Time Series Prediction Using Hybridization of AR, SETAR and ARM Models

Dursun AYDIN
Öznur İŞÇİ GÜNERİ[§]
Department of Statistics
Science Faculty
University of Muğla Sıtkı Koçman
Muğla, 48000, Turkey

Abstract

It is known that parametric and nonparametric methods are used for nonlinear time series. In recently, hybrid models are also considered in time series forecasting. In this paper we present the hybrid models whose components are parametric and nonparametric models. Of the parametric methods, autoregressive (AR) model and self-threshold value (SETAR) model and, of the nonparametric methods, additive regression model (ARM) and hybrid AR&AAR, AR&SETAR, AAR&SETAR, SETAR&AR, SETAR&AAR and AAR&AR models are used in this study. In this context, back fitting algorithm based on smoothing spline method in the existing literature is discussed. A comparison has been made for the performance of the models obtained for the export volume index numbers and domestic producer price index data for Turkey. These results showed that AAR&SETAR hybrid model has denoted the best performance among the all models in time series forecasting.

Keywords: Time series, AR model, SETAR model, ARM model, Hybrid models

1. Introduction

In recently, nonparametric regression methods called as semi parametric and additive regression have become a very useful tool for non-linear data such as time series (Eva et. al., 2000). However, these approaches do not perform well when seasonality is present. The errors of a regression model are generally possible to be used in the estimations of auto correlated time series. The studies regarding to this subject is performed by Engle, Granger, Rice and Weiss (1986); Harvey and Koopman (1993). When the errors are auto correlated, non-parametric estimation methods can be used. Altman (1990); Hurvich and Zeger (1990); Hart (1991, 1994) have correlated the time with single-variable non-parametric methods as the independent variable. Autoregressive model for the errors (Smith, Wong and Kohn, 1998),

$$y_t = f(x_t) + u_t, \quad t = 1, ..., n \tag{1}$$

can be identifies as above. Here y_t is the dependent variable, $f(x_t)$ is the unknown regression function of the independent variable x_t and u_t is a constant autocorrelation error series. The errors are modeled in zero mean s.th level constant autoregressive process as below:

$$u_t = \theta_1 u_{t-1} + ... + \theta_s u_{t-s} + \varepsilon_t, \ \varepsilon_t \sim NIID(0, \sigma^2) \tag{2}$$

Most of the approaches used frequently in the development of time series models are obtained from a linear Gaussian process (Box, Jenkins and Reinsel, 1994). Among the important causes of such concerns interested in linear Gaussian models providing various attractive properties; physical interpretations that are not successful in producing various nonlinear models, frequency analysis, asymptotic results, statistical inference can be stated.

Despite such advantages, real existence system generally consists of numerous nonlinear properties and such properties cannot be explained completely by the linear statistical models. In other words, linear models are insufficient in explaining specific properties of economical and financial data.

Because economical and financial systems consist of both structural and behavioral changes, it is required to use different time series in order to explain the experimental data at different times. In order to model the nonlinear behavior in the time series; it is required to get different dynamics in different models and to have different models exist. Nonlinear time series have started to gain value in the later years of 1970 and have become famous because of the necessity to model the nonlinear dynamics with real data (Tong, 2007). Starting from this point of view; first of all threshold autoregressive model (TAR) which is formed of different models and is defined by autoregressive (AR) model or simply the Self-Exciting Threshold Model (SETAR) model is taken into consideration. This model provides an important point of view for constant time series. TAR model is formed of piecewise linear models and its general opinion is to change the parameters of a linear AR model according to the value of a visibly single variable which can be named as threshold variable. If this variable is a lagged variable of the time series, it is called as the SETAR model. Another approach in the time series can be given as an additive regression model (ARM) estimation which is a nonparametric autocorrelation model.

Each of the nonlinear components is modeled as a regression spline using more than one nodal point, on the other hand the errors are modeled by a high-degree constant autoregressive process which is parameterized according to the partial autocorrelations. Every non-parametric function forming model (10) which is defined as the additive type of the model given in equation (1) is based on the non-parametric correction methods such as local weighed polynomial (Fan and Gijbels, 1996) and spline correction (Green and Silverman, 1994) or penalized regression spline (Eilers and Marx, 1996) etc and can be estimated by using the "*backfitting algorithm*" given by Hastie and Tibshirani (1990). Every non-parametric term in Model (10) is the functions which are represented by penalized cubic spline regressions. Such functions can also be estimated by using the "*mgcv package*" in R environment, as specified in the study of Wood (2000). A hybrid approach that Tseng et al.(2002) and Zhang(2003) in their studies recommend a hybrid approach that uses autoregressive integrated moving average (ARIMA) . In addition, Aslanargun et al.(2007) demonstrated that hybrid models combines models with two nonlinear components have had the best performance for time series forecasting. Zhang (2003) explains the reasons of using hybrid models in detail. This paper generalizes the hybrid models studied by Zhang(2003) for parametric and nonparametric regression models. Consequently, hybrid model has indicated the best performance among all the models for time series forecasting.

2. Methods Used In the Prediction of Time Series

In modeling the non-linear behaviors, it is possible to show economical and financial time series of different countries with different equations. This section emphasizes on the models assuming the time series determined by an autoregressive AR, self-exciting threshold autoregressive SETAR and ARM models behave differently in every time series.

2.1 AR Model

If the values of the dependent y_t variable regarding to the previous periods consist of $y_{t-1}, y_{t-2}, \ldots, y_{t-p}$ then such models are called as autoregressive model. First-degree autoregressive AR(1) model is given by

$$\begin{aligned} y_t &= \beta_0 + \beta_1 y_{t-1} + u_t \\ u_t &= \phi u_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim NIID(0,\sigma^2) \end{aligned} \tag{3}$$

In first-degree autoregressive statistical model; ϕ is the unknown autocorrelation parameter that is assumed to have a value between -1 and +1 and ε_t is an independent error term with zero mean and a constant σ_ε^2 variance. Statistical model structure given in equation (3) is defined as the AR (1) time series model or the AR (1) process. Similarly, second-degree autoregressive AR (2) model can be written as

$$\begin{aligned} y_t &= \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + u_t \\ u_t &= \phi_1 u_{t-1} + \phi_2 u_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim NIID(0,\sigma^2) \end{aligned} \tag{4}$$

More generally it is possible to define a statistical model stating the *p*-level autoregressive AR process. Thus AR(p) model can be written as follows,

$$\begin{aligned} y_t &= \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \ldots + \beta_p y_{t-p} + u_t \\ u_t &= \phi_1 u_{t-1} + \phi_2 u_{t-1} + \ldots + \phi_p u_{t-p} + \varepsilon_t, \quad \varepsilon_t \sim NIID(0, \sigma^2) \end{aligned} \tag{5}$$

Here, $\phi_1, \phi_2, \ldots, \phi_P$ are the unknown autoregressive parameters (Dagum and Giannerini, 2006).

2.2 SETAR Model

One of the popular methods used in non-linear time series models, TAR (Threshold Auto Regressive) model, was firstly offered by Tong (1978) and then it was considered in detail. SETAR (Self-Exciting TAR) model is a special case of TAR model. Piecewise linear models known as SETAR, develops the simplest class of non-linear models. Simple AR models constituting the SETAR model can be easily estimated by using the regression methods. In order to consider the non-linear behaviors, AR models are expanded and non-linear models become as easily understandable and interpretable. Detailed scope of the SETAR model and its statistical properties can be found in Tong (1990). This model is widely used to asymmetrical modeling of the economical series. For example, Pfann, Schotman and Tschernig (1996) have discussed that the interest ratios in America have more equations than one and estimated such series via the SETAR model in their article. Again such models are used by Henry, Olekalns and summers (2001) for the exchange rate modeling. Also detailed information and examples about the usage of non-linear SETAR models in the time series analysis and the applications in various areas can be found in the articles of Tong and Lim (1980); Tong (1983, 1990); Granger and Terasvirta (1993); Franses and Dijk (2000), Chan and Tong (2001).

Despite the simplicity of TAR model form, numerous parameters should be predicted and the variables should be selected in the formation of a TAR model. This situation has prevented the early usage of the method. However, in recent years, an improvement is shown in the properties and prediction of TAR models (Zivot, 2005). SETAR models have several types. In this article, only the two-equation SETAR model which is formed of two linear sub models related to the situation of the threshold and which is defined by Tong (1990) is taken as a basis. For the ease of presentation, SETAR (1) shows the one-equation linear AR model for k = 1 and SETAR (2) shows the two-equation TAR model for k = 2. For one-equation SETAR (1) model, $-\infty = r_0 < r_1 < \cdots < r_k = \infty$ and unknown parameters $\Theta = (\phi^{(1)}, \sigma^{(1)})$; and for two-equation SETAR (2) model, single threshold value $-\infty < r_1 < \infty$ and unknown parameters $\Theta = (\phi^{(1)}, \phi^{(2)}, \sigma^{(1)}, \sigma^{(2)})$ are shown. Lagging parameter *d* is a positive integer and the threshold variable for the SETAR model is a known lagged value of the process itself. First-degree SETAR model is defined as follows:

$$y_t = \begin{cases} \phi_{1,0} + \phi_{1,1} y_{t-1} + \sigma_1 e_t, \; y_{t-1} \le r \\ \phi_{2,0} + \phi_{2,1} y_{t-1} + \sigma_2 e_t, \; y_{t-1} > r \end{cases} \tag{6}$$

Here ϕ 's are the autoregressive parameters. σ Is the noise standard deviation, r is the threshold value parameter and e_t is the unit variance and zero mean independent and uniform random error terms. Conditional distribution of y_t is the same as the first AR(1) sub model which is ready to use with $\phi_{1,0}$ constant term and $\phi_{1,1}$ autoregressive coefficient and σ_1^2 error variance. On the other hand, if the 1st lagging value of y_t exceeds the r threshold value, second AR(1) period with the parameter $\phi_{2,0}, \phi_{2,1}, \sigma_2^2$ is ready to use.

First-degree SETAR model can be easily expanded to a high-degree with an integer lagging value:

$$y_t = \begin{cases} \phi_{1,0} + \phi_{1,1} y_{t-1} + \ldots + \phi_{1,p_1} y_{t-p_1} + \sigma_1 e_t, \; y_{t-d} \le r \\ \phi_{2,0} + \phi_{2,1} y_{t-1} + \ldots + \phi_{2,p_2} y_{t-p_2} + \sigma_2 e_t, \; y_{t-d} > r \end{cases} \tag{7}$$

Here, it is not required to have p_1 *ve* p_2 autoregressive levels of two sub models the same and d lagging parameter can be bigger than the maximum autoregressive level. However in order to simplify the notation, it can be assumed that $p_1 = p_2 = p \; ve \; 1 \le d \le p$. TAR model defined with the equation (7) is shown as the d lagged $TAR(2, p_1, p_2)$ model.

2.3 ARM Model

Another method used in the prediction of non-linear time series is the ARM model. This model is a more flexible method than the standard regression model.

Here, ARM Model can be defined as follows,

$$y_t = f(x_t) + g(z_t) + u_t \tag{8}$$

for *x* and *z* independent variables. In equation (8); when the u_t errors are independent from each other, "*backfitting algorithm*" developed by Hastie and Tibshirani (1990) is used for the prediction of *f* and *g* functions. However, there is no backfitting algorithm for the auto correlated incorrect additive regression model. Thus, when u_t incorrect (8) model developed by $u_t = \phi u_{1-1} + \varepsilon_t$ first degree autocorrelation is taken into consideration, this model will be equivalent to the model specified as follows:

$$y_t - \phi y_{t-1} = f(x_t) + g(z_t) - \phi f(x_{t-1}) - \phi g(z_{t-1}) + \varepsilon_t \tag{9}$$

Here ε_t errors have independent and same distribution with zero mean and σ^2 variance ($\varepsilon_t \sim NIID(0, \sigma^2)$). In equation (9); if it is written as $v_t = x_{t-1}$, $w_t = z_{t-1}$, $f_1(x) = f(x)$ and $g_1(x) = g(x)$, the said model is defined as follows:

$$y_t - \phi y_{t-1} = f(x_t) + g(z_t) - \phi f_1(v_t) - \phi g_1(w_t) + \varepsilon_t \tag{10}$$

"*Backfitting algorithm*" can be applied to the equation (10) which gives the predictive of f and g non-parametric functions by considering f , g , f_1 and g_1 as four different functions.

2.4 The Hybrid Methodology

Hybrid methodology that has both linear and nonlinear modeling capabilities can be a good strategy for practical use. By combining different models, different aspects of the underlying patterns may be captured. Hence, hybrid model structure of Zhang (2003) can be generalized. In this paper, hybrid models whose components are different parametric and nonparametric regression models are evaluated for time series. The hybridized model which was suggested is defined as follows:

$$y_t = y_t^1 + y_t^2 \tag{11}$$

In this model, y_t is the observation value at time point t, y_t^1 and y_t^2 are linear or nonlinear model components, superscripts denote the row number of the model. Firstly the model with 1-indiced is applied to the observation data and $e_t^1 = y_t - \hat{y}_t^1$, then the others are calculated. Here $\hat{y}_t^1$ is the forecast value of the first model at time point t. If the first model contains m_1 input units, the number of e_t^1 units will be $N - m_1$. If the second model contains m_2 input units, the number of $\hat{y}_t^2$ forecast values will be $N - m_1 - m_2$. In this case, the forecast values appropriate for the second model are calculated as follows:

$$\hat{y}_t^2 = f_2(e_{t-1}, e_{t-2}, \ldots, e_{t-m_2}) \tag{12}$$

where f_2 is the function obtained from the second model. Thus, the forecast for the combined model is defined as follows:

$$\hat{y}_t = \hat{y}_t^1 + \hat{y}_t^2 \qquad (13)$$

The adjusted forecasts are calculated as the sums of the first model and the second model. The hybrid model with good performance is obtained by the model evaluation criteria for the forecasting.

3. Experimental Evaluations

We illustrate our methods with applications to two univariate time series. The first application is monthly export volume index data for the time period 1997-2014. The second is the monthly domestic producer price index data for the time period 2006-2014. We follow Ghaddar and Tong (1981) and make a square-root transformation $y = 2\left(\left(\sqrt{1+N}\right)-1\right)$, where N denotes the raw data series. For two real data sets y values indicated as observations are displayed in Figures 1 and 3, respectively.

In order to evaluate the predictions of a model with observations, the following statistical performance measures, which include the mean square error (MSE) or the root mean square error (RMSE), the Mean absolute error (MAE) and the mean absolute percentage error (MAPE) have been used (Goh and Law, 2002). Forecast evaluation measurements are defined as following way:

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2 \qquad RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(y_t - \hat{y}_t\right)^2}$$

$$MAE = \frac{1}{n}\sum_{t=1}^{n}\left| y_t - \hat{y}_t \right| \qquad MAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{\left|y_t - \hat{y}_t\right|}{\left|y_t\right|}\ (\%100)$$

where y_t is represent the observed values; $\hat{y}_t$ is indicate the forecasted values and $Mean$ is the arithmetic mean value .

A perfect model would have MSE or RMSE, MAE and MAPE $\cong 0.00$; naturally, because of the influence of random errors, there is no such thing as a perfect model in time series modeling. For numerical calculations, R statistical packaged program is used.

3.1 Real Data Example 1

Within the scope of this study; the data of the monthly export volume index numbers, between 1997:12-2014:12 (2010=100) are used. Export volume index data are taken from the web page of TUİK. The data set is divided into two parts for the use in training and forecasting. In the first part, 192 monthly data are taken into account for the period of the January 1997–December 2012 period. These data are used in training stage to construct the models. In the second part, with the help of the models constructed in the first part, the performances of those models are calculated using the 24 monthly test data for the January 2013–December 2014 period. We used a nonparametric regression model called as additive regression model (ARM). Estimation of the ARM in equation (8) is obtained by using backfitting algorithm. There is a selection problem of the smoothing parameter λ included in this method. The parameter λ is selected by using generalized cross validation (GCV). The observed and estimation values from nonparametric ARM and parametric AR2 in equation (4) and SETAR in equation (7) are demonstrated in Figure 1.

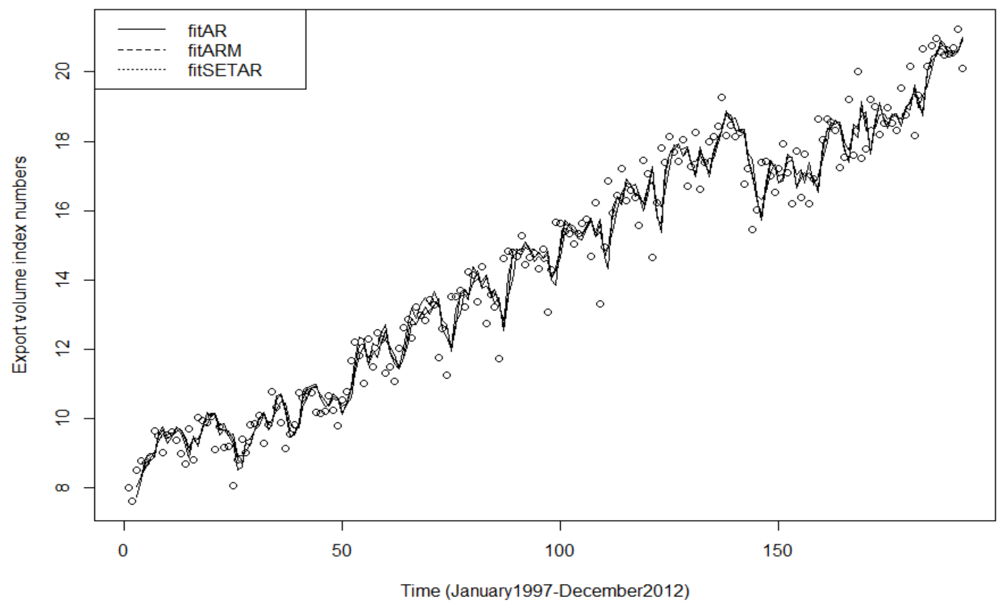

Figure 1: Observed monthly export volume index values and their estimated values obtained by AR2, ARM and SETAR models

In the determining hybrid models whose first component is nonparametric regression, firstly, the nonparametric regression model was applied to the real data set composed of 192 cases, and then, the 192 residuals were obtained for the first data sets. At the next step, to construct the second component of hybrid model, parametric model was applied to the 192 residuals data. Thus, hybrid model is obtained by combining the estimations in these two steps. Similarly, it is obtained 6 different hybrid models such as AR2&AAR, AR2&SETAR, AAR&SETAR, SETAR&AR, SETAR&AAR and AAR&AR models.

Table 1: Performance Values of the Models for Export Volume Index Numbers

Models	MAE	MAPE	MSE	RMSE
AR2	0.7832594	3.799851	0.8460767	0.9198243
SETAR	0.5584736	2.779853	0. 603419	0.7768005
AAR	0.5679738	2.832754	0.6229199	0.7892527
AR2&AAR	0.8611482	4.162941	0.9438735	0.9715315
AR2&SETAR	0.8122842	3.934331	0.9035898	0.9505734
AAR&SETAR	0.5564938	2.767973	0.5908856	0.7686908
SETAR&AR2	0.5619007	2.797629	0.6162787	0.7850342
SETAR&AAR	0.6258759	3.086900	0.6384326	0.7990198
AAR&AR2	0.5680712	2.833124	0.6233602	0.7895316

An evaluation of the models was made depending on the forecasts for the 24 monthly data between 2013–2014 period. For the 24 monthly test data, the performances of these models are carried out and compared by using the MAE, MAPE, MSE and RMSE values. The mentioned these values are given in Table 1. As can be seen from Table 1, according to the export volume index data sets, SETAR model has a good performance among the models such as AR2 and ARM. On the other hand, the hybrid models whose first component is parametric regression, SETAR&AR2 has demonstrated a good performance, whereas the hybrid models whose first component is nonparametric regression, AAR&SETAR model, has demonstrated the best performance. For test data composed of the 24 values, the observed and forecasted values obtained by nine different models are calculated, but they are only given graphically for the three models because of space limitations. The only observed and forecasted values by the AR2, ARM and the hybrid models are given in Figure 2.

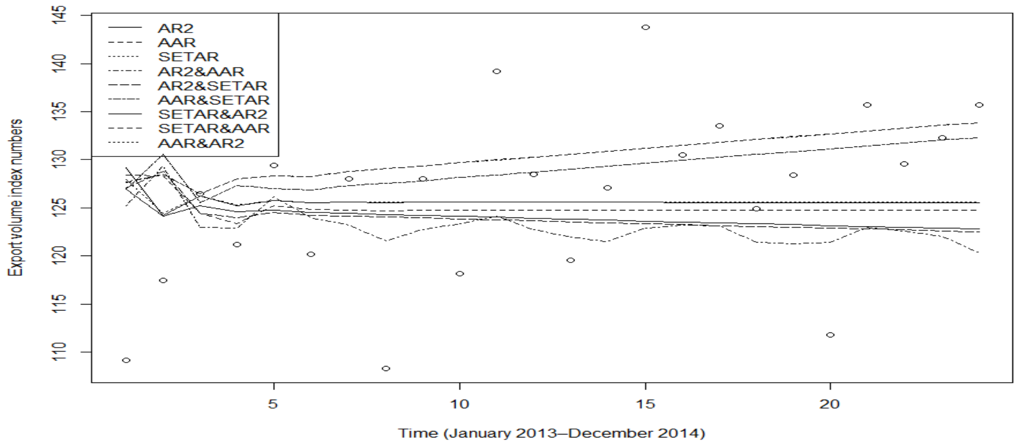

Figure 2: Observed data specified by sign circle and their forecasted values obtained by different parametric, nonparametric and hybrid models for the January 2013–December 2014 period

3.2 Real Data Example 2

The monthly domestic producer price index data sets are used for the second application. As in first data set, this data set is also divided into two parts for the use in training and forecasting. In the first part, 84 monthly data are taken into account for the period of the January 2006–December 2012 period.

These data are used in training stage to construct the models. In the second part, with the help of the models constructed in the first part, the performances of those models are calculated using the 24 monthly test data for the January 2013–December 2014 period. As in the first application, ARM is considered as nonparametric regression, whereas AR2 and SETAR are used parametric regression models. Estimation of the ARM in equation (6) is obtained by using back fitting algorithm. There is a selection problem of the smoothing parameter λ included in this method. The parameter λ is selected by using generalized cross validation (GCV). The observed and estimation values from nonparametric ARM and parametric AR2 in equation (4) and SETAR in equation (7) are demonstrated in Figure 2.In the determining hybrid models whose first component is nonparametric regression, firstly, the nonparametric regression model was applied to the real data set composed of 84 cases, and then, the 84 residuals were obtained for the first data sets. At the next step, to construct the second component of hybrid model, parametric model was applied to the 84 residuals data. Thus, hybrid model is obtained by combining the estimations in these two steps. Similarly, it is obtained 6 different hybrid models such as AR2&AAR, AR2&SETAR, AAR&SETAR, SETAR&AR, SETAR&AAR and AAR&AR models.

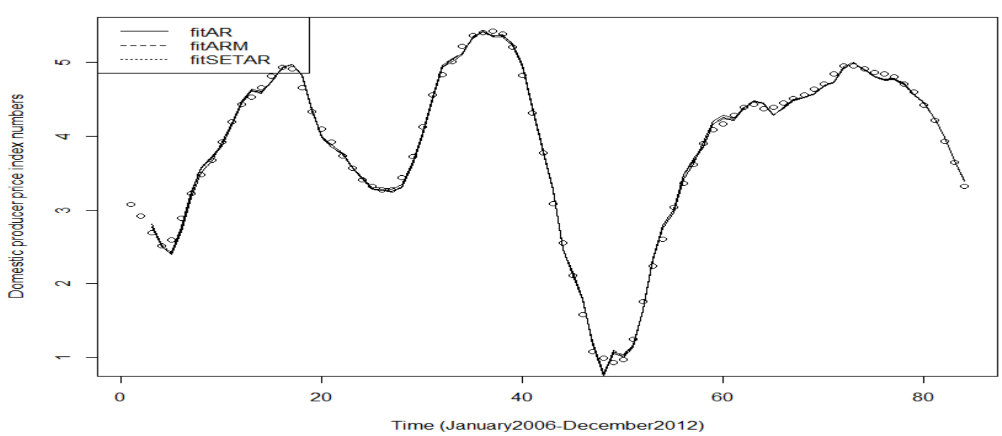

Figure 3: Observed monthly domestic producer price index values and their estimated values obtained by AR2, SETAR and ARM models

Table 2: Performance Values of the Models for the Domestic Producer Price Index Data

Models	MAE	MAPE	MSE	RMSE
AR2	0.09953025	2.745957	0.01632611	0.1277737
AAR	0.09394687	2.667006	0.01269961	0.1126925
SETAR	0.10588890	2.939485	0.01794904	0.1339740
AR2&AAR	0.10027430	2.777185	0.01638269	0.1279949
AR2&SETAR	0.10143740	2.803090	0.01659169	0.1288087
AAR&SETAR	0.09232388	2.617244	0.01234716	0.1111178
SETAR&AR2	0.10482240	2.905108	0.01778102	0.1333455
SETAR&AAR	0.10706090	2.978552	0.01815183	0.1347287
AAR&AR2	0.09375183	2.660544	0.01269344	0.1126652

An evaluation of the models was made depending on the forecasts for the 24 monthly data between 2013–2014 periods. For the 24 monthly test data, the performances of these models are carried out and compared by using the MAE, MAPE, MSE and RMSE values. The mentioned these values are given in Table 2. As can be seen from Table 2, according to the export volume index data sets, AAR model has a good performance among the models such as AR2 and ARM. On the other hand, the hybrid model whose second component is parametric model AAR&AR2 has demonstrated a good performance.

As in application 1, the hybrid models whose first component is nonparametric AAR&SETAR model has demonstrated the best performance for the domestic producer price index data. For test data composed of the 24 values, the observed and forecasted values obtained by nine different models are calculated, but they are only given graphically for the three models because of space limitations. The only observed and forecasted values by the AR2, ARM and the hybrid models are given in Figure 3.

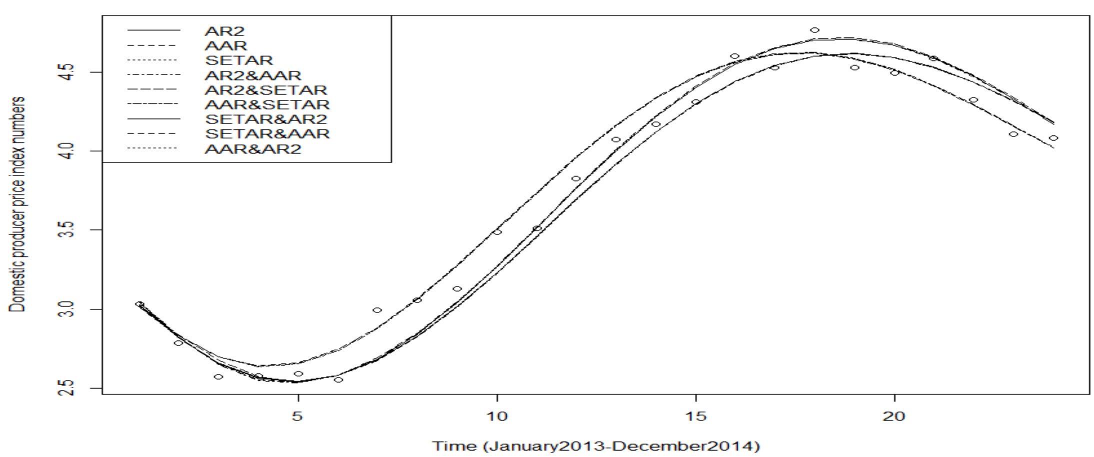

Figure 4: Observed data specified by sign circle, the forecasted values specified by different parametric, nonparametric and hybrid models for January 2013–December 2014 period

4. Result and Evaluation

It is known that hybrid models indicate very good performance in time series forecasting problems. Zhang (2003) reported that hybrid models where a component is linear and the other is nonlinear have demonstrated a good performance in time series forecasting. Then, it is discussed in Aslanargun (2007) that using hybrid models, whose components are nonlinear, more effective. As specified in the introduction; AR, SETAR and ARM models are commonly used in the prediction of non-linear economical time series. Non-parametric regression techniques try to specify the relationship between the variables without considering any functional status of the model. These techniques try to determine the functional type of the model directly instead of calculating the model parameters.

In the study; the export volume index numbers and domestic producer price index and rate of change for Turkey are predicted by the AR, SETAR and ARM models and hybrid models as the performance indicators of such models, MAE, MAPE, MSE and RMSE values are calculated. When Table 1 and Table 2 are examined; the lowest MAE, MAPE, MSE and RMSE are obtained for the AAR&SETAR model.

These values are found as 6.463981, 5.371849, 79.64720 and 8.924528 for the export volume index numbers and 0.09232388, 2.617244, 0.01234716 and 0.1111178 domestic producer price index data sets, respectively. In this case, we observed that hybrid models make more better forecasting in the time series forecasting problems based on the export volume index numbers and domestic producer price index data sets for Turkey. As a result, our opinion is that, using hybrid models, whose components are nonparametric regression can be useful in time series forecasting problems included seasonality and trend.

References

Altman, N. S. (1990). Kernel Smoothing of Data with Correlated Errors. *Journal of the American Statistical Association*, 85: 749-759.

Aslanargun, A., Mammadov, M., Yazıcı, B., and Yolacan, S. (2007). Comparison of ARIMA, neural networks and hybrid models in time series: tourist arrival forecasting, Journal of Statistical Computation and Simulation, 77, pp. 29-53.

Box, G. E. P., Jenkins, G. M. and Reinsel, G. C. (1994). *Time Series Analysis*. New Jersey: Prentice Hall.

Chan, K. S. and Tong, H. (2001). *Chaos: A Statistical Perspective*. Springer Verlag.

Dagum, E. B. and Giannerini, S. (2006). A Critical İnvestigation On Detrending Procedures For Non Linear Processes. *Journal of Macroeconomics*, Elsevier, Vol. 28(1), 175-191.

Engle, R., Granger, W., Rice, J. and Weiss, A. (1986). Semi Parametric Esitmates of The Relation Between Weather and Electricity Sales. *J. Am. Statist. Ass*., 81, 310-320.

Eilers, P. H. C. and Marx, B. D. (1996). Flexible Smoothing with B-Splines and Penalties (with discussion). *Statist. Sci*., 89, 89–121.

Eva, F., Vicente, N.A., Juan, R.P.(2000). Semi-parametric approaches to signal extraction problems in economic time series, Computational Statistics & Data Analysis, 33(2000), pp. 315-333.

Fan, J. and Gijbels, I. (1996). *Local Polynomial Modelling and its Applications*. Chapman and Hall: London.

Franses, P. H. and Dijk, V. D. (2000). *Nonlinear Time Series Models in Empirical Finance*. Cambridge University Press, Cambridge.

Ghaddar, D. K. and Tong, H. (1981) Data transformation and self-exciting threshold autoregression. Journal of the Royal Statistical Society, C30, 238±248.

Goh, C. and Law, R. (2002). Modelling and forecasting tourism demand for arrivals with stochastic nonstationary seasonality and intervention. Tourism Management. 23 (5), 499-510

Granger, C. W. J. and Terasvirta, T. (1993). *Modelling Nonlinear Economic Relationships*. Oxford University Press, Oxford.

Green, P. J. and Silverman, B. W. (1994) . *Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach*. Chapman and Hall, London.

Hart, J. D. (1991). Kernel Regression Estimation with Time Series Errors. *Journal of the Royal Statistical Society*, B 53: 173-187.

Hart, J. D. (1994). Automated Kernel Smoothing of Dependent Data By Using Time Series Cross-Validation. *Journal of the Royal Statistical Society*, B 56: 529-542.

Harvey, A.C. and S.J. Koopman (1993) . Short Term Forecasting of Periodic Time Series Using Time-Varying Splines. *J. Amer. Statist. Assoc*., to appear.

Hastie, T. J. and Tibshirani, R. J. (1990). *Generalized Additive Models*. New York: Chapman and Hall.

Henry, Ó., Olekalns, N. and Summers, P. (2001). Exchange Rate İnstability, A Threshold Autoregressive Approach. *Economic Record*, 237, 160-166.

Hurvich, C. M. and Zeger, S. L. (1990). A Frequency Domain Selection Criterion for Regression with Autocorrelated Errors. *Journal of the American Statistical Association*, 85: 705-714.

Pfann, G. A., Schotman, P. C. and Tschernig, R. (1996). Nonlinear Interest Rate Dynamics and Implications for The Term Structure. *Journal of Econometrics*, 74, 149–176.

Smith, M., Wong, C. M. and Kohn, R. (1998) "Additive Nonparametric Regression with Autocorrelated Errors" *Journal of the Royal Statistical Society*: Series B: Statistical Methodology, Vol. 60, 2, 311-331.

Tseng, F.M., Yu, H.C. and Tzeng, G.H.(2002). Combining neural network model with seasonal time series ARIMA model. Technological Forecasting & Social Change, 69(2002), pp. 71–87.

Tong, H. (1978). *On a Thresold Model. In Pattern Reconition and Signal Processing*. (Edited by C. H. Chen), Sijthoff and Noordhoff, Amsterdam, 101-41.

Tong, H. and Lim, K. S. (1980). Threshold Autoregression, Limit Cycles and Cyclical Data (with discussion). *Journal of the Royal Statistical Society*, Ser. B, 42, 245-292.

Tong, H. (1983). *Threshold Models in Nonlinear Time Series Analysis*. Lecture Notes in Statistics, Springer-Verlag.

Tong, H. (1990). *Nonlinear Time Series: A Dynamical System Approach*. Oxford University Press, Oxford.

Tong, H. (2007). Birth of The Threshold Time Series Model. *Statist. Sinica*, 17, 8–14.

TUİK (2012). *Turkish Statistics Institute*. http://www.tuik.gov.tr/VeriBilgi.do?alt_id=13

Wood, S. N. (2000). Modelling and Smoothing Parameter Estimation with Multiple Quadratic Penalties. *J.R. Statist. Soc*. B 62, 413-428.

Zhang, G.P. (2003). Time-series forecasting using a hybrid ARIMA and neural network model. Neurocomputing, 50, pp. 159–175.

Zivot, E. (2005). Nonlinear Time Series Models. http://faculty.washington.edu/ezivot/econ584/notes/nonlinear.pdf