

Comparison between nonlinear statistical time series forecasting and neural network forecasting

Inkyu Kim¹⁾, Cheolyoung Lee²⁾ and Sungduck Lee³⁾

Abstract

Nonlinear time series prediction is derived and compared between statistic of modeling and neural network method. In particular, mean squared errors of prediction are obtained in generalized random coefficient model and generalized autoregressive conditional heteroscedastic model, and compared with them by neural network forecasting.

1. 서 론

Box-Jenkins의 ARIMA 모형으로 대변되는 선형 시계열 모형들은 실제 시계열 자료들에 잘 적합 될 뿐만 아니라 확률모형에 대한 이론 전개에 있어서도 무리가 없어서 많은 연구가 이루어져 왔다. 그러나 강한 비대칭을 나타내는 시계열 자료들이나 또는 시간 가역성(time invertibility)을 갖지 않는 자료들에 대해서는 일반적으로 선형 시계열 모형으로 정형화하는 것이 적당하지 않다고 알려져 있다(Tong(1990)). 여러 분야에서 얻어지는 다양한 시계열 자료의 일반적인 도약 현상(jump phenomena)과 한계 순환(limit cycles)같은 것들은 선형 시계열 모형으로는 잘 설명되어 질 수가 없어서 이러한 선형 시계열 모형들의 제한들과 비선형 시계열 모형들의 이점들을 이용한 비선형 시계열 모형들에 관한 많은 연구는 필요가 절실하나 아직까지 그에 대한 자료분석과 예측에 관한 연구는 아주 미흡한 편이다.

한편 신경망(neural network) 방법은 신호의 전달을 수학적으로 모형화한 것이며 이는 여러 응용 분야에 폭 넓게 적용되고 있다. Sharda와 Patil(1990)은 여러 가지 시계열 자료들에 대하여 신경망 방법과 전통적인 통계 예측방법인 Box-Jenkins방법의 예측을 비교하였는데 이 들은 신경망을 이용한 예측결과가 Box-Jenkins 예측방법만큼 효율적임을 보였다. 그러나 시계열 자료에 주기성(periodicity)이 있는 경우에는 신경망을 이용한 예측방법이 Box-Jenkins 예측방법보다 별 효과가 없었다. Tang, Almeida와Fishwish(1991)는 신경망과 Box-Jenkins모형의 예측력을 비교하면서, 신경망 이론의 예측에 역전파(backpropagation) 프로그램을 이용하였다. 이 연구에 의하면 시계열 자료가 변동의 폭이 작거나 일관된 추세를 보이는 경우에 두 방법사이에는 별 차이가 없는 것으로 나타났다. 그러나 일반적으로 단기 예측에는 Box-Jenkins방법이 더 효과가

1) Assistant Professor, Dept. of Computer Information, Woosong Information College, Taejeon, Korea (300-715)

2) Consultant, KMA Consultants., Youngdeungpo-ku, Seoul, Korea (150-010)

3) Professor, Dept. of statistics, Chungbuk National University, Cheongju, Korea (361-763)

있었고, 장기 예측에서는 신경망 방법을 이용한 예측이 더 효과가 있는 것으로 나타났다. 시계열의 변동폭이 크면서 불규칙한 추세를 보이는 경우에는 Box-Jenkins방법보다 효율적인 것으로 나타났다. 또한 신경망 방법은 Box-Jenkins방법에 대해 표본의 크기가 작아도 효율적인 것으로 실증분석을 통해 보여 주었다.

본 논문에서는 이분산성을 갖는 비선형 시계열 자료를 가지고, 비선형 시계열 모형중 일반화 확률계수 자기회귀모형(Generalized random coefficient autoregressive model)과 일반화 조건부 이분산 자기회귀모형(Generalized autoregressive conditional heteroscedastic model) 그리고 자료의 형태에 상관없이 적용할 수 있는 신경망(neural networks) 이론을 이용하여 예측을 해서 어느 것이 최소 평균예측오차제곱의 기준에서 더 정확한 예측값을 제공하는지를 비교분석한다.

제 2장에서는 비선형 시계열 모형중에서 일반화 확률계수 자기회귀모형과 일반화 조건부 이분산 자기회귀모형에 대한 소개와 예측에 대하여 다루었고 제 3장에서는 신경망 이론을 이용한 시계열 자료분석을 연구하였다. 제 4장에서는 이들 모형과 분석방법들에 대하여 실제 비선형 모형에 따른다고 알려져 있는 자료를 가지고 예측 비교 분석하였다.

2. 비선형 시계열모형

2.1 GRCA모형

Nicholls와 Quinn(1982)은 처음으로 확률계수자기회귀모형을 소개하였는데 이 모형은 조건부 이분산형태를 보이는 자료에 적합한 모형이라 설명하였고 또한 추정량의 일치성(strong consistency)과 점근적인 정규성(asymtotic normality)을 조사하였다. Hwang과 Basawa(1998)는 일반화된 p차 확률계수 모형, GRCA(p)을 소개하였는데 그 모형은 다음과 같다.

$$X_t = \Phi'_t X(t-1) + \varepsilon_t . \tag{2.1}$$

여기서 $\Phi_t = (\phi_{t1}, \dots, \phi_{tp})'$ 는 $(p \times 1)$ 인 확률계수의 벡터, $X(t-1) = (X_{t-1}, X_{t-2}, \dots, X_{t-p})'$

는 시점 t에서의 $(p \times 1)$ 인 과거 관측치의 벡터이고, $\{\varepsilon_t\}$ 는 확률 오차들의 수열이다.

$(\Phi_t, \varepsilon_t)'$, $t=0, \pm 1, \pm 2, \dots$,는 독립이며 동일한 분포를 갖는 $((p+1) \times 1)$ 인 확률 벡터들의 수열이며 평균과 분산은 다음과 같다.

$$E \begin{pmatrix} \Phi_t \\ \varepsilon_t \end{pmatrix} = \begin{pmatrix} \phi \\ 0 \end{pmatrix} , \quad Var \begin{pmatrix} \Phi_t \\ \varepsilon_t \end{pmatrix} = \begin{pmatrix} \Sigma_\phi & \sigma_{\phi\varepsilon} \\ \sigma'_{\phi\varepsilon} & \sigma_\varepsilon^2 \end{pmatrix}$$

여기서 $\phi = (\phi_1, \dots, \phi_p)'$ 는 $(p \times 1)$ 모수 벡터이고, $\Sigma_\phi = Var(\Phi_t)$ 는 $(p \times p)$ 행렬이고,

$\sigma_{\phi\varepsilon} = Cov(\Phi_t, \varepsilon_t)$ 는 $(p \times 1)$ 벡터이고, $\sigma_\varepsilon^2 = Var(\varepsilon_t)$ 이다. 이 모형은 확률계수 Φ_t 와 오차항

ε_t 가 종속되어 있다는 점에서 RCA모형과 다르다.

ϕ 의 조건부 최소제곱 추정량 $\hat{\phi}$ 은 다음과 같다.

$$\hat{\phi}_{(p \times 1)} = \left[\sum_{t=1}^n X(t-1)X'(t-1) \right]^{-1} \left[\sum_{t=1}^n X_t X(t-1) \right]. \quad (2.2)$$

식 (2.2)로부터 우리는 확률 계수 자기회귀 과정, $\{X(t)\}$ 의 예측치를 다음과 같이 구할수 있다.

$$\widehat{X}(t) = E(X_t | X(t-1)) = \hat{\phi}' X(t-1). \quad (2.3)$$

2.2 GARCH모형

ARCH모형에 대한 연구는 Engle(1982)이 처음으로 자기회귀 조건부 이분산 모형(ARCH : autoregressive conditional heteroscedastic model)을 개발하였으며, 이 모형은 오차의 조건부 분산이 시차를 갖는 오차항들의 제곱에 의해 설명된다고 하였으며 영국의 인플레이션 자료를 이 모형으로 적합시켜 분석하였고, Bollerslev(1986)는 ARCH모형을 일반화시킨 GARCH모형을 제시하였다. ARCH(1)모형은 다음과 같다.

$$\begin{aligned} y_t &= \varepsilon_t h_t^{\frac{1}{2}}, \\ h_t &= \alpha_0 + \alpha_1 y_{t-1}^2. \end{aligned} \quad (2.4)$$

여기서 ε_t 의 평균은 0이고, 분산은 1이다. 또한 ε_t 들은 서로 독립적이다. 이 모형은 정확히 중선형모형은 아니지만 매우 근사한 형태를 갖는다.

ARCH모형에 정규성 가정을 추가하면, ARCH모형은 시간 t에 이용 가능한 정보집합인 ϕ_t 의 항들로 직접적으로 표현될 수 있다. 조건부 밀도함수를 이용하여 ARCH모형을 다시 쓰면 다음과 같다.

$$y_t | \phi_{t-1} \sim N(0, h_t), \quad (2.5)$$

$$h_t = \alpha_0 + \alpha_1 y_{t-1}^2. \quad (2.6)$$

여기서 $\phi_{t-1} = \{y_{t-1}, \alpha_0, \alpha_1\}$ 이다. 그리고 분산의 함수 h_t 를 더 일반적으로 표현할 수 있는데 다음과 같다.

$$\begin{aligned}
 h_t &= \alpha_0 + \alpha_1 y_{t-1}^2 + \dots + \alpha_p y_{t-p}^2 \\
 &= h(y_{t-1}, y_{t-2}, \dots, y_{t-p}, \alpha) .
 \end{aligned}
 \tag{2.7}$$

여기서 p 는 ARCH모형의 차수이고, α 는 알려지지 않은 모수들의 벡터이다.

ARCH모형을 보다 일반화시킨 것이 일반화된 조건부 이분산 자기회귀모형이다. Bollerslev(1986)는 조건부 이분산의 개념을 오차항들의 제곱뿐만 아니라 시차를 갖는 과거의 조건부 이분산에 의하여 영향을 받는 GARCH모형을 도출하였는데 그 모형은 다음과 같이 표현된다.

$$\begin{aligned}
 y_t &= x_t \beta + \varepsilon_t , \\
 \varepsilon_t | \phi_{t-1} &\sim N(0, h_t) , \\
 h_t &= \alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^q \beta_j h_{t-j} .
 \end{aligned}
 \tag{2.8}$$

단, $p > 0, q \geq 0, \alpha_0 > 0, \alpha_i \geq 0 (i = 1, \dots, p), \beta_j \geq 0 (j = 1, \dots, q)$ 이다.

식(2.8)에서 $q = 0$ 이면 ARCH(p)모형이고, $p = q = 0$ 이면 ε_t 는 단순히 백색잡음이 된다. 위에서 보는 바와 같이 ARCH(p)모형은 조건부 분산이 단지 과거 오차항의 선형함수로 표시되어 있지만, GARCH(p, q)모형에서는 과거조건부 분산도 반영되어 있는 것을 알 수 있다.

예를 들어 정상 AR(1)모형에서 오차가 GARCH(1,1)에 따르는 다음과 같은 모형에서 예측값을 구하면,

$$y_t = \phi_1 y_{t-1} + \varepsilon_t, \quad |\phi_1| < 1 .$$

여기서 $\varepsilon_t \sim \text{GARCH}(1,1)$ 모형을 따른다. 최적의 예측치는 다음과 같다.

$$E(y_{t+s} | \Psi_t) = \phi_1 E(y_{t+s-1} | \Psi_{t-1}) .$$

여기서 Ψ_t 는 (y_1, y_2, \dots, y_t) 로 만들어지는 σ -field이다. 첫 번째 예측치는 다음과 같다.

$$E(y_{t+1} | \Psi_t) = \hat{\phi}_1 y_t .$$

GARCH모형의 예측과 AR모형의 예측의 차이는 단지 다른 추정량 $\hat{\phi}$ 을 사용하는 것이다.

3. 신경망 모형

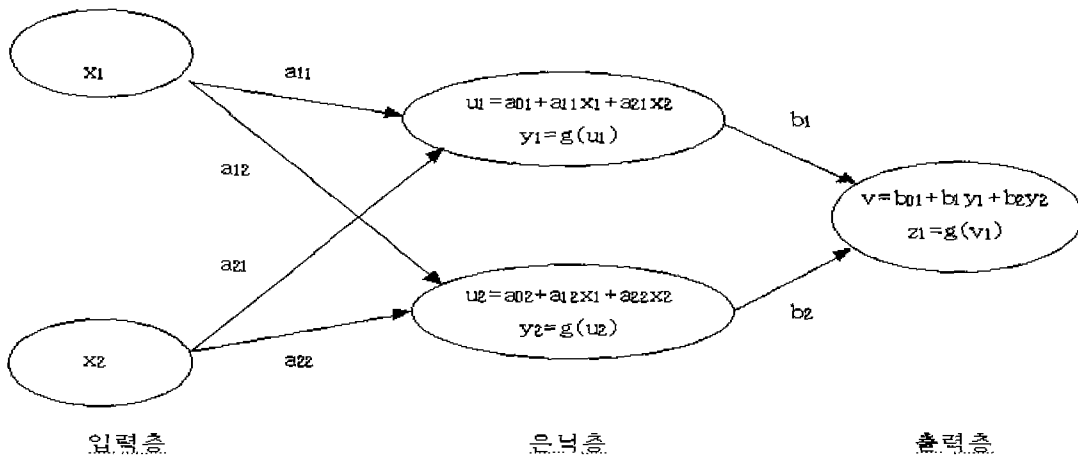
신경망은 입력층, 은닉층, 출력층으로 구성되며, 각 층의 노드의 개수로 신경망 구조를 표현한다. 본 논문의 자료에 적절한 신경망 구조를 찾기 위하여 환율자료를 사용하여 여러 가지 신경망구조의 정확도를 RMSE로 비교하여 본 결과는 다음의 표와 같다. 각 층의 노드의 수가 2이상

인 신경망 구조에서는 계산시간과 복잡성에 비교하여 볼 때 RMSE가 2×2×1 신경망의 구조와 큰차이 없었다.

<표 3.1 노드별 RMSE>

$\kappa=0.1 \quad \phi=0.5 \quad \theta=0.7 \quad \mu=0.9 \quad \text{epoch}=200$			
Input	Hidden	RMSE	
		학습오차	예측오차
1	1	0.021008	0.016444
1	2	0.019610	0.016030
1	3	0.019253	0.016074
1	4	0.019354	0.016269
2	1	0.018602	0.015159
<u>2</u>	<u>2</u>	<u>0.016294</u>	<u>0.013733</u>
2	3	0.016552	0.014886
2	4	0.016409	0.014767

따라서 본 논문에서의 신경망 구조는 2×2×1 신경망의 구조가 가장 적절한 구조임을 확인할 수 있었고, 그림으로 표현하면 다음과 같은 구조를 가진다



<그림 3.1> 2×2×1 신경망의 구조

본 논문에서는 다음과 같은 변환 함수를 사용하였다.

$$g(u) = \frac{1}{1 + e^{-u}} \quad , \quad 0 < g(u) < 1 \tag{3.1}$$

한편 출력값(z_n)과 목표값(t_n)간의 학습오차를 최소화시킬 수 있는 오차판단기준은 다음과 같은 식을 이용한다.

$$MSE = \frac{\sum_{n=1}^N (z_n - t_n)^2}{N}$$

여기서 N 은 예제 총수이다.

신경망 모형을 이용하여 실제 시계열자료를 적용시키는 경우에는 관측치들의 값들을 0과 1사이의 값을 가지도록 선형변환하여 사용하게 된다.

$$\text{변환값} = \frac{\text{관측치} - \text{관측치의최소값}}{\text{관측치의최대값} - \text{관측치의최소값}}$$

학습과정은 신경망 모형의 초기가중치에 임의의 값을 준 후 선형변환된 입력값을 대입하여 선형식과 시그모이드(sigmoid)변환함수를 거쳐서 계산된 출력 값을 실제값과 비교하여 MSE를 계산한다. 계산된 MSE를 이용하여 가중치를 바꾸는 과정인 역전파 알고리즘을 사용한 후, 다시 선형 변환된 입력값을 신경망모형에 적용하여 새로운 MSE를 계산한다.

본 논문에서는 Smith(1993)의 제안을 근거로 하여 최대학습수를 10,000으로 지정하고, 허용오차의 범위를 0.001이하로 정하고, 두 조건중 하나만 만족하면 학습과정을 멈추고, 그때의 신경망 모형을 최적의 모형으로 사용하였다. 학습 후에 $t+1$ 시점 이후의 k 개 예측값 $\hat{Z}_{t+1}, \hat{Z}_{t+2}, \dots, \hat{Z}_{t+k}$ 은 다음과 같이 구하였다.

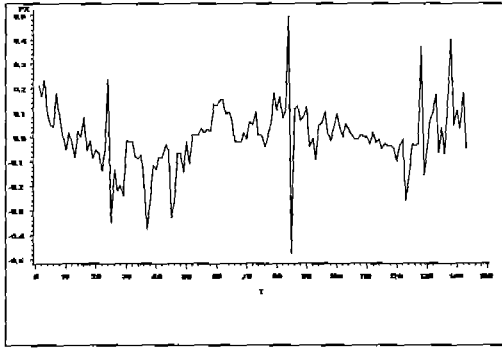
< 예측값 >	←	< 입력값 >
\hat{Z}_{t+1}		Z_t 와 Z_{t-1}
\hat{Z}_{t+2}		\hat{Z}_{t+1} 와 Z_t
⋮		⋮
\hat{Z}_{t+k}		\hat{Z}_{t+k-1} 와 \hat{Z}_{t+k-2}

4. 예측결과비교

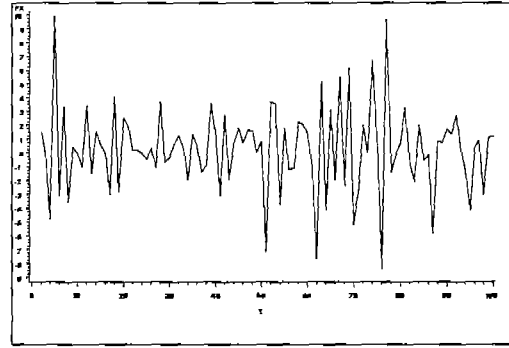
김명기와 문소양(1998)의 연구에 의하면 환율, 주가, 금리변동의 분산은 시간의 흐름에 따라 변화하고, 이러한 변동성을 분석하는데 있어서 비선형모형인 ARCH모형을 활용하였다.

환율변동율은 1985년 1월부터 1996년 12월까지의 우리나라의 매월말 원/달러 환율변동율을 이용하여 1997년 1월부터 12월까지의 매월말 원/달러 환율변동율을 예측하였고, 미국의 담배생산량의 변동율은 1871년부터 1972년까지 자료를 이용하여 1973년부터 12년간의 연간 담배생산량의 변동율을 예측하였고, 종합주가지수는 1990년 1월 3일부터 1996년 10월 31일까지의 종가기준 일일 한국종합주가지수의 변동율을 이용하여 이후 12일간의 변동율을 예측하여, 실제 변동율과 비교하

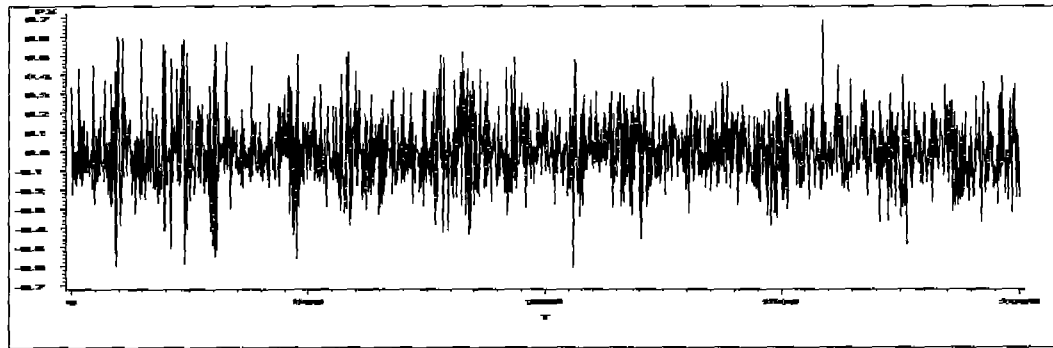
였다.



<그림 4.1 > 환율의 변환



<그림4.2> 담배생산량의 변환



<그림 4.3 > 종합주가지수변환

이와 같은 환율변동율과 담배생산량 변동율 그리고 종합주가지수변동율의 원자료에 대한 그래프에서는 평균과 분산이 변하는 비정상시계열임을 확인할 수 있다. 각 자료에 대한 변동율 그래프를 통하여 평균이 일정하나, 분산이 시간에 따라 바뀌는 조건부 이분산성을 시각적으로 확인할 수 있다. 그래프를 통한 이분산성을 포트맨토우 검정(Portmanteau test)와 Engle의 라그랑지 승수 검정(Lagrange Multiplier test)을 사용하여 이분산성을 검정하여본 결과 환율, 담배생산량, 종합주가지수의 변동율이 이분산성이 존재한다는 것을 알 수 있었다.

각각의 변동율자료에 대하여 비선형모형인 GRCA모형과 GARCH모형을 적용시켜서 모수절약의 원칙과 오차등을 고려하여 차수를 선택한 결과 GRCA(1)모형과 GARCH(1,1)모형이 적절함을 알 수 있었다. 두 개의 선택된 비선형모형과 신경망 구조로 각각의 자료에 대하여 12시점 이후까지 예측하여 실제값과 비교하여 각 시점까지의 MSE를 구하였다.

<표 4-1> 환율 변동율의 MSE비교 *는 각 모형중 가장 작은 MSE를 나타냄

	GARCH	GRCA	Nerual-Network
1	6.2108E-04	6.2731E-04	* 1.3009E-05
2	1.6239E-03	1.6408E-03	* 8.8925E-05
3	2.1178E-03	2.1435E-03	* 8.1471E-05
4	4.3410E-03	4.3856E-03	* 1.1389E-04
5	5.3848E-03	5.4419E-03	* 1.2647E-04
6	6.0699E-03	6.1367E-03	* 1.3081E-04
7	6.4135E-03	6.4872E-03	* 2.0464E-03
8	6.8082E-03	6.8885E-03	* 2.2927E-03
9	7.4596E-03	7.5474E-03	* 3.0129E-03
10	8.4405E-03	8.5369E-03	* 3.2690E-03
11	1.1509E-02	1.1622E-02	* 7.7645E-03
12	* 3.1198E-02	3.1357E-02	5.9617E-02

<표 4-2> 담배생산량 변동율의 MSE비교 *는 각 모형중 가장 작은 MSE를 나타냄

	GARCH	GRCA	Nerual-Network
1	1.3745E-02	1.2681E-02	* 1.3009E-05
2	1.2621E-02	1.1209E-02	* 8.8925E-05
3	1.2860E-02	1.1112E-02	* 8.1471E-05
4	9.6456E-03	8.3651E-03	* 1.1389E-04
5	9.2643E-03	8.7365E-03	* 1.2647E-04
6	8.4531E-03	8.3422E-03	* 1.3081E-04
7	7.4602E-03	7.2403E-03	* 2.0464E-03
8	6.5498E-03	6.4270E-03	* 2.2927E-03
9	1.1242E-02	1.0476E-02	* 3.0129E-03
10	1.1101E-02	1.0155E-02	* 3.2690E-03
11	1.0182E-02	9.4198E-03	* 7.7645E-03
12	9.3338E-03	* 8.6478E-03	5.9617E-02

<표 4-3> 종합주가지수 변동율의 MSE비교 *는 각 모형중 가장 작은 MSE를 나타냄

	GARCH	GRCA	Nerual-Network
1	2.2623E-04	2.2650E-04	* 1.3009E-05
2	1.8371E-04	1.8398E-04	* 8.8925E-05
3	1.9805E-04	1.9812E-04	* 8.1471E-05
4	4.6036E-04	4.6021E-04	* 1.1389E-04
5	4.1480E-04	4.1461E-04	* 1.2647E-04
6	4.3681E-04	4.3656E-04	* 1.3081E-04
7	4.8075E-04	* 4.8045E-04	2.0464E-03
8	6.1550E-04	* 6.1513E-04	2.2927E-03
9	7.3736E-04	* 7.3692E-04	3.0129E-03
10	6.6929E-04	* 6.6888E-04	3.2690E-03
11	6.2946E-04	* 6.2906E-04	7.7645E-03
12	6.1558E-04	* 6.1517E-04	5.9617E-02

각각의 자료에 대한 MSE를 비교하여 보면 비선형모형인 GRCA모형과 GARCH모형의 예측값은 표에서 보는 것과 같이 거의 비슷하여 두 모형의 비교는 의미가 없는 것으로 보여진다. 두 비선형모형과 신경망모형은 조금은 다른 예측치를 나타냄을 알 수 있다. 환율 변동율과 담배생산량 변동율의 예측값의 MSE비교에서는 시점 1에서 시점 11까지의 예측에서는 신경망모형이 비선형모형보다 좋은 예측값을 보였고 시점 12부터는 비선형모형이 신경망모형보다 좋은 예측값을 나타냄을 알 수 있다. 종합주가지수 변동율에서도 초기 시점 6까지는 비선형모형이 좋은 예측값을 보였으나 시점 7부터는 비선형모형인 GRCA모형이 더 좋은 예측값을 보였다. 이러한 예측결과로 보아 비선형 시계열자료인 경우는 Tang 등(1991)의 과거 논문들의 선형자료에 대한 선형시계열모형과 신경망모형과의 비교와 상이한 결과를 보였다. 즉, 선형자료에 대한 선형시계열모형과 신경망모형의 예측력 비교에서 단기 예측의 경우 시계열모형의 예측력이 신경망모형보다 좋았으며 장기 예측의 경우에는 신경망모형의 예측력이 더욱더 좋았다. 그러나 본 논문의 비선형 시계열자료에 대한 비선형모형과 신경망모형의 비교에서는 단기 예측의 경우에는 신경망모형이 더욱더 좋은 예측력을 보였으며 장기 예측의 경우에는 비선형 모형이 더 좋은 예측력을 보였다.

5. 결론

두 개의 비선형 시계열모형과 신경망모형의 예측방법을 논하였으며, 비선형 시계열 자료인 조건부 이분산자료에 대하여 두 개의 비선형모형과 신경망모형을 적용하여 보았다. 조건부 이분산모형에 따르는 자료로 확인된 환율변동율, 담배생산량 변동율, 종합주가지수 변동율에 대한 사례 분석결과를 보면 비선형모형인 GARCH모형과 GRCA모형의 예측은 서로 비슷한 예측력을 가짐으로 두 모형의 예측력을 비교하는 것은 무의미한 것으로 보여졌으며, 비선형 모형과 신경망 방법의 예측력 비교를 하여본 결과 신경망 방법은 단기예측에 비선형 모형보다 좋은 예측력을 가짐을 알 수 있었고, 비선형 모형은 장기예측에 신경망 방법보다는 좋은 예측력을 가짐을 볼 수 있었다. 이러한 결과로부터 신경망 방법이 관찰가능한 자료의 끝부분에 영향을 받아 초기 예측에 더 적절한 예측값을 제공하는 데 비하여 비선형모형은 관찰된 자료의 전체적인 추세와 경향을 그대로 반영하므로 초기예측값 보다는 중·장기의 예측에 있어서 보다 좋은 예측값을 제공함을 알 수 있다.

참 고 문 헌

- [1] 김명기, 문소양(1998) 환율, 금리, 주가변동의 상호관련성 분석. 「경제분석」 제 4권 제2호, 한국은행 금융경제연구소.
- [2] Bollerslev, T. (1986) Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics* 31, 307-327.
- [3] Engle, R. F. (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of U. K. inflation. *Econometrica* 50, 987-1008.
- [4] Hwang, S. Y. and Basawa, I. V. (1998). Parameter estimation for generalized random coefficient autoregressive processes. *Journal of Statistical Planning and Inference* 68, 323-337.
- [5] Nicholls, D. F. and Quinn, B. G. (1982) *Random Coefficient Autoregressive*

Models: An Introduction. Lecture Notes in Statistics, Vol. 11, Springer-Verlag, New York.

- [6] Sharda, R. and Patil, R. B. (1990). Neural Networks as Forecasting Experts : An Empirical Test, IJCNN Washington, 2, 491-494.
- [7] Smith, M.(1993) Neural Networks for Statistical Modeling. New York, Van Nostrand Reinhold.
- [8] Tang, Z., Almeida, C., Fishwick, P.A. (1991). Time Series Forecasting Using Neural Networks vs. Box-Jenkins Methodology, Simulation, 57(5), 303~310.
- [9] Tong, H. (1990) Non-linear Time Series, Oxford University Press, Oxford.