



Daily basis mid-term demand forecast of city natural gas using univariate statistical techniques

Mustafa Akpınar^{1*} , Nejat Yumuşak² 

¹Department of Software Engineering, Sakarya University, Sakarya, 54050, Turkey

²Department of Computer Engineering, Sakarya University, Sakarya, 54050, Turkey

Highlights:

- Mid-term natural gas demand forecast with daily basis consumption
- 365 days forecast at one time thanks to statistical techniques
- Improving monthly forecasts using daily forecast results

Keywords:

- Demand forecasting
- Natural gas
- Time series decomposition
- Holt-Winters exponential smoothing
- ARIMA/SARIMA

Article Info:

Research Article
Received: 09.01.2019
Accepted: 27.09.2019

DOI:

10.17341/gazimmfd.494094

Correspondence:

Author: Mustafa Akpınar
e-mail:
akpinar@sakarya.edu.tr
phone: +90 264 295 32 38

Graphical/Tabular Abstract

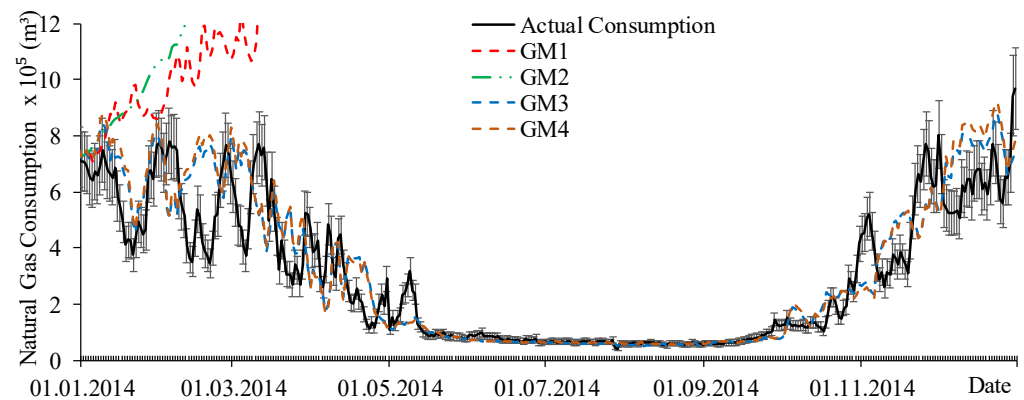


Figure A. Mid-term forecasting results for SARIMA models in daily-basis

Purpose: The aim of this study is to find results of applying univariate techniques among statistical time series methods used in the literature, which can perform one-year daily forecast at a single step (which can estimate 365 days at once). These results will be both daily basis year ahead forecasts, and can be aggregated monthly and show the forecast of medium-term natural gas demand. Since the techniques applied in this study have a structure for the consumption of natural gas in the city, it can be easily used as a guiding study by decision makers.

Theory and Methods:

This study forecasts year 2014 using 3-year data models on a daily basis between 2011 and 2013. The research forecasts mid-term natural gas demand on daily and monthly basis applying time series decomposition (TSD), Holt-Winters and ARIMA/SARIMA models, which are univariate statistical methods including seasonality. The important feature of these methods is that they can perform long-term forecasts using only historical data. The study forecasts 365-day on a daily basis and 12-month on a monthly basis at once.

Results:

Natural gas estimation, the lowest daily error rate is performed by ARIMA(1,0,1)1(0,1,1)³⁶⁵ model with 24.6% MAPE and highest compliance 0.802 R² (Figure A, GM₄). In the monthly conversion, the lowest estimation is found by the same model with 11.32% MAPE and highest compliance (R² - 0.981). These results show that seasonal ARIMA models are the most suitable models among the univariate techniques. Other results are shown in the graphs and figures.

Conclusion:

Decision-makers can test the consistency of their predictions and their methods. According to the results of the study, models with high computational complexity have lower error rates. The lowest error rate is found with SARIMA, then Holt-Winters and lastly TSD.



Günlük temelli orta vadeli şehir doğal gaz talebinin tek değişkenli istatistik teknikleri ile tahmini

Mustafa Akpınar^{1*}, Nejat Yumuşak²

¹Sakarya Üniversitesi, Bilgisayar ve Bilişim Bilimleri Fakültesi, Yazılım Mühendisliği Bölümü, 54050 Serdivan Sakarya, Türkiye

²Sakarya Üniversitesi, Bilgisayar ve Bilişim Bilimleri Fakültesi, Bilgisayar Mühendisliği Bölümü, 54050 Serdivan Sakarya, Türkiye

Ö N E Ç İ K A N L A R

- Günlük temelde doğal gaz talep tahmininin orta vadeli yapılması
- İstatistiki mevsimsel teknikler sayesinde bir anda 365 gün tahmini
- Günlük tahmin kullanılarak aylık tahminlerin iyileştirilmesi

Makale Bilgileri

Araştırma Makalesi
Geliş: 09.01.2019
Kabul: 27.09.2019

DOI:

10.17341/gazimmfd.494094

Anahtar Kelimeler:

Talep tahmini,
doğal gaz,
zaman serilerinin
ayrıştırılması,
Holt-Winters modeli,
ARIMA/SARIMA

ÖZET

Doğalgaz sektöründe yüksek tüketime sahip şirketler (sanayi kuruluşları gibi) ve şehir dağıtım şirketleri yıl öncesinde aylık, yıl içinde de günlük tüketim talep tahminlerini bildirmekle yükümlüdür. Bu çalışma günlük ve aylık temelde orta vadeli doğalgaz talep tahminini tek değişkenli mevsimsellik içeren istatistiki yöntemler (zaman serileri ayrıştırılması, Holt-Winters üstel düzeltme, ARIMA/SARIMA) ile gerçekleştirmiştir. Bu tekniklerin ortak özelliği mevsimsellik içeren zaman serilerinde geçmiş bilgiyi barındıracak biçimde model kurmaları ve bir anda çok sayıda tahmine izin vermeleridir. Yapılan bu çalışmada günlük temelde 365 günlük, aylık temelde de 12 aylık tahmin bir anda gerçekleştirilmiştir. İstatistiksel olarak uygun tüm tahmin modellerinde, günlük temelde yıl öncesi doğalgaz tüketimini, 2014 yılı için en düşük hata, en yüksek uyum %24,6 MAPE ve 0,802 R² değeri ile ARIMA(1,0,1)1(0,1,1)³⁶⁵ modelinde olmuştur. Bu modelin katsayıları istatistiksel olarak da anlamlı olup, kalıntıları beyaz gürültü olarak bulunmuştur. Aynı model aylık tahminlerde de en düşük hataya (MAPE) ve en yüksek uyuma (R²) sahip olduğu görülmüştür. Aylık tahminde, bu modelin MAPE ve R² değeri sırasıyla %11,32 ve 0,981 olmuştur. Bu sonuçlar mevsimsel ARIMA modellerinin tek değişkenli teknikler arasında en uygun tahmin tekniği olduğunu göstermiştir. Bir anda çok sayıda tahmin yapılabilmesi ve sonuçlarının kabul edilebilir olması, bu tekniklerin yıl öncesinde aylık ve günlük tahminlerde kullanılabilmesine olanak tanımaktadır.

Daily basis mid-term demand forecast of city natural gas using univariate statistical techniques

H I G H L I G H T S

- Mid-term natural gas demand forecast with daily basis consumption
- 365 days forecast at one time thanks to seasonal statistical techniques
- Improving monthly forecasts using daily forecast results

Article Info

Research Article
Received: 09.01.2019
Accepted: 27.09.2019

DOI:

10.17341/gazimmfd.494094

Keywords:

Demand forecasting,
natural gas,
time series decomposition,
Holt-Winters model,
ARIMA/SARIMA models

ABSTRACT

City distribution companies or companies with high consumption are required to report monthly consumption demand forecasts for the year ahead and year based daily consumption demand forecasts in natural gas sector. This paper studies forecasting daily and monthly demand for mid-term natural gas as contract estimations using statistical methods (time series decomposition, Holt-Winters exponential smoothing, ARIMA/SARIMA), include univariate seasonality. In the study, 365-day forecast is performed on a daily basis and 12-month forecast is performed on a monthly basis at once. Among all statistically appropriate forecasting models, ARIMA(1,0,1)1(0,1,1)³⁶⁵ model found daily basis year ahead natural gas consumptions the best with the lowest error, highest compliance with 24.6% MAPE and 0.802 R², for the year 2014. The coefficients of this model were statistically significant and the residuals were found as white noise. The same model has the lowest error (MAPE - 11.32%) and highest compliance (R² - 0.981) in the monthly estimations as well. The results show that seasonal ARIMA models are the most appropriate estimation technique among the univariate techniques. The fact that many predictions can be made at a time and the results are acceptable allow these techniques to be used in the year ahead monthly and daily forecasting.

*Sorumlu Yazar/Corresponding Author: akpinar@sakarya.edu.tr, nyumusak@sakarya.edu.tr / Tel: +90 264 295 32 38

1. GİRİŞ (INTRODUCTION)

Doğal gaz dünyadaki başlıca enerji kaynakları arasında yer almaktadır. Günümüzde ısınma, elektrik üretimi, pişirme, üretim ve ulaşımda kullanılan doğal gazın tüketimi enerji ihtiyaçlarının artmasıyla paralellik göstererek yükselmektedir. Doğal gazın yoğun biçimde kullanımı, elektrik sektöründe olduğu gibi piyasa yapısının oluşmasını sağlamaya başlamıştır. Elektrik enerjisinde saatlik olarak yapılan enerji alım-satım işlemlerinin bir benzeri doğal gaz enerjisinde de günlük olarak kendini göstermektedir. Doğal gaz sektöründeki piyasa yapısında, bir veya birkaç gün öncesinden tüketilmesi beklenen doğal gazın tahmini yapılarak, ülke içine girmesi gereken gaz miktarı belirlenmeye çalışılmaktadır. Böylece eksik ithal edilen gaz sebebiyle sistemde meydana gelen dengesizliklerin önüne geçilmeye çalışılmaktadır. Fazla veya eksik ithalatın günlük olarak en düşük hata ile gerçekleştirilmesi istenen durumdur. Gün öncesinde gerçekleştirilen tahminin dışında, orta vadede yıl öncesi tahmin de ikili sözleşmelere konu olmaktadır. Yıl öncesi tahminde yıl boyu şehirdeki tüketicilerin doğal gaz kullanımı belirlenmeye çalışılmaktadır ve toptan satış şirketleri ile gaz dağıtım şirketleri arasında yıllık doğal gaz sözleşmeleri imzalanmaktadır. Yıl öncesi aylık doğal gaz tahmini ne kadar az hata ile gerçekleşirse, gün öncesi doğal gaz tahminine o oranda yardımcı olmaktadır. Yıl öncesi yapılan aylık tahmin ile gün öncesi tahmin arasında önemli bir ilişki söz konusudur. Tahmin hatasını düşürmek, %75'i ithal edilen temel enerji kaynaklarında da maliyetleri düşürecektir [1]. Bu sebeple düşük tahminler karar vericiler tarafından önemli bir konu haline gelmektedir. Akademik olarak bu hatayı düşürmek için farklı çalışmalar yapılmaktadır. Doğal gaz sektörünün özelinde birçok çalışma bulunmaktadır [2-9]. Wang vd. yaptığı çalışmada Çin'deki 30 şehrin senelik doğal gaz tüketim tahmini çok yönlü en iyileme aracı (multiverse optimizer) kullanılarak Nash doğrusal olmayan Gray Bernoulli tekniği (Nash Nonlinear Gray Bernoulli) ile gerçekleştirilmiştir [2]. Bu çalışmada 2 sene doğrulama amacıyla kullanılmış, tahmin hataları ARIMA (Autoregressive Integrated Moving Average – Özbağlımsal Bütünleşik Hareketli Ortalama) ve yapay sinir ağları (YSA) ile karşılaştırarak gösterilmiştir. Diğer bir çalışmada ise gün öncesi şehir doğal gaz tüketim tahmini yapay arı kolonisi (YAK) algoritması ile eğitilmiş YSA modeli ile gerçekleştirilmiş ve geri yayılım ile eğitilmiş YSA modeli ile karşılaştırılmıştır [3]. Tahminde sadece geçmiş veri kullanılmış ve YSA-YAK modelinin daha düşük hatalı tahmin yaptığı görülmüştür. Scarpa ve Bianco ise doğal gazı yıllık olarak kalman filtresi ve regresyon kullanarak gerçekleştirmiştir [4]. Nan vd. ise faktör seçim algoritması, yaşam genetik algoritması ve karar destek makinası ile model oluşturarak doğal gaz tahminini Yunanistan'daki farklı bölgelerdeki şehirler için günlük olarak gerçekleştirmişlerdir [6]. Yine bu çalışmada, hava durumunu faktörleri belirlemek için kullanmışlar ve YSA ile karşılaştırmışlardır. Naim ve Mahari ise tek değişkenli tahmin yöntemlerini kullanarak aylık doğal gaz tahmini gerçekleştirmişlerdir. Kullandıkları veri seti mevsimsel

değildir ve buna uygun olarak basit, Holt, Hata-Trend-Mevsim (ETS), basit üstel düzeltme, ARIMA ve YSA teknikleri kullanmışlardır. Yaptıkları 3 aylık tahminlerde ARIMA(1,0,0) modelinin en uygun sonuçları verdiğini belirlemişlerdir [7]. Şehir tüketiminin belirlendiği diğer bir çalışmada takvimsel olaylar ve hava durumu verisi kullanılarak gün öncesi doğal gaz tahmini yapılmıştır [8]. Bu çalışmada çoklu doğrusal regresyon denklemi kullanılarak düşük hata değeri elde edilmiştir. Enerji üzerine tek ve çok değişkenli tekniklerin kullanıldığı çalışmalar da bulunmaktadır [10-14]. Es vd. yıllık temelde toplam enerji tahminini Türkiye için YSA modeli kullanarak gerçekleştirmişler ve çoklu regresyon sonuçları ile karşılaştırmışlardır [10]. Bu çalışmada ekonomik ve demografik veriler bağımsız değişken olarak kullanılmıştır. Jiang vd. ise mevsimsellik olmayan veride yıllık Hindistan enerji talep tahminini ARIMA, metabolic Gray model ve YSA modelleri kullanarak yapmışlardır [11]. Enerji ve doğal kaynaklardan biri olan su tüketimi üzerine de tahmin çalışmalarını bulunmaktadır. Bu çalışmalardan birinde uyarlamalı sinirsel bulanık mantık yaklaşımı kullanılarak ekonomik ve iklim verileri kullanılması ile su tüketimi tahmini gerçekleştirilmiştir [12]. Hamzaçebi ve Kutay ise çalışmalarında Türkiye'nin elektrik yükünü YSA, regresyon ve ARIMA teknikleri ile yıllık temelde tahmin etmiştir [13]. Çalışmalarında kullandıkları YSA ve regresyon modellerinde yıl ve nüfus bilgisini bağımsız değişken olarak kullanmışlardır. Başoğlu ve Bulut ise elektrik talebi için 24 saatlik ve günlük ortalama saatlik tahmin modelini YSA kullanarak gerçekleştirmişlerdir [15]. Rehman vd. ise Pakistan'ın yıllık enerji talebini belirlemeye çalışmıştır [14]. Çalışmalarında ARIMA ve Holt-Winters tekniklerini kullanmışlardır ve mevsimsellik verilerinde bulunmamaktadır. Yine aynı çalışmada literatürdeki çalışmaların özet tablosu bulunmaktadır. Doğal gaz tahmini üzerine farklı kriterler üzerinden literatürdeki çalışmaları gösteren en güncel yayın 2017 yılında Şebalj vd. tarafından gerçekleştirilmiştir [16]. Prema ve Rao rüzgâr hızı tahmininde Holt-Winters üstel düzeltme, ARIMA ve zaman serilerinin ayrıştırmasını kullanırken [17], Gelažanskas ve Gamage ise evsel sıcak su tüketiminin talep tahminini üstel düzeltme, SARIMA, zaman serilerinin ayrıştırılması ile gerçekleştirmişlerdir [18]. Wu ve Peng ise rüzgâr gücü üretme tahminini ayrıştırma, temel bileşen analizi ve yasa algoritmaları destekli en küçük kareler destek vektör makinaları ile gerçekleştirmiş, sonuçlarını ARIMA yöntemi ile karşılaştırmıştır [19]. Elektrik tarafında ise yük tahmini ARMA modeli ve Adaptif Ağ Tabanlı Bulanık Çıkarım Sistemi (Adaptive Network Based Fuzzy Inference Systems-ANFIS) ile yapılarak karşılaştırılmıştır [20]. Genel olarak literatür incelendiğinde zaman serileri yöntemlerinin temel alınarak, farklı tahmin teknikleri ile karşılaştırılması yapılmaktadır. Bunun sebebi zaman serileri tekniklerinin temel tahmin modelleri olmalarıdır.

Bu çalışmanın amacı, literatürde kullanılan istatistiki tek değişkenli zaman serileri teknikleri ile bir yıllık günlük doğal gaz tüketim tahminini tek aşamada gerçekleştirmek (365

günlük tahmini bir anda gösterebilecek) ve bu tahminlerin sonuçlarını bulmaktadır. Bu sonuçlar hem günlük olarak bir sene sonrasının öngörüsü olacak, hem de aylık hale getirilerek orta vadeli doğal gaz talep tahminini gösterebilecektir. Şehir tüketimi için yapılan bu çalışmadaki teknikler, karar vericiler tarafında da rahatlıkla kullanılabilir yapıları bulunduğundan dolayı, yol gösterici bir çalışma olmaktadır. Bu teknikler daha önce günlük temelde yıllık tahminde hem enerji sektörü, hem de diğer sektörler için kullanılmamıştır ve sonuçları da gösterilmemiştir. Orta vadede yüksek yoğunlukta tahmine, kullanılan tekniklerin uygunluğunun tespit edilmesi bu çalışma ile literatüre özgün katkı olarak sunulmuştur. Çalışmanın devamı şu şekilde düzenlenmiştir: Verinin elde edilmesi, tahmin modelleri ve performans kriterleri 2. bölüm olan Metotlarda, çalışmada hazırlanan model bilgileri 3. bölüm de, tahmin öncesi çalışmalar 4. bölümde, bulgular ve tartışma 5. bölümde, çalışmanın değerlendirmesi ise 6. bölüm olan sonuçlarda yapılmıştır.

2. METOTLAR (METHODOLOGIES)

2.1. Veri (The Data)

Türkiye'nin doğal gaz piyasası BOTAŞ ve EPDK tarafından yönetilmektedir. Saatlik olarak ölçülen ve hesaplanan RMS-A (reducing and measuring stations – RMS) tüketim bilgileri, Türkiye'deki doğal gaz piyasasını yöneten Boru Hatları ile Petrol Taşıma Anonim Şirketi (BOTAŞ) ile mutabakat sağlanarak günlük hale getirilmektedirler. Böylece saatlik verideki oluşan anlık gaz değişiminden kaynaklanan hatalar günlük veriye dönüştürülürken arındırılmış olur. RMS-A bazındaki veriler telemetri sistemi ile uzak veri tabanına aktarılmaktadır. Pilot şehir olarak Sakarya seçilmiştir. Altı farklı RMS-A istasyonundan elde edilen günlük tüketim verisi toplanarak, RMS-A'dan bağımsız şehir tüketimi belirlenmiş olur. Daha sonra, yüksek tüketime sahip sanayi kuruluşlarının günlük tüketimleri (RMS-B ve RMS-C istasyonlarına sahip kuruluşlar) RMS-A toplam tüketiminden çıkartılarak şehrin kendi tüketimi elde edilmiştir.

2.2. Tahmin Modelleri (Forecasting Models)

Bu bölümde zaman serilerinin ayrıştırılması (Time Series Decomposition – TSD), Holt-Winters Üstel Düzeltme (Holt-Winters Exponential Smoothing – HWES), ARIMA ve SARIMA (Seasonal Autoregressive Integrated Moving Average – Mevsimsel Özbağlımsal Bütünleşik Hareketli Ortalama) modelleri hakkında bilgiler verilecektir. Zaman serileri, mevsimsel, eğilim, konjonktür ve düzensiz olarak dört farklı bileşenden oluşmaktadır [17, 21-23]. TSD'nin iki yaklaşımından biri olan çarpımsal modelinde her bileşen bir diğeriyle çarpılarak bulunurken, toplamsal modelde her bileşen bir diğeriyle toplanmaktadır [17, 21-24]. Modelde uzun süreli davranışı eğilim bileşeni ile belirlerken, uzun süreli çevrimsel etkiler konjonktür bileşeni ile gösterilir. Mevsimsel bileşen kısa süreli ve sürekli salınımları, düzensiz

bileşen ise beklemedik veya tahmin edilemeyen değerleri göstermektedir [25]. İki dönem tahmin yapmak konjonktür etkisini barındırabilir [26]. Bunun sebebi bu sürenin uzun olması ve iş süresince değişikliklerin meydana gelmesidir. Ayrıştırma yöntemini uygulamanın avantajı; anlaşılması ve uygulamasının kolay olmasıdır. Üç dönemlik veri modelin kurulması için yeterlidir. Üstel düzeltme yönteminde en son gerçekleşme verisinin etkisi ağırlığınca olmaktadır. En son gerçekleşme verisi dışında daha eski verilerin de tahminde etkisi azalarak bulunmaktadır. Üstel düzeltme denkleminde, kendinden önceki tahminin özinelemeli olarak alfa katsayısı ile modele eklenmesi ile ağırlıklı bir denklem oluşmaktadır [17, 21-25, 27]. Holt-Winters düzeltme yönteminde de basit üstel düzeltme yöntemine benzer çalışma yapılmaktadır. Buradaki önemli nokta, HWES yöntemi, mevsimsel tahmin yapmaya uygun model oluşturmak için γ katsayısını ve bunu içeren denklemi barındırmaktadır. HWES; seviye (level), eğilim ve mevsimsel denklemler kullanarak tahmin denklemini uygun biçimde hesaplar. Seviye, eğilim ve mevsimsel bileşenlerin ağırlıkları sırasıyla α , β , γ olarak gösterilir. HWES metodu üstel düzeltme sürecinde eğilim, mevsimsellik ve rastlantısallığın etkisini göstermenin etkili bir yoludur [22]. Bu tekniğin tek dezavantajı üç parametreyi belirlemenin uzun zaman alabilmesidir. Bu yöntemde ayrıştırma gibi toplamsal ve çarpımsal olabilmektedir.

ARIMA zaman serisi analiz ve tahmininde popüler bir tekniktir [28-34]. Bu yöntemin üç bileşeni bulunmaktadır. Model oluşturulurken kendinden önceki veriler ile ilişkisi özbağlımsal kısım (autoregressive - AR), hareketli ortalamalar kısmı (moving average - MA) ve seriyi durağanlaştırmak için bütünleştirme kısmı (integration - I) bulunmaktadır. ARIMA yönteminin kullanımında seride boşluk olmaması gerekmektedir ve seri durağan olmalıdır [21-24, 28, 32-35]. ARIMA modelleri ARIMA(p, d, q) biçimi ile gösterilir. Burada, p AR parametresini, (ϕ_p), d geriye doğru fark işlem değerini, q ise MA parametresini (θ_q) göstermektedir. ARIMA yönteminin mevsimsellik içeren şekilde modellenmesinde mevsimsel ARIMA (Seasonal ARIMA - SARIMA) kullanılmaktadır. ARIMA ile aynı şekilde oluşturulan modelde mevsim uzunluğuna bağlı olarak mevsimsel AR, mevsimsel I ve mevsimsel MA parametreleri de modele eklenerek kurulmaktadır [22, 24]. Genel olarak ARIMA(p,d,q)T(P,D,Q)^s şeklinde gösterilmektedir. Bu gösterimde ilk p, d, q değerleri ARIMA yönteminde gösterilenlerdir. T, dönüştürme parametresidir. Dönüştürme işlemi yoksa sıfırdır. Eğer dönüştürme işlemi logaritmik ise 1 olur. Bunun dışında yazılan sayı ne ise değerler bu sayı ile üstlendirilir. İkinci P, D, Q değerleri ise ARIMA modelinin mevsimsel kısmıdır. Bu değerlerin üstünde bulunan "s" değeri ise mevsim uzunluğunu göstermektedir.

2.3. Box-Jenkins Yaklaşımı (Box-Jenkins Approach)

ARIMA modelleri Box-Jenkins yaklaşımı ile belirlenebilmektedir ve üç aşamadan oluşmaktadır [24]. Bunlar; "model tanımlama", "parametre tahmini ve model

teşhisi" ve "uygulama" aşamalarıdır. Model tanımlama kısmında serinin grafiği, özilinti (autocorrelation - ACF) ile parçalı özilinti (partition autocorrelation - PACF) fonksiyonları incelenir ve durağanlık testleri yapılır. Serideki veri değerleri yüksek olduğunda normalize etmek için logaritma, doğal logaritma gibi dönüşümler yapılır. Bu dönüşümler sonucunda veri daha tahmin edilebilir duruma gelmektedir. Model tanımlamanın bir sonraki aşaması ise durağanlık testleri yapılmaktadır. Durağanlık testleri üzerine birçok çalışma yapılmıştır ve literatürde en sık iki durağanlık (birim kök) testi kullanılmaktadır [36-38]; Bu testler, Genişletilmiş Dickey-Fullder (Augmented Dickey-Fuller - ADF) [37] ve Phillips-Perron (PP) [38] testleridir. Durağanlık testleri sonrasında, durağan olduğu belirlenen seride ikinci aşama olan parametre tahmini ve model teşhisi aşamasına geçilir. Bu aşamada potansiyel modeller için serilere ARIMA yöntemi uygulanır ve parametreler tahminleri yapılır. Bu aşamada kalıntı serileri ile bu kalıntı serilerinin ACF ve PACF grafikleri incelenir, portmanteau testi yapılarak artıkların beyaz gürültü olması durumunda bir sonraki aşama olan uygulamaya geçilir. Üçüncü aşama olan uygulama aşamasında modeller belirlenmiş olur, tahminler elde edilir.

2.4. Performans ve Tahmin Hataları (Performance and Estimation Errors)

MAPE gerçek değer tahmin değerinden çıkartılarak mutlak değeri alınır ve gerçekleşen değere bölünmesiyle elde edilir. Böylece negatif ve pozitif hataların birbirini götürmesini engellemek mutlak değer ile sağlanmış olur [21, 22, 31]. MAPE (Mean Absolute Percent Error – Ortalama Mutlak Yüzdesel Hata) Eş. 1’de gösterilmiştir. Burada \hat{y}_i ; i . tahmin değerini, y_i ; i . gerçekleşen değeri, n ; eleman sayısını göstermektedir.

$$MAPE = \left(\frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \right) \cdot 100\% \quad (1)$$

MAPE değerinin 10%'dan küçük olması yüksek doğrulukta tahmini, 10-20% değerleri arasında olması iyi tahmini, 20-50% değerleri arasında makul tahmini, 50%'nin üstünde ise yanlış tahmini göstermektedir [39].

Gerçekleşen ile tahminin tüm seri boyunca uyumu R^2 ile gösterilmektedir. Bu çalışmada R^2 biçiminde belirtilen ifade düzeltilmiş belirlilik katsayısıdır [22, 23]. Bu değer korelasyon katsayısının karesidir ve yünden bağımsız olarak [0 1] aralığında değer alabilir. Modeldeki veri ve parametre sayısına göre düzeltilerek düzeltilmiş belirlilik katsayısı elde edilir. R^2 ; Eş. 2’de gösterilmiştir. Burada \hat{y}_i ; i . tahmin değerini, y_i ; i . gerçekleşen değeri, n ; eleman sayısını, k ; modeldeki değişken sayısını göstermektedir.

$$R^2 = 1 - \left[1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2} \right] \frac{n-1}{n-k-1}, 0 \leq R^2 \leq 1 \quad (2)$$

Serideki hatalarda MAPE ve R^2 aynı anda yüksek olabilir. Bu durum tahminin sürekli olarak kötü olduğunu, kötü tahminin gerçekleşen değerler ile uyumlu olduğunu gösterir. Çalışmadaki sonuçlarda değerlendirme önceliği MAPE üzerindedir ve uyum terimi olarak R^2 kullanılmaktadır. Diğer bir ifade ile isabet dereceleri R^2 , hata ölçümleri MAPE ile gösterilmiştir. Bu ölçütler ile tüm seri üzerindeki genel güvenilirlik belirlenmiştir.

3.DURUM ÇALIŞMASININ MODEL TANIMLARI (MODEL DEFINITIONS OF CASE STUDY)

Şehir doğal gaz talebi mevsimlere göre değişiklik göstermektedir. Modelleme aşamasında mevsimsel değişimlerin de içinde bulunduğu modellerin kullanılması önemli olmaktadır. Tek değişkenli uzun vadeli tahmin için kullanılan modeller; zaman serisi ayrıştırması, Holt-Winters üstel düzeltme, ARIMA ve SARIMA yöntemleridir. Diğer yapay zekâ ve öğrenme algoritmalarının aksine, modeller mevsimsel etkiyi tutabildikleri için uzun süreli tahminleri bir anda ve daha az gözlem sayısı ile alabilmektedirler. Sakarya, Türkiye’de bulunan şehir giriş istasyonu RMS-A’dan sağlanan veriler günlük temelde 2011-2014 yıllarına ait olarak bulunmaktadır. Bu veri seti toplam 1461 günlük doğal gaz tüketim verisinden oluşmaktadır. İlk üç yıllık veri (2011-2013) olan 1096 adet veri modellerin oluşturulmasında kullanılmıştır. 2014 yılına ait olan 365 günlük veri ise modellerin sınanmasında kullanılmıştır. Yapılan tahminlerin sonuçları bir sonraki bölümde ele alınmış olup, hata terimi olarak MAPE, uyum terimi olarak R^2 kullanılmıştır.

Çalışmanın bundan sonraki kısmında hazırlanan modellerin yöntemleri kısaltmalar ile ifade edilmiştir. Buna göre zaman serilerinin ayrıştırılması için “Dxy” ile ifade edilmiştir. Buradaki “x” modelin ayrıştırma yöntemini, “y” ise bileşenlerini göstermektedir. Holt-Winters üstel düzeltme için “Wxycy” ifadesi kullanılmıştır. Buradaki “x” model türünü, “c” katsayıların bilgisinin başladığını, “y” ise katsayıların durumunu göstermektedir. ARIMA için “GM_x” kısaltması kullanılmıştır. Burada “GM” günlük model olduğunu, “x” değeri ise kaçınıcı model olduğunu göstermektedir. SARIMA/ARIMA modellerinin detay bilgisi sonuçlar tablosunda verilmiştir.

Zaman serilerinin ayrıştırma yönteminde tüketimin davranışına göre sadece “mevsimsel” etki içeren model kurulabileceği gibi, tüketimde artış söz konusu ise “eğilim ve mevsimsellik” içeren model kurulmaktadır. Nüfus artışı, evlere döşenen doğal gaz tesisat sayısının artması, farklı mahallelere doğal gaz götürülmesi tüketimi arttıran etkenlerdir. Bu etkenler kendini eğilim olarak gösterirken, kış ve yaz aylarındaki tüketim davranışları da kendini mevsimsel etki olarak göstermektedir. Zaman serilerinin ayrıştırması altında kullanılacak modeller çarpımsal mevsimsel model için DMS (decomposition, multiplicative, seasonal), çarpımsal eğilim-mevsimsel model için DMTS (decomposition, multiplicative, trend - seasonal), toplamsal mevsimsel model için DAS, toplamsal eğilim-mevsimsel model için ise DATS kısaltması kullanılmıştır.

Mevsimsellik barındıran ve bu çalışmada kullanılan ikinci yöntem Holt-Winters üstel düzeltmedir. Burada α , β ve γ parametreleri modeli oluşturmada kullanılmaktadır sırasıyla seviye, eğilim ve mevsimselliği ifade etmede kullanılır. Varsayılan olarak bu değerler 0,2 olarak alınmaktadır [32] fakat probleme özgü olarak optimize edilmeleri gerekmektedir. Diğer bir ifadeyle 0,2 değeri en iyi sonucu vermemektedir. Katsayıların belirlenmesi çeşitli yaklaşımlar söz konusudur. Bu çalışmada sıradan en küçük kareler (ordinary least squares - OLS) yakınsama tekniği olarak kullanılmıştır. Yakınsaklık değeri 10^{-7} ve iterasyon sayısı 500.000'dir. Oluşturulan modeller Tablo 1'de gösterilmiştir. Çarpımsal ve katsayıların 0,2 alındığı model "WMc0.2" şeklinde, katsayıların optimize edildiği model "WMcOpt" şeklinde gösterilmiştir. Toplamsal modelde de benzer yaklaşım söz konusudur ve sadece "M" yerine "A" ifadesi gelmektedir. Böylece toplamsal ve katsayıların 0,2 ve optimize edildiği modeller "WAc0.2" ve "WAcOpt" şeklinde gösterilmektedir. Eğilimin modelde olmadığı düşünüldüğünde α ve γ katsayıları optimize edilerek $\beta=0$ olması da hesaplanmıştır ve "WAcOpt-E" ve "WMcOpt-E" biçiminde gösterilmiştir. Burada "c" ifadesi katsayı anlamında kullanılmaktadır.

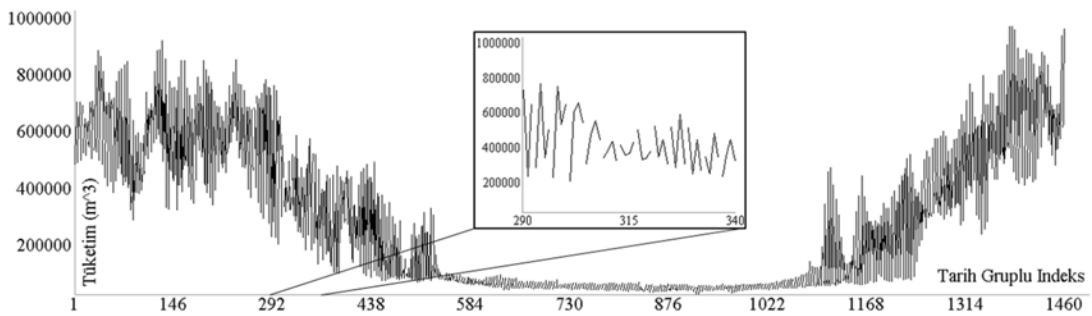
ARIMA/SARIMA yöntemleri bu çalışmada kullanılan diğer yöntemlerdir. ARIMA/SARIMA yöntemleri durağan seriler üzerinde tahmin modeli oluşturabildiğinden dolayı, kullanmadan önce seri durağan değilse, durağanlaştırılması gerekmektedir. Serinin durağanlaştırılması için fark alma işlemi yapılmaktadır. Fark alma işlemi, serinin durağanlığı sağlanıncaya kadar tekrar edilebilmektedir. Daha sonra ARIMA yöntemindeki AR ve MA parametreleri, SARIMA yönteminde ise ARIMA yöntemindeki AR ve MA parametrelerine ek SAR ve SMA (seasonal AR ve seasonal MA) parametreleri belirlenir. Farklı alınmış seriler ARIMA/SARIMA yönteminde I ile ifade edilmektedir.

Seride büyük değerler bulunmasından dolayı (Tablo 2, en yüksek tüketim değeri 974.960 ve en düşük tüketim değeri 27.765) logaritmik ölçeğe dönüştürülmüştür. Dört farklı fark alınmış seri üzerinden tahminler gerçekleştirilmiştir. Bu farkları alınmış seriler sırasıyla: Birincil fark; $\Delta \log(\text{Tüketim})$; I(1)1 veya ARIMA(0,1,0)1, İkincil fark; $\Delta^2 \log(\text{Tüketim})$; I(2)1 veya ARIMA(0,2,0)1, Birincil mevsimsel fark; $\Delta_{365} \log(\text{Tüketim})$, I(0)1(1) veya ARIMA(0,0,0)1(0,1,0)³⁶⁵, Birincil fark ve birincil mevsimsel fark; $\Delta_{365,1} \log(\text{Tüketim})$; I(1)1(1) veya ARIMA(0,1,0)1(0,1,0)³⁶⁵ biçiminde gösterilmiştir.

Doğal gaz tüketimi; hava durumu, abone sayısı ve takvimsel olaylar (tatiller, bayramlar, hafta sonları gibi) ile ilişkilidir. Bu etkiler tüketimin biçimini değiştirmektedir. Bu çalışmanın ana amacı, günlük temelde doğal gaz tüketiminin, iyi bilinen tek değişkenli istatistik teknikler ile şehir seviyesinde bağımsız değişken kullanmadan tahmin edilmesidir. Bu tekniklerin güçlü yanı, bağımsız değişken olmadan 365 günlük tahmini tek seferde getirebilmeleridir. Çalışmanın diğer önemli noktası; tahminlerin rahatlıkla aylık tahmin biçimine getirilebilir ve modellerin aylık performansının da belirlenebilir olmasıdır.

4. TAHMİN ÖNCESİ (PRE-FORECAST)

Doğal gaz talep tahminini gerçekleştirmeden önce verinin incelenmesi ve davranışının görülmesi faydalı ve gereklidir. Şekil 1'de günlük doğal gaz tüketimi alt grafiği görülmektedir. Grafikte her gün için bir seri, diğer bir ifadeyle 365 farklı seri bulunmaktadır. Burada 1-4 arası tarih gruplu indeks sırasıyla 2011, 2012, 2013 ve 2014 yıllarındaki 1 Ocak tarihinin tüketimlerini göstermektedir. Farklı senelerdeki aynı günlerdeki tüketimler incelendiğinde mevsimsel geçişlerin yakın tarihlerde olduğu görülmektedir. Örneğin 290-500 tarih grup indeksleri arasında mevsim



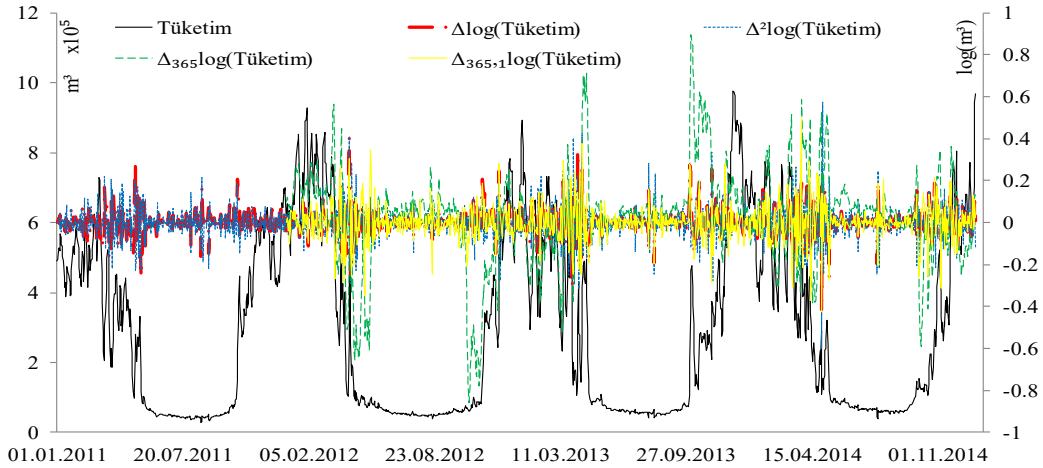
Tablo 1. Holt-Winters üstel düzeltme modellerinin kısaltmaları (Holt-Winters exponential smoothing model abbreviations)

Holt-Winters Modeli	Yöntem	Parametreler	Kısaltma
Toplamsal	$\alpha, \beta, \gamma = 0,2$	$\alpha, \beta, \gamma = 0,2$	WAc0.2
	Tümü optimizasyon	α, β, γ optimizasyon	WAcOpt
	Eğilimsiz optimizasyon	α, γ optimizasyon $\beta=0$	WAcOpt-E
Çarpımsal	$\alpha, \beta, \gamma = 0,2$	$\alpha, \beta, \gamma = 0,2$	WMc0.2
	Tümü optimizasyon	α, β, γ optimizasyon	WMcOpt
	Eğilimsiz optimizasyon	α, γ optimizasyon $\beta=0$	WMcOpt-E

geçiş her yıl için bulunmaktadır. Diğer taraftan farklı yılların bazı günlerinde her yıl için tamamen farklı tüketimler kendini göstermektedir (120-150 arası, 280-300 arası) ve örnek bir bölüm detay grafiğinde verilmiştir. Bunun aksine davranış birbirine yakın olduğu günler de bulunmaktadır (520-560 arası ve 940-1.000 arası). Genel olarak seride mevsimselliğin bulunduğu ilk incelemede görülmektedir.

ARIMA ve SARIMA yöntemlerinde durağanlaştırma işleminin gerçekleştirilebilmesi için öncelikli seri logaritmik tabanda gösterilmeli (serinin değer aralığının düşürülmesi) ve ardından durumu incelenmelidir (Şekil 2). Şekil 2’de fark alma işlemi “ Δ ” ifadesi ile gösterilmiştir. Bu çalışmada tek fark, çift fark, tek mevsimsel fark, mevsimsel ve birincil fark alma işlemleri yapılmış ve sırasıyla, Δ , Δ^2 , Δ_{365} ve $\Delta_{1,365}$ ile ifade edilerek, Şekil 2’de gösterilmiştir. Buradaki mevsimsellik yıllık temeldedir ve 365 günlük fark alma işlemi gerçekleştirilmiştir. Şekil 2’deki sağ eksen logaritmik tüketimi, sol eksen ise gerçek tüketimi göstermektedir. Siyah çizgili tüketim serisi sol eksende gösterilirken, logaritması alınmış olan diğer seriler sağ eksende gösterilmiştir. Tablo 2’de doğal gaz tüketim serisinin tanımlayıcı istatistikleri

verilmiştir. Logaritma işlemi günlük tüketimleri [4,4435, 5,9890] aralığına indirgiş olup, tüketim aralığının standart sapması 0,45 olmuştur. Fark alınmış serilerin ortalaması ve standart sapmaları sıfır civarındadır. Bu durum serilerin temel istatistikler açısından durağanlık şartlarını sağladığını göstermektedir. Temel istatistik ve görsel araçlar ile serinin durağan olduğuna karar verildikten sonra, durağanlık testleri serilere uygulanmaktadır. Durağanlık testlerinden en bilindik testler ADF ve PP testleridir [23, 36-38]. Tablo 3, bu çalışmada hazırlanan serilerin ADF ve PP test sonuçlarını göstermektedir. Bu testler seride birim kök durumunu araştırmaktadır. Birim kök seride yapısal kırılmaları incelemektedir. Diğer bir ifadeyle serinin birbirinden bağımsız serilere bölünebilmesini inceler. Böylece serinin birim kökü bulunuyorsa, durağanlıktan söz edilemez. Birim kök testlerinde hipotez en az bir birim kökün bulunması üzerine kurulur. Daha sonra test uygulanır ve seçilen olasılık değerine göre birim kökün varlığı için ya hipotez reddedilir ve seri durağandır denir, ya da hipotez reddedilemez ve serinin durağan olmadığı ortaya çıkar. Bu çalışmada birim kök testi için p değeri, $\alpha=0,05$ alınmıştır. ADF ve PP testleri üç şekilde kurulabilmektedir. Bunlar sabit ve eğilim olmadan “Yok”, sadece sabit varken “C”, hem sabit hem de eğilim



Şekil 2. Tüketim ve fark serileri (Consumption with integrated series)

Tablo 2. Günlük tüketimin açıklayıcı istatistikleri (Descriptive statistics of daily consumptions)

Değişken (m ³)	Gözlem	Minimum	Maksimum	Ortalama	Standart Sapma
(Tüketim)	1461	27,765	974,960	290,541	252,416
log(Tüketim)	1461	4,44349	5,988987	5,253828	0,453635
$\Delta\log(\text{Tüketim})$	1460	-0,42328	0,402571	0,000202	0,061085
$\Delta^2\log(\text{Tüketim})$	1459	-0,59924	0,570782	0,000005	0,078176
$\Delta_{365}\log(\text{Tüketim})$	1096	-0,86111	0,908935	0,020006	0,231417
$\Delta_{365,1}\log(\text{Tüketim})$	1095	-0,45898	0,351911	0,000023	0,085488

Tablo 3. Tüketim serisi için durağanlık testleri (Stationary test for consumption series)

Durağanlık için p -değeri Değişken (m ³)	ADF Test			PP Test		
	Yok	C	C + T	Yok	C	C + T
Tüketim	0,0251	0,0045	0,0250	0,1403	0,0556	0,2014
$\Delta\log(\text{Tüketim})$	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
$\Delta^2\log(\text{Tüketim})$	0,0000	0,0000	0,0000	0,0001	0,0001	0,0001
$\Delta_{365}\log(\text{Tüketim})$	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
$\Delta_{365,1}\log(\text{Tüketim})$	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000

varken “C+T” şeklinde gösterilmiştir. Buna göre tüketim serisi fark alınmadan hipotezi reddedememiştir ve durağan olmadığını göstermiştir. Diğer tüm fark alınmış serilerde ise durağanlık söz konusudur.

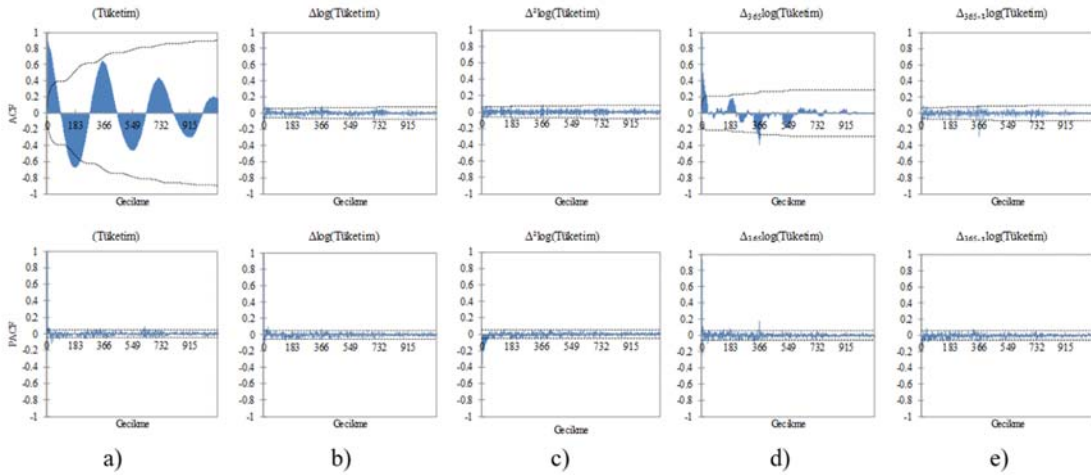
ACF ve PACF grafikleri (özilinti grafikleri (autocorrelogram)) hem durağanlığı belirlemede hem de serinin kendisiyle ilişkisi hakkında bilgi vermede kullanılmaktadır. Tüketimin kendi içindeki ilişkisinin anlamlı olduğu bölge, kesikli çizgiler ile ayrılmaktadır. Bu sınırların dışına çıkılan noktalarda serinin kendisiyle ilişkisinin güçlü olduğu görülür [22-24]. Burada fark alınmamış tüketim serisindeki mevsimsel örüntü rahatlıkla görülmektedir (Şekil 3). Fakat fark işlemi sonrasında oluşturulmuş diğer tüm serilerde durağanlık görülmektedir. Mevsimsel fark alınmış serinin (Şekil 3d) başlangıç, 183. gün, 365. gün ve 549. gün civarında ilişki olduğu görülse bile bunlardan sadece 365. gün civarındaki gecikmenin anlamlı olduğu görülebilmektedir. Genel olarak durağan serilerde, senelik olarak anlamlı ilişki olduğu grafiklerde kendini göstermektedir. Böylece ARIMA yöntemi oluşturulan doğal gaz tüketim serilerinde kolaylıkla kullanılabilir. Serilerdeki mevsimsellik bu yol ile de görülebilir. AR ve MA katsayılarında uygun tahmin modelinin belirlenmesi için mevsimsel ve mevsimsel olmayan katsayıları sıfır ile üç arasında farklı değerler için bulunmaya çalışılmıştır. Böylece her seri için en fazla 256 farklı tahmin modeli kurulmuştur. Bu serilerden bazılarında fark alma işleminden dolayı modeller oluşturulamamıştır. Örneğin çift fark almasından dolayı ARIMA(0,1,0)I(1,1,3) modeli kurulamayacaktır. Çünkü kendisinden önce iki mevsimsel veri bulunmaktadır ve mevsimsel model tarafında MA katsayısı için 3 değerine ilişkin sonuç belirlenmemektedir. Farklı ARIMA modelleri için (256 tane) en uygun tahmin sonuçları belirlenmiştir. ARIMA ve SARIMA yöntemlerinde fark alma operatörü “I” olarak gösterilmektedir. Sadece birincil fark ve mevsimsel fark olan seri ARIMA(0,1,0)I(0,1,0)³⁶⁵ veya I(1)I(1)³⁶⁵ biçiminde gösterilmiştir.

5. SONUÇLAR VE TARTIŞMALAR (RESULTS AND DISCUSSIONS)

Bu çalışmada günlük veri kullanılarak doğal gaz tüketimi tahmini, tek değişkenli mevsimsellik içeren istatistikî teknikler kullanılarak gerçekleştirilmiştir. Doğal gaz talep tahmin sonuçlarının günlük olması ve bir anda 365 günlük tahminin yapılabilmesi, hem günlük tahmin hem de aylık tahmin sonuçlarının gösterilmesine imkân vermiştir. Gün öncesi talep tahmininde 1096 gün kullanılarak 365 gün tahmin edilmeye çalışılmıştır. Kullanılan yöntemlerinin sonuçları günlük olarak bölüm 5.1’de gösterilecektir. Çalışmanın 5.2 bölümünde yapılan günlük tahminler aylık hale getirilerek aylık doğal gaz tahmin sonuçları gösterilecektir.

5.1. Günlük Doğal Gaz Talep Tahmin Sonuçları (Daily Natural Gas Demand Forecasting Results)

Öncelikli olarak günlük tahminler için “G” kısaltması ile modeller gösterilmiştir. Daha sonra modelin kendi kısaltması gelmektedir. Örneğin günlük zaman serileri çarpımsal mevsimsel model için kısaltma DMS’den GDMS’ye dönüşmektedir. Günlük tahmin modellerinde ilk kullanılan teknik zaman serilerinin ayrıştırılmasıdır ve daha önce belirtilen DAS, DATS, DMS ve DMTS modelleri ile tahminler yapılmıştır. Buna göre zaman serilerinin ayrıştırma sonuçları GDAS, GDATS, GDMS ve GDMTS için sırasıyla hatalar %26,23, %29,47, %26,95, %27 MAPE ve uyum değerleri 0,8028, 0,8053, 0,7989, 0,8039 R² olmuştur. MAPE ile R² değerlerinin birbiriyle uyumlu olduğu tespit edilmiştir. Tahmin grafiği olan Şekil 4’de zaman serilerinin ayrıştırma yöntem sonuçları görülmektedir. Tahmin serileri ve bu serilerin MAPE değerlerine bakıldığında, dört modelin de birbirine yakın davranış sergilediği görülmektedir. Şekildeki siyah çizgi gerçekleşen tüketimi, bu çizginin üzerindeki hata bantları %15 sınırlarını göstermektedir. Yazın tüketim daha belirgin ve mevsimsel davranış daha durağan olduğu için bu

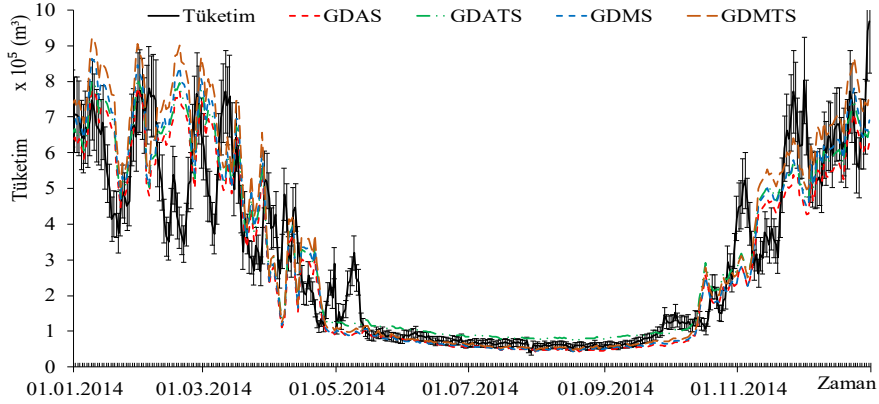


Şekil 3. Tüketim değerleri için ACF ve PACF Grafikleri: a) Yok (b) $\Delta \log$ (c) $\Delta^2 \log$ (d) $\Delta_{365} \log$ (e) $\Delta_{365,1} \log$
(ACF and PACF plots of consumption values for (a) None (b) $\Delta \log$ (c) $\Delta^2 \log$ (d) $\Delta_{365} \log$ (e) $\Delta_{365,1} \log$ of consumption)

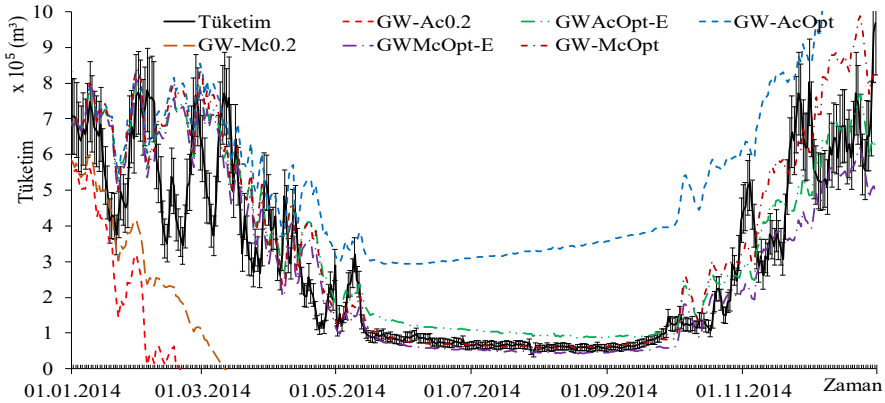
kısımdaki hatalar daha düşük, R^2 değerleri daha yüksek olmuştur. Ocak – Nisan ve Ekim – Aralık ayları arasında ise tüketim tahminin yükselmesi modeller tarafından elde edilmiş fakat dalgalanmalar tespit edilememiştir.

Holt-Winters üstel düzeltme yönteminde eklenen γ parametresi, üstel düzeltme yöntemlerinde bulunmayan mevsimselliğin de modele dâhil edilmesini sağlamıştır. Holt-Winters üstel düzeltme yöntemi için mevsimselliğin toplamsal ve çarpımsal modelleri mümkün olabilmektedir Tablo 4’de ve Şekil 5’de Holt-Winters üstel düzeltme yönteminin sonuçları gösterilmiştir. Sonuçlarda parametrelerin 0,2 olduğu durumda hataların en yüksek olduğu görülmektedir. Eğilimli modellerin geçmiş veriye bakarak artan bir eğilim ($\beta=0,02$) tespit ettiği görülmektedir. Artan eğilim modelin 365 günü yukarı yönde tahmin etmesine sebep olmuştur. Eğilimsiz modellerde ($\beta=0$) hataların eğilimli modellere göre daha düşük olduğu görülmüştür. Bu durum belirlenen $\beta=0,02$ katsayının çok daha düşük olması gerektiğini göstermektedir. Modeldeki diğer bir parametre olan mevsimsellik katsayısı γ , mevsimsellik eleman sayısı düşük olan serilerde (4 çeyrek, 7 gün, 12 ay gibi) yüksek değerlerde olabilmektedir. Bu çalışmada mevsimsellik çevrimi 365 gün olduğundan dolayı, bu katsayıda 365 gün ile ters orantılı olarak düşük olmuştur. Mevsimsellik katsayısı yüksek olduğu durumda ise WAc0.2 ve WMc0.2 modellerinde olduğu gibi mevsimselliğin aşırı

değerlenmesinden dolayı seriyi uygun biçime getiremediğinden, uç değerlerde tahmin gerçekleştirilebilmektedir. Şekil 5’de görüldüğü gibi parametrelerin 0,2 olduğu modellerin (eğilim içeren ve optimize edilmemiş modeller) sonuçları üçüncü ay civarında eksi değerlere gitmektedir. WAcOpt modelinin de çok yüksek sapma ile tahmin ettiği belirlenmiştir. Toplamsal olan bu modelde eğilimin etkisi rahatlıkla görülebilmektedir. Eğilimsiz modellerden toplamsal modelin çarpımsal modele göre daha yüksek tahminler gerçekleştirdiği de gözle görülebilecek kadar açıktır. WAcOpt-E, WMcOpt-E ve WMcOpt modellerinin tahminde yaklaşık sonuçlar verdiği görülmektedir. Tablo 4’deki diğer modellere göre WAcOpt-E, WMcOpt-E ve WMcOpt modellerinin nispeten yüksek R^2 değerleri şeklindeki sonuçları doğrulamaktadır. HWES modellerinin MAPE ve R^2 değerleri incelendiğinde ters orantı olduğu rahatlıkla görülebilmektedir. En düşük MAPE ve en yüksek R^2 değerleri WAcOpt-E ve WMcOpt-E modellerinde görülmüştür. WAcOpt-E modelinde R^2 değeri, WMcOpt-E modelindeki R^2 değerine göre %5 daha yüksek olsa da, MAPE değerinde buna ters orantılı olarak düşmesi beklenirken, %46 oranında artmıştır. Günlük tahminde en yüksek R^2 değeri WMcOpt modelinde bulunmaktadır. WMcOpt-E modelinin MAPE değeri WMcOpt modelinden %0,2 düşük, R^2 değeri 0,05 yüksektir. Diğer bir ifadeyle WAcOpt-E modeli daha uyumlu tahmin serisi elde ederken, hata oranı daha yüksek kalmıştır. Bunun sebebi, WAcOpt-E



Şekil 4. Zaman serisi ayrıştırılması yöntemi tahmin sonuçları (Forecasting results for time series decomposition)



modelinin yaz tahmininde (15 Mayıs – 20 Eylül arası), tüm mevsim boyunca yüksek tahminde bulunmasıdır. Bu durumda WMcOpt-E modelinin daha iyi performans gösterdiği görülmüştür.

ARIMA/SARIMA yöntemleri, bu çalışmada kullanılan diğer yöntemlerdir. 2011-2013 yılları arasındaki veriler kullanılarak öncelikle ARIMA yöntemi üzerinde durulmuştur. 4. bölümde belirtildiği gibi, serideki değişkenlik yüksek olduğu için öncelikle doğal gaz tüketiminin logaritması alınmıştır. Elde edilen I(0)1 serisinin durağan olmadığı Tablo 3 ve Şekil 2 ile Şekil 3a üzerinde görülmüştür. Bu sebeple modeller ilk aşamada sadece I(1)1 ve I(2)1 için çalıştırılmıştır. AR ve MA parametrelerinin 0, 1, 2 ve 3 olması durumu için 16 farklı model elde edilebilmektedir. Birincil ve ikincil fark serileri bulunduğu için model sayısı 32 olmuştur. Burada hazırlanan 32 farklı modelde en düşük MAPE değeri ARIMA(2,2,3)1 modelinde 2014 yılı için %1837,8 olmuştur. Bu modeldeki R² uyum değeri ise 2014 yılı için 0,0006 olmuştur. ARIMA(2,2,3)1 modeli oluşturulmasında kullanılan 2011-2013 yıllarının MAPE değeri ise %8,81, R² değeri ise 0,972 olmuştur. Modelin katsayıları Tablo 5’de gösterilmiştir ve hem AR hem de MA tarafındaki katsayıların anlamlı olduğu görülmüştür. ARIMA teknikleri birkaç adım sonrasında tahmin etmek için hazırlanan modellerdir. Uzun süreli tahminde hatalar çarpıcı biçimde yükselmektedir. Çalışmanın bu kısmı ARIMA tekniğinin kullanılabilir olmadığını göstermiştir. 32 model için yapılan kalıntı testlerinde, Ljung-Box testinin sonuçlarında kalıntıların beyaz gürültü olmadığını, diğer bir ifadeyle anlamlılık içerdiğini ve modellerin tahminde yetersiz kaldığı görülmüştür. Bu sebeple mevsimsel olmayan ARIMA yöntemleri çalışmada uygun olmamıştır ve ARIMA yönteminde mevsimsellik içeren modellere yönelim

olmuştur. SARIMA yönteminde ise hem mevsimsel AR ve MA bileşenleri hem de mevsimsel fark alma işlemleri önemli hale gelmektedir. Mevsimsel AR ve MA bileşenlerinin yanında mevsimsel olmayan AR ve MA bileşenleri de eklenerek modeller oluşturulmaktadır. Burada I(1)1(0), I(2)1(0), I(0)1(1) ve I(1)1(1) serileri üzerinden modeller oluşturulmuştur. AR, MA, SAR (mevsimsel AR) ve SMA (mevsimsel MA) parametreleri sırasıyla 0, 1, 2 ve 3 için modelleri kurulmuş ve sınanmıştır. Dört farklı seri için dört parametrenin, dört farklı değer alması sonucu 1024 model elde edilmiştir. Önceki adımda sadece ARIMA yöntemi sonuçları bulunmuştur. Burada SAR ve SMA parametreleri de belirlenebileceği için birincil ve ikincil fark serilerinin de modelleri mevsimsel bileşen içerecek biçimde bulunabilir. Bu modellerden 32 tanesi ARIMA modelleridir ve bir üst paragrafta sonuçları gösterilmiştir. Nihai SARIMA model sayısı böylece 992’ye düşmüştür. Diğer taraftan seride üç mevsimsel uzunluk bulunduğu için mevsimsel fark bulunan serilerde SAR ve SMA parametreleri en fazla 2 olabilmektedir. Bu durumda model sayısı 452’ye düşürmüştür. Tüm modeller belirlendikten sonra kalıntıların incelenmesine geçilmiştir. Bu aşamada öncelikle ACF/PACF grafikleri incelenmiştir [31]. Daha sonra kalıntıların ilişkisi ve beyaz gürültü grafiği kullanarak inceleme işlemi yapılmıştır. Bir sonraki aşamada ise olağanlık grafiği incelenmiştir. Üç farklı yapı ile kalıntılar inceleniş olup anlamlı model sayısının 96 olduğu görülmüştür. Örneğin; ARIMA(1,0,1)1(1,1,1) modelinin kalıntı ACF/PACF grafikleri, Ljung-Box beyaz gürültü olasılık grafiği ve olağanlık (dağılım) grafiği Şekil 6’da sırasıyla Şekil 6a, Şekil 6b, Şekil 6c şekillerinde gösterilmiştir. Bu modelin gerek ACF/PACF grafikleri, gerek beyaz gürültü olasılık grafiği, gerek de kalıntı dağılım grafiği incelendiğinde sonucun anlamlı model olduğu görülmektedir.

Tablo 4. Holt-Winters üstel düzeltme yöntemi ile günlük tahmin hata sonuçları ve belirlenen katsayıları (Error measurements and coefficients for Holt-Winters exponential smoothing forecast)

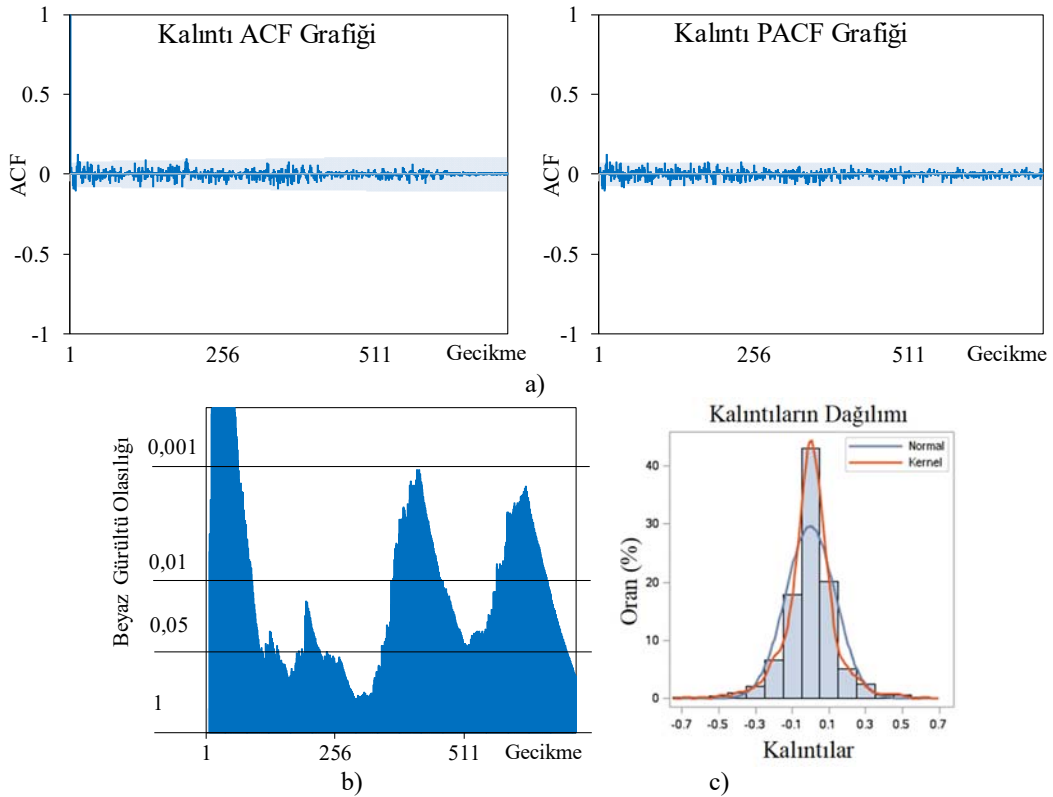
	Yöntem	Kısaltması	Parametre değerleri	MAPE (%)	R ²
Toplamsal Model	$\alpha, \beta, \gamma = 0,2$ Eğilimsiz	WAc0.2	$\alpha, \beta, \gamma = 0,2$	1962	0,1032
	optimizasyon	WAcOpt-E	$\alpha=0,9797, \gamma=0,0004$	39,5	0,8149
	Tümü optimize	WAcOpt	$\alpha=0,9916, \beta=0,02$ $\gamma=0,062$	197,4	0,7227
Çarpımsal Model	$\alpha, \beta, \gamma = 0,2$ Eğilimsiz	WMc0.2	$\alpha, \beta, \gamma = 0,2$	251,7	0,1047
	optimizasyon	WMcOpt-E	$\alpha=0,9797, \gamma=0,0004$	26,9	0,7782
	Tümü optimize	WMcOpt	$\alpha=0,9916, \beta=0,02,$ $\gamma=0,062$	27,1	0,8280

Tablo 5. ARIMA için belirlenen parametreler ve anlamlılıkları (Estimated parameters and significances for ARIMA)

Parametre	Katsayı	Standart Hata	t değeri	Olasılık, $p > t $
Kesişim	0,0007168	0,0055008	0,13	0,8963
MA(1)	-0,19921	0,02811	-7,09	<,0001
MA(2)	0,79567	0,02568	30,98	<,0001
AR(1)	0,96912	0,03884	24,95	<,0001
AR(2)	-0,18889	0,03762	-5,02	<,0001
AR(3)	-0,04335	0,02762	-1,57	0,1165

Çalışma sonrası elde edilen 96 modelin SARIMA yönteminde dört farklı ölçüt için en iyi tahmin sonuçları bulunmuş ve iki bölümde gösterilmiştir. İlk bölümde, 2011-2013 yıllarındaki günlük doğal gaz tüketim verileri ile model kurularak, en düşük MAPE ($MAPE_{2011-2013}$) ve en yüksek R^2 ($R^2_{2011-2013}$) değerleri bulunmuştur. İkinci bölümde ise, 2014 yılının doğal gaz tüketiminin en düşük MAPE ($MAPE_{2014}$) ve en yüksek R^2 (R^2_{2014}) değerleri belirlenmiştir. Seçim kriteri olarak $MAPE_{2011-2013}$ veya $R^2_{2011-2013}$ kullanıldığında, mevsimsel olmayan tek farkın modelinin (I(1)1(0)), temsil etmede daha düşük hatalı ve daha uyumlu olduğu görülmüştür (Tablo 6). Bu modellerde mevsimsellik kendi SAR ve SMA parametrelerinde göstermektedir. Fakat uzun vadeli 365 günlük tahminde GM_1 ve GM_2 olarak gösterilen bu modellerin hata değerleri %10,345 ve %8,711, R^2 değerleri ise 0,248 ve 0,090 olmuştur (Tablo 6, Şekil 7). Bu değerler uç tahminlerin olduğunu göstermektedir ve başarısızdırlar. Ancak bu modellerin sonuçları öndeki 7 gün için (01/01/2014 – 07/01/2014) incelendiğinde MAPE değerleri GM_1 ve GM_2 için sırasıyla %7,55 ve %11,65, R^2 değerleri sırasıyla 0,046 ve 0,374 olmuştur. GM_1 ve GM_2 için 7 günlük tahminde MAPE değerlerinin düşük olmasından dolayı R^2 değerlerinin yüksek olması beklenmiştir, fakat R^2 değerleri aksine düşük çıkmıştır. Bu durum serinin uyum tahmininde kötü olduğunu fakat 7 günlük tüketim değerlerinin birbirine yakın olmasından dolayı MAPE hatasının düşük olduğunu göstermektedir.

Burada gün sayısı 21 güne çıkartıldığında ise GM_1 ve GM_2 modellerinin tahmin sonuçları sırasıyla %41,69 ve %42,36 MAPE olurken, R^2 değerleri sırasıyla 0,467 ve 0,539 olmuştur. Diğer bir ifadeyle kısa dönemli tahminlerde tatmin edici sonuçlar elde edilirken, orta, orta-uzun ve uzun dönemde çok büyük sapmalar göstermişlerdir. Seçim kriteri olarak $MAPE_{2014}$ veya R^2_{2014} kullanıldığında ise mevsimsel tek fark alınmış serinin (I(0)1(1)) daha doğru tahmin sonucu ürettiği görülmüştür. AR, MA, SAR ve SMA parametrelerinin kullanımı, mevsimsellik ve kendinden önceki değer tahmin modelinde öneme sahip olduğunu göstermiştir. Etkin olan SAR(1) ve SMA(1) parametreleri kendini Şekil 3d üzerinde ACF/PACF grafiklerinde göstermektedir ve beklenen bir durumdur. Seçim kriteri olarak $MAPE_{2014}$ veya R^2_{2014} kullanıldığı GM_3 ve GM_4 modellerinin uzun vadeli 365 günlük tahminde kullanılabileceği görülmüştür. Yine GM_3 ve GM_4 modellerinin 7 günlük tahminleri MAPE üzerinden incelenmiş olup, sırasıyla %7,87 ve %8,76 değerleri bulunmuştur. GM_3 ve GM_4 modellerinin 7 günlük tahminleri R^2 üzerinden incelenecek olursa, GM_1 ve GM_2 modellerine göre daha kötü performans sergileyerek 0,002 olmuştur. Burada gün sayısı 21 güne çıkartıldığında ise GM_3 ve GM_4 modellerinin tahmin sonuçları sırasıyla %19,64 ve %23,98 MAPE olurken, R^2 değerleri sırasıyla 0,704 ve 0,596 olmuştur. Bu durum, mevsimsel fark içeren ve mevsimsel katsayı içeren modellerin, hem kısa, hem orta, hem de uzun



Şekil 6. ARIMA(1,0,1)1(1,1,1) modelinin kalıntı grafikleri a) ACF ve PACF grafiği b) Beyaz gürültü olasılık grafiği c) Kalıntı olağanlık (dağılım) grafiği.

(Residuals plots of ARIMA(1,0,1)1(1,1,1) model (a) ACF and PACF plot (b) White noise probability plot (c) Distribution of residuals plot)

vadede tahminde yetenekli olduğunu göstermiştir. Elde edilen bu dört modelin parametrelerinin anlamlılıkları Tablo 7’de verilmiştir. Tablo 7 incelendiğinde GM₁ modeli olan ARIMA(1,1,0)1(2,0,3) modelinin mevsimsel parametrelerinin anlamlı olmadığı görülmüştür. GM₂ modelinde ise SMA parametresi kaldırılması ile SAR parametresinin anlamlılığı değişmiştir. SAR parametresinin anlamlılığı her ne kadar artsa da yine de anlamsız olduğu ($p = 0,2023$) Tablo 7’de görülmüştür. 2014 yılı için, en düşük MAPE ve en yüksek R² değerini veren GM₃ modelinin SAR ve SMA parametrelerinin anlamsız olduğu görülmüştür. Bu parametrelerden anlamsız olan SAR parametresi kaldırıldığında, 2014 yılı için, ikinci en düşük MAPE ve en yüksek R² değerini veren GM₄ modeli elde edilmiş olur. Bu modelde ise tüm parametreler anlamlı olarak elde edilmiştir. GM₃ ve GM₄ modellerindeki tahminlerin yaz mevsimi boyunca uyumlu ve düşük hatalı olduğu görülmektedir. GM₃ ve GM₄ modelleri kendi aralarında karşılaştırıldığında, kış mevsiminde farkları ortaya çıkmaktadır. GM₃ modeli GM₄ modeline göre kışın daha düşük hata ile tahmini gerçekleştirmiştir. Tablo 7 üzerinde görüldüğü gibi GM₃ modelinde SAR ve SMA parametreleri anlamsız olarak belirlendiği için GM₄ modelini kullanmak gerekmektedir.

5.1.1. Günlük doğal gaz talep tahmin sonuçlarının değerlendirilmesi

(Evaluation of daily natural gas demand forecasting results)

Günlük talep tahmininde üç farklı tek değişkenli istatistik yöntemleri kullanılarak elde edilen sonuçlara göre GM₄ modeli (ARIMA(1,0,1)1(0,1,1)) 2014 yılı için en düşük

MAPE değerine (MAPE %24,57, R² 0,802), WMcOpt modeli ise en yüksek R² değerine (MAPE %27,1, R² 0,83) sahiptir. MAPE düşük, R² değeri yüksek olduğu durumda hatası düşük, uyumu yüksek modeller elde edilmiş olur. MAPE serideki genel ortalama hatayı, R² değeri ise serinin tüm değerlerde örtüşme eğilimini gösterir. MAPE ve R² değerinin düşük olduğu durumlarda keskin ve yüksek hataların olduğu görülmüştür. MAPE ve R² yüksek olduğu durumda tahmin ile gerçekleşen serilerin birbirine uyumu yüksek olmasına rağmen, seri neredeyse hiç örtüşmemektedir. Üç teknik için sonuçlar sırasıyla; Holt-Winters yöntemi için WMcOpt (MAPE %27,1, R² 0,83), ARIMA/SARIMA yöntemi için GM₄ (MAPE %24,57, R² 0,802), zaman serilerinin ayrıştırılması için GDAS (MAPE %26,23, R² 0,80) modellerinin en uygun modeller olduğu belirlenmiştir. Yazdan kışa geçerken GDAS, kıştan yaza geçerken GM₄ modelinin daha iyi tahmin gerçekleştirdiği de yine sonuçlarda görülmüştür. Günlük tahminler sene öncesindeki o senenin genel davranışını belirlemede önemli olmaktadır. GDAS, WMcOpt ve GM₄ model tahmin sonuçları da genel davranışı yansıtacak biçimde iyi bir performans sergilemişlerdir. Eğer modeller iyi performans gösterdiler ise aylık hale getirildiğinde de performansının artması beklenmektedir (Bölüm 5.2).

5.2. Günlük Temelli Veri İle Aylık Doğal Gaz Talep Tahmin Sonuçları

(Daily Basis Monthly Natural Gas Demand Forecasting Results)

Önceki bölümde günlük tahminler için modelin önüne “G” simgesi eklenmişti. Buradaki model sonuçlarının başında

Tablo 6. ARIMA/SARIMA için en iyi sonuç veren modeller ve bunların günlük tahmin hata ve uyum sonuçları
(Best results of ARIMA/SARIMA models and daily forecasting errors and coefficient of determination of these models)

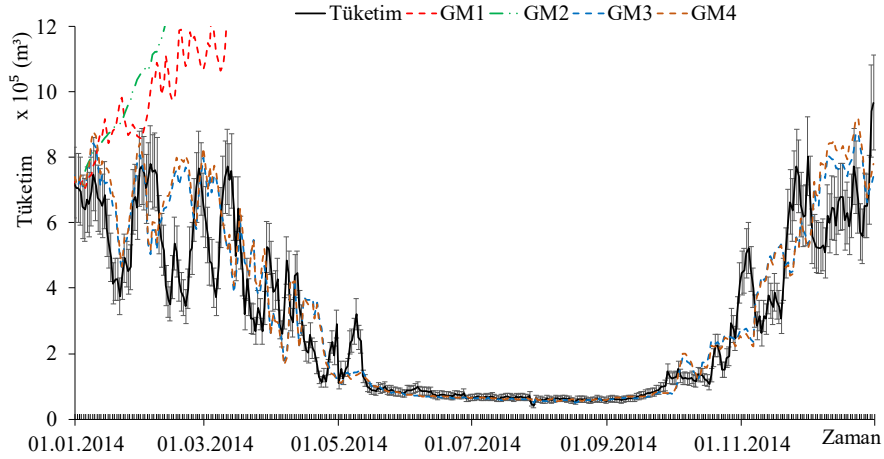
Kısaltma	Seçim Kriteri	Model	MAPE ₂₀₁₁₋₂₀₁₃	MAPE ₂₀₁₄	R ² ₂₀₁₁₋₂₀₁₃	R ² ₂₀₁₄
GM ₁	MAPE ₂₀₁₁₋₂₀₁₃	ARIMA(1,1,0)1(2,0,3) ³⁶⁵	%8,80	%10,345	0,954	0,248
GM ₂	R ² ₂₀₁₁₋₂₀₁₃	ARIMA(3,1,1)1(1,0,0) ³⁶⁵	%8,82	%8,711	0,972	0,090
GM ₃	MAPE ₂₀₁₄ , R ² ₂₀₁₄ - 1	ARIMA(1,0,1)1(1,1,1) ³⁶⁵	%12,16	%24	0,954	0,811
GM ₄	MAPE ₂₀₁₄ , R ² ₂₀₁₄ - 2	ARIMA(1,0,1)1(0,1,1) ³⁶⁵	%12,16	%24,6	0,954	0,802

Tablo 7. SARIMA modellerimin parametreleri ve anlamlılıkları
(Parameters and significances for SARIMA models)

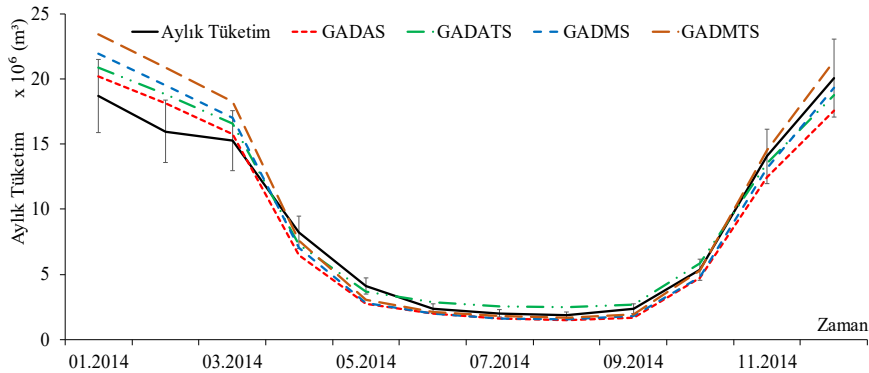
	Parametre	Katsayı	Standart Hata	t değeri	Olasılık, $p > t $
ARIMA(1,1,0)1(2,0,3) (GM ₁)	Kesişim	0,0003737	0,0038522	0,1	0,9227
	SMA(3)	0,99313	303,43418	0	0,9974
	AR(1)	0,25238	0,02085	12,1	<,0001
	SAR(2)	-0,03742	0,04113	-0,91	0,3629
ARIMA(3,1,1)1(1,0,0) (GM ₂)	Kesişim	0,000343	0,0049414	0,07	0,9447
	MA(1)	-0,27541	0,02932	-9,39	<,0001
	AR(3)	-0,0781	0,03024	-2,58	0,0098
	SAR(1)	0,05156	0,04043	1,28	0,2023
ARIMA(1,0,1)1(1,1,1) (GM ₃)	Kesişim	0,03812	0,05591	0,68	0,4954
	MA(1)	-0,28974	0,03025	-9,58	<,0001
	SMA(1)	0,99984	14,088,8	0	0,9999
	AR(1)	0,90835	0,01304	69,67	<,0001
ARIMA(1,0,1)1(0,1,1) (GM ₄)	SAR(1)	0,14821	3,01161	0,05	0,9607
	Kesişim	0,03713	0,0624	0,6	0,5518
	MA(1)	-0,28976	0,0348	-8,33	<,0001
	SMA(1)	0,56003	0,10548	5,31	<,0001
	AR(1)	0,90829	0,01495	60,74	<,0001

“GA” biçimde ek konulacaktır. Buna göre kullanılan ilk teknik zaman serilerinin ayrıştırılmasıdır. Burada günlük tahmin sonuçlarının aylık biçimde yazılarak bulunan sırasıyla GADAS, GADATS, GADMS ve GADMTS modelleri için 16,78%, 14,75%, 16,23%, 14,04% MAPE, 0,964, 0,969, 0,965, 0,973 R^2 şeklinde tahmin hataları belirlenmiştir (Şekil 8). Şekilde de GADATS ve GADMTS tahminleri %15 hata bandının içinde diğer modellere göre daha fazla kalarak düşük tahminli olduklarını göstermişlerdir. Bu modeller incelendiğinde eğilim barındırdıkları rahatlıkla görülebilmektedir. Günlük tahminlerde eğilimsiz modeller daha düşük MAPE ile tahmin ederken, aylık veride eğilim olan modeller daha düşük MAPE ile tahmini gerçekleştirmiştir. R^2 üzerinden incelendiğinde ise hem günlük hem de aylık tahminlerde eğilim olan modellerin çok az farkla da olsa daha uyumlu olduğu görülmüştür. Holt-Winters üstel düzeltme yöntemi sonuçları aylık tahminlere dönüştürüldüğünde en uygun modellerin çarpımsal modeller olduğu görülmektedir. Bunun sebebi her üç yapıda da toplamsal eş değer model ile karşılaştırıldığında (GAWAc0.2 ile GAWMc0.2, GAWAcOpt-E ile GAWMcOpt-E, GAWAcOpt ile GAWMcOpt) daha düşük MAPE ve daha yüksek R^2 değeri vermesidir (Tablo 8). Burada görülen diğer bir nokta aylık

dönüşüm sonrasında günlük tahminlere göre daha düşük MAPE ve daha yüksek R^2 değerlerinin belirlenmesidir. Günlük tahminlerdeki WMcOpt-E modeli, aylık tahmine dönüştürüldüğünde de (GAWMcOpt-E), tıpkı günlük tahminde olduğu gibi en düşük MAPE sonucunu vermiştir. Holt-Winters üstel düzeltme modellerinin sonuçları Şekil 9’da gösterilmiştir. Bu şekilde de görüleceği gibi γ değerinin aşırı değerlenmesi sonucunda modeller uç noktalara ulaşmaktadır. 2014 yılının hem başı hem de sonunda %15 güven bandına yakın olan GAWMcOpt-E modelidir. GAWAcOpt-E modeli kış aylarında çok iyi sonuç verse de yaz aylarında performansı kötü olmuştur. Yaz aylarında en iyi performans sergileyen model GAWMcOpt olmuştur. GAWAcOpt modelinin R^2 değeri GAWMcOpt modelinden daha yüksek olmasına karşın, MAPE tarafında %40 oranında daha kötü uyuma sahiptir. Bu durum kendini Şekil 9 üzerinde açık biçimde göstermektedir. Yine Şekil 9’da GAWMcOpt-E ile GAWMcOpt modellerinin yakın tahminler yaptıkları, sadece son iki ay tahminlerde farklılaştıkları görülmektedir. Eğilim etkisini son iki ayda ortaya çıkartmıştır. ARIMA yönteminde diğer tekniklerden farklı olarak iki farklı durum oluşmaktadır. Birinci durum günlük tahminler sonrasında en düşük hataya sahip modellerin aylık tüketime dönüştürülerek sonuçlarının



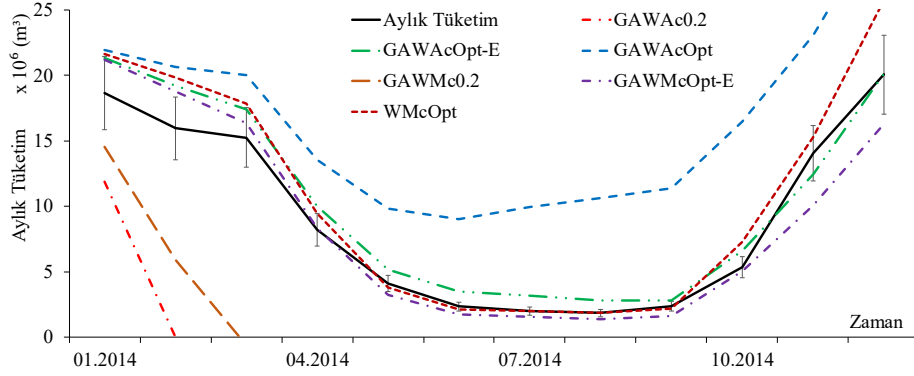
Şekil 7. SARIMA modelleri tahmin sonuçları (Forecasting results for SARIMA models)



Şekil 8. Zaman serisi ayrıştırılması yöntemi için günlük tahminin aylık tahmin haline getirilmiş sonuçları (Forecasting results of daily basis monthly forecasts for time series decomposition)

Tablo 8. Holt-Winters ile günlük tahminin aylık tahmin haline getirilmiş hata sonuçları ve belirlenen katsayıları
(Error measurements and coefficients for Holt-Winters exponential smoothing results of daily basis monthly forecasts)

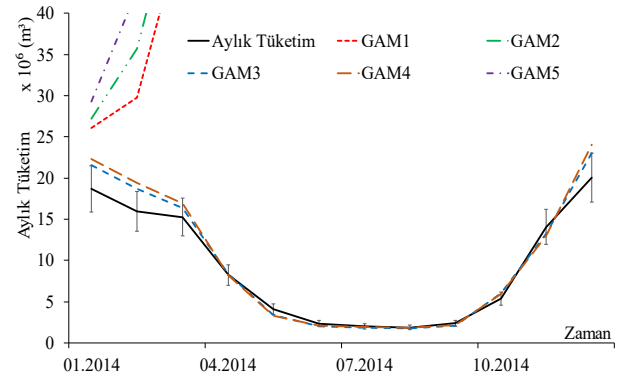
	Yöntem	Kısaltması	Parametre değerleri	MAPE (%)	R ²
Toplamsal Model	$\alpha, \beta, \gamma = 0,2$	GAWAc0.2	$\alpha, \beta, \gamma = 0,2$	1.858	0,111
	Eğilimsiz optimizasyon	GAWAcOpt-E	$\alpha=0,9797, \gamma =0,0004, \beta=0$	179,8	0,856
	Tümü optimize	GAWAcOpt	$\alpha=0,9916, \beta=0,02, \gamma=0,062$	25,74	0,972
Çarpımsal Model	$\alpha, \beta, \gamma = 0,2$	GAWMc0.2	$\alpha, \beta, \gamma = 0,2$	243,1	0,131
	Eğilimsiz optimizasyon	GAWMcOpt-E	$\alpha=0,9797, \gamma =0,0004, \beta=0$	14,36	0,990
	Tümü optimize	GAWMcOpt	$\alpha=0,9916, \beta=0,02, \gamma=0,062$	18,31	0,926

**Şekil 9.** Holt-Winters üstel düzeltme yöntemi için günlük tahminin aylık tahmin haline getirilmiş sonuçları
(Forecasting results of daily basis monthly forecasts for Holt-Winters exponential smoothing)**Tablo 9.** SARIMA modelleri için en iyi sonuç veren modeller ve bunların aylık tahmin hata ve uyum sonuçları
(Best results of SARIMA models and monthly forecasting errors and coefficient of determination of these models)

Günlük Kısaltma	Aylık Kısaltma	Model	MAPE ₂₀₁₁₋₂₀₁₃	MAPE ₂₀₁₄	R ²	R ² ₂₀₁₄
GM ₁	GAM ₁	ARIMA(1,1,0)1(2,0,3) ³⁶⁵	%1,55	%9772,62	0,999	0,218
GM ₂	GAM ₂	ARIMA(3,1,1)1(1,0,0) ³⁶⁵	%1,57	%8267,69	0,999	0,017
GM ₃	GAM ₃	ARIMA(1,0,1)1(1,1,1) ³⁶⁵	%2,86	%10,93	0,999	0,988
GM ₄	GAM ₄	ARIMA(1,0,1)1(0,1,1) ³⁶⁵	%2,85	%11,32	0,999	0,981
-	GAM ₅	ARIMA(1,1,0)1(1,0,0) ³⁶⁵	%1,53	%21337,14	0,999	0,075

gösterilmesidir. İkinci durum ise günlük tahminlerin öncelikle aylık tahminlere dönüştürülmesi ve bu tahminlerde en düşük hataya sahip modelin belirlenmesidir. Tablo 9, kalıntılar arasında ilişki olmayan modeller seçilerek yapılan aylık SARIMA tahmin sonuçlarını her iki durum için göstermektedir. Günlük tahminlerde dört farklı kritere göre en iyi performans modeller Tablo 6'da gösterilmiştir. Bu modeller aylık hale getirildiğinde ortaya çıkan tahmin hataları ve uyum değerleri Tablo 9'dadır. GAM₅ modeli (ARIMA(1,1,0)1(1,0,0)), aylık hale getirildikten sonra, 2011-2013 yılları için en düşük MAPE ve en yüksek R² değerine sahip model olarak ortaya çıkmıştır. Bu modelin AR(1) parametresi 0,25225, SAR(1) parametresi 0,06667 olarak belirlenmiştir. Tablodaki diğer ilk dört model günlük tahminde en iyi sonuç veren modellerdir. 2014 yılı aylık tahmin sonuçlarındaki durumda, günlük tahmin sonuçlarındaki gibi, GM₃ (GAM₃) ve GM₄ (GAM₄) modelleri en düşük MAPE ve en yüksek R² ile tahminlerini gerçekleştirmişlerdir. Her ne kadar GAM₃ modeli daha düşük MAPE ve daha yüksek R² sahip olsa da anlamsız

parametre barındırdığı için kullanılamamaktadır. Şekil 10 üzerinde iki modelin yakın sonuç verdiği görülmektedir.

**Şekil 10.** ARIMA/SARIMA yöntemi için günlük tahminin aylık tahmin haline getirilmiş sonuçları
(Forecasting results of daily basis monthly forecasts for ARIMA/SARIMA models)

5.2.1. Günlük temelli aylık doğal gaz talep tahmin sonuçlarının değerlendirilmesi

(Evaluation of daily basis monthly natural gas demand forecasting results)

Günlük temelli aylık talep tahmininde üç farklı tek değişkenli istatistik yöntemleri kullanılarak elde edilen sonuçlara göre GM₄ modeli (ARIMA(1,0,1)₁(0,1,1)) 2014 yılı için en düşük MAPE değerine (MAPE %11,32, R² 0,981), GAWMcOpt-E modeli ise en yüksek R² değerine (MAPE %13,36, R² 0,990) sahiptir. Üç teknik için en tahminde kullanılabilir değerlere sahip modeller sırasıyla; TSD için GADMTS modeli 14,04% MAPE, 0,973 R²; HWES için GAWMcOpt-E modeli %13,36 MAPE, 0,990 R², ARIMA/SARIMA yöntemi için GM₄ modeli %11,32 MAPE, 0,981 R² ile belirlenmiştir. Aylık tahmin sonuçları genel olarak günlük tahminlere göre daha düşük MAPE ve daha yüksek R² değerleri elde etmiştir. Yoğun veri üzerinden elde edilen tahminler her ne kadar fazla sapma gösterse de değer aralığındaki zaman boyutu genişledikçe hata oranının düştüğü görülmektedir. Tahminler tüm yılda düşünüldüğünde toplam 2014 yılı için 110.211.986 m³ olan tüketime en yakın toplam tüketimi GDMS %1,97 hata ile 112.385.846,36 m³ olarak bulmuştur. İkinci sırada GAWMcOpt-E modeli -%4,19 hata ile tahminde bulunurken, ardından GDAS modeli -%4,97 hata ile takip etmektedir. Görüldüğü tüm 2014 yılı toplam tahmininde ARIMA/SARIMA yöntemlerinin önüne TSD ve HWES yöntemleri geçmiştir. Ancak veri kısımlı zaman bağlı olarak arttıkça daha detaylı tahmin modelleri daha doğru sonuç üretmiştir.

5.3. Sonuçların Tartışılması (Discussion of Results)

Bu çalışmada kullanılan üç teknik; rüzgâr hızı, sıcak su ve rüzgâr enerjisi tahmininde kullanılmıştır. Örneğin Prema ve Rao Holt-Winters, TSD ve ARIMA yöntemlerini kullanarak gün öncesi tahmini 10 dakikalık veriler üzerinden gerçekleştirmiştir [17]. Üç aylık tahmin sonuçlarının gün öncesi tahminde 10 dakikalık veriler üzerinden gerçekleştirmesi sebebiyle 144 adet veri tahmin bir kerede gerçekleştirilmiştir. Bizim çalışmamızda bu değer 365'dir ve bizim çalışmamıza en yakın çalışma budur. Yine aynı çalışmada bir yıldan 1 aya kadar farklı zaman periyotlarında veri ile model kurulduğunda gün öncesi 144 noktanın tahminindeki en düşük hata ARIMA, Holt-Winters ve TSD yöntemleri için sırasıyla %23,26, %28,63 ve %18,24 MAPE olmuştur. Bu değerler bizim çalışmamızda en düşük sırasıyla ARIMA, Holt-Winters ve TSD için %23,68, %26,9 ve %26,23 MAPE olmuştur. Prema ve Rao 144 adet tahmin yaparken, bizim çalışmamızda 365 adet tahmin yapılmış olup, tahmin sayısı bakımından 2,5 kat fazla tahmin yapılmasına rağmen MAPE hatalarının Prema ve Rao'nun çalışmasına yaklaşık olduğunu göstermiştir. Ayrıca Lewis tarafından gösterilmiş olan MAPE sonuçlarına göre [39], tahmin değerlerimizin %20 çok az üzerinde olması sebebiyle iyi ve makul arasında olduğu söylenebilir. Diğer bir çalışmada Gelažanskas ve Gamage 2006 ve 2007 yıllarının 1 saatlik yoğunlukta verisi kullanılarak sıcak su tüketimini gün öncesinde saatlik bazda (24 değer) TSD, Holt-Winters

ve ARIMA yöntemleri ile tahmin etmişlerdir ve sırasıyla 0,863, 0,811, 0,872 R² değerlerini bulmuşlardır [18]. Bizim çalışmamızda aynı yöntemler sırasıyla en yüksek 0,805, 0,828 ve 0,907 R² değerleriyle günlük, 0,973, 0,990 ve 0,990 R² değerleriyle aylık bazda tahminleri gerçekleştirmiştir. Wu ve Peng ise rüzgâr gücünü tahminde ARIMA yöntemini günlük veride kullanmışlar ve model oluşturmada 219 gün, tahminde 82 gün verisi ile çalışmışlardır [19]. Çalışmalarında tahmin hatası %38,57 MAPE olarak bulunmuştur ve bizim çalışmamızda modellerimizin daha iyi performans sergilediği görülmektedir. Akpinar ve Yumuşak ise çalışmalarında sadece aylık veri ile doğal gaz tüketimini tahmin etmiştir [22]. Yaptıkları çalışmada ARIMA(1,1,1)₁(0,1,1)₂ modelinde %12,9 ile en düşük MAPE ve 0,956 ile en yüksek R² değerinin bulunduğunu göstermişlerdir. Bizim çalışmamızda aylık tahminlerdeki MAPE ve R² değeri sırasıyla %11,32 ve 0,981 olmuştur. Diğer bir ifadeyle MAPE hatasının %10 oranında düştüğü görülmüştür. Üç çalışma da bu çalışma ile karşılaştırıldığında, yöntem olarak kullanılan üç yöntemin de yoğun veri ile doğal gaz tüketim tahmini için uygunluğunu göstermiştir. Karar vericiler, yıl öncesinde tahminlerini yaparken hesap tablosu gibi basit araçlar ile TSD tekniği üzerinden hızlıca tahminlerini gerçekleştirebilirler. Şayet HWES yöntemi kullanılmak isteniyorsa, modeller hakkında bilgi gerektirmekte olup, hesap tabloları için geliştirilmiş ücretli ve/veya ücretsiz araçlar yardımıyla modeller oluşturulabilmektedir. ARIMA/SARIMA tekniklerinin kullanımında ise karar vericiler ancak bu konuda bir uzmandan yardım alarak ve SAS, R, JMP, NCSS, MATLAB gibi yazılımlar ile gerçekleştirebilir.

6. SONUÇLAR (CONCLUSIONS)

Evsel ve düşük tüketimli ticari kullanıcılar doğal gazı pişirme ve sıcak su kaynağı olarak sürekli kullanırken, ısıtma amacıyla kış mevsimi ve civarında kullanmaktadır. Diğer bir ifadeyle doğal gaz tüketimi mevsimsellik içermektedir. Bu çalışmada da Türkiye'deki Sakarya ili için doğal gaz talep tahmini orta vadede günlük veri ile gerçekleştirilmiştir. 4 yıllık verinin ilk üç yılı olan 2011-2013 arasında günlük tüketimler 2014 yılını tahminde model oluşturmada kullanılmıştır. Bu tüketim davranışı genelde yaz mevsiminde düşük, kış mevsiminde yüksek ve kışa girişte artan ve kıştan çıkışta düşen biçimdedir. Bu çalışmada 365 günlük tahmini bir kerede gösterebilen modeller üzerine çalışılmıştır. Bu modellerin özelliği ise istatistiksel zaman serilerini kullanmalarıdır. Bu çalışmada tek değişkenli istatistik zaman serileri tahmin yöntemlerinden mevsimsellik içerebilen tekniklerden TSD, Holt-Winters üstel düzeltme ve ARIMA/SARIMA kullanılmıştır. Çalışma neticesinde en düşük MAPE sırasıyla ARIMA/SARIMA, Holt-Winters ve TSD için %24,60, %26,9 ve %26,23 olarak bulunmuştur. Sonuçlar literatürde en benzer çalışma olan [17] ile kıyaslandığında MAPE değerlerine yakın olduğu tespit edilmiştir. [17] çalışması, 144 noktanın tahminindeki en düşük MAPE değerini ARIMA, Holt-Winters ve TSD yöntemleri için sırasıyla %23,26, %28,63 ve %18,24 olarak bulmuştur. İyi bilinen bu teknikler karar vericiler tarafından

rahatlıkta bir karar destek sistemi ile tahminde kullanılabilir. Böylece kendi yöntemleri ile yaptıkları tahminlerin tutarlılığını sınavabilirler. Çalışmadaki sonuçlara göre hesaplama karmaşıklığı yüksek olan modellerin hata oranı daha düşüktür. Buna göre hata oranı en düşük SARIMA, sonrasında Holt-Winters ve en son zaman TSD'dir. Gelecek çalışmalarda bağımsız değişkenlerin modele eklenmesi, yapay zekâ algoritmaları, doğrusal regresyon ve ARIMAX yöntemlerinin kullanılmasının tahmini iyileştirebileceği düşünülmektedir. Bu iki durum için (veri seti değişikliği ve kullanılan yöntem değişikliği) tahminlerin iyileşebileceği öngörülmektedir.

TEŞEKKÜR (ACKNOWLEDGEMENT)

AGDAŞ – Adapazarı Gaz Dağıtım Anonim Şirketine veri hazırlamadaki destekleri için ve çalışmaya desteğinden dolayı Sayın Kevser Ovaz Akpınar 'a teşekkür ederiz

KAYNAKLAR (REFERENCES)

1. M. Özcan and M. Yıldırım, The Impact of Capital Subsidy Incentive on Renewable Energy Deployment in Long-Term Power Generation Expansion Planning, *Sakarya University J. Comput. Inf. Sci.*, (1)3, 1–19, 2018.
2. X. Wang, D. Luo, J. Liu, W. Wang, and G. Jie, Prediction of Natural Gas Consumption in Different Regions of China Using a Hybrid MVO-NNGBM Model, *Math. Probl. Eng.*, 2017, 1–10, 2017.
3. M. Akpınar, M. Adak, and N. Yumuşak, Day-Ahead Natural Gas Demand Forecasting Using Optimized ABC-Based Neural Network with Sliding Window Technique: The Case Study of Regional Basis in Turkey, *Energies*, 10 (6), 781, 2017.
4. F. Scarpa and V. Bianco, Assessing the Quality of Natural Gas Consumption Forecasting: An Application to the Italian Residential Sector, *Energies*, 10 (11) 1879, Nov. 2017.
5. O. Laib, M. T. Khadir, and L. Mihaylova, A Gaussian Process Regression for Natural Gas Consumption Prediction Based on Time Series Data, in 2018 21st International Conference on Information Fusion (FUSION), 2018, 55–61.
6. N. Wei et al., Short-Term Forecasting of Natural Gas Consumption Using Factor Selection Algorithm and Optimized Support Vector Regression, *J. Energy Resour. Technol.*, 141 (3), 032701, Oct. 2018.
7. I. Naim and T. Mahara, Comparative Analysis of Univariate Forecasting Techniques for Industrial Natural Gas Consumption, *Int. J. Image, Graph. Signal Process.*, 10 (5), 33–44, May 2018.
8. M. Akpınar and N. Yumuşak, Naïve forecasting of household natural gas consumption with sliding window approach, *TURKISH J. Electr. Eng. Comput. Sci.*, 25 (1), 30–45, 2017.
9. Y. E. Özger, M. Akpınar, Z. Musayev, and M. Yaz, Electrical Load Forecasting Using Genetic Algorithm Based Holt-Winters Exponential Smoothing Method, *Sakarya University J. Comput. Inf. Sci.*, 2 (2), 108–123, Aug. 2019.
10. H. A. Es, F. Y. Kalender, and C. Hamzaçebi, Forecasting the Net Energy Demand of Turkey by Artificial Neural Networks, *Journal of the Faculty of Engineering and Architecture of Gazi University*, 29 (3), 495-504, 2014.
11. F. Jiang, X. Yang, and S. Li, Comparison of Forecasting India's Energy Demand Using an MGM, ARIMA Model, MGM-ARIMA Model, and BP Neural Network Model, *Sustainability*, 10 (7), 2225, Jun. 2018.
12. M. Fırat, M. A. Yurdusev, and M. Mermer, Monthly Water Demand Forecasting by Adaptive Neuro-Fuzzy Inference System Approach, *Journal of the Faculty of Engineering and Architecture of Gazi University*, 23 (2), 449–457, 2008.
13. C. Hamzaçebi and F. Kutay, Electric Consumption Forecasting of Turkey Using Artificial Neural Networks Up to Year 2010, *Journal of the Faculty of Engineering and Architecture of Gazi University*, 19 (3), 227–233, 2004.
14. S. Rehman, Y. Cai, R. Fazal, G. Das Walasai, and N. Mirjat, An Integrated Modeling Approach for Forecasting Long-Term Energy Demand in Pakistan, *Energies*, 10 (11), 1868, Nov. 2017.
15. M. Bulut and B. Başoğlu, Development of a Hybrid System Based on Neural Networks and Expert Systems for Short-Term Electricity Demand Forecasting, *Journal of the Faculty of Engineering and Architecture of Gazi University*, 32 (2), 575-583, 2017.
16. D. Šebalj, J. Mesarić, and D. Dujak, Predicting Natural Gas Consumption – A Literature Review, in *Central European Conference on Information and Intelligent Systems 2017*, 2017, 293–300.
17. V. Prema and K. U. Rao, Time series decomposition model for accurate wind speed forecast, *Renewables Wind. Water, Sol.*, 2 (1), 18, Dec. 2015.
18. L. Gelažanskas and K. Gamage, Forecasting Hot Water Consumption in Residential Houses, *Energies*, (8) 11, 12702–12717 2015.
19. Q. Wu, C. Peng, Wind Power Generation Forecasting Using Least Squares Support Vector Machine Combined with Ensemble Empirical Mode Decomposition, Principal Component Analysis and a Bat Algorithm, *Energies*, (9) 4, 261, 2016.
20. Ö. Demirel, A. Kakilli, and M. Tektaş, Electric Energy Load Forecasting Using ANFIS and ARMA Methods, *Journal of the Faculty of Engineering and Architecture of Gazi University*, 25 (3), 601–610, 2010.
21. R. A. Yaffee and M. McGee, *Introduction to Time Series Analysis and Forecasting: With Applications of SAS and SPSS*, 1st ed. Orlando, FL: Academic Press, 2000.
22. Stephen A. DeLurgio, *Forecasting principles and applications*. Boston: Irwin McGraw-Hill, 1998.
23. M. Akpınar and N. Yumuşak, Year Ahead Demand Forecast of City Natural Gas Using Seasonal Time Series Methods, *Energies*, 9 (9), 727, Sep. 2016.

24. S. Makridakis, S. C. Wheelwright, and H. R. J., *Forecasting : methods and applications*. New York: John Wiley, 2008.
25. S. Delurgio and C. Bhame, *Forecasting Systems for Operations Management*. New York: Irwin Professional Pub, 1991.
26. H. Ma and Y. Wu, 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, WMWA 2009*, 2009, 91–94.
27. P. R. Winters, Forecasting Sales by Exponentially Weighted Moving Averages, *Manage. Sci.*, 6 (3), 324–342, Apr. 1960.
28. G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time Series Analysis*, 5th ed. Hoboken, NJ: John Wiley & Sons, Inc., 2008.
29. M. Valipour, Long-term runoff study using SARIMA and ARIMA models in the United States, *Meteorol. Appl.*, 22 (3), 592–598, Jul. 2015.
30. R. Burtiev, F. Greenwell, and V. Kolivenko, Time Series Analysis of Wind Speed and Temperature in Tiraspol, Moldova, 12 (1), 23–33, 2013.
31. S. I. Inc., *SAS/ETS® 13.2 User's Guide*. North Carolina: SAS Institute Inc., 2014.
32. P. S. P. Cowpertwait and A. V. Metcalfe, *Introductory Time Series with R*. New York: Springer, 2009.
33. J. D. Hamilton, *Time Series Analysis*, Prentice Hall New Jersey 1994, vol. SFB 373, no. Chapter 5, 837–900, 1994.
34. W. W. S. Wei, *Time Series Analysis: Univariate And Multivariate Methods*, 2nd Ed. Boston: Pearson, 2006.
35. G. Wypych, Solvent Use in Various Industries: Asphalt Compounding, in *Handbook of Solvents: Second Ed.*, vol. 2, Elsevier, 2014, 13–14.
36. D. A. Dickey and W. A. Fuller, Distribution of the Estimators for Autoregressive Time Series With a Unit Root, *J. Am. Stat. Assoc.*, 74, 366, 427, Jun. 1979.
37. E. Said and D. A. Dickey, Testing for unit roots in autoregressive-moving average models of unknown order, *Biometrika*, 71 (3), 599–607, 1984.
38. P. C. B. Phillips and P. Perron, Testing for a unit root in time series regression, *Biometrika*, (75) 2, 335–346, 1988.
39. C. D. Lewis, B. Green, and K. Sevenoaks, *Industrial and business forecasting methods: A practical guide to exponential smoothing and curve fitting*, First. London: Butterworth-Heinemann, 1982.

