

## HCBKA를 이용한 Interval Type-2 퍼지 논리시스템 기반 예측 시스템 설계

### Prediction System Design based on An Interval Type-2 Fuzzy Logic System using HCBKA

방 영 근\*      이 철 희\*\*  
Bang, Young-Keun      Lee, Chul-Heui

---

#### Abstract

To improve the performance of the prediction system, the system should reflect well the uncertainty of nonlinear data. Thus, this paper presents multiple prediction systems based on Type-2 fuzzy sets. To construct each prediction system, an Interval Type-2 TSK Fuzzy Logic System and difference data were used, because, in general, it has been known that the Type-2 Fuzzy Logic System can deal with the uncertainty of nonlinear data better than the Type-1 Fuzzy Logic System, and the difference data can provide more steady information than that of original data. Also, to improve each rule base of the fuzzy prediction systems, the HCBKA (Hierarchical Correlation Based K-means clustering Algorithm) was applied because it can consider relationship and statistical characteristics between data at a time. Subsequently, to alleviate complexity of the proposed prediction system, a system selection method was used. Finally, this paper analyzed and compared the performances between the Type-1 prediction system and the Interval Type-2 prediction system using simulations of three typical time series examples.

키워드 : Type-2 퍼지 집합, Interval Type-2 TSK 퍼지 논리시스템, 차분데이터, HCBKA, 시계열 데이터

Keywords : Type-2 fuzzy sets, Interval Type-2 Fuzzy Logic System, difference data, HCBKA, time series

---

#### 1. 서론

시간의 흐름에 따라 일정한 간격으로 기록된 시계열 데이터들은 인간의 일상생활에 밀접한 관련을 가지고 있다. 대표적인 예로는 기상 관측 데이터, 추가지수 데이터, 전력 수요량 등을 들 수 있을 것이다. 이러한 시계열 데이터들은 관측된 자료를 분석하여 미래의 값을 예측함으로써 기상 예보,

경제전망, 전력생산량의 결정 등 인간생활과 밀접한 다양한 분야에서 편의를 제공할 수 있다. 하지만, 이러한 데이터들은 자연현상이나 인간의 생활 패턴 변화 등과 같은 다양한 인자들에 의하여 강한 비선형성(non-linearity)을 가진다. 따라서 예측 시스템의 성능은 데이터의 비선형성을 얼마큼 잘 반영할 수 있는가에 의해 결정된다고 볼 수 있다. 시계열 예측 시스템으로 빈번히 사용되는 고전 선형 회귀모형들의 경우 예측분야의 대표적인 분석 모형으로 간주 될 수 있으나, 이들은 데이터의 비선형성을 반영하기에는 한계점이 존재했다. 이러한 선형 회귀모형들의 문제점들을 해결하기 위해 Zadeh가 처음 제안한 퍼지 이론은 비선형 시스템

---

\* 강원대학교 전기전자공학과 박사과정

\*\* 강원대학교 전기전자공학부 교수, 공학박사, 교신저자

의 설계에 있어 인간의 언어체계의 애매성을 표현할 수 있어 시계열 예측 분야뿐만 아니라 다양한 분야에서 높은 성과를 나타내었다 [1-3]. 하지만 기존의 Type-1 퍼지 논리 시스템은 데이터들이 포함하는 불확실성(uncertainty)을 표현하는데 한계가 있다. 따라서 Zadeh는 Type-2 퍼지 논리 시스템을 제안하여 이러한 불확실성을 효과적으로 취급할 수 있도록 하였으며 [4], Mendel에 의해 Type-2 퍼지 논리시스템은 더욱 체계화되었다.

Type-2 퍼지 논리 시스템은 크게 General Type-2 퍼지논리와 Interval Type-2 퍼지 논리로 구분할 수 있으며, Interval Type-2 퍼지 논리 시스템은 General Type-2 퍼지논리 시스템에 비해 연산의 과정이 간편하면서도 데이터의 불확실성을 충분히 반영할 수 있어 보다 효과적이다 [5-6]. 또한 Takagi와 Sugeno에 의해 제안된 TSK 퍼지 논리 시스템은 인간의 언어적 애매함과 고전 선형모델의 이점을 동시에 취할 수 있어 비선형 시스템 설계에 매우 효과적이다 [7-8].

따라서, 본 논문에서는 비선형 데이터의 불확실성을 적절히 취급할 수 있으며, 언어적 애매성과 선형모델의 이점을 동시에 취할 수 있는 Interval Type-2 TSK 퍼지 논리시스템을 이용하여 예측 시스템을 구현하였다. 또한, 시스템의 설계에 있어 보다 안정된 형태의 정보를 제공할 수 있는 차분 데이터를 이용하여 다수의 퍼지 예측 시스템을 구현하고 [9-10], 각각의 예측 시스템의 규칙기반의 형성을 위해선 HCBKA (Hierarchical Correlation Based K-means Clustering Algorithm)를 이용하여 데이터들의 상관성과 통계적 특성을 동시에 고려할 수 있도록 하였다 [11]. 또한, 시스템 선택을 통해 다중 예측 시스템들 중 하나의 최적 시스템이 예측을 수행하게 함으로써 구조적 복잡성을 최소화 하였다. 마지막으로 3개의 시계열 데이터를 이용하여 기존의 Type-1 TSK 퍼지 예측 시스템과 제안된 예측 시스템의 성능을 비교 분석하였다.

## 2. 차분 데이터의 생성

서론에서 언급되었듯이 본 논문에서는 원형 시계열 데이터의 사용보다 통계적 특성이 안정된 차분데이터들을 이용하여 다중 퍼지 예측 시스템을 구현한다. 하지만 원형 데이터에서 생성가능한 모든 차분데이터들이 원형데이터의 특성보다 우수할 수 없기 때문에 적절한 판별이 요구된다. 따라서 본 논문에서는 다음과 같은 방법으로 적절한 차분 데이터들을 선택한다.

**1단계)** 각각의 차분간격에 따른 상관성 분석

$$r_j = \frac{\frac{1}{N-j} \sum_{i=1}^{N-j} (y(i) - \bar{y})(y(i+j) - \bar{y})}{\frac{1}{N} \sum_{i=1}^N (y(i) - \bar{y})^2}} \quad (1)$$

여기서,  $y(i)$ 는  $i$ 번째 훈련 데이터이며,  $\bar{y}$ 는 훈련 데이터의 평균이다. 또한,  $N$ 은 훈련 데이터의 길이이고,  $j$ 는 차분 간격 값이다.

**2단계)** 0이하의 상관계수 값에 상응하는 차분간격 값 삭제

**3단계)** 상관계수 값의 내림차 정렬

**4단계)** 상위 5개의 상관계수 값 선택

**5단계)** 나머지 상관계수 값들 사이의 차 연산

$$CH_s = (cof_s - cof_{s+1}), \text{ where } s = [1, N-7-\varphi] \quad (2)$$

여기서,  $cof$ 는 (1)에 의해 계산된 상관계수 값들이고,  $\varphi$ 는 2단계에서 제거된 차분 간격 값들의 수이다. 따라서  $s$ 는 남겨진 상관 계수 값들 사이의 차 연산을 위한 총 수행 길이를 의미한다.

**6단계)**  $\max(CH_s)$ 상위의 상관계수 값들 선택

**7단계)** 선택된 상관계수 값들에 상응하는 차분간격 값들을 최적 차분간격 값으로 선택

**8단계)** 선택된 차분간격들을 이용한 차분 데이터의 생성

$$\begin{aligned} d_{m(i)}t_1 &= y(N) - y(N-m(i)) \\ d_{m(i)}t_2 &= y(N-1) - y(N-m(i)-1) \\ &\vdots \\ d_{m(i)}t_n &= y(N-n-1) - y(N-m(i)-n-1) \\ &\vdots \\ d_{m(i)}t_{N-m(i)} &= y(m(i)+1) - y(1) \end{aligned} \quad (3)$$

여기서,  $m(i)$ 는 각각의 선택된 차분 간격 값들의  $i$ 번째 값이며,  $d_{m(i)}t_n$ 은 생성된  $n$ 번째 차분 데이터이다. (3)에 의해 생성된 차분 데이터들은 다중 퍼지 시스템의 구현을 위해 각각 사용된다.

위의 방법들에 의해 선택된 차분간격 값들의 수는 최소 6개 이상이 될 것이며, 이를 통해 구현되는 퍼지 예측 시스템의 수 또한 6개 이상이 된다. 이러한 방법은 각각의 차분 특성들을 충분히 반영함으로써 원형 시계열 데이터가 가지고 있는 다양한 특성들을 충분히 고려하기 위한 것이다 [9-10]. 또한 퍼지 시스템의 설계를 위해 사용될 입력 데이터 쌍은 기존 논문 [9-10]에서 언급된 것과 같이 3개의 차분데이터를 하나의 입력쌍으로 사용하였으며, 따라서 차분 간격에 따라 생성 가능한 입력 쌍은 다음과 같이 주어지게 된다.

$$D_{m(i)} = \{d_{m(i)}t_{k+1}, d_{m(i)}t_{k+2}, d_{m(i)}t_{k+3}, d_{m(i)}t_k\} \quad (4)$$

where,  $k = [1, N - m(i) - 3]$

여기서,  $N - m(i) - 3$ 은 총 생성될 수 있는 입력 데이터 쌍의 개수이다.

### 3. Interval Type-2 TSK 퍼지 논리 시스템

#### 3.1 Type-2 퍼지 논리 시스템

그림 1은 본 논문에 사용된 Type-2 퍼지 논리 시스템의 구조를 보여준다.

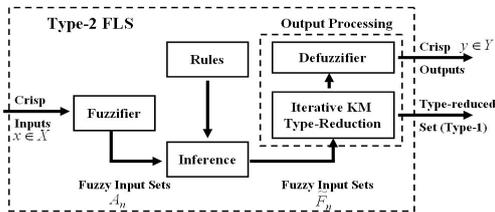


그림 1. Type-2 퍼지 논리 시스템

Type-2 퍼지 논리 시스템의 구조는 Type-1 퍼지 논리 시스템과 유사하나 출력을 얻는 과정에 차이가 있다. 이는 Type-2 퍼지 논리가 불확실성을 표현하기 위하여 FOU(Footprint of Uncertainty)의 개념을 이용하기 때문이다.

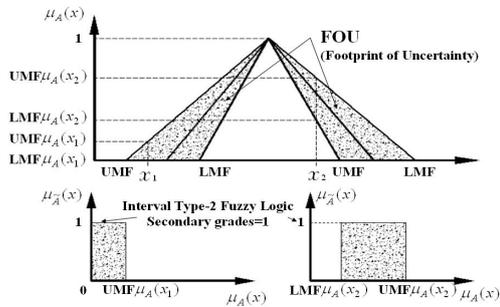


그림 2. Interval Type-2 퍼지 집합의 FOU

그림 2는 불확정 편차에 의한 Interval Type-2 퍼지 집합의 FOU를 보여준다. UMF는 상위 멤버쉽함수를 의미하고, LMF는 하위 멤버쉽함수를 의미한다. Type-2 퍼지 집합  $\tilde{A}$ 는 Type-2 멤버쉽함수  $\mu_{\tilde{A}}(x, \mu_A)$ 로 완전히 기술될 수 있다. 수식 (5)는 Type-2 멤버쉽 함수를 표현한 것이다.

$$\tilde{A} = \{((x, \mu_A), \mu_{\tilde{A}}(x, \mu_A)) \mid \forall x \in X, \forall \mu_A \in J_x \subseteq [0, 1]\} \quad (5)$$

여기서  $\mu_{\tilde{A}}(x, \mu_A)$ 는 0과 1사이의 값을 가지며, 식(5)는 다음과 같이 다시 표현될 수 있다.

$$\tilde{A} = \int_{x \in X} \int_{\mu_A \in J_x} \mu_{\tilde{A}}(x, \mu_A) / (x, \mu_A) \quad J_x \subseteq [0, 1] \quad (6)$$

Type-2 퍼지 집합 식(6)는 다시 식 (7)처럼 간략화 될 수 있다. 이는 본 논문에서 Interval Type-2 퍼지 집합을 사용하므로 Secondary grades  $\mu_{\tilde{A}}(x, \mu_A)$ 가 모두 1이기 때문이다.

$$\tilde{A} = \int_{x \in X} \int_{\mu_A \in J_x} 1 / (x, \mu_A) \quad J_x \subseteq [0, 1] \quad (7)$$

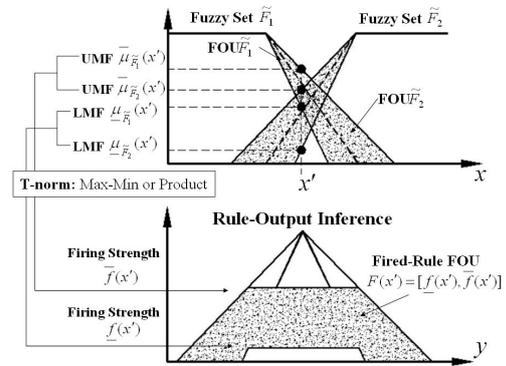


그림 3. Type-2 퍼지 논리 시스템의 추론

그림 3은 Type-2 퍼지 논리 시스템의 추론과정을 보여준다. Type-2 퍼지논리 시스템의 추론 과정은 그림에서 보여 지듯이 규칙의 상위 멤버쉽함수의 T-norm연산과 하위 멤버쉽 함수의 T-norm 연산으로 기술될 수 있다.

#### 3.2 제안된 HCBKA의 구조

그림 4는 본 논문에 사용된 Interval Type-2 퍼지 논리 시스템을 이용한 HCBKA (Hierarchical Correlation Based K-means Clustering Algorithm)의 구조를 보여준다.

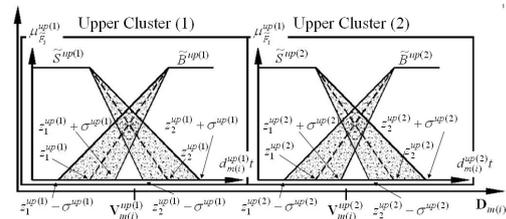


그림 4. 제안된 HCBKA의 구조

그림과 같이 HCBKA는 crisp한 상위 클러스터들과 각각의 클러스터들을 구성하는 퍼지 집합으로 이루어져 있다. 상위 클러스터를 구현하기 위해선, 먼저 임의의 상위 클러스터 중심  $V_{m(i)}^{up}$ 가  $[v1_{m(i)}^{up}, v2_{m(i)}^{up}, v3_{m(i)}^{up}]$ 이라면, 각각의 입력데이터 쌍은 다음의 적합도 함수를 더 크게 만족하는 클러스터 쪽으로 분류된다.

$$\rho_{TV}^{up} = \frac{\sum_{l=1}^3 (d_{m(i)} t_{k+l} - \bar{T}_k) (V_l^{up} - \bar{V}^{up})}{\sqrt{\sum_{l=1}^3 (d_{m(i)} t_{k+l} - \bar{T}_k)^2} \sqrt{\sum_{l=1}^3 (V_l^{up} - \bar{V}^{up})^2}} \quad (8)$$

여기서,  $\bar{T}_k$ 는 각각의 입력데이터 쌍의 평균을 의미하고,  $\bar{V}^{up}$ 는 각각의 상위 클러스터 중심 값들의 평균을 의미한다. 이렇게 분류된 상위 클러스터의 중심은 다음과 같은 방법에 의해 갱신된다.

$$V_{m(i)}^{up} = \frac{1}{m^{up}} \sum_{n=1}^{m^{up}} V^{up}(M) \quad (9)$$

여기서,  $V^{up}(M)$ 는  $V^{up}$ 에 분류된 입력 데이터 쌍들이고,  $m^{up}$ 는 데이터의 개수이다. 이러한 중심 값의 갱신은 다음과 같이 정의되는 왜곡을 만족할 때까지 반복된다.

$$V_D = \left| \frac{V_{pre} - V_{curr}}{V_{pre}} \right| < 10^{-4} \quad (10)$$

여기서,  $V_{pre}$ 는 이전의 중심 값들이며,  $V_{curr}$ 는 현재의 갱신된 중심 값을 의미한다.

이렇게 상위 클러스터에 분류된 데이터들은 각각의 하위 퍼지집합의 구현을 위한 데이터로 사용된다. 먼저, 각각의 상위 클러스터에 분류된 데이터들 중 최소값과 최대값을 전체 집합으로 하여  $\tilde{S}$ ,  $\tilde{B}$ 의 두 개의 단일 표준편차를 이용한 불확정 편차 기반 Type-2 퍼지집합으로 구현된다. 그림 4에서  $\mathbf{z}$ 는 각 퍼지집합의 중심을 의미하며, FOU를 위한  $\sigma$ 는 다음과 같이 정의된다.

$$\sigma^{up} = \arg \sum_{i=1}^b \left( \sqrt{\sum_{h=1}^g (d_{m(i)}^u t_h^i - \bar{d}_{m(i)}^{up} t^i) / (g-1)} \right) \quad (11)$$

여기서  $b$ 는 퍼지집합의 총 수이며,  $g$ 는 각각의 퍼지집합에 포함된 데이터의 수이다. 이렇게 구현된 Type-2 퍼지 집합들은 K-means 클러스터링 알고리즘에 의해 보다 적합한 형태로 구현된다. 본 논문에서는 2개의 상위 클러스터와 2개의 퍼지집합을 이용함으로써 최소규칙이 생성될 수 있도록

하였으며, 또한 계층적 분류로 나타낼 수 있는 데이터의 결핍현상을 완화할 수 있도록 하였다. 따라서 3입력 데이터에 의한 Type-2 TSK 퍼지 논리 시스템의  $up$ 번째 클러스터의  $j$ 번째 Type-2 TSK 퍼지 규칙은 다음과 같이 일반적으로 정의된다.

$Rule(R_j^{up})$

If

$d_{m(i)}^{up} t_{k+1}^j$  is  $A_j$  and  $d_{m(i)}^{up} t_{k+2}^j$  is  $B_j$  and  $d_{m(i)}^{up} t_{k+3}^j$  is  $C_j$

Then

$$d_{m(i)}^{up} t_k^j = p_0^{j(up)} + p_1^{j(up)} d_{m(i)}^{up} t_{k+1}^j + p_2^{j(up)} d_{m(i)}^{up} t_{k+2}^j + p_3^{j(up)} d_{m(i)}^{up} t_{k+3}^j \quad (12)$$

여기서  $j=1, \dots, Q$  이며  $P$ 는 추정되어야 할 파라미터이다. 따라서 위와 같은 퍼지규칙에서의 firing set  $F^j(\mathbf{x})$ 는 다음과 같이 정의될 수 있다.

$$F^j(\mathbf{x}) = \prod_{k=1}^l \mu_{\tilde{F}_k}(x_k) \quad (13)$$

여기서,  $\Pi$ 은 Type-2 퍼지집합의 meet 연산을 의미한다. 따라서 식 (13)은 Interval Type-2 집합에서 minimum T-norm에 의한 meet연산에 의해 다음과 같이 다시 정의 될 수 있다.

$$F^j(\mathbf{x}) = [f^j(\mathbf{x}), \bar{f}^j(\mathbf{x})] \quad (14)$$

여기에서,

$$f^j(\mathbf{x}) = \underline{\mu}_{\tilde{F}_1}(x_1) \star \dots \star \underline{\mu}_{\tilde{F}_l}(x_l) \quad (15)$$

$$\bar{f}^j(\mathbf{x}) = \bar{\mu}_{\tilde{F}_1}(x_1) \star \dots \star \bar{\mu}_{\tilde{F}_l}(x_l) \quad (16)$$

여기에서  $\star$ 는 T-norm 연산자를 의미한다.

### 3.3 Type-2 퍼지논리 시스템의 파라미터추정

만약 차분간격  $m(i)$ 에 대한  $i$ 번째 TSK 퍼지 시스템의  $up$ 번째 상위 클러스터의  $j$ 번째 퍼지규칙  $R_j^{up}$ 을 만족하는 입력 데이터쌍의 수가  $n$ 개이면, 후반부의 선형 식은 다음과 같이  $n$ 개의 연립 방정식으로 표현된다.

$$\begin{aligned} d_{m(i)}^{up} t_k^{j(1)} &= p_0^{j(up)} + p_1^{j(up)} d_{m(i)}^{up} t_{k+1}^{j(1)} + p_2^{j(up)} d_{m(i)}^{up} t_{k+2}^{j(1)} \\ &\quad + p_3^{j(up)} d_{m(i)}^{up} t_{k+3}^{j(1)} \\ d_{m(i)}^{up} t_k^{j(2)} &= p_0^{j(up)} + p_1^{j(up)} d_{m(i)}^{up} t_{k+1}^{j(2)} + p_2^{j(up)} d_{m(i)}^{up} t_{k+2}^{j(2)} \\ &\quad + p_3^{j(up)} d_{m(i)}^{up} t_{k+3}^{j(2)} \\ &\quad \vdots \\ d_{m(i)}^{up} t_k^{j(n)} &= p_0^{j(up)} + p_1^{j(up)} d_{m(i)}^{up} t_{k+1}^{j(n)} + p_2^{j(up)} d_{m(i)}^{up} t_{k+2}^{j(n)} \\ &\quad + p_3^{j(up)} d_{m(i)}^{up} t_{k+3}^{j(n)} \end{aligned} \quad (17)$$

이 연립방정식은 최소자승법에 의한 파라미터 추정을 위해 다음과 같이 변형 될 수 있다.

$$\begin{bmatrix} d_{m(i)}^{up} t_k^{j(1)} \\ d_{m(i)}^{up} t_k^{j(2)} \\ \vdots \\ d_{m(i)}^{up} t_k^{j(n)} \end{bmatrix} = \begin{bmatrix} 1 & d_{m(i)}^{up} t_{k+1}^{j(1)} & d_{m(i)}^{up} t_{k+2}^{j(1)} & d_{m(i)}^{up} t_{k+3}^{j(1)} \\ 1 & d_{m(i)}^{up} t_{k+1}^{j(2)} & d_{m(i)}^{up} t_{k+2}^{j(2)} & d_{m(i)}^{up} t_{k+3}^{j(2)} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & d_{m(i)}^{up} t_{k+1}^{j(n)} & d_{m(i)}^{up} t_{k+2}^{j(n)} & d_{m(i)}^{up} t_{k+3}^{j(n)} \end{bmatrix} \begin{bmatrix} p_0^{j(up)} \\ p_1^{j(up)} \\ p_2^{j(up)} \\ p_3^{j(up)} \end{bmatrix} \quad (18)$$

$$Y_j^{up} = D_j^{up} P_j^{up} \quad (19)$$

여기서,  $Y_j^{up}$ 는 출력벡터,  $D_j^{up}$ 는 입력벡터,  $P_j^{up}$ 는 파라미터 벡터를 의미한다. 따라서 파라미터들은 식(15)의 최소자승법을 통해 추정되게 된다.

$$\hat{P}_j^{up} = ((D_j^{up})^T D_j^{up})^{-1} (D_j^{up})^T Y_j^{up} \quad (20)$$

### 3.4 Type-2 퍼지논리 시스템의 출력

식 (12)의 Interval Type-2 TSK 퍼지 논리 시스템의 출력 역시 Interval Set이 되므로  $Q$ 개의 퍼지규칙을 만족한 입력쌍에 대한 Interval Type-2 TSK 논리시스템의 최종 출력은 Karnik and Mendel이 제안한 Iterative KM 알고리즘 [12]을 이용하여 다음과 같이 구할 수 있다.

$$y^L = \min_{\theta_j} \frac{\sum_{j=1}^Q \hat{d}_{m(i)} y_k^j}{\sum_{j=1}^Q \theta_j} \quad (21)$$

$$y^R = \max_{\theta_j} \frac{\sum_{j=1}^Q \hat{d}_{m(i)} y_k^j}{\sum_{j=1}^Q \theta_j} \quad (22)$$

위 식(21)과 (22)는 Iterative KM 알고리즘에 적용하여 구해지는 최소 가중 평균 무게중심  $y^L$ 과 최대 가중 평균 무게중심  $y^R$ 로 일반적으로 다음과 같이 다시 정의 될 수 있다.

$$y^L = \frac{\sum_{j=1}^W \bar{f}^j \hat{d}_{m(i)} y_k + \sum_{j=W+1}^Q \underline{f}^j \hat{d}_{m(i)} y_k}{\sum_{j=1}^W \bar{f}^j + \sum_{j=W+1}^Q \underline{f}^j} \quad (23)$$

$$y^R = \frac{\sum_{j=1}^H \underline{f}^j \hat{d}_{m(i)} y_k + \sum_{j=H+1}^Q \bar{f}^j \hat{d}_{m(i)} y_k}{\sum_{j=1}^H \underline{f}^j + \sum_{j=H+1}^Q \bar{f}^j} \quad (24)$$

따라서 Interval Type-2 퍼지논리 시스템의 출력은 다음과 같이 정의된다.

$$\hat{d}_{m(i)} y_k^{TSK,2} = (y^L + y^R)/2 \quad (25)$$

또한, 본 논문에서는 차분데이터를 퍼지 시스템의 입력으로 사용하였으므로 식 (21)의 출력은 현재값과 미래값 사이의 증가분을 의미하게 된다. 따라서 제안된 시스템의 최종 출력은 다음과 같이 현재값과 증가분의 합으로 주어진다.

$$\hat{y}(k+p) = y(k) + \hat{d}_{m(i)} y_k^{TSK,2} \quad (26)$$

본 논문에서  $p$ 는 1로 정의되며, 따라서 시스템의 출력 값은 현재 값에 대한 한 스텝 전망 예측 값을 의미하게 된다.

## 4. 예측 시스템의 선택

6개 이상으로 구현된 다중 퍼지 예측시스템 중에서 최종 예측을 수행할 하나의 예측시스템을 선택하기 위하여 MSE(mean squared error)를 이용하였으며 다음과 같다.

$$MSE = \frac{1}{N} \sum_{n=m(i)+4}^N (y(n) - \hat{y}(n))^2 \quad (27)$$

여기서,  $n=m(i)+4$ 는 3입력 데이터에 의해 실제로 예측 되는 데이터는 4번째 데이터부터이므로 정확한 평균을 구하기 위함이다. 따라서 구현된 다중 퍼지 예측 시스템들은 훈련데이터를 이용하여 예측을 수행한 후 MSE를 최소화한 하나의 예측 시스템이 선택되며, 선택된 예측 시스템이 전체 구간의 예측을 수행하게 된다.

## 5. 시뮬레이션

제안된 방법의 성능을 검증하기 위하여 3개의 시계열 데이터를 시뮬레이션 하여 기존의 Type-1 퍼지 논리시스템 기반 예측 시스템의 성능과 제안된 시스템의 성능을 비교하였다. 아래의 표 1은 시뮬레이션에 사용된 각각의 데이터 지표들을 보여준다.

표 1. 시뮬레이션 데이터 지표 [13]

지표 \ 데이터	전체 데이터	훈련길이	예측길이	성능지표
우유생산량	166	80	86	MRE
다우지수	292	150	142	MRE
전력생산량	155	70	85	MRE

각각의 데이터의 성능평가 지표로 사용된 MRE(mean relative error)는 다음과 같이 정의된

다.

$$MRE = \left( \frac{1}{N} \sum_{i=1}^N \frac{|y(i) - \hat{y}(i)|}{|y(i)|} \right) \times 100\% \quad (28)$$

아래의 그림들은 각각의 데이터들에 대한 제안된 예측 시스템의 Type-2 퍼지 집합의 동조 과정과 예측결과 및 오차를 보여준다.

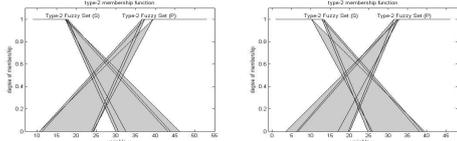


그림 5. 우유생산량 데이터의 퍼지집합 동조

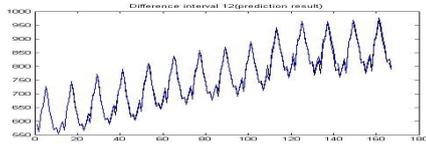


그림 6. 우유생산량 데이터의 예측 결과

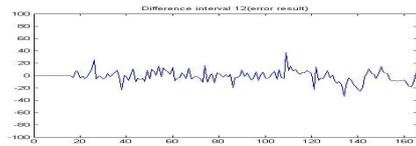


그림 7. 우유생산량 데이터의 예측 오차

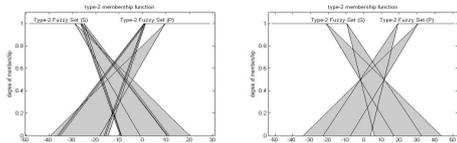


그림 8. 다우지수 데이터의 퍼지집합 동조

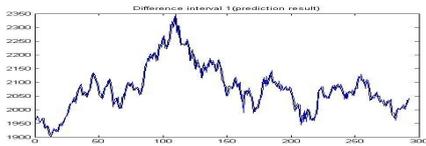


그림 9. 다우지수 데이터의 예측 결과

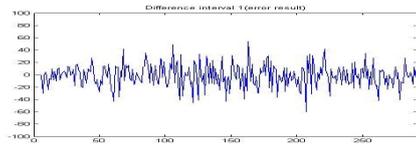


그림 10. 다우지수 데이터의 예측 오차

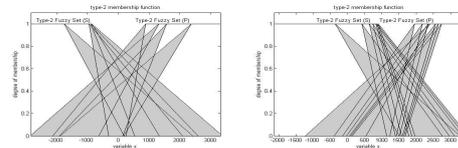


그림 11. 전력생산량 데이터의 퍼지집합 동조

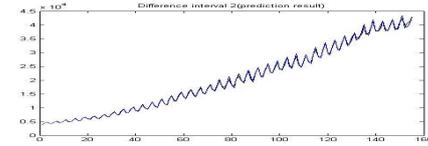


그림 12. 전력생산량 데이터의 예측 결과

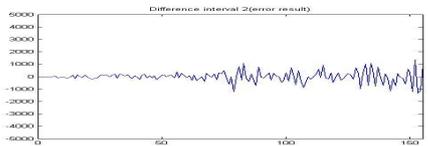


그림 13. 전력생산량 데이터의 예측 오차

시뮬레이션 결과의 그림 6, 9, 12에서 검은색은 원 시계열 데이터 값을 의미하고 파란색은 예측된 값을 의미한다. 그림을 살펴보면 검은색선과 파란색선이 거의 중복되어 있음을 알 수 있으며, 이는 제안된 예측 시스템의 성능이 비교적 우수함을 보여주는 결과이다. 이러한 결과들은 그림 7, 10, 13에서 보여주는 예측 오차를 살펴봐도 비교적 적은 예측 오차를 가짐을 알 수 있다. 또한, 그림 5, 8, 11은 상위 두 개의 클러스터내의 Type-2 퍼지 집합들이 K-means 클러스터링 알고리즘에 의해 동조되는 과정을 보여준다. 아래의 표 2는 기존의 Type-1 퍼지 논리 시스템 기반의 예측 시스템의 결과와 제안된 Interval Type-2 퍼지 논리 시스템 기반 예측 시스템의 성능을 비교한 것으로 성능 비교 지수로는 식(28)의 MRE를 사용하였다.

표 2. Type-1과 Type-2 퍼지 논리 시스템 기반 다중예측 시스템의 성능 비교

시스템 데이터	based on Type-1 FLS		based on Type-2 FLS	
	훈련구간 (MSE)	예측구간 (MRE)	훈련구간 (MSE)	예측구간 (MRE)
우유생산량	65.2452	0.9951	65.0955	0.9918
다우지수	332.4111	0.6941	322.0540	0.6929
전력생산량	2.2317E+4	1.5315	2.2298E+4	1.5310

표 2를 살펴보면 기존의 Type-1 퍼지 논리 시스템 기반 다중 예측시스템의 결과보다 제안된

Interval Type-2 퍼지 논리시스템 기반 다중 예측 시스템의 결과가 훈련구간이나 예측 구간 모두에서 우수하게 나타났으며, 예측 결과 MRE의 값들이 비교적 작은 좋은 예측 성능을 보였다.

## 6. 결론

본 논문에서는 차분데이터와 Interval Type-2 TSK 퍼지 논리 및 HCBKA를 이용하여 비선형 시계열 데이터 예측을 위한 예측 시스템을 구현하였다. 먼저, 시뮬레이션 결과에서 보여주듯이 Interval Type-2 TSK 퍼지 논리시스템 기반 예측 시스템의 성능이 기존의 Type-1 퍼지 시스템에 비해 좋음을 알 수 있다. 또한, 최종 예측 결과들은 제안된 예측 시스템의 성능이 우수함을 보여 주는 것으로, 이는 차분데이터를 이용하여 시스템 모델링에 있어 안정된 정보를 제공할 수 있었을 뿐만 아니라 HCBKA를 통해 각각의 퍼지 예측 시스템의 규칙기반의 적합성을 높일 수 있었던 결과로 볼 수 있다. 따라서 제안된 예측 시스템의 설계 기법들은 비선형 시계열 예측 시스템 구현 분야뿐만 아니라 비선형 제어시스템의 설계 등 다양한 분야로의 접근이 가능할 것으로 생각된다.

## 참 고 문 헌

- [1] George J. Klir and Bo Yuan, *Fuzzy Sets and Fuzzy Logic Theory and Applications*, Prentice-Hall, 1995.
- [2] G.Janazcek, L.Swift, *Time Series Forecasting, Simulation, Applications*, Ellis Horwood, 1993.
- [3] K.Ozawa, T.Niimura, "Fuzzy Time-Series Model of Electric Power Consumption", *IEEE Canadian Conference on Electrical and Computer Engineering*, pp. 1195-1198, 1999.
- [4] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning-1", *Information Science*, vol. 8, pp. 199-249, 1975.
- [5] J. M. Mendel, "Uncertain Rule-Based Fuzzy Logic System: Introduction and New Directions", *Prentice-Hall, Upper Saddle River, Nj 07458*, 2001.
- [6] J. M. Mendel, R. I. John, F. Lui "Interval Type-2 Fuzzy Logic Systems Made Simple", *IEEE Trans. on Fuzzy System*, vol. 14, pp. 808-821, 2006.
- [7] T. Takagi, M. Sugeno, "Fuzzy Identification of Systems and its Applications to modeling and control", *IEEE Trans. Syst, Man, Cybern*, vol. 15, pp. 116-132, 1985.
- [8] K. Tanaka, M. Sugeno, "A robust stabilization problem of fuzzy control systems and its application to backing up control of a truck-trailer", *IEEE Trans. Fuzzy Systems*, vol. 2, pp. 119-134, 1994.
- [9] Y. K. Bang, C. H. Lee, "Design of Multiple Model Fuzzy Predictors using Data Preprocessing and its Application", *Transactions of KIEE*, vol. 58, No. 1, pp. 173-180, 2009.
- [10] Y. K. Bang, C. H. Lee, "Multiple Model Fuzzy Prediction Systems with Adaptive Model Selection Based on Rough Sets and its Application to Time Series Forecasting", *Journal of KIIS*, vol. 19, No. 1, pp. 25-33, 2009.
- [11] Y. K. Bang, C. H. Lee, "Design of Fuzzy System with Hierarchical Classifying Structures and its Application to Time Series Prediction", *Journal of KIIS*, vol. 19, No. 5, pp. 595-602, 2009.
- [12] N. N. Karnik, J. M. Mendel, "Centroid of a Type-2 Fuzzy Set", *Information Sciences*, vol. 132, pp. 195-200, 2001.
- [13] <http://www-personal.buseco.monash.edu.au>