

최적 TS 퍼지 모델 기반 다중 모델 예측 시스템의 구현과 시계열 예측 응용

Multiple Model Prediction System Based on Optimal TS Fuzzy Model and Its Applications to Time Series Forecasting

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

Abstract

In general, non-stationary or chaos time series forecasting is very difficult since there exists a drift and/or nonlinearities in them.

To overcome this situation, we suggest a new prediction method based on multiple model TS fuzzy predictors combined with preprocessing of time series data, where, instead of time series data, the differences of them are applied to predictor's as input. In preprocessing procedure, the candidates of optimal difference interval are determined by using correlation analysis and corresponding difference data are generated. And then, for each of them, TS fuzzy predictor is constructed by using k-means clustering algorithm and least squares method. Finally, the best predictor which minimizes the performance index is selected and it works on hereafter for prediction.

Computer simulation is performed to show the effectiveness and usefulness of our method.

키워드 : 데이터 전처리, 상관분석, 최적 차분 간격, 다중 모델 퍼지 예측기, 모델선택
Keywords : data preprocessing, correlation analysis, optimal difference interval, multiple model fuzzy predictor, model selection

1. 서론

시계열 데이터는 전력의 생산량이나 전력 수요(부하), 태양 흑점 변동, 지구의 기온 변화, 주가지수, 환율 변화 등과 같이 시간의 흐름에 따라 일정한 간격으로 관측하여 기록한 자료로 인간의 생활과 밀접한 다양한 분야에서 나타난다. 이러한 시계열 데이터의 예측은 미래에 대한 정보를 현재에서 얻을 수 있는 것이므로 매우 유용하다. 하지만 미래의 정보를 정확하게 예측하는 것은 쉬운 문제가 아니다. 또한 이런 시계열 데이터의 특징은 같은 기간이나 상황에서 독립적인 반복 관측이 허용

되지 않고 시간에 따라 관측된 한 계열만을 이용하기 때문에 더욱 필요한 정보를 얻는 것이 어렵다. 따라서 미래의 정보를 좀 더 정확하게 예측하기 위해서는 관측된 시계열 데이터를 정확하게 분석할 수 있는 적절한 분석 기법이 필요하다.

시계열 예측을 목적으로 관측된 자료를 분석하기 위해서는, 먼저 시계열 데이터가 가지고 있는 특징들을 분석하는 것이 필요하다. 일반적으로 시계열 데이터들은 자료가 관측된 순서에 따라 배열되고 인접된 값들이 서로 독립이 아니고 종속적인 특징을 가진다. 또한 비정상 또는 혼돈 시계열의 경우에는 불확실성(uncertainty)과 비선형적(non-linear) 특성들이 내재되어 있다. 따라서 이러한 시계열 데이터를 분석하는 것은 시계열 데이터에 내재된 확률적, 통계적 특성들을 파악하거나, 추세 성분, 순환 성분, 계절 변동 성분, 확률 변동 성분

* 강원대학교 대학원 전기전자공학과 박사과정
** 강원대학교 전기전자공학부 교수, 교신저자

등이 일부 또는 모두 포함되어 시계열 데이터가 생성되었다고 보고 이에 대한 모델을 개발함으로써 수행할 수 있다. 시계열 분석을 위한 모델로는 AR(Auto Regressive), MR(Moving Average), ARMA(Auto Regressive Moving Average)와 같은 선형 회귀모델이나 Box-Jenkins 방법(Auto Regressive Integrated Moving Average : ARIMA) 등이 널리 사용되어 왔다[1-3]. 하지만 이와 같은 선형적 접근 방법들이 비정상 시계열이나 혼돈 시계열 모두에 대하여 적합한 것은 아니며 최근에는 이러한 선형적 접근 방법들로 해결하지 못하는 문제점들을 극복하고자 퍼지이론, 신경망, 유전 알고리즘 등과 같은 soft computing 기법을 이용하여 시계열 데이터의 불확실성과 비선형적 특성을 적절하게 취급할 수 있는 방법들에 대한 연구가 많이 이루어지고 있다[4-5]. 그중에서 특히 퍼지 이론은 시계열 데이터에 내재된 부정확성(imprecision)과 불확실성(uncertainty)을 수리적으로 취급 가능하도록 하여 시계열 데이터들의 특징들을 반영하는 부분에 있어 매우 효과적이다.

그렇지만 이원 퍼지이론만으로는 모든 비정상 또는 혼돈시계열의 특징들을 표현하는 것에는 많은 한계가 따른다. 그러므로 이러한 한계를 극복하기 위하여 퍼지-신경망, 유전자-퍼지 등 soft computing 기법들을 혼합한 방법[6-8]들과 시계열 데이터를 적절하게 가공하여 사용하는 방법[9-10]들에 대한 연구가 진행되고 있다.

이 논문에서는 예측 모델의 구현의 효율성을 위하여 퍼지 이론을 적용하여 예측 시스템을 구현하며, 대신 퍼지 이론의 한계를 극복하기 위하여 위에서 설명된 후자의 데이터의 적절한 가공법을 제시한다. 먼저, 이 논문에서는 시계열 예측을 위한 예측 시스템을 구현하기 위하여 T-S 퍼지 모델을 이용하였다. 시계열 예측 시스템은 하나의 독립된 T-S 퍼지 모델로 구현되는 것이 아니라 비정상 또는 혼돈 시계열 데이터의 특성을 잘 표현할 수 있는 다수의 최적 T-S 퍼지 모델들로 구성된 다중 퍼지 예측 시스템을 이용한다. 다수의 최적 T-S 모델 구현을 위하여서는 상관분석을 기반으로 하는 데이터 전처리 과정을 통해 비정상 시계열을 잘 표현할 수 있는 적절한 수의 차분 데이터들을 생성하고 이들 각각에 대하여 T-S 퍼지 모델들을 구현하게 된다. 또한 T-S 퍼지 모델들의 퍼지 분할과 파라미터 추정에 k-means 클러스터링 알고리즘[11-12]과 LSM(Least Squares Method) 이용하여 최적의 T-S 퍼지 모델들을 구현하고, 이들을 이용하여 다중 퍼지 예측시스템을 구성한다. 이렇게 최적 T-S 퍼지 모델들로 구성된 다중 퍼지 예측 시스템은 검증을 통해 최종 예측을 수행하는 하나의 최상의 T-S 모델을 선택하게 되고, 선택된 T-S 모델은 미래의 정보를 얻기 위하여 예측을 수

행한다. 따라서 제안된 논문의 방법은 다중 퍼지 예측 시스템의 구현, 검증(모델선택), 예측과 같이 크게 3단계로 이루어지며 시뮬레이션을 통하여 제안된 방법의 유용성을 입증하였다.

2. 제안된 퍼지 예측 시스템의 개요

2.1 T-S 퍼지 모델

제안된 방식에서 사용하는 예측 모델로는 T-S 퍼지 모델을 이용하였다. 기존의 연구에서는 Mamdani 퍼지 모델을 많이 사용하였으나, Mamdani 모델의 경우 일반적으로 원하는 예측 정확도를 달성하기 위해서는 입력력 변수에 대한 퍼지 분할 영역(퍼지 집합)과 입력의 수가 많아야 하며, 데이터에 기반한 퍼지 규칙 도출이 까다롭고 복잡하다. 이에 반해 제안된 방법에서 사용한 T-S 퍼지 모델은 전건부에는 퍼지 집합을 이용함으로써 데이터가 지나고 있는 불확실성이나 비선형적 특성을 적절하게 취급할 수 있으며 후건부의 선형 수식을 통해 고선적인 선형 회귀 모델의 이점도 동시에 취할 수 있는 장점이 있다. 또한, 데이터 표현에 대한 적합성과 예측기 구현의 용이성도 포함할 수 있다. 아래의 식(1)은 제안된 방법에 사용된 T-S 퍼지 모델의 R번째 규칙의 표현 방식이다.

$$\begin{aligned}
 R: & \text{ If } x_1 \text{ is } A_1 \text{ and } x_2 \text{ is } A_2 \text{ and } \dots & (1) \\
 & \text{ and } x_n \text{ is } A_n \\
 \text{Then } & y = a_0 + a_1x_1 + a_2x_2 + \dots + a_nx_n
 \end{aligned}$$

위 식(1)에서 x_1, x_2, \dots, x_n 은 예측을 위하여 입력되는 사전에 관측된 자료이며 A_1, A_2, \dots, A_n 은 이 자료들에 적합한 언어적 변수 값이다. 또한 a_0, a_1, \dots, a_n 은 학습을 통하여 추정되는 파라미터 값들이며, y 는 추정된 파라미터들과 관측된 입력들에 의해 예측된 미래의 정보 값이다. 따라서 T-S 퍼지 모델 규칙은 데이터의 언어적 변수 값을 부여하기 위하여 사전 정보와 데이터로부터 규칙의 전건부에 사용되는 입력 공간의 퍼지 분할과 후건부 선형 수식의 파라미터를 식별하는 과정을 필요로 하게 된다. 입력 공간의 퍼지 분할에는 경험적 방법, 템플릿 기반 방법, clustering, 신경망 학습이나 유전 알고리즘을 이용한 최적화 등 여러 가지 방법을 적용할 수 있으나 본 논문에서는 k-means 알고리즘을 이용한 clustering 기법을 적용하여 데이터 표현에 대한 적합도를 향상시켰으며, 후건부 파라미터의 식별에는 가장 대표적인 최소 자승법(Least Squares Method)을 이용하여 파라미터 식별의 부정확성을 최소화하도록 하였다. 자세한 설명은 3장에서 다루기로 한다.

2.2 다중 퍼지 예측 시스템

비정상 또는 혼돈 시계열 데이터의 예측을 위하여 원 시계열 데이터를 그대로 예측기의 입력으로 사용할 경우 시계열에 내재된 부정확성과 불확실성을 적절히 표현할 수 없으므로 정확한 예측이 어렵다. 따라서 제안된 논문에서는 이러한 비정상 또는 혼돈 시계열의 통계적 특성을 안정화 하여 시계열 데이터에 내재된 패턴이나 규칙성들을 잘 표현할 수 있도록 하기 위하여 차분 데이터를 이용한다. 이러한 차분 데이터는 하나의 차분 시계열이 아니라 다수의 차분 시계열로 표현되며 이렇게 표현된 다수의 차분 시계열을 이용하여 2.1절에서 설명된 T-S퍼지 모델의 입력과 파라미터 추정에 사용함으로써 최적의 T-S퍼지 모델을 구현하게 된다. 이렇게 구현된 최적 T-S모델들은 제안된 방식에 표현된 하나의 다중 퍼지 예측 시스템을 구성하게 된다. 아래의 그림은 제안된 방법에 쓰인 다중 퍼지 예측 시스템의 구조도이다.

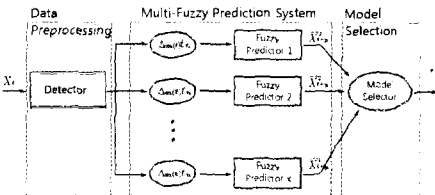


그림 1 다중 퍼지 예측 시스템의 구조

3. 제안된 다중 퍼지 예측시스템의 구현과 예측 모델 선택

제안된 알고리즘의 다중 퍼지 예측 시스템을 구현하기 위하여 먼저 다중 퍼지 예측 시스템을 구성할 최적의 T-S퍼지 모델의 구현이 필요하다. 이를 위해선 최적의 T-S퍼지 모델들 각각에 쓰일 입력과 파라미터 추정에 필요한 데이터들이 필요하게 된다. 제안된 알고리즘에서는 이러한 각각의 최적 T-S퍼지 모델을 구현하기 위하여 입력과 파라미터 추정을 위한 데이터로 원 시계열 데이터를 적절히 가공하여 비정상 또는 혼돈 시계열에 내재된 불확실성과 비선형성이 잘 표출될 수 있는 차분 데이터를 사용하였다. 이러한 차분 데이터들은 그림 1의 데이터 전처리 과정을 통하여 선별될 수 있으며, 이렇게 선별된 차분데이터들로부터 다중 퍼지 예측 시스템을 구성할 T-S 퍼지 모델을 구현하게 된다. 최적 T-S퍼지 모델의 입력으로 쓰일 차분데이터를 선별하는 방법으로 제안된 알고리즘에서는 상관함수를 이용하였고 이렇게 선별된 차분데이터를 이용한 최적의 T-S퍼지 모델기반 다중퍼지 예측 시스템에서 최종 예측을 수행할 최적 예측 T-S퍼지 모델 선택에서는 MSE(mean

square error)를 성능지표로 사용하였다. 아래의 그림2는 제안된 알고리즘의 순서도 이다.

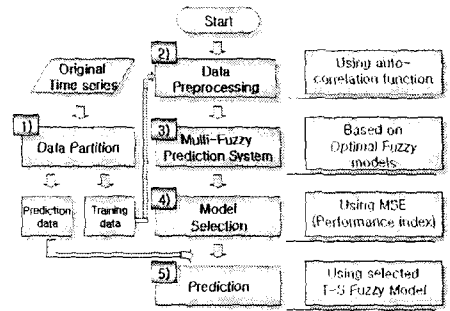


그림 2 제안된 알고리즘의 순서도

그림 2를 살펴보면 제안된 알고리즘은 크게 3단계로 구현 할 수 있다. 먼저 예측 수행에 필요한 최적 T-S퍼지 모델 기반 다중 퍼지 예측 시스템을 구현하는 단계(그림에서 1,2,3 단계)와 검증을 통해 최종 예측을 수행할 최종 T-S 퍼지 모델 선택(그림에서 4)단계, 그리고 마지막으로 선택된 T-S 퍼지 모델을 이용하여 최종예측을 수행하는 예측단계(그림에서 5)로 구성된다.

3.1 다중 퍼지 모델의 구현

먼저 제안된 알고리즘을 실현하기 위해서는 그림2의 1)에 해당하는 다중 퍼지 모델의 구현을 위한 데이터(즉, 예측시스템을 학습시킬 구간)와 실제 예측을 수행할 데이터를 설정하기 위한 원시계열 데이터를 적합하게 분할하여야 한다. 그러므로 원시계열 데이터는 위의 그림 2처럼 훈련데이터와 예측 데이터로 분할된다. 이때 훈련데이터의 크기는 다중 퍼지 예측 시스템이 예측을 수행함에 있어 학습이 충분히 이루어 질 수 있을 정도로 적당한 길이로 분할하여야 한다. 다음은 이렇게 분할된 훈련데이터를 이용하여 다중 퍼지 예측시스템을 구성할 최적의 T-S퍼지 모델들을 구현한다. 최적의 T-S퍼지 모델들을 구현하기 위해서는 이 퍼지 모델들의 입력으로 쓰일 데이터가 필요하게 되는데, 이를 위해 제안된 알고리즘에서는 데이터의 전처리 과정을 통하여 T-S 퍼지 모델의 구현에 적합한 차분데이터들을 생성하고, 이렇게 생성된 차분데이터들을 각각의 T-S 퍼지 모델들의 입력 및 파라미터 추정에 사용하였다. 그러므로 선별된 차분데이터들의 개수와 구현되는 최적 T-S모델의 개수는 같아지게 된다. 데이터의 전처리 과정은 그림2의 2)에 해당하는 처리 과정이다. 이 과정은 데이터 분할과정을 통해 적절히 분할된 훈련데이터를 이용하여 최적의 T-S퍼지 모델들을 구현하기 위한 차분 데이터를 생성하는 과정이다. 만약 훈련 데이터의 길이가 N 이라면 이 훈련데이터에 의하

여 생성되는 차분데이터의 개수는 $N-1$ 가 될 것이다. 이 많은 차분데이터들로 T-S퍼지 모델을 구현하게 되면 예측모델의 수가 너무 많고 또한 예측의 성능이나 시간에 많은 문제점이 생긴다. 따라서 제안된 알고리즘에서는 생성되는 차분데이터 전부를 사용하는 것이 아니라 원시계열에 내재된 비선형성이나 불확실성을 가장 잘 표현할 수 있는 적절한 양의 차분데이터들만 사용하여 다중 퍼지 예측시스템을 구성할 T-S퍼지 모델들을 구현하였다. 이를 위하여 적합한 차분간격 값, 즉 원시계열 데이터의 특징을 잘 표현할 수 있는 차분데이터의 생성을 위해 다음과 같이 주어지는 자기상관함수를 이용한다.

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

여기서, N 은 훈련 데이터의 길이이고, j 는 차분간격 값이다. 또한, $y(i)$ 는 i 번째 훈련 데이터이며, \bar{y} 는 훈련 데이터의 평균이다. 이렇게 구하여진 자기 상관계수 값은 총 $N-1$ 개를 가질 것이다. 이렇게 구하여진 상관계수는 상관 계수 값이 큰 순서대로 나열한 후 인접한 상관 계수 값들 사이의 차를 이용하여 가장 큰 폭의 변화를 보이는 구간에서 절단하여 그 이상이 되는 상관계수 값을 가지는 차분 간격 값들을 선택한다. 이때 문제가 되는 것인 가장 높은 상관계수 값과 두 번째 상관계수 값 사이의 차가 다른 인접한 상관 계수 값들의 차 보다 크다면 선택되는 차분 간격 값은 하나가 될 것이다. 이렇게 되면 다중 퍼지 예측시스템을 구성할 최적의 T-S퍼지 모델이 하나가 되기 때문에 좋은 예측성능을 나타낼 수 없다. 왜냐하면 상관 값이 큰 차분간격을 가지는 차분데이터라고 해서 전체 구간에 대한 원시계열 데이터의 특징을 전부 잘 표현할 수 있는 것은 아니기 때문이다. 따라서 제안된 알고리즘에서는 여러 개의 시계열 데이터를 이용한 시뮬레이션 결과나 기존의 연구결과를 토대로 다중 퍼지 예측 시스템을 구성할 최적의 T-S퍼지 모델 구현에 이용될 차분데이터의 차분 간격 값을 최소한 5개 이상이 되도록 구현하였다. 이를 위해 상관 계수 값이 큰 순서대로 랭킹에 따라 5개의 차분 간격 값을 먼저 선택하고 6번째 상관계수에서부터 전체 상관계수 값들까지 인접한 상관계수의 차를 이용하여 가장 큰 폭의 변화를 보이는 곳에서 절단하여 그이상의 차분 간격 값을 사용하여 최적의 T-S퍼지 모델들을 구현하였다. 이렇게 선별된 차분 간격을 이용하여 다중 퍼지 예측시스템을 구성할 최적의 T-S퍼지 모델들에 대한 각각의 입력으로 사용될 차분 데이터들

을 생성하며 방법은 다음과 같다.

$$\begin{aligned} \Delta_{m(i)}t_0 &= y(t) - y(t - m(i)) \\ \Delta_{m(i)}t_1 &= y(t-1) - y(t-1 - m(i)) \\ &\vdots \\ \Delta_{m(i)}t_n &= y(t-n) - y(t-n - m(i)) \\ &\vdots \\ \Delta_{m(i)}t_{t-m(i)-1} &= y(m(i)+1) - y(1) \end{aligned} \quad (3)$$

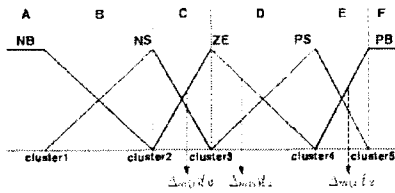
여기에서, $\{m(i)\}$ 는 선별된 차분간격 값들의 집합이고 $\Delta_{m(i)}t$ 는 차분 간격 $m(i)$ 에 대해 시간 t 순간에 생성되는 차분 데이터들이다. 식 (3)에서 보면, 차분 간격 값 $m(i)$ 에 따라 N 개의 훈련 데이터로부터 생성되는 차분 데이터의 수가 달라짐을 알 수 있다. 이는 유효한 데이터 값을 이용하여 차분 데이터를 생성하기 위한 것으로 $y(n)=0, n \leq 0$ 라는 가정을 배제하여 예측 과정에 그로 인한 편향(bias)이 생기는 것을 방지하는 효과를 가져다준다.

이렇게 생성된 각 차분 간격 값에 대한 차분 데이터들은 최적의 T-S 퍼지 모델의 입력으로 사용되며, 또한 T-S 퍼지 규칙 기반의 후건부 선형 수식 파라미터의 식별에 이용된다. 따라서 선별된 차분 값의 개수와 구현된 T-S모델의 개수는 같으며 이 T-S모델들을 이용하여 다중 퍼지 예측시스템을 구현하게 된다. 또한 2.1 절에서 설명 되었듯이 T-S퍼지 모델의 R 번째 규칙을 표현하기 위해서는 입력데이터의 특징을 언어적으로 표현할 수 있는 선건부의 퍼지 분할과 후건부의 선형수식을 위한 파라미터 식별이 필요하다. 먼저 선별된 차분 간격 값의 수가 k 개라면 전건부의 퍼지 분할을 위하여 제안된 논문에서는 선정한 k 개의 차분 간격 $m(i)$ 에 대해 예측이 수행되는 시간 t 에서 가장 최근의 차분 데이터 $\Delta_{m(i)}t_0, \Delta_{m(i)}t_1, \Delta_{m(i)}t_2$ 의 3개를 입력 변수로 하고 이들에 대해 각각 NB, NS, ZE, PS, PB의 5개 퍼지 집합으로 입력 공간을 퍼지 분할하였다. 따라서 차분 간격 $m(i)$ 에 해당하는 j 번째 T-S 퍼지 예측 모델의 언어적 규칙은 다음과 같은 형태가 된다.

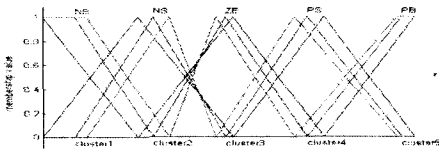
$$\begin{aligned} R_j : & \text{ if } \Delta_{m(i)}t_0 \text{ is } A_j \text{ and } \Delta_{m(i)}t_1 \text{ is } B_j \text{ and} \\ & \Delta_{m(i)}t_2 \text{ is } C_j \\ & \text{ then } \hat{y}_t^j = a_0^j \Delta_{m(i)}t_0 + a_1^j \Delta_{m(i)}t_1 + a_2^j \Delta_{m(i)}t_2 + e^j \end{aligned} \quad (4)$$

여기서 후건부의 출력은 $\hat{y}_t^j = \hat{y}^j(t+p) - y(t)$ 로서 현재 순간의 데이터 $y(t)$ 와 j 번째 규칙에 의해 결정되는 예측하고자 하는 p 스텝 앞의 예측 값 $\hat{y}^j(t+p)$ 와의 차분 값이다. 다시 말해 과거와 현재의 데이터로부터 예측되는 미래 값의 증가분의 예측 값이다. 또한, 퍼지 예측기의 입력 변수

$\Delta_{m(i)}t_0, \Delta_{m(i)}t_1, \Delta_{m(i)}t_2$ 에 대해 다음과 같이 퍼지 영역으로 분할한다. 먼저 N 개의 훈련 데이터로부터 차분 간격 $m(i)$ 에 대해 생성된 차분 값 $\Delta_{m(i)}N_0, \dots, \Delta_{m(i)}N_{N-m(i)-1}$ 의 최소값과 최대값 사이를 퍼지 분할의 전체 영역(universe of discourse)으로 하고, 입력 변수 데이터에 대해 k-means 클러스터링 알고리즘을 이용하여 구한 클러스터 중심 값을 삼각형 소속 함수의 중심 값으로 채택하여 그림 3과 같이 NB, NS, ZE, PS, PB의 5개 퍼지 집합으로 분할하였다.



(a) k-means 클러스터링에 의한 퍼지 집합



(b) k-means 클러스터링 알고리즘 실행 예
 그림 3 입력 공간 퍼지 분할

그림 3의 b)는 k-means 알고리즘을 이용해 차분 데이터에 적합화 시킨 소속 함수의 한 예로서, 클러스터링이 진행됨에 따라 소속 함수가 계속 바뀌는 것을 볼 수 있다. 이러한 입력 공간의 퍼지 분할은 다중 퍼지 예측 시스템을 구성하는 k 개의 퍼지 모델 각각에 대해 독립적으로 수행된다. 각각의 퍼지 모델마다 사용되는 차분 데이터가 다르므로 k-means 알고리즘에 의해 구해지는 퍼지 집합의 소속 함수도 당연히 달라진다. 그림 3의 a)와 같은 소속 함수에 대해서 퍼지 모델의 입력 값이 A, F 구간에 위치할 경우에는 하나, B, C, D, E 구간에 위치할 경우에는 두 개의 퍼지 집합에 속하게 된다. 따라서 입력 값 x 에 대한 소속 함수 값은 다음과 같이 구할 수 있다.

$$\begin{aligned} \text{A, F 구간} : \mu_{NB}(x) \text{ or } \mu_{PB}(x) &= 1 \\ \text{B, C, D, E 구간} : \mu_L(x) &= \frac{C_R - x}{C_R - C_L} \\ \mu_R(x) &= \frac{x - C_L}{C_R - C_L} \end{aligned} \quad (5)$$

여기서 $\mu_L(x)$ 는 입력 값 x 이 속한 구간의 왼쪽에 위치한 클러스터 중심 값 C_L 을 중심 값으로 하

는 퍼지 집합에 대한 소속 함수 값을 나타내며, $\mu_R(x)$ 는 오른쪽에 위치한 클러스터 중심 값 C_R 을 중심 값으로 하는 퍼지 집합에 대한 소속 함수 값을 나타낸다. 예를 들어, 그림 3(a)에서 입력 값이 $x = \Delta_{m(i)}t_1$ 인 경우, $C_R = \text{cluster4}$ 가 되고 $\mu_L(x) = \mu_{ZE}(x)$, $\mu_R(x) = \mu_{PS}(x)$ 가 된다. 이렇게 분할된 퍼지 집합을 이용하여 그림 3의 a)와 같은 입력에 대한 퍼지 규칙은 다음과 같이 8개를 생성할 수 있다.

- R_1 : if f_0 is NS and f_1 is ZE and f_2 is PS then
- R_2 : if f_0 is NS and f_1 is ZE and f_2 is PB then
- R_3 : if f_0 is NS and f_1 is PS and f_2 is PS then
- R_4 : if f_0 is NS and f_1 is PS and f_2 is PB then
- R_5 : if f_0 is ZE and f_1 is ZE and f_2 is PS then
- R_6 : if f_0 is ZE and f_1 is ZE and f_2 is PB then
- R_7 : if f_0 is ZE and f_1 is PS and f_2 is PS then
- R_8 : if f_0 is ZE and f_1 is PS and f_2 is PB then

(6)

여기서 f_0 는 $\Delta_{m(i)}t_0$, f_1 은 $\Delta_{m(i)}t_1$, f_2 는 $\Delta_{m(i)}t_2$ 이며 규칙생성을 위한 하나의 입력집합을 만들게 된다. 이러한 규칙은 훈련 데이터로부터 생성된 차분 값 $\Delta_{m(i)}N_0, \dots, \Delta_{m(i)}N_{N-m(i)-1}$ 의 전체에 대하여 하나씩 증가하면서 입력집합을 형성하게 되며, 각각의 입력집합들에 대하여 중복 생성되는 규칙은 제거하면서 T-S퍼지 모델의 규칙 기반을 만들면 된다. 또한 이 과정은 각각의 T-S퍼지 모델들에 대하여 독립적으로 수행된다.

마지막으로 생성된 각각의 규칙들에 대한 파라미터 식별을 위해 LSM(least square method)방법을 이용하였다. 예를 들어 차분 간격 $m(i)$ 에 대한 i 번째 T-S 퍼지 예측기의 j 번째 퍼지 규칙 R_j 의 생성에 기여한 입력집합이 n 개로 이루어져 있다면, 이들에 의한 후건부 선형 식은 다음과 같이 n 개의 연립 방정식이 된다.

$$\begin{aligned} \widehat{\nabla_{t_1}^j} &= \alpha_0^j \Delta_{m(i)}t_1^j + \alpha_1^j \Delta_{m(i)}t_1^j + \alpha_2^j \Delta_{m(i)}t_1^j + e^j \\ \widehat{\nabla_{t_2}^j} &= \alpha_0^j \Delta_{m(i)}t_2^j + \alpha_1^j \Delta_{m(i)}t_2^j + \alpha_2^j \Delta_{m(i)}t_2^j + e^j \\ &\vdots \\ \widehat{\nabla_{t_n}^j} &= \alpha_0^j \Delta_{m(i)}t_n^j + \alpha_1^j \Delta_{m(i)}t_n^j + \alpha_2^j \Delta_{m(i)}t_n^j + e^j \end{aligned} \quad (7)$$

이를 벡터-행렬식으로 표현하면

$$\begin{bmatrix} \widehat{\nabla_{t_1}^j(1)} \\ \widehat{\nabla_{t_2}^j(2)} \\ \vdots \\ \widehat{\nabla_{t_n}^j(n)} \end{bmatrix} = \begin{bmatrix} \Delta_{m(i)}t_0^j(1) & \Delta_{m(i)}t_1^j(1) & \Delta_{m(i)}t_2^j(1) & 1 \\ \Delta_{m(i)}t_0^j(2) & \Delta_{m(i)}t_1^j(2) & \Delta_{m(i)}t_2^j(2) & 1 \\ \vdots & \vdots & \vdots & \vdots \\ \Delta_{m(i)}t_0^j(n) & \Delta_{m(i)}t_1^j(n) & \Delta_{m(i)}t_2^j(n) & 1 \end{bmatrix} \begin{bmatrix} \alpha_0^j \\ \alpha_1^j \\ \alpha_2^j \\ e^j \end{bmatrix} \quad (8a)$$

$$Y_j = X_j \Theta_j \quad (8b)$$

여기서 Y_j 는 출력 벡터, X_j 는 입력 데이터 행렬, Θ_j 는 계수 벡터이다. 식(8.b)의 계수 벡터 Θ_j 는 최소 자승법을 이용하여 다음과 같이 구할 수 있다.

$$\hat{\Theta}_j = (X_j^T X_j)^{-1} X_j^T Y_j \quad (9)$$

식(9)과 같이 데이터로부터 구한 추정 계수는 다음과 같이 정의되는 오차 파위의 합을 최소화하는 최적해이다.

$$E_j = (Y_j - X_j \hat{\Theta}_j)^T (Y_j - X_j \hat{\Theta}_j) \quad (10)$$

이렇게 하여 총 q 개의 퍼지 규칙을 갖는 퍼지 예측기의 입력 X_j 에 대한 출력 \hat{Y}_j 은 각각의 규칙 R_j 의 전건부에서 결정되는 직함도 μ_j 와 후건부에서 (4)를 이용하여 결정된 \hat{V}_j^i 로부터 가장 함으로 구해지고, \hat{V}_j^i 은 시간 t 에서 예측된 미래 값의 증가분이므로 원하는 시계열의 예측 값 $\hat{y}(t+p)$ 은 다음과 같이 구할 수 있다.

$$\hat{V}(t) = \frac{\sum_{i=1}^q \mu_i \hat{V}_i^j}{\sum_{i=1}^q \mu_i}, \quad \hat{y}(t+p) = y(t) + \hat{V}(t) \quad (11)$$

3.2 모델 선택 및 예측

구현된 다중 퍼지 예측 시스템을 구성하고 있는 최적 T-S 퍼지 모델들은 훈련데이터에 대한 예측 성능을 검증하여 최종 예측을 수행할 T-S퍼지 모델을 선택한 후 최상의 성능을 나타내는 T-S퍼지 모델을 이용하여 나머지 예측 데이터들의 예측을 수행하게 된다. 이런 최상의 T-S퍼지 모델을 선택하기 위한 모델선택의 성능지표로는 다음과 같은 자승 오차 평균(Mean Square Error : MSE)를 이용하였다.

$$MSE = \frac{1}{N-m(i)} \sum_{n=m(i)+1}^N (y(n) - \hat{y}(n))^2 \quad (12)$$

여기서, $\hat{y}(n)$ 은 퍼지 예측기에서 출력한 $y(n)$ 의 예측 값이며, 차분 간격 $m(i)$ 인 퍼지 예측기 모델에서는 예측에 사용되는 차분 데이터가 $N-m(i)$ 개이므로 위와 같이 평균이 구해진다. 그러므로 최종 예측을 수행하는 T-S퍼지 모델은 최적의 T-S 퍼지 모델로 구성된 다중 퍼지 예측시스템에서 MSE가 최소인 오직 하나의 T-S퍼지 모델이 선택되어 예측을 수행하게 된다.

4. 시뮬레이션

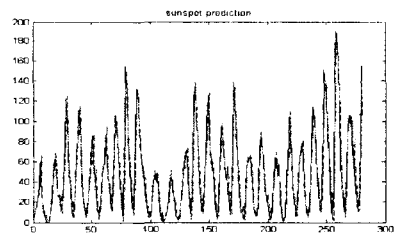
제안된 기법의 유효성을 검증하기 위하여 컴퓨터 시뮬레이션을 수행하였다. 시뮬레이션 예로는 비정상 및 혼돈 비선형 특성을 갖는 대표적인 시계열 데이터를 선택하여 예측 결과를 분석, 검토해 보고 기존의 연구 결과들과 성능을 비교하였다.

첫 번째 사례는 1700년도부터 1979년 까지 관측된 태양흑점 데이터[13]를 이용하여 제안된 알고리즘의 성능을 검증하였다.

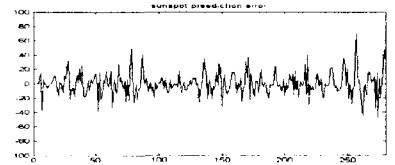
데이터의 총 길이는 280개이며, 이중 1700년도에서 1920년까지의 데이터를 훈련데이터 및 검증 데이터로 사용하였으며, 1921- 1956, 1957-1979년 까지 두 단계로 분류하여 예측을 수행하였다. 성능 평가를 위한 성능평가 지수로는 다음과 같이 정의되는 NMSE(normalized mean square error)를 이용하였다.

$$NMSE = \frac{\sum_{i=1}^L (y(i) - \hat{y}(i))}{\sum_{i=1}^L (y(i) - \bar{y})} \quad (13)$$

여기에서, $y(i)$ 는 실제의 데이터이고 $\hat{y}(i)$ 는 예측된 데이터이다. 또한, L 은 예측구간의 데이터의 길이이고, \bar{y} 는 예측 구간에서 실제 데이터들의 평균을 의미한다.



(a) 흑점 데이터의 예측 결과



(b) 흑점 데이터의 예측 오차

그림 4 제안된 방식의 흑점 데이터 예측 결과

위 그림 4에서, (a)의 검은색 선은 실제 데이터 값들이고 파란색 선이 예측된 값들이다. 그림 4에서 보듯이 제안된 알고리즘의 예측성능이 양호함을 알 수 있다.

아래의 표 1은 제안된 알고리즘의 성능을 Tong and Lim(1980), Weigend(1990) 과 Geva(1999)가 제안했던 방법들과 비교한 것이다[7][14].

표 1 기존 기법들과의 예측 성능 비교

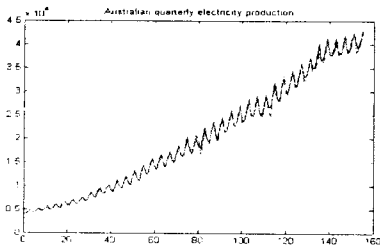
사용 방식	입력 수	퍼지 분할	검증 구간별 RMSE		평균 RMSE
			1921-1956	1957-1979	
			1921	1979	1921-1979
Tong and Lim	12	-	0.097	0.28	0.1885
Weigend	12	-	0.086	0.35	0.218
Geva	11	8	0.0803	0.1211	0.1007
ours	3	5	0.1461	0.1075	0.1268

표 1을 살펴보면 제안된 알고리즘의 성능이 구간별로 예측성능이 크게 차이가 나지 않음을 알 수 있다. 이는 비정상 상태의 원 시계열 데이터가 가지고 있는 비선형적 특성을 데이터의 전처리 과정을 통해 평균이 안정화된 적당한 차분 데이터들과 부정확성을 최소화 하는 퍼지 모델 구현 방법에 의한 최적의 T-S퍼지 모델들로 구성된 예측 시스템을 이용함으로써 일관적인 예측 결과를 나타낼 수 있음을 보인다.

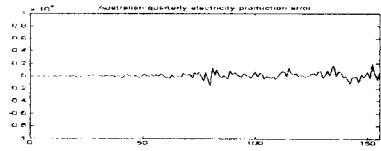
또한, 제안된 알고리즘은 다른 알고리즘에 비해 적은 수의 입력데이터를 이용하면서도 퍼지분할을 최적화시킴으로써 좋은 성능을 나타내었다. 특히 Geva에 의해 제안된 알고리즘에 비해 예측성능은 조금 떨어지지만 입력 개수는 훨씬 적은 수의 데이터를 사용하였으며, 또한 퍼지 규칙을 생성하는 퍼지집합의 개수 또한 적게 사용하여 유사한 예측 결과를 도출하였으므로 예측모델의 구현이 Geva의 방식보다 용이하면서도 효율적이다.

두 번째 시뮬레이션 예는 호주의 전력생산량 데이터[13]를 이용한 것이다. 호주의 전력생산량 데이터는 총 155개의 데이터로 구성되어 있으며, 이 중 70개의 데이터를 훈련 및 검증데이터로 나머지 데이터를 예측데이터로 사용하였다.

아래의 그림 5는 호주 전력생산량 데이터에 대한 제안된 알고리즘의 예측결과를 보여준다.



(a) 호주의 전력 생산량 예측 결과



(b)호주의 전력 생산량 예측 오차
그림 5 제안된 방식의 호주 전력 생산량 예측

그림 5의 (a)는 제안된 알고리즘의 예측결과로, 검은색 선은 실제의 호주 전력생산량을 나타내며 파란색 선은 제안된 알고리즘의 예측 값이다. 또한 그림 (b)은 제안된 알고리즘의 예측 오차이다. 그림 5에서 보듯 제안된 알고리즘의 예측성능이 양호함을 알 수 있다. 제안된 알고리즘의 성능을 비교하기 위하여 성능지표로는 다음과 같이 정의 되는 MRE(mean square error)를 이용하였다.

$$MRE = \frac{1}{L} \sum_{i=1}^L \left(\frac{|y(i) - \hat{y}(i)|}{|y(i)|} \right) \times 100\% \quad (14)$$

여기에서, $y(i)$ 는 실제의 데이터이고 $\hat{y}(i)$ 는 예측된 데이터이다. 또한, L 은 예측 데이터의 길이이다. 아래의 표 2는 제안된 알고리즘과 다른 알고리즘의 성능을 비교한 것이다[15].

표 2 기존 기법들과의 예측 성능 비교

	Mamdani 퍼지 모델	다중 퍼지 모델	Fuzzy AR	GA-RS 방식	ours
MRE	7.8123	2.7125	3.1254	1.8100	1.7077

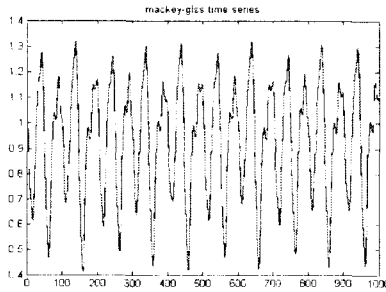
표 2에서 보듯이 제안된 알고리즘이 다른 알고리즘에 비해 성능이 우수함을 알 수 있다. 표 2를 살펴보면 제안된 알고리즘이 원 시계열의 시간 지연 데이터를 이용하여 구현된 다중 퍼지 모델보다 성능이 우수함을 알 수 있다. 이것은 제안된 방법이 비정상 시계열의 예측에 있어 불확실성과 비선형성에 좀 더 능동적으로 대처할 수 있음을 나타낸다. 또한 최근에 발표된 GA-RS방식은 퍼지모델에 유전자 알고리즘과 러프집합 이론을 포함한 방식을 선택하여 다른 알고리즘들에 비해 좋은 예측 성능을 나타내었지만 원시계열을 그대로 사용함으로써 예측 모델이 비정상 시계열의 특성을 잘 반영하지 못하여 제안된 방식에 비해 예측성능이 좋지 못함을 보인다. 뿐만 아니라 모델의 구현에 있어서도 제안된 방식이 좀 더 간편하며 효율적임을 알 수 있다.

마지막으로 시뮬레이션을 수행한 데이터는 비정상 시계열 데이터의 검증에 빈번하게 쓰이는 mackey-glass시계열 데이터로 다음과 같이 정의되

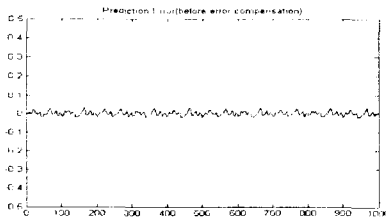
는 식에 의하여 생성된다.

$$\frac{dx(t)}{dt} = \frac{ax(t-\tau)}{1+x^{10}(t-\tau)} - bx(t) \quad (15)$$

위 식의 파라미터 값으로 $a=0.2$, $b=0.1$, $\tau=17$ 을 사용하였다. 데이터는 $x(124)$ 부터 $x(1123)$ 까지 1000개의 데이터를 추출하여 500개의 데이터를 훈련 데이터 및 검증데이터로, 나머지를 예측 데이터로 사용하였다. 그림 6은 제안된 알고리즘이 예측을 수행한 결과이다. 그림 (a)에서 검은색 선이 실제의 데이터 값이고 파란색 선이 제안된 알고리즘의 예측 값이다. 또한 그림 (b)는 예측 오차를 나타낸다.



(a) Mackey-Glass 시계열 예측 결과



(b) Mackey-Glass 시계열 예측 오차

그림 6 제안된 방법의 Mackey-Glass 예측

아래의 표 3은 이 시계열 예측에 대해 Wang과 Mendel의 방법[16], MCM-1과 MCM-2[17]와 제안된 방법의 성능을 비교한 것으로서 성능평기 지수는 다음과 같이 정의되는 RMSE(root mean square error)을 이용하였다.

$$RMSE = \sqrt{\frac{1}{L} \sum_{i=1}^L (y(i) - \hat{y}(i))^2} \quad (16)$$

수식의 표기는 이전 시뮬레이션들과 같은 표기를 사용하였다. 표 3을 살펴보면 제안된 방식이 훨씬 적은 수의 퍼지 집합을 이용하여 보다 좋은 예측 성능을 나타냄을 알 수 있다.

표 3 기존 기법들과의 예측 성능 비교

입력 변수 퍼지 집합	RMSE			
	WANG	MCM-1	MCM-2	제안된 방법
5	-	-	-	0.0110
7	0.0372	0.0374	0.0340	-
11	0.0253	0.0265	0.0235	-
15	0.0191	0.0197	0.0187	-
19	0.0161	0.0162	0.0159	-
23	0.0135	0.0142	0.0131	-
27	0.0115	0.0115	0.0113	-
31	0.0108	0.0108	0.0105	-

이러한 결과는 제안된 방법의 퍼지 분할이나 파라미터 추정, 그리고 비정상 시계열 데이터의 특성들을 잘 반영하는 적당한 차분 데이터를 이용한 최적 T-S 퍼지 모델기반 다중 퍼지 예측시스템의 구현이 비정상 시계열 예측에 효율적이며 유용할 수 있음을 보이는 것이다.

5. 결론

비정상 또는 혼돈 시계열에 대한 예측은 시계열 데이터가 가지고 있는 불확실성과 비선형적 특성들로 인해 많은 어려운 문제점들이 존재한다. 특히 예측 모델을 구현함에 있어 이러한 시계열의 특성과 모델파라미터들의 부정확성을 잘 반영하지 못할 경우 예측결과는 좋을 수가 없다. 따라서 이러한 문제를 어떻게 적절히 취급하는가에 따라 예측 성능이 크게 다르게 나타날 것이다.

제안된 방법은 비정상 시계열의 특징을 잘 표현할 수 있는 적절한 차분데이터들과 T-S퍼지 모델의 입력 공간 분할과 파라미터 추정에 k-means 클러스터링 알고리즘과 LSM(least square method)를 이용하여 최적의 T-S퍼지 모델들을 구현하고 이를 기반으로 다중 퍼지 예측시스템을 구성하여 성능을 검증한 후, 검증된 성능을 이용하여 최종 예측을 수행할 최상의 T-S퍼지 모델을 선택하여 예측을 수행하도록 하였다. 시뮬레이션 결과는 제안된 방식이 비정상 또는 혼돈 시계열 데이터에 대해서도 시계열 데이터에 내재된 불확실성과 비선형성에 능동적으로 대처할 수 있으며, 또한 예측의 부정확성을 최소화 할 수 있어 좋은 예측 결과를 나타내었음을 보인다. 그러므로 제안된 방법은 비정상 시계열이나 혼돈 시계열의 예측 문제에 있어 매우 효과적으로 사용될 수 있을 것이다.

향후 연구과제로는 다중 퍼지 예측시스템을 구성하고 있는 최적의 T-S퍼지 모델의 구현을 위하여 평균이나 분산과 같은 시계열의 통계적 특성을 좀 더 잘 반영할 수 있는 방법이나 파라미터 추정에서 발생하는 부정확성을 좀 더 감소시킬 수 있

는 방법에 관한 연구가 필요할 것으로 사료된다.

참 고 문 헌

- [1] George E. P. Box and Gwilym M. Jenkins, "Time Series Analysis : Forecasting and Control", *Holden-Day*, 1970.
- [2] George J. Klir and Bo Yuan, "Fuzzy Sets and Fuzzy Logic Theory and Applications", *Prentice-Hall*, 1995.
- [3] 김해경, 김태수, "시계열 분석과 예측 이론", 문운당, 2003.
- [4] K.Ozawa, T.Niimura, "Fuzzy Time-Series Model of Electric Power Consumption", *IEEE Canadian Conference on Electrical and Computer Engineering*, pp.1195-1198, 1999.
- [5] 진강규, "유전자 알고리즘과 그 응용", 교우사, 2000.
- [6] Dai-jin Kim, Chul-hyun Kim, "Forecasting Time Series with Genetic Fuzzy Predictor Ensemble". *IEEE Trans. on Fuzzy Systems*, Vol.5, pp.523-535, 1997.
- [7] Juhong Nie, "Nonlinear Time-Series Forecasting: A Fuzzy Neural Approach", *Neuro Computing*, Vol.16, pp.63-76, MacMaster University, 1997.
- [8] Yang Gao, Meng Joo Er, "NARMAX Time Series Model Prediction: Feedforward and Recurrent Fuzzy Neural Network Approaches", *Fuzzy Sets and Systems*, Vol 150, pp. 331-350, 2005.
- [9] Inteak Kim, Song-Rock Lee, "A Fuzzy Time Series Prediction Method Based on Consecutive Values", *IEEE International Conerence on Fuzzy Systems*, Vol.2, pp.703-707, 1999.
- [10] Chul-Heui Lee, Sang-Hun Yoon, "Fuzzy Nonlinear Time Series Forecasting with Data Preprocessing and Model Selection", *Journal of Telecommunications and Information*, Vol.5, pp.232-238, 2001.
- [11] Carl G. Looney, "Pattern Recognition using Neural Networks", *Oxford University Press*, 1997.
- [12] Stephen J. Redmond. Conor Heneghan, "A Method for Initialising the K-means Clustering Algorithm using kd-Trees", *Pattern Recognition Letters*, Vol. 28, pp. 965-973, 2007.
- [13] <http://www-personal.buseco.monash.edu.au/hyndman/TSDL/>
- [14] Amir B. Geva, "Hierarchical - Fuzzy Clustering of Temporal-Patterns and Its Application for Time-Series Prediction", *Pattern Recognition Letters*, Vol. 20, pp. 1519 -1532, 1999.
- [15] 주용석, "유전알고리즘과 러프집합을 이용한 퍼지시스템 모델링", *강원대학교 석사학위논문*, 2003.
- [16] L. X. Wang, J. M. Mendel, "Generating Fuzzy Rules from Numerical Data, with Applications", *IEEE Trans. on Systems, Man, and Cybernetics*, Vol. 22 No.6, pp. 1414-1427, 1992.
- [17] 김인택, 공창욱, "시계열 예측을 위한 퍼지 학습알고리즘", *한국 퍼지 지능시스템 학회*, Vol.7, No.3, pp34-42, 1997.