

* 주저자 : 서울시립대학교 교통공학과 연구교수
 ** 공저자 : 서울시립대학교 교통공학과 석사과정
 *** 공저자 : 전남대학교 불류통상학부 교수
 **** 공저자 : 서울시 자전거정책과 주무관
 ***** 교신저자 : 서울시립대학교 교통공학과 교수

Jungyeol Hong* · Eunryong Han* · Changho Choi** · Minseo Lee***
 · Dongjoo Park*

* Dept. of Transportation Eng., Univ. of Seoul
 ** Division of Logistics and International Trade, Chonnam National Univ.
 *** Division of Bicycle Policy, Seoul Metropolitan Government
 † Corresponding author : Dongjoo Park, djpark@uos.ac.kr

Vol.20 No.1(2021)

February, 2021
 pp.10~21

pISSN 1738-0774
 eISSN 2384-1729
<https://doi.org/10.12815/kits.2021.20.1.10>

Received 17 December 2020
 Revised 31 December 2020
 Accepted 11 January 2021

© 2021. The Korea Institute of
 Intelligent Transport Systems. All
 rights reserved.

요 약

COVID-19 발병으로 세계는 심각한 위기에 직면해 있으며, 각 국에서는 전염병 확산을 감소시키고 안전한 통행환경을 조성하기 위하여 사회적 거리두기가 가능한 공공자전거 활성화에 대한 관심이 높아지고 있다. 본 연구는 COVID-19 시대에서 공공자전거의 공급정책, 전염병의 확산 등의 외생요인들이 공공자전거 이용수요에 어떠한 영향을 미치는지 분석하고, 이를 반영한 장래수요예측 방법론을 제시하는데 주요 목적이 있다. 기존 시계열 모형이 가지고 있는 외생요인 미반영의 한계점을 보완하기 위하여 SARIMAX 방법론을 제시하였다. 본 분석을 통하여 외생변수인 COVID-19 발병률과 공공자전거 공급량이 공공자전거 이용수요와 양(+)의 관계에 있다는 것이 통계적으로 유의하게 나타났다. 2022년까지 4만5천대의 공공자전거가 공급되고, COVID-19의 발병률이 현재 수준으로 지속될 경우, 서울시 공공자전거는 2021년 연간 약 3천2백만 대, 2024년에 약 4천6백만 대의 이용수요가 발생할 것으로 예측되었다.

핵심어 : 공공자전거, COVID-19, 외생변수, 장래수요예측, SARIMAX

ABSTRACT

This study analyzed how external variables, such as the supply policy of shared bicycles and the spread of infectious diseases, affect the demand for shared bicycle use in the COVID-19 era. In addition, this paper presents a methodology for more accurate predictions. The Seasonal Auto-Regulatory Integrated Moving Average with Exogenous stressors methodology was applied to capture the effects of exogenous variables on existing time series models. The exogenous variables that affected the future demand for shared bicycles, such as COVID-19 and the supply of public bicycles, were statistically significant. As a result, from the supply volume and COVID-19 outbreak according to the scenario, it was estimated that approximately 46,000 shared bicycles would be supplied by 2022, and the COVID-19 cases would continue to be at the current level. In addition,

approximately 32 million and 45 million units per year will be needed in 2021 and 2024, respectively.

Key words : Shared bicycle, COVID-19, Exogenous variables, Demand estimation, SARIMAX

I. 서 론

도시 내 교통 혼잡 및 환경오염 문제를 해결하기 위하여 세계 주요 도시들은 공유 퍼스널 모빌리티(shared personal mobility)를 활성화하는데 노력하고 있다. 특히 공공자전거는 퍼스널 모빌리티의 대표적 통행수단으로 단거리 통행, 대중교통 시설로의 접근을 위해 주로 이용되며, 지하철, 버스, 자가용의 보완재 또는 대체재 역할도 하고 있다(Choi and Jung, 2020). 과거에는 서울시를 비롯한 대부분의 국내 대도시들이 버스, 지하철 등 대중교통 네트워크를 체계적으로 구축하여 시민들의 중·장거리 이동성을 향상하고자 하였다. 하지만 출발지로부터 대중교통 이용 장소까지 보행거리가 길거나 대중교통 서비스 공급이 부족한 지역은 연계 교통수단이 매우 필요한 실정이다. 이에 따라 공공자전거와 같은 공유 퍼스널 모빌리티 수단을 활용하여 라스트 마일을 실현할 수 있는 교통 시스템 마련이 요구되는 추세이다.

2015년 서울시는 도시 내 공유경제 서비스를 제공함과 동시에 라스트마일 및 교통·환경문제를 해결하기 위한 방안으로서 공공자전거 ‘따릉이’를 도입하였다. 도입 당시 연간 약 114천 건에 불과하던 공공자전거 이용량은 꾸준히 증가하여 2019년에는 19,071천 건으로 나타났다(Seoul open data portal, 2020). 이와 같은 공공자전거 이용량의 증가로 시민들은 운영·관리의 효율화, 안전성 향상, 고품질의 서비스 제공의 확대를 기대하게 되었고 관리자 및 의사결정자는 적절한 공급 및 수요관리를 위해 더욱 효율적인 운영 전략 수립에 노력을 기울이고 있다.

일반적으로 공공자전거의 장래 이용수요를 예측하고 적절한 공급량을 결정하기 위해서는 공공자전거의 이용 추세뿐만 아니라 요금, 계절성, 인구 등 다양한 외생 및 내생적 요인들이 고려되어야 한다. 특히 공공자전거는 공급량이 증가할수록 이용수요도 증가하는 선순환 패턴을 보이므로 장래 이용수요를 예측하기 위해서는 공급량 수준에 대한 반영이 반드시 필요하다. 그러나 2020년 1월 중순 국내에 COVID-19 확진자가 발생하고 이후 전국적으로 확산하면서 사람들의 통행 빈도와 통행행태, 교통수단 선택 등에 변화가 발생하였고, 이는 장래 교통수요 예측에 불확실성을 키우고 있다. COVID-19 확산이 통행행태에 미친 변화를 분석한 연구 결과에 따르면(Choi and Jung, 2020), 도로교통 이용자는 소폭 감소한 것에 비해 대중교통 이용자는 급격히 감소한 것으로 나타났다. 다만, 공공자전거와 같은 공유교통수단은 대중교통과 비교하였을 때 COVID-19의 영향을 크게 받지 않고 있다(Seoul open data portal, 2020). 하지만 이러한 결과만으로 앞으로도 서울시 공공자전거를 포함한 공유교통수단이 COVID-19와 같은 전염병 확산에 영향을 받지 않을 것이라는 결론을 내리기도 어렵다.

따라서 본 연구는 COVID-19 시대에 장래 공공자전거 이용수요예측 모형을 추정하고, 이를 기반으로 적절한 공급 수준을 제시하는 데 목적이 있다. 따라서 공간적 범위는 서울시 전체로 설정하고, 2015년부터 2020년까지의 공공자전거 이용내역 데이터를 활용하여 2024년까지의 공공자전거 수요를 예측하도록 시간적 범위를 정하였다. 수요예측모형은 현재의 COVID-19 상황과 공공자전거 공급 수준의 변화, 그리고 공공자전거 이용량 추이 등을 반영하여 도출코자 하였다.

II. 선행연구 고찰

1. 공공자전거 수요예측 연구

최근에 국내외에서 수행된 공공자전거 수요예측 연구는 <Table 1>과 같다. Kaltenbrunner et al.(2010), Vogel et al.(2011), O'Brien et al.(2014), Rudloff and Lackner(2014) 등이 수행한 초기 단계 연구는 이용 데이터에 기반한 패턴분석으로 장래 수요를 예측하였다. 하지만 <Table 1>에 정리한 최근 연구에서는 고도화된 통계 방법과 머신러닝 기법 등을 적용하여 대여소별, 시간대별 등 미시적인 시공간 단위의 수요예측이 이루어지고 있다.

Fourier et al.(2017)은 공공자전거 이용 자료의 확보 어려움뿐만 아니라 이용자들의 성향과 계절적 특성 등에 영향을 크게 받아 수요예측과정이 매우 복잡함을 보였다. 특히 분석대상 대여소의 공공자전거 이용량이 적을 경우 예측의 어려움을 극복하고자 사인파함수(sinusoidal function)를 이용하여 데이터를 보정한 다음 월 평균이용량과 연평균이용량을 예측하는 모델을 개발하였다. Lim and Chung(2019)는 계절성을 반영한 Holt winters 평활기법을 이용하여 공공자전거 일별 이용수요를 예측하는 모델을 개발하였다. 이 연구는 공공자전거 이용수요가 계절적 특성을 보이는 것에 초점을 두었으며, 이러한 특성을 지닌 공공자전거의 장래 수요를 예측하기 위해서는 더욱 강건성(robustness)을 지닌 시계열 모형이 필요하다고 제시하였다. Cho et al.(2020)는 딥러닝모형과 시계열분석기법을 이용하여 서울시 공공자전거의 대여량을 예측하는 모델을 개발하였는데, 시간별 이용수요를 예측하기 위한 영향요소로서 대여소별 대여량, 기상정보, 지하철 이용량 등을 반영하였다.

<Table 1> Previous studies for the shared bicycle demand estimations

Author	Purpose	Results	Methods
Fournier et al.(2017)	Estimation of the monthly average daily bicycle counts and average annual daily bicycle counts using sinusoidal function	In most cities used as samples, the estimate of each monthly average daily bicycle counts results in a margin of error of less than 30%. Developing a prediction model using the sine wave function of the demand for public bicycles, which was difficult to predict when the bicycle counts was low.	Seasonal Demand Model Using Sine Wave Function
Kim and Jang (2017)	Estimation of Monthly Demand for Subway in Seoul Metropolitan Area Using ARIMA Model	Forecasting the Demand for the Use of One Transportation (urban railway) Network with the Change of Macro-Traffic Conditions	SARIMA
Lim and Chung (2019)	Development of a Model for Predicting Demand for Public Bicycles Using Holt Winters Method with Seasonality	Development of a Demand Prediction Model with Robustness Reflecting the Seasonal Characteristics of Public Bicycles	Holt-Winters method
Cho et al.(2020)	Development of a Model for Predicting Demand of Public Bicycles in Seoul Using Deep Learning Models	Data sets were composed of rental data, weather data, and subway usage data, and forecast errors were compared using ARIMA and LSTM-based deep learning techniques.	Exponential smoothing, ARIMA, LSTM

2. COVID-19와 교통수요 변화 연구

COVID-19와 교통 패턴 변화 간의 관계를 고찰한 국내외 연구들을 <Table 2>과 같다. COVID-19확산 이후 대중교통과 코로나의 영향 관계를 분석한 연구들은 다수 있었으나(Wilbur et al., 2020; Tirachini and Cats, 2020; Zheng et al., 2020), COVID-19와 공공자전거 이용량 간의 연관을 분석한 연구는 매우 미미한 실정이다.

Chai et al.(2020)의 연구에서는 군집분석과 시계열분석을 이용해 COVID-19가 발생한 이후 중국 베이징의 전년 동월일 공공자전거 수요를 비교하였고, 그 결과 사회활동과 생산활동을 위해 이용되던 공공자전거 수요가 약 40% 감소한 것으로 나타났다. Teixeira and Lopes(2020)의 연구는 미국 뉴욕시에 COVID-19 발생 이후 지하철 수요는 크게 감소한 반면 공공자전거 수요는 비교적 적게 감소한 것을 발견하였으며, 이를 통해 공공자전거가 COVID-19 대응에 효과적인 교통수단으로 향후 지하철 수요를 대체할 가능성을 제안하였다.

이상의 선행연구 사례를 볼 때, 보다 체계적이고 효율적인 공공자전거 운영 전략을 수립하기 위해서는 먼저 도시 전체의 이용수요를 예측하고, 예측한 수요의 범위 내에서 권역별, 대역소별 공급의 재분배가 필요한 것으로 평가된다. 또한 과거부터 현재까지 거시적으로 공공자전거 이용수요에 영향을 준 요인들을 찾아 미래 수요예측에 적용해야 할 필요성도 제기된다.

이에 따라 본 연구는 장래 서울시 공공자전거 이용수요를 정책적 요인(공공자전거 공급), 시계열적 요인(계절 및 요일), 사회적 요인(COVID-19 확산) 등을 반영하여 예측하고, 이를 토대로 공공자전거 정책을 수행하는 의사결정자 측면에서 적절한 공급 규모를 제시하고자 한다.

<Table 2> Previous studies for the impact of COVID-19 on transportation

Author	Purpose	Results	Methods
Musselwhite et al. (2020)	The Effect of COVID-19 on Transportation and Health	The danger of COVID-19 propagation in public transportation used by countless people is significant.	Correlation analysis
Wilbur et al. (2020)	Effect of COVID-19 on Accessibility and Demand for Public Transportation	Bus demand fell from 66.9 percent to 65.1 percent. The difference in public-traffic demand for use between high-income area and low-income area is significant.	Data description
Tirachini et al. (2020)	the Relationship between COVID-19 and Public Transportation	It is necessary to adjust public transportation hours, make it mandatory to wear masks when using public transportation, and to implement public transportation hygiene and disinfection	Data description
Cho et al. (2020)	The Change of Traffic due to COVID-19 and the Direction of Transportation Policy in Seoul for Post Corona	Due to the influence of COVID-19, road traffic has decreased slightly and public transportation users have decreased dramatically.	Data description
Teixeira et al. (2020)	Identification of connection between public bicycles and subways due to COVID-19	When traffic volume has decreased overall due to COVID-19, the demand for subways has decreased very significantly, but the demand for public bicycles has decreased relatively small	Regression analysis
Chai et al. (2020)	Analysis of Behavioral Patterns of Public Bicycles by COVID-19	Compared to the demand for public bicycles on the same day last year since the outbreak of COVID-19, the demand for bicycles used in social and production activities has decreased by nearly 40 percent	Co-location analysis
Ruizhi Zheng et al.(2020)	Changes in the Use of Public and Private Transportation by the Spread of COVID-19	Public transportation has a high risk of spreading the COVID-19	Data description
Kim et al.(2020)	Changes in Transportation Policies since COVID-19	Due to the COVID-19, policies are in place to activate bicycles as a substitute for public transportation in foreign countries. Increased utilization of shared mobility during rush hour during weekdays after COVID-19	Data description

III. 분석 방법론

공공자전거는 계절 및 시간의 영향에 민감하므로 장래수요 추정시 이러한 추세, 순환, 계절변동을 반영할 수 있는 시계열분석에 기반을 두어야 한다. 시계열분석은 시간변수의 흐름에 따라 종속변수의 변화를 진단하고 예측하는 확률과정이며, 다양한 교통수단의 수요예측에 적용되어 왔다. Hong(2019)은 자기회귀이동평균(ARIMA: autoregressive integrated moving average)을 이용하여 대구국제공항의 항공수요를 예측하였으며, Cho and Baik(2013)은 해상교통 여객수요를 예측하는데 이동평균법(moving average)과 지수평활법(exponential smoothing)을 적용하였다. Kim and Park(2016)은 관광객 수요예측을 위해 시계열모형과 인과모형 방법론을 제시하였는데 다른 방법론에 비해 시계열 요인을 반영한 지수평활법(winters exponential smoothing)이 성능적으로 우수하다는 것을 주장하였다.

공공자전거는 계절과 시간 같은 시계열 요인 이외에 신규 공공자전거 공급량과 같은 외생적 요인에도 영향을 받는다. 특히 최근 발병이 시작되어 공공자전거 이용에 지속적인 영향을 미치는 COVID-19도 중요한 외생적 요인이므로 이들을 반영할 수 있는 방법론이 적용되어야 한다. 이러한 요인들을 반영하기 위해 본 연구에서는 시계열분석 기반의 SARIMAX(seasonal auto-regressive integrated moving average with exogenous regressors)모형을 적용하고자 하였다.

자기회귀(AR)모형은 현재 관측자료(y_t)가 과거자료(y_{t-1}, y_{t-2}, \dots)들과 비관측영향(α_t)의 부분의 선형결합으로 표시되는 모형이며 식(1)과 같이 시차변수(lagged variables)를 회귀변수와 같은 형태의 선형결합으로 구성될 수 있다는 가정에 기반을 두어 표현된다. 여기서 ϕ_i 는 자기회귀계수로 현재부터 i 번째 이전의 시간에 대한 가중치를 의미하며, 차수가 p 인 자기회귀모형을 $AR(p)$ 로 나타낸다.

$$AR(p): y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \alpha_t = \sum_{i=1}^p \phi_i y_{t-i} + \alpha_t \quad \dots \dots \dots (1)$$

이동평균(MA)모형은 현재 관측자료(y_t)가 현재와 과거의 비관측영향(α_t)의 선형결합으로 설명되는 모형이며 모형에서 비관측영향(α_t)는 IID(independent and identically distributed)를 따른다. MA모형은 AR모형과 유사하게 시차변수를 선형결합하는 형태이며, 현재부터 i 번째 이전 시간에 대한 오차항 가중치를 나타낸다. ϕ_i 는 이동평균계수이며 차수가 q 인 $MA(q)$ 모형은 식(2)와 같이 나타낸다.

$$MA(q): y_t = \alpha_t - \theta_1 \alpha_{t-1} - \dots - \theta_q \alpha_{t-q} \quad \dots \dots \dots (2)$$

ARIMA(auto-regressive integrated moving average)모형은 과거의 관측 값과 오차를 사용해서 현재의 값을 설명하되, 시간의 추이에 따라 평균 및 분산이 변하거나 시점 간의 공분산이 기준시점과 연관성을 보이는 비정상시계열(non-stationary time series)을 안정화하기 위하여 차분(d)과정을 포함한 모형이다. 또한 시계열 데이터가 계절성 또는 주기적 특성을 가지고 있는 경우 비계절성과 계절성 연산자 모두를 포함하는 SARIMA(seasonal ARIMA)모형을 적용한다. 식 (3)과 (4)는 각각 ARIMA모형과 SARIMA모형을 보여준다.

$$ARIMA(p, d, q): \phi_p(B)(1-B)^d y_t = \theta_0 + \theta_q(B) \alpha_t \quad \dots \dots \dots (3)$$

여기서, $\phi_p(B) = (1 - \phi_1 B - \dots - \phi_p B^p)$, $\theta_q(B) = (1 - \theta_1 B - \dots - \theta_q B^q)$, θ_0 는 deterministic trend이다.

$$SARIMA(p, d, q)(P, D, Q) : \phi_p(B)\Phi_P(B^s)(1 - B)^d(1 - B^s)y_t = \theta_q(B)\Theta_Q(B^s)\alpha_t \dots (4)$$

s는 계절주기, $\phi_p(B)$ 와 $\Phi_P(B)$, $\theta_q(B)$ 와 $\Theta_Q(B^s)$ 는 각각 비계절성 및 계절성 AR, MA의 연산자이다.

SARIMAX 모형은 SARIMA모형 과정에 외생변수들의 영향력을 반영하기 위해 선형회귀모형을 포함한 것이다. 이 방법론은 비관측된 외생변수들의 영향으로 계절성 패턴에 바이어스 관측이 의심될 때 유용하게 적용가능하며 다음 식(5)와 같이 정의된다(Hamilton, 1994; Tsay and Ruey, 2005)

$$w_t = y_t - \beta_1 x_{1,t} - \beta_2 x_{2,t} - \dots - \beta_b x_{b,t}$$

$$(1 - \sum_{i=1}^p \phi_i L^i)(1 - \sum_{j=1}^P \Phi_j L^{j \times s})(1 - L)^d(1 - L^s)^D w_t - \eta = (1 + \sum_{i=1}^q \theta_i L^i)(1 + \sum_{j=1}^Q \Theta_j L^{j \times s})\alpha_t \quad (5)$$

L은 시차 연산자, y_t 는 t시점에 관측자료, $x_{k,t}$ 는 t시점의 k번째 외생입력변수, β_k 는 k번째 외생 입력 변수의 계수값, b는 외생변수의 수, w_t 는 자기회귀 오차, p는 비계절성 AR 차수, P는 계절성 AR 차수, q는 비계절성 MA 차수, Q는 계절성 MA 차수, d는 비계절성 시계열의 차분 횟수, D는 계절성 시계열의 차분 횟수, η 는 SARIMA 모형의 상수항, α_t 는 t 시차의 오차항이다.

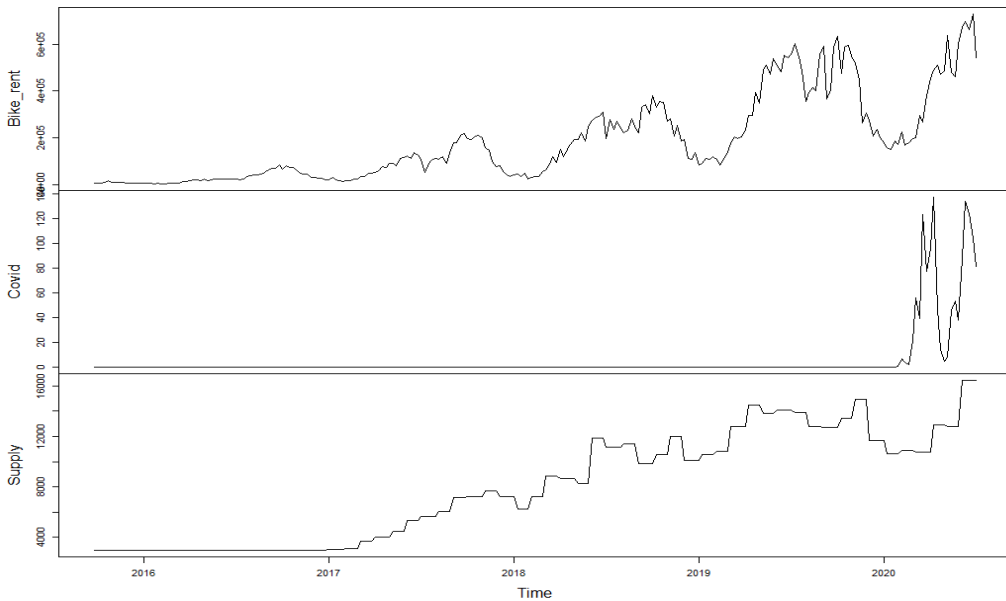
본 연구에서는 COVID-19 발병데이터, 공공자전거 공급량 데이터를 외생변수($x_{k,t}$)로 설정하여 SARIMAX 모형을 구축하고, 장래 공공자전거 이용량(y_t)을 예측하였다. 또한 모형으로부터 도출된 COVID-19 변수 및 공급량 변수의 통계적 유의성을 바탕으로 공공자전거 이용과 COVID-19 확산변화의 영향 관계, 공공자전거 공급량 정책과 공공자전거 수요와의 관계를 분석하고자 하였다.

IV. 자료수집 및 가공

SARIMAX를 이용한 장래 공공자전거 수요예측모형 구축을 위하여 서울시의 공공자전거 대여·반납건수, 공공자전거 공급대수, COVID-19 확진자 수 등을 수집하였다. 또한 데이터 분석결과의 일관성을 위하여 각 변수를 주 단위로 정리하였다. 자료 구축 기간은 2015년 9월부터 2020년 6월까지이나 변수별로 서울시 공공 데이터 제공 시점에 차이가 있다. 공공자전거 공급대수는 2015년 9월부터, COVID-19 확진자 수는 2020년 1월부터 제공된다. 수집된 외생변수들의 기초정보는 <Table 3>에서 설명되고 있다.

<Table 3> Descriptive statistics for exogenous variables

Variables	Min	1st Qu.	Mean	3rd Qu.	Max
Shared bike usage (bike/week)	2,570	34,767	186,109	275,745	729,923
Covid-19 patient (person/week)	0.00	0.00	5.24	0.0	137
Shared bike supply (bike/week)	5,316	5,316	14,333	21,018	29,500



<Fig. 1> Exogenous variables changes by time

<Fig. 1>에서 보여지듯이 공공자전거 이용은 도입 이후 꾸준히 증가해 오고 있으며 겨울철에 수요가 감소하였다가 봄, 가을철에 증가하는 계절성 패턴을 보인다. 특히 2018년부터는 이용수요가 급격히 증가하고 있는데, 이는 2017년 이후에 공공자전거 공급량 증가와의 관련성이 있는 것으로 해석되며, 두 자료는 시간적 간격을 두고 유사한 패턴을 보이고 있다. 2020년 1월 이후 공공자전거 이용량은 전년 대비 소폭 증가하였으나 COVID-19의 확산에 따라 과거의 증가 폭만큼 이용량이 늘어나지는 않은 것으로 나타났다. 이처럼 선택된 외생변수들은 공공자전거 이용수요에 영향을 미치며, 이러한 추세와 변화를 기반으로 비관측 영향요소들을 반영한 장래 수요예측이 가능한 것으로 판단하였다.

V. 모형 추정 및 분석

서울시 공공자전거 이용량과 각 외생변수들의 관계를 설명하는 시계열 기반의 SARIMAX모형을 구축하였다. 본 연구에서는 먼저 공공자전거 공급 시나리오에 따른 이용수요를 예측하고, 적정규모를 도출한 후 이를 바탕으로 COVID-19에 대한 시나리오를 적용하여 장래 2024년까지의 공공자전거 이용수요 변화를 예측하도록 하였다. 편자기상관계수, 자기상관계수, 정상성 여부, 계절성 등의 차수를 조합하여 여러 가지 모형을 구축하였으며, 이 중 각 계수들이 통계적으로 유의하고 모형의 AIC(Akaike's information criterion) 값을 기준으로 가장 적합한 모형을 선정하였다. AIC 추정식은 식(6)과 같다.

$$AIC = -2 \times \ln(\text{likelihood}) + 2 \times k \dots\dots\dots (6)$$

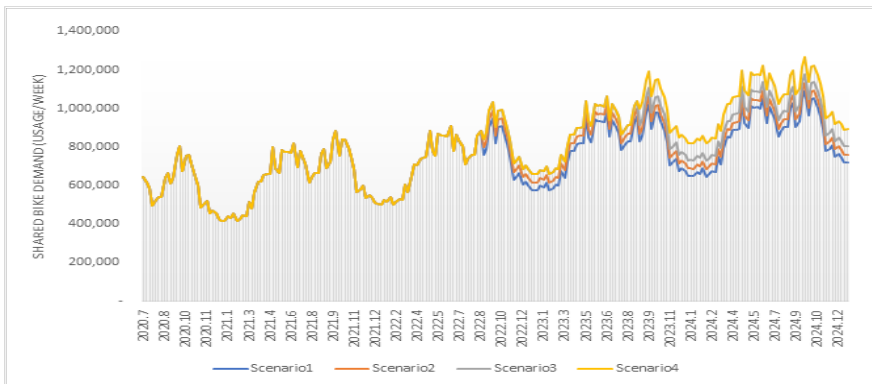
여기서 k 는 모수의 개수이다.

1. 공공자전거 공급시나리오 설정

공공자전거 공급에 따른 수요예측을 위하여 서울시 공공자전거 공급계획을 바탕으로 총 4가지 공급 시나리오(I)를 설정하였다. 2020년 현재까지 29,500대의 공공자전거가 공급되고 있는 상황에서 시나리오 I-1은 2020년까지 8,000대, 2021년까지 3,000대를 추가하여 총 40,500대를 공급하는 것이며, 시나리오 I-2는 2020년까지 8,000대, 2021년까지 3,000대, 2022년까지 4,500대를 추가하여 총 45,000대를 공급하는 방안이다. 시나리오 I-3은 2020년과 2021년은 시나리오 I-2와 동일하게 추가하지만 2022년에는 공급대수를 9,500대로 하여 총 50,000대를 공급하는 방안이다. 그리고 시나리오 I-4는 2022년까지는 시나리오 I-3과 같으나 2023년에 10,000대를 추가 공급하여 총 60,000대의 공공자전거가 공급되도록 하였다.

이상의 4가지 공급시나리오 설정에 따라 2021년에 예측되는 공공자전거의 연간 수요는 31,779,146건으로 도출되었으며 2024년에 시나리오 I-1은 약 4천6백만 건, 시나리오 I-2는 약 4천 5백9십만건, 시나리오 I-3은 약 4천7백만 건, 그리고 시나리오 I-4는 약 4천8백3십만 건으로 으로 추정되었다. 이와 같은 시나리오에 따라 향후 4년간 공공자전거 수요를 주단위로 나타내면 <Fig. 2>와 같다.

다음, 공공자전거 이용요금 수준이 현재와 같다는 전제로 공공자전거 1대당 운영비용 단가, 자전거 한 대당 구입비용, 거치대 1개소 당 비용단가를 각각 652,842원/대, 488,856원/대, 220,000원/대로 적용하여 비용대비 수입을 산출하였다. 결과는 시나리오 I-2의 비용대비 수입이 평균 1.12로 가장 효율성 있는 것으로 나타났다.



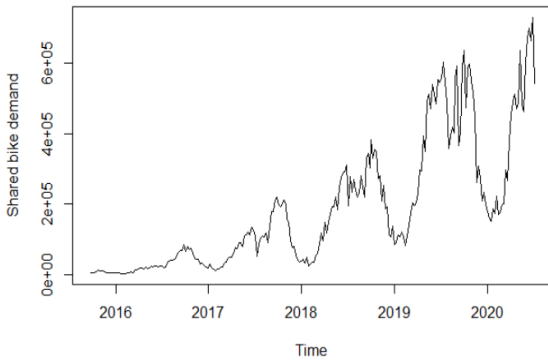
<Fig. 2> Results for shared bike demand estimation by supply scenarios

2. SARIMAX 모델 구축 결과

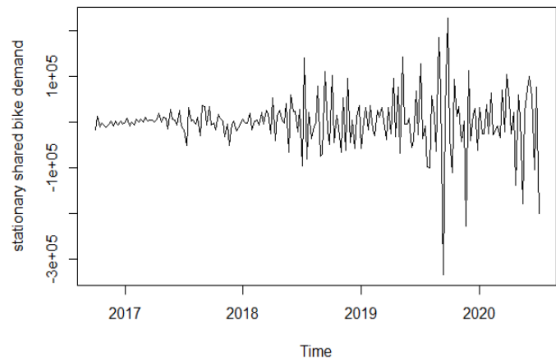
시계열 분석을 위해서는 종속변수와 관련 없는 요인들의 영향력이 배제되어야 하므로 시간에 따라 통계적 특징이 변화하지 않는 정상성(stationary) 시계열데이터를 사용해야 한다. <Fig. 3>에서 보여주고 있는 서울시 공공자전거 이용의 계절성 패턴, 추세성은 공공자전거 시계열데이터가 비정상성(non-stationary)의 특성을 가지고 있으며 모형구축 및 분석을 위하여 정상성 데이터로의 변환과정이 필요하다는 것을 시사하고 있다. 일반적으로 정상성 데이터로의 변환은 차분(differencing)으로 하며, 이를 통해 평균 변화를 일정하게 조절하고 추세와 계절성을 제거 또는 감소시키는 기능을 한다. 본 연구에서 차분한 정상데이터는 <Fig. 4>와 같다. 또한 시계열의 정상성을 검정하는 Dickey-Fuller Test를 통해 P-value 값 0.01을 도출하였으며, 이는 y_t 는 정상

1) 서울시시설공단 내부자료, 2020.11

성을 가지지 않는다는 귀무가설을 기각하여 데이터의 변환으로 정상성을 가지는 것을 확인할 수 있다.



<Fig. 3> Non-stationary shared bike data



<Fig. 4> Differenced and stationary data

공공자전거 공급 시나리오를 바탕으로 서울시 COVID-19 확진자 수를 외생변수로 추가한 SARIMAX 모델을 구축하였고, 최종 모형은 우연변동과 추세변동, 주기변동 등이 중첩되는 계절성 비정상시계열 모형인 SARIMA(0,1,2)(0,1,1)로 제시되었다. <Table 4>에 정리한 SARIMAX 모형 추정 결과와 같이 MA와 Seasonal MA항 등은 통계적으로 유의한 반면에, AR, Seasonal AR항들은 유의하지 않아 최종모형에 포함하지 않았다. 모형을 통하여 공공자전거의 현재 이용수요는 과거 시점의 수요(AR)로부터 받는 영향은 유의미한 반면 과거 2개 시점 이전부터 오차 혹은 변동값을 반영하여(MA) 현재 수요를 추정하는 것이 유의미한 것을 알 수 있다. 외생변수인 COVID-19와 공급량은 각각 신뢰수준 80%, 99%에서 통계적으로 유의하며, 공공자전거 수요와 양(+)의 관계를 갖는 것으로 나타났다. 즉, COVID-19 바이러스로 확진자 수가 증가할수록, 그리고 공공자전거의 공급량이 증가할수록 공공자전거 이용수요는 증가한다는 유의미한 결과가 도출되었다.

<Table 4> SARIMAX model results with exogenous variables

Model	MA1	MA2	SMA1	COVID-19	SUPPLY
Coefficients	-0.4426	-0.1197	-0.3349	312.6311	16.0822
Standard error	0.0849	0.0792	0.0937	184.2386	5.8292
AIC = 4,815.71, AICc = 4,816.16, BIC = 4,835.38, log likelihood = =2,401.86					

3. COVID-19 시나리오별 공공자전거 장래수요예측

최적 모형을 활용하여 장래 공공자전거 수요를 추정하기 위해서는 COVID-19 발병률의 변화에 대한 정보가 필요하다. 이를 위해 예측 불가능한 COVID-19 발병률에 대하여 3가지 추가 시나리오(II)를 설정하였다. 장래 COVID-19 확진자 수를 현재까지 발생한 확진자의 평균 수준(100명 내외)으로 향후 1년간 확진자 수 유지 후 종결(시나리오 II-1), 2021년부터 백신개발로 확진자 수 0명, 즉 COVID-19 바이러스의 영향력이 없는 경우(시나리오 II-2), 2021년부터 확진자수의 급격한 증가로 500명대를 유지, 2022년 이후 확진자 수 0(시나리오 II-3) 등으로 설정하였다. 그리고 3가지 추가 시나리오에 의해 도출된 서울시 공공자전거 이용수요 예측 결과는 <Table 5>와 같다.

<Table 5> Demand forecastign by COVID-19 cases scenarios

(unit: usage/year)

Year	COVID-19 Scenario1	COVID-19 Scenario2	COVID-19 Scenario3
2021년	32,221,236	31,779,146	39,004,893
2022년	38,484,054	38,484,054	38,484,054
2023년	43,663,408	43,663,408	43,663,408
2024년	45,959,445	45,959,445	45,959,445

SARIMAX 모형에 의한 향후 4년간 공공자전거 이용수요 예측 결과는 COVID-19 확진자 수가 많이 발생할수록 공공자전거 수요가 다소 증가하는 것으로 나타났다. 2022년 이후로는 COVID-19 확진자 수가 없다고 가정하였으므로 시나리오별 예측 결과는 같게 도출되었으나 2021년에는 COVID-19 발병률의 시나리오에 따라 예측수요는 상이하게 나타났다. 100명 내외의 확진자가 있을 경우 2021년 공공자전거 수요는 연간 3천 2백만 건이며, 확진자가 발생하지 않을 때(시나리오 II-2)와 비교하여 1.4% 증가 수준을 보였다. 확진자 수가 500명대를 유지하는 경우 약 3천 9백만 건의 연간 수요가 예측되었으며, 이는 확진자가 발생하지 않는 경우와 비교하여 약 25.8%로 급격히 이용률이 증가하는 것을 알 수 있다. 이러한 결과는 COVID-19의 확진자 수가 증가하고 발병률이 높아질수록 통행자들이 사회적 거리두기나 언택트의 교통수단 이용을 선호하는 경향을 반영하는 것으로 추정할 수 있으며, 이러한 패턴은 최근 COVID-19와 교통패턴 간의 관계를 도출한 기존 연구 결과와 일관성을 보인다.

Musselwhite et al.(2020)의 연구에서는 COVID-19로 인하여 시민들의 근무, 모임을 지양함으로써 통행량 축소를 일으켰다고 주장한 반면, 김채만 외(2020)의 연구에서는 COVID-19로 시민들의 대중교통 선호도가 감소하였으나 단독으로 이용되는 교통수단 선호도는 증가하여 대중교통의 대체재로 공공자전거를 활성화시키는 정책이 시행되는 사례가 증가함을 제시하였다. Cho et al.(2020)의 연구에서도 COVID-19 발생 이후 서울시 대중교통 이용량이 대폭 감소하고 승용차 이용량은 소폭 감소한 반면에 공공자전거, 나눔카 등 공유교통수단 이용량은 증가하였다. 이는 전염병 전파로 인하여 시민들의 전체적인 통행량이 감소하였음에도 불구하고 공공자전거는 대중교통수단에 비해 사회적 거리두기가 가능하여 이용률 감소의 영향을 상대적으로 적게 받고, 일부는 대중교통 수요가 공공자전거 이용수요로 대체되어 COVID-19 발생 이후 공공자전거의 수담 분담률이 다소 증가했다는 것을 보여준다. 또한 장래 COVID-19가 더욱 확산될 경우, 공공자전거는 본 연구에서 예측한 결과와 같이 수요가 증가되어 현재 접근교통 수단으로서의 제한적 용도에서 대중교통수단의 대체 역할을 하는 개인교통수단으로까지 확대될 가능성도 기대된다.

V. 결 론

본 연구는 공공자전거의 장래 이용수요를 예측하기 위한 방법론을 제시하고, 공공자전거의 공급정책에서 전염병의 확산과 같은 외생변수들이 이용수요에 미치는 영향을 파악하였다. 현재 공공자전거 수요예측은 빅데이터와 머신러닝 기법들을 중심으로 활발히 진행되고 있으나 대역소별, 시간대별 등 미시적인 시공간적 범위 위주로 분석이 수행되고 있으며, 특히 상위계획 수립을 위한 도시 수준의 거시적 차원에서는 연구가 미흡하여 공공자전거 효율화를 위한 정책 결정에 어려움이 있다.

이에 따라 본 연구는 SARIMAX 모형을 구축하여 주요한 외생변수들을 반영하여 서울시의 장래 공공자전

거 수요를 예측하는 접근법을 사용하였다. SARIMAX 모형은 ARIMA 등 일반적인 시계열모형이 비관측된 외생변수들의 영향을 반영하지 못해 결과를 왜곡하는 문제를 보완할 수 있다. 본 연구 결과를 통하여 COVID-19의 영향과 공급량 수준을 고려한 모형을 구축·분석함으로써 장래 이용수요에 영향을 주는 외생변수들이 통계적으로 유의하다는 것을 입증하였으며, 최종 모형을 활용하여 2021년부터 2024년까지의 서울시 공공자전거 수요를 예측하였다.

연구 결과, 공공자전거 이용수요는 공공자전거 공급에 직접적인 영향을 받는 것으로 나타났다. 예컨대, 국내 공공자전거 시장이 정착기에 도달하지 않아 공급량이 증가할수록 이용수요 역시 증가하는 패턴이다. 또한 COVID-19의 확산 정도는 공공자전거 이용수요와 양(+)의 상관관계를 갖는 것으로 나타나 전염병의 확산세가 심각해질수록 통행자들의 공공자전거 이용이 증가하는 것을 알 수 있었다. 서울시의 경우 2022년까지 4만5천 대의 공공자전거가 공급되고 2022년 COVID-19가 종식된다는 가정 하에 추정된 이용수요는 약 3천8백만 대이며, 2024년에는 약 4천 6백만 대까지 증가하는 것으로 추정되었다. 이는 다소 과다추정으로 생각될 수 있으나 우리나라 공공자전거 공급지수²⁾는 1.23수준으로 선진외국과 비교했을 때 낮으며(미국 16.33, 독일 24.43, 덴마크 17.33), 아시아의 중국 8.1, 일본 2.85 보다도 낮은 수준이다.

본 연구에서는 2024년까지 단기 공공자전거의 수요를 예측하였기 때문에 공공자전거 이용수요에 영향을 줄 수 있는 인구수, 자동차등록대수 등 중장기 외생변수들을 고려하지 않았다. 따라서 향후 중장기적인 공공자전거 이용수요를 예측할 때는 현재 반영한 공급량과 COVID-19 발병자수 외에 영향력 있는 외생변수들을 추가하여 본 연구 결과와 비교하는 후속 연구를 기대한다.

ACKNOWLEDGEMENTS

이 논문은 2020년도 서울시립대학교 교내학술연구비에 의하여 지원되었습니다.

REFERENCES

- Chai X., Guo X., Xiao J. and Jiang J.(2020), “Spatiotemporal Analysis of Share Bike Usage during the COVID-19 Pandemic: A Case Study of Beijing,” *Physics and Society*, arXiv preprint arXiv:2004.12340.
- Cho H., Yun S. and Jeong Y.(2020), “Seoul Transportation Changes and Strategies After COVID-19,” *Transportation Technology and Policy*, vol. 17, no. 3, pp.46-51.
- Cho K. M., Lee S. S. and Nam D.(2020), “Forecasting of Rental Demand for Public Bicycles Using a Deep Learning Model,” *The Journal of the Korea Institute of Intelligent Transport Systems*, vol. 19, no. 3, pp.28-37.
- Cho M. and Baik D.(2013), “A Study on Demand Forecasting Model for the Passengers of Marine Traffic using Time-Series Analysis Techniques,” *Journal of the Korea Society for Simulation Conference 2013*, pp.56-93.
- Choi M. and Jung H.(2020), “A Study on the Influencing Factor of Intention to Use Personal Mobility

2) 공급 지수는 인구 10만명 당 공공자전거 대수의 최소 및 최대값을 기준으로 정규화한 값임. Coya, The 2019 Bicycle Cities Index.

- Sharing Services,” *The Journal of Korean Society of Transportation*, vol. 38, no. 1, pp.1-13. Coya, <https://www.coya.com/bike/index-2019>, 2020.11.05.
- Fournier N., Christofa E. and Knodler M. A.(2017), “A sinusoidal model for seasonal bicycle demand estimation,” *Transportation Research Part D: Transport and Environment*, vol. 50, pp.154-169.
- Hamilton J.(1994), *Time series econometric*, Princeton U. Press, p.816.
- Hong J.(2019), “Air Demand Forecasting using Time Series Data: Focusing on Daegu International Airport,” *International Journal of Tourism and Hospitality Research*, vol. 34, no. 3, pp.61-77.
- Kaltenbrunner A., Meza R., Grivolla J., Codina J. and Banchs R.(2010), “Urban cycles and mobility patterns: Exploring and predicting trends in a bicycle-based public transport system,” *Pervasive and Mobile Computing*, vol. 6, no. 4, pp.455-466.
- Kim C. and Han A.(2020), “Since Corona 19, transportation policy has changed from efficiency to safety,” *Issue & Analysis*, Vol. 2020, no. 417, pp.1-26.
- Kim J. and Jang K.(2017), “Subway Demand Forecast using Seasonal Autoregressive Integrated Moving Average,” *Korean Society of Transportation, the 76th Academic Conference*, pp.450-453.
- Kim S. and Park M.(2016), “A Comparison of Accuracy among Tourism Demand Forecasting Models: Suwon-City, Gyeonggi-Do,” *Northeast Asia Tourism Research*, vol. 12, no. 4, pp.121-142.
- Lim H. and Chung K.(2019), “Development of Demand Forecasting Model for Seoul Shared Bicycle,” *Journal of Korea Contents Association*, vol. 19, no. 1, pp.132-140.
- Musselwhite C., Avineri E. and Susilo Y.(2020), “Editorial JTH 16-The Coronavirus Disease COVID-19 and implications for transport and health,” *Journal of Transport & Health*, vol. 16, 100853.
- O’Brien O., Cheshire J. and Batty M.(2014), “Mining bicycle sharing data for generating insights into sustainable transport systems,” *Journal of Transport Geography*, vol. 34, pp.262-273.
- Rudloff C. and Lackner B.(2014), “Modeling demand for bikesharing systems: Neighboring stations as source for demand and reason for structural breaks,” *Transportation Research Record*, vol. 2430, no. 1, pp.1-11.
- Seoul Open Data Portal, <http://data.seoul.go.kr>, 2020.11.01.
- Teixeira J. F. and Lopes M.(2020), “The link between bike sharing and subway use during the COVID-19 pandemic: The case-study of New York’s Citi Bike,” *Transportation Research Interdisciplinary Perspectives*, vol. 6, 100166.
- Tirachini A. and Cats O.(2020), “COVID-19 and public transportation: Current assessment, prospects, and research needs,” *Journal of Public Transportation*, vol. 22, no. 1, pp.1-21.
- Tsay R. S.(2005), *Analysis of financial time series*, John Wiley & Sons, p.720.
- Vogel P., Greiser T. and Mattfeld D. C.(2011), “Understanding bike-sharing systems using data mining: Exploring activity patterns,” *Procedia-Social and Behavioral Sciences*, vol. 20, pp.514-523.
- Wilbur M., Ayman A., Ouyang A., Poon V., Kabir R., Vadali A. and Dubey A.(2020), “Impact of COVID-19 on Public Transit Accessibility and Ridership,” *Physics and Society*, arXiv preprint arXiv:2008.02413.
- Zheng R., Xu Y., Wang W., Ning G. and Bi Y.(2020), “Spatial transmission of COVID-19 via public and private transportation in China,” *Travel Medicine and Infectious Disease*, vol. 34, 101626.