

# 뉴로-퍼지 모델 기반 단기 전력 수요 예측시스템의 신뢰도 계산

論 文  
54A-10-1

## Reliability Computation of Neuro-Fuzzy Model Based Short Term Electrical Load Forecasting

沈 炫 正<sup>†</sup> · 王 普 賢<sup>\*</sup>  
(Hyun-Jeong Shim · Bo-Hyeun Wang)

**Abstract** - This paper presents a systematic method to compute a reliability measure for a short term electrical load forecasting system using neuro-fuzzy models. It has been realized that the reliability computation is essential for a load forecasting system to be applied practically. The proposed method employs a local reliability measure in order to exploit the local representation characteristic of the neuro-fuzzy models. It, hence, estimates the reliability of each fuzzy rule learned. The design procedure of the proposed short term load forecasting system is as follows: (1) construct initial structures of neuro-fuzzy models, (2) store them in the initial structure bank, (3) train the neuro-fuzzy model using an appropriate initial structure, and (4) compute load prediction and its reliability. In order to demonstrate the viability of the proposed method, we develop an one hour ahead load forecasting system by using the real load data collected during 1996 and 1997 at KEPCO. Simulation results suggest that the proposed scheme extends the applicability of the load forecasting system with the reliably computed reliability measure.

**Key Words** : 단기 전력 수요 예측, 신뢰도, 뉴로-퍼지 모델, 구조 학습

### 1. 서 론

최근 정확하고 신뢰할 수 있는 단기 수요 예측시스템을 개발하기 위해서 우수한 학습 및 근사화 능력을 갖고, 유연한 입력 선택이 가능한 신경회로망을 예측 모델에 적용하기 위한 연구가 활발하게 수행되어 왔다 [1-3]. 그러나 이들 연구에서 예측의 정확도가 매우 우수한 예측시스템을 얻을 수 있었으나, 예측 시스템의 불확실성 또는 신뢰성에 관한 고려는 상대적으로 미흡하였다. 특히, 신경회로망과 같이 수집된 데이터의 학습에만 전적으로 의존하여 구현된 경험적 모델을 실제 활용하기 위해서는 예측시스템의 정확도뿐만 아니라 예측시스템의 신뢰도에 관한 연구가 필수적이다.

이러한 이유에서 신경회로망의 출력별 신뢰도를 계산하는 연구가 활발하게 진행되고 있다. Cross-validation 개념에 기초한 Wolpert의 stacked generalization 방법과 신경회로망의 출력의 오차의 분산을 추정하는 predictive error bar 방법 등이 대표적인 예이다 [4,5]. 1992년 Leonard 등은 레이다일 베이스 함수 네트워크(RBFN)과 같이 지역 표현하는 신경회로망에서 베이스 함수에 대한 지역적 오차 특성을 이용하여 신뢰 구간을 추정하는 지역 신뢰도 측정 (local reliability measure)이라는 통계적 방법을 제안하였다 [6].

그러나, 이처럼 다양한 신경회로망의 출력별 신뢰도 계산 방법 중에서 아직까지 일반적으로 받아들여지는 표준적인 방

법이 존재하지 않기 때문에, 지금까지 제안된 신뢰도 계산의 결과를 비교하기 위한 연구가 수행되어 왔다 [7,8]. [8]에서는 stacked generalization, predictive error bar, local reliability measure 등의 세 가지 신뢰도 계산 방법을 정성적, 정량적으로 비교한 결과 Leonard가 제안한 local reliability measure가 일관되게 좋은 성능을 유지함을 밝혔다.

본 논문은 뉴로-퍼지 모델을 이용한 단기 전력 수요 예측 시스템에서 예측치 별로 신뢰 구간을 계산하는 방법을 제안한다. 기본적으로 신뢰 구간 계산이 가능한 예측시스템을 개발을 위해서 [9]에서 제안한 뉴로-퍼지 모델의 구조 학습을 이용하여 요일별 시간별 초기 구조를 생성하고, 이를 활용하여 예측 모델을 실시간 학습하는 방법을 그대로 사용한다. 개발된 뉴로-퍼지 모델 기반 예측시스템의 출력별 신뢰 구간 계산은 [8]에서 우수한 성능이 입증된 Leonard의 지역 신뢰도 측정 기법을 사용한다. 지역 표현하는 신경회로망의 특성을 활용하기 위해서 고안된 이 방법은 뉴로-퍼지 모델에서 학습된 퍼지 규칙 각각에 대해 신뢰도를 추정하고, 이를 토대로 예측 모델의 출력에 대한 신뢰 구간을 계산하기 때문에 학습된 예측 모델의 신뢰도를 규칙 단위에서 추정할 수 있다. 제안된 방법의 타당성을 검증하기 위해서 1996년부터 1997년까지 한국 전력에서 수집한 실제 부하 데이터를 이용하여 한시간 앞의 수요를 예측하는 예측시스템을 구현하고 예측치의 신뢰 구간을 계산하는 모의 실험을 수행한다.

본 논문의 구성은 다음과 같다. 2장에서는 뉴로-퍼지 모델의 구조학습과 뉴로-퍼지 모델의 출력별 신뢰 구간 계산 방법을 소개하고, 3장에서는 예측치와 예측치의 신뢰도를 동시에 계산하는 수요 예측 시스템 개발 방법을 제안한다. 4장에서는 실제 부하 데이터를 이용한 모의실험을 수행하고, 그 결과를 분석한다. 마지막으로 5장에서 결론을 제시한다.

<sup>†</sup> 교신저자, 正會員 : 팬택 & 큐리텔 전임연구원  
E-mail : shim.hyunjeong@curitel.com

<sup>\*</sup> 正會員 : 강릉대 공대 전자공학과 부교수 · 공학박사  
接受日字 : 2005年 3月 3日  
最終完了 : 2005年 9月 9日

2. 뉴로-퍼지 모델의 출력별 신뢰도 계산 방법

2.1 뉴로-퍼지 모델과 구조 학습

본 절에서는 먼저 단기 전력 수요 예측시스템의 예측 엔진을 구현하는 뉴로-퍼지 모델에 관하여 간략하게 소개한다. 뉴로-퍼지 모델은 다음과 같은 형태의 퍼지 규칙을 학습하고 추론한다:

- 규칙 1: If ( $x_1$  is  $A_1^1$ ) and ... and ( $x_n$  is  $A_1^n$ ), then  $y$  is  $w^1$ .
- ⋮
- 규칙  $p$ : If ( $x_1$  is  $A_p^1$ ) and ... and ( $x_n$  is  $A_p^n$ ), then  $y$  is  $w^p$ .

(1)

여기서  $x_j$  ( $1 \leq j \leq n$ )는 입력 변수이고,  $y$ 는 출력 변수이다.  $A_j^i$  와  $w^i$  ( $1 \leq i \leq p$ )는 소속 함수에 의해 정의되는 퍼지 집합이며, 본 논문에서는 식 (1)에서 후건부 퍼지 집합이 상수의 형태를 갖는 간단한 구조의 퍼지 규칙을 고려한다.

그림 1은 두 개의 입력 변수와 하나의 출력 변수를 갖는 뉴로-퍼지 시스템의 구조이다. 뉴로-퍼지 시스템에 임의의 입력  $x$ 가 주어지면 뉴로-퍼지 모델의 출력은 다음 식으로 계산된다:

$$F(x, w) = \frac{\sum_{i=1}^p w^i R_i(x)}{\sum_{i=1}^p R_i(x)} \quad (2)$$

여기서 Gaussian 함수로 전건부 소속 함수  $A_j^i$ 를 구현하면,  $i$  번째 규칙의 적합도  $R_i(x)$ 는 다음과 같이 구해진다:

$$R_i(x) = \prod_{j=1}^n \exp \left[ - \left( \frac{x_j - c_j^i}{\sigma_j^i} \right)^2 \right] \quad (3)$$

뉴로-퍼지 시스템의 학습은 학습 데이터를 이용하여 입력 공간을 분할하는 구조 학습과 전건부 소속 함수  $A_j^i$ 와 후건부 인자  $w^i$ 를 설정된 목적 함수를 최소화하도록 정하는 인자 학습으로 구성된다. 입력 공간 분할 문제를 해결하기 위한 다양한 방법이 제안되었는데 [10,11], 본 논문은 결정 트리를 이용하여 입력 공간을 분할하는 방법을 이용한다. 결정 트리를 이용한 입력공간 분할 방법과 모델의 인자 학습에 관한 구체적인 내용은 [9]를 참조하기 바란다.

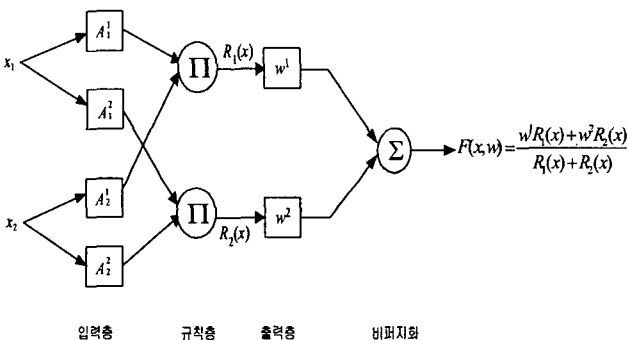


그림 1 뉴로-퍼지 시스템의 구조  
Fig. 1 Structure of neuro-fuzzy system

2.2 신뢰도의 기본 개념

신경회로망 모델의 신뢰도를 이해하기 위해서 함수 근사

화 문제를 고려하자. 함수 근사화 문제는 근사 함수  $F(x, w)$ ,  $x \in R^n$ 를 이용하여 미지의 연속 함수  $f(x)$ 를 근사화하는 것이다. 특정한 근사 함수를 선정하면, 함수 근사화 문제는 데이터를 이용하여 근사 함수의 인자  $w$ 를 학습하는 문제가 된다. 근사 함수를 학습하기 위하여 다음과 같이  $m$ 개의 데이터로 구성된 학습 집합을 고려하자:

$$L = \{(x_i, f(x_i)) | i = 1, 2, \dots, m\} \quad (4)$$

성공적인 학습은 학습 데이터에 포함되지 않은 입력  $x \in R^n - L$ 에 대하여 학습된 근사 함수  $F(x, w^* \leftarrow L)$ 가 미지의 함수  $f(x)$ 의 근사치가 됨을 의미하며, 이를 근사 함수의 일반화라 한다.

함수 근사화 문제에서 근사 함수의 신뢰도를 추정하는 가장 기본적인 간단한 방법은 근사 함수의 일반화 성능을 추정하는 것이다. 이를 위해서 학습 집합  $L$ 에 포함되지 않은 데이터로 구성된 테스트 집합  $T$ 를 수집한다

$$T = \{(x_i \in R^n - L, f(x_i)) | i = 1, 2, \dots, l\} \quad (5)$$

수집된 테스트 데이터에 대하여 식 (5)와 같이 일반화 오차를 이용한 근사 함수의 정확도를 계산한다:

$$MSE_T = \frac{1}{l} \sum_{i=1}^l \| f(x_i) - F(x_i, w^* \leftarrow L) \|^2 \quad (6)$$

테스트 집합에 대한 일반화 제곱 오차의 평균과 함께 일반화 제곱 오차의 분산을 계산하는 방법 또한 근사 함수의 신뢰도를 추정하기 위해서 일반적으로 사용하는 방법이다. 그러나 모델별로 신뢰도를 계산하는 이와 같은 방법은 특정한 입력에 대한 모델의 출력별 신뢰도를 계산할 수 없다는 단점을 갖는다.

모델의 출력별 신뢰도를 계산하기 위해서 통계학의 구간 추정 이론에 기반한 신뢰 구간을 이용 한다 [12]. 학습된 근사 함수에 입력  $x \in R^n - L$ 를 가하고, 연속 함수  $f(x)$ 가 오차 구간에 존재할 확률을  $r$ 이라 가정하자:

$$Pr[F(x, w^* \leftarrow L) - CL \leq f(x) \leq F(x, w^* \leftarrow L) + CL] = r \quad (7)$$

여기서 확률  $r$ 을 신뢰 상수라 하고, 오차 구간  $[F(x, w^* \leftarrow L) - CL, F(x, w^* \leftarrow L) + CL]$ 를  $(r \times 100)\%$  신뢰 구간이라 한다.

신뢰 구간의 추정은 주어진  $r$ 에 대하여  $CL$ 을 구하는 작업이다. 만약 신뢰 구간이 정확하게 추정되었다고 가정하면, 단일 모델에서 그림 2와 같이 입력에 따라 지역적으로 변하는  $(r \times 100)\%$  신뢰 구간의 넓이가 모델의 지역적 신뢰도를 반영한다. 본 논문에서 수행하는 모의실험은 편의상 95% 신뢰 구간을 계산한다.

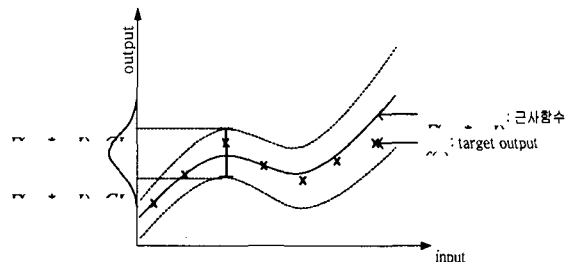


그림 2 예측 모델의 신뢰 구간  
Fig. 2 Confidence interval of a prediction model

2.3 Leonard의 지역 신뢰도 측정

1992년에 Leonard는 RBFN, 뉴로-퍼지 모델과 같이 지역 표현에 기반한 신경회로망에 적용 가능한 출력별 신뢰 구간 계산 방법을 제안하였다 [6]. 제안된 신뢰 구간 계산 방법은 각 지역에 해당하는 베이스 함수에 대한 지역 신뢰 구간을 계산하고, 임의의 테스트 입력에 대해서 해당하는 베이스 함수의 신뢰 구간의 가중 평균을 모델 출력의 신뢰 구간으로 출력한다.

그림 1의 구조를 갖는 뉴로-퍼지 모델을 식 (4)와 같이 주어진 학습 집합을 이용하여 학습하고, 모델의 오차를 구한다:

$$e(x_i) = F(x_i, w^* \leftarrow L) - f(x_i), \text{ for } x_i \in L, i = 1, 2, \dots, m \quad (8)$$

여기서 모델의 오차  $e(x_i)$ 는 Gaussian 랜덤 변수라 가정한다. 구해진  $m$ 개의 오차 샘플을 이용하여  $h$ 번째 퍼지 규칙에 대한 모델 오차의 지역 평균과 지역 분산을 추정한다:

$$\hat{\mu}_h = \frac{1}{n_h} \sum_{i=1}^m R_h(x_i) e(x_i) \quad (9)$$

$$s_h^2 = \frac{1}{(n_h - 1)} \left[ \sum_{i=1}^m R_h(x_i) e^2(x_i) \right] \quad (10)$$

여기서  $n_h$ 는  $h$  번째 규칙과 연관된 입력 데이터의 수를 의미하며 다음과 같이 계산된다:

$$n_h = \sum_{i=1}^m R_h(x_i) \quad (11)$$

$h$ 번째 규칙에 대한 모델 오차의 지역 평균을 0이라 가정하면, 식 (12)와 같이 정의된 랜덤 변수  $t$ 는 자유도가  $(n_h - 1)$ 인 Student  $t$ -분포를 따른다:

$$t(n_h - 1) \sim \frac{\hat{\mu}_h - \mu_h}{s_h / \sqrt{n_h}} = \frac{-\mu_h}{s_h / \sqrt{n_h}} \quad (12)$$

여기서  $\mu_h$ 는  $h$  번째 규칙에 대한 모델 오차의 실제 지역 평균이다.  $h$  번째 규칙에서 모델 오차의 실제 지역 평균값이 존재할 95% 신뢰 구간은  $h$  번째 규칙과 연관된 모델 오차의 샘플 평균  $\hat{\mu}_h$ 의 분산  $s_h^2/n_h$ 을 이용하여 다음과 같이 구할 수 있다:

$$CL_h = [\hat{\mu}_h - CL_h \leq \mu_h \leq \hat{\mu}_h + CL_h] \quad (13)$$

여기서  $CL_h = t_{0.95} \frac{s_h}{\sqrt{n_h}}$  이고,  $t_{0.95}$ 는 Student  $t$ -분포에서  $(n_h - 1)$  자유도를 갖는 95% 신뢰도의 결정 값이다.

일반적으로 모델 오차의 평균에 대한 신뢰 구간보다 우리는 임의의 입력에 대한 모델의 출력 오차에 대한 신뢰 구간이 필요한 경우가 많다. 이를 위해서 평균의 분산과 모델 오차의 분산의 합을 사용하여  $h$  번째 규칙에 대한 오차 한계를 계산한다 [13]:

$$CL_h' = t_{0.95} s_h \sqrt{(1 + 1/n_h)} \quad (14)$$

임의의 테스트 입력  $x$ 에 대한 모델의 오차 한계는 각 규칙에 대하여 식 (14)에서 계산한 규칙별 오차 한계를 가중 평균하여 계산된다:

$$CL(x) = \frac{\sum_{h=1}^k CL_h' R_h(x)}{\sum_{h=1}^k R_h(x)} \quad (15)$$

모델의 오차 한계를 이용하여 실제 출력  $f(x)$ 가 존재하는

95% 신뢰 구간을 다음과 같이 추정한다:

$$CI = [F(x, w^* \leftarrow L) - CL(x) \leq f(x) \leq F(x, w^* \leftarrow L) + CL(x)] \quad (16)$$

식 (15)와 식 (16)에서 알 수 있듯이 모델의 출력 오차에 대한 신뢰 구간은 각 퍼지 규칙의 신뢰 구간을 가중 평균하여 구할 수 있으며, 뉴로-퍼지 모델에 은닉층을 공유하는 별도의 출력층을 추가하여 그림 3과 같이 간단하게 구현할 수 있다.

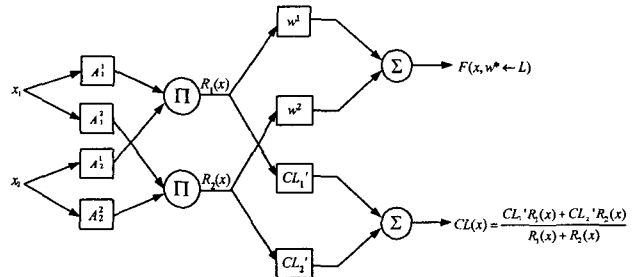


그림 3 뉴로-퍼지 모델의 출력의 신뢰도 계산

Fig. 3 Reliability computation of an output of a NF model

3. 출력별 신뢰도 계산이 가능한 단기 전력 수요 예측시스템

본 논문은 뉴로-퍼지 모델을 이용하여 한 시간 앞의 전력 수요를 예측하고, 계산된 예측치에 대한 신뢰 구간을 예측치와 함께 제공하는 단기 전력 수요 예측시스템을 제안한다. 제안된 예측시스템에서 예측치의 계산은 [9]에서 제안된 뉴로-퍼지 모델의 구조 학습을 이용한 설계 방법을 그대로 사용하고, 예측 모델의 출력별 신뢰 구간을 계산하기 위해서는 2.3절에서 설명한 Leonard의 지역 신뢰도 측정 방법을 적용한다. 그림 4는 출력별로 신뢰 구간을 계산하는 단기 전력 수요 예측시스템의 전체 구성도이다. 그림 4에서 알 수 있듯이 예측시스템을 설계하는 작업은 1) 초기 구조 बैं크의 설계, 2) 뉴로-퍼지 시스템의 학습, 3) 신뢰 구간의 계산으로 구성된다.

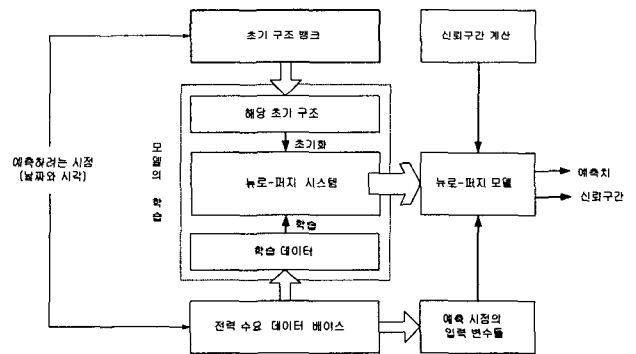


그림 4 뉴로-퍼지 기반 전력 수요 예측 시스템의 구조

Fig. 4 Architecture of neuro-fuzzy model based short-term electrical load forecasting system

초기 구조 बैं크는 뉴로-퍼지 모델의 초기화에 필요한 초기 구조를 저장한다. 일반적으로 전력 수요 패턴은 특수일을 제외하면 화, 수, 목, 금의 평일, 월요일, 토요일, 일요일과 같이 4개의 유형으로 분류된다. 초기 구조 बैं크는 4개의 요일 유형 각각에 대해서 매 시간별로 24개의 서로 다른 초기 구조를 갖기 때문에 모두 96개의 초기 구조로 구성된다.

초기 구조 बैं크에서 하나의 초기 구조를 생성하는데 필요한 학습 데이터는 해당 요일 유형 및 해당 시간을 참조하여 개별적으로 수집된다. 예를 들어, 평일 유형 01시에 대한 초기 구조를 만들기 위한 학습 데이터는 어느 특정한 한 주의 화, 수, 목, 금요일의 01시에 대한 학습 데이터를 모두 합하여 구성한다. 나머지 토, 일, 월요일의 초기 구조를 위한 학습 데이터는 연속된 특정한 2주의 데이터로부터 준비한다.

초기 구조 생성을 위한 학습 데이터가 준비되면, 그림 5의 초기 구조 생성 알고리즘을 이용하여 96개의 초기 구조를 작성한다. 초기 구조의 생성 절차는 먼저 군집의 수를 2로 하고, 결정트리를 이용하여 초기 구조를 만든다. 만약 생성된 규칙의 수가 허용 규칙의 수보다 적으면, 뉴로-퍼지 모델을 초기화하고 학습한다. 학습된 뉴로-퍼지 모델의 오차가 허용 오차보다 작을 경우 작업을 종료하고, 그렇지 않은 경우 군집의 수를 하나 늘려 초기 구조를 다시 만든다. 만약 규칙의 수가 허용 규칙의 수보다 크면, 결정트리를 pruning하여 허용 규칙의 수보다 적은 수의 퍼지 규칙을 갖는 초기 구조를 만들고, 뉴로-퍼지 모델을 학습한 후 작업을 종료한다.

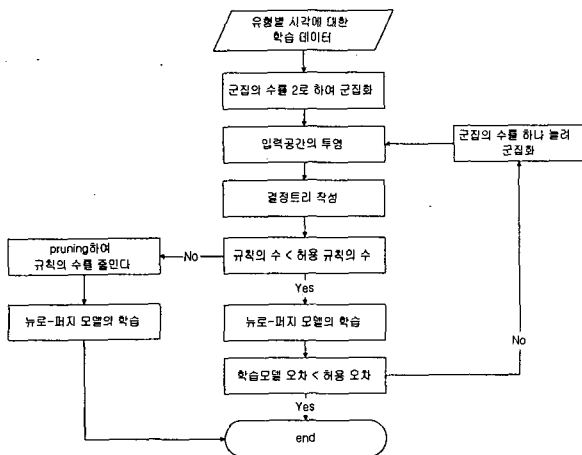


그림 5 초기 구조 생성 알고리즘  
Fig. 5 Algorithm of initial structure bank

예측시스템의 입력 변수 선정은 예측 모델의 성능을 좌우한다. 제안된 방법으로 한 시간 앞의 전력 수요를 예측하는 문제에서 [9]는 14개의 입력 변수를 사용하여 식 (18)로 정의된 MAPE이 약 1.06% 정도의 성능을 얻었고, [14]에서는 3개의 입력 변수를 사용하여 1.05% 정도의 성능을 얻었다. 본 논문에서 표 1에서 정리한 단지 두 개의 입력 변수만으로도 [9,14]와 예측의 정확도 측면에서 비교할 만한 수준의 결과를 얻기 때문에 이를 입력 변수로 선정하여 사용하였다. 표 1에서  $i$ 는 예측이 수행되는 날을,  $t$ 는 예측하는 시점을 각각 의미한다. 전력 수요 예측 문제에서 입력 변수 선정에 관한 논의는 [15]를 참조하기 바란다.

표 1 입력 변수와 출력 변수

Table 1 Input and output variables

입력 변수	$p(i, t-1), p(i-1, t)$
출력 변수	$p(i, t)$

한 시간 앞의 전력 수요를 예측하는 예측시스템을 학습하기 위해서 먼저 초기 구조 बैं크에 미리 저장된 초기 구조 중에서 예측 시점에 해당하는 초기 구조를 로드하여 뉴로-퍼지 모델을 초기화한다. 그림 6과 같이  $m$  개의 학습 데이터로 구성된 학습 집합을 이용하여 초기화된 뉴로-퍼지 모델의 인자를 학습한다. 이 과정에서 그림 1의 구조를 갖는 뉴로-퍼지 모델의 전건부 인자와 후건부 인자가 모두 학습된다.

인자 학습이 완료된 뉴로-퍼지 모델에서 학습 집합  $L$ 을 이용하여 그림 7과 같이  $m$  개의 모델 오차를 확보한다.

$$e_k = p(i, t) - F(x_k, w^* - L) \text{ for } k = 1, 2, \dots, m \quad (17)$$

여기서 뉴로-퍼지 모델의 입력  $x_k$ 는  $[p(i-1, t), p(i, t-1)]$ 이다. 확보한  $m$  개의 모델 오차를 이용하여  $h$  번째 퍼지 규칙과 연관된 입력 데이터 수  $n_h$ 와 모델 오차의 지역 분산  $s_h^2$ 을 식 (11)과 식 (10)을 이용하여 추정한다. 여기서 퍼지 규칙의 수는 예측 시점에 해당되는 초기 구조에 따라 결정된다. 추정된  $n_h$ 와  $s_h^2$ 을 이용하여  $h$  번째 퍼지 규칙에 대한 오차 한계  $CL_h'$ 을 식 (14)를 이용하여 계산한다. 계산된  $CL_h'$ 을 이용하여 임의의 입력에 대한 예측치의 신뢰 구간을 그림 3과 같은 구조를 갖는 뉴로-퍼지 모델을 이용하여 추정한다.

제안된 예측시스템에서 한 시간 앞의 전력 수요를 예측하기 위해서 매 시간 뉴로-퍼지 모델을 초기화하고, 인자 학습 및 신뢰 구간 추정 과정을 반복하여 수행한다. 이 과정을 수행하기 위해 필요한 학습 데이터는 표 1에서 언급한 2개의 입력 변수를 기초로 전력 수요 데이터 베이스에서 구성한다. 학습 데이터의 수는 임의로 정할 수 있지만 본 논문 4장의 모의 실험에서는 50개의 학습 데이터를 사용하였다. 만약 학습 데이터를 준비할 때, 데이터가 특수일에 속하게 되는 경우에는 모순되는 데이터의 학습을 피하기 위해서 한 주일 전의 데이터로 대체한다.

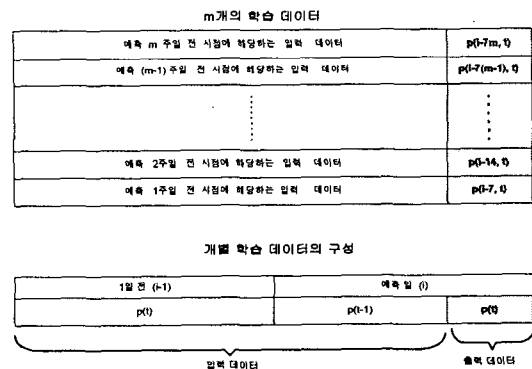


그림 6 학습 데이터의 생성 방법  
Fig. 6 Method to collect a set of training data

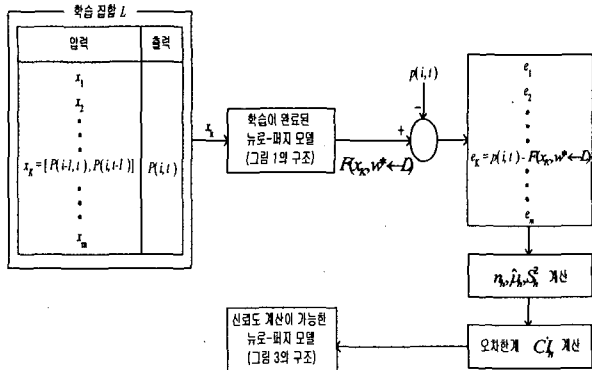


그림 7 예측시스템의 출력별 신뢰도 계산 방법  
Fig. 7 Reliability computation of an output of the system

4. 모의 실험 결과 및 분석

제안된 예측시스템의 실효성을 검증하기 위해서 97년 1월부터 97년 12월까지의 전력 수요를 예측하였다. 매 시간 뉴로-퍼지 모델을 학습하기 위해서 표 1에서 제시한 입출력 변수를 토대로 그림 6, 7과 같이 50개의 학습 데이터를 시간 별로 구성하였다. 96개의 초기 구조를 학습하는데 필요한 학습 데이터는 시간 별로 준비된 학습 집합 중에서 97년 1월 11일부터 1월 20일에 해당되는 학습 집합을 이용하여 표 2와 같이 구성하였다. 표 2에서 보여지는 것처럼 평일 유형의 초기 구조에서는 200개의 학습 데이터가 사용되었고, 나머지 유형의 초기 구조는 100개의 학습 데이터가 사용되었다.

그림 5에서 허용 규칙의 수를 10, 허용 오차를 1%로 정하고, 초기 구조 학습 데이터를 이용하여 초기 구조 뱅크를 생성하였다. 표 3은 생성한 초기 구조의 규칙 수와 결정된 초기 구조를 이용하여 뉴로-퍼지 모델을 학습한 경우 얻어진 학습 데이터에 대한 오차를 나타낸다. 평일 유형에 대한 초기 구조는 평균 3.5개의 규칙을 갖고 토요일, 일요일, 월요일 유형에 대해서는 평균 4.2, 4.2, 4.4개의 규칙을 갖는다. 모델의 성능 평가를 위해서 평균 퍼센트 오차(Mean Absolute Percentage Error, MAPE)를 사용하였다.

$$MAPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{\text{실제부하값} - \text{예측값}}{\text{실제부하값}} \right| \times 100(\%) \quad (18)$$

표 2 초기 구조 생성을 위한 학습 데이터  
Table 2 Training data for generating initial structures

유형	날짜	기간	학습데이터 수
평일 (화, 수, 목, 금)	97년 1월 14일 ~97년 1월 17일	4일	200
토요일	97년 1월 11일 97년 1월 18일	2일	100
일요일	97년 1월 12일 97년 1월 19일	2일	100
월요일	97년 1월 13일 97년 1월 20일	2일	100

예측을 수행하고자 하는 요일 유형 및 시각에 해당하는 초기 구조를 선택하여 뉴로-퍼지 모델의 구조와 인자를 초기화한다. 준비된 50개의 학습 데이터를 이용하여 초기화된 뉴로-퍼지 모델을 학습하면, 뉴로-퍼지 모델의 전진부와 후진부 인자가 결정된다. 학습이 완료된 뉴로-퍼지 모델에 50개의 학습 데이터를 인가하여 모델 오차를 확보한다. 확보된 모델 오차를 이용하여 퍼지 규칙의 오차 한계를 표현하는 인자  $CL_h$ 를 구하고 신뢰도 계산이 가능한 예측시스템의 구축을 완료한다. (그림 3 참조)

표 3 결정된 초기구조의 규칙 수와 학습 오차율  
Table 3 Number of rules and training error for the initial structure bank

유형	평일		토요일		일요일		월요일	
	규칙 수	학습 오차율	규칙 수	학습 오차율	규칙 수	학습 오차율	규칙 수	학습 오차율
1	2	0.55	4	0.90	7	0.88	4	0.73
2	2	0.50	2	0.48	5	0.52	2	0.49
3	2	0.57	4	0.57	4	0.46	2	0.64
4	2	0.45	4	0.43	4	0.36	2	0.59
5	2	0.53	2	0.51	2	0.50	5	0.48
6	6	0.47	4	0.47	4	0.46	6	0.57
7	2	0.57	4	0.57	4	0.41	4	0.65
8	2	0.81	6	0.60	8	0.85	5	0.78
9	8	0.81	4	0.88	4	0.92	8	1.02
10	2	0.96	7	1.10	4	0.80	7	0.65
11	4	0.61	2	0.67	2	0.64	2	0.56
12	4	0.42	4	0.64	2	0.53	2	0.69
13	4	0.62	2	0.82	2	0.64	6	0.78
14	2	0.73	4	0.82	2	0.65	2	0.70
15	2	0.63	4	0.73	2	0.65	2	0.60
16	2	0.53	4	0.55	4	0.57	4	0.53
17	6	0.54	3	0.77	5	0.77	4	0.40
18	8	0.76	7	0.85	7	0.94	8	0.87
19	6	0.85	6	0.94	6	1.15	6	1.00
20	2	1.11	6	0.86	9	0.98	8	0.86
21	4	0.70	7	0.69	4	0.81	6	0.83
22	4	0.41	2	0.40	4	0.55	2	0.49
23	4	0.51	4	0.64	2	0.81	2	0.53
24	2	0.63	4	0.70	2	0.82	4	0.47
평균	3.5	0.64	4.2	0.69	4.2	0.69	4.4	0.66

구축된 예측시스템을 사용하여 한 시간 앞의 부하를 예측하고, 계산된 예측치의 신뢰 구간을 추정하여 제안된 예측시스템의 성능을 분석하였다. 이를 위하여 97년 1월 21일부터 97년 12월 31일까지의 테스트 결과를 월별, 요일별, 시간별로 표 4, 5, 6에 요약하였다. 본 논문에서 개발한 예측시스템은 특수일에 대해 전혀 고려하지 않았기 때문에, 특수일에 대한 예측 결과를 제외하고 성능을 분석하였다. 예측치의 정확도 성능을 분석한 결과 MAPE의 평균은 1.12이고, MAPE의 표준 편차는 1.20정도였다. 동일한 예측시스템 설계 방법으로 3개의 입력 변수를 사용한 [14]의 결과와 비교하면, 2개의 입력 변수를 사용한 경우가 3개의 입력 변수를 사용한 경우 보다 조금 성능이 떨어지지만 그래도 비교할 만한 수준의 결과를 얻을 수 있다. 또한 표 4의 월별 분석 결과에서 알 수 있듯이 여름철에 해당하는 6월에서 9월까지의 MAPE의 평균이 특히 나빠지는 것을 알 수 있다. 이는 제안된 수요 예측시스템의 성능을 향상하기 위해서 여름철의 수요 예측에 대한 보완책이 필요함을 입증하는 결과이다.

시간별로 뉴로-퍼지 모델을 학습하고 예측치를 구함과 동시에 구해진 예측치에 대한 신뢰 구간을 추정하였다. 계산된 95% 신뢰 구간의 성능을 정량적으로 분석하기 위하여 실제 부하 값이 신뢰 구간에 포함되는 확률을 계산하였다. 이렇게 계산된 결과를 표 4, 5, 6에 요약하였다. 97년 1월 21일부터 12월 31일까지 계산된 신뢰 구간에 실제 부하가 존재하는 확률은 87.9%였다. 이는 계산된 신뢰 구간이 실제의 신뢰 구간보다 좁게 추정되었음을 의미한다.

신뢰 구간의 정량적 분석을 통해서 예측치의 정확도와 계산된 신뢰 구간이 서로 밀접하게 연관되어 있음을 알 수 있었다. 예를 들어 MAPE의 평균이 2% 수준까지 높아지는 7월과 8월에서는 신뢰 구간이 매우 부정확하게 추정되었음을 알 수 있다. 또한 6월과 9월 그리고 1월과 2월에서도 이와 비슷한 경향을 관찰 할 수 있다.

표 4 예측 오차와 신뢰도 계산 결과 (월별 분석)

Table 4 Performance analysis of the system (Monthly)

월	1월	2월	3월	4월	5월	6월
MAPE의 평균	1.17	1.01	0.67	0.69	0.71	1.45
MAPE의 표준 편차	0.77	0.90	0.61	0.61	0.66	1.26
정량적 분석 (%)	83.7	91.7	98.0	97.9	98.2	76.4
월	7월	8월	9월	10월	11월	12월
MAPE의 평균	2.00	2.13	1.16	0.79	0.80	0.80
MAPE의 표준 편차	1.91	1.72	1.15	0.72	0.72	0.70
정량적 분석 (%)	68.1	63.2	89.3	96.0	97.0	96.0

표 5 예측 오차와 신뢰도 계산 결과 (요일별 분석)

Table 5 Performance analysis of the system (Daily)

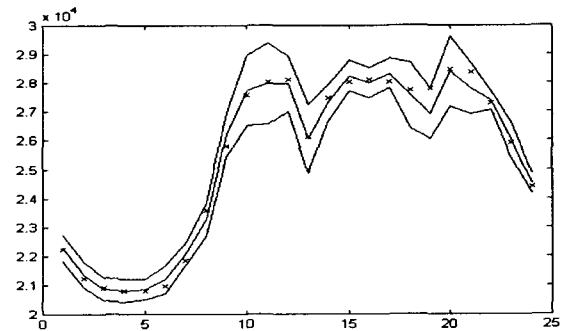
요일	월	화	수	목	금	토	일
MAPE의 평균	1.23	1.00	1.10	1.08	1.12	1.14	1.18
MAPE의 표준 편차	1.35	1.02	1.33	1.13	1.28	1.13	1.13
정량적 분석 (%)	84.5	90.8	83.8	80.3	77.5	73.4	67.4

표 6 예측 오차와 신뢰도 계산 결과 (시간별 분석)

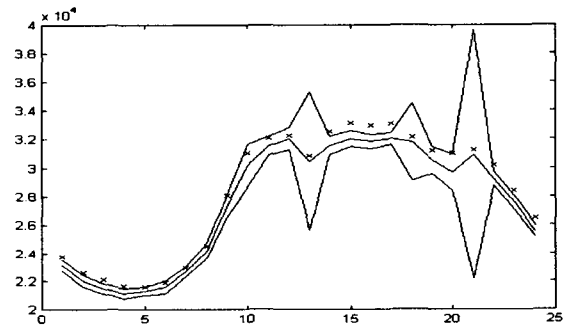
Table 6 Performance analysis of the system (Hourly)

시각	1	2	3	4	5	6	7	8
MAPE의 평균	0.91	0.88	0.82	0.78	0.71	0.86	0.79	1.17
MAPE의 표준 편차	0.86	0.86	0.88	0.89	0.77	0.77	0.71	0.90
정량적 분석 (%)	87.9	85.9	85.2	84.9	84.9	86.6	84.9	79.9
시각	9	10	11	12	13	14	15	16
MAPE의 평균	1.48	1.20	0.99	1.00	1.53	1.04	1.00	0.94
MAPE의 표준 편차	1.10	1.04	1.09	1.18	1.92	1.07	1.13	1.12
정량적 분석 (%)	90.9	92.0	91.3	90.6	95.3	86.6	84.9	83.9
시각	17	18	19	20	21	22	23	24
MAPE의 평균	1.07	1.97	1.50	1.58	1.47	0.87	1.06	0.87
MAPE의 표준 편차	1.20	2.30	1.37	1.51	1.36	1.04	1.17	0.88
정량적 분석 (%)	86.2	93.6	94.6	92.0	93.3	83.6	87.3	85.2

그림 8에서 임의로 선택된 2일간의 예측 결과와 추정된 신뢰 구간을 표시하였다. 제안된 예측시스템의 특성을 파악하기 위해서 비교적 정확한 예측이 이루어진 경우 (그림 8 (a))와 그렇지 못한 경우 (그림 8 (b))를 선택하였다. 그림 8에서 가로축은 시간을 세로축은 총 부하를 나타낸다. x로 표시된 것이 실제 수요이고, 신뢰 구간내의 실선은 예측된 수요를 의미한다. 그림 8의 (a)에서 알 수 있듯이 예측 오차가 큰 경우 신뢰 구간 또한 넓게 추정되기 때문에, 시스템 사용자가 추정된 신뢰 구간을 예측의 정확도에 대한 지표로 활용할 수 있다. 그러나 그림 8의 (b)의 경우에는 추정된 신뢰 구간이 예측치의 정확도를 잘 반영하고 있지 못하며, 결국 예측 성능에 대한 신뢰도 정보를 제공하지 못하게 된다.



(a) MAPE = 0.76% 정량적 결과 = 96% (4월 29일, 화)

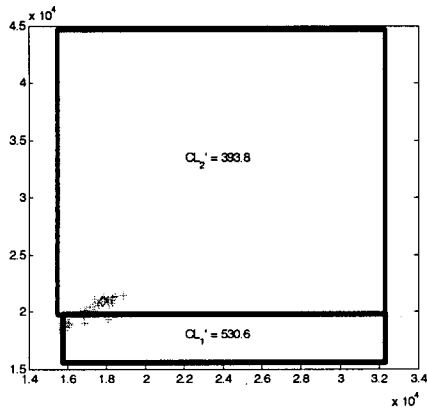


(b) MAPE = 1.63%, 정량적 결과 = 46% (6월 17일, 화)

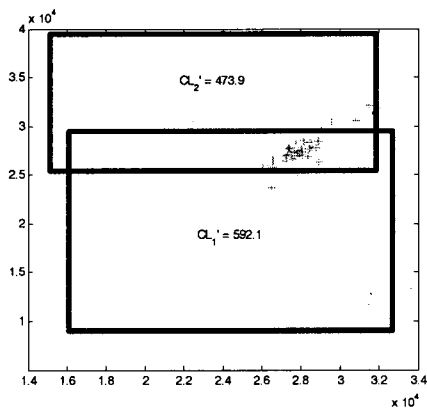
그림 8 예측치와 신뢰 구간

Fig. 8 Predicted output and confidence interval

그림 9는 정확한 예측이 수행된 4월 29일 04시와 그렇지 못한 6월 17일 15시에 대한 뉴로-퍼지 모델의 퍼지 규칙을 입력 공간에 표시한다. 퍼지 규칙은 입력 공간에서 직사각형으로 표현되고, 두께는 퍼지 규칙의 오차 한계  $CL_h$ 에 비례한다. 그림 9에서 알 수 있듯이 두 경우 모두 단지 두 개의 퍼지 규칙으로 예측을 수행하였고, 예측이 정확한 (a)의 경우 실제 수요와 학습에 사용된 데이터가 서로 근접하게 위치하나, (b)의 경우에는 실제 입력 데이터가 학습 데이터의 밀도가 큰 부분에서 벗어나 있다. 이와 같은 규칙 단위의 해석은 향후 보다 신뢰할 수 있는 예측시스템을 개발하기 위한 가이드 역할을 함과 동시에 계산된 신뢰 구간의 신뢰성을 예측시스템의 사용자가 미리 판단할 수 있는 방법을 제공할 수 있을 것이라 기대한다.



(a) MAPE=0.10,  $CL' = 405.7$  (4월 29일 화요일 04시)



(b) MAPE=4.17,  $CL' = 489.9$  (6월 17일 화요일 15시)

그림 9 학습데이터와 실제 수요 데이터의 분포와 규칙  
 ◇:실제 수요, +:m개의 학습데이터(m = 50), □:퍼지 규칙  
 Fig. 9 Distribution of training data and load data and rules

마지막으로 개발된 예측 시스템이 장기간에 걸쳐 천천히 변화하는 데이터의 특성에 효과적으로 대처할 수 있는지 확인하기 위하여 2001년과 2002년의 실제 전력 수요 데이터를 이용하여 동일한 예측 실험을 수행하였다. 표 7에서 예측 실험을 통해서 얻어진 결과를 월별로 분석하였다. 표 7과 표 4의 결과를 비교하면 1997년 MAPE의 평균이 1.12이고 2002년 MAPE의 평균은 1.33이다. 또한 신뢰구간의 정량적 분석은 1997년이 87.9%이고 2002년은 88.7%이다. 여름철에 예측 시스템의 성능이 현저하게 떨어지는 것은 공통적으로 같았으나, 2002년 11월과 12월의 예측 결과가 1997년의 11월과 12월보다 매우 좋지 않았다. 예측 시스템의 실용화를 위하여 온도 변수가 필요함을 설명하고 있다고 이해된다. 결론적으로 말하자면 11월과 12월의 예측을 제외하면 1997년의 예측 결과와 2002년의 예측 결과는 거의 동일한 수준이고, 이를 통하여 본 논문에서 제안된 바와 같이 예측시점에서 매번 재학습을 수행하고, 보다 장기간에 걸쳐 천천히 변화하는 데이터의 특징에 대처하기 위하여 1년에 한번씩 초기 구조 뱅크를 갱신하는 재학습 방법이 매우 효과적이라 판단된다.

표 7 예측시스템의 성능 (월별 분석, 2001년/2002년)

Table 7 Performance of the prediction system (Monthly, Year 2001 and Year 2002 data)

월	1월	2월	3월	4월	5월	6월
MAPE의 평균	0.91	0.86	0.89	0.92	0.85	1.48
MAPE의 표준 편차	0.58	0.62	0.86	0.83	0.65	1.74
정량적 분석 (%)	94.6	98.4	96.1	94.6	96.2	86.6
월	7월	8월	9월	10월	11월	12월
MAPE의 평균	2.17	2.11	1.31	0.81	1.58	2.09
MAPE의 표준 편차	2.05	2.03	1.21	0.72	1.19	1.80
정량적 분석 (%)	73.4	78.9	93.1	99.0	79.4	74.5

### 5. 결 론

본 논문은 예측시스템의 효율적 운용을 위해서는 신뢰도 계산이 필수적이라는 인식 하에 뉴로-퍼지 모델 기반 단기 전력 수요 예측시스템에서 예측치별로 신뢰도를 계산하는 체계적인 방법을 제안하였다. 이를 위해 먼저 뉴로-퍼지 모델의 구조 학습을 설명하고, 모델의 신뢰도를 나타내는 신뢰구간의 개념과 출력별 신뢰 구간 계산 방법을 설명하였다. 출력별 신뢰 구간 계산 방법은 지역 표현하는 뉴로-퍼지 모델의 특성을 활용하여 학습된 퍼지 규칙 각각에 대해 신뢰도를 추정하는 지역 신뢰도 측정 기법을 사용하였다.

제안된 모델의 실효성을 검증하기 위해 한국 전력에서 1996년부터 1997년까지 수집한 실제 부하 데이터를 이용하여 한 시간 앞의 총 부하와 신뢰 구간을 추정하는 모의실험을 수행하였다. 예측치의 정확도와 추정된 신뢰 구간을 분석하여 제안된 예측 모델의 성능을 평가하였다. 모의 실험 결과 예측 성능은 MAPE의 평균이 1.12, MAPE의 표준 편차의 평균이 1.20이었고, 신뢰 구간의 정량적 분석 결과 실제 수요가 추정된 신뢰 구간에 존재할 확률은 87.9%이었다. 특히 여름철에 해당되는 6월부터 9월까지 상대적으로 매우 높은 MAPE를 얻었는데, 향후 온도와 같은 날씨 변수를 사용하여 단기 전력 수요 예측시스템의 성능을 향상시킬 수 있을 것이라 생각된다.

단기 전력 수요 예측시스템의 운용 측면에서 제안된 시스템은 예측치 뿐만 아니라 예측치에 대한 신뢰 구간을 함께 제공하기 때문에, 예측시스템의 사용자가 예측이 수행된 시점에서 예측치가 존재하는 범위를 미리 추정할 수 있게 된다. 만약 추정된 신뢰 구간이 매우 넓다면, 계산된 예측치를 신뢰 할 수 없다는 경고로 활용할 수 있다. 이를 위해서 향후 예측시스템의 출력별 신뢰 구간과 함께 예측의 정확도와 신뢰 구간 사이의 상관 관계를 이용하여 예측 결과에 대한 불확실성을 알릴 수 있는 효과적인 방법이 연구되어야 할 것이다.

향후 연구 과제로써 본 논문에서는 다루지 않았지만 특수일과 계절 (온도) 요인에 대한 연구를 수행할 예정이다. 특수일에 대한 예측과 계절 요인을 고려하여 예측하는 것은 단기 전력 수요 예측시스템의 실용화에 필수적이다. 이러한 이유에서 최근까지도 이에 관한 다양한 연구가 진행되어 왔다 [16,17]. 본 논문에서 제안된 예측 시스템은 기본적으로 모델

화가 용이하게 개발되었기 때문에 예를 들어 계절별 초기 구조 또는 특수일별 초기 구조를 설계하여 예측 모듈에 추가하는 방법이 이 단계에서 고려 될 수 있을 것이다.

**감사의 글**

본 연구가 가능하도록 전력 수요 데이터를 제공해 준 한국 전력에 감사드립니다.

**참 고 문 헌**

- [1] D. C. Park, M. El-Sharkawi, R. Marks, A. Atlas, and M. Damborg, "Electrical Load Forecasting Using an Artificial Neural Network," IEEE Trans. on Power Systems, Vol. 6, No. 2, pp. 442-449, May 1991.
- [2] A. G. Bakirtzis, J. B. Theocharis, S. J. Kiartzis, and K. J. Satsios, "Short Term Load Forecasting Using Fuzzy Neural Networks," IEEE Trans. on Power Systems, Vol. 10, No. 3, pp. 1518-1524, Aug. 1995.
- [3] T. Senjyu, H. Takara, K. Uezato, and T. Funabashi, "One Hour Ahead Load Forecasting Using Neural Network," IEEE Trans. on Power Systems, Vol. 17, No. 1, pp. 113-118, Feb. 2002.
- [4] D. H. Wolpert, "Stacked Generalization," Neural Network 5, pp. 241-259, 1992.
- [5] C. Satchwell, "Finding error bars (the easy way)", Neural Computing Applications Forum, Edition 5, 1994.
- [6] J. A. Leonard, M. A. Karner and L. H. Ungar, "A Neural Network Achitecture that Compute Its Own Reliability," Computer Chem. Engng. Vol. 16, No. 9 pp. 819-835, 1992.
- [7] R. Tibshirani, "A Comparison of Some Error Estimates for Neural Network Models," Neural computation 8, pp. 152-163, 1996.
- [8] 심현정, 박래정, 왕보현, "뉴로-퍼지 모델의 신뢰도 계산: 비교 연구," 한국 퍼지 및 지능 시스템 학회 논문지, Vol. 11, No. 4, pp. 293-301, 2001.
- [9] 박영진, 심현정, 왕보현, "뉴로-퍼지 모델을 이용한 단기 전력 수요 예측 시스템," 대한전기학회 논문집, 49A권 3호, Mar. 2000.
- [10] M. Kubat, "Decision Trees can Initialize Radial Basis Function Networks," IEEE Trans. Neural Networks, Vol. 9, No. 5, pp. 813-821, Sept. 1998.
- [11] H. Pomares, I. Rojas, J. Gonzalez, and A. Prieto, "Structure Identification In Complete Rule-based Fuzzy Systems," IEEE Trans. Fuzzy Systems, Vol. 10, No. 3, pp. 349-359, Jun. 2002.
- [12] H. Stark and J. W. Wood, "Probability, Random Process and Estimation Theory for Engineers," pp. 270-288, 1994.
- [13] N. R. Draper and H. Smith, Applied Regression Analysis, 2nd Eds., John Wiley & Sons, Inc., 1966.
- [14] 박영진, 심현정, 왕보현, "뉴로-퍼지 모델을 이용한 단기 전력 수요 예측 시스템," 대한전기학회 하계학술대회 논문집. A권, pp. 102-106, July 2000.
- [15] I. Drezga and S. Rahman, "Input Variable Selection for ANN-based Short-term Load Forecasting," IEEE Trans. on Power Systems, Vol. 13, No. 4, pp. 1238-1244, Nov. 1998.
- [16] 김광호, "특수일 전력 수요 예측을 위한 퍼지 전문가 시스템," 전기학회논문지, 47권 7호, pp. 886-891, 1998.
- [17] 구본석, 백영식, 송경빈, "추석과 설날 연휴에 대한 전력수요예측 알고리즘 개선," 전기학회논문지, 51권 10호, pp. 453-459, 2002.

**저 자 소 개**



**심현정 (沈 炫 正)**

1999년 강릉대학교 전자공학과 공학사.  
 2001년 강릉대학교 전자공학과 공학석사.  
 2001년~2003년 (주)보이스웨어 연구원.  
 2003년~현재 (주)팬택 & 큐리텔 전임연구원.  
 Tel : 02-580-5176  
 E-mail : shim.hyunjeong@curitel.com



**왕보현 (王 普 賢)**

1987년 연세대학교 전기공학과 공학사.  
 1992년 Georgia Institute of Technology 공학박사. 1992년~1998년 LG 종합기술원 책임연구원. 1997년~현재 강릉대학교 전자공학과 부교수.  
 Tel : 033-640-2384  
 Fax : 033-640-2244  
 E-mail : bhw@kangnung.ac.kr