

서울시 공유자전거의 수요 예측 모델 개발

Development of Demand Forecasting Model for Seoul Shared Bicycle

임희종*, 정광현**

서울시립대학교 경영대학*, 홍익대학교 경영대학**

Heejong Lim(limh@uos.ac.kr)*, Kwanghun Chung(khchung@hongik.ac.kr)**

요약

최근 전 세계 많은 도시는 교통량 및 대기오염을 감축하기 위해 공유자전거 시스템을 도입하여 운영하고 있고, 서울시에서도 2015년부터 따릉이 공유자전거 서비스를 제공하고 있다. 공유자전거의 사용이 확산됨에 따라 대여소별로 자전거의 수요는 증가하고 있으나, 제한된 예산 하에서 대여소별로 수요를 관리하기 때문에 운영 및 관리상의 어려움이 존재한다. 현재 자전거 재배치를 통해 대여소별로 수요의 변동을 해결하려고 노력하고 있으나, 불확실한 미래의 사용자 수요를 정확히 예측하는 것이 보다 근본적인 방안이다. 본 연구에서는 통계적 시계열 분석을 통해 서울시 따릉이의 수요를 예측하는 모델을 개발하고, 이를 실제 데이터를 통해 분석하고자 한다. 특히, 전기 사용량의 수요에 사용했던 Holt-Winters 방법을 따릉이 수요 예측을 위해 변형하여 적용하였고, 또한 파라미터들의 변동이 실제 수요예측에 어떠한 영향을 미치는지 민감도 분석도 수행하였다.

■ 중심어 : | 공유자전거 | 수요예측 | 지수평활법 | Holt-Winters 기법 | 시계열분석 |

Abstract

Recently, many cities around the world introduced and operated shared bicycle system to reduce the traffic and air pollution. Seoul also provides shared bicycle service called as "Ddareungi" since 2015. As the use of shared bicycle increases, the demand for bicycle in each station is also increasing. In addition to the restriction on budget, however, there are managerial issues due to the different demands of each station. Currently, while bicycle rebalancing is used to resolve the huge imbalance of demands among many stations, forecasting uncertain demand at the future is more important problem in practice. In this paper, we develop forecasting model for demand for Seoul shared bicycle using statistical time series analysis and apply our model to the real data. In particular, we apply Holt-Winters method which was used to forecast electricity demand, and perform sensitivity analysis on the parameters that affect on real demand forecasting.

■ keyword : | Shared Bicycle | Demand Forecasting | Exponential Smoothing | Holt-Winters Method | Time-Series Analysis |

1. 서론

공유 자전거(Shared Bicycle)는 전 세계 많은 도시들이 도심 내 교통량 감축 및 대기 환경 오염을 줄이기 위해 도입하여 현재 널리 활용되고 있다. 서울시에서도 2015년에 “따릉이”라는 이름의 공공 자전거 시스템을 도입하였고, 현재는 서울시 모든 자치구에서 쉽게 사용할 수 있도록 많이 보급되었다[1]. 하지만 따릉이 도입 초기에는 공유 자전거 시스템의 초기 구축 비용이 상대적으로 적으나, 사용자가 증가함에 따라 수요/공급의 불일치로 인한 시스템의 효율적 관리가 필요하게 되었다. 특히, 도입 초기에는 시스템 확산을 위해 자전거 확대 보급 및 대여소 설치에 관심을 기울였다면, 이제는 주어진 예산의 범위 내에서 대여소별로 사용자 수요를 만족시키기 위해 충분히 자전거를 배치해야 하는 문제가 발생하고 있다. 이러한 자전거 배치 문제의 전제는 신뢰할 수 있는 데이터 분석을 통한 사용자의 수요예측이 있어야 한다.

공유 자전거 시스템에서는 이러한 수요/공급의 불일치를 해결하기 위해 자전거 재고의 재배치(Inventory Rebalancing)를 통해 해결하고 있다. 예를 들면, 홍대입구역의 대여소는 아침 출근시간대에는 직장인들이 지하철역에서 사무실로 이동을 위해 따릉이를 많이 사용하기 때문에 자전거가 부족한데, 난지한강공원의 대여소는 이용자가 적어서 자전거가 남아 있는 경우가 많다. 반대로 주말 낮에는 반대의 상황이 발생하는데, 이러한 문제를 해결하기 위해 마포구에서는 대여소 간의 남은 자전거를 주기적으로 트럭을 사용하여 재배치하여 수요를 원활히 충족하도록 운영하고 있다.

자전거 재배치와 관련된 기존의 연구는 여러 가지 방향으로 진행되어 왔다. Raviv와 Forma는 전통적인 재고관리 기법에서 사용되는 서비스 레벨을 자전거 대여 요구와 반납 요구의 만족도를 기준으로 정의하고, 사용자 요구가 만족되지 않는 경우에 대해 비용으로 계산하여 운송 차량의 운영비용의 총합을 최소화하는 자전거 재배치 관리 구역을 설정하였다[2]. Schuijbroek과 van Hove는 사용자 서비스 레벨을 정의하고, 자전거 재배치 관리구역의 설정과 재배치를 위한 트럭 운송경로를

동시에 결정하는 혼합정수계획(Mixed-Integer Programming) 모형을 개발하였다[3]. 이들은 서비스 레벨을 정의에서 반납 가능한 보관대가 없으면 반납이 불가능하다고 가정하였으나, 이상복, 임희중, 정광현의 연구에서는 서울시 따릉이 시스템의 특성을 반영하여 대여소 보관대의 수와 관계없이 반납이 가능한 상황을 고려하여 관리구역을 설정하는 최적화 모형을 개발하였다[4]. Dell’Amico와 Novellani도 공유자전거 재배치 문제에 대한 최적화 모형을 전통적인 최적화 문제인 One-commodity pickup-and-delivery vehicle routing problem으로 모형화하였고, 이탈리아의 한 도시 데이터를 사용하여 분석하였다[5].

이러한 자전거 재배치 연구들에서 공통으로 문제가 되는 것은 대여소별로 사용자의 요구, 즉 수요 데이터가 필요한데, 이는 과거 데이터를 기반으로 통계적으로 예측하는 방법이 가장 널리 사용되고 있다. 본 연구는 영등포구와 마포구의 총 수요를 과거 실적을 이용하여 예측하였다. 향후 연구 결과를 개별 대여소의 수요 예측 문제에 적용하여, 대여소별 재배치 문제에 활용하고자 한다.

본 논문에서는 서울시 공유자전거 시스템인 따릉이의 수요데이터를 기반으로 하여 사용자 수요를 예측하는 모델을 개발하고자 한다. 공유자전거의 수요예측에 대한 기존 연구들도 여러 가지 방향으로 진행되어 왔다. O’Brien과 Batty은 여러 도시의 공공자전거 시스템의 사용자 데이터를 분석하여 지역별 특성 및 공통점들을 파악하고, 이를 기반으로 향후 확산 및 보급을 위한 방향을 제안하였다[6]. Vogel과 Mattfeld은 오스트리아 빈의 공공자전거 사용자 데이터를 기반으로 시각화(Visualization) 기법을 사용하여 사용자의 이용 패턴을 분석하였다[7]. Rudloff와 Lackner도 오스트리아 빈의 사용자 데이터를 사용하여 수요를 예측하였고, 특히 날씨와 근접한 대여소의 재고 수준의 영향을 고려하였다[8]. Kaltenbrunner와 Banchs도 스페인 바르셀로나의 데이터를 분석하여 그 결과를 대여소별 사용자 수요예측에 적용하였다[9]. 하지만, 이러한 연구들은 전통적으로 생산/운영관리에서 사용하는 시계열 분석과 같은 통계적 분석방법을 사용하지 않아 향후 수요를 정확히 예

측하는 데 사용하기에는 여러 가지 한계가 존재한다. 따라서 본 연구에서는 따릉이 사용자 데이터의 시계열 분석에 기반을 둔 통계적 수요예측을 수행하고자 한다.

전통적인 생산/운영관리에서 적절한 수요예측 방법의 선택은 수요예측의 목적, 가능한 관측의 수, 그리고 예측하고자 하는 변수의 수에 따라 달라진다[10]. 이제까지 제안된 다양한 예측방법 중 예측하고자 하는 것에 대해 가장 적절한 방법을 선택하는 것은 어려운 방법이다. 이에 대한 길잡이를 위해 Newbold와 Granger은 Box-Jenkins, Holt-Winters, 그리고 Stepwise Regression 등의 기법들을 다양한 데이터에 적용하여 성과지표로 평균제곱오차(MSE, Mean Squared Error)를 비교하였다[11]. 그 결과 Holt-Winters 방법이 많은 경우 가장 우수하다고 결론지었다. 게다가 실용적인 측면에서도 Holt-Winters가 가장 우수하다고 알려져 있다[10]. 가법적인(additive) Holt-Winters 방법은 ARIMA(Autoregressive Integrated Moving Average) 방법의 한 형태로 볼 수 있는 반면, 승법적인(multiplicative) Holt-Winters 방법은 ARIMA의 방법으로 볼 수 없다. 게다가 가법적인 Holt-Winters 방법은 실제 데이터에 적용하기에는 너무나 복잡해지기 때문에 실제 데이터를 규명하는 데 사용하지 않는다[10].

단기간의 전력사용량을 예측하는데, 이중 계절성을 가진(double seasonal) Holt-Winters 방법을 사용하여 기존의 단일 계절성을 가진 Holt-Winters 방법과 승법적 이중계절성을 가진 ARIMA 방법보다 더 나은 결과를 가져왔다[12][13]. 본 연구는 공유자전거의 과거 사용패턴 분석 결과로 계절과 요일에 의한 이중 계절성(double seasonality)을 가지고 있다고 판단하고, double seasonal Holt-Winters 방법을 사용하여 수요예측 모델을 개발하였다.

본 논문의 구성은 다음과 같다. II.장에서는 수요예측 기법 중 시계열분석을 소개하고, 본 논문에서 사용한 Holts-Winters Exponential Smoothing에 대해 살펴본다. III.장에서는 본 논문의 수요예측을 위해 사용한 서울시 공유 자전거 시스템 따릉이의 사용자 수요 데이터를 소개하고, II.장에서 설명한 수요예측기법을 적용한 결과를 제시한다. 특히, 지수평활상수에 대한 결정을 위

해 사용한 최적화 기법과 민감도 분석의 결과도 함께 제시한다. 마지막으로 IV.장에서는 본 논문의 연구결과를 정리하고, 향후 연구 방향에 대해 제시한다.

II. 수요예측 기법

1. 시계열 분석

시계열 분석(Time Series Analysis)은 과거의 수요 데이터를 분석하여 수요에 영향을 주는 요소들을 파악하고, 이를 기반으로 하여 미래의 수요를 예측하는 방법이다. 대표적인 시계열 수요예측 기법으로는 이동 평균법(Moving Average), 지수 평활법(Exponential Smoothing), 추세 분석법(Trend Analysis), 시계열 분해법(Time Series Decomposition) 등이 있다[14]. 시계열 분석을 위해 과거 수요를 수준(Level), 추세(Trend), 계절성(Seasonality), 우연 변동(Random Variation)의 요소로 구분할 수 있다. 수준은 수요의 평균을 뜻하며, 수요의 증가 또는 감소는 반영되지 않는다. 추세는 수요가 시간이 지남에 따라 증가하거나 감소하는 경향을 나타낸다. 계절성은 일정한 시기에 따라 수요가 증가 또는 감소하는 규칙적인 변동을 뜻한다. 계절성은 일반적으로 가법적(Additive) 반영과 승법적(Multiplicative) 반영으로 나뉜다. 우연 변동은 잡음(Noise)나 예측에 반영하지 못한 요소에 의한 영향을 뜻하며, 이것은 일반적으로 예측이 불가능한 것으로 가정한다.

본 연구에서는 시계열 수요예측 방법 중 지수평활법에 기반한 모델을 사용하고자 한다. 지수평활법은 과거 데이터가 추세나 계절성이 없는 경우에 적합한 수요예측 기법으로, 수준에 의해 수요를 예측한다. 특히, t 시점의 수준(L_t)을 t 시점에서 관측된 수요(D_t)와 $t-1$ 시점의 수준(L_{t-1})의 가중평균으로 아래의 식(1)과 같이 추정한다.

$$L_t = \alpha D_t + (1 - \alpha)L_{t-1} \quad (1)$$

2. Holt-Winters 기법

과거 데이터에 추세나 계절성이 있는 경우에는

Holt-Winters Exponential Smoothing을 사용한다. Holt-Winters 기법은 지수평활법을 수준(level), 추세(trend), 계절성(seasonality)의 세 가지 요소로 구분하여 예측하는 방법으로 Winters에 의해서 제안된 방법이다[15]. 이 방법은 아래의 수식 (3) 과 같이 가법적으로 추세를 반영한다. 계절성은 아래의 수식 (4) 와 같이 승법적으로 반영하며, 새로운 수요 데이터(D_t)를 반영할 때마다 smoothing constant α, β, γ 를 갖는 지수평활법을 사용하여 수준, 추세, 계절성을 아래의 수식 (2)-(4) 와 같이 다시 계산하고, 이를 통해 수요예측값($\hat{D}_t(k)$)을 계산한다.

수준

$$L_t = \alpha(D_t / (S_{t-s_1})) + (1 - \alpha)(L_{t-1} + T_{t-1}) \quad (2)$$

추세

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta) T_{t-1} \quad (3)$$

계절성

$$S_t = \gamma(D_t / (L_t)) + (1 - \gamma) S_{t-s_1} \quad (4)$$

수요예측

$$\hat{D}_t(k) = (L_t + kT_t) S_{t-s_1+k} \quad (5)$$

특히, 수식 (2)-(5)는 t 시점까지 예측된 수준(L_t), 추세(T_t), 계절성(S_{t-s_1+k})에 기반하여 k 시점 이후의 수요예측값($\hat{D}_t(k)$)을 계산한 결과이다.

Holt-Winters exponential smoothing은 수식 (4) 와 달리 계절성을 가법적으로 반영할 수 있다. 하지만, 수식 (4) 와 같은 승법적인 방법이 보다 널리 사용되며, 본 연구에서는 승법적인 계절성을 이용한다.

3. Double Seasonal Holt-Winters 기법

II.2.에서 설명한 기본적인 Holt-Winters 모델이 시계열 수요예측 기법에서 많이 사용되지만, 계절성의 요인이 하나인 경우에만 적용할 수 있는 한계가 있다. 이러한 한계를 극복하기 위해 Taylor는 이중 계절성(Double Seasonality)을 반영한 지수평활법을 제안했다[12]. 특히, 이 기법은 두 가지의 계절적 패턴이 있는 경우 계절지수를 두 개의 S_t, W_t 값으로 나타낸다. 또한,

기본적인 Holt-Winters 모델에 두 번째 계절성을 반영하기 위한 smoothing constant ω 를 추가하여 아래의 수식 (6)-(10) 과 같이 예측한다.

수준

$$L_t = \alpha(D_t / (S_{t-s_1} W_{t-s_2})) + (1 - \alpha)(L_{t-1} + T_{t-1}) \quad (6)$$

추세

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta) T_{t-1} \quad (7)$$

계절성 1

$$S_t = \gamma(D_t / (L_t W_{t-s_2})) + (1 - \gamma) S_{t-s_1} \quad (8)$$

계절성 2

$$W_t = \omega(D_t / (L_t S_{t-s_1})) + (1 - \omega) W_{t-s_2} \quad (9)$$

수요예측

$$\hat{D}_t(k) = (L_t + kT_t) S_{t-s_1+k} W_{t-s_2+k} \quad (10)$$

수식 (6)-(10) 에서 제안한 이중 계절성은 3개 이상의 더 많은 계절성을 가진 경우에도 추가적인 Smoothing Constant와 수식을 추가하여 수요를 예측하는데 사용할 수 있다[12].

III. 수요예측 결과

1. 따릉이 사용자 데이터

III.장에서는 II.장에서 설명한 Double Seasonal Holt-Winters 기법을 사용하여 따릉이 수요를 예측하고자 한다. 본 연구에 사용된 서울시 공공자전거 따릉이의 사용자 데이터는 2015년 10월 15일부터 2017년 4월 30일까지의 사용자별 이용 실적을 사용하였다. 하나의 데이터는 사용자의 대여시각, 대여장소, 반납시각, 반납장소에 관한 정보로 구성되어 있다. 본 연구에서는 자전거의 일간 수요를 해당 지역구에서 하루 동안 대여된 자전거의 대수를 일별로 합한 값으로 정의한다.

본 연구에서는 서울시 마포구와 영등포구에 초점을 맞춰 데이터 분석 및 수요예측을 진행하였다. 먼저 마포구와 영등포구의 과거 일간 수요 데이터를 그래프로 그려 보면 아래의 [그림 1]과 같다.

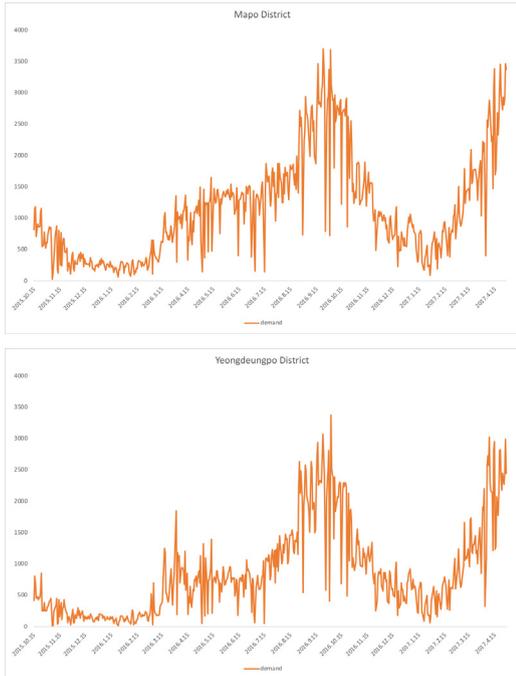


그림 1. 자전거 대여 일간 수요 추세

[그림 1]을 통해 관찰할 수 있는 점은 자전거 일간 수요에 계절성(Seasonality)이 강함을 알 수 있다. 특히, 사계절에 의한 패턴뿐만 아니라 어린이날과 같은 특정일에 의해 일간 수요가 반복됨을 알 수 있다. 추가적으로 요일별 편차가 큰 것은 요일별 수요의 계절성이 동시에 있다고 추정하게 한다. 마지막으로 시간이 흐름에 따라 전반적으로 수요량이 증가하고 있음을 알 수 있다.

[그림 1]의 분석 결과에 기반을 두어 본 연구에서는 II.3.에서 설명한 Double Seasonal Holt-Winters 기법을 사용하는 것이 적절하고, 수요예측 시 계절별, 요일별 수요의 이중 계절성과 추세, 수요의 수준 등의 요인들을 추정하여 지수평활법을 적용하여 예측하고자 한다.

본 연구에서는 제한된 크기의 데이터(구별로 564개의 데이터 포인트)를 이용해야 하는 상황에서, 계절별 계절성을 예측해야 하므로, Training, Testing, Validating data set으로 나누지 못했으나, 향후 추가적인 데이터를 수집하여, Training, Testing, Validating data set으로 나누어 분석할 예정이다.

ARIMA를 사용하여 시계열 분석을 하는 경우 단위

근 검정(Unit root test)를 수행하는 것이 일반적이지만 [16], Holt-Winters 방법을 사용하는 경우에는 Unit root test를 기존의 많은 연구들[10][12-14]에서 생략하였다. 향후 본 연구결과를 확장하여 ARIMA와 Holt-Winters 방법의 성능을 비교 및 분석할 예정이며, ARIMA를 수행할 때, 단위근 검정을 실시할 계획이다.

2. 따름이 수요예측 결과

III.1.절에서 설명한 바와 같이 과거 따름이 사용자 수요 데이터의 분석 결과에 기반하여 본 연구에서는 수준(level), 추세(trend), 그리고 두 개의 계절성(seasonality)의 네 가지 요소를 갖는 Double Seasonal Holt-Winters 기법을 사용해 따름이 수요를 예측하였다. 특히 사계절의 계절별 환경에 따라 변화하는 연간 계절성과 요일별 환경에 따라 변화하는 일간 계절성 등의 두 가지 계절성을 고려하고자 한다. 본 연구에서는 전력수요 예측을 위해 이중 계절성을 고려한 Taylor에서 제안한 방법을 활용하여 서울시 공공자전거의 영등포구와 마포구의 수요를 예측하였다[12].

수식 (10) 을 이용하여 t 시점까지 예측된 수준(L_t), 추세(T_t), 사계절에 의한 계절성(S_{t-s_1+k}), 요일별 계절성(W_{t-s_2+k})에 의한 k 시점 이후의 수요($\hat{D}_t(k)$)를 예측하도록 하였다.

각 요소의 초기값(Initial Value)는 수요의 이동평균을 구해서 일간 수요의 비계절화(Deseasonalized)된 수요를 계산하고, 이를 단순선형회귀법(Simple Linear Regression)을 통해서 수준의 초기값을 정하였다. 이를 바탕으로 추세값과 계절성을 나타내는 두 개의 계절지수 S_{t-s_1+k} , W_{t-s_2+k} 를 구하였다. 비계절화의 방법은 Chopra에서 제시된 아래의 수식 (11) 을 사용하였다[14].

$$\bar{D}_t = \begin{cases} \frac{D_{t-p/2} + D_{t+p/2} + \sum_{i=t+1-p/2}^{t-1+p/2} 2D_i}{2p} & \text{for } p \text{ even} \\ \frac{\sum_{i=t-(p-1)/2}^{t+(p-1)/2} D_i}{p} & \text{for } p \text{ odd} \end{cases} \quad (11)$$

다음으로 비계절화된 수요(\overline{D}_t)를 아래의 선형회귀식 (12)를 활용하여 수준(L_0), 추세(T_0)를 추정한다. 이를 통해서 수준과 추세에 따라 추정된 수요 직선을 이용하여 두 개의 계절지수에 대한 초기값 S_k, W_k 를 구하였다.

$$D_t = L_0 + T_0t + \epsilon \tag{12}$$

본 연구의 수요예측 모델의 성과지표(Performance Measure)는 예측오차를 측정하는 방법의 하나인 평균 제곱오차(Mean Squared Error, MSE)로 다음과 같이 정의된다.

$$MSE_t = \frac{\sum_{k=1}^t (\widehat{D}_k - D_k)^2}{t} \tag{13}$$

평균제곱오차는 지수평활법에서 사용된 상수 $\alpha, \beta, \gamma, \omega$ 에 대해 비선형식(non-linear)의 형태를 가진다. 평균제곱오차를 최소화하는 지수평활 상수들을 구하기 위해서 비선형 최적화 문제를 풀어야 하는데, 본 논문에서는 마이크로소프트 엑셀의 GRG(Generalized Reduced Gradient method) 비선형 해찾기 기능을 사용해서 [표 1]과 같이 구했다.

표 1. 최적 지수평활상수

| 상수 | 요인 | Mapo | Yeongdeungpo |
|----------|-------|--------|--------------|
| α | Level | 0.7938 | 0.7828 |
| β | Trend | 0 | 0 |
| γ | Day | 0.1874 | 0.1299 |
| ω | Year | 0.5613 | 0.8421 |

[표 1]의 결과는 지수평활법을 통해서 추세를 추정하는 것이 수요예측에 큰 이점이 없음을 나타낸다. 수준(level)은 거의 비슷한 상수를 사용하면 된다. 그리고 마포구는 요일별 계절성이 영등포구보다 더 민감하게 반영되었고, 반면에 사계절에 의한 연간계절성은 영등포구가 더 민감하게 반영하는 것이 MSE를 최소화하는데 도움을 준다. [표 1]의 결과를 이용하여 수요예측을 할 경우, 평균제곱오차는 아래의 [표 2]와 같이 마포구

의 경우 183358.312, 영등포구의 경우 161750.415로 수요예측을 하였다.

수요예측 모델의 다른 성과지표로, 평균절대백분율 오차(Mean Absolute Percentage Error, MAPE)를 계산하였다.

$$MAPE_t = \frac{100}{t} \times \sum_{k=1}^t \frac{|\widehat{D}_k - D_k|}{D_k} \tag{14}$$

MAPE의 계산결과 [표 2]와 같이 마포구는 MAPE가 38.71%, 영등포구는 조금 더 오차가 큰 60.32%로 계산되었다.

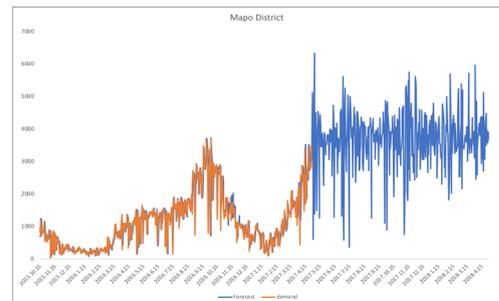
표 2. 최적 평활상수 사용시 평균제곱오차

| | Mapo | Yeongdeungpo |
|------|------------|--------------|
| MSE | 183358.312 | 161750.415 |
| MAPE | 38.71% | 60.32% |

MAPE를 최소화하는 지수평활상수를 구하여, 수요예측을 할 경우, 너무 크게 예측되는 경향이 있고, 영등포구와 마포구의 MAPE차이가 너무 커서, 본 연구에서는 MSE를 최소화하는 지수평활상수를 이용하여 수요를 예측하였다.

다음으로 최적화된 지수평활 상수를 사용하여 2017년 4월 30일 이후부터 1년간의 수요를 예측한 결과는 다음 [그림 2]와 같다.

과거 데이터와 비교하면, 본 연구에서 제시한 수요예측 모형은 신뢰할 만한 수준의 예측치를 보여준다. 즉, 수요예측 값과 실제 수요가 매우 비슷하게 계산되었다. 2017년 5월 1일부터의 수요예측 값은 [그림 1]을 통해서 변동과 추세를 파악할 수 있다.



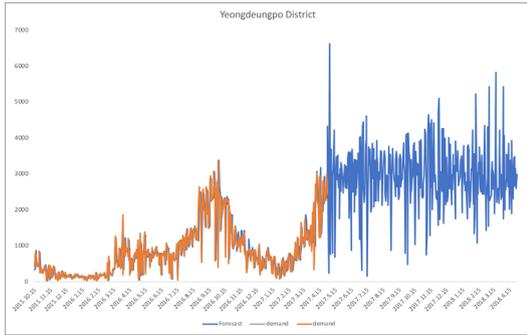


그림 2. 수요예측 결과

3. 민감도 분석

지수평활법에서는 지수평활상수의 결정이 모형의 성능에 큰 영향을 미친다. Holt-Winters 모형 또한 수요예측 모형의 요소들에 대해서 지수평활법을 사용하므로, 지수평활상수에 대한 평균제곱오차(MSE)의 민감도를 확인할 필요가 있다.

본 연구에서는 민감도 분석을 수행하기 위해서 다른 모든 조건을 고정하고, 한 개의 지수평활상수를 변화시켜가면서 평균제곱오차(MSE)에 미치는 영향을 관찰하였다. 즉 어떤 1개의 지수평활상수의 민감도를 분석하기 위해서 다른 3개의 지수평활상수를 고정하고, 해당 지수평활 상수를 변화시키면서 수요예측을 하고, 수요예측 결과와 실제 데이터와의 차이를 평균제곱오차로 측정하였다. 민감도 분석의 결과는 아래의 [그림 3]과 같다.

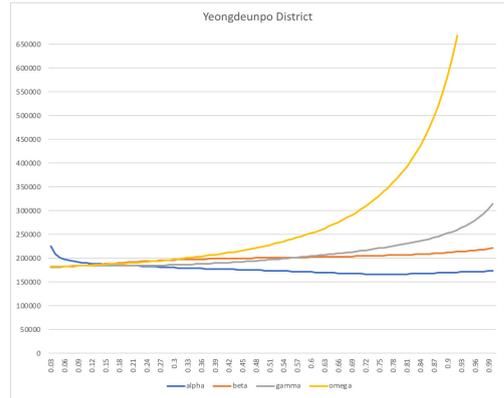
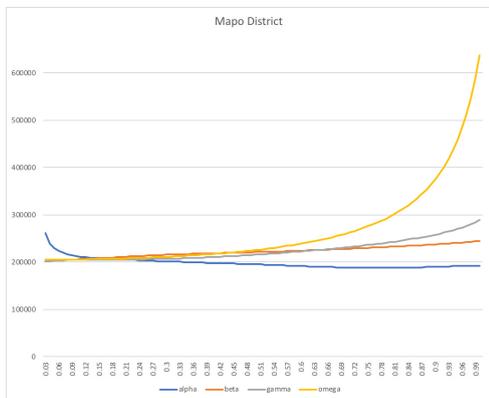


그림 3. 민감도 분석

수준에 대한 지수평활상수(α)의 변화에 따라 평균제곱오차(MSE)는 전반적으로 아래로 볼록한(Convex) 형태로 변화하였다. 특히, 0.7~0.8사이의 값에서 평균제곱오차를 최소화시키고, 0.12 이상의 값부터는 상대적으로 평균제곱오차에 적게 영향을 주는 것으로 마포구와 영등포구 데이터를 통해서 확인할 수 있다. 나머지 지수평활상수는 증가할수록 평균제곱오차는 커지는 양상을 보였다. 추세에 대한 지수평활상수(β)는 커지면 커질수록 평균제곱오차를 증가시키지만, 수준에 대한 지수평활상수(α)처럼 평균제곱오차에 적게 영향을 주었다. 반면, 계절성에 관한 두 지수평활상수(γ, ω)는 증가함에 따라 평균제곱오차를 증가시키고, 0.6이상의 값부터는 급격히 평균제곱오차가 증가하였다.

이러한 지수평활상수변화에 따른 평균제곱오차의 변화 양상을 통해서, 수준에 대해서는 실제 관측된 수요의 변화 수준을 의미 있게 반영할 필요가 있지만, 수요의 추세변화, 계절성 변화는 상대적으로 적게 반영하는 것이 좋을 것 같다.

IV. 결론

공유 자전거의 사용수요 예측은 적절하고, 강건성(robustness)를 갖춘 방법이 필요하다. 특히 시간의 변화에 따른 사용수요예측을 하는 경우, 시간을 단일 변수로 보는 단변량(univariate) 예측방법이 필요하다.

ARIMA와 같은 기준에 개발된 여러 가지 방안들이 존재하지만, 지수평활법의 강건성을 고려해 볼 때, Holt-Winters방법은 사용수요 예측에 적절한 방법이라고 판단된다. 하지만, 기존의 Holt-Winters 방법은 단일 계절성을 상정한 모델이므로, 본 연구에서는 Taylor가 제안한 이중 계절성을 반영한 Holt-Winter 방법의 변형을 공유자전거의 사용수요 예측에 적용하였다[12].

본 연구에서는 마포구와 영등포구의 서울시 공유자전거 시스템인 따릉이의 수요실적 데이터를 이용해서 미래의 수요 변화를 예측하였다. 본 연구에서는 2015년 10월 15일부터 2017년 4월 30일까지의 사용자별 이용실적 데이터에 Holt-Winter방법을 적용하였다. 또한, 비선형 최적화 기법을 통해서 예측모형의 요소인 수준, 추세, 그리고 두 가지의 계절성에 대해서 평균제곱오차(MSE)를 최소화하는 지수평활상수를 계산하였다. 예측 결과로 마포구의 경우 183358.312, 영등포구의 경우 161750.415의 평균제곱오차를 갖는 수요예측을 도출하였다.

본 연구에서 개발한 예측모델을 이용하여, 대역소별 수요예측에 적용하여, 공유자전거의 재배치에 활용할 수 있다. 또한, 추가적인 자전거 투자를 위한 의사결정 자료로 사용될 수 있을 것으로 기대된다.

본 연구는 서울시 공용자전거의 수요를 이중 계절성을 예측하는데 적절한 방법으로 인정받는 Holt-Winter 방법을 이용하여 예측하였고, 특히 이중 계절성을 반영한 것이 의미가 있다. 하지만 몇 가지 한계점도 존재한다. 첫째, 계절성을 보다 현실에 적합하게 반영하기 위해서는 더 많은 데이터를 이용하여 예측할 필요가 있다. 현재는 1년 6개월 동안의 공용자전거 이용실적을 이용한 한계가 있다. 둘째, ARIMA와 최근 주목받는 기계학습(Machine Learning)과의 예측 성능을 비교한다면 보다 풍부한 통찰과 예측 방법별 특성을 밝힐 수 있을 것이며 이는 추후 연구과제로 고려할 예정이다.

참 고 문 헌

- [1] 장재민, 김태형, 이무영, “서울시 공공자전거 이용 특성에 관한 연구 - 여의도 및 상암 지구를 사례

- 로,” 서울시연구, 제17권, 제4호, pp.77-91, 2016.
- [2] T. Raviv, M. Tzur, and I. A. Forma, “Static repositioning in a bike-sharing system: models and solution approaches,” *EURO Journal on Transportation and Logistics*, Vol.2, No.3, pp.187-229, 2013.
- [3] J. Schuijbroek, R. C. Hampshire, and W. J. van Hoes, “Inventory rebalancing and vehicle routing in bike sharing systems,” *European Journal of Operational Research*, Vol.257, No.3, pp.992-1004, 2017.
- [4] 이상복, 임희중, 정광현, “서울시 공공자전거 시스템 운영을 위한 효율적 관리 구역 설정,” *경영과학*, 제35권, 제1호, pp.55-67, 2018.
- [5] M. Dell’Amico, E. Hadjicostantinou, M. Iori, and S. Novellani, “The bike sharing rebalancing problem: Mathematical formulations and benchmark instances,” *Omega*, Vol.45, pp.7-19, 2014.
- [6] O. O’Brien, J. Cheshire, and M. Batty, “Mining bicycle sharing data for generating insights into sustainable transport systems,” *Journal of Transport Geography*, Vol.34, pp.262-273, 2014.
- [7] P. Vogel, T. Greiser, and D. Mattfeld, “Understanding bike-sharing systems using Data Mining: Exploring activity patterns,” *Procedia Social and Behavioral Sciences*, Vol.20, pp.514-523, 2011.
- [8] C. Rudloff and B. Lackner, “Modeling demand for bikesharing system,” *Transportation Research Record: Journal of the Transportation Research Board*, Vol.2430, pp.1-11, 2014.
- [9] A. Kaltenbrunner, R. Meza, J. Grivolla, J. Codina, and R. Banchs, “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, 2010.

- [10] C. Chatfield, "The Holt-Winters Forecasting Procedure," Journal of the Royal Statistical Society, Series C(Applied Statistics), Vol.27, No.3, pp.264-279, 1978.
- [11] P. Newbold and C. Granger, "Experience with Forecasting Univariate Time Series and the Combination of Forecasts," Journal of the Royal Statistical Society Series A(General), Vol.137, No.2, pp.131-165, 1974.
- [12] J. Taylor, "Short-term electricity demand forecasting using double seasonal exponential smoothing," Journal of the Operational Research Society, Vol.54, pp.799-805, 2003.
- [13] J. Taylor and P. McSharry, "Short-term load forecasting methods: An evaluation based on european data," IEEE Transactions on Power Systems, Vol.22, No.4, pp.2213-2219, 2007.
- [14] S. Chopra and P. Meindl, *Supply Chain Management Strategy, Planning & Operation*, Gabler, 2007.
- [15] P. Winters, "Forecasting sales by exponentially weighted moving averages," Management Science, Vol.6, pp.324-342, 1960.
- [16] 서정민, "시계열 모델을 활용한 사회서비스 수요 · 공급모형 구축," 한국콘텐츠학회논문지, 제15권, 제6호, pp.399-410, 2015.

저 자 소 개

임 희 중(Heejong Lim)

정회원



- 2006년 8월 : KAIST 산업공학과(공학사)
- 2008년 8월 : Purdue University, Industrial Engineering(공학석사)
- 2016년 8월 : Purdue University, Management(경영학 박사)

▪ 현재 : 서울시립대학교 경영대학 조교수

<관심분야> : 공급사슬관리, 운영관리

정 광 현(Kwanghun Chung)

정회원



- 1997년 2월 : 서울대학교 산업공학과(공학사)
- 1999년 2월 : 서울대학교 산업공학과(공학석사)
- 2010년 8월 : University of Florida(공학박사)

▪ 현재 : 홍익대학교 경영대학 조교수

<관심분야> : 경영과학, 최적화