

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

SİSTEM TANILAMA

MAKİNE ÖĞRENMESİ VE DERİN ÖĞRENME

MODELLERİ İLE GÜNEŞ RADYASYONU TAHMİNİ

YÜKSEK LİSANS TEZİ

Göksel KADER

Enerji Bilim ve Teknoloji Anabilim Dalı

Enerji Bilim ve Teknoloji Programı

ŞUBAT 2022

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

SİSTEM TANILAMA

MAKİNE ÖĞRENMESİ VE DERİN ÖĞRENME

MODELLERİ İLE GÜNEŞ RADYASYONU TAHMİNİ

YÜKSEK LİSANS TEZİ

Göksel KADER

301141015

Enerji Bilim ve Teknoloji Anabilim Dalı

Enerji Bilim ve Teknoloji Programı

Tez Danışmanı: Prof. Dr. Önder GÜLER

ŞUBAT 2022

İTÜ, Lisansüstü Eğitim Enstitüsü'nün 301141015 numaralı Yüksek Lisans Öğrencisi Göksel KADER, ilgili yönetmeliklerin belirlediği gerekli tüm şartları yerine getirdikten sonra hazırladığı “SİSTEM TANILAMA MAKİNE ÖĞRENMESİ VE DERİN ÖĞRENME MODELLERİ İLE GÜNEŞ RADYASYONU TAHMİNİ” başlıklı tezini aşağıda imzaları olan jüri önünde başarı ile sunmuştur.

Tez Danışmanı : **Prof. Dr. Önder GÜLER**

Jüri Üyeleri : **Dr. Öğr. Üy. M. Berker YURTSEVEN**

Doç.Dr. Cenk YAVUZ

Teslim Tarihi : **12 Ocak 2022**

Savunma Tarihi : **08 Şubat 2022**





Aileme,



ÖNSÖZ

Yenilenebilir enerji üretiminde, operasyon ve yatırım maliyetlerini azaltmak için kullanılabilecek tahmin modelleri oluşturularak Türkiye'nin Güneydoğu Anadolu bölgesi için yapılacak yatırımlarda yardımcı olacağını düşünmekteyim.

Bitirme çalışmam boyunca kapısını her çaldığımda yardımlarıyla karşılaştığım, rehberliği ile bana her türlü desteği sağlayan değerli hocam Prof. Dr. Önder GÜLER'e teşekkürü bir borç bilirim. Çalışmam süresince bilgi ve deneyimlerini paylaşarak desteğini esirgemeyen Ömer GÖNÜL'e çok teşekkür ederim. Yüksek lisans eğitimim boyunca beni cesaretlendiren ve desteklerini esirgemeyen FORD OTOSAN ailesindeki dostlarıma teşekkür ederim.

Desteğini hiçbir zaman esirgemeyen ve her daim bana inanan başta annem; Güler DEMİRKOL KADER olmak üzere ailemin bütün üyelerine ve her türlü desteği sağlayan arkadaşlarıma sonsuz teşekkürlerimi sunarım.

Şubat 2022

Göksel Kader

Kontrol ve Otomasyon Mühendisi



İÇİNDEKİLER

Sayfa

ÖNSÖZ.....	vii
İÇİNDEKİLER	ix
KISALTMALAR.....	xi
SEMBOLLER	xiii
ÇİZELGE LİSTESİ.....	xv
ŞEKİL LİSTESİ.....	xvii
ÖZET.....	xix
SUMMARY	xxi
1. GİRİŞ.....	1
1.1 Literatür Araştırması	3
1.2 Tezin Amacı ve Kapsamı	7
2. METODOLOJİ	9
2.1 Sistem Tanılama Modelleri	9
2.1.1 Doğrusal modeller.....	10
2.1.2 Doğrusal olmayan modeller	12
2.1.2.1 Dışsal kaynaklı otoregresif modeli.....	12
2.1.2.2 Hammerstein – Wiener modeli	15
2.2 Makine Öğrenmesi Modelleri	20
2.3 Derin Öğrenme Modelleri	26
2.3.1 Tekrarlayan sinir ağları	27
2.3.2 Uzun kısa süreli hafıza.....	29
3. PERFORMANS METRİKLERİ	33
3.1 R^2	33
3.2 Ortalama Mutlak Hata (MAE)	33
3.3 Kök Ortalama Hata (RMSE)	34
3.4 Normalleştirilmiş Kök Ortalama Hata (nRMSE).....	34
4. SİMÜLASYON SONUÇLARI.....	35
4.1 Meteorolojik Verilerin Hazırlanması	35
4.1.1 Şanlıurfa meteorolojik veri analizi.....	35
4.1.2 Adıyaman meteorolojik veri analizi.....	48
4.1.3 Gaziantep meteorolojik veri analizi	51
4.1.4 Diyarbakır meteorolojik veri analizi	53
4.1.5 Hatay meteorolojik veri analizi	56
4.2 Sistem Tanılama Modelleri	59
4.2.1 Doğrusal modeller.....	59
4.2.1.1 Dışsal kaynaklı otoregresif modeli.....	59
4.2.1.2 Box – Jenkins modeli	65
4.2.2 Doğrusal olmayan modeller	71
4.2.2.1 Dışsal kaynaklı otoregresif modeli.....	71
4.2.2.2 Hammerstein-Wiener modeli	77

4.3 Makina Öğrenmesi Modelleri.....	85
4.4 Derin Öğrenme Modelleri	90
5. SONUÇLAR	93
KAYNAKLAR.....	99
ÖZGEÇMİŞ	103



KISALTMALAR

ADF	: Geliştirilmiş Dickey-Fuller test
ARIMA	: Otoregresif entegre hareketli ortalama
ADF	: Geliştirilmiş Dickey-Fuller test
BJ	: Box-Jenkins
DVR	: Destek vektör regresyonu
DKDO	: Dışsal kaynaklı doğrusal otoregresif
DKDOO	: Dışsal kaynaklı doğrusal olmayan otoregresif
DP	: Doğrusal parçalı
ET0	: Evatranpirasyon
HW	: Hammerstein-Wiener
ÖB	: Ölü bölge
RMSE	: Kök ortalama kare hata
RTF	: Radyal tabanlı fonksiyon
LIBSVM	: Destek vektör makineleri kütüphanesi
TEİAŞ	: Türkiye Elektrik İletim Anonim Şirketi
TSA	: Tekrarlayan sinir ağları
nRMSE	: Normalleştirilmiş kök ortalama kare hata
MAE	: Ortalama mutlak hata
MATLAB	: Matris laboratuvarı
NA	: Uygun değil
UKSH	: Uzun kısa süreli hafıza



SEMBOLLER

n_a, n_b, n_c	: Doğrusal Modellerin Mertebeleri
n_d, n_f	: Doğrusal Modellerin Öteleme Mertebeleri
n	: Giriş Sayısı
n_k	: Gecikme Katsayısı
w	: Ağırlık Katsayısı
b	: Sapma Katsayısı
t	: Zaman
e	: Hata değeri
$\emptyset$	: Aktivasyon Fonksiyonu
P, Q	: Projeksiyon Matrisleri
D_a	: x Uzay Bölümlemeye Aktif Bölümlenmiş Kısım
D_k	: x Uzay Bölümlemesi
J	: Seviye
N	: Ağaçtaki Yaprak Sayısı
L	: Doğrusallık katsayısı
d	: Çıkış Offseti
b	: Sigmoid Genişleme Katsayısı
c	: Öteleme Katsayısı
a	: Çıkış Katsayısı
Q	: Doğrusal Olmayan Alt Uzay
P	: Doğrusal Alt Uzay
ε	: DVR'nin Duyarsız Bölgesinin Genişliği
C	: Maliyetin Ağırlık Faktörü
b	: DVR Bayas Terimi
s_i, s_i^*	: Marjinin Dışında Kalan En Yakın Veriler İçin Gevşek Değişkenler
$\eta_i, \eta_i^*, \alpha_i, \alpha_i^*$	: Lagrange Çarpanları
A	: Sinir Ağı Yığını

C_t	: Yeni Hücre Değeri
$\widetilde{C}_t$	: Yeni Hücre Aday Değeri
$\tanh$	: Tanjant Hiperbolik Operatörü
σ	: Sigmoid Operatörü



ÇİZELGE LİSTESİ

Sayfa

Çizelge 4.1: Ölçüm merkezleri ve veri detayları.....	35
Çizelge 4.2: Pearson matrisi yorumlanması.....	36
Çizelge 4.3: Meteorolojik parametrelerin açıklanması.....	46
Çizelge 4.4: Şanlıurfa meteorolojik parametrelerin analizi.....	47
Çizelge 4.5: Adıyaman meteorolojik parametrelerin analizi.....	50
Çizelge 4.6: Gaziantep meteorolojik parametrelerin analizi.....	53
Çizelge 4.7: Diyarbakır meteorolojik parametrelerin analizi.....	56
Çizelge 4.8: Hatay için meteorolojik parametrelerin analizi.....	58
Çizelge 4.9: Şanlıurfa dışsal kaynaklı otoregresif model sonuçları.....	59
Çizelge 4.10: Adıyaman dışsal kaynaklı otoregresif model sonuçları.....	60
Çizelge 4.11: Diyarbakır dışsal kaynaklı otoregresif model sonuçları.....	61
Çizelge 4.12: Gaziantep dışsal kaynaklı otoregresif model sonuçları.....	62
Çizelge 4.13: Hatay dışsal kaynaklı otoregresif model sonuçları.....	63
Çizelge 4.14: Beş ölçüm merkezi için en iyi DKDO seçimi.....	64
Çizelge 4.15: Şanlıurfa Box-Jenkins model sonuçları.....	65
Çizelge 4.16: Adıyaman Box-Jenkins model sonuçları.....	66
Çizelge 4.17: Diyarbakır Box-Jenkins model sonuçları.....	67
Çizelge 4.18: Gaziantep Box-Jenkins model sonuçları.....	68
Çizelge 4.19: Hatay Box-Jenkins model sonuçları.....	69
Çizelge 4.20: Beş bölge için en iyi Box-Jenkins model seçimi.....	71
Çizelge 4.21: Şanlıurfa DKDOO model sonuçları.....	71
Çizelge 4.22: Adıyaman DKDOO model sonuçları.....	72
Çizelge 4.23: Diyarbakır DKDOO model sonuçları.....	73
Çizelge 4.24: Gaziantep DKDOO model sonuçları.....	74
Çizelge 4.25: Hatay DKDOO model sonuçları.....	75
Çizelge 4.26: Beş ölçüm merkezi için en iyi DKDOO model seçimi.....	76
Çizelge 4.27: Şanlıurfa Hammerstein-Wiener model sonuçları.....	77
Çizelge 4.28: Adıyaman Hammerstein-Wiener model sonuçları.....	79
Çizelge 4.29: Diyarbakır Hammerstein-Wiener model sonuçları.....	80
Çizelge 4.30: Gaziantep Hammerstein-Wiener model sonuçları.....	82
Çizelge 4.31: Hatay Hammerstein-Wiener model sonuçları.....	83
Çizelge 4.32: Beş ölçüm merkezi için en iyi HW model seçimi.....	85
Çizelge 4.33: Şanlıurfa RTF kernel DVR model sonuçları.....	86
Çizelge 4.34: Model 4 RTF kernel DVR doğrulama sonuçları.....	86
Çizelge 4.35: Model 6 RTF kernel DVR doğrulama sonuçları.....	86
Çizelge 4.36: Model 12 RTF kernel DVR doğrulama sonuçları.....	87
Çizelge 4.37: Şanlıurfa sigmoid kernel ile DVR model sonuçları.....	87
Çizelge 4.38: Model 2 sigmoid kernel ile DVR doğrulama sonuçları.....	88
Çizelge 4.39: Şanlıurfa UKSH küme büyüklüğü ayarlaması.....	90
Çizelge 4.40: Şanlıurfa UKSH nöron sayısı ayarlaması.....	90

Çizelge 4.41: Şanlıurfa UKSH epoch sayısı ayarlaması.....	91
Çizelge 4.42: Şanlıurfa UKSH doğrulama sonuçları.....	92
Çizelge 5.43: Şanlıurfa için model doğrulama sonuçları karşılaştırması.	94
Çizelge 5.44: Adıyaman için model doğrulama sonuçları karşılaştırması.	95
Çizelge 5.45: Diyarbakır için model doğrulama sonuçları karşılaştırması.....	95
Çizelge 5.46: Gaziantep için model doğrulama sonuçları karşılaştırması.	96
Çizelge 5.47: Hatay için model doğrulama sonuçları karşılaştırması.....	97



ŞEKİL LİSTESİ

Sayfa

Şekil 2.1: Tek çıkışlı DKDOO model gösterimi (Alcan ve diğ, 2019).	13
Şekil 2.2: Hammerstein-Wiener modeli.	15
Şekil 2.3: Ölü bölge aralıkları.....	17
Şekil 2.4: Doygunluk aralıkları.....	18
Şekil 2.5: Sigmoid fonksiyonu katmanları.	19
Şekil 2.6: Yapay zeka alt uzayları.	26
Şekil 2.7: Basit nöron yapısı (Alam, 2018).	28
Şekil 2.8: Tekrarlayan sinir ağı (Alam, 2018).	28
Şekil 2.9: Yinelenen TSA modülü (Alam, 2018).	29
Şekil 2.10: Basit uzun kısa süreli hafıza yapısı (Lysfjord, 2017).	30
Şekil 2.11: UKSH unutma kapısı (Lysfjord, 2017).	31
Şekil 2.12: UKSH giriş kapısı (Lysfjord, 2017).	31
Şekil 2.13: UKSH yeni bilginin aktarılması (Lysfjord, 2017).	32
Şekil 2.14: UKSH Çıkış kapısı (Lysfjord, 2017).	32
Şekil 4.1: Şanlıurfa verisi pearson matrisi.	37
Şekil 4.2: Şanlıurfa sıcaklık dış ve güneş radyasyonu ilişkisi.	38
Şekil 4.3: Şanlıurfa sıcaklık1 ve güneş radyasyonu ilişkisi.	38
Şekil 4.4: Şanlıurfa rüzgar hızı ve güneş radyasyonu ilişkisi.	39
Şekil 4.5: Şanlıurfa yağış ve güneş radyasyonu ilişkisi.	39
Şekil 4.6: Şanlıurfa azami yağış hızı ve güneş radyasyonu ilişkisi.	40
Şekil 4.7: Şanlıurfa nem hava ve güneş radyasyonu ilişkisi.	40
Şekil 4.8: Şanlıurfa nem1 ve güneş radyasyonu ilişkisi.	41
Şekil 4.9: Şanlıurfa hava basıncı ve güneş radyasyonu ilişkisi.	41
Şekil 4.10: Şanlıurfa sıcaklık toprak ve güneş radyasyonu ilişkisi.	42
Şekil 4.11: Şanlıurfa nem toprak ve güneş radyasyonu ilişkisi.	42
Şekil 4.12: Şanlıurfa ET0 ve güneş radyasyonu ilişkisi.	43
Şekil 4.13: Şanlıurfa sıcaklık yaprak ve güneş radyasyonu ilişkisi.	43
Şekil 4.14: Şanlıurfa yaprak ıslaklık ve güneş radyasyonu ilişkisi.	44
Şekil 4.15: Şanlıurfa çiğlenme ve güneş radyasyonu ilişkisi.	44
Şekil 4.16: Şanlıurfa güneş enerji miktarı ve güneş radyasyonu ilişkisi.	45
Şekil 4.17: Şanlıurfa güneşlenme süresi ve güneş radyasyonu ilişkisi.	45
Şekil 4.18: Şanlıurfa eğitim verisi.	47
Şekil 4.19: Şanlıurfa test verisi.	48
Şekil 4.20: Adıyaman sıcaklık dış ve güneş radyasyonu ilişkisi.	48
Şekil 4.21: Adıyaman nem hava ve güneş radyasyonu ilişkisi.	49
Şekil 4.22: Adıyaman hava basıncı ve güneş radyasyonu ilişkisi.	49
Şekil 4.23: Adıyaman ET0 ve güneş radyasyonu ilişkisi.	50
Şekil 4.24: Gaziantep sıcaklık dış ve güneş radyasyonu ilişkisi.	51
Şekil 4.25: Gaziantep nem hava ve güneş radyasyonu ilişkisi.	51
Şekil 4.26: Gaziantep hava basıncı ve güneş radyasyonu ilişkisi.	52

Şekil 4.27: Gaziantep ET0 ve güneş radyasyonu ilişkisi.....	52
Şekil 4.28: Diyarbakır sıcaklık dış ve güneş radyasyonu ilişkisi.	54
Şekil 4.29: Diyarbakır nem hava ve güneş radyasyonu ilişkisi.	54
Şekil 4.30: Diyarbakır hava basıncı ve güneş radyasyonu ilişkisi.	55
Şekil 4.31: Diyarbakır ET0 ve güneş radyasyonu ilişkisi.....	55
Şekil 4.32: Hatay sıcaklık dış ve güneş radyasyonu ilişkisi.	56
Şekil 4.33: Hatay nem hava ve güneş radyasyonu ilişkisi.....	57
Şekil 4.34: Hatay hava basıncı ve güneş radyasyonu ilişkisi.	57
Şekil 4.35: Hatay için ET0 ve güneş radyasyonu ilişkisi.....	58
Şekil 4.36: Şanlıurfa arx220 model sonucu.	60
Şekil 4.37: Adıyaman arx220 model sonucu.	61
Şekil 4.38: Diyarbakır arx220 model sonucu.	62
Şekil 4.39: Gaziantep arx220 model sonucu.	63
Şekil 4.40: Hatay arx220 model sonucu	64
Şekil 4.41: Şanlıurfa bj20110 model sonucu.....	66
Şekil 4.42: Adıyaman bj20110 model sonucu.	67
Şekil 4.43: Diyarbakır bj20110 model sonucu.	68
Şekil 4.44: Gaziantep bj20110 model sonucu.	69
Şekil 4.45: Hatay bj20110 model sonucu.	70
Şekil 4.46: Şanlıurfa ikili ağaç model sonucu.	72
Şekil 4.47: Adıyaman sigmoid model sonucu.	73
Şekil 4.48: Diyarbakır ikili ağaç model sonucu.	74
Şekil 4.49: Gaziantep sigmoid model sonucu.....	75
Şekil 4.50: Hatay dalgacık model sonucu.....	76
Şekil 4.51: Şanlıurfa giriş DP ve çıkış DP olan HW modeli sonucu.	78
Şekil 4.52: Adıyaman giriş ÖB ve çıkış doygunluk olan HW modeli sonucu.	79
Şekil 4.53: Diyarbakır giriş sigmoid ve çıkış ÖB olan HW modeli sonucu.....	81
Şekil 4.54: Gaziantep giriş DP ve çıkış DP olan HW modeli sonucu.	82
Şekil 4.55: Hatay giriş DP ve çıkış ÖB olan HW modeli sonucu.....	84
Şekil 4.56: Şanlıurfa Model 6 RTF kerneli ile DVR.	88
Şekil 4.57: Şanlıurfa UKSH model sonucu.	91
Şekil 4.58: Şanlıurfa UKSH eğitim ve test kayıpları.....	92

SİSTEM TANILAMA

MAKİNE ÖĞRENMESİ VE DERİN ÖĞRENME

MODELLERİ İLE GÜNEŞ RADYASYONU TAHMİNİ

ÖZET

Dünyadaki teknolojik gelişmelerle birlikte enerji ihtiyacı artmaktadır. Modern toplumda artan insan nüfusu ile birlikte insanların günlük ihtiyaçları ve faaliyetlerini yerine getirilebilmesi büyük ölçüde enerjiye bağımlıdır. Fosil yakıtların sonlu enerjiler olması bilinmektedir ve enerji ihtiyacını karşılamak için yeni ve yenilenebilir enerji kaynaklarına ihtiyaç duyulmaktadır. Yenilenebilir enerji kaynakları olarak güneş enerjisi, rüzgar enerjisi, jeotermal enerji ve hidroelektrik enerji başlıca doğadan elde edilebilen yenilenebilen kaynaklardır. Özellikle güneş enerjisi ve rüzgar enerjisi, konvansiyonel kaynakların neden olduğu küresel ısınmayı azaltırken, elektrik enerji talebinin önemli bir kısmını karşılayabilmektedir. Son yıllarda çevresel kaygılar, teknolojik gelişmeler ve kurulum maliyetlerinin düşmesine paralel olarak rüzgar ve güneş enerjisi kurulumları büyük bir artış göstermiştir. Yapılan yatırımların fizibilitesi açısından rüzgar hızlarının ve güneş radyasyonunun doğru tahmin edilmesi, güç sistemi tasarımcıları ve şebeke operatörleri için son derece önemlidir. Güneş radyasyonu ölçümü oldukça maliyetli ve sensör bakım işlemlerinin operasyon, planlama ve yatırım maliyeti gibi güçlükleri vardır. Bu çalışma kapsamında güneş radyasyonunun tahmini meteorolojik veriler kullanılarak sistem tanılama, makine öğrenmesi ve derin öğrenme tahmin modelleri yardımıyla yapılmıştır.

Çalışma kapsamında 2012-2017 yılları arasında Şanlıurfa, Adıyaman, Diyarbakır, Gaziantep ve Hatay illerindeki ölçüm merkezlerinden alınan 10 dakikalık veri setleri tahmin modellerinde ve doğrulamalarda kullanılmıştır. Veri setindeki sıcaklık, nem, hava basıncı, ET0 ve güneş radyasyonu değişkenleri tez kapsamında yapılan analizler sonucunda modelleme ve doğrulama için en uygun değişkenler olarak seçilmiştir. Veri setindeki eksik veriler doğrusal enterpolasyon ile tamamlanmış ve uygun olmayan veriler elemine edilmiştir. Bu tez kapsamında sistem tanılama, makine öğrenmesi ve derin öğrenme tahmin modelleri Şanlıurfa ölçüm merkezi baz alınarak elde edilmiştir.

İlk olarak doğrusal sistem tanılama modellerinden olan dışsal kaynaklı otoregresif ve Box-Jenkins yöntemleri kullanılarak doğrusal olarak güneş radyasyonu tahmin modelleri oluşturulmuştur. Doğrusal olmayan tahmin modellerinin sonuçları göz önüne alınarak problemin doğrusal olmayan modeller kullanılarak tanımlanmasına karar verilmiştir. Doğrusal olmayan sistem tanılama modellerinden olan dışsal kaynaklı otoregresif modeli ve Hammerstein-Wiener modeli kullanılarak doğrusal olmayan güneş radyasyonu modelleri elde edilmiştir.

İkinci yöntem olarak makine öğrenmesi modellerinden olan destek vektör makinelerin zaman serisi tahmin modellerine uygulanabilen regresyon modeli olan destek vektör regresyonu ile farklı kerneller kullanılarak ve parametreleri optimize edilerek doğrusal olmayan güneş radyasyonu modelleri oluşturulmuştur.

Üçüncü yöntem olarak derin öğrenme yöntemlerinden olan tekrarlayan sinir ağlarının geliştirilmesi ile kullanılmaya başlayan ve güçlü bir doğrusal olmayan modellemeye sahip uzun kısa süreli hafıza modelleri oluşturulmuş, uzun kısa süreli hafıza modeli için hiperparametreler optimize edilmiştir.

Performans metrikleri olarak istatistiki çalışmaların hata analizinde kullanılan R kare, ortalama mutlak hata, kök ortalama kare hata ve normalleştirilmiş kök ortalama kare hata kullanılmıştır.

Bu tez kapsamında oluşturulan modellerin, Gaziantep, Diyarbakır, Adıyaman ve Hatay verileri kullanılarak doğrulamaları yapılmıştır. Doğrulama sonuçlarına bakıldığında performans metrikleri kriterine göre uzun kısa süreli hafıza modelinin en iyi performansı sergilediği görülmüştür.



SOLAR RADIATION FORECASTING USING SYSTEM IDENTIFICATION MACHINE LEARNING AND DEEP LEARNING MODELS

SUMMARY

The world's technological advancements have increased the need for energy. Fossil fuels are known to be finite energies, as a result, alternative and renewable energy sources are needed to meet the energy consumption. New technologies such as electric motor vehicles, bitcoin, and robotics-based manufacturing, for example, are increasing the need for electrical energy. Renewable energy sources such as solar energy, wind power energy, geothermal energy and hydroelectric energy are the sources which can be mainly obtained from nature. Solar radiation measurement is quite expensive and sensor maintenance comes with its own set of issues, including operational, planning, and investment expenses.

Within the scope of this thesis, the importance of time series forecasting models for solar radiation, which is highly critical for the accuracy of measurement and forecast in solar plants, are emphasized. Time series forecasting is a technique which is used to predict events through a time sequence. This technique predicts future events by analyzing past trends with the assumption that future trends will be similar to previous trends. By their nature, meteorological variables are time-dependent variables.

For Sanliurfa, Adıyaman, Diyarbakır, Gaziantep, and Hatay, data with 10-minute observations acquired from Turkey's Meteorological Services between 2012 and 2017 were utilized. The data set has 19 separate variables. The correlations of variables with sun radiation were investigated using the Pearson's Correlation matrix and the relations of the variables with the solar radiation were analyzed by figures. Temperature, humidity, air pressure, ET0, and solar radiation data were used in the thesis as a consequence of the analyses. Linear interpolation was used to fill in the missing data in our data set, and improper data was eliminated.

Forecast models were acquired based on Sanliurfa's data and those forecast models were applied using the variables selected before for Adıyaman, Diyarbakır, Gaziantep and Hatay data. Those forecast models were validated for the Adıyaman, Diyarbakır, Gaziantep, and Hatay. Solar radiation estimation was discussed with three different methods commonly used in the estimation of time series.

As the first method, system identification forecast models were produced with using MATLAB Identification Toolbox. First of all, the autoregressive model with exogenous input, which is one of the linear system identification models which produces better results thanks to external input, was used than the classical autoregressive models used with large data. By varying polynomial degrees systematically, ten alternative autoregressive models with exogenous input were created.

The Box-Jenkins model was then utilized, which is a linear system identification model derived from autoregressive integrated moving average models. Changing the polynomial degrees yielded ten different Box-Jenkins models.

It was determined to use the problem with nonlinear models as a result of the outputs of linear models. Initially, a nonlinear system identification model called an autoregressive model with exogenous input was utilized. Using tree partition, sigmoid, and wavelet network activation functions, as well as modifying polynomial degrees, three nonlinear autoregressive with exogenous input models were created for forecasting solar radiation. Second, models were generated using the Hammerstein-Wiener model, which combines linear and nonlinear structures from nonlinear system identification models to create models. For input and output non-linearity, all combinations of piecewise linear, sigmoid, saturation, dead zone, and wavelet network activation functions are constructed, yielding twenty-five alternative forecast models.

For the second method, forecast models were created with the regression model support vector regression, which can be applied to the time series prediction models of support vector machines, which is one of the machine learning models, by using the Python programming language. Three kernels were employed to give the nonlinear structure: radial-based function, sigmoid function, and polynomial function. the C , γ and ε parameters for the radial basis function kernel were optimized systematically to generate 12 forecast models. For the sigmoid kernel, 10 prediction models were created by optimizing the C and γ parameters. by making optimizations for the polynomial kernel, models were generated, however the results were not given due to the fact that the polynomial kernel's inability to converge. For the C parameter, the values range of 0.01, 0.1, 10 and 1000 were utilized. For the γ parameter, the values range of 0.001, 0.01, 0.1 and 1 were utilized. For the ε parameter, the values range of 0.001 and 0.01 were utilized.

The third method is the long-short term memory model, which is one of the deep learning methods and developed over recurrent neural networks, with a strong nonlinear modeling background. Hyperparameters such as neuron size, epoch size, and batch size were tuned. Initially, ten models were developed with batch size parameters ranging from 1 to 256 with all other characteristics remaining constant. It was discovered that when the Batch size was set to 8, the best results were obtained. Second, nine models were developed, with the neuron size ranging from 1 to 128 and all other characteristics remaining constant. When the Neuron sizes was set to 32, the best results were obtained. Finally, eight models were developed, with epoch sizes varying from 4 to 256 and all other parameters remaining constant. When the epoch size was set to 256, the best results were obtained. The hyperparameters for batch size, neuron size, and epoch size were identified as 8 for batch size, 32 for neuron size, and 256 for epoch size as a result of the optimization research.

R square is a statistical measure that represents the proportion of the variance for a dependent variable that's explained by an independent variable or variables in a regression model., mean absolute error (MAE) is measures the average magnitude of the errors in a set of forecasts, without considering their direction, root mean square error (RMSE) is a quadratic scoring rule which measures the average magnitude of the error., and normalized root mean square error (nRMSE) is the comparison between datasets or models with different scales were employed as performance metrics for comparing forecast model errors.

According to Şanlıurfa validation results, the best result in linear system identification models is Box-Jenkins with 0.91314 R square, the best result in nonlinear system

identification models is Hammerstein-Wiener with 0.95546 R square, the best result in machine learning model is 0.8890 R square and the best result in deep learning model is 0.8890 R square

Adiyaman validation results were seen that Box-Jenkins reached 0.91154 R square in linear system identification models, Hammerstein-Wiener reached 0.9339 R square in nonlinear system identification models, machine learning model reached 0.8764 R square and LSTM model reached 0.9723 R square as the best forecast results.

Diyarbakir validation results were showed that the best forecast result was Box-Jenkins with 0.92292 R square in linear system identification models, Hammerstein-Wiener with 0.94438 R square in nonlinear system identification models, machine learning model with 0.8751 R square, and LSTM model with 0.9680 R square.

Gaziantep validation results were obtained that Box-Jenkins achieved 0.92322 R square in linear system identification models, Hammerstein-Wiener achieved 0.94333 R square in nonlinear system identification models, machine learning model achieved 0.8289 R square, and LSTM model achieved 0.9614 R square as the best forecast result.

Hatay validation results were observed that Box-Jenkins achieved 0.9306 R square value in linear system identification models, Hammerstein-Wiener achieved 0.93641 R square value in nonlinear system identification models, machine learning model achieved 0.8159 R square value, and LSTM model achieved 0.9763 R square value as the best forecast result.



1. GİRİŞ

Enerji kavramı Leibniz'in bir maddenin kütlesi ve onun yer çekimi alanı olarak tanımladığı yaşam gücünden gelmektedir. İnsanların ilk enerji kaynağı olan ateş ve onun en önemli yakıtı odun olmuştur. Teknolojinin gelişmesi ile enerji kaynakları da çeşitlenmeye başlamıştır. 1700'lü yıllarda başlayan sanayi devrimi enerji ihtiyacının artmasını ve çeşitliliğin çoğalmasını sağladı. 1860'lı yıllarda Amerika'nın petrolü bulması ile fosil yakıtlarla beraber yeni bir döneme geçildi. 1880'li yıllarda elektrik enerjisinin bulunması enerji alanında önemli adımlardan biri olmuştur. 1900'lü yılların başlarında içten yanmalı motorların kullanılmaya başlaması ve yeni makinaların icat edilmesi ile fosil yakıtların kullanımı oldukça artmıştır. 1940'lı yıllarda Amerika'da ilk nükleer reaktör devreye alınmıştır. Fosil yakıtların sonlu olması ve çevreye verdikleri zararları gözönünde bulundurularak ve nükleer enerji reaktörlerinde çevreye verdiği zarar ve risklerinin yüksek olması sebebiyle yeni ve yenilenebilen enerji kaynakları aranmaya başlamıştır.

Yenilenebilir enerji, doğada kendini tekrarlayabilen ve süreklilik sağlayabilen enerji kaynaklarıdır. Yenilenebilir enerjiler; güneş, rüzgar, biyokütle, jeotermal, dalga enerjisi gibi doğada kendiliğinden var olan kaynaklardan elde edilmektedir. Bu kaynaklar fosil yakıtların tersine enerji ihtiyacını daha temiz ve sürdürülebilir olarak sağlayabilmektedir. Rüzgar enerjisi, rüzgar bakımından uygun coğrafi alanlara kurulan turbinlerle havanın kinetik enerjisi elektrik enerjisine dönüştürülerek elde edilir. Biyokütle enerjisi; hayvan ve insan atıkları, atık sular ve tarım atıkları gibi organik atıklardan anaerobik fermentasyonla elde edilen enerjinin elektrik enerjisine dönüştürülmesi ile kullanım alanı oluşturmaktadır. Jeotermal enerji; jeotermal kaynakların bulunduğu bölgelerde turbinler ile elektrik enerjisi üretiminde bulunmaktadır. Tüm yenilenebilir enerji kaynaklarının ortak özelliği çevre dostu olmaları, sürdürülebilir olmaları ve yerlilik oranlarının yüksek olmasıdır. Güneş enerjisi ise toplayıcılar ile direkt yada dolaylı olarak güneş radyasyonunu toplayarak fotovoltaik pillerle elektrik enerjisine dönüştürülmesidir. Güneş enerjisi, kaynağı

güneş olan ısı ve radyasyondur. Güneş'in çekirdeğinde yer alan füzyon süreci ile açığa çıkan ışıyım enerjisidir. Güneş oluşumunun temel kaynağı hidrojen atomunun helyuma parçalanmasıdır. Güneş'in yüzeyinde güneş radyasyonunun yoğunluğu yaklaşık $6,33 \times 10^7 \text{ W/m}^2$ dir. Ancak yeryüzüne $0-1100 \text{ W/m}^2$ arası güneş radyasyonu ulaşmaktadır (YEGM, 2015). Açık hava da öğlen vaktinde 1100 W/m^2 kadar ulaşabilmektedir.

1954 yılında Amerika'da Bell laboratuvarlarında yapılan çalışmalar ile önceleri %4, daha sonrasında ise %11 verim ile çalışabilen silikon güneş hücreleri üretilmiştir ve bu konudaki devrim niteliğindeki çalışmalardan olmuştur. 1960 yılında Hoffman elektronik firması %14 verime ulaşmıştır. 1977 yılında dünyada toplam kurulu güneş enerjisi 500kW 'a ulaşmıştır. 1982 yılında dünyada toplam kurulu güneş enerjisi 10MW 'a ulaşmıştır. 1992 yılında %15 verimle çalışan ince film fotovoltaik güneş hücreleri kullanılmaya başlanmıştır. 1999 yılında dünyada kurulu güneş enerjisi 1000MW 'a ulaşmıştır (Durak, 2016). Günümüzde PV hücre verimleri %5 ile %20 arasında değişmektedir ve %15 verimlilik iyi olarak değerlendirilmektedir (Öztürk, 2014). Günümüzde TEİAŞ'ın verilerine göre, Ağustos 2021 dönemi itibariyle Türkiye'de 8064 kurulu güneş enerjisi santrali ile $7435,2 \text{ MW}$ 'lık kurulu güç mevcuttur. Bununla birlikte Türkiye'de güneş enerjisinin toplam kurulu güç içindeki payı %7,5'e yükselmiştir. Enerji, günümüzde ülkelerin gelişmişlik düzeyinin önemli bir göstergesidir. Türkiye, 2010'da yayınlanan yenilenebilir enerji kaynakları yasası ile güneş enerjisi kullanımını arttırılmasına yönelik kararlar almıştır.

Güneş radyasyonu, fotovoltaik enerji üretimini etkileyen en önemli faktördür (Lave ve diğ, 2010). Güneş enerjisinin doğru tahmini, güneş enerjisi kayıplarıyla ilişkili entegrasyon ve işletme maliyetlerini azaltarak güneş enerjisi üretiminin şebekeye entegrasyonunu kolaylaştırır (Chow ve diğ, 2011). Güneş enerjisinin doğru tahmini, fotovoltaik güneş hücreleri sistemlerinin verimini artıracak gibi, Türkiye'de yapılacak güneş santrali yatırımlarının doğru yönlendirilmesini sağlamaktadır.

Güneş enerjisi ölçümünde kullanılan ekipmanlarda kalibrasyon hatası, elektrik kesintisi ve yanlış konumlandırma gibi sebeplerle kayıplar yaşanmaktadır. Güneş enerjisinin ölçümünde kullanılan ekipmanların yüksek maliyeti ve bakımlarının yüksek maliyetli olması sebebiyle güneş enerjisi tahmin modellerine ihtiyaç

duyulmaktadır. Bu nedenle, özellikle daha az ölçüm kullanabilen etkili güneş radyasyonu tahmin yöntemleri geliştirmek çok önemlidir. Güneş radyasyonu dizileri, rastgele süreçler tarafından üretilen bir zaman dizisi olarak ele alınabilir; bu nedenle matematiksel modeller, temel süreçlere uyum sağlamak ve gelecekteki değerleri tahmin etmek için kullanılabilir (Wei, 2017; Langella ve diğ, 2016).

1.1 Literatür Araştırması

Bu bölümde, güneş radyasyonu verileri için zaman serisi tahmin yöntemlerini kullanan ilgili çalışmaları inceleyeceğiz. Güneş radyasyonu tahmininde doğrusal ve doğrusal olmayan sistem tanımlama modelleri kullanıldığı gibi makine öğrenme modelleri ve derin öğrenme modelleri de sıkça kullanılmıştır. Zaman serisi modelleri, ekonometri biliminden ilham alan ve kökleri istatistikteki gelişmelere dayanan modellerdir (Ljung, 2010).

Güneş radyasyonunun tahmini çeşitli meteorolojik değişkenlerden etkilenmektedir. Güneş radyasyonu ile çeşitli meteorolojik değişkenler arasındaki bağlantıyı göstermek için birçok çalışma yapılmıştır. Yapılan çalışmalar da güçlü bağlantılar elde edilmiştir. Sfetsos ve Coonick çalışmalarında güneş radyasyonu tahmini için tek değişkenli ve çok değişkenli yaklaşımları karşılaştırmışlardır. Modellerinde sıcaklık, basınç, rüzgar hızı ve rüzgar yönü gibi ek parametreler kullanılmış ve verinin durumuna göre farklı parametrelerde güneş radyasyonunun tahminini arttırılabileceğini göstermişlerdir (Sfetsos ve Coonick, 2000). Chen ve Li, destek vektör makineleri modellerinde değişkenler olarak güneşlenme süresi, sıcaklık, nem ve basıncı kullanmışlardır ve sıcaklık ile oldukça önemli bir etkileşimin olduğunu belirtmişlerdir (Chen ve Li, 2016).

Kolombiya, Bogota'daki yüksek teknolojili bir izleme istasyonunda elde edilen radyasyonu ve sıcaklık kalıpları üzerinde dışsal girdiler ile otoregresif model uygulamasının sonuçlarını ve analizini sunmaktadır. Güneş radyasyonu tanımlama ve tahmini için Chica'nın (2012) yaptığı dışsal kaynaklı otoregresif model çalışmalarında yedi yıllık bir süre için elde edilen sonuçlar ve Bogata, Kolimbiya için izleme istasyonundan alınan gerçek güneş radyasyonu değeri ile %84 korelasyonlu olarak tahmin edildiği gözlemlenmiştir.

Box-Jenkins modelini zaman serisinde güneş radyasyonu tahmininde kullanılmasının oldukça uygun olduğu düşünülmektedir ve zaman serisi tarafından oluşturulan bir verinin önceki değerlerine göre davranışını tahmin etmek için uygun bir model olduğu araştırmalarda görülmüştür (Taneja ve diğ, 2016).

Box-Jenkins modeli Malezya'da Mersing ve Kuala Terengganu olmak üzere iki farklı lokasyondan 2000 ile 2007 yılları arasında alınan güneş radyasyonu verilerine uygulanmıştır. Ljung-Box istatistiği modellerin yeterli sonuçlar verdiğini göstermiştir. Kuala Terengganu için kök ortalama kareler hatası değeri olarak 0,047 ve Mersing için 0,044 olarak tahmin edilmiştir. Güneş radyasyonu bilgisi birçok gelişmekte olan ülke için kolayca mevcut olmadığından, elde edilen modeller gelecekteki güneş radyasyonu verilerinin tahmin edilmesine de yardımcı olabileceği belirtilmiştir (Kamil ve Razali, 2015).

Yatay bir yüzey üzerinde doğrudan güneş radyasyonu tahmini için 10 günlük ve 30 dakikalık aralıklarla alınan bir veriseti ile doğrusal olmayan dışsal kaynaklı otoregresif modeli önerilmektedir. Değişik değerlendirme kriterleri ile simulasyonlar yapılmış ve güneş radyasyonu tahmini için ortalama kareler hatası olarak 0,00279 hata değerine ulaşılmıştır (Boussada ve diğ, 2018).

Mutah, Ürdün için 2015 yılına ait sıcaklık, rüzgar hızı ve nem gibi hava durumu verilerini kullanarak doğrusal olmayan dışsal kaynaklı otoregresif modelin doğrulanması ve valide edilmesini sağlamıştır. Farklı girdi değişken kombinasyonlarına sahip yedi doğrusal olmayan dışsal kaynaklı otoregresif modeli, ortalama kare hata değeri açısından incelenmiştir. Sadece sıcaklık verisi kullanılarak ortalama kare hatanın 0,00269 olduğu ve yedi model arasında en iyisi olduğu gözlemlenmiştir. Buna ek olarak sıcaklık, rüzgar hızı ve nem verileri beraber kullanıldığında ortalama kare hatanın 0,00273 olduğu gözlemlenmiştir (Al-Sbou ve Alawasa, 2017).

Bhubaneswar, Hindistan için 2002-2006 arası dört yıllık veri ile güneşlenme süresi, sıcaklık ve nem verileri kullanarak doğrusal olmayan dışsal kaynaklı otoregresif model ile güneş radyasyonu tahminleri yapmıştır. Kullanılan üç tane doğrusal olmayan dışsal kaynaklı otoregresif model arasında en iyi sonuca ortalama kare hata değerinin 0,0094 olduğu durumda ulaşılmıştır (Mohanty ve diğ, 2015).

Ayaş, Ankara, Türkiye için bir yıllık sıcaklık, nem ve yağış verilerini kullanılarak doğrusal olmayan dışsal kaynaklı otoregresif model ile güneş radyasyonu tahminleri yapmıştır. Sıcaklık verisi kullanılarak 0,00457 ortalama kare hatası ile en iyi sonuç elde edilmiştir. Nem verisi için 0,00507 ortalama kare hatası ve yağış verisi için 0,0056 ortalama kare hatası elde edilmiştir (Sarı ve diğ, 2021).

Lityum iyon pil hücrelerinin dinamik davranışı için doğrusal olmayan bir model üzerine yapılan çalışmada Hammerstein-Wiener modeli kullanılmıştır. Dört adet model de farklı doğrusalsızlık fonksiyonları kullanılarak sonuçlar elde edilmiştir. Giriş için sigmoid aktivasyon fonksiyonu ve çıkış için dalgacık aktivasyon fonksiyonu kullanılan tahmin modeli sonuçları ortalama kare hata olarak 0,0010368, giriş için parçacıklı doğrusal aktivasyon fonksiyonu ve çıkış için dalgacık aktivasyon fonksiyonu kullanılan tahmin modeli sonuçları ortalama kare hata olarak 0,009902, giriş için parçacıklı doğrusal aktivasyon fonksiyonu ve çıkış için polinom aktivasyon fonksiyonu kullanılan tahmin modeli sonuçları ortalama kare hata olarak 0,004988, giriş ve çıkış için polinom aktivasyon fonksiyonu kullanılan model sonuçları ortalama kare hata olarak 0,0002777 gösterilmiş ve en iyi performansa ulaşılmıştır (Khalfi ve diğ, 2020).

Beton baca modelinin yer hareketleri ve yapısal tepki ölçümleri ile Hammerstein-Wiener modeli kullanılarak tahmin modeli elde edilmiştir. Beton baca modelinin doğrusal olmayan sistem tanımlaması sonuçları, tahmin verilerine uyumun %100 olduğunu göstermiştir. Hammerstein-Wiener sistem tanımlama yönteminin beton baca matematiksel modelini tanımlamada etkili ve doğru yöntem olduğu sonucuna varılmıştır (Tuhta ve diğ, 2020).

New alipore, Kolkata, Hindistan için iki yıllık (2011-2013) sıcaklık, hava basıncı ve nem verileri kullanılarak destek vektör regresyonu yönteminin güneş radyasyonunun tahmini için kullanılması ve doğruluğu üzerine bir çalışma yapmıştır. Polinom ve radyal tabanlı fonksiyon kernelleri kullanılmıştır. Elde edilen sonuçlar, RTF kerneli ile eğitim için R^2 değeri 0,7976 ve RMSE değeri 1,0564 iken, test için R^2 değeri 0,7845 ve RMSE değeri 1,0532'dir. Polinom çekirdeği eğitim için R^2 değeri 0,3993 ve RMSE değeri 1,1975 ulaşırken, test için R^2 değeri 0,9060 ve RMSE değeri 1,1594'ulaşılmıştır. Polinom kerneline sahip DVR, ele alınan çalışma alanı için RTF

çekirdeğinden daha yüksek performans sağladığı gözlemlenmiştir (Shaw ve Prakash, 2019).

Gurugram, Hindistan için 3 yıllık (2009-2011) sıcaklık, hava basıncı ve nem verileri kullanılarak destek vektör regresyonu yöntemi ile güneş radyasyonu tahmini yapılmıştır. Radyal tabanlı fonksiyon kerneli performansının diğer destek vektör regresyonunda kullanılan kernellerden daha başarılı olduğu görülmüştür. Destek vektör regresyonu sonuçlarının kök ortalama kareler hatası değeri olarak 0,143 olarak gösterilmiştir (Bhola ve Bhardwaj, 2019).

Sharifzadeh ve diğ (2019) Canterbury, İngiltere için 29 yıllık (1985-2014) veri kullanarak makine öğrenmesi modellerinden destek vektör regresyonunun güneş tahmininde kullanılmasını ve kerneller arasında karşılaştırmalar yapılmıştır. Polinom kernelinin yakınsayamadığı ve güneş veri setinin temel modelini yakalayamadığı görülmüştür. Radyal tabanlı fonksiyon kernelin 0,015 ortalama hata değeri ile tahminde bulunduğu görülmüştür.

Ajacio, Fransa için 1998 yılının Ocak ayı ile 2007 yılının Aralık ayı arasındaki sıcaklık, nem, bulutluluk, güneşlenme süresi, hava basıncı ve yağmur verilerinden oluşan 87600 örnek ile uzun kısa dönemli hafıza metodu kullanılarak farklı girişler ve farklı modeller ile güneş radyasyonu tahminleri yapmıştır. Güneş radyasyonu, sıcaklık, bulutluluk ve nem verileri ile oluşturulan modelin tahmin sonuçlarının en iyi performansa sahip olduğu ve normalize kök ortalama hata değerinin 0,159261 olduğu görülmüştür. Ayrıca çalışmanın içeriğinde tek değişkenli ve çok değişkenli modeller arasındaki tahmin değerleri karşılaştırılmış ve güneş radyasyonun diğer meteorolojik verilere oldukça bağlı bir değişken olduğu ve tahmininde mümkün olan diğer meteorolojik verilerin kullanılması gerektiği belirtilmiştir (Sorkun ve diğ, 2020).

Gifu, Japonya'nın 2014, 2015 ve 2016 yıllarının Ocak ayından Aralık ayına kadar olan sıcaklık, rüzgar hızı, rüzgar yönü ve nem verileri ile uzun kısa dönemli hafıza methodu kullanılarak güneş radyasyonu tahminleri yapılmıştır. Yapılan tahmin modeli ve çıktılar arasındaki ilk üç aylık dönem için yapılan karşılaştırmada normalize kök ortalama hata değeri 0,28 olarak elde edilmiştir. Bu çalışmada hava durumunun önemi üzerinde durulmuş ve yağmurlu günlerde, güneşli günlerde ve bulutlu günlerde sonuçlarda farklılık oluşabileceği gösterilmiştir (Arujo, 2020).

LI-200S pironometre ile 10 ms'lik alınan 4 günlük 20736000 örnek ile uzun kısa dönemli hafıza ve destek vektör regresyonu yöntemleri ile güneş radyasyonu tahminleri yapılmıştır. Destek vektör regresyon tahmin modeli için kök ortalama hata değeri 0,11 ve uzun kısa süreli hafıza tahmin modeli için kök ortalama hata değeri 0,086 olarak elde edilmiştir. Uzun kısa süreli hafıza yönteminin büyük verilerde diğer yöntemlere göre daha başarılı olduğu gösterilmiştir. Güneş radyasyonu verilerinin örnekleme periyodu ve hacmi göz önüne alındığında büyük veri kapsamında değerlendirilmektedir (Alzahrani ve diğ, 2017).

1.2 Tezin Amacı ve Kapsamı

Bu tez çalışmasında Türkiye'nin güneş enerjisi potansiyelinden daha verimli ve etkin bir şekilde yararlanılabilmesi için meteorolojik veriler kullanılarak Güneydoğu Anadolu bölgesi için uygulanabilecek zaman serisi tahmin modellerinin oluşturulması amaçlanmaktadır. Güneş enerjisi ölçümlerinde ekipmanların ve operasyonların maliyetinin yüksek olması tahmin modellerine ihtiyacı arttırmaktadır. Güneş radyasyonu tahmin modelleri oluşturulurken zaman serisi tahmin modellerinde sıkça kullanılan yöntemler olan sistem tanılama, makine öğrenmesi ve derin öğrenme yöntemleri kullanılmaktadır. Bu yöntemler ile Şanlıurfa verileri baz alınarak oluşturulan tahmin modellerinin Gaziantep, Diyarbakır, Adıyaman ve Hatay verileri ile doğrulamaları yapılmış ve sonuçlar karşılaştırılarak en iyi performansı veren tahmin modelleri belirlenmiştir.

Öncelikle doğrusal sistem tanılama modelleri çok değişkenli olarak oluşturulmakta ve optimize edilen tahmin modellerinin doğrulamaları yapılmıştır. Literatür araştırmalarında güneş radyasyonu tahmininde Hammerstein-Wiener modellerin yaygın olarak kullanılmadığı görülmüştür. Tez kapsamında giriş ve çıkış doğrusalsızlığının oluşturabileceği kombinasyonlar üzerinde çalışılarak çok değişkenli Hammerstein-Wiener modelleri oluşturulmuştur. Oluşturulan modellerin doğrulamaları yapılarak en iyi performansa ulaşan giriş ve çıkış doğrusalsızlığı kombinasyonu elde edilmiştir. HW modelleri, sistem tanılama tahmin modellerinden bir başkası olan dışsal kaynaklı doğrusal olmayan otoregresif modeli tahmin sonuçları ile karşılaştırılmıştır.

İkinci olarak makine öğrenmesi tahmin modellerinden olan DVR modelleri farklı kernel fonksiyonları ile oluşturulup, doğrulamaları yapılmıştır. DVR tahmin modelleri literatür çalışmalarında genellikle veri sayısı yüksek olmayan tahmin problemlerinde sıkça kullanıldığı görülmüştür. Tez kapsamında yüksek veri sayısı ile hiperparametreler optimize edilmiştir.

Üçüncü olarak derin öğrenme modellerinden UKSH modelleri oluşturulmuştur. Literatür araştırmalarında tek değişkenli tahmin modelleri sıkça kullanılırken, çok değişkenli tahmin modelleri daha az kullanılmış ve doğrulama aşamalarına sıkça rastlanmadığı görülmüştür. Tez kapsamında çok değişkenli UKSH modelleri optimize edilerek ve doğrulamalar yapılmıştır.

Tez kapsamında tahmin modellerinin metodolojisine ve simulasyon sonuçlarına yer verilmiştir.

2. METODOLOJİ

2.1 Sistem Tanılama Modelleri

Sistem tanılama, sistemin giriş ve çıkış sinyallerinin ölçümlemlerini kullanarak dinamik sistemlerin matematiksel modellerini oluşturmaya yönelik bir metodolojidir. Gerçek dünya uygulamaları ile kontrol teorisi ve model soyutlamalarının matematiksel dünyası arasındaki arayüz olarak görülebilir. Sistem tanılama, tahmin yapabilecek modellerin karakterine bağlı olarak farklı tekniklerle çok geniş bir konudur: doğrusal, doğrusal olmayan, hibrit, parametrik, parametrik olmayan, vb. Aynı zamanda, sistem tanılama alanı az sayıda öncü ilke ile karakterize edilebilir. Örneğin model karmaşıklıklarına uygun sürdürülebilir açıklamalar, verilerdeki bilgi içerikleri ve etkili doğrulama olarak gösterilebilir (Ljung, 2010).

Sistem tanılama, verilerin değişiminden oluşturulan iteratif modellerdir. Basit model yapılarından karmaşık yapılara doğru iteratif olarak devam eden karmaşıklık değeri ile en iyi sonucu veren basit model oluşturulur. Yüksek dereceli polinom değerleri her zaman en iyi sonucu vermeyebilir. Karmaşıklığın artması ile daha çok veriye ihtiyaç duyulabilir ve parametrelerdeki belirsizlik artabilir (Ljung, 2008a,b).

Sistem tek giriş-tek çıkışlı ise SISO, çok giriş bir çıkışlı ise MISO, çok girişli-çok çıkışlı ise MIMO sistem olarak adlandırılır. Deneysel verileri kullanarak dinamik sistemlerin matematiksel modellerini oluşturma sanatıdır. Sistem tanılama, iteratif bir metodolojidir.

Sistem tanılama da fiziksel denklemler ve ideal bileşenlerle oluşturulan modeller beyaz kutu modelleri olarak adlandırılır. Fiziksel denklemler ve verilerin kullanımı ile denklemlerde eksik kalan kısımların tamamlandığı modeller gri kutu modeller olarak adlandırılmaktadır. Giriş ve çıkış bilgisi kullanılarak oluşturulan ve bilinen matematiksel denklemleri olmayan modeller kara kutu modeller olarak adlandırılmakta ve tahmin modellerinde sıkça kullanılmaktadır.

2.1.1 Doğrusal modeller

2.1.1.1 Dışsal kaynaklı otoregresif modeli

Otoregresif modeller, veride bulunan gözlem değerinin ondan önceki tanımlanmış sayıdaki dönemin gözlem değerinin ve hata teriminin doğrusal olarak ifade etmektedir. Aynı değişkenin tanımlı bir zaman dönemindeki gözlem değeri ile ondan önceki dönemlere ait gözlem değerleri arasındaki doğrusal ilişkiyi tanımladığı için otoregresif model olarak adlandırılır. Zaman serisi kendi gecikmeli değerlerinin bir fonksiyonu ile ifade edilir. Bağımsız değişkenler $y(t - 1 \dots n)$ aynı değişkenin önceki dönemlere ait gecikmiş değerleridir. Bozucular dinamik modele bağlanmıştır ve bozucular ile dinamiği birbirinden bağımsız olarak modellenemez. Dışsal değişkenler, bağımsız değişkenler değildir. Dışsal değişkenler, modellemekte olduğunuz sürecin dışında belirlenir. Dışsal kaynaklı otoregresif model denklem 2.1'deki gibi ifade edilmektedir.

$$\begin{aligned} y(t) + a_1(t - 1) + \dots + a_{na}y(t - na) \\ = b_1u(t - nk) + \dots + b_{nb}u(t - nb - nk + 1) + e(t) \end{aligned} \quad (2.1)$$

Burada $y(t)$ t anındaki çıkış değeri, n_a kutup sayısı, n_b sıfır sayısı, n_k sistemdeki ölü zaman, $u(t - n_k) \dots u(t - n_k - n_b + 1)$ önceki ve geciktirilmiş zaman, $e(t)$ bozucu hatası olarak tanımlanmaktadır.

Dışsal kaynaklı otoregresif model denkleminin daha kompakt ifadesi 2.2 denklemindeki gibi ifade edilir.

$$A(q)y(t) = B(q)u(t - n_k) + e(t) \quad (2.2)$$

2.1.1.2 Box-Jenkins modeli

1970 yılında İki ünlü istatistikçi olan George E. Box (Wisconsin Üniversitesi) ve Gwilym M. Jenkins (Lancaster Üniversitesi) çalışmaları ile oluşmuştur. Box-Jenkins gelecek tahmini için kullanılan ARIMA bazlı bir modeldir. Box-Jenkins modellemesi, uygun bir ARIMA sürecini tanımlamayı, onu verilere uydurmayı ve ardından tahmin için uygun modeli kullanmayı içerir. Box-Jenkins'in tahmine yaklaşımının önemli özelliklerinden biri, ARIMA süreçlerinin zengin bir olası model sınıfı olmasıdır.

Box-Jenkins modelleri için eşit zaman aralıkları ile alınan verinin durağan olması oldukça önemli bir parametre olmasına rağmen modelin yaygın olarak kullanılmasının nedeni ise ele alınan herhangi bir seri durağan olsun veya olmasın, mevsimsellik içersin veya içermesin, kullanılan denklemlerle bir sonuca kavuşturma yeteneğidir. Zaman serilerinin yapısı veya akışı ilgili önceden bilgiye ihtiyaç duymaması sebebiyle karmaşık zaman serilerinde kullanılması uygundur. Geçmiş verileri açıklayıcı değişken olarak kullanılmaktadır. Zaman serilerindeki verilerin zaman içinde aldığı değerler arasında iç bağımlılık etkili bir şekilde kullanılmaktadır. Zaman serilerini yalnızca kendi geçmiş değerleri ve olasılıksal hata terimi ile açıklamaktadır. Genel Box-Jenkins yapısı denklem 2.3'te gösterilmiştir.

$$y(t) = \sum_{i=1}^{nu} \frac{B_i(q)}{F_i(q)} u_i(t - nk_i) + \frac{C(q)}{D(q)} e(t) \quad (2.3)$$

nu giriş sayısını ve nk ise gecikmeyi ifade etmektedir. Box-Jenkins modelindeki mertebeler aşağıdaki gibi tanımlanabilir;

$$n_b: B(q) = b_1 + b_2 q^{-1} + \dots + b_{n_b} q^{-n_b+1} \quad (2.4)$$

n_b mertebesi, denklem 2.4'te gösterildiği gibidir.

$$n_c: C(q) = 1 + c_1 q^{-1} + \dots + c_{n_c} q^{-n_c} \quad (2.5)$$

n_c mertebesi, denklem 2.5'te gösterildiği gibidir.

$$n_d: D(q) = 1 + d_1 q^{-1} + \dots + d_{n_d} q^{-n_d} \quad (2.6)$$

n_d mertebesi, denklem 2.6'da gösterildiği gibidir.

$$n_f: F(q) = 1 + f_1 q^{-1} + \dots + f_{n_f} q^{-n_f} \quad (2.7)$$

n_f mertebesi, denklem 2.7'de gösterildiği gibidir.

2.1.2 Doğrusal olmayan modeller

2.1.2.1 Dışsal kaynaklı otoregresif modeli

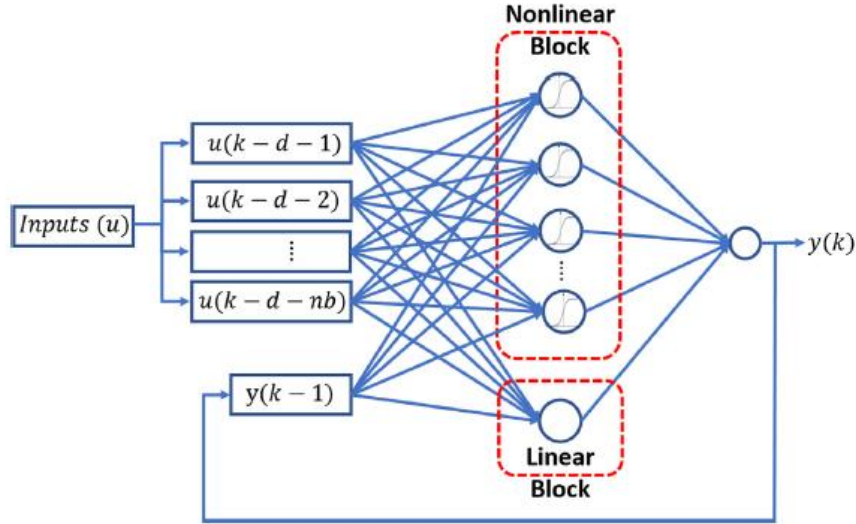
Bu doğrusal olmayan dışsal kaynaklı otoregresif model yaklaşımı, zaman serisi modellemede yaygın olarak kullanılan doğrusal dışsal kaynaklı otoregresif model yaklaşımına dayanmaktadır. Doğrusal olmayan dışsal kaynaklı otoregresif modelleri, doğrusal olmayan yapıları tanımlarken bir doğrusal ve bir doğrusal olmayan bloğun paralel kombinasyonu kullanır. Doğrusal ve doğrusal olmayan fonksiyonlar, ölçülen giriş ve çıkış sinyallerinin belirtilen modelin çıktılarına göre dönüşümleri gerçekleştirilerek hesaplanan ve regresör adı verilen değişkenler cinsinden ifade edilir. Dizinin son değeri; önceki değerlerinin bir sabit terimin ve bir hata teriminin doğrusal bir fonksiyonudur, seriyi etkileyen harici veriler ve geçmiş değerlerle beraber regresyon uygulanarak tahmin yapılmaktadır.

Matematik model olarak tek katmanlı dışsal kaynaklı doğrusal olmayan otoregresif model denklem 2.8'deki gibi tanımlanmaktadır.

$$y(k) = w_0x(k) + b_0 + \sum_{i=1}^m \phi_i(w_ix(k) + b_i) + e(k) \quad (2.8)$$

Denklemden $y(k)$ çıkışı, $x(k)$ regresör setini, doğrusal olmayan modelde w_i ağırlıkları ve b_i sapmaları, doğrusal modelde w_0 ağırlıkları ve b_0 sapmaları, e model hatalarını, m doğrusal olmayan modelde kullanılan birim sayısını ve ϕ_i aktivasyon fonksiyonunu göstermektedir. Tek çıkışlı dışsal kaynaklı doğrusal olmayan otoregresif model gösterimi Şekil 2.1'deki gösterilmektedir.

Regresör seti, girdilerin ve çıktıların geçmiş değerlerinden oluşur. Regresörler, parametre tahmin işlemi sırasında optimizasyon algoritmasının çok sık sapmasına neden olur. Bu nedenle, çıktının tek bir geçmiş değeri, regresör kümesine dahil edilir.



Şekil 2.1: Tek çıkışlı DKDOO model gösterimi (Alcan ve diğ, 2019).

Doğrusal olmayan dışsal kaynaklı otoregresif modelindeki parametreler, aşağıdaki satır arama algoritmaları olarak, altuzay Gauss-Newton en küçük kareler araması, Levenberg-Marquardt en küçük kareler araması, uyarlanabilir altuzay Gauss-Newton araması ve dik inişli en küçük kareler araması kullanılarak tahmin edilirler.

Model tahminlerinde farklı aktivasyon fonksiyonları kullanılabilir. Bunlar sigmoid ağı, dalga ağı ve ikili arama ağacı ağı olabileceği gibi kendi tanımlayabileceğiniz yapay ağlar da kullanılabilir.

Dalga ağı aktivasyon fonksiyonu

Dalga ağı, doğrusal olmayan dalgacık ağını depolayan bir nesnedir. Dalga ağı fonksiyonu doğrusal olmayan bağlantıyı denklem 2.9'daki gibi tanımlamaktadır.

$$\begin{aligned}
 F(x) = & (x - r)PL + a_{s1}f \left((b_{s1}(x - r))Q - c_{s1} \right) \\
 & + \dots a_{sn}f \left((b_{sn}(x - r))Q - c_{sn} \right) \\
 & + a_{w1}g \left((b_{w1}(x - r))Q - c_{wn} \right) + d
 \end{aligned} \tag{2.9}$$

Burada f ölçeklendirme fonksiyonu ve g ise dalga ağı fonksiyonudur ve denklem 2.10 ve 2.11'deki gibi tanımlanır.

$$f(x) = e^{-0.5x'x} \quad (2.10)$$

$$g(x) = (\dim(x) - x'x)e^{-0.5x'x} \quad (2.11)$$

Tahmin verilerinin temel bileşen analizi, P ve Q projeksiyon matrislerini belirlemek için gerçekleştirilir. Genellikle $p = p.m$ değeri m'den küçük ise tahmin verilerindeki x'in bileşenleri doğrusal bağımlıdır. Ölçekleme ve dalgacık fonksiyonlarında kullanılan x bileşenlerinin sayısı, Q matrisinin sütunlarının sayısını gösterir. r , tahmin verilerinden hesaplanan regresör vektörünün ortalama değerini temsil eder. d, a_s, b_s, a_w, b_w değerleri skaler değerlerdir. Ölçeklendirme parametreleri alt indis olarak s'yi ve dalgacık parametreleri de alt indis olarak w ile tanımlanmaktadır.

Sigmoid ağı aktivasyon fonksiyonu

Sigmoid ağı, doğrusal olmayan sigmoid ağını depolayan bir nesnedir. Sigmoid ağı fonksiyonu doğrusal olmayan bağlantıyı denklem 2.12'deki gibi tanımlamaktadır.

$$F(x) = (x - r)PL + a_1f((x - r)Qb_1 - c_1) + \dots \quad (2.12)$$

$$+ a_nf((x - r)Qb_n - c_n) + d$$

Burada sigmoid fonksiyonu f denklem 2.13'teki gibi tanımlanır.

$$f(z) = \frac{1}{e^{-z} + 1} \quad (2.13)$$

P ve Q, tahmin verilerinin temel bileşen analizi kullanılarak belirlenen projeksiyon matrisleridir. Genellikle $p = p.m$ değeri m'den küçük ise tahmin verilerindeki x'in bileşenleri doğrusal bağımlıdır. Sigmoid işlevinde kullanılan x bileşenlerinin sayısı, Sigmoid ağı nesnesinin doğrusal olmayan regresyonlar özelliğinin boyutuna eşit olan Q matrisi ve q sütunlarının sayısına karşılık gelir.

İkili arama ağacı aktivasyon fonksiyonu

İkili arama ağacı ağı, doğrusal olmayan ikili ağacı depolayan bir nesnedir. Doğrusal olmayan bir $y = F(x)$ işlevi tanımlamak için ağaç bölümü kullanıldığında, F ikili ağaç arama ile x -uzay bölümlemede yerel bir doğrusal eşlemedir. Burada ikili ağaç arama fonksiyonu f denklem 2.14'teki gibidir.

$$F(x) = xL + [1, x]C_a + d \quad (2.14)$$

Burada x , D_a 'nın aktif bölümlenmiş kısmına aittir ve D_k , x uzayının bölümleridir. D_a 'nın aktif bölümlerini hesaplamak için N düğümlü ağaç ve J seviyesi başlatılır. J seviyesindeki düğüm bir sonlandırma yaprağıdır ve J seviyesinde iki alt ögeye sahiptir. N , ağaçtaki yaprak sayısıdır ve $2^{(J+1)} - 1$ olarak tanımlanır. r düğümündeki bölümleme $[1, x]^* B_r > 0$ veya ≤ 0 temelindedir. B_r , ağaçların geçmişinden gelen düğümlerdeki bölümlerin en küçük kareler yöntemi ile hesaplanarak kararlılığı arttırmak için kullanılır. Her düğümde r , D_r üzerindeki bilinmeyen regresyon fonksiyonunu C_r katsayıları ile en iyi doğrusal yaklaşımı hesaplarken cezalandırılmış en küçük kareler yöntemi kullanılır.

2.1.2.2 Hammerstein – Wiener modeli

Hammerstein-Wiener modelleri, bir doğrusal blok ile seri halinde bir veya iki statik doğrusal olmayan blok kullanan dinamik sistemleri tanımlar. Doğrusal blok, modelin dinamik bileşenini temsil eden ayrı bir transfer fonksiyonudur. Doğrusal blok, ayrık zamanlı bir transfer fonksiyonudur ve doğrusal olmayan bloklar, doygunluk, dalgacık, sigmoid, parçacıklı doğrusal ve ölü bölge gibi doğrusal olmayan tahmin ediciler kullanılarak uygulanır. Hammerstein-Wiener modeli Şekil 2.2'de gösterildiği gibidir.



Şekil 2.2: Hammerstein-Wiener modeli.

Giriş sinyali, çıkış sinyalini üretmek için birinci doğrusal olmayan blok, bir doğrusal blok ve bir ikinci doğrusal olmayan bloktan geçer. Hammerstein-Wiener modeli denklemler 2.15, 2.16 ve 2.17'deki gibi tanımlanır.

$$u(t) = f(u(t)) \quad (2.15)$$

$$x(t) = \frac{B_{j,i}(q)}{F_{j,i}(q)} w(t) \quad (2.16)$$

$$y(t) = h(x(t)) \quad (2.17)$$

$u(t)$ sistemin girişi ve $y(t)$ sistemin çıkışıdır. f giriş için doğrusal olmayan bir fonksiyonu tanımlar ve h çıkış için doğrusal olmayan fonksiyonları tanımlamaktadır ve çoklu giriş ve çıkış için her kanal için bağımsız olarak tanımlanır. $w(t)$ ve $x(t)$ doğrusal bloğun giriş ve çıkışlarını tanımlayan iç değişkenlerdir. $B(q)$ ve $F(q)$ doğrusal dinamik bloktaki doğrusal fonksiyonlardır. Eğer doğrusalsızlık sadece giriş fonksiyonunda kullanılırsa bu modele Hammerstein modeli denir ve eğer doğrusalsızlık sadece çıkış fonksiyonunda kullanılırsa bu modele Wiener modeli denir.

Hammersteinn – Wiener modeli için aktivasyon fonksiyonları ölü bölge fonksiyonu, parçalı doğrusal fonksiyonu, doygunluk fonksiyonu, sigmoid ağ fonksiyonu ve dalgacık ağ fonksiyonu olarak sıralanabilir.

Ölü bölge fonksiyonu

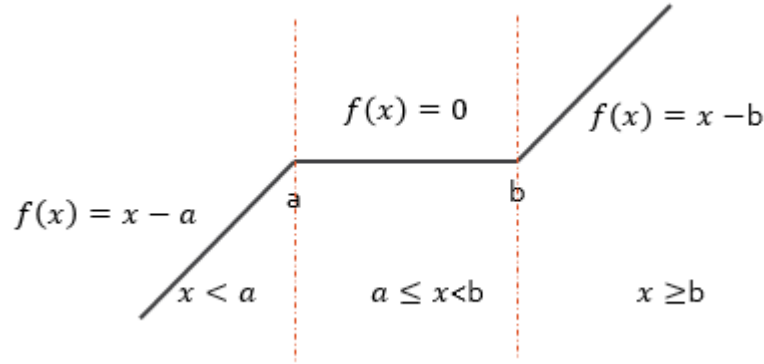
Ölü bölge fonksiyonu, ölü bölge olarak adlandırılan belirli bir bölgede sıfır çıktı üretir. Ölü bölgenin alt ve üst limitleri sırasıyla parametrelerin başlangıç ve bitişi olarak belirtilir. Ölü bölge, doğrusal olmayan bir $y = f(x)$ fonksiyonunu tanımlar ve tanımlanan üç aralık denklemler 2.18, 2.19 ve 2.20' de gösterilmiştir.

$$f(x) = x - a \quad x < a \quad (2.18)$$

$$f(x) = 0 \quad a \leq x < b \quad (2.19)$$

$$f(x) = x - b \quad x \geq b \quad (2.20)$$

$f(x) = 0$ olduğunda, x 'in değeri a ile b arasında olduğunda, bu bölgeyi, sıfır aralık bölgesi olarak tanımlanır. Şekil 2.3'te gösterildiği gibi aralıklar vardır.



Şekil 2.3: Ölü bölge aralıkları.

Parçalı doğrusal fonksiyon

Parçalı doğrusal fonksiyonu, doğrusal olmayan bir $y = f(x)$ fonksiyonu olarak tanımlanır; burada f x 'in parçalı doğrusal bir fonksiyonudur.

$k = 1, \dots, n$ ve $y_k = f(x_k)$ olmak üzere n adet kesme noktası (x_k, y_k) vardır. f , kesme noktaları arasında doğrusal olarak enterpole olmaktadır. y ve x scaler değerlerdir.

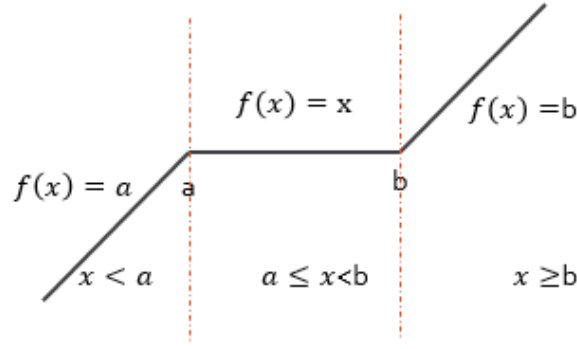
Doygunluk fonksiyonu

Doygunluk fonksiyonu, doğrusal olmayan bir $y = f(x)$ fonksiyonunu tanımlar ve tanımlanan üç aralık denklem 2.21, 2.22 ve 2.23'teki gibidir ve Şekil 2.4'teki gibi gösterilir.

$$f(x) = a \quad x < a \quad (2.21)$$

$$f(x) = x \quad a \leq x < b \quad (2.22)$$

$$f(x) = b \quad x \geq b \quad (2.23)$$



Şekil 2.4: Doygunluk aralıkları.

Sigmoid ağ fonksiyonu

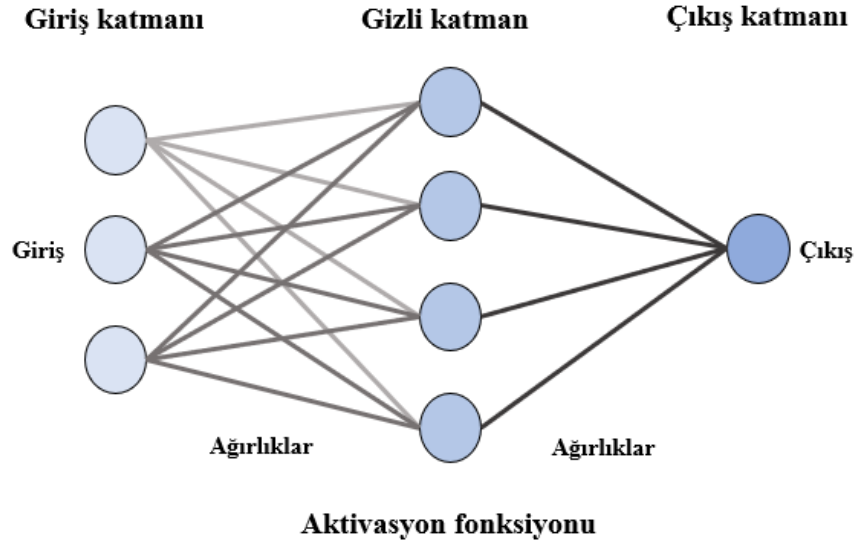
Sigmoid ağ fonksiyonu, bir giriş katmanı, bir çıkış katmanı ve sigmoid aktivasyon fonksiyonlarını kullanan bir gizli katmandan oluşan sinir ağıdır.

Aktivasyon fonksiyonu olarak bir sigmoid kullanan radyal tabanlı sinir ağı fonksiyonunu birleştirir. Tahmin fonksiyonu denklem 2.24'e dayanmaktadır.

$$y(t) = (u - r)PL + \sum_i^n a_i f((u - r)Qb_i - c_i) + d \quad (2.24)$$

Denklemden u girişi, y çıkışı, r regresörü, Q doğrusal olmayan alt uzayı, P doğrusal alt uzayı, L doğrusal katsayısı, d çıkışın offsetini, b genişleme katsayısını, c öteleme katsayısını, a çıkış katsayısını göstermektedir. f sigmoid fonksiyonu denklem 2.25 gibi tanımlanmaktadır. Sigmoid fonksiyonu katman yapısı Şekil 2.5'te gösterilmiştir.

$$f(z) = \frac{1}{e^{-z} + 1} \quad (2.25)$$



Şekil 2.5: Sigmoid fonksiyonu katmanları.

Dalgacık ağ fonksiyonu

Dalgacık ağ fonksiyonu, dalgacık teorisini ve sinir ağlarını birleştiren doğrusal olmayan bir fonksiyondur. Dalgacık ağları, aşağıdaki ifadeye dayalı olarak bir aktivasyon fonksiyonu olarak bir dalgacık kullanan ileri beslemeli sinir ağlarıdır. Dalgacık ağ yapısı denklem 2.26’da tanımlanan gibidir.

$$\begin{aligned}
 F(x) = & (x - r)PL + a_{s1}f\left((b_{s1}(x - r))Q - c_{s1}\right) \\
 & + \dots a_{sn}f\left((b_{sn}(x - r))Q - c_{sn}\right) \\
 & + a_{w1}g\left((b_{w1}(x - r))Q - c_{wn}\right) + d
 \end{aligned} \tag{2.26}$$

Denklemde u girişi, y çıkışı, r regresörü, Q doğrusal olmayan alt uzayı, P doğrusal alt uzayı, L doğrusal katsayıyı, d çıkışın offsetini, b_s genişleme katsayısını, b_w dalgacık genişleme katsayısını, c_s öteleme katsayısını, c_w dalgacık öteleme katsayısını, a_s ölçeklendirme katsayısı ve a_w dalgacık ölçeklendirme katsayısı göstermektedir. Burada f ölçeklendirme fonksiyonu ve g ise dalga ağ fonksiyonudur ve denklem 2.27 ve 2.28’deki gibi tanımlanır.

$$f(x) = e^{-0.5x'^2} \tag{2.27}$$

$$g(x) = (\dim(x) - x'x)e^{-0.5x'x} \quad (2.28)$$

Tahmin süreçlerinde en iyi performans modelini elde etmek için model öğrenme adımları sırasında dalgacık katsayısı a , genişleme katsayısı b ve öteleme katsayısı c optimize edilir.

2.2 Makine Öğrenmesi Modelleri

Makine öğrenmesi, insanların öğrenme şekillerini taklit etmek için veri ve algoritmaların kullanımına odaklanıp doğruluğunu kademeli olarak artıran bir bilgisayar bilimi dalıdır. Makine öğrenmesi , bir bilgisayarın doğrudan yönergeler olmadan öğrenmesine yardımcı olmak için matematiksel modelleri kullanma işlemidir. Makine öğrenmesi, verilerdeki kalıpları belirlemek için algoritmaları kullanır. Tahmin yapabilen bir veri modeli oluşturmak için de bu kalıplar kullanılır. Tıpkı insanların daha fazla alıştırma yaptıkça gelişmesi gibi, veri ve deneyim miktarı arttıkça makine öğrenmesinin sonuçları da daha doğru hale gelir. Uyarlanabilirliği sayesinde makine öğrenmesi verilerin, isteklerin veya görevlerin sürekli değiştiği senaryolarda veya bir çözümün etkili bir şekilde kodlanmasının mümkün olmadığı durumlarda kullanılabilen bir seçenektir.

Destek vektör regresyonu

Destek vektör makinelerin temelleri Vladimir Vapnik ve arkadaşları tarafından atılmıştır. Vapnik-Chervonekis istatistiksel öğrenme teorisine dayanan bir öğrenme yöntemidir. Destek vektörü regresyonu, destek vektör makinelerinin başka bir şeklidir. Destek vektör makineleri sınıflandırma amacıyla kullanılmaya ve iki doğrusal veriyi sınıflandırmak amacıyla oluşturulmuştur. Destek vektör makineleri, regresyon ile veriyi sadece sınıflandırmak için değil aynı zamanda tahmin yapmak için de kullanılabileceğini gösteren çalışmalar yapılmıştır (Vapnik, 2000). Destek vektör makineleri ampirik riski azaltmak yerine yapısal riski azaltır. Ampirik risk ile gürültüyü ve diğer bozucuların etkisini tanımlayan modeller sayesinde tahmin hatasını en aza düşürmeye odaklanırken yapısal riski azaltarak mümkün olduğunca verileri takip edecek bir fonksiyon eğitilir.

Destek vektör regresyonu ise doğrusal olmayan verilerin haritalanması ile aralarındaki korelasyonu bulmak için kullanılır. Destek vektör regresyonu, değişkenler arasında hiperdüzlem oluşturur ve veri noktaları ile hiperdüzlem arasındaki mesafeyi ortaklayarak veriler arasındaki korelasyonu hesaplar. Değişkenler arasındaki hiperdüzlemi her iki değişkenin en yakın veri noktalarından eşit uzaklıkta kalacak şekilde oluşturur. Hiperdüzlem sayesinde, destek vektör regresyonunun marjinal limitleri ihlal etmeden mümkün olduğu kadar çok veri noktası sığdırmaya çalıştığı bir tüp oluşturur. Bu tüpün genişliği ε parametresi yardımıyla tanımlanır. Regresyon için hiperdüzlemdeki tahmin hatası ve oluşturulan doğru arasındaki bağlamı etkileyen iki ana parametre ε ve C kalibre edilerek kullanılır. ε parametresi, tahmin hatasının olmadığı duyarsız bölgeyi tanımlar ve C parametresi, duyarsız bölge dışındaki hatanın ağırlığını tanımlar.

Bu tüpün altında kalan veri noktaları iyi bir tahmin doğruluğuna sahipken, tüpten uzak olan veri noktaları kötü bir doğruluğa sahiptir. Yolun sınır değerleri üzerinde veya ona yakın olan veri noktaları, destek vektörleri fonksiyonu olarak görülür. Bu destek vektörleri temelinde sadece veri korelasyonu matrisini oluştur. Destek vektör makinelerinde, veri kümesi doğada doğrusal kaldığından hiperdüzlem oluşumu kolaylaşır ancak destek vektör regresyonunda esas olarak hiperdüzlemin oluşturulmasının zorlaştığı doğrusal olmayan veri kümesi kullanılır.

Destek vektör regresyonunun destek vektör makinelerinden ayıran ana özellik kernel fonksiyonlarıdır. Vapnik ve diğ. (1996) tarafından önerilen çekirdek fonksiyonlarının tanıtılmasıyla doğrusal olmayan sınıflandırma mümkün olmuştur. Kernel fonksiyonları, veri kümesinin boyutunu daha yüksek boyuta dönüştürür ve hiperdüzlemi oluşturur. Kernel fonksiyonları kullanıldığında, C parametresi için büyük değerler kullanılması algoritmanın tüm eğitim hatalarını duyarsız bölge içine koymasını götürmesine sebep olur ve destek vektör regresyonun eğitim verilerini ezberlemesine neden olur. ε parametresi için küçük değerler kullanılması verilerde gürültü olmadığını varsayar ve destek vektör regresyonun aşırı uyum sağlamasına neden olur. Buna karşılık, C parametresi için küçük değerler veya ε parametresi için büyük değerler tanımlanması destek vektör regresyonun yetersiz uyum sağlamasına neden olur.

Destek vektör makineleri Smola ve diğ. (2003) tarafından önerilen bir kayıp fonksiyonunun tanıtılmasıyla regresyon problemlerini çözmek için güncellendi. Destek Vektör Regresyonu ile ilgili ayrıntılı yayınlar, Smola ve Schölkopf (2003) tarafından yayınlanmıştır. Chang ve Lin (2013) tarafından birçok programlama diline ve veri madenciliği yazılımına entegre edilebilen LIBSVM kütüphanesi geliştirilmiş ve yayınlanmıştır.

Destek vektör regresyonu için temel algoritma veri seti için giriş olarak x_n ile çıkış olarak y_n arasında $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ doğrusal bir yaklaşım elde etmektir. Çıkış değerini, verilerden en az sapmaya sahip, mümkün olduğu kadar düz ve en az model karmaşıklığına sahip bir $f(x)$ fonksiyonu ile bir ağırlıklandırma vektörü ve bayas değeri yardımıyla denklem 2.29'da gösterildiği gibi tahmin edilmektedir.

$$f(x) = \langle w, x \rangle + b \quad (2.29)$$

w ve x arasındaki noktasal çarpım özel bir uzay oluşturmaktadır ve bu uzay sayesinde doğrusal olmayan tahmin fonksiyonu için doğrusal bir regresyon sağlar ve tabiki kernel fonksiyonlarını destekler.

Destek vektör regresyonu tahmin hatasını, yani tahmin edilen ve istenen çıktılar arasındaki mesafeyi en aza indirirken, yüzey etrafında merkezlenmiş en dar tüpü bulmaya çalışan bir optimizasyon problemi olarak formüle eder. Bunun için tahmin hatasının katı bir kısıtlaması olacaktır ve regresyon için optimizasyon problemi denklem 2.30'daki gibi yazılabilir.

$$\min \frac{1}{2} \|w\|^2 \quad (2.30)$$

burada $\|w\|$, yaklaşılan yüzeye normal vektörün büyüklüğüdür. İyi bir doğrusal regresyon için bir dışbükey fonksiyon kullanılır.

$$\frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (s_i + s_i^*) \quad (2.31)$$

Denklem 2.31’de tanımlanan ikinci terim genelleme terimidir. s_i ve s_i^* değerleri marjinin dışında kalan en yakın veriler için gevşek değişkenler olarak tanımlanır. Doğrusal fonksiyonun etrafında marjinler olduğu sürece yukarıdaki denklemi en aza indirirken şartlar vardır ve bunlar denklem 2.32, 2.33 ve 2.34’deki gibi tanımlanabilir.

$$y_i - \langle w, x_i \rangle - b \leq \varepsilon + s_i \quad (2.32)$$

$$\langle w, x_i \rangle + b - y_i \leq \varepsilon + s_i^* \quad (2.33)$$

$$s_i, s_i^* \geq 0 \quad (2.34)$$

Bu eşitsizlikler, ana çizginin hem üstündeki hem de altındaki çizgileri temsil eder. Denklem 2.35’te belirtilen ε -duyarsız kayıp fonksiyonunda gösterildiği gibi, bu marjlar içindeki verilerle ilgili tüm hatalar göz ardı edilir.

$$|s|_\varepsilon = \begin{cases} 0 & , \text{eğer } |s| \leq \varepsilon \\ |s| - \varepsilon & , \text{aksi halde} \end{cases} \quad (2.35)$$

Minimizasyon yada maksimizasyon optimizasyon problemlerinde, kısıtlamalara sahip fonksiyonlar mevcutsa Lagrange fonksiyonu bu tür problemlerin çözümünde etkili bir yöntemdir. Yukarıda bahsedilen kısıtlama ve minimizasyon fonksiyonları ile Langrange fonksiyonu tanımlanmak istenirse denklem 2.36’daki gibi oluşturabilir.

$$\begin{aligned} L := & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (s_i + s_i^*) - \sum_{i=1}^n (\eta_i s_i + \eta_i^* s_i^*) \\ & - \sum_{i=1}^n \alpha_i (\varepsilon + s_i - y_i + \langle w, x_i \rangle + b) \\ & - \sum_{i=1}^n \alpha_i^* (\varepsilon + s_i^* - y_i + \langle w, x_i \rangle + b) \end{aligned} \quad (2.36)$$

$\eta_i, \eta_i^*, \alpha_i, \alpha_i^*$ Langrange çarpanlarıdır. L ’nin türevlerini minimize edip bulduktan sonra denklem 2.37’nin maksimize edilmesi gerekir.

$$-\frac{1}{2} \sum_{i,j=1}^n (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) \langle x_i, x_j \rangle - \varepsilon \sum_{i=1}^n (\alpha_i + \alpha_i^*) + \sum_{i=1}^n y_i (\alpha_i - \alpha_i^*) \quad (2.37)$$

Denklem 2.38 ve 2.39 ile tanımlanan aralıklarda gevşek değerler kullanılır.

$$\sum_{i=1}^n (\alpha_i - \alpha_i^*) = 0 \quad (2.38)$$

$$\alpha_i, \alpha_i^* \in [0, C] \quad (2.39)$$

w 'nin türevi denklem 2.40'daki gibidir.

$$w = \sum_{i=1}^n (\alpha_i - \alpha_i^*) x_i \quad (2.40)$$

Yukarıdaki denklemler ile ana nihai model bir ikili doğrusal problem olacaktır. Bu model, denklem 2.41'de gösterildiği gibi öğrenme tamamlandıktan sonra yeni değerleri tahmin etmek için kullanılacaktır.

$$f(x) = \sum_i^n (\alpha_i - \alpha_i^*) \langle x_i, x \rangle + b \quad (2.41)$$

Kernel fonksiyonları

Doğrusal olarak çözelemeyecek problemler için verileri doğrusal olmayan fonksiyonlarla eşleyerek, verilerin küçük boyutlu özelliklerini daha yüksek boyutlara taşıyan kernel fonksiyonları denklem 2.42 yardımıyla doğrusal olarak tahmin edilebilir.

$$k(x, y) = (\varphi(x) \cdot \varphi(y)) \quad (2.42)$$

$\varphi(x)$ ve $\varphi(y)$, x ve y için aralarında noktasal çarpım içeren doğrusal olmayan haritalamalardır. Buradan hareketle, doğrusal olmayan problemler için $\langle x_i, x \rangle$ yerine denklem 2.43'teki gibi kernel fonksiyonları kullanılabilir (Smola ve Schölkopf, 2003).

$$f(x) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) k(x_i, x) + b \quad (2.43)$$

Destek vektör regresyona uygulanabilecek çeşitli kernel türleri vardır ve bu kernellerin kombinasyonları mümkündür. Doğrusal olmayan çekirdekler, ε ve C 'nin destek vektör regresyonun ana parametreleri olarak kullanıcı tarafından önceden tanımlanmış parametrelerdir.

Doğrusal kerneli

Doğrusal kernel de vektörlerin nokta çarpımı alınır ve hiçbir eşleme uygulanmaz. Destek vektör regresyonu bir kernel fonksiyonu ile çalıştırılacağı için doğrusal kernel seçenekler arasında olacaktır. Denklem 2.44'deki gibi tanımlanabilir.

$$k(x_i, x_j) = \langle x_i, x_j \rangle \quad (2.44)$$

Polinom kerneli

Polinom kerneli doğrusal olmayan veri tahmin yöntemlerinde kullanılabilir. Denklem 2.45'deki gibi tanımlanır.

$$k(x_i, x_j) = (\langle x_i, x_j \rangle + 1)^d \quad (2.45)$$

d , kernel parametresi için polinomun derecesini tanımlar.

Radyal tabanlı fonksiyon kerneli

Radyal tabanlı fonksiyon kerneli, iki vektör arasındaki öklid mesafesinin karesini kullanan bir fonksiyondur ve denklem 2.46'daki gibi tanımlanır.

$$k(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2) \quad (2.46)$$

Burada $\gamma = 1/2\sigma^2$ ve $\gamma > 0$ olarak kernel parametrelerini gösterir.

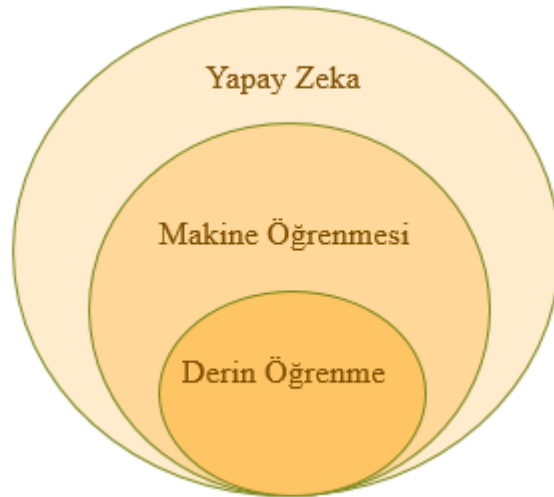
Sigmoidal kerneli

Sigmoidal kernel, iki vektör arasındaki hiperbolik tanjant ilişkisini kullanır. Denklem 2.47'deki gibi gösterilir.

$$k(x_i, x_j) = \tanh(\gamma \cdot x_i^T x_j) + r \quad (2.47)$$

2.3 Derin Öğrenme Modelleri

Derin öğrenme, dijital sistemlerin yapılandırılmamış ve etiketlenmemiş veri seti ile sonuçları tahmin eden birden fazla katmanı kullanarak kararlar almasını sağlayan bir makine öğrenme yöntemidir. Derin öğrenme, makine öğrenmesi ve yapay zekâ birbiri ile bağlantılı farklı terimlerdir. Derin öğrenme, makine öğrenmesinin; makine öğrenmesi ise yapay zekânın alt dalı olarak özetlenebilir ve Şekil 2.6'daki gösterilebilir.



Şekil 2.6: Yapay zeka alt uzayları.

Yapay zeka, sürekli kendini geliştiren ve insan zekasına benzer şekilde görevler getiren bilgisayar modelleridir. Makine öğrenmesi, bir veri kümesini işlemekten geçirip tahminlerde bulunmak ya da sınıflandırmak için kullanılırlar. Makine öğrenme

algoritmalarında üç tür öğrenme biçimi vardır: gözetimli öğrenme, gözetimsiz öğrenme ve pekiştirmeye dayalı öğrenmedir.

Gözetimli öğrenmede algoritma, etiketlenmiş verilere göre eğitilir. Algoritma karar verdiğinde, verilere dahil edilen etiketler kullanılarak bu kararın doğruluğu kontrol edilir. Tahmin hakkındaki doğruluk geri bildirimi ve verilerin etiketlenmesi ise kullanıcı tarafından yapılır.

Gözetimsiz öğrenmede algoritma, etiketlenmemiş verilere göre eğitilir. Algoritmanın kararlarını kontrol edebilecek etiketler veya bilgiler olmadan eğitilir. Bunun yerine algoritmanın kendisi veriden keşfederek sonuca varır.

Pekiştirmeye dayalı öğrenmede algoritma deneme yanılma yöntemi ile hedeflenen amaca göre kararlar alır. Veri kümeleri yerine kendi oluşturduğu bilgiler ile karara ulaşır.

Derin öğrenme birçok alanda aktif olarak kullanılmaktadır. Yüz tanıma sistemleri, ses tanıma sistemleri, görüntü iyileştirilmesinde ve doğal dil işleme gibi alanlarda yaygın kullanıma sahiptir.

Üç ana derin öğrenme modeli vardır; bunlar evrişimli sinir ağları, üretken çekişmeli ağlar ve tekrarlayan sinir ağlarıdır. Bu tezin kapsamında tekrarlayan sinir ağlarından bahsedilecektir.

2.3.1 Tekrarlayan sinir ağları

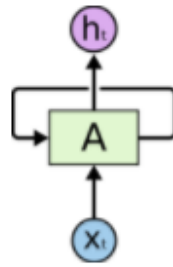
Tekrarlayan sinir ağları, mevcut zaman adımındaki tahminleri etkilemek için ağa girdi olarak önceki bir zaman adımından ağ aktivasyonlarını besleyen döngüler içerir. Bu aktivasyonlar, prensipte uzun vadeli zamansal bağlamsal bilgiyi tutabilen ağın dahili durumlarında depolanır. Bu mekanizma, TSA'ların giriş dizisi geçmişi üzerinde dinamik olarak değişen bir bağlamsal pencereden yararlanmasına izin verir (Sak ve diğ, 2014).

Bu durum TSA yapısının temelini oluşturur. Klasik yapay ileri beslemeli sinir ağları eski bilgilerden bağlamsal pencereden yararlanmazlar. Klasik nöron yapısına geri besleme ekleyerek geçmiş bilgilerden faydalanılması sağlanır. Nöron yapısı Şekil 2.7'deki gibi tanımlanır.

Burada A sinir ağı yığını, x_t giriş değeri ve h_t çıkış değeridir. Bir döngü halinde geribeslemeli olarak bilgiyi tekrarlayarak kullanır. Bu nöronların bir arada kullanılması ile bilgiler nöronlara aktırılarak tekrarlayan sinir ağını oluştururlar. Şekil 2.8'deki gibi kullanılırlar.

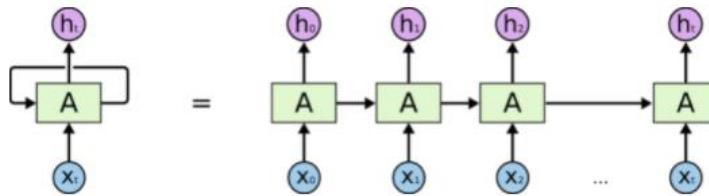
Standart TSA'larda, yinelenen modül tek bir tanh katmanı gibi çok basit bir yapıya sahip olacaktır. Şekil 2.9'te gösterildiği gibidir.

TSA güçlü bir tahmin modelidir. TSA'larından doğruluğu yüksek ve iyi tahmin değerleri alabilmek için mimari de dikkat edilmesi gereken iki önemli husus vardır. Bu hususlar kaybolan veya patlayan gradyan problemi ve uzun süreli bağımlılık problemleridir.



Şekil 2.7: Basit nöron yapısı (Alam, 2018).

Geri yayılım ile amacımız ağıdaki ağırlıkların her birini güncellemek ve bu sayede gerçek çıktının hedef çıktıya daha yakın olmasına neden olmasını sağlamaktır

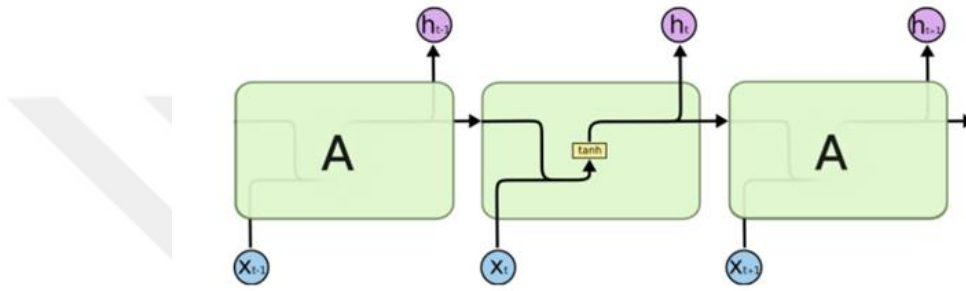


Şekil 2.8: Tekrarlayan sinir ağı (Alam, 2018).

Ancak geriye dönük olarak sürekli yenilenen gradyanların azalması sebebiyle tahmin sonuç giderek küçülmektedir ve bu durum kaybolan gradyan problemi olarak adlandırılmaktadır. Patlayan gradyan problemi ise gradyanların büyümesi sonucunda tahmin sonuçlarının giderek büyümesi olarak adlandırılır. Bu sorunu çözmek için

gradyan azalması optimizasyon algoritması kullanılır. Bu algoritma denklemin local minimuma yakınsamak için birinci türev kullanır.

Uzun süreli bağımlılık probleminde ise tüm girdiler nöronların dahili hafızasında tutulmaktadır ve bu hafıza sayesinde geçmiş değerler ile bağlantılar kurarak zaman serisi problemlerini anlamlandırmaktadır. Ancak dahili hafızada bu bilgilerin saklanma süreleri bilinmemektedir. Dahili hafızadaki bilgilerin değerleri önemli değildir ve bu nedenle önemsiz verilerinde olduğu tüm geçmişin tutulmasına ihtiyaç yoktur.



Şekil 2.9: Yinelenen TSA modülü (Alam, 2018).

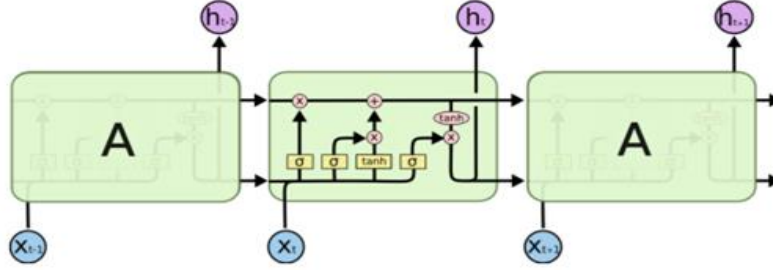
Yakın geçmişte olan bilgiler bağlamsal anlamları kolayca oluşturabiliyorken araya giren farklı bilgiler ile daha geçmiş zaman bilgilerine ihtiyaç duyulduğunda TSA'ların tahminlerinin kötüleştiği görülmektedir. Bu sorunun çözümü için uzun kısa süreli hafıza modeli geliştirilmiştir.

2.3.2 Uzun kısa süreli hafıza

Uzun süreli bağımlılık problemin önüne geçmek için TSA mimarisi üzerinden tasarlanan uzun kısa süreli hafıza modeli 1997 yılında Hochreiter ve Schmidhuber tarafından yayınlanmıştır. UKSH, TSA'lar gibi birbiri ardına kullanılan nöronlardan oluşan bir zincir ile temsil edilmektedir ancak bu zincir daha fazla katman ve özel bağlamlar ile kullanılmaktadır. Standart TSA'lar 5-10 ayrı zaman adımından daha büyük geçmişe bakmalardan fazlasında öğrenmesi güçleşir ve bu uzun süreli bağımlılık problemi standart TSA'ların zaman serisi modellerinde avantaj sağlayıp sağlayamayacağı konusunda şüpheler oluşturuyor. Ancak UKSH bu sorundan etkilenmez, sabit hata döngüleri aracılığıyla sabit hata akışını zorlayarak 1000 ayrı zaman adımını aşan geçmiş değerleri köprülemeyi öğrenebilir (Gers ve diğ, 2000). Tek

bir sinir ağı katmanını yerine, tekrar eden ve içeriğinde özel bir şekilde etkileşimde olan unutmama kapısı, giriş kapısı ve çıkış kapısı vardır (Lysfjord, 2017). Basit UKSH yapısı Şekil 2.10'daki gibidir.

UKSH, kullandığı hücrelerdeki kapılar ile bilgiyi tanımlar. Kapılar isteğe bağlı olarak bilgi vermenin bir yoludur. Kullanılan sigmoid katmanında bilgi ile sigmoidin seçimi noktasal olarak çarpılarak bilginin geçişine karar verilir.



Şekil 2.10: Basit uzun kısa süreli hafıza yapısı (Lysfjord, 2017).

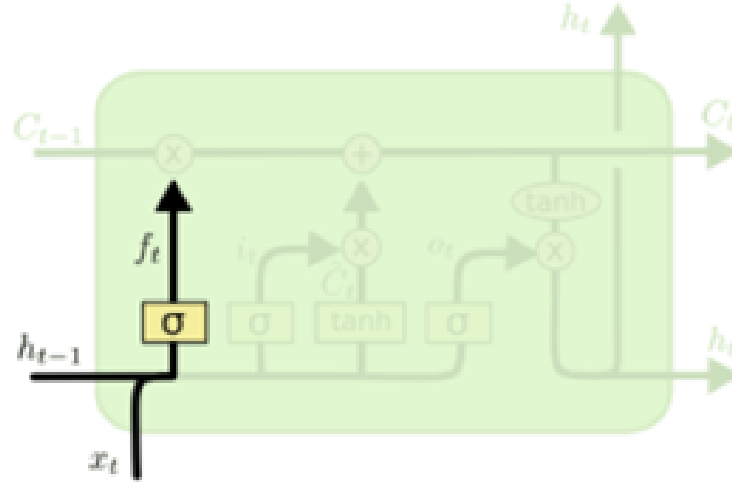
Sigmoid katmanında, 0 değerinde bilginin geçmesine izin verilmezken 1 değeri bilginin geçmesine izin verilir. Böylelikle bilgi sigmoid katmanı sayesinde 0 ile 1 arasına sıkıştırılmış olur.

UKSH hücre yapısını incelemeye başladığımızda ilk olarak hücre içine alınacak bilgilere karar veren unutmama kapısını ele alalım. f_t unutmama kapısı denklem 2.48'deki gibi tanımlanır.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (2.48)$$

Unutmama kapısı, h_{t-1} ve x_t 'ye bakarak C_{t-1} hücre durumu için bilginin alınıp alınmayacağına sigmoid katmanı olarak karar verir. Denklemden W_f kapının ağırlıklandırma değeri ve b_f ön bilgiyi temsil eder. Unutmama kapısı Şekil 2.11'deki gibi temsil edilir.

Unutmama kapısı sonrası giriş kapısı olarak kullanılan sigmoid katmanı ile hangi bilgilerin güncelleneceğine denklem 4.1 de temsil edilen şekilde karar verilir. Denklem 2.49 ve 2.50'de gösterildiği şekilde tanh katmanı ile yeni aday değerler vektörü oluşturulur ve hücre durumunu güncellemek için giriş kapısı değerleri ile birleştirilir. Şekil 2.12'deki gibi temsil edilir.



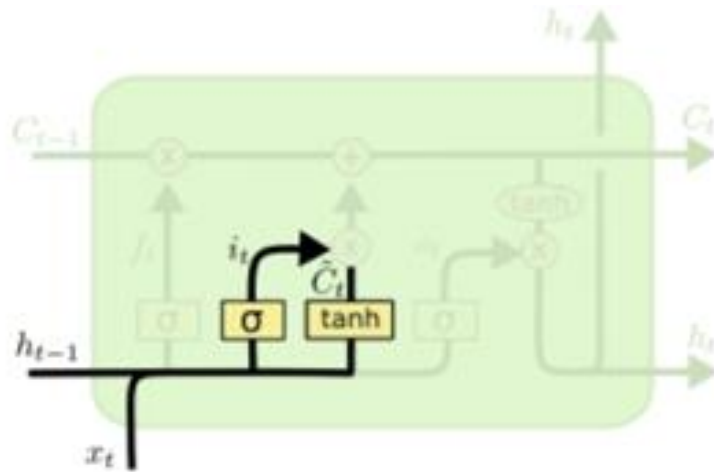
Şekil 2.11: UKSH unutma kapısı (Lysfjord, 2017).

C_t artık yeni hücre durumu olmuştur. Unutmak istediğimiz değerleri belirlediğimiz f_t unutma kapısı ile eski hücre değerleri olan C_{t-1} ile çarpılarak unutulur. i_t giriş kapısı değerleri ile $\tilde{C}_t$ yeni hücre aday değeri çarpılarak giriş kapısından geçecek bilgiler elde edilir.

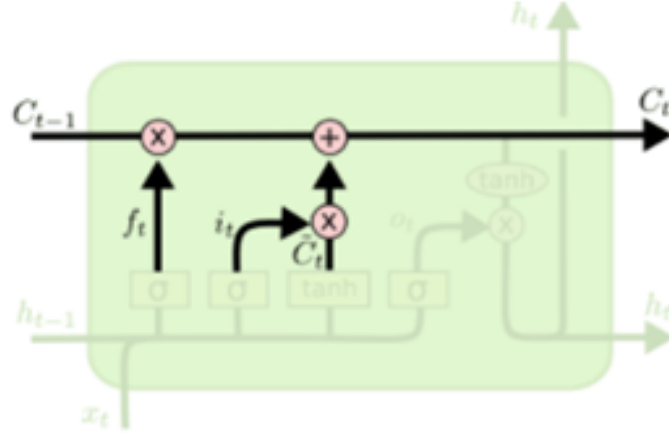
Güncellene hücre değeri ve giriş kapısından geçmesine karar verdiğimiz değerler toplanarak dahili hafızaya eklenir. Yeni bilginin aktarılması Şekil 2.13 ile temsil edilir ve denklem 2.51 ile gösterilir.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2.49)$$

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (2.50)$$

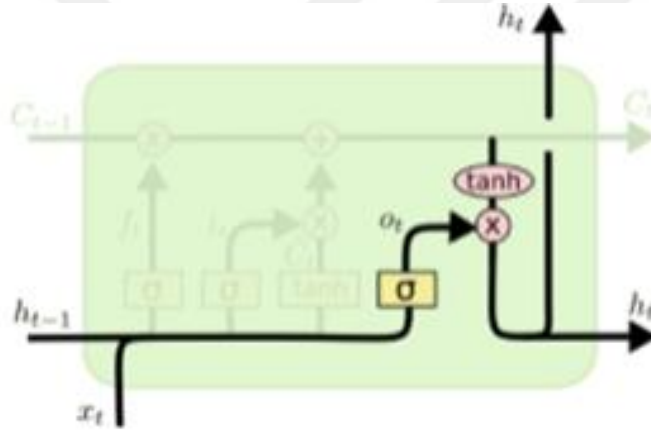


Şekil 2.12: UKSH giriş kapısı (Lysfjord, 2017).



Şekil 2.13: UKSH yeni bilginin aktarılması (Lysfjord, 2017).

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \quad (2.51)$$



Şekil 2.14: UKSH Çıkış kapısı (Lysfjord, 2017).

$$o_t = W_o \cdot [h_{t-1}, x_t] + b_o \quad (2.52)$$

$$h_t = o_t \cdot \tanh(C_t) \quad (2.53)$$

Çıktının ne olması gerektiği çıkış kapısı durumuna göre belirlenir. Hangi bölümlerin çıkacağına karar verilen bir sigmoid katmanından geçerek hücre durumuna tanh fonksiyonu uygulanarak bilginin -1 ile 1 arasında tanımlanır ve son sigmoid katmanı ile çarpılarak h_t çıkış değeri belirlenir. Çıkış değerinin belirlenmesi Şekil 2.14 ile temsil edilir, denklem 2.52 ve denklem 2.53 ile gösterilir.

3. PERFORMANS METRİKLERİ

Yapılan tahminlerin ölçülmesinde ve karşılaştırılmasında performans metrikleri kullanılır. Kullanılan tahmin yöntemlerinin veriye uygunluğu ve algoritmaların verimliliğini ölçmek için farklı yöntemler kullanılır. Literatürde yoğunlukla kullanılan performans metrikleri bu tez kapsamında kullanılacaktır. Tez kapsamında 4 temel performans metriği kullanılacaktır; R^2 , ortalama mutlak hata (MAE), kök ortalama kare hata (RMSE) ve normalleştirilmiş kök ortalama hata (nRMSE) metrikleri kullanılacaktır.

3.1 R^2

Tahmin hatası karelerinin toplamının, hesaplanan tahmini ortalama ile değiştiren kareler toplamının toplamına bölünmesiyle hesaplanır. R^2 değeri 0 ile 1 arasındadır ve daha büyük bir değer, tahmin ile gerçek değer arasında daha iyi bir uyumu gösterir. Modelin bağlı değişkenlere uyumunu gösterir. Değerin 1 yakınlığı daha iyi olduğunu gösterir. Denklem 3.1'deki gibi tanımlanır.

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2} \quad (3.1)$$

3.2 Ortalama Mutlak Hata (MAE)

Gerçek değer ile tahmin edilen değer arasındaki her bir farkın mutlak değerinin ortalamasıdır. Bir dizi tahmindeki hataların yönlerini dikkate almadan ortalama büyüklüğünü ölçen, tüm tekil hataların ortalamada eşit olarak ağırlıklandırıldığı doğrusal bir metriktir. 0'a yakın değerler daha iyi tahmin edildiğini gösterir. Denklem 3.2'deki gibi tanımlanır.

$$MAE = \frac{1}{n} \sum_{i=1}^n |e_i| \quad (3.2)$$

n örnek sayısını ve e her tahminin hatasını göstermektedir.

3.3 Kök Ortalama Hata (RMSE)

Ortalama hatanın kareköküdür. Tahmin hatalarının standart sapmasıdır. Büyük hataları daha fazla cezalandırmanın avantajına sahiptir. Tahmin hataları, regresyon çizgisi veri noktalarından ne kadar uzakta olduğunun ve bu tahmin hatalarını ne kadar yayıldığının bir ölçüsüdür. Bu ölçü, verilerin en uygun çizgi etrafında ne kadar yoğun olduğunun analizini gösterir. 0'a yakın değerler daha iyi tahmin edildiğini gösterir. Denklem 3.3'teki gibi tanımlanır.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n e_i^2} \quad (3.3)$$

n örnek sayısını ve e her bir tahminin hatasını göstermektedir.

3.4 Normalleştirilmiş Kök Ortalama Hata (nRMSE)

Kök mutlak hatanın gerçek değerin en büyük ve en küçük değerlerin farkı ile standartlaştırılmış halidir. Büyük hataları yüksek dereceden cezalandırır. 0'a yakın değerler daha iyi tahmin edildiğini gösterir. Denklem 3.4'teki gibi tanımlanır.

$$nRMSE = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n e_i^2}}{y_{max} - y_{min}} \quad (3.4)$$

n örnek sayısını, y gerçek değeri ve e her bir tahminin hatasını göstermektedir.

4. SİMÜLASYON SONUÇLARI

Simulasyonların koşulabilmesi için ilk aşama elimizdeki verileri sınıflandırmak ve yeterliliklerini analiz etmektir. Analizler sonucunda seçilen değişkenler ile sistem tanılama, makine öğrenmesi ve derin öğrenme Şanlıurfa verileri kullanılarak modeller oluşturulmuştur. Oluşturulan modeller Adıyaman, Gaziantep, Diyarbakır ve Hatay verilerine uygulanarak tahmin modellerinin hataları simulasyon sonuçları kısmında detayları ile gösterilmiştir.

4.1 Meteorolojik Verilerin Hazırlanması

Tez kapsamında Güneydoğu Anadolu bölgesinde bulunan 5 ilden 10 dakikalık veriler kullanılmıştır. Verilerin alındığı il ve ilçeler şu şekildedir, Şanlıurfa (Keçili), Adıyaman (Eskihüsni), Gaziantep (Kopcağız), Diyarbakır (GAPUATEM), Hatay (Dr. Güneş Bahadırılı) olarak ölçüm merkezlerinden alınan veriler kullanılmıştır. Veri sayıları ve tarih aralıkları Çizelge 4.1 de gösterildiği gibidir.

Çizelge 4.1: Ölçüm merkezleri ve veri detayları.

Ölçüm Merkezi	Tarih aralığı	Veri sayısı
Şanlıurfa	2012-2017	277791
Adıyaman	2013-2017	231516
Gaziantep	2013-2017	244810
Diyarbakır	2012-2017	232536
Hatay	2013-2017	244626

4.1.1 Şanlıurfa meteorolojik veri analizi

Bu tez kapsamında tahmin modelleri Şanlıurfa verilerine göre yapılacaktır. Şanlıurfa verisi, sayı ve kullanılacak olan veri özellikleri olarak en uygun verilere sahiptir. Verilerin yapısı aşağıda detaylı şekilde verilmektedir.

5 ilin verileri de aynı meteorolojik bilgilere sahiptir. Meteorolojik verilerin içeriği sıcaklık verileri, rüzgar hızı, ultraviyole ışın, güneş radyasyonu, yağış verileri, nem verileri, hava basıncı verisi, ET0 verisi, ıslaklık verileri, çiğlenme verisi, güneş enerji miktarı verisi, güneşlenme süresi verisi, sıcaklık kızılötesi verisi olmak üzere 19 farklı veri içermektedir. Bu verilerden ultraviyole ışın, nem toprak8 ve nem toprak9 verileri yetersiz olduğundan kullanılmamıştır.

Korelasyon, iki sayısal ölçüm arasında doğrusal bir ilişki olup olmadığını, varsa bu ilişkinin yönünü ve şiddetinin ne olduğunu belirlemek için kullanılan bir istatistiksel yöntemdir. Korelasyon değeri -1 ve 1 arasında değişmektedir. Korelasyon katsayısı negatif ise iki değişken arasında ters ilişki vardır, yani değişkenlerden biri artarken diğeri azalmaktadır denir. Korelasyon katsayısı pozitif ise değişkenlerden biri artarken diğeri de artmaktadır yorumu yapılır. Pearson korelasyon katsayısı, iki değişkenin kovaryans değerinin, her bir değişkenin standart sapmasının çarpımına bölünmesi ile elde edilebilir. Pearson korelasyon katsayısının yorumlamak için Çizelge 4.2 kullanılabilir.

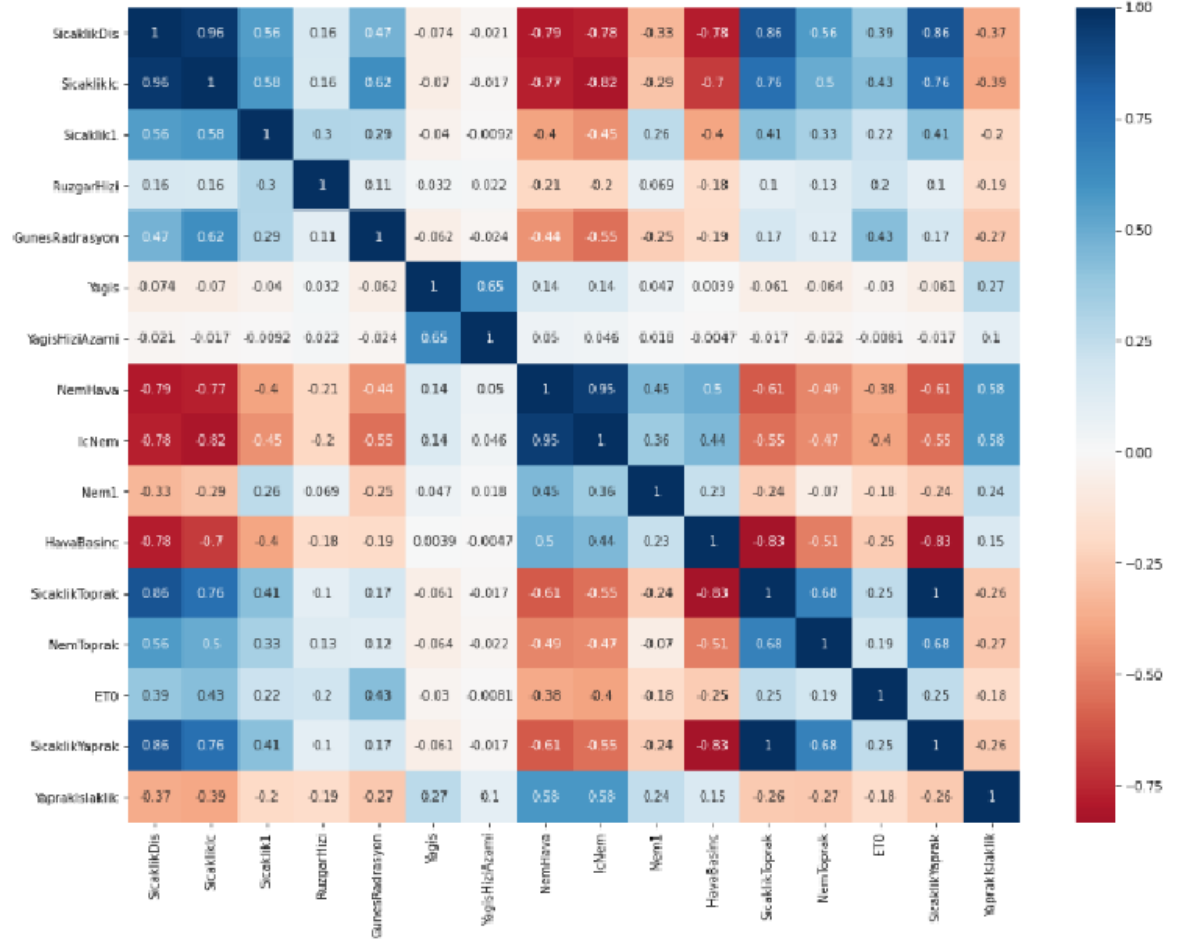
Çizelge 4.2: Pearson matrisi yorumlanması.

Korelasyon katsayısı	Yorumu
$r < 0,2$	Çok zayıf ilişki
$0,2 < r < 0,4$	Zayıf ilişki
$0,4 < r < 0,6$	Orta ilişki
$0,6 < r < 0,8$	Yüksek ilişki
$r > 0,8$	Çok yüksek ilişki

Güneş radyasyonu ve sıcaklık dış arasındaki korelasyon 0,47 olarak orta ilişkili olarak tanımlanıyor. Güneş radyasyonu sıcaklık iç arasındaki korelasyon 0,62 olarak yüksek ilişkili olarak tanımlanıyor. Güneş radyasyonu ve nem hava arasındaki korelasyon -0,44 olarak orta ilişkili olarak tanımlanıyor. Güneş radyasyonu ve iç nem arasındaki korelasyon -0,55 olarak orta ilişkili tanımlanıyor.

Güneş radyasyonu ve ET0 arasındaki korelasyon 0,43 olarak orta ilişkili tanımlanıyor. Güneş radyasyonu ve hava basıncı arasındaki korelasyon -0,19 olarak zayıf ilişkili tanımlanıyor. Güneş radyasyonu ve yaprak ıslaklık arasındaki korelasyon 0,27 olarak zayıf tanımlanıyor. Diğer verilere baktığımız zaman zayıf ve çok zayıf ilişkili olarak

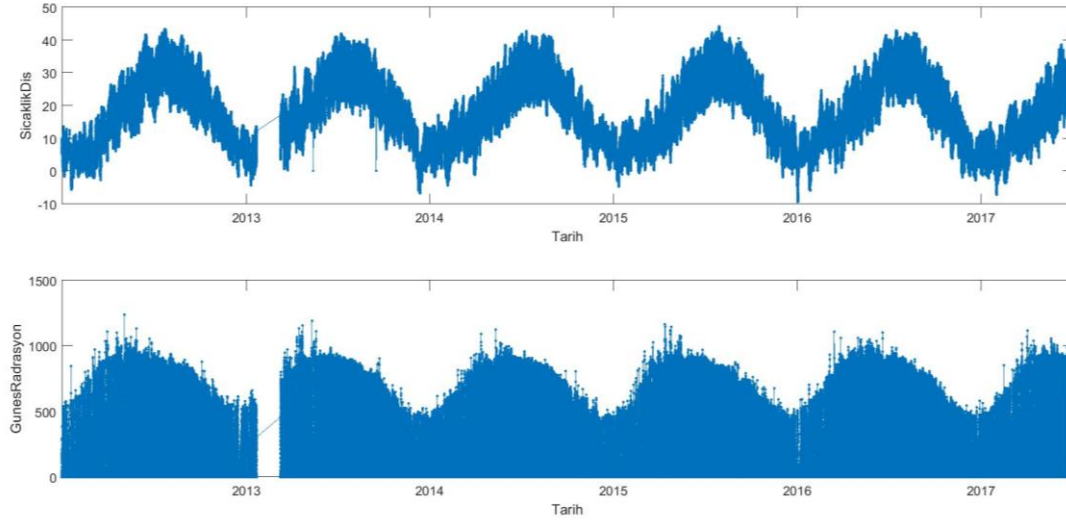
tanımlanıyor. İç nem ve iç sıcaklık ölçüm cihazının kendine ait değerler olması sebebiyle kullanılmayacaktır. Pearson matrisine göre güneş radyasyonu ile anlamlı ilişkiye sahip veriler sıcaklık dış, nem hava, ET0, hava basıncı ve yaprak ıslaklık olarak tanımlanabilir.



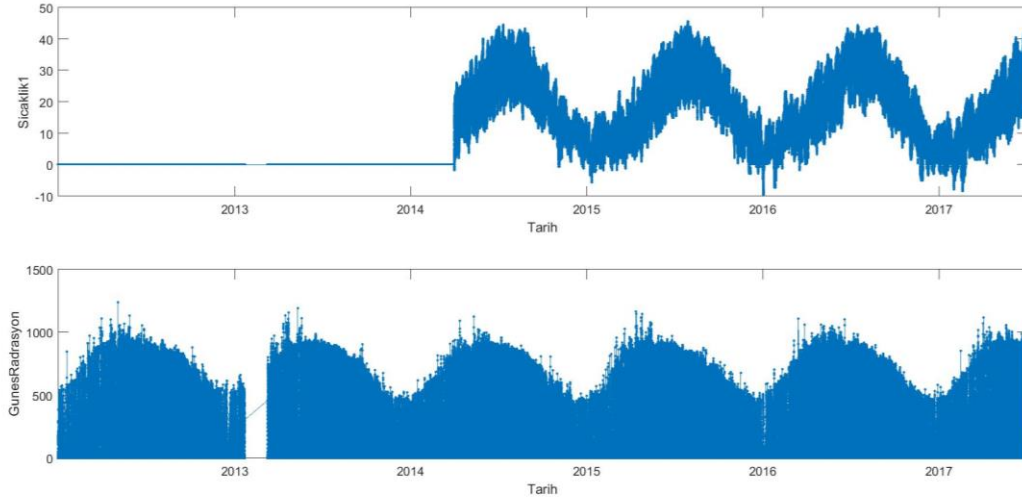
Şekil 4.1: Şanlıurfa verisi pearson matrisi.

Grafik üzerinden verilerin güneş radyasyonu ile ilişkilerini ve verilerin yeterliliğini incelersek Şekil 4.2 de sıcaklık dış ve güneş radyasyonu arasındaki ilişki gösterilmiştir. Sıcaklık dış ve güneş radyasyonu arasında benzer bir veri yapısı Şekil 4.2’de görülüyor. Sıcaklık dış değerinde aykırı değerler görünmüyor ve veri sayısı yeterli görülüyor. Bu sebeple kullanıma uygundur.

Sıcaklık1 ile güneş radyasyonu arasındaki ilişki Şekil 4.3'te görünmesine rağmen verilerin eksik olması sebebiyle kullanılmaz bir ölçüm oluyor.



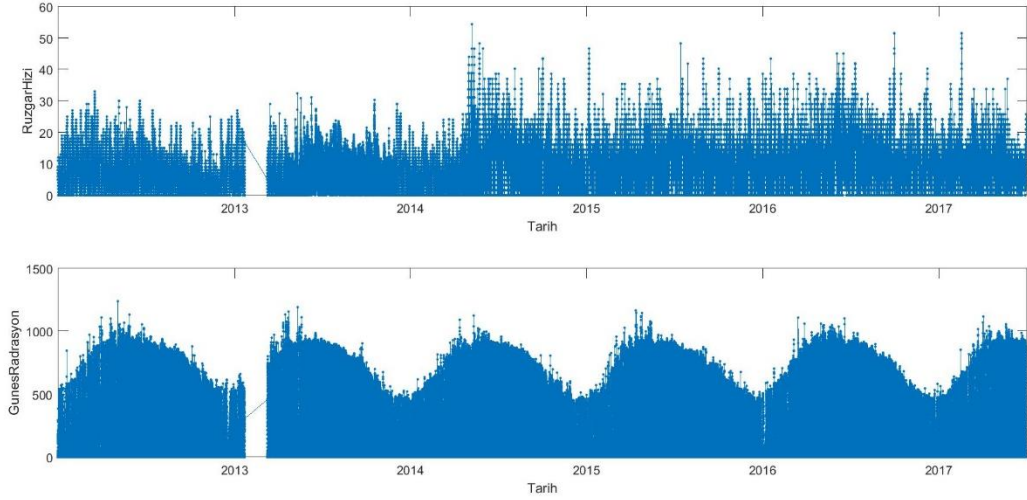
Şekil 4.2: Şanlıurfa sıcaklık dış ve güneş radyasyonu ilişkisi.



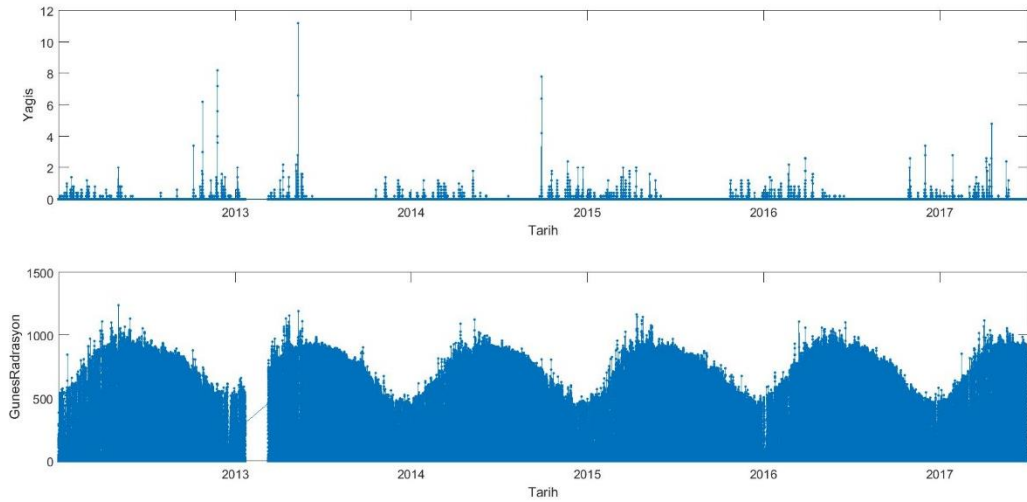
Şekil 4.3: Şanlıurfa sıcaklık1 ve güneş radyasyonu ilişkisi.

Rüzgar hızlı ve güneş radyasyonu arasında uygun bir ilişki Şekil 4.4'te görülmemektedir. Verilerde aykırı değerler bulunuyor ve rüzgar yönü bilgisi olmadan rüzgar hızı tek başına anlamlı olmayabilir. Bu sebeple rüzgarı hızı verisi kullanılmayacaktır.

Yağış ve güneş radyasyonu arasındaki Şekil 4.5’te gösterildiği gibidir. Yağış verilerin yetersiz olduğu ve ayrık değerlere sahip olduğu gözükmemektedir. Bu sebeple yağış verisi kullanılmayacaktır.



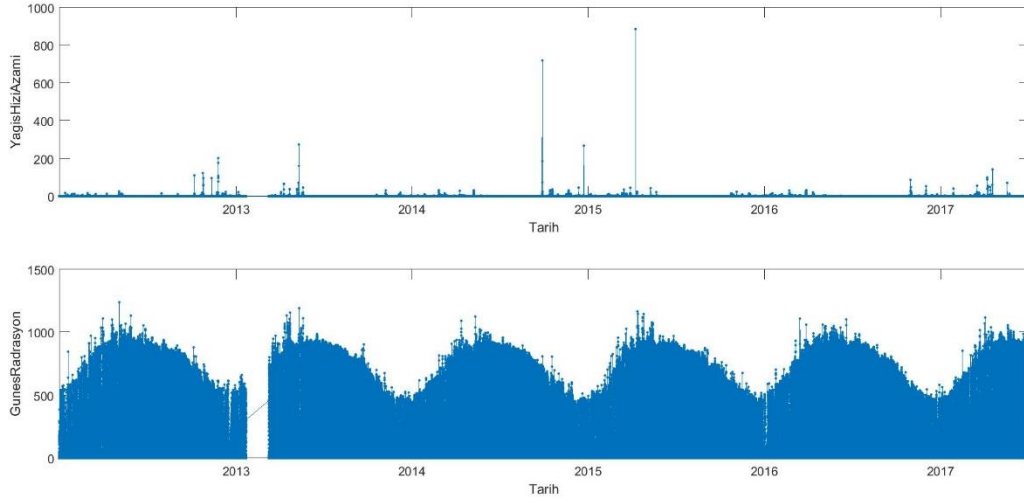
Şekil 4.4: Şanlıurfa rüzgar hızı ve güneş radyasyonu ilişkisi.



Şekil 4.5: Şanlıurfa yağış ve güneş radyasyonu ilişkisi.

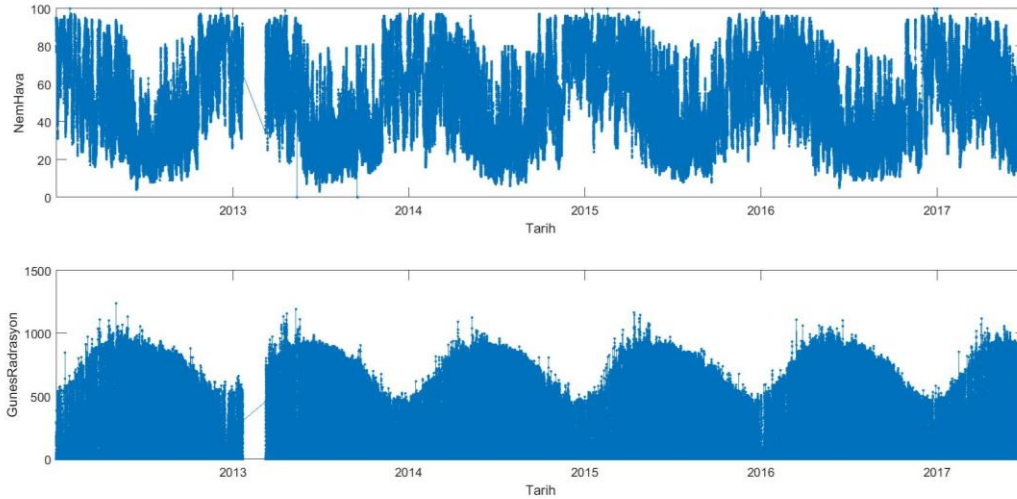
Azami yağış hızı ve güneş radyasyonu arasındaki ilişki Şekil 4.6’da gösterildiği gibidir. Yağış verisinde olduğu gibi azami yağış hızı verisi de yeterli sayıda olmadığı ve ayrık değerlere sahip olduğu görülmektedir. Ancak yağış hızı verisinin Şekil 4.5’te görülen yağış verisi ile korelasyonda olduğu görülmektedir. Yağış ile ilgili ölçümlerin

yeterli olmadığını söyleyebiliriz. Bu sebeple azami yağış hızı verisi kullanılmayacaktır.



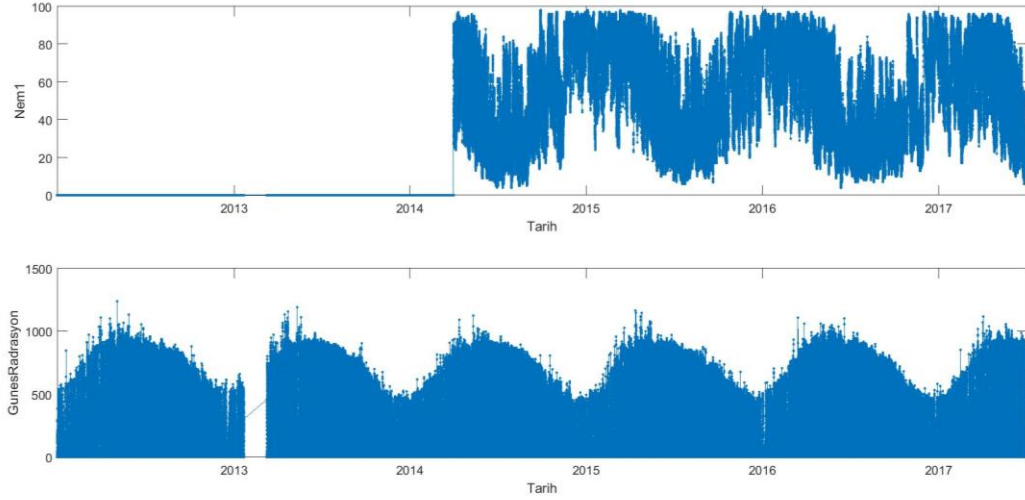
Şekil 4.6: Şanlıurfa azami yağış hızı ve güneş radyasyonu ilişkisi.

Nem hava ve güneş radyasyonu arasındaki ilişki Şekil 4.7’de gösterildiği gibidir. Aralarında ters bir ilişki vardır ve nem hava değeri azaldıkça güneş radyasyonu değeri artmaktadır. Nem hava değeri ayırık değerlere sahip olmadığı ve yeterli veriye sahip olduğu görülmektedir.



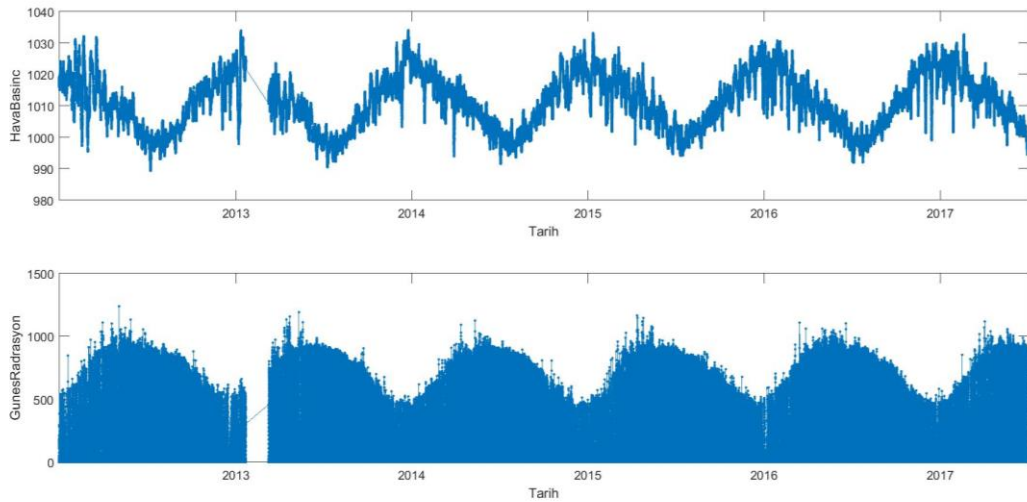
Şekil 4.7: Şanlıurfa nem hava ve güneş radyasyonu ilişkisi.

Nem1 ve güneş radyasyonu arasındaki ilişkisi Şekil 4.8’de gösterildiği gibidir. Aralarında ters bir ilişki vardır. Ancak 2014 öncesi verileri eksik olması sebebiyle kullanılmayacaktır.



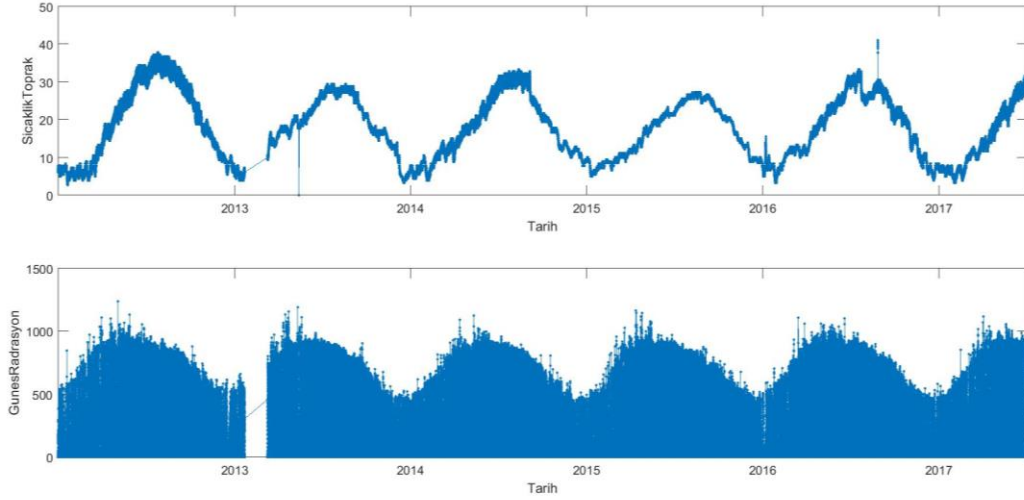
Şekil 4.8: Şanlıurfa nem1 ve güneş radyasyonu ilişkisi.

Nem1 ve güneş radyasyonu arasındaki ilişkisi Şekil 4.9’da gösterildiği gibidir. Aralarında ters bir ilişki vardır. Verilerde ayrıık değerlerin fazla olmaması ve yeterli veri sayısı olduğu için kullanıma uygun olarak görünüyor.



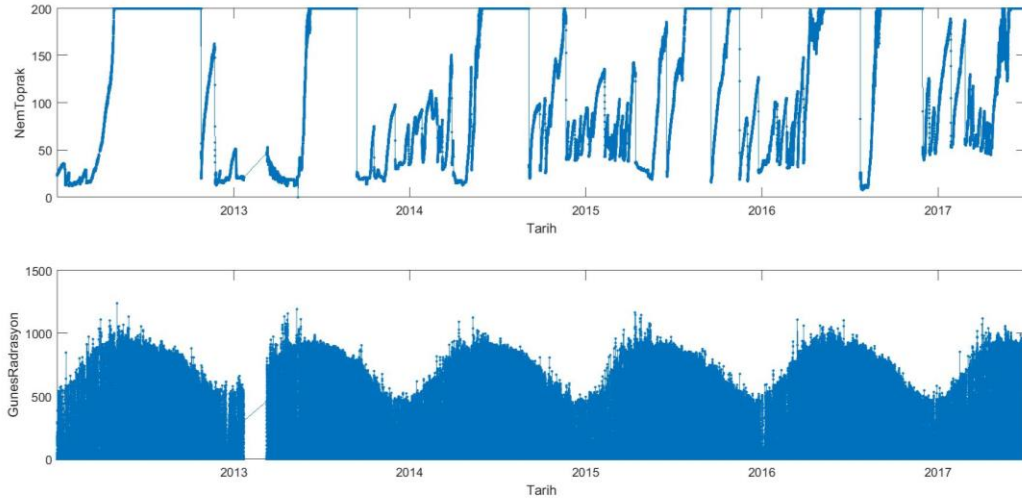
Şekil 4.9: Şanlıurfa hava basıncı ve güneş radyasyonu ilişkisi.

Sıcaklık toprak ve güneş radyasyonu arasındaki ilişkisi Şekil 4.10'da gösterildiği gibidir. Aralarında benzer bir örüntüye sahip olduğu görülüyor, ayırık değerlere sahip olmaması ve yeterli veri olmasından dolayı uygun bir veridir.



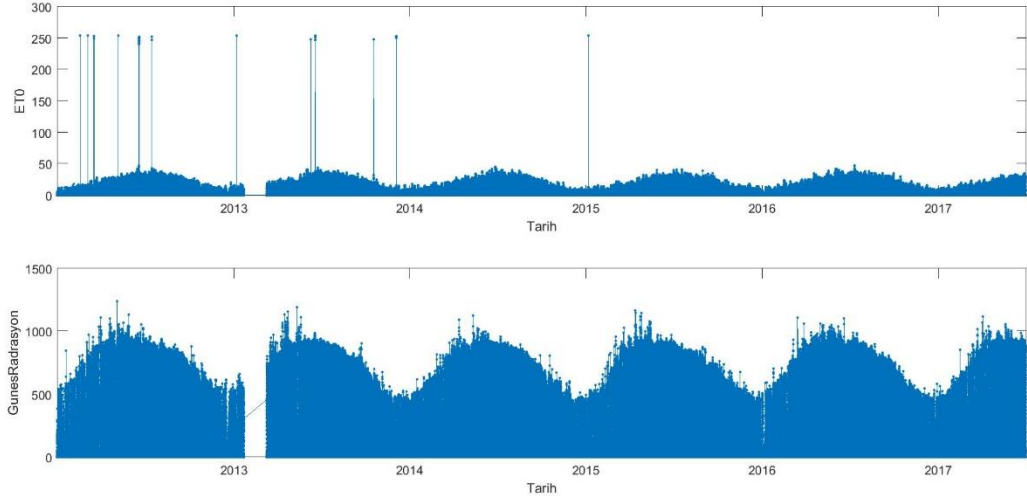
Şekil 4.10: Şanlıurfa sıcaklık toprak ve güneş radyasyonu ilişkisi.

Nem toprak ve güneş radyasyonu arasındaki ilişkisi Şekil 4.11'de gösterildiği gibidir. Ölçüm cihazındaki hatalardan dolayı değerler doygunluğa ulaşmıştır. Güneş radyasyonu ile ilişkisinin düşük olduğu görülmektedir.



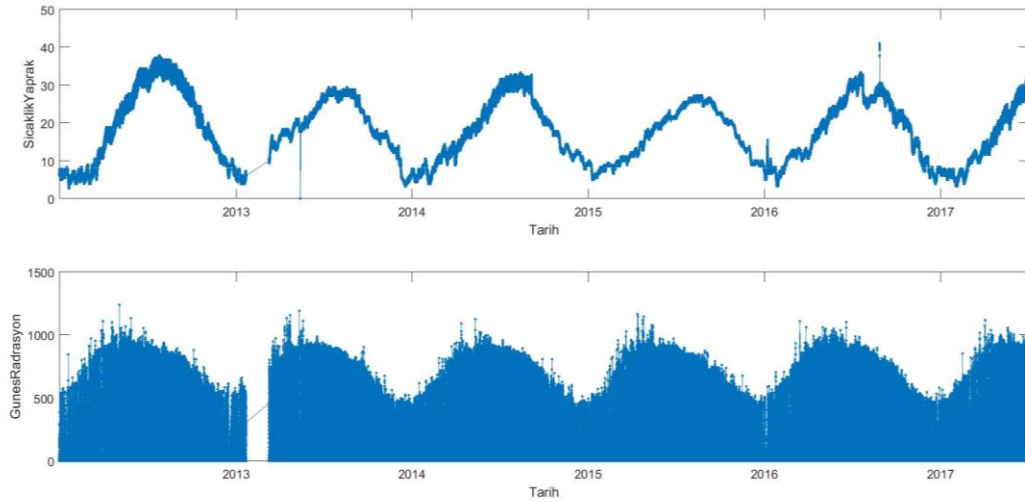
Şekil 4.11: Şanlıurfa nem toprak ve güneş radyasyonu ilişkisi.

ET0 ve güneş radyasyonu arasındaki ilişkisi Şekil 4.12'de gösterildiği gibidir. ET0 ile güneş radyasyonu arasında benzer bir örüntü vardır.



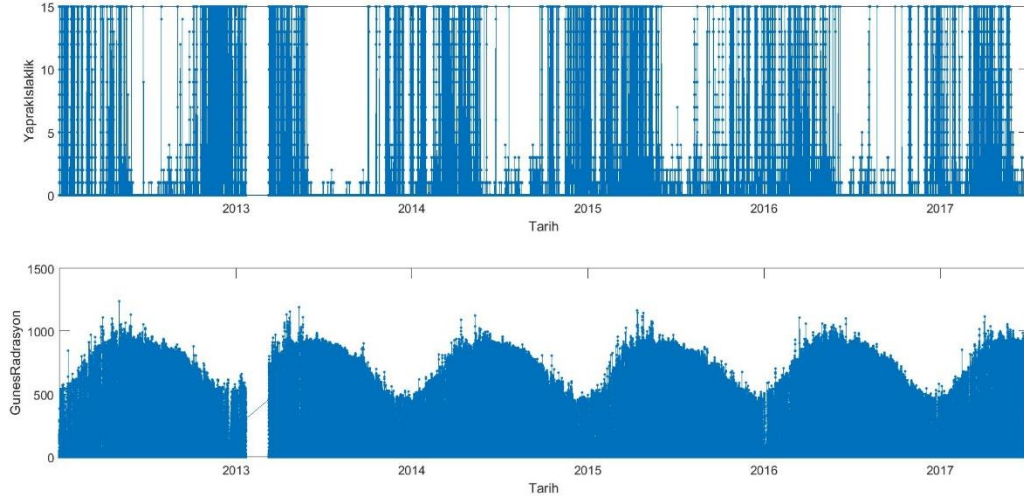
Şekil 4.12: Şanlıurfa ET0 ve güneş radyasyonu ilişkisi.

ET0 arttıkça güneş radyasyonu artmakta ve ET0 azaldıkça güneş radyasyonu azalmaktadır. Veride ayırık değerler vardır ancak genel olarak verilerin tamlığı ve güneş radyasyonu ile olan ilişkisi olarak kullanıma uygundur. Sıcaklık yaprak ve güneş radyasyonu arasındaki ilişkisi Şekil 4.13’de gösterildiği gibidir.



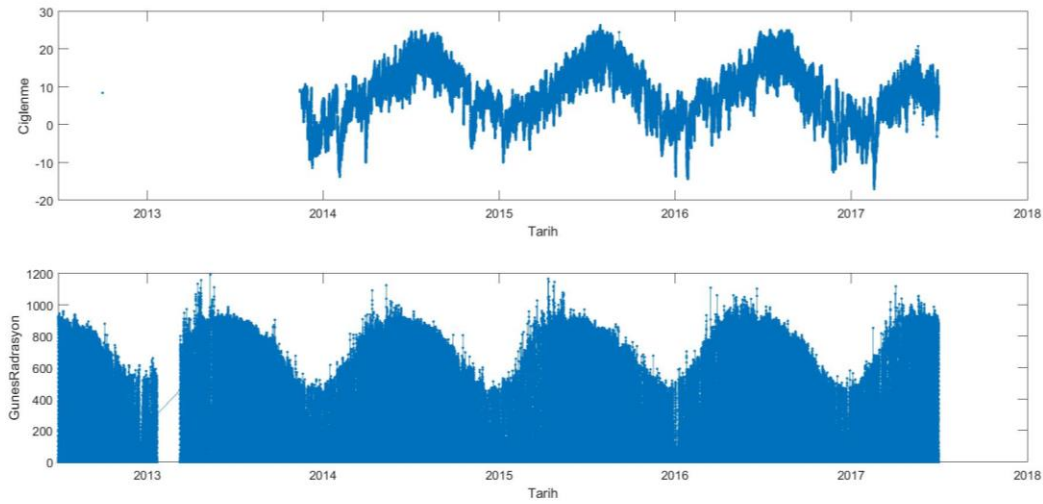
Şekil 4.13: Şanlıurfa sıcaklık yaprak ve güneş radyasyonu ilişkisi.

Aralarındaki ilişki benzer bir örüntü içinde gerçekleşmektedir. Veri eksikliğinin olmaması ve ayırık değerlerin azlığı sebebiyle kullanıma uygun olarak değerlendirilir.



Şekil 4.14: Şanlıurfa yaprak ıslaklık ve güneş radyasyonu ilişkisi.

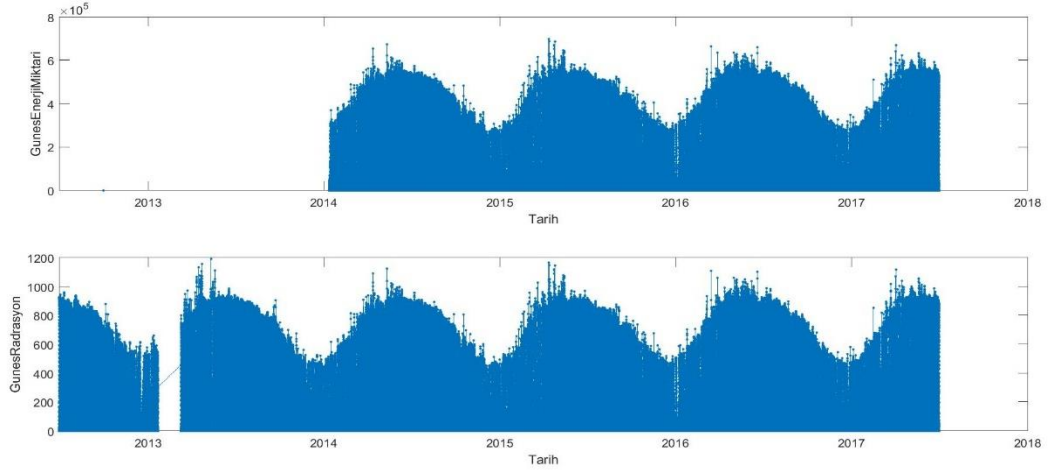
Yaprak ıslaklık ve güneş radyasyonu arasındaki ilişkisi Şekil 4.14’de gösterildiği gibidir. Yaprak ıslaklık verileri ölçüm hataları sebebiyle doygunluğa ulaşmış ve verilerde sapmalar olmuştur. Verilerin yeterli olmaması sebebiyle yaprak ıslaklık kullanılmayacaktır. Çiğlenme ve güneş radyasyonu arasındaki ilişkisi Şekil 4.15’de gösterildiği gibidir.



Şekil 4.15: Şanlıurfa çiğlenme ve güneş radyasyonu ilişkisi.

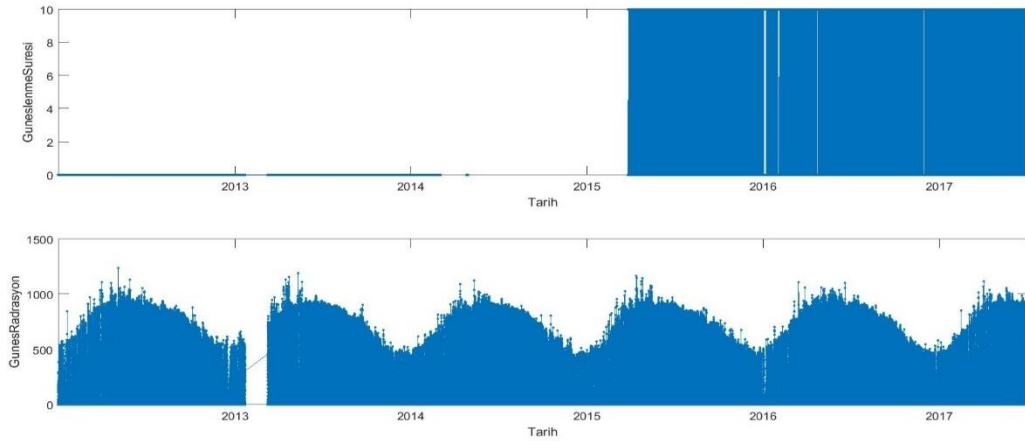
Aralarındaki ilişki güneş radyasyonu artarken çiğlenme de artıyor ve azalırken beraber azalıyor. Çiğlenme verisinin yapısı uygun olmasına rağmen 2014 öncesi verilerin eksik olması kullanılmamasını sağlıyor.

Güneş enerji miktarı ve güneş radyasyonu arasındaki ilişkisi Şekil 4.16’de gösterildiği gibidir. Güneş enerji miktarı doğrudan güneş radyasyonu bağlı bir veri olduğunu anlayabiliyoruz ancak 2014 öncesi verilerin eksik olması sebebiyle kullanılamayacaktır.



Şekil 4.16: Şanlıurfa güneş enerji miktarı ve güneş radyasyonu ilişkisi.

Güneşlenme süresi ve güneş radyasyonu arasındaki ilişkisi Şekil 4.17’de gösterildiği gibidir. Güneşlenme süresinde ölçüm hataları olduğu için kullanılmayacaktır



Şekil 4.17: Şanlıurfa güneşlenme süresi ve güneş radyasyonu ilişkisi.

Pearson korelasyon matrisi sonucunda sıcaklık dış, nem hava, ET0, hava basıncı ve yaprak ıslaklık verileri kullanılabilir ve anlamlı veriler olarak gözüküyor. Şekiller ile yapılan veri analizinde ise sıcaklık dış, nem hava, hava basıncı, sıcaklık toprak, ET0 ve sıcaklık yaprak verileri kullanıma uygun olarak analiz edilmiştir. Sıcaklık

yaprak ve sıcaklık toprak verilerinin Pearson korelasyon matrisi sonucunda ilişkilerinin yeterli olmaması sebebiyle kullanılmayacaktır. Yaprak ıslaklık verisi Pearson korelasyon matrisi sonucunda uygun bir veri olarak görünüyorsa da grafik analizi sonucunda yeterli olmadığı anlaşılmıştır ve kullanılmayacaktır.

Tez kapsamında kullanılacak veriler; sıcaklık dış, nem hava, ET0 ve hava basıncı olarak karar verilmiştir. Kullanılan verilerin açıklamaları ve tez kapsamında kullanılacak değişken isimleri Çizelge 4.3'te verilmiştir.

Çizelge 4.3: Meteorolojik parametrelerin açıklanması.

Parametre	Açıklama	Birimi
Sıcaklık	Atmosferik sıcaklık	°C
Nem	Bağıl nem	%
ET0	Evapotranspirasyon	mm
Hava basıncı	Atmosferik basınç	Pa
Güneş radyasyonu	Global güneş ışınımı	W/m ²

Çizelge 4.4'te Şanlıurfa için meteorolojik verilerin analizi yapılmıştır. Analiz sonucunda ayırık verilere rastlanmamıştır.

Veri hazırlama kısmında, eksik olan verileri MATLAB'in 'fillmissing' fonksiyonu 'spline' methodu kullanarak tamamlanmıştır. ET0 verisi, 45 değeri ile limitlenerek ayırık verilerden uzaklaştırılmıştır. Güneş radyasyonu verisi, 0 ile 1200 arasında limitlenmiştir ve ayırık verilerden uzaklaştırılmıştır.

1970 yılında D. A. Dickey ve W. A. Fuller tarafından tanımlanan geliştirilmiş Dickey-Fuller yöntemi zaman serilerinde birim kök içerip içermediğini test etmeye yarayan istatistiksel bir analizdir. Boş hipotezi, zaman serisinin bir birim kökü olduğunu, yani durağan olmadığını gösterir ve zamana bağlı bir yapısı vardır. Alternatif hipotez, boş hipotez reddedilir ve zaman serisinin birim kökü olmadığını, yani durağan olduğunu gösterir. Zamana bağlı bir yapıya sahip değildir. Bu sonucu ADF testindeki p değerini kullanarak yorumluyoruz. Eğer p değeri 0,5'ten küçükse boş hipotezi reddedilir ve verilerde birim kök yoktur, zaman serisi durağan yapıdadır. Eğer p değeri

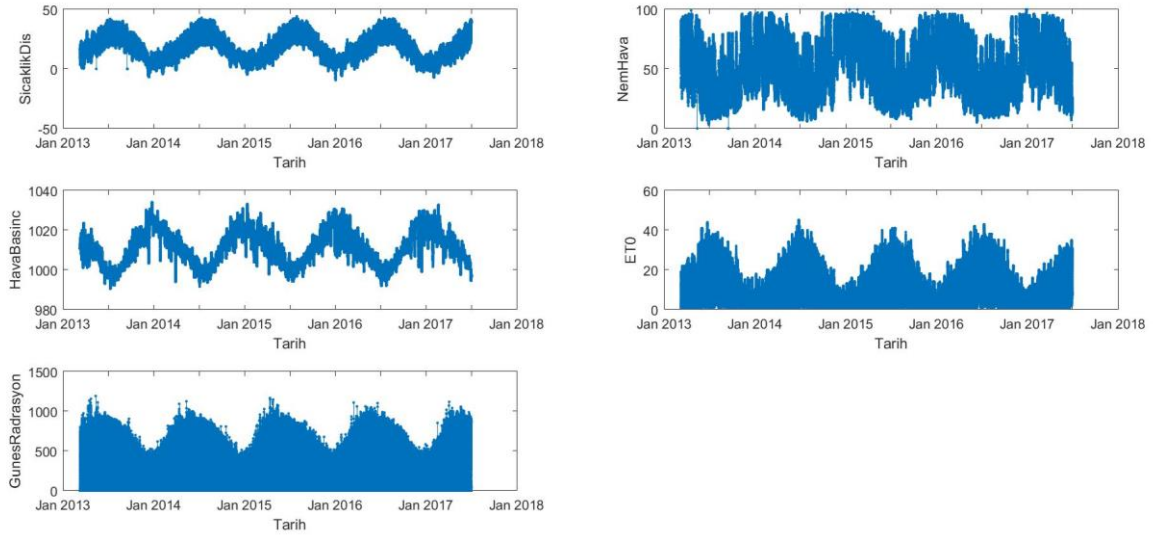
0,5'ten büyükse boş hipotezini reddemeyiz ve verilerde birim kök vardır, zaman serisi durağan olmayan yapıdadır.

Çizelge 4.4: Şanlıurfa meteorolojik parametrelerin analizi.

Parametre	Minimum	Maksimum	Standart sapma	Ortalama
Sıcaklık	-9,27	44,2	10,47	18,37
Nem	3	100	24,29	52,22
ET0	0	45	8,65	8,34
Hava basıncı	989,4	1034,1	8,41	1011,07
Güneş radyasyonu	0	1192	281,52	199,73

Veriler, %20 test ve %80 eğitim verisi olarak ayrılmıştır. Şanlıurfa verisindeki ölçüm hatası bulunan kısımdan eğitim ve test olarak ayrılmıştır. Şekil 4.18'de eğitim verisi ve Şekil 4.19'da test verisi olarak gösterilmiştir.

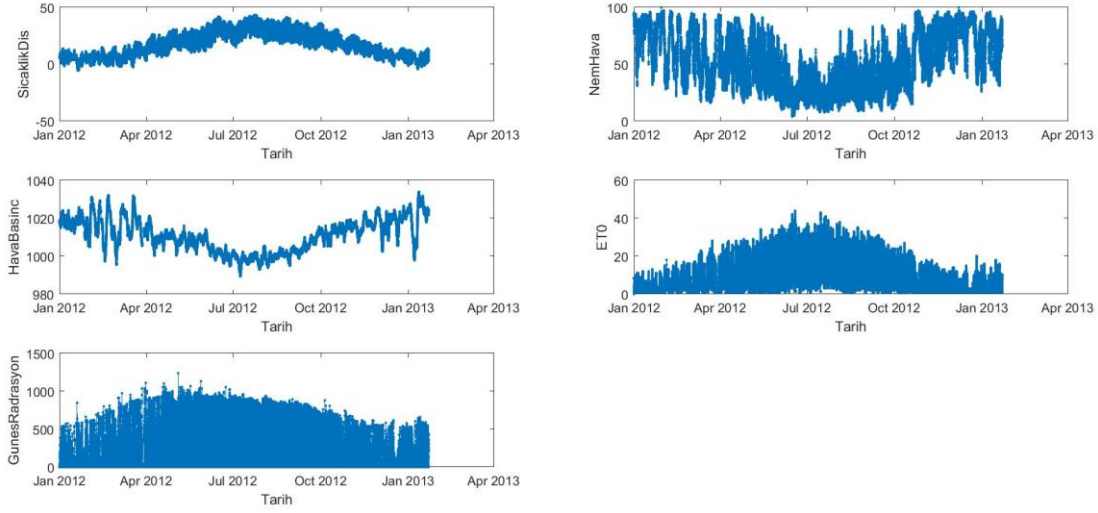
Şanlıurfa verilerine ADF testi uygulandığında 5 değişken değeri için p değeri 0,5'in altında olduğu ve zaman serisinin durağan olduğu analizi yapılabilir.



Şekil 4.18: Şanlıurfa eğitim verisi.

Şanlıurfa eğitim verisi 4 yıllık sıcaklık, nem, hava basıncı, ET0 ve güneş radyasyonu verileri içermektedir.

Şanlıurfa test verisi 1 yıllık sıcaklık, nem, hava basıncı, ET0 ve güneş radyasyonu verilerini içermektedir.

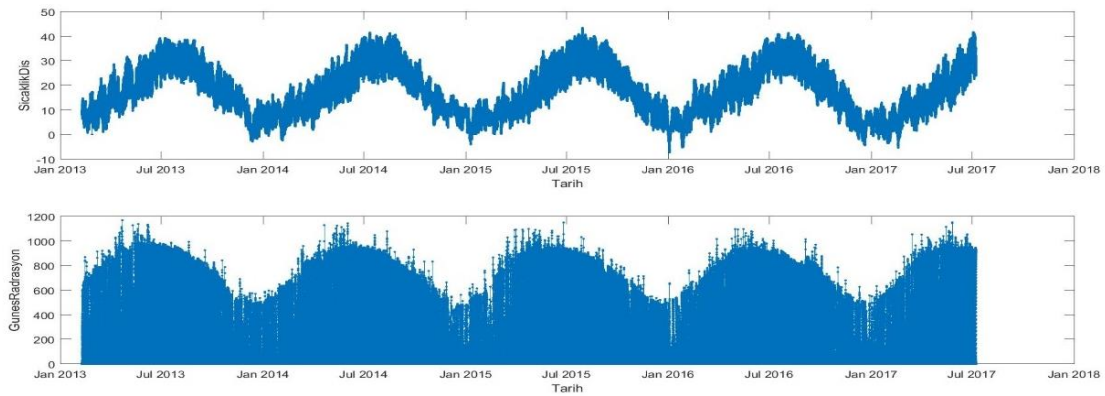


Şekil 4.19: Şanlıurfa test verisi.

Şanlıurfa verileri kullanılarak oluşturulan tahmin modelleri Adıyaman, Diyarbakır, Gaziantep ve Hatay ölçüm merkezlerinden alınan veriler ile doğrulaması yapılacaktır. Bu sebeple diğer ölçüm merkezleri için sıcaklık, nem, hava basıncı, ET0 ve güneş radyasyonu verilerinin analizi yapılacaktır.

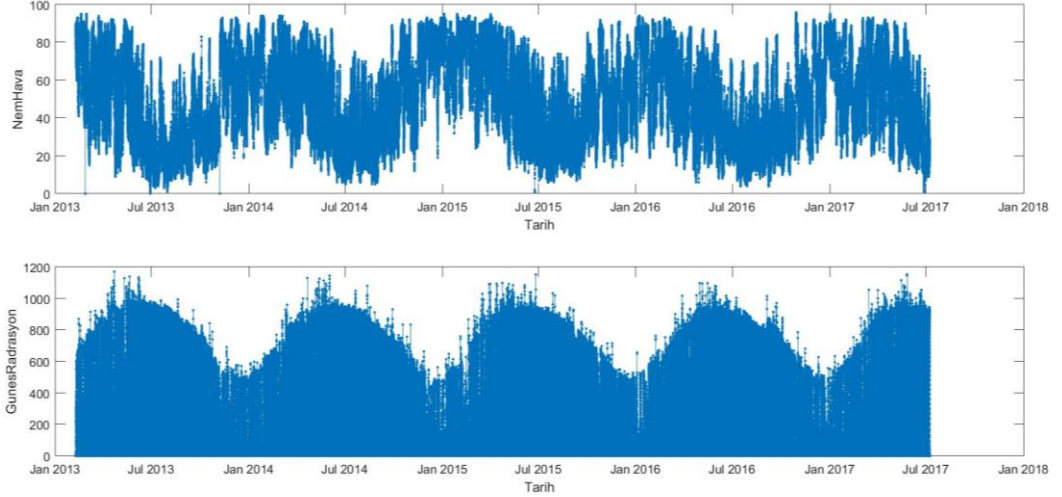
4.1.2 Adıyaman meteorolojik veri analizi

Adıyaman için sıcaklık ve güneş radyasyonu arasındaki ilişki Şekil 4.20'de görüldüğü gibi aralarında bir örüntü vardır.



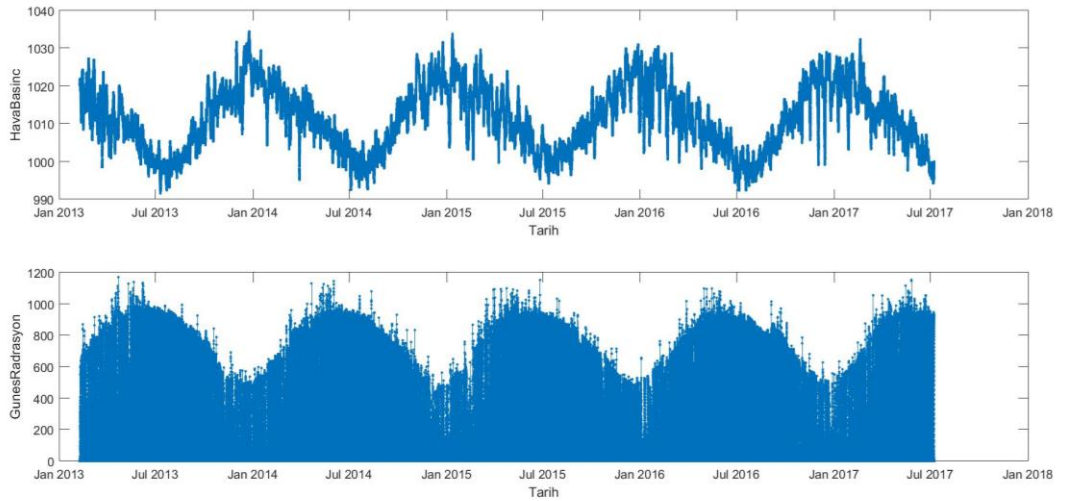
Şekil 4.20: Adıyaman sıcaklık dış ve güneş radyasyonu ilişkisi.

Eksik veri olmaması ve ayırık değerlerin olmaması kullanım olarak uygun olduğunu gösteriyor. Adıyaman için nem hava ve güneş radyasyonu arasındaki ilişki Şekil 4.21’de görüldüğü gibi aralarında ters bir ilişki vardır. Eksik veri olmaması ve ayırık değerlerin olmaması kullanım olarak uygun olduğunu gösteriyor.



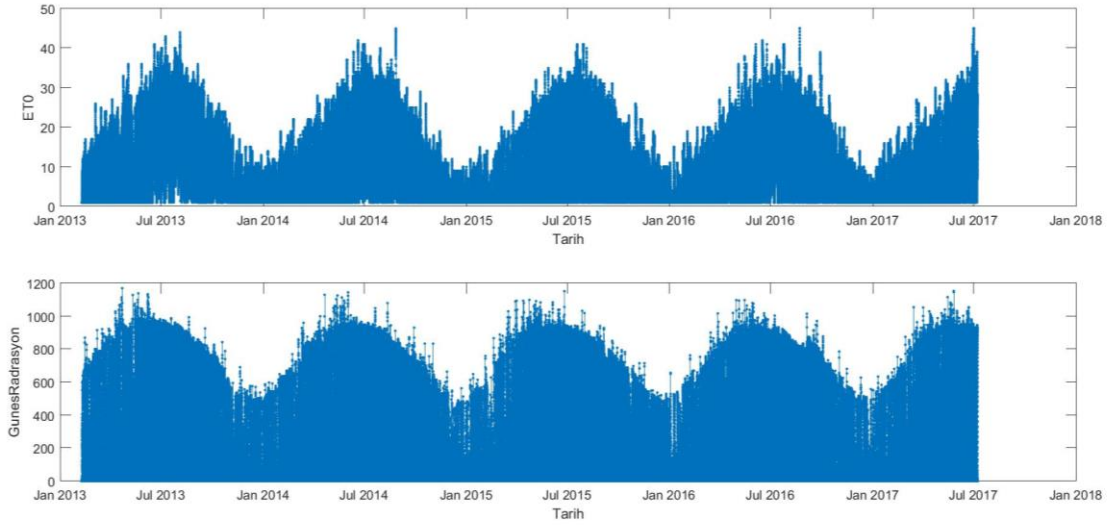
Şekil 4.21: Adıyaman nem hava ve güneş radyasyonu ilişkisi.

Adıyaman için hava basıncı ve güneş radyasyonu arasındaki ilişki Şekil 4.22’de görüldüğü gibi aralarında ters bir ilişki vardır. Eksik veri olmaması ve ayırık değerlerin olmaması kullanım olarak uygun olduğunu gösteriyor.



Şekil 4.22: Adıyaman hava basıncı ve güneş radyasyonu ilişkisi.

Adıyaman için ET0 ve güneş radyasyonu arasındaki ilişki Şekil 4.23’de görüldüğü gibi aralarında benzer bir ilişki vardır. Eksik veri olmaması ve ayrık değerlerin olmaması kullanım olarak uygun olduğunu gösteriyor.



Şekil 4.23: Adıyaman ET0 ve güneş radyasyonu ilişkisi.

Çizelge 4.5’te Adıyaman için meteorolojik verilerin analizi yapılmıştır. Analiz sonucunda ayrık verilere rastlanmamıştır. Şanlıurfa bölgesine mesafe olarak uzak olmaması sebebiyle ortalama değerlerin birbirine yakın olduğu görülmektedir.

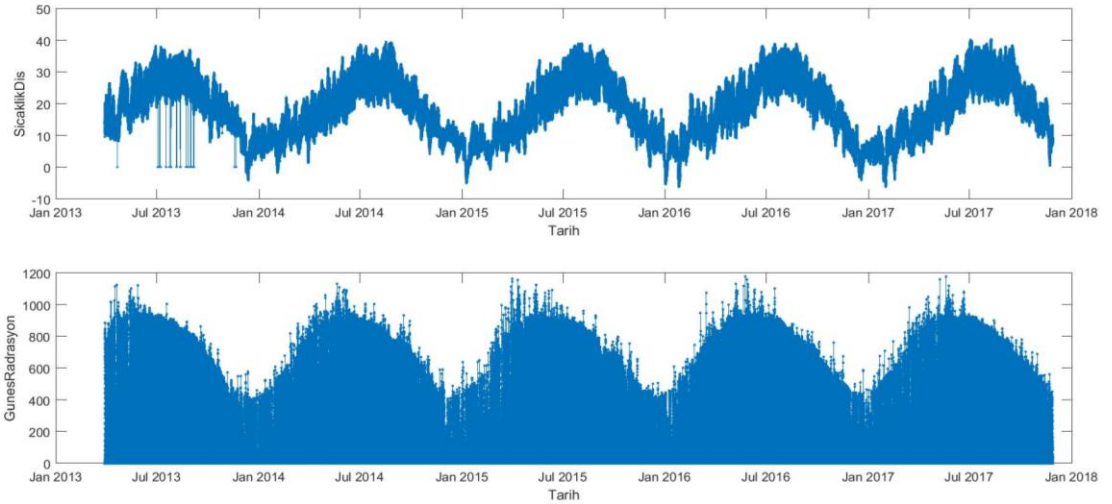
Çizelge 4.5: Adıyaman meteorolojik parametrelerin analizi.

Parametre	Minimum	Maksimum	Standart sapma	Ortalama
Sıcaklık	-7,22	43,5	9,99	18,03
Nem	0	96	22,96	49,39
ET0	1	45	8,69	8,13
Hava basıncı	991,56	1034,47	8,35	1011,31
Güneş radyasyonu	0	1170	289,71	204,60

Adıyaman verilerine ADF testi uygulandığında 5 değişken değeri için p değeri 0,5’in altında olduğu ve zaman serisinin durağan olduğu analizi yapılabilir.

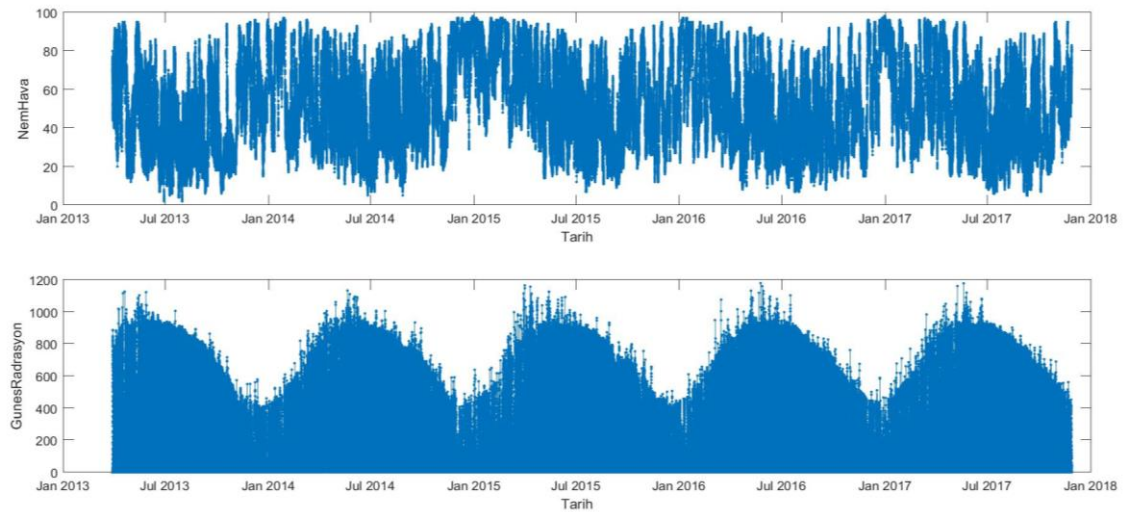
4.1.3 Gaziantep meteorolojik veri analizi

Gaziantep için sıcaklık dış ve güneş radyasyonu arasındaki ilişki Şekil 4.24’de görüldüğü gibi aralarında bir örüntü vardır. Eksik veri olmaması ve ayırık değerlerin olmaması kullanım olarak uygun olduğunu gösteriyor. 2013 yılında ölçüm hataları görülmektedir.



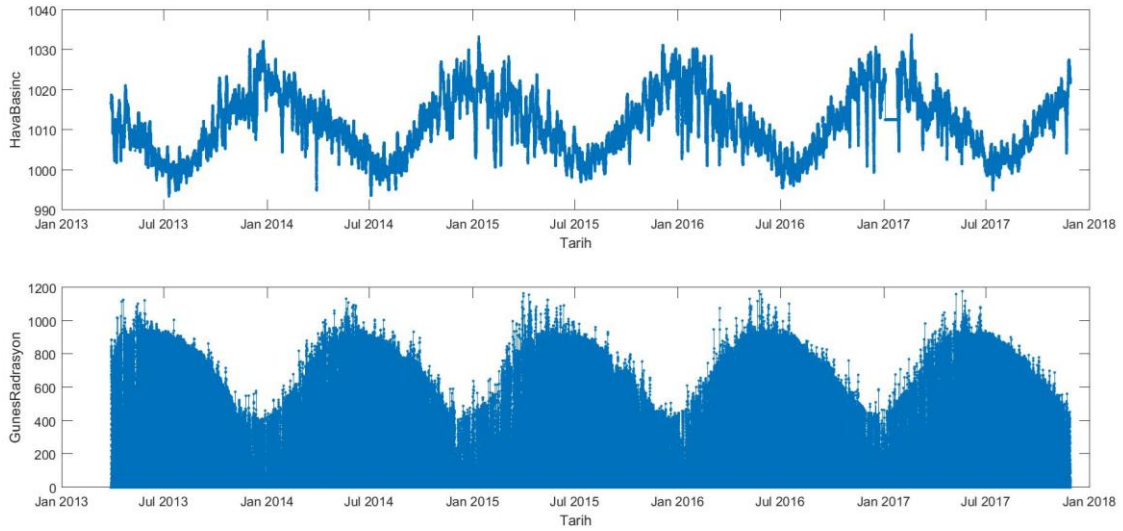
Şekil 4.24: Gaziantep sıcaklık dış ve güneş radyasyonu ilişkisi.

Gaziantep için nem hava ve güneş radyasyonu arasındaki ilişki Şekil 4.25’de görüldüğü gibi aralarında ters bir ilişki vardır. Eksik veri olmaması ve ayırık değerlerin olmaması kullanım olarak uygun olduğunu gösteriyor. Şanlıurfa’ya göre daha nemli bir bölge olduğu görülmektedir.



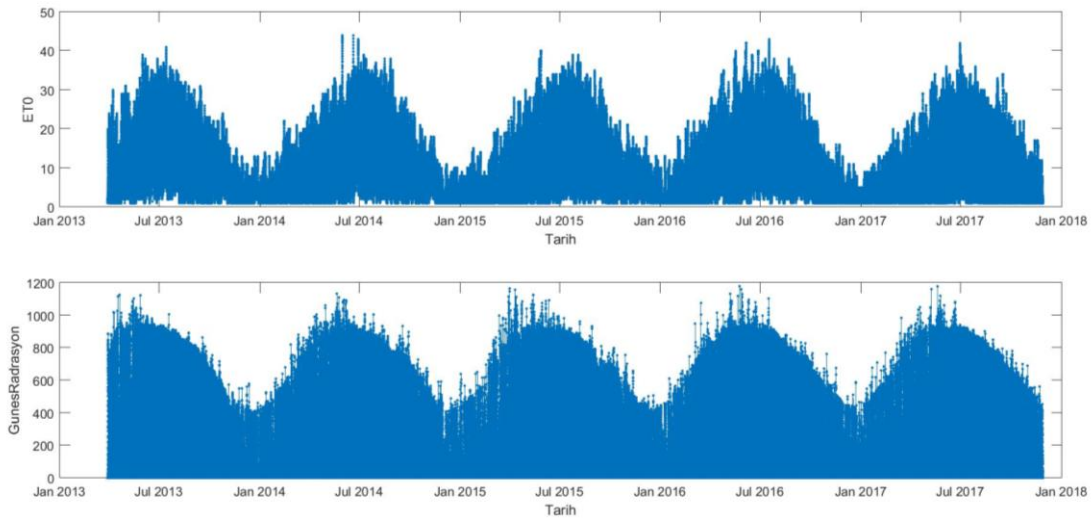
Şekil 4.25: Gaziantep nem hava ve güneş radyasyonu ilişkisi.

Gaziantep için hava basıncı ve güneş radyasyonu arasındaki ilişki Şekil 4.26’de görüldüğü gibi aralarında ters bir ilişki vardır. Eksik veri olmaması ve ayırık değerlerin olmaması kullanım olarak uygun olduğunu gösteriyor. 2017 yılında kısa süreli ölçüm hataları görülmektedir.



Şekil 4.26: Gaziantep hava basıncı ve güneş radyasyonu ilişkisi.

Gaziantep için ET0 ve güneş radyasyonu arasındaki ilişki Şekil 4.27’de görüldüğü gibi aralarında benzer bir ilişki vardır. Eksik veri olmaması ve ayırık değerlerin olmaması kullanım olarak uygun olduğunu gösteriyor.



Şekil 4.27: Gaziantep ET0 ve güneş radyasyonu ilişkisi

Çizelge 4.6’da Gaziantep için meteorolojik verilerin analizi yapılmıştır. Analiz sonucunda ayırık verilere rastlanmamıştır. Şanlıurfa bölgesine göre minimum sıcaklığın daha yüksek olduğu ve maksimum sıcaklığın daha düşük olduğu gözlemlenmiştir. Hava basıncı ve güneş radyasyonu olarak benzer özelliklere sahip olduğu görülmüştür. Nem hava bakımından Gaziantep’in, Şanlıurfa’dan görece daha nemli olduğu söylenebilir.

Çizelge 4.6: Gaziantep meteorolojik parametrelerin analizi.

Parametre	Minimum	Maksimum	Standart sapma	Ortalama
Sıcaklık	-6,11	40,33	9,12	18,22
Nem	2	98	22,35	52,7
ET0	1	44	8,59	8,42
Hava basıncı	993,39	1033,83	7,57	1011,45
Güneş radyasyonu	0	1179	276,35	195,5

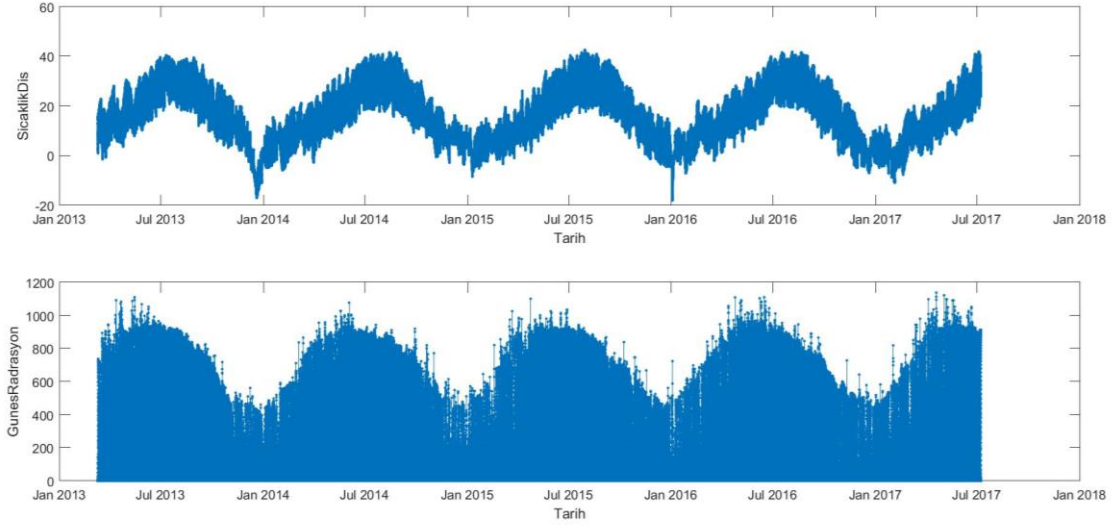
Gaziantep verilerine ADF testi uygulandığında 5 değişken değeri için p değeri 0,5’in altında olduğu ve zaman serisinin durağan olduğu analizi yapılabilir.

4.1.4 Diyarbakır meteorolojik veri analizi

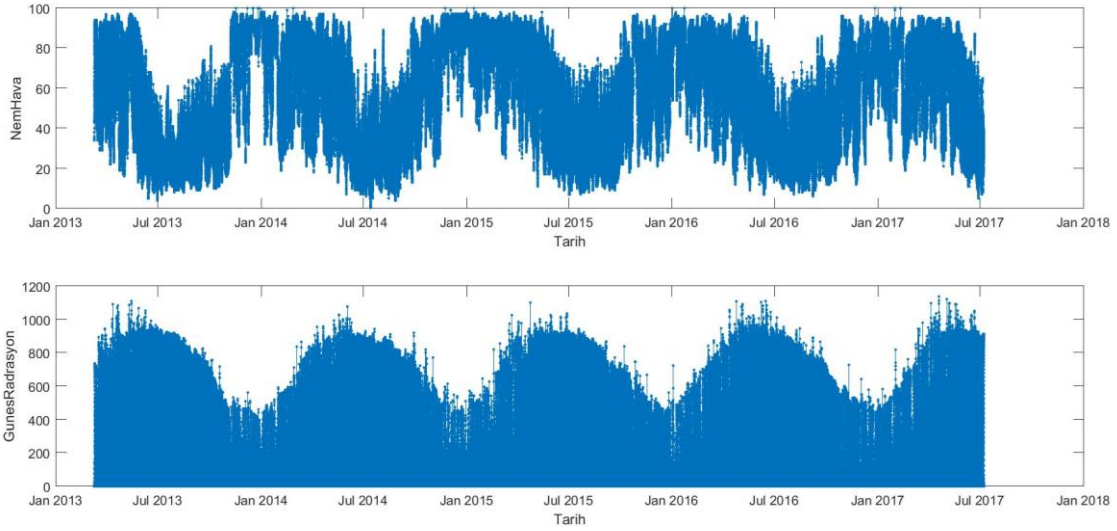
Diyarbakır için sıcaklık dış ve güneş radyasyonu arasındaki ilişki Şekil 4.28’de görüldüğü gibi aralarında bir örüntü vardır. Eksik veri olmaması ve ayırık değerlerin olmaması kullanım olarak uygun olduğunu gösteriyor.

Diyarbakır için nem hava ve güneş radyasyonu arasındaki ilişki Şekil 4.29’da görüldüğü gibi aralarında ters bir ilişki vardır. Eksik veri olmaması ve ayırık değerlerin olmaması kullanım olarak uygun olduğunu gösteriyor. Nem oranının diğer bölgelerden yüksek olduğu gözükmemektedir.

Diyarbakır için hava basıncı ve güneş radyasyonu arasındaki ilişki Şekil 4.30’da görüldüğü gibi aralarında ters bir ilişki vardır. Eksik veri olmaması ve ayırık değerlerin olmaması kullanım olarak uygun olduğunu gösteriyor.



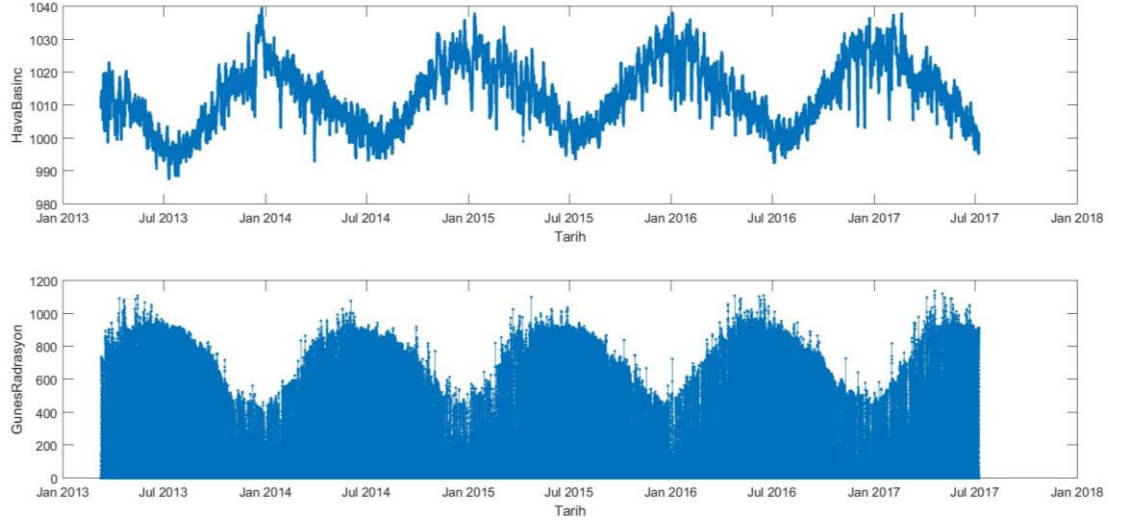
Şekil 4.28: Diyarbakır sıcaklık dış ve güneş radyasyonu ilişkisi.



Şekil 4.29: Diyarbakır nem hava ve güneş radyasyonu ilişkisi.

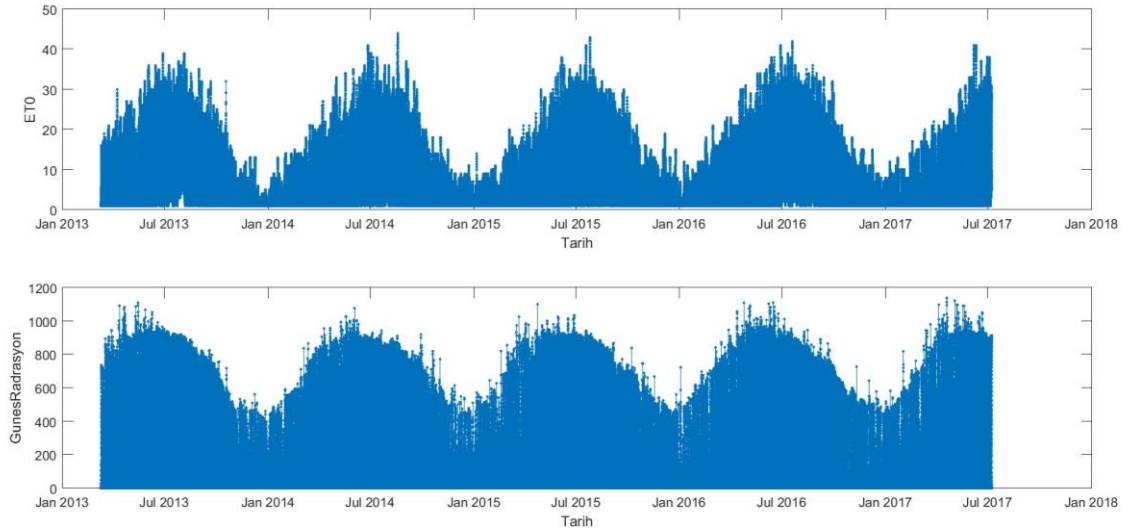
Diyarbakır için ET0 ve güneş radyasyonu arasındaki ilişki Şekil 4.31’de görüldüğü gibi aralarında benzer bir ilişki vardır. Eksik veri olmaması ve ayırık değerlerin olmaması kullanım olarak uygun olduğunu gösteriyor.

Çizelge 4.7’te Diyarbakır için meteorolojik verilerin analizi yapılmıştır. Analiz sonucunda ayırık verilere rastlanmamıştır.



Şekil 4.30: Diyarbakır hava basıncı ve güneş radyasyonu ilişkisi.

Şanlıurfa bölgesine göre minimum sıcaklığın daha düşük olduğu ve maksimum sıcaklığın da daha düşük olduğu gözlemlenmiştir ve Diyarbakır bölgesinde soğukların daha etkili olduğu söylenebilir.



Şekil 4.31: Diyarbakır ET0 ve güneş radyasyonu ilişkisi.

Güneş radyasyonu ve ET0 olarak Şanlıurfa bölgesine göre daha düşük olduğu görülmektedir. Nem hava bakımından Diyarbakır'ın, Şanlıurfa'dan görece daha nemli olduğu söylenebilir.

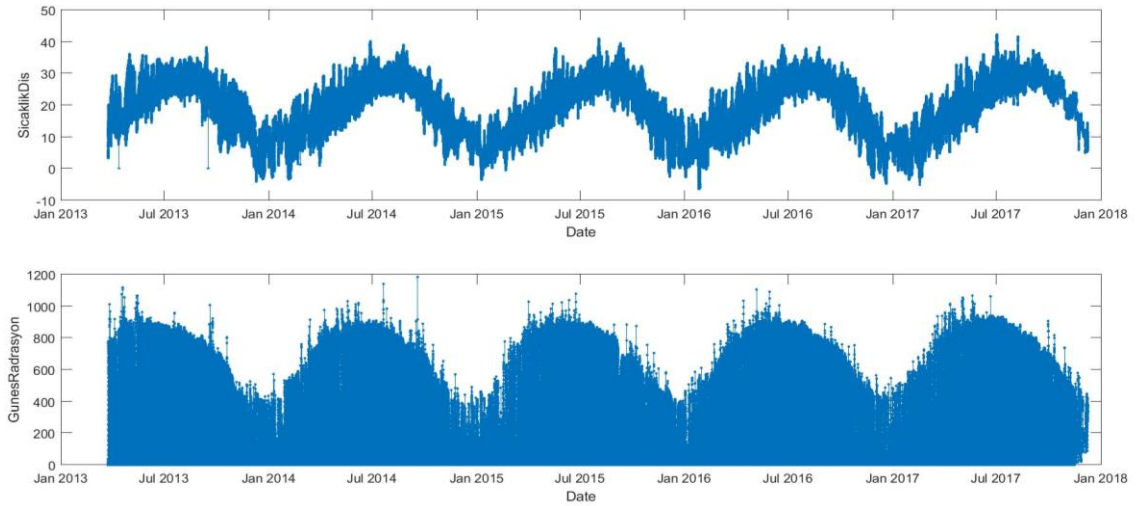
Çizelge 4.7: Diyarbakır meteorolojik parametrelerin analizi.

Parametre	Minimum	Maksimum	Standart sapma	Ortalama
Sıcaklık	-18,05	42,6	11,25	16,36
Nem	0,26	100	26,03	56,03
ET0	1	44	8,14	7
Hava basıncı	987,60	1039,51	9,66	1013,05
Güneş radyasyonu	0	1138	274,75	190,96

Diyarbakır verilerine ADF testi uygulandığında 5 değişken değeri için p değeri 0,5'in altında olduğu ve zaman serisinin durağan olduğu analizi yapılabilir.

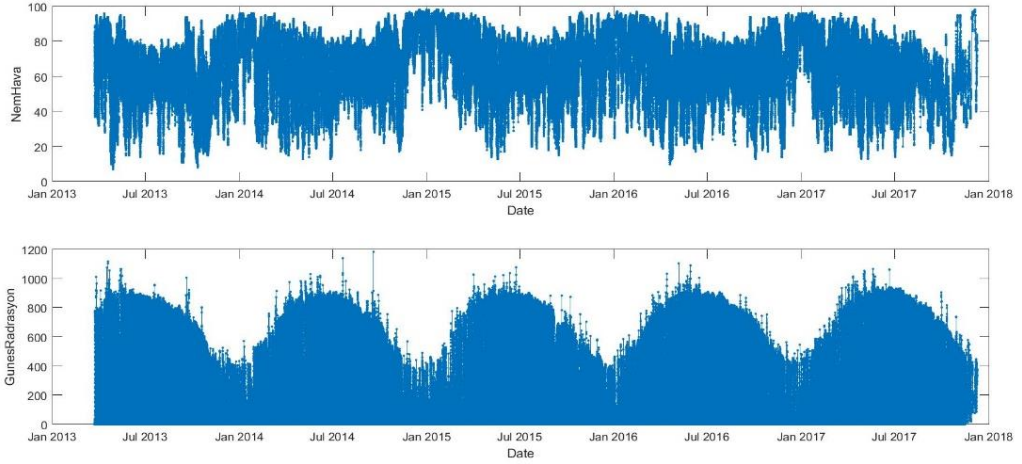
4.1.5 Hatay meteorolojik veri analizi

Hatay için sıcaklık dış ve güneş radyasyonu arasındaki ilişki Şekil 4.32'de görüldüğü gibi aralarında bir örüntü vardır.



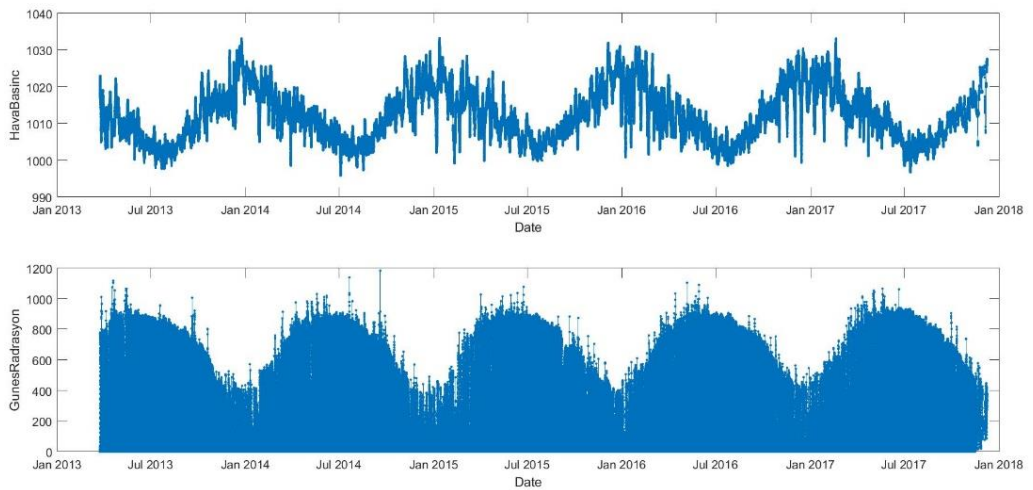
Şekil 4.32: Hatay sıcaklık dış ve güneş radyasyonu ilişkisi.

Hatay için nem hava ve güneş radyasyonu arasındaki ilişki Şekil 4.33’de görüldüğü gibi aralarında ters bir ilişki vardır. Eksik veri olmaması ve ayırık değerlerin olmaması kullanım olarak uygun olduğunu gösteriyor. Nem hava değerinin oldukça gürültülü olduğu görünüyor ancak bölgenin nem durumunun değişimini Akdeniz iklimi etkisi ile açıklayabiliriz.



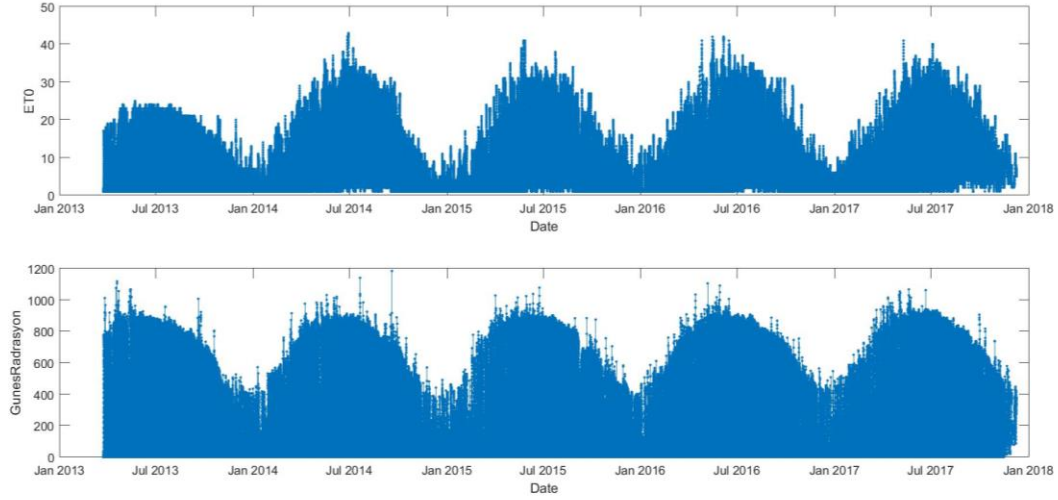
Şekil 4.33: Hatay nem hava ve güneş radyasyonu ilişkisi.

Hatay için hava basıncı ve güneş radyasyonu arasındaki ilişki Şekil 4.34’de görüldüğü gibi aralarında ters bir ilişki vardır. Eksik veri olmaması ve ayırık değerlerin olmaması kullanım olarak uygun olduğunu gösteriyor. Nem hava ile ilişkili olarak gürültülü olarak verilerin iklim ile alakalı değişimini görebiliyoruz.



Şekil 4.34: Hatay hava basıncı ve güneş radyasyonu ilişkisi.

Hatay için ET0 ve güneş radyasyonu arasındaki ilişki Şekil 4.35’de görüldüğü gibi aralarında benzer bir ilişki vardır. Eksik veri olmaması ve ayrık değerlerin olmaması kullanım olarak uygun olduğunu gösteriyor.



Şekil 4.35: Hatay için ET0 ve güneş radyasyonu ilişkisi.

Çizelge 4.8’te Hatay için meteorolojik verilerin analizi yapılmıştır. Analiz sonucunda ayrık verilere rastlanmamıştır. Sıcaklık dışı değeri olarak Şanlıurfa’ya yakın değerlere sahip bir bölge olduğu görülmüyor. Nem hava verisi olarak oldukça nemli bir bölge olduğu görülmüyor. Güneş radyasyonu ve ET0 olarak Şanlıurfa bölgesine yakın değerler olduğu görülmüyor.

Hatay verilerine ADF testi uygulandığında 5 değişken değeri için p değeri 0,5’in altında olduğu ve zaman serisinin durağan olduğu analizi yapılabilir.

Çizelge 4.8: Hatay için meteorolojik parametrelerin analizi.

Parametre	Minimum	Maksimum	Standart sapma	Ortalama
Sıcaklık dışı	-6,38	42,27	8,57	19
Nem hava	7	98	17,99	64,97
ET0	1	43	8,38	7,4
Hava basıncı	995,86	1033,28	6,94	1012,27
Güneş radyasyonu	0	1184,06	274,89	197,89

4.2 Sistem Tanılama Modelleri

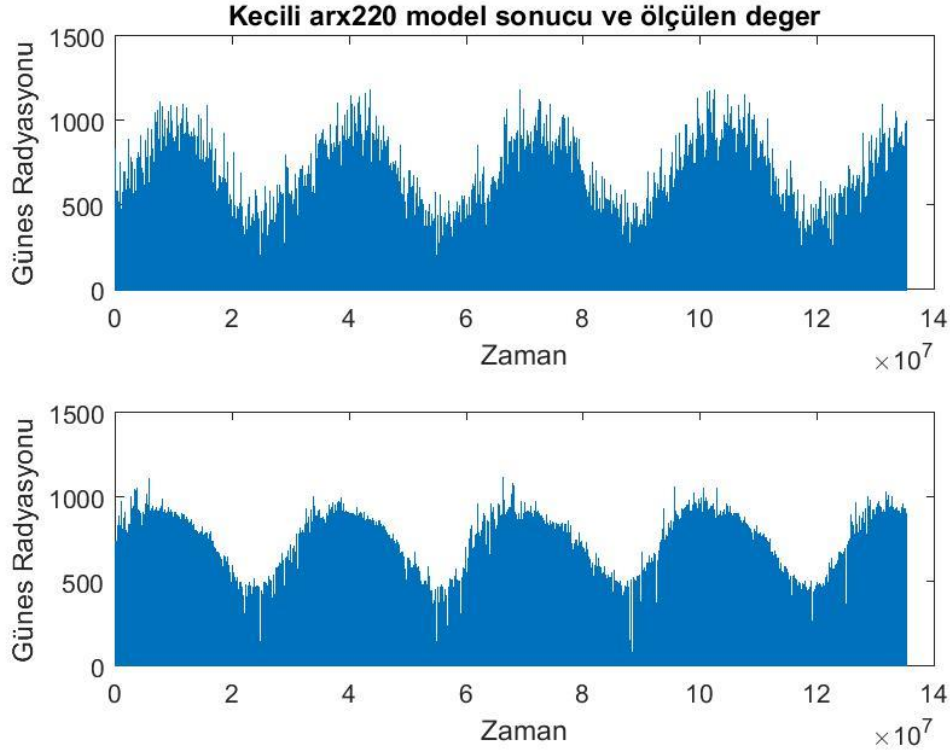
4.2.1 Doğrusal modeller

4.2.1.1 Dışsal kaynaklı otoregresif modeli

Veri hazırlama bölümünde bahsedilen eğitim ve test verileri kullanılarak dışsal kaynaklı otoregresif modeller MATLAB'in System Identification Toolbox'ı kullanılarak oluşturulmuştur. Deneme yanılma yöntemi ile 10 farklı model oluşturulmuş ve bunlar diğer illerin verileri ile beslenerek doğrulama yapılmıştır. Doğrulama için her modele belli bir puan verilerek en son olarak hangi modelin 5 il için daha uygun olduğuna karar verildi. Model isimlendirmesi mertebeler kullanılmaktadır ve örnek olarak 'arx220' ismi n_a kutup sayısını 2, n_b sıfır sayısını 2, n_k sistemdeki ölü zamanın 0 olduğunu gösterir. Şanlıurfa için oluşturulan modellere Çizelge 4.9'dan incelendiğinde arx220 diğerlerinden daha iyi sonuç verdiği gözlemlenmiştir. arx120 modelinde çok yakın sonuçlar elde ettiği gözlenmektedir. Şekil 4.36'de Şanlıurfa verisi ile arx220 model sonucu ve ölçülen değer arasındaki farkı görebiliriz.

Çizelge 4.9: Şanlıurfa dışsal kaynaklı otoregresif model sonuçları.

Model	R^2	RMSE	nRMSE	MAE	Puan
arx220	0.91314	82.9718	0.074481	45.5902	10
arx120	0.91042	84.2627	0.07564	45.8121	9
arx10102	0.87486	99.5925	0.089401	56.2126	8
arx223	0.8516	108.4516	0.097353	58.8837	7
arx762	0.84772	109.8602	0.098618	62.8265	6
arx222	0.8392	112.8934	0.10134	65.744	5
arx791	0.8328	115.1181	0.10334	62.5963	4
arx221	0.79502	127.4618	0.11442	69.9324	3
arx933	0.74017	143.5047	0.12882	83.3316	2
arx224	0.72537	147.5354	0.13244	86.6864	1



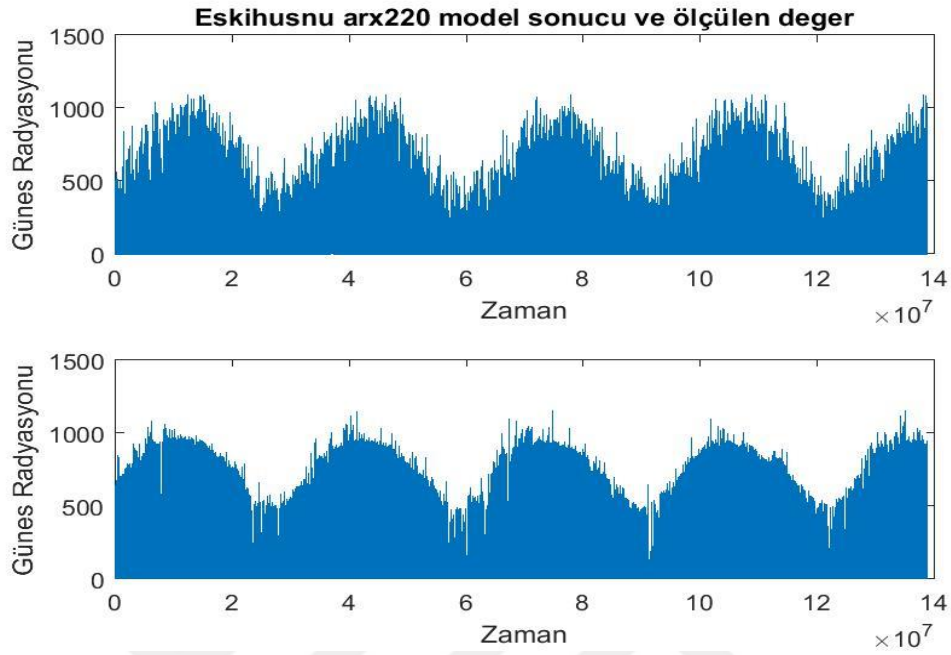
Şekil 4.36: Şanlıurfa arx220 model sonucu.

Oluşturulan dışsal kaynaklı otoregresif tahmin modeli Adıyaman verilerine uygulandığında Çizelge 4.10'dan görüldüğü gibi arx220 modelinin en iyi sonucu verdiği ve arx120 modelinin en iyiye yakın sonuçlar verdiği görülmektedir.

Çizelge 4.10: Adıyaman dışsal kaynaklı otoregresif model sonuçları.

Model	R^2	RMSE	nRMSE	MAE	Puan
arx220	0.8474	113.2648	0.098235	61.6989	10
arx120	0.8472	113.3394	0.0983	60.778	9
arx223	0.84662	113.5536	0.098485	61.8705	8
arx791	0.80876	126.7953	0.10997	69.3771	7
arx221	0.78108	134.3714	0.11654	73.917	6
arx10102	0.75666	143.0291	0.12405	78.3378	5
arx222	0.72313	152.5668	0.13232	83.4594	4
arx762	0.71617	154.471	0.13397	83.6024	3
arx933	0.56689	190.8173	0.1655	101.3227	2
arx224	0.56491	191.2532	0.16587	103.0826	1

Şekil 4.37’de Adıyaman verisi ile arx220 model sonuçları ve ölçülen değerler arasındaki ilişkiyi görebiliriz.



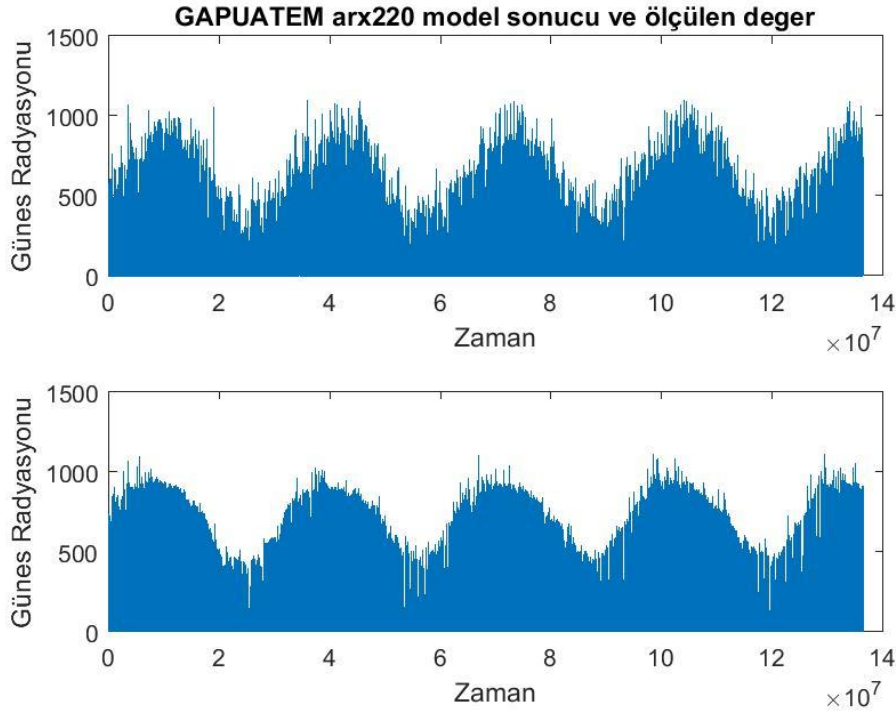
Şekil 4.37: Adıyaman arx220 model sonucu.

Oluşturulan model Diyarbakır verilerine uygulandığında Çizelge 4.11’de görüldüğü gibi arx220 modelinin en iyi sonucu verdiği ve arx120 ve arx224 modelinin en iyiye yakın sonuçlar verdiği görülmektedir.

Çizelge 4.11: Diyarbakır dışsal kaynaklı otoregresif model sonuçları.

Model	R^2	RMSE	nRMSE	MAE	Puan
arx220	0,85313	105,3063	0,094615	56,843	10
arx120	0,85284	105,4112	0,094709	56,3852	9
arx223	0,85054	106,2314	0,095446	57,2632	8
arx791	0,80641	120,9015	0,10863	65,7061	7
arx221	0,78033	128,3518	0,11532	70,3014	6
arx10102	0,74349	139,1709	0,12504	75,5382	5
arx222	0,72099	145,1448	0,13041	79,8264	4
arx762	0,71482	146,7419	0,13184	80,366	3
arx224	0,57092	179,9974	0,16172	98,0591	2
arx933	0,55501	183,3036	0,16469	99,1435	1

Şekil 4.38’de Diyarbakır verisi ile arx220 model sonuçlarını ve gerçek değerler arasındaki ilişkiyi görebiliriz.



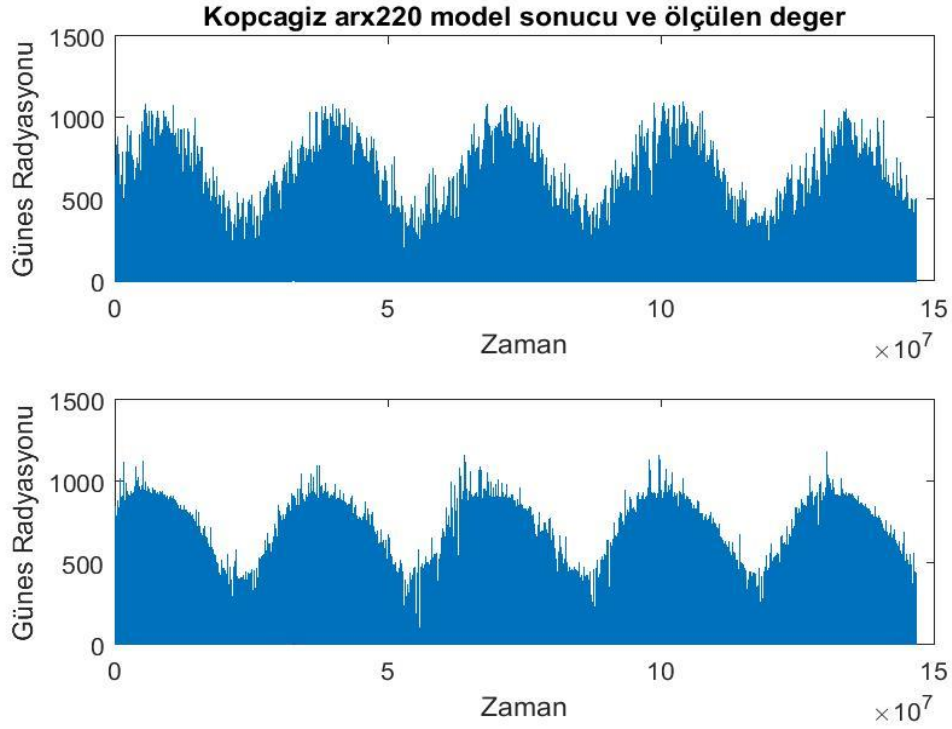
Şekil 4.38: Diyarbakır arx220 model sonucu.

Oluşturulan model Gaziantep verilerine uygulandığında Çizelge 4.12’den görüldüğü gibi arx220 modelinin en iyi sonucu verdiği ve arx120 ve arx224 modelinin en iyiye yakın sonuçlar verdiği görülmektedir.

Çizelge 4.12: Gaziantep dışsal kaynaklı otoregresif model sonuçları.

Model	R ²	RMSE	nRMSE	MAE	Puan
arx220	0,84485	108,9257	0,092467	59,4418	10
arx120	0,8428	109,646	0,093078	59,5342	9
arx223	0,84146	110,1097	0,093472	59,9066	8
arx791	0,8115	120,0644	0,10192	65,4989	7
arx221	0,7898	126,7862	0,10763	69,6947	6
arx10102	0,75703	136,3132	0,11572	74,6496	5
arx222	0,72654	144,612	0,12276	79,9564	4
arx762	0,71324	148,0885	0,12571	80,6892	3
arx224	0,56325	182,7588	0,15514	99,0576	2
arx933	0,54622	186,2859	0,15814	99,8795	1

Şekil 4.39’da Gaziantep verisi ile arx220 model sonuçlarını ve gerçek değerler arasındaki ilişkiyi görebiliriz.



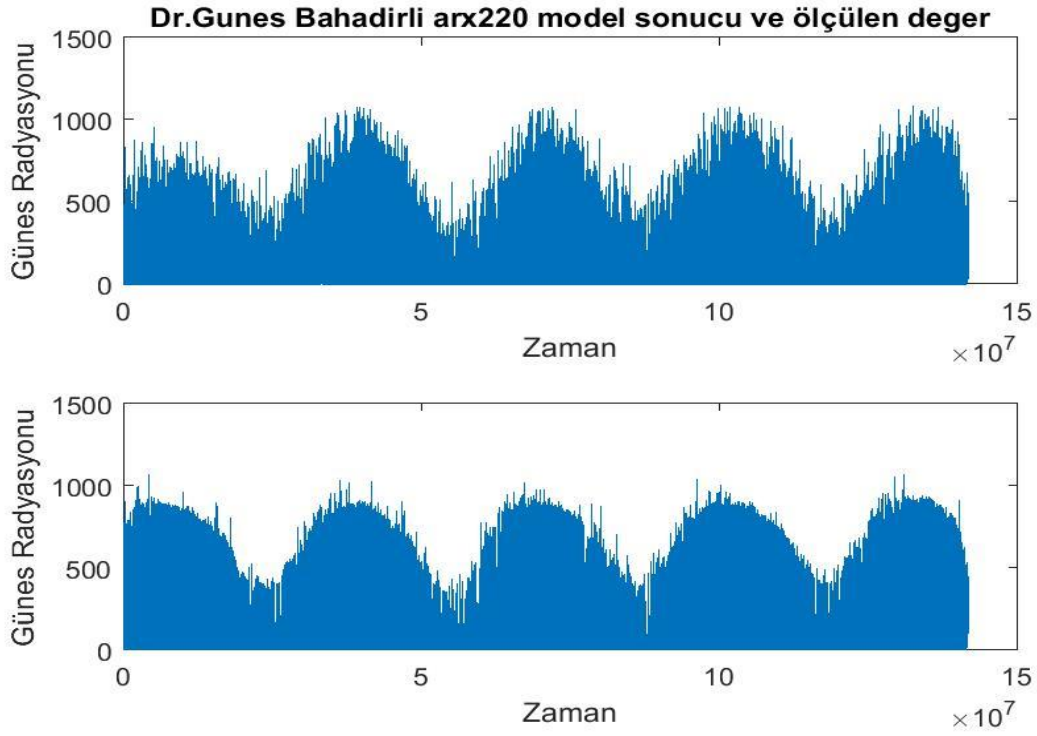
Şekil 4.39: Gaziantep arx220 model sonucu.

Oluşturulan model Hatay verilerine uygulandığında Çizelge 4.13’den görüldüğü gibi arx220 modelinin en iyi sonucu verdiği ve arx223 ve arx120 modelinin en iyiye yakın sonuçlar verdiği görülmektedir.

Çizelge 4.13: Hatay dışsal kaynaklı otoregresif model sonuçları.

Model	R^2	RMSE	nRMSE	MAE	Puan
arx220	0,85762	103,695	0,097184	56,8412	10
arx223	0,85543	104,4874	0,097926	57,2241	9
arx120	0,8527	105,4702	0,098847	57,5309	8
arx791	0,82478	115,0339	0,10781	63,9756	7
arx221	0,80094	122,6083	0,11491	67,8816	6
arx10102	0,765	133,219	0,12485	73,9551	5
arx222	0,74107	139,8378	0,13106	78,0345	4
arx762	0,73391	141,7585	0,13286	79,113	3
arx933	0,58563	176,8985	0,16579	98,5705	2
arx224	0,58536	176,9572	0,16585	98,2106	1

Şekil 4.40'ta Hatay verisi ile arx220 model sonuçlarını ve gerçek değerler arasındaki ilişkiyi görebiliriz.



Şekil 4.40: Hatay arx220 model sonucu

Beş ölçüm merkezi için en iyi modeli bulmak için verilen puanlar ile bir sıralama yapıldığında Çizelge 4.14'de görüldüğü gibi arx220 bütün bölgeler için iyi sonuçlar vermektedir.

Çizelge 4.14: Beş ölçüm merkezi için en iyi DKDO seçimi.

Model	Puan
arx220	50
arx120	44
arx223	40
arx791	32
arx10102	28
arx221	27
arx222	21
arx762	18
arx933	8
arx224	7

4.2.1.2 Box – Jenkins modeli

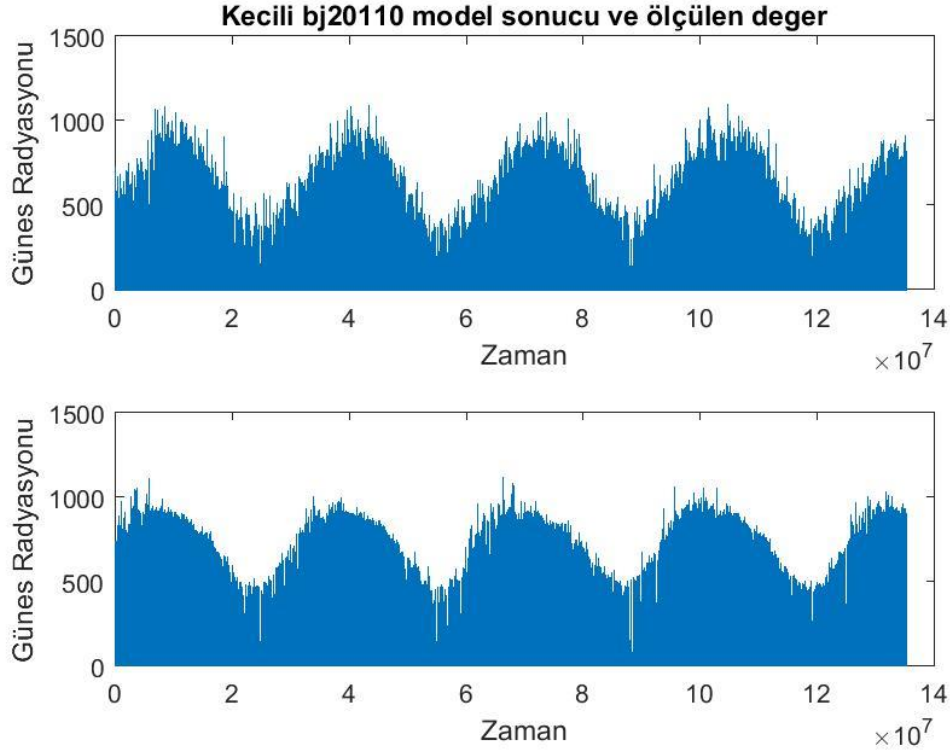
Veri hazırlama bölümünde bahsedilen eğitim ve test verileri kullanılarak Box-Jenkins modeller MATLAB'ın System Identification Toolbox'ı kullanılarak oluşturulmuştur. Deneme yanılma yöntemi ile 10 farklı model oluşturulmuş ve bunlar diğer illerin verileri ile beslenerek doğrulama yapılmıştır. Doğrulama için her modele belli bir puan verilerek en son olarak hangi modelin beş il için daha uygun olduğuna karar verilmiştir. Model isimlendirmesi mertebeler kullanılmaktadır ve örnek olarak 'bj20110' ismi n_b, n_c, n_d, n_f ve n_k değerlerini göstermektedir.

Şanlıurfa için oluşturulan modellere Çizelge 4.15'den görüldüğü gibi bj20110 diğerlerinden daha iyi sonuç verdiği gözlemlenmiştir. Bj20310 modelinde çok yakın sonuçlar elde ettiği gözlenmektedir.

Çizelge 4.15: Şanlıurfa Box-Jenkins model sonuçları.

Model	R ²	RMSE	nRMSE	MAE	Puan
bj20110	0,92711	76,0172	0,068238	40,8621	10
bj20310	0,92071	79,2816	0,071168	43,3993	9
bj20520	0,91923	80,0184	0,07183	44,1716	8
bj21510	0,91777	80,7366	0,072475	43,903	7
bj20510	0,91743	80,9038	0,072625	44,058	6
bj22220	0,91652	81,3475	0,073023	44,2597	5
bj221020	0,91483	82,1675	0,073759	44,9842	4
bj22510	0,91016	84,3925	0,075756	46,687	3
bj22520	0,90835	85,2383	0,076515	47,0329	2
bj225202	0,90708	85,8263	0,077043	47,6282	1

Şekil 4.41'de Şanlıurfa verisi ile bj20110 model sonucu ve ölçülen değer arasındaki farkı görebiliriz.

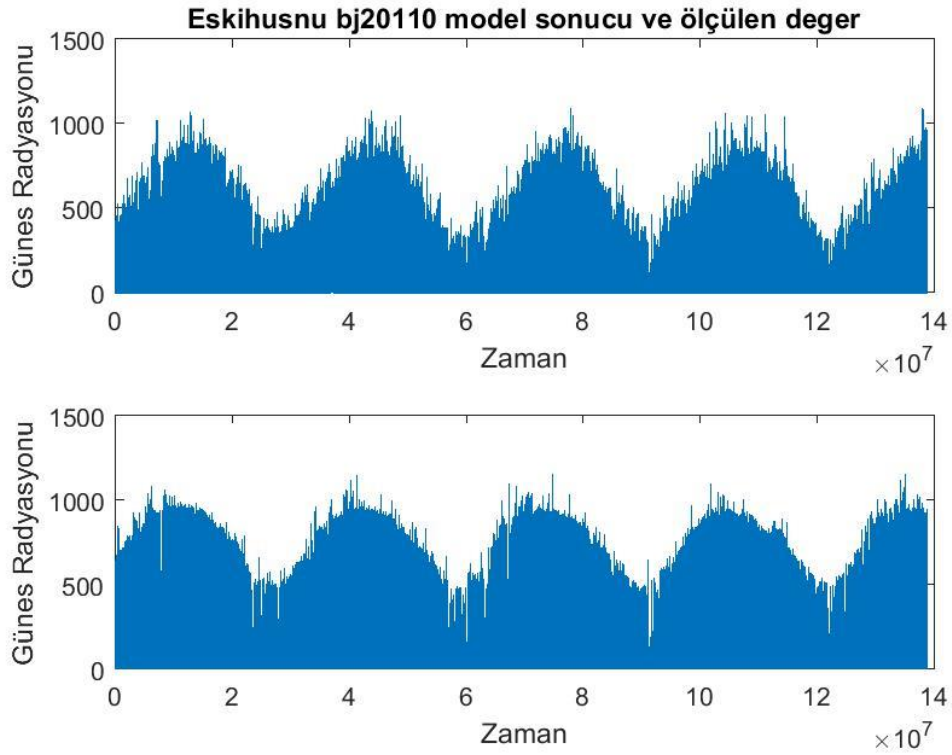


Şekil 4.41: Şanlıurfa bj20110 model sonucu

Oluşturulan Box-Jenkins tahmin modeli Adıyaman verilerine uygulandığında Çizelge 4.16'dan görüldüğü gibi bj20110 modelinin en iyi sonucu verdiği ve bj20310 ile beraber bj21510 modelinin en iyiye yakın sonuçlar verdiği görülmektedir. Şekil 4.42'de Adıyaman verisi ile bj20310 model sonuçları ve ölçülen değerler arasındaki ilişkiyi görebiliriz.

Çizelge 4.16: Adıyaman Box-Jenkins model sonuçları

Model	R ²	RMSE	nRMSE	MAE	Puan
bj20110	0,91154	86,2364	0,074793	47,9061	10
bj20310	0,90242	90,5728	0,078554	51,0461	9
bj21510	0,9012	91,1361	0,079043	51,1275	8
bj20520	0,90072	91,3577	0,079235	51,7856	7
bj22220	0,90048	91,4695	0,079332	51,3197	6
bj20510	0,90001	91,6865	0,07952	51,4601	5
bj221020	0,89312	94,7893	0,082211	53,888	4
bj22520	0,88886	96,6636	0,083837	54,6844	3
bj22510	0,88879	96,6935	0,083863	54,9015	2
bj225202	0,8841	98,7105	0,085612	55,996	1



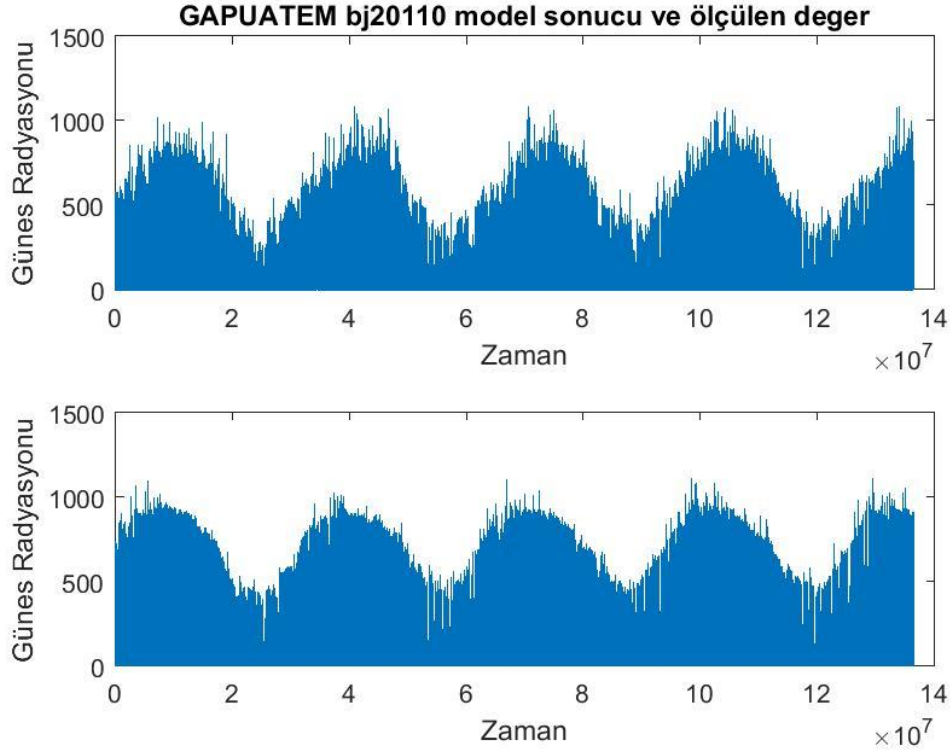
Şekil 4.42: Adıyaman bj20110 model sonucu.

Oluşturulan Box-Jenkins tahmin modeli Diyarbakır verilerine uygulandığında Çizelge 4.17’den görüldüğü gibi bj20110 modelinin en iyi sonucu verdiği ve bj20310 ile beraber bj21520 modelinin en iyiye yakın sonuçlar verdiği görülmektedir.

Çizelge 4.17: Diyarbakır Box-Jenkins model sonuçları.

Model	R ²	RMSE	nRMSE	MAE	Puan
bj20110	0,92292	76,2904	0,068545	41,0225	10
bj20310	0,91412	80,529	0,072353	44,0774	9
bj20520	0,91222	81,4106	0,073145	44,9818	8
bj21510	0,91091	82,0179	0,073691	44,4346	7
bj20510	0,91019	82,3508	0,07399	44,6971	6
bj221020	0,90966	82,5914	0,074206	45,2533	5
bj22220	0,90943	82,6983	0,074302	44,8295	4
bj22510	0,90159	86,2031	0,077451	47,689	3
bj22520	0,90036	86,7403	0,077934	47,7307	2
bj225202	0,89935	87,1765	0,078326	48,5286	1

Şekil 4.43’te Diyarbakır verisi ile bj20110 model sonuçları ve ölçülen değerler arasındaki ilişkiyi görebiliriz.



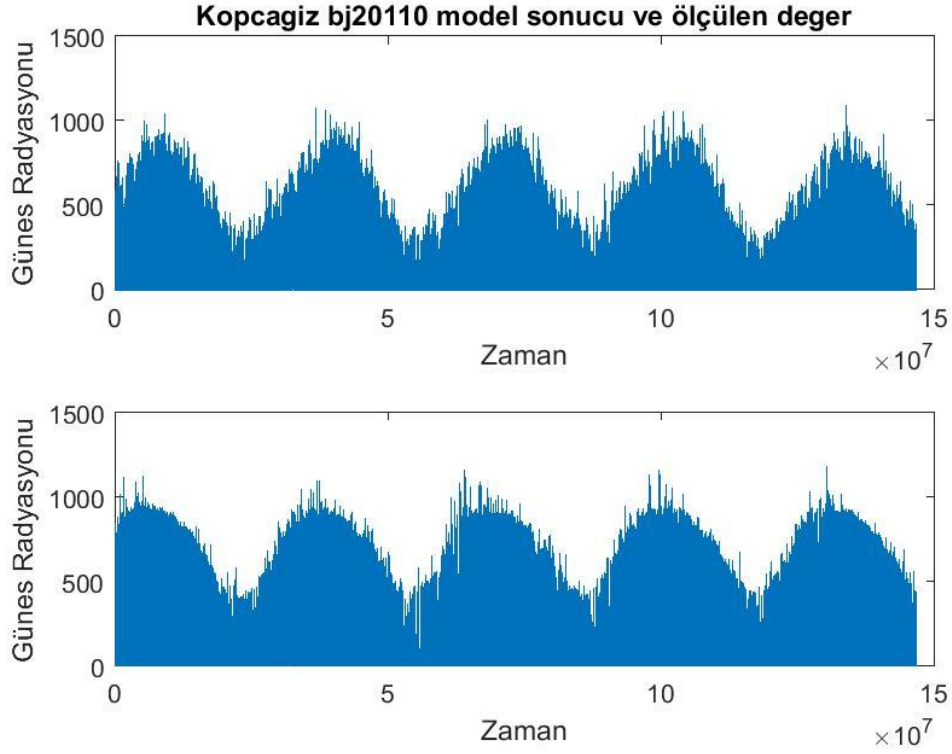
Şekil 4.43: Diyarbakır bj20110 model sonucu.

Oluşturulan Box-Jenkins tahmin modeli Gaziantep verilerine uygulandığında Çizelge 4.18’den görüldüğü gibi bj20110 modelinin en iyi sonucu verdiği ve bj20310 ile beraber bj21520 modelinin en iyiye yakın sonuçlar verdiği görülmektedir.

Çizelge 4.18: Gaziantep Box-Jenkins model sonuçları.

Model	R ²	RMSE	nRMSE	MAE	Puan
bj20110	0,92322	76,6263	0,065048	41,093	10
bj20310	0,91678	79,7755	0,067721	43,5467	9
bj21510	0,91536	80,4533	0,068297	43,7468	8
bj20520	0,91536	80,4548	0,068298	44,2321	7
bj20510	0,91503	80,6088	0,068429	43,844	6
bj22220	0,91498	80,6346	0,06845	43,8845	5
bj221020	0,91442	80,8995	0,068675	44,4356	4
bj22510	0,90597	84,7979	0,071985	46,88	3
bj22520	0,90588	84,8394	0,07202	46,7899	2
bj225202	0,9041	85,6379	0,072698	47,3408	1

Şekil 4.44’te Gaziantep verisi ile bj20110 model sonuçları ve ölçülen değerler arasındaki ilişkiyi görebiliriz.



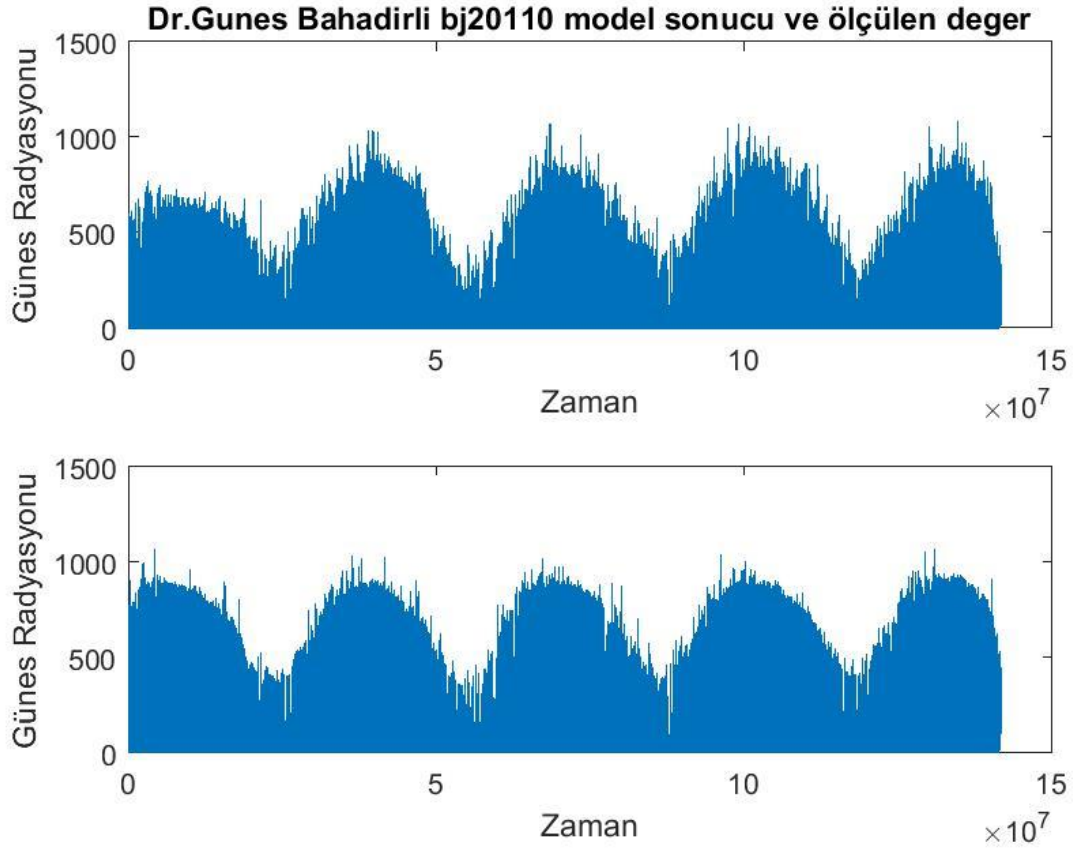
Şekil 4.44: Gaziantep bj20110 model sonucu.

Oluşturulan Box-Jenkins tahmin modeli Gaziantep verilerine uygulandığında Çizelge 4.19’den görüldüğü gibi bj20110 modelinin en iyi sonucu verdiği ve bj20520 ile beraber bj20310 modelinin en iyiye yakın sonuçlar verdiği görülmektedir.

Çizelge 4.19: Hatay Box-Jenkins model sonuçları.

Model	R^2	RMSE	nRMSE	MAE	Puan
bj20110	0,9306	72,3972	0,067851	39,0273	10
bj20520	0,92394	75,7882	0,071029	42,1402	9
bj20310	0,92289	76,3106	0,071519	42,0754	8
bj221020	0,91783	78,7744	0,073828	43,6222	7
bj20510	0,91617	79,5652	0,074569	44,0914	6
bj21510	0,91579	79,7454	0,074738	44,2706	5
bj22220	0,91501	80,1163	0,075086	44,4149	4
bj225202	0,91221	81,4222	0,076309	45,6743	3
bj22510	0,91009	82,4011	0,077227	46,1684	2
bj22520	0,91003	82,4308	0,077255	46,0284	1

Şekil 4.45'te Gaziantep verisi ile bj20110 model sonuçları ve ölçülen değerler arasındaki ilişkiyi görebiliriz.



Şekil 4.45: Hatay bj20110 model sonucu.

5 bölge için en iyi modeli bulmak için verilen puanlar ile bir sıralama yapıldığında Çizelge 4.20'de görüldüğü gibi bj20110 bütün bölgeler için iyi sonuçlar vermektedir. Modellere bakıldığı zaman bj20310, bj20520 ve bj21510 modellerinin de diğer 5 il için Box-Jenkins doğrulama sonuçlarının oldukça yakın olduğunu görmekteyiz. bj20310 modeli, 5 il için de en iyi doğrulama sonucuna sahip bj20110 modeline en yakın doğrulama sonucuna sahip olduğu görülmektedir.

Çizelge 4.20: Beş bölge için en iyi Box-Jenkins model seçimi.

Model	Puan
bj20110	50
bj20310	44
bj20520	39
bj21510	35
bj20510	29
bj22220	24
bj221020	24
bj22510	13
bj22520	10
bj225202	7

4.2.2 Doğrusal olmayan modeller

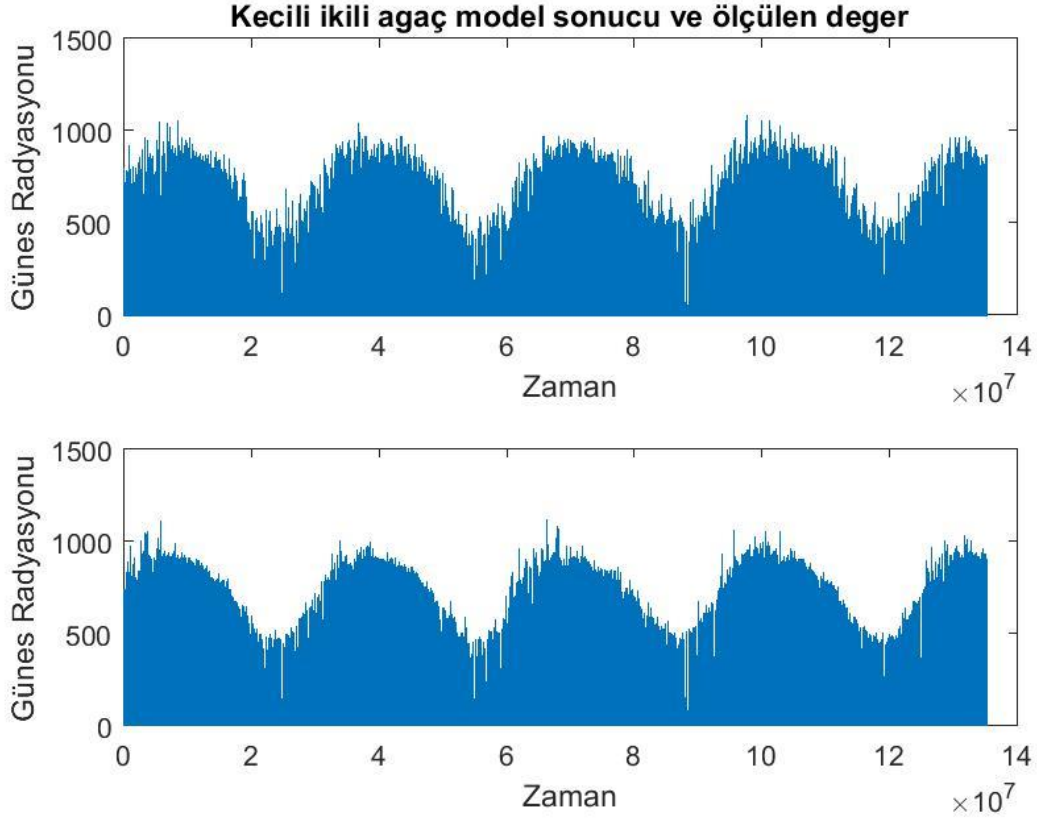
4.2.2.1 Dışsal kaynaklı otoregresif modeli

Veri hazırlama bölümünde bahsedilen eğitim ve test verileri kullanılarak doğrusal olmayan dışsal kaynaklı otoregresif modeller MATLAB'in System Identification Toolbox'ı kullanılarak oluşturulmuştur. Dalgacık, ikili ağaç ve sigmoid aktivasyon fonksiyonu kullanılarak 3 farklı model oluşturulmuş ve bunlar diğer illerin verileri ile beslenerek valide edilmiştir. Doğrulama için her modele belli bir puan verilerek en son olarak hangi modelin 5 ölçüm merkezi için daha uygun olduğuna karar verilmiştir.

Çizelge 4.21: Şanlıurfa DKDOO model sonuçları.

Model	R ²	RMSE	nRMSE	MAE	Puan
İkili ağaç	0,95538	59,4713	0,053385	31,1725	3
Sigmoid	0,937	70,6645	0,063433	37,2337	2
Dalgacık	0,93032	74,312	0,066707	39,464	1

Şanlıurfa için oluşturulan modellere bakıldığında İkili ağaç aktivasyon fonksiyonu ile oluşturulan model Çizelge 4.21'de gösterildiği gibi diğerlerinden daha iyi sonuç verdiği gözlemlenmiştir. Şekil 4.46'da Şanlıurfa verisi ile ikili ağaç aktivasyon fonksiyonu modeli sonucu ve ölçülen değer arasındaki farkı görebiliriz.



Şekil 4.46: Şanlıurfa ikili ağaç model sonucu.

Oluşturulan tahmin modeli Adıyaman verilerine uygulandığında Çizelge 4.22’de gösterildiği gibi sigmoid aktivasyon fonksiyonu ile en iyi sonucun alındığı görülüyor.

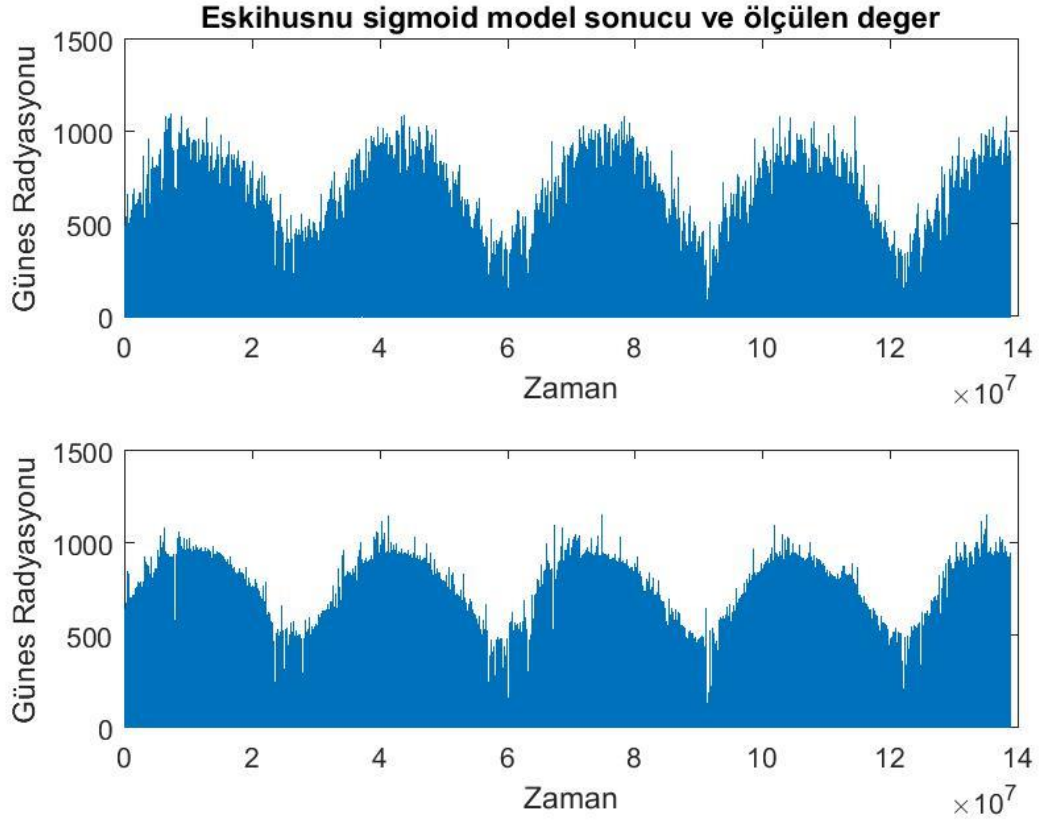
Çizelge 4.22: Adıyaman DKDOO model sonuçları.

Model	R^2	RMSE	nRMSE	MAE	Puan
Sigmoid	0,92414	79,8586	0,069262	42,7484	3
Dalgacık	0,91932	82,3592	0,07143	44,5887	2
İkili ağaç	0,918	83,0262	0,072009	42,8858	1

Şekil 4.47’de Adıyaman verisi ile sigmoid aktivasyon fonksiyonu modeli sonucu ve ölçülen değer arasındaki farkı görebiliriz.

Oluşturulan tahmin modeli Diyarbakır verilerine uygulandığında Çizelge 4.23’de gösterildiği gibi ikili ağaç aktivasyon fonksiyonu ile en iyi sonucun alındığı görülüyor ancak 3 aktivasyon fonksiyonunda yakın değerler aldığı görülmektedir. Şekil 4.48’de

Diyarbakır verisi ile ikili ağaç aktivasyon fonksiyonu modeli sonucu ve ölçülen değer arasındaki farkı görebiliriz.

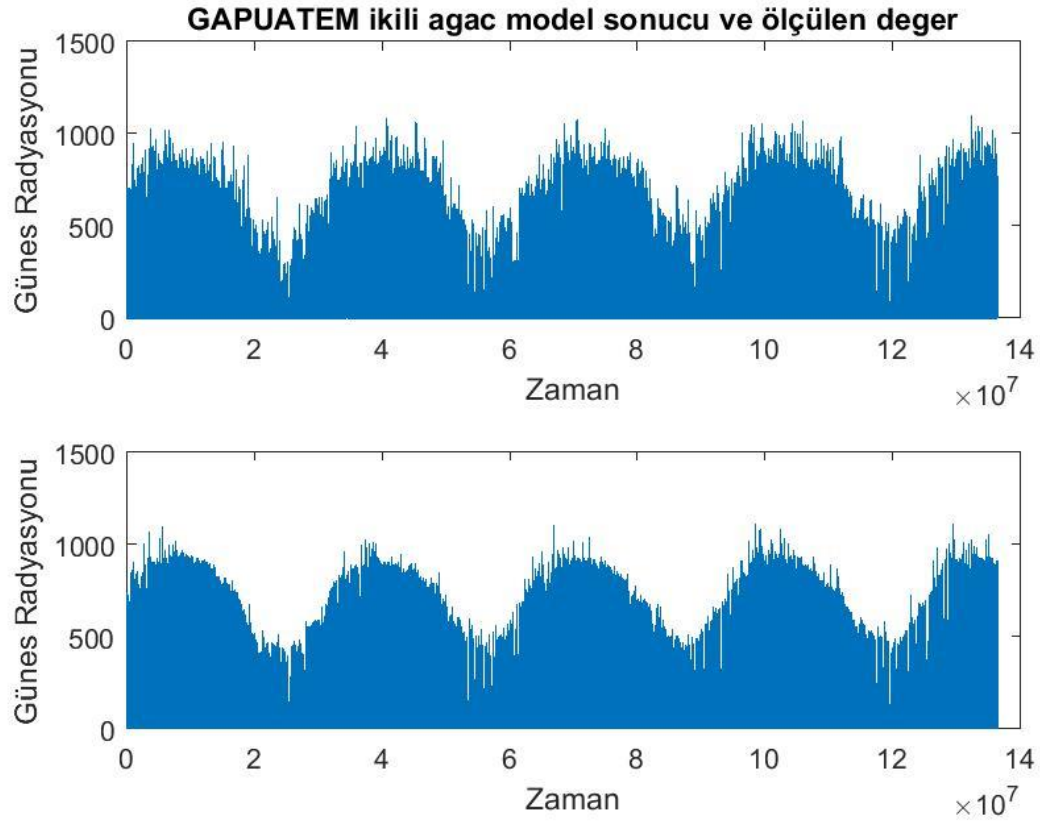


Şekil 4.47: Adıyaman sigmoid model sonucu.

Oluşturulan tahmin modeli Gaziantep verilerine uygulandığında Çizelge 4.24’de gösterildiği gibi sigmoid aktivasyon fonksiyonu ile en iyi sonucun alındığı görülüyor.

Çizelge 4.23: Diyarbakır DKDOO model sonuçları.

Model	R^2	RMSE	nRMSE	MAE	Puan
İkili ağaç	0,93878	67,9884	0,061086	34,6468	3
Dalgacık	0,93565	69,7035	0,062627	36,7557	2
Sigmoid	0,93504	70,0341	0,062924	35,7192	1

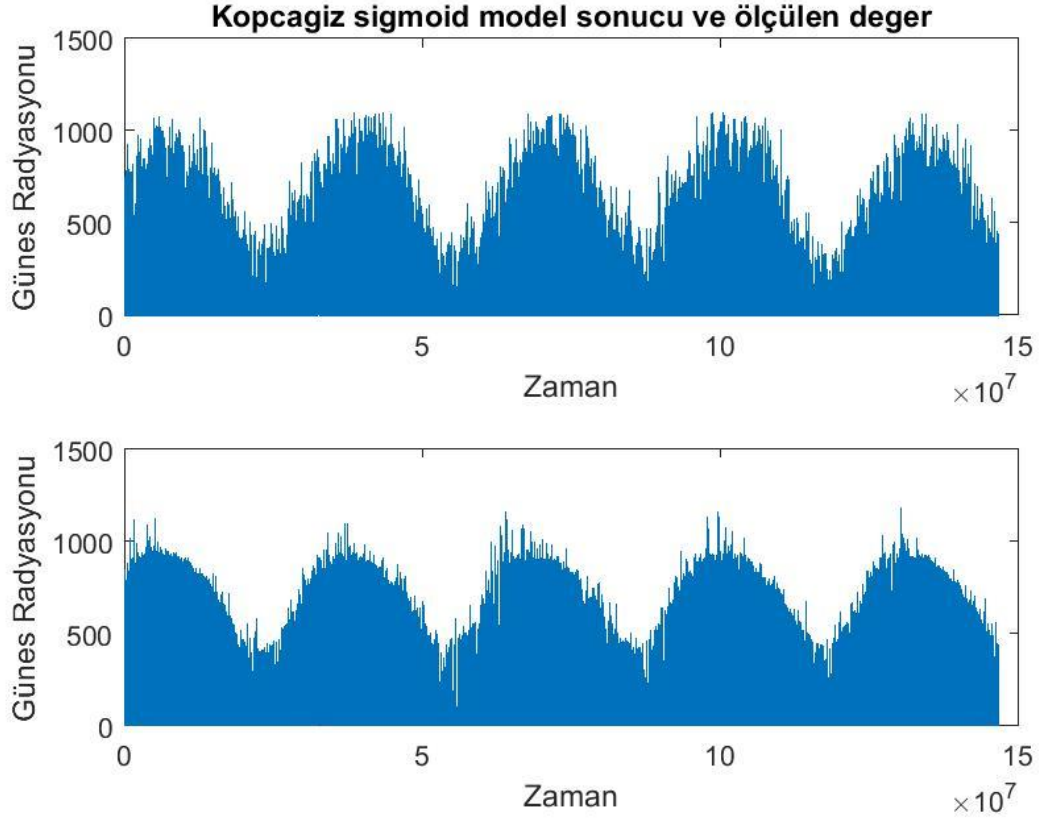


Şekil 4.48: Diyarbakır ikili ağaç model sonucu.

Çizelge 4.24: Gaziantep DKDOO model sonuçları.

Model	R^2	RMSE	nRMSE	MAE	Puan
İkili ağaç	0,92262	76,9244	0,065301	40,9162	3
Dalgacık	0,92035	78,0461	0,066253	41,7371	2
Sigmoid	0,91895	78,7283	0,066832	40,4376	1

Şekil 4.49’da Gaziantep verisi ile sigmoid aktivasyon fonksiyonu modeli sonucu ve ölçülen değer arasındaki farkı görebiliriz.



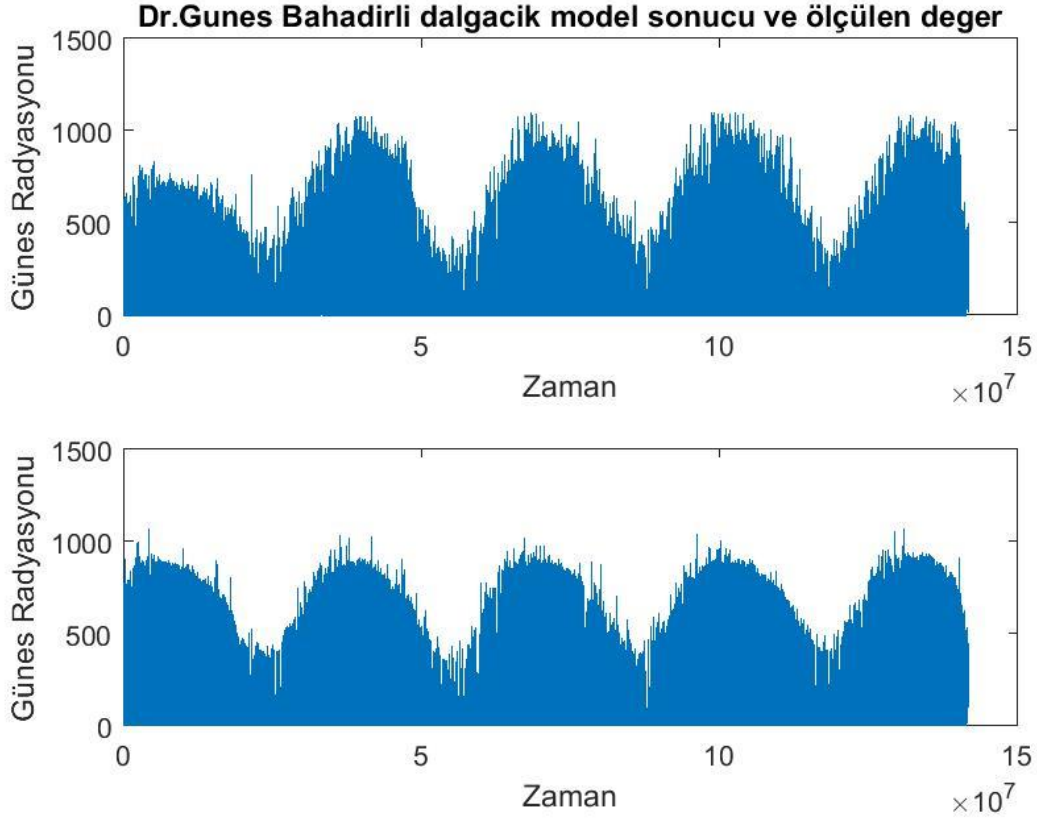
Şekil 4.49: Gaziantep sigmoid model sonucu.

Oluşturulan tahmin modeli Hatay verilerine uygulandığında Çizelge 4.25’de gösterildiği gibi dalgacık aktivasyon fonksiyonu ile en iyi sonucun alındığı görülmüyor.

Çizelge 4.25: Hatay DKDOO model sonuçları.

Model	R^2	RMSE	nRMSE	MAE	Puan
Dalgacık	0,93089	72,2447	0,067708	40,0399	3
İkili ağaç	0,92632	74,594	0,06991	38,9497	2
Sigmoid	0,91956	77,9393	0,073045	42,4833	1

Şekil 4.50’de Hatay verisi ile dalgacık aktivasyon fonksiyonu modeli sonucu ve ölçülen değer arasındaki farkı görebiliriz.



Şekil 4.50: Hatay dalgacık model sonucu.

Beş ölçüm merkezi için en iyi modeli bulmak için verilen puanlar ile bir sıralama yapıldığında Çizelge 4.26’da görüldüğü gibi 3 aktivasyon koşulu ile eşit sonuçlar elde ediliyor. Bu durumda doğrulamaları koşarken daha hızlı çözüme ulaşan ikili ağaç aktivasyonu en iyi model olarak seçilmiştir.

Çizelge 4.26: Beş ölçüm merkezi için en iyi DKDOO model seçimi.

Model	Puan
İkili ağaç	10
Sigmoid	10
Dalgacık	10

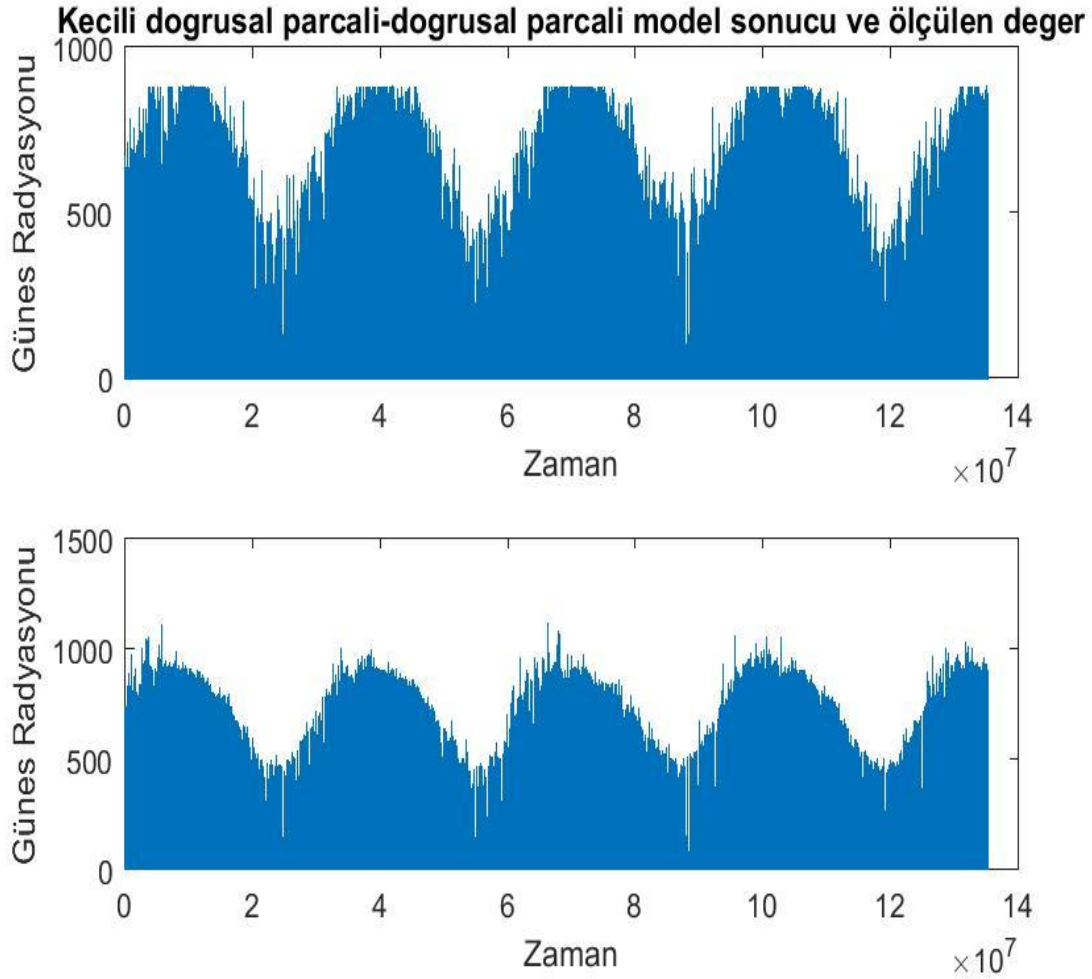
4.2.2.2 Hammerstein-Wiener modeli

Veri hazırlama bölümünde bahsedilen eğitim ve test verileri kullanılarak Hammerstein-Wiener modelleri MATLAB'ın System Identification Toolbox'ı kullanılarak oluşturulmuştur. Giriş ve çıkış doğrusalsızlığı için 5 farklı aktivasyon fonksiyonu olarak parçalı doğrusal, ölü bölge, doyunluk, sigmoid ve dalgacık aktivasyon fonksiyonu birbirleri ile oluşturulan kombinasyonlarından 25 farklı model elde edildi ve modeller diğer illerin verileri ile beslenerek valide edilmiştir. Doğrulama için her modele belli bir puan verilerek en son olarak hangi modelin 5 ölçüm merkezi için daha uygun olduğuna karar verilmiştir.

Çizelge 4.27: Şanlıurfa Hammerstein-Wiener model sonuçları.

Giriş doğrusalsızlığı	Çıkış doğrusalsızlığı	R ²	RMSE	nRMSE	MAE	Puan
Parçalı doğrusal	Parçalı doğrusal	0,95546	59,4179	0,053337	30,1109	10
Parçalı doğrusal	Sigmoid	0,94983	63,0612	0,056608	31,9289	4
Parçalı doğrusal	Doygunluk	0,95265	61,2582	0,054989	30,9852	7
Parçalı doğrusal	Ölü bölge	0,95411	60,3071	0,054136	31,2052	8
Parçalı doğrusal	Dalgacık	0,95053	62,6166	0,056209	31,9294	5
Sigmoid	Parçalı doğrusal	0,94122	68,2564	0,061271	35,4402	-
Sigmoid	Sigmoid	0,94952	63,2507	0,056778	32,2559	2
Sigmoid	Doygunluk	0,95085	62,4113	0,056025	31,4199	6
Sigmoid	Ölü bölge	0,95414	60,2864	0,054117	31,2578	9
Sigmoid	Dalgacık	0,92576	76,7083	0,068858	38,6326	-
Doygunluk	Parçalı doğrusal	0,94211	67,739	0,060807	35,4085	-
Doygunluk	Sigmoid	0,93948	69,2557	0,062168	36,2712	-
Doygunluk	Doygunluk	0,94966	63,1666	0,056702	32,7099	3
Doygunluk	Ölü bölge	0,94382	66,7278	0,059899	34,9193	-
Doygunluk	Dalgacık	0,93027	74,3427	0,066735	39,1648	-
Ölü bölge	Parçalı doğrusal	0,94419	66,5101	0,059704	34,4142	-
Ölü bölge	Sigmoid	0,94417	66,5191	0,059712	34,4103	-
Ölü bölge	Doygunluk	0,94351	66,9125	0,060065	35,0179	-
Ölü bölge	Ölü bölge	0,94943	63,3085	0,05683	33,0911	1
Ölü bölge	Dalgacık	0,92724	75,9408	0,068169	39,6636	-
Dalgacık	Parçalı doğrusal	0,43324	211,944	0,19026	118,602	-
Dalgacık	Sigmoid	0,48372	202,283	0,18158	114,650	-
Dalgacık	Doygunluk	0,29123	237,013	0,21276	135,941	-
Dalgacık	Ölü bölge	0,84416	111,138	0,099765	62,1376	-
Dalgacık	Dalgacık	0,47007	204,942	0,18397	116,468	-

Şanlıurfa için oluşturulan Hammerstein-Wiener modellerine bakıldığında Çizelge 4.27’de görüldüğü gibi giriş ve çıkış doğrusalsızlığının parçalı doğrusal olan model diğerlerinden daha iyi sonuç verdiği gözlemlenmiştir. Giriş doğrusalsızlığında sigmoid ve çıkış doğrusalsızlığında ölü bölge kullanılan modelde çok yakın sonuçlar elde ettiği gözlenmektedir. Şekil 4.51’de Şanlıurfa verisi ile giriş ve çıkış doğrusalsızlığının parçalı doğrusal olduğu Hammerstein-Wiener modeli sonucu ve ölçülen değer arasındaki farkı görebiliriz.

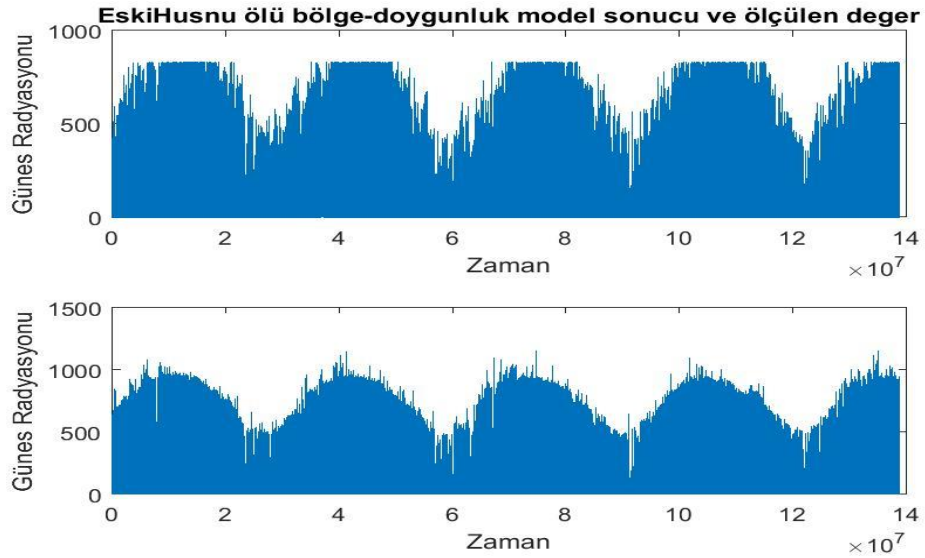


Şekil 4.51: Şanlıurfa giriş DP ve çıkış DP olan HW modeli sonucu.

Oluşturulan Hammerstein-Wiener modelleri Çizelge 4.28’deki gibi Adıyaman verisine uygulanınca giriş doğrusalsızlığı ölü bölge ve çıkış doğrusalsızlığının doygunluk olduğu model diğerlerinden daha iyi sonuç verdiği gözlemlenmiştir.

Çizelge 4.28: Adıyaman Hammerstein-Wiener model sonuçları.

Giriş doğrusalsızlığı	Çıkış doğrusalsızlığı	R ²	RMSE	nRMSE	MAE	Puan
Parçalı doğrusal	Parçalı doğrusal	0,93215	75,5278	0,065505	41,6756	7
Parçalı doğrusal	Sigmoid	0,92287	80,5275	0,069842	43,4681	-
Parçalı doğrusal	Doygunluk	0,92229	80,8263	0,070101	44,3326	-
Parçalı doğrusal	Ölü bölge	0,93269	75,2219	0,06524	41,5315	9
Parçalı doğrusal	Dalgacık	0,91794	83,0568	0,072035	44,794	-
Sigmoid	Parçalı doğrusal	0,91122	86,3936	0,074929	46,7536	-
Sigmoid	Sigmoid	0,92491	79,4555	0,068912	44,1326	-
Sigmoid	Doygunluk	0,92593	78,9121	0,068441	43,1804	-
Sigmoid	Ölü bölge	0,93269	75,2222	0,06524	41,4925	8
Sigmoid	Dalgacık	0,90344	90,0989	0,078143	47,6556	-
Doygunluk	Parçalı doğrusal	0,92913	77,1861	0,066944	42,4687	4
Doygunluk	Sigmoid	0,92849	77,5347	0,067246	42,5826	3
Doygunluk	Doygunluk	0,92841	77,5767	0,067283	42,8479	2
Doygunluk	Ölü bölge	0,92291	80,5042	0,069821	44,0395	-
Doygunluk	Dalgacık	0,91994	82,0396	0,071153	44,8286	-
Ölü bölge	Parçalı doğrusal	0,92948	76,9997	0,066782	42,6361	5
Ölü bölge	Sigmoid	0,92949	76,9893	0,066773	42,6266	6
Ölü bölge	Doygunluk	0,9339	74,5447	0,064653	40,5262	10
Ölü bölge	Ölü bölge	0,92627	78,7316	0,068284	43,4256	1
Ölü bölge	Dalgacık	0,916	84,0345	0,072883	44,3653	-
Dalgacık	Parçalı doğrusal	0,40922	222,859	0,19329	124,968	-
Dalgacık	Sigmoid	0,43301	218,326	0,18936	123,258	-
Dalgacık	Doygunluk	0,264	248,746	0,21574	142,651	-
Dalgacık	Ölü bölge	0,83354	118,296	0,1026	65,7458	-
Dalgacık	Dalgacık	0,43733	217,493	0,18863	123,122	-



Şekil 4.52: Adıyaman giriş ÖB ve çıkış doygunluk olan HW modeli sonucu.

Giriş doğrusalsızlığında parçalı doğrusal ve çıkış doğrusalsızlığında ölü bölge kullanılan modelde çok yakın sonuçlar elde ettiği gözlenmektedir. Şekil 4.52’de Adıyaman verisi ile giriş doğrusalsızlığı ölü bölge ve çıkış doğrusalsızlığının doygunluk olduğu Hammerstein-Wiener modeli sonucu ve ölçülen değer arasındaki farkı görebiliriz.

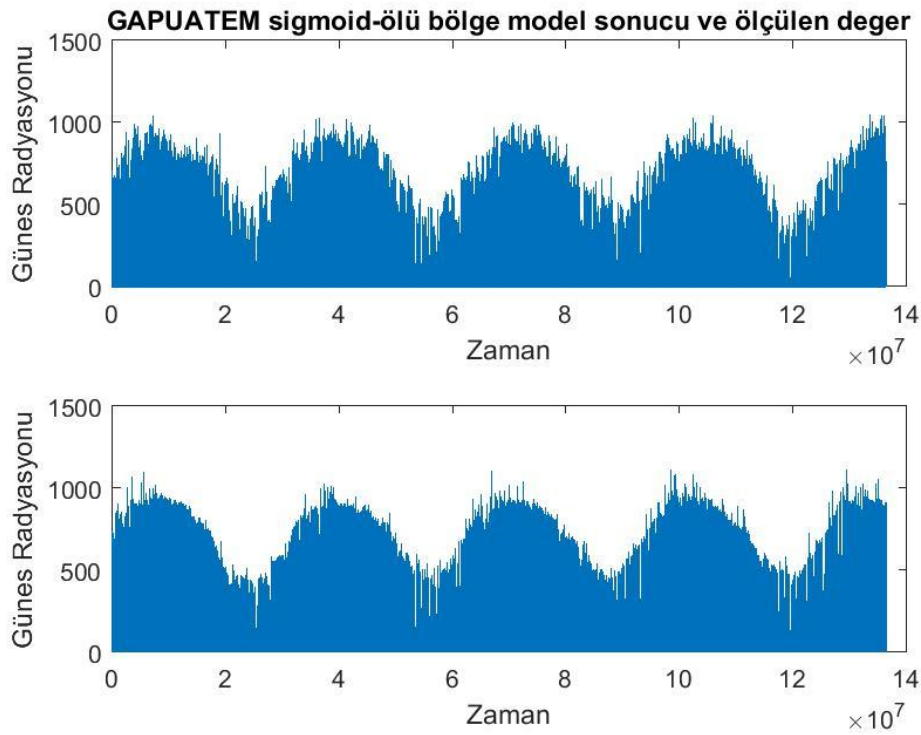
Çizelge 4.29: Diyarbakır Hammerstein-Wiener model sonuçları.

Giriş doğrusalsızlığı	Çıkış doğrusalsızlığı	R ²	RMSE	nRMSE	MAE	Puan
Parçalı doğrusal	Parçalı doğrusal	0,93705	68,9416	0,061942	34,796	3
Parçalı doğrusal	Sigmoid	0,93097	72,1966	0,064867	35,8417	-
Parçalı doğrusal	Doygunluk	0,9145	80,3488	0,072191	40,0535	-
Parçalı doğrusal	Ölü bölge	0,94237	65,9667	0,059269	34,1945	8
Parçalı doğrusal	Dalgacık	0,93116	72,0955	0,064776	35,7325	-
Sigmoid	Parçalı doğrusal	0,8913	90,5961	0,081398	47,1976	-
Sigmoid	Sigmoid	0,94003	67,2923	0,06046	33,4483	6
Sigmoid	Doygunluk	0,93311	71,0658	0,063851	35,3215	-
Sigmoid	Ölü bölge	0,94438	64,8059	0,058226	33,6453	10
Sigmoid	Dalgacık	0,91644	79,4303	0,071366	39,8287	-
Doygunluk	Parçalı doğrusal	0,93485	70,1388	0,063018	36,6411	2
Doygunluk	Sigmoid	0,93395	70,6232	0,063453	37,2955	-
Doygunluk	Doygunluk	0,94145	66,4893	0,059739	34,8009	7
Doygunluk	Ölü bölge	0,93848	68,1549	0,061235	36,2319	4
Doygunluk	Dalgacık	0,92692	74,2841	0,066742	39,7947	-
Ölü bölge	Parçalı doğrusal	0,93392	70,6362	0,063465	36,2869	-
Ölü bölge	Sigmoid	0,934	70,5932	0,063426	36,2468	1
Ölü bölge	Doygunluk	0,93969	67,4819	0,060631	35,6409	5
Ölü bölge	Ölü bölge	0,94278	65,7296	0,059056	34,4284	9
Ölü bölge	Dalgacık	0,92329	76,1052	0,068378	39,3399	-
Dalgacık	Parçalı doğrusal	0,42187	208,933	0,18772	115,336	-
Dalgacık	Sigmoid	0,48022	198,108	0,178	110,557	-
Dalgacık	Doygunluk	0,30192	229,587	0,20628	130,386	-
Dalgacık	Ölü bölge	0,78436	127,601	0,11465	73,2525	-
Dalgacık	Dalgacık	0,46317	201,333	0,18089	113,497	-

Oluşturulan Hammerstein-Wiener modelleri Çizelge 4.29'daki gibi Diyarbakır verisine uygulanınca giriş doğrusalsızlığı sigmoid ve çıkış doğrusalsızlığının ölü bölge olan model diğerlerinden daha iyi sonuç verdiği gözlemlenmiştir.

Giriş doğrusalsızlığında ölü bölge ve çıkış doğrusalsızlığında ölü bölge olarak kullanılan modelde çok yakın sonuçlar elde ettiği gözlenmektedir.

Şekil 4.53'te Diyarbakır verisi ile giriş doğrusalsızlığı sigmoid ve çıkış doğrusalsızlığının ölü bölge olduğu Hammerstein-Wiener modeli sonucu ve ölçülen değer arasındaki farkı görebiliriz.

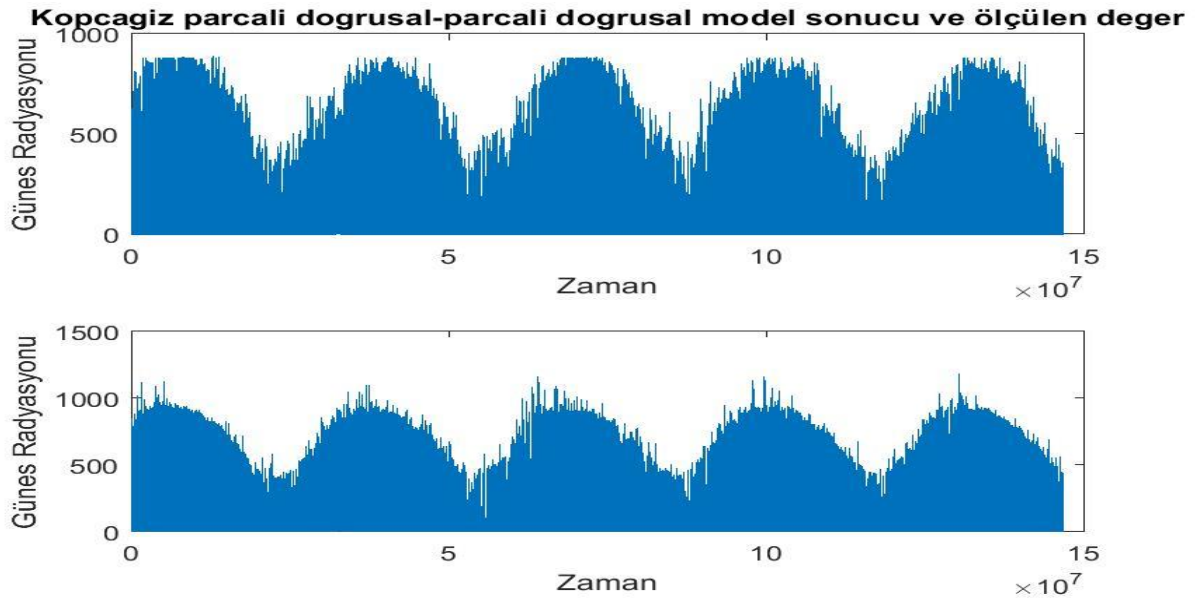


Şekil 4.53: Diyarbakır giriş sigmoid ve çıkış ÖB olan HW modeli sonucu

Oluşturulan Hammerstein-Wiener modelleri Çizelge 4.30'daki gibi Gaziantep verisine giriş ve çıkış doğrusalsızlığının parçalı doğrusal olan model diğerlerinden daha iyi sonuç verdiği gözlemlenmiştir. Giriş doğrusalsızlığında parçalı doğrusal ve çıkış doğrusalsızlığında ölü bölge olarak kullanılan model de yakın sonuçlar elde ettiği gözlenmektedir.

Çizelge 4.30: Gaziantep Hammerstein-Wiener model sonuçları.

Giriş doğrusalsızlığı	Çıkış doğrusalsızlığı	R ²	RMSE	nRMSE	MAE	Puan
Parçalı doğrusal	Parçalı doğrusal	0,94333	65,8314	0,055884	34,0973	10
Parçalı doğrusal	Sigmoid	0,93075	72,7731	0,061777	37,5878	-
Parçalı doğrusal	Doygunluk	0,92773	74,3449	0,063111	38,3542	-
Parçalı doğrusal	Ölü bölge	0,9429	66,0807	0,056096	34,735	9
Parçalı doğrusal	Dalgacık	0,92927	73,5467	0,062433	38,1783	-
Sigmoid	Parçalı doğrusal	0,89765	88,4714	0,075103	45,3491	-
Sigmoid	Sigmoid	0,93401	71,039	0,060305	37,0794	-
Sigmoid	Doygunluk	0,93285	71,6629	0,060834	37,0786	-
Sigmoid	Ölü bölge	0,94271	66,193	0,056191	34,9011	8
Sigmoid	Dalgacık	0,91487	80,6869	0,068495	41,2847	-
Doygunluk	Parçalı doğrusal	0,93563	70,1603	0,059559	36,3599	2
Doygunluk	Sigmoid	0,93472	70,6584	0,059982	36,6314	1
Doygunluk	Doygunluk	0,94112	67,1045	0,056965	34,5622	7
Doygunluk	Ölü bölge	0,93462	70,7118	0,060027	37,2917	-
Doygunluk	Dalgacık	0,92775	74,3303	0,063099	38,9452	-
Ölü bölge	Parçalı doğrusal	0,93815	68,7741	0,058382	35,4818	3
Ölü bölge	Sigmoid	0,93824	68,7255	0,058341	35,4382	4
Ölü bölge	Doygunluk	0,94046	67,4777	0,057282	35,253	6
Ölü bölge	Ölü bölge	0,93926	68,1522	0,057854	36,064	5
Ölü bölge	Dalgacık	0,91951	78,4546	0,0666	40,3928	-
Dalgacık	Parçalı doğrusal	0,41877	210,830	0,17897	116,258	-
Dalgacık	Sigmoid	0,46802	201,701	0,17122	113,522	-
Dalgacık	Doygunluk	0,29144	232,780	0,19761	132,964	-
Dalgacık	Ölü bölge	0,82726	114,935	0,097568	64,0964	-
Dalgacık	Dalgacık	0,17605	207,386	0,17605	115,904	-



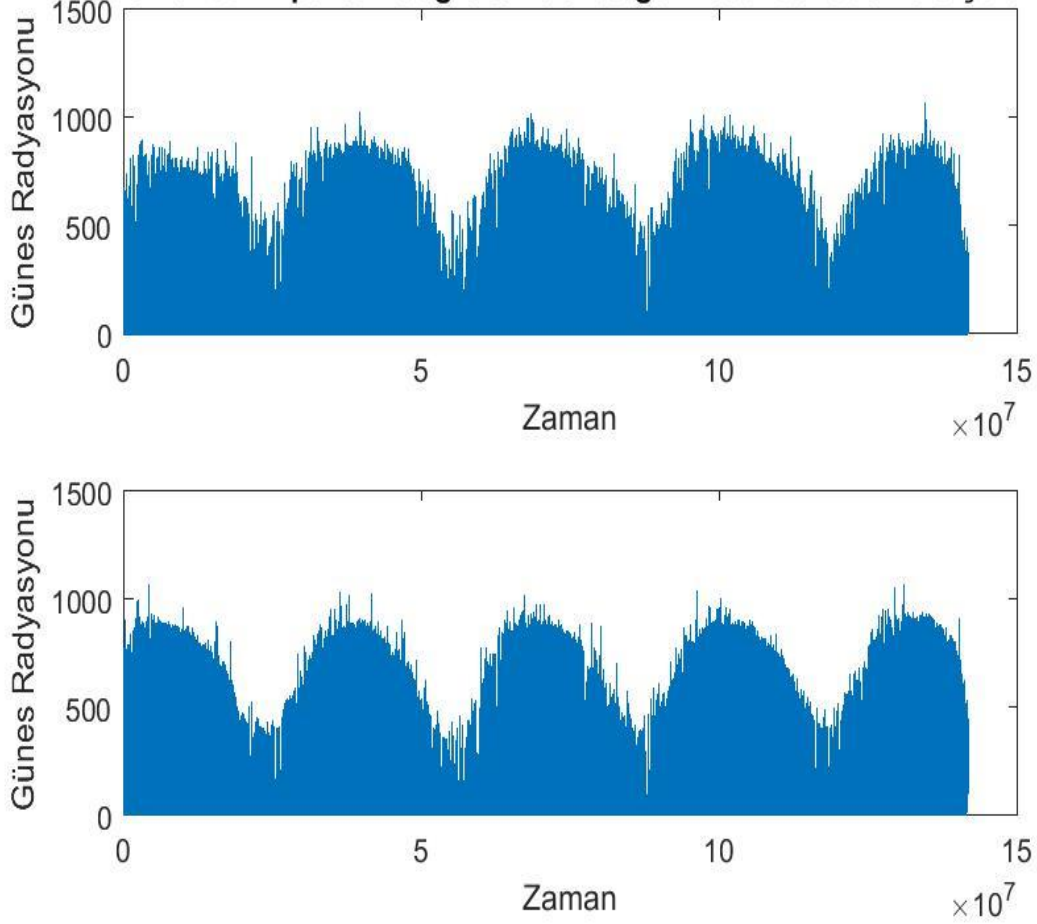
Şekil 4.54: Gaziantep giriş DP ve çıkış DP olan HW modeli sonucu.

Şekil 4.54’de Gaziantep verisi ile giriş ve çıkış doğrusalsızlığının parçalı doğrusal olan Hammerstein-Wiener modeli sonucu ve ölçülen değer arasındaki farkı görebiliriz. Oluşturulan Hammerstein-Wiener modelleri Çizelge 4.31’teki gibi Hatay verisine giriş doğrusalsızlığının parçalı doğrusal ve çıkış doğrusalsızlığı ölü bölge olan model diğerlerinden daha iyi sonuç verdiği gözlemlenmiştir. Giriş doğrusalsızlığında ölü bölge ve çıkış doğrusalsızlığında ölü bölge olarak kullanılan model de yakın sonuçlar elde ettiği gözlenmektedir.

Çizelge 4.31: Hatay Hammerstein-Wiener model sonuçları.

Giriş doğrusalsızlığı	Çıkış doğrusalsızlığı	R ²	RMSE	nRMSE	MAE	Puan
Parçalı doğrusal	Parçalı doğrusal	0,92777	73,8552	0,069218	39,7768	-
Parçalı doğrusal	Sigmoid	0,91972	77,864	0,072975	42,6351	-
Parçalı doğrusal	Doygunluk	0,91284	81,1307	0,076036	44,3525	-
Parçalı doğrusal	Ölü bölge	0,93641	69,2996	0,064948	36,5529	10
Parçalı doğrusal	Dalgacık	0,91634	79,485	0,074494	43,7842	-
Sigmoid	Parçalı doğrusal	0,90253	85,7947	0,080407	45,4031	-
Sigmoid	Sigmoid	0,92976	72,8326	0,068259	38,2027	1
Sigmoid	Doygunluk	0,90935	82,7393	0,077544	44,8432	-
Sigmoid	Ölü bölge	0,93527	69,9171	0,065527	37,2376	7
Sigmoid	Dalgacık	0,90034	86,7553	0,081308	45,4813	-
Doygunluk	Parçalı doğrusal	0,93003	72,6909	0,068126	38,7719	2
Doygunluk	Sigmoid	0,93087	72,2561	0,067719	38,1995	5
Doygunluk	Doygunluk	0,93111	72,1304	0,067601	38,2977	6
Doygunluk	Ölü bölge	0,93027	72,5653	0,068009	37,9121	3
Doygunluk	Dalgacık	0,9353	69,9003	0,065511	37,2325	8
Ölü bölge	Parçalı doğrusal	0,9278	73,8399	0,069203	38,9399	-
Ölü bölge	Sigmoid	0,92799	73,7439	0,069113	38,8938	-
Ölü bölge	Doygunluk	0,92785	73,8166	0,069181	39,7467	-
Ölü bölge	Ölü bölge	0,93577	69,6493	0,065276	37,1368	9
Ölü bölge	Dalgacık	0,93044	72,4768	0,067926	37,2422	4
Dalgacık	Parçalı doğrusal	0,4072	211,586	0,1983	118,767	-
Dalgacık	Sigmoid	0,40831	211,387	0,19811	120,395	-
Dalgacık	Doygunluk	0,28894	231,731	0,21718	133,421	-
Dalgacık	Ölü bölge	0,80126	122,511	0,11482	70,596	-
Dalgacık	Dalgacık	0,38709	215,144	0,20163	122,777	-

Dr.Gunes Bahadirli parcali dogrusal-ölü bölge model sonucu ve ölçülen degeri



Şekil 4.55: Hatay giriş DP ve çıkış ÖB olan HW modeli sonucu

Şekil 4.55’te Hatay verisi giriş doğrusalsızlığının parçalı doğrusal ve çıkış doğrusalsızlığı ölü bölge olan Hammerstein-Wiener modeli sonucu ve ölçülen değer arasındaki farkı görebiliriz.

Beş ölçüm merkezi için en iyi modeli bulmak için verilen puanlar ile bir sıralama yapıldığında Çizelge 4.32’de görüldüğü gibi giriş doğrusalsızlığı parçalı doğrusal ve çıkış doğrusalsızlığı ölü bölge olan model bütün bölgeler için iyi sonuçlar vermektedir.

Sistem tanılama tahmin modelleri oluşturulurken MATLAB 2016B versiyonu kullanıldı. DELL LATITUDE E5450 bilgisayar modeli kullanılmıştır ve Intel core i5-5300U işlemci seti ve 16 GB RAM’e sahiptir.

Çizelge 4.32: Beş ölçüm merkezi için en iyi HW model seçimi.

Giriş doğrusalsızlığı	Çıkış doğrusalsızlığı	Puan
Parçalı doğrusal	Ölü bölge	44
Sigmoid	Ölü bölge	42
Parçalı doğrusal	Parçalı doğrusal	30
Ölü bölge	Ölü bölge	25
Doygunluk	Doygunluk	25
Ölü bölge	Doygunluk	21
Ölü bölge	Sigmoid	11
Doygunluk	Parçalı doğrusal	10
Sigmoid	Sigmoid	9
Doygunluk	Sigmoid	9
Ölü bölge	Parçalı doğrusal	8
Doygunluk	Dalgacık	8
Parçalı doğrusal	Doygunluk	7
Doygunluk	Ölü bölge	7
Sigmoid	Doygunluk	6
Parçalı doğrusal	Dalgacık	5
Parçalı doğrusal	Sigmoid	4
Ölü bölge	Dalgacık	4

4.3 Makina Öğrenmesi Modelleri

Destek vektör regresyon

Destek vektör regresyonu kullanılarak radyal tabanlı fonksiyon, sigmoid ve polinom kernelleri ile 5 ölçüm merkezi için modeller oluşturulmuştur. Radyal temelli kernel için farklı C , γ ve ε değerleri ile 12 model oluşturuldu ve en iyi olan 3 tanesi 4 ölçüm merkezi için valide edilmiştir. Sigmoid kerneli için farklı C ve γ değerleri 10 farklı model oluşturuldu ve en iyi olan model 5 ölçüm merkezi için valide edildi. Polinom kerneli için farklı C ve γ değerleri 6 farklı model oluşturuldu ancak polinom kerneli yakınsıyamadığı için doğrulama işlemi yapılmamıştır.

Radyal tabanlı fonksiyon kerneli ile model 4,6 ve 12'nin iyi sonuçlar verdiği görülüyor. Model 11 ve 12 arasında ε olarak fark olmasına rağmen değişiklik olmadığından büyük olan seçilmiştir. Model 3 ve 4 arasında ε olarak fark olmasına tahmin değerlerinde ufak farklar vardır bu sebeple model 4 seçilmiştir. Genel olarak bakıldığında seçilen ε değerleri ile oluşturulan modellerde fark olmadığı görülmüştür. Çizelge 4.33'te görüldüğü gibi en iyi sonucu $C = 10$ ve $\gamma = 1$ değerleri ile elde edilmiştir.

Çizelge 4.33: Şanlıurfa RTF kernel DVR model sonuçları.

Model	C	γ	ε	R^2	RMSE	nRMSE	MAE
1	0,1	1	0,0001	0,8291	116,3665	0,0976	72,0161
2	0,1	1	0,01	0,8291	116,3672	0,0976	72,0163
3	0,1	0,1	0,0001	0,8467	110,2309	0,0925	71,5634
4	0,1	0,1	0,01	0,8467	110,2296	0,0925	71,5623
5	10	1	0,0001	0,8890	93,7735	0,0787	51,4030
6	10	1	0,01	0,8890	93,7737	0,0787	51,4031
7	10	0,1	0,0001	0,8817	96,8199	0,0812	57,0562
8	10	0,1	0,01	0,8817	96,8204	0,0812	57,0563
9	1000	1	0,0001	0,8291	116,3673	0,0976	72,0163
10	1000	1	0,01	0,8291	116,3672	0,0976	72,0163
11	1000	0,1	0,0001	0,8846	95,6339	0,0802	54,2426
12	1000	0,1	0,01	0,8846	95,6339	0,0802	54,2427

Çizelge 4.34’te model 4 için 5 ölçüm merkezi tahmin sonuçlarına bakıldığında Adıyaman ve Gaziantep doğrulama değerleri Şanlıurfa’dan daha düşük çıkmasına rağmen Diyarbakır verileri ile daha iyi sonuç vermiştir.

Çizelge 4.34: Model 4 RTF kernel DVR doğrulama sonuçları.

Model 4	R^2	RMSE	nRMSE	MAE
Şanlıurfa	0,8467	110,2296	0,0925	71,5623
Adıyaman	0,8274	120,3620	0,1029	78,0009
Diyarbakır	0,8510	106,0606	0,0932	69,7348
Gaziantep	0,8277	114,7241	0,0973	76,7725
Hatay	0,8362	111,2673	0,0940	74,4278

Çizelge 4.35’te model 6 için 5 ölçüm merkezi tahmin sonuçlarına görülmektedir.

Çizelge 4.35: Model 6 RTF kernel DVR doğrulama sonuçları.

Model 6	R^2	RMSE	nRMSE	MAE
Şanlıurfa	0,8890	93,7737	0,0787	51,4031
Adıyaman	0,8764	101,8700	0,0871	59,0789
Diyarbakır	0,8751	97,1153	0,0853	59,0023
Gaziantep	0,8272	114,8639	0,0974	66,3907
Hatay	0,8106	119,6271	0,1010	72,0269

Model 6 için 5 ölçüm merkezi tahmin sonuçlarına bakıldığında Diyarbakır ve Adıyaman doğrulama değerlerinin Şanlıurfa'ya yakın olmasına rağmen Hatay ve Gaziantep da düşük sonuçlar vermiştir.

Çizelge 4.36'da model 12 için 5 ölçüm merkezi tahmin sonuçlarına bakıldığında Diyarbakır ve Adıyaman doğrulama değerlerinin Şanlıurfa'ya yakın olmasına rağmen Hatay ve Gaziantep da düşük sonuçlar vermiştir.

Çizelge 4.36: Model 12 RTF kernel DVR doğrulama sonuçları.

Model 12	R^2	RMSE	nRMSE	MAE
Şanlıurfa	0,8846	95,6339	0,0802	54,2427
Adıyaman	0,8745	102,6294	0,0877	60,4756
Diyarbakır	0,8740	97,5141	0,0857	60,5769
Gaziantep	0,8289	114,3031	0,0969	67,3028
Hatay	0,8159	117,9577	0,0996	72,3945

Model 4, 6 ve 12 arasındaki karşılaştırma ile birini seçmemiz gerekirse model 6 seçilebilir çünkü Şanlıurfa, Adıyaman ve Diyarbakır'da diğer modellerden daha iyi sonuç vermektedir ve Gaziantep ile Hatay için de kabul edilebilir farklar olduğu görülmektedir.

Çizelge 4.37'ye bakıldığı zaman model 2 ve model 7 dışında yakınsayabilen model olmadığı görülüyor. Model 7'in değerleri de yeterli bulunmadığı için model 2 ile doğrulamalar yapıldı.

Çizelge 4.37: Şanlıurfa sigmoid kernel ile DVR model sonuçları.

Model	C	γ	R^2	RMSE	nRMSE	MAE
1	0,1	1	NA	NA	NA	NA
2	0,1	0,1	0,7775	132,7795	0,1114	93,7210
3	10	1	NA	NA	NA	NA
4	10	0,1	NA	NA	NA	NA
5	1000	1	NA	NA	NA	NA
6	1000	0,1	NA	NA	NA	NA
7	0,1	0,01	0,4248	213,4887	0,1831	140,1759
8	0,01	0,01	NA	NA	NA	NA
9	0,1	0,001	NA	NA	NA	NA
10	1	0,1	NA	NA	NA	NA

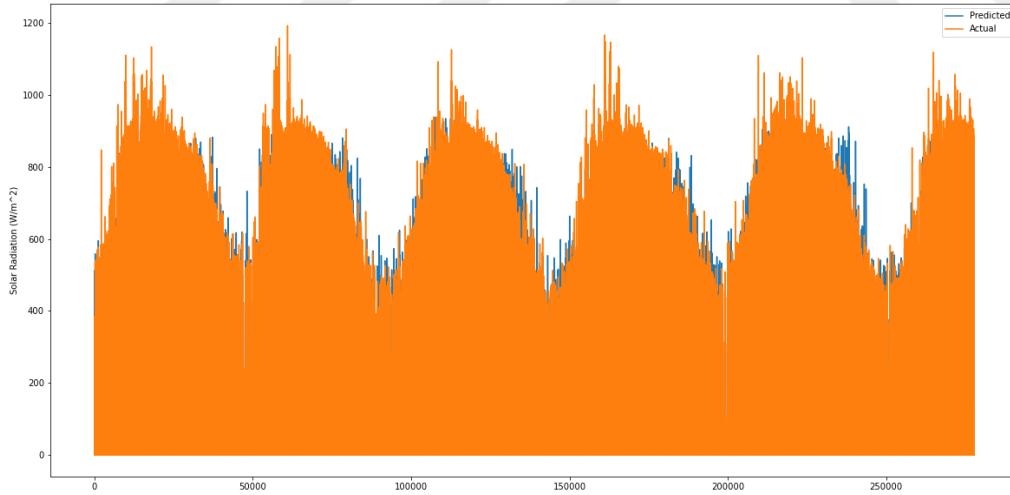
Model 2 için Çizelge 4.38’deki doğrulama sonuçlarına bakıldığı zaman RTF da olduğu gibi Gaziantep ve Hatay doğrulama sonuçları düşük görünüyor ancak Adıyaman ve Diyarbakır sonuçları Şanlıurfa sonuçlarına yakın olduğu görünüyor.

Çizelge 4.38: Model 2 sigmoid kernel ile DVR doğrulama sonuçları.

Model 2	R^2	RMSE	nRMSE	MAE
Şanlıurfa	0,7775	132,7795	0,1114	93,7210
Adıyaman	0,7670	139,8401	0,1195	96,1000
Diyarbakır	0,7785	129,3113	0,1136	91,7692
Gaziantep	0,7469	139,0318	0,1179	95,5877
Hatay	0,7780	129,5221	0,1094	88,8199

Polinom kerneli modelleri yakınsayamadı bu sebeple doğrulaması yapılamadı. Radyal temelli fonksiyon modellerinin sigmoid modellerinden düşük farklarla da olsa daha iyi olduğunu söyleyebiliriz.

Destek vektör regresyon modellerinden model 6, 5 ölçüm merkezi için en iyisi olarak seçilmiştir.



Şekil 4.56: Şanlıurfa Model 6 RTF kerneli ile DVR.

Şekil 4.56’da model 6’nın oluşturulduğu Şanlıurfa tahmin değeri ve ölçülen değerin arasındaki farkı görebiliriz.

DVR modelleri oluşturulurken Phyton 3.8 versiyonu sklearn kütüphanesi 1.0 versiyonu kullanılmıştır. ASUS K53SV bilgisayar modeli kullanılmıştır ve Intel Core i7-2670QM işlemci seti ve 8 GB RAM'e sahiptir.

Çapraz doğrulama, uygulanan modellerin kullanılmayan verilerdeki değerlendirmesinin doğru bir şekilde ve yansız olarak ele alınmasını sağlayan istatistiksel bir örnekleme yöntemidir. Uygulanan modellerin yüksek performanslarının rastgele olup olmadığını kontrol etmemizi sağlamaktadır. Çapraz doğrulamanın amacı, eğitim verilerinden doğrulama kümeleri olarak bilinen örnek gözlem bölümlerini oluşturmaktır. Kullanılan modellerin seçilen eğitim verilerine uygulanması ile performansı, doğrulama kümelerine karşı ölçülür ve yeni tahminler yapılmak istendiğinde doğruluğu yüksek bir sonuç elde edilebilir. Kullanılacak bölüm sayısı, örnek veri sayısına ve önyargı ile varyans arasında bir ödünleşme dengesini gözönüne alarak objektifliğin azalmasına sebep olacak ve varyans göre değişimin artmasına bağlıdır.

Destek vektör regresyonunda kullanılan en yaygın yöntem ise k katlamalı çapraz doğrulama methodudur. Veri kümesi, k gruba ayrılır ve bekletme yöntemi k kere tekrarlanmaktadır. Her grup için, seçilen grup doğrulama seti olarak kullanılır ve k-1 alt grupları bir eğitim grubu oluşturulmak için kullanılır. Daha sonra tüm k denemelerinde ortalama hata, standart sapma, maksimum ve minimum gibi istatistiksel değerlendirmeler hesaplanır. Veriler en az bir kez test grubunda ve k-1 kez ise eğitim grubunda değerlendirilirler. k arttıkça sonuç tahmini varyansı azaltılır. k değerinin veri sayısına uygun olarak seçilmemesi modelin performansının yanlış değerlendirilmesine sebep olabilmektedir. k değerini seçmek için kesin bir kural kullanılmamaktadır. Bu tez kapsamında destek vektör regresyonu için k değeri 5 olarak seçilmiştir. Çapraz doğrulama için Phyton programlama dilinin Sklearn kütüphanesinden cross_val_score fonksiyonu kullanılmıştır.

4.4 Derin Öğrenme Modelleri

Uzun kısa süreli hafıza modeli

Veri hazırlama kısmında bahsedilen sıcaklık dış, nem hava, hava basıncı, ET0 ve güneş radyasyonu kullanarak çok değişkenli uzun kısa süreli hafıza modelleri oluşturulmuştur. UKSH'nin hiperparametreleri optimizasyonu ve 5 bölge için doğrulama yapılmıştır.

Çizelge 4.39'a bakıldığında küme büyüklüğü sayısı 8 olana kadar tahmin değerinin arttığı daha sonra düştüğü görülmektedir. En uygun küme büyüklüğü sayısı 8 olarak seçilmiştir.

Çizelge 4.39: Şanlıurfa UKSH küme büyüklüğü ayarlaması.

Nöron sayısı	Epoch sayısı	Küme büyüklüğü	R^2	RMSE	nRMSE	MAE
16	40	1	0,9680	51,0147	0,0464	21,0236
16	40	2	0,9679	51,0365	0,0464	23,6775
16	40	4	0,9724	47,3786	0,0431	19,1285
16	40	8	0,9729	46,9611	0,0427	18,3212
16	40	16	0,9290	75,9754	0,0691	55,5548
16	40	32	0,9192	81,0053	0,0736	59,6702
16	40	40	0,9294	75,7314	0,0688	55,5744
16	40	64	0,8754	100,6337	0,0915	77,1771
16	40	128	0,8385	114,5500	0,1041	93,4501
16	40	256	0,2956	239,2310	0,2175	197,260

Çizelge 4.40'a bakıldığında nöron sayısı 32 olana kadar tahmin değerinin arttığı daha sonra düştüğü görülmektedir. En uygun nöron sayısı 32 olarak seçilmiştir.

Çizelge 4.40: Şanlıurfa UKSH nöron sayısı ayarlaması.

Nöron sayısı	Epoch sayısı	Küme büyüklüğü	R^2	RMSE	nRMSE	MAE
1	40	8	0,9265	77,2963	0,0703	60,8847
2	40	8	0,9392	70,2931	0,0639	43,2750
4	40	8	0,9535	61,4920	0,0559	38,2103
8	40	8	0,9701	49,3014	0,0448	24,6436
16	40	8	0,9725	47,2496	0,0430	20,6471
32	40	8	0,9733	46,5873	0,0424	17,1140
40	40	8	0,9693	49,9473	0,0454	23,9335
64	40	8	0,9726	47,1404	0,0429	18,0992
128	40	8	0,9727	47,1346	0,0428	18,0826

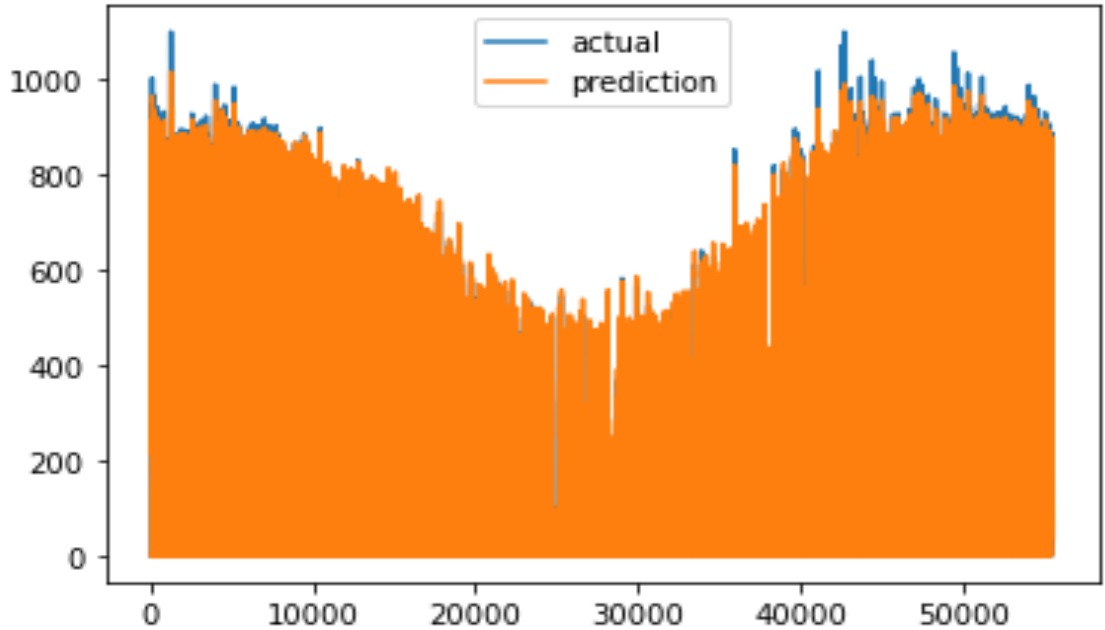
Çizelge 4.41'e bakıldığında epoch sayısı 256 olana kadar tahmin değerinin arttığı daha sonra düştüğü görülmektedir. En uygun epoch sayısı 256 olarak seçilmiştir.

Çizelge 4.41: Şanlıurfa UKSH epoch sayısı ayarlaması.

Nöron sayısı	Epoch sayısı	Küme büyüklüğü	R^2	RMSE	nRMSE	MAE
32	4	8	0,8268	118,6365	0,1079	96,2454
32	8	8	0,9326	74,0094	0,0673	51,7303
32	16	8	0,9325	74,0602	0,0673	52,5942
32	32	8	0,9727	47,0853	0,0428	19,5756
32	40	8	0,9717	47,9750	0,0436	20,6446
32	64	8	0,9733	46,5866	0,0424	16,9668
32	128	8	0,9734	46,5118	0,0423	16,7580
32	256	8	0,9742	45,7469	0,0416	15,6318

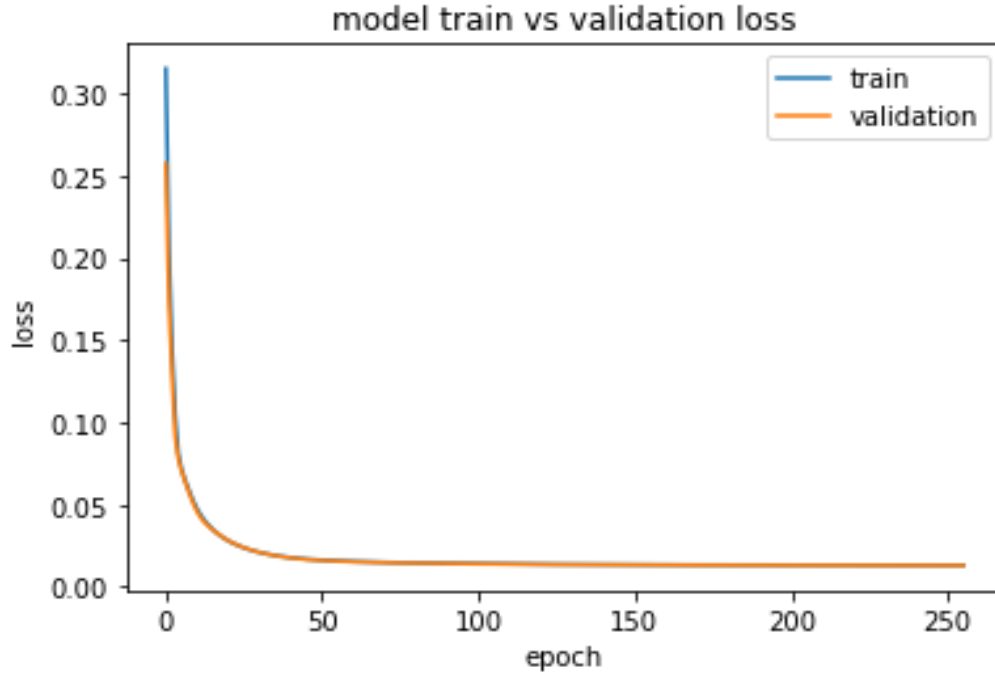
Hiperparametrelerin ayarlanması ile noron sayısı 32, küme büyüklüğü sayısı 8 ve epoch sayısı 256 olarak UKSH modeli oluşturulmuştur.

Oluşturulan modelin doğrulama sonucu Şekil 4.57'de verildiği gibidir. Şekil 4.58'de eğitim ve test arasındaki ilişki gösterilmektedir.



Şekil 4.57: Şanlıurfa UKSH model sonucu.

Şekil 4.58 ‘de görüldüğü gibi tahmin değeri gerçek değeri oldukça başarılı takip etmektedir.



Şekil 4.58: Şanlıurfa UKSH eğitim ve test kayıpları.

Oluşturulan UKSH modeli 4 ölçüm merkezi verilerine uygulandığında sonuçları Çizelge 4.42’te görebiliriz. UKSH, 5 ölçüm merkezi içinde oldukça iyi performans gösterdiği görülmektedir.

Çizelge 4.42: Şanlıurfa UKSH doğrulama sonuçları.

Ölçüm Merkezi	R ²	RMSE	nRMSE	MAE
Şanlıurfa	0,9742	45,7469	0,0416	15,6318
Adıyaman	0,9723	48,2496	0,0439	17,3216
Diyarbakır	0,9680	49,1183	0,0447	17,1007
Gaziantep	0,9614	54,7078	0,0497	18,3098
Dr.Günes Bahadırlı	0,9763	42,2268	0,0384	15,6975

UKSH modelleri oluşturulurken Phyton 3.7 versiyonu Tensorflow 2.41 versiyonu ve Keras kütüphanesi için 2.2 versiyonu kullanılmıştır. ASUS K53SV bilgisayar modeli kullanılmıştır ve Intel Core i7-2670QM işlemci seti ve 8 GB RAM’e sahiptir.

5. SONUÇLAR

İstatistiki yöntemler geleceğe dair öngörülerin istikrarlı olması için kullanılmaktadır. Geçmiş verilere istatistiksel modeller uygulanarak gelecek verileri belli hata oranları ile elde edilebilmektedir. Hata oranlarının azalması ve tahmin değerlerinin doğruluğunun artması ile insan yaşamına dair alanlarda engellerin ortadan kalkmasını ve kolaylaşmasını sağlayacaktır. Güneş radyasyonu tahminin doğruluğu kurulacak güneş enerji santralleri için maliyet hesapları, planlama ve organizasyon için oldukça önemlidir.

Bu tez kapsamında güneş radyasyonu değerinin sistem tanılama ve makine öğrenmesi modelleri ile güneş radyasyonu verisine ihtiyaç duymadan ve sensör maliyetlerinden kaçınarak temel meteorolojik değişkenler ile tahmin modelleri ve güneş radyasyonu verisi ölçümler arasında varsa derin öğrenme yöntemleri ile tahmin modellerine yer verilmiştir.

Çalışma kapsamında, 10 dakikalık aralıklarla ölçülen veriler üzerinde zaman serisi için tahmin modelleri uygulanmıştır. 5 ölçüm merkezi olarak kullanılmak üzere Şanlıurfa, Adıyaman, Diyarbakır, Gaziantep ve Hatay için seçilen sıcaklık dış, nem hava, hava basıncı, ET0 ve güneş radyasyonu verileri ile Keçili, Şanlıurfa verileri baz alınarak çok değişkenli sistem tanılama, makine öğrenmesi ve derin öğrenme tahmin modelleri güneş radyasyonu tahmini için hazırlanmış ve 5 ölçüm merkezi için doğrulamaları yapılmıştır. Doğrulama yapılırken tahmin modellerinde sıkça kullanılan 4 farklı performans metriği olarak R^2 , ortalama mutlak hata, kök ortalama kare hata ve normalleştirilmiş kök ortalama kare hata kullanılmıştır. Öncelikle doğrusal sistem tanılama tahmin modelleri olan dışsal kaynaklı otoregresif model ve bir otoregresif entegre hareketli ortalama modeli olan Box-Jenkins tahmin modeli kullanılarak deneme yanılma yöntemi ile sonuçlar elde edilmiş ve problemin doğasının doğrusal olmadığına karar verilmiştir. Daha sonra doğrusal olmayan tanılama yöntemleri olan dışsal kaynaklı otoregresif model ve Hammerstein-Wiener modeli kullanarak tahmin

sonuçları elde edilmiş, sonuçların doğrusal modellerden daha iyi olduğu görülmüştür. Sistem tanılama yöntemlerinden sonra makine öğrenme modeli olan ve doğrusal yapısını kerneller kullanarak doğrusal olmayan yapı da kullandığımız destek vektör regresyon yöntemi sigmoid, radyal tabanlı fonksiyon ve polinom kernelleri kullanılarak tahminler yapıldı ancak sonuçların sistem tanımlama yöntemlerinden daha iyi olmadığı görüldü ve bunun sebebi olarak veri sayısının fazlalığının regresyon modellerinin doğru aralıklar elde edemediğine ve destek vektör regresyonunun az veri ile daha iyi sonuçlar verebileceğine çıkarımı yapılmıştır. 3 kernelin arasında polinom kernelinin yakınsayamadığı ve bazı C ve γ değerlerinde sigmoid fonksiyonunda problemler yaşadığı görülmüştür. Radyal tabanlı fonksiyon kernelinde en iyi sonuçlar elde edilmiştir. Derin öğrenme yöntemlerinden tekrarlayan sinir ağlarının uzun süreli bağımlılık problemini gideren karmaşık doğrusal olmayan problemleri çözme ve sonucu genelleştirme yeteneğine sahip olan uzun kısa süreli hafıza kullanılarak en iyi hiperparametreler olan nöron sayısı, epoch sayısı ve küme büyüklüğü aranarak oluşturulan modeller kullanılmıştır. 5 ölçüm merkezi için oluşturulan modellerin sonuçlarının yorumları aşağıda verilmiştir.

Çizelge 5.43’de Şanlıurfa için tanımlanan modellerin sonuçları verilmiştir. DKDO ile BJ doğrusal modellerinin %1,5’luk tahmin farkı ve DKDOO ile HW modellerinin oldukça yakın sonuçlar verdiği görülmektedir ancak doğrusal olmayan modeller doğrusal modellerden hata metrikleri kıyasla daha iyi sonuçlar elde edilmiştir. DVR modelinin oluşturulan modeller arasında R^2 , RMSE, nRMSE ve MAE değerlerine bakarak en kötü performansı gösterdiği ve UKSH modelinin en iyi sonucu verdiği analizi yapılabilir.

Çizelge 5.43: Şanlıurfa için model doğrulama sonuçları karşılaştırması.

Şanlıurfa modelleri	R^2	RMSE	nRMSE	MAE
DKDO	0,91314	82,9718	0,074481	45,5902
BJ	0,92711	76,0172	0,068238	40,8621
DKDOO	0,95538	59,4713	0,053385	31,1725
HW	0,95546	59,4179	0,053337	30,1109
DVR	0,8890	93,7737	0,0787	51,4031
UKSH	0,9742	45,7469	0,0416	15,6318

UKSH en yakın tahmin modeli olan HW modelinden %1,9'luk tahmin farkı ile daha iyi olduğu görülmektedir. DVR modeli en kötü performansı gösterirken UKSH modeli ile %9,5'luk tahmin farkı olarak iyileştirme sağlanmıştır. Çizelge 5.44'te Adıyaman için doğrulama sonuçları verilmiştir.

Çizelge 5.44: Adıyaman için model doğrulama sonuçları karşılaştırması.

Adıyaman modelleri	R^2	RMSE	nRMSE	MAE
DKDO	0,8474	113,2648	0,098235	61,6989
BJ	0,91154	86,2364	0,074793	47,9061
DKDOO	0,92414	79,8586	0,069262	42,7484
HW	0,9339	74,5447	0,064653	40,5262
DVR	0,8764	101,8700	0,0871	59,0789
UKSH	0,9723	48,2496	0,0439	17,3216

Doğrusal modeller olan DKDO ve BJ arasındaki fark göze çarpmaktadır. Buradaki farkın sebebi verilerdeki gürültünün BJ modellerinde daha iyi modellenmesi olarak gösterilebilir. DKDOO ile HW modellerinin yakın sonuçlar verdiği görülmektedir ve aradaki farkın HW modelinin çıkış doğrusalsızlığını kullanarak model hatasını azaltması olarak belirtebiliriz. UKSH modelinin R^2 , RMSE, nRMSE ve MAE değerlerine bakarak en iyi performansı gösterdiği ve DKDO ile beraber DVR modellerinin yeterli sonuçlar göstermediği analizi yapılabilir. UKSH en yakın tahmin modeli olan HW modelinden %4,1'lik tahmin farkı ile daha iyi olduğu görülmektedir. DKDO modeli en kötü performansı gösterirken UKSH modeli ile %14,7'lik tahmin farkı olarak iyileştirme sağlanmıştır. Çizelge 5.45'te Diyarbakır için doğrulama sonuçları verilmiştir.

Çizelge 5.45: Diyarbakır için model doğrulama sonuçları karşılaştırması.

Diyarbakır modelleri	R^2	RMSE	nRMSE	MAE
DKDO	0,85313	105,3063	0,094615	56,843
BJ	0,92292	76,2904	0,068545	41,0225
DKDOO	0,93878	67,9884	0,061086	34,6468
HW	0,94438	64,8059	0,058226	33,6453
DVR	0,8751	97,1153	0,0853	59,0023
UKSH	0,9680	49,1183	0,0447	17,1007

Doğrusal modeller olan DKDO ve BJ arasındaki %8,1'lik tahmin farkı model yapılarının farklılığıyla alakalıdır. DKDOO ile HW modellerinin %0,59'luk tahmin farkı ile yakın sonuçlar verdiği görülmektedir. UKSH modelinin R kare, RMSE, nRMSE ve MAE değerlerine bakarak en iyi performansı gösterdiği ve DKDO ile beraber DVR modellerinin yeterli sonuçlar göstermediği analizi yapılabilir. UKSH en yakın tahmin modeli olan HW modelinden %2,5'luk tahmin farkı ile daha iyi olduğu görülmektedir. DKDO modeli en kötü performansı gösterirken UKSH modeli ile %13,4'lük tahmin farkı olarak iyileştirme sağlanmıştır. Çizelge 5.46'da Gaziantep için doğrulama sonuçları verilmiştir.

Çizelge 5.46: Gaziantep için model doğrulama sonuçları karşılaştırması.

Gaziantep modelleri	R ²	RMSE	nRMSE	MAE
DKDO	0,84485	108,9257	0,092467	59,4418
BJ	0,92322	76,6263	0,065048	41,093
DKDOO	0,92262	76,9244	0,065301	40,9162
HW	0,94333	65,8314	0,055884	34,0973
DVR	0,8289	114,3031	0,0969	67,3028
UKSH	0,9614	54,7078	0,0497	18,3098

Doğrusal modeller olan DKDO ve BJ arasındaki %9,2'lik tahmin farkı Adıyaman ve Diyarbakır doğrulama sonuçları ile aynı model yapısının farklılığıyla alakalıdır. DKDOO ile HW modellerinin yakın sonuçlar verdiği görülmektedir aralarındaki R² olarak %2,2'lik tahmin farkı doğrusalsızlık fonksiyonların kullanımı ile açıklayabiliriz. UKSH modelinin R², RMSE, nRMSE ve MAE değerlerine bakarak en iyi performansı gösterdiği ve DKDO ile beraber DVR modellerinin yeterli sonuçlar göstermediği analizi yapılabilir. UKSH en yakın tahmin modeli olan HW modelinden %2'lik tahmin farkı ile daha iyi olduğu görülmektedir. DVR modeli en kötü performansı gösterirken UKSH modeli ile %15,9'luk tahmin farkında iyileştirme sağlanmıştır. Çizelge 5.47'de Hatay için modellerin doğrulama sonuçları verilmiştir.

Çizelge 5.47: Hatay için model doğrulama sonuçları karşılaştırması.

Hatay modelleri	R ²	RMSE	nRMSE	MAE
DKDO	0,85762	103,695	0,097184	56,8412
BJ	0,9306	72,3972	0,067851	39,0273
DKDOO	0,93089	72,2447	0,067708	40,0399
HW	0,93641	69,2996	0,064948	36,5529
DVR	0,8159	117,9577	0,0996	72,3945
UKSH	0,9763	42,2268	0,0384	15,6975

Doğrusal modeller olan DKDO ve BJ arasındaki %8,5'luk tahmin farkı model yapısının farklılığıyla alakalıdır. DKDOO ile HW modellerinin %0,59'luk tahmin farkı ile yakın sonuçlar verdiği görülmektedir. UKSH modelinin R², RMSE, nRMSE ve MAE değerlerine bakarak en iyi performansı gösterdiği ve DKDO ile beraber DVR modellerinin yeterli sonuçlar göstermediği analizi yapılabilir. UKSH en yakın tahmin modeli olan HW modelinden %4,2'lik tahmin farkı ile daha iyi olduğu görülmektedir. DVR modeli en kötü performansı gösterirken UKSH modeli ile %19,6'luk tahmin farkında iyileştirme sağlanmıştır.

Doğrulama sonuçları genel olarak analiz edildiğinde doğrusal olmayan sistem tanılama yöntemlerinin doğrusal sistem tanılama yöntemlerinden daha iyi performans gösterdiği, makine öğrenmesi yöntemi olan destek vektör regresyonun büyük boyutlu verilerde kötü performans sergilediğini ve uzun kısa süreli hafıza modelinin en iyi performansı gösterdiği görülmektedir.

Güneş radyasyonu tahminlerinde UKSH kullanımı yeterli sonuçlar elde edilmesini sağlayacaktır. Çünkü güneş radyasyonu çıkış değeri aynı zamanda giriş değeri olarak kullanılmaktadır. Güneş radyasyonu sensör verisi olmadığında HW modelinin uygun sonuçlar vereceği analizlerde ortaya çıkmıştır.



KAYNAKLAR

- Alam, S. (2018).** Recurrent neural networks in electricity load forecasting.
- Alcan, G., Unel, M., Aran, V., Yilmaz, M., Gurel, C., & Koprubasi, K. (2019).** Predicting NOx emissions in diesel engines via sigmoid NARX models using a new experiment design for combustion identification.
- Al-Sbou, Y. A., & Alawasa, K. M. (2017).** Nonlinear Autoregressive Recurrent Neural Network Model For Solar Radiation Prediction.
- Alzahrana, A., Shamsi, P., Dagli, C., & Ferdowsi, M. (2017).** Solar Irradiance Forecasting Using Deep Neural Networks.
- Araujo, J. M. (2020).** Performance comparison of solar radiation forecasting between WRF and LSTM in Gifu, Japan.
- Bhola, P., & Bhardwaj, S. (2019).** Estimation of solar radiation using support vector regression.
- Boser, B. E., Guyon, I., & Vapnik, V. N. (1996).** A Training Algorithm for Optimal Margin Classifier.
- Boussaada, Z., Curea, O., Remaci, A., Camblong, H., & Bellaaj, N. M. (2018).** A Nonlinear Autoregressive Exogenous (NARX) Neural Network Model for the Prediction of the Daily Direct Solar Radiation.
- Chang, C.-C., & Lin, C.-J. (2013).** LIBSVM: A Library for Support Vector Machines.
- Chen, J.-L., & Li, G.-S. (2013).** Evaluation of support vector machine for estimation of solar radiation from measured meteorological variables.
- Chica, A., Rey, F., & Aristizábal, J. (2012).** Application of autoregressive model with exogenous inputs to identify and analyse patterns of solar global radiation and ambient temperature.

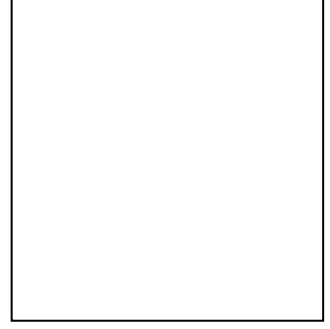
- Chow, W. C., Urquhart, B., Lave, M., Dominguez, A., Kleiss, J., Shields, J., & Washom, B. (2011).** Intra-hour forecasting with a total sky imager at the UC San Diego solar energy testbed.
- Drucker, H., Kaufman, L., Smola, A., Vapnik, V., & Chris, J. C. (2003).** Support Vector Regression Machines.
- Durak, M. (2016).** GÜNEŞ ENERJİSİ: TEORİ VE UYGULAMA.
- Gers, F. A., Schmidhuber, J., & Cummins, F. (2000).** Learning to Forget: Continual Prediction with LSTM.
- Hochreiter, S., & Schmidhuber, J. (1997).** LONG-SHORT TERM MEMORY.
- Kamil, M. S., & Razali, A. M. (2015).** Time Series SARIMA Models for Average Monthly Solar Radiation in Malaysia.
- Khalfi, J., Boumaaz, N., Soulmani, A., & Laadissi, E. (2020).** Nonlinear Modeling of Lithium-Ion Battery Cells for Electric Vehicles using a Hammerstein–Wiener Model.
- Langella, R., Proto, D., & Testa, A. (2016).** Solar Radiation Forecasting, Accounting for Daily Variability.
- Lave, M., & Kleissl, J. (2010).** Solar variability of four sites across the state of Colorado.
- Ljung, L. (2008a).** System Identification Toolbox™ 7 Getting Started Guide. The MathWorks Inc.
- Ljung, L. (2008b).** System Identification Toolbox™ 7 User's Guide. The MathWorks Inc.
- Ljung, L. (2010).** Perspectives on system identification.
- Lysfjord, M. J. (2017).** Modeling and Forecasting the Nord Pool Day-Ahead Power Market through Deep-Learning.
- Mohanty, S., Patra, P. K., & Sahoo, S. S. (2015).** Prediction of Global Solar Radiation using Nonlinear auto Regressive Network with Exogenous inputs.

- Öztürk, H. (2014).** GÜNEŞ ENERJİSİNDEN FOTOVOLTAİK YÖNTEMLE ELEKTRİK ÜRETİMİNDE GÜÇ DÖNÜŞÜM VERİMİ VE ETKİLİ ETMENLER.
- Sak, H., Senior, A., & Beaufays, F. (2014).** Long Short-Term Memory Recurrent Neural Network Architectures for Large Scale Acoustic Modeling.
- Sarı, A., Aktaş, M., Yönetken, A., & Doğan, R. (2021).** Güneş Işınım Tahmininde NARX Modelinin Uygulanması.
- SFETSOS, A., & COONICK, A. H. (2000).** UNIVARIATE AND MULTIVARIATE FORECASTING OF HOURLY SOLAR RADIATION WITH ARTIFICIAL INTELLIGENCE TECHNIQUES.
- Sharifzadeh, M., Sikinioti-Lock, A., & Shah, N. (2019).** Machine-learning methods for integrated renewable power generation: A comparative study of artificial neural networks, support vector regression and Gaussian Process Regression.
- Shaw, S., & Prakash, M. (2019).** Solar Radiation Forecasting Using Support Vector Regression.
- SMOLA, A. J., & SCHOLKOPF, B. (2003).** A tutorial on support vector regression.
- SORKUN, M. C., DURMAZ İNCEL, Ö., & PAOLI, C. (2020).** Time series forecasting on multivariate solar radiation data using deep learning (LSTM).
- Taneja, K., Ahmad, S., Ahmad, K., & Attri, S. D. (2016).** Time series analysis of aerosol optical depth over New Delhi using Box-Jenkins ARIMA modeling approach.
- Tuhta, S., Günday, F., & Aydin, H. (2020).** NONLINEAR SYSTEM IDENTIFICATION OF MODEL CONCRETE CHIMNEY USING HAMMERSTEIN-WIENER MODELS.
- Vapnik, V. N. (2000).** The Nature of Statistical Learning Theory. Springer.
- Wei, C.-C. (2017).** Predictions of Surface Solar Radiation on Tilted Solar Panels using Machine Learning Models: A Case Study of Tainan City, Taiwan.
- YEGM. (2015, 03 17).** http://www.eie.gov.tr/yenilenebilir/g_enj_tekno.aspx adresinden alındı



ÖZGEÇMİŞ

Ad-Soyad : Göksel Kader



ÖĞRENİM DURUMU:

- **Lisans** : 2013, İstanbul Teknik Üniversitesi, Elektrik-Elektronik Fakültesi, Kontrol Mühendisliği bölümü

MESLEKİ DENEYİM VE ÖDÜLLER:

- 2013-2014 yılları arasında SAFER A.Ş firmasında Saha Mühendisi olarak görev aldım.
- 2014-2017 yılları arasında Honeywell Sanayi A.Ş firmasında Proje ve Yazılım Mühendisi olarak görev aldım.
- 2017 yılında Ford Otomotiv Sanayi A.Ş firmasında Güç Aktarma Sistemleri Yazılım Mühendisi olarak çalışmaya başladım ve halen devam etmekteyim.