

A Study of Air Freight Forecasting Using the ARIMA Model

ARIMA 모델을 이용한 항공운임예측에 관한 연구

Sang-Sok Suh(서상석)*, Jong-Woo Park(박종우)**, Gwang Suk Song(송광석)***, Seung-Gyun Cho(조승균)****

Received: December 23, 2013. Revised: January 20, 2014. Accepted: February 16, 2014.

Abstract

Purpose – In recent years, many firms have attempted various approaches to cope with the continual increase of aviation transportation. The previous research into freight charge forecasting models has focused on regression analyses using a few influence factors to calculate the future price. However, these approaches have limitations that make them difficult to apply in practice: They cannot respond promptly to small price changes and their predictive power is relatively low. Therefore, the current study proposes a freight charge-forecasting model using time series data instead a regression approach. The main purposes of this study can thus be summarized as follows. First, a proper model for freight charge using the autoregressive integrated moving average (ARIMA) model, which is mainly used for time series forecast, is presented. Second, a modified ARIMA model for freight charge prediction and the standard process of determining freight charge based on the model is presented. Third, a straightforward freight charge prediction model for practitioners to apply and utilize is presented.

Research design, data, and methodology – To develop a new freight charge model, this study proposes the ARIMAC(p,q) model, which applies time difference constantly to address the correlation coefficient (autocorrelation function and partial autocorrelation function) problem as it appears in the ARIMA(p,q) model and materialize an error-adjusted ARIMAC(p,q). Cargo Account Settlement Systems (CASS) data from the International Air Transport Association (IATA) are used to predict the air freight charge. In the modeling, freight charge data for 72 months (from January 2006 to December 2011) are used for the training set, and a prediction interval of 23 months (from January 2012 to November 2013) is used for the validation set. The freight charge from November 2012 to November 2013 is pre-

dicted for three routes – Los Angeles, Miami, and Vienna – and the accuracy of the prediction interval is analyzed using mean absolute percentage error (MAPE).

Results – The result of the proposed model shows better accuracy of prediction because the MAPE of the error-adjusted ARIMAC model is 10% and the MAPE of ARIMAC is 11.2% for the L.A. route. For the Miami route, the proposed model also shows slightly better accuracy in that the MAPE of the error-adjusted ARIMAC model is 3.5%, while that of ARIMAC is 3.7%. However, for the Vienna route, the accuracy of ARIMAC is better because the MAPE of ARIMAC is 14.5% and the MAPE of the error-adjusted ARIMAC model is 15.7%.

Conclusions – The accuracy of the error-adjusted ARIMAC model appears better when a route's freight charge variance is large, and the accuracy of ARIMA is better when the freight charge variance is small or has a trend of ascent or descent. From the results, it can be concluded that the ARIMAC model, which uses moving averages, has less predictive power for small price changes, while the error-adjusted ARIMAC model, which uses error correction, has the advantage of being able to respond to price changes quickly.

Keywords: Forecasting of Time Series, ARIMA(p,q), ARIMAC, Error Adjusted ARIMAC.

JEL Classifications: C15, M21, M31, M39.

1. 서론

최근 세계은행은 전 세계 155개 국가의 물류성과지수(Logistics Performance Index, LPI)를 발표하였다. 발표결과에 의하면 한국은 5점 만점에 3.7점으로 150여개 국가 중 21위를 차지하는 것으로 나타났다. 물류관련 리서치 기관인 Armstrong & Associates가 2011년 발표한 세계 제3자 물류기업(3rd Party Logistics) 순위 발표에 있어서도 국내 기업은 50위에 단 2개 기업만이 포함된 것으로 나타났다. 이러한 원인에 대해 대한상공회의소에서 발표한 연구결과에 의하면, 기업운영의 영세성에 대한 부분을 원인으로 지적하며 기업들이 고부가가치 서비스를 제공하기보다 주로 가격에 의한 시장경쟁을 추구하고 있어 산업의 지속적인 성장을 저해

* First Author, Hyundai GLOVIS, Korea.

** Corresponding Author, Professor, Division of Business Administration, Soongsil University, Korea. Tel:+82-02-828-7251, E-mail: jongpark7@ssu.ac.kr.

*** Division of Business Administration, Soongsil University.

**** Graduate School of Business Administration, Soongsil University.

한다고 지적하였다. 전반적으로 국내 물류의 경쟁력 수준이 경쟁 상대인 홍콩이나 싱가포르에 비해 떨어지고 있는 상황이다(KCCI, 2013). 하지만 최근 물류환경의 변화와 더불어 기업 공급망의 글로벌 확대에 의해 전문 물류기업들은 제조 및 유통기업들과 협업을 통해 전 세계를 대상으로 영업망을 확대하고 있다. 특히, 글로벌 항공시장 전망에 대한 에어버스 보고서에 의하면 전 세계 항공 화물시장은 연평균 5%이상의 성장을 예상하고 있으며 항공화물의 주종을 이루는 것이 유행과 시기에 민감한 고부가가치 IT상품들이 주를 이루고 있는 것으로 나타났다(www.koila.or.kr).

국내 항공물류 시장도 LCD, 반도체 등 주로 첨단제품의 수출물량이 크게 증가하고 있으며 환적화물도 지속적인 성장세를 나타내고 있다. 하지만 항공운임의 경우 다양한 외부 영향요인에 의해 가격이 결정되기 때문에 기간별 변동성이 크게 나타나고 있으며 이러한 운임의 변동성은 운영기반이 약한 영세 물류기업에 있어 시장에 대한 불확실성을 가중시키고 기업운영에 있어 치명적인 문제를 야기할 수 있다. 또한 공급망을 운영하는 기업들의 경우 공급망의 채찍효과를 야기할 수 있으며 제조기업에 있어서는 생산계획 및 물류계획을 반영한 수요계획 전반에 영향을 미치게 된다. 따라서 시장 글로벌화에 따른 효율적인 물류 환경을 구축하고 협업적 공급망 관리체계를 구축하기 위해서는 물동량 정보 예측과 더불어 운임에 대한 예측은 기업의 생존을 위한 필수적인 활동이라 할 수 있다. 특히, 공급중심의 대량생산방식이 아닌 수요중심의 맞춤 생산방식 환경에서는 다른 어떤 활동보다 수요예측에 대한 정보는 중요하다 할 수 있다. 그럼에도 불구하고 물류기업의 운임 예측에 대한 연구는 그리 많지 않은 실정이다.

최근 국내 항공물동량의 지속적인 증가에 맞춰 많은 기업들은 항공운임을 예측하기 위한 다각적인 접근을 시도하고 있는데 기존에 제시된 운임예측모델의 경우 소수의 영향요인을 활용하여 가격을 예측하는 회귀분석방법을 주로 이용하고 있다. 하지만 이러한 분석방법은 다양한 외부환경요인에 의해 결정되는 항공운임의 예측에 적용할 경우 민감한 가격변화에 빠르게 대응하지 못하며 예측력도 상대적으로 낮아 실무적으로 활용하는데 한계점을 가지고 있다. 따라서 본 연구에서는 기존 연구에서 주로 활용한 회귀분석 접근방법이 아닌 시계열 자료를 활용한 운임예측 모형을 제시하고자 한다. 특히, 시계열자료 분석시 주로 활용하는 기존의 ARIMA 모형을 개선하여 실무적으로 활용 가능한 예측모형을 제시하고자 한다. 이러한 측면에서 본 연구의 목적을 제시하면 다음과 같다.

첫째, 시계열 예측에 주로 활용하는 ARIMA모형을 이용하여 항공운임예측에 적합한 모형을 구축하고자 한다. 특히, 노선별 항공운임정보를 기반으로 운임예측 모형을 제시하여 관련기업들에게 활용 가능한 정보를 제시하고자 한다.

둘째, 기존 ARIMA모형의 단순 적용보다 항공운임 예측에 특화된 ARIMA모형을 개발하고 이를 기반으로 운임예측의 표준 프로세스를 제시하고자 한다.

셋째, 현업의 실무자들이 쉽게 적용 및 활용 가능한 항공운임 예측모형을 제시하고자 한다. 일반적으로 시계열 예측모형은 복잡한 수식과 예측과정의 어려움으로 인해 실무자들에게 많이 활용되지 않고 있는 한계점이 있는데 본 연구에서는 스프레드시트를 활용하여 실무에 적용 가능한 운임예측 모형을 제시하고자 한다.

이를 위해 항공운임 예측시 활용한 운임정보는 국제항공운송협회(International Air Transport Association, IATA)에서 제공하는 CASS(Cargo Account Settlement Systems)자료를 이용하였다.

예측모형 구축시 활용할 분석용 표본(training set)은 2006년 1월부터 2011년 12월까지 72개월의 항공운임을 이용하였으며 일종

의 검증표본(validation set)인 예측구간은 2012년 1월부터 2013년 11월까지 23개월을 이용하였다. 분석노선은 LA, Miami, Vienna의 3개 노선에 대해 2012년부터 2013년 11월까지의 항공운임을 예측하고자 한다.

2. 이론적 배경

2.1. 수요예측모델에 관한 연구

수요 및 가격 예측 모델은 크게 정성적인 방법과 정량적인 방법으로 분류할 수 있다. 정성적인 분석방법은 주로 계량적인 분석방법을 적용하기 어렵거나 전문가들의 정성적인 지식을 활용하여 미래 기간을 예측할 경우에 주로 활용하게 된다.

대표적인 기법으로는 델파이기법이 있다. 그러나 정성적인 예측 기법은 전문가 지식베이스를 활용하기 때문에 주관적인 특성이 반영될 소지가 있으며 예측기간이 길수록 예측의 정확도가 낮아지는 한계점이 있다. 이러한 문제점으로 인해 최근에는 정성적인 기법과 정량적인 기법을 결합한 전문가 평가방법이 활용되고 있다(Hibon & Evgeniou, 2005). 반면 정량적인 분석방법은 인과분석과 시계열 분석으로 나눌 수 있는데 인과분석은 하나 혹은 그 이상의 독립변수와 종속변수와의 함수적 관계를 가정하여 예측하는 방법으로 일반적으로 예측모델로 폭넓게 활용되고 있다.

시계열 예측방법은 과거의 시계열 형태가 미래에도 같은 형태로 반복되리라는 기본 가정 하에 시계열 자료간의 상호관계를 밝히고, 이를 기반으로 미래 시계열에 대하여 예측하는 방법을 활용하게 된다. 이러한 시계열 분석기법은 크게 사용변수의 수에 따라 단변량 시계열 분석방법과 다변량 시계열 분석방법으로 분류할 수 있는데 이 중 단변량 기법은 시간에 따른 변수의 움직임을 변수의 과거정보만으로 미래를 예측한 기법으로 분석의 주목적이 시간적인 변동에 따른 예측일 경우 다변량 기법보다 예측 정확도가 높을 수 있다(Li et al., 2005; Song & Lee, 2009). 또한 예측기법 간 예측정확도에 대한 연구(Witt et al., 1992; Chu, 2004)에 의하면 대부분의 연구에서 정성적인 예측기법보다 정량적인 예측기법이 우수한 것으로 나타났으며, 복잡한 예측모형보다 단순한 예측모형이 예측오차를 줄이는데 더 효과적인 것으로 나타났다. 따라서 본 연구에서는 단일변량기법으로 가장 광범위하게 사용되고 있는, Box and Jenkins (1976) approach 기반의 ARIMA(Auto Regressive Integrated Moving Average) 모형을 활용하여 항공운임을 예측하고자 한다.

수요 및 가격예측에 관한 대부분의 연구는 예측을 위해 새롭게 모델을 구축하여 검증하는 방법이 아닌 기존의 제시된 모형을 예측모델로 활용한 연구들이 주를 이루고 있다.

우선 국내의 관광수요를 예측한 Lee and Song (2007)의 연구에서는 다양한 시계열 모델과 ARIMA 개입모형에 대해 예측정확도를 분석하였다. 분석결과 단계적 회귀모델, Winters 지수평활모델, ARIMA모델, ARIMA 개입모델 등을 비교하였다. 분석모델 모두 MAPE 값이 0%에서 10% 내에 있어 매우 정확한 예측결과를 나타냈다. 그 중에서도 ARIMA 개입모델의 정확도가 가장 우수한 것으로 나타났다.

Totamane et al. (2009)은 seasonal ARIMA 모델과 기간별 가중치를 결합한 모델을 제시하여 항공화물의 수요를 예측하였으며 Chen et al. (2009)는 Holt-Winters모형과 GRE 모형, 그리고 Seasonal ARIMA 모형을 활용하여 일본, 홍콩, 미국에서 대만을

방문한 관광객수를 예측하였다.

Wu and Morrell (2007)은 단계적 회귀분석을 활용하여 중국내 주요 공항과 홍콩, 대만의 항공화물에 대한 장기수요를 예측하였다. 특히, 단계적 회귀분석의 통해 개별 공항의 월별 예측치를 제시하였는데 일반적인 회귀분석 예측치보다 정확도가 17%이상 높게 나타났다.

Coshall (2006)은 ARIMA 방법론을 이용하여 영국의 항공여객 수요에 대해 예측하였는데 다변량 시계열 예측방법보다 단변량 시계열 예측방법이 예측정확도가 높게 나타났으며 다양한 예측변수를 활용한 방법보다 예측모형의 구조적 관점에서 간단한 모형이 더 좋은 결과를 나타낼 수 있다는 결과를 제시하였다.

Coshall (2006)의 연구와 유사한 국내 연구로 Hur (2010)는 국제항공 여객수요와 국제항공 화물수요예측에 회귀분석과 일변량 시계열 모형 및 다변량 시계열 모형을 활용한 수요예측결과를 비교 분석하였다. 분석결과 회귀모형보다는 일변량 시계열 모형이 우수하였으며 일변량 시계열 모형보다는 다변량 시계열 모형이 예측에 있어 더 우수한 결과를 보이는 것으로 나타났다.

2.2. ARIMA 모형의 특징

ARIMA 모형은 기본적으로 시계열 자료의 과거 특성에 기초한 모형구축 작업을 통해 미래 경향을 예측하는 방법으로 알려져 있다. 이러한 분석방법은 Box and Jenkins (1976)가 제시한 ARIMA 모형을 통해 구체화 되었다. 이러한 ARIMA 모형은 시계열 자료의 구조적 관계 특성을 분석하기보다 시계열 자체에 대한 통계적 성질을 분석하는데 초점을 두고 있는 방법으로 t기간의 예측 값인 y_t 를 k개의 설명변수인 $x_1, x_2, \dots, x_k$ 에 의해 설명하지 않고 y_t 의 과거시차 변수들과 확률적 오차항들만으로 설명하는 방법이다. 이러한 이유로 인해 ARIMA 모형은 무이론적 접근방법(atheoretic approach)으로 불리어지기도 한다.

ARIMA 모형은 AR, MA, ARMA, ARIMA 모형으로 다양하게 설명할 수 있는데 시계열 자료에서 한 시점의 관측치는 그 이전 관측치들의 영향을 받거나 과거 관측치의 오차에 영향을 받게 되는데 이를 이용한 것이 자기회귀모형(autoregressive model; AR)이며 n기간의 이동평균과정의 오차항 특성을 이용한 것이 이동평균 모형(moving average model, MA)이다. 이러한 두 모형이 결합된 방법이 ARMA모형이다. 이를 수식으로 정리하면 다음과 같이 나타낼 수 있다.

$$Z_t = \phi_1 z_{t-1} + \dots + \phi_p z_{t-p} + a_t - \phi_1 a_{t-1} - \dots - \phi_q a_{t-q} \quad (1)$$

$\phi_p z_t$ 부분은 AR특성을 반영한 부분이고 $\phi_q a_{t-q}$ 부분은 오차항에 대한 이동평균특성을 나타내고 있다. 즉, AR모형에 기간별 오차항이 특정기간 동안 연속적으로 결합되어 있는 특성이 나타나게 된다. 일반적으로 일변량 ARIMA 방법은 정상적인(stationary) 시계열 자료에만 적용할 수 있는데 ARIMA(Auto-Regressive Integrated Moving Average)모형은 AR모형, MA모형, 그리고 정상성 처리부분에 대한 세부분이 조합을 통해 미래 시점을 예측하게 된다.

정상적인 시계열의 경우 평균, 분산, 자기상관함수는 시간의 흐름에 따라 일정하게 유지되는 특성이 있다. 즉, 시계열 자료에 추세나 분산변동 등이 안정적일 때 적용할 수 있는 모형이다. 이러한 ARIMA 모형은 단일 시계열분석일 경우 ARIMA(p,d,q)로 표현할 수 있는데 p는 ARIMA 모형의 차수를 나타내며 d는 분석시계열 자료의 차분(differencing) 차수를 의미한다. 차분(differencing)은

시계열 자료의 반복적인 상승 혹은 하강의 패턴이 있을 경우 차분하여 정상시계열로 변환할 수 있다. 끝으로 q는 이동평균인 MA모형의 기간차수를 나타낸다. 이동평균인 MA(n)는 시계열 자료를 n 기간 동안 시차를 발생시켜 시계열 자료의 안정성을 추구하는 방법이다. 항공운임예측의 경우 과거의 운임자료들이 미래의 운임에 지속적으로 영향을 미치고 있어 시계열 자료 분석시 MA 접근 방법을 고려할 수 있다.

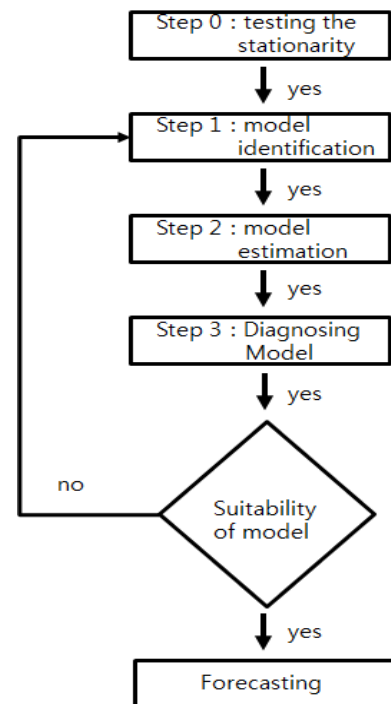
따라서 이러한 특성을 고려한 ARIMA모형의 추정과정은 시계열 패턴을 반영해 줄 수 있는 차수 p, d, q를 적절히 찾아내는 과정이라고 할 수 있다(Lee & Song, 2007).

ARIMA 모형의 특성을 수학적인 식으로 나타낼 때 사용하는 기호는 후향연산자(Backshift operator)와 차분연산자(differencing operator)로 일반적으로 후향연산자는 B로 나타내며 $B^k y_t = y_{t-k}$ 로 표기할 수 있다. 예를 들면 $B y_t = y_{t-1}$, $B^2 y_t = y_{t-2}$ 로 나타낼 수 있다. 차분연산자는 1-B나 '∇'로 나타낸다.

이러한 ARIMA 모형에 대해 Box et al. (1994)은 좋은 모형을 찾기 위해서 실전적인 3단계 절차를 제안하였다. 다음은 ARIMA 모형 설정을 위해 필요한 절차를 나타낸 것이다(Jung, 2009).

[0단계: 평균과 분산의 정상성 판단]

ARIMA 모형을 구축하기 위해서는 우선 주어진 시계열 자료에 대한 정상성(stationarity)을 만족시켜 주어야 한다. 여기에서 정상성(stationarity)이란 시계열을 일정한 주기로 나누었을 때 각 주기에 해당하는 평균과 분산이 일정하다는 의미로서 만약 시계열의 분산과 평균이 비정상적일 경우, 정상성을 만족시키기 위해서 각각 변수변환 및 차분을 취하여 분석 시계열 자료의 정상성을 판단해야 한다.



<Figure 1> ARIMA Model Building Process

[1단계: 모델 식별]

모델 식별단계에서는 시계열 내 관측값들 사이에 존재하는 상

관계를 측정하여 ARIMA(p,d,q) 모델을 구성하는 자기회귀(AR) 요소인 p와 이동평균(MA)요소인 q를 임시적으로 결정한다. 이 단계에서 상관관계를 측정하는 도구는 자기상관함수(Autocorrelation function, ACF)와 편자기상관함수(Partial Autocorrelation function, PACF)이다.

자기상관계수(ACF)는 시차가 증가함에 따라 싸인곡선이나 지수 곡선으로 0을 향해 점차적으로 감소하는 형태를 나타내며 이러한 시계열 패턴은 정상시계열임을 판단하는 근거가 된다.

자기회귀모형(AR Model)의 편자기상관계수(PACF)는 시차 p까지만 유의한 값이며, 나머지 시차에서는 신뢰구간 내에 있는 형태로 모형의 차수를 결정할 수 있다(Bowerman et al., 2005). 이러한 두 개의 상관함수를 이용하여 적절해 보이는 ARIMA 모형을 선택한 후 자료로부터 추정된 두 상관함수를 각각 비교하여 모형의 특성을 판정하게 된다. 모델식별 단계에서 적용한 모형은 최종 모형을 구축하기 위해서 추정, 모형진단 단계를 거쳐야 한다.

[2단계: 모형 추정]

모형추정 단계에서는 이전 단계에서 선택한 모형의 계수를 정확히 추정하여, 추정된 모수가 통계적으로 유의한지 여부를 판가름할 수 있기 때문에 1차적으로 모형의 적절성을 탐색할 수 있다. 특히, 추정된 계수들의 절대값 크기를 통해서 정상성(stationarity)과 가역성(invertibility)의 만족 여부를 판가름할 수 있다.

추정된 계수의 통계적인 판단은 t-통계량의 절대값이 2.0 이상이어야 한다. 또한 정상성과 더불어 ARIMA 모형이 만족해야 할 가역성 조건이 있다. 이는 MA프로세스와 AR프로세스가 결합된 모형에서만 고려되는 조건으로, 수학적으로는 자기회귀(AR) 계수에 관련된 정상성 조건과 동일한 특징을 가지고 있다. 하지만 MA 프로세스가 없는 q=0인 AR(p)모형은 가역성 조건을 점검할 필요가 없다. 따라서 가역성 조건을 점검할 모형은 가장 기본적인 모형인 ARIMA(p)에 이동평균(MA)이 반영된 ARMA 모형에서 점검해야 한다.

MA(1) 프로세스의 가역성 조건은 ϕ 의 절대값이 1보다 작아야 하며 MA(2) 이상의 가역성 조건은 $|\phi_2| < 1$, $\phi_2 + \phi_1 < 1$, $\phi_2 - \phi_1 < 1$ 이다(Box & Jenkins, 1976; Box et al., 1994; Jung, 2009).

ARIMA 모형이 가역성 조건을 충족하지 않는다는 의미는 먼 과거 관측값이 갖는 가중치와 예측시점에서 가까운 관측값이 갖는 가중치에 차이가 없다는 것을 의미한다. 그러나 기본적으로 단기 예측을 위한 ARIMA 모형은 최근에 가까운 관측 값일수록 가중값이 더 커져야 한다는 기본전제를 가정으로 구축된 특징이 있다. 따라서 가역성 조건이 만족되면 최근 시점에서 가까운 과거 관측 값일수록 더 큰 가중치를 가지고 현재 시점에 영향을 미친다는 것을 의미한다. 이러한 가역성 조건이 만족되면 다음과 같이 MA(1) 프로세스는 자기회귀(AR) 프로세스로 전환하여 다음과 같이 전개할 수 있다(Jung, 2009).

$$\tilde{y} = (1 - \theta_1 B) \epsilon_t \Rightarrow (1 - \theta_1 B)^{-1} \tilde{y} = \epsilon_t \quad (2)$$

$$\Rightarrow (1 + \theta_1 B + \theta_1^2 B^2 + \dots) \tilde{y} = \epsilon_t, |\theta_1| < 1 \quad (3)$$

$$y_t = C + \theta_1 y_{t-1} - \theta_1^2 y_{t-2} + \theta_1^3 y_{t-3} \dots \quad (4)$$

즉, ARMA모형을 기반으로 ARIMA모형을 구축한 경우 최종 구

축모형은 식(4)와 같이 AR 추정계수 값인 C에서 t-1기의 실제치와 시차계수를 연속적으로 곱하여 추정하게 된다.

여기서 추정된 모수가 유의하지 않을 경우, 다시 식별단계로 돌아가서 또 다른 ARIMA모형을 임시로 선정하여 재추정과정을 진행해야 한다.

[3단계: 모형 진단]

이 단계에서는 추정된 모형이 통계적으로 적절한지 여부를 결정하는 단계이며, 몇 가지 검정법에 근거하여 부합하지 않는 모형은 기각시킬 수 있다. 모형을 검정하는 단계에서 가장 중요한 검정은 백색잡음(white noise)의 독립성 가정에 대한 검정이라 할 수 있다. 백색잡음(white noise)들이 독립이라는 의미는 자기상관성이 없다는 의미이다. 백색잡음에 대한 독립성 여부와 더불어 잔차(residual)에 대한 자기상관성을 검정하는 방법과 Box-Ljung검정 방법들이 주로 이용 된다. 이러한 검정 도구들은 모형진단 단계에서 추정된 모형이 통계적으로 적합하지 않더라도 재식별 및 재추정 단계의 중요한 도구로 활용할 수 있다(Box et al., 1994; Li et al., 2005; Jung, 2009). 즉, 최종모형을 찾을 때까지 식별, 모형추정, 모형진단의 3단계를 되풀이하여 반복한다. 특히, 추정 및 모형검진단계를 통해 만족하는 모형을 찾을 때까지 재식별, 재추정, 재모형 진단을 계속 반복하여 실행한다.

2.3. ARIMA 예측값 추정

추정된 ARIMA 모형에 근거하여 얻은 예측 값인 부트스트랩 예측 값에 대한 추정 과정을 AR(1)모델로 제시하면 다음과 같다.

우선 시점 t를 현 시점이라 할 때 시계열 변수 y_{t+l} , $l \geq 1$ 에 대해 시점 t는 예측원점이라 하고, l을 예측 선행시차라고 하면 AR(1) 모형은 다음과 같이 제시할 수 있다.

$$(1 - \phi B) \tilde{y}_t = \epsilon_t \quad (5)$$

$$\text{단 } \tilde{y}_t = (y_t - \bar{y})$$

이를 기반으로 다시 예측원점으로 변환하면 다음과 같이 변환할 수 있다.

$$y_t = \mu(1 - \phi) + \phi_1 y_{t-1} + \epsilon_t \quad (6)$$

즉, 예측원점의 경우 AR추정상수인 $\mu(1 - \phi)$ 와 t-1기의 실제 값을 반영하여 예측하게 되며 t+1기간은 식(6)을 활용하여 $\mu(1 - \phi_1) + \phi_1 y_t + \epsilon_{t+1}$ 로 산출하게 된다.

여기서 예측시점이 t=3인 예측 값을 구하면 다음과 같이 나타낼 수 있다. $\mu(1 - \phi)$ 는 AR추정상수인 C로 나타낼 수 있으며 AR계수인 $\phi \times y_{t-2} - \theta_2 \times \epsilon_{t-2} + \epsilon_t$ 로 산출된다.

기본적인 AR(1)모형에 MA(5)가 반영된 예측 값 추정 예를 들어 설명할 경우 ARIMA(1,0,5)모델의 예측 값은 일련의 과정을 거쳐 식 (9)와 같이 제시할 수 있다.

$$\tilde{y}_t = (1 - \theta_1 B - \theta_2 B^2 - \theta_3 B^3 - \theta_4 B^4 - \theta_5 B^5) \epsilon_t \quad (7)$$

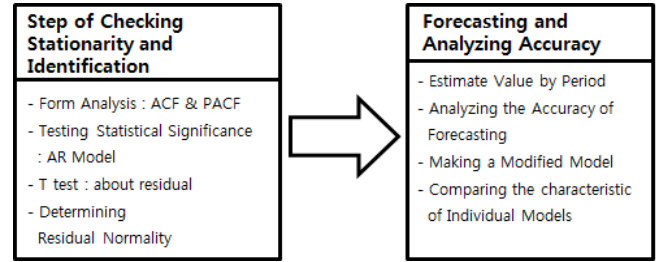
$$= (1 - \theta_5 B^5) \epsilon_t$$

$$(1 - \theta_1 B) \tilde{y} = (1 - \theta_5 B^5) \epsilon_t \quad (8)$$

$$\text{단 } \tilde{y}_t = (y_t - \bar{y})$$

$$y_t = C + \phi_1 y_{t-1} - \theta_5 \epsilon_{t-1} + \epsilon_t \quad (9)$$

식(9)을 통해 ARIMA모형을 분해하면 상수항인 C와 AR추정 값인 $\phi_1 y_{t-1}$, t-1기간의 오차를 차감하여 추정하는 형태로 구성되었음을 알 수 있다. ϵ_t 는 t기의 백색잡음(white noise)을 의미한다.



<Figure 2> Forecasting model building process

2.4. 예측의 정확도 평가기준

예측값의 정확도를 판정하는 도구는 누적예측오차(CFE), 평균절대오차(MAE), 평균제곱오차(MSE), 평균절대백분율오차(MAPE) 등 다양한 기준이 있지만 일반적으로 MAPE가 가장 많이 활용되고 있다(Wong et al., 2007; Burger et al., 2001).

이러한 MAPE 이점으로는 예측오차의 모형간 비교가 간단하고, 상대적인 오차크기를 직접적으로 고려하지 않고 상대적 비율로 비교하기 때문에 다른 측정기준에 비해 신뢰할 수 있는 이점이 있다. MAPE의 판정기준을 제시하면 다음과 같다(Lewis, 1982; Lee & Song, 2007).

- 0% ≤ MAPE < 10% - 매우 정확한 예측
- 10% ≤ MAPE < 20% - 비교적 정확한 예측
- 20% ≤ MAPE < 50% - 비교적 합리적 예측
- MAPE ≥ 50% - 부정확한 예측

3. 실증분석

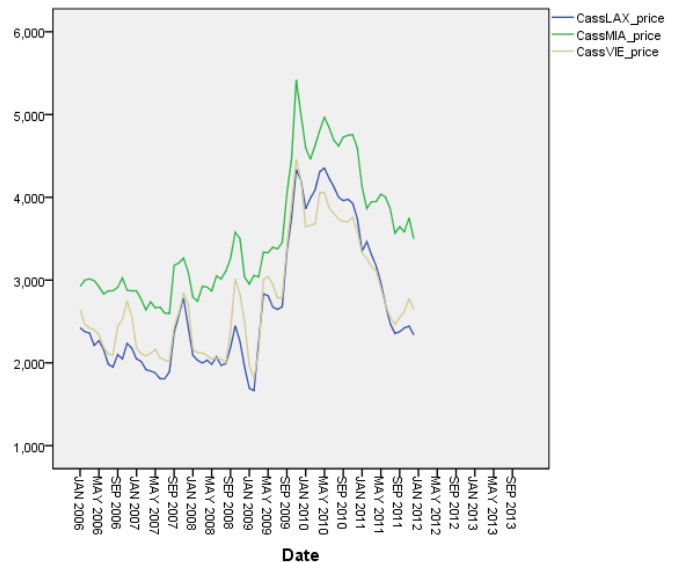
3.1. ARIMA모형의 구축 절차

IATA의 2006년부터 2011년간 LA, Miami, Vienna의 4개 노선에 대한 CASS 데이터를 이용하여 항공운임을 예측하기 위한 모형구축과정을 제시하면 다음과 같다. 우선 모형의 구축과정은 앞서 제시한 Box and Jenkins (1976)의 모형구축방법을 기반으로 현업의 담당자들이 실무적으로 활용 가능한 모형을 제시한다는 연구목적에 기반하여 기존 모형의 단점을 개선한 예측모형을 제시하고자 한다. 특히, ARIMA 모형을 적용하여 미래 시점을 예측할 경우 전문적인 소프트웨어를 사용하여 매번 기간을 선정하고 복잡한 분석작업을 수행하는 것은 현업의 실무자들에게 매우 어려운 작업일 수 있으며 이러한 복잡하고 난해한 과정을 생략하고 간단한 스프레드시트 프로그램을 이용하여 활용 가능한 모델을 제시하는 것이 실무적 활용성 측면에서 유용한 방법일 수 있다. 또한 ARIMA모델의 기간별 오차항에 해당하는 잔차(residual)나 백색잡음(white noise)에 대한 추정은 지극히 어려운 작업일 수 있다. 특히, 백색잡음(white noise)은 관측할 수 없는 값이지만 예측모형 구축시 자기상관성을 야기하며 예측값의 정확도를 낮추는데 많은 영향을 주어 자료에 대한 예측치의 적합도를 낮추게 한다. 따라서 본 연구에서는 기존 시계열 예측방법과 달리 실무적으로 이용 가능한 예측 모형 구축 프로세스를 제시하고자 한다. 즉, 기존 연구에서 제시된 ARIMA모형을 수정하여 실무적으로 쉽게 적용 가능한 운임 예측모형을 제시하고자 한다.

3.2. ARIMA 모형 추정

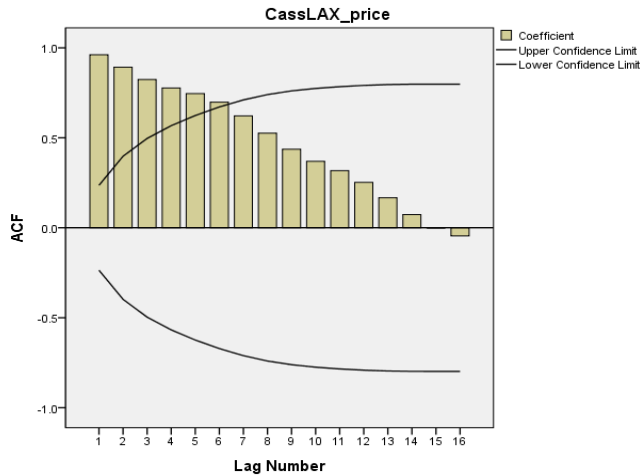
예측모형의 구축에 앞서 분석대상 시계열 자료의 특성을 분석하고자 한다. 분석 시계열 자료는 LA, Miami, Vienna의 2006년부터 2011년까지의 항공운임에 대한 자료이다.

6년간(72개월)의 월별 데이터를 기반으로 3개 노선에 대한 자료를 비교한 결과 MIA의 항공운임이 다른 3개 노선에 비해 전반적으로 높게 나타나고 있다. 전반적으로 전체 시계열곡선의 특성을 볼 때 추세나 계절성이 나타나고 있지 않은 것으로 판단된다. 따라서 차분인 d=0인 ARIMA(p) 혹은 ARIMA(p,q)모형을 고려해야 한다.

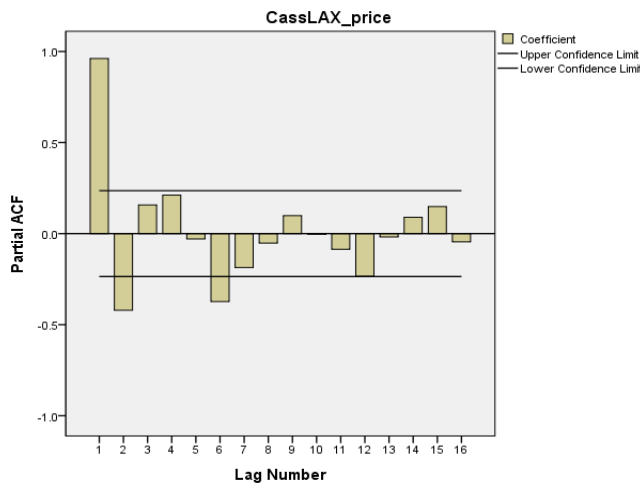


<Figure 3> Bar graph of air fares on 3 routes

개별 노선의 예측모형을 구축하기 위해 첫 번째로 LA의 노선의 항공운임 특성에 대한 분석을 실시하였다. 우선 시간변화에 따른 항공운임의 자기상관성(ACF)과 편자기상관함수(PACF)를 비교하고자 한다. 항공운임의 자기상관함수(ACF)와 편자기상관함수(PACF)는 과거의 항공운임이 미래의 특정시점에 영향을 주는지에 대한 여부를 판단하는 방법이다. 이를 위해 가장 기본적인 모형이라 할 수 있는 ARIMA(1)모형을 통해 관측치들의 특성을 분석하고자 한다. 세부적인 분석에 앞서 상관계수들(ACF, PACF)에 대한 신뢰구간을 비교하는 방법을 통해 시계열 자료의 특성을 파악하였다.



<Figure 4> ACF of LA air fares



<Figure 5> PACF of LA air fares

ARIMA(1)모형을 통해 LA노선에 대한 자기상관함수(ACF)를 분석한 결과 7개월 이후 전체 상관함수가 모두 95% 신뢰구간 안에 포함되는 것으로 나타났다. 또한 월별 운임의 변화가 두드러지게 반복되는 패턴은 없는 것으로 나타나 정상성 조건에 해당되어 차분계수인 d 를 0으로 결정하였다($d=0$). 특히, 자기상관함수(ACF)의 변화량을 분석한 결과 제공하는 형태로 감소하는 패턴이 없는 것으로 나타났다. 따라서 월별 시차를 발생시키기 위해 이동평균인 MA(7)을 고려하였다.

자기상관함수(ACF)와 편자기상관함수(PACF)를 사용하여 자기회귀요소(p)와 이동평균요소(q)를 결정하게 되는데 PACF분석에서도 나타나듯이 6개월 이후 안정적으로 95% 신뢰구간 안에 PACF가 모두 포함되고 있다. 따라서 ARIMA 모형 구축시 ARIMA(1,0,7)모형을 분석하고자 한다.

LA노선위 항공운임에 대한 ARIMA(1)모형과 ARIMA(1,0,7)모형을 분석하여 구축모형의 통계적 유의성을 검정하였다.

우선 구축된 ARIMA(1) 모형의 경우 잔차에 대한 ACF와 PACF의 t-통계량 절대값이 단기시차(1, 2, 3)에서는 1.25 이하, 그 밖의 시차에서는 1.6보다 작아야 한다(Box et al., 1994; Jung, 2009).

분석결과 기간1, 2, 3, 6의 경우 잔차들이 서로 상관되어 있어 오차항에 대한 독립성이 인정되지 않는 것으로 나타났다. 즉, 각 기간별로 통계적인 상관관계가 있어 백색잡음(white noise)이 독립

적이라는 귀무가설이 기각되었음을 의미한다. 따라서 이동평균이 반영된 새로운 모델을 추정해야 한다.

<Table 1> T-test of ACF and PACF

	Period 1	Period 2	Period 3	Period 4	Period 5	Period 6
ACF	3.71	0.09	1.91	1.25	1.62	2.71
PACF	3.71	1.88	1.92	0.54	3.30	1.06

<Table 2> Calculation air fares using AR(1)

Period	Actual Fare	Forecast Fare
2006/1	2425	2575.6
2006/2	2376	$2431.9 = 118.48 + 0.954 \times 2425$
2006/3	2361	$2385.1 = 118.48 + 0.954 \times 2376$
2006/4	2211	$2370.9 = 118.48 + 0.954 \times 2361$

LA노선에 대한 ARIMA(1) 모형은 통계적으로 적합하지 않아 다른 모형을 추정해야 하지만 시계열 분석의 가장 기본적 모형인 ARIMA(1)모형을 통해 예측값이 어떻게 산출되는지에 대해 분석하였다. ARIMA(1) 모형의 기간별 예측값을 산출하기 위한 분석결과는 다음과 같다.

AR시차계수는 0.954, 상수항은 2575.6로 나타났으며 유의확률인 $p < 0.01$ 이하로 나타났다. 이러한 분석결과를 통해 ARIMA(1)의 예측값을 수식으로 변환하여 나타내면 다음과 같다. 우선 식(6)에 제시한 것처럼 AR상수는 $2575.6 \times (1 - 0.954) = 118.48$ 로 계산되며 초기 예측치는 AR(1)의 상수항에 해당된다.

2006년 1월부터 3월까지의 실제치와 AR(1) 모델을 통해 예측한 값을 정리하면 다음과 같다.

분석표본의 경우 앞 기간의 가격정보를 통해 예측할 수 있지만 예측구간에서는 앞 기간의 실제운임 정보를 모르기 때문에 앞 기간의 예측치를 활용하여 차기의 운임을 예측하게 된다.

앞서 제시한 <Figure 4>, <Figure 5>에 나타난 것과 같이 ARIMA(1)모형에서 6개월의 자기상관성이 발생하고 있어 ARIMA(1,0,7) 모형을 제시하여 통계적인 모델 검정을 실시하였다.

ARIMA모형의 통계적 검정은 백색잡음(white noise)에 대한 유의성 검정을 통해 시계열 자료에 대한 정상성을 평가한 후 자기상관성을 분석하여 모델을 구축하는 방법을 적용할 수 있다.

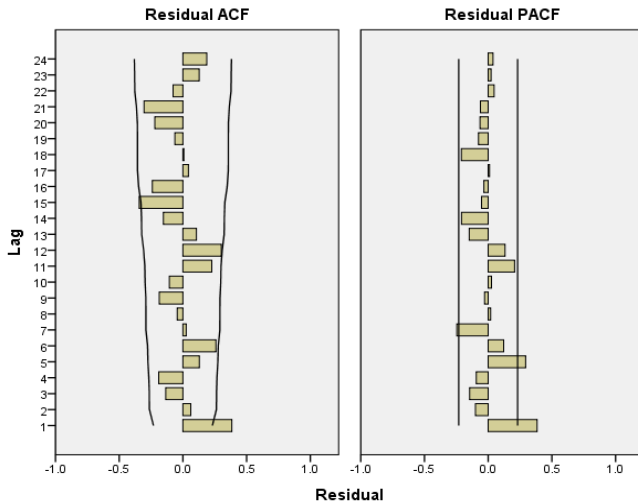
여기서 백색잡음(white noise)에 대한 가설은 " H_0 : 백색잡음(white noise)항은 독립이다." 라는 귀무가설을 검정하게 되는데 분석방법은 Box-Ljung Q통계량을 활용하게 된다.

백색잡음(white noise)에 대한 독립성 여부를 평가한 후 자기상관성을 분석하기 위한 방법으로 자기상관함수(ACF)와 편자기상관함수(PACF)의 형태가 신뢰구간 밖으로 벗어난 스파이크가 존재하는지를 점검한 후 t-통계량을 이용하여 절대값이 '2'이상 나타난 것이 존재하는지를 분석하였다.

ARIMA(1,0,7)의 모형통계량 분석결과 MAPE는 5.94, Box-Ljung Q통계량(Ljung & Box, 1978)은 58.820($df=16$, $p=0.01$)로 p 값이 0.05보다 작게 나타나 "백색잡음(white noise)항이 독립이다." 라는 귀무가설(H_0)은 5% 유의수준에서 기각되었다. 즉, 기간별 자기상

관성이 존재하는 것으로 나타났다. 백색잡음(white noise)에 대한 독립성 여부를 판단한 후 이동평균인 MA(7)을 반영한 모형의 자기상관성을 비교한 후 관측데이터의 특성을 분석하였다.

<Figure 6>은 잔차(residual)에 대한 자기상관함수(ACF)와 편자기상관함수(PACF)를 도표로 제시한 것으로 신뢰구간을 벗어난 스파이크가 여러 기간에 걸쳐 나타나고 있다. 이러한 특징을 t-통계량 분석을 통해 명확하게 분석할 수 있는데 t-통계량은 각 상관함수를 표준오차로 나누고 절대값을 취하여 구할 수 있다.



<Figure 6> ACF and PACF on residual of ARIMA(1,0,7) model

우선 <Figure 6>에 나타난 것과 같이 기간1, 기간5에서 신뢰구간을 벗어난 스파이크가 발견되었다. 또한 특정 기간에서 t-통계량의 절대값이 '2'이상을 초과하는 기간이 나타나 모형구축에 있어 여전히 자기상관성이 발생하는 것으로 나타났다.

기간 7까지의 t-통계량의 절대값은 <Table 3>과 같으며 t-통계량이 유의하지 않은 구간은 *로 제시하였다.

<Table 3> T-test of ACF and PACF

	Period1	Period2	Period3	Period4	Period5	Period6	Period7
ACF	3.26*	0.46	-1.00	-1.39	0.92	1.83	0.17
PACF	3.26*	-0.85	-1.23	-0.79	2.49*	1.03	-2.09*

하지만 보수적인 기준으로 판정시 t-통계량의 절대값이 '2'이상으로 나타나도 비교대상 모형과의 차이를 분석하여 모형을 구축한 Li et al. (2005)의 연구와 같이 본 연구에서도 전체 모형을 구축하기 위한 과정인 점과 또 다른 개선 모형과의 비교를 통해 최종 모형을 결정하기 위해 ARIMA(1,0,7) 모형을 활용하고자 한다.

정상성 조건에 대한 분석에 이어 AR모형과 이동평균인 MA(n)이 반영된 시차모델의 경우 추가적으로 가역성 조건을 만족시켜야 한다. ARIMA(1,0,7)모형의 분석결과 AR시차계수($\hat{\phi}_1$) = 0.94, MA시차계수($\hat{\theta}_7$) = -0.442로 나타나 $\hat{\phi}_1$ 와 $\hat{\theta}_7$ 의 절대값이 모두 1보다 작은 것으로 나타났으며 $\hat{\phi}_1$ 와 $\hat{\theta}_7$ 의 t-통계량의 절대값이 '2'보다 크고 유의확률인 p값이 0.01 이하로 나타나 통계적으로 모두 유의한 것으로 판정하였다(Box et al., 1994; Kulendran & Witt, 2003; Jung, 2009). 전체 분석결과는 다음과 같다.

<Table 4> Parametric analysis result of ARIMA(1,0,7)

	Estimate Values	SE	t	p
Constant Term	2580.53	458.925	5.623	0.01
AR	0.94	0.038	24.558	0.01
MA	-0.442	0.119	-3.708	0.01

이한 과정을 거쳐 구축된 ARIMA(1,0,7)모형의 특징을 종합하면 우선적으로 시차에 대한 상관성을 들 수 있다. 즉, 시차를 발생하여 ARIMA모형을 구축하였음에도 불구하고 지속적으로 특정 기간에 자기상관성이 발생하는 있어 관측치에 대한 적합도를 낮추고 있다. 이러한 연속적인 자기상관성이 나타나는 원인에 대해 분석하면 항공운임 결정에 대한 데이터의 특수성에 기인하는 것으로 볼 수 있다. 즉, 기간을 7개월 이동시켜도 과거의 운임이 미래의 운임에 영향을 지속적으로 미치고 있어 특정기간에 대한 시차를 발생시켜도 운임에 대한 자기상관성이 발생하는 것으로 판단된다. 즉, 12개월의 운임은 단순히 7개월의 시차를 발생시켜도 항공운임 예측시 과거의 운임이 지속적으로 미래운임 결정에 관여한다는 의미로 해석할 수 있다. 따라서 시차를 특정구간에 대해서만 발생시키는 전통적인 ARIMA(p,q)모형과 달리 개별 기간들의 운임을 연속적으로 7기간까지 시차를 발생시켜 자기상관성을 상쇄시킬 필요가 있다. 따라서 본 연구에서는 특정 기간만 시차를 발생시키는 방법이 아니라 설정된 기간별로 과거의 개별 기간이 다음 기간에 연속적으로 영향을 미치는 방법을 제안하고자 한다. 즉, 연속적으로 과거운임이 미래의 운임에 영향을 미치는 현실적인 특성이 반영된 모형을 제시하고자 한다. 이러한 특징을 좀 더 구체화시켜 식으로 제시하기 위해 앞서 제시한 모형인 ARIMA(1,0,5)모형인 식 (7), (8), (9)에 연속적으로 영향을 미치는 방법을 적용하여 제시하면 다음과 같다.

$$\tilde{y}_t = 1 - [\theta_1 B + (\theta_1 B + \theta_2 B^2) \quad (10)$$

$$+ (\theta_1 B + \theta_2 B^2 + \theta_3 B^3) + (\theta_1 B + \theta_2 B^2 + \theta_3 B^3 + \theta_4 B^4)$$

$$+ (\theta_1 B + \theta_2 B^2 + \theta_3 B^3 + \theta_4 B^4 + \theta_5 B^5)] + \epsilon_t$$

$$= (1 - 5\theta_1 B - 4\theta_2 B^2 - 3\theta_3 B^3 - 2\theta_4 B^4 - \theta_5 B^5) \tilde{y} \quad (11)$$

$$\text{단 } \tilde{y}_t = (y_t - \bar{y})$$

$$y_t = C + \phi_1 y_{t-1} \quad (12)$$

$$- (5\theta_1 \epsilon_{t-1} + 4\theta_2 \epsilon_{t-2} + 3\theta_3 \epsilon_{t-3} + 2\theta_4 \epsilon_{t-4} + \theta_5 \epsilon_{t-5}) + \epsilon_t$$

일반적으로 ARIMA(p,q)모형은 시차가 발생할수록 지수적으로 감소하는 자기상관함수의 특징이 있는데 이러한 특징에 연속적으로 과거 기간의 영향정도를 반영하는 방법을 적용하였다. 따라서 기존의 ARIMA(p,q)모형과 구분하고 시차가 연속적으로 다음 기간에 영향을 미치는 'ARIMA with Continuous MA'의 특징을 반영하여 ARIMAC(1,0,7)로 정의하고 모형을 구축하였다.

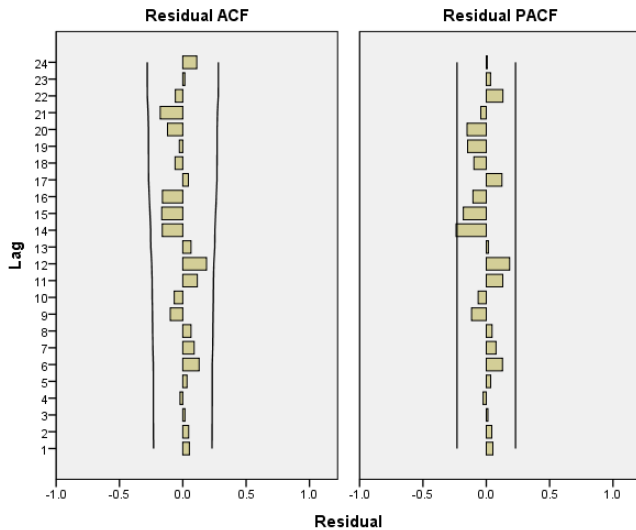
구축된 모형을 기반으로 ARIMA(1,0,7)과 ARIMAC(1,0,7) 모형을 비교하였다.

우선 ARIMAC(1,0,7)의 정상성 판정시 분석하는 백색잡음(white noise)에 대한 분석결과 Ljung-Box Q통계량의 p값이 0.05보다 크

게 나타나 백색잡음(white noise)에 대한 독립성이 인정되었다. 또한 잔차에 대한 자기상관성을 분석한 결과 <Figure 7>에 나타난 것처럼 95% 신뢰구간에 전 기간의 모든 상관계수가 포함되는 것으로 나타났으며 t-통계량의 절대 값이 모두 매우 낮게 나타났다. 즉, ARIMAC(1,0,7)의 경우 정상성 조건을 모두 충족하는 것으로 나타났다.

<Table 5> T-test of ACF and PACF using ARIMAC(1,0,7) model

Accuracy	Ljung-Box Q		
MAPE	Statistics	df	p
5.041	17.068	10	0.073



<Figure 7> Residual analysis of ARIMA(1,0,7) model

<Table 6> T-test of ACF and PACF using ARIMAC(1,0,7) model

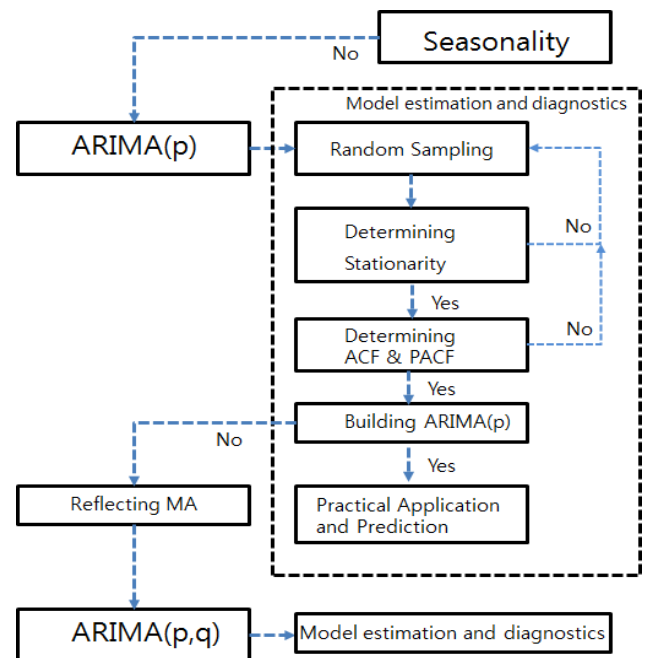
A	Period	1	2	3	4	5	6
	t-value	0.13	0.03	-0.25	-0.20	0.09	0.74
C	Period	7	8	9	10	11	12
	t-value	-0.03	-0.11	-1.33	-1.09	0.45	1.68
F	Period	1	2	3	4	5	6
	t-value	0.13	0.03	-0.25	-0.19	0.10	0.74
C	Period	7	8	9	10	11	12
	t-value	-0.06	-0.12	-1.31	-1.09	0.50	1.76

다음은 ARIMAC(1,0,7)의 가역성 조건을 통해 모델의 특성을 판정하였다. ARIMAC(1,0,7)의 경우 연속적인 시차를 발생시켜 7개 기간의 시차계수인 θ 가 7개 나타나는데 이러한 시차계수의 절대 값은 0.154~0.439 사이에 나타나고 있어 가역성 조건을 모두 만족하는 것으로 나타났다.

<Table 7> Reversibility condition test of ARIMAC(1,0,7)

MA	Estimate Values(θ)	Absolute Value
시차 1	-0.439	0.439
시차 2	-0.148	0.148
시차 3	0.135	0.135
시차 4	0.219	0.219
시차 5	-0.154	0.154
시차 6	-0.200	0.200
시차 7	-0.239	0.239

지금까지 LA노선의 항공운임에 대한 ARIMA모형구축 과정을 요약하였다 이러한 분석절차를 기반으로 효율적인 ARIMA모형 구축을 위한 실무적 프로세스를 제시하고자 한다. 이는 현업의 실무자들에게 ARIMA모형 구축시 참고하기 위한 표준적인 가이드라인을 제시하는데 있다.



<Figure 8> Practical process of ARIMA model

ARIMA 모형 구축시 다른 분석보다 가장 우선적으로 수행해야 할 부분은 관측 데이터에 계절성이 있는지를 판정하는 부분으로 LA를 비롯한 Miami, Vienna 노선의 경우 계절성이 나타나지 않았다. 따라서 비계절성 ARIMA 모형을 구축을 위해 임의로 모형을 선정한 후 통계적인 분석과정을 실시하여 모형을 구축하게 된다.

임의로 모형을 선정할 때에는 일반적으로 가장 기본이 되는 ARIMA(1)모형을 기준으로 분석하여 통계적인 분석과정을 거치게 되는데 분석방법을 나타내면 우선 정상성 분석을 위해 잔차에 대한 독립성을 판정하기 위한 방법으로 Box-Ljung의 통계량을 통해 분석할 수 있다.

독립성에 대한 가설검정을 통해 정상성을 분석한 후 잔차에 대한 ACF와 PACF에 대한 검정을 실시한다. 이때 잔차도표와 함께 t-통계량을 활용하여 분석할 수 있다. 이러한 일련의 과정을 거쳐 모형을 구축한 결과가 만족스럽지 못할 경우 이동평균인 MA(n)을 반영하여 ARIMA(p,q)모형을 구축해야 한다. 그 다음 앞서 제시한

통계적인 도구를 이용하여 모형을 적합시켜야 한다. 특히 MA(n) 값을 수정하여 모형을 적합시키기 위한 일련의 과정을 반복해야 한다.

ARIMA모형을 분석하기 위한 실무적 프로세스와 더불어 ARIMA(1), ARIMA(1,0,7), ARIMAC(1,0,7)의 특징을 비교하고자 한다. 특히, ARIMAC모형의 경우 앞서 분석한 ARIMA(1,0,7)의 문제점으로 나타난 자기상관성의 문제를 무력화시켜 잔차의 독립성을 확보하였지만, 실제 이러한 분석방법을 스프레드시트로 계산하는 것은 지극히 어려운 문제이다. 특히 기간별 예측 값을 산출할 때 부트스트랩에 의한 예측 값을 통해 평균과 신뢰구간을 자동적으로 산출하게 되는데 이러한 부트스트랩 예측 값은 반복적인 연산과정을 통해 오차 항에 대한 추정치를 산출하고 있다. 또한 ARIMAC의 경우 연속적으로 시차에 대한 연산을 반복하기 때문에 ARIMA(p,q)모형보다 더 복잡한 과정을 거치게 된다. 즉, ARIMAC(p,q)의 경우 모형의 설명력은 좋지만 산출과정이 기존 모형에 비해 매우 어렵다는 문제점이 있다. 이러한 측면에서 LA의 항공운임에 대한 3개 모형의 특징을 요약하면 다음과 같다.

<Table 8> Comparison with ARIMA(1), ARIMA(1,0,7), and ARIMAC(1,0,7) model

	ARIMA(1,0,0)	ARIMA(1,0,7)	ARIMAC(1,0,7)
ACF&PACF	Not Satisfied	Not Satisfied Partially	Satisfied
Residual Analysis	Not Satisfied	Not Satisfied	Satisfied
T-value	Not Satisfied	Not Satisfied Partially	Satisfied
Condition of Stationarity	Not Satisfied	Not satisfied	Satisfied
Condition of Reversibility	Unnecessary	Not Satisfied	Satisfied
MAPE	5.982	5.288	5.041
Characteristics	Easier to Calculating Forecasting Value	Relatively Difficult to Calculating Forecasting Value	Very difficult to Calculating Forecasting Value

<Table 8>에 나타난 것처럼 ARIMAC모형이 전반적으로 우수한 것으로 나타났지만 ARIMA(1,0,0)과 ARIMA(1,0,7)의 경우 모델의 기본적 추정과 진단단계를 통과하지 못해 MAPE의 추정은 의미가 없다고 할 수 있다. 따라서 <Table 8>를 기반으로 연구모형을 제시하기 위해 본 연구에서는 ARIMAC(p,q)를 기반으로 하고 실무적으로 쉽게 적용하기 위한 방법을 제안하였다.

연구모형의 구축에 대한 방향을 제시하면 우선 실무적 적용의 용이성과 구축된 모형을 기반으로 수요변화에 따라 변형 가능한 적응적인(adaptive) 특징을 고려하였다. 즉, 간단한 스프레드시트 프로그램을 이용하여 예측모형을 구축할 수 있으며 구축된 모형의 시차계수를 적절히 조정하여 관측값의 적합도를 높일 수 있는 이점이 있다.

본 연구에서 제시한 연구모형은 ARIMAC모형처럼 예측값과 오차항에 대한 부트스트랩 예측값을 활용하지 않고 기간별 ARIMAC 모형과 실제값의 차이에 대한 오차와 MA시차계수를 반영하여 산

출하였다. MA의 시차계수는 기간에 영향 정도를 나타내는 가중치 역할을 수행하고 있어 실제 수요의 변화에 따라 적절히 변형하여 활용할 수 있다.

4. 연구모형

4.1. 연구모형의 구축

연구모형의 제시에 앞서 본 연구모형의 근간이 되는 ARIMAC(1,0,7)모형을 수식으로 제시하면 다음과 같이 제시할 수 있다.

$$y_t = C + \phi_1 y_{t-1} - (5\theta_1 \epsilon_{t-1} + 4\theta_2 \epsilon_{t-2} + 3\theta_3 \epsilon_{t-3} + 2\theta_4 \epsilon_{t-4} + \theta_5 \epsilon_{t-5}) + \epsilon_t \quad (13)$$

ARIMAC모형의 경우 각 기간의 월별시차인 θ 와 앞 기간의 오차를 곱하여 AR예측 값에서 기간별 오차를 차감하는 방법으로 계산할 수 있다. 또한 각 기간별 오차항을 계산하여 예측치에 반영하는 방법을 통해 예측 값을 추정한다.

이러한 반복적인 계산과정의 복잡함으로 인해 본 연구에서는 실무적으로 오차를 쉽게 반영하기 위한 방법을 제시하여 연구모형으로 제시하고자 한다.

우선 연구모형의 근간이 되는 모형은 ARIMAC(1,0,7)로 2006년부터 2011년의 72개월에 해당하는 월별 백분율오차를 산출하여 각 월별로 6개년의 평균을 산출한다. 즉, 연구모형은 ARIMAC모형의 오차를 보정하기 위한 방법으로 6개년의 1월부터 12월까지 월별 백분율오차의 평균을 추정하여 AR예측치에서 차감하는 방법을 통해 오차를 보정하는 방법을 적용하였다. 이러한 특성으로 인해 연구모형은 Error adjusted ARIMAC로 정의하여 연구모형을 제시하였다. 연구모형인 Error adjusted ARIMAC를 수식으로 제시하면 다음과 같다.

$$y_t = C + \phi_1 y_{t-1} - \left(\frac{\sum_{j=1}^6 \theta_j PE_{jt}}{6} \right) \quad (14)$$

j : 예측연도 ($j = 1, 2, 3, \dots, 6$)

θ_i : 월별 MA시차 계수 ($i = 1, 2, 3, \dots, 7$)

PE_t : t 월의 백분율 오차 ($t = 1, 2, 3, \dots, 12$)

Error adjusted ARIMAC에서 활용한 6개년(72개월)의 월별 백분율오차를 나타내면 다음과 같다.

<Table 9> Monthly percentage error of research model

Period	1	2	3	4	5	6
PE	-0.057	0.015	-0.001	-0.023	-0.011	-0.017
Period	7	8	9	10	11	12
PE	-0.026	0.002	0.084	0.016	0.023	-0.063

앞서 제시한 ARIMAC모형을 기반으로 실무적인 활용성을 고려한 연구모델인 Error adjusted ARIMAC를 제시하였다.

식(13)과 <Table 9>를 활용하여 Error adjusted ARIMAC의 기간별 예측치를 산출하는 과정을 제시하였다. 우선 연구모형에서 제시된 $y_t = C + \phi_1 y_{t-1}$ 산출식은 앞서 제시한 ARIMAC(1,0,7)모형에서 제시한 결과값으로 상수항은 2602.633이고 AR계수는 0.921, AR상수인 C는 $2602.633 \times (1 - 0.921)$ 의 결과인 205.608로 나타낼 수 있다. 또한 월별 MA시차계수인 θ 는 <Table 5>에 제시하였다. 이를 기반으로 2006년 2월 이후의 예측치 계산과정을 제시하면 다음과 같다. 2월의 실제 가격은 2431.9이다.

1) AR계산:

$$t_2 = 205.608 + (1 - 0.921) \times t_1 \text{실제가격}(2575.6) = 2439$$

2) 오차계산

$$e_2 = \text{ARIMAC } t_1 \text{ 예측치} \times 1 \text{월PE} \times \text{MA1계수}(-0.439) + \text{ARIMAC } t_2 \text{ 예측치} \times 2 \text{월PE} \times \text{MA2계수}(-0.148) = 59.04$$

3) 예측치

$$\text{AR계산} - \text{오차합} = 24.39 - 59.03 = 2380$$

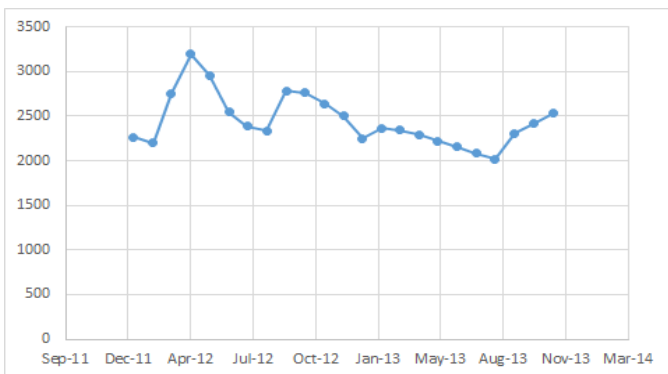
분석표본(training set)을 통해 기간별 예측을 수행할 경우는 앞기간의 실제 값을 활용하여 가격을 예측하지만 예측구간에서는 실제 가격정보가 없어 앞기간의 예측정보를 활용하게 된다. 즉, 예측구간이 일종의 검증표본(validation set) 역할을 수행하게 된다. 이러한 특징은 ARIMA 모형의 기본적인 특성이다.

4.2. ARIMAC와 Error adjusted ARIMAC의 예측력 비교

연구모형인 Error adjusted ARIMAC는 ARIMAC를 기반으로 예측 오차를 보정한 모형이다. 즉, Error adjusted ARIMAC모형은 실무적인 활용성에 초점을 맞춘 ARIMAC모형이라 할 수 있는데 반복적인 연산과정의 복잡성을 비교적 간단하게 계산하여 예측치를 추정할 수 있으며 실무적으로 활용도가 높다고 할 수 있다.

모형의 실무적 활용성과 더불어 예측의 정확도는 예측모형 구축시 매우 중요한 평가요소이다 따라서 ARIMAC모형과 Error adjusted ARIMAC모형의 정확도를 비교하였다.

모형의 예측구간은 2012년 1월부터 2013년 11월까지를 대상으로 비교하였다. 예측의 정확도는 평균절대백분율오차인 MAPE를 이용하였다.



<Figure 9> Prediction interval of air fares (LA route -2012/1 ~ 2013/11)

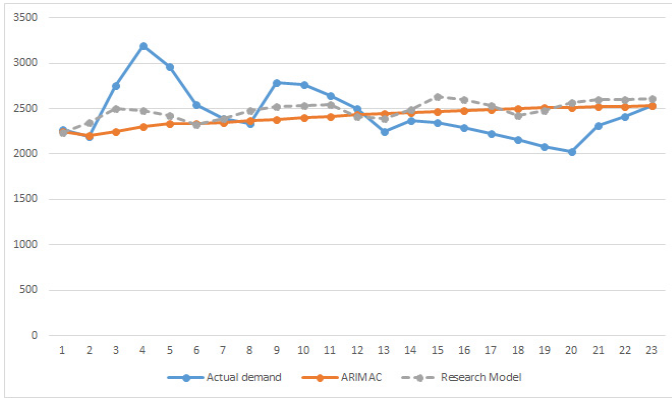
<Table 10>의 e1, e2는 ARIMAC의 오차(e1)와 연구모형인 Error adjusted ARIMAC의 오차(e2)이며 M1, M2도 ARIMAC와 Error adjusted ARIMAC의 절대백분율오차(APE)이다. LA노선의 예측구간의 항공운임은 서서히 감소하는 추세가 나타나고 있다.

<Table 10> Comparison in prediction accuracy (LA route air fare)

Period	Fare	ARIMAC	Research Model	e1	e2	M1	M2
2012/1	2265	2251	2240	14	25	0.6	1.1
2012/2	2197	2200	2351	-3	-154	0.1	7
2012/3	2756	2248	2501	508	255	22.4	9.3
2012/4	3199	2303	2478	896	721	39.6	22.5
2012/5	2959	2331	2422	628	537	27.7	18.2
2012/6	2546	2336	2327	210	219	9.3	8.6
2012/7	2387	2343	2386	44	1	2	0.1
2012/8	2337	2363	2476	-26	-139	1.1	6
2012/9	2785	2382	2520	403	265	17.8	9.5
2012/10	2760	2399	2530	361	230	15.9	8.3
2012/11	2642	2415	2539	227	103	10	3.9
2012/12	2505	2430	2410	75	95	3.3	3.8
2013/1	2,252	2444	2391	-192	-139	8.5	6.2
2013/2	2,363	2456	2491	-93	-128	4.1	5.4
2013/3	2,342	2468	2635	-126	-293	5.6	12.5
2013/4	2,292	2478	2601	-186	-309	8.2	13.5
2013/5	2,223	2488	2531	-265	-308	11.7	13.8
2013/6	2,159	2497	2423	-338	-264	14.9	12.2
2013/7	2,086	2506	2475	-420	-389	18.5	18.7
2013/8	2,022	2513	2563	-491	-541	21.7	26.8
2013/9	2,308	2520	2601	-212	-293	9.4	12.7
2013/10	2,416	2527	2604	-111	-188	4.9	7.8
2013/11	2,529	2533	2608	-4	-79	0.2	3.1

ARIMAC모형과 개선된 ARIMAC모형인 Error adjusted ARIMAC 모형의 MAPE를 비교한 결과 ARIMAC는 11.2%로 나타났으며 연구모형인 Error adjusted ARIMAC모형은 10%로 근소하지만 Error adjusted ARIMAC모형의 정확도가 높게 나타나고 있다. 또한 예측구간의 기간별 예측운임에 대해 두 모형을 비교하면 모형 간 특성을 좀 더 명확하게 알 수 있다.

<Figure 10>에 나타난 것처럼 예측구간에 대한 ARIMAC모형은 거의 직선에 가까운 선형성을 나타내고 있지만 실제 가격은 변동성이 크게 나타나고 있다. 하지만 연구모형인 Error adjusted ARIMAC모형의 경우 ARIMAC모형보다 좀 더 탄력적으로 실제 가격곡선에 일치하는 것으로 나타났다. 하지만 연구모형인 Error adjusted ARIMAC모형은 ARIMAC모형을 기반으로 추정된 모형이기 때문에 변동하는 특성이 유사한 형태를 나타내고 있다.



<Figure 10> Comparison of forecasting graph between ARIMAC and research model (LA)

4.3. 노선별 운임 예측모델 구축

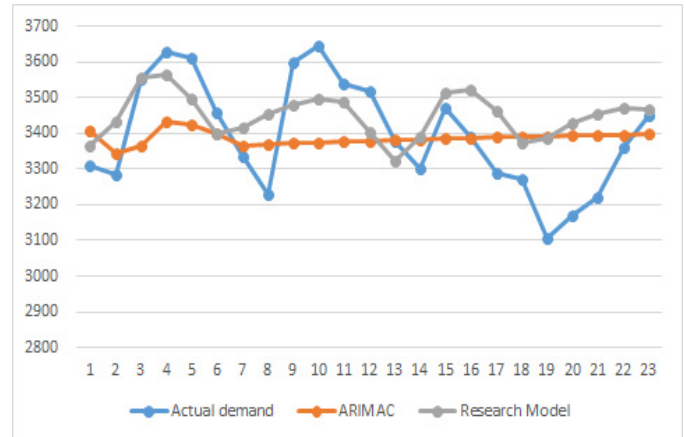
노선별 항공운임을 예측하기 위해 Miami 노선의 운임을 분석하였다. 분석결과 앞서 제시한 LA와 동일한 ARIMAC(1,0,7) 모델로 나타났다. LA와 같이 동일한 북미지역에 속해 운임의 변동성이 유사한 것으로 보인다. 앞서 분석한 LA와 동일한 방법으로 Miami 노선에 대해 분석한 결과를 제시하면 다음과 같다.

Miami노선에 대해 ARIMAC(1,0,7)모형에서 도출한 상수항은 2411.921이고 AR계수는 0.928, AR상수인 C는 $3411.921 \times (1-0.928)$ 의 결과인 245.66으로 나타났다. 2012년 1월부터 2013년 11월까지 예측구간에 대해 정확도를 분석하였다.

<Table 11> Comparison in prediction accuracy (Miami route air fare)

Period	Fare	ARIMAC	Research Model	e1	e2	M1	M2
2012/1	3311	3407	3366	-96	-55	2.9	1.7
2012/2	3283	3343	3431	-60	-148	1.8	4.5
2012/3	3552	3366	3557	186	-5	5.2	0.1
2012/4	3628	3433	3567	195	61	5.4	1.7
2012/5	3613	3426	3499	187	114	5.2	3.2
2012/6	3458	3401	3400	57	58	1.7	1.7
2012/7	3336	3365	3416	-29	-80	0.9	2.4
2012/8	3229	3368	3455	-139	-226	4.3	7.0
2012/9	3597	3372	3479	225	118	6.3	3.3
2012/10	3644	3374	3497	270	147	7.4	4.0
2012/11	3539	3377	3490	162	49	4.6	1.4
2012/12	3518	3380	3403	138	115	3.9	3.3
2013/1	3380	3382	3324	-2	56	0.1	1.6
2013/2	3303	3384	3391	-81	-88	2.5	2.7
2013/3	3470	3386	3513	84	-43	2.4	1.2
2013/4	3391	3388	3524	3	-133	0.1	3.9
2013/5	3289	3390	3464	-101	-175	3.1	5.3
2013/6	3270	3391	3372	-121	-102	3.7	3.1
2013/7	3108	3393	3386	-285	-278	9.2	9.0
2013/8	3169	3394	3427	-225	-258	7.1	8.2
2013/9	3220	3395	3453	-175	-233	5.4	7.2
2013/10	3361	3397	3473	-36	-112	1.1	3.3
2013/11	3452	3398	3467	54	-15	1.6	0.4

예측구간에 대한 정확도를 분석한 결과 ARIMAC모형은 MAPE가 3.7%로 나타났으며 연구모형인 Error adjusted ARIMAC모형은 3.5%로 나타나 LA노선처럼 연구모형의 예측력이 근소한 차로 높게 나타났다. 예측구간에 모형별 예측곡선을 제시하여 예측 값에 대한 특성을 비교한 결과 다른 지역에 비해 실제 운임의 변동성이 매우 높은 것으로 나타났으며 연구모형이 ARIMAC모형보다 민감하게 변동하고 있는 것으로 나타났다.



<Figure 11> Comparison of forecasting graph between ARIMAC and research model (Miami)

Miami 노선에 이어 유럽지역인 Vienna 노선의 항공운임에 대해 분석하였다. Vienna노선도 ARIMAC(1,0,7)모형에서 자기상관성과 잔차에 대한 검정통계량을 모두 만족하는 것으로 나타났다.

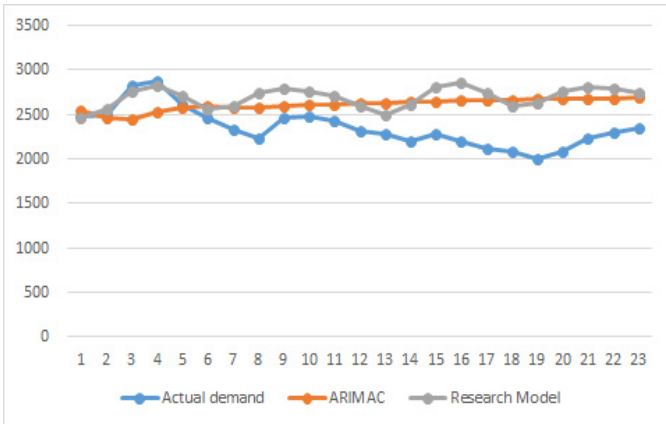
분석결과 상수항은 2723.07이고 AR계수는 0.916, AR상수인 C는 $2723.07 \times (1-0.916)$ 의 결과인 228.74로 나타났다. 모형구축 결과를 기반으로 2012년 1월부터 2013년 11월까지 예측구간에 대해 정확도를 분석하였다.

<Table 12> Comparison in prediction accuracy (Vienna route air fare)

Period	Fare	ARIMAC	Research Model	e1	e2	M1	M2
2012/1	2481	2547	2460	-66	21	2.7	0.8
2012/2	2494	2455	2566	39	-72	1.6	2.9
2012/3	2826	2443	2767	383	59	13.6	2.1
2012/4	2875	2529	2820	346	55	12.1	1.9
2012/5	2608	2576	2703	32	-95	1.2	3.6
2012/6	2454	2601	2560	-147	-106	6.0	4.3
2012/7	2338	2571	2597	-233	-259	10.0	11.1
2012/8	2224	2584	2738	-360	-514	16.2	23.1
2012/9	2457	2596	2790	-139	-333	5.6	13.6
2012/10	2484	2606	2765	-122	-281	4.9	11.3
2012/11	2425	2616	2717	-191	-292	7.9	12.0
2012/12	2318	2625	2589	-307	-271	13.2	11.7
2013/1	2278	2633	2500	-355	-222	15.6	9.8
2013/2	2197	2641	2606	-444	-409	20.2	18.6
2013/3	2282	2648	2809	-366	-527	16.0	23.1
2013/4	2205	2654	2857	-449	-652	20.4	29.6
2013/5	2118	2660	2737	-542	-619	25.6	29.2
2013/6	2088	2665	2590	-577	-502	27.6	24.0

2013/7	1998	2670	2623	-672	-625	33.6	31.3
2013/8	2085	2675	2763	-590	-678	28.3	32.5
2013/9	2226	2679	2814	-453	-588	20.3	26.4
2013/10	2296	2682	2787	-386	-491	16.8	21.4
2013/11	2351	2686	2736	-335	-385	14.2	16.4

분석결과 ARIMAC 모형의 MAPE 값은 14.5%로 나타났으며 연구모형인 Error adjusted ARIMAC모형의 MAPE 값은 15.7%로 나타났다. 앞서 두 개 노선의 분석결과와는 달리 연구모형보다 ARIMAC 모형의 정확도가 다소 높게 나타났다. 이러한 원인은 Vienna 노선의 항공운임이 예측구간에서 서서히 감소하는 형태로 나타나고 있는데 ARIMAC 모형은 이동평균을 순차적으로 적용한 모형적 특성으로 인해 가격의 변동성에 상대적으로 둔감하게 변동하지만 연구모형인 Error adjusted ARIMAC모형은 ARIMAC모형을 기반으로 실제 가격의 변동성을 기간별로 반영하였기 때문에 좀 더 민감한 운임의 예측에 대해서는 정확도가 높지만 실제 가격이 서서히 변화하거나 하강 또는 상승이 특징을 나타낼 때는 빠르게 실제 가격을 반영하지 못하는 것으로 보인다.

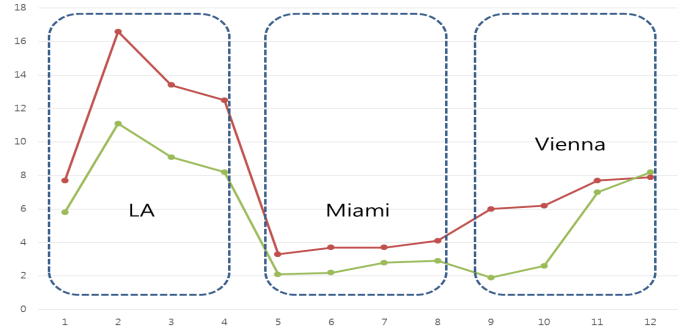


<Figure 12> Comparison of forecasting graph between ARIMAC and research model (Vienna)

지금까지 Error adjusted ARIMAC모형을 활용한 LA, Miami, Vienna노선의 항공운임을 비교하였는데 일반적으로 물류기업의 경우 1년 이상의 장기 예측을 활용하기보다 분기 혹은 1년 이내의 예측구간을 이용하는 특징이 있으며 ARIMA(p,q)모형의 경우 이동평균적인 특성이 나타나기 때문에 장기적인 예측에는 평균직선을 추정하는 특성이 나타난다. 따라서 3개 노선의 MAPE를 3, 6, 9, 12개월의 기간으로 구분하여 예측 정확도를 비교하는 것이 실무적으로 활용할 수 있는 정보를 제공할 수 있다.

<Table 13> Comparison in MAPE of 3 routes

Period	LA		Miami		Vienna	
	ARIMAC	Research Model	ARIMAC	Research Model	ARIMAC	Research Model
3	7.7	5.8	3.3	2.1	6.0	1.9
6	16.6	11.1	3.7	2.2	6.2	2.6
9	13.4	9.1	3.7	2.8	7.7	7.0
12	12.5	8.2	4.1	2.9	7.9	8.2



<Figure 13> Comparison in MAPE of 3 routes

<Table 13>은 3개 노선에 대해 예측 정확도를 비교한 것으로 연구모형인 Error adjusted ARIMAC모형이 전 노선에 있어서 전반적으로 예측 정확도가 높게 나타나고 있다. 이러한 결과를 통해 유추할 수 있는 특징으로 연구모형인 Error adjusted ARIMAC 모형은 1년 이상의 경우 장기예측보다 1년 미만의 단기예측에 더 우수한 예측력을 가지고 있는 것으로 나타났다. 이러한 연구결과와 더불어 3개 노선의 경우 모두 과거운임에 대해 7개월 이동평균이 반영된 노선으로 7개월간의 가격운임이 미래에도 지속적으로 영향을 미치는것으로 나타났다. 이러한 운임의 시차특성 정보는 물류기업의 물동량 및 운임 계약에 있어 중요 정보로 활용할 수 있으며 연구모형인 Error adjusted ARIMAC 모형의 단기예측 정확도를 결합하여 중요한 의사결정정보로 활용할 수 있다.

5. 요약 및 결론

본 연구는 국제항공운송협회(International Air Transport Association, IATA)에서 제공하는 CASS(Cargo Account Settlement Systems)자료를 이용하였으며 LA, Miami, Vienna의 3개 노선 항공운임에 대한 예측모형을 제시하였다.

일반적으로 예측모형 구축 시 폭넓게 이용되는 회귀분석은 소수의 영향요인을 활용하여 미래 시점의 가격을 예측하는데 이러한 분석방법은 다양한 외부환경요인에 의해 결정되는 항공운임의 예측에 적용할 경우 민감한 가격변화에 민첩하게 대응하지 못하고 예측력도 상대적으로 낮아 실무적으로 활용하는데 한계점을 가지고 있다. 또한 시계열 자료를 활용하여 예측하는 대표적인 도구로 ARIMA 모형이 있는데 ARIMA 모형의 경우 차분(Differencing)과 이동평균(MA)을 반영하여 예측하게 되는데, 활용하는 모수가 많아 질수록 계산과정이 복잡하여 현업에서 실무에 적용하는데 많은 어려움이 있다. 따라서 본 연구에서는 기존의 ARIMA 모형의 단점을 보완한 운임예측모형인 Error adjusted ARIMAC 모형을 제시하였다. 우선 3개(LA, Miami, Vienna) 노선에 대한 ARIMA 모형을 구축하기 위해 자기상관함수와 잔차에 대한 통계적 검정을 통해 ARIMA(1,0,7) 모델을 제시하였다. 또한 ARIMA(p,q)모형에서 발생하는 지속적인 자기상관성을 제거하기 위해 연속적으로 시차를 발생시켜 자기상관성을 제거한 ARIMAC모형을 제시하였으며 이를 기반으로 실무적으로 활용가능하고 ARIMAC의 단점을 보완한 연구모형 Error adjusted ARIMAC를 제시하였다. 또한 ARIMA 모형을 실무적으로 활용하기 위한 프로세스를 제시하여 모형 구축의 가이드라인을 제시하였다.

ARIMAC와 연구모형인 Error adjusted ARIMAC모형의 예측력에 대한 정확도분석을 비교하기 위해 MAPE를 비교한 결과 LA노선

에서 연구모형인 Error adjusted ARIMAC 모형의 MAPE는 10%로 나타났으며 ARIMAC의 MAPE는 11.2%로 나타나 연구모형의 예측정확도가 높은 것으로 나타났다.

Miami노선에서도 연구모형인 Error adjusted ARIMAC 모형의 MAPE는 3.5%인 반면 ARIMAC의 경우 3.7%로 나타나 근소한 차로 연구모형의 예측력이 높게 나타났다. 반면 Vienna 노선에서는 ARIMAC의 MAPE가 14.5%이며 연구모형의 MAPE는 15.7%로 나타나 ARIMAC의 예측 정확도가 다소 높게 나타났다. 하지만 이러한 분석결과를 기간별로 구분하여 MAPE를 분석한 결과 전반적으로 전 노선에 있어 12개월 미만의 예측 정확도는 연구모형인 Error adjusted ARIMAC 모형의 정확도가 매우 높게 나타나고 있다. 이러한 특징은 연구모형인 Error adjusted ARIMAC 모형의 이동평균적인 특징이 반영된 결과로 연구모형을 단기예측에 활용할 경우 유용한 도구로 활용할 수 있다.

3개 노선의 분석결과를 기반으로 연구모형인 Error adjusted ARIMAC 모형의 특징을 제시하면 다음과 같다.

우선 실제 가격의 변동성이 큰 경우의 가격예측에는 ARIMAC의 오차를 보정하는 특징이 있는 Error adjusted ARIMAC 모형의 정확도가 높게 나타나고 있으며 가격변동성이 작거나 서서히 상승 혹은 하강하는 가격예측에는 ARIMAC의 예측력이 높게 나타나는 것으로 판단된다. 이러한 특징은 이동평균을 순차적으로 적용한 ARIMAC 모형의 이동평균적 특성으로 인해 가격이 민감한 변화에 둔감하게 예측하지만 ARIMAC 모형을 기반으로 실제 가격의 변동성(오차)을 반영한 연구모형은 가격변동성에 대해 빠르게 반응하는 특징이 나타나고 있다. 이러한 연구결과를 기반으로 본 연구의 기여점을 제시하면 다음과 같다.

첫째, 물류기업의 항공운임 예측에 활용할 수 있는 연구모형인 Error adjusted ARIMAC 모형을 제시하였으며 실무적으로 활용 가능한 모형 개발 프로세스의 가이드라인을 제시한 점은 기존 연구와 차별화되는 실무적 기여점이라 할 수 있다.

둘째, Error adjusted ARIMAC 모형을 통해 장기예측보다 단기예측에 민감하게 반응하는 모델을 구축한 점과 항공운임이 7개월 이동평균적 특성을 나타내고 있다는 점을 제시한 것은 물류기업에 있어 물동량과 가격예측 측면의 중요한 관리적 시사점을 제시한다고 할 수 있다. 끝으로 본 연구의 주요 한계점을 제시하면 항공운임의 3개노선에 대한 특성만을 분석하였기에 Error adjusted ARIMAC 모형 적용시 일반화의 문제가 있을 수 있어 향후 다양한 노선에 대한 운임예측 모델을 제시하고자 한다.

References

Bowerman, B., O'Connell, R., & Koehler, A. (2005). *Forecasting, Time Series, and Regression* (4th ed.). Belmont, U.S.A.: Thomson Brooks.

Box, G. E. P., & Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control*. San Francisco, U.S.A.: Holden-Day.

Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (1994). *Time Series Analysis: Forecasting and Control* (3rd ed.). Englewood Cliffs, U.S.A.: Prentice Hall.

Burger, C. J. S. C., Dohnal, M., Kathrada, M., & R. Law (2001). A practitioners guide to time-series methods for tourism demand forecasting - a case study of Durban, South

Africa. *Tourism Management*, 22(4), 403-409.

Chen, C., Chang, Y., & Chang, Y. (2009). Seasonal ARMA forecasting of inbound air travel arrivals to Taiwan. *Transportmetrica*, 5(2), 124-140.

Chu, F. L. (2004). Forecasting tourism demand: a cubic polynomial approach. *Tourism Management*, 25(2), 209-218.

Coshall, J. (2006). Time Series Analyses of UK Outbound Travel by Air. *Journal of Travel Research*, 44(3), 335-347.

Hibon, M., & Evgeniou, T. (2005). To combine or not to combine: Selecting among forecasts and their combinations. *International Journal of Forecasting*, 21(1), 15-24.

Hur, Nam-kyun (2010). A Study on the Air Travel Demand Forecasting using Time-Series Model. Go-yang, Korea: Thesis for Doctorate in Korea AeroSpace University.

Jung, Dong-bin (2009). *Demand Forecasting of time series I*. Seoul, Korea: Hannarae Academy.

Korea Chamber of Commerce & Industry (2013). *SCM CEO REPORT Vol.12*. Seoul, Korea: KCCI. Retrieved June 21, 2013, from http://www.korcham.net/EconNews/PublishData/CRE07102R.asp?m_SITEID=13602&m_BOARDSEQ=A001&m_cci-Code=&m_query=&m_queryText=&m_page=10

Kulendran, N., & Witt, S. F. (2003). Forecasting the demand for international business tourism. *Journal of Travel Research*, 41(3), 265-271.

Lee, Choong-Ki, & Song, Hak-jun (2007). Selecting most Appropriate time Series Forecasting Model. *Journal of tourism Science*, 64(6), 289-311.

Lewis, C. D. (1982). *Industrial and Business Forecasting Methods*. London, U.K: Butterworth Scientific

Li, G., Song, H., & Witt, S. F. (2005). Recent developments in econometric modeling and forecasting. *Journal of Travel Research*, 44(1), 83-99.

Ljung, G. M., & G. E. P. Box (1978). On a Measure of Lack of Fit Time Series Models. *Biometrika*, 65, 297-303.

Song, Keun-Seok, & Lee, Choong-Ki (2009). Tourism Demand Forecasting by Combining Forecasts. *The Korea Academic Society of Tourism and Leisure*, 46(1), 183-202.

Totamane, R., Dasgupta, A., Mulukutla, R.N., & Rao, S. (2009). Air Cargo Demand Prediction. Proceedings of 3rd Annual IEEE International Systems Conference (pp.23-26). Vancouver, Canada: IEEE(International Air Transportation Conference)

Witt, S. F., Newbould, G. D., & Watkins, A. J. (1992). Forecasting domestic tourism demand: Application to Las Vegas arrivals data. *Journal of Travel Research*, 31(1), 36-41.

Wong, K. K. F., Song, H., Witt, S. F. & Wu, D. C. (2007). Tourism forecasting: To combine or not to combine?. *Tourism Management*, 28(4), 1068-1078.

Wu, Ming-Cheng, & Morell, P. (2007). China's Future Market Development and Implications. Proceedings of International Air Transportation Conference (pp.146-153). Irving, U.S.A.: ASCE(American Society of Civil Engineers)