



Model-Ağacı (M5-tree) yaklaşımı ile HELIOSAT tabanlı güneş radyasyonu tahmini

Taha Demirgöl¹, Vahdettin Demir^{*1}, Mehmet Faik Sevimli¹

¹KTO Karatay Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, İnşaat Mühendisliği Bölümü, Konya, Türkiye

Anahtar Kelimeler

Solar Radyasyon
HELIOSAT
M5-Tree
Tahmin
Türkiye

Araştırma Makalesi

Geliş: 29.06.2022
Reviz: 03.08.2022
Kabul: 12.08.2022
Yayınlanma: 15.02.2023



Öz

Solar radyasyon (SR), enerji dönüşümü, meteoroloji, tarım ve hayvancılık ile ilgili çalışmalar için önemli bir parametredir. Solar radyasyonun belirlenebilmesi için gereken alıcıların tüm noktalar için temin edilememesinden dolayı bu parametre çeşitli yöntemlerle tahmin edilebilir. Bu çalışmada, Türkiye'nin aylık ortalama SR değerleri (kWsa/m^2) M5 model ağacı (M5-tree), sezgisel regresyon tekniği kullanılarak MATLAB platformunda tahmin edilmiştir. SR modellemesinde komşu ölçüm istasyonlarına ait konum bilgileri ve periyodiklik bileşeni olan ay değerleri kullanılarak tahminler gerçekleştirilmiştir. Modellerde 2004-2018 yıllarını kapsayan uzun dönem aylık ortalama SR verileri kullanılmıştır. Bu veriler Meteoroloji Genel Müdürlüğü'nden temin edilen ve uydu tabanlı hibrit bir model olan HELIOSAT model verileridir. Çalışmada Türkiye'nin 81 iline ait ölçüm ortalamaları kullanılmıştır. 81 noktadan alınan 12 aylık verilerin %75'i eğitim sürecinde kullanılırken %25'i test sürecinde kullanılmıştır. Rastgele seçilen test istasyonlarından elde edilen solar radyasyon tahminleri gözlenen verilerle karşılaştırılmıştır. Bu karşılaştırmalarda, Karekök Ortalama Karesel Hata (KOKH), Ortalama Mutlak Hata (OMH), Ortalama Mutlak Bağlı Hata (OMBH) ve Determinasyon Katsayısı (R^2) kullanılmıştır. M5-tree kullanılarak elde edilen modellerde en başarılı sonuçlar; KOKH= 0.3604, OMH= 0.1451, OMBH= 3.6029 ve R^2 = 0.9879 olarak elde edilmiştir. Böylece M5-tree yönteminin literatürde yer alan yöntemlere alternatif bir yöntem olabileceği ortaya konmuştur.

Forecasting of HELIOSAT-Based solar radiation by Model-Tree (M5-tree) approach

Keywords

Solar Radiation
HELIOSAT
M5-Tree
Forecast
Turkey

Research Article

Received: 29.06.2022
Revised: 03.08.2022
Accepted: 12.08.2022
Published: 15.02.2023

Abstract

Solar radiation (SR) is an important parameter for studies on energy conversion, meteorology, agriculture and animal husbandry. Since the receivers required for the determination of solar radiation are not available for all points, this parameter can be estimated by various methods. In this study, Turkey's monthly average SR values (kWh/m^2) were estimated on the MATLAB platform using the M5 model tree (M5-tree), a heuristic regression technique. SR was estimated using the location information of neighboring measurement stations and the month values, which are the periodicity component. Long-term monthly average SR data covering the years 2004-2018 were used in the models. These data are HELIOSAT model data, which is a satellite-based hybrid model obtained from the General Directorate of Meteorology. In the study, the measurement averages of Turkey's 81 provinces were used. While 75% of the 12-month data obtained from 81 points were used in the training process, 25% was used in the testing process. Solar radiation estimates from randomly selected test stations were compared with the observed data. In these comparisons, Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Relative Error (MARE) and Coefficient of Determination (R^2) were used. The most successful results in the models obtained using the M5-tree; RMSE= 0.3604, MAE= 0.1451, MARE= 3.6029, and R^2 = 0.9879. Thus, it has been revealed that the M5-tree method can be an alternative method to the methods in the literature.

*Sorumlu Yazar

(demirgultaha@gmail.com) ORCID ID 0000-0003-0164-058X
*(vahdettin.demir@karatay.edu.tr) ORCID ID 0000-0002-6590-5658
(mehmet.faik.sevimli@karatay.edu.tr) ORCID ID 0000-0002-4676-8782

Kaynak Göster (APA)

Demirgöl, T., Demir, V., Sevimli, M. F. (2023). Model-Ağacı (M5-tree) yaklaşımı ile HELIOSAT tabanlı güneş radyasyonu tahmini. *Geomatik*, 8(2), 124-135

1. Giriş

Güneş yerkürenin en önemli enerji kaynaklarından biri olmakla beraber aynı zamanda temiz ve ucuz bir enerji kaynağıdır (Kılıç ve Kumaş, 2016). Genellikle solar radyasyon (SR) ölçüm cihazları yardımıyla ölçülür. Ancak, bu cihazların kurulum-bakım maliyetleri ve kalibrasyon gereksinimleri yüksektir. Bu nedenle, dünya çapında çoğu bölgede ölçümü yapılamamaktadır.

Güneş enerjisi, Dünya’da mevcut olan çoğu enerjinin kaynağıdır ve yakın gelecekte, özellikle gelişmekte olan ülkelerde çok önemli bir rol oynaması beklenmektedir. Bu enerji kaynağının kullanımına olan ilginin ortaya çıkması, özellikle konvansiyonel kaynakların artan maliyetleri nedeniyle son yıllarda artmıştır (Guermoui ve ark., 2020). Dünya yüzeyinde mevcut olan solar radyasyon, yerel iklim koşullarına bağlıdır. Yerel solar radyasyon bilgisi, bina enerji sistemlerinin, güneş enerjisi sistemlerinin doğru tasarımı ve binalardaki termal ortamın verimli şekilde değerlendirilmesi için esastır. En iyi veri tabanı, önerilen güneş sisteminin sahasında uzun vadeli ölçülen veriler olacaktır. Bununla birlikte, radyasyon ölçüm ağlarının sınırlı kapsamı, güneş radyasyonu modellerinin geliştirilmesi ihtiyacını ortaya çıkarmıştır (Ahmad ve Tiwari, 2011).

Güneşten dünya yüzeyine aktarılan elektromanyetik enerji olarak tanımlanabilen SR, günümüzde güneş ışınlarıyla doğrudan ilişkisi bulunan, başta elektrik üretiminde kullanılan sistemlerde olmak üzere tarımsal ürünlerin yetiştirilmesi, meteorolojik araştırmalar, doğal ışıklandırma sistemleri, seralar ve hayvan yetiştiriciliğinde ısıtma ve soğutma katkısının belirlenmesi gibi pek çok alanda kullanılmaktadır (Jiang, 2009; Şeker, 2021). Günümüzde Türkiye gibi gelişmekte olan ülkeler ve gelişmiş ülkelerde elektrik enerjisine olan ihtiyaç katlanarak artmaktadır (Taktak ve İli, 2018). SR, tarımda bitkilerin büyüme süreci, büyümesi ve biyokütle depolanması için referans evapotranspirasyonun hesaplanmasında önemli bir yere sahip olup (Çıtakoğlu, 2015), günümüzde iklim değişikliği ile ilgili senaryo gruplarının incelenmesinde SR değerleri kullanılmaktadır (Bora ve ark., 2015). Ayrıca güneş enerjisi çalışmalarının tasarımı ve incelenmesi aşamasında solar radyasyon hakkında bilgi sahibi olmak gereklidir (Çıtakoğlu ve Demir, 2021).

SR'nin tahmin edilebilir hale gelmesi, güneş enerjisi sistemlerinin tasarlanma ve geliştirilme aşamaları için oldukça önemlidir. Yapılarda ısınma, atmosferik enerji dengesi çalışmaları, binalarda termal analizler ve meteorolojik tahminler gibi çeşitli alanlarda SR verilerine ihtiyaç duyulmaktadır (Özgören ve ark., 2012). Ayrıca, seçilen bir coğrafi bölgedeki solar radyasyon miktarının isabetli tahmini, birçok mühendislik, akademik ve çevresel uygulama için çok değerlidir. Türkiye'nin yıllık ortalama solar radyasyonuna baktığımızda 3,6 kWsa/m² gün değeri gözlenmiş olup toplam yıllık radyasyon periyodu yaklaşık olarak 2610 saat olarak hesaplanmıştır (Sözen ve Arcaklıoğlu, 2005). Yerleşim bölgelerinde ölçülen rüzgâr hızı, nispi nem, buhar basıncı, hava basıncı, hava sıcaklığı, güneşlenme süresi gibi meteorolojik verilerin güneş radyasyonunun doğru hesaplanmasında önemli bir etken olduğu araştırmalar sonucunda ortaya çıkmıştır (Kılıç ve Kumaş,

2016). Ancak solar radyasyonun ölçümü hem maliyet hem de kullanılan gereçler, kalibrasyon ve bakım aşamaları bakımından güç olduğu için gözlemlenen ölçüm verilerinden faydalanarak tahmin metodlarının geliştirilmesine ve test edilmesine yönelim artmaktadır (Gül ve Çelik, 2017).

Solar radyasyon, günümüzde pek çok alanda yapılan çalışmalar için ihtiyaç duyulan önemli bir parametre haline gelmiştir. Ama elimizde gözlenmiş verinin kısıtlı olmasından dolayı bu gözlenmiş verileri referans olarak solar radyasyon ölçümü olmayan istasyonların verilerinin oluşturulmasına ihtiyaç duyulmaktadır. Son yıllarda SR tahmini için gözlem verileri baz alınarak uygulanan yöntemler artmaktadır (Alizamir ve ark., 2020).

SR verileri nem, sıcaklık, rüzgâr hızı, bulut örtüsü gibi kolay ölçülebilen iklim parametreleri ile tahmin edilebilmektedir. Bu bakış açısıyla, SR verilerini tahmin etmek için çok sayıda model önerilmiştir. Bunlardan bazıları matematiksel formüllere dayanmakta ve ampirik modeller olarak adlandırılmaktadırlar. Ampirik modellerin hesaplanması kolaydır ve güneş radyasyonu verilerini tahmin etmek için yararlı bir teknik olarak kabul edilir. Aylık ortalama günlük küresel güneş radyasyonunun tahmininde ampirik modeller sıklıkla kullanılsa da bu modeller bulut örtüsü, yağışlı günler gibi hava koşullarındaki hızlı değişiklikler nedeniyle kısa vadeli güneş radyasyonu verilerini doğru bir şekilde tahmin edemezler. Buna paralel olarak, bazı araştırmacılar, bu modellerin, yağışlı günlerde yoğun bulutlardan güneş ışınımının güçlü bir şekilde etkilendiği nemli bölgelerde hem bağımlı hem de bağımsız değişkenler arasındaki karmaşık ve doğrusal olmayan ilişkileri yansıtamadıklarını bildirmiştir (Ağbulut ve ark., 2021). Bu nedenle diğer parametre ölçümlerine bağımlı olmadan yapılan ölçümler, örneğin konum bilgilerine göre hem daha pratik hem de daha ekonomiktir.

Günümüzde matematiksel formüllere alternatif olan makine öğrenimi modelleri, yüksek doğrulukları nedeniyle solar radyasyon tahminleri için umut verici yöntemler olup, bunlardan en yaygın olarak uygulananı yapay sinir ağlarıdır (YSA). Örneğin çok katmanlı algılayıcı sinir ağları, evrimsel yapay sinir ağları, genelleştirilmiş regresyon sinir ağları ve geri yayılım sinir ağları bunlardan bazılarıdır (Feng ve ark., 2019).

SR tahmini ile ilgili literatürde yer alan bazı önemli çalışmalar incelendiğinde, Sözen ve Arcaklıoğlu (2005) Türkiye’de yer alan 12 adet istasyona ait ortalama günlük SR tahmini gerçekleştirmişlerdir. SR potansiyelini belirlemek için Yapay Sinir Ağları metoduyla 3 farklı algoritma (Scaled konjugat gradyan, Pola-Ribiere eşlenik gradyan ve Levenberg-Marquardt öğrenme algoritmaları) ve lojistik sigmoid transfer fonksiyonu kullanmışlardır. Algoritmalarda 6 farklı giriş (enlem, boylam, yükseklik, ay, ortalama güneşlenme süresi ve ortalama sıcaklık) parametresi ile SR tahmini gerçekleştirmişlerdir (Sözen ve Arcaklıoğlu, 2005).

Özgören ve ark. (2012) Türkiye'nin herhangi bir yerindeki aylık ortalama günlük toplam küresel güneş ışınımını tahmin etmek için çoklu doğrusal olmayan regresyon (MNL) yöntemine dayalı bir yapay sinir ağı (YSA) modeli geliştirmişlerdir. Türkiye'nin 31 istasyonunun verilerinden hareketle 1 giriş 1 çıkışlı

kombinasyondan 10 giriş 1 çıkışlı kombinasyona kadar analizler yaparak SR tahmini yapmışlardır (Özgören ve ark., 2012).

Bhardwaj ve ark. (2013) yaptıkları çalışmada solar radyasyon tahmini için Gizli Markov Modeli ve genelleştirilmiş Bulanık Modelin bir kombinasyonunu kullanarak şekil tabanlı kümeleme modelini kullanmışlardır. Model için kullanılacak parametreleri hava durumu izleme istasyonundan kaydedilen ortam sıcaklığı, bağıl nem, atmosfer basıncı, rüzgâr hızı, rüzgâr yönü, güneş spektrumu ve güneş radyasyonu olarak belirlemişlerdir. Çalışmalarında solar radyasyonun tahmini için model benzerliğine dayalı kümeleme algoritmasını geliştirerek hibrit bir teknik kullanmışlardır. Sonuç olarak önerilen modelin yaygın olarak kullanılan modellere kıyasla daha iyi tahmin sağladığını belirtmişlerdir (Bhardwaj ve ark., 2013).

Wang ve ark. (2017) Çin'deki 21 farklı bölgedeki günlük küresel SR tahminini ızgara bölmeli Adaptive-Network Based Fuzzy Inference Systems (ANFIS)-SC, ANFIS-GP ve M5 Model ağacı olmak üzere 3 farklı makine öğrenme tekniği kullanarak gerçekleştirmişlerdir. Bu modeller kullanılarak kalibre edilmiş bir ampirik Ångström modeli ile karşılaştırılmıştır. Çalışmalarında 7 giriş (güneşlenme süresi, hava basıncı, maksimum sıcaklık, minimum sıcaklık, ortalama sıcaklık, su buharı basıncı ve bağıl nem) parametresi ile SR tahmini gerçekleştirmişlerdir (Wang ve ark., 2017).

Kuncan ve Şahin (2017) uydu verilerini kullanarak Türkiye'deki aylık ortalama günlük güneş radyasyon tahmini gerçekleştirmişlerdir. Çalışmalarında yöntem olarak YSA'yı (farklı transfer fonksiyonları, öğrenme algoritmaları ve 40'ı bulan ara tabaka hücre sayıları ile) kullanmışlar ve 53 istasyon için 20 farklı model oluşturmuşlardır. Bu modellerde 5 farklı giriş (solar radyasyon ay değerleri, yeryüzü sıcaklığı, yükseklik, enlem ve boylam) parametresi ile SR tahmin edilmeye çalışılmıştır (Kuncan ve Şahin, 2017).

Aksoy ve Bolat (2017) Akdeniz Bölgesi'nde yedi farklı lokasyon için aylık NOAA/AVHRR uydu verilerini kullanarak YSA metodu ile aylık ortalama günlük SR tahmini gerçekleştirmişlerdir. Çalışmalarında Yapay Sinir Ağları yöntemini kullanarak solar radyasyon değerleri tahmin etmişlerdir. Modellemede 5 farklı giriş (ay, yükseklik, yeryüzü sıcaklığı, enlem, boylam) ile SR tahmin edilmeye çalışılmıştır (Aksoy ve Bolat, 2017).

Gül ve Çelik (2017) Tunceli ilinin aylık ortalama günlük global güneş radyasyonunu tahmin etmek için adaptif ağ tabanlı bulanık çıkarım sistemi (ANFIS) yöntemini kullanmışlardır. Bu çalışmada uygulanan ANFIS modelinde, 3 giriş parametresi (yıllar, aylar ve aylık ortalama günlük toplam güneşlenme süresi [saat]) kullanarak aylık ortalama günlük toplam global güneşlenme şiddeti (AOGGR) [$kWsa/m^2$] tahmini gerçekleştirmişlerdir (Gül ve Çelik, 2017).

Kaba ve ark. (2018) Türkiye'de yer alan 34 adet istasyon verisi ile ortalama günlük SR tahmini gerçekleştirmişlerdir. Derin Öğrenme yöntemini kullandıkları çalışmalarında günlük küresel SR'yi tahmin etmek için en iyi ilişkili ve en yaygın olarak kullanılan bir astronomik parametre (dünya dışı radyasyon) (H_0), ve dört meteorolojik parametre (güneşlenme süresi (SD), bulut örtüsü (CC), maksimum sıcaklık (T_{max}) ve minimum

sıcaklık (T_{min})) kullanarak 16 farklı kombinasyon ile modellerini eğitmişlerdir (Kaba ve ark., 2018).

Keshtegar ve ark. (2018) Türkiye'nin Doğu Akdeniz Bölgesi'nde bulunan Adana ve Antakya istasyonlarında SR tahmini gerçekleştirmişlerdir. Çalışmalarında Kriging, yanıt yüzeyi yöntemi (RSM), çok değişkenli adaptif regresyon (MARS) ve M5 model ağacı (M5-tree) gibi dört farklı sezgisel regresyon yöntemi kullanmışlardır. Modellemelerinde 5 giriş parametresi (maksimum sıcaklık (T_{max}), minimum sıcaklık (T_{min}), güneş süresi (H_s), rüzgâr hızı (WS) ve bağıl nem (RH)) ile SR tahminini gerçekleştirmişlerdir (Keshtegar ve ark., 2018).

Ruşen (2018), Konya ve Karaman bölgesinde yatay bir yüzey üzerinde günlük SR bileşenlerini araştırdığı çalışmada uydu görüntüleri ile HELIOSAT yöntemini kullanmıştır. Çalışma sonucunda HELIOSAT yönteminin iyi tahminlere sahip olduğunu belirtmiştir (Rusen, 2018).

Cornejo-Bueno ve ark. (2019) İspanya'nın Toledo şehri için Meteosat uydusundan elde ettikleri bir yıllık saatlik solar radyasyon verileriyle Destek Vektör Regresyonu (SVR), Çok Katmanlı Algılayıcı (MLP), Aşırı Öğrenme Makinesi (ELM) ve Gauss Süreçleri (GPR) olmak üzere birkaç son teknoloji makine öğrenmesi yöntemiyle solar radyasyon tahmin modellerini çalışmışlardır. Bu çalışmada seçilen farklı makine öğrenmesi regresyon tekniklerinin performansını değerlendirmek için 4 senaryoyu içeren ve her senaryo için çıkış katmanında solar radyasyon değeri tahmin eden denemeler yapmışlardır. Sonuç olarak karşılaştırılan regresörler arasında en iyi performansı Aşırı Öğrenme Makineleri (ELM)'nin sergilediğini belirtmişlerdir (Cornejo-Bueno ve ark., 2019).

Alizamir ve ark. (2020) 6 farklı makine öğrenme tekniği (Gradyan Artırma Ağacı, Çok Katmanlı Algılayıcı Sinir Ağı, Sınıflandırma ve Regresyon Ağacı, Çok Değişkenli Uyarlamalı Regresyon Eğrisi, ANFIS-FCM, ANFIS-SC) kullanarak Türkiye ve ABD olmak üzere 2 farklı lokasyondaki iki istasyondan SR tahmini yapmışlardır. Yöntemlerde 4 giriş parametresi (rüzgâr hızı, maksimum sıcaklık, minimum sıcaklık, bağıl nem) kullanarak SR tahmini gerçekleştirmişlerdir (Alizamir ve ark., 2020).

Çıtakoğlu ve Demir (2021) Türkiye'nin aylık ortalama solar radyasyon değerlerini üç farklı derin öğrenme algoritması kullanarak gerçekleştirmişlerdir. Çalışmada tahminler 3 farklı derin öğrenme eğitim-optimizasyon algoritması olan ADAM (Adaptive Moment Estimation), SGDM (Stochastic Gradient Descent with Momentum) ve RMSProp (Root Mean Square propagation) optimizasyon algoritmaları kullanmışlardır. Algoritmalarda 4 giriş (Enlem, Boylam, Yükseklik, Periyodiklik) parametresi kullanarak SR tahmini gerçekleştirmişlerdir (Çıtakoğlu ve Demir, 2021). Çalışmalarında konum bilgilerine periyodiklik bileşeninin eklemesi durumunda daha başarılı sonuçlar elde edildiğini belirtmişlerdir.

Ağbulut ve ark. (2021) Türkiye'de farklı güneş radyasyon dağılımlarına sahip 4 istasyonun (Kırklareli, Tokat, Nevşehir ve Karaman) günlük küresel güneş radyasyonu verilerini tahmin etmek için 4 farklı makine öğrenme algoritması (destek vektör makinesi (SVM), yapay sinir ağı (ANN), çekirdek ve en yakın komşu (k-NN) ve derin öğrenme (DL))

kullanmışlardır. Çalışmada 5 giriş parametresi (maksimum sıcaklık, minimum sıcaklık, bulut örtüsü, atmosfer dışı SR, gün uzunluğu) kullanarak SR tahmini gerçekleştirmişlerdir (Ağbulut ve ark., 2021).

Literatür taramasında regresyon ağaçları, destek vektör makineleri, yapay sinir ağları, derin öğrenme, bulanık mantık gibi yapay zekâ kullanılarak tahmin etme yöntemlerinin oldukça popüler olduğu ve yüksek doğruluklu sonuçlar verdiği görülmektedir. Bu yöntemlerin en önemli özellikleri başlangıçta giriş parametresi olarak verilen somut veriler (bağımsız değişken) ile çıkış parametresi olarak elde edilmek istenilen veri (bağımlı değişken) belirtildiğinde kendini eğiten sistemlerin kullanılmasıdır. Çalışmalarda genellikle KOKH, OMH, R, R² gibi karşılaştırma kriterlerinin kullanıldığı görülmektedir. Ayrıca SR tahmini için genellikle konumsal, astronomik ve meteorolojik verilerin çeşitli kombinasyonları girdi parametresi olarak kullanılmaktadır.

Bu çalışmada, Türkiye'nin 81 iline ait ölçüm istasyonlarının 2004-2018 yıllarına ait aylık ortalama solar radyasyon değerleri (kWs/m²) kullanılarak, SR tahminleri gerçekleştirilmiştir. Çalışmada Model Ağacı (M5-tree) sezgisel regresyon yöntemi kullanılmış olup, 81 istasyondan alınan 12 aylık verilerin (972 adet) %75'i (729 adet) eğitim sürecinde, %25'i (273 adet) ise test sürecinde kullanılmıştır. Rastgele seçilen test istasyonlarından elde edilen solar radyasyon tahminleri gözlenen verilerle karşılaştırılmıştır. Bu karşılaştırmalarda, Karekök Ortalama Karesel Hata (KOKH), Ortalama Mutlak Hata (OMH), Ortalama Mutlak Bağlı Hata (OMBH) ve Determinasyon Katsayısı (R²) kullanılmıştır. Tahminler komşu ölçüm istasyonlarına ait konum bilgileri ve periyodiklik bileşeni olan ay değerleri kullanılarak gerçekleştirilmiştir. SR'nin belirlenmesinde M5-tree yönteminin literatüre alternatif bir yöntem olma durumu araştırılmıştır.

Çalışmanın literatürden ayrılan yönleri; 81 adet ölçüm istasyonuna ait ve 2004-2018 yıllarını kapsayan verilerin kullanılması (1), Türkiye'de HELIOSAT tabanlı SR tahminin gerçekleştirilmesi (2), Türkiye çapında ilk kez M5-tree yönteminin SR tahmininde kullanılıyor olmasıdır (3).

Çalışmanın gelecek bölümünde HELIOSAT modeli hakkında bilgiler verilmiş, çalışma alanı tanıtılmış ve çalışmada kullanılan veriler ve yöntemler açıklanmıştır. Üçüncü bölümde sonuçları değerlendirirken kullanılan karşılaştırma kriterleri açıklanmış, bulgular gidış-saçılma grafiklerine ek olarak çalışma alanı için oluşturulmuş harita ve test istasyonları için oluşturulmuş uzun dönem aylık ortalama grafikleri üzerinden yorumlanmıştır. Tüm bulgular bu bölümde literatürle karşılaştırılmıştır. Son bölümde ise sonuçlar özetlenmiş çalışma kapsamı ile gelecek çalışmalar açıklanmıştır.

2. Yöntem

Türkiye, global güneş radyasyon dağılımı hesaplama yöntemi olarak HELIOSAT modelini tercih etmektedir. HELIOSAT hibrit bir modelleme yöntemidir. Çalışma prensibi bir radyasyon transfer denkleminin analizi ve basit istatistiksel ilişkilerdir. İlk olarak açık hava için

direkt ve difüz bileşenlerin ayrı ayrı hesaplandığı model ile global güneş radyasyon değişkeni hesaplanmaktadır. Sonraki aşamada uydu verileri ikinci nesil METEOSAT (Meteosat Second Generation) (MSG'den elde edilmektedir. Temin edilen uydu verileri kullanılarak ilk aşamada elde edilen global güneş radyasyon değerleri bulut değişkenine (parametre) göre düzenlenmektedir. 15 dakikada bir yenilenen uydu verilerinin işlenmesiyle elde edilen değerlerden bulut indeksine geçiş yapılarak bu değerler açık hava güneş radyasyon değerleri ile işlenmekte ve bu işlem sonucunda mevcut hava koşullarına göre güneş radyasyon değerleri elde edilmektedir. Saatlik verilerden günlük, aylık, mevsimlik, yıllık veriler elde eden Meteoroloji Genel Müdürlüğü (MGM) bu verilerle 2004 yılından başlayan bir veri arşivi oluşturmuştur. 0,2⁰ x 0,2⁰ çözünürlüğe sahip olan HELIOSAT modeli ile Türkiye için toplam 3610 grid noktasında veri üretilmektedir. MGM tarafından yapılan verifikasyon çalışmalarında modelin yaklaşık %2 hata ile radyasyon verisi tahmininde bulunduğu belirtilmektedir (URL-1, 2022).

Çalışmada kullanılan verilere ait tanımlayıcı istatistiksel bilgiler Tablo 1' de yer almaktadır.

Tablo 1. Verilere ait tanımlayıcı bilgiler

Kriter	Eğitim	Test
Ortalama	4,424	4,418
Standart Hata	0,069	0,117
Ortanca	4,565	4,584
Standart Sapma	1,840	1,853
Basıklık	-1,346	-1,354
Çarpıklık	0,059	0,055
En Büyük	7,777	7,685
En Küçük	1,282	1,205
Veri sayısı	720	252

2.1. Çalışma alanı

Türkiye, 26⁰-45⁰ doğu boylamları ve 36⁰-42⁰ kuzey enlemleri arasında yer almaktadır (İşcan ve İlğaz, 2017). Tipik Akdeniz iklimine sahip olmasından dolayı her mevsim güneş alan Türkiye'nin güneş enerjisi potansiyeli oldukça yüksektir (Sözen ve Arcaklioğlu, 2005). Çalışmada yer alan 81 ile ait istasyon konumları, Şekil 1'de verilmiştir. Bu illere ait ortalama aylık güneş radyasyonu verileri Meteoroloji Genel Müdürlüğü'nün resmî sitesinden elde edilmiştir (URL-2, 2022). Şekil 1 incelendiğinde rastgele seçilen eğitim istasyonları olarak Türkiye'nin 60 iline ait istasyonlar yeşil olarak gösterilmiştir. Kırmızı ile işaretlenen 21 il istasyonu ise test istasyonlarını simgelemektedir. Ayrıca eğitim ve test istasyonlarına ait konum bilgileri ile ortalama SR değerleri (kWh/m²) Tablo 2 ve Tablo 3'te sırasıyla yer almaktadır.

2.2. M5 Model Ağacı (M5-Tree)

1992 yılında Quinlan, M5 model ağacı algoritmasını alternatif bir regresyon yöntemi olarak geliştirmiştir (Quinlan, 1992). Demirci (2019), karar ağacını, bir bağımlı değişkenin değerinin tahminini bağımsız değişken kümesinin değerlerini kullanarak nasıl tahmin edilebileceğini gösteren bir ikili (iki yönlü bölünmüş) ağaç olarak gösterilen mantıksal bir model olarak ifade etmiştir ve sınıflama ağacı ve regresyon ağacı olmak

üzere iki tür karar ağacı olduğunu belirtmiştir. Bu türlerden ilki ve en uygun olanı sınıflama ağaçlarıdır ve bir sayısal niteliğin değerini tahmin etmek için kullanılan bir sembolik sınıftır. İkinci tür ise regresyon ağaçlarıdır ve öngörü (tahmin) için kullanılmaktadır.

Tablo 2. Eğitim istasyonları

İstasyon	Enlem (°K)	Boylam (°D)	Yükseklik (m)	SR (kWsa/m ²)
Adana	37	35.32	23	4.77
Adıyaman	37.76	38.27	669	4.92
Afyon	38.74	30.55	1013	4.51
Ağrı	39.72	43.05	1640	4.43
Amasya	40.65	35.83	392	4.17
Ankara	39.97	32.86	870	4.34
Antalya	36.9	30.79	43	5.1
Artvin	41.18	41.82	597	3.99
Aydın	37.84	27.84	597	4.72
Balıkesir	39.63	27.92	101	4.05
Bilecik	40.14	29.97	526	4.16
Bingöl	38.88	40.49	1177	4.61
Bitlis	38.39	42.12	1545	4.7
Bolu	40.74	31.6	732	4.11
Burdur	37.68	30.33	1025	4.91
Bursa	40.23	29.01	100	4.1
Çanakkale	40.15	26.41	3	3.95
Çankırı	40.61	33.61	730	4.19
Çorum	40.54	34.94	798	4.26
Denizli	37.76	29.09	450	4.7
Diyarbakır	37.92	40.22	677	4.78
Edirne	41.68	26.56	48	3.87
Elazığ	38.67	39.22	1015	4.85
Erzincan	39.74	39.5	1214	4.54
Erzurum	39.95	41.17	1893	4.36
Eskişehir	39.78	30.58	732	4.32
Gaziantep	37.07	37.39	840	4.9
Giresun	40.92	38.39	84	3.89
Gümüşhane	40.46	39.47	1210	4.18
Hakkari	37.57	43.75	1720	4.62
Hatay	36.36	36.28	85	4.87
Isparta	37.76	30.55	1043	4.76
İstanbul	40.98	28.82	30	4.1
İzmir	38.41	27.14	25	4.47
K.maraş	37.57	36.91	568	4.85
Kars	40.59	43.08	1750	4.27
Kastamonu	41.37	33.77	800	4.02
Kayseri	38.72	35.49	1071	4.65
Kırklareli	41.73	27.21	203	4.37
Kırşehir	39.16	34.15	985	4.45
Kocaeli	40.76	29.91	76	3.98
Konya	37.99	32.56	1026	4.7
Kütahya	39.42	29.99	969	4.31
Malatya	38.35	38.31	977	4.83
Manisa	38.61	27.4	42	4.38
Mardin	37.31	40.73	1150	4.91
Mersin	36.8	34.62	6	4.88
Muğla	37.29	28.37	646	5.09
Muş	38.74	41.49	1300	4.54
Nevşehir	38.63	34.71	1250	4.59
Niğde	37.97	34.69	1208	4.75
Ordu	40.98	37.88	10	3.76
Rize	41.04	40.5	4	3.6
Sakarya	40.77	30.39	31	4.01
Samsun	41.35	36.24	44	3.96
Siirt	37.93	41.94	895	4.83
Sinop	42.02	35.15	32	4.01
Sivas	39.74	37.02	1285	4.46
Tekirdağ	40.99	27.49	3	3.92
Tokat	40.3	36.56	623	4.13

Tablo 3. Test istasyonları

İstasyon	Enlem (°K)	Boylam (°D)	Yükseklik (m)	SR (kWsa/m ²)
Aksaray	38.37	34.03	900	4.72
Ardahan	41.11	42.7	2200	3.99
Bartın	41.62	32.35	25	4.12
Batman	37.89	41.12	550	4.79
Bayburt	40.25	40.43	1550	4.33
Düzce	40.84	31.15	160	3.96
Iğdır	39.92	44.06	858	4.5
Karabük	41.2	32.62	278	4.07
Karaman	37.17	33.22	1250	4.91
Kırıkkale	39.85	33.31	700	4.11
Kilis	36.72	37.12	640	4.89
Osmaniye	37.1	36.25	120	4.81
Şanlıurfa	37.16	38.79	547	4.94
Şırnak	37.52	42.45	1350	4.8
Trabzon	41	39.78	37	3.58
Tunceli	39.1	39.55	914	4.68
Uşak	38.67	29.4	921	4.47
Van	38.49	43.39	1661	4.62
Yalova	40.66	29.21	2	4.11
Yozgat	39.82	34.81	1418	4.39
Zonguldak	41.45	31.78	136	3.99

Regresyon analizinden kısaca bahsetmek gerekirse iki ya da daha fazla sayısal değişken arasındaki ilişki seviyesini belirlemek için kullanılan çözümlenme yöntemidir. Tek değişkenli regresyon, bir değişkenin kullanıldığı çözümlenmedir. Çok değişkenli regresyon ise iki ya da daha fazla değişkenin kullanıldığı çözümlenmelerdir. Regresyon modelleri akademik çalışmalarda elde edilen veriler kullanılarak bir öngöründe bulunmak için kullanılmaktadır. M5 modeli ise regresyon ağacı modellerinin birleşiminden oluşan bir karar ağacı yöntemidir. M5 model ağacı yönteminin ana unsuru iki bileşenli karar ağacı prensibidir. Karar ağacı metodu, bir bağımlı değişkenin değerini tahminini bağımsız değişken kümesindeki değerleri kullanarak nasıl bir tahminde bulunabileceğini gösteren iki yönlü ayrık ağaç olarak temsil edilen mantıksal bir modeldir (Demirci, 2019). Karar ağacı modelinde kategorik veriler kullanılabilirken, M5 model ağacı yönteminde ise kategorik verilerin yanı sıra nicel veriler de kullanılabilir (Quinlan, 1992). Quinlan (1992), M5 karar ağacını üç çeşit düğüm içeren ve model ağacının oluşturulması ve dalların budanması şeklinde iki ana aşamadan oluşan bir regresyon ağacı olarak ifade etmektedir. Bu düğümler kök, dal ve yaprak düğümleri olarak düşünülmektedir. İlk aşamada veriler kök düğümüne aktarılır burada işlenir ve daha sonra kök düğümünün bölünmesi sonucunda düğüm sağ ve sol olmak üzere ikiye bölünerek birbirine benzeyen iki düğüm oluşur (Ji ve ark., 2016). Bu aşamadan sonra düğümler sürekli bölünerek dallara ve yapraklara kadar uzanarak bir model ağacı oluşur. Dolayısıyla kök üst uzayı oluşturur ve bu üst uzay dal ve yapraklara doğru alt uzaylara bölünür. Bu modelde her düğüm bir tahmin değişkenini temsil etmektedir ve her düğüm sağ ve sol olmak üzere iki dala ayrılır. M5 karar ağacı yönteminde iki ana unsurdan ilk unsur bir kök düğümünden başlayarak her seferinde iki alt düğüme bölünerek bir ağaç modeli inşa etmek, ikinci unsur ise inşa edilen bu ağacın dallarını budama işlemidir. Ağaç modelinin inşa edildiği ilk aşamada girdi verileri alt uzaylara bölünerek

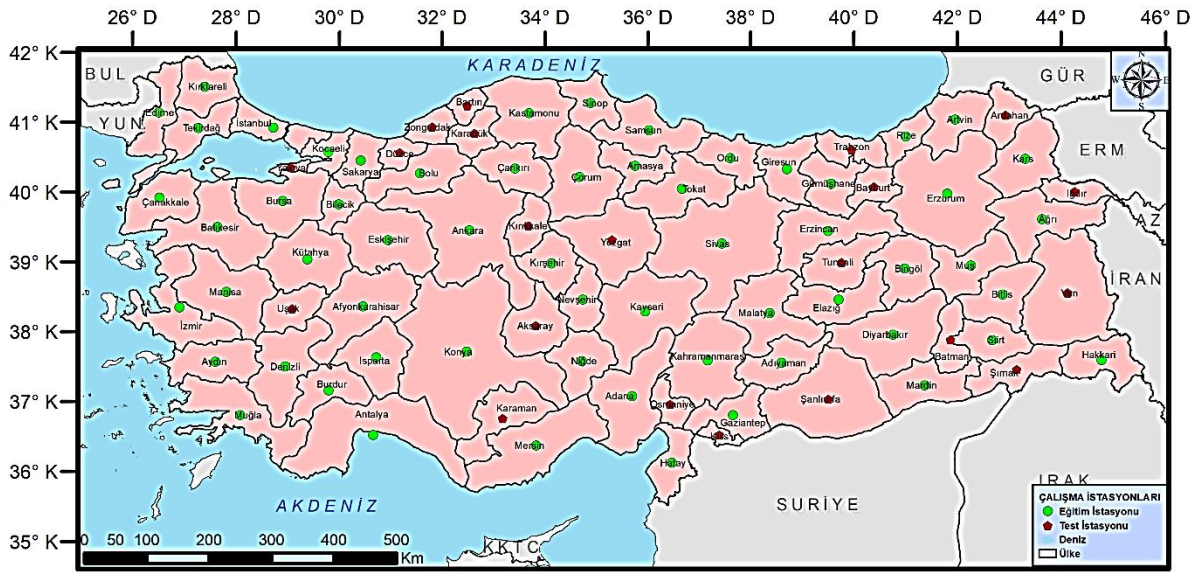
bir karar ağacı oluşturulur ve her alt uzay için bir doğrusal regresyon modeli tanıtılır (Granata ve ark., 2017; Kişi, 2015). Bölme kriteri, bir düğümde elde edilen sınıf değerlerinin standart sapması kullanılarak kriterlere ayrılması için kullanılarak her kriterin o düğümde test edilmesi sonucunda ortaya çıkan hataya bağlı beklenen azalmanın hesaplanmasına bağlıdır (Kisi, 2015).

Bölme kriterinin amacı, düğümdeki hata oranını açıklayarak düğümlerin standart sapmasını minimuma indirmektir. Eğer düğümdeki standart sapmayı minimuma indirmek mümkün değilse düğüm yaprak olarak biter (Bahadırılı, 2021). Yönteme ait standart sapma azalması (SDR) Eşitlik 1’de verilmiştir (Pal ve Deswal, 2009).

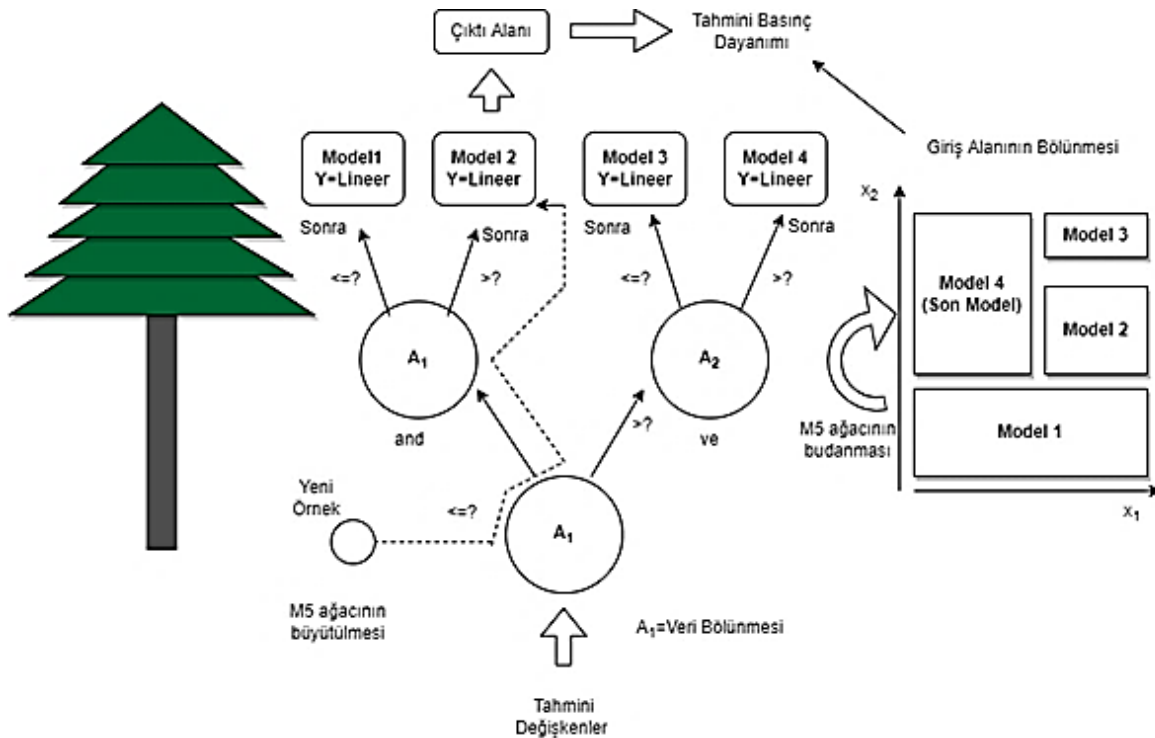
$$SDR = sd(T) - \sum \frac{|T_i|}{|T|} sd(T_i) \quad (1)$$

Formüldeki sd standart sapmayı göstermektedir. Düğüme tesir eden bir dizi bütünü ise T simgesi ifade etmektedir. Düğümdeki bölünmeden oluşan potansiyel verilerin i inci sonuçlarının bulunduğu alt uzaylar ise Ti olarak ifade edilmektedir.

Bölünme işlemi sırasında eğitim aşamasında eğitim için verilen verilere yüksek doğrulukta uyumalara neden olabilecek çok fazla dalları olan çok büyük bir ağaç oluşmaktadır. Bu nedenle dalları budamak gerekir. Quinlan’ın algoritması budama işlemine çok uygundur. Şekil 2’de Model ağacı yapısı yer almaktadır.



Şekil 1. Çalışma alanı Türkiye’nin 81 ili



Şekil 2. Model ağacı yapısı

Şekil 2’te gösterilen model ağacı diyagramında bir kök hüresinin her seferinde ikiye bölünerek ağacın dallarını düğüm noktalarını ve yapraklarını oluşturup daha sonra çok büyüyen bu ağacın budanması anlatılmaktadır. İlk adımda, algoritma ağacın olabildiğince büyümesine izin verir. Daha sonra, modelin daha yüksek doğruluğa ulaştıramayan dalları budanır. Son adımda ise, budama işleminden sonra, keskin hatlara sahip (ayırışma) ağaç modelini yumuşatmak için düzleştirme işlemi uygulanır (Bahadırli, 2021).

2.3. IDW Metodu

Ters Mesafe Ağırlıklandırma (IDW) Metodu, örnek noktalar tarafından örneklenemeyen noktalara veri oluşturmak için kullanılan bir enterpolasyon metodudur (Hastaoğlu ve ark., 2022; Yılmaz ve ark., 2023). Bu verilerin oluşturulması, çeşitli noktalarla komşuluklar değerlendirilerek uygulanan interpozisyonel mesafeye ve formüle bağlıdır (Taylan ve Damçayiri, 2016; Demirgöl ve ark., 2022). Yönteme ait genel eşitlikler Eşitlik 2 ve Eşitlik 3’te gösterilmektedir.

$$f(x, y) = \sum_{i=1}^n w_i f_i \quad (2)$$

$$w_i = \frac{h_i^{-p}}{\sum_{j=1}^n h_j^{-p}} \quad (3)$$

Burada; w_i ; ağırlıkları temsil eder ve değerlerinin toplamı Eşitlik 1’deki gibi olmalıdır
 f_i ; bilinen yükseklik değerini temsil eder,
 p ; bir güç parametresi olarak alınır ve üs ile gösterilir,
 h_i ; örnek noktalar ile enterpolasyon noktaları arasındaki uzamsal mesafeyi temsil eder (Krige, 1951; Shepard, 1968).

3. Bulgular ve Tartışma

Hidrolojik ve meteorolojik tahminler değerlendirilirken gözlenmiş değer ile modelde tahmin edilen değer arasındaki hata ölçütlerinin kullanımına sıklıkla rastlanılmaktadır (Yaseen, 2021). Literatürde yaygın olarak kullanılan karşılaştırma kriterlerinden Karekök Ortalama Karesel Hata (KOKH), Ortalama Mutlak Hata (OMH), Ortalama Mutlak Bağlı Hata (OMBH), determinasyon katsayısı (R^2) bu çalışmada karşılaştırma kriteri olarak tercih edilmiştir. Karşılaştırma kriterlerinin formülleri Eşitlik 4-7’de gösterilmektedir.

$$KOKH = \sqrt{\frac{1}{N} \sum_{i=1}^N (SR_p - SR_o)^2} \quad (4)$$

$$OMH = \frac{1}{N} \sum_{i=1}^N |SR_p - SR_o| \quad (5)$$

$$OMBH = \frac{1}{N} \left(\frac{\sum_{i=1}^N |SR_p - SR_o|}{SR_o} * 100 \right) \quad (6)$$

$$R^2 = \left(\frac{N * (\sum SR_o * SR_p) - (\sum SR_o) * (\sum SR_p)}{\sqrt{(N * \sum SR_o^2) - (\sum SR_o)^2} * \sqrt{(N * \sum SR_p^2) - (\sum SR_p)^2}} \right)^2 \quad (7)$$

Eşitliklerde kullanılan N veri sayısı, SR_p modelde tahmin edilen solar radyasyon değerini, SR_o gerçekte gözlenen solar radyasyon değerini ifade etmektedir. Bu çalışmada kullanılan veriler MGM’den temin edilen solar radyasyon verileridir ve birimler $kWsa/m^2$ cinsindedir. Karşılaştırma kriterlerimizden Karekök Ortalama Karesel Hata, Ortalama Mutlak Hata ve Ortalama Mutlak Bağlı Hata, hata sonuçları olduğundan modeldeki hata sonuçlarının birimleri de aynı şekilde $kWsa/m^2$ ’dir. Hata sonuçlarının yüksek olması modelin gerçek veriye göre uzakta yani kötü sonuç verdiğini anlatır. Düşük olması ise modelin gerçek veriye göre yakın, hatası düşük ve modelin uygun olduğunun göstergesidir. R^2 ’nin 1 olması, deneysel verilerin kusursuz bir doğrusal eğri sağlandığının kanıtıdır. R^2 ’nin 0 olduğu durumlar testin tamamen başarısız olduğunu göstermektedir. M5 Model Ağacı modellemesinde 1 giriş (ay) 1 çıkış (SR tahmini), 2 giriş (ay ve enlem) 1 çıkış (SR tahmini) ve 3 giriş (ay-enlem-boylam) 1 çıkış (SR tahmini) şeklinde üç farklı giriş veri seti oluşturulmuştur. Modelleme sonucunda elde edilen karşılaştırma kriterleri Tablo 4 ve Tablo 5’te verilmiştir. Çalışmada kullanılan verilerin %75’i eğitim aşamasında, %25’i ise test aşamasında kullanılmıştır.

Tablo 4. Model 5 Karar Ağacı (M5-Tree) modeli eğitim sonuçları

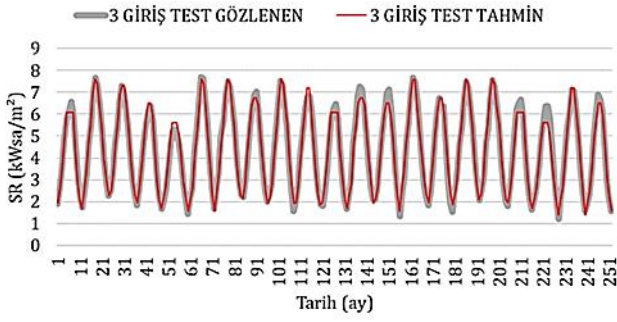
Değerlendirme Kriterleri	Girişler (eğitim)		
	1 Giriş	2 Giriş	3 Giriş
KOKH	0.4227	0.1997	0.1749
OMH	0.3359	0.1300	0.1085
OMBH	8.5477	3.2634	2.6475
R^2	0.9472	0.9882	0.9909

Tablo 5. Model 5 Karar Ağacı (M5-Tree) modeli test sonuçları

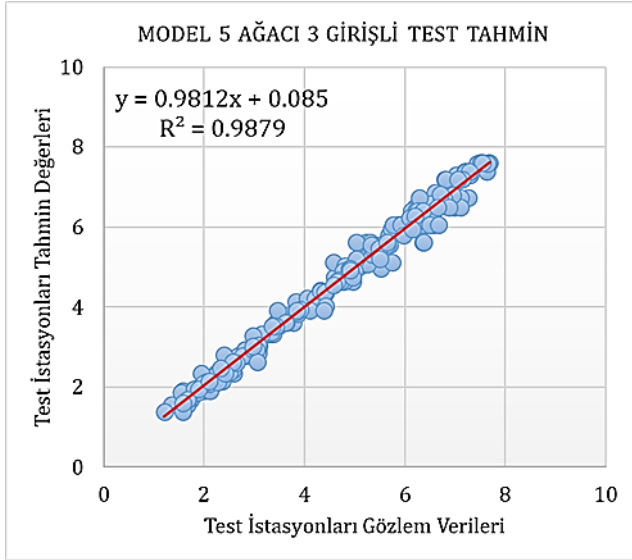
Değerlendirme Kriterleri	Girişler (Test)		
	1 Giriş	2 Giriş	3 Giriş
KOKH	0.4353	0.2638	0.3604
OMH	0.3543	0.1748	0.1451
OMBH	9.4156	4.2398	3.6029
R^2	0.9448	0.9797	0.9879

Analiz sonucunda test aşamasında en iyi tahmini 3 giriş-1 çıkışa ait model vermiştir. Bu modele ait gidış grafiği ve saçılma grafiği Şekil 3 ve Şekil 4’te yer almaktadır.

Şekil 3’te görüldüğü üzere model tahminlerinin gözlenen verilerle üst üste geldiği fakat bazı uç noktaları yakalayamadığı tespit edilmiştir.



Şekil 3. 3 girişli test aşamasına ait gidiş grafiği



Şekil 4. 3 girişli test aşamasına ait saçılma grafiği

Şekil 4 incelendiğinde modelin denklemi en uygun doğruya ($y=x$) oldukça benzerdir ve R^2 değeri 0.9879 olarak elde edilmiştir. Bu değer gözlenen ile tahmin edilen değerlerin birbirlerine oldukça yakın olduğunu göstermektedir. Türkiye'nin 2004-2018 yıllarını kapsayan gözlenmiş aylık ortalama SR değerleriyle IDW enterpolasyon haritası oluşturularak Şekil 5'te verilmiştir.

Şekil 5'te görüldüğü üzere Türkiye'nin kuzeyinden güneye doğru çizgisel şekilde SR değerleri artmaktadır. Özellikle Güneydoğu Anadolu Bölgesinde Ortadoğu ikliminin etkisiyle havanın nem oranı denize sınırı olan Akdeniz Bölgesine göre daha az olduğu için SR değerlerinin nispeten daha fazla olduğu tespit edilmiştir.

Çalışmada kullanılan 21 adet test istasyonuna ait gözlenen ve tahmin edilen verilerin saçılma grafikleri ise Şekil 6'da verilmiştir.

Grafiklerde de görüldüğü gibi M5-tree yöntemiyle yapılan tahmin çalışmalarında tahmin değerleri gözlenen değerlerle çoğu noktada örtüşse de özellikle pik noktalarda gözlenen değeri yakalayamamıştır.

Literatürde yer alan SR tahmin çalışmalarının sonuçları incelendiğinde, Sözen ve Arcaklıoğlu (2005) YSA kullanarak oluşturdukları 3 farklı algoritmayla gerçekleştirdikleri çalışmada 12 istasyon içinden en iyi sonuç veren istasyonun (Artvin ist.) R^2 değerlerine bakıldığında 0,999 olduğu en kötü sonuç veren istasyonun (Çanakkale ist.) R^2 değerlerine bakıldığında 0,998 olduğu görülmüştür. Özgören ve ark. (2012) YSA kullanarak gerçekleştirdikleri çalışmalarında 10 model

inden en iyi sonucu veren modelin R^2 değerinin 0,9872, en kötü sonucu veren R^2 değerinin ise 0,9771 olduğu görülmüştür. Wang ve ark. (2017) ANFIS-SC, ANFIS-GP ve M5 Model Ağacı kullanarak oluşturdukları 3 farklı model ile gerçekleştirdikleri çalışmada M5 Model Ağacı tahmini için 21 bölge içinden en iyi sonuç veren istasyonun 51777 olduğu ve OMH değerinin 2,1 ($\text{MJ}/\text{m}^2/\text{gün}$), KOKH değerinin 2,79 (MJ/m^2) ve R^2 değerinin 0,85 olduğu, en kötü sonuç veren istasyonun 53487 olduğu ve OMH 2,95 ($\text{MJ}/\text{m}^2/\text{gün}$), KOKH değerinin 3,87 (MJ/m^2) ve R^2 değerinin 0,74 olduğu görülmüştür.

Kuncan ve Şahin (2017) YSA kullanarak gerçekleştirdikleri çalışmada 20 model içinden en iyi sonuç veren modelin KOKH değerine bakıldığında 1,550 (MJ/m^2) olduğu ve çalışmalarındaki en başarısız tahminine ait KOKH değerinin 1,782 (MJ/m^2) olduğu görülmüştür. Aksoy ve Bolat (2017) YSA kullanarak gerçekleştirdikleri çalışmada 13 model içinden en iyi sonuç veren modelin KOKH değerine bakıldığında 1,277 (MJ/m^2) olduğu ve çalışmalarındaki en başarısız tahminine ait KOKH değerinin 1,669 (MJ/m^2) olduğu görülmüştür.

Ruşen (2018) SR bileşenlerini araştırdığı çalışmasında HELIOSAT yöntemini, en yaygın uydu tabanlı küresel güneş ışınımı tahmin modeli olduğunu belirtmiştir.

Keshtegar ve ark. (2018) Kriging, yanıt yüzeyi yöntemi (RSM), çok değişkenli adaptif regresyon (MARS) ve M5-tree gibi dört farklı sezgisel regresyon ile gerçekleştirdikleri çalışmada en iyi M5-tree modelinin Antakya istasyonunda 5 girişli veri setinde elde etmişlerdir. Modellerine periyodiklik bileşeninin de eklenmesiyle en iyi OMH değerinin 5,81 ($\text{MJ}/\text{m}^2/\text{gün}$), en iyi KOKH değerinin 7,63 (MJ/m^2) olduğu belirtmişlerdir. Periyodiklik parametresinin eklenmediği modellemelerde ise en iyi M5-tree modelinin Antakya İstasyonunda 5 girişli modelinde en iyi OMH değerinin 9,26 ($\text{MJ}/\text{m}^2/\text{gün}$), en iyi KOKH değerinin 12,35 (MJ/m^2), olduğu belirtmişlerdir. Bu çalışmadan da anlaşılacağı üzere periyodiklik bileşeni tahmin performansını artırmaktadır.

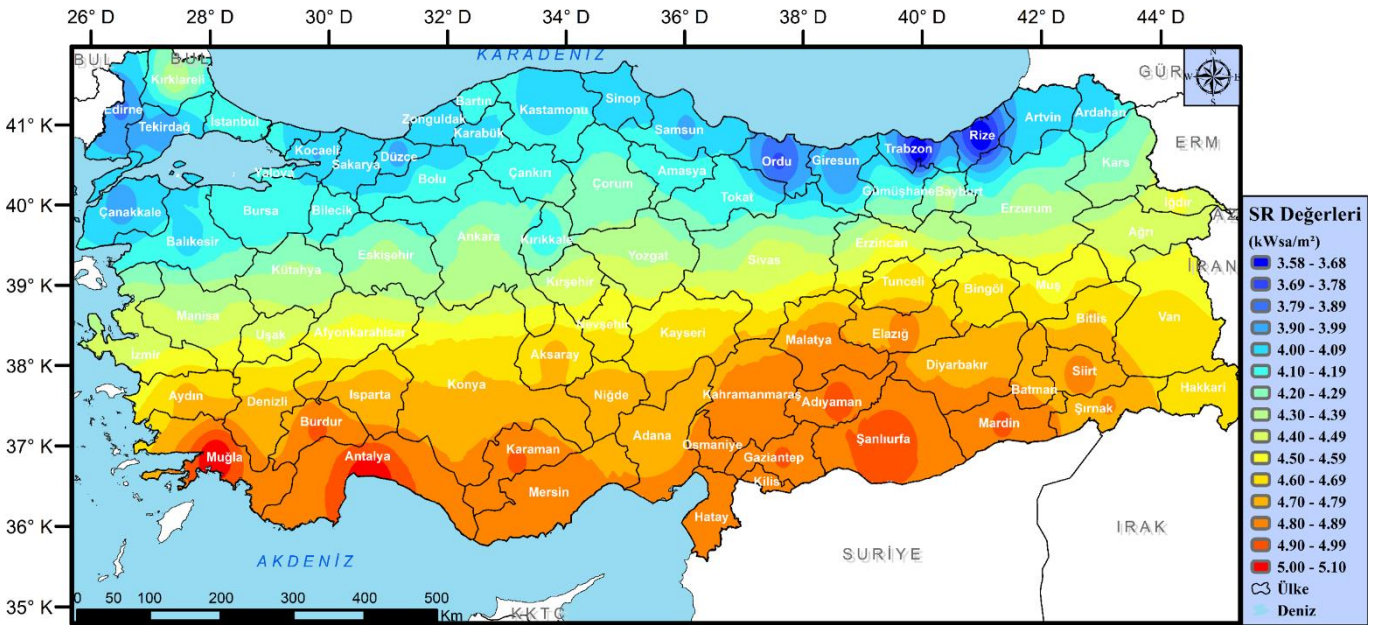
Kaba ve ark. (2018) Derin öğrenme kullanarak gerçekleştirdikleri çalışmada 16 kombinasyon içinden en iyi sonuç veren kombinasyonu olan C16 kombinasyonunun OMH, KOKH ve R^2 değerlerine bakıldığında OMH:0,61 ($\text{MJ}/\text{m}^2/\text{gün}$), KOKH:0,78 (MJ/m^2), R^2 :0,98, en kötü kombinasyonun tahminine bakıldığında ise C01 kombinasyonunun OMH, KOKH ve R^2 değerlerine bakıldığında OMH:1,52 ($\text{MJ}/\text{m}^2/\text{gün}$), KOKH:1,94 (MJ/m^2), R^2 :0,867 olduğu görülmüştür. Alizamir ve ark. (2020) 6 farklı makine öğrenmesi tekniği kullanarak gerçekleştirdikleri çalışmalarında Antalya/Mersin İstasyonlarında en iyi sonuçların Gradyan Artırma Ağacı yönteminde olduğu (KOKH:8,747/7,352 (MJ/m^2), OMH:7,056/5,61 ($\text{MJ}/\text{m}^2/\text{gün}$), R^2 :0,946/0,93) en kötü sonuçların ise Sınıflandırma ve Regresyon Ağacı yönteminde (KOKH:17,073/16,282 (MJ/m^2), OMH:14,471/11,141 ($\text{MJ}/\text{m}^2/\text{gün}$), R^2 :0,839/0,72) olduğu görülmüştür. Çıtakoğlu ve Demir (2021) Derin Öğrenme kullanarak gerçekleştirdikleri çalışmada 81 istasyonu tahmininde en iyi sonuç veren istasyonun değerlerine bakıldığında OMH değerinin 0,17 ($\text{MJ}/\text{m}^2/\text{gün}$), KOKH değerinin 0,228

(MJ/m²) ve R² değerinin 0,987, en kötü sonucu veren istasyonun ise OMH değerinin 0,174 (MJ/m²gün), KOKH değerinin 0,238 (MJ/m²) ve R² değerinin 0,985 olduğu görülmüştür. Ağbulut ve ark. (2021) 4 farklı makine öğrenmesi algoritması kullanarak gerçekleştirdikleri çalışmalarında en iyi sonucu veren modelin KOKH değerinin 2,157 (MJ/m²) ve R² değerinin 0,936, en kötü sonucu veren modelin KOKH değerinin 2,776 ve R² değerinin 0,883 olduğu görülmüştür.

Model ağaçlarının diğer yumuşak hesaplama araçlarından farklı olarak en büyük avantajı, daha kolay ve hızlı kullanımları yanı sıra anlaşılır matematiksel kuralları olmasıdır. Ayrıca çok büyük veri kümelerinde bile kullanılabilir olmalarının yanında kabul edilebilir sonuçlar vermesi yönüyle de diğer yöntemlere göre oldukça avantajlıdır (Bonakdar ve Etemad-Shahidi, 2011). Ayrıca Pal (2007) model ağacı yöntemlerinden

M5-tree model ağacının, sürekli sınıfları tahmin etmede oldukça başarılı olduğunu belirtmektedir (Pal, 2007).

Bu çalışmada elde edilen sonuçlar literatürle karşılaştırıldığında gerek KOKH gerekse R² değerlerinin literatürle uyumlu olduğu ve M5-tree'nin alternatif bir yöntem olabileceği düşünülmektedir. M5-tree yöntemi özellikle literatürde yer alan makine öğrenmesi yöntemlerinde olduğu gibi bir parametreye bağımlı değildir ve bir optimizasyon süreci barındırmamaktadır. Bu da onu diğer modellere göre daha kararlı yapmaktadır. Başka bir ifade ile başlangıç ağırlık ataması hücreler arasında yapılmadığı için model kaç defa çalıştırılırsa çalıştırılsın sonuçlar değişmemektedir. Ayrıca derin öğrenme gibi yöntemlere göre model başarısı düşük olsa da modelleme hızı oldukça fazladır. Son olarak model sonucu model hiper parametresine bağlı olmayıp sınırlamalara bağlı denklemlerden oluşmaktadır.



Şekil 5. SR ortalama değerleri için hazırlanmış IDW haritası

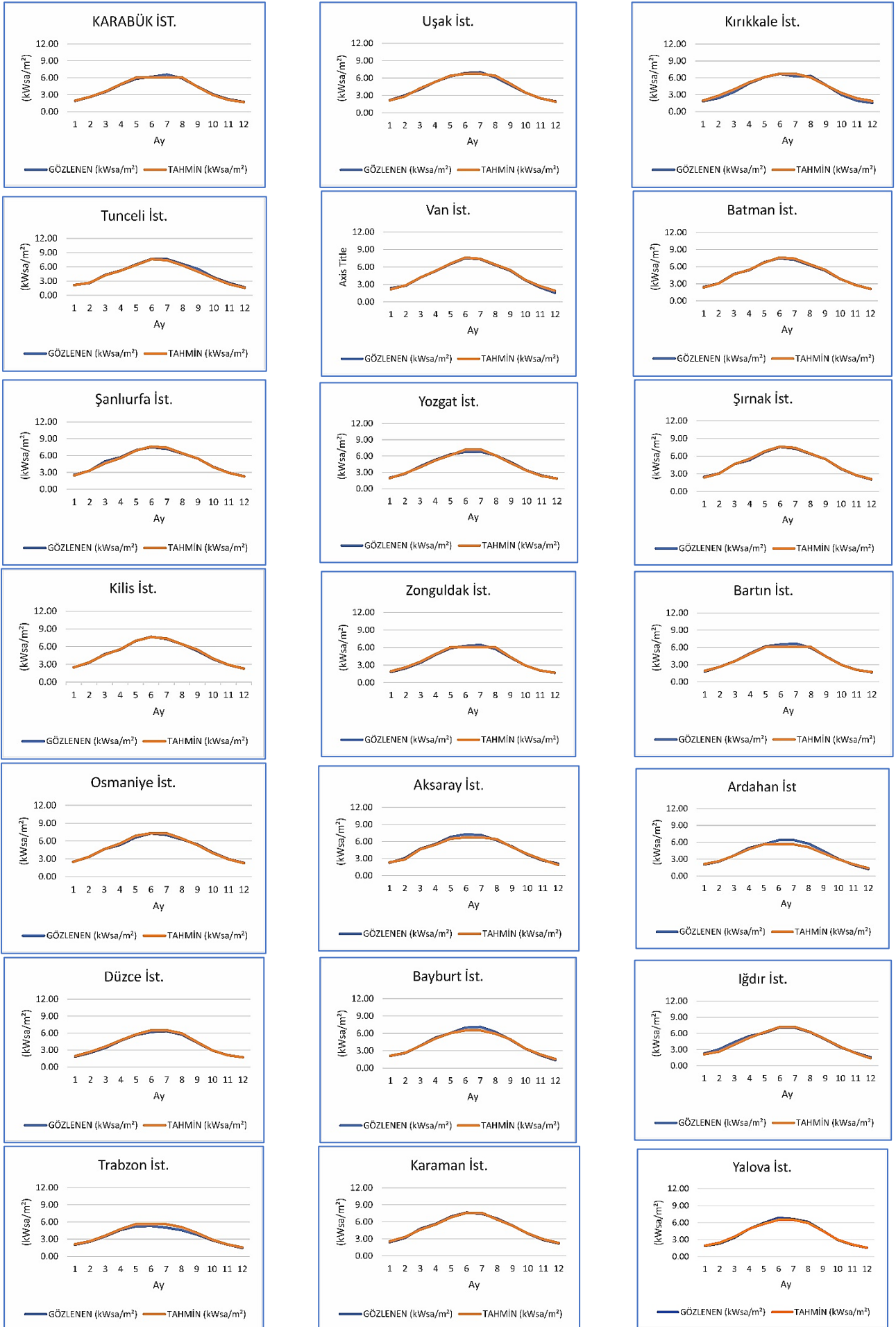
4. Sonuç

Bu çalışmada, Türkiye'nin 2004-2018 yılları arasındaki aylık ortalama solar radyasyon değerleri 1 giriş 1 çıkış, 2 giriş 1 çıkış, 3 giriş 1 çıkış olmak üzere çeşitli 3 farklı giriş veri seti (ay (1 giriş), ay-enlem (2 giriş), ay-enlem-boylam (3 giriş)) (kWsa/m²) ve M5 tree sezgisel regresyon tekniği kullanılarak tahmin edilmiştir. Çalışmada istasyonların %75'lik kısmı eğitim (60 adet), %25'lik kısmı ise (21 adet) test aşamalarında kullanılmıştır. Karşılaştırma kriteri olarak Karekök ortalama karesel hata (KOKH), ortalama mutlak hata (OMH), Ortalama Mutlak Bağlı Hata (OMBH), ve determinasyon katsayısı (R²) kullanılmıştır.

Sonuçlar incelendiğinde; M5-tree sezgisel regresyon tekniğinde en iyi sonuç 3 giriş (ay-enlem-boylam) 1 çıkışlı (SR tahmin değeri) algoritmasıyla yapılan analizde görülmektedir. M5-tree yönteminde 1 giriş sonuçları;

(KOKH= 0.4353, OMH= 0.3543, OMBH=9.4156, R²= 0.9448), incelendiğinde sonuçlar tatmin edici olsa da en az hataya sahip tahminleri 3 girişli model vermiştir ve sonuçları; (KOKH= 0.3604, OMH= 0.1451, OMBH= 3.6029, R²= 0.9879) şeklindedir. Bu sonuçlar literatürle karşılaştırıldığında literatürde kullanılan yöntemlerle uyumlu olduğu görülmüştür.

Gelecek çalışmalarda, her bir istasyonun ayrı ayrı modellenmesi ve model performansına etkisi araştırılacaktır. Farklı yöntemlerle (destek vektör makineleri, derin öğrenme, rastgele orman vb.) modellemeler yapılacak, sonuçlar hem hata kriterleri bakımından hem de model süreleri bakımından karşılaştırılacaktır. Ayrıca eğitim ve test oranları değiştirilerek eğitim-ters oranlarının performansa etkisi araştırılacaktır. Son olarak test istasyonlarının sürekli değiştirilmesi suretiyle en başarılı yöntem araştırılacaktır.



Şekil 6. Test istasyonlarında özlenen ve tahmin edilen SR değerleri

Bilgilendirme/Teşekkür

Yazarlar desteklerinden dolayı KTO Karatay Üniversitesine teşekkür etmektedir.

Araştırmacıların katkı oranı

Taha Demirgöl: Literatür taraması, Modelleme, Makale yazımı; **Vahdettin Demir:** Kurgu, Revizyonlar, **Mehmet Faik Sevimli:** Son kontrol.

Çatışma Beyanı

Herhangi bir çıkar çatışması bulunmamaktadır.

Kaynakça

Ağbulut, Ü., Gürel, A. E., & Biçen, Y. (2021). Prediction of daily global solar radiation using different machine learning algorithms: Evaluation and comparison. *Renewable and Sustainable Energy Reviews*, 135, 110114.

Ahmad, M. J., & Tiwari, G. N. (2011). Solar radiation models—A review. *International Journal of Energy Research*, 35(4), 271-290.

Aksoy, S., & Bolat, M. (2017). Akdeniz Bölgesinin Noaa/Avhrr Uydu Verileri Kullanılarak Solar Radyasyonun YSA Metodu ile Hesaplanması. *International Conference on Multidisciplinary, Science, Engineering and Technology (IMESET'17 Bitlis)*, 654–660.

Alizamir, M., Kim, S., Kisi, O., & Zounemat-Kermani, M. (2020). A comparative study of several machine learning based non-linear regression methods in estimating solar radiation: Case studies of the USA and Turkey regions. *Energy*, 197, 117239.

Bahadırılı, Z. M. (2021). *Yeraltı suyu seviye değişiminin M5 karar ağacı ve yapay sinir ağları yöntemleriyle belirlenmesi* (Master's thesis, İskenderun Teknik Üniversitesi/Lisansüstü Eğitim Enstitüsü/İnşaat Mühendisliği Ana Bilim Dalı).

Bhardwaj, S., Sharma, V., Srivastava, S., Sastry, O. S., Bandyopadhyay, B., Chandel, S. S., & Gupta, J. R. P. (2013). Estimation of solar radiation using a combination of Hidden Markov Model and generalized Fuzzy model. *Solar Energy*, 93, 43-54.

Bonakdar, L., & Etemad-Shahidi, A. (2011). Predicting wave run-up on rubble-mound structures using M5 model tree. *Ocean Engineering*, 38(1), 111-118.

Bora, E., Tekelli, M., & Etöz, M. (2015). Menemen Ovasında Referans Bitki Su İhtiyacı ve Ölçülen Solar Radyasyonun Zamansal Değişimi. *Anadolu Ege Tarımsal Araştırma Enstitüsü Dergisi*, 24(2), 65-76.

Cornejo-Bueno, L., Casanova-Mateo, C., Sanz-Justo, J., & Salcedo-Sanz, S. (2019). Machine learning regressors for solar radiation estimation from satellite data. *Solar Energy*, 183, 768–775.

Citakoglu, H. (2015). Comparison of artificial intelligence techniques via empirical equations for prediction of solar radiation. *Computers and Electronics in Agriculture*, 118, 28-37.

Çitakoğlu, H., & Demir, V. (2021). Solar Radyasyonun Derin Öğrenme Kullanılarak Tahmin Edilmesi:

Türkiye Örneği. *International World Energy Conference*, 167–175.

Demirci, M. (2019). Destek Vektör Makineleri ve M5 Karar Ağacı Yöntemleri Kullanılarak Yağış Akış İlişkinin Tahmini. *Dicle Üniversitesi Mühendislik Fakültesi Mühendislik Dergisi*, 10(3), 1113-1124.. <https://doi.org/10.24012/dumf.525658>

Demirgöl, T., Yılmaz, C. B., Zıpır, B. N., Kart, F. S., Pehriz, M. F., Demir, V., & Sevimli, M. F. (2022). Investigation of Turkey's climate periods in terms of precipitation and temperature changes. *Engineering Applications*, 1(1), 80-90.

Feng, Y., Gong, D., Zhang, Q., Jiang, S., Zhao, L., & Cui, N. (2019). Evaluation of temperature-based machine learning and empirical models for predicting daily global solar radiation. *Energy conversion and management*, 198, 111780. <https://doi.org/10.1016/J.ENCONMAN.2019.111780>

Granata, F., Papirio, S., Esposito, G., Gargano, R., & De Marinis, G. (2017). Machine learning algorithms for the forecasting of wastewater quality indicators. *Water*, 9(2), 105. <https://doi.org/10.3390/w9020105>

Guermoui, M., Melgani, F., Gairaa, K., & Mekhalfi, M. L. (2020). A comprehensive review of hybrid models for solar radiation forecasting. *Journal of Cleaner Production*, 258, 120357. <https://doi.org/10.1016/j.jclepro.2020.120357>

Gül, M., & Çelik, E. (2017). ANFIS kullanılarak Tunceli ili için global güneş radyasyonu tahmini. *Dicle Üniversitesi Mühendislik Dergisi*, 8(4), 891–899.

Hastaoğlu, K. Ö., Göğsu, S., & Gül, Y. (2022). Determining the relationship between the slope and directional distribution of the UAV point cloud and the accuracy of various IDW interpolation. *International Journal of Engineering and Geosciences*, 7(2), 161-173. <https://doi.org/10.26833/ijeg.940997>

İşcan, F., & Ilgaz, A. (2017). Analysis Of Geographic/Urban Information System Web Presentations Of Local Governments In Turkey. *International Journal of Engineering and Geosciences*, 2(3), 75–83. <https://doi.org/10.26833/ijeg.317088>

Ji, Y., Lei, X., Cai, S., & Wang, X. (2016). Application of a classifier based on data mining techniques in water supply operation. *Water (Switzerland)*, 8(12). <https://doi.org/10.3390/w8120599>

Jiang, Y. (2009). Computation of monthly mean daily global solar radiation in China using artificial neural networks and comparison with other empirical models. *Energy*, 34(9), 1276–1283. <https://doi.org/10.1016/J.ENERGY.2009.05.009>

Kaba, K., Sarıgül, M., Avcı, M., & Kandırmaz, H. M. (2018). Estimation of daily global solar radiation using deep learning model. *Energy*, 162, 126–135. <https://doi.org/10.1016/j.energy.2018.07.202>

Keshtegar, B., Mert, C., & Kisi, O. (2018). Comparison of four heuristic regression techniques in solar radiation modeling: Kriging method vs RSM, MARS and M5 model tree. In *Renewable and Sustainable Energy Reviews* (Vol. 81, pp. 330–341). Elsevier Ltd. <https://doi.org/10.1016/j.rser.2017.07.054>

- Kılıç, B., & Kumaş, K. (2016). Burdur İli Güneşlenme Değerlerinin Yapay Sinir Ağları Metodu ile Tahmini. *SDU Teknik Bilimler Dergisi*, 6(1), 38–44.
- Kisi, O. (2015). Pan evaporation modeling using least square support vector machine, multivariate adaptive regression splines and M5 model tree. *Journal of Hydrology*, 528, 312–320. <https://doi.org/10.1016/j.jhydrol.2015.06.052>
- Krige, D. G. (1951). A Statistical Approach To Some Basic Mine Valuation Problems On The Witwatersrand. *Chemical Journal of the Metallurgical & Mining Society of South Africa*, 52(6), 119–139.
- Kuncan, F., & Şahin, M. (2017, September). Yapay sinir ağı ve uydu dataları kullanılarak güneş radyasyonunun tahmini. In 2017 International Artificial Intelligence and Data Processing Symposium (IDAP) (pp. 1-10). IEEE.
- Özgören, M., Bilgili, M., & Şahin, B. (2012). Estimation of global solar radiation using ANN over Turkey. *Expert Systems with Applications*, 39(5), 5043–5051. <https://doi.org/10.1016/j.eswa.2011.11.036>
- Pal, M. (2007). M5 model tree for land cover classification. *International Journal of Remote Sensing*, 27(4), 825–831. <https://doi.org/10.1080/01431160500256531>
- Pal, M. & Deswal, S. (2009). M5 model tree based modelling of reference evapotranspiration. *Hydrological Processes*, 23(10), 1437–1443. <https://doi.org/10.1002/hyp.7266>
- Rusen, S. E. (2018). Modeling and Analysis of Global and Diffuse Solar Irradiation Components Using the Satellite Estimation Method of HELIOSAT. *CMES*, 115(3), 327–343. <https://doi.org/10.3970/cmcs.2018.00159>
- Quinlan, J. R. (1992). Learning with Continuous Classes. *Proceedings of Australian Joint Conference on Artificial Intelligence*, Hobart 16-18 November 1992, 343-348.
- Shepard, D. (1968). A two-dimensional interpolation function for irregularly-spaced data. In *Proceedings of the 1968 23rd ACM national conference (ACM '68)*. Association for Computing Machinery, New York, NY, USA, 517–524. <https://doi.org/10.1145/800186.810616>
- Sözen, A. & Arcaklioğlu, E. (2005). Solar potential in Turkey. *Applied Energy*, 80(1), 35–45. <https://doi.org/10.1016/j.apenergy.2004.02.003>
- Şeker, M. (2021). Yapay Sinir Ağı (YSA) Kullanılarak Meteorolojik Verilere Dayalı Solar Radyasyon tahmini. *DEU Mühendislik Fakültesi Fen ve Mühendislik*, 23(69), 923–935. <https://doi.org/10.21205/deufmd.2021236920>
- Taktak, F., & İli, M. (2018). Güneş Enerji Santrali (GES) Geliştirme: Uşak Örneği. *Geomatik*, 3(1), 1-21.
- Taylan, E. D., & Damçayırı, D. (2016). Isparta bölgesi yağış değerlerinin IDW ve Kriging enterpolasyon yöntemleri ile tahmini. *Teknik Dergi*, 27(3), 7551-7559.
- URL-1: <https://www.mgm.gov.tr/FTPDATA/arastirma/radyasyon/heliosat.pdf>
- URL-2: https://mgm.gov.tr/kurumici/radyasyon_iller.aspx
- Wang, L., Kisi, O., Zounemat-Kermani, M., Zhu, Z., Gong, W., Niu, Z., Liu, H., & Liu, Z. (2017). Prediction of solar radiation in China using different adaptive neuro-fuzzy methods and M5 model tree. *International Journal of Climatology*, 37, 1141–1155. <https://doi.org/10.1002/joc.4762>
- Yaseen, Z. M. (2021). An insight into machine learning models era in simulating soil, water bodies and adsorption heavy metals: Review, challenges and solutions. *Chemosphere*, 277. <https://doi.org/10.1016/j.chemosphere.2021.130126>
- Yılmaz, C. B., Bodu, H., Yüce, E. S., Demir, V., & Sevimli, M. F. (2023). Türkiye'nin uzun dönem ortalama sıcaklık (°C) değerlerinin üç farklı enterpolasyon yöntemi ile tahmini. *Geomatik*, 8(1), 9-17. <https://doi.org/10.29128/geomatik.984310>

