

**T.C.
SAKARYA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

**EVSEL DOĞAL GAZ KULLANIMINDA KONJONKTÜR
ETKİSİNDEN ARINDIRILMIŞ TÜKETİMİN ARIMA VE
ÇOKLU REGRESYON YÖNTEMLERİ İLE TAHMİNİ**

YÜKSEK LİSANS TEZİ

Mustafa AKPINAR

**Enstitü Anabilim Dalı : BİLGİSAYAR VE BİLİŞİM
MÜHENDİSLİĞİ**
Tez Danışmanı : Prof. Dr. Nejat YUMUŞAK

Ocak 2014

T.C.
SAKARYA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

EVSEL DOĞAL GAZ KULLANIMINDA KONJONKTÜR
ETKİSİNDEN ARINDIRILMIŞ TÜKETİMİN ARIMA VE
ÇOKLU REGRESYON YÖNTEMLERİ İLE TAHMİNİ

YÜKSEK LİSANS TEZİ

Mustafa AKPINAR

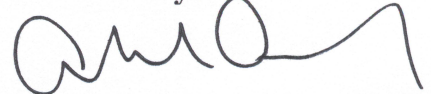
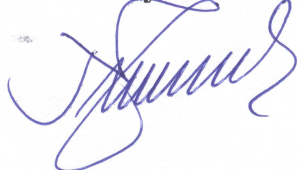
Enstitü Anabilim Dalı : BİLGİSAYAR VE BİLİŞİM
MÜHENDİSLİĞİ

Bu tez 03 / 01 /2014 tarihinde aşağıdaki jüri tarafından Oybirliği ile kabul edilmiştir.

Doç. Dr. Mehmet FIRAT
Jüri Başkanı

Prof. Dr. Nejat YUMUŞAK
Üye

Doç. Dr. Celal ÇEKEN
Üye



İÇİNDEKİLER

TEŞEKKÜR.....	ii
İÇİNDEKİLER	iii
SİMGELER VE KISALTMALAR LİSTESİ.....	v
ŞEKİLLER LİSTESİ	vi
TABLolar LİSTESİ	vii
ÖZET.....	viii
SUMMARY	ix
BÖLÜM 1.	
GİRİŞ	1
1.1. Enerji Tahminine Yönelik Literatür Çalışmaları.....	2
1.2. Çalışmanın Ana Hatları.....	3
BÖLÜM 2.	
TAHMİN YÖNTEMLERİ VE TANIMLARI.....	5
2.1. İstatistik Bilgisi	5
2.1.1. Tek değişkenli istatistik	5
2.1.2. İki değişkenli istatistik.....	6
2.2. Zaman Serileri Analizi	7
2.2.1. Trendin belirlenmesi	8
2.2.2. Konjonktür etkisi	10
2.2.3. Sezonsallık (Mevsimsellik) etkisi.....	10
2.2.4. Düzensiz hareketler (Kalıntılar)	11
2.3. Zaman Serileri.....	11
2.3.1. Hareketli ortalamalar	11
2.3.2. Otoregresif modeller	12

2.3.2.1. Otokorelasyon katsayısı	12
2.3.2.2. Kısmi otokorelasyon katsayısı.....	13
2.3.2.3. Birleşik otoregresif hareketli ortalama yöntemi (Box-Jenkins)	15
2.4. Çoklu Regresyon Yöntemi.....	16
2.4.1. Çoklu regresyon yönteminin güvenilirliği.....	17
2.5. Hata ve Ortalama Mutlak Yüzdesel Hata.....	17
BÖLÜM 3.	
DOĞAL GAZ TÜKETİM VERİLERİ.....	19
3.1. Tüketim Verisi	19
3.2. İklim Verisi	20
3.3. Tüketici Verisi	22
3.4. Takvimsel Veriler	23
BÖLÜM 4.	
DOĞAL GAZ TÜKETİMİNİN TAHMİNİ VE SONUÇLARI	25
4.1. Konjonktür Etkisinin Kaldırılması.....	25
4.2. Hata Türleri ve Hata Oranları	27
4.3. ARIMA Yöntemi İle Tahmin.....	29
4.4. Çoklu Regresyon Yöntemi ile Tahmin.....	31
4.5. Gerçekleşen Tüketimler ile Tahmin Sonuçlarının Karşılaştırılması ...	36
4.5.1. Tüketim tahmininin karşılaştırılması.....	37
4.5.2. Kalıntıların karşılaştırılması	39
4.5.3. Ortalama yüzdesel hata oranlarının karşılaştırılması.....	41
BÖLÜM 5.	
SONUÇLAR VE ÖNERİLER	45
KAYNAKLAR.....	47
ÖZGEÇMİŞ	49

SİMGELER VE KISALTMALAR LİSTESİ

AR	: Otoregresyon
ARIMA	: Birleşik otoregresif hareketli ortalamalar
BBS	: Bağımsız birim sayısı
c_k	: Otokovaryans katsayısı
MA	: Hareketli ortalamalar
OMYH	: Ortalama mutlak yüzdesel hata
r_k	: Otokorelasyon katsayısı
X	: Bağımsız değişken
Y	: Bağımlı değişken
Y'	: Bağımlı değişken tahmini

ŞEKİLLER LİSTESİ

Şekil 2.1. Zaman serisinin ayrıştırılması [20,22]	8
Şekil 4.1. Mevsimsel etkiden arındırılmış modellerin oluşturulması.....	26
Şekil 4.2. Tüketim aralıklarında model verilerinin frekansları [20,22]	27
Şekil 4.3. Birleştirilmiş modelin oluşturulması.....	28
Şekil 4.4. Modelleme ve sonuç adımları [20,22].	29
Şekil 4.5. Model 1 için en düşük hataya sahip tüketim tahminleri [20].....	31
Şekil 4.6. Model 1 için gerçekleşme ve çoklu regresyon tahmini [22].....	36
Şekil 4.7. Tüketim tahmini sonuçları grafiği [20,22].....	38
Şekil 4.8. Tahmin kalıntıları grafiği [20,22]	40
Şekil 4.9. Örnek zaman aralığında detaylandırılmış kalıntı grafiği	41
Şekil 4.10. Ortalama yüzdesel hata grafiği	42
Şekil 4.11. Mutlak hata oranları grafiği	43

TABLolar LİSTESİ

Tablo 2.1. Otokorelasyon katsayısının hesaplanması	13
Tablo 3.1. Gerçekleşen saatlik örnek verisi	21
Tablo 3.2. Tahmin günlük örnek verisi	22
Tablo 3.3. Tüketici örnek verisi	23
Tablo 3.4. Takvimsel Olaylar Örnek Verisi.....	24
Tablo 4.1. Tüm modeller için en düşük 5 azaltılmış hata [20].	30
Tablo 4.2. Model 1 için tüm bağımsız değişkenler ile çoklu regresyon sonuçları....	32
Tablo 4.3. Tüm bağımsız değişkenler ve etkisi düşük olan bağımsız değişkenler çıkartılması ile bulunan güven düzeyleri	34
Tablo 4.4. Çoklu regresyon hata oranları [22].	36
Tablo 5.1. Tahmin sonuçları	46

ÖZET

Anahtar kelimeler: Doğal gaz talep tahmini, ARIMA, zaman serileri, istatistiksel analiz, çoklu regresyon, konjonktür etkisi

Doğal gaz enerji sektöründe önemli yer tutmaktadır. Doğal gazın ısınma, soğutma, buhar üretimi, kojenerasyon gibi geniş kullanım alanları bulunmaktadır. Sanayide imalat, buhar ve kojenerasyon için kullanılan doğal gaz, evsel tüketimde pişirme, ısınma ve sıcak su ihtiyaçlarını karşılamaktadır. Farklı müşteri türlerinin tüketimi de farklıdır. Evsel tüketicilerin mevsimlere bağlı kışın artan yazın azalan tüketimleri bulunurken sanayide bu durum söz konusu olmamaktadır.

Bu çalışmada evsel doğal gaz tüketimi konjonktürel bileşenin etkisi kaldırılarak ARIMA ve çoklu regresyon yöntemleri kullanılarak tahmin edilmeye çalışılmıştır. Konjonktür bileşenin etkisinden arındırmak için mevsimselliği eş değer olan aylar bir araya getirilerek 6 farklı model oluşturulmuştur. Oluşturulan bu modeller bağımsız değişken gerektirmeyen ARIMA yöntemi ve bağımsız değişken gerektiren çoklu regresyon yöntemi ile tahmin edilmiştir. Bu tahminlerden 2012 yılı için yapılan tahminler birleştirilerek (Birleştirilmiş Model), konjonktür etkisi kaldırılmamış (Model 7) tahminler ile karşılaştırılmıştır.

ARIMA yöntemi ile Birleştirilmiş Modelde OMYH (ortalama mutlak yüzdesel hata) %8,48 olurken Model 7' de %8,67 olmuştur ve Birleştirilmiş Model, Model7' ye göre %2 oranında hatayı azaltmıştır. Çoklu regresyon yönteminde ise Birleştirilmiş Model OMYH %14,38 olurken, Model 7' de %55,10 olmuştur ve Birleştirilmiş Model, Model 7' ye göre yaklaşık %74 oranında hatayı azaltmıştır. ARIMA yöntemi her iki model türünde de çoklu regresyon tahminlerine göre daha düşük hata oranına sahip olmuştur.

CYCLE REMOVED HOUSEHOLD NATURAL GAS CONSUMPTION PREDICTION WITH ARIMA AND MULTIPLE REGRESSION

SUMMARY

Key Words: Natural gas demand estimation, ARIMA, time series, statistical analysis, multiple regression, cycle effect

Natural gas has a wide usage range like heating, processing, cooling, steam and cogeneration. For example in factories it is used for processing, steam and cogeneration. In household, natural gas is used for cooking, and heating water and house. Consumptions of different types of consumers are also different. Household consumption has a strong relationship with seasonality. It increases in winter and decreases in summer. Unlike household consumption the industrial consumption does not change from season to season.

In this study, ARIMA and multiple linear regression methods were used for estimation of gas consumption and effects of cycle were examined. To examine the cycle, the data were divided into 6 pieces. Each piece of data was called as "Model". The data which is included in these models were gathered from the same months of successive years. In addition a new model named, Model 7 was created by using all the data. The Merged Model which was created by using predictive values of models for 2012 and Model 7 were compared.

In ARIMA method, mean absolute percent errors (MAPE) of Merged Model and Model 7 were %8,48 and %8,67 respectively. Merged Model is reduced the error rate %2 from Model 7. In multiple regression MAPE of Merged Model and Model 7 were 14.38% and 55.10% respectively. Merged Model is reduced the error rate nearly %74 from Model 7. All of the two model types, ARIMA method MAPE's are become lower than multiple regression predictions.

BÖLÜM 1. GİRİŞ

Tahminde bulunmak yönetimin karar verme faaliyetlerinin ayrılmaz bir parçasıdır. Yönetim karar verirken şansa olan bağıllığını azaltmaya çalışır ve bilimsel tahmin ihtiyacı ortaya çıkar. Tahminlemede; planlama, sonuç istendiği gibi giderse elde edilecek sonuçlar ve kaynak gereksinimlerini belirlemek önemli rol oynar. Üç çeşit tahmin tipi vardır. Bunlar kısa, orta ve uzun dönemli tahminlerdir. Çalışmaya göre bu tahmin türleri değişmektedir. Günümüzde doğru ve hızlı tahmin yapmak için istatistik hesaplamaları bilgisayar üzerinde yapılmaktadır [1].

Literatürde geleceği tahmin modelleri nicel ve nitel olmak üzere iki kısımda incelenmektedir [2]. Nicel kısım da kendi içerisinde İlişkiye Dayanan Tahmin Teknikleri” ve “Zaman Serileri Analizleri” olmak üzere iki kısımda incelenmektedir [2,3]. İlişkiye Dayanan Tahmin Tekniği (Regresyon Korelasyon Analizi), matematiksel fonksiyon ile bir değişkenin tahminine dayanmaktadır. Zaman Serileri Analizinde ise değişkenin yakın geçmişi ve eğilimi izlenerek tahminde bulunulur [3]. Regresyon analizinde tahmin edilecek değişken, tahmin edilecek değişken dışındaki değişkenlere (bağımsız değişken) bağlı olmaktadır. Böylece hangi bağımsız değişkenlerin, tahmin edilecek değişken için artış veya azalış etkisi ortaya çıkacaktır [1,3,4,5]. Zaman serileri analizinde ise tahmin, geçmiş dönem verilerinin çeşitli yöntemler ile tahminde kullanılmasıdır ve sadece tahmin edilecek değişkenin zamana bağlı olması yeterlidir [1,5,6,7].

Doğal gaz kullanımı son yıllarda bütün dünyada önemli oranda artmaktadır. Bu artış beraberinde sektörde gelişimi getirmektedir. Bu gelişim enerji talebinde artışın dengelenmesinde doğal gazın kullanılmasından kaynaklandığı gibi (doğal gaz çevrim santralleri), çevresel etkilere de bağlıdır [8,9,10]. Doğal gaz kullanımının artması son

yıllarda sorunları da beraberinde getirmektedir. Türkiye' ye doğal gaz Rusya, Türkmenistan, İran ve Azerbaycan' dan boru gazı ile, Cezayir ve Nijerya' dan ise LNG (Liquid Natural Gas – Sıvılaştırılmış Doğal Gaz) olarak girmektedir. 2012 yılında tedarikçi ile transit ülkeler arasındaki sorunlar ve teknik problemlerden dolayı kış aylarında bazı dönemlerde gaz arzında sıkıntı olmuştur [10]. Bu sıkıntı enerji sektörüne yansımış, doğal gaz elektrik dönüşüm santralleri üretimini azaltmıştır ve elektrik kesintileri meydana gelmiştir. Burada tüketilecek doğal gaz miktarının yanlış tahmini evsel tüketiciler dahil tüm sektörleri etkilemiştir.

1.1. Enerji Tahminine Yönelik Literatür Çalışmaları

Enerji tahmini ile ilgili ARIMA yöntemi kullanılarak bazı çalışmalar yapılmıştır.

M. Blanchard ve G. Desrochers enerji sektöründe ARIMA yöntemini kullanan ilk çalışmalardandır ve ARIMA yöntemini kullanarak rüzgar hızını tahmin etmeye çalışmışlardır. Saatlik rüzgar hızı kullanarak yaptıkları çalışmada iyi sonuçlar almışlardır [11].

M. T. Hagan ve S. M. Behr kısa dönem elektrik tüketiminin tahmini için ARIMA, transfer fonksiyonu ve doğrusal olmayan model kullanmıştır. ARIMA yöntemini doğrusal olmayan modele çok yakın sonuçlar vermiştir ve transfer fonksiyonundan daha düşük hataya sahip olmuştur [12].

L. Kamal ve Y. Z. Jafri saatlik rüzgar hızını ARIMA yöntemi ile tahmin etmiştir. Bir yıldan daha fazla verinin tahminlerde doğruluğu arttırdığını göstermiş, kısa dönem tahminlerinde yaklaşımın uygun olduğunu belirtmiştir [13].

R. Burtiev ve arkadaşları rüzgar hızına ve sıcaklığına zaman serileri analizini yaparak 2012 yılındaki rüzgar hızını ve sıcaklığını tahmin etmiştir. Winters ve ARIMA yöntemlerinin kullanıldığı çalışmada yöntemler için en uygun parametreler bulunmuştur [14].

Enerji tahminine yönelik çoklu regresyon yöntemi kullanılarak da bazı çalışmalar yapılmıştır.

Brian E. P., ve MO-YUEN C. elektrik tüketimi ve şebeke yükünü açıklamışlardır ve tüketim tahmininde çoklu regresyon tekniğini etkinliğini göstermiş ve belirtmişlerdir [15].

Perry C. kısa dönemli elektrik tüketimini çoklu regresyon yöntemi ile tahmin etmiştir. Tahmin için modeli kış ve yaz olarak bölmüştür. Kısa dönemli tahminlerde iki saatlik tahminlerde doğru sonuç verdiğini göstermiştir [16]. Hong T. ve arkadaşları takvimsel olayları da kullanarak evsel kullanıcıların elektrik tüketimini saatlik olarak tahmin etmiştir [17].

Ediger V.S. ve arkadaşları Türkiye’ de fosil kaynakların üretimini ile alakalı kübik regresyon ve ARIMA yöntemlerini uygulamıştır. Uygulanan yöntemlerle Türkiye’nin fosil kaynaklı yakıtlarının üretim ve tüketim tahminini 2039 yılına kadar yapmıştır [18].

Bu tez çalışmasında birleşik otoregresif hareketli ortalama yöntemi (ARIMA) ve çoklu regresyon yöntemi kullanılarak doğal gaz tüketimi tahmini konjonktür etkili ve konjonktür etkisi arındırılarak tahmin edilmiştir. ARIMA yönteminde sadece tüketim verisi yeterli olurken, çoklu regresyon yönteminde bağımsız değişkenlere ihtiyaç duyulmuştur. Konjonktürel etkiden arındırılmış modelin, konjonktür etkisi olan modele göre hem ARIMA hem de çoklu regresyon yöntemine göre daha iyi sonuç verdiği görülmüştür.

1.2. Çalışmanın Ana Hatları

İkinci bölümde, temel istatistik bilgisi verilmiş, daha sonra zaman serisinin tanımı yapılmıştır. Zaman serilerinde kullanılan ARIMA yöntemi ile ilişkiye dayanan tahmin yöntemi olan çoklu regresyon tanıtılmıştır.

Üçüncü bölümde, doğal gaz tüketimi, tahminde kullanılan değişkenlerin grupları ve bu değişkenlerin hangi yöntemler ile nerelerden alındığı gösterilmiştir. Değişkenler burada detaylandırılmıştır.

Dördüncü bölümde, konjonktür etkisinin kaldırılması anlatılmış ve tahmin yöntemlerinin uygulaması yapılmıştır. Sonuçlar ve tahmin hataları incelenmiştir ve konjonktür etkisi kaldırılmamış durum ile arasındaki farklar irdelenmiştir.

Beşinci bölümde, tez içeriği ana hatlarıyla belirtilip sonuçlar değerlendirilerek sonraki çalışmalar için öneriler getirilmiştir.

BÖLÜM 2. TAHMİN YÖNTEMLERİ VE TANIMLARI

Tahmin etmek, eldeki verileri kullanarak gerçekleşmemiş durumun belirlenmesidir. Bu bölümde temel istatistik bilgisi, sonrasında ise kantitatif (nicel) yöntemlerden “Zaman Serileri Analizi” ve “İlişkiye Dayanan – Nedensel Tahmin Teknikleri” hakkında bilgiler verilecektir.

Şartlarda çok büyük değişim olmayacağı düşünüldüğünde gerçekçi tahmin yapmak için zaman serileri kullanılması daha uygundur [5]. Fakat şartlarda değişim olacağı (mevsim değişimi, kampanya, fiyat artışı vb.) bilirse zaman serileri yetersiz kalacaktır. Bu durumda ilişkiye dayanan tahmin yöntemleri tekniği kullanılması daha uygun olacaktır.

2.1. İstatistik Bilgisi

Zaman serileri ve istatistik yaklaşımı anlamak için temel düzeyde istatistik bilgisi gerekmektedir. Günlük sıcaklığı aynı olan günlerin tüketim değerlerini ele alacak olursak; tek değişkenli veri elde edilir. Bu sıcaklık için tek değer yerine sıcaklık aralığı veya birden fazla sıcaklık için tüketim değerleri ele alındığı durumda, iki (veya çoklu) değişkenli veri elde edilmiş olacaktır.

2.1.1. Tek değişkenli istatistik

Belirli bir sıcaklıktaki tüketim değerleri ele alındığında birbirinden farklılık gösterebilmektedir. n adet farklı tüketim verisinin (x) ortalaması;

$$Ort_x = \frac{1}{n} \sum_{i=1}^n X_i \quad (2.1)$$

olacaktır. [1,19]. Bu tüketimlerin ortalaması, eldeki veri için sadece genel bir bilgi verir. Eldeki tüketim verisinin bu ortalama ile olan ilişkisi (değerlerin ortalamaya yakın veya uzak olması) varyans ile tanımlanmaktadır [1,19].

$$Var(X) = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \quad (2.2)$$

şeklinde ifade edilir.

Varyans bulunurken tüketim değeri ile ortalama tüketimin farkının karesi alınır. Böylece farkların toplamının sıfır olması engellenir. Burada çıkan sonuç X^2 boyutunda bulunduğundan tekrar X türüne dönüştürmek gerekir. Bu dönüşüm sonucu;

$$S_x = \sqrt{Var(X)} \quad (2.3)$$

olur. Burada çıkan sonuç standart sapma olarak adlandırılır [1,19].

2.1.2. İki değişkenli istatistik

Tek değişkenli istatistikte belirli sıcaklık için tüketim değerleri ile ilgili istatistiki veriler bulunur. İki değişkenli istatistikte ise sıcaklık farklı değerler alabilmekte ve her farklı sıcaklıkta farklı tüketim değeri oluşabilmektedir.

Tüketim için X , sıcaklık için Y denildiğinde, X ile Y arasında bir ilişki bulunabilir. Bu ilişki kovaryans olarak adlandırılır [1].

$$Cov(XY) = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y}) \quad (2.4)$$

şeklinde ifade edilir. Bu denklemde \bar{X} , X değişkeninin aritmetik ortalaması, \bar{Y} , Y değişkeninin aritmetik ortalamasıdır. Kovaryans X ve Y 'nin birimleri cinsinden sonuç

döndürecekler. X ve Y arasındaki ilişkiyi bulmak için birimsiz bir değer bulunmalıdır. Bu değer;

$$r_{XY} = \frac{\text{Cov}(XY)}{S_X S_Y} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (2.5)$$

olarak bulunur ve *korelasyon katsayısı* olarak adlandırılır [1,3].

Kovaryans ve korelasyon katsayısı ikili istatistiğin temellerini oluşturur. Kovaryans ve korelasyon katsayısı iki değişken arasındaki lineer ilişkiyi göstermektedir.

2.2. Zaman Serileri Analizi

Şartlarda büyük değişim olmadığı durumlarda gerçekçi tahmin yapmak için zaman serileri kullanılır. Zaman serileri dört farklı bileşenden meydana gelmektedir [1,2,5,6] (Şekil 2.1).

- Trend Etkisi (T): Serinin doğrusal, parabolik vb. eğilimini gösterir. Zaman serisinin uzun dönemde belirli bir yöne doğru gösterdiği gelişmedir.
- Konjonktür Etkisi (C): Zaman serilerinde kendini tekrar eden fakat bu tekrarların düzeni belirli olmayan ve trendin etrafında oluşan değişimlerdir.
- Sezonallık Etkisi (S): Seride tekrar eden devir hareketinin tümüne mevsim dalgalanmaları denir. Serideki ölçütlere göre haftalık, aylık, yıllık olabilir.
- Düzensiz Hareketler (I): Trend, konjonktür ve sezonallık ile açıklanamayan sonuçlar düzensiz hareketleri göstermektedir. Bunların ne zaman, nasıl bir şiddette ortaya çıkacağı önceden belirlenemez.

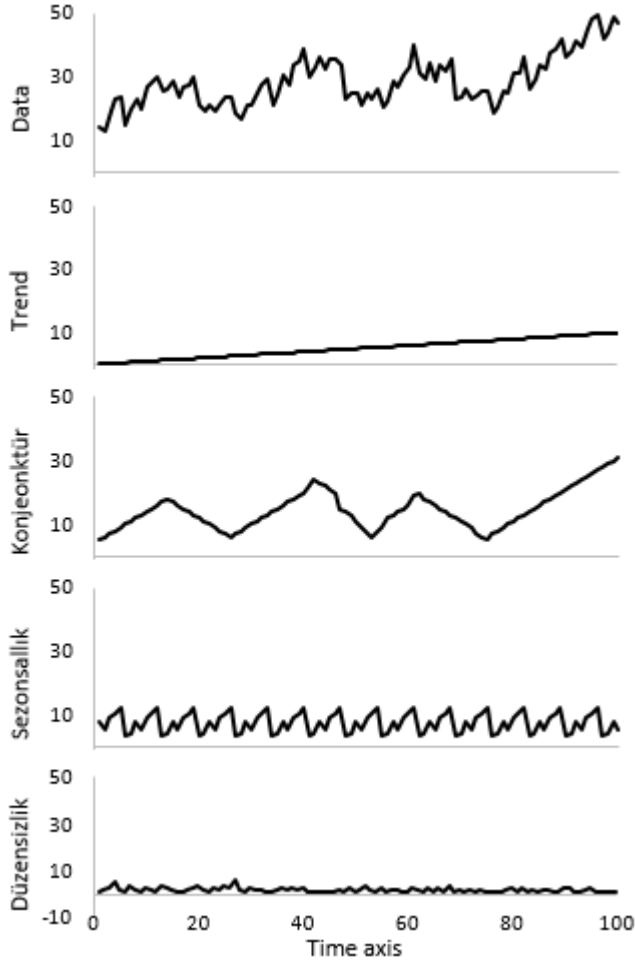
Bu bileşenler zaman serilerinde iki farklı şekilde model oluşturulurken kullanılmaktadır.

- Toplamsal Model

$$Y = T + C + S + I$$

- Çarpımsal Model

$$Y = T \times C \times S \times I$$



Şekil 2.1. Zaman serisinin ayrıştırılması [20,22]

2.2.1. Trendin belirlenmesi

Trendin belirlenmesi için farklı yöntemler bulunmaktadır. Bunlar hareketli ortalamalar yöntemi ve en küçük kareler yöntemidir. En küçük kareler yönteminde farklı trend fonksiyonları bulunmaktadır. Bu fonksiyonlardan bazıları doğrusal, ikinci dereceden, üstel şeklinde belirtilebilir.

Hareketli ortalamalar yönteminde geçmiş dönem verilerinin ortalaması alınır [1,5]. Bu ortalama ya içinde bulunulan dönemin (merkezi hareketli ortalama) ya da bir sonraki dönemin (basit hareketli ortalama) tahmini olur. Doğrusal trend eğilimi;

$$T = a + bt + e \quad (2.6)$$

şeklinde hesaplanır. Bu en küçük kareler yöntemi kullanılarak a ve b tahmin edilir. e ise gerçekleşme ile tahmin arasında kalan kalıntılardır. t zaman serisinin zaman adımını, Y ise bağımlı değişkenin (trend değişkeni) değerini göstermektedir. Hata gerçekleşen değer ile tahmin arasındaki farktır ve aşağıdaki gibi ifade edilir;

$$e = Y - T \quad (2.7)$$

Hatanın (e) minimum olması istenmektedir. Diğer taraftan hatanın yönden bağımsız olması gerekmektedir [5].

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (Y_i - T_i)^2 = \sum_{i=1}^n (Y_i - a - bt_i)^2 = \text{minimum} \quad (2.8)$$

şeklinde yazılabilir. Bu yöntemle “en küçük kareler yöntemi” denmektedir. Bu denklemin minimum olması için a ve b 'ye göre türevlerinin sıfır olması gerekmektedir [5].

$$\frac{de}{da} = 2 \sum_{i=1}^n -(Y_i - a - bt_i) = -\sum Y + na + b \sum t = 0 \quad (2.9)$$

$$\frac{de}{db} = 2 \sum_{i=1}^n -t_i (Y_i - a - bt_i) = -\sum Y t + a \sum t + b \sum t^2 = 0 \quad (2.10)$$

$$\sum t = na + b \sum t \quad (2.11)$$

$$\sum Y t = a \sum t + b \sum t^2 \quad (2.12)$$

İki bilinmeyenli iki denklem çözülerek a ve b bulunur. Böylece başlangıç anındaki değeri olan “ a ” ve denklemin trendini gösteren “ b ” bulunmuş olur.

2.2.2. Konjonktür etkisi

Konjonktür etkisi bulunmadan önce verideki trend bulunmalıdır. Bu trend sayesinde konjonktür daha rahat anlaşılır. Konjonktürde; dönemsel olarak artış ve azalışlar olmaktadır. Bu etkinin süresi uzundur ve kendini seride belli eder. Doğal gaz tüketiminde kış ve yaz mevsimleri konjonktürel etkiyi göstermektedir. Konjonktür bulunurken verinin çevrim zamanını belirlemek önemlidir. Çevrim zamanına göre yürüyen ortalamalar yöntemi kullanılarak bulunan ortalama değerler ile ilgili zamandaki değerlerin oranı konjonktürel etkiyi verecektir. 12 aylık merkezi hareketli ortalamaları hesabı;

$$Y'_{t+6_{H.O.}} = \frac{Y_t / 2 + Y_{t+1} + Y_{t+1} + \dots + Y_{t+11} + Y_{t+12} / 2}{12} \quad (2.13)$$

şeklinde olur. Konjonktür etkisini bulmanın yöntemi de hareketli ortalamaların ($\overline{Y'_{H.O.}}$) trende bölünmesidir.

$$C = \frac{\overline{Y'_{H.O.}}}{T} \quad (2.14)$$

2.2.3. Sezonsallık (Mevsimsellik) etkisi

Mevsimsel etki, konjonktür etkisine göre kısa sürede ve daha sık görünür. Mevsimsel etki; konjonktür etkisine göre daha küçük değişimlere sebep olur [1,5,6]. Buna doğal gaz tüketiminde incelediğimiz zaman hafta sonu oluşan tüketim değişimleri sezonsallık ile ifade edilir. Mevsimselliği ayırtmak için gerçek serinin, hareketli ortalamalar serisine bölümüyle bulunmaktadır.

$$SI = \frac{TCSI}{TC} \quad (2.15)$$

2.2.4. Düzensiz hareketler (Kalıntılar)

Zaman serisinde trend, sezonsallık, konjonktür bileşenleri ile açıklanamayan ve gerçek değer ile bileşenlerin arasında kalan değerler düzensiz bileşenleri oluşturmaktadır. Her zaman serisinde düzensiz bileşen bulunur [5].

2.3. Zaman Serileri

Zamana bağlı verilerin oluşturduğu seriler kullanılarak farklı yöntemler ile tahmin yapılabilmektedir. Bu yöntemler farklı biçimlerde olabileceği gibi birleşik yöntemler de olabilmektedir. Bu yöntemlerden bazıları [1,2,3,5,6,21];

- Yürüyen (Hareketli) ortalamalar (MA Modelleri)
- Üstel düzgünleştirme yöntemi
- Otoregresif modeller (AR Modelleri)
- Bileşik otoregresif hareketli ortalamalar (Box Jenkins- ARIMA)
- Çoklu regresyon yöntemi (Multiple regression)
- Gözlenemeyen bileşen analizi (Unobserved component analysis)

şeklinde. Bu tez çalışmasında ARIMA ve çoklu regresyon yöntemleri uygulanmıştır. Hareketli ortalamalar ve otoregresif modelleri ARIMA yönteminin özel durumlarıdır.

2.3.1. Hareketli ortalamalar

Bu yöntemde (k) sayıda geçmiş dönem verisinin aritmetik ortalaması alınarak, bir sonraki değer tahmini olarak kullanılmaktadır. Burada dönem geçtikçe bir sonraki dönemi içinde barındırırken, kendinden önceki k. dönemi dışlar [1,5].

$$Y'_{t+1} = \frac{1}{k} \sum_{i=t-k+1}^t Y_i \quad (2.16)$$

olur.

2.3.2. Otoregresif modeller

Bu yöntemde regresyon analizinin zaman serilerine uygulaması söz konusudur. Regresyon analizinde bağımlı değişken olan Y ; $X_1, X_2, X_3, \dots, X_n$ gibi bağımsız değişkenler ile temsil edilirken (2.20), zaman serilerinde Y_t kendinden geçmiş dönemlerin $Y_{t-1}, Y_{t-2}, \dots, Y_{t-n}$ değerleri ile temsil edilir (2.21) [1,5].

$$Y = a + bX_1 + cX_2 + \dots + kX_n \quad (2.20)$$

$$Y_t = a + bY_{t-1} + cY_{t-2} + \dots + kY_{t-k} + e_t \quad (2.21)$$

Bu denklemler $t-1$; t anına göre bir periyodluk bekletme (lag) ifadesidir. Bu bekletme; periyodik farkı arttıkça değişecektir. Zaman serilerinde yakın periyodlar arasında ilişkiler bulunmaktadır. Bu ilişkilerin belirlenmesinde tahmin edilecek değişkenin geçmiş dönem değerleri arasındaki hesaplanan korelasyon katsayıları “Otokorelasyon katsayısı” olarak adlandırılır. Bu katsayı zaman serisinin durgun veya tesadüfi olup olmadığını belirtir [1,5].

2.3.2.1. Otokorelasyon katsayısı

Y_t değerinin Y_{t-1} değerinden Y_{t-k} değerleri arasındaki k dönemlik ilişki gücünü göstermek için belirlenmektedir.

Otokorelasyon katsayısı hesaplanmadan önce otokovaryans hesaplanır [5] (2.22).

$$C_k = \frac{1}{N} \sum_{t=k+1}^N (Y_t - \bar{Y})(Y_{t-k} - \bar{Y}) \quad (2.22)$$

hesaplanır. Daha sonra otokovaryans serinin varyansına bölünür ve otokorelasyon katsayısı hesaplanmış olur (2.23).

$$r_k = \frac{\frac{1}{N} \sum_{t=k+1}^N (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\frac{1}{N} \sum_{t=1}^N (Y_t - \bar{Y})^2} \quad (2.23)$$

Dönem farkı arttıkça, her dönem için farklı otokorelasyon katsayısı hesaplanacaktır [5].

Tablo 2.1 üzerinde otokovaryans ve otokorelasyon katsayılarının basit bir şekilde hesaplanması için örnek veriler ile birlikte tablo oluşturulmuştur. Bu tablo üzerinde 1 dönem gecikme için otokovaryans ve otokorelasyon katsayıları hesaplanmıştır.

Tablo 2.1. Otokorelasyon katsayısının hesaplanması

t	Y _t	Y _{t-1}	(Y _t - \bar{Y})	(Y _{t-1} - \bar{Y})	(Y _t - \bar{Y}) ²	(Y _t - \bar{Y}) (Y _{t-1} - \bar{Y})
1	1500		112,8333		12731,36	
2	1468	1500	80,83333	112,8333	6534,028	9120,694444
3	1415	1468	27,83333	80,83333	774,6944	2249,861111
4	1358	1415	-29,1667	27,83333	850,6944	-811,8055556
5	1304	1358	-83,1667	-29,1667	6916,694	2425,694444
6	1278	1304	-109,167	-83,1667	11917,36	9079,027778
Toplam	8323				39724,833	22063,47222

$$\bar{Y} = \frac{8323}{6} = 1387,167$$

$$C_1 = \frac{22063,47222}{6} = 3677,245$$

$$r_k = \frac{22063,4722}{39724,833} = 0,5554$$

Otokorelasyon analizinde tesadüflük araştırılır, durgunluk saptanır. Eğer seri durgun değilse durgun hale dönüştürülür ve mevsim etkisi ortaya çıkartılır.

2.3.2.2. Kısmi otokorelasyon katsayısı

Kısmi otokorelasyon katsayısı Y_t değer ile Y_{t-k} değeri arasındaki ilişki derecesini Y_{t-1}, Y_{t-2}, ..., Y_{t-k+1} değerleri olmadan bulunmasıdır. Y_t ile Y_{t-1} arasında anlamlı bir

otokorelasyon olduğu bilindiğinde Y_t ile Y_{t-2} arasında da bir ilişki olur. Bu sebeple Y_t ile Y_{t-2} arasındaki gerçek korelasyonu ölçmek için, ara değer olan Y_{t-1} etkisini ortadan kaldırmak gerekmektedir. Kısmi otokorelasyon tam anlamıyla bunu yapar ve ARMA yönteminin tanımlanmasında kullanılmaktadır [1,5]

$k = 1$ için $r_{kk} = r_1$ olacaktır.

$$\text{Eğer } t = 2, 3, \dots \quad \text{ise} \quad r_{kk} = \frac{r_k - \sum_{j=1}^{k-1} r_{k-1,j} r_{k-j}}{1 - \sum_{j=1}^{k-1} r_{k-1,j} - r_j} \quad (2.24)$$

$$r_{kj} = r_{k-1,j} - r_{kk} r_{k-1,k-j} \quad j = 1, 2, \dots, k-1 \quad (2.25)$$

r_k : k dönem kaydırmalı otokorelasyon katsayısı

r_{kk} : k dönem kaydırmalı iki zaman serisi arasındaki kısmi otokorelasyon katsayısı olmaktadır.

Gerekli işlemler yapıldığında;

$$r_{22} = \frac{r_2 - \sum_{j=1}^{2-1} r_{2-1,j} r_{2-j}}{1 - \sum_{j=1}^{2-1} r_{2-1,j} - r_j} = \frac{r_2 - r_{11} r_1}{1 - r_{11} r_1} \quad (2.26)$$

$$r_{21} = r_{11} - r_{22} r_{11} \quad (2.27)$$

$$r_{33} = \frac{r_3 - (r_{21} r_2 + r_{22} r_1)}{1 - (r_{21} r_1 + r_{22} r_2)} \quad (2.28)$$

olacaktır.

Otoregresif modeller AR(m) şeklinde ifade edilmektedir. Bu ifadelerde $\hat{\phi}_1, \hat{\phi}_2, \dots, \hat{\phi}_m$, AR modelinin m kısmi otokorelasyon katsayılarıdır.

$$\begin{aligned}
Y_t &= \hat{\phi}_1 Y_{t-1} + e_t \\
&\dots \\
Y_t &= \hat{\phi}_1 Y_{t-1} + \hat{\phi}_2 Y_{t-2} + \dots + \hat{\phi}_m Y_{t-m} + e_t
\end{aligned} \tag{2.29}$$

Bu denklemler çözümlenerek $\hat{\phi}$ değerleri bulunur. $\hat{\phi}_1$ değeri anlamlı, diğer $\hat{\phi}$ değerleri anlamsız ise AR(1) modeli oluşur. AR(2) modeli geçerli ise $\hat{\phi}_1$ ve $\hat{\phi}_2$ değerleri anlamlıdır.

Kısmi otokorelasyon katsayısındaki değerler, AR modelinin katsayısını vermektedir. Bu durumda otokorelasyon değerleri üstel olarak azalır. Tam tersi durumda ise (otokorelasyon katsayıları üstel değil fakat kısmi otokorelasyon katsayıları üstel azalıyor ise) MA modellerini kullanmak uygundur.

2.3.2.3. Birleşik otoregresif hareketli ortalama yöntemi (Box-Jenkins)

Literatürde bu yöntem ARIMA şeklinde geçmektedir. AR, MA, ARMA modellerini kapsamaktadır. ARIMA(1,0,0) şeklinde gösterilen AR(1) modeldir. ARIMA(0,0,2) şeklinde gösterilen ise MA(2) modeldir. Aynı şekilde ARIMA(2,0,3); ARMA(2,3) modelini göstermektedir. ARIMA içinde sırasıyla p, d, q değişkenleri bulunmaktadır [1,5,6,7]. “p” değişkeni otoregresif modelde bulunan “m” değişkenini göstermektedir. “q” değişkeni hareketli ortalama kısmını belirtir. “d” değişkeni ise serinin durağanlaştırmak için kaç kez farkının alındığını göstermektedir.

Bu modelde MA(q) kısmını açıklamak gerekirse, MA(1) modeli için,

$$Y_t = e_t - \theta_1 e_{t-1} \tag{2.30}$$

olduğu bilinmektedir.

$$Y_{t-1} = e_{t-1} - \theta_1 e_{t-2} \tag{2.31}$$

$$Y_t = e_t - \theta_1(Y_{t-1} + \theta_1 e_{t-2}) \quad (2.32)$$

olur. İç içe e_{t-2} hatası e_{t-3} cinsinden ifade edilirse,

$$Y_t = e_t - \theta_1 Y_{t-1} - \theta_1^2 e_{t-2} - \theta_1^3 e_{t-3} \quad (2.33)$$

olur. MA(2) modeli ise;

$$Y_t = e_t - \theta_1 Y_{t-1} - \theta_2 e_{t-2} \quad (2.34)$$

MA(3) modeli için ise;

$$Y_t = e_t - \theta_1 Y_{t-1} - \theta_2 e_{t-2} - \theta_3 e_{t-3} \quad (2.35)$$

şeklinde olur. Genel olarak MA(q)

$$Y_t = e_t - \theta_1 Y_{t-1} - \dots - \theta_q e_{t-q} \quad (2.36)$$

olur.

2.4. Çoklu Regresyon Yöntemi

Bu yöntemde bağımlı değişkeni bağımsız değişkenler ile ifade edilmeye çalışılmaktadır. Dondurma satışlarını ele aldığımızda hem sıcaklıkla hem de GSYİH ile ilişkili olduğunu biliyorsak, dondurma satışını tahmin etmek için kullanılacak yöntem çoklu regresyon olabilmektedir.

Regresyon analizinden farklı olarak diğer değişkenlerde denkleme eklenmektedir [1,3].

$$Y = k_0 + k_1 X + e \quad (2.37)$$

(2.37) denklemi basit doğrusal regresyon denklemidir. Denklemde yer alan Y bağımlı değişken, X bağımsız değişkenidir. Çoklu regresyonda ise k_0 ve k_1 katsayılarına k_2 , k_3 şeklinde bağımsız değişken sayısı kadar katsayı eklenir [1,3,4].

$$Y = k_0 + k_1X_1 + k_2X_2 + e \quad (2.38)$$

(2.37) ve (2.38) denklemlerinde regresyon doğrusaldır. Bu denklem türü dışında doğrusal olmayan (eğrisel) regresyon denklemi de bulunmaktadır.

2.4.1. Çoklu regresyon yönteminin güvenilirliği

Bu çalışmada kısmi regresyon katsayılarının testleri güvenilirliği belirlerken kullanılmıştır. Genel olarak regresyon katsayısı, değişkenin standart sapmasına bölünür ve t elde edilir. “ t ” literatürde “Student – t ” olarak da bilinir. Bu değer yüksek olması, bağımsız değişken etkisinin modelde anlamlı olacağını göstermektedir. Bu değer dağılım tablosu yardımıyla incelenerek olasılık değeri olan “ p ” bulunur. Bulunan bu olasılık değeri düşükse anlamlı olduğu ve sifira yaklaştıkça bağımsız değişkenin anlamının arttığı söylenir [1,4].

2.5. Hata ve Ortalama Mutlak Yüzdese Hata

Tahminler ile gerçekleşen değerler arasında her zaman bir farklılık bulunmaktadır. Bu farklılık hata olarak adlandırılır. Hata değeri doğal gaz tüketiminde miktar olarak kışın yüksek, yazın ise düşük olacaktır. Hatanın miktar olarak incelenmesinde sakınca bulunmaktadır. Bu sakıncayı ortadan kaldırmak için bulunduğu şartlara göre oranlamak hatanın birbirine olan bağımlılığını azaltacaktır. Bunu sağlama için tahmin edilen değer ile gerçekleşen değer farkı alındıktan sonra gerçekleşen değere bölerek oranlamak gerekir. Bu oranlamayı yüzdese olarak ifade edebilmek için 100 ile çarpılması gerekmektedir. Böylece Ortalama Mutlak Yüzdese Hata ortaya (OMYH) çıkacaktır [1,23]. OMYH;

$$OMYH = 1/n \sum_{i=1}^n \left| \frac{Y_i - Y}{Y_i} \cdot 100\% \right| \quad (2.39)$$

şeklinde ifade edilir [1,20,22,23].

BÖLÜM 3. DOĞAL GAZ TÜKETİM VERİLERİ

Doğal gaz tüketimi çeşitli etkilere bağlı olarak değişmektedir. Meteorolojik değişkenler, tüketici davranışları, tüketici sayıları, takvimsel olaylar gibi zamanla değişen veriler doğal gaz tüketimini etkilemektedir. Bu bölümde doğal gaz tüketimini etkileyebilecek bu değişkenler ve verilerin toplanma aşamaları gösterilecektir.

3.1. Tüketim Verisi

Doğal gaz tüketimleri hesaplanırken çeşitli yöntemler bulunmaktadır. Doğal gaz istasyona belirli basınç altında gelmekte şehir içi dağıtım hattına iletilmektedir. Doğal gaz şehir içine dağıtılırken basıncı düşürülmekte ve tüketim ölçümü yapılmaktadır. Bu ölçümlerde basınç düşürme işlemi regülatör ile, birim enerji miktarı ise kromatograf ile yapılmaktadır. Genel olarak şehir girişleri olan bu istasyonlar RMS-A (Reducing & Measuring Station - Basınç Düşürme ve Ölçüm İstasyonu) şeklinde adlandırılır. “A” istasyon tipini göstermektedir. “B” tipi RMS yüksek tüketime sahip sanayi kuruluşlarında bulunurken, “C” tipi RMS ise RMS-B tipine nazaran daha az tüketim veya basınca ihtiyaç duyan istasyonlardır. RMS üzerinde mobil iletişim üzerinden saatlik olarak okunan tüketim değerleri ilgili veri tabanına aktarılmaktadır. Bu veriler daha sonra gaz günü bittiğinde toplanarak günlük tüketimi oluşturmaktadır. Doğal gaz sektöründe gaz günü kavramı ilgili günün sabah 08:00 saatinden (08:00 09:00 arası dahil) bir sonraki günün 08:00 saatine kadar (08:00 – 09:00 hariç) geçen süre olarak tanımlanmaktadır.

Tüketimin bulunmasında verilerin geldiği diğer bir kaynak ise OSB (Organize Sanayi Bölgesi) tüketimleridir. OSB müdürlüğünün doğal gaz tüketimi kendi içinde bulunan uzaktan okuma sistemi ile olmaktadır. OSB içindeki tüm sayaçlar uzaktan okunmakta ve OSB veri tabanına kaydedilmektedir.

Bu çalışmada geçen evsel doğal gaz tüketimi RMS-A, RMS-B ve RMS-C istasyonları ve OSB tüketimlerinin ilişkisi ile bulunmuştur. RMS-A üzerinden çekilen doğal gaz miktarı, RMS-B, RMS-C ve OSB tüketimi çıkartılarak kalan miktar ile evsel tüketim gösterilmiştir. 2011 yılı ve sonrasında uzaktan okuma sistemi üzerinden gelen tüketim verileri, 2009 ve 2010 yıllarında günlük olarak gezilerek okunmuş ve ölçülmüştür. Uzaktan okuma sistemi üzerinde hazırlanmış olan *VIEW* yapıları ile bu bilgiler görülmektedirler. Tahmin sunucusuna bu verilerin aktarımına gerek yoktur. Hazırlanmış olan *VIEW* yapıları günlük olarak değişmemektedir ve hergün gerçekleşen tüketimleri getirebilmektedir.

3.2. İklim Verisi

İklimsel olarak değişen veriler Meteoroloji Genel Müdürlüğü'nden alınmaktadır. Bu veriler iki türde alınmaktadır. Bunlar gerçekleşen ve tahmin edilen meteorolojik verilerdir ve ftp sunucusu üzerinden XML (Extensible Markup Language - Genişletilebilir İşaretleme Dili) formatında ayrı dosyalarda alınmaktadır. Gerçekleşen meteorolojik veriler saatlik olarak alınırken, tahmin verileri günde bir kez ve tahminleri günlük olarak içerir. Tahminin alındığı gün dahil 5 günlük tahminler, günlük tahmin XML dosyasında bulunmaktadır. XML dosyalarından okunan veriler, tahmin için hazırlanmış sunucuda bulunan veritabanına anlık olarak aktarılmaktadır.

Saatlik olarak gerçekleşen verileri

- Ortalama Sıcaklık (°C),
- Nem (%),
- Islak Termometre Sıcaklığı (°C),
- Rüzgar Hızı (m/s),
- Rüzgar Yönü,
- Hadise,
- Basınç (mbar),
- Görüş mesafesi (m)

olarak alınmaktadır.

Tablo 3.1. Gerçekleşen saatlik örnek verisi

Hava Tarih	Ort Sıcaklık	Nem	Islak Termometre Sıcaklık	Rüzgar Hizi	Rüzgar Yonu	Durum	Basinc	Gorus Mesafesi
7.12.2011 11:00	14,7	78	10,9	2	300	SAĞANAK YAĞMURLU	1012,6	10
7.12.2011 10:00	14,7	80	11,4	4	350	SAĞANAK YAĞMURLU	1013	10
7.12.2011 09:00	18,2	51	7,9	3	260	ÇOK BULUTLU	1013,4	15
7.12.2011 08:00	17,5	52	7,7	5	230	PARÇALI BULUTLU	1013,4	15
7.12.2011 07:00	16,9	55	7,8	4	240	ÇOK BULUTLU	1013,5	15
7.12.2011 06:00	17,1	52	7,2	4	240	ÇOK BULUTLU	1013,6	15
7.12.2011 05:00	17,7	48	6,6	7	230	ÇOK BULUTLU	1013,1	15
7.12.2011 04:00	16,8	51	6,6	3	220	ÇOK BULUTLU	1012,9	15
7.12.2011 03:00	16,5	51	6,5	3	200	ÇOK BULUTLU	1012,7	15
7.12.2011 02:00	16,7	51	6,6	3	200	ÇOK BULUTLU	1013	15
7.12.2011 01:00	17,7	50	7,2	5	220	ÇOK BULUTLU	1013,8	15
7.12.2011 00:00	18,1	44	5,9	4	220	ÇOK BULUTLU	1013,6	15
6.12.2011 23:00	18,1	44	5,9	3	210	ÇOK BULUTLU	1013,6	15

Günlük olarak yapılan tahmin verileri ise;

- En Yüksek Sıcaklık (°C),
- En düşük Sıcaklık (°C),
- En yüksek Nem (%),
- En düşük Nem (%),
- Rüzgar hızı (m/s),
- Rüzgar yönü (Derece),
- Hadise

olarak alınmaktadır.

Tablo 3.2. Tahmin günlük örnek verisi

Tahmin Tarihi	Yüksek Sıcaklık	Düşük Sıcaklık	Yüksek Nem	Düşük Nem	Ruzgar Hız	Ruzgar Yönü	Hadise
28.12.2011	11	-1	94	65	7	8	SIS
27.12.2011	9	2	94	62	7	43	PB
26.12.2011	5	0	94	75	5	281	AB
25.12.2011	6	2	94	87	16	346	CB
24.12.2011	5	3	96	80	24	354	KKY
23.12.2011	12	8	93	54	7	114	Y
22.12.2011	15	6	91	55	12	234	SY
21.12.2011	14	10	83	54	11	186	Y
20.12.2011	17	10	94	56	12	178	SY

3.3. Tüketici Verisi

Günlük tüketici verileri, doğal gaz tüketen müşterilerin bilgisinden oluşmaktadır. İlgili gün içinde tüketim yapabilecek durumda olan müşteriler her gün değişebilmektedir. Geciken faturalar sebebiyle gaz kesintileri, abone isteğiyle kapatmalar, ilk abonelikler, kiraya verilen eve yerleşen kiracıların abonelik açtırması, vb. birçok durum tüketici verisini etkileyecektir.

Abone Bilgi Yönetim Sistemi veritabanında oluşturulmuş olan *VIEW* yapıları ile oluşturulmuş sanal tablolar, tahmin sistemi için geliştirilmiş olan veritabanına bir yazılım aracılığı her akşam 22:00 – 23:00 saatleri arasında günlük olarak veritabanına aktarılmaktadır.

Tüketici verileri RMS-A ve abone türü kırımında aktarılmaktadır. İlgili çalışmada dört farklı RMS-A üzerinden evsel doğal gaz tüketimi olmaktadır. Abone kırımında ise evsel, merkezi sistem, sanayi, üretim amaçlı ticarethane, resmi kurumlar, okul ve yurtlar, sağlık kurumları, serbest tüketiciler bulunmaktadır.

Bu veri içinde belirtilmiş olan kırımlar ile abone sayısı, BBS (Bağımsız Birim Sayısı) ve birim fiyat gelmektedir. BBS tanımı; abone binasının 200 m² ye kadar alanı için değeri 1 olan ve bundan sonraki ilave her 100 m² ye kadar brüt alan için değeri bir artırılan sayıdır. 125 m² alana sahip bina için BBS “1” olurken, 345 m² alana sahip bina için BBS “3” olmaktadır. Metrekare bilgisi sistemden alınmakla

birlikte eski sistem üzerinden gelen değerlerde sorun bulunduğundan dolayı değişken olarak kullanılmamıştır. Birim fiyat; sayaç üzerinden okunan değer düzeltilmesi ve taahhüt edilen enerji miktarı ile düzeltilmiş halde 1 m³ karşılığına gelen ödeme tutarıdır.

Tablo 3.3. Tüketici örnek verisi

Tarih	RmsA	Abone Türü	Abone Sayısı	BBS Sayısı	m2	BirimFiyat
25.2.2012 22:48	1	10	5	9	449	0,680709
25.2.2012 22:48	1	9	17	157	9934	0,680709
25.2.2012 22:48	1	6	39	69	2215	0,680709
24.2.2012 22:48	5	5	28	1041	15620	0,680709
24.2.2012 22:48	5	1	4502	4743	102176	0,680709
24.2.2012 22:48	5	6	74	229	9907	0,680709
24.2.2012 22:48	5	9	13	187	1626	0,680709
24.2.2012 22:48	5	8	43	331	3921	0,680709
24.2.2012 22:48	5	16	7	464	6405	0,571565
24.2.2012 22:48	4	9	195	3841	59278	0,680709
24.2.2012 22:48	4	8	193	3038	15993	0,680709
24.2.2012 22:48	4	2	5	57	434	0,680709
24.2.2012 22:48	4	6	328	1066	20049	0,680709
24.2.2012 22:48	4	10	36	1470	2599	0,680709
24.2.2012 22:48	4	5	3	27	727	0,680709
24.2.2012 22:48	4	16	3	20	20	0,571565
24.2.2012 22:48	4	1	65210	70144	945924	0,680709

3.4. Takvimsel Veriler

Takvimsel olaylara bağlı olarak her sene tekrarlanabileceği gibi, her sene yeri değişen olaylarda bu tablo üzerinde belirtilmiştir. Takvimsel olaylar gün boyu değerlendirilmektedir. Burada tatil olayları üç durumda incelenmektedir;

- Hafta sonu tatili
- Resmi tatiller: Her sene aynı gün tatil olmaktadır.
- Dini tatiller: Her sene farklı günlerde tatil olmaktadır.

Bu tatiller;

- Hafta Sonu (Cumartesi ve Pazar tatilleri)
- Resmi Tatil (29 Ekim, 19 Mayıs, 1 Mayıs, 23 Nisan, 30 Ağustos, vb.)
- Seçim & Referandum Tatili
- Ramazan Bayramı
- Kurban Bayramı

Belirtilen tatiller tüm abone türleri için her gün olacak şekilde düzenlenmiştir. Bazı özel durumlar da bu takvimsel olaylara eklenerek veriler oluşturulmuştur. Buna göre “Okul ve Yurtlar” abone türünde yaz dönemi boyunca öğrenci olmadığı için “Resmi Tatil” olarak belirlenmiştir.

Takvimsel verilerde bulunan diğer bir veri türü ise kampanyalardır. Kampanyalar yeni abonelerin sisteme dâhil olmasını hızlandırmaktadır. Takvimsel olaylar üzerinde kampanyalar da bulunmaktadır.

Tablo 3.4. Takvimsel Olaylar Örnek Verisi

Tarih	AT	IGM	TN	K
12.06.2011	1	0	4	1
12.06.2011	2	0	4	NULL
12.06.2011	8	0	4	NULL
12.06.2011	9	0	4	NULL
12.06.2011	10	1	NULL	NULL
12.06.2011	5	0	4	NULL
17.06.2011	2	1	NULL	NULL
17.06.2011	8	1	NULL	NULL
17.06.2011	9	1	NULL	NULL
17.06.2011	10	1	NULL	NULL
17.06.2011	5	1	NULL	NULL
17.06.2011	16	1	NULL	NULL

Tarih	AT	IGM	TN	K
18.06.2011	1	1	NULL	1
18.06.2011	2	1	NULL	NULL
18.06.2011	8	0	1	NULL
31.08.2011	1	1	NULL	NULL
31.08.2011	2	1	NULL	NULL
31.08.2011	8	0	3	NULL
31.08.2011	9	0	3	NULL
15.09.2011	5	1	NULL	NULL
15.09.2011	16	1	NULL	NULL
15.09.2011	6	1	NULL	NULL
16.09.2011	1	1	NULL	1
16.09.2011	2	1	NULL	NULL

Tablo 3.4 üzerinde AT; abone türünü, IGM; iş günü olup olmadığını, TN; tatil nedenini, K ise kampanya durumunu göstermektedir. Her gün için abone türüne bağlı olarak tatil ve kampanya durumları bu tablo üzerinde bulunmaktadır.

BÖLÜM 4. DOĞAL GAZ TÜKETİMİNİN TAHMİNİ VE SONUÇLARI

Bu bölümde; Bölüm 3’ de anlatılan veriler ve Bölüm 2’ de bahsedilen istatistik yöntemleri olan birleşik otoregresif hareketli ortalama yöntemi ile çoklu regresyon yöntemi kullanılarak doğal gaz tüketimi tahmin edilmeye çalışılmıştır. Bu tahminler iki şekilde yapılmıştır. Bunlar;

- Tüm tüketim verisi kullanılarak yapılan tahminler
- Konjonktürel etki kaldırılarak yapılan tahminler

olmak üzere iki kısımda gösterilmektedir.

4.1. Konjonktür Etkisinin Kaldırılması

Konjonktürel etki, zaman serilerinde bulunan 4 etkiden biridir. Kendisini uzun dönemli olarak göstermektedir ve büyük dalgalanmalara sebep olmaktadır. Bu etki azaltılırsa tahminlerin daha doğru yapılacağı düşünülmüştür.

Konjonktürel etki doğal gaz tüketimi için mevsimlerin değişmesi ile kendini göstermektedir. Bu çalışmada 01.01.2009 ile 31.12.2012 tarihleri arasında günlük olarak bulunan verileri konjonktürel etkiden arındırmak için 6 farklı model oluşturulmuştur. Bu modellerdeki temel amaç birbirleri ile benzer tüketim davranışı gösteren zaman aralıklarını aynı model içinde tutarak daha doğru tahmin yapılmasıdır [20,22].

Oluşturulan modeller “Model” ön adı ile adlandırılmıştır. Buna göre;

- Model 1: Ocak Şubat,
- Model 2: Mart, Nisan,

- Model 3: Mayıs,
- Model 4: Haziran, Temmuz,
- Model 5: Ağustos, Eylül, Ekim,
- Model 6: Kasım, Aralık

olarak belirlenmiştir. Bu modeller oluşturulurken tüketimler üzerinde bulunan artış, azalış veya durağanlıklar dikkate alınmıştır.

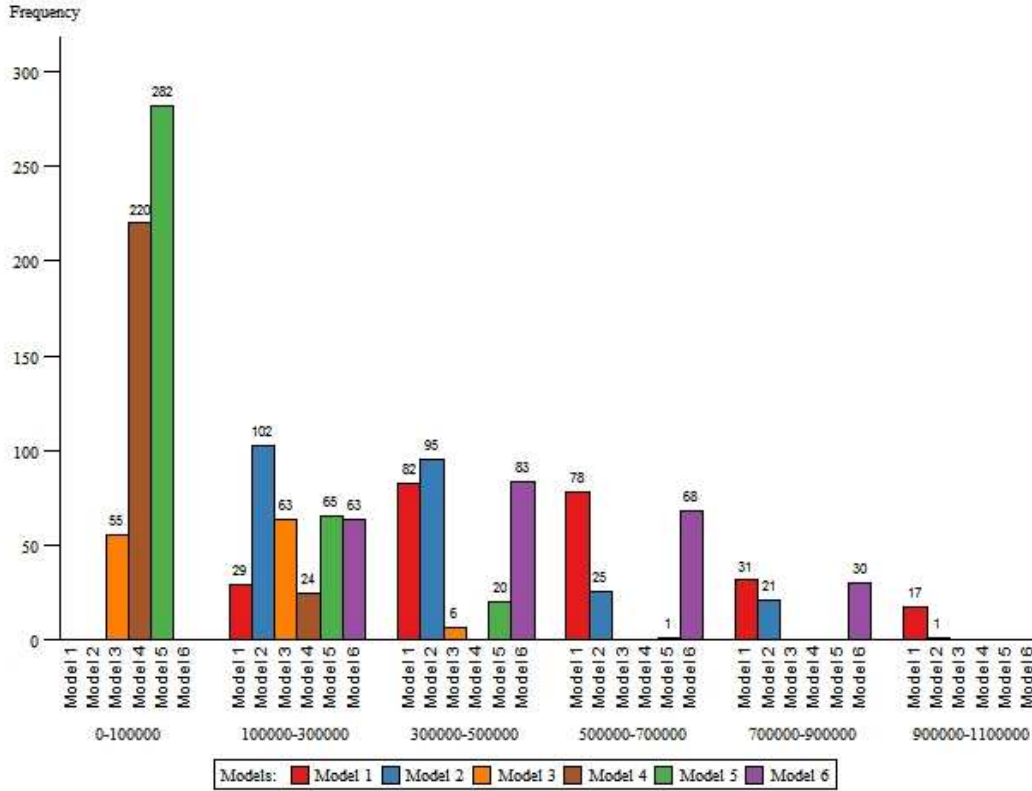
Model 1 veri setinde 2009, 2010, 2011 ve 2012 yıllarının Ocak ve Şubat ayları bulunmaktadır (Şekil 4.1). Aynı şekilde diğer modellerinde veri setleri oluşturulmuştur. Burada konjonktür etkisinden arındırıldıktan sonra sezonsallık etkisinin de modeller üzerinden kaldırıldığı görülmektedir. Bu etkinin model üzerinde bulunması gerekmektedir. Sezonsallığın modellerin içinde bulunması sağlamak amacıyla en son eklenen yıl için modelde yer alan ilk aydan önceki üç haftalık veri modellere eklenmiştir. Model 1 için bu işlem 2011 yılındaki son günden başlayarak 21 günlük verinin 2012 yılı verisinden önce araya eklenmesiyle olmuştur. Model 2 için ise bu durum Mart 2012 ayından önceki 21 günlük Şubat ayı verisinin modele dahil edilmesiyle gerçekleştirilmiştir.



Şekil 4.1. Mevsimsel etkiden arındırılmış modellerin oluşturulması

Bu 6 model için oluşturulmuş verilerin frekansları Şekil 4.2' de gösterilmiştir. Burada tüketim aralıkları en düşük 0 m^3 , en yüksek $1.1 \times 10^6 \text{ m}^3$ olarak alınmış ve 10^5 , 3×10^5 , 5×10^5 , 5×10^7 , $5 \times 10^9 \text{ m}^3$ tüketimlerden ayrıştırılmıştır [20,22]. Model 1, Model

3 ve Model 6 normal dağılım sergilerken, Model 2, Model 4 ve Model 5 üstel dağılım sergilemiştir. Model 1; 234, Model 2; 243, Model 3; 124, Model 4; 244, Model 5; 367 ve Model 6; 244 zaman adımına sahiptir.



Şekil 4.2. Tüketim aralıklarında model verilerinin frekansları [20,22]

Konjonktürel etkinin kaldırılmadığı model ise tüm verileri içermektedir ve “Model 7” olarak adlandırılmaktadır.

4.2. Hata Türleri ve Hata Oranları

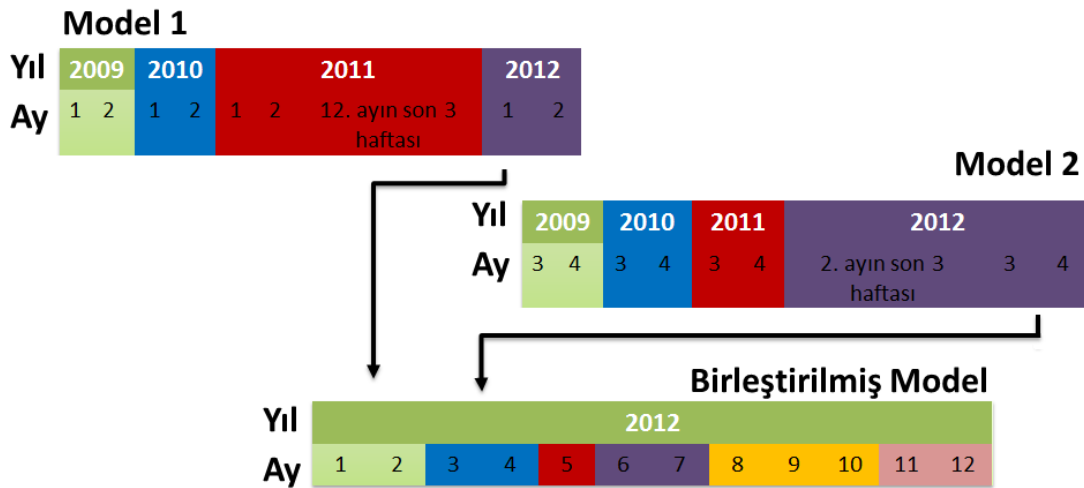
Tahminler ve sonuçların ortalama mutlak yüzdesel hata oranları bulunurken üç farklı hata oranı belirlenmiştir. Bunlar;

- Model ortalama mutlak yüzdesel hataları: Her model için bulunan ve modeldeki tüm veriler kullanılarak bulunan ortalama mutlak yüzdesel hata oranları,

- Modellerin azaltılmış ortalama mutlak yüzdesel hataları: Tüm veriler kullanılarak yapılan tahminlerde sadece 2012 yılı için bulunan ortalama mutlak yüzdesel hata oranı,
- Birleştirilmiş Model ortalama mutlak yüzdesel hatası: 2012 yılı için yapılan tahminlerin birleştirilmesi sonucunda bulunan ortalama mutlak yüzdesel hata oranı

olarak gösterilmektedir.

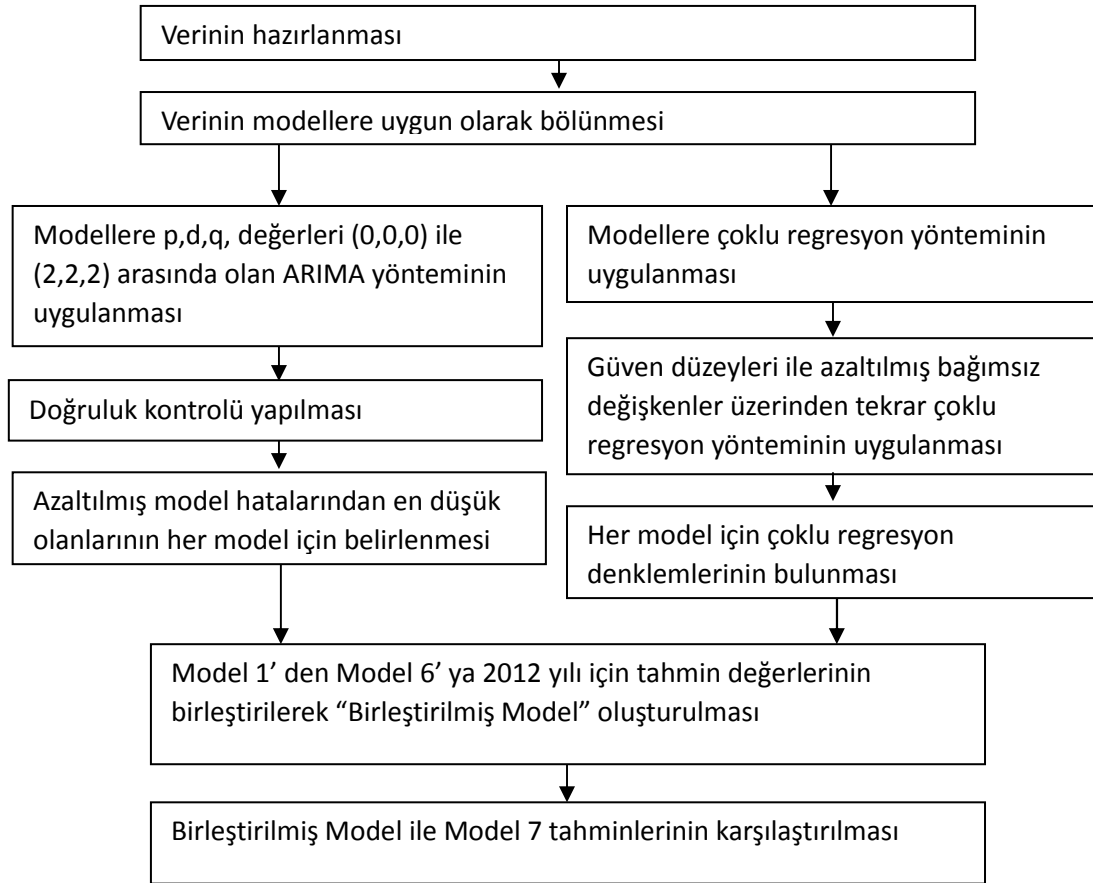
Birleştirilmiş Model ortalama mutlak yüzdesel hatası konjonktürel etkisi kaldırılmış modellerde 2012 yılı için yapılan tahminlerin birleştirilmesi ile bulunan hatadır (Şekil 4.3).



Şekil 4.3. Birleştirilmiş modelin oluşturulması

Oluşturulan modeller kullanılarak iki farklı yöntem ile doğal gaz tüketimleri tahmin edilmeye çalışılmıştır (Şekil 4.4). Bunlar ARIMA ile çoklu regresyon yöntemleridir. ARIMA yönteminde modellere farklı p,d,q değerleri uygulanmış ve doğruluk kontrolü yapılmıştır. Daha sonra modellerin azaltılmış ortalama mutlak yüzdesel hatalarına göre sıralanarak en düşük hatalaya sahip değerler ile yapılan tahmin sonuçları ele alınarak Birleştirilmiş Model ARIMA yöntemi için elde edilmiştir. Çoklu regresyon yönteminde ise regresyon denklemleri her model için bulunmuş, güven düzeyi kontrolü yapılarak birbirini etkileyecek veya etkisi düşük olan değişkenler modelden çıkartılarak yöntem tekrar uygulanmıştır. Uygulanan yöntem

sonucunda her model için çoklu regresyon denklemleri bulunmuştur ve tüketimler tahmin edilmiştir. Her model için yapılan tüketim tahminleri ile Birleştirilmiş Model çoklu regresyon için elde edilmiştir. ARIMA ve çoklu regresyon kullanılarak oluşturulmuş Birleştirilmiş Model sonuçları ile Model 7 kullanılarak yapılan ARIMA ve çoklu regresyon yöntemi tahminleri, gerçekleşme verileri ile karşılaştırılmıştır.



Şekil 4.4. Modelleme ve sonuç adımları [20,22].

Tahminler; Statistical Analysis System (SAS®) yazılımı kullanarak kodlanmış ve yapılmıştır.

4.3. ARIMA Yöntemi İle Tahmin

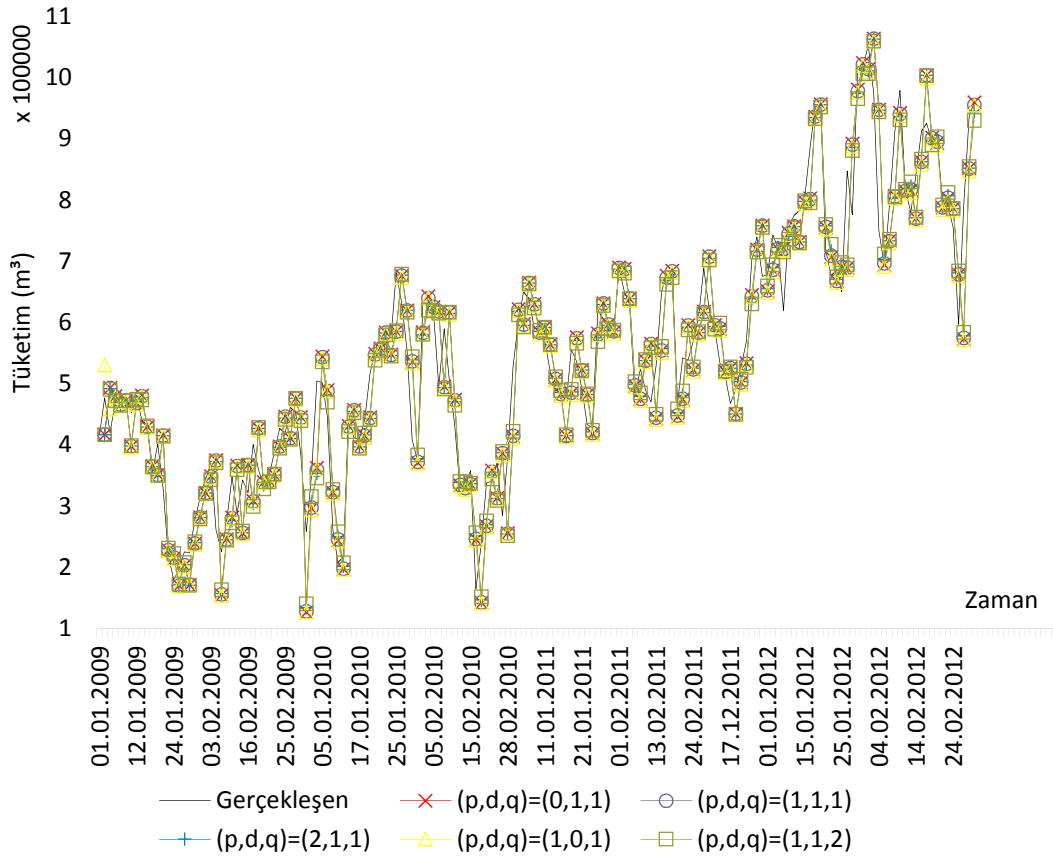
ARIMA yöntemi zaman serileri üzerinde tahminleme işlemini sadece tahmin edilecek veri üzerinden yapmaktadır. Bu nedenle bağımsız değişkenlere gerek duymamaktadır. ARIMA, yapısı gereği farklı p,d,q değerleri alabilmektedir. Her farklı p,d,q durumu için tahmin sonuçları değişecektir. Bu çalışmada p,d,q değerleri

sırası ile 0,0,0 ile 2,2,2 arasında seçilerek tahmin sonuçları çıkartılmıştır. Buna göre her model için 27 farklı tahmin sonucu üretilmiştir. 7 model için 189 tahmin oluşturulmuştur. Modeller için çıkartılan bu tahminlerde doğruluk kontrolü yapıldıktan sonra model hataları ve azaltılmış model hataları bulunmuştur. Azaltılmış model hatalarına göre en düşük hata oranına sahip 5 model çıkartmıştır (Tablo 4.1).

Tablo 4.1. Tüm modeller için en düşük 5 azaltılmış hata [20].

Model 1	(p,d,q)	(0,1,1)	(1,1,1)	(2,1,1)	(1,0,1)	(1,1,2)
	<i>Azaltılmış OMYH</i>	6.76%	6.76%	6.81%	6.82%	6.87%
	<i>OMYH</i>	9.87%	9.87%	9.89%	9.84%	9.92%
Model 2	(p,d,q)	(1,0,1)	(1,0,0)	(0,1,1)	(1,1,0)	(1,1,1)
	<i>Azaltılmış OMYH</i>	15.45%	15.81%	16.01%	16.05%	16.11%
	<i>OMYH</i>	15.80%	16.05%	16.26%	16.28%	16.28%
Model 3	(p,d,q)	(0,2,1)	(0,1,0)	(1,2,1)	(2,2,1)	(1,1,0)
	<i>Azaltılmış OMYH</i>	6.98%	7.00%	7.07%	7.20%	7.24%
	<i>OMYH</i>	18.51%	15.61%	18.53%	18.72%	15.58%
Model 4	(p,d,q)	(1,0,1)	(2,0,1)	(1,0,0)	(1,1,1)	(0,1,1)
	<i>Azaltılmış OMYH</i>	6.54%	6.58%	6.62%	6.74%	6.79%
	<i>OMYH</i>	10.67%	10.74%	10.76%	10.98%	10.95%
Model 5	(p,d,q)	(0,1,2)	(2,1,0)	(0,1,0)	(1,1,2)	(2,1,1)
	<i>Azaltılmış OMYH</i>	6.74%	6.75%	6.83%	6.88%	6.88%
	<i>OMYH</i>	13.57%	13.57%	13.56%	13.82%	13.81%
Model 6	(p,d,q)	(1,1,0)	(1,1,1)	(2,1,1)	(1,1,2)	(0,1,1)
	<i>Azaltılmış OMYH</i>	8.54%	8.56%	8.56%	8.56%	8.61%
	<i>OMYH</i>	12.25%	12.27%	12.27%	12.27%	12.28%
Model 7	(p,d,q)	(1,0,1)	(1,1,0)	(0,1,1)	(1,1,1)	(2,1,1)
	<i>Azaltılmış OMYH</i>	8.67%	8.71%	8.73%	8.75%	8.77%
	<i>OMYH</i>	11.73%	11.86%	11.85%	11.84%	11.84%

Model 1 için bulunan azaltılmış ortalama mutlak yüzdesel hata (Azaltılmış OMYH) için bulunan sonuçlarından en düşük beş sonuç Şekil 4.5 üzerinde gösterilmiştir. Tablo 4.1 incelendiğinde OMYH oranlarının Azaltılmış OMYH' ya göre daha yüksek olduğu görülmektedir. Bunun sebebi temel olarak son yıl verisinde sezonsallık etkisinin modele dâhil edilmesi sonucu daha doğru tahminlerin 2012 yılı için yapılmış olmasıdır.



Şekil 4.5. Model 1 için en düşük hataya sahip tüketim tahminleri [20]

4.4. Çoklu Regresyon Yöntemi ile Tahmin

Çoklu regresyon yönteminde tüketim tahmini, bağımsız değişkenler üzerinden yapılmaktadır. Bu yöntem ile kullanılan bağımsız değişkenler [22];

- Hava durumu değişkenleri; En yüksek, en düşük, ortalama nem ve sıcaklık, rüzgar hızı, bir gün gecikmeli rüzgar hızı, ortalama sıcaklık ve nem,
- Tüketici değişkenleri; abone sayısı ve BBS,
- Takvimsel değişkenler; Ramazan, kurban bayramları ve hafta sonları (cumartesi ve pazar etkisi),
- Sahte değişkenler; Pazartesi, Salı, Çarşamba, Perşembe, Cuma, Cumartesi, 2010 yılı, 2011 yılı, 2012 yılı

şeklinde hazırlanmıştır. Modeller üzerinde bağımsız değişkenlerde çoklu doğrusallığı engellemek için 2009 yılı ve Pazar günü sahte değişken olarak eklenmiştir [1,3].

Bu değişkenler 7 Model için uygulanmıştır. Model 1 için uygulanan çoklu regresyon yönetiminin sonuçları incelenirse (Tablo 4.2), bağımsız değişkenler için bulunan tahminler görülmektedir. 24 bağımsız değişkenin tahmin modellerinde yer alması mümkündür fakat bu bağımsız değişkenlerden bazıları tahminleri yanlış etkileyebilmektedir.

Tablo 4.2. Model 1 için tüm bağımsız değişkenler ile çoklu regresyon sonuçları

Bağımsız değişkenler	Parametre Tahmini	Standart Sapma	t değeri	Pr > t
Kesişim			2,45	0,0151
Ramazan bayramı	0	.	.	.
Kurban bayramı	0	.	.	.
Hafta sonu	36858	29040	1,27	0,2056
Ortalama nem	-1536,3	1390,256	-1,11	0,2703
Rüzgar hızı	-81,8592	1057,694	-0,08	0,9384
Ortalama sıcaklık	-7273,4	5767,687	-1,26	0,2085
En yüksek sıcaklık			-3,27	0,0012
En düşük sıcaklık	-1745,66	3197,198	-0,55	0,5856
En düşük nem	741,0208	789,2721	0,94	0,3488
En yüksek nem	133,3575	894,4721	0,15	0,8816
Abone sayısı			2,88	0,0043
BBS	-16,0214	10,14561	-1,58	0,1156
Lag1 rüzgar hızı	-2967,15	1011,268	-2,93	0,0037
Lag1 ortalama sıcaklık	-8995,24	1315,691	-6,84	<.0001
Lag1 ortalama nem	-1086,58	264,9795	-4,1	<.0001
Pazartesi	83984	29873	2,81	0,0053
Salı	89720	29851	3,01	0,0029
Çarşamba	88204	29989	2,94	0,0036
Perşembe		29847	3,2	0,0016
Cuma	84704	29788	2,84	0,0049
Cumartesi	-8214,96	9357,313	-0,88	0,3809
2010 yılı	5625,67	45524	0,12	0,9018
2011 yılı	-49524	87254	-0,57	0,5709
2012 yılı	-13722	95527	-0,14	0,8859

Fazla bağımsız değişken sayısı modelin esnekliğini de azaltmaktadır. Diğer bir ifade ile bir bağımsız değişkenin etkisi daha anlamlı ve fazla olabilecek iken anlamsız bağımsız değişkenler bunları gölgeleyebilmektedir. Tablo 4.2 üzerinde t değerleri parameter tahminlerinin, standart sapmaya bölünmesi ile bulunmaktadır. Bu t

değerleri olasılık dağılım tablosu üzerinden değerleri bulunarak anlamlılık seviyeleri ölçülmektedir. Güven düzeyinin bağımsız değişkenler için düşük olması istenmektedir. Buna göre 0.1 ve altında güven düzeyine sahip bağımsız değişkenler anlamlı olarak ele alınmıştır (Tablo 4.2). Bu güven düzeyinde bulunan bağımsız değişkenler ile tekrar çoklu regresyon yöntemi uygulanmıştır [22]. Tabloda önemli olan diğer bir nokta ise yıllara bağlı bağımsız değişkenlerin güven düzeylerine bakılarak modelde yer almaması gerektiği görülmüştür. Bu, modellerde konjonktürel etkinin önemli ölçüde etkili olduğu ve yıllardan bağımsız olduğunu göstermektedir. Model 1 için tabloya bakıldığında bir gün gecikmeli sıcaklık, nem ve rüzgarın o gün gerçekleşen nem, sıcaklık ve rüzgardan daha anlamlı olduğu görülmektedir. Bu da evsel kullanıcıların tüketimlerinde hava durumunu daha sonra algılayabildiklerini göstermektedir.

Güven düzeyi kullanılarak her Model için bağımsız değişkenler azaltılmıştır (Tablo 4.3). Bağımsız değişken sayısı azaltılmış olan veri setleri kullanılarak her Model için çoklu regresyon tekrar uygulanmıştır. Böylece etkisi düşük olan değişkenler modelden arındırılmıştır. Tablo 4.3 üzerinde görülen diğer önemli nokta ise güven düzeylerinde veri seti azaltıldıktan sonra iyileşme meydana gelmiştir. Bu beklenen bir durumdur. Böylece bağımsız değişkenlerin etkisi ve esnekliği artmıştır. Veri seti azaltılmış değişkenler incelendiğinde ramazan ve kurban bayramlarının Model 5 ve Model 6' da etkili olduğu görülmektedir. 2009 – 2012 yılları ve arasındaki tarihlerde Ramazan bayramı Ağustos ve Eylül aylarında bulunurken, Kurban bayramı Ekim ve Kasım aylarında bulunmaktadır. Diğer modellerde bu bayramlar ile ilişkinin bulunmaması beklenen durumdur. Güven düzeyleri incelendiğinde abone sayısı 6 modelde bulunmaktadır. Bu durum, abone sayısının etkili bağımsız değişken olduğunu göstermektedir. Diğer taraftan tablo üzerinde bulunan 2010 yılı değişkeninde de 6 model içinde bulunması, 2010 senesinin diğer senelerdeki tüketim ile karşılaştırıldığında farklı bir davranışa sahip olduğunu göstermektedir. Model 4; yıl ve gün sahte değişkenlerini kullanması ise tüketimlerin tamamen dengeli; sadece yıl ve günlere göre tüketimde değişiklik olduğu göstermektedir.

Tablo 4.3. Tüm bağımsız değişkenler ve etkisi düşük olan bağımsız değişkenler çıkartılması ile bulunan güven düzeyleri

Bağımsız Değişkenler	Çoklu regresyon güven düzeyleri							Veri seti azaltılmış çoklu regresyon güven düzeyleri						
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Ramazan bayramı	<.0001	.	0,0048					<.0001		0,0055
Kurban bayramı	0,0020	0,0003	0,0164					0,0012	0,0019	0,0148
Hafta sonu	0,2056	0,5259	0,7532	0,4026	0,3293	0,0141	0,8742						<.0001	
Ortalama nem	0,2703	0,0100	0,2446	0,4351	0,4233	0,4408	0,7099		<.0001					
Rüzgar hızı	0,9384	0,0369	0,8055	0,1364	0,7467	0,8668	0,0018		0,0329					0,0002
Ortalama sıcaklık	0,2085	0,0023	0,8277	0,8683	0,8004	0,0173	0,8569		0,0034				<.0001	
En yüksek sıcaklık	0,0012	0,0655	0,0228	0,9440	0,0907	0,6228	<.0001	<.0001	0,0398	<.0001		<.0001		<.0001
En düşük sıcaklık	0,5856	0,0391	0,5995	0,3224	0,7196	0,6395	0,6469		0,0381					
En düşük nem	0,3488	0,0004	0,3384	0,9264	0,9778	0,1214	0,6754		0,0001					
En yüksek nem	0,8816	0,8866	0,0788	.	0,0264	0,8138	0,5312			0,0104		0,5310		
Abone sayısı	0,0043	<.0001	<.0001	0,9880	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
BBS	0,1156	<.0001	<.0001	0,4778	<.0001	<.0001	<.0001		<.0001	<.0001		<.0001	<.0001	<.0001
Lag1 rüzgar hızı	0,0037	0,7987	0,0479	0,2859	0,6709	0,1340	0,0625	0,0037		0,0687				0,0255
Lag1 ortalama sıcaklık	<.0001	<.0001	0,2956	0,9489	<.0001	<.0001	<.0001	<.0001	<.0001			<.0001	<.0001	<.0001
Lag1 ortalama nem	<.0001	<.0001	0,4160	0,7117	0,8719	0,0266	<.0001	<.0001	<.0001				0,0012	<.0001
Pazartesi	0,0053	0,0840	0,7767	0,0087	0,7699	0,5188	0,0998	<.0001	0,0003		<.0001			0,0008
Salı	0,0029	0,0358	0,4279	0,0048	0,7751	0,5483	0,0610	<.0001	<.0001		<.0001			<.0001
Çarşamba	0,0036	0,0253	0,5760	0,0027	0,9740	0,5314	0,0945	<.0001	<.0001		<.0001			0,0006
Perşembe	0,0016	0,0673	0,6267	0,0015	0,8071	0,3392	0,1327	<.0001	<.0001		<.0001			
Cuma	0,0049	0,1812	0,8930	0,0369	0,6987	0,2785	0,3070	<.0001			<.0001			
Cumartesi	0,3809	0,3593	0,6638	0,0338	0,6353	0,5326	0,4035				0,0753			
2010 yılı	0,9018	<.0001	0,0329	<.0001	0,0199	0,0002	<.0001		<.0001	0,0002	<.0001	<.0001	<.0001	<.0001
2011 yılı	0,5709	<.0001	0,0264	0,0006	0,7555	0,4011	<.0001		<.0001	0,0001	<.0001			<.0001
2012 yılı	0,8859	<.0001	0,0120	0,0007	0,5734	0,6742	0,0054		<.0001	<.0001	<.0001			0,0086

Anlamalı (güven düzeyi düşük) bağımsız değişkenler kullanılarak oluşturulan çoklu regresyon denklemlerinin parameter tahminleri ve standart sapmaları bulunmuştur. Buna göre Model 1 için oluşturulan tüketim tahmini denklemi [22];

$$\begin{aligned} \text{Tahmini} &= 391415 - 17949 (V_1) + 8.5634 (V_2) + 2497.34123 (V_3) \\ \text{Tüketim} &- 14461 (V_4) - 1964.50287 (V_5) + 60042 (V_6) \\ &+ 61365 (V_7) + 61818 (V_8) + 18727 (V_9) + 58633 (V_{10}) \end{aligned}$$

Şeklinde olmuştur. Burada V_1 en yüksek sıcaklığı, V_2 abone sayısını, V_3 bir gün gecikmeli (lag1) rüzgar hızını, V_4 bir gün gecikmeli (lag1) ortalama sıcaklığı,, V_5 bir gün gecikmeli nemi, V_6 pazartesi, V_7 salı, V_8 çarşambayı, V_9 perşembeyi ve V_{10} Cuma değişkenini göstermektedir.

7 model için çoklu regresyon denklemi oluşturulmuştur. Oluşturulan bu denklemler ile tüketimler tahmin edilmiştir ve Tablo 4.4 üzerinde hata oranları gösterilmiştir. Her model için ortalama mutlak yüzdesel hata ve azaltılmış ortalama mutlak yüzdesel hata bulunmuştur. Her iki hata durumu da incelendiğinde Model 3 ve Model 5 yüksek hata oranları görülmüştür. OMYH ile azaltılmış OMYH karşılaştırıldığında ise Model 2 ve Model 4 haricinde azaltılmış OMYH azaldığı görülmüştür.

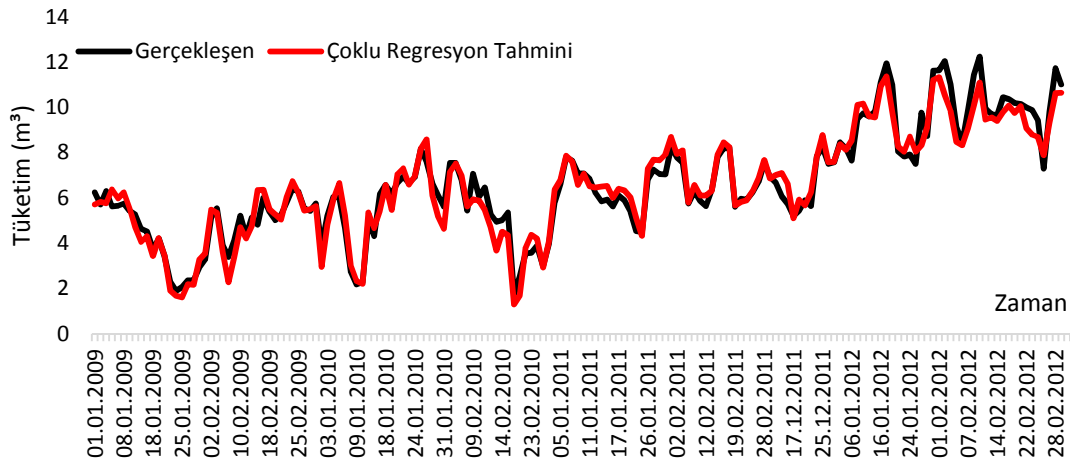
Model 7 tüm verileri içeren model olarak ele alındığında hata oranı 53.03% OMYH olmuştur. Fakat Model 7 için azaltılmış OMYH incelendiğinde 55.32% olarak bulunmuştur. Bunun en temel sebebi diğer yıllardaki verinin etkisinin 2012 yılına göre daha fazla olmasıdır.

Model 1 için bulunan ortalama mutlak yüzdesel hata 7.34% oranının gerçekleşme ve çoklu regresyon tahmin grafiği Şekil 4.6' da gösterilmektedir. Bu grafik üzerinde veriler incelenirse günlere bağlı tüketimin etkili olduğu ilk anda görülebilmektedir. Bu nedenle hafta sonu günleri olan cumartesi ve Pazar hariç diğer günlerin sahte değişkenleri model için bulunmaktadır. Hafta içi günlerinin modelde tüketimi artırıcı etki etmesi, hafta sonunda gerçekleşen tüketim düşüş sebebinin bir bölümünü açıklamaktadır. Bunun dışında gerçekleşen anlık tüketim artış ve azalışları sıcaklık

Tablo 4.4. Çoklu regresyon hata oranları [22].

	<i>OMYH</i>	<i>Azaltılmış OMYH</i>
Model 1	7,34%	5,84%
Model 2	12,54%	17,43%
Model 3	24,81%	23,37%
Model 4	8,82%	10,03%
Model 5	30,69%	26,97%
Model 6	16,61%	10,20%
Model 7	53,03%	55,32%

ve nem üzerinden etkili olduğu görülmektedir. Burada bir gün sonra havadaki değişimin evsel kullanıcılarda daha etkili olduğu, modeldeki lag1 değişkenlerinden anlaşılmaktadır. Yıl bazında eklenmiş olan sahte değişkenler modelde bulunmamaktadır. Bu da kış mevsiminde tüketimin mevsimsel etkilerden fazlasıyla etkilendiği göstermektedir. Genel olarak veri üzerinde tüketimlerdeki artış trendi ise abone sayısı ile çoklu regresyon modelinde kendini göstermektedir. Burada abone başına günlük olarak Model 1 için 8.56 m^3 tüketim öngörmüştür.



Şekil 4.6. Model 1 için gerçekleşme ve çoklu regresyon tahmini [22].

4.5. Gerçekleşen Tüketimler ile Tahmin Sonuçlarının Karşılaştırılması

Tüketim ve bağımsız değişkenlerden oluşturulmuş olan 7 farklı model ve veri setine ARIMA ile çoklu regresyon yöntemi uygulanmıştır. ARIMA yöntemi için yapılan çalışmalarda her model için 27 farklı tahmin üretilmiş, bunlardan en düşük azaltılmış hataya sahip p,d,q değerinin tahmini ele alınmıştır [20]. Çoklu regresyon yöntemi

kullanılarak yapılan tahminde ise öncelikle tüm bağımsız değişkenler ile tahmin yapılmıştır. Daha sonra her model için etkisi olmayan bağımsız değişkenler modelden çıkartılarak çoklu regresyon yöntemi tekrar uygulanmıştır [22].

Gerçekleşen tüketim ile tahmin edilen tüketim arasında karşılaştırma için 2012 yılı kullanılmıştır. 2012 yılı için aylara bölünmüş olan 6 modelden ilgili aylar birleştirilerek önerilmiş olan konjunktür etkisi kaldırılmış model “Birleştirilmiş Model” olarak elde edilmiştir (Şekil 4.3). Model 7 tüm verileri içeren modeldir.

Sonuçlar;

- Tüketim tahmininin karşılaştırılması,
- Kalıntıların karşılaştırılması,
- OMYH karşılaştırması

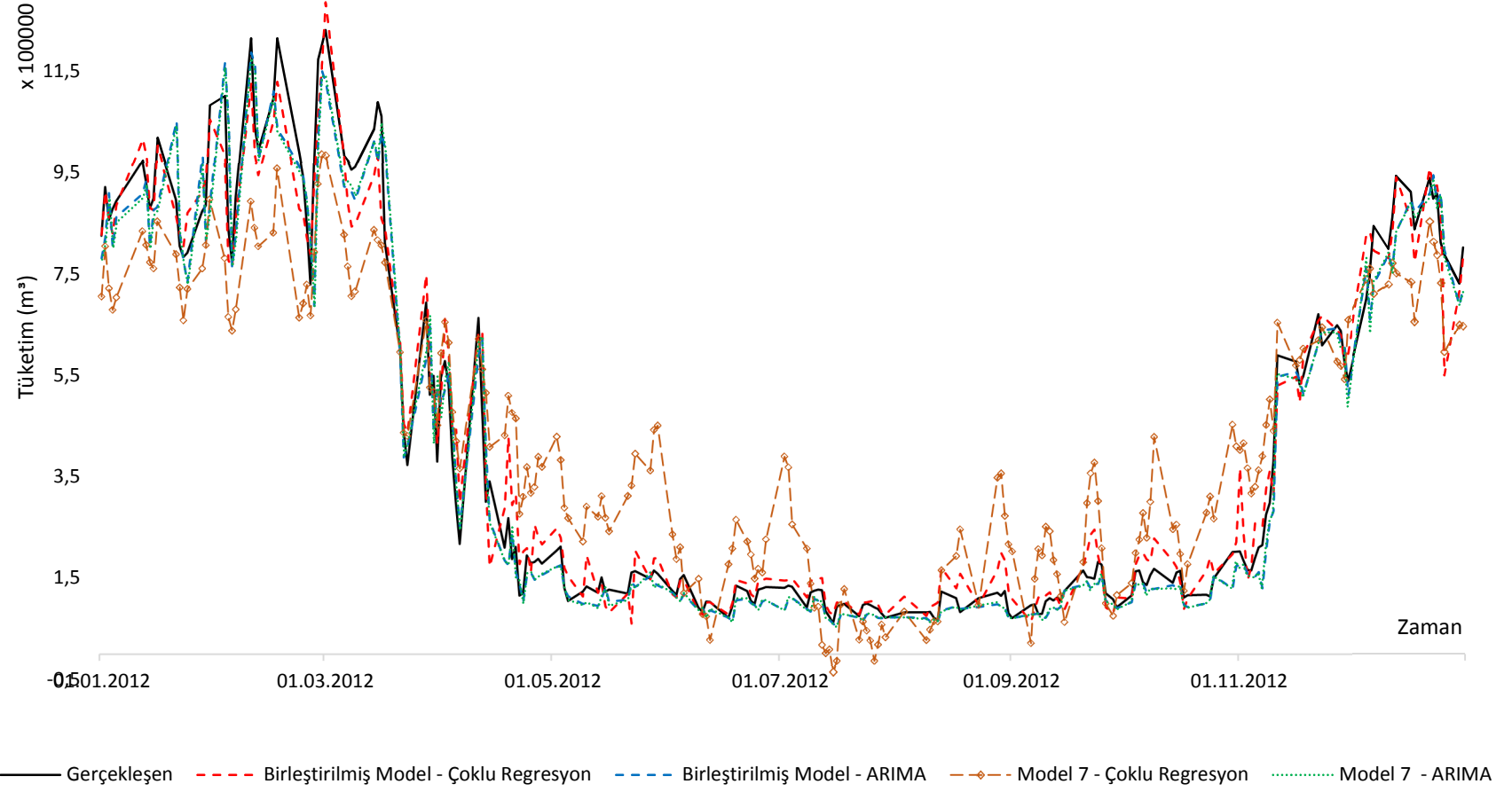
olarak gösterilmiştir.

4.5.1. Tüketim tahmininin karşılaştırılması

Tüketim tahminleri bu çalışmanın temel amacıdır. Bu bölümde konjunktür etkisi kaldırılmış ve kaldırılmamış modellerin karşılaştırılması söz konusudur. Şekil 4.7 üzerinde doğal gaz tüketimleri ve tahminlerin karşılaştırılması görülmektedir. Sonuçlar incelendiğinde ARIMA yöntemlerinin en iyi sonuçları verdiği görülmüştür. Çoklu regresyon yöntemlerinde ise Model 7, Birleştirilmiş Model ile karşılaştırıldığında hatanın çok fazla olduğu görülmektedir [20,22].

ARIMA yönteminde tüketim tahminleri genel olarak gerçekleşen tüketime göre daha düşük çıkmıştır. Grafik üzerinde Model 7 ve Birleştirilmiş Model için ARIMA yönteminin karşılaştırılması birbirine yakın sonuçlar bulunması sebebiyle zorlaşmaktadır.

Çoklu regresyon yönteminde ise tüketim tahminleri, gerçekleşen tüketimlerin hem altında hem de üstünde bulunmaktadır. Bu bağımsız değişkenlerin etkisinden



Şekil 4.7. Tüketim tahmini sonuçları grafiği [20,22]

kaynaklanmaktadır. Model 7 ile yapılan tahminlerde 2012 yılı dışındaki yılların etkisinin fazla olması, sonuçlarda yüksek deęişimlere sebep olmaktadır. Birleřtirilmiř Model incelendięinde konjonktür etkisinin kaldırılması ile aynı tüketim grupları bir araya gelmektedir. Bu durumda bağımsız deęişkenlerde her birim deęişimin etkisi 6 model için de farklı olmaktadır. Bu durum sonuçların iyileřmesini ve tahminlerin doęruluęunu arttırmaktadır.

4.5.2. Kalıntıların karřılařtırılması

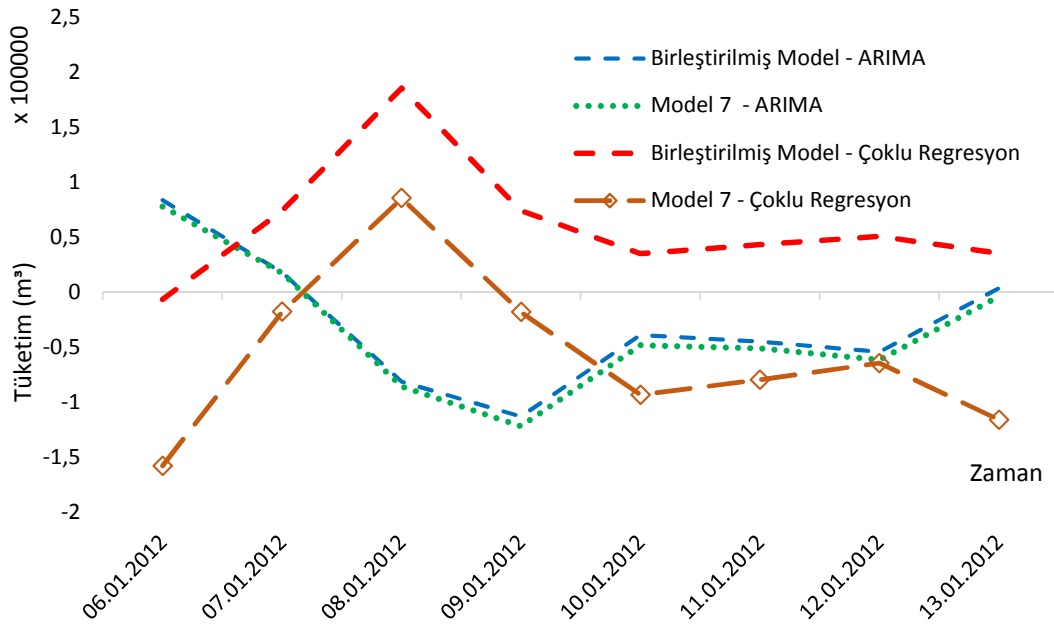
Kalıntılar tahmin deęerleri ile gerçekleřen deęerler arasındaki farktır ve açıklanamayan kısımlardır. Kalıntılar aylara göre farklılık göstermektedir. Kış aylarında bulunan kalıntılar ile yaz aylarında bulunan kalıntıların birbirine eř deęer olmadığı bilinmelidir. Buna göre kışın artan kalıntı miktarlarının yazın azalması beklenmektedir.

Kalıntılar Őekil 4.8 üzerinde gösterilmiřtir. Buna göre Model 7 – çoklu regresyon tahmini yüksek kalıntı bırakan model olmuřtur. En yüksek kalıntılar incelendięinde, 07.01.2012 tarihinde (gerçekleřen 571.550 m³) Birleřtirilmiř Model içinde regresyon yöntemi ile 185.589 m³, 12.05.2012 tarihinde (gerçekleřen 97.445 m³) Model 7 içinde regresyon yöntemi ile 196.504 m³, 12.04.2012 tarihinde (gerçekleřen 321.660 m³) Birleřtirilmiř Model içinde ARIMA yöntemi ile 195.292 m³ ve aynı tarihte Model 7 ARIMA yöntemi ile 200.497 m³ kalıntı olmuřtur. Aynı Őekilde en düşük kalıntılar incelendięinde ise, 26.12.2012 tarihinde (gerçekleřen 716.363 m³) Birleřtirilmiř Model içinde regresyon yöntemi ile -216.128 m³, 04.12.2012 tarihinde (gerçekleřen 624.414 m³) Model 7 içinde regresyon yöntemi ile -288.376 m³, 02.02.2012 tarihinde (gerçekleřen 1.094.755 m³) Birleřtirilmiř Model içinde ARIMA yöntemi ile -340.706 m³ ve 04.12.2012 tarihte (gerçekleřen 624.414 m³) Model 7 ARIMA yöntemi ile -289.913 m³ kalıntı olmuřtur. En yüksek ve düşük kalıntılar incelendięinde Model 7 kalıntılarının Birleřtirilmiř Model kalıntılarından daha yüksek olduęu görölmektedir.



Şekil 4.8. Tahmin kalıntıları grafiği [20,22]

Sonuçlarda örnek bir tarih aralığı için detaylandırılan kalıntı grafiği Şekil 4.9 üzerinde gösterilmiştir. Burada ARIMA yöntemi ile yapılan tahmin değerlerinin ve kalıntılarının birbirine yakın olduğu görülmektedir. Fakat detay grafiğinde görüldüğü gibi Birleştirilmiş Model için kalıntılar daha düşük olmuştur. Aynı grafik incelendiğinde çoklu regresyon kalıntılarının da kendi aralarında benzer davranış sergilediği görülmüş fakat değişkenlerin etkisinin farklı olması nedeniyle sonuçlar değişmiştir.



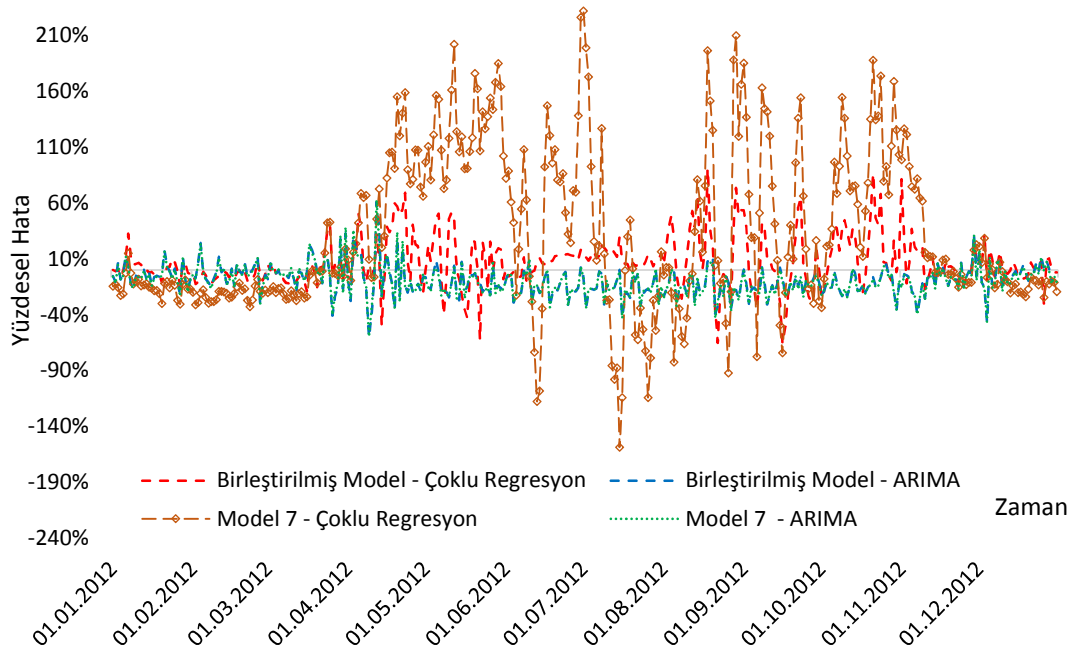
Şekil 4.9. Örnek zaman aralığında detaylandırılmış kalıntı grafiği

İncelenen örnek zaman aralığında görülen diğer bir farklı durum ise çoklu regresyon ile ARIMA yönteminin birbirlerinin tersi kalıntı durumları bırakmasıdır. ARIMA yönteminde kalıntılar azalırken çoklu regresyon yönteminde artmakta ve ARIMA yönteminde kalıntılar artarken ise çoklu regresyon yönteminde kalıntılar azalmaktadır.

4.5.3. Ortalama yüzdesel hata oranlarının karşılaştırılması

Yapılan tahminler ve kalıntılarının birbirinden farklı zamanlarda farklı değerlerde olmaları, tahminlerin karşılaştırılmasını yetersizleştirmektedir. Karşılaştırılma için

ortak bir seviyeye getirilmesi gereken sonuçlar için yüzdesel hataların incelenmesi doğru yaklaşım olacaktır. Ortalama yüzdesel hatalar Şekil 4.10 üzerinde gösterilmiştir. Hata oranı en yüksek çoklu regresyon yöntemi uygulanmış Model 7 olmuştur.

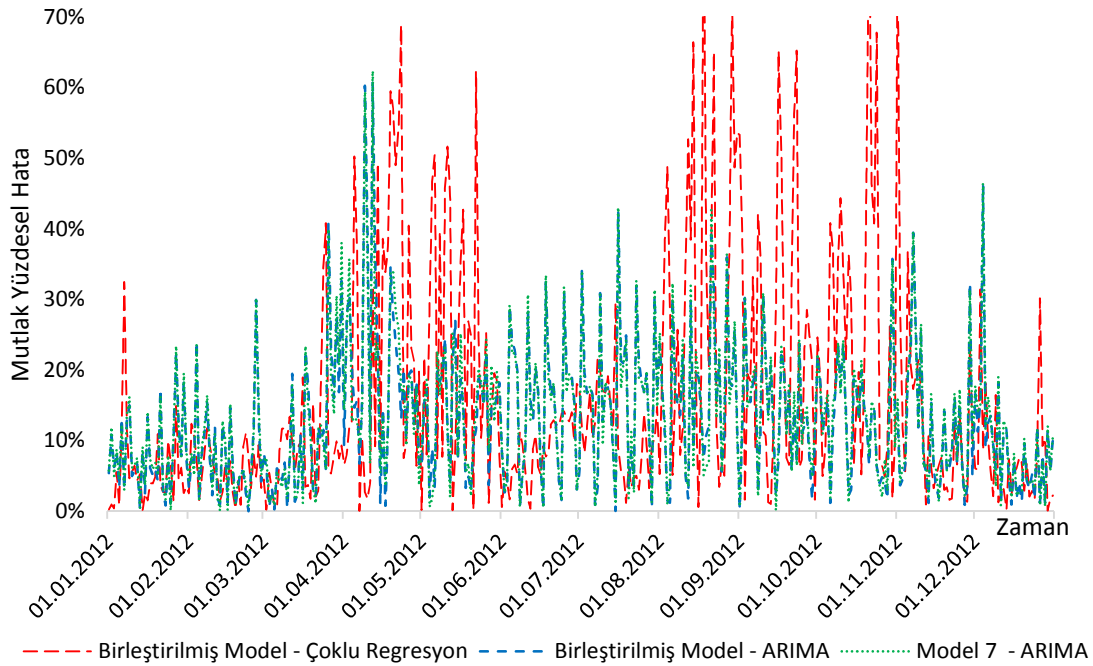


Şekil 4.10. Ortalama yüzdesel hata grafiği

ARIMA yönteminde hata oranlarının genelde negative olması düşük tahmin yapıldığını göstermektedir. Hata oranları incelendiğinde en yüksek hata oranı %231 ile 01.07.2012 tarihinde Model 7 çoklu regresyon yönteminde olmuştur. Sırasıyla model en yüksek hata oranları (tüketimin tahminden düşük olması), çoklu regresyon Birleştirilmiş Model 18.08.2012 tarihinde %90,53, ARIMA Birleştirilmiş Model ve ARIMA Model 7 için 12.04.2012 tarihinde %60,71 ve %62,33 olmuştur. En düşük hata oranları (tüketimin tahminin üzerinde gerçekleşmesi) incelendiğinde Çoklu regresyon yönteminde Birleştirilmiş Model için 22.08.2012 tarihinde -%65, Model 7 için 15.07.2012 tarihinde -%158; ARIMA yönteminde ise Birleştirilmiş Model için 09.04.2012 tarihinde -%60,26, Model 7 için 09.04.2012 tarihinde -%59,42 olmuştur.

Yüzdesel hatalar incelendiğinde Model 7 ile çoklu regresyon yönteminin kullanılması grafikte bulunan diğer modellerin gösterimini etkilemektedir. Bunun

dışında hatanın yönden bağımsız olarak alınması için mutlak değer ile gösterilmesi sonuçların daha doğru yorumlanmasını sağlayacaktır. İlgili değişiklikler ile Şekil 4.11 oluşturulmuştur. Bu grafikte çoklu regresyonun Model 7 için uygulanması çıkartılmıştır ve hataların mutlak değerleri alınmıştır.



Şekil 4.11. Mutlak hata oranları grafiği

Mutlak hata oranları grafiği incelendiğinde (Şekil 4.11) kış ayları ve yaz aylarında hata oranlarının düşük, ilkbahar ve sonbahar aylarında ise yüksek olduğu görülmektedir. Kış aylarında (kasım, aralık, ocak, şubat, mart) doğal gaz tüketimi sıcaklığa bağlıdır. Bu aylarda sıcaklık ile bağlantılı yöntem olan Birleştirilmiş Model ile çoklu regresyon yöntemi kullanılması, model farketmeksizin ARIMA yöntemi tahminlerine göre daha düşük hata oranına sahip olduğu görülmektedir. Aynı şekilde yaz aylarında (haziran, temmuz, ağustos) doğal gaz tüketiminde sıcaklık etkisi olmamaktadır. Bu dönemlerde tüketim genel olarak günler, tatiller ve abone sayılarından etkilenmektedir. Bunun sonucu olarak çoklu regresyon yöntemi Birleştirilmiş Model üzerinde hata oranı daha düşük sonuçlar vermektedir. Bu iki kısım dışında bulunan aylarda (nisan, mayıs, eylül, ekim) her iki modeldeki ARIMA yöntemini tahminlerinin Birleştirilmiş Model çoklu regresyon tahminlerinden daha iyi sonuç verdiği görülmektedir. Bunun sebebi tüketimin geçmiş tüketim ile olan

ilişkisidir. Bir gün önceki tüketim ve hareketli ortalamalar kullanan ARIMA yöntemi ilgili zamanlarda daha doğru sonuç vermektedir.

Ortalama mutlak yüzdesel hatalar incelendiğinde ARIMA yönteminde Model 7 ile %8,67 hata oranı bulunmuşken, Birleştirilmiş Model ile %8,48 hata oranı bulunmuştur [20]. Aynı hata oranı çoklu regresyon için ise Model 7' de %55,10 olurken, Birleştirilmiş Model için %14,38 olmuştur[22].

BÖLÜM 5. SONUÇLAR VE ÖNERİLER

Bu tezin amacı evsel doğal gaz tüketimi ve tahmininde konjonktür bileşeninin etkisini bulmaktır. Bu amaçla zaman serisi şeklinde bulunan veriyi eş mevsimsel zaman dilimlerine bölünmüştür ve mevsimsel modeller oluşturulmuştur. Oluşturulan modellere zaman serileri yöntemlerinden ARIMA yöntemi ile istatistiksel yöntemlerden olan çoklu regresyon yöntemi uygulanmıştır. Böylece konjonktürün evsel doğal gaz tüketimde etkisi bulunmuştur. ARIMA yöntemi uygulamasında p,d,q değerleri 0,0,0 ile 2,2,2 arasında alınmıştır ve hata oranı en düşük tahminler bulunmuştur. Çoklu regresyon yönteminde ise öncelikle bağımsız değişkenlerden anlamlı olan değişkenler her model için bulunmuştur. Daha sonra her model için anlamlı olan değişkenler ile çoklu regresyon yöntemi uygulanıp regresyon denklemleri bulunmuştur.

Konjonktür etkisinden arındırılmış modeller içinde 2012 yılı için yapılan tahminlerin birleştirilmesi ile bir model oluşturulmuş ve Birleştirilmiş Model olarak adlandırılmıştır. Konjonktür etkisinin kaldırılmadığı model ise Model 7 olarak adlandırılmıştır.

Birleştirilmiş Model ve Model 7' de ARIMA ile çoklu regresyon yöntemi sonuçları tüketim tahminleri, kalıntılar ve yüzdesel hatalar olmak üzere 3 durumda incelenmiştir. Tüketim tahminlerinde ARIMA yöntemi ile bulunan sonuçların gerçekleşen tüketimlerden daha düşük olduğu görülmüştür (Şekil 4.7). Model 7 için uygulanmış olan çoklu regresyon yönteminde en yüksek miktarda ve zaman adımında kalıntı oluşmuştur. Bu durum en kötü tahmin sonuçlarının konjonktür etkisi kaldırılmamış model olan Model 7 ile çoklu regresyon kullanımında olduğunu göstermektedir. Hava meteorolojik verilerin dengeli gittiği ve ani değişimlerin fazla yaşanmadığı aylar olan kasım, aralık, ocak, şubat, mart, haziran, temmuz ve ağustosda çoklu regresyon uygulanmış Birleştirilmiş Model ARIMA tahminlerine

göre daha iyi sonuç vermiştir. Mevsim değişimi ayları olan nisan, mayıs, eylül, ekim aylarında ise ARIMA yöntemi Model 7 ve Birleştirilmiş Model için çoklu regresyon yöntemine göre daha doğru tahminler yapmıştır. Bunun en temel sebebi ARIMA yönteminde tahminler kendinden önceki değer ile yüksek ilişki içindedir. Bir önceki değer tahmin edilen değeri etkilemekte ve hata oranını düşürmektedir. Yüzdesele hatalar incelendiğinde Birleştirilmiş Model ARIMA yöntemiyle en düşük hata oranına sahip model elde edilmiştir (Tablo 5.1). Model 7 ile ARIMA yöntemi, Birleştirilmiş Model ile karşılaştırıldığında %2 daha kötü sonuç verdiği görülmüştür. Birleştirilmiş Model ile çoklu regresyon yöntemi ise mevsim değişimindeki aylarda hata oranlarının yüksek olması sonucunda yüzdesele hatasını arttırarak %14,58 olmuştur ve ARIMA tahminlerinden daha yüksek hata oranı bulunmuştur.

Tablo 5.1. Tahmin sonuçları

	Birleştirilmiş Model	Model 7
Çoklu Regresyon	%14,38	%55,10
ARIMA	%8,48	%8,67

Bu çalışmada daha doğru tahmin sonuçları elde edebilmek için konjonktür etkisinin kaldırılması hata oranını azaltmıştır. Burada kullanılan yöntemler dışında öğrenme yöntemlerinden olan yapay sinir ağları ve bulanık mantık teknikleri kullanılabileceği gibi zaman serileri yöntemlerinden ARIMAX (Tahmin edilecek zaman serisinden farklı değişkenler içerebilen ARIMA yöntemi – ARIMA with exogenous variable) ve gözlenemeyen bileşen analizi (unobserved component component models) de kullanılabilir. Kullanılabilecek bu farklı yöntemler daha doğru tahmin yapabilmesi sağlanabilir.

KAYNAKLAR

- [1] MAKRIDAKIS, S., WHEELWRIGHT, S. C., HYNDMAN, R. J., Forecasting Methods and Applications, John Wiley & Sons, 1998.
- [2] ÖZCUMUR, S. Geleceği Tahmin Yöntemleri, İstanbul Sanayi Odası, 1990.
- [3] ORHUNBİLGE, N., Uygulamalı Regresyon ve Korelasyon Analizi, İ. Ü. İşletme Fakültesi Yayın No: 267, 1996.
- [4] O'ROURKE N., HACTHER L., STEPANSKI E. J., A Step by Step Approach to Using SAS® for Univariate & Multivariate Statistics, SAS Institute and Wiley, 2005.
- [5] ORHUNBİLGE, N., Zaman Serileri Analizi Tahmin ve Fiyat İndeksleri, İ. Ü. İşletme Fakültesi Yayın No: 277, 1999.
- [6] YAFFEE, R., MCGEE, M., Introduction to Time Series Analysis and Forecasting with Applications of SAS and SPSS, Academic Press, 2000.
- [7] BOX, G. E. P., JENKINS, G. M., Time Series Analysis Forecasting and Control, Holden-Day, 1976.
- [8] Enerji Piyasaları Denetleme Kurumu (EPDK), 2009 Yılı Doğal Gaz Piyasası Sektör Raporu, EPDK, 2010.
- [9] Enerji Piyasaları Denetleme Kurumu (EPDK), 2011 Yılı Doğal Gaz Piyasası Sektör Raporu, EPDK, 2012.
- [10] Enerji Piyasaları Denetleme Kurumu (EPDK), 2012 Yılı Doğal Gaz Piyasası Sektör Raporu, EPDK, 2013.
- [11] BLANCHARD, M., DESROCHERS, G., Generation of autocorrelated wind speeds for wind energy conversion system studies, Solar Energy, 33, 6, 571-579, 1984.
- [12] HAGAN, M. T., BEHR S. M., The time series approach to short term load forecasting, IEEE Transactions on Power Systems, PWRS-2(3), 785-791, 1987.

- [13] KAMAL, L., JAFRI, Y.Z., Time series models to simulate and forecast hourly averaged wind speed in Quetta, Pakistan, *Solar Energy*, 61, 1, 23-32, 1997.
- [14] BURTIEV, R., GREENWELL, F., KOLIVENKO, V., Time series analysis of wind speed and temperature in Tiraspol, Moldova, *Environmental Engineering and Management Journal*, 12, 1, 23-33, 2013.
- [15] PURVIS, B.E., CHOW, M., Analysis of direct load control using multiple linear regression techniques, *Southeastcon '91*, 1, 454-458, 1991.
- [16] PERRY C., Short-term load forecasting using multiple regression analysis, *Rural Electric Power Conference*, B3/1-B3-8, 1999.
- [17] HONG, T., WANG, P., WILLIS, H.L., A Naïve multiple linear regression benchmark for short term load forecasting, *Power and Energy Society General Meeting*, 1-6, 2011.
- [18] EDIGER, V.S., AKAR, S., UGURLU, B. Forecasting production of fossil fuel sources in Turkey using a comparative regression and ARIMA model, *Energy Policy*, 34,3836-3846, 2006.
- [19] ORHUNBİLGE, N., *Tanımsal İstatistik Olasılık ve Olasılık Dağılımları*, Avcıol Basım Yayın, 2000.
- [20] AKPINAR, M., YUMUSAK N., Forecasting household natural gas consumption with ARIMA model: A case study of removing cycle, *International Conference on Application of Information and Communication Technologies*, 319-324, 2013
- [21] GHARIBVAND, L, Using unobserved components model (UCM) for a stock price fluctuation, *Western Users of SAS Software 2008*, 2008.
- [22] AKPINAR, M., YUMUSAK N., Estimating household natural gas consumption with multiple regression: Effect of cycle, *International Conference on Electronics Computer and Computation*, 192-195, 2013
- [23] ARMSTRONG, J. S., COLLOPY, F., Error measures for generalizing about forecasting methods: Empirical comparisons. *International Journal of Forecasting*, 8.1: 69-80, 1992.

ÖZGEÇMİŞ

Mustafa AKPINAR, 1983 yılında Sakarya' da doğdu. 1994 yılında Figen Sakallıođlu Anadolu Lisesi öğrencisi oldu. 7 yıl süren orta ve lise öğreniminden sonra 2001 yılı Sakarya Üniversitesi, Makina Mühendisliđi bölümünü kazandı. Aynı bölümden 2006 yılında mezun oldu. 2007 yılında Yüksek Lisans programına başladı. 2009 yılı Sakarya Üniversitesi, Bilgisayar Mühendisliđi bölümünü kazandı ve 2011 yılında mezun oldu. Aynı senede Yüksek Lisans programına başladı Ford Tool & Die ve Coşkunöz Metal Form firmalarının Tübitak projelerinde araştırma mühendisi, AGDAŞ Adapazarı Gaz Dağıtım A.Ş. şirketinde Doğal Gaz Talep Tahmin Projesinde araştırma ve uygulama mühendisi olarak çalıştı. Sakarya Üniversitesi Bilgisayar Mühendisliđi bölümünde araştırma görevlisi olarak çalışmaktadır ve halen eğitimine devam etmektedir.