

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

**AÇIKLANABİLİR YAPAY ZEKA YÖNTEMLERİ İLE TOPRAK TUZLULUĞU
HARİTALARININ ÜRETİLMESİ**

YÜKSEK LİSANS TEZİ

Samet AKSOY

Geomatik Mühendisliği Anabilim Dalı

Geomatik Mühendisliği Programı

OCAK 2023

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

**AÇIKLANABİLİR YAPAY ZEKA YÖNTEMLERİ İLE TOPRAK TUZLULUĞU
HARİTALARININ ÜRETİLMESİ**

YÜKSEK LİSANS TEZİ

**Samet AKSOY
(501191640)**

Geomatik Mühendisliği Anabilim Dalı

Geomatik Mühendisliği Programı

Tez Danışmanı: Prof. Dr. Elif SERTEL

OCAK 2023

ISTANBUL TECHNICAL UNIVERSITY★ GRADUATE SCHOOL

**SOIL SALINITY MAPPING USING EXPLAINABLE ARTIFICIAL
INTELLIGENCE METHODS**



M.Sc. THESIS

**Samet AKSOY
(501191640)**

Department of Geomatics Engineering

Geomatics Engineering Programme

Thesis Advisor: Prof. Dr. Elif SERTEL

JANUARY 2023

İTÜ, Lisansüstü Eğitim Enstitüsü'nün 501191640 numaralı Yüksek Lisans Öğrencisi Samet AKSOY, ilgili yönetmeliklerin belirlediği gerekli tüm şartları yerine getirdikten sonra hazırladığı “AÇIKLANABİLİR YAPAY ZEKA YÖNTEMLERİ İLE TOPRAK TUZLULUĞU HARİTALARININ ÜRETİLMESİ” başlıklı tezini aşağıda imzaları olan jüri önünde başarı ile sunmuştur.

Tez Danışmanı : **Prof. Dr. Elif SERTEL**
İstanbul Teknik Üniversitesi

Jüri Üyeleri : **Prof. Dr. Ayşe Gül TANIK**
İstanbul Teknik Üniversitesi

Doç. Dr. Beyza USTAOĞLU
Sakarya Üniversitesi

Teslim Tarihi : 30 Aralık 2022
Savunma Tarihi : 16 Ocak 2023





Canım aileme,



ÖNSÖZ

Öncelikle her konuda bana destek olan, tüm akademik çalışmalarımnda yanımda olan danışman hocam Prof. Dr. Elif SERTEL'e sayısız fikirleri, sınırsız teşvikleri, geniş vizyonu ve motivasyonu için, özellikle de bu tezi tamamlamadaki destekleri için en içten şekilde sonsuz teşekkürlerimi sunarım.

Ayrıca Prof. Dr. Ayşegül TANIK'a çalışmalarımız boyunca olan önerileri ve bilgeliği için teşekkür ederim. Sayın Dr. Nikou HAMZEHPUR'a da sağladığı veriler ve yorumları için teşekkür ederim.

Benim bu günlere gelmemi sağlayan, eğitim hayatım boyunca her türlü fedakarlığı gösteren ve hep arkamda duran aileme de sonsuz teşekkürlerimi sunarım.

Son olarak bu tezi hazırlama konusunda manevi desteğini esirgemeyen başta Yük. Müh. Esra ŞENGÜN olmak üzere tüm arkadaşlarıma teşekkür ederim.

Ocak 2023

Samet Aksoy
(Geomatik Mühendisi)



İÇİNDEKİLER

Sayfa

ÖNSÖZ.....	ix
İÇİNDEKİLER	xi
KISALTMALAR	xiii
SEMBOLLER	xv
ÇİZELGE LİSTESİ.....	xvii
ŞEKİL LİSTESİ.....	xix
ÖZET.....	xxi
SUMMARY	xxv
1. GİRİŞ	1
1.1 Tezin Amacı	4
1.2 Literatür Araştırması	5
1.2.1 Toprak tuzluluğunun uzaktan algılama yöntemleri ile tespiti.....	5
1.2.2 Makine öğrenme algoritmalarıyla toprak tuzluluğunun tespiti.....	7
1.2.3 Açıklanabilir yapay zeka yöntemleri ile modellerin açıklanması	9
2. ÇALIŞMA BÖLGELERİ VE KULLANILAN VERİLER.....	13
2.1 Çalışma Bölgeleri	13
2.1.1 Batı Urmiye playası	13
2.1.2 Bonab bölgesi.....	14
2.2 Kullanılan Veriler.....	15
2.2.1 Uydu görüntüleri	15
2.2.2 Toprak örnekleri.....	15
3. YÖNTEM.....	19
3.1 Google Earth Engine	19
3.2 Makine Öğrenmesi için Özellikler	20
3.3 Aşırı Örneklemeye	21
3.4 Makine Öğrenmesi Yöntemleri.....	25
3.4.1 Rastgele orman (random forest).....	25
3.4.2 XGBoost (aşırı gradyan arttırma, extreme gradient boosting).....	25
3.5 Açıklanabilir Yapay Zeka	26
3.5.1 SHAP yöntemi	27
3.5.1.1 SHAP özet grafiği	27
3.5.1.2 SHAP özellik önem grafiği	27
3.5.1.3 SHAP bağımlılık grafiği	28
3.5.1.4 SHAP kuvvet grafiği	29
3.6 Doğruluk Analizi.....	30
4. BULGULAR VE TARTIŞMA.....	31
5. SONUÇLAR VE ÖNERİLER	85
KAYNAKLAR	87
ÖZGEÇMİŞ.....	93



KISALTMALAR

ANN	: Yapay Sinir Ağları
dS/m	: Desisiemens / metre
EC	: Elektriksel İletkenlik
FAO	: Birleşmiş Milletler Gıda ve Tarım Örgütü
RF	: Rastgele Orman
GEE	: Google Earth Engine
LIME	: Yerel Yorumlanabilir Model-agnostik Açıklamalar
RMSE	: Karesel Ortalama Hata
ROS	: Rastgele Aşırı Örnekleme
SHAP	: Shapley katkı açıklamaları
SI	: Tuzluluk İndisi
SMOTE	: Sentetik Azınlık Aşırı Örnekleme Tekniği
SVM	: Destek Vektör Makineleri
TOA	: Atmosfer Üstü Reflektans
XGBoost	: Aşırı Gradyan Arttırma
XAI	: Açıklanabilir Yapay Zekâ
XML	: Açıklanabilir Makine Öğrenmesi



SEMBOLLER

X_i	: Ölçülen toprak tuzluluk değeri
Y_i	: Tahmin edilen toprak tuzluluk değeri
$\bar{X}$	: Ortalama ölçülen toprak tuzluluk değeri
R^2	: Determinasyon katsayısı



ÇİZELGE LİSTESİ

Sayfa

Çizelge 2.1 : Landsat 8 uydusunun bantlarının spektral ve mekânsal özellikleri.....	15
Çizelge 3.1 : Toprak tuzluluğunu modellemek için kullanılan özellikler.	23
Çizelge 4.1 : Batı Urmiye Playası makine öğrenme modellerinin doğruluk sonuçları.	40
Çizelge 4.2 : Batı Urmiye Playası test örneklerinin tuzluluk seviyeleri incelenmesi.	40
Çizelge 4.3 : Bonab bölgesi makine öğrenme modellerinin doğruluk sonuçları.....	42
Çizelge 4.4 : Bonab bölgesi test örneklerinin tuzluluk seviyeleri incelenmesi.	42



ŞEKİL LİSTESİ

Sayfa

Şekil 1.1 : Toprak tuzluluğunun dünya yüzeyinde dağılımı (Stavi ve diğ., 2021).	2
Şekil 1.2 : Toprak tuzluluğunun uzaktan algılama ile tespiti literatür kelime bulutu..	7
Şekil 1.3 : Makine öğrenme yöntemleri ile toprak tuzluluğunun uzaktan algılama ile tespiti literatür kelime bulutu.	10
Şekil 1.4 : Açıklanabilir yapay zeka yöntemleri literatür kelime bulutu.	12
Şekil 2.1 : Çalışma alanları: a) genel bakış, b) Urmiye Gölü, c) Batı Urmiye Playası, d) Bonab Bölgesi.....	14
Şekil 2.2 : Batı Urmiye Playası toprak örneklerinin mekansal dağılımı.	16
Şekil 2.3 : Bonab bölgesi toprak örneklerinin mekansal dağılımı.....	16
Şekil 3.1 : Tez çalışmasının iş akış diyagramı.....	20
Şekil 3.2 : Rastgele Orman Regresyonu genel mimarisi.	25
Şekil 3.3 : XGBoost Regresyonu genel mimarisi.	26
Şekil 3.4 : Örnek SHAP özet grafiği.	28
Şekil 3.5 : Örnek SHAP özellik önem grafiği.	29
Şekil 3.6 : Örnek SHAP kuvvet grafiği.	29
Şekil 3.7 : Örnek SHAP bağımlılık grafiği.....	30
Şekil 4.1 : Batı Urmiye Playası için eğitim örneklerinin tuzluluk seviyelerine göre dağılımı.....	32
Şekil 4.2 : Batı Urmiye Playası için test örneklerinin tuzluluk seviyelerine göre dağılımı.....	32
Şekil 4.3 : Bonab bölgesi için eğitim örneklerinin tuzluluk seviyelerine göre dağılımı.	33
Şekil 4.4 : Bonab bölgesi için test örneklerinin tuzluluk seviyelerine göre dağılımı.....	33
Şekil 4.5 : Batı Urmiye Playası için aşırı örnekleme sonucu eğitim örneklerinin tuzluluk seviyelerine göre dağılımı.	34
Şekil 4.6 : Bonab bölgesi için aşırı örnekleme sonucu eğitim örneklerinin tuzluluk seviyelerine göre dağılımı.	34
Şekil 4.7 : Batı Urmiye Playası için eğitim verileri üzerinde modellerin dağılım grafiği	35
Şekil 4.8 : Bonab bölgesi için eğitim verileri üzerinde modellerin dağılım grafiği. .	36
Şekil 4.9 : Batı Urmiye Playası için test verileri üzerinde modellerin dağılım grafiği.	38
Şekil 4.10 : Bonab bölgesi için test verileri üzerinde modellerin dağılım grafiği.	39
Şekil 4.11 : Batı Urmiye Playası tuzluluk haritaları.	41
Şekil 4.12 : Bonab bölgesi tuzluluk haritaları.	43
Şekil 4.13 : Batı Urmiye Playası için geliştirilen modellerin mutlak SHAP değerleri.	45
Şekil 4.14 : Bonab bölgesi için geliştirilen modellerin mutlak SHAP değerleri.	46
Şekil 4.15 : Batı Urmiye Playası için SHAP değerlerinin özelliklere göre modele olan etkisi.	47

Şekil 4.16 : Bonab bölgesi için SHAP değerlerinin özelliklere göre modele olan etkisi.	48
Şekil 4.17 : Batı Urmiye Playası NDVI ve ENDVI özelliklerinin bağımlılık grafiği.	50
Şekil 4.18 : Bonab bölgesi Mavi ve EVI özelliklerinin bağımlılık grafiği.	51
Şekil 4.19 : Batı Urmiye Playası 1. test örneği SHAP değerleri.	53
Şekil 4.20 : Batı Urmiye Playası 2. test örneği SHAP değerleri.	54
Şekil 4.21 : Batı Urmiye Playası 3. test örneği SHAP değerleri.	55
Şekil 4.22 : Batı Urmiye Playası 4. test örneği SHAP değerleri.	57
Şekil 4.23 : Batı Urmiye Playası 5. test örneği SHAP değerleri.	58
Şekil 4.24 : Batı Urmiye Playası 6. test örneği SHAP değerleri.	59
Şekil 4.25 : Batı Urmiye Playası 7. test örneği SHAP değerleri.	61
Şekil 4.26 : Batı Urmiye Playası 8. test örneği SHAP değerleri.	62
Şekil 4.27 : Batı Urmiye Playası 9. test örneği SHAP değerleri.	63
Şekil 4.28 : Batı Urmiye Playası 10. test örneği SHAP değerleri.	64
Şekil 4.29 : Bonab bölgesi 1. test örneği SHAP değerleri.	66
Şekil 4.30 : Bonab bölgesi 2. test örneği SHAP değerleri.	67
Şekil 4.31 : Bonab bölgesi 3. test örneği SHAP değerleri.	68
Şekil 4.32 : Bonab bölgesi 4. test örneği SHAP değerleri.	70
Şekil 4.33 : Bonab bölgesi 5. test örneği SHAP değerleri.	71
Şekil 4.34 : Bonab bölgesi 6. test örneği SHAP değerleri.	72
Şekil 4.35 : Bonab bölgesi 7. test örneği SHAP değerleri.	74
Şekil 4.36 : Bonab bölgesi 8. test örneği SHAP değerleri.	75
Şekil 4.37 : Bonab bölgesi 9. test örneği SHAP değerleri.	76
Şekil 4.38 : Bonab bölgesi 10. test örneği SHAP değerleri.	78
Şekil 4.39 : Bonab bölgesi 11. test örneği SHAP değerleri.	79
Şekil 4.40 : Bonab bölgesi 12. test örneği SHAP değerleri.	80
Şekil 4.41 : Bonab bölgesi 13. test örneği SHAP değerleri.	81

AÇIKLANABİLİR YAPAY ZEKA YÖNTEMLERİ İLE TOPRAK TUZLULUĞU HARİTALARININ ÜRETİLMESİ

ÖZET

Toprak çok önemli bir doğal kaynak olmanın yanı sıra yaşadığımız ekosistemin değişmez bir parçasıdır. Her doğal kaynak gibi toprakta tüketimi ve kullanımı konusunda dikkat edilmesi gereken bir kaynaktır. Toprak özellikle tarımsal faaliyetler için vazgeçilmez bir konumdadır. Ancak tarımsal faaliyetlerde aşırı sulama, yer altı sularının fazla kullanımı, iklim değişikliği, tozla taşınma gibi sebepler yüzünden tuzlanmalar meydana gelmektedir. Bu tuzlanma durumu toprak tuzluluğu olarak incelenmektedir.

Toprak tuzluluğu birçok nedenden dolayı meydana gelse de insani faaliyetler toprağın bu tuzlanma sürecini hızlandırmaktadır. Özellikle de iklim değişikliğinin son yıllarda etkisinin artması ile birlikte toprak tuzluluğu problemi küresel bir boyuta ulaşmıştır. Bu problemin küresel ölçekte anlamak için Birleşmiş Milletler Gıda ve Tarım Örgütü'nün yaptığı araştırmalar incelenebilir. Örgüte göre dünya üzerinde 397 milyon hektar alan tuzlu topraklar ile kaplıdır. Bu alanların artması ise dünyayı gıda krizi, ekonomik kriz gibi sıkıntılara sokacağı öngörülmektedir.

Toprak tuzluluğu ile mücadelede izleme ve tespit etme müdahale kadar önemlidir. Doğru alanlara erken teşhis ile gerekli düzenlemeler ile toprak tuzluluğunun artması önlenir. Toprak tuzluluğunun tespitinde uzaktan algılama sistemleri sıklıkla kullanılmaktadır. Özellikle optik sistemler toprağın spektral yansıtım değerlerini inceleyerek tuzluluk seviyesi hakkında bilgi verebilirler.

Makine öğrenme algoritmaları çevre problemlerinin modellenmesi için literatürde sıklıkla kullanılmaktadır. Özellikle rastgele orman, destek vektör makineleri gibi popüler yöntemler toprak tuzluluğu tespiti içinde kullanılmaktadır. Bu çalışma kapsamında rastgele orman yöntemine ek olarak XGBoost algoritması da kullanılmıştır.

Toprak tuzluluğunu seçilen iki çalışma alanında incelemek için yersel toprak örnekleri ile mekânsal ve zamansal olarak eşleşen Landsat 8 uydu görüntüleri kullanılmıştır. Batı Urmiye Playası için 2018 yılında son bahar mevsiminde 71 adet toprak örneği toplanmışken, Bonab bölgesi için ise 74 adet toprak örneği 2014 yılında son bahar mevsiminde toplanmıştır. Batı Urmiye Playasında toprak örnekleri 1:2.5 süspansiyonda hazırlanan solüsyonlar ile elektriksel iletkenlikleri laboratuvar ortamında ölçülmüştür. Bonab bölgesi için ise iletkenlikler doğrudan Jenway 4510 iletkenlik ölçer ile ölçülmüştür.

Toprak tuzluluk örnekleri tuzluluk seviyelerine göre eşit olarak dağılıma sahip olmadıkları için veri çeşitlendirme yöntemleri kullanılmıştır. Bu yöntemler ise rastgele aşırı örnekleme ve sentetik azınlık aşırı örnekleme tekniği kullanılmıştır. Makine öğrenme problemlerinde olan veri dengesizliği bu şekilde giderilmesi amaçlanmıştır.

Google Earth Engine (GEE) sunduğu petabaytlarca analize hazır mekânsal veri ve hesaplama gücü ile kullanıcılarına büyük verileri hızlıca analiz edip görselleştirmesini sağlar. Bu çalışma kapsamında GEE uydu görüntülerinin işlenmesi, makine öğrenme modeli için özelliklerin üretimi ve analizi için kullanılmıştır.

Toprak tuzluluğunu modellemek için SCORPAN formülü baz alınarak özellikler tanımlanmıştır. Bu değişkenler toprağın parametrelerini, iklimini, bitki örtüsünü, litolojisi vb. modellemek için kullanılır. Toplamda 26 adet özellik oluşturulmuştur.

Geliştirilen modelleri açıklayabilmek için açıklanabilir yapay zekâ yöntemlerinden olan SHAP yöntemi tercih edilmiştir. Modelden bağımsız olarak sonuçları açıklayabilmesi yöntemi öne çıkarmaktadır. Yöntem ile modelleri açıklarken 3 temel başlık üzerinde durulmuştur. Bunlar sırasıyla özelliklerin modele olan mutlak katkıları, özelliklerin değerlerine göre modele katkısının hangi yönde olduğu ve test örnekleri bazında değişkenlerin modele katkıları incelenmiştir.

Çalışmanın bulguları yani toprak tuzluluk haritaları 6 farklı kombinasyona göre üretildi, doğruluk analizi yapıldı ve açıklandı. Bu kombinasyon 2 farklı makine öğrenme algoritmasının ve 3 farklı veri çeşitlendirme yönteminden ileri gelmektedir. Doğruluk analizi sonuçlarına göre Batı Urmiye Playasında XGBoost yöntemi rastgele aşırı örnekleme metodu ile birlikte en iyi sonuca ulaşmıştır. R^2 değeri olarak 0.76 değerine ve RMSE olarak 16.60 değerine ulaşmıştır. Bonab bölgesi için ise yine XGBoost yöntemi bu kez SMOTE aşırı örnekleme yöntemi ile en yüksek doğruluğa ulaşmıştır. R^2 değeri olarak 0.83'e ulaşırken RMSE değeri olarak ise 12.29 olmuştur.

Görsel açıdan haritalar yorumlandığında ise, XGBoost algoritmasının Batı Urmiye Playasında öne çıktığı görülürken, Bonab bölgesinde rastgele orman algoritması daha iyi sonuçlar vermektedir. Özellikle Bonab bölgesinde bulunan tuz yapılarını rastgele orman kolaylıkla ayırt edebilmesi ile öne çıkmıştır.

Açıklanabilir yapay zekâ için yapılan SHAP analizleri incelendiğinde ise, Batı Urmiye Playası için ENDVI, NDVI, EVI gibi bitki indislerinin daha çok katkı verdiği görülmüştür. Bu alanda yapılan önceki çalışmaya benzer olarak CRSI indisinde yine yüksek katkı verdiği görülmüştür. Bonab bölgesi için ise Landsat 8 uydu görüntüsünün kendi bantlarından olan mavi ve yeşil bantların modele en çok katkı veren değişkenler olduğu göze çarpmaktadır. Bitki indisleri Batı Urmiye Playası'ndaki kadar katkı vermemiştir.

Batı Urmiye Playası için SHAP değerleri göz önüne alındığında bitki örtüsünü gösteren indislerin artması tuzluluğu azaltırken, NDWI indisi arttıkça tuzluluğun arttığını göstermektedir. Bonab bölgesindeki SHAP değerleri incelendiğinde ise parlak tuz yapılarının görünür bölge bantlarının yansıtımını arttırdığı ve parlaklık ile doğru bir orantıya sahip olduğu görülmüştür. Test örneklerinde de yukarıdaki sonuçları doğrulamak mümkündür.

Sonuç olarak toprak tuzluluğu makine öğrenme algoritmaları ile tespit edilebildiği, bu algoritmaların ise veri çeşitleme yöntemleri kullanılarak doğruluklarının artırılabilirdiği açıktır. Aynı şekilde SHAP değerleri bu makine öğrenme algoritmalarını açıklayarak hangi parametrenin toprak tuzluluğunu belirlemede nasıl bir katkı sunduğu genel ve örnek bazında açıklayabilmiştir. Birleşmiş Milletlerin 2030 Yılı Sürdürülebilir Kalkınma Hedefleri kapsamında gıda güvenliği, su kullanımı ve sürdürülebilir ekosistemler çerçevesinde toprak tuzluluğu önemli bir yer işgal etmektedir. Özellikle de gıda ile olan ilişkisi nedeniyle toprak tuzluluğu yakın gözlem altında tutulmalı ve karar vericilere destek olunarak üstesinden gelinmelidir.

Gelecekteki toprak tuzluluęu alıřmaları iin farklı uydu sistemleri, daha yksek mekânsal özünrlkl uydular, klasik makine ęrenme yntemleri yerine derin ęrenme yntemleri kullanılabilir. Trkiye zerinde de hali hazırda bu problemin bař gsterdięi bir alan yeni bir alıřma iin seilebilir. Optik uydular yeni radar uyduları da yine kullanılmak iin nerilebilir.





SOIL SALINITY MAPPING USING EXPLAINABLE ARTIFICIAL INTELLIGENCE METHODS

SUMMARY

Soil is a very important natural resource and a permanent part of the ecosystem we live in. Like every natural resource, soil is a resource that needs attention in terms of its consumption and use. Soil is especially indispensable for agricultural activities. Healthy soil is important for many reasons such as vegetation growth, food supply chain, and water filtration. In addition to that, healthy soil is more resilient to flood, drought, and fire events. Especially, it will also regulate climate because of consist carbon inside. However, salinization occurs due to reasons such as excessive irrigation in agricultural activities, excessive use of groundwater, climate change, dust transportation. This salinization situation is examined as soil salinity.

Although soil salinity occurs for many reasons, human activities accelerate this process of soil salinization. Especially with the increasing impact of climate change in recent years, the problem of soil salinity has reached a global dimension. In order to understand this problem on a global scale, the research conducted by the Food and Agriculture Organization of the United Nations can be examined. According to the organization, 397 million hectares of land in the world are covered with saline soils. If these areas increase, it is predicted that the world will face problems such as food crisis and economic crisis.

Monitoring and detection are as important as prevention in the fight against soil salinity. Increasing soil salinity can be prevented by early diagnosis and necessary arrangements in the right areas. Remote sensing systems are frequently used in the detection of soil salinity. Especially optical systems can provide information about the salinity level by examining the spectral reflectance values of the soil.

Machine learning algorithms are frequently used in the literature for modeling environmental problems. Especially popular methods such as random forest and support vector machines are used in soil salinity detection. In this study, in addition to the random forest method, the extreme gradient boosting (XGBoost) algorithm was also used.

Landsat 8 satellite imagery, spatially and temporally matched with situ soil samples, was used to determine soil salinity in the two selected study areas. For West Playas Urmia, 71 soil samples were collected in 2018, while 74 soil samples were collected for Bonab region in 2014. In West Playas Urmia, the electrical conductivity of the soil samples was measured in the laboratory using solutions prepared in 1:2.5 suspension. For the Bonab region, conductivities were measured directly with a Jenway 4510 conductivity meter.

Since soil salinity samples were not evenly distributed according to salinity levels, oversampling methods were used. These methods are random oversampling (ROS)

and Synthetic Minority Oversampling Technique (SMOTE). In this way, it is aimed to overcome the data imbalance in machine learning problems.

Google Earth Engine (GEE) offers petabytes of analysis-ready spatial data and computational power, allowing users to quickly analyze and visualize big geo-spatial data. In this study, GEE was used for processing satellite imagery, generating and analyzing environmental covariates.

To model soil salinity, environmental covariates are defined based on the SCORPAN formula. These variables are used to model soil parameters, climate, vegetation, lithology, etc. In total, 26 environmental covariates were created.

In order to explain the models developed, SHAP method, which is one of the explainable machine learning methods, was preferred. Its ability to explain the results independently of the model makes the method stand out. While explaining the models with the method, 3 main topics were emphasized. These are, the absolute contribution of environmental covariates to the model, the direction in which environmental covariates contribute to the model according to their values, and the contribution of variables to the model on the basis of test samples, respectively.

The findings of the study, namely soil salinity maps, were produced, accuracy analyzed and explained according to 6 different combinations of analysis. These combinations come from 2 different machine learning algorithms and 3 different oversampling methods. According to the results of the accuracy analysis, the XGBoost method together with the random oversampling method achieved the best result in the West Playas Urmia. The R^2 value is 0.76 and the RMSE is 16.60. For the Bonab region, the XGBoost method again achieved the highest accuracy, this time with the SMOTE oversampling method. The R^2 value reached 0.83 and the RMSE value was 12.29.

When the maps are interpreted visually, it is seen that the XGBoost algorithm gives better results in the West Playas Urmia, while the random forest algorithm gives better results in the Bonab region. Especially in the Bonab region, the random forest stands out with its ability to easily distinguish the salt structures.

When SHAP analyses for explainable machine learning were analyzed, it was seen that plant indices such as ENDVI, NDVI, EVI contributed more for the West Playas Urmia. Similar to the previous study conducted in this area, it was observed that the CRSI index also made a high contribution. For the Bonab region, the blue and green bands, which are the bands of Landsat 8 satellite image, are the variables that contribute the most to the model. Vegetation indices did not contribute as much as in the other study area.

Considering the SHAP values for the West Playas Urmia, the increase in the indices indicating vegetation cover decreases the salinity, while the increase in the NDWI index increases the salinity. When the SHAP values in the Bonab region are analyzed, it is seen that the bright salt-crust structures increase the reflection of the visible region bands and have a direct proportion with brightness. It is possible to confirm the above results in the test samples.

As a result, we explained that soil salinity can be detected by machine learning algorithms and that the accuracy of these algorithms can be increased by using oversampling methods. Likewise, by explaining these machine learning algorithms, SHAP values were able to explain which parameter contributes to the determination of soil salinity on a general and example basis. Soil salinity occupies an important place within the framework of food security, water use and sustainable ecosystems

within the United Nations Sustainable Development Goals. Soil salinity should be kept under close observation, especially because of its relationship with food production and security, and should be tackled by supporting decision-makers.

For future soil salinity studies, different satellite systems, satellites with higher spatial resolution, deep learning methods instead of classical machine learning methods can be used. An area in Türkiye where this problem already exists could be selected for a new study. Optical satellites and new radar satellites can also be proposed for use.

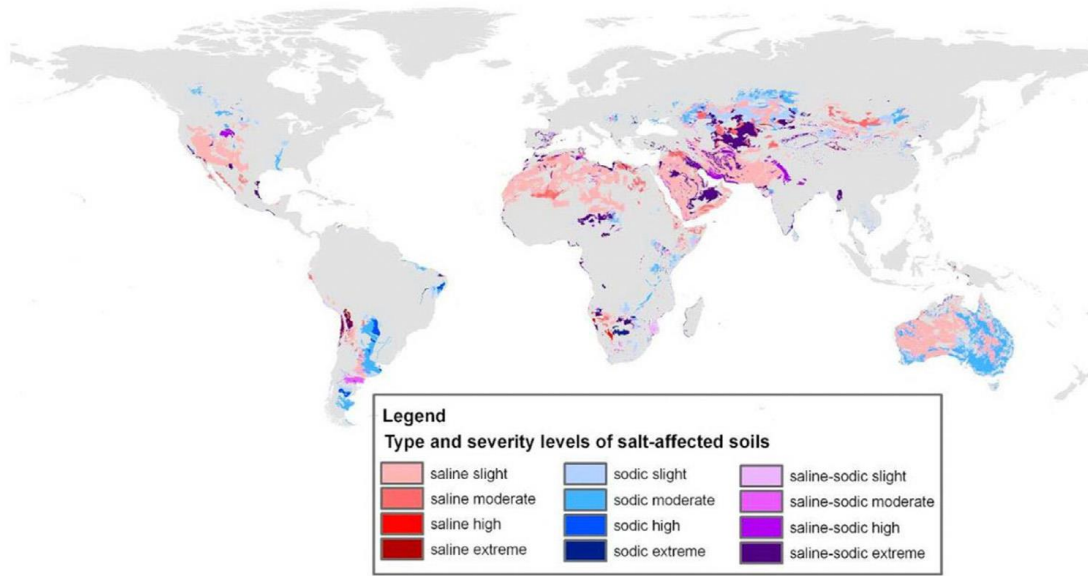




1. GİRİŞ

Toprak çok önemli bir doğal kaynak olmanın yanı sıra mikro ölçekten makro düzeye geniş biyo çeşitliliğe sahip habitatlara ortam olmaktadır. Toprak altı ve üzerinde birçok ekosistemde canlılar bulunmaktadır. Burada bulunan canlı yaşam için su önemli bir rol oynamaktadır. Toprağın su içeriği toprak üzerinde ve içinde yaşayan canlıları doğrudan etkilese de dolaylı olarak insanları da etkiler (Brevik ve diğ., 2015). Sağlıklı toprak; bitki gelişimi, insanların besin ihtiyacının karşılanması ve suyun filtrasyonu açısından son derece önemlidir. Ayrıca sağlıklı toprak koşullarının olması sel, kuraklık ve yangın gibi afetlere karşı arazinin daha dirençli olmasını destekler ve bünyesinde barındırdığı karbon ile iklimin düzenlenmesinde rol oynar. Toprakların ve biyo çeşitliliğin korunması son derece önemlidir. Fakat, doğal ya da insan kaynaklı nedenlerle topraklarda bozulmalar veya köklü değişimler meydana gelmektedir. Özellikle tarım için kullanılan topraklarda, ekim ve sulama davranışlarının kontrollü yönetilmemesi sonucunda, toprak yapısında, ağırlıklı olarak insan kaynaklı, değişimler meydana gelmektedir. Bu değişimlerden birisi de toprak tuzluluğudur (Stavi ve diğ., 2021). Tuzlu topraklar halomorfik topraklar altında incelenmektedir. Solonçaklar olarak da ifade edilen bu topraklar eğer kabarık ise bünyesinde %20 – 60 arasında $\text{Na}_2\text{SO}_4 \cdot \text{H}_2\text{O}$ içerir. Bu gibi alanlarda tuzlar yüzeyde birikerek nem ile şişer ve kümelenmeler oluşur. Dahili solonçaklar ise yüzeyden yıkanan tuzun belirli derinliklerde birikmesi ile oluşmaktadır. Alkali ve tuzlu alkali olarak da ifade edilen sodyum tuzlarını içeren topraklarda bu kapsamda incelenmektedir (Atalay, 2006).

Toprak tuzluluğu insan kaynaklı faktörler nedeniyle oluşabilirken aynı zamanda da doğal faktörlerden dolayı da gerçekleşebilmektedir. Doğal nedenlerden dolayı zamanla toprak tuzluluğu artsa da; tarımsal faaliyetler ve sulama aktiviteleri gibi insan kaynaklı faktörler bu süreci hızlandırır. Özellikle kurak arazileri toprak tuzluluğu daha çok etkiler. Son yıllarda, iklim değişikliği nedeniyle sıklaşan kuraklık olayları ise bu sorunu küresel ölçeğe taşımaktadır. Küresel ölçekte toprak tuzluluğunun dağılımı Şekil 1.1’de gösterilmiştir. (Stavi ve diğ., 2021). Kurak ve yarı-kurak iklim bölgelerinde bulunan alanlar toprak tuzluluğu probleminden daha çok etkilenmektedir.



Şekil 1.1 : Toprak tuzluluğunun dünya yüzeyinde dağılımı (Stavi ve diğ., 2021).

Birleşmiş Milletler Gıda ve Tarım Örgütü (Food and Agriculture Organization of the United Nations, FAO) dünya üzerindeki toprakların %3'ün üst tabakasının, %6'sının ise alt tabakalarının tuzluluktan etkilendiğini belirtmektedir (Url-1). Tuzlu toprakların artması, özellikle de tarım alanlarındaki toprakların tuzlanması büyük ekonomik kayıplara neden olmaktadır. Örneğin Umman'da yapılan bir çalışmaya göre az tuzlu toprakların orta tuzlu seviyeye gelmesi 28% oranında tarımsal geliri düşürürken, az tuzlu toprakların çok tuzlu toprak seviyesine gelmesi ise tarımsal geliri 76% oranında düşürdüğü tespit edilmiştir. Toprak tuzluluğunun şiddeti arttıkça, bitkilerin verimi düşmekte belirli bir seviyeden sonra bitkiler canlılığını kaybetmektedir. Toprak tuzluluğu toprağın yapısına, hidrolojik, iklim, tarımsal, sosyal ve ekonomik dengelere bağlı olan dinamik bir problemdir (Allbed ve Kumar, 2013; Gorji ve diğ., 2020; Aksoy ve diğ., 2022).

Toprak tuzluluğunu uzaktan algılama ile tespit etmek için öncelikle tuzlu toprağın özelliklerini belirlenmelidir. Tuzlu topraklar üzerindeki tuz yapıları, bitki örtüsünün azlığı ve tuza dayanıklı bitkilerin varlığı gibi özellikler ile göze çarparlar (Allbed ve Kumar, 2013). Günümüzde uzaktan algılama teknolojilerinde gerçekleşen gelişmeler bölgesel ve küresel ölçekte toprak tuzluluğunun belirlenmesinde önemli bir rol oynamıştır. Landsat 8, Sentinel 2 gibi uydular yüksek mekânsal, spektral ve zamansal çözünürlüklerinin sağladığı avantajlar ve ücretsiz erişilebiliyor olmaları nedeni ile, sıklıkla bu problemin tespitinde ve izlenmesinde kullanılmaktadır. Uzaktan algılama verilerine ek olarak arazide yapılmış yersel tuzluluk ölçümleri de doğru modeller

geliştirmek için önemlidir. Toprak örnekleri ile uydu görüntüleri arasında oluşturulan regresyon ve sınıflandırma modelleri için istatistiksel ve makine öğrenmesi modelleri de sıklıkla kullanılmaktadır (Gorji ve diğ., 2020; J. Wang ve diğ., 2021; Aksoy ve diğ., 2022).

Literatürde farklı modelleme teknikleri olmasına karşın; özellikle son yıllarda, performans ve doğrulukları istinaden makine öğrenme modelleri öne çıkmaktadır. Özellikle yapay sinir ağları (Artificial Neural Networks, ANN), rastgele orman (Random Forest, RF), ve destek vektör makineleri (Support Vector Machines, SVM) yöntemleri toprak tuzluluğu probleminin incelenmesinde yaygın olarak kullanılmaya başlanmıştır. Rastgele orman algoritması özellikle büyük veri setlerinde, çok fazla değişkeni efektif olarak kullanabildiği için tercih sebebi olmaktadır (Yildirim, 2021; Kabiraj ve diğ., 2022). Makine öğrenme algoritmalarının ek olarak derin öğrenme yöntemleri de toprak tuzluluğu konusunda kullanılmaya başlanmıştır. Özellikle çok yüksek mekânsal çözünürlüklü uydu görüntülerinde klasik makine öğrenme yönteminden biri olan destek vektör makinelerinden çok daha iyi sonuçlar verdiği bulunmuştur. U-NET mimarisi bu alanda kullanılmaktadır (Akca ve Gungor, 2022). Fakat, derin öğrenme tabanlı yaklaşımların uygulanmasında ihtiyaç duyulan etiketlenmiş veri seti gerekliliği ve modellerin genelleştirilebilir olması dikkat edilmesi gereken önemli hususlardır.

Makine öğrenme ve derin öğrenme yöntemlerinde her ne kadar doğruluk artsa da kapalı kutu olarak kalmaya devam etmektedirler. Bu yöntemlerin kararları nasıl verdiği, doğru özellikleri dikkate alıp, doğru kararlar alıp almadığı günümüzün önemli çalışma alanlarından birisidir. Uygulanan bir yapay zekâ modelini yorumlamak ve açıklamak, geliştirilen modelin güvenilirliği, transfer edilebilirliği ve genelleştirilmesi açısından çok önemlidir. Bu çalışmalar açıklanabilir yapay zekâ (Explainable Artificial Intelligence-XAI) olarak tarif edilmektedir. Açıklanabilir yapay zekâ literatürde farklı başlıklar altında incelenmektedir. Bu başlıklar şeffaflık, yorumlanabilir, açıklanabilir, alan bilgisi ve bilimsel tutarlılıktır (Roscher ve diğ., 2020). Toprak tuzluluğu kapsamında değerlendirildiğinde; tuzluluk kestirimine yönelik geliştirilen makine öğrenme modellerinin, toprak tuzluluğunu belirlerken nasıl karar verdiği, hangi model girdilerinin sonucu ne kadar etkilediği ve toprak tuzluluğunu belirlemenin yanında doğru niteliklere göre belirlenip belirlenmediğini de

görmekte önemlidir. Literatürde, toprak tuzluluğunun modellenmesinin açıklanabilir yapay zekâ yaklaşımları ile ele alındığı çalışmalara henüz rastlanmamıştır.

1.1 Tezin Amacı

Bu tezin ana amacı, uzaktan algılama verileri ve yersel toprak ölçümlerini kullanarak; geniş alanların farklı makine öğrenmesi tabanlı regresyon yöntemler ile toprak tuzluluğun mekânsal dağılımını belirlemektir. Bu ana amaç çerçevesinde tezin hedefleri ise;

- Landsat-8 uydu görüntülerinin farklı iki çalışma alanı için toprak tuzluluğunu belirlemedeki performansını araştırmak,
- Rastgele Orman ve XGBoost makine öğrenme modellerinin, toprak tuzluluğu belirlemedeki performansını karşılaştırmak,
- Makine öğrenmesi modellerine girdi olarak sunulabilecek farklı özelliklerin, tahmin doğruluğuna etkisi incelemek,
- Uygulanan farklı makine öğrenmesi modellerinin sonuçlarını, açıklanabilir yapay zeka yöntemleri kullanarak anlamak ve detaylı açıklamak,
- Tüm aşamaları farklı açık kaynak kodlu platform, kütüphane ve veriler ile gerçekleştirip, çözüm önerisinin yaygın etkisinin ve uygulanabilirliğin yüksek olmasını sağlamaktır.

Bu hedefler doğrultusunda, Landsat-8 görüntüsünün farklı spektral bantları ve bu bantlardan oluşturulan farklı indisler, 26 adet özellik (feature) olarak düzenlenmiş ve bu özellikler ve yersel tuzluluk ölçümleri kullanılarak, iki farklı makine öğrenme tabanlı regresyon modeli eğitilmiştir. Referans veri sayısındaki kısıtlardan kaynaklı problemleri en aza indirmek amacı ile farklı veri çeşitlendirme yöntemleri denenmiştir. Son olarak, eğitilen makine öğrenme modelleri, açıklanabilir yapay zekâ yöntemleri (XAI) uygulanarak özelliklerin katkıları yönünden açıklanmıştır. Farklı XAI fonksiyon değerleri ile, hangi özelliklerin, kestirim değerine ne kadar katkı verdiği irdelenmiş, modellerin daha anlaşılabilir ve açıklanabilir hale getirilmesi hedeflenmiştir. Tüm bu analizler Google Earth Engine ve Google COLAB platformları üzerinde sklearn, xgboost, ee, matplotlib, PIL, pandas ve imblearn kütüphaneleri

kullanılarak gerçekleştirilmiştir. Bu platformlar sayesinde gerçekleştirilen analizler kolaylıkla tekrar edilebilir ve sonuçlar yaygınlaştırılabilir.

1.2 Literatür Araştırması

Bu tez çalışma kapsamında, tezin gelişim aşamaları üzerinden kurgulanarak literatür araştırması üç ana başlık altında incelenmiştir. Öncelikle toprak tuzluluğunun Uzaktan Algılama sistemleri ile tespiti ve izlenebilirliğinin araştırması yapılmıştır. Daha sonra ise, makine öğrenme algoritmaları ile toprak tuzluluğunun tespiti üzerine yoğunlaşmıştır. Literatürde, toprak tuzluluğunun belirlenmesinde, farklı makine öğrenme yöntemlerini kullanan çok sayıda çalışma olmasına karşın; XGBoost yöntemi ile yapılan çalışma sayısı kısıtlıdır. Ayrıca, literatürdeki çalışmalarda, incelenen alanlarının farklı olması da tuzluluk probleminin irdelenmesi açısından önemli bir husustur. Son aşamada, makine öğrenme algoritmalarının bizlere sunduğu modellerin açıklanabilmesi için açıklanabilir yapay zekâ yöntemleri ile ilgili literatür araştırması yapılmıştır. Literatürde toprak tuzluluğu, uzaktan algılama ve makine öğrenme algoritmaları ile ilgili birçok eser bulunmaktadır. Bu nedenle, 2018 ve 2023 yılları arasındaki makaleler literatür taraması kapsamında incelenmiştir. Ancak açıklanabilir yapay zekâ ve toprak tuzluluğu arasında literatürde çalışmaya rastlanmamıştır. Tez çalışması bu anlamda, literatürdeki bu önemli boşluğu doldurmayı da hedeflemektedir. Açıklanabilir yapay zekâ ile ilgili literatür araştırmasında benzer makine öğrenme algoritmaları ile çözülen çevre problemlerini açıklayan çalışmalar üzerinde durulmuştur.

1.2.1 Toprak tuzluluğunun uzaktan algılama yöntemleri ile tespiti

Racetin ve diğ. (2020) Hırvatistan'da bulunan Neretva nehri deltasında 2009 yılında toplanmış yersel toprak tuzluluğu verilerini ve Landsat 5 görüntülerini kullanarak çoklu lineer regresyon yöntemini kullanarak, toprak tuzluluğunu tahmin eden bir model üretmişlerdir. Landsat 5 görüntülerinde 27 tane farklı indis üretilmiş ve bu indisler ile tuzluluk değerleri arasında korelasyonlar bulunmuştur. 2 yıl süresince de toprak tuzluluğu Landsat 5 görüntüleri kullanılarak incelenmiştir. En yüksek korelasyon tuzluluk indisi 8 ((NIR-SWIR1) / (NIR+SWIR1)) ile 0,93 değeri elde etmiştir.

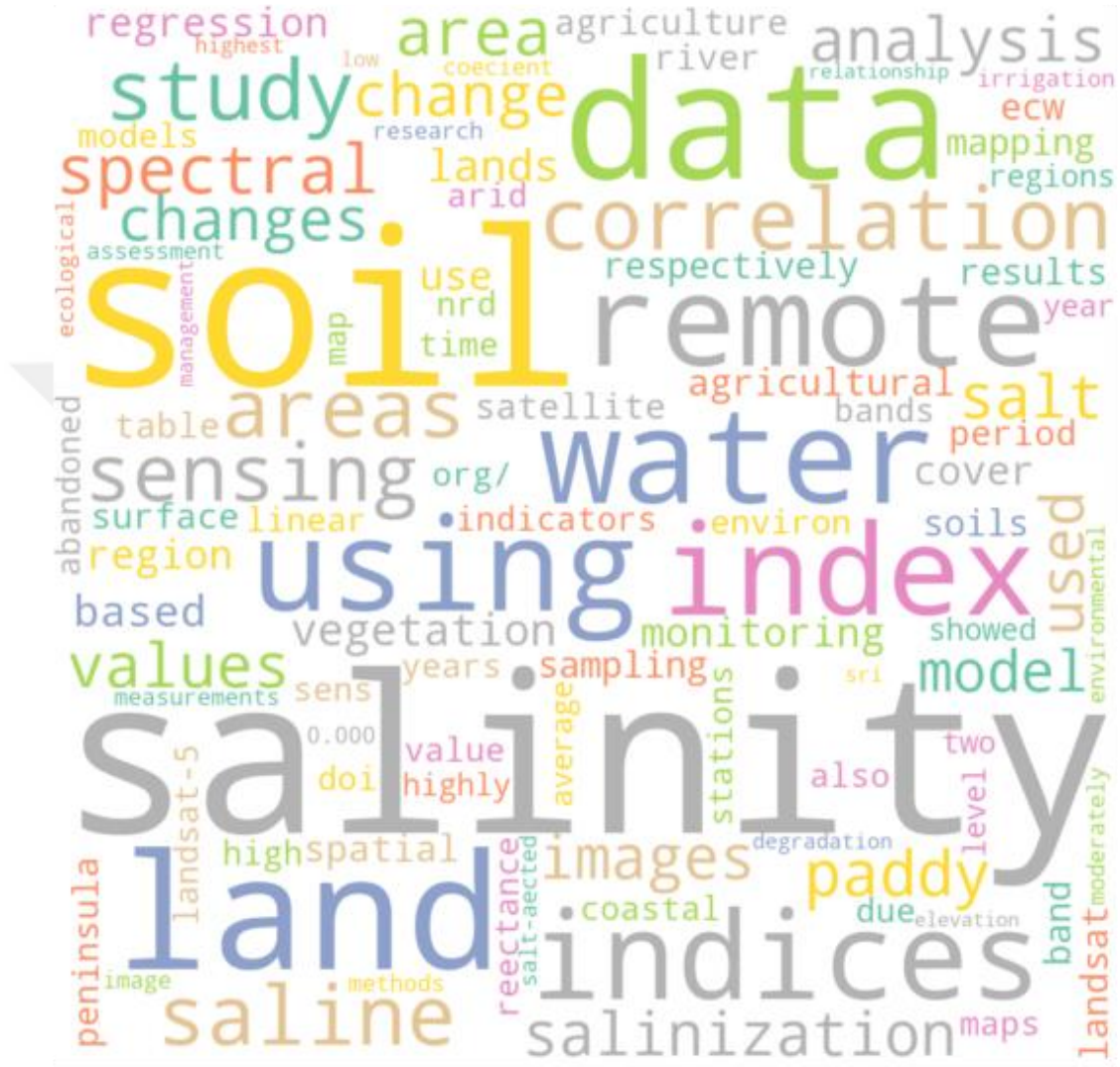
Gopalakrishnan ve Kumar (2021) Sri Lanka’da terk edilen çeltik arazilerinde toprak tuzluluğunu Landsat 5 ve Landsat 8 uydu görüntüleri ve 2019 yılında yersel toplanan örnekler ile incelemiştir. Metot olarak kısmi en küçük kareler regresyonu (Partial Least Square Regression, PLSR) kullanmıştır. Bu yöntemin daha önceki çalışmalarda başarılı sonuçlar verdiğini vurgulamışlardır. PLSR modeline girdi olarak 6 tuzluluk indeksi ve uydu görüntülerinin mavi, yakın kızıl ötesi, kısa dalga kızıl ötesi bantları ve sayısal yükseklik modeli kullanılmıştır. En yüksek R^2 değeri olarak ise 0.63 bulunmuştur.

Thiam ve diğ. (2021) Senegal’in kıyı şeridinde 2017 yılında toplanan 300 adet toprak örneği ve 1984, 1994, 2007 ve 2017 yıllarına ait Landsat görüntüleri kullanılarak toprak tuzluluğunun arazi üzerine yayılımı incelenmiştir. Landsat görüntülerinden elde edilen ıslaklık indisi (wetness index, TWI), normalize edilmiş fark bitki indisi (Normalized Difference Vegetation Index, NDVI), tuzluluk indisi ($\sqrt{Kırmızı * NIR}$) ve ayarlanmış toprak bitki indisine ek olarak kıyaya olan uzaklık ve yükseklik gibi parametrelerde çoklu lineer regresyon modeline girdi olarak kullanılmıştır. En yüksek R^2 değeri olarak ise 0.73 bulunurken RMSE değeri ise 0.68 olarak hesaplanmıştır.

Gorji ve diğ. (2017) Türkiye’de bulunan Tuz Gölü ve çevresindeki alandan 2002 yılında toplanan 322 toprak örneği ve 1990, 2002, 2006, 2011 ve 2015 yıllarına ait Landsat görüntüleri ile toprak tuzluluğu incelemiştir. 5 adet toprak tuzluluk indisi Landsat uydu görüntülerinden hesaplanmış ve üstel ve lineer regresyon kullanılarak modellenmiştir. Üstel regresyon modeli R^2 değeri olarak 0,93 değerine ulaşırken, lineer regresyon modeli ise 0.83 R^2 değerine ulaşmıştır. Bu değere ulaşan indis ise tuzluluk indisi $1 (\sqrt{Mavi * Kırmızı})$ olmuştur.

Kasim ve diğ. (2020) Çin’in Uygur bölgesinde bulunan Keriya vadisinde EM38-MK2 iletkenlik ölçer ile topladıkları 353 toprak örneği ve WordView-2 uydu görüntüleri ile toprak tuzluluğunu incelemişlerdir. Model olarak PLSR seçilmiştir. Tuzluluk değeri ile en iyi korelasyona sahip indis bulabilmek için optimize edilmiş spektral indisler kullanmıştır. Bu indisler seçilen 2 bandın oranını, normalleştirilmiş farkını ve karelerinin toplamın karekökünü bulurken, diğer bir indis ise seçilen 3 bandın karelerinin toplamının karekökünü bulmaktadır. Tüm bant kombinasyonları denenerek 2 ve 3 boyutlu korelasyon hesaplanmıştır. Bunun yanı sıra uydu görüntülerine 5 farklı matematiksel transformasyonlar uygulanmış. En yüksek doğruluk değerine RF-PLSR modeli R^2 0.79 ve RMSE 1.51 olarak ulaşmıştır.

Toprak tuzluluğunun uzaktan algılama ile tespiti için yapılan literatür araştırması aynı zamanda bir kelime bulutuna dönüştürülüp incelenmiştir. Kelime bulutu Şekil 1.2’de gösterilmektedir. Kelime bulutu incelendiğinde korelasyon analizlerinin yaygın olduğu, tarım ile olan ilişkilerinin incelendiği göze çarpmaktadır.



Şekil 1.2 : Toprak tuzluluğunun uzaktan algılama ile tespiti literatür kelime bulutu.

1.2.2 Makine öğrenme algoritmalarıyla toprak tuzluluğunun tespiti

Hateffard ve diğ. (2019) İran’ın Alborz şehrinde Landsat 8 uydu görüntülerini ve sayısal yükseklik modelini kullanarak toprak parametrelerini modellemişlerdir. Tehran Üniversitesi’nin Toprak Bilimi Bölümü’nün hazırladığı 362 toprak örneği verilerini kullanmışlardır. Toprak parametrelerinden elektriksel iletkenliği değerinin tuzluluk göstergesi olarak kabul etmişlerdir. Toprak tuzluluğunu modellerken karar ağacı ve yapay sinir ağları kullanmışlardır. Yapay sinir ağı modeli olarak 17 nörondan

oluşan giriş katmanına ek olarak 10 nörondan oluşan bir gizli katman ve 1 nörondan oluşan çıktı katmanı vardır. Karar ağacı modeli ile R^2 değeri 0.59 ve RMSE değeri 13.74'e ulaşırken yapay sinir ağları ile bu değerler R^2 için 0.95'e ve RMSE için 6.27'ye ulaşmıştır.

Aksoy ve diğ. (2022) İran'ın Urmiye gölü kenarında bulunan bir çalışma alanında 2016 yılında toplanan 70 adet toprak tuzluluğu örneklerini kullanarak makine öğrenme algoritmalarını karşılaştırarak incelemişlerdir. Bu makine öğrenme algoritmaları ise karar ağacı, rastgele orman ve destek vektör makineleridir. Landsat 8 ve Sentinel 2 uydu görüntülerini kullanarak 26 adet çevresel değişken/özellik (feature) üretmişlerdir. Bu özellikleri değişken azaltma yöntemi uygulayarak azaltmış ve makine öğrenme algoritmalarını seçilen daha az sayıdaki özellikler ile eğitmişlerdir. Tüm bu eğitimleri ve analizleri Google Earth Engine platformu üzerinde yapmışlardır. En iyi sonucu rastgele orman algoritması vermiştir. Landsat 8 uydu görüntüleri ile R^2 0,84 ve RMSE 6.85 değerine ulaşmışlarken, Sentinel 2 uydu görüntüleri ile R^2 0,90 ve RMSE 5.91 değerine ulaşmışlardır. Bu durum, tuzluluk belirlemede, mekânsal çözünürlük artışının olumlu etkisini de göstermektedir.

Kabiraj ve diğ. (2022) Google Earth Engine platformunu kullanarak Hindistan'ın Tamil Nadu bölgesinde arazinin toprak tuzluluğunu rastgele orman makine öğrenme algoritmasını kullanarak modellemişlerdir. 2020 yılında 593 adet toplanan yersel toprak tuzluluk örneği Sentinel 2 ve Landsat 8 uydu görüntüleri birleştirilerek analizler yapılmıştır. Rastgele orman algoritmasında uygun ağaç sayısını bulmak için 1 ile 500 arasında 10'ar aralıklarla denemeler yapılmıştır. Uydu görüntülerinin kendi bantlarının yanına 4 adet bitki indisi ve 11 adet tuzluluk indisi çalışma kapsamında kullanılmıştır. Landsat 8 ile geliştirilen modelde CRSI, SI-III((Yeşil*Kırmızı)/Mavi), SI-I(Mavi/Kırmızı) gibi indisler önemli çıkarken, Sentinel 2 ile geliştirilen modelde NDVI, SI5((Kırmızı*NIR)/Yeşil), SAVI gibi indisler öne çıkmıştır. Genel doğruluk değeri olarak %98'e ulaşmışlardır.

F. Wang ve diğ. (2019) 5 adet makine öğrenme metodunu 3 farklı çalışma alanında toplanan 390 adet toprak örneği ile toprak tuzluluğunu incelemişlerdir. Bu yöntemler ise en küçük mutlak küçültme ve operatör seçimi (Least Absolute Shrinkage and Selection Operator, LASSO), çok değişkenli uyarlanabilir regresyon uzanımları (Multiple Adaptive Regression Splines, MARS), sınıflandırma ve regresyon ağaçları (Classification and Regression Trees, CART), rastgele orman (Random Forest, RF) ve

stokastik gradyan ağaç desteğidir (Stochastic Gradient Treeboost, SGT). Landsat 8 uydu görüntülerinden üretilen çevresel değişkenleri kullanarak bu makine öğrenme algoritmaları eğitilmiştir. Çevresel değişkenlerden en önemlileri ise EEVI, EVI2, ENDVI, CRSI olmuştur. SGT algoritması R^2 0.63 ve RMSE 11.55 değeri ile en iyi sonuçları verirken onu izleyen algoritması ise R^2 0.61 ve RMSE 11.95 değeri ile RF'dir.

X. Wang ve diğ. (2020) kesirli derece filtresi (Fractional Order Filter) algoritması uyguladıkları Sentinel 2 uydu görüntüleri ve 2015 yılında toplanan 60 adet yersel veri ile destek vektör makineleri makine öğrenmesi modeli eğitmişlerdir. Destek vektör makineleri için en uygun parametreleri bulmak için grid arama yöntemini kullanmışlar. Çapraz doğrulama yöntemi ile de en uygun gamma ve cost değerini bulmuşlardır. Eğitilen optimum makine öğrenme algoritması ile de tuzlu alanların seviyeleri belirlenmiştir. Genel doğruluk olarak %90 ve üzeri elde edilmiştir.

J. Wang ve diğ. (2021) Çin'in Wensu şehrindeki kurak alanlarda toprak tuzluluğunu Sentinel 2 uydu görüntüleri ve yersel olarak 2019 yılında toplanan 160 adet toprak örneğini kullanarak modellemiştir. Model olarak 3 farklı makine öğrenme algoritması seçilmiştir bunlar; destek vektör makineleri, rastgele orman ve yapay sinir ağlarıdır. Modelleri eğitmek için girdi olarak tuzluluk indislerinin yanı sıra uydu görüntüsünün kendi bantları ve bu bantlara uygulanan temel bileşen analizi kullanılmıştır. Doğruluk analizi sonucu destek vektör makineleri R^2 değeri 0.88 ve RMSE değeri 4.89 ile en yüksek doğruluğu vermiştir.

Makine öğrenme yöntemleri ile toprak tuzluluğu literatürü de bir kelime bulutu olarak Şekil 1.3'te gösterilmiştir. Kelime bulutu incelendiğinde tuzluluk, uzaktan algılama gibi çıkmasını beklediğimiz kelimelerin dışında çevresel değişkenler, bazı makine öğrenme yöntemleri, kurak alanlar, Landsat 8 ve Sentinel 2 gibi uydu görüntülerinin öne çıktığı görülmüştür.

1.2.3 Açıklanabilir yapay zeka yöntemleri ile modellerin açıklanması

Açıklanabilir yapay zekâ konusunun, yer gözlem uydu verilerinin kullanıldığı çalışmalara yönelik uygulanması ve temel kavramları açıklayan ilk makalelerden biri 2020 yılında Roscher ve diğ. tarafından yayınlanmıştır. Doğa bilimlerindeki uygulamalar dikkate alınarak, açıklanabilir makine öğrenmesi (Explanaible Machine Larning-XML) yaklaşımının ele alındığı çalışmada, şeffaflık, yorumlanabilirlik ve

açıklanabilirlik temel öğeleri vurgulanmıştır. Bu çalışma sonrasında, uzaktan algılama alanında, özellikle 2021 ve 2022 yıllarında XML/XAI yaklaşımlarını içeren çalışmaların arttığı belirlenmiştir.



Şekil 1.3 : Makine öğrenme yöntemleri ile toprak tuzluluğunun uzaktan algılama ile tespiti literatür kelime bulutu.

Al-Najjar ve diğ. (2022) Butan Krallığındaki Chukka şehrinde sıklıkla gerçekleşen heyelanları, radar uydu görüntülerinin, Landsat 8 uydu görüntülerini ve sayısal yükseklik modeli verileri ve makine öğrenmesi yöntemlerini kullanarak modellemiştir. Makine öğrenme algoritması olarak rastgele orman ve destek vektör makineleri kullanılarak, heyelan haritaları üretilmiştir. Makine öğrenme algoritmalarını açıklayabilmek için ise Shapley değerleri (Shapley additive explanations, SHAP) kullanılmıştır. SHAP her iki model içinde aynı girdileri öne çıkarmıştır. Ancak birebir olarak aynı önemde değildirler.

Kakogeorgiou ve Karantzas (2021) derin öğrenme yöntemlerini kullanarak BigEarthNet ve SEN12MS veri setleri üzerinde sınıflandırma problemini ele almıştır. Bu veri setleri Sentinel 2 uydu görüntülerinden oluşmaktadır. CORINE arazi örtüsü/kullanımı verisi ise etiket olarak kullanılmıştır. Eğitilen modelleri açıklayabilmek için birçok yöntem test edilmiştir. Bu yöntemler ise Saliency, Input x Gradient, Integrated Gradients, Guided Backpropagation, Grad-CAM, Guided Grad-CAM, Occlusion, DeepLift, Lime ve SmoothGrad'dir. Bu yöntemlerden Occlusion, Grad-CAM ve Lime yöntemleri en yorumlanabilir olarak belirlenmiştir. Guided Backpropagation ise en az güvenilir metot olmuştur.

Stomberg ve diğ. (2022) Sentinel 2 uydu görüntüleri ile ürettikleri veri setinden insan eli değmiş ve değmemiş alanları açıklanabilir yapay zekâ ile tespit etmişler. Derin öğrenme yöntemi ile eğittikleri model uydu görüntüsünü iki sınıfta sınıflandırarak insan eli değip değmediği ile ilgili bilgi vermektedir. Derin öğrenme modelinin aktivasyonlarını inceleyerek hassaslık haritaları oluşturulmuş, bu haritalardan hangi alanların insan eli değdiğini hangi alanlara değmediğini mekânsal olarak tespit etmişler.

Betancourt ve diğ. (2022) troposferik ozon gazının rastgele orman makine öğrenme algoritması ile modellemiştir. 30'dan fazla çevresel değişken kullanarak eğittikleri modeli SHAP değerleri kullanarak açıklamıştır. Çevresel değişkenleri 6 ana başlık altına toplamışlardır. Bunlar; genel, arazi örtüsü, tarımsal değişkenler, ozon ile ilgili değişkenler, nüfus ve ışık kirliliğidir. Troposferik ozon modellemesinde, enlem ve yükseklik değişkenleri modeli en çok etkileyen özellikler olarak belirlemişlerdir.

Abdollahi ve Pradhan (2021) piksel tabanlı derin sinir ağları kullanarak, hava fotoğraflarındaki az ve çok yoğunluklu bitki örtüsü alanlarını sınıflandırmışlardır. Modele girdi olarak görüntülere ek olarak görüntülerden doku ayırt etmeyi sağlayan indislerde kullanılmıştır. Hangi değişkenlerin sınıflara nasıl katkılar sunduğunu incelemek için yani bir başka deyişle modeli açıklamak için SHAP değerlerini kullanmışlardır. Sınıf bazında değişkenlerin sırası değişse de bazı değişkenler bariz olarak tüm sınıflara yüksek oranda etki ettiği görülmektedir.

Açıklanabilir yapay zekâ yöntemleri için yapılan literatür araştırması için de kelime bulutu oluşturulmuştur. Şekil 1.4'te gösterilen kelime bulutuna göre makine öğrenme yöntemleri için kullanılan özelliklerin yanı sıra SHAP metodu da öne çıkmıştır.

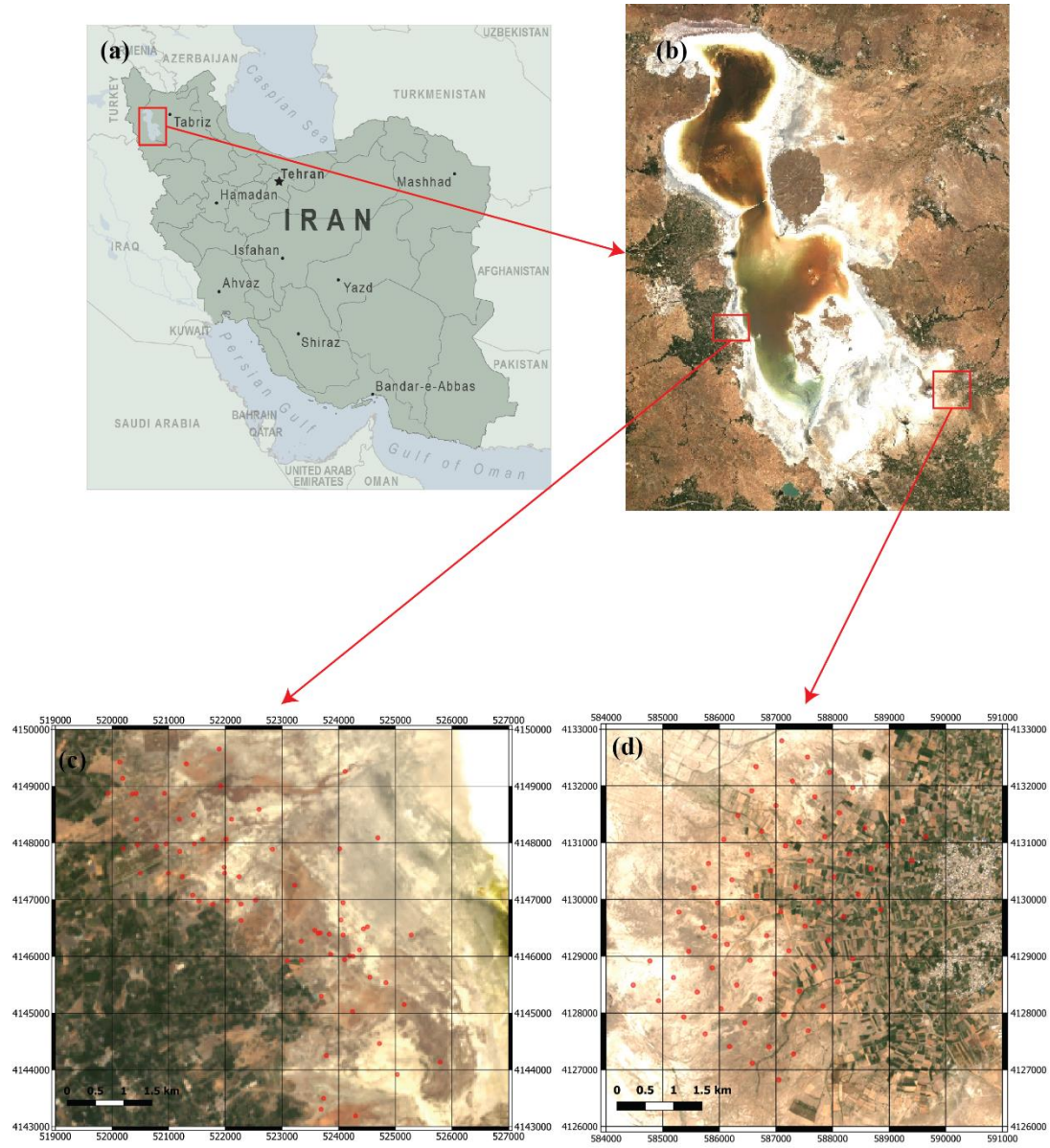
2. ÇALIŞMA BÖLGELERİ VE KULLANILAN VERİLER

2.1 Çalışma Bölgeleri

Bu tez kapsamında İran'ın Urmıye Gölü çevresindeki iki çalışma alanı incelenmiştir. Bu çalışma alanları sırasıyla Batı Urmıye Playası ve Bonab bölgesidir. Çalışma alanlarının İran'ın hangi bölgesinde kaldığı ve seçilen alanlar Şekil 2.1'in (a) ve (b) kısımlarında gösterilmiştir. Urmıye Gölü 6000 km² alanı kaplayan çok tuzlu büyük bir göldür. Bu göl 1976 yılından beri doğal park olarak ilan edilmiş ve UNESCO koruma listesinde dir. Urmıye Gölü 51,876 km²'lik bir havzaya sahiptir. Bu havzadan 6,5 milyon insan tarımsal ve sanayi faaliyetleri ile faydalanmaktadır. Tarım alanları alanın %10'unu kaplamasına rağmen yenilenebilir su kaynaklarının %90'ını tüketmektedir (Haghighi ve diğ., 2018). Fazla su tüketimi, yağışların yetersizliği gibi sebepler yüzünden gölün su ile kaplı yüzey alanı zaman içerisinde küçülmüştür. Su tüketimindeki artış yer altı sularının fazla tüketilmesine ve tuzlu yeraltı sularının tarımsal verimliliği azaltmasına neden olmuştur. Verimliliği azalan tarım arazileri ise terk edilmeye başlanmıştır (Hamzhepour ve diğ., 2018; Gorji ve diğ., 2020; Aksoy ve diğ., 2022).

2.1.1 Batı Urmıye playası

Şekil 2.1'in (c) kısmında ise Batı Urmıye Playasının Landsat 8 doğal renkli uydu görüntüsü üzerinde kullanılan toprak örnekleri gösterilmiştir. Çalışma alanı yaklaşık 82 km²'dir. 71 adet toprak örneği toplanmış ve laboratuvar da analiz edilmiştir. Toprak örneklerinin alana dağılımı Şekil 2.1 (c) de görülmektedir. Örnekler 2 – 12 Ekim 2018 tarihin de toprağın en üst 20 cm olan tabakasından toplanmıştır. Toprak örnekleri tuz birikiminin en yüksek olduğu kurak sezonun sonuna doğru toplanmıştır. Toplanan örnekler 1:2,5 oranında su ile süspansiyon oluşturulmuş ve elektriksel iletkenlik değeri (Electrical Conductivity, EC) ölçülmüştür. Veriler İran'ın Maragheh Üniversitesindeki Dr. Nikou Hamzhepour tarafından sağlanmıştır.



Şekil 2.1 : Çalışma alanları: a) genel bakış, b) Urmie Gölü, c) Batı Urmie Playası, d) Bonab Bölgesi

2.1.2 Bonab bölgesi

Bonab bölgesi Urmie Gölü'nün güneydoğu kısmında yer almaktadır ve Şekil 2.1'in (d) kısmında gösterilmiştir. Rakım 1270 m ile 1300 m arasında değişmektedir. Toplamda 74 adet toprak elektriksel iletkenlik örneği 0 – 25 cm derinlikten Jenway 4510 iletkenlik ölçer aleti ile 2014 yılında son bahar mevsiminde toplanmıştır. Batı Urmie Playası bölgesinde olduğu gibi kurak dönemin sonu örnek toplamak için tercih edilmiştir. Toprak örnekleri Dr. Nikou Hamzeshpour tarafından sağlanmıştır.

2.2 Kullanılan Veriler

2.2.1 Uydu görüntüleri

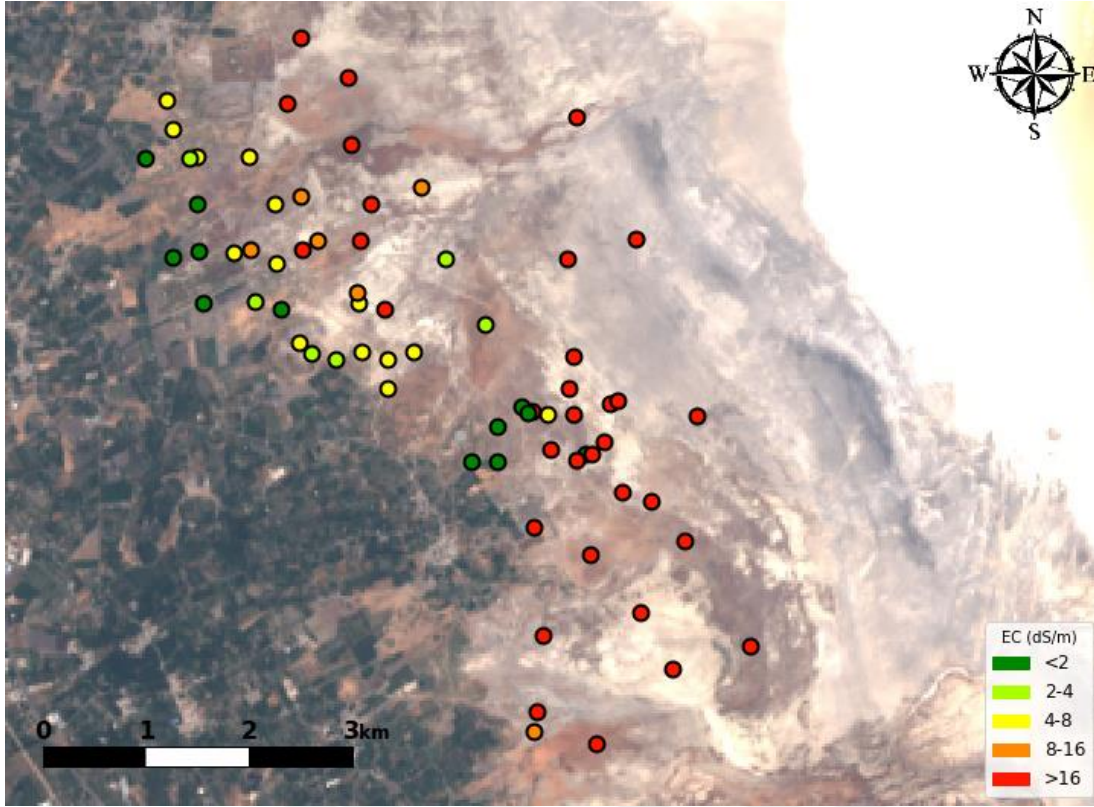
Çalışma kapsamında Bonab bölgesi için 15 Eylül 2014 tarihli Landsat 8 uydu görüntüsü kullanılırken, Batı Urmiye Playası için ise 3 Ekim 2018 tarihli uydu görüntüsü kullanılmıştır. Landsat 8 uydusu 705 km yükseklikteki yörüngesini 99 dakika da tamamlamaktadır. Zamansal çözünürlüğü ise 16 gündür. Landsat 8 uydu görüntüsünün içerdiği bantlar Çizelge 2.1’de gösterilmiştir. Landsat 8 uydusunun koleksiyon 2 seviye 1 görüntüleri bu çalışma kapsamında kullanılmıştır. Bu seviyedeki görüntüler atmosfer üstü yansıtım değerlerini (top of atmosphere, TOA) ifade etmektedir. Uydu görüntüleri geometrik ve radyometrik olarak düzeltilmiş uydu seviyesinde yansıtım değerleri hesaplanmıştır (Url-2). Landsat 8 koleksiyon 2 seviye 2 görüntüleri ise atmosferik düzeltme yapılmış olarak sunulmaktadır. Ancak bu çalışma kapsamında atmosferik düzeltmenin neden olduğu eksi yansıtım değerlerinin göl ve etrafında fazlaca olması nedeniyle bu seviye kullanılmamıştır. Bu problem Landsat 8 için bilinen problemler arasında yer almaktadır (Url-3).

Çizelge 2.1 : Landsat 8 uydusunun bantlarının spektral ve mekânsal özellikleri.

Bant Numarası	Bant İsmi	Dalga Boyu (µm)	Mekânsal Çözünürlük
1	Kıyı Aerosolü	0.43 – 0.45	30 m
2	Mavi	0.45 – 0.51	30 m
3	Yeşil	0.53 – 0.59	30 m
4	Kırmızı	0.64 – 0.67	30 m
5	Yakın Kızılötesi	0.85 – 0.88	30 m
6	SWIR 1	1.57 – 1.65	30 m
7	SWIR 2	2.11 – 2.29	30 m
8	Pankromatik	0.50 – 0.68	15 m
9	Sirrus	1.39 – 1.38	30 m
10	TIRS 1	10.6 – 11.19	100 m
11	TIRS 2	11.5 – 12.51	100 m

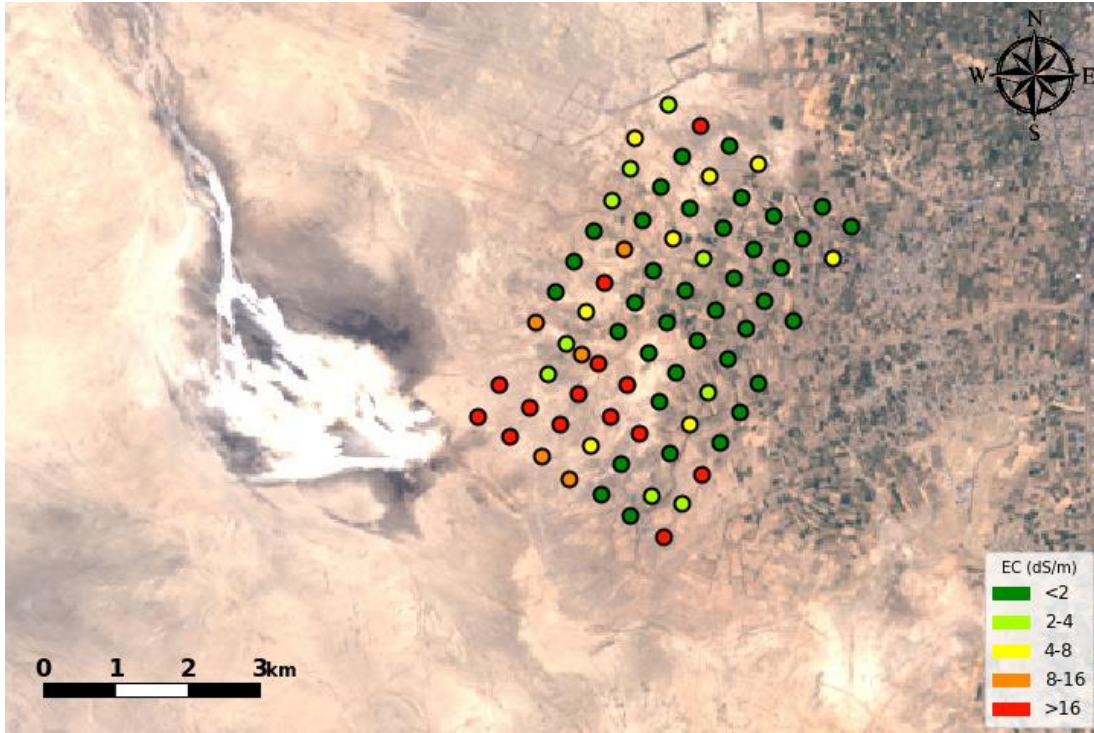
2.2.2 Toprak örnekleri

Çalışma kapsamında iki bölge için farklı iki yöntem ile toprak elektriksel iletkenliği değerleri ölçülmüştür. Laboratuvar da yapılan ölçümlerin daha doğru sonuçlar verdiği bilinmektedir. Toprak örneklerinin dağılımı Batı Urmiye Playası için Şekil 2.2’de gösterilmiştir. Örneklerin göle doğru elektriksel iletkenliği değerinin arttığı görülürken tarım arazilerine doğru ise daha düşük değerlere sahip olduğu tespit edilmiştir.



Şekil 2.2 : Batı Urmieh Playası toprak örneklerinin mekansal dağılımı.

Bonab bölgesi için ise Batı Urmieh Playasından farklı olarak örnekler bir grid ağı şeklinde düzenli olarak toplanmıştır (Şekil 2.3). Ancak çalışma alanında bulunan tuz



Şekil 2.3 : Bonab bölgesi toprak örneklerinin mekansal dağılımı.

yapılarını da haritalandırmak için çalışma alanı daha geniş tutulmuştur. Örneklerin dağılımında Batı Urmiye Playasına benzer olarak göle doğru yaklaşıldıkça tuzluluk seviyelerin arttığı tarım alanlarına doğru ise azaldığı görülmüştür. Toprak örnekleri tuzluluğa göre 5 ayrı kategoride incelenmiştir. İlk olarak elektriksel iletkenliği değeri 0 – 2 dS/m arasında olan toprak örnekleri tuzsuz topraklar olarak ifade edilirken, ikinci seviyede ise 2 – 4 dS/m arasında olan topraklar az tuzlu toprak, üçüncü seviyede 4 – 8 dS/m arasındaki topraklar orta derecede tuzlu toprak, dördüncü seviyede ise 8 – 16 dS/m arasında iletkenliğe sahip olan topraklar yüksek derece tuzlu toprak olarak tanımlanırken, son olarak, 16 dS/m üzeri elektriksel iletkenliğine sahip topraklar aşırı derecede tuzlu topraklar olarak ifade edilmektedir (Brown ve diğ., 1954).





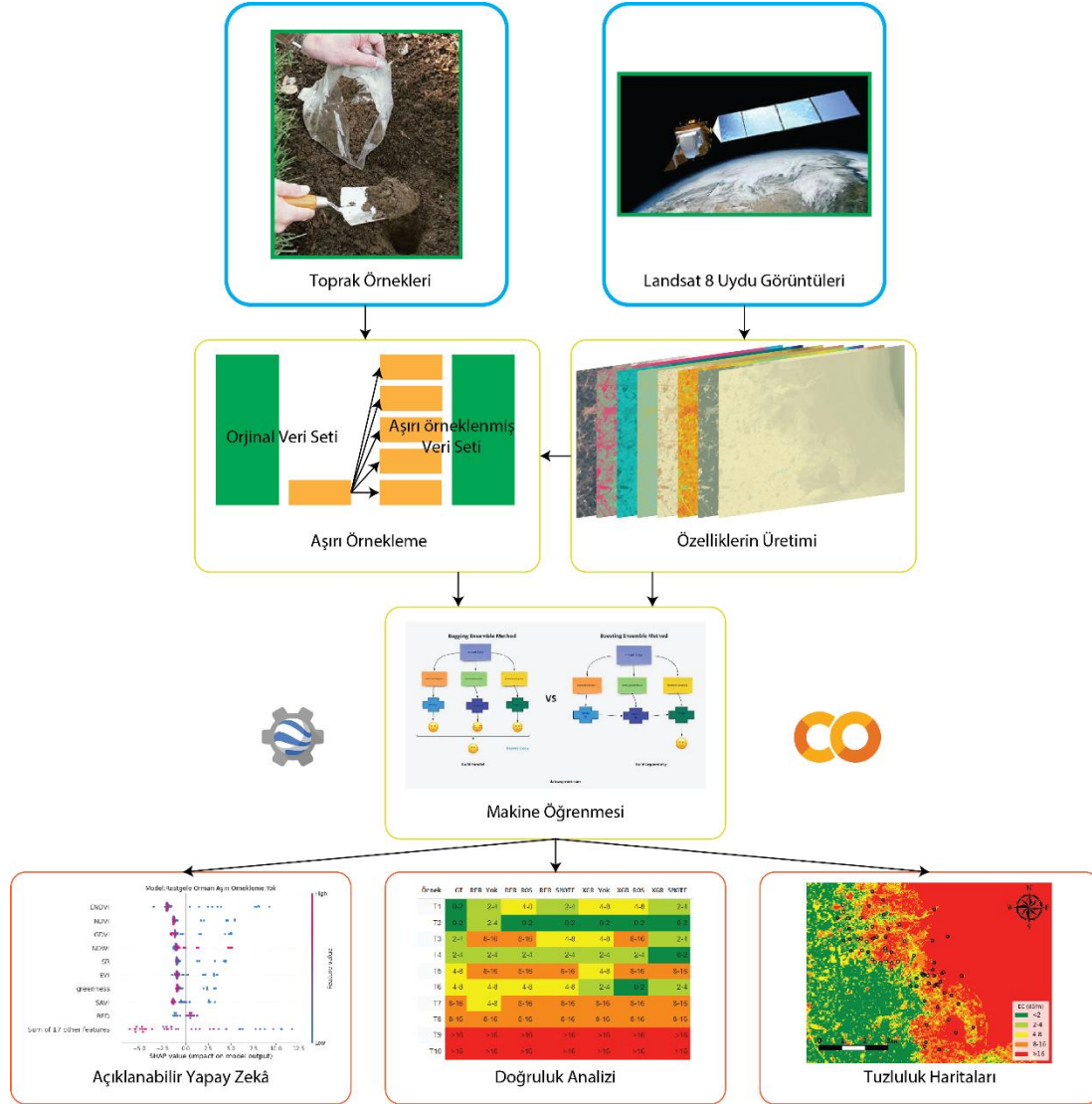
3. YÖNTEM

Bu çalışmada öncelikle çalışma bölgeleri ve kullanılan veriler başlığı altında tanıtılan toprak örneklerinin toplanma tarihi ile uyumlu Landsat 8 uydu görüntüleri temin edilmiştir. Uydu görüntüleri radyometrik ve geometrik olarak düzeltilmiş olduğu için herhangi bir ön işleme adımına ihtiyaç duyulmamıştır. Daha sonra uydu görüntüleri ve toprak örnekleri arasında bir model oluşturabilmek için özellikler üretilmiştir. Üretilen özellikler makine öğrenme algoritmaları için girdi olurken, elektriksel iletkenliği değeri ise modelin çıktısı olarak kullanılmıştır. Eğitilen modeller ile haritalar üretilmesinin yanı sıra bu modeller açıklanabilir yapay zekâ yöntemleriyle açıklanmıştır. Tüm bu yöntemlerin iş akışı ise Şekil 3.1’de gösterilmiştir.

3.1 Google Earth Engine

Google Earth Engine (GEE) platformu, içerdiği petabaytlarca analize hazır mekânsal veri ve paralel hesaplama yetenekleri ile kullanıcılarına hızlıca mekânsal analizler yapabilmesini, bu analizleri görselleştirebilmesini sağlar. GEE bulut tabanlı bir sistem olarak sunulduğu için kullanıcılara herhangi bir hesaplama gücü olmadan da yüksek performans gerektiren analizleri yapabilmeye olanak sunar. Üst seviye programlama bilgisi olmadan da yüksek performanslı hesaplama araçlarını kolaylıkla kullanılabilir olarak sunulur. GEE verilerinin yanında birçok mekânsal algorithmada platform çerçevesinde sunulmaktadır. GEE fonksiyonel programlama yöntemini temel alarak analizler için boru hattı oluşturmamıza olanak sağlar. Analiz algoritmaları fonksiyonel programlama gereği içerisine aldıkları verileri değiştirmez düzenlenmiş kopyalarını verir. Oluşturulan akışlar ise yönlendirilmiş döngüsüz grafik (directed acyclic graph) halini alır. GEE oluşturulan bu grafik analizlerinde tembel değerlendirme yöntemini kullanarak sadece ihtiyaç duyulan kısımları hesaplar. Veri analizinde mekânsal veriler için her zaman bir ölçek girdi olarak verilir ki analiz sonuçları bu mekânsal ölçeğe göre işlenebilsin. Ayrıca GEE birden fazla kişinin aynı analiz üzerinde çalışmasına olanak sağlamaktadır (Gorelick ve diğ., 2017). Bu yönüyle de ekip çalışmalarına oldukça uygundur. GEE son yıllarda birçok araştırma için kullanılmıştır. Bu

araştırmaların hatırı sayılır bir kısmı yapay zekâ çalışmalarını içermektedir. Özellikle tarım arazilerinde ürün desen haritaların oluşturulması, arazi örtüsü / arazi kullanım haritalarının oluşturulması, ormanlık alanlardaki değişimlerin incelenmesi, sulak alanların tespiti gibi konulardır (Yang ve diğ., 2022). Bu çeşitli sorunlarda sunduğu çözüm ve olanaklar düşünüldüğünde bu çalışma kapsamında GEE kullanımının faydalı olduğu görülmüştür.



Şekil 3.1 : Tez çalışmasının iş akış diyagramı

3.2 Makine Öğrenmesi için Özellikler

Literatürde, makine öğrenmesine girdi olarak sunulan veriler için iki farklı terminoloji kullanılmaktadır. Bunlardan bir tanesi Türkçeye çevresel ortak değişken olarak çevrilmiş olan environmental co-variate; diğeri ise özellik olarak çevrilen feature terimidir. Bu tez kapsamında, özellik teriminin kullanılması tercih edilmiştir; fakat

literatür taraması bölümünde yazarların tercih ettiği terim kullanılmıştır. Bu çalışma kapsamında birçok özellik toprak tuzluluğunun tahmini için kullanılmıştır. İlk olarak 1941 yılında Jenny tarafından ortaya atılan toprak gelişimi için mekanik model, bu çalışmada özelliklerin seçiminde önemli bir rol almıştır. Model kısaca SCORPAN formülü olarak bilinir. Bu formül toprağın parametrelerini, iklimini, bitki örtüsünü topoğrafyasını, litolojisini, yaşını ve mekânsal özelliklerini modellemek için kullanılır (Hateffard ve diğ., 2019; McBratney ve diğ., 2003). Bu çerçevede çalışmada kullanılmak üzere toprak, bitki ve mekânsal özellikleri belirlemek için Çizelge 3.1’de gösterilen özellikler üretilmiştir. Urmiye gölü etrafındaki çalışma alanlarında topoğrafya fazla değişmediği için bu formül kapsamında üretilen özelliklere eklenmemiştir (Aksoy ve diğ., 2022). F. Wang ve diğ. (2019) SCORPAN formülünü kullanarak çok kapsamlı özellikler tanımlamıştır. Ancak bu özellikler her çalışma alanı için iyi sonuçlar vermemektedir. Aksoy ve diğ. (2022) Urmiye Gölü ve çevresi için bu özelliklerinden 26 tane seçerek başarılı sonuçlar elde etmişlerdir. Seçilmiş olan bu özellikler bu tez kapsamında da kullanılmıştır.

Bu çalışmada, Landsat 8 uydu görüntülerine tasseled cap dönüşümü yapılarak elde edilen üç bant, orjinal spektral bantlar, farklı bitki örtüsü ve tuzluluk indisleri özellik olarak tercih edilmiştir. Yedi adet tuzluluk indisi bu çalışma kapsamında kullanılmak üzere seçilmiştir. Bu indislerden bir kısmı bu bölgede iyi sonuç verdiği bilinen önceki çalışmalardan seçilmiştir (Aksoy ve diğ., 2022; Gorji ve diğ., 2020). Uydu görüntülerinin orijinal bantları da çalışma kapsamında kullanılmıştır. Seçilen bantlar ise mavi, yeşil, kırmızı, yakın kızıl ötesi (Near infrared, NIR), kısa dalga kızıl ötesi (short wave near infrared, SWIR) bir ve kısa dalga kızıl ötesi iki bantları seçilmiştir (Aksoy ve diğ., 2022). Tasseled cap dönüşüm ile üretilen parlaklık, yeşillik ve ıslaklık bilgileri kullanılmıştır. Bunlara ek olarak 10 adet indiste modellenmek üzere seçilmiştir. Toplamda yirmi altı adet özellik bu çalışma kapsamında kullanılmıştır. Tüm özellikler, en küçük ve en büyük değerleri kullanılarak normalize edilmiştir. Normalize edilen bu özellik görüntüleri 3 x 3 boyutunda alçak geçirgenli bir filtre kullanılarak yumuşatılmıştır (Aksoy ve diğ., 2022; Gopalakrishnan ve Kumar, 2021).

3.3 Aşırı Örnekleme

Makine öğrenme algoritmalarında veri setlerinin her bir sınıf için eşit miktarda veya yeterli sayıda örnek içermediği durumlar oldukça sıktır. Verideki bu sınıf dengesizliği

(class unbalance) çoğunluk olan sınıfların baskın olmasına neden olabilmektedir. Bu dengesiz veri dağılımını düzeltmek için örnekleme yöntemleri kullanılabilmektedir. Aşırı örnekleme yöntemleri ise azınlık olan sınıftaki verilerin artırılarak çoğunluk olan sınıfla aynı dağılıma getirmeyi amaçlamaktadırlar. Bu yöntemlerden rastgele aşırı örnekleme (random over sampling, ROS) ve sentetik azınlık aşırı örnekleme tekniği (Synthetic Minority Oversampling Technique, SMOTE) uzaktan algılama çalışmalarında sıklıkla kullanılmaktadır (Fonseca ve diğ., 2021).

Rastgele aşırı örnekleme yöntemi azınlık olan sınıftaki verilerin rastgele kopyalanması ile çoğunluk sınıftaki örnek sayısına erişilmesi ile yapılır (Hounkpatin ve diğ., 2018). Kolay bir yöntem olduğu için tercih sebebidir; ancak bu yöntem makine öğrenme algoritmalarında aşırı öğrenmesine (over fitting) neden olduğu ve aşırı örnekleme yapılmadan eğitilen algoritmalarla göre daha kötü sonuçlar verdiği belirlenmiştir (Fonseca ve diğ., 2021).

SMOTE yöntemi Chawla ve diğ. (2002) tarafından literatüre kazandırılmıştır. SMOTE yöntemi rastgele olarak azınlık olan örnekleri çoğaltmak yerine en yakın komşu yöntemi ile belirlediği komşu örnekler arasında birim vektörler oluşturur. Daha sonra 0 ile 1 arasında seçtiği rastgele bir sayı ile bu vektör çarpılır ve özellik uzayında yeni bir noktayı ifade eder. Bu yeni nokta sentetik olarak üretilmiş olur (Chawla ve diğ., 2002).

Bu çalışma kapsamında toprak tuzluluğu örnekleri daha önce de bahsedildiği gibi (tuzsuz, az tuzlu, orta derece tuzlu, yüksek derece tuzlu ve aşırı tuzlu olmak üzere) 5 sınıfa ayrılmıştır. Bu sınıflara düşen örneklere yukarıda belirtilen özellikler de eklenerek rastgele aşırı örnekleme ve SMOTE aşırı örnekleme yöntemleri uygulanmıştır. İki yöntemden elde edilen yeni örnek dağılımı ve orijinal verilerin olduğu örnek dağılımı ile makine öğrenme algoritmaları eğitilmiş ve analizlere devam edilmiştir.

Çizelge 3.1 : Toprak tuzluluğunu modellemek için kullanılan özellikler.

Veriler	İndis	Kısaltma	Formül	Referans
Landsat 8 OLI	Mavi, Yeşil, Kırmızı, NIR, SWIR1, SWIR2			
	Tasseled Cap	TC1,2,3		(Baig ve diğ., 2014)
Bitki İndisleri	Normalized Difference Vegetation Index	NDVI	$(NIR - Kırmızı) / (NIR + Kırmızı)$	(Rouse ve diğ., 1973)
	Soil Adjusted Vegetation Index	SAVI	$[(NIR - Kırmızı) * (1 + L)] / (NIR + Kırmızı + L); L = 0.5$	(A. R. Huete, 1988)
	Enhanced Vegetation Index	EVI	$[2.5 * (NIR - Kırmızı)] / [NIR + 6 * Kırmızı - 7.5 * Mavi]$	(A. Huete ve diğ., 2002)
	Generalized Difference Vegetation Index	GDVI	$(NIR^2 - Kırmızı^2) / (NIR^2 + Kırmızı^2)$	(Wu, 2014)
	Canopy Response Salinity Index	CRSI	$[(NIR * Kırmızı) - (Yeşil * Mavi)] / [(NIR * Kırmızı) + (Yeşil * Mavi)]^{0.5}$	(Scudiero ve diğ., 2014)
	Simple Ratio vegetation index	SR	$NIR / Kırmızı$	(Jordan, 1969)
	Two-band enhanced vegetation index	EVI2	$[2.5 * (NIR - Kırmızı)] / (NIR + 2.4 * Kırmızı - 1)$	(Jiang ve diğ., 2008)
	Extended NDVI	ENDVI	$(NIR + SWIR2 - Kırmızı) / (NIR + SWIR2 + Kırmızı)$	(H. Chen ve diğ., 2019)
	Extended EVI	EEVI	$2.5 * (NIR + SWIR1 - Kırmızı) / [NIR + 2.5 * (SWIR1 + 6 * NIR + 7.5 * SWIR1) * Mavi + 1]$	(H. Chen ve diğ., 2015)
	Normalized Difference Water Index	NDWI	$(Yeşil - NIR) / (Yeşil + NIR)$	(Mcfeeters, 1996)
Tuzluluk İndisleri	Tuzluluk İndisi 1	SI1	$(Mavi * Kırmızı) / Yeşil$	(Abbas ve Khan, 2007)
	Tuzluluk İndisi 2	SI2	$(Mavi * Kırmızı)^{0.5}$	(Khan ve diğ., 2005)

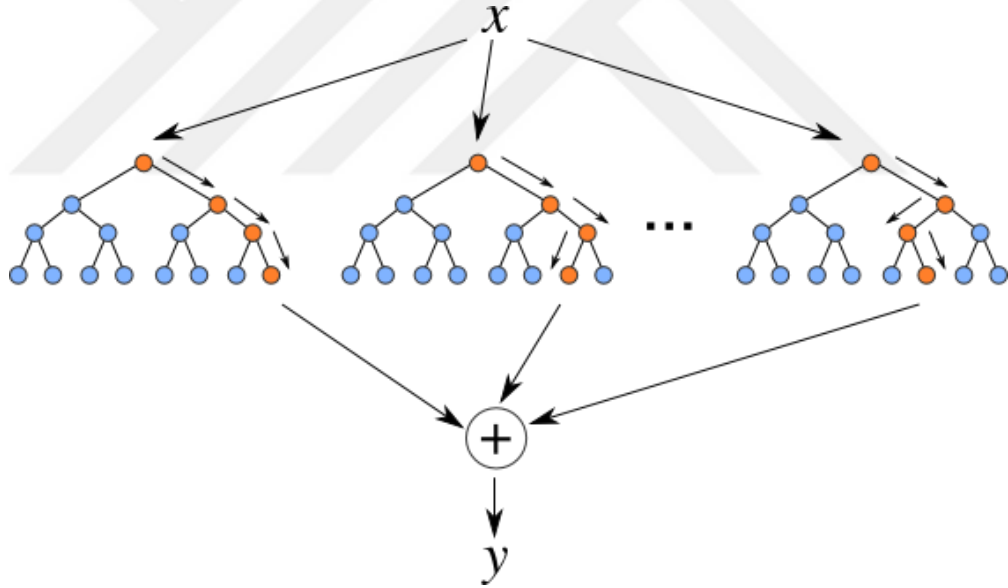
Çizelge 3.1 (devam): Toprak tuzluluğunu modellemek için kullanılan özellikler.

Veriler	İndis	Kısaltma	Formül	Referans
	Tuzluluk İndisi 3	SI3	$(\text{Yeşil} + \text{Kırmızı})/2$	(Douaoui ve diğ., 2006)
	Tuzluluk İndisi 4	SI4	$(\text{Yeşil} * \text{Kırmızı})^{0.5}$	(Khan ve diğ., 2005)
	Tuzluluk İndisi 5	SI5	Mavi / Kırmızı	(Abbas ve Khan, 2007)
	Tuzluluk İndisi 6	SI6	$(\text{Kırmızı} * \text{NIR}) / \text{Yeşil}$	(Abbas ve Khan, 2007)
	Tuzluluk İndisi 7	SI7	$2 * \text{Yeşil} - 5 * (\text{Kırmızı} + \text{NIR})$	(Jiang ve diğ., 2008)

3.4 Makine Öğrenmesi Yöntemleri

3.4.1 Rastgele orman (random forest)

Rastgele Orman (Random Forest, RFR) makine öğrenme algoritması sınıflandırma ve regresyon ağaçları algoritmasına bağlı bir topluluk öğrenme algoritmasıdır. Rastgele orman algoritması Breiman (2001) tarafından sunulmuştur. Karar ağaçlarını temel alan yapıda birbirleriyle korelasyonu olmayan farklı karar ağaçlarının birleştirilmesi ile oluşturulmuştur. Karar ağaçları tek başlarına zayıf bir makine öğrenmesi metodu olmasına karşın, birleştirildiklerinde güçlü bir makine öğrenme algoritması elde edilebilmektedir (S. Wang ve diğ., 2019). Rastgele orman algoritması sınıflandırma problemlerinin yanı sıra regresyon problemlerine de çözüm sunmaktadır (Li ve diğ., 2019). Rastgele orman algoritması regresyon problemlerini çözerken birden fazla ağacın sonucunun ortalamasını alarak hesaplamaktadır. Bu yöntemin mimarisi Şekil 3.2’de gösterilmiştir.



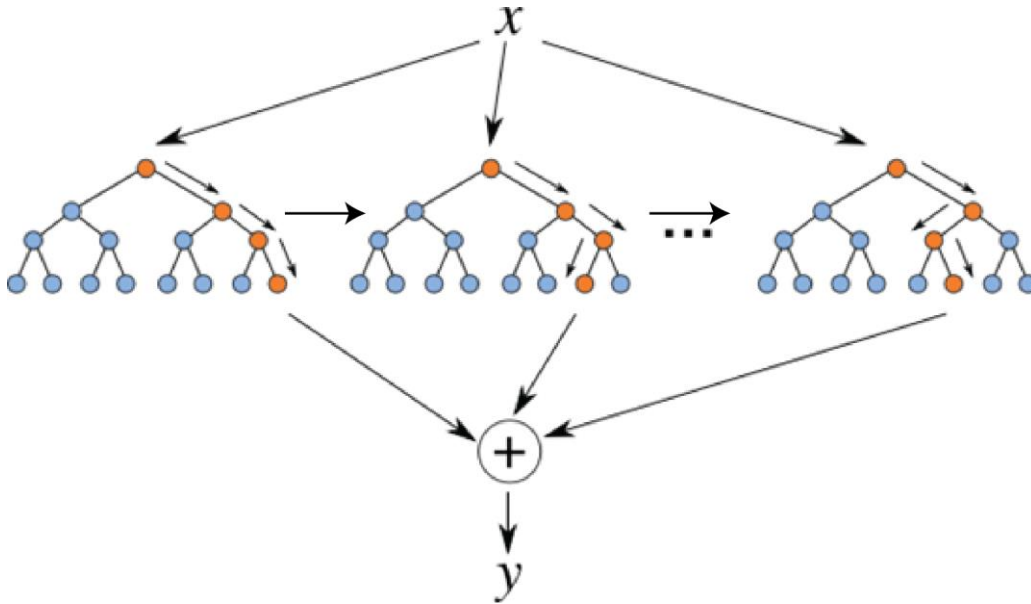
Şekil 3.2 : Rastgele Orman Regresyonu genel mimarisi.

Rastgele orman algoritması regresyon ve sınıflandırma olarak Python programlama dilinde sklearn ve spark kütüphanelerinde kullanılabilir olarak sunulmaktadır (Pedregosa ve diğ., 2011). Bu çalışmada sklearn kütüphanesi kullanılmıştır.

3.4.2 XGBoost (aşırı gradyan arttırma, extreme gradient boosting)

XGBoost (Aşırı Gradyan arttırma, eXtreme Gradient Boosting, XGR) T. Chen ve Guestrin (2016) tarafından sunulmuştur. XGBoost çok hızlı çalışan gradyan arttırma

ağaç algoritmalarından birisidir. Bu hızla gradyan destekli ağaçların önemli bir yetersizliğini gidererek ulaşır. Ağaç için yeni bir ayırım oluştururken olası hataları tüm önceki üye ayırımlar için tahmin edilir. XGboost bu eksikliği ağacın yaprağındaki tüm veri noktalarının dağılımını inceleyerek özellik ayırımını aradığı uzayı daraltır. Çeşitli düzenleme yöntemleri XGBoost içerisinde çalışsa da yöntemin hızlı ve kullanışlı olması nedeniyle bir çok hiper parametre olasılığı hızlıca bulunabilir (Zarei ve diğ., 2021). XGBoost genel mimarisi rastgele orman algoritmasına benzer ancak ağaçlar Rastgele Orman algoritmasındaki gibi paralel olarak bağımsız değildirler. Mimari Şekil 3.3'te gösterilmiştir. Bu çalışmada XGBoost'un Python programlama dili ile T. Chen ve Guestrin (2016) tarafından yazılmış olan kütüphanesi kullanılmaktadır.



Şekil 3.3 : XGBoost Regresyonu genel mimarisi.

3.5 Açıklanabilir Yapay Zeka

Açıklanabilir yapay zeka yöntemleri kara kutu olarak ifade edilen makine öğrenme algoritmalarının davranışlarını açıklamak üzere kurgulanmış algoritmalar (Stadler ve diğ., 2022). Literatürde çeşitli açıklanabilir yapay zeka yöntemleri sunulmuştur. Bunlardan pertürbasyon temelli Local Interpretable Model-agnostic Explanations (LIME) (Ribeiro ve diğ., 2016) ve Shapley Additive Explanations (SHAP) (Lundberg ve Lee, 2017) gibi açıklama yöntemleri ve gradyan tabanlı Gradient times Input (Selvaraju ve diğ., 2020), SmoothGrad (Smilkov ve diğ., 2017), Integrated Gradients (Sundararajan ve diğ., 2017), Gradient-weighted Class Activation Mapping (GradCAM) (Selvaraju ve diğ., 2020) gibi yöntemlerdir. Bu yöntemlerin birbirleriyle

aynı sonuç vermediği sıklıkla karşılaşılan problemler arasındadır. Bu problem karşısında en çok tercih edilen algoritma ise SHAP yöntemidir (Krishna ve diğ., 2022). Bu yüzden tez kapsamında SHAP yöntemi tercih edilmiştir.

3.5.1 SHAP yöntemi

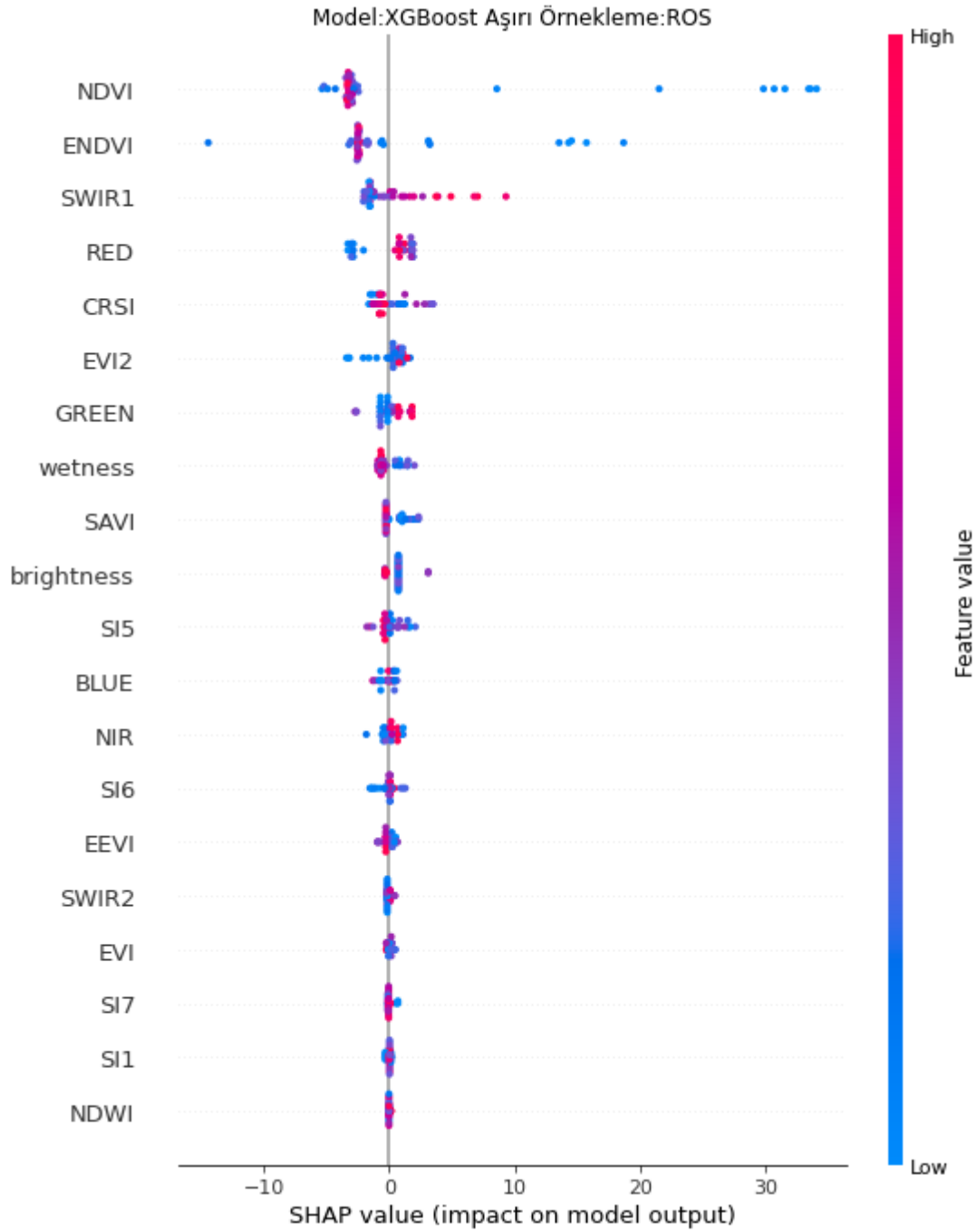
SHAP yöntemi ilk olarak Lundberg ve Lee (2017) tarafından geliştirilmiştir, bu yaklaşım SHAP değerlerini modelden bağımsız olarak sonuçlara bağlı olarak açıklayan bir yapay zeka yöntemidir. SHAP değerleri her bir özelliğin modelin sonuca olan katkısını gösterir. Eğitilen veri seti için modelden beklenen bir çıktı değeri vardır. Bu değere taban değer denir. SHAP değerlerinin modele olan katkısı ise modelin bu taban değerden ne kadar uzaklaştırdığını göstermektedir. Daha fazla katkı sunan özellikler model için önemli olan özellikler olarak kabul edilir. Bu katkılar negatif ve pozitif de olabilir. Mutlak SHAP değerleri özelliklerin önemini belirtirken, tüm sonuçlar için mutlak SHAP değerlerinin ortalaması ise genel önem değerlerini belirtmektedir (Stadtler ve diğ., 2022). SHAP yöntemi ile yapılan analizler genellikle 4 grafik üzerinden sunulmaktadır. Bu grafikler sırasıyla özet grafiği, özellik önem grafiği, bağımlılık grafiği ve kuvvet grafiğidir.

3.5.1.1 SHAP özet grafiği

SHAP özet grafiği modeli girdi olan özelliklerin modelin çıktısına olan etkilerini her bir eğitim örneği için toplandığı bir grafikdir. Girdilerin değerlerine göre renklendirilerek hangi girdinin hangi değerlerinin modele nasıl katkı sunduğunu göstermektedir. Şekil 3.4'te gösterilen örnek grafik incelendiğinde NDVI özelliğinin yüksek değerlerinin modele girdiğinde modelin çıktısını azaltma yönünde bir etki yaparken düşük değerlerinin ise modeli çıktısını yükselttiği göze çarpmaktadır. Burada gösterilen her bir nokta açıklama için kullanılan eğitim verilerinin özelliklerinin modele olan etkisini ve özelliğin değerlerini içermektedir.

3.5.1.2 SHAP özellik önem grafiği

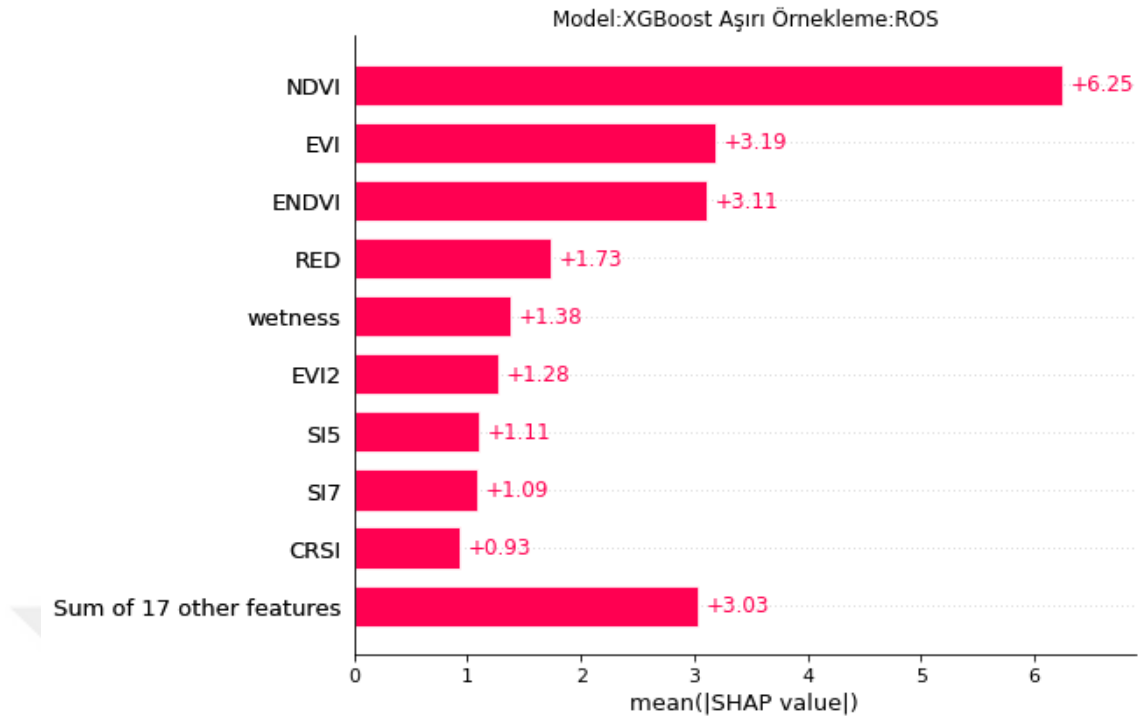
SHAP özellik önem grafiği ise özet grafiğinde gösterilen her bir verinin katkısı yerine ortak mutlak katkıları bir çubuk grafiği olarak gösterilmektedir. Bu grafik modele olan toplam katkıyı öne çıkarmaktadır. Modele hangi yönde katkı verdiği göz ardı edilmiştir. Şekil 3.5'te gösterilen örnek grafik incelendiğinde NDVI özelliğinin modele en yüksek katkıyı verdiği görülürken bunu EVI özelliği takip etmektedir.



Şekil 3.4 : Örnek SHAP özet grafiği.

3.5.1.3 SHAP bağımlılık grafiği

Bağımlılık grafiği ise iki özellik arasındaki ilişkiyi model üzerinden açıklamaktadır. Bir özelliğin modele olan katkısını hem o özelliğin değerleri üzerinden hem de o özellik ile bağımlı olan başka bir özelliğin değeri üzerinden açıklamaktadır. Şekil 3.7’de gösterilen örnek grafik incelendiğinde NDVI özelliğinin düşük değerlerinin modele yüksek katkı sunarken, SWIR2 özelliğinin de NDVI düştükçe değerlerin yükseldiğini bu grafik üzerinden anlaşılabilir.



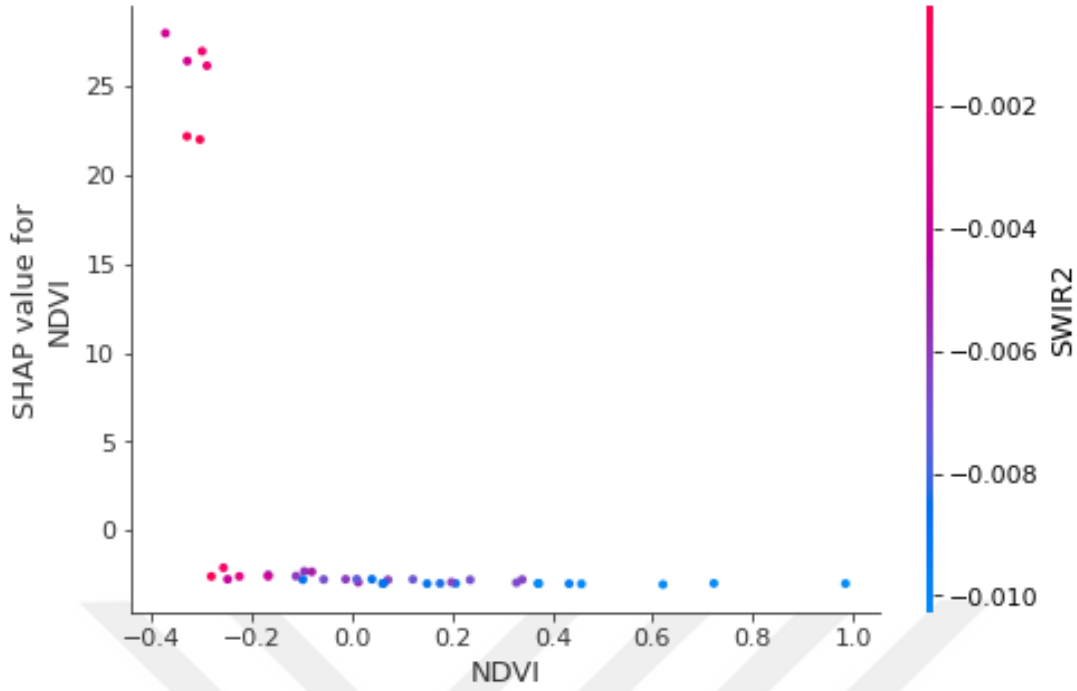
Şekil 3.5 : Örnek SHAP özellik önem grafiği.

3.5.1.4 SHAP kuvvet grafiği

Kuvvet grafiği ise her bir özelliğin tek bir örnek için modele yaptıkları katkıları gösteren bir grafikdir. Regresyon modeli için tahmin edilen taban değere, seçilen örneğin özelliklerinin modele olan katkılarını ve toplamı sonunda oluşan model çıktısını göstermektedir. Şekil 3.6’de gösterilen örnek grafikte modelin taban değeri 14 dS/m iken, EVI2 -25 değeri ile taban değere pozitif dS/m eklerken, EVI özelliği 12.86 değeri ile negatif dS/m olacak şekilde katkı sunmuştur. Bu özelliklerin tüm katkıları toplandığında ise modelin tahmini olan 2.28 dS/m değerine ulaşılmıştır. Özelliklerin katkılarına göre boyutlar belirlenmiş, mavi renkler negatif katkıyı ve kırmızı renkler ise pozitif katkıları göstermektedir.



Şekil 3.6 : Örnek SHAP kuvvet grafiği.



Şekil 3.7 : Örnek SHAP bağımlılık grafiği.

3.6 Doğruluk Analizi

Makine öğrenme modellerinin doğruluğunu test etmek için iki çalışma alanında da eğitim verilerinin dışında doğruluk analizi için test örnekleri ayrılmıştır. Test örnekleri üzerinden doğruluğu hesaplayabilmek için karesel ortalama hata (root mean square error, RMSE) ve determinasyon katsayısı (coefficient of determination, R^2) kullanılmıştır. Karesel ortalama hata ve determinasyon katsayısı denklem 3.1 ve 3.2 de gösterilmiştir.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - Y_i)^2} \quad (3.1)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (X_i - Y_i)^2}{\sum_{i=1}^n (X_i - \bar{X})^2} \quad (3.2)$$

Denklemden gösterilen n örnek sayısını, X_i ölçülen toprak tuzluluk değerini, Y_i tahmin edilen toprak tuzluluk değerini ve $\bar{X}$ ortalama ölçülen tuzluluk değerini ifade etmektedir (Hateffard ve diğ., 2019; F. Wang ve diğ., 2019).

4. BULGULAR VE TARTIŞMA

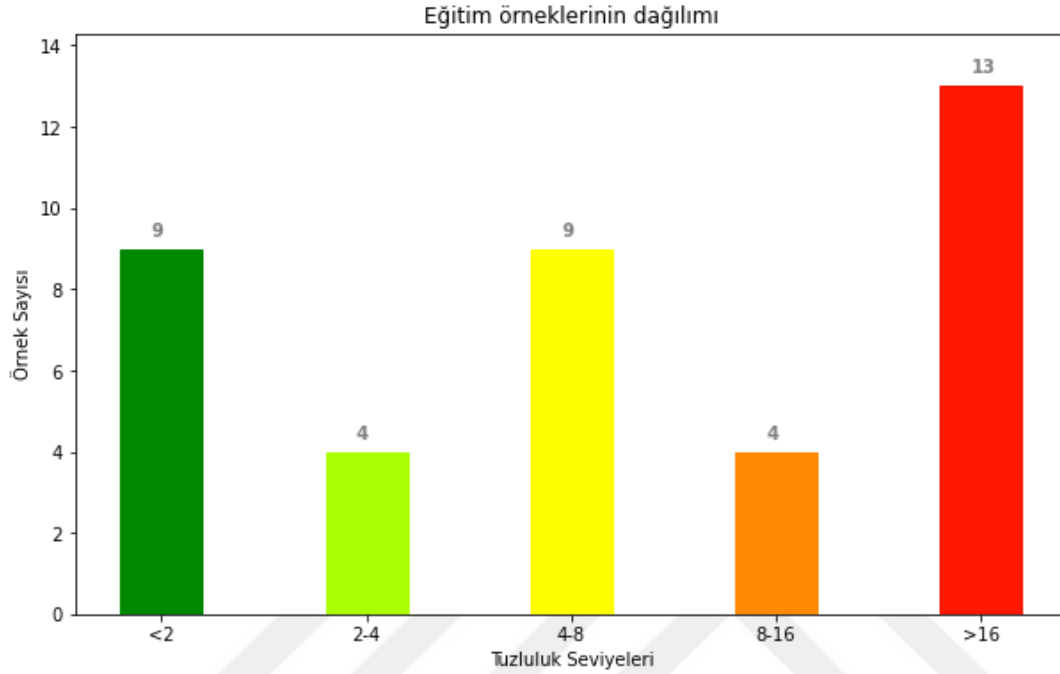
Toprak tuzluluğu haritaları 6 adet senaryo için üretilmiştir. Bu senaryo kombinasyonları iki adet makine öğrenme algoritmasının üç adet aşırı örnekleme çeşidi ile birleşiminden üretilmiştir. Makine öğrenme algoritmaları olarak rastgele orman ve XGBoost kullanılırken, aşırı örnekleme yöntemleri için ise rastgele aşırı örnekleme, SMOTE ve aşırı örnekleme olmayan şeklinde denemeler yapılmıştır. Özetle aşağıdaki model ve veri kombinasyonları oluşturularak testler gerçekleştirilmiştir.

- 1. Test: Rastgele Orman modeli ve orjinal veriler,
- 2. Test: Rastgele Orman modeli ve rastgele aşırı örnekleme ile oluşturulan veriler,
- 3. Test: Rastgele Orman modeli ve SMOTE aşırı örnekleme ile oluşturulan veriler,
- 4. Test: XGBoost modeli ve orjinal veriler,
- 5. Test: XGBoost modeli ve rastgele aşırı örnekleme ile oluşturulan veriler,
- 6. Test: XGBoost modeli ve SMOTE aşırı örnekleme ile oluşturulan veriler.

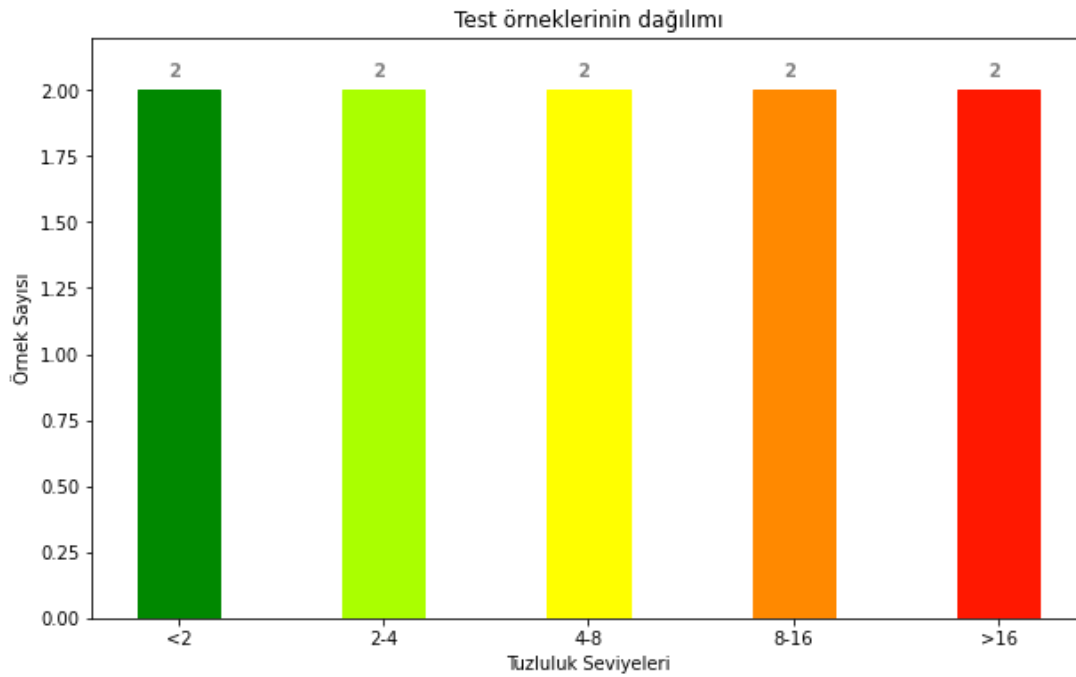
Modeller kullanılarak üretilen toprak tuzluluğu haritalarının görsel olarak incelenmesine ek olarak modellerin eğitim ve test örnekleri üzerinde performansları değerlendirilmiştir. Modellere ait eğitim ve test örneklerinin tuzluluk seviyelerinin dağılımı Şekil 4.1, Şekil 4.2, Şekil 4.3 ve Şekil 4.4'te gösterilmiştir. Aşırı örnekleme yöntemleri ile düzenlenen dağılımlar ise Şekil 4.5 ve 4.6'da her iki bölge için gösterilmiştir. Testler sonucunda üretilen modelleri doğrulamak ve açıklamak için dağılım grafiklerinin yanı sıra SHAP yöntemi kullanılarak özet, önem, bağımlılık ve kuvvet grafikleri üretilmiştir.

Batı Urmiye Playası eğitim verileri üzerinde modellerin performansını aşırı örnekleme yöntemlerine göre incelemek için Şekil 4.7'de gösterilen grafikler üretilmiştir. XGBoost modelinin eğitim veri setini aşırı öğrendiği göze çarpmaktadır. Rastgele

orman algoritması ise aşırı örnekleme yöntemlerinin kullanılması ile veri seti üzerinde daha fazla öğrendiği göze çarpmaktadır. Daha önce de bahsedildiği gibi rastgele aşırı örnekleme yöntemi rastgele orman ve XGBoost için aşırı öğrenmeye neden olduğu R^2 ve RMSE değerleri üzerinden kolaylıkla anlaşılmaktadır.

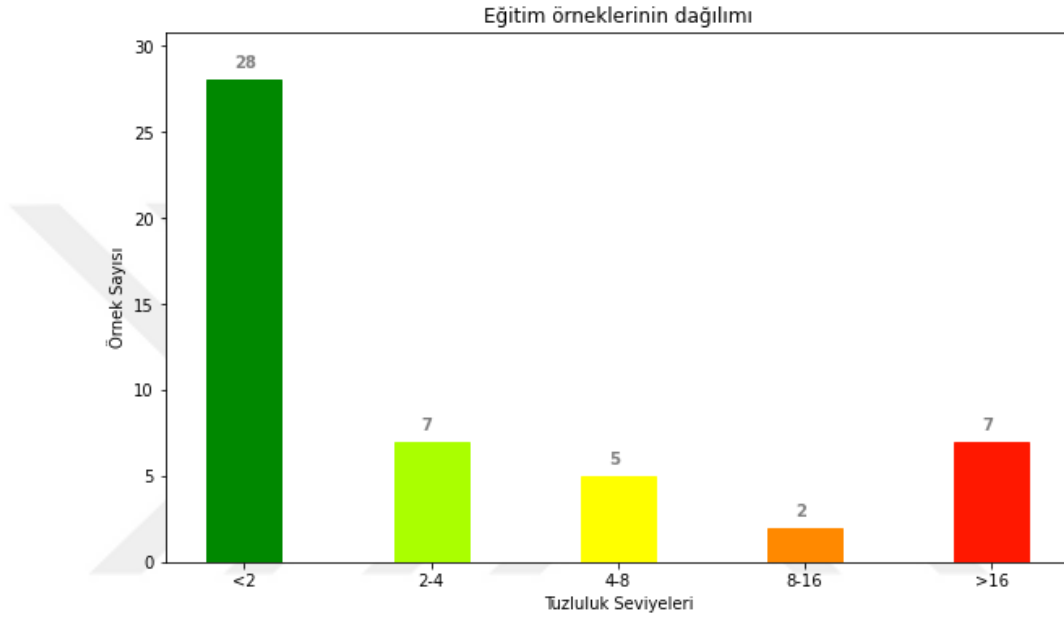


Şekil 4.1 : Batı Urmiye Playası için eğitim örneklerinin tuzluluk seviyelerine göre dağılımı.

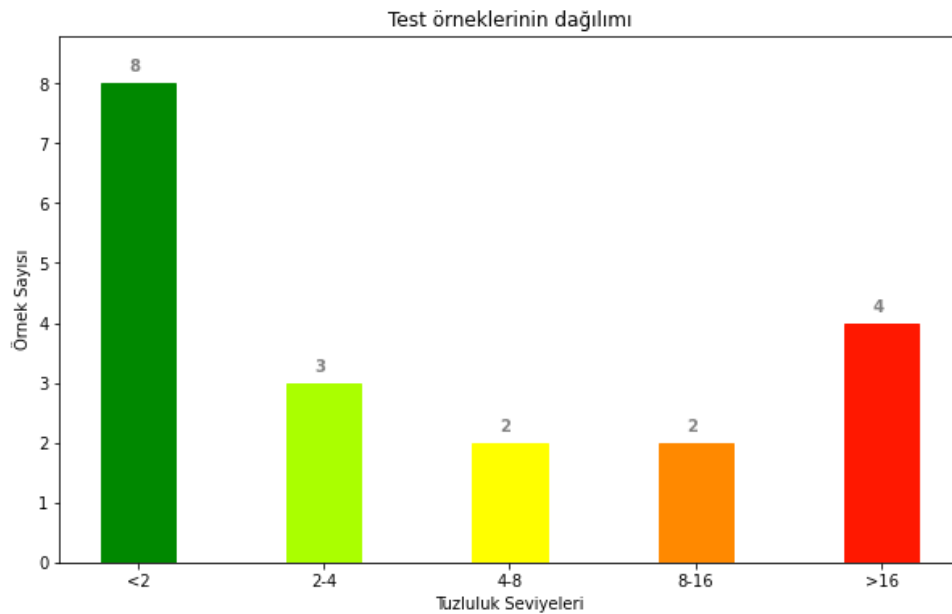


Şekil 4.2 : Batı Urmiye Playası için test örneklerinin tuzluluk seviyelerine göre dağılımı.

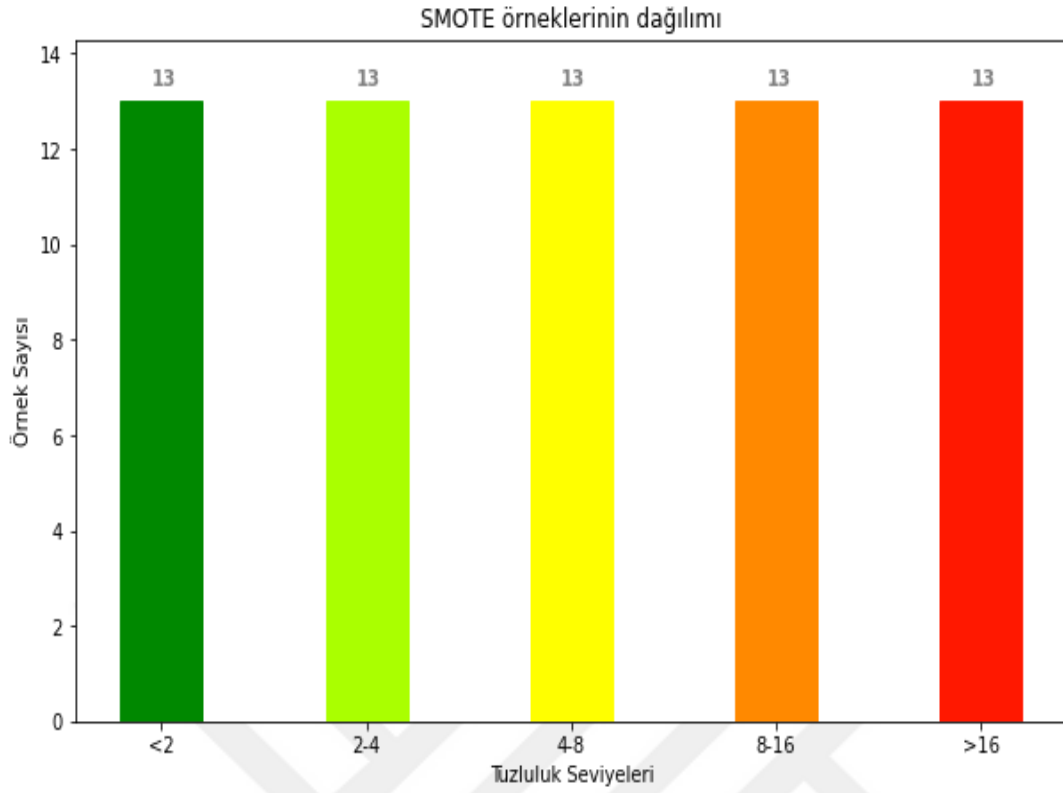
Bonab bölgesi için de Batı Urmiye Playasına benzer grafikler eğitim verileri için üretilmiştir. Grafikler Şekil 4.8’de görülebilir. Grafikler incelendiğinde, yine benzer olarak XGBoost algoritmasının eğitim verilerine rastgele orman algoritmasına göre aşırı öğrendiği göze çarpmaktadır. Rastgele aşırı örnekleme yöntemi yine SMOTE ve aşırı örnekleme olmayan yöntemlere göre aşırı öğrenmeye sebebiyet vermiş. Genel olarak her iki bölgede de modeller eğitim veri setini iyi öğrenmiş olarak gözükmektedirler.



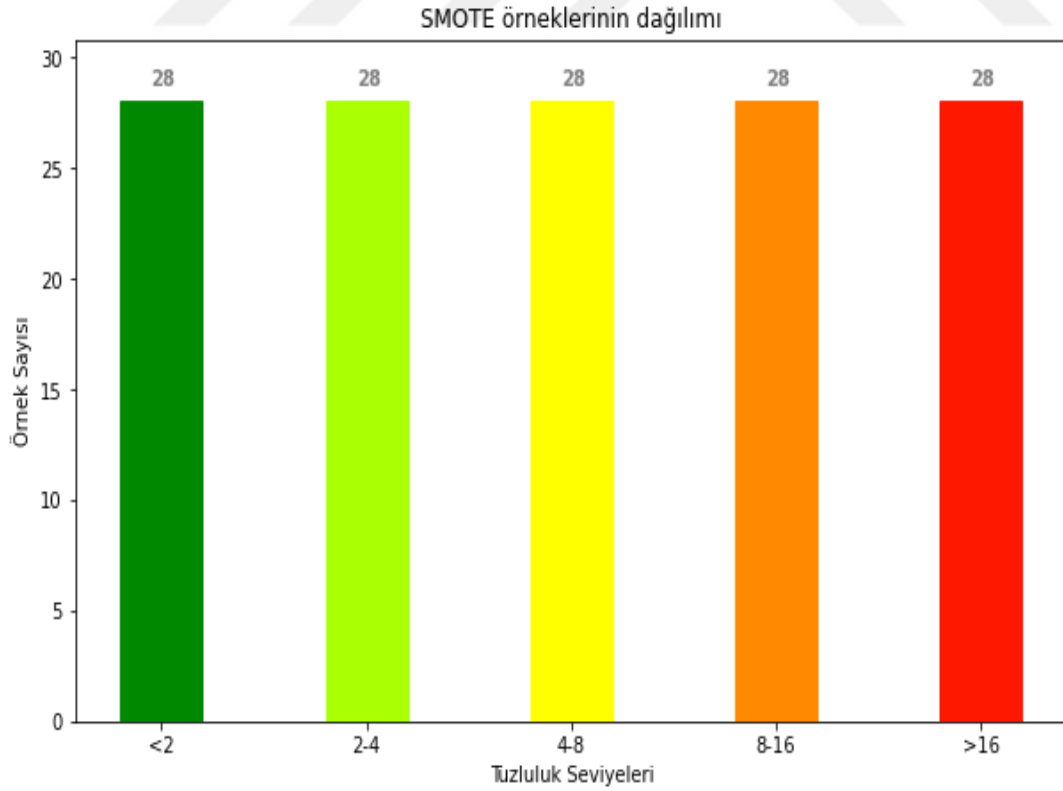
Şekil 4.3 : Bonab bölgesi için eğitim örneklerinin tuzluluk seviyelerine göre dağılımı.



Şekil 4.4 : Bonab bölgesi için test örneklerinin tuzluluk seviyelerine göre dağılımı.

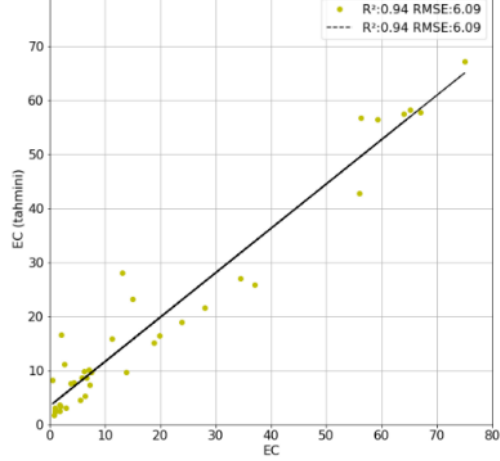


Şekil 4.5 : Batı Urmiye Playası için aşırı örnekleme sonucu eğitim örneklerinin tuzluluk seviyelerine göre dağılımı.

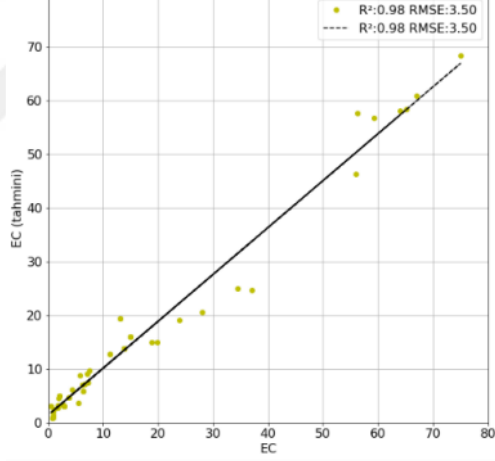


Şekil 4.6 : Bonab bölgesi için aşırı örnekleme sonucu eğitim örneklerinin tuzluluk seviyelerine göre dağılımı.

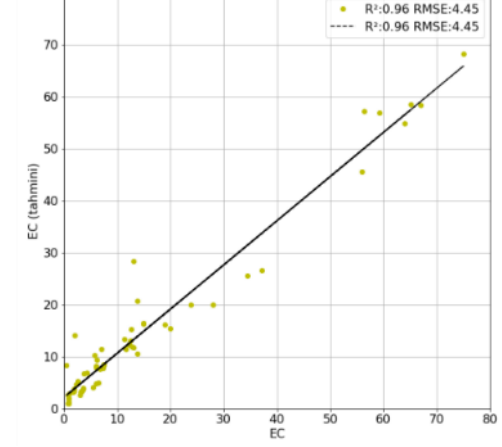
Bölge: West_Urmia_2018 Model: Rastgele Orman Aşırı Örnekleme: Yok



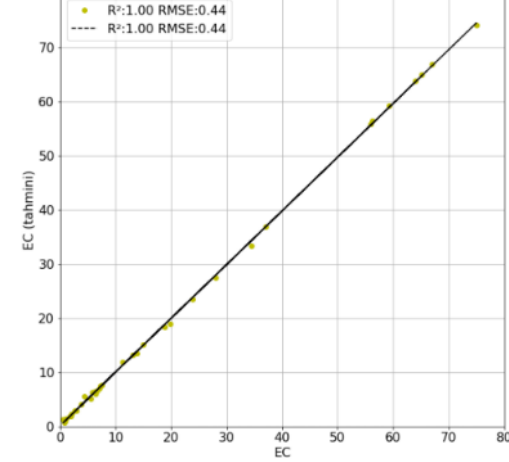
Bölge: West_Urmia_2018 Model: Rastgele Orman Aşırı Örnekleme: ROS



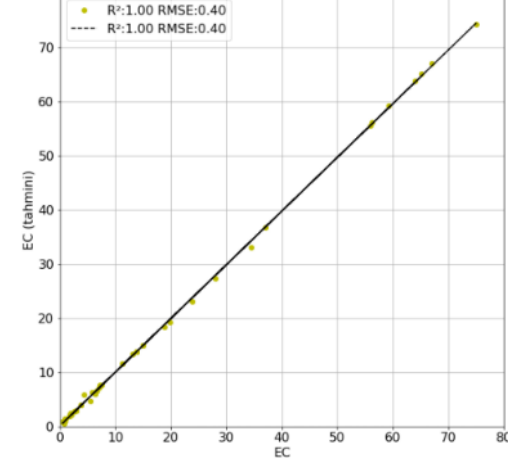
Bölge: West_Urmia_2018 Model: Rastgele Orman Aşırı Örnekleme: SMOTE



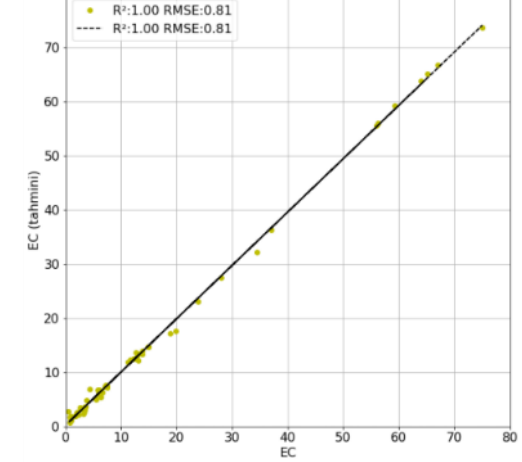
Bölge: West_Urmia_2018 Model: XGBoost Aşırı Örnekleme: Yok



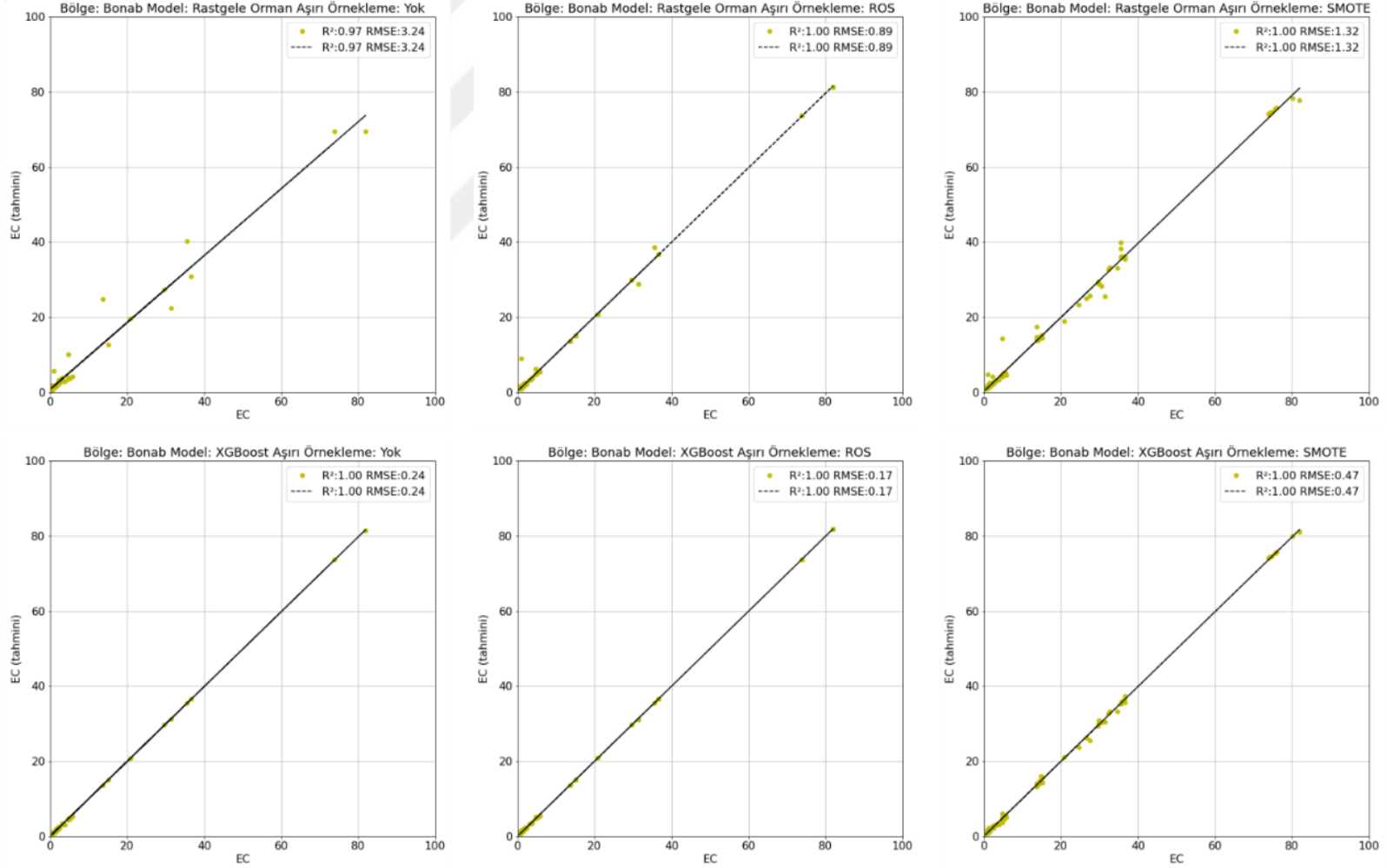
Bölge: West_Urmia_2018 Model: XGBoost Aşırı Örnekleme: ROS



Bölge: West_Urmia_2018 Model: XGBoost Aşırı Örnekleme: SMOTE



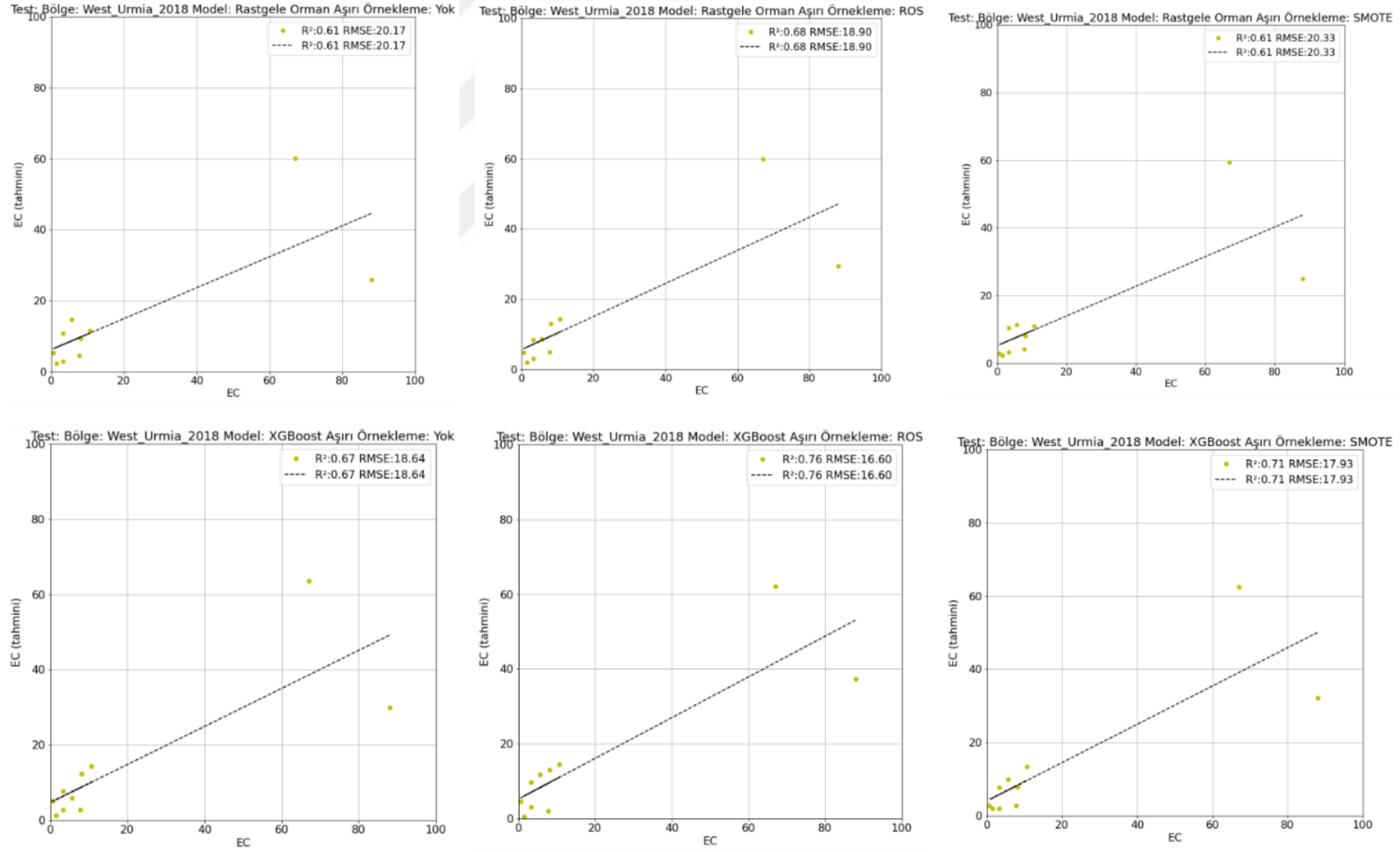
Şekil 4.7 : Batı Urmie Playası için eğitim verileri üzerinde modellerin dağılım grafiği



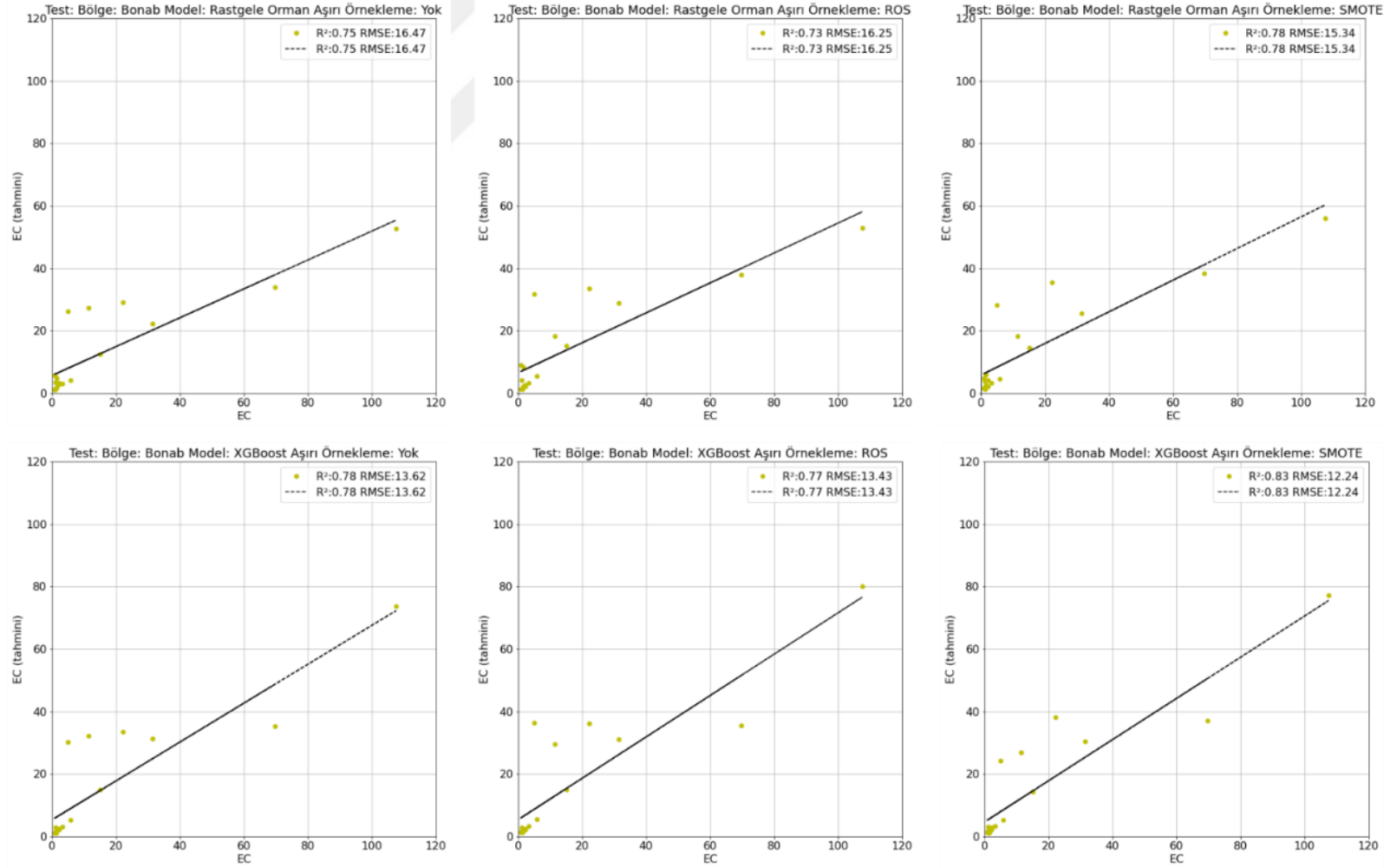
Şekil 4.8 : Bonab bölgesi için eğitim verileri üzerinde modellerin dağılım grafiği.

Batı Urmiye Playası için 10 adet test noktası seçildi. Test örnekleri kullanılarak eğitilen modellerin doğrulukları incelenmiştir. Doğruluk analizine ait grafikler Şekil 4.9'te gösterilmiştir. Grafikler incelendiğinde, XGBoost modelinin Rastgele Orman algoritmasına göre daha iyi sonuçlar verdiği görülmüştür. XGBoost modeli R^2 değeri 0.06 ila 0.10 arasında daha iyi sonuç vermiştir. Rastgele aşırı örnekleme metodu ise modelin doğruluğunu en çok arttıran yöntem olmuştur. SMOTE yöntemi de doğruluğa katkı sunmuştur. Test örnkTest örnekleri için modeller ile tahmin edilen örneklerin dağılım grafiği Şekil 4.8'da görülebilmektedir.

Bonab bölgesi için ise 19 test noktası seçilmiştir. Test noktaları Batı Urmiye Playası gibi her tuzluluk seviyesinden eşit olarak alınmadı, onun yerine Eğitim örneklerinin dağılımına benzer bir oran seçilmiştir. Doğruluk analizine ait üretilen grafikler ise Şekil 4.10'da gösterilmiştir. Grafikler incelendiğinde, XGBoost algoritması diğerine göre daha yüksek doğruluk vermiştir. Rastgele aşırı örnekleme yöntemi Batı Urmiye Playası bölgesinin aksine doğruluğu bu bölgede düşürmüştür. SMOTE aşırı örnekleme yöntemi ve XGBoost bu bölgede en yüksek doğruluğu sunmuştur. R^2 değeri 0.83 ve RMSE değeri 12.24 ile bu model en iyi sonucu vermiştir.



Şekil 4.9 : Batı Urmie Playası için test verileri üzerinde modellerin dağılım grafiği.



Şekil 4.10 : Bonab bölgesi için test verileri üzerinde modellerin dağılım grafiği.

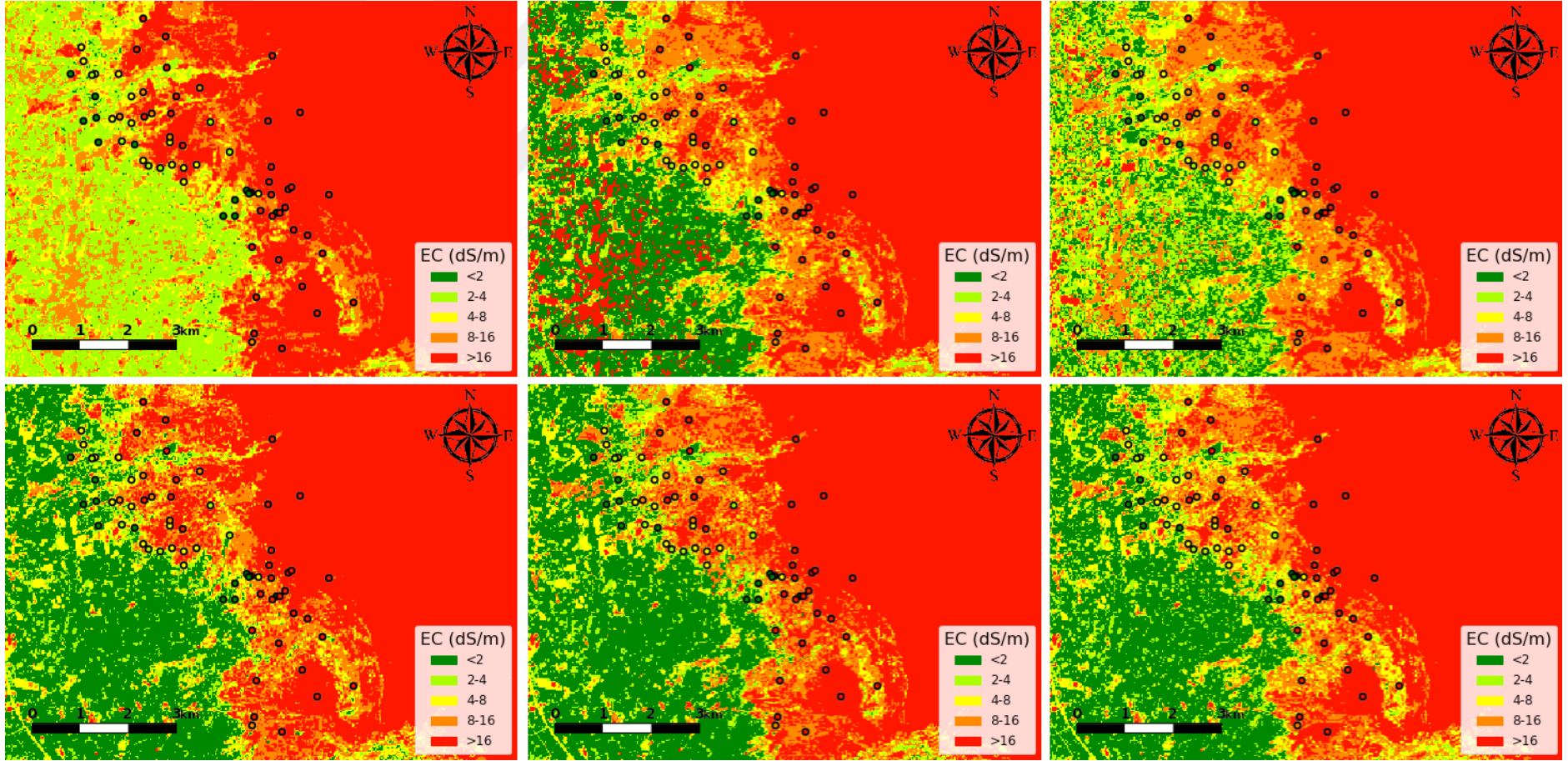
Toprak tuzluluğu haritaları Batı Urmiye Playası için Şekil 4.11’de gösterilmiştir. Her iki makine öğrenme modeli de göle yakın olan alanları aşırı tuzlu bulmayı başarmıştır. Ancak rastgele orman algoritması gölden uzakta olan tarım alanlarını yüksek tuzlu olarak belirlemiştir. Aslında bu alanlar düşük veya tuzlu olmayan alanlar olmasına karşın ekili olmadıkları için regresyon analizi sonucu yanlış olarak tahmin edilmişlerdir. SMOTE aşırı örnekleme yönteminin bu alanlarda hatalı regresyonu daha da kötü etkilediği her iki model incelendiğinde de görülmektedir. Bu alanda yapılan önceki çalışmalar göz önüne alındığında (Aksoy ve diğ., 2022) XGBoost algoritmasının daha doğru olduğu gözlemlenmektedir. Görsel olarak tespit edilen bu doğruluk istatistiksel analizlere de yansımaktadır. Çizelge 4.1’de gösterildiği üzere R^2 değerleri XGBoost için rastgele orman algoritmasına göre daha yüksek iken, RMSE değerleri daha düşüktür. 10 adet test örneğinin regresyon sonrası düştükleri tuzluluk seviyeleri Çizelge 4.2’de gösterilmiştir. Çizelge incelendiğinde XGBoost algoritmasının aynı veya bir yakın seviye sınıfına düşen örnekler ile daha baskın ve doğru bir sonuç verdiği görülmüştür.

Çizelge 4.1 : Batı Urmiye Playası makine öğrenme modellerinin doğruluk sonuçları.

Makine Öğrenme Modeli	Aşırı Örnekleme Yöntemi	R^2	RMSE
RFR	Yok	0.61	20.17
RFR	ROS	0.68	18.90
RFR	SMOTE	0.61	20.33
XGR	Yok	0.67	18.64
XGR	ROS	0.76	16.60
XGR	SMOTE	0.71	17.93

Çizelge 4.2 : Batı Urmiye Playası test örneklerinin tuzluluk seviyeleri incelenmesi.

Örnek	GT	RFR Yok	RFR ROS	RFR SMOTE	XGR Yok	XGR ROS	XGR SMOTE
1	0-2	4-8	4-8	2-4	4-8	4-8	2-4
2	0-2	2-4	0-2	2-4	0-2	0-2	0-2
3	2-4	8-16	8-16	8-16	4-8	8-16	4-8
4	2-4	2-4	2-4	2-4	2-4	2-4	0-2
5	4-8	8-16	8-16	8-16	4-8	8-16	8-16
6	4-8	4-8	4-8	4-8	2-4	0-2	2-4
7	8-16	8-16	8-16	8-16	8-16	8-16	4-8
8	8-16	8-16	8-16	8-16	8-16	8-16	8-16
9	>16	>16	>16	>16	>16	>16	>16
10	>16	>16	>16	>16	>16	>16	>16



Şekil 4.11 : Batı Urmıye Playası tuzluluk haritaları.

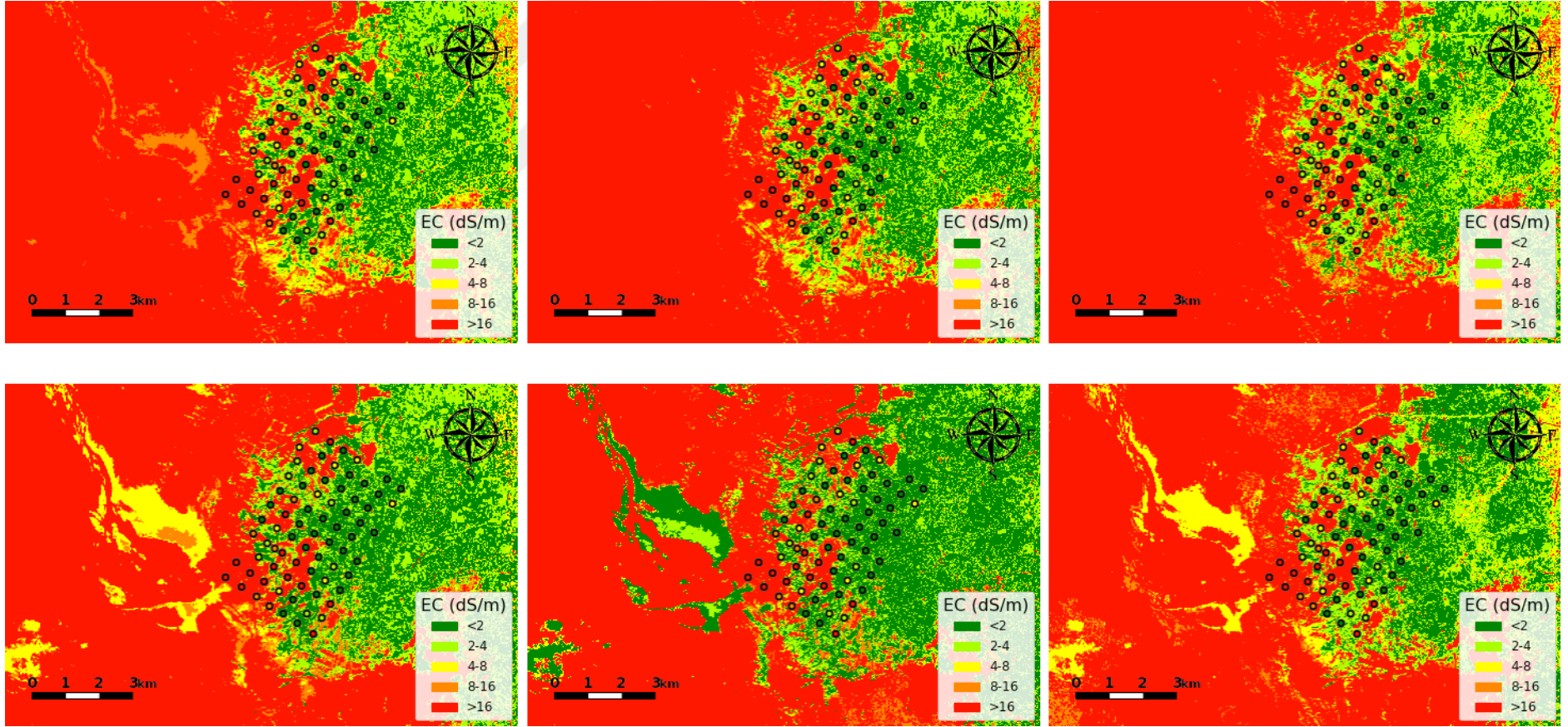
Bonab bölgesi için üretilen tuzluluk haritaları Şekil 4.12’de gösterilmiştir. Rastgele orman algoritması göle yakın olan tuz yapılarını aşırı tuzlu olarak belirlemeyi başarırken XGBoost bu alanı az tuzlu olarak belirlemiştir. Rastgele aşırı örnekleme metodunda ise tuzsuz olarak bulmuştur. Tarım alanlarında toprak tuzluluğunun düşük olduğu bilinmekte, bu yüzden tüm modellerin bu alanları az tuzlu veya hiç tuz olmayan şeklinde belirlemesi önemli bir bulgudur. Çizelge 4.3’te gösterilen, Bonab bölgesi için doğruluk sonuçları incelendiğinde ise XGBoost algoritmasının rastgele orman algoritmasına göre daha iyi sonuç verdiği görülmektedir. SMOTE aşırı örnekleme yöntemi ile XGBoost en iyi sonucu vermiştir. Tuz yapıları dışındaki alanlarda görsel analiz bu bulgular ile tutarlıdır. Çizelge 4.4’te gösterilen test örneklerini modellere göre incelendiğinde, XGBoost algoritmasının tuzlu olmayan örneklerde daha yüksek başarımlar gösterdiği göze çarpmaktadır.

Çizelge 4.3 : Bonab bölgesi makine öğrenme modellerinin doğruluk sonuçları.

Makine Öğrenme Modeli	Aşırı Örnekleme Yöntemi	R ²	RMSE
RFR	Yok	0.75	16.47
RFR	ROS	0.73	16.25
RFR	SMOTE	0.78	15.34
XGR	Yok	0.78	13.62
XGR	ROS	0.77	13.43
XGR	SMOTE	0.83	12.24

Çizelge 4.4 : Bonab bölgesi test örneklerinin tuzluluk seviyeleri incelenmesi.

Örnek	GT	RFR Yok	RFR ROS	RFR SMOTE	XGR Yok	XGR ROS	XGR SMOTE
1	0-2	0-2	2-4	2-4	0-2	0-2	0-2
2	0-2	0-2	0-2	2-4	0-2	0-2	2-4
3	0-2	4-8	8-16	4-8	0-2	0-2	0-2
4	2-4	2-4	2-4	2-4	2-4	2-4	2-4
5	2-4	2-4	2-4	4-8	2-4	2-4	2-4
6	2-4	2-4	2-4	2-4	2-4	2-4	2-4
7	4-8	4-8	4-8	4-8	4-8	4-8	4-8
8	4-8	>16	>16	>16	>16	>16	>16
9	8-16	>16	>16	>16	>16	>16	>16
10	8-16	8-16	8-16	8-16	8-16	8-16	8-16
11	>16	>16	>16	>16	>16	>16	>16
12	>16	>16	>16	>16	>16	>16	>16
13	>16	>16	>16	>16	>16	>16	>16



Şekil 4.12 : Bonab bölgesi tuzluluk haritaları.

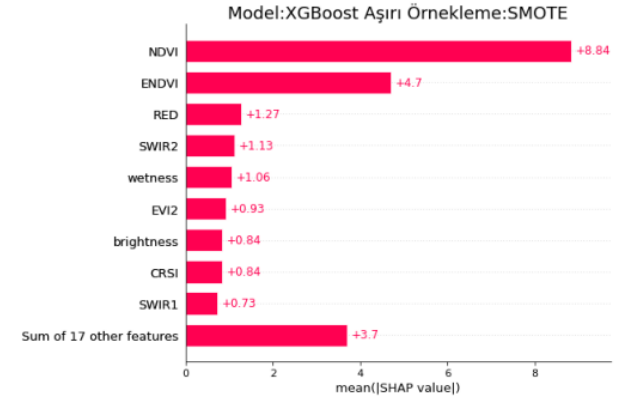
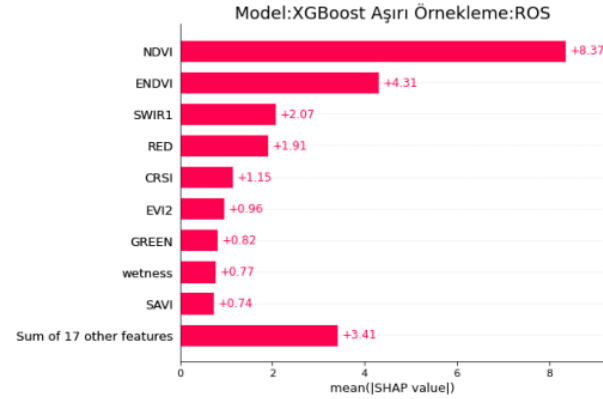
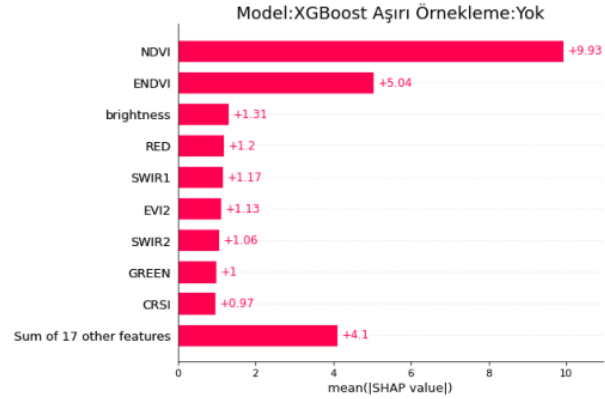
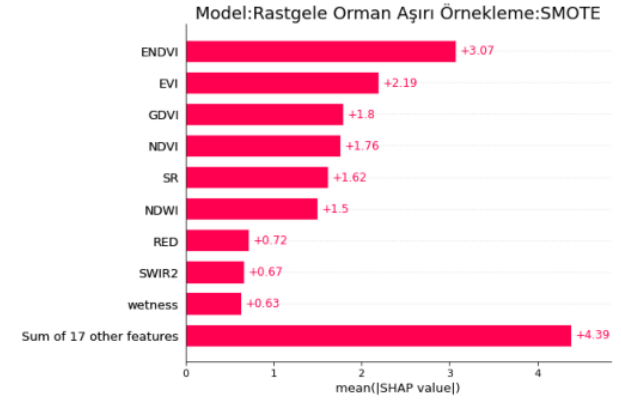
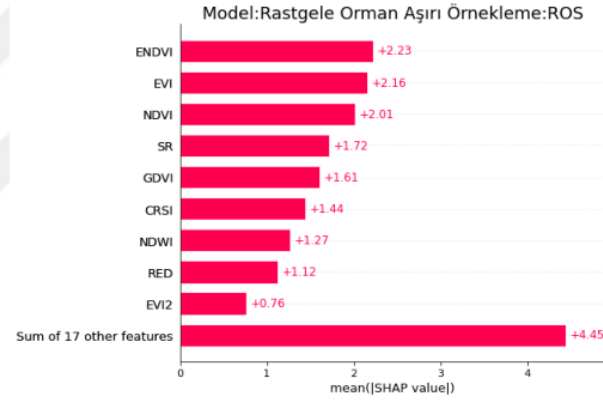
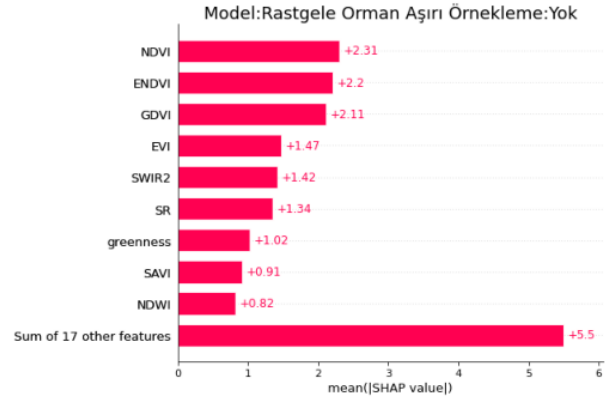
Bu çalışmanın bir diğer aşaması olan açıklanabilir yapay zekâ tarafı ise SHAP kütüphanesi kullanılarak analizler gerçekleştirilmiştir. Yapay zekâ modelleri 4 farklı şekilde açıklanmıştır. İlk olarak özelliklerin her birinin modelin sonucuna olan mutlak katkısı incelenmiştir. İkinci olarak ise özelliklerin modele olan katkılarının hangi yönde olduğu, hangi değerlerinin modellere nasıl katkılar verdiği incelenmiştir. Üçüncü olarak ise en çok katkı sunan iki özellik arasındaki bağımlılık model üzerinde incelendi. Son olarak, her iki alandan seçilen ve Çizelge 4.2 ve Çizelge 4.4'te gösterilen test örnekleri üzerinde sadece örnek bazında özelliklerin modellere olan katkıları irdelenmiştir.

Batı Urmiye Playası için mutlak SHAP değerleri Şekil 4.13'te gösterilmiştir. Mutlak SHAP değerleri incelendiğinde ENDVI, NDVI, EVI gibi bitki indislerinin model sonucuna daha baskın olarak katkı verdiği görülmektedir. CRSI indisinde önceki çalışmaya (Aksoy ve diğ., 2022) benzer olarak bu alanda yine modele yüksek katkı verdiği görülmektedir.

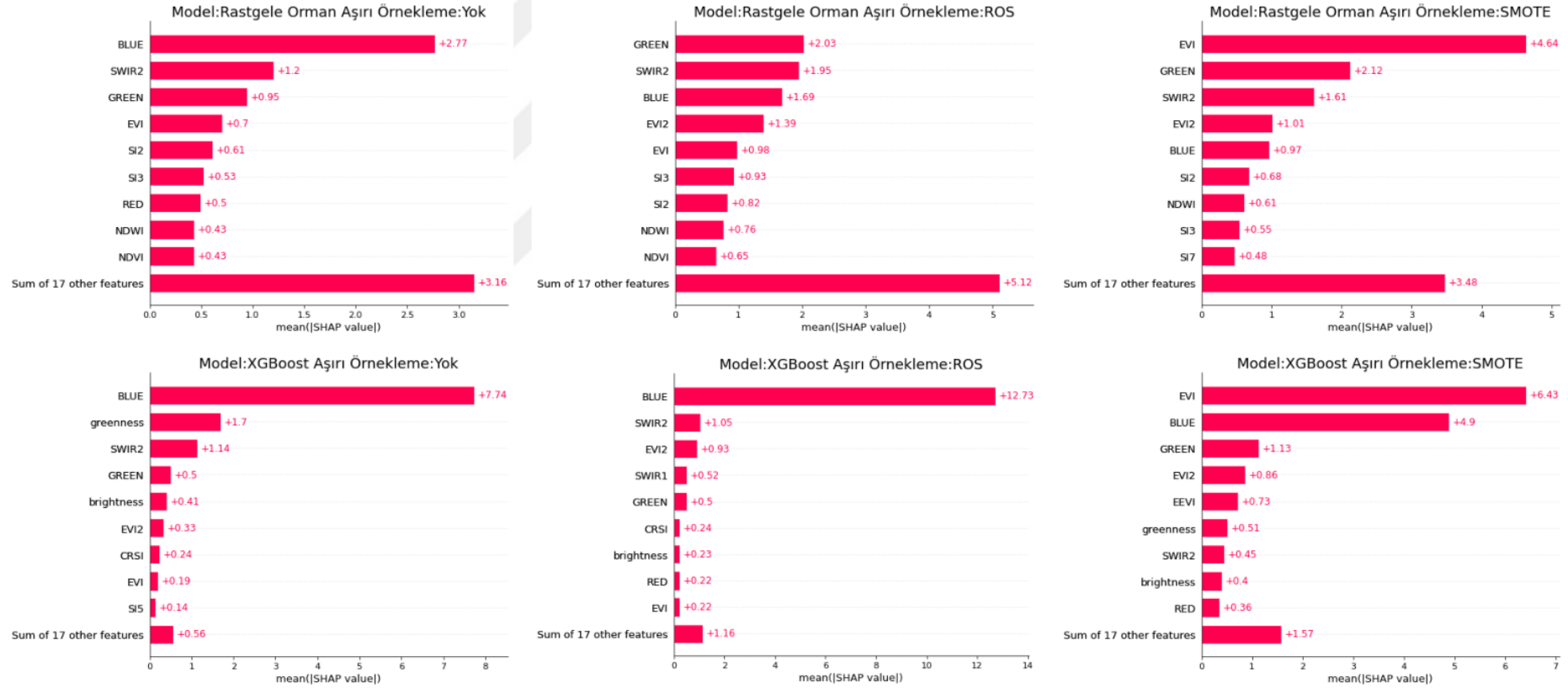
Bonab bölgesi için mutlak SHAP değerleri ise Şekil 4.14'te gösterilmiştir. Bu değerler incelendiğinde uydu görüntüsünün kendi bantlarının daha baskın olduğu özellikle de mavi ve yeşil bantların modele en çok katkı yapan özellikler olduğu görülmektedir. Bitki indisleri Batı Urmiye Playası kadar olmasa da yine modele katkı sunmuşlardır.

SHAP değerleri ile özelliklerin modele olan katkıları pozitif ve negatif olarak görülebilir. Batı Urmiye Playası için SHAP değerlerinin özelliklere göre modele etkisi Şekil 4.15'te gösterilmiştir. Özelliklerin modele olan katkıları incelendiğinde ENDVI, NDVI gibi indislerin yüksek değerleri tuzluluğun az olması yönündeki tahmine katkı sunarken düşük değerleri ise tuzluluğun fazla tahmin edilmesine katkı sunmaktadır. NDWI indisi ise tam tersi katkı sunmuştur. Bu bulgulardan çıkarılacak sonuç ise bitki örtüsü toprak tuzluluğu arttıkça azalmaktadır.

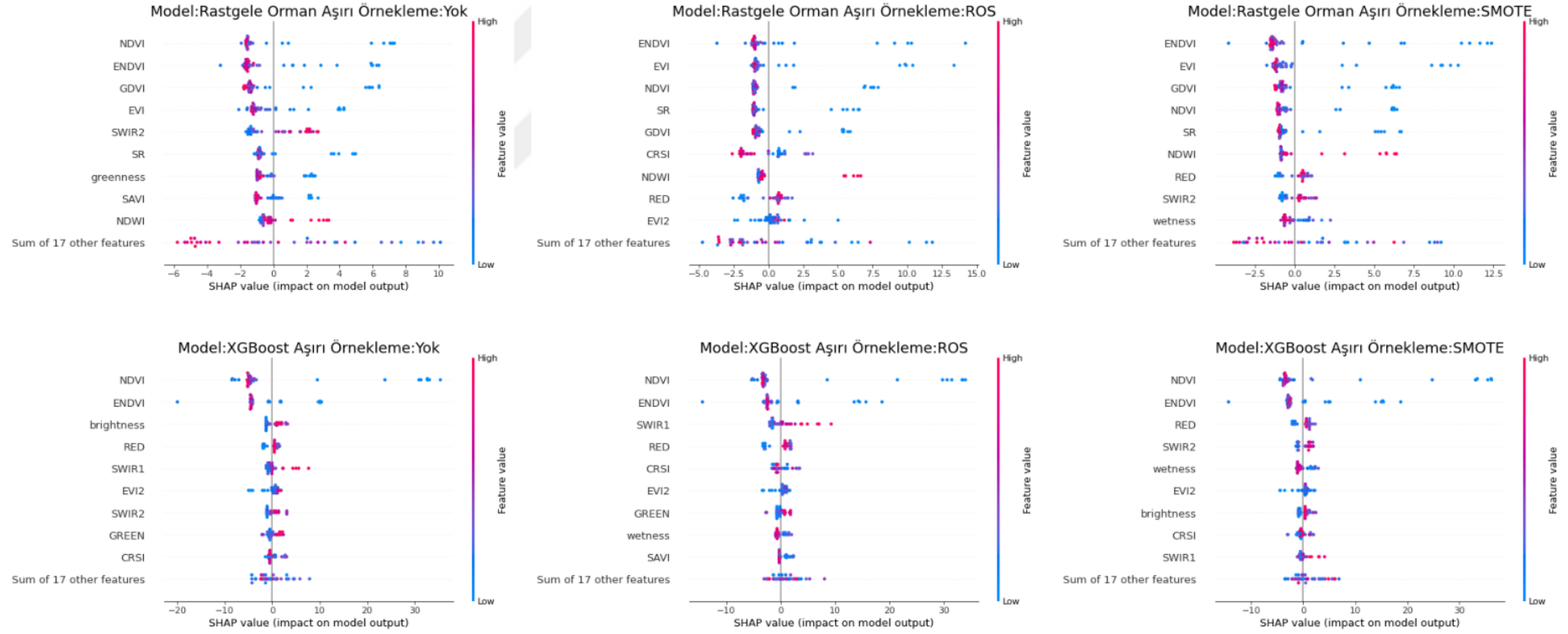
Bonab bölgesi için aynı analizin sonucu Şekil 4.16'da gösterilmiştir. Mavi, yeşil gibi uydu görüntüsünün kendi bantlarının düşük değerleri toprak tuzluluğunun az olmasına katkı sunarken, yüksek değerler toprak tuzluluğunun fazla olarak tahmin edilmesine katkı sunmaktadır. Bu bulgular doğrultusunda tuzlu alanların görünür bölgede daha parlak alanlar olduğu sonucuna ulaşılabilir. Tuz yapıları beyaz ve parlak olmaları ile uydu görüntüleri üzerinde kolaylıkla ayırt edilebilirler.



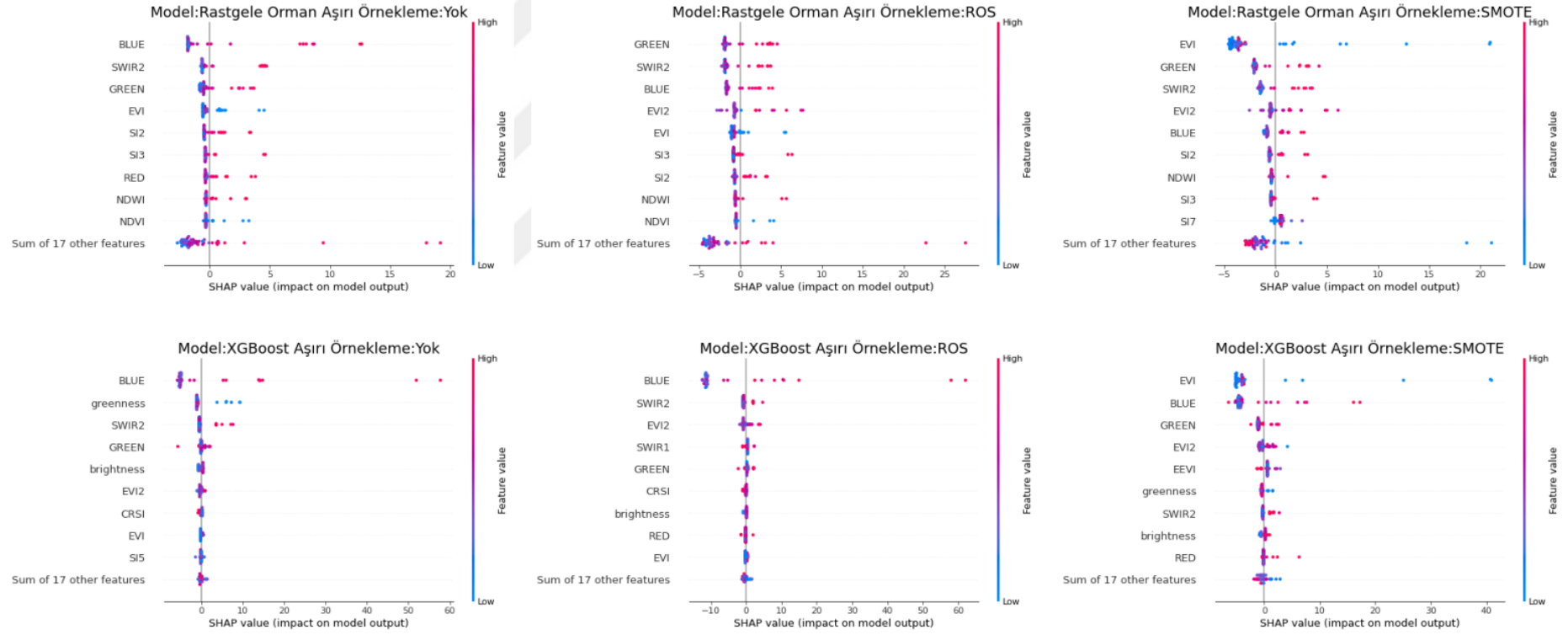
Şekil 4.13 : Batı Urmiye Playası için geliştirilen modellerin mutlak SHAP değerleri.



Şekil 4.14 : Bonab bölgesi için geliştirilen modellerin mutlak SHAP değerleri.



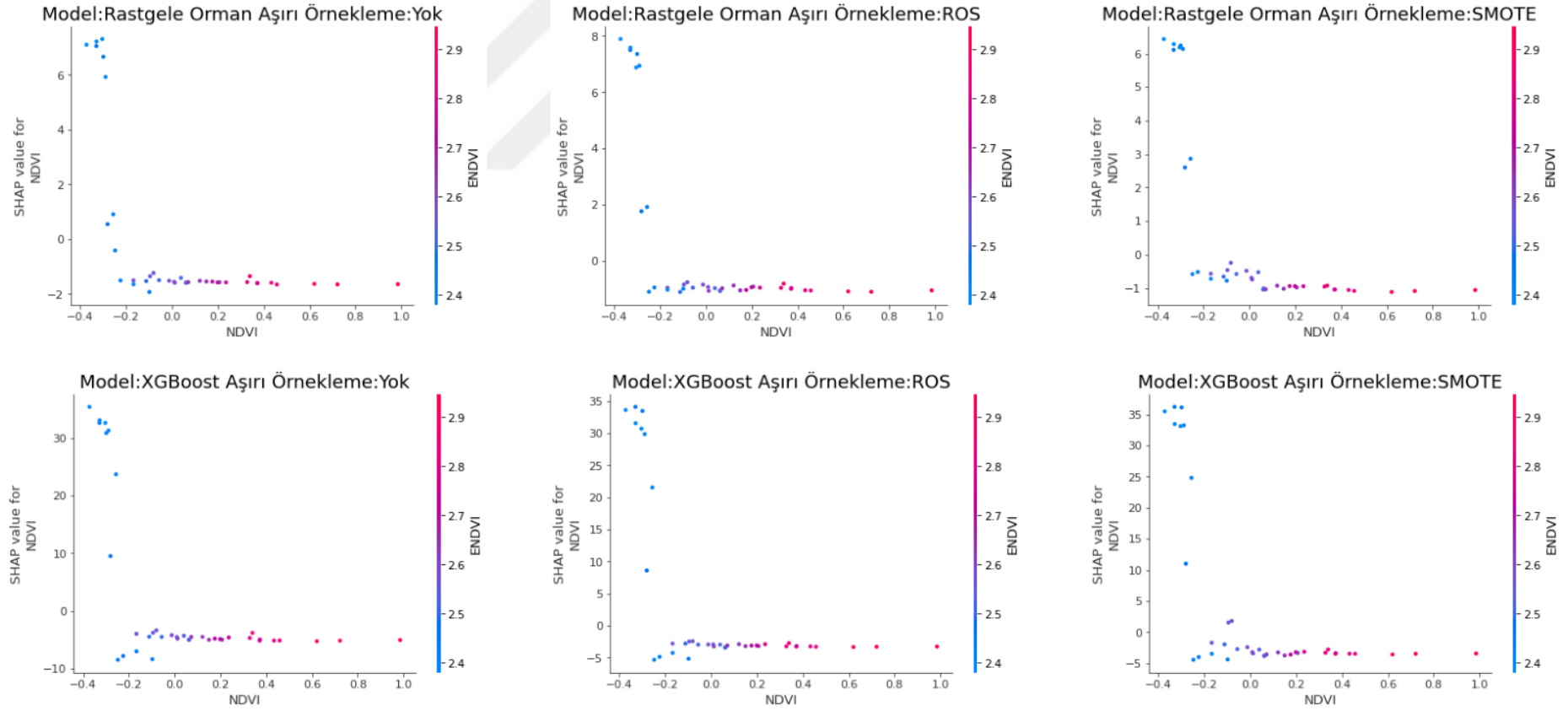
Şekil 4.15 : Batı Urmiye Playası için SHAP değerlerinin özelliklere göre modele olan etkisi.



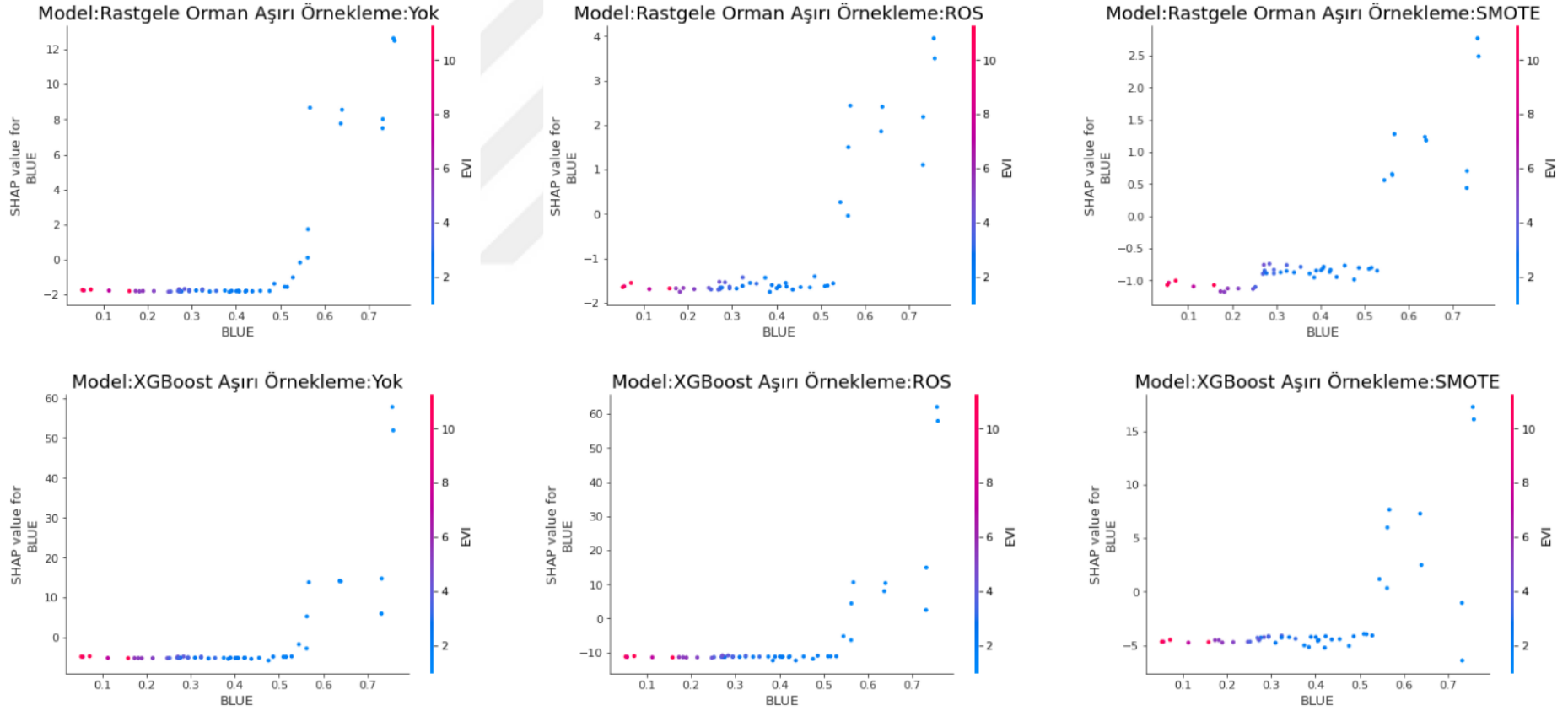
Şekil 4.16 : Bonab bölgesi için SHAP değerlerinin özelliklere göre modele olan etkisi.

Özellikler arasındaki bağımlılığı incelemek için 6 test arasında en çok katkı sunan iki özellik bağımlılık için karşılaştırılmıştır. Batı Urmıye Playası için oluşturulan bağımlılık grafiğinde NDVI ve ENDVI özellikleri incelenmiştir. Oluşturulan grafik Şekil 4.17’de gösterilmiştir. Grafik incelendiğinde hem NDVI özelliğinin hem de ENDVI özelliğinin düşük değerlerinin modele olumlu katkı gösterdiği görülmektedir. Model ve aşırı örnekleme yöntemlerinden bağımsız olarak bu sonuç bulunmuştur. Bunun sebebi ise NDVI ve ENDVI özellikleri çoğunluk olarak aynı özelliklerin birleştirilmesi ile üretilmeleridir.

Bonab bölgesi için üretilen bağımlılık grafikleri ise Şekil 4.18’de gösterilmiştir. Bonab bölgesi için ise Batı Urmıye Playası bölgesinden farklı olarak Mavi ve EVI özellikleri en çok katkı vermiştir. Bu yüzden bu iki özellik arasındaki bağımlılık incelenmiştir. Grafik incelendiğinde ise Mavi özelliğın düşük değeri modele negatif yönde katkı sunarken, EVI özelliğın düşük değeri modele pozitif bir yönde katkı sunduđu göze çarpmaktadır. İki özellik arasında ters bir orantı olduđu tüm model ve aşırı örnekleme testlerinde ortak olarak tespit edilmiştir. Mavi değeri 0.55’ten büyük olan değeri de modele olan katkısı pozitive doğru geçiş yapmaktadır.



Şekil 4.17 : Batı Urmıye Playası NDVI ve ENLVI özelliklerinin bağımlılık grafiği.



Şekil 4.18 : Bonab bölgesi Mavi ve EVI özelliklerinin bağımlılık grafiği.

Test örnekleri bazında modelleri açıklamak için ise Şekil 4.19 – 4.28 arasında Batı Urmıye Playası için olan SHAP grafikleri incelenmiştir. Özellikler için genel analizden çok farklı olmayan sonuçlar elde edilmiştir. Grafikler temel olarak bir taban değerinin üzerine her bir özelliğin eklediği ya da çıkardığı elektriksel iletkenliği değerlerini ifade ederek açıklamaktadır. Taban değeri ise modelin herhangi bir özellik kullanmadan tahmin ettiği değerdir. Batı Urmıye Playası için test edilen 10 örnek incelendiğinde ENDVI, NDVI gibi indislerin aşırı tuzlu örnekler haricinde regresyon sonucunu düşürücü etkide olduğu bulunmuştur.

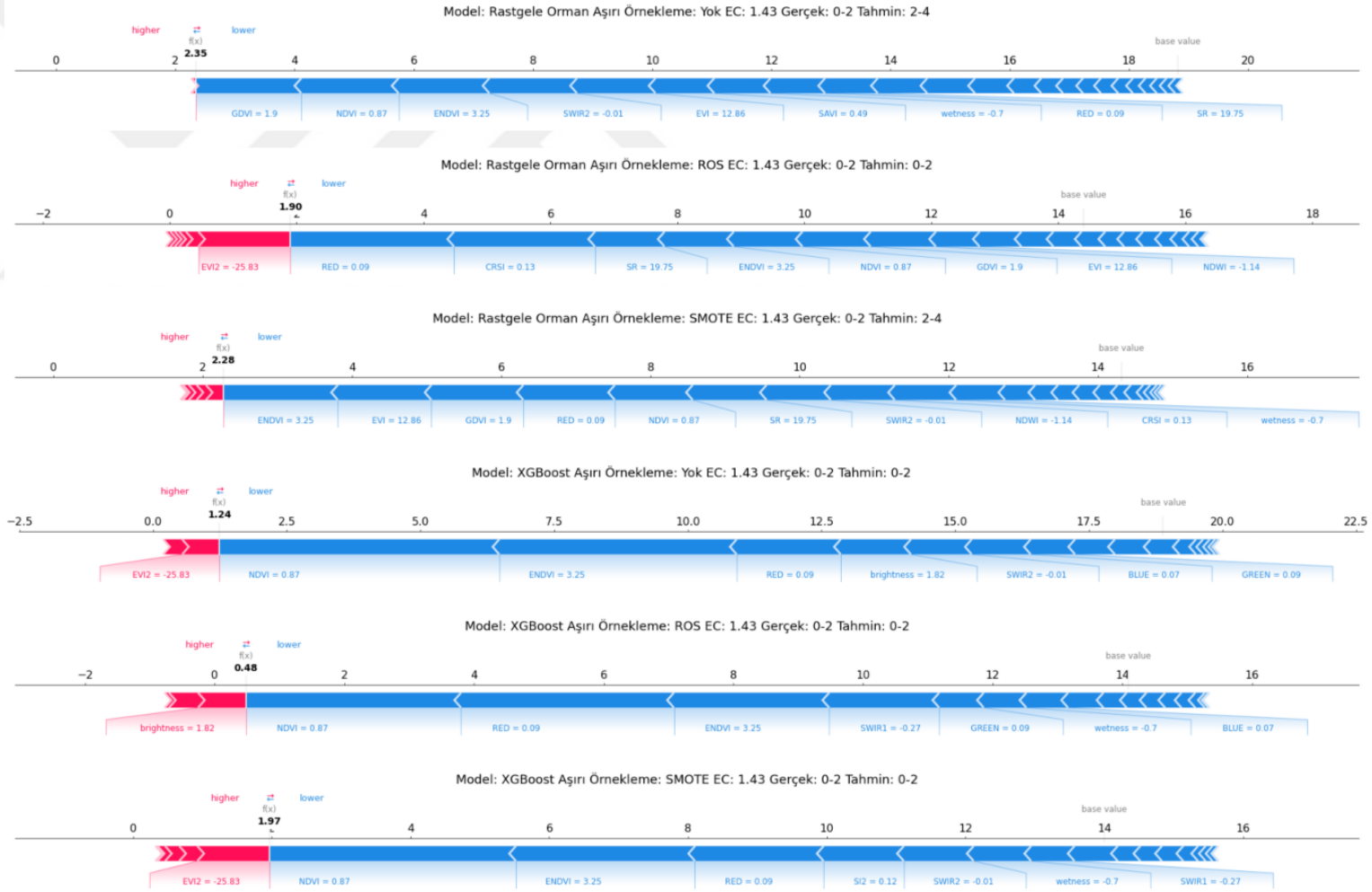
Şekil 4.19’da gösterilen test örneği incelendiğinde gerçeğe en yakın değeri XGBoost algoritması ve SMOTE aşırı örnekleme yöntemi ile elde edilen modelin bulunduğu görülmüştür. Model 0.57 dS/m değerini 2.64 dS/m olarak tahmin etmiştir. NDVI değerinin 0.53 ve ENDVI değerinin 3.04 ile model sonucuna en çok katkıyı negatif olarak verirken CRSI -0.13 değeri ile pozitif bir katkı vermiştir.

2. test örneği olan Şekil 4.20 incelendiğinde ise gerçeğe en yakın değeri tahmin eden model 1.24 dS/m ile XGBoost olmuştur. Aşırı örneklemenin kullanılmadığı bu modelde sonuca NDVI ve ENDVI özellikleri en yüksek negatif katkıyı verirken, EVI2 özelliği pozitif bir katkı vermiştir.

3. test örneğine ait kuvvet grafiği Şekil 4.21’de gösterilmiştir. Bu grafik incelendiğinde ise hiçbir model doğru tuzluluk seviyesinde tahmin yapamamıştır. En yakın model ise XGBoost ve SMOTE aşırı örnekleme yöntemi olmuştur. NDVI ve EVI2 özellikleri en çok negatif katkıyı sunarken, tuzluluk indislerinden olan SI6 özelliği modele en çok pozitif katkı veren özellik olmuştur.



Şekil 4.19 : Batı Urmiye Playası 1. test örneği SHAP değerleri.



Şekil 4.20 : Batı Urmiye Playası 2. test örneği SHAP değerleri.



Şekil 4.21 : Batı Urmiye Playası 3. test örneği SHAP değerleri.

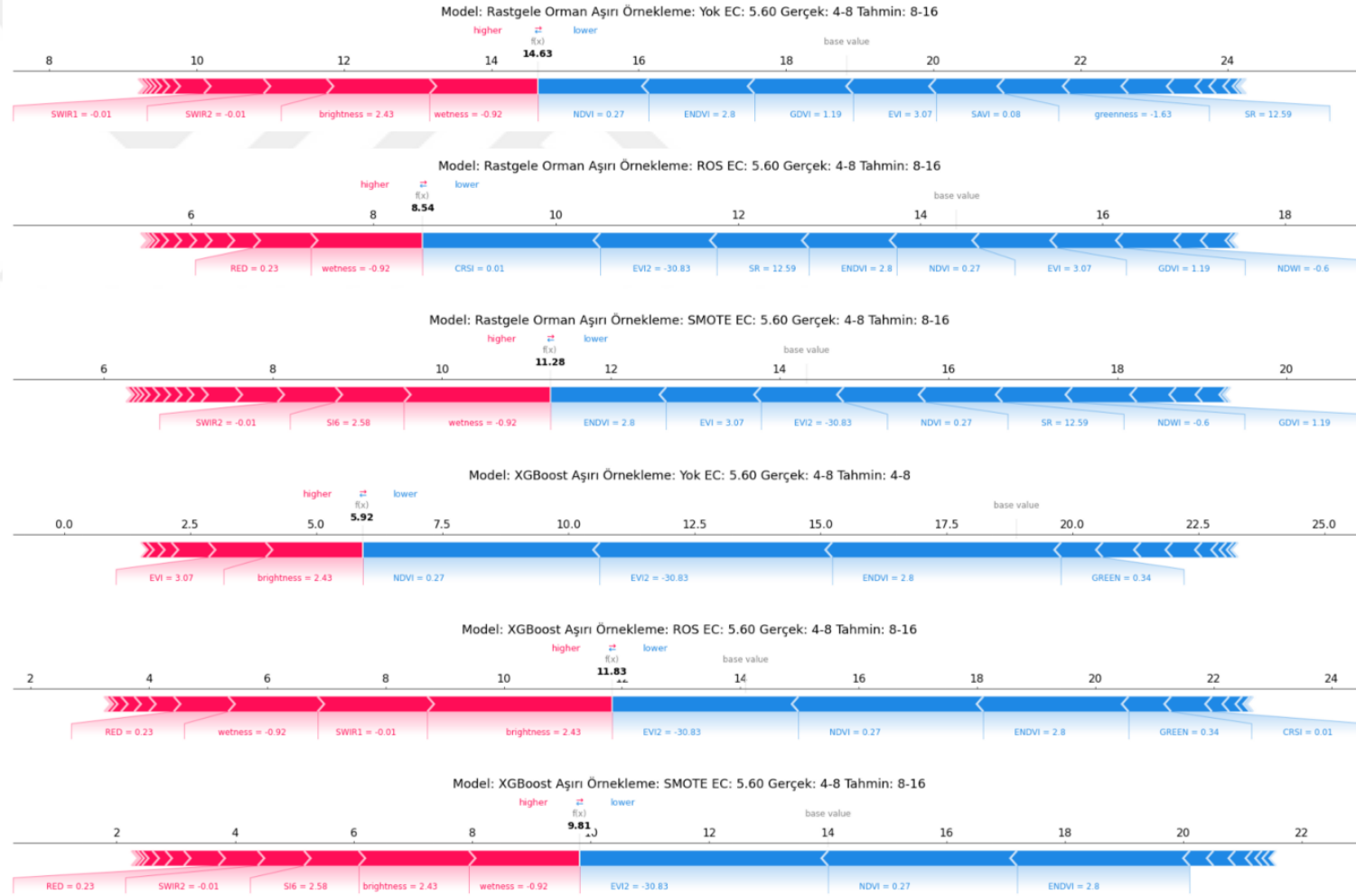
Şekil 4.22’de 4. test örneğine ait kuvvet grafikleri gösterilmiştir. Bu grafikler incelendiğinde ise Rastgele Orman algoritması ve SMOTE aşırı örnekleme yöntemi ile oluşturulan model en yakın tahmini 3.26 dS/m ile vermiştir. ENDVI ve GDVI özellikleri en çok negatif katkılar yapan özellikler olurken, pozitif katkı yapan özellikler yeterince katkı yapmadıkları için grafikte gösterilmemişlerdir.

Şekil 4.23’te gösterilen 5. test örneği için yapılan tahminler incelendiğinde en yakın sonucu 5.92 dS/m değeri ile XGBoost modeli yakalamıştır. Model sonucuna katkı sunan özellikler incelendiğinde ise NDVI ve EVI2 özellikleri en çok negatif katkı sunan özellikler olurken, Parlaklık ve EVI özellikleri ise en çok pozitif katkı sunan özellikler olmuşlardır.

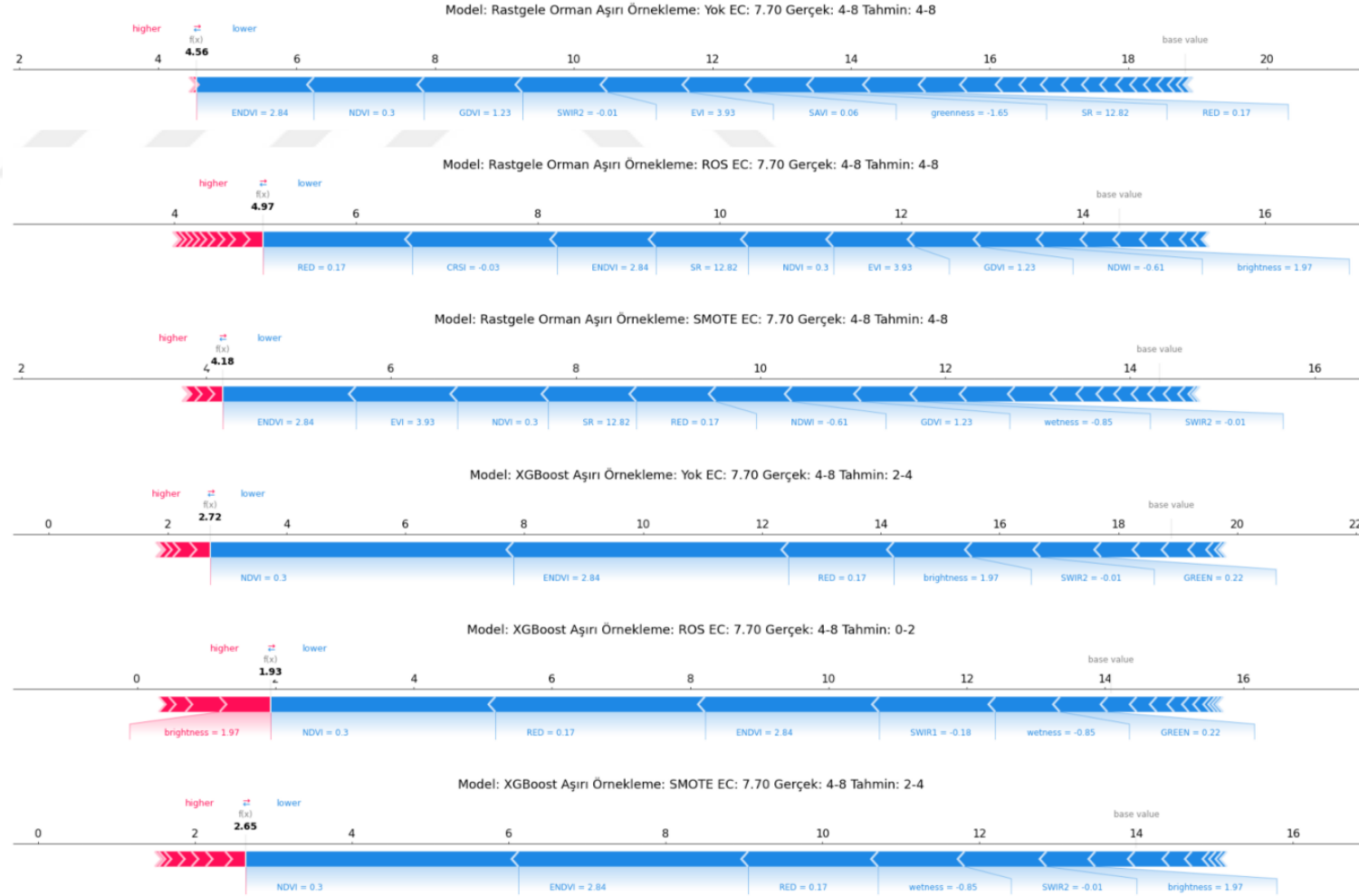
6. test örneği Şekil 4.24’te incelendiğinde hemen hemen her modelin büyük hatalar yaptığı göze çarpmaktadır. En yakın tahmini yapan model ise 4.97 dS/m değeri ile Rastgele Orman ve ROS örnekleme metodu olmuştur. Kırmızı ve CRSI özellikleri model sonucuna en çok negatif katkı veren özellikler olmuştur.



Şekil 4.22 : Batı Urmiye Playası 4. test örneği SHAP değerleri.



Şekil 4.23 : Batı Urmiye Playası 5. test örneği SHAP değerleri.



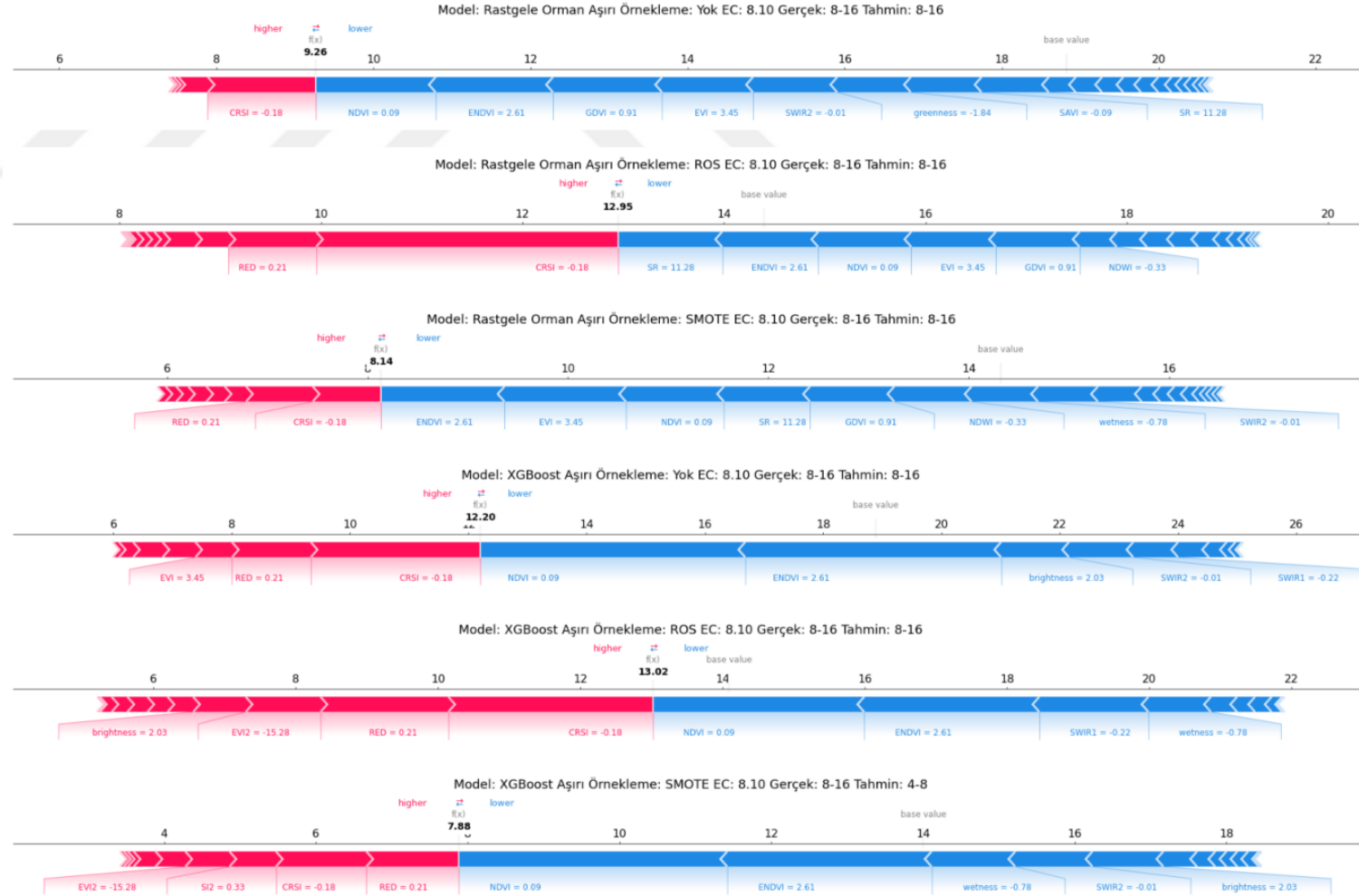
Şekil 4.24 : Batı Urmiye Playası 6. test örneği SHAP değerleri.

7. test örneği Şekil 4.25'te incelendiğinde SMOTE aşırı örnekleme ile eğitilen modellerin daha yakın sonuçlar verdiği göze çarpmaktadır. En yakın sonucu ise 8.14 dS/m değeri ile Rastgele Orman yöntemi vermiştir. ENDVI ve EVI özellikleri en çok negatif katkı veren özellikler olurken, Kırmızı ve CRSI özellikleri ise en çok pozitif katkı veren özellikler olmuşlardır.

Şekil 4.26'da gösterilen 8. test örneğinde ise en yakın tahmini veren model 11.04 dS/m değeri ile Rastgele Orman ve SMOTE aşırı örnekleme yöntemi olmuştur. EVI ve ENDVI en çok negatif katkıyı veren özellikler olurken, Kırmızı ve CRSI ise en çok katkı veren negatif özellikler olmuştur.

Şekil 4.27 incelendiğinde en iyi sonucu veren model 63.65 dS/m değeri ile XGBoost modeli olmuştur. -0.33 NDVI değeri ve ENDVI özelliği en çok pozitif katkı veren özellik olurken, tuzluluk indislerinden olan SI5 negatif katkı vermiştir.

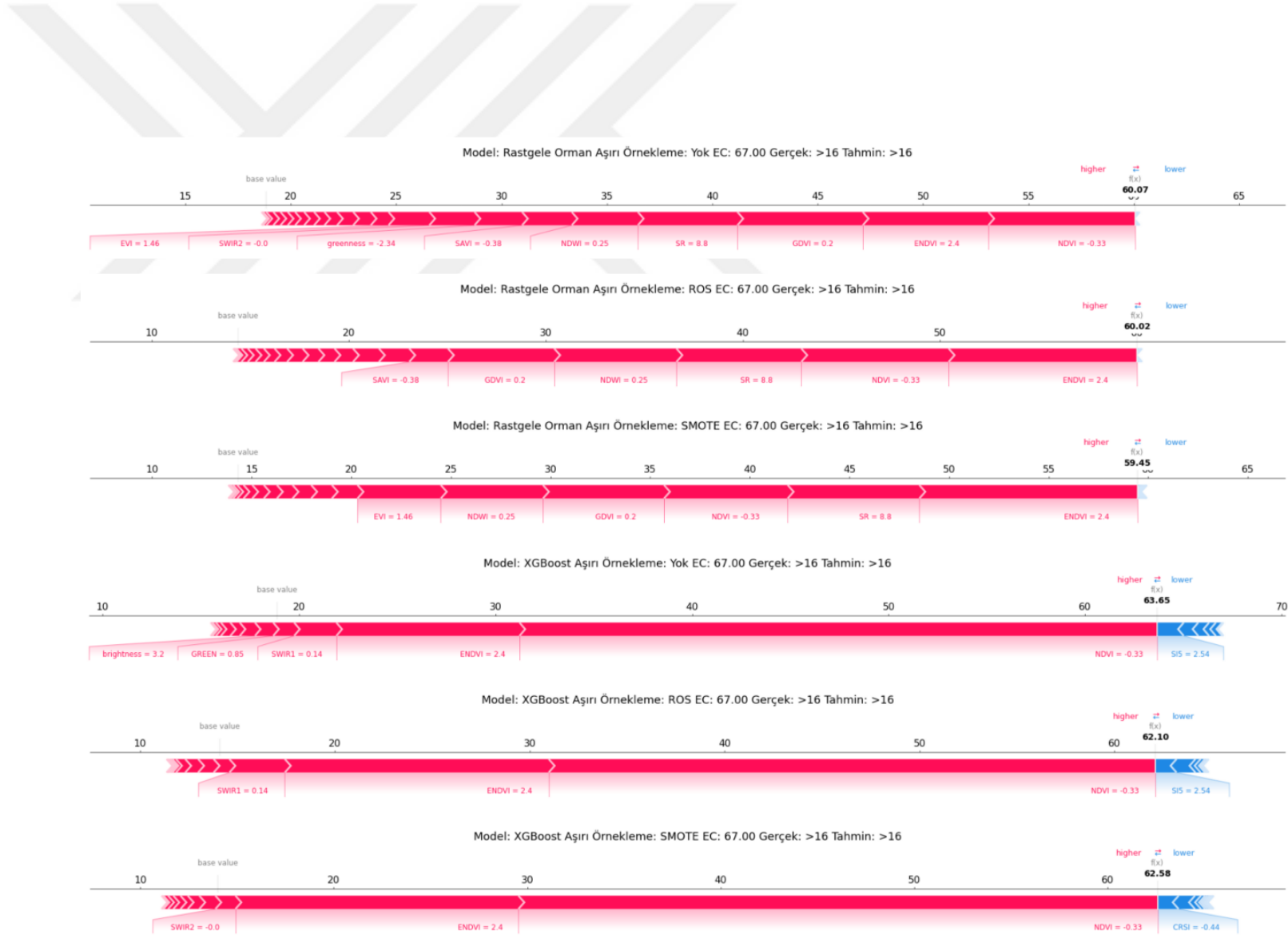
Batı Urmiye Playası için son örnek olan 10. test örneği için oluşturulan kuvvet grafikleri Şekil 4.28'de gösterilmiştir. Tüm modeller gerçek EC değerinden farklı sonuçlar vermesine karşın en yakın sonucu 37.56 dS/m değeri ile XGBoost algoritması ROS aşırı örnekleme metodu ile vermiştir. Ancak yine de tüm modeller doğru tuzluluk seviyesinde değerler tahmin etmişlerdir. ENDVI ve Yeşil özellikleri en çok pozitif katkıyı veren özellikler olurken, NDVI bu test örneğinde negatif katkı vermiştir.



Şekil 4.25 : Batı Urmiye Playası 7. test örneği SHAP değerleri.



Şekil 4.26 : Batı Urmiye Playası 8. test örneği SHAP değerleri.



Şekil 4.27 : Batı Urmiye Playası 9. test örneği SHAP değerleri.



Şekil 4.28 : Batı Urmıye Playası 10. test örneği SHAP değerleri.

Bonab bölgesi için seçilen 13 test örneği için ise üretilen grafikler Şekil 4.29 – Şekil 4.41 arasında gösterilmiştir. Uydu görüntüsünün mavi bandı toprak tuzluluğu regresyonuna negatif olarak katkı yaparken, yeşil bandı tuzluluk seviyesi ile doğru orantılı katkı yapmaktadır. Tuzluluk indisleri ise özellikle tuzluluk arttıkça pozitif yönlü katkılar yapmaktadırlar. Parlaklık değeri birçok tuzluluk seviyesinde regresyon sonucuna pozitif olarak katkı sunmuştur.

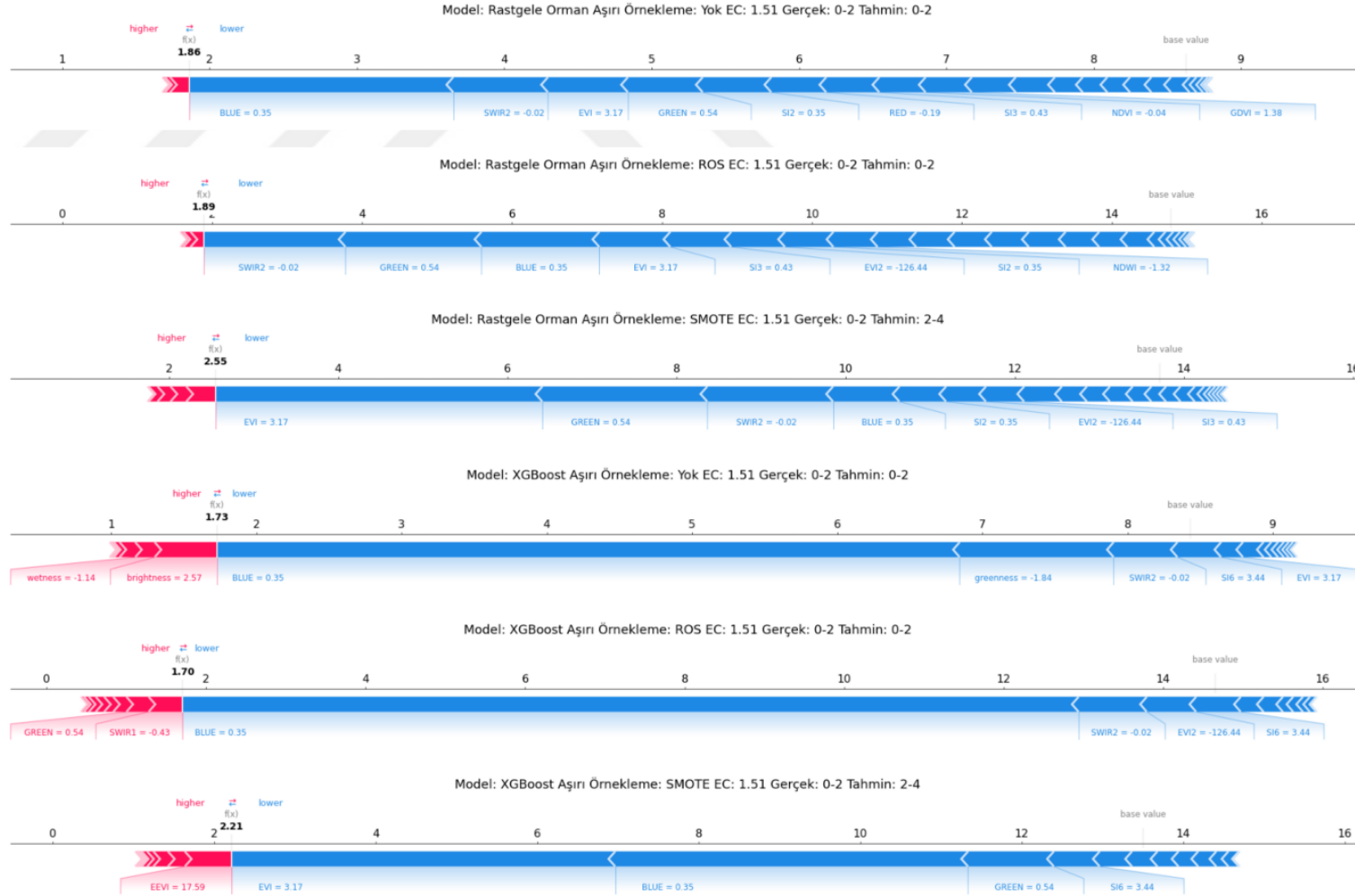
Şekil 4.29’da Bonab bölgesi için 1. test örneğine ait kuvvet grafikleri gösterilmiştir. En yakın tahmini veren model 1.81 dS/m değeri ile XGBoost ve ROS aşırı örnekleme yöntemi olmuştur. Yeşil ve SWIR1 özellikleri modele pozitif katkı yaparken Mavi ve EVI2 özellikleri ise en çok negatif katkı yapan özellikler olmuşlardır.

Şekil 4.30’da gösterilen 2. test örneğine ait kuvvet grafikleri incelendiğinde en iyi sonucu 1.70 dS/m değeri ile XGBoost algoritması ROS ile başarmıştır. Model sonucuna katkı yapan özellikler incelendiğinde ise Mavi ve SWIR2 özellikleri negatif katkı yaparken, SWIR1 ve Yeşil özellikleri en çok pozitif katkı yapan özellikler olmuşlardır.

3. test örneğine ait kuvvet grafikleri Şekil 4.31’de gösterilmiştir. XGBoost algoritmasının bariz bir şekilde Rastgele Orman algoritmasına göre daha yakın tahminler ürettiği göze çarpmaktadır. En iyi tahmini yapan ise 1.29 dS/m değeri ile XGBoost algoritması aşırı örnekleme olmadan yapmıştır. Mavi ve Yeşillik özellikleri en çok katkı yapan özellikler olmanın yanı sıra model sonucuna negatif katkı yapmışlardır. Aksine parlaklık ve Yeşil özellikleri ise modele en çok pozitif katkı yapan özellikler olmuşlardır.



Şekil 4.29 : Bonab bölgesi 1. test örneği SHAP değerleri.



Şekil 4.30 : Bonab bölgesi 2. test örneği SHAP değerleri.



Şekil 4.31 : Bonab bölgesi 3. test örneği SHAP değerleri.

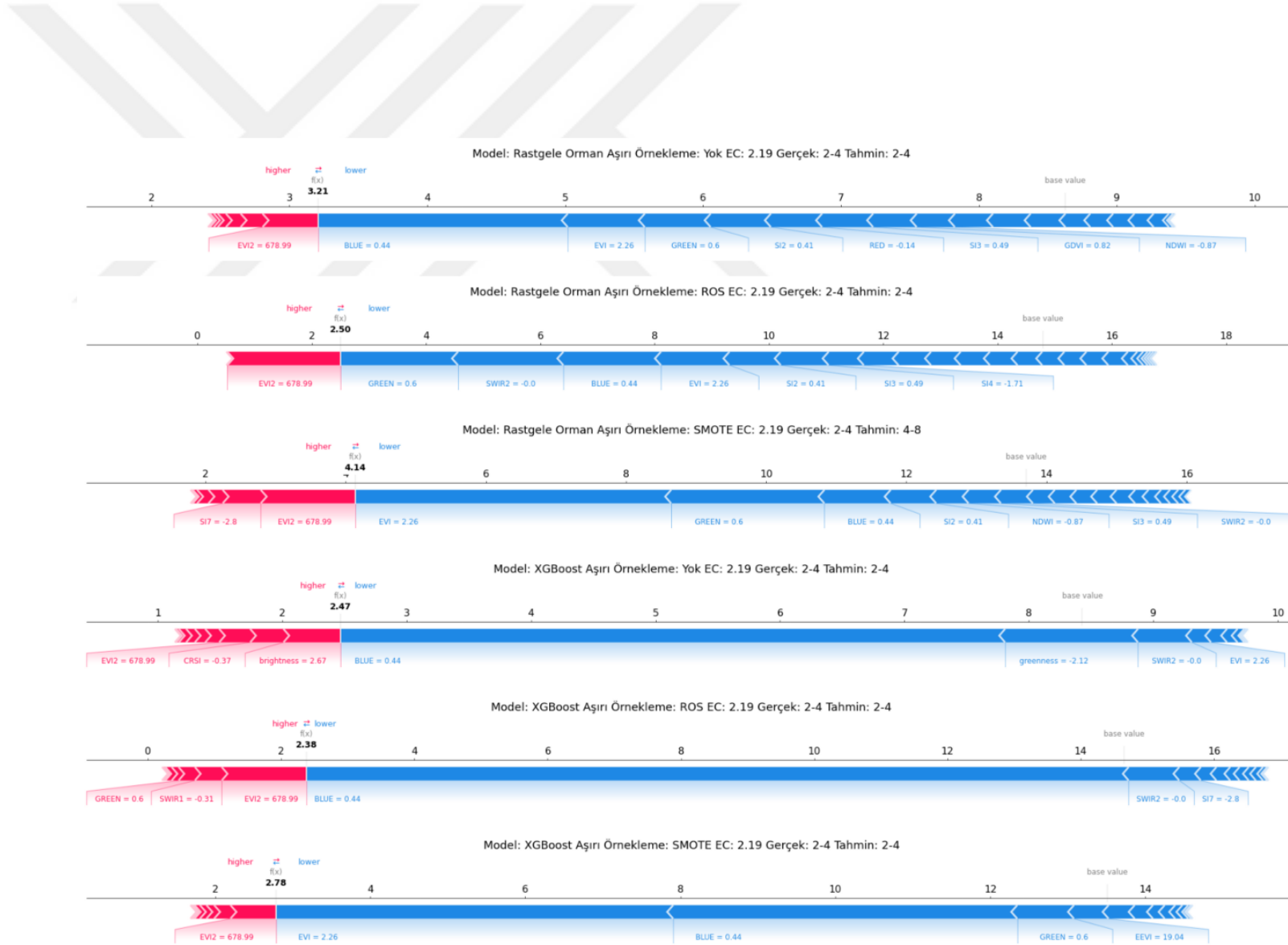
4. test örneğine ait Şekil 4.32’de gösterilen kuvvet grafikleri incelendiğinde en yakın tahmini 2.16 dS/m değeri ile Rastgele Orman algoritması SMOTE aşırı örnekleme yöntemi ile ulaşmıştır. Diğer test örneklerinden farklı olarak EVI ve Yeşil özellikleri model sonucuna negatif katkı sunarken, tuzluluk indislerinden olan SI7 özelliği ise pozitif olarak katkı sunmuştur.

Şekil 4.33’te 5. test örneğine ait kuvvet grafiği gösterilmiştir. 2.38 dS/m değeri ile en yakın tahmini XGBoost algoritması ROS aşırı örnekleme yöntemi ile vermiştir. Mavi ve SWIR2 özellikleri modele sonucuna en çok negatif katkı veren özellikler olurken, EVI2 ve SWIR1 özellikleri ise modele sonucuna en çok pozitif katkı veren özellikler olmuşlardır.

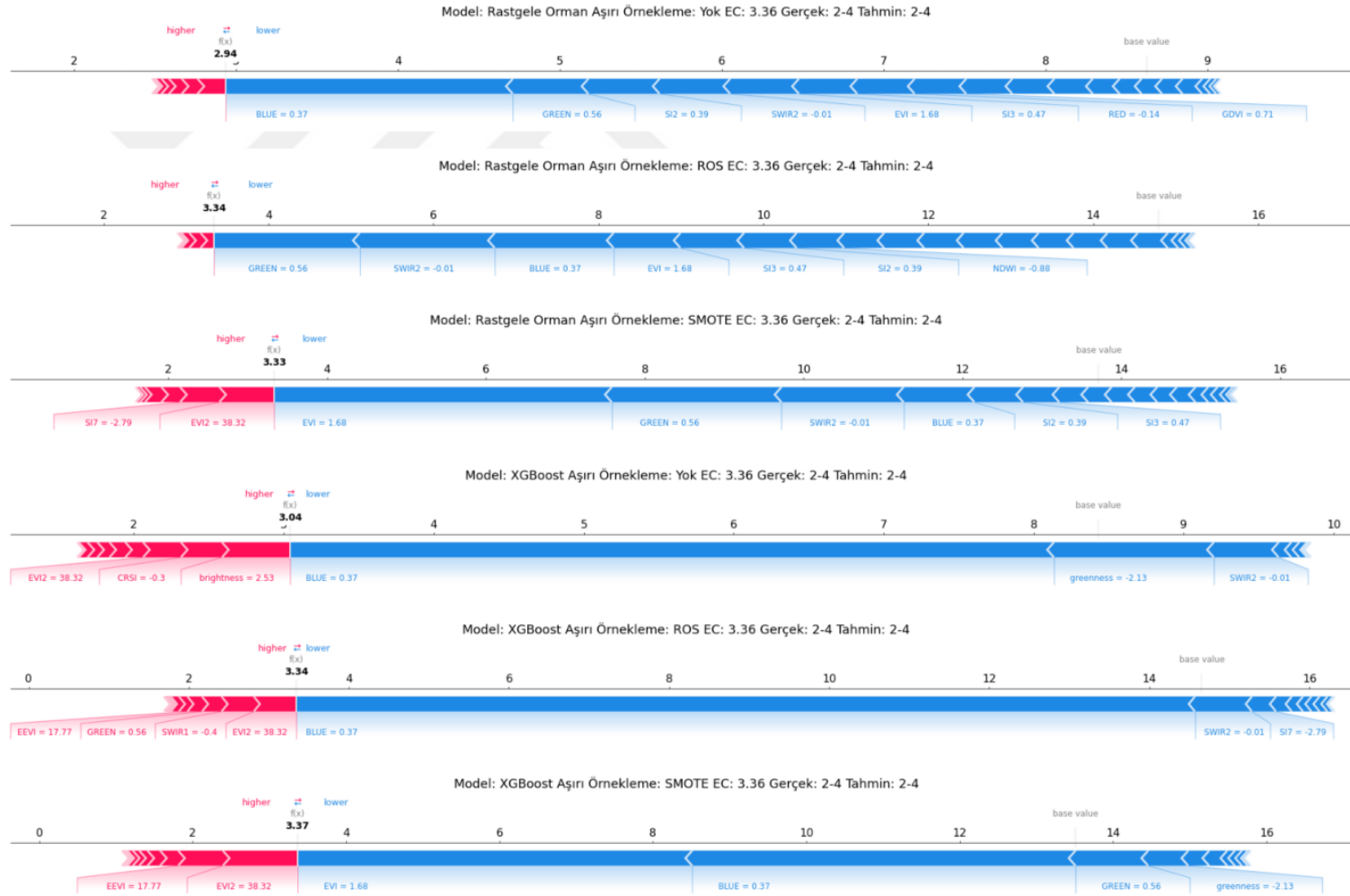
Şekil 4.34’te gösterilen 6. test örneğine ait kuvvet grafiği incelendiğinde 3.37 dS/m değeri ile en yakın tahmini veren modelin XGBoost ve SMOTE aşırı örnekleme metodu olduğu görülmüştür. EVI ve Mavi özellikleri model sonucuna en negatif katkıyı verirken, EVI2 ve EEVI özellikleri model sonucuna en pozitif katkı veren özellikler olmuşlardır.



Şekil 4.32 : Bonab bölgesi 4. test örneği SHAP değerleri.



Şekil 4.33 : Bonab bölgesi 5. test örneği SHAP değerleri.

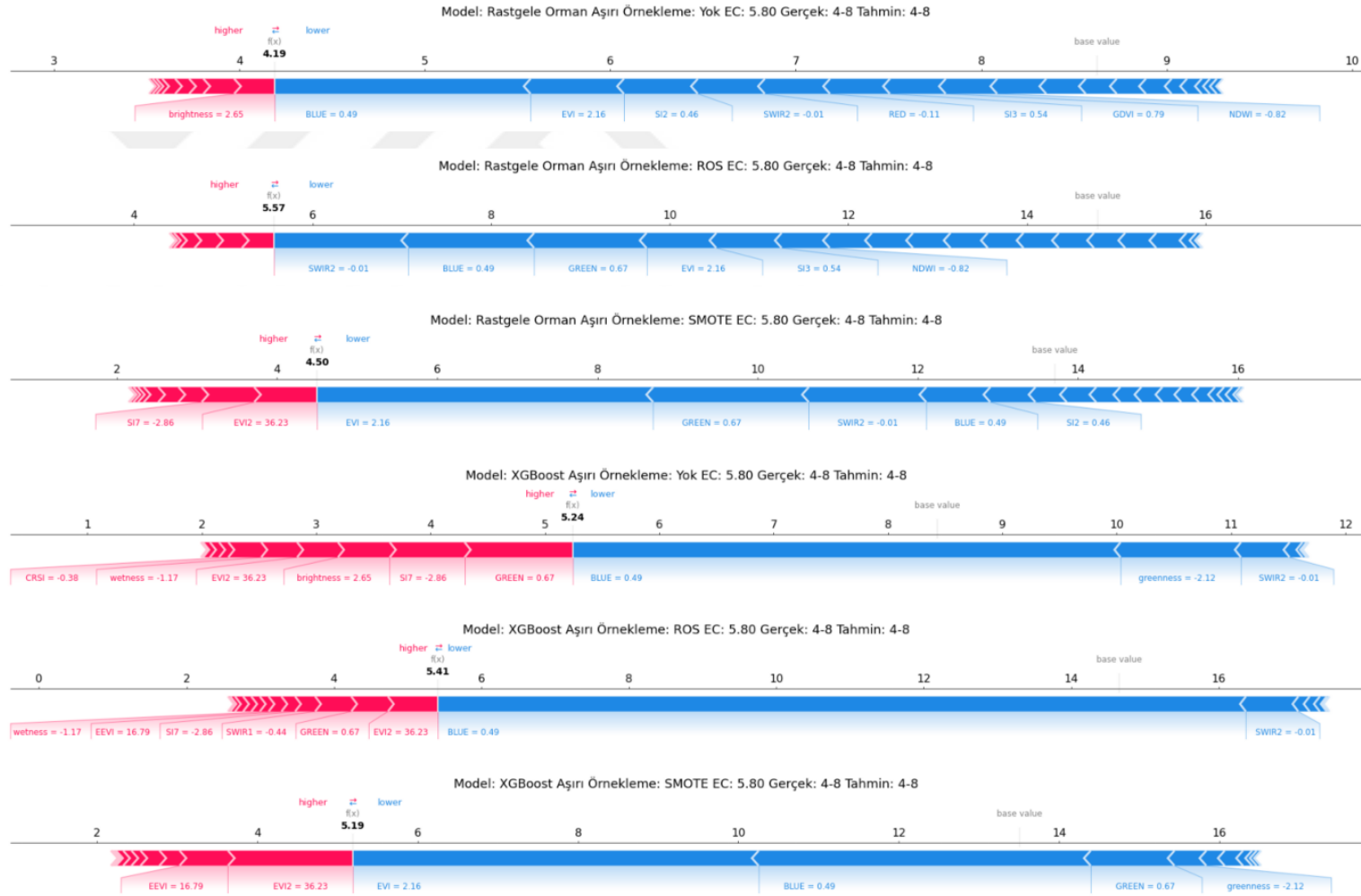


Şekil 4.34 : Bonab bölgesi 6. test örneği SHAP değerleri.

7. test örneğine ait kuvvet grafikleri Şekil 4.35'te gösterilmiştir. Kuvvet grafikleri incelendiğinde 5.57 dS/m değeri ile en yakın tahmini Rastgele Orman algoritması ROS aşırı örnekleme yöntemi ulaşmıştır. En çok katkı veren özellikler incelendiğinde ise, SWIR2 ve Mavi özellikleri en negatif katkıyı verirken pozitif katkı veren özelliklerden ayrışan yoktur.

8. test örneğine ait Şekil 4.36'da gösterilen kuvvet grafikleri incelendiğinde ise hiçbir modelin gerçeğe yakın bir tahminde bulunamadığı görülmüştür. En yakın tahmini 24.16 dS/m değeri ile XGBoost algoritması SMOTE aşırı örnekleme yöntemi ile yakalamıştır. EVI özelliği model sonucuna negatif katkı sunarken, Mavi ve Yeşil özellikleri ise model sonucuna pozitif olarak katkı sunmuştur.

9. test örneğine ait kuvvet grafikleri Şekil 4.37'de gösterilmiştir. 8. test örneğine benzer olarak yine hiçbir model gerçeğe yakın bir tahminde bulunamamıştır. En yakın tahmini ise bu sefer Rastgele Orman yöntemi yine SMOTE aşırı örnekleme yöntemi ile 18.24 dS/m değeri ile yakalamıştır. Model sonucuna EVI2 ve ENDVI özellikleri en negatif katkıyı verirken, Yeşil ve Mavi özellikleri ise en pozitif katkıyı vermiştir.



Şekil 4.35 : Bonab bölgesi 7. test örneği SHAP değerleri.



Şekil 4.36 : Bonab bölgesi 8. test örneği SHAP değerleri.



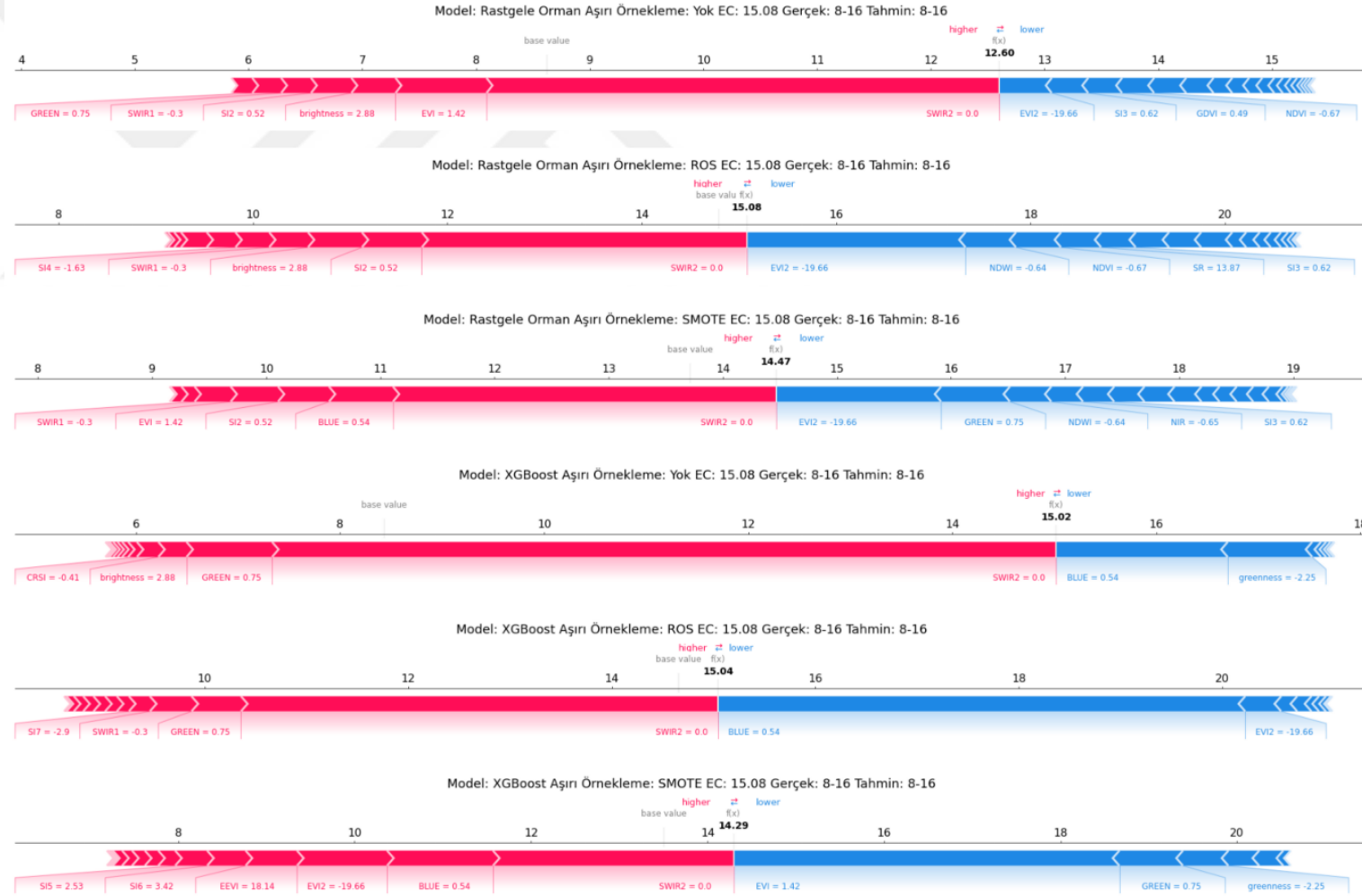
Şekil 4.37 : Bonab bölgesi 9. test örneği SHAP değerleri.

Şekil 4.38’de gösterilen 10. test örneğine ait kuvvet grafikleri incelendiğinde gerçekte aynı değeri yani 15.08 dS/m tahmin eden model Rastgele orman ve ROS aşırı örnekleme yöntemi olmuştur. Tuzluluk indislerinden olan SI2 ve SWIR2 özellikleri model sonucuna en pozitif katkı veren özellikler olurken, EVI2 ve NDWI özellikleri ise model sonucuna en negatif katkı veren özellikler olmuşlardır.

Şekil 4.39’da gösterilen 11. test örneğine ait kuvvet grafikleri gösterilmiştir. Hiçbir model gerçeğe yakın olarak sonuç vermemesine karşın hepsi doğru tuzluluk seviyesine düşen değerler tahmin etmişlerdir. En yakın tahmini veren model 29 dS/m değeri ile Rastgele Orman olmuştur. Mavi ve SWIR2 özellikleri model sonucuna en pozitif katkı veren özellikler olurken, negatif katkı veren özelliklerden öne çıkan olmamıştır.

Şekil 4.40’ta gösterilen kuvvet grafikleri 12. test örneğine aittir. Bu grafikler incelendiğinde ise XGBoost algoritmasının bariz olarak daha yakın sonuçlar tahmin ettiği göze çarpmaktadır. En yakın tahmini veren model ise 31.27 dS/m değeri ile XGBoost yöntemi olmuştur. Mavi ve yeşillik özellikleri model sonucuna en pozitif katkı veren özellikler olurken, EVI2 özelliği ise en negatif katkı veren özellik olmuştur.

Bonab bölgesine ait son test örneğinin kuvvet grafikleri Şekil 4.41’de gösterilmiştir. Grafikler incelendiğinde, hiçbir modelin gerçek değere yakın sonuçlar tahmin edemediği görülmüştür. Buna rağmen tüm modeller doğru tuzluluk seviyesinde değerler tahmin etmişlerdir. En yakın değeri 38.45 dS/m değeri ile Rastgele Orman algoritması SMOTE aşırı örnekleme yöntemi ile elde etmiştir. EVI ve EVI2 model sonucuna en pozitif katkısı olan değerler olurken, negatif katkı yapan ve öne çıkan bir özellik olmamıştır.



Şekil 4.38 : Bonab bölgesi 10. test örneği SHAP değerleri.



Şekil 4.40 : Bonab bölgesi 12. test örneği SHAP değerleri.



Şekil 4.41 : Bonab bölgesi 13. test örneği SHAP değerleri.

Toprak tuzluluğunun tahmini için geliştirilen yöntemler karşılaştırıldığında, makine öğrenme yöntemlerinin lineer regresyon gibi daha basit yöntemlere göre daha iyi sonuçlar verdiği gözlenmektedir. Yersel örneklerin kısıtlı olmasına rağmen makine öğrenme algoritmaları yine de iyi tahminler yapabilmektedir. Özellikle XGboost algoritması rastgele orman algoritmasına göre daha yüksek doğrulukta sonuçlar vermiştir. Sadece Bonab bölgesinde bulunan aşırı tuzlu toprak olan tuz yapısını tespit etmekte güçlük çekmiştir. SMOTE aşırı örnekleme yöntemi de makine öğrenme modellerinin doğruluğuna en çok katkı yapan yöntem olmuştur.

Aynı bölge için daha önce Aksoy ve diğ. (2022) tarafından yapılan çalışma göz önüne alındığında ise Batı Urmiye Playası için daha düşük doğruluk elde edilmiştir. Bunun sebebi ise değişen yersel örneklerle, farklı tarihli uydu görüntüsüne ve çapraz doğrulama ile doğruluk analizi yapılmış olmasına bağlanabilir.

Zarei ve diğ. (2021) tarafından yapılan toprak tuzluluğunun makine öğrenme yöntemleri ile incelenmesi çalışmasında XGBoost yöntemini rastgele orman ile karşılaştırmışlardır. Çalışmayı İran'ın Eshtehard şehrinde gerçekleştirmişlerdir. Toprak tuzluluğunu belirlemede bu çalışmadaki sonuca benzer olarak XGBoost R^2 değeri 0.76 ile en iyi sonucu vermiştir. 46 adet özellik kullanarak modellerini eğitmişlerdir.

Zhou ve diğ. (2022) küresel çapta yaptıkları tuzluluk çalışmalarında rastgele orman XGBoost makine öğrenme algoritmalarını karşılaştırmışlardır. Yine bu çalışmaya benzer olarak XGBoost algoritmasının en iyi sonucu verdiğini bulmuşlardır. XGBoost yöntemi R^2 0.71 ile en iyi sonuç iken onu izleyen yöntem ise R^2 0.69 ile rastgele orman algoritması olmuştur.

XGBoost, rastgele orman ve karar ağaçlarının toprak tuzluluğunu modellemede karşılaştırıldığı başka bir çalışma Ma ve diğ. (2021) tarafından yapılmıştır. Çalışma Çin'in Tarım Havzasında 2018 yılında toplanan toprak örnekleri ile yapılmıştır. Bu çalışmada sonucunda %5'lik bir fark ile XGBoost yöntemi rastgele ormana göre daha iyi bir sonuç vermiştir.

Abedi ve diğ. (2021) tarafından İran'ın Fars şehrinde toplanan toprak örnekleri ile yapılan çalışması incelendiğinde diğer çalışmalardan farklı olarak elektriksel iletkenlik tahmininde en düşük doğruluğu R^2 olarak vermiştir. Bu çalışma XGBoost algoritmasının her zaman rastgele ormana veya diğer algoritmalara göre iyi

olamayacağının güzel bir göstergesidir. Yapılan çalışmada R^2 değerleri XGBoost için 0.45, rastgele orman için 0.48 ve Cubist modeli için 0.61 olarak bulunmuştur.

Nguyen ve diğ. (2020) tarafından 2016 ve 2020 yılları arasında Vietnam'da bulunan Mekong Deltasında toplanan toprak örnekleri ile analizler gerçekleştirilmiştir. Makine öğrenme yöntemleri olarak XGBoost, SVM, derin sinir ağları kullanılmıştır. XGBoost R^2 0.76 değeri ile bu modeller arasında en iyi sonucu veren yöntem olmuştur.

Özelliklerin modellere ne kadar katkı verdikleri ise bu çalışmalar XGBoost algoritmasının sunduğu özellik önemi derecelendirmesine bakarak incelemişlerdir (Abedi ve diğ., 2021; Ma ve diğ., 2021; Nguyen ve diğ., 2020; Zhou ve diğ., 2022). Özellik önem dereceleri SHAP özellik önem grafiklerine benzemektedir. Ancak SHAP sadece özelliklerin önemini değil, aynı zamanda modele hangi yönde katkı yaptığını hem genel model için hem de örnek bazlı olarak açıklayabilmektedir. Özellikler arasında bağımlılık bu çalışmalarda korelasyonlar olarak ifade edilmiştir. SHAP, özelliklerin bağımlılıklarını modele olan katkılarına göre incelemek için de imkân sunmaktadır.



5. SONUÇLAR VE ÖNERİLER

Toprak tuzluluğu birçok ekosistemi etkilemesiyle bilinen çok önemli çevresel bir problemdir. Özellikle yarı kurak ve kurak alanları daha çok etkileyen bu problem, tarım arazilerinin terk edilmesine yol açmaktadır. Hızla artan nüfus düşünüldüğünde tarım arazilerinin terk edilmesi kabul edilemez bir durumdur. Gıda krizi gibi çok kritik olaylara neden olmaması için toprak tuzluluğu karar vericiler tarafından problem olduğu fark edilmeli ve bu konuda önlemler alınmalıdır. Bu kavramda Birleşmiş Milletlerin Sürdürülebilir Kalkınma Hedefleri (Url-4) kapsamında gıda güvenliği, su kullanımı ve sürdürülebilir ekosistemler çerçevesinde toprak tuzluluğu izlenmeli ve gerekli müdahaleler yapılmalıdır.

Birçok çevresel sorunun belirlenmesinde kullanıldığı gibi, uzaktan algılama ve makine öğrenme yöntemleri toprak tuzluluğunun belirlenmesi ve izlenmesi için önemli roller oynamaktadır. Rastgele orman ve XGBoost algoritmaları literatürde bulunan makine öğrenme algoritmalarından sadece ikisidir. Temel olarak makine öğrenme algoritmaları bir veri seti üzerinde eğitilip, daha sonra hiç görmediği veriler üzerinde tahminler yapması ile bilinirler. Makine öğrenme algoritmaları birçok girdiye almasına karşın, girdilerin problemi doğru yansıtıp yansıtmadığı ise büyük bir ikilemdir. Bu yüzden alan bilgisi bu konuda uzaktan algılama ile bu problemi çözerken oldukça faydalı olmaktadır.

Makine öğrenme algoritmalarının kapalı bir kutu olması ve verdikleri kararları doğru sebeplere göre mi verdikleri her zaman tartışma konusu olmuştur. Açıklanabilir yapay zekâ yöntemleri ise bu soruna bir çözüm getirmişlerdir. SHAP değerleri literatürde de sıklıkla kullanıldığı için bu çalışma kapsamında tercih edilmiştir. SHAP değerleri toprak tuzluluğunun hangi özelliklerin nasıl etkilediğini bu çalışma kapsamında gözler önüne sermiştir.

Urmiye Gölü havzası ekonomik ve ekolojik olarak Urmiye ili için çok önemli bir yere sahiptir. Ancak insan etkisiyle, aşırı yer altı sularının kullanımı gibi nedenlerden dolayı havza üzerindeki tarım arazilerinde toprak tuzlanma sorunu baş göstermektedir. Bu

sorun havza üzerindeki tarımsal faaliyetleri olumsuz etkilemektedir. Son yıllarda artan toprak tuzluluğu problemi, tarım arazilerinin terk edilmesine varacak kadar kritik duruma gelmiştir. Bu bölgede toprak tuzluluğunun tespiti, izlenmesi ve müdahale edilmesi bu yüzden oldukça önemlidir.

Çalışmanın bulguları incelendiğinde XGBoost makine öğrenme algoritmasının daha iyi sonuç verdiği görülürken, sadece tuz yapıları için Bonab bölgesinde aşırı tuzlu alanı yakalamakta güçlük çektiği de görülmüştür. Veri çeşitlendirme yöntemlerinin model doğruluklarını iyileştirdiği tespit edilmiştir. SHAP değerleri incelendiğinde ise iki çalışma alanı için farklı olsa da yakın kızıl ötesi ve görünür bölge bantlarının yanı sıra onlardan türetilmiş bitki indislerinin de modellere yüksek katkılar verdiği görülmüştür.

Bu çalışmayı ileriye taşımayı limitleyen ise yersel verilerin kısıtlı olmasıdır. Makine öğrenme yöntemleri nihayetinde elde olan veriler ile eğitilmektedir. Daha fazla veri ile eğitilmiş ve test edilmiş modeller daha güvenilir sonuçlar vereceği aşikardır. Yersel örnekler toplanmadan önce uydu üzerinden muhtemel alanlar tercih edilerek toprak tuzluluğu için toplanan verilerin niteliği de arttırılabilir.

Gelecekteki çalışmalar için farklı uydu sistemleri, daha yüksek mekânsal çözünürlüklü uydular, derin öğrenme yöntemleri, alternatif veri çeşitlendirme yöntemleri kullanılabilir. Türkiye üzerinde hali hazırda toprak tuzluluğu problemi yaşayan Tuz Gölü havzası çalışma alanı olarak seçilebilir. Özellikle Urmiye Gölüne olan benzerliği, tarım alanları üzerindeki insan kaynaklı faktörler ile benzer bölgelerdir Aynı yöntemlerin Tuz Gölü içinde faydalı olacağı düşünülmektedir. Optik uydu sistemleri verilerinin yanı sıra yapay açıklı radar sistemleri, pasif mikrodalga uydu sistemleri, arazinin yapısı ile ilgili lazer tarama verileri, toprak ile ilgili veriler, meteorolojik veriler kullanılabilir.

KAYNAKLAR

- Abbas, M., ve Khan, S.** (2007). Using Remote Sensing Techniques for Appraisal of Irrigated Soil Salinity. *MODSIM07 - Land, Water and Environmental Management: Integrated Systems for Sustainability, Proceedings*.
- Abdollahi, A., ve Pradhan, B.** (2021). Urban Vegetation Mapping from Aerial Imagery Using Explainable AI (XAI). *Sensors*, 21(14), Art. 14. <https://doi.org/10.3390/s21144738>
- Abedi, F., Amirian-Chakan, A., Faraji, M., Taghizadeh-Mehrjardi, R., Kerry, R., Razmjoue, D., ve Scholten, T.** (2021). Salt dome related soil salinity in southern Iran: Prediction and mapping with averaging machine learning models. *Land Degradation & Development*, 32(3), 1540-1554. <https://doi.org/10.1002/ldr.3811>
- Akca, S., ve Gungor, O.** (2022). Semantic segmentation of soil salinity using in-situ EC measurements and deep learning based U-NET architecture. *CATENA*, 218, 106529. <https://doi.org/10.1016/j.catena.2022.106529>
- Aksoy, S., Yildirim, A., Gorji, T., Hamzehpour, N., Tanik, A., ve Sertel, E.** (2022). Assessing the performance of machine learning algorithms for soil salinity mapping in Google Earth Engine platform using Sentinel-2A and Landsat-8 OLI data. *Advances in Space Research*, 69(2), 1072-1086. <https://doi.org/10.1016/j.asr.2021.10.024>
- Allbed, A., ve Kumar, L.** (2013). Soil Salinity Mapping and Monitoring in Arid and Semi-Arid Regions Using Remote Sensing Technology: A Review. *Advances in Remote Sensing*, 2013. <https://doi.org/10.4236/ars.2013.24040>
- Al-Najjar, H. A. H., Pradhan, B., Beydoun, G., Sarkar, R., Park, H.-J., ve Alamri, A.** (2022). A novel method using explainable artificial intelligence (XAI)-based Shapley Additive Explanations for spatial landslide prediction using Time-Series SAR dataset. *Gondwana Research*. <https://doi.org/10.1016/j.gr.2022.08.004>
- Atalay, İ.** (2006). *Toprak oluşumu, Sınıflandırılması ve Coğrafyası* (3. bs). Meta Basım Matbaacılık Hizmetleri.
- Baig, M. H. A., Zhang, L., Shuai, T., ve Tong, Q.** (2014). Derivation of a tasselled cap transformation based on Landsat 8 at-satellite reflectance. *Remote Sensing Letters*, 5(5), 423-431. <https://doi.org/10.1080/2150704X.2014.915434>
- Betancourt, C., Stomberg, T. T., Edrich, A.-K., Patnala, A., Schultz, M. G., Roscher, R., Kowalski, J., ve Stadtler, S.** (2022). Global, high-resolution mapping of tropospheric ozone – explainable machine learning and impact of uncertainties. *Geoscientific Model Development*, 15(11), 4331-4354. <https://doi.org/10.5194/gmd-15-4331-2022>
- Breiman, L.** (2001). Random Forests. *Machine Learning*, 45(1), 5-32. <https://doi.org/10.1023/A:1010933404324>

- Brevik, E. C., Cerdà, A., Mataix-Solera, J., Pereg, L., Quinton, J. N., Six, J., ve Van Oost, K.** (2015). The interdisciplinary nature of *SOIL*. *SOIL*, 1(1), 117-129. <https://doi.org/10.5194/soil-1-117-2015>
- Brown, J. W., Allison, L. E., Hayward, H. E., Richards, A., Bernstein, L., Fireman, M., Pearson, G. A., Bower, C. A., Hatcher, J. T., ve Reeve, R. C.** (1954). *United States Salinity Laboratory Staff*. 166.
- Chawla, N. V., Bowyer, K. W., Hall, L. O., ve Kegelmeyer, W. P.** (2002). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, 16, 321-357. <https://doi.org/10.1613/jair.953>
- Chen, H., Zhao, G., Chen, J., Wang, R., ve Gao, M.** (2015). Remote sensing inversion of saline soil salinity based on modified vegetation index in estuary area of Yellow River. *Transactions of the Chinese Society of Agricultural Engineering*. <https://doi.org/10.3969/j.issn.1002-6819.2015.05.016>
- Chen, H., Zhao, G., Li, Y., Wang, D., ve Ma, Y.** (2019). Monitoring the seasonal dynamics of soil salinization in the Yellow River delta of China using Landsat data. *Natural Hazards and Earth System Sciences*, 19(7), 1499-1508. <https://doi.org/10.5194/nhess-19-1499-2019>
- Chen, T., ve Guestrin, C.** (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785-794. <https://doi.org/10.1145/2939672.2939785>
- Douaoui, A. E. K., Nicolas, H., ve Walter, C.** (2006). Detecting salinity hazards within a semiarid context by means of combining soil and remote-sensing data. *Geoderma*, 134(1), 217-230. <https://doi.org/10.1016/j.geoderma.2005.10.009>
- Fonseca, J., Douzas, G., ve Bacao, F.** (2021). Improving Imbalanced Land Cover Classification with K-Means SMOTE: Detecting and Oversampling Distinctive Minority Spectral Signatures. *Information*, 12(7), Art. 7. <https://doi.org/10.3390/info12070266>
- Gopalakrishnan, T., ve Kumar, L.** (2021). Linking Long-Term Changes in Soil Salinity to Paddy Land Abandonment in Jaffna Peninsula, Sri Lanka. *Agriculture*, 11(3), Art. 3. <https://doi.org/10.3390/agriculture11030211>
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., ve Moore, R.** (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202, 18-27. <https://doi.org/10.1016/j.rse.2017.06.031>
- Gorji, T., Sertel, E., ve Tanik, A.** (2017). Monitoring soil salinity via remote sensing technology under data scarce conditions: A case study from Turkey. *Ecological Indicators*, 74, 384-391. <https://doi.org/10.1016/j.ecolind.2016.11.043>
- Gorji, T., Yildirim, A., Hamzehpour, N., Tanik, A., ve Sertel, E.** (2020). Soil salinity analysis of Urmia Lake Basin using Landsat-8 OLI and Sentinel-2A based spectral indices and electrical conductivity measurements. *Ecological Indicators*, 112, 106173. <https://doi.org/10.1016/j.ecolind.2020.106173>
- Haghighi, A., Fazel, N., Hekmatzadeh, A. A., ve Klöve, B.** (2018). Analysis of Effective Environmental Flow Release Strategies for Lake Urmia Restoration. *Water Resources Management*, 32(11), 3595-3609. <https://doi.org/10.1007/s11269-018-2008-3>

- Hamzehpour, N., Eghbal, M. K., Abasiyan, S. M. A., ve Dill, H. G.** (2018). Pedogenic evidence of Urmia Lake's maximum expansion in the late Quaternary. *CATENA*, 171, 398-415. <https://doi.org/10.1016/j.catena.2018.07.019>
- Hateffard, F., Dolati, P., Heidari, A., ve Zolfaghari, A. A.** (2019). Assessing the performance of decision tree and neural network models in mapping soil properties. *Journal of Mountain Science*, 16(8), 1833-1847. <https://doi.org/10.1007/s11629-019-5409-8>
- Hounkpatin, K. O. L., Schmidt, K., Stumpf, F., Forkuor, G., Behrens, T., Scholten, T., Amelung, W., ve Welp, G.** (2018). Predicting reference soil groups using legacy data: A data pruning and Random Forest approach for tropical environment (Dano catchment, Burkina Faso). *Scientific Reports*, 8(1), Art. 1. <https://doi.org/10.1038/s41598-018-28244-w>
- Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., ve Ferreira, L. G.** (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, 83(1), 195-213. [https://doi.org/10.1016/S0034-4257\(02\)00096-2](https://doi.org/10.1016/S0034-4257(02)00096-2)
- Huete, A. R.** (1988). A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, 25(3), 295-309. [https://doi.org/10.1016/0034-4257\(88\)90106-X](https://doi.org/10.1016/0034-4257(88)90106-X)
- Jiang, H., Ding, J., Tahiplot, T., Zhao, R., ve Zhang, F.** (2008). EXTRACTION OF SALINIZED SOIL INFORMATION IN ARID AREA BASED ON ETM+ DATA. *ACTA PEDOLOGICA SINICA*, 45(2), 222-228. <https://doi.org/10.11766/trxb200607260236>
- Jordan, C. F.** (1969). Derivation of Leaf-Area Index from Quality of Light on the Forest Floor. *Ecology*, 50(4), 663-666. <https://doi.org/10.2307/1936256>
- Kabiraj, S., Jayanthi, M., Vijayakumar, S., ve Duraisamy, M.** (2022). Comparative assessment of satellite images spectral characteristics in identifying the different levels of soil salinization using machine learning techniques in Google Earth Engine. *Earth Science Informatics*. <https://doi.org/10.1007/s12145-022-00866-9>
- Kakogeorgiou, I., ve Karantzalos, K.** (2021). Evaluating explainable artificial intelligence methods for multi-label deep learning classification tasks in remote sensing. *International Journal of Applied Earth Observation and Geoinformation*, 103, 102520. <https://doi.org/10.1016/j.jag.2021.102520>
- Kasim, N., Maihemuti, B., Sawut, R., Abliz, A., Dong, C., ve Abdumutallip, M.** (2020). Quantitative Estimation of Soil Salinization in an Arid Region of the Keriya Oasis Based on Multidimensional Modeling. *Water*, 12(3), Art. 3. <https://doi.org/10.3390/w12030880>
- Khan, N. M., Rastoskuev, V. V., Sato, Y., ve Shiozawa, S.** (2005). Assessment of hydrosaline land degradation by using a simple approach of remote sensing indicators. *Agricultural Water Management*, 77(1), 96-109. <https://doi.org/10.1016/j.agwat.2004.09.038>
- Krishna, S., Han, T., Gu, A., Pombra, J., Jabbari, S., Wu, S., ve Lakkaraju, H.** (2022). *The Disagreement Problem in Explainable Machine Learning: A Practitioner's Perspective* (arXiv:2202.01602). arXiv. <https://doi.org/10.48550/arXiv.2202.01602>

- Li, Z., Li, Y., Xing, A., Zhuo, Z., Zhang, S., Zhang, Y., ve Huang, Y.** (2019). Spatial Prediction of Soil Salinity in a Semiarid Oasis: Environmental Sensitive Variable Selection and Model Comparison. *Chinese Geographical Science*, 29(5), 784-797. <https://doi.org/10.1007/s11769-019-1071-x>
- Lundberg, S. M., ve Lee, S.-I.** (2017). A Unified Approach to Interpreting Model Predictions. *Advances in Neural Information Processing Systems*, 30. <https://proceedings.neurips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html>
- Ma, G., Ding, J., Han, L., Zhang, Z., ve Ran, S.** (2021). Digital mapping of soil salinization based on Sentinel-1 and Sentinel-2 data combined with machine learning algorithms. *Regional Sustainability*, 2(2), 177-188. <https://doi.org/10.1016/j.regsus.2021.06.001>
- McBratney, A. B., Mendonça Santos, M. L., ve Minasny, B.** (2003). On digital soil mapping. *Geoderma*, 117(1), 3-52. [https://doi.org/10.1016/S0016-7061\(03\)00223-4](https://doi.org/10.1016/S0016-7061(03)00223-4)
- Mcfeeters, S. K.** (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, 17(7), 1425-1432. <https://doi.org/10.1080/01431169608948714>
- Nguyen, K.-A., Liou, Y.-A., Tran, H.-P., Hoang, P.-P., ve Nguyen, T.-H.** (2020). Soil salinity assessment by using near-infrared channel and Vegetation Soil Salinity Index derived from Landsat 8 OLI data: A case study in the Tra Vinh Province, Mekong Delta, Vietnam. *Progress in Earth and Planetary Science*, 7(1), 1. <https://doi.org/10.1186/s40645-019-0311-0>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., ve Cournapeau, D.** (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 6.
- Racetin, I., Krtalic, A., Srzic, V., ve Zovko, M.** (2020). Characterization of short-term salinity fluctuations in the Neretva River Delta situated in the southern Adriatic Croatia using Landsat-5 TM. *Ecological Indicators*, 110, 105924. <https://doi.org/10.1016/j.ecolind.2019.105924>
- Ribeiro, M. T., Singh, S., ve Guestrin, C.** (2016). “Why Should I Trust You?”: Explaining the Predictions of Any Classifier (arXiv:1602.04938). arXiv. <http://arxiv.org/abs/1602.04938>
- Roscher, R., Bohn, B., Duarte, M. F., ve Garcke, J.** (2020). Explainable Machine Learning for Scientific Insights and Discoveries. *IEEE Access*, 8, 42200-42216. <https://doi.org/10.1109/ACCESS.2020.2976199>
- Rouse, J. W., Haas, R. H., Schell, J. A., ve Deering, D. W.** (1973). Monitoring vegetation systems in the great plains with ERTS. *SP-3511*, 309-317.
- Scudiero, E., Skaggs, T. H., ve Corwin, D. L.** (2014). Regional scale soil salinity evaluation using Landsat 7, western San Joaquin Valley, California, USA. *Geoderma Regional*, 2-3, 82-90. <https://doi.org/10.1016/j.geodrs.2014.10.004>
- Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., ve Batra, D.** (2020). Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization. *International Journal of Computer Vision*, 128(2), 336-359. <https://doi.org/10.1007/s11263-019-01228-7>

- Smilkov, D., Thorat, N., Kim, B., Viégas, F., ve Wattenberg, M.** (2017). *SmoothGrad: Removing noise by adding noise* (arXiv:1706.03825). arXiv. <http://arxiv.org/abs/1706.03825>
- Stadtler, S., Betancourt, C., ve Roscher, R.** (2022). Explainable Machine Learning Reveals Capabilities, Redundancy, and Limitations of a Geospatial Air Quality Benchmark Dataset. *Machine Learning and Knowledge Extraction*, 4(1), Art. 1. <https://doi.org/10.3390/make4010008>
- Stavi, I., Thevs, N., ve Priori, S.** (2021). Soil Salinity and Sodicity in Drylands: A Review of Causes, Effects, Monitoring, and Restoration Measures. *Frontiers in Environmental Science*, 9. <https://www.frontiersin.org/articles/10.3389/fenvs.2021.712831>
- Stomberg, T. T., Stone, T., Leonhardt, J., Weber, I., ve Roscher, R.** (2022). *Exploring Wilderness Characteristics Using Explainable Machine Learning in Satellite Imagery* (arXiv:2203.00379). arXiv. <https://doi.org/10.48550/arXiv.2203.00379>
- Sundararajan, M., Taly, A., ve Yan, Q.** (2017). *Axiomatic Attribution for Deep Networks* (arXiv:1703.01365). arXiv. <http://arxiv.org/abs/1703.01365>
- Thiam, S., Villamor, G. B., Faye, L. C., Sène, J. H. B., Diwediga, B., ve Kyei-Baffour, N.** (2021). Monitoring land use and soil salinity changes in coastal landscape: A case study from Senegal. *Environmental Monitoring and Assessment*, 193(5), 259. <https://doi.org/10.1007/s10661-021-08958-7>
- Wang, F., Yang, S., Yang, W., Yang, X., ve Jianli, D.** (2019). Comparison of machine learning algorithms for soil salinity predictions in three dryland oases located in Xinjiang Uyghur Autonomous Region (XJUAR) of China. *European Journal of Remote Sensing*, 52(1), 256-276. <https://doi.org/10.1080/22797254.2019.1596756>
- Wang, J., Peng, J., Li, H., Yin, C., Liu, W., Wang, T., ve Zhang, H.** (2021). Soil Salinity Mapping Using Machine Learning Algorithms with the Sentinel-2 MSI in Arid Areas, China. *Remote Sensing*, 13(2), Art. 2. <https://doi.org/10.3390/rs13020305>
- Wang, S., Chen, Y., Wang, M., ve Li, J.** (2019). Performance Comparison of Machine Learning Algorithms for Estimating the Soil Salinity of Salt-Affected Soil Using Field Spectral Data. *Remote Sensing*, 11(22), Art. 22. <https://doi.org/10.3390/rs11222605>
- Wang, X., Zhang, F., Kung, H., Johnson, V. C., ve Latif, A.** (2020). Extracting soil salinization information with a fractional-order filtering algorithm and grid-search support vector machine (GS-SVM) model. *International Journal of Remote Sensing*, 41(3), 953-973. <https://doi.org/10.1080/01431161.2019.1654142>
- Wu, W.** (2014). The Generalized Difference Vegetation Index (GDVI) for Dryland Characterization. *Remote Sensing*, 6(2), Art. 2. <https://doi.org/10.3390/rs6021211>
- Yang, L., Driscoll, J., Sarigai, S., Wu, Q., Chen, H., ve Lippitt, C. D.** (2022). Google Earth Engine and Artificial Intelligence (AI): A Comprehensive Review. *Remote Sensing*, 14(14), Art. 14. <https://doi.org/10.3390/rs14143253>
- Yildirim, A.** (2021). *HIGH-RESOLUTION SOIL SALINITY MAPPING USING MACHINE LEARNING BASED REGRESSION AND*

CLASSIFICATION METHODS. ISTANBUL TECHNICAL UNIVERSITY.

- Zarei, A., Hasanlou, M., ve Mahdianpari, M.** (2021). A COMPARISON OF MACHINE LEARNING MODELS FOR SOIL SALINITY ESTIMATION USING MULTI-SPECTRAL EARTH OBSERVATION DATA. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, V-3-2021, 257-263. <https://doi.org/10.5194/isprs-annals-V-3-2021-257-2021>
- Zhou, Y., Chen, S., Hu, B., Ji, W., Li, S., Hong, Y., Xu, H., Wang, N., Xue, J., Zhang, X., Xiao, Y., ve Shi, Z.** (2022). Global Soil Salinity Prediction by Open Soil Vis-NIR Spectral Library. *Remote Sensing*, 14(21), Art. 21. <https://doi.org/10.3390/rs14215627>
- Url-1** <<https://www.fao.org/global-soil-partnership/gasmap/en>>, erişim tarihi 23.12.2022.
- Url-2** <<https://www.usgs.gov/landsat-missions/landsat-8>>, erişim tarihi 21.11.2022.
- Url-3** <<https://www.usgs.gov/faqs/why-are-negative-values-observed-over-water-some-landsat-surface-reflectance-products>>, erişim tarihi 21.11.2022.
- Url-4** <<https://sdgs.un.org/2030agenda>>, erişim tarihi 20.12.2022.

ÖZGEÇMİŞ

Ad-Soyad : Samet Aksoy

ÖĞRENİM DURUMU:

- **Lisans** : 2019, İstanbul Teknik Üniversitesi, İnşaat Fakültesi, Geomatik Mühendisliği Bölümü

MESLEKİ DENEYİMLER:

- 2020 – Devam Ediyor: Araştırma Görevlisi
İTÜ İnşaat Fakültesi Geomatik Mühendisliği Bölümü
- 2019 – 2020: Geomatik Mühendisi
Uydu Haberleşmesi ve Uzaktan Algılama UYG-AR Merkezi

AKADEMİK YAYINLAR

SCI Makaleler

- Aksoy, S.,** Yildirim, A., Gorji, T., Hamzehpour, N., Tanik, A., & Sertel, E. (2022). Assessing the performance of machine learning algorithms for soil salinity mapping in Google Earth engine platform using sentinel-2a and landsat-8 oli data. *Advances in Space Research*, 69(2), 1072–1086. DOI: 10.1016/j.asr.2021.10.024
- Bahşi, K., Ustaoglu, B., **Aksoy, S.,** & Sertel, E. (2023). Estimation of emissions from Crop Residue Burning in Türkiye using remotely sensed data and the Google Earth Engine Platform. *Geocarto International*, 38(1). <https://doi.org/10.1080/10106049.2022.2157052>

Konferans Bildirileri

- Aksoy, S.,** Görücü, Ö. & Sertel, E. (2019). *Drought Monitoring using MODIS derived indices and Google Earth Engine Platform. 2019 8th International Conference on Agro-Geoinformatics (Agro-Geoinformatics)*, Istanbul, Turkey, 2019, pp. 1-6. doi: 10.1109/Agro-Geoinformatics.2019.8820209
- Akçay H. M. , Kaya Ş. , Sertel E. , Algancı U. , **Aksoy S.** (2021) *Sentinel-1 and -2 time-series data-fusion for olive tree identification in heterogeneous land surfaces using Google Earth Engine. 2nd Intercontinental Geoinformation Days (IGD)*, Mersin, Turkey, 2021, pp. 156-162

- Aksoy S.,** Sertel E. (2021) *Comparision of drought monitoring indices derived from MODIS and CHIRPS data using Google Earth Engine*, 9th Global Conference on Global Warming (GCGW-2021), (pp. 12-15)
- Aksoy S.,** Sertel E. (2021) *Comparison of Landsat and MODIS derived vegetation health indices for drought monitoring using Google Earth Engine Platform*, International Symposium on Applied Geoinformatics 2021, Riga, Latvia, 2021, doi: 10.15659/isag2021.12604
- Şengün E., Algancı U. & **Aksoy S** (2022) *Long term monitoring of land use - land cover change and its effect on surface temperature by use of Landsat images and Google Earth Engine Platform*, GIS Ostrava 2022 – Earth Observation for Smart City and Smart Region, Ostrava, Czechia, 2022, doi: 10.31490/9788024846026-8
- Arıkan C., Tümer İ. N., **Aksoy S.,** Sertel E. (2022) *Determination of burned areas using Sentinel-2A imagery and machine learning classification algorithms*, 4th Intercontinental Geoinformation Days (IGD), Tabriz, Iran, 2022, pp. 43 – 46