

Ege Denizi'nde makine öğrenimi yöntemleri ile anlık deniz seviyesi değişimlerinin tahmini

Ahmet Yavuzdoğan^{1*} , Emine Tanır Kayıkcı² 

¹Gümüşhane Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, Harita Mühendisliği Bölümü, Gümüşhane, Türkiye.

²Karadeniz Teknik Üniversitesi, Mühendislik Fakültesi, Harita Mühendisliği Bölümü, Trabzon, Türkiye.

Öz: Anlık deniz seviyesinin tahmini, jeodezik dişey datumun belirlenmesi ve güncellenmesi, kıyı alanlarının korunması, kıyı ekosistemlerinin izlenmesi, kıyı yapılarının planlanması ve bakımı, iklim deęişikliği etkilerinin gözlenmesi açısından büyük önem taşımaktadır. Anlık deniz seviyesi tahmini için kullanılan geleneksel yöntemler genellikle doğrusal varsayımlara dayanmaktadır. Ancak deniz seviyesini etkileyen faktörler çok çeşitlidir ve etkileri bölgeden bölgeye deęişmektedir. Genellikle doğrusal olmayan ve karmaşık bağımlılık yapılarına sahiptirler. Bu nedenle, doğrusal olmayan deniz seviyeleri doğrusal modeller kullanılarak yüksek duyarlılıkta belirlenemez. Makine öğrenimi tahmin yöntemleri ise, son zamanlarda deęişkenler arasındaki karmaşık bağımlılık yapılarının modellenmesinde sıklıkla kullanılmaktadır. Bu çalışma kapsamında, anlık deniz seviyesini yüksek doğrulukta tahmin etmek ve doğrusal tahmin yöntemleri ile doğrusal olmayan tahmin yöntemlerini karşılaştırmak amacıyla makine öğrenimi tahmin yöntemlerinden Çoklu Doğrusal Regresyon (ÇDR) doğrusal modeli, Destek Vektör Regresyonu (DVR) doğrusal olmayan model ve Rastgele Orman Regresyonu (ROR) doğrusal olmayan model algoritmaları kullanılmış ve tahmin performansları karşılaştırılmıştır. Çalışma sonucunda anlık deniz seviyesi için en yüksek tahmin performansı ROR ile elde edilmiş olup, en düşük tahmin performansı ise ÇDR yöntemi ile elde edilmiştir. Sonuç olarak anlık deniz seviyelerinin çalışmada kullanılan öncül bilgiler ile ROR kullanılarak yüksek hassasiyette tahmin edilebileceği ve doğrusal tahmin modelinin anlık deniz seviyesinin karmaşık bağımlılık yapısının modellenmesinde yetersiz olduğu gösterilmiştir.

Anahtar Sözcükler: Deniz seviyesi, Makine öğrenimi, Ege Denizi, Çoklu doğrusal regresyon, Destek vektör, Rastgele orman

Forecasting instantaneous sea level changes with machine learning methods in Aegean Sea

Abstract: Forecasting instantaneous sea-level is of great importance in terms of determination of geodetic vertical datum and updating, conservation of coastal areas, monitoring coastal ecosystems, maintenance, and planning of coastal structures, monitoring of climate change effects. Traditional methods used for instantaneous sea level estimation are often based on linear assumptions. However, contributors to sea levels are very various and their effects vary from region to region. Generally, they have complex and nonlinear dependence structures. Therefore, nonlinear sea-level cannot be determined with high precision using linear models. Recently, machine learning prediction methods have been frequently used in the modelling of complex dependency structures between variables. Within the scope of this study, to predict the instantaneous sea level with high accuracy and to compare linear estimation methods with nonlinear estimation methods, the Multiple Linear Regression (MLR) linear model, Support Vector Regression (SVR) non-linear model, and Random Forest Regression (RFR) non-linear model algorithms were used, and their prediction performances were compared. As a result of the study, the highest prediction performance for instantaneous sea level was obtained with RFR, and the lowest prediction performance was obtained with the MLR method. As a result, it has been shown that instantaneous sea level can be predicted with high precision using RFR with the features used in this study, and the linear prediction models are insufficient in modelling the complex dependency structure of instantaneous sea level.

Keywords: Sea level, Machine learning, Aegean Sea, Multiple linear regression, Support vector, Random forest

* Sorumlu Yazar/Corresponding Author: Tel: +90 456 233 1754

Geliş Tarihi/Received: 04.07.2020
Kabul Tarihi/Accepted: 21.12.2020



1. Giriş

Geleceđe yönelik deniz seviyelerinin yüksek doğrulukta tahmin edilmesi Türkiye gibi üç tarafı denizler ile çevrili bir ülke için kıyı alanlarının korunması, kıyı ekosistemlerinin izlenmesi, kıyı yapılarının bakımı ve planlanması, küresel ısınmanın etkilerinin gözlenmesi ve düşey kontrol ağlarının datumunun belirlenmesi açısından son derece önemlidir. Ayrıca deniz seviyesinin yüksek doğrulukta tahmin edilmesi geleceđe yönelik oluşturulacak etkin su ve kıyı politikalarının hazırlanması açısından da büyük önem taşımaktadır. Deniz seviyesi tahmini için kullanılan geleneksel yöntemler genellikle doğrusal varsayımlara dayanmaktadır. Deniz seviyesi deđişimini etkileyen sıcaklık, akış, su tuzluluđu, buharlaşma, yağış gibi hem cođrafi hem de meteorolojik birçok deđişken ise doğrusal olmayan bir yapıya sahiptir (Imani, You, & Kuo, 2014a). Üstelik bu etkiler bölgeden bölgeye deđişebilen ve oldukça karmaşık bağımlılık yapılarına sahiptir. Bu nedenle, doğrusal olmayan deniz seviyesi zaman serileri doğrusal modeller kullanılarak doğru şekilde temsil edilemez. Üstelik birçok deđişkene bađlı olarak deđişen deniz seviyesinin yüksek doğrulukta kestirimi zorlaşmaktadır (Talebizadeh & Moridnejad, 2011).

Geçmişte deniz seviyesi tahmini için birçok çalışma yapılmıştır. Irvine ve Eberhardt (1992), deniz seviyesini tahmin etmek için aylık ortalama deniz seviyesi verilerini kullanarak Erie Gölü ve Ontario Gölü için otoregresif bütünleşik hareketli ortalama (Autoregressive Integrated Moving Average, ARIMA) modellerini geliştirdi. Brundrit (1995), 8 mareograf istasyonunda deniz seviyesini modellemek için ARIMA modellerini kullandı. Vaziri (1997), 12 aylık Hazar Denizi seviyesini tahmin etmek için ARIMA ve Yapay Sinir Ağları (YSA) modellerini kullandı. Meshkani ve Meshkani (1997) Hazar Denizi'ndeki stokastik deniz seviyesindeki dalgalanmaları modellemek için ARIMA modellerini kullandı. Şen, Kadođlu ve Batur (2000) göl seviyesi kayıtlarındaki deterministik kısmı modellemek için basit doğrusal ve periyodik doğrusal olmayan modeller kullanmıştır. Söz konusu çalışmada kullanılan doğrusal model, basit bir doğrusal regresyondur. Göl seviyesi kayıtlarının doğrusal olmayan modellemesi için ikinci dereceden bir Markov modeli ve Fourier serisi kullanılmıştır. Srivastava vd. (2016), üstel yumuşatma durum-uzay modelleri (Exponential smoothing state-space, ESM) ve ARIMA modelini kullanarak, deniz seviyesindeki yükselme için nicel tahmin yöntemlerinin birleşik kullanımını araştırmışlardır. Fu, Zhou, Sun ve Tang (2019), deneysel mod ayrıştırma, tekil spektrum analizi ve uydudan türetilen deniz seviyesi anomali tahmini için en küçük kareleri birleştiren bir hibrit model uygulamıştır. Ancak ARIMA veya Box-Jenkins gibi bazı geleneksel yöntemler (Box, Jenkins, Reinsel, & Ljung, 2015), bir zaman serisinin doğrudan doğrusal bir süreçten üretildiđini varsaydıđından, doğrusal olmayan deniz seviyesi zaman serilerini iyi modelleyemeyebilir (Imani, You, & Kuo, 2014b).

YSA, Destek Vektör Makineleri (DVM), Karar Ağacı teknikleri ve bulanık mantık gibi yapay zekâ (Artificial Intelligence, AI) teknikleri; hidroloji ve su kaynakları mühendisliğinde doğrusal olmayan karmaşık olayları modellemek için sıklıkla kullanılmaktadır (Huang, Peng, Zhang, & Zhang, 2006; Kisi, Shiri, & Nikoofar, 2012; More & Deo, 2003; Wu & Chau, 2010). Son zamanlarda özellikle DVM'ler ve Rastgele Orman (Random Forest, RO) gibi bazı doğrusal olmayan makine öğrenimi yöntemleri de zaman serilerinin tahmininde sıklıkla kullanılmaktadır (Ghorbani, Khatibi, AYTEK, Makarynsky, & Shiri, 2010; Güven & Günal, 2008; Kim, 2003; Rajasekaran, Gayathri, & Lee, 2008; Yu, Chen, & Chang., 2006).

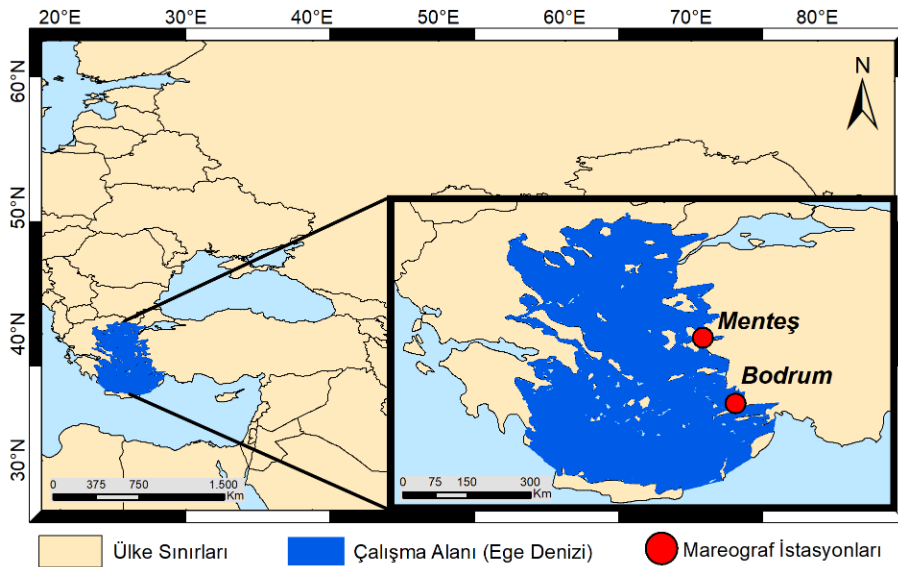
Ayrıca makine öğrenimi yöntemleri birçok kez anlık deniz seviyelerini modellemek amacıyla kullanılmıştır. Imani vd. (2014a) yaptığı çalışmada Hazar Denizi'nde DVM kullanarak anlık deniz seviyesini 35 mm karesel ortalama hata (Root Mean Square Error, RMSE) ile belirlemiştir. Ertuđrul ve Tađluk (2017) tarafından yapılan çalışmada ise ortalama deniz seviyeleri Genelleştirilmiş Davranışsal Öğrenme Metodu (Generalized Behavioral Learning Method, GBLM) yöntemi ile tahmin edilmiştir. Bu çalışmada GBLM ile ortalama deniz seviyeleri 69.152 mm ortalama hata ile belirlenmiştir. Imani, Kao, Lan ve Kuo (2018) yaptığı çalışmada Taivan Chiayi kıyılarında anlık deniz seviyesini belirlemek için radyal temelli fonksiyon (Radial Basis Function, RTF), özel bir vektör makinesi yöntemi olan RVM (Relevance Vector Machine), Aşırı

Öğrenme Makinesi (Extreme Learning Machine, ELM) ve DVM yöntemlerini kullanmıştır. En iyi performansı 34 mm RMSE ile RVM algoritması göstermiştir. [Roshni, Samui ve Drisya \(2019\)](#) de çalışmalarında anlık deniz seviyelerini tahmin etmek için 3 farklı makine öğrenimi algoritması kullanmıştır. Bu algoritmalar RVM, ELM ve Gauss Süreç Regresyonu (Gaussian Process Regression, GPR) algoritmalarıdır. Anlık deniz seviyesini tahmin edebilmek amacıyla meteorolojik değişkenlerden faydalanmışlardır. En iyi tahmin performansı 75 mm RMSE ve 0.96 korelasyon katsayısı ile GPR algoritması ile elde edilmiştir. Anlık deniz seviyesinin tahmini için birçok çalışmada YSA'lar da kullanılmıştır ([Imani, You, & Chung-Yen, 2013](#); [Imani vd., 2014b](#); [Makarynsky, Makarynska, Kuhn, & Featherstone, 2004](#); [Zhao, Fan, & Mu, 2019](#)).

Bu çalışma kapsamında anlık deniz seviyesini yüksek hassasiyetle tahmin edebilmek amacıyla 3 makine öğrenimi algoritması uygulanmış ve tahmin performansları karşılaştırılmıştır. Çalışma kapsamında kullanılan yöntemlerden Çoklu Doğrusal Regresyon (Multiple Linear Regression, ÇDR) algoritması ile anlık deniz seviyesi doğrusal bir model ile tahmin edilmiştir. Böylece doğrusal bir model ile karmaşık bağımlılık yapısına sahip anlık deniz seviyesi değişimlerinin ne kadar hassasiyetle tahmin edileceği gösterilmiştir. Deniz seviyesi tahmini için kullanılan bir diğer fonksiyon olan Destek Vektör Regresyonu (Support Vector Regression, DVR) algoritması ile radyal temelli fonksiyon içeren bir çekirdek fonksiyonu kullanılarak, doğrusal olmayan bir model ile anlık deniz seviyeleri modellenmiştir. Son olarak kullanılan RO Regresyonu (ROR) algoritması ile ağaç algoritma yapısındaki bir fonksiyon ile doğrusal olmayan bir model kullanılarak anlık deniz seviyesi tahmin etmek amaçlanmıştır. Ağaç algoritma yapısı ile çok karmaşık bağımlılıklar modellenilebilmekte ve doğrusal olmayan ilişkiler açığa çıkarılabilmektedir.

2. Çalışma Alanı ve Kullanılan Veriler

Bu çalışma kapsamında kullanılan iki mareograf istasyonuna (Menteş ve Bodrum) ait deniz seviyesi, hava sıcaklığı, hava nemi, hava basıncı, rüzgâr hızı ve rüzgâr yönü verileri Türkiye Ulusal Deniz Seviyesi İzleme Sistemi (TUDES) kapsamında Harita Genel Müdürlüğü tarafından sağlanan internet adresinden elde edilmiştir ([URL-1](#)) (Şekil 1). Ege Denizi'nde bulunan Mentеш ve Bodrum istasyonlarından elde edilen ham veriler saatlik alınmış olup, Mentеш istasyonu için veriler 1999 – 2014 aralığını, Bodrum istasyonu için ise 1999 – 2010 aralığını kapsamaktadır.



Şekil 1: Çalışma alanı ve mareograf istasyonları

3. Yöntem

3.1 Veri Önışleme İşlemleri

Mareograf istasyonlarında kaydedilen anlık deniz seviyesi ölçümlerinin çeşitli analiz algoritmalarında kullanılması için uyumsuz ölçü testlerinin yapılması gerekmektedir. Zaman serilerindeki uyumsuz ölçüler genellikle zaman serisinin diğer değerleri ile belirgin derecede farklılık gösterirler. Bu nedenle ilk önışleme adımı olarak deniz seviyesi zaman serilerindeki uyumsuz ölçülerin tespit edilerek ölçü kümesinden çıkarılması gerekmektedir. Ancak zaman serilerinde uyumsuz ölçü testi; trend, mevsimsel, dönemsel bileşen etkisi giderilmiş rastgele bileşene uygulanır. Bu çalışma kapsamında Bölüm 3.2.'de anlatıldığı şekilde deniz seviyesi zaman serilerinde trend bileşeni, basit doğrusal regresyon modeli ile mevsimsel değişimler ve dönemsel değişimler harmonik analiz yöntemiyle belirlenmiştir. Ardından trend, mevsimsel ve dönemsel bileşenden arındırılmış rastgele bileşene uyumsuz ölçü testi uygulanmıştır. Zaman serilerindeki uyumsuz ölçülerin tespit edilmesi için literatürde en sık kullanılan yöntemlerden biri k-sigma kuralıdır. Yöntemle, zaman serisinde rastgele bileşene ait ölçünün düzeltilmesi, birim ağırlıklı ölçünün ortalama hatasının k katını aşıyor ise, o ölçü uyumsuz ölçü olarak ele alınır. Uyumsuz ölçü testinde k sınır değeri olarak α yanılma olasılıklarına karşılık gelen ve normal dağılım tablosundan alınan 1.645, 2.58 ve 3.29 değerleri kullanılır. Genellikle k sınır değeri 3 olarak uygulanır. Bu durumda yöntem 3-sigma yöntemi olarak da anılır.

Bu çalışma kapsamında, mareograf istasyonlarından elde edilen anlık deniz seviyesi zaman serilerindeki uyumsuz ölçüler 3-sigma yöntemiyle tespit edilerek ölçü kümesinden çıkarılmıştır. Ardından deniz seviyesi zaman serilerindeki veri alınamayan zamanlara ait veri boşlukları hareketli medyan (moving median) yöntemine göre enterpolasyon ile tamamlanmıştır. Hareketli medyan yöntemi lineer enterpolasyon yöntemlerine göre zaman serisindeki doğrusal olmayan bağımlılıkları koruması nedeniyle tercih edilmektedir. Bu çalışmada ise deniz seviyesi zaman serilerindeki doğrusal olmayan bağımlılıkların korunması ve öğrenme algoritmalarının doğrusal olmayan ilişkileri de öğrenebilmeleri amacıyla veri boşlukları hareketli medyan yöntemi ile enterpolasyon yapılarak doldurulmuştur.

Çalışmada, bağımlı ve bağımsız değişkenlere ait veriler herhangi bir işlemde kullanılmadan önce ağırlık merkezine ötelenmiş ve normlandırılmıştır. Bu sayede matrislerde oluşabilecek kondisyon bozuklukları giderilmiştir. X_i bağımsız değişkenleri ve Y_i bağımlı değişkenleri için X_i' ve Y_i' ötelenmiş değerleri, n veri sayısını göstermek üzere;

$$X_{i_{ort}} = \frac{[X_i]}{n} \quad Y_{i_{ort}} = \frac{[Y_i]}{n} \quad (1)$$

$$X_i' = X_i - X_{i_{ort}} \quad Y_i' = Y_i - Y_{i_{ort}} \quad (2)$$

şeklinde elde edilir. X'' ve Y'' normlandırılmış değerler;

$$c_x = \sqrt{[X_i'^2]} \quad c_y = \sqrt{[Y_i'^2]} \quad (3)$$

$$X_i'' = \frac{X_i'}{c_x} \quad Y_i'' = \frac{Y_i'}{c_y} \quad (4)$$

formülleri ile hesaplanmıştır. Böylece X_i bağımsız değişkenleri ve Y_i bağımlı değişkenleri için X_i'' ve Y_i'' ötelenmiş ve normlandırılmış değerleri elde edilmiş ve makine öğrenimi algoritmalarında ötelenmiş ve normlandırılmış değişkenler kullanılmıştır.

3.2 Öncül Bilgi (Öznitelik) Çıkarımı

Uyuşumsuz ölçülerden arındırılmış deniz seviyesi değişim verileri ile anlık deniz seviyelerini tahmin etmek için kullanılacak öncül bilgilerin (öznitelik) elde edilmesi için en küçük kareler doğrusal regresyon ve klasik harmonik analiz yöntemleri kullanılmıştır. Burada;

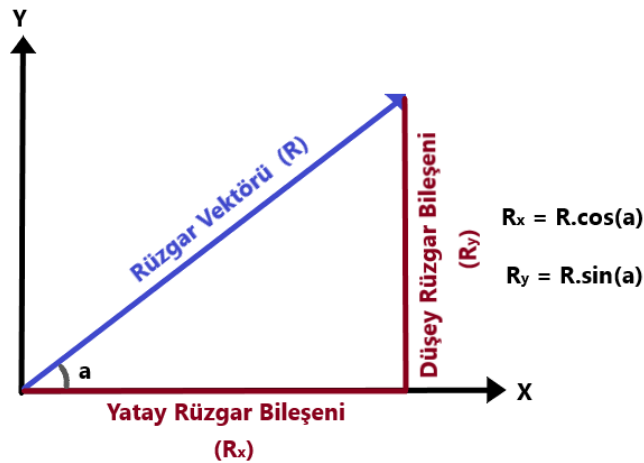
Doğrusal Model: Anlık deniz seviyesini tahmin etmek için kullanılan çok değişkenli tahmin modeli,

Doğrusal Trend: Anlık deniz seviyesinin uzun zaman içerisinde gösterdiği 1. dereceden olan eğilimi,

Doğrusal Regresyon: Doğrusal trendi ortaya çıkarmak için kullanılan yöntemi,

ifade etmektedir. Deniz seviyesi zaman serilerinin trend bileşeni, basit doğrusal regresyon modeli ile elde edilirken, mareograf istasyonlarındaki gelgit bileşenleri, mevsimsel değişimler ve dönemsel değişimler harmonik analiz yöntemi ile elde edilmiştir. Bu öncül bilgilerin Ege Denizi'ndeki anlık deniz seviyesinin temel bileşen özellikleri hakkında makine öğrenimi algoritmalarına öncül bilgi sağlaması amaçlanmaktadır. Bu nedenle makalenin geri kalanında deniz seviyesi zaman serilerinin trend ve harmonik bileşenleri “deniz seviyesi temel bileşenleri” olarak adlandırılmıştır.

Deniz seviyesini etkileyen önemli etkilerden biri olan rüzgarlar, vektörel bir büyüklük olduğundan, ham veri setinde bulunan rüzgar hızı ve rüzgar yönü değişkenleri ile mareograf istasyonundaki düşey ve yatay yönlerindeki rüzgar bileşenleri Şekil 2'deki gibi çıkarılmıştır.



Şekil 2: Yatay ve düşey rüzgar bileşenleri

3.2.1 Deniz Seviyesi Zaman Serisinin Doğrusal Trendinin Çıkarılması

Trend, uzun dönemdeki eğilimleri yansıtmaları bakımından zaman serisinin temel bileşenlerinden biridir. Deniz seviyeleri için ise tarihsel verilerden yola çıkılarak geleceğe yönelik deniz seviyesi değişimi hakkında önemli bilgiler içerir. Bu nedenle Ege Denizi'ndeki genel deniz seviyesi değişimi hakkında makine öğrenimi algoritmalarına bilgi sağlayabilmek amacıyla doğrusal trend, En Küçük Kareler Regresyonu (EKKR) ile elde edilmiştir. EKKR ile doğrusal trend denklemi;

$$Y_i + v_i = a_0 + a_1 X_i \quad (5)$$

şeklinde yazılır. Burada a_0 katsayısı doğrunun y ekseninde kestiği noktayı ve a_1 katsayısı ise doğrunun eğimini göstermektedir. v_i regresyon modeline ait düzeltme değerini, X_i bağımsız değişken, Y_i ise bağımlı değişken değerlerini göstermektedir. EKKR'ye göre regresyon katsayılarının hesaplanmasında regresyon modeli için düzeltme denklem sistemi,

$$\mathbf{y} + \mathbf{v} = \mathbf{A}\mathbf{x} \quad (6)$$

olarak kurulur. Burada \mathbf{A} katsayılar matrisi, \mathbf{x} bilinmeyenler vektörü ve \mathbf{v} düzeltme vektörüdür;

$$\mathbf{A} = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_n \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} \quad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \quad (7)$$

Burada n veri boyutudur. (6) eşitliğindeki düzeltme denklem sisteminin EKK ilkesine göre çözümü sonucunda bilinmeyenler

$$\mathbf{x} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{y} \quad (8)$$

olarak elde edilir.

3.2.2 Mevsimsel, Dönemsel ve Gelgit Bileşenlerinin Çıkarılması

Deniz seviyesi değişimlerinin en önemli nedenlerinden biri gelgitlerdir. Gelgitler deniz seviyeleri üzerinde oldukça düzenli ve harmonik etkilere sahiptir. Bu nedenle deniz seviyesi tahmin çalışmalarında gelgit modelleri sıklıkla kullanılmaktadır. Astronomik gelgit bileşenleri Tablo 1'deki bileşenlerle ifade edilir (Yüksel & Esin, 2016; Yüksel vd., 2018).

Tablo 1: Astronomik gelgit bileşenleri

Kaynak	Periyot (saat)
Main lunar, yarı-günlük	12.42
Main solar, yarı-günlük	12.00
Lunar elliptic, yarı-günlük	12.66
Lunar-solar, yarı-günlük	11.97
Lunar solar, günlük	23.93
Main lunar, günlük	25.82
Main solar, günlük	24.07
Lunar fortnightly (15 günlük), uzun periyotlu	327.86
Lunar monthly (aylık), uzun periyotlu	661.30

Bu etkilerin dışında gelgit hareketleri, meteorolojik etkiler, sığ sularda meydana gelen taban sürtünmeleri gibi birçok farklı faktörle de ilişkilidir. Zaman serilerinde gelgit sinyalleri birçok farklı genlik ve frekansa sahip harmonik sinyallerin toplamı şeklinde ifade edilmektedir. Bu sinyaller yarı-günlük, günlük veya daha uzun periyotlarda olabilir. Bir harmonik model;

$$h_i = Z_o + at_i + \sum_{j=1}^M A_j \cos(\omega_j t_i - \theta_j) \quad (9)$$

şeklinde ifade edilir (Teferle, 2003). Gelgiti oluşturan her bir bileşene ait genlik ve faz değerleri klasik harmonik analizlerle belirlenir. Bir sinyale ait genlik ve faz açıları (9) eşitliğinde verilen formül genişletilerek elde edilir (Sezen, 2006; Teferle, 2003).

$$A_j \cos(\omega_j t_i - \theta_j) = A_j [\cos(\omega_j t_i) \cos(\theta_j) + \sin(\omega_j t_i) \sin(\theta_j)] \quad (10a)$$

$$A_j \cos(\theta_j) = B_j \quad (10b)$$

$$A_j \sin(\theta_j) = C_j \quad (10c)$$

$$\omega_j = 2\pi f_j \quad (10d)$$

kabulleri yapılarak harmonik analiz denkleminde yerine konulursa;

$$h_i = Z_0 + a(t_i - t_0) + \sum_{j=1}^M B_j \cos(2\pi f_j(t_i - t_0)) + C_j \sin(2\pi f_j(t_i - t_0)) \quad (11)$$

denklemini elde edilir. Bu denkleme göre t_0 başlangıç epogunda **A** katsayılar matrisi ve **x** bilinmeyenler vektörü aşağıdaki gibi yazılır. Her periyodik etkiye karşılık **A** matrisine 2 sütun eklenir (Sezen, 2006; Teferle, 2003).

$$\mathbf{A} = \begin{bmatrix} 1 & t_1 - t_0 & \sin(2\pi f_1(t_1 - t_0)) & \cos(2\pi f_1(t_1 - t_0)) \\ 1 & t_2 - t_0 & \sin(2\pi f_1(t_2 - t_0)) & \cos(2\pi f_1(t_2 - t_0)) \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 1 & t_{N-1} - t_0 & \sin(2\pi f_1(t_{N-1} - t_0)) & \cos(2\pi f_1(t_{N-1} - t_0)) \\ 1 & t_N - t_0 & \sin(2\pi f_1(t_N - t_0)) & \cos(2\pi f_1(t_N - t_0)) \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} Z_0 \\ a \\ B_1 \\ C_1 \end{bmatrix} \quad (12)$$

EKK yöntemine göre yapılan dengeleme işlemi sonucunda her bir periyodik bileşen için B ve C katsayıları ile genlik ve faz açısı;

$$A_j = \sqrt{B_j^2 + C_j^2} \quad \text{ve} \quad \theta_j = \arctan \frac{C_j}{B_j} \quad (13)$$

formüllerini ile elde edilir. Burada;

- h_i : t_i zamanındaki ortalama deniz seviyesi değeri
- t_i : başlangıç epogundan itibaren geçen gün sayısı
- a : doğrusal trend
- Z_0 : başlangıç epogundaki ortalama deniz seviyesi
- M : modelde kullanılan anlamlı gelgit bileşenlerinin toplam sayısı
- N : aylık toplam ortalama değer sayısı
- A_j, ω_j, θ_j : j. bileşenin genliği, frekansı ve faz açısıdır.

Klasik harmonik analizle ilgili ayrıntılı bilgiler Pugh (1996) çalışmasından elde edilebilir. Bu çalışma kapsamında klasik harmonik analiz yöntemi ile mareograf istasyonlarında kaydedilen deniz seviyesi zaman serilerinde mevsimsel, dönemsel ve gelgit bileşenleri Pawlowicz, Beardsley ve Lentz (2002) tarafından geliştirilen, MATLAB tabanlı T_TIDE yazılımı kullanılarak elde edilmiştir. Bu yazılım ile gelgitlere sebep olan en temel 64 farklı bileşene ait sinyallerin faz, genlik ve frekans değerleri hesaplanabilmektedir. Bu nedenle oşinografi ile ilgili çalışmalarda özellikle genlik bileşenlerinin modellenmesi için T_TIDE yazılımı sıklıkla kullanılmaktadır (Arns vd., 2020; Costa vd., 2020; Juva, Flögel, Karstensen, Linke, & Dullo, 2020). Bu çalışma kapsamında harmonik analiz yöntemi ile belirlenen gelgit bileşenleri, mevsimsel bileşenler ve dönemsel bileşenler ‘‘Harmonik Bileşenler’’ olarak adlandırılmıştır.

3.3 Deniz Seviyesi Tahmini İçin Makine Öğrenimi Algoritmaları

Bu çalışmada Ege Denizi'ndeki anlık deniz seviyesi 3 farklı makine öğrenimi algoritması kullanılarak tahmin edilmiştir. Bu yöntemler regresyon temelli algoritmalarıdır. Regresyon problemlerinde amaç bağımsız değişkenlerden bağımlı değişkene giden yolun bulunmasıdır. Bu yol sadece bir korelasyon katsayısı ile ifade edilemeyecek kadar kompleks bir yapıda olabilir. Bu nedenle regresyon problemleri için birçok farklı algoritma geliştirilmiştir. Algoritmaların başarısı kullanılan veri yapısına ve değişkenler arasındaki ilişkinin karmaşıklığına göre değişmektedir. İlgili çalışma kapsamında deniz seviyesi tahmini için CDR, DVR, ROR algoritmaları kullanılmıştır.

3.3.1 Çoklu Doğrusal Regresyon

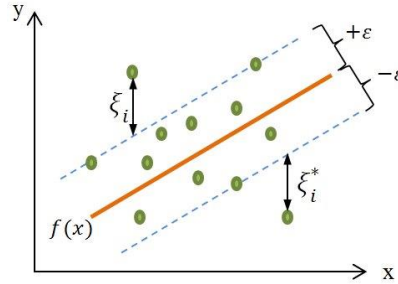
ÇDR algoritması temelde birçok bağımsız değişken kullanılarak bir bağımlı değişkeni bulacak çok boyutlu doğru denkleminin elde edilmesi işlemidir. ÇDR algoritması genellikle bağımlı ve bağımsız değişkenler arasında doğrusal bir ilişki varsa doğru sonuçlar verebilmektedir. ÇDR ile değişkenler arasındaki ilişkiyi gösteren doğru denklemi;

$$y = a + b_1 * x_1 + b_2 * x_2 + \dots + b_n * x_n \quad (14)$$

şeklinde yazılır. Burada y bağımlı değişkeni b_n ise n . bağımsız değişkeni ifade etmektedir.

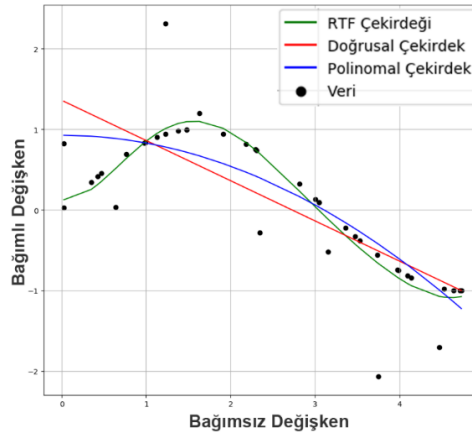
3.3.2 Destek Vektör Regresyonu

DVR algoritması ise DVM algoritmasına dayalı sınıflandırma algoritmasının tahmin uygulamalarında kullanılmak üzere geliştirilmiş özel bir formudur. Günümüzde birçok sınıflandırma ve regresyon problemlerinde kullanılmakta olan DVM algoritması ilk olarak Cortes ve Vapnik (1995) tarafından literatüre kazandırılmıştır. Regresyon problemlerinde kullanılan DVR algoritması ise ilk olarak Drucker, Burges, Kaufman, Smola ve Vapnik (1997) çalışması ile sunulmuştur. DVR algoritmasında amaç fonksiyonu, bir marjın aralığına (ϵ) maksimum sayıda noktayı en küçük hata ile alabilecek doğru veya eğri fonksiyonunun belirlenmesidir (Müller vd., 1997) (Şekil 3).



Şekil 3: DVR algoritması (Chanklan vd., 2018)

DVR algoritmasının önemli bir parametresi de çekirdek (kernel) fonksiyonudur. Çekirdek fonksiyonu ile tahmin edilecek regresyon çizgisinin şekli belirlenebilmektedir. Çekirdek fonksiyonunun bir doğru denklemi seçilmesi halinde değişkenler arasındaki doğrusal ilişkiler DVR ile belirlenir. Çekirdek fonksiyonu polinomal bir eğri fonksiyonu olabileceği gibi DVR uygulamalarında sıklıkla kullanılan RTF de olabilir (Şekil 4).



Şekil 4: DVR algoritmasında çekirdek fonksiyonları

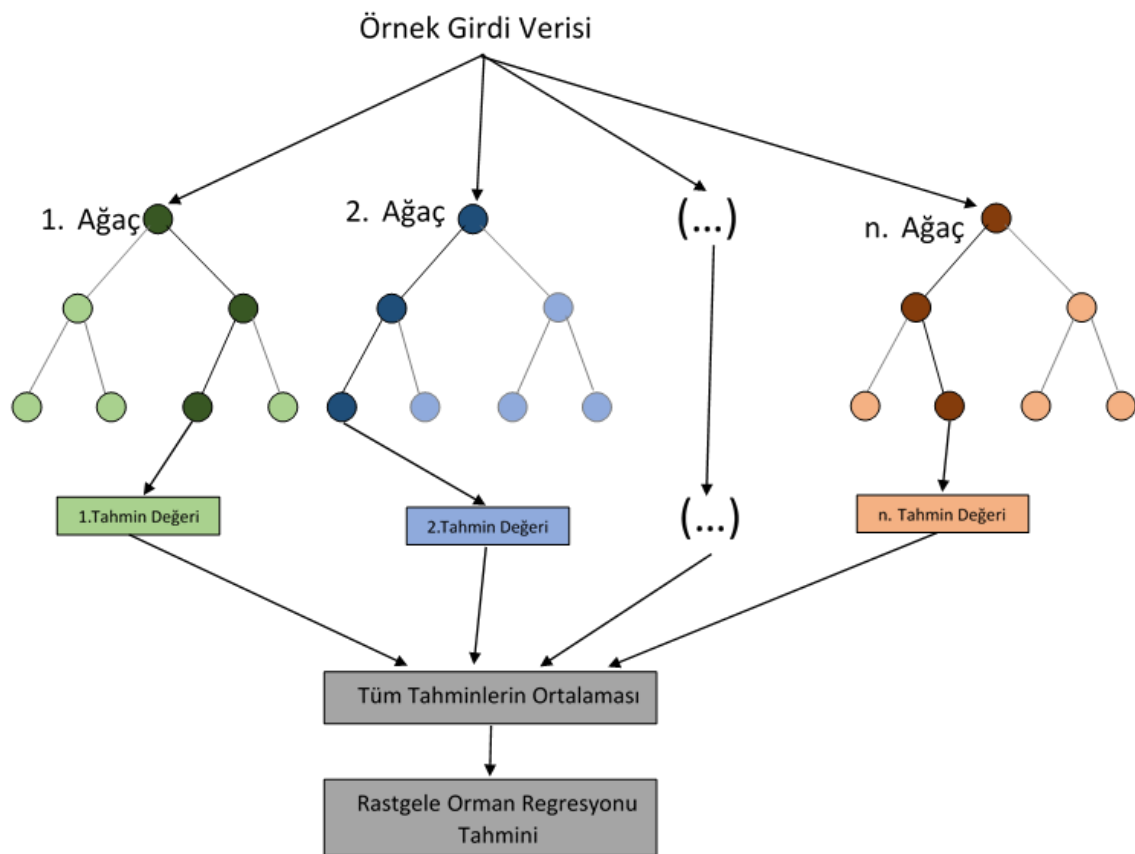
3.3.3 Rastgele Orman Regresyonu

ROR algoritması, tahmin uygulamalarında bağımlı değişkeni tahmin etmek için birçok karar ağacı yapısının kullanıldığı bir regresyon yöntemidir. Karar ağaçları ise düğümlerinde karar değişkenlerini, yapraklarında tahmin değerlerini bulunduran bir yapıya sahiptir. Karar ağaçlarında amaç, hata fonksiyonunun değerini minimuma indirmektir. Hata fonksiyonu bir hata metriği yardımıyla ölçülebilir. Bu çalışmada RO'da bulunan karar ağaçlarının hata fonksiyonları için karesel hataların toplamı (Sum of Squared Errors, SSE) metriği kullanılmıştır. \bar{X} , bağımsız x değişkenlerinin ortalamasını göstermek üzere SSE ve Hata Fonksiyonu (HF);

$$SSE = \sum_{i=1}^N (x_i - \bar{X})^2 \quad (15a)$$

$$HF = \frac{M_{sol}}{M} MSE_{sol} + \frac{M_{sag}}{M} MSE_{sag} \quad (15b)$$

formülleri ile hesaplanır. Burada M örnek sayısını göstermektedir (Gülaçar, 2018). ROR'da oluşturulan rastgele alt kümeler ile çok sayıda karar ağacının çıktı verilerinin ortalaması alınarak son tahmin değeri belirlenir (Şekil 5).



Şekil 5: ROR algoritma yapısı

3.4 Algoritmaların Performanslarının Karşılaştırılması

Bu çalışmada algoritmaların hem eğitim hem de test veri setlerinde model performanslarını karşılaştırmak amacıyla zaman serisi tahmin çalışmalarında sıklıkla kullanılan r (Pearson korelasyon katsayısı) ve $RMSE$ metrikleri kullanılmıştır. r korelasyon katsayısı model verilerinin gerçek veriler ile uyumluluğunun bir ölçümü iken, $RMSE$ gerçek veriler ile model verileri arasındaki farkı gösteren bir metriktir. r ve $RMSE$ metrikleri;

$$r = \frac{\sum_{i=1}^n (y_i - \bar{y})(y_i^p - \bar{y}^p)}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2 \sum_{i=1}^n (y_i^p - \bar{y}^p)^2}} \quad (16a)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - y_i^p)^2}{n}} \quad (16b)$$

formüllerini kullanılarak hesaplanabilir. Burada y_i ve y_i^p sırasıyla i . zamanda gözlemlenen ve tahmin edilen verileri, n veri sayısını, \bar{y} ve \bar{y}^p ortalama gözlem ve ortalama tahmin değerlerini göstermektedir.

4. Sayısal Uygulama

Bu çalışma kapsamında Ege Denizi'nde bulunan 2 mareograf istasyonundaki anlık deniz seviyesi değişimlerini tahmin etmek amacıyla, meteorolojik değişkenler (hava sıcaklığı, hava basıncı, hava nemi, düşey ve yatay rüzgâr bileşenleri) ve deniz seviyesi temel bileşenleri (doğrusal trend ve harmonik bileşenler) kullanılarak 3 farklı makine öğrenimi algoritması ile tahmin modelleri geliştirilmiş ve uygulanan modellerin performansları karşılaştırılmıştır.

4.1 Deniz Seviyesi Temel Bileşenlerinin Öncül Tahmini

Deniz seviyesi temel bileşenleri olan trend ve harmonik bileşenlerinin öncül tahminleri makine öğrenimi algoritmalarında öncül bilgi (öznitelik) olarak kullanılmıştır. Bodrum ve Menteş istasyonları zaman serilerinde trend bileşeninin öncül tahmini için uygulanan basit doğrusal regresyon modelinden elde edilen a_1 eğim katsayısı için parametre anlamlılık testi uygulanmıştır. a_1 eğim katsayısı için

$$t_{a_1} = \frac{|a_1|}{m_{a_1}} \quad (17)$$

olarak elde edilen test değerleri $\alpha = 0.05$ yanılma olasılığı ve $f = n - 2$ serbestlik derecesi ile t dağılım tablosundan alınan tablo değeri 1.96 değerleriyle karşılaştırılmıştır (Tablo 2-3).

Tablo 2: Bodrum istasyonu doğrusal trend bileşeni katsayı anlamlılık testi

a_1 katsayısı	Test değeri	T-Tablo değeri	Sonuç
0.1234	38.3583	1.96	a_1 anlamlı

Tablo 3: Menteş istasyonu doğrusal trend bileşeni katsayı anlamlılık testi

a_1 katsayısı	Test değeri	T-Tablo değeri	Sonuç
0.1461	54.1474	1.96	a_1 anlamlı

Basit doğrusal regresyon modeli ile trend bileşeninin öncül tahmini yapıp trend etkisi giderilmiş deniz seviyesi zaman serileri elde edilmiştir. Daha sonra, harmonik bileşenlerin öncül tahmini için trend etkisi giderilmiş zaman serilerine harmonik analiz uygulanmıştır. İstasyonlardaki bu harmonik bileşenlerin çıkarılması için zaman serisinde bulunabilecek uzun dönemli, yıllık, yarı yıllık, aylık, 15 günlük, günlük ve yarı-günlük ve daha kısa süreli sinyaller incelenmiştir. Yapılan harmonik analiz sonucunda istasyonlardaki saatlik deniz seviyesi zaman serilerinden elde edilen sinyallerin genlik, faz ve frekans değerleri Tablo 4 ve Tablo 5'teki gibidir.

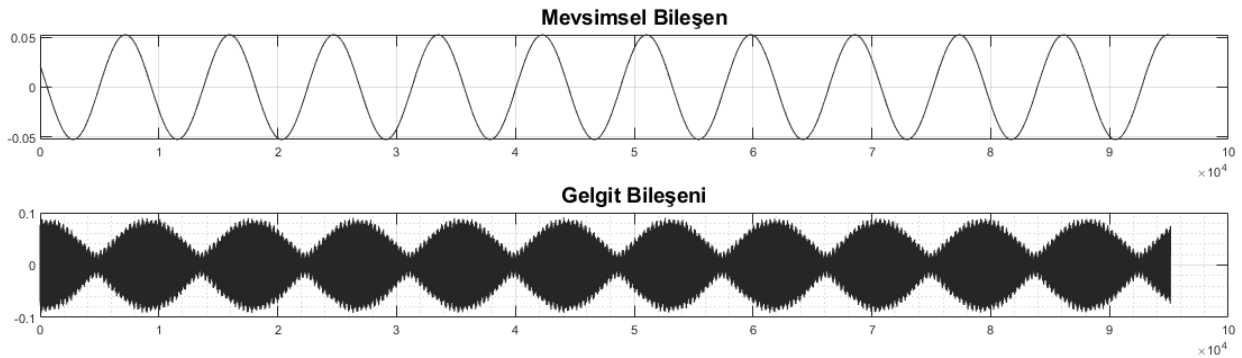
Tablo 4: Menteş istasyonunda anlamlı sinyallerin frekans, genlik ve faz değerleri

Sinyal adı	Frekans	Genlik	Genlik duyarlılığı	Faz	Faz duyarlılığı
SA	0.000114	0.2238	0.090	147.03	25.37
P1	0.041553	0.0227	0.011	347.13	29.63
PSII	0.041895	0.0213	0.013	105.85	30.58
M2	0.080511	0.0336	0.021	105.96	37.65
H2	0.080626	0.1434	0.022	96.71	8.82
T2	0.083219	0.0348	0.020	286.82	36.62
S2	0.083333	0.0336	0.023	300.17	36.40
R2	0.083447	0.0391	0.023	337.69	29.18
2MN6	0.240022	0.0051	0.003	60.33	37.08

Tablo 5: Bodrum istasyonunda anlamlı sinyallerin frekans, genlik ve faz değerleri

Sinyal adı	Frekans	Genlik	Genlik duyarlılığı	Faz	Faz duyarlılığı
SA	0.000114	0.0524	0.009	109.53	9.72
O1	0.038731	0.0027	0.001	162.68	32.98
PSII	0.041895	0.0034	0.001	230.73	22.32
H2	0.080626	0.0060	0.002	352.52	19.84
R2	0.083447	0.0061	0.002	179.95	22.85
M4	0.161023	0.0001	0.000	203.53	44.66

Burada SA bir yıllık periyoda sahip mevsimsel bileşeni, P1 - PSII - O1 günlük periyoda sahip günlük gelgit bileşenini, M2 - H2 - T2 - S2 - R2 sinyalleri yarım günlük gelgit bileşenlerine ait sinyalleri, M4 6 saatlik gelgit bileşenini ve 2MN6 sinyali ise 4 saatlik gelgit bileşenini göstermektedir. Bodrum istasyonu için bulunan mevsimsel bileşen ve anlamlı gelgit sinyallerinin toplamı olan gelgit bileşeni Şekil 6'da verilmiştir.

**Şekil 6:** Bodrum istasyonu için mevsimsel bileşen ve gelgit bileşeni

Öncül bilgiler;

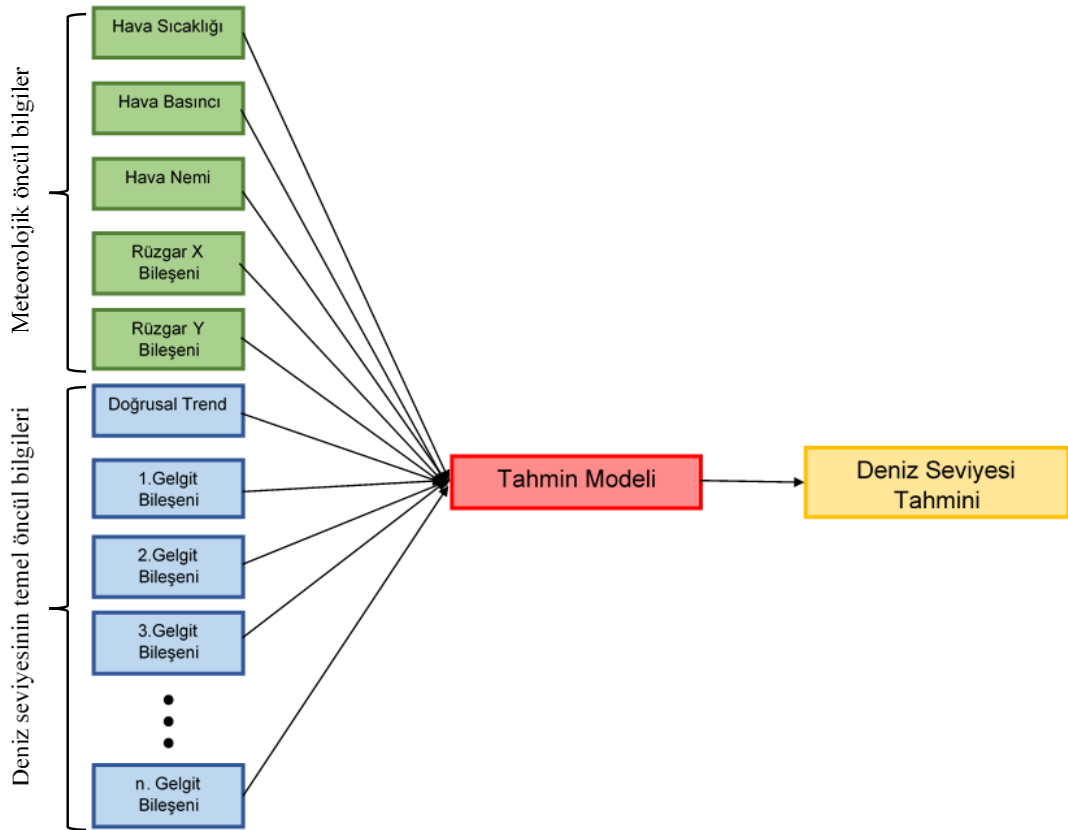
- Tablo 2 ve Tablo 3'te verilen doğrusal trend bileşenleri,
- Tablo 4, Tablo 5'te verilen ve Şekil 6'da gösterilen harmonik bileşenleri,
- İstasyonlarda kaydedilen meteorolojik bileşenleri (hava sıcaklığı, hava basıncı, hava nemi, x yönündeki rüzgar bileşeni, y yönündeki rüzgar bileşeni)

olarak belirlenmiştir. Bu şekilde trend ve harmonik bileşenlerden oluşan temel öncül bileşenler ve meteorolojik bileşenlerle,

Bodrum istasyonu için 12 adet, Menteş istasyonu için ise 15 adet farklı öncül bilgi deđiřkeni kullanılmıřtır. İstasyonlardaki öncül bilgi (öznitelik) çıkarımı tamamlandıktan sonra makine öğrenimi algoritmalarının eğitimi ve test edilmesi aşamasına geçilmiştir.

4.2 Makine Öğrenimi Algoritmalarının Eğitilmesi ve Test Edilmesi

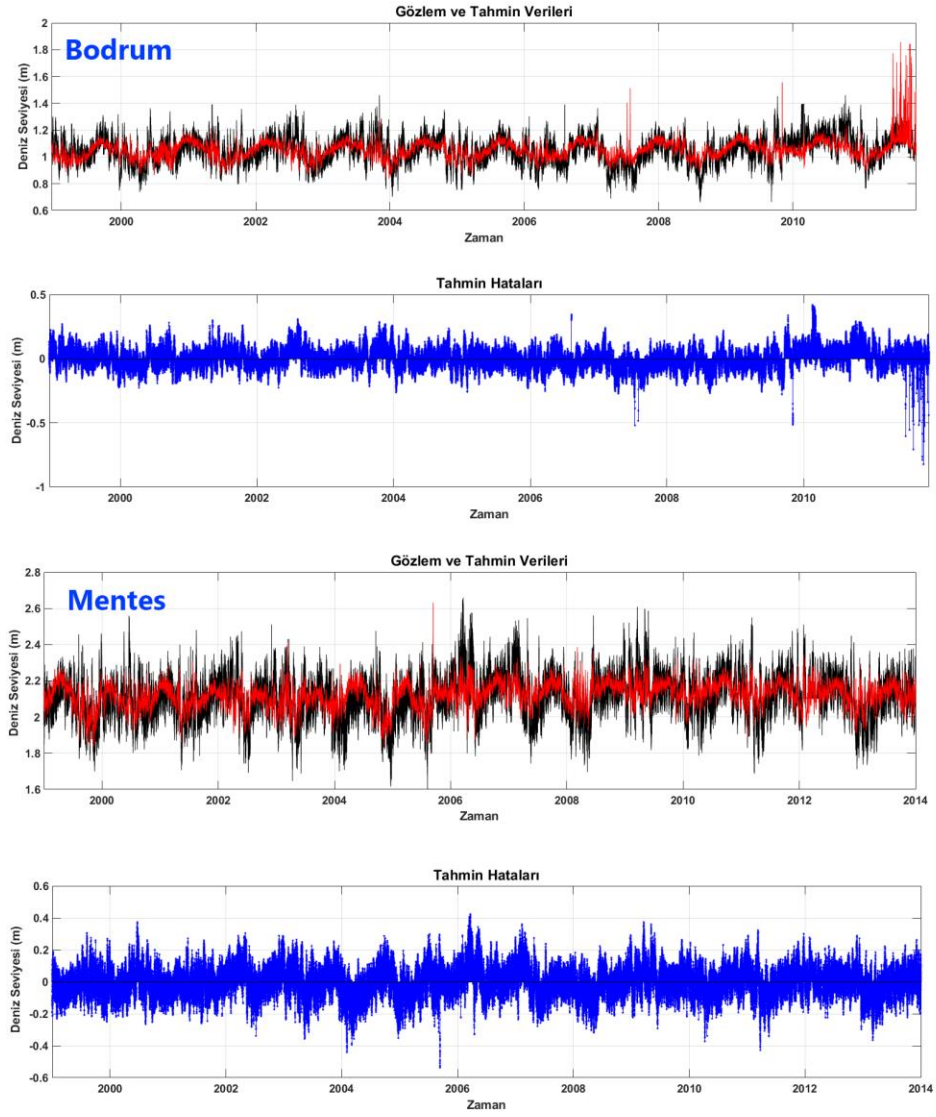
İstasyon bölgelerine ait deniz seviyesi temel bileşenlerine ait öncül bilgilerin elde edilmesinin ardından ÇDR, DVR ve ROR algoritmaları eğitim verileri kullanılarak eğitilmiştir. Algoritmaların eğitimi için kullanılan öncül bilgiler Şekil 7’de verilmiştir. Öncül bilgi verilerinin %70’i makine öğrenimi algoritmalarının eğitimi için eğitim verisi olarak, %30’u ise tahmin model performanslarını değerlendirmek üzere test verisi olarak rastgele seçilmiştir. Makine öğrenimi algoritmalarının eğitimi ve test edilmesi için Pedregosa vd. (2011) tarafından geliştirilen Python ortamındaki SciKit-Learn kütüphanesi kullanılmıştır.



Şekil 7: Eğitim için kullanılan öncül bilgiler

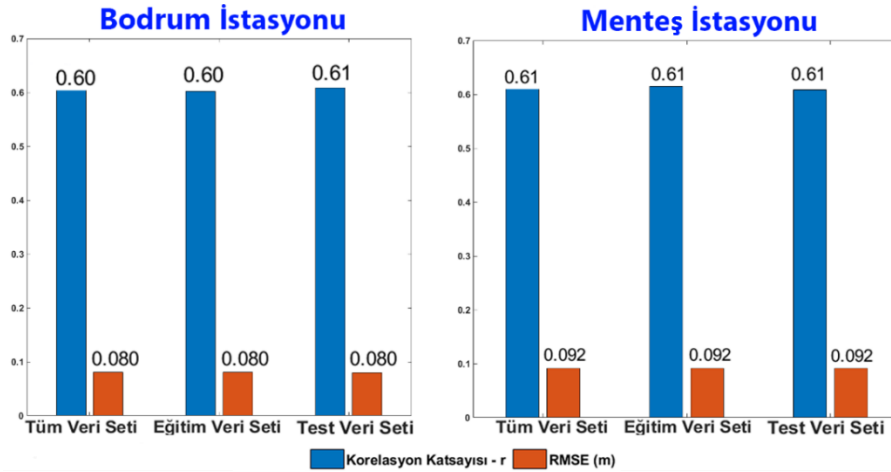
4.2.1 Çoklu Doğrusal Regresyon Algoritması Eğitim ve Test Performansı

ÇDR algoritması eğitim ve test verileri kullanılarak eğitilmiştir. Eğitilmiş model kullanılarak eğitim, test ve tüm dönemler için deniz seviyesi tahminleri yapılmıştır. Yapılan tahmin değerleri gerçek değerlerle karşılaştırılarak, modellerin eğitim, test ve tüm veri seti için tahminlerinin doğruluk performansları değerlendirilmiştir (Şekil 8).



Şekil 8: ÇDR algoritması tahminleri

ÇDR algoritmasının eğitim ve test veri setindeki deniz seviyesi tahmin performansları r Pearson korelasyon katsayısı ve RMSE metrikleri kullanılarak ölçülmüştür. Elde edilen tahmin performansları Şekil 9'da verilmiştir.

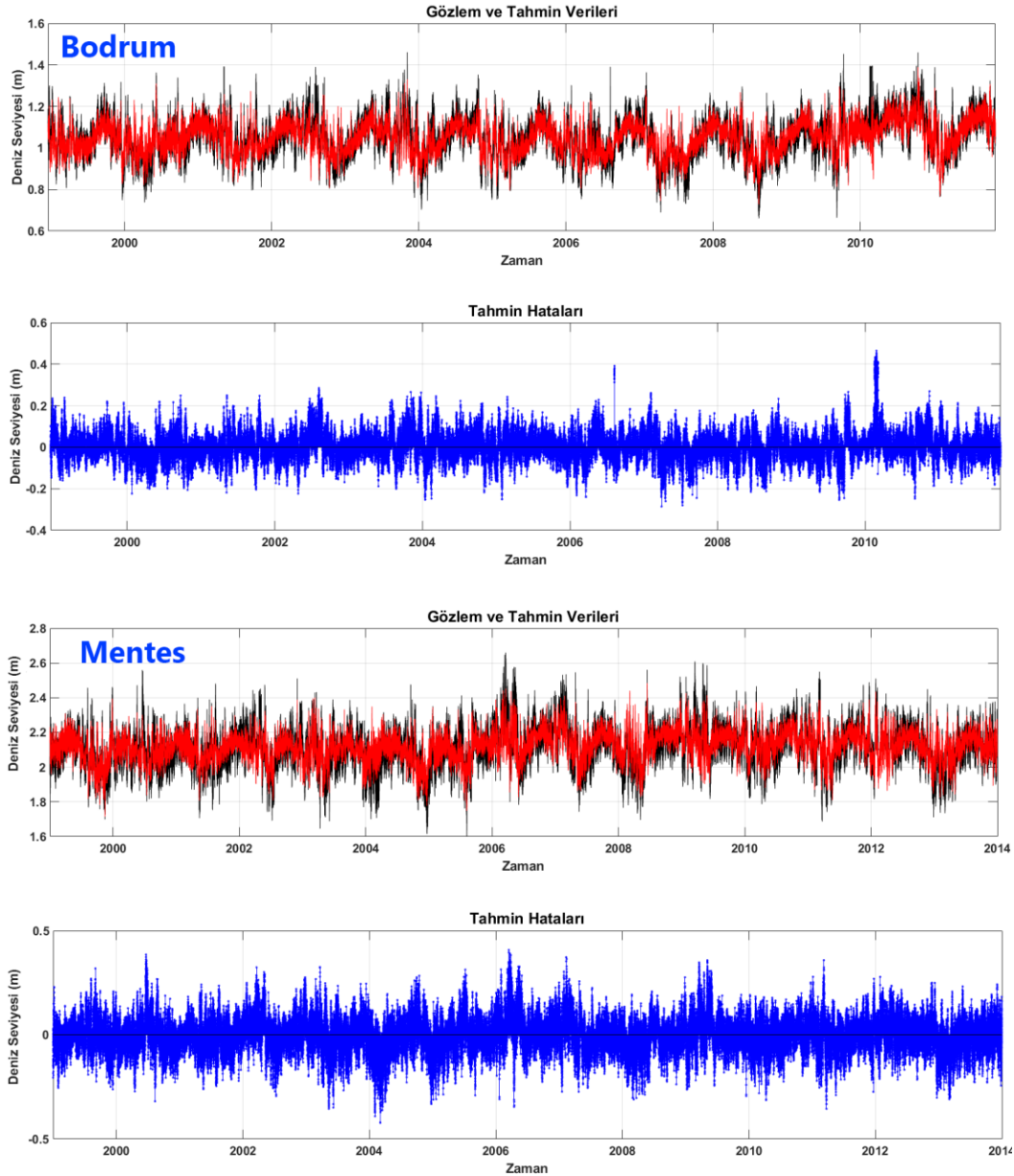


Şekil 9: ÇDR algoritması deniz seviyesi tahmin performansı

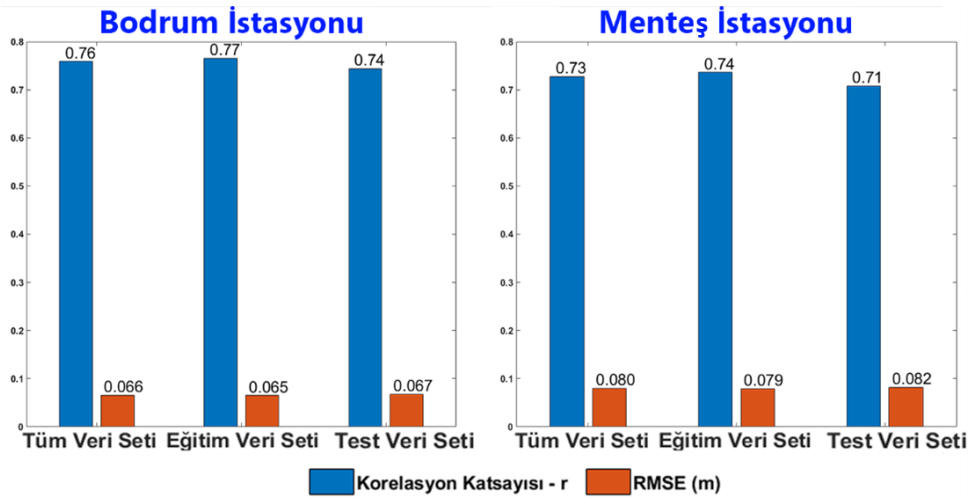
Şekil 8 ve Şekil 9’da görüldüğü gibi ÇDR algoritması hem eğitim hem test döneminde her iki istasyondaki deniz seviyesi tahminlerinde benzer sonuçlar vermektedir. Özellikle Bodrum istasyonunun son dönem tahminleri oldukça yüksek olup, gerçek veriden uzaktır. Bu durum karmaşık bağımlılık yapılarına sahip veri setlerinde doğrusal fonksiyonlarının tahminlerinin ne kadar yetersiz olduğunu göstermesi bakımından önemlidir.

4.2.2 Destek Vektör Regresyonu Algoritması Eğitim ve Test Performansı

RTF çekirdeği kullanılan DVR algoritmasının istasyonlardaki deniz seviyesi tahminleri Şekil 10’da verilmiştir. Ayrıca DVR algoritmasının eğitim ve test veri setindeki deniz seviyesi tahmin performansları r ve RMSE metrikleri cinsinden Şekil 11’de verilmiştir.



Şekil 10: DVR algoritması tahminleri (tüm veri seti için)

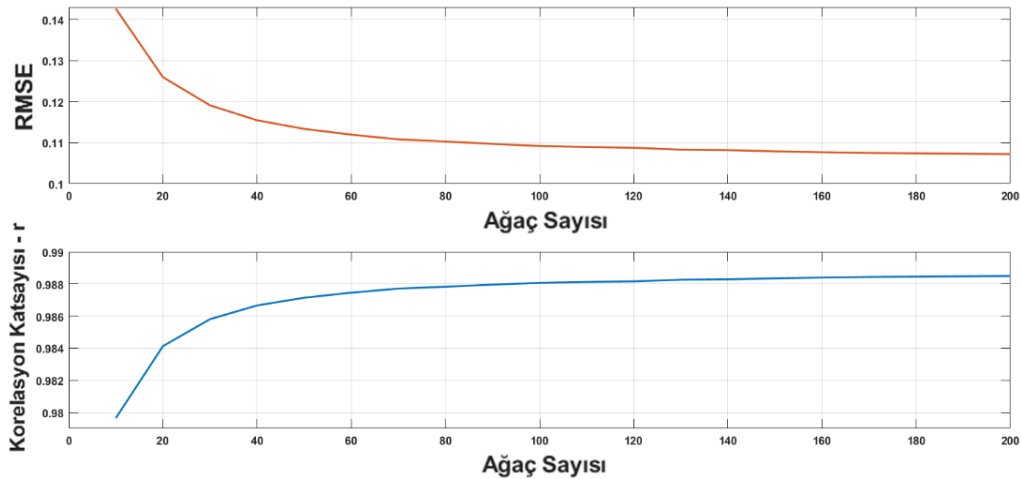


Şekil 11: DVR algoritması deniz seviyesi tahmin performansı

Şekil 10 ve Şekil 11'de görüldüğü gibi DVR algoritmasının deniz seviyesi tahmin performansı hem daha yüksek korelasyon hem de daha düşük RMSE metrikleri ile ÇDR algoritmasına göre daha iyidir.

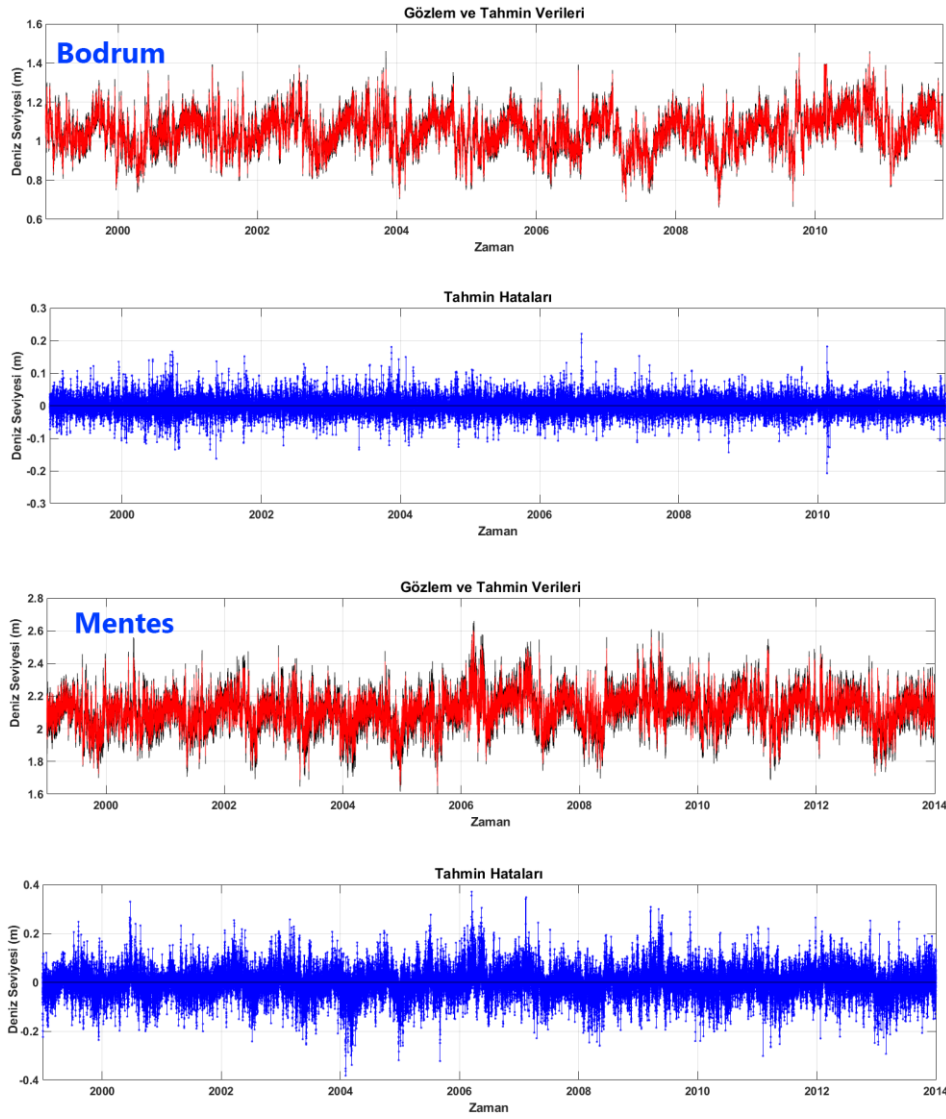
4.2.3 Rastgele Orman Regresyonu Algoritması Eğitim ve Test Performansı

ROR algoritması ile deniz seviyesi tahmininin yüksek doğrulukta yapılabilmesi için karar ağacı ormanındaki ağaç sayısının doğru seçilmesi oldukça önemlidir. Bu çalışmada ROR algoritmasının içereceği ağaç sayısının belirlenmesi için RMSE ve r doğruluk analizlerine dayanan bir yaklaşım kullanılmıştır. Algoritmadaki ağaç sayısının model doğruluğuna etkisi incelenerek kullanılacak ağaç sayısı belirlenmiştir. 10 ağaçtan başlayarak 200 ağaca kadar ROR algoritması ile deniz seviyesi tahminleri yapılmıştır. Bodrum istasyonu için yapılan tahminlere ilişkin model doğrulukları test veri seti için Şekil 12'deki gibidir.

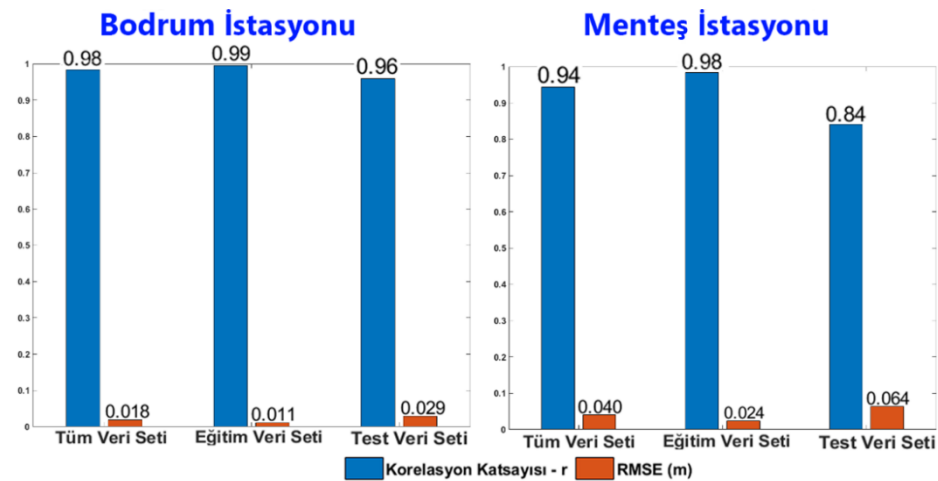


Şekil 12: ROR algoritmasında ağaç sayısının tahmin performansına etkisi

Şekil 12'de görüldüğü gibi ağaç sayısı arttıkça ROR modelinin tahmin performansı yükselmektedir. Ancak yaklaşık 100 ağaçtan sonra tahmin performansının hem korelasyon hem de RMSE metriği açısından fazla bir değişim göstermediği belirlenerek ROR algoritmasında kullanılacak ağaç sayısı 100 olarak seçilmiştir. ROR algoritmasının deniz seviyesine ilişkin tahminleri ve tahmin performansı sırasıyla Şekil 13 ve Şekil 14'te verilmiştir.



Şekil 13: ROR algoritması tahminleri (tüm veri seti için)



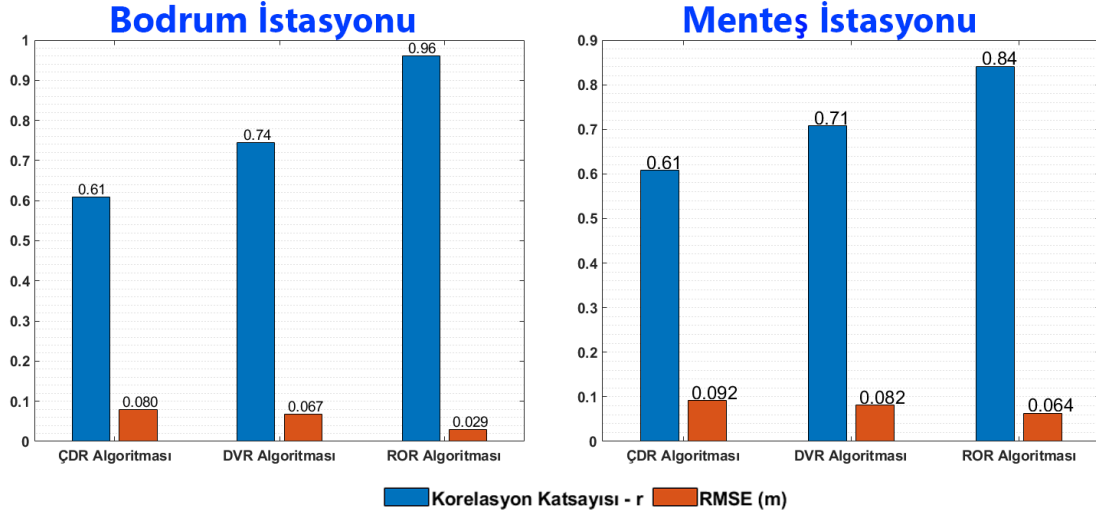
Şekil 14: ROR algoritması deniz seviyesi tahmin performansı

Şekil 13 ve Şekil 14'te görüldüğü gibi ROR algoritmasının deniz seviyesi tahmin performansı hem eğitim hem de test veri seti için ÇDR ve DVR algoritmalarına kıyasla oldukça yüksektir. Bu yüksek doğrulukta tahmin performansının, ROR algoritmasının karmaşık bağımlılık yapılarını kendi ağaç algoritma yapısı sayesinde modelleyebilmesinden kaynaklandığı

değerlendirilmektedir.

4.3 Makine Öğrenimi Algoritmalarının Karşılaştırılması

Makine öğrenimi algoritmalarının eğitim verileri kullanılarak eğitilmesinin ardından test veri seti kullanılarak yapılan tahmin sonuçlarına ilişkin performans analizleri sonuçları Şekil 15'de verilmiştir.



Şekil 15: Makine öğrenimi algoritmalarının karşılaştırılması

Her iki istasyonda da ÇDR ve DVR algoritmalarının test performansları ROR algoritmasına göre düşüktür. Ancak ROR algoritmasının performansı Bodrum istasyonunda daha yüksek olarak bulunmuştur. Bu durumun Mentеш istasyonundaki eksik verilerin fazla olmasından ve daha fazla enterpolasyon verisi içermesinden kaynaklandığı değerlendirilmektedir. Eksik veriler her ne kadar enterpolasyon işlemleri ile tamamlansa da makine öğrenimi tahmin algoritmalarında model performansına etki etmektedirler.

5. Sonuçlar

Uzun dönemlere ait deniz seviyesi verileri ve meteorolojik veriler kullanılarak deniz seviyesinin yüksek doğrulukta kestirimi sonucunda kıyı yapılarının planlanması, küresel ısınmanın etkilerinin tahmin edilmesi ve geleceğe yönelik etkin su politikalarının oluşturulması konularına katkı sağlayacağı değerlendirilmektedir. Bu çalışmada, Ege denizindeki iki mareograf istasyonundan elde edilen anlık deniz seviyelerinin tahmini için makine öğrenimi yöntemleri kullanılarak tahmin modelleri oluşturulmuş ve tahmin modellerinin performansları karşılaştırılmıştır. Makine öğrenimi algoritmalarının anlık deniz seviyesini tahmin etmek için kullanacağı öncül bilgiler istasyonlarda ölçülen meteorolojik veriler ve anlık deniz seviyesi zaman serilerinin doğrusal trend ve harmonik bileşenleridir. Bu kapsamda ÇDR, DVR ve ROR makine öğrenimi algoritmaları kullanılmıştır. Makine öğrenimi algoritmaları eğitim verileri ile eğitilmiş, test verileri kullanılarak deniz seviyesi tahmin performansları korelasyon katsayısı ve RMSE metrikleri ile değerlendirilmiştir. Yapılan analizler sonucunda mareograf istasyonlarında deniz seviyesi ROR algoritması kullanılarak %96 ila %84 korelasyon ve 29 ila 64 mm RMSE değerleri ile oldukça yüksek doğrulukta tahmin edilmiştir. En yüksek tahmin performansı ROR algoritması ile elde edilirken en düşük tahmin performansı ÇDR algoritması ile elde edilmiştir. Çalışma kapsamında elde edilen sonuçlar geçmişte yapılan çalışmalar ile karşılaştırıldığında, her ne kadar aynı veri setleri ve aynı öncül bilgiler kullanılsa da okuyuculara ROR algoritmasının ve kullanılan öncül bilgilerin anlık deniz seviyesini belirlemedeki başarısı hakkında fikir vereceği düşünülmektedir. Bu nedenle geçmiş çalışmalar incelendiğinde, anlık deniz seviyeleri makine öğrenimi yöntemleri

kullanılarak Imani vd. (2018) tarafından 34 mm, Imani vd. (2014a) tarafından 35 mm, Roshni vd. (2019) tarafından 75 mm RMSE hassasiyetle belirlenirken, bu çalışmada anlık deniz seviyeleri Bodrum istasyonunda 29 mm, Menteş istasyonunda ise 64 mm RMSE ile belirlenmiştir.

Ayrıca elde edilen sonuçlar doğrusal tahmin modellerinin deniz seviyesini tahmin etmekte ne kadar yetersiz olduğunu göstermesi bakımından da oldukça önemlidir. Çünkü ÇDR algoritması doğrusal bir tahmin modeli olarak hem korelasyon hem de RMSE metrikleri bakımından doğrusal olmayan DVR ve ROR algoritmalarına göre düşük tahmin performansı göstermiştir. ROR algoritmasının ağaç algoritma yapısı sayesinde meteorolojik veriler, doğrusal trend ve harmonik bileşenler bağımsız değişkenleri kullanılarak deniz seviyesi yüksek doğrulukta kestirilmiştir. Anlık deniz seviyelerinin karmaşık bağımlılık yapılarının modellenmesinde ROR algoritmasının diğer algoritmalara göre daha başarılı olduğu belirlenmiştir.

Teşekkür

Çalışmada kullanılan TUDES verilerini <https://tudes.harita.gov.tr/> linki üzerinden sağlayan Harita Genel Müdürlüğüne teşekkür ederiz.

Yazar Katkısı

Ahmet Yavuzdođan: Literatür taraması, Veri işleme, Analiz ve yorumlama, Yazım. **Emine Tanır Kayıkçı:** Fikir, Tasarım, Denetleme, Makale değerlendirme.

Çıkar Çatışması Beyanı

Yazarlar, bu çalışmada bilinen ilgili herhangi bir finansal veya finansal olmayan çıkar çatışması olmadığını beyan ederler.

Kaynaklar

- Arns, A., Wahl, T., Wolff, C., Vafeidis, A. T., Haigh, I. D., Woodworth, P., Niehüser, S., & Jensen, J. (2020). Non-linear interaction modulates global extreme sea levels, coastal flood exposure, and impacts. *Nature communications*, 11(1), 1-9.
- Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.
- Brundrit, G. B. (1995). Trends of southern African sea level: statistical analysis and interpretation. *South African journal of marine science*, 16(1), 9-17.
- Chanklan, R., Kaoungku, N., Suksut, K., Kerdprasop, K., & Kerdprasop, N. (2018). Runoff Prediction with a combined artificial neural network and support vector regression. *International Journal of Machine Learning and Computing*, 8(1).
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20(3), 273-297.
- Costa, C. G., Leite, J. R. B., Castro, B. M., Blumberg, A. F., Georgas, N., Dottori, M., & Jordi, A. (2020). An operational forecasting system for physical processes in the Santos-Sao Vicente-Bertioga Estuarine System, Southeast Brazil. *Ocean Dynamics*, 70(2), 257-271.
- Drucker, H., Burges, C. J., Kaufman, L., Smola, A., & Vapnik, V. (1997). Support vector regression machines. *Advances in neural information processing systems*, 9, 155-161.
- Ertuğrul, Ö. F., & Tağluk, M. E. (2017). Forecasting local mean sea level by generalized behavioral learning method. *Arabian Journal for Science and Engineering*, 42(8), 3289-3298.
- Fu, Y., Zhou, X., Sun, W., & Tang, Q. (2019). Hybrid model combining empirical mode decomposition, singular spectrum analysis, and least squares for satellite-derived sea-level anomaly prediction. *International journal of remote sensing*, 40(20), 7817-7829.

- Ghorbani, M. A., Khatibi, R., AYTEK, A., Makarynsky, O., & Shiri, J. (2010). Sea water level forecasting using genetic programming and comparing the performance with artificial neural networks. *Computers & Geosciences*, 36(5), 620-627.
- Gülaçar, H. (2018). *Nesnelerin İnterneti Platformları İçin Makine Öğrenmesi Tabanlı Bir Tahmin Modülü*. (Yüksek Lisans Tezi), İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul, Türkiye.
- Güven, A., & Günel, M. (2008). Genetic programming approach for prediction of local scour downstream of hydraulic structures. *Journal of Irrigation and Drainage Engineering*, 134(2), 241-249.
- Huang, M., Peng, G., Zhang, J., & Zhang, S. (2006). Application of artificial neural networks to the prediction of dust storms in Northwest China. *Global and Planetary Change*, 52(1-4), 216-224.
- Imani, M., You, R. J., & Chung-Yen, K. (2013). Accurate Forecasting of the satellite-derived seasonal Caspian sea level anomaly using polynomial interpolation and holt-winters exponential smoothing. *Tao: terrestrial, atmospheric and oceanic sciences*, 24(4), 521.
- Imani, M., You, R. J., & Kuo, C. Y. (2014a). Forecasting Caspian Sea level changes using satellite altimetry data (June 1992–December 2013) based on evolutionary support vector regression algorithms and gene expression programming. *Global and planetary change*, 121, 53-63.
- Imani, M., You, R. J., & Kuo, C. Y. (2014b). Caspian Sea level prediction using satellite altimetry by artificial neural networks. *International journal of environmental science and technology*, 11(4), 1035-1042.
- Imani, M., Kao, H. C., Lan, W. H., & Kuo, C. Y. (2018). Daily sea level prediction at Chiayi coast, Taiwan using extreme learning machine and relevance vector machine. *Global and planetary change*, 161, 211-221.
- Irvine, K. N., & Eberhardt, A. J. (1992). Multiplicative, Seasonal Arima Models for Lake Erie And Lake Ontario Water Levels 1. *JAWRA Journal of the American Water Resources Association*, 28(2), 385-396.
- Juva, K., Flögel, S., Karstensen, J., Linke, P., & Dullo, W. C. (2020). Tidal dynamics control on cold-water coral growth: A high-resolution multivariable study on eastern Atlantic cold-water coral sites. *Frontiers in Marine Science*, 7(132).
- Kim, K. J. (2003). Financial time series forecasting using support vector machines. *Neurocomputing*, 55(1-2), 307-319.
- Kisi, O., Shiri, J., & Nikoofar, B. (2012). Forecasting daily lake levels using artificial intelligence approaches. *Computers & Geosciences*, 41, 169-180.
- Makarynsky, O., Makarynska, D., Kuhn, M., & Featherstone, W. E. (2004). Predicting sea level variations with artificial neural networks at Hillarys Boat Harbour, Western Australia. *Estuarine, Coastal and Shelf Science*, 61(2), 351-360.
- Meshkani, M. R., & Meshkani, A. (1997). Stochastic modelling of the Caspian Sea level fluctuations. *Theoretical and applied climatology*, 58(3), 189-195.
- More, A., & Deo, M. C. (2003). Forecasting wind with neural networks. *Marine structures*, 16(1), 35-49.
- Müller, K. R., Smola, A. J., Rätsch, G., Schölkopf, B., Kohlmorgen, J., & Vapnik, V. (1997). Predicting time series with support vector machines. *International Conference on Artificial Neural Networks* (s. 999-1004). Berlin, Heidelberg: Springer.
- Pawłowicz, R., Beardsley, B., & Lentz, S. (2002). Classical tidal harmonic analysis including error estimates in MATLAB using T_TIDE. *Computers & Geosciences*, 28(8), 929-937.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perot, M., & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *the Journal of machine Learning research*, 12, 2825-2830.
- Pugh, D. T. (1996). Tides, surges and mean sea-level (reprinted with corrections). Chichester, UK: John Wiley & Sons Ltd.
- Rajasekaran, S., Gayathri, S., & Lee, T. L. (2008). Support vector regression methodology for storm surge predictions. *Ocean Engineering*, 35(16), 1578-1587.
- Roshni, T., Samui, P., & Drisya, J. (2019). Operational use of machine learning models for sea-level modeling. *Indian Journal of Geo Marine Sciences*, 48(9), 1427-1434.
- Sezen, E. (2006). *Antalya-I (1935-1977) ve Antalya-II (1985-2005) Mareograf İstasyonlarında Deniz Seviyesi Değişimlerinin Araştırılması*. (Yüksek Lisans Tezi), Afyon Kocatepe Üniversitesi, Afyonkarahisar, Türkiye.
- Srivastava, P. K., Islam, T., Singh, S. K., Petropoulos, G. P., Gupta, M., & Dai, Q. (2016). Forecasting Arabian Sea level rise using exponential smoothing state space models and ARIMA from TOPEX and Jason satellite radar altimeter data. *Meteorological applications*, 23(4), 633-639.
- Şen, Z., Kadioğlu, M., & Batur, E. (2000). Stochastic modeling of the Van Lake monthly level fluctuations in Turkey. *Theoretical and*

applied climatology, 65(1), 99-110.

- Talebzadeh, M., & Moridnejad, A. (2011). Uncertainty analysis for the forecast of lake level fluctuations using ensembles of ANN and ANFIS models. *Expert Systems with applications*, 38(4), 4126-4135.
- Teferle, F. N. (2003). *Strategies for Long Term Monitoring of Tide Gauges Using GPS*. (Doktora Tezi). University of Nottingham, Nottingham, İngiltere.
- Vaziri, M. (1997). Predicting Caspian Sea surface water level by ANN and ARIMA models. *Journal of waterway, port, coastal, and ocean engineering*, 123(4), 158-162.
- Wu, C. L., & Chau, K. W. (2010). Data-driven models for monthly streamflow time series prediction. *Engineering Applications of Artificial Intelligence*, 23(8), 1350-1367.
- Yu, P. S., Chen, S. T., & Chang, I. F. (2006). Support vector regression for real-time flood stage forecasting. *Journal of Hydrology*, 328(3-4), 704-716.
- Yüksel, Y., & Esin, Ç. (2016). *Kıyı Mühendisliđi*. İstanbul, Türkiye: BETA Yayınevi.
- Yüksel, Y., Öztürk, M., Şahin, C., Halat, O., Dođan, U., Yüksel, Z. T., & Karova, C. (2018). Türkiye Denizlerinde Su Seviyesi Deđişimi. *9. Kıyı Mühendisliđi Sempozyumu*, Adana, Türkiye.
- Zhao, J., Fan, Y., & Mu, Y. (2019). Sea level prediction in the Yellow Sea from satellite altimetry with a combined least squares-neural network approach. *Marine geodesy*, 42(4), 344-366.
- URL-1: <https://tudes.harita.gov.tr/>, (Erişim Tarihi : 06 Haziran 2019).