

Article

심해저 망간단괴에서 추출되는 금속가격 예측 및 적합도 분석

권석재 · 정선영*

한국해양과학기술원 해양정책연구소
(426-744) 경기도 안산시 상록구 해안로 787

Analysis of Price Forecasting and Goodness-of-Fit of the Metals Extracted from Deep Seabed Manganese Nodules

Suk-Jae Kwon and Sun-Young Jeong*

Ocean Policy Institute, KIOST
Ansan 426-744, Korea

Abstract : The development of deep seabed manganese nodules has been carried out with the aim of commercial development in 2023. It is important to forecast the price of the four metals (copper, nickel, cobalt, and manganese) extracted from manganese nodules because price change is a criterion for investment decision. The main purpose of the study is to forecast the price of four metals using the ARIMA model and VAR model, and calculate the MAPE to compare a goodness-of-fit between the two models. The estimated results of the two models reveal statistical significance and are in keeping with economic theory. The results of MAPE for goodness-of-fit show that the VAR model is between 0.1 and 0.2, and the ARIMA model is between 0.4 and 0.6. That is, the VAR model is better than the ARIMA model in forecasting changes in the price of metals.

Key words : deep seabed manganese nodule, price forecasting, ARIMA, VAR, MAPE

1. 서 론

지난 30년 간 금속가격은 가격의 변화 추이를 예측하기 어려울 정도로 상승과 하락이 불규칙적으로 나타나고 있다. 현재 금속가격이 하락추세를 보이고 있다고 하지만, 갑작스럽게 가격이 반등할 가능성도 크기 때문에 금속수입의존도가 높은 우리나라는 이에 대한 대비가 필요하다.

4대 금속(구리, 니켈, 코발트, 망간)을 전량 수입에 의존하고 있는 우리나라는 금속가격의 상승 및 금속공급 불안 등이 발생하면 다른 나라보다 더 큰 타격을 받게 되므로 금속가격의 중·장기 전망은 필수적이라 할 수 있다. 즉, 정교한 금속가격 예측 모형이 이루어진다면 이를 활용하

여 신뢰성 있는 중·장기 금속가격 수준을 전망하고 이 결과는 태평양 심해저 광물자원 개발사업 등과 같은 경제성 분석에 기초자료로 제공되어 정확한 분석이 이루어질 수 있도록 해야 한다.

금속가격 예측에 관한 연구는 한국해양연구원에서 1992년과 2002년에 수행되었었다. 두 연구 모두 OLS (Ordinary Least Square) 모형을 이용하여 예측하였으며, 생산량, 산업생산지수, 더미변수 등을 사용하였다. 그러나 연구가 이루어진 시점이 10년이 넘게 지났으며 사용된 변수들도 현재시점에 적용하기에는 무리가 있어 새로운 예측모형이 필요한 시점이라고 할 수 있다.

금속가격은 광산을 소유하고 있는 국가의 정치적 상황이나 광산을 소유한 자본의 성격에 따라 비정상적 영향을 끼칠 요인이 많아 금속가격을 정확히 예측한다는 것은 쉽

*Corresponding author. E-mail : syjeong@kiost.ac

지 않다. 그럼에도 불구하고 금속가격의 예측이 중요한 이유는 다음과 같다. 태평양 광물자원 개발사업은 2023년 상용화를 목표로 연구가 수행 중에 있다. 한국해양과학기술원(2012)에 따르면 본 사업은 초기 자본비용은 약 4조 원, 연간 운영비용은 약 7천억원에 달하는 막대한 비용이 필요할 것으로 예측되었다. 이는 2012년 기준으로 산정된 비용이기 때문에 개발예정인 2023년에는 비용이 더욱 증가할 것으로 예상된다. 이처럼 개발비용이 매우 클 것으로 예상되는 태평양 심해저 광물자원 개발사업에서 얻을 수 있는 유일한 수익은 금속을 판매함으로써 얻는 수익이다.

본 연구에서는 ARIMA(Autoregressive Moving Average) 모형과 VAR(Vector Autoregressive) 모형을 사용하여 금속가격예측모형을 구축하고자 한다. 또한 MAPE(Mean Absolute Percentage Error)를 계산하여 모형간의 적합도를 비교하고자 한다.

2. 분석방법 및 설명

ARIMA 모형

ARIMA 모형은 Box-Jenkins 모형이라 하기도 하며 시계열의 확률적 특성 분석이 강조되며 종속변수 Y 가 자신의 과거 값들과 확률적 오차항으로 결정되는 모형이다(최 등 2007). ARIMA 모형은 AR 모형과 MA 모형이 결합된 형태로 AR 모형과 MA 모형은 다음과 같다.

AR(Autoregressive) 과정은 t 기의 Y 값이 과거의 Y 값에 직접적으로 의존하는 구조이며 일반적인 AR(p) 모형은 다음과 같은 형태로 표현된다.

$$Y_t = \delta + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t \quad (1)$$

위 식에서 t 시점에서의 Y 값은 과거 p 기간 동안의 Y 값 가중평균과 t 기에 있어서 무작위로 발생하는 오차항의 합으로 구성된다. 예를 들어 AR(1) 모형인 다음 식에서 t 기의 시계열 Y_t 값은 시차변수 Y_{t-1} 과 백색잡음 ε_t 에 의해서 설명된다.

$$Y_t = \delta + \beta_1 Y_{t-1} + \varepsilon_t \quad (2)$$

MA(Moving-average) 과정은 시계열 자료 Y_t 가 과거 백색잡음의 가중 합으로 형성되는 구조로 일반적인 MA(q) 모형은 다음과 같다.

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

t 시점에서의 Y 값은 과거 q 기간 동안 발생한 오차항의 합으로 표시된다. 예를 들어 다음의 MA(1) 모형에서 t 기의 Y 값은 t 기와 $t-1$ 기에 발생한 오차항에 의해서는 직접 영향을 받지만 그 이전기간의 오차항에 의해서는 영향을 받지 않는다.

$$Y_t = \delta - \theta_1 \varepsilon_{t-1} + \varepsilon_t \quad (4)$$

안정적인 확률과정들이 순수하게 AR이나 MA의 특성만을 갖는 경우도 있으나 많은 경우 두 과정의 특성을 동시에 지니고 있는 과정을 발견할 수 있다. 이처럼 AR(p)와 MA(q)가 혼합된 형태의 과정을 ARMA(p, q)라 표기하며 다음과 같다.

$$Y_t = \delta + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (5)$$

이 모형은 AR(p) 모형과 MA(q) 모형의 결합형태로서 시계열 Y_t 가 과거 p 시차의 시계열 관측값 및 q 시차의 확률오차항에 의해서만 영향을 받는다. 예를 들어 다음의 ARMA(1,1) 모형에서 t 기의 시계열 Y_t 값은 시차변수 Y_{t-1} 과 t 기와 $t-1$ 기에 발생한 오차항에 의해서 직접적인 영향을 받는다.

$$Y_t = \delta + \phi_1 Y_{t-1} - \theta_1 \varepsilon_{t-1} + \varepsilon_t \quad (6)$$

ARMA 모형은 안정적 시계열에만 적용되는 모형인데 반하여 ARIMA 모형은 불안정한 시계열을 차분(differencing) 과정을 거쳐 안정적 과정으로 전환한 후 AR과 MA유형을 동시에 사용하여 표현한 모형이다. 불안정한 시계열 Y_t 를 d 회 차분하여 구해진 시계열이 안정적인 시계열이 될 경우 ARMA(p, q)로 표현할 수 있게 되는데, 이 때 원래의 시계열 Y_t 를 ARIMA(p, d, q)라 하고 이는 AR과 MA라는 확률적보행과 안정적인 시계열 확보를 위한 차분의 과정이 복합된 시계열모형을 의미한다. 이를 후방전위연산자(backward shift operator) B 를 이용하여 표현하면 다음과 같다.

$$\phi(B)(1-B)^d Y_t = c + \theta(B)\varepsilon_t \quad (7)$$

예를 들어 ARIMA(1,1,1)의 경우

$$(1 - \phi_1 B)(1 - B)Y_t = \mu + (1 - \theta_1 B)\varepsilon_t \quad (8)$$

로 나타낼 수 있으며 여기서 ϕ_1 은 AR(1) 계수를 나타내고 θ_1 은 MA(1) 계수를, d 는 차분횟수를 나타낸다.

VAR 모형

VAR 모형은 일변량 AR 모형을 다변량 AR 모형으로 확장시킨 모형으로 예측 및 내생변수의 변화에 따른 효과 분석 등과 관련하여 자주 활용되고 있다.

전통적인 회귀모형에 의한 구조방정식 모형은 변수간의 인과관계를 통하여 종속변수 Y 를 몇 개의 설명변수(X_1, X_2, \dots)에 의해서 설명하고 있다. 그러나 회귀모형에서는 설명변수의 영향이 시간 t 가 변하더라도 항상 일정하다는 가정을 하고 있어 구조적 변화가 급속히 진행되어 설명변수의 영향이 변한 경우 이를 적절히 반영하지 못한다는

약점이 있다. 또한 구조모형(structure model)은 경제이론에 의해서 모형을 구축하고 있어 변수선택 및 모형의 내·외생변수의 선정이 모형설계자의 주관에 의해서 결정된다는 단점이 있다(정 2006).

따라서 이러한 시간에 대한 경직성과 주관성을 극복할 수 있는 방법이 위에서 살펴본 Box and Jenkins (1976)의 ARIMA 모형이라고 할 수 있다. ARIMA 모형은 현재의 관측치는 과거의 어떠한 규칙성에 의해서 재현되며, 이러한 규칙성은 미래에도 유지된다고 가정하고 미래를 예측하고자 한 것이다. 이러한 방법은 모형 설정이 용이한 반면 변수들 사이의 상호작용을 무시하고 있어 일변량분석이라는 한계에 부딪치게 된다. 이들 회귀모형과 시계열분석의 한계를 보완한 모형이 바로 Sims (1980)가 고안한 VAR 모형이다.

VAR 모형의 장점으로는 첫째, 모든 변수가 내생변수이므로 단순하다는 점을 들 수 있다. 둘째, 통상적인 OLS 방법이 각 개별 회귀식에 적용될 수 있어 추정이 용이하다. 셋째, 많은 경우에 있어서 VAR 모형으로 구한 예측결과가 더 복잡한 연립방정식 모형으로 구한 예측결과보다 낫다. 넷째, 일반적으로 모형 내에 포함되는 변수가 많지 않아 실제 예측을 수행하는데 비용과 시간이 절약되고 대규모 모형에서 제기되는 자료의 수집과 변형의 번거로운 문제가 없다.

반면 VAR 모형의 단점으로는 첫째, 연립방정식 모형에 비해 VAR 모형은 사전적인 정보가 적으므로 비이론적이다. 둘째, 예측을 강조하기 때문에 정책효과 분석에 적합하지 않다. 셋째, VAR 모형의 가장 큰 실제 운용상의 난관은 적절한 시차길이의 선택이다. 넷째, 추정결과가 모형 내에 포함된 변수의 선택 및 시차의 길이에 따라 민감하게 변화하는 등의 문제점을 지니고 있다(이 2010).

VAR 모형을 통하여 도출될 수 있는 구체적인 결과는 첫째, 한 변수의 변화가 다른 내생변수들에 미치는 동태적 효과를 충격반응분석(Impulse Response Analysis)을 통해 계산할 수 있다는 것이다. 둘째, 분산분해(Variance Decomposition)를 통하여 각각의 내생변수의 변동 가운데 이들 변수들이 전체변동에 기여한 부분의 상대적 크기 또한 분석할 수 있다. 시차가 p 인 일반적인 VAR 모형의 함수형태 $VAR(p)$ 를 살펴보면 다음 식과 같다.

$$\begin{aligned} Y_t &= C + \theta_1 Y_{t-1} + \dots + \theta_p Y_{t-p} + BX_t + \varepsilon_t \\ &= C + \sum_{i=1}^p \theta_i Y_{t-i} + BX_t + \varepsilon_t \end{aligned} \quad (9)$$

where $E(\varepsilon_t) = 0$

$$\sum (\varepsilon_t \varepsilon_s') = \begin{cases} \sigma^2 & \text{if } t = s \\ 0 & \text{else } t \neq s \end{cases}$$

위 식을 행렬로 나타내면 다음과 같이 표현된다.

$$\begin{pmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{Kt} \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_K \end{pmatrix} + \sum_{i=1}^p \begin{pmatrix} \theta_{11} & \theta_{12} & \dots & \theta_{1K} \\ \theta_{21} & \theta_{22} & \dots & \theta_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \theta_{K1} & \theta_{K2} & \dots & \theta_{KK} \end{pmatrix} \begin{pmatrix} y_{1,t-i} \\ y_{2,t-i} \\ \vdots \\ y_{K,t-i} \end{pmatrix} + \begin{pmatrix} \beta_{11} & \beta_{12} & \dots & \beta_{1L} \\ \beta_{21} & \beta_{22} & \dots & \beta_{2L} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{K1} & \beta_{K2} & \dots & \beta_{KL} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \vdots \\ \varepsilon_{Kt} \end{pmatrix} \quad (10)$$

위 식에서 Y_t 는 내생변수, X_t 는 외생변수, ε_t 는 오차항을 나타낸다. 한편 VAR 모형에서는 적정차수 p 를 결정하기 위한 모형 선정 기준으로는 Akaike Information Criterion (AIC)와 Schwarz Bayesian Criterion(SIC)이 있다.

$$AIC(p) = \ln \left| \sum \right| + \frac{2(K^2 p + 1)}{T - p} \quad (11)$$

$$SIC(p) = \ln \left| \sum \right| + \frac{(K^2 p + 1) \ln(T - p)}{T - p} \quad (12)$$

두 가지 기준 모두 차수 p 가 증가함에 따라 추가적으로 추정해야 할 계수의 수가 많아지는 것에 대한 불이익(penalty)을 추가한 것이라는 점에서 공통점을 가진다. 두 기준의 차이점으로는 AIC가 시차를 증가시킬 때마다 변수 개수의 자승에 2를 곱하는 방식으로 벌점을 부과하는 반면, SIC는 변수 개수의 자승에 관찰치 수의 자연로그 값을 곱하여 부과하는 방식을 취한다. 따라서 통상적으로 SIC는 AIC보다 더 작은 최적차수를 선정하도록 하여 모형의 간결성(Parsimony)를 달성하도록 한다.

국의 선행연구로써 Bekiros and Diks (2008)은 VECM (Vector Error Correction Model; 벡터오차수정모형) 모형을 사용하여 WTI 유가의 선물가격과 현물가격의 상호가격으로부터 WTI 유가를 예측하려는 시도를 하였다. Kaufmann and Ullman (2009)도 VECM 모형을 이용하여 WTI 유가, Brent 유가, Maya 유가, Bonny Light 유가를 예측하였다. 국내의 경우 최와 최 (2007)는 VAR 모형과 ARIMA, GARCH 모형을 이용하여 오이, 호박, 토마토 등의 농산물의 가격예측을 수행하였으며, 추정된 예측치와 실측치를 비교하여 각 품목에 적합한 가격예측모형을 찾고자 하였다.

3. 금속가격 변화 분석

사용변수 및 자료

변수를 설정하기 전에 우선 그동안의 금속가격의 변화

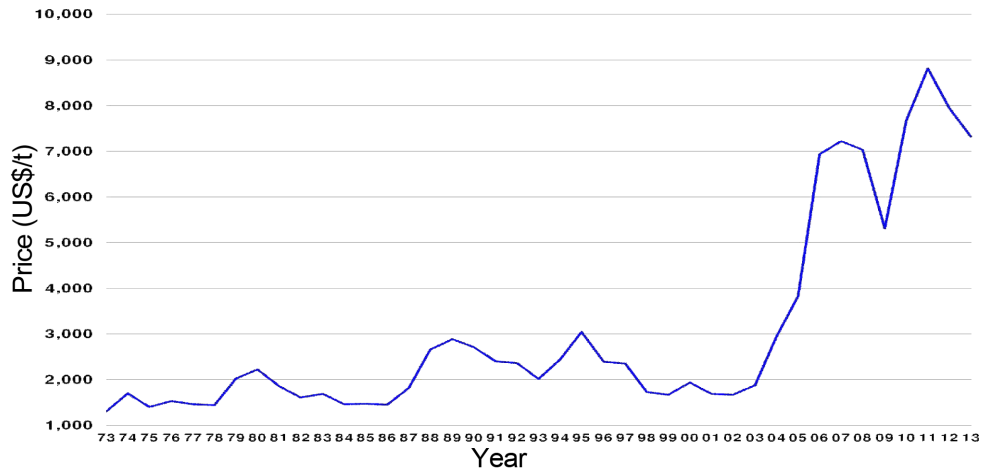


Fig. 1. Trend of copper price

를 살펴보기로 하자. Fig. 1은 동 가격의 변화추이를 나타내고 있다. 동 가격은 2003년까지는 변화폭이 상대적으로 작았지만, 2003년 이후 큰 폭으로 가격이 상승하였다. 이는 미국의 달러가치 하락과 저금리 정책, 중국의 인프라 건설 활성화 및 IP 산업성장에 따른 수요 급증 그리고 2003년 후반부터 광산사고 등으로 인하여 수급이 공급부족으로 전환되면서 거래소 재고가 급격히 감소하며 가격이 급등한 것으로 나타났다. 다만 최근에는 동의 수급이 안정화 되고 공급과잉 상태로 전환되며 가격이 다시 하락하는 추세로 전환되었다.

니켈 가격은 1987년에 가격이 큰 폭으로 상승하였다가 이내 다시 변화폭이 작아졌지만, 2005년을 기점으로 가격이 큰 폭으로 상승하였다. 2006년 가격급등의 원인은 니켈의 주 수요처인 스테인리스가 중국 등의 수요증가로 인하여 시장이 확대된 반면, 뉴칼레도니아 및 보이지베이의 파업, PT Inco사의 화재, PT Antam 신규 제련소의 누수

사고, 웨스턴오스트레일리아의 사이클론 등으로 공급불안감을 키웠기 때문으로 분석되었다. 이후 니켈가격은 공급이 늘어나며 다시 원래의 가격수준으로 하락하는 추세를 보이고 있다.

코발트의 가격은 4대 금속 중에서 가장 변동이 심하게 나타난다. 다른 금속들은 가격이 상승과 하락을 반복하면서 꾸준히 상승추세를 보인 반면, 코발트의 경우에는 상승 추세를 보이기는 하지만 변동폭이 매우 크게 나타난다. 코발트의 경우에는 주요 생산국인 콩고가 잦은 내전이 발생하여 공급이 안정적으로 수급하지 못하였기 때문으로 분석되며, 사상 최고치를 기록하였던 2008년의 가격급등은 일본, 중국의 2차전지 수요 증가로 공급부족이 발생하였으며, 팰컨브릿지사·서드베리 사업소의 파업에 의해 공급이 불안해지면서 가격이 급등한 것으로 분석되었다.

망간의 가격은 2008년 가격급등이 발생하기 전까지 4대 금속 중에서 가장 안정적인 상승추세를 보이고 있었다.

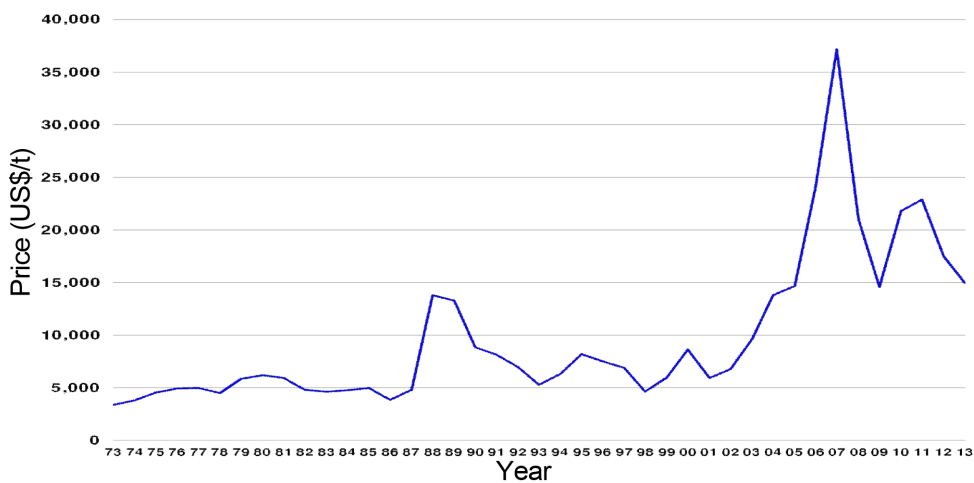


Fig. 2. Trend of nickel price

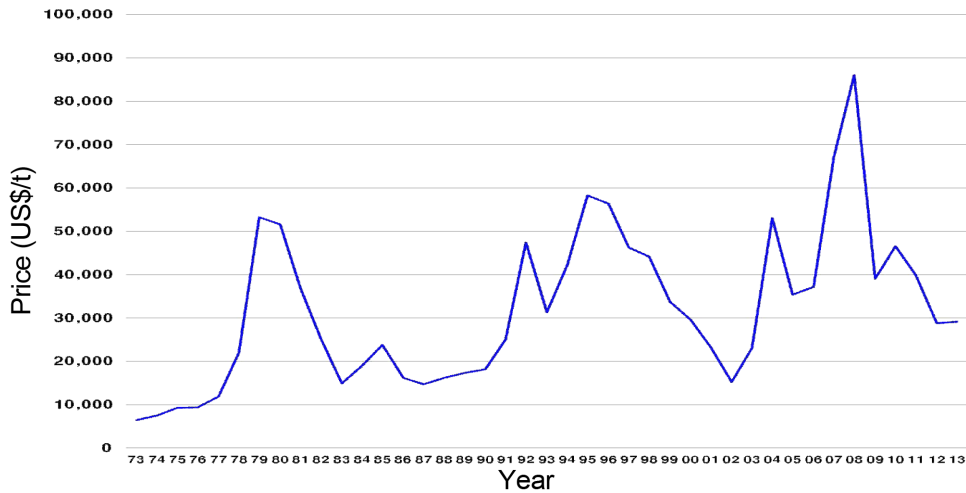


Fig. 3. Trend of cobalt price

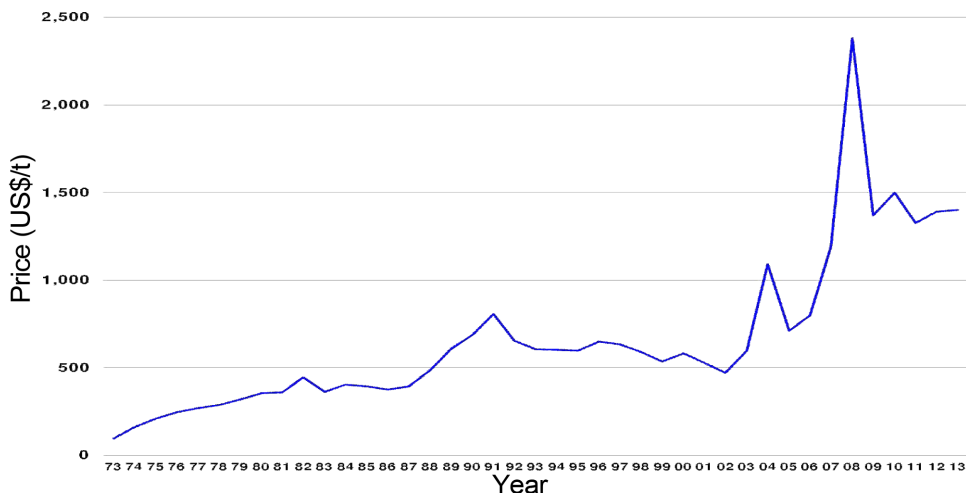


Fig. 4. Trend of manganese price

2008년 가격급등의 원인은 중국의 철강생산량 증가로 인하여 망간의 수요가 증가하는 속도를 공급이 따라잡지 못하였기 때문이고 다른 금속들과 마찬가지로 재고가 부족해지면서 공급에 대한 불안감이 커져 가격이 급등하였다. 그러나 이내 가격이 큰 폭으로 하락하였지만 가격급등 이전의 수준까지는 하락하지는 않았다.

일반적으로 가격은 수요와 공급에 의해서 결정된다. 수요측면에서 보면 최근 중국과 같은 신흥공업국들의 수요가 급속도로 증가하고 있는 반면 미국과 같은 기존 선진국들의 수요 증가율은 감소하고 있는 상황이다. 공급측면을 살펴보면 특정국가가 특정금속의 생산량을 차지하는 비율이 높은 경우가 많다. 동의 경우 칠레가 세계생산량의 32% 정도를 차지하며 코발트는 콩고가 세계생산량의 55%를 차지하고 있다. 칠레나 콩고에서 해당금속의 생산이 안정적이지 못하다면 가격은 급속히 상승할 것이다. 따라서 세계금속생산량을 가격에 미치는 요인으로 변수에

포함시킨다. 다만 본 연구에서는 세계수요량을 변수에서 제외하였다. 그 이유는 두 변수 간의 상관관계가 높아 모형의 유의성을 감소시켰으며, 공급측면인 생산량과 수요측면인 소비량은 큰 차이가 나지 않을 것이라 가정하였기 때문이다.

최근 달러의 가치가 하락하면서 안전자산의 선호로 인하여 금이나 원유 같은 실물자산은 가격이 상승하였다. 다시 말하면 달러가치와 실물자산 간에 밀접한 관계가 있다는 것이다. 금속시장에서도 마찬가지로 달러의 가치가 하락하면서 대부분의 금속은 가격이 상승하였음을 확인할 수 있었다. 이러한 현상은 2005년 이후부터 두드러진 현상으로 미확인된 정보의 확산과 함께 미국경제의 악화로 달러의 가치가 급속도로 하락하고 있고 그로 인하여 실물자산의 선호가 증가하였기 때문이다. 또한 금속과 관련된 파생상품 개발 등이 이루어지면서 투자자금이 금속시장에 유입되고 있는 이유도 있다. 따라서 본 연구에서는 금융요

인을 반영하기 위한 변수로 달러가치에 대한 자료를 사용하였다. 달러가치의 하락으로 투기세력의 자금이 실물 관련 금융상품에 투자할 것이기 때문이다. 그리하여 본 연구에서는 대리변수로 달러인덱스를 고려하였다.

일반적으로 완전경쟁시장에서 제품의 가격은 한계비용과 같아야 한다는 것이 경제학적 이론이다. 그러나 금속과 같은 고갈자원은 아무리 경쟁적이어도 가격은 한계비용보다 높게 설정된다. 따라서 금속가격은 한계비용과 상관관계가 있으므로 한계비용은 시간이 갈수록 광산 개발의 심부화로 생산비용이 증가하므로 금속가격도 이에 대한 영향을 받는다. 그리하여 육상광산 생산비용을 변수로 채택하였다. 다만 코발트는 동과 니켈 생산에 따른 부산물로 얻어지는 경우가 많으므로 이와 같은 경우에는 변수를 사용하지 않았다.

다음으로 세계 GDP를 변수로 사용하였다. 일반적으로 경기가 호황이라는 것은 GDP가 증가한다는 것이고 경기가 호황일수록 건설, 제조 등에서 금속의 수요가 많아져 가격이 상승할 것이기 때문이다. 따라서 세계 GDP를 변수로 채택하였다.

마지막으로 더미변수를 고려하기로 한다. 경제위기로 인한 가격급등, 콩고와 같은 치안이 불안한 국가에서 발생한 내전으로 인한 가격의 급등, 자연재해로 인한 생산량의 급감 등은 모두 금속가격에 영향을 미치는 요인들이다. 그러나 최근에는 주요 국가들은 각각의 금속들에 대한 비축분을 보유하고 있기 때문에 초대형 규모의 재난이나 대규모 전쟁 등이 발생하지 않는 한 단기적인 수급조절은 이루어질 수 있다. 따라서 본 연구에서는 금속가격에 심각한 영향을 미치는 요인들을 더미변수(요인을 미치는 경우 1, 그렇지 않은 경우는 0)로 설정하였다.

즉, 금속가격의 설명변수로는 세계생산량, 달러화 가치, 육상광산 생산비용, 세계 GDP, 더미변수를 고려하였다

Table 1. Variables of model

| | Variables | Remark |
|---------------------|-----------|--------------------------|
| Price of Metal | PRICE | Average for Annual Price |
| World Production | PROD | Production per Year (t) |
| Dollar Index | DOL | |
| Cost of Land Mining | COST | |
| World GDP | GDP | |
| Dummy Variable | D | |

(Table 1).

동 가격의 기초통계량을 살펴보면, 평균가격은 톤당 약 3,251달러이며 최댓값, 최솟값은 약 8,821달러, 1,460달러이었다. 약 30년간 최댓값과 최솟값의 차이가 7,400달러에 달하는 것을 보면 가격변화가 매우 심하게 이루어졌음을 알 수 있다(Table 2).

니켈 가격의 기초통계량을 살펴보면, 평균가격, 최댓값, 최솟값은 톤당 약 11,062달러, 37,230달러, 3,880달러이었다. 니켈은 약 30년간 최댓값과 최솟값의 차이가 약 10배에 달하는 것을 알 수 있다. 이는 니켈의 가격변화는 동보다 더욱 크게 이루어졌음을 알 수 있다(Table 3).

코발트 가격의 기초통계량을 살펴보면, 평균가격, 최댓값, 최솟값은 톤당 약 29,241달러, 86,106달러, 6,480달러이었다. 코발트는 약 30년간 최댓값과 최솟값의 차이가 약 13배에 정도로 4개 금속 중에서 차이가 가장 크게 나타났다. 이는 코발트의 가격변화가 4개 금속 중에서 가장 크게 나타났음을 의미한다(Table 4).

망간 가격의 기초통계량을 살펴보면, 평균가격, 최댓값, 최솟값은 톤당 약 742달러, 2,380달러, 354달러이었다. 망간은 약 30년간 최댓값과 최솟값의 차이가 약 6배에 정도로 코발트와 니켈에 비해서는 작게 나타났다(Table 5).

Table 2. Basic statistics for copper price

| | PRICE | PROD | DOL | COST | GDP |
|--------------|----------|----------|----------|----------|----------|
| Mean | 3250.894 | 11479091 | 94.08582 | 1319.252 | 36059494 |
| Median | 2370 | 11000000 | 90.4271 | 1199.227 | 31990498 |
| Maximum | 8821 | 16600000 | 133.5541 | 2289.501 | 72440449 |
| Minimum | 1460 | 7200000 | 70.8345 | 963.001 | 20029995 |
| Std. Dev. | 2237.477 | 3182999 | 15.38687 | 380.0032 | 14435465 |
| Skewness | 1.367536 | 0.196099 | 0.799341 | 1.382481 | 1.095155 |
| Kurtosis | 3.320999 | 1.494362 | 3.240528 | 3.684417 | 3.390671 |
| Jarque-Bera | 10.42753 | 3.328551 | 3.593753 | 11.15598 | 6.806359 |
| Probability | 0.005441 | 0.189328 | 0.165816 | 0.00378 | 0.033267 |
| Sum | 107279.5 | 3.79E+08 | 3104.832 | 43535.3 | 1.19E+09 |
| Sum Sq. Dev. | 1.60E+08 | 3.24E+14 | 7576.182 | 4620877 | 6.67E+15 |
| Observations | 33 | 33 | 33 | 33 | 33 |

Table 3. Basic statistics for nickel price

| | PRICE | PROD | DOL | COST | GDP |
|--------------|----------|----------|----------|----------|----------|
| Mean | 11062.1 | 1194258 | 93.70078 | 4724.726 | 37080991 |
| Median | 8160 | 1140000 | 89.9118 | 4399.986 | 33241788 |
| Maximum | 37230 | 2127000 | 133.5541 | 8331.71 | 72440449 |
| Minimum | 3880 | 621000 | 70.8345 | 2901.945 | 20648355 |
| Std. Dev. | 7727.129 | 358024.1 | 15.76504 | 1406.467 | 14299698 |
| Skewness | 1.5983 | 0.684058 | 0.857501 | 0.925044 | 1.076774 |
| Kurtosis | 5.41303 | 3.074793 | 3.217206 | 3.056944 | 3.315561 |
| Jarque-Bera | 20.71958 | 2.424894 | 3.860034 | 4.425341 | 6.119071 |
| Probability | 0.000032 | 0.297469 | 0.145146 | 0.109408 | 0.046909 |
| Sum | 342925 | 37022000 | 2904.724 | 146466.5 | 1.15E+09 |
| Sum Sq. Dev. | 1.79E+09 | 3.85E+12 | 7456.093 | 59344502 | 6.13E+15 |
| Observations | 31 | 31 | 31 | 31 | 31 |

Table 4. Basic statistics for cobalt price

| | PRICE | PROD | DOL | GDP |
|--------------|----------|----------|----------|----------|
| Mean | 32090.38 | 34765.9 | 95.36711 | 32829990 |
| Median | 29241.5 | 27400 | 94.8218 | 28385438 |
| Maximum | 86106 | 82247 | 133.5541 | 72440449 |
| Minimum | 6480 | 16600 | 70.8345 | 16015152 |
| Std. Dev. | 18172.68 | 17157.62 | 14.31166 | 14889506 |
| Skewness | 0.754755 | 1.362899 | 0.583966 | 1.135283 |
| Kurtosis | 3.279715 | 3.9392 | 3.306124 | 3.596771 |
| Jarque-Bera | 3.9281 | 13.85346 | 2.429629 | 9.186007 |
| Probability | 0.140289 | 0.000981 | 0.296765 | 0.010122 |
| Sum | 1283615 | 1390636 | 3814.684 | 1.31E+09 |
| Sum Sq. Dev. | 1.29E+10 | 1.15E+10 | 7988.121 | 8.65E+15 |
| Observations | 40 | 40 | 40 | 40 |

Table 5. Basic statistics for manganese price

| | PRICE | PROD | DOL | GDP |
|--------------|----------|----------|----------|----------|
| Mean | 742.062 | 9337424 | 94.08582 | 36059494 |
| Median | 602 | 8600000 | 90.4271 | 31990498 |
| Maximum | 2380 | 16840000 | 133.5541 | 72440449 |
| Minimum | 353.7987 | 6390000 | 70.8345 | 20029995 |
| Std. Dev. | 439.0901 | 2580306 | 15.38687 | 14435465 |
| Skewness | 1.96297 | 1.488466 | 0.799341 | 1.095155 |
| Kurtosis | 7.094208 | 4.609996 | 3.240528 | 3.390671 |
| Jarque-Bera | 44.24138 | 15.74954 | 3.593753 | 6.806359 |
| Probability | 0 | 0.00038 | 0.165816 | 0.033267 |
| Sum | 24488.05 | 3.08E+08 | 3104.832 | 1.19E+09 |
| Sum Sq. Dev. | 6.17E+06 | 2.13E+14 | 7576.182 | 6.67E+15 |
| Observations | 33 | 33 | 33 | 33 |

4. 금속가격 예측모형 추정결과

ARIMA 모형

ARIMA 모형을 추정하기에 앞서 각 금속의 가격변수가 안정적인 자료인지 살펴보기 위하여 단위근 검정을 실시하였다. 단위근 검정에는 ADF(Augmented Dickey-Fuller) 검정을 사용하였다. 검정결과 각 금속의 t-Statistics가 동: -0.0958, 니켈: -1.8919, 코발트: -2.8974, 망간: -1.2118로 나타나 금속의 가격변수에서 단위근이 존재한다고 판정되었다. 따라서 각 가격변수들에 차분을 1회 실시하여 단위근 검정을 실시하였으며 그 결과 모두 단위근이 존재하지 않는 것으로 나타났다. 그러므로 본 연구에서는 안정적인 시계열 자료에서만 사용할 수 있는 ARMA 모형 대신 불안정한 시계열 자료에서 사용하는 ARIMA 모형을 사용하여 추정하였다.

모든 금속에서 ARIMA(1,1,1)의 경우 가장 안정적으로 나타났으며, 추정결과도 유의하게 나타났다. 또한 각 금속별로 단위근 검정을 실시한 결과 모든 금속의 가격변수들

Table 6. Estimated result of the ARIMA

| | Copper | Nickel | Cobalt | Manganese |
|----------|----------------------|------------------------|----------------------|----------------------|
| D(PRICE) | 0.513*** (21.623) | 0.526*** (7.451) | 0.513*** (23.668) | 0.529*** (36.792) |
| AR(1) | 0.988*** (14.388) | 0.739*** (5.568) | 0.331*** (5.283) | 780.99*** (8.277) |
| MA(1) | 0.947*** (23.93) | 0.972*** (23.31) | 0.661*** (18.87) | 0.817*** (17.803) |
| C | 15911.49 (0.206) | 12414.23*** (3.201) | 0.951*** (5.090) | 0.957*** (3.268) |

Note) t-Statistics are in (). Coefficients with *** are significant statistically at 1% level (two sided test) and with ** are significant at 5% level, those with * are significant at 10% level.

은 단위근이 존재한 것으로 나타났다. 일반적으로 불안정한 시계열을 안정적으로 만들기 위해 변수에 차분을 적용하거나 로그차분을 적용한다. 그리하여 가격변수를 1차 차분을 적용하여 추정하였다. 추정결과는 각 추정계수에 대한 통계치는 전반적으로 유의하게 나타났다.

VAR 모형

ARIMA모형은 변수 추정 시 안정성에 관한 단위근 검정을 실시하였는데, 검정결과 모든 금속의 가격변수들은 단위근이 존재한 것으로 나타났다. 그리하여 ARIMA 모형에서는 1차 차분을 적용하였지만 VAR 모형에서는 차분을 적용하지 않기로 하였다. 그 이유는 차분을 적용할 경우 데이터가 가지고 있는 본연의 정보가 사라지게 되어 예측력이 감소할 수 있다는 연구가 있기 때문이다(Gujarati 1995).

다음 단계로 변수들 사이에 장기균형관계가 있는지 파악하기 위해 공적분(cointegration) 검정을 실시하였다. 공적분 여부는 Johansen Cointegration Test로 검정하였으며 변수들 사이에 공적분 관계가 나타나는 경우에는 벡터오차수정모형(Vector Error Correction Model)을 사용하여 모형을 추정하였다. 공적분 검정을 시행한 결과 p -값이 동: 0.0046, 니켈: 0.0011, 코발트: 0.5392, 망간: 0.1479로 나타나 니켈의 경우에만 공적분 관계가 나타났다.

다음으로 VAR모형은 변수 간의 순서와 시차의 길이에 의해 결과가 달라질 수 있으므로 Granger 인과검정을 시행하여 변수 간의 순서를 정하였다. 각 Granger 인과검정 결과를 토대로 동시 경우에는 생산량, 세계 GDP, 달러화 가치, 가격, 생산비용의 순서로 VAR 모형을 설정하기로 하며 니켈은 세계 GDP, 생산량, 달러화 가치, 가격, 생산비용의 순서로 설정한다. 코발트의 경우에는 세계 GDP, 생산량, 달러화의 가치, 가격의 순서로 VAR 모형을 설정하였고 망간의 경우에는 GDP, 생산량, 달러화 가치, 가격의 순서로 VAR 모형을 설정하였다.

한편, VAR 모형의 시차는 LR, AIC, SIC 등과 같은 정보기준 통계량을 사용하여 적정시차를 선택하는데, 시차 길이 결정을 위한 추정결과는 동 3년, 니켈, 코발트, 망간은 2년으로 나타났다. 그러나 동시 경우 시차를 3년으로 할 경우 시차로 인하여 과다추정이 되어 자료의 손실이라는 문제가 발생하고 본 연구에서 사용한 자료는 연간자료로 시계열의 갯수가 많지 않으므로 2년 시차를 선택하였다.

위의 과정들을 통하여 변수의 순서와 시차의 길이가 결정되었으므로 각 금속별 시계열 자료를 사용하여 VAR(2) 모형을 설정하였으며 추정결과는 Table 7과 같다. 본 연구에서는 더미변수를 제외한 모든 변수를 내생변수로 설정하여 추정하였으며, 선형모형과 마찬가지로 더미변수를

Table 7. Estimated result of the VAR

| | Copper | Nickel | Cobalt | Manganese |
|----------------|--------------------|---------------------|--------------------|--------------------|
| log(PRICE(-1)) | -0.313 (-1.574) | 0.402 (2.885) | -0.115 (-0.600) | -0.469 (-4.017) |
| log(PRICE(-2)) | -0.544 (-2.798) | -0.255 (-1.435) | -0.204 (-1.061) | -0.363 (-3.279) |
| log(PROD(-1)) | -0.613 (-0.459) | -1.634 (-1.966) | 0.739 (1.219) | -0.192 (-0.688) |
| log(PROD(-2)) | -1.882 (-1.382) | 0.359 (0.630) | -0.872 (-1.697) | 0.144 (0.610) |
| log(DOL(-1)) | -0.411 (-0.583) | -1.276 (-2.372) | -1.569 (-1.494) | 0.110 (0.280) |
| log(DOL(-2)) | -1.112 (-1.924) | 1.456 (2.742) | -0.044 (-0.038) | -0.266 (-0.665) |
| log(COST(-1)) | -0.139 (-0.276) | 0.930 (2.268) | - | - |
| log(COST(-2)) | 0.961 (2.328) | -0.639 (-2.250) | - | - |
| log(GDP(-1)) | -0.249 (-0.212) | -2.595 (-1.397) | 1.644 (0.598) | -1.044 (-0.940) |
| log(GDP(-2)) | 2.627 (1.729) | 4.684 (2.024) | -5.328 (-1.441) | -3.153 (-2.210) |
| C | 0.00064 (0.009) | -13.863 (-3.076) | 0.098 (0.543) | 0.224 (3.694) |
| D | 0.582 (4.031) | 0.772 (7.598) | 0.635 (3.175) | 0.724 (7.037) |

Note) t-Statistics are in ().

제외한 모든 변수에 로그를 취하여 추정하였다.

적합도 분석

모형이 변수의 예측에 이용되는 경우에 피설명변수의 예측능력으로 모형을 평가하는 것이 바람직할 것이다. 그리하여 전체 데이터를 가격예측모형 추정에 사용한 데이터 기간(시작시점부터 t 시점까지)과 가격예측이 이루어진 기간(t 시점 이후)으로 나눈 후에 가격예측 추정치와 실제 데이터와의 오차를 계산하여 수요예측의 정확성을 평가한다. 가격예측치와 실제치의 오차를 측정하기 위하여 본 연구에서는 MAPE(Mean Absolute Percentage Error)를 이용하였으며, MAPE는 다음과 같이 계산된다.

$$MAPE = \frac{\sum_{t=1}^L \frac{|P_t - \hat{P}_t|}{P_t}}{L} \times 100 \quad (13)$$

P_t 는 실적치, \hat{P}_t 는 추정된 전망 예측치, L 은 관찰기간을 나타내며, 위 식을 토대로 계산한 결과는 다음과 같다.

결과를 살펴보면 ARIMA 모형의 경우 0.4~0.6, VAR 모형의 경우 0.1~0.2 사이로 나타나 대체적으로 ARIMA 모형이 예측력이 떨어지는 것으로 나타났으며 모든 금속

Table 8. Predictive power of the price forecasting models from MAPE

| | Copper | Nickel | Cobalt | Manganese |
|--------------|--------|--------|--------|-----------|
| ARIMA(1,1,1) | 0.6194 | 0.5318 | 0.4530 | 0.4005 |
| VAR | 0.1432 | 0.2042 | 0.1848 | 0.1493 |

에서 VAR 모형의 예측력이 높은 것으로 나타났다.

5. 결 론

본 연구에서는 망간단괴 함유 4개 금속의 가격예측모형을 설정하였다. 가격예측모형은 전통적인 ARIMA 모형과 VAR 모형을 사용하여 예측모형을 구축하였다. 사용된 변수로는 각 금속가격을 종속변수로 하여 세계 생산량, 달러화 가치, 육상광산 생산비용, 세계 GDP와 더미변수를 사용하였다.

ARIMA 모형을 추정하면서 시계열의 안정성을 검정하기 위하여 단위근 검정을 실시하였으며 검정결과 I(1)인 불안정한 시계열인 것으로 나타나 차분을 사용한 ARIMA (1,1,1) 모형을 사용하여 예측하였다. VAR 모형을 추정하기에 앞서 시계열들이 단위근을 가지는 불안정한 시계열이었으므로 불안정시계열 간의 선형결합이 안정적인지 확인하기 위해 Johansen 공적분 검정을 실시하여 동, 코발트, 망간의 경우 공적분 관계가 성립하는 것을 확인할 수 있었다.

다음으로 변수 간의 순서를 정하기 위하여 Granger 인과검정을 실시하여 각 금속별로 변수의 순서를 정하였으며, 최적시차의 길이를 추정하였다. 최적시차 추정 결과 동은 3년 시차가 적정한 것으로 나타났지만 사용된 시계열 자료가 연간자료기 때문에 많은 시차를 택하게 되면 과다추정으로 인하여 자료가 손실되는 문제가 발생하여 2년 시차를 선택하였다. 니켈과 코발트의 경우에는 2년 시차가 적정시차로 추정되어 2년 시차를 그대로 선택하였고 망간의 경우에는 1년 시차가 적정시차로 추정되었지만 너무 적은 시차를 선택하면 과다추정으로 자료의 손실이라는 문제가 발생하기 때문에 망간도 2년 시차를 선택하여 모형을 추정하였다.

각 모형별 예측력을 평가하기 위해 각 금속의 기간을 대상으로 MAPE를 계산하였다. 계산결과 VAR 모형이 ARIMA 모형에 비해 예측력이 더 우수한 것으로 나타났다. 그러나 한 가지 유의해야 할 점은 VAR 모형이나 ARIMA 모형의 경우 단기예측에 적합한 모형이며 장기예측에는 예측력이 떨어진다는 것이다. 특히 VAR 모형의 경우 경제이론에 근거하지 않는 모형이기 때문에 이는 큰 약점이 될 수 있다. 그렇지만 지금까지 VAR 모형을 사용한 연구들을 보면 VAR 모형의 예측력이 우수한 것으로

나타났으며 본 연구에서도 VAR 모형의 예측력이 더 뛰어난 것으로 나타났다.

본 연구는 망간단괴 개발사업의 경제성에 가장 큰 영향을 미치는 금속가격에 대한 연구가 2000년대 초반 이후로 수행되지 않았던 점을 고려할 때, 2000년대 이후의 금속가격을 반영하고 기존 모형에서 사용하지 않았던 모형을 사용하여 금속가격을 예측하였다는 점에서 의의가 있다. 또한 본 연구에서 수행한 금속가격 예측이 망간단괴 개발사업의 경제성을 평가할 때 좀 더 객관적이고 유용하게 활용할 수 있는 자료가 될 수 있을 것으로 기대한다.

사 사

본 연구는 해양수산부 “태평양 심해저 광물자원 개발사업(PM57950)”의 지원을 받아 수행된 연구임.

참고문헌

- 문건순 (1997) 벡터자기회귀(VAR)모형의 이해. 통계분석연구 2(1):23-56
- 박완규, 홍성표 (2009) Gujarati 계량경제학(제5판). 지필미디어
- 이철용 (2010) 국제 유가 중·장기예측모형 개발-베이지안추론 이용 보고서
- 정선경 (2006) VAR모형을 이용한 분기GDP 예측모형 연구 보고서
- 최병욱, 최익창 (2007) 시계열 분석방법을 이용한 과채류 월별가격 예측. 농촌경제 30(1):129-148
- 한국과학기술정책연구원 (2007) 심해저 광물자원(망간단괴) 개발 경제성 분석 평가 보고서
- 한국해양과학기술원 (2012) 심해저 광물자원 개발사업 보고서 III. 법제도분석
- 한국해양과학기술원 (2013) 심해저 광물자원 개발사업 보고서 III. 법제도분석
- 한국해양연구소 (1990) 심해저 광물자원 개발전략연구(II) 보고서
- 한국해양연구소 (1992) 심해저 광물자원 개발 연구 보고서
- 한국해양연구원 (2002) 심해저 광물자원탐사 보고서(III. 심해저 자원개발 제도 분석)
- Bekiros SD, Diks CGH (2008) The relationship between Crude oil spot and future price: Cointegration, Linear and Nonlinear Causality. Energy Economics 30(5):2673-2685
- Box GEP, Jenkins GM (1976) Time Series Analysis: Forecasting and control. revised edition, Holden-Day Inc. San Francisco.
- Dickey DA, Fuller WA (1979) Distribution of the estimators for autoregressive time series with a unit root. J Am Stat Assoc 74(366):426-431
- Granger CWJ (1969) Investigating causal relationships by

- econometric models and cross-spectral method. *Econometrica* **37**(3):424-438
- Gujarati D. (2002) *Basic Econometrics*, 4th edition, McGraw-Hill, New York
- Johansen S (1991) Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica* **59**(6):1551-1580
- Kaufmann RK, Ulman B (2009) Oil prices, speculation, and fundamentals, interpreting causal relations among spot and future prices. *Energy Economics* **31**:550-558
- Phillips PCB (1987) Testing for a Unit Root in Time Series Regression. *Biometrika* **75**(2):335-346
- Phillips PCB (1987) Time series regression with a unit root. *Econometrica* **55**:277-301
- Said SE, Dickey DA (1985) Hypothesis Testing in ARIMA (p, 1, q) Models. *J Am Stat Asso* **80**:369-374
- Sims CA (1980) Macroeconomics and Reality, *Econometrica* **48**:1-48
- USGS <http://www.usgs.gov/> Accessed 10 Oct 2014
- 국문 참고자료의 영어 표기**
English translation / Romanization of references originally written in Korean
- Choi B-O, Choi I-C (2007) Monthly price forecasting of fruit-type vegetables using time-series analysis. *J Rural Development* **30**(1):129-148 (in Korean)
- Jung S-K (2006) A Study on Quarterly GDP Forecasting Model Using the VAR Model. Statistical Research Institute (in Korean)
- KIOST (2012) A Study on the Strategy for the Development of Deep Seabed Mineral Resources (III. Legal Regime Analysis) (in Korean)
- KIOST (2013) A Study on the Strategy for the Development of Deep Seabed Mineral Resources (III. Legal Regime Analysis) (in Korean)
- KORDI (1990) A Study on the Strategy for the Development of Deep Seabed Mineral Resources (II) (in Korean)
- KORDI (1992) A Study on the Strategy for the Development of Deep Seabed Mineral Resources (in Korean)
- KORDI (2002) A Study on the Strategy for the Development of Deep Seabed Mineral Resources (III. The Development of Deep Seabed Resources Regime Analysis) (in Korean)
- Lee C-Y (2010) Development of Long-term Oil Price Forecasting Model-Using Bayesian Inference. Korea Energy Economics Institute (in Korean)
- Moon K-S (1997) A Understanding of Vector Autoregressive Model. *Journal of The Korean Official Statistics* **2**(1):23-56 (in Korean)
- Park W-G, Hong S-P (2009) *Gujarati Basic Econometrics 5/E*. JIPHIL (in Korean)
- STEPI (2007) A Economic Analysis for the Development of Deep Seabed Mineral Resources (Manganese Nodules) (in Korean)

Received Oct. 26, 2014

Revised Nov. 17, 2014

Accepted Dec. 9, 2014