

전기 자동차 전력 수요 예측 연구[†]

최수임¹ · 손흥구² · 김삼용³

¹³중앙대학교 응용통계학과 · ²한국교통연구원 항공교통연구본부

접수 2018년 7월 11일, 수정 2018년 9월 10일, 게재확정 2018년 9월 13일

요 약

최근 들어 지구 온난화, 미세먼지 등의 환경 오염 문제로 인하여 정부의 친환경 정책에 많은 관심이 쏟아지고 있다. 친환경 에너지뿐만 아니라 전기 자동차 사용을 활성화 시키기 위하여 보조금 지원 정책 등 여러 정책을 펼치고 있는 실정이다. 전기 자동차가 활성화되면서 자동차 충전을 위한 전력 수요량이 증가하게 될 것이므로 원활한 전력 공급 계획을 수립하기 위해서는 전기 자동차 충전을 위한 전력량을 정확히 예측하는 것이 중요하다. 본 논문에서는 서울특별시와 제주특별자치도의 전기 자동차 충전소 데이터를 수집하여 지역별 충전 전력량 1일 예측 연구를 하였다. 충전 전력 예측을 위하여 시계열 모형인 ARIMA, ARIMAX, ARIMA-GARCH, ARIMAX-GARCH, SARIMA, SARIMAX 모형을 비교하였다. 모형들의 비교하는 기준으로는 MAPE, SMAPE를 사용하였다.

주요용어: 시계열 모형, 전기 자동차, 충전 전력량.

1. 서론

전세계적으로 가장 많이 사용되는 화석 연료가 고갈되고 있을 뿐만 아니라 화석 연료를 사용함에 따라 방출되는 이산화탄소 배출량으로 인하여 지구 온난화 등이 심각한 문제로 대두되었다. 심각해지는 환경 문제에 대비하기 위하여 여러 국가들이 힘을 모아 기후변화협약을 맺으며 각 국가별로 여러 정책을 펼치고 있다. 특히, 많은 국가들이 전기 자동차에 지대한 관심을 쏟으며 나라별로 전기 자동차 관련 여러 정책을 펼치고 있는 실정이다.

OECD와 IEA (2016)에 따르면 2015년 기준 주요 국가들의 전기 자동차 신규 등록 현황을 조사하였을 때, 중국 (20.7만 대), 미국 (11.4만 대), 네덜란드 (4.4만 대) 순으로 중국이 가장 많은 신규 등록 건수를 보였다. 중국의 경우, 국가에서 자동차 산업에 개입을 하여, 내연 기관 자동차에 대한 여러 규제들을 가하면서 전기 자동차 사용을 장려하기 위하여 많은 보조금을 지급한다. 한국의 경우에도, 보조금을 지급하고 홍보하며 전기 자동차 사용을 장려하는 추세이고 정부에서도 점차적으로 전기 자동차 수를 증가시키기 위한 여러 방법을 찾고 있는 중이다.

현재까지 진행 중인 연구 결과를 보면 우선, Arias와 Bae (2016)은 전기 자동차의 경우, 이산화탄소 배출에 따른 환경문제를 해결하는 것 이외에도 현재 한국이 석유 수입 의존도가 높은 국가인 만큼 에너

[†] 이 논문은 2018년도 정부 (교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (NRF-2016R1D1A1B01014954).

¹ (06974) 서울특별시 동작구 흑석로 84, 중앙대학교 응용통계학과, 석사과정.

² (30147) 세종 시청대로 370 나라키움 세종국책연구단지, 한국교통연구원 항공교통연구본부, 박사후연구원.

³ 교신저자: (06974) 서울특별시 동작구 흑석로 84, 중앙대학교 응용통계학과, 교수.

E-mail: sahm@cau.ac.kr

지 안보를 높일 수 있는 중요한 역할을 할 수 있다고 하였다. 전 세계의 추세에 따라, 한국도 전기 자동차 사용을 증가시키려고 하고 있다. 이로 인하여 전기 자동차 사용자가 점점 증가하게 되면서 전기 자동차 충전을 위한 전력 수요가 점점 상승할 것이다. 또한 전기 자동차가 증가함에 따라 충분한 인프라 구축도 필요한 실정이다. 효율적인 전기 자동차 정책 수행을 위해서는 전기 자동차 충전을 위한 전력 수요에 대한 예측이 시급하다. 그래서 현재까지 충전 인프라 구축을 위한 전력 수요 예측 연구가 진행되어 왔다.

Mo (2012)는 ARIMA모형을 이용하여 전기 자동차가 확산되었을 상황의 전력 수요를 예측하기 위해 여러 시나리오를 짜고 이를 시뮬레이션 하였다. 전체 전력 수요의 3%에 해당하는 양으로 영향이 미비하였지만 V2G 프로그램 도입 시에는 심야 시간대의 전력 수요량이 크게 증가하여 기존의 발전소 용량으로는 감당하지 못 할 수도 있음을 시사했다. Majidpour 등 (2015)는 고객 프로파일 데이터와 충전소 측정 데이터를 이용하여 전기 자동차 충전을 위한 전력 수요를 예측했다. TWDP-NN (time weighted dot product based nearest neighbor), MPSF (modified patten sequence forecasting), SVR (support vector regression), RF (random forest) 모형을 사용하고 예측 정확도를 평가하는 기준으로 SMAPE (syemmetric mean absolute percentage error)를 사용했다. 이 논문에서는 TWDP-NN과 MPSF가 가장 좋은 예측력을 보였다. Qian 등 (2010)은 전기 자동차 사용 용도별 비율과 배터리 타입, 충전 특성을 구별하여 전기 자동차 충전 전력 예측하는 연구를 했다. 여름철과 겨울철의 영향만을 확인하였고 전기 요금 구조가 충전할지 말지 결정한다고 가정하였다. Arias 등 (2017)은 서울 도로망 CCTV를 이용한 데이터를 가지고 시공간적이 충전 전력 수요 예측 모형을 연구하였다. Markov chain traffic model과 순간이동 접근법을 이용하여 충전 위치, 주행 상황을 통하여 전기 자동차 충전소 도착률을 계산하였다. 시간의 경우, 아침, 오후, 저녁 3가지 범위로 나누었다. 충전 전의 SOC (state of charging) 상태를 가우시안 분포의 난수 생성을 통하여 결정하였다. 이를 바탕으로 하여 여러 시나리오 별로 분석을 실시하였고 교통 데이터를 사용하는 것이 예측 정확도를 높일 수 있다는 것을 증명했다. Xydas 등 (2013)는 support vector machine (SVM)과 인공 지능 기술을 이용하여 전기 자동차 충전 전력 수요에 관하여 단기 예측 연구를 진행했다. 기존의 Monte Carlo 방법과 결과를 비교하였는데 1년 중 50주 정도의 시간별 데이터가 사용되었고 MAPE와 RMSE를 기준으로 하여 모형의 성능을 평가했다. MAPE를 기준으로 하였을 때, SVM이 3.69, Monte Carlo가 8.99로 SVM이 우수한 성능을 보였다.

본 논문에서는 여러 시계열 모형을 이용하여 서울, 제주 지역 전기 자동차 충전 전력 수요를 예측하며, 예측을 위한 외생변수로는 공휴일, 전기 자동차 등록 대수, 냉난방지수, 계절 (여름, 겨울) 더미 변수를 고려할 것이다. 다음 2장에서는 전기 자동차 충전 전력 예측을 위한 시계열 모형을 설명하고자 한다. 3장에서는 본 논문에서 활용된 전기 자동차 충전 전력 데이터 외에 외생변수들을 소개하고 적합한 모형 및 분석 결과에 대해 논할 것이다. 4장에서는 3장의 분석 결과를 바탕으로 하여 모형들의 성능 차이를 확인하고 5장에서는 결론과 향후 연구 방향에 대하여 제안할 것이다.

2. 예측 모형

본 논문에서는 서울, 제주 지역의 전기 자동차 충전 전력을 일별로 예측하기 위하여 시계열 모형 중 ARIMA, ARIMAX, Seasonal ARIMA, Seasonal ARIMAX, ARIMA-GARCH, ARIMAX-GARCH 모형을 이용하였다. 2장에서는 각 모형별 성능 분석 및 비교를 위하여 사용된 모형을 소개하고자 한다.

2.1. ARIMA 모형

ARIMA (auto-regressive integrated moving average) 모형은 비정상 시계열을 d차 차분하여 정상 시계열로 변환한 이후에 정상 시계열 모형은 ARMA (auto-regressive moving average)모형에 적합시

킨 모형을 말한다. 시계열 Y_t 가 평균이 μ 인 $ARIMA(p,d,q)$ 모형식은 다음과 같이 표현된다.

$$\phi_p(B)(1-B)^d(Y_t - \mu) = \theta_q(B)\epsilon_t.$$

Y_t 는 t 시점에서의 관측값을, B 는 후진 연산자 (back-shift operator)를 의미한다. p 는 AR 차수, q 는 MA 차수이며, d 는 차분 차수를 의미한다. $\phi_p(B)$, $\theta_q(B)$ 는 각각 p 차, q 차 다항식으로 각각의 식은 식 (2.1), (2.2) 와 같다.

$$\phi_p(B) = 1 - \phi_1 B - \dots - \phi_p B^p, \quad (2.1)$$

$$\theta_q(B) = 1 - \theta_1 B - \dots - \theta_q B^q. \quad (2.2)$$

2.2. ARIMAX 모형

ARIMAX (auto-regressive integrated moving average with exogeneous variable) 모형은 2.1절에서 설명한 ARIMA 모형에 외생변수를 추가한 모형을 말한다. 시계열 Y_t 가 평균이 μ 이고 외생변수 x_{it} 가 k 개인 ARIMAX(p,d,q) 모형식은 다음과 같이 표현된다.

$$\phi_p(B)(1-B)^d(Y_t - \mu) = \theta_q(B)\epsilon_t + \sum_{i=1}^k \gamma_i x_{it}.$$

각각의 요소들의 2.1절과 동일하며 γ_i 는 외생변수 x_{it} 의 계수들을 의미한다. 본 논문에서 외생변수로 공휴일, 전기 자동차 등록대수, 냉난방지수, 여름, 겨울 더미 변수를 사용하였다. $\phi_p(B)$, $\theta_q(B)$ 는 각각 p 차, q 차 다항식으로 각각의 식은 식 (2.1), (2.2) 와 같다.

2.3. ARIMA-GARCH 모형

ARIMA-GARCH (auto-regressive integrated moving average generalized autoregressive conditionally heteroskedastic) 모형은 평균 모형은 ARIMA이고, 분산 모형은 GARCH인 경우 이다. GARCH (generalized auto-regressive conditionally heteroskedastic) 모형은 Engle (1982)이 기존 가정인 등분산이 아니라 이분산성을 고려하기 위하여 제안한 ARCH (autoregressive conditionally heteroskedastic) 모형을 현재 분산이 과거 분산의 영향을 받는 모형으로 확장하여 일반화한 것으로 Bollerslev (1986)이 제안하였다. GARCH(r,m)에서 r 이 0인 경우, ARCH(m)과 같아진다. 이를 고려한 것으로 ARIMA(p,d,q)-GARCH(r,m) 모형의 식은 다음과 같다.

$$\begin{aligned} \phi_p(B)(1-B)^d(Y_t - \mu) &= \theta_q(B)\epsilon_t, \\ \epsilon_t = e_t \sqrt{h_t}, \quad h_t &= \alpha_0 + \sum_{i=1}^m \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^r \beta_i h_{t-i}, \end{aligned}$$

여기서, m 은 ARCH, r 은 GARCH 부분의 차수를 뜻한다. ARIMA 모형 부분의 $\phi_p(B)$, $\theta_q(B)$ 는 각각 p 차, q 차 다항식으로 각각의 식은 식 (2.1), (2.2) 와 같이 나타난다.

GARCH의 경우, $e^t \sim WN(0, \sigma^2)$ 이며 모수들이 음이 아니어야하는 비음조건 (non-negative condition, $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_i \geq 0$)을 만족해야한다.

2.4. ARIMAX-GARCH 모형

ARIMAX-GARCH (auto-regressive integrated moving average with exogenous variables generalized autoregressive conditionally heteroskedastic) 모형은 평균 모형은 기존의 ARIMA 모형에서 외생 변수의 효과를 고려한 ARIMAX 모형이고, 분산 모형은 GARCH인 경우이다. 이 모형의 경우에는 시계열 모형의 기존 가정인 등분산이 아니라 이분산성을 고려한 것으로 ARIMAX(p, d, q)-GARCH(r, m) 모형의 식은 다음과 같다.

$$\begin{aligned}\phi_p(B)(1-B)^d(Y_t - \mu) &= \theta_q(B)\epsilon_t + \sum_{i=1}^k \gamma_i x_{it}, \\ \epsilon_t &= e_t \sqrt{h_t}, \quad h_t = \alpha_0 + \sum_{i=1}^m \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^r \beta_i h_{t-i}.\end{aligned}$$

각 요소들은 2.3절과 동일하다. $\phi_p(B)$, $\theta_q(B)$ 는 각각 p 차, q 차 다항식으로 각각의 식은 식 (2.1), (2.2)와 같이 나타난다. GARCH의 경우, 모수들이 음이 아니어야하는 비음조건 (non-negative condition, $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_i \geq 0$)을 만족해야한다.

2.5. SARIMA 모형

SARIMA (seasonal auto-regressive integrated moving average) 모형은 Box와 Jenkins (1994)에 의하여 제안된 모형으로 주기성이 존재하는 비정상 시계열에 적용하는 모형이다. 시계열 Y_t 가 평균이 μ 인 ARIMA(p, d, q)(P, D, Q)_s를 따를 때의 모형에 대한 식은 다음과 같다.

$$\phi_p(B)\Phi_P(B^S)(1-B)^d(1-B^S)^D(Y_t - \mu) = \theta_q(B)\Theta_Q(B^S)\epsilon_t.$$

Y_t 는 t 시점에서의 관측값을, B 는 후진 연산자 (back-shift operator)를 의미한다. p 는 AR 차수, q 는 MA 차수, P 는 seasonal AR 차수, Q 는 Seasonal MA 차수이며, S 는 계절 주기를 의미한다. $\phi_p(B)$, $\theta_q(B)$, $\Phi_P(B^S)$, $\Theta_Q(B^S)$ 는 각각 p 차, q 차, P 차, Q 차 다항식을 의미하며, d 는 차분 차수, S 는 계절 차분 차수를 의미한다. 각각의 식은 식 (2.1), (2.2), (2.3), (2.4) 와 같다.

$$\Phi_P(B^S) = 1 - \Phi_1 B^S - \dots - \Phi_P B^{PS}, \quad (2.3)$$

$$\Theta_Q(B^S) = 1 - \Theta_1 B^S - \dots - \Theta_Q B^{QS}. \quad (2.4)$$

2.6. SARIMAX 모형

SARIMAX (seasonal auto-regressive integrated moving average with exogenous variable) 모형은 기존의 SARIMA 모형에 외생변수를 추가로 고려한 모형이다. 시계열 Y_t 가 평균이 μ 인 ARIMAX(p, d, q)(P, D, Q)_s의 식은 다음과 같다.

$$\phi_p(B)\Phi_P(B^S)(1-B)^d(1-B^S)^D(Y_t - \mu) = \theta_q(B)\Theta_Q(B^S)\epsilon_t + \sum_{i=1}^k \gamma_i x_{it}.$$

각 요소들의 의미는 2.5절과 같다. γ_i 는 외생변수 x_{it} 의 계수이다. 본 논문에서 외생변수로 공휴일, 전기 자동차 등록대수, 냉난방지수, 여름, 겨울 더미 변수를 사용하였다. $\phi_p(B)$, $\theta_q(B)$ 는 각각 p 차, q 차 다항식으로 각각의 식은 다음과 같다. $\phi_p(B)$, $\theta_q(B)$, $\Phi_P(B^S)$, $\Theta_Q(B^S)$ 는 각각 p 차, q 차, P 차, Q 차 다

항식을 의미하며, d 는 차분 차수, S 는 계절 차분 차수를 의미한다. 각각의 식은 식 (2.1), (2.2), (2.3), (2.4) 와 같이 나타낼 수 있다.

3. 자료 분석

본 논문에서 사용된 데이터는 서울특별시와 제주특별자치도의 전기 자동차 충전 전력 데이터이다. 전기 자동차 충전 전력 데이터는 한국환경공단에서 제공받았으며 전국의 충전소에서 전기 자동차 사용자가 전기 자동차를 1번 충전할 때 축적된 데이터로 전국의 시, 군, 구 충전소 별로 수집된 데이터이다. 본 논문에서는 일별 예측을 하기 위하여 시·도 별로 데이터를 충전 시작 시간을 기준으로 일별로 충전 전력을 합하였으며 전국 시·도 중 일별 충전 전력이 연속적으로 존재하는 서울 특별시와 제주특별 자치도의 경우만 고려하기로 하였다. 기간은 2014년 1월 1일에서 2017년 6월 30일까지이고 2014년 1월에서 2016년 12월까지는 트레이닝셋으로 2017년 1월에서 2016년까지는 테스트셋으로 사용했다. 트레이닝 셋의 경우에는 R에서 `tsoutlier` 함수를 이용하여 이상치 검증 및 보정을 실시하였다. Figure 3.1에서 검정 그래프가 실제 데이터이고 회색 그래프가 보정된 데이터로 시계열 데이터의 분산을 안정화 시키기 위하여 로그변환을 실행하였다.

Table 3.1 Variables

Variable	Explanation
Y_t	Daily charging power for electric vehicles, log transformation
$\ln.ev$	Number of electric vehicles registration, log transformation
CDD	Cooling index
HDD	Heating index
$Holiday$	Weekday is 0, weekend or holiday is 1.
$Summer$	Summer is 1, others are 0.
$Winter$	Winter is 1, others are 0.

$$CDD = \begin{cases} T_t - 24, & T_t \geq 24 \\ 0, & T_t < 24, \end{cases} \quad HDD = \begin{cases} 0, & T_t > 18 \\ 18 - T_t, & T_t \leq 18. \end{cases} \quad (3.1)$$

전력 충전량 이외에 외생 변수로 사용된 변수는 Table 3.1과 같다. 기존 논문들에서와 같이 전기 자동차 사용자들의 충전 패턴이 주말 및 공휴일에 따른 변화가 있을 것이라 예상하여 공휴일 더미 변수 (평일은 0, 주말 및 공휴일은 1)를 외생 변수로 추가하였고 여름철, 겨울철에 따라 다른 영향이 있다는 것이 나타나 여름철, 겨울철에 대한 영향을 추가하기 위해 여름, 겨울에 관하여 각각 더미 변수를 만들었다. 또한 일반적으로 전력 사용량에 영향을 주는 냉·난방지수를 일 평균 기온 (T_t)을 사용하는 식 (3.1)을 이용하여 계산하였다. 냉방 지수와 난방 지수는 실내에서 사람이 쾌적하게 지낼 수 있는 기온을 기준 온도로 하여 계산된다. 사회적, 경제적 여건, 목적, 연령 등에 따라 다르게 나타나는데 국내의 경우에는 냉방 지수의 기준 온도는 24°C, 난방 지수의 기준 온도는 18°C로 사용한다. 전기 자동차 충전량의 경우, 겨울철이나 여름철의 경우 냉난방 장치 사용으로 배터리 전력 소모량이 큰 것을 고려하여 변수로 추가하였는데 기온 이외에 하루 동안 충전한 전기 자동차 수등 다른 외부적인 요인으로 인하여 변동이 크기 때문에 plot으로는 큰 상관성이 보이지 않아 생략하였다. 전기 자동차 충전 전력에 가장 크게 영향을 줄 것이라 고려되는 전기 자동차 등록 대수를 추가하려 하였으나 일별 전기 자동차 등록 대수는 데이터가 존재하지 않아 월별 전기 자동차 등록 대수를 기준으로 계산한 일별 전기 자동차 수를 전기 자동차 충전 전력 예측을 위해 외생 변수로 추가하였다. 전기 자동차 대수 또한 충전 전력량과 같이 로그 변환된 것을 사용하였다.

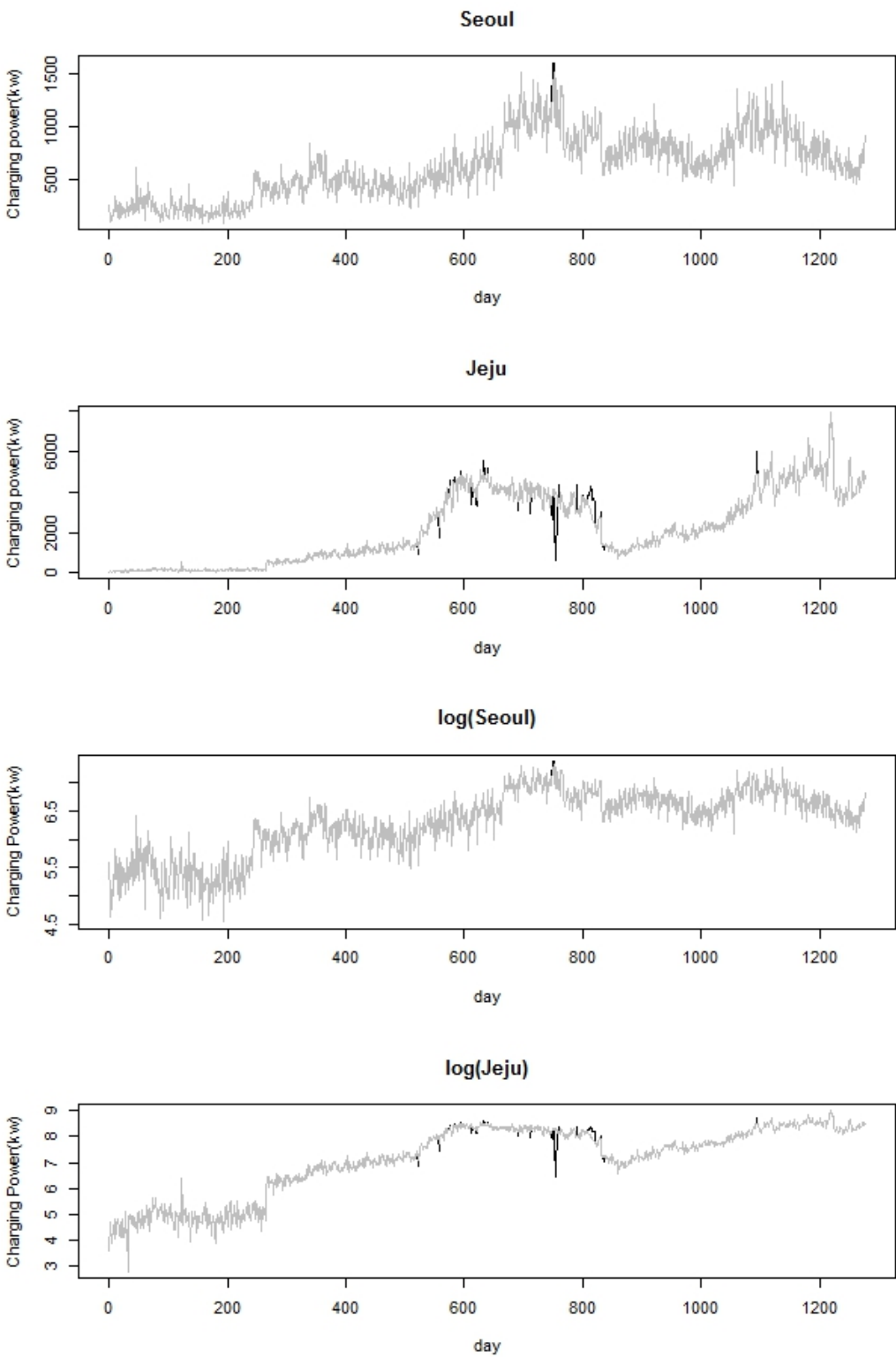


Figure 3.1 Charging power and transformation by region

일별 전기 자동차 충전 전력량 데이터를 이용하여 ARIMA, ARIMA-GARCH, ARIMAX, ARIMAX-GARCH, SARIMA, SARIMAX 모델을 이용한 전기 자동차 충전을 위해 사용된 일별 전력을 예측하였다. 전기 자동차 충전의 특성상 충전기 종류가 AC상, DC콤보, DC차데모 총 3가지의 종류가 있다. 그러나 충전 전력량에는 차이가 없었으므로 충전기 종류에 상관없이 일별로 모두 합하여 예측하는데 사용하였다. Figure 3.1에서 전기 자동차 충전량 데이터의 plot을 보면 추세가 있으므로 이를 제거하기 위해 차분을 한 모델을 적합시켰다. 지역별로 각각 분석결과를 보이고 다음 절에서 모형별로 비교하였다. 모형별 적합한 차수의 기준은 AIC (Akaike information criterion)를 기준으로 하여 결정하였다. 모형 적합 결과에서 외생변수들의 계수들의 값이 gamma로 나타나 있는데 gamma1은 전기 자동차 등록 대수, gamma2는 냉방지수, gamma3는 난방지수, gamma4은 여름 더미 변수, gamma5는 겨울 더미 변수를 나타낸다.

3.1. 서울 지역

서울 지역의 경우, ARIMA(6,1,5)가 가장 적합한 모형으로 나타났다. Table 3.2에 해당 모형의 계수들에 대해 나타나 있다.

Table 3.2 Parameter estimations of the ARIMA(6,1,5) model in Seoul

Parameter	Estimation	S.E.	p-value
ar1	-0.398	0.040	0.0000
ar2	0.351	0.047	0.0000
ar3	-0.581	0.020	0.0000
ar4	-0.916	0.022	0.0000
ar5	0.163	0.049	0.0008
ar6	0.021	0.037	0.5716
ma1	-0.309	0.027	0.0000
ma2	-0.694	0.018	0.0000
ma3	0.760	0.009	0.0000
ma4	0.517	0.010	0.0000
ma5	-0.828	0.017	0.0000

다음은 ARIMA-GARCH 모형을 적합한 결과이다. ARIMA 모형과는 달리 이분산성을 고려한 모형으로 서울 지역에서는 ARIMA(6,1,5)-GARCH(1,1)이 가장 적합한 모형으로 나타났다. Table 3.3는 ARIMA(6,1,5)-GARCH(1,1) 모형에 대한 계수가 나타나있다.

Table 3.4는 ARIMAX 모형에 대한 결과가 나타나있다. 사용된 입력변수들을 여러 가지로 조합하여 ARIMAX 모형을 적합하였다. 가장 적합한 ARIMAX모형은 ARIMAX(6,1,5) 모형이다. 유의수준 0.05 하에서 냉난방 지수와 여름 변수가 유의하게 나타났다.

ARIMAX-GARCH 모형의 적합한 결과, ARIMAX(6,1,5)-GARCH(1,1) 모형이 가장 적합한 것으로 나타났다. Table 3.5는 적합하다고 판단된 ARIMAX(6,1,5)-GARCH(1,1) 모형의 계수에 대하여 나타나있다. ARIMAX-GARCH 모형에서는 전기자동차 등록 대수와 냉난방지수 모두 유의한 변수로 나타났다.

Table 3.3 Parameter estimations of the ARIMA(6,1,5)-GARCH(1,1) model in Seoul

Parameter	Estimation	S.E.	p-value
mu	5.578	0.062	0.0000
ar1	1.242	0.002	0.0000
ar2	-0.319	0.004	0.0000
ar3	-0.884	0.008	0.0000
ar4	0.638	0.011	0.0000
ar5	0.067	0.010	0.0000
ar6	-0.087	0.003	0.0000
ma1	-1.977	0.000	0.0000
ma2	1.223	0.000	0.0000
ma3	0.618	0.004	0.0000
ma4	-1.304	0.000	0.0000
ma5	0.509	0.000	0.0000
omega	0.000	0.000	0.0910
alpha1	0.019	0.002	0.0000
beta1	0.978	0.001	0.0000

Table 3.4 Parameter estimations of the ARIMAX(6,1,5) model in Seoul

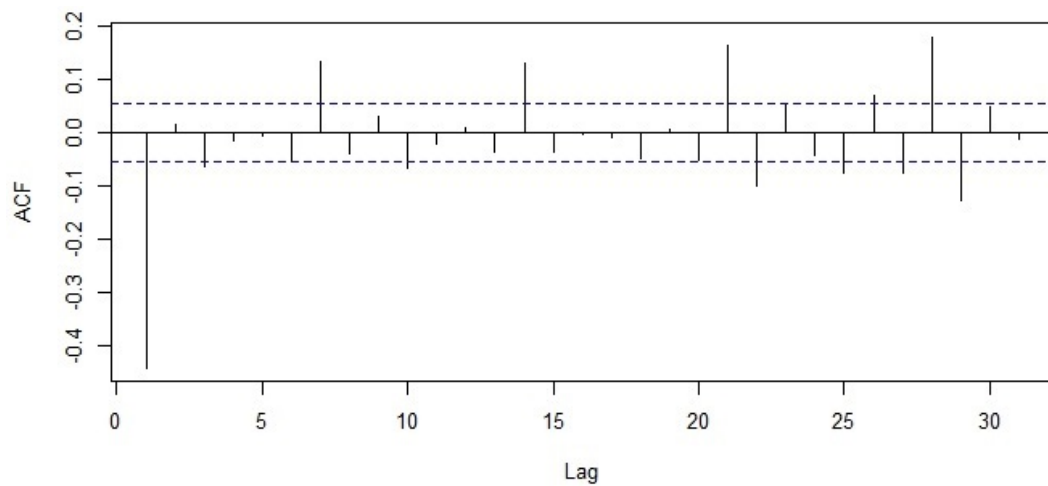
Parameter	Estimation	S.E.	p-value
ar1	-0.477	0.040	0.0000
ar2	0.275	0.047	0.0000
ar3	-0.580	0.020	0.0000
ar4	-0.954	0.021	0.0000
ar5	0.069	0.049	0.0057
ar6	-0.014	0.036	0.4299
ma1	-0.293	0.026	0.0000
ma2	-0.704	0.013	0.0000
ma3	0.757	0.009	0.0000
ma4	0.532	0.016	0.0000
ma5	-0.833	0.024	0.0000
gamma1	1.238	0.650	0.0553
gamma2	0.042	0.009	0.0000
gamma3	0.017	0.002	0.0000
gamma4	-0.123	0.046	0.0071
gamma5	0.042	0.047	0.3509

Seasonal ARIMA (SARIMA) 모형을 적합하기에 앞서 주기를 알아보기 위하여 차분된 데이터의 sample autocorrelation plot을 그렸다. Figure 3.2를 보면 주기가 7 인 것으로 나타났다. Table 3.6은 seasonal ARIMA (SARIMA) 모형을 적합한 결과로 가장 적합한 모형으로는 ARIMA(2,1,1)(0,1,1)[7]모형으로 나타났다.

외생 변수를 고려하여 seasonal ARIMAX-SARIMAX (SARIMAX) 모형을 적합한 결과, ARIMAX(1,1,1)(0,1,1)[7] 모형이 가장 적합한 것으로 나타났다. 그에 대한 계수의 값은 Table 3.7과 같다. 서울 지역의 SARIMAX 모형의 경우, 난방지수와 여름 더미 변수에 대하여 유의하게 나타났다.

Table 3.5 Parameter estimations of the ARIMAX(6,1,5)-GARCH(1,1) model in Seoul

Parameter	Estimation	S.E.	p-value
mu	0.215	0.451	0.6330
ar1	1.223	0.007	0.0000
ar2	-0.308	0.013	0.0000
ar3	-0.827	0.010	0.0000
ar4	0.588	0.018	0.0000
ar5	0.079	0.017	0.0000
ar6	-0.048	0.009	0.0000
ma1	-2.039	0.00	0.0000
ma2	1.332	0.000	0.0000
ma3	0.525	0.003	0.0000
ma4	-1.281	0.000	0.0000
ma5	0.517	0.000	0.0000
gamma1	0.824	0.079	0.0000
gamma2	0.044	0.008	0.0000
gamma3	0.017	0.002	0.0000
gamma4	-0.080	0.044	0.0668
gamma5	0.024	0.040	0.5430
alpha1	0.022	0.002	0.0000
beta1	0.974	0.002	0.0000

**Figure 3.2** Sample autocorrelation plot of Seoul data**Table 3.6** Parameter estimations of the ARIMA(2,1,1)(0,1,1)[7] model in Seoul

Parameter	Estimation	S.E.	p-value
ar1	0.160	0.040	0.0001
ar2	0.047	0.036	0.1694
ma1	-0.853	0.025	0.0000
sma1	-0.975	0.011	0.0000

Table 3.7 Parameter estimations of the ARIMAX(1,1,1)(0,1,1)[7] model in Seoul

Parameter	Estimation	S.E.	p-value
ar1	0.084	0.038	0.0246
ma1	-0.841	0.021	0.0000
sma1	-0.976	0.011	0.0000
gamma1	0.946	-	-
gamma2	0.041	-	-
gamma3	0.017	0.001	0.0000
gamma4	-0.123	0.046	0.0074
gamma5	0.050	0.047	0.2958

3.2. 제주 지역

제주 지역의 충전 전력량 데이터를 이용하여 서울 지역과 동일하게 모형을 적합하였다. 먼저, 제주 지역에 대하여 ARIMA 모형을 적합하였다. 가장 적합한 모형으로는 ARIMA(1,1,0)로 나타났고 Table 3.8에 ARIMA(1,1,0) 모형에 대한 계수가 나타나있다.

Table 3.8 Parameter estimations of the ARIMA(1,1,0) model in Jeju

Parameter	Estimation	S.E.	p-value
ar1	-0.382	0.028	0.0000

이분산성을 고려한 ARIMA-GARCH 모형 적합 결과이다. 가장 적합한 모형으로는 ARIMA(3,1,4)-GARCH(1,1) 모형으로 나타났고 Table 3.9는 ARIMA(3,1,4)-GARCH(1,1) 모형의 계수이다.

Table 3.9 Parameter estimations of the ARIMA(3,1,4)-GARCH(1,1) model in Jeju

계수	Estimation	S.E.	p-value
mu	3.791	0.095	0.0000
ar1	0.392	0.029	0.0000
ar2	0.064	0.034	0.0563
ar3	-0.833	0.027	0.0000
ma1	-0.929	0.039	0.0000
ma2	0.096	0.041	0.0184
ma3	0.875	0.034	0.0000
ma4	-0.467	0.030	0.0000
omega	0.000	0.000	0.1282
alpha1	0.075	0.017	0.0000
beta1	0.923	0.015	0.0000

제주 지역의 ARIMAX 모형 적합 결과, ARIMAX(3,1,3)이 가장 적합한 모형으로 나타났다. Table 3.10은 ARIMAX(3,1,3) 모형의 계수이다. 그 결과, 유의수준 0.05하에서 전기 자동차 등록 대수와 난방지수가 유의한 변수로 나타났다.

다음으로 ARIMAX-GARCH 모형의 적합한 결과, ARIMAX(6,1,3)-GARCH(1,1) 모형이 가장 적합한 것으로 나타났고 그에 대한 계수가 Table 3.11에 나타나있는데 제주 지역의 경우 서울과 달리 전기 자동차 등록 대수만을 고려하였고 p-value를 기준으로 0.05보다 작으므로 유의한 변수이다.

제주 지역에 관하여 seasonal ARIMA (SARIMA) 모형을 적합하기 전에 주기를 알아보기 위하여 차분한 데이터의 sample autocorrelation plot을 그렸다. Figure 3.3을 보면 주기가 7 인 것으로 나타났다. SARIMA 모형을 적합한 결과 ARIMA(0,1,0)(0,1,1)[7] 모형으로 나타났다. Table 3.12에 SARIMA 모형의 계수가 나타나있다.

Table 3.10 Parameter estimations of the ARIMAX(3,1,3) model in Jeju

Parameter	Estimation	S.E.	p-value
ar1	1.393	0.028	0.0000
ar2	-1.179	0.035	0.0000
ar3	0.144	0.028	0.0000
ma1	-2.036	-	-
ma2	1.980	-	-
ma3	-0.780	-	-
gamma1	1.344	0.439	0.0021
gamma2	0.008	0.010	0.4458
gamma3	0.012	0.003	0.0000
gamma4	0.062	0.056	0.2540
gamma5	-0.021	0.057	0.6455

Table 3.11 Parameter estimations of the ARIMAX(6,1,3)-GARCH(1,1) model in Jeju

Parameter	Estimation	S.E.	p-value
mu	-15.447	0.011	0.0000
ar1	1.661	0.001	0.0000
ar2	-1.250	0.001	0.0000
ar3	0.311	0.000	0.0000
ar4	0.095	0.000	0.0000
ar5	0.077	0.000	0.0000
ar6	0.105	0.000	0.0000
ma1	-2.217	0.000	0.0000
ma2	2.151	0.000	0.0000
ma3	-0.938	0.000	0.0000
gamma1	3.338	0.002	0.0000
omega	0.073	0.004	0.0992
alpha1	0.073	0.004	0.0000
beta1	0.926	0.002	0.0000

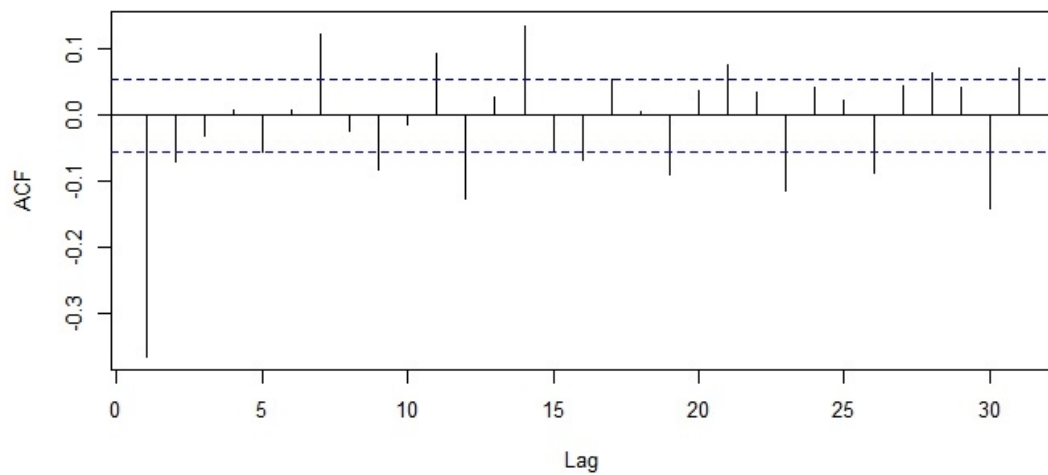
**Figure 3.3** Sample autocorrelation plot of Seoul data

Table 3.12 Parameter estimations of the ARIMA(0,1,0)(0,1,1)[7] model in Jeju

Parameter	Estimation	S.E.	p-value
sma1	-0.957	0.009	0.0000

외생변수와 주기성을 고려한 SARIMAX 모델을 적합한 결과, 가장 적합한 모형은 전기 자동차 등록 대수와 냉방지수, 난방지수를 고려한 ARIMAX(0,1,0)(0,1,1)[7] 모형이다. Table 3.13은 SARIMAX에 관한 계수가 나타나있다. 유의한 변수는 난방지수인 것으로 나타났다.

Table 3.13 Parameter estimations of the ARIMAX(0,1,0)(0,1,1)[7] model in Jeju

Parameter	Estimation	S.E.	p-value
sma1	-0.957	0.009	0.0000
gamma1	1.890	7.511	0.8014
gamma2	0.005	0.028	0.8551
gamma3	0.011	0.005	0.0229

4. 모형의 성능 비교

앞의 절에서 적합한 모형들을 비교하기 위한 오차의 척도로는 *MAPE*와 *SMAPE* 두 가지 척도를 사용하였다. 일반적으로 모형을 평가하는데 *MAPE*가 가장 많이 사용되는데 참고 논문 Majidpour 등 (2015)에 따르면 *MAPE*가 이상치에 영향을 많이 받아 제안된 *SMAPE*를 사용하여 모형을 비교하였다. *MAPE*와 *SMAPE*의 식은 식 (4.1)과 (4.2)와 같다.

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{Y_t - F_t}{F_t} \right| \times 100(\%), \quad (4.1)$$

$$SMAPE = \frac{\sum_{t=1}^n |F_t - Y_t|}{\sum_{t=1}^n (Y_t + F_t)} \times 100(\%). \quad (4.2)$$

식 (4.1)과 (4.2)에서 n 은 예측에 사용한 데이터 수이고, Y_t 는 t 시점에서의 실제값, F_t 는 t 시점에서의 예측값을 의미한다. *MAPE*와 *SMAPE*는 모두 값이 작을 수록 우수한 모형임을 뜻한다.

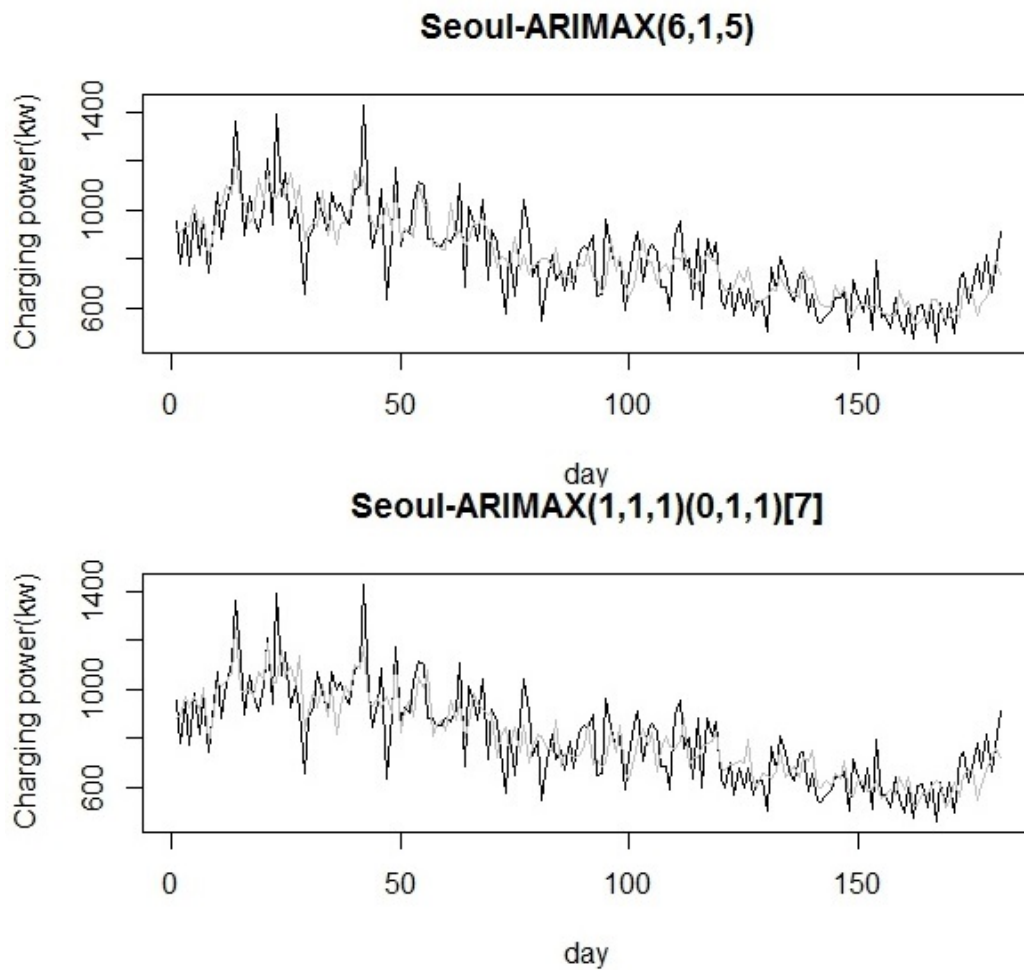
4.1. 서울 지역 비교

서울 지역에 관하여 여러 모형을 가지고 전기 자동차 충전 전력을 예측해보았다. 그 결과는 Table 4.1과 같다. 서울 지역의 경우, *MAPE*가 가장 작은 전기 자동차 등록 수, 냉·난방 지수, 여름 더미 변수, 겨울 더미 변수를 외생변수로 추가한 ARIMAX(6,1,5) 모형이 가장 높은 정확도를 보였고, *SMAPE*를 기준으로 보았을 때에는 계절적 영향을 고려한 ARIMAX(1,1,1)(0,1,1)[7]이 가장 좋은 정확도를 보였다. 외생변수는 전기 자동차 등록 수, 냉·난방 지수, 여름 더미 변수, 겨울 더미 변수를 고려하였다.

예측 결과, 공휴일을 추가하였을 때보다 공휴일을 외생변수로 추가하지 않은 모형들이 더 좋은 예측력을 보였다. 서울 지역의 전기 자동차 충전량을 예측하는데 공휴일은 큰 영향을 끼치지 않아 고려할 필요는 없을 것으로 보인다. 또한 앞 3.1절에서 적합 결과를 보면 전반적으로 여름 변수에 대한 계수값이 음수로 나타났다. 이는 다수의 사람들이 여름에 휴가를 많이 떠나 상대적으로 전기 자동차 충전 전력량이 떨어지는 것으로 추측된다. 서울 지역에서는 전체적으로 ARIMA 모형과 ARIMAX 모형에 비하여 ARIMA-GARCH, ARIMAX-GARCH 모형의 정확도가 더 떨어지는 것으로 나타났다. 그러므로 이분

Table 4.1 Forecast evaluation in Seoul

Model	<i>MAPE</i>	<i>SMAPE</i>
ARIMA(6,1,5)	11.443	11.514
AIRMA(6,1,5)-GARCH(1,1)	11.931	11.982
ARIMAX(6,1,5)	10.926	10.982
AIRMAX(6,1,5)-GARCH(1,1)	11.302	11.285
ARIMA(2,1,1)(0,1,1)[7]	11.625	11.682
ARIMAX(1,1,1)(0,1,1)[7]	11.022	10.969

**Figure 4.1** Best model in Seoul

산성을 고려한 ARIMA-GARCH, ARIMAX-GARCH 모형보다 ARIMA, ARIMAX 모형을 적합시키는 것이 더 효율적이고 적절한 모형을 만들어 낼 수 있는 것으로 보인다. Figure 4.1는 가장 예측력이 좋았던 ARIMAX(6,1,5) 모형, ARIMAX(1,1,1)(0,1,1)[7] 모형에 대한 plot이다. 검정 그래프가 실제

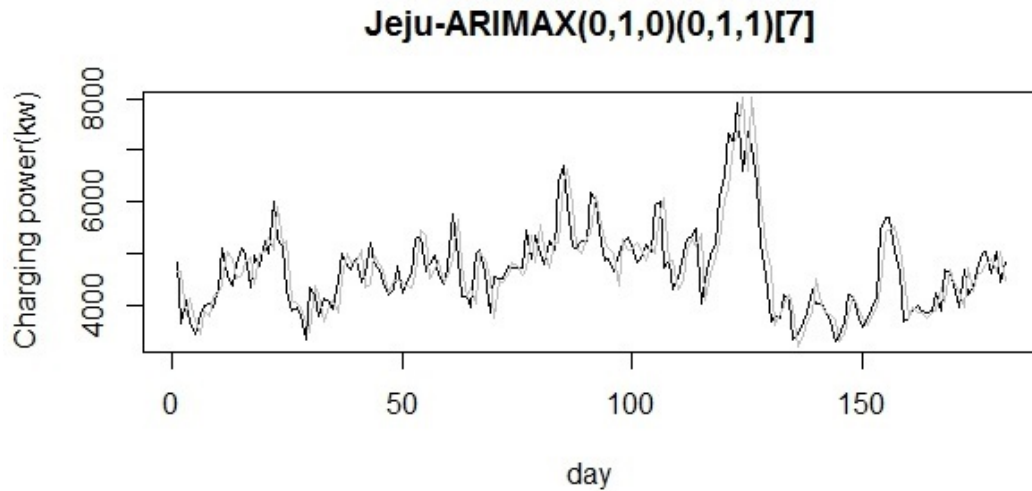


Figure 4.2 Best model in Jeju

값이고, 회색 그래프가 예측값이다.

4.2. 제주 지역 비교

제주 지역의 전기 자동차 충전 전력량 예측 모형에 대한 예측값과 실제값에 대한 비교 결과이다. *MAPE*와 *SMAPE*를 기준으로 모형을 평가했을 때, *MAPE*는 7.806, *SMAPE*는 7.785로 두 지수가 모두 낮은 ARIMAX(0,1,0)(0,1,1)[7] 모형이 가장 좋은 예측력을 보였다. 전기 자동차 등록 대수와 냉·난방 지수를 외생 변수로 고려한 모형이다. 제주 지역 또한, 서울 지역과 마찬가지로 공휴일 변수를 추가하였을 때가 공휴일 변수를 추가하지 않은 모형보다 정확도가 떨어지는 것으로 나타나 공휴일 변수를 추가하지 않기로 하였다. Figure 4.2는 제주 지역 예측 모형 중 가장 좋은 예측력을 보인 모형인

Table 4.2 Forecast evaluation in Jeju

Model	<i>MAPE</i>	<i>SMAPE</i>
ARIMA(1,1,0)	8.417	8.411
ARIMA(3,1,4)+GARCH(1,1)	8.205	8.186
ARIMAX(3,1,3)	8.523	8.513
ARIMAX(6,1,3)+GARCH(1,1)	8.166	8.195
ARIMA(0,1,0)(0,1,1)[7]	7.928	7.937
ARIMAX(0,1,0)(0,1,1)[7]	7.806	7.785

ARIMAX(0,1,0)(0,1,1)[7] 모형에 대한 plot이다. 검정 그래프가 실제값이고, 회색 그래프가 예측값이다. 서울 지역의 예측 모형보다 훨씬 더 좋은 예측력을 보였다. 서울 지역의 예측 모형과의 차이점은 서울에 비하여 ARIMA 모형이나 ARIMAX 모형보다 기존의 가정과 달리 이분산성을 고려한 ARIMA-GARCH 모형, ARIMAX-GARCH 모형이 더 좋은 성능을 보인다는 것이다. 상대적으로 *MAPE*와 *SMAPE* 모두 낮아졌다. 하지만 그 수준은 미미한 것으로 나타나 효율적인 면으로만 평가한다면 시간이 될 소요되는 ARIMA, ARIMAX 모형이 더 적합한 모형으로 볼 수 있다.

5. 결론

최근 지하 자원의 한계, 환경 보호의 중요성이 커지면서 전세계적으로 신재생 에너지에 대한 관심과 같이 전기 자동차에 관한 관심도 높아지고 있다. 그러나 전기 자동차가 증가함에 따라 전기 자동차를 충전하기 위한 전력 수요가 증가할 것이다. 따라서, 전력을 안정적으로 공급하기 위하여 전기 자동차 충전 전력에 관한 예측도 중요해 질 것이다. 이에 대비하여 시계열 모형을 이용한 전기 자동차 충전을 위한 전력 사용량을 예측하는 연구를 실시하였다. 보다 정확한 예측을 하기 위하여 외생 변수로는 전기 자동차 등록 수, 냉·난방 지수, 공휴일 더미 변수 외에도 일반적으로 전력 수요량이 높은 여름과 겨울에 대한 특성을 고려하기 위해 여름 더미 변수, 겨울 더미 변수를 추가로 고려하였다. 특히 모형을 적합한 결과 서울과 제주 지역 대부분 난방 지수가 높을수록 충전 전력량이 높아지는 것을 알 수 있었다. 전기차의 경우, 히터 등의 난방기기 사용시 전력 사용량이 크게 늘어나며, 이에 따라 난방지수에 영향이 크게 발생함을 확인할 수 있었다. 서울 특별시와 제주특별자치도의 전기 자동차 충전 전력량을 지역별로 예측하였는데 시계열 모형을 이용하여 예측한 본 논문에서는 서울 지역을 예측한 경우 *MAPE*, *SMAPE*가 각각 10.926, 10.969이고 제주 지역을 예측한 경우에는 각각 7.806, 7.785로 서울 지역보다 제주 지역의 결과에 더 좋은 예측력을 보였다. 따라서, 지역별로 지역 특성을 고려하여 적절한 모형이 각각 다를 것으로 예상되므로 추후 연구에서도 이를 고려하여 서울, 제주 지역 이외에도 다른 지역들에도 각각 적합한 모형을 적용시켜 보는 것이 가장 좋은 정확도를 보일 수 있을 것이다. 서울 지역의 경우, ARIMA-GARCH, ARIMAX-GARCH 모형보다 ARIMA, ARIMAX 모형이 더 좋은 예측력을 보였으며 *MAPE*가 10.926인 ARIMAX 모형, *SMAPE*가 10.969인 SARIMAX 모형이 가장 좋은 성능을 보였다. 서울과 달리, 제주 지역의 경우에는 ARIMA, ARIMAX 모형 보다 ARIMA-GARCH, ARIMAX-GARCH 모형이 더 좋은 예측력을 보였으며 가장 좋은 성능을 보인 모형은 계절성과 외생변수의 영향을 고려한 SARIMAX 모형이다. 특히, 서울과 다르게 제주는 관광 특구로 지정되어 있어 서울 지역의 데이터보다 변동성이 크므로 서울 지역보다 GARCH 모형이 더 좋은 결과를 얻을 수 있었다.

전기 자동차는 보통 주간 시간대에 충전을 하기보다는 퇴근 후인 야간 시간대에 충전할 확률이 더 높다고 알려져 있다. 그러나 본 논문에서 사용된 데이터는 전기 자동차가 아직 활성화되지 않았던 2014년부터의 데이터로 매 시간마다 충전이력이 나타나 있지 않았다. 그러므로 일별 총 충전 전력 데이터를 이용하여 예측하였기 때문에 주간, 야간 등의 시간대는 고려하지 못하였다. 최근 들어 전기 자동차 활성화 정책이 지역별로 실시됨에 따라 전기 자동차 사용이 활성화 되고 있어 전기 자동차가 활성화되어 충전이력도 활발하게 늘어날 것이라고 예상된다. 따라서 추후에는 충전 시간대를 고려한 연구가 필요하다고 사료된다.

References

- Arias, M. B. and Bae, S. (2016). Electric vehicle charging demand forecasting model based on big data technologies. *Applied Energy*, **183**, 327-339.
- Arias, M. B., Kim, M. and Bae, S. (2017). Prediction of electric vehicle charging-power demand in realistic urban traffic networks. *Applied Energy*, **195**, 738-753.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, **31**, 307-327.
- Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of U. K. Inflation, *Econometrica*, **50**, 987-1008.
- IEA (2016). *Global EV outlook 2016: Beyond one million electric cars*, OECD Publishing, Paris/IEA, Paris, <http://dx.doi.org/10.1787/9789264279469-en>.
- Majidpour, M., Qiu, C., Chu, P., Pota, H. R. and Gadh, R. (2016). Forecasting the EV charging load based on customer profile or station measurement? *Applied Energy*, **163**, 134-141.

- Mo, J. (2012). Estimation of electricity hourly demand and policy implications by introduction of electric vehicle. *KIET*, 2012-294.
- Qian, K., Zhou, C., Allan, M. and Yuan, Y. (2010). Load model for prediction of electric vehicle charging demand, *In Power System Technology (POWERCON)*, 2010 International Conference on IEEE.
- Xydas, E. S., Marmaras, C. E., Cipcigan, L. M., Hassan, A. S. and Jenkins, N. (2013). Forecasting electric vehicle charging demand using support vector machines. *In Power Engineering Conference (UPEC)*, 2013 48th International Universities' IEEE.

A study on electricity demand forecasting for electric vehicles in KOREA[†]

Suim Choi¹ · Heung-gu Sohn² · Sahm Kim³

¹³Department of Applied Statistics, Chung-Ang University

²Department of Aviation, The Korea Transport Institute

Received 11 July 2018, revised 10 September 2018, accepted 13 September 2018

Abstract

Environmental pollution issues such as global warming and particular matter (PM) have recently brought a lot of attention to the government's eco-friendly policies all. In addition to environmentally friendly energy, various policies like subsidy support policies are being implemented to promote the use of electric vehicles. As the use of electric vehicles become active, the demand of electric power for charging electric vehicles will increase. Therefore, it is important to predict accurately the electric demand for charging electric vehicles in order to establish a smooth demand supply plan. In this paper, electric vehicle charging station data from Seoul and Jeju Island were collected and the research was conducted to forecast the amount of charging power for each region. The time series prediction models (ARIMA, ARIMAX, ARIMA-GARCH, ARIMAX-GARCH, SARIMA, SARIMAX) for electric vehicles charging power amount prediction were compared on the basis of MAPE, SMAPE.

Keywords: Charging power, EV, time series model.

[†] This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2016R1D1A1B01014954).

¹ Graduate student, Department of Applied Statistics, Seoul, Korea.

² Postdoctoral researcher, Department of Aviation, The Korea Transport Institute, Sejong, Korea.

³ Professor, Department of Applied Statistics, Seoul, Korea. E-mail: sahm@cau.ac.kr