

항공화물수요예측에서 계절 ARIMA모형 적용에 관한 연구: 인천국제공항발 미주항공노선을 중심으로

서보현 · 양태웅 · 하한구*
인하대학교 물류전문대학원

Application of SARIMA Model in Air Cargo Demand Forecasting: Focussing on Incheon-North America Routes

SUH, Bo Hyoun · YANG, Tae Woong · HA, Hun-Koo*

Graduated School of Logistics, Inha University, Incheon 22212, Korea

*Corresponding author: hkha@inha.ac.kr

Abstract

For forecasting air cargo demand from Incheon National Airport to all of airports in the United States (US), this study employed the Seasonal Autoregressive Integrated Moving Average (SARIMA) method and the time-series data collected from the first quarter of 2003 to the second quarter of 2016. By comparing the SARIMA method against the ARIMA method, it was found that the SARIMA method performs well, relatively with time series data highlighting seasonal periodic characteristics. While existing previous research was generally focused on the air passenger and the air cargo as a whole rather than specific air routes, this study emphasized on a specific air cargo demand to the US route. The meaningful findings would support the future research.

Keywords: air cargo, air cargo volume forecasting, ARIMA, demand forecasting, seasonal ARIMA

초록

본 연구는 2003년 1사분기부터 2016년 2사분기 까지 인천국제공항에서 미주노선을 통하여 미주 내 공항에 도착하는 항공화물의 시계열 자료를 통하여 SARIMA 모형을 활용하여 항공화물 수요예측을 시행하였다. 또한 SARIMA 모형을 활용하여 만들어진 수요예측 모형과 기존 연구에 주로 활용되어졌던 ARIMA 모형을 활용하여 만들어진 수요예측 모형과 비교분석함으로써, 주기적인 특성 및 계절성을 가진 시계열 자료에 대한 SARIMA 모형의 상대적으로 우수한 예측 정확성을 입증하였다. 기존의 항공 관련 연구는 주로 여객에 관한 연구가 상대적으로 많았다. 또한 화물과 관련된 연구에서도 특정노선이 아닌 공항이나 전체에 대한 연구가 대부분이었다. 이러한 상황에서, SARIMA 모형을 활용하여 미주지역이라는 특정 노선에 대한 항공화물의 수요를 예측한 본 연구는 큰 의의가 있다고 생각된다.

주요어: 계절성ARIMA, 수요예측, 항공화물, 항공물동량 예측

J. Korean Soc. Transp.
Vol.35, No.2, pp.143-159, April 2017
<https://doi.org/10.7470/jkst.2017.35.2.143>

pISSN : 1229-1366
eISSN : 2234-4217

Received: 31 October 2016

Revised: 1 February 2017

Accepted: 24 April 2017

Copyright ©
Korean Society of Transportation

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

서론

산업기술이 점점 발전함에 따라 소형 고부가가치 제품인 전자기기, 반도체, 등의 수출비중이 계속해서 증가하는 추세이다. 이러한 제품 품목의 운송은 항공을 통한 운송이 주를 이룬다. 이에 대한 가장 큰 이유는 항공운송의 특성상 해상운송에 비해 운송 속도가 빠른 것이다. 따라서 이러한 제품들의 수출량이 증가함에 따라 항공운송을 통한 교역량이 꾸준히 증가하는 추세이다. 실제로 한국무역협회에 따르면 지난 20여 년 동안 연 평균 약 5.3%씩 우리나라의 전체 수출입 물동량은 계속해서 증가하였다. 이 가운데 항공운송을 통한 교역량은 전체 교역량의 1%도 미치지 못하지만, 교역액을 기준으로 하면 전체 교역액의 20%이상을 차지하고 있다. 이는 항공을 통해 운송되는 화물이 고부가가치 제품이 주를 이루기 때문이다. 또한 이러한 특성의 품목은 계속해서 수출이 증가할 추세이며, 이와 더불어 항공화물 운송의 중요성이 높아지고 있다.

항공화물 운송의 중요성이 높아지면서 항공화물의 수요예측은 중요성이 커지고 있는 연구주제이다. 그렇기에 본 연구는 인천국제공항에서 출발하여 미주지역에 도착하는 항공화물에 대한 시계열 자료를 통해 수요예측을 연구하였다. 미주 항공노선의 예측력을 강화시킴으로써, 기대수요에 맞춰 공급량을 조절할 수 있고, 공급의 부족이나 과잉 공급으로 일어나는 손실을 최소화시킬 수 있을 것이다. 미주항공노선의 미래수요 예측을 통해 향후 항공사들의 경영전략의 수립 및 항공과 관련된 교통정책의 수립에 기여하고자 한다.

본 연구는 시계열 분석기법(Time Series Analysis)기법 중 계절성(Seasonality)을 가진 시계열 자료의 수요예측으로 적합한 SARIMA(Seasonal Autoregressive Integrated Moving Average)기법을 사용하여 항공화물의 수요를 예측하였다. 사용된 시계열 자료로는 인천국제공항에서 출발하여 미주노선을 통해 전체 미주 공항에 도착한 항공화물에 대한 분기별 시계열 자료이며, 2003년 1사분기부터 2014년 4분기까지 총 48개의 분기별 항공화물 물동량 자료를 활용하였다. 또한 이번 연구는 SARIMA 모형의 예측결과와 계절성(Seasonality)을 포함하지 않은 예측모형인 ARIMA(Autoregressive Integrated Moving Average)모형의 결과를 비교하고자 한다. 이는 다음과 같은 이유에서 비롯된다. 기존연구에서는 대부분 ARIMA 모형을 사용하여 계절성을 반영하지 않아 예측의 정확도에 대해 의문이 제기된다. 따라서 이번 연구에서는 SARIMA모형과 ARIMA모형의 비교를 통해 SARIMA를 사용하는 것이 더 좋다는 것을 보일 것이다.

이번 연구는 기존 연구에 비해 다음과 같은 차별성을 지닌다. 첫째, 현존하는 연구는 항공여객 부문에 집중되어 있으며 항공화물에 대한 연구는 상대적으로 많지 않았다. 예를 들어, Erma et al.(2010)은 System Dynamics를 활용하여 항공 여객 수요 및 공항 터미널 수용력을 예측하였다. 최종해 et al.(2015)는 유형, 직항, 이원, 배후 및 브리지 수요로 구분하여 수요를 유형별로 통계적으로 유의한 설명변수를 선정하는 회귀식을 통하여 인천국제공항의 신규 항공노선에 대한 수요예측을 제시하였다.

Arnaldo(2013)은 Integrated mixture of local experts model을 통하여 상파울로 국제공항을 이용한 항공 여객 수요를 예측하였다. Pablo(2004)는 Box-jenkins 방법론을 통하여 Santander 공항을 이용하는 항공여객을 비즈니스 여객과 레저 여객으로 구분하여 각 수요에 대해 예측연구를 시행하였다. 이러한 연구들은 다양한 방법론을 통하여 항공에 대한 수요를 예측하고 미래 수요를 제시하였으나, 항공여객 분야에 중점적으로 연구되었다. 둘째, 기존의 선행연구들 중 인천국제공항에서 출발하여 미주노선을 통해 미주 전체 공항에 도착하는 항공화물에 대한 연구는 전무하다. Chou et al.(2011)은 Fuzzy regression을 활용하여 항공화물 전체 물동량에 대한 예측을 시행하였으나 특정 공항을 기준으로 하는 화물 수요를 예측하여, 전체 미주공항에 대한 분석을 실시하는 이번 연구에 비해선 범위가 매우 한정적이다.

방법론적인 측면으로 살펴보았을 때, ARIMA, Seasonal ARIMA 모형을 이용한 항공과 관련한 선행연구는 찾을 수 있으나, 대부분 항공여객에 초점을 맞추었다. 특히 국내에서는 교통과 관련하여 SARIMA 모형을 이용한 선행연구는 매우 찾기 어려웠다. 아리온토야(2013)은 인천-울란바토르 노선의 항공여객에 대하여 Seasonal ARIMA, Holt-Winters모형을 이용하여 수요를 예측함으로써, 인천-울란바토르 노선의 미래 수요를 예측하였다. Chen et

al.(2009)는 Holt-Winter 모형과 SARIMA, GRE 모형을 활용하여 대만에 방문한 항공여객 수에 대한 예측을 연구하였다. Chang et al.(2010)은 SARIMA 모형, 이동평균법, 지수평활법을 활용하여 대만에서 일본, 미국, 홍콩으로 향하는 관광객 수에 대한 예측을 연구하였다. Yoon et al.(2010)은 계절형 VAR, 계절형 ARIMA를 통하여 국제항공 여객과 화물의 수요예측에 관하여 연구하였다. Kim et al.(2014)는 계절 ARIMA 모형을 활용하여 국내선 제주권과 내륙권의 저가항공 수요를 예측 비교분석하였다. Kim et al.(2015)는 계절 ARIMA 모형을 활용하여 인천공항에서 저가항공을 이용하는 여객의 수요예측에 관한 연구를 시행하였다. 앞서 언급하였듯이, Seasonal ARIMA 모형을 활용한 국내·외 항공 관련 연구들은 대부분 항공화물이 아닌 여객부문에 대한 수요 예측을 연구하였으며, SARIMA 모형을 활용하여 항공화물에 대한 연구를 한 경우는 그 수가 적었으며, 항공화물에 관한 연구에서도 특정 지역이 아닌 전체적인 화물 물동량에 대한 수요예측 연구가 주를 이루었다.

특정 노선을 중심으로 한 항공화물에 대한 연구는 Min et al.(2013)이 유일하다. 그들은 계절 ARIMA 모형을 활용하여 인천국제공항에서 출발하여 유럽노선을 통해 유럽 전체 공항에 도착하는 항공화물 물동량을 예측하였지만, 유럽노선에 한정시켜 분석을 실시했다. 한국무역협회에 따르면 현재 미국과 우리나라간의 교역량은 중국 다음으로 많은 부분을 차지하고 있다. 교역량은 중국이 가장 많은 부분을 차지하지만, 항공사 입장에서는 Revenue Tonne Kilo-meter가 교역량보다 더욱 중요한 지표이며, Revenue Tonne Kilometer를 지표로 하게 된다면 미국노선이 매우 중요하다고 판단되어진다. 그렇기에 인천국제공항에서 미주노선을 통해 미주 전체 공항에 도착하는 화물의 수요를 예측하는 연구는 향후 우리나라 경제정책 수립에 있어서 많은 도움을 줄 것으로 기대된다. 또한 SARIMA 모형은 기존 ARIMA 모형의 단점을 보완한 모형으로 계절성을 가진 항공화물 물동량에 유용한 수요예측 방법론으로 활용될 수 있다. 아직까지 국내에서는 교통 화물수요 예측부문에 있어서 SARIMA 모형이 자주 활용되어지지 않고 있기에 상대적으로 연구가 적은 노선을 중심으로 한 항공화물에 대한 수요를 SARIMA 모형을 활용하여 예측함으로써 이후 항공화물 수요예측부문과 더불어 다른 교통 분야의 수요예측에서도 SARIMA 모형이 활용될 것이라 기대된다.

방법론

1. ARIMA 모형 및 SARIMA 모형

ARIMA(Autoregressive Integrated Moving Average) 모형은 ‘종속변수 Y_t 는 독립변수 X_t 에 의해 설명된다.’라는 일반적인 회귀방법론이 아니라 ‘종속변수 Y_t 는 Y_t 의 과거값과 확률적 오차항에 의해 설명된다.’는 경제 시계열의 확률적 특성에 기반을 둔 예측방법론이다(Box G. E. P, 1976). ARIMA모형을 활용하여 수요를 예측하기 위해서는 주어진 시계열 자료의 안정성(stationary)을 우선적으로 확인하여야 한다. 만약 시계열 자료가 불안정적일 경우에는 차분(differencing) 혹은 적분(Integrated)의 과정을 통하여 불안정적인 시계열 자료를 안정적인 시계열 자료로 바꾸는 과정을 거쳐야 한다. ARIMA(p,d,q) 모형의 기본형태는 Equation(1)로 나타낼 수 있다.

$$\begin{aligned} (1 - \Phi_1 L - \Phi_2 L^2 - \dots - \Phi_p L^p)(1 - L)^d Z_t \\ = \delta + (1 - \Theta_1 L - \Theta_2 L^2 - \dots - \Theta_q L^q) \epsilon_t \end{aligned} \tag{1}$$

Z_t : 원시계열 자료

t : 시간

ϵ_t : $N(0, \sigma^2)$ 을 따르는 오차항, 백색잡음

L : 후향연산자

p : 자기회귀(Autoregressive)항의 차수

q : 이동평균(Moving average)항의 차수

d : 차분의 차수

δ : 상수

Equation(1)에서 $1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p$ 은 자기회귀(Autoregressive; AR)항을 의미하고, $1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q$ 은 이동평균(Moving average; MA)항을 의미한다. 또한 위의 후향연산자를 사용하여 ARIMA(p,d,q) 모형을 간단하게 표현하면 Equation(2)의 형태를 가진다.

$$\Phi_p(L)(1-L)^d Z_t = \delta + \Theta_q(L)\epsilon_t \tag{2}$$

ARIMA(p,d,q) 모형에서 만약 $d=0$ 이라면, ARIMA(p,q)모형이 된다.

ARIMA(p,d,q) 모형을 통한 수요예측은 다양한 분야에서 널리 사용되고 있다. 하지만 ARIMA모형은 원 시계열 자료가 계절성과 같은 주기적인 특성을 가진 경우에는 시계열 자료의 주기적인 특성까지 반영하기 힘들다는 문제점이 존재한다. SARIMA(Seasonal Autoregressive Integrated Moving Average)모형은 ARIMA모형의 이러한 문제점을 보완한 모형이다. 또한 SARIMA 모형은 역시 ARIMA 모형과 동일하게 ‘종속변수 Y_t 는 Y_t 의 과거값과 확률적 오차항에 의해 설명된다.’는 경제 시계열의 확률적 특성에 기반을 둔 예측방법론이다. SARIMA(p,d,q)(P,D,Q)s 모형은 Equation(3)으로 나타낼 수 있다.

$$\begin{aligned} \phi_p(L)\Phi_P(L^s)(1-L)^d(1-L^s)^D Z_t \\ = \delta + \theta_q(L)\Theta_Q(L^s)\epsilon_t \end{aligned} \tag{3}$$

Z_t : 원시계열 자료

t : 시간

ϵ_t : $N(0, \sigma^2)$ 을 따르는 오차항, 백색잡음

L : 후향연산자

P : SAR(Seasonal Autoregressive)항의 차수

Q : SMA(Seasonal Moving Average)항의 차수

D : 계절차분의 차수

p : 자기회귀(Autoregressive)항의 차수

q : 이동평균(Moving average)항의 차수

d : 차분의 차수

δ : 상수

Equation(3)에서 $\Phi_P(L^s)$ 는 계절자기회귀(Seasonal Autoregressive; SAR)항을 의미하고, $\Theta_Q(L^s)$ 는 계절이동평균(Seasonal Moving Average; SMA)항을 의미한다. 또한 $\phi_p(L)$ 과 $\theta_q(L)$ 는 각각 자기회귀(Autoregressive ; AR)항과 이동평균(Moving Average ; MA)항을 나타낸다. SARIMA 모형은 앞서 언급한바와 같이 ARIMA 모형을 보완하여 개발된 기법이다. SARIMA 모형이 어떠한 형태로 보완이 되었는지 확인하기 위해서는 앞서 설명한 ARIMA 모형을 나타낸 Equation(1)과의 비교를 통하여 확인 가능하다.

Equation(1)과 Equation(3)을 비교하여 보면, SARIMA 모형은 ARIMA 모형에 SAR항과 SMA항이 추가되어 있는 모형임을 알 수 있다. 즉, 기존 ARIMA 모형에서 SAR항과 SMA항을 추가함으로써, SARIMA모형은 원 시계

열자료의 주기적인 특성 혹은 계절성을 고려한 수요예측을 할 수 있도록 기존 ARIMA모형을 보완한 수요예측 방법론이다.

2. 모형 추정 절차

ARIMA 모형과 SARIMA 모형은 Box-jenkins 분석과정에 따라 주어진 시계열 자료를 통하여 모형 추정 및 미래 수요를 예측한다. Box-jenkins 분석과정은 Figure 1과 같다.

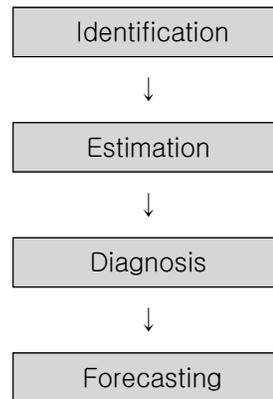


Figure 1. Flow of box-jenkins method

Figure 1과 같이 Box-jenkins 방법론은 크게 식별(Identification), 추정(Estimation), 검진(Diagnosis), 예측(Forecasting) 이렇게 4단계를 거친다. 식별(Identification)단계에서는 SARIMA(p,d,q)(P,D,Q)s모형을 추정하게 된다(Box G. E. P, 1976).

우선 시계열 자료가 안정적인지 여부를 판단하기 위해 안정성 검정을 실시해야 한다. 안정성 검정으로는 그래프 분석, 상관도표(Corrleogram) 분석, 단위근 검정 등의 방법이 있으며, 이러한 안정성 검정을 통하여 'd', 'D'의 값이 결정된다. 다음으로는 안정성 검정을 통해 변화된 안정적인 시계열 자료의 상관도표(Correlogram)를 활용하여 자기상관 함수(Autocorrelation function; ACF) 및 편 자기상관함수(Partial Autocorrelation function; PACF)를 분석하거나 DF-GLS 방법론을 활용하여, AR(Autoregressive)항의 차수 'p', MA(Moving Average)항의 차수 'q', SAR(Seasonal Autoregressive)항의 차수 'P', SMA(Seasonal Moving Average)항의 차수 'Q'의 값을 결정하게 된다. 이러한 과정을 통해 SARIMA(p,d,q)(P,D,Q)s모형의 각각의 차수인 'p', 'd', 'q', 'P', 'D', 'Q'를 결정하게 된다.

추정(Estimation) 단계는 앞선 단계를 통하여 결정된 'p', 'd', 'q', 'P', 'D', 'Q'을 활용하여, AR항, MA항, SAR항, SMA항 각각의 모수를 추정하게 된다. 모수를 추정하는데 있어서 널리 사용되는 방법으로는 최소제곱 추정법(Least Squares Estimation Method), 최대우도 추정법(Maximum Likelihood Estimation Method), 적률 추정법(Method of Moment Estimation), 비선형 추정법(Nonlinear Estimation Method) 등이 있다. 이러한 방법들을 통하여 모형을 추정할 때, AIC(Akaike Information Criterion,) BIC(Bayesian Information Criterion)의 값과 동시에 유의성 검정 및 모델의 간결성 등을 고려하여 모형의 최종 차수를 결정하여야 한다.

진단(Diagnosis)의 단계란, 추정된 SARIMA모형이 자료와 잘 부합하는지 여부를 점검하는 단계이다. 잔차분석(Residual Analysis), 과대적합진단(Overfitting Diagnosis)등의 유의성 검정을 활용하여 판단을 한다. 진단의 단계에서 최종으로 결정된 모형이 이상이 없다고 판단될 시, 예측(Forecasting)의 단계로 넘어가 추정모형을 통해 미래에 대한 예측이 시행되어지고, 추정 모형에 이상이 있다고 판단되어지면, 다시 식별의 단계로 돌아가 계속해서 반복 과정을 거치게 된다.

예측(Forecasting)의 단계는 Box-jenkins 분석과정의 마지막 단계로서, 앞서 이루어진 단계를 통하여 최종으로 결정된 모형을 활용하여 미래수요를 예측하는 단계이다. 예측된 수요는 시간이 흐른 후, 실제 나타나는 수치와의 비교를 통하여 최종 모형의 검증에 할 수 있다.

모형추정

1. 분석자료

본 연구에 사용된 자료는 인천국제공항에서 출발하는 미주 항공노선을 통해 도착한 항공화물의 분기별 시계열 자료로서, 인천국제공항공사로부터 습득한 월별 항공 통계 자료를 분기별 자료로 통합하였다. 구체적인 기간은 2003년 1월부터 2016년 6월까지의 시계열 자료를 2003년 1사분기부터 2016년 2사분기까지 총 52개의 분기별 항공화물 물동량으로 통합하였다. 이 중 2003년 1사분기부터 2014년 4사분기까지의 항공화물 물동량을 모형 추정에 대한 자료로 사용하였으며, 이렇게 추정된 예측 모형을 통해 2015년 1사분기부터 2016년 2사분기까지의 예측치를 구한 후, 실제 물동량과 비교하여 그 예측정확도를 비교할 계획이다. Figure 2는 2003년 1사분기부터 2016년 2사분기까지의 인천공항 발 미주 항공노선의 항공화물 물동량에 대한 그래프이다. Figure 2를 보면 2008년 1사분기부터 화물 물동량이 급격하게 감소하였다. 이는 미국 서브프라임 모기지 사태 등으로 인한 미국 경제시장의 불황이 주 원인이다. 이후 2009년 3사분기부터 화물물동량이 다시 회복세를 보였으나, 2010년 3사분기부터 다시 항공 화물물동량이 전체적으로 감소하는 경향을 나타내었다. 이는 그리스 사태 여파로 인한 세계적인 경제불황이 원인으로 판단된다. 또한 본 연구에서 활용하는 항공화물 물동량의 계절성을 파악하기 위하여 원자료에 1계 차분을 통하여 추세를 제거하였더니 Figure 3과 같이 나타났다. Figure 3을 보면 매년 1사분기마다 계절 스파이크가 나타났으며, 이를 통해 직관적으로 4분기 단위로 계절주기가 있다고 판단하였다.

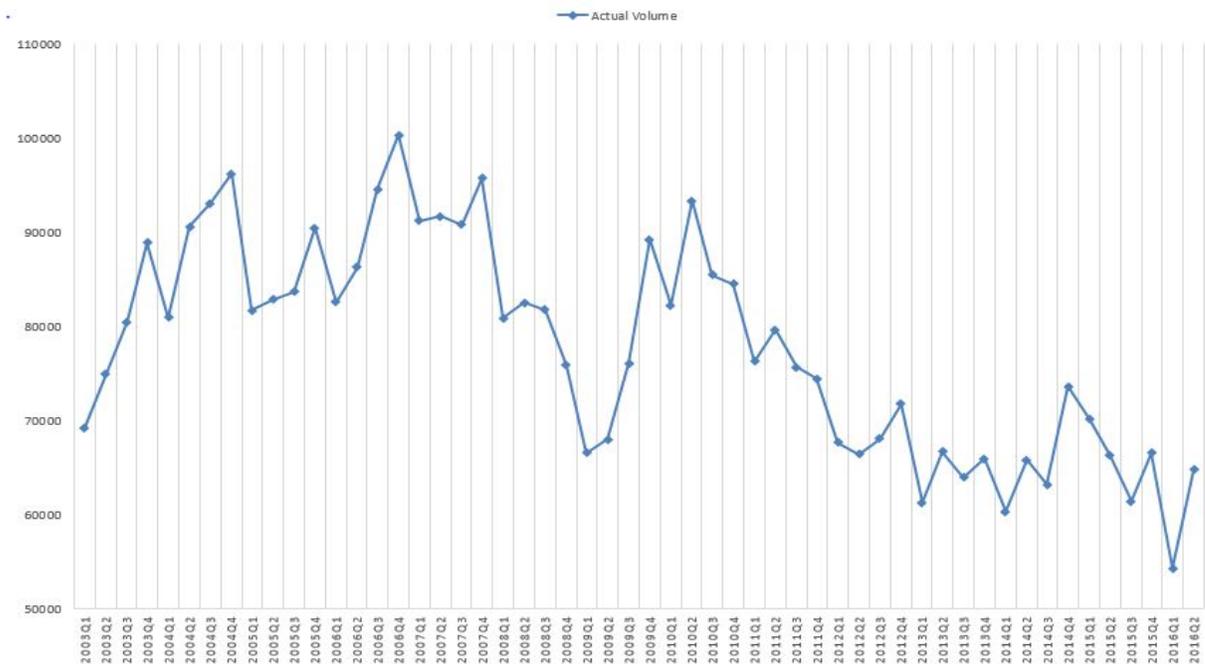


Figure 2. Air cargo volume from ICN to America quarterly (2003 Q1 - 2016 Q2)

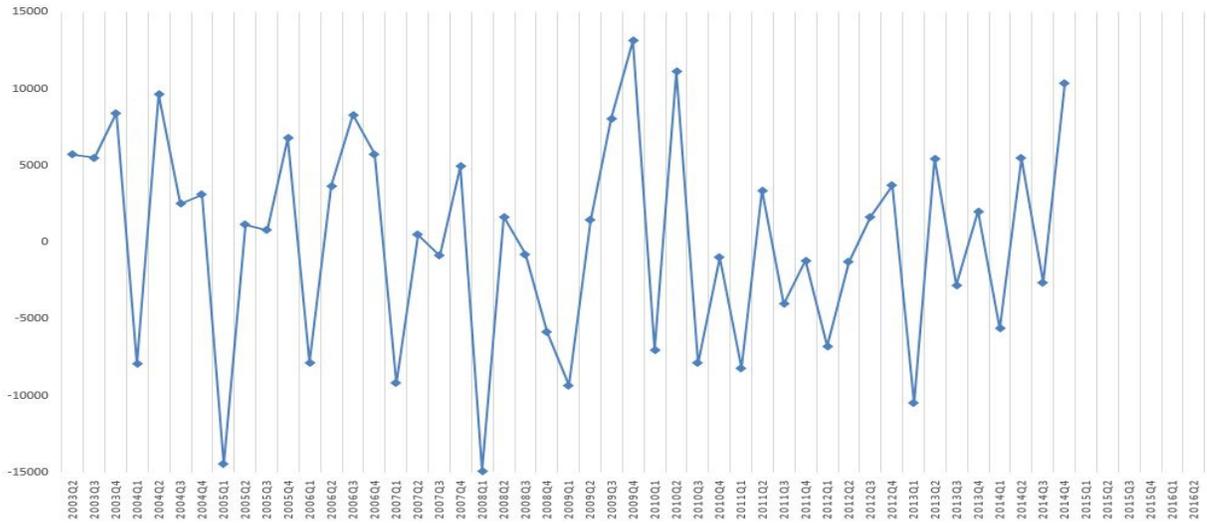


Figure 3. Timeseries data of 1 order differencing (2003 Q2 - 2014 Q4)

구체적인 수치적 분석은 이후 모형추정에서 다룰 계획이다. Figure 3을 통해 본 연구에 활용되는 인천국제공항에서 출발하여 미주노선을 통해 미주 전체 공항에 도착하는 항공화물 물동량은 계절성 및 주기적 특성을 가지고 있기에 기존 ARIMA모형보다는 계절성을 감안하는 SARIMA모형을 통한 수요예측이 적합함을 알 수 있다.

2. 안정성 검증

앞서 언급한 바와 같이, SARIMA모형을 추정하기 위해서는 box-jenkins 분석과정에 따라 시계열 자료의 정상성 여부를 확인하여야한다. 그렇기에 시계열 자료가 안정적인지 여부를 우선적으로 확인하고 모형을 추정하여야한다. 또한 본 연구는 SARIMA 방법론을 사용하며, 활용되어지는 시계열 자료가 분기별 자료이기에, 계절 단위근 검정(Seasonal Unit root test)도 시행하였다.

본 연구에서는 상관도표(Correlogram) 분석과 Augmented Dickiey-Fuller(ADF) 검정을 활용하여 안정성의 여부를 확인하고, Hylleberg-Engle-Granger-Yoo(HEGY)검정을 통하여 계절 단위근의 여부를 확인하였다.

Table 1을 살펴보면 자기상관계수가 lag 1에서 높은 값(0.7862)으로 시작하여 lag가 길어질수록 감소하는 것으로 확인할 수 있다. 이를 통하여 원 시계열 자료가 평균 또는 분산에 대하여 불안정적이란 것을 의미한다. Table 2에서는 원 시계열 자료에 대한 ADF 검정결과를 알 수 있다. ADF 검정은 임의행보(Random walk)과정에 표류경향(Drift)이 없을 경우와 임의행보(Random walk)과정에 표류경향(Drift)이 있을 경우, 그리고 임의행보(Random walk)과정에 표류경향(Drift)과 추세(Trend)가 존재할 수 있는 경우에 대하여 실시하였다.

Table 2를 보면 표류경향이 없는 임의행보모형, 표류경향이 있는 임의행보모형 2가지 모형에 대하여 5% 수준 하에서 ‘시계열 자료가 단위근을 갖는다.’ 귀무가설을 기각할 수 없다. 그러므로 원 시계열 자료는 안정적이지 않은 시계열 자료라고 판단할 수 있다.

원 시계열 자료가 안정적이지 않은 시계열 자료임을 확인 하였기에, 차분의 과정을 통해 원 시계열 자료를 안정성을 가지도록 변환시키는 단계를 거쳐야 한다. Table 3은 원 시계열 자료를 1계 차분과 4계 계절 차분을 통하여 변환된 시계열 자료에 대한 상관도표이며, Table 4는 원 시계열 자료를 1계 차분과 4계 계절 차분을 통해 변환된 시계열 자료의 ADF검정 결과를 나타낸다. Table 3을 확인해보면 자기상관이 거의 모든 시차에서 0주변에 머물고 있음을 알 수 있다. 또한 Table 4를 살펴보면 표류경향이 없는 임의행보모형, 표류경향이 있는 임의행보모형, 추세가 존재하는 임의행보모형 이 3가지 모형 모두 1% 수준 하에서 ‘시계열 자료가 단위근을 갖는다.’는 귀무가설이 기각됨을

Table 1. Correlogram of raw data

Lag	AC	PAC	Q	Prob>Q	Autocorrelagtion			Partial Autocorr.		
					-1	0	1	-1	0	1
1	0.7862	0.7911	31.564	0.0000						
2	0.6739	0.1785	55.258	0.0000						
3	0.5165	-0.1663	69.485	0.0000						
4	0.4991	0.2639	83.070	0.0000						
5	0.2819	-0.5118	87.505	0.0000						
6	0.2123	0.1791	90.080	0.0000						
7	0.1657	0.2628	91.687	0.0000						
8	0.2523	0.2361	95.507	0.0000						
9	0.1443	-0.3570	96.788	0.0000						
10	0.1451	0.2265	98.118	0.0000						
11	0.1188	0.1304	99.033	0.0000						
12	0.2104	0.2023	101.98	0.0000						
13	0.0800	-0.4267	102.42	0.0000						
14	0.0472	0.0896	102.58	0.0000						
15	-0.0186	-0.0002	102.61	0.0000						
16	0.0072	0.0954	102.61	0.0000						
17	-0.1208	-0.0630	103.74	0.0000						

Table 2. Result of ADF unit root test

(raw data)

	tau value	P> t	lags
Random Walk	-0.98	0.335	8
Random Walk with Drift	-0.98	0.335	8
Random Walk with Drift and Trend	-2.86	0.008	8

알 수 있다.

추가적으로 1계 차분과 4계 계절 차분을 통해 변환된 시계열 자료에 대하여 HEGY 검정을 추가로 실시하여 변환된 시계열 자료가 계절 단위근을 가지고 있는 지 여부를 확인하였다.

Table 5에서는 변환된 시계열 자료의 HEGY 검정 결과를 알 수 있다. HEGY 검정 결과 모든 귀무가설이 기각됨을 알 수 있다. 이를 통하여 1계 차분 및 4계 계절 차분된 시계열 자료는 단위근, 계절 단위근 모두를 갖지 않는 안정적인 시계열 자료임을 판단 할 수 있다.

3. 모형 추정

SARIMA(p,d,q)(P,D,Q)s 모형을 추정하는 것에 있어서 ‘p’, ‘d’, ‘q’, ‘P’, ‘D’, ‘Q’의 값을 찾는 것이 가장 핵심적인 사항이다. 우선적으로 ‘d’와 ‘D’의 값은 1계 차분과 4계 계절 차분한 자료를 단위근 검정을 통해, 안정성을 가진다고 판단하였기에 ‘d=1’, 그리고 ‘D=1’임을 확인 가능하다. 다음으로 ‘p’, ‘q’, ‘P’, ‘Q’의 값을 추정하여야 하는데, ‘p’, ‘q’, ‘P’, ‘Q’의 각각의 값을 결정하는데 있어서 크게 2가지의 요소를 확인하여야 한다. 첫 번째 요소는 Akaike Information Criterion(AIC)와 Bayesian Information Criterion(BIC)의 값을 확인하는 것이다. 두 번째 요소로는,

Table 3. Correlogram of regular and seasonal differencing

Lag	AC	PAC	Q	Prob>Q	Autocorrelagtion			Partial Autocorr.		
					-1	0	1	-1	0	1
1	0.1867	0.1937	1.6052	0.2052						
2	0.1620	0.1346	2.8430	0.2413						
3	-0.0476	-0.1113	2.9526	0.3990						
4	-0.5398	-0.6086	17.409	0.0016						
5	-0.1045	0.1103	17.965	0.0030						
6	-0.2604	-0.0650	21.511	0.0015						
7	-0.0297	0.0783	21.558	0.0030						
8	-0.0675	-0.5781	21.81	0.0053						
9	-0.0859	0.0134	22.229	0.0082						
10	0.0718	-0.1434	22.532	0.0126						
11	-0.0334	-0.0353	22.599	0.0201						
12	0.2682	-0.2104	27.09	0.0075						
13	0.1563	0.0199	28.666	0.0073						
14	0.1424	-0.2170	30.02	0.0076						
15	0.1665	0.6665	31.936	0.0066						
16	-0.1974	-0.3604	33.289	0.0068						
17	-0.0615	-0.2907	33.571	0.0095						

Table 4. Result of ADF unit root test (data after regular and seasonal differencing were applied)

	tau value	P> t	lags
Random Walk	-4.26	0.000	7
Random Walk with Drift	-4.26	0.000	7
Random Walk with Drift and Trend	-4.37	0.000	7

Table 5. Result of HEGY test (data after regular and seasonal differencing were applied)

	Frequency			lags
	Zero ($t: \pi_1 = 0$)	Biannual ($t: \pi_2 = 0$)	Annual ($F: \pi_3 = 0 \cap \pi_4 = 0$)	
Noconstant	-5.697 (0.000)	-5.252 (0.000)	-8.996 (0.000)	0
Constant	-5.592 (0.000)	-5.176 (0.000)	-8.865 (0.000)	0
Constant with Trend	-5.587 (0.000)	-5.123 (0.000)	-8.773 (0.000)	0

각 항들의 계수의 z-value를 확인해야 한다.

첫 번째 요소인 AIC와 BIC의 추정식은 Equation(4)와 같다.

$$\begin{aligned}
 AIC &= -2 \times \ln(\text{likelihood}) + 2 \times k \\
 BIC &= -2 \times \ln(\text{likelihood}) + \ln(N) \times k
 \end{aligned}
 \tag{4}$$

k : 모수의 수

N : 관측수

Equation(4)을 통하여 AIC와 BIC의 값은 우도값이 높아질수록 작아지며, 모형을 추정하는 모수의 수가 많을수록 AIC와 BIC의 값이 높아짐을 알 수 있다. 그러므로 각 AIC와 BIC는 값이 낮을수록(값이 더욱 '음(-)'일수록), 모형의 적절성이 높아진다는 것을 의미한다.

두 번째 요소인 각 항들의 계수의 z-value에 관해서는, 모형추정을 통해 SAR항, SMA항, AR항, MA항들의 계수를 추정해낼 수 있다. 그리고 추정된 각 항들의 계수에 대한 유의성 검정을 통하여 각 항들의 계수가 적절한지 여부를 확인함으로써, 전체 모형의 적절성 여부를 판단할 수 있다.

앞선 안정성 검정을 통해 1계 차분 및 4계 계절차분화된 시계열 자료가 안정적인 것을 확인하였고, 이를 확인하는 과정에서 최적의 lag값을 도출하였다. 단위근의 존재 여부를 확인하기 위해 시행한 ADF 검정과정을 통하여 최적의 lag값이 '7'임을 알아냈고, 계절 단위근의 존재 여부를 확인하기 위해 시행한 HEGY 검정과정을 통하여 계절파트의 최적의 lag값이 '0'임을 알아냈다.

앞서 언급한 바와 같이 SARIMA모형은 ARIMA(p,d,q)모형에 계절파트인 (P,D,Q)가 포함된 모형이다. 즉 ADF 검정과정으로 도출한 최적의 lag값을 바탕으로 'p'와 'q'의 값을 결정할 수 있으며, HEGY 검정과정으로 도출한 최적의 lag값을 통해 'P'와 'Q'의 값을 결정할 수 있다. 앞선 과정을 통해 'p'와 'q'의 추정값은 '7'이며 'P'와 'Q'의 추정값은 '0'이라는 결론을 얻을 수 있다. 이러한 과정을 통해 최적의 모형으로 우선적으로 판단되는 모형은 SARIMA(7,1,7)(0,1,0)4 이다. 이는 우선적으로 판단되어진 모형이며, 다음 모형이 반드시 최적의 모형이 된다고 할 수는 없다.

그렇기에 앞서 언급한 AIC 및 BIC의 값과 계수에 대한 z-value를 토대로 우선적으로 추정되어진 모형이 적절하게 추정되었는지 확인하여야 한다.

앞서 모형을 추정하는 과정에서 'p'와 'q'의 최적의 값이 '7'이고, 'P'와 'Q'의 최적의 값이 '0'이라는 것을 확인하였고 이에 모델의 간결성을 추구하기 위하여 각 항들의 값의 범위를 Equation(5)과 같이 정의하였다.

$$p: 0 \leq p \leq 7$$

$$q: 0 \leq q \leq 7$$

$$P: 0 \leq P \leq 0$$

$$Q: 0 \leq Q \leq 0$$

(5)

여기서 간결성을 추구하는 이유는 최적의 lag값보다 더 높은 차수를 통해 높은 예측력을 가지게 된다 하더라도, AR항, MA항, SAR항, SMA항들의 차수가 너무 높게 되면, 실제로 이를 활용하는데 있어서 실용성이 떨어지기 때문이다. 즉, 모델이 간결할수록 더욱 실용적인 모델이 될 수 있음을 의미한다.

Table 6은 Equation(5)에서 정한 조건에 따르면 총 64개의 모형이 선택 가능하다. 이 중 앞서 추정된 Model(1)과 63개의 모형 중 AIC와 BIC의 값 및 각 계수들의 z-value 결과에 의해 최적으로 선택된 Model(2)를 나타낸 것이다. 앞서 우선적으로 추정된 Model(1)의 경우, z-value에 따라 AR항들과 MA항들 대부분이 무의미한 것으로 나타났다. 반면 Model(2)의 경우, MA항들은 유의미한 것으로 나타났고 AR항도 유의미한 것으로 나타났다. 또한 AIC와 BIC 값도 큰 차이를 나타내지 않았다. 또한 간결성의 원칙을 보게되면, Model(1)은 AR항의 차수가 7인 반면에, Model(2)는 AR항의 차수가 1이다. 즉, Model(2)가 Model(1)에 비해 간결성이 뛰어나며 실용성도 높을 것이라고 판단된다. 이와 같은 이유로 Model(2)가 앞서 추정된 Model(1)과 비교하여 보다 적합한 모형이라고 판단되어진다.

Table 6. Estimated model

		Model(1)		Model(2)	
		SARIMA(7,1,7)(0,1,0)4			
		coef (Std. Err.)	z-value (p> z)	coef (Std. Err.)	z-value (p> z)
AR	AR(1)	1.246839 (0.7583356)	1.64 (0.100)	-0.4329935 (0.1739031)	-2.49 (0.013)
	AR(2)	0.4440717 (1.151816)	0.39 (0.700)		
	AR(3)	-1.014181 (0.5538603)	-1.83 (0.067)		
	AR(4)	-0.2045279 (0.8415566)	-0.24 (0.808)		
	AR(5)	0.6124432 (0.4091371)	1.50 (0.134)		
	AR(6)	-0.1307822 (0.8946647)	-0.21 (0.831)		
	AR(7)	-0.269257 (0.602025)	-0.04 (0.964)		
MA	MA(1)	-1.058774 (164.166)	-0.01 (0.995)	0.8887458 (0.3067116)	2.90 (0.004)
	MA(2)	-1.031704 (4.821421)	-0.21 (0.831)	1.092804 (0.3079684)	3.55 (0.000)
	MA(3)	1.397871 (151.4704)	0.01 (0.993)	1.204056 (0.2563751)	4.70 (0.000)
	MA(4)	-0.8041588 (179.8907)	-0.00 (0.996)	-	-
	MA(5)	0.0426991 (·)	·	-	-
	MA(6)	0.8634631 (130.5797)	0.01 (0.995)	-	-
	MA(7)	-0.7936918 (0.5250291)	-1.51 (0.131)	-	-
SAR		-	-	-	-
SMA		-	-	-	-
Constantterm		-186.4957 (352.8503)	-0.61 (0.544)	87.65321 (2007.538)	0.04 (0.965)
Number of obs		43		43	
Likelihood		-418.933		-428.5832	
AIC		865.8659		867.1664	
BIC		890.5227		875.9724	

마찬가지로 다른 선택 가능한 모델들과 비교해본 결과, Model(2)가 가장 적합한 모형이라고 판단되었다. 따라서 본 연구에서는 인천국제공항에서 출발하여 미주노선을 통해 도착한 분기별 항공화물량의 수요를 예측하기 위해서 추정된 모형으로 SARIMA(1,1,3)(0,1,0)4 모형이 최적의 모델로 나타났다.

최적의 모델로 추정된 SARIMA(1,1,3)(0,1,0)4 모형은 Equation(6)와 같이 나타낼 수 있다.

$$\begin{aligned}
 &\Delta Y_t - \Delta Y_{t-4} \\
 &= -0.4329935(\Delta Y_{t-1} - \Delta Y_{t-5}) + \epsilon_t \\
 &\quad - 0.8887458\epsilon_{t-1} - 1.092804\epsilon_{t-2} \\
 &\quad - 1.204056\epsilon_{t-3} + 87.65321
 \end{aligned} \tag{6}$$

$$\begin{aligned}
 &\Delta Y_t = Y_t - Y_{t-1} \\
 &\epsilon_t = N(0, \sigma^2) \text{을 따르는 오차항, 백색잡음}
 \end{aligned}$$

Equation(6)를 Y_t 에 관해 나타내면, Equation(7)로 나타낼 수 있다.

$$\begin{aligned}
 Y_t = & 0.5670065 Y_{t-1} + 0.4329935 Y_{t-2} \\
 & + Y_{t-4} - 0.5670065 Y_{t-5} \\
 & - 0.4329935 Y_{t-5} + \epsilon_t - 0.8887458\epsilon_{t-1} \\
 & - 1.092804\epsilon_{t-2} - 1.204056\epsilon_{t-3} \\
 & + 87.6531
 \end{aligned} \tag{7}$$

Equation(7)에 나타난 바와 같이, SARIMA(1,1,3)(0,1,0)₄에서 물동량 Y_t 는 Y_{t-1} , Y_{t-2} , Y_{t-4} , Y_{t-5} , Y_{t-6} 과 ϵ_{t-1} , ϵ_{t-2} , ϵ_{t-3} 의 값에 의해 결정된다. 즉 시계열 자료가 분기별 자료인 것을 감안하였을 때 예측시점을 기준으로 6분기 전까지의 화물 물동량이 예측시점의 물동량을 결정짓는 것으로 추정된다고 할 수 있다. 앞서 언급한 바와 같이 본 연구에서 활용한 시계열자료는 계절성을 띄고 있다. 이러한 시계열 자료를 활용한 수요 예측은 최근 과거값보다 전년도 동분기의 값이 더 높은 설명력을 가지게 되는데, SARIMA(1,1,3)(0,1,0)₄의 경우에도 전년도 동분기 값인 Y_{t-4} 이 추정에 활용되어짐을 알 수 있다. 또한, SARIMA(1,1,3)(0,1,0)₄모형의 경우, Y_{t-4} 의 계수가 '1'것을 알 수 있다. 즉, SARIMA(1,1,3)(0,1,0)₄으로 예측한 물동량은 전년도 동 분기 값인 Y_{t-4} 가 유지되면서, Y_{t-1} , Y_{t-2} , Y_{t-5} , Y_{t-6} , ϵ_{t-1} , ϵ_{t-2} , ϵ_{t-3} 의 값들에 의하여 변동된다고 추론할 수 있다.

4. 모형 검증

1) 잔차항의 자기상관여부 진단

앞선 단계를 통하여 SARIMA(1,1,3)(0,1,0)₄ 모형을 추정하였다. 검진의 단계에서는 추정된 SARIMA(1,1,3)(0,1,0)₄ 모형이 자기상관(Autocorrelation)이 존재하는지 여부를 검사하고 자기상관이 존재한다면, 다시 추정의 단계에서부터 모형을 새로이 추정하게 되고, 자기상관이 나타나지 않는다면 추정된 모형이 유의미하다고 판단할 수 있다.

Table 7은 SARIMA(1,1,3)(0,1,0)₄의 잔차항(Residuals)를 Q-test를 통하여 자기상관의 여부를 나타낸 표이다. 표에 따르면 시차에 따른 Q-value가 5% 유의수준에서 잔차항간의 자기상관이 존재한다는 귀무가설을 기각할 수 있다. 즉 SARIMA(1,1,3)(0,1,0)₄ 모형은 적절한 모형이라고 판단할 수 있다.

Table 7. Q-value of SARIMA(1,1,3)(0,1,0)₄

Lag	Q	Prob>Q
1	0.535	0.8171
2	0.3426	0.8426
3	0.6453	0.8860
4	6.1861	0.1856
5	6.3739	0.2715
6	7.0108	0.3198
7	7.4367	0.3849
8	7.5579	0.4778
9	7.6498	0.5698
10	7.6557	0.6624
11	7.8048	0.7307
12	10.3943	0.5814
13	10.4470	0.6570
14	10.6703	0.7117
15	11.5912	0.7097

2) ARIMA 모형과의 비교

추정된 SARIMA 모형을 통한 미주노선에 대한 항공화물의 수요예측이 설득력을 가지기 위해서는 방법론의 적절성 여부를 파악하는 것이 중요하다. 그렇기에 우리는 시계열 분석 기법 중 가장 널리 사용되어지는 방법론인

ARIMA 모형과의 비교를 통하여 SARIMA 모형의 적절성 여부를 파악해 보았다. 적절성 여부를 파악하기 위한 구체적 비교방법은 SARIMA 모형 추정에 사용되어진 2003년도 1사분기부터 2014년도 4사분기까지 총 48개의 인천국제공항에서 출발한 분기별 미주노선 항공화물 물동량을 동일하게 활용하여 ARIMA 모형을 추정한 후, 추정된 ARIMA 모형과 SARIMA 모형을 이용한 예측치를 실제 항공화물 물동량과의 비교를 통하여 SARIMA 모형이 적합한지 그 여부를 판단하였다. Table 8은 앞선 SARIMA 모형 추정 시 사용한 항공 화물 분기별 물동량을 동일하게 사용하여 추정된 ARIMA(2,1,0) 모형을 나타낸 것이다. Table 8을 살펴보면, 본 연구에서 최적으로 모형으로 판단된 SARIMA(1,1,3)(0,1,0)₄ 모형과 비교하여 AIC와 BIC 값 모두 높게 나타났다. 또한 ARIMA(2,1,0) 모형의 경우 AR항들의 z-value가 상대적으로 높게 나타났기에 이 모형이 유의미하다고 판단하기 어렵다. ARIMA 모형 추정에 유의미한 모형이 나오지 않는 이유는 본 연구에서 사용된 시계열 자료가 주기적인 특성을 가지기 때문이라고 판단되어지며, 그렇기에 원 시계열 자료가 자체적으로 계절주기를 가지고 있다는 것을 통해 유의미한 모형 추정이 불가능한 ARIMA(2,1,0) 모형보다는 계절주기를 감안해서 수요예측이 가능한 SARIMA(1,1,3)(0,1,0)₄ 모형이 더 적합하다고 판단 가능하다.

Table 9은 SARIMA(1,1,3)(0,1,0)₄ 모형과 ARIMA(2,1,0)모형을 활용하여 예측한 2015년 1사분기부터 2016년 2사분기까지의 화물물동량의 예측치를 실측치와 비교하여 각 분기마다 오차와 예측정확도를 나타낸 것이다. 2015년 1사분기부터 2016년 2사분기까지 총 6분기동안의 실제 물동량과 예측치를 비교하여보면, 2015년도 3사분기, 2015년도 4사분기, 2016년도 1사분기에서 SARIMA(1,1,3)(0,1,0)₄ 모형의 예측정확도가 더 높게 나타났다. 또한, ARIMA(2,1,0) 모형의 가장 낮은 예측정확도는 81.55%이며 가장 높은 예측정확도가 99.3%인 반면에 SARIMA(1,1,3)(0,1,0)₄ 모형의 가장 낮은 예측정확도는 84.39%이며 가장 높은 예측정확도는 99.92%이다. 또한 2015년도

Table 8. ARIMA(2,1,0) model

		ARIMA ARIMA(2,1,0)	
		coef (Std. Err.)	z-value (p) z
AR	AR(1)	-0.240726 (0.1878167)	-1.28 (0.200)
	AR(2)	0.0864448 (0.1825435)	0.47 (0.636)
Constant term		51.23255 (935.6964)	0.05 (0.956)
Number of obs		47	
Likelihood		-479.4807	
AIC		966.9613	
BIC		974.3619	

Table 9. Comparison between SARIMA(1,1,3)(0,1,0)₄ and ARIMA(2,1,0)

	Actual Volume	SARIMA(1,1,3)(0,1,0) ₄			ARIMA(2,1,0)		
		Forecasting	Accuracy	Error	Forecasting	Accuracy	Error
2015 Q1	70196	59240.93	84.39%	10955.07	63930.62	91.07%	6265.38
2015 Q2	66283	63148.83	95.27%	3134.17	64197.36	96.85%	2085.64
2015 Q3	61407	61358.07	99.92%	48.93	64153.78	95.53%	2746.78
2015 Q4	66562	64603.85	97.06%	1958.15	64246.46	96.52%	2315.54
2016 Q1	54267	57381.17	94.26%	3114.17	64279.52	81.55%	10012.52
2016 Q2	64795	61317.82	94.63%	3477.18	64338.71	99.30%	456.29
MAPE (2015 1Q - 2016 2Q)		5.743572%			6.529774%		



Figure 4. Comparison of forecasting between SARIMA(1,1,3)(0,1,0)₄ and ARIMA(2,1,0)

1사분기의 예측정확도가 상대적으로 낮은 이유는 시계열자료의 2012년도부터 2014년도까지 1사분기마다 화물물동량이 급격히 떨어지는 주기적인 특성이 감안되어졌기에 SARIMA(1,1,3)(0,1,0)₄ 모델의 2016년도 1사분기 예측치도 급격히 떨어진다고 예측되어졌다고 판단된다. 이를 제외하고는 2015년도 1사분기부터 2016년도 2사분기까지의 MAPE(Mean absolute percentage error)를 비교해보아도 SARIMA(1,1,3)(0,1,0)₄ 모형의 MAPE가 더 낮은 값을 나타냈다. 이를 통하여 SARIMA 방법론으로 추정된 모형이 ARIMA 방법론으로 추정된 모형보다 예측의 정확도가 높다는 것을 알 수 있다.

Figure 4는 SARIMA(1,1,3)(0,1,0)₄ 모형의 예측치와 ARIMA(2,1,0) 모형의 예측치를 실제 항공 화물 물동량과 비교한 그래프이다. Figure 4를 분석해보면, ARIMA모형의 예측치는 실제 물동량에 1시점씩 뒤처지지만 값에 있어서는 SARIMA모형보다 더 정확하게 나타나있다. 하지만, ARIMA모형으로 나타난 예측치의 시점을 따로 조정하는 것이 아니며, 실제로 구하고자 하는 2015년도 1사분기부터 2016년도 2사분기까지의 예측치에서 SARIMA(1,1,3)(0,1,0)₄ 모형의 예측치가 ARIMA(2,1,0) 모형을 활용한 예측치보다 실제 물동량에 더 근접하면서 높은 예측정확도를 보이고 있음을 알 수 있다. 이렇듯 상대적으로 SARIMA 모형이 ARIMA 모형보다 상대적으로 정확도가 높은 이유는 다음과 같이 예상된다. ARIMA 모형이 자체가 갖는 모형이 항공화물의 계절성을 감안하지 못한다는 한계 때문이다. 서론에서 언급하였듯이, ARIMA 모형은 예측을 하는 시점으로부터 가까운 과거 값들에 높은 설명력을 부여하여 예측을 하기에, 전년도 동분기와 같은 주기적인 특성을 가진 값에는 높은 설명력을 부여하지 못하기 때문이다. 그렇기에 항공화물처럼 계절성을 가지는 시계열 자료의 경우에는 상대적으로 이러한 계절성 및 주기적인 특성을 고려 가능한 SARIMA 모형이 ARIMA 모형에 비해 예측정확도가 높을 수밖에 없다. 이러한 ARIMA 모형의 한계점으로 인해 ARIMA(2,1,0) 모형의 예측 정확도가 SARIMA(1,1,3)(0,1,0)₄ 모형에 비해 떨어지게 나타난 것으로 예상된다.

5. 시사점

본 연구는 2003년 1사분기부터 2014년 4사분기 까지 인천국제공항에서 출발하여 미주에 도착한 항공화물의 시계열 자료를 바탕으로 SARIMA 모형을 활용, 수요예측 모형을 구축하였다. 앞서 서론에 언급하였듯이 현재 국내에서는 교통이나 화물수요를 예측하는 연구에 있어서 SARIMA 모형을 활용한 연구는 많지 않다. 또한 특정 공항의 여

객 수요 및 화물물동량이 아닌 특정 노선으로 향하는 화물 물동량의 수요를 예측하거나 이와 관련된 연구 역시 제한적이었다. 그렇기에 본 연구는 SARIMA 모형을 활용하여 인천국제공항에서 출발하여 미주노선을 통해 미주지역에 도착하는 항공화물의 수요를 예측을 연구하였다는 점에서 상당히 큰 의의가 있다고 생각된다.

본 연구를 통해 계절성을 보이는 항공화물 물동량과 같은 시계열 자료에 있어서는 기존 연구에서 자주 활용되어진 ARIMA 모형에 비하여 SARIMA 모형이 상대적으로 예측정확도가 높다는 것을 알 수 있었다. 하지만 본 연구의 결과만을 가지고 ARIMA 모형보다 SARIMA 모형이 항상 우월한 방법론이라고 판단할 수는 없다. 오히려 계절성을 띄지 않는 시계열 자료의 경우에는 ARIMA 모형이 추정과정이 더욱 간편하기에 실용성이나 예측면에서 더 나을 수도 있기 때문이다. 하지만, 본 연구를 통해 계절성을 띄는 시계열 자료가 대해서는 SARIMA 모형과 ARIMA 모형의 비교를 통하여 SARIMA 모형의 우수함을 알 수 있었으며, 향후 수요예측 연구에 있어서 기존에 자주 활용되어진 ARIMA 모형과 함께 SARIMA 모형이 수요 예측에 있어서 유용한 방법임을 확인하였기에, 향후 수요예측 연구에서 SARIMA 모형이 더욱 자주 활용되어 질 것으로 기대된다.

결론

항공운송은 해상운송과 함께 주요 운송수단으로써 국가 간의 교역에서 매우 큰 역할을 수행하고 있으며, 특히 항공운송은 신속성이라는 장점으로 인해, 고부가가치 제품을 운송하면서 교역량에 비해 높은 수준의 교역액 비중을 차지면서 그 중요성이 계속해서 높아지고 있다.

그동안의 항공수요에 대한 예측은 주로 화물이 아닌 여객에 초점을 맞춰 이루어져 왔다. 항공 화물수요에 대한 연구는 대부분 공항을 기준으로 하는 연구였으며, 노선을 기준으로 하는 연구는 찾아보기 힘들다. 물론 이러한 연구들은 공항의 운영방침에는 큰 도움이 된다. 하지만, 항공사의 입장에서 실제 노선에 대한 수요 예측이 가능하다면 항공사는 예측된 수요에 따라 공급량을 조절하면서 공급과잉과 공급부족으로 인해 일어나는 손실을 최소화 시킬 수 있다. 그렇기에 기존 연구들과 함께 특정 노선을 통하는 항공화물에 대한 수요 예측 연구 또한 관심을 가지고 비중있게 다루어져야 한다.

본 연구에서는 SARIMA 모형을 이용하여, 인천국제공항에서 출발하여 미주노선을 통해 미주에 도착하는 분기별 항공화물량을 바탕으로 수요를 예측해보았으며, 이를 기존의 ARIMA 모형의 결과와 비교함으로써 본 연구에서 제시하고 있는 SARIMA 방법론의 예측정확도를 검증해 보았다. 모형의 추정에 활용되어진 시계열 자료는 2003년 1사분기부터 2014년 4사분기까지이며 이를 활용하여 2015년 1사분기부터 2016년 2사분기까지의 예측치와 실적치를 비교하여 보았을 때, 본 논문에서 제시한 SARIMA모형의 예측치가 ARIMA 모형의 예측치보다 MAPE 값을 통해 예측정확도가 더 높았다. 이는 SARIMA 모형이 계절성 및 주기적인 특성을 감안할 수 있다는 특성을 가졌기에 나타난 것으로 판단되며, 항공화물과 같이 계절성을 가진 시계열 자료의 경우, 기존의 대표적 시계열 분석기법인 ARIMA 모형보다 SARIMA 모형이 더 적합하다고 할 수 있다.

본 연구는 공항을 중심으로 실행되어진 기존 연구와는 차별성을 두어 특정 노선을 중심으로 한 수요예측 연구이나, 하나의 노선이 아닌 미주로 통하는 모든 노선을 통합한 물동량을 활용하여 연구하였기에, 이에 따른 한계점을 지니고 있다. 그러므로 향후 특정 세부 노선간의 항공화물 수요를 예측하는 연구에 있어서 SARIMA 모형을 활용하는 연구가 계속해서 이루어져야 할 것이다. 또한 본 연구는 미주노선에 중요성을 강조하여 미주노선에 대한 수요예측을 시행하고 SARIMA 모형이 높은 예측정확도를 가짐을 보였다. 하지만 미주 노선 예측만으로 모든 항공화물에서 SARIMA 모형이 우수하다고 판단하기는 어렵다. 따라서 미주노선에 관한 항공화물이 아닌 전체 항공화물시장이 계절성을 가지고 SARIMA모형이 높은 예측 정확도를 가짐을 검증하기 위해선 추가적인 다양한 노선의 수요 예측 결과를 비교하여 정확도를 검증하는 연구가 필요하다고 보여진다.

본 연구는 계절성을 주목하여 SARIMA모형을 활용하였지만 본 연구에서는 항공화물을 모두 합쳐서 분석하였다. 향후에는 항공화물의 성격이나 품목별로 분류하여 SARIMA모형을 활용한 수요예측에 대한 연구가 진행된다.

면, 더욱 항공화물에 계절성을 부각될 것이며, SARIMA 모형의 예측정확성을 입증하는데 많은 도움을 줄 수 있다고 본다.

본 연구에서 활용되어진 SARIMA 모형인 경우 다른 시계열 분석 모형들과 마찬가지로 과거 물동량 수치를 바탕으로 미래 물동량을 예측하는 것이므로 다른 GDP, 유가, 환율 등의 경제요인들의 변화에 따른 미래 물동량에 대한 구체적인 변화를 설명할 수 없다는 단점을 가지고 있다. 이 때문에 장기수요를 예측하는데 있어서는 회귀분석에 비해 예측정확도가 떨어진다는 단점이 존재한다. 따라서 항공화물의 수요를 예측하는 데 있어서, SARIMA 모형이 가장 적합한 모형이라고 단정할 수는 없다. 연구의 목적에 따라, 다양한 변수를 고려한 분석모형과 같은 다른 방법론이 더 적합할 수 있다.

그럼에도 불구하고 본 연구는 기존 항공화물 관련 연구와는 차별성을 둔 노선중심의 연구라는데 있어서 그 의의를 두고 있으며, 계절성이 뚜렷한 항공화물의 특성을 감안하여 기존 ARIMA 모형이 아닌 SARIMA 모형을 활용하였다는 점에서 또한 연구가치가 있다고 판단된다. 다른 경제요인들의 변화에 따른 구체적 변화를 설명할 수는 없지만 본 연구를 통하여 SARIMA 모형이 상대적으로 단기간의 수요예측에 있어서 높은 정확도를 보인다는 점에서 연구가치가 있다고 판단되어지며, 향후 본 연구를 통해 항공화물을 포함한 계절성을 가진 다른 교통 수요를 예측하는데 있어서 SARIMA 모형이 보다 유용하고 자주 활용되어지기를 기대한다.

ACKNOWLEDGEMENT

This work was supported by Inha University.

REFERENCES

- Ariuntuya T. I. (2013), A study on the Airline Demand Forecasting ICN-ULN International Routes, Thesis of Korea Aerospace University(Master).
- Arnaldo S. R. (2013), Forecasting air passengers at São Paulo International Airport Using a Mixture of Local Experts Model, *Journal of Air Transport Management*, 26, 35-39.
- Box G. E. P., Jenkins G. M., Reinsel G. C. (1976), *Time Series Analysis, Forecasting and Control*, San Francisco, Holden Day.
- Chang Y. W., Liao M. Y. (2010), A Comparison of Time Series Models for Forecasting Outbound Air Travel Demand, *Journal of Aeronautics, Astronautics and Aviation, Series A*, 42(2), 73-78.
- Chen C. F., Chang Y. H., Chang Y. W. (2009), Seasonal ARIMA Forecasting of Inbound Air Travel Arrivals to Taiwan, *Transportmetrica*, 5(2), 125-140.
- Choi J. H., Yoo K. W., Lee S. Y. (2015), Demand Estimation Methodology for a New Air Route, *J. Korean Soc. Transp.*, 33(2), Korean Society of Transportation, 145-153.
- Chou T. Y., Liang G. S., Han T. C. (2011), Application of Fuzzy Regression on Air Cargo Volume Forecast, *Quality and Quantity* Volume 47.
- Dickey D. A., Fuller W. A. (1979), Distribution of the Estimators for Autoregressive Time Series With a Unitroot, *Journal of the American Statistical Association*, 74(366), 427-431.
- Erma S. Y., Chou S. Y., Chen C. H. (2010), Air Passenger Demand Forecasting and Passenger Terminal Capacity Expansion: A System Dynamics Framework, *Expert Systems With Applications*, 37(3), 2324-2339.

- Kim S. M., Bae H. J., Park Y. H. (2015), Forecasting the LCC Air Passenger Demand in Incheon National Airport with Seasonal ARIMA Model, *Journal of Aviation Management Society of Korea*, 2015(02015), 45-45.
- Kim Y. J. (2014), Study on Low Cost Carrier Demand Forecasting Using Seasonal ARIMA Model: Domestic Jeju Route Inland Route Forecast Comparison Analysis, *Journal of Tourism Studies*, 26(1), 3-25.
- Min K. C., Jun Y. I., Ha H. K. (2013), Forecasting the Air Cargo Demand With Seasonal ARIMA Model: Focusing on ICN to EU Route, *J. Korean Soc. Transp.*, 31(3), Korean Society of Transportation, 3-18.
- Pablo C. M., Gema C. G., Vicente I. L., Rubén S. G. (2004), An Air Traffic Demand Forecasting Model Applied To Santander Airport, *International Journal of Transport Economics*, 31(2), 197-208.
- Yoon J. S., Huh N. K., Kim S., Hur H. Y. (2010), A Study on International Passenger and Freight Forecasting Using the Seasonal Multivariate Time Series Models, *Communications of the Korean Statistical Society*, 17(3), 473-381.