

Neural network AR model with ETS inputs

Minjae Kim^a, Byeongchan Seong^{1,a}

^aDepartment of Applied Statistics, Chung-Ang University

Abstract

This paper evaluates the performance of the neural network autoregressive model combined with an exponential smoothing model, called the NNARX+ETS model. The combined model utilizes the components of ETS as exogenous variables for NNARX, to forecast time series data using artificial neural networks. The main idea is to enhance the performance of NNAR using only lags of the original time series data, by combining traditional time series analysis methods with the neural networks through NNARX. We employ two real data for performance evaluation and compare the NNARX+ETS with NNAR and traditional time series analysis methods such as ETS and ARIMA (autoregressive integrated moving average) models.

Keywords: ETS, ARIMA, NNAR, NNARX, combined forecast

1. 서론

최근 인공 신경망을 활용하여 시계열 자료를 예측하는 많은 방법들이 등장하고 있다. 예를 들면, 다층 퍼셉트론(multi-layer perceptron)과 같은 순방향 네트워크부터 순환 네트워크 계열인 순환신경망(recurrent neural network; RNN), 장단기 메모리(long short-term memory; LSTM), 게이트 순환 유닛(gated recurrent unit; GRU)과 시퀀스 데이터에 유용한 트랜스포머(transformer) 등 여러 모형들을 시계열 자료에 접목하려는 노력이 이어지고 있다. 이 중 순방향 네트워크로 시계열 자료를 예측하는 모형들은 계산이 빠르지만 시간에 따라 변화하는 추세와 계절성 등 시계열 성질을 충분히 학습하지 못하여 상대적으로 낮은 성능을 보여주며, 주로 과거의 정보를 기억하여 시퀀스 단위로 처리가 가능한 순환 네트워크 또는 트랜스포머 모형들은 시계열 성질을 잘 학습하지만 계산이 오래 걸리고 그 예측력이 자료에 로버스트하지 않다는 단점을 가진다 (Joseph, 2022).

신경망 모형들은 상대적으로 높은 예측력을 가지고 있으나 블랙박스 형태로서 시계열의 주요 특성들이 출력값에 어떤 영향을 미치는지 파악하기 어렵다. 따라서, 지수평활법(exponential smoothing model; ETS) (Brown, 1959; Holt, 1957)과 1970년대 대중화된 ARIMA 모형(autoregressive integrated moving average model) (Box 등, 1976)과 같은 전통적인 시계열 분석 방법과 접목에 대한 필요성이 부각되고 있다 (Crone 등, 2011). 그러나, ETS 모형은 외생변수 또는 입력변수를 따로 받아 들이기 어려우며, ARIMA 모형은 정상성(stationarity)을 확보하기 위하여 차분과 같은 변환에 의존해야 하는 단점이 존재한다.

최근 여러 연구들에서 신경망 모형의 블랙박스 특성을 보완하기 위하여 전통적 시계열 모형과 결합하는 형태가 나타나고 있다. 예를 들면, Babu와 Reddy (2014)는 ARIMA와 인공 신경망을 결합하였으며 Panigrahi와 Behera (2017)는 ETS 모형의 예측값과 그 잔차를 인공 신경망으로 합성하여 최종 예측값을 생성하였다.

This research was supported by the Chung-Ang University Research Scholarship Grants in 2022.

¹Corresponding author: Department of Applied Statistics, Chung-Ang University, 84 Heukseok-ro, Dongjak-Gu, Seoul 06974, Korea. E-mail: bseong@cau.ac.kr

Table 1: Additive error model

Trend	Seasonality		
	N	A	M
N	ETS(A,N,N)	ETS(A,N,A)	ETS(A,N,M)
A	ETS(A,A,N)	ETS(A,A,A)	ETS(A,A,M)
Ad	ETS(A,Ad,N)	ETS(A,Ad,A)	ETS(A,Ad,M)

Table 2: Multiplicative error model

Trend	Seasonality		
	N	A	M
N	ETS(M,N,N)	ETS(M,N,A)	ETS(M,N,M)
A	ETS(M,A,N)	ETS(M,A,A)	ETS(M,A,M)
Ad	ETS(M,Ad,N)	ETS(M,Ad,A)	ETS(M,Ad,M)

Smyl (2020)은 대표적인 시계열 예측 대회인 M4 Competition에서 ETS 모형과 RNN을 결합하여 높은 성능을 이끌어 내었다.

본 논문에서는 NNARX 모형(neural network autoregressive with exogenous model)과 ETS 모형의 결합을 또다른 방식으로 고려한다. NNARX 모형은 외생변수가 존재하는 시계열 자료를 예측하는 인공 신경망 모형으로서, ARIMA 모형과 유사한 구조를 가지고 있으며 과거 관측값들을 순방향 네트워크로 예측한다. 본 연구에서는 NNARX 모형의 외생변수로서 ETS 모형의 구성 성분을 활용한다. 원시계열의 과거값들을 입력값으로 사용하는 동시에 ETS 구성 성분을 직접적으로 학습시킴으로써 NNAR 모형의 블랙박스 특성을 개선하고 ETS 모형의 예측력을 제고할 수 있다.

본 논문의 2, 3, 4절에서는 신경망 모형을 포함하여 본 연구에서 고려된 시계열 예측 모형들을 간단하게 설명하며 5절에서는 실제 데이터를 통해 예측력을 평가하고 6절에서 결론을 맺는다.

2. 지수평활 모형

지수평활법은 1950년대 제안된 시계열 예측 기법으로 예측값은 과거 관측값들의 가중 평균이며 그 가중치는 과거 관측값일수록 지수적으로 감소하는 형태를 가진다. 특히, 수준(level), 추세(trend), 계절(seasonal), 오차(remainder)의 시계열 구성 성분의 다양한 조합을 이용하여 시계열 자료의 예측 작업을 빠르게 수행할 수 있다. 가장 기본적인 모형인 ETS(A,A,A) 모형의 형태는 다음과 같다 (Gardner, 1985).

$$y_t = \ell_{t-1} + b_{t-1} + s_{t-m} + \varepsilon_t,$$

$$\ell_t = \ell_{t-1} + b_{t-1} + \alpha\varepsilon_t,$$

$$b_t = b_{t-1} + \beta\varepsilon_t,$$

$$s_t = s_{t-m} + \gamma\varepsilon_t,$$

여기서 α, β, γ 는 평활상수(smoothing constants)이며 ε_t 는 iid $N(0, \sigma^2)$ 를 따르며 ℓ_t, b_t, s_t 는 시계열 구성 성분인 수준, 추세, 계절을 각각 의미하며 m 은 계절 주기를 의미한다. ETS(A,A,A)는 시계열 자료에 수준, 추세와 계절성이 존재하고 각 성분들이 선형결합의 형태이며 특히 계절 변동이 시계열 전반에 걸쳐 일정한 경우 주로 사용된다.

Table 1과 Table 2는 지수평활 모형의 주요 형태들을 정리한 것으로서, Table 1은 오차항이 가법적인 경우, Table 2는 오차항이 승법인 경우를 나타낸다. 4가지 알파벳 N, A, Ad, M은 각각 해당 성분이 없는 경우

(none), 가법인 경우(additive), 축소 가법인 경우(damped additive), 승법인 경우(multiplicative)를 나타낸다. 여기서, 축소 가법은 추세에 대한 가법의 영향을 줄이기 위해서 $0 < \phi < 1$ 를 이용하여 b_t 대신 ϕb_t 를 사용하는 경우를 의미한다. 승법의 경우에는 추정의 어려움 때문에 축소 승법(damped multiplicative)의 형태를 잘 사용하지 않는다. Table 1과 Table 2와 같은 다양한 모형 중에서 모형을 선택하는 것은 정보량 기준인 AIC (Akaike information criterion), BIC (Bayesian information criterion) 또는 corrected AIC (AICC)를 주로 사용한다 (Hyndman 등, 2008).

3. NNAR model

3.1. AR model

AR 모형(autoregressive model)은 ARIMA의 가장 기본적인 모형 형태로서 시계열 변수의 과거값들의 선형결합으로 시계열 자료를 예측하는 모형이다. 차수 p 의 자기회귀모형은 $AR(p)$ 또는 $ARIMA(p,0,0)$ 로 표기하며, 계절성이 존재하는 경우 주로 승법 형태를 이용하여 $ARIMA(p,0,0)(P,0,0)_m$ 로 확장하며 모형식은 다음과 같이 표현된다.

$$\begin{aligned} \phi_p(B)\Phi_P(B^m)y_t &= c + \varepsilon_t, \\ \phi_p(B) &= 1 - \phi_1 B - \dots - \phi_p B^p, \\ \Phi_P(B^m) &= 1 - \Phi_1 B^m - \dots - \Phi_P B^{Pm}, \end{aligned}$$

여기서 B 는 $B^j y_t = y_{t-j}$ 를 만족하는 시차연산자, p 는 비계절 AR 차수, P 는 계절 AR 차수이며, $\phi_p(B)$ 는 비계절성을 위한 자기회귀다항식, $\Phi_P(B^m)$ 는 계절성을 위한 자기회귀다항식이다.

보통 AR 모형은 정상성 조건 하에서 식별 및 추정되며, 예를 들어 $AR(1)$, $AR(2)$ 의 정상성 조건은 아래의 식과 같다.

$$\begin{aligned} AR(1) : & -1 < \phi_1 < 1, \\ AR(2) : & -1 < \phi_2 < 1, \quad \phi_1 + \phi_2 < 1, \quad \phi_2 - \phi_1 < 1. \end{aligned}$$

다만, 정상성 조건은 p 의 값이 커질수록 매우 복잡한 형태를 가진다.

3.2. NNAR model

NNAR 모형(nerual network autoregressive model)은 시계열 자료의 과거 관측값들을 신경망의 입력값(inputs)으로 사용하는 모형으로 반응 변수와 입력 변수 또는 자기회귀 변수 사이의 복잡한 비선형 관계를 다룰 수 있다. 특히, 신경망 모형 특성상 비선형 함수로 출력값을 수정하여 극단적인 입력값의 효과를 줄일 수 있다. NNAR 모형의 구조는 기본적으로 입력층(input layer), 은닉층(hidden layer), 출력층(output layer)으로 구성되어 있으며 각 층들의 노드들이 가중치와 편향을 결합하여 다음 층으로 신호를 보내는 순방향 네트워크 구조이다. NNAR 모형은 $NNAR(p,P,k)_m$ 으로 표기하며 이는 $ARIMA(p,0,0)(P,0,0)_m$ 와 유사한 입력값들을 가지며 t -시점의 시계열 예측값 $\hat{y}_{t|t-1}$ 을 다음과 같이 계산한다.

$$\hat{y}_{t|t-1} = F(y_{t-1}, y_{t-2}, \dots, y_{t-p}, y_{t-m}, y_{t-2m}, \dots, y_{t-pm}),$$

여기서 함수 F 는 신경망 모형의 블랙박스 구조를 나타내는 함수로서 비선형 형태로 고려된다. 이 모형은 ARIMA와는 달리 정상성 조건없이 적합이 가능하며 기존 시계열 예측 모형들과는 다르게 가중치 초기값에 따라 예측값이 변동하므로 여러 번의 훈련을 통해서 구한 예측값들의 평균을 이용하여 예측한다.

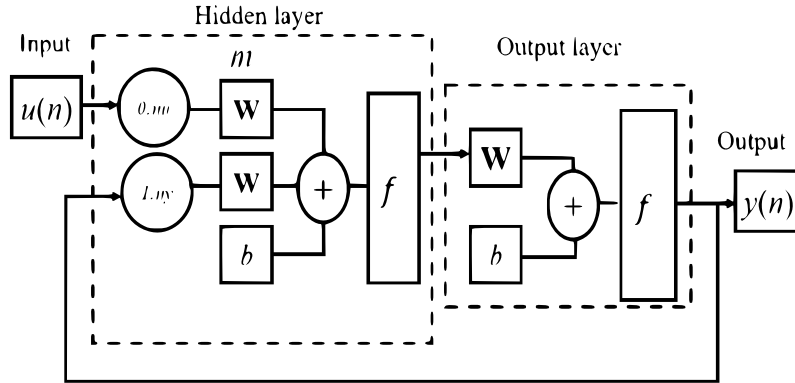


Figure 1: Architecture of NNARX model.

4. NNARX+ETS model

4.1. NNARX model

NNARX 모형(neural network autoregressive with exogenous model)은 시계열 자료의 예측을 위하여 사용하는 신경망 모형 중 하나로 신경망의 입력값으로 원시계열 자료의 시차값뿐만 아니라 외생변수(u_t)까지 고려하여 시계열 자료를 예측한다 (Khaled 등, 2020). t -시점까지의 자료를 토대로 미래 h -step 예측값은 다음과 같이 표현된다.

$$\hat{y}_{t+h|t} = F_X(u_t, \dots, u_{t-n_u}, y_t, \dots, y_{t-n_y}),$$

여기서, F_X 는 외생변수를 포함하는 신경망 모형의 블랙박스 구조를 나타내며, n_u 는 외생변수의 최대 과거시차, n_y 는 원시계열 변수의 최대 과거시차를 나타낸다.

NNARX 모형의 구조는 Series-Parallel architecture와 Parallel architecture로서 두 가지로 존재한다. 먼저 Series-Parallel architecture는 미래 시점을 예측할 때 원시계열의 참값만을 사용하여 더 정확한 값을 예측할 수 있게 되며 순방향 방식으로 훈련 할 수 있다는 장점이 있다. 반면에 Parallel architecture는 이전 시점의 예측값을 입력값으로 사용하는 순환 신경망 구조를 가지고 있으며, 주로 Series-Parallel를 통해 훈련을 마친 이후 미래 시점을 예측할 때에는 Parallel로 전환하여 진행된다. 하나의 은닉층을 가진 NNARX 모형은 Figure 1과 같은 구조를 가지며 W , b , f 는 각각 가중치, 편향, 활성화 함수(activation function)를 의미하며 출력값이 다시 원시계열의 입력값으로 들어가는 루프를 열고 닫음으로써 그 구조를 변경하게 된다.

4.2. NNARX+ETS model

순수한 NNAR 모형은 순방향 네트워크로서 추세와 계절성과 같은 시계열의 주요 구성성분들에 대한 학습이 미흡하여 예측 성능이 다른 신경망 모형보다 떨어진다. 또한 블랙박스 형태인 신경망 모형 특성상 학습 과정을 알 수 없기에 본 연구에서는 이러한 단점을 보완하기 위하여 NNAR 모형과 지수평활 모형을 결합해서 사용하는 방법을 제안하고자 한다. NNARX+ETS 모형은 NNARX의 외생변수를 지수평활법의 성분들로 사용하여 시계열 자료의 특징들을 직접적으로 학습시킴으로써 기존 NNAR 모형의 단점을 보완하며 예측값을 보정해준다. 또한 전통적인 시계열 예측 모형들은 외생변수를 추가하는 경우 예측시 외생변수의 미래값들을 따로 추정 또는 예측해야 하지만 신경망 모형은 이런 과정을 피할 수 있어서 편리함을 더해 준다. t -시점에서의

미래 h -step 예측값은 다음과 같이 표현된다.

$$\hat{y}_{t+h|t} = F_{XE}(\ell_t, \dots, \ell_{t-n_\ell}, b_t, \dots, b_{t-n_b}, s_t, \dots, s_{t-n_s}, y_t, \dots, y_{t-n_y}),$$

여기서, F_{XE} 는 지수평활 모형과 결합된 신경망 모형의 블랙박스 구조를 나타내며, n_ℓ, n_b, n_s, n_y 는 각각 ℓ_t, b_t, s_t, y_t 의 최대 과거시차를 나타낸다.

5. 예측 성능 평가

본 절에서는 NNARX+ETS 모형의 성능을 실제 데이터에 적용하여 평가하였다.

5.1. 평가 지표

예측 성능 평가를 위하여 다음과 같은 4가지 지표를 사용하였으며 평가 방법은 시계열 교차 검증을 사용하여 처음 훈련 데이터와 평가 데이터를 일정 비율로 나누고 rolling을 통하여 미래 예측 시점(forecast horizon) 1-step부터 12-step까지의 예측값을 구하여 비교하였으며, 신경망 모형들의 경우 100번의 출력값 평균을 예측값으로 사용하였다.

먼저, 시계열 자료의 예측 원점(forecast origin) $T + j$ 와 예측 시점 $h = 1, 2, \dots, H$ 에 대하여 다음과 같은 MAPE (mean absolute percentage error)를 사용한다 (Hyndman과 Koehler, 2006).

$$\text{MAPE} = \frac{1}{J} \sum_{j=1}^J \left| \frac{y_{T+j+h} - \hat{y}_{T+j+h|T+j}}{y_{T+j+h}} \right|,$$

여기서, $T + j$ 는 rolling에 따른 예측 원점을 나타내며 J 는 rolling에 의한 반복 예측수를 나타낸다. MAPE는 오차를 실제값으로 나누어 오차가 실제값에서 차지하는 상대적인 비율로 성능을 평가하여 직관적인 이해가 쉽다. 하지만 실제값에 0이 존재하면 정의되지 않고 오차의 절대값이 같더라도 실제값과 예측값의 대소 관계에 따라 과대 추정하는 단점이 있다. 두 번째는 MAPE의 단점을 보완하기 위한 지표인 sMAPE (symmetric MAPE)를 사용한다 (Svetunkov, 2023).

$$\text{sMAPE} = \frac{2}{J} \sum_{j=1}^J \frac{|y_{T+j+h} - \hat{y}_{T+j+h|T+j}|}{|y_{T+j+h}| + |\hat{y}_{T+j+h|T+j}|}$$

하지만 이 측도는 분모에 예측값이 들어가기 때문에 예측값에 의존한다는 단점이 존재한다. 세 번째로는 MASE (mean absolute scaled error)를 사용한다 (Hyndman 등, 2008).

$$\text{MASE} = \frac{1}{J} \sum_{j=1}^J \frac{|y_{T+j+h} - \hat{y}_{T+j+h|T+j}|}{1/(T-m) \sum_{i=m+1}^T |y_i - y_{i-m}|},$$

이는 예측값의 절대 오차를 차분을 통한 훈련 샘플의 변동 평균값으로 나눈 것으로 데이터 척도와 무관하다. 여기서, 계절성이 없는 경우 $m = 1$ 로 간주한다. 마지막으로 다음과 같은 sMAE (scaled mean absolute error)를 사용한다.

$$\text{sMAE} = \frac{1}{J} \sum_{j=1}^J \frac{|y_{T+j+h} - \hat{y}_{T+j+h|T+j}|}{1/(T+j) \sum_{i=1}^{T+j} y_i},$$

이는 예측오차의 절대값 평균(mean absolute error; MAE)을 다시 훈련 데이터의 평균으로 나누어 MASE와 마찬가지로 척도가 다른 예측값들을 비교할 때 사용할 수 있다.

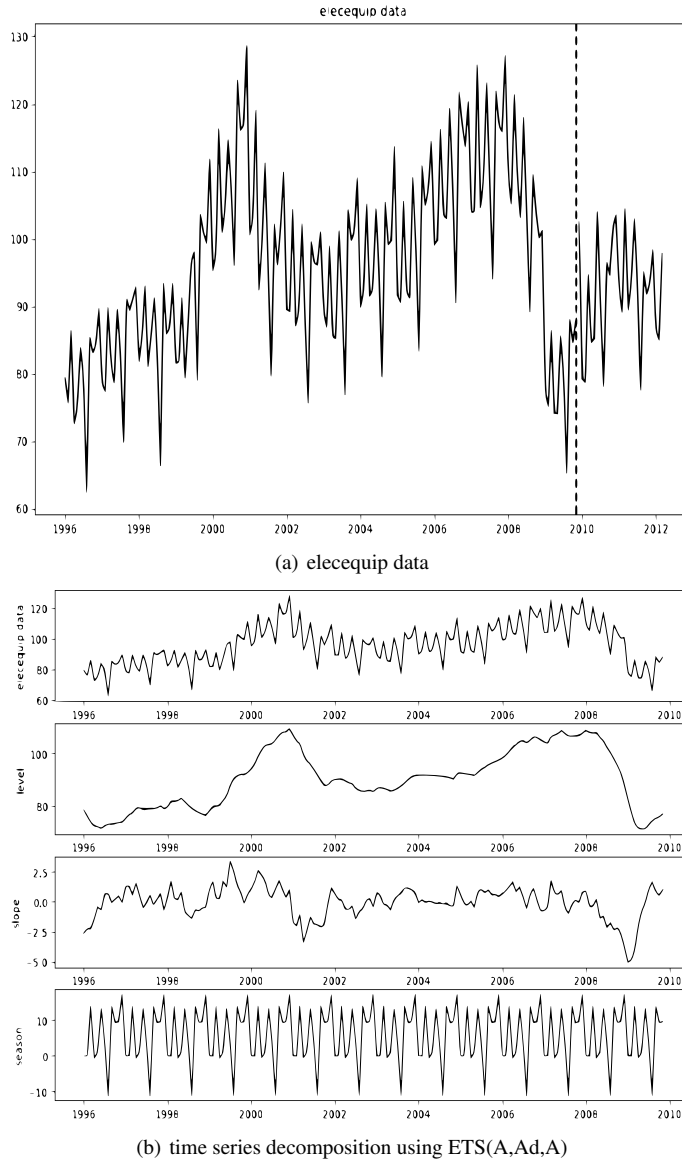


Figure 2: *Electrical equipment manufactured in the Euro area.*

5.2. 예측 성능 비교

본 연구에서는 NNARX+ETS 모형의 성능을 비교하기 위하여 ETS, ARIMA, NNAR 모형을 사용하였으며 시계열 자료는 R의 fpp2 패키지에서 제공하는 데이터 2가지이다. 4가지 지표로 모두 사용하여 비교하였으나 대체로 비슷한 결과를 보였기 때문에 각 시계열 자료에 대하여 1개의 지표만을 대표적으로 설명하고 나머지 지표들에 대한 예측결과는 Appendix에 Tables A.1–A.6으로 첨부하고 설명을 생략하였다.

Table 3: Accuracy of rolling forecasting using MASE

<i>h</i> -step	ETS	ARIMA	NNAR	NNARX+ETS
1	<u>3.3328</u>	3.7338	3.7278	2.9879
2	<u>3.4622</u>	3.5081	3.8464	3.1570
3	4.1654	4.0161	<u>3.5204</u>	3.4061
4	4.5350	4.6367	<u>3.9006</u>	3.4922
5	4.9893	5.1156	<u>4.3696</u>	3.6564
6	4.5288	4.9476	<u>4.4789</u>	4.4069
7	5.4061	5.3884	5.1414	<u>5.1639</u>
8	5.4899	<u>5.4433</u>	5.1336	5.4731
9	6.0926	<u>5.1621</u>	5.3909	5.0912
10	6.3093	5.2950	<u>5.0536</u>	5.0143
11	5.6337	5.0261	5.2624	<u>5.1434</u>
12	6.3387	5.8752	<u>5.2854</u>	5.2714
Total	5.0237	4.8457	<u>4.5926</u>	4.3553

Note: Bold numbers indicate that the corresponding model has the best forecasting performance in the given horizon while underlined numbers indicate the second best performance. The bottom row represents forecasting performances averaged over all the steps $h = 1 \sim 12$.

5.2.1. Electrical equipment manufactured in the Euro area

첫 번째 자료는 유럽지역 전자 장비 신규 주문 수에 대한 1996년 1월부터 2012년 3월까지의 월별 시계열 자료이며 Figure 2(a)는 시계열 그림을 나타낸다. 지수평활법은 ETS(A,Ad,A)로 적합하였으며 시계열 구성 성분들을 도식화한 Figure 2(b)로 보았을 때 시간이 지남에 따라 일정한 계절성을 보이므로 덧셈 기법의 활용이 적절해 보인다. 또한, 이 성분들은 NNARX+ETS 모형을 위한 외생변수로 사용되어 NNARX의 예측값을 보정해주는 역할을 해준다. 또한 ARIMA 모형은 ARIMA(1,0,0)(0,1,2)_m으로 적합하였고 NNAR과 NNARX+ETS 모형의 최대 과거시차값은 원시계열 변수 및 외생 변수 모두 편의상 12로 설정하였다.

Table 3은 MAPE를 이용하여 3개의 단변량 모형(ETS, ARIMA, NNAR) 및 결합 모형 NNARX+ETS의 성능을 비교한 결과를 나타낸다. 표에서 가장 우수한 성능을 보이는 경우는 진하게, 두 번째로 우수한 성능을 보이는 경우는 밑줄로 표시하였으며 표 맨아래의 Total 값은 $h = 1 \sim 12$ 에 대한 성능의 평균을 구한 것이다. 전반적으로 NNARX+ETS 모형이 가장 우수한 성능을 보이며 NNAR 모형에 비해 대략 5.17%의 예측 성능이 개선되었으며 ETS 모형과 비교한 경우 또한 13.30% 개선되었다.

5.2.2. Quarterly Australian Portland cement production

두 번째 시계열 자료는 호주 포틀랜드 지역의 시멘트 생산량에 대한 1956년 1사분기부터 2014년 1사분기까지의 분기별 자료이며 시계열 그래프는 Figure 3(a)에서 확인할 수 있다. 마찬가지로 ETS(A,Ad,A)로 적합하였으며 Figure 3(b)를 통해 계절성의 변동이 점차로 커지고 전반적으로 시계열의 수준이 높아짐을 볼 수 있다. 또한 ARIMA 모형은 ARIMA(2,0,3)(0,1,2)_m로 적합하였으며 NNAR과 NNARX+ETS 모형의 과거 시차값은 12로 설정하였다.

Table 4로 확인한 결과 전반적으로 NNARX+ETS 모형이 우수한 성능을 보이고 있다. 전체 평균으로 확인한 결과 NNAR 모형에 비해 대략 4.77%의 성능이 개선되었으며 ETS 모형과 비교한 경우는 3.60% 개선되었다. ARIMA 모형의 경우 예측시점 8-step 이전에는 성능이 좋지 않다가 그 이후부터는 ARIMA 모형의 성능이 대폭 개선되고 있다.

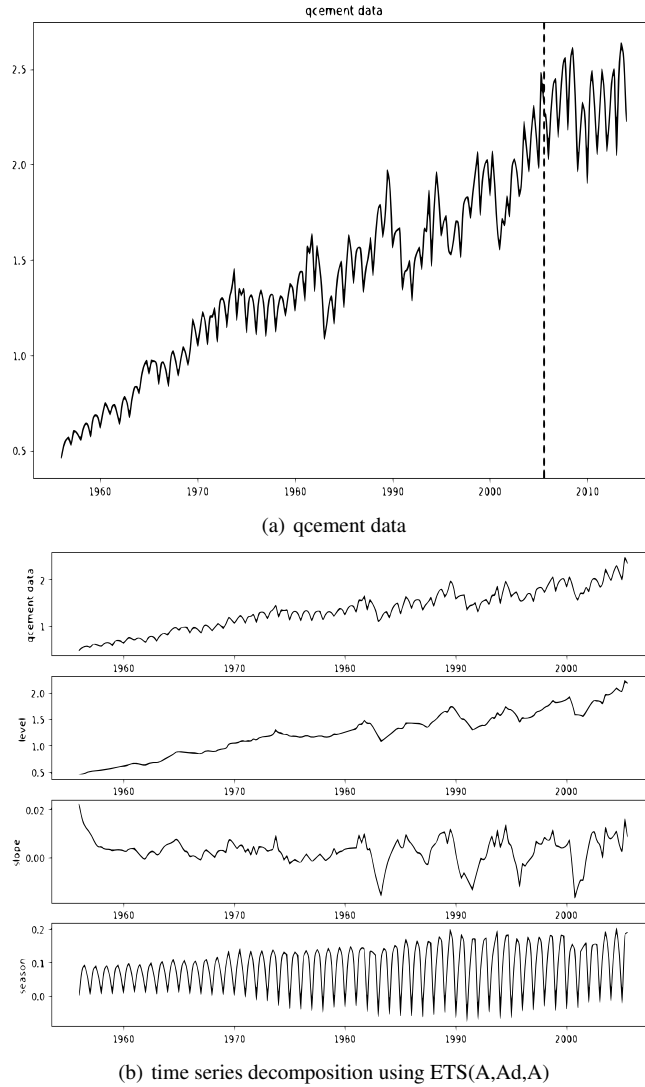


Figure 3: Quarterly Australian Portland cement production.

6. 결론

최근 다양한 신경망 모형을 통해 시계열 자료를 예측하는 방법들이 연구되고 있다. 본 연구에서는 ETS 모형의 시계열 구성 성분들을 외생변수로 사용하여 기존 NNAR 모형의 예측값을 보정하고 예측력 성능을 개선해주는 NNARX+ETS 모형을 연구하였다. 실제 시계열 자료들로 성능을 비교한 결과 ETS 성분들을 활용하여 시계열의 주요 패턴을 직접적으로 학습시킨 NNARX+ETS 모형의 성능이 원시계열의 과거 시차만을 사용한 NNAR 모형 뿐만 아니라 ETS와 ARIMA와 같은 전통적 시계열 예측 모형보다 우수한 성능을 보여주었다. 이는 순방향 네트워크에서의 학습력을 높이고 계산 시간을 줄일 수 있다는 장점을 준다.

본 연구의 제한된 예측 성능 평가에도 불구하고, NNARX+ETS 모형의 우수한 성능은 신경망 모형이 스스

Table 4: Accuracy of rolling forecasting using sMAPE

<i>h</i> -step	ETS	ARIMA	NNAR	NNARX+ETS
1	4.2853	4.0156	<u>3.9094</u>	3.4307
2	4.8650	4.7888	<u>4.4851</u>	4.2835
3	<u>4.8558</u>	5.2544	4.7867	4.9236
4	5.6141	5.7753	<u>5.7225</u>	5.8488
5	<u>6.5519</u>	6.7196	6.6302	6.3413
6	6.5815	6.8903	6.9232	<u>6.5843</u>
7	6.6472	6.9233	7.1994	<u>6.8700</u>
8	<u>7.0159</u>	6.9535	7.3891	7.0328
9	6.9371	6.8048	7.1212	<u>6.8730</u>
10	7.3765	6.8433	<u>6.6818</u>	6.5112
11	6.8081	5.9620	7.0336	<u>6.3668</u>
12	6.0899	5.0395	6.6512	<u>5.8888</u>
Total	6.1357	<u>5.9975</u>	6.2111	5.9150

Note: Bold or underlined numbers and the bottom row denote the same as in Table 3.

로 숨겨져 있는 피쳐(feature)를 찾아내도록 하는 것이 아니라 ETS의 구성 성분들을 입력해 줌으로써 신경망 모형이 추세와 계절성과 같은 피쳐를 더 잘 훈련할 수 있도록 도와 준다는 점에서 비롯된 것으로 추론된다.

향후 본 연구에서 제안된 NNARX+ETS 모형의 일반적 예측 성능을 검증하기 위하여 다양한 방향으로의 실험 및 시뮬레이션이 필요하다. 예를 들어, 대규모의 시계열 자료에 적용하여 그 성능을 평가할 필요가 있다. M3와 M4 competition의 대규모 자료(각각 3,003개, 100,000개 시계열 종류)가 그 예가 될 수 있을 것이다 (Makridakis와 Hibon, 2000; Petropoulos와 Svetunkov, 2020). 또한, ETS의 시계열 구성 성분을 가법 기법에 국한하지 않고 승법까지 확장할 수 있는 방법을 고안할 필요가 있다. 마지막으로, 본 연구에서 제안된 외생변 수로서의 조합 방식은, 트리 기반의 머신러닝 기법에도 확장할 수 있을 것으로 기대된다.

Appendix A: Accuracy of rolling forecasting using error measures

Table A.1: Accuracy of elecequipment data using MASE

<i>h</i> -step	ETS	ARIMA	NNAR	NNARX+ETS
1	<u>0.3788</u>	0.4202	0.4043	0.3286
2	<u>0.3781</u>	0.3853	0.4124	0.3440
3	0.4603	0.4399	<u>0.3833</u>	0.3755
4	0.4996	0.5220	<u>0.4238</u>	0.3867
5	0.5497	0.5637	<u>0.4728</u>	0.4029
6	0.5089	0.5459	<u>0.4871</u>	0.4839
7	0.6037	0.6025	0.5586	0.5679
8	0.6159	0.6063	0.5483	0.5986
9	0.6754	<u>0.5595</u>	0.5799	0.5538
10	0.7030	0.5965	<u>0.5545</u>	0.5521
11	0.6225	0.5641	0.5731	<u>0.5653</u>
12	0.5592	0.6594	0.5775	<u>0.5780</u>
Total	0.5592	0.5388	<u>0.4980</u>	0.4781

Note: Bold or underlined numbers and the bottom row denote the same as in Table 3.

Table A.2: Accuracy of elecequipment data using sMAPE

<i>h</i> -step	ETS	ARIMA	NNAR	NNARX+ETS
1	<u>3.3416</u>	3.7480	3.7209	2.9735
2	<u>3.4504</u>	3.5258	3.8155	3.1158
3	4.1662	4.0241	<u>3.4890</u>	3.3435
4	4.5177	4.6437	<u>3.8590</u>	3.4085
5	4.9761	5.1428	<u>4.2987</u>	3.5477
6	4.4571	4.9440	<u>4.3827</u>	4.2798
7	5.3514	5.3738	<u>5.0419</u>	5.0123
8	5.4128	5.3890	4.9875	5.2994
9	6.0088	<u>5.1347</u>	5.2551	4.9216
10	6.2224	5.2004	<u>4.9642</u>	4.8434
11	5.5808	4.9507	5.1288	<u>4.9753</u>
12	6.3163	5.8056	<u>5.1746</u>	5.1326
Total	4.9835	4.8236	<u>4.5098</u>	4.2378

Note: Bold or underlined numbers and the bottom row denote the same as in Table 3.

Table A.3: Accuracy of elecequipment data using sMAE

<i>h</i> -step	ETS	ARIMA	NNAR	NNARX+ETS
1	<u>0.0324</u>	0.0366	0.0352	0.0286
2	<u>0.0329</u>	0.0335	0.0359	0.0299
3	0.0400	0.0383	<u>0.0333</u>	0.0327
4	0.0435	0.0454	<u>0.0369</u>	0.0336
5	0.0478	0.0490	<u>0.0411</u>	0.0351
6	0.0443	0.0475	<u>0.0424</u>	0.0421
7	0.0525	0.0524	0.0486	0.0494
8	0.0536	0.0527	0.0477	<u>0.0521</u>
9	0.0587	<u>0.0487</u>	0.0504	0.0482
10	0.0612	0.0519	0.0482	0.0480
11	0.0541	0.0491	0.0499	<u>0.0492</u>
12	0.0621	0.0574	0.0502	<u>0.0503</u>
Total	0.0486	0.0469	<u>0.0433</u>	0.0416

Note: Bold or underlined numbers and the bottom row denote the same as in Table 3.

Table A.4: Accuracy of qcement data using MAPE

<i>h</i> -step	ETS	ARIMA	NNAR	NNARX+ETS
1	4.2978	4.0056	3.9039	3.4643
2	4.8845	4.7710	<u>4.5156</u>	4.3136
3	4.8898	5.2325	4.8239	4.9575
4	5.6205	<u>5.7144</u>	5.7573	5.8825
5	6.5917	6.6473	6.6993	6.3809
6	6.6166	6.8130	7.0108	<u>6.6182</u>
7	6.6681	<u>6.8143</u>	7.2684	6.9070
8	<u>7.0039</u>	6.8099	7.4837	7.0586
9	6.9726	6.6826	7.2102	6.9161
10	7.4037	<u>6.7265</u>	6.8452	6.6315
11	6.8000	5.8344	7.2390	6.4960
12	6.1146	4.9590	6.8794	<u>6.0592</u>
Total	6.1553	5.9175	6.3031	<u>5.9738</u>

Note: Bold or underlined numbers and the bottom row denote the same as in Table 3.

Table A.5: Accuracy of qcement data using MASE

<i>h</i> -step	ETS	ARIMA	NNAR	NNARX+ETS
1	1.0153	0.9506	<u>0.9267</u>	0.8181
2	1.1613	1.1383	<u>1.0698</u>	1.0217
3	<u>1.1611</u>	1.2426	1.1464	1.1818
4	1.3443	<u>1.3834</u>	1.3701	1.4012
5	<u>1.5715</u>	1.6070	1.5783	1.5167
6	1.5656	1.6230	1.6459	<u>1.5674</u>
7	1.5711	<u>1.6220</u>	1.7092	1.6372
8	1.6743	1.6469	1.7528	<u>1.6722</u>
9	1.6523	1.6123	1.6799	<u>1.6244</u>
10	1.7478	1.6078	<u>1.5714</u>	1.5331
11	1.6121	1.4083	1.6569	<u>1.5020</u>
12	1.4429	1.1930	1.5593	<u>1.3826</u>
Total	1.4600	<u>1.4196</u>	1.4722	1.4049

Note: Bold or underlined numbers and the bottom row denote the same as in Table 3.

Table A.6: Accuracy of qcement data using sMAE

<i>h</i> -step	ETS	ARIMA	NNAR	NNARX+ETS
1	0.0739	0.0692	<u>0.0674</u>	0.0600
2	0.0845	0.0828	<u>0.0779</u>	0.0743
3	<u>0.0845</u>	0.0904	0.0834	0.0860
4	0.0978	0.1007	<u>0.0997</u>	0.1020
5	<u>0.1144</u>	0.1169	0.1149	0.1104
6	0.1139	0.1181	0.1198	0.1141
7	0.1143	<u>0.1180</u>	0.1244	0.1191
8	0.1218	0.1198	0.1276	0.1217
9	0.1202	0.1173	0.1223	<u>0.1182</u>
10	0.1272	0.1170	0.1144	0.1116
11	0.1173	0.1025	0.1206	<u>0.1093</u>
12	0.1050	0.0868	0.1135	<u>0.1006</u>
Total	0.1056	<u>0.1033</u>	0.1071	0.1023

Note: Bold or underlined numbers and the bottom row denote the same as in Table 3.

References

- Brown RG (1959). *Statistical Forecasting for Inventory Control*, McGraw/Hill, New York.
- Box GEP, Jenkins GM, and Reinsel GC (1976). *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco.
- Babu CN and Reddy BE (2014). A moving-average filter based hybrid ARIMA-ANN model for forecasting time series data, *Applied Soft Computing*, **23**, 27–38.
- Crone SF, Hibon M, and Nikolopoulos K (2011). Advances in forecasting with neural networks? Empirical evidence from the M3 competition on time series prediction, *International Journal of Forecasting*, **27**, 635–660.
- Gardner ES (1985). Exponential smoothing: The state of the art, *Journal of Forecasting*, **4**, 1–28.
- Holt CC (1957). Forecasting seasonals and trends by exponentially weighted moving averages, *International Journal of Forecasting*, **20**, 5–10.
- Hyndman RJ and Koehler AB (2006). Another look at measures of forecast accuracy, *International Journal of Forecasting*, **22**, 679–688.
- Hyndman RJ, Koehler AB, Ord JK, and Snyder RD (2008). *Forecasting with Exponential Smoothing: The State*

Space Approach, Springer, Berlin.

Joseph M (2022). *Modern Time Series Forecasting with Python*, Packt Publishing, Birmingham.

Khaled S, Fakhry MM, and Ahmed S (2020). Classification of PCG signals using a nonlinear autoregressive network with exogenous inputs (NARX). In *Proceedings of 2020 International Conference on Innovative Trends in Communication and Computer Engineering (ITCE)*, Aswan, Egypt, 98–102.

Makridakis S and Hibon M (2000). The M3-Competition: Results, conclusions and implications, *International Journal of Forecasting*, **16**, 451–476.

Panigrahi S and Behera HS (2017). A hybrid ETS–ANN model for time series forecasting, *Engineering Applications of Artificial Intelligence*, **66**, 49–59.

Petropoulos F and Svetunkov I (2020). A simple combination of univariate models, *International Journal of Forecasting*, **36**, 110–115.

Smyl S (2020). A hybrid method of exponential smoothing and re-current neural networks for time series forecasting, *International Journal of Forecasting*, **36**, 75–85.

Svetunkov I (2023). *Forecasting and Analytics with the Augmented Dynamic Adaptive Model (ADAM)*, Chapman and Hall/CRC, Florida.

Received January 27, 2024; Revised February 1, 2024; Accepted February 1, 2024

지수평활법을 외생변수로 사용하는 자기회귀 신경망 모형

김민재^a, 성병찬^{1,a}

^a중앙대학교 응용통계학과

요 약

본 논문에서는 자기회귀 신경망 모형과 지수평활법을 결합(NNARX+ETS 모형)하고 그 성능을 평가한다. 제안된 결합 모형은 시계열 자료를 예측하기 위하여 NNARX 모형의 외생변수로서 ETS 모형의 구성 성분을 활용한다. 이 모형의 주요 아이디어는, 신경망 모형이 원시계열 자료의 과거 시차만을 고려하는 것을 한계를 넘어서서 전통적 시계열 예측 방법인 지수평활법에 의해서 추출된 정제된 시계열 구성 성분까지도 추가로 신경망 모형의 입력값으로 사용하는 것이다. 예측 성능 평가는 2가지 실제 시계열 자료를 사용하였으며 제안된 모형을 NNAR 모형 및 전통적 시계열 분석 방법인 ETS와 ARIMA 모형과 비교하였다.

주요용어: 지수평활법, ARIMA, 자기회귀 신경망, 자기회귀 외생변수 신경망, 결합 모형

이 논문은 2022년도 중앙대학교 연구장학기금 지원에 의한 것임.

¹교신저자: (06974) 서울시 동작구 흑석로 84, 중앙대학교 경영경제대학 응용통계학과. E-mail: bcseong@cau.ac.kr