

ARIMA를 이용한 항공기 수리부속의 수요 예측

(A Demand Forecasting for Aircraft Spare Parts using ARIMA)

박 영 진(Young Jin Park)*, 전 건 옥(Geonwook Jeon)**

초 록

신형/장기 운영 무기체계 수리부속의 불균형적 수요발생에 따른 항공기 불가동을 해소하기 위한 수요예측 기법 개선의 필요성이 대두되고 있다. 항공기 수리부속들은 고단가이고, 청구에 소요되는 기간이 길어 사전에 예측하지 못한다면 작전지원에 문제가 발생하게 된다. 신뢰성 있는 수요 예측은 과보유로 인한 재고비용을 줄일 수 있으며, 수요를 예측하기 위한 방법은 회귀분석, 단/다변량 시계열분석, 데이터 마이닝 기법 등이 있다. 항공기 부품의 수요 예측은 그 부품의 수가 8만 가지 이상이며 각 부품 간의 관계를 분석하기에 어려움이 있어 시간에 종속적인 단변량 시계열 분석을 통해 수요예측을 실시하였다.

본 연구에서는 이러한 문제점을 해결하기 위하여 첫째, 자료를 기존 AMMIS 체계에서 고장 자료를 실수요로 가정하여 수집하였다. 고장이나 주기 검사, 시한성 기술지시 등으로 부품을 장·탈착하게 되면 정비부서에서 이를 전산프로그램인 AMMIS에 입력하도록 하고 있다. 따라서 실제 정비부서에서 부품을 사용한 현황을 실수요라고 인정할 수 있다. 둘째, 1999년 1월부터 2007년 2월까지의 월별 자료(98개)를 수집하였다. 자료의 수가 충분하므로 예측 정확성 향상을 위하여 ARIMA 기법에 적용이 가능하다. 고장빈도가 높은 부품 50여개를 추출하여 Box-Jenkins의 ARIMA기법을 적용하여 예측을 실시하였다. 실시 결과 적합한 모형식을 도출하였으며, 현용기법보다 예측 정확성이 높다는 결론을 얻었다.

Abstract

This study is for improvement of repair part demand forecasting method of Republic of Korea Air Force aircraft. Recently, demand prediction methods are Weighted moving average, Linear moving average, Trend analysis, Simple exponential smoothing, Linear exponential smoothing. But these use fixed weight and moving average range. Also, NORIS(Not Operationally Ready supply) is increasing. Recommended method of Box-Jenkins' ARIMA can solve problems of these method and improve estimate accuracy. To compare recent prediction method and ARIMA that use mean squared error(MSE) is reacted sensitively in change of error.

ARIMA has high accuracy than existing forecasting method. If apply this method of study in other several Items, can prove demand forecast Capability.

* 공군 제5전술공수비행단 야전정비대대 통제실장

** 국방대학교 운영분석전공 부교수

1. 서론

최근 신형/장기 운영 무기체계 수리부속의 불균형적 수요발생에 따른 항공기 불가동을 해소하기 위한 수요예측 기법 개선의 필요성이 대두되고 있다. 이러한 맥락에서 공군에서는 '02년부터 '04년까지 창 수리부속의 수요예측기법을 다시 검토하여 좋은 결과를 제시하였다. 하지만 실무에서의 적용이 현재 시점에서 효과적으로 운영되고 있는지 확인되지 않고 있으며, 적용기법이 단순하며 논리적인 근거가 희박한 실정이다. 그런 의미에서 현실태를 재고해 보고 향상된 예측 기법을 적용하여 공군 항공기 수리부속품의 효과적인 기법을 추천하는데 연구의 목적이 있다.

공군의 현용 수요예측기법은 다음과 같다. 첫째, 가중 이동평균법은 3개년 최근 수요량 중 X년 50%, X-1년 30%, X-2년 20%의 가중치를 적용하여 예측수요량을 산출하는 방법이다. 둘째, 추세 분석법은 실제 수요량 발생 추세를 가장 잘 나타내는 대표값을 통해 예측하는 방법이다. 셋째, 지수 평활법은 최근의 자료에 더 많은 가중치를 두어 현 시점의 예측값에 과거의 예측오류를 가감하여 예측하는 방법이다. 넷째, 선형 이동평균법은 이동평균값을 다시 이동 평균하여 추세요인을 고려 예측하는 방법이다. 다섯째, 선형 지수평활법은 지수평활법에 추세요인으로 발생하는 과소/과대 추정분을 조정하는 방법이다. 시계열 자료는 각 부품별 5개년 분기별 자료 20개를 분석하여 예측오차가 가장 작은 모형을 적절한 모형으로 가정하고 다음 분기의 수요를 예측하고 있다.

현용기법에서는 평활에 사용된 가중치와 이동평균의 길이가 전 부품에 동일하게 적용되는 문제점을 지적하고, 예측 정확도가 높은 ARIMA 기법을 추천하였다. ARIMA 기법은 최소 50개 이상의 자료가 필요한데, 실수요를 반영하기 위하여 실제 수리부속을 교체한 실적을 시계열 자료로 제시하였다.

본 연구에서는 제00전투비행단의 00항공기 수리부속 중 50여개 품목을 Sampling하여 기존 5가지 예측 기법으로 예측한 결과와 확률적 예측 기법인 ARIMA(계절 ARIMA 포함)기법으로 예측한 결과를 비교분석하였다. 비교분석의 기준은 오차의 변화에 민감하게 변하는 특성이 있는 평균 제곱오차(Mean Squared Error : MSE)를 사용하였다.

연구의 제한사항은 급격한 변화들의 발생 시점을 분석하기 위한 과거 자료의 수집이 용이하지 않아 개입모형은 분석에서 제외하였다. 또한 8600여개의 품목 중 50여개의 품목으로 제한되어 있어 품목 특성별 예측기법에 대한 분석이 이루어지지 않았다.

2. 시계열 분석의 이론 검토

2.1. 시계열의 정의

시계열이란 한 사상 또는 여러 사상에 대하여 시간의 흐름에 따라 일정한 간격으로 이들을 관찰하여 기록한 자료를 말한다. 시간적 변화를 나타내는 역사적 계열이므로 어느 한 시점에서 관측된 시계열자료는 그 이전까지의 자료들에 주로 의존하게 된다. 따라서 시계열분석을 통한 예측에서는 관측된 과거의 자료들을 분석하여 법칙성을 발견하여 이를 모형화하여 추정하고, 이 추정된 모형을 사용하여 미래에 관측될 값들을 예측하게 된다.

시계열 자료는 크게 다음의 4가지의 변동 요인으로 분해하여 분석할 수 있다.

- (1) 추세변동(trend variation : T)
- (2) 계절변동(seasonal variation : S)
- (3) 순환변동(cyclical variation : C)
- (4) 불규칙변동(irregular variation : I)

이때 시점 t 에서 관측된 시계열을 Z_t 라고 하면 Z_t 는 일반적으로 이틀 네 가지 성분들의 곱(승법 모형)이나 합(가법모형)의 관계가 성립하는 것으로 가정하여 나타낸다.

- (1) 승법모형 : $Z_t = T_t \cdot S_t \cdot C_t \cdot I_t$
- (2) 가법모형 : $Z_t = T_t + S_t + C_t + I_t$

2.2. 시계열의 예측방법

2.2.1 정성적 예측방법

정성적 예측방법은 1950년대 이래 크게 발전되어 왔다. 이 방법은 예측가들의 주관적인 판단이나 의견을 근거로 예측하는 방법으로 명확한 수치적 기술을 고려하지 않고 주관적인 것을 입력하여 예측한다. 정성적 예측방법에는 시장조사법, 패널동의법, 중역의견법, 델파이기법 및 역사적 유추법 등이 있다.

2.2.2 정량적 예측방법

일정기간 동안 변화해 가는 어떤 변수의 움직임의 형태를 분류하고 연구하여 예측에 이용되는 것을 말한다.

2.2.2.1 이동평균법

일정기간의 시계열 자료를 대상으로 하여 산술평균 또는 가중평균치를 계산하여 계절적 및 불규칙 요인을 제거하는 기법으로서 이동식으로 산출하는데 그 특징이 있다. 이를 구하는 방법은 가장 과거의 관측치 하나를 삭제하는 대신 새로운 관측치 하나를 포함시켜서 계속적으로 이동하면서 산술평균치를 계산하기 때문에 이동평균법이라고 한다. 예측의 정확도는 단기·중기·장기 그 어느 것도 타기법에 비하여 미흡하지만 예측방법이 용이하다는 점에서 널리 사용되고 있으며, 세부방법으로는 전기수요법, 단순 이동평균법(simple

moving average method), 가중 이동평균법(weighted moving average method)이 있다.

2.2.2.2. 지수평활법

지수평활법은 이동평균법의 장점인 가중치 선정과 대상기간 n 을 정하는 기준의 주관적 성격을 보다 합리적으로 개선한 방법이다. 즉, n 기간을 과거 시계열 자료가 있는 모든 기간에 대하여 무제한 연장하여 대상을 삼으며, 가중치는 예측하려는 기간에 가까울수록 많이, 점차 멀어질수록 일정한 비율(α)에 따라 점진적으로 적게 n 기간에 걸쳐 적용하는 방법이다. 지수평활법에는 단순 지수평활법(Simple Exponential Smoothing Method)과 이중 지수평활법(Double Exponential Smoothing Method)이 있다.

2.2.2.3 인과형 예측기법

인과형 예측기법은 독립변수가 시간에 의존해 변화하는 정도가 기록된 특성을 가지고 있다. 회귀분석(regression analysis)은 인과분석을 통한 예측법으로서 예측대상변수(종속변수)와 밀접한 관계를 가진 변수(독립변수 혹은 설명변수)를 찾아서 이들 변수사이의 의존관계를 이용하여 예측하는 방법이다. 여기서 시계열 분석기법이 수요량 대 시간(년, 월 등)과의 분석을 바탕으로 하는 것임에 비하여, 회귀분석은 타 변수와의 인과관계를 분석하는 방법이다.

2.2.3 Box-Jenkins의 확률적 예측방법

분해법은 시계열의 변동요인이 각각 분해될 수 있다는 가정하에서 시계열을 분석하는 방법으로써 이론적으로 많은 약점을 가지고 있고, 또한 대부분의 시계열 자료는 변동요인들이 복합적으로 얽혀있어 쉽게 분해될 수 없는 것이 사실이다. 이에 시계열 분석에 대한 새로운 접근 방법을 정립했고, 또 이 방법을 더욱 일반화시킴으로써 통합

자기회귀이동평균(Autoregressive integrated moving average : ARIMA) 모형화를 토대로 하는 통계적인 이론 체계를 구축하게 되었다. 이 분석 방법은 관찰된 시계열 자료를 하나의 시계열 모형 집단으로부터 구축된 표본으로 간주하여 이들이 어떤 확률적 성질을 만족하는가를 조사하고 통계적 추정 및 검정을 통하여 적절한 시계열 모형을 수립하는 것이다. 구체적으로 Box-Jenkins 방법은 관찰된 시계열 자료가 어떤 확률적 성질을 가지고 있으며 어떤 시계열 모형(ARIMA 모형)이 적합한가를 찾기 위한 모형 식별(Model Identification), 모수 추정(Parameter Estimation) 그리고 모형의 적합성 진단(Model Diagnostic Checking)의 세 단계를 거친다. 모형식별은 관찰된 시계열 자료에 대하여 적절한 것이라고 생각되는 몇 개의 모형을 선정된 후 자료의 그래프, 관련 통계량 등을 비교하여 가장 적절한 하나의 모형을 선정하는 것이다. 이렇게 선정된 모형은 잠정적인 것으로 계속적인 분석을 통하여 개선시켜야 한다.

모형을 선정하는데 주의해야 할 점은 소위 모수 절약의 원칙(Principle of parsimony)에 충실해야 한다는 것이다. 이것은 관찰된 자료를 적절히 표현하면서 모수의 수가 가장 작은 모형을 선정하는 것으로써 가능한 한 단순하게 표현하려는 것이다.

선정된 모형은 하나 이상의 모수를 포함하고 있는데 이 모수는 관찰된 자료로부터 추정할 수 있다. 모수의 추정방법으로는 적률법, 최우추정법, 그리고 최소제곱법 등이 주로 사용된다.

모형이 선정되면 원래의 시계열 자료를 설명하는데 문제가 없는지 검정해야 하는데 이를 모형의 적합성 진단이라고 한다. 만일 모든 가정과 조건이 만족한다면 모형은 적절한 것이며 그 모형을 기초로 미래값을 예측할 수 있다. 그러나 가정 및 조건이 만족되지 않는다면 다시 모형식별 단계로 되돌아가서 적절한 모형이 찾아질 때까지 위의 세 단계를 반복하여야 한다.

2.2.3.1 확률과정과 정상성

임의의 시점 t 에서의 시계열을 Z_t 라고 표시하고 이를 하나의 확률변수(random variable)라고 하면 실제로 관찰된 자료는 확률변수 Z_t 의 모집단으로부터 각 시점에서 얻어진 표본으로서 이를 실현값(realization : z_t)이라고 한다. 일반적으로 확률변수들의 집합인 $\{Z_t, t = 0, \pm 1, \pm 2, \dots\}$ 를 확률과정(Stochastic Process)라고 하고, 특히 t 가 시간을 나타낼 때 이 확률과정을 시계열과정 또는 시계열 모형이라고 한다.

유한개의 관찰된 시계열 자료로부터 모집단인 시계열모형에 관한 통계적 추정 및 검정을 하기 위하여서는 분석을 단순화시키는 어떤 가정이 필요한데 이 중 가장 중요한 것이 정상성(Stationarity) 가정이다. 정상시계열이란 확률과정에서 발생하는 시계열 $\{Z_t\}$ 의 평균($E(Z_t)$)과 분산(σ_t^2)이 시점 t 에 관계없이 일정하고, 이들의 자기 공분산 γ_k 와 자기상관 함수 ρ_k 는 단지 시차 k 에 의존하고 실제 시점들인 t 와 $t+k$ 에 의존하지 않는다는 것이다.

2.2.3.2 자기상관함수

Z_t 와 Z_{t+k} 의 공분산(covariance)은 다음 식(1)과 같다.

$$\begin{aligned} \gamma_k &= \text{Cov}(Z_t, Z_{t+k}) \\ &= E(Z_t - \mu)(Z_{t+k} - \mu) \end{aligned} \quad (1)$$

한편 Z_t 와 Z_{t+k} 의 상관함수(Correlation Function)은 다음 식(2)와 같다.

$$\begin{aligned} \rho_k &= \frac{\text{cov}(Z_t, Z_{t+k})}{\sqrt{\text{Var}(Z_t) \text{Var}(Z_{t+k})}} = \frac{\gamma_k}{\gamma_0}, \quad (2) \\ \gamma_0 &= \text{Var}(Z_t) = \text{Var}(Z_{t+k}) \end{aligned}$$

여기서 γ_k 는 k 의 함수이므로 자기공분산함수 (autocovariance function)이라 하고, ρ_k 는 자기상관함수 (autocorrelation function : ACF)라고 한다.

Box-Jenkins 모형을 이용한 시계열 분석은 자기상관함수로서 시계열의 특성을 파악하는 것이다. 그러나 실제 모집단의 자기상관함수를 알 수 없으므로 관찰된 자료(z_t)로부터 추정해야 하는데, 이 추정량을 표본자기상관함수 (Sample autocorrelation function : SACF)라고 하고 다음 식 (3)과 같이 정의한다.

$$\gamma_k = \hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} = \frac{\sum_{t=1}^{n-k} (z_t - \bar{z})(z_{t+k} - \bar{z})}{\sum_{t=1}^n (z_t - \bar{z})^2}, \quad (3)$$

$k = 0, 1, 2, \dots$

시차 k 에 대한 $\hat{\rho}_k$ 의 그래프를 표본상관도 (Sample correlogram)라고 하며, 이것은 시계열의 특성을 파악하고 모형을 식별하는데 유용하게 쓰인다. 또한 n 이 큰 경우 γ_k 는 근사적으로 평균 ρ_k 를 갖는다.

2.2.3.3 부분자기상관함수

부분자기상관함수란 임의의 시차 k 에 대하여 Z_t 와 Z_{t+k} 의 상관관계에서 그것들 사이에 끼어 있는 $Z_{t+1}, Z_{t+2}, \dots, Z_{t+k-1}$ 의 상호 선형 의존성을 제거한 후의 Z_t 와 Z_{t+k} 만의 자기상관함수로 다음 식(4)와 같은 조건부 자기상관함수로 표현된다.

$$\phi_{kk} = \text{Corr}(Z_t, Z_{t+k} \mid Z_{t+1} = z_{t+1}, \dots, Z_{t+k-1} = z_{t+k-1}) \quad (4)$$

이와 같은 부분자기상관함수 ϕ_{kk} 는 시차 k 의 함수이며 다음과 같은 회귀 모형을 고려함으로써

얻을 수 있다. 즉, 평균이 0인 정상과정에서 종속변수 Z_{t+k} 을 k 개의 시차변수 $Z_{t+k-1}, Z_{t+k-2}, \dots, Z_t$ 로 회귀분석하면 다음 식(5)와 같다.

$$Z_{t+k} = \phi_{k1}Z_{t+k-1} + \phi_{k2}Z_{t+k-2} + \dots + \phi_{kk}Z_t + \varepsilon_{t+k} \quad (5)$$

이때 $\{\phi_{kk}, k = 0, 1, 2, \dots\}$ 를 부분 자기상관함수라고 한다. 정상 시계열 $\{Z_t\}$ 에 대하여 $\rho_k = \phi_{k1}\rho_1 + \phi_{k2}\rho_2 + \dots + \phi_{kk}\rho_k$ 와 같은 방정식을 생각해 보자. 만일 $\{\rho_k\}$ 가 주어졌다면 방정식의 해 중에서 ϕ_{kk} 는 시차 k 인 부분자기상관함수 값이 되는 것이다. 따라서 시차 k 인 부분자기상관관계수 ϕ_{kk} 는 다음 식(6)과 같다.

$$\phi_{kk} = \frac{|P_k^*|}{|P_k|}, \quad (6)$$

P_k : $k \times k$ 인 자기상관 행렬

P_k^* : 마지막 열을 (ρ_1, \dots, ρ_k) 로 대치한 행렬

부분상관관계수는 자기상관값을 이용하여 구하게 되며 모형에 포함되지 않은 시차 이후의 부분상관관계수는 '0'의 값을 갖게 된다.

표본 편자기상관함수 (Sample partial autocorrelation function : SPACF) $\hat{\phi}_{kk}$ 는 ρ_k 대신에 $\hat{\rho}_k$ 를 대체하여 구한다. 복잡한 행렬식을 계산하는 대신 다음 식(7)과 같은 Durbin과 Levinson의 알고리즘을 이용한 축소모형을 사용한다.

$$\hat{\phi}_{k+1, k+1} = \frac{\gamma_{k+1} - \sum_{j=1}^k \hat{\phi}_{kj} \gamma_{k+1-j}}{1 - \sum_{j=1}^k \hat{\phi}_{kj} \gamma_j} \quad (7)$$

$$\hat{\phi}_{k+1, j} = \hat{\phi}_{kj} - \hat{\phi}_{k+1, k+1} \hat{\phi}_{k, k+1-j},$$

$$j = 1, \dots, k$$

귀무가설이 $\phi_{kk} = 0$ 인 경우에는 $\hat{\phi}_{kk}$ 의 분산은 근사적으로 다음 식(8)과 같다.

$$\text{Var}(\hat{\phi}_{kk}) \simeq \frac{1}{n} \quad (8)$$

따라서 $\phi_{kk} = 0$ 의 가설을 검정할 때 95% 신뢰 구간은 $\hat{\phi}_{kk} \pm 1.96 \sqrt{1/n}$ 을 사용할 수 있다.

2.3 ARIMA 모형 이론

2.3.1 자기회귀모형(AR)

시계열유형을 나타내는 기본모형 중 첫 번째 형태는 자기회귀모형으로 예를 들어 자기회귀모수 한 개를 갖는 AR(1)모형은 다음 식(11)과 같다.

$$Z_t = \phi_1 Z_{t-1} + a_t \quad (11)$$

Z_t 는 정상계열이며 $\phi_1 Z_{t-1}$ 은 원시계열 Z_t 에 적합된 적합치이고, ϕ_1 을 차수 1의 자기회귀모수로 표시한다. 또한 a_t 는 t 시점의 시계열자료에서 가정된 확률오차를 말한다.

위 식의 모형은 t 시점의 계열치가 단지 t-1 시점의 계열치와 t 시점의 확률오차 a_t 를 합한 것에 좌우되는데 이는 t 시점에 발생한 자료가 t-1시점에 발생한 자료와 t시점의 확률오차 a_t 만을 더한 것에 따라 영향을 받는다는 것을 의미한다. 그러나 물론 Z_t 가 바로 한 시점 이전의 과거치 뿐만 아니라 그 이전의 과거치와 관련될 수도 있다. 예를 들어 AR 모수 두 개를 갖는 AR(2)모형은 다음 식(12)와 같다.

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + a_t \quad (12)$$

이때 ϕ_1 과 ϕ_2 를 각각 차수 1과 2의 AR모수라고 한다. 또한 위 식의 모형은 t시점의 계열치 Z_t 가 t-1시점과 t-2시점의 계열치 Z_{t-1} 과 Z_{t-2} 와 t시점의 확률오차 a_t 를 합한 적합치와 관련되어 있음을 뜻한다.

이 개념을 일반화시켜 p개의 AR 모수를 갖는 일반 AR(p)모형은 다음 식(13)과 같다.

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t \quad (13)$$

$\phi_1, \phi_2, \dots, \phi_p$ 를 AR 모수로서 ϕ 의 첨자 1, 2, ..., p를 AR 모수의 차수(order)라고 하며, 가장 높은 차수 p를 모델의 차수로 간주한다.

위 식의 일반 AR 모형에서 계열치 Z_t 는 p개의 이전시점의 계열치와 확률오차 a_t 를 합한 적합치의 조합형태이다.

단순평활화 모형을 위의 AR모형과 비교해 보면 사실상 3차 AR 모형: AR(3)과 같은 형태를 갖지만 엄격한 의미로는 AR 모형과 단순평활화 모형에는 차이가 있다. 즉, Box-Jenkins 모형에서는 시계열형태의 AR 유형과 이의 모수를 계산하는 체계적인 방법을 갖는 반면에 평활화모형은 그러한 체계적인 과정을 갖고 있지 못하다는 점이다.

2.3.2 이동평균모형(MA)

Box-Jenkins 기본모형의 두 번째 형태는 이동평균모형으로 MA 모수 한개를 갖는 MA(1) 모형은 다음 식(14)와 같다.

$$Z_t = a_t - \theta_1 a_{t-1} \quad (14)$$

$-\theta_1 a_{t-1}$ 은 계열치 Z_t 에 적합된 적합치이고, θ_1 은 차수 1의 MA모수라고 부르는데 AR모형을 표시한 것과 비슷하게 MA(1)으로 표시하며, θ_1

계수의 '-'부호는 다만 편의상 이용한 것으로 아무런 의미는 없다.

MA 모형의 외관상 AR 모형과 거의 비슷하지만 의미는 전혀 다르다. 즉 MA 모수는 t시점에 발생한 계열치는 단지 그 이전시점에 발생한 확률오차 a_{t-1}, a_{t-2}, \dots 에만 관련된 것이다. 위 식의 모형은 t시점의 계열치 Z_t 가 바로 이전의 t-1시점의 확률오차 a_{t-1} 과 현재 t시점의 확률오차 a_t 만을 더한 것에 따라 영향을 받는다는 것이다.

MA 모형도 AR모형처럼 일반화시켜 q개의 모수를 갖는 일반 MA(q)모형은 다음 식(15)와 같다.

$$Z_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad (15)$$

$\theta_1, \theta_2, \dots, \theta_q$ 를 각각 차수 1, 2, ..., q의 MA 모수라고 하며 가장 높은 차수 q를 MA모형의 차수로 간주한다. 위식의 일반 MA모형에서 계열치 X_t 는 q개의 이전시점의 확률오차와 현재시점의 확률오차 a_t 를 합한 것으로 표현한다. 또한 q차수의 MA(q)모형은 다음 식(16)과 같다.

$$Z_t = a_t - \theta_q a_{t-q} \quad (16)$$

차수 q의 MA 모수 한 개만을 갖는 q번째 차수의 모형을 나타내므로 q개 MA 모수보다 적게 갖고 있어 q보다 낮은 차수항이 제외되어 있다.

MA라는 용어는 MA모형이 단순히 t가 증가함에 따라 시간상 이동하는 이전시점의 확률오차 크기의 가중평균을 나타내는 것으로 유도된 것이다. MA개념이 임의의 계열치가 실제로 과거시점의 확률오차를 이용한 것이며 이것이 Box-Jenkins 모형과 종래의 일반적인 예측모형과 구분하여 설명할 수 있는 개념이 된다. 따라서 Box-Jenkins 모형에서 확률오차성분은 모형의 형태를 결정하는 아주 중요한 요소가 된다.

2.3.3 자기회귀이동평균모형(ARMA)

AR모수와 MA모수 모두를 갖는 모형 형태를 자기회귀이동평균모형이라고 부르며, 이의 일반형은 다음 식(17)과 같다.

$$Z_t = (\phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p}) + a_t - (\theta_1 a_{t-1} + \theta_2 a_{t-2} + \dots + \theta_q a_{t-q}) \quad (17)$$

ARMA 모형의 차수는 p와 q이며 대개 ARMA(p, q)로 표시한다.

정상계열이 주어졌을 때 분석대상의 계열의 형태를 가장 잘 나타내는 적합모형을 선택하는 것으로 AR모형인지 MA모형인지를 판단하는 것이다. 판단 기준은 다음과 같다. 자기회귀 과정의 자기상관 함수와 이동 평균과정의 부분 자기상관 함수는 시차 k가 커짐에 따라 각각 감소하는 형태를 갖는다. 또 자기회귀과정의 부분 자기상관 함수와 이동평균과정의 자기상관 함수는 차수 p, q보다 큰 시차 이후는 0의 값을 갖고 소멸해 간다. 즉, 자기상관 함수와 부분 자기상관 함수는 모형의 식별에 대단히 유용하다. AR과정, MA과정, ARMA과정의 자기상관 함수와 부분 자기상관 함수와의 관계를 정리하면 다음 [표 1]과 같다.

[표 1] 자기상관함수들의 성질

구분	AR(p)	MA(q)	ARMA
자기상관 함수 (ACF)	시나브로	시차 q+1에서 절단형태	시나브로
부분자기상관함수 (PACF)	시차 p+1에서 절단형태	시나브로	시나브로
역자기상관함수 (IACF)	시차 p+1에서 절단형태	시나브로형태	시나브로형태
※ 시나브로형태 : 시차 k가 커짐에 따라 지수적으로 감소하는 형태 또는 지수적으로 감소하는 사인함수형태 ※ 절단형태 : 시차 k+1에서부터 자기상관함수값이 0인 형태			

위 표와 같이 Box-Jenkins 식별 방법으로는 ARMA 모형의 차수들을 식별하지 못한다. 이 모형은 자기상관함수와 부분자기상관함수 모두가 시나브로형태를 갖기 때문에 추정함수가 아닌 모수함수들이 주어진다 해도 p값과 q값을 정할 수 없으므로 모형 식별을 객관적으로 하기 위하여 다음 식(18)와 같은 통계량을 기준으로 모형을 식별하게 된다.

$$AIC(p, q) = \ln \hat{\sigma}_e^2(p, q) + \frac{2(p+q)}{T} \quad (18)$$

이 함수를 최소화하는 (p, q)를 최소 AIC 추정량이라고 한다. 하지만 통계적 불일치성이 있기 때문에 다음 식(19)과 같이 BIC를 사용하여 차수를 결정한다.

$$BIC(p, q) = \ln \hat{\sigma}_e^2(p, q) + \frac{2 \ln \ln T}{T}(p+q) \quad (19)$$

함수를 최소화하는 (p, q)를 최소 BIC 추정량이라고 한다. 이 최소 BIC 추정량은 일치성을 가짐이 증명되었다.

SAS/ETS의 PROC ARIMA에서 estimate statement를 사용하면 AIC값과 BIC값이 출력된다. 이 통계량을 통해 ARMA 모형을 추정하게 된다.

2.3.4 자기회귀누적이동평균(ARIMA)

자기회귀의 차수가 p이고, 이동평균의 차수가 q이며 차분차수가 d일때 시계열은 ARIMA(p, d, q)모형을 가지며 다음 식(20)와 같이 표현할 수 있다.

$$\phi(B)(1-B)^d Z_t = \theta(B)a_t \quad (20)$$

$$\begin{aligned} \phi(B) &= (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) \\ \theta(B) &= (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) \end{aligned}$$

Where

B: 후진연산자, $BZ_t = Z_{t-1}$

d: 일반 시계열 차분차수 (differencing order)

p: 자기회귀 차,

ϕ : 자기회귀 모수

q: 이동평균 차수

θ : 이동평균 모수

ARIMA(p, d, q) 모형에서 p=0, d=1, q=1 이면 일반차분 이동평균모형 IMA(1,1)과정이라고 하는데, 일반식은 $Z_t(1-B) = a_t(1-\theta_1 B)$ 이고 다시 정의하면 다음 식(21)과 같다.

$$Z_t = Z_{t-1} + a_t - \theta_1 a_{t-1} \quad (21)$$

특히 ARIMA(p, d, q) 모형에서 p=0, d=1, q=0 이면 확률보행과정(random walk)이라고 하고, $Z_t(1-B) = a_t$, 정리하면 $Z_t = Z_{t-1} + a_t$ 가 된다.

ARIMA(p, d, q)의 자기상관함수와 부분자기상관함수는 ARMA(p, q) 모형의 자기상관함수와 부분자기상관함수의 패턴과 동일하다.

2.3.5 계절성모형

월별 혹은 분기별 시계열 자료들은 일정한 시간 간격을 두고 매년 동일한 현상이 반복되는 경향을 가지고 있다. 강수량은 여름에 가장 많고 겨울에 가장 적은 경향을 띠고 있으며, 아이스크림은 여름철에 가장 많이 판매된다. 군의 수리부속에도 계절성을 띠고 있는 품목이 있는지 알아보았으나 표본 집단의 수가 50개로 제한되어 있기 때문에 확인해 볼 수는 없었다. 하지만 윤회계통의 품목과 계절에 따라 많이 교체되는 품목은 반

드시 있을 것으로 예상되며 본 연구에서는 한 개 품목이 계절형 승법 모형을 가짐을 확인하였다.

계절성을 가지는 시계열의 경우에는 계절에 해당되는 항만이 모형에 필요한 경우, 부분모형을 이용하여 모형을 만든다. AR연산자 또는 MA 연산자들이 비계절형 연산자와 계절형 연산자의 곱의 형태로 표현되는 경우 다음 식(22)와 같이 승법모형으로 적합된다.

$$(1-B)^d(1-B^s)^D\phi_p(B)\Phi_p(B^s)W_i \quad (22)$$

$$= \mu + \theta_q(B)\Theta_Q(B^s)\varepsilon_t$$

s : 계절주기

2.4 예측정확도 측정방법

수요를 예측할 때에는 예측치(forecast value)와 실제치(actual value)사이의 오차(error)가 존재하기 마련이다. 예측의 정확성 평가는 예측치(X_t)와 실제치(x_t)간의 차이인 예측오차(ε_t)의 크기에 따라 예측의 질을 평가하는 것으로 여러 가지 평가척도가 사용된다. 예측오차의 측정방법에는 평균오차, 평균절대편차, 평균자승오차, 평균절대비율오차가 있다. 이 중에서 평균제곱오차가 가장 많이 사용된다.

평균자승오차는 예측된 각 기간의 오차의 자승을 모두 합하여 예측기간 수로 나눈 것이다. 평균자승오차는 평균오차가 가지고 있는 단점인 양과 음의 오차의 상쇄효과를 제거할 뿐만 아니라 오차를 자승함으로써 단위오차는 각각 다른 가중치를 가지고 있는 것과 같다. 즉, 작은 오차의 경우는 자승함으로써 작은 가중치를 더욱 큰 가중치를 가지고 있는 것과 같이 보인다. 가장 많이 사용되는 측정방법이다. 이를 수식으로 표시하면 다음 식(23)와 같다.

$$MSE = \frac{\sum_{t=1}^n (x_t - X_t)^2}{n-1} \quad (23)$$

3. 공군의 수요예측 기법

공군은 작전요구도와 군수지원 능력을 고려하여 적정 수준의 기종별 목표운영률을 결정한다. 항공기 성능별 작전임무와 조종사의 자격 유지를 위한 작전요구도를 만족하기 위하여서는 정비지원과 보급지원 측면에서 그 능력이 충족되어야 한다.

이러한 측면에서 보급 지원능력 향상 및 경제적인 군 운영을 위하여 2003년 초에 공군은 수요예측 기법 개선 계획을 수립하였다. 기존의 수요예측 방법은 분기별로 산정하며, 과거 3개년 자료를 이용하여 가중 이동평균법으로 소요량을 산출하여 왔다. 일률적 예측기법 적용으로 인한 문제점들은 품목별 수요발생 특성이 미고려되어 수요예측 정확성이 떨어지고, 예산 및 조달계획 수립 시 자료로서의 타당성이 저하되었다. 이에 대한 개선 방안으로 품목별 수요발생 특성을 고려한 현용 5가지 수요 예측 기법을 적용하여 예측의 정확성을 높이려는 노력을 시작한 것이다.

적용 예측기법으로는 일정한 수요 추세를 가지는 품목은 가중 이동평균법을 적용하고, 지속적 증가 또는 감소 추세의 것은 추세분석법과 선형 이동평균법, 급격한 수요변동이 있는 품목에 대해서는 지수평활법, 선형 지수평활법을 적용하였다. 통계적 예측이 불가능한 3년간 1~2회의 수요가 발생한 품목에 대해서는 과거 운영 자료를 비교

[표 2] 예측정확성 판단을 위한 모의검증 실시 결과

구 분	1차 검증 ('02. 11)	2차 검증 ('03. 6)	3차 검증 ('03. 6)
대상기간	10개년 ('01. 6. 30기준)	10개년 ('02. 12. 31 기준)	5개년 ('02. 12. 31기준)
기본 자료수 (분기별)	40개	40개	20개
대상품목	2,903 품목	86,620 품목	86,620 품목
MSE 증감	46% 감소	30% 감소	34%감소

분석하여 타당성을 검증하였다. 또한 3개년 연간 수요실적(3개)을 5개년 분기별 수요실적(20개)으로 기본 자료를 재구축하였다.

3차에 걸친 검증 결과 [표 2]와 같이 평균제곱 오차(MSE : Mean Square Error)는 34% 감소한 것으로 분석되었고, 예측 정확성이 향상되었다. 검증 절차 및 방법은 부록 1에 첨부하였다. 이러한 기법의 변화로 수요 저빈도 품목 삭제에 따른 청구목표 운영 효율성이 증가하였고, 다빈도 품목 위주의 RO 설정 및 운영으로 보급지원 능력을 제고하였다.

하지만 재고의 부족으로 항공기 불가동률은 계속해서 증가하고 있으며, 이를 줄이기 위한 청구 시기 설정과 수요량 예측기법의 부정확성이 여전히 문제가 되고 있다. 또한 소요 제기부서에서 보급지원, 물량부족, 수요급증 시 선확보를 위한 과다청구로 가수요가 발생되고 있어 예측에 사용되는 자료의 신뢰성이 저하된다.

이에 본 연구에서는 실수요 반영을 위하여 분석을 위한 자료의 획득 방법을 달리하고 현용 수요예측기법의 단점을 보완할 대안으로 ARIMA 기법을 추천한다. 먼저 공군의 5가지 예측기법을 소개한다.

3.1 공군의 현용 예측 기법

3.1.1 가중이동평균법

최근 3개년간 수요량에 가중치를 적용하여 예측수요량을 산출하는 방법으로 다음 식(24)과 같이 수요를 예측하고 있다. 고정된 가중치 값을 사용하고 있다.

$$Z_{t+1} = 0.2 \times Z_{t-2} + 0.3 \times Z_{t-1} + 0.5 \times Z_t \quad (24)$$

3.1.2 선형이동평균법

관측된 시계열이 선형 추세성을 갖는 경우에 적합시킬 수 있는 평활법인 선형 이동평균 방법은 이중이동평균(double moving average: MA')을 이용하여 현재 시점이 n인 경우에 n+1시점의 값을 다음 식(25)와 같이 예측하는 방법이다.

$$\begin{aligned} Z_{n+l} &= a_n + b_n \cdot l \quad (25) \\ &= (2MA_n - MA_n') \\ &\quad + \frac{2}{m-1}(MA_n - MA_n') \cdot l \end{aligned}$$

여기서 MA_n 은 이동평균을 MA_n' 은 이중이동평균을 나타낸다. 이중이동평균이란 원계열의 이동평균을 다시 이동평균 한 것으로 다음 식(26)과 같이 정의된다.

$$MA_n' = \frac{1}{m}(MA_n + MA_{n-1} + \dots + MA_{n-m+1}) \quad (26)$$

3.1.3 추세분석법

실제 수요량 발생 추세를 가장 잘 나타내는 대표 값을 통해 예측하는 방법으로 다음 식(27)와 같이 선형 회귀 모형을 만들어 예측값을 구한다.

$$Y = bX + a \quad (27)$$

$$b = \frac{n \sum_{t=1}^n t Y_t - \sum_{t=1}^n t \sum_{t=1}^n Y_t}{n \sum_{t=1}^n t^2 - \left(\sum_{t=1}^n t \right)^2}$$

$$a = \frac{1}{n} \sum_{t=1}^n Y_t - b \left(\frac{1}{n} \sum_{t=1}^n t \right) = \bar{Y} - b \cdot \bar{t}$$

3.1.4 단순지수평활법

과거의 관측값을 가지고 미래의 값을 예측할 때 최근의 자료들에 대하여 더 많은 가중치를 부

여하여 예측하는 방법이다. 단순 지수평활법에서는 최근의 관측치 z_{t-1} , 바로 전 시점의 단순 지수평활 예측치 F_{t-1} 와 평활상수 값만으로 예측이 가능하다. 실제로 단순 지수평활예측치는 과거 시점의 시계열자료들의 가중평균이다. 즉, 단순 지수평활법은 식(28)와 같은 방법으로 예측치를 계산할 수 있다.

$$F_t = F_{t-1} + \alpha(z_{t-1} - F_{t-1}) \quad (28)$$

$$= \alpha z_{t-1} + (1 - \alpha)F_{t-1}$$

where F_t : t기의 단순 지수평활법의 예측치

α : 지수평활 상수 ($0 \leq \alpha \leq 1$)

z_{t-1} : t-1기의 실제치

F_{t-1} : t-1기의 단순지수평활법의 예측치

위 식에서 단순 지수평활(Simple Exponential Smoothing) F_t 를 평활통계량(smoothed statistic)이라고 한다. 단순 지수평활값 F_t 는 $F_0, z_1, z_2, \dots, z_{t-1}$ 의 가중평균값이다. 그리고 시점 t에 대한 평활추정값 F_t 는 모든 과거의 관측치의 선형결합이다. F_0 는 β_0 에 대한 초기추정값(initial estimate)이다. α 의 선택에 따라 F_t 의 부드러운 정도가 결정된다. 0에 가까운 값을 선택하게 되면 시계열의 이전(과거)값에 보다 큰 가중치가 부여되므로 이전 값이 보다 더 강조된다. 따라서 시계열은 더 평활화된다. 반면에 1에 가까운 값을 선택하게 되면 시계열의 현재 값에 보다 큰 가중값이 부여된다. 여기서 주목하여야 할 것은 초기 추정값 F_0 를 어떻게 정하느냐 하는 것이다. 만일에 과거의 자료가 주어졌다면 전체평균값으로 F_0 를 정할 수 있으며, 그러한 자료가 없으면 시점 1의 관측치 z_{t-1} 이나 시점 1에서 시점 2까지의 관측값들의 산술평균을 사용한다.

3.1.5 선형지수평활법

지수평활법에 추세요인으로 발생하는 과소/과대 추정분을 조정하는 방법이며 다음 식(29)과 같은 공식에 의해 예측값을 구한다.

$$n \text{ 시점의 예측치} \quad (29)$$

$$= F_n + (F_n - F_n') = 2F_n - F_n'$$

$$F_{n+1}' = \alpha F_n + (1 - \alpha)F_n'$$

F_n' : n기간의 이중지수 평활치

F_n : n기간의 단순 평활 예측치

3.2 현용기법 분석 결과

공군의 수요예측기법을 이용하여 산출된 평균 제곱오차(MSE) 값은 다음 [표 3]과 같다. 자료는 정보정보체계에서 산출된 월별 정보실적치 자료를 이용하였다. 진하게 표시된 값이 최적의 기법으로 산출된 값이며 가장 오른쪽 열의 예측값은 최적 기법으로 예측한 1개월 수요를 예측한 값이다. 샘플에서는 선형이동평균에 적합한 품목은 나타나지 않았다.

우선 이동평균법은 산술평균의 특성을 가지고 있어서 수요추세가 상승하고 있거나 하강하고 있을 때에는 과소예측 또는 과다예측하는 단점을 발견하였다. 또한 몇 개월의 평균치를 사용할 것 인지를 예측자는 부품별 적용 기준이 동일하여 평균제곱오차를 최소화 하는 부품별 이동평균의 길이를 산출하지 못한다. 또한 가중이동평균에서 가중치에서도 같은 문제로 인해 부품별 가중치 적용을 동일하게 하고 있다. 다음으로 추세분석법은 회귀선을 식별하여 시계열의 추세를 분석하기는 용이하나 선형회귀에 제한되어 있으며, 분산이 큰 시계열의 경우에는 적합하지 않다. 지수평활법 역시 가중이동평균법과 같이 α 의 선정에 있어서 예측가의 주관이 개입되어 있다. 실제 α 값을 0.1 이하로 설정하여 분석한 결과 평균제곱오차가 작게 나오는 품목들이 발생하였다. α 값 선정의 절차를 거쳐 부품별 가중치를 재설정할 필요가 있다.

[표 3] 현용기법을 적용한 결과

순번	최적 기법	MSE	예측값 ('07. 3)
1	선형지수평활	33.8	1.56
2	지수평활($\alpha:0.2$)	12.35	3
3	가중이동평균	240.47	22.2
4	지수평활($\alpha:0.2$)	4.86	5.12
5	지수평활($\alpha:0.2$)	2.71	0.173
6	추세분석	3.358	1.78
7	추세분석	1.14	0.69
8	지수평활($\alpha:0.4$)	5.53	2.8
9	지수평활($\alpha:0.1$)	10.28	1.01
10	추세분석	1.33	1.34
11	추세분석	2.56	1.63
12	가중이동평균	59.59	25.5
13	추세분석	4.75	1,841
14	선형지수평활	1.03	1.02
15	지수평활($\alpha:0.3$)	23.14	0.43
16	지수평활($\alpha:0.2$)	1.02	0.186
17	가중이동평균	4.07	6.7
18	추세분석	1.84	1.39
19	가중이동평균	47.06	12.8
20	추세분석	4.23	4.89
21	지수평활($\alpha:0.1$)	3.98	1.78
22	지수평활($\alpha:0.1$)	4.08	2.35
23	선형지수평활	11.20	1.77
24	지수평활($\alpha:0.2$)	10.9	3.92
25	지수평활($\alpha:0.3$)	12.63	2.88
26	선형지수평활	16.36	3.25
27	가중이동평균	1.64	3.3
28	추세분석	22.48	5.35
29	추세분석	8.29	7.63
30	지수평활($\alpha:0.1$)	44.77	4.83
31	추세분석	2.76	2.73
32	추세분석	1.58	2.12
33	추세분석	1.43	0.88
34	선형지수평활	20.62	1.68
35	추세분석	1.09	0.87
36	지수평활($\alpha:0.2$)	8.33	1.71
37	선형지수평활	10.45	3.67
38	추세분석	4.58	4.63
39	지수평활($\alpha:0.5$)	284.97	3.73
40	추세분석	1.30	0.875
41	선형지수평활	27.17	5.30
42	추세분석	23.87	14.66
43	추세분석	1.22	1.18
44	지수평활($\alpha:0.1$)	4.18	5.29
45	지수평활($\alpha:0.2$)	13.28	7.25
46	선형지수평활	4.23	4.18
47	지수평활($\alpha:0.1$)	6.08	4.77
48	지수평활($\alpha:0.1$)	10.73	6.75
49	지수평활($\alpha:0.2$)	2.43	2.38
50	추세분석	3.83	5.63
51	지수평활($\alpha:0.1$)	4.42	5.32
52	지수평활($\alpha:0.1$)	3.82	4.1
53	지수평활($\alpha:0.1$)	2.26	1.84
54	지수평활($\alpha:0.1$)	5.19	0.74

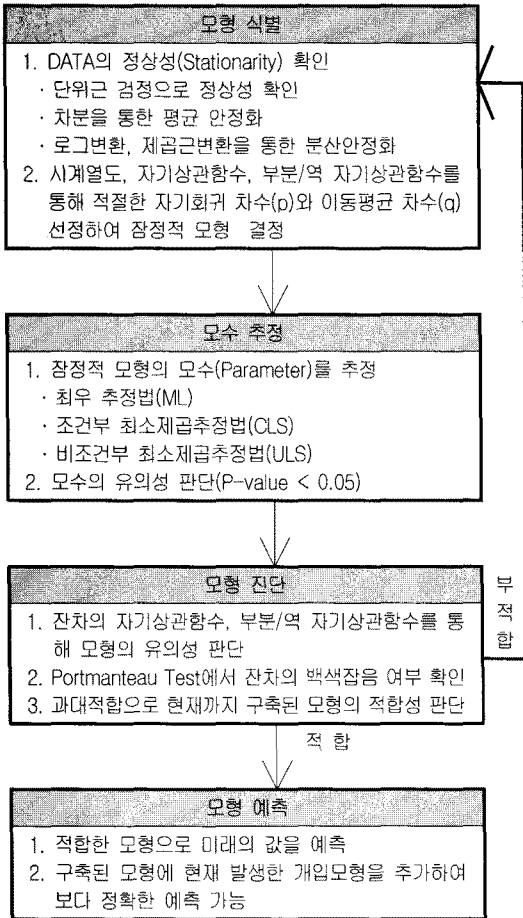
4. 실증분석

4.1 ARIMA 모형 연구를 위한 자료의 처리

군의 수요 관리는 S/N(Stock Number)별로 이루어지고 있으며, 청구되는 양에 따라서 5가지 기법으로 예측하고 있다. 정비부서의 소요에 따라 수요를 예측하는 것은 실수요 반영이 불가하다. 다시 말해 수요예측 기법으로 수요를 예측하기 위하여서는 항공기 부품의 사용 여부를 판단해야 하지만 정비부서에서 신청하는 부품 수는 실수요라고 보기는 힘들다.

본 연구에서 사용한 자료는 다음과 같다. 소요란 당연히 실소요만을 뜻하나, 시장에서 인플레이션 등의 경우 매점, 매석 등에 의한 가수요가 발생, 군에서도 보급지연, 물량부족, 수요급증 시 선확보를 위한 과다청구로 가수요가 발생되고 있다. 가수요를 포함한 수요실적에 의해 익년도 조달량은 실제 필요한 소요보다 많게 되고, 보급이 원활하면 가수요는 감소되어 다음 해의 소요량은 적어지게 되는데 이러한 과정이 반복되면 수요 추세는 안정될 수 없다. 수요추세를 안정시키는 가장 중요한 방법은 실소요를 반영하는 것이다. 따라서 수리부속품의 실제 교체·정비된 현황이 실수요라고 가정하였다. 자료는 항공기 정비정보체계(Aircraft Maintenance Management Information System : AMMIS, 이하 AMMIS)에서 산출된 값이며, 제00전투비행단의 항공기 수리순환 품목의 교체된 현황을 토대로 수집한 다수 결함이 발생한 50여개 품목의 DATA이다.

정비사들은 실시간으로 정비내역을 항공기정비정보체계(AMMIS) 상에 기록하게 되어 있는데, 부품의 장탈착한 시간이 곧 그 부품의 교체 시간이 되고, 월별 교체된 실적이 곧 실수요인 것이다. 항공기정비정보체계에서는 부품의 P/N(Part Number)별로 관리를 하고 있으며, 분기별로 교체한 시간을 종합하면 해당 분기당 부품의 사용한 실제 수요치를 얻을 수 있게 된다.



[그림 1] Box-Jenkins 수요예측 세부 절차

실증분석에 앞서 Box-Jenkins의 ARIMA 모형 구축 절차를 다음 [그림 1]과 같이 요약하였다.

4.2 수리부속의 ARIMA 모형

4.2.1 모형의 식별

모형의 식별은 ARIMA(p,d,q)의 차수를 결정함으로써 적절한 모형을 찾는 것이다. 주어진 데이터의 시계열도과 자기상관함수(ACF), 부분자기상관함수(PACF), 역 자기상관함수(IACF)를 통해 데이터의 특성을 알 수 있다. 부분자기상관함수와 역 자기상관함수는 AR모형의 차수를 결정하는데 주

Lag	Covariance	Correlation	Autocorrelations																										
			-1	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
0	330.911	1.00000																											
1	189.237	0.57187																											
2	160.911	0.48627																											
3	102.679	0.31029																											
4	83.45647	0.25220																											
5	104.883	0.31698																											
6	73.74795	0.22288																											
7	65.09682	0.19673																											
8	29.63821	0.08956																											
9	24.14021	0.07295																											
10	73.75322	0.22288																											
11	72.36736	0.21859																											
12	105.310	0.31624																											
13	54.591224	0.16497																											
14	46.068234	0.13828																											
15	13.124941	0.03906																											
16	6.104255	0.01845																											
17	24.130546	0.07292																											
18	-2.797675	-0.0845																											
19	-23.389728	-0.7068																											
20	-52.380373	-1.5829																											
21	-54.907710	-1.6693																											
22	-56.245598	-1.6698																											
23	-49.173225	-1.4860																											
24	-42.090455	-1.2720																											

[그림 2] 특정부품의 자기상관함수

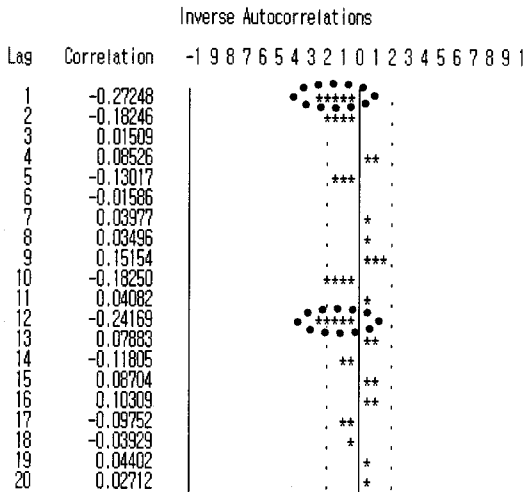
로 쓰이며, 자기 상관함수는 MA모형의 차수를 결정하는데 사용된다. 공군의 특정 부속품('99.1~'07.2)의 수요예측을 실증분석으로 제시한다.

4.2.1.1 상관함수를 이용한 모형식별

시계열의 자기상관함수(ACF)는 다음 [그림 2]와 같다.

[그림 3]에서 자기상관함수는 곡선을 그리며 점차 감소하다가 12시점마다 튀어 나오는 형태를 가지고 있다. 절단(Cut-Off)된 모습이 아니므로 비정상성을 가진 시계열이다. 6시점(Lag) 이후부터는 $\pm 5\%$ 신뢰구간(점선) 내에 드는 형태를 보이고 있어 12시점마다 순환성을 보이지만 계절성이 있다고 하기에는 객관성이 떨어진다. 부분자기상관함수와 역자기상관함수를 확인하여 계절성 여부를 판단한다. 역자기상관함수는 다음 [그림 3]와 같다.

위 그림에서 보는 바와 같이 1시점과 12시점에서 계절성이 보인다. 부분자기상관함수는 역자기상관함수와 역의 관계를 가지고 있으므로 생략한다. 역자기상관함수와 부분자기상관함수를 통해 비정상성과 계절성을 확인하였다. 자기상관함수와 부분자기상관함수, 역자기상관함수의 목적으로도 시계열이 백색잡음 과정이 아님을 알 수 있



[그림 3] 특정부품의 역자기상관함수

Autocorrelation Check for White Noise

To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations																			
6	85.04	6	<.0001	0.579	0.467	0.321	0.238	0.284	0.192														
12	104.51	12	<.0001	0.143	0.082	0.056	0.162	0.187	0.274														
18	109.97	18	<.0001	0.163	0.124	0.051	-0.015	0.022	-0.041														
24	130.16	24	<.0001	-0.084	-0.170	-0.184	-0.179	-0.163	-0.164														

[그림 4] 포트맨토우 검정 결과

지만 보다 객관성을 갖기 위하여 다음 [그림 4]의 포트맨토우 검정으로 확인한다.

포트맨토우 검정(Portmanteau Test)은 잔차가 백색잡음과정을 따르는지를 검정하기 위하여 다음 식(30)과 같은 검정 통계량을 제시한다.

$$Q^* = n(n+2) \sum_{j=1}^K \frac{\hat{\rho}_a^2(j)}{(n-j)} \quad (30)$$

유의수준 α 에서 만일 $Q^* > \chi_\alpha^2(K-p-q)$ 이면 a_t 가 백색잡음이라는 귀무가설은 기각되어 잠정 모형은 부적합하며, $Q^* < \chi_\alpha^2(K-p-q)$ 이면 귀무가설을 기각할 수 없으므로 잠정모형을 적합하다고 판정한다. SAS에서는 통계량이 Chi Square로 표시된다. Q^* 값이 유의하다면 ($\text{Prob} < \alpha$) 잠정모형이 부적합함을 나타내고 있는 것이다. [그림 4-5]

에서 유의수준이 모두 0.05보다 작게 나타나고 있다. 따라서 원시계열 자료의 잔차가 백색잡음이 아니라는 사실이 확인된다. 만일 백색잡음이면 추정된 모형의 잔차가 확률적 요소만 남기 때문에 모형이 적합하므로 모형추정을 중단한다. 포트맨토우 검정은 모형을 구축하는 과정에서 계속에서 모형의 적합성을 판단하는 기준이 된다.

4.2.1.2 시계열의 정상성 확인

확률적 추세가 존재하는 경우에는 시계열도와 자기상관함수 만으로는 정규성 판단이 어려운 경우가 많다. 정규성 판단에는 대표적으로 Dickey-Fuller(이하 DF검정)검정이 있다. DF검정은 순수한 확률보행모형(Zero Mean), 절편이 있는 확률보행모형(Single Mean), 결정적 추세와 절편이 있는 확률보행이 혼합된 모형에 대한 검정이다. 시차변수 Y_{t-1} 에 대한 Y_t 의 회귀모형을 고려하여 시차변수의 회귀계수가 1인지 아닌지에 대한 유의성 검정이라 할 수 있다. 회귀계수가 1이면 단위근을 가지고 있으므로 차분으로 정상시계열로 변환해야 한다. D-F 검정 결과는 다음 [그림 5]와 같다.

H_0 는 Y_t 가 단위근에 의한 비정상성을 가지고 있다는 것이고, H_1 은 Y_t 가 정상성을 가지고, 잔차는 정상적인 오차(백색잡음)로 구성되어 있다는 것이다. 만일 이 확률이 ≥ 0.05 이면 H_0 는 기각될 수 없으며, 1차 차분이 계열을 정상화하는데 필요하게 된다. 위 그림에서 원 시계열은 단위근

Augmented Dickey-Fuller Unit Root Tests

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau
Zero Mean	0	-6.1616	0.0047	-1.80	0.0683
	1	-3.3459	0.0067	-1.39	0.1510
Single Mean	0	-39.2477	0.0009	-4.82	0.0002
	1	-24.5073	0.0024	-3.31	0.0171
Trend	0	-40.5322	0.0003	-4.82	0.0009
	1	-25.9251	0.0135	-3.32	0.0686

[그림 5] D-F 검정 결과

을 가진 비정상 시계열임을 알 수 있다. t 검정은 모형에 상수항과 결정적 추세에 있을 때 사용되는 유의성 검정으로 $Pr<Tau=0.0616$ 은 $H_0(\delta = 0)$ 를 기각할 수 없음을 보여준다.

DF검정에서 단위근에 의한 비정상성을 가진 것으로 판단되어 1차 차분을 실시하고, 상관함수에서 보였던 계절성을 제거하기 위하여 계절차분을 실시하였다. 로그변환이 유의한지를 판단하는 로그테스트에서는 통계량이 유의하지 않다. 다시 D-F검정을 통해 확인 결과 차분만으로 정상성을 확보하였다.

4.2.1.3 모형의 차수결정

SAS에서는 AR(p+d)와 MA(q)모형의 차수를 0부터 5(사용자 지정)까지 시뮬레이션하여 모형의 차수를 보여준다. 차수결정의 기준이 되는 통계량은 AIC(Akaike 정보판단기준)와 SBC(Schwarz 베이지안판단기준)이다. 차수가 높은 ARIMA 모형에 시계열을 적합시키면 오차항의 분산에 대한 추정량인 평균제곱오차(MSE)는 작아지지만 모형의 차수가 커져 모형 식별에서의 모수절약의 원칙에 위배된다. 따라서 평균제곱오차에 높은 차수를 택하는 것에 대한 위험을 추가시킨 통계량으로 시뮬레이션 한다. 그 결과는 다음 [그림 6]과 같다.

```

ARMA(p+d, q)
Tentative
Order
Selection
Tests

----ESACF----
p+d      q
  0       1
  1       1
  2       1
  3       1
  4       0
  5       0
    
```

(5% Significance Level)

[그림 6] 잠정적 모형의 차수 결정

위 그림은 유의수준 5%에서 모형의 차수를 식별한 결과이다. 6가지 경우로 축약되어 있는데, 모수 절약의 원리를 이용하여 차수가 낮은 MA(1)모형을 우선 검토한다.

잠정적인 모형은 다음 식(31)와 같다.

$$ARIMA(0,1,1)(0,1,1)_{12} \Rightarrow (1-B)(1-B^{12})Z_t = (1-\theta B)(1-\Theta B^{12})a_t \quad (31)$$

4.2.2 모수 추정

ARIMA모형의 자기회귀차수(p), 차분차수(d), 이동평균차수(q)가 결정되면 모수를 추정한다. 이 과정은 추정된 모수가 얼마나 정확히 추정되는냐에 따라서 시계열의 자료를 분석하는데 많은 영향을 미친다.

모수의 추정방법에는 적률 추정법, 조건부 최소 제곱법, 비조건부 최소 제곱법, 최대 우도법등의 다양한 방법이 있으나 자료에 계절차분과 또 다른 차분을 동시에 고려할 경우 많은 관측값들이 결측치가 되기 때문에 조건부 최소제곱법은 사용하지 않기로 한다. 따라서 본 연구에서는 최대 우도법(ML)을 이용하여 모수를 추정한다. 최대우도법을 이용한 추정 결과는 다음 [그림 7]과 같다.

MU는 질편을 의미하는데 유의 확률이 0.05보다 크기 때문에 “평균이 ‘0’이다” 라는 귀무가설을 기각할 수 없다. 따라서 평균이 ‘0’인 모형으로 최대 우도법을 적용하여 추정한다. 추정 결과는 다음 [그림 8]과 같다.

Maximum Likelihood Estimation

Parameter	Estimate	Standard Error	t Value	Approx Pr > t
MU	-0.12968	0.33015	-0.39	0.6945
MA1,1	0.47774	0.09404	5.08	<.0001
MA2,1	0.85592	0.22439	3.81	0.0001

[그림 7] 최대우도법 모수 추정 결과

Maximum Likelihood Estimation

Parameter	Estimate	Standard Error	t Value	Approx Pr > t
MA1,1	0.48111	0.09294	5.18	<.0001
MA2,1	0.86410	0.23403	3.69	0.0002

[그림 8] 최대우도법 모수 추정 결과

평균값의 추정치를 제거한 모형에서 나머지 모수들의 유의 확률이 모두 0.05보다 작기 때문에 추정된 모수는 통계적으로 유의하다. 추정된 모수의 값은 다음과 같다.

$$\theta = 0.48111, \Theta = 0.8641$$

4.2.3 모형의 진단

모형의 진단은 주로 우차, 즉 백색잡음의 추정량인 잔차를 중심으로 이루어진다. 잔차가 오차에 주어졌던 백색잡음 가정을 만족하게 되면 시계열 분석은 종료된다. 그렇지 않을 경우 추정된 모형이 자료를 잘 설명하지 못한다는 의미이므로 새로운 모형을 식별하는 처음 단계로 돌아간다.

4.2.3.1 잔차분석

먼저 모수 추정과정에서 유의하다고 판단된 모형의 잔차를 분석하여 추정된 모형이 자료를 잘 설명하는지 확인한다. 잔차 분석에서 좋은 결과를 얻으면, 추정된 모형에서 자기회귀 차수와 이동평균의 차수를 한 단계 높여 과대적합을 실시한다. 이상의 모형 진단이 끝난 모형은 예측에 사용할

Autocorrelation Check of Residuals

To Lag	Chi-Square	DF	Pr > ChiSq	-----Autocorrelations-----									
6	2.27	4	0.6966	0.034	0.014	-0.094	-0.044	0.032	-0.064				
12	7.36	10	0.6907	-0.108	-0.076	-0.113	0.102	-0.011	0.106				
18	8.93	16	0.9162	-0.014	0.043	-0.002	0.022	0.106	-0.025				
24	16.03	22	0.8142	-0.000	-0.119	-0.021	-0.069	-0.020	-0.200				

[그림 9] 포트맨투우 검정 결과

Autocorrelation Plot of Residuals

Lag	Covariance	Correlation	-1	9	8	7	6	5	4	3	2	1	0	1	2	3	4	5	6	7	8	9	
0	218.861	1.00000																					
1	7.416197	0.03389										*											
2	3.107320	0.01420																					
3	-20.598070	-.09411										**											
4	-9.646199	-.04407																					
5	20.150880	0.09207																			**		
6	-14.116130	-.06450										*											
7	-23.532413	-.10780										**											
8	-16.717198	-.07638																			**		
9	-24.709684	-.11290										**											
10	22.394031	0.10232											**										
11	-2.480269	-.01133																					
12	23.301703	0.10647											**										
13	-3.128468	-.01429																					
14	9.478863	0.04331											*										

[그림 10] 자기상관함수

수 있는 적합한 모형이다. 잔차의 포트맨투우 검정 결과는 다음 [그림 9]와 같다.

위 그림에서 유의 수준이 모두 0.05보다 높다는 것을 발견할 수 있다. 이는 '잔차의 상관계수가 0이다'라는 귀무가설을 기각할 수 없다. 따라서 상관계수가 '0'이면 잔차가 비자기상관의 성질을 보여주기 때문에 기본 가정을 만족하여 모형이 잘 선택되고 있다. 잔차의 자기상관함수는 다음 [그림 10]과 같다.

자기상관함수가 1시점부터 5% 유의수준 이내에 존재하므로 백색잡음과정이 되었다. 원시계열에서 차분과 계절차분을 하여 추세와 계절성을 제거하고 이동평균을 실시하여 얻은 모형의 잔차는 자기상관이 없는 확률적 요소이다. 잔차의 부분자기상관함수는 다음 [그림 11]과 같다. 역자기상관함수는 생략한다.

Partial Autocorrelations

Lag	Correlation	-1	9	8	7	6	5	4	3	2	1	0	1	2	3	4	5	6	7	8	9		
1	0.03389										*												
2	0.01306																						
3	-0.09516										**												
4	-0.09821										*												
5	0.09882											**											
6	-0.08002										**												
7	-0.11736										**												
8	-0.05009										*												
9	-0.11324										**												

[그림 11] 부분자기상관함수

Maximum Likelihood Estimation					
Parameter	Estimate	Standard Error	t Value	Approx Pr > t	Lag
MA1,1	-0.03999	0.10921	-0.37	0.7142	2
MA2,1	0.68198	0.13016	5.23	<.0001	12

[그림 12] 최대우도법 결과

Maximum Likelihood Estimation					
Parameter	Estimate	Standard Error	t Value	Approx Pr > t	Lag
MA1,1	0.59672	0.17819	3.35	0.0008	1
MA2,1	0.83200	0.19572	4.25	<.0001	12
AR1,1	0.13573	0.21153	0.64	0.5211	1

[그림 13] 최대우도법 결과

4.2.3.2 과대적합

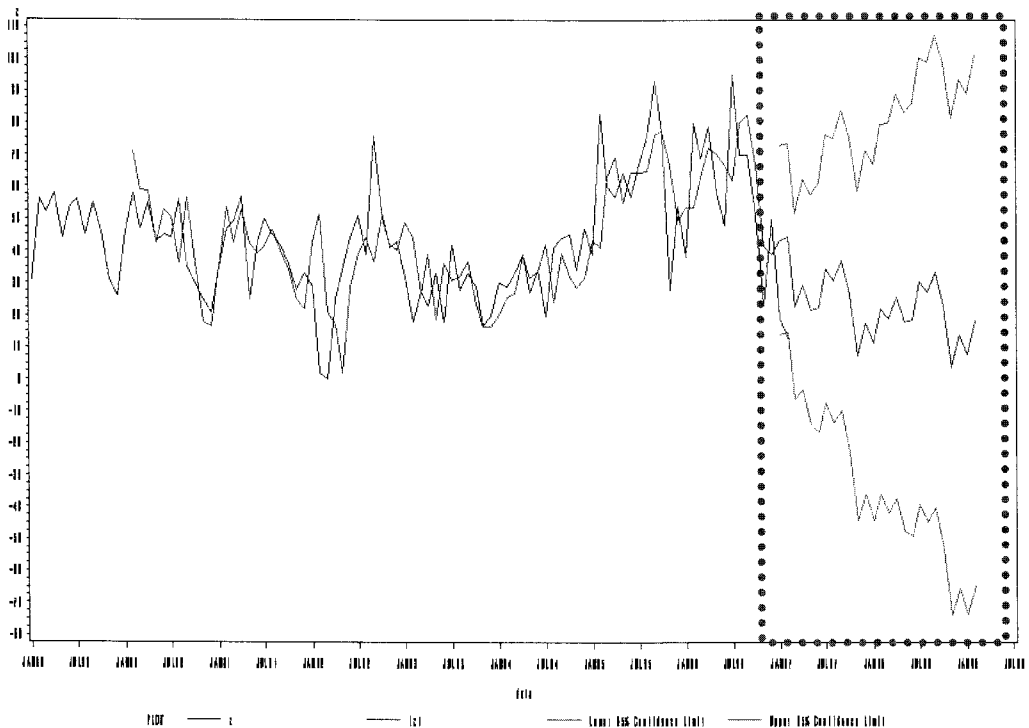
차분의 차수는 그대로 유지하고 이동평균의 차수를 한 단계 높은 모형($p=0, q=2$)과 자기회귀차수를 함께 높은 모형($p=1, q=2$)을 고려하여 과대적합을 실시한다. 이동평균의 차수만을 높은 모형의 최대우도법 추정 결과는 다음 [그림 12]와 같다.

모수의 추정결과 유의 수준이 0.05보다 높은 0.7142로 나타나게 된다. 이는 모수가 유의하지 않은 것이므로 이동평균 차수($q=2$)가 적합하지

않다. 자기회귀차수 및 이동평균 차수를 함께 고려한 과대적합 결과는 다음 [그림 13]과 같다.

과대적합된 자기회귀차수에서 유의 수준이 0.05보다 높아 유의하지 않다.

따라서 각각의 과대적합에서 P-value가 유의수준보다 높게 나타나기 때문에 과대추정된 모형이 적합하지 않다. 그러므로 과대적합을 하기 이전에 추정된 모형이 적합하며 추정된 모형은 다음 식(32)와 같다.



[그림 14] 예측결과 포함된 시계열도

$$ARIMA(0,1,1)(0,1,1)_{12} \Rightarrow \quad (32)$$

$$(1-B)(1-B^{12})Z_t = (1-\theta B)(1-\Theta B^{12})a_t$$

$$Z_t = Z_{t-1} + Z_{t-12} - Z_{t-13} + a_t \\ -0.48111a_{t-1} - 0.8641a_{t-12} + 0.41573a_{t-13}$$

4.2.4 모형의 예측

시계열 자료를 분석하는 궁극적인 목적은 미래의 값을 예측하는 것이다. 예측의 정확도는 실제 관측값과 예측값의 오차가 작은 것으로 판단할 수 있다. 현용 기법과 ARIMA 기법의 비교에 앞서 예측을 실시하고 평균제곱오차(MSE)를 비교한다. 예측된 값을 포함한 시계열도는 다음 [그림 14]와 같다.

오른 쪽의 점선으로 표시된 곳이 예측된 값의 시계열 그림이다. 가운데를 중심으로 위, 아래에 그려진 선은 95%의 신뢰구간 내의 예측값을 표현하고 있다. 그림에서 예측된 값들은 하강 추세를 보이고 있으며, 계절성을 가진다. 현용 기법과의 비교를 위하여 예측오차를 산출한 결과는 다음 [그림 15]와 같다.

OBS	_TYPE_	_FREQ_	sa_p1
1	0	86	235.108

[그림 15] 평균제곱오차 산출 결과

86개의 관측치에서 평균제곱오차(MSE)를 산출한 값은 235.108이다. 이 품목에서 현용예측기법 중 가장 작은 평균제곱오차는 가중이동평균법으로 계산된 240.47로 ARIMA 기법의 평균제곱오차 235.108보다 큰 값을 가진다. 표본으로 추출한 품목에 대한 적용결과는 다음 [표 4]와 같다.

54개 부품의 모형 구축 과정을 전부 게재하는 것은 지면 관계상 생략한다. 한 부품에서 여러 가지 모형이 가능할 경우에는 모형 절약의 원칙과 최소 예측오차(MSE)를 만족하는 모형을 채택하

[표 4] 표본의 모형 구축 결과

순번	ARIMA (p,d,q)	MSE (B)	예측값 (07.3)
1	(0,0,1)	21.6879	3.41
2	(0,1,1)	12.175	9.60
3	(0,1,1)(0,1,1)12	235.108	22.86
4	(0,1,1)	4.9996	2.22
5	(0,0,1)	2.42402	1.41
6	백색잡음 ($\sigma^2 2.672354$)	3.36415	1.38
7	(1,0,1)	1.11389	0.944
8	(2,1,2)	4.34997	5.6879
9	(1,0,0)	10.998	1.487
10	백색잡음 ($\sigma^2 1.405926$)	1.23056	0.8969
11	(0,0,1)	2.49748	1.2867
12	(0,1,1)	53.3748	19.18
13	백색잡음 ($\sigma^2 1.582426$)	4.83394	1.16
14	(0,0,1)	0.79706	1.0768
15	(1,0,1)	21.3697	1.5998
16	(5,1,0)	0.75166	1.763
17	(1,0,0)	3.76715	4.96
18	백색잡음 ($\sigma^2 1.85175$)	1.90176	1.3829
19	(0,1,1)	43.659	14.97
20	(0,0,1)	4.0037	3.922
21	백색잡음 ($\sigma^2 3.686$)	3.66943	1.7245
22	백색잡음 ($\sigma^2 4.076$)	4.06705	2.0816
23	(2,0,2)	9.40117	3.331
24	(1,0,1)	9.34558	3.967
25	(0,1,1)	12.4148	2.4507
26	(1,0,0)	14.7038	1.7199
27	(0,1,1)	1.6091	4.0175
28	(0,0,1)	20.7524	13.81
29	(0,1,1)	8.96715	7.2
30	(0,0,1)	37.645	5.27
31	백색잡음 ($\sigma^2 2.9708$)	2.96236	1.908
32	(0,1,1)	1.5784	1.279
33	(1,0,1)	1.31091	1.402
34	(0,0,2)	18.9559	2.495
35	백색잡음 ($\sigma^2 1.1136$)	1.10383	1.0309
36	(1,0,1)	7.92051	2.7725
37	(0,0,1)	10.465	2.71
38	(0,1,1)	4.89126	4.767
39	(1,0,1)	247.516	4.32
40	백색잡음 ($\sigma^2 1.3158$)	1.15586	1.0612
41	(1,0,0)	22.7848	7.73
42	(0,1,1)	25.1939	10.75
43	(1,0,0)	1.12825	1.0324
44	(0,0,1)	3.31493	4.352
45	(1,0,1)	12.5062	6.384
46	(0,0,1)	3.31909	4.00
47	(0,0,1)	5.14773	4.1525
48	(1,0,0)	9.9748	5.0887
49	(0,1,1)	2.43558	2.512
50	(0,0,1)	3.5766	4.82
51	(2,0,1)	3.47657	4.89
52	(0,0,1)	3.20037	0.7673
53	백색잡음 ($\sigma^2 2.1474$)	2.1261	1.556
54	(1,0,1)	4.72534	1.282

였다. 백색잡음으로 판명된 부품은 월별 수요 발생 실적이 추세가 없고, 분산이 일정하기 때문인데, 현용기법 중 추세분석과 지수평활($\alpha=0.1$)로 예측되는 부품들이다. 이 부품들의 다수에서 추세 분석법이 가장 작은 예측오차를 나타내고 있지만 기울기가 거의 없는 평이한 모습을 보인다. 계절성을 가진 품목은 한 가지 품목이다. 1차 차분으로 정상성을 확보할 수 있는 품목이 다수이며, 로그변환과 제곱근변환 등을 활용하여 정상 시계열로 변환이 가능하다. 정상성 검정(DF검정)에서 단위근이 없는 품목은 자기회귀(AR)모형이나 이동평균(MA)모형으로 구축된다.

정비 소요가 갑자기 늘어나는 경우는 확률적인 추세가 아닌 결정적 추세를 포함한 것으로 그 시점을 개입모형으로 해석하여야 한다. 개입이 일어난 시계열 자료를 분석하기 위하여 개입의 발생 시점, 지속기간 및 종료시점을 알고 있어야 한다. 펄스의 형태를 띠게 되는 이런 현상은 확률적 모형이 아닌 결정적 변수로 모형화 할 수 있다. 하지만 본 연구에서는 정보의 부족으로 개입이 추가된 모형을 구축하지 못한다.

일반적인 개입모형의 형태는 다음 식(33)와 같다.

$$Y_t = \sum_t f(z_t) + N_t \quad (33)$$

Y_t 는 개입 전 ARIMA 모형의 함수와 각 개입에 대한 결정적 개입 척도의 입력함수의 합이라는 것이다. $f(x_t)$ 는 시점 t에서의 개입변수의 함수로 합으로 표시한 것은 모든 가능한 개입변수가 모형에 포함되어 있다는 의미이고, N_t 는 개입 전 ARIMA 모형이다. Box-Jenkins 방법이 현용기법보다 우수한 이유는 이러한 구축된 ARIMA 모형에 가까운 미래에 발생하는 외부의 개입에 대한 반응을 추가할 수 있다.

4.3 현용 기법과 ARIMA 기법의 결과 비교

[표 5] 현용기법과 ARIMA 기법 종합

구 분	예측오차 합 (평균제곱 오차)	'07. 3월 예측치 합	'07. 3월 관측치 합(C)
현용기법	1,056.3(A)	509.6(A')	208
ARIMA	957.8(B)	225.8(B')	
A - B	98.5		
A' - B'		283.7	
A' - C		301.6	
B' - C		17.8	

4.3.1 결과 종합

현용기법과 ARIMA 기법의 비교 결과는 다음 [표 5]와 같다.

이상과 같이 ARIMA 기법으로 예측을 실시하여 MSE를 절감할 수 있었다. 총 8만여개의 부품 중 50여개의 부품에 대한 실증 분석에서 예측 오차를 줄일 수 있었다. 예측 오차(MSE)는 현용기법이 1,056.268이고, ARIMA기법이 957.7531이다. 예측오차 중 ARIMA 기법보다 현용 기법이 좋은 결과를 나타내는 부품이 12품목이지만 예측 오차(MSE)는 큰 차이를 보이지 않는다.

'07년 2월까지의 자료를 활용하여 3월의 예측값을 확인한 결과 현용기법에 의해 산출된 예측치('07. 3월)의 합은 509.5739(A')이며, ARIMA기법은 225.8318(B')로 두 기법의 예측치 차이의 합은 283.742(A'-B')이다. 양의 값이 의미하는 것은 현용기법이 과다수요를 예측하고 있어 재고 비용이 많다는 것이다. 실제 관측된 값과의 비교를 해서 과다수요인지의 여부를 확인한다. '07. 3월 관측된 부품의 수요 합(C)와 현용기법의 3월 예측값 합(A'), ARIMA기법의 3월 예측값합(B')의 차이는 각각 현용기법이 301.5739(A'-C)이고, ARIMA기법이 17.83178(B'-C)이다. 이는 ARIMA 기법이 현용기법보다 실제 수요변동을 잘 설명한다는 증거이다. 이에 본 연구에서는 ARIMA 기법을 재고 비용을 줄일 수 있는 기법으로 추천한다. ARIMA 기법으로 예측한 값이 관측치 보다 작게 나타나

는 경우 해당 부품의 보급지연으로 항공기 불가동이 예상된다. 특히 고단가 품목일 경우 청구기간이 길고, 예산 확보가 되어 있지 않은 경우가 많다. 이는 평소 적정 예비품(Spare)을 확보하는 정책에 의해 해결 할 수 있다.

ARIMA 기법으로 추정된 모형과 기존 기법이 일관성을 가지고 연결되지는 않는다. 예를 들어 선형지수평활법에 적합한 부품이 ARIMA기법에서는 (1,0,0), (2,0,2), (0,0,1) 등으로 나타난다. 단지 부품 별 예측오차가 작은 모형이 예측의 효율을 높일 수 있다.

종합하면 예측오차의 합은 ARIMA 기법이 작기 때문에 현용기법은 ARIMA 기법에 비해 예측 정확성이 낮다(A > B). 또한 현용기법 및 ARIMA기법 예측치 합과 실제 관측치 합과의 차를 비교하면 현용기법에서 수요가 과다추정 되었음을 알 수 있다(A' > B' ⇔ A'-C > B'-C).

ARIMA 기법의 우수성은 다음과 같다.

현용 기법이 과거의 관측값들의 결합으로 표현되는 반면 ARIMA 기법은 과거의 관측값들과 백색잡음이라고 부르는 오차들의 선형함수를 설명 변수로 사용하고 있다. ARIMA 기법은 부품별로 적절한 모형을 구축하여 차후 개입모형을 추가할 수 있어 예측력을 향상시킬 수 있다는 장점이 있다. 또한 급격한 변화를 보이는 시계열에서 예측의 정확성이 타 기법에 비해 높은 것으로 알려져 있다. 항공기 수리부속품 중 수요의 변화가 급격하고, 고단가이며 다수결합이 발생하는 품목에 대해 ARIMA 기법을 적용할 것을 추천한다. 결합 발생률이 낮은 품목에 대한 예측은 현용 기법 중 가장 예측 오차가 적은 지수평활법을 적용할 것을 추천한다[3].

4.3.2 적합도 검정(쌍체 T-검정)

본 연구에서 Box & Jenkins의 ARIMA기법이 현용 수요예측기법보다 정확도가 높은 기법임을

검정하기 위한 방법으로 쌍체 T검정(paired T test)을 통한 적합도 검정을 실시하였다. 쌍체 T검정은 하나의 모집단으로부터 추출된 짝을 이룬 두개의 표본을 상호 비교하는 검정법이다.

표본의 크기 N=54이며 표본의 평균과 표준편차는 식(34)과 식(35)와 같다.

$$\bar{d} = \frac{\sum d_i}{n} = \frac{98.5149}{54} = 1.82435 \quad (34)$$

$$S_d = \sqrt{\frac{\sum (d_i - \bar{d})^2}{n-1}} \quad (35)$$

$$= \sqrt{\frac{1566.161}{53}} = 5.436$$

ARIMA기법의 MSE가 작다는 것을 검정해야 하므로 가설은 다음과 같다.

$$\cdot H_0 : \mu_d (= \mu_1 - \mu_2) \leq 0 \quad (36)$$

(ARIMA기법의 예측오차가 기존기법의 예측오차보다 크다)

$$\cdot H_1 : \mu_d (= \mu_1 - \mu_2) > 0 \quad (37)$$

(ARIMA기법의 예측오차가 기존기법의 예측오차보다 작다)

여기서 μ_1 과 μ_2 는 각각 기존기법과 ARIMA기법의 예측오차(MSE)이다.

검정통계량(test statistic)은 다음 식(38)과 같다.

$$t = \frac{\bar{d} - \mu_{d_0}}{S_d / \sqrt{n}} = \frac{1.82435 - 0}{5.436 / 7.348} = 2.466 \quad (38)$$

유의수준(significance level) $\alpha=0.05$, 자유도(degrees of freedom) $n-1=53$ 에 대한 기각역은 다음 식(39)와 같다.

$$t_{53, 0.05} = t_{\infty, 0.05} = 1.645 < t = 2.466 \quad (39)$$

통계량 $t = 2.466$ 은 기각역 내에 있으므로 귀무가설은 기각(reject)된다.

즉, ARIMA에 의한 예측치의 MSE가 이동평균법에 의한 MSE보다 작다고 할 수 있다. 따라서 현용 기법의 분석 보다 ARIMA기법이 우수한 예측기법임을 추천 할 수 있다.

5 결론 및 향후 연구방향

전·평시 성공적인 공군의 작전요구도를 만족시키기 위하여서는 효과적인 정비지원과 보급지원이 이루어져야 한다. 본 연구는 보급지원 측면에서 항공기 목표운영률을 달성하기 위하여 보급지원의 측면에서 현용 수요예측기법을 개선하는데 목적이 있다.

공군의 수요예측 기법은 가중/선형 이동평균법, 추세분석법, 지수/선형지수 평활법 5가지이다. 본 연구에서는 5가지 방법으로 수요예측을 실시하고 있는 86,620개 부품 중 다수 결함이 발생하는 부품 50여개를 선정하여 적합한 ARIMA 모형을 산출하고 수요예측을 실시하였다.

자료는 실수요를 반영하고자 수요 신청자의 현황을 자료로 사용하는 기존의 방식에서 항공기 정비정보체계(AMMIS)으로부터 실제 수리부속의 교체가 일어난 정비실적을 자료로 수집하였다. 또한 현용 수요예측기법은 분기별 자료 20개(과거 5년치)를 이용하여 예측치를 산출하지만, ARIMA 기법에 적합시키기 위하여서는 자료의 수가 최소 50개 이상이 되어야 하므로 월별 고장 실적 98개의 자료(1999년 1월 ~ 2007년 2월)로 수요를 예측하여 정확도를 제고 하였다.

또한 현용기법으로는 수리 모형을 만들어 수요를 예측할 수 없었지만, ARIMA 기법으로 최적의 모형을 식별하여 수요를 예측할 수 있었으며, 현용 기법들 보다 예측 오차가 작다는 결과를 도출할 수 있었다. ARIMA 모형을 구축하는 가장 중요한 단계는 자료를 정상시계열로 변형하는 단계이다. Trend가 존재하는 경우 차분을 통하여 비

정상성을 상쇄시킬 수 있으며, 분산이 불안정 한 경우 자료에 대해 LOG 변환, 제곱근 변환 등을 통해 정상성을 가지게 하였다. 정상성을 만족하는 자료를 만들어 준 다음에는 시계열도와 자기상관 함수를 이용하여 자기회귀차수(p)와 이동평균차수(q)를 결정하였다. 모형의 차수를 결정한 이후 최대우도법(ML)으로 유의성을 가진 모수를 구할 수 있었다. 마지막으로 예측을 실시하고, 평균제곱오차(MSE)를 산출하여 기존 기법과 비교하여 ARIMA 기법이 적절한 예측모형임을 제시하였다.

본 연구의 제한 사항은 다음과 같다. 첫째, 자료의 수는 50여개의 부품으로 분석을 실시하였기 때문에 그 결과가 상당히 제한적으로 적용된다고 판단된다. 그러므로 운용기간이 짧은 항공기의 경우에는 ARIMA 기법을 적용할 수 없다. 그 대안으로 단순 지수평활법을 추천한다[3]. 자료의 수가 25개 내외로 작은 경우 예측오차가 다른 기법에 비해 예측오차가 작은 것으로 밝혀졌다. 둘째, 갑작스러운 수요의 증가를 파악하지 못하여 개입 모형을 추가하지 못하였다. 결국 항공기의 운영률에 영향을 미치는 보급지원 요소는 수요추세를 예측하지 못해서가 아니라 갑작스런 수요의 발생 원인을 파악하지 못하거나 행정상의 착오, 보급라인에서 발생하는 지연 등이다.

많은 학자들이 예측기법의 연구를 계속 해오고 있지만 완벽한 예측기법이란 존재하지 않는다. 다만 예측의 오차를 줄일 수 있는 기법들을 통해 최소한의 비용을 지불하고자 하는 것이다. 이에 본 연구는 공군의 수리부속품 중 다수결함이 발생하는 일부 품목에 대해 현용 예측기법보다 예측 정확도가 높은 ARIMA기법을 추천하였다.

향후 연구의 방향은 다음과 같다. 첫째, 고단가이며 청구기간이 길어 항공기 불가동의 원인이 되고 있는 주요부품에 대한 품목 특성별 예측방법의 연구가 이루어져야 한다. 주요 품목과 단순 소모성 품목은 분리하여 관리되어야 하며, 실무에서는 관심을 갖고 정확한 재고조사와 품목별 특성파악을 위한 노력이 병행되어야 한다. 둘째, 예

측의 정확성을 향상시킬 수 있는 방향으로 항공기 부품(LRU)별 신뢰도 모형을 추가한 예측모형을 구축할 것을 제안한다. 현재 공군에서는 각 부품별 신뢰도를 결정하기 위하여 고장간평균시간(MTBF)를 수집 중이다. 실제 당시의 부품의 신뢰도는 운영환경에 많은 영향을 받는다. 항공기 부품은 대부분이 수입되기 때문에 한국의 환경(계절, 습도, 기온)과 한국 공군의 작전 임무도에 따라 설계당시의 신뢰도는 달라진다. 그러므로 공군에서 수집된 부품별 수명주기를 고려한 신뢰도 모형을 수요예측 모형과 접목하여 부품의 수요 예측력을 향상시킬 것으로 기대된다.

끝으로 최적의 예측 결과를 얻을 수 있는 논리적 근거를 제시한다면 어떠한 분석 기법을 사용하든지 무관하다. 예를 들어 회귀분석의 설명력이 70%이고 시계열이 75%라고 한다면 회귀분석에 적절한 독립변수를 추가하여 5% 향상된 결과를 얻을 수도 있을 것이다. 따라서 수요예측의 정확도를 높이기 위해서는 통계적 모형 보다는 현실을 반영할 수 있는 올바른 데이터의 축적과 수집이 무엇보다 중요하다.

참고문헌

[1] 고영복, "Box-Jenkins Model을 이용한 합정 보조기관의 경제적 교체시기 연구", 국방대학교, 1992.
 [2] 김동복외 2명, "국내 중급항공기 수요예측", 경영과학, Vol. 8, No2, 1991, pp37~47.
 [3] 김수연, "시계열 분석에 의한 수리부속 수요예측기법에 관한 연구", 국방대학교 석사학위논문, 2005.
 [4] 김진용, "보급품의 수요예측에 관한 발전 방안", 전남대학교 석사학위논문, 1998.
 [5] 김희철, "K-1전차 수리부속품 최적소요산정에 관한 연구", 국방대학교 석사학위논문, 1999.
 [6] 박미경, "수요예측기법의 신뢰성에 대한 사례 연구", 상명여대 석사학위논문, 1993.
 [7] 박창형, "승법계절 ARIMA 모형에 의한 에너지 수요예측", 국방대학교 석사학위논문, 1998.
 [8] 이성덕, 이우리, "SAS를 이용한 시계열자료분

석", 탐진, 2006.
 [9] 이종협, "시계열분석과 응용", 자유아카데미, 2007.
 [10] 임진한, "군 수요예측 기법의 효과적 적용방안에 관한 연구", 국방대학교 석사학위논문, 2001.
 [11] 전덕빈, "임의의 수준변화에 적절히 반응할 수 있는 지수이동가중평균법", 대한산업공학회, Vol. 16, No.2, 1990, pp129~134.
 [12] 정우익, "회전의 항공기 수리부속 수요예측에 관한 연구", 국방대학교 석사학위논문, 1998.
 [13] 조성익, "군수물자의 수요예측기법 선정에 관한 연구", 한양대학교 석사학위논문, 1998.
 [14] Brown R. G., "Smoothing Forecasting and Prediction of Discrete Time Series", Prentice-Hall:New-Jersey, 1962.
 [15] Brown R. G. and Meyer R. F., "The Fundamental Theorem of Exponential Smoothing", Operations Research, Vol. 9, 1961, pp.673~685.
 [16] Douglas C. Montgomery and Lynwood A. Johnson, "Forecasting and Time Series Analysis". New York ; MCGRAW-HILL, 1976.
 [17] George E. P. Box and Gwilym M. Jenkins, "Time Series Analysis Forecasting and Control", Holden-Day, 1976.
 [18] Gordon T. J. and Helmer O., "Report on a Long-range Forecasting Study", The Rand Corp, 1964.
 [19] Graner C. W. J. and Newbold P., "Forecasting Economic Time Series", Academic Press, New-York, 1986.
 [20] Holt C. C., "Forecasting Trends and Seasonals by Exponentially Weighted Moving Averages", O. N. R. Memorandum, No.52, Carnegie Institute of Technology, 1957.
 [21] Segura J. V. and Vercher E. A., "Spreadsheet Modeling Approach to the Holt-Winters Optimal Forecasting", European Journal of Operatinal Research Vol. 131, 2001, pp375~388.
 [22] Shim J. K., Siegel J. G., and Liew C. J., "Strategic Business Forecasting", Probus Publishing Compny, 1994.
 [23] William H. Greene, "Econometric Analysis 2nd ed.", Macmillan, 1993.

..... | 저자 소개 |

박 영 진 (E-mail : pyj0029@hanmail.net)

2002 공군사관학교 졸업(학사)
2008 국방대학교 졸업(운영분석 석사)
현재 공군 제5전술공수비행단 근무
관심분야 시계열분석, Queueing, 신뢰도분석

전 건 옥 (E-mail : gwjeon@kndu.ac.kr)

1984 공군사관학교 졸업(학사)
1993 고려대학원 졸업(산업공학 석사)
1999 미국 Univ. of Louisville 졸업 (산업공학 박사)
현재 국방대학교 운영분석전공 부교수
관심분야 최적화기법 응용, 일정계획, 신뢰도분석, 셀형 제조시스템