

Article

Year Ahead Demand Forecast of City Natural Gas Using Seasonal Time Series Methods

Mustafa Akpınar * and Nejat Yumusak

Computer Engineering Department, Faculty of Computer and Information Sciences, University of Sakarya, 2nd Ring Street, Esentepe Campus, Serdivan, Sakarya 54187, Turkey; nyumusak@sakarya.edu.tr

* Correspondence: akpinar@sakarya.edu.tr; Tel.: +90-264-295-3238

Academic Editor: Francisco Martínez-Álvarez

Received: 13 May 2016; Accepted: 6 September 2016; Published: 8 September 2016

Abstract: Consumption of natural gas, a major clean energy source, increases as energy demand increases. We studied specifically the Turkish natural gas market. Turkey's natural gas consumption increased as well in parallel with the world's over the last decade. This consumption growth in Turkey has led to the formation of a market structure for the natural gas industry. This significant increase requires additional investments since a rise in consumption capacity is expected. One of the reasons for the consumption increase is the user-based natural gas consumption influence. This effect yields imbalances in demand forecasts and if the error rates are out of bounds, penalties may occur. In this paper, three univariate statistical methods, which have not been previously investigated for mid-term year-ahead monthly natural gas forecasting, are used to forecast natural gas demand in Turkey's Sakarya province. Residential and low-consumption commercial data is used, which may contain seasonality. The goal of this paper is minimizing more or less gas tractions on mid-term consumption while improving the accuracy of demand forecasting. In forecasting models, seasonality and single variable impacts reinforce forecasts. This paper studies time series decomposition, Holt-Winters exponential smoothing and autoregressive integrated moving average (ARIMA) methods. Here, 2011–2014 monthly data were prepared and divided into two series. The first series is 2011–2013 monthly data used for finding seasonal effects and model requirements. The second series is 2014 monthly data used for forecasting. For the ARIMA method, a stationary series was prepared and transformation process prior to forecasting was done. Forecasting results confirmed that as the computation complexity of the model increases, forecasting accuracy increases with lower error rates. Also, forecasting errors and the coefficients of determination values give more consistent results. Consequently, when there is only consumption data in hand, all methods provide satisfying results and the differences between each method is very low. If a statistical software tool is not used, time series decomposition, the most primitive method, or Winters exponential smoothing requiring little mathematical knowledge for natural gas demand forecasting can be used with spreadsheet software. A statistical software tool containing ARIMA will obtain the best results.

Keywords: demand forecasting; natural gas; univariate methods; time series decomposition; Holt-Winters model; autoregressive integrated moving average (ARIMA); seasonal ARIMA

1. Introduction

Natural gas has the fastest growing consumption rates among clean energy resources in the world. It is considered a common resource and is used in different sectors such as heating, electricity generation, transportation, cooking, and cooling. Large investments are needed for forwarding, transporting and using natural gas. Some factors that affect these investments are whether the investment matches the consumption amount or how much air pollution occurs. Contracts have been made between various countries for natural gas supply whereby usage started at the beginning

of the 1990s and demand has rapidly increased since then [1,2]. These contracts are mainly take-or-pay type contracts, focusing on estimating long-term natural gas consumption. One of the characteristics of take-or-pay contracts is that for lower consumption than estimated the price of estimated consumption must still be paid. However, for higher consumption than estimated, reducing the gas supply by slightly closing the valve or paying an extra price per unit m^3 occurs. In order to avoid these situations and reduce economic and social losses, demand forecasting with some minimum acceptable error should be used. Country governments should have preliminary information on regional consumption levels for demand forecasting. The main objective of gathering the preliminary information is that regional consumers use natural gas for various reasons and together they form a country's consumption capacity. For instance, large factories use natural gas for electricity generation and manufacturing. These factories consume similar amounts both in winter and summer seasons, so they mostly show stationary behavior. Likewise, as long as there is no failure, daily high consumption of electricity generation plants depending on electricity generation staying at the same levels. Besides high consumption factories, organizations and plants, there are low consuming corporations and residential consumers as well. Their consumption patterns are mainly affected by seasonal variations. For instance, consumption levels decrease in the summer season and noticeably increase in the winter period. Even if each region in the country has different consumption levels, low consuming consumers always have a critical amount of natural gas consumption at the national level. Since natural gas unit costs of these consumers need extra investment costs, their unit m^3 charges are higher compared to the high consumption sectors. Seasonal influences (change in consumption behavior) and high unit costs have made these consumers more important. Our case study is based on a city's consumption in Turkey and this situation is also applied to cities in Turkey. Seasonally affected consumption behaviors of city consumers have an impact on the natural gas market in Turkey.

The natural gas market in Turkey is shown in Figure 1. The companies placed on dotted lines perform an annual forecast based on regulations [3]. The producers are generally outside Turkey [4]. Import/export and wholesale companies could import natural gas to Turkey through pipelines or as liquid natural gas (LNG). At each level except the bottom, companies report their year-ahead forecasts in a hierarchical manner to the companies that have a contract. Finally, import or wholesale companies make a final estimation using bottom-up collected data. Each month, these forecasts are checked and if the mean absolute percent error is higher than 10%, penalties occur [3]. The natural gas market is inspected by the Energy Market Regulatory Authority (EMRA, known by its Turkish acronym EPDK) and controlled by the Petroleum Pipeline Corporation (PPC, known as BOTAS). According to the EMRA report, in 2014, 49.262 billion m^3 natural gas was imported by nine long-term and two spot (LNG) import licensed entities [4]. In the same report, it is stated that 7.281 billion m^3 of LNG was imported, which equals 14.78% of total imports. At the national level, the household ratio of consumption is nearly 20% of total consumption [4]. This consumption amount is noticeably high, affecting penalties, and it is forecasted based on the bottom part of the market, from the residential/and small commercial end users who are subscribers of the City Distribution Company. As mentioned in the report, the sum of household and low consumption consumers comprises nearly 26% of total consumption at the national level [4]. Therefore, the main objective of this application paper is to show the possibility of monthly forecasting natural gas demand for household and low consumption consumers by applying well-known univariate methods in the literature at the city level. Thus, it is expected that a low error rate and local error-free prediction results will be obtained.

The rest of the paper is organized as follows: related studies are presented in Section 2. The data and theoretical description of methods are described in Section 3. Section 4 gives detailed information about modeling, definitions, scenario analysis and error benchmarks. Section 5 presents pre-forecasting steps and Section 6 shows the forecasting results and discussion. The key findings and next studies are given at the end of the paper as conclusions in Section 7.

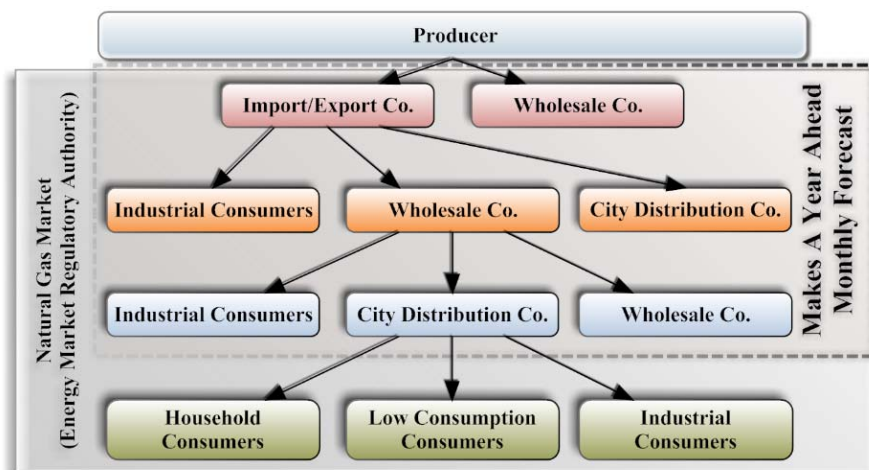


Figure 1. Turkey natural gas market and users.

2. Related Work

Time series forecasting is an important area of forecasting in which past observations of the same variable area are collected and analyzed to develop a model describing the underlying relationship [5]. Natural gas consumption predictions are being made with several approaches in different fields. These studies can be investigated as daily, monthly, national level, regional level, residential area, industrial area, use of an independent variable and no use of an independent variable.

In the first group, publications can be divided according to the use of timeframes which apply the time series method in daily periods [6–19] and monthly periods [20–24]. In the second group, publications can be grouped as regional [8–12,15,18,19,21–24] or national [6,7,13,14,20,24–31] consumptions are investigated. In the third group, papers are investigated by consumer types. This group includes household consumers [6–12], commercial consumers [11,13,25] and consumers where all consumption sectors are included [14–18,24–31]. In the fourth group, where studies are categorized by data used, papers are divided with respect to the use of only consumption data using univariate approaches [28–31] or independent variable [6–18,20,22–27] included studies. Investigation of these studies showed that, mostly independent variable included regional-based natural gas consumption prediction is done. The summary of this research is published by Soldo [32].

Univariate techniques have a broad usage area in time series forecasting. Ediger et al. used autoregressive moving average (ARIMA), seasonal ARIMA (SARIMA) and comparative regression techniques to forecast the production of fossil fuel sources in Turkey, which include natural gas [33]. They made annual forecasts from 2004 to 2038 and used different regression types such as linear, logarithmic, inverse, quadratic, cubic, compound, power, growth, exponential, and logistic. They concluded that ARIMA is a suitable technique for natural gas consumption. Gutiérrez et al. used the Gompertz-type innovation diffusion process as a stochastic growth model to forecast annual natural gas in Spain from 1973 to 1997 [29]. They compared the results between 1998 and 2000 with stochastic logistic innovation modelling and the Gompertz model was found to be more suitable. Ma and Wu studied China's annual natural gas consumption and production prediction with the Grey model. They used data from 1990 to 2003 and generated forecasts for 2004 to 2007 [30]. In a comparison between the Grey model (with one variable and rank 1 differential equation—GM(1,1)) and Grey-Markov model, the Grey-Markov gave better results. Xie and Li also used the Grey model to predict China's annual natural gas forecasting. Unlike Ma and Wu, they used a genetic algorithm for optimizing the GM(1,1) model [31]. They used data from 1996 to 2002 and predict for the forecast range of 2003 to 2005. Here, genetic optimization performed better results. In the literature, only Liu and Lin studied forecasting monthly natural gas [34]. Their research is on a national level,

and they made predictions on monthly and quarterly periods. They formed ARIMAX (ARIMA with exogenous) models by adding temperature and price into ARIMA models.

Univariate statistical forecasting could also be applied in other sections of the energy sector such as electrical, water, solar, wind etc. Yalcintas et al. studied water management through demand and supply forecasting for Istanbul city [35]. They applied ARIMA to forecast annual demand 2015–2018 by using data from 2006 to 2014 and suggested that sustainable management of water can be achieved by reducing residential water use through the use of water efficient technologies. Gelažanskas and Gamagė used time series seasonal decomposition, exponential smoothing and SARIMA methods for predicting hot water demand [36]. They found that the most significant part in the accuracy of forecasting is the seasonal decomposition method of the time series. Prema and Rao forecasted wind speed using time series decomposition, exponential smoothing and back propagation neural networks [37]. They observed that decomposition of time series and ARIMA methods gave more accurate results. The ARIMA method has been frequently used for daily and hourly load prediction of electricity consumption. Research on time series applying electricity prediction is presented as a survey [38]. It is observed that ARIMA is one of the most used linear prediction techniques. For instance, Wang et al. studied electricity price estimation with Winters' exponential smoothing and SARIMA methods [39].

3. Our Contribution

This application paper studies forecasting of natural gas demand. As mentioned above, in previous studies time series decomposition, seasonal exponential smoothing (Holt-Winters exponential smoothing), ARIMA and SARIMA methods are frequently used ones used to study predictions of electricity load, price, hot water demand, and wind speed. To the best of our knowledge, ARIMAX, includes exogenous variables, has been used for natural gas monthly predictions [34]. In our research, three methods are applied for the monthly demand prediction of natural gas, which is a sub-branch of the energy sector. These methods have not been used previously together on natural gas and monthly predictions. The existence of seasonality is the power of these methods as they do not contain any information with other variables except its own past data.

4. Methodologies

This paper examines time series decomposition, Holt-Winters exponential smoothing and ARIMA methods. The common feature of these techniques is they do not need extra data besides their own data. Other than this, each model has its own characteristics. The models consisting of seasonality and trends can also be used in natural gas forecasting. These methods are briefly introduced in this section.

4.1. The Data

Natural gas is distributed with pipelines to the end-users in Turkey. Pipelines are connected to the cities with reducing and measuring stations (RMS). In these stations, the natural gas pressure decreases and the volume of natural gas is calculated. RMSs are divided into three categories. The first is RMS-A, which connects pipelines from national distribution to regional distribution. Those connections are steel lines because of the pressure (40–75 bar is reduced to 12–25 bar and the consumption range in an hour is 10,000 to 300,000 m³). RMS-A type stations are administrated by city natural gas distribution companies. The other two kinds of RMS are B and C, which reduce pressure from 6–25 bar to 4 bar and from 1–4 bar to 0.3 bar. RMS-B stations are also steel lines while RMS-C stations are polyethylene lines. All RMS' consumptions are measured and calculated hourly. In this study, the natural gas consumption data is collected from the Natural Gas Distribution Company of Sakarya province. One hour resolution data consumption is first converted to daily resolution and then monthly data for all RMSs. 90% of industrial subscribers have an RMS and telemetry system for remote measurement. Monthly converted telemetry consumption of RMS-B and RMS-C are subtracted from RMS-A consumption. The remaining 10% of industrial subscribers' invoices are billed the first day of the month like

the other 90%. Also those 10% of subscribers' consumptions are subtracted from the remaining RMS-A consumptions so that the monthly consumption of households and low-consuming consumers are found.

4.2. Forecasting Models

This section gives brief information about time series decomposition, Winter exponential smoothing, ARIMA and SARIMA models. The time series has four different elements, which are the seasonal component (S), trend component (T), cycle component (C) and irregular component (E) [37,40,41]. The decomposition of time series (D) has two approaches, namely the additive and multiplicative model. The multiplicative decomposition model multiplies the components one by another, whereas in the additive decomposition model component predictor variables are added to one another [37,40–43]. In the model, trend component denotes long-term tendencies while cycle component indicates longer periodic seasonal movements. The seasonal component represents short periodic oscillations and the irregular component represents not expected or could not be predicted values. Forecasting two periods may contain a cycle effect [43]. The main reason for this effect is the long term and changes during the process. The advantage of applying the decomposition method is that the method is easy to understand and applicable. Compared to other methods, it needs at least three-term data to process.

In the exponential smoothing method, the effect of last actual data is related to its weight. Besides the last actual data, older data has an influence on the prediction as well. In fact, the prediction has an impact on all data in descending order. In order to obtain the exponential smoothing equation, the previous prediction is added to the model with the alpha coefficient recursively and a new weighted equation is formed [37,40–46].

The Holt-Winters exponential smoothing (W) method also has a similar process chain with a basic exponential smoothing method. A key point here is having a coefficient, which is suitable to make seasonal predictions and an equation. Holt-Winter computes the prediction by using the level, trend, and seasonal equations and putting them in appropriate positions in the forecast equation. Weights of level, trend and seasonal components are α , β and γ , respectively. The Holt-Winter method is able to present trend, seasonality and randomness in an effective way by the exponential smoothing process [41]. The only disadvantage of this approach is taking a long time duration to determine three parameters.

ARIMA is a popular technique used for time series analysis and prediction [47–53]. This method has three components. These components are the autoregressive part (AR), which shows a relationship between previous data during modeling; the moving average part (MA) and integration part (I), which is used to make the series stationary. During the ARIMA process, the series should not have a missing value and it should be stationary [40–42,47,50–53]. ARIMA models can be defined as an ARIMA(p,d,q) notation. In the notation, p is the AR parameter (ϕ_p), d is the order of the backward-difference value and q is the MA parameter (θ_q).

If seasonality is contained in the ARIMA model, the seasonal ARIMA (SARIMA) method is used. In this method, the model can be initiated the same as ARIMA, and related to the duration of the season AR, I and MA parameters are applied to the model [41,42]. Usually, the model is represented as ARIMA(p,d,q)T(P,D,Q)^s notation. In the first part of the notation, the p,d,q values are the same parameters modeled in ARIMA method. T is a transformation parameter and if there is no transformation operation, the coefficient is zero. If the transformation operation is logarithmic, the parameter is assigned to 1. Except this assignment, whatever number is written, values are exponentiated to that number. The second P,D,Q value is the seasonality part of the ARIMA model. The value of the power “s” indicates the duration of the season.

Box-Jenkins Approach

The ARIMA Box-Jenkins approach consists of four stages [41], which are model identification, parameter estimation, model diagnostics and forecast verification, respectively. During the model identification phase, series graphics, autocorrelation and partial autocorrelation functions (ACF-PACF) are investigated and stationary tests are completed. If data values are high in a series, logarithm, natural logarithm etc. transformations are done for normalization. As a result of this step, data will be converted into a more predictable version. The next phase in the ARIMA process is stationary tests. Various unit root tests are applied in the literature [54–56]; however, more frequently used ones are two tests, namely the Augmented Dickey-Fuller (ADF) [55] and Phillips-Perron (PP) [56] tests. In the ADF test, the model can be formed by whether the coefficient and trend exist in the equation or not exist. On the other hand, the PP test is applied on large datasets. Another difference between the PP and ADF test is that since the lag value does not exist in PP, there is no reduction in the degree of freedom. In the Dickey Fuller test, in order to remove the autocorrelation problem, lag lengths of dependent variables are added. This operation yields a decrease in the degree of freedom as well. However in the Philips Perron test, instead of adding an additional lag, non-parametric corrections are done. Therefore, the degree of freedom is not lost. Although it seems like an advantage that the PP test gives better results in large samples compared to the ADF test, the major disadvantage is the PP gives worse results in small samples. After the stationary test, the ARIMA method is applied to a stationary series. In this step, parameters are determined. In the next step, model diagnostics are completed.

4.3. Performance Evolution and Error Estimations

In the diagnostics phase, the success of the ARIMA method is related to some particular components such as the adjusted coefficient of determination (\hat{R}), mean absolute percent error (MAPE), Akaike information criterion (AIC), Bayesian information criterion (BIC) and whether residuals are white noise or not [41,42]. These terms are defined as diagnosis tools and used in the determination process [40–42,57].

Other success criterion is MAPE. MAPE is calculated by subtracting the actual value from the forecasting value and then dividing by the actual value. The absolute value of the division is multiplied by 100 and divided by the number of observations. The absolute value is critical in the equation to ensure that negative and positive percent errors should not eliminate each other.

AIC and BIC are other information criteria of this study [40,41,50]. Relevant to model complexity, they significantly help to select more accurate and low error models in out-of-sample forecasting.

The errors for each method are measured by the coefficient of determination and mean absolute percent error. By examining these two criteria, the estimation satisfaction level on the overall general series is determined.

5. Case Study

The case study models low consuming commercial and residential consumers' natural gas consumption by seasonal univariate statistical methods (time series decomposition, Holt-Winters exponential smoothing, ARIMA, SARIMA). The natural gas consumption data is collected from the city of Sakarya, Turkey. Industrial hourly consumption data (102 users) was summarized as monthly and another 12 industrial subscribers' consumption was prepared from manually billed invoices. The consumption of RMS-A is prepared between the years 2011–2014 (48 month-long) and this data was later divided into two parts. In the first part, 2011–2013 (36 month-long) data is used for monthly forecasts, and in the second part, 2014 (12 month-long) data. In the next section, the evaluation and comparison of results will be presented for all series and for 2014 separately. The error and estimation accuracy will be also investigated with MAPE and R^2 .

Hereafter, time series decomposition, Holt-Winters exponential smoothing and autoregressive integrated moving average are referred to as “D”, “W” and “ARIMA”, respectively. Each applied method has its own sub-techniques. These three methods are preferred as they all include the seasonality effect.

The time series decomposition method can be additive or multiplicative as shown in Table 1. If the values increase or decrease proportionally over time, the “multiplicative model” gives better results. However, if values increase or decrease additively over time the “additive model” gives better results. This study applies both additive and multiplicative models in order to show which technique has more influence on natural gas prediction. The decomposition process of this method can contain both trend and seasonality or only the seasonality estimation method. Consumptions that increase during winter and decrease during summer indicate the seasonality effect. Therefore, all decomposition methods contain seasonality. Besides seasonality, this research investigates both trends and seasonality included situations to determine the trend effect on consumption. Thus, two situations (additive/multiplicative or trend/seasonal) can generate up to four different estimations. These four techniques can be easily computed with simple mathematical operations on spreadsheet software.

Table 1. Time series decomposition model abbreviations.

Method D	Components	Abbreviation
Additive	Seasonal	D-AS
	Trend Seasonal	D-ATS
Multiplicative	Seasonal	D-MS
	Trend Seasonal	D-MTS

Holt-Winters Exponential Smoothing is the second method used. This approach can be additive or multiplicative as well (Table 2). The determination of α , β and γ parameters is important in this method. Before the determination step, default values for α , β and γ are taken as 0.2 as stated in [51]. However, these values will not give the best results. In order to obtain the best result, they should be optimized. Various methods are used in convergence. In this work, ordinary least squares (OLS) is applied as the convergence technique. The convergence value is 10^{-7} and iteration number is 500,000. In abbreviation, coefficients are divided by “c” char. The left side of “c” is the Holt-Winters method such as “A” for additive and “M” for multiplicative. The right side of “c” is parameters. If default values are taken, it is 0.2 or else it is written as “Opt” in generally. This technique can be implemented on spreadsheet software with programming experience.

Table 2. Holt-Winters exponential smoothing model abbreviations.

Method W	Parameters	Abbreviation
Additive	$\alpha, \beta, \gamma = 0.2$	W-Ac0.2
	α, β, γ optimized	W-AcOpt
Multiplicative	$\alpha, \beta, \gamma = 0.2$	W-Mc0.2
	α, β, γ optimized	W-McOpt

Other forecasting techniques used in the study are ARIMA and SARIMA methods. The stationarity of the series is investigated by these methods. In this respect, necessary conversion operations are done and the series is stationarized. In order to stationarize the series, differentiation is applied to the ARIMA method. On this new stationary series, particular parameters are determined such as AR, MA, seasonal AR (SAR) and seasonal MA (SMA). Data generated by taking the difference can be described as:

- Secondary difference; $\Delta^2 \log(\text{Consumption})$; $I(2)1$ or $\text{ARIMA}(0,2,0)1$.
- Primary seasonal difference; $\Delta_{12} \log(\text{Consumption})$; $I(0)1(1)$ or $\text{ARIMA}(0,0,0)1(0,1,0)^{12}$.
- Primary difference and primary seasonal difference; $\Delta_{12,1} \log(\text{Consumption})$; $I(1)1(1)$ or $\text{ARIMA}(0,1,0)1(0,1,0)^{12}$.

Equations are formed containing these parameters and later estimations are done.

Natural gas consumption values are related to weather conditions, the number of consumers and calendar effects. Eventually, the consumption is formed by these facts. The main objective of this paper is to show the possibility of monthly forecasting natural gas demand on the city level by performing well-known univariate methods in the literature. Thus, the low error rate and local error-free prediction results will be obtained.

6. Pre-Forecast

Natural gas consumption changes depend on weather conditions. Seasonal impact mainly causes the weather change. Thus, natural gas consumption is influenced by seasonal facts (Figure 2). In D and W methods, seasonality is mandatory in a series. In this way, forecasting is more accurate with the help of seasonal effects. On the contrary, ARIMA and SARIMA methods require stationarity in a series. In other words, seasonality should not exist in a series. Since stationarity is an important fact, preprocesses need to be accomplished. The first step is setting the T value to 1. This process takes the logarithm of a series. There is no difference between the series and logarithm. Only the range of the series is narrowed.

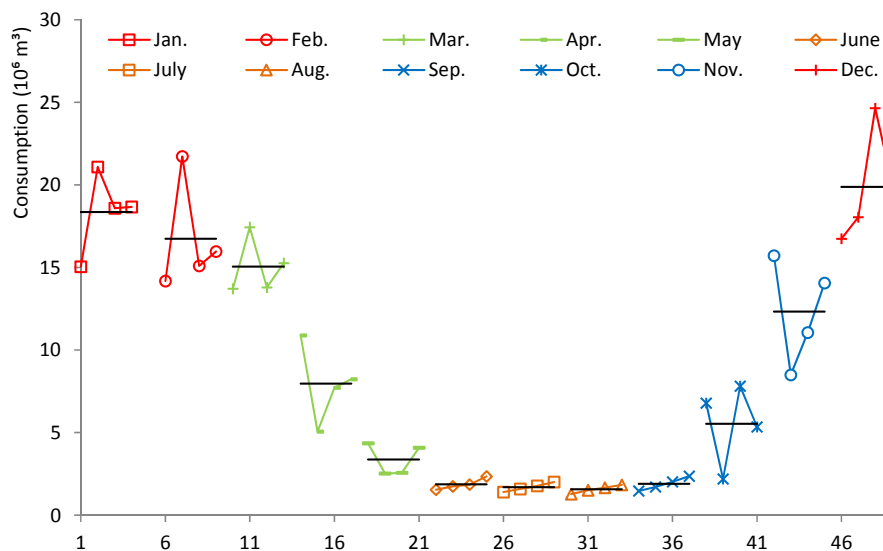


Figure 2. Monthly consumption subplot by years.

In order to stationarize the series, differentiation processes should be completed. The differentiation operation is represented with the “ Δ ” operant. The difference between the series itself and the previous value of it is called the first difference, and it is presented singly. When the differentiation is applied again to the differentiated series, the secondary difference is generated and it is denoted as Δ^2 . The seasonality impact takes place on 12-month data in this study. Therefore, the seasonal difference of a series is shown as Δ_{12} . Both the seasonality and primary difference included series is shown as a $\Delta_{12,1}$. This representation expresses that first seasonal difference, and the next primary difference is performed. In Figure 3, consumption is presented on the left axis and logarithmic values are presented on the right axis of the graph. Results show that consumption series and logarithmic primary difference series are not stationarized. Other differentiation processes applied to the logarithmic series have stationarity.

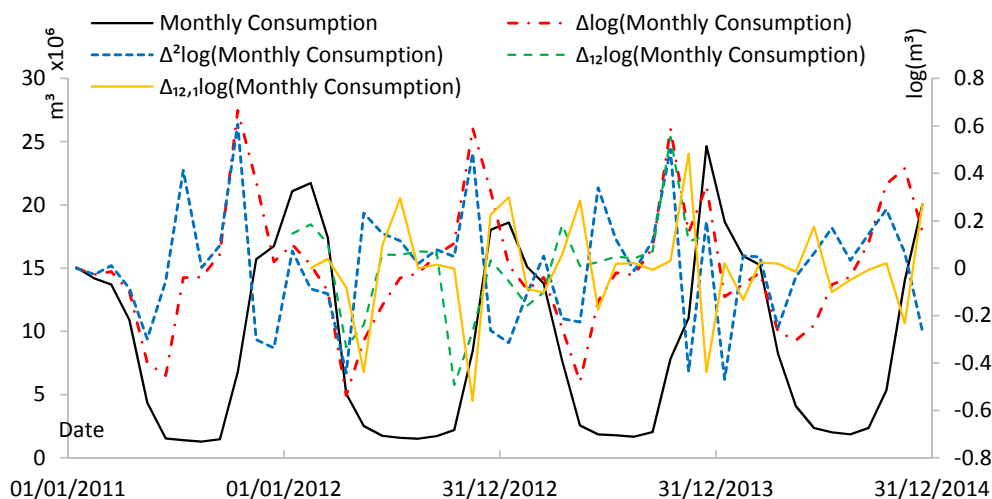


Figure 3. Consumption with integrated series.

Descriptive statistics in Table 3 indicate that the log operation exposes consumption values between 6.108 and 7.392 on an average of 6.757 and standard deviation of 0.443. The mean of the first differentiated series is around zero and its standard deviation reduces by half according to the $\log(\text{consumption})$ series. The second differentiated, seasonal differentiated and both seasonal and non-seasonal differentiated log series have a zero mean and reduced their standard deviations.

Table 3. Descriptive statistics of monthly consumptions.

Variable (m^3)	Obs	Minimum	Maximum	Mean	Std. Dev.
(Consumption)	48	1,281,594	24,651,621	8,854,694	7,203,095
$\log(\text{Consumption})$	48	6.108	7.392	6.757	0.443
$\Delta\log(\text{Consumption})$	48	−0.537	0.665	0.003	0.263
$\Delta^2\log(\text{Consumption})$	48	−0.469	0.606	0.003	0.243
$\Delta_{12}\log(\text{Consumption})$	36	−0.492	0.554	0.019	0.176
$\Delta_{12,1}\log(\text{Consumption})$	36	−0.561	0.483	−0.007	0.210

Even though stationarity can be seen visually (Figure 3), the stationarity of a series is examined by ADF and PP tests [54,55,58]. Table 4 shows probability values of the prepared series on ADF and PP tests. The tests basically investigate the existence of the unit root. This existence proves the series is not stationary. According to this, the hypothesis is prepared. Hypothesis H_0 defines that there is a unit root in a series while the alternative hypothesis shows there is not a unit root, meaning the series is stationary. In this computation, the significance degree of the p probability value is taken as 0.05. If the p value is less than 0.05, then the series is called stationary. Three situations of unit root tests are represented in the table, which are no constant and no trend, only constant, constant and trend. For these three situations, unit root tests are performed. “C” value in the table represents a constant in the equation, while “T” represents the trend. Regarding the outcome of the tests, ADF and PP tests show that all series are stationary except the raw consumption series.

Table 4. Stationary test for consumption series.

p -Value for Stationary	ADF Test			PP Test		
Variable (m^3)	None	C	C + T	None	C	C + T
Consumption	0.715	0.000	0.001	0.077	0.062	0.22
$\Delta\log(\text{Consumption})$	0.000	0.000	0.000	0.001	0.012	0.048
$\Delta^2\log(\text{Consumption})$	0.000	0.000	0.000	0.000	0.000	0.000
$\Delta_{12}\log(\text{Consumption})$	0.000	0.002	0.006	0.000	0.002	0.008
$\Delta_{12,1}\log(\text{Consumption})$	0.079	0.453	0.644	0.000	0.000	0.000

Another way used to determine the stationarity of a series is analyzing ACF and PACF graphs. The ACF graph is demonstrated as an autocorrelogram and PACF is demonstrated as a partial autocorrelogram (Figure 4). Depending on lags, correlograms vary between -1 and 1 , and the significance degree of the relationship between the variable and its history is shown with a dotted line. The area above the dotted line is accepted as a significant relation [41,42]. When ACF and PACF are analysed, it is seen that seasonal patterns of a 12-month series of Consumption and $\Delta\log(\text{Consumption})$ consumptions are noteworthy in ACF graphs. Although ADF and PP test results show stationarity, because of the seasonal patterns of these series, it is not appropriate to use them in forecasting. The $\Delta^2\log(\text{Consumption})$, $\Delta_{12}\log(\text{Consumption})$ and $\Delta_{12,1}\log(\text{Consumption})$ series do not have seasonal patterns. Thus, the ARIMA method can be easily applied to the generated series. It is observed that the autocorrelograms and partial autocorrelograms of these three series have a relation in the 12th month. The seasonality of the series can be seen by this way. The AR, MA coefficients in the non-seasonal and the seasonal parts are used between zero to three for finding the optimum forecast results. Different ARIMA models (256) were prepared and the forecast results are examined.

The equivalents of differentiation operations of ARIMA and SARIMA methods are called “I” values. Log of differentiate and seasonal differentiate are formulated as $\text{ARIMA}(0,1,0)1(0,1,0)^{12}$ or $I(1)1(1)^{12}$.

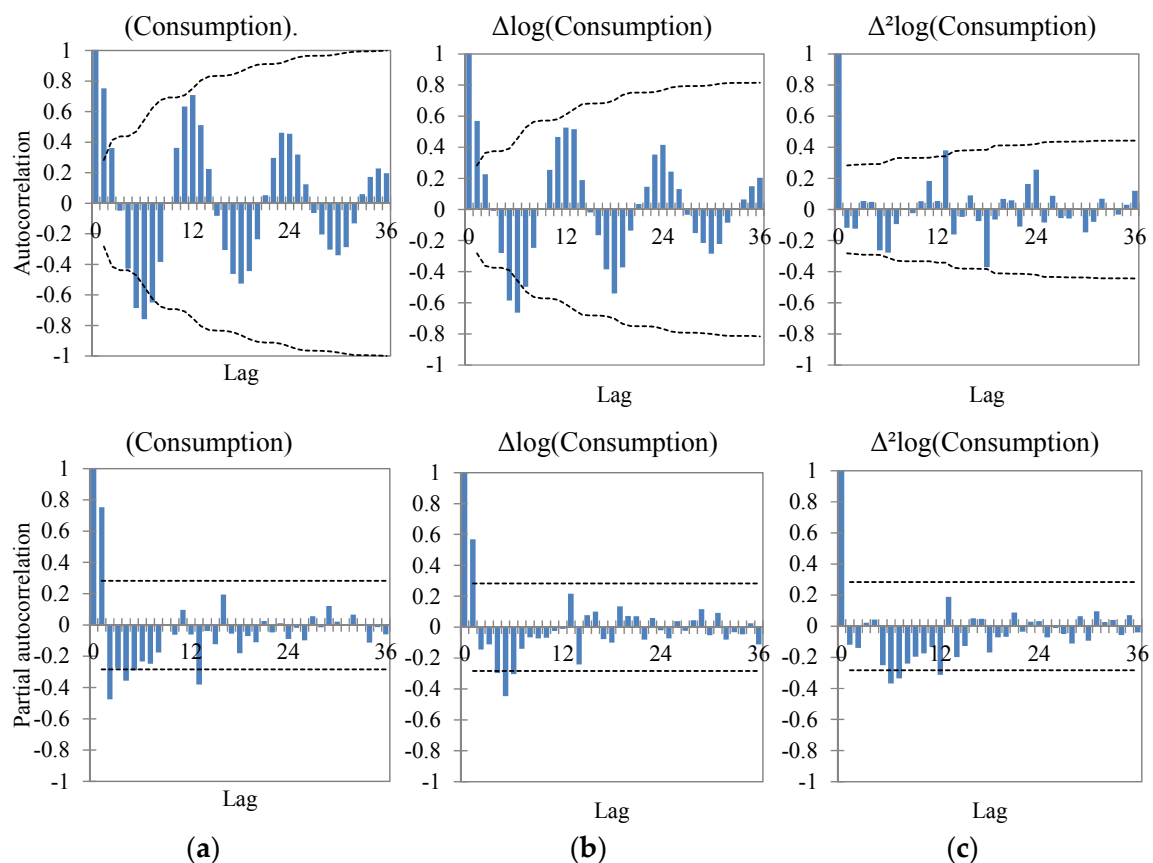


Figure 4. Cont.

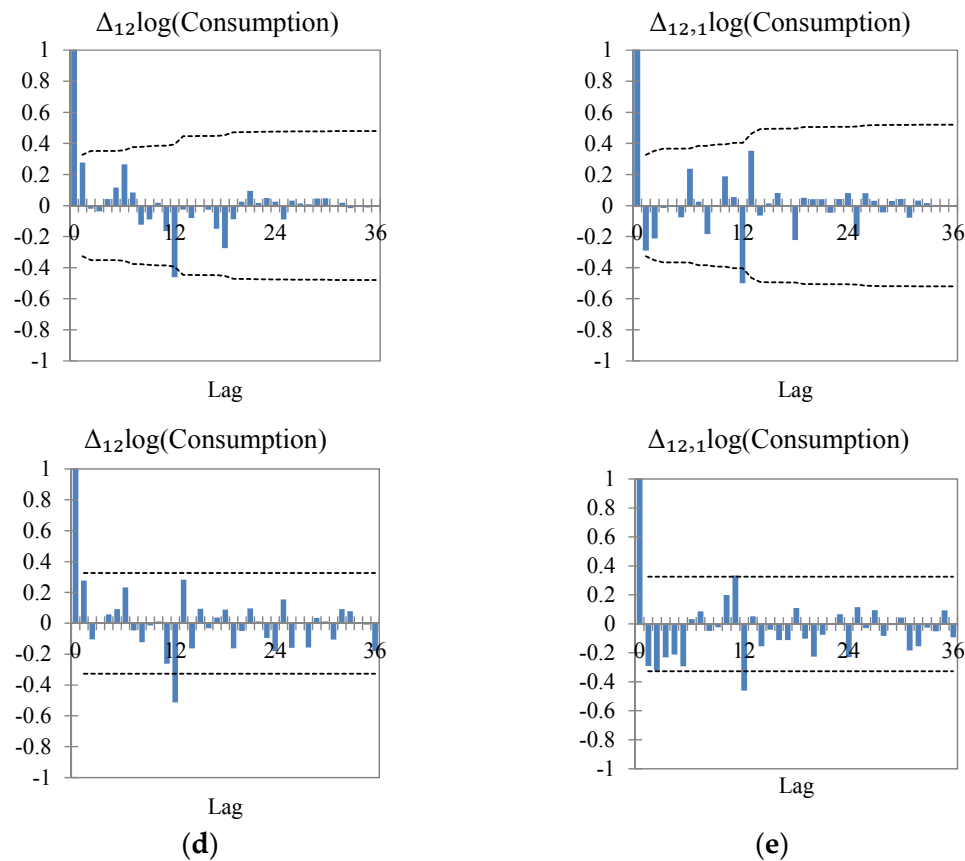


Figure 4. ACF and PACF plots of logged transformed consumption with differenced values for: (a) None; (b) $\Delta\log$; (c) $\Delta^2\log$; (d) $\Delta_{12}\log$; (e) $\Delta_{12,1}\log$ of consumption.

7. Results and Discussion

The forecast results are shown based on the method in Figure 5. The year 2014 is shown in grey and a transparent box in graphs. In this way, the difference between the fitted historical data and the forecast can be easily seen.

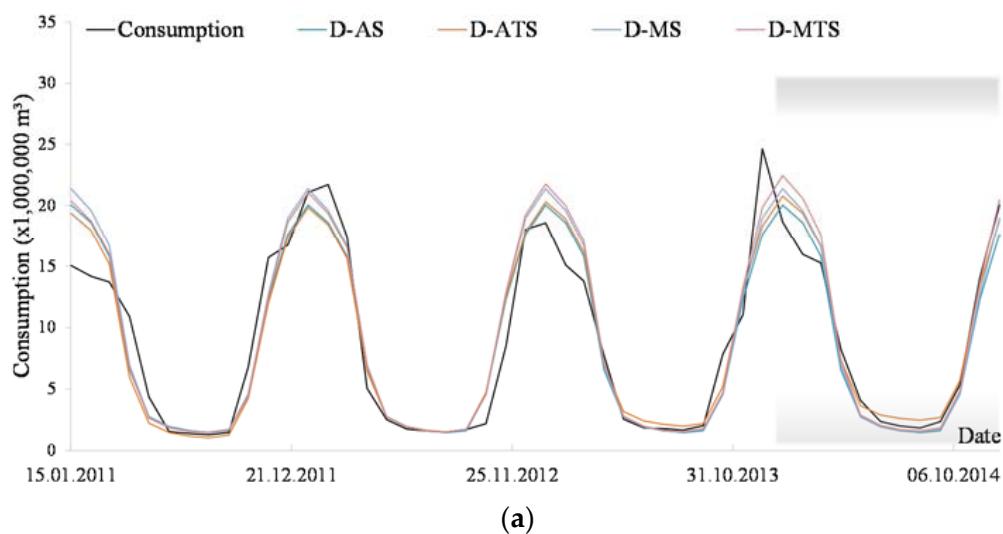


Figure 5. Cont.

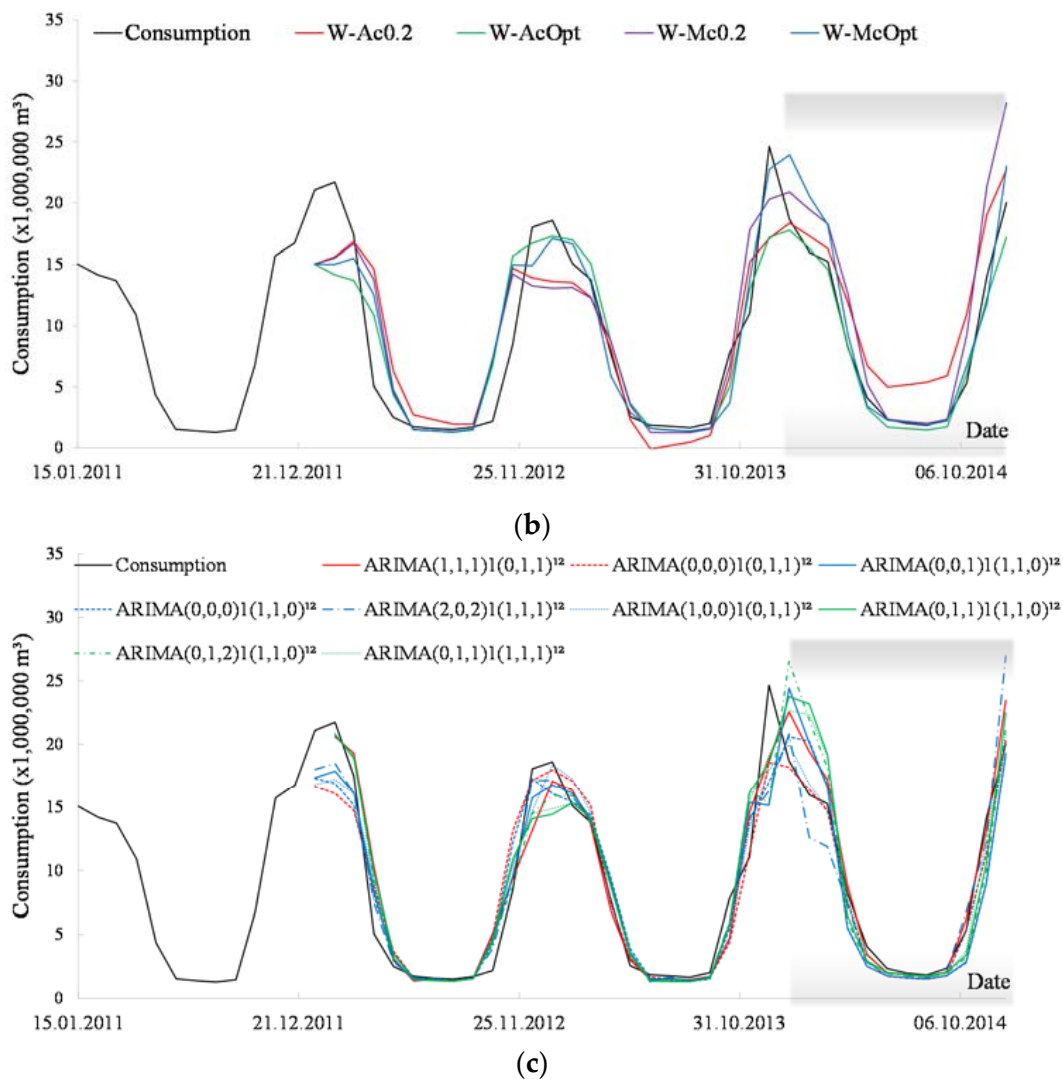


Figure 5. Forecasting results for: (a) Time series decomposition; (b) Holt-Winters exponential smoothing; (c) ARIMA and SARIMA model.

For the decomposition model, the first method in the graph, at the beginning of 2014, the consumption and forecast difference is considerably large. However, this difference decreases gradually by the end of the year. \hat{R}^2 and MAPE values for D-AS, D-ATS, D-MS, D-MTS are 0.907, 0.910, 0.909, 0.915 and 19%, 20%, 20%, 19%, respectively. It is clearly seen that high consumption in January affected forecasting in 2014. Even though spring and autumn seasons are difficult parts of the forecast because of climate changes, the decomposition forecast is well fitted here. The best outcome for 2014 is 15% MAPE at additive trend seasonal decomposition.

Holt-Winters exponential smoothing is applied with two sub-models, additive and multiplicative. Both sub-models' coefficients are taken as 0.2 in the first study of the Holt-Winters model. After gathering results, the second study coefficients of the model are calculated with least square regression to find the optimized solution. The first study gives a weaker estimation than the optimized coefficients. Considering the additive and multiplicative models, the multiplicative models generate higher estimation values than additive models. The major reason behind this outcome is that the seasonal impact in the Holt-Winters method increases multiplicatively. Although additive stays at similar levels, the consumption forecast rises steadily with the multiplicative influence. Another point is that the Holt-Winters method has negative value estimations. Since parameters are taken as 0.2, the June 2013 consumption is around $-95 \times 10^3 \text{ m}^3$ (Figure 5b). Since the estimation values cannot

be negative, it can be evidently seen that the parameter selection is important. The optimized α , β , γ parameters are 0, 0.01, 0.375 and 0.15, 0, 0.74 for the additive and multiplicative methods, respectively. However in the Holt-Winters method, the optimized parameters α and β are very close to zero. Parameter γ is on average above 0.5, such that it proves the seasonal influence on the series, clearly. Table 5 presents the results based on the data range and methodology. In cases where parameters are not optimized according to the optimized situation, (α , β , $\gamma = 0.2$) MAPE values are high and \hat{R}^2 values are low (Table 5). Non-optimized results present worse performance. For optimized results, however, the additive method has better outcomes than the multiplicative method, both on the entire series and the 2014 estimation. The lowest MAPE over the 2011–2014 period with the additive method is 28.81% and the highest \hat{R}^2 value is 0.846. In 2014, the lowest MAPE is 14.01% while the highest \hat{R}^2 is 0.983. The additive methodology obtains the lower MAPE values in 2014 for the time series decomposition and Holt-Winters methods. Moreover, Holt-Winters has 1% less MAPE value than the time series composition.

Table 5. Error measurement for Holt-Winters exponential smoothing forecast.

Date Range	Methodology	$\alpha, \beta, \gamma = 0.2$		Optimized Parameters	
		MAPE	\hat{R}^2	MAPE	\hat{R}^2
January 2011—December 2014	Additive	62.02%	0.747	28.81%	0.846
	Multiplicative	36.86%	0.758	29.60%	0.835
January 2014—December 2014	Additive	74.63% ¹	0.946	14.01% ¹	0.983
	Multiplicative	26.67% ¹	0.954	15.05% ¹	0.960

¹ MAPE values are for 2014 is shown as MAPE₂₀₁₄.

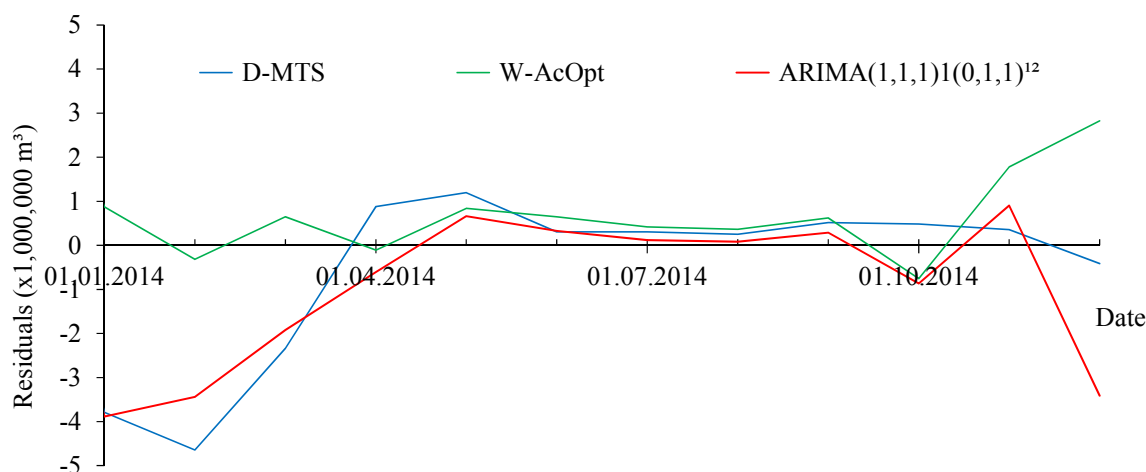
The third method used in this study is ARIMA. The estimation of natural gas consumption with the ARIMA method needs stationary data. Therefore, as the first stage, ACF and PACF graphs should be explored. The second stage of determination of stationarity applies ADF and PP tests. The results of stationarity represented in Section 5 are named Pre-Forecast. In ACF and PACF graphs (Figure 4c–e), the 12th lag state of the stationary series, which is formed by taking both the secondary difference and seasonal difference, it is seen that the significance crosses the boundary in ACF and PACF autocorrelograms and there is a relation between the two. Therefore, the results prove that seasonality is critical. In order to identify the forecast accuracy, AIC, BIC, \hat{R}^2 , MAPE and MAPE₂₀₁₄ (MAPE in 2014) are determined as selection criteria. For each selection criteria, the best ARIMA model results and result values are presented in Table 6. The outcome graphs are also visualized in Figure 5c. For each selection criteria of I(2)1 series of the ARIMA method, the best results are different. AIC and BIC have the same ARIMA(1,2,1)1; however, the MAPE series has ARIMA(0,2,2)1 as the best result. For the I(0)1(1) series, the best outcome on the AIC and BIC criteria is the ARIMA(0,0,1)1(1,1,0)¹² model while \hat{R}^2 and MAPE are ARIMA(3,0,q)1(1,1,1)¹² where q is the non-seasonal parameter of MA and its values are 2 and 3. For MAPE₂₀₁₄, the best model is ARIMA(0,0,0)1(0,1,1)¹², which shows the effectiveness of the first seasonal MA parameter. The I(1)1(1) series performs nearly similar results to the I(0)1(1) series on \hat{R}^2 , MAPE and MAPE₂₀₁₄ criteria. In the I(1)1(1) series, the ARIMA(3,1,3)1(1,1,1)¹² model is the most suitable model on \hat{R}^2 . ARIMA(1,1,0)1(1,1,1)¹² and ARIMA(1,1,1)1(0,1,1)¹² models are found to be the best models for MAPE and MAPE₂₀₁₄ criteria.

The SARIMA series has at least 1 seasonal parameter. This shows the strength of the seasonal aspect of the series. Since seasonality is not contained in the I(2)1 series, the results are considerably weak. In seasonality included models, the \hat{R}^2 value is around 0.95, proving that the influence of the method is precisely high during the estimation. AIC and BIC values can vary between −600 and 340,000 [59]. In this research, the minimum AIC is observed as −17.26, while the minimum BIC is observed as −14.82 among obtained results. The lowest AIC and BIC indicate that I(0)1(1) models give more accurate results.

Table 6. Best results of error and forecasting measurement for ARIMA and SARIMA.

Differencing	AIC	BIC	\hat{R}^2	MAPE	MAPE ₂₀₁₄
$\Delta^2 \log(\text{Consumption})$	(1,2,1)1 5.54	(1,2,1)1 10.12	(3,2,3)1 0.753	(0,2,2)1 335%	(0,2,2)1 1186%
$\Delta_{12} \log(\text{Consumption})$	(0,0,1)1(1,1,0) ¹² (−17.26)	(0,0,0)1(1,1,0) ¹² (−14.82)	(3,0,3)1(1,1,1) ¹² 0.956	(3,0,2)1(1,1,1) ¹² 17.9%	(0,0,0)1(0,1,1) ¹² 12.9%
$\Delta_{12,1} \log(\text{Consumption})$	(0,1,1)1(1,1,0) ¹² (−10.09)	(0,1,1)1(1,1,0) ¹² (−6.69)	(3,1,3)1(1,1,1) ¹² 0.950	(1,1,0)1(1,1,1) ¹² 18.9%	(1,1,1)1(0,1,1) ¹² 12.9%

Figure 6 presents the 2014 residual graph of the lowest MAPE₂₀₁₄ values observed using the three estimation methods. The time series decomposition is represented in blue, Holt-Winters is represented in green and SARIMA is represented in red in the series. The fact that winter consumption is 10 times greater than summer consumption, is also seen in the estimation errors. Although residuals are low during the summer period, in winter 5 times more residuals occur than in summer.

**Figure 6.** Residuals of Consumption for MAPE₂₀₁₄.

The methods studied in this paper are currently used on research related to the energy sector, mainly, in wind speed, hot water demand and wind power generation. For instance, Prema and Rao applied Holt-Winters, ARIMA and time series decomposition methods, and they found 28.63%, 23.26%, 18.24% MAPE', respectively. They also mentioned that 30% MAPE is acceptable by the Government of India [37]. We found MAPE for the time series decomposition, Holt-Winters and ARIMA methods at 19%, 14% and 12.9% respectively. Gelažanskas and Gamage found the R value for the time series decomposition, Holt-Winters and ARIMA to be 0.863, 0.811, 0.872 respectively [36] to forecast hot water demand. In our study, the \hat{R}^2 value of the time series decomposition, Holt-Winters and ARIMA are 0.915, 0.846 and 0.956, respectively. Wu and Peng introduced a wind power generation forecasting model and they compared their result with the ARIMA method [60]. They found 38.57% MAPE with ARIMA forecasting whereas we achieved a three times lower MAPE. Our results prove that these methods are suitable for the natural gas demand forecast over the mid-term range, over a year measured on a monthly basis.

8. Conclusions

The main reasons for households and low consuming commercial users to use natural gas are heating, cooking and water heating. Even though cooking and water heating routinely occur, heating only appears in the winter period. Natural gas consumption also increases related to infrastructure investments and growth. The data used for forecasting in this study is prepared for the Sakarya province of Turkey. Households and low consuming commercial users' 4-year consumption data

between years 2011–2014 are gathered in monthly periods. This consumption data decreases in the summer while it increases in the winter. Therefore, the study researches natural gas demand forecasting by applying univariate seasonal and statistical methods. Well-known techniques of Time Series Decomposition and Winters Exponential Smoothing can be easily applied with spreadsheet software in daily life. However, ARIMA models need moderate knowledge and software containing ARIMA methods. Decision makers can use the natural gas demand forecasting results obtained from forecasting models as decision support systems. Therefore, they can comfortably use the supporting system for determining year-ahead demand and show the consistency of forecasts by comparing their prediction and statistical method results.

Based on the results, the main conclusions of the paper are as follows: the stationarity of the series cannot be accepted after applying one time differencing because the series still include seasonality. Taking advantage of applying various methods such as time series decomposition, Holt-Winters exponential smoothing, and ARIMA-SARIMA, it is evaluated that the error rates decrease as the computation complexity of the method increases. Also the fact that infrastructure investments of the region where the data is gathered are almost completed, the investigated dataset does not show an increasing trend in consumption. However, the time series decomposition method, such as the simplest method, can be used by decision makers by calculating one by one, manually without using any statistical software. Moreover, even the worst case in the D-AS model is $0.907 R^2$, 20% MAPE, 15% MAPE₂₀₁₄ which is a better result than [36,37]. This outcome shows the possibility of forecasting natural gas demand.

Future research of this study will be in three different directions. The first case will use independent variables such as temperature, humidity, wind speed, number of subscribe and unit price if applicable. The second case applies methods such as ARIMAX (ARIMA with exogenous), regression models, learning algorithms, etc. The last case involves changing the time density such as using daily forecasts to make monthly estimations for a year.

Acknowledgments: The authors would like to acknowledge the AGDAŞ—Adapazarı Natural Gas Distribution Company for information about sector and valuable supports to prepare the data.

Author Contributions: Mustafa Akpınar designed the models and prepared the forecasts. Data analysis and interpretation of results are achieved by Nejat Yumusak and Mustafa Akpınar. Flow of the paper is organized by Nejat Yumusak and written by Mustafa Akpınar and Nejat Yumusak.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Brown, S.P.A.; Yucel, M.K. Deliverability and regional pricing in U.S. natural gas markets. *Energy Econ.* **2008**, *30*, 2441–2453. [[CrossRef](#)]
2. Evans, P.C.; Farina, M.F. *The Age of Gas and the Power of Networks*; General Electric Company: Fairfield, NY, USA, 2013.
3. BOTAS İletim Sebekesi İşleyiş Düzenlemelerine İlişkin Esaslar (*The Basis of Regulatory Process on Transmission Network*); BOTAS: Ankara, Turkey, 2013; pp. 1–50.
4. *Turkish Natural Gas Market Report 2014*; Energy Market Regulatory Authority: Ankara, Turkey, 2015.
5. Zhang, G.P. Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing* **2003**, *50*, 159–175. [[CrossRef](#)]
6. Harold, J.; Lyons, S.; Cullinan, J. The determinants of residential gas demand in Ireland. *Energy Econ.* **2015**, *51*, 475–483. [[CrossRef](#)]
7. Potočnik, P.; Soldo, B.; Šimunović, G.; Šarić, T.; Jeromen, A.; Govekar, E. Comparison of static and adaptive models for short-term residential natural gas forecasting in Croatia. *Appl. Energy* **2014**, *129*, 94–103. [[CrossRef](#)]
8. Taspınar, F.; Celebi, N.; Tutkun, N. Forecasting of daily natural gas consumption on regional basis in Turkey using various computational methods. *Energy Build.* **2013**, *56*, 23–31. [[CrossRef](#)]
9. Soldo, B.; Potočnik, P.; Šimunović, G.; Šarić, T.; Govekar, E. Improving the residential natural gas consumption forecasting models by using solar radiation. *Energy Build.* **2014**, *69*, 498–506. [[CrossRef](#)]

10. Dagher, L. Natural gas demand at the utility level: An application of dynamic elasticities. *Energy Econ.* **2012**, *34*, 961–969. [[CrossRef](#)]
11. Tavakoli, E.; Montazerin, N. Stochastic analysis of natural gas consumption in residential and commercial buildings. *Energy Build.* **2011**, *43*, 2289–2297. [[CrossRef](#)]
12. Akpınar, M.; Yumusak, N. Naïve forecasting household natural gas consumption with sliding window approach. *Turk. J. Electr. Eng. Comp. Sci.* **2016**, *24*. [[CrossRef](#)]
13. Sánchez-Úbeda, E.F.; Berzosa, A. Modeling and forecasting industrial end-use natural gas consumption. *Energy Econ.* **2007**, *29*, 710–742. [[CrossRef](#)]
14. Zhu, L.; Li, M.S.; Wu, Q.H.; Jiang, L. Short-term natural gas demand prediction based on support vector regression with false neighbours filtered. *Energy* **2015**, *80*, 428–436. [[CrossRef](#)]
15. Szoplik, J. Forecasting of natural gas consumption with artificial neural networks. *Energy* **2015**, *85*, 208–220. [[CrossRef](#)]
16. Yu, F.; Xu, X. A short-term load forecasting model of natural gas based on optimized genetic algorithm and improved BP neural network. *Appl. Energy* **2014**, *134*, 102–113. [[CrossRef](#)]
17. Aramesh, A.; Montazerin, N.; Ahmadi, A. A general neural and fuzzy-neural algorithm for natural gas flow prediction in city gate stations. *Energy Build.* **2014**, *72*, 73–79. [[CrossRef](#)]
18. Demirel, O.F.; Zaim, S.; Çalışkan, A.; Ozuyar, P. Forecasting natural gas consumption in Istanbul using neural networks and multivariate time series methods. *Turk. J. Electr. Eng. Comp. Sci.* **2012**, *20*, 695–711.
19. Akpınar, M.; Adak, M.F.; Yumusak, N. Forecasting Natural Gas Consumption with Hybrid Neural Networks—Artificial Bee Colony. In Proceedings of the IEEE 2016 2nd International Conference on Intelligent Energy and Power Systems, Kiev, Ukraine, 7–11 June 2016.
20. Aras, N. Forecasting Residential Consumption of Natural Gas Using Genetic Algorithms. *Energy Explor. Exploit.* **2008**, *26*, 241–266. [[CrossRef](#)]
21. Izadyar, N.; Ghadamian, H.; Ong, H.C.; Moghadam, Z.; Tong, C.W.; Shamshirband, S. Appraisal of the support vector machine to forecast residential heating demand for the District Heating System based on the monthly overall natural gas consumption. *Energy* **2015**, *93*, 1558–1567. [[CrossRef](#)]
22. Izadyar, N.; Ong, H.C.; Shamshirband, S.; Ghadamian, H.; Tong, C.W. Intelligent forecasting of residential heating demand for the District Heating System based on the monthly overall natural gas consumption. *Energy Build.* **2015**, *104*, 208–214. [[CrossRef](#)]
23. Timmer, R.P.; Lamb, P.J. Relations between temperature and residential natural gas consumption in the Central and Eastern United States. *J. Appl. Meteorol. Climatol.* **2007**, *46*, 1993–2013. [[CrossRef](#)]
24. Iranmanesh, H.; Abdollahzade, M.; Miranian, A. Mid-term energy demand forecasting by hybrid neuro-fuzzy models. *Energies* **2012**, *5*, 1–21. [[CrossRef](#)]
25. Huntington, H.G. Industrial natural gas consumption in the United States: An empirical model for evaluating future trends. *Energy Econ.* **2007**, *29*, 743–759. [[CrossRef](#)]
26. Shaikh, F.; Ji, Q. Forecasting natural gas demand in China: Logistic modelling analysis. *Int. J. Electr. Power Energy Syst.* **2016**, *77*, 25–32. [[CrossRef](#)]
27. Melikoglu, M. Vision 2023: Feasibility analysis of Turkey’s renewable energy projection. *Renew. Energy* **2013**, *50*, 570–575. [[CrossRef](#)]
28. Siemek, J.; Nagy, S.; Rychlicki, S. Estimation of natural-gas consumption in Poland based on the logistic-curve interpretation. *Appl. Energy* **2003**, *75*, 1–7. [[CrossRef](#)]
29. Gutiérrez, R.; Nafidi, A.; Gutiérrez Sánchez, R. Forecasting total natural-gas consumption in Spain by using the stochastic Gompertz innovation diffusion model. *Appl. Energy* **2005**, *80*, 115–124. [[CrossRef](#)]
30. Ma, H.; Wu, Y. Grey predictive on natural gas consumption and production in China. In Proceedings of the 2009 2nd Pacific-Asia Conference on Web Mining and Web-Based Application, Wuhan, China, 6–7 June 2009.
31. Xie, Y.; Li, M. Research on Prediction Model of Natural Gas Consumption Based on Grey Modeling Optimized by Genetic Algorithm. In Proceedings of the 2009 IITA International Conference on Control, Automation and Systems Engineering (CASE 2009), Zhangjiajie, China, 11–12 July 2009.
32. Soldo, B. Forecasting natural gas consumption. *Appl. Energy* **2012**, *92*, 26–37. [[CrossRef](#)]
33. Ediger, V.S.; Akar, S.; Ugurlu, B. Forecasting production of fossil fuel sources in Turkey using a comparative regression and ARIMA model. *Energy Policy* **2006**, *34*, 3836–3846. [[CrossRef](#)]
34. Liu, L.-M.; Lin, M.-W. Forecasting residential consumption of natural gas using monthly and quarterly time series. *Int. J. Forecast.* **1991**, *7*, 3–16. [[CrossRef](#)]

35. Yalcintas, M.; Bulu, M.; Kucukvar, M.; Samadi, H. A Framework for Sustainable Urban Water Management through Demand and Supply Forecasting: The Case of Istanbul. *Sustainability* **2015**, *7*, 11050–11067. [[CrossRef](#)]
36. Gelažanskas, L.; Gamage, K. Forecasting Hot Water Consumption in Residential Houses. *Energies* **2015**, *8*, 12702–12717. [[CrossRef](#)]
37. Prema, V.; Rao, K.U. Time series decomposition model for accurate wind speed forecast. *Renew. Wind Water Sol.* **2015**, *2*, 18. [[CrossRef](#)]
38. Martínez-Álvarez, F.; Troncoso, A.; Asencio-Cortés, G.; Riquelme, J. A Survey on Data Mining Techniques Applied to Electricity-Related Time Series Forecasting. *Energies* **2015**, *8*, 13162–13193. [[CrossRef](#)]
39. Wang, J.; Zhu, S.; Zhang, W.; Lu, H. Combined modeling for electric load forecasting with adaptive particle swarm optimization. *Energy* **2010**, *35*, 1671–1678. [[CrossRef](#)]
40. Yaffee, R.A.; McGee, M. *Introduction to Time Series Analysis and Forecasting: With Applications of SAS and SPSS*, 1st ed.; Academic Press: Orlando, FL, USA, 2000.
41. DeLurgio, S.A. *Forecasting Principles and Applications*; Irwin McGraw-Hill: Boston, MA, USA, 1998.
42. Makridakis, S.; Wheelwright, S.C.; Hyndman, R.J. *Forecasting: Methods and Applications*; John Wiley: New York, NY, USA, 2008.
43. Delurgio, S.; Bhame, C. *Forecasting Systems for Operations Management*; Irwin Professional Pub: New York, NY, USA, 1991.
44. Winters, P.R. Forecasting Sales by Exponentially Weighted Moving Averages. *Manag. Sci.* **1960**, *6*, 324–342. [[CrossRef](#)]
45. Sudheer, G.; Suseelatha, A. Short term load forecasting using wavelet transform combined with Holt-Winters and weighted nearest neighbor models. *Int. J. Electr. Power Energy Syst.* **2015**, *64*, 340–346. [[CrossRef](#)]
46. Son, Y.S. Prediction of Electricity Sales by Time Series Modelling. *Korean J. Appl. Stat.* **2014**, *27*, 419–430. [[CrossRef](#)]
47. Box, G.E.P.; Jenkins, G.M.; Reinsel, G.C. *Time Series Analysis*, 5th ed.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2008.
48. Valipour, M. Long-term runoff study using SARIMA and ARIMA models in the United States. *Meteorol. Appl.* **2015**, *22*, 592–598. [[CrossRef](#)]
49. Burtiev, R.; Greenwell, F.; Kolivenko, V. Time Series Analysis of Wind Speed and Temperature in Tiraspol, Moldova. *Environ. Eng. Manag. J.* **2013**, *12*, 23–33.
50. SAS/ETS®13.2 User's Guide; SAS Institute Inc.: Cary, NC, USA, 2014.
51. Cowpertwait, P.S.P.; Metcalfe, A.V. *Introductory Time Series with R*; Springer: New York, NY, USA, 2009.
52. Hamilton, J.D. *Time Series Analysis*; Princeton University Press: Princeton, NJ, USA, 1994.
53. Wei, W.W.S. *Time Series Analysis: Univariate And Multivariate Methods*, 2nd ed.; Pearson Addison Wesley: Boston, MA, USA, 2006.
54. Dickey, D.A.; Fuller, W.A. Distribution of the Estimators for Autoregressive Time Series With a Unit Root. *J. Am. Stat. Assoc.* **1979**, *74*, 427.
55. Said, E.; Dickey, D.A. Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika* **1984**, *71*, 599–607. [[CrossRef](#)]
56. Phillips, P.C.B.; Perron, P. Testing for a unit root in time series regression. *Biometrika* **1988**, *75*, 335–346. [[CrossRef](#)]
57. O'Rourke, N.; Hatcher, L.; Stepanski, E.J. *A Step-by-Step Approach to Using SAS for Univariate & Multivariate Statistics*; SAS Institute Inc.: Cary, NC, USA, 2005.
58. Cavaliere, G.; Harvey, D.I.; Leybourne, S.J.; Taylor, A.M.R. Testing for Unit Roots Under Multiple Possible Trend Breaks and Non-Stationary Volatility Using Bootstrap Minimum Dickey-Fuller Statistics. *J. Time Ser. Anal.* **2015**, *36*, 603–629. [[CrossRef](#)]
59. Burnham, K.P. Multimodel Inference: Understanding AIC and BIC in Model Selection. *Sociol. Methods Res.* **2004**, *33*, 261–304. [[CrossRef](#)]
60. Wu, Q.; Peng, C. Wind Power Generation Forecasting Using Least Squares Support Vector Machine Combined with Ensemble Empirical Mode Decomposition, Principal Component Analysis and a Bat Algorithm. *Energies* **2016**, *9*, 261. [[CrossRef](#)]

