Mühendislik Bilimleri ve Tasarım Dergisi
9(1), 1 – 13, 2021
e-ISSN: 1308-6693

Journal of Engineering Sciences and Design
DOI: 10.21923/jesd.837788

Araştırma Makalesi | Research Article

MULTI-STEP FORWARD FORECASTING OF ELECTRICAL POWER GENERATION IN LIGNITE-FIRED THERMAL POWER PLANT

Alper KEREM[1*], İsmail KIRBAŞ[2]

[1]Kahramanmaras Sütçü İmam University, Electrical Electronics Engineering, Kahramanmaraş, Turkey
[2]Burdur Mehmet Akif Ersoy University, Computer Engineering, Burdur, Turkey

Keywords
ARIMA,
NAR,
Artificial Neural Network,
Time Series,
Multi-Step Forward Forecast.

Abstract
This paper presents multi-step forward forecasting studies using real-time generated electrical power time series. Nonlinear Automatic Regression (NAR) and Autoregressive Integrated Moving Average (ARIMA) models were created and applied to the generator power time series produced in Afşin-Elbistan Thermal B Plant. The data were divided into three categories as raw, 10-moving average and 20-moving average while the number of forwarding steps has been established as 6-step forward, 12-step forward and 20-step forward. Performance results of NAR and ARIMA models were presented with 6 scenarios, and then, the results were compared with tables and graphs. As a result of all studies, it has been observed that the model's success was greatly affected by moving average and forward steps parameters.

LİNYİT YAKITLI TERMİK SANTRALDE ELEKTRİK ENERJİSİ ÜRETİMİNİN ÇOK ADIMLI İLERİ TAHMİNİ

Anahtar Kelimeler
ARIMA,
NAR,
Yapay Sinir Ağları,
Zaman Serisi,
Çok-Adımlı İleri Tahmin.

Öz
Bu çalışmada gerçek zamanlı üretilen elektriksel güç zaman serilerinin kullanılmasıyla çok adımlı ileriye dönük tahmin çalışmaları anlatılmaktadır. Doğrusal Olmayan Otoregresif (NAR) ve Otoregresif Hareketli Ortalama (ARIMA) modelleri oluşturulmuş ve Afşin-Elbistan Termik B Santralinde üretilen generatör güç zaman serilerine uygulanmıştır. Veriler ham, 10 hareketli ortalama ve 20 hareketli ortalama olarak üç kategoriye ayrılırken, adım sayısı 6 adım ileri, 12 adım ileri ve 20 adım ileri olarak belirlenmiştir. NAR ve ARIMA modellerinin performans sonuçları 6 senaryo ile oluşturulmuş, ardından sonuçlar tablo ve grafikler ile karşılaştırılmıştır. Tüm çalışmalar sonucunda, hareketli ortalama ve ileri adım sayısı parametrelerinin model başarısını büyük ölçüde etkilediği görülmüştür.

Alıntı / Cite
Kerem, A., Kırbaş, İ., (2021). Multi-Step Forward Forecasting of Electrical Power Generation in Lignite-Fired Thermal Power Plant, Journal of Engineering Sciences and Design, 9(1), 1-13.

Yazar Kimliği / Author ID (ORCID Number)
A. Kerem, 0000-0002-9131-2274
İ. Kirbaş, 0000-0002-1206-8294

Makale Süreci / Article Process	
Başvuru Tarihi / Submission Date	08.12.2020
Revizyon Tarihi / Revision Date	22.12.2020
Kabul Tarihi / Accepted Date	03.02.2021
Yayım Tarihi / Published Date	30.03.2021

1. Introduction

Electrical energy is vital for human life to survive. However, the rapidly growing global population and developing technology brings several concerns for energy supplies (Kırbaş and Kerem, 2016). Coal, natural gas, nuclear, oil and renewable energy sources are among the main sources of power production. And, energy consumption can be detailed according to residences, industry, trade and other user communities (Rahman and Esmailpour, 2015). Energy supply, demand and prices have become more changeable and unpredictable in recent competitive

* İlgili yazar / Corresponding author: alperkerem@ksu.edu.tr, +90-344-300-1702

conditions (Hong et al., 2016). Thus, energy forecasting studies are carried out in order to make these complexities more stable. Furthermore, it's aimed to balance the energy demand; determining the amount of imported power (Rahman and Esmailpour, 2015); managing the electrical power systems with accurately and reliably; ensuring the efficient operation of electrical power systems (Bracale et al., 2014); integrating power plants into the market; balancing the demand for auxiliary service and cost; increasing to power quality, transmission capacity, stability and reliability (Yuan-Kang and Jing-Shan, 2007) using forecasting studies.

Forecasting models consist of physical method, statistical method and hybrid method (Mao and Shaoshuai, 2016). Physical method is a method that uses weather data such as temperature, pressure, surface roughness and obstacles (Chang, 2014). The statistical method based on time series (Mao and Shaoshuai, 2016) that uses the difference between actual data and recently estimated data (Yuan-Kang and Jing-Shan, 2007). Time series forecasting is a technique that uses historical data to predict future problems in science, engineering, economics and many other applications (Niu et al., 2010). The hybrid method aims to increase the prediction accuracy and decrease the error rate by taking advantages of each model that creates it (Bhaskar and Singh, 2012).

This study is based on forecasting of real-time generated electrical power time series. Thus, forecasting studies which use time-series in related literature, have been examined. Accordingly, Rahman and Esmailpour (2015) designed Backpropagation Neural Network (BP+NN) model for electricity generation forecasting. Obradovic and Tomsovic (1999) used time series model to estimate the electricity market price. Amjady et al. (2010) designed the Short-Term Load Forecast (STLF) model for short-term load estimates. They created this new hybrid model using Enhanced Differential Evolution (EDE), Feature Selection (FS) and NN+EA models. Hossain and Mahmood (2020) designed Long Short-Term Memory Neural Network (LSTM+NN) model for electrical load estimations. They compared model performance with Extreme Learning Machine (ELM) and Generalized Regression Neural Network (GRNN) models. Jeong (2020) used Gaussian Process Regression (GPR) model to predict energy trade in microgrid. Liu et al. (2020) performed the electricity market price prediction using LSTM model. Bálint et al. (2019) have forecasted power generation of solar panels using a weather forecast for microgrid application. Sahu et al. (2019) used Nonlinear Automatic Regression (NAR) and Nonlinear Automatic Regression Exogenous (NARX) models in the wind and solar generation forecasting studies. Sfetsos (2000) has applied the models for wind speed forecasting separately, Adaptive Neuro-Fuzzy Inference Systems (ANFIS), Autoregressive Integrated Moving Average (ARIMA), Backpropagation Neural Network (BP+NN), Levenberg Marquardt Neural Network (LM+NN), Neural Logic Network Logic Rules (NLN+LR) and Radial Basis Function (RBF). Taylor et al. (2009) designed Autoregressive Moving Average Autoregressive Conditional Heteroskedasticity (ARMA+GARCH) and Autoregressive Fractionally Integrated Moving Average Autoregressive Conditional Heteroskedasticity (ARFIMA+GARCH) models for wind power density estimations. Shi (2012) performed wind speed and wind power forecast using ARIMA+ANN and ARIMA+Support Vector Machine (SVM) models. Kırbaş (2018) developed ARIMA and NAR models for wind speed forecasting. Liu et al. (2013) designed Wavelet Packet Broyden Fletcher Goldfarb Shanno (WP+BFGS), WP+ARIMA+BFGS and Wavelet+BFGS models for wind speed forecast. They compared model forms with Neuro Fuzzy (NF), ANFIS, Wavelet Radial Basis Function (W+RBF) and Persistent Model (PM). Hu et al. (2013) developed a hybrid model for wind speed estimates using Ensemble Empirical Mode Decomposition (EEMD) and SVM. Abdel-Aal et al. (2009) used Abductive Networks (AN) and NN models for wind speed estimation. Kerem and Kırbaş (2019) designed NAR model for multi-step forward wind speed forecasting. The related literature studies are presented in Table 1.

In this study multi-step forward forecasting studies were carried out using real-time electrical power generated in Afşin-Elbistan B Thermal Power Plant. Performances results of designed ARIMA and NAR models were presented with 6 different scenarios.

Table 1. Literature for time series applications in electricity studies

method	problem	reference
BP+NN	electricity generation forecast	(Rahman and Esmailpour, 2015)
time series	electricity market price forecast	Obradovic and Tomsovic (1999)
STLF EDE NN+EA ARIMA ARIMA+W AWNN	electricity load forecast of microgrids	Amjady *et al.* (2010)
LSTM+NN GRNN ELM	electrical load forecast	Hossain and Mahmood (2020)
GPR	energy trading forecast of microgrids	Jeong (2020)
LSTM	electricity market price forecast	Liu *et al.* (2020)
time series	solar generation forecast	Bálint et al. (2019)
NAR NARX BR	wind and solar generation forecast	Sahu et al. (2019)
ANFIS ARIMA BP+NN LM+NN NLN+LR RBF	wind speed forecast	Sfetsos (2000)
ARMA+GARCH ARFIMA+GARCH	wind power density forecast	Taylor et al. (2009)
ARIMA+ANN ARIMA+SVM	wind speed forecast wind power forecast	Si (2012)
ARIMA NAR	wind speed forecast	Kırbaş (2018)
W+BFGS WP+BFGS WP+ARIMA+BFGS NF ANFIS WP+RBF PM	wind speed forecast	Liu *et al.* (2013)
EEMD+SVM	wind speed forecast	Hu *et al.* (2013)
AN NN	wind speed forecast	Abdel-Aal *et al.* (2009)
NAR	wind speed forecast	Kerem and Kırbaş (2019)

Abbreviations - AN:Abductive Networks, ANN:Artificial Neural Network, ANFIS:Adaptive Neuro-Fuzzy Inference Systems, ARIMA:Autoregressive Integrated Moving Average, ARMA+GARCH:Autoregressive Moving Average Autoregressive Conditional Heteroskedasticity, ARFIMA+GARCH:Autoregressive Fractionally Integrated Moving Average Autoregressive Conditional Heteroskedasticity, AWNN: Adaptive Wavelet Neural Network, BFGS:Broyden Fletcher Goldfarb Shanno, BPNN:Backpropagation Neural Network, BR:Bayesian Regularization, DE:Differential Evolution, EA:Evolutionary Algorithm, EDE:Enhanced Differential Evolution, EEMD:Ensemble Empirical Mode Decomposition, ELM:Extreme Learning Machine, GPR:Gaussian Process Regression, GRNN:Generalized Regression Neural Network, LM:Levenberg Marquardt, LSTM+NN:Long Short Term Memory Neural Network, LR:Logic Rules, NAR:Nonlinear Automatic Regression, NARX:Nonlinear Automatic Regression Exogenous, NF:Neuro+Fuzzy, NLN:Neural Logic Network, NN:Neural Network, PM:Persistent Model, RBF: Radial Basis Function, STLF:Short-Term Load Forecast, SVM:Support Vector Machine, W:Wavelet, WP:Wavelet Packet

2. Data From Lignite-Fired Thermal Power Plant

Turkey's installed electrical power is 91.342MW as of 2020. The biggest one is the thermal power plants with 46500,09MW power (and 51% ratio) among the other sources. Table 2 shows Turkey's installed electrical power.

Table 2. Installed Power of Turkey, 2020 (TEY, 2020)

Source	Installed Power (MW)
thermal	46500,09
hydroelectric	28508,06
wind	7609,34
solar	6032,07
biomass, waste heat, geothermal	2692,14
total	91342

Afşin-Elbistan B Thermal Power Plant is located in Kahramanmaraş city of Turkey. It consists of 4 units and each unit has 360MW installed power. The total installed power of the plant, which uses lignite as fuel, is 1440MW. Geographic location of Afşin-Elbistan B Thermal Power Plant is given in Figure 1.

Figure 1. Geographic location of Afşin-Elbistan B thermal power plant

In our study, raw production data obtained from thermal power plant were analyzed as raw and moving averages of 10 and 20. Descriptive statistics data related to the gathered production information are shown in detail in the Table 3.

Table 3. Data obtained from Afşin-Elbistan B thermal power plant

name	length	max	min	mean	mode
raw	298	899,92	135	705,014	811,000
ma 10	298	873,336	276,6	704,983	476,400
ma 20	298	870,276	439,8	704,657	761,700
name	median	skewness	kurtosis	var	std
raw	768,540	-1,206	3,826	26750,289	163,555
ma 10	739,800	-1,090	3,673	16566,712	128,712
ma 20	739,333	-0,894	2,824	10555,958	102,742

3. Methodology

In this study, the electricity production data of the thermal power plant was considered as a time series and 6 different models were developed and the future value of the production was estimated. In order to measure the performance of the prediction models, the raw data has been converted into 10 and 20 moving averages and the effect of using the moving average on the model performance has also been examined. For the forecasting, 3 different levels (6, 10 and 12) are determined as prediction steps. Three main scenarios, consisting of different input data types and different estimation steps, were adapted to ARIMA and NAR methods and a total of 6 different study scenarios were obtained. The scenarios and forecast models used are shown in Figure 2.

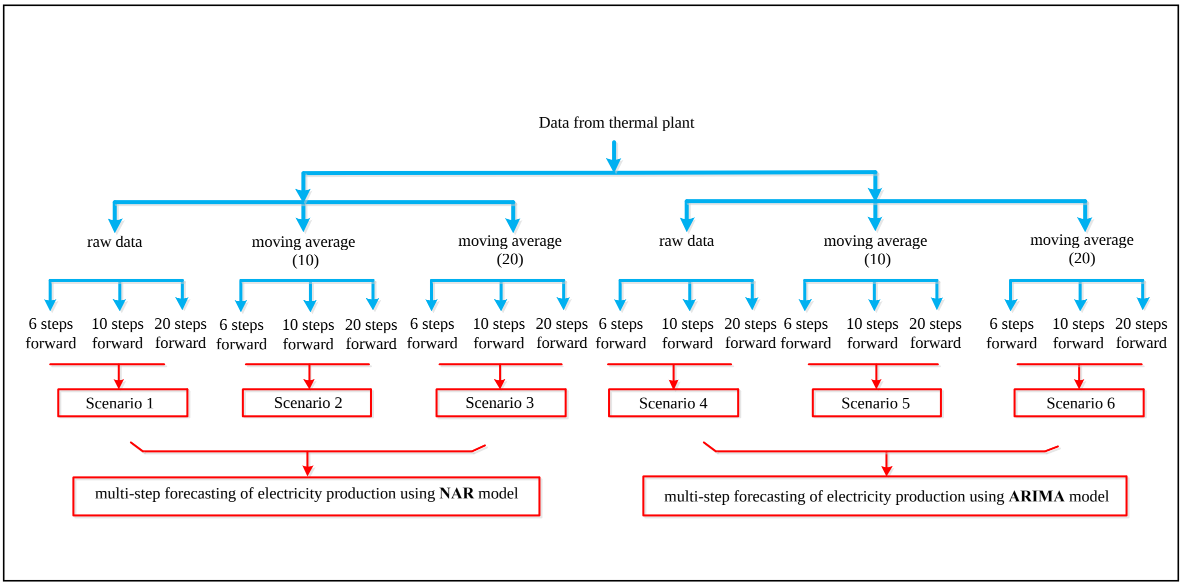

Figure 2. Scenarios for NAR and ARIMA models

Time series are numerical quantities in which the values of variables are observed sequentially from one period to another. It is not a condition that the observed data occur sequentially over time, but it is necessary to see the development of the sequence at regular intervals. The data used in business, economics, engineering, environmental sciences, medicine and many other scientific researches are collected in time series form. The series of observations for such a series, for example, measuring hourly temperatures, daily stock prices, weekly stocks monitoring, monthly diesel consumption or annual growth rates are compiled or collected at regular intervals of time. One of the distinctive features of the time series data being different from other series data is that the observed values in the series are interdependent during the time period. Therefore, statistical methods and techniques that assume that the observations are independent of each other are not valid for time series (Sevüktekin and Çınar, 2017).

Relationship between realization (observed sample values) and process (a specific stochastic process) in time series analysis; similar to studies in statistical hypothesis testing, it is similar to the relationship between sample and population. Therefore, it can be said that a time series is a sample obtained from a certain stochastic process that constitutes the series.

A special aspect of time series analysis is the fact that consecutive observations take into account the time series. When sequential observations are dependent, it is possible to predict the values they will receive in the future from their previous observations. If a time series is fully predictable, it is expressed as a deterministic time series. However, most of the time series are stochastic (probabilistic); that is, the data that the series may receive in the future can be partially defined by its historical values. Precise predictions of the stochastic series can not be achieved, and future values have a distribution of probability conditioned by the knowledge of past values.

Analysis of time series summarizes the characteristics of a series and tries to reveal the outstanding structure of the series. The situation frequently encountered in time series is autocorrelated structures. As it is known, correlation is a measure of acting together between two variables or a non-causal relationship. Autocorrelation, on the other hand, conceptually implies the relationship between the value of a series in any period and the value of moving together between the value of the previous or next period. If a series has autocorrelation, it can be said that there is a correlation between the observations of the series (Sevüktekin and Çınar, 2017).

In the time series, sometimes instantaneous and mostly non-repeating sling-valued (outlier) observations can be encountered. The deviating values (outliers) are, on average, either too small or too large than the observation values present in the series. Therefore, they do not represent the pattern of the past or the future. The deviating values are generally observation values resulting from unusual events and /but not repeated.

ARIMA (AutoRegressive Integrated Moving Average) approach is one of the widely used methods for time series analysis. The reason why the method is so popular is that it can be solved with computer software, whether or not any series is stationary, with or without seasonal factors.

In the ARIMA model, there are basically three parameters *p, q* and *d*. The AR component of ARIMA implies that the changing interest variable regresses to its own previous values. The mean of the error term is zero in a stationary time series and the deviation is σ^2. As in equation 1, the representation of this time series as a p-order autoregressive phase is shown as AR(p) if Y_t indicates the value of the time series at t-time (Kırbaş, 2018).

$$Y_t = \delta + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \cdots + \varphi_P Y_{t-P} + \varepsilon_t \quad (1)$$

where δ is a constant value and ε_t is an error term. If we want to express the time series as a q^{th} degree of moving average process MA(q), we can use equation 2.

$$Y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} \quad (2)$$

If we combine two AR(p) and MA(q) equations we obtain ARMA(p,q) equation 3.

$$Y_t = \delta + \varphi_1 Y_{t-1} + \cdots + \varphi_P Y_{t-P} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} \quad (3)$$

If the time series seeking to process is not constant, it can be done as a stationary by taking the *d* times the difference shape to make it constant. Once the difference is taken from of the non-stationary Y_t series, the ΔY series, which indicates a stationary feature, is obtained in equation 4.

$$\Delta Y_t = Y_t\text{-}Y_{t-1} = Y_t - LY_t = Y_t' \quad (4)$$

The ARIMA (p, d, q) process can be generally expressed by equation 5 (Sevüktekin and Çınar, 2017)

$$\left(1 - \varphi_1 L - \varphi_1 L^2 - \cdots - \varphi_p L^q\right)\Delta^d Y_t = \delta + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} \quad (5)$$

PACF is used to determine the AR parameter value of the ARIMA model, and the correlogram graphs of the ACF functions are used to determine the parameter value of the MA. Step numbers above the specified threshold level are tested for both parameters AR and MA. Figure 3 shows the ACF correlogram used to determine the MA parameter. Figure 4 shows the PACF correlogram used to determine the AR parameter.

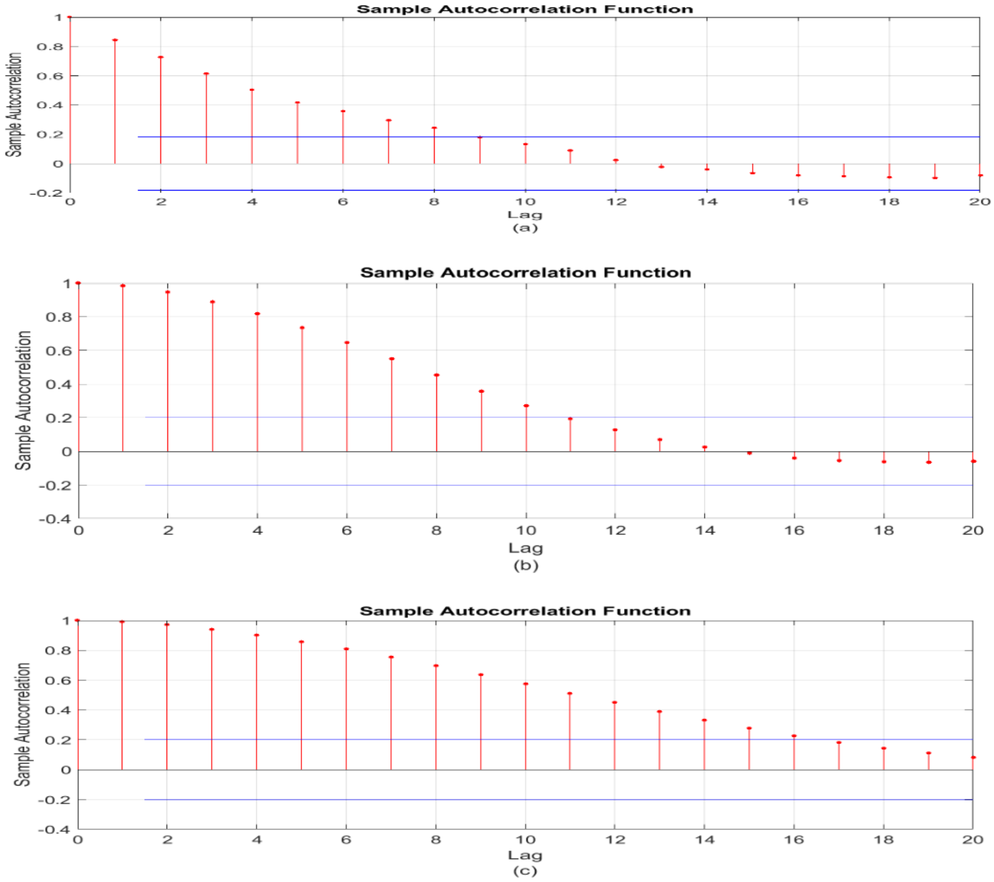

Figure 3. Sample Autocorrelation (a) raw (b) mov avg.10 (c) mov avg. 20 ARIMA

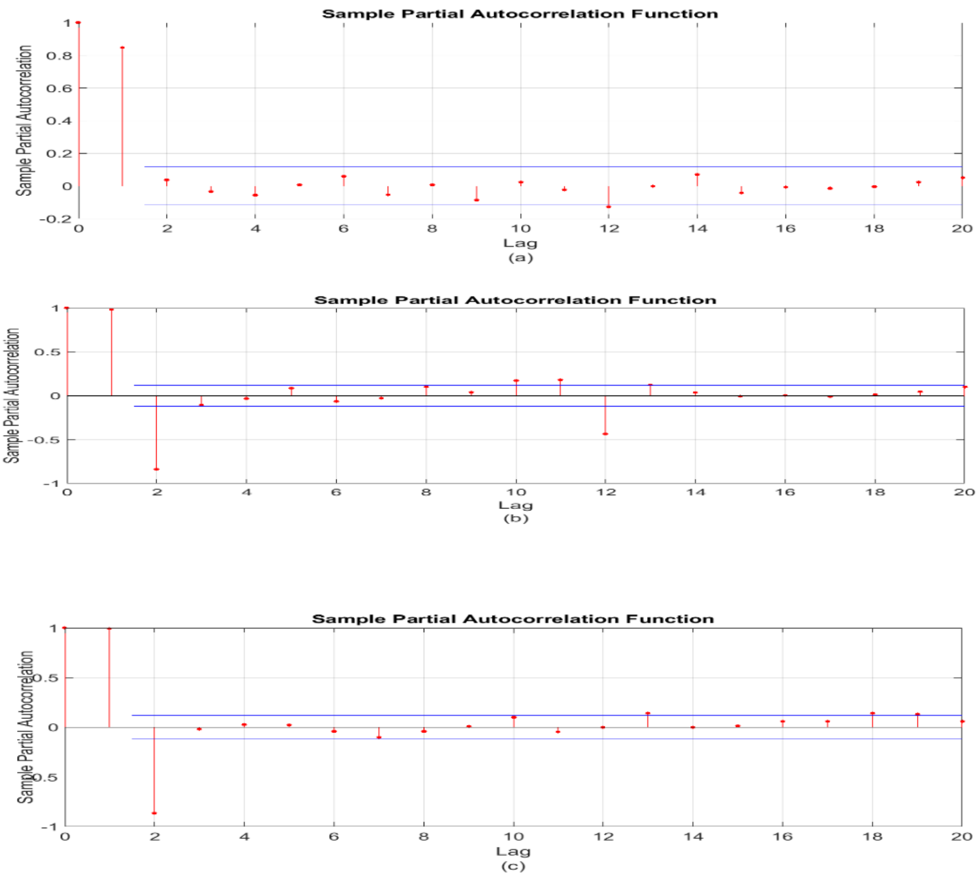

Figure 4. Sample Partial Autocorrelation (a) raw (b) mov avg. 10 (c) mov avg. 20 ARIMA

The efficiency of the model is generally measured by the Akaike Information Criteria (AIC) equation in order to determine the most appropriate parameter in the ARIMA method and is written as (Kırbaş, 2018)

$$AIC = -2\log(L) + 2(p + q + k) \quad (6)$$

where L is the probability of the data, p is the autoregressive portion order, and q is the moving average portion order. The k represents the ARIMA model's intercept. The model with the lowest AIC criterion is deemed to be more successful than the others, according to this metric. The parameters demonstrating the best efficiency were obtained from the ARIMA (2,2,5) model in our analysis.

In the scope of the study, NAR ANNs structure was used as an alternative to ARIMA model. Artificial neural network structure is based on expressing the working and learning behaviour of the human brain with a mathematical model. They are used successfully in many different problems such as pattern recognition, classification, regression, optimization.

The multi-layered artificial neural network structure consists of synapses as in the biological sample. Here a mathematical model of a nerve cell is created. The model includes neural weights and a nonlinear transfer function. There are basically three layers in artificial neural networks. The first layer is the input layer where the data of the input variables are entered. The second layer is called the hidden layer and is fully connected with the neurons in the input layer. Data from the input layer is multiplied by certain weight values and transmitted to the transfer functions of neurons in the hidden layer (Kirbaş and Kerem, 2016).

Similarly, the transfer function outputs of the neurons in the hidden layer form the input data of the neurons in the output layer. After the transfer function of the output layer is performed, the output data of the artificial neural network is obtained. In order to determine the multiplier weights in the artificial neural network, the backpropagation algorithm is used. For this process, data with known input and output values are used for training and an increase or decrease is performed on the multiplier weights until the artificial neural network value calculates the correct output value against the known input value. After obtaining the determined performance criteria, the training phase is completed.

NAR is a frequently used approach especially in time series estimations. As training data and multiplier weights in the artificial neural network are calculated, this artificial neural network uses a certain part of the time series (Kırbaş, 2018).

In order to measure model performance, the values that neural network does not know during training are given as an input and asked to estimate the result. The NAR approach assumes that the value of y in time t (y_t) is a function of the past d number, as seen in equation 7 (Kerem and Kırbaş, 2019).

$$y_t = f(y_{t-1}, \dots y_{t-d}) \quad (7)$$

and this unknown f function tried to be modelled with artificial neural network. Figure 5 shows the NAR 2-delay model consisting of 10 neurons. The future value is estimated by this neural network by focusing on two historical data.

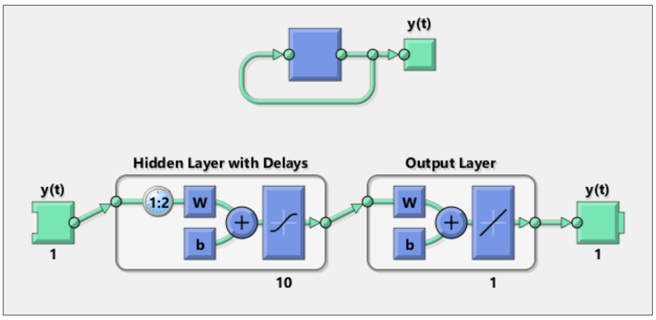

Figure 5. Main structure of NAR neural network model.

The values determined by the neural network are compared with the results previously known and the difference is looked at in order to assess the model's efficiency. For high efficiency, it is required that the value of the difference be close to zero. The seven criteria used to measure the performance of the prediction models in Table 4 are given together with their formulas.

Table 4. Model performance criteria (Kerem *et al.,* 2016; Kırbaş and Kerem, 2016; Kırbaş, 2018; Kerem and Kırbaş, 2019)

Model	Equation
Mean squared error	$MSE = \frac{1}{n}\sum_{i=1}^{n}(\hat{Y}_i - Y_i)^2$
Peak signal-to-noise ratio	$PSNR = 20\log_{10}\left(\frac{MAX_f}{\sqrt{MSE}}\right)$
Root-mean-square error	$RMSE = \sqrt{\frac{SSE}{n}} = \sqrt{\frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{n}}$
Normalized root-mean-square error	$NRMSE = \frac{RMSD}{Y_{max} - Y_{min}}$
Mean absolute percentage error	$MAPE = \frac{100}{n} \times \sum_{i=1}^{n}\left\lvert\frac{Y_i - \hat{Y}_i)}{Y_i}\right\rvert$
Symmetric mean absolute percentage error	$SMAPE = \frac{200}{n} \times \sum_{i=1}^{n}\left\lvert\frac{Y_i - \hat{Y}_i}{Y_i + \hat{Y}_i}\right\rvert$

4. Performance Results

The question of how successful the prediction models developed in the study is very important. In order to make a fair comparison and examine the factors affecting the success of the forecast in more detail, a time series consisting of 298 measurements was created and forward predictions were made using 6 different scenarios. The performance values obtained for the forward estimation made in 6, 10 and 20 forward steps are given in Table 5.

Table 5. Scenario 1 results (NAR)

forward steps	values	MSE	PSNR	R	RMSE	NRMSE	MAPE	SMAPE
6 steps	best	3,06E+03	13,269	0,995	55,342	1,100	6,526	6,776
	average	1,25E+04	9,137	0,984	93,276	1,855	10,575	11,374
	worst	2,13E+04	4,851	0,966	145,876	2,901	16,280	18,128
10 steps	best	1,98E+03	15,171	0,996	44,458	0,413	4,832	5,001
	average	4,07E+03	10,132	0,987	83,580	0,777	9,285	9,630
	worst	1,48E+03	6,423	0,975	121,717	1,131	15,135	16,459
20 steps	best	2626,000	13,936	0,995	51,251	0,140	6,415	6,384
	average	2,14E+03	8,028	0,978	104,241	0,285	12,274	12,033
	worst	2319,00	4,477	0,955	152,288	0,416	16,658	18,747

Scenario 1: (10 run), raw data
delay of training: 8, number of neurons: 12, number of data: 298

When the results of Scenario 1 NAR model are examined, it is seen that the best estimate is obtained in 10 steps. Figure 6 shows the estimates and actual values of the NAR model for scenario 1 together. Table 6 also shows scenario 2 NAR results.

Table 6. Scenario 2 results (NAR)

forward steps	values	MSE	PSNR	R	RMSE	NRMSE	MAPE	SMAPE
6 steps	best	1,98E+02	25,172	0,999	14,057	0,417	1,689	1,673
	average	7,64E+02	20,436	0,998	25,974	0,771	3,132	3,071
	worst	1,39E+03	16,703	0,997	37,271	1,107	4,461	4,352
10 steps	best	9,95E+01	28,151	0,999	9,976	0,100	1,094	1,097
	average	1,74E+03	18,700	0,996	36,234	0,366	4,028	4,092
	worst	4,11E+03	11,996	0,992	64,077	0,647	7,106	7,463
20 steps	best	551,091	20,718	0,998	23,475	0,124	2,900	2,951
	average	2,12E+03	15,921	0,995	43,542	0,230	5,066	4,941
	worst	4185,00	11,913	0,991	64,693	0,341	8,024	7,618

Scenario 2: (10 run), moving average: 10,
delay of training: 8, number of neurons: 12, number of data: 298

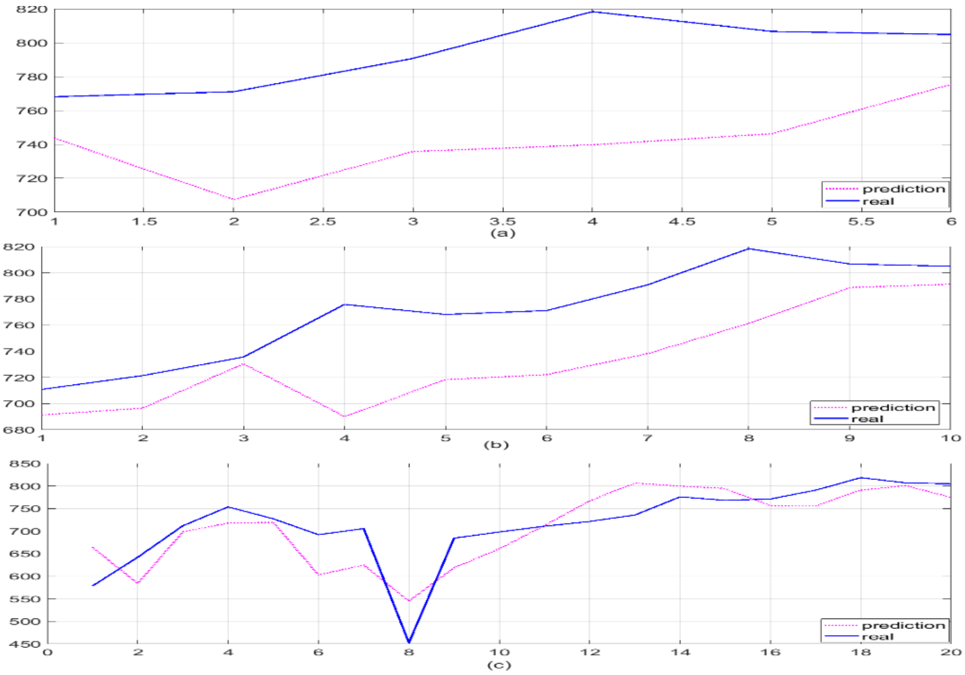

Figure 6. Scenario 1 results (NAR) (a) 6 step-forward (b) 10 step-forward (c) 20 step-forward

When the results are evaluated, it is seen that the 10-step estimate includes the lowest error value. The results of NAR model for Scenario 2 are given in Figure 7.

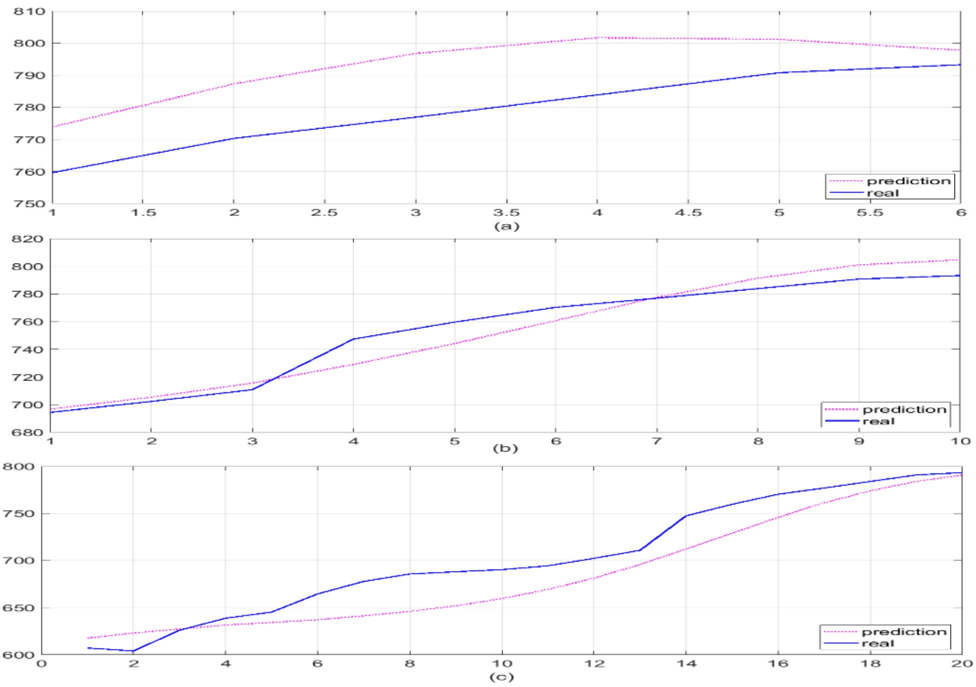

Figure 7. Scenario 2 results (NAR) (a) 6 step-forward (b) 10 step-forward (c) 20 step-forward

Scenario 5 results are given in Table 7. Accordingly, the lowest estimation error is seen in the 6-step estimate.

Table 7. Scenario 3 results (NAR)

forward steps	values	MSE	PSNR	R	RMSE	NRMSE	MAPE	SMAPE
6 steps	best	42,717	31,824	0,999	6,535	0,187	0,843	0,842
	average	4,89E+02	25,357	0,998	17,298	0,496	1,880	1,919
	worst	2,46E+03	14,221	0,995	49,598	1,424	5,323	5,549
10 steps	best	1,53E+02	26,283	0,999	12,369	0,267	1,020	1,033
	average	2,88E+02	24,158	0,999	16,389	0,354	1,824	1,807
	worst	5,71E+02	20,563	0,998	23,897	0,516	2,635	2,587
20 steps	best	1,65E+02	25,966	0,999	12,830	0,087	1,527	1,535
	average	2,48E+03	16,607	0,994	44,417	0,303	5,284	5,096
	worst	7,10E+03	9,616	0,985	84,281	0,576	9,857	9,408

Scenario 3: (10 run), moving average: 20
delay of training: 8, number of neurons: 12, number of data: 298

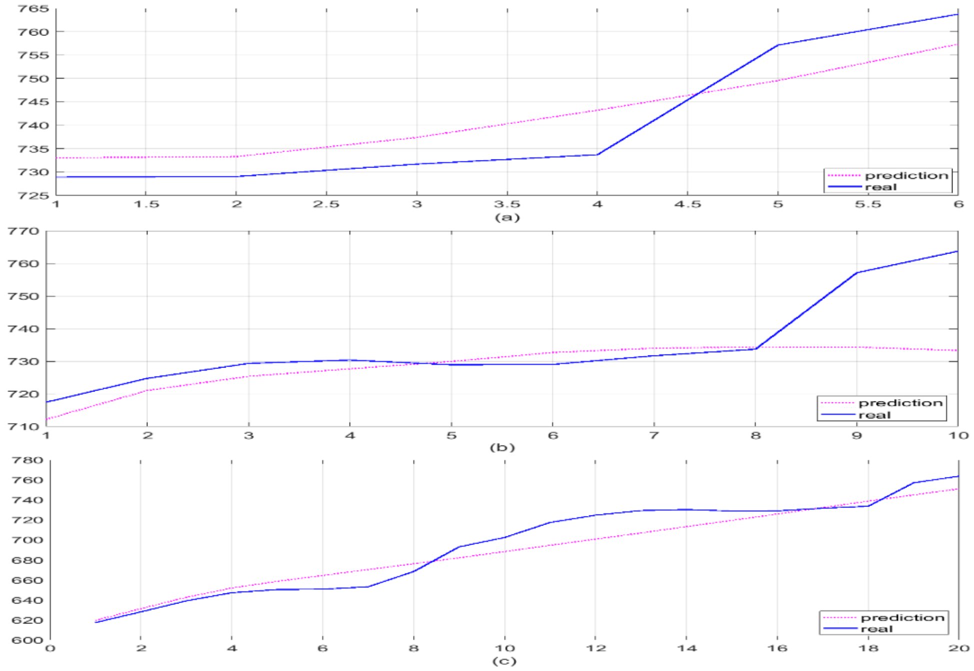

Figure 8. Scenario 3 results (NAR) (a) 6 step-forward (b) 10 step-forward (c) 20 step-forward

Table 8 shows the results for scenarios 4, 5 and 6. Accordingly, 10-step estimation of scenario 6 gives the most successful result, while the worst estimate was realized in the 20-step process of scenario 4-5-6.

Table 8. Scenario 4-5-6 results ARIMA(2,2,5)

	frd step	MSE	PSNR	R value	RMSE	NRMSE	MAPE	SMAPE	SMAPE
Scenario 4	6 step	11110	7,674	0,982	105,402	2,096	11,273	12,203	1,392
	10 step	21904	4,726	0,963	147,999	1,376	17,180	19,135	1,375
	20 step	27570	3,726	0,947	166,043	0,454	21,687	24,177	2,432
Scenario 5	step 6	480,51	21,314	0,999	21,921	0,651	2,470	2,432	3,751
	step 10	4660,6	11,446	0,992	68,268	0,690	7,516	7,921	6,764
	step 20	4137,8	11,963	0,992	64,326	0,340	6,396	6,764	7,921
Scenario 6	step 6	152,59	26,296	1,000	12,353	0,355	1,383	1,392	10,026
	step 10	188,5	25,378	1,000	13,729	0,296	1,359	1,375	10,140
	step 20	6191,2	10,213	0,987	78,684	0,538	9,404	10,026	12,203

Scenario 4: (10 run), raw data
Scenario 5: (10 run), moving average: 10
Scenario 6: (10 run), moving average: 20
number of data: 298

Figures 9 show 3 different scenario results for the 10 and 11 ARIMA models.

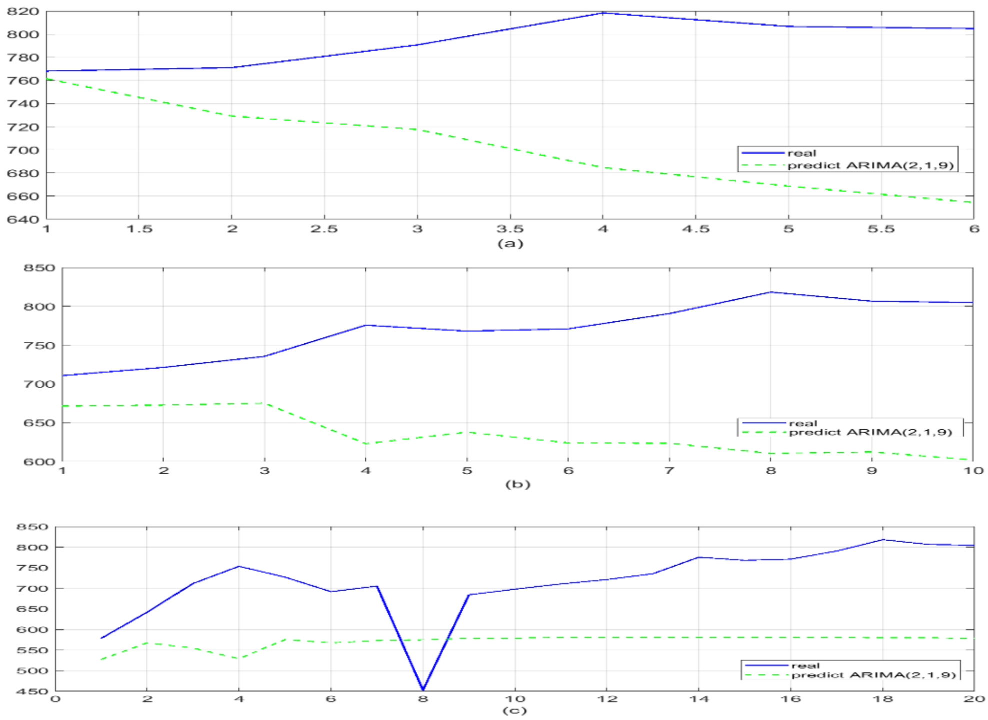

Figure 9. Scenario 4 results (ARIMA) (a) 6 step-forward (b) 10 step-forward (c) 20 step-forward

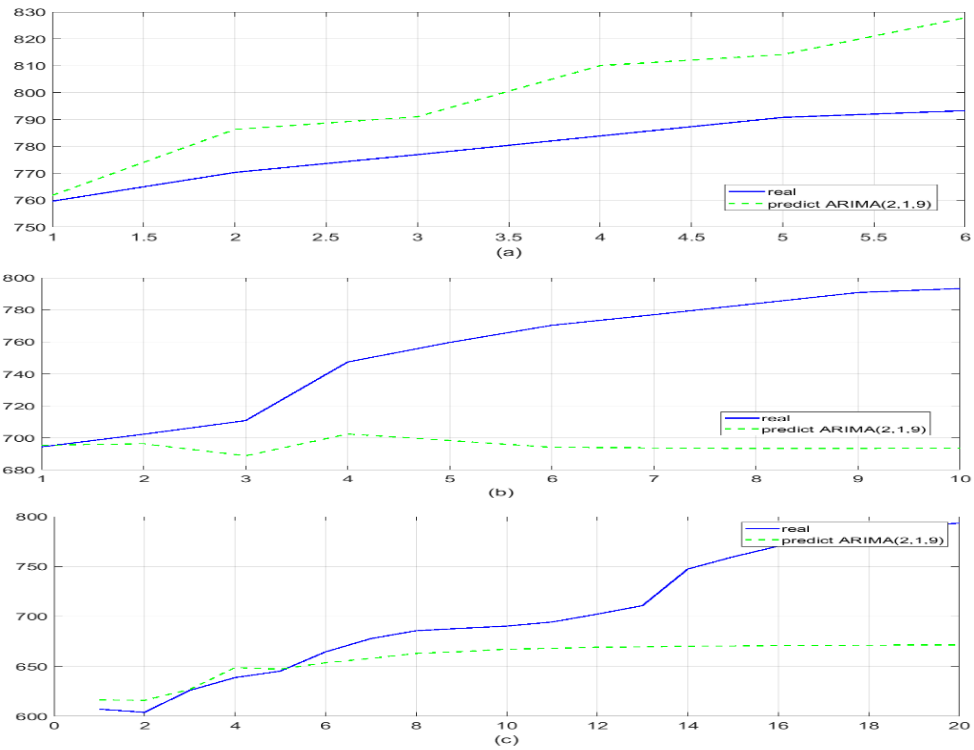

Figure 10. Scenario 5 results (ARIMA) (a) 6 step-forward (b) 10 step-forward (c) 20 step-forward

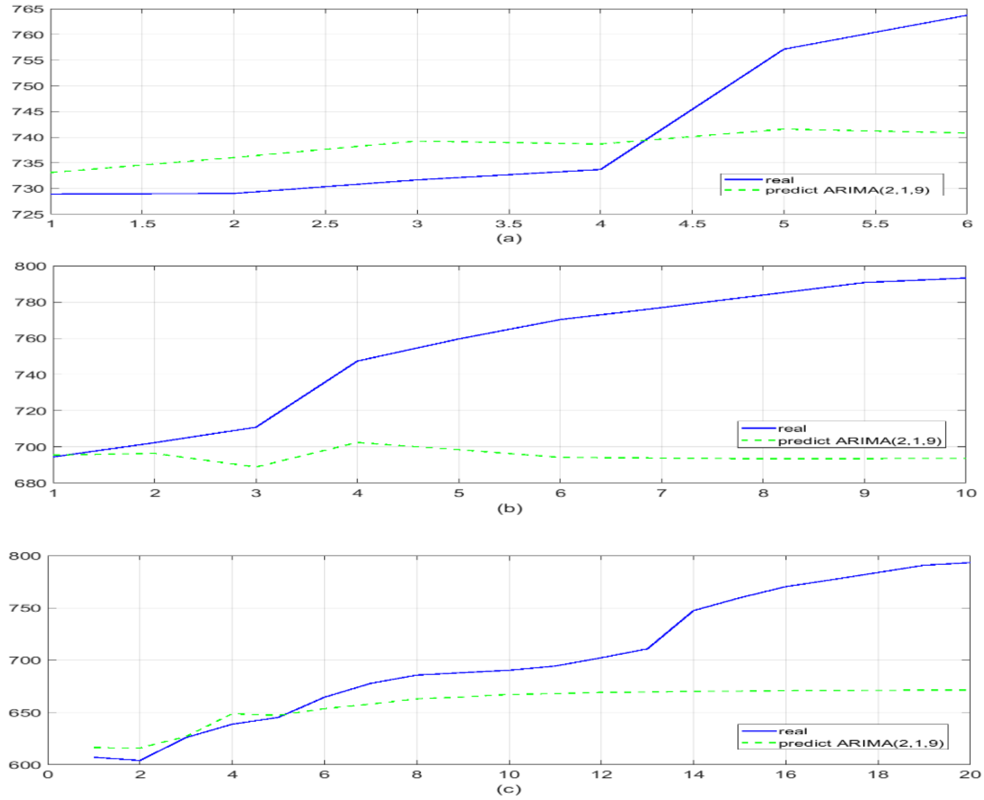

Figure 11. Scenario 6 results (ARIMA) (a) 6 step-forward (b) 10 step-forward (c) 20 step-forward

5. Conclusion

In this study, it is aimed to forecast the production data from a thermal power plant in 6, 10 and 20 steps prospectively. For this aim, two different prediction models based on ARIMA and NAR models have been developed. In order to compare the performance of the models, six different scenarios have been created and the results are compared with tables and graphs.

NAR model seems to be more successful than ARIMA model when tables and graphs are examined. The number of prediction steps is also vital in multi-step estimation. Estimation performance decreases as the number of steps increases. It is quite difficult to find an exact number for the most suitable prediction number. When the model results are examined, it is seen that the best model is the NAR model used in scenario 3. Since the NAR model is based on artificial neural network, it contains randomness and different performance rates can be observed in each training steps. For a fairer evaluation, the same scenario was run 20 times in a row, and the best, worst and average values obtained from these 20 studies were found. Another outcome obtained from the results is that the NAR model can make more flexible predictions than the ARIMA model. The use of moving averages in the time series also clearly affects the forecast performance positively.

The results of the study show that the thermal power plant production data can be estimated prospectively using ARIMA or NAR based time series prediction models.

Acknowledgements

The authors thank to Electricity Generation Company (EÜAŞ) and directorate of Afşin-Elbistan B Power Plant for for their cooperation in providing data.

Conflict of Interest

No conflict of interest was declared by the authors.

References

Abdel-Aal R.E., Elhadidy M.A., Shaahid S.M., 2009. Modeling And Forecasting the Mean Hourly Wind Speed Time Series Using GMDH-Based Abductive Networks, Renewable Energy 34: 1686-1699.

Amjady A., Keynia F., Zareipour H. 2010., Short-Term Load Forecast of Microgrids by a New Bilevel Prediction Strategy, IEEE Transactıons on Smart Grid, 1(3):286-294.

Bálint R., Fodor A., Magyar A. 2019., Model-Based Power Generation Estimation of Solar Panels Using Weather Forecast for Microgrid Application, Acta Polytechnica Hungarica, 16 (7): 149-165.

Bhaskar K. and Sing S. N. 2012., AWNN-Assisted Wind Power Forecasting Using Feed-Forward Neural Network. The Institute of Electrical and Electronics Engineers Transactions on Sustainable Energy, 3(2): 306-315.

Bracale A., Carpinelli G., Rizzo R., Russo A. 2014., Advanced Method and Cost-Based Indices for Probabilistic Forecasting the Generation of Renewable Power, 3rd Renewable Power Generation Conference (RPG 2014), 24-25 Sept. 2014, Naples, Italy.

Chang W.Y. 2014., A Literature Review of Wind Forecasting Methods, Journal of Power and Energy Engineering, 2:161-168.

Hong T., Pinson P., Fan S., Zareipour H., Troccoli, A., Hyndmanc, R. J. 2016., Probabilistic Energy Forecasting: Global Energy Forecasting Competition 2014 and Beyond, International Journal of Forecasting, 32: 896-913.

Hossain M.S., Mahmood H. 2020., Short-Term Load Forecasting Using an LSTM Neural Network, 2020 IEEE Power and Energy Conference at Illinois (PECI), 27-28 Feb. 2020 Champaign, IL, USA.

Hu J., Wang J., Zeng G. 2013., A Hybrid Forecasting Approach Applied to Wind Speed Time Series, Renewable Energy, 60:185-194.

Jeong G., Park S., Hwang G. 2020., Time Series Forecasting Based Day-Ahead Energy Trading in Microgrids: Mathematical Analysis and Simulation, IEEE Access, 10.1109/ACCESS.2020.2985258, 8(20): 63885-63900.

Kerem A., Kirbas İ., Saygın A. 2016., Performance Analysis of Time Series Forecasting Models for Short Term Wind Speed Prediction, International Conference on Engineering and Natural Sciences (ICENS), 24-28 May 2016, Sarajevo, Bosnia-Herzegovina.

Kerem A., Kırbaş İ. 2019., Doğrusal Olmayan Otoregresif Ağ (NAR) Modelinin Gerçek Zamanlı Rüzgar Hızı Zaman Serilerine Uygulanması, International Symposium on Advanced Engineering Technologies (ISADET), 02-04 May, Kahramanmaraş, Turkey.

Kırbaş I., Kerem A. 2016., Short-Term Wind Speed Prediction Based on Artificial Neural Network Models, Measurement and Control, 49 (6) 183-190.

Kırbaş İ. 2018., İstatistiksel Metotlar ve Yapay Sinir Ağları Kullanarak Kısa Dönem Çok Adımlı Rüzgâr Hızı Tahmini, Sakarya Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 22(1): 24-38.

Liu H., Tian H., Pan D., Li Y. 2013., Forecasting Models for Wind Speed Using Wavelet, Wavelet Packet, Time Series and Artificial Neural Networks, Applied Energy, 107:191-208.

Liu S., Zhang L., Zou B. 2020., Study on Electricity Market Price Forecasting With Large-Scale Wind Power Based on LSTM, 6th International Conference on Dependable Systems and Their Applications (DSA), 3-6 Jan. 2020, Harbin, China.

Mao Y. and Shaoshuai W. 2016., A Review of Wind Power Forecasting & Prediction. Probabilistic Methods Applied to Power Systems (PMAPS), 2016 International Conference, Beijing, China.

Niu D., Shi H., Li J., Wei Y. 2010., Research on Short-Term Power Load Time Series Forecasting Model Based on BP Neural Network, 2nd International Conference on Advanced Computer Control, 27-29 March 2010, Shenyang, China.

Obradovic Z., Tomsovic, K. 1999., Time Series Methods for Forecasting Electricity Market Pricing, IEEE Power Engineering Society Summer Meeting, Conference Proceedings (Cat. No.99CH36364), 18-22 July 1999, Edmonton, Alta, Canada.

Rahman M.N., Esmailpour A. 2015., An Efficient Electricity Generation Forecasting System Using Artificial Neural Network Approach With Big Data, 2015 IEEE First International Conference on Big Data Computing Service and Applications, 213-217.

Sahu M. K., Sahoo B., Khatoi M., Behera S. 2019., Short-Term Wind And PV Generation Forecasting of Time-Series Using ANN, International Conference on Intelligent Computing and Control Systems (ICCS), 15-17 May 2019, Madurai, India.

Sevüktekin M., Çınar M. 2017., Ekonometrik Zaman Serileri Analizi: Eviews Uygulamalı, Dora Yayıncılık, 1- 667.

Sfetsos A. 2000., A Comparison of Various Forecasting Techniques Applied to Mean Hourly Wind Speed Time Series, Renewable Energy, 21: 23-35

Shi J., Guo J., Zheng S. 2012., Evaluation of Hybrid Forecasting Approaches for Wind Speed and Power Generation Time Series, Renewable and Sustainable Energy Reviews,16: 3471-3480.

Taylor J.W., McSharry P.E., Buizza R. 2009., Wind Power Density Forecasting using Ensemble Predictions and Time Series Models, IEEE Transactions on Energy Conversion, 24(3): 775-782.

Türkiye Elektrik Yatırımları 2020 Yılı Ocak Ayı Özet Raporu, Enerji İşleri Genel Müdürlüğü Yatırımlar Dairesi Başkanlığı, Ocak 2020, 1-4, Ankara, Turkey

Yuan-Kang W., Jing-Shan H. 2007., A Literature Review of Wind Forecasting Technology in The World, 2007 IEEE Lausanne Power Tech, 1-5 July 2007, Lausanne, Switzerland.