

# 계층적 군집분석방법을 활용한 건물 부하의 전력수요예측

## Load Forecasting using Hierarchical Clustering Method for Building

황혜미\* · 이성희† · 박종배\*\* · 박용기\*\* · 손성용\*\*\*

(Hye-Mi Hwang · Sung-Hee Lee · Jong-Bae Park · Yong-Gi Park · Sung-Yong Son)

**Abstract** - In recent years, energy supply cases to take advantage of EMS(Energy Management System) are increasing according to high interest of energy efficiency. The important factor for essential and economical EMS operation is the supply and demand plan the hourly power demand of building load using the hierarchical clustering method of variety statistical techniques, and use the real historical data of target load. Also the estimated results of study are obtained the reliability through separate tests of validity.

**Key Words** : Load forecasting, Cluster analysis, Hierarchical clustering method, Load pattern, Energy management system (EMS)

### 1. 서론

일반적으로 전력 수요의 증가는 공급 인프라를 확충함으로써 해결할 수 있으나, 최근 국내 전력계통은 환경 및 지역적 제약 등으로 인하여 공급 설비의 추가적 확보가 어려워지고 있어 매년 증가하고 있는 전력 수요에 대한 대응 방안이 요구되고 있다. 특히, 최근의 에너지 효율향상에 대한 관심이 증대됨에 따라 공급 측면이 아닌 수요 측면에서의 수요 반응 및 수요 관리에 대한 중요성이 더욱 커지고 있다. 이와 관련하여 부하에 에너지 저장 장치(ESS, Energy storage system) 및 신재생에너지 등을 활용하여 에너지관리시스템(EMS, Energy management system)을 적용하는 사례가 늘고 있는데, 이러한 시스템에서는 관리 대상이 되는 전력 부하의 정확한 부하 사용 패턴을 파악하는 것이 중요하다. 전력 부하 예측에 관한 연구는 주로 예측 정확도를 향상시키기 위하여 통계적, 수리적 모형의 입장에서 다양한 방법으로 진행되어 왔으며, 예측 대상을 세분화하여 예측력이 떨어지는 구간에 대한 추정방법론을 제시하거나 고도의 예측 모형을 제안, 또는 예측 기법을 다양하게 활용하여 비교하고 있다. 단기 전력 수요 예측은 크게 과거 시계열적 특성을 고도화하기 위한 연구와 기상과 같은 요인들(Casual Factors)과의 관계를 규명하기 위한 연구로 나눌 수 있다[1, 11]. 과거 시계열적 특성을 활용하여 이중계절 지수평활법이나 삼중계절 방식을 활용하여 전력 수요를

예측해왔으며, ARIMA 모형을 사용하여 예측 방법의 정확도를 향상하였다[2, 3]. 또한 온도, 기간, 기온 등을 고려하거나 날씨와 같은 기상 상태를 고려한 회귀모형 등을 사용하여 국내 전력 수요를 다양한 방식으로 예측하고 있다[4, 10].

수요 예측은 평일, 주말, 특수일 및 특수 경부하 기간으로 세분화하며, 예측 대상에 대한 세부적인 운영은 특수 경부하 기간에 대한 수요 예측은 통계적인 예측방법으로 수행하기 힘들기 때문에 예년 해당기간 실적치의 평일대비 수요 저감률 및 수요경향을 분석한 후 전년대비 부하증가율과 최근 휴일의 평일대비 수요 저감률을 반영하여 예측한다. 주말이나 특수일과 같이 불연속적인 시계열 특성을 나타내는 단기 전력수요예측에서는 지수평활화법의 오차율이 높기 때문에 회귀분석법, 퍼지선형 회귀분석법, 인공신경망, 전문가 시스템 등과 같은 고도의 예측 기법들을 적용하고 있으며 오차율을 낮추기 위한 연구가 계속적으로 진행되고 있다. 또한, 전력부하의 세분화를 위해 데이터마이닝(Data Mining) 기법을 이용한 유형분류에 대한 연구도 진행되고 있는데, 미리 그룹을 정해놓고 이에 해당되는 개체들을 분류하는 k-means 알고리즘을 활용하거나 분류된 예측 대상을 의사결정나무(Decision Tree) 형태로 표현하여 비전문가의 경우도 이해가 용이하도록 운영 아이디어를 제시하기도 한다[5-9].

본 연구에서는 특정 건물부하의 전력부하예측을 위하여 데이터마이닝의 기법 중, 사전 정보가 없거나 타깃(종속변수)이 없는 경우 수행되는 자율예측(Unsupervised Prediction) 기법의 하나인 계층적 군집분석(Hierarchical Cluster Analysis)을 적용하여 예측대상(전력부하)을 도출하고자 한다. 기존 연구들은 대부분 국내 전력수요 전체를 대상으로 하기 때문에 특정 건물에 대한 전력 부하를 제대로 설명하지 못하거나 향후 예측에 대한 운영이 용이하지 않을 수 있다. 하지만 계층적 군집분석은 기존 연구와는 달리 예측대상을 사전에 정의하지 않고 자율적으로 그룹핑하도록 하여 목표 건물의 전력부하에 가장 적합한 예측 대상을 도출하는데 적절한 방법이라 할 수 있다.

† Corresponding Author : Dept. of Consulting, E3 EXPERT Inc., Korea.

E-mail : leevaluator@naver.com

\* Photovoltaic Group, Korea Institute of Energy Research, Korea.

\*\* Dept. of Electrical Engineering, Konkuk University, Korea

\*\*\* Dept. of Electrical Engineering, Gachon University, Korea

Received : May 14, 2014; Accepted : December 09, 2014

## 2. 목표부하의 전력수요 예측방법

### 2.1 연구데이터 및 수요예측절차

본 연구에서는 건물 부하의 전력수요 히스토리를 활용하여 일별, 시간대별 전력수요를 예측하였다. 이를 위하여, 우선 그림 1과 같이 전력부하를 효과적으로 예측하기 위한 예측방법론을 수립하고 수요 예측에 활용할 데이터들의 군집분석을 수행하기 위하여 사전에 수집된 자료들을 일별, 시간대별로 구분하였다.



그림 1 전력부하 예측 절차

Fig. 1 Load forecasting procedure

수요 예측을 위한 학습데이터는 2012년 1월 1일부터 2012년 12월 31일까지 총 366일에 해당되는 실제 전력부하 측정 데이터를 활용하였다. 또한, 해당 데이터 중 일별 전력부하에는 그림 2에서 보는 바와 같이 연, 월, 일, 요일 등의 기본정보와 휴일명, 휴일여부와 같은 캘린더 정보들을 포함하였으며, 시간별 전력부하의 속성은 on peak(08:00~19:00), off peak(19:00~08:00)로 구분하고 시간은 1시간 간격과 15분 간격으로 정의하였다(그림 3 참조). 이후, 본 연구에서는 일차별 부하의 특성을 도출하고 해당 부하의 수요 예측 패턴 정립을 위한 군집분석을 수행하여 일별, 시간대별 부하 예측치를 산정하였고 이를 바탕으로 부하의 연간 패턴을 예측하였다.

또한 예측된 부하 수요의 검증을 위하여 수집된 2013년 전력부하를 바탕으로 검증을 실시하였다. 이것은 2013년 1월 1일부터 2013년 10월 25일까지 측정된 총 300일 간의 전력 부하와 2012년 수요를 바탕으로 예측된 패턴에 의한 전력 수요 결과를 비교하여 정확도를 분석함으로써, 군집분석을 통한 부하 패턴 도출 결과를 검증하여 보았다.

그림 2와 그림 3은 군집분석을 통한 그룹핑 결과로서의 예측 날짜 반영한 것으로, 수요예측 패턴 정립을 위하여 필요한 일간 및 시간대별 전력부하의 작성 양식을 나타낸다. 일간 데이터의 경우, 그림 2에서 보는 바와 같이 실제 전력부하(사용량) 실적과 예측용 날짜의 실적을 비교하여 일별 정확도를 확인할 수 있다. 이 중, 해당 데이터의 예측용 날짜는 군집분석을 통하여 요일 및 휴일 여부 등을 고려하여 결정하였다. 시간대별 데이터의 경우도 마찬가지로, 그림 3에서 보는 바와 같이 실제 날짜의 데이터를

24시간 전체에 대하여 군집분석을 한 후 예측 값과의 비교를 통하여 시간대별 예측 데이터의 정확도를 확인하였다. 이와 같은 방법에서 데이터의 예측 정확도는 시간당 부하패턴이 유사한 날짜를 정확히 찾아내는 방법과 밀접한 관련을 갖는다. 따라서 본 논문에서는 여러 가지 군집분석 방법 중 초기 그룹 설정 및 해석이 용이하고 그룹 간의 연관 관계를 쉽게 파악할 수 있는 계층적 군집분석을 활용한 전력부하의 예측 결과를 제시하고자 한다.

날짜	연도	월	구분	및완전?	예측용 날짜	요일	휴일명	휴일여부	실적	예측	정확도	Check
2012년 02월 17일	2012	2	7	1	2012년 02월 16일	6		0	399	402	99%	0
2012년 02월 18일	2012	2	4	7	2012년 02월 11일	7		1	290	300	96%	0
2012년 02월 19일	2012	2	4	7	2012년 02월 12일	1		1	321	299	96%	0
2012년 02월 20일	2012	2	8	3	2012년 02월 17일	2		0	447	399	96%	0
2012년 02월 21일	2012	2	7	1	2012년 02월 20일	3		0	412	417	99%	0
2012년 02월 22일	2012	2	7	7	2012년 02월 15일	4		0	378	394	96%	0
2012년 02월 23일	2012	2	7	7	2012년 02월 16일	5		0	392	402	97%	0
2012년 02월 24일	2012	2	7	7	2012년 02월 17일	6		0	361	399	96%	1
2012년 02월 25일	2012	2	4	7	2012년 02월 18일	7		1	297	290	92%	0
2012년 02월 26일	2012	2	4	1	2012년 02월 25일	1		1	295	287	97%	0
2012년 02월 27일	2012	2	6	7	2012년 02월 20일	2		0	398	417	95%	0
2012년 02월 28일	2012	2	7	7	2012년 02월 21일	3		0	401	412	97%	0
2012년 02월 29일	2012	2	7	7	2012년 02월 22일	4		0	359	378	95%	0
2012년 03월 01일	2012	3	1	366	2011년 03월 01일	5	삼일절	1	246	284	85%	1
2012년 03월 02일	2012	3	7	7	2012년 02월 24일	6		0	331	361	91%	0
2012년 03월 03일	2012	3	4	7	2012년 02월 25일	7		1	243	297	96%	0
2012년 03월 04일	2012	3	4	7	2012년 02월 26일	1		1	243	265	91%	0
2012년 03월 05일	2012	3	6	3	2012년 03월 02일	2		0	357	331	93%	0
2012년 03월 06일	2012	3	7	1	2012년 03월 05일	3		0	382	357	96%	0
2012년 03월 07일	2012	3	7	1	2012년 03월 06일	4		0	352	362	97%	0
2012년 03월 08일	2012	3	7	1	2012년 03월 07일	5		0	382	352	97%	0
2012년 03월 09일	2012	3	7	1	2012년 03월 08일	6		0	343	362	94%	0

그림 2 일간 전력부하 데이터

Fig. 2 Daily load data

year	month	day	DT	hour	minute	별번호	전력사용량(kWh)	예측	정확도	Check	일별 CHECK	구분	On-Peak
2012	1	11	11-Jan-12	7	45	36033	458.84	455.84	100%	0	0	7	2.0#
2012	1	11	11-Jan-12	7	60	36034	489.44	489.28	100%	0	0	7	2.0#
2012	1	11	11-Jan-12	8	15	36035	504	500.04	99%	0	0	7	2.0#
2012	1	11	11-Jan-12	8	30	36036	563.76	562.86	100%	0	0	7	2.0#
2012	1	11	11-Jan-12	8	45	36037	597.42	595.08	100%	0	0	7	2.0#
2012	1	11	11-Jan-12	8	60	36038	623.7	599.58	96%	0	0	7	2.0#
2012	1	11	11-Jan-12	9	15	36039	621.72	627.66	99%	0	0	7	1.0n
2012	1	11	11-Jan-12	9	30	36040	630.72	631.62	100%	0	0	7	1.0n
2012	1	11	11-Jan-12	9	45	36041	645.84	645.56	100%	0	0	7	1.0n
2012	1	11	11-Jan-12	9	60	36042	636.66	641.16	99%	0	0	7	1.0n
2012	1	11	11-Jan-12	10	15	36043	628.22	624.6	100%	0	0	7	1.0n
2012	1	11	11-Jan-12	10	30	36044	617.76	627.12	98%	0	0	7	1.0n
2012	1	11	11-Jan-12	10	45	36045	589.5	621	95%	0	0	7	1.0n
2012	1	11	11-Jan-12	10	60	36046	589.14	610.38	96%	0	0	7	1.0n
2012	1	11	11-Jan-12	11	15	36047	587.88	589.5	100%	0	0	7	1.0n
2012	1	11	11-Jan-12	11	30	36048	575.28	580.86	99%	0	0	7	1.0n
2012	1	11	11-Jan-12	11	45	36049	589.14	583.38	99%	0	0	7	1.0n
2012	1	11	11-Jan-12	11	60	36050	577.44	562.5	97%	0	0	7	1.0n
2012	1	11	11-Jan-12	12	15	36051	528.48	542.34	97%	0	0	7	1.0n
2012	1	11	11-Jan-12	12	30	36052	525.6	547.92	96%	0	0	7	1.0n
2012	1	11	11-Jan-12	12	45	36053	541.44	551.88	98%	0	0	7	1.0n

그림 3 시간대별 전력부하 데이터

Fig. 3 Hourly load data

### 2.2 계층적 군집분석 방법

군집분석(Cluster Analysis)은 여러 개체들을 분류하는 방법론으로서, p개의 변수로 구성된 N개의 개체들에 대하여 p차원 공간에 흩어진 N개의 점으로 생각하고 이들이 어떤 의미의 조밀성을 가지고 군집을 이루고 있는지에 관한 정보를 통하여 다변량 자료의 구조를 이해하는 것이다. 즉, 군집 그룹의 개수, 내용, 구조 등이 온전히 알려지지 않은 상태에서 데이터의 특성을 파악할 수 있는 방법이다. 분석 그룹의 유형은 각 개체가 상호배반적인 여러 군집들 중 어느 하나에만 속하는 상호 배반적 그룹과 하나의 그룹이 다른 그룹에 포함은 되나 그룹간의 중복이 허용되지

않고 계층형식으로 나뉘어가지(Tree)와 같은 구조를 취하게 되는 계층적(Hierarchical) 군집, 그리고 하나의 개체가 두 개 이상의 그룹에 동시에 소속되는 상황을 허용하는 중복 군집, 마지막으로 각 개체가 어떤 그룹에 속할 확률이나 자격을 어떠한 지표로 표현하는 퍼지(Fuzzy) 군집이 있다.

군집분석은 군집의 형태와 사용되는 유사도(Similarity) 혹은 비유사도(Dissimilarity)의 척도와 연관된 다양한 방법이 존재하는데, 군집분석을 위해 고려되는 변수가 3개 이하인 경우에는 산점도 등을 활용한 목적(目測)에 의한 군집 관계 파악이 바람직하지만 변수의 수가 늘어나게 되면 이러한 방식을 사용할 수 없게 되므로 여러 가지 군집분석 방법의 특성을 적절히 활용해야 한다. 따라서 본 논문에서는 연간 전력사용 데이터의 12개월 각각에 대한 특성을 도출하기 위하여 월별로 군집분석을 수행하였다. 특히 각 월별 날짜를 개체로 간주하고 24시간의 시간 변수에 따라 시간별 날짜의 유사도에 따른 그룹을 찾기 위한 계층적(Hierarchical) 군집분석을 실시하였으며, 이때 일간 평균전력의 유사도에 따라 그룹이 분류될 수 있도록 평균연결법(Average Linkage)을 활용하였다.

평균연결법(Average Linkage Method)은 유사도를 표현하는 방식 중의 하나로서 두 군집  $U, V$  사이의 거리를 각 군집에 속하는 모든 개체들의 평균거리로 정의하여 가장 유사성이 큰 군집을 묶어 나가는 방식이다. 즉, 크기가 각각  $N_1, N_2$  인 두 군집  $U, V$ 에서 각 군집에 속해 있는 개체 하나씩을 선택하여  $N_1 \times N_2$  가지의 거리  $d_{ij}$ 의 평균을 정의한 것으로 식(1)과 같이 나타낼 수 있다.

$$d(U, V) = 1/(N_1 N_2) \sum_i \sum_j d_{ij} \quad (1)$$

여기서, 거리  $d_{ij}$ 는 Euclidean 거리, Mahalanobis 거리, Minkowski 거리 등이 있으며 두 개체의 유사성(Similarity,  $s_{ij}$ )을 이용하여 정의할 수도 있다. 유사성은 두 개체에 대한 변수들 사이의 상관계수가 사용될 수 있고 식(2)과 같이 나타낸다.

$$s_{ij} = \frac{\sum_{k=1}^p (X_{ik} - \bar{X}_i)(X_{jk} - \bar{X}_j)}{\sqrt{\sum_{k=1}^p (X_{ik} - \bar{X}_i)^2} \sqrt{\sum_{k=1}^p (X_{jk} - \bar{X}_j)^2}} \quad (2)$$

여기서,  $\bar{X}_i = \frac{1}{n_i} \sum_{k=1}^p X_{ik}$

이때 비유사성으로 표현되는 거리는 식(3)~(5)과 같이 유사성  $s_{ij}$ 를 이용하여 구할 수 있다.

$$d_{ij} = 1 - s_{ij} \quad (3)$$

$$d_{ij} = \sqrt{1 - s_{ij}} \quad (4)$$

$$d_{ij} = \sqrt{2(1 - s_{ij})} \quad (5)$$

위와 같은 방법으로 부하패턴 그룹을 분류하면, 그에 따른 타

당성의 검증이 필요하다. 따라서 본 논문에서는 앞서 예측된 결과의 타당성 검증을 위하여 각 그룹별 절대오차율(APE, Absolute Percent Error)과 실적치( $y$ ), 예측치( $\hat{x}$ ) 간의 단순회귀 모형(Simple Regression Model)의 결정계수(R-squared)를 계산하였다. 절대오차율은 0에서 1 사이의 값을 가지며 0에 가까울수록 오차가 작음을 의미한다. 또한, 결정계수는 1에 가까울수록 실적치와 예측치가 선형적으로 일치함을 나타낸다. 절대오차율과 결정계수는 식 (6) ~ 식 (7)과 같이 계산할 수 있다.

$$APE = \left| \frac{\sum_{k=1}^{n_k} \hat{y}_k}{\sum_{k=1}^{n_k} y_k} - 1 \right| \quad (6)$$

$$R^2 = 1 - \frac{\left[ \sum_{k=1}^n (y_k - \hat{y}_k)^2 \right]}{\left[ \sum_{k=1}^n (y_k - \bar{y})^2 \right]} \quad (7)$$

- 여기서,  $R^2$  : 결정계수
- $k$  : 그룹별 날짜
- $y$  : 전력부하 실적치 (kWh)
- $\hat{y}$  : 전력부하 예측치 (kWh)
- $\bar{y}$  : 전력부하 실적치의 평균 (kWh)

### 3. 부하패턴 도출 및 타당성 검증 결과

#### 3.1 군집분석 수행결과

전술하였듯이 본 논문에서는 수집된 데이터를 통하여 24시간 및 On Peak, Off Peak에 대한 월별 군집분석을 수행하여 유사도를 찾아내고, 이를 바탕으로 수요예측 패턴을 수립하여보았다. 먼저 24시간 전체에 관한 월별 군집분석을 위하여 24시간 전체를 변수로 정의하고 각 월별 날짜들의 그룹핑 방식을 분석하였다. 계층적 군집분석에서 평균연결법, 최단연결법, WARD 연결법 등 알고리즘 별로 분석을 수행하였고, 각 알고리즘별 그룹핑 결과는

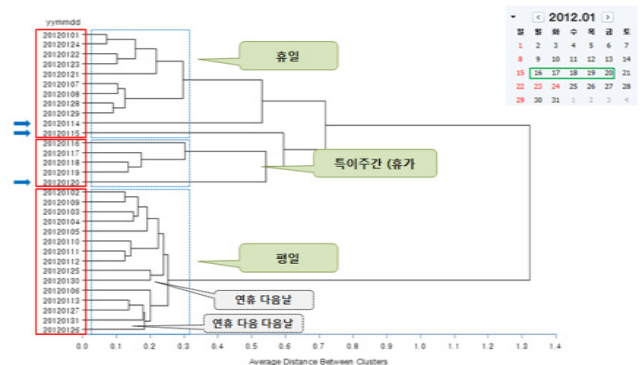


그림 4 24시간 전력부하의 군집분석 수행결과 (2012년 1월)  
Fig. 4 Cluster analysis of 24 hours power load (Jan. 2012)

평균연결법에 의한 결과와 동일하였다.

다음은 월별 24시간 군집 분석 중 2012년 1월과 2월의 군집 분석 결과이다. 분석 결과, 그림 4에서 보는 바와 같이 크게 휴일(신정연휴, 구정연휴, 토·일요일)과 평일로 구분되었으며, 그 외에도 연휴 다음날이 평일이나 휴일과는 다른 패턴을 나타냈다. 또한 캘린더 상의 녹색으로 표시된 날짜와 같이 특이성을 갖는 패턴도 확인되었다. 이러한 경우는 주로 연휴 전날(2월의 경우 24일, 29일)이라는 공통점이 있었으며, 이는 연휴 전 휴가를 사용하는 사람들이 증가하는 경향이 있으므로 이로 인하여 전력 사용량의 변화가 발생한 것으로 판단된다.

위와 같은 방식으로 1년간 월별 24시간 군집분석을 실시한 결과, 평일의 경우에도 연휴 전후의 수요 특성에 따라 그룹이 나뉘었고, 휴일은 토요일과 일요일, 국경일, 연휴 등으로 그 특성이 구분되어 나타났다. 또한 평일은 요일에 따른 특성이 상이하게 나타났으며 월요일, 화요일~목요일, 금요일이 각 하나의 그룹으로 구성되었다. 목표 부하의 특성에 따라 혹서기와 혹한기의 경우는 수요관리로 인하여 통상적인 특성과는 매우 다른 양상이 나타남을 확인할 수 있었다.

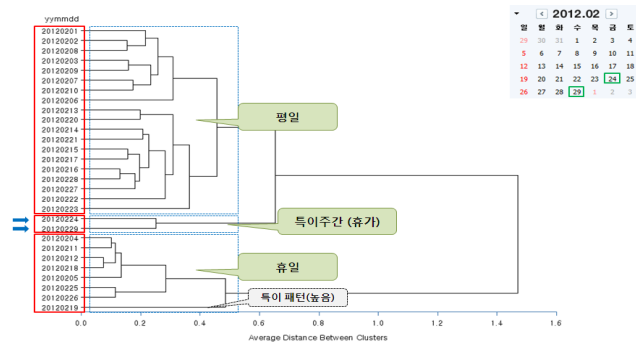


그림 5 24시간 전력부하의 군집분석 수행결과 (2012년 2월)  
Fig. 5 Cluster analysis of 24 hours power load (Feb. 2012)

본 논문에서는 24시간에 대한 군집분석 이외에 On peak와 Off peak에 관한 군집분석을 함께 수행하여 각 월별 그룹의 분류가 유사한지를 검증하여 보았다. On peak 시간대는 오전 8시부터 오후 7시 사이, Off peak는 오후 7시부터 다음날 오전 8시 사이이며, 24시간에 관한 분석과 마찬가지로 2012년 1월~12월 까지 1년간의 시간대별 전력부하 사용량 데이터를 활용하여 군집 분석을 실시하였다. 다음은 2012년 1월의 On peak/ Off peak 군집분석 결과이다. 같은 기간의 24시간 군집분석과 비교하여 그룹이 형성되는 순서는 약간씩 다르지만 패턴의 결과는 유사한 것을 확인할 수 있다.

분석 결과에서 볼 수 있듯이, 군집분석을 수행하여 나타난 부하의 기본 패턴은 크게 휴일, 주말, 평일(일반), 평일(연휴 후)로 구분할 수 있으며 이외에도 부하의 특성에 따라 충분히 발생 가능한 별도 패턴(앞의 4가지 중 어느 쪽에도 속하지 않고 정의가 불가능한 특이패턴)도 구분하여 수요예측패턴을 정의하였다.

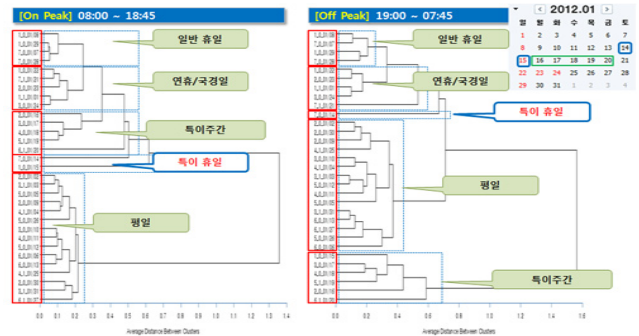


그림 6 On peak/Off peak에 대한 군집분석 수행결과 (2012년 1월)  
Fig. 6 Cluster analysis for On peak/Off peak (Jan. 2012)

### 3.2 수요예측패턴 및 예측치 산정

분석 결과를 활용하여 수요예측패턴을 정립하기 위하여, 수행된 군집분석 결과는 그림 7과 같이 의사결정 나무모형을 활용하여 나타내었다. 우선 1차적으로 휴일의 여부와 휴일의 유형에 따라서 분류한 후, 평일을 월요일과 화요일~금요일까지로 2차 분류하고 연휴 전후에 따라 수요 패턴이 다르게 나타난 특이일, 마지막으로 정의되지 않은 기타 특수일로 분류하였다. 이중 예측에 활용하는 참조수요는 1주일 전 데이터, 1일 전 데이터, 전년도 동일 데이터를 기본으로 하였으며 예외적으로 연휴 전후에 대한 예측은 일반적으로 1주일 전 부하 변동의 폭이 크므로 2주 전 데이터를 활용하였다.

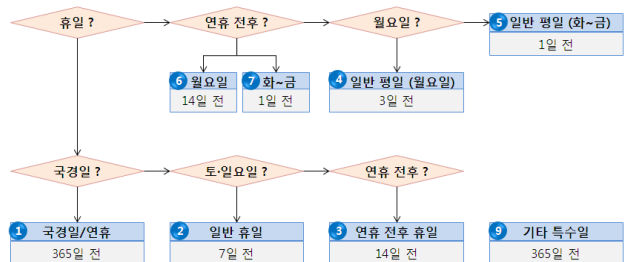


그림 7 군집분석 수행결과에 대한 의사결정나무  
Fig. 7 Decision tree for result of cluster analysis

이를 바탕으로 본 수요패턴에 따른 규칙과 전년도 대비 예측정확도는 표 1과 같이 정리되었다. 표 1에서 보는 바와 같이 분석 결과는 특수 패턴(그림 7에서 9번)을 포함하여 총 8가지 패턴으로 구분되었으며 2012년을 기준으로 예측정확도를 비교해 본 결과, 2012년 전체 366일 중에서 82%에 해당하는 300일에 대한 정확도가 90%를 초과하는 것을 확인하였다. 그 외 나머지 66일은 사무실 용도인 건물의 특성에 따라 혹서기 및 혹한기의 수요관리기간 및 건물 자체의 정비 및 점검, 또는 행사 등과 같은 이벤트에 의하여 예측 정확도가 떨어진 것으로 판단된다. 이중 그림 7의 5번에 해당하는 일반 평일(화요일~금요일)의 경우는 주로 수요관리에 의한 경부하가 발생되어 참조수요를 따라가지 못하는 결과를 나타내었고, 그림 7의 2번에 해당

표 1 수요패턴에 따른 규칙과 예측정확도

Table 1 Rules and predictive accuracy based on demand patterns

구분	번호	상세	정확도 90%이상		전체	활용자료		비고
			일수	비중				
휴일	1	국경일, 연휴	10	53%	19	1년前	-365일	명절, 샌드위치
토,일	2	일반	71	81%	88	1주前	-7일	
	3	연휴전주, 다음주	16	94%	17	2주前	-14일	
평일 (일반)	4	월요일	36	77%	47	3일前	-3일	국경일후 월요일은 1주일前
	5	화~금요일	163	86%	190	1일前	-1일	
평일 (연휴 後)	6	월요일	3	100%	3	2주前	-14일	
	7	화~금요일				1일前	-1일	
그외	9	점검일,개원일,휴가,프로 젝트마감일, 기타 특성일	2	100%	2	1년前	-365일	특정 패턴
합계			300일 (82.0%)		366일			

하는 일반 휴일(토·일요일)은 주말을 이용한 건물 자체 정비 및 점검 등으로 인하여 5번과 마찬가지로 참조수요와는 다른 양상을 나타내었다. 특히, 연휴가 집중되는 국경일 또는 휴일의 경우, 건물 내 전력 사용자들의 휴가 사용 등으로 인하여 전력부하 패턴이 매우 다르게 발생함을 확인할 수 있었다.

3.3 예측타당성 검증결과

앞서 예측된 결과의 타당성 검증을 위하여 정의된 7개 그룹과 특수 패턴(9번)에 대한 예측 결과는 전술한 바와 같이 식(6)~ 식(7)을 활용하여 절대오차율과 결정계수를 계산하여 비교하여 보았다. 산정된 결과는 표 2와 같다. 표에서 보는 바와 같이, 각 수요패턴들에 대한 전반적인 검증결과는 절대오차율(APE) 1.3%, 결정계수(R-squared) 100.0%로 전체적으로 매우 타당함을 나타내었다. 패턴 그룹별로 살펴보면 절대오차율이 5%를 초과하는 경우는 국경일 및 연휴(1번)의 경우 7.3%, 그 외 특성일(9번)에서 5.1%로 나타났다. 특히 국경일 및 연휴와 같은 특수일의 경우는 결정계수가 매우 낮게 계산된 것을 볼 수 있는데, 이는 전체적인 전력수요는 적지만 상대적으로 변동이 매우 크기 때문으로 판단된다.

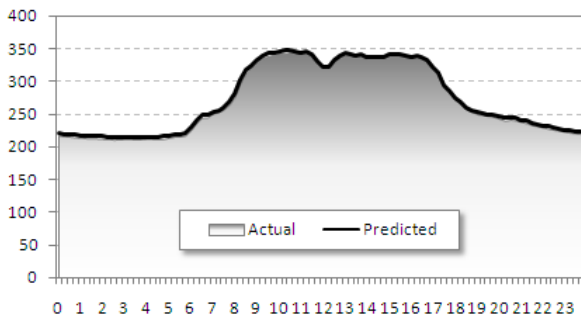


그림 8 전체 타당성 검증결과 (2012년)  
Fig. 8 Full validation results (2012)

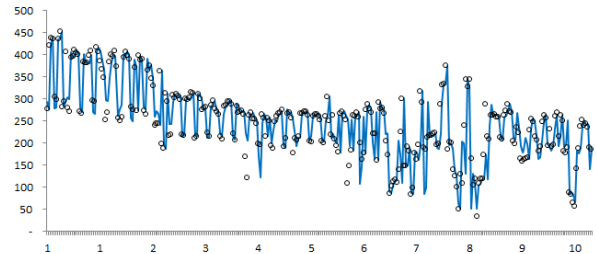


그림 9 2013년 전력부하 예측결과 (x축 날짜(월), y축 kWh)  
Fig. 9 Load forecasting results in 2013 (x-axis: month, y-axis: kWh)

같은 방식으로 전력부하 패턴에 따라 2013년 데이터에 적용하여 타당성 검증을 수행하여 보았다. 2013년 1월 1일부터 2013년 10월 25일까지 총 300일에 대해 일별 데이터를 활용하여 검증한 결과, 절대오차율(APE)은 0.7%, 결정계수는 70.9%로 나타났다. 패턴 그룹별로 절대오차율이 5% 이상인 경우는 2012년과 마찬가지로 국경일 및 연휴(1번)의 경우 8.8%, 연휴 전후 토·일요일(3번) 8.4%, 평일(7번) 5.3%였다. 2013년은 2012년과 비교하여 큰 차이를 나타내지 않은 것을 확인할 수 있었으며 절대오차율의 평균은 0.7%에 불과하여 본 논문에서 사용한 방식의 예측정확도는 매우 높은 것으로 분석되었다.

4. 결 론

본 논문에서는 목표 건물을 대상으로 전력부하의 연간 수요를 예측하기 위하여 계층적 군집분석을 활용한 전력수요 예측방법을 제시하였다. 이를 위하여 수립된 예측방법론에 따라 예측에 활용할 데이터들의 군집분석을 수행하기 위하여 수집된 자료들을 일별, 시간대별로 구분하여 정리한 후 계층적 군집분석을 수행하여 각 데이터들의 그룹핑을 통한 일별 수요특성을 도출하였다. 해당



**표 2** 수요패턴에 따른 타당성 검증 결과 (2012년)

**Table 2** Validation results based on demand patterns (2012)

구분	번호	상세	일수	APE	R-squared
휴일	1	국경일, 연휴	19	7.3%	28.0%
토,일	2	일반	88	1.0%	89.8%
	3	연휴전주, 다음주	17	2.7%	94.8%
평일 (일반)	4	월요일	47	4.6%	99.6%
	5	화~금요일	190	1.2%	99.9%
평일 (연후)	6	월요일	3	5.0%	91.5%
	7	화~금요일			
그 외	9	점검일, 개원일, 휴가, 프로젝트마감일, 기타 특성일	2	5.1%	0.4%
전체			366일	1.3%	100.0%

**표 3** 수요패턴에 따른 타당성 검증 결과 (2013년 테스트자료)

**Table 3** Validation results based on demand patterns (test data in 2013)

구분	번호	상세	일수	APE	R-squared
휴일	1	국경일, 연휴	13	8.8%	95.4%
토,일	2	일반	79	0.2%	51.1%
	3	연휴전주, 다음주	4	8.4%	48.5%
평일 (일반)	4	월요일	40	2.5%	49.0%
	5	화~금요일	159	1.4%	73.9%
평일 (연후)	6	월요일			
	7	화~금요일	5	5.3%	87.0%
그 외	9	점검일, 개원일, 휴가, 프로젝트마감일, 기타 특성일			
전체			300일	0.7%	70.9%

부하패턴은 크게 휴일 및 평일, 토요일/일요일, 연휴 후 평일 등 총 7가지로 수요패턴이 분류되었으며(부하 고유의 이벤트에 의한 비정상적 수요패턴은 기타로 분류하여 총 8가지 패턴을 적용함), 이를 통하여 2012년의 일별, 시간대별 부하수요 예측치를 산정하여 보았다.

그 결과, 2012년 366일에 대한 예측 정확도는 평균 82%로 나타났다. 이 중 300일의 정확도가 90%를 초과하였다. 그 외 정확도가 90% 이하로 나타난 66일은 군집분석 시 이상수요 패턴을 나타냈다. 이러한 특수일은 대부분 부하의 특성에 따른 것으로서, 향후 지속적인 학습에 의하여 해당 일자의 특수성이 예측 패턴에 반영되어 예측 결과의 정확도에 기여할 수 있게 된다.

또한 예측 방식의 타당성을 검증하기 위하여 예측의 절대오차율(APE) 및 단순회귀모형의 결정계수(R-squared)를 계산해 본 결과 2012년의 경우 절대오차율 1.3%, 결정계수 100.0%로 나타났다. 2013년 300일수에 대한 경우도 각각 0.7%, 80.9%로 나타남으로서 해당 예측은 매우 타당한 것으로 판단되었다. 본 연구의 결과를 통하여 얻어진 목표 건물에 대한 전력수요 예측패턴은

향후 목표 부하의 양적인 변동 및 실시간 예측오차 등을 고려한 보정에 관한 연구가 더해질 예정이며, 이를 활용하여 더욱더 정밀하고 정확한 부하수요예측을 통하여 목표 건물의 EMS 운영에 반영할 예정이다.

### 감사의 글

본 연구는 2014년 한국에너지기술연구원의 주요사업으로 수행한 연구 결과입니다(B4-2472-01).

### References

- [1] Amjady, N. (2001). Short-Term Hourly Load Forecasting Using Time-Series Modeling with Peak Load Estimation Capability, IEEE Transactions on Power Systems, 16, pp.498~505.
- [2] D.H. Hong, S.H. Lee, and H.Y. Do (2011), "Fuzzy linear regression analysis for fuzzy input-output data using shape preserving operations", Fuzzy Sets and Systems 122, pp.513-526.
- [3] Kyung-Bin Song, Young-Shik Baek, Dug-Hun Hong, Gil-Soo Jang (2005), "Short-term load forecasting for the holidays using fuzzy linear regression method", IEEE Trans. on Power Systems, Vol.20, No.1, pp.96-101.
- [4] Taylor, J. W. (2010). Triple seasonal methods for short-term electricity demand forecasting, European Journal of Operational Research, 204, pp.139~152.
- [5] B.G. Ku, C.H. Kim, J.H. Park, and H.S. Lee (2009), Daily electric load classification using data mining, KIEE summer conference 2009, pp.11-112.
- [6] J.D. Park, K.B. Song, H.W. Lim, and H.S. Park (2012), Short-Term Load Forecast for Near Consecutive Holidays Having The Mixed Load Profile Characteristics of Weekdays and Weekends, The Transactions of the Korean Institute of Electrical Engineers, 61(12), pp.1765-1773.
- [7] J.D. Park, and K.B. Song (2013), Short-Term Load Forecast for Summer Special Light-Load Period, The Transactions of the Korean Institute of Electrical Engineers, 62(4), pp. 482~488.
- [8] J.S. Lee, H.G. Son, and S.Y. Kim (2013), Daily Peak Load Forecasting for Electricity Demand by Time Series Models, The Korean Journal of Applied Statistics, 26(2), pp.349-360.
- [9] S.Y. Choi, and H.J. Kim (2007), Short-term Demand Forecasting Using Data Mining Method, Journal of Korean Institute of Illuminating and Electrical Installation Engineers. 21(10), pp.126-133.

- [10] Taylor, J. W. (2003). Short-term electricity demand forecasting using double seasonal exponential smoothing, Journal of the Operational Research Society, 54, pp.799~805.
- [11] Weron, R. (2006). Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach, Wiley, Chichester.



### 손 성 용(Sung-Yong Son)

1968년 1월 28일생. 1990년 KAIST 생산공학과 졸업. 1992년 동 대학원 정밀공학과 졸업(석사). 2000년 University of Michigan, Ann Arbor 기계공학과 졸업(박사). 현재 가천대학교 전기공학과 부교수.

## 저 자 소 개



### 황 혜 미(Hye-Mi Hwang)

1979년 7월 9일생. 2002년 강원대 전기공학과 졸업. 2004년 동 대학원 전기공학과 졸업(석사). 2013년 건국대 전기공학과 박사수료. 현재 한국에너지기술연구원 태양광연구실 선임연구원



### 이 성 희(Sung-Hee Lee)

1970년 6월 27일생. 1992년 목원대 응용통계학과 졸업. 1994년 충남대 대학원 통계학과 졸업(석사). 1999년 岡山대학 대학원 환경수리학과 졸업(박사). 현재 (주)E3 EXPERT 부사장



### 박 종 배(Jong-Bae Park)

1963년 11월 24일생. 1987년 서울대 전기공학과 졸업. 1989년 동 대학원 전기공학과 졸업(석사). 1998년 동 대학원 전기공학과 졸업(박사). 현재 건국대학교 전기공학과 교수



### 박 용 기(Yong-Gi Park)

1979년 11월 14일생. 2005년 건국대 전기공학과 졸업. 2009년 동 대학원 전기공학과 졸업(석사). 2014년 동 대학원 전기공학과 졸업(박사). 현재 건국대학교 전력시장신기술연구센터 연구원