

DNN과 LSTM 활용한 일일 전력수요모델 개발 및 예측

김영수* · 박호정**†

*고려대학교 에너지환경대학원 박사과정, **고려대학교 식품자원경제학과 교수

Modeling and Predicting South Korea's Daily Electric Demand Using DNN and LSTM

Kim, Youngsoo* and Park, Hojeong**†

*Ph.D. candidate, Graduate School of Energy and Environment, Korea University, Seoul, Korea

**Professor, Department of Food and Resource Economics, Korea University, Seoul, Korea

ABSTRACT

Demand for electricity is influenced by factors, such as economic structure, industrial sector, and environmental volatility. This paper aims to capture characteristics pertaining to electricity demand by developing a daily forecast model. Key variables were selected from socio-economic and environmental perspectives. Installed capacity alongside socio-economic metrics including the consumer composite sentiment index (CCSI), trade balance, unemployment rate, and day of the week were added. Environmental variables for the model were average daily temperature and COVID-19 case count.

The Deep Neural Network (DNN) model, an Artificial Neural Network (ANN) model, was used to compensate for lack of definitive linear relationships between variables for electricity demand. The trained model was performed with rRMSE of 3.74% and MAPE of 2.67%.

Further scenario analysis helped to shed light on the utility of this model. The scenario aims to explain how energy demand is affected by supply-centric electricity policies and demand management policies, as well as macro-economic expansion and contraction. A recurrent network model, LSTM, was used to forecast average daily temperatures and COVID-19 cases before the DNN electricity demand forecast model interpreted the forecast results.

Key words: Electricity Demand, Covid-19, ANN, Deep Neural Network, LSTM, Scenario Analysis

1. 서론

1.1 전력수요 예측의 중요성

현대사회에서 전기는 산업 및 일상의 필수재이다. 이런 전기의 원활한 공급을 위해선 적절한 규모의 전력설비 확보가 선행되어야 한다. 전력수요에 비해 공급이 초과될 경우 예산낭비를 초래하고 수요에 비해 증설이 더딜 경우 전력공급에 부정적인 영향을 끼친다. 따라서 정확한 전력수요 예측은 에너지계획, 전력수립, 에너지정책의 결정의 주요 판단자료로 그 중요성이 매우 크다(García-Ascanio and

Maté, 2010; Saravanan et al., 2012). 한국에선 특히 다음의 상황을 고려할 때 정확한 장단기 전력수요예측이 필요하다(Jang and Jo, 2018).

- 한국은 저탄소에너지 사회로의 전환을 위해 신재생에너지 확대 정책을 적극 추진 중이다. 이런 기조에 따라 제9차 전력수급계획에서는 재생에너지를 2020년의 20.1 GW에서 2034년까지 77.8 GW 규모로 확대할 계획이다. 2034년 기준으로 태양광은 45.6 GW, 풍력은 24.9 GW로서 전체 신재생의 91%를 차지할 전망이다.
- 재생에너지는 기상조건에 따라 발전량이 불규칙한 특성을 가지므로 대규모 신재생 발전설비가 전력계통에

†Corresponding author : hjeongpark@korea.ac.kr (02841, Room 208, Life Sciences Bld. (East), 145, Anam-ro, Seongbuk-gu, Seoul, Republic of Korea, Tel. +82-2-3290-3039)

ORCID 김영수 0000-0001-5784-4771

박호정 0000-0003-2747-1905

투입될 경우 전력수급 균형을 유지하기 어려워진다.

- 즉, 이런 급변하는 수급변동에 적절한 발전계획을 수립하기 위해서는 정확한 장단기 전력수요 예측이 특히 중요하다.

1.2 전력수요에 영향을 미치는 변수

전력수요는 경제지표 및 산업구조, 급격한 환경변화 등 다양한 변수에 영향을 받는다. 따라서 정확한 전력수요예측을 위해서는 해당 국가의 전력수요에 영향을 미치는 변수를 올바르게 선택해야 한다. 이를 위해 본 연구에서는 기존문헌 연구를 통해 보편적으로 전력수요에 영향을 미치는 변수와 국내 전력수요의 특수성을 반영한 변수, 그리고 경제활동의 가장 중요한 변수로 등장한 코로나19 세가지로 구분한다.

1.3 일반 변수

기존문헌을 통해 전력수요에 영향을 미치는 일반적 요인을 살펴본다.

1.3.1 기온

여러 국가를 대상으로 진행된 실증연구를 통해 기온이 전력수요에 영향을 미침을 확인하였다. 김혜민 외(2015)는 코익(Koyck)모형을 활용하여 전력의 소비자물가지수, 분기별 국내총생산, 기온을 변수로 전력수요를 추정하였으며 기온이 포함된 모형의 설명력이 더 높음을 확인하였다(Kim et al., 2015). Lee and Chiu (2011)은 panel smooth transition regression (PSTR)을 활용하여 OECD 24개국의 기온과 전력수요 관계를 실증연구하였으며 U자형 비선형 관계가 존재함을 확인하였다.

Fung et al. (2005)는 기온과 에너지수요(전력, 가스, 석유)와의 관계를 회귀분석을 통해 확인하였으며 1°C 증가할 때 가정, 상업, 산업에서의 전력소비가 각각 9.2%, 3.0%, 그리고 2.4% 임을 확인하였다(Fung et al., 2006). Bessec and Fouquan (2008)은 유럽의 15개 국가를 대상으로 Panel Smooth Transition Regression (PSTR)을 이용하여 전력수요와 기온의 비선형관계를 연구하였으며, 두 변수간의 U자형 관계가 있음을 확인하였다(Bessec and Fouquan, 2008).

선행연구로부터 기온은 전력수요에 영향을 끼치는 주요 변수이며 전력수요와 기온은 U자형의 비선형 관계를 이룸을 확인할 수 있다(Kim et al., 2015).

1.3.2 휴일

공휴일이나 명절연휴, 주말에는 나들이 인파 및 생산활동 저감으로 전력수요가 줄어든다. 백종관 외(2015)은 회귀모형을 활용하여 전력수요를 예측하였으며 휴일을 고려하는 경우 고려하지 않는 경우에 비해 오차율이 감소하는 효과를 확인하였다(Baek and Ha, 2015). 또한, Kim(2016)은 정준 공적분 회귀분석을 활용하여 산업별로 공휴일이 전력수요에 미치는 영향을 실증하였으며 전자 및 전자기 산업과 금속제품산업에서는 공휴일 효과가 매우 큰 것에 비해 숙박 및 음식점과 부동산 및 임대업에서는 공휴일 효과가 거의 나타나지 않음을 확인하였다(Kim, 2016).

1.3.3 전력요금

가격변화에 따른 수요변화는 가격탄력성으로 측정할 수 있다. 가격 탄력성이란, 수요에 영향을 미치는 다른 변수는 일정하다는 가정하에 가격 변화분의 수요변화로 정의된다. “전력수요가 전력요금에 비탄력적이다”라는 가설은 국외 다수의 연구논문 통해 실증되었다.

Shu Fan et al. (2011)은 준모수가법모형 (semi-parametric additive model)을 활용하여 오스트레일리아의 전력수요의 가격탄력성을 추정하였으며 가격탄력성은 -0.363부터 -0.428이다. 또한 이 논문에서 정리한 다른 연구결과를 살펴보면 단기 탄력성은 -0.2에서 -0.4, 장기적으로는 -0.5에서 -0.7의 값을 갖고 있음을 확인하였다. 단기에 비해 장기 탄력성이 강하게 나타나는 이유는 단기에 비해 장기적인 경우 대체품을 찾을 가능성이 높기 때문인 것으로 예상하였다(Fan and Hyndman, 2011).

Korea Energy Economics Institute(1996)는 1980년부터 1994년까지 자료를 활용하여 가정용 전력수요의 장단기 가격탄력성을 각각 -0.160, -0.380으로 추정하였다(Yu, 1996). 또한, 2016년 연구에서도 1985년부터 2011년까지 산업용 전력수요에 대한 단기 가격탄력성을 -0.445로 추정하였다.

단, Lee (2014)는 국내 데이터를 활용하여 전력사용량이 상대적으로 높은 산업군인 화학물 및 화학제품 산업을 통해 전력수요의 가격탄력성 -1.14임을 추정하였다(Lee, 2014).

1.3.4 GDP(국내 총생산)와 실업률

경제성장과 전력수요 증가와는 높은 양의 상관관계가 있음을 다수의 연구를 통해 확인할 수 있다. Ahn et al.

(2016)은 1991년부터 2014년까지 전력명목가격, 실질 GDP를 통해 전력수요량을 추정하는 수요함수에 대한 연구를 수행하였으며 이를 통해 각 변수가 통계적으로 유의하며 장단기 소득탄력성 각각 0.866, 1.287임을 확인하였다. S. Saravanan et al. (2012)는 ANN을 통해 전력수요를 예측하였는데 입력값으로 일인당 GDP를 사용하였다. 인도의 1981년부터 2010년까지의 일인당 GDP와 전력수요 간의 결정계수는 0.966으로 높게 나타났다(S. Saravanan, 2012).

또한, 해외 다수의 연구로부터 전력수요와 GDP 간의 상관관계가 있음을 실증분석하였다(Jamil and Ahmad, 2010). Narayan et al. (2005)는 Cointegration Granger causality 방법을 활용하여 호주의 1966~1999년 데이터를 실증분석한 결과 전력소비와 GDP는 전력소비에 영향을 미침을 확인하였다.

1.4 국내 변수

한국 전력소비는 지속적으로 증가하였다. 대부분의 OECD 국가들은 전력수요 증가율이 경제성장률보다 낮지만 한국은 2001년까지 전력수요 증가율이 경제성장률보다 높았다. 특히, 2000년대 중후반 이후 고유가 시기에 대부분의 OECD 국가들은 전력소비가 감소하거나 횡보한 반면 한국은 이 시기에 전력소비가 급증하는 양상을 보였다.

최근 전력소비 증가율이 둔화되었으나 한국의 1인당 전력소비는 다른 OECD 국가들에 비해 여전히 높게 나타난다. 한국과 유사하게 제조업의 비중이 높고 4계절 기후가 있는 미국, 독일, 일본 3개국의 1인당 전력소비와 OECD 국가 평균 1인당 전력소비를 비교하면 한국이 월등히 높고 여전히 증가함을 확인할 수 있다.

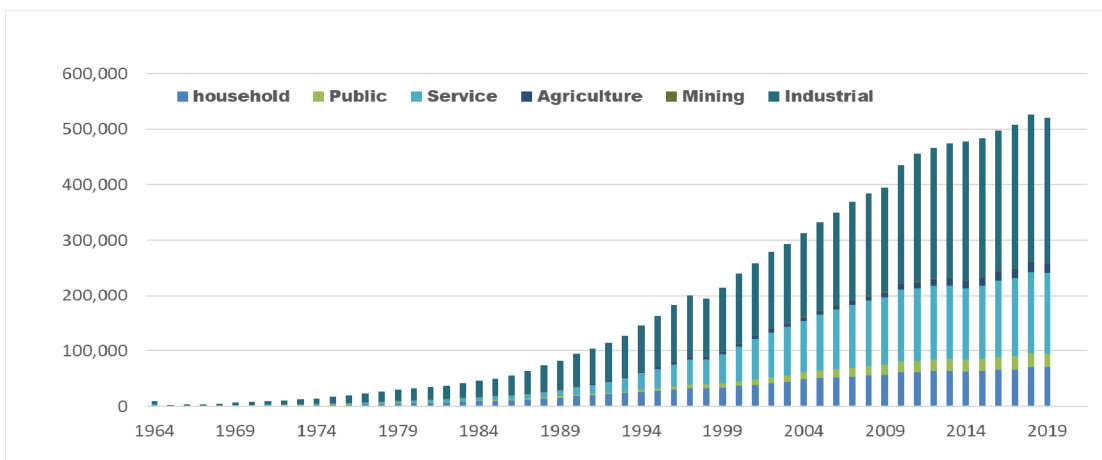


Fig. 1. Electricity consumption by sector (IEA, 2019)

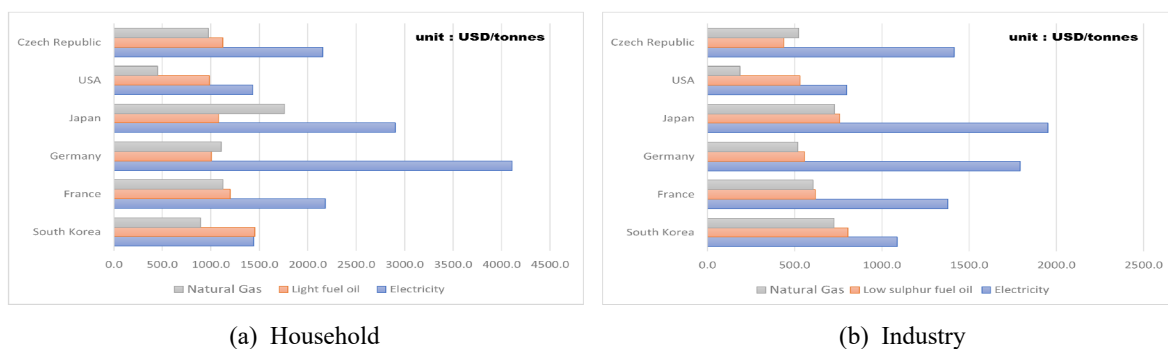


Fig. 2. Comparison by energy source (IEA, 2019)

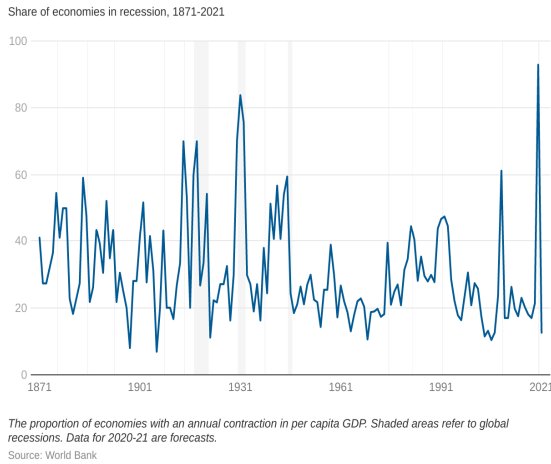


Fig. 3. The proportion of economic with and annual contraction in per capita GDP (World Bank, 2020)

Oil price hits 18-year low

Brent crude, US dollars per barrel

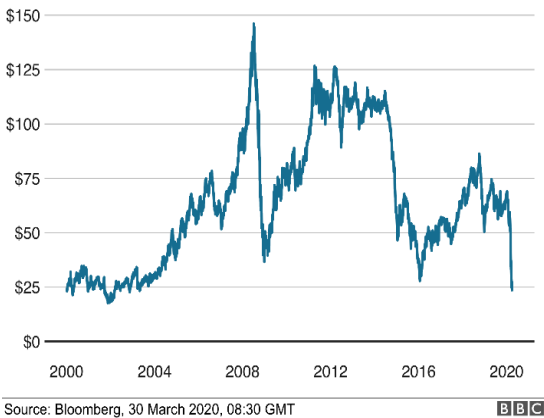


Fig. 4. Oil price hits 18-year low (Bloomberg, 2020)

이런 현상은 한국의 산업구조와 에너지 정책에서 그 원인을 찾아볼 수 있다.

전체 전력수요에서 산업용이 약 50%를 차지한다. 한국과 유사한 제조업 중심의 산업구조를 가진 일본 및 독일 등과 비교해 봐도 월등히 높은 수치를 보여준다.

한국과 산업구조가 유사한 독일, 일본, 프랑스 등과 비교했을 시에도 한국의 전력가격은 저렴하였다. 전력은 2차 에너지원인 바, 모든 에너지 중에 가장 비싼 것이 일반적

이나, 한국은 저렴한 전력요금을 유지하기 위해 원자력발전과 석탄발전에 세금 우대와 같은 각종 지원과 보조금 정책을 시행하였고, 정부가 적극적으로 요금을 통제하였다. 그 결과 전력요금은 다른 나라 수준보다 낮을 뿐만 아니라 타 에너지원에 비해서도 상대적으로 낮게 유지되고 있다.

따라서 국내 변수는 다음과 같이 설정한다.

국내 산업구조는 수출 중심인 바, 산업구조와 전력소비 효율이 현재와 동일하다는 가정 시 수출증가는 전력수요 증가를 불러온다. 수출증가를 대표하는 지표로 월간 수출 지수를 활용한다.

단, 최근 들어 전력기후환경비용 반영 및 K-RE100으로 녹색프리미엄 제도 등으로 산업용 전력요금이 빠르게 증가하고 있다.

또한, 전력요금을 변수 중 하나로 선정한다. 기존 문헌에선 전력수요는 전력요금에 비탄력적이라고 실증되었으나 국내는 낮은 전기요금으로 인한 비효율적 전력기기 사용과 타 국가와의 유사한 전력요금 수준으로 변화시 상대가격 변화가 크게 나타나 개인의 선호에 영향을 미칠 수 있다.

또한 소비자심리지수를 변수로 추가한다. 만약 개개인이 향후 경제가 좋아질 것이라 전망한다면 개인은 소비 등 경제 활동에 적극적으로 참여할 것이다. 즉, 개인의 향후 경기에 대한 전망은 민간 경제를 예상하는 선행지표로 활용가능하다.

1.5 코로나19 영향

1.5.1 경제 악화

2019년 12월 31일 중국 우한에서 중국 WHO 사무국에서 불명확한 원인에 의한 폐렴이 보고된 이후(WHO Novel Coronavirus (2019-nCoV) Situation report-1, 2020.1.21), 전 세계로 급속히 확대되어 2020년 1월 30일에 코로나19에 대해 국제공중보건비상사태(PHEIC)를 선포하고(WHO Director-General’s statement on IHR Emergency Committee on Novel Coronavirus) 3월 11일에는 “팬데믹”을 선언하였다(WHO Director-General’s opening remarks at the media briefing on COVID-19 March 11).

세계는 국가 및 지역간 이동을 통제하고 대면 경제활동을 제약하는 등 코로나의 급격한 확산을 막기위해 안간힘을 쓰고 있다. 이런 일련의 조치로 경제활동은 거의 중단되었으며 경제는 심각하게 악화되고 있다.

무역 감소와 관광업의 중단, 취약계층과 기업 지원을 위한 재정부채에 따른 정부부채 증가 등은 모든 국가가 당면한 문제이며 개인 또한 체감하고 있다. World

Bank (2020)는 1870년 이후 전 세계적으로 가장 많은 국가에서 1인당 소득이 감소할 것으로 전망하고 있다 (World Bank, The Global Economic Outlook During the COVID-19 Pandemic: A Changed World).

1.5.2 전력수요의 영향

이동통제 및 경제활동의 둔화는 전력수요에 영향을 미쳤다. 인도 및 프랑스, 스페인 영국 등 각 국가의 데이터를 보면 이동통제 정책으로 경제활동이 제한됨에 따라 전력수요는 주말 수준으로 떨어졌다. 4월 이탈리아와 독일의 통제정책 완화 이후 전력수요가 회복되는 모습을 보였다. 단, 인도는 8월부터 전력수요가 2019년보다 높은 수준으로 확인되었다(IEA, 2020).

코로나19가 가장 먼저 확인되어 이동통제 정책을 처음으로 적용한 중국은 초기 이동제한 등 강력한 정책 및 코로나19 추적을 통해 확산을 효과적으로 통제하고 있다. 첫 이동금지 정책은 2020년 1월 23일 우한에서 시행되어 전국으로 확대되었고 2020년 2월 7일 더욱 강력한 락다운(lockdown) 조치를 시행하여 코로나 19 확산을 막았다. 이러한 조치에 힘입어 3월 이후부터 확진자수가 적절히 통제되고 있다. 락다운 조치가 시행되었던 2020년 1월과 2월에는 전력수요가 전년대비 약 13%가 하락하였으나, 통제가 완화된 이후 차츰 회복하고 전년대비 전력수요가 확대되는 모습을 보이고 있다(IEA, 2020; Silvia Weko, 2020). 이에 따라 본 연구에서는 코로나19의 확산자수를 전력변수에 영향을 미치는 변수로 선정한다.

2. 문헌연구

전력수요 연구는 주로 시계열모형, 회귀모형 등 전통적 예측 모형이 활용되었다. 단일시계열 데이터를 활용한 통계모형과(ABDEL-AAL, 1997; Samer Saab, 2001) 전력수요와 관련된 다변수를 활용한 다변량선형회귀모형이 주로 활용되었다(Papalexopoulos, 1990). 다변량선형회귀모형은 여러 변수를 예측에 활용할 수 있다는 장점은 있으나 비선형적인 전력데이터에 대한 예측 정확도가 낮다는 문제점이 있다. 또한, ARIMA나 AR모형, Holt-Winter's seasonal 모형 등 단일변량의 시계열 패턴만을 고려한다는 점에서 다양한 변수가 상호작용하여 영향을 미치는 전력수요에 활용하기에는 근본적인 문제가 있다.

Makridakis (1982)의 연구에서는 이미 선형예측 방법의 중 어느 경우에도 우수한 모형이라고 할 수 있는 방법

은 없으며 각 데이터의 계형에 적합한 모형만이 있을 뿐임을 확인하였다. 이런 기존 방법론의 한계는 데이터의 비선형정도를 선형 통계 방법으로는 적절히 처리할수 없다는 데에 기인한다(Zhang, 2001).

인공지능신경망(이하 ANN)은 일반화된 비선형 예측이라는 점에서 전력수요예측 연구에서 그 활용빈도가 증가하고 있다.

Pao (2006)은 선형 및 비선형회귀모형과 ANN을 활용하여 대만의 전력수요를 예측하였다(Pao, 2006). Saravanan (2012)은 GDP와 인구수를 입력값으로 하는 ANN을 통해 인도의 연도별 전력수요를 예측하였다(Saravanan et al., 2012). Table 1은 기존의 전력수요예측에 대한 연구목록이다.

기존 문헌연구를 통해 확인한 결과 전력수요는 각 요인별 선형관계로만 나타내긴 어렵다고 판단되므로 비선형 모형인 인공지능신경망 모형을 통해 본 전력수요 예측을 위한 주요 모형으로 선정한다.

3. 연구방법

본 연구는 한국의 일별 전력수요를 예측하는 인공신경망 모형 개발하고 이를 통한 시나리오별 전력수요 예측을 목적으로 한다.

기존 연구, 최근 상황 및 국내 전력수요 특성을 반영하여 7개의 변수를 선정하였으며 이를 통해 일일 전력수요를 예측한다.

일일 전력수요 예측은 위의 DNN (Deep Neural Network) 모형을 사용한다. 일일평균기온 및 코로나19 확진자는 Recurrent Deep learning Model인 LSTM을 사용하여 단기 시계열을 예측한다.

각 상황별 시나리오를 설정하고 각 시나리오별 전력수요를 예측한다.

3.1 DNN 설명

인공신경망은 사람의 뉴런세포에서 비롯된 통계적 학습 알고리즘으로, 입력층, 은닉층, 그리고 출력층으로 구분된다. 입력층과 은닉층, 은닉층과 은닉층 사이를 연결하는 가중치에 따라 결과값이 실제값에 가까워질 수도 결과값이 수렴하지 않을 수도 있다. 인공신경망에는 실제값과 예측값의 오차를 줄이기 위한 가중치 조절을 위해 대표적으로 2개의 알고리즘을 사용한다.

3.1.1 순방향 신경망(Feed forward neural network) 알고리즘

순방향 신경망은 인공신경망의 기본적인 구조를 가진 모델로, 하나 이상의 은닉층을 가지고 있으며 활성화함수를

활용하여 비선형 특징을 추출한다.

인공신경망은 뉴런과 유사한 퍼셉트론(perceptron)을 기본으로 한다. 퍼셉트론은 다수의 신호를 받아 하나의 신호를 출력한다. 이 출력값은 활성화 함수(activation function)을 통해 비선형함수로 변형된다. 활성화함수는

Table 1. Summary of related researches

| Authors | Method used | Independent variables | Forecast energy type | Forecast countries |
|---|---|---|--|-----------------------------------|
| Amarawickrama and Hunt (Amarawickrama and Hunt, 2008) | Six econometric methods, Dynamic Engle and Granger (Dynamic EG), fully modified ordinary least squares (FMOLS), Pesaran, Shin and Smith (PSS), Johansen method (Johansen), structured time series method (STSM) | Per capita gross of domestic production, electricity price and underlying energy demand trend | Annual electricity demand and annual peak electricity demand | Sri Lanka |
| Saab et al. (Samer Saab, 2001) | Autoregressive models | Monthly electricity consumption data | Monthly electric energy consumption | Lebanon |
| Salman Al-Shobaki et al. (Al-Shobaki and Mohsen, 2008). | Regression models | Generated and consumed energy data | Monthly electricity demand | Jordan |
| Ghiassi et al. (Bedi and Toshniwal, 2019) | SVM, DNN, RNN, LSTM | season, day and interval electricity consumption data | season, day and interval electricity consumption | Union Territory Chandigarh, India |
| Pao (Pao, 2006) | ANN, linear and non-linear statistical models | National income, GDP, consumer price index | Electricity consumption | Taiwan |
| Bianco et al. (Bianco et al., 2013) | Regression models | GDP, population, import, export | consumption Energy demand | Italy |
| Geem and Roper (Geem and Roper, 2009) | ANN | GDP, population, import, export | Energy demand | South Korea |
| Mohamed and Bodger (Mohamed and Bodger, 2005) | Multiple linear regression analysis | GDP, average price of electricity, population | Electricity consumption | New Zealand |
| Ekonomou | ANN | Ambient temperature, installed power capacity, yearly per resident electricity consumption, GDP | Energy consumption | Greek |
| Abdel-Aal and Al-Garni (R. E. ABDEL-AAL, 1997) | ARIMA Models and univariate Box-Jenkins time series analysis | | Monthly electric energy | Easter Saudi Arabia |
| Zhou et al. | Trigonometric grey prediction | Past consumption values, degree days, total imports | Electric energy consumption Electricity | Romania |
| S. Saravanan et al. (Saravanan et al., 2012) | ANN | GDP, population | Electricity load forecasting | India |
| Kadir Kavaklioglu et al. (Kavaklioglu et al., 2009) | ANN | GDP, population, import and export | Electricity consumption | Turkey |

출력값을 활성화할지 여부를 결정하는 역할을 한다.

다수의 퍼셉트론과 여러 개의 은닉층을 가진 경우 다층 퍼셉트론이라 하며 이를 이용한 인공지능신경망 모델을 DNN이라 한다.

$Z^{[i]} = W^{[i]T} A^{[i-1]} + b^{[i]}$ 는 각 은닉층 퍼셉트론의 입력데이터가 되고 활성화함수인 $\psi^{[i]}(Z^{[i]})$ 를 통해 다음 은닉층의 입력데이터로 사용된다.

$W^{[i]}$, $b^{[i]}$ 는 각 레이어의 가중치 행렬이며, 활성화함수 $\psi^{[i]}$ 는 Relu 함수를 사용한다.

3.1.2 역전파(Backforward) 알고리즘

딥러닝의 학습은 실제치와 예측치의 차이 즉, 오차를 낮추는 방향으로 진행된다. 이는 비용함수가 최저가 되는 가중치를 구하는 것과 동일하다. 비용함수를

$$L = \sum \frac{1}{2} (\text{예측값} - \text{실제값})^2$$

으로 정의하면 이는 오목함수의 형태를 띠므로 기울기가 0에 가까운 경우 L은 최소값을 가지게 된다. 즉, 학습을 반복할 때 마다 기울기가 낮은 방향으로 연속적으로 이동시켜 최소가 되는 점을 찾는다. 이 최적화 방법을 경사하

강법(Gradient descent method)이라 한다. 이를 기본개념으로 손실함수의 편도함수를 활용하여 각 변수의 변화율을 적용하면 각각 다음의 수식을 얻을 수 있다.

$$W^{[i]} = W^{[i]} - \alpha \frac{\partial L}{\partial W^{[i]}} \tag{1}$$

$$b^{[i]} = b^{[i]} - \alpha \frac{\partial L}{\partial b^{[i]}} \tag{2}$$

여기서 α 는 학습률을 나타내며 적절한 학습률 설정을 통해 위의 식을 활용하여 비용함수가 최소가 되는 최적의 변수를 얻을 수 있다. 본 연구에서는 RMSProp Optimizer을 활용하며 하이퍼파라미터는 0.001로 한다.

3.1.3 전력수요예측 DNN 모델

전력수요는 위의 모델을 활용하여 문헌 조사를 통해 도출한 10개의 변수를 인풋값으로하는 Fig-5 DNN모델을 구성한다.

모형은 python tensorflow로 작성한다.

은닉층은 총 3개로 하며 각 은닉층별 퍼셉트론 개수는 각각 30개, 20개, 20개로 구성한다. 본 모형의 목적은 추정인 바 output 함수는 선형으로 설정한다.

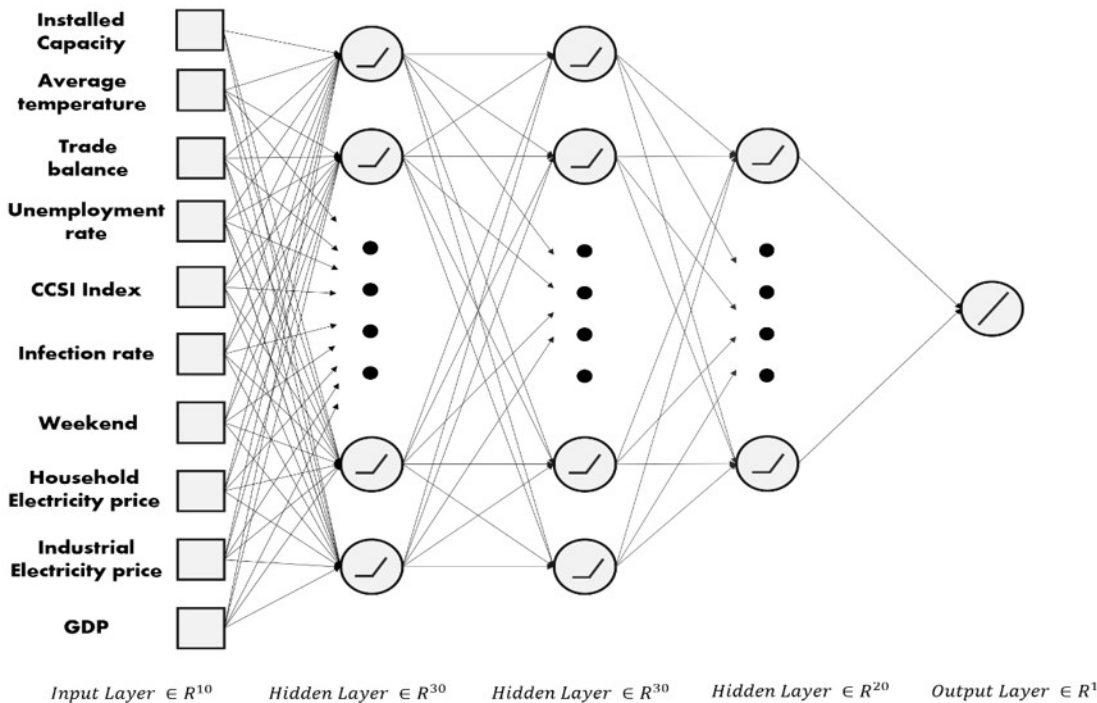


Fig. 5. The scheme of DNN model

3.2 순환신경망

순환 신경망은 입력층에서 출력층으로 한 방향으로만 신호가 흐르는 일반적인 인공신경망(artificial neural network, ANN)과 달리 히든 노드가 방향을 가진 엣지로 연결된 순환구조(directed cycle)를 이뤄 앞선 시간의 정보가 뒤에 있는 정보에 영향 및 효과를 반영할 수 있도록 설계되어 있다.

하지만 일반적인 순환신경망의 경우 이전 정보와 그 정보를 사용하는 지점 사이의 거리가 멀 경우 역전파시 그 그래디언트(gradient)가 점차 줄어들어 모형의 학습력이 감소하는 기울기 사라짐 현상이 발생한다. 이를 해결하기 위해 LSTM이 개발되었다. 히든 레이어에 활성화 함수 하나만 존재한 RNN 에 비하여 Forget, input, output 3개의 노드를 추가하여 정보의 전달을 통제함으로써 학습이 진행되는 도중 그래디언트가 사라지는 것을 해결하였다.

본 연구에서는 Tensorflow를 활용하여 본 모델을 구성하며 시계열 데이터인 일별온도변화와 코로나19 감염자수를 예측하는데 사용한다.

모형은 200개의 메모리 셀로 구성된 LSTM layer와 1개의 Dense로 구성된 예측값 layer로 구성된다.

4. 데이터 특징 및 전처리

4.1 데이터 특징

본 모형에 사용한 자료는 2008년 07월 1일부터 2020년 09월 30일까지 총 4,514개이다. 최대전력수요는 종속변수로 하며 7개의 변수는 모두 입력값으로 설정한다.

본 모형에서의 데이터 정제 시 데이터는 다음의 특징과 한계를 가진다.

코로나19 확진자수는 WHO 자료를 사용하며 확진자수는 2020년 1월 22일부터 집계되어 총 253개의 관측값을 활용하며 집계되지 않은 기간은 0으로 한다.

일별평균온도는 전국평균온도를 사용한다.

소비자심리지수와 무역수지, 실업률은 월별 데이터만 있으므로 해당 달에는 일별로 모두 동일한 값을 가지는 것으로 가정하여 데이터를 정제한다.

휴일은 더미변수로 처리하며 2013년 이후 한글날은 휴일로 지정된 바, 2013년 이후부터 반영하며 국회의원 및 대통령 선거일도 휴일로 한다.

4.2 데이터 전처리

본 모형의 데이터는 모두 전처리 과정을 하여 스케일링을 조절한다.

DNN 모형의 경우 총 4,514개 데이터 중 약 90%인 4,192개는 Train set으로 설정한다.

Table 2. Summary of variables

| No. | Variables | Unit | Source | Variable explanation |
|-----|---------------------------|----------------|-------------------------------------|--|
| X1 | Installed capacity | 10,000 kW | KPX | Installed electricity generating capacity |
| X2 | Epidemic Status | person | WHO | the number of daily infected people |
| X3 | Average temperature) | Celsius degree | Korea Meteorological Administration | The daily average tempertarue |
| X4 | CCSI index | | Korea Bank | - Psychological indicators that consumers feel about economic status - As the value decreases based on 100, consumers judge the perceived economy worse |
| X5 | Trade balance | USD 1,000 | Korea Customs Service | Montly substract between Export and import |
| X6 | Weekend and Holiday dummy | dummy | | |
| X7 | Monthly employment rate | % | Korea Bank | Monthly unemployment rate |
| Y | Daily electricity load | 10,000 kW | KPX | Daily electricity load |

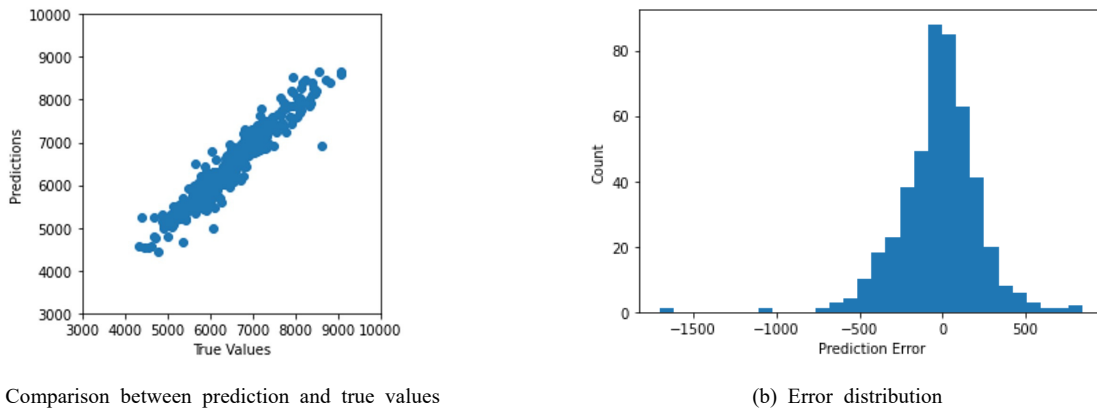


Fig. 6. Test prediction and Prediction error

시계열예측을 위한 LSTM 모델 중 연평균기온모델은 200일 간의 날씨가 영향을 미친다고 가정하고 코로나19 관련 5일간의 감염자수가 다음 감염자수에 영향을 미친다고 가정한다. Train set은 전체 중 30%로 연평균기온모델은 3,159개, Covid 모델은 177개로 한다.

5. 결과 분석

5.1 DNN 결과

DNN 모형은 그래프에서 보듯이 예측값과 실제값이 높은 상관관계를 가지며, 오차 또한 0을 중심으로 정규분포의 형태를 띠고 있다. 본 모형의 예측정확도를 수치로 표현하기 위해 4개의 평가지표(RMSE, rRMSE, MAPE, R^2)를 이용하였다.

추정된 값과 실제값 사이의 차이인 RMSE은 240.89 MW이며 이를 실제평균값으로 나눈 rRMSE값은 3.74%로 추정되었다. rRMSE는 10% 미만일 경우 매우 우수한 모델로 평가하는 바, 본 모형은 우수한 예측성능을 가진다고 평가된다.

Table 3. The result of DNN model

| Evaluation metric | value |
|-------------------|--------|
| RMSE | 240.89 |
| rRMSE | 3.74% |
| MAPE | 2.67% |
| R^2 | 0.922 |

MAPE는 단위와 무관하게 모형의 예측 성능을 비교할 수 있어 모형간 성능비교 시 주로 이용된다. 본 모형의 MAPE는 2.67%이며 결정계수 R^2 값은 0.922를 얻었다.

5.2. LSTM 결과

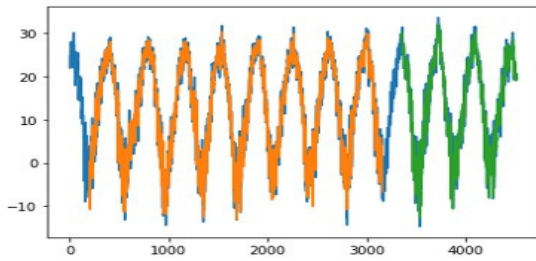
LSTM은 날씨와 코로나19모형을 중심으로 분석하였다. 날씨는 주기성을 가지는 바, Table 4에서 보듯 RMSE와 rRMSE 모두 준수한 성능을 가짐을 확인하였다. 오차항 분포 또한 정규분포의 형태를 띤다. MAPE의 경우 수식의 특성상 분모가 0에 가까울수록 무한대 값이 나오는 바, 본 모형에서는 MAPE는 예측정확도 평가를 위한 지표에서 제외한다.

Table 4. The result of daily temperature model

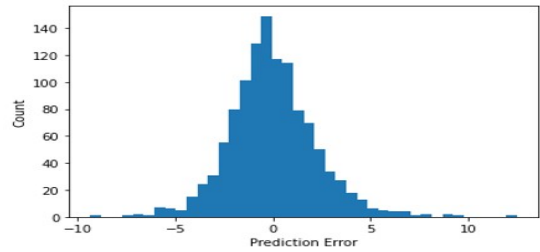
| Evaluation metric | value |
|-------------------|-------|
| RMSE | 2.23 |
| rRMSE | 16.2% |

코로나19 모형은 상대적으로 낮은 예측 정확도를 보여 준다. RMSE는 48.6이며 rRMSE는 31.8%로 보통 수준의 예측역량을 보여준다. 코로나19의 경우 확진자수는 과거의 데이터보다 봉쇄정책, 백신 등 정책적 요인에 영향을 더 많이 받는 바, LSTM 모델이 위의 평균온도 모형 대비 상대적으로 낮은 예측 성능을 보이는 이유는 일면 타당하다.

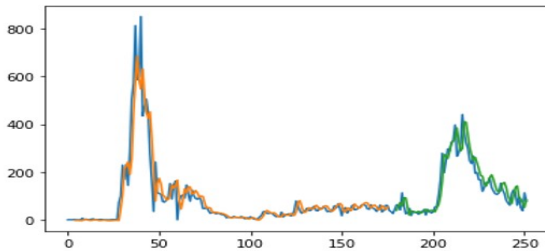
단, 본 논문은 새로운 방법의 활용이라는 측면이 강조되는 바, 성과 지표의 미흡에도 불구하고 본 연구결과를 예측을 위한 모형으로 활용한다.



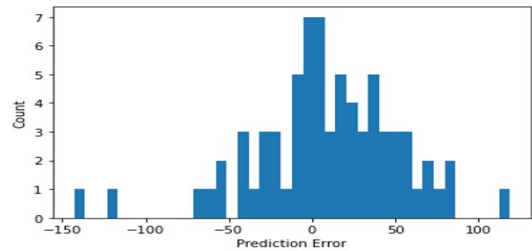
(a) Comparison between prediction and true values – average temperature



(b) Error distribution - average temperature



(c) Comparison between prediction and true values – The number of people infected with Covid



(d) Error distribution - The number of people infected with Covid

Fig. 7. LSTM model –test prediction and prediction error

Table 5. The explanation of model evaluation metrics

| Evaluation metric | Definition | Explanation | Evaluation standard |
|-------------------|--|---|---|
| RMSE | $\sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - x_i)^2}$ | Affected by the absolute value of the predicted value | - The closer to 0, the better the prediction |
| rRMSE (%) | $\frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - x_i)^2}}{x_i} \times 100$ | RMSE divided by actual value | - excellent: rRMSE < 10% - good: 10% < rRMSE < 20% - fair: 20% < rRMSE < 30% - poor: rRMSE > 30% |
| R ² | $1 - \frac{\sum (y_i - x_i)^2}{\sum (x_i - \bar{x}_i)^2}$ | The most commonly used performance evaluation metric | - The closer to 1, the better the prediction |
| MAPE | $\frac{1}{n} \sum_{i=1}^n \left \frac{x_i - y_i}{x_i} \right \times 100$ | Measure the size of the error as a percentage | - The lower the value, the better prediction |

Table 6. The result of daily temperature model

| Evaluation metric | value |
|-------------------|-------|
| RMSE | 48.6 |
| rRMSE | 31.8% |

6. 시나리오 분석

DNN모형은 요인과 결과를 통해 범칙을 추론하는 방법인 바, 방법론을 먼저 결정한 후 결과를 예측하는 기존의 연구방법에 비해 연구방법이 블랙박스로 존재한다. 따라

서 본 모델의 높은 예측력과는 무관하게 적합한 정책을 도출하기 위한 방안으로 활용하기는 어려움이 있다. 따라서 본 연구에서는 시나리오 수립 및 각 시나리오에 알맞은 데이터를 생산하여 모델의 입력값으로 사용하고 결과를 예측함으로써 본 모형의 활용가능성을 제고하고자 한다.

각 시나리오는 다음과 같이 구성한다.

- 공급중심전력정책과 수요관리정책이 전력수요에 미치는 영향
 - 경제성장 또는 하락이 전력수요에 미치는 영향
- 시나리오 시 관련 데이터만 변경하며 데이터가 미치는 영향을 확인한다.

6.1 공급중심정책과 수요관리정책이 전력수요에 미치는 영향

본 시나리오는 전력정책이 공급중심인 경우 및 수요관리정책으로 구분하여 전력수요가 어떻게 변화하는지 확인한다. 각 정책은 다음과 같이 정의한다.

- 공급중심정책 : 설비용량은 6개월마다 5%씩 증가
- 수요관리정책 : 설비용량 변화 없음

Fig. 8에서 보듯 공급중심전력정책은 수요관리정책에 비해 전력수요가 소폭 낮게 나타나지만 시간이 지남에 따라 전력설비용량이 증가하면서 전력수요 또한 상승하는 것으로 예측되었다.

6.2 경제성장 또는 하락이 전력수요에 미치는 영향

시나리오는 경제성장 또는 하락이 전력수요에 어떤 영향을 미치는지 확인한다. 각 경우는 다음과 같이 구분한다.

- 경제성장 : 기존 소비자심리지수와 무역수지의 최댓값에서 매달 1% 상승
- 경제하락 : 기존 소비자심리지수와 무역수지의 최댓값에서 매달 1% 하락

경제성장과 경제하락의 경우를 명확히 구분한 바, 각 경우 예측값들의 차이가 명확하게 나타난다. 경제성장의 경우 경제하락에 비해 지속적으로 높은 전력수요를 보이거나 경제가 성장할수록 그 격차가 커지는 등의 경향성은 낮게 나타난다.

또한 경제하락 시 전력수요는 하락과는 무관하게 일정 수준의 전력수요가 지속적으로 필요함에 예측되었다.

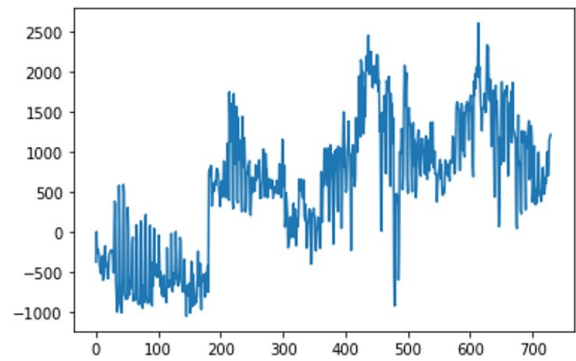


Fig. 8. The gap between supply oriented policy

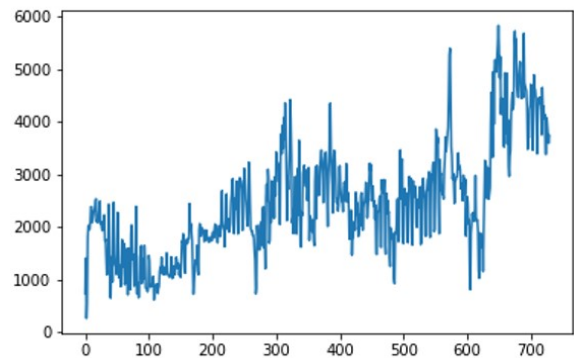


Fig. 9. The gap between economic growth and economic recession

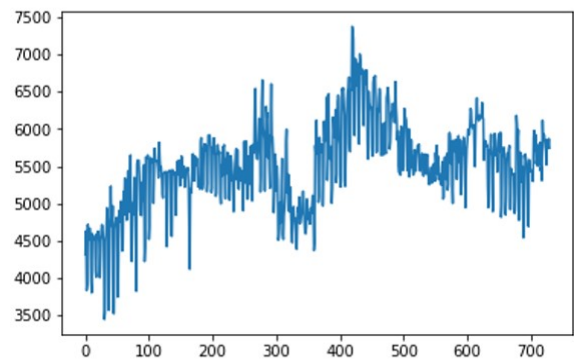


Fig. 10. The prediction of economic recession

7. 한계

본 논문의 한계는 크게 3가지로 구분할 수 있다.

- 수집 데이터의 한계: 소비자심리지수나 무역수지 등 경제상황을 나타내는 지표들은 주로 월별 데이터만을 제공한다. 따라서 일일 데이터 예측을 위해 주기가 동일한 데이터를 사용하는 것으로 가정하였다.
- 코로나19 예측 모형의 한계: 확진자수는 과거 추세 뿐만 아니라 거리두기, 백신접종 등 보건정책에 큰 영향을 받는다. LSTM 모형은 보건정책 시행에 따른 변화 등을 보여주기에는 아직 한계가 있었다.
- 각 변수간의 상관관계 연구 부족: 딥러닝 모형은 데이터와 결과를 통해 법칙을 도출하는 방법인 바, 정책적 활용을 위해서는 새로운 데이터 입력을 통해 결과를 예측하는 민감도 분석과 유사한 방법을 통해 그 결과를 확인한 후 정책에 활용할 수 있다. 본 논문에서는 단순한 시나리오를 가정한 바, 변수간의 상관관계를 고려하지 않았다.

위의 한계를 통해 추후 연구는 아래의 내용을 보강하고자 한다.

- 일일 경제 지표 선정: 주가지수 또는 통화량 등 경제 상황을 반영하는 지표를 선정하여 모형의 예측도를 제고하도록 한다.
- 코로나19 예측 모형 개선: LSTM 모형에서 대표적인 질병확산 모형인 SIR 모형으로 변경하여 정책적 특징을 반영한 정교한 코로나19 확산 예측결과를 활용한다.
- 각 변수간의 상관관계 분석: 입력 변수간의 상관관계를 분석하여 시나리오 작성 시 더욱 정교한 데이터 생산에 활용한다.
- RE100 등 신재생에너지 수요 정책 반영: 태양광을 비롯한 신재생에너지에 대한 수요가 지속적으로 확대되고 있고 이에 따른 제도가 도입되고 있는 바, 정책분석을 통해 입력변수를 다변화한다.

향후 위의 내용을 반영한 추가연구를 통해 본 연구보다 정교한 전력수요예측 모형 개발이 가능하리라 예상된다.

References

Al-Shobaki S, Mohsen M. 2008. Modeling and forecasting of electrical power demands for capacity planning.

- Energy Conversion and Management. 49(11):3367-3375.
- Alex D. Papalexopoulos TCH. 1990. A regression-based approach to short-term system load forecasting. IEEE Transactions on Power Systems. 5(4):1535-1559.
- Amarawickrama H, Hunt L. 2008. Electricity demand for sri lanka: A time series analysis. Energy. 33(5):724-739.
- Baek J-K, Ha J-H. 2015. A study on calibrating the forecasted load of electric power considering special day fact. Journal of Industrial Economics and Business. 28(1):191-203.
- Bakay MS, Ağbulut Ü. 2020. Electricity production based forecasting of greenhouse gas emissions in turkey with deep learning, support vector machine and artificial neural network algorithms. Journal of Cleaner Production.
- Bedi J, Toshniwal D. 2019. Deep learning framework to forecast electricity demand. Applied Energy. 238:1312-1326.
- Bessec M, Fouquau J. 2008. The non-linear link between electricity consumption and temperature in europe: A threshold panel approach. Energy Economics. 30(5):2705-2721.
- Bianco V, Manca O, Nardini S. 2013. Linear regression models to forecast electricity consumption in italy. Energy Sources, Part B: Economics, Planning, and Policy. 8(1):86-93.
- Covid-19 impact on electricity. 2020. IEA, Paris; [accessed 2020]. <https://www.iea.org/reports/covid-19-impact-on-electricity>.
- Deep learning's mathematics. 2020. [accessed 2021 Feb. 19]. <https://towardsdatascience.com/deep-learnings-mathematics-f52b3c4d2576>.
- Fan S, Hyndman RJ. 2011. The price elasticity of electricity demand in south australia. Energy Policy. 39(6):3709-3719.
- Fung W, Lam K, Hung W, Pang S, Lee Y. 2006. Impact of urban temperature on energy consumption of hong kong. Energy. 31(14):2623-2637.
- G. Peter Zhang BEP, Michael Y. Hu. 2001. A simulation study of artificial neural networks for nonlinear time-series forecasting. Computers & Operations

- Research. 28:381-396.
- García-Ascanio C, Maté C. 2010. Electric power demand forecasting using interval time series: A comparison between var and impl. *Energy Policy*. 38(2):715-725.
- Geem ZW, Roper WE. 2009. Energy demand estimation of south korea using artificial neural network. *Energy Policy*. 37(10):4049-4054.
- IEA. 2019. Energy prices and taxes. International Energy Agency.
- Jamil F, Ahmad E. 2010. The relationship between electricity consumption, electricity prices and gdp in pakistan. *Energy Policy*. 38(10):6016-6025.
- Jang H, Jo J. 2018. Analysis of the economic feasibility of flexible facilities according to the expansion of the proportion of renewable energy generation in the era of energy transition. Korea Energy Economics Institute. Research Paper 18-02.
- Kavaklioglu K, Ceylan H, Ozturk HK, Canyurt OE. 2009. Modeling and prediction of turkey's electricity consumption using artificial neural networks. *Energy Conversion and Management*. 50(11):2719-2727.
- Kim H-m, Kim I-g, Park K-J, Yoo S-H. 2015. The effect of temperature on the electricity demand: An empirical investigation. *Journal of Energy Engineering*. 24(2):167-173.
- Kim I-M, Lee Y-J, Lee S, Kim D. Holiday Effects of Disaggregated Sectoral Demand for Electricity. *Korean Energy Economic Review* 15(1): 99-137.
- Lee M. 2014. An analysis on the effects of higher power rates on supply price and power savings for korean manufacturing sector. *Environmental and Resource Economics Review*. 23(1):43-65.
- Mohamed Z, Bodger P. 2005. Forecasting electricity consumption in new zealand using economic and demographic variables. *Energy*. 30(10):1833-1843.
- Pao H. 2006. Comparing linear and nonlinear forecasts for taiwan's electricity consumption. *Energy*. 31(12):2129-2141.
- R. E. ABDEL-AAL AZA-G. 1997. Forecasting monthly electric energy consumption in eastern saudi arabia using univariate time-series analysis. *Energy*. 22(11):1059-1069.
- S. MAKRIDAKIS AA, R. CARBONE, R. FILDES, M. HIBON, R. LEWANDOWSK, J. NEWTON, E. PARZEN, R. WINKLER. 1982. The accuracy of extrapolation (lime series) methods: Results of a forecasting competition. *Journal of Forecasting*. 1:111-153.
- S. Saravanan SKaCT. 2012. Forecasting india's electricity demand using artificial neural network. *IEEE-International Conference On Advances In Engineering*. 79-83.
- Samer Saab EB, George Nasr. 2001. Univariate modeling and forecasting of energy consumption: The case of electricity in lebanon. *Energy*. 26(1):1-14.
- Silvia Weko LE, Rainer Quitzow, Germán Bersalli, Flávio Lira, Adela Marian, Diana Süsser, Sapan Thapar, Bing Xue. 2020. Covid-19 and carbon lock-in : Impacts on the energy transition. In: *Studies IfAS*, editor.: IASS POTSDAM.
- Yu B. 1996. Price elasticity of demand and the tariff adjustment plan. Korea Energy Economics Institute.