

# 퍼지 예측 시스템을 이용한 전력 부하 예측

## Electric Power Load Forecasting using Fuzzy Prediction System

방 영 근\* · 심 재 선†  
(Young-Keun Bang · Jae-Sun Shim)

**Abstract** - Electric power is an important part in economic development. Moreover, an accurate load forecast can make a financing planning, power supply strategy and market research planned effectively. This paper used the fuzzy logic system to predict the regional electric power load. To design the fuzzy prediction system, the correlation-based clustering algorithm and TSK fuzzy model were used. Also, to improve the prediction system's capability, the moving average technique and relative increasing rate were used in the preprocessing procedure. Finally, using four regional electric power load in Taiwan, this paper verified the performance of the proposed system and demonstrated its effectiveness and usefulness.

**Key Words** : Electric power load, Prediction system, Correlation-based clustering algorithm, TSK fuzzy logic system

### 1. 서 론

전력 부하 예측은 전력 사업에 있어서 전력 수급 계획이나 전력 설비 시설 및 운용을 위한 가장 중요한 요소 중 하나이며, 전력 사업은 한 국가의 전력 공급 전략뿐만 아니라 경제 계획에도 중요한 영향을 미치게 된다[1]. 따라서 보다 정확한 예측은 전력 산업 전반뿐만 아니라 큰 범위에선 국가 경제 계획에서도 보다 효과적으로 적용될 수 있다. 전력 공급 전략이나 계획에서는 지역에 분포된 인구 밀도, 지역별 경제 성장, 지역별 기후 등도 중요한 요소가 된다. 이는 각 지역에서 소비되는 전력량이 인구밀도나 지역의 경제 개발 상태, 또는 지역별 기후 특색에 따라 각기 다르게 나타나기 때문이다. 따라서 지역적 특색에 맞게 전력 공급계획을 세우게 되면 보다 효율적인 전력 설비 운용도 가능할 것이며, 이를 통해 통합 전력 운용 망을 능동적으로 구현할 수 있을 것이다[2-4].

일반적으로 전력 부하를 예측하기 위해 많이 사용되는 방법들에는 ARMA (Autoregressive moving average model) 이나 ARIMA (Box-Jenkins model) 같은 통계적 수학 모델을 기반으로 하는 방법과 ANN (Artificial neural network), FIS (Fuzzy inference system), SVM (Support vector machines)과 같은 지능형 시스템을 기반으로 하는 방법들이 많이 사용되고 있다[5-7]. Ping-Feng과 Wei-Chang은 그들의 논문에서 SVM을 이용하여 대만 전체 전력 부하를 예측하여 좋은 성능을 보여 주었으며[5], Jagadish H는 퍼지 시스템을 이용하여 장기 부하 예측에서 매우 좋은 결과를 얻을

수 있음을 보여 주었다[6]. 또한, Jiann-Fuh은 통계적 수학 모델인 ARIMA을 단기 부하 예측을 위해 사용하였으며[2], Mohsen Hayati는 ANN 모델을 이란 서부지역의 전력 부하를 예측하기 위해 적용하였다[7]. 이러한 지능형 시스템들 중에 ANN은 비록 모델의 구조를 설명하는 것이 어렵거나 잘 제약된 조건들에서 좋은 성능을 나타내지만, 비선형 데이터들을 취급함에 있어 매우 강인한 특성을 보이며, 이러한 ANN의 단점을 보완한 SVM은 예측 모델 구현에 최근 빈번히 사용되고 있으며 좋은 성과들을 보여주고 있다. 또한 퍼지 추론 시스템의 경우에는 언어적 규칙기반을 통해 다양한 데이터들의 특성을 취급할 수 있어 많은 분야에서 그 성능을 검증 받아왔다[8-9].

한편, 전력 부하 예측 모델의 구현에서 사용되는 입력 데이터는 모델의 성능에 상당한 영향을 미친다. 논문 [1]에서는 대만을 크게 4개 지역으로 분할하고, 입력으로 각 지역별 인구 수, 최고 온도 및 GDP를 사용하였으며, 논문 [7]에서는 이란 서부 지역의 온도, 풍속 및 습도를 사용하여 전력 부하 예측에서 매우 좋은 결과를 얻었다. 반면에 논문 [5]는 과거의 수집된 전력부하 데이터를 기반으로 계절성이나 추세와 같은 통계적 기법에 SVM을 적용하여 효과적인 예측을 수행하였다.

본 논문에서는 논문 [1]의 대만 4개 지역 최대 전력 부하 예측을 수행하였다. 먼저 예측 시스템의 구현을 위해 논문 [1]에서 사용된 ANN 대신 본 논문에서는 FIS를 이용하여 설계하였으며, 시스템의 입력을 위한 데이터들은 논문 [1]에서 사용한 인구수, 최대온도, GDP가 아닌 4개 지역 중 한 지역의 전력부하를 예측하기 위해 나머지 세 지역의 전력부하 값을 입력으로 사용하였다. 이러한 방법은, 물론, 지역별 기온, 인구 수, GDP 데이터 보단 시스템 설계를 위해 적합성이 떨어진다고 볼 수 있다. 이는 지역별 인구의 증가나 온도의 변화 및 GDP는 해당 지역의 전력부하에 밀접한 관계적 특성을 나타내지만, 각 지역별 전력부하 특성은 각 지역의 개발 특성이나 기후 특색에 따라 서로 차이를 보일 수

\* Dept. of Electrical Engineering, Kangwon National University, Korea

† Corresponding Author : Dept. of Electrical Engineering, Kangwon National University, Korea

E-mail : goosung@kangwon.ac.kr

Received : September 5, 2013; Accepted : October 18, 2013

있기 때문이다. 하지만 본 논문에서는 이러한 유동적 데이터 특성을 일차로 전처리하여 통계적 특성을 안정화 시키고, 그 후 퍼지 시스템을 이용해 불명확 특성들을 기술할 수 있도록 함으로써 극복할 수 있도록 하였다. 또한 퍼지 규칙기반에 사용된 퍼지 군집화 기법에는 K-평균 기법을 개선한 상관성 기반 군집화를 수행하여 퍼지 규칙기반이 입력데이터의 특성을 잘 반영할 수 있도록 하였다. 마지막으로 시물레이션을 통해 제안된 기법이 매우 우수한 성능으로 각 지역의 전력 부하를 예측할 수 있음을 검증하였다.

## 2. 제안된 시스템의 구조

그림 1은 제안된 시스템의 전체 구조를 보여준다.

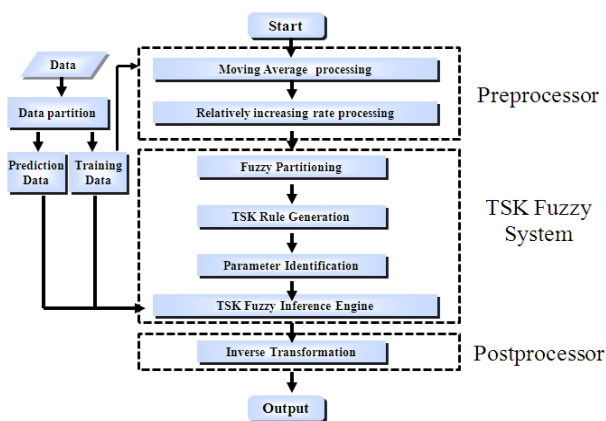


그림 1 제안된 퍼지 시스템의 구조

Fig. 1 Structure of the proposed fuzzy system

그림 1을 살펴보면 주어진 전력 데이터는 시스템의 설계를 위한 훈련데이터와 시스템 성능 검증을 위한 예측 데이터로 분할된다. 본 논문에 사용된 전력데이터의 크기는 20행 4열로 구성되어 있으며 이들 중 16행까지의 데이터들이 시스템의 설계 및 학습을 위해 사용되며, 나머지 데이터들이 시스템의 검증을 위해 사용되었다. 이러한 데이터 분할은 기존 연구된 논문 [1]에서와 같은 크기로 정의 되었다.

서론에서 설명하였듯이 시스템에 사용되는 입력데이터의 특성은 시스템의 성능에 큰 영향을 미칠 수 있다. 따라서 제안된 시스템은 전처리를 이용하여 데이터의 구조를 단순화 하고, 또한 편향적 특성을 완화하기 위해 이동평균(MA : Moving Average) 기법과 상대 증가율(RI : Relatively Increasing rate)을 적용하여 후에 설계되는 TSK 퍼지 예측 시스템의 성능이 개선될 수 있도록 하였으며 이는 3장에서 다루어진다.

퍼지 예측 시스템을 설계에는 TSK 퍼지 모델이 이용되었으며, 규칙기반의 생성을 위한 퍼지 분할에는 상관성 기반 군집화 기법을 적용하였으며, 각 규칙의 파라미터 추정에서는 구조가 단순한 최소 자승법을 적용하였다. 이들은 각각 4.1절과 4.2절에서 다루어지게 된다.

마지막으로 시스템의 출력에는 TSK 퍼지 추론과정과 가공된 데이터의 역변환 과정을 거치게 되며, 이는 4.3절에서 다루어지게 된다. 따라서 그림 1에서 보이듯 제안된 시스템

은 크게 전처리기와 TSK퍼지 시스템, 그리고 후처리로 구성되어진다. 마지막으로 그림 1의 제안된 시스템은 Matlab 프로그램을 이용하여 설계 되었으며, 시스템의 결과들은 시각화를 높이기 위해 Excel 파일로 변환 작업 되었다.

## 3. 전력 데이터의 전처리

일반 적으로 시계열 데이터(전력데이터 포함)들은 비선형적 특성을 포함하고 있으며, 서로 다른 지역에서 수집된 데이터의 경우 데이터 상호간의 상관성이 매우 낮을 수 있다. 따라서 이들을 원형 그대로 입력 데이터로 사용하게 되면 설계된 시스템이 데이터의 특성을 기술함에 있어 많은 제약점이 따를 수 있다.

특히, 본 논문에 사용된 TSK 퍼지 논리 [10]의 경우 후견항의 파라미터 추정에서 부정확성이 증가되어 예측 성능을 저해할 수 있게 된다. 또한, 주어진 시계열 데이터와 유사한 특성, 즉 인구나 GDP 등과 같이 지속적으로 증가하는 추세성을 가진 데이터의 경우 어느 시점 이후에는 규칙기반의 한쪽으로는 편중되는 현상이 발생한다. 예를 들어 2개의 군집에 대해 한쪽은 작은 수로 구성되고 다른 쪽은 큰 수로 구성되어 있다면, 지속적인 증가추세의 데이터의 경우 훈련과정 이후의 데이터들은 모두 큰 수로 구성된 군집에 포함되게 될 것이며, 따라서 작은 수로 구성된 규칙기반은 사실상 예측을 위해선 무의미해질 수 있다. 따라서 이러한 경우에는 생성된 규칙기반이 예측을 위해 효율적으로 동작하지 못하는 예가 되며, 시스템의 성능을 저하시키는 원인이 된다.

데이터의 편중 현상을 완화하기 위해 우리는 기존 논문 [11-12]에서 차분데이터를 사용함으로써 이를 극복하려고 노력한바 있다. 차분데이터의 경우 원형데이터의 분산과 같은 통계적 특성을 안정화 시킬 수 있는 장점과 규칙기반의 편중 현상을 극복할 수 있었지만, 반면에 통계적 특성의 안정화를 위한 차분 분석법과 또한 과 차분으로 야기되는 데이터 량의 결핍, 즉 차분 간격이 증가되면 될수록 데이터 량은 감소하는 단점이 발생한다. 결국, 논문 [11-12]에서는 1차 차분데이터들로만 구성된 여러 개의 모델을 병렬로 운용하여 이러한 문제점을 극복하였지만, 다중 모델 운용에 의해 구조적으로 복잡한 형태를 야기하였다.

본 논문에서는 이동평균 기법을 사용하여 일종의 필터 역할을 선행함으로써 노이즈성 데이터, 즉 기복이 심한 데이터의 특성을 완화하여 일차로 데이터의 통계적 특성을 안정화 시키고, 이 이동평균 데이터의 상대 증가율을 계산하여 입력으로 사용함으로써 데이터의 편중 현상을 극복할 수 있도록 하였다.

따라서 제안된 퍼지 시스템의 입력으로 사용되는 데이터는 원형 데이터로부터 다음과 같은 방법으로 생성된다.

예를 들어, N개의 관측된 데이터  $x_1, x_2, \dots, x_N$ 이 주어진다면, 1차 이동평균 데이터의 i번째 데이터는 다음과 같이 표현될 수 있다.

$$m_i = \frac{x_1 + x_2 + \dots + x_i}{i} = \frac{1}{i} \sum_{j=1}^i x_j, \quad i = 1, 2, 3, \dots, N \quad (1)$$

여기서  $m_i$ 는 원시계열의  $i$ 번째 이동 평균 값을 의미한다.

이렇게 1차 처리된 이동 평균 데이터들은 그들 상호간 상대 증가율에 의한 2차 전처리 과정을 거치게 되며,  $i$ 번째 상대 증가율 값은 다음과 같이 표현될 수 있다.

$$d_i = \frac{m_{i+1} - m_i}{m_i} = \frac{\frac{1}{i+1} \sum_{j=1}^{i+1} x_j - \frac{1}{i} \sum_{j=1}^i x_j}{\frac{1}{i} \sum_{j=1}^i x_j}, \quad i = 1, 2, 3, \dots, N-1 \quad (2)$$

여기서  $d_i$ 는  $i+1$ 번째 이동 평균값과  $i$ 번째 이동 평균 값 사이의 상대 증가율을 의미한다.

위와 같은 방법으로 처리된 데이터들은 결국 -1에서 1 사이의 값으로 분포될 것이고, 따라서 동일한 범위 내에서 퍼지 규칙기반이 생성될 수 있도록 할 수 있으며, 규칙기반의 편중 현상을 극복할 수 있게 된다.

#### 4. 퍼지 예측 시스템의 설계

본 논문에서는 TSK 퍼지 모델을 이용하여 예측 시스템을 구현하였으며, TSK 퍼지 예측시스템의  $r$ 번째 일반식은 다음과 같이 표현 된다[10].

$$R_r : \text{If } x_1 \text{ is } A_1 \text{ and } x_2 \text{ is } A_2 \text{ and } \dots \text{ and } x_n \text{ is } A_n \quad (3) \\ \text{Then } y = p_0 + p_1x_1 + p_2x_2 + \dots + p_nx_n$$

식 (3)에서 보이 듯 TSK 퍼지 모델은 데이터의 특성을 기술하는 언어적 규칙기반인 전건부와 각 규칙의 결과를 유도하는 후건부로 구성되어 있으며, 전건부의 규칙기반을 생성하기 위해선 퍼지분할 기법이 요구되며, 각 규칙의 국부 출력을 산출하는 후건부에서는 파라미터를 추정하기 위한 기법이 요구된다. 전건부의 퍼지 분할을 위해, 본 논문에서는 상관성 기반 군집화 기법 (Cross-correlation based K-means)을 적용하였으며[12], 후건부의 파라미터 추정에는 구조가 간단하면서도 효율적인 최소 자승법 (least square method)을 사용하였다[10]. 한 지역의 최대 전력 값을 예측하기 위해 다른 세 지역의 전력 최대값이 입력으로 사용되므로 퍼지 규칙기반을 생성하기 위한 입출력 데이터 쌍은 다음과 같이 표현할 수 있다.

$$\text{Sets} = [d_i^1, d_i^2, d_i^3], \quad \text{where } i = [1, N-1] \quad (4)$$

여기서,  $d_i^k$ 는 예측을 하기 위한 지역의 데이터 값으로 퍼지 시스템의 학습을 위한 출력 값이 되며, 나머지 값들은 같은 시기의 다른 지역의 데이터 값들을 의미한다.

따라서 TSK 퍼지 규칙의 일반식은 본 논문에서 다음과 같이 수정된다.

$$\text{Rule}(R_r) \quad (5) \\ \text{If } d_i^1 \text{ is } FS_r^1 \text{ and } d_i^2 \text{ is } FS_r^2 \text{ and } d_i^3 \text{ is } FS_r^3 \\ \text{Then } d_i^4 = c_0 + c_1d_i^1 + c_2d_i^2 + c_3d_i^3$$

여기서  $FS_r^1$ 는 첫 번째 지역에 해당하는 데이터들에 대한 퍼지 집합들 중 규칙  $R_r$ 에 해당하는 퍼지 집합을 의미하고,  $c^k$ 들은 규칙  $R_r$ 의 국부 출력을 위해 추정되어야 할 미지의 파라미터들을 의미한다.

#### 4.1 상관성 기반 군집화 기법

식 (4)에서 보이듯 하나의 입력 데이터 쌍에는 하나의 출력 값과 세 개의 입력 값으로 구성된다. 따라서 군집화 과정에서는 세 개로 구성된 입력들의 특성이 서로 유사한 데이터 쌍들을 분류하여, 퍼지 규칙기반의 생성이나 후건부의 파라미터 추정이 효과적으로 이루어질 수 있도록 한다. 군집화 기법에는 구조가 단순한 k-평균 기법이나 c-평균 기법들이 빈번하게 사용되며, k-평균 기법은 단순 유클리언 거리를 사용하는 방법으로 군집의 중심과 최소 거리에 위치한 데이터를 분류하는 구조를 가지고 있으며, c-평균 기법은 군집과 데이터사이의 소속 정도를 이용하여 군집을 수행하는 구조를 가지고 있다[10]. 본 논문에서는 k-평균 기법과 유사한 구조를 이용하지만, 단순 유클리언 거리를 사용하지 않고 군집의 중심과 데이터 입력 쌍과의 교차 상관성을 이용한 군집화를 수행하였다[12].

먼저, 임의의 군집의 중심이  $[z_u^1, z_u^2, z_u^3]$ 로 정의 되었다고 하자. 여기서  $z_u^1$ 는 첫 번째 지역에 해당하는 입력데이터에 대한  $u$ 번째 군집의 중심을 의미한다. 이렇게 군집의 중심이 정의되면, 이 중심과 각 데이터 쌍들의 교차 상관성을 검증하여 상관성이 보다 높은 군집으로 데이터를 분류하는 구조를 가지며, 이때 사용된 교차 상관함수는 다음과 같이 정의된다.

$$\rho_{DZ}^u = \frac{C_{DZ}^u}{\sqrt{C_D^u} \sqrt{C_Z^u}} \quad (6)$$

여기서,  $C_D$ 는 군집을 위한 각각의 입력데이터 쌍의 공분산이고,  $C_Z^u$ 는 각 군집의 중심  $Z_u$ 의 공분산을 의미한다. 또한,  $Z_u$ 는  $u$ 번째 군집의 중심에 해당하는 값들을 의미한다.  $C_{DZ}$ 는 군집의 중심들과 각각의 입력 데이터 쌍 상호간의 교차 공분산을 의미한다. 식 (6)에 주어진 각각의 공분산들은 다음과 같이 정의된다.

$$C_D = \sum_{k=1}^3 (d_i^k - \bar{d}_i)^2 \quad (7)$$

$$C_Z^u = \sum_{k=1}^3 (z_u^k - \bar{z}_u)^2 \quad (8)$$

$$C_{DZ}^u = \sum_{k=1}^3 (d_i^k - \bar{d}_i)(z_u^k - \bar{z}_u) \quad (9)$$

여기서,  $d_i^k$ 는  $i$ 번째 입력데이터 쌍에서  $k$ 번째 지역의 데이터 값을 의미하고,  $\bar{d}_i$ 는  $i$ 번째 입력데이터 쌍에서의 평균을 의미하며,  $z_u^k$ 는  $u$ 번째 군집의 중심값들 중에서  $k$ 번째 지역에 해당하는 값을 의미하고,  $\bar{z}_u$ 는  $u$ 번째 군집의 중심들의 평균

값을 의미한다.

이러한 교차상관함수를 이용하여 본 논문에서는 군집 중심들의 패턴과 높은 상관성을 나타내는 패턴의 입력 데이터 쌍을 각각의 군집으로 분류함으로써 군집화의 유효성을 높일 수 있도록 하였다. 각 군집들의 중심 값들은 k-평균 기법과 유사하게 다음과 같은 방법으로 갱신된다.

$$z_u^k = \frac{1}{n_u} \sum_{p=1}^{n_u} d_u^{k(p)} \quad (10)$$

여기서,  $d_u^{k(p)}$ 는  $u$ 번째 군집에 분류된 입력 데이터 쌍의  $k$ 번째 지역 값들 중  $p$ 번째 값을 의미하고,  $n_u$ 는  $u$ 번째 군집에 분류된 입력 데이터 쌍의  $k$ 번째 지역 값들의 개수를 의미한다. 이러한 군집의 중심 값들은 다음의 조건을 만족할 때까지 반복 수행되어지며, 이 역시 k-평균 기법과 유사한 구조를 나타낸다.

$$Z_u = \left| \frac{Z_{pre} - Z_{curr}}{Z_{pre}} \right| \leq 10^{-4} \quad (11)$$

여기서,  $Z_{pre}$ 는 이전의 중심 값들이며,  $Z_{curr}$ 는 현재의 갱신된 중심 값들을 의미한다. 결국 중심의 변동 폭이  $10^{-4}$  이하이면 더 이상의 변동이 없는 것으로 간주하여 중심 값의 갱신을 종료하는 것으로 간주될 수 있다.

각각의 중심들이 정의되면, 각 입력 데이터 쌍이 각 군집에 포함되는 정도를 계산할 수 있게 된다. 먼저, 군집의 개수는 수집된 데이터의 양이 지역별로 각각 20개이며 이들 중 일부를 훈련과정을 위해 사용하였으므로 본 논문에서는 각 지역별 데이터들에 대하여 각각 2개의 군집을 이용하였다. 따라서  $u$ 는 1 또는 2의 값을 가지게 되며, 각 지역별 입력데이터가 각 지역별 군집(퍼지집합)들을 만족하는 퍼지 소속 정도는 다음과 같은 가우시안 소속 함수를 이용하여 정의하였다.

$$\text{If } d_i^k \leq z_1^k \text{ or } d_i^k \geq z_2^k \quad (12)$$

$$\mu_L(d_i^k) \text{ or } \mu_R(d_i^k) = 1$$

Else

$$\mu_L(d_i^k) = \exp\left(-\frac{1}{2} \frac{d_i^k - z_1^k}{\{z_2^k - \min(D^k)\}^2}\right)$$

$$\text{or } \exp\left(-\frac{1}{2} \frac{d_i^k - z_1^k}{\{\max(D^k) - z_1^k\}^2}\right)$$

$$\mu_R(d_i^k) = \exp\left(-\frac{1}{2} \frac{d_i^k - z_2^k}{\{z_2^k - \min(D^k)\}^2}\right)$$

$$\text{or } \exp\left(-\frac{1}{2} \frac{d_i^k - z_2^k}{\{\max(D^k) - z_1^k\}^2}\right)$$

여기서,  $z_1^k$ 는  $k$ 번째 지역의 최대 전력데이터에 대한 첫 번째 군집의 중심을 의미하고,  $z_2^k$ 는 같은 지역 데이터의 두 번째 군집의 중심 값을 의미한다. 또한,  $d_i^k$ 는  $k$ 번째 지역의  $i$ 번째 데이터를 의미하고  $D^k$ 는  $k$ 번째 지역의 전체 데이터 군을 의미한다. 마지막으로  $\mu_L(d_i^k)$ 는  $d_i^k$ 가 첫 번째 군집에 포함되는 소속 정도를 의미하고  $\mu_R(d_i^k)$ 는  $d_i^k$ 가 두 번째 군집에 포

함되는 소속 정도를 의미한다.

이와 같은 방법으로 모든 입력 데이터들에 대해 소속 정도를 계산하게 된다.

### 4.2 TSK 퍼지 시스템의 규칙기반

식 (13)과 같이 세 지역에 해당하는 입력 데이터와 각 지역별 입력데이터에 대해 2개의 군집(퍼지집합)을 사용하였으므로  $i$ 번째 입력 데이터 쌍에 의해 생성될 수 있는 퍼지 규칙기반은 최대 다음과 같이 8개로 정의 될 수 있다.

$$\begin{aligned} R_1: d_i^1 \text{ is } FS_1^1 \text{ and } d_i^2 \text{ is } FS_1^2 \text{ and } d_i^3 \text{ is } FS_1^3 \\ R_2: d_i^1 \text{ is } FS_1^1 \text{ and } d_i^2 \text{ is } FS_1^2 \text{ and } d_i^3 \text{ is } FS_2^3 \\ R_3: d_i^1 \text{ is } FS_1^1 \text{ and } d_i^2 \text{ is } FS_2^2 \text{ and } d_i^3 \text{ is } FS_1^3 \\ R_4: d_i^1 \text{ is } FS_1^1 \text{ and } d_i^2 \text{ is } FS_2^2 \text{ and } d_i^3 \text{ is } FS_2^3 \\ R_5: d_i^1 \text{ is } FS_2^1 \text{ and } d_i^2 \text{ is } FS_1^2 \text{ and } d_i^3 \text{ is } FS_1^3 \\ R_6: d_i^1 \text{ is } FS_2^1 \text{ and } d_i^2 \text{ is } FS_1^2 \text{ and } d_i^3 \text{ is } FS_2^3 \\ R_7: d_i^1 \text{ is } FS_2^1 \text{ and } d_i^2 \text{ is } FS_2^2 \text{ and } d_i^3 \text{ is } FS_1^3 \\ R_8: d_i^1 \text{ is } FS_2^1 \text{ and } d_i^2 \text{ is } FS_2^2 \text{ and } d_i^3 \text{ is } FS_2^3 \end{aligned} \quad (13)$$

여기서  $FS_1^k$ 는  $k$ 번째 지역 데이터에 대하여 분할된 첫 번째 퍼지집합(군집)을 의미하고  $FS_2^k$ 는 같은 지역 데이터에 대한 두 번째 퍼지 집합을 의미한다. 모든 입력 데이터 쌍에 대해서도 시스템 운영에 사용될 수 있는 퍼지 규칙기반은 위와 같이 최대 8개로 정의 되며, 즉 제한된 예측시스템의 규칙기반은 8개의 퍼지 규칙으로 운영되는 것으로 간주될 수 있다.

규칙기반의 전건부가 생성되면, 식 (5)에서 보이듯 후건부의 파라미터를 추정해야만 TSK 퍼지 시스템을 이용할 수 있게 된다. 이러한 파라미터를 추정하기 위해선, 먼저, 어떤 하나의 규칙의 전건부를 만족하는 모든 입력 데이터 쌍을 탐색하게 된다. 식 (13)에서 알 수 있듯 하나의 입력 데이터 쌍은 최대 8개의 퍼지 규칙 생성이 가능하며, 따라서 역으로 의미하면, 하나의 규칙을 만족하는 입력 데이터 쌍이 여러 개임을 알 수 있다. 만약, 첫 번째 퍼지 규칙  $R_1$ 을 만족하는 입력 데이터쌍이  $n$ 개라면 TSK 퍼지 규칙의 후건부는 다음과 같이  $n$ 개의 연립방정식으로 표현될 수 있으며,

$$\begin{aligned} d_1^1 &= c_0^1 + c_1^1 d_1^1 + c_2^1 d_1^2 + c_3^1 d_1^3 \\ d_2^1 &= c_0^1 + c_1^1 d_2^1 + c_2^1 d_2^2 + c_3^1 d_2^3 \\ &\vdots \\ d_n^1 &= c_0^1 + c_1^1 d_n^1 + c_2^1 d_n^2 + c_3^1 d_n^3 \end{aligned} \quad (14)$$

이 연립방정식은 다음과 같이 행렬 및 벡터 형태로 표현 가능해 진다.

$$\begin{bmatrix} d_1^1 \\ d_2^1 \\ \vdots \\ d_n^1 \end{bmatrix} = \begin{bmatrix} 1 & d_1^1 & d_1^2 & d_1^3 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & d_n^1 & d_n^2 & d_n^3 \end{bmatrix} \begin{bmatrix} c_0^1 \\ c_1^1 \\ c_2^1 \\ c_3^1 \end{bmatrix} \quad (15)$$

$$Y_1 = D_1 C_1 \quad (16)$$

여기서,  $Y_1$ 는 퍼지 규칙  $R_1$ 에 대한 출력 값, 다시 말해 예측을 수행하고자 하는 데이터들 중 해당 규칙을 만족하는

데이터들을 의미하고,  $D_1$ 는 그 때의 다른 지역에 대한 입력 데이터 쌍을 의미한다. 마지막으로  $C_1$ 은 규칙 1의 후건부를 표현하기 위한 파라미터들을 의미하며, 다음과 같이 최소자승법에 의해 쉽게 추정될 수 있다.

$$\hat{C}_1 = (D_1^T D_1)^{-1} D_1^T Y_1 \quad (17)$$

식 (17)의 최소 자승법은 다음과 같이 오차의 자승을 최소화 하는 파라미터  $\hat{C}$ 를 탐색하게 된다.

$$E_1 = (Y_1 - D_1 \hat{C}_1)^T (Y_1 - D_1 \hat{C}_1) \quad (18)$$

### 4.3 시스템의 출력

TSK 퍼지 모델의 후건부의 파라미터 추정이 완료되면, 이 후 입력데이터 쌍을 이용하여 각 규칙에 대한 국부 출력을 얻을 수 있다. 앞서 서술했듯 하나의 입력 데이터 쌍은 여러 개의 퍼지 규칙을 중복으로 만족할 수 있으므로 만족되는 퍼지 규칙으로부터 얻어진 국부 출력들과 입력쌍이 각 규칙을 점화시키는 강도를 이용하여 전체 시스템의 출력을 얻게 된다.

먼저,  $i$ 번째 입력 데이터 쌍이 식 (13)의  $r$ 번째 퍼지 규칙을 만족하였다면, 이때 입력 데이터 쌍에 의한 규칙의 점화 강도  $f_r$ 은 다음과 같이 정의된다.

$$f_r = \min[\mu(a_i^L) \mu(a_i^R) \mu(b_i^R)] \quad (19)$$

여기서  $f_r$ 은  $i$ 번째 입력데이터 쌍에 의한  $r$ 번째 규칙이 점화되는 강도를 의미하며, 본 논문에서는 식 (19)와 같이 이 각 입력 데이터의 소속 정도의 최소값으로 정의하였다.

또한,  $i$ 번째 입력 데이터 쌍이 만약  $m$ 개의 퍼지 규칙을 만족하였다면, 각 규칙의 점화 강도와 각 규칙의 후건부의 국부 출력을 이용하여 시스템의 전체 출력 값은 다음과 같이 정의된다.

$$\hat{d}_i^4 = \frac{\sum_{r=1}^m f_r d_i^4}{\sum_{r=1}^m f_r} = \frac{\sum_{r=1}^m f_r (c_0 + c_1 d_i^L + c_2 d_i^R + c_3 d_i^R)}{\sum_{r=1}^m f_r} \quad (20)$$

여기서  $\hat{d}_i^4$ 는 시스템의 전체 출력 값을 의미하지만, 본 논문에서는 시스템의 학습을 위해 2절에서 원형 데이터를 전처리하여 사용하였으므로 시스템의 최종 출력은 다음과 같이 정의된다.

$$\hat{x}_{i+1}^4 = \left\{ \left(1 + \frac{1}{i}\right) (\hat{d}_i^4 + 1) - 1 \right\} \sum_{j=1}^i x_j^4 \quad (21)$$

여기서  $\hat{x}_{i+1}^4$ 는 예측하고자 하는 4번째 지역의 원형 데이터

에 대한 예측 값을 의미하며, 처음의 전처리 과정의 역으로 상대증가율  $\hat{d}_i^4$ 와 과거 데이터들의 누적 평균값들의 조합으로 구성됨을 알 수 있다.

## 5. 시뮬레이션

제안된 예측 시스템의 성능 검증을 위해 본 논문에서는 서론에서 언급하였듯 대만의 4개 지역 연간 최대 전력 부하 데이터를 이용하였다. 그림 2는 대만을 크게 4개 지역으로 분할한 모습을 보여 준다[1].

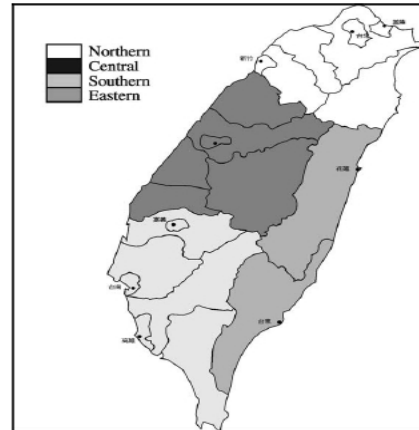


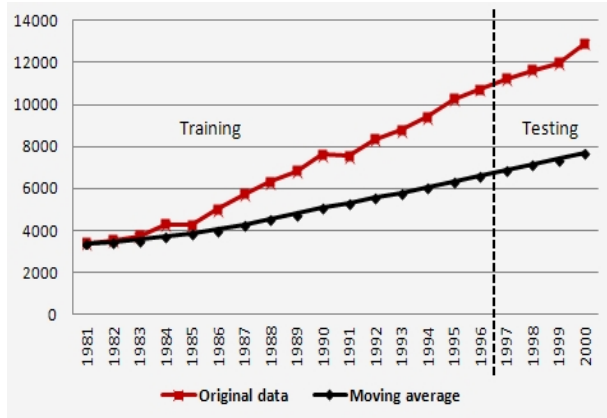
그림 2 대만의 4개 지역 분할  
Fig. 2 The four regions of Taiwan

그림 2에서 수집된 각 지역별로 20개의 데이터(1981년~2000년)들 중에서 제안된 시스템의 학습을 위해 각 지역별로 16개의 데이터(1981년~1996년)를 이용하였으며, 나머지 데이터들을 시스템의 성능 검증을 위해 이용하였다. 이러한 데이터의 분할은 기존 논문 [1]과 동일하게 하여 시스템들의 성능을 비교하기 위함이다.

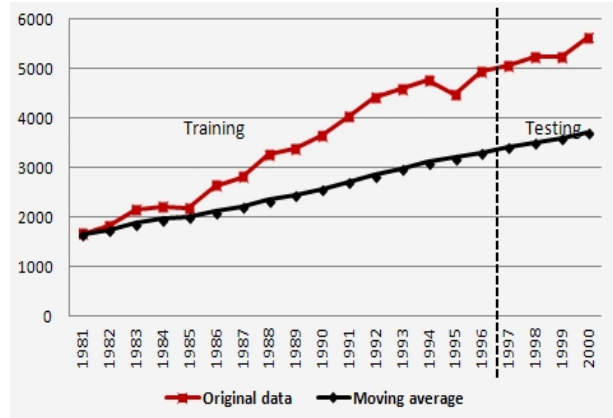
그림 3은 4개 지역의 원 시계열 데이터와 본 논문에서 사용된 전처리 과정 중 이동 평균 기법에 의한 변화된 계열의 모습을 보여준다. 그림 3을 살펴보면 이동 평균 데이터의 경우 저역 통과 필터에 의해 처리된 형태로 원형 계열의 특성 보다 더 선형 근사화 된 특성을 보이며, 이는 결국 시스템의 규칙 기반이나 파라미터 추정에 긍정적인 요소로 작용될 것임을 기대할 수 있게 한다.

그림 4는 상대 증가율에 따른 4개 지역 전력부하의 전처리된 모양을 보여준다. 그림 4를 살펴보면 3절에서 설명하였듯, 그림 3에서 보여주는 데이터의 크기에 상관없이 전체 데이터의 변화 폭이 동일 범위 내에 존재하며, 데이터의 변화폭이 뒤로 갈수록 작아지는 추세를 나타내고 있다. 결국 이러한 데이터 구조는 훈련과정에서 학습된 시스템의 규칙 기반이나 추정된 파라미터들이 이후의 예측 과정에서의 데이터 특성들도 잘 기술할 수 있을 것으로 기대할 수 있다.

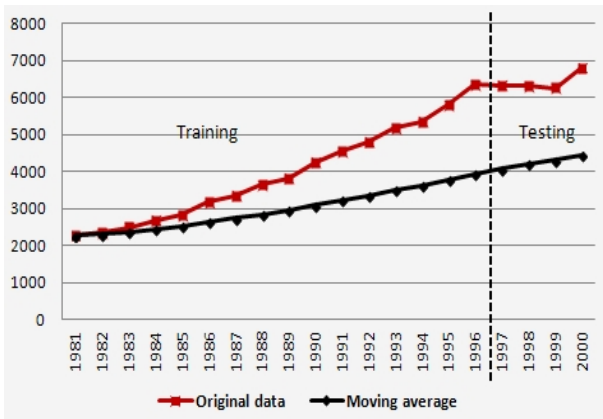
그림 5는 제안된 시스템의 예측 성능을 보여준다. 그림 5에서 빨간색은 실제 대만의 지역별 연간 최대 전력 부하를 의미하고, 검은 색은 제안된 시스템의 예측 모습을 보여준다. 제안된 시스템의 예측 값들이 4개 지역 모두에서 실제 전력부하 값을 거의 정확하게 예측하고 있음을 알 수 있다.



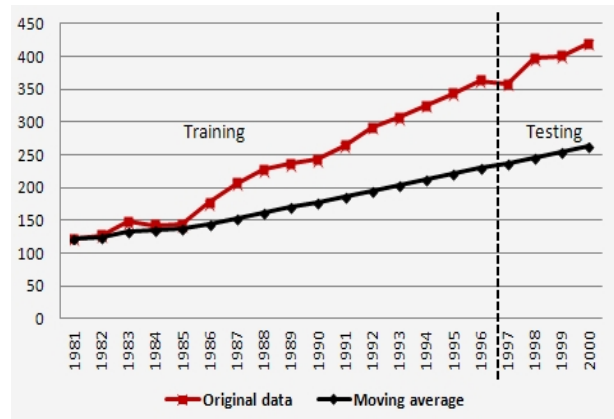
a) 북부 지역 전력 부하  
a) Electric power load of Northern region



b) 중부 지역 전력 부하  
b) Electric power load of Central region



c) 서부 지역 전력 부하  
c) Electric power load of Southern region



d) 동부 지역 전력 부하  
d) Electric power load of Eastern region

그림 3 지역별 전력부하와 이동평균 부하 (단위 : MW)  
Fig. 3 Regional electric power load and moving average (Unit : MW)

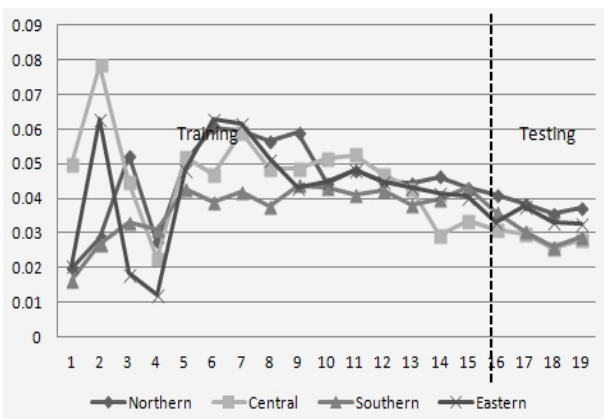
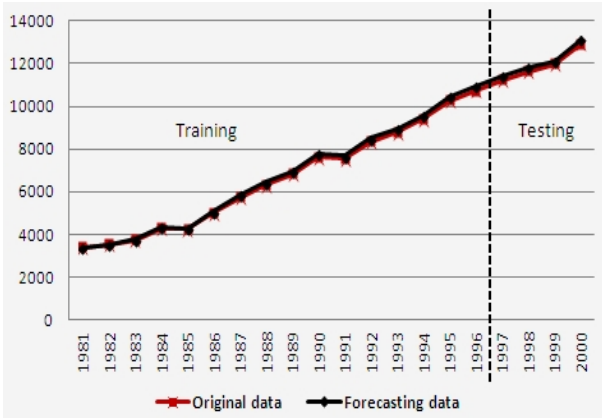


그림 4 전력부하의 상대 증가율  
Fig. 4 Relative increasing rate of the preprocessed data

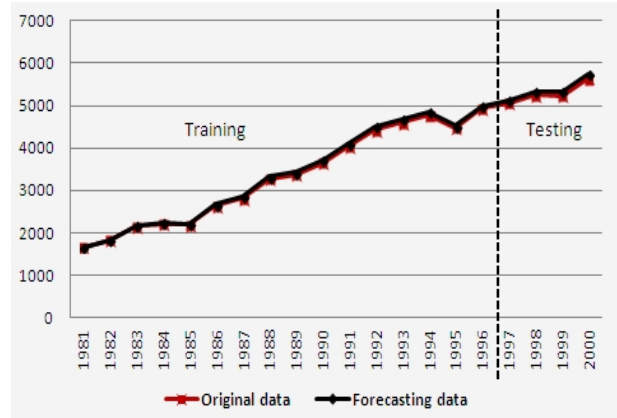
표 1은 제안된 시스템의 성능과 다른 기법들과의 성능을 비교한 것으로 성능 비교에는 동일한 검증 데이터(1997

년~2000년)를 활용하였다. 또한, 성능 평가에는 MAPE(mean absolute percentage error)를 이용하였다. 표 1에서 제안된 시스템의 MAPE가 회귀 시스템의 성능보다 매우 우수하였으며, ANN 시스템에 대해서는 북부 지역에 대한 예측 성능을 제외하고 모든 지역에서 보다 정확한 예측을 수행하였음을 알 수 있다. 특히, 각 지역의 예측 성능을 살펴보면 회귀 시스템의 경우 중부와 서부 지역의 예측 성능이 매우 나쁘게 나타났으며, ANN 시스템의 경우에는 서부와 동부 지역의 예측이 다른 지역의 예측 보다 나쁜 결과를 나타냈으나, 제안된 시스템의 경우 4개 지역 모두에 대하여 상당히 우수한 예측을 수행하였음을 알 수 있다.

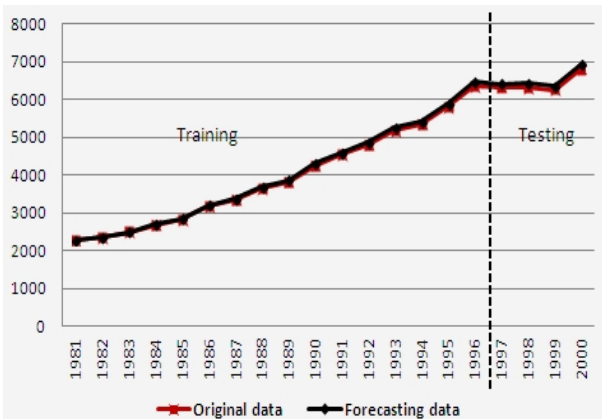
이러한 시뮬레이션 결과는 결국 제안된 시스템의 설계에 사용된 방법들이 다양한 데이터의 특성을 잘 반영할 수 있는 구조를 가지고 있다는 것으로 간주 할 수 있으며, 4개 지역의 평균 예측 성능을 살펴봐도 제안된 시스템이 다른 시스템들에 비해 매우 우수하므로 결국, 제안된 기법이 예측 시스템의 설계를 위해 매우 유효할 수 있음을 보여준다고 할 수 있다



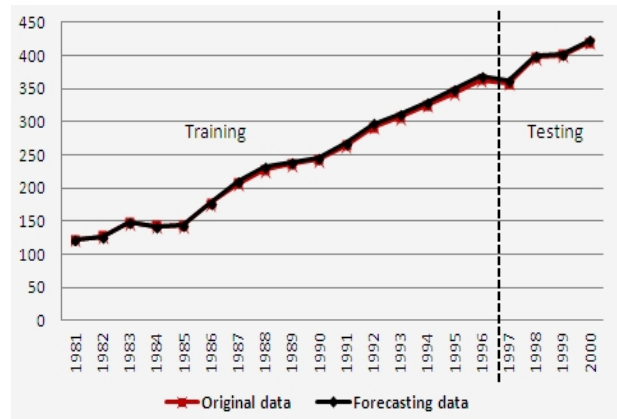
a) 북부 지역 전력 부하  
a) Electric power load of Northern region



b) 중부 지역 전력 부하  
b) Electric power load of Central region



c) 서부 지역 전력 부하  
c) Electric power load of Southern region



d) 동부 지역 전력 부하  
d) Electric power load of Eastern region

그림 5 제안된 시스템의 예측 성능 (단위 : MW)  
Fig. 5 Prediction performance of the proposed system (Unit : MW)

표 1 제안된 시스템과 다른 시스템들과의 성능 비교 (지표 : MAPE (%))  
Table 1 Performance comparison between the proposed system and other systems (Index: MAPE(%))

System \ Index	Region				average	input
	Northern	Central	Southern	Eastern		
Regression	2.42	8.52	8.29	4.10	5.83	Regional GDP, population, highest temperature
ANN	1.06	1.73	2.48	3.62	2.22	
Proposed	1.47	1.62	1.62	0.77	1.37	Other regions

## 6. 결 론

본 논문에서는 대만의 각 지역별 최대 전력부하를 예측할 수 있는 퍼지 시스템을 설계하였고, 이를 이용하여 실제 지역별 전력 부하 예측을 수행하였다. 시뮬레이션 결과에서 보여주듯 제안된 시스템의 예측 성능은 북부, 중부, 서부 및 동부 지역의 전력부하 예측에 대하여 MAPE가 각각 1.47%, 1.62%, 1.62% 및 0.77%로 매우 우수한 예측을 수행하였다. 또한, 결과를 살펴볼 때 제안된 시스템이 특정한 기록 없이 4개 지역 모두에서 우수한 예측을 수행하였음을 알 수 있으

며, 이는 결국 제안된 시스템의 설계 기법들이 전력 부하 데이터에 내재된 다양한 특성들을 충분히 수용하거나 기술할 수 있는 구조를 가지고 있는 것으로 생각할 수 있다. 마지막으로, 전체 평균 예측성능이 타 시스템들에 비해 매우 우수한 결과를 나타내었으며, 이는 제안된 시스템이 전력부하 예측에 매우 효과적으로 적용될 수 있을 것으로 간주할 수 있다. 뿐만 아니라, 제안된 시스템은 전력부하 데이터의 특성과 유사한 구조, 즉 지속적인 증가 추세의 시계열 데이터의 예측 분야에서도 매우 효과적으로 적용될 수 있을 것으로 생각된다.



제안된 시스템이 보다 큰 데이터양을 가지는 계열에 적용되기 위해선 데이터의 크기에 적합한 개수의 퍼지집합을 효과적으로 선정할 수 있는 방법이 연구되어야 할 것이며, 이를 통해 방대한 양의 데이터를 취급하는 분야에도 좋은 결과를 유도할 수 있을 것으로 생각된다.

### References

- [1] C. C. Hsu, C. Y. Chen, "Regional load forecasting in Taiwan-applications of artificial neural networks", Energy Conversion and Management, vol. 44, pp. 1941-1949, 2003.
- [2] J. F. Chen, W. M. Wang, C. M. Huang, "Analysis of an adaptive time-series autoregressive moving-average(ARMA) model for short-term load forecasting", Electric Power Systems Research, vol. 34, pp. 187-196, 1995.
- [3] A. K. Wadhwani, M. S. Wadhwani, "Short Term Load Forecasting Using Multi Parameter Regression", IJCSIS, vol. 6, no. 2, pp. 303-306, 2009.
- [4] Z. Aung, M. Toukhy, J. R. Williams, A. Sanchez, S. Herrero, "Towards Accurate Electricity Load Forecasting in Smart Grids", DBKDA 2012, pp. 51-57, 2012.
- [5] P. F. Pai, W. C. Hong, "Support vector machines with simulated annealing algorithms in electricity load forecasting", Energy Conversion and Management, vol. 46, pp. 2669-2688, 2005.
- [6] J. H. Pujar, "Fuzzy Ideology based Long Term Load Forecasting", World Academy of Science, Engineering and Technology, vol. 40, pp. 640-645, 2010.
- [7] M. Hayati, Y. Shirvany, "Artificial Neural Network Approach for Short Term Load Forecasting for Illam Region", IJECSE, vol. 1, no. 2, pp. 121-125, 2007.
- [8] H. K. Alfares, M. Nazeeruddin, "Electric load forecasting: literature survey and classification of methods", International Journal of Systems Science, vol. 33, no. 1, pp. 23-34, 2002.
- [9] E. B. Ul. Islam, "Comparison of Conventional and Modern Load Forecasting Techniques Based on Artificial Intelligence and Expert Systems", IJCSI, vol. 8, no. 3, pp. 504-513, 2011.
- [10] J. M. Mend, Uncertain Rule-based Fuzzy Logic System: Introduction and New Directions, Prentice-Hall, Upper Saddle River, Nj 07458, 2001.
- [11] Y. K. Bang, C. H. Lee, "Design of Multiple Model Fuzzy Predictors using Preprocessing and its Application", Trans. KIEE, vol. 58, no. 1, pp. 178-180, 2009.
- [12] Y. K. Bang, C. H. Lee, "Fuzzy Time Series Prediction using Hierarchical Clustering Algorithms", Expert Systems with Applications, vol. 38, pp. 4312-4325, 2011.

## 저 자 소 개



### 방 영 근 (方 寧 根)

1974년 4월 7일

2000년 삼척대학교 전자공학과(학사)

2003년 삼척대학교 전기공학과(석사)

2010년 강원대학교 전기공학과(박사)

E-mail : b2y2c1@hanmail.net



### 심 재 선 (沈 在 先)

1950년 7월 23일

1973년 성균관대학교 전기공학과(학사)

1979년 단국대학교 전기공학과(석사)

1989년 성균관대학교 전기공학과(박사)

1975년~현재 : 강원대 전기공학과 교수

E-mail : goosung@kangwon.ac.kr