

## 예측치 결합을 위한 PNN 접근 방법

전덕빈<sup>1</sup> · 신효덕<sup>1</sup> · 이정진<sup>2</sup>

<sup>1</sup>KAIST 테크노경영대학원 / <sup>2</sup>University of Hartford, Department of Computer Science PH.D

### A PNN approach for combining multiple forecasts

Duk-Bin Jun<sup>1</sup> · Hyo-Duk Shin<sup>1</sup> · Jung-Jin Lee<sup>2</sup>

In many studies, considerable attention has been focussed upon choosing a model which represents underlying process of time series and forecasting the future. In the real world, however, there may be some cases that one model can not reflect all the characteristics of original time series. Under such circumstances, we may get better performance by combining the forecasts from several models.

The most popular methods for combining forecasts involve taking a weighted average of multiple forecasts. But the weights are usually unstable. In cases the assumptions of normality and unbiasedness for forecast errors are satisfied, a Bayesian method can be used for updating the weights. In the real world, however, there are many circumstances the Bayesian method is not appropriate. This paper proposes a PNN(Probabilistic Neural Net) approach as a method for combining forecasts that can be applied when the assumption of normality or unbiasedness for forecast errors is not satisfied.

In this paper, PNN method, which is similar to Bayesian approach, is suggested as an updating method of the unstable weights in the combination of the forecasts. The PNN method has been usually used in the field of pattern recognition. Unlike the Bayesian approach, it requires no assumption of a specific prior distribution because it gets probabilities by using the distribution estimated from given data. Empirical results reveal that the PNN method offers superior predictive capabilities.

#### 1. 서론

시계열 예측모형은 시계열이 갖는 특성들 가운데 어떤 특성을 고려하는가에 따라 여러 가지 형태의 모형이 존재할 수 있다. 기존의 많은 연구들은 주어진 시계열에 적합한 몇 가지 모형 가운데 시계열의 내재적 과정(underlying process)을 가장 잘 표현하는 하나의 모형을 선택하고 미래를 예측하는 데 관심을 가져왔다. 그러나 원시계열(original time series)의 내재적 과정은 정확하게 파악하기 어려울 정도로 복잡한 특성이 있기 때문에 내재적 과정과 가장 유사한 모형일지라도 하나의 예측모형이 자료의 특성을 모두 반영할 수는 없는 경우가 존재한다(Newbold and Granger, 1974). 이와 같은 상황에서는 여러 모형으로부터 얻어진 예측치를 조합하여 자료의 다양한 특성을 반영함으로써 개별 모형들의 설명력 이상의 효과를 얻을 수 있다.

기존의 예측치 조합에 관한 연구들은 다양한 방법으로 시도 되어 왔다. 특히, 두 예측모형의 통계적 방법에 의한 선형결합

을 통하여 예측치를 얻는 방법이 많이 연구되어져 왔다. Kang(1986)에 의하면 여러 개의 모형에서 나온 예측치들을 선형결합할 때 각 모형의 예측치에 대한 weight로 사용되는 가중치들이 대부분 불안정하게 나타난다. 이러한 가중치의 불안정성은 매 시점마다 베이지안 방법(Diebold and Pauly, 1990; Min and Zellner, 1993)으로 가중치를 갱신함으로써 해결해 줄 수 있다. 그러나 베이지안 방법은 각 모형의 예측오차에 대한 분포가 가정되어야만 하는 한계점이 있다. 반면에 현실적으로 시계열 자료의 예측오차는 특정 분포를 따르지 않거나 편의가 있는 경우도 많으므로 이와 같은 상황에 적용할 수 있는 새로운 예측치 조합 방법이 필요하다.

이에 본 논문에서는 시계열 자료의 예측오차가 특정 분포를 따르지 않거나 편의가 존재하는 경우에 베이지안 방법을 적용할 수 없는 한계점을 지적하고, 그 대안으로서 사후확률을 이용한 예측치들의 선형 결합을 통하여 예측치를 조합하는 PNN(Probabilistic Neural Net) 방법을 제시하였다. PNN 방법은 주로 유형인식(pattern recognition)에 활용되어 온 방법으로서

자료로부터 분포를 추정하여 추정된 분포에서 확률값을 얻기 때문에 베이지안 접근방법에 비하여 사전에 특정 분포를 가정할 필요가 없는 장점이 있다. 또한, 본 논문에서는 PNN 방법의 장점을 확인하기 위하여 실제 자료에 베이지안 방법과 PNN 방법을 적용시켜 예측치를 조합하고 예측오차가 어떠한 형태를 가질 때 어떤 방법이 더 우세한지를 비교 분석하였다.

## 2. 예측치 결합 모형

이 절에서는 여러 개의 모형에서 나온 예측치들을 결합하는 접근 방법으로서 베이지안 모형과 PNN 모형의 비교를 위하여 두 모형에 대해 자세히 설명하고자 한다.

### 2.1 베이지안 모형

베이지안 결합 모형은 매 시점마다 시계열 자료가 특정 모형을 따른다는 조건하에서 조건부 확률을 구하고 2개 이상의 모형에 대한 조건부확률로부터 얻어진 우열비율을 통하여 사후확률을 구해 예측치들을 결합하는 방법이다(Min and Zellner, 1993).

예를 들어 시계열자료가 정규분포를 따르고 2개의 예측모형, 모형1과 모형2가 있을 때 베이지안 방법을 이용하여 예측치를 결합시키는 방법은 다음과 같다.

$$f_y = P_1 \times f_{y|I_1} + (1 - P_1) \times f_{y|I_2} \quad (1)$$

이 때,

- $y$ : 자료(data)
- $f_y$ : 결합 모형의 예측치
- $f_{y|I_1}$ : 모형1의 예측치
- $f_{y|I_2}$ : 모형2의 예측치
- $I_1$ : 모형1 하에서의 사전정보(prior information)
- $I_2$ : 모형2 하에서의 사전정보(prior information)
- $P_1$ : 내재적 과정이 모형1일 확률

내재적 모형이 모형1일 확률을 우열비율을 이용하여 구하면 다음과 같다.

$$P_1 = \frac{\Pr(M_1 | y, I_1)}{\Pr(M_1 | y, I_1) + \Pr(M_2 | y, I_2)} \quad (2)$$

$\Pr(M_1 | y, I_1)$ : 자료 및 모형1하에서의 사전정보가 주어졌을 때 모형1을 택할 확률

$\Pr(M_2 | y, I_2)$ : 자료 및 모형2하에서의 사전정보가 주어졌을 때 모형2를 택할 확률

여기서  $M_1$ 은 모형1을,  $M_2$ 는 모형2를 나타낸다.

이 때, 우변의 분자와 분모를  $\Pr(M_1 | y, I_1)$ 로 나누면,  $P_1 = \frac{1}{1 + K}$ 이다.

$K$ 는 우열비율에 해당하며 다음과 같이 분해된다.

$$\begin{aligned} K &= \frac{\Pr(M_1 | y, I_1)}{\Pr(M_2 | y, I_2)} \\ &= \frac{\Pr(M_1 | I_1) \times f(y | M_1, I_1)}{\Pr(M_2 | I_2) \times f(y | M_2, I_2)} \end{aligned} \quad (3)$$

여기서  $\frac{\Pr(M_1 | I_1)}{\Pr(M_2 | I_2)}$ 는 두 모형에 관한 사전확률분포비율(prior odd)이며,  $f(y | M_1, I_1)$ 와  $f(y | M_2, I_2)$ 은 다음과 같이 시계열 자료가 정규분포를 따른다는 가정으로부터 구할 수 있다.

$$\begin{aligned} y | M_1, I_1 &\sim N(f_{y|M_1, I_1}, \sigma_{y|M_1, I_1}) \\ y | M_2, I_2 &\sim N(f_{y|M_2, I_2}, \sigma_{y|M_2, I_2}) \end{aligned} \quad (4)$$

베이지안 결합 모형은 내재적 과정을 주어진 모형들이 모두 포함하고 있을 경우에는 예측오차의 제곱합을 최소로 만들지만, 주어진 모형들이 내재적 과정을 모두 포함하고 있지 못할 경우에는 그렇지 않을 수 있다(Min and Zellner, 1993). 이 경우에는 주어진 모형 중에 한 모형만을 선택해 주는 것이 예측오차의 제곱합을 최소로 해 줄 수 있다. 따라서 본 논문에서는 예측치의 결합 방법 이외에도  $K$ 가 1보다 클 경우에는 모형1의 예측치를 선택하고 그렇지 않은 경우에는 모형2의 예측치를 선택하는 예측치의 선택 방법도 함께 제안한다.

### 2.2 Probabilistic Neural Network 모형

PNN 모형(Specht, 1990)은 패턴 인식에 주로 사용된 방법으로서 비모수 분포 추정을 통하여 얻어진 확률 분포와 각 패턴에 해당하는 주어진 손실을 이용하여 기대손실을 최소화하는 모형이다. 이 모형은 측정형(metric)이나 범주형(category) 변수를 사용하여 범주형 변수를 예측하는 데 사용된다. 예측치 조합 모형에서는 PNN을 통하여 각 특정 모형의 과거의 예측오차 및 기타 다른 설명변수들을 이용하여 매 시점 내재적 과정과 유사한 모형이 어떤 모형인지를 예측하게 된다.

본 논문에서 예측치 결합모형으로 제안한 PNN 모형을 간단한 예를 들어 설명하면 다음과 같다(<그림 1> 참조). 예측모형은 모형1과 모형2 두 개이며, 시점 1에서 시점 4까지 매 시점마다 각 모형을 사용하여 구한 예측치(one-step ahead forecast)와 실제치를 비교하여 얻은 예측오차가 있다고 하자. 현재, 시점 5에서 모형1과 모형2의 예측치 중 어느 것을 선택할 것인지를 PNN 모형을 통해 결정하려고 한다.

<그림 1>에서 설명변수(input)는 현재 시점에서 가장 최근에 도출된 예측오차, 즉 시점 4에서 각 모형으로부터 얻은 예측오차들이 되고, 그 이전 시점의 예측오차들은 유형화 단계(pattern unit)로 가는 weight로 사용한다. 만약, 시점 1에서는 모

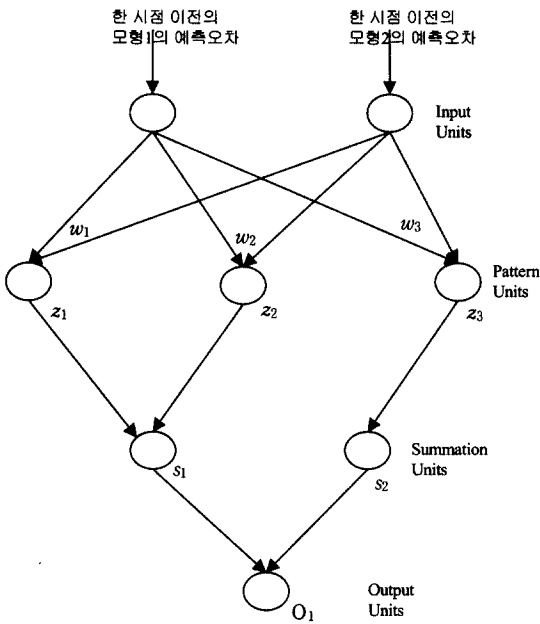


그림 1. PNN의 구조(Specht, 1990).

형2의 예측오차가 상대적으로 작고 시점 2와 3에서는 모형1의 예측오차가 작았다면, 이 세 시점의 과거 예측오차 자료를 사용하여 다음과 같이 weight를 구성한다.

- $W_1 = (w_{11}, w_{12})$  : 시점 3에서의 모형1, 2에서 얻은 예측오차들로 구성된 벡터
  - $w_{11}$  : 시점 3에서의 모형1의 예측오차
  - $w_{12}$  : 시점 3에서의 모형2의 예측오차
- $W_2 = (w_{21}, w_{22})$  : 시점 2에서의 모형1, 2에서 얻은 예측오차들로 구성된 벡터
  - $w_{21}$  : 시점 2에서의 모형1의 예측오차
  - $w_{22}$  : 시점 2에서의 모형2의 예측오차
- $W_3 = (w_{31}, w_{32})$  : 시점 1에서의 모형1, 2에서 얻은 예측오차들로 구성된 벡터
  - $w_{31}$  : 시점 1에서의 모형1의 예측오차
  - $w_{32}$  : 시점 1에서의 모형2의 예측오차

이 때, input(설명변수)로 사용한 시점 4에서의 각 모형에서 얻은 예측오차들로 구성된 벡터를 X라고 하면, 벡터 X와 가중치 벡터(weight vector),  $W_1, W_2, W_3$ 들과의 거리를 다음과 같은 지수함수를 이용하여 변환시켜 유희화 단계(pattern unit)의 값,  $Z_1, Z_2, Z_3$ 를 얻는다.

$$Z_i = \exp \left[ -\frac{(W_i - X)'(W_i - X)}{2\alpha^2} \right], \quad i = 1, 2, 3 \quad (5)$$

이 때,  $X = (x_1, x_2)$  : 시점 4에서 모형1, 2로부터 얻은 예측

오차들로 구성된 벡터

- $x_1$  : 시점 4에서의 모형1의 예측오차
- $x_2$  : 시점 4에서의 모형2의 예측오차

식 (5)에서  $\alpha$ 는 평활화 변수(smoothing parameter)로서 추정되는 분포의 평활화(smooth) 정도를 결정하게 된다. 즉, 자료로부터 도출되는 값이 아니고 분포의 평활화 정도를 고려하여 주어지는 상수이다.

합산 단계(summation unit)에서는 다음과 같이 각 패턴별로  $Z_i$ 를 합한다. 이 때, 패턴이란 모형1의 예측오차가 작은 경우와 모형2의 예측오차가 작은 경우를 구분한 것이다.

$$\begin{aligned} S_1 &= Z_1 + Z_2 \\ S_2 &= Z_3 \end{aligned} \quad (6)$$

마지막으로 최종 결과치,  $O_1$ 은 합산단계에서 얻은 값에  $C_k$ 를 다음과 같이 곱한 다음 합산하여 구한다.

<표 1>에서  $l_1$ 은 내부구조(underlying process)는 모형1임에도 불구하고 모형2를 선택할 경우 발생하는 손실을 의미하고,  $l_2$ 는 반대 상황에서 발생하는 손실을 의미한다.

$$O_1 = S_1 + C_k \times S_2 \quad (7)$$

이 때,

$$C_k = -\frac{h_2 \times l_2}{h_1 \times l_1} \times \frac{n_2}{n_1}$$

- $h_2$  : 내재적 과정(underlying process)이 모형2 일 경우에 대한 사전확률(prior)
- $h_1$  : 내재적 과정이 모형1 일 경우에 대한 사전확률
- $l_2$  : 내재적 과정이 모형2일 때 모형1을 선택한 경우 발생하는 손실
- $l_1$  : 내재적 과정이 모형1일 때 모형2를 선택한 경우 발생하는 손실
- $n_2$  : 이전에 모형2가 모형1에 비해 예측오차가 작았던 경우의 수
- $n_1$  : 이전에 모형1이 모형2에 비해 예측오차가 작았던 경우의 수

최종 결과치,  $O_1$ 이 0보다 큰 경우에는 모형1로 인식을 하게 되며, 그렇지 않은 경우에는 모형2로 인식하게 된다.

표 1. PNN에서의 손신표

실제 \ 선택	모형1	모형2
모형1	0	$l_1$
모형2	$l_2$	0

PNN에서는 비모수 분포 추정을 통해서 패턴 인식을 하기 때문에 그 결과로 사후 확률을 추정할 수 있다. 확률을 도출할 때 베이지안 방법에서는 정규분포와 같은 특정 분포를 가정하는 반면, PNN 방법에서는 특정 분포를 가정하지 않고 자료를 통해서 분포를 추정한 후 추정된 분포로부터 확률을 구한다. 만약, 오차항의 분포가 한쪽으로 치우쳐 있다거나 대칭적이지 않은 경우에는 PNN을 이용하여 분포를 직접 추정함으로써 이런 문제를 해결할 수가 있다. PNN에서 내재적 모형이 모형1일 확률은 다음과 같다.

$$P_1 = \frac{\Pr(M_1|y, I_1)}{\Pr(M_1|y, I_1) + \Pr(M_2|y, I_2)} \quad (8)$$

이 때,

- $\Pr(M_1|y, I_1) \propto \Pr(M_1|I_1) \times \Pr(y|M_1, I_1)$
- $\Pr(M_2|y, I_2) \propto \Pr(M_2|I_2) \times \Pr(y|M_2, I_2)$
- $\Pr(M_1|I_1)$  : Prior, (7)의  $h_1 \times n_1$ 에 해당함
- $\Pr(M_2|I_2)$  : Prior, (7)의  $h_2 \times n_2$ 에 해당함
- $\Pr(y|M_1, I_1) = S_1$
- $\Pr(y|M_2, I_2) = S_2$

PNN 모형에서도 베이지안 모형에서와 같이 식 (8)을 가지고 식 (1)과 같은 방법으로 예측치를 결합시켜 예측력을 높일 수 있다.

베이지안 방법과 PNN 방법에서 사후확률을 도출할 때의 차이점은 베이지안의 경우 각 모형이 옳을 것이라는 가정하에서 자료에 대해 정규분포와 같은 특정 분포를 가정하는 반면에, PNN의 경우 특정 분포를 가정하지 않고 자료로부터 분포를 추정하여 추정된 분포로부터 확률값을 사용한다는 점이다. 반면에 사후확률 분포의 계산 및 가중치의 갱신 방법은 PNN 방법과 베이지안 방법이 동일하다.

### 3. Empirical Results

이 절에서는 예측치 결합 모형으로서 PNN 방법의 장점을 확인하기 위하여 실제 자료에 대해 베이지안 방법과 PNN 방법을 이용하여 예측치를 조합하고 예측오차가 어떠한 형태를 가질 때 어떤 모형이 더 우세한지를 비교 분석하였다.

첫 번째 분석 자료는 Fildes(1992)에 실린 263개의 통신 시계열 가운데 105번째 시계열을 임의로 선택하였으며 이 시계열에 대한 자세한 설명은 구할 수 없었다(<그림 2> 참조). 71개의 자료 가운데 앞부분의 30개 자료를 이용하여 모형을 적합시키고, 나머지 41개 자료는 예측에 사용하였다.

자료를 적합시킬 때 사용한 모형은 무작위 보행 모형(random walk model), 이중 지수 평활 모형(double exponential smoothing model), 감쇠 지수 평활 모형(damped exponential smoothing model),

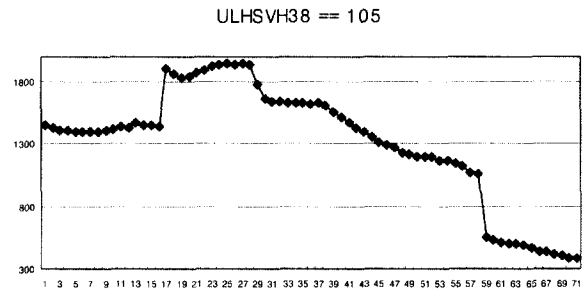


그림 2. Fildes(1992)의 105번째 통신 시계열 자료.

그리고 Robust Trend model이며, 1기 앞 예측치부터 18기 앞 예측치까지를 서로 비교 분석하였다(부록 참조). 오차의 비교 기준으로는 오차제곱평균(MSE), 절대오차평균(MAE) 그리고 상대오차평균(MAPE)을 이용하였다. 앞에서 적합에 사용한 여러 모형 가운데 이중 지수 평활 모형과 Robust Trend 모형의 적합 및 예측력이 대부분 높게 나타났다. 따라서, 예측치를 조합할 때에는 이중 지수 평활 모형과 Robust Trend 모형에서 나온 예측치를 사용하였다.

예측오차의 정규분포에 대한 검정은 적합도 검정(chi-square goodness-of-fit statistic)을 이용하여 이론적인 정규분포와 관측된 상대도수간의 차이를 비교하였는데 부록에서 보는 바와 같이 대부분의 경우에 정규분포를 따른다는 가설이 기각되었다(부록 <표 3> 참조).

예측치 조합 모형으로서는 베이지안 결합모형(bayesian combination model), 베이지안 선택모형(bayesian selection model), PNN 결합모형(PNN combination model) 그리고 PNN 선택모형(PNN selection model)을 사용하였다. 각각에 대한 상대오차평균은 다음 <표 2>와 같다(절대오차평균 및 오차지승평균에 대한 결과는 부록 <표 1>, <표 2> 참조).

<표 2>에서 보는 바와 같이 대부분의 경우 여러 가지 예측치 결합 모형 가운데 PNN 선택모형의 예측력이 우수하였다. 6기 앞, 7기 앞 및 8기 앞 예측치에서는 베이지안 모형의 예측력이 우수하였으나, 이 경우에는 예측오차의 정규분포에 대한 검정결과 정규분포에 가까움을 알 수 있었다.

두 번째 분석 자료는 1998년 1월 3일부터 1998년 11월 24일까지 267개의 국내 91일물 양도성 예금증서(Certificate of Deposit: CD) 이자율이다(<그림 3> 참조). 7월 4일까지의 117개의 자료를 모형 적합에 사용하고 이후의 자료는 예측에 사용하였으며 분석결과는 <표 3>의 (a)와 같다. 적합에 사용된 모형은 무작위 보행 모형, 이중 지수 평활 모형 및 Robust Trend 모형이다. 단기 예측치로 1기 앞 예측치와 5기 앞 예측치를, 중기 예측치로 10기 및 15기 앞 예측치를, 장기 예측치로 20기 앞 예측치를 사용하였다. 분석결과에서 보는 바와 같이 무작위 보행 모형의 예측력이 높은 것을 알 수 있다. 원시계열에 하강하는 추세가 있는 것처럼 보이므로, 장기예측으로 갈수록 Robust Trend 모형의 예측력이 향상되는 것으로 나타났다. 이중 지수 평활 모형은 무작위 보행 모형에 비해 비교적 예측력이 낮게

표 2. 통신 시계열 자료에 대한 예측치의 상대오차평균(%)의 결과

	DES	Damp	RT	RW	Bcom	Bsel	Psel	Pcom
1-step	3.70	4.10	3.60	4.40	3.58	3.74	3.45	3.59
2-step	6.49	7.89	6.72	8.87	6.76	7.39	6.37*	6.50
3-step	10.30	12.12	10.34	13.84	7.07	7.06	6.96	7.86*
4-step	14.00	17.41	13.86	19.15	13.67	13.85	13.48	13.48
5-step	17.95	22.90	17.60	24.70	17.40	17.60	17.29	17.24
6-step	22.05	28.65	20.68	30.70	20.54	20.68*	20.01	20.26
7-step	26.69	34.98	23.96	36.95	23.93	23.96	23.92	24.25
8-step	31.96	41.59	27.43	43.78	27.40	27.43	27.55	27.55
9-step	36.97	48.65	30.58	50.95	30.59	30.58	30.12	30.71
10-step	41.77	56.24	33.38	58.80	33.38	33.38	32.69*	33.55
11-step	45.44	64.64	35.76	67.12	35.83	35.76	34.67**	35.29
12-step	50.48	73.67	39.05	76.15	39.17	39.05	37.72***	37.76**
13-step	52.50	83.24	42.28	85.59	42.45	42.28	36.04*	36.04*
14-step	55.09	88.90	42.71	91.89	43.00	42.71	41.17*	41.17*
15-step	57.40	94.85	43.79	96.90	44.06	43.79	42.07*	42.11
16-step	59.29	100.46	44.76	102.70	44.69	44.76	38.42	38.42
17-step	57.16	106.19	46.27	108.83	46.28	46.27	37.23**	37.23**
18-step	54.11	112.51	48.38	115.39	48.37	48.38	35.16**	35.16**

DES : 이중 지수 평활 모형 Damp : 감쇠 지수 평활 모형 RT : Robust Trend model RW : 무작위 보행 모형 Bcom : Bayesian Combination

Bsel : Bayesian Selection Psel : PNN Selection Pcom : PNN Combination

\* : Robust Trend와의 비교에서 10% 유의수준에서 유의한 경우

\*\* : Robust Trend와의 비교에서 5% 유의수준에서 유의한 경우

\*\*\* : Robust Trend와의 비교에서 1% 유의수준에서 유의한 경우

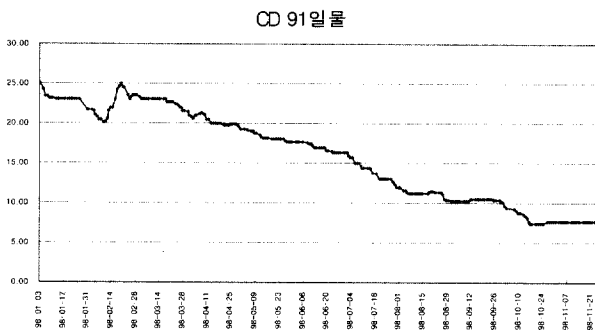


그림 3. 91일물 CD 이자율 자료.

나타났다. 따라서, 예측치 조합 모형은 무작위 보행 모형과 Robust Trend 모형을 사용하여 분석하였다.

주어진 무작위 보행 모형과 Robust Trend 모형의 예측오차의 분포는 정규분포를 따르지 않는 것으로 나타났다(부록 <표 4> 참조). 따라서, 베이지안 모형보다는 PNN 모형을 이용하여 예측치를 조합하는 것이 적절할 것으로 판단된다. 예측치 조합 모형의 분석결과는 <표 3>의 (b)에서 보는 바와 같이 대부분 PNN 모형의 예측력이 베이지안 모형보다 우수했다. 그리고 증기를 제외한 장, 단기 예측에서 PNN 모형의 예측력이 무작위 보행 모형에 비해 우수하였다.

#### 4. 결론 및 향후 연구방향

본 연구에서는 현실적으로 시계열 자료의 예측오차가 특정 분포를 따르지 않거나 편이가 있는 경우에 베이지안 방법을 적용할 수 없는 한계점을 지적하고 그와 같은 경우에 적용할 수 있는 예측치 결합 방법으로서 PNN(Probabilistic Neural Net) 방법을 제시하였다. PNN 방법은 베이지안 방법에 비해 주어진 자료로부터 분포를 추정하여 추정된 분포에서 확률값을 얻기 때문에 사전에 특정 분포를 가정할 필요가 없는 장점이 있다. 실증 분석에서는 통신 시계열과 양도성 예금증서 이자율 자료 각각에 대해서 베이지안 방법과 PNN 방법을 이용하여 매 시점 예측치를 결합 또는 선택한 결과를 비교 분석하였다. 그 결과 예측오차가 특정 분포를 따르지 않는 경우, PNN 방법을 사용하여 예측오차의 분포까지 함께 추정하면서 가중치를 갱신하면 특정 분포를 가정하는 베이지안 방법에 비해서 예측력이 더 우수함을 알 수 있었다. 반면에 예측오차가 정규 분포를 따르면 PNN 방법을 사용하여 예측오차의 분포를 추정하는 과정을 생략하고, 직접 베이지안 방법을 사용하는 것이 예측오차의 제곱합을 더 줄이는 것으로 나타났다.

예측치 결합 모형의 한계는 기존의 예측치들이 고정된 것으로 가정한다는 것이다. 매시점마다 각 모형의 가중치는 갱신

표 3. CD 이자율에 대한 모형 분석 결과

(a) 무작위 보행 모형, 이중 지수 평활 모형 및 Robust Trend 모형 적합 결과

		RW	DES1	DES2	DES3	RT1	RT2	RT3	RT4
1step ahead	MAE	0.081	0.108	0.088	0.084	0.101	0.092	0.088	0.082
	MSE	0.029	0.035	0.029	0.029	0.034	0.026	0.028	0.029
	MAPE	0.753	0.990	0.820	0.790	0.900	0.827	0.814	0.757
5step ahead	MAE	0.386	0.451	0.412	0.421	0.509	0.425	0.405	0.390
	MSE	0.338	0.411	0.350	0.388	0.651	0.401	0.346	0.338
	MAPE	3.740	4.360	4.010	4.130	4.803	4.079	3.951	3.783
10step ahead	MAE	0.748	0.929	0.836	0.900	1.085	0.920	0.829	0.761
	MSE	1.042	1.607	1.248	1.506	2.612	1.628	1.292	1.054
	MAPE	7.514	9.390	8.580	9.240	10.722	9.297	8.587	7.627
15step ahead	MAE	1.034	1.426	1.269	1.357	1.745	1.480	1.233	1.048
	MSE	1.884	3.690	2.703	3.208	6.393	3.924	2.645	1.898
	MAPE	10.746	15.170	13.690	14.580	17.831	15.596	13.376	10.866
20step ahead	MAE	1.327	1.893	1.667	1.764	2.404	2.031	1.588	1.282
	MSE	2.853	6.390	4.622	5.344	11.985	7.194	4.202	2.799
	MAPE	14.133	20.890	18.620	19.490	25.345	21.992	17.858	13.701

RW : 무작위 보행 모형  
 DES : 이중 지수 평활 모형(1, 2, 3은 모형적합에 쓰인 자료의 수에 따라 정해짐)  
 RT : Robust Trend 모형(1, 2, 3, 4는 모형적합에 쓰인 자료의 수에 따라 정해짐)

(b) 예측치 결합모형의 분석 결과

		Bcom	Bsel	Pcom	Psel
1step ahead	MAE	0.081	0.081	0.081	0.081
	MSE	0.029	0.029	0.029	0.029
	MAPE	0.754	0.754	0.750	0.750
5step ahead	MAE	0.386	0.386	0.384	0.384
	MSE	0.338	0.338	0.337	0.337
	MAPE	3.742	3.742	3.731	3.731
10step ahead	MAE	0.748	0.748	0.749	0.749
	MSE	1.042	1.042	1.043	1.043
	MAPE	7.514	7.514	7.521	7.521
15step ahead	MAE	1.034	1.034	1.036	1.036
	MSE	1.884	1.884	1.885	1.885
	MAPE	10.746	10.746	10.763	10.763
20step ahead	MAE	1.327	1.327	1.282	1.282
	MSE	2.853	2.853	2.799	2.799
	MAPE	14.133	14.133	13.701	13.701

Bcom : Bayesian Combination      Bsel : Bayesian Selection  
 Pcom : PNN Combination          Psel : PNN Selection

되지만 갱신된 가중치가 각 모형의 예측치에는 영향을 미치지 못한다. 따라서 예측모형은 알 수 없고 예측치만 알고 있다면 다른 방법이 없지만, 가능한 예측 모형들을 알고 있다면 그 모형들을 분석하여 각 모형을 포함할 수 있는 모형을 개발하는 노력이 선행되어야 할 것이다.

참고문헌

Bates, J. M., and Granger, C. W. J. (1969), The Combination of Forecasts, *Operations Research Quarterly*, 20(4), 451-468.  
 Clemen, R. T. (1989), Combining Forecasts: A review and annotated bibliography with Discussion, *International Journal of Forecasting*, 5, 559-609.  
 Collopy, F., and Armstrong, J. S. (1992), Rule-Based Forecasting: Development and Validation of an Expert Systems Approach to Combining Time Series Extrapolations, *Management Science*, 38(10), 1394-1414.  
 De Menezes, L. M., and Bunn, D. W. (1998), The Persistence of Specification Problems in the Distribution of Combined Forecast Errors, *International Journal of Forecasting*, 14, 415-426.  
 Diebold, F. X., and Pauly, P. (1990), The Use of Prior Information in Forecast Combination, *International Journal of Forecasting*, 6, 503-508.  
 Fildes, R. (1992), The evaluation of extrapolative forecasting methods, *International Journal of Forecasting*, 8, 81-98.  
 Foster, D. P., and Vohra, R. V. (1993), A Randomization Rule for Selecting Forecasts, *Operations Research*, 41(4), 704-709.  
 Gardner Jr., E. S. (1985), Exponential Smoothing: The State of the Art, *Journal of Forecasting*, 4, 1-28.  
 Grambsch, P., and Stahel, W. A. (1990), Forecasting demand for special telephone services A case study, *International Journal of Forecasting*, 6, 53-64.  
 Granger, C. W. J., and Ramanathan, R. (1984), Improved Methods of Combining Forecasts, *Journal of Forecasting*, 3, 197-204.  
 Greene, William, H. (1993), *Econometric Analysis*, 2nd Ed., 216-218, Macmillan Publishing Company, New York.  
 Harvey, D. I., Leybourne, S. J., and Newbold, P. (1998), Tests for Forecast Encompassing, *Journal of Business & Economic Statistics*, 16(2), 254-259.  
 Kang, H. J. (1986), Unstable Weight in The Combination of Forecasts, *Management Science*, 32(6), 683-695.  
 Min, C., and Zellner, A. (1993), Bayesian and non-Bayesian methods for combining models and forecasts with applications to forecasting international growth rates, *Journal of Econometrics*, 56, 89-118.  
 Newbold, P., and Granger, C. W. J. (1974), Experience with Forecasting Univariate Time Series and the Combination of Forecasts with Discussion, *Journal of the Royal Statistical Society A*, 137, Part 2, 131-165.  
 Specht, D. F. (1990), Probabilistic Neural Networks, *Neural Networks*, 3, 109-118.

부 록

1. 감쇠 지수 평활 모형(Gardner, Jr., 1985)

$$S_t = \alpha \times X_t + (1 - \alpha) \times (S_{t-1} + \phi \times T_{t-1})$$

$$T_t = \gamma \times (S_t - S_{t-1}) + (1 - \gamma) \times \phi \times T_{t-1}$$

$$\hat{X}_t(m) = S_t + \sum_{i=0}^m \phi^i \times T_t$$

이 때,  $X_t$  : 원 시계열,  $\hat{X}_t(m)$  :  $m$ 기 앞 예측치

2. Robust Trend Model(Grambsch and Stahel, 1990)

$$Z_t = X_t - X_{t-1} = \mu + \sigma \times \epsilon_t$$

이 때,  $\epsilon_t$  : i. i. d. stable Random variable

$T$ 기에서의  $k$ 기 앞 예측치:  $\hat{X}_{T+k} = X_T + k \times \hat{\mu}_T$

robust estimate of the trend :  $\hat{\mu}_T = M_T + \frac{m_i}{T} \times \sum_{i=1}^T \psi\left(\frac{Z_i - M_T}{m_T}\right)$

$M_T$  : the median of  $(Z_1, \dots, Z_T)$

$m_T$  : the median of  $(|Z_1 - M_T|, \dots, |Z_T - M_T|)$

$\Psi(x) = \text{sign}(x) \times \text{Max}\left(\left|\frac{2 \times x}{3}\right|, 1.0, 2 - \left|\frac{x}{3}\right|\right), 0]$

3. Empirical Results

표 1. 통신 자료에 대한 오차자승 평균 결과

	DES	Damp	RT	RW	Bcom	Bsel	Psel	Pcom
1-step	5562	5901	5959	6825	5690	6003	5555	5648
2-step	12542	14271	12508	15761	12131	12729	11755	11990
3-step	19264	25175	20370	27870	12157	12145	12128	13510
4-step	28164	39078	28669	42517	28358	28669	27664	27664
5-step	38346	54895	37053	59280	36922	37053	37390	37596
6-step	47804	72860	43785	78175	43587	43785	41805	41816
7-step	61782	93670	50482	99888	50411	50482	53762	53863
8-step	76556	117066	57100	124611	56918	57100	58898	58898
9-step	89871	142938	62656	151627	62478	62656	58964	61915
10-step	102637	172134	68265	182190	67696	68265	62927	65189
11-step	111346	203698	72733	215511	71827	72733	64559	65532
12-step	124092	239276	80743	253176	79490	80743	70437	70460
13-step	131441	278350	90081	294031	88485	90081	68310	68310
14-step	135569	308141	121776	323615	92700	94912	80588	80588
15-step	141934	339670	102021	355870	97921	102021	83516	83896
16-step	147536	370686	109575	389135	107228	109575	77492	77492
17-step	141413	403302	120889	425096	117221	120889	78430	78430
18-step	132092	441011	137374	464625	131655	137374	80794	80794

DES : 이중 지수 평활 모형 Damp : 감쇠 지수 평활 모형  
 RT : Robust Trend model RW : 무작위 보행 모형  
 Bcom : Bayesian Combination Bsel : Bayesian Selection  
 Psel : PNN Selection Pcom : PNN Combination

표 2. 통신 자료에 대한 절대오차 평균 결과

	DES	Damp	RT	RW	Bcom	Bsel	Psel	Pcom
1-step	26.70	28.69	25.44	32.44	25.51	26.93	24.86	25.82
2-step	55.45	56.16	48.52	63.70	47.19	52.27	44.94	45.43
3-step	70.68	84.01	73.11	97.03	47.81	47.74	47.31	52.64
4-step	95.75	119.35	97.28	131.63	94.28	97.28	91.80	91.80
5-step	121.70	156.03	123.11	167.59	120.61	123.11	117.60	116.57
6-step	150.67	193.88	145.01	205.56	142.95	145.01	137.08	137.70
7-step	179.53	234.33	166.64	244.09	165.43	166.64	157.98	159.43
8-step	210.95	275.75	188.15	284.82	187.19	188.15	178.30	178.30
9-step	242.98	319.27	206.66	327.06	206.46	206.66	200.20	197.51
10-step	270.67	365.17	222.02	370.44	221.65	222.09	212.71	213.81
11-step	290.02	412.01	233.73	414.55	233.58	233.73	219.11	220.68
12-step	316.94	459.86	250.47	459.77	250.55	250.47	232.26	232.42
13-step	332.10	507.23	269.12	505.93	269.29	269.12	231.23	231.23
14-step	338.57	542.16	315.55	536.14	277.42	277.83	256.32	256.32
15-step	349.10	573.28	290.60	566.41	287.57	290.60	264.76	263.48
16-step	368.07	604.01	304.42	595.46	302.06	304.42	256.86	256.86
17-step	358.31	634.81	320.94	625.04	318.81	320.94	258.46	258.46
18-step	343.32	666.67	339.39	655.17	336.90	339.39	254.17	254.17

DES : 이중 지수 평활 모형 Damp : 감쇠 지수 평활 모형  
 RT : Robust Trend model RW : 무작위 보행 모형  
 Bcom : Bayesian Combination Bsel : Bayesian Selection  
 Psel : PNN Selection Pcom : PNN Combination

표 3. 통신 자료에 대한 예측 오차의 정규분포 검정 결과

		DES	Damp	RT	RW
2step	chi-square	54.35***	66.25***	69.65***	83.25***
3step	chi-square	53.85***	52.97***	45.13***	66.92***
4step	chi-square	58.0***	54.63***	37.8***	48.7***
5step	chi-square	46.5***	48.2***	28.3***	42.1***
6step	chi-square	42.2***	40.4***	19.1	53.8***
7step	chi-square	26.71**	34.0***	21.2*	42.3***
8step	chi-square	19.65	46.94***	18.71	70.47***
9step	chi-square	18.9	56.7***	20.8**	65.4***
10step	chi-square	14.9	48.6***	27.1***	58.0***
11step	chi-square	6.3	54.6***	28.5***	60.5***
12step	chi-square	21.0*	71.0***	44.0***	70.0***
13step	skewness	0.78	0.16	1.4	0.17
	kurtosis	0.62	-5.91***	-0.24	-6.80***
	Wilk's W	0.97	0.82***	0.87***	0.78***
14step	skewness	0.34	0.13	1.28	0.13
	kurtosis	0.77	-6.08***	-0.43	-7.12***
	Wilk's W	0.98	0.81***	0.89***	0.77***
15step	skewness	0.56	0.04	1.05	0.07
	kurtosis	1.47	-6.96***	-1.18	-7.49***
	Wilk's W	0.96	0.78***	0.89***	0.76***
16step	skewness	1.25	0.015	0.75	0.07
	kurtosis	1.99**	-7.04***	-2.28**	7.17***
	Wilk's W	0.90**	0.77***	0.89***	0.77***
17step	skewness	0.41	0.094	0.58	0.056
	kurtosis	2.14**	-6.69***	-2.90***	-6.59***
	Wilk's W	0.91**	0.77***	0.87***	0.77***
18step	skewness	1.52	0.14	0.61	0.14
	kurtosis	1.96**	-5.20***	-1.99**	-5.46***
	Wilk's W	0.88**	0.81***	0.87***	0.79***

13기 앞 예측오차부터는 자료의 수가 부족하여 다른 검정 방법 이용하였음.

DES : 이중 지수 평활 모형 Damp : 감쇠 지수 평활 모형

RT : Robust Trend model RW : 무작위 보행 모형

\* : 10% 유의수준에서 정규분포를 따른다는 가설 기각

\*\* : 5% 유의수준에서 정규분포를 따른다는 가설 기각

\*\*\* : 1% 유의수준에서 정규분포를 따른다는 가설 기각

표 4. 91일물 이자율 자료에 대한 각 모형별 예측오차의 정규분포 검정

	RT	RW
1-step	1478.33***	1581.19***
5-step	967.01***	1006.57***
10-step	1449.13***	1467.13***
15-step	1686.96***	1703.61***
20-step	1978.42***	2144.64***

RT : Robust Trend model RW : 무작위 보행 모형

\* : 10% 유의수준에서 정규분포를 따른다는 가설 기각

\*\* : 5% 유의수준에서 정규분포를 따른다는 가설 기각

\*\*\* : 1% 유의수준에서 정규분포를 따른다는 가설 기각