



Dünya dışı ışınimlarla iyileştirilmiş ARIMA, ridge regresyon ve lasso regresyon yöntemlerinin saatlik ışınım tahmininde kullanılması

Use of ARIMA, ridge regression and lasso regression methods enhanced by extraterrestrial radiation in hourly radiation estimation

Burak Arseven^{1,*}, Said Mahmut Çınar²

^{1,2} Afyon Kocatepe Üniversitesi, Elektrik Mühendisliği Bölümü, 03200, Afyonkarahisar Türkiye

Öz

Yeşil enerjiye artan ilgi, özellikle güneş enerjisinden elektrik enerjisi üretimini popüler hale getirmiştir. Güneş enerjisinin kesikli yapısı nedeniyle gelecekteki enerji üretim tahmini, elektrik enerjisinin kesintisiz ve doğaya dost biçimde karşılanması açısından oldukça önemlidir. Bu nedenle, güneş ışınımının tahmin edilmesinin, elektrik üretimi, iletimi, dağıtımı ve tüketimine kadar bütün süreçlerin planlanması ve yönetilmesinde hayati öneme sahip olacağı açıktır. Bu çalışmada Meteoroloji 5. Bölge Müdürlüğü'nün (MBM) Afyon Bölge isimli istasyonundan alınan saatlik bazdaki dört yıllık güneş ışınım verisi üzerinde, zaman serisi analiz yöntemi olan Oto Regresif Entegreli Hareketli Ortalama (Auto Regressive Integrated Moving Average-ARIMA) ve regresyon modelleri olan ridge ve lasso regresyon tahmin modelleri kullanılarak bir yıllık, altı aylık, dört aylık ve iki aylık saatlik güneş ışınımı tahmini işlemi gerçekleştirilmiştir. Ayrıca yöntemlerin yapılarında Dünya Dışı Işınım Filtresi (DDIF) uygulanarak tahmin sonuçları iyileştirilmiştir. Kullanılan yöntemler ile elde edilen tahmin değerleri ve gerçek ışınım değerleri arasındaki ilişkinin yakınlığı grafikler ve istatistiksel yöntemler ile ortaya koyulmaya çalışılmıştır.

Anahtar kelimeler: Yenilenebilir enerji, Güneş ışınım tahmini, ARIMA modeli, Ridge regresyon, Lasso regresyon

1 Giriş

Fosil tabanlı kaynakların gerek doğaya verdikleri zararlar gerekse limitli kaynaklar olmaları yenilenebilir ve sürdürülebilir enerji kaynaklarına olan ilgiyi oldukça artırmıştır. Bu kaynaklardan biri olan güneş, günümüzde en çok tercih edilen enerji kaynaklarından biridir. Güneş enerjisini elektrik enerjisine çevirmek adına yapılan yoğun çalışmalar enerji verimliliğini artırma noktasında tüm dünyaya adeta bir umut ışığı sunmuştur. T.C. Enerji ve Tabii Kaynaklar Bakanlığı verilerine göre Türkiye'de 2014 yılında 40 MW olan güneş enerjisi kurulu gücü 2020 yılı itibarıyla 6667 MW seviyelerine yükselmiştir. Yine bakanlık verilerine göre güneş enerjisinin Türkiye'deki toplam elektrik üretimi içerisindeki yüzdelik payı 2014 yılında %0.01 seviyesinde iken 2020 yılı itibarıyla bu pay %3.66 seviyelerine kadar çıkmıştır [1].

Abstract

Increasing interest in green energy has popularized the generation of electrical energy, especially from solar energy. Due to the discrete nature of solar energy, future energy production forecasting is very important in terms of uninterrupted and environmentally friendly supply of electrical energy. Therefore, it is clear that estimating solar radiation will be vital in planning and managing all processes, from electricity generation, transmission, distribution and consumption. In this study, one-year, six months, four months and two months hourly solar radiation estimation was performed on the hourly basis of four years' solar radiation data obtained four years of hourly data from the station named Afyon Region of the 5th Regional Directorate of Meteorology (MBM), using the Auto Regressive Integrated Moving Average (ARIMA), a time-series analysis method, ridge regression, and lasso regression estimation models, which are regression models. In addition, the estimation results were improved by applying the Extraterrestrial Radiation Filter (DDIF) in the structures of the methods. The closeness of the relationship between the estimated values obtained by the methods used and the actual radiation values has been tried to be revealed with graphics and statistical methods.

Keywords: Renewable energy, Solar radiation estimation, ARIMA model, Ridge regression, Lasso regression

Güneş enerjisinden elektrik enerjisi üretiminde en sık kullanılan aygıtlar güneş panelleridir. Güneş panelleri üzerlerine düşen güneş ışınımını elektrik enerjisine dönüştüren yarı iletken devre elemanlarıdır. Güneş panelinden elde edilecek elektrik enerjisi miktarı, panele gelen güneş ışınımı şiddeti ile doğrudan doğruya ilişkilidir. Güneş ışınımının süresi ve şiddeti esas olarak atmosferik koşullarca belirlenmektedir ve güneş enerjisi kesikli bir yapıya sahiptir. Bu kesikli yapının en temel sebebi ise dünyanın günlük ve mevsimlik hareketleridir. Işınım miktarı güneş enerjisinin kesikli yapısı ile birlikte gölgelenme, nem ve sıcaklık gibi değişkenlerin de devreye girmesi ile doğrusal olmayan bir değişim gösterir. İşte bu kesikli ve doğrusal olmayan yapının mevcut verilerle modellenmesi ile gelecekte oluşması muhtemel durumların tahmini yapılabilir. Yapılan tahminlerin başarısı, gelecekte güneş

* Sorumlu yazar / Corresponding author, e-posta / e-mail: burakarseven89@gmail.com (B. Arseven)
Geliş / Received: 09.02.2023 Kabul / Accepted: 10.05.2023 Yayınlanma / Published: 15.07.2023
doi: 10.28948/ngumuh.1249503

enerjisinden ne kadar elektrik enerjisi üretimi gerçekleşeceği konusunda bir projeksiyon sunar ve arz-talep planlamasının daha doğru bir şekilde yapılmasına yardımcı olur.

Güneş ışınımının doğru tahmini için çeşitli modelleme yaklaşımları, pek çok araştırmacı tarafından kapsamlı bir şekilde kullanılmıştır [2-5]. Akarşlan vd., saatlik bazda bir yıllık güneş ışınımlarını, sıcaklıkları, dünya dışı ışınımları ve zaman içindeki türevlerini çok boyutlu görüntüler olarak değerlendirip, bu görüntüleri önerilen optimal katsayılı doğrusal tahmin filtreleri yardımıyla birbirine bağlayarak ışınım tahmini yapmıştır [6]. Bhardwaj vd., çeşitli iklim koşulları altında gizli Markov modeli ile genelleştirilmiş bulanık modelin bir kombinasyonunu kullanmış ve meteorolojik parametrelere dayalı güneş ışınımının kısa vadeli tahmininde başarılı sonuçlar elde etmiştir [7]. Yine başka bir çalışmada saatlik güneş ışınımı tahmini için Mycielski-Markov modeli adı verilen yeni bir hibrit yöntem kullanılmıştır [8]. Bir yapay sinir ağı (YSA) türü olan doğrusal olmayan Oto Regresif Eksojen (Nonlinear Auto Regressive exogenous-NARX) modelinin kullanılması ile sıcaklık, yağış ve nem miktarlarının değişkenlik gösterdiği durumlar için güneş ışınım tahmini başarıyla gerçekleştirilmiştir [9].

Güneş ışınımı tahmininde regresyon tabanlı yöntemlerin kullanıldığı çalışmalar da mevcuttur. İbrahim ve arkadaşları 2006 yılı için Malezya/Perlis'te güneş ışınımı ve sıcaklık arasındaki ilişkiyi regresyon modeli ile ortaya koyan bir çalışma sunmuşlardır. Çalışmalarının sonucunda doğrusal korelasyon katsayısı değerini 0.7473 bulmuş olmaları güneş ışınımı ile sıcaklık arasında güçlü bir doğrusal ilişki olduğunu göstermektedir [10]. Başka bir çalışmada ise doğrusal, ridge, kement, smoother, kısmi en az, K En Yakın Yomşu (K Nearest Neighbor-KNN) ve Gauss proses regresyon yöntemleri kullanılarak Türkiye'de değişen iklim koşullarına sahip 53 konum için güneş ışınımı tahmini gerçekleştirilmiştir. Sonuç olarak kullanılan yöntemler arasında en iyi sonuç veren Gaussian süreç regresyon yönteminin Ortalama Yanılma Hatası (Mean Bias Error-MBE) değeri 0.274 MJ/m² iken, aynı yöntemin Karesel Ortalama Hataların Karekökü (Root Mean Square Error-RMSE) değeri 2.260 MJ/m² olarak hesaplanmıştır. Ayrıca yine ilgili yöntemin korelasyon katsayısı 0.941 olarak hesaplanmıştır [11].

Bir tür zaman serisi olan güneş ışınımı verilerinin modellenmesi için zaman serisi analiz ve tahmin yöntemleri de literatürde sıkça kullanılmıştır [12-14]. Bir çalışmada, Ajaccio (Fransa) meteoroloji istasyonundan alınan 19 yıllık küresel saatlik ışınım verileri için Oto Regresif (Auto Regressive-AR) ve Oto Regresif Hareketli Ortalama (Auto Regressive Moving Average-ARMA) zaman serisi analiz ve tahmin yöntemlerini kullanarak tahmin gerçekleştirmiş ve sonuçlar istatistiksel ölçütlerle karşılaştırıldığında AR modelinin daha iyi sonuç verdiği ortaya çıkmıştır [15]. Paris'in Alfortville banliyösünün 144 aya ait saatlik güneş ışınımı verileri üzerinde Naive ve ARMA zaman serisi tahmin modellerinin yanı sıra Benzerlik (Similarity-SIM), Destek Vektör Makineleri (Support Vector Machine-SVM) ve YSA yöntemlerini kullanarak tahmin gerçekleştiren Touati ve arkadaşları, ARMA modelinin diğer modellerle

oldukça yakın ve iyi sonuçlar verdiğini göstermiştir [16]. Yine bir başka çalışmada Hindistan'ın belirli bir bölgesindeki uzaktan algılama yöntemi ile elde edilen güneş ışınımı verileri kullanılarak aylık güneş ışınım tahmini yapılmaya çalışılmıştır. Bahsedilen bu çalışmada Mevsimsel ARIMA (Seasonal ARIMA-SARIMA) modelleri, NASA'nın POWER veri arşivinden elde edilene güneşlenme verilerine ait zaman serileri tahmin etmekte kullanılmıştır [17].

Bu makalede, Meteoroloji 5. Bölge Müdürlüğü'nden (MBM) alınan 2018, 2019, 2020 ve 2021 yıllarına ait dört yıllık saatlik bazda güneş ışınım veri seti üzerinde sırasıyla 3 yıllık, 42 aylık, 44 aylık ve 46 aylık verilerle Oto Regresif Entegreli Hareketli Ortalama (Auto Regressive Integrated Moving Average-ARIMA), ridge regresyon ve lasso regresyon yöntemleri eğitilmiş, daha sonra eğitilen modeller ile yine sırasıyla bir yıllık, altı aylık, dört aylık ve iki aylık dönemler için ışınım tahminleri yapılmıştır. Literatürde ARIMA, ridge ve lasso regresyon ile yapılan birçok güneş ışınım tahmin çalışması mevcuttur. Fakat bu çalışmaların hiç birinde Dünya Dışı Işınım (DDI) değerleri ile filtreleme/karşılaştırma yapılmamıştır. Bu çalışmada ise ARIMA, ridge regresyon ve lasso regresyon yöntemlerinin yapılarında yeryüzünde herhangi bir konum için güneş geometrisinden kolayca DDI şiddetleri kullanılarak oluşturulan ve Dünya Dışı Işınım Filtresi (DDIF) adını verdiğimiz filtre modellerin yapılarında kullanılarak güneş ışınımı tahmin performansları iyileştirilmiştir. Burada DDIF sayesinde tahmin değerlerinde yer alan ve normal şartlarda oluşması muhtemel olmayan aşırı yüksek değerler böylece sönmülmüştür. Böylelikle ARIMA, ridge regresyon ve lasso regresyon yöntemlerine güneş ışınımı tahmini başarısını arttıran yeni bir özellik kazandırılmıştır.

Makale dört bölüm halinde düzenlenmiş olup, ikinci bölümde kullanılan veri setine, ARIMA yöntemine, ridge regresyona ve lasso regresyona ait detaylar verilmiştir. Üçüncü bölümde kullanılan yöntemlerle elde edilen tahmin değerleri ve gerçek değerleri karşılaştırır grafikler ve istatistiksel analizler ile yöntemlerin performansı ortaya koyulmaya çalışılmıştır. Son olarak dördüncü bölümde ise tartışma ve sonuçlar verilmiştir.

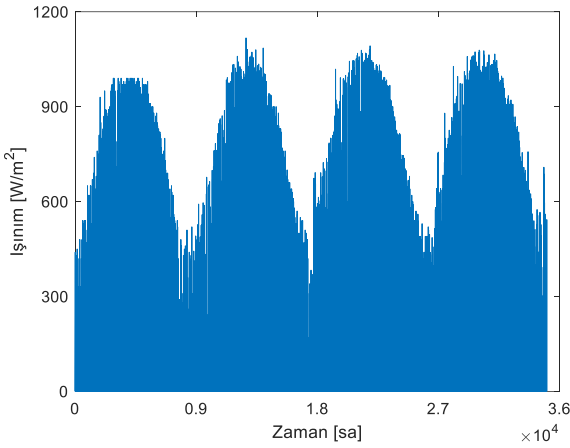
2 Materyal ve metod

Bu bölümde kullanılan veri setine, bu veri seti üzerinde kullanılan tahmin modellerine ve bunlarla birlikte dünya dışı ışınım konusuna ait detaylar verilmiştir.

2.1 Tahmin işleminde kullanılan veri seti

Tahmin işleminde kullanılan veri seti 5. MBM'nin Afyon Bölge İstasyonu'ndan alınan 1 Ocak 2018 tarihinden 31 Aralık 2021 tarihine kadarki (dört yıllık) saatlik güneş ışınımı verilerinden oluşmaktadır. Veri setinin grafiği Şekil 1'de verilmiştir.

Şekil 1'den de görüleceği üzere yaz aylarında güneş ışınımı maksimum düzeyde iken kış aylarında minimum düzeydedir. En yüksek ışınım değeri 13 Haziran 2019 tarihinde saat 12'de 1117 W/m² şeklinde kayıtlara geçmiştir.



Şekil 1. Dört yıllık saatlik güneş ışınımı değerleri

2.2 ARIMA modeli

Zaman serileri zamanla ölçülen gözlem değerlerinin yer aldığı serilerdir. Zaman serilerinin analizini/tahminini yapabilmek için serideki verilerin eşit aralıklarla ve belirli sıklıklarla alınmış olması şarttır. ARIMA modelleri, zaman serisi modelleme ve tahmininde kullanılan matematiksel modellerdir. Esasen ARIMA modelinin temelleri Box-Jenkins yaklaşımına dayanır [17]. AR ve hareketli ortalama (Moving Average – MA) modellerinin entegre edilmesi ile ortaya çıkmış bir modeldir. AR modelinde şimdiki zaman serisi gözlemi, önceki zaman serisi gözlemleri ile bir miktar hata teriminin doğrusal bir kombinasyonudur. p . dereceden bir AR modelinin denklemi aşağıda Denklem (1)'de verilmiştir.

$$X_t = \sum_{i=1}^p \phi_i X_{t-i} + \varepsilon_t \quad (1)$$

Burada;

p : AR modeli derecesi,

ε_t : tahmin hatası,

$\phi = (\phi_1, \phi_2, \dots, \phi_p)$: AR modeli katsayıları vektörüdür.

MA modelinde ise mevcut zaman serisi değeri geçmiş hataların bir fonksiyonudur. q . dereceden bir MA modelinin denklemi aşağıda Denklem (2)'de verilmiştir.

$$X_t = \sum_{j=0}^q \theta_j \varepsilon_{t-j} \quad (2)$$

Burada;

q : MA modeli derecesi,

$\theta = (\theta_1, \theta_2, \dots, \theta_q)$: MA modeli katsayıları vektörü,

$\varepsilon = (\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-q})$: geçmiş hata değerleridir.

AR ve MA modelleri teker teker kullanılabilirler gibi, birleştirilerek ARMA modeli olarak da kullanılabilir. p . Ve q . derecelerden bir ARMA modeline ait Denklem (3) şu şekildedir:

$$X_t = \sum_{i=1}^p \phi_i X_{t-i} + \sum_{j=0}^q \theta_j \varepsilon_{t-j} \quad (3)$$

AR, MA ve ARMA modelleri durağan zaman serilerinde kullanılabilirler. Fakat güneş ışınımı veri seti gibi durağan olmayan zaman serilerinde öncelikle bir durağanlaştırma işlemi gerçekleştirilmelidir. Durağan olmayan bir zaman serisinin durağanlaştırılıp daha sonra bu seriye ARMA modeli uygulanmış haline ARIMA modeli denir. ARIMA modeli genellikle ARIMA (p, d, q) olarak gösterilir. Burada p oto-regresif sürecin derecesini, d fark alma işlemi derecesini ve q ise hareketli ortalama sürecinin derecesini temsil eder. Herhangi bir ARIMA(p, d, q) modeline ait Denklem (4) şu şekildedir:

$$X_t = \sum_{i=1}^p \phi_i X_{t-i} S^d + \sum_{j=0}^q \theta_j \varepsilon_{t-j} \quad (4)$$

Burada;

S^d : fark alma operatörüdür.

2.3 Ridge regresyon

Lineer regresyon yöntemi verilere en çok uyum sağlayan bir çizgi denklemi oluşturmayı hedefler. Lineer regresyonda, tahmin edicilerin sayısı gözlemlerin sayısını aştığında model hiçbir değer hesaplayamaz. Bu durum bir modele gereğinden fazla uymaya ve elde veri yok iken çözüm bulamamaya, yani başka bir deyişle tahmin yapamamaya yol açar. Yine lineer regresyonda veriler arasında çoklu bağıntılar var ise yöntem problemler çıkarabilir. Ridge regresyon ise tüm bu sorunları önler [18-19].

Ridge regresyonda verilerin bir kısmı doğru modeli oluşturmak adına eğitim verisi olarak kullanılırken kalan diğer kısmı ise oluşturulan modelin performansını ortaya koyabilmek için test verisi olarak kullanılır. Ridge regresyonun arkasında yatan ana fikir, eğitim verilerine de uymayan yeni bir çizgi denklemi elde etmektir. Diğer bir deyişle, ridge regresyon modelinin verilerle nasıl uyum sağlayacağı ile alakalı lineer regresyon modeline bir miktar sapma değeri eklenir. Bu sapma değerinin eklenmesi ile varyans değerinde önemli ölçüde bir azalma gerçekleşir. Yani ridge regresyon çizgisi lineer regresyona göre eğitim verilerine daha kötü uyum sağlasa da test verilerinde daha iyi uzun vadeli tahminler gerçekleştirir. Ridge regresyon modeline ait Denklem (5) şu şekildedir:

$$\tilde{X}_i = \beta_0 + \beta_1 X_i + \lambda(\beta_1^2) \quad (5)$$

Burada;

\tilde{X}_i tahmin değeri,

β_0 : doğrunun y-eksenini kestiği nokta,

β_1 : doğrunun eğimi,

$\beta_0 + \beta_1 X_i$: bileşeni lineer regresyon bileşeni,

$\lambda(\beta_1^2)$: ridge cezası,

λ : ceza şiddeti çarpanıdır.

2.4 Lasso regresyon

Lasso regresyonda, ridge regresyondaki gibi verilerin bir kısmı doğru modeli oluşturmak adına eğitim verisi olarak kullanılırken kalan diğer kısmı ise oluşturulan modelin performansını ortaya koyabilmek için test verisi olarak kullanılır. Lasso regresyon modeli karakteristik olarak ridge regresyon ile aynı mantığa uyarak oluşturulur. Burada lasso regresyonu, ridge regresyondan ayıran kısım lasso cezasıdır. Lasso regresyon modeline ait [Denklem \(6\)](#) şu şekildedir:

$$\tilde{X}_i = \beta_0 + \beta_1 X_i + \lambda |\beta_1| \quad (6)$$

Burada;

\tilde{X}_i tahmin değeri,

β_0 : doğrunun y-eksenini kestiği nokta,

β_1 : doğrunun eğimi,

$\beta_0 + \beta_1 X_i$: bileşeni lineer regresyon bileşeni,

$\lambda |\beta_1|$: lasso cezası,

λ : ceza şiddeti çarpanıdır.

2.5 Dünya dışı ışıınım

Dünya dışı ışıınım (DDI), atmosferin tepesindeki güneşin yoğunluğudur. Dünya'nın eliptik yörüngesi nedeniyle DDI'nın aldığı değer yıl boyunca değişir fakat her yıl için yılın aynı günlerinin aynı saatlerinde aldıkları değerler her zaman aynıdır. Bu sebeple DDI değerini bir denklem vasıtasıyla hesaplamak mümkündür. DDI değerleri saatlik bazda ampirik olarak aşağıda verilen [Denklem \(7\)](#) ile temsil edilebilir:

$$I_0 = I_{sc} \left[1 + 0.33 \cos \left(\frac{360n}{365} \right) \right] (\cos \emptyset \cos \partial \cos \omega + \sin \emptyset \sin \partial) \quad (7)$$

Burada;

I_{sc} : güneş sabiti,

\emptyset : enlem açısı,

∂ : deklinasyon açısı,

ω : güneşin saat açısı,

n : gün değeridir.

Güneş'in kendi iç reaksiyonları sonucunda iç yapısında bulunan hidrojen elementi devamlı surette füzyona uğrayarak helyuma dönüşür ve gerçekleşen bu dönüşümle birlikte çeşitli dalga boylarında enerji açığa çıkar. Bu enerji yeryüzüne ulaşana kadar 150 milyon kilometre yol alır ve bu süreçte bahsedilen bu enerjinin sadece iki milyarda biri yeryüzüne ulaşabilir. Dünya atmosferinin üstüne ulaşan güneş ışığının birim zamandaki enerji yoğunluğu ile ilgili yapılan hesaplamalara göre I_{sc} değeri 1367 W/m^2 'dir [20].

Enlem açısı (\emptyset), ekvator düzlemi ile yerden Dünya'nın merkezine olan radyal çizgi arasındaki açıdır. \emptyset değeri -90° ile 90° arasında değişir. \emptyset değeri hesaplanacak yer kuzey yarımkürede ise \emptyset pozitif değerler, güney yarımkürede ise \emptyset negatif değerler alır.

Güneş ışınlarının geldiği doğrultu ile ekvator düzlemi arasındaki açı değeri ise deklinasyon açısını (∂) verir. ∂ , -23.45 ila 23.45 arasındadır ve \emptyset değerinin hesaplanmasında da olduğu gibi kuzey yarımküredeki yerler için pozitif,

güney yarımkürede yerler için ise negatif değer alır. ∂ değeri aşağıdaki [Denklem \(8\)](#) ile temsil edilebilir:

$$\partial = 23.45 \sin \left(360 \frac{284 + n}{365} \right) \quad (8)$$

Güneşin saat açısı (ω), zamanın açılal ölçüsüdür ve bir saat 15° boylama eşittir. ω değeri gün doğumundan öğlen saat 12'ye kadar negatif, sonrasında ise pozitif değerler alır. ω değeri aşağıdaki [Denklem \(9\)](#) ile temsil edilebilir:

$$\omega = \left(\frac{360}{24} \right) (h - 12) \quad (9)$$

Burada;

h : Saat değeridir.

3 Bulgular ve tartışma

Bu bölümde ilk olarak en uygun ARIMA, ridge ve lasso regresyon modellerinin seçiminin nasıl yapıldığı üzerinde durulmuştur. Daha sonra tahmin edilen verilerin alabileceği maksimum değer ile ilgili yapılan filtreleme işleminin ayrıntıları verilmiştir. Son olarak gerçek değerler ile tahmin değerleri arasındaki benzerlik grafikler ve istatistik analizler ile ortaya koyulmuştur.

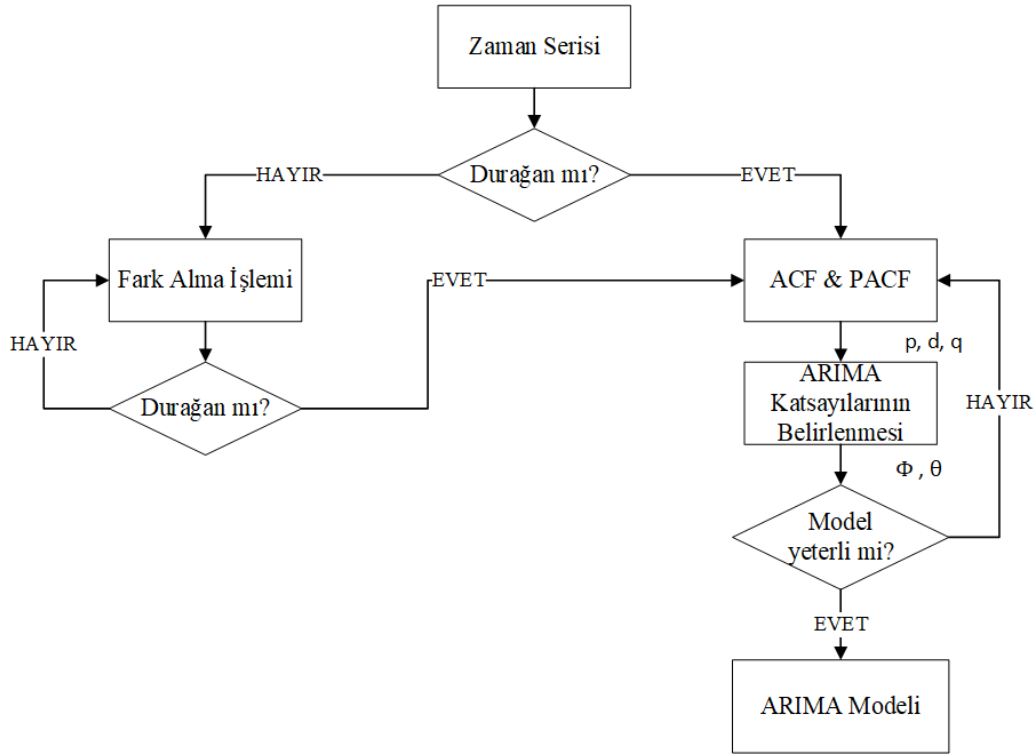
3.1 ARIMA modeli seçimi

ARIMA modelleri ARIMA(p,d,q) olarak gösterilirler. Burada parantez içinde verilen değerler sırasıyla AR derecesini, fark alma işlemi derecesini (I) ve MA derecesini temsil ederler. ARIMA modelini belirlemek için kullanılan akış şeması [Şekil 2](#)'de verilmiştir.

En iyi ARIMA modelini seçmek için gerçek veriler üzerinde öncelikle trend ve mevsimsellik analizi yapmak gerekir. Eğer verilerde bir mevsimsellik veya artan ya da azalan bir trend durumu söz konusu ise veriler üzerinde ilk olarak fark alma işlemi yapılarak veriler üzerindeki bu etki ortadan kaldırılır. Fark alma işlemi genel olarak bir kez, ya da maksimum düzeyde ise iki kez yapılır. Daha sonra otokorelasyon fonksiyonu (Auto Correlation Function-ACF) ve kısmi otokorelasyon fonksiyonu (Partial Auto Correlation Function-PACF) grafiklerine bakılarak AR ve MA dereceleri kestirilmeye çalışılır. Bu süreç esasen oldukça tecrübe isteyen sezgisel bir süreçtir.

[Şekil 3](#)'te yer alan üç güne ait saatlik ışıınım verilerinin yer aldığı grafikten de görüleceği üzere ışıınım verileri belli bir trendi olmayan mevsimsel verilerdir. Bu durumda bu tür bir veri seti ile güneş ışıınımı tahmini yapmak isteniyorsa fark alma işlemi mutlaka uygulanmalı ve mevsimsellik etkisi ortadan kaldırılmalıdır.

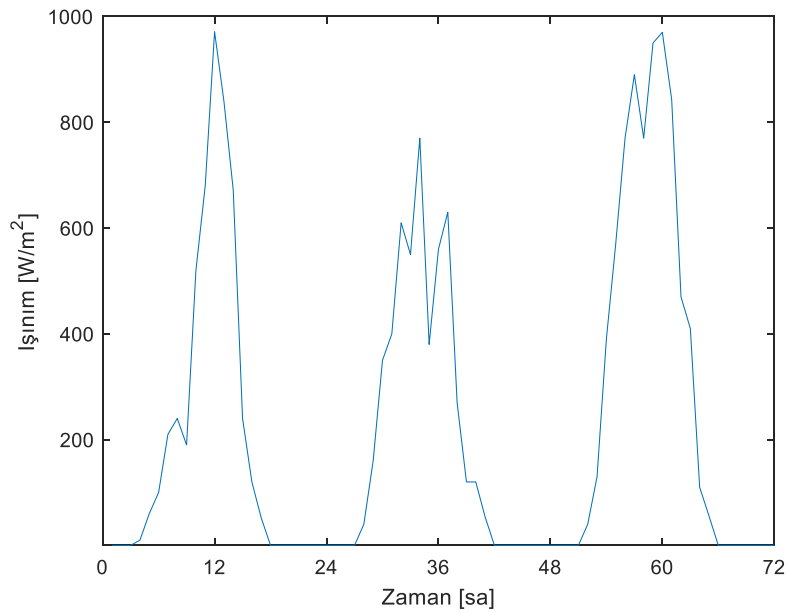
p ve q değerlerini belirlemek için ise ACF ve PACF grafiklerine bakılır. ACF grafiği şimdiki veri ile diğer tüm verilerin korelasyonunu verir. ACF grafiğine bakılarak MA modeli derecesi hakkında, yani q değeri hakkında, yorum yapılabilir. PACF grafiği bir önceki veri ile şimdiki verinin diğer tüm verilerden arındırılmış olarak korelasyonunu verir. PACF grafiğine bakılarak ise AR modeli derecesi hakkında, yani p değeri hakkında, yorum yapılabilir.



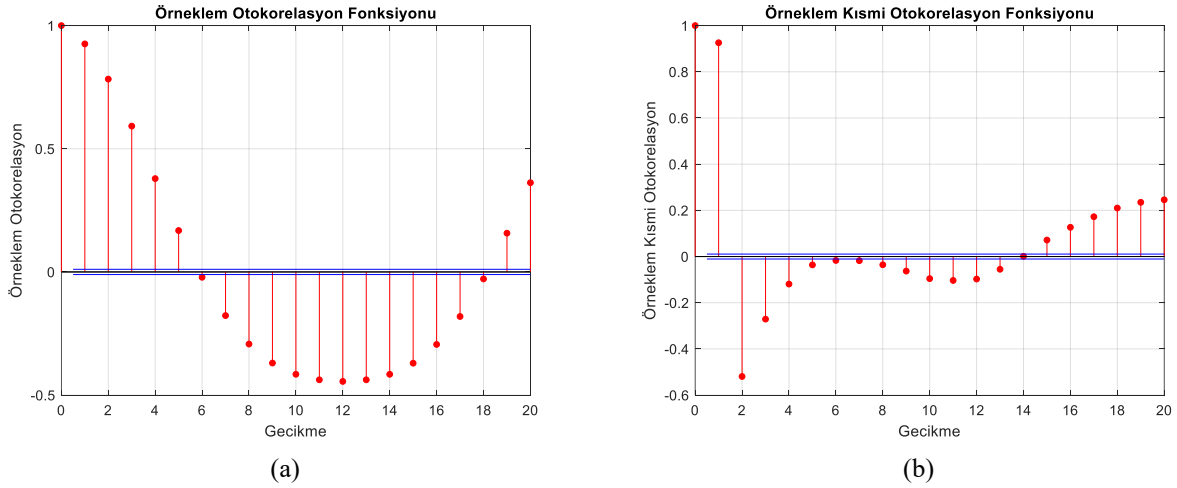
Şekil 2. ARIMA modelinin belirlenmesi aşamaları

Grafiklere bakıldığında anlamlı gecikme sayılarının oldukça fazla olması sebebiyle bu çalışmada p ve q değerlerine deneme-yanılma yöntemiyle karar verilmiştir. Yapılan denemeler sonucunda model olarak ARIMA(2,1,1) modelinin en iyi tahmini gerçekleştirdiği belirlenmiştir. Ayrıca gerçek değerlerin yer aldığı veri setindeki sırasıyla ilk 36 aylık (3 yıllık), 42 aylık, 44 aylık ve 46 aylık veriler AR

ve MA katsayı vektörlerinin değerlerini belirlemek için eğitim veri seti olarak kullanılmış ve elde edilen ARIMA katsayılar modelde kullanılarak yine sırasıyla 12 aylık (1 yıllık), 6 aylık, 4 aylık ve 2 aylık tahmin işlemleri gerçekleştirilerek test verileri ile karşılaştırma yapılmıştır.



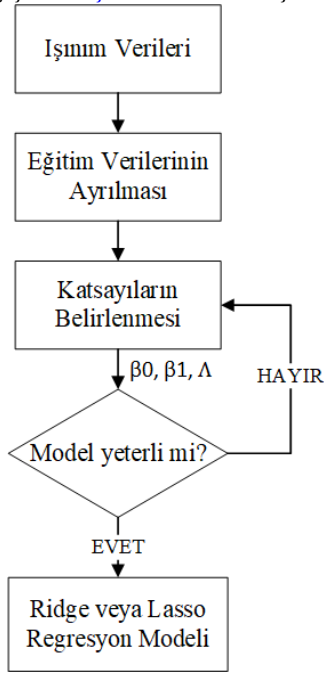
Şekil 3. Rastgele seçilmiş 3 güne ait saatlik ışınım değerleri



Şekil 4. ACF grafiği (a) PACF grafiği (b)

3.2 Ridge ve lasso regresyon parametrelerinin belirlenmesi

Ridge ve lasso regresyon parametrelerini belirlemek için kullanılan akış şeması Şekil 5'te verilmiştir.



Şekil 5. Ridge veya lasso regresyon modelinin belirlenmesi aşamaları

MATLAB programında yazılan kod yardımıyla 46 aylık test verileriyle elde edilen ridge regresyon modeli için β_0 değeri 14.5648 ve β_1 değeri de 0.9257 olarak elde edilmiştir. Yine 46 aylık test verileriyle elde edilen lasso regresyon modeli için ise lasso regresyon için β_0 değeri 19.2315 ve β_1 değeri de 0.9255 olarak elde edilmiştir λ değeri ise 0.5 olarak belirlenmiştir. Bu durumda ridge regresyon modeli Denklem (10)'daki hale, lasso regresyon modeli ise Denklem (11)'deki hale bürünmüştür.

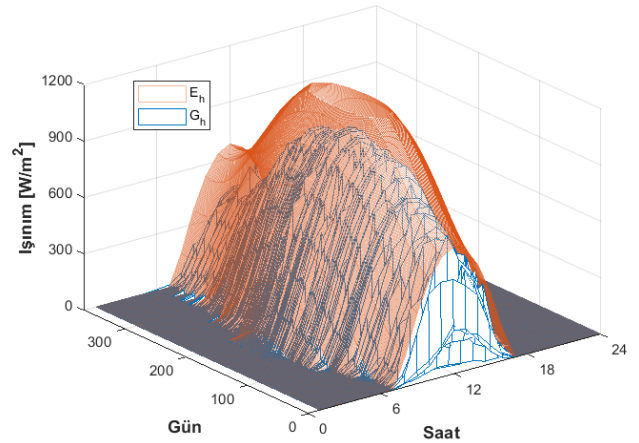
$$\tilde{X}_t = 14.5648 + 0.9257X_t + 0.5(0.9257^2) \quad (10)$$

$$\tilde{X}_t = 19.2315 + 0.9255X_t + 0.5|0.9255| \quad (11)$$

Şekil 5'teki ridge ve lasso regresyon modelleri akış şeması kullanılarak 1 yıllık tahmin işlemi gerçekleştirilerek, tahmin sonuçları test verileri ile karşılaştırılmıştır.

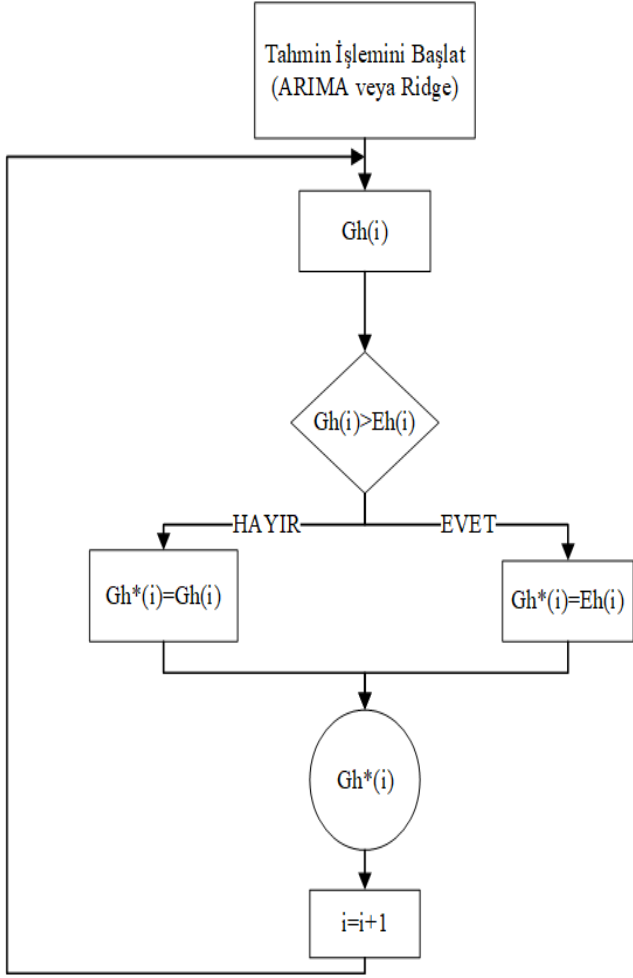
3.3 Dünya dışı ışınım filtresinin uygulanması

Dünya'nın hiçbir yerinde herhangi bir anda alınan ışınım değeri DDI değerinden büyük olamaz. Bu durumdan yola çıkarak bu çalışmada öncelikle gerçek veriler ile DDI verileri arasında bir katsayı değeri belirlenmiş ve tahmin değerlerinin bu katsayı değeri ile çarpılmış olan DDI değerinin üzerinde çıkmaması için Dünya Dışı Işınım Filtresi (DDIF) uygulanmıştır. Şekil 6'da 2020 yılına ait saatlik güneş ışınımı verileri (G_h) ve aynı yıl için saatlik DDI (E_h) değerlerinin yer aldığı grafik verilmiştir.



Şekil 6. 2020 yılına ait saatlik dünya dışı ışınım verilerinin iki boyutlu yüzey (veya ağ) grafiği

Yukarıdaki şekilden de görüleceği üzere DDI değeri ölçülen değerlerden (gerçek değerler) oldukça yüksektir. Bu yüzden DDI değerleri bu çalışma için optimum değer olarak bulunan 0.85 değeri ile çarpılmıştır. DDIF'nin uygulanması ile ilgili akış diyagramı Şekil 7'de verilmiştir.



Şekil 7. DDIF akış diyagramı

Burada:

$G_h(i)$: ARIMA veya ridge regresyon ile gerçekleştirilen i . saate ait tahmin değerini,

$E_h(i)$: DDI vektörünün i . saate ait değerini,

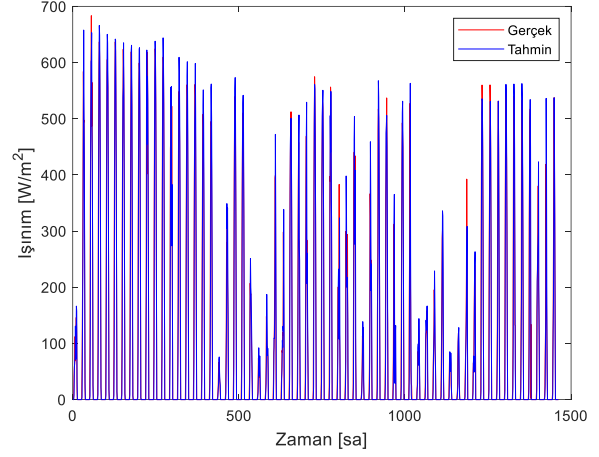
$G_h^*(i)$: DDIF uygulanmış nihai tahmin vektörünün i . saate ait tahmin değerini, ifade eder.

Şekil 7'den de görüleceği üzere ARIMA veya ridge regresyon yöntemleri ile ilk olarak $G_h(i)$ bulunmuştur. Daha sonra elde edilen $G_h(i)$ değeri $E_h(i)$ değeri ile karşılaştırılıyor, yani DDIF'den geçirilmiş ve $G_h(i)$ değerinin $E_h(i)$ değerinden büyük olması durumunda $G_h(i)$ değeri $E_h(i)$ değerine eşitlenerek $G_h^*(i)$ elde edilmiştir. Elde edilen bu $G_h^*(i)$ değerleri bir sonraki tahmin değerinin ($i + 1$). saatin) elde edilmesinde kullanılmıştır. Böylece ARIMA, ridge ve lasso regresyon modellerinin tahmin döngüsüne eklenen DDIF tahmin değerlerinin DDI'ya geçmesini engellemiş ve bir sonraki saatin tahmin değeri için de daha başarılı bir referans değeri sunmuştur.

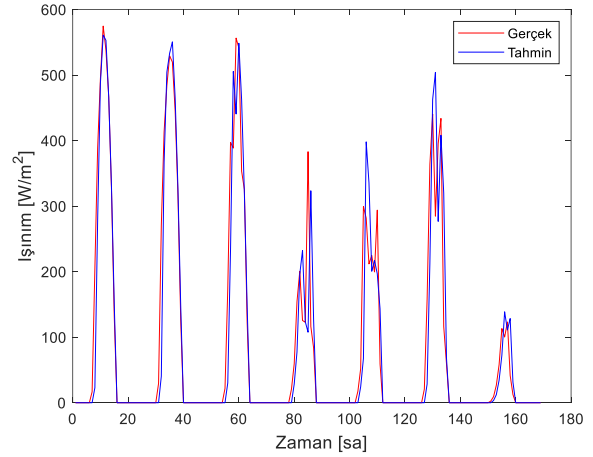
3.4 Tahmin sonuçları

Tahminin performansını incelemek için öncelikle gerçek değerler ile tahmin değerlerinin grafiklerinden yararlanılmıştır. Şekil 8'de gerçek değerlerin ve ARIMA(2,1,1) modeli kullanılarak bulunan 2 aylık tahmin

değerlerinin yer aldığı grafik verilmiştir. Şekil 9'da ise ARIMA(2,1,1) modeli ile gerçekleştirilen tahminin başarısını daha görünür kılabilmek adına Şekil 8'deki grafiğe yakınlaştırma işlemi uygulanmıştır.



Şekil 8. 2 aylık gerçek veriler ve ARIMA(2,1,1) modeli ile elde edilen tahmin değerleri

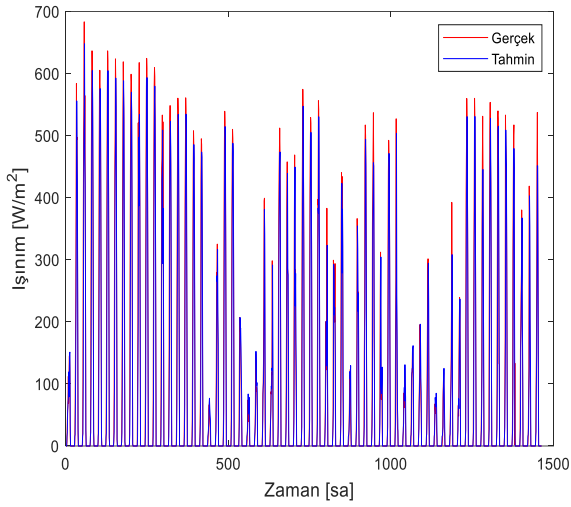


Şekil 9. Bir haftalık gerçek değerler ve ARIMA(2,1,1) modeli ile elde edilen tahmin değerleri

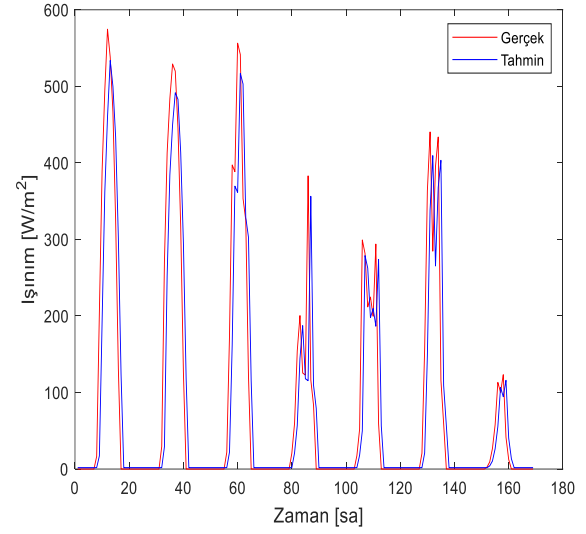
Şekil 10'da gerçek değerlerin ve ridge regresyon ile bulunan 2 aylık tahmin değerlerinin yer aldığı grafik verilmiştir. Şekil 11'de ise ridge regresyon ile gerçekleştirilen tahminin başarısını daha görünür kılabilmek adına Şekil 11'deki grafikteki bir haftalık değerlere yakınlaştırma işlemi uygulanmıştır.

Şekil 12'de gerçek değerlerin ve lasso regresyon ile bulunan 2 aylık tahmin değerlerinin yer aldığı grafik verilmiştir. Şekil 13'te ise lasso regresyon ile gerçekleştirilen tahminin başarısını daha görünür kılabilmek adına Şekil 13'teki grafikteki bir haftalık değerlere yakınlaştırma işlemi uygulanmıştır.

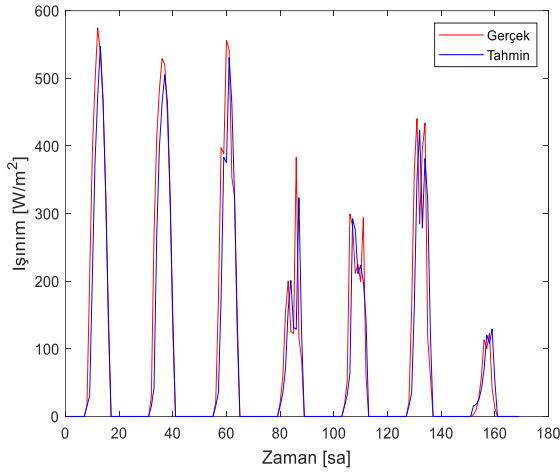
Şekillerden de görüleceği üzere ARIMA(2,1,1) yöntemi, ridge ve lasso regresyon yöntemleri kullanılarak elde edilen tahmin sonuçları gerçek değerler ile oldukça yüksek benzerlik göstermektedir.



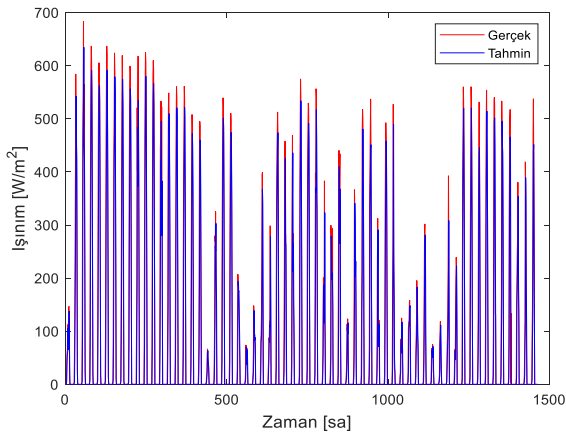
Şekil 10. 2 aylık gerçek veriler ve ridge regresyon modeli ile elde edilen tahmin değerleri



Şekil 13. Bir haftalık gerçek değerler ve lasso regresyon modeli ile elde edilen tahmin değerleri



Şekil 11. Bir haftalık gerçek değerler ve ridge regresyon modeli ile elde edilen tahmin değerleri



Şekil 12. 2 aylık gerçek veriler ve lasso regresyon modeli ile elde edilen tahmin değerleri

Kullanılan yöntemlerin ne derece başarılı olduğunu ortaya koymak adına en sık kullanılan performans ölçütü RMSE ölçütüdür. RMSE, artıkların (tahmin hatalarının) standart sapmasıdır. RMSE artıkların ne kadar yayıldığına bir ölçüsüdür ve bize tahmin değerlerinin en uygun form etrafında ne kadar yoğun olduğu hakkında nicel bir bilgi sağlar. RMSE formülü [Denklem \(12\)](#)'de verilmiştir.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\tilde{x}_i - x_i)^2}{n}} \quad (12)$$

Normalleştirilmiş karesel kök ortalama hataların karekökü (Normalized Root Mean Square Error-NRMSE) farklı ölçeklerdeki modeller arasında karşılaştırmayı kolaylaştırır. Bu nedenle, NRMSE, model tarafından tipik olarak çözülen genel aralığın bir kısmı olarak yorumlanabilir. [Denklem \(13\)](#)'te formülü yer alan NRMSE değeri performans değerlendirme ölçütü olarak kullanılabilir.

$$NRMSE = \frac{\sqrt{\frac{\sum_{i=1}^n (\tilde{x}_i - x_i)^2}{n}}}{\max(x_i)} \quad (13)$$

Yine performansı ölçmede kullanılacak bir başka performans metriği normalleştirilmiş ortalama mutlak hata (Normalized Mean Absolute Error-NMAE) ölçütüdür. Bu metrik, farklı ölçeklerdeki veri kümelerinin ortalama mutlak hatasına ilişkin karşılaştırmayı kolaylaştırmak için kullanılır. Normalleştirilmenin bir ortalaması olarak, model performans değerlendirme aracı, ölçülen verilerin ortalamasını kullanır. NMAE formülü [Denklem \(14\)](#)'te verilen şekildedir.

$$NMAE = \frac{1}{n} \sum_{i=1}^n \frac{|\tilde{x}_i - x_i|}{\max(x_i)} \quad (14)$$

Tablo 1. Yöntemlerin RMSE, NRMSE ve NMAE değerleri (bir yıllık tahmin sonuçları için)

	ARIMA(2,1,1)		Ridge Regresyon		Lasso Regresyon	
	DDIF Yok	DDIF Var	DDIF Yok	DDIF Var	DDIF Yok	DDIF Var
RMSE	87.4627	70.8897	99.7243	83.5277	100.5084	87.7335
NRMSE	0.0811	0.0657	0.0924	0.0774	0.0932	0.0813
NMAE	0.0491	0.0340	0.0664	0.0434	0.0614	0.0459

Tablo 2. Yöntemlerin RMSE, NRMSE ve NMAE değerleri (altı aylık tahmin sonuçları için)

	ARIMA(2,1,1)		Ridge Regresyon		Lasso Regresyon	
	DDIF Yok	DDIF Var	DDIF Yok	DDIF Var	DDIF Yok	DDIF Var
RMSE	78.0489	60.9495	94.6652	77.3289	95.5272	82.0097
NRMSE	0.0734	0.0623	0.0890	0.0727	0.0898	0.0771
NMAE	0.0457	0.0300	0.0651	0.0408	0.0598	0.0435

Tablo 3. Yöntemlerin RMSE, NRMSE ve NMAE değerleri (dört aylık tahmin sonuçları için)

	ARIMA(2,1,1)		Ridge Regresyon		Lasso Regresyon	
	DDIF Yok	DDIF Var	DDIF Yok	DDIF Var	DDIF Yok	DDIF Var
RMSE	70.1925	54.8147	81.7854	66.0007	82.1229	70.0715
NRMSE	0.0751	0.0586	0.0875	0.0706	0.0879	0.0750
NMAE	0.0448	0.0304	0.0628	0.0383	0.0557	0.0408

Tablo 4. Yöntemlerin RMSE, NRMSE ve NMAE değerleri (iki aylık tahmin sonuçları için)

	ARIMA(2,1,1)		Ridge Regresyon		Lasso Regresyon	
	DDIF Yok	DDIF Var	DDIF Yok	DDIF Var	DDIF Yok	DDIF Var
RMSE	55.7012	41.8375	62.9689	48.7017	62.6598	52.2766
NRMSE	0.0815	0.0562	0.0931	0.0713	0.0917	0.0765
NMAE	0.0470	0.0322	0.0634	0.0375	0.0557	0.0404

Yukarıdaki tablolarda ARIMA(2,1,1), ridge ve lasso regresyon yöntemleri ile yapılan tahminler için DDIF'nin kullanıldığı ve kullanılmadığı durumların RMSE, NRMSE ve NMAE değerleri verilmiştir. Tablo 1'de 1 yıllık, Tablo 2'de altı aylık, Tablo 3'te dört aylık ve Tablo 4'te ise 2 aylık tahmin sonuçlarına ait değerler verilmiştir. Değerler incelendiğinde eğitim verisinin test verisine oranı yükseldikçe daha iyi tahmin sonuçları elde edilmiştir. Kullanılan yöntemler arasında ise ARIMA(2,1,1) yöntemi ile yapılan tahmin işleminin ridge ve lasso regresyon yöntemlerine göre çok daha başarılı olduğu görülmüştür. Ayrıca yine DDIF kullanıldığında her üç yöntemde de tahmin sonuçları bir miktar daha iyileşmiştir. Sonuç olarak ise en başarılı sonuçların DDIF'nin kullanıldığı ARIMA(2,1,1) yöntemiyle elde edildiği görülmüştür.

4 Sonuçlar

Bu çalışma ile birlikte zaman serisi tahmin modellerinden biri olan ARIMA modeli ve regresyon yöntemlerinden olan ridge ve lasso regresyon yöntemleri kullanılarak 1 yıllık, 6 aylık, 4 aylık ve 2 aylık güneş ışınımı tahmini başarıyla gerçekleştirilmiştir. Çalışmayı diğer benzer çalışmalardan farklı kılan özelliklerinden birisi bu çalışmada DDIF adı verilen bir filtrenin yöntemlerin iç yapılarında/tahmin döngülerinde kullanılmış olmasıdır. Bu filtreleme işlemi ile birlikte modellerin tahmin performansı daha da artırılmıştır. Ayrıca tahmin modellerinin gün içerisinde bulutlanma veya başka sebeplerle meydana gelen

ani ışınım değişimlerini oldukça başarılı bir şekilde yakalaması da çalışmanın bir başka başarısıdır.

Bundan sonraki çalışmalarda kullanılacak veri boyutunun artırılması ile daha uzun vadeli tahminler gerçekleştirmek ve bu yöntemlerin yanı sıra başka bir takım makine öğrenmesi tabanlı yöntemler, derin öğrenme tabanlı yöntemler, sezgisel algoritmalar gibi başka birtakım yöntemler ile hibrit tahmin modelleri oluşturularak tahmin performansını daha da yükseltmek amaçlanmaktadır.

Teşekkür

Bu çalışmada kullanılan verilerin Meteoroloji 5. Bölge Müdürlüğü'nden alınmış olması sebebiyle kendilerine teşekkür ederiz. Ayrıca makalenin inceleme ve değerlendirme aşamalarında yapmış oldukları katkılardan dolayı editör ve hakemlere de teşekkür ederiz.

Çıkar çatışması

Yazarlar çıkar çatışması olmadığını beyan etmektedir.

Benzerlik oranı (iThenticate): %11

Kaynaklar

- [1] Güneş. <https://enerji.gov.tr/eigm-yenilenebilir-enerji-kaynaklar-gunes>, Erişim 6 Şubat 2023.
- [2] R. H. Inman, H. T. Pedro and C. F. Coimbra, Solar forecasting methods for renewable energy integration. Progress in Energy and Combustion Science, 39 (6), 535-576, 2013. <https://doi.org/10.1016/j.pecs.2013.06.002>.

- [3] R. C. Deo and M. Şahin, Forecasting long-term global solar radiation with an ANN algorithm coupled with satellite-derived (MODIS) land surface temperature (LST) for regional locations in Queensland. *Renewable and Sustainable Energy Reviews*, 72, 828-848, 2017. <https://doi.org/10.1016/j.rser.2017.01.1141>.
- [4] R. Blaga, A. Sabadus, N. Stefu, C. Dughir, M. Paulescu and V. Badescu, A current perspective on the accuracy of incoming solar energy forecasting. *Progress in Energy and Combustion Science*, 70, 119-144, 2019.
- [5] A. Masoom, Y. Kashyap and A. Bansal, Solar radiation assessment and forecasting using satellite data, *Advances in Solar Energy Research*. *Advances in Solar Energy Research*, Springer, London, pp. 45-71, 2019.
- [6] E. Akarslan, F. O. Hocaoglu and R. Edizkan, A novel M-D (multi-dimensional) linear prediction filter approach for hourly solar radiation forecasting. *Energy*, 73, 978-986, 2014. <https://doi.org/10.1016/j.energy.2014.06.113>.
- [7] S. Bhardwaj, V. Sharma, S. Srivastava, O. S. Sastry, B. Bandyopadhyay, S. S. Chandel and J. R. P. Gupta, Estimation of solar radiation using a combination of hidden Markov model and generalized fuzzy model. *Solar Energy*, 93, 43-54, 2013.
- [8] F. O. Hocaoglu and F. Serttas, A novel hybrid (Mycielski-Markov) model for hourly solar radiation forecasting. *Renewable Energy*, 108, 635-643, 2017. <https://doi.org/10.1016/j.renene.2016.08.058>.
- [9] A. Sarı, M. Aktaş, A. Yönetken ve R. Doğan, Güneş ışınım tahmininde NARX modelinin uygulanması. *Afyon Kocatepe Üniversitesi Uluslararası Mühendislik Teknolojileri ve Uygulamalı Bilimler Dergisi*, 4, 1-6, 2021.
- [10] S. Ibrahim, I. Daut, Y. M. Irwan, M. Irwanto, N. Gomesh and Z. Farhana, Linear regression model in estimating solar radiation in Perlis. *Energy Procedia*, 18, 1402-1412, 2012. [doi: 10.1016/j.egypro.2012.05.156](https://doi.org/10.1016/j.egypro.2012.05.156).
- [11] Ö. Kızıltan and M. Şahin, Calculation of solar radiation by using regression methods. *International Physics Conference at the Anatolian Peak (IPCAP 2016)*, sayfa 1-9, Erzurum, Türkiye, 2016.
- [12] O. Kärner, ARIMA representation for daily solar irradiance and surface air temperature time series. *Atmospheric and Solar-Terrestrial Physics*, 71, 841-847, 2009. [doi: 10.1016/j.jastp.2009.03.018](https://doi.org/10.1016/j.jastp.2009.03.018).
- [13] J. Wu and C. K. Chan, Prediction of hourly solar radiation using a novel hybrid model of ARMA and TDNN. *Solar Energy*, 85, 808-817, 2011. [doi: 10.1016/j.solener.2011.01.013](https://doi.org/10.1016/j.solener.2011.01.013).
- [14] M. B. Shams, S. Haji, A. Salman, H. Abdali and A. J. Alsaffar, Time series analysis of Bahrain's first hybrid renewable energy system. *Energy*, 103, 1-15, 2016. [doi: 10.1016/j.energy.2016.02.136](https://doi.org/10.1016/j.energy.2016.02.136).
- [15] C. Paoli, C. Voyant, M. Muselli and M. L. Nivet, Forecasting of preprocessed daily solar radiation time series using neural networks. *Solar Energy*, 84, 2146-2160, 2013. <https://doi.org/10.1016/j.solener.2010.08.011>.
- [16] F. N. Melzi, T. Touati, A. Same and L. Oukhellou, Hourly solar irradiance forecasting based on machine learning models. *15th IEEE International Conference on Machine Learning and Applications (ICMLA 2016)*, sayfa 441-446, California, USA, 2016. [doi: 10.1109/ICMLA.2016.0078](https://doi.org/10.1109/ICMLA.2016.0078).
- [17] A. Shadab, S. Ahmad and S. Said, Spatial forecasting of solar radiation using ARIMA model. *Remote Sensing Applications: Society and Environment*, 20, 2020. <https://doi.org/10.1016/j.rsase.2020.100427>.
- [18] A. E. Hoerl and R. W. Kennard, Ridge regression biased estimation for nonorthogonal problems. *Technometrics*, 12, 55-67, 2000. <https://doi.org/10.2307/1271436>.
- [19] D. W. Marquardt and R. D. Snee, Ridge regression in practice. 29, 3-20, 1975. <https://doi.org/10.2307/2683673>
- [20] Solar Constant. <https://www.britannica.com/science/solar-constant> , Erişim 7 Şubat 2023.

