

# Prediction of Electric Power on Distribution Line Using Machine Learning and Actual Data Considering Distribution Plan

## 배전계획을 고려한 실데이터 및 기계학습 기반의 배전선로 부하예측 기법에 대한 연구

Junhyuk Kim, Byung-Sung Lee

### Abstract

In terms of distribution planning, accurate electric load prediction is one of the most important factors. The future load prediction has manually been performed by calculating the maximum electric load considering loads transfer/switching and multiplying it with the load increase rate. In here, the risk of human error is inherent and thus an automated maximum electric load forecasting system is required. Although there are many existing methods and techniques to predict future electric loads, such as regression analysis, many of them have limitations in reflecting the nonlinear characteristics of the electric load and the complexity due to Photovoltaics (PVs), Electric Vehicles (EVs), and etc. This study, therefore, proposes a method of predicting future electric loads on distribution lines by using Machine Learning (ML) method that can reflect the characteristics of these nonlinearities. In addition, predictive models were developed based on actual data collected at KEPCO's existing distribution lines and the adequacy of developed models was verified as well. Also, as the distribution planning has a direct bearing on the investment, and amount of investment has a direct bearing on the maximum electric load, various baseline such as maximum, lowest, median value that can assesses the adequacy and accuracy of proposed ML based electric load prediction methods were suggested.

*Keywords: Distribution Planning, Machine Learning, Load Prediction, Actual Data, SOMAS*

### 1. Introduction

배전계획은 미래에 발생할 수 있는 잠재적 위험을 미리 예측하여 배전계를 보강하는 등의 고품질 전력을 제공하기 위한 투자사결정 행위라고 볼 수 있다. 이러한 배전계획은 각 배전선로의 최대부하 예측값을 기반으로 수행되는데, 현재 최대 부하 예측은 배전 계획을 수립하는 직전 년도의 최대 부하 값에 부하증감율을 곱하여 수행되며, 직전 년도의 최대 부하 값을 산정하는 과정에서 부하전환이나 고장에 의한 일시적인 부하변화에 의한 최대 부하값은 제외하여 계산하고 있다 [1]. 아울러 부하증감율의 경우에도, 과거 10개년 이상의 행정구역별 전력판매량 이력과 각 배전선로의 점유율을 고려하여 도출되는 직관적인 방법이 적용되고 있다 [1][2]. 현행과 같이 배전 계획을 위한 부하 예측 과정에서 부하 절체나 고장에 의한 일시적인 최대 부하를 제외하는 행위는 일정부분 행위자의 경험과 지식에 의존해야 하므로, 인적 오류가 발생하거나 행위자에 따라서 상이한 결과가 야기될 가능성이 있다. 부하증감율을 산출하는 과정도 행정구역 별 전력판매량과 해당 배전선로의 점유율 만을 고려하므로 정밀한 배전선로 부하예측에는 한계를 가질 수 있다.

이러한 한계점을 보완하기 위하여 다양한 기법과 연구들이 수행되어 왔다. 회귀분석 기반의 부하 예측 방법도 연구되었으며 [3]-[5], 부하데이터를 시계열 패턴화하여 부하예측을 수행한 연구도 존재한다 [6]. 그러나 대부분의 기존 부하 예측 방법이나 연구는 배전 계통의 운전환경 변화 등에 기인하여 부하 패턴이 바뀌는 경우 능동적으로 대응하기 어렵다는 한계를 가진다. 특히 미래 배전 계통의 운전환경은 이동식 부하인 태양광발전이나 전기자동차의 빈번한 연계가 예상되며, 이는 배전계통 해석의 측면에서 복잡성 및 가변성을 가중시킬 것이 자명하므로, 이러한 능동적 대응 능력은 부하 예측의 측면에서 극복해야할 이슈로 부각되고 있다. 능동적 대응이라는 문제를 해소하기 위하여 기계학습을 접목한 부하 예측 방법이 대두되고 있다 [7]-[17].

기계학습의 경우 세부적인 모델의 구동원리나 목적은 다소 상이할 수 있으나, 과거의 데이터를 학습하여 미래를 예측한다는 측면에서는 모두 공통점을 가지고 있다. 그러나 이는 달리 말하면 과거의 데이터가 높은 신뢰성을 가지는 경우에만 예측 결과가 신뢰성을 가질 수 있음을 의미한다. 전력 부하의 특성상 이상 및 노이즈가 데이터 내에 다소 포함되어 있으므로, 신뢰성 있는 예측 결과를 도출하기 위해서는 이에 대한 적절한 처리가 전제되어야 한

### Article Information

Manuscript Received July 28, 2020, Accepted October 29, 2020, Published online June 30, 2021

The Authors are with KEPCO Research Institute, Korea Electric Power Corporation, 105 Munji-ro Yuseong-gu, Daejeon 34056, Republic of Korea.

Correspondence Author: Junhyuk Kim (kim\_jh@kepcoco.kr)

ORCID: 0000-0002-2849-9291 (J. Kim)



This paper is an open access article licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International Public License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0>  
This paper, color print of one or more figures in this paper, and/or supplementary information are available at <http://journal.kepcoco.kr>.

TABLE 1  
연도별 SOMAS 선로부하 데이터 구성

선로명	정상	누락	이상	가용율
동방	43453	11	240	99.43
죽동	40339	8	213	99.46
목원	43378	12	266	99.36
선경#1	43450	11	243	99.42
유통	43449	11	244	99.42
구암	43443	11	250	99.4
한신	43248	13	275	99.34
대화#1	43010	13	297	99.28
대평	43089	12	267	99.36
두계	42792	14	322	99.22

TABLE 2  
최적화된 Hyper Parameter 예시

Number of Node	200	Batch Size	16
Number of Layer	5	Epochs	20
Activation Function	relu	Loss Function	MSE
Optimiser	adam		

다. 이러한 이유로 대부분의 기존연구들은 신뢰도 높은 센서에서 고품질의 데이터를 수집할 수 있는 빌딩이나 몇몇 가구를 대상으로 미래부하를 예측하거나, 노이즈나 이상데이터의 영향이 적은 도시 단위 또는 국가 단위의 부하예측을 수행하였다.

본 연구에서는 배전계획 측면에서 필요한 배전선로 단위로 한전의 각 배전선로에서 수집된 실제 데이터를 기계학습 예측모델에 접목시켜 배전선로 부하 예측 가능성을 검토하고, 예측된 결과에 대하여 배전 계획 측면의 정확도 평가 수식을 제안하여 그 적정성을 평가 및 검증하고자 한다.

## II. 본문

본 연구에서는 배전선로에서 수집된 실제 데이터를 활용하여 미래부하를 예측하는 방법으로, 널리 알려진 LSTM (Long Short-Term Memory)과 해당 방법의 한계점을 극복하기 위한 군집화 기반의 하이브리드 예측기법을 제안 및 적용하고, 두 가지 방법을 분석하였다. 본문에서는 활용된 학습데이터 및 각 학습 모델에 대한 설명과 적용결과에 대하여 기술한다.

### A. 학습데이터

배전선로의 부하 예측을 위하여 변전소 내 배전선로 인출단에서 측정되는 배전선로 부하값을 사용하였다. 해당 데이터는 한국 전력의 변전소 운영관리 자동화 시스템(Substation Operation Management Automation System, 이하 SOMAS)에서 수집하였으며, 2015년부터 2019년까지 총 5개년의 전국 배전선로 데이터를 활용하여 예측을 수행하였으며, 부하예측 및 정확도 검증 등은 대전충남본부 내 대덕유성지사의 10개 배전선로를 대상으로 수행하였

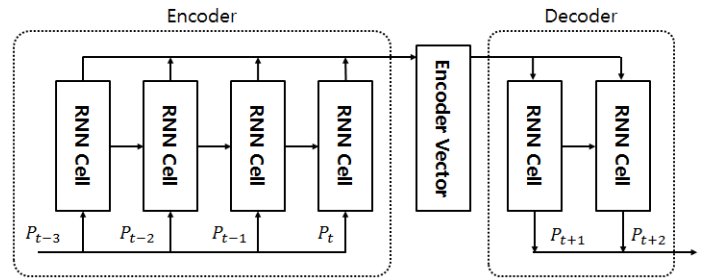


Fig. 1. Encoder-Decoder 기반 RNN의 일반적인 구조.

다. SOMAS에서 수집되는 데이터의 경우, 전력 계획 수립에도 활용되는 신뢰도 높은 데이터로 누락 및 이상데이터가 상대적으로 낮은 빈도로 포함되었다는 특성을 가지며, 시간별 최대부하값을 기록하여 제공하고 있다. 실제 수집된 데이터의 연도별 이상 및 누락데이터 비율은 TABLE 1과 같다.

### B. LSTM 예측방법

본 연구에서는 기존에 존재하는 다양한 기계학습 방법 중에서 전력데이터의 특성 중 하나인 시계열 데이터 예측에 특화된 RNN (Recurrent Neural Network) 알고리즘을 적용하였다. 또한, 특정 시점의 데이터 패턴이 과도하게 반영되지 않고 과거의 전력데이터 패턴을 충분히 반영하여 미래 부하를 예측할 수 있도록, 일반적인 RNN이 아닌 과거의 데이터를 더 효과적으로 반영할 수 있는 LSTM 모델을 적용하였다 [18][19].

일반적으로 시계열 데이터와 같은 순차데이터의 예측을 위해서 사용하는 RNN 구조는 Many to One 또는 Many to Many 형태로 구현한다. Many to One은 일련의 입력 시퀀스를 입력받아 하나의 결과를 생성하며, Many to Many는 일련의 입력 데이터를 입력받아 복수 개의 결과를 생성한다. 이러한 방식은 시계열 데이터 예측에 사용할 수 있으나, 결국 예측값을 입력값으로 활용하여 다시 예측하게되는 구조를 가지므로, 예측 오차의 누적으로 전체 학습결과에 큰 오차를 야기할 수 있는 합성 오차(compounding error) 등을 야기할 수 있다는 문제를 내포하여 시계열 데이터 예측에는 최적의 모델이라 할 수 없다. 또한, 기계학습 기반의 시계열 데이터 예측은 직관적으로 1단계에서 입력데이터를 관측하고 시스템의 현재 상태를 추론하며, 2단계에서 미래를 예측하는 방법을 이해하고 적용하게 되는데, 일반적으로 사용되는 Single RNN 모델에서 이를 구현하기 위해서는 하나의 cell이 과거의 중요한 이벤트를 기억하고, 동시에 기억한 이벤트를 미래 예측에 사용할 수 있어야 한다. 해당 모델에서 이를 반영하는 것은 분명한 한계가 있으므로, 본 연구에서는 이러한 일련의 단계를 2개의 특별한 cell에 각각 할당할 수 있는 Encoder-Decoder 기반의 RNN LSTM 모델을 적용하였다. Encoder-Decoder LSTM 모델에서 과거의 중요한 이벤트를 기억하는 cell을 Encoder라 하며, 기억된 정보를 미래 예측값으로 변환하는 cell을 Decoder라 한다. Fig. 1에 Encoder-Decoder 기반 RNN 모델의 간략한 구조를 나타냈다 [20].

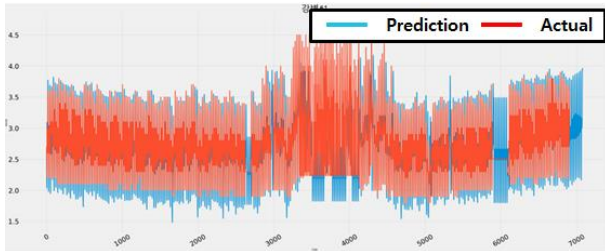
적용한 Encoder-Decoder LSTM 모델은 널리사용되는 기계학습 라이브러리 중 하나인 케라스(Keras)를 이용하여 구축하였으며, 이를 Fig. 2에 나타냈다. 아울러 해당 모델을 사용하여 학습하고 예측을 수행하기 위하여 적용한 Hyper Parameter와 예측 결과를 아

```

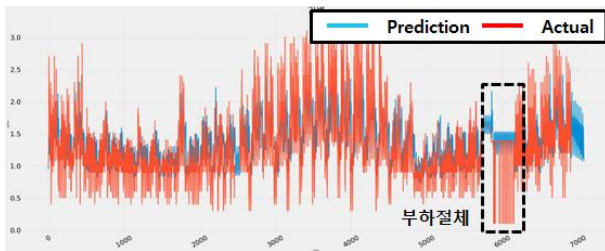
from keras.models import Sequential
from keras.layers import Dense, Activation

model = Sequential()
model.add(LSTM(200, activation = 'relu', input_shape = (n_timesteps, n_features)))
model.add(RepeatVector(n_outputs))
model.add(LSTM(200, activation = 'relu', return_sequence = True))
model.add(TimeDistributed(Dense(100, activation = 'relu')))
model.add(TimeDistributed(Dense(1)))
model.compile(loss = 'mse', optimizer = 'adam', metrics = ['accuracy'])
    
```

Fig. 2. 케라스(Keras)를 활용하여 구축한 Encoder-Decoder LSTM 모델.



(a)



(b)

Fig. 3. Encoder-Decoder LSTM 기반 7 일 부하예측 결과. (a) 강변#1 D/L. (b) 기병 D/L.

래의 TABLE 2 및 Fig. 3에 나타냈다.

Fig. 3에서 확인할 수 있는 것과 같이, 강변#1 배전선로의 경우 전체적인 패턴을 잘 추종하여 예측하는 것을 확인할 수 있다. 반면, 기병 배전선로의 경우 예측값과 실측값의 오차가 상당함을 확인할 수 있다. 이러한 오차는 적용된 Hyper Parameter가 강변#1 배전선로에 최적화되었기 때문이다. 달리 말하면, 최적화된 Hyper Parameter를 도출할 수 있다면 Encoder-Decoder LSTM 방식을 적용하여 납득할 만한 수준의 부하 예측이 가능하나, 이를 위하여 모든 배전 선로에 별도로 최적화를 수행하여야 한다는 것을 의미한다. 또한, Fig. 3(b)의 점선으로 표기된 부분과 같이 부하절체에 의한 부하패턴 변화가 발생하는 경우, 절체여부를 자동으로 파악하기 어려우므로 이를 학습하여 예측결과에 해당 패턴이 반영될 수 있다는 문제점을 가질 수 있다. 이러한 한계점을 극복하기 위하여 군집화 기반의 부하예측 기법을 적용하는 방안에 대한 추가 연구를 수행하였다.

C. 군집화 기반 부하예측방법

앞서 언급한 것과 같이, Encoder-Decoder LSTM 모델을 위시한 RNN 계열 모델의 예측을 위해서는 해당 배전 선로별 최적화된

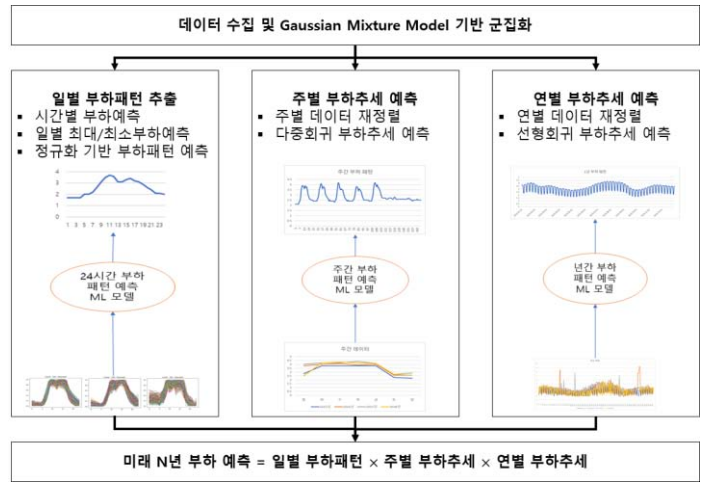


Fig. 4. 군집화 기반 학습데이터 구축 및 부하예측 프로세스.

Hyper Parameter 도출이 필요하다. 해당 과정은 프로그래밍을 통하여 일부 자동화를 시킬 수 있지만 별도의 연산을 수행하여야 하며, 경우에 따라 최적화에 많은 시간이 소요될 수 있다. 아울러, 부하절체나 고장 등 일시적인 부하패턴 변화 등 이상 상황에 대한 선별적 학습이 어려워 이를 해소하기 위해서는 이상 상황에 의한 일시적 패턴변화가 학습에 사용되지 않도록 하는 등의 추가 설정이 필요하다.

그러므로, 본 연구에서는 이러한 문제점을 완화할 수 있는 군집화 기반 예측기법을 적용하였다. 군집화(Clustering)란 비지도 학습으로 적용할 수 있는 한 분야로 데이터의 유사성을 기준으로 군집을 형성하는 기법이다. 해당 방식을 적용하는 경우, 군집을 구성하는 객체 수가 작은 군집을 일시적 패턴변화 등에 따른 대표성이 낮은 패턴으로 분류하여 학습데이터에서 제외하는 등의 처리가 가능하다. 군집화를 적용한 부하예측 방법을 Fig. 4에 정리하였다.

제안한 군집화 기반 부하예측방법은 먼저 사용 데이터별 최적 군집화 기법을 선정하며, 선정된 군집화 기법을 적용하여 대표성을 가지는 학습 데이터를 구축하고, 해당 데이터를 학습하여 일별 부하의 패턴을 도출한다. 이렇게 도출된 일별 대표패턴을 조합하여 연 대표패턴을 도출할 수 있다. 다음 단계로는 다중회귀기법을 기반으로 연별 또는 월별 최대값 증감추이를 도출하여 이를 기-도출한 연 대표패턴에 적용하여 미래부하예측을 수행한다.

1) 군집화 기법 선정

시계열 부하예측에 적용할 군집화 기법 선정을 위하여 널리 쓰이는 군집화 알고리즘인 DBScan, K-Means, Bayesian Mixture Model, Gaussian Mixture Model을 적용하여 각 방법의 비교 검증을 수행하였다. DBScan알고리즘을 적용한 결과, 군집화에 따른 군집이 명확히 나타나지 않아 시계열 데이터 분류로는 적용이 적합하지 않음을 확인할 수 있었고, K-means알고리즘과 Bayesian Mixture Model의 경우 상위로 분류된 군집 내의 데이터 간 유사도가 높지 않으며, 이상데이터가 다수 포함되어 부하예측을 위한 군집화 기법으로는 한계가 있음을 확인하였다.

이와 같은 결과는 TABLE 1에 기술한 10개 배전선로를 위시하여 대전충남본부 내 모든 배전선로에서 유사하게 나타남을 확인하였다. 반면, Gaussian Mixture Model을 적용한 결과, 군집 내의 유

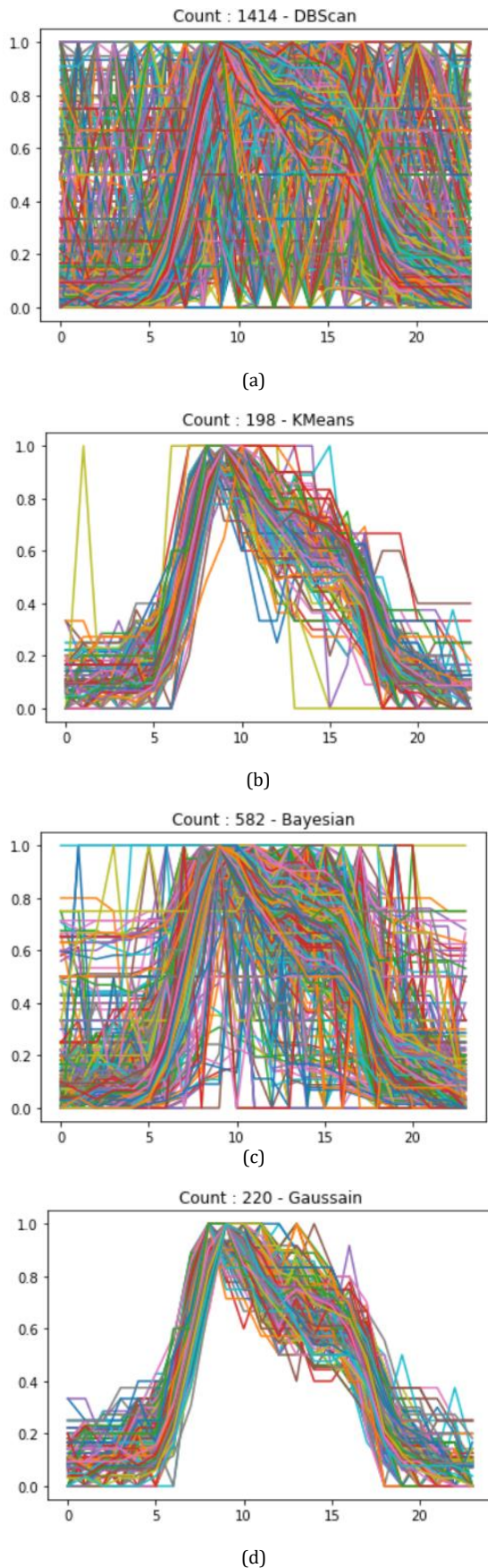


Fig. 5. 대전충남 홍명 D/L 군집화 대표결과: (a) DBSCAN. (b) K-means. (c) Bayesian Mixture (d) Gaussian Mixture Model.

사도가 명확하며 이상데이터의 군집을 명확히 구분하는 것을 확인할 수 있다. 특히, 상위 군집에서 명확한 패턴을 나타내고 있어, 이를 활용하여 대표성을 가지는 시간별 부하패턴 추출이 가능함을 확인하였다. 이러한 결과는 배전선로의 부하데이터가 표준 분포를 따르지 않으며, 다른 군집화 기법에 비하여 Gaussian Mixture Model이 이러한 특성을 효과적으로 반영하기 때문으로 분석된다.

### 2) 주중 및 주말 패턴 도출

본 연구에서는 주중 및 주말의 상이한 부하패턴을 반영한 배전선로 부하예측을 수행하기 위하여, 앞서 도출된 군집화 기법 기반으로 주중 및 주말 부하패턴의 변화를 분석하였다. 실제 내황 배전선로의 부하데이터를 활용하여 분석한 결과, 요일별로 도출된 패턴은 예상한 것과 같이, 주중에도 요일에 따라 다소 차이가 발생하지만 주중과 주말의 패턴 변화가 더욱 확연한 것을 확인할 수 있다. 이처럼, 주중과 주말을 모두 고려한 부하예측은 예측결과의 정확도를 개선하는데 일조할 수 있으며, 아울러 이 과정에서 일별 최대 및 최소값을 도출해낼 수 있는데 이는 전체 학습 기간 동안 가이드로 활용되어 일별부하예측 정확도를 개선하고 최대부하 추정을 위한 방향성을 정립하는데 활용될 수 있다. Fig. 6은 군집화 기법으로 도출된 내황 배전선로의 요일별 군집화 결과를 나타낸다.

### 3) 군집화 기반 배전선로 부하예측

군집화 기반 예측모델 도출을 위하여 아래와 같이 크게 다섯 단계의 과정을 거친다.

① 대표 군집 도출: 앞서 기술한 것과 같이 Gaussian Mixture Model을 적용하여 배전선로의 데이터를 분석하였으며, 그 결과로 휴리스틱(Heuristic) 분석을 통하여 도출된 도수의 합이 상위 70% 이내에 포함되는 클래스에 속한 데이터만을 대표성을 가지는 것으로 판단하였다.

② 일별 대표패턴 도출: 대표 군집 내 복수 객체에 대한 최대값 또는 평균값을 기준으로 해당 군집의 일별 대표패턴을 도출한다. 주중과 주말의 패턴이 상이하므로 일별 대표패턴 도출 및 조합을 통하여 부하예측의 정확도 개선의 효과를 기대할 수 있다.

③ 연 패턴 도출: 앞서 도출된 일별 패턴의 조합을 통하여 일별 및 주별 패턴을 도출할 수 있다. 이렇게 도출된 일별 및 주별 패턴을 조합한다면 1년에 대한 패턴을 도출해낼 수 있다.

④ 월별 및 연별 추세 도출: 학습 데이터의 일별 대표값을 기반으로 다중회귀분석을 적용하여 월별 또는 연별 부하증감추세를 도출해낸다. 아울러 일별 대표값의 경우 중앙값, 평균값, 최빈값, 최대값 등이 고려될 수 있으나, 본 연구에서는 최대값을 기반으로 의사결정이 이루어지는 배전계획의 특성을 반영하여 최대값을 일별 대표값으로 설정하였다.

⑤ 부하예측: 도출된 연별 대표패턴에 월별 또는 연별 추세를 조합하여 미래부하를 예측한다.

다만 제안 방법의 경우, 과거의 패턴이 미래에도 유사하게 지속될 것이라는 전제를 가지는데 이의 타당성을 검증하기 위하여 금남 배전선로의 2015년부터 2019년 데이터를 분석하였다. Fig. 7에서 확인할 수 있는 것과 같이, 부하절체 등으로 사료되는 일시적인 부하 급증을 제외한다면, 전체 5년의 기간 동안 유사한 일별패턴을 가지는 것을 확인할 수 있다. 또한, 기술한 다섯 단계를 통해 금남 배전선로의 부하예측을 수행한 결과를 Fig. 8에 나타냈다.

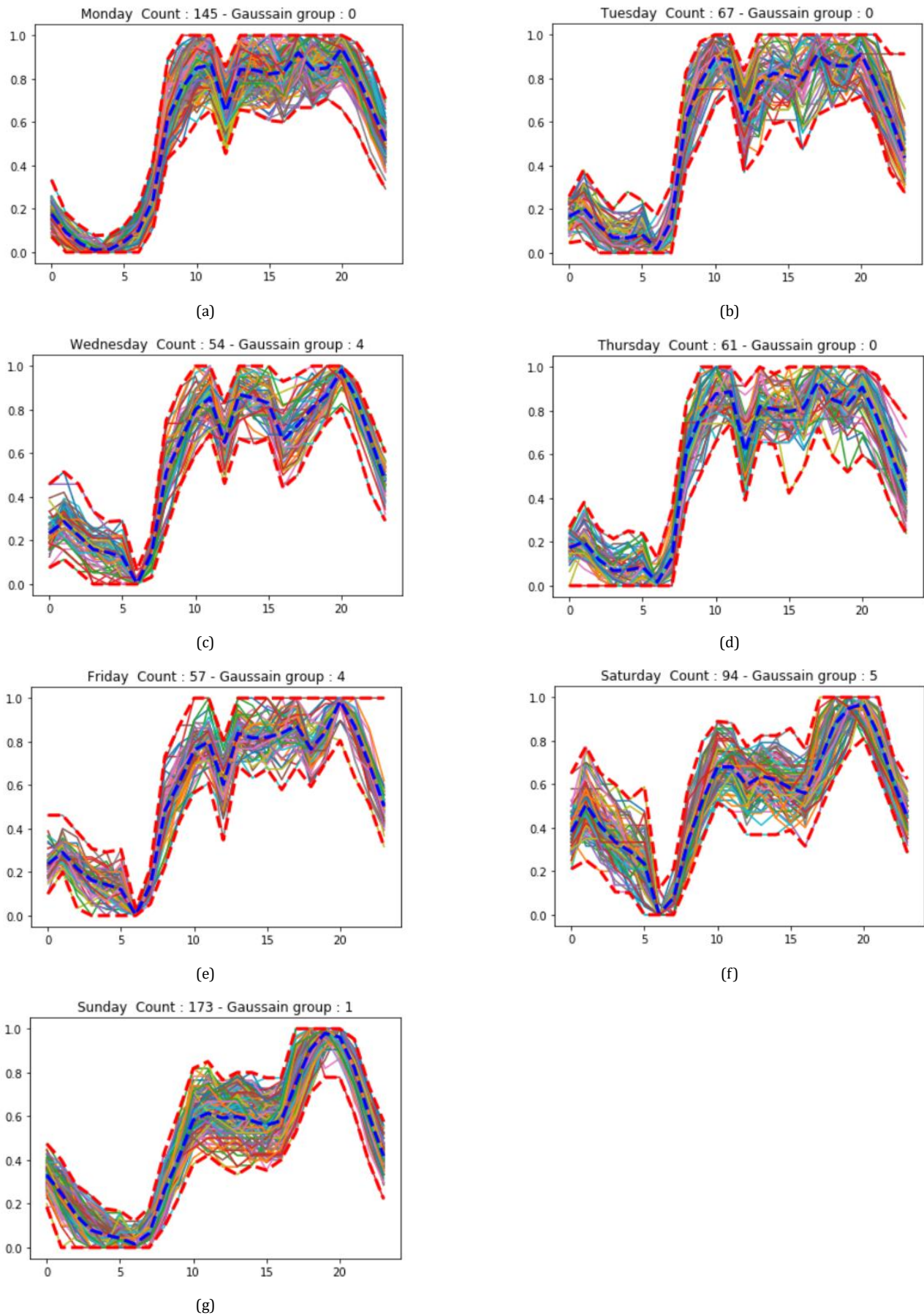


Fig. 6. 요일별 Gaussian Mixture Model 기반 클러스터링 결과(대전충남 내황 D/L): (a) 월요일, (b) 화요일, (c) 수요일, (d) 목요일, (e) 금요일, (f) 토요일, (g) 일요일.

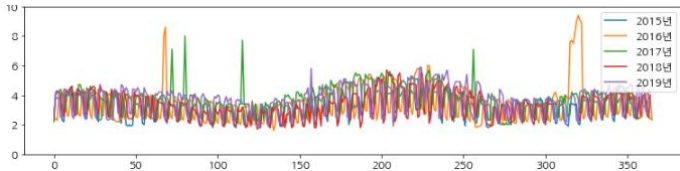


Fig. 7. 일별 부하패턴 예시(금남 D/L, 2015~2019).

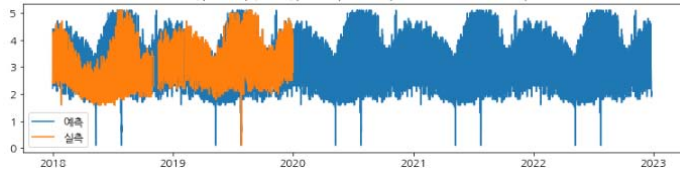


Fig. 8. 다년간 부하예측 예시(금남 D/L, 2020~2022).

4) 예측결과 평가

기계학습 모델의 예측정확도의 평가는 확보된 실측 데이터를 분리해서 사용하는 것이 일반적이다. 이 경우, 데이터 구성에 따라서 정확도 검증 결과가 달라질 수 있는데, 본 연구에서는 실측 데이터 전체를 학습데이터이자 검증데이터로 활용하였다. 이 경우 과적합 문제가 발생할 수 있다고 알려져 있으나, 본 연구에서 사용한 학습데이터의 경우 데이터의 양이 충분하여 해당 문제를 일부 상쇄시킬 수 있다. 아울러 일반적으로 사용되는 정확도 검증식은 Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), Normalized Root Mean Square Error (NRMSE) 등이 있는데, RMSE는 모델의 예측값과 실제값의 차이를 숫자로 표현할 수 있어 기계학습의 Loss Function으로 쓰이기도하나, 크기 의존적 에러(Scale Dependent Error)를 포함할 수 있다는 한계가 있으며, MAPE는 이러한 크기 의존적 에러를 보완하기 위한 수식이나, 실제 값이 0인 경우 MAPE 값 자체를 계산할 수 없다는 한계를 가진다. 또한, 이러한 문제 외에도 배전계획 측면에서의 의사결정은 선로의 최대부하값을 기준으로 수립되므로, 전체 데이터를 얼마나 잘 추종하고 예측하느냐 외에 최대값을 얼마나 잘 예측하는지가 경우에 따라 더 중요할 수 있다. 그러므로, 본 연구에서는 백분율로 표현하여 직관적이며 계산에 의한 제약이 적은 NRMSE에 별도의 무게합수(weight)를 추가로 부과하여 선로부하 최대값 예측의 중요도를 높여 제안모델의 적정성을 검증할 수 있는 아래의 식을 제안하였다.

$$NRMSE_{max} = \frac{\sqrt{\frac{1}{n} \sum \left( (y_{actual} - y_{predict}) \times \left( \omega + \frac{y_{actual}}{y_{max}} \right) \right)^2}}{y_{max} - y_{min}} \quad (1)$$

여기서,  $y_{actual}$ : 실측값,  $y_{predict}$ : 예측값,  $y_{max}$ : 실측 최대값,  $y_{min}$ : 실측 최소값,  $\omega$ : 무게합수,  $n$ : 관측수이다.

Eq. (1)을 적용하는 경우, 실측값과 예측값의 차이에 최대 부하 대비 실측값의 비율만큼이 추가되어 최대 부하값과 가까운 값들에 대한 중요도를 높이는 효과가 있다. Eq. (1)에서  $\omega$ 는 부여하고자 하는 무게값으로 의사결정에 따라 조정이 가능하다. Fig. 9에  $\omega$ 는 1로 일괄 적용하였을 때, 대전충남본부 내 대덕유성지사의 10

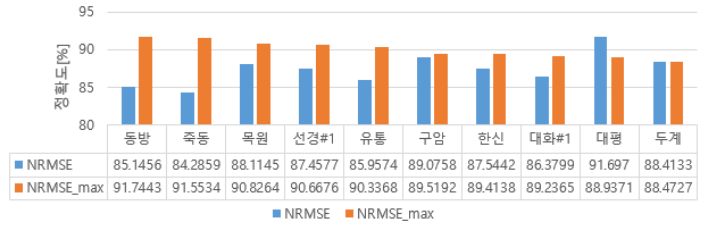


Fig. 9. 정확도평가방법에 따른 비교결과 (NRMSE vs. NRMSEmax).

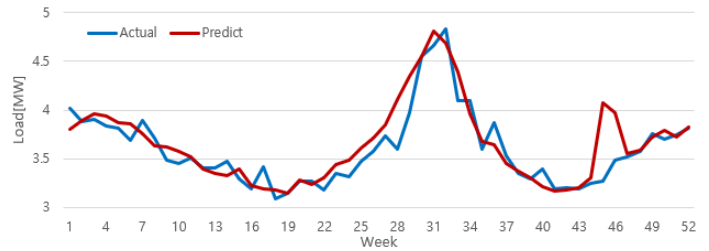


Fig. 10. 대평 D/L 2019년 실측 및 예측값 비교결과.

개 배전선로에 대한 기존 NRMSE 방식과 NRMSEmax에 따른 정확도 분석 결과를 나타내었다. 분석결과, 대평 배전선로의 경우 NRMSEmax 적용 시 정확도가 저하되는 것을 확인할 수 있는데, 이는 해당 배전선로의 경우 전체 부하추세는 정확도 91.69%로 매우 잘 추정하나 최대부하 예측의 경우 88.89%로 상대적으로 예측정확도가 낮음을 의미한다. 하지만 이 역시도 Fig. 10에서 확인할 수 있는 것처럼 미소한 차이로, 실제 실측 최대부하값은 4.833 MW이고 예측 최대부하값은 4.812 MW로 약 0.02 MW정도의 오차를 가지는 것을 확인하였다.

또한, 제안 모델의 검증에 사용한 10개 배전선로의 경우, 전체 패턴을 얼마나 잘 예측하는지를 나타내는 NRMSE가 85%이상의 결과를 가짐을 확인할 수 있다. 아울러 최대값을 얼마나 잘 예측하는지를 나타내는 NRMSEmax의 경우 대부분의 선로에서 오히려 NRMSE보다 더 높은 정확도를 가지는 것을 확인할 수 있는데, 이는 앞서 일별 부하패턴 도출과정에서 명시한 것과 같이 일별 대표값 선정에 최대값을 활용하였기 때문이다. 이처럼, 제안방법의 경우 군집화를 기반으로 대표성을 가지는 데이터를 선별하고 선별된 데이터를 대상으로 패턴을 도출하여 미래부하를 예측함으로써 신뢰할만한 예측결과를 가지는 것을 확인할 수 있다.

IV. Conclusion

근래에 기계학습의 장점에 기인하여 다양한 산업현장에 기계학습을 적용하기위한 연구들이 수행되고 있다. 전력분야에서도 마찬가지로, 특히 기계학습 기반의 미래부하 예측에 초점이 맞춰진 상황이다. 그러나 기존에 수행된 대부분의 연구는 고품질의 데이터를 전제한 연구로 실제 현장에서 기계학습을 적용하는 방법과 실효성 분석에는 한계를 가진다. 본 연구에서는 실제 현장에서 운전 중인 배전선로로부터 데이터를 수집하여, 기계학습 기반으로 미래부하를 예측하고 정확도를 검증을 통해 기계학습의 효용성을 분석하였다.

미래부하 예측을 위한 다양한 기계학습 모델 중에서 노이즈나 부하절제 및 고장에 의한 일시적 부하증가에 강건한 특성을 보일 수 있는 군집화 기반의 예측방법을 적용하였으며, 주중 및 주말의 부하패턴을 반영하기 위해 요일별 대표패턴을 도출하고 예측하는 알고리즘을 적용하였다. 아울러, 2년 이상의 부하를 예측하기 위하여 연도별 부하추세를 분석 및 예측하고, 해당 결과를 요일별 대표패턴과 접목시켜 미래부하를 예측하는 하이브리드형 미래부하 예측모델을 개발하였다. 예측모델을 통하여 도출된 결과는 일반적인 정확도 분석 수식인 RMSE, MAPE, NRMSE 뿐 아니라 배전계획 측면에서 매우 중요한 최대부하의 추종성 및 예측 정확성을 판단하기 위한 RMSEmax 수식으로 검증하였다. 제안한 예측방법과 정확도 평가 수식을 대전충남본부 내 대덕유성지사의 10개 배전선로에 적용한 결과, 모든 배전선로에서 85%를 상회하는 정확도를 보이는 것을 확인할 수 있었다. 특히, 최대부하를 얼마나 잘 예측하는지를 파악할 수 있는 제안한 NRMSEmax로 평가 시, 오히려 대부분의 선로에서 NRMSE보다 더 높은 정확도를 가지는 것을 확인할 수 있고, 이는 제안방법이 최대부하값을 기반으로 의사결정이 이루어지는 배전계획 측면에서 높은 활용성을 가질 수 있음을 의미한다. 아울러 제안한 기계학습 및 군집화 기반의 배전선로 부하 예측 알고리즘을 통하여, 기존 배전계획을 위하여 담당자가 수행하던 업무를 경감시키고 인적오류로 인한 편차를 최소화하면서 자동화가 가능하다는 점은 기계학습 기법의 현장적용이라는 관점에서 큰 장점으로 작용될 것으로 기대된다.

## Acknowledgment

This research was supported by Korea Electric Power Corporation under Grant R17DA02

본 연구는 한국전력공사의 주력연구개발과제 연구비에 의해 지원되었음

## References

- [1] 한국전력, "배전설비계획 수립지침", 2014.
- [2] 한국전력, "배전선로 수요예측 절차서", 2013.
- [3] Bong-Woo Nam, Kyung-Bin Song, Kyu-Ho Kim, Jun-Min Cha, "The Spatial Electric Load Forecasting Algorithm using the Multiple Regression Analysis Method," *Journal of the Korean Institute of Illuminating and Electrical Installation Engineers*, 22(2), pp. 63-70, 2008. <https://doi.org/10.5207/JIEIE.2008.22.2.063>
- [4] Kyung-Bin Song, "The 24 Hourly Load Forecasting of the Election Day Using the Load Variation Rate," *The transactions of The Korean Institute of Electrical Engineers*, 59(6), pp.1041-1045, 2010. <https://dx.doi.org/10.5370/KIEE.2010.59.6.1041>
- [5] Kyung-Bin Song, "Various Models of Fuzzy Least-Squares Linear Regression for Load Forecasting," *Journal of the Korean Institute of Illuminating and Electrical Installation Engineers*, 21(7), pp.61-67, 2007. <https://doi.org/10.5207/JIEIE.2007.21.7.061>
- [6] Seong Seung Park, Ho Sun Sho, Dong Gyu Lee, Eun Mi Ji, Hi-seok Kim, Keun Ho Ryu, "Short-term Power Load Forecasting using Time Pattern for u-City Application," *Journal of Korea Spatial Information System Society*, 11(2), pp.177-181, 2009. <https://www.koreascience.or.kr/article/JAKO200925265918394.page>
- [7] Dong-Ha Shin, Chang-Bok Kim, "A Study on Deep Learning Input Pattern for Summer Power Demand Prediction," *The Journal of Korean Institute of Information Technology*, 14(11), pp.127-134, 2016. <https://doi.org/10.14801/jkiit.2016.14.11.127>
- [8] Jun-Young Ahn, Sang-Min Park, Chang-Bok Kim, "A Study on Neural Network Model for Winter Electric Power Demand Prediction," *The Journal of Korean Institute of Information Technology*, 15(9), pp.1-9, 2017. <https://doi.org/10.14801/jkiit.2017.15.9.1>
- [9] Jinwoong Park, Jihoon Moon, Eenjun Hwang, "An Electric Load Forecasting Scheme with High Time Resolution Based on Artificial Neural Network," *KIPS Transaction on Software and Data Engineering*, 6(11), pp.527-536, 2017. <http://10.3745/KTSDE.2017.6.11.527>
- [10] Jihoon Moon, Sungwoo Park, Eenjun Hwang, "A Multilayer Perceptron-Based Electric Load Forecasting Scheme via Effective Recovering Missing Data," *KIPS Transaction on Software and Data Engineering*, 8(2), pp.67-78, 2019. <https://doi.org/10.3745/KTSDE.2019.8.2.67>
- [11] C. Fan, F. Xiao, S. Y. Hea, and Y. Zhao, "A short-term building cooling load prediction method using deep learning algorithms," *Applied Energy*, Vol.195, pp.222-233, 2017. <https://doi.org/10.1016/j.apenergy.2017.03.064>
- [12] S. Ryu, J. Noh, and H. Kim, "Deep Neural Network Based Demand Side Short Term Load Forecasting," *Energies*, 10(1), pp.1-20, 2017. <https://doi.org/10.3390/en10010003>
- [13] Salah Bouktif, Ali Fiaz, Ali Ouni and Mohamed Adel Serhani, "Optimal Deep Learning LSTM Model for Electric Load Forecasting using Feature Selection and Genetic Algorithm: Comparison with Machine Learning Approaches," *Energies*, 11(7), 2018. <https://doi.org/10.3390/en11071636>
- [14] Yanbing Lin, Hongyuan Luo, Deyun Wang, Haixiang Guo and Kejun Zhu, "An Ensemble Model Based on Machine Learning Methods and Data Preprocessing for Short-Term Electric Load Forecasting," *Energies*, 10(8), 2017. <https://doi.org/10.3390/en10081186>
- [15] Salah Bouktif, Ali Fiaz, Ali Ouni and Mohamed Adel Serhani, "Single and Multi-Sequence Deep Learning Models for Short and Medium Term Electric Load Forecasting," *Energies*, 12(1), 2019. <https://doi.org/10.3390/en12010149>
- [16] Heng Shi, Minghao Xu, and Ran Li, "Deep Learning for Household Load Forecasting—A Novel Pooling Deep RNN," *IEEE TRANSACTIONS ON SMART GRID*, 9(5), 2018. <https://doi.org/10.1109/tsg.2017.2686012>
- [17] Mehdi Rafiei, Taher Niknam, Jamshid Aghaei, Miadreza Shafie-Khah, and João P. S. Catalão, "Probabilistic Load Forecasting Using an Improved Wavelet Neural Network Trained by Generalized Extreme Learning Machine," *IEEE TRANSACTIONS ON SMART GRID*, 9(6), 2018. <https://doi.org/10.1109/tsg.2018.2807845>
- [18] Klaus Greff, Rupesh K. Srivastava, Jan Koutník, Bas R. Steunebrink, Jürgen Schmidhuber, "LSTM: A Search Space Odyssey," *IEEE Transaction on Neural Networks and Learning Systems*, 28(10), pp.2222-2232, 2017. <https://doi.org/10.1109/tnnls.2016.2582924>
- [19] F.A. Gers, J. Schmidhuber, F. Cummins, "Learning to forget: continual prediction with LSTM," 9th International Conference on Artificial Neural Networks, pp.850-855, 1999. <https://doi.org/10.1049/cp:19991218>
- [20] Luke Tonin, (2019. Sep. 1), "Keras implementation of a sequence to sequence model for time series prediction using an encoder-decoder architecture", <https://github.com/LukeTonin/keras-seq-2-seq-signal-prediction>