



69

계절성 ARIMA를 이용한 지하철 수요예측

교통운영 김재환 (한국과학기술원 박사과정)
장기태 (한국과학기술원 조교수)

수도권 지하철은 자치단체 및 민간기업 소유의 다수 기관이 운영하고 있다. 따라서 승객들의 빈번한 환승 및 노선 확충으로 인한 정확한 수요예측이 필요하다. 본 연구에서는 시계열 분석 방법을 적용한 지하철 수요예측을 수행하였다. 이를 위하여 서울메트로와 서울도시철도공사의 월별 수송실적을 계절성 ARIMA(Seasonal Autoregressive Integrated Moving Average)를 이용하여 모형화하였다. 이렇게 추정된 모형으로 실제 수요를 예측했을 때 2년간 95 ~ 98 % 대의 정확도를 기록하여 모형의 정확성을 입증하였다. 이 예측모형은 장래 연락운임의 정산·배분, 안전관리 등을 아우르는 통합적 운영 계획 수립 전반에 활용 가능할 것으로 예상된다.

70

교통부문 미세먼지 배출 현황 및 주요 저감대책의 효과 평가

교통환경 박상준 (한국교통연구원 부연구위원)
천승훈 (한국교통연구원 부연구위원)
전민택 (한국교통연구원 ???)

우리나라의 주기적인 고농도 미세먼지가 사회적 문제로 등장한 이래 정부는 미세먼지 저감을 위해 다양한 미세먼지 저감 대책을 수립하고 많은 예산을 투입하였으나 아직까지 가시적인 성과는 나타나지 않고 있다. 미세먼지의 발생원인에 대해서는 명확한 정의가 내려져 있지 않지만 중국에서 유입되는 비율이 약 30~40% 정도이며 나머지는 국내에서 발생하고 있는 것으로 추산되며 국내 발생요인으로는 이견이 있으나 교통부문 중 도로에서 약 25%가 발생하고 있다. 특히 도로부문 미세먼지의 주 원인으로 지목된 경유차의 규제와 관련한 다양한 정책이 제시되었으나 경유차 규제에 의한 미세먼지 저감 효과를 객관적으로 평가한 자료는 구축되어 있지 않다. 이러한 측면에서 본 연구는 경유차의 미세먼지 저감효과를 객관적으로 평가하여 저감대책의 효과를 추정하였다. 저감대책 평가대상은 수도권에 등록된 10년 이상 노후경유차 중 저공해화 장치(DPF) 부착이 가능한 차량을 대상으로 미세먼지 저감효과를 산정하였고 DPF 장착으로 저감된 미세먼지 감축에 대한 추정량을 연간 미세먼지로 인한 피해금액으로 환산하여 저감대책의 효과평가가 계량적으로 이루어질 수 있도록 하였다.

71

온실가스 감축모형 LEAP을 활용한 교통부문 장기 온실가스 감축잠재량 분석 연구

교통환경 박상준 (한국교통연구원 부연구위원)
김병협 (한국교통연구원 연구원)

2015년 12월 파리에서 개최된 제21차 기후변화협약(COP21)에서 신기후체제(POST-2020)가 합의됨에 따라 그동안 선진국에 대한 부과된 온실가스 감축의무가 2020년부터는 개도국을 포함한 전세계 모든 국가에게 부과된다. 이에 대응하기 위해 우리 정부는 2015년 6월 UN에 자발적 감축목표인 INDC를 제출하였는데, 국가 온실가스 감축목표를 2030년 BAU 대비 37%로 결정하였다. 이후 정부는 공동작업반을 운영하여 2030년까지 부문별 온실가스 감축잠재량 및 감축목표를 설정하였는데, 교통부문은 24.6%로 결정되었다. 따라서 본 연구에서는 정부가 제시한 목표가 달성가능한지, 어떤 수단(정책)을 통해 달성가능한지에 대한 검토를 위해 온실가스 감축모형을 구축하여 평가를 시도하였다.

72

위험물질 운송관리 시스템 구축을 위한 실태분석 및 관리 체계 구축 연구

교통환경 계동민 (한국교통연구원 연구원)
노홍승 (한국교통연구원 연구위원)
신승진 (한국교통연구원 Post-Doc)

지속적인 국제 경기불황과 기업간 경쟁 심화에 의해 정부와 기업은 물류비 절감에 많은 노력을 기울이고 있다. 특히 물류비의 70%이상을 차지하는 운송비 절감에 대한 많은 관심과 연구가 지속되고 있는 실정이다. 하지만 운송대상이 폭발물 및 독극물, 유해화학물질 등 위험물질일 경우 사고시 발생할 수 있는 인적·물적·환경적 피해가 막대하고 추가적이고 지속적인 피해를 야기할 수 있기 때문에 무분별한 비용의 절감은 지양해야 한다. 이처럼 위험물 운송사고는 사고빈도가 낮다고 해서 정책 우선순위를 낮게 평가할 수는 없을 것이다. 그럼에도 불구하고 현재 국내 위험물 운송관리는 국토교통부, 환경부 등 다양한 정부부처들이 각 담당분야와 역할을 달리하여 11개(부령, 고시 포함 23개)에 달하는 복잡한 법령체계에 얽혀 체계적이고 신속한 대응이 어려운 현실이다. 본 연구는 국내 위험물질 운송·관리실태 및 법·제도 현황을 조사하고 국내 실정에 맞는 위험물질 운송관리 체계 및 관리방안을 제언하고자 한다. 이를 위해 각 위험물 운송 단계별 주요 추진 과제를 설정하고 시급성과 중요도가 높은 과제 순으로 우선순위를 평가하여 제시하였다.

계절성 ARIMA를 이용한 지하철 수요예측

김재환¹ · 장기태^{2*}

¹한국과학기술원 조천식녹색교통대학원, ²한국과학기술원 조천식녹색교통대학원

Subway Demand Forecast using Seasonal Autoregressive Integrated Moving Average

KIM, Jae Hwan · JANG, Ki Tae*

¹The CCS Graduate School for Green Transportation, KAIST, Daejeon 305-701, Korea

²The CCS Graduate School for Green Transportation, KAIST, Daejeon 305-701, Korea

Abstract

수도권 지하철은 자치단체 및 민간기업 소유의 다수 기관이 운영하고 있다. 따라서 승객들의 빈번한 환승 및 노선 확충으로 인한 정확한 수요예측이 필요하다. 본 연구에서는 시계열 분석 방법을 적용한 지하철 수요예측을 수행하였다. 이를 위하여 서울메트로와 서울도시철도공사의 월별 수송실적을 계절성 ARIMA(Seasonal Autoregressive Integrated Moving Average)를 이용하여 모형화하였다. 이렇게 추정된 모형으로 실제 수요를 예측했을 때 2년간 95 ~ 98 % 대의 정확도를 기록하여 모형의 정확성을 입증하였다. 이 예측모형은 장래 연락운임의 정산·배분, 안전관리 등을 아우르는 통합적 운영계획 수립 전반에 활용 가능할 것으로 예상된다.

Key Words

지하철, 수요예측, 운영계획, 운임정산, 안전관리

I. 서론

1. 연구의 배경 및 목적

수요예측은 시장에서 공급되는 서비스나 물품의 수요 통계를 분석하여, 장래 특정 기간 동안의 수요 총량, 구성, 특징 및 변화 양상을 추정하는 것을 의미한다. 여객철도에서는 이용객 수의 추정이 이에 해당한다. 철도 운영기관은 장래 예측수요를 충족하기 위한 인프라 확보, 서비스 향상 및 안전관리대책 등의 계획을 수립하며, 특히 다수의 운영기관으로 구성된 수도권 도시철도망에서 수요예측은 기관 간 연락운임 정산·배분, 직결운행 및 통합적 운영전략 입안 등에 활용할 참고자료로서의 가치가 매우 높다.

교통 수요예측에 관한 기존 연구에서는 특정 출발지(Origin, O)와 목적지(Destination, D) 간 개인통행을 배정하기 위해 통행발생, 통행배분, 수단선택 그리고 경로설정을 거치는 4단계 수요 추정법(Four - Step Model (FSM))을 널리 활용하고 있다. 그러나 FSM은 차량 통행을 기반으로 개발되어 도로 등급에 따른 교통망 분류, 가구별 소득수준, 차량 수 등 대중교통 수요 예측에는 적합하지 않은 가정들도 포함되어 있다. 또한 통행발생 단계에서 통행숫자가 이미 고정되어 인프라 등 거시적 교통여건 변화에 따른 통행패턴 변화를 반영하기 어렵다. 아울러 단계마다 개인의 통행양상을 논리적으로 추정하기 위한 알고리즘(Logit Model, Equilibrium Model 등)을 선택해야 하는데 이 때 수집하는 자료의 형태, 알고리즘의 전제조건 및 특징이 실제 여건과 부합하는 정도에 따라 결과가 달라질 가능성도 있다. 본 연구에서는 통행수단이 이미 도시철도로 고정되어 있고, 세부경로와 무관하게 네트워크 전체 수요를 예측하는 것을 목표로 하므로 FSM을 구성하는 네 단계를 굳이 거쳐야 할 필요성이 적다.

따라서 본 연구에서는 통행발생부터 차례대로 분석하는 FSM보다 최종적으로 산출된 합산 수요자료만으로 단기간의 패턴과 중·장기적인 추세를 분석, 이를 근거로 장래 수요를 예측하는 방식이 더 적합할 것으로 판단하였다.

한편 국내 도시철도 운영기관에서는 최근 이용자료를 바탕으로 간단한 회귀분석을 수행, 기존 노선의 수요를 예측하여 왔다.

그러나 일정 수준 이상의 신뢰도를 보장하기에는 표본의 숫자가 10 개 내외로 과도하게 적고, 기본적으로 회귀분석은 독립변수인 수집된 자료 내의 상관관계가 없음을 전제로 하기 때문에, 과거 수요에 영향을 받아 다중공선성문제를 야기할 수 있는 장래 수요를 예측하기에는 적합하지 않다.

이러한 점을 감안할 때 이용수단이 단일하며, 경로와 무관한 수도권 도시철도 네트워크 전체의 수요를 예측하기 위해서는 기간에 따라 축적된 수요 자료 내의 상관관계를 활용한 시계열 분석이 타당할 것으로 판단하였다.

따라서 본 연구에서는 수도권 도시철도 수요예측 방법론으로 시계열 분석을 채택, 이 중 계절을 고려한 ARIMA(SARIMA, Seasonal Autoregressive Integrated Moving Average Model) 기반 수요예측을 목표로 한다. 이를 위해서 SARIMA에 대해 설명하고, 수집한 자료(2006 ~ 2014 월별 수송실적)에 적합한 모형을 표준화된 과정(Box - Jenkins Method)을 거쳐 설정하여 2 년치(2013, 2014) 예측 결과를 제시한다. 이렇게 도출된 모형을 실제 운영에 적용 가능한 방안을 모색하고 이를 위한 요건과 향후 연구주제를 도출하고자 한다.

II. 본론

1. 수요예측

수요예측 대상은 수도권 양대 도시철도 운영기관인 서울메트로(Seoul Metro)와 서울도시철도공사(SMRT)로, 두 기관의 2006 ~ 2014 월별 수송실적 통계자료를 분석대상으로 설정하였다. 월별 수송실적은 해당 기관이 운영하는 각 역을 한 달 동안 이용한 승객들의 총합으로, 각 역에서 직접 승차한 승객의 총합인 승차인원과 다른 기관에서 승차 후 해당 기관으로 갈아타 각 역에서 하차한 승객의 총합인 유입인원의 합으로 구성된다. 이 중에서 2012년까지의 자료(표본 총 12 개월·7 년 = 84 개)로 모형을 수립하여 마지막 두 해 (2013, 2014년 12 개월·2 년 = 24 개) 실제 수요와 예측결과를 비교하였다.

그림 1은 서울메트로와 서울도시철도공사의 월별 수송실적을 선형그래프로 나타낸 것으로 시계열 자료(Time - Series Data)의 특성인 계절적인 패턴과 장기적인 추세를 동시에 포함하고 있다. 일정한 패턴만 주기적으로 반복되는 시계열 자료에서는 과거

* : 교신저자: kitae.jang@kaist.ac.kr, Phone: +82-42-350-1264, Fax: +82-42-350-1250

패턴을 모듈화하여 유사한 상황에 대입하는 예측기법(Pattern Matching)을 적용할 수 있지만, 이처럼 추세가 함께 나타나는 경우에는 두 요소를 모두 고려한 예측모형의 선택이 필요하다.

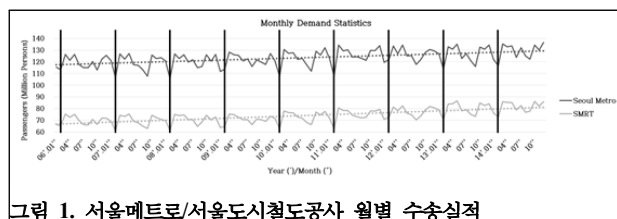


그림 1. 서울메트로/서울도시철도공사 월별 수송실적

시계열 분석의 기본적인 방법은 과거 자료만을 기반으로 미래 수치를 예측하는 것이다. 대표적인 모형으로 현재 오차와 과거 자료의 가중평균을 반영하는 자기상관(Autoregressive, AR) 모형과 현재 오차와 과거 오차의 가중합산을 반영하는 이동평균(Moving Average, MA)모형이 있으며 이 두 모형을 바탕으로 여러 모형이 파생 가능하며 대표적인 파생모형은 두 모형의 조합인 자기상관이동평균(Autoregressive Moving Average, ARMA)모형과 이것에 차분(Differentiation)을 더해 일반화한 자기상관통합이동평균(Autoregressive Integrated Moving Average, ARIMA)모형이다.

최근 대중교통 수요예측 관련 연구에서는 ARIMA를 주어진 조건에 맞게 수정하여 폭넓게 활용하고 있다. ARIMA로 철도 여객수요를 예측한 국내 문헌에서는 전체 이용객을 종속변수로 설정하고 이 추이를 개략적으로 예상하기 위한 모형을 수립하였다. 그러나 도시철도 수요는 단일 네트워크 내에서도 복수의 이용기관 간 승객의 이동이 빈번하여, 수요가 순수 승차와 타 기관에서의 유입으로 분류되는 점에서 기존 주요 ARIMA 적용 연구 대상인 간선철도(단일 운영기관 내 이동)나 항공물류 수요(네트워크로의 성격이 약함)와는 구별되는 특성을 지닌다.

본 연구에서는 계절에 따른 패턴을 보이는 자료의 특성을 감안하여, ARIMA 중에서도 계절 요인을 고려한 계절 ARIMA(SARIMA, Seasonal Autoregressive Integrated Moving Average)를 적합한 모형으로 선택하였다.

SARIMA 모형을 수요예측에 널리 활용한 타 교통수단과 달리 도시철도는 이러한 사례가 드물고, 기존 기관의 수요예측은 전술한 바와 같이 제한된 숫자의 표본을 바탕으로 한 단순회귀분석에 국한되어있다. 따라서 본 연구에서는 타 교통수단 수요예측에서 검증된 SARIMA 모형을 도시철도에도 적용, 승차와 유입 두 수요에 대한 예측을 고도화하고자 한다.

표 1·2의 ARIMA 전개식에서 예측 결과(X_t)는 과거 자료($X_{t-1}, X_{t-2}, \dots, X_1$), 특정 시점의 과거 자료가 영향을 미치는 시간차를 의미하는 변수(L, Lag Operator), 그리고 확률오차(ϵ_t)에 좌우된다. ARIMA를 구성하는 AR/MA의 전개식이 양변에 위치한다. AR/MA에서 과거 자료 또는 오차가 결과에 영향을 미치는 범위(p, q)가 AR(p)와 MA(q)의 모수(parameter)이며, 여기에 차분 횟수(d)를 더해 ARIMA 모수 조합(p, d, q)을 완성한다. 각 항을 단순화한 전개식은 표 2 맨 아래와 같다.

표 1. ARIMA 기본 형태 및 모수(p, d, q)

Form	$(1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p)(1 - L)^d X_t = \delta + (1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q) \epsilon_t$
Term	X_t : Raw Data, t : Time Operator, δ : Constant ϵ_t : Probabilistic Error with $N(0, \sigma^2)$ as a White Noise(No Autocorrelation) L : Lag Operator(Time Gap Component in the Influence of Previous Data) $1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p$: AR Component(Simplified as $\phi_p(L)$) $1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q$: MA Component(Simplified as $\theta_q(L)$) p : AR Degree(Range of Previous Terms with Output Contribution) q : MA Degree(Range of Previous Errors with Output Contribution)

표 2. ARIMA 모수 및 간략화된 전개식

Term (Cont'd.)	d : Differential Degree(Necessary Frequency for Stationary Achievement)
Simplification	$\phi_p(L)(1 - L)^d X_t = \delta + \theta_q(L) \epsilon_t$

SARIMA는 전술한 ARIMA(p, d, q)에 ARIMA 구성요소의 계절적 변형인 계절(Seasonal) AR(P) 및 MA(Q) (SAR(P)/SMA(Q)), 차분(D)과 시간차 변수(S)를 더하여 구성된다. 이를 종합하여 SARIMA의 모수 조합(p, d, q)(P, D, Q)s가 완성되며 전개식은 표 3과 같다.

표 3. SARIMA 기본 형태 및 모수(p, d, q)(P, D, Q)s

Form	$\phi_p(L)\Phi_P(L)(1 - L)^d(1 - L^S)^D Z_t = \delta + \theta_q(L)\Theta_Q(L^S)\epsilon_t$
Term	P : SAR($1 - \Phi_1 L^S - \Phi_2 L^{2S} - \dots - \Phi_P L^{PS}$) Degree Q : SMA($1 - \Theta_1 L^S - \Theta_2 L^{2S} - \dots - \Theta_Q L^{QS}$) Degree D : Seasonal Differential Degree S : Seasonal Lag Degree

이처럼 SARIMA는 모수를 통하여 정의되므로, 예측모형을 확정하기 위해서는 모수를 추정해야 한다. 이를 위하여 표준화된 추정과정인 Box - Jenkins Method를 거친다. Box - Jenkins Method는 모형채택(Model Identification), 모수추정(Parameter Estimation), 모형진단(Model Diagnosis) 그리고 이를 활용한 예측으로 구성된다.

모형 채택은 자료를 SARIMA 적용에 적합한 형태로 가공하기 위해 필요한 최소 차분(일반/계절) 횟수를 정하고 SARIMA를 구성하는 세부모형(AR/MA/SAR/SMA) 중 필요한 종류를 가려내는 과정이다. 모수 추정은 각 세부모형의 적정 모수 범위를 설정하여 후보 예측모형들을 도출하고, 이 중 적정 신뢰도를 만족하는 모형을 가려낸 다음, 가장 적합한 모형을 추정하는 과정이다. 모형 진단은 해당 모형이 자료를 설명하기에 적합하게 설계되었는지를 최종 검증하는 과정이다. 이렇게 구성된 Box - Jenkins Method의 전 과정은 표 4와 같이 정리된다.

표 4. Box-Jenkins Method

Model Identification	Data Stationary Achievement: Differentiation AR/MA/SAR/SMA: (Partial) Autocorrelative Function(ACF/PACF) Pattern Matching
Parameter Estimation	Parameter Range: Augmented Dickey-Fuller (ADF) Unit Root Test Parameter Candidate: Rule of Parsimony Validity: Z-Value Test
Model Diagnosis Fitness Check	Likelihood: Akaike/Bayesian Information Criteria (AIC/BIC) Similarity: Ordinary Least Square (OLS) Regression Analysis Statistics Fitness: Residual White Noise(P-Value) Test
Model Selection and Forecast	Accuracy: Mean Absolute Percentage Error (MAPE)

Box - Jenkins Method에서 시계열 자료는 무작위 확률과정(Random Process)을 거친 결과로 간주되며 따라서 평균과 분산이 시간에 무관하게 불변인 정상상태(Stationary)임을 전제로 한다. 이러한 가정이 없을 경우, 표본 수(n)만큼의 확률과정에 따라 도출된 각 총 2n 개(평균(n) + 분산(n))의 모수를 구하는 것은 원천적으로 불가능하기 때문이다. 차분으로 자료를 정상상태로 가공하여, 구해야 할 모수를 2 개로 줄이고 선형 자기상관식을 유도할 수 있다. 정상상태 여부는 특정 횟수의 차분을 거친 자료가 시간차 횟수(Lag Frequency)에 따른 (부분)자기상관함수((Partial) Autocorrelative Function, (P)ACF)의 이론적인 유형 중 하나 이상에 해당되는지를 확인하여 알 수 있다. 이 과정에서 일반 및 계절차분 횟수에 해당하는 모수(d, D)를 구할 수 있다. 또한 유사한 상관함수의 이론적인 유형에 따라 p, q, P, Q 중 필요한 모수의 종류를 가려낼 수 있다. ACF와 PACF의 이론적인 패턴은 아래 표 5와 같다.

표 5. 모형별 ACF/PACF 상관함수 패턴

Model	ACF	PACF
AR(p)	Exponential Decrease Gradual Decrease by Sine Curve Rapid Convergence to 0 by Lag Increase	Disappearance from p+1
MA(q)		Exponential Decrease Gradual Decrease by Sine Curve Rapid Convergence to 0 by Lag Increase
ARMA(p, q)	The Same Pattern with AR(p) from q+1	The Same Pattern with MA(q) from p+1

이 자료에서는 일반 및 계절차분을 각 1 회씩 거쳤을 때 그림 2와 같이 위 표 4에서 제시한 것과 유사한 정상화된 패턴을 얻을 수 있으므로, 유력한 d와 D의 값은 1이다. 또한 월별로 집계된 해당 자료는 1 년마다 유사한 값이 반복되므로, 일반적인 계절 시간차 모수인 12를 적용할 수 있다.

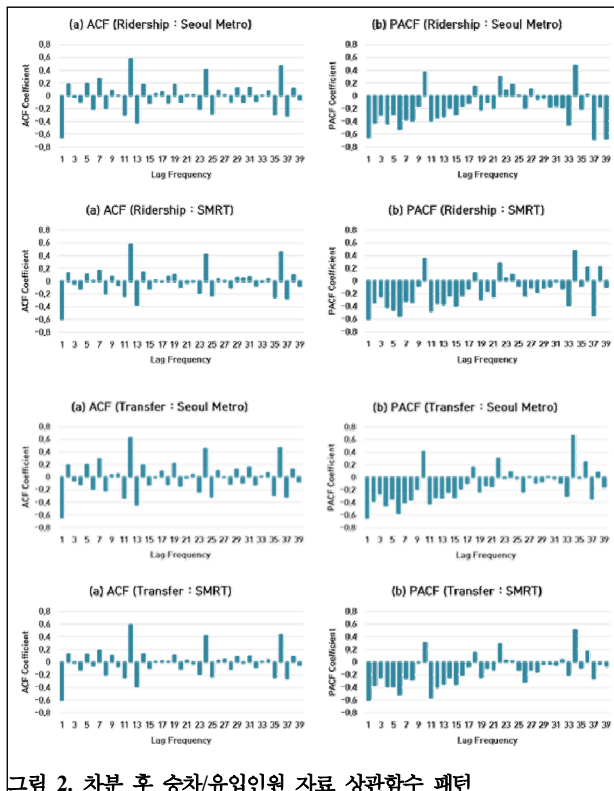


그림 2. 차분 후 승차/유입인원 자료 상관함수 패턴

이어서 Augmented Dickey - Fuller(ADF) 단위근 검사(Unit Root Test)를 통하여 각 모수의 범위를 설정한다. 모수의 숫자는 간절성의 원리(Rule of Parsimony)에 따라 같은 조건에서 적을수록 적합하다. 본 연구에서는 이를 통하여 유력한 모형이 (0, 1, 1)(P, 1, D)₁₂ 또는 (1, 1, 1)(P, 1, D)₁₂의 형태를 지닐 것으로 추정하였다. 이어 이론적인 (P)ACF 패턴과의 대조를 통하여 가능한 P와 D의 조합으로 (0, 0), (0, 1), (1, 0), (1, 1)을 도출하였다.

따라서 수요 예측모형으로 총 8 개((0,1,1)(0,1,0)₁₂, (0,1,1)(0,1,1)₁₂, (0,1,1)(1,1,0)₁₂, (0,1,1)(1,1,1)₁₂, (1,1,1)(0,1,0)₁₂, (1,1,1)(0,1,1)₁₂, (1,1,1)(1,1,0)₁₂, (1,1,1)(1,1,1)₁₂)의 후보군을 도출할 수 있으며, 이들 모형의 적절성 여부를 검증하기 위해 신뢰도 검사(Z - Value < 0.05)를 거쳐 승차는 3 개((0,1,1)(1,1,0)₁₂, (1,1,1)(1,1,0)₁₂, (1,1,1)(1,1,1)₁₂), 유입은 2 개((0,1,1)(1,1,0)₁₂, (1,1,1)(1,1,0)₁₂)의 유력한 모형을 추려낼 수 있다.

Z - Value 검사를 통과한 후보군 중 가장 예측에 적합한 모형을 선택할 때 최우선으로 참고할 기준은 우도(Likelihood)를

나타내는 척도이다. 이를 위하여 각 모형의 Akaike/Bayesian Information Criteria (AIC/BIC) 값을 비교하여 가장 낮은 경우를 선택한다. 표 6·7과 같이 AIC/BIC 검정을 실시한 결과, 승차와 환승에 대하여 각각 (1,1,1)(1,1,1)₁₂와 (1,1,1)(1,1,0)₁₂를 최적의 예측모형으로 추정하였다.

표 6. 승차인원 예측모형 후보군 AIC/BIC 검정 결과

Operator	Seoul Metro		SMRT	
SARIMA/IC	AIC	BIC	AIC	BIC
(0,1,1)(1,1,0) ₁₂	2344.632	2353.682	2262.524	2271.575
(1,1,1)(1,1,0) ₁₂	2341.955	2353.268	2257.575	2268.889
(1,1,1)(1,1,1) ₁₂	2319.743	2333.319	2238.116	2251.692
Estimation	(1,1,1)(1,1,1) ₁₂			

표 7. 유입인원 예측모형 후보군 AIC/BIC 검정 결과

Operator	Seoul Metro		SMRT	
SARIMA/IC	AIC	BIC	AIC	BIC
(0,1,1)(1,1,0) ₁₂	2189.577	2198.628	2146.630	2155.681
(1,1,1)(1,1,0) ₁₂	2188.369	2199.683	2145.506	2156.819
Estimation	(1,1,1)(1,1,0) ₁₂			

추정 모형의 적절성을 재차 확인하기 위하여 최소자승법(Ordinary Least Square(OLS)) 기반 회귀분석 결과가 후보군 중 가장 우수함을 확인한다. 실제로 해당 모형을 적용하였을 때의 회귀분석 결과를 나타내는 척도인 비례상수(Coefficient of Correlation(Coef.)), (다중)상관계수(R²), t - Value, 표준오차(Standard Error(Std. Err.)) 및 평균제곱근편차(Root Mean Square Error(RMSE))는 아래 표 8·9와 같이 각각 실제 자료에 가장 근접한 것을 암시한다.

표 8. 승차인원 예측모형 후보군 회귀분석 결과 비교

SARIMA	Operator	Coef.	R ²	t - Value	Std. Err.	RMSE
(0,1,1)(1,1,0) ₁₂	Metro	0.776	0.517	8.59	0.090	3.4·10 ⁵
	SMRT	0.897	0.688	12.32	0.073	2.0·10 ⁵
(1,1,1)(1,1,0) ₁₂	Metro	0.798	0.538	8.96	0.089	3.3·10 ⁵
	SMRT	0.917	0.710	13.01	0.071	1.9·10 ⁵
(1,1,1)(1,1,1) ₁₂	Metro	0.934	0.609	10.38	0.090	3.0·10 ⁵
	SMRT	0.998	0.761	14.81	0.067	1.7·10 ⁵

표 9. 유입인원 예측모형 후보군 회귀분석 결과 비교

SARIMA	Operator	Coef.	R ²	t - Value	Std. Err.	RMSE
(0,1,1)(1,1,0) ₁₂	Metro	0.872	0.685	12.26	0.071	1.2·10 ⁵
	SMRT	0.906	0.730	13.67	0.066	8.6·10 ⁵
(1,1,1)(1,1,0) ₁₂	Metro	0.866	0.700	12.70	0.068	1.1·10 ⁵
	SMRT	0.915	0.740	14.01	0.065	8.4·10 ⁵

마지막으로 예측모형이 자료를 설명하기에 적합하려면 모형에서 제외된 부분은 어떤 자기상관도 없는 순수 백색잡음이어야 한다. 이를 확인하기 위해 잔여백색잡음검사(Residual White Noise Test)를 실시, 가능한 시간차 범위(Lag Frequency) 내 P - Value가 0.05보다 항상 큰 것을 확인한다. 검사를 통과한 모형으로 수요를 예측, 자료와 예측의 평균절대오차율(Mean Absolute Percentage Error)을 월별로 산출한다.

Box - Jenkins Method를 거쳐 선택된 예측모형에 2006-2014 통계 중 2012년까지의 자료를 대입, 마지막 두 해(2013-2014)의 예측 수요를 산출하여 표 10·11과 같이 실제 수요와의 오차를 월별 MAPE의 평균으로 정리하였다. 이 과정에서 최종모형이 대개 후보군 중 가장 정확한 예측(= 가장 낮은 MAPE)를 도출함을 알 수 있다. 또, 예측기간을 1년 연장하여 도출한 결과를 같은 방법으로 비교한 결과 본 연구에서 채택한 SARIMA 모형은 최소 2년 간 매우 정밀한 수준의 수요예측이 가능함이 입증되었다. 도시철도 운영기관에서는 통상적으로 중기 운영 및 재정계획 기간을 향후 2년으로 설정하므로, 이에 참고자료로 활용하기 위한 정확성을 보장한다고 할 수 있다.

본 연구에서 승차인원과 유입인원의 최적 예측모형이 상이한 것은 두 통계자료의 통계적 양상이 다름을 의미한다. 승차인원은

순수하게 해당 운영기관에 1차로 접근한 수요에 해당하는 반면, 유입인원은 타 기관에서 추가로 유입된 수요에 해당하므로, 타 운영기관 측의 환경 변화 (신규 운영기관 및 노선의 출범/개통, 기존 노선 연장, 기타 대체 교통수단의 부상 및 주변지역 수요 유발 요인의 변화 등)에 따라 승차와 다른 패턴을 지닐 수 있다. 승차인원의 경우 서울메트로의 MAPE가 서울도시철도공사에 비하여 더 높는데 이것은 서울메트로 산하 노선이 수요에서 차지하는 비중이 더 크고 (약 5/3 수준), 인구와 각종 기능이 밀집한 지역을 더 많이 통과하여 수요 관련각종 환경의 변화에 더 큰 영향을 받기 때문으로 해석할 수 있다. 반대로 유입인원의 경우 서울도시철도공사의 MAPE가 서울메트로에 비하여 더 높는데 이것은 통계에서 제외된 타 노선(9호선, 신분당선 등)의 개통/확장으로 인한 통행패턴 변화로 인한 여파가 더 컸을 암시한다. 실제 수요는 전 기간에 걸쳐 대개 예측보다 다소 낮은 수준이었으며, 이는 서울시내 인구의 정체 및 감소와 광역철도 및 타 노선 확충에 따른 수요의 분산을 암시한다.

표 10. 승차인원 예측모형 후보군 평균절대오차를 비교

SARIMA	2013-2014			2015
	2013	2014	Total	
(0.1.1)(1.1.0) ₁₂	2.482	2.607	2.545	6.170
(1.1.1)(1.1.0) ₁₂	2.567	2.691	2.629	6.490
(1.1.1)(1.1.1) ₁₂	2.257	2.500	2.379	6.975

SARIMA	2013-2014			2015
	2013	2014	Total	
(0.1.1)(1.1.0) ₁₂	2.159	2.265	2.212	7.902
(1.1.1)(1.1.0) ₁₂	2.061	2.291	2.176	8.090
(1.1.1)(1.1.1) ₁₂	1.984	2.195	2.090	8.719

표 11. 유입인원 예측모형 후보군 평균절대오차를 비교

Seoul Metro				
SARIMA	2013~2014			2015
	2013	2014	Total	
(0.1.1)(1.1.0) ₁₂	2.898	5.041	3.969	10.293
(1.1.1)(1.1.0) ₁₂	2.280	3.288	2.784	7.370

SMRT				
SARIMA	2013~2014			2015
	2013	2014	Total	
(0.1.1)(1.1.0) ₁₂	2.767	6.261	4.570	14.460
(1.1.1)(1.1.0) ₁₂	2.766	6.375	4.514	14.675

III. 결론

통합적인 운영계획 수립의 필요성은 다양한 운영기관이 존재하는 철도교통망의 복잡성, 그리고 철도교통에 영향을 미치는 다양한 외부 요인에 기인한다. 또한 정확한 운영계획을 수립하기 위해서는 핵심 근거인 수요예측모형의 정확성이 우선적으로 담보되어야 한다. 따라서 본 연구에서는 기존 도시철도 수요통계자료 특성에 적합한 예측모형으로 계절성 ARIMA(SARIMA)를 채택하여 구체적인 모형을 수립, 실제 수요예측에 적용하고 결과를 비교하여 해당 모형이 유용함을 입증하였다. 이를 바탕으로 도시철도 운영계획의 신뢰성 제고가 가능하게 되었다. 본 연구에서 도출한 보다 정확한 수요예측모형은 연간 운영 및 재정계획 수립에 참고자료로 활용가능하며, 장래 타 운영기관 설립 시에도 확대 적용할 수 있다. 또 예측수요와 실제 결과에 대한 심층 분석, 도시철도 사고통계자료를 조합하여 이용객 서비스 개선 및 안전관리에도 참고할 수 있다. 향후 연구로는 보다 정밀한 통계수집 방안과 장기적인 예측을 위한 모형의 확장, 사회경제적 변수를 감안한 예측모형의 다각화 등이 필요할 것으로 보인다.

후기

본 연구는 국토교통부 철도 특성화대학원 지원사업에 의해 수행되었습니다.

참고문헌

- Boyce, D.E., Zhang, Y.F., Lupa, M.R., (1994), Introducing "Feedback" into Four-Step Travel Forecasting, Transport Research Record: Journal of the Transportation Research Board, 1443(9), TRB, pp. 65-74.
- Zhao, Y., Kockelman, K.M., (2002), The Propagation of Uncertainty through Travel Demand Models: An Exploratory Analysis, The Annals of Regional Science, 36(1), Western Regional Science Association, pp. 145-163.
- McNally, M.G., (2007), The Four Step Model, University of California, Irvine(Irvine, The United States of America), 18 pages.
- Lim, K.K., Kim, S.S., Chung, S.B., (2013), Activity - based Approaches for Travel Demand Modeling: Reviews on Developments and Implementations, Journal of the Korean Society of Civil Engineers, 33(2), KSCE, pp. 719-727.
- Client Service Headquarter(2006-2014), Annual Plan, Seoul Metro(Seoul, Republic of Korea)
- Client Service Headquarter(2006-2014), Annual Plan, Seoul Metropolitan Rapid Transit(SMRT) Corporation(Seoul, Republic of Korea)
- Achieved at www.seoulmetro.co.kr/page.action?mCode=E030050000&cidX=721
- Achieved at www.smrt.co.kr/main/publish/view.jsp?menuID=001007004004010
- Kim, B.S., Lee, C.W., Sim, J.H., Lee, S.K., Kim, J.H., Park, H.G. (2013) Forecasting Passenger Transport Demand of Joongang Line using Seasonal ARIMA Model, Proceedings of the 2013 Autumn Conference & Annual Meeting of the Korean Society for Railway, the Korean Society for Railway, pp. 71-82.
- Kim, K.H., Kim, H.S., (2011), KTX Passenger Demand Forecast with Intervention ARIMA Model, Journal of the Korean Society for Railway, 14(5), the Korean Society for Railway, pp. 470-476.
- Williams, B.M., Hoel, L.A., (2003), Modelling and Forecasting Vehicular Traffic Flow as a Seasonal ARIMA Process: Theoretical Basis and Empirical Results, Journal of Transportation Engineering, 129(6), American Society of Civil Engineers, pp. 664-672.
- Brockwell, P.J., Davis, R.A., (1996), Introduction to Time Series and Forecasting, Springer(New York, The United States of America)
- Box, G.E.P., Jenkins, G.M., (1970), Time Series Analysis: Forecasting and Control, Holden - Day, (San Francisco, The United States of America)