

SARIMA 시계열 모형을 이용한 환동해 물동량 예측*

송민주
영남대학교 무역학과 박사과정

이희용
영남대학교 무역학과 부교수

Forecasting the East Sea Rim Container Volume by SARIMA Time Series Model

Min-Ju Song^a, Hee-Yong Lee^b

^aPh.D Student, Department of International Economics and Business, Yeungnam University, South Korea.

^bAssociate Professor, Department of International Economics and Business, Yeungnam University, South Korea.

Received 20 September 2020, Revised 26 October 2020, Accepted 29 October 2020

Abstract

The purpose of this paper was to analyze the trend of container volume using the Seasonal Autoregressive Intergrated Moving Average (SARIMA) model. To this end, this paper used monthly time-series data of the East Sea Rim from 2001 to 2019.

As a result, the SARIMA(2,1,1)12 model was identified as the most suitable model, and the superiority of the SARIMA model was demonstrated by comparative analysis with the ARIMA model. In addition, to confirmed forecasting accuracy of SARIMA model, this paper compares the volume of predict container to the actual volume.

According to the forecast for 24 months from 2020 to 2021, the volume of containaer increased from 60,100,000Ton in 2020 to 64,900,000Ton in 2021

Keywords: Time Series analysis, SARIMA, ARIMA, Seasonal ARIMA, Container Volume Forecasting, East Sea Rim

JEL Classifications: C22, C53, F10, F17, F47

* This work was supported by the 2019 Yeungnam University Research Grant

^a First Author, E-mail: tomboysmj@ynu.ac.kr

^b Corresponding Author, E-mail: ilugit@ynu.ac.kr

© 2020 The Korea Trade Research Institute. All rights reserved.

I. 서론

환동해권(環東海權)은 동해를 중심으로 접하고 있는 중국 동북지역, 러시아 극동지역, 일본 서해지역을 하나의 권역으로 묶은 지역으로 동북아시아 국가들의 지역별 개발정책을 통해 환동해 경제권의 중심으로 부상하고 있다. 실제 문제인 정부의 신북방정책, 중국 일대일로정책, 러시아 신동방정책 등의 지역개발정책으로 중국 동북 3성과 러시아 북방지역의 물류네트워크 확보에 대한 필요성이 증대되고 있으며, 환동해권 교류·협력에서 가장 큰 영향을 미치는 항만물류분야의 중요성이 커지고 있는 상황이다. 또한 국가경제발전의 핵심 요인 중 하나인 국제무역의 활성화로 물류활동의 역할이 커지면서 미시적 수준 정확한 물동량 예측은 계획 수립의 핵심이 되었다 (Dragan and Kramberger (2014)). 항만의 경우, 국가 경제와 사회 전반에 영향을 미치는 공공재적 성격이 강한 사회간접자본시설 중 하나로 항만수요 및 물동량에 대한 예측은 필수적(Go Yong-Gi, et al (2008))이며, 이러한 물동량 예측은 서비스 품질, 계획, 시설 및 운영을 향상시키는 결정적 구성요소로 작용하기 때문이다 (Wei, Yu and Mu-Chen Chen (2012)).

항만분야의 수요예측은 시설 계획의 대규모 투자에 따른 예측이 주목적을 이루기 때문에 대부분 장기수요에 초점이 맞추어져 있었다. 항만시설은 많은 시간과 비용이 소모되는데 반해 상대적으로 그 효과는 길기 때문에 장기수요예측을 우선적으로 진행하는 것이 이상적이기 때문이다. 하지만 항만 운영계획에 직접적으로 영향을 미칠 수 있는 단기수요예측 또한 항만·물류정책의 운영 및 관리의 효율성을 높일 수 있기 때문에 수요예측에서 중요한 요소로 작용하고 있다.

이에 본 연구에서는 정치적, 경제적, 문화적, 사회적 이질성으로 어려웠던 환동해권 국가 간 협력과 세부 발전방안을 모색하고 향후 환동해 경제활성화를 위한 정책적 시사점 도출을 위한 수요예측을 진행하고자 한다. 이를 위해 환동해권 컨테이너 물동량에 대한 시계열 자료를 통해 수요분석을 실시하였으며, 물동량 예측에

서 가장 많이 사용되고 있는 시계열분석(Time Series Analysis)을 통해 물동량을 예측하였다. 시계열분석은 연속적으로 이어진 관찰치를 바탕으로 시간에 흐름에 따라 나타나는 시계열 패턴을 파악하여 미래 특정 시점의 값을 예측하는 방법으로 ARIMA모형(Autoregressive Intergrated Moving model, ARIMA)과 지수평활법(Exponential Smoothing Method, ESM)이 대표적이다 (Yoo et al (2014)).

앞서 언급한 내용과 같이 본 연구는 단기수요예측에 대한 중요성을 인식하고 단기 예측의 신뢰성이 높은 ARIMA모형 중 계절성과 주기성을 지닌 SARIMA(Seasonal Autoregressive Integrated Moving Average) 모델을 통해 환동해 물동량 추이를 분석하고자 한다. 연구의 범위는 2001년부터 2019년까지 환동해권 5개 항만인 부산항, 동해항, 속초항, 포항항, 울산항 물동량을 중심으로 분석하였으며 연구는 다음과 같이 구성된다. II장에서는 수요예측을 위한 기존 선행연구들을 살펴보고 III장에서 시계열모형에 대한 이론적 배경을 살펴본다. 이후 IV장에서 SARIMA모형 추정을 통해 적합한 모형을 설정한 다음 환동해 물동량 수요를 예측하고 마지막 V장에서 결론 및 시사점을 제시한다.

향후 본 연구의 결과는 환동해 물동량 현황 파악과 함께 경제 활성화를 위한 정책적 시사점 도출 및 향후 전략수립의 기초 자료로 활용될 수 있을 것이라 기대된다.

II. 선행연구

사회경제의 불확실성과 변화의 가속화로 수요예측에 대한 필요성이 증가하고 있는 상황에서 수요예측은 중요한 연구 분야로 인식되고 있다 (G. De Jong, H. Gunn and W. Walker, (2004)). 시계열자료를 통한 수요예측은 이미 많은 연구들에 의해 진행되고 있으며(Jang et al (1998)), 경제와 경영학을 중심으로 한 사회과학의 다양한 분야에서 활용되고 있다.

수요예측의 경우, 동일한 데이터를 활용하더라도 예측 방법에 따라 그 결과가 달라질 수 있

기 때문에 다양한 모형을 통해 예측력을 비교·분석하는 연구가 주를 이루고 있다 (Silva, Emmanuel Sirimal, and Hossein Hassani (2015)). 일반적인 수요예측 방법으로는 정성적기법과 정량적 기법 그리고 시스템 기법으로 구분할 수 있다. 정성적 기법은 데이터를 활용한 분석이 어려울 때 효과적인 방법으로 델파이기법, Judgemental Bootstrapping, 컨조인트 분석방법이 있으며, 정량적 기법으로는 상관관계를 분석하는 회귀분석방법과 시계열자료를 이용하는 시계열분석방법, 마지막으로 시스템 다이나믹스, 인공지능망 등의 시스템적 기법으로 구분된다. 기존 연구를 통해 확인된 항만, 물류, 교통분야의 주요 예측방법으로는 정량적 기법으로 과거 관측치에 대한 상호관계 모형을 구축하여 수요를 예측하는 시계열 분석방법을 주로 활용하고 있었다.

시계열 예측기법 중 주기성과 계절성을 반영한 SARIMA모형의 경우, 기존의 ARIMA모형과 SARIM모형의 비교분석을 통해 SARIMA모형의 우수성을 검증하고 있다.

Yi Ghae-Deug (2013)는 1992년부터 2011년까지 부산항의 월별 물동량 통계자료를 사용하여 물동량 수요를 추정하여 예측하였으며, Suh Bo-Hyoun, Tae-Woong YANG and Hun-Koo HA(2017)는 2003년 1사분기부터 2016년 2사분기까지 인천국제공항 미주 항공 화물 시계열 자료를 통하여 장래물동량을 예측하였다.

그 외 Yoon, Han-Young and Sung-Sik Park (2019)은 외부 충격이 특정 시계열에 미치는 영향을 추정하는 모형인 ARIMA-Intervention(개입모형)을 통해 한국방문 자제를 요청한 한한령을 개입변수로 설정하여 인천국제공항의 식음료 매출액을 예측하였으며, Jeong, Su-hyun, and Chang-Hoon Shin (2016) 또한 아시아, 글로벌 금융위기를 영향을 분석하기 위해 1990년 1월부터 2011년 12월까지 월별자료를 활용한 ARIMA-Intervention(개입모형)을 적용하여 수요를 추정·분석한 것을 확인할 수 있었다.

하지만 앞에서 언급한 내용과 같이 수요를 예측함에 있어 어떠한 모형이 적합한지에 대한 정확한 기준이 모호하기 때문에 다양한 모형을

함께 비교·분석하여 예측의 정확도를 검증하는 연구 또한 증가하는 추세를 보이고 있다. Yoo, Sang-Lok, et al (2014)는 1996년부터 2013년까지 인천항 월별 관측값을 대상으로 시계열 분석(ARIMA, ESM)과 ANN 모형, ESM-ANN 총 3가지 예측 모델을 구축하여 장래 해상 교통량을 예측하였으며, Kim, Chang-beom (2015)은 1994년 1월부터 2014년 12월까지 해상 수출입물동량을 이용하여 외환위기와 글로벌 금융위기 기간을 개입변수 적용시점으로 설정하여 2019년까지의 예측결과를 도출하였다. 이 외에도 항만 물동량 예측력 제고를 위해 1991년~2006년까지 부산항과 인천항의 월별자료를 이용하여 ARIMA모형과 인공지능망(Artificial Neural Network, ANN), 그리고 이를 결합한 하이브리드모형을 활용해 다른 모형들과의 예측성과를 비교(Shin Chang-Hoon and Su-Hyun Jeong (2011)하는 등 예측력을 높이기 위한 다양한 연구방법들이 함께 활용되고 있다.

기존의 선행연구들의 경우, 거시적 관점의 장기수요예측에 대한 연구가 주를 이루고 있으나 항만 간 경쟁심화, 항만기능의 고도화 등으로 항만 운영효율성 제고를 위한 단기수요예측 또한 비중 있게 다루어져야 할 필요가 있다 (Min, Kyung-Chang and Hun-Koo Ha (2014)). 항만의 물동량 예측은 물류회사, 운송 네트워크, 항만 등에서 중요한 부분으로 인식되고 있으며, 이러한 예측 결과는 향후 경쟁력 제고 및 단·장기 전략수립을 가능하게 해주기 때문이다 (Farhan, J. and G. P. Ong (2018)).

이에 본 연구에서는 환동해권 내 항만의 월별 데이터(2001년~2019년)를 활용하여 물동량 분석을 진행하였으며, 시계열분석 방법 중에서도 계절성 및 주기성을 반영한 SARIMA(Seasonal Autoregressive Integrated Moving Average) 모형을 활용하여 물동량을 예측하고자 하였다. 또한 추정된 예측치를 실제 물동량 수치와 비교·분석하여 예측의 정확도를 검증하였으며 이를 통해 향후 물동량 예측에서 SARIMA모형의 활용가능성을 확인하고자 한다.

Table 1. Port's Cargo Volume

(Unit: Ton)

	Total	The East Sea Rim
2001	610,910,071	308,681,303
2002	635,545,272	323,306,176
2003	667,608,033	345,389,644
2004	733,377,114	391,511,117
2005	754,936,409	396,689,780
2006	809,829,604	415,213,806
2007	862,523,478	440,390,344
2008	894,693,317	446,555,314
2009	848,298,644	421,884,895
2010	966,193,342	468,609,149
2011	1,069,565,588	526,038,830
2012	1,108,538,270	542,105,482
2013	1,123,205,055	550,662,457
2014	1,184,641,194	579,257,178
2015	1,216,781,726	585,632,527
2016	1,242,597,061	592,292,648
2017	1,312,489,474	628,160,688
2018	1,405,925,305	700,473,439
2019	1,429,149,574	707,410,015

Source: Port-Mis

Ⅲ. 분석방법

1. 분석자료

본 연구에서 사용한 시계열자료는 해양수산부 해운항만물류정보시스템(PORT-MIS)의 2001년 1월부터 2019년 12월까지 환동해권 내 항만인 부산항, 동해항, 속초항, 포항항, 울산항의 월별 물동량 데이터를 활용하였으며, SARIMA 모형을 통해 예측된 물동량의 검증에 위해 2020년 실제 관측값과 예측값의 비교·분석을 통해 예측력을 검증하였다.

우리나라 전체 컨테이너 물동량은 2001년 10,910,071톤을 시작으로 2019년 1,429,149,574톤으로 연평균 4.8% 증가하여 지속적인 성장세를 유지하고 있으며, 이 중 환동해권 물동량은

2019년 기준 전체 물동량의 49%를 차지하는 707,410,015톤으로 나타나고 있다.

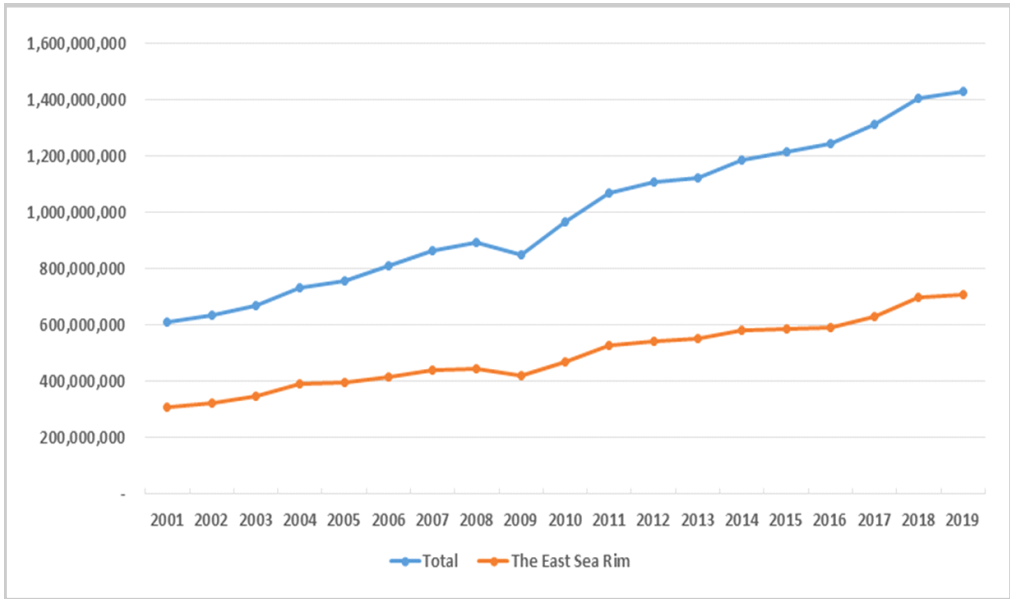
2. 시계열분석

1) ARIMA 모형

ARIMA모형은 모형의 식별(Model identification), 모수의 추정(parameter estimation), 모형의 진단(model diagnostic checking), 예측(forecasting) 과정으로 이루어진 시계열분석 기법 중 하나로 과거의 관측값과 오차를 활용하여 현재의 관측값 또는 미래를 예측하는 모형이다.

ARIMA모형을 통해 수요를 예측하기 위해서는 가장 먼저 시계열자료의 정상성(Stationarity)를 확보해야한다. 정상시계열자료가 아닐 경우

Fig. 1. Port Trade Volume



Source: Port-Mis

우, 시계열의 평균과 분산이 시점에 따라 변화하는 추세나 계절적 변동이 있다는 의미로 경향성, 계절성과 같은 불안정한 시계열 자료의 차분(Difference)과정을 통해 자료의 안정성을 확보하여야 한다.

ARIMA모형은 과거값을 이용하여 회귀식을 구성하는 자기회귀(AR : Autoregressive)와 모형 추정 잔차의 이동평균(MA : Moving Average)을 결합한 함수로 추정하며 시계열자료가 가지는 특성(추세, 계절성)에 따라 적합한 모형을 선택하여 예측값을 산출한다. 기회귀(AR : Autoregressive)는 현 관측치는 이전의 관측과 자기회귀 관계에 있다는 의미로 이전의 시계열 값이 다음의 시계열 값에 영향을 주는 경우를 나타내며 차수가 p인 AR(p) 모형을 식으로 표현하면 식(1)와 같다.

$$Y_t = \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + \dots + \Phi_p Y_{t-p} + \varepsilon_t \quad (1)$$

AR(p) 모형을 후방연산자를 사용하여 표현

하면 식 (2)와 같으며, $\Phi_p(L) = 1 - \Phi_1 L - \Phi_2 L^2 - \dots - \Phi_p L^p$ 는 특성함수이다.

$$(1 - \Phi_1 L - \Phi_2 L^2 - \dots - \Phi_p L^p) Y_t = \varepsilon_t \quad (2)$$

$$\Phi_p(L) Y_t = \varepsilon_t$$

특정시점에 발생한 사건 이후 시계열자료에 편향성을 초래하는 이동평균(Moving Average) 또한 AR(p) 모형으로 설명할 수 있지만 오차항(백색잡음)이 중요한 의미를 가진다는 점에서 AR(p)와 차이점을 가진다. MA(q) 모형의 경우 현 관측치는 현시점과 과거시점의 백색잡음의 선형결합으로 이루어지며 수식으로 나타내면 식(3)과 같다.

$$Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_p \varepsilon_{t-p} \quad (3)$$

후방연산자를 사용하면 식 (4)와 같이 표현할

수 있으며, $\theta_q(L) = 1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q$ 는 MA(q) 모형의 특성함수라 부른다.

$$Y_t = (1 - \theta_1 L - \dots - \theta_q L^q) \epsilon_t \quad (4)$$

$$= \theta_q(L) \epsilon_t$$

AR(p)과 MA(q) 모형을 결합한 ARIMA(p, d, q) 모형의 기본형태는 식(5)과 같이 나타낼 수 있다.

$$Y_t = \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + \dots + \Phi_p Y_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \dots + \theta_q \epsilon_{t-q} \quad (5)$$

$$\Phi_p(L) (1 - L)^d Y_t = \delta + \theta_q(L) \epsilon_t$$

$$\Phi_p(L) = 1 - \Phi_1 L - \Phi_2 L^2 - \dots - \Phi_p L^p$$

: AR 연산자

$$\theta_q(L) = 1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q$$

MA 연산자

Y_t : 시계열자료

t : 시간

ϵ_t : 오차항, 백색소음(white noise)

p : 자기회귀(Autoregressive, AR) 항 차수

q : 이동평균(Moving Average, MA)항 차수

d : 차분의 차수

2) SARIMA 모형

SARIMA 모형은 시계열자료의 계절성 및 주기성을 반영할 수 없는 ARIMA(Autoregressive Integrated Moving Average)모형의 한계점을 보완한 방법으로 SAR항과 SMA항을 추가한 SARIMA(p, d, q)(P, D, Q) 형태를 나타내고 있다. ARIMA 모형의 경우 계절성과 같은 주기적인 특성을 가진 시계열 자료의 특성을 반영하기 어렵다는 문제점이 존재하기 때문이다.

SARIMA 모형 또한 ARIMA 모형과 같이 시계열자료에 확률적 특성을 반영한 분석방법으로 관련 모형을 수식으로 나타내면 식 (6)와 같다.

$$\Phi_P(L^S) \phi_p(L) (1 - L)^d (1 - L^S)^D Y_t = \delta + \Theta_Q(L^S) \theta_q(L) \epsilon_t \quad (6)$$

Y_t : 시계열자료

t : 시간

ϵ_t : 오차항, 백색소음(white noise)

L : 후향연산자

p : 비계절자기회귀(Autoregressive, AR) 항 차수

q : 비계절이동평균(Moving Average, MA) 항 차수

d : 차분의 차수

p : 계절자기회귀(Seasonal Autoregressive, SAR) 항 차수

Q : 계절이동평균(Seasonal Moving Average, SMA) 항 차수

D : 계절차분의 차수

3) 모형추정 절차

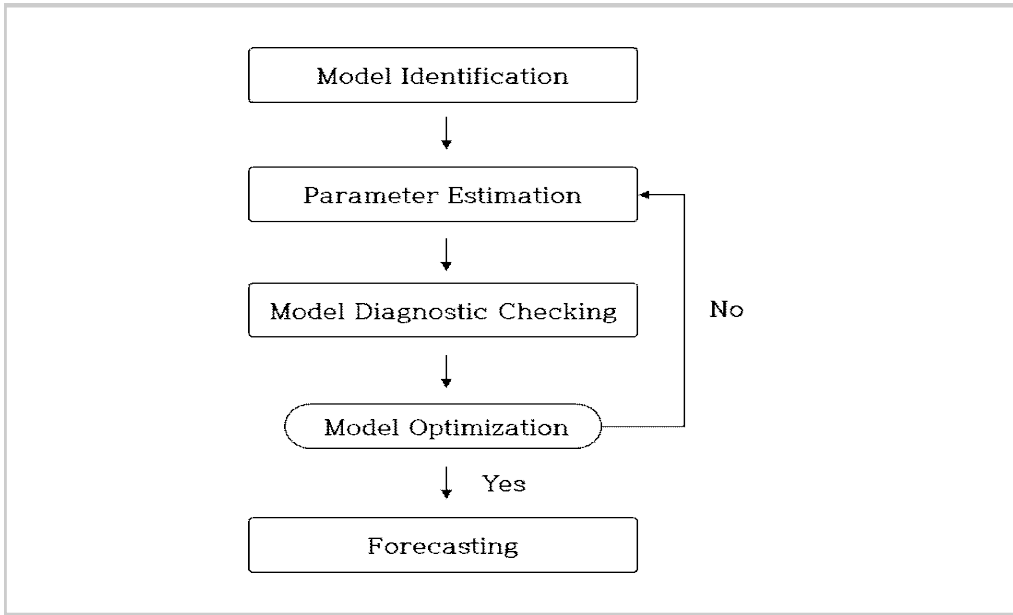
시계열모형은 시간의 흐름에 따라 나타나는 시계열 패턴이 미래의 관측치에 영향을 준다는 가정에서 예측하는 방법으로 시계열모형과 확률적 시계열모형 두 가지로 구분할 수 있다.

고전적 방법인 시계열모형은 지수평활모형으로 시계열자료에 확률적 방법론을 반영한 확률적 시계열모형은 ARIMA모형으로 수요를 예측한다. 특히 ARIMA 모형은 Box-Jenkins 분석 방법에 따라 주어진 시계열 자료의 식별, 추정, 검증, 예측과정을 통해 모형을 추정하여 수요를 예측하게 된다.

모형의 식별단계에서는 원시계열자료의 그래프를 통해 시계열의 추세와 계절성을 확인하며 필요할 경우 차분과 로그변환을 통해 시계열자료의 안정성을 확보한다. 시계열의 정상성(Stationary)을 검증하기 위해서는 원시계열자료의 그래프를 이용해 직관적으로 파악할 수 있으나 통계적 검증이 필요하기 때문에 단위근검정(Unit Root test)이 사용된다.

이후 정상성(Stationary)검정이 완료된 시계열 자료는 상관도표(correlogram)을 통해 자기상관함수(Autocorrelation Function, ACF)와

Fig. 2. Time series modelling steps



부분자기상관함수(Partial Autocorrelation Function, PACF)을 분석하여 ARIMA(p, d, q) 모형을 구성하는 AR항의 차수(p), MA항의 차수(q), 차분의 차수(d) 또는 SARIMA(p, d, q)(P, D, Q) 모형의 각항의 차수(d)를 임시적으로 결정하게 된다.

모형추정 단계에서는 식별단계에서 추정된 각 항의 모수를 추정하게 되는데 모수를 추정하는 방법으로는 최소제곱추정법, 최대우도추정법 등이 있으며 이와 함께 AIC(Akaike's Information Criterion)통계량, BIC(Baysian Information Criterion)통계량, 유의성 검정 등을 고려하여 모형의 모수를 최종결정하게 된다.

이후 진단단계에서는 예측된 SARIMA모형에 대한 검증으로 잔차분석, 과대적합진단 등 유의성 검정을 활용하여 판단하며 최종적으로 결정된 모형의 이상이 없다고 판단될 경우 수요 예측을 진행하게 된다.

IV. 실증분석

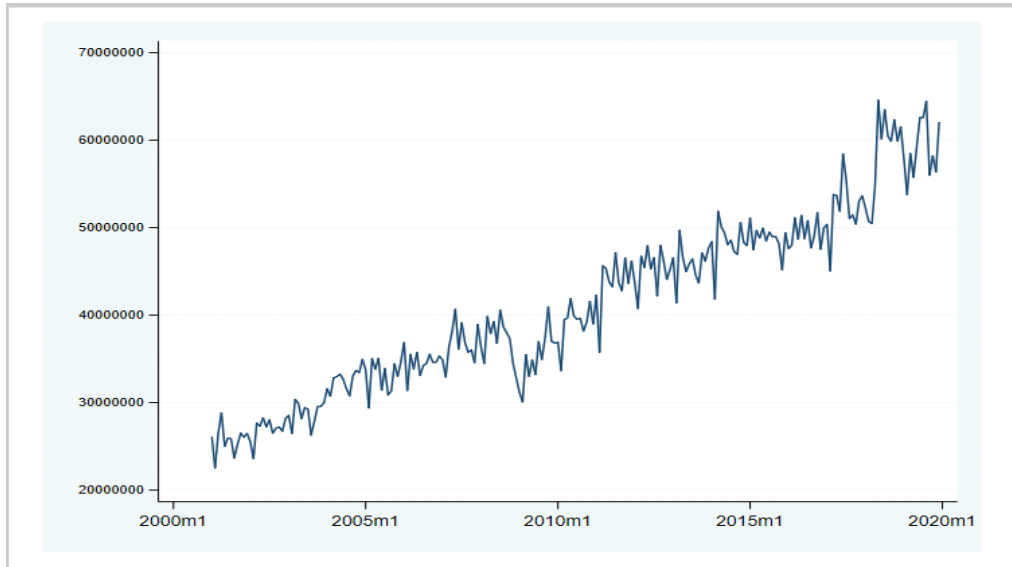
1. 모형식별

본 연구에 사용된 시계열자료는 해양수산부 해운항만물류정보(Port-MIS)에서 제공하고 있는 월별 컨테이너 물동량 자료로 2001년 1월부터 2019년 12월까지 환동해권 내 항만(부산항, 동해항, 속초항, 포항항, 울산항)의 월별 자료를 활용하였다.

일반적으로 항공이나 화물의 수요는 계절적 특성을 지니고 있는 것으로 알려져 있으며, 본 연구의 분석대상인 컨테이너 물동량 시계열자료의 정상성 검증을 위해 1차적으로 그래프를 통해 직관적으로 파악해보고자 한다.

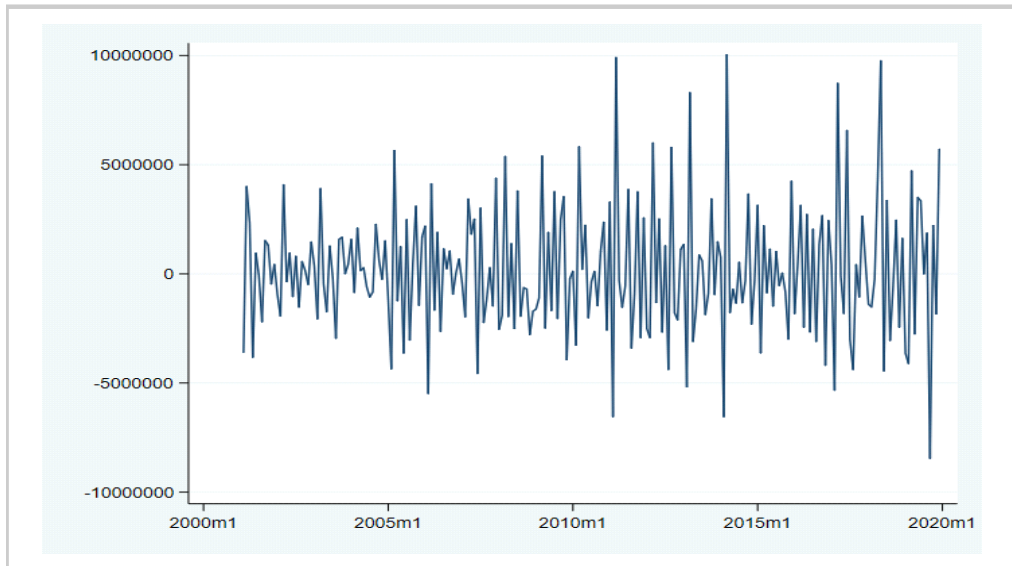
〈Fig. 3〉은 환동해권 내 항만의 컨테이너 물동량에 대한 그래프로 전체 기간에 걸쳐 점차 증가하는 추세를 보이고 있어 차분을 통한 시계열데이터의 안정화가 필요한 것으로 나타났다. 이에 차분을 통해 추세를 제거 한 결과 〈Fig. 4〉와 같이 나타났으며, 일정한 주기 및 계절성을 가진다고 판단하였다.

Fig. 3. Trade volume (2010~2019)



Source:

Fig. 4. Time series data of Differencing (2010~2019)

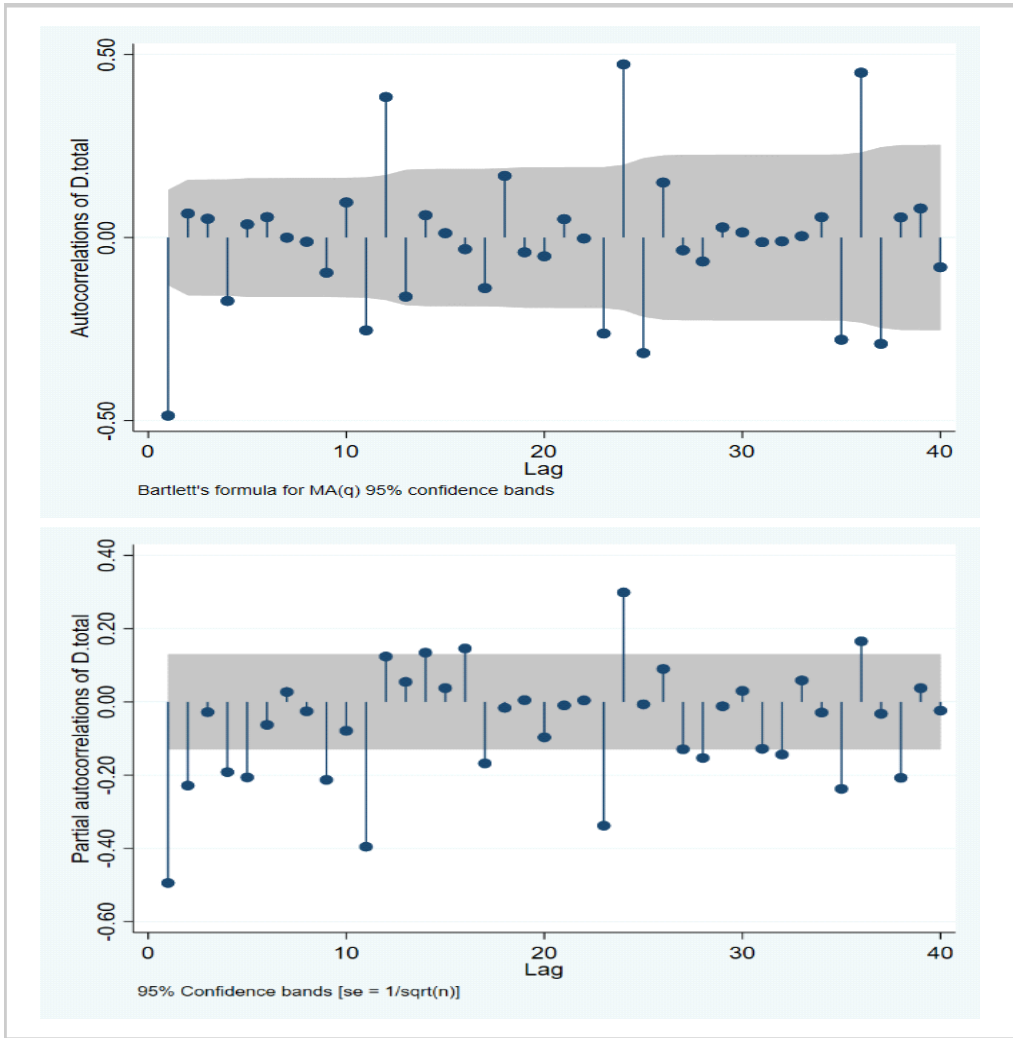


Source:

또한 안정성 여부를 확인하기 위한 통계적 검증을 위해 상관도표(Correlogram)분석과 단위근검정을 함께 활용하였으며, 단위근 검정은

Dickey-Fuller 검정기법으로 분석을 진행하였다. 상관도표를 통해 자기상관계수가 지속적으로 감소하는 불안정한 시계열자료라는 것을 확

Fig. 5. Auto Correlation Function and Partial Auto Correlation Function



인하였으며, 이에 차분과정을 통해 안정성을 가지도록 하였다. 그 결과는 <Fig 5>의 상관도 표로 나타내었다.

AIC와 BIC 값은 우도값이 높아질수록 작아지고 모형을 추정하는 모수의 수가 많을수록 높아지기 때문에 AIC와 BIC는 값이 작을수록 모형의 적절성이 높아진다는 것을 의미하며 각 모형에 대한 추정치는 <Table 2>와 같으며, 환동해권 내 물동량 수요를 예측하기 위한 최적의 모형은 SARIMA(2,1,1)(0,1,1)₁₂로 나타났다.

2. 예측 정확도 (검증)

본 연구는 2001년 1월부터 2019년 12월까지 환동해권 내 항만 물동량 자료를 바탕으로 예측된 SARIMA(2,1,1)(0,1,1)₁₂ 모형을 활용하여 2021년 12월까지의 물동량을 예측하였다. 이렇게 추정된 모형을 대상으로 물동량 수요예측 분야에서 가장 많이 활용되는 ARIMA 모형과의 비교분석을 통해 모형의 상대적 적절성을 검증하였으며, 예측정확도를 측정하는 MAPE값으

Table 2. ARIMA model estimation

		Model(1)	Model(2)	Model(3)	Model(4)
		SARIMA(2,1,0)(0,1,1)12	SARIMA(2,1,1)(1,1,0)12	SARIMA(2,1,1)(0,1,1)12	SARIMA(0,1,1)(0,1,1)12
AR	L1	-.51178933***	-1.1481607***	.32751328***	
	L2	-.22289681***	-.48549807***	.17642719**	
MA	L1		.6831292***	-.87782525***	-.56553997***
SAR	L1		-.55556865***		
SMA	L1.	-.81014831***		-.86802887***	-.83124333***
sigma_cons		2069124.6***	2234102.7***	2009967.7***	2046440.5***
N		215	215	215	215
aic		6884.5438	6911.2906	6878.7247	6879.2262
bic		6898.0264	6928.1438	6895.5779	6889.3381
ll		-3438.2719	-3450.6453	-3434.3624	-3436.6131

* p<0.05; ** p<0.01; *** p<0.001

Table 3. Comparison SARIMA(2,1,1)(0,1,1,12) with ARIMA(2,1,1)

		ARIMA(2,1,1)		SARIMA(2,1,1)(0,1,1)12	
		Coef.	z-value	Coef.	z-value
		(Std. Err.)	P>z	(Std. Err.)	P>z
AR	L1.	.21517875** (0.069645)	3.09 (0.002)	.32751328*** (0.0932428)	3.51 (0.000)
	L2.	.21405186** (0.0709215)	3.02 (0.003)	.17642719** (0.0618946)	2.85 (0.004)
MA	L1.	-.90897489*** (0.0419572)	-21.66 (0.000)	-.87782525*** (0.0652583)	-13.45 (0.000)
SMA	L1.			-.86802887*** (0.0659458)	-13.16 (0.000)
sigma_cons		2413311.6*** (86953.50)	27.75 (0.000)	2009967.7*** (65241.96)	30.81 (0.000)
N		227		215	
aic		7327.4252		6878.7247	
bic		7344.5500		6895.5779	
ll		-3658.7126		-3434.3624	

* p<0.05; ** p<0.01; *** p<0.001

로 예측모형 검증을 추가적으로 진행하였다.

〈Table 3〉은 두가지 예측 모형을 비교 분석한 표로 AIC, BIC값의 경우 SARIMA(2,1,1)(0,1,1)12모형에서 작은 값을 나타내고 있고, 설명력을 나타내는 R2값 또한 SARIMA(2,1,1)(0,1,1)12모형이 0.95으로 ARIMA(2,1,1)모형보다 높게 나

타나 SARIMA(2,1,1)(0,1,1)12모형 좀 더 적합하다는 판단이 가능하다.

또한 분석된 SARIMA(2,1,1)(0,1,1)12모형 예측 정확도는 예측치와 실측치를 통해 MAE(Mean Absolute Error), MSE(Mean Squared Error), RMSE(Root Mean Squared Performance),

Table 4. Measure of Forecast Error

Accuracy Measure	Calculation
MAE	$\frac{\sum_{t=1}^n A_t - F_t }{n}$
MSE	$\frac{\sum_{t=1}^n (A_t - F_t)^2}{n}$
RMSE	$\sqrt{\frac{\sum_{t=1}^n (A_t - F_t)^2}{n}}$
MAPE	$\frac{\sum_{t=1}^n A_t - F_t }{n} \times 100$

Table 5. Interpretation of MAPE Results for Forecasting Accuracy

MAPE	Accuracy of forecast
< 10%	Highly Accurate Forecasting
10% ~ 20%	Good Forecasting
20% ~ 50%	Reasonable Forecasting
> 50%	Inaccurate Forecasting

Source: 5Lewis (1982)

MAPE(Mean Absolute Percentage Error)를 이용하여 측정할 수 있다.

RMSE,와 MAE의 경우 상대적으로 큰 예측 오류를 밝히는데 좋은 척도가 되기 때문에 본 연구에서는 다른 평가지표보다 신뢰도가 높고 이상치에 대한 민감도가 낮은 MAPE 적합도 계수를 활용하여 모형의 정확도를 확인하였다.

$$MAPE = \frac{\sum_{t=1}^n |A_t - F_t|}{n} \times 100 \quad (7)$$

식 (7)에서 A_t 는 실측치 F_t 는 예측값, n 은 관측기간(월)을 나타내며, MAPE 적합도 계수

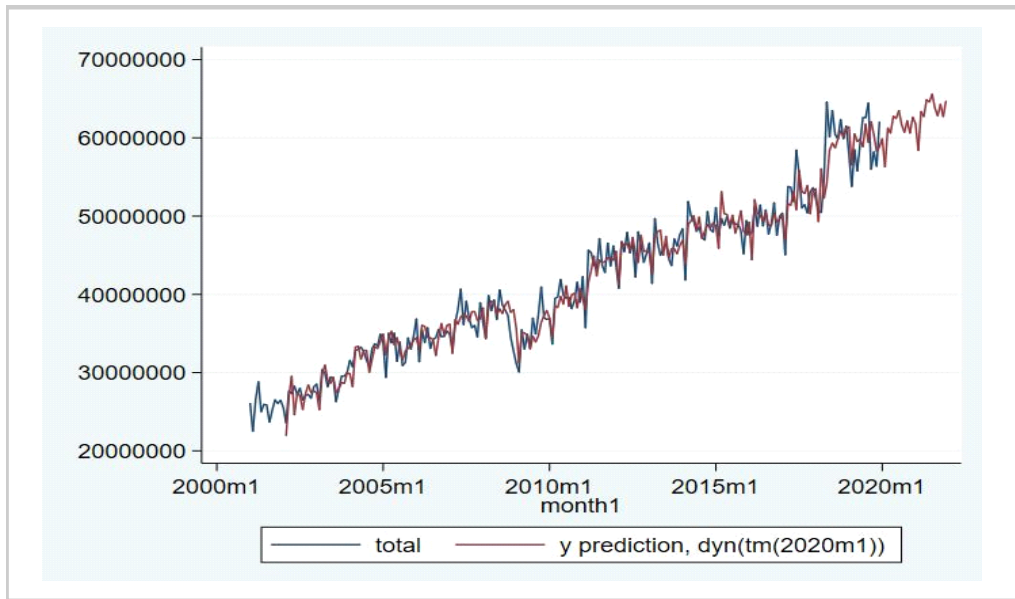
는 $0\% \leq MAPE < 10\%$ 는 매우 적합한 예측, $10\% \leq MAPE < 20\%$ 은 비교적 정확한 예측, $20\% \leq MAPE < 50\%$ 비교적 합리적 예측, $MAPE \geq 10\%$ 는 부적절함을 기준으로 모형의 예측 적합도를 판단한다.

이에 SARIMA(2,1,1)(0,1,1)₁₂ 모형의 예측치 자료와 2020년 1월부터 9월까지 실측치 자료에 대한 MAPE 값을 분석하였으며, 그 결과 7%로 정확한 예측력을 보이는 것으로 판단 할 수 있다.

3. 물동량 예측

상기 분석 내용을 바탕으로 SARIMA(2,1,1)(0,1,1)₁₂ 모델로 추정된 환동해권 내 물동량

Fig. 6. Prediction Trade volume by SARIMA(2,1,1)(0,1,1)12



수요 추정 결과는 <Fig. 6>과 같으며 실제 환동해권 항만 물동량과 SARIMA(2,1,1)(0,1,1)12 모형의 예측치를 비교하여 나타내었다.

SARIMA(2,1,1)(0,1,1)12의 경우, 2010년 예측시점부터 시계열데이터의 계절성 및 주기성을 반영한 물동량을 예측하고 있는 것으로 나타나고 있어 계절적 특성을 지닌 환동해권 내 물동량을 추정하는 예측모델로 적합한 모형이라는 것을 확인할 수 있다.

SARIMA(2,1,1)(0,1,1,12) 모형으로 추정된 월별 물동량은 Table 6과 같이 제시하였으며 2020년 1월 60,100,000TEU를 시작으로 2021년 12월 64,900,000TEU까지 월별 증감은 있지만 2030년까지 점진적으로 증가하는 추세를 보이고 있다.

V. 시사점 및 결론

본 연구는 환동해 경제권 활성화 방안 및 세부 전략방안 구축을 위해 필요한 기본자료인 물동량수요를 예측하고 추정모델을 구축하는

데 연구의 목적이 있다. 이를 위해 2001년부터 2019년 12월까지 환동해권 항만의 월별 물동량 시계열자료를 바탕으로 SARIMA모형을 구축하여 2021년까지 장래 물동량을 예측하였다. 이후 SARIMA모형의 적합성 확인을 위해 ARIMA 모형과의 비교분석을 진행하였으며 2020년 1월부터 7월까지의 예측치와 실측치를 비교하여 예측정확도를 검증하였다.

이를 위해 먼저, 상관도표(Correlogram)의 AC, PAC 분석을 통해 추세와 계절성이 있는 비정상성을 보이는 자료임을 확인하였으며, 차분(Differencing)을 통해 안정적 시계열로 전환하여 연구분석을 진행하였다. 연구결과 환동해권 물동량과 같이 계절성 및 주기성을 가진 시계열 자료의 경우 ARIMA(2,1,1)모형보다 SARIMA(2,1,1)(0,1,1)12모형이 적합하다는 것을 확인할 수 있었으며 구축된 모형으로 향후 물동량을 예측하였다.

환동해권 항만의 물동량 예측결과 점차 증가하는 추세를 보이고 있지만 그 증가율은 크지 않은 것으로 나타났다. 이를 통해 향후 물동량 제고를 위한 전략수립 시 새로운 물동량 창출

Table 6. Prediction of Trade Volume

(Unit: Ton)

	Forecasting
2020.01	60,100,000
2020.02	57,000,000
2020.03	61,800,000
2020.04	61,100,000
2020.05	63,000,000
2020.06	62,600,000
2020.07	63,600,000
2020.08	61,800,000
2020.09	61,000,000
2020.10	62,500,000
2020.11	61,000,000
2020.12	62,800,000
2021.01	62,200,000
2021.02	58,800,000
2021.03	63,800,000
2021.04	63,200,000
2021.05	65,100,000
2021.06	64,600,000
2021.07	65,700,000
2021.08	63,900,000
2021.09	63,100,000
2021.10	64,600,000
2021.11	63,000,000
2021.12	64,900,000

을 위한 신성장동력육성, 항만배후단지 내 기업유치 등 항만의 효율성을 높일수 있는 전략에 초점을 맞추어야 할것으로 판단된다.

앞서 언급한 내용과 같이 수요예측의 경우 연구의 목적과 주제에 따라 다양한 연구방법론을 활용할 수 있지만 본 연구에서는 단기 예측의 정확도가 높은 SARIMA모형을 통해 미래 물동량을 예측하고 모델의 적합성 및 예측정확도를 확인하였다는 점에서 연구의 의의를 둘 수 있다. SARIMA모형의 경우, 다른 시계열모형과 마찬가지로 거시경제변수를 반영한 인과관

계를 분석하기에는 어려운 단점을 가지고 있지만 시계열자료만으로 예측모형을 구축하여 분석할 수 있으며 단기예측의 정확성과 유용성이 높다는 점에서 큰 이점을 가지고 있기 때문이다.

본 연구는 연구범위를 환동해권 항만을 대상으로 설정하고 SARIMA모형을 통해 수요를 예측했다는 점에서 기존연구와의 차별성을 가지고 있지만 각 항만의 차별화된 특성을 확인하기에는 어려운 한계점을 지니고 있다. 이에 향후 연구에서는 환동해권 내 각 항만별 물동량 분석 및 수요예측이 이루어져야 할 것이며, 이

를 통해 SARIMA모형의 정확도 검증은 물론 세부노선별 특성을 반영한 다양한 분석모형과의 비교분석을 통해 예측정확도가 높은 모형을 파악하는 연구가 필요할 것으로 판단된다. 또한 환동해권 물동량을 품목별, 국가별로 나누어

추가 연구를 진행한다면 환동해권 경제 활성화를 위한 구체적인 전략방안을 마련할 수 있는 의미 있는 자료로 활용될 수 있을 것이라 기대한다.

References

- Box, George EP, and George C. Tiao (1975), "Intervention analysis with applications to economic and environmental problems", *Journal of the American Statistical association*, 70(349), 70-79.
- Dragan, D. and T. Kramberger (2014), "Forecasting the Container Throughput in the Port of Koper using Time Series ARIMA model", *Proceedings of the conference ICLST*, 14.
- Farhan, J. and G. P. Ong (2018), "Forecasting seasonal container throughput at international ports using SARIMA models", *Maritime Economics&Logistics*, 20(1), 131-148.
- G. De Jong, H. Gunn and W. Walker, (2004), "National and international freight transport models: An overview and ideas for future development", *Transport Reviews*, 24(1), 103-124.
- Go. Yong-Gi, et al (2008), "Forecasting Export Loaded Container Throughput of Incheon Port", *Journal of Korea Port Economic Association* 24(3), 57-77.
- Jeong, Su-hyun, and Chang-Hoon Shin (2016), "A Study on the Impact of the Financial Crises on Container Throughput of Busan Port", *Journal of Korea Port Economic Association*, 32(2), 25-37.
- Jo, Jun-Ho, Je-Seop Byon and Hee-Cheul Kim (2017), "Analysis of Global Shipping Market Status and Forecasting the Container Freight Volume of Busan New port using Time-series Mode", *jkiiect*, 10, 295-303.
- Kim, Chang-Beom (2015), "Forecasting the Seaborne Trade Volume using Intervention Multiplicative Seasonal ARIMA and Artificial Neural Network Model", *Journal of Korea Port Economic Association*, 31(1), 69-84
- Lewis, Colin David (1982), *Industrial and business forecasting methods: A practical guide to exponential smoothing and curve fitting*. Butterworth-Heinemann
- Min, Kyung-Chang and Hun-Koo Ha (2014). "Forecasting the Korea's Port Container Volumes With SARIMA Model", *Journal of Korean Society of Transportation* 32(6), 600-614.
- Suh, Bo-Hyoun, Tae-Woong YANG and Hun-Koo HA (2017), "Application of SARIMA Model in Air Cargo Demand Forecasting: Focussing on Incheon-North America Routes", *Journal of Korean Society of Transportation* 35(2), 143-159.
- Shin, Chang-Hoon and Su-Hyun Jeong (2011), "A study on application of ARIMA and neural networks for time series forecasting of port traffic", *Journal of Navigation and Port Research*, 35(1), 83-91.
- Silva, Emmanuel Sirimal, and Hossein Hassani (2015), "On the use of singular spectrum analysis for forecasting US trade before, during and after the 2008 recession", *International Economics* 141, 34-49.

-
- Wei, Yu and Mu-Chen Chen (2012), "Forecasting the short-term metro passenger flow with empirical mode decomposition and neural networks", *Transportation Research Part C: Emerging Technologies*, 21(1), 148-162.
- Yi, Ghae-Deug (2013), "Forecasting the Container Throughput of the Busan Port using a Seasonal Multiplicative ARIMA Model", *Journal of Korea Port Economic Association*, 29(3), 1-23.
- Yoo, Sang-Lok, et al (2014), "A prediction of marine traffic volume using artificial neural network and time series analysis", *Journal of the Korean Society of Marine Environment & Safety*, 20(1), 33-41.
- Yoon, Han-Young, and Sung-Sik Park (2019), "Analysis and Estimation of Food and Beverage Sales at Incheon Int'l Airport by ARIMA-Intervention Time Series Model", *Journal of the Korea Academia-Industrial cooperation Society*, 20(2), 458-468.