

Development of System Marginal Price Forecasting Method Using ARIMA Model

金大容* · 李瓚柱** · 鄭允源† · 朴宗培*** · 愼重麟§
 (Dae-Yong Kim · Chan-Joo Lee · Yun-Won Jeong · Jong-Bae Park · Joong-Rin Shin)

Abstract - Since the SMP(System Marginal Price) is a vital factor to the market participants who intend to maximize their profit and to the ISO(Independent System Operator) who wish to operate the electricity market in a stable sense, the short-term marginal price forecasting should be performed correctly. In an electricity market the short-term market price affects considerably the short-term trading between the market entities. Therefore, the exact forecasting of SMP can influence on the profit of market participants. This paper presents a new methodology for a day-ahead SMP forecasting using ARIMA(Autoregressive Integrated Moving Average) model based on the time-series method. And also the correction algorithm is proposed to minimize the forecasting error in order to improve the efficiency and accuracy of the SMP forecasting. To show the efficiency and effectiveness of the proposed method, the case studies are performed using historical data of SMP in 2004 published by KPX(Korea Power Exchange).

Key Words : SMP(System Marginal Price) Forecasting, Day-Ahead, ARIMA(Autoregressive Integrated Moving Average), Time-Series, Electricity Market

1. 서 론

오늘날 전력산업은 급속한 속도로 변하여 전력에너지도 일반 상품과 동일하게 시장에서 거래될 수 있는 경쟁적 전력시장이 형성되어 가고 있다. 따라서 경쟁적 전력시장에서 발전회사들은 자신의 수익을 최대화할 수 있는 입찰전략을 필요로 하며, 이를 해결하기 위한 많은 노력들이 시도되고 있다[1,2]. 전력산업에 시장원리가 도입됨으로써 계통계획은 가격신호를 바탕으로 수립되며, 발전확충계획, 송전망확충계획, 배전계획 등은 장기 전력가격 예측에 의해 결정된다. 반면, 시장참여자들 사이의 전력거래는 대부분 단기 전력가격 예측에 의해 결정되므로[3], 단기 전력가격 예측은 시장참여자들의 이익 극대화를 위한 중요한 요소가 된다. 따라서 정확한 계통한계가격 예측은 시장참여자들에게 있어서 이익을 극대화 할 수 있는 전략수립의 바탕이 된다.

한국전력거래소(Korea Power Exchange)는 전력시장 운영규칙에 따라 입찰, 정산, 계량, 시장감시, 정보공개, 분쟁조정 등 공정하고 투명한 시장 운영업무를 맡고 있다. 전력거래소에서 결정하는 계통한계가격은 발전사업자가 제출한 입

찰 자료를 바탕으로 계통 상황을 고려하여 결정되어지며 기본적으로 계통한계가격은 각 시간대별로 필요한 전력수요와 공급의 균형이 이루어지는 상태에서 마지막으로 투입된 한계 발전기의 가격으로 결정된다. 계통한계가격은 전력의 수요, 발전기의 고장정지확률, 발전기의 설비용량, 발전기의 보수상태, 발전기의 입·출력 특성, 발전기들의 투입 우선순위 등 여러 가지 요소들에 의해 영향을 받는다[4]. 이러한 계통한계가격의 정확한 예측은 투명하고 효율적인 시장형성에 매우 중요한 요소이며 이에 대한 예측 알고리즘들이 연구되어 왔다[3].

기존에 사용된 예측 방법론은 시계열 자료 예측기법인 자기회귀누적이동평균(Autoregressive Integrated Moving Average ; ARIMA)모형[1], 신경회로망(Artificial Neural Network)[3], 혼합 시스템(Hybrid System)[5] 등이 있다. ARIMA 모형은 단기 예측의 정확성이 매우 높지만 새로운 자료가 투입될 때 마다 모형에 적합한 모수를 선정해야 하는 단점이 있다. 신경회로망은 복잡하고 다양한 문제에 대한 적용이 가능하나 느린 학습속도와 학습과정에서 비선형성과 오차 역전파 학습(back-propagation learning) 알고리즘을 사용하기 때문에 오차 표면상에서 바람직하지 못한 국부해에 수렴하는 단점이 있다. 혼합 시스템은 예측 방법론에 다양한 수학적인 도구를 적용하거나 예측 방법론들을 결합하여 서로의 단점을 보완한다. 이와 같은 예측 방법론 이외에도 다양한 예측기법들이 시도되고 있다. 하지만 시장가격에는 많은 가변요소가 내포되어 있기 때문에 정확한 예측이 어렵다. 또한 대부분의 예측 알고리즘이 과거 자료의 일정 패턴에 의존하기 때문에 시변의 가격변동에 대한 불규칙적 동특성을 반영하기 어렵고, 특정일 또는 특정 시간대에

† 교신저자, 正會員 : 建國大學 電氣工學科 博士課程

E-mail : ywjeong@konkuk.ac.kr

* 學生會員 : 建國大學 電氣工學科 碩士課程

** 正會員 : 建國大學 電氣工學科 博士課程

*** 正會員 : 建國大學 電氣工學科 助教授 · 工博

§ 正會員 : 建國大學 電氣工學科 教授 · 工博

接受日字 : 2005年 11月 12日

最終完了 : 2005年 12月 22日

심각한 예측오차가 발생하는 것을 볼 수 있다. 따라서 원활한 시장형성 및 시장참여자의 공정한 경쟁을 위한 보다 정확한 시장가격 예측 알고리즘이 필요하다.

본 논문에서는 계통한계가격 예측을 위해서 시계열 자료 예측기법 가운데 하나인 ARIMA 모형 기반의 계통한계가격 예측 알고리즘을 제시한다. 본 논문에서 제시하는 예측 알고리즘은 최소 시계열 자료를 가지고 예측에 필요한 주요 입력요소를 이용하여 시간대별(hourly), 주간대별(weekly), 월별(monthly) 각각의 계통한계가격을 예측한다. 시간대별 예측 방법론은 t 시간의 계통한계가격을 예측하기 위해 과거 연속된 시계열 자료 ($t-1, t-2, \dots$)를 입력 자료로 사용하여 예측하고, 주간대별 예측 방법론은 주말 특성을 예측 알고리즘에 반영하기 위한 방법으로 이전 주의 연속된 시계열 자료 ($w-1, w-2, \dots$)를 입력 자료로 사용한다. 그리고 월별 예측 방법론은 계절별 특성을 반영하기 위한 알고리즘으로서 이전 달의 연속된 시계열 자료 ($m-1, m-2, \dots$)를 입력 자료로 사용하여 예측한다. 또한 본 논문에서 제안한 보정 알고리즘을 통해 각 입력요소별로 예측된 시장가격의 예측오차를 최소화하여 계통한계가격 예측의 정확성을 향상시킨다.

본 논문에서 제안한 계통한계가격 예측 알고리즘의 타당성을 검증하기 위하여 2004년 8월(첨두기간) 한 달간의 계통한계가격을 예측하였다. 또한 예측값의 정확성을 평가하기 위하여 평균절대비율오차(Mean Absolute Percentage Error ; MAPE)를 사용하여 실제값과의 예측오차를 분석하였다.

2. 일반적인 ARIMA 모형

경제규모의 팽창과 변화 등으로 미래에 대한 불확실성이 증가함에 따라 특정 변수의 예측을 수행하고자 할 경우 보다 정확하고 효율적인 방법을 찾게 된다. 이러한 현실적 요구에 부응하는 방법론으로 등장한 시계열분석(time series analysis)은 특정 변수의 미래 값을 예측하기 위해 단지 해당변수의 과거 변동내용과 유형을 활용하는 방법이다[6]. 시계열분석을 위해서 시계열은 각 항이 서로 비교할 수 있는 성질이 있어야 되기 때문에 다음과 같은 요인이 갖추어져야 한다. 첫째, 시계열에서의 시간의 경과는 동일한 시간간격에 따라야 한다. 둘째, 시계열을 구성하는 숫자들은 동일한 정의와 방법에 의하여 조사된 것이어야 한다. 이러한 시계열 자료를 분석하는 목적은 과거와 현재의 자료에 대한 성격을 보다 자세히 이해하고, 그 이해를 바탕으로 미래를 보다 합리적으로 예측하기 위함이다[7].

전력계통분야에서 시계열분석을 이용한 장기간·단기간 시장가격예측, 수요예측 등 다수의 연구사례가 이미 보고된 바 있다[8-11]. 전력계통에서 측정되는 계통한계가격 자료는 평균, 분산 및 자기공분산이 시간의 변화에 영향을 받는 비정상시계열 자료이다. Box-Jenkins는 이 같은 비정상시계열 자료를 연속적인 차분(differencing)과 확률과정(stochastic process)을 통해 정상시계열로 변환을 일반화하였다. 시계열 분석 방법론은 Box-Jenkins에 의해 보다 체계화되었고 자기회귀이동평균(Autoregressive Moving Average ; ARMA)모형을 중심으로 하는 모형 식별(identification)과 추정(estimation) 그리고 추정된 모형의 검진(diagnostic checking)

과 예측을 수행하는 방법론이 제시되어 일반적으로 사용되고 있다[6,12]. 이러한 시계열분석 절차에 따라 과거 계통한계가격 자료를 이용하여 전일 계통한계가격 예측이 가능하다. 다음 절에서는 Box-Jenkins분석방법의 기초가 되는 개념을 간략히 정리하여 보고, 본 논문에서의 예측모형인 ARIMA모형에 관해 살펴보기로 한다.

2.1 자기회귀(Autoregressive ; AR)모형

시계열의 현재값(y_t)은 현재 관측값을 설명해 주는 시계열의 과거값들과 설명해 주지 못하는 부분인 오차항 백색잡음(white noise)의 선형결합으로 구성할 수 있다. 즉, 과거의 실제값들을 기반으로 미래를 예측하는 방법이다. 일반적으로 차수(lagged variables)가 p 인 자기회귀모형은 다음과 같다.

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \epsilon_t \quad (1)$$

여기서,

ϕ_i : 자기회귀계수

y_{t-i} : 자기회귀모형의 차수

ϵ_t : 평균이 0이고 분산이 일정한 백색잡음

p -차 자기회귀모형에서 자기상관함수(Autocorrelation Function ; ACF)는 근이 실수인 경우 지수곡선 또는 근이 복소수인 경우에는 사인곡선의 형태로 시차(k)가 증가함에 따라 0으로 급속히 감소한다. 부분자기상관함수(Partial Autocorrelation Function ; PACF)는 시차(k)가 차수 p 보다 크면 0으로 소멸한다.

근본적으로 관측된 시계열자료가 자기회귀모형의 어느 차수에 적합한지 모르지만 정칙 (1)과 같이 차수 p 보다 큰 시차(k)를 가진 시계열 사이의 부분자기상관은 0이 됨을 알 수 있다. 이것을 자기상관함수의 형태와 병행하여 사용해서, 자기회귀모형의 차수를 결정한다.

2.2 이동평균(Moving Average ; MA)모형

E.Slutzky가 제안한 이동평균모형 역시 선형결합 형태이고, 이 모형은 과거 시계열의 오차값들을 기반으로 미래를 예측하는 방법이다. 일반적으로 차수가 q 인 이동평균모형은 다음과 같다.

$$y_t = \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \dots - \theta_q \epsilon_{t-q} \quad (2)$$

여기서,

θ_i : 이동평균계수

ϵ_{t-i} : 이동평균모형의 차수

정칙 (2)에서 ϵ_t 는 t 시점의 확률오차이며 $\epsilon_{t-1}, \epsilon_{t-2}, \dots, \epsilon_{t-q}$ 들은 전시점들의 확률오차를 나타낸다[7]. q -차 이동평균모형에서 자기상관함수는 시차(k)가 차수 q 보다 크면 0으로 소멸하고 부분자기상관함수는 근이 실수인 경우 지수

곡선 또는 근이 복소수인 경우에는 사인곡선의 형태로 시차(k)가 증가함에 따라 0으로 감소한다. 따라서 자기상관함수의 정보를 이용하여 이동평균모형의 차수결정에 큰 역할을 한다.

2.3 자기회귀이동평균모형

정식 (1)과 (2)의 두 가지 모형은 시계열 y_t 가 과거의 시계열자료 값들로 표현될 수 있다는 것과 과거 시계열자료의 오차값들로 표현될 수 있다는 것을 나타내고 있다. 하지만 경우에 따라서 시계열 y_t 는 과거의 시계열값들과 과거의 시계열 오차값들을 동시에 고려하여 영향을 받는다고 할 수도 있다. 그러므로 이렇게 설정된 자기회귀모형과 이동평균모형의 혼합형태를 H.O.Wold의 자기회귀이동평균모형 또는 ARMA(p, q)모형이라 한다[13]. 차수가 p 와 q 인 자기회귀이동평균모형은 다음과 같다.

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \dots - \theta_q \epsilon_{t-q} \quad (3)$$

정식 (3)의 차수(p, q)가 높아지면 그들 모형에 포함된 계수(ϕ_i, θ_i)가 많아져서 모형이 복잡하게 된다. 따라서 이를 일반적으로 적용할 수 있는 단순화된 형태의 조건으로 표현하는 것이 필요하다. 이를 위하여 후향전위연산자(backward shift operator)를 이용하기로 한다. 후향전위연산자는 B^m 으로 정의하며 연산의 지수 m 은 차수를 나타내고 다음과 같이 표현한다[14].

$$(B^m)y_t = y_{t-m} \quad (4)$$

후향전위연산자의 정의에 의해 정식 (3)의 ARMA(p, q)모형은 다음과 같다.

$$\begin{aligned} (1 - \phi_1 B^1 - \phi_2 B^2 - \dots - \phi_p B^p) y_t \\ = (1 - \theta_1 B^1 - \theta_2 B^2 - \dots - \theta_q B^q) \epsilon_t \\ \phi_p(B) y_t = \theta_q(B) \epsilon_t \\ y_t = \frac{\theta_q(B)}{\phi_p(B)} \epsilon_t \end{aligned} \quad (5)$$

여기서,

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

ARMA(p, q)모형에서 자기상관함수와 부분자기상관함수는 시차(k)가 차수 p, q 보다 크면 지수곡선 또는 사인곡선의 형태로 감소한다.

3. 계통한계가격 예측 ARIMA 모형

3.1 ARIMA 모형

지금까지 살펴본 시계열모형들은 추세요인이 없는 정상시계열 자료에만 적용될 수 있는 모형이다. 그러나 현실적으로 존재하는 전력계통의 계통한계가격은 시간의 변화에 영향을 받는 비정상시계열이다. 이를 한 두번의 차분과정을 거치면 대부분 추세요인이 제거된 정상시계열로 전환될 수 있으며 일단 전환된 시계열은 AR, MA 또는 ARMA유형으로 표현이 가능하다.

일반적으로 원계열 y_t 의 연속적인 차분을 차분연산자(difference operator)인 $\nabla = (1-B)$ 를 사용하면 각 차수의 차분은 다음과 같다.

$$\begin{aligned} \text{1차 차분: } \nabla y_t &= y_t - y_{t-1} = (1-B)y_t \\ \text{2차 차분: } \nabla^2 y_t &= (y_t - y_{t-1}) - (y_{t-1} - y_{t-2}) \\ &= y_t - 2y_{t-1} + y_{t-2} = (1-B)^2 y_t \\ &\vdots \\ \text{d차 차분: } \nabla^d y_t &= (1-B)^d y_t \end{aligned} \quad (6)$$

여기서,

d : 차분차수

비정상적인 시계열 y_t 를 적절히 차분시켜 얻은 새로운 계열을 w_t 라 표시하면 다음과 같다.

$$w_t = \nabla^d y_t = (1-B)^d y_t \quad (7)$$

정식 (7)은 추세요인이 제거된 정상시계열이며 이것은 정상확률과정인 ARMA(p, q)를 따른다면 다음과 같다.

$$\begin{aligned} w_t = \phi_1 w_{t-1} + \phi_2 w_{t-2} + \dots + \phi_p w_{t-p} + \epsilon_t \\ - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \dots - \theta_q \epsilon_{t-q} \end{aligned} \quad (8)$$

이를 모형식으로 표현하면 아래의 식과 같다.

$$\begin{aligned} (1-B)^d (1 - \phi_1 B^1 - \phi_2 B^2 - \dots - \phi_p B^p) w_t \\ = (1 - \theta_1 B^1 - \theta_2 B^2 - \dots - \theta_q B^q) \epsilon_t \\ (1-B)^d \phi_p(B) w_t = \theta_q(B) \epsilon_t \\ y_t = \left(\frac{1}{1-B} \right)^d \frac{\theta_q(B)}{\phi_p(B)} \epsilon_t \end{aligned} \quad (9)$$

이처럼, ARMA 모형과 차분연산기법을 합한 정식 (9)와

같은 모형을 ARIMA(p, d, q) 모형이라 한다.

Box-Jenkins는 이론적으로 도출된 자기회귀모형, 이동평균모형, 및 ARIMA 모형의 자기상관함수와 부분자기상관함수가 차수 p 또는 q 의 값에 따라 독특한 형태를 가짐을 이용하여 시계열자료의 적합한 모형과 차수를 선정한다. 또한 ACF와 PACF는 정상성 여부를 점검하는데 유용하게 사용하고 과소차분(under-differencing)과 과대차분(over-differencing)을 점검하는 데에도 사용된다[7,12-15]. 이러한 ARIMA 모형의 일반적인 ACF와 PACF의 특징을 정리하면 표 1과 같다.

표 1 ARIMA 모형의 자기상관함수와 부분자기상관함수
Table 1 Autocorrelation and Partial autocorrelation function of ARIMA Model

모형	자기상관함수	부분자기상관함수
AR(1)	지수감소함수	$\phi_{kk}=0, k>1$
AR(2)	지수감소함수 또는 감소사인곡선	$\phi_{kk}=0, k>2$
AR(p)	지수감소함수 또는 감소사인곡선	$\phi_{kk}=0, k>p$
MA(1)	$\rho_k=0, k>1$	지수감소함수
MA(2)	$\rho_k=0, k>2$	지수감소함수 또는 감소사인곡선
MA(q)	$\rho_k=0, k>q$	지수감소함수 또는 감소사인곡선
ARMA (1,1)	시차 $k=1$ 에서부터 지수감소함수	시차 $k=1$ 에서부터 지수감소함수
ARMA (p, q)	시차 $p-q$ 이후부터 지수감소함수 또는 감소사인곡선	시차 $p-q$ 이후부터 지수감소함수 또는 감소사인곡선

단, ρ_k 는 자기상관함수, ϕ_{kk} 는 부분자기상관함수

3.2 계통한계가격 예측 ARIMA 모형 유도

일반적으로 계통한계가격은 시간대별(hourly), 주간대별(weekly), 월별(monthly)에 따라 계통한계가격함수(y_t)에 대해 각각 서로 다른 영향을 받는다. 때문에 예측 계통한계가격함수는 각각의 인자에 따라서 일반화가 될 수 있으며, 이는 예측모형의 차수설정에 중요한 요소로 적용된다. 따라서 각각의 인자에 대한 t 시간대 계통한계가격 예측모형은 정식(9)를 이용하여 다음과 같이 나타낼 수 있다.

• 시간대별 모형 $y_t^h = \left(\frac{1}{1-B}\right)^d \frac{\theta_a^h(B)}{\phi_b^h(B)} \epsilon_t$ (10)

• 주간대별 모형 $y_t^w = \left(\frac{1}{1-B}\right)^d \frac{\theta_a^w(B)}{\phi_b^w(B)} \epsilon_t$ (11)

• 월별 모형 $y_t^m = \left(\frac{1}{1-B}\right)^d \frac{\theta_a^m(B)}{\phi_b^m(B)} \epsilon_t$ (12)

여기서,

- $\phi_b^h(B)$: 차수 p 를 갖는 시간대별 자기회귀모형
- $\phi_b^w(B)$: 차수 p 를 갖는 주간대별 자기회귀모형
- $\phi_b^m(B)$: 차수 p 를 갖는 월별 자기회귀모형
- $\theta_a^h(B)$: 차수 q 를 갖는 시간대별 이동평균모형
- $\theta_a^w(B)$: 차수 q 를 갖는 주간대별 이동평균모형
- $\theta_a^m(B)$: 차수 q 를 갖는 월별 이동평균모형

4. 보정 알고리즘

본 논문에서는 시간대별, 주간대별, 월별 예측모형으로 예측된 결과를 이용하여 정확한 계통한계가격 예측을 위해 보정 알고리즘을 개발하였다. 시계열 자료 예측에 있어서 가장 큰 영향을 미치는 것은 바로 직전의 계통한계가격이다. 따라서 시간대별과 주간대별 예측모형은 월별 예측모형에 비해 일반적으로 더 정확한 예측값을 보인다. 반면, 월별 예측모형은 시간대별 및 주간대별 예측모형에 의해 발생할 수 있는 돌출적인 예측값을 보정한다.

확장된 ARIMA 모형을 이용하여 시간대별(y_t^h), 주간대별(y_t^w), 월별(y_t^m)에 대한 각각의 예측값을 다음과 같이 선정한다. 여기서 y_t^{\max} 는 t 시간대 예측값(y_t^h, y_t^w, y_t^m)들 가운데 가장 큰 예측값을 나타내며 $\max[y_t^h, y_t^w, y_t^m]$ 으로 정의하고, y_t^{\min} 는 t 시간대 예측값들 가운데 중간 크기 값을 나타내며 $\text{mid}[y_t^h, y_t^w, y_t^m]$ 으로 정의한다. 또한 y_t^{\min} 는 t 시간대 예측값들 가운데 가장 작은 크기 값을 나타내며 $\min[y_t^h, y_t^w, y_t^m]$ 으로 정의한다. 이를 식으로 표현하면 다음과 같다.

$$y_t^{\max} = \max[y_t^h, y_t^w, y_t^m] \quad (13-a)$$

$$y_t^{\min} = \min[y_t^h, y_t^w, y_t^m] \quad (13-b)$$

$$y_t^{\text{mid}} = \text{mid}[y_t^h, y_t^w, y_t^m] \quad (13-c)$$

이 경우 모형에 의한 최선의 예측은 선정된 예측값들을 이용하여 예측오차를 최소화하는 것이다. 본 논문에서는 예측 오차를 다음과 같은 식으로 최소화한다.

$$Y_t = \begin{cases} \frac{y_t^{\max} + y_t^{\text{mid}}}{2} & (y_t^{\max} - y_t^{\text{mid}} < y_t^{\text{mid}} - y_t^{\min}) \\ \frac{y_t^{\text{mid}} + y_t^{\min}}{2} & (y_t^{\max} - y_t^{\text{mid}} > y_t^{\text{mid}} - y_t^{\min}) \\ y_t^{\text{mid}} & (y_t^{\max} - y_t^{\text{mid}} = y_t^{\text{mid}} - y_t^{\min}) \end{cases} \quad (14)$$

여기서,

Y_t : t 시간대 최종 계통한계가격 보정값

제한한 모형과 보정 알고리즘을 토대로 ARIMA 모형을 이용한 SMP 예측 알고리즘 순서도는 그림 1과 같다.

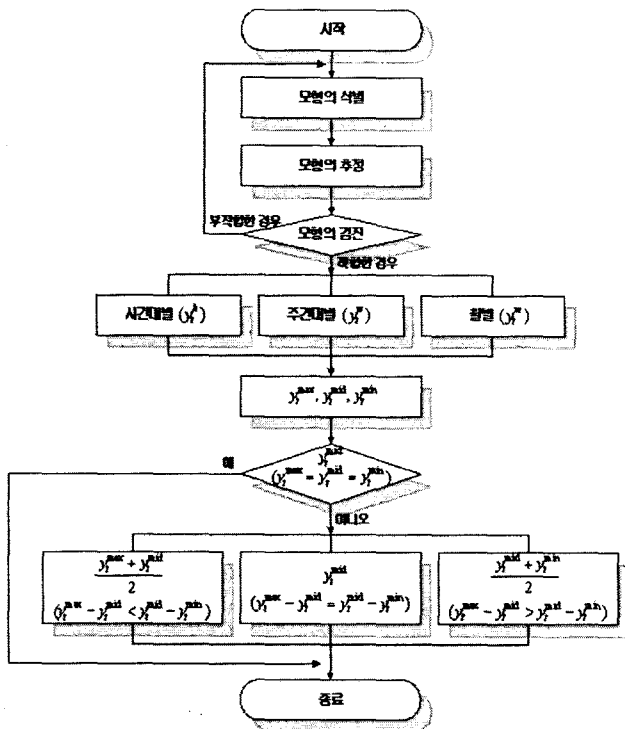


그림 1 ARIMA 모델을 이용한 SMP 예측 알고리즘 순서도
 Fig. 1 Flow-chart of SMP forecasting algorithm using ARIMA model

5. 사례 연구

본 논문에서는 한국전력거래소에서 제공하는 시간대별 계통한계가격 자료를 이용하여 2004년 8월 한 달간을 예측하여 제시한 알고리즘의 타당성을 검증한다. 정확한 예측은 모형에서 산출된 예측값과 실제값의 차이인 예측오차가 작은 것을 의미한다. 일정기간 동안 예측오차의 분석은 사용된 예측 알고리즘이 이 패턴에 적합한가를 지적해 준다. 만약 일정기간 동안 예측오차가 예측방법으로 오차의 확률분포를 이용했다면, 정확한 예측을 가능하게 하기 위해서는 오차의 크기를 측정하는 것이 중요하다[16]. 따라서 상대적인 오차의 크기를 나타내는 평균절대비율오차는 다음의 식으로 정의된다.

$$MAPE = \frac{\sum_{i=1}^n \left(\frac{|z_i - \hat{z}_i|}{z_i} \right) \times 100}{n} \quad (15)$$

여기서,

- n : 시계열 자료의 수
- z_i : 어떤 변수의 실제값
- \hat{z}_i : 어떤 변수의 예측값

제시한 방법론의 ARIMA 모형의 모수추정 및 예측치는 통계학적 범용 소프트웨어인 SAS Ver. 8.1을 사용하여 모의 시험한다. SAS의 ARIMA 프로시저에서는 ESTIMATE 문에서 CLS(Conditional Least Squares), ULS(Unconditional

Least Squares), ML(Maximum Likelihood) 중 하나를 지정할 수 있으며 디폴트는 CLS이다[13]. 이때, 시계열을 분석하기 위해서 일반적으로 필요한 시계열 자료 수는 최소 30개 이상이어야 정확한 예측을 할 수 있다.

5.1 2004년 8월 계통한계가격 예측

그림 2~9는 예측에 대한 모든 입력요소를 고려하여 2004년 8월 한 달간을 예측한 계통한계가격과 한국전력거래소에서 제공하는 실제 계통한계가격의 비교 차트를 나타낸다. 그림은 정확한 분석을 위해서 4주로 나누어서 보여주며 각각의 인자에 대한 예측값을 비교한다. 또한 예측오차를 최소화하는 보정 알고리즘을 통한 보정값을 실제 계통한계가격과 비교한다.

8월 첫째주의 실제 계통한계가격과 각각의 인자에 대한 예측값의 비교는 그림 2와 같다.

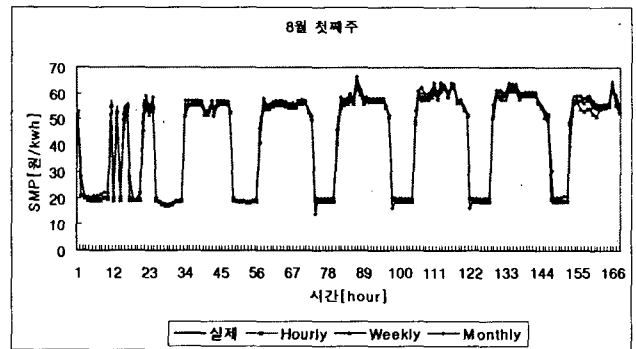


그림 2 2004년 8월 첫째주에 대한 실제값과 예측값 비교
 Fig. 2 Comparison of real and forecasting data (for the 1st week, Aug. 2004)

8월 첫째주의 실제 계통한계가격과 보정 알고리즘을 통한 보정값의 비교는 그림 3과 같다.

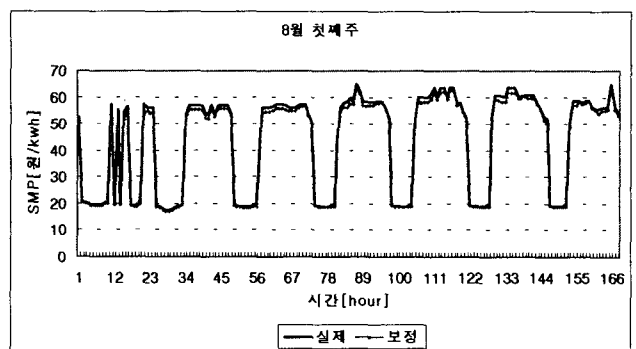


그림 3 2004년 8월 첫째주에 대한 실제값과 보정값 비교
 Fig. 3 Comparison of real and correction data (for the 1st week, Aug. 2004)

8월 둘째주의 실제 계통한계가격과 각각의 인자에 대한 예측값의 비교는 그림 4와 같다.

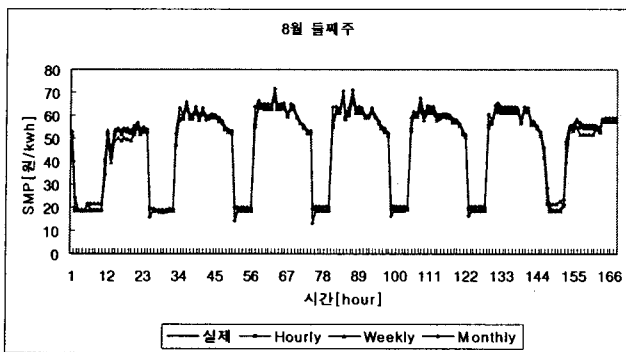


그림 4 2004년 8월 둘째주에 대한 실제값과 예측값 비교
 Fig. 4 Comparison of real and forecasting data (for the 2nd week, Aug. 2004)

8월 둘째주의 실제 계통한계가격과 보정 알고리즘을 통한 보정값의 비교는 그림 5와 같다.

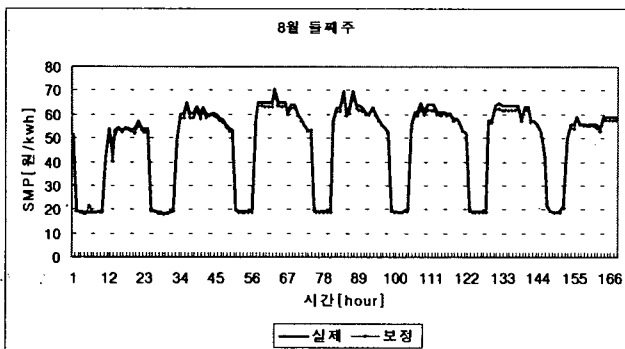


그림 5 2004년 8월 둘째주에 대한 실제값과 보정값 비교
 Fig. 5 Comparison of real and correction data (for the 2nd week, Aug. 2004)

8월 셋째주의 실제 계통한계가격과 각각의 인자에 대한 예측값의 비교는 그림 6과 같다.

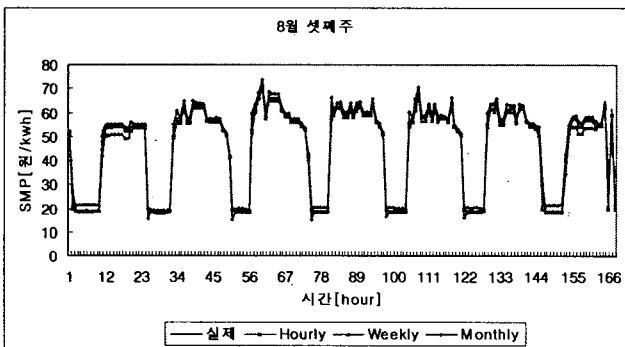


그림 6 2004년 8월 셋째주에 대한 실제값과 예측값 비교
 Fig. 6 Comparison of real and forecasting data (for the 3rd week, Aug. 2004)

8월 셋째주의 실제 계통한계가격과 보정 알고리즘을 통한 보정값의 비교는 그림 7과 같다.

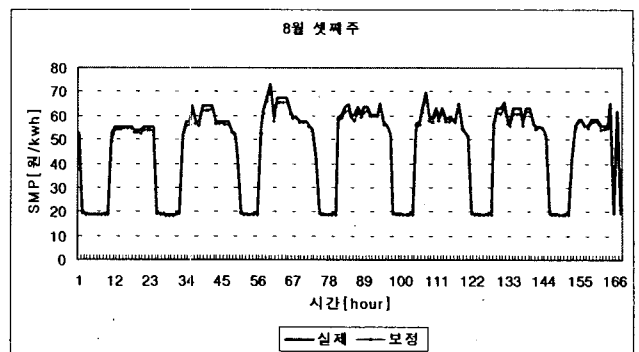


그림 7 2004년 8월 셋째주에 대한 실제값과 보정값 비교
 Fig. 7 Comparison of real and correction data (for the 3rd week, Aug. 2004)

8월 넷째주의 실제 계통한계가격과 각각의 인자에 대한 예측값의 비교는 그림 8과 같다.

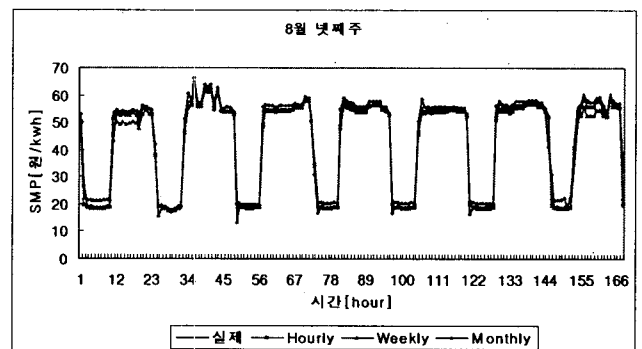


그림 8 2004년 8월 넷째주에 대한 실제값과 예측값 비교
 Fig. 8 Comparison of real and forecasting data (for the 4th week, Aug. 2004)

8월 넷째주의 실제 계통한계가격과 보정 알고리즘을 통한 보정값의 비교는 그림 9와 같다.

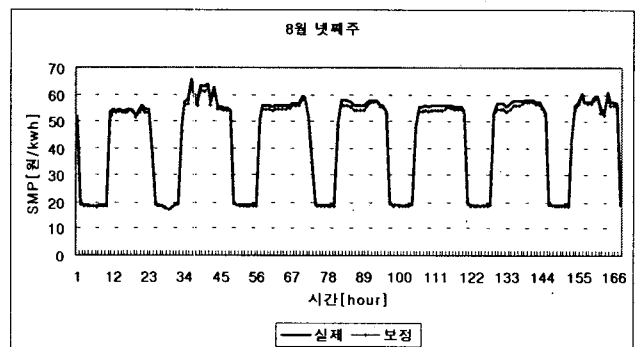


그림 9 2004년 8월 넷째주에 대한 실제값과 보정값 비교
 Fig. 9 Comparison of real and correction data (for the 4th week, Aug. 2004)

본 논문에서 제시한 예측보정 알고리즘의 타당성을 검증한 결과에 대해서 8월 한 달간의 요일별 분석은 표 2, 3과

표 2 2004년 8월 요일별 분석(1)

Table 2 Analysis of SMP forecasting data (Aug. 2004)

		일요일				월요일				화요일			
날짜		1일				2일				3일			
오차율	방법론	y_i^1	y_i^2	y_i^3	보정	y_i^1	y_i^2	y_i^3	보정	y_i^1	y_i^2	y_i^3	보정
	최대	4.19	14.34	46.92	13.94	4.13	3.22	6.65	3.62	3.89	3.42	9.38	3.45
	최소	2.00	2.05	0.00	0.05	2.20	1.60	0.00	1.03	2.19	1.18	0.00	0.88
	평균	2.72	6.01	11.77	2.95	2.48	2.31	0.69	2.33	2.47	2.22	0.71	2.27
날짜		8일				9일				10일			
오차율	방법론	y_i^1	y_i^2	y_i^3	보정	y_i^1	y_i^2	y_i^3	보정	y_i^1	y_i^2	y_i^3	보정
	최대	3.88	15.30	26.79	14.98	4.05	3.26	17.44	3.04	4.21	4.67	29.18	3.22
	최소	1.65	1.73	0.03	0.08	2.16	1.42	0.00	1.04	2.14	0.58	0.00	0.53
	평균	2.45	7.12	3.70	2.77	2.47	2.46	1.84	2.13	2.44	2.68	2.16	1.89
날짜		15일				16일				17일			
오차율	방법론	y_i^1	y_i^2	y_i^3	보정	y_i^1	y_i^2	y_i^3	보정	y_i^1	y_i^2	y_i^3	보정
	최대	4.12	15.10	41.98	7.94	4.11	3.25	18.37	2.96	4.15	5.88	20.85	3.65
	최소	2.03	2.39	0.00	0.52	2.06	1.31	0.00	1.11	0.96	0.22	0.00	0.37
	평균	2.44	8.78	3.06	2.02	2.47	2.61	1.89	2.11	2.37	2.83	2.13	1.81
날짜		22일				23일				24일			
오차율	방법론	y_i^1	y_i^2	y_i^3	보정	y_i^1	y_i^2	y_i^3	보정	y_i^1	y_i^2	y_i^3	보정
	최대	4.14	14.62	40.98	8.06	4.03	3.48	18.68	3.00	4.09	6.28	31.91	4.32
	최소	1.23	0.97	0.00	0.13	0.36	1.34	0.00	1.12	1.99	1.08	0.00	1.13
	평균	2.48	8.79	3.50	1.95	2.37	2.65	1.95	2.04	2.46	3.36	2.04	2.20
날짜		29일				30일				31일			
오차율	방법론	y_i^1	y_i^2	y_i^3	보정	y_i^1	y_i^2	y_i^3	보정	y_i^1	y_i^2	y_i^3	보정
	최대	4.12	14.90	23.06	17.93	4.16	3.66	18.85	3.65	4.14	4.99	20.53	2.89
	최소	2.24	2.35	0.00	0.48	1.93	1.49	0.05	1.14	0.91	0.39	0.00	0.88
	평균	2.46	7.67	2.59	2.79	2.48	2.68	2.15	2.22	2.36	2.92	1.57	2.05

표 3 2004년 8월 요일별 분석(2)

Table 3 Analysis of SMP forecasting data (Aug. 2004)

		수요일				목요일				금요일				토요일			
날짜		4일				5일				6일				7일			
오차율	방법론	y_i^1	y_i^2	y_i^3	보정	y_i^1	y_i^2	y_i^3	보정	y_i^1	y_i^2	y_i^3	보정	y_i^1	y_i^2	y_i^3	보정
	최대	3.88	5.49	28.27	3.54	4.09	5.09	15.15	3.44	4.13	5.36	15.18	3.48	4.09	9.44	59.60	3.20
	최소	2.19	0.34	0.00	0.48	2.10	0.65	0.00	0.61	2.23	0.69	0.00	0.51	1.97	1.53	0.00	0.37
	평균	2.46	3.06	2.04	1.78	2.44	3.04	1.47	2.02	2.45	3.11	1.14	2.01	2.44	4.95	4.33	1.91
날짜		11일				12일				13일				14일			
오차율	방법론	y_i^1	y_i^2	y_i^3	보정	y_i^1	y_i^2	y_i^3	보정	y_i^1	y_i^2	y_i^3	보정	y_i^1	y_i^2	y_i^3	보정
	최대	4.21	8.14	30.50	5.32	4.17	8.73	15.48	5.44	4.21	8.31	15.11	5.22	3.94	17.97	33.96	5.70
	최소	1.98	0.01	0.00	0.10	2.20	0.43	0.00	0.14	2.09	0.45	0.00	0.32	0.28	0.37	0.00	0.09
	평균	2.45	3.63	2.32	1.78	2.45	3.71	1.65	1.80	2.44	3.74	1.20	1.99	2.34	6.60	3.50	2.08
날짜		18일				19일				20일				21일			
오차율	방법론	y_i^1	y_i^2	y_i^3	보정	y_i^1	y_i^2	y_i^3	보정	y_i^1	y_i^2	y_i^3	보정	y_i^1	y_i^2	y_i^3	보정
	최대	4.24	9.26	20.67	5.06	4.20	8.11	14.30	5.15	4.20	9.03	15.03	5.62	4.27	15.54	83.56	9.40
	최소	0.95	0.34	0.00	0.25	1.87	0.36	0.00	0.57	2.00	0.30	0.00	0.47	2.22	1.03	0.00	0.14
	평균	2.35	3.88	2.05	1.87	2.43	4.02	1.79	1.99	2.44	4.20	1.39	2.08	2.68	8.64	12.06	2.62
날짜		25일				26일				27일				28일			
오차율	방법론	y_i^1	y_i^2	y_i^3	보정	y_i^1	y_i^2	y_i^3	보정	y_i^1	y_i^2	y_i^3	보정	y_i^1	y_i^2	y_i^3	보정
	최대	4.07	9.78	13.39	4.61	4.03	8.95	14.17	5.09	4.13	9.03	15.95	5.47	3.88	18.13	67.91	8.73
	최소	0.17	0.64	0.00	0.58	2.28	0.81	0.00	0.42	2.30	1.16	0.00	0.74	2.17	1.43	0.01	0.11
	평균	2.31	4.44	1.74	2.01	2.46	4.41	1.67	2.44	2.46	4.58	1.22	2.27	2.49	7.91	8.36	2.53

같다. 여기서, y_t^h 는 시간대별 ARIMA 모형을 나타내고, y_t^w 는 주간대별 ARIMA 모형을 나타내며, y_t^m 는 월별 ARIMA 모형을 나타낸다.

표 2, 3에서 예측보정값의 최대 오차일은 1일(일요일)이고 2.95[%]의 오차율을 갖는다. 또한 예측보정값의 최소 오차일은 11일(수요일)이며 1.78[%]의 작은 오차율을 갖는다. 여기서 오차율은 정식 (15)에 의해 계산된 상대적인 오차의 크기를 나타낸다.

5.2 사례연구 고찰

본 논문에서 제시한 ARIMA 모형별 예측값은 평균 3.27[%]의 비교적 작은 오차율을 가지고 있지만 특정 시간대에 큰 오차율을 가지는 것을 확인할 수 있다. 반면, 본 논문에서 제시한 보정 알고리즘을 통해 보정된 최종 계통한계가격의 오차율은 2.15[%]이다. 예측보정 알고리즘의 타당성을 검증한 결과 요약은 표 4와 같다.

표 4 각 입력 자료별 결과 요약

Table 4 Comparison of results for SMP forecasting

오차율 \ 방법론	시간대별	주간대별	월별	산술평균	보정 알고리즘
8월 평균	2.45	4.48	2.89	3.27	2.15
주중 평균	2.43	3.30	1.67	2.47	2.04
주말 평균	2.50	7.38	6.00	5.29	2.40

표 4에서 보정 알고리즘 구간에 대한 실제값과 예측값의 오차율은 8월 한 달간 평균 2.15[%]로 매우 작은 오차율을 나타내고 있으며, 2.04[%]의 주중 평균 오차율과 2.40[%]의 주말 평균 오차율을 나타내었다. 그림 2~9에서 볼 수 있듯이 예측보정 알고리즘을 통해 예측된 예측값의 오차율은 매우 작은 값을 갖는 것을 알 수 있으며, 이를 통해 예측 결과값이 매우 정확하다는 것을 확인할 수 있다.

5. 결론

본 논문에서는 정확한 계통한계가격 예측을 위해서 시계열 자료 예측기법의 하나인 자기회귀누적이동평균(ARIMA) 모형을 확장하여 사용하였다. 또한 예측보정 알고리즘을 통해 각 입력요소별로 예측된 계통한계가격의 예측오차를 최소화하였다. 본 논문에서 제안한 계통한계가격 예측 알고리즘의 타당성을 검증하기 위하여 한국전력거래소에서 제공하는 과거의 계통한계가격을 입력 자료로 사용하여 2004년 8월 한 달간을 예측하였으며, 이를 실제 계통한계가격과 비교·분석하였다. 제시한 방법론은 2004년 8월 한 달간 평균 오차율이 약 2.15[%]의 예측 결과값을 나타내었다. 하루 평균 최대 오차일은 1일(일요일)이고 2.95[%]의 오차율을 가지며, 최소 오차일은 11일(수요일)이고 1.78[%]의 매우 정확한 예측을 보였다. 이를 통해서 제시한 기법을 이용한 계통한계가격 예측값은 신뢰할 수 있으며 여러 시장참여들에게 그들의 목적에 부합할 수 있는 유용한 정보의 제공이 가능할 것으로 판단된다. 본 논문에서 제시한 예측보정 알고리즘은

안정적인 전력시장을 위한 단기간 수요예측에 활용될 수 있으며, 부하별로 다양하고 신뢰성 있는 수요예측에도 적용이 가능할 것이다. 또한 제시한 방법론을 통해 각 발전회사는 경쟁적 전력시장에서 자신의 이익을 최대화하기 위한 최적의 입찰전략을 수립하는데 이용할 수 있으며, 시장운영자는 전력시장을 예측하고 분석할 수 있을 것으로 판단된다. 향후 연구에서는 주말 또는 특정일의 예측 정확도가 주중에 비해 상대적으로 떨어지는 단점을 보완하고 가끔 돌출적인 예측값이 발생하는 문제를 해결하기 위한 새로운 예측 방법론이 모색되어야 할 것으로 생각된다.

감사의 글
이 논문은 산업자원부에서 시행한 전력산업 인프라구축지원 사업으로 수행된 논문입니다.

참 고 문 헌

- [1] J. Contreras, R. Espinola, F.J. Nogales, and A.J. Conejo, "ARIMA Models to Predict Next-Day Electricity Prices", *IEEE Transactions on Power Systems*, Vol. 18, No. 3, pp. 1014-1020, Aug. 2003.
- [2] J.M. Arroyo and A.J. Conejo, "Optimal Response of a Thermal Unit to an Electricity Spot Market", *IEEE Transactions on Power Systems*, Vol. 15, pp. 1098-1104, Aug. 2000.
- [3] B.R. Szkuta, L.A. Sanabria, and T.S. Dillon, "Electricity Price Short-Term Forecasting Using Artificial Neural Network", *IEEE Transactions on Power Systems*, Vol. 14, No. 3, pp. 851-857, Aug. 1999.
- [4] A. Martini, P. Pelacchi, L. Pellegrini, M.V. Cazzol, A. Garzillo, and M. Innorta, "A Simulation Tool for Short Term Electricity Market", *PICA 2001. 22nd IEEE Power Engineering Society International Conference on Innovative Computing for Power Electric Energy Meets the Market*, pp. 112 -117, May 2001.
- [5] A.A.El Desouky and M.M.EIKateb, "Hybrid adaptive techniques for electric-load forecast using ANN and ARIMA", *IEE Proc.-Gener. Transm. Distrib.*, Vol. 147, No. 4, pp. 213-217, July 2000.
- [6] 이상래, *환율예측모형의 비교연구*, 한국과학기술원 석사학위논문, 2005.
- [7] 김연형, *시계열 분석*, 자유아카데미, 1994.
- [8] M.T. Hagan and S.M. Behr, "The Time Series Approach to Short Term Load Forecasting", *IEEE Transactions on Power Systems*, Vol. 2. pp. 785-791, Aug. 1987.
- [9] 박종배, 이찬주, 이재용, 신중린, 이창호, "ARIMA 모형에 기초한 수요실적자료 보정기법 개발", *대한전기학회 논문지*, 제 53권, 7호, pp. 405-413, 2004

- [10] Chan-Joo Lee, Jong-Bae Park, Jae-Yong Lee, Joong-Rin Shin, and Chang-Ho Lee, "A Hybrid Correction Technique of Missing Load Data Based on Time Series Analysis", *KIEE Transaction on System and Control*, Vol. 3-A, No. 1, pp. 123-132, 2003.
- [11] A.J. Conejo, M.A. Plazas, R. Espinola, and A.B. Molina, "Day-Ahead Electricity Price Forecasting Using the Wavelet Transform and ARIMA Models", *IEEE Transactions on Power Systems*, Vol. 20, No. 2, pp. 1035-1042, May 2005.
- [12] G.E.P. Box, G.M. Jenkins, and G.C. Reinsel, *Time Series Analysis : Forecasting and Control*, Prentice Hall, 1994.
- [13] A. Pole, M. West, and J. Harrison, *Applied Bayesian Forecasting and Times Series Analysis*, Chapman & Hall, 1994.
- [14] 박유성, 김기환, *SAS/ETS를 이용한 시계열자료분석 I*, 자유아카데미, 2002.
- [15] Wei and W.S. William, *Time Series Analysis : Univariate And Multivariate Methods*, Addison-Wesley, 2005.
- [16] K.L. Lo and Y.K. Wu, "Risk Assessment Due to Local Demand Forecast Uncertainty in The Competitive Supply Industry", *IEE Proc.-Gener. Transm. Distrib.*, Vol. 150, No. 5, pp. 573-581, Sept. 2003.

저 자 소 개



김 대 용(金 大 容)

1978년 8월 26일생. 2004년 한국교육개발원 전기공학과 졸업. 현재 건국대 대학원 전기공학과 석사과정.
Tel.: 02-458-4778 , Fax : 02-444-1418
E-mail: sports@konkuk.ac.kr



이 찬 주(李 瓊 柱)

1974년 12월 15일생. 2000년 안양대 공대 전기공학과 졸업. 2002년 건국대 대학원 전기공학과 졸업(석사). 현재 동 대학원 전기공학과 박사과정.
Tel.: 02-458-4778 , Fax : 02-444-1418
E-mail: jubily@konkuk.ac.kr



정 윤 원(鄭 允 源)

1977년 5월 8일생. 2002년 건국대 공대 전기공학과 졸업. 2004년 동 대학원 전기공학과 졸업(석사). 현재 동 대학원 전기공학과 박사과정.
Tel.: 02-458-4778 , Fax : 02-444-1418
E-mail: ywjeong@konkuk.ac.kr



박 종 배(朴 宗 培)

1963년 11월 24일생. 1987년 서울대 공대 전기공학과 졸업. 1989년 동 대학원 전기공학과 졸업(석사). 1998년 동 대학원 전기공학과 졸업(박사). 현재 건국대학교 공과대학 전기공학과 조교수.
Tel.: 02-450-3483, Fax: 02-444-1418
E-mail: jbaepark@konkuk.ac.kr



신 중 린(慎 重 麟)

1949년 9월 22일생. 1977년 서울대 공대 전기공학과 졸업. 1984년 동 대학원 전기공학과 졸업(석사). 1989년 동 대학원 전기공학과 졸업(박사). 현재 건국대학교 공과대학 전기공학과 교수.
Tel.: 02-450-3487, Fax: 02-444-1418
E-mail: jrshin@konkuk.ac.kr