

T.C.
FIRAT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ



**DÜŞÜK HESAPLAMA MALİYETİNE SAHİP MİYOPİK
OLMAYAN ADAPTİF ÖRNEKLEME VE ÇEVRESEL İZLEME
İÇİN MONTE CARLO ARAMA AĞACI VE DALLANMA
SINIRLAMA YÖNTEMLERİNİN UYGULANMASI**

Perihan KARAKÖSE

Doktora Tezi

MEKATRONİK MÜHENDİSLİĞİ ANABİLİM DALI

Mekanik Sistemler Bilim Dalı

TEMMUZ 2024

T.C.
FIRAT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

Mekatronik Mühendisliği Anabilim Dalı

Doktora Tezi

**DÜŞÜK HESAPLAMA MALİYETİNE SAHİP MİYOPİK OLMAYAN
ADAPTİF ÖRNEKLEME VE ÇEVRESEL İZLEME İÇİN MONTE
CARLO ARAMA AĞACI VE DALLANMA SINIRLAMA
YÖNTEMLERİNİN UYGULANMASI**

Tez Yazarı

Perihan KARAKÖSE

Danışman

Doç. Dr. Cafer BAL

İkinci Danışman

Dr. Öğr. Üyesi Harun YETKİN

TEMMUZ 2024

ELAZIĞ

T.C.
FIRAT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

Mekatronik Mühendisliği Anabilim Dalı

Doktora Tezi

Başlığı: Düşük Hesaplama Maliyetine Sahip Miyopik Olmayan Adaptif Örnekleme ve Çevresel İzleme için Monte Carlo Arama Ağacı ve Dallanma Sınırlama Yöntemlerinin Uygulanması

Yazarı: Perihan KARAKÖSE

İlk Teslim Tarihi: 11.06.2024

Savunma Tarihi: 18.07.2024

TEZ ONAYI

Fırat Üniversitesi Fen Bilimleri Enstitüsü tez yazım kurallarına göre hazırlanan bu tez aşağıda imzaları bulunan jüri üyeleri tarafından değerlendirilmiş ve akademik dinleyicilere açık yapılan savunma sonucunda OYBİRLİĞİ ile kabul edilmiştir.

Danışman:	Doç. Dr. Cafer BAL Fırat Üniversitesi, Teknoloji Fakültesi	<i>İmza</i> Onayladım
Başkan:	Prof. Dr. Mehmet ÇAVAŞ Fırat Üniversitesi, Teknoloji Fakültesi	Onayladım
Üye:	Doç. Dr. Deniz KORKMAZ Turgut Özal Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi	Onayladım
Üye:	Doç. Dr. Gürkan KAVURAN Turgut Özal Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi	Onayladım
Üye:	Doç. Dr. Orhan YAMAN Fırat Üniversitesi, Teknoloji Fakültesi	Onayladım

Bu tez, Enstitü Yönetim Kurulunun/...../20..... tarihli toplantısında tescillenmiştir.

İmza

Prof. Dr. Burhan ERGEN
Enstitü Müdürü

BEYAN

Fırat Üniversitesi Fen Bilimleri Enstitüsü tez yazım kurallarına uygun olarak hazırladığım “Düşük Hesaplama Maliyetine Sahip Miyopik Olmayan Adaptif Örnekleme ve Çevresel İzleme için Monte Carlo Arama Ağacı ve Dalların Sınırlama Yöntemlerinin Uygulanması” Başlıklı Doktora Tezimin içindeki bütün bilgilerin doğru olduğunu, bilgilerin üretilmesi ve sunulmasında bilimsel etik kurallarına uygun davrandığımı, kullandığım bütün kaynakları atıf yaparak belirttiğimi, maddi ve manevi desteği olan tüm kurum/kuruluş ve kişileri belirttiğimi, burada sunduğum veri ve bilgileri unvan almak amacıyla daha önce hiçbir şekilde kullanmadığımı beyan ederim.

18.07.2024

Perihan KARAKÖSE



ÖNSÖZ

Çevresel izleme, çevrenin kalitesini karakterize etmek ve izlemek için gerçekleşmesi gereken süreçleri ve faaliyetleri tanımlar. Bu tez çalışması gerçek zamanlı çevresel izleme yapmak için bilgilendirici güzergah planlama algoritmasında karşılaşılan sorunlara çözüm önermektedir. Elde edilen sonuçlar karşılaştırmalı olarak kıyaslanmıştır.

Bu tez çalışması süresince yardımlarını esirgemeyen ve değerli fikirleriyle bana yol gösteren, her daim yardımcı olan danışman hocam, Sayın Doç. Dr. Cafer BAL' a çok teşekkür ederim. Tez çalışmamın başından sonuna kadar çok desteğini gördüğüm, her konuda desteğini hiçbir zaman esirgemeyen ikinci danışmanım Dr. Öğr. Üyesi Harun YETKİN' e teşekkürü bir borç bilirim. Son olarak tüm hayatım boyunca yanımda olmaya çalışan başta anne ve babam olmak üzere yanımda olan tüm sevdiğlerime teşekkür ederim.

Bu tez çalışması, Yüksek Öğretim Kurulu'nun (YÖK) 100/2000 doktora projesi programı "İleri Robotik Sistemler ve Mekatronik" öncelikli alanı ve Türkiye Bilimsel ve Teknik Araştırma Kurumu (TÜBİTAK) 2211-A genel doktora burs programı kapsamında desteklenmiştir.

Perihan KARAKÖSE
ELAZIĞ, 2024

İÇİNDEKİLER

	Sayfa
ÖNSÖZ.....	iv
İÇİNDEKİLER	v
ÖZET	vii
ABSTRACT	viii
ŞEKİLLER LİSTESİ	ix
TABLolar LİSTESİ	xiii
SİMGELER	xiv
KISALTMALAR	xv
1. GİRİŞ	1
2. MATERYAL VE METOT	13
2.1. Makine Öğrenmesi Yöntemleri	13
2.1.1. Gauss Süreçleri ile Ortamın Olasılıksal Modelinin Oluşturulması	15
2.1.2. Çok Çıktılı Gauss Süreci	16
2.1.3. Seyrek Gauss Süreci ile Makine Öğrenimi	16
2.1.4. Gauss Süreçlerinde Hiperparametre Optimizasyonu.....	17
2.2. Eğitim Verilerinin Konumlarının Belirlenmesi	17
2.2.1. Hiper parametreleri Öğrenmede Eğitim Verisinin Rolü.....	18
2.2.2. Uzamsal Alanda Aday Yol Noktalarının Dizilimi	20
2.2.3. Eğitim Verisi Toplamak İçin Geliştirilen Yeni Bir Güzergah Planlama Algoritması	21
2.3. Keşif ve Sömürü Dengesinde Kullanılan Fayda Fonksiyonlarının Önemi	24
2.3.1. Keşif ve Sömürü Dengesinde Yaygın Olarak Kullanılan Ödül Fonksiyonları	25
2.3.2. Ödül Fonksiyonları Kullanılarak Yeni Örnek Noktasının Seçimi	26
2.3.3. ROI Alanlarından Örnek Toplamak için Geliştirilen Güzergah Planlama Algoritması	29
2.3.4. Çapraz Doğrulama Yöntemi.....	30
2.4. Adaptif Yörünge Planlama Algoritmaları.....	30
2.4.1. Monte Carlo Arama Ağacı ile Adaptif Örnekleme	31
2.4.2. Yorucu Arama Algoritması ile Adaptif Örnekleme	34
2.4.3. Önerilen Hibrit bir Model Olan MCTS-BnB Algoritması ile Adaptif Örnekleme.....	36
2.5. Seyrek Sözcük Girişlerin Belirlenmesi için Kullanılan Kümeleme Yöntemleri.....	41
2.5.1. K-Means Kümeleme Algoritması	42
2.5.2. Agglomerative Hierarchical	42
2.5.3. Birch.....	43
2.5.4. MiniBatch K-Means	43
2.5.5. Spektral Kümeleme	44
2.6. Siluet Skor İndeksi (SSI).....	44
3. BULGULAR VE TARTIŞMALAR	45
3.1. Eğitim Verilerini Belirlemek Geliştirilen Güzergah Planlama Algoritması	45
3.1.1. Hiper parametreleri Öğrenmek için Kullanılan Veri Seti.....	46
3.1.2. Eğitim Verilerini Toplamak İçin Önerilen Güzergah Planlama Algoritmasıyla Elde Edilen Sonuçlar.....	47
3.2. ROI Alanlarından Örnek Toplamak için Geliştirilen Güzergah Planlama Algoritması	51
3.2.1. ROI Alanlarını Belirlemek için Oluşturulan Test Ortamı	51

3.2.2. ROI Alanlarında Adaptif Örnekleme	55
3.2.3. Keşif Verimliliğinde Varyans Etkisi	73
3.2.4. Tahmin performansını değerlendirmek için, RMSE'deki değişimin incelenmesi	94
3.3. Gerçek Zamanlı Denizdeki Klorofil Miktarını Belirlemek için Geliştirilen Hibrit MCTS-BnB Algoritması	102
3.3.1. Saildrone SSS ve SST Veri Seti	102
3.3.2. Klorofil-a parametresini bir özellik seçimi yöntemi kullanarak modelleme	103
3.3.3. MCTS-BnB Algoritmasının Başarımının Değerlendirilmesi	111
3.3.4. MCTS-BnB Algoritmasının Başarımının Farklı Veri Setlerinde Değerlendirilmesi.....	114
3.3.5. Seyrek Sözde Girişleri Kullanarak MCTS-BnB Algoritması ile Adaptif Örnekleme	123
4. SONUÇLAR.....	126
KAYNAKLAR.....	129
ÖZGEÇMİŞ	



ÖZET

Düşük Hesaplama Maliyetine Sahip Miyopik Olmayan Adaptif Örnekleme ve Çevresel İzleme için Monte Carlo Arama Ağacı ve Dallanma Sınırlama Yöntemlerinin Uygulanması

Perihan KARAKÖSE

Doktora Tezi

FIRAT ÜNİVERSİTESİ
Fen Bilimleri Enstitüsü

Mekatronik Mühendisliği Anabilim Dalı

Temmuz 2024, Sayfa: xv + 135

Çevresel izleme, doğal kaynakların sürdürülebilir kullanımı ve ekosistem sağlığının korunması açısından son derece önemlidir. Bu tez çalışmasında, su parametrelerinin gerçek zamanlı izlenmesi için Gauss süreçleri (GP) yöntemi kullanılarak gerçekleştirilen çalışmalar incelenmiştir. Tez, dört farklı aşamada ele alınmıştır. Birinci aşamada, eğitim verilerinin konumlarının, GP algoritmasının tahmin performansı üzerinde kritik bir rol oynadığı belirlenmiş ve sınırlı bir bütçeyle tüm alandan örnek toplamak için yeni bir güzergah planlama algoritması geliştirilmiştir. Bu algoritma, 10 farklı veri kümesinde test edilmiş ve literatürde mevcut diğer yaklaşımlarla karşılaştırıldığında üstün performans sergilediği görülmüştür. İkinci aşamada, yaygın olarak kullanılan ödül fonksiyonları (Entropi, UCB ve Seviye Belirleme) kullanılarak İlgi Alanı (ROI) bölgeleri tespit edilmiştir. Bu fonksiyonların keşif ve sömürü dengelerindeki başarıları değerlendirilmiştir. Entropi ödül fonksiyonunun, karmaşık olmayan veri setlerinde belirsizliği en aza indirerek ROI alanlarını daha hızlı keşfettiği belirlenmiştir. Seviye Belirleme ödül fonksiyonunun ise kompleks veri setlerinde daha başarılı sonuçlar verdiği görülmüştür. UCB ödül fonksiyonunun, yüksek belirsizlik içeren haritalarda daha başarılı olduğu ancak düşük belirsizlik durumlarında yerel maksimumlara takılma riskinin bulunduğu tespit edilmiştir. Üçüncü aşamada, gezgin satıcı problemini çözmek amacıyla Monte Carlo arama ağacı (MCTS) ve Dallan ve Sınırla Tekniği (BnB) bir araya getirilmiştir. BnB algoritmasında alt sınır, MCTS algoritması kullanılarak belirlenmiş, bu sayede optimal rota belirleme süresi önemli ölçüde kısaltılmıştır. Algoritma, gerçek bir veri seti üzerinde test edilmiş ve MCTS ile yorucu arama ve adaptif örnekleme algoritmaları (ESAS) karşısında üstün performans sergilemiştir. Bu algoritma, üç farklı hiperparametreye sahip veri setlerinde de başarılı sonuçlar elde etmiştir. Ayrıca, önerilen algoritma literatürdeki farklı alt sınır ve üst sınır belirleme yöntemleriyle karşılaştırılmıştır. Bu karşılaştırmalar sonucunda hesaplama süresinin en kısaya yakın çıkarken, doğruluğun da en düşük RMSE oranına yaklaştığı gözlemlenmiştir. Son aşamada ise MCTS-BnB algoritmasının hesaplama süresini azaltmak amacıyla seyrek temsili girişler kullanılmış ve bu girişler 5 farklı kümeleme yöntemi ile belirlenmiştir. K-Means kümeleme algoritması, en iyi kümeleme performansını göstererek hesaplama süresini önemli ölçüde azaltmıştır. Tez çalışmasının genel bulguları, Gauss süreçleri ve ödül fonksiyonları kullanılarak su parametrelerinin etkili bir şekilde gerçek zamanlı olarak izlenebileceğini göstermektedir. Ayrıca, MCTS ve BnB yöntemlerinin birleştirilmesi, güzergah planlamada önemli iyileştirmelere yol açarak çevresel izlemeye yönelik daha optimal çözümler sunabileceğini ortaya koymaktadır.

Anahtar Kelimeler: Gauss Süreçleri, Çevresel İzleme, Dallan ve Sınırla, Monte Carlo Arama Ağacı, Güzergah Planlama

ABSTRACT

Application of Monte Carlo Tree Search and Branch-and-Bound Methods for Non-Myopic Adaptive Sampling and Environmental Monitoring with Low Computational Cost

Perihan KARAKÖSE

Ph.D. Thesis

FIRAT UNIVERSITY
Graduate School of Natural and Applied Sciences
Department of Mechatronics Engineering

July 2024, Pages: xv + 135

Environmental monitoring is of critical importance for the sustainable use of natural resources and the preservation of ecosystem health. This thesis examines studies conducted using Gaussian Processes (GP) methods for real-time monitoring of water parameters. The thesis is structured into four distinct phases. In the first phase, the importance of the locations of training data is emphasized. It has been determined that the locations of training data play a crucial role in the prediction performance of the GP algorithm. To address this, a new path planning algorithm was developed to collect samples from the entire area within a limited budget. This algorithm was tested on 10 different datasets and was found to outperform existing approaches in the literature. In the second phase, commonly used reward functions (Entropy, UCB, and Level Setting) were employed to identify Regions of Interest (ROI). The success of these functions in balancing exploration and exploitation was evaluated. It was found that the Entropy reward function minimized uncertainty more effectively in non-complex datasets, allowing for faster discovery of ROI areas. The Level Setting reward function yielded better results in more complex datasets, while the UCB reward function performed well in high-uncertainty maps but risked getting stuck in local maxima in low-uncertainty situations. The third phase involved combining the Monte Carlo Tree Search (MCTS) and Branch and Bound (BnB) techniques to solve the Traveling Salesman Problem. In the BnB algorithm, the lower bound was determined using the MCTS algorithm, significantly reducing the time required to identify the optimal route. The algorithm was tested on real datasets and demonstrated superior performance compared to MCTS and exhaustive search and adaptive sampling algorithms (ESAS). This algorithm was also successful across datasets with three different hyperparameters. Additionally, the proposed algorithm was compared with different lower and upper bound determination methods in the literature. The comparison revealed that the calculation time was close to the shortest while achieving an accuracy near the lowest RMSE rate. In the final phase, sparse representative inputs were used to reduce the computation time of the MCTS-BnB algorithm, with these inputs determined using five different clustering methods. The K-Means clustering algorithm showed the best clustering performance, significantly reducing the computation time. The overall findings of the thesis demonstrate that water parameters can be effectively monitored in real-time using Gaussian Processes and reward functions. Moreover, combining the MCTS and BnB methods leads to significant improvements in path planning, offering more optimal solutions for environmental monitoring.

Keywords: Gaussian Processes, Environmental Monitoring, Branch and Bound, Monte Carlo Tree Search, Path Planning

ŞEKİLLER LİSTESİ

Sayfa

Şekil 2.1. Makine Öğrenmesi Yöntemleri	14
Şekil 2.2. 4 farklı rastgele senaryo için hiper parametre değişimi	20
Şekil 2.3. Yol Noktası Dizilimleri a) Halton Dizilimi b) Sobol Dizilimi c) Düzenli Dizilim d) Rastgele Dizilim	21
Şekil 2.4. 10 farklı veri setinde softmax fonksiyonuna bağlı RMSE değişimi	23
Şekil 2.5. Yeni Örnekleme Noktalarının Belirlenmesi Oluşturulan Akış Diyagramı	27
Şekil 2.6. Ödül fonksiyonları kullanılarak 10 örnek noktanın seçimi	28
Şekil 2.7. Ödül fonksiyonları kullanılarak 10 örnek noktanın seçimiyle varyans haritasının değişimi	29
Şekil 2.8. MCTS Algoritması	32
Şekil 2.9. Örnek alt limit ve üst limit hesaplama şeması	40
Şekil 2.10. MCTS-BnB Algoritmasında hesaplanan alt limit ve üst limit değerleri a) Alt limit b) Üst limit	41
Şekil 3.1. Eğitim verilerinin önerilen yöntemle öğrenilmesi için belirlenen 10 farklı bölge	46
Şekil 3.2. Nasa Veri Setinden seçilen bölgelerdeki veri haritaları	47
Şekil 3.3. 1100 Km'de Oluşturulan Yörüngeler a) Halton Dizilimi b) Sobol Dizilimi c) Düzenli Dizilim d) Rastgele Dizilim e) Pilot f) Pilot ₂ g) Çim Biçme Rotası h) Rastgele Seçilen Yol Noktalarıyla Oluşturulan Güzergah	49
Şekil 3.4. Seçilen Yörüngelerdeki RMSE değişimi	50
Şekil 3.5. Gauss eğrisi	51
Şekil 3.6. a) Tek ROI alanı olan test ortamı b) Başlangıçta Varyans Haritası c) Çapraz Doğrulama yöntemi uygulanan varyans haritası	52
Şekil 3.7. a) İki ROI Gerçek Haritası b) Başlangıçta Varyans Haritası c) Çapraz Doğrulama yöntemi uygulanan varyans haritası	53
Şekil 3.8. a) Üç ROI Gerçek Haritası b) Başlangıçta Varyans Haritası c) Çapraz Doğrulama yöntemi uygulanan varyans haritası	54
Şekil 3.9. a) Dört ROI Gerçek Haritası b) Başlangıçta Varyans Haritası c) Çapraz Doğrulama yöntemi uygulanan varyans haritası	55
Şekil 3.10. Tek ROI için adaptif örnekleme a) <i>LPI</i> ödül fonksiyonu kullanılarak toplanan 150 örnek b) <i>LPI</i> ödül fonksiyonu kullanılarak toplanan 300 örnek c) a) <i>LLevelSet</i> ödül fonksiyonu kullanılarak toplanan 150 örnek d) <i>LLevelSet</i> ödül fonksiyonu kullanılarak toplanan 300 örnek e) a) <i>LUCB</i> ödül fonksiyonu kullanılarak toplanan 150 örnek f) a) <i>LUCB</i> ödül fonksiyonu kullanılarak toplanan 300 örnek	57
Şekil 3.11. Çapraz doğrulama yöntemi uygulanan Tek ROI alanında adaptif örnekleme a) <i>LPI</i> ödül fonksiyonu kullanılarak toplanan 150 örnek b) <i>LPI</i> ödül fonksiyonu kullanılarak toplanan 300 örnek c) a) <i>LLevelSet</i> ödül fonksiyonu kullanılarak toplanan 150 örnek d) <i>LLevelSet</i> ödül	

- fonksiyonu kullanılarak toplanan 300 örnek e) a) *LUCB* ödül fonksiyonu kullanılarak toplanan 150 örnek f) a) *LUCB* ödül fonksiyonu kullanılarak toplanan 300 örnek..... 60
- Şekil 3.12.** İki ROI için adaptif örnekleme a) *LPI* ödül fonksiyonu kullanılarak toplanan 150 örnek b) *LPI* ödül fonksiyonu kullanılarak toplanan 300 örnek c) a) *LLevelSet* ödül fonksiyonu kullanılarak toplanan 150 örnek d) *LLevelSet* ödül fonksiyonu kullanılarak toplanan 300 örnek e) a) *LUCB* ödül fonksiyonu kullanılarak toplanan 150 örnek f) a) *LUCB* ödül fonksiyonu kullanılarak toplanan 300 örnek..... 62
- Şekil 3.13.** Çapraz doğrulama yöntemi uygulanan İki ROI'li alanda adaptif örnekleme a) *LPI* ödül fonksiyonu kullanılarak toplanan 150 örnek b) *LPI* ödül fonksiyonu kullanılarak toplanan 300 örnek c) a) *LLevelSet* ödül fonksiyonu kullanılarak toplanan 150 örnek d) *LLevelSet* ödül fonksiyonu kullanılarak toplanan 300 örnek e) a) *LUCB* ödül fonksiyonu kullanılarak toplanan 150 örnek f) a) *LUCB* ödül fonksiyonu kullanılarak toplanan 300 örnek..... 64
- Şekil 3.14.** Üç ROI için adaptif örnekleme a) *LPI* ödül fonksiyonu kullanılarak toplanan 150 örnek b) *LPI* ödül fonksiyonu kullanılarak toplanan 300 örnek c) a) *LLevelSet* ödül fonksiyonu kullanılarak toplanan 150 örnek d) *LLevelSet* ödül fonksiyonu kullanılarak toplanan 300 örnek e) a) *LUCB* ödül fonksiyonu kullanılarak toplanan 150 örnek f) a) *LUCB* ödül fonksiyonu kullanılarak toplanan 300 örnek..... 66
- Şekil 3.15.** Çapraz doğrulama yöntemi uygulanan üç ROI alanında adaptif örnekleme a) *LPI* ödül fonksiyonu kullanılarak toplanan 150 örnek b) *LPI* ödül fonksiyonu kullanılarak toplanan 300 örnek c) a) *LLevelSet* ödül fonksiyonu kullanılarak toplanan 150 örnek d) *LLevelSet* ödül fonksiyonu kullanılarak toplanan 300 örnek e) a) *LUCB* ödül fonksiyonu kullanılarak toplanan 150 örnek f) a) *LUCB* ödül fonksiyonu kullanılarak toplanan 300 örnek..... 68
- Şekil 3.16.** Dört ROI için adaptif örnekleme a) *LPI* ödül fonksiyonu kullanılarak toplanan 150 örnek b) *LPI* ödül fonksiyonu kullanılarak toplanan 300 örnek c) a) *LLevelSet* ödül fonksiyonu kullanılarak toplanan 150 örnek d) *LLevelSet* ödül fonksiyonu kullanılarak toplanan 300 örnek e) a) *LUCB* ödül fonksiyonu kullanılarak toplanan 150 örnek f) a) *LUCB* ödül fonksiyonu kullanılarak toplanan 300 örnek..... 70
- Şekil 3.17.** Çapraz doğrulama yöntemi uygulanan dört ROI alanı için adaptif örnekleme a) *LPI* ödül fonksiyonu kullanılarak toplanan 150 örnek b) *LPI* ödül fonksiyonu kullanılarak toplanan 300 örnek c) a) *LLevelSet* ödül fonksiyonu kullanılarak toplanan 150 örnek d) *LLevelSet* ödül fonksiyonu kullanılarak toplanan 300 örnek e) a) *LUCB* ödül fonksiyonu kullanılarak toplanan 150 örnek f) a) *LUCB* ödül fonksiyonu kullanılarak toplanan 300 örnek..... 72
- Şekil 3.18.** Tek ROI için Varyans Haritası: a) $L_{entropi}$ ödül fonksiyonu kullanılarak toplanan 150 örnek b) $L_{entropi}$ ödül fonksiyonu kullanılarak toplanan 300 örnek c) L_{LS} ödül fonksiyonu kullanılarak toplanan 150 örnek d) $L_{LevelSet}$ ödül fonksiyonu kullanılarak toplanan 300 örnek e) L_{UCB} ödül fonksiyonu kullanılarak toplanan 150 örnek f) L_{UCB} ödül fonksiyonu kullanılarak toplanan 300 örnek..... 74
- Şekil 3.19.** Çapraz doğrulama yöntemi uygulanan tek ROI alanında varyans haritası: a) $L_{entropi}$ ödül fonksiyonu kullanılarak toplanan 150 örnek b) $L_{entropi}$ ödül fonksiyonu kullanılarak toplanan 300 örnek c) L_{LS} ödül fonksiyonu kullanılarak toplanan 150 örnek d) L_{LS} ödül fonksiyonu kullanılarak toplanan 300 örnek e) L_{UCB} ödül fonksiyonu kullanılarak toplanan 150 örnek f) L_{UCB} ödül fonksiyonu kullanılarak toplanan 300 örnek 76

Şekil 3.20.	İki ROI için Varyans Haritası: a) $L_{entropi}$ ödül fonksiyonu kullanılarak toplanan 150 örnek b) $L_{entropi}$ ödül fonksiyonu kullanılarak toplanan 300 örnek c) L_{LS} ödül fonksiyonu kullanılarak toplanan 150 örnek d) L_{LS} ödül fonksiyonu kullanılarak toplanan 300 örnek e) L_{UCB} ödül fonksiyonu kullanılarak toplanan 150 örnek f) L_{UCB} ödül fonksiyonu kullanılarak toplanan 300 örnek	78
Şekil 3.21.	Çapraz doğrulama yöntemi uygulanann iki ROI alanı için varyans Haritası: a) $L_{entropi}$ ödül fonksiyonu kullanılarak toplanan 150 örnek b) $L_{entropi}$ ödül fonksiyonu kullanılarak toplanan 300 örnek c) L_{LS} ödül fonksiyonu kullanılarak toplanan 150 örnek d) L_{LS} ödül fonksiyonu kullanılarak toplanan 300 örnek e) L_{UCB} ödül fonksiyonu kullanılarak toplanan 150 örnek f) L_{UCB} ödül fonksiyonu kullanılarak toplanan 300 örnek	80
Şekil 3.22.	Üç ROI için varyans haritası: a) $L_{entropi}$ ödül fonksiyonu kullanılarak toplanan 150 örnek b) $L_{entropi}$ ödül fonksiyonu kullanılarak toplanan 300 örnek c) L_{LS} ödül fonksiyonu kullanılarak toplanan 150 örnek d) L_{LS} ödül fonksiyonu kullanılarak toplanan 300 örnek e) L_{UCB} ödül fonksiyonu kullanılarak toplanan 150 örnek f) L_{UCB} ödül fonksiyonu kullanılarak toplanan 300 örnek	81
Şekil 3.23.	Çapraz doğrulama yöntemi uygulanan üç ROI alanı için varyans haritası: a) $L_{entropi}$ ödül fonksiyonu kullanılarak toplanan 150 örnek b) $L_{entropi}$ ödül fonksiyonu kullanılarak toplanan 300 örnek c) L_{LS} ödül fonksiyonu kullanılarak toplanan 150 örnek d) L_{LS} ödül fonksiyonu kullanılarak toplanan 300 örnek e) L_{UCB} ödül fonksiyonu kullanılarak toplanan 150 örnek f) L_{UCB} ödül fonksiyonu kullanılarak toplanan 300 örnek	83
Şekil 3.24.	Dört ROI için varyans haritası: a) $L_{entropi}$ ödül fonksiyonu kullanılarak toplanan 150 örnek b) $L_{entropi}$ ödül fonksiyonu kullanılarak toplanan 300 örnek c) L_{LS} ödül fonksiyonu kullanılarak toplanan 150 örnek d) L_{LS} ödül fonksiyonu kullanılarak toplanan 300 örnek e) L_{UCB} ödül fonksiyonu kullanılarak toplanan 150 örnek f) L_{UCB} ödül fonksiyonu kullanılarak toplanan 300 örnek	85
Şekil 3.25.	Çapraz doğrulama yöntemi uygulanan dört ROI alanı için varyans haritası: a) $L_{entropi}$ ödül fonksiyonu kullanılarak toplanan 150 örnek b) $L_{entropi}$ ödül fonksiyonu kullanılarak toplanan 300 örnek c) L_{LS} ödül fonksiyonu kullanılarak toplanan 150 örnek d) L_{LS} ödül fonksiyonu kullanılarak toplanan 300 örnek e) L_{UCB} ödül fonksiyonu kullanılarak toplanan 150 örnek f) L_{UCB} ödül fonksiyonu kullanılarak toplanan 300 örnek	87
Şekil 3.26.	Farklı ROI sayıları için elde edilen veri setlerinde ortalama varyans a) 1 ROI alanı olan veri seti b) 2 ROI alanı olan veri seti c) 3 ROI alanı olan veri seti d) 4 ROI alanı olan veri seti	90
Şekil 3.27.	Farklı ROI sayıları için elde edilen veri setlerinde tahmin edilen varyansın standart sapması a) 1 ROI alanı olan veri seti b) 2 ROI alanı olan veri seti c) 3 ROI alanı olan veri seti d) 4 ROI alanı olan veri seti	93
Şekil 3.28.	Farklı ROI sayıları için elde edilen veri setlerinde RMSE değeri a) 1 ROI alanı olan veri seti b) 2 ROI alanı olan veri seti c) 3 ROI alanı olan veri seti d) 4 ROI alanı olan veri seti	97
Şekil 3.29.	Veri Setinin Örnekleme Güzergahı	103
Şekil 3.30.	Denizdeki klorofil miktarını etkileyen parametreler	105
Şekil 3.31.	Su Parametrelerinin Çoklu Çıktılı Gauss Regresyon Algoritması Kullanılarak Tahmini Tuzluluk, b) Su Sıcaklığı, c) İletkenlik, d) Rüzgar Hızı	108
Şekil 3.32.	Çoklu Çıktılı Gauss Regresyon Kullanılarak Tahmin Edilen Varyans Haritası	109

Şekil 3.33. a) Tahmin Edilen Klorofil Haritası b) Tahmin Edilen Varyans Haritası.....	111
Şekil 3.34. Tahmin Edilen Klorofil Haritası a) MCTS b) MCTS-BnB c) ESAS	113
Şekil 3.35. Farklı hiper parametrelere sahip 3 veri setinde RMSE değerinin değişimi a) $\ell = 1$ b) $\ell = 3$ c) $\ell = 9$	116
Şekil 3.36. Toplam Entropi Değişimi	118
Şekil 3.37. Zamana bağlı toplanan örnek miktarı.....	120
Şekil 3.38. MCTS-BnB algoritmasının literatürdeki diğer çalışmalarla zaman performansı açısından karşılaştırılması.....	122
Şekil 3.39. MCTS-BnB algoritmasının literatürdeki diğer çalışmalarla RMSE performansı açısından karşılaştırılması.....	122
Şekil 3.40. Alt örneklem kümesinin RMSE değeri değişimi.....	124
Şekil 3.41. Alt örneklem kümesinin entropi değeri değişimi	124
Şekil 3.42. Alt örneklem kümesinin zamanın değişimi	125

TABLULAR LİSTESİ

	Sayfa
Tablo 1.1. Literatür Taraması.....	9
Tablo 3.1. Ödül Fonksiyonlarının Karşılaştırılması.....	100
Tablo 3.2. Çapraz doğrulama yöntemi uygulanan ödül fonksiyonlarının karşılaştırılması.....	101
Tablo 3.3. Literatürdeki çalışmayla UCB ödül fonksiyonunun performans metriklerinin karşılaştırılması.....	101
Tablo 3.4. Pearson Korelasyon Katsayısı.....	107
Tablo 3.5. Örnek miktarına bağlı harcananan zaman (dk)	121
Tablo 3.6. MCTS-BnB algoritmasının literatürdeki diğer BnB yöntemleriyle karşılaştırılması	123

SİMGELER

$f(x)$: Ölçüm verileri
x	: Giriş
y	: Çıkış
ϵ	: Sensör ölçümlerinden kaynaklı hata oranı
w	: Ağırlık Oranı
σ_n^2	: Varyans
\mathcal{N}	: Normal Dağılım
I	: Birim Matris
$p(y x, w)$: Olasılık Dağılımı
$p(w y, X)$: Sonsal Dağılım
$p(w)$: Önsel Dağılım
$p(y X)$: Marjinal Olasılık
\bar{f}_*	: Tahmin edilen ortalama değer
$\mathbb{V}[f_*]$: Tahmin edilen varyans değeri
K_M	: Bütünleşik kovaryans fonksiyonu
$K(X, X)$: Kovaryans fonksiyonu
k_{SE}	: Kareli Üstel Kovaryans Fonksiyonu
σ^2, l	: Hiper parametreler
k_{RQ}	: Rasyonel İkinci Dereceden Kovaryans Fonksiyonu
k_{per}	: Periyodik kovaryans fonksiyonu
k_{Matern}	: Matern kovaryans fonksiyonu
h	: Eşik değeri
$R(i)$: Beklenen risk
s_v	: Çocuk düğümün ziyaret edilme sayısı
N_v	: Mevcut düğümün ziyaret edilme sayısı
C	: Keşif ve sömürü katsayısı
H	: Robotun seyahat bütçesi
n	: Örnek sayısı
x_0	: Başlangıç konumu
y_0	: Başlangıç konumundan alınan ölçüm
$L_{entropi}$: Entropi ödül fonksiyonu
L_{UCB}	: Üst güven sınırı ödül fonksiyonu
L_{LS}	: Seviye belirleme ödül fonksiyonu

KISALTMALAR

UGV	: İnsansız Kara Araçları (Unmanned Ground Vehicles)
UAV	: İnsansız Hava Araçları (Unmanned Aerial Vehicles)
ASV	: Otonom Yüzey Araçları (Autonomous Surface Vehicles)
AUV	: Otonom Su Altı Araçları (Autonomous Underwater Vehicles)
GP	: Gauss süreçleri (Gaussian Process)
SE	: Kareli Üstel (Squared Exponential)
RQ	: Rasyonel İkinci Dereceden (Rational Quadratic)
ROI	: İlgi Alanı (Region of Interest)
UCB	: Güvenli Üst Sınır(Upper Confidence Bound)
TSP	: Gezici Satıcı Problemi(Traversing Salesman Problem)
KP	: Total Skoru Maksimize Etme (Knapsack Problem)
RRT	: Hızlıca Keşfeden Rastgele Arama Ağacı (Rapidly Exploring Random Trees)
PRM	: Olasılıksal Yol Haritası (Probabilistic Road Map)
CE	: Çapraz Entropi (Cross Entropy)
MCTS	: Monte Carlo Arama Ağacı (Monte Carlo Tree Search)
BnB	: Dallan ve Sınırla (Branch and Bound)
ESAS	: Yorucu Arama ile Adaptif Örnekleme (Exhaustive Search Adaptive Sampling)
RMSE	: Kök ortalama kare hatası (Root Mean Square Error)
PI	: Beklenen Bilgi (Probability Improvement)
UCT	: Ağaçtaki üst güven sınırı (Upper Confidence Bound-Tree)
MOGP	: Çoklu çıkış gauss süreçleri (Multi Output Gaussian Process)
SSI	: Siluet Skor İndeksi
PSO	: Parçacık sürü optimizasyonu (Particle Swarm Optimization)
MSE	: Ortalama hatanın karesi (Mean Square Error)

1. GİRİŞ

Çevresel izleme, çevrenin kalitesini karakterize etmek ve izlemek için gerçekleştirilmesi gereken süreçleri ve faaliyetleri tanımlar. Çevresel değerlendirmelerinin hazırlanmasında ve insan faaliyetlerinin doğal çevre üzerinde zararlı riskler taşıdığı birçok durumda çevresel izleme yapılır. Tüm izleme stratejileri ve programlarının temelinde, ortamın mevcut durumunu belirlemek ve olası noktalar için belirsizlik tahmini yapmak vardır. Her durumda, izleme sonuçları gözden geçirilecek, makine öğrenmesi yöntemleriyle olasılıksal bir tahmin yapılır. Bu tahmin algoritmalarının geliştirilmesi için hava, su ve toprak örneklerinin düzenli olarak toplanması gerekmektedir. İnsansız Kara Araçları (Unmanned ground vehicles -UGV), İnsansız Hava Araçları (Unmanned Aerial Vehicles-UAV), Otonom Yüzey Araçları (Autonomous Surface Vehicles- ASV) ve İnsansız Su Altı Araçları (Autonomous Underwater Vehicles-AUV) şeklindeki mobil algılama platformları sıklıkla örnek toplamak için kullanılmaktadır [1]. Mobil Algılama Platformlarıyla,

- Sudaki artan plankton ekolojisinin [2,3]
- Suda yaşayan canlılar ve insanlar için zararlı, büyüyen alg çiçeklerinin [4–7]
- Su kalitesinin [8–16]
- Sudaki klorofil miktarının [17]
- Su altı mayın lokasyonlarının [18,19]
- Su altı ekolojisinin [20]
- Tarım arazilerine ve yerel ekosisteme zarar veren istilacı yabancı otların [21]
- Toprak parametrelerinin [22,23]
- Yangın ve deprem konumlarının [23,24]
- Volkanik patlama noktalarının [25]
- Mercan Reziflerinin [1,26]
- Maden yataklarının konumunun [27]
- Mars'ta lokalize edilen metan kaynaklarının [28]
- Orman ve açık denizler gibi ulaşılması zor alanlarda arama kurtarma çalışmalarının [29]
- Hava kirliliğinin [30,31]
- Yüzey bilgisinin [32]

izlenmesi ve örneklerin toplanması ortamın olasılıksal modelinin oluşturulmasında önem arz etmektedir.

Bu tez çalışmasında denizdeki klorofil miktarının yoğun olduğu bölgelerdeki örnekleri toplayabilmek için bir güzergah planlama algoritması geliştirilmiştir. Denizlerdeki klorofil miktarı, fitoplanktonun büyüme düzeyini yansıtarak deniz ekosisteminin temel dinamiklerinden biridir [33]. Fitoplankton, deniz ekosistemindeki birincil üreticiler olarak önemli bir rol oynar ve fotosentez yoluyla atmosferden karbonu alarak oksijen üretirler. Klorofil, bu fotosentez sürecinde temel bir

pigmenttir ve fitoplanktonun büyüme aktivitesini izlemek için yaygın bir gösterge olarak kullanılır [34].

Denizdeki klorofil miktarının artışı, fitoplanktonun yoğunluğunun ve biyokütlesinin arttığını gösterir. Bu durum, deniz ekosistemlerindeki besin döngüsünü etkiler ve çeşitli canlı gruplarını doğrudan etkiler. İşte denizdeki klorofil miktarının artmasının önemli etkileri:

Fitoplankton ve Zooplankton Popülasyonları: Denizdeki klorofil artışı, fitoplankton popülasyonlarının artmasına neden olur. Bu durum, zooplankton gibi fitoplanktonları tüketen organizmaların da artmasına yol açar [35]

Balık Popülasyonları: Fitoplanktonlar, besin zincirinin temelini oluşturduğundan, klorofil artışı balık popülasyonları üzerinde doğrudan etki eder. Balıklar, zooplanktonları tüketerek beslenir, bu nedenle fitoplankton artışı balık popülasyonlarının artmasına neden olabilir [34]

Deniz Memelileri: Deniz memelileri, özellikle balinalar ve deniz memelisi yavruları, doğrudan veya dolaylı olarak fitoplankton ve balıklara bağlı olarak beslenir. Bu nedenle, denizdeki klorofil miktarındaki değişiklikler, deniz memelilerinin besin kaynakları üzerinde etkili olabilir [36]

Oksijen Üretimi ve Karbon Denge: Fitoplanktonlar, fotosentez sırasında karbon dioksidi emer ve oksijen üretir. Bu nedenle, artan klorofil miktarı, atmosferdeki karbon dengesini etkileyebilir ve oksijen üretimini artırabilir [37]

Deniz Ekosistemi Dengesi: Denizdeki klorofil miktarındaki artışlar, ekosistem dengesini etkileyebilir. Özellikle aşırı ötrofikasyon durumlarında, alg çoğalmaları ve siyanobakteri patlamaları gibi olaylar deniz ekosistemlerinde bozulmalara neden olabilir [36]

Ötrofikasyon etkisi: Ötrofikasyon, besin elementlerinin (genellikle fosfor ve azot bileşikleri) neden olduğu tatlı su ve kıyusal denizde, biyolojik üretimin doğal işleyişinin zenginleşmesidir. Ötrofikasyon, görülebilir mavi-yeşil alg (siyanobakteri) veya algal patlamalar, yüzey birikintileri, yüzey bitki yığınları ve bentik macrofitlerin toplanmasıyla sonuçlanabilir. Bu nedenle sudaki klorofil miktarının artışı sudaki ötrofikasyon etkisinden de kaynaklanabilir [38].

Çeşitli çevresel parametreler, deniz ekosistemlerindeki klorofil-a miktarını etkileyebilir. İşte belirttiğiniz bazı parametrelerin klorofil-a miktarına muhtemel etkileri:

1. **Sıcaklık:** Sıcaklık, fitoplankton büyümesini etkiler. Genellikle, sıcaklık arttıkça fitoplankton aktivitesi ve klorofil a üretimi de artar. Ancak, aşırı yüksek sıcaklıklar veya aşırı düşük sıcaklıklar fitoplankton büyümesini olumsuz etkileyebilir.
2. **Tuzluluk:** Deniz suyunun tuzluluğu, fitoplankton türlerinin dağılımını etkileyebilir. Bazı türler belirli bir tuzluluk seviyesini tercih ederken, diğerleri daha farklı koşullarda gelişebilir. Bu, klorofil-a miktarını etkileyebilir.

3. **İletkenlik:** İletkenlik, suyun içinde çözülmüş tuz miktarını gösterir ve genellikle tuzluluk ile ilişkilidir. Bu nedenle, suyun iletkenliği, fitoplanktonun büyümesini ve dolayısıyla klorofil-a miktarını etkileyebilir.
4. **Oksijen Doygunluğu:** Oksijen, deniz suyu içinde önemli bir faktördür. Fitoplanktonlar, fotosentez sırasında oksijen üretirler. Oksijen doygunluğundaki değişiklikler, fitoplankton aktivitesini ve dolayısıyla klorofil-a miktarını etkileyebilir.
5. **Rüzgar Hızı:** Rüzgar, deniz yüzeyindeki karışımı etkileyebilir. Bu, besin maddelerinin yüzeye çıkmasına ve fitoplanktonların bu besin maddelerini kullanmasına olanak tanır, bu da klorofil-a üretimini etkileyebilir.
6. **Nem:** Hava neminin, yağış miktarının ve su buharının atmosferdeki döngünün, deniz suyu kalitesini ve besin içeriğini etkileme potansiyeli vardır. Bu da klorofil a miktarını etkileyebilir [39].

Sonuç olarak, denizdeki klorofil miktarının kontrolü, deniz ekosistemlerinin sağlığı, balıkçılık endüstrisi ve atmosferik karbon dengesi gibi önemli faktörlere yönelik stratejik çabaların bir parçasıdır. Bilim insanları ve çevre koruma uzmanları, bu parametreleri izleyerek, deniz ekosistemlerindeki değişiklikleri anlamaya ve sürdürülebilir deniz kaynaklarını korumaya çalışmaktadır. Bu çalışmada da NASA tarafından paylaşılan açık kaynak kod bir veri setindeki, sıcaklık, tuzluluk, iletkenlik, oksijen doygunluğu, rüzgar hızı ve nem parametreleri kullanılarak denizdeki klorofil miktarı tahmin edilmiştir.

Çevresel izleme problemlerinde mevcut örnekleme kümesini kullanarak ortamın olasılıksal modelinin oluşturulmasında uygun ve yaygın olarak Gauss süreçleri (Gaussian Process-GP) kullanılmaktadır [40–43]. GP, çekirdek fonksiyonunu kullanarak ve eğitim verisinden yola çıkarak ortam konumlarının birbiriyle olan ilişkilerini matematiksel olarak hesaplar. Örnekleme kümesi, çekirdek fonksiyonu içerisinde yer alan hiper parametreleri optimize etmek için kullanılır. Hiper parametre optimizasyonu genellikle marjinal olasılık fonksiyonu ile yapılır[1,12,29,40,42]. Çekirdek fonksiyonu için yaygın olarak Kareli üstel kovaryans fonksiyonu (Squared-Exponential-SE), Rasyonel ikinci dereceden (Rational Quadratic-RQ), Matern ve lineer fonksiyonlar kullanılır [44–51]. GP ile veri haritasının oluşturulmasında toplanan eğitim verilerinin konumu da önem arz etmektedir. Bu veri setinin konumunu belirlemek için iki yöntem kullanılır. Birinci yöntemde çevrim dışı olarak bütün örnekler toplandıktan sonra belirli bir kısmı eğitim verisi olarak kullanılır. İkinci öğrenme yönteminde ise, keşif öncesinde ve sırasında sürekli olarak yapılır, aktif öğrenme olarak adlandırılır. Çevrim dışı öğrenme yönteminde hiper parametre optimizasyonu yapmak için kullanılacak eğitim verileri ya köşelerden [52] yada ortamdaki rastgele [51] alınmaktadır. Fakat köşelerden alınan eğitim verileri ortamın tamamı hakkında detaylı bilgi vermemektedir. Rastgele toplanan verilerde bazı alanlarda belirsizliğe neden olmaktadır. Aktif öğrenme sırasında hiper parametre optimizasyonunda başlangıçta alınan eğitim verilerin konumları önemlidir. Aktif

öğrenme yönteminde de eğitim verileri genellikle rastgele seçilir [51]. Bu, hiper parametrelerin yanlış tahmin edilmesine yol açar. Bunun yanında Ericsson vd. [53] yaptıkları çalışmada ilk sensör konumlarının nerelere yerleştirileceği araştırmıştır. Bunun için Halton dizilimi ile rastgele dizilimi karşılaştırmışlardır. Halton diziliminin alanın tamamından bilgi toplaması nedeniyle daha iyi sonuç verdiği görülmüştür. Gribic vd. [54] giriş parametrelerini belirlemek için karşılıklı bilgi maksimizasyonu yöntemini kullanmışlardır. İlk değişkeni belirledikten sonra diğer değişkenlerin bilgilendiriciliği hesaplanmış ve sırasıyla bilgilendiriciliği en yüksek olan veriler seçilmiştir. Verilerin bilgilendiriciliği ölçülürken gecikmeli hava sıcaklığı ve nehrin akış hızı gibi parametreler kullanılarak hesaplanmıştır. Fakat su sıcaklığını etkileyen rüzgâr hızı, basınç, nem ve güneş radyasyonu gibi ölçümlerinde alınması gerekmektedir. Kemna vd. [52] yaptıkları çalışmada ise aday yol noktaları ile mevcut yörünge arasında sıcaklık faktörüne bağlı olarak bir softmax fonksiyonu tanımlamışlardır. Daha sonra her yol noktasına bağlı hesaplanan olasılık oranına bağlı olarak yeni örnekler seçilmiştir.

GP ile öğrenilen ortamdaki yeni örnekler keşif ve sömürü arasında bir denge kuracak şekilde toplanır. Keşif mevcut alan içerisinde farklı noktalardan örnekler alarak ilgi alanı (Region of Interest-ROI)'ni belirlemek için kullanılır. ROI alanları genellikle kirlilik gibi takip edilmesi istenilen özelliğin fazla olduğu alanlar olarak nitelendirilmektedir. Sömürü ROI alanlarda alınan örnekleme miktarının yoğunluğunu belirlemektedir [45]. Keşif ve sömürü arasındaki bu dengeyi sağlayacak yeni örneklem noktasını belirlemek için bir fayda fonksiyonu tanımlanır. Bu fayda fonksiyonu genellikle GP'den elde edilen ortalama ve varyans değerleri kullanılarak hesaplanır. Literatürde yaygın olarak, iki bilginin karşılıklı bağımlılığını ölçen karşılıklı bilgi (mutual information) [48,49,55–60], ortamdaki belirsizliği ölçen varyans [30,46,47], bir üst sınır tanımlayan UCB (Upper Confidence Bound-UCB) [61,62] ya da rastgele değişkenlerinin belirsizliğinin bir ölçüsü olan entropi [32,43,45,57,58,63] gibi ölçekler kullanılarak fayda fonksiyonu tanımlanır. Gotovas vd. [64] ise yeni bir fayda fonksiyonu olan seviye belirleme optimizasyonu olarak adlandırdıkları bir ödül fonksiyonu önermişlerdir. Önerilen fayda fonksiyonunun amacı alanı sınıflandırmaktır.

Belirli bir eşik değerinin üzerindeki bir küme altındaki ise başka bir küme atamışlardır. Yeni örnekler keşfedilmemiş örneklerden toplanmaktadır. Munir vd. [62] yapmış oldukları çalışmada sömürü yapmak için tahmin değerini bir katsayı ile çarpmışlardır. Katsayının değerini, 0, 0.25, 0.5, 0.75 ve 1 olarak belirlemişlerdir. En uygun sonucu 0.25 olduğunda elde etmişlerdir. Bunların yanından Petrillo vd. [65] GP ile eksenlerin konumlarını tahmin ederek eksenlerdeki hata oranına bağlı bir ödül fonksiyonu tanımlamışlardır. Fayda fonksiyonu örneklerin bilgilendiriciliğinin bir ölçüsüdür ve maksimum fayda sağlayan örnekler seçilir. Gerçek uygulamalarda birden fazla yerel maksimum noktası olabilir. Yani kirliliğin birden fazla lokasyonda arttığı bölgeler olabilir. Önerilen ölçekler keşif ve sömürü arasındaki dengeyi

tanımlarken çoğu problemde başarılı sonuç sergilese de bazı uygulamalarda birden fazla yerel maksimum noktası bulunması nedeniyle bu alanlardaki örnekleme yoğunluğu artmaktadır. Buda genel maksimum noktasının kaçırılmasına neden olmaktadır. Genel maksimum noktası kirlilik gibi parametrelerin maksimum olduğu alanlardır. Bu nedenle ortamın genel maksimum noktasının tahmin edilmesiyle ilgili çalışmalar yapılmaktadır [66]. Gerçek ortam uzamsal bir alanda modellenir. Literatürde uzamsal alanı modellemek için yol integralleri [67] gibi uygulamalar kullanılmaktadır. Fakat uzamsal alanda güzergah planlama algoritması geliştirmek oldukça karmaşıktır. Bu nedenle genellikle güzergah planlama algoritmaları skaler alanda yapılmaktadır. Otonom aracın örnekleri toplaması için gerekli güzergah planlama algoritmalarında olası yollar arasından en kısa yolu bulma (Traversing Salesman Problem-TSP) ve total skoru maksimize etme (Knapsack Problem-KP) problemlerinin bir kombinasyonu olan oryantiring problemi oluşmaktadır. Bu nedenle güzergah planlama algoritmaları NP-Hard bir problemdir [67]. En kısa yolu bulmak için RRT (Rapidly Exploring Random Trees-RRT) ve PRM(Probaplistic Road Map-PRM) olarak adlandırılan iki yöntem kullanılmaktadır. PRM algoritmasıyla olası yol noktaları üzerinden güzergah planlama yapılır. Petrillo vd. [65] sezgisel olarak hesaplanan N tane yol noktası belirlemişlerdir. Cheng vd. [68], PRM algoritmasını yine GP ile birleştirerek robotun kinematik modelini ve hedef konuma ulaşmak için olası rotalardan optimal rotayı tahmin etmek istemişlerdir. Wang vd. [69] RRT algoritmasıyla GP'yi birleştirmişlerdir. RRT algoritması ile ortamı ızgaralara bölerek bir güzergah oluşturulur. Çalışmada GP, engellerden ve insanlardan uzak bir çember çapı tahmin etmek için kullanılmaktadır. Olası tüm yolları belirleyerek en uygun yolu seçmek küçük menzilli alanlarda alanın hızla keşfedilmesini sağlamaktadır. Fakat alan büyüdükçe, keşif hızı eksponansiyel olarak düşmektedir [61,70]. Xiong vd. [71] RRT algoritmasından ilham alarak RAST adını verdikleri yeni bir algoritma sunmuşlardır. Geliştirdikleri algoritma turnuva seçimi yöntemi ve bilgilendirici sezgiseller kullanarak sürekli uzayda bilgilendirici verilerin verimli bir şekilde aranmasını hedefler.

Uzun menzilli alanları hızlıca keşfetmek ve güzergah planlama algoritmalarında oluşan oryantiring problemini azaltmak için olası yolların sınırlandırılması gerekmektedir [67]. Sınırlama yapılırken güzergah planlama algoritmaları, planlama uzunluğu açısından miyop ve miyop olmayan algoritmalar olarak ikiye ayrılmaktadır. Miyop olan yaklaşımda komşu düğümler arasından en kısa yola karar verilir [56]. Miyop olan güzergah planlama algoritmalarındaki hesaplama süresini azaltmak için görüntü işleme yöntemlerini kullanarak ortamın iskeleti çıkarılıp bu iskelet üzerinden en kısa komşu hücrelere ulaşılan çalışmalarda yer almaktadır [67]. Miyop olmayan yaklaşımda ise görev süresi ve otonom aracın kalan enerji miktarı gibi kısıtlamalar hesaba katılarak izlenmesi gereken yola karar verilir [72]. Gerçek uygulamalarda ufuk sınırı bilinmemektedir. Ufuk sınırını belirlemek için ya görüntü işleme yöntemleri yada sensörler gibi veriler kullanılarak ufuk sınırı tahmin edilebilir [72,73]. Fakat tek bir ufuk sınırı belirlemek

hesaplama yoğunluğunu arttırabilir. Ayrıca uzak menzilli alanlarda bunu tahmin etmek oldukça güçtür. Bu nedenle alan, probleme göre bölmelere ayrılarak her bölmede ufuk sınırını belirlemek hesaplama yoğunluğunu önemli ölçüde düşüreceklerdir. Literatürde Belman Ford algoritmasını kullanarak uzamsal ortamı bölümlendiren uygulamalar bulunmaktadır [19]. Burdaki en büyük etmen planlama uzunluğunun seçimidir. Planlama uzunluğu arttıkça daha doğru bir güzergah seçilecek fakat hesaplama yoğunluğu artacaktır. Jones vd. [74] yorucu arama algoritması kullanarak ortamı keşfetmişlerdir. Planlama uzunluğunun göreve bağlı olarak değiştiğini belirterek ve görev gereksiniminin maliyeti ile ortam entropisiyle ilişkili bir ödül fonksiyonu tanımlayarak güzergahı planlamışlardır. Görev gereksinimi bu makalede hedef bölgeye olan yakınlık olarak nitelendirilmiştir. Yani gerçek dünya problemleriyle bütünleşik bir ödül fonksiyonunun daha doğru olacağını belirtilmiştir. Yorucu arama algoritması ile bu ödül fonksiyonunu maksimize eden bir güzergah planlanmıştır. Nguyen vd. [75] uzun vadeli güzergah planlamak için Monte Carlo simülasyonlarını kullanarak kümeleme yapmışlardır. Yaptıkları kümeleme algoritmasında her küme içinde yorucu arama algoritmasını kullanıp güzergah planı yapmışlardır. Her planlama uzunluğu tamamlandıktan sonra ortam öğrenilmiş ve yeni yollar oluşturulmuştur. Fayda fonksiyonu olarak belirsizliği minimize etmişlerdir. Yorucu arama algoritmasında planlama uzunluğu boyunca olası yollar değerlendirilerek maksimum fayda sağlayan yollar seçilir. Tiwari vd. [76] sürü robotiği kullanarak ROI alanını belirleyen bir çalışma yapmışlar ve ortamı keşfetmek için geçen süreyi azaltmak istemişlerdir. Literatürde yapılan çalışmalar incelendiğinde Yorucu arama algoritması genellikle belirsizliği minimize eden algoritmalarda kullanılmıştır. Bu nedenle sadece keşif için kullanılmıştır. Güzergah planlama algoritmalarında olası yolları sınırlamak için genellikle Çapraz Entropi (Cross Entropy-CE) ve Dallon ve Sınırla Tekniği (Branch and Bound-BnB) kullanılır [45,49,61,70,77,78]. CE yönteminde belirsizliği yüksek yollar tercih edilir [45]. BnB yönteminde ise teorik olarak bir alt sınır ve üst sınır hesaplanır. Alt sınırın altında kalan ve üst sınırın üstünde kalan yollar elenir [46,49,79].

Belirlenen alt sınır ve üst sınır probleme göre değişmektedir. Yapılan çalışmalar incelendiğinde bu alt sınır ve üst sınır için anlık risk ve beklenen risk arasındaki farkı baz alan sınırlar belirlenmiştir. Alt sınır olarak olası tüm lokasyonlar için aradaki bu farkın toplamı olarak belirlenmiş, robotun bir sonraki adımda gidebileceği bütün lokasyonlar arasından maksimum değere sahip olan konum üst sınır olarak belirlenmiştir. Üst sınır sürekli olarak hesaplanmaktadır. Bunun önüne geçebilmek yaklaşık bir çözüm önerisi sunulmuştur. Yaklaşık çözümde üst sınır olarak anlık risk ile beklenen risk arasındaki farkı baz alan bir ödül fonksiyonu hesaplanmıştır. Her yeni hücrede ödül değeri arttıkça üst sınır güncellenmiştir. İki çözüm karşılaştırıldığında hesaplama süresinin büyük oranda düştüğü, fakat bu düşüşe oranla arama performansının daha az düştüğü gözlenmiştir [79]. Best ve Fitch [80] alt sınırı aç gözlü bir güzergah planlama algoritmasıyla belirlemişlerdir. Üst sınırı ise robotun mevcut bütçeyle gidebileceği düğümlerden elde edebileceği

maksimum ödül değerini kullanarak hesaplamışlardır. Ayrıca yapılan çalışmada arama yapılırken maksimum alt sınır, maksimum üst sınır, maksimum ortalama sınır gibi farklı özelliklerdeki düğümleri sıralayarak arama yapmışlardır. En çok sınırlama maksimum üst sınır kullanılarak elde edilmiştir.

Literatürde oryantiring probleminin önüne geçebilmek için deterministik kadar sezgisel yöntemlerde mevcuttur. MCTS yöntemi ise büyük durum uzay modellerinde optimale yakın sonuçlar vermektedir. Nguyen vd. [75] ses sinyalinin konumlandırılmasını sağlamak için MCTS algoritmasını kullanmışlardır. Oluşturulan güzergah açgözlü bir yaklaşım ve rastgele arama yöntemiyle kıyaslanmış ve sonuçlar daha iyi çıkmıştır. Chen [81] ise MCTS algoritmasını pareto planlamasıyla beraber kullanmıştır. Pareto planlaması içinde bir çok görev olan veri setini temsil etmektedir. Bu veri setinde çoklu robotlar için görev dağılımı yapmak için MCTS algoritmasını kullanmışlardır. Kurzer vd. [82] işbirliği gerektiren durumlarda otomatikleştirilmiş araçlar için çok etmenli yörünge planlama algoritmalarını hızlandıracak bir yöntem önermektedir. Trafikğin olası durumunu tahmin ederek en iyi trafik seneryosuna MCTS ile karar vermişlerdir. Elde edilen sonuçlarda seneryo sayısı 2000'i aştığında performans değerinin düştüğü gözlenmiştir. Jang vd. [78] ortamın belirsizliğini düşürmek çoklu robotlarda yörünge planlama algoritması olarak MCTS algoritmasını kullanmışlardır. Kartal vd. [83] çoklu robotlar için görev ataması yaparken MCTS algoritması ile BnB tekniğini birleştirmişlerdir. MCTS'nin arama sırasında gereksiz arama yapmaması için BnB ile arama ağacını sınırlandırmışlardır. Kathen vd. [84] genetik bir algoritma olan PSO (Partical Swarm Optimization-PSO) yöntemi kullanarak deneysel bir veri setinde adaptif örnekleme yapmışlardır. Ypacarai gölünde su kalitesini belirlemek için güzergah oluşturmuşlar ve düşük bir MSE (Mean Square Error-MSE) oranı ile ortamın öğrenildiği görülmüştür. Ayrıca başka bir çalışmada Kathen vd. [85] PSO algoritması yerine geliştirilmiş GP tabanlı bir PSO algoritması kullanarak başarımını PSO algoritması ile kıyaslamışlardır. Sonuçlarda geliştirilmiş GP tabanlı PSO algoritmasının çevresel izlemede daha iyi olduğu görülmüştür.

Assenine vd. [31] sezgisel ve genetik algoritmalara alternatif olarak takviyeli öğrenme algoritmasını kullanarak performanslarını genetik ve rastgele yöntemlerle kıyaslamışlardır. Yapılan çalışmada, takviyeli öğrenme algoritması olarak Q öğrenme algoritmasından yararlanılmıştır. Barrionuevo vd.[14] iki katmanlı Q öğrenme algoritmasını kullanarak adaptif örnekleme yapmışlardır. Elde edilen sonuçlarda çim biçme rotası ve PSO algoritmasına oranla başarımının %82 arttığını belirtmişlerdir. Lui vd. [16] iki katmanlı Q öğrenme algoritmasında Q değerini tahmin etmek için derin öğrenme yöntemini kullanmışlardır. Elde edilen sonuçlarda PSO ve iki katmanlı Q öğrenme algoritmasına göre başarımının %20 arttığı söylenmiştir. Önerilen yöntemlerde, rastgele oluşturulan yol ve genetik algoritmadan daha iyi sonuçlar vermiş olmasına rağmen, oluşturan güzergahın optimale yakınlığı üzerine durulmamıştır. Bunun yanında güzergah planlama

algoritmalarında dikkat edilmesi gereken diğer nokta statik ve dinamik engellerdir. Bu engeller güzergah planlama algoritmalarında yol haritalarına köşe olarak eklenmiştir [86,87].

Bu tez çalışmasında GP ile bilinmeyen bir ortamın gerçek zamanlı ve yüksek doğrulukla öğrenilmesi amaçlanmaktadır. GP, skaler bir ortamın olasılıksal modelinin oluşturulmasında oldukça uygun bir yöntemdir. Fakat dinamik ortamda çevresel izleme yapılacak alanın boyutunun artması örnekleme kümesinin büyüklüğünün artmasına neden olmaktadır. GP’de her örnek alındıktan sonra hiper parametre optimizasyonu yapılmaktadır. Hiper parametre optimizasyonunda her örnek noktasının diğer örnek noktalarıyla ilişkisini matematiksel olarak tanımlayan kovaryans fonksiyonunun tersinin alınması hesaplama yoğunluğunu önemli ölçüde arttırmaktadır. Literatürde bu hesaplama gücünü azaltmak için her örnek alındıktan sonra hiper parametre optimizasyonu gerçekleştirmek yerine k adet örnek alındıktan sonra hiper parametre optimizasyonu yapılmıştır [2]. Ayrıca GP hesaplama gücünü azaltmak için Seyrek Temsili Giriş’ler (Sparse Pseudo Input) kullanılmıştır [47,55]. Bu örnekleme kümesinin seçimi probleme göre değişmektedir. Yapılan çalışmalarda, sözde girişler doğrusal zamanla değişmeyen sistemler için stokastik diferansiyel denklemler kullanılarak üretilmiş [88], genel ve yerel maksimum noktalarından [89], kullback leibler sapmasını en aza indirgeyecek fayda fonksiyonunu kullanarak [90], ve bir limit belirlenerek, bu limitte maksimum fayda sağlayan örnekler seçilmiştir [91]. Ayrıca Kovaryans matrisinin tersinin alınırken oluşan hesaplama yoğunluğunu azaltmak için Mahalanobis mesafesi kullanılmıştır. Bu mesafenin 1’den büyük olması durumunda kovaryans matrisindeki ilgili değeri 0 alınmıştır [92]. Bu çalışmaların yanında Jakkala vd. [93] sensör konumlarını belirlemek için seyrek temsili girişler kullanmışlardır. Çalışmada veri uyumu, karmaşıklık ve izleme terimlerini içeren bir optimizasyon denklemi kullanılmıştır. Bu değeri 0 yapan sensör konumları oluşturulmuştur. Literatürde Gauss regresyon probleminin hesaplama gücünü azaltmak için yapılan çalışmalar incelendiğinde, genellikle bilgilendiriciliği yüksek sınırlı sayıda örnekler seçilerek alt örnekleme kümesi oluşturulmuştur. Yapılan çalışmalarda alt kümelerin optimuma yakınlığıyla ilgili bir sonuç yer almamaktadır [4].

Bu nedenle bu tez çalışmasında optimuma yakın bir alt örnekleme kümesinin seçilmesi hedeflenmektedir. Bunu belirlemek için kümeleme yöntemleri kullanılmıştır. Kümeleme denetimsiz bir makine öğrenmesi tekniğidir. Benzer girişleri gruplandırmak için kullanılır [94]. Bu çalışmada yaygın olarak kullanılan K-Means, Agglomerative Hierarchical, DBSCAN, MeanShift, Birch, MiniBatchKMeans ve Spektral gibi kümeleme algoritmaları eğitim verilerini kümelemek için kullanılmıştır. Kümeleme algoritmalarının performansı, SSI (Siluet Skor İndeksi) kullanılarak değerlendirilir, bu, -1 ile +1 arasında değişen bir kümeleme metriği olup, +1’e daha yakın bir değer daha iyi kümeleme performansını gösterir. En iyi kümeleme yapan algoritmanın seçtiği küme üzerinden sözde girişler belirlenmiştir. Seçilen sözde girişlerden optimale en yakın olanı belirlemek için her kümeden 1,2,3,4,5 gibi sayılarda veriler alınmış ve optimale yakınlığı test edilmiştir.

Özetle, çevresel izleme yapılırken kullanılan güzergah planlama algoritmaları ve ortamı öğrenmek için kullanılan fayda fonksiyonları, modern teknolojinin çevresel analiz ve inceleme süreçlerinde önemli bir rol oynamaktadır. Bu algoritmalar ve fonksiyonlar sayesinde, çeşitli çevresel faktörlerin izlenmesi, analiz edilmesi ve gerektiğinde müdahale edilmesi mümkün hale gelmektedir. Güzergah planlama algoritmaları, en etkili ve verimli rotaların belirlenmesine yardımcı olurken, fayda fonksiyonları da bu rotaların optimize edilmesine katkı sağlamaktadır. Bu sayede çevresel izleme süreçleri daha hassas, hızlı ve doğru bir şekilde gerçekleştirilebilmektedir. Dolayısıyla, çevresel izleme yapılırken kullanılan güzergah planlama algoritmaları ve fayda fonksiyonları, çevresel koruma, izleme ve müdahale alanında önemli bir araç haline gelmiştir. Bu tekniklerin geliştirilmesi ve kullanımı, çevre bilimlerindeki ilerlemeleri desteklemekte ve doğal kaynakların sürdürülebilir bir şekilde yönetilmesine yardımcı olmaktadır. Tablo 1.1 çevresel izleme yapılırken kullanılan güzergah planlama algoritmalarını ve ortamı öğrenmek için kullanılan fayda fonksiyonlarını vermektedir.

Tablo 1.1. Literatür Taraması

Güzergah Planlama Algoritması	Kullanılan Fayda Fonksiyonu	Çevresel İzleme Yöntemi	Kaynak	Yıl
Özyinemeli açgözlü güzergah planlama algoritması	Karşılıklı bilgi maksimizasyonu	Su altı aracıyla trafikten uzak yörünge tahmini	[56]	2010
ESAS	Entropi	Hava kirliliği	[74]	2011
Aç gözlü en iyi ilk arama (Belman Ford algoritması ile ortam bölmelendirme)	Belirsizlik azaltma	Okyanus altı izleme	[19]	2011
BnB	Karşılıklı bilgi maksimizasyonu	Sıcaklık, Tuzluluk, İletkenlik	[46]	2012
CE	-	Engelli ortamda yörünge tahmini	[61]	2012
Maksimum fayda sağlayan yörünge seçimiyle güzergah planlama	Entropi	Sıcaklık	[70]	2013
Maksimum fayda sağlayan düğüm seçimiyle güzergah planlama	Karşılıklı bilgi maksimizasyonu(Düğümleri seçmek için seviye belirleme ödül fonksiyonu kullanılmıştır)	Göldeki siyano bakteri haritası	[72]	2014

Tablo 1.1. (Devamı) Literatür Taraması

ESAS	Belirsizlik azaltma	Hava kirliliği ve orman yangını izleme	[75]	2014
BnB	Beklenen olasılık	-	[80]	2016
Özyinelemeli adaptif örnekleme (Recursive Adaptive Identification – RAId)	Hedefe olan uzaklık	Hedef bulma	[73]	2016
BnB	Varyans azaltma	Çoklu robot ile aktif algılama stratejisi	[76]	2017
BnB	Beklenen bilgi kazancı	Su altı hedef bulmak	[79]	2017
MCTS	Entropi azaltma	Ses lokalizasyonu	[77]	2017
MCTS	Karşılıklı bilgi maksimizasyonu	Çoklu robot ile aktif algılama stratejisi	[60]	2018
CE	UCB	Su kirliliği	[49]	2018
CE	Karşılık bilgi maksimizasyonu	Zararlı alg çiçekleri	[4]	2019
Aç gözlü sezgisel arama algoritması	Seviye belirleme, Entropi	pH	[67]	2019
GMR-RRT*	Heuristik örnekleme	Engelli ortamda yörünge tahmini	[69]	2020
MCTS	Entropi azaltma	Zararlı alg çiçekleri	[81]	2020
RAST	Belirsizlik azaltma	Sıcaklık	[71]	2020
PRM	-	Engelli ortamda yörünge tahmini	[68]	2021
CE ve BnB	UCB	Hipoksik bölge tahmini	[45]	2021
MCTS	Entropi	Sıcaklık	[78]	2021
Aç gözlü en iyi ilk arama	Entropi	Yüzey bilgisi	[32]	2021
Aç gözlü en iyi ilk arama	Beklenen iyileşme	Su kalitesi	[15]	2021
PSO	GP'den elde edilen ortalama ve varyansı kullanan bir ödül fonksiyonu	Su kalitesi	[84]	2021
RRT	UCB	Tehlikeli gaz lokalizasyonu	[95]	2022
PRM	Tahmini öklid mesafesi	Lokalizasyon	[65]	2023
Q Öğrenme	Karşılıklı bilgi maksimizasyonu	Hava kirliliği	[31]	2024
İki Katmanlı Q öğrenme	GP'den elde edilen ortalama ve varyansı kullanan bir ödül fonksiyonu	Su kalitesi	[14]	2024
Geliştirilmiş GP tabanlı PSO	GP'den elde edilen ortalama ve varyansı kullanan bir ödül fonksiyonu	Su kalitesi	[85]	2024
Derin iki katmanlı Q öğrenme	Karşılıklı bilgi maksimizasyonu	Su kalitesi	[16]	2024

Bu tez çalışması çevresel izleme alanında bilgilendirici yol planlama algoritmasında karşılaşılan zorlukları ele almakta ve çözüm önerileri sunmaktadır. Çalışmanın katkıları şu şekilde sıralanabilir:

1.Eğitim Verilerinin Önemi: Eğitim verilerinin konumu GP ile çevresel izleme yaparken tahmin performansını önemli ölçüde etkilemektedir. Fakat literatürde bu konudaki çalışma sayısı sınırlıdır. Bu zorlukla başa çıkmak için, sınırlı bir bütçeyle tüm alandan örnekler toplamak amacıyla yeni bir güzergah planlama algoritması geliştirilmiştir.

2. Algoritmanın Performansı: Geliştirilen güzergah planlama algoritması, 10 farklı veri kümesinde test edilerek literatürdeki çalışmalarla kıyaslanmıştır. Sonuçlar, algoritmanın ortamı daha etkili bir şekilde öğrendiğini göstermektedir.

3. ROI Alanı Keşfi: Bu tez çalışmasında, ödül fonksiyonlarının ortamı öğrenme ve ROI keşfetme başarısı incelenmiştir. Bu amaçla, içinde 1, 2, 3 ve 4 ROI alanı bulunan bir test ortamı oluşturulmuş ve mevcut konumdan 5 farklı yörünge seçilerek maksimum fayda sağlayan yörüngelerle yeni örnekler toplanmıştır. Literatürde gerçek bir veri setinde UCB ödül fonksiyonu ile yeni örnekler toplamak için bir güzergah oluşturulmuştur[45]. Bu tez çalışmasında, oluşturulan test ortamının doğruluğu literatürdeki çalışma ile karşılaştırılarak doğrulanmıştır ve farklı ödül fonksiyonlarının performansları değerlendirilmiştir. Bu tez çalışmasında seviye belirleme, UCB ve entropi ödül fonksiyonları kullanılmıştır. Sonuçlar, entropi ödül fonksiyonunun diğer ROI alanlarını daha hızlı keşfettiğini göstermiştir. UCB ödül fonksiyonunun ise belirsizliği yüksek haritalarda ortamı daha iyi öğrendiği, ancak belirsizliği düşük haritalarda yerel maksimum noktalarına takılarak ortamın belirsizliğini düşüremediği gözlemlenmiştir. Başlangıçta, mevcut eğitim kümesinin fazla öğrenmeye neden olduğu ve bu durumun yüksek belirsizliğe sahip bir varyans haritası ile sonuçlandığı tespit edilmiştir. Bu tez çalışmasında, aynı eğitim verisi miktarıyla alanın başlangıçtaki belirsizlik değeri çapraz doğrulama yöntemi kullanılarak azaltılmıştır.

4. Monte Carlo Arama Ağacı ve BnB Birleşimi: Adaptif örnekleme yaparken en kısa sürede en bilgilendirici örnekleri toplamak gezgin satıcı problemi olarak bilinmektedir. Üçüncü aşamada, gezgin satıcı problemine çözüm sunmak için MCTS algoritması ile BnB algoritması birleştirilmiştir. MCTS algoritması genellikle sezgisel çalışırken, BnB algoritması kesin ve optimize bir şekilde çalışır. Bu çalışmada, bu iki algoritma birleştirilerek hibrit bir bilgilendirici yol planlama algoritması oluşturulmuştur. MCTS algoritması alt limit değerini hesaplarken, üst limit sınırı ise sınırlı bütçeyle toplanabilecek maksimum faydayı temsil eder.

5. Algoritma Performansı ve Karşılaştırılması: Tezde geliştirilen algoritmanın performansı, MCTS ve ESAS ile karşılaştırılmıştır. Elde edilen sonuçlar, MCTS-BnB algoritmasının, MCTS kadar hızlı rota belirleyip ESAS algoritması gibi deterministik yöntemler kadar optimale yakın rotalar oluşturduğunu göstermektedir.

6. Farklı Veri Setlerinde Algoritmanın Performansının Araştırılması: MCTS-BnB algoritmasında üst limit ve alt limit hesaplamaları, veri setinin özelliklerine bağlı olarak büyük önem taşımaktadır. Veri setindeki değişkenliğin fazlalığı kovaryans fonksiyonundaki ℓ hiper parametresini düşük çıkartırken, az değişkenlik ise yüksek değerlerle sonuçlanmaktadır. Bu

nedenle, NASA tarafından sağlanan sıcaklık veri setinde ($\ell=1$ ve 3) ve deneysel bir veri setinde ($\ell=9$) MCTS-BnB algoritmasının başarımı test edilmiştir.

7. Önerilen MCTS-BnB Algoritmasının Başarımının Literatüdeki Çalışmalarla Karşılaştırılması: BnB algoritmasının başarısında, üst ve alt limitlerin hesaplanması büyük önem taşımaktadır. Bu iki sınır birbirine yakın olduğunda, daha fazla dal budanmakta ve dolayısıyla hesaplama süresi azalmaktadır. Ancak, bu durum optimal düğümün kaçırılma riskini artırmaktadır. Üst ve alt limit aralığının genişlemesi ise daha az dalın budanmasını sağlamakta ve hesaplama süresini artırmaktadır. Bu tez çalışmasında, alt ve üst limit hesaplamalarının literatürdeki diğer yöntemlerle karşılaştırılması gerçekleştirilmiştir. Algoritmanın performansını test etmek amacıyla, mevcut veri setine literatürdeki çalışmaların programları yazılarak uygulanmıştır. Elde edilen sonuçlarda, önerilen yöntemin daha kısa sürede ve optimal rotayı kaçırmadan mevcut güzergahı belirlediği görülmüştür.

8. Seyrek Temsili Girişlerin Belirlenmesi: Gauss regresyonunda eğitim veri kümesi büyüdükçe hesaplama süresi de artmaktadır. Bu çalışmada, eğitim veri kümesini azaltmak için seyrek temsili girişler belirlenmiştir. Alt örneklem kümesi oluşturmak için veri setine 5 farklı kümeleme yöntemi uygulanmış ve maksimum SSI (Siluet Skor İndeksi) skoruna sahip kümeleme yöntemi seçilmiştir. Eğitim veri seti bu kümelere ayrılmış ve her kümeden 1, 2, 3 ve 4 örnek alınarak hesaplama süresi ve sonuçların optimal rotaya yakınlığı incelenmiştir. Kümelerden alınan örnek sayısı arttıkça, rota belirleme doğruluğu ve hesaplama süresi artmıştır.

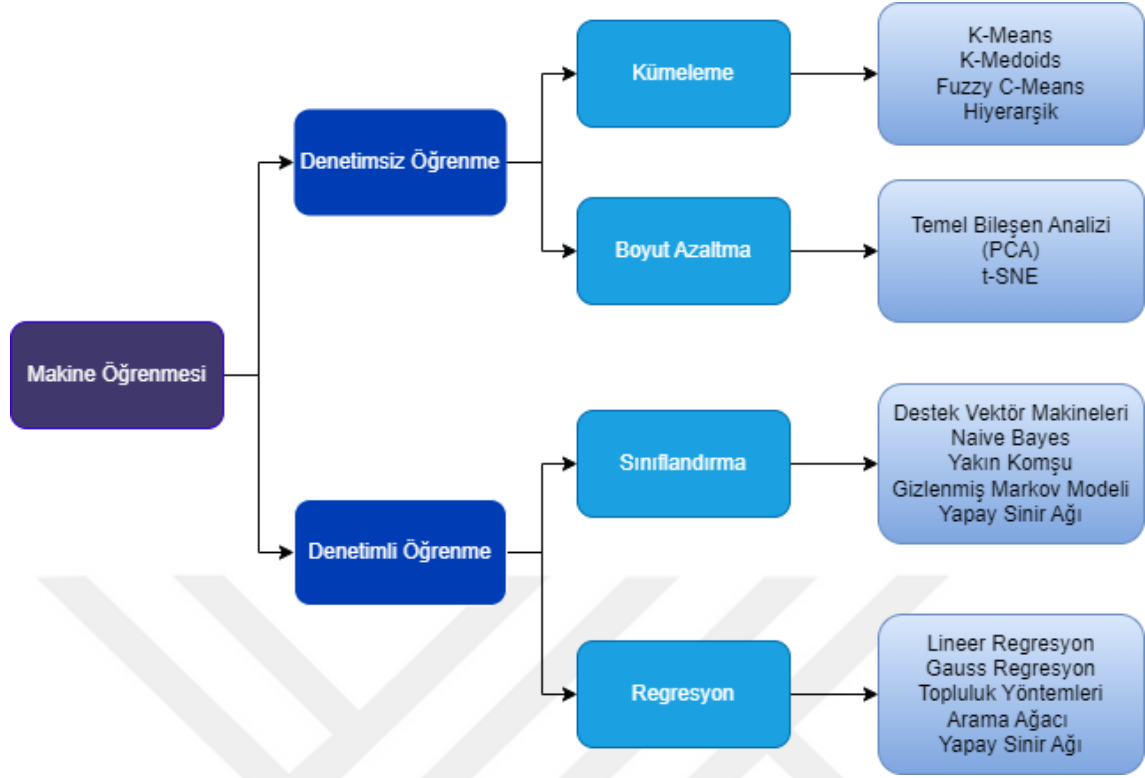
Bu tez çalışması, çevresel izleme alanında önemli sorunlara odaklanarak yeni bir yol planlama algoritması geliştirmiş ve literatürdeki eksiklikleri gidermeye yönelik değerli katkılar sağlamıştır. Yapılan çalışmaların sonuçları, çevresel izleme sistemlerinin verimliliğini artırmak ve kaynak kullanımını optimize etmek için önemli bir adım olabilir.

2. MATERYAL VE METOT

Bu tez çalışmasında, Gauss süreçleri çerçevesinde oluşturulan veri haritaları ve güzergâh planlama algoritmaları kullanılarak çevresel veri toplama süreçlerinin geliştirilmesi amaçlanmıştır. Ortamın olasılıksal modelini oluşturmak için Gauss süreçlerinden (GP) faydalanılmıştır. GP ile ortamın olasılıksal modelini oluşturma aşamaları Bölüm 2.1'de detaylandırılmıştır. GP, ortamı öğrenmek için eğitim verilerini kullanmaktadır. Bu çalışmada, ortamı öğrenmek amacıyla yeni bir güzergâh planlama algoritması önerilmiştir. Önerilen güzergâh planlama algoritması, Bölüm 2.2'de ayrıntılı bir şekilde açıklanmıştır. Ancak, değişken çevresel parametreler gibi faktörlerden ötürü ortamdaki sürekli olarak yeni örneklerin toplanması gerekmektedir. Gauss süreçleri, ortamı öğrenmek ve özellik yoğunluğunun yüksek olduğu ROI bölgelerinden yeni örnekleri toplamak için bir güzergâh planlama algoritmasıyla entegrasyon halinde çalışmaktadır. Ödül fonksiyonlarının ROI alanındaki performansını değerlendirmek amacıyla 1, 2, 3 ve 4 ROI içeren bir veri seti oluşturulmuştur. Bölüm 2.3, kullanılan ödül fonksiyonlarını, bu fonksiyonların performansını değerlendirmek için geliştirilen güzergâh planlama algoritmasını ve ROI alanındaki başarımlarını artırmak için uygulanan çapraz doğrulama yöntemini ele almaktadır. Bir sonraki bölümde, adaptif örnekleme sürecinde ortaya çıkan gezgin satıcı problemi üzerinde durulmaktadır. Gezgin satıcı problemi, maksimum bilgiye sahip örnekleri en kısa sürede toplamayı hedefleyen bir optimizasyon problemidir. Bu çalışmada, gezgin satıcı problemi için hibrit bir güzergâh planlama algoritması önerilmiştir. Algoritmanın performansını değerlendirmek amacıyla literatürdeki diğer çalışmalar, mevcut veri setine uygulanacak şekilde kodlanmıştır. Bölüm 2.4, uygulanan adaptif güzergâh planlama algoritmaları hakkında detaylı bilgileri içermektedir. Son aşamada, önerilen yöntemin hesaplama süresini azaltmak için alt örneklem kümeleri, kümeleme yöntemleri kullanılarak seçilmiştir. Bölüm 2.5, bu örneklem kümeleri hakkında ayrıntılı bilgiler sunmaktadır. Ayrıca, Bölüm 2.6'da anlatılan siluet skor indeksi kullanılarak kümeleme yöntemlerinin performansı ölçülmüştür.

2.1. Makine Öğrenmesi Yöntemleri

Makine öğrenmesi yöntemleri girişler ve çıkışlar arasındaki örüntüyü tanımlamak için kullanılmaktadır. Öğrenme yöntemleri denetimli ve denetimsiz öğrenme olarak ikiye ayrılmaktadır. Denetimli öğrenmede, öğrenme algoritmasına hem giriş hem de çıkışlar verilir. Denetimsiz öğrenmede ise sadece girişler sisteme verilerek tahmin yapılır. Denetimsiz öğrenme yöntemleri kümeleme yapmak için kullanılmaktadır. Denetimli öğrenme yöntemleriyle ise sınıflandırma ve regresyon yapılmaktadır. Şekil 2.1'de yaygın olarak kullanılan makine öğrenmesi algoritmaları yer almaktadır.



Şekil 2.1. Makine Öğrenmesi Yöntemleri

Gauss Süreçleri (GP) , regresyon ve olasılıksal sınıflandırma problemlerini çözmek için tasarlanmış genel bir denetimli öğrenme yöntemidir.

Gauss süreçlerinin avantajları,

- Gözlemlenen noktalar arasında enterpolasyon yaparak tahmin yapar.
- Tahmin olasılıksaldır, böylece ampirik güven aralıkları hesaplanabilir.
- Tahmin için kullanılan kovaryans matrisi özelleştirilebilir.

Gauss süreçlerinin dezavantajları,

- Seyrek değildirler, yani tahmini gerçekleştirmek için tüm örnekleri/özellik bilgilerini kullanırlar.
- Yüksek boyutlu uzaylarda, yani öznitelik sayısı birkaç düzineyi aştığında verimliliklerini kaybederler.

Çevresel izleme problemlerinde, GP kompleks değişken ilişkilerini modellemek için yaygın olarak kullanılmaktadır. Bu tez çalışmasında Gauss regresyon algoritması ile ortamın olasılıksal modeli oluşturulmaktadır. Ortamın olasılıksal modeli oluşturulurken girişler kadar çıkışlarda önem arz etmektedir. Bu nedenle çoklu çıktılı Gauss regresyon algoritmasında kullanılmıştır. Ayrıca Gauss regresyonun hesaplama yükünü azaltmak için alt örneklem kümesi seyrek sözde girişler kullanılarak seçilmektedir. Bu bölüm bu çalışmalar hakkında detaylı bilgi vermektedir.

2.1.1. Gauss Süreçleri ile Ortamın Olasılıksal Modelinin Oluşturulması

Gauss regresyonun sahip olduğu avantajlardan dolayı çevresel izleme problemlerinde, ortamın olasılıksal modelinin oluşturulmasında sıklıkla kullanılmaktadır. Standart lineer regresyon modelinde $f(x)$, Denklem 1’de olduğu gibi ifade edilir. x girişleri, y çıkışları, w ise ağırlık oranını temsil eder. ϵ , sensör ve ölçümlerden kaynaklı hata oranını temsil eder. Bu hata oranı Gauss regresyonda belirli bir ortalama ve varyansa sahip normal dağılım ile ifade edilir. Denklem 2’de ortalaması 0, varyansı σ_n^2 ile gösterilen bir dağılım olarak verilmiştir.

$$f(x) = x^T w, \quad y = f(x) + \epsilon \quad (1)$$

$$\epsilon \sim \mathcal{N}(0, \sigma_n^2) \quad (2)$$

Eğitim verisi kullanılarak hesaplanan olasılık dağılımı Denklem 3’de verilmektedir. Ortalaması $f(x)$, varyansı ise hatanın değişimini vermektedir.

$$p(y|x, w) \sim \mathcal{N}(x^T w, \sigma_n^2 I) \quad (3)$$

Gözlemlenen eğitim verisi ise Denklem 4’de gösterildiği gibi bir dağılıma sahiptir.

$$w \sim \mathcal{N}(0, \Sigma_p) \quad (4)$$

Bu olasılık ve önsel dağılımdan yola çıkarak, sonsal dağılım Bayes teoremiyle Denklem 5’de hesaplanmaktadır.

$$\text{Sonsal} = \frac{\text{Olasılık} \times \text{Önsel}}{\text{Marjinal Olasılık}}, \quad p(w|y, X) = \frac{p(y|X, w)p(w)}{p(y|X)} \quad (5)$$

Özetle, GP ile belirli bir hatayla gelen ölçüm verilerinden yola çıkarak belirli ortalama ve varyansa sahip bir dağılım elde edilmektedir. Denklem 6’da bu dağılımın gösterimi yer almaktadır.

$$f(x) \sim GP(m(x), k(x, x')) \quad (6)$$

Eğitim verileri kullanılarak öğrenilen ortamın doğruluğunu tespit etmek için test verileri kullanılmaktadır. Denklem 7’de $K(X_*, X_*)$ test verilerinin kovaryans fonksiyonunu vermektedir. $K(X, X)$ ise eğitim verilerinin kovaryans fonksiyonunu vermektedir. σ_n sensörlerdeki gürültünün bir ölçüsüdür.

$$\begin{bmatrix} y \\ f_* \end{bmatrix} \sim \mathcal{N} \left(0, \begin{bmatrix} K(X, X) + \sigma_n^2 I & K(X, X_*) \\ K(X_*, X) & K(X_*, X_*) \end{bmatrix} \right) \quad (7)$$

Denklem 7’de yer alan matrisin indirgenmesiyle elde edilen test verilerinin ortalama ve varyans değerleri sırasıyla, Denklem 8 ve 9’da yer almaktadır.

$$\bar{f}_* = k_*^T (K + \sigma_n^2 I)^{-1} y \quad (8)$$

$$\mathbb{V}[f_*] = k(x_*, x_*) - k_*^T (K + \sigma_n^2 I)^{-1} k_* \quad (9)$$

2.1.2. Çok Çıktılı Gauss Süreci

Çoklu Çıktılı Gauss Regresyonu, özellikle tek bir konumdan alınan çok sayıda ölçümü içeren senaryolarda (örneğin, su parametreleri gibi), birden çok bağımlı değişkenin (çıktıların) eş zamanlı tahmin edilmesi için özel olarak kullanılan bir regresyon tekniğidir. Bu tür durumlarda, sadece girdiler arasında değil, çıktılar arasında da bir korelasyon bulunmaktadır. Çoklu Çıktılı Gauss Regresyonu, bu çıktıları karma bir kovaryans fonksiyonu kullanarak hesaplar. Bu fonksiyon, Denklem 10'da [96] verilmiştir.

$$K_M = \begin{bmatrix} K_{11} & K_{12} & K_{13} & \dots & K_{1T} \\ K_{21} & K_{22} & K_{23} & \dots & K_{2T} \\ & & & \ddots & \\ & & & & \ddots \\ K_{T1} & K_{T2} & K_{T3} & \dots & K_{TT} \end{bmatrix} \quad (10)$$

$$f(x) \sim GP(0, K_M + \Sigma_M) \quad (11)$$

"T" sayıda çıktı arasındaki ilişkiyi hesaplayan " $K_M \in R^{T \times T}$ " ifadesi, " $K_T \in R^{1 \times T}$ " ifadesi ise, her bir çıktı ile kendi arasındaki kovaryans fonksiyonunu tanımlar. Denklem 11'de, $f(x) = \{f_1(x), \dots, f_t(x)\}$ çıktılar x örnek noktasında bulunur. " $\Sigma_M \in R^{T \times T}$ " ifadesi ise, $\{\sigma_i^2\}_{i=1}^T$ öğelerini içeren diyagonal bir matristir.

$$\hat{f}(x_*) = K_{M_*}^T (K_M + \Sigma_M)^{-1} y \quad (8) \quad (12)$$

$$\Sigma_* = K_{M_{**}} - K_{M_*}^T (K_M + \Sigma_M)^{-1} K_{M_*} \quad (13)$$

Denklemler 12 ve 13, sırasıyla, Denklem 10'da hesaplanan karmaşık kovaryans fonksiyonunu kullanarak tahmin edilen ortalama ve tahmin edilen varyansı hesaplar.

2.1.3. Seyrek Gauss Süreci ile Makine Öğrenimi

Bir robot sürekli olarak keşif yaparken ve örnek sayısı arttıkça, hesaplama yükü büyür ve benzer ölçümlerin kullanılması tahmin performansını düşürür. Bu durumu önlemek için, seyrek girişler kullanılabilir. Seyrek girişler, veri noktalarını daha az sayıda örnekle temsil etmeyi sağlar

ve böylece hesaplama yükünü azaltır. Seyrek sözde girişler kullanılarak hesaplanan GP’de ortalama ve varyans değerleri Denklem 14 ve 15’de verilmiştir [88,97].

$$\bar{f}_* = k_{*M}^T (K_{MM} + \sigma_n^2 I)^{-1} F_M \quad (14)$$

$$\mathbb{V}[f_*] = k(x_*, x_*) - k_{*M} (K_{MM} + \sigma_n^2 I)^{-1} k_{M*} + k_{*M} (K_{MM} + \sigma_n^2 I)^{-1} A_M (K_{MM} + \sigma_n^2 I)^{-1} k_{M*} \quad (15)$$

$$A_M = K_{MM} \sum K_{MM} \quad (16)$$

Denklemde K_{MM} seyrek sözde girişlerin kovaryans fonksiyonunu vermektedir. k_{*M} ise test girişleriyle indirgenen alt kümenin kovaryans fonksiyonunu vermektedir. Alt örneklem kümesini belirlemek için ise 5 farklı kümeleme yöntemi kullanılmıştır. Kullanılan kümeleme yöntemleri Bölüm 2.5’de anlatılmıştır. Bu kümeleme yöntemleri, veri noktalarını belirli özelliklere göre gruplandırmak için kullanılan istatistiksel ve veri madenciliği teknikleridir. Kümeleme performansını ölçmek ve en uygun kümeyi seçmek için siluet skor indeksi kullanılmıştır. Siluet skoru, bir veri noktasının kendi kümelerine ne kadar benzediğini ve diğer kümelerden ne kadar farklı olduğunu ölçen bir indekstir. Maksimum siluet skor değerine sahip olan veri seti, eğitim verisi olarak kullanılmıştır. Başlangıçta tüm veri seti kullanılarak bir model oluşturulmuş ve daha sonra modelin etkinliğini değerlendirmek amacıyla 1, 2, 3, 4 ve 5 adet veri noktası seçilmiştir. Bu şekilde, modelin farklı örneklem büyüklüklerinde ne kadar optimale yakın olduğu test edilmiştir.

2.1.4. Gauss Süreçlerinde Hiperparametre Optimizasyonu

Kovaryans fonksiyonunda yer alan hiper parametreler ise literatürde yaygın olarak log marjinal olasılık fonksiyonu ile hesaplanmaktadır. Denklem 17’de yer alan log marjinal olasılık fonksiyonunun kısmi türevini 0 yapan değerler hiper parametreleri vermektedir.

$$\log p(y|X) = -\frac{1}{2} y^T (K_M + \Sigma_M I)^{-1} y - \frac{1}{2} \log |K_M + \Sigma_M I| - \frac{n}{2} \log 2\pi \quad (17)$$

Denklem 18, gradyan tabanlı optimizasyon çözüm denklemini verir [96].

$$\frac{\partial}{\partial \theta_i} \log(Y|P) = \frac{1}{2} \text{tr} \left(((K_M^{-1} Y)(K_M^{-1} Y)^T - K_M + \Sigma_M^{-1}) \frac{\partial K_M + \Sigma_M}{\partial \theta_i} \right) \quad (18)$$

2.2. Eğitim Verilerinin Konumlarının Belirlenmesi

Gauss regresyon algoritmasının tahmin performansında eğitim verilerinin konumları büyük önem arz etmektedir. Bu bölümde öğrenilen farklı hiper parametrelerin etkisiyle tahmin

performansının deęiřimi arařtırılmıř ve eęitim verilerini toplamak iin bir guzergah planlama algoritması nerilmiřtir.

2.2.1. Hiper parametreleri ğrenmede Eęitim Verisinin Rol

Kovaryans fonksiyonu, iki konum arasındaki iliřkiyi lmek iin kullanılır. Gauss regresyonunda, bu kovaryans fonksiyonundaki hiper parametreler eęitim verilerinden faydalanılarak tahmin edilebilir. Kullanılan kovaryans fonksiyonu problemde farklılık gstermektedir. Literatrde yaygın olarak SE (Squared Exponential), RQ (Rational Quaternions), lineer ve Matern kovaryans fonksiyonu kullanılmaktadır. Sırasıyla Denklem 19, 20, 21 ve 22 bu fonksiyonları vermektedir. Denklem 19'da σ^2 ve l^2 hiper parametreleri vermektedir. SE, verilerin daha yumuřak deęiřtięi durumlar iin kullanılmaktadır. RQ ise deęerlerin daha hızlı deęiřtięi verileri tahmin etmek iin kullanılmaktadır. Denklem 20'de yer alan α bu deęiřimin lęini vermektedir. Periyodik kovaryans fonksiyonunda periyodik olarak veriler deęiřmektedir. Denklem 21'de yer alan p , periyodun deęiřimini vermektedir. Denklem 22'de yer alan K_v ise Bessel fonksiyonundan tretilmektedir. v deęeri 1, 3/2, 5/2 gibi deęerler almaktadır. v , 3/2 olduęu zaman denklemin dzenlenmiř hali Denklem 23'de verilmektedir. Kovaryans fonksiyonları, beraberinde kullanılabilmektedir. Denklem 24'de SE ile periyodik kovaryans fonksiyonunun denklemini almaktadır. Bu durumda verilerde yksek deęiřim periyodik olurken, her periyotta yumuřak deęiřimler olmaktadır. Bu alıřmada en iyi sonular Matern kovaryans fonksiyonu ile elde edilmiřtir. Bu alıřmada en iyi sonular Matern kovaryans fonksiyonu ile elde edilmiřtir.

$$k_{SE}(x, x') = \sigma^2 \exp\left(-\frac{(x - x')^2}{2l^2}\right) \quad (19)$$

$$k_{RQ}(x, x') = \sigma^2 \left(1 + \frac{(x - x')^2}{2\alpha l^2}\right)^{-\alpha} \quad (20)$$

$$k_{Per}(x, x') = \sigma^2 \exp\left(-\frac{2\sin^2(\pi|x - x'|/p)}{l^2}\right) \quad (21)$$

$$k_{Matern}(x, x') = \frac{2^{1-v}}{\Gamma(v)} \left(\frac{\sqrt{2}v(x - x')}{l} K_v \exp\left(-\frac{\sqrt{2}v(x - x')}{l}\right)\right) \quad (22)$$

$$k_{v=\frac{3}{2}}(x, x') = \left(1 + \frac{\sqrt{3}(x - x')}{l}\right) \exp\left(-\frac{\sqrt{3}(x - x')}{l}\right) \quad (23)$$

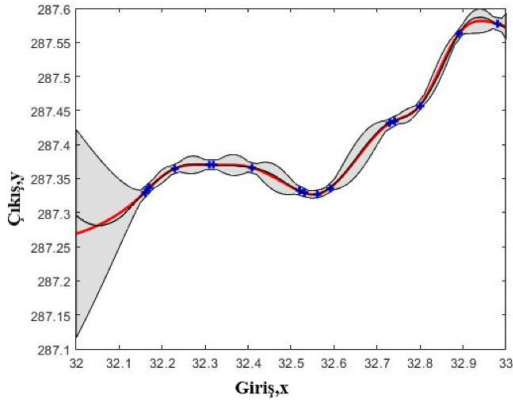
$$k_{LocalPer(x-x')}k_{Per}(x, x')k_{SE}(x, x') = \sigma^2 \exp\left(-\frac{2\sin^2(\pi|x-x'|/p)}{l^2}\right)k_{SE}(x, x')\exp\left(-\frac{(x-x')^2}{2l^2}\right) \quad (24)$$

Eğitim verileri kullanılarak σ^2 ve ℓ hiper parametrelerinin tahmin edilmesiyle, Gauss regresyon modelleri, kovaryansı etkin bir şekilde belirleyebilir ve farklı konumlar arasındaki ilişkiyi nicelendirir. Bu hiper parametreler, kovaryans fonksiyonunun davranışını şekillendirir ve Gauss regresyon modelinin genel tahmin gücünü ve genelleme yeteneklerini etkiler.

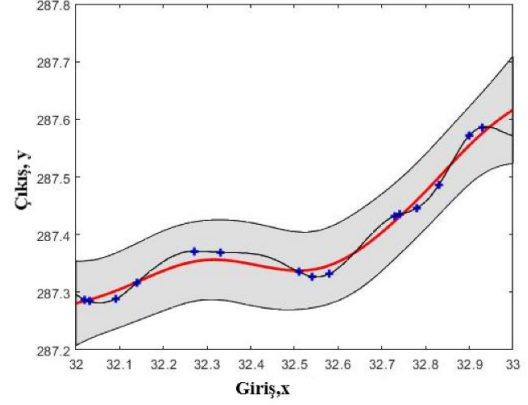
Bu bölüm, eğitim verilerini edinmek için uygun konumların seçilmesinin önemini vurgular. Şekil 2.2. (a-d)'de gösterilen örnekler aracılığıyla, farklı konum kümelerinden aynı sayıda eğitim girişinin Gauss Süreci (GP) modelleri için önemli ölçüde farklı performansa yol açabileceğini gösteriyoruz. ℓ uzunluk ölçeği parametresi, iki konum arasındaki korelasyonu belirler. Daha yüksek değerler daha büyük bir mekansal korelasyonu ve dolayısıyla temel parametrenin daha yavaş değişimini, daha düşük değerler ise daha az korelasyonu ve dolayısıyla kısa aralıklarda daha büyük değişiklikleri gösterir. Modelin belirsizliği, σ^2 hiper parametresinin değerinden ölçülür. Daha yüksek değerler, daha yüksek bir belirsizlik derecesine sahip konumları gösterir. Şekil 2.2.b ve Şekil 2.2.c incelendiğinde yanlış tahmin edilen ℓ değeriyle alanın tamamında yüksek belirsizlik olduğu görülmüştür. Şekil 2.2.a ve Şekil 2.2.d'de ise ℓ değerinin daha doğru tahmin edilmesi nedeniyle gerçek ölçümler ile tahmin edilen değerlerin birbirine daha yakın olduğu görülmüştür. Böylece ölçüm noktalarında ki belirsizlik minimuma inerken ölçüm noktasından uzaklaştıkça belirsizliğinde arttığı gözlenmiştir.

Şekil 2.2'deki sıcaklık veri seti, NASA'dan alınmıştır [98] ve boylam 32-33 arasına ve enlem değeri 42. paralele karşılık gelir. Şekilde, kırmızı çizgi tahmini ortalama değeri, gri alan belirsizliği ve siyah çizgi veri setindeki gerçek sıcaklık veri noktalarını temsil eder. Şekil 2.2. c, en yüksek ℓ değerini gösterir ve bu da daha yüksek bir hata oranına ve daha az değişen bir eğriye yol açar. ℓ değeri, Şekil 2.2.a ve Şekil 2.2. d'de aynı olmasına rağmen, Şekil 2.2. d'de daha düşük bir RMSE değeri elde edilir. En yüksek belirsizliğe sahip konum, σ^2 parametresi tarafından belirtildiği gibi, Şekil 2.2. c'de gözlemlenir ve bu da daha büyük bir gri alanı ortaya çıkarır.

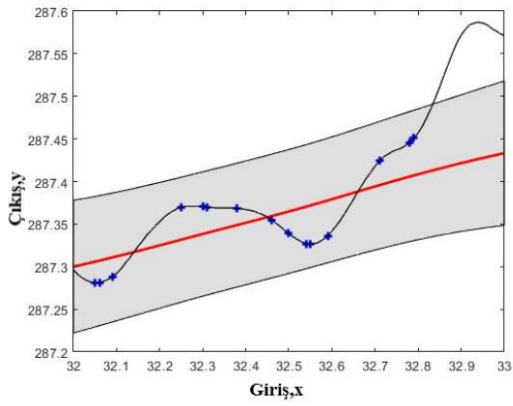
Şekil 2.2'deki sonuçlar, hiper parametrelerin doğru öğrenilmesinin daha düşük hata oranlarına ve daha kesin tahminlere yol açabileceğini vurgular. Uygun eğitim konumlarını seçerek ve hiper parametreleri etkili bir şekilde tahmin ederek, GP modeli, verideki temel desenleri yakalayabilir, belirsizliği azaltabilir ve tahmin performansını iyileştirebilir. Şekil 2.2, tek değişkenli analiz için eğitim verilerinin konumlarını gösterir ve veri konumunun tahmin performansı üzerindeki etkisini vurgular.



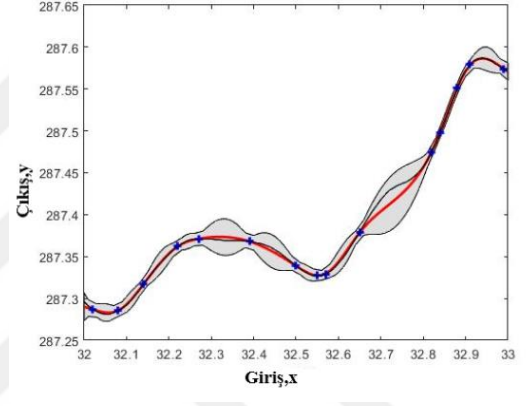
(a) $\ell=-0.56$, $\sigma=1.34$, $RMSE=0.039$



(b) $\ell=-0.11$, $\sigma=1.42$, $RMSE=0.069$



(c) $\ell=0.88$, $\sigma=2.25$, $RMSE=0.091$



(d) $\ell=-0.56$, $\sigma=1.39$, $RMSE=0.013$

Şekil 2.2. 4 farklı rastgele senaryo için hiper parametre değişimi

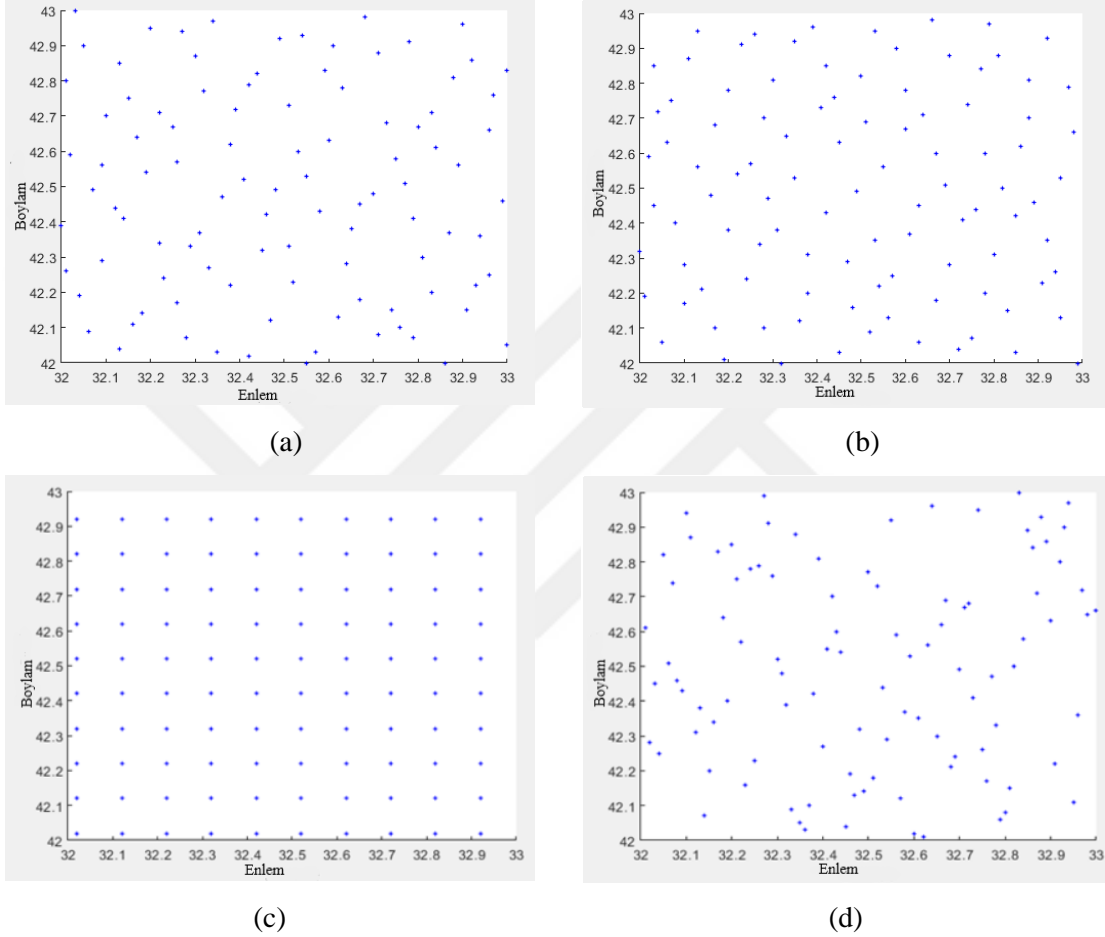
2.2.2. Uzamsal Alanda Aday Yol Noktalarının Dizilimi

Büyük bir veri setinde her veri noktası için fayda fonksiyonunun hesaplanması, hesaplama yoğunluğunu önemli ölçüde artırır. Bu nedenle, bu çalışma, farklı aday yol noktası dizileri kullanarak algoritmanın etkinliğini incelemiştir. Rastgele, yarı rastgele Halton ve Sobol dizileri, ve düzenli bir dağılım, bu aday rota noktalarını belirlemek için kullanılmıştır. Halton ve Sobol dizileri, sayıların ikili gösterimlerini kullanarak bir yarı rastgele olmayan bir dağılım elde ederler.

Halton dizisinde, dağılım 2, 3, 4 veya 5 gibi aralık değerlerini seçerek gerçekleştirilir. Örneğin, aralık değeri 2 olarak ayarlandığında, dizi 0 ile 100 arasında eşit aralıklı örnekler olan 2, 4, 8, 16 gibi örnekler üretir. Aralık değeri arttıkça, örnekler birbirlerine daha yakın hale gelir, ancak gruplar arasında boşluklar oluşur. Ancak, bu çalışmada dengeli bir dağılım elde etmek amacıyla daha düşük bir aralık değeri seçilmiştir. Bu nedenle, yatay ekseninde aralık değeri olarak 2 ve dikey ekseninde 3 seçilmiştir. Bu seçim, dengesiz bir dağılıma neden olacak olan aday rota noktalarını yalnızca çaprazlardan seçmeyi önler.

Öte yandan, Sobol dağılımı, istenen bir alan boyutu içinde dengeli bir dağılım üretir. Bu çalışmada, Halton, Sobol, Düzenli ve Rastgele Diziler olmak üzere dört farklı yöntem kullanılarak 100 aday rota noktası seçilmiştir. Şekil 2.3, bu dizilerin görsel temsillerini sağlar.

Halton, Sobol, Düzenli ve Rastgele Dizileri kullanarak, bu çalışma 100x100 bir alanda aday rota noktalarının dengeli bir dağılımını sağlamayı amaçlamaktadır. Bu seçim süreci, farklı dağılım yöntemlerinin ve sonuçlar üzerindeki etkilerinin kapsamlı bir şekilde araştırılmasına olanak tanır.



Şekil 2.3. Yol Noktası Dizilimleri a) Halton Dizilimi b) Sobol Dizilimi c) Düzenli Dizilim d) Rastgele Dizilim

2.2.3. Eğitim Verisi Toplamak İçin Geliştirilen Yeni Bir Güzergah Planlama Algoritması

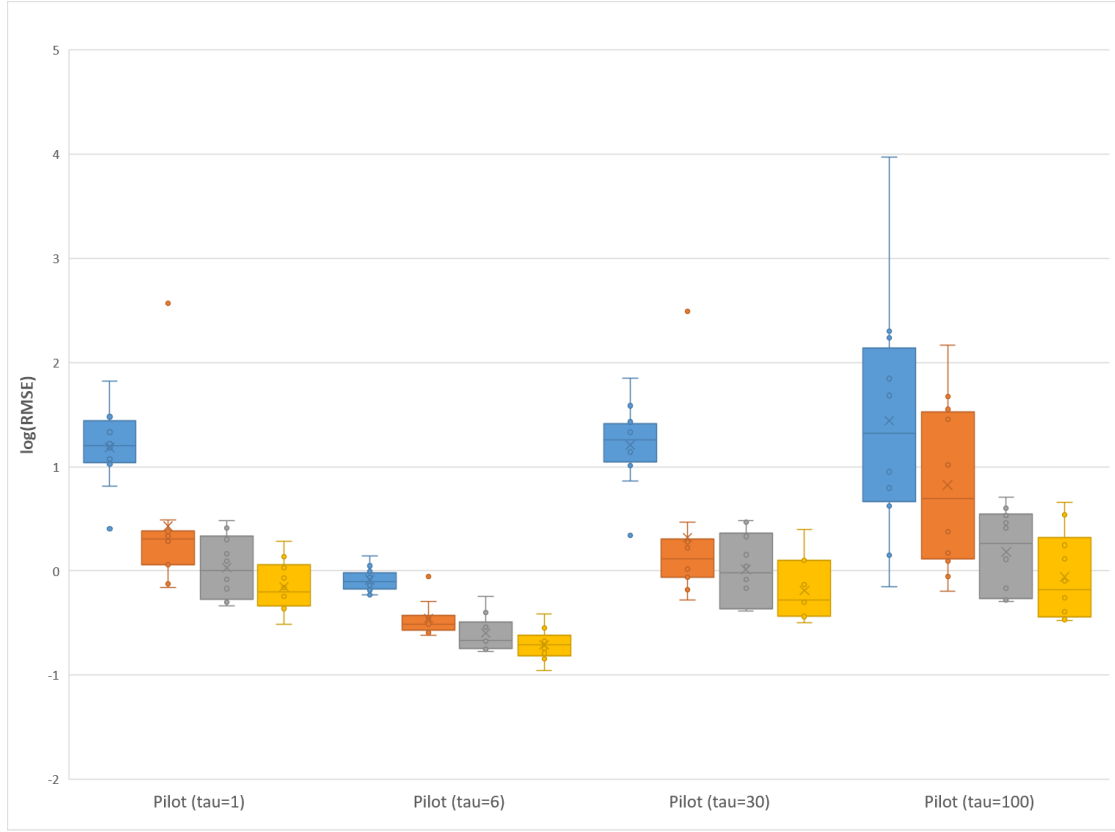
Öğrenme hiper parametrelerinde tahmin performansı, eğitim verilerinin konumları büyük bir önem arz etmektedir. Ancak, eğitim verilerini elde etmek için örnekleme konumlarını seçme sorunu literatürde nadiren ele alınmaktadır. Bu boşluğu doldurmak için, bu bölümde robotun sınırlı bütçesi dikkate alarak bilgilendirici eğitim verileri toplamayı amaçlayan bir rota planlama algoritması önerilmektedir. Yaklaşımı, seyahat süresi kısıtlamalarını dikkate alarak tüm alan boyunca dengeli bir örnek dağılımı elde etmeye odaklanmaktadır.

Algoritmanın ilk aşamasında, verilen seyahat süresi içinde ulaşılabilen aday yol noktaları arasından en uzak noktaları seçilmektedir. Bu, ilgi alanı içinde geniş bir konum yelpazesini kapsadığımızı garanti etmektedir. İkinci aşamada, sol alt köşeden başlayarak seçilen yol noktalarından en yakınına ziyaret ederek miyop bir güzergah planlama algoritması oluşturulmuştur. Böylece, verilen kısıtlar içinde mümkün olduğunca fazla ilgili yol noktasının ziyaret edilmesi sağlanmaktadır. Toplanan bilgiyi maksimize etmek ve alanı kapsamlı bir şekilde keşfetmek için, robotun bütçesi tamamen kullanılmış olana kadar ziyaret edilen aday yol noktalarından en uzak nokta rota planlama algoritmasına dahil edilmektedir. Sonuçlar Kemna vd.[52] tarafından önerilen yöntemle karşılaştırılmaktadır. Kemna vd. yol noktalarını rastgele oluşturarak softmax fonksiyonu ile hesaplanan yol noktasını ziyaret etmektedir. Denklem 25’de kullanılan softmax fonksiyonunu vermektedir.

$$p(x_i) = \frac{e^{D(x_i)/\tau}}{\sum_j e^{D(x_j)/\tau}} \quad (25)$$

Denklemden $D(x_i)$ oklid mesafesini vermektedir. τ ise rastgeleliğin bir ölçüsü olarak verilmektedir. Bu değer arttıkça $p(x_i)$ değeri yani aday yol noktasının seçim ihtimali azalacaktır. Başka bir deyişle τ değeri aday yol noktasının seçiminin rastgeleliğini belirlemektedir. Bu çalışmada farklı τ değerleri 10 veri setine uygulanarak farklı τ değerlerinde ortamın öğrenilme performansı ölçülmüştür. Şekil 2.4’de dört farklı "Pilot (τ)" değeri için hesaplanan log(RMSE) değerleri kutu grafiklerle görselleştirilmiştir. X ekseninde dört farklı "Pilot (τ)" değeri ($\tau = 1, \tau = 6, \tau = 30, \tau = 100$) yer alırken, Y ekseninde ise log(RMSE) değerleri bulunmaktadır. Sırasıyla 110, 330, 550 ve 1100 km bütçeyle hesaplanan RMSE değeri, mavi, turuncu, gri ve sarı kutu grafiğiyle gösterilmiştir. Pilot ($\tau = 1$) için kutu grafikleri, yüksek log(RMSE) değerlerine sahip olduğunu ve bazı grupların aykırı değerlere sahip olduğunu göstermektedir. Özellikle mavi renkli grup, geniş bir dağılıma ve birkaç aykırı değere sahiptir. τ değeri 6 iken elde edilen sonuçlar genellikle daha tutarlıdır ve medyan değeri düşüktür, bu da modelin stabil performans gösterdiğini işaret etmektedir. τ değeri 30 için mavi grup yine yüksek log(RMSE) değerleriyle dikkat çekerken, diğer gruplar daha düşük ve tutarlı sonuçlar vermektedir. τ değerinin 100 olduğu durumda, log(RMSE) değerleri yine oldukça dağınık olmakla birlikte, özellikle turuncu grup birkaç aykırı değere sahiptir. Bu grafik, farklı τ değerlerinin model performansını nasıl etkilediğini anlamak için kullanışlıdır. τ değeri arttıkça bazı modellerin daha fazla varyasyon içerebileceği ve hatta daha yüksek hata oranları gösterebileceği gözlemlenmektedir. Bu durum, modelin belirli τ değerleri altında daha stabil veya daha performanslı çalışıp çalışmadığını değerlendirmek için önemlidir. Genel olarak, grafik, log(RMSE) değerlerinin varyasyonunu, median ve ortalama değerlerini etkili bir şekilde sunarak, model optimizasyon çalışmalarına ışık tutmaktadır. Bu veriler, farklı τ

değerlerinin model doğruluğu üzerindeki etkilerini karşılaştırmak ve ideal τ değerini seçmek için değerlidir. Elde edilen sonuçlarda mevcut veri seti için ideal τ değeri 6 olarak belirlenmiştir.



Şekil 2.4. 10 farklı veri setinde softmax fonksiyonuna bağlı RMSE değişimi

Kemna vd.'nin önermiş olduğu pilot yöntemde yol noktalarına giderken bir güzergah planlama algoritması oluşturulmamıştır. Bu nedenle önerilen yönteme miyop bir güzergah planlama algoritması uygulanarak yöntemin başarımının artırılması hedeflenmiştir. Yöntemlerin başarısı, kök ortalama hatanın karesi (Root Mean Square Error –RMSE) kullanılarak hesaplanan Denklem 26 kullanılarak değerlendirilmiştir.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - \widehat{Y}_i)^2} \quad (26)$$

Bu tez çalışmasında eğitim verilerini toplamak için önerilen yöntemin sözde kodu Algoritma 1'de verilmiştir. Algoritma, başlangıç düğümü olarak belirtilen X_s ile başlar. Satır 6'da, aday yol noktaları ile başlangıç düğümü arasındaki mesafe hesaplanır. Sonra satır 8'de, başlangıç düğümünden en uzak yol noktası seçilir ve oluşturulan yolu temsil eden düğüm kümesine, P olarak belirtilen, eklenir. Başlangıçta boş olan X_a kümesi, bu aşamada seçilen düğümleri temsil eder. Satır 10'da, seçilen rota noktası X_a kümesine eklenir. Ek olarak, bu noktaya seyahat etmek için gereken

bütçe, satır 11'de robotun seyahat bütçesine eklenir. Sürekli olarak, robotun seyahat bütçesi tükenene kadar X_a kümesinden en uzak aday rota noktaları seçilerek işlem devam eder. Rota planlama aşaması, P kümesine en yakın örneklerin toplandığı satırlar 13'ten 24'e kadar gerçekleşir. Satır 21'de, P kümesinin boş olup olmadığı kontrol edilir. Eğer boşsa ve robotun seyahat bütçesi henüz bitmemişse, algoritma X_a kümesinden en uzak en yeni örneği seçmek ve onu P kümesine eklemek için satır 4'e geri döner. Bu süreç, seyahat bütçesi tükenene kadar devam eder, seyahat süresi dolunca döngü sona erer.

Algoritma 1: Eğitim Verilerini Toplamak için Önerilen Yöntem

```

1: Olası 100 yol noktasının belirlenmesi (Bölüm 2.2.2)
2:  $P \leftarrow X_S$ 
3:  $X_S \leftarrow$  Başlangıç Düzümü
4: Seyahat mesafesi  $\leftarrow 0$ 
5: döngü Seyahat mesafesi < bütçe yap
6:   for Aday yol noktası  $\leftarrow X_C$  yap
7:      $m \leftarrow$  Oklid ( $X_S, X_C$ )
8:      $m^* \leftarrow$  argmak(Oklid ( $X_S, X_C$ ))
9:    $P \in m^*$ 
10:   $X_S \leftarrow X_C$ 
11:  Seyahat mesafesi  $\leftarrow m^*$ 
12: döngü bitti
13:  $X_S \leftarrow$  Başlangıç Düzümü
14: Seyahat süresi  $\leftarrow 0$ 
15: döngü Seyahat süresi  $\leq$  bütçe yap
16:   for P güzergahındaki yol noktası yap
17:     Mesafe  $\leftarrow$  Oklid ( $X_S, P$ )
18:     Seyahat süresi  $\leftarrow$  argmin(Oklid ( $X_S, P$ ))
19:     Optimal yol  $\leftarrow P$ 
20:     eğer isempty (P) sonra
21:       4. Satıra git
22: döngü bitti

```

2.3. Keşif ve Sömürü Dengesinde Kullanılan Fayda Fonksiyonlarının Önemi

Literatürde yeni örnek noktalarını belirlemek için çeşitli ödül fonksiyonları kullanılmaktadır. Gerçek bir veri setinde 2 ROI alanı olan durumda, UCB ödül fonksiyonunun ROI alanlarındaki performansı araştırılmıştır [45]. Bu tez çalışmasında ise UCB, seviye belirleme ve entropi ödül fonksiyonlarının test ortamlarındaki performansı incelenmiştir. 1, 2, 3 ve 4 ROI alanına sahip veri setleri kullanılarak test ortamları oluşturulmuştur. Bu bölüm, kullanılan ödül fonksiyonları ve yeni örnekleri belirlemek için kullanılan güzergah planlama algoritması ve başlangıçta tahmin edilen haritanın belirsizliğini azaltmak için uygulanan çapraz doğrulama yöntemi hakkında ayrıntılı bilgi içermektedir.

2.3.1. Keşif ve Sömürü Dengesinde Yaygın Olarak Kullanılan Ödül Fonksiyonları

ROI alanlarını belirlemek için Gauss regresyon kullanılarak çeşitli ödül fonksiyonları kullanılmaktadır. Literatürde yaygın olarak entropi azaltma, UCB ve seviye belirleme ödül fonksiyonları kullanılmaktadır. Bu çalışmada bu ödül fonksiyonları kullanılarak güzergah planlama algoritması geliştirilmiştir.

Denklem 27 UCB ödül fonksiyonunu vermektedir. Denklemde μ_{n-1} ve σ_{n-1} Gauss regresyon ile hesaplanan ortalama ve varyans değerlerini vermektedir. β_n Denklem 28 ile hesaplanmaktadır. Burada “n” örnek sayısını temsil etmektedir. “ δ ” ise 0 yada 1 değerini almaktadır. D örneklem uzayının boyutunu temsil etmektedir. Denklemde yer alan “ $q(x_n)$ ” ise Denklem 29 ile hesaplanmaktadır. Denklemde “o” engel olup olmama durumunu kontrol etmektedir. “ $q(x_n)$ ” ortamda engel olması durumunda ∞ değerini almaktadır. Denklemde μ_{n-1} ROI alanlarından örnek toplamak için kullanılmaktadır. Bu değer Gauss regresyonda tahmin edilen veriyi verdiği için maksimum olması durumu ROI alanlarını yani ölçülen değer maksimum olduğu konumu vermektedir. σ_{n-1} ise Gauss regresyondan elde edilen belirsizliğin bir ölçüsüdür. Bu nedenle keşif yapmak için kullanılan bir parametredir. Yani ortamın belirsizliğinin yüksek olduğu zamanlarda ilk olarak yeni örnek noktaları ortamın belirsizliğini düşürecek konulardan seçilirken düşük olması durumunda yeni örnek noktası ortalamanın yüksek olduğu ROI alanlarından toplanacaktır.

$$L_{UCB} = \mu_{n-1}(p_n) + \sqrt{\beta_n \varphi_{n-1}(p_n)} - q(x_n) \quad (27)$$

$$\beta_n = 2 \log \left(\frac{|d|n^2\pi^2}{6\delta} \right) \quad (28)$$

$$q(x_n) = \begin{cases} \infty & \text{if } p_n \notin D \text{ or } x_n \in o \\ 0 & \text{otherwise} \end{cases} \quad (29)$$

Denklem 30 ise seviye belirleme ödül fonksiyonunu vermektedir [99]. “h” ROI alanlarının eşik değerini temsil etmektedir. Eşik değeri klorofil gibi belirlenmek istenen alanın özelliğine göre değişmektedir. Zhanh vd. [35] yaptıkları çalışmada klorofil miktarının yüksek olduğu bölgeyi belirleyebilmek için HSCS(Heterogeneous Strategy Cooperative Sampling) yöntemini kullanmışlardır. Bu yöntemde deniz ortamı için alanın tamamını keşfetmesi ve exploration exploitaion arasında bir denge kurması için eşik değerini 0.5’den küçük bir değer olarak seçilmesi gerektiğini belirtip 0.4 olarak seçmişlerdir. Bu nedenle kullanılan veri seti için h değeri 0.4 olarak alınmıştır. Bu ödül fonksiyonun ortam öğrenildikten sonra eşik değerinin üzerindeki alanlarda örnekler toplanır. Denklem de amaç alanı sınıflandırmaktır.

Hesaplanan ödül fonksiyonunun değeri eşik değerinden yüksekse eşik değerinin üzerindeki konuları barındıran bir kümeye yeni örneğin konumu atılmaktadır. Eğer belirlenen eşik değerinin

altındaysa eşik değerinin altındaki konumları barındıran bir kümeye atılmaktadır. Yeni örnekler seçilirken bu iki kümeye ait olmayan yeni örnekler toplanmaktadır. Bu nedenle genellikle düşük belirsizliğe sahip haritalarda yeni örnekler ROI alanının etrafından toplanmaktadır. Bu çalışmada seviye belirleme ödül fonksiyonunun ROI alanındaki performansı araştırılmıştır.

$$L_{LS} = \beta_t^2 \varphi_{n-1}(x) - |\mu_{t-1}(x) - h| - q(x_n) \quad (30)$$

Entropi ödül fonksiyonu ise Denklem 31 kullanılarak hesaplanır. Denklemde . σ_{n-1} Gauss regresyondan elde edilen belirsizlik miktarını temsil etmektedir. Bu nedenle maksimum ödüle sahip olan konum belirsizliğin en yüksek olduğu konumu temsil etmektedir. Bu nedenle bu ödül fonksiyonuyla genellikle alanın tamamından örnekler toplanmaktadır.

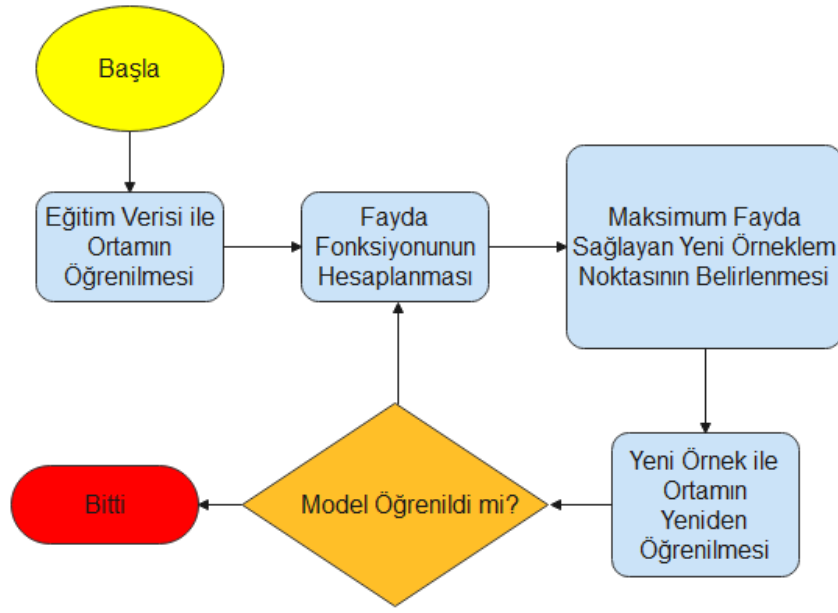
$$L_{entropi} = -\varphi_{n-1} * \log(\varphi_{n-1}) \quad (31)$$

2.3.2. Ödül Fonksiyonları Kullanılarak Yeni Örnek Noktasının Seçimi

Veri bilimi ve makine öğrenmesi süreçlerinde, örnek verilerin seçimi ve optimize edilmesi modelin performansını doğrudan etkiler. Ödül fonksiyonları kullanılarak örnek veri noktalarının seçimindeki önemi belirlemek amacıyla yapılan çalışmalarda, başlangıç aşamasında mevcut eğitim verisi kullanılarak model, ortamın temel özelliklerini öğrenir. Bu süreç, modelin mevcut verilerle eğitilmesi ile başlar ve devam eder. Model eğitildikten sonra, yeni örneklerin bilgilendiriciliğini ölçmek için fayda fonksiyonu hesaplanır. Fayda fonksiyonu, yeni veri noktalarının modele katacağı bilgi değerini değerlendirir; Bölüm 2.3.2'de literatürde yaygın olarak kullanılan fayda fonksiyonları verilmektedir.

Fayda fonksiyonu hesaplandıktan sonra, maksimum fayda sağlayacak olan yeni örnek belirlenir. Bu, modele en fazla katkıda bulunacak olan veridir. Bu yeni örnek, mevcut eğitim veri kümesine eklenir ve model bu yeni verilerle birlikte yeniden eğitilir. Modelin yeniden eğitilmesi, bilgi tabanını genişleterek genel performansını artırır.

Her yeni örnek eklendiğinde ve model yeniden eğitildiğinde, modelin istenilen öğrenme seviyesine ulaşip ulaşmadığı kontrol edilir. Bu aşamada belirli durdurma kriterleri kullanılır; bu kriterler genellikle hata oranları veya belirli başarı ölçütleri olabilir. Eğer model belirlenen kriterleri sağlıyorsa, döngü sonlandırılır. Aksi takdirde, yeni örneklerin seçimi ve modelin yeniden eğitilmesi döngüsü devam eder. Bu süreç, modelin optimize edildikçe performansının arttığı ve öğrenme kapasitesinin genişletildiği bir döngüdür. Şekil 2.5 yeni örnek noktasının belirlenmesi kullanılan yöntemin akış diyagramını vermektedir.



Şekil 2.5. Yeni Örnekleme Noktalarının Belirlenmesi Oluşturulan Akış Diyagramı

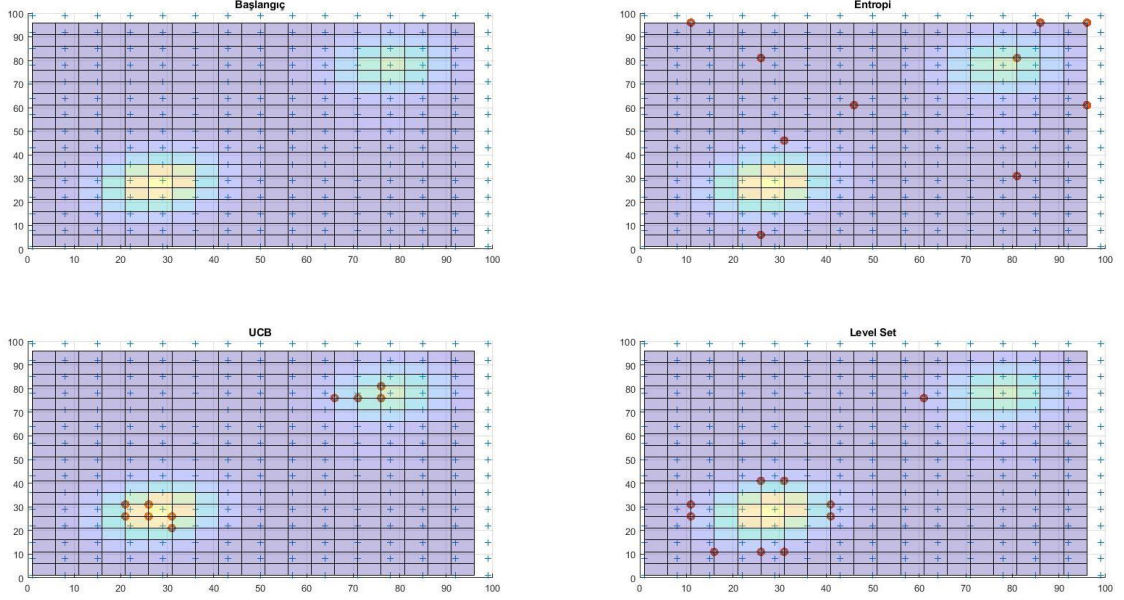
Şekil 2.6'da gösterilen grafikler, UCB, entropi ve seviye belirleme ödül fonksiyonları kullanılarak 10 örnek noktasının seçimini göstermektedir. Sol üst köşe mevcut tahmin edilen başlangıç haritasını vermektedir. Sağ üst köşede entropi ödül fonksiyonuyla toplanan 10 örnek, sol alt köşede UCB ödül fonksiyonuyla toplanan 10 örnek ve sağ alt köşede seviye belirleme ödül fonksiyonuyla toplanan 10 örnek verilmektedir. Bu grafikler, her bir ödül fonksiyonunun farklı stratejilerle örnek noktalarını nasıl seçtiğini görselleştirir. Entropi, belirsizliğin bir ölçüsü olarak kullanılır ve belirsizliği minimize edecek şekilde örnekler seçilir. Şekilde görüldüğü gibi, entropi ödül fonksiyonu tüm alana yayılan örnekler seçmiştir. Bu strateji, belirsizliğin en yüksek olduğu bölgelerden örnekler toplayarak belirsizliği azaltmayı hedefler. Bu seçim sürecinde, bir örneğin toplanması ile entropinin azalması hesaplanır ve ardından ikinci örnek ona göre seçilir. Bu nedenle, örnekler genellikle birbirine en uzak konumlardan seçilir.

UCB ödül fonksiyonu ise örnekleri doğrudan ROI alanlarından toplar. ROI alanları, modelin yüksek belirsizlik veya bilgi kazanımı sağlayabileceği alanlar olarak tanımlanır. UCB, bu alanlardan örnekler toplayarak modelin belirsizliğini ve doğruluğunu optimize etmeyi amaçlar. Şekil 2.6'da gösterildiği gibi, UCB ödül fonksiyonunda örnekler, belirli yoğunluk bölgelerinden seçilmiştir.

Seviye belirleme ödül fonksiyonu ise ROI alanının etrafında örnekler toplar ve belirlenen bir eşik değerine göre seçim yapar. Eşik değeri, modelin belirli bir güven seviyesinde bilgi kazanmasını sağlar.

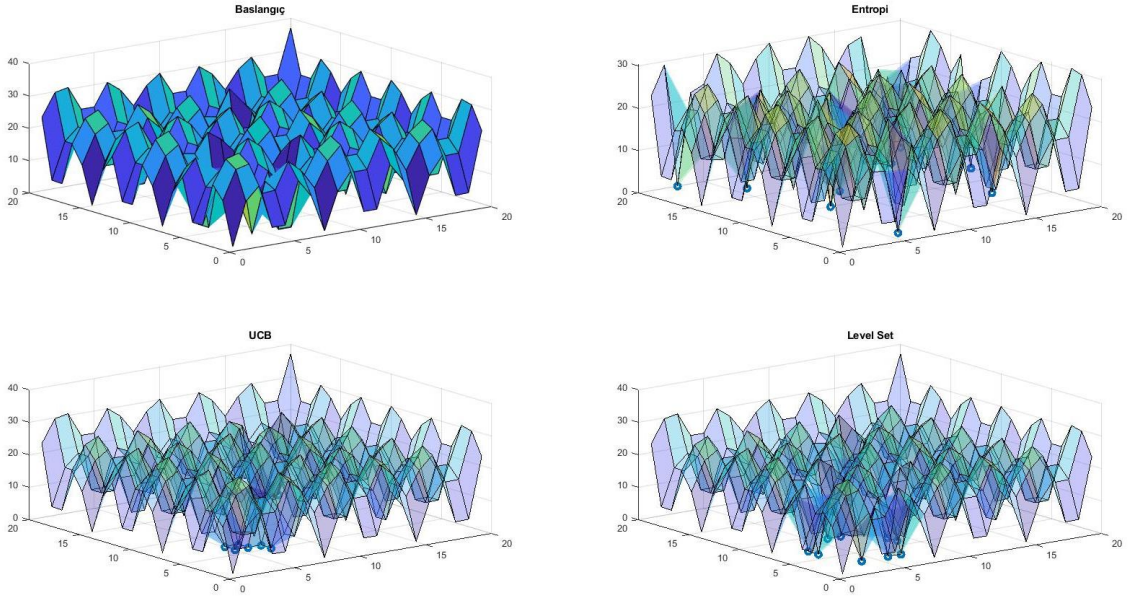
Grafikler, seviye belirleme ödül fonksiyonunun, örnek seçiminde eşik değerine göre örneklerin toplandığını ve bu örneklerin ROI alanının çevresinde yoğunlaştığını göstermektedir.

Seviye belirleme ödül fonksiyonu, seçilen noktaların etrafındaki yerel bilgiyi optimize ederek modelin doğruluk ve güven düzeyini artırmayı hedefler. Bu stratejiler, örnek seçiminde kullanılan ödül fonksiyonlarının farklı avantajlarını ve kullanım alanlarını vurgulamaktadır.



Şekil 2.6. Ödül fonksiyonları kullanılarak 10 örnek noktanın seçimi

Şekil 2.7'de, farklı ödül fonksiyonları kullanılarak seçilen 10 örneklem noktası ile belirsizliğin nasıl değiştiği gözlemlenmiştir. Entropi ödül fonksiyonu kullanıldığında, başlangıçta maksimum değeri 40 olan belirsizliğin 30'a düştüğü gözlemlenmiştir. Bu, entropi fonksiyonunun belirsizliği minimize etme konusunda etkili olduğunu ve tüm veri alanında daha dengeli bir belirsizlik azaltma sağladığını göstermektedir. Entropi ödül fonksiyonu, belirsizliği düşürmek için geniş bir alan boyunca örnekler seçerek genel belirsizlik profilini iyileştirmektedir. Buna karşın, UCB ödül fonksiyonu ile yapılan seçime baktığımızda, yalnızca örneklerin toplandığı konumlarda belirsizliğin düştüğü ancak genel belirsizlik ortalamasının önemli ölçüde değişmediği görülmüştür. UCB fonksiyonu, örnekleme alanlarında yüksek performans sağlarken genel belirsizlik üzerinde sınırlı bir etkiye sahiptir. Benzer şekilde, seviye belirleme ödül fonksiyonu kullanıldığında da sadece örneklerin toplandığı konumlarda belirsizliğin düştüğü ve genel belirsizlik ortalamasının belirgin bir şekilde değişmediği gözlemlenmiştir. Seviye belirleme, belirli bir eşik değerine göre örnekler seçer ve ROI alanının etrafında yoğunlaşır. Bu yöntem, modelin belirli bölgelerde daha hassas olmasını sağlarken, genel belirsizlik azaltma konusunda UCB gibi sınırlı bir etkiye sahiptir. Bu grafikler, Entropi ödül fonksiyonunun belirsizliği genel olarak düşürmede daha etkili olduğunu, UCB ve seviye belirleme ödül fonksiyonlarının ise belirli bölgelerde yoğunlaşarak belirsizliği azalttığını göstermektedir.



Şekil 2.7. Ödül fonksiyonları kullanılarak 10 örnek noktanın seçimiyle varyans haritasının değişimi

2.3.3. ROI Alanlarından Örnek Toplamak için Geliştirilen Güzergah Planlama Algoritması

Çevresel izlemede, kirlilik, tuzluluk ve sıcaklık gibi çevresel özelliklerin en yüksek değerlere sahip alanları sürekli olarak izlemek esastır. Bu yüksek özellik değerlerine sahip alanlar, ROI alanı olarak bilinir. Bir USV kullanarak bölgeyi keşfederken, hem ROI alanlarındaki örnek sayısını artırabilen (sömürü) hem de alanı düşük bir hata oranıyla öğrenebilen (keşif) bir yol planlama algoritmasına ihtiyaç vardır. Çevreyi öğrenmek için genellikle Gauss regresyonu kullanılır. Bu bölümde, Gauss regresyonu ile hesaplanan ödül fonksiyonları, belirsizliği en aza indiren ve ROI alanlarını belirleyen bir yol planlama algoritması geliştirmek için bir güzergah planlama algoritması geliştirilmiştir.

Bu çalışmada, olası bir yol belirlemek için 10.000 veri kümesinden 2000 aday örnek noktası rastgele seçilmiştir. Bu örnek noktalarının ödül değerleri hesaplandı ve aday örnek noktaları, ödül değerlerine göre azalan sıraya göre sıralanmıştır. Ödül fonksiyonları bölüm 2.3.1'de tanımlanmıştır. Daha sonra, 10-oklid uzaklığındaki en yüksek değerli örneğe ulaşmak için gereken yörünge parametreleri v ve θ hesaplanmıştır. " t_s ", başlangıç konumunu belirtmektedir [45].

$$\theta(t) = \theta(t_s) + \theta \quad (32)$$

$$p_x(t) = p_x(t_s) + v(t)\cos\theta(t) \quad (33)$$

$$p_y(t) = p_y(t_s) + v(t)\sin\theta(t) \quad (34)$$

Denklemler 32, 33 ve 34 kullanılarak θ değeri hesaplandıktan sonra, beş yörünge oluşturmak için $\theta \in \{[\theta - \pi \quad \theta - \pi/2 \quad \theta \quad \theta + \pi/2 \quad \theta + \pi]\}$ aralığında yeni bir p_x ve p_y değer seti hesaplanmıştır. Daha sonra maksimum ödüle sahip yörünge seçilerek bütçe tamamlanana kadar seçim devam etmiştir.

2.3.4. Çapraz Doğrulama Yöntemi

Çapraz Doğrulama, makine öğrenimi modellerinin performansını değerlendirmek için kullanılan bir yöntemdir. Bu yaklaşım, modelin farklı veri setleri üzerinde nasıl performans gösterdiğini değerlendirerek, genel bir genelleme kapasitesine sahip olup olmadığını belirlemeyi amaçlar. Temel olarak, modelin aşırı öğrenme (overfitting) yapmasını veya yetersiz öğrenme (underfitting) sorununu tespit etmeye yardımcı olur.

Çapraz doğrulama yönteminin adımları şu şekildedir [100]:

Veriyi Bölme: Verilerinizi belirli sayıda alt küme veya "fold" olarak adlandırılan parçalara bölünür. En yaygın kullanılan yöntem 10-fold çapraz doğrulama yöntemidir. Ancak 5-fold, 20-fold gibi farklı sayılar da kullanılabilir.

Model Eğitme ve Doğrulama: İlk adımda, verilerin bir kısmını eğitim (training) seti, kalan kısmını doğrulama (validation) seti olarak kullanılır. Örneğin, verilerinizi 5 parçaya bölünürse, 4 parçası modeli eğitmek için kullanılır ve kalan 1 parça modeli doğrulamak için kullanılır. Bu adımı model tüm parçalardan geçene kadar tekrarlanır. Her seferinde farklı bir parça doğrulama seti olarak kullanılır ve geri kalan parçalar eğitim seti sağlar.

Performans Ölçümü: Her bir adımda modelin doğrulama seti üzerindeki performansı kaydedilir. Bu çalışmada model performansını değerlendirmek için RMSE değeri hesaplanmıştır.

Sonuçları Birleştirme: Tüm adımlardan elde edilen performans ölçütlerinin ortalamasını alınarak, modelin genel performansı hakkında bir fikir edinilir. Bu, modelin yeni ve görülmemiş veriler üzerinde nasıl performans göstereceğini tahmin etmeye yardımcı olur.

Bu çalışmada çapraz doğrulama yöntemi kullanılarak en düşük RMSE oranına sahip veri seti eğitim veri seti olarak seçilmiştir.

2.4. Adaptif Yörünge Planlama Algoritmaları

Bölüm 2.3'de ödül fonksiyonlarının ROI alanlarındaki başarımını değerlendirmek için maksimum ödüle sahip yörünge seçilerek güzergah planlanmıştır. Fakt bu miyop olan bir yaklaşım olup bir sonraki hamleye karar vererek yeni örneklerin toplanmasını sağlamaktadır. Bunun yanında bütün örneklerle başlangıçta karar verip bu örnekleri en kısa sürede toplama problemi gezgin satıcı problemi olarak bilinmektedir. Bu tez çalışmasında maksimum fayda sağlayan örnekleri en kısa

süreden toplayan hibrit bir yaklaşım önerilmiştir. Bu bölüm hem önerilen hibrit model hemde diğer yöntemler hakkında bilgi vermektedir. Literatürdeki diğer yöntemler algoritmanın başarımını değerlendirmek için mevcut veri setine uygulanmıştır.

Bölüm 2.4.1'de, sezgisel bir yöntem olarak Monte Carlo arama ağacı algoritması kullanılarak adaptif bir güzergah planlama algoritması geliştirilmiştir. Monte Carlo arama ağacı, optimal rotayı garantilememekle birlikte optimale yakın çözümler sunabilmesi nedeniyle tercih edilmiştir. Bu yöntem, özellikle büyük ve karmaşık problem alanlarında hesaplama yükünü azaltırken, makul süreler içinde kabul edilebilir çözümler üretme potansiyeline sahiptir.

Bölüm 2.4.2'de ise deterministik bir yöntem olan yorucu arama algoritması (ESAS) ile adaptif örnekleme gerçekleştirilmiştir. Yorucu arama algoritması, olası tüm çözüm yollarını değerlendirme yeteneği sayesinde optimal çözüme ulaşmayı garanti eder. Ancak, bu yöntem, hesaplama yoğunluğunun oldukça artmasına neden olabilecek bir yapıya sahiptir. Özellikle büyük veri setlerinde, hesaplama süresi ve kaynak tüketimi açısından ciddi zorluklar yaşanabilmektedir.

Bu çalışmada, optimal çözümü elde etmek amacıyla, BnB algoritması kullanılmıştır. BnB algoritması, çözüm uzayını etkili bir şekilde daraltarak optimal çözüm arayışını hızlandırmaktadır. Çalışmada, alt limit Monte Carlo arama ağacı kullanılarak hesaplanmış ve bu sayede algoritmanın etkinliği artırılmıştır. Geliştirilen algoritmanın detayları ve performans analizleri ise Bölüm 2.4.3'te detaylandırılmıştır.

2.4.1. Monte Carlo Arama Ağacı ile Adaptif Örnekleme

MCTS algoritması en iyi arama algoritmasını kullanarak başlangıç düğümünden bir arama ağacı oluşturur. Monte Carlo ağaç araması (MCTS) terimi, rastgele yaklaşımı ilk kez Go-oyun programı Crazy Stone'da kullanan Rémi Coulom [101] tarafından icat edildi. MCTS algoritması go oyunuyla 2015, 2016 ve 2017 dünya şampiyonlarını yenmeyi başardı [102]. Fu ise MCTS'nin genel mekaniğini tic-tac-toe oyunuyla göstermektedir [102]. Yapılan çalışmalarda MCTS algoritması oyunlarda bir sonraki hamleyi belirlemek için kullanılmaktadır. Bu çalışmada güzergah planlama algoritmasında kullanılarak entropiyi minimize edecek bir algoritma geliştirilmiştir.

MCTS algoritması genel olarak seçim, genişletme, simülasyon ve geri yayılım olmak üzere 4 bölümden oluşmaktadır.

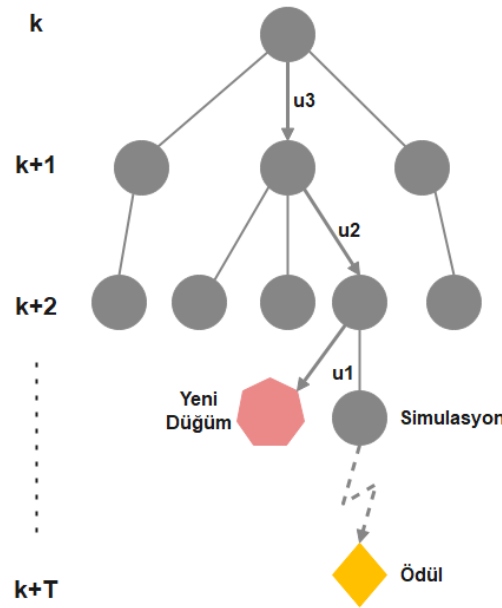
Seçim (Selection): Ağaçtaki düğümleri dolaşarak, en umut verici düğüme (genellikle bir tür seçim stratejisi ile) ulaşılır. Bu UCT adı verilen bir dengeleme stratejisi kullanılarak yapılır. Bu strateji, mevcut bilgiyle keşfedilmemiş düğümler arasında denge kurar.

Genişletme (Expansion): Seçim aşamasında ulaştığımız düğümde, henüz keşfedilmemiş olan çocuk düğümlerden birini seçerek ağacı genişletiriz. Bu, oyun veya problem alanına göre yeni olası hareketleri veya durumları simüle eder.

Simülasyon (Simulation): Genişletme adımında eklenen düğümü kullanarak rastgele oyun veya problem simülasyonları yapılır. Bu simülasyonlar, mevcut durumdan başlayarak rastgele hareketler yaparak bir sonuç elde etmek için kullanılır. Genellikle bu simülasyonlar, belirli bir hedefe ulaşma olasılığını veya başarı ölçütünü belirlemek için kullanılır.

Geriye Yayılma (Backpropagation): Elde edilen sonuçlar, ağaçtaki yolu geri izleyerek, her düğümün başarı oranını güncellemek için kullanılır. Bu geri yayılma, her düğümün ziyaret sayısını ve başarı oranını güncelleyerek gelecekteki seçimleri etkiler.

Şekil 2.8’de MCTS algoritmasının akışı yer almaktadır. k . Adımda başlangıç düğümünden başlanarak ağaca UCT denklemine göre ekleme yapılır. UCT denklemi Denklem 35’de verilmiştir. Daha sonra Algoritma 2’nin 6. satırında yer alan RolloutPolicy ile 2000 ms boyunca rastgele simülasyonlar gerçekleştirilir. Daha sonra geri yayılım ile her bir düğümün ödül değeri ve ziyaret sayısı güncellenir.



Şekil 2.8. MCTS Algoritması

Algoritma 2 MCTS algoritmasının sözde kodunu vermektedir. 2. Satırda belirtilen bütçe kadar örnek toplanıp toplanmadığını kontrol etmektedir. Eğer istenilen kadar örnek toplanmadıysa mevcut düğümde ağaca ekleme yapılır. Ekleme yapılırken robot için “sağ”, “sol” ve “ileri” etiketleri kullanılmıştır. Genişleme yapılırken düğümler UCT denklemi kullanılarak seçilmektedir. UCT denklemi Denklem 35’de verilmiştir. 6. Satırda en iyi arama algoritması kullanılarak oluşturulan ağaçtan rastgele güzergah oluşturulmaktadır. Her bir güzergahtaki ödül o güzergahın seçilmesiyle muhtemel entropi düşüşü olarak hesaplanmaktadır. 10. Satırda UCT algoritmasına

göre ağaca eklenen düğüm sonrasında her bir düğümün ödül değeri, ve ziyaret edilme sayısı geri yayılım yapılarak güncellenir. Son olarak en iyi rota maksimum ödüle sahip rota olarak seçilir.

$$UCT(s) = \frac{R(i)}{s_v} + C \times \sqrt{\frac{\ln N_v}{s_v}} \quad (35)$$

Denklemden çocuk için hesaplanan $R(i)$ değeri, entropi düşüşünü vermektedir. Denklem 36 ile hesaplanmaktadır [103]. s_v değeri çocuk düğümün ziyaret edilme sayısını vermektedir. N_v değeri mevcut düğümün ziyaret edilme sayısını vermektedir.

$$R = \sum_{i=g}^m \sum_{j=1}^m P(E = w_j) \rho(i|w_j) \quad (36)$$

MCTS fonksiyonunda ödül değeri belirlenirken oradan alınan ölçümün ne olacağı bilinmemektedir. Fakat GP ile tahmin edilebilmektedir. GP ile aday noktadan alınan ölçümün aralığının $\mu - 2 \times \sigma$ ile $\mu + 2 \times \sigma$ arasında olduğu bilinmektedir. Denklem 36'da belirtilen m değeri, bu aralıktan alınan örneklerin sayısını temsil eder ve bu çalışmada 20 adet eşit aralıkta örnekler alınmıştır. Bu örnekler, alınan ölçüm değerinin entropi değişimini nasıl etkileyeceğini hesaplamak için kullanılan Gauss regresyon algoritmasında eğitime dahil edilmiştir. Denklem i , güzergahın uzunluğunu verir ve güzergah uzunluğu arttıkça hesaplama süresi de eksponansiyel olarak artmaktadır. Bu durum, güzergah planlama probleminin karmaşıklığını artıran önemli bir faktördür.

MCTS algoritması, başlangıç konumunda yapılan ilk ölçüm ve bu ölçümün GP modeline eklenmesiyle başlar. Algoritma, belirlenen zaman sınırı boyunca (örneğin 2000 birim zaman) çalışarak, her döngüde olası bir eylem (ileri, sağa ve sola gitme gibi) seçer ve bu eylemin sonuçlarını simüle eder. Seçilen eylemler, UCT fonksiyonu kullanılarak değerlendirilir. UCT fonksiyonu, deneme-yanılma yöntemiyle yeni alanları keşfederken, önceki tecrübeleri de göz önünde bulundurarak optimal rotayı belirlemeye yardımcı olur. Belirlenen zaman sınırı içinde en uygun eylemi belirlemek amacıyla arama ağacı genişletilir ve en yüksek ödül değerine sahip olan eylem seçilir.

Algoritma, belirlenen bütçe dahilinde (H) simülasyonları gerçekleştirir. Her bir simülasyonda, belirli bir yolu izleyerek (RolloutPolicy) bu yolun ödülünü (R yolu) hesaplar ve bu ödülü geri yayılım yöntemiyle arama ağacına ekler. Eğer hesaplanan ödül, mevcut maksimum ödülde düşürülürse, bu yeni yol optimal yol olarak güncellenir. Bu süreç, algoritmanın en iyi yolu bulmak için sürekli olarak kendini iyileştirdiği bir döngü oluşturur.

MCTS ile adaptif örnekleme entegrasyonu, algoritmanın performansını artıran kritik bir bileşendir. Adaptif örnekleme, GP modelinin belirli alanlarda daha fazla veri toplamasını sağlayarak ortam hakkında daha fazla bilgi edinilmesine olanak tanır. Bu sayede, algoritma,

belirsizlikleri azaltarak ve daha doğru kararlar vererek karmaşık ortamlarda daha yüksek performans gösterebilir. Ölçüm değerlerinin entropi değişimini etkileyen örnekleri GP modeline dahil ederek, algoritmanın karar verme sürecinde daha bilinçli ve güvenilir adımlar atması sağlanır.

Algoritma 2 MCTS

```
1: fonksiyon MCTS(GP, x0, n, H)
2:   Başlangıç konumundan başla
3:   y0 = Ölçüm(x0)
4:   GP'ye başlangıç noktasındaki ölçümü ekle
5:   GP0 = GP (GP, x0, y0)
6:   mak_odul= sonsuz
7:   döngü Zaman < 2000 yap
8:     Kuyruk = Aksiyon (Sağ, sol, ileri)
9:     UCT(Kuyruk) // Denklem 26
10:    Arama_Ağacına_Ekle ← mak(kuyruk.UCT)
11:  döngü bitti
12:  döngü n < H yap
13:    Güzergah ← RolloutPolicy(Arama Ağacı)
14:    Güzergah ödülü = R(Güzergah) // Denklem 27
15:    Geri Yayılım(Ağaç)
16:    eğer Güzergah ödülü < mak_odul sonra
17:      Optimal Güzergah = Güzergah
18:    n=n+1
19:  döngü bitti
20: fonksiyon bitti
```

2.4.2. Yorucu Arama Algoritması ile Adaptif Örnekleme

MCTS algoritması, problemleri çözmek için sezgisel ve rastgele bir yaklaşıma dayanır. Bu algoritma, mevcut tüm olasılıkları sistematik bir şekilde değerlendirerek, maksimum ödül değeri elde edecek en uygun güzergahı belirlemeye çalışır. Sezgisel yapısı, her aşamada bir dizi rastgele hareketler gerçekleştirilmesini ve bu hareketlerin sonucunda elde edilen ödüllerin değerlendirilmesini içerir.

Buna karşılık, deterministik bir algoritma olan Yorucu Arama Algoritması (Exhaustive Search) kesin sonuçlar sağlar. Bu algoritma olası tüm rotaları sistemli bir şekilde inceleyerek, her bir rotayı değerlendirir ve maksimum ödüle sahip en uygun rotayı seçer. Yorucu aramanın doğası gereği, tüm olasılıkları hesaba kattığı için her zaman kesin çözüm sunar, ancak bu kapsamlı değerlendirme süreçleri hesaplama açısından oldukça zahmetli olabilir.

Yorucu Arama Algoritması ve Adaptif Örnekleme'yi birleştiren Algoritma 3, belirsiz ortamlardaki güzergah planlama problemlerine çözüm getiren etkili bir strateji sunar. Exhaustive Search Adaptive Sampling (ESAS) olarak adlandırılan bu algoritma, yorucu aramanın kapsamlı doğası ile adaptif örnekleme tekniklerini entegre ederek maksimum ödüle sahip güzergahları bulmayı hedefler. Algoritmanın başlangıcında, GP, başlangıç konumu (x0), simülasyon sayısı (n) ve bütçe (H) gibi temel parametreler girilir.

İlk adımda, başlangıç konumunda bir ölçüm gerçekleştirilir ve bu ölçüm verisi GP modeline dahil edilir. Bütçe (H) ve maksimum ödül (max_reward) değişkenleri sıfırlanarak algoritmanın işlemeye başlaması sağlanır. Yorucu arama aşamasında, algoritma iki temel liste kullanır: openlist ve closedlist.

Openlist henüz keşfedilmemiş komşu konumları içerirken, closedlist ziyaret edilmiş konumları barındırır. Her döngüde, bütçenin tükenmesine kadar olası bir eylem seçilir. Olası eylem olarak openlist içerisinde ilk eleman seçilir (Alg. 3 Satır 12). Ziyaret edilmemiş bir konum seçildiğinde, openlist güncellenir ve konumun ödülü hesaplanır. Eğer seçilen konumun ödülü, mevcut mak_odul değerini aşıyorsa, bu konum yeni optimal konum olarak işaretlenir. Böylece, en yüksek ödülü sunmaya potansiyel en yakın konum belirlenmiş olur.

Adaptif örnekleme aşamasına geçildiğinde, algoritma yeni başlangıç konumu ile eski başlangıç konumu arasında bir bağlantı kurar. Bu adım, algoritmanın daha önce keşfedilen bölgeler arasındaki belirsizlikleri azaltıp daha verimli arama yapmasına yardımcı olur. Yeni başlangıç konumu, eski başlangıç konumu olan startNode'un yerini alır ve H sayacı bir artırılır. Bu işlem, GP modelinin yeni verilerle sürekli olarak güncellenmesini sağlar, böylece model, ortam hakkında daha doğru tahminlerde bulunabilir.

ESAS algoritmasının en büyük avantajı, yorucu arama ile ortamı detaylı bir şekilde keşfetmesi ve adaptif örnekleme ile belirsizlikleri minimize edebilmesidir. Bu kombinasyon, potansiyel olarak daha yüksek ödül sunma ihtimali olan bölgelerde ek ölçümler olarak Gauss işlemini iyileştirir. Böylece, karmaşık ve belirsiz ortamlarda optimal güzergahı bulma şansı artar.

Bu algoritmanın gücü, yorucu aramanın kapsamlı doğası ile adaptif örneklemenin verimliliğinden kaynaklanır. Ancak, bu kombinasyonun büyük ve karmaşık problem alanlarında hesaplama yükünün artması gibi dezavantajları da bulunmaktadır. Büyük veri setlerinde veya çok karmaşık ortamlar için algoritmanın performansı, artan hesaplama süreleri ve kaynak kullanımı nedeniyle zor olabilir. Ancak, ESAS algoritması, belirsizlikleri azaltarak ve potansiyel olarak daha iyi rotaları keşfederek belirlenen hedefe en uygun ve optimal güzergahı bulmayı amaçlar. Bu nedenle, yüksek doğruluk ve verimlilik gerektiren senaryolarda önemli avantajlar sunar.

Algoritma 3 Yorucu Arama Algoritması ile Adaptif Örnekleme (Exhaustive Search Adaptive Sampling-ESAS)

```
1: ESAS(GP, x0, n, H)
2: Başlangıç konumundan başlayarak ölçümleri toplama
3: y0 = Olcum(x0)
4: Başlangıç konumunu kullanarak GP'yi eğit
5: GP0 = GP (GP, x0, y0)
6: H = 0
7: mak_odul ← -∞
8: eğer H < T sonra
9:   openlist = {x0}, closedlist = ∅
10:  döngü n < bütçe yap
11:    Aksiyon = Aksiyon al {sağ, sol, ileri}
12:    kuyruk = openlist(1)
13:    eğer kuyruk ziyaret edilmediyse sonra
14:      openlist = kuyruk ve openlist(1).odul = R (openlist(1)) [Denklemler 19] ve kuyruk.veli =
openlist(1)
15:      eğer kuyruk.odul > mak_odul sonra
16:        optimal_dugum = mevcut_dugum, mak_oduk = optimal_dugum.odul
17:      eğer baslangic_dugumu = x0 sonra
18:        p0 = mevcut_dugum, baslangic_dugumu = mevcut_dugumu_veli
19:        H = H + 1
20:        y0 = Olcum(p0), GP0 = GP (GP, x0, y0), n = n + 1
21:      closedlist = openlist(1)
```

2.4.3. Önerilen Hibrit bir Model Olan MCTS-BnB Algoritması ile Adaptif Örnekleme

Algoritma 4, BnB-MCTS Adaptif Örnekleme olarak adlandırılan bir yaklaşımı tanımlamaktadır. Belirtilen bu strateji, optimizasyon problemlerinde en iyi çözümü aramak amacıyla MCTS algoritması ve BnB algoritmasını bir araya getirir. MCTS algoritması, sezgisel ve rastgele bir yaklaşım kullanarak optimal çözüme yakın sonuçlar elde etmeye çalışır. Ancak bu yöntem her zaman kesin optimal çözümü garanti etmez. Buna karşın, Yorucu Arama Algoritması, problem uzayındaki tüm olası yolları değerlendirerek kesin olarak en iyi sonucu bulur. Fakat bu süreç, büyük ve karmaşık problem alanlarında yüksek hesaplama maliyetlerine yol açar. BnB-MCTS adaptif örnekleme algoritması, belirsiz ortamlarda en iyi güzergahı bulmak amacıyla BnB, MCTS ve adaptif örnekleme yöntemlerini birleştirerek avantajlar sunar.

Algoritmanın başlangıç adımında, belirli bir başlangıç konumunda bir ölçüm gerçekleştirilerek bu ölçümü GP modelinde saklar. Bu işlem, başlangıç durumunu GP₀ olarak tanımlar. Algoritmanın çalıştığı süre boyunca, H sayacı sıfırlanmış olarak tutulur ve maksimum ödül negatif sonsuz olarak ayarlanır (Alg.4 Satır 7). Yani başlangıçta herhangi bir ödül elde edilmemiştir ve çözüm arama işlemi sadece H değeri belirli bir eşik değeri olan T'den küçük olduğu sürece devam eder.

Bir eşik değeri olan T'ye ulaşana kadar, belirli bir bütçe sınırı çerçevesinde işlem yapılır. Algoritma, farklı hareketler seçer (sağa gitme, sola gitme, ileri gitme) ve oluşturulan yeni sorguların henüz ziyaret edilip edilmediğini kontrol eder.

Eğer sorgu daha önce ziyaret edilmemişse, bu sorguya ait üst sınır (kuyruk_üst sınır) ve alt sınır (kuyruk_altsınır), hesaplanır. Alt sınır MCTS algoritması kullanılarak hesaplanır. Üst sınır Algoritma 5 kullanılarak hesaplanır. Bu süre zarfında sorguya ait üst ve alt sınırlar hesaplanırken yukarıya doğru geri yayılım işlemi gerçekleştirilir. Bu süreç, her düğümün ana düğümlerine kadar ilerleyerek üst ve alt sınırların güncellenmesini içerir. Böylece, her sorgunun en güncel ve kesin değerlendirilmesi sağlanır.

Algoritma, en iyi alt sınıra sahip sorguyu belirler ve bu sorgunun üst sınırının en iyi alt sınırdan daha düşük olduğu durumlarda, ilgili sorguyu kırpır (Alg.4 Satır 24). İlgili sorgunun eliminasyonu, gereksiz hesaplamaların önüne geçilmesi ve algoritmanın daha etkin hale gelmesini sağlar. Bu şekilde belirlenen optimal düğüm ve ödülü, H sayacına bağlı olarak başlatılan ve güncellenen GP modeli aracılığıyla ölçülür.

Algoritmanın adaptif örnekleme stratejisi, belirlenen eşik değeri olan T'yi aşmadığı sürece devam eder ve her adımda en uygun hareketi seçer. Bu yöntem, veri toplama işlemini optimize ederken toplanan verilerin sayısını ve niteliğini, hedefe daha etkin bir şekilde ulaşmak için belirlenen bütçe sınırları içinde tutmayı amaçlar. Her adımda yapılan ölçümler, GP modelinin sürekli olarak güncellenmesini sağlar ve bu güncellemeler, algoritmanın daha isabetli tahminler yapmasına ve daha uygun adımlar belirlemesine katkıda bulunur.

MCTS algoritması sezgisel bir yaklaşım olmasına rağmen optimale yakın bir sonuç vermektedir. Fakat optimal çözümü garanti etmemektedir. Yorucu Arama algoritması ise dallandıktan sonra bütün olası yolları değerlendiği için optimal çözümü kesin bulmasına rağmen hesaplama maliyeti oldukça yüksektir. BnB-MCTS adaptif örnekleme algoritması, belirsiz ortamlarda en iyi güzergahı bulmak için Branch-and-Bound (BnB), Monte Carlo Tree Search (MCTS) ve adaptif örnekleme birleştirir.

Algoritma 4 BnB-MCTS Adaptif Örnekleme

```
1: BnB-MCTS (GP, x0, n, H)
2: Başlangıç konumundan başlayarak ölçümleri topla
3: y0 = Olcum(x0)
4: Başlangıç konumunu kullanarak GP'yi eğit
5: GP0 = GP(GP, x0, y0)
6: H = 0
7: mak_odul ← -∞
8: eğer H < T sonra
9:   openlist = {x0}, closedlist = ∅
10:  döngü n < bütçe yap
11:    Aksiyon = Aksiyon al{sağ, sol, ileri}
12:    kuyruk = openlist(1)
13:    eğer kuyruk ziyaret edilmediyse sonra
14:      kuyruk_üstsınır ← UP(kuyruk, n)
15:      kuyruk_altsınır ← MCTS(GP, kuyruk, n, H)
16:      mevcut_dugum = kuyruk
17:      döngü mevcut düğüm olmayana kadar yap //Üst limit ve alt limiti belirle (geri yayılım)
18:        mevcut_dugum_üstsınır ← UP(kuyruk, n)
19:        mevcut_dugum_altsınır ← MCTS(GP, kuyruk, n, H)
20:        mevcut_dugum ← mevcut_dugum.veli
21:        eğer mevcut_dugum_altsınır < optimal_altsınır sonra
22:          optimal_dugum = kuyruk
23:        eğer kuyruk_üstsınır < kuyruk_altsınır sonra
24:          buda (kuyruk)
25:        değilse
26:          openlist = kuyruk ve openlist(1).odul = R(openlist(1)) [Denklem 19] ve kuyruk_veli =
openlist(1)
27:        eğer kuyruk_odul > odul sonra
28:          optimal_dugum = mevcut_dugum, odul = optimal_dugum.odul
29:        eğer baslangic_dugumu = x0 sonra
30:          p0 = mevcut_dugum, baslangic_dugumu = mevcut_dugum.veli
31:          y0 = Olcum(p0), GP0 = GP (GP, x0, y0), n = n + 1
32:          closedlist = openlist(1)
```

Algoritma 5 üst sınır hesaplarırken, belirli bir bütçe içinde en iyiyi seçmek için belirlenen yol noktaları arasından en iyi seçeneği seçer. Bu sayede, robotun sınırlı bütçeyle toplayabileceği maksimum bilgiyi hesaplar. Algoritma 4'teki BnB algoritmasında alt sınır değerini hesaplarırken MCTS algoritmasını kullanılır. Bu algoritma, rastgele seçilen yol noktaları arasından en iyiyi seçerek alt sınır değerini belirler. Üst sınır değerini belirlemek için ise Algoritma 5'i kullanılır. Bu algoritma, 2. Adımda başlangıç noktasından belirli bir bütçeyle gidebileceği yol noktaları belirlenir. Ardından, belirlenen yol noktaları arasından en iyi seçeneği seçerek maksimum ödül değerini hesaplanır. Bu şekilde, hem alt sınır değerini Monte Carlo Ağaç Araması ile hem de üst sınır değerini Algoritma 5 ile hesaplayarak, Bnb algoritmasını daha verimli bir şekilde çalıştırabilir. Bu sayede, robotun hareket edebileceği en iyi yolu belirleyerek maksimum ödül değerini elde edebilir.

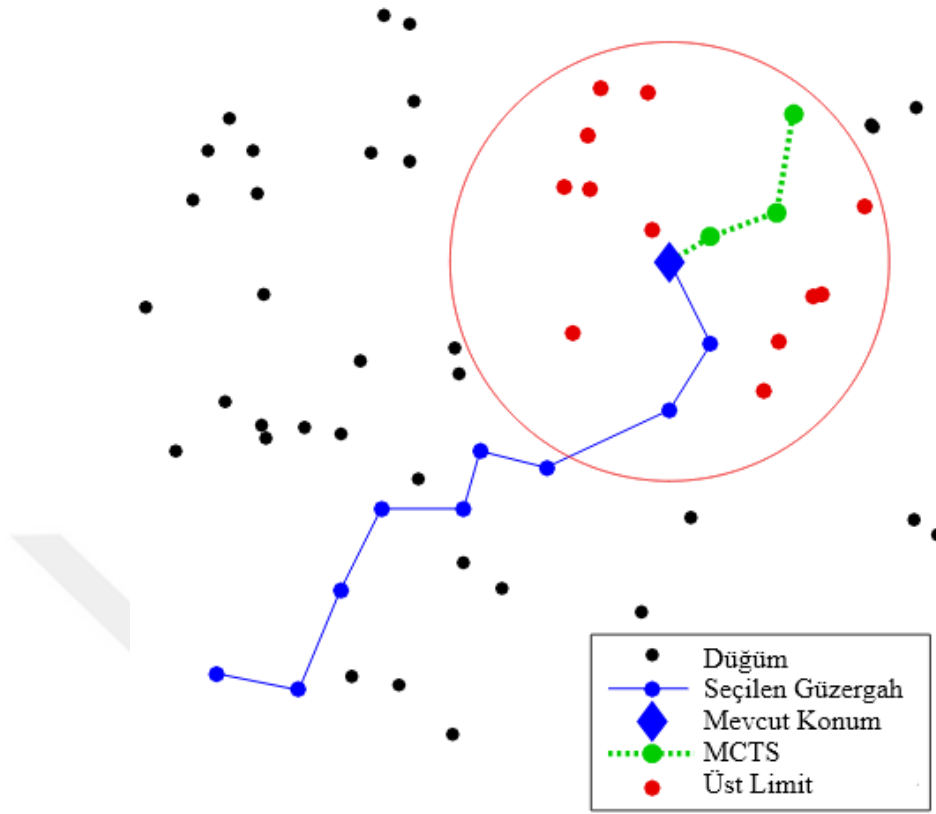
Algoritma 5 Üst Limit(Upper Bound-UP)

1: UP (x_0, n)
2: **eğer** OKLİD(x_0 , Aday Yol Noktası) $<H$ **sonra**
3: Yol Noktası kümesi \leftarrow Aday Yol Noktası
4: **döngü** $n < H$ **yap**
5: $R(\text{mak}(\text{Yol Noktası}))$ //Denklem 19
6: $n = n + 1$
7: **döngü bitti**

Şekil 2.9'da, BnB-MCTS algoritmasının alt ve üst limitleri hesaplama süreci ayrıntılı olarak gösterilmektedir. Şekil 2.9, algoritmanın çalışma prensiplerini anlamada kritik bir rol oynamaktadır. Mavi baklava dilimi, algoritmanın mevcut konumunu gösterirken, bu konumdan ulaşılacak diğer düğümler, çözüm alanının sınırlarını oluşturmaktadır. Üst limit hesaplanırken, mavi baklava diliminden gidilebilecek potansiyel konumlar belirlenir ve bu küme üzerinde entropi değerini maksimum şekilde düşüren 10 örnek nokta seçilir. Bu 10 nokta, üst limit değerlerini temsil eder ve kırmızı ile işaretlenmiştir.

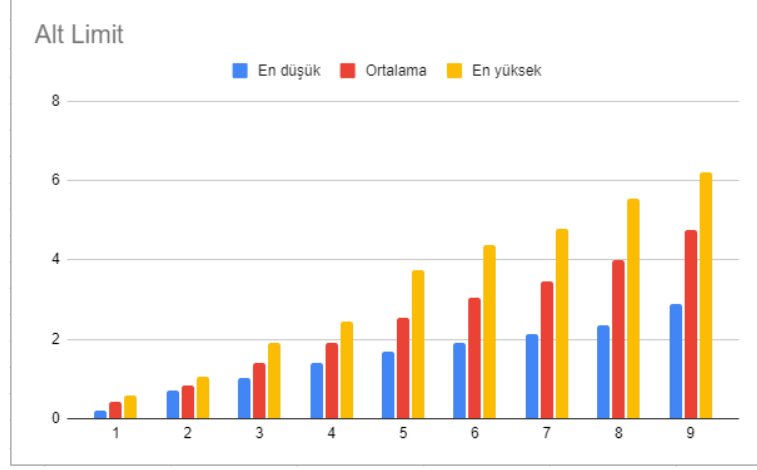
Yeşil ile belirlenen güzergah, MCTS algoritmasının 10 adımda ulaşabileceği en iyi güzergahı temsil etmektedir. MCTS, birçok olası güzergahı test eder ve sezgisel olarak maksimum ödüle sahip olan güzergahı seçer. Bu seçilen güzergah, algoritmanın verimliliğini ve doğruluğunu artırmak amacıyla sürekli olarak güncellenir. Siyah noktalar, harita üzerindeki potansiyel düğüm pozisyonlarını, mavi çizgiyle gösterilen güzergah ise, algoritmanın şimdiye kadar seçmiş olduğu rotayı göstermektedir.

Bu güzergah üzerindeki her bir düğümün alt ve üst limit değeri, mevcut konuma yeni düğümler eklendikçe sürekli olarak güncellenir. Güncelleme işlemi, algoritmanın adaptif ve dinamik bir şekilde en uygun rotayı belirlemesini sağlar. Ayrıca, algoritma sınırlama yaparken, aday düğümün üst limit değeri, gidilebilecek sağ, sol, yukarı ve aşağı etiketlerindeki diğer düğümlerin alt limit değerlerinden küçükse o düğüm sistematik olarak elenmektedir. Bu işlem, algoritmanın daha etkili ve verimli bir arama yapmasını sağlar ve gereksiz düğümlerin işleme alınmasını engeller.

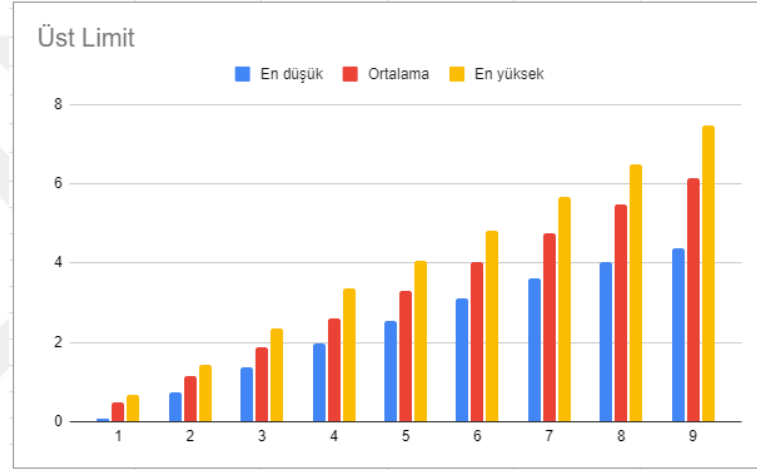


Şekil 2.9. Örnek alt limit ve üst limit hesaplama şeması

Şekil 2.10’da, BNB algoritmasının 9 adımlık süreç boyunca alt ve üst limit değerlerinin nasıl değiştiğini göstermektedir. Her iki grafik de mavi (en düşük), kırmızı (ortalama) ve sarı (en yüksek) olmak üzere üç farklı değeri içermektedir. Grafikler başlangıç düğümünün alt limit ve üst limitini göstermektedir. Her yeni düğümün alt limit ve üst limit hesaplandığında Algoritma 4’de verildiği gibi geri yayılım yaparak başlangıç düğümünün alt limit ve üst limit değeri güncellenmektedir. Böylece başlangıç düğümünden elde edilebilecek maksimum ve minimum ödül değerleri hesaplanmaktadır. Grafikler incelendiğinde üst limitin alt limitten sürekli daha yüksek çıktığı, fakat buna nazaran aradaki farkın çok yüksek olmadığı görülmektedir. Bu durum daha fazla sınırlama yaparak daha kısa sürede optimal rotanın bulunmasını sağlamıştır.



a)



b)

Şekil 2.10. MCTS-BnB Algoritmasında hesaplanan alt limit ve üst limit değerleri a) Alt limit b) Üst limit

2.5. Seyrek Sözde Girişlerin Belirlenmesi için Kullanılan Kümeleme Yöntemleri

Gauss regresyon algoritması, tahmin yaparken eğitim verilerinin birbirleriyle olan ilişkisini matematiksel olarak hesaplamaktadır. Ancak, eğitim veri kümesinin büyümesiyle birlikte hesaplama süresi de üstel şekilde artmaktadır. Literatürde yapılan çalışmalar incelendiğinde, GP algoritmasındaki hesaplama yükünü azaltmak amacıyla alt örneklem kümesinin kullanıldığı görülmektedir. Bununla birlikte, bu çalışmaların optimal sonuçlara ne kadar yakın olduğu yeterince değerlendirilmemiştir. Bu tez çalışmasında, GP'nin hesaplama süresini azaltmak için alt örneklem kümesi, çeşitli kümeleme algoritmaları kullanılarak belirlenmiştir. Mevcut veri seti, iki ile yirmi aralığında farklı sayılarda kümelemeye tabi tutulmuş ve kümeleme algoritmalarının siluet skor indeksi, Bölüm 2.6'da açıklandığı üzere hesaplanmıştır. En yüksek siluet skor indeksine sahip olan kümeleme yöntemi, alt örneklem kümesi olarak seçilmiştir. Sonrasında, elde edilen kümelerden 1, 2, 3 ve 4 örnek alınarak eğitim veri sayısı düşürülmüş ve Bölüm 2.4.3'de detaylı olarak anlatılan

hibrit güzergah planlama algoritmasının hesaplama süresi azaltılmaya çalışılmıştır. Böylece, oluşturulan güzergahın optimal çözüme ne kadar yakın olduğu araştırılmıştır. Bu bölümde, kullanılan beş farklı kümeleme yöntemi detaylı şekilde incelenmiştir.

2.5.1. K-Means Kümeleme Algoritması

K-Means, kümeleme problemine çözüm sunan yaygın olarak bilinen denetimsiz öğrenme algoritmalarından biridir. Bu algoritma, verilen bir veri kümesini önceden belirlenmiş sayıda küme (k kümeleri olarak adlandırılır) ayırmak için basit ve doğrudan bir yöntem benimser. Temel prensip, her bir küme için bir merkez atamak ve bu merkezlerin birbirinden olabildiğince uzak konumlandırılarak farklı sonuçlardan kaynaklanan hataları minimize etmektir [104]. Algoritma, her bir veri noktasını belirli bir kümeye atayarak başlar ve bu adım, tüm noktalar atanana kadar devam eder ve böylece ilk gruplandırma işlemi gerçekleşir. Sonrasında, elde edilen k kümelere ait merkezlerin yeniden hesaplanması gerekmektedir. Bu yeni k merkezlerin hesaplanmasını takiben, aynı veri noktaları en yakın yeni merkeze tekrar atanır. Bu süreç, bir döngü oluşturur ve k merkezlerinin artık hareket etmediği veya sadece ihmal edilebilir ölçüde hareket ettiği ana kadar devam eder. Sonuç olarak, bu algoritma, Denklem 37 ile ifade edilen bir amaç fonksiyonunu minimize etmeyi hedefler.

$$W(S, C) = \sum_{k=1}^K \sum_{i \in S_k} \|y_i - c_k\|^2 \quad (37)$$

S, varlık kümesinin bir alt kümesini temsil eder ve y_i ($i \in I$) vektörleri ile ifade edilir. Bu alt küme, m-boyutlu özellik uzayında yer alır ve birbirini örtmeyen, boş olmayan S_k kümelerinden oluşur. Her bir küme, c_k [105] ($k = 1, 2, \dots, K$) olarak adlandırılan bir merkeze sahiptir.

2.5.2. Agglomerative Hierarchical

Agglomeratif Hiyerarşik Kümeleme, nesnelere benzerliklerine göre kümeler halinde gruplandırmak için sıklıkla kullanılan etkili bir yöntemdir. Bu algoritma, AGNES (Agglomerative Nesting) olarak da bilinir ve temel olarak her bir nesneyi başlangıçta bağımsız bir küme olarak ele alır [106]. Sürekli bir süreç olarak, her adımda en yüksek benzerlik gösteren iki küme birleştirilerek daha büyük bir küme oluşturulur. Bu birleştirme işlemi, tüm nesnelere tek bir büyük küme oluşturana kadar devam eder. Algoritmanın operasyonel süreci, başlangıçta her bir nesneyi tek elemanlı bir küme olarak kabul etmeyi ve daha sonra benzerlik metriği (örneğin, Öklidyen mesafe, Manhattan mesafesi veya korelasyon katsayısı) temel alınarak en benzer olan iki kümeyi birleştirmeyi içerir. Bu birleştirme adımları, veri yapısının büyüklüğüne ve karmaşıklığına bağlı olarak tekrarlanır. Nihayetinde, bu algoritma, veri kümesine ait nesnelere ağaç tabanlı bir temsili olan ve dendrogram olarak adlandırılan hiyerarşik bir yapı oluşturur [107]. Dendrogram, veri

nesnelerinin hiyerarşik ilişkilerini ve benzerliklerini görselleştirmek için kullanılır. Bu yapıda, yatay eksen benzerlik/distance ölçüsünü, dikey eksen ise kümeleme adımlarını gösterir. Her bir dal, kümeleme sürecinin herhangi bir aşamasında birleştirilen kümeleri temsil eder ve bu dalların birleşim yerleri, kümelerin birleştirme noktalarını ve açıklıklarını görselleştirir.

2.5.3. Birch

Birch, büyük hacimli veri kümeleriyle başa çıkmak için özel olarak geliştirilmiş bir kümeleme algoritmasıdır. 1996 ve 1997'de T. Zhang, Ramakrishnan ve Livny tarafından geliştirilen bir algoritma olan Birch, bu alandaki öncü algoritmayı temsil eder ve hiyerarşik kümeleme analizinde karşılaşılan ölçekleme sorunlarına ve önceki adıma geri dönememe sorunlarına çözüm getirir [108,109]. Büyük veri kümelerinde karşılaşılan bir zorluk, $N \times N$ boyutunda bir mesafe matrisi kullanılmasından kaynaklanır. Bu durum mekansal karmaşıklık sorunlarına yol açar. Birch, bu soruna bir çözüm sunar ve kümelerin özelliklerini temsil eden bir ağaç oluşturarak kümeleme işlemini bu ağaç üzerinden gerçekleştirir. Bu yaklaşım, mekansal karmaşıklık sorunlarına bir çözüm olarak kabul edilir. Birch, büyük veri kümelerinden alınan bir örnekleme temelinde kümelenmeye dayanan birçok hiyerarşik kümeleme algoritmasının aksine, tüm veri kümesini işler. Veri kümesindeki tüm kümelerin özelliklerini temsil eden kümeleme özelliğini kullanarak, Birch, ilişkili bilgilerin özetini çıkarır ve kümeleme işlemini bir kümeleme özellik ağacı aracılığıyla gerçekleştirir.

2.5.4. MiniBatch K-Means

MiniBatch K-Means, benzer özelliklere sahip verileri gruplamak için kullanılan bir kümeleme algoritmasıdır. K-Means algoritması genel olarak verimli performansı nedeniyle tercih edilen bir yöntem olsa da, analiz edilen veri kümesinin boyutu arttıkça, ana bellekte tüm veri kümesine ihtiyaç duyması ve buna bağlı olarak işlem süresinin uzaması gibi zorluklarla karşılaşabilir. Bu zorluğa bir çözüm olarak önerilen MiniBatch K-Means algoritması, zaman ve bellek maliyetlerini azaltmayı amaçlar. Bu algoritmanın temel fikri, sabit boyuttaki küçük rastgele seçilmiş veri gruplarını kullanarak bellek depolama alanını kolaylaştırmaktır [110]. Her bir iterasyonda, veri kümesinden yeni bir rastgele örnek çekilir, kümeler güncellenir ve bu işlem, yakınsama elde edilene kadar tekrarlanır. Bu yaklaşım, büyük veri kümelerini bellekte daha yönetilebilir hale getirmeyi ve K-Means algoritmasının zaman ve bellek karmaşıklıklarını ele almayı amaçlar.

2.5.5. Spektral Kümeleme

Spektral kümeleme, veri noktalarını grafikteki düğümler olarak işleyerek kümeleme işlemini bir grafik bölme problemi olarak ele alır. Kümelerin şekli veya formu hakkında herhangi bir önyargıya sahip olmayan bu yaklaşım, veri noktalarını düşük-boyutlu bir uzaya yansıtarak kümeleme işlemini gerçekleştirir [111]. İş akışı genellikle bir benzerlik grafiği, örneğin ϵ -komşuluk grafiği veya tam bağlantılı bir grafik, hesaplanarak başlar. Veri daha sonra Graf Laplacian matrisi kullanılarak düşük boyutlu bir uzaya yansıtılır, bu da grafiğin önemli özelliklerini belirlemeye yardımcı olur [112]. Spektral kümeleme süreci, veri noktalarını daha düşük boyutlu bir uzaya yerleştirme ve ardından özdeğerler ve özvektörler üzerinden kümeleme yapma adımlarını içerir. Genellikle bu aşamada K-Means kümeleme kullanılır. Spektral kümeleme, basit uygulaması ve etkili sonuçlarıyla bilinir. Özellikle seyrek veri kümeleri için hızlı ve etkili olmasına rağmen, büyük veri kümeleri için hesaplama açısından maliyetli olabilir. Bu nedenle, Spektral kümeleme algoritması genellikle veri setinin boyutuna ve özelliklerine göre dikkatlice değerlendirilmelidir.

2.6. Siluet Skor İndeksi (SSI)

SSI (Silhouette Score), kümeleme algoritmalarında sonuçlar arasındaki ayırım mesafesini keşfetmek ve anlamak için kullanılan bir metriktir. Diğer performans değerlendirme yöntemlerinden farklı olarak, SSI bir eğitim setine ihtiyaç duymadan kümeleme sonuçlarını değerlendirebilir, bu da onu kümeleme görevleri için daha uygun hale getirir. SSI, bir kümenin her bir nesnenin diğer kümelerdeki nesnelere ne kadar yakın olduğunu değerlendirmek amacıyla kullanılır [113]. SSI skoru -1 ile +1 arasında değişir ve kümeleme kalitesini gösterir. +1 puanı doğru kümeleme anlamına gelirken, -1 puanı nesnelere düzgün bir şekilde kümeleştirilmediğini gösterir. SSI değerini hesaplarırken, her bir nesne için öncelikle aynı kümedeki diğer nesnelere olan ortalama mesafe belirlenir ve bu mesafe $a(i)$ olarak gösterilir. Daha sonra, farklı kümelerdeki tüm nesnelere olan ortalama mesafe, $b(i)$ olarak temsil edilir. Son olarak, SSI değeri, Denklem 38 kullanılarak türetilir [114]. Bu metrik, kümeleme algoritmalarının performansını nesnelere kümeler arasındaki tutarlılığına göre değerlendirmeye yardımcı olur.

$$S(i) = (a(i) - b(i)) / \max(a(i), b(i)) \quad (38)$$

3. BULGULAR VE TARTIŞMALAR

Bu tez çalışmasının sonuçları üç bölümden oluşmaktadır. Birinci bölümde, eğitim verilerini toplamak için bir güzergah planlama algoritması geliştirilmiştir. Geliştirilen güzergah planlama algoritması, NASA tarafından açık olarak paylaşılan 10 farklı bölgeden alınan sıcaklık veri setinde test edilmiştir ve algoritmanın başarımı literatürdeki diğer çalışmalarla karşılaştırılmıştır. İkinci bölümde, farklı ödül fonksiyonlarının ROI alanlarındaki performansını araştırmak için 4 farklı test ortamı oluşturulmuştur. Bu test ortamında, Gauss dağılımı kullanılarak 1,2,3 ve 4 ROI alanı olan bölgeler bulunmaktadır. Üçüncü bölümde, çevresel izleme problemlerinde karşılaşılan gezgin satıcı problemine çözüm bulmak için hibrit bir güzergah planlama algoritması önerilmiştir. Önerilen güzergah planlama algoritması, 3 farklı veri setinde 2 farklı güzergah planlama algoritmasıyla karşılaştırılmış ve BnB algoritmasında kullanılan alt limit ve üst limit belirleme yöntemleri Bölüm 3.3.1’de anlatılan deneysel veri setine uygulanarak önerilen MCTS-BnB algoritmasının başarımı incelenmiştir. Son olarak, algoritmanın hesaplama süresini azaltmak için 5 farklı kümeleme yöntemi kullanılarak alt örneklem kümesi oluşturulmuş ve bu alt örneklem kümesinin optimale yakınlığı araştırılmıştır.

3.1. Eğitim Verilerini Belirlemek Geliştirilen Güzergah Planlama Algoritması

Çevresel izleme problemlerinde, çevresel eğitim ve başlangıç haritasının tahmin edilmesi için eğitim verileri kritik bir role sahiptir. Gauss regresyon, eğitim verileri üzerindeki ilişkileri kovaryans fonksiyonunu kullanarak analiz eder. Yakın örnek noktalarının seçilmesi, bu alanların daha iyi tahmin edilmesine yol açarken, eğitim verilerinin hiç alınmamış olması ise yanlış öğrenmeye yol açabilir. Bu nedenle, bu çalışmada, bir robotun sınırlı bütçesiyle alanın tamamından örneklerin toplanabilmesi için bir yörünge planlama algoritması önerilmiştir. Önerilen güzergah planlama algoritması Bölüm 2-2’de detaylı olarak anlatılmıştır.

Önerilen algoritmanın etkinliği NASA’nın açık olarak paylaştığı 10 farklı bölgedeki sıcaklık veri setinde test edilmiştir [98]. Alanın tamamının gidilecek yol noktalarının belirlenmesi önemli bir husustur. Bu bağlamda, alanın tamamını yol noktası olarak belirlemek yerine farklı dizilimlere sahip 100 aday yol noktası seçilmiş ve hesaplamalar bu noktalar üzerinden yapılmıştır. Bu 100 nokta, Halton, Sobol ve düzenli dizilim gibi farklı dağılım teknikleri kullanılarak belirlenmiş, rastgele seçim ise karşılaştırmak amacıyla yapılmıştır.

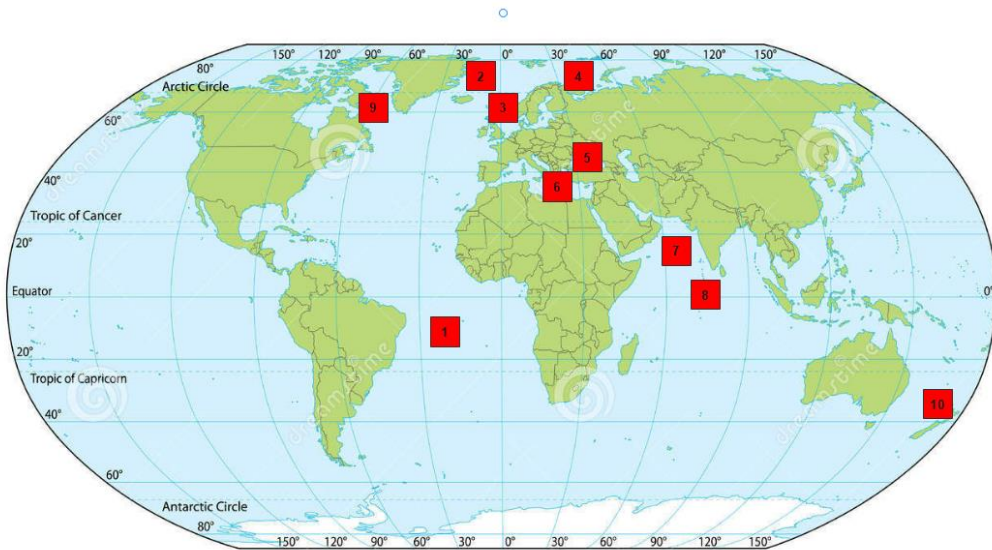
Ayrıca önerilen yöntemin etkinliğini değerlendirmek için literatürde yer alan Kemna vd. [52] tarafından önerilen pilot yöntem üzerinde de testler yapılmıştır. Kemna vd. çalışmasında planlama planlamal arı uygulamamışken, bu çalışmada pilot yöntem rota planlama algoritması entegre edilerek robotun bütçesinin daha etkin kullanılması sağlanmıştır.

Önerilen yöntem, önce mevcut bütçeyle toplamak üzere belirlenen örneklerden başlayarak, sonra artan bütçeyle gidilebilecek diğer örnekleri de toplamaktadır. Bu sayede mevcut bütçeyle daha fazla örnek toplanması ve alanın geniş bir yelpazede örnekleme yapılması sağlanmaktadır.

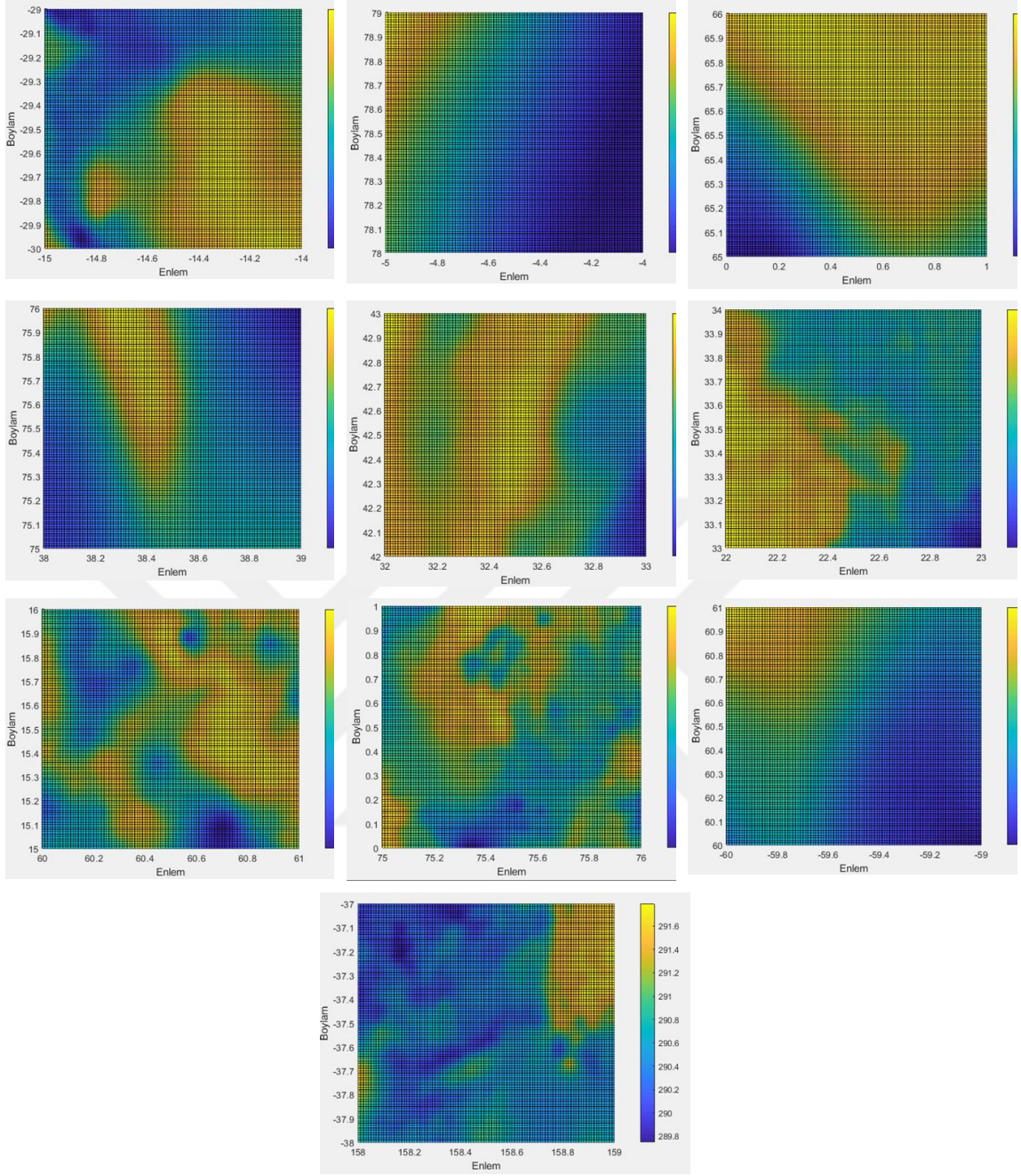
Çalışmada ayrıca önerilen yöntemin başarımı, çim biçme rotası ve rastgele yöntemlerle karşılaştırılarak değerlendirilmiştir. Elde edilen yolların başarımı RMSE değeri ile test edilmiş ve aynı sayıdaki örnekleme durumunda alanın tahmin başarımı ölçülmüştür.

3.1.1. Hiper parametreleri Öğrenmek için Kullanılan Veri Seti

Bu çalışmada, NASA tarafından sağlanan açık kaynak sıcaklık verileri, Gauss regresyonunda önerilen yöntemle eğitim verilerinden tahmin performansını ölçmek için kullanılmıştır. Önerilen yöntemin doğruluğunu test etmek için 10 farklı bölgede yöntemin başarımı test edilmiştir. NASA'dan elde edilen veriler, 23 Kasım 2021'de saat 09:00'da kaydedilen sıcaklık verilerini içermektedir. Gauss regresyon yaklaşımı, kovaryans fonksiyonu olarak ortalama sıfır ve Matern kovaryans fonksiyonunu kullanır. Önerilen yöntemin doğruluğunu değerlendirmek için, 10 farklı bölge üzerinden sonuçlar alınmıştır. Bu seçilen bölgeleri gösteren, kırmızı karelerle vurgulanmış, Şekil 3.1'de açıklanmaktadır. Her bölge, 111.111 km'lik bir alanı temsil eden 100*100'lük matrislerle temsil edilir. Bu bölgelerin gerçek sıcaklık haritaları, Şekil 3.2'de gösterilmiştir. Şekil 3.2'deki veri haritaları, Şekil 3.1'de kırmızı kare ile işaretlenen alanları temsil eder. Şekil 3.2'deki şekillerde turuncu olarak görünen bölgeler sıcaklığın arttığı bölgeleri temsil etmektedir. Şekillerde bazı verilerde tek bir ROI alanı olan sıcaklığın arttığı bölgeler bulunurken bazı şekillerde ise keskin bir ROI alanının bulunmadığı görülmektedir. Bu nedenle seçilen algoritmanın etkinliği farklı ROI alanlarında sahip verilerde test edilmiş olacaktır.



Şekil 3.1. Eğitim verilerinin önerilen yöntemle öğrenilmesi için belirlenen 10 farklı bölge.

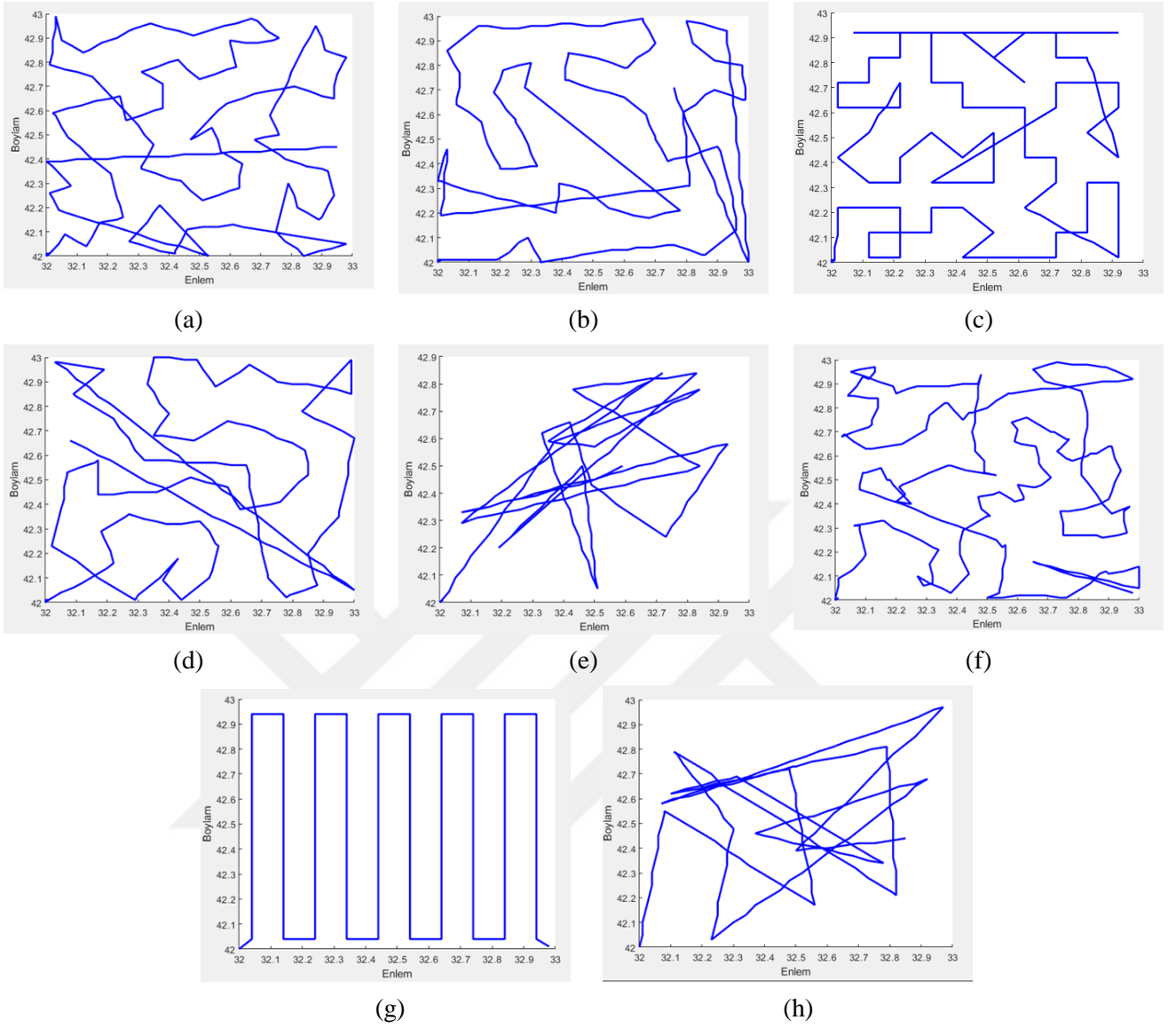


Şekil 3.2. Nasa Veri Setinden seçilen bölgelerdeki veri haritaları

3.1.2. Eğitim Verilerini Toplamak İçin Önerilen Güzergah Planlama Algoritmasıyla Elde Edilen Sonuçlar

Literatürde eğitim verilerini toplamak için geliştirilen güzergah planlama algoritmaları sınırlı ve az sayıdadır. Bu nedenle bu çalışmada eğitim verilerini toplamak için bir güzergah planlama algoritması önerilmiştir. Bölüm 2.2.3’de eğitim verilerini toplamak için geliştirilen güzergah planlama algoritması detaylı olarak verilmiştir. Şekil 3.3’de, örnek toplamak için 1100 km’lik bir bütçe ile robot tarafından oluşturulan yörüngeler gösterilmektedir. Bu analiz için maksimum 1100

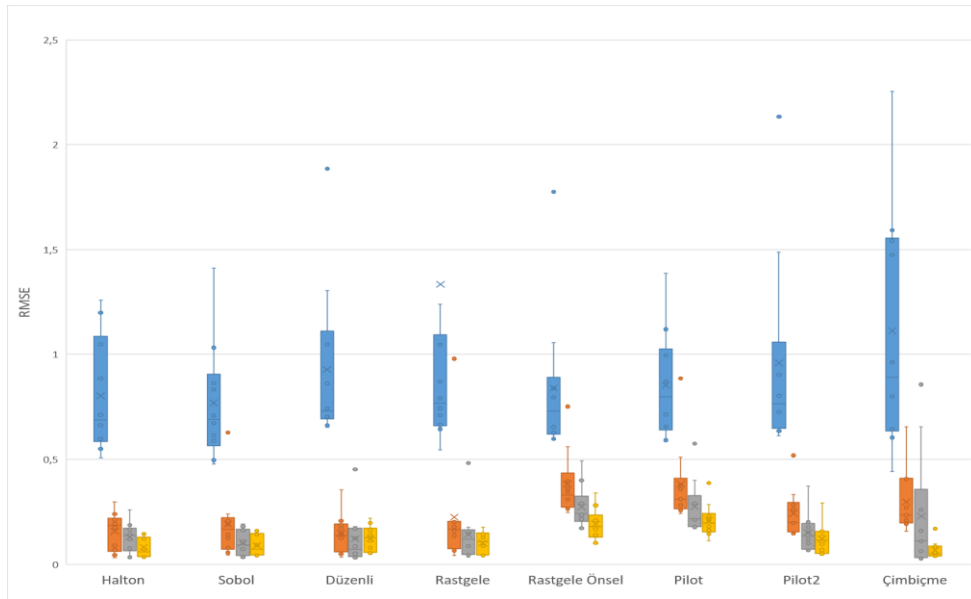
km'lik bir bütçe seçilmiştir. Şekil 3.3.a, Şekil 3.3.b, Şekil 3.3.c ve Şekil 3.3.d, sırasıyla, 1100 km bütçe altında Halton, Sobol, Düzenli ve Rastgele dizilimler kullanılarak elde edilen yörüngeler gösterilmektedir. Şekil 3.3.e, Kemna vd.[52] tarafından önerilen pilot yöntemi ile oluşturulan güzergahı göstermektedir. Ayrıca, Şekil 3.3.f, önerilen yöntemle miyop bir güzergah planlama algoritmasının entegre edilmesiyle elde edilen yöntemi göstermektedir. Miyop bir güzergah planlama algoritması ile mevcut bütçe azaldıkça daha fazla alan tarandığı gözlemlenmektedir. Ayrıca, Şekil 3.3.g, çim biçme rotasını temsil ederken, Şekil 3-3.h, oluşturulan yol noktalarını rastgele ziyaret ederek elde edilen yolu belirtmektedir. Toplamda, 1100 km bütçe içinde tüm yörüngelerden eşit aralıklarla 200 örnek toplanmıştır. Şekil 3.3.a, Şekil 3.3.b, Şekil 3.3.c ve Şekil 3.3.d incelendiğinde yol noktası dizilimleri farklı olmasına rağmen 4 güzergahtada alanın tamamından örnek toplandığı görülmüştür. Fakat Şekil 3-3.e'de ve Şekil 3.3.h'da rastgele yol noktalarının belirlenmesi robotun alanda rastgele hareket etmesine neden olmuş ve mevcut bütçesini verimli kullanmayarak ileri geri hareket etmesine neden olmuştur. Bu güzergahların aksine Şekil 3.3.d'de rastgele aday yol noktaları oluşturulmuştur. Fakat rastgele oluşturulan aday yol noktalarından birbirine en uzak yol noktalarının seçilmesi ve miyop bir güzergah planlama algoritmasının uygulanması alana yayılan bir güzergahın tanımlanmasını sağlamıştır. Şekil 3.3.f de ise Kemna vd. nin önermiş olduğu yöntemle miyop güzergah planlama algoritmasının eklenmesinin robotun mevcut bütçesini efektif kullanmasını sağlamış ve daha fazla örneğin toplanmasını sağlamıştır. Böylece alanda ileri geri hareketler yerine en yakınındaki örnekleri toplayacak güzergahı belirlemiştir.



Şekil 3.3. 1100 Km'de Oluşturulan Yörüngeler a) Halton Dizilimi b) Sobol Dizilimi c) Düzenli Dizilim d) Rastgele Dizilim e) Pilot f) Pilot₂ g) Çim Bıçme Rotası h) Rastgele Seçilen Yol Noktalarıyla Oluşturulan Güzergah

Yörünge planlama algoritmalarının performans değerlendirmesi, RMSE değerinin hesaplanmasına dayanmaktadır. 10 belirlenmiş bölge için 8 farklı yörünge belirlendi ve RMSE değeri dört farklı bütçe senaryosu için hesaplandı: sırasıyla mavi, turuncu, gri ve sarı kutular, 110 km, 330 km, 550 km ve 1100 km'lik bütçeleri temsil etmektedir. Halton, Sobol ve Düzenli aday yol noktaları kullanılarak oluşturulan yörüngeler analiz boyunca sabit kaldı ve yalnızca bir kez çalıştırıldı. Bununla birlikte, rastgele aday rota noktaları kullanılarak oluşturulan yörünge, 100 kez çalıştırılarak ortalama değer alındı. Elde edilen sonuçlar Kemna vd.[52] tarafından önerilen pilot yöntemi ve önerilen pilot yöntemine eklenen güzergah planlama algoritmasıyla (pilot₂) kıyaslandı. Bu yöntemlerde de rastgele bir seçim yapıldığı için 100 kez çalıştırıldı ve ortalaması alındı.

Şekil 3.4, 10 bölge için 8 yörünge planlama algoritmasının kullanımı durumunda RMSE değerinin değişimini göstermektedir. Kutular, 10 bölge boyunca gözlemlenen minimum ve maksimum RMSE değerleri arasındaki aralığı gösterir. Aykırı değerler noktalarla temsil edilirken, 10 bölge için ortalama RMSE değeri çarpı işareti ile gösterilir. Sonuçlar, Sobol dizileri kullanılarak oluşturulan güzergahta en düşük ortalama RMSE değerine ulaştığını göstermektedir. Önerilen yöntemde, örnekleme sayısı arttıkça örneklerin bilgilendiriciliğinin azaldığı gözlemlenmiştir. Dahası, 4 dizinin tümü için ortalama sonuçlar, rastgele dizide gözlemlenen aykırı değerler dışında genellikle uyumludur. Bu, aday yol noktalarının rastgele seçilmesi durumunda bile sonuçların diğer yöntemlerle karşılaştırıldığında gelişme gösterdiğini göstermektedir. Örnekleme sayısının artmasıyla örneklerin bilgilendiriciliğinde gözlemlenen azalma, örneklerin bilgilendiriciliğinin belirli bir miktardan sonra düştüğünü göstermektedir. . Ancak, algoritmanın gereksinim duyduğu hesaplama zamanı ile örnek sayısı arasındaki dengeyi göz önünde bulundurmak önemlidir. Ayrıca, rastgele oluşturulan aday yol noktalarının varlığı, belirli alanların keşfedilememesine yol açar. Önerilen algoritmanın genellikle çoğu alanı keşfetmeyi başardığı, ancak hala keşfedilmemiş bölgelerin bulunduğu göz önüne alındığında, aykırı değerlerin ortaya çıkmasına neden olur. Bu nedenle, aday yol noktalarının oluşturulmasında düzenli bir dağılımın sağlanması, algoritmanın etkinliğini garanti eder. Buna rağmen, rastgele dizilim yaklaşımının genel performansı, birçok diğer yöntemi hala geride bırakmaktadır. Genel olarak, Şekil 3.4'deki bulgular, Sobol dizileri tabanlı güzergah planlama algoritmasının daha iyi sonuç verirken 4 farklı dizilimde de düşük bütçede literatürdeki çalışmalara oranla daha iyi sonuçlar elde edilmiştir.



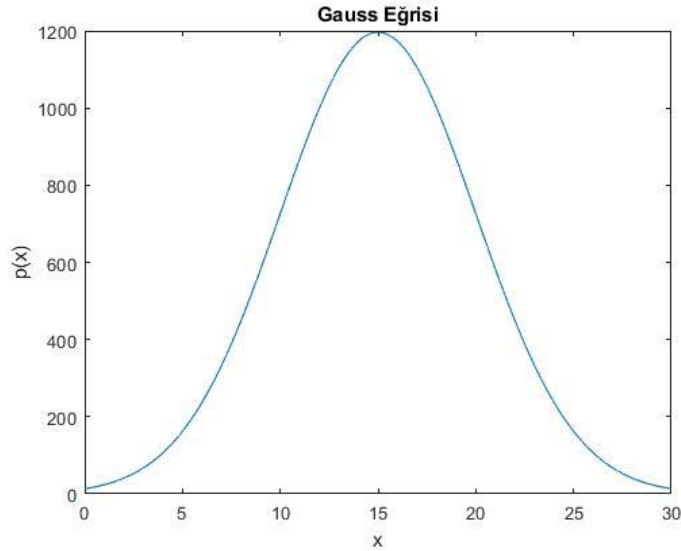
Şekil 3.4. Seçilen Yörüngelerdeki RMSE değişimi

3.2. ROI Alanlarından Örnek Toplamak için Geliştirilen Güzergah Planlama Algoritması

Bölüm 2.3.3'de fayda fonksiyonlarının ROI alanlarındaki başarımını karşılaştırmak için 1,2,3 ve 4 ROI alanı bulunan veri setine uygulanan güzergah planlama algoritması detaylı olarak verilmiştir. Bu bölümde dört test ortamında seviye belirleme, entropi ve UCB ödül fonksiyonları kullanılarak oluşturulan güzergahlar verilmiştir.

3.2.1. ROI Alanlarını Belirlemek için Oluşturulan Test Ortamı

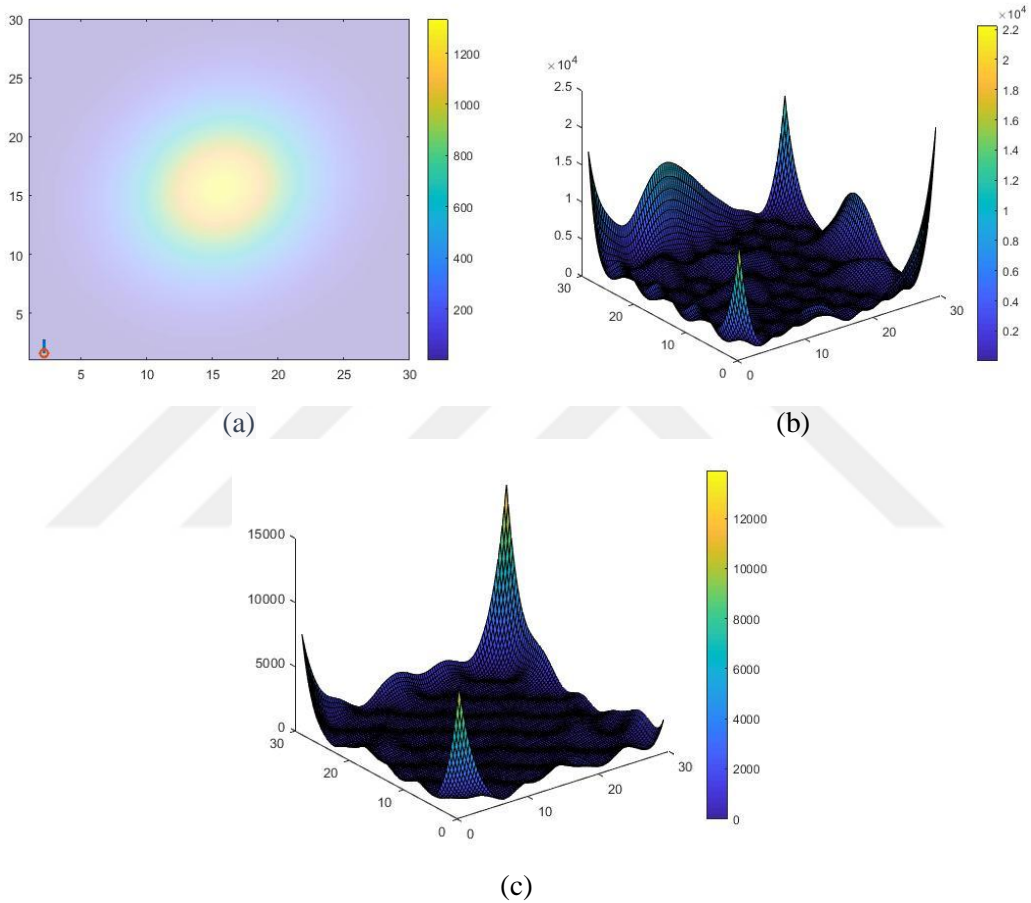
Bu çalışmanın amacı, kirlilik, tuzluluk, sıcaklık ve diğer ilgili faktörlerin yüksek seviyelerine sahip İlgi Alanı (ROI) bölgelerini tespit etmek için çevresel izleme yapmaktır. Bu amaçla, ROI bölgelerinde örnekleme yoğunluğunu artırmak üzere bir güzergah planlama algoritması geliştirilmiştir. Algoritmanın etkililiğini değerlendirmek için Gauss dağılımı kullanılarak oluşturulan bir veri haritasından faydalanılmıştır. Şekil 3.5, tek boyutlu Gauss dağılımını göstermektedir. Tek boyutlu veri grafiğinde, ortalaması 15 olan konumdan başlayan değerlerin nasıl değiştiği görülebilmektedir. Bu dağılımın merkezinde veri değeri 12000'e ulaşırken, 6 varyans mesafesinde 0'a yaklaşmaktadır. Test ortamı, 30×30'luk bir alanda oluşturulmuş olup, Gauss dağılımının merkezi ROI alanını temsil etmektedir.



Şekil 3.5. Gauss eğrisi

Şekil 3.6.a, tek bir ROI veri kümesi için gerçek haritayı sunmaktadır. İlk olarak, tek bir ROI haritası için eğitim verileri, belirli bir aralık içinde Halton dizisi kullanılarak belirlenen 50 veri noktasıyla oluşturulmuştur. Halton dizisi, sayıların ikili sayı sistemindeki karşılıklarını kullanarak yarı düzenli bir dizilim elde eder. Şekil 3.6.b, bu haritanın varyans grafiğini sunmaktadır. Varyans

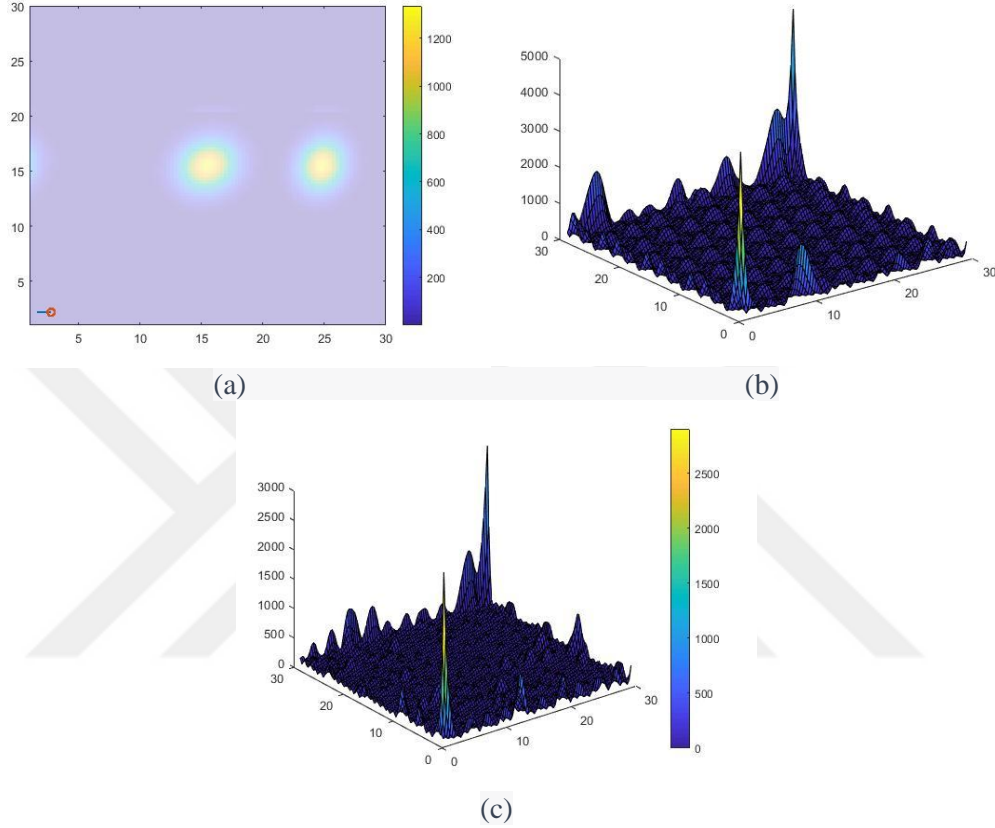
ortamın belirsizliğini temsil etmektedir. Başlangıçta varyans grafiğinin oldukça yüksek olduğunu göstermiştir. Bu çalışmada, başlangıç varyans değerini azaltmak için çapraz doğrulama yöntemi uygulanmıştır. Çapraz doğrulama yönteminde, Halton dizilimi ile belirlenmiş olan 60 eğitim verisinden 50 veri seçilmiştir. 60 verinin 50'si eğitim verisi, 10'u ise test verisi olarak ayrılmıştır. Şekil 3.6.b'de maksimum varyans değeri yaklaşık olarak 2.2×10^4 elde edilirken, çapraz doğrulama yöntemiyle bu değer yaklaşık olarak 12000'e düşürülmüştür. Tek bir ROI için tahmin edilen harita, Şekil 3.6'da gösterildiği üzere üç farklı ödül fonksiyonu için aynıdır.



Şekil 3.6. a) Tek ROI alanı olan test ortamı b) Başlangıçta Varyans Haritası c) Çapraz Doğrulama yöntemi uygulanan varyans haritası

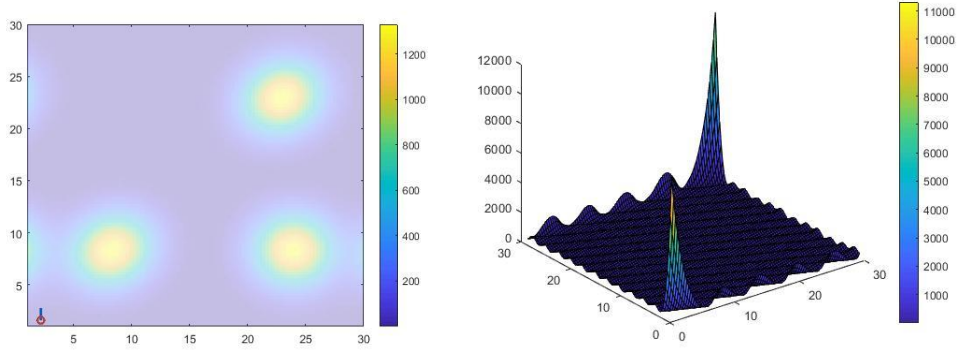
Şekil 3.7, iki Gauss dağılımı kullanılarak oluşturulan bir haritayı sunmaktadır. ROI alanlarının ortalaması (15,15) ve (15,25)'dir. Varyansı ise 6'dır. Şekil 3.7.a, gerçek veri haritası olarak tanımlanmıştır. Bu haritayı tahmin etmek, tek bir ROI alanına kıyasla daha zorlayıcıdır. Bu nedenle, bu veri kümesinde eğitim başlangıçta 100 eğitim veri noktası kullanılarak gerçekleştirilmiştir. Bu eğitim veri noktaları da Halton dizisi kullanılarak elde edilmiştir. Şekil 3-7.b, varyans haritasını göstermektedir. Şekil 3.7.c ise çapraz doğrulama yöntemi kullanılarak tahmin edilen varyans haritasını temsil etmektedir. Bu grafikte, 110 Halton dizilimi oluşturulmuş

ve 100 veri çapraz doğrulama yöntemi ile seçilmiştir. Üç ödül fonksiyonunun başlangıç koşulları ise aynıdır. Şekil 3.7.b’de maksimum varyans 5000 iken Şekil 3.7.c’de 3000’e düşürülmüştür. Gauss regresyon algoritması ile elde edilen varyans değeri ne kadar yüksekse belirsizlikte aynı oranda yüksektir. Çapraz doğrulama yöntemi ile maksimum belirsizlik %40 düşürülmüştür.

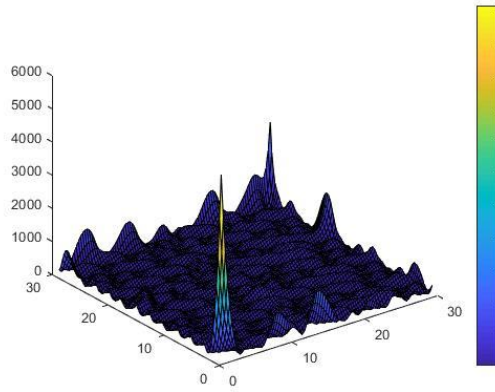


Şekil 3.7. a) İki ROI Gerçek Haritası b) Başlangıçta Varyans Haritası c) Çapraz Doğrulama yöntemi uygulanan varyans haritası

Şekil 3.8, üç ROI alanı bulunan bir ortam için oluşturulan veri setini göstermektedir. Bu veri seti, (10, 10), (10, 25) ve (25, 25) ortalama noktalarına sahip ve her biri 6 varyans değeri ile tanımlanan Gauss dağılımları kullanılarak oluşturulmuştur. Eğitim sürecinin başlangıcında, modelin öğrenmesi için 150 eğitim verisi Halton dizisi yöntemine göre üretilmiştir. Şekil 3.8.c’de ise çapraz doğrulama yöntemi uygulanan veri setinin başlangıç varyans haritasını vermektedir. 160 halton dizilimine ait eğitim veri setinden 150 en iyi örnek seçilmiştir. Grafikte varyans grafiğinin yarı yarıya düştüğü görülmektedir.



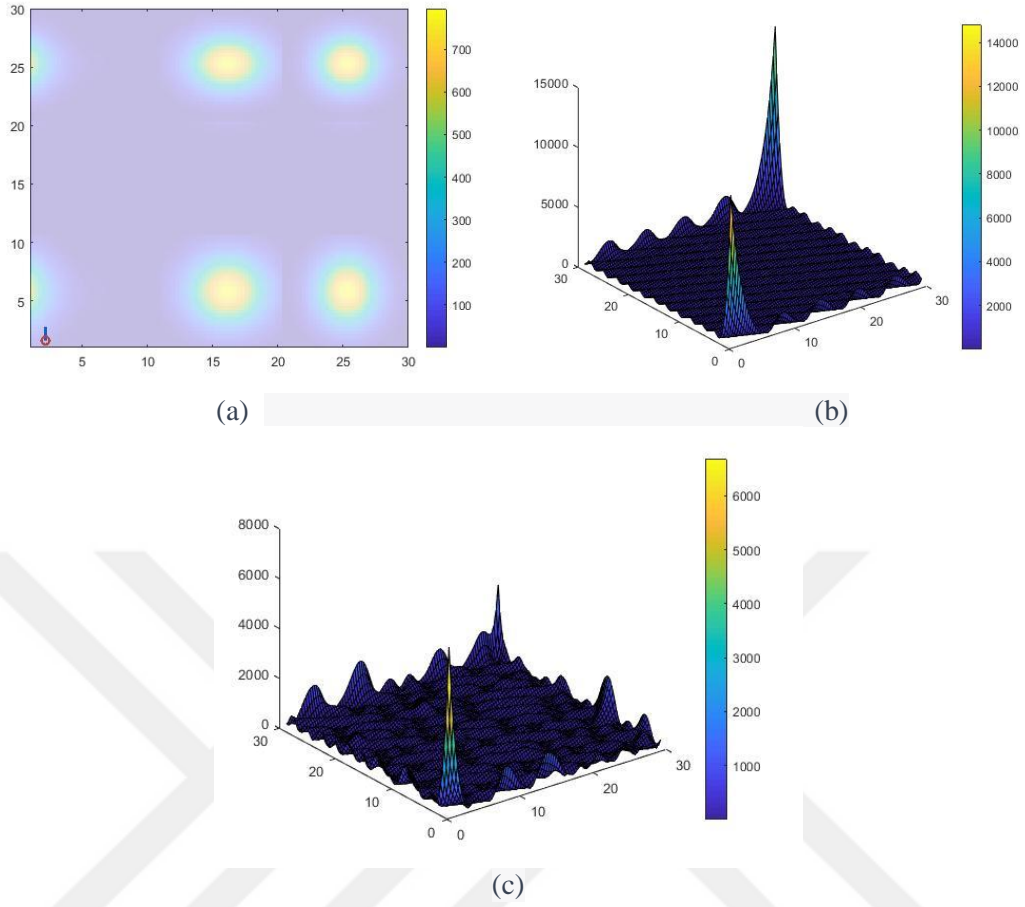
(a) (b)



(c)

Şekil 3.8. a) Üç ROI Gerçek Haritası b) Başlangıçta Varyans Haritası c) Çapraz Doğrulama yöntemi uygulanan varyans haritası

Şekil 3.9’da, dört ROI alanı bulunan test ortamı görülmektedir. Bu veri seti, her biri farklı konumlarda bulunan ve belirli varyans değerlerine sahip Gauss dağılımlarından oluşmaktadır. Veri seti oluşturulurken (10, 10), (10, 25), (25, 10) ve (25, 25) ortalama noktalarına sahip Gauss dağılımları kullanılmıştır ve her bir dağılımın varyansı 6 olarak belirlenmiştir. İki tahminde de 200 veri kullanılarak ortam öğrenilmiştir. Şekil 3.9.a ‘da başlangıç varyans değerinde maksimum değer 15000’e yaklaşırken çapraz doğrulama yöntemi uygulanan veri setinde başlangıç değeri 8000 civarındadır. Şekil 3.6, Şekil 3.7, Şekil 3.8 ve Şekil 3.9 incelendiğinde veri setinin karmaşıklığı arttıkça başlangıç varyans haritasının daha çok düştüğü gözlenmiştir.



Şekil 3.9. a) Dört ROI Gerçek Haritası b) Başlangıçta Varyans Haritası c) Çapraz Doğrulama yöntemi uygulanan varyans haritası

3.2.2. ROI Alanlarında Adaptif Örnekleme

Bu bölümde, farklı ödül fonksiyonlarının keşif ve sömürü dengesindeki performansı araştırılmıştır. Keşif, belirli bir alandaki bilgi eksikliğini gidermek için örnek toplamayı ifade ederken, sömürü, zaten bilinen bölgelerde daha yoğun bir örnekleme yaparak ROI alanlarının değerlendirilmesine odaklanmaktadır. Bu çalışmadaki hedef, ödül fonksiyonlarının ROI alanlarında performansını değerlendirmek üzere RMSE ve varyans değerleri konularında analiz yapmaktır. Şekil 3.10, tek bir ROI veri seti için belirlenen yolları göstermektedir ve her bir şekil farklı ödül fonksiyonları ve örnekleme sayıları ile elde edilen yol haritalarını sunmaktadır.

Şekil 3.10.a ve Şekil 3.10.b'de L_{entropi} ödül fonksiyonu (Denklem 31) kullanılmıştır. Şekil 3.10.a, 150 örnek toplamak için izlenen yolu gösterirken, Şekil 3.10.b aynı fonksiyonla 300 örnek toplamak için oluşturulan yolu göstermektedir. Görüldüğü üzere, artan örnekleme ile daha geniş bir alan üzerinde daha seyrek ama daha iyi bir dağılım sağlanmıştır.

Şekil 3.10.c ve Şekil 3.10.d ise L_{LS} ödül fonksiyonu (Denklem 30) ile elde edilen yolları göstermektedir. L_{LS} fonksiyonu, ROI alanlarının sınırlarından örnekler toplayarak alanı

sınıflandırmaktadır. Eşik değerinin üzerinde bir ödül değeri elde eden düğümler, ROI alanı olarak kabul edilerek ilgili kümelere eklenir. Bu, alanın daha iyi bir şekilde tanımlanmasına ve doğru tahmin edilmesine yardımcı olur. Şekil 3.10.c'de 150 örnek, 3-10.d'de ise 300 örnek toplanmıştır.

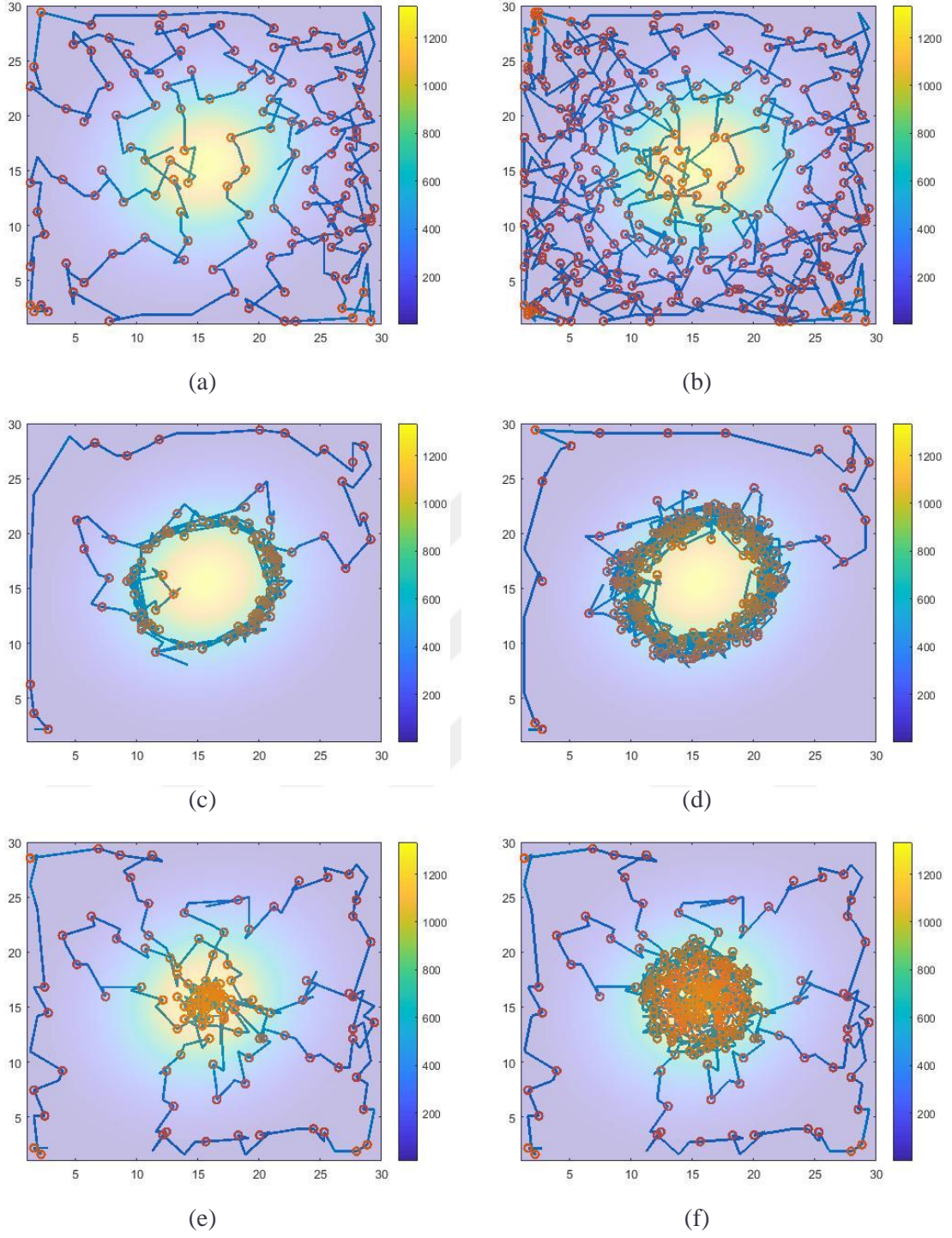
Şekil 3.10.e ve Şekil 3-10.f, L_{UCB} ödül fonksiyonunu (Denklem 27) kullanarak toplanmış örnekleri göstermektedir. L_{UCB} fonksiyonu, hem keşif hem de sömürü arasında denge kurarak, GP'den elde edilen ortalama ve varyans değerlerini kullanır. Bir bölgenin düşük belirsizlik düzeyine sahip bir haritada maksimum değere sahip olması, o bölgenin ROI olarak kabul edilmesini sağlar. Bu şekilde, Şekil 3.10.e 150 örnekle gerçekleştirilen keşif yolunu, Şekil 3.10.f ise 300 örnekle bu strateji kullanılarak elde edilen yolu göstermektedir.

Entropiye dayalı ödül fonksiyonu, tüm alanı verimli bir şekilde öğrenmeyi sağlarken, ROI alanlarında daha yoğun bir örnekleme sağlar. Bu fonksiyon, yeni örnekler toplarken maksimum belirsizlik içeren noktaları ödüllendirir. Bu, belirsizliklerin büyük ölçüde azaltılmasına ve alanın daha iyi tahmin edilmesine yardımcı olur.

Analizin sonucunda, üç ödül fonksiyonu da ROI alanlarından örnekler toplamakta başarılı olmuştur, ancak farklı stratejiler izlemeleri, çeşitli keşif ve sömürü dengeleri sunar. Entropi ödül fonksiyonu, yeni örneklemlerin maksimum belirsizliğe sahip noktalardan toplanmasını hedefler. Bu yaklaşım, ROI alanlarının keşfinde daha geniş bir bilgi tabanı oluşturulmasına ve genel belirsizliğin azaltılmasına yardımcı olur. Şekilde gözlemlendiği üzere, entropi ödül fonksiyonu kullanılarak elde edilen dağılımlar, süreç boyunca alanın daha iyi anlaşılmasına katkıda bulunur.

Seviye belirleme ödül fonksiyonu, ROI alanlarının sınırlarını belirlemeyi ve bu sınır bölgelerinden örnekler toplamayı amaçlar. Eşik değerinin belirlenmesi ve bu değerin üzerinde ödül değerine sahip düğümlerin kümelenmesi, ROI alanının daha detaylı bir şekilde tanımlanmasına imkan tanır. Bu şekilde, ROI alanının etrafındaki verilerin toplanması sağlanır ve alan daha etkili bir şekilde sınıflandırılır. Şekil 3.10.c ve Şekil 3.10.d'ye bakıldığında, seviye belirleme fonksiyonunun bu strateji ile başarılı sonuçlar elde ettiği görülmektedir.

UCB ödül fonksiyonu ise özellikle keşif ve sömürü arasındaki dengeyi kurmada güçlü bir araçtır. GP'den elde edilen ortalama ve varyans değerlerini kullanarak, keşif yapmak için varyansın yüksek olduğu bölgeler üzerine odaklanır ve belirsizliğin azaltılması amaçlanır. Öte yandan, bilinen ROI alanlarında ise tahmin edilen maksimum değerlere sahip bölgelerde yoğunlaşarak veri toplanmasını sağlar. Şekil 3.10.e ve Şekil 3.10.f incelendiğinde, UCB fonksiyonunun bu dengeyi kurarak ROI alanlarından verimli örnekler topladığı görülmektedir.



Şekil 3.10. Tek ROI için adaptif örnekleme a) L_{PI} ödül fonksiyonu kullanılarak toplanan 150 örnek b) L_{PI} ödül fonksiyonu kullanılarak toplanan 300 örnek c) a) $L_{LevelSet}$ ödül fonksiyonu kullanılarak toplanan 150 örnek d) $L_{LevelSet}$ ödül fonksiyonu kullanılarak toplanan 300 örnek e) a) L_{UCB} ödül fonksiyonu kullanılarak toplanan 150 örnek f) a) L_{UCB} ödül fonksiyonu kullanılarak toplanan 300 örnek

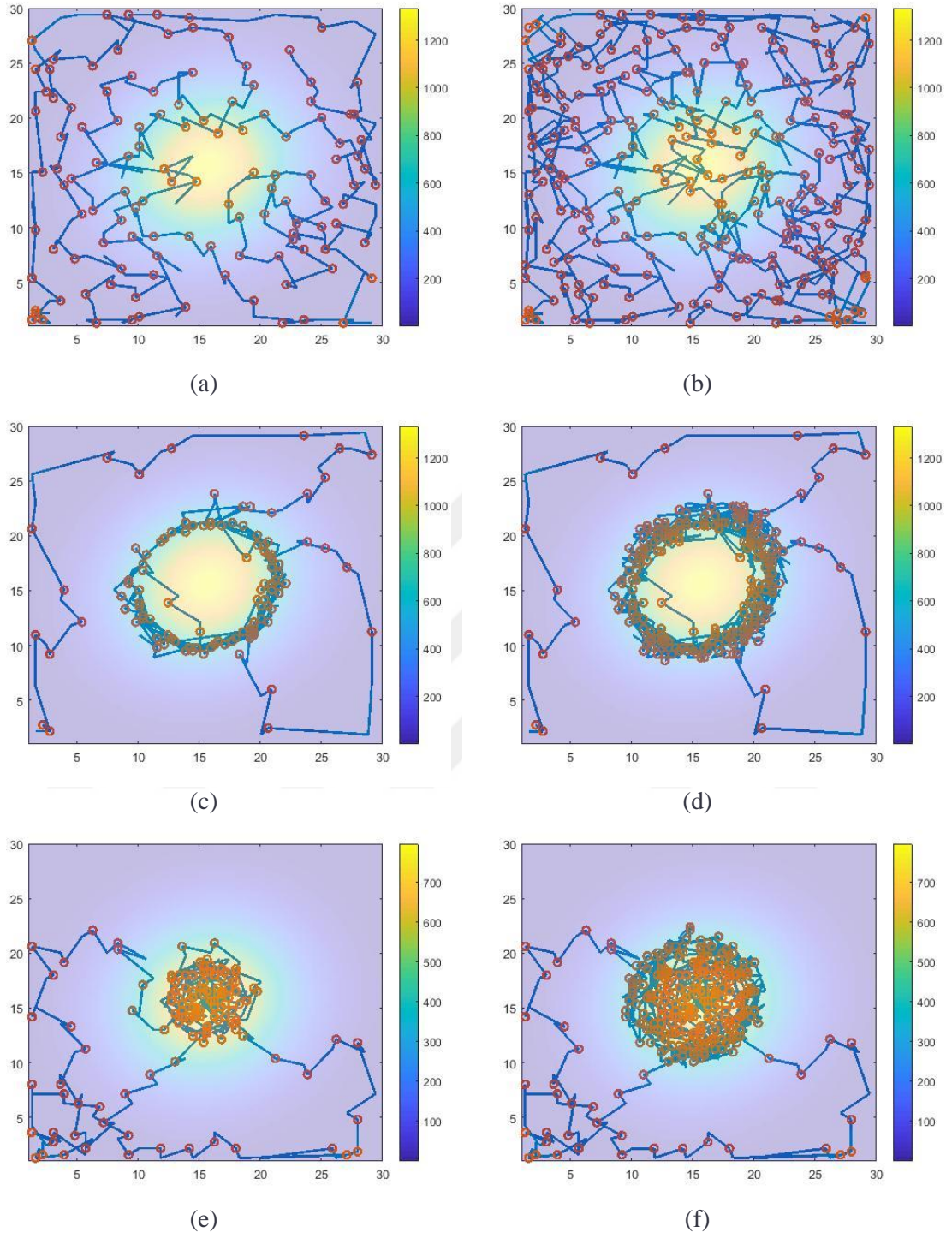
Yüksek varyansa sahip olan varyans haritasında elde edilen rotayı gösteren Şekil 3.10'daki sonuçlar, başlangıçta gözlemlenen veri dağılımıyla ilgilidir. Bu haritada elde edilen rota, veri setinin geniş varyans aralığı göz önünde bulundurularak oluşturulmuştur. Buna karşın, çapraz doğrulama yöntemi eklendikten sonra düşük varyanslı veri haritasında ortaya çıkan rotayı temsil eden Şekil 3.11, daha optimize edilmiş ve güvenilir sonuçlar sunmaktadır. Şekil 3.6'da ise, başlangıç varyans haritası ile çapraz doğrulama yöntemi uygulandıktan sonra elde edilen yeni varyans haritası karşılaştırmalı olarak incelenmektedir. İlk aşamada, veri setindeki maksimum varyans değeri $2,5 \times 10^4$ gibi oldukça yüksek bir seviyede iken, çapraz doğrulama yönteminin uygulanmasıyla bu değer büyük ölçüde azalmakta ve 15000 değerine düşmektedir. Bu önemli varyans düşüşü, verideki belirsizliklerin azalmasını ve daha doğru tahminlerin yapılmasını sağlamaktadır. Ancak, bu düşüşe rağmen maksimum varyans değerinin hala yüksek olması, modelin bazı bölgelerdeki belirsizliklerinin tamamen çözülemediğini göstermektedir. Hem Şekil 3.10 hem de Şekil 3.11'de benzer rotaların oluşması, varyansın hala yüksek olmasının bir sonucudur. Yüksek varyans, modelin bazı bölgelerde belirsizlik taşımasını ve rotanın optimize edilmesini zorlaştırmaktadır. Bununla birlikte, görsel analizler UCB ödül fonksiyonu kullanılarak oluşturulan güzergah planlama algoritmasında büyük oranda sömürüye (exploitation) odaklanıldığını göstermektedir. Özellikle Şekil 3.10.e ve Şekil 3.10.f ile kıyaslandığında, Şekil 3.11.e ve Şekil 3.11.f'de daha fazla sayıda ROI alanlarından örnek toplandığı açıkça görülmektedir. Bu, modelin yüksek getirili alanlardan daha fazla veri toplamak suretiyle belirsizlikleri azaltma ve daha isabetli tahminlerde bulunma stratejisini benimsediğini göstermektedir.

Başlangıç varyans haritasında belirsizliğin neden olduğu geniş dağılım ve olası veri noktalarının rastgele seçimi, rotanın optimize edilerek güzergahtaki her bir noktanın doğrulama sürecine katkıda bulunma gerekliliğini vurgulamaktadır. Çapraz doğrulama yöntemi, çeşitli alt setler üzerinden modelin performansını değerlendirmekte ve daha düşük varyansla birlikte daha sağlam sonuçlar sağlamaktadır. Bu yöntem sayesinde varyansın düşmesi, daha iyi kalibrasyonların yapılmasına olanak tanırken, modelin belirsiz alanlara daha fazla odaklanarak veri toplamasını sağlar. Örneğin, Şekil 3.6'daki karşılaştırmalarda görüldüğü üzere, başlangıç varyans haritasında belirlenen yüksek varyanslı bölgeler, metodun uygulanması sonrası daha düşük varyanslı olarak yeniden hesaplanmaktadır. Bu durum, modelin genel doğruluğunu arttırmakta ve belirsiz bölgelerin daha az varyans içermesini sağlamaktadır. Çapraz doğrulama stratejisiyle yapılan iyileştirmelerin sonucunda, Şekil 3.11'deki güzergah planlaması, modelin öğrenme sürecinin optimizasyonuna büyük katkı sağlamaktadır. Yine de maksimum varyans değerinin yüksek kalması, modelin belirli bölgelerde halen tam anlamıyla kesinlik sağlayamadığını ve bu nedenle daha fazla veri toplama ihtiyacını işaret eder. UCB ödül fonksiyonunun kullanıldığı güzergah planlama algoritması, belirsiz ve yüksek getirili alanlardan daha fazla veri toplayarak bu bölgelerdeki belirsizliği en aza indirme amacını güder. Şekil 3.10.e ve Şekil 3-10.f ile kıyaslandığında, Şekil 3.11.e ve Şekil 3.11.f'de

belirlenen rotaların daha fazla ROI alanını kapsamaları, bu stratejinin etkinliğini göstermektedir. Model, maksimum varyansın yüksek kaldığı alanlarda veri toplamaya devam ederek, rotanın optimize edilmesine ve belirsizliklerin giderilmesine katkı sağlar.

Sonuç olarak, çapraz doğrulama yöntemi sayesinde önemli ölçüde varyans düşüşü elde edilmiş olmasına rağmen, maksimum varyans seviyelerinin hala yüksek olması, modelin sonuçlarının tam optimize edilemediğini göstermektedir. Yüksek varyansa sahip alanların veri toplama sürecinde belirsizlik yaratması, rotaların benzer şekillerde oluşmasına neden olurken, çapraz doğrulama yöntemi sayesinde daha düşük varyansla elde edilen veri setlerinde, rota daha doğrusal ve odaklı hale gelmektedir.





Şekil 3.11. Çapraz doğrulama yöntemi uygulanan Tek ROI alanında adaptif örnekleme a) L_{PI} ödül fonksiyonu kullanılarak toplanan 150 örnek b) L_{PI} ödül fonksiyonu kullanılarak toplanan 300 örnek c) a) $L_{LevelSet}$ ödül fonksiyonu kullanılarak toplanan 150 örnek d) $L_{LevelSet}$ ödül fonksiyonu kullanılarak toplanan 300 örnek e) a) L_{UCB} ödül fonksiyonu kullanılarak toplanan 150 örnek f) a) L_{UCB} ödül fonksiyonu kullanılarak toplanan 300 örnek

Şekil 3.12'de sunulan üç farklı ödül fonksiyonu kullanılarak, 2-ROI veri kümesi ile planlan güzergah planlama çalışmasının sonuçları verilmiştir. Şekil 3.12.a, L_{entropi} ödül fonksiyonu kullanılarak 150 örnek toplanmış ve bu örnekler belirgin şekilde birinci ROI alanına odaklanmıştır. Şekil 3.12.b'de aynı L_{entropi} ödül fonksiyonu kullanılmış ve 300 örnek toplanmıştır. Artan örnek sayısının etkisiyle, daha geniş bir alan üzerinde örnekleme yapılmış ve veri setinin daha homojen bir şekilde temsil edildiği gözlemlenmiştir.

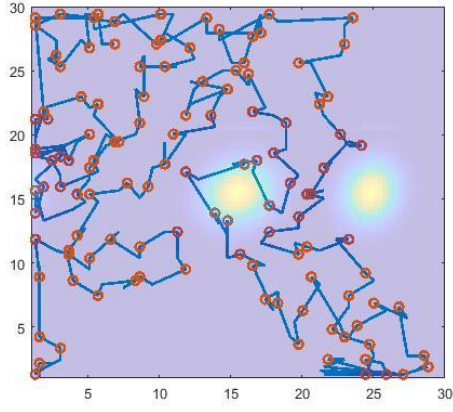
Şekil 3.12.c, L_{LS} ödül fonksiyonunu kullanarak ilk ROI alanında 150 örnek toplanmasını göstermektedir. Burada, örneklerin yoğunlukla belirli bölgelerde toplandığı ve ROI alanının etrafında yoğunlaştığı görülmektedir. Şekil 3.12.d'de ise, aynı L_{LS} ödül fonksiyonu ile bu sefer 300 örnek toplanmıştır. Bu durumda, örneklemenin yapıldığı alanın genişlediği ve veri setinin daha kapsamlı bir şekilde örneklendiği anlaşılmaktadır.

Şekil 3.12.e ve Şekil 3.12.f, L_{UCB} ödül fonksiyonunun kullanıldığı güzergahları temsil etmektedir. Şekil 3.12.e'de, 150 örnekle elde edilen güzergah planı gösterilmiştir. Burada, örneklerin ilk ROI alanında yoğunlaştığı ve diğer alanlarda daha az örnek toplandığı görülür. Şekil 3.12.f'de ise, 300 örnekle daha geniş bir alana yayıldığı gözlemlenmiştir. Ancak, yine de ilk ROI alanında yoğunlaşma devam etmektedir.

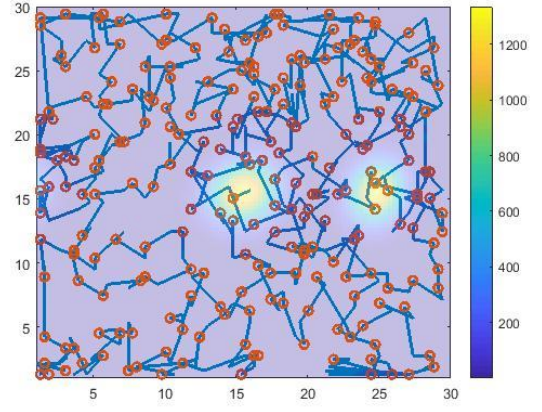
150 örnek ile üç ödül fonksiyonunda ilk ROI alanını başarıyla örneklemediği anlaşılmaktadır. L_{entropi} fonksiyonu, ilk ROI alanında makul sayıda örnek toplamıştır. Diğer taraftan, L_{UCB} fonksiyonuyla daha yoğun bir örnek yoğunluğuna ulaşılmıştır. L_{LS} fonksiyonu kullanıldığında ise, örneklerin daha iyi dağıldığı ve ROI alanının etrafında yoğunlaştığı gözlemlenmiştir. Bu durum, L_{LS} fonksiyonunun potansiyel olarak daha iyi bir keşif yapabildiğini göstermektedir.

Hem L_{entropi} hem de L_{LS} ödül fonksiyonlarında, 150 örnekle ikinci ROI alanından da bazı örnekler toplanmış olup bu durum, bu fonksiyonların keşif kapasitesinin iyi olduğunu ortaya koymaktadır. Ancak, L_{UCB} fonksiyonu kullanılarak yapılan örnekleme, ilk ROI alanına daha fazla odaklandığından, ikinci ROI alanının tespit edilmesinde yetersiz kalmıştır. Bu, L_{UCB} fonksiyonunun keşif ve sömürü arasında dengenin korunmasında zorluk yaşayabileceğini göstermektedir.

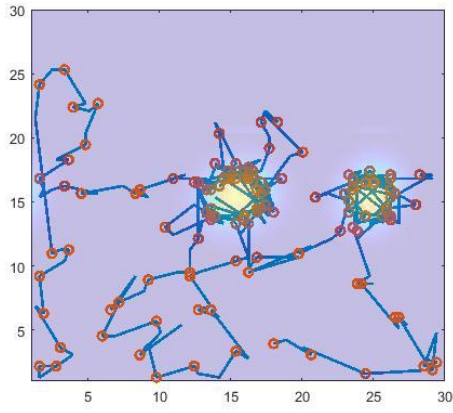
Bu durum, L_{UCB} yönteminin çoklu ROI veri kümelerinde, ROI alanlarını tanımlarken yerel maksimumlara sıkışma eğilimi gösterebileceği ve bu nedenle yerel maksimumları göz ardı edebileceği fikrini destekler niteliktedir. Sonuç olarak, L_{LS} ve L_{entropi} ödül fonksiyonlarının daha iyi performans gösterdiği ve daha olumlu sonuçlar verdiği açıkça görülmektedir.



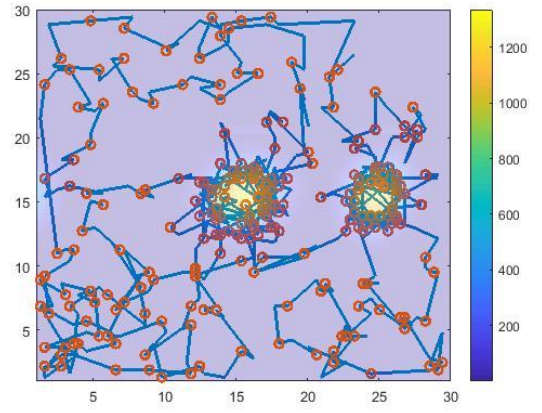
(a)



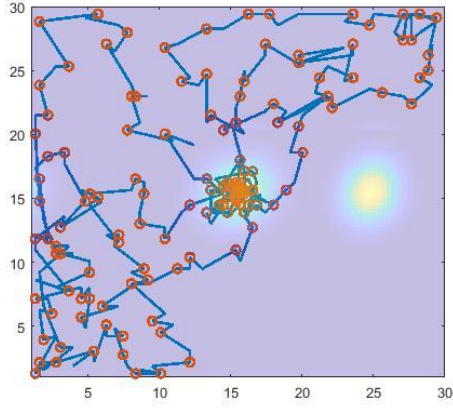
(b)



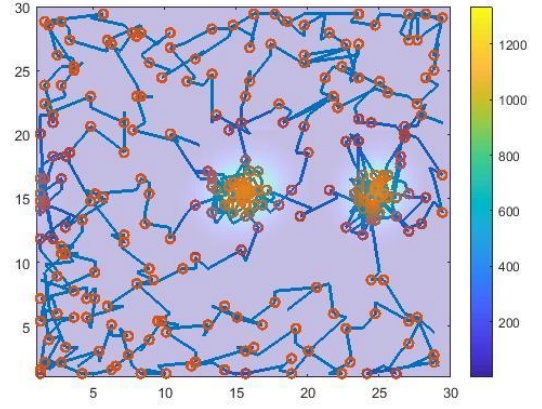
(c)



(d)



(e)

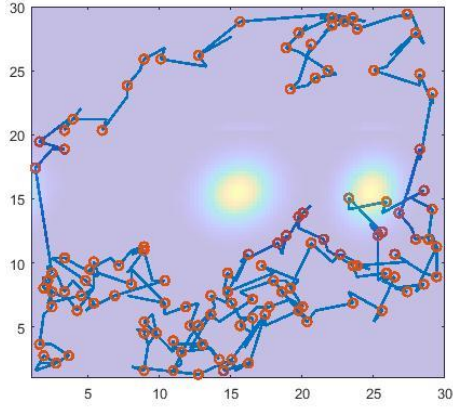


(f)

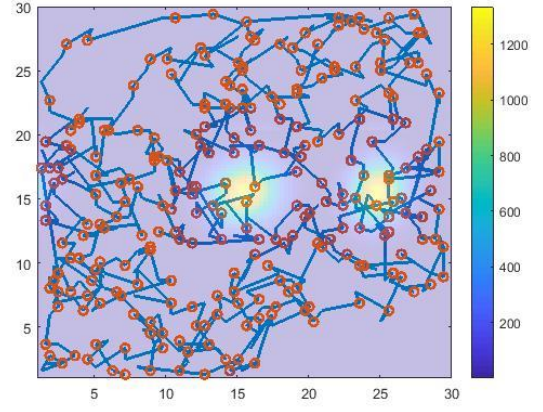
Şekil 3.12. İki ROI için adaptif örnekleme a) L_{PI} ödül fonksiyonu kullanılarak toplanan 150 örnek b) L_{PI} ödül fonksiyonu kullanılarak toplanan 300 örnek c) a) $L_{LevelSet}$ ödül fonksiyonu kullanılarak toplanan 150 örnek d) $L_{LevelSet}$ ödül fonksiyonu kullanılarak toplanan 300 örnek e) a) L_{UCB} ödül fonksiyonu kullanılarak toplanan 150 örnek f) a) L_{UCB} ödül fonksiyonu kullanılarak toplanan 300 örnek

Şekil 3.13'te, iki ROI alanı bulunan bir veri setine çapraz doğrulama yöntemi uygulanmıştır. Bu yöntem, veri setinde bulunan iki ROI ile Şekil 3.7'de gösterildiği gibi maksimum varyansı yaklaşık 5000'den 3000'e düşürmeyi başarmıştır. Şekil 3.13.a ve Şekil 3.13.b'de, entropi ödül fonksiyonu kullanılarak Şekil 3.12.a ve Şekil 3.12.b'ye benzer bir rota oluşturulmuş ve bu sayede maksimum varyans değeri düşürülmeye çalışılmıştır. Entropi ödül fonksiyonunun burada gösterilen etkisi, veri setini daha dengeli ve anlaşılır hale getirmeye yönelik olmuştur. Seviye belirleme ödül fonksiyonunun performansı, 150 örnek alındığında Şekil 3.12.c ile kıyaslandığında farklı bir yaklaşım benimsemiştir. Şekil 3.13.c'de, seviye belirleme ödül fonksiyonu doğrudan ROI alanlarına odaklanmıştır. Bu durum, veri toplama sürecindeki örneklerin daha stratejik ve hedefe yönelik olmasını sağlamıştır. Şekil 3.13.d'ye geldiğimizde, ROI alanında daha fazla örnek toplandığı, ancak bunun yanında keşif işlemlerinin de yapıldığı görülmektedir. Bu, veri setinin daha geniş bir yelpazede incelenmesini ve çeşitli varyasyonların anlaşılmasını sağlamıştır.

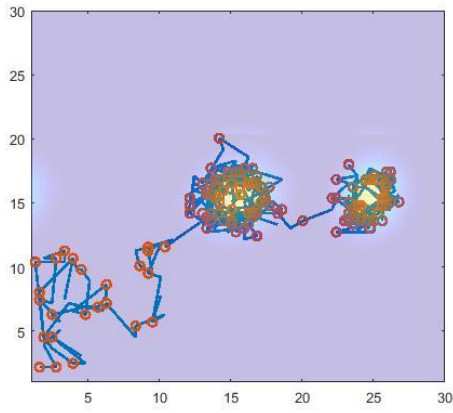
Öte yandan, UCB ödül fonksiyonunun 150 örnek için yalnızca ilk ROI alanından veri topladığı görülmüştür (Şekil 3.12.e). Bu, UCB ödül fonksiyonunun düşük varyans değerinde yerel maksimumda takılı kalmasına ve genel maksimumu kaçırmasına neden olmuştur. Ancak, Şekil 3.13.e'de bu sorunun aşılabildiği ve 150 örnek ile iki ROI alanından da veri toplanabildiği gösterilmiştir. Bu, UCB ödül fonksiyonunun daha iyi optimize edilerek performansının artırılabilirliğini kanıtlamaktadır. Son olarak, Şekil 3.13.f'de, Şekil 3.12.f'ye oranla daha az keşif yapıldığı ve ROI alanındaki örnek miktarının arttırıldığı gözlemlenmektedir. Bu durum, veri toplamanın daha verimli hale getirildiğini ve ROI alanına daha fazla odaklanıldığını göstermektedir. Keşif miktarının azaltılması, veri toplama sürecinde daha hedefe yönelik bir yaklaşım benimsenmesine olanak tanımıştır. Bu sayede, ROI alanı hakkında daha detaylı ve spesifik bilgi elde edilmiştir.



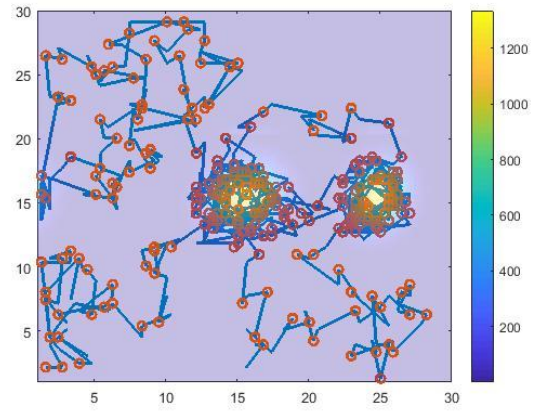
(a)



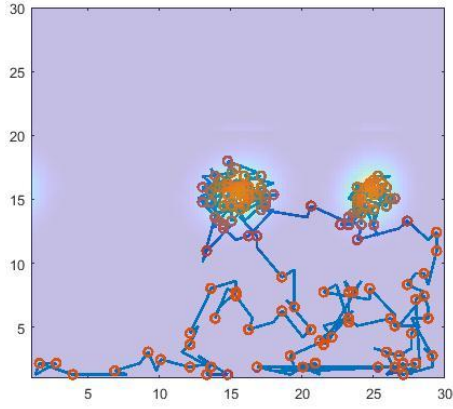
(b)



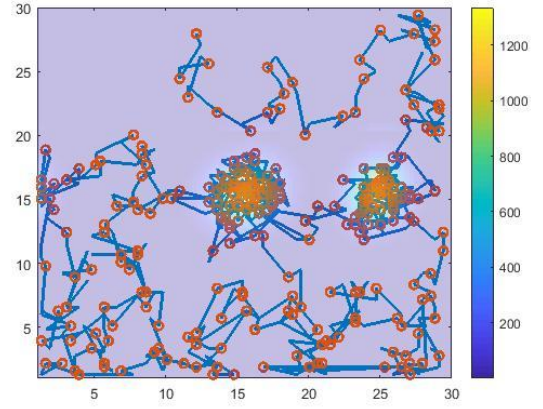
(c)



(d)



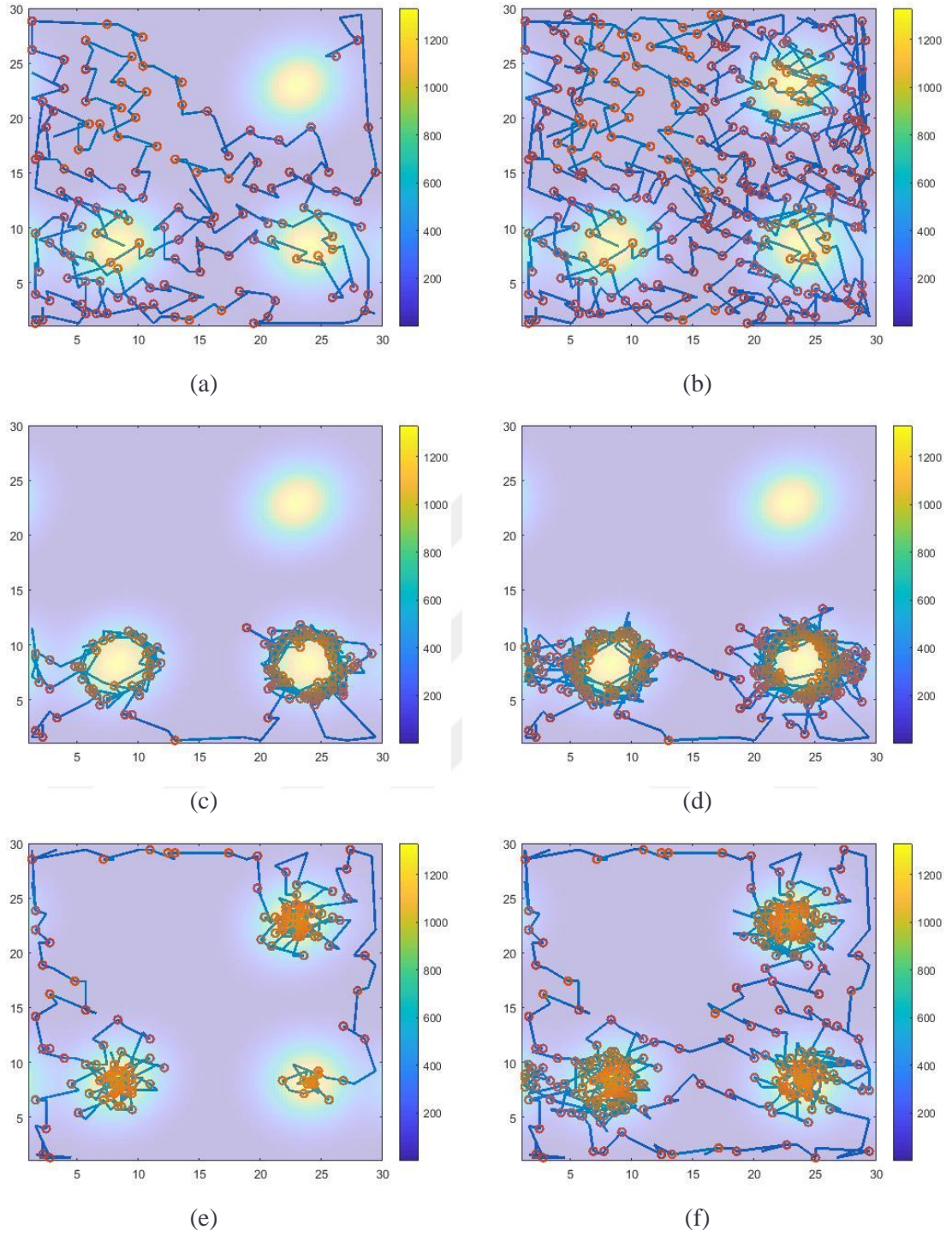
(e)



(f)

Şekil 3.13. Çapraz doğrulama yöntemi uygulanan İki ROI'li alanda adaptif örnekleme a) L_{PI} ödül fonksiyonu kullanılarak toplanan 150 örnek b) L_{PI} ödül fonksiyonu kullanılarak toplanan 300 örnek c) a) $L_{LevelSet}$ ödül fonksiyonu kullanılarak toplanan 150 örnek d) $L_{LevelSet}$ ödül fonksiyonu kullanılarak toplanan 300 örnek e) a) L_{UCB} ödül fonksiyonu kullanılarak toplanan 150 örnek f) a) L_{UCB} ödül fonksiyonu kullanılarak toplanan 300 örnek

Şekil 3.14'de 3 ROI alanı olan veri seti için 3 ödül fonksiyonunu kullanmıştır. Şekil 3.14.a ve Şekil 3.14.b'de, entropi ödül fonksiyonu kullanılarak toplanan verilerin güzergahı gösterilmektedir. Şekil 3.14.a'da 150 örnekle 2 ROI alanından veri toplandığı gözlenmektedir. Ancak, 3. ROI alanından bu örnekleme sürecinde hiç veri toplanmamıştır. Bu durum, entropi ödül fonksiyonunun belirli bölgelerdeki belirsizliği azaltmaya odaklanarak, diğer potansiyel önemli ROI bölgelerini ihmal edebileceğini göstermektedir. Şekil 3.14.b'ye bakıldığında ise, 3. ROI alanından 300 örnek toplandıktan sonra veri toplanmaya başlanmıştır. Yine de, bu yaklaşımın eksikliklerini ve sadece belirli bölgelere odaklanarak genel optimum çözümü ihmal etme riskini açıkça ortaya koymaktadır. Seviye belirleme ödül fonksiyonuyla oluşturulan rotalar Şekil 3.14.c ve Şekil 3.14.d'de gösterilmektedir. Şekil 3.14.c'de, 150 örnekle yine yalnızca 2 ROI alanı etrafından örneklerin toplandığı ve bu örneklerin 3. ROI alanını kapsamadan gerçekleştiği görülmektedir. Daha fazla örnek toplandığında (Şekil 3.14.d'de 300 örnek), yine aynı iki ROI alanına odaklanmış ve 3. ROI alanından örnek toplamadan devam etmiştir. Bu sonuçlar, seviye belirleme ödül fonksiyonunun yerel maksimum noktasına odaklanarak genel maksimum noktayı kaçırma riskine işaret etmektedir. Dolayısıyla, seviye belirleme ödül fonksiyonu yerel çözümlere odaklanmakta ve daha büyük ölçekte optimal sonuçları göz ardı etmektedir. Son olarak, UCB ödül fonksiyonu ile toplanan verilerin rotaları Şekil 3.14.e ve Şekil 3.14.f'de gösterilmektedir. Bu fonksiyonla çalışıldığında, daha dengeli ve kapsamlı bir veri toplama stratejisi izlenmiştir. Özellikle Şekil 3.14.e'de, 150 örnekte bile üç ROI alanından da örneklerin toplandığı dikkat çekmektedir. Bu durum, UCB ödül fonksiyonunun diğer iki fonksiyona kıyasla daha etkili bir şekilde hem yerel hem de genel maksimum noktaları yakalayabildiğini göstermektedir. Şekil 3.14.f'de, 300 örnek toplandığında dahi üç ROI alanından da örneklerin etkin bir şekilde toplandığı gözlenmektedir. UCB ödül fonksiyonunun geniş bir keşif-stratejisine sahip olması, hem çeşitli ROI alanlarında veri toplanmasını hem de genel optimum çözüme daha hızlı ulaşılmasını sağlamaktadır.

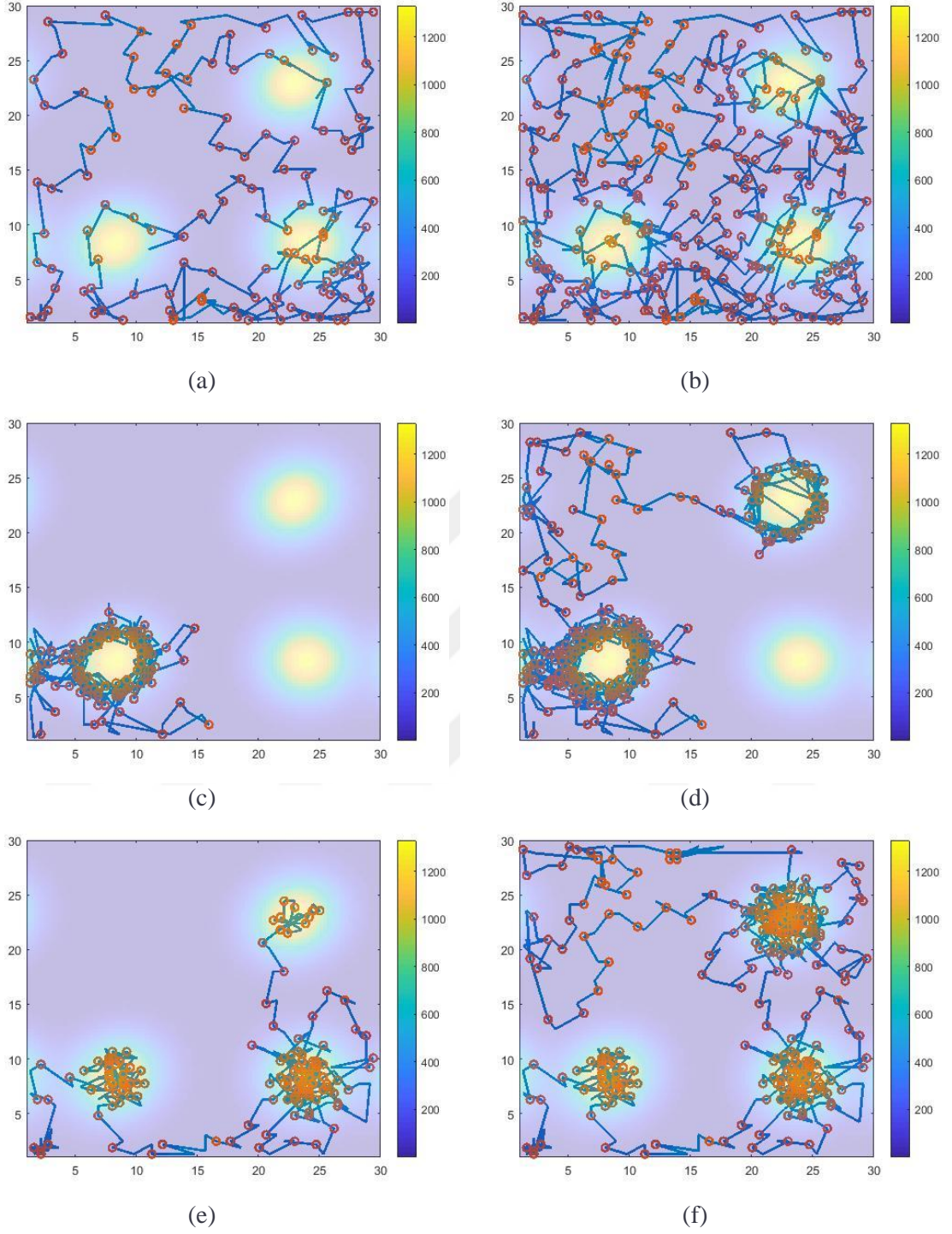


Şekil 3.14. Üç ROI için adaptif örnekleme a) L_{PI} ödül fonksiyonu kullanılarak toplanan 150 örnek b) L_{PI} ödül fonksiyonu kullanılarak toplanan 300 örnek c) $L_{LevelSet}$ ödül fonksiyonu kullanılarak toplanan 150 örnek d) $L_{LevelSet}$ ödül fonksiyonu kullanılarak toplanan 300 örnek e) L_{UCB} ödül fonksiyonu kullanılarak toplanan 150 örnek f) L_{UCB} ödül fonksiyonu kullanılarak toplanan 300 örnek

Şekil 3.15'te, çapraz doğrulama yöntemi kullanılarak 3 ROI alanı içeren bir veri setindeki güzergahlar gösterilmektedir. Şekil 3.8'de başlangıç varyans değerinin 12000'den 6000'e düşürüldüğü açıkça görülmektedir. Bu, veri setinde yapılan optimizasyonların varyansı önemli ölçüde azalttığı ve daha tutarlı bir tahmin elde edildiğini göstermektedir. Şekil 3.14.a'da, entropi ödül fonksiyonunun uygulandığı durumda yalnızca 2 ROI alanına odaklanıldığı görülmektedir. Ancak, Şekil 3.15.a'da, entropi ödül fonksiyonu kullanılarak veri setinin alanının tamamından örneklerin toplandığı gözlemlenmiştir. Bu durum, entropi ödül fonksiyonunun alanın daha geniş bir şekilde keşfedilmesini sağladığını ve varyansın azalmasına katkıda bulunduğunu göstermektedir. Entropi ödül fonksiyonunun veri setinin keşfi açısından daha etkin bir yöntem olduğu bu sonuçtan çıkarılabilir.

Seviye belirleme ödül fonksiyonunun performansı ise Şekil 3.15.c'de gözlemlenmiştir. Burada, seviye belirleme ödül fonksiyonunun yalnızca 1 ROI alanına odaklandığı açıkça görülmektedir. Bu durum, seviye belirleme ödül fonksiyonunun keşiften ziyade sömüreye daha fazla ağırlık verdiğini göstermektedir. Yani, mevcut örneklerin detaylandırılmasına ve optimize edilmesine odaklanmakta, ancak yeni örneklerin ve ilgi alanlarının keşfini ihmal etmektedir. Bu, düşük varyans değerine sahip veri setlerinde ROI alanlarını yeterince keşfetmemesi durumunda belirli riskler taşıdığını kanıtlamaktadır.

UCB ödül fonksiyonuyla oluşturulan güzergahlar ise Şekil 3.15.e ve Şekil 3.15.f'de gösterilmiştir. Bu güzergahlar, Şekil 3.14.e ve Şekil 3.14.f'de izlenen rotalara oldukça benzerdir. Bu benzerlik, UCB ödül fonksiyonunun veri toplama stratejisinde belirli bir tutarlılığı ve tekrarlanabilirliği olduğunu göstermektedir. UCB ödül fonksiyonu ile oluşturulan rotaların, veri setinin dengeli bir şekilde keşfedilmesine katkıda bulunduğu ve varyansın azaltılmasına yardımcı olduğu görülmektedir. Ancak, yine de UCB ödül fonksiyonunun bazen yeterli keşif yapmadan belirli alanlara odaklanma eğiliminde olabileceği dikkate alınmalıdır.



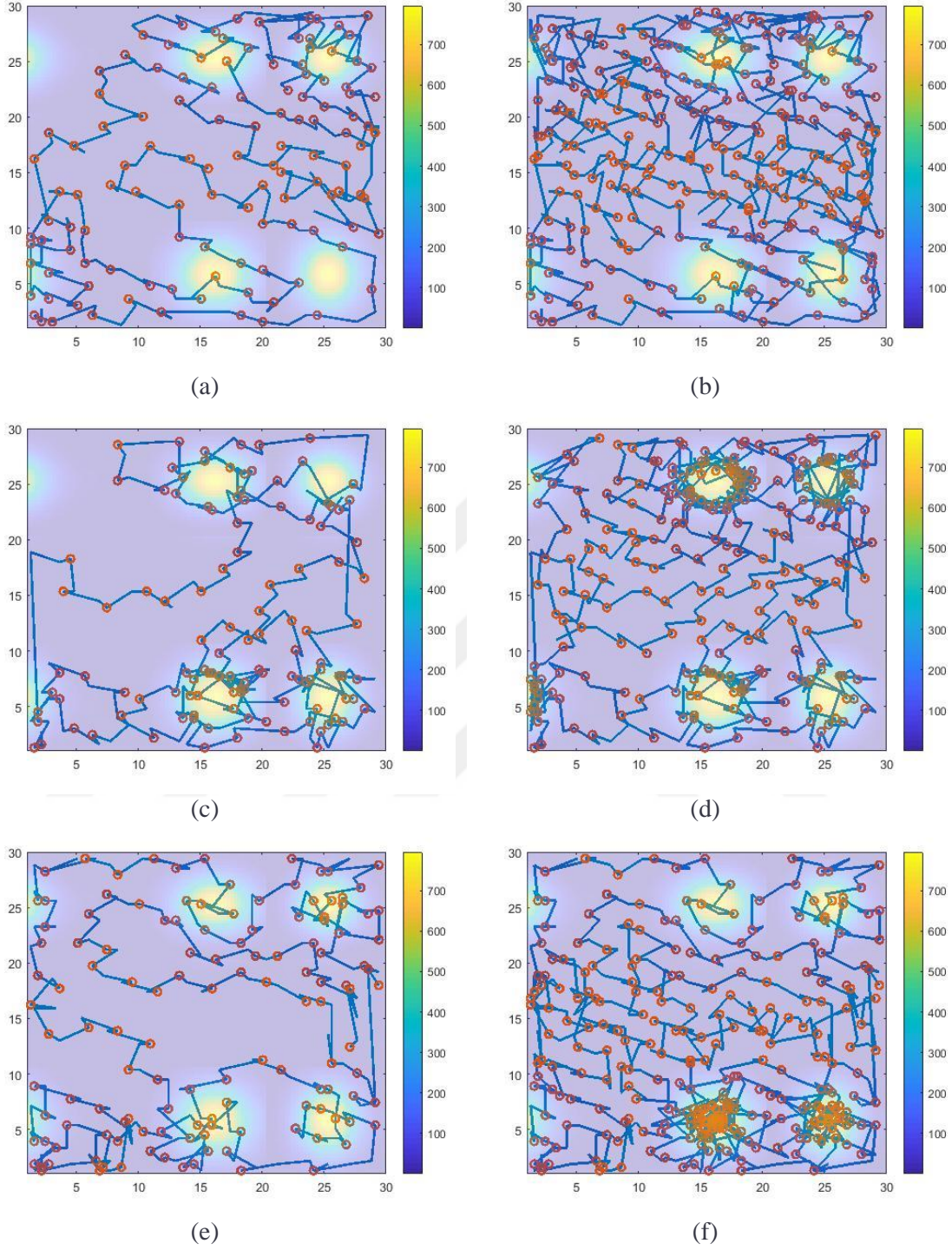
Şekil 3.15. Çapraz doğrulama yöntemi uygulanan üç ROI alanında adaptif örnekleme a) L_{PI} ödül fonksiyonu kullanılarak toplanan 150 örnek b) L_{PI} ödül fonksiyonu kullanılarak toplanan 300 örnek c) a) $L_{LevelSet}$ ödül fonksiyonu kullanılarak toplanan 150 örnek d) $L_{LevelSet}$ ödül fonksiyonu kullanılarak toplanan 300 örnek e) a) L_{UCB} ödül fonksiyonu kullanılarak toplanan 150 örnek f) a) L_{UCB} ödül fonksiyonu kullanılarak toplanan 300 örnek

Şekil 3.16, farklı ödül fonksiyonlarının 4 ROI alanı içeren bir veri seti üzerinde nasıl farklı performanslar gösterdiğine dair önemli gözlemler yapılabilir. Şekil 3.16.a ve Şekil 3.16.b'de entropi ödül fonksiyonu ile oluşturulan güzergahlar incelendiğinde, 150 örnekle sadece 3 ROI alanından veri toplandığı ve 1 ROI alanının tamamen göz ardı edildiği görülmektedir. Ancak, örnek sayısının 300'e çıkarılmasıyla, entropi ödül fonksiyonu tüm ROI alanlarından veri toplayabilmiştir. Bu durum, entropi ödül fonksiyonunun başlangıç aşamasında belirsizlik ve karmaşıklık nedeniyle bazı ROI alanlarını kaçırabileceğini, ancak örnek sayısı arttıkça daha dengeli bir veri toplama gerçekleştirebildiğini göstermektedir.

Şekil 3.14 ve Şekil 3.16'da gösterilen seviye belirleme ödül fonksiyonu ile tüm ROI alanlarından 150 örnekle bile veri toplanabilmesi, bu fonksiyonun karmaşık ve engel yoğunluğu yüksek veri setlerinde üstün performans gösterdiğini ve kapsamlı bir veri toplama kabiliyetine sahip olduğunu doğrular niteliktedir.

UCB ödül fonksiyonu ile, 150 örnekle tüm ROI alanlarından veri toplanabilmiş olması, bu fonksiyonun keşif ve sömürü arasındaki dengeyi başlangıçta iyi kurduğunu göstermektedir. Ancak, 300 örnekle sadece 2 ROI alanına odaklanması ve diğer 2 ROI alanının ihmal edilmesi, UCB fonksiyonunun yerel maksimum noktalarına takılıp genel maksimum noktalarını kaçırabileceğini göstermektedir.

Sonuç olarak, elde edilen veriler doğrultusunda karmaşık veri setleri ve çok sayıda ROI alanı içeren veri setlerinde, seviye belirleme ödül fonksiyonunun daha güvenilir ve mantıklı bir seçim olduğu görülmektedir. Entropi ödül fonksiyonu, veri toplama sürecinde ROI alanlarını kaçırma eğiliminde olup, UCB ödül fonksiyonu ise yerel maksimumlara odaklanarak genel optimum çözümleri göz ardı edebilmektedir. Bu bulgular, ödül fonksiyonu seçiminde, veri setinin özelliklerine ve ihtiyaç duyulan dengeye göre dikkatli bir değerlendirme yapılmasının önemini vurgulamaktadır. Karmaşık ve belirsizlik içeren veri setlerinde seviye belirleme ödül fonksiyonu, daha geniş kapsamlı ve tutarlı örnek toplama yeteneği sayesinde öne çıkarken, belirsizliğin belirli bir seviyeye kadar yönetilebilir olduğu durumlarda entropi ödül fonksiyonu daha az sıkıntı yaratabilir. UCB ödül fonksiyonu ise keşif ve sömürü dengesini korumada başlangıçta iyi performans gösterebilir, ancak uzun vadede yerel optimumlara takılma riski dikkate alınmalıdır. Bu farklı performans dinamikleri, veri toplama stratejilerinin optimize edilmesinde, ödül fonksiyonlarının özelliklerine göre özelleştirilmesi gerektiğini göstermektedir.



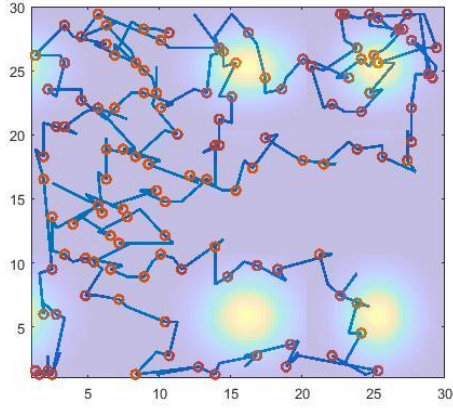
Şekil 3.16. Dört ROI için adaptif örnekleme a) L_{PI} ödül fonksiyonu kullanılarak toplanan 150 örnek b) L_{PI} ödül fonksiyonu kullanılarak toplanan 300 örnek c) $L_{LevelSet}$ ödül fonksiyonu kullanılarak toplanan 150 örnek d) $L_{LevelSet}$ ödül fonksiyonu kullanılarak toplanan 300 örnek e) L_{UCB} ödül fonksiyonu kullanılarak toplanan 150 örnek f) L_{UCB} ödül fonksiyonu kullanılarak toplanan 300 örnek

Şekil 3.9’da dört farklı ROI alanına çapraz doğrulama yöntemi uygulandığında maksimum varyans değerinin 15.000’den 8.000’e düştüğü gözlemlenmiştir. Şekil 3.16.c ve Şekil 3.16.d’de, seviye belirleme ödül fonksiyonunun kullanılmasıyla oluşturulan güzergahlar detaylı bir şekilde sunulmaktadır. Benzer şekilde, Şekil 3.17.c ve Şekil 3.17.d’de düşük varyans değerinde seviye belirleme ödül fonksiyonuyla oluşturulan güzergahlar gösterilmektedir. Bu karşılaştırmalar ışığında, Şekil 3.16.c’de seviye belirleme ödül fonksiyonunun dört farklı ROI alanından da örnekler topladığı görülmektedir. Ancak, düşük varyanslı harita olan Şekil 3.17.c’de ise sadece iki ROI alanından örnekler toplanabilmiştir. Bu durum, yüksek varyanslı haritada daha geniş çaplı veri toplama eğilimini ve düşük varyanslı haritada ise daha sınırlı veri toplama eğilimini ortaya koymaktadır. Şekil 3.16.f’de, UCB ödül fonksiyonu incelendiğinde, bu fonksiyonun iki ROI alanından yoğun bir şekilde örnek toplarken diğer iki ROI alanından daha az miktarda örnek topladığı gözlemlenmiştir.

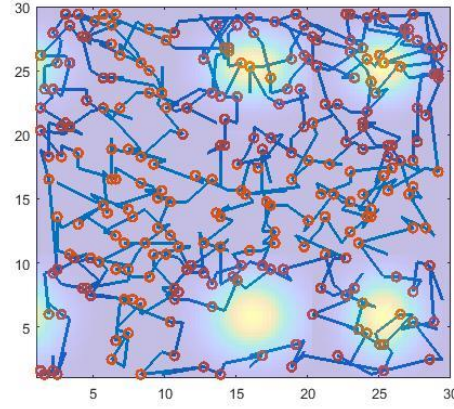
Fakat, Şekil 3.17.f’de UCB ödül fonksiyonunun dört ROI alanından da yoğun miktarda örnek topladığı görülmüştür. Bu durum, UCB fonksiyonunun düşük varyanslı haritada daha dengeli bir veri toplama stratejisi izlediğini işaret etmektedir.

Bu karşılaştırmalar ışığında, yüksek varyanslı haritada seviye belirleme ödül fonksiyonunun keşif ve sömürü dengesinde daha etkin bir performans sergilediği söylenebilir. Diğer yandan, düşük varyanslı haritada ise UCB ödül fonksiyonunun keşif ve sömürü dengesinde daha iyi bir performans gösterdiği gözlemlenmiştir. Bu farklılıklar, her iki ödül fonksiyonunun çeşitli varyans koşulları altında nasıl davrandığını anlamamız açısından oldukça önemlidir. Entropi ödül fonksiyonu incelendiğinde ise, düşük varyanslı haritada alanın tamamından örnek topladığı, yüksek varyanslı haritada ise belirsizliğin çok yüksek olduğu alanlara odaklandığı gözlemlenmiştir.

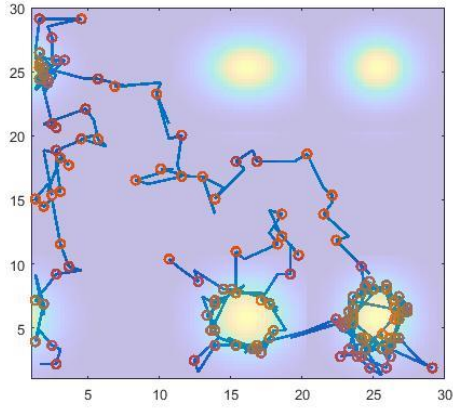
Sonuç olarak, varyans seviyesine bağlı olarak farklı ödül fonksiyonlarının keşif ve sömürü dengelerinde belirgin farklılıklar sergilediği görülmektedir. Yüksek varyanslı haritalarda seviye belirleme ödül fonksiyonunun daha dengeli bir yaklaşım sergileyerek hem keşif hem de sömürüyü optimize ettiği, düşük varyanslı haritalarda ise UCB ödül fonksiyonunun benzer bir dengeyi koruduğu anlaşılmaktadır. Entropi ödül fonksiyonu ise, her iki durumda da belirsizliğin en yüksek olduğu alanlara odaklanarak keşif sürecini yönlendirmektedir.



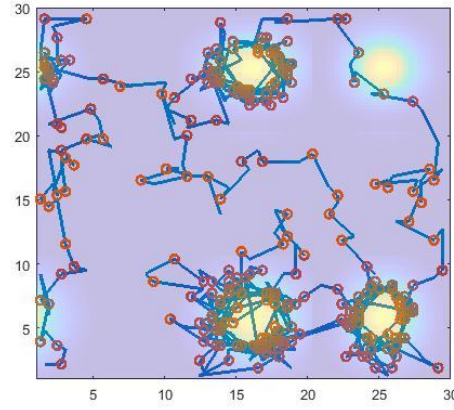
(a)



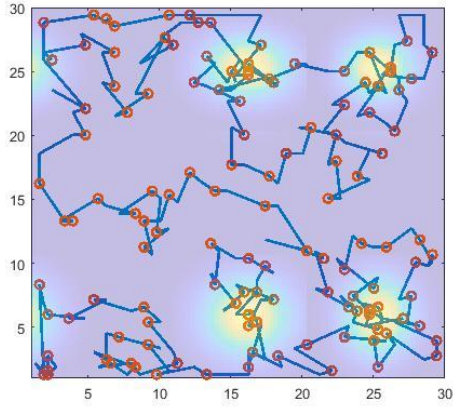
(b)



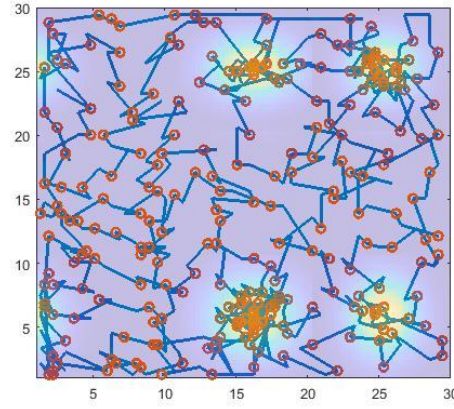
(c)



(d)



(e)



(f)

Şekil 3.17. Çapraz doğrulama yöntemi uygulanan dört ROI alanı için adaptif örnekleme a) L_{PI} ödül fonksiyonu kullanılarak toplanan 150 örnek b) L_{PI} ödül fonksiyonu kullanılarak toplanan 300 örnek c) a) $L_{LevelSet}$ ödül fonksiyonu kullanılarak toplanan 150 örnek d) $L_{LevelSet}$ ödül fonksiyonu kullanılarak toplanan 300 örnek e) a) L_{UCB} ödül fonksiyonu kullanılarak toplanan 150 örnek f) a) L_{UCB} ödül fonksiyonu kullanılarak toplanan 300 örnek

3.2.3. Keşif Verimliliğinde Varyans Etkisi

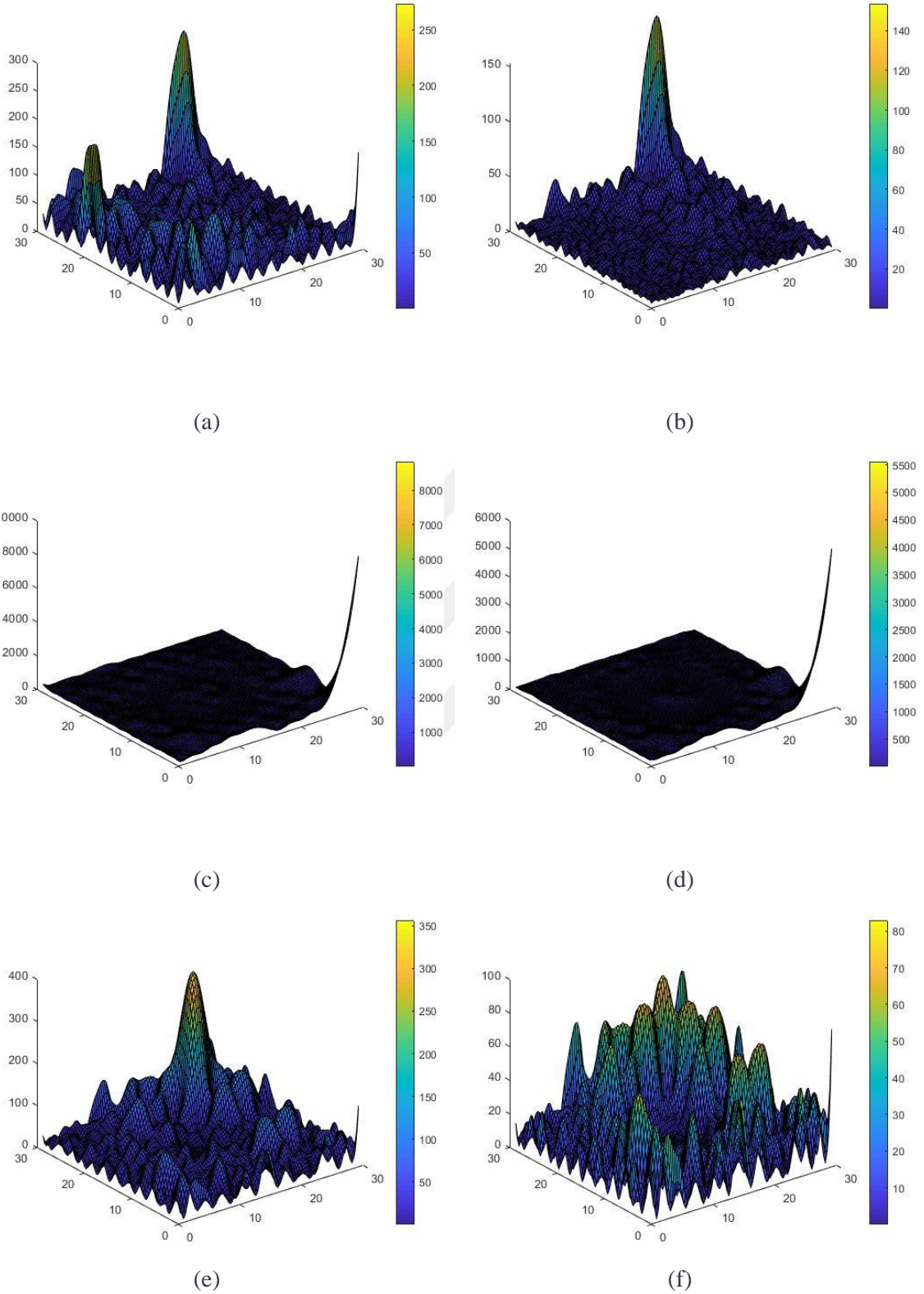
GP, çevre izleme çalışmalarında çokça tercih edilen bir yöntemdir; çünkü sadece tahminler yapma kabiliyeti değil, aynı zamanda bu tahminlerin belirsizliklerini ölçme yeteneği de sağlar. Bu yöntemde çevresel parametreler, gözlem noktalarından elde edilen verilerle tahmin edilir ve belirsizlikler varyans haritalarıyla görselleştirilir. Varyans haritaları, tahminlerdeki belirsizliği nicel olarak ifade eder ve çevre izleme çalışmalarında kritik bir rol oynar.

Bu çalışmada, farklı ödül fonksiyonlarının performansını değerlendirmek amacıyla varyans haritaları kullanılmıştır. Şekil 3.18, üç farklı ödül fonksiyonuyla (L_{entropi} , L_{LS} ve L_{UCB}) yapılan denemelerde, 150 ve 300 örnek toplandıktan sonraki varyans haritalarını göstermektedir. L_{entropi} ödül fonksiyonu ile elde edilen varyans haritaları (Şekil 3.18.a ve Şekil 3.18.b), L_{LS} ödül fonksiyonuyla elde edilen varyans haritaları (Şekil 3.18.c ve Şekil 3.18.d), ve L_{UCB} ödül fonksiyonuyla elde edilen varyans haritaları (Şekil 3.18.e ve Şekil 3.18.f) karşılaştırılmıştır.

Sonuçlarda en yüksek varyans değeri, seviye belirleme ödül fonksiyonuyla oluşturulan varyans haritasında gözlemlenmiştir. Bu, seviye belirleme ödül fonksiyonunun diğer ödül fonksiyonlarına kıyasla daha fazla varyansa neden olduğunu gösterir. Ancak, L_{UCB} ve L_{entropi} ödül fonksiyonları, öğrenme sürecinde alanı seviye belirleme ödül fonksiyonuna göre daha iyi temsil ederek belirsizliği yüksek oranda azaltmışlardır. Bu, L_{UCB} ve L_{entropi} ödül fonksiyonlarının daha etkili bir şekilde çalıştığını ve daha güvenilir sonuçlar verdiğini gösterir.

Buna karşın, entropi ödül fonksiyonuyla kıyaslandığında, UCB ödül fonksiyonunun genel haritasında belirsizlik daha yüksek çıkmıştır. UCB, keşif ve sömürü arasında denge kurmada bazı zorluklarla karşılaşabilir ve bu da daha fazla belirsizliğe yol açabilir.

Entropi ödül fonksiyonu ise, genel haritada belirsizliği diğer ödül fonksiyonlarına kıyasla daha düşük tutmayı başarmıştır. Her ne kadar entropi ödül fonksiyonu belirli bir bölgede belirsizliği artırsa da, genel belirsizlik seviyesini düşük tutarak daha homojen ve güvenilir bir öğrenme sağlamıştır. Bu, entropi ödül fonksiyonunun keşif sürecinde daha etkin olduğunu ve sonuçların genel belirsizliğini düşürmede başarılı olduğunu göstermektedir.



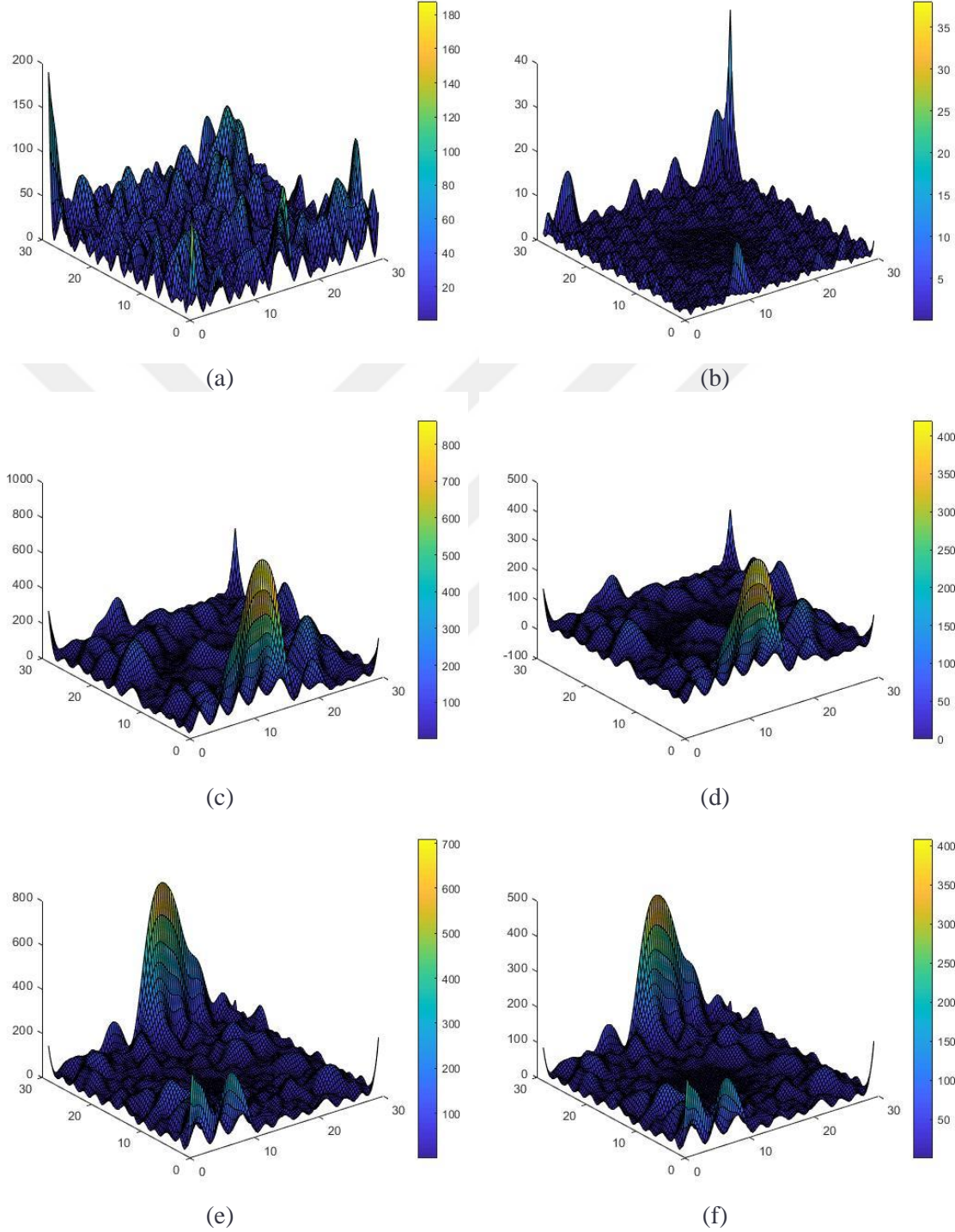
Şekil 3.18. Tek ROI için Varyans Haritası: a) $L_{entropy}$ ödül fonksiyonu kullanılarak toplanan 150 örnek b) $L_{entropy}$ ödül fonksiyonu kullanılarak toplanan 300 örnek c) L_{LS} ödül fonksiyonu kullanılarak toplanan 150 örnek d) $L_{LevelSet}$ ödül fonksiyonu kullanılarak toplanan 300 örnek e) L_{UCB} ödül fonksiyonu kullanılarak toplanan 150 örnek f) L_{UCB} ödül fonksiyonu kullanılarak toplanan 300 örnek

Şekil 3.19'da, çapraz entropi yöntemi uygulanmış tek ROI alanı bulunan haritada varyans değişimi görselleştirilmiştir. Başlangıç varyans haritası, Şekil 3.6'da belirtildiği üzere $2,5 \times 10^4$ seviyesinde bulunmaktadır ve süreç içerisinde bu varyans değeri 15000 seviyesine kadar azalmıştır. Bu azalma göz önünde bulundurularak, entropi ödül fonksiyonu kullanılarak varyans değerindeki düşüş maksimum seviyeye ulaştırılmış ve 300 örnek toplandıktan sonra alandaki belirsizlik, Şekil 3.19.b'de görülebileceği üzere, 40'a kadar düşürülmüştür. Öte yandan, seviye belirleme ödül fonksiyonu kullanıldığında belirsizlik değeri, başlangıçta yaklaşık olarak 10000 seviyesindeyken, 150 örnek ile bu değer yaklaşık onda birine indirildiği Şekil 3.19.c'de görülebilmektedir. Ayrıca, Şekil 3.18.d ve Şekil 3.18.f'de UCB ödül fonksiyonu ile seviye belirleme ödül fonksiyonunun varyans haritaları arasında belirgin bir farklılık görünmesine rağmen, başlangıç varyans değeri azaltıldığında bu varyans haritalarının Şekil 3.19.d ve Şekil 3.19.f'de benzer sonuçlar verdiği gözlemlenmiştir. Ancak UCB ödül fonksiyonu, yüksek varyans değerindeki başlangıç koşulunda daha etkili bir keşif yaparak, 150 örnek toplandıktan sonra bile maksimum varyansı Şekil 3.18.e'de görüldüğü gibi 400'e kadar düşürebilmiştir. Buna karşın, Şekil 3.18.f'de görüldüğü üzere, 300 örnek toplandığında bile maksimum varyans değeri yaklaşık 500'e yaklaşmıştır. Tüm bu veriler, kullanılan yöntem ve fonksiyonların varyans azaltmadaki etkinliğini, başlangıçtaki yüksek varyans koşullarında dahi etkili bir keşif stratejisi uygulayarak belirsizliğin nasıl en aza indirilebileceğini göstermektedir. Özellikle, entropi ödül fonksiyonunun optimize edilmesiyle elde edilen sonuçlar, örnekleme sürecinin ne kadar etkili yönetilebileceğini ve belirsizliğin nasıl azaltılabileceğini ortaya koymaktadır. Bu değerlendirmede dikkat edilmesi gereken kritik bir diğer husus, seviye belirleme ve UCB ödül fonksiyonlarının karşılaştırılmasıdır. Başlangıç varyans değerleri düşürüldüğünde, Şekil 3.19.d ve Şekil 3.19.f'de her iki yöntemin varyans haritalarının benzer sonuçlar verdiği gözlemlenmiştir.

Ancak, UCB ödül fonksiyonunun üstün olduğu nokta, yüksek varyans başlangıç koşullarında daha iyi bir performans sergilemesi ve varyansı daha hızlı azaltabilmesidir. Şekil 3.18'deki karşılaştırmalar bu durumu açıkça ortaya koymuştur. Ek olarak, Şekil 3.18.d ve Şekil 3.18.f'de görüldüğü gibi, UCB ödül fonksiyonu, entropi ödül fonksiyonuna kıyasla yüksek varyans koşullarında daha etkin çalışmış ve 150 örnekte bile maksimum varyansı 400'e kadar düşürebilmiştir. Bu durum, UCB'nin daha az örnek ile daha etkili varyans azaltma yeteneğine işaret ederken, entropi ödül fonksiyonu ile karşılaştırıldığında benzer sayıda örnek ile daha yüksek varyans değerlerinde kalınmıştır.

Bu analizler ve sonuçlar, varyans azaltımında kullanılan yöntemlerin etkinliğini kıyaslarken, başlangıç koşulları ve toplama sayılarının önemini vurgulamaktadır. Özellikle, yöntem seçiminde yüksek varyans başlangıç koşullarında UCB'nin daha uygun olduğu, ancak düşük varyans başlangıç koşullarında entropi ödül fonksiyonunun tercih edilebileceği sonucuna varılmaktadır. Ayrıca, belirsizliğin minimize edilmesi ve keşif süreçlerinin optimizasyonu için farklı ödül

fonksiyonlarının nasıl kullanılacağını anlamak, gelecekteki çalışmalarda ve gerçek dünya uygulamalarında önemli bir avantaj sağlayacaktır.



Şekil 3.19. Çapraz doğrulama yöntemi uygulanan tek ROI alanında varyans haritası: a) L_{entropi} ödül fonksiyonu kullanılarak toplanan 150 örnek b) L_{entropi} ödül fonksiyonu kullanılarak toplanan 300 örnek c) L_{LS} ödül fonksiyonu kullanılarak toplanan 150 örnek d) L_{LS} ödül fonksiyonu kullanılarak toplanan 300 örnek e) L_{UCB} ödül fonksiyonu kullanılarak toplanan 150 örnek f) L_{UCB} ödül fonksiyonu kullanılarak toplanan 300 örnek

Şekil 3.20, iki ROI'li bir veri kümesi için varyans haritalarını göstermektedir. Şekil 3.18'dekine benzeyen bulgulara ek olarak, L_{LS} ödül fonksiyonu kullanıldığında en yüksek varyans değerlerinin Şekil 3.20.c ve Şekil 3.20.d'de gözlemlendiğini belirtmektedir. Bu gözlem, $L_{LevelSet}$ ödül fonksiyonunun yalnızca belirli bölgelere odaklanarak, tüm alanı yeterince öğrenemediğini bir kez daha doğrulamaktadır.

Bunun yanında, Şekil 3.20.a ve Şekil 3.20.c'de $L_{entropi}$ ve L_{UCB} ödül fonksiyonları kullanılarak elde edilen varyans haritalarına bakıldığında, ilginç bir farklılık gözlenir. L_{UCB} ödül fonksiyonunun varyans haritası, $L_{entropi}$ ödül fonksiyonu ile elde edilen varyans haritasına göre daha düşük varyans değerleri göstermektedir. Bu durum, L_{UCB} ödül fonksiyonunun da belirli bölgelerde etkili olduğunu, ancak diğer alanları ihmal etme riskini taşıdığı anlamına gelir.

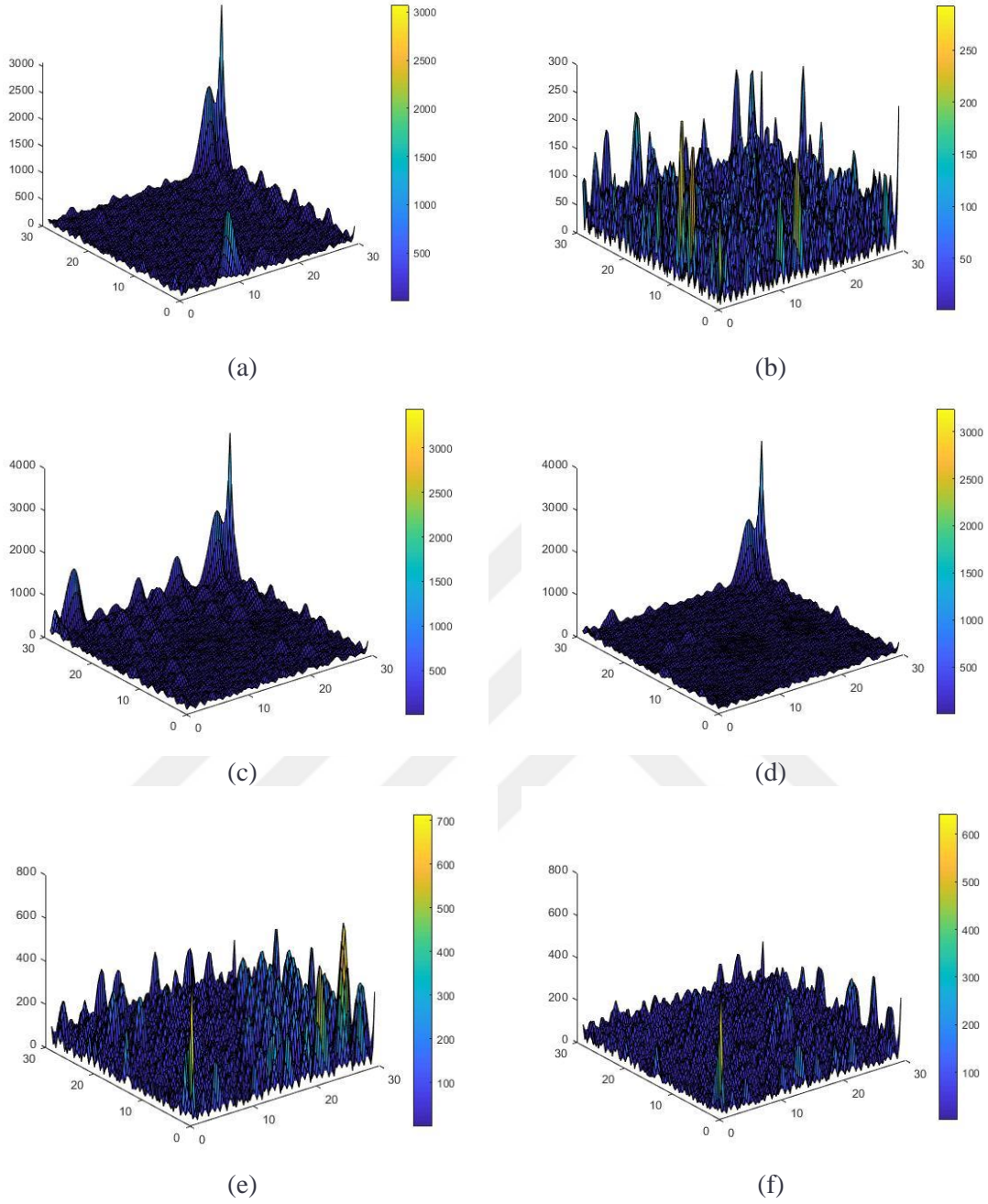
Özellikle Şekil 3.20.b'de, 300 örnek toplandıktan sonra öğrenmenin daha hızlı gerçekleştiği açıkça görülmektedir. Bu durum, örnek sayısının artmasıyla birlikte $L_{entropi}$ ödül fonksiyonunun tüm alanı daha iyi öğrenme yeteneğini desteklemektedir. $L_{entropi}$ ödül fonksiyonu, çevresi daha geniş bir şekilde tarayarak ve daha dengeli bir öğrenme süreci sağlayarak diğer ödül fonksiyonlarına göre daha etkili performans göstermektedir.

L_{LS} ödül fonksiyonu, L_{UCB} 'ye benzer bir stratejiyle ilk ROI alanına odaklanıp daha sonra onu tespit eder; bu da diğer ROI alanlarını göz ardı etmesine neden olabilir. Bu stratejinin bir sonucu olarak, 150 örnek ile elde edilen sonuçlar iyi olmasına rağmen, tüm alan $L_{entropi}$ ödül fonksiyonu kullanılarak daha etkin bir şekilde öğrenilmektedir. $L_{entropi}$ ödül fonksiyonu, düşük varyans değerleri ve daha geniş bir alanda daha kesin tahminler yapma yeteneği ile ön plana çıkmaktadır.

Özetle, bu çalışma, iki ROI'li bir veri kümesi üzerindeki ödül fonksiyonlarının performansını varyans haritaları aracılığıyla karşılaştırmaktadır. L_{LS} ödül fonksiyonu, belirli bölgelere odaklanarak genelde en yüksek varyans değerlerine yol açar ve tüm alanı yeterince kapsayamaz; bu durum Şekil 3.20.c ve Şekil 3.20.d'de bariz şekilde gözlemlenebilir.

Öte yandan, L_{UCB} ödül fonksiyonu, varyans değerleri açısından $L_{entropi}$ 'ye göre çekişmeli bir performans sergilese de, L_{UCB} ilk ROI'ye odaklandığı ve onu tespit ettikten sonra diğer alanları ihmal edebildiği için Şekil 3.18.f'deki durum ortaya çıkar. Bu, özellikle ilk aşamada öğrenme sürecini etkiler ve $L_{entropi}$ ödül fonksiyonuna kıyasla daha düşük bir performansla neden olabilir.

Bu bağlamda, varyans haritaları ve örnek sayıları dikkate alındığında, $L_{entropi}$ ödül fonksiyonunun çevresel izleme çalışmalarında daha güvenilir ve etkin sonuçlar sunduğu söylenebilir. Tüm bulgular, çevre izlemede kullanılan ödül fonksiyonlarının seçiminde varyans haritalarının ve örnekleme stratejilerinin dikkate alınmasının önemini bir kez daha vurgulamaktadır. Yani, $L_{entropi}$ ödül fonksiyonu, daha düşük varyans ve daha geniş kapsama alanıyla çevresel izleme uygulamaları için daha uygun bir seçimdir.



Şekil 3.20. İki ROI için Varyans Haritası: a) $L_{entropi}$ ödül fonksiyonu kullanılarak toplanan 150 örnek b) $L_{entropi}$ ödül fonksiyonu kullanılarak toplanan 300 örnek c) L_{LS} ödül fonksiyonu kullanılarak toplanan 150 örnek d) L_{LS} ödül fonksiyonu kullanılarak toplanan 300 örnek e) L_{UCB} ödül fonksiyonu kullanılarak toplanan 150 örnek f) L_{UCB} ödül fonksiyonu kullanılarak toplanan 300 örnek

Şekil 3.7'deki başlangıç varyans haritasında iki ROI alanı bulunan veri setinde, başlangıç maksimum varyans değeri çapraz doğrulama yöntemi ile 5000'den 3000'e düşürülmüştür. Bu önemli bir azalmadır, ancak bu varyans düşüşü bile, entropi ödül fonksiyonu kullanılarak 150 örnekle düşük varyanslı haritada maksimum varyansın büyük ölçüde azaltılmasıyla karşılaştırıldığında daha etkili bir şekilde varyans yönetimi sağlanmıştır. Şekil 3.21.a'da maksimum

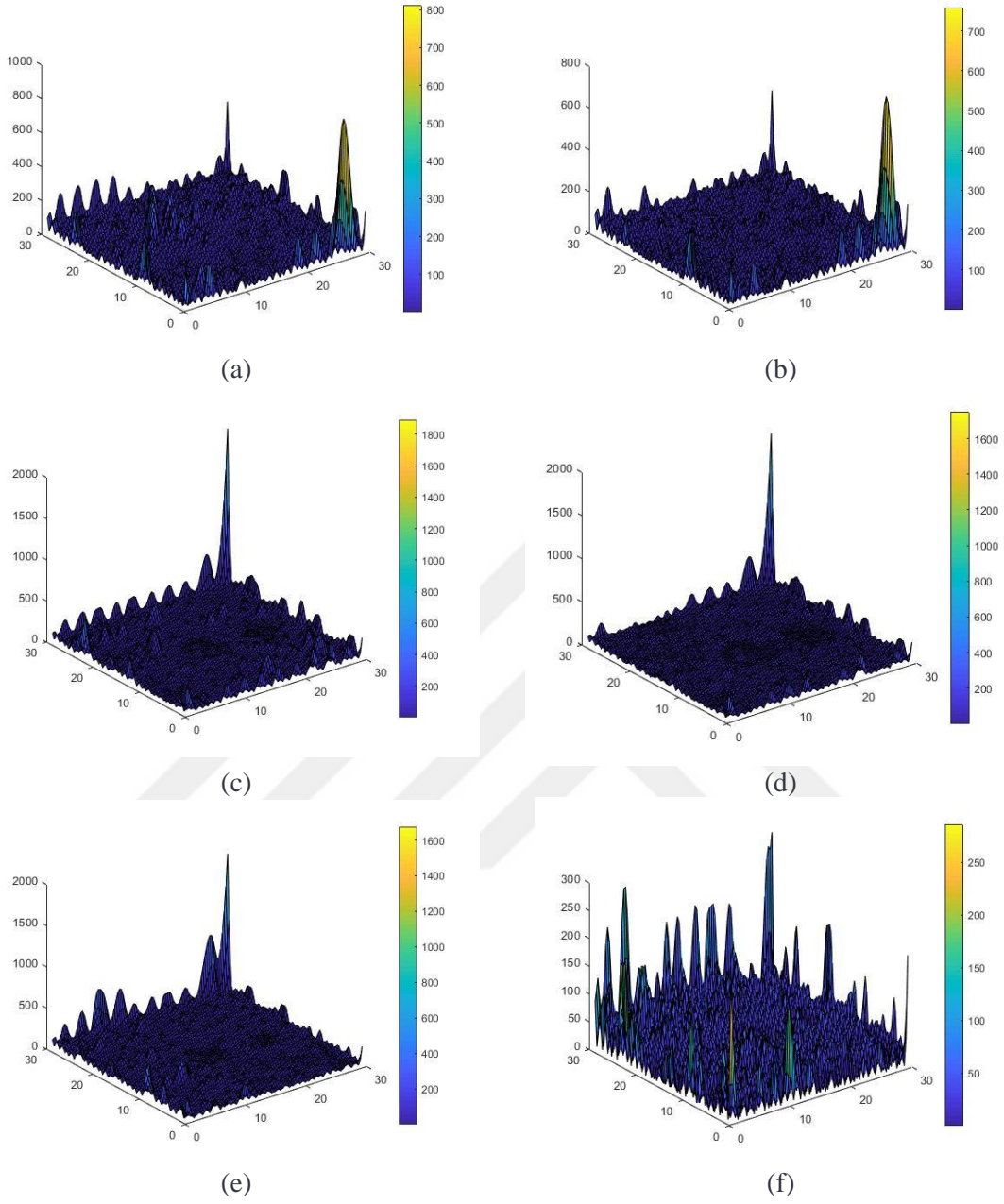
varyans değeri 800 iken, Şekil 3.20.a'da bu değer 3000'dir. Bu durum, başlangıç varyans değerinin yarı yarıya düşmemesine rağmen entropi ödül fonksiyonunun optimum örnekleme stratejileri ile oldukça etkili olduğunu göstermektedir. Ancak, 300 örnek toplandığında (Şekil 3.21.b), maksimum varyans değeri Şekil 3.20.a'dakinden daha yüksek çıkmıştır. Bu sonuç, örnek sayısının artması ile birlikte belirli bölgelerdeki belirsizliğin azalmasının yanı sıra, daha geniş bir alanda belirsizliğin yayılmasının sonuçlarından biri olabilir. Bu strateji, ROI dışındaki bölgelere de daha fazla keşif yapılmasına neden olabilir. Benzer durum, seviye belirleme ödül fonksiyonu için de geçerlidir. Şekil 3.20.c'de maksimum varyans 4000 civarındayken, Şekil 3.21.c'de bu değer 2000'e düşmüştür. Ancak, 300 örnek toplandığında, genel belirsizlik düşmesine rağmen maksimum varyans değeri daha az oranda düşmüştür. Bu durum, seviye belirleme ödül fonksiyonunun belirsizliğin genel paylaştırılmasına odaklanması ve böylelikle bazı kritik bölgelerdeki maksimum varyans değerinin yeterince düşürülememesine bağlanabilir. Bu fonksiyonun örnekleme stratejisi, maksimum varyansın verimli bir şekilde düşürülmesini zorlaştırabilir, çünkü belirsizliğin yayılması yerine belirli noktalarda yoğunlaşma olabilir.

UCB ödül fonksiyonu ise farklı bir yaklaşımla çalışır ve düşük belirsizlikli haritalarda ROI alanlarına odaklanarak, daha fazla keşif yapmaktan ziyade belirli bölgelerde detaylı örnekleme yapar.

Şekil 3.20.f ile karşılaştırıldığında, UCB ödül fonksiyonu kullanılarak belirsizliğin daha yüksek çıkmasının nedeni, özellikle ROI dışındaki bölgelerde daha az keşif yapılması olabilir. Bu strateji, ROI'lerde örnekleme derinlemesine yapılmasına olanak tanısa da, diğer alanlardaki belirsizlikleri yeterince keşfedememe riski taşır. Bu nedenle, belirsizliklerin yüksek kaldığı bölgelerde maksimum varyansın düşürülmesi daha zayıf olur.

Sonuç olarak, entropi ve seviye belirleme ödül fonksiyonları genel belirsizliği azaltmada etkili olabilirken, maksimum varyansı düşürme performansları farklılık göstermektedir. Entropi fonksiyonu, örnekleme sayısı arttıkça belirsizlik yayılmasına yol açabilir ve bu da maksimum varyansın yüksek kalmasına neden olabilir. Seviye belirleme fonksiyonu ise genel belirsizlik düşüşü sağlasa da bazı bölgelerdeki maksimum varyans yeterince düşürülemeyebilir. UCB fonksiyonu da farklı bir strateji izleyerek ROI alanlarına yoğunlaşırken, belirsizlik dengesi ve keşif arasında zayıflıklar gösterebilir.

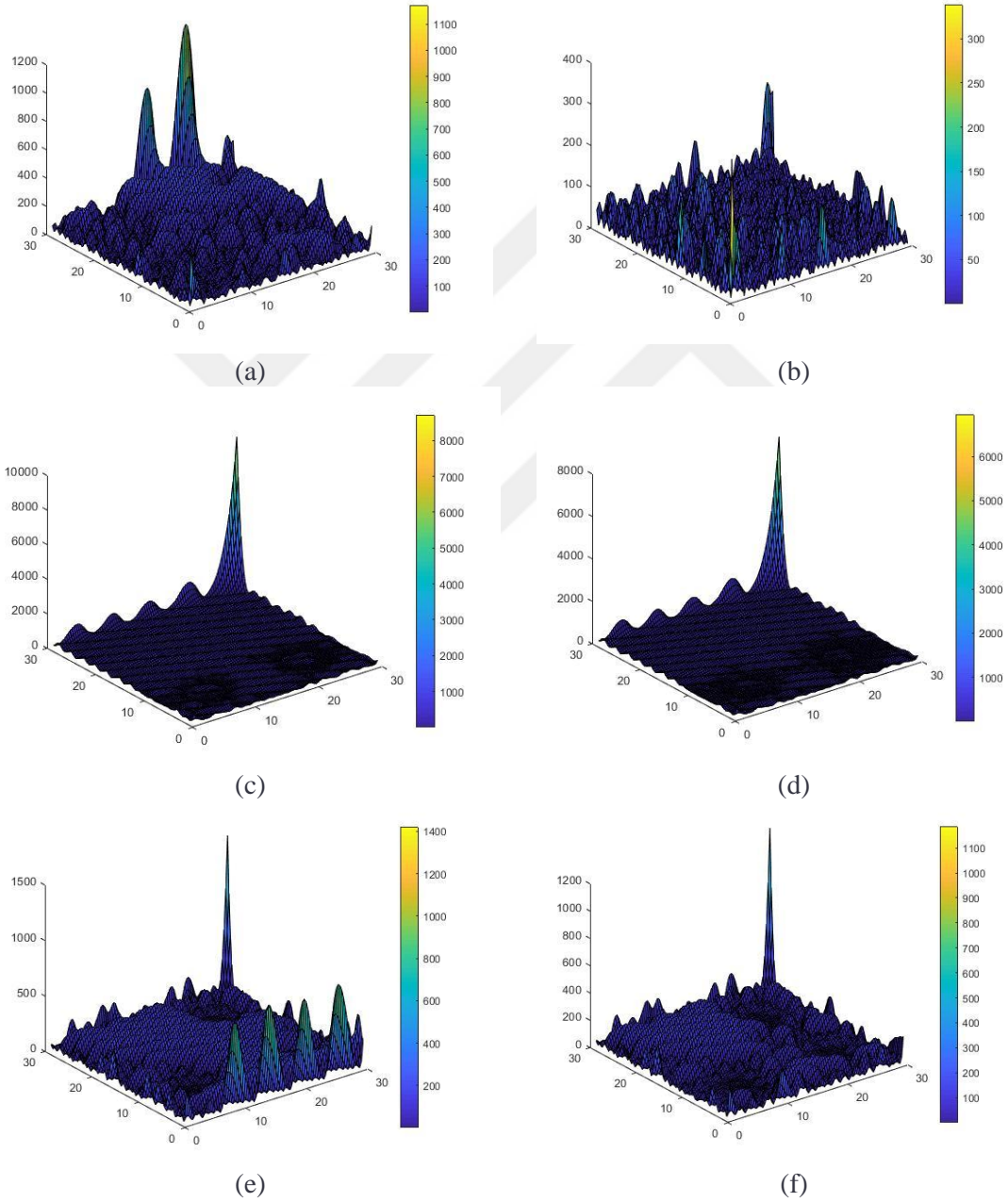
Bu veriler ışığında, hangi ödül fonksiyonunun kullanılacağına karar verirken varyans kontrolü, belirsizlik dağılımı ve keşif stratejileri göz önünde bulundurulmalıdır. Hedeflenen belirsizlik azaltma amacına göre ödül fonksiyonları ve örnekleme stratejileri arasında optimal denge sağlanmalıdır. Örneğin, belirli bölgelerdeki detayların derinlemesine incelenmesi gerekiyorsa UCB fonksiyonu tercih edilebilirken, geniş alanlardaki belirsizlikleri daha etkili şekilde yönetmek için entropi veya seviye belirleme fonksiyonları daha uygun olabilir.



Şekil 3.21. Çapraz doğrulama yöntemi uygulananın iki ROI alanı için varyans Haritası: a) $L_{entropi}$ ödül fonksiyonu kullanılarak toplanan 150 örnek b) $L_{entropi}$ ödül fonksiyonu kullanılarak toplanan 300 örnek c) L_{LS} ödül fonksiyonu kullanılarak toplanan 150 örnek d) L_{LS} ödül fonksiyonu kullanılarak toplanan 300 örnek e) L_{UCB} ödül fonksiyonu kullanılarak toplanan 150 örnek f) L_{UCB} ödül fonksiyonu kullanılarak toplanan 300 örnek

Şekil 3.22'de, 3 ROI alanı içeren bir veri seti için varyans haritası gösterilmektedir. Bu şekilde, 150 örnek kullanıldığında bile en düşük varyans haritasının entropi ödül fonksiyonu ile elde edildiği görülmektedir. Bu, entropi ödül fonksiyonunun belirsizliği en aza indirmede etkili olduğunu göstermektedir. UCB ödül fonksiyonu ile de benzer bir şekilde, 150 örnek kullanıldığında belirsizlik önemli ölçüde azaltılmıştır. Ancak, 150 örnekten sonra UCB ödül fonksiyonu ROI alanlarına odaklandığı için belirsizlikte büyük bir düşüş gözlemlenmemiştir. Bu durum, UCB'nin

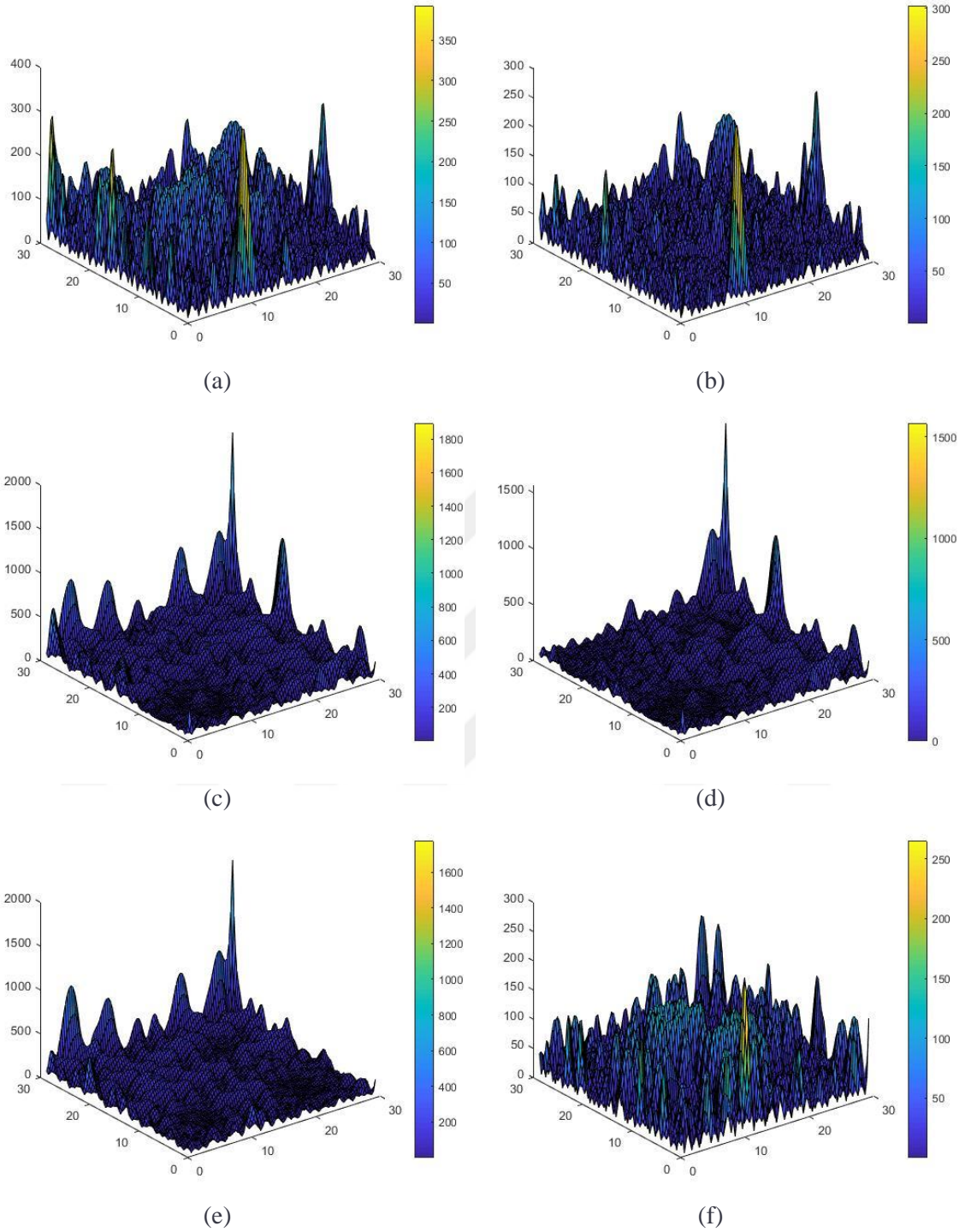
başlangıçta belirsizliği azaltmada etkili olmasına rağmen, daha sonra belirli alanlara odaklandığında bu etkinliğin azaldığını göstermektedir. Öte yandan, seviye belirleme ödül fonksiyonu, belirsizliği diğer ödül fonksiyonlarına kıyasla çok fazla düşürememiştir. Bu, seviye belirleme ödül fonksiyonunun varyans haritasında daha yüksek belirsizlik seviyeleriyle sonuçlandığını ve dolayısıyla veri setindeki belirsizliği etkili bir şekilde azaltmada yetersiz kaldığını göstermektedir.



Şekil 3.22. Üç ROI için varyans haritası: a) $L_{entropi}$ ödül fonksiyonu kullanılarak toplanan 150 örnek b) $L_{entropi}$ ödül fonksiyonu kullanılarak toplanan 300 örnek c) L_{LS} ödül fonksiyonu kullanılarak toplanan 150 örnek d) L_{LS} ödül fonksiyonu kullanılarak toplanan 300 örnek e) L_{UCB} ödül fonksiyonu kullanılarak toplanan 150 örnek f) L_{UCB} ödül fonksiyonu kullanılarak toplanan 300 örnek

Şekil 3.23'teki çapraz doğrulama yöntemi uygulanan ve 3 ROI alanı bulunan veri seti için hazırlanan varyans haritalarında gözlemlenen sonuçlar şu şekildedir: Çapraz doğrulama yöntemiyle en yüksek varyans değerinin başlangıçta yaklaşık 12.000 iken 6.000'e düştüğü, yani iki kat azaldığı görülmektedir. Bu azalma sürecinde, entropi ödül fonksiyonunda 150 örnekte maksimum varyansın önemli ölçüde düştüğü ve genel belirsizlik seviyesinin de büyük oranda azalmasına rağmen, 300 örnekte bu maksimum varyans değerinde belirgin bir değişiklik olmadığını belirtmek önemlidir. Aynı şekilde, seviye belirleme ödül fonksiyonu da 150 örnek kullanıldığında belirsizliği ciddi oranda düşürmesine karşın, 300 örnekte dikkat yüksek bir ilave düşüş gözlemlenmemiştir.

Bununla birlikte, UCB ödül fonksiyonu 2-ROI alanı bulunduran veri setiyle karşılaştırıldığında, belirsizliği daha büyük oranda azaltarak maksimum varyans değerini 300 örnekte entropi ödül fonksiyonuna yaklaştırmış, fakat bu durumda genel belirsizlik seviyesi entropi ödül fonksiyonununkinden daha yüksek kalmıştır. Ayrıca, 150 örnekte Şekil 3.22.e'deki belirsizlik, Şekil 3.23.e'ye kıyasla daha düşük çıkmıştır. Bu sonuç, UCB ödül fonksiyonunun diğer veri setlerinde olduğu gibi düşük belirsizlik seviyelerinde sömürüye daha fazla odaklanıp, keşfetme sürecini ikinci plana attığını doğrulamaktadır. Bu stratejik yaklaşım, Çapraz Doğrulama Yöntemi ve ödül fonksiyonları arasındaki farkların ve stratejik yönelimlerin veri seti davranışları üzerindeki etkilerini anlamak için önemli bir bulgudur.



Şekil 3.23. Çapraz doğrulama yöntemi uygulanan üç ROI alanı için varyans haritası: a) $L_{entropi}$ ödül fonksiyonu kullanılarak toplanan 150 örnek b) $L_{entropi}$ ödül fonksiyonu kullanılarak toplanan 300 örnek c) L_{LS} ödül fonksiyonu kullanılarak toplanan 150 örnek d) L_{LS} ödül fonksiyonu kullanılarak toplanan 300 örnek e) L_{UCB} ödül fonksiyonu kullanılarak toplanan 150 örnek f) L_{UCB} ödül fonksiyonu kullanılarak toplanan 300 örnek

Şekil 3.24'de, 4 ROI alanı içeren bir veri seti için varyans haritası verilmiştir. Bu sonuçlar, ROI sayısı arttıkça model karmaşıklığının da etkisiyle varyans haritalarında bir artış olduğunu

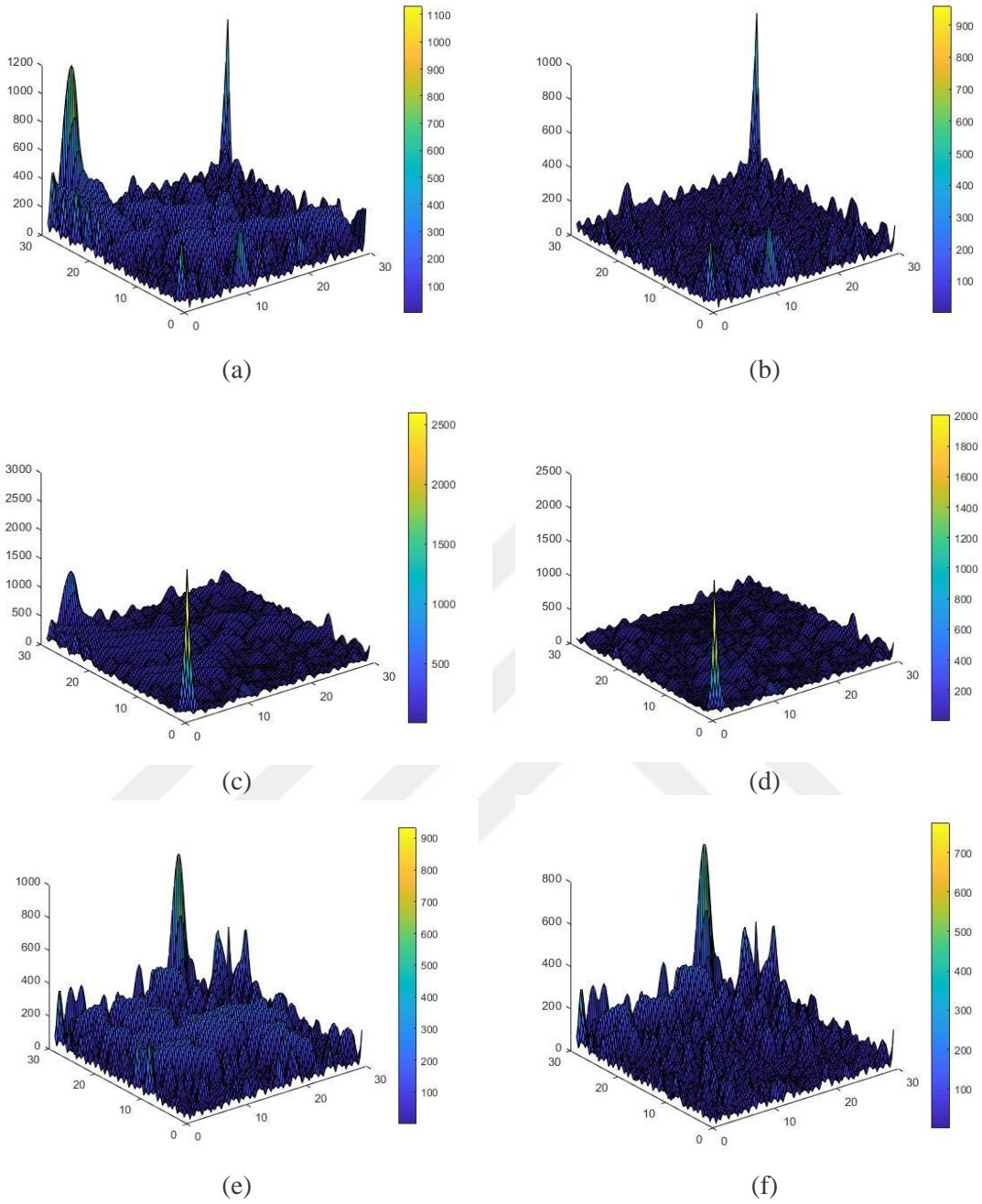
göstermektedir. Model karmaşıklığı artmasına rağmen, seviye belirleme ödül fonksiyonu bu veri setinde en iyi belirsizlik düşüşünü sağlamıştır. Bu durum, model karmaşıklığı arttıkça seviye belirleme ödül fonksiyonunun daha etkili olduğunu ortaya koymaktadır.

UCB ödül fonksiyonu da, seviye belirleme ödül fonksiyonuna benzer şekilde, ortamın belirsizliğini entropi ödül fonksiyonuna kıyasla daha fazla azaltmıştır. Bu, UCB'nin keşif ve sömürü arasındaki dengeyi iyi kurduğunu ve belirsizliği etkili bir şekilde azaltabildiğini göstermektedir. Yüksek belirsizliği olan veri setlerinde, entropi ödül fonksiyonu alanı öğrenmeye çalışırken ROI alanlarındaki örnek miktarını azaltmıştır. Bu durum, genel maksimum noktalarının kaçırılmasına neden olmuş ve alandaki belirsizliğin artmasına yol açmıştır. Elde edilen sonuçlar, entropi ödül fonksiyonunun ROI alanı az olan veri setlerinde daha iyi performans gösterdiğini, ancak karmaşık modellerde performansının düştüğünü göstermektedir.

Seviye belirleme ödül fonksiyonu ise, entropi ödül fonksiyonunun aksine, karmaşık olmayan modellerde ROI alanlarına odaklandığı için alanın belirsizliğini düşürmekte başarısız kalmış ve yeterince alanı öğrenememiştir. Bu, seviye belirleme ödül fonksiyonunun basit modellerde etkili olmadığını, ancak karmaşıklık arttıkça daha iyi performans gösterdiğini ortaya koymaktadır.

UCB ödül fonksiyonu, 1 ve 2 ROI alanı olan veri setlerinde başlangıçta alanı öğrenip daha sonra ROI alanlarına odaklanmıştır. Ancak, yerel maksimum noktalarına takılan UCB ödül fonksiyonu, ROI alanlarını keşfettikten sonra belirsizliği azaltmada performansını kaybetmiştir. Bu, UCB'nin başlangıçta güçlü bir keşif yeteneğine sahip olduğunu, ancak belirli alanlara odaklandığında etkinliğini yitirdiğini göstermektedir.

Şekil 3.24'deki sonuçlar, seviye belirleme ödül fonksiyonu, karmaşıklık arttıkça daha iyi performans gösterirken, entropi ödül fonksiyonu basit modellerde daha etkili olmuştur. UCB ödül fonksiyonu ise başlangıçta güçlü bir keşif yapabilse de, belirli bölgelere odaklandığında performansını kaybetmektedir. Bu bulgular, keşif ve sömürü arasındaki dengenin önemini ve ödül fonksiyonlarının farklı veri setleri ve model karmaşıklıkları karşısında nasıl performans gösterdiğini vurgulamaktadır.

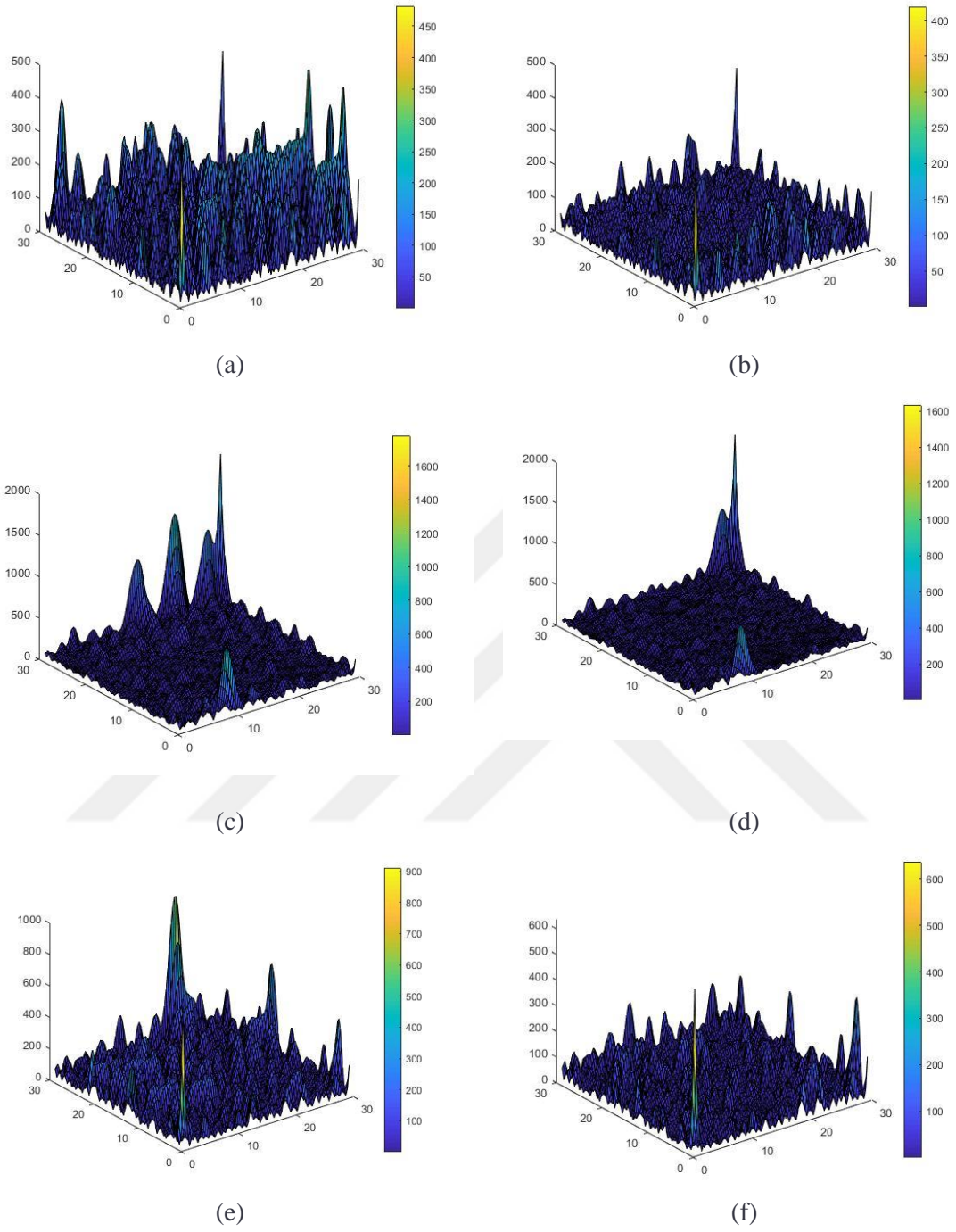


Şekil 3.24. Dört ROI için varyans haritası: a) L_{entropi} ödül fonksiyonu kullanılarak toplanan 150 örnek b) L_{entropi} ödül fonksiyonu kullanılarak toplanan 300 örnek c) L_{LS} ödül fonksiyonu kullanılarak toplanan 150 örnek d) L_{LS} ödül fonksiyonu kullanılarak toplanan 300 örnek e) L_{UCB} ödül fonksiyonu kullanılarak toplanan 150 örnek f) L_{UCB} ödül fonksiyonu kullanılarak toplanan 300 örnek

Şekil 3.25'te çapraz doğrulama yöntemi uygulanan ve 4 ROI bulunan veri seti için varyans haritalarındaki değişim ayrıntılı şekilde incelenmiştir. Çapraz doğrulama yöntemi kullanıldığında, en yüksek varyans başlangıçta yaklaşık olarak 15.000 iken 8.000'e düşürülmüştür. Entropi ödül fonksiyonu, 150 örnek ile bu değeri 500'e kadar azaltmayı başarmıştır. Ancak, daha önceki veri setlerinde gözlemlendiği gibi, 300 örnek toplandığında entropi ödül fonksiyonunun ana odağı genel

belirsizliđi dűşűrmek olduđundan, maksimum varyans deđerinde dikkate deđer bir deđiřiklik meydana gelmemiřtir. Seviye belirleme dűl fonksiyonu ise varyansın dűřűrűlmesi aısından daha karmařık bir davranıř sergilemiřtir. Őekil 3.24'teki gzlemlerle karřılařtırıldıđında, 4 ROI bulunan bu veri setinde belirsizliđi daha fazla azaltmıř olmasına rađmen, 2 ve 3 ROI bulunan veri setlerinde gzlemlenen bűyűk dűřűřleri tekrarlayamamıřtır. Bu sonu, seviye belirleme dűl fonksiyonunun keřiftan ziyade, ROI alanlarındaki rnek miktarını artırmaya ynelik bir strateji izlediđini ve bu nedenle, belirsizliđi dűřűrme performansının tam olarak optimize edilemediđini dođrular niteliktedir.

UCB dűl fonksiyonu ise benzer Őekilde davranarak 3-24.e ve 3-25.e haritalarında benzeyen varyans dađılımları oluřturmuřtur. UCB dűl fonksiyonunun dűřűk belirsizlik seviyelerinde sműrű miktarını artması, bunun temel nedeni olarak n plana ıkmaktadır. Bununla birlikte, 300 rnek toplandıđında maksimum varyans deđerini, entropi dűl fonksiyonunun 300 rnekte ulařtıđı deđere yaklařmıřtır. Bu durum, UCB dűl fonksiyonunun yntemsel olarak dűřűk belirsizlikli alanlarda sműrűye daha fazla odaklanırken, yeterli veri toplandıktan sonra entropi dűl fonksiyonuna benzer belirsizlik azaltma seviyelerine ulařabileceđini gstermektedir.



Şekil 3.25. Çapraz doğrulama yöntemi uygulanan dört ROI alanı için varyans haritası: a) $L_{entropi}$ ödül fonksiyonu kullanılarak toplanan 150 örnek b) $L_{entropi}$ ödül fonksiyonu kullanılarak toplanan 300 örnek c) L_{LS} ödül fonksiyonu kullanılarak toplanan 150 örnek d) L_{LS} ödül fonksiyonu kullanılarak toplanan 300 örnek e) L_{UCB} ödül fonksiyonu kullanılarak toplanan 150 örnek f) L_{UCB} ödül fonksiyonu kullanılarak toplanan 300 örnek

Şekil 3.26, farklı ödül fonksiyonlarının (Seviye belirleme, UCB, Entropi) artan örnek sayısı ile birlikte tahmin edilen ortalama varyans üzerindeki etkilerini ve performanslarını dört farklı senaryoda (1, 2, 3 ve 4 ROI) incelemektedir.

Bu kapsamda, Şekil 3.26.a incelendiğinde, tek bir ROI'ye odaklanıldığında seviye belirleme (1_ROI_LS) ve UCB (1_ROI_UCB) arasında başlangıçta benzer bir varyans düşüşü görülürken, Entropi (1_ROI_ENT) fonksiyonunun varyansı daha farklı ve daha uzun vadede azalttığı görülmektedir. Çapraz doğrulama yöntemi uygulandığında tüm ödül fonksiyonlarının varyansı sürekli bir şekilde azalttığı ve minimum seviyelere ulaştığı belirlenmiştir.

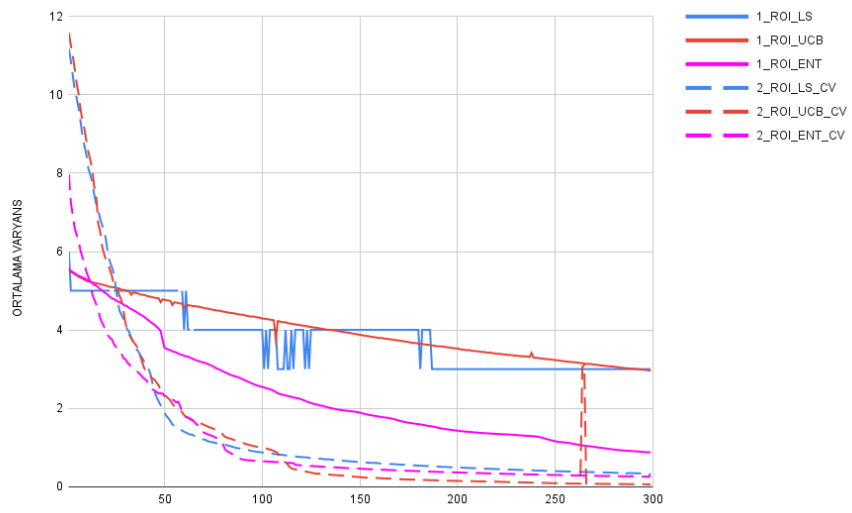
Şekil 3.26.b'de ortalama varyansın genel olarak düşme eğiliminde olduğu, çapraz doğrulama yönteminin ise daha düşük bir başlangıç varyansı ile başlamasına rağmen düşme eğilimini değiştirmedeği gözlemlenmiştir.

Şekil 3.26.c, seviye belirleme (3_ROI_LS) ile UCB (3_ROI_UCB) arasındaki performans farklılıklarını daha net bir şekilde ortaya koymaktadır. Seviye belirleme, ortalama varyans başlangıçta hızlı bir şekilde azaltırken, UCB daha değişken bir performans göstermiştir. Entropi (3_ROI_ENT) fonksiyonu ise 150 örnekten sonra en düşük varyans seviyesine ulaşarak daha tutarlı bir düşüş sergilemiştir. Çapraz doğrulama yöntemi ile iniş-çıkışların azaldığı ve daha tutarlı bir düşüş gerçekleştiği gözlemlenmiş, fakat sıralamada değişiklik olmamıştır.

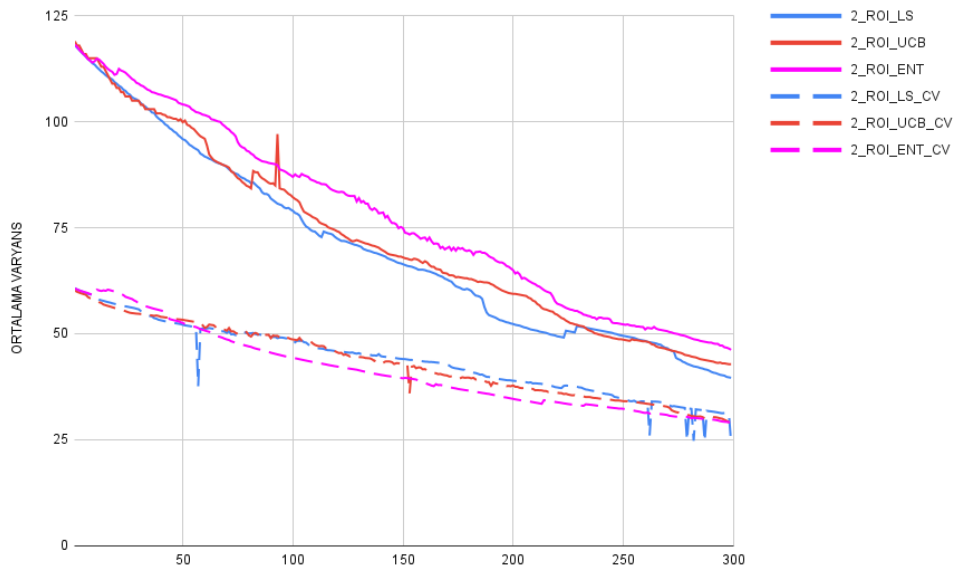
Şekil 3.26.d ise benzer bir trendi sürdürmektedir. Başlangıçta tüm ödül fonksiyonları benzer bir varyans düşüşü gösterirken, çapraz doğrulama yöntemi uygulandığında seviye belirleme ödül fonksiyonu ile ortalama varyans en düşük seviyeye ulaşmıştır. Entropi ve UCB ödül fonksiyonları ise benzer bir varyans düşüşü göstermiştir.

Bu grafikler genel olarak, ödül fonksiyonlarının farklı örnek sayılarında nasıl performans sergilediğini ve varyans üzerindeki etkilerini açıkça göstermektedir. İlk bakışta seviye belirleme ödül fonksiyonunun hızlı ROI tespiti ve erken dönemde hızlı varyans düşüşü sağlama yeteneği ön plana çıkmaktadır. Ancak, daha uzun vadede Entropi fonksiyonunun özellikle çapraz doğrulama (CV) yöntemi ile varyansı minimize ettiği dikkat çekmektedir. UCB ise genel olarak daha dengeli ve tutarlı bir performans sergileyerek, varyansı sürekli olarak azalan bir eğilime sahiptir.

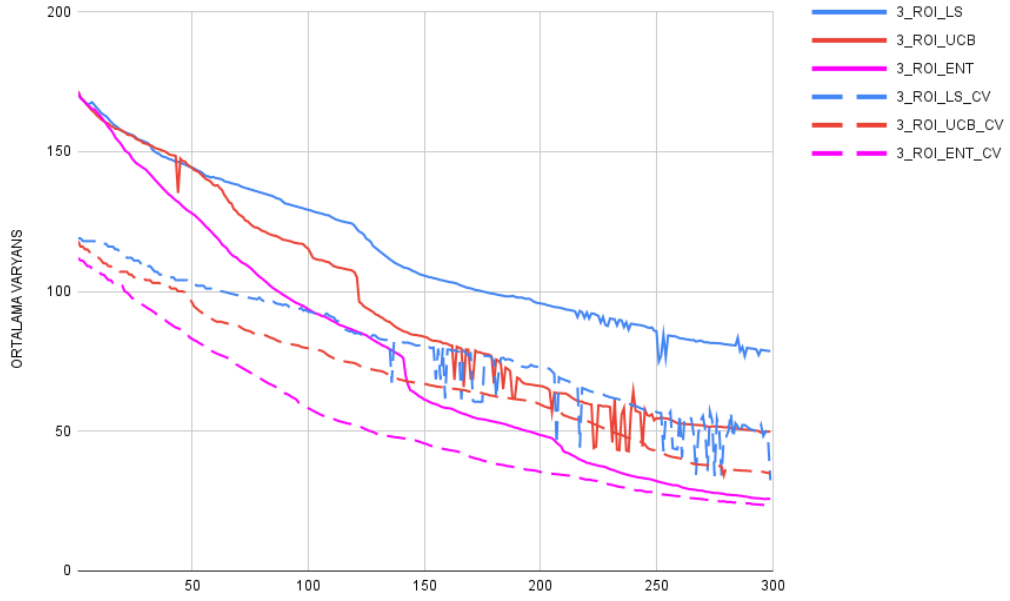
Bu analizler, çevre izleme ve benzeri uygulamalarda hangi ödül fonksiyonunun seçileceğine dair önemli ipuçları sunmaktadır. Örneğin, hızlı kararların gerektiği senaryolarda seviye belirleme ödül fonksiyonu tercih edilebilirken, daha kapsamlı ve derinlemesine bir analiz gereken durumlarda Entropi fonksiyonu daha uygun olabilir. UCB ödül fonksiyonu ise dengeli ve güvenilir bir performans arayan uygulamalar için ideal bir seçenek olarak değerlendirilebilir. Her bir ödül fonksiyonunun avantaj ve dezavantajlarının dikkatlice değerlendirilmesi, izleme süreçlerinin verimliliğini artırmada kritik bir rol oynamaktadır.



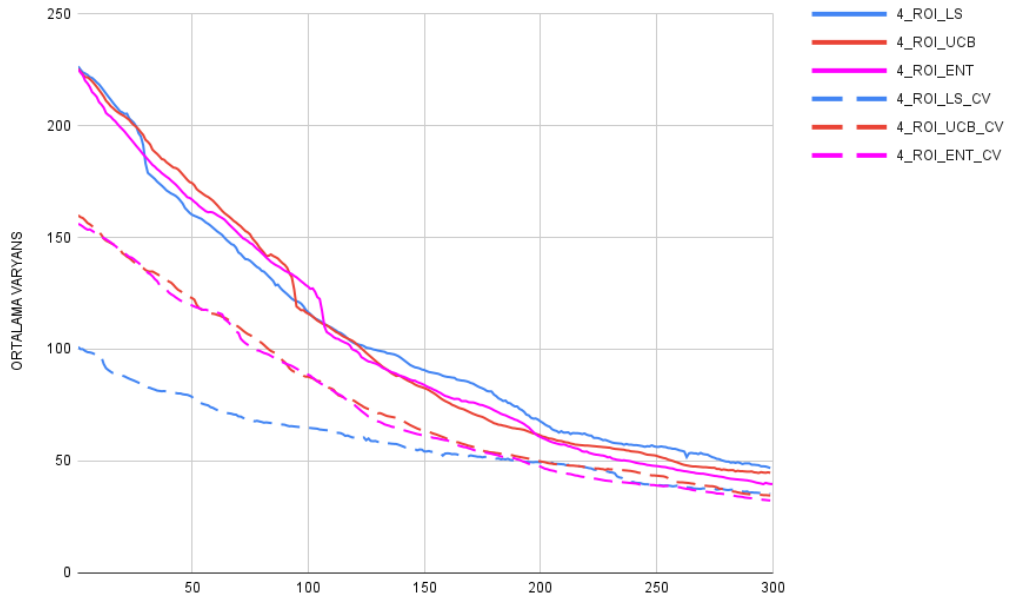
(a)



(b)



(c)



(d)

Şekil 3.26. Farklı ROI sayıları için elde edilen veri setlerinde ortalama varyans a) 1 ROI alanı olan veri seti b) 2 ROI alanı olan veri seti c) 3 ROI alanı olan veri seti d) 4 ROI alanı olan veri seti

Şekil 3.27’de sırasıyla 1,2,3, ve 4 ROI alanına sahip veri setlerinde test edilen ödül fonksiyonlarının performansını incelemek için varyansın standart sapması verilmiştir. Çevresel izleme problemlerinde belirsizliği azaltmak kadar alanın tamamının belirli bir seviyede belirsizliğe ulaşmasında önem arz etmektedir. Yani ortalama varyans kadar maksimum varyansında

düşürülmesi gerekmektedir. Bu nedenle varyansın standart sapması ödül fonksiyonlarının başarımını değerlendirirken önem arz etmektedir.

Şekil 3.27.a'da tek ROI odaklı analizde, Entropi (1_ROI_ENT) fonksiyonu hızlı bir varyans düşüşü sağlayarak başlangıçta en düşük varyans seviyesine ulaşmaktadır. Fakat çapraz doğrulama yöntemi uygulandığında en geç minimum değere ulaşmıştır. Bu durum entropinin yüksek varyanslı haritalarda alanı daha erken öğrendiğini fakat düşük varyanslı haritalarda alanın tamamını öğrenmeye çalışırken maksimum varyansı azaltmakta geç kaldığını göstermektedir. Seviye belirleme (1_ROI_LS) ve UCB (1_ROI_UCB) ise başlangıçta dalgalı fakat genel olarak sürekli bir varyans azalımı sunmaktadır. Fakat çapraz doğrulama yöntemi uygulanan grafiklerde (1_ROI_LS_CV, 1_ROI_UCB_CV, 1_ROI_ENT_CV) minimum değere ilk seviye belirleme ödül fonksiyonu ile ulaşılmıştır.

Şekil 3.27.b'de ise iki ROI odaklı analizde, seviye belirleme (2_ROI_LS) ödül fonksiyonu hızlı başlangıcıyla ortamı en erken öğrenir. Bu, seviye belirleme ödül fonksiyonunun efektif bir ROI tespit yeteneğine sahip olduğunu göstermektedir. UCB (2_ROI_UCB) daha dengeli ve stabilize bir varyans azalması sunarken, Entropi (2_ROI_ENT) yüksek varyans değerindeki haritada ortamı öğrenirken başarısız olmuştur. Fakat çapraz doğrulama yöntemi uygulandığında en düşük varyansın standart sapması değerine ilk entropi ödül fonksiyonuyla ulaşılmıştır.

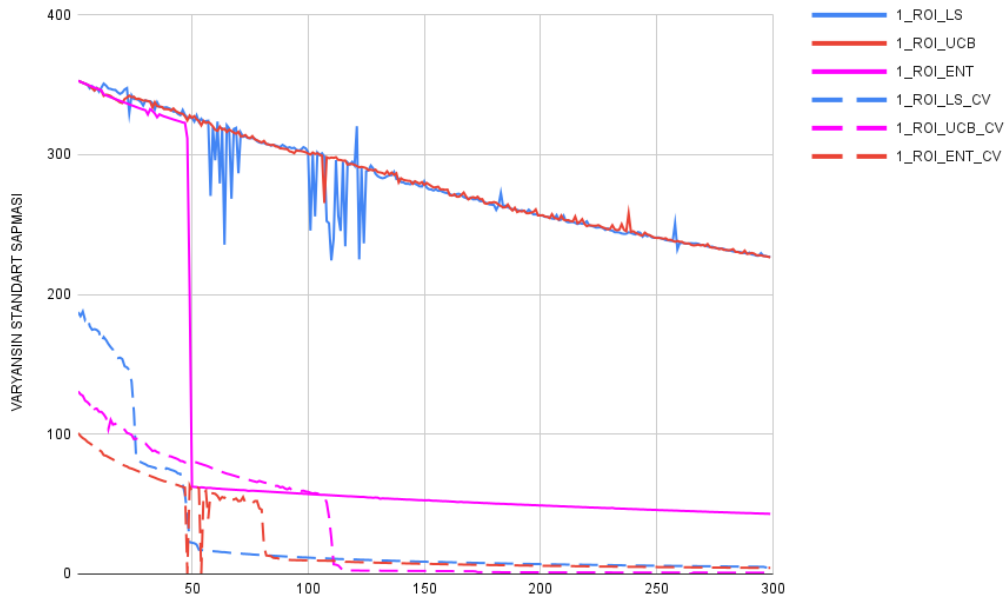
Şekil 3.27.c'de üç ROI'li veri setinde, seviye belirleme (3_ROI_LS) ve UCB (3_ROI_UCB) arasındaki fark daha belirgin hale gelmektedir. Seviye belirleme hızlı varyans düşüşü sağlarken, UCB inişli çıkışlı bir performans sergilemektedir. Entropi (3_ROI_ENT) ise daha tutarlı bir varyans azaltımı performansı göstermekte, 150 örnekten sonra en düşük varyans seviyesine ulaşarak derinlemesine keşifle başarılı olmaktadır. Fakat çapraz doğrulama yönteminin uygulanmasıyla en hızlı keşfi yine entropi ödül fonksiyonu yapmıştır. Seviye belirleme ödül fonksiyonu minimum değere entropi ödül fonksiyonundan daha geç ulaşmıştır. UCB ödül fonksiyonu ile daha yüksek bir varyansın standart sapması değeri elde edilmiştir.

Şekil 3.27.d.'de dört ROI alanı olan veri seti için varyansın standart sapması değişimi verilmektedir. Seviye belirleme (4_ROI_LS) ve UCB (4_ROI_UCB), ortalama varyansı etkin bir şekilde azaltmada başarılı olmuşlardır. Seviye belirleme, hem yüksek varyanslı haritada hem düşük varyanslı haritada ilk en düşük varyansın standart sapması değerine ulaşması ile dikkat çekmektedir. Bu durum seviye belirleme ödül fonksiyonunun ROI alanlarına odaklanarak ortamı öğrenme çabasının komplike veri setlerinde daha başarılı olacağını göstermektedir. UCB ödül fonksiyonu ise ortamı entropi ödül fonksiyonuna göre daha erken öğrenmesine rağmen ortalama varyansın standart sapması değerini entropi ödül fonksiyonu kadar düşürememiştir.

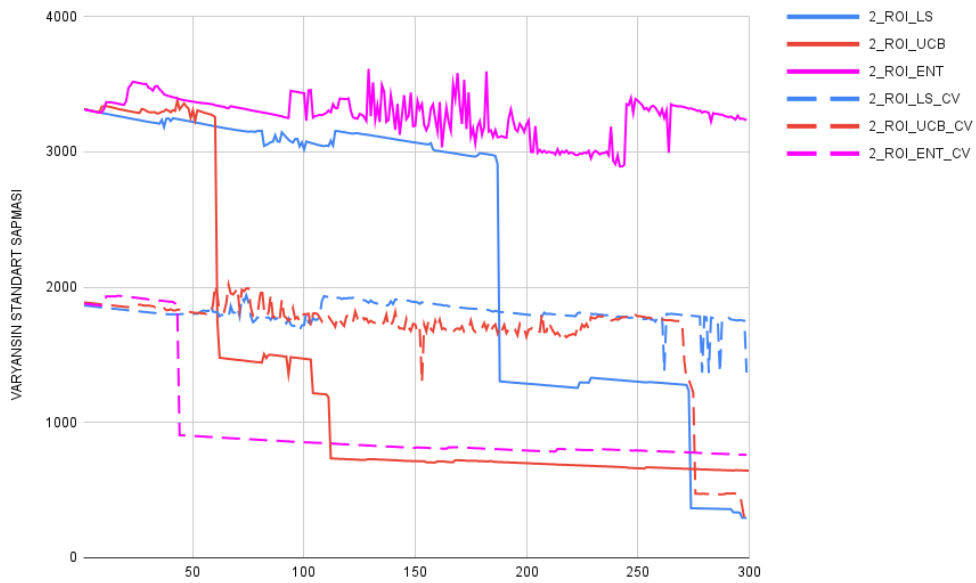
Şekil 3.27'nin sunduğu dört senaryo da ödül fonksiyonlarının farklı örnek sayılarında nasıl performans gösterdiğini ve varyans üzerinde nasıl bir etki bıraktığını açıkça ortaya koymaktadır. Seviye belirleme ödül fonksiyonunun hızlı ROI tespiti yaparak erken dönemde belirgin varyans

düşüşü sağlaması, özellikle kritik zaman faktörünün olduğu durumlarda büyük bir avantaj sağlamaktadır.

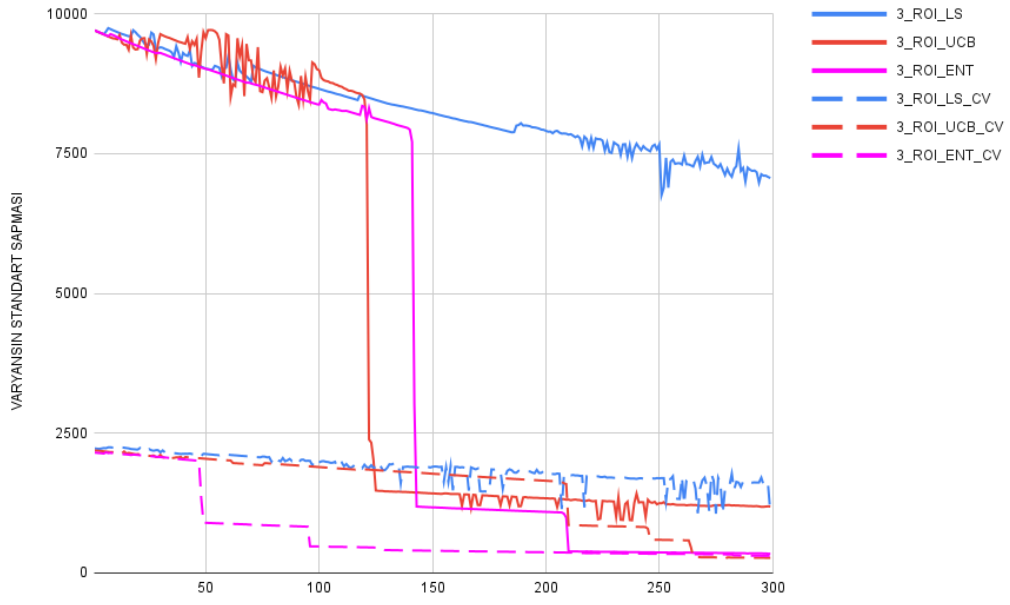
UCB fonksiyonu, daha dengeli ve sürdürülebilir bir performans sunarken, inişli çıkışlı bir varyans azaltma stratejisi izlemektedir. Entropi fonksiyonu ise daha geniş veri setleri üzerinde uzun vadede en düşük varyans seviyesine ulaşarak kapsamlı bir keşif süreci sunmaktadır. Bu bulgular, çevre izleme çalışmalarında hangi ödül fonksiyonunun kullanılacağına karar verirken dikkate alınması gereken önemli bilgiler sunmaktadır.



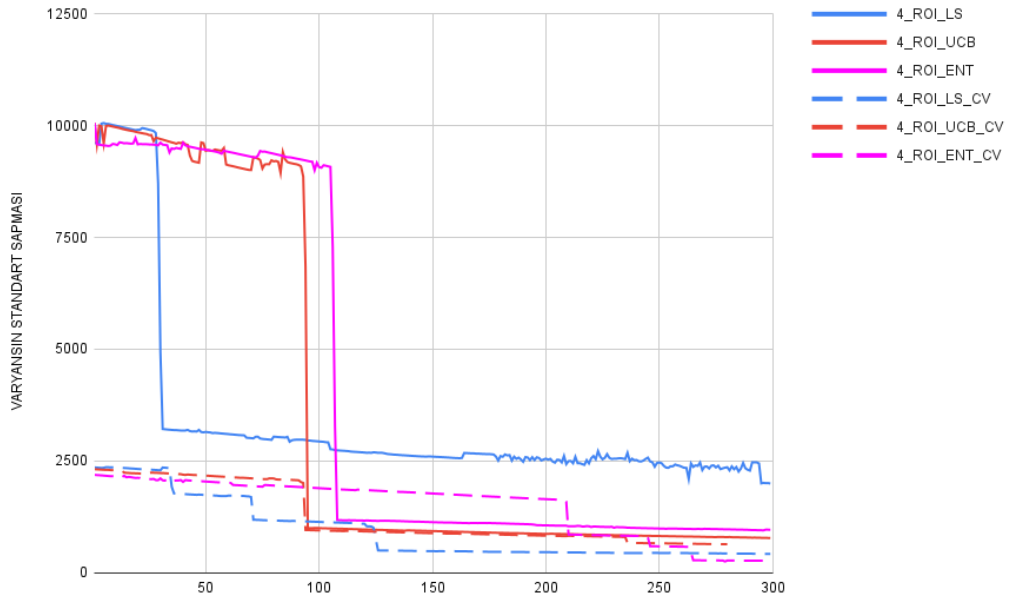
a)



b)



c)



d)

Şekil 3.27. Farklı ROI sayıları için elde edilen veri setlerinde tahmin edilen varyansın standart sapması a)1 ROI alanı olan veri seti b) 2 ROI alanı olan veri seti c) 3 ROI alanı olan veri seti d) 4 ROI alanı olan veri seti

3.2.4. Tahmin performansını değerlendirmek için, RMSE'deki değişimin incelenmesi

RMSE değeri, tahmin performansını değerlendirmek amacıyla kullanılan önemli bir metriktir ve Denklem 26'da hesaplanan formülü kullanılarak belirlenmiştir. RMSE'nin düşük olması, modelin tahminlerinin gerçek değerlere daha yakın olduğunu ve dolayısıyla daha iyi bir performans sergilediğini göstermekte olup, model doğruluğunu ölçmek için yaygın olarak kullanılmaktadır.

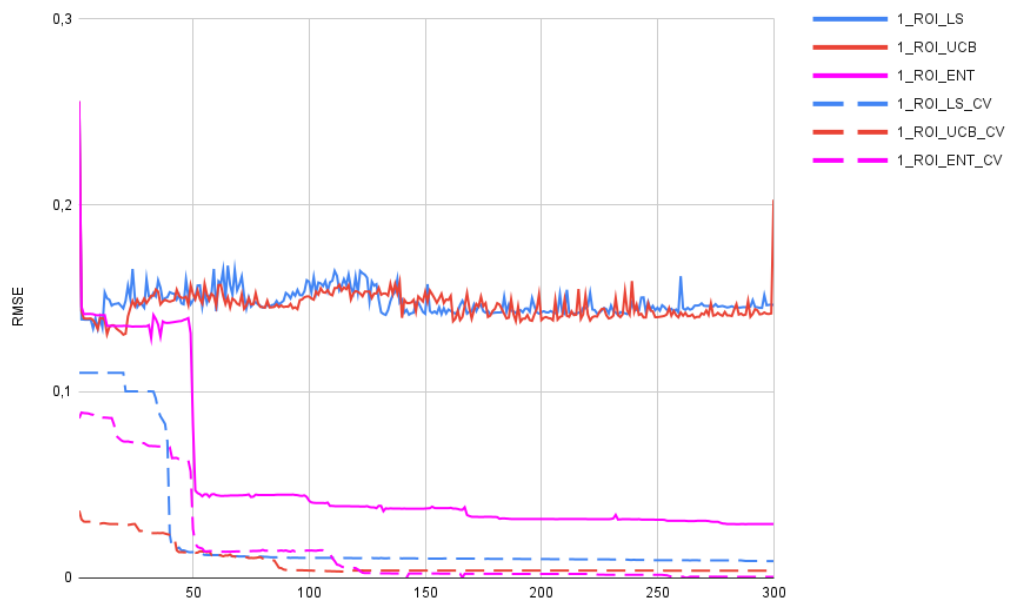
Şekil 3.28'de, dört farklı veri setinde çeşitli ödül fonksiyonlarının RMSE değerleri üzerindeki etkisi incelenmiştir. Bu inceleme, veri setlerinin sırasıyla 1, 2, 3 ve 4 ROI alanına sahip olduğu durumlarda yapılmıştır. Tek ROI içeren veri setinde, yüksek Entropi (1_ROI_ENT) fonksiyonu başlangıçta belirgin bir RMSE düşüşü sağlamış ve yaklaşık 50 örnek sonrasında en düşük RMSE seviyelerine ulaşarak performansını sabitlemiştir. Seviye belirleme (1_ROI_LS) ve UCB (1_ROI_UCB) fonksiyonları ise başlangıçta daha yüksek RMSE değerlerine sahip olmasına rağmen, yaklaşık 150 örnek sonrasında benzer RMSE seviyelerine ulaşarak dengelenmişlerdir. Çapraz doğrulama yöntemi uygulandığında en hızlı RMSE düşüşü UCB ödül fonksiyonu tarafından sağlanmış, bunu seviye belirleme ve entropi ödül fonksiyonları takip etmiştir. Yüksek varyansla düşük RMSE değerine sadece entropi fonksiyonu ile ulaşılmıştır.

İki ROI içeren veri setinde, seviye belirleme (2_ROI_LS) fonksiyonu başlangıçta daha dalgalı bir performans sergileyip, yaklaşık 50 örnekten sonra belirgin bir azalma göstermiş ve daha sonra daha stabil bir düşüş trendi yakalamıştır. UCB (2_ROI_UCB) fonksiyonu, yaklaşık 70-80 örnekten sonra ani bir RMSE düşüşü gösterip, ardından daha stabil bir eğilimde devam etmiştir. Entropi (2_ROI_ENT) fonksiyonu ise yaklaşık 75 örnekte dikkat çekici bir azalma gerçekleştirip, performansını sabitlemiştir. Her üç fonksiyon yaklaşık 150 örnekten sonra benzer RMSE seviyelerinde stabilize olmuştur. Fakat çapraz doğrulama yapıldıktan sonra en düşük RMSE, entropi ödül fonksiyonuyla elde edilmiştir. UCB ve ardından seviye belirleme ödül fonksiyonları takip etmiştir. Bu, düşük varyansla seviye belirleme fonksiyonunun ROI alanlarına odaklanmada geç kaldığını göstermektedir.

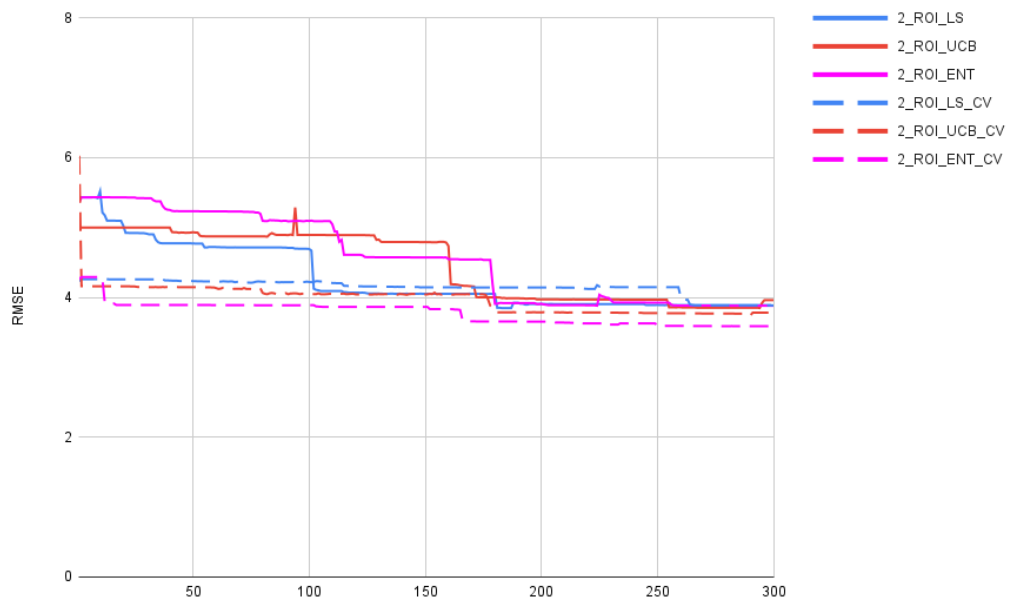
Üç ROI içeren veri setinde entropi fonksiyonu başlangıçta yüksek RMSE değerlerine sahip olmasına rağmen, yaklaşık 50 örnek sonrasında hızlı bir azalma sergileyerek 150 örnek sonrasında en düşük seviyelere ulaşmıştır. UCB fonksiyonu ise başlangıçta daha değişken bir performans göstermekte ve sonrasında stabilize olarak sürekli bir RMSE azalışı sağlamaktadır. Seviye belirleme fonksiyonu daha yavaş ancak düzenli bir azalma eğilimi göstermiş ve uzun vadede tutarlı bir RMSE azalışı sergilemiştir. Çapraz doğrulama yöntemi uygulandıktan sonra, seviye belirleme ödül fonksiyonu ortamı en iyi öğrenmiş olmasına rağmen en düşük RMSE değerine ulaşmada yetersiz kalmıştır. Entropi ödül fonksiyonu, RMSE değerini en son düşürmesine rağmen minimum seviyeye bununla ulaşmıştır.

Dört ROI içeren veri setinde, seviye belirleme ödül fonksiyonu hızlı ve erken RMSE düşüşü sağlayarak yaklaşık 75 örnek sonrasında en düşük seviyelere ulaşmıştır. Entropi ödül fonksiyonu diğer veri setlerinin aksine RMSE değerini çok daha yavaş düşürmüştür. Bu durum, modelin karmaşıklığı arttığında Entropi ödül fonksiyonunun belirsizliği azaltma eğilimi göstererek ROI alanlarını kaçırdığını ve ortamın genel öğrenme oranını yavaşlattığını belirtmektedir. UCB fonksiyonu ise 75 örnekten sonra ani bir düşüş göstererek, 150 örnek sonrasında stabil bir performansa ulaşmıştır. Çapraz doğrulama yöntemi uygulandıktan sonra, yine en düşük RMSE değeri Entropi ödül fonksiyonu ile elde edilmiştir. UCB ve seviye belirleme ödül fonksiyonları, RMSE değerini Entropi ödül fonksiyonu kadar düşürememiştir.

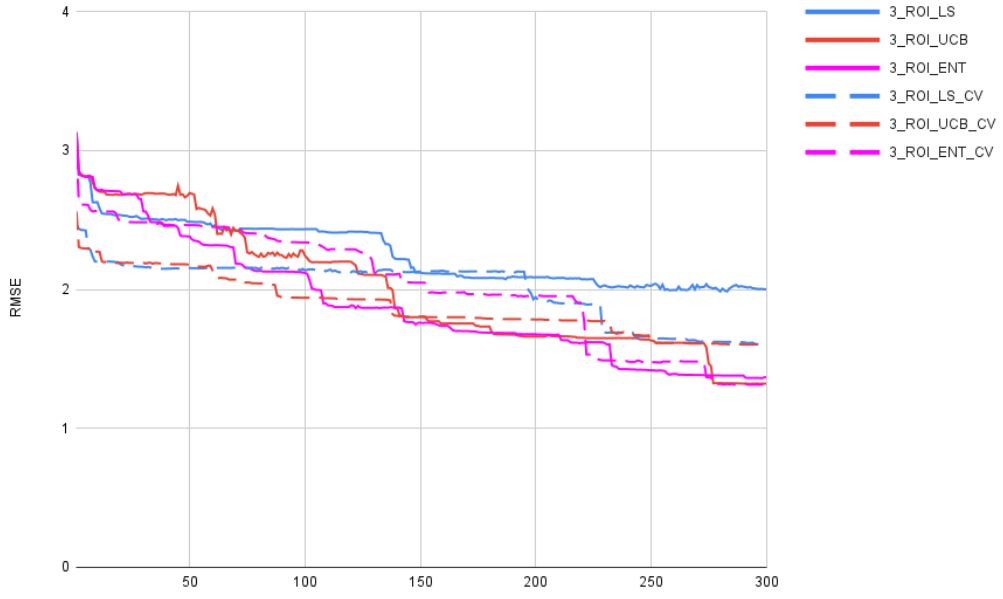
Özetle, ödül fonksiyonlarının RMSE değerleri üzerindeki etkileri, veri setlerinin ROI sayısına ve modelin karmaşıklığına bağlı olarak değişiklik göstermektedir. Bu analiz, ödül fonksiyonlarının başlangıç haritasının ve varyansın önemini vurgulamaktadır. Farklı ROI sayılarında ve veri setlerinde, entropi ödül fonksiyonu genellikle en düşük RMSE değerine ulaşırken, seviye belirleme ve UCB fonksiyonları değişken başarılı görülmüştür. Analiz, modelin kompleksliği arttıkça ödül fonksiyonlarının performansında farklılıklar ortaya koymakta ve bu, ödül fonksiyonlarının seçiminde dikkat edilmesi gereken kriterleri vurgulamaktadır.



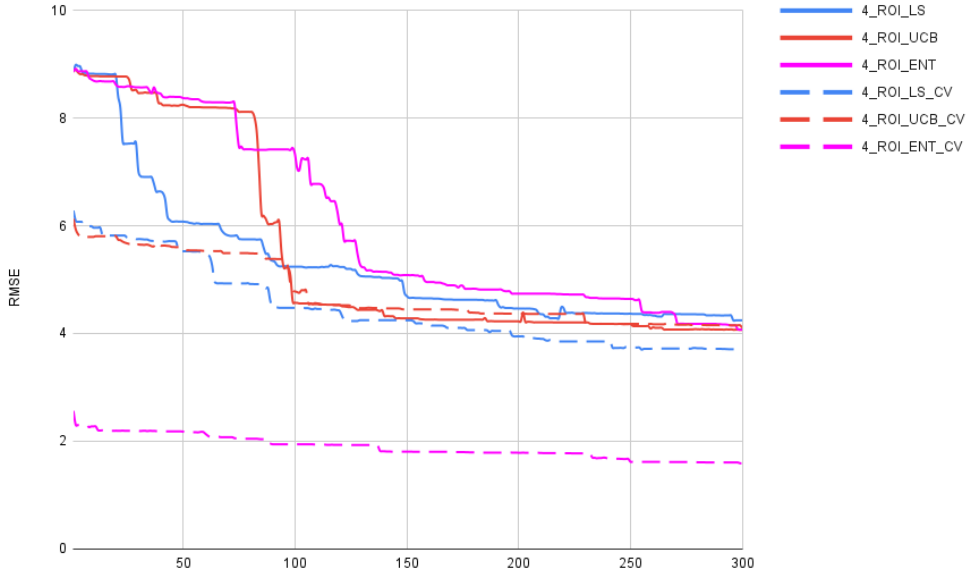
a)



b)



c)



d)

Şekil 3.28. Farklı ROI sayıları için elde edilen veri setlerinde RMSE değeri a) 1 ROI alanı olan veri seti b) 2 ROI alanı olan veri seti c) 3 ROI alanı olan veri seti d) 4 ROI alanı olan veri seti

Tablo 3.1. Entropi, UCB ve seviye belirleme ödül fonksiyonlarının performansını tek ROI, iki ROI, üç ROI ve dört ROI koşullarında çeşitli metriklerle değerlendirerek kapsamlı bir analiz sunmaktadır. Bu metrikler arasında ROI alanındaki görev oranı (%), Ortalama RMSE, Ortalama Varyans ve Ortalama Varyansın Standart Sapması gibi önemli performans göstergeleri yer almaktadır.

ROI Alanındaki Görev Oranı (%)

ROI alanındaki görev oranları incelendiğinde, entropi fonksiyonunun tek ROI durumunda %37.00 ile nispeten düşük bir görev oranı sergilediği görülmektedir. İki ROI (%7.33), üç ROI (%33.67) ve dört ROI (%20.67) koşullarında bu oranlarda belirgin dalgalanmalar gözlemlenmiştir. UCB fonksiyonu, tek ROI'de %84.34 gibi çok yüksek bir görev oranı ile öne çıkarken, iki ROI (%30.0), üç ROI (%78.0) ve dört ROI (%39.00) durumlarında bu oranlarda belirgin düşüşler görülmektedir. Bu, UCB'nin özellikle tek ROI durumunda yüksek bir başarı yakalasa da çoklu ROI koşullarında performansının düştüğünü ortaya koymaktadır. Seviye belirleme ödül fonksiyonu ise tek ROI'de %39.34 ile orta düzey bir başarı gösterirken, iki ROI (%39.34), üç ROI (%48.67) ve dört ROI (%48.67) durumlarında daha istikrarlı ve göreceli olarak artan görev oranlarına sahiptir.

Ortalama RMSE

Ortalama RMSE değerleri, tahmin doğruluğunu değerlendirmek için önemli bir göstergedir. Entropi, UCB ve seviye belirleme fonksiyonlarının performansı, farklı ROI stratejileri altında karşılaştırıldığında belirgin farklar ortaya çıkmaktadır. Tek ROI durumunda, Entropi fonksiyonu 0,05 gibi çok düşük bir RMSE değeri ile en yüksek doğruluğu sergiler. Bu, Entropi fonksiyonunun daha homojen ve basit bir veri yapısında etkili olduğunu gösterir. UCB (0,146) ve seviye belirleme (0,148) fonksiyonları da düşük RMSE değerleri sunar ancak entropi'ye kıyasla biraz daha yüksek hata oranlarına sahiptir. İki ROI durumunda, verinin karmaşıklığı artarken hata oranları da yükselir. Seviye belirleme fonksiyonu, 4,27 RMSE değeri ile en iyi performansı gösterir, bu da onun veri karmaşıklığı ile daha iyi başa çıkabildiğini gösterir. UCB (4,46) ve entropi (4,55) fonksiyonları ise daha yüksek RMSE değerleri sergiler. Üç ROI durumu, daha fazla karmaşıklık ve belirsizlik yaratabilir. Bu durumda, Entropi fonksiyonu 1,89 RMSE değeri ile en iyi performansı sunar, UCB (1,99) ve seviye belirleme (2,25) ise daha yüksek hata oranlarına sahiptir. Dört ROI durumu ise en karmaşık strateji olup, hata oranları genel olarak en yüksek seviyede izlenir. Seviye belirleme ödül fonksiyonu, 5,30 RMSE değeri ile en düşük hatayı sunar ve en etkili tahmin doğruluğunu sağlar. UCB fonksiyonu bu duruma oldukça yakın bir performans gösterir (5,48) ancak entropi fonksiyonu burada en yüksek hata oranına (6,03) sahiptir. Bu sonuçlar, ROI stratejilerinin ve ödül fonksiyonlarının tahmin doğruluğu üzerindeki etkilerini açıkça ortaya koyar.

Ortalama Varyans

Ortalama varyans değerleri, modelin tahminlerindeki belirsizliği değerlendirir. Entropi fonksiyonu tek ROI'de 2.32 gibi düşük bir varyans değeri gösterirken, iki ROI (70.21), üç ROI (76.34) ve dört ROI (101.37) durumlarında belirgin şekilde artan varyans değerleri, belirsizliğin arttığını ortaya koymaktadır. UCB ödül fonksiyonu için, tek ROI'de 3.97 olan varyans değeri, iki ROI'de 72.17'ye, üç ROI'de 94.00'e ve dört ROI'de 103.04'e yükselmiştir. Bu, UCB'nin tek ROI'de

daha kararlı tahminlerde bulunduğunu, fakat çoklu ROI koşullarında belirsizliğin arttığını göstermektedir. Seviye belirleme fonksiyonunun ortalama varyans değeri de benzer eğilimler göstermekte, tek ROI'de 3.77 iken, iki ROI (77.45), üç ROI (114.78) ve dört ROI (104.82) gibi daha yüksek değerlere ulaşmaktadır. Bu, belirsizliğin çoklu ROI koşullarında arttığını belirtir.

Ortalama Varyansın Standart Sapması

Ortalama varyansın standart sapması, varyans dağılımının genişliğini ölçer ve bu metrik de belirsizliği daha derinlemesine ele alır. Entropi fonksiyonu için, tek ROI'de 97,46 olan varyansın standart sapması, iki ROI (2359,79), üç ROI (4508,47) ve dört ROI (4028,19) durumlarında büyük artışlar sergilemektedir. Bu, Entropi fonksiyonunun tek ROI'de daha kararlı bir performans sergilerken, çoklu ROI koşullarında belirsizliğin büyük ölçüde arttığını göstermektedir. UCB fonksiyonu için, tek ROI'de 282,87 olan standart sapma, iki ROI, üç ROI ve dört ROI koşullarında sırasıyla (1338,8), (4479,25) ve (3573,64) olarak artış göstermektedir. Bu, UCB fonksiyonunun belirsizliklerinin de çoklu ROI koşullarında ciddi şekilde arttığını ortaya koymaktadır.

Seviye belirleme ödül fonksiyonu, benzer bir eğilim sergileyerek, tek ROI'de 279,11 olan standart sapma değerini, iki ROI (3269,09), üç ROI (8297,25) ve dört ROI (3380,61) koşullarında belirgin şekilde artırmaktadır. Bu, tüm ödül fonksiyonları için çoklu ROI stratejilerinin daha fazla belirsizlik ve geniş varyans dağılımıyla ilişkili olduğunu ortaya koymaktadır.

Tablo 3.1'de görüldüğü üzere, ROI alanındaki görev oranı metriklerinde, UCB tek ROI'de en yüksek başarıyı sağlarken, çoklu ROI koşullarında önemli düşüşler yaşamaktadır. Entropi ve seviye belirleme fonksiyonları ise daha istikrarlı ve dengeli görev oranları sergilemektedir. Ortalama RMSE analizleri, tek ROI koşullarında tüm fonksiyonların daha düşük hata oranları sunduğunu, ancak çoklu ROI koşullarında tahmin doğruluğunun azaldığını göstermektedir. Ortalama Varyans ve Ortalama Varyansın Standart Sapması değerleri ise, çoklu ROI stratejilerinin daha yüksek belirsizlikle ilişkili olduğunu ve daha geniş varyans dağılımları sergilediğini ortaya koymaktadır.

Tablo 3.1. Ödül Fonksiyonlarının Karşılaştırılması

		ROI alanındaki görev oranı (%)	Ortalama RMSE	Ortalama Varyans	Ortalama Varyansın Standart Sapması
Tek ROI	Entropi	37,00	0,05	2,32	364,92
	UCB	80,67	0,146	3,97	282,87
	Seviye Belirleme	92,34	0,148	3,77	279,11
İki ROI	Entropi	7,33	4,55	70,21	2359,79
	UCB	30,0	4,46	72,17	1338,8
	Seviye Belirleme	39,34	4,27	77,45	3269,09
Üç ROI	Entropi	33,67	1,89	76,34	4508,47
	UCB	78,0	1,99	94,00	4479,25
	Seviye Belirleme	80,34	2,25	114,78	8297,25
Dört ROI	Entropi	20,67	6,03	101,37	4028,19
	UCB	39,00	5,48	103,04	3573,64
	Seviye Belirleme	48,67	5,30	104,82	3380,61

Tablo 3.1 ve Tablo 3.2 arasında yapılan karşılaştırma, verilere çapraz doğrulama uygulandığında ödül fonksiyonlarının performansında genel bir iyileşme olduğunu göstermektedir. Çapraz doğrulama sırasında ROI alanındaki görev oranlarında genellikle bir artış gözlemlenirken, UCB fonksiyonunun özellikle tek ve dört ROI koşullarında yüksek performans sergilediği, entropi ve seviye belirleme fonksiyonlarının ise daha dengeli ve istikrarlı görev oranlarına sahip olduğu görülmektedir. Ortalama RMSE değerlerinde çapraz doğrulama uygulandığında önemli bir düşüş yaşanmakta, bu da doğrulama sürecinin modelin tahmin doğruluğunu artırdığını göstermektedir. Ortalama varyans ve varyansın standart sapması değerlerinde ise yine çapraz doğrulama ile belirgin azalmalar görülmekte olup, belirsizliğin ve varyans dağılımının daraldığını işaret etmektedir. Bu sonuçlar, çapraz doğrulamanın modeli daha sağlam ve güvenilir hale getirdiğini ve çoklu ROI koşullarında daha kararlı performans sağladığını doğrulamaktadır.

Tablo 3.2. Çapraz doğrulama yöntemi uygulanan ödül fonksiyonlarının karşılaştırılması

		ROI alanındaki görev oranı (%)	Ortalama RMSE	Ortalama Varyans	Ortalama Varyansın Standart Sapması
Tek ROI	Entropi	36,40	0,016	1,36	97,46
	UCB	84,34	0,008	1,14	22,45
	Seviye Belirleme	84,67	0,022	1,48	26,68
İki ROI	Entropi	9,67	3,78	41,27	974,23
	UCB	39,67	3,97	43,19	1658,38
	Seviye Belirleme	49,34	4,15	43,81	1810,39
Üç ROI	Entropi	36,00	2,02	52,69	723,023
	UCB	70,67	2,60	69,97	1498,96
	Seviye Belirleme	73,00	2,00	79,78	1858,64
Dört ROI	Entropi	21,00	4,68	77,93	980,0
	UCB	40,00	5,15	77,07	1342,09
	Seviye Belirleme	66,00	4,85	58,52	1896,81

Ödül fonksiyonlarının ROI alanlarındaki başarımını değerlendiren çalışmaların sayısı oldukça sınırlıdır. Stankiewicz vd. [45], UCB ödül fonksiyonunun başarımını incelenmiştir. Çalışmada 2 farklı ROI alanı kullanılmış ve elde edilen sonuçlar karşılaştırılmıştır. Veri setinde çözünmemiş oksijen miktarı tahmin edilmiş ve 0 ile 1 arasında değişen bu değerler için 130 örnek toplanmıştır. Tablo 3.3'te iki ROI alanı bulunan test ortamından elde edilen sonuçlar, Stankiewicz ve arkadaşları [45] tarafından yapılan deneysel veri seti ile elde edilen sonuçlarla karşılaştırılmıştır. Deneysel veri setinde ROI alanındaki görev oranı %17 olarak belirlenmiştir. Çalışmada yapılan test ortamında ise çapraz doğrulama yöntemi uygulanmadan önce %12 görev oranı elde edilmiştir. Gauss regresyon analizinde, veri setinin değişkenliği arttıkça hata oranının da arttığı gözlemlenmiştir. Dolayısıyla, deneysel veri setinde RMSE oranı ve ortalama varyans düşük çıkmıştır. Ancak, çapraz doğrulama yöntemi uygulandıktan sonra ROI alanındaki görev oranı %32'ye çıkarılmış ve RMSE oranı ile ortalama varyans daha düşük değerler elde edilmiştir. Bu sonuçlar, UCB ödül fonksiyonunun başarımını arttıracak çapraz doğrulama yönteminin önemini vurgulamaktadır.

Tablo 3.3. Literatürdeki çalışmayla UCB ödül fonksiyonunun performans metriklerinin karşılaştırılması

	RMSE	Ortalama Varyans	ROI alanındaki görev oranı (%)
Stankiewicz vd. [45]	0.056	0.098	17
İki ROI	5.77	182.19	12
Çapraz Entropi Yöntemi Uygulanan İki ROI	5.39	85.95	32

3.3. Gerçek Zamanlı Denizdeki Klorofil Miktarını Belirlemek için Geliştirilen Hibrit MCTS-BnB Algoritması

Bu bölümde, çevresel izleme süreçlerinde karşılaşılan gezgin satıcı problemi için hibrit bir MCTS-BnB algoritması önerilmektedir. Gezgin satıcı problemi, bilgilendiriciliği yüksek örneklerin en kısa sürede toplanmasını içeren bir optimizasyon problemidir. Bu optimizasyon sorunu için önerilen algoritmanın performansı, MCTS algoritması ve ESAS algoritması ile karşılaştırılmıştır. Başlangıçta deneysel bir veri seti üzerinde kıyaslanan performans, daha sonra diğer veri setlerinde de incelenmiştir. Çevresel izleme süreçlerinde, GP ortamı öğrenirken eğitim verilerinin birbiriyle olan ilişkisi, kovaryans fonksiyonu kullanılarak hesaplanmaktadır. Bu kovaryans fonksiyonu sayesinde ortam öğrenilir ve belirsizlik hesaplanır. Kovaryans fonksiyonunun hesaplanmasında kullanılan hiperparametreler, veri setinin değişkenliğine bağlı olarak farklılık göstermektedir. Rüzgar hızı gibi anlık değişimin yüksek olduğu veri setlerinde bu değer daha düşük iken, sıcaklık gibi anlık değişimin çok olmadığı veri setlerinde ise daha yüksek olmaktadır. Bu nedenle, modelin performansı, üç farklı hiperparametreye sahip veri setinde de test edilmiştir. Önerilen yöntemde, alt ve üst limitler kullanılarak yollara sınırlamalar getirilmiştir. Bu bölümde, literatürdeki diğer alt ve üst limit hesaplama yöntemleri kullanılarak mevcut veri setine uygulanmış ve modelin başarımı detaylı bir şekilde araştırılmıştır.

Son aşamada ise, önerilen yöntemin hesaplama gücünü azaltmak amacıyla alt örneklem kümesi, beş farklı kümeleme yöntemi kullanılarak belirlenmiş ve bu kümelerden elde edilen örnek sayısına bağlı olarak güzergahın optimale yakınlığı değerlendirilmiştir.

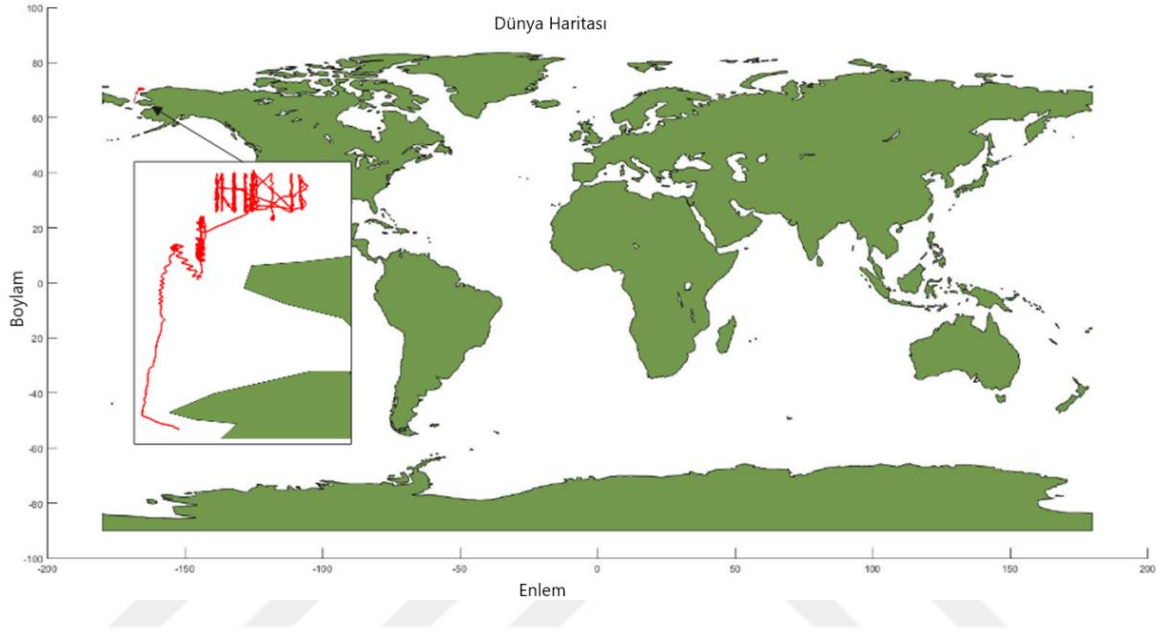
3.3.1. Saildrone SSS ve SST Veri Seti

Saildrone Arktik 2021 seferinde, Bering ve Çukçi Denizleri'nde gerçek zamanlı çevresel ölçümler almak amacıyla insansız yüzey araçları (İYA'lar) kullanılmıştır. Bu keşif, 6 Temmuz - 20 Eylül 2021 tarihleri arasında, toplam 76 gün süresince gerçekleştirilmiş olup, iki adet Saildrone'un kritik çevresel değişkenleri ölçebilecek şekilde gelişmiş sensörlerle donatıldığı bir kampanyayı kapsamaktadır [115]. Kampanyanın ana amacı, deniz ortamındaki klorofil konsantrasyonlarını tahmin etmektir. Bu hedef doğrultusunda, seferin 5. gününde alınan ölçümler kullanılmıştır.

Söz konusu veri seti, çeşitli ortam parametrelerini içermektedir. Bunlar arasında suyun tuzluluğu, sıcaklık, oksijen doygunluğu ve konsantrasyonu, iletkenlik, hava nem oranı ve rüzgar hızı gibi önemli değişkenler yer almaktadır. Günlük olarak yapılan ölçümlerde, her bir Saildrone'un aldığı 240 veri ölçümü kaydedilmiştir. Bu veriler, deniz ortamını daha doğru bir şekilde analiz etmeye olanak tanımaktadır.

Şekil 3.29'da, sefer süresince takip edilen örnekleme güzergahı görsel olarak sunulmuştur. Bu güzergah, belirli çevresel koşullara ve hedeflenen araştırma bölgesine uygun olarak seçilmiştir. Saildrone'ların rotası, Bering ve Çukçi Denizleri boyunca kritik noktaları kapsamaktadır.

Bu keşif, denizel ekosistemlerin daha iyi anlaşılması açısından büyük bir önem taşımaktadır. Klorofil konsantrasyonlarını tahmin etmeye yönelik ölçümler, deniz ortamının biyolojik ve fizikokimyasal özelliklerini anlama ve bu sistemlerde meydana gelen değişiklikleri izleme şansı sunmaktadır. Toplanan veriler, çevresel modellerin doğruluğunu artırmak ve deniz bilimleri alanında yeni bulgular sunmak için değerlidir.



Şekil 3.29. Veri Setinin Örnekleme Güzergahı

3.3.2. Klorofil-a parametresini bir özellik seçimi yöntemi kullanarak modelleme

Sudaki klorofil varlığı, deniz ekosistemleri için büyük öneme sahip pek çok temel ekolojik fonksiyonu yerine getirir. Klorofil, deniz ekosistemlerinde fotosentez yapan mikroskobik bitkiler olan fitoplanktonların enerji üretiminde merkezi bir role sahiptir. Fitoplanktonlar, güneş ışığını kullanarak organik maddeler üretirken oksijen salar ve bu oksijen sucul yaşamın devamı için gereklidir. Ayrıca, fitoplanktonlar besin zincirinin temel taşlarından birini oluşturur ve bu nedenle zooplankton ve balıklar gibi daha büyük organizmalar için temel bir besin kaynağıdır. Bu, denizlerdeki biyolojik çeşitliliğin ve ekosistem dengesinin korunmasına yardımcı olur.

Klorofil seviyelerinin doğru bir şekilde ölçülmesi genellikle laboratuvar analizlerini gerektirir. Ancak, geniş deniz alanlarından su örneklerinin toplanması, taşınması ve analiz edilmesi süreci, lojistik zorluklar taşır. Su örneklerinin laboratuvara ulaştırılması zaman alıcı olabilir ve bu süre zarfında örneklerin kalitesinde bozulmalar olabilir. Dahası, bu süreç gerçek zamanlı izleme yeteneklerini de azaltır ve denizdeki klorofil konsantrasyonunu sürekli ve anlık olarak izlemeyi zorlaştırır. Bu nedenle, laboratuvar ortamına ihtiyaç duymadan anlık ölçülebilen verileri kullanarak denizdeki klorofil miktarını tahmin etmek önemli bir ihtiyaç haline gelmiştir.

Denizdeki klorofil miktarını etkileyen çeşitli çevresel parametreler vardır. Şekil 3.30'da özetlendiği gibi, bu parametreler sıcaklık, tuzluluk, iletkenlik, oksijen doygunluğu, nem ve rüzgar hızıdır. Sıcaklık, fitoplankton aktivitesini ve klorofil üretimini doğrudan etkiler. Tuzluluk, suyun kimyasal bileşimini ve dolayısıyla mikroorganizmaların yaşam koşullarını belirler. İletkenlik, suyun içerisindeki iyonların hareketini ve fitoplanktonların besin alımını etkiler. Oksijen doygunluğu, suyun ne kadar oksijen içerdiğini ve dolayısıyla sucul organizmaların solunum faaliyetlerini gösterir.

Yine nem, atmosferdeki su buharı miktarı ile ilişkilidir ve deniz yüzeyinde buharlaşma gibi süreçleri etkileyebilir. Rüzgar hızı ise deniz yüzeyindeki karışımı etkileyerek, besin maddelerinin dağılımına ve fitoplanktonların üretkenliğine katkıda bulunur.

Bu parametrelerin her biri, denizdeki klorofil konsantrasyonunu ve dolayısıyla fitoplankton popülasyonlarını etkileyen doğrudan veya dolaylı faktörlerdir. Anlık verilerle bu parametrelerin izlenmesi, klorofil miktarının tahmini için kullanılan modellerin doğruluğunu artırabilir. Böylece, söz konusu ekosistemlerin sağlığının ve işleyişinin izlenmesi ve değerlendirilmesi daha kolay hale gelir. Örneğin, sıcaklık ve tuzluluk ölçümleri, deniz yüzeyine yakın alanlardaki fitoplankton üretimini ve dağılımını tahmin etmek için kullanılabilir. Yüksek sıcaklıklar genellikle metabolik hızları artırarak fitoplanktonların daha aktif hale gelmesine neden olabilir, ancak aynı zamanda stres faktörü de olabilir. Tuzluluk, çeşitli fitoplankton türlerinin yaşama ve büyüme yeteneklerini sınırlayabilir.

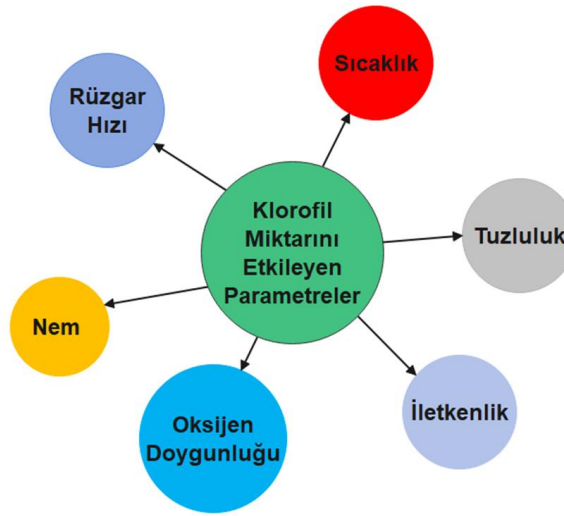
Nem ve rüzgar hızı ölçümleri, atmosferik koşulların deniz üzerindeki etkilerini anlamak için önemlidir. Rüzgar, özellikle, deniz suyunun dikey olarak karışmasını sağlar ve bu da derinlerde bulunan besin maddelerinin yüze çıkmasına yardımcı olur. Böylece, üst su tabakalarında besin zenginliği artar ve fitoplanktonlar için uygun koşullar oluşur.

Son olarak, iletkenlik ve oksijen doygunluğu gibi parametrelerle ilişkili kimyasal süreçler de fitoplankton popülasyonlarının sağlığını ve dağılımını anlamak açısından büyük önem taşır. Yüksek oksijen doygunluğu, fitoplanktonların sağlıklı bir şekilde fotosentez yapabildiğini gösterirken, düşük oksijen seviyeleri, ekosistemde olumsuz değişikliklerin ve potansiyel risklerin işareti olabilir.

Genel olarak, denizdeki klorofil miktarının tahmini için kullanılan bu parametrelerin tümü, deniz ekosistemlerinin dinamiklerini ve sağlığını izlemek için hayati öneme sahiptir. Geniş ölçekli deniz alanlarında klorofil miktarının anlık olarak tahmin edilebilmesi, yalnızca lojistik zorlukları ve laboratuvar bağımlılığını azaltmakla kalmaz, aynı zamanda çevresel değişikliklere yanıt verme hızını ve doğruluğunu da artırır. Dolayısıyla, denizdeki klorofil miktarının anlık izlenmesi ve tahmin edilmesi için geliştirilmiş teknolojiler ve yaklaşımlar, sürdürülebilir deniz yönetimi ve koruma stratejilerinin oluşturulması adına önemli bir rol oynar. Bu şekilde, ekosistemlerin ve deniz biyolojik çeşitliliğinin uzun vadeli sağlığı ve dengesinin korunmasına katkıda bulunulabilir ve aynı

zamanda balıkçılık gibi sucul kaynaklardan sürdürülebilir bir şekilde faydalanılabilir. Şekil 3.30'da özetlenen parametrelerin sürekli izlenmesi ve değerlendirilmesi, bilim insanlarına ve çevre yöneticilerine daha sağlıklı kararlar alabilme olanağı sağlar ve kapsamlı bir ekosistem yönetimi için temel oluşturur.

Suda bulunan klorofil varlığı, ekosistem için büyük öneme sahip birçok temel ekolojik fonksiyonu yerine getirir. Bu fonksiyonlar arasında fitoplankton ve zooplankton popülasyonlarını düzenleme, balık popülasyonlarını sürdürme, oksijen üretimini kolaylaştırma ve karbon döngüsüne katkıda bulunma gibi kritik roller yer almaktadır. Klorofil seviyelerinin doğru bir şekilde değerlendirilmesi genellikle laboratuvar analizlerini gerektirir. Ancak geniş deniz alanlarından su örneklerinin toplanması, örneklerin laboratuvara taşınması sürecinde ağırlığın artması ve bu süreçte gerçek zamanlı izleme yeteneklerinin azalması gibi lojistik zorlukları beraberinde getirir. Bu nedenle bu çalışmada laboratuvar ortamına ihtiyaç duymadan anlık ölçülebilen verileri kullanarak denizdeki klorofil miktarını tahmin etmeyi amaçlamaktadır. Şekil 3.30 denizdeki klorofil miktarını etkileyen parametreleri özetlemektedir.



Şekil 3.30. Denizdeki klorofil miktarını etkileyen parametreler

Bu çalışma, NASA tarafından toplanan gerçek zamanlı verileri kullanarak klorofil seviyelerini tahmin etmeyi amaçlamaktadır. Ancak mevcut tüm ölçümleri dahil etmek ağır bir hesaplama yükü yaratabilir ve gereksiz parametrelerin dahil edilmesi yanlış tahminlere yol açabilir. Bu nedenle, her bir parametre ile klorofil seviyeleri arasındaki ilişkiyi değerlendirmek üzere özellik seçimi yöntemi kullanılmıştır.

Özellik seçimi, makine öğrenimi modellerinin performansını ve yorumlanabilirliğini artırmak için kritik bir adımdır. En uygun ve bilgilendirici özellikleri seçerek, veri boyutunu

azaltmak mümkündür. Bu da daha hızlı eğitim süreleri, artırılmış doğruluk ve genelleme kabiliyeti sağlar. Özellik seçimi yöntemleri üç ana kategoriye ayrılır: filtreleme, sarmalama (wrapper) ve gömülü (embedded). Sarmalama yöntemi, özelliklerin bir alt kümesini elde etmek için kullanılırken, gömülü yöntem, modelin öğrenme sürecinde özelliklerin önemini hesaplar.

Filtreleme yönteminde ise, iki değişken arasındaki korelasyonu matematiksel olarak hesaplayan Pearson korelasyon katsayısı sıkça kullanılır. Pearson korelasyon katsayısı, iki değişken arasındaki doğrusal ilişkinin gücünü gösteren istatistiksel bir ölçüdür ve hedef değişkenle yüksek derecede korelasyon gösteren özellikleri tanımlamak için kullanılabilir. Bu katsayı genellikle -1 ile +1 arasında değerler alır ve 0'a yakın değerler, değişkenler arasında zayıf bir ilişki olduğunu gösterir.

Denklem 39, Pearson korelasyon katsayısını hesaplamak için kullanılan matematiksel formülü sunmaktadır. Bu katsayı, (r), aşağıdaki şekilde hesaplanır:

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}} \quad (39)$$

Burada, r pearson korelasyon katsayısını göstermektedir. x_i ve y_i gözlemlenen iki değişkeni temsil etmektedir. \bar{x} ve \bar{y} ise bu değerler ortalamayı temsil etmektedir. Özetle, Pearson korelasyon katsayısı, iki değişken arasındaki doğrusal ilişkiyi ölçer. Değerler arasında pozitif bir korelasyon varsa, katsayı pozitif olacaktır (yaklaşık olarak +1). Eğer negatif bir ilişki varsa, katsayı negatif olacaktır (yaklaşık olarak -1). Değişkenler arasında bir ilişki yoksa, katsayı yaklaşık olarak 0 olacaktır.

Çalışmada, NASA tarafından ölçülen sekiz farklı parametre ile klorofil seviyeleri arasındaki ilişki matematiksel olarak hesaplanmıştır. Bu hesaplamalar, her bir parametrenin klorofil seviyeleri üzerindeki etkisini değerlendirmek amacıyla yapılmıştır.

Tablo 3.4, bu parametrelerin klorofil seviyeleri ile arasındaki Pearson korelasyon katsayılarını göstermektedir. Bu tablo, hangi parametrelerin klorofil seviyeleri ile daha güçlü ilişkili olduğunu belirlemeye yardımcı olur. Klorofil seviyeleri ile güçlü bir korelasyon gösteren parametreler, özellik seçimi sürecinde tercih edilen özellikler olacaktır. Bu sayede, modelin hem doğruluğu artırılır hem de gereksiz parametrelerin etkisi minimize edilir.

Tablo 3.4'de sunulan bulgulara dayanarak, bu çalışmada klorofil seviyelerini tahmin etmek için tuzluluk, deniz sıcaklığı, iletkenlik ve rüzgar hızı seçilmiştir.

Tablo 3.4. Pearson Korelasyon Katsayısı

Ölçülen su parametresi	Pearson Korelasyon Katsayısı
O ₂ Konsantrasyonu	0.3097
Tuzluluk	-0.3300
Deniz Sıcaklığı	-0.4103
O ₂ Saturasyonu	-0.0921
İletkenlik	-0.4699
Nem	0.1887
Rüzgar Hızı	-0.5568
Hava Sıcaklığı	-0.1995

Bu çalışmada, NASA tarafından sağlanan açık kaynak veri seti, Bölüm 2.1.12’de anlatılan çoklu çıkış Gauss regresyon algoritmasında kullanılmıştır. Algoritmada enlem ve boylam, modeldeki girdi değişkenleri olarak seçilirken, rüzgar hızı, deniz sıcaklığı, tuzluluk ve iletkenlik ise çıktı değişkenleri olarak belirlenmiştir. Veri seti, deniz yüzeyi koşullarını modellemek amacıyla çeşitli meteorolojik ve oşinografik değişkenleri içermektedir. Şekil 3.31’de görüldüğü gibi, kırmızı 'o' sembolleri ile temsil edilen veri noktaları eğitim setini işaret ederken, mavi '+' sembolleri gerçek ölçümleri göstermektedir. Bu ayırım, modelin performansını ve doğruluğunu değerlendirmek amacıyla oldukça önemlidir.

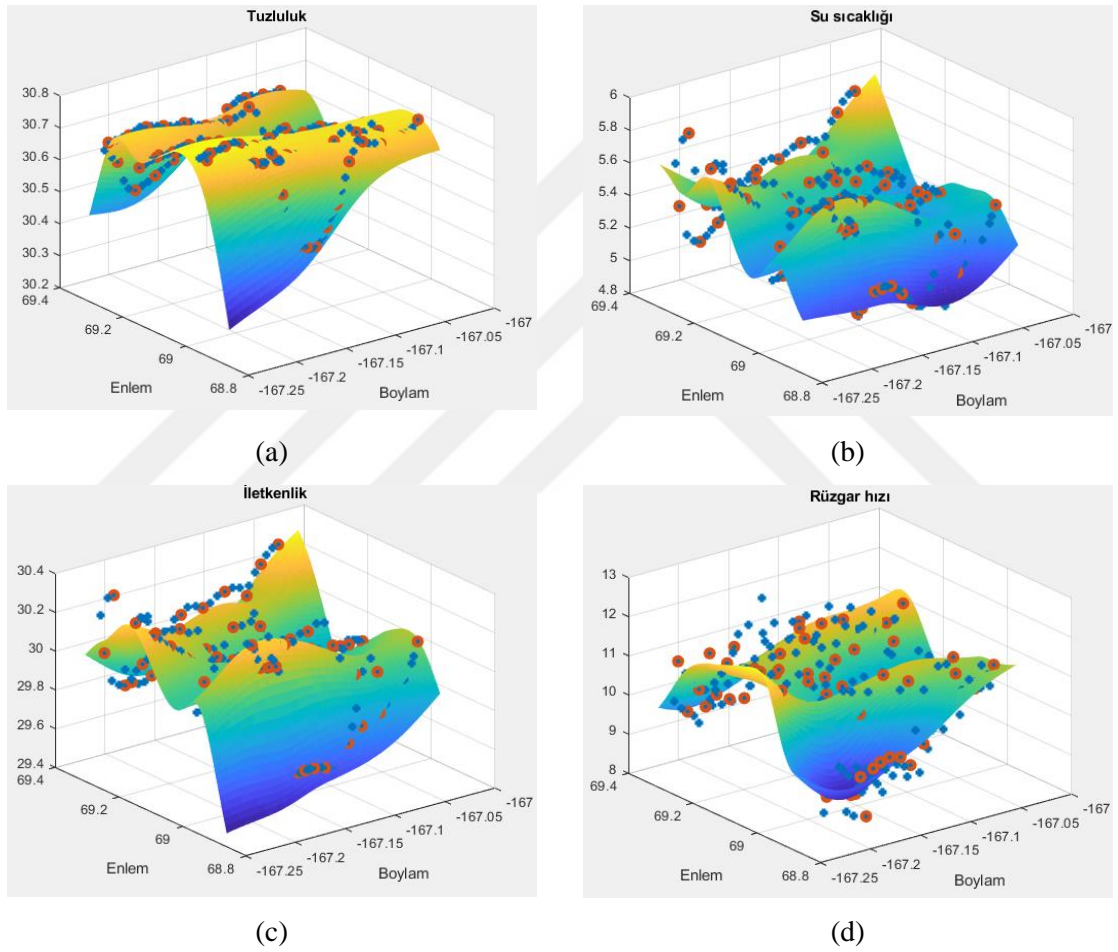
Seçilen bölgeye ait toplam 280 veri noktası, modele eğitim amaçlı olarak sağlanmıştır. Bu veri noktaları, modeli eğitmek ve farklı enlem ve boylam değerlerine karşılık gelen rüzgar hızı, deniz sıcaklığı, tuzluluk ve iletkenlik değerlerini tahmin edebilmesi için kullanılmıştır. Test verileri ise, belirli bir alanda modelin tahmin yeteneklerini değerlendirmek için kullanılmıştır. Test verileri, boylam için '-167.25, -167' ve enlem için '68.8, 69.4' koordinatlarında bir 100x100 ızgara oluşturacak şekilde 10.000 veri noktası olarak seçilmiştir.

Grafiklerde ise elde edilen sonuçlar detaylı bir şekilde gösterilmiştir. Grafiklerde, modelin tahmin ettiği değerler ile gerçek ölçümler arasındaki farklar görsel olarak sunulmuştur. Tuzluluk, su sıcaklığı ve iletkenlik değişkenlerinin ölçümleri, test verileriyle oldukça yakından eşleşmiştir. Bu, modelin bu değişkenleri doğru bir şekilde tahmin edebildiği anlamına gelmektedir. Ancak, rüzgar hızı değişkeni için modelin tahmin performansı nispeten daha düşüktür. Bunun nedeni, rüzgar hızının diğer değişkenlere göre daha fazla anlık değişkenliğe sahip olmasıdır. Yani, rüzgar hızı hızlı bir şekilde ve düzensiz bir şekilde değişebildiği için modelin bu değişkeni tahmin etme kapasitesi daha sınırlı olmuştur.

Bu çalışma, çoklu çıkış Gauss regresyon algoritmasının farklı çevresel koşulları nasıl modelleyebildiğini göstermektedir. Modelin performansı, özellikle tuzluluk, iletkenlik ve deniz suyu sıcaklığı gibi değişkenler için oldukça başarılıdır. Grafiklerde görüldüğü gibi, model bu

değişkenleri yüksek bir doğrulukla tahmin edebilmiştir. Eğitim verileri ile gerçek ölçümlerin uyum içinde olması, modelin güvenilirliğini ve doğruluğunu arttırmaktadır.

Öte yandan, rüzgar hızı gibi daha fazla değişiklik gösteren faktörlerin modellenmesinde bazı zorluklarla karşılaşıldığı görülmektedir. Rüzgar hızı, kısa süre içinde büyük değişikliklere uğrayabilen bir meteorolojik parametre olduğundan, doğal bir şekilde modelleme süreci bu dinamik yapıyı tam anlamıyla yakalayamayabilir. Bu da sonuçlarda daha geniş bir hata payı olarak kendini göstermiştir. Ancak yine de bu durum, çalışmanın genel bulgularının ve modelin uygulanabilirliğinin önemini azaltmamaktadır.



Şekil 3.31. Su Parametrelerinin Çoklu Çıktılı Gauss Regresyon Algoritması Kullanılarak Tahmini Tuzluluk, b) Su Sıcaklığı, c) İletkenlik, d) Rüzgar Hızı

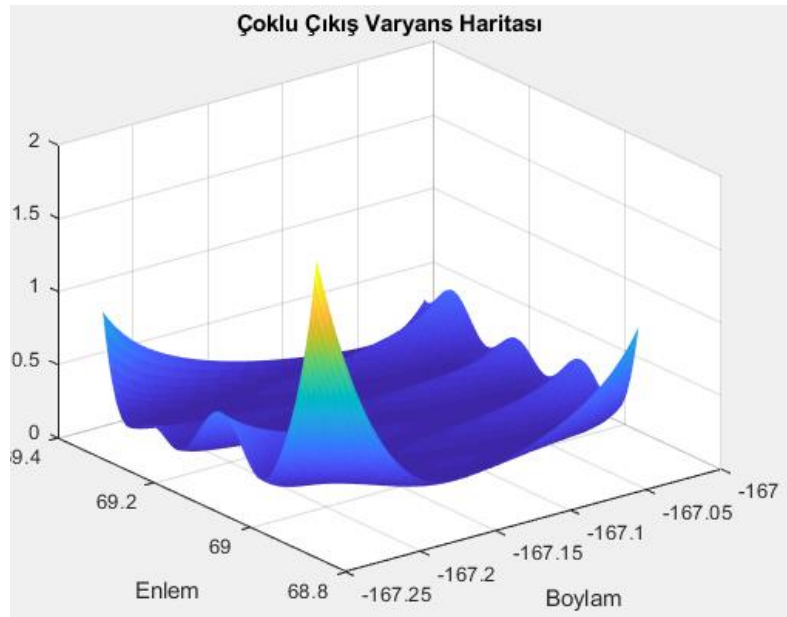
Çoklu Çıktılı Gauss Regresyon Algoritması, farklı değişkenlerin öngörülmesinde entegre bir kovaryans fonksiyonu kullanarak belirsizlik hesaplamalarını daha kapsamlı hale getirir. Bu entegre yapı sayesinde, modelin tahmin ettiği her bir çıktı için belirsizlikler birlikte değerlendirilir ve böylece tahminlerin doğruluğu daha etkin bir şekilde ölçülür. Şekil 3.32, çoklu çıktı Gauss regresyon algoritması kullanılarak hesaplanan varyans haritasını detaylı bir şekilde sunmaktadır.

Bu harita, modelin sunduğu tahminlerin hangi noktalarda ne kadar belirsizlik içerdiğini görsel olarak ifade eder. Grafik üzerinden, 8 farklı değişken çıktısının tahmin edilebilirliği üzerine yapılan analizde maksimum belirsizliğin yaklaşık olarak 2'ye ulaştığı gözlemlenmiştir. Bu maksimum belirsizlik değeri, modelin bazı bölgelerde değişken tahminlerinde oldukça güvenilir olduğunu, ancak belirli konumlarda tahmin belirsizliğinin arttığını göstermektedir.

Grafikte, x ve y eksenleri enlem ve boylam değerlerini temsil ederken, z eksenini ise hesaplanan varyans değerlerini göstermektedir. Varyansın yüksek olduğu bölgeler, modelin o bölgelerde daha fazla belirsizliğe sahip olduğunu ve tahminlerin daha az kesin olduğunu belirtir. Özellikle belirli bir enlem ve boylam kombinasyonunda varyansın zirve yaptığı, yani belirsizliğin maksimum değerlere ulaştığı görülmektedir. Bu zirve, tahmin edilen 8 çıkışın ortak belirsizlik noktalarını temsil eder ve çoklu çıktı Gauss regresyonunun ne kadar hassas olduğunu vurgular.

Varyans haritası, modelin doğruluk ve güvenilirlik analizi için kritik bir rol oynar. Harita, duyarlı bölgeleri ve dolayısıyla modelin performansının düşük olduğu alanları belirlemeye yardımcı olur. Bu bilgi, gelecekteki model geliştirme ve iyileştirme çalışmalarında yönlendirici olabilir. Belirsizliklerin minimize edilmesi, modelin genel doğruluğunu artırarak, tahminlerin daha güvenilir hale gelmesini sağlar.

Çoklu çıktı regresyon algoritması, gerçek dünyadaki karmaşık veri setleri ile çalışırken büyük avantajlar sunar. Özellikle entegre kovaryans fonksiyonları sayesinde, modelin tüm çıktı değişkenlerinde aynı ölçüde belirsizlik hesaplaması sağlanır. Bu, modelin genelleme kapasitesini artırarak çeşitli çıktılarının tahmininde tutarlılık sağlar. Şekil 3.32'de gösterilen varyans haritası, bu belirsizliklerin mekansal dağılımını ve değişkenlerin tahmin edilebilirliklerini görselleştirir.



Şekil 3.32. Çoklu Çıktılı Gauss Regresyon Kullanılarak Tahmin Edilen Varyans Haritası

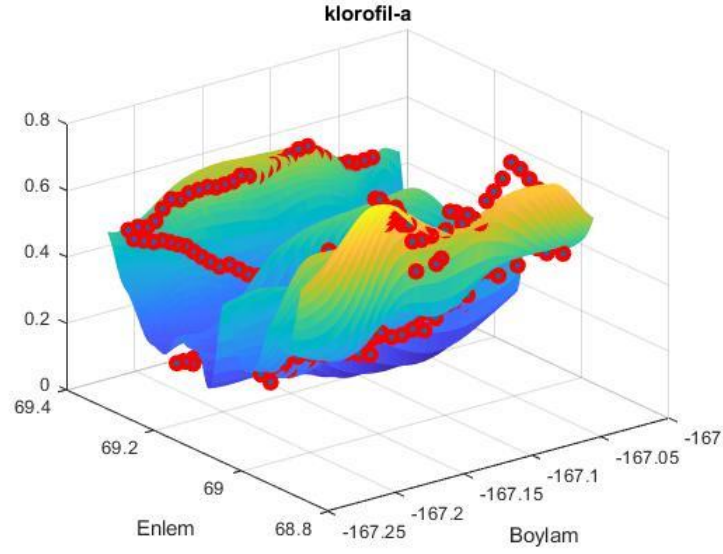
Şekil 3.33.a'de tahmin edilen su sıcaklığı, tuzluluk, iletkenlik ve rüzgar hızı parametrelerini vermektedir. Şekil 3.33.b'de ise bu parametreler kullanılarak tahmin edilen denizdeki klorofil miktarını varyans haritası verilmektedir. Varyans haritasındaki renkli yüzey, farklı enlem ve boylamlar boyunca varyans değerlerini göstermektedir ve maksimum varyans değeri 0.03 olarak belirlenmiştir. Bu varyans değeri, tahminlerin genel olarak güvenilir ve tutarlı olduğunu, fakat bazı bölgelerde küçük varyasyonlar olabileceğini göstermektedir.

Daha ayrıntılı olarak incelendiğinde, Şekil 3.33'de klorofil-a miktarının enlemler ve boylamlar boyunca nasıl değiştiği gözlemlenebilmektedir. Kırmızı noktalar, ölçülen verilerin topografyasını ve yüzeydeki değişiklikleri açıkça göstermektedir. Bu veriler, su sıcaklığı, tuzluluk, iletkenlik ve rüzgar hızının klorofil miktarını doğrudan etkileyen önemli faktörler olduğunu ortaya koymaktadır.

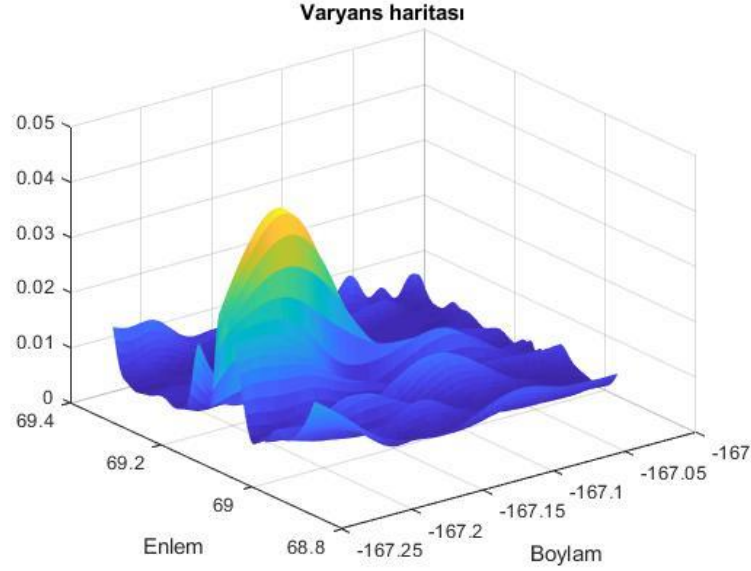
Şekil 3.33.b'deki varyans haritası ise, tahmin edilen klorofil değerlerinin ne kadar kararlı ve güvenilir olduğunu görselleştirerek modelin performansını değerlendirmektedir. Bu harita, belirli bir bölgede yapılan tahminlerin sapmalarını ve olası belirsizlikleri göstermektedir. Tahmin edilen varyansın maksimum değeri 0.03 olarak gözlemlenmiştir ki bu oldukça düşük bir değerdir. Düşük varyans değeri, kullanılan modelin tahminlerde yüksek doğruluk ve güvenilirlik sağladığını ve dolayısıyla güvenilir sonuçlar elde edildiğini gösterir.

Bu bulgular, su ekosistemlerinin durumunu izlemek ve yönetmek için kullanılan modelin güvenilirliğini artırmaktadır. Klorofil-a, su kütlelerindeki fitoplankton biyokütlesinin önemli bir göstergesi olduğundan, bu tür tahminlerin doğru yapılması, deniz bilimcileri ve çevre yönetimi ile ilgilenen profesyoneller için oldukça değerlidir. Dolayısıyla, su sıcaklığı, tuzluluk, iletkenlik ve rüzgar hızı gibi fiziksel parametrelerin ölçümü ve bu verilerin doğru analiz edilmesi, su kalitesinin izlenmesi ve deniz ekosistemlerinin sağlığına yönelik kararların alınmasında hayati rol oynamaktadır.

Genel olarak, bu modellerin ve haritaların kullanımı, deniz yüzeyi gözlemlerine dayalı klorofil tahminlerinin yapılmasını ve bu tahminlerin doğruluğunun sürekli olarak değerlendirilmesini sağlar. Bu da hem çevresel araştırmaların ilerlemesine katkıda bulunur hem de çevre yönetimi stratejilerinin daha etkili bir şekilde uygulanmasını olanaklı kılar.



(a)



(b)

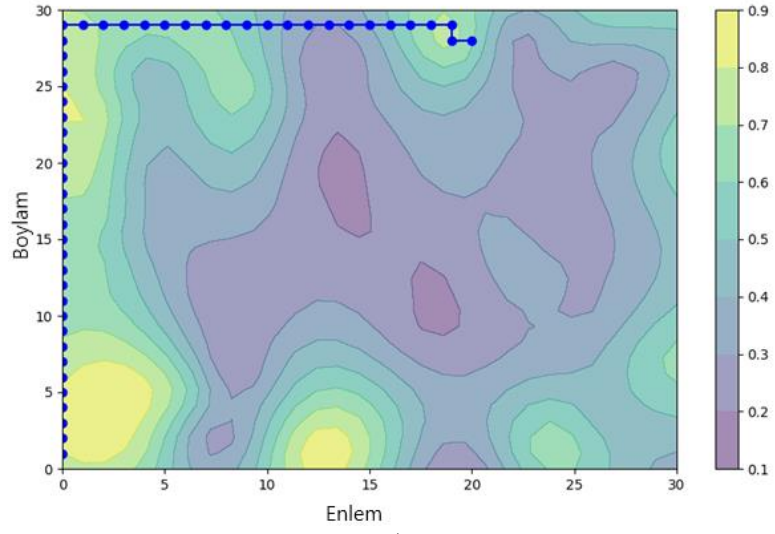
Şekil 3.33. a) Tahmin Edilen Klorofil Haritası b) Tahmin Edilen Varyans Haritası

3.3.3. MCTS-BnB Algoritmasının Başarımının Değerlendirilmesi

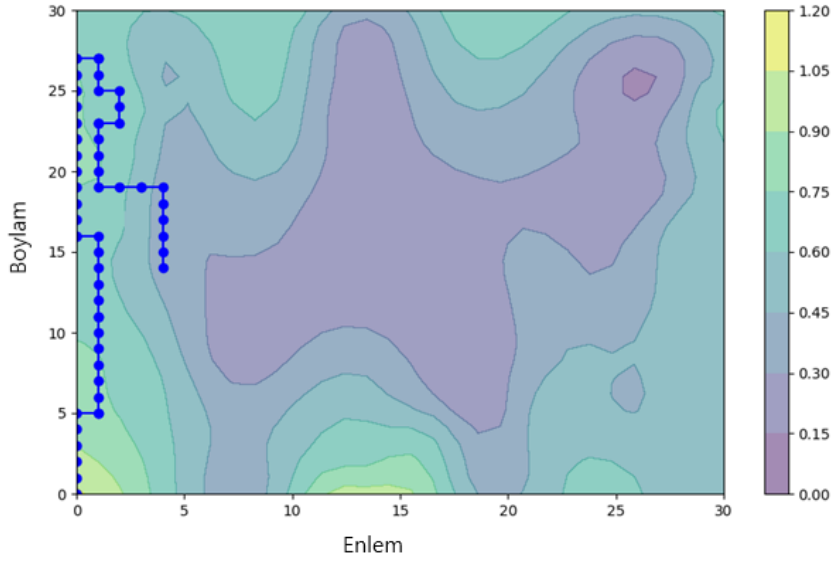
Adaptif güzergah planlama algoritmalarında, öğrenme süreci sürekli devam eder ve zamanla algoritmanın performansı gelişir. Algoritmanın başlangıçta karşılaştığı temel sorun, yeni örneklerin hangi konumdan toplanacağı ve bu konumlara giden en iyi rotayı bulmaktır. Bu sorun, genellikle gezgin satıcı problemine (TSP) benzer bir yapıya sahiptir. Gezgin satıcı problemi, belirli bir dizi şehir (veya durak) arasında en kısa toplam mesafeyi kapsayan ideal rotayı bulmayı amaçlayan klasik bir optimizasyon problemidir. Bu bağlamda, adaptif güzergah planlama algoritmalarında

kullanılan iki önemli yöntem olan MCTS ve ESAS algoritmaları sezgisel ve deterministik yöntemler olarak karşımıza çıkmaktadır. Bölüm 2.4'de bu algoritmaların teorik altyapısına ve işleyişlerine dair detaylı bilgiler sunulmuştur. MCTS algoritması, karar verme süreçlerinde rastgele örneklem almayı ve bu örneklem üzerinde olasılık bazlı analizler yaparak en iyi kararı vermeyi amaçlar. ESAS algoritması ise olası tüm yollar için arama ağacı oluşturarak maksimum ödüle sahip rotayı seçmeyi amaçlar. Bu sezgisel ve deterministik yöntemler, Bölüm 3.3.2'de anlatılan gerçek bir veri seti üzerinde test edilmiştir. Ayrıca, MCTS algoritmasının bir başka varyantı olan MCTS-BnB yöntemi de Bölüm 2.4.3'de ele alınmış ve aynı veri setine uygulanmıştır. MCTS-BnB, geleneksel MCTS yöntemine ek olarak, çözümleri daha verimli bir şekilde kesebilen ve daha hızlı optimal ya da alt-optimal çözümler üretebilen BnB yöntemini içerir.

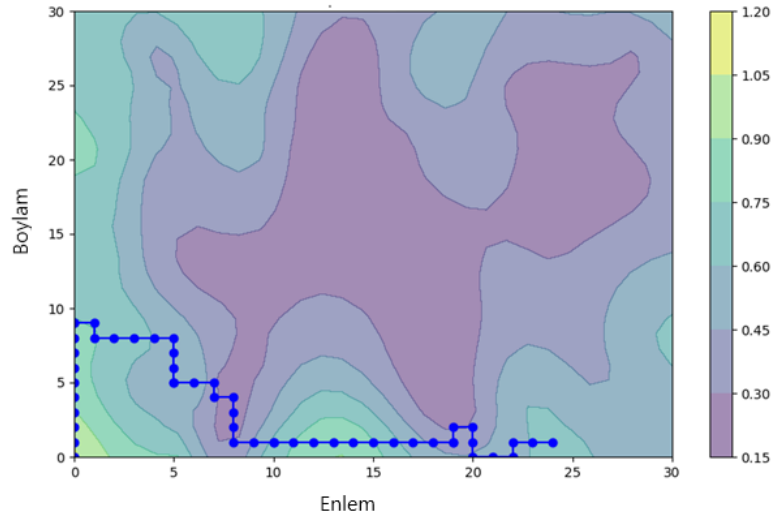
Şekil 3.34'de BnB-MCTS, MCTS ve ESAS algoritmalarıyla oluşturulan 50 adımlık yollar incelendiğinde, MCTS algoritmasının belirli bir davranış yolu sergilediği gözlemlenmektedir. ESAS algoritmasının performansı, başlangıçta nispeten düşük hata oranları sergileyerek MCTS'ye göre daha istikrarlı bir başlangıç yapmaktadır. ESAS algoritması, düşük maliyetli bölgelerde dolaşarak seyahat maliyetini minimize etmeye çalışır. Ancak bu süreçte alanın tamamını taramaya gereksinim duyması, algoritmanın öğrenme kapasitesini artırıp, daha güvenilir rotalar oluşturmasına yardımcı olur. MCTS-BnB algoritması daha stratejik ve detaylı bir yönlendirme sağlar. MCTS-BnB algoritmasında, x ekseninde belirli kıvrımlar olduğu gözlemlenmektedir. Bu kıvrımlar, algoritmanın alanın farklı kısımlarını daha iyi öğrenmek için daha fazla esneklik tanıdığı anlamına gelir. MCTS-BnB algoritması, rota oluştururken hem düşük seyahat maliyetini hedeflerken hem de alanı daha iyi öğrenmeyi amaçlamaktadır. Bu algoritmanın rotaları, özellikle y ekseninde görülen kıvrımlarla, alanın çeşitli yerlerine daha dengeli bir şekilde dağılmaktadır. Y ekseninde görülen bu kıvrımlar, algoritmanın hem kısa vadeli hem de uzun vadeli optimizasyon hedeflerini göz önünde bulundurarak daha ayrıntılı ve stratejik bir planlama yapmasını sağlamaktadır. Bu kıvrımlar ve rota stratejileri sayesinde MCTS-BnB, diğer algoritmalara kıyasla daha iyi bir performans sergilemektedir.



a)



b)



c)

Şekil 3.34. Tahmin Edilen Klorofil Haritası a) MCTS b) MCTS-BnB c) ESAS

3.3.4. MCTS-BnB Algoritmasının Başarımının Farklı Veri Setlerinde Değerlendirilmesi

Bu çalışma kapsamında, algoritmaların başarımını değerlendirmek amacıyla Bölüm 3.3.2’de anlatılan deneysel veri setine ($\ell=9$) ek olarak NASA tarafından sağlanan uydudan alınan veriler bilgileri doğrultusunda oluşturulan Bölüm 3.1.1.’de anlatılan sıcaklık veri setleride ($\ell=1$ ve $\ell=3$) kullanılarak kapsamlı testler gerçekleştirilmiştir. Sıcaklık veri setleri içerisinde dikkate alınan ℓ değeri, veri setinin iç dinamiklerini ve verilerin birbirleriyle olan ilişkisini anlamada kritik bir öneme sahiptir. Yüksek ℓ değerleri, veri noktalarının arasındaki ilişkinin yüksek olduğunu ve ani değişimlerin nadiren görüldüğünü gösterirken, düşük ℓ değerleri ani değişimlerin sıkça yaşandığı bir veri seti profili sunmaktadır. Bu bağlamda, üç farklı ℓ değeri üzerinden model performanslarının incelenmesi, algoritmaların çeşitli veri senaryolarındaki adaptif kapasitesini ortaya koymaktadır.

Şekil 3.35, farklı hiper parametre değerlerine sahip üç modelin (MCTS, ESAS ve MCTS-BnB) RMSE performansını göstermektedir. Şekil 3.35.a’da, $\ell =1$ değeri için performanslar karşılaştırılmıştır. MCTS modeli başlangıçta yüksek bir RMSE değeri ile başlamakta ve zamanla hafif dalgalanmalar yaşasa da yaklaşık 0.014 seviyesinde sabitlenmektedir. ESAS modeli ise başlangıçta biraz daha düşük bir RMSE değeri ile başlamış ve kademeli olarak yaklaşık 0.0125 seviyesine düşmüştür. MCTS-BnB modeli aynı şekilde başlangıçta 0.014 seviyesinde olup, zamanla en düşük RMSE değeri olan yaklaşık 0.012’ye kadar inmiştir. Bu durum, MCTS-BnB modelinin bu hiper parametre değerinde daha başarılı olduğunu göstermektedir.

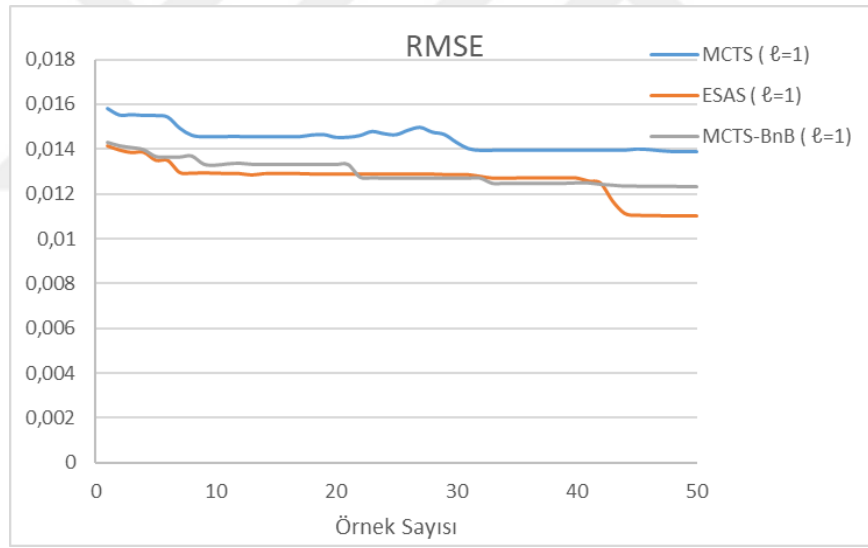
İkinci grafikte, $\ell =3$ değeri için performanslar incelenmiştir. Burada MCTS modeli, başlangıçta oldukça yüksek bir RMSE değeriyle başlamakta ve ani bir düşüşle 0.0014 seviyesine sabitlenmektedir. ESAS modeli daha düşük bir başlangıç RMSE değeriyle başlamış ve yaklaşık 0.0008 seviyesine kadar azalmıştır. MCTS-BnB modeli ise, yaklaşık 40. örnekte ani bir düşüş yaşayarak ESAS ile benzer düşük seviyeye ulaşmıştır. Bu durumda ESAS ve MCTS-BnB modelleri, MCTS modeline göre daha başarılı performans göstermektedir.

Üçüncü grafikte, $\ell =9$ değeri için tüm modellerin performansı incelenmiştir. MCTS modeli, başlangıçta yüksek bir RMSE değeri sergilemiş ve zamanla bu değeri yaklaşık 0.065 seviyesine sabitlemiştir. ESAS modeli başlangıçta daha düşük bir değer ile başlamış, ancak 20. örnek civarında ani bir zirve yaparak daha sonra istikrarlı bir şekilde 0.055 seviyesine düşmüştür. MCTS-BnB modeli ise, başlangıçta nispeten düşük bir RMSE değeri göstermiş ve kademeli olarak yaklaşık 0.055 seviyesinde sabitlenmiştir. Bu grafikte, ESAS ve MCTS-BnB modelleri genel olarak daha düşük ve stabil RMSE değerleri sergileyerek öne çıkmaktadır.

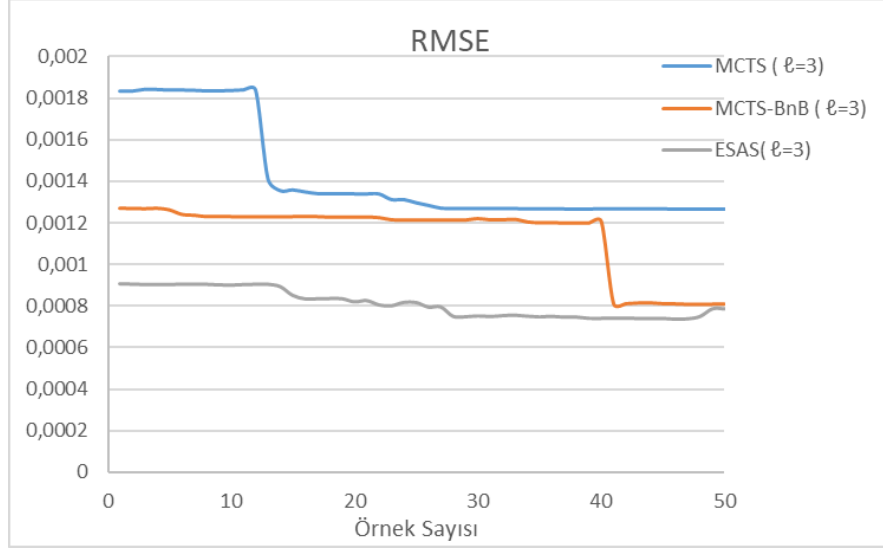
Bu analizler, adaptif güzergah planlama algoritmalarının başarımını değerlendirirken düşük RMSE değerlerini elde etme kapasitelerinin kritik öneme sahip olduğunu ortaya koymaktadır. İlk grafik, MCTS-BnB’nin ve ESAS algoritmalarının $\ell =1$ parametresi ile en düşük RMSE değerine sahip olduğunu ve dolayısıyla daha iyi bir performans sergilediğini göstermektedir.

İkinci grafik ise, $\ell = 3$ parametresi kullanıldığında hem ESAS hem de MCTS-BnB modellerinin MCTS modeline göre daha düşük RMSE değerlerine ulaştığını ve daha başarılı performanslar sergilediğini ortaya koymaktadır. Üçüncü grafik ise, $\ell = 9$ parametresi ile ESAS ve MCTS-BnB'nin yine MCTS modeline göre daha düşük ve stabil RMSE değerlerine sahip olduğunu göstermektedir.

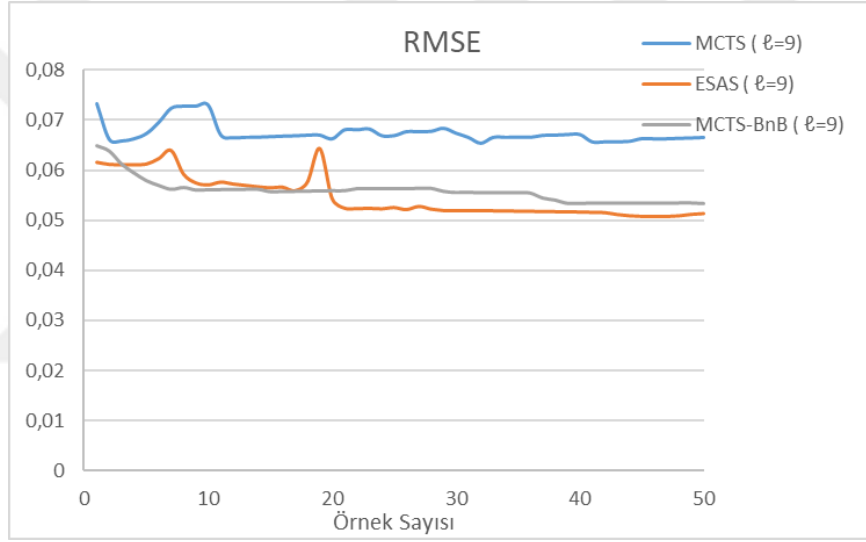
Bu durum, ESAS ve MCTS-BnB modellerinin, özellikle düşük hata oranlarının önemli olduğu veri yoğun ve hassas uygulamalarda tercih edilebileceğini göstermektedir. MCTS modeli, başlangıçta daha yüksek hata oranları sergileyebilse dahi, belirli kullanım senaryolarında stabil performans sağlayabilir. Ancak genel olarak, ESAS ve MCTS-BnB modellerinin daha geniş bir uygulama yelpazesinde etkin bir şekilde kullanılabilmesi ve adaptif güzergah planlama algoritmalarında daha güvenilir sonuçlar üretebileceği görülmektedir. Bu nedenle, özellikle gerçek veri setleri üzerinde yapılan testlerde, bu iki modelin sağladığı düşük RMSE değerleri, onların adaptif güzergah planlama alanında önemli bir yere sahip olduğunu pekiştirmektedir.



(a)



(b)



(c)

Şekil 3.35. Farklı hiper parametrelere sahip 3 veri setinde RMSE değerinin değişimi a) $\ell = 1$ b) $\ell = 3$ c) $\ell = 9$

Şekil 3.36'daki grafikler, üç farklı adaptif güzergah planlama algoritmasının (MCTS, ESAS, MCTS-BnB) entropi değerlerinin örnek sayısına bağlı olarak nasıl değiştiğini göstermektedir. Her bir grafik farklı " ℓ " değerleri ($\ell = 1$, $\ell = 3$, $\ell = 9$) için entropi değişimlerini sergilemektedir.

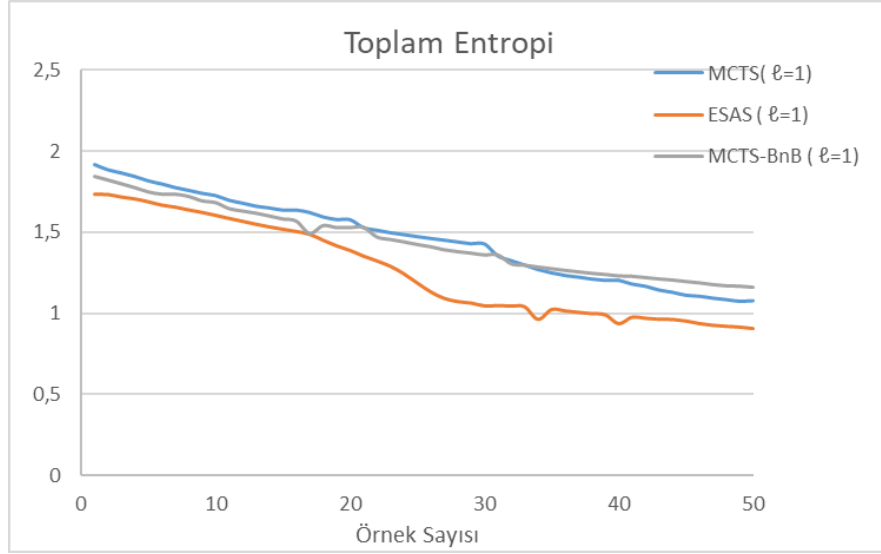
İlk grafikte ($\ell = 1$), MCTS algoritması başlangıçta yaklaşık 2.2 entropi değeri ile başlamakta ve zamanla belirgin dalgalanmalar göstererek azalma eğilimi sergilemektedir. Ancak bu düşüş oldukça yavaş olmaktadır ve süreç boyunca yüksek entropi değerleri korunduğu gözlemlenmektedir.

ESAS algoritması ise başlangıçta yaklaşık 1.6 entropi değeri ile başlamakta ve hızlı bir düşüş göstererek 10. örnekten itibaren daha düşük ve stabil bir seviyeye ulaşmaktadır. MCTS-BnB algoritması ise kademeli bir düşüş göstermekte ve 20. örnekten itibaren stabilize olarak yaklaşık 0,85 entropi değeri seviyesine ulaşmaktadır.

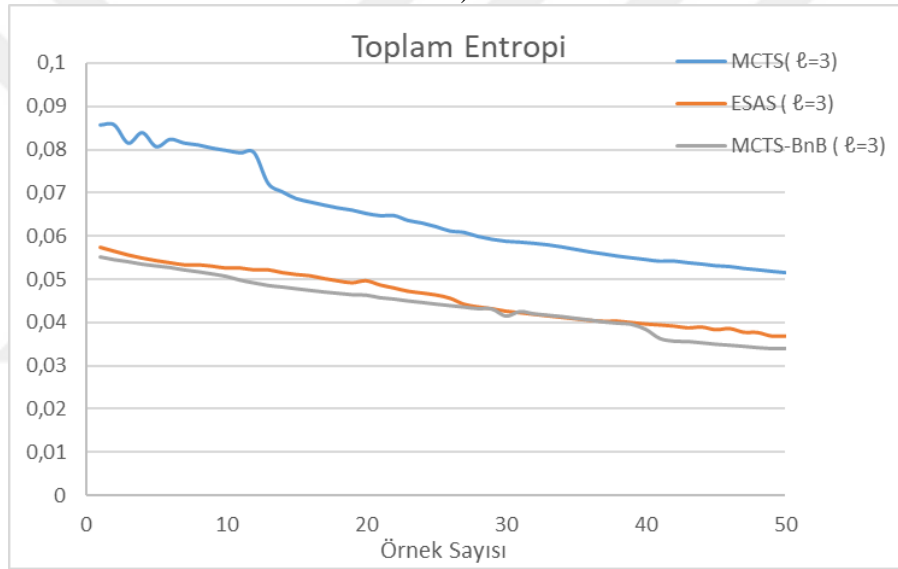
İkinci grafikte ($\ell = 3$), MCTS algoritması başlangıçta yaklaşık 0,09 entropi değeri ile başlamakta ve zamanla azalmakta ancak diğer algoritmalara göre daha yüksek kalmaktadır. ESAS ve MCTS-BnB algoritmaları başlangıçta sırasıyla yaklaşık 0,057 ve 0,054 değerleri göstererek hızlı bir düşüş sergilemekte ve 20. örnekten itibaren yaklaşık 0.04 civarında stabilize olmaktadır. Bu durum, MCTS algoritmasının yine daha yavaş bir iyileşme süreci sergilediğini göstermektedir.

Üçüncü grafikte ($\ell = 9$), MCTS algoritması başlangıçta yaklaşık 22 entropi değeri ile başlayarak süreç boyunca dalgalanmalar göstermekte ve genel olarak yüksek değerler korumaktadır. ESAS algoritması yaklaşık 18 değerinden başlayarak hızlı bir düşüş göstermekte ve 10. örnekten itibaren yaklaşık 15 seviyesinde stabilize olmaktadır. MCTS-BnB algoritması ise süreç boyunca kademeli bir düşüş göstererek benzer şekilde 15 seviyesinde stabilize olmaktadır ve genel olarak ESAS ile benzer performans sergilemektedir.

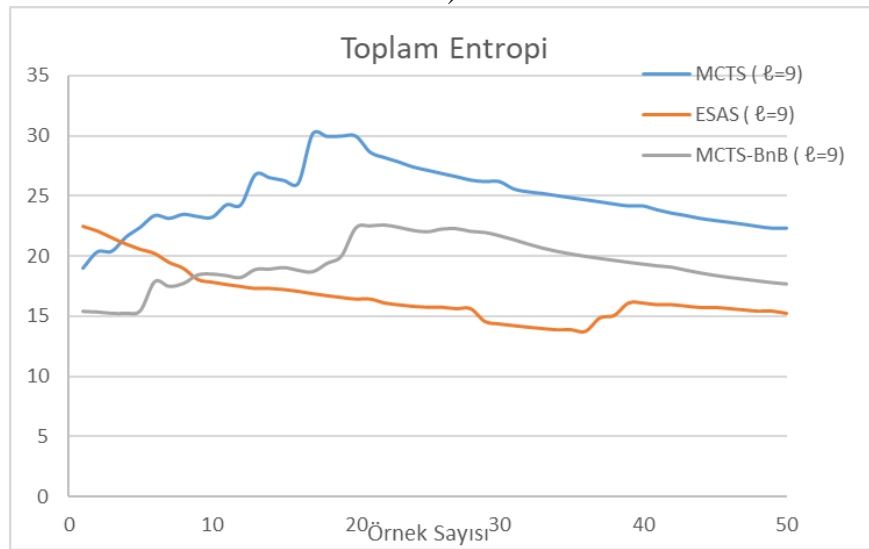
Genel olarak, MCTS algoritmasının başlangıçta ve süreç boyunca daha yüksek entropi değerlerine sahip olduğu ve daha yavaş bir azalma eğilimi gösterdiği açıkça görülmektedir. Bu, MCTS'in entropiyi azaltma konusunda daha az etkili olduğunu ve diğer algoritmalara kıyasla daha uzun bir iyileşme süreci gerektirdiğini göstermektedir. Buna karşılık, ESAS ve MCTS-BnB algoritmaları, hızlı bir başlangıç düşüşü sergileyerek düşük ve stabil entropi seviyelerine daha çabuk ulaşmaktadır. Özellikle, ESAS algoritması başlangıçtaki hızlı düşüş ile dikkat çekmekte ve düşük seviyelerde stabil kalmaktadır. MCTS-BnB algoritması ise daha kontrollü ve kademeli bir düşüş göstererek performansını optimize etmektedir.



a)



b)



c)

Şekil 3.36. Toplam Entropi Değişimi

Şekil 3.37, üç farklı adaptif güzergah planlama algoritmasının (MCTS, ESAS ve MCTS-BnB) zaman performansını karşılaştırmaktadır. Grafikte, X ekseninde artan örnek sayısı ve Y ekseninde bu işlemler için harcanan zaman yer almaktadır ve logaritmik bir ölçek kullanılarak algoritmaların zaman performansındaki eğilimler detaylı bir şekilde incelenmektedir.

MCTS algoritması, örnek sayısı arttıkça zaman harcamasında nispeten ılımlı bir artış sergileyen bir eğri ortaya koymaktadır. $\ell = 1$ için MCTS, başlangıçta yaklaşık 12 dakika harcamakta ve örnek sayısı arttıkça bu süre hafif bir artış göstermektedir. $\ell = 3$ için zaman harcaması yaklaşık 25 dakikadan başlamakta ve sabit veya hafif bir artışla devam etmektedir. $\ell = 9$ için ise başlangıçta yaklaşık 45 dakika olan zaman harcaması, örnek sayısı arttıkça daha belirgin bir artış göstermektedir, ancak genel olarak ılımlı bir artış eğilimi sergilemektedir.

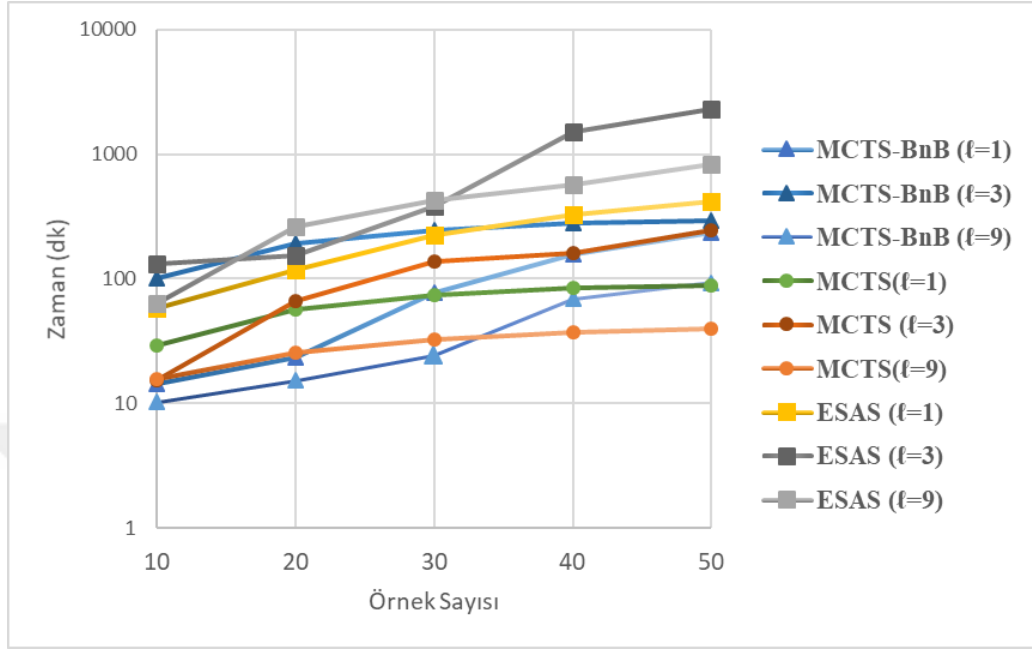
ESAS algoritması, örnek sayısı arttıkça zaman harcamasında belirgin bir artış göstermekte ve büyük veri setlerinde maliyetli hale gelmektedir. $\ell = 1$ için ESAS algoritması başlangıçta yaklaşık 10 dakika harcarken, 50 örnek noktasına ulaşırken bu süre yaklaşık 180 dakikaya çıkmaktadır. $\ell = 3$ için ise zaman harcaması başlangıçta yaklaşık 12 dakikadan 50 örnek noktasında yaklaşık 5000 dakikaya ulaşarak eksponansiyel bir artış sergilemektedir. $\ell = 9$ için ise başlangıçta yaklaşık 20 dakika olan zaman harcaması, 50 örnek noktasında yaklaşık 9000 dakikaya ulaşmaktadır, bu da ESAS algoritmasının zaman açısından oldukça maliyetli olduğunu göstermektedir.

MCTS-BnB algoritması, MCTS algoritmasına benzer bir eğri göstermekte ancak genel olarak her iki algoritmadan da daha düşük zaman harcamaktadır. $\ell = 1$ için MCTS-BnB, başlangıçta yaklaşık 10 dakika harcayarak 50 örnek noktasına kadar nispeten düşük bir artış sergilemektedir. $\ell = 3$ ve $\ell = 9$ değerleri için de benzer şekilde düşük zaman harcamaları gözlemlenmekte, başlangıçta sırasıyla 10 ve 15 dakika civarında başlayan bu süreler, 50 örnek noktasında maksimum 300 dakikayı bulmaktadır. Bu durum, MCTS-BnB algoritmasının daha düşük zaman maliyetleri ile dikkat çektiğini ve bu nedenle daha etkili ve verimli olduğunu ortaya koymaktadır.

Genel olarak, MCTS ve MCTS-BnB algoritmaları, zaman açısından önemli avantajlar sunarken, ESAS algoritmasının büyük örnek sayılarında daha zaman maliyetli hale geldiği net bir şekilde gözlemlenmektedir. ESAS algoritması, düşük hata oranları ve yüksek doğruluk gerektiren uygulamalarda avantaj sağlasa da, operasyonel maliyetler dikkate alındığında MCTS ve MCTS-BnB algoritmaları daha verimli çözümler sunmaktadır. MCTS-BnB algoritmasının, ESAS algoritmasına yakın bir doğruluk sağlarken, zaman açısından MCTS algoritmasına benzer performans göstermesi, adaptif güzergah planlama ve çevresel izleme problemlerinde hem yüksek doğruluk hem de zaman verimliliği açısından başarılı olacağını doğrulamaktadır.

Bu bulgular, adaptif güzergah planlama sistemlerinin zaman performansındaki farklılıkları net bir şekilde ortaya koymakta ve pratik uygulamalarda MCTS ve MCTS-BnB algoritmalarının daha uygun olduğu sonucunu desteklemektedir. Grafik, MCTS-BnB'nin özellikle zaman verimliliği

konusunda üstünlüğünü gösterirken, ESAS'ın yüksek doğrulukla birlikte gelen zaman maliyetlerinin değerlendirilmesi gerektiğini vurgulamaktadır.



Şekil 3.37. Zamana bağlı toplanan örnek miktarı

Tablo 3.5 ve Şekil 3.37, MCTS, ESAS ve MCTS-BnB olmak üzere üç farklı adaptif güzergah planlama algoritmasının zaman performansını karşılaştırmaktadır. Tabloda farklı örnek sayıları için algoritmaların harcadığı zaman detaylı olarak incelenirken, Şekil ise zaman performansındaki eğilimleri logaritmik bir ölçekte görselleştirmektedir. Buna göre, MCTS-BnB algoritması genel olarak deterministik bir yaklaşım olmasına rağmen $\ell=1$, $\ell=3$ ve $\ell=9$ olan data setlerinde, ESAS algoritmasına oranla sırasıyla %43,38, %87,30 ve %88,65 oranlarında daha düşük hesaplama süresi sunmaktadır. MCTS-BnB algoritması zaman tüketimi açısından MCTS algoritmasına göre sırasıyla %38, %85 ve %42 daha yakın sonuçlar vermektedir.

Şekil 3.35'de ise, elde edilen sonuçların ESAS algoritmasına yakın bir optimallik düzeyinde olduğu görülmektedir. Bu bulgular, adaptif güzergah planlama sistemlerinde zaman verimliliği açısından MCTS-BnB algoritmasının tercih edilebilirliğini desteklemektedir. Dolayısıyla, hızlı ve etkili sonuçlar elde etmek istendiğinde MCTS-BnB algoritması adaptif güzergah planlama veya benzer uygulamalarda tercih edilebilir bir seçenek olabilir.

Tablo 3.5. Örnek miktarına bağlı harcananan zaman (dk)

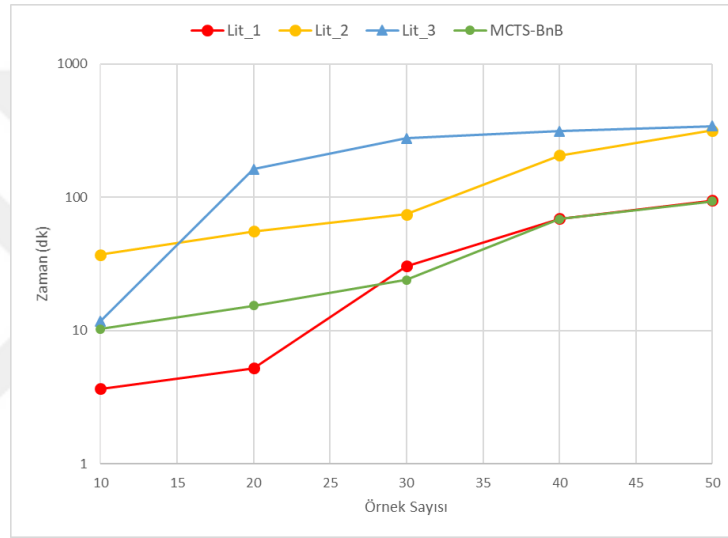
Örnek Sayısı	MCTS-BnB			MCTS			ESAS		
	$\ell=1$	$\ell=3$	$\ell=9$	$\ell=1$	$\ell=3$	$\ell=9$	$\ell=1$	$\ell=3$	$\ell=9$
10	14,48	100,09	10,25	29,20	15,23	15,72	58,24	131,62	63,25
20	23,51	193,07	15,30	57,19	65,85	25,90	117,69	154,02	258,85
30	77,4	247,23	23,91	73,17	136,88	32,68	222,45	380,65	423,13
40	156,23	280,3	58,60	84,60	160,53	37,47	327,79	1500,75	566,05
50	235,38	289,38	93,09	88,02	245,65	39,63	415,72	2276,92	820,56

Şekil 3.38 ve Şekil 3.39, MCTS-BnB algoritmasının literatürdeki diğer BnB algoritmalarıyla karşılaştırılmasını içermektedir. Binney ve Sukhatme [46], kendi çalışmalarında alt limit belirlemek yerine sadece üst limit belirlemişlerdir. Bu yöntemde, robotun seyahat bütçesi ve ziyaret edebileceği düğümler üzerinden bir üst limit tanımlanmış ve bütçenin altında kalan düğümler elenmiştir. McMahan vd. [79] önerilen yaklaşık çözümü, gerçek bir çözümle karşılaştırmıştır. Gerçek çözümde, alt limit olarak çim biçme rotası, üst limit olarak ise robotun seyahat bütçesinde bir hücreden elde edilebilecek maksimum risk azaltımının, robotun belirli bir bütçede ziyaret edebileceği hücre sayısı ile çarpılması ile hesaplanmıştır. Çalışmada ödül fonksiyonu olarak maksimum risk azaltımı kullanılmıştır. Araştırmacılar, önerilen yöntemi denizaltında bilinmeyen sayıda hedefi bulmak amacıyla kullanmış ve ödül fonksiyonu olarak bu hedefleri tespit etmeye yönelik bir fonksiyon benimsemişlerdir. Bu çalışmada, alanın tamamını öğrenmek ve alanın tümünün klorofil haritasındaki belirsizliği azaltmak amaçlanmıştır. Bu nedenle, entropi ödül fonksiyonu kullanılarak belirlenen yöntem ile BnB algoritması mevcut veri setine entropi ödül fonksiyonu ile uygulanmıştır. Best ve Fitch [80], alt limit olarak açgözlü bilgilendirici yol planlama algoritmasını kullanmışlar ve üst limit olarak ise robotun seyahat bütçesine göre toplayabileceği maksimum ödül olarak tanımlamışlardır.

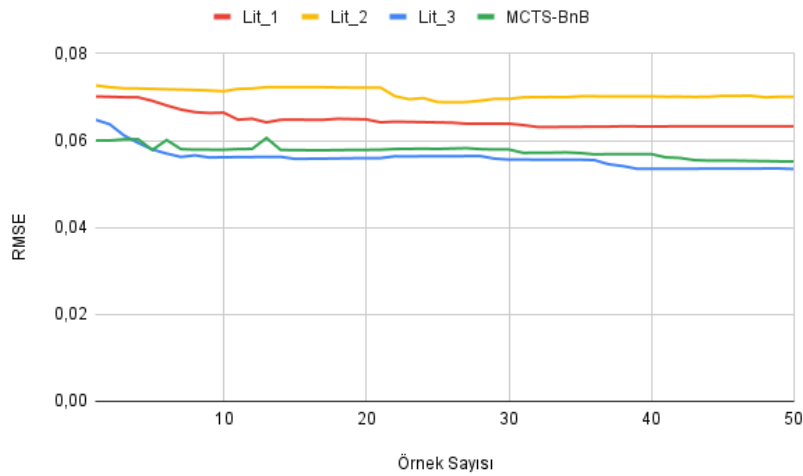
Bu çalışmada, [46,80] çalışmalarında önerilen alt ve üst limit hesaplama yöntemleri ve [79] çalışmasında kullanılan gerçek çözüm yöntemi, önerilen MCTS-BnB algoritmasının başarımını değerlendirmek için deneysel veri setine uygulanmıştır. Şekil 3.38 ve Şekil 3.39'da Lit_1 olarak isimlendirilen grafik [46] tarafından önerilen BnB yöntemiyle elde edilen sonuçları, Lit_3 [79]'da kullanılan gerçek çözüm yöntemiyle elde edilen sonuçları ve Lit_1 ise [80]'de önerilen yöntem ile elde edilen sonuçları vermektedir. Literatürdeki çalışmalar farklı ödül fonksiyonları kullanmış olsa da, bu çalışmada MCTS-BnB algoritmasının başarımını değerlendirmek için üç çalışmada da entropi ödül fonksiyonu kullanılmıştır.

Elde edilen sonuçlarda Lit_3 50 adımı en uzun zaman diliminde tamamladığı görülmüştür. Lit_2 ise 30 adımı Lit_3'e göre 30 adımı daha hızlı sürede tamamlamasına rağmen 50 adımı Lit_3 ile yaklaşık aynı sürede hesaplamıştır. Bu alt limit ve üst limit değerinin ilk 30 adımda daha fazla

sınırlama yaparken 50 adımda daha uzun sürede hesaplama yapıldığını göstermektedir. Fakat Şekil 3.39’da en yüksek RMSE oranının Lit_2 ile elde edildiği görülmüştür. Bu durum başlangıçta fazla sınırlama yapmasının optimal olabilecek düğümleride elemesinden kaynaklıdır. Lit_3 ise en uzun hesaplama süresine sahip olmasına rağmen en düşük RMSE oranına sahiptir. Lit_1’de ise sadece üst limitin hesaplanması hesaplama süresini büyük ölçüde düşürmüştür. Fakat RMSE oranı yüksek çıkmıştır. Önerilen MCTS-BnB algoritması ise başlangıçta Lit_1’e oranla daha yüksek bir hesaplama süresine sahip olmasına rağmen 50 adımı Lit_1’e yakın sürede hesaplamıştır. Fakat RMSE oranı Lit_3’de elde edilen en düşük RMSE oranına sahip bir değer elde edilmiştir. Bu durum sınırlama oranının artmasına rağmen optimal düğümün elimine edilmediğini göstermektedir.



Şekil 3.38. MCTS-BnB algoritmasının literatürdeki diğer çalışmalarla zaman performansı açısından karşılaştırılması



Şekil 3.39. MCTS-BnB algoritmasının literatürdeki diğer çalışmalarla RMSE performansı açısından karşılaştırılması

Tablo 3.6 literatürdeki çalışmalarla önerilen yöntemi karşılaştırmak için sunulmuştur. Genel olarak, MCTS-BnB algoritması hem zaman tüketimi hem de doğruluk (RMSE) açısından literatürdeki diğer BnB yöntemlerine kıyasla üstün bir performans sergilemektedir. Zaman performansında Lit_3 ([80]) ve Lit_2 ([79]) yöntemine göre çok daha hızlı olan MCTS-BnB, Lit_1 ([46])'e yakın bir hesaplama süresi elde edilmiştir. Doğruluk performansında ise, MCTS-BnB algoritmasının, daha düşük hesaplama süresine rağmen, en düşük RMSE oranına sahip Lit_3 ([80])'e yakın sonuçlar elde edilmiştir.. Bu da MCTS-BnB algoritmasının, optimal rotayı hızlı bir şekilde üretme kapasitesine sahip olduğunu göstermektedir.

Tablo 3.6. MCTS-BnB algoritmasının literatüdeki diğer BnB yöntemleriyle karşılaştırılması

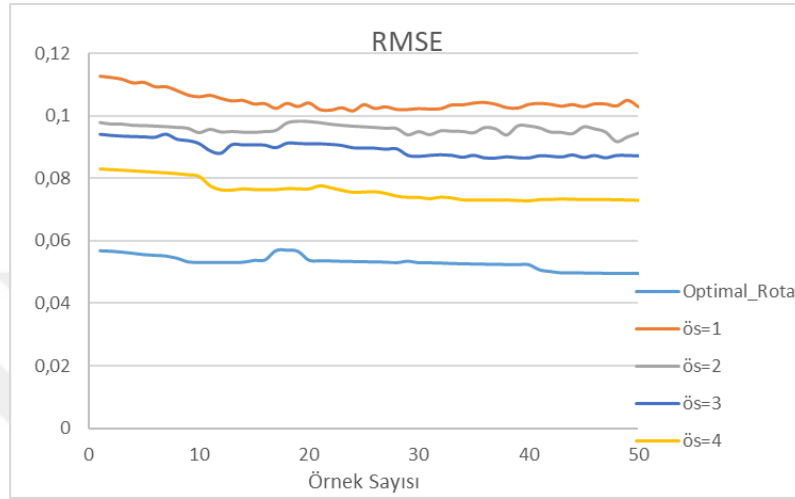
Örnek Sayısı	[46]		[79]		[80]		MCTS-BnB	
	Zaman (dk)	RMSE	Zaman (dk)	RMSE	Zaman (dk)	RMSE	Zaman (dk)	RMSE
10	3,65	0,0664	37,11	0,0713	11,77	0,0562	10,25	0,0578
20	5,22	0,0648	55,47	0,0721	162,57	0,0555	15,30	0,0578
30	30,47	0,0638	74,15	0,0695	276,92	0,0555	23,92	0,0578
40	68,84	0,0632	206,29	0,0701	312,62	0,0534	68,59	0,0568
50	94,32	0,0632	315,76	0,0700	340,20	0,0533	93,09	0,0552

3.3.5. Seyrek Sözde Girişleri Kullanarak MCTS-BnB Algoritması ile Adaptif Örnekleme

Adaptif örnekleme yaparken örneklem uzayının artması hesaplama yükünü önemli ölçüde değiştirmektedir. Fakat örneklem sayısının artışına bağlı olarak sistemin öğrenme oranı da artmaktadır. Bu çalışmada, seyrek sözde girişler kullanılarak optimal bir rotanın hesaplanması hedeflenmiştir. Örneklem uzayının genişlemesi, hesaplama yükünü arttırmakla birlikte, örneklem sayısının artışına paralel olarak sistemin öğrenme oranı da artmaktadır. Bir önceki bölümlerde, 100 eğitim verisi kullanılarak ortamın öğrenilmesi tamamlanmıştı. Bu çalışma kapsamında, 100 örnek arasından seyrek temsil eden girişler belirlenerek, optimale yakın bir rota bulunması amaçlanmıştır. Bölüm 2.5'te, seyrek sözde girişleri belirlemek için kullanılan kümeleme yöntemleri detaylı bir şekilde anlatılmıştır. Bu yöntemler, mevcut 100 eğitim verisine uygulanarak 1 ile 20 arasında farklı kümeleme yapılmıştır. Bu kümelemenin performansı ise Bölüm 2.6'da açıklanan SSI yöntemi ile ölçülmüştür. En yüksek SSI skoruna sahip olan kümeleme şekli tercih edilerek, alt örneklem kümesi oluşturulmuştur.

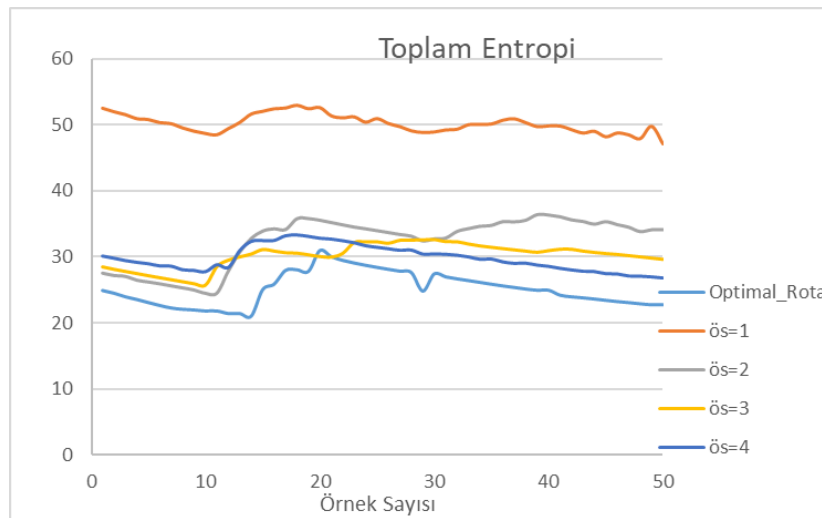
Oluşturulan bu kümelerden sırasıyla 1, 2, 3 ve 4 örnek seçilerek farklı alt örneklem kümeleri oluşturulmuştur. Şekil 3.40'da, bu alt örneklem kümelerinin optimale yakınlığını araştırmak

amacıyla RMSE değerlerindeki değişimler incelenmiştir. Bulunan sonuçlara göre, kümeden alınan eleman sayısı arttıkça, mevcut rotadan hesaplanan RMSE değeri, optimal rota üzerinde hesaplanan RMSE değerine yaklaşmıştır. Böylece, örneklem sayısının artışı ile elde edilen alt örneklem kümelerinin, daha optimal sonuçlara yaklaştığı ve hesaplama yükünü makul bir seviyede tutarak daha iyi öğrenme performansı sağladığı görülmüştür.



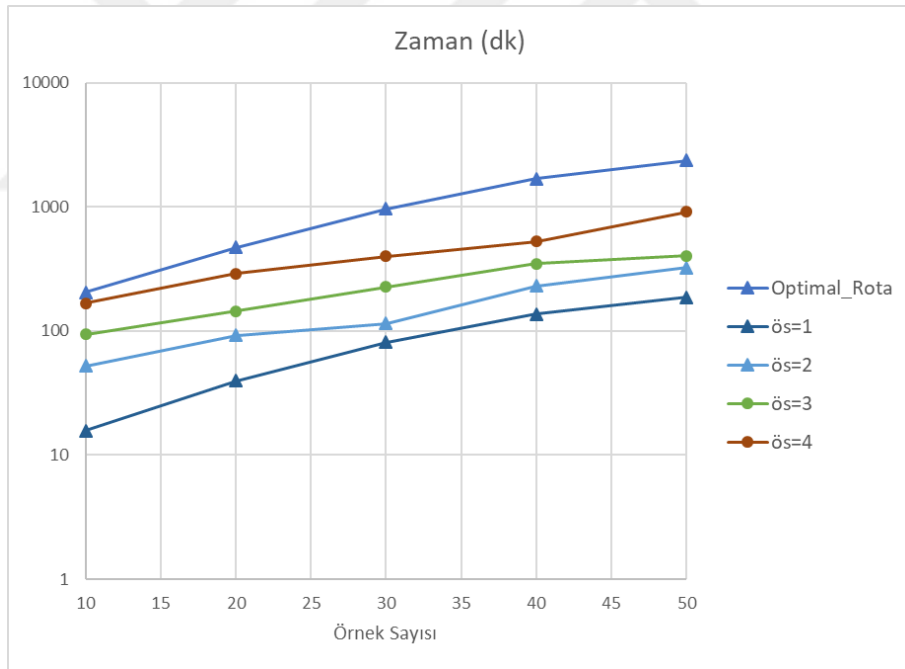
Şekil 3.40. Alt örneklem kümesinin RMSE değeri değişimi

Şekil 3.41’de ise entropi değeri değişimi verilmiştir. Toplam entropi değeri değişimi belirsizliğin bir ölçüsünü verdiğinden RMSE değerine benzer bir grafik elde edilmiştir. Kümelere alınan örnek sayısı arttıkça optimal rotaya olan yakınlıkta artmıştır.



Şekil 3.41. Alt örneklem kümesinin entropi değeri değişimi

Genel olarak, kümelerden alınan örnek sayısı arttıkça hesaplama gücü de eğitim veri sayısının artışına bağlı olarak artar. Gauss regresyon tahmin yapmak için kovaryans fonksiyonunu kullanır. Kovaryans fonksiyonu her bir eğitim verisinin diğer verilerle olan ilişkisini matematiksel olarak hesaplar. Bu durum hesaplama yoğunluğunu büyük ölçüde etkiler. Başlangıçta 30×30 'luk bir alanda 900 veriden 100'ü eğitim verisi olarak seçilmiştir. Şekil 3.42'de en iyi kümeleme SSI sonucu K-means kümeleme algoritmasıyla 4 küme için gerçekleştirilmiştir. Her bir kümeden sırasıyla 1,2,3 ve 4 örnek alınarak eğitim veri sayısı 4, 8, 12, 16 olarak azaltılmıştır. Bu durumda Gauss regresyon algoritması ile 100×100 'lük bir kovaryans fonksiyonu hesaplamak yerine 4×4 , 8×8 , 12×12 ve 16×16 'lık bir kovaryans fonksiyonu hesaplanır. Eğitim veri sayısının bu derece düşürülmesi hesaplama süresini önemli ölçüde azaltır. Örneğin, 4 eğitim verisi kullanılarak 50 örnek toplandığında 186,92 dakika sürerken, 8, 12 ve 16 eğitim verisi kullanıldığında sırasıyla 322,84, 401,68 ve 909,25 dakika sürer. 100 eğitim verisi kullanılarak başlangıçta 50 örnek toplandığında ise 2360,47 dakika sürmüştür. Bu da eğitim veri sayısının artmasının hesaplama süresini üstel olarak arttırdığını göstermektedir.



Şekil 3.42. Alt örneklem kümesinin zamanın değişimi

4. SONUÇLAR

Bu tez çalışmasının temel amacı, çevresel izleme süreçlerinde bilgilendirici yol planlama algoritmalarının etkinliğini artırmak için yeni yöntemler geliştirmek ve bu yöntemlerin performansını çeşitli veri kümelerinde test etmektir. Çalışma kapsamında geliştirilen algoritmalar ve kullanılan ödül fonksiyonlarının performansı ayrıntılı olarak incelenmiş ve sonuçlar ışığında çeşitli bulgular elde edilmiştir.

Bu çalışmanın ilk aşamasında, eğitim verilerinin konumlarının çevresel izleme algoritmalarının performansını büyük ölçüde etkilediği vurgulanmıştır. Mevcut literatürde, eğitim verilerinin düzgün seçilmesinin önemine rağmen, bu konuda yeterince çalışma yapılmamıştır. Bu tez çalışmasında, sınırlı bir bütçeyle tüm alandan örnekler toplamak üzere yeni bir güzergah planlama algoritması geliştirilmiştir. Bu algoritma, 10 farklı veri kümesi üzerinde test edilmiş ve literatürdeki mevcut çalışmalarla kıyaslandığında daha başarılı sonuçlar elde etmiştir. Algoritmanın, ortamı daha etkili bir şekilde öğrendiği ve daha doğru tahminler yaptığı görülmüştür.

Çalışmanın ikinci aşamasında, ödül fonksiyonlarının ortamı öğrenme ve ROI alanlarını keşfetmedeki başarısı değerlendirilmiştir. Bu amaçla, içinde çeşitli sayılarda ROI alanı bulunan test ortamları oluşturulmuş ve mevcut konumdan maksimum fayda sağlayan yörüngeler seçilerek yeni örnekler toplanmıştır. Kullanılan ödül fonksiyonları arasında entropi, seviye belirleme ve UCB bulunmuştur.

Entropi Ödül Fonksiyonu: Daha az karmaşık veri setlerinde, ortamın belirsizliğini etkili bir şekilde düşürmüştür. Ancak örnekleme sayısı arttıkça belirsizliğin azalması yavaşlamıştır. Ayrıca, belirli bölgelerde fazla odaklanma yapmadan genel belirsizliği azaltmada başarılı olmuştur.

Seviye Belirleme Ödül Fonksiyonu: Özellikle karmaşık veri setlerinde ROI alanlarını daha başarılı bir şekilde keşfetmiştir. Bu fonksiyon, entropi ödül fonksiyonundan farklı olarak, belirli bölgelerde odaklanarak belirsizlik yönetiminde üstün performans sergilemiştir.

UCB Ödül Fonksiyonu: UCB ödül fonksiyonu, başlangıçta ortamın belirsizliğini hızlı ve etkili bir şekilde düşürmüştür. Ancak, belirsizliği düşük haritalarda yerel maksimum noktaya takılma riski taşıdığı gözlemlenmiştir. Bu nedenle, geniş alanlarda keşif yapma eksikliği nedeniyle bazı bölgelerde belirsizlik yüksek kalabilmiştir.

Üçüncü aşamada, gezgin satıcı problemine yönelik olarak MCTS ve BnB algoritmaları birleştirilerek yeni bir bilgilendirici yol planlama algoritması geliştirilmiştir. Önerilen algoritmanın başarımını değerlendirmek için deneysel bir veri seti kullanılmıştır. Kullanılan veri setinde sıcaklık, tuzluluk iletkenlik gibi ölçümler bulunmaktadır. Mevcut veriler kullanılarak bu parametrele çoklu çıktılı Gauss regresyon algoritması ile tahmin edilmiştir. Data setinde klorofil ölçümleride bulunmaktadır. Klorofil ölçümleri laboratuvar ortamında su numunesi toplanarak yapılabilmektedir. Bu durum gerçek zamanlı çevresel izleme probleminde sorun oluşturmakla beraber su

numunelerini toplayıp laboratuvar ortamına taşımak robotun enerji bütçesinin fazla harcanmasına neden olmaktadır. Bu nedenle veri setindeki uydudan yada anlık olarak sensörlerle ölçülebilen parametreler kullanılarak klorofil miktarı çok düşük bir belirsizlik değeri ile tahmin edilebilmiştir. Veri setindeki parametrelerin klorofil miktarıyla olan ilişkisi pearson korelasyon katsayısı kullanılarak ölçülmüştür. En ilişkili parametreler sıcaklık, tuzluluk, rüzgar hızı ve iletkenlik olarak bulunmuştur. Fakat çevresel izleme problemlerinde örneklerin çevresel koşullarında değişiminin etkisiyle sürekli toplanması gerekmektedir. Bu aşamada yeni örnekler toplanırken ortaya çıkan gezgin satıcı problemi üzerine odaklanılmıştır. Gezgin satıcı problemi olası en bilgilendirici örneklerin en kısa sürede toplanmasını içeren bir optimizasyon problemidir. Bu çalışmada bunun için hibrit bir güzergah planlama algoritması önerilmiştir. Geliştirilen MCTS-BnB algoritmasının performansı, MCTS ve ESAS algoritmaları ile karşılaştırılmıştır. MCTS, sezgisel olarak çalışarak olası güzergahları rastgele oluşturmuş ve maksimum ödüle sahip olan güzergah seçmiştir. ESAS algoritması ise deterministik bir yaklaşımla tüm olası yolları değerlendirerek maksimum ödüle sahip güzergahı belirlemiştir. Bu birleşimin, hem sezgisel hem de deterministik yaklaşımların avantajlarını bir araya getirerek daha optimal ve hızlı çözümler sunduğu görülmüştür.

Sonuçlar, MCTS-BnB algoritmasının MCTS kadar hızlı rota belirleyebildiğini ve ESAS algoritmasına yakın düzeyde optimal rotalar oluşturabildiğini göstermiştir. Bu durum, algoritmanın yüksek performansını ve geniş uygulama potansiyelini kanıtlamaktadır.

MCTS-BnB algoritması, çeşitli veri setlerinde (ℓ değeri 1 ve 3 olan NASA'nın sağladığı sıcaklık veri seti, ℓ değeri 9 olan deneysel bir veri seti) test edilmiştir. Veri setindeki değişkenliğin, ℓ hiper parametresini nasıl etkilediği incelenmiş ve algoritmanın farklı şartlar altındaki başarımı değerlendirilmiştir. Algoritma, değişkenliği yüksek veri setlerinde de başarılı sonuçlar elde etmiştir, bu da algoritmanın esnekliğini ve genelleme yeteneğini göstermektedir. Ayrıca literatürdeki diğer alt limit ve üst limit hesaplama yöntemleride deneysel veri setine uygulanarak MCTS-BnB algoritmasının başarımı araştırılmıştır. Elde edilen sonuçlarda literatürdeki diğer çalışmalar hesaplama süresinin kısa ama optimale yakın bir rotanın bulunduğu görülmüştür.

Son olarak Gauss regresyon algoritmasında eğitim veri kümesinin boyutu arttıkça hesaplama süresinin de arttığı göz önüne alınarak, seyrek temsili girişlerin belirlenmesi için beş farklı kümeleme yöntemi uygulanmıştır. Bu yöntemler arasında maksimum SSI değerine sahip olan kümeleme yöntemi seçilmiş ve eğitim veri seti kümelere ayrılmıştır. Her bir kümeden 1, 2, 3 ve 4 örnek alınarak hesaplama süresinin ve sonuçların optimal rotaya yakınlığı araştırılmıştır. Sonuçlar, kümelere alınan örnek sayısının artmasıyla birlikte oluşturulan güzergahların optimal rotaya yakınlığının arttığını, ancak hesaplama süresinin de aynı oranda yükseldiğini göstermektedir. Bu bulgu, daha az sayıda örnek çalışıldığında hesaplama süresinin düşürülebileceğini, ancak belirli durumlarda çözümün optimalden sapabileceğini doğrulamaktadır.

Bu çalışmanın bulguları, çevresel izleme ve bilgilendirici yol planlama alanlarında önemli katkılar sunmaktadır. Eğitim verilerinin stratejik seçimi ve efektif ödül fonksiyonları kullanımı, çevresel izleme algoritmalarının başarısını büyük ölçüde artırmaktadır. Bu bulgular, gelecekte bu alanda yapılacak çalışmalara yol gösterici olabilir. Özellikle ödül fonksiyonlarının ve güzergah planlama algoritmalarının kombinasyonu, mevcut sınırlamaların aşılmasını ve daha verimli örnek toplama stratejilerinin geliştirilmesine imkan tanımaktadır.

Bir diğer önemli bulgu, MCTS ve BnB algoritmalarının kombinasyonunun başarısıdır. Bu çalışma, sadece gezgin satıcı problemine değil, çeşitli yol planlama problemlerine de adapte edilebilecek bir model sunmaktadır. Bu kombinasyonun farklı veri setlerinde test edilmesi, algoritmanın genelleme yeteneğini ve geniş uygulama potansiyelini göstermektedir.

Son olarak, seyrek temsili girişlerin belirlenmesi, büyük veri kümeleriyle çalışan algoritmalar için hesaplama süresinin iyileştirilmesi yönünde önemli bir adımdır. Bu yöntemlerin başarılı bir şekilde uygulanması, özellikle büyük ölçeklerde uygulama yapılacak çevresel izleme projelerinde maliyet ve zaman tasarrufu sağlamaktadır.

Genel olarak, bu tez çalışması çeşitli ödül fonksiyonları ve algoritmaların çevresel izleme süreçlerinde nasıl performans gösterdiğini ve bu performansların veri setine, örnekleme stratejisine ve problem yapılarına nasıl bağlı olduğunu detaylı bir şekilde analiz etmektedir. Bu analizlerin sonuçları ve sonraki araştırmalar için sunduğu stratejik öneriler, çevresel izleme algoritmalarının geliştirilmesi ve iyileştirilmesi için güçlü bir temel oluşturur.

KAYNAKLAR

- [1] S. Manjanna, G. Dudek, Data-driven selective sampling for marine vehicles using multi-scale paths, *IEEE Int. Conf. Intell. Robot. Syst.* 2017-Septe (2017) 6111–6117. <https://doi.org/10.1109/IROS.2017.8206511>.
- [2] J. Das, F. Py, J.B.J. Harvey, J.P. Ryan, A. Gellene, R. Graham, D.A. Caron, K. Rajan, G.S. Sukhatme, Data-driven robotic sampling for marine ecosystem monitoring, *Int. J. Rob. Res.* 34 (2015) 1435–1452. <https://doi.org/10.1177/0278364915587723>.
- [3] M.J. Stealey, A. Singh, M.A. Batalin, B. Jordan, W.J. Kaiser, NIMS-AQ: A novel system for autonomous sensing of aquatic environments, *Proc. - IEEE Int. Conf. Robot. Autom.* (2008) 621–628. <https://doi.org/10.1109/ROBOT.2008.4543275>.
- [4] C. Denniston, A. Kumaraguru, G.S. Sukhatme, Comparison of path planning approaches for harmful algal bloom monitoring, *Ocean. 2019 MTS/IEEE Seattle, Ocean. 2019* (2019). <https://doi.org/10.23919/OCEANS40490.2019.8962687>.
- [5] P.V. Alvarado, T. Taher, H. Kurniawati, G. Weymouth, R.R. Khan, J. Leighton, G. Papadopoulos, G. Barbastathis, N. Patrikalakis, A coastal distributed autonomous sensor network, *Ocean. - MTS/IEEE Kona, Progr. B.* (2011). <https://doi.org/10.23919/oceans.2011.6106998>.
- [6] G. Hitz, F. Pomerleau, M.È. Garneau, C. Pradalier, T. Posch, J. Pernthaler, R.Y. Siegwart, Autonomous inland water monitoring: Design and application of a surface vessel, *IEEE Robot. Autom. Mag.* 19 (2012) 62–72. <https://doi.org/10.1109/MRA.2011.2181771>.
- [7] C.R. Anderson, M.R.P. Sapiano, M.B.K. Prasad, W. Long, P.J. Tango, C.W. Brown, R. Murtugudde, Predicting potentially toxigenic *Pseudo-nitzschia* blooms in the Chesapeake Bay, *J. Mar. Syst.* 83 (2010) 127–140. <https://doi.org/10.1016/j.jmarsys.2010.04.003>.
- [8] Y. Kaizu, M. Iio, H. Yamada, N. Noguchi, Development of unmanned airboat for water-quality mapping, *Biosyst. Eng.* 109 (2011) 338–347. <https://doi.org/10.1016/j.biosystemseng.2011.04.013>.
- [9] C.W. Powers, Monitoring the Transport of Microorganisms in Aquatic Environments Using Unmanned Surface Vehicles, (2017).
- [10] J. Moulton, N. Karapetyan, S. Bukhsbaum, C. McKinney, S. Malebary, G. Sophocleous, A.Q. Li, I. Rekleitis, An Autonomous Surface Vehicle for Long Term Operations, *Ocean. 2018 MTS/IEEE Charleston, Ocean 2018* (2019) 1–10. <https://doi.org/10.1109/OCEANS.2018.8604718>.
- [11] H. Bayram, Design and Implementation of Autonomous Surface Vehicle for Inland Water, *Iğdır Üniversitesi Fen Bilim. Enstitüsü Derg.* 10 (2020) 101–111. <https://doi.org/10.21597/jist.642503>.
- [12] D. Madeo, A. Pozzebon, C. Mocenni, D. Bertoni, A Low-Cost Unmanned Surface Vehicle for Pervasive Water Quality Monitoring, *IEEE Trans. Instrum. Meas.* 69 (2020) 1433–1444. <https://doi.org/10.1109/TIM.2019.2963515>.
- [13] F.P. Samaniego, D.G. Reina, S.L.T. Marin, M. Arzamendia, D.O. Gregor, A Bayesian Optimization Approach for Water Resources Monitoring through an Autonomous Surface Vehicle: The Ypacarai Lake Case Study, *IEEE Access* 9 (2021) 9163–9179. <https://doi.org/10.1109/ACCESS.2021.3050934>.
- [14] A.M. Barrionuevo, S. Yanes Luis, D. Gutierrez Reina, S.L. Toral Marin, Informative Deep Reinforcement Path Planning for Heterogeneous Autonomous Surface Vehicles in Large Water Resources, *IEEE Access* 12 (2024) 71835–71852. <https://doi.org/10.1109/ACCESS.2024.3402980>.
- [15] F. Peralta, D.G. Reina, S. Toral, M. Arzamendia, D. Gregor, A bayesian optimization approach for multi-function estimation for environmental monitoring using an autonomous surface vehicle: Ypacarai lake case study, *Electron.* 10 (2021). <https://doi.org/10.3390/electronics10080963>.
- [16] S. Yanes Luis, D. Shutin, J. Marchal Gómez, D. Gutiérrez Reina, S. Toral Marín, Deep Reinforcement Multiagent Learning Framework for Information Gathering with Local Gaussian Processes for Water Monitoring, *Adv. Intell. Syst.* 2300850 (2024). <https://doi.org/10.1002/aisy.202300850>.
- [17] S. Kemna, H. Heiarsson, G.S. Sukhatme, On-board Adaptive Informative Sampling for AUVs: A

- Feasibility Study, Ocean. 2018 MTS/IEEE Charleston, Ocean 2018 (2019). <https://doi.org/10.1109/OCEANS.2018.8604838>.
- [18] A.J. Shafer, M.R. Benjamin, J.J. Leonard, J. Curcio, Autonomous cooperation of heterogeneous platforms for sea-based search tasks, Ocean. 2008 (2008). <https://doi.org/10.1109/OCEANS.2008.5152100>.
- [19] R.N. Smith, M. Schwager, S.L. Smith, B.H. Jones, D. Rus, G.S. Sukhatme, Persistent ocean monitoring with underwater gliders: Adapting sampling resolution, J. F. Robot. 28 (2011) 714–741.
- [20] P. Tokekar, D. Bhadauria, A. Studenski, V. Isler, A robotic system for monitoring carp in Minnesota lakes, J. F. Robot. 27 (2010) 779–789.
- [21] M. Bryson, A. Reid, F. Ramos, S. Sukkarieh, Airborne vision-based mapping and classification of large farmland environments, J. F. Robot. 27 (2010) 632–655.
- [22] P.N. Eze, S.K. Kumahor, Gaussian process simulation of soil Zn micronutrient spatial heterogeneity and uncertainty – A performance appraisal of three semivariogram models, Sci. African 5 (2019) e00110. <https://doi.org/10.1016/j.sciaf.2019.e00110>.
- [23] C. Ballabio, E. Lugato, O. Fernández-Ugalde, A. Orgiazzi, A. Jones, P. Borrelli, L. Montanarella, P. Panagos, Mapping LUCAS topsoil chemical properties at European scale using Gaussian process regression, Geoderma 355 (2019) 113912. <https://doi.org/10.1016/j.geoderma.2019.113912>.
- [24] M.F. Mysorewala, D.O. Popa, F.L. Lewis, Multi-scale adaptive sampling with mobile agents for mapping of forest fires, J. Intell. Robot. Syst. Theory Appl. 54 (2009) 535–565. <https://doi.org/10.1007/s10846-008-9246-1>.
- [25] G. Muscato, F. Bonaccorso, L. Cantelli, D. Longo, C.D. Melita, Volcanic environments: Robots for exploration and measurement, IEEE Robot. Autom. Mag. 19 (2012) 40–49. <https://doi.org/10.1109/MRA.2011.2181684>.
- [26] S. Manjanna, N. Kakodkar, M. Meghjani, G. Dudek, Efficient terrain driven coral coverage using Gaussian processes for mosaic synthesis, Proc. - 2016 13th Conf. Comput. Robot Vision, CRV 2016 (2016) 448–455. <https://doi.org/10.1109/CRV.2016.63>.
- [27] K.H. Low, G.J. Gordon, J.M. Dolan, P. Khosla, Adaptive Sampling for Multi-Robot Wide-Area Exploration, (2024) 10–14.
- [28] V. Formisano, S. Atreya, T. Encrenaz, N. Ignatiev, M. Giuranna, Detection of methane in the atmosphere of Mars, Science (80-.). 306 (2004) 1758–1761.
- [29] M. Meghjani, S. Manjanna, G. Dudek, Multi-target search strategies, SSRR 2016 - Int. Symp. Safety, Secur. Rescue Robot. (2016) 328–333. <https://doi.org/10.1109/SSRR.2016.7784323>.
- [30] O. Velasco, J. Valente, A.Y. Mersha, An Adaptive Informative Path Planning Algorithm for Real-time Air Quality Monitoring Using UAVs, 2020 Int. Conf. Unmanned Aircr. Syst. ICUAS 2020 (2020) 1121–1130. <https://doi.org/10.1109/ICUAS48674.2020.9214013>.
- [31] M.S. Assenine, W. Bechkit, H. Rivano, Informative and Communication-Efficient Multi-Agent Path Planning for Pollution Plume Monitoring, Proc. - 2024 IEEE 25th Int. Symp. a World Wireless, Mob. Multimed. Networks, WoWMoM 2024 (2024) 329–338. <https://doi.org/10.1109/WoWMoM60985.2024.00060>.
- [32] H. Zhu, J.J. Chung, N.R.J. Lawrance, R. Siegwart, J. Alonso-Mora, Online Informative Path Planning for Active Information Gathering of a 3D Surface, Proc. - IEEE Int. Conf. Robot. Autom. 2021-May (2021) 1488–1494. <https://doi.org/10.1109/ICRA48506.2021.9561963>.
- [33] O.İ. Çelik S. Çelik, C. Gazioglu, Evaluation on 2002-2021 CHL-A Concentrations in the Sea of Marmara with GEE Enhancement of Satellite Data, Int. J. Environ. Geoinformatics 9 (2022) 68–77. <https://doi.org/10.30897/ijegeo.1066168>.
- [34] G.T. Kayaalp, S. Polat, Tüm Gözlemler ve Eksik Gözlemler Regresyon Modelinde Klorofil-a Miktarının Tahmini, 18 (2001) 529–535.
- [35] M. Papenfus, B. Schaeffer, A.I. Pollard, K. Loftin, Exploring the potential value of satellite remote sensing to monitor chlorophyll-a for US lakes and reservoirs, Environ. Monit. Assess. 192 (2020). <https://doi.org/10.1007/s10661-020-08631-5>.

- [36] C. Karul, S. Soyupak, A.F. Çilesiz, N. Akbay, E. Germen, Case studies on the use of neural networks in eutrophication modeling, *Ecol. Modell.* 134 (2000) 145–152. [https://doi.org/10.1016/S0304-3800\(00\)00360-4](https://doi.org/10.1016/S0304-3800(00)00360-4).
- [37] F.A. Almomani, B. Örmeci, Monitoring and measurement of microalgae using the first derivative of absorbance and comparison with chlorophyll extraction method, *Environ. Monit. Assess.* 190 (2018). <https://doi.org/10.1007/s10661-018-6468-y>.
- [38] N. Doğan-Sağlamtimur, B. Sağlamtimur, Sucul Ortamlarda Ötrofikasyon Durumu Ve Senaryolari, Ömer Halisdemir Üniversitesi Mühendislik Bilim. Derg. 7 (2018) 75–82. <https://doi.org/10.28948/ngumuh.386307>.
- [39] O. Yüzügüllü, Determination of chlorophyll-a distribution in Lake Eymir using regression and artificial neural network models with hybrid inputs, Middle East Technical University, 2011.
- [40] C.E. Rasmussen, H. Nickisch, Gaussian processes for machine learning (GPML) toolbox, *J. Mach. Learn. Res.* 11 (2010) 3011–3015.
- [41] M.P. Deisenroth, Efficient reinforcement learning using Gaussian processes, KIT Scientific Publishing, 2010.
- [42] P. Boyle, Gaussian processes for regression and optimisation, Open Access Te Herenga Waka-Victoria University of Wellington, 2007.
- [43] B. Nguyen, L. Nguyen, T.X. Nghiem, H. La, J. Baca, P. Rangel, M.C. Montoya, T. Nguyen, Spatially temporally distributed informative path planning for multi-robot systems, (n.d.).
- [44] M. Quann, L. Ojeda, W. Smith, D. Rizzo, M. Castanier, K. Barton, Off-road ground robot path energy cost prediction through probabilistic spatial mapping, *J. F. Robot.* 37 (2020) 421–439. <https://doi.org/10.1002/rob.21927>.
- [45] P. Stankiewicz, Y.T. Tan, M. Kobilarov, Adaptive sampling with an autonomous underwater vehicle in static marine environments, *J. F. Robot.* 38 (2021) 572–597. <https://doi.org/10.1002/rob.22005>.
- [46] J. Binney, G.S. Sukhatme, Branch and bound for informative path planning, *Proc. - IEEE Int. Conf. Robot. Autom.* (2012) 2147–2154. <https://doi.org/10.1109/ICRA.2012.6224902>.
- [47] R. Mishra, T.B. Koay, M. Chitre, S. Swarup, Multi-USV Adaptive Exploration Using Kernel Information and Residual Variance, *Front. Robot. AI* 8 (2021) 1–21. <https://doi.org/10.3389/frobt.2021.572243>.
- [48] G. Hitz, E. Galceran, M.È. Garneau, F. Pomerleau, R. Siegwart, Adaptive continuous-space informative path planning for online environmental monitoring, *J. F. Robot.* 34 (2017) 1427–1449. <https://doi.org/10.1002/rob.21722>.
- [49] Y.T. Tan, A. Kunapareddy, M. Kobilarov, Gaussian Process Adaptive Sampling Using the Cross-Entropy Method for Environmental Sensing and Monitoring, *Proc. - IEEE Int. Conf. Robot. Autom.* (2018) 6220–6227. <https://doi.org/10.1109/ICRA.2018.8460821>.
- [50] T. Chen, B. Dai, R. Wang, D. Liu, Gaussian-process-based real-time ground segmentation for autonomous land vehicles, *J. Intell. \& Robot. Syst.* 76 (2014) 563–582.
- [51] D.A. Duecker, A.R. Geist, E. Kreuzer, E. Solowjow, Learning environmental field exploration with computationally constrained underwater robots: Gaussian processes meet stochastic optimal control, *Sensors (Switzerland)* 19 (2019). <https://doi.org/10.3390/s19092094>.
- [52] S. Kemna, O. Kroemer, G.S. Sukhatme, Pilot Surveys for Adaptive Informative Sampling, *Proc. - IEEE Int. Conf. Robot. Autom.* (2018) 6417–6424. <https://doi.org/10.1109/ICRA.2018.8460488>.
- [53] P. Erickson, M. Cline, N. Tirpankar, T. Henderson, Gaussian processes for multi-sensor environmental monitoring, *IEEE Int. Conf. Multisens. Fusion Integr. Intell. Syst.* 2015-Octob (2015) 208–213. <https://doi.org/10.1109/MFI.2015.7295810>.
- [54] R. Grbić, D. Kurtagić, D. Slišković, Stream water temperature prediction based on Gaussian process regression, *Expert Syst. Appl.* 40 (2013) 7407–7414. <https://doi.org/10.1016/j.eswa.2013.06.077>.
- [55] K.C. Ma, L. Liu, H.K. Heidarrson, G.S. Sukhatme, Data-driven learning and planning for environmental sampling, *J. F. Robot.* 35 (2018) 643–661. <https://doi.org/10.1002/rob.21767>.

- [56] J. Binney, A. Krause, G.S. Sukhatme, Informative path planning for an autonomous underwater vehicle, *Proc. - IEEE Int. Conf. Robot. Autom.* (2010) 4791–4796. <https://doi.org/10.1109/ROBOT.2010.5509714>.
- [57] R. Cui, Y. Li, W. Yan, Mutual Information-Based Multi-AUV Path Planning for Scalar Field Sampling Using Multidimensional RRT*, *IEEE Trans. Syst. Man, Cybern. Syst.* 46 (2016) 993–1004. <https://doi.org/10.1109/TSMC.2015.2500027>.
- [58] K.C. Ma, L. Liu, G.S. Sukhatme, An information-driven and disturbance-aware planning method for long-term ocean monitoring, *IEEE Int. Conf. Intell. Robot. Syst.* 2016-Novem (2016) 2102–2108. <https://doi.org/10.1109/IROS.2016.7759330>.
- [59] K.C. Ma, L. Liu, G.S. Sukhatme, Informative planning and online learning with sparse Gaussian processes, *Proc. - IEEE Int. Conf. Robot. Autom.* (2017) 4292–4298. <https://doi.org/10.1109/ICRA.2017.7989494>.
- [60] N. Fung, J. Rogers, C. Nieto, H.I. Christensen, S. Kemna, G. Sukhatme, Coordinating multi-robot systems through environment partitioning for adaptive informative sampling, *Proc. - IEEE Int. Conf. Robot. Autom.* 2019-May (2019) 3231–3237. <https://doi.org/10.1109/ICRA.2019.8794103>.
- [61] M. Kobilarov, Cross-entropy motion planning, *Int. J. Rob. Res.* 31 (2012) 855–871. <https://doi.org/10.1177/0278364912444543>.
- [62] A. Munir, R. Parasuraman, Exploration–Exploitation Tradeoff in the Adaptive Information Sampling of Unknown Spatial Fields with Mobile Robots, *Sensors* 23 (2023). <https://doi.org/10.3390/s23239600>.
- [63] A. Viseras, D. Shutin, L. Merino, Online information gathering using sampling-based planners and GPs: An information theoretic approach, in: *2017 IEEE/RSJ Int. Conf. Intell. Robot. Syst.*, 2017: pp. 123–130.
- [64] A. Gotovos, N. Casati, G. Hitz, A. Krause, Active learning for level set estimation, *IJCAI Int. Jt. Conf. Artif. Intell.* (2013) 1344–1350.
- [65] M. De Petrillo, D. Ross, J.N. Gross, Gaussian Process Regression for Learning Environment Impacts on Localization Accuracy of a UAV with Respect to UGV for Search Planning, *2023 IEEE/ION Position, Locat. Navig. Symp. PLANS 2023* (2023) 260–271. <https://doi.org/10.1109/PLANS53410.2023.10139936>.
- [66] G. Flaspohler, V. Preston, A.P.M. Michel, Y. Girdhar, N. Roy, Information-Guided Robotic Maximum Seek-and-Sample in Partially Observable Continuous Environments, *IEEE Robot. Autom. Lett.* 4 (2019) 3782–3789. <https://doi.org/10.1109/lra.2019.2929997>.
- [67] L. Bottarelli, M. Bicego, J. Blum, A. Farinelli, Orienteering-based informative path planning for environmental monitoring, *Eng. Appl. Artif. Intell.* 77 (2019) 46–58. <https://doi.org/10.1016/j.engappai.2018.09.015>.
- [68] Q. Cheng, W. Zhang, H. Liu, Y. Zhang, L. Hao, Research on the path planning algorithm of a manipulator based on gmm/gmr-mpm, *Appl. Sci.* 11 (2021). <https://doi.org/10.3390/app11167599>.
- [69] J. Wang, T. Li, B. Li, M.Q.H. Meng, GMR-RRT*: Sampling-Based Path Planning Using Gaussian Mixture Regression, *IEEE Trans. Intell. Veh.* 7 (2022) 690–700. <https://doi.org/10.1109/TIV.2022.3150748>.
- [70] J. Suh, S. Oh, Efficient environmental monitoring using cost-aware path planning, *Int. Conf. Control. Autom. Syst.* 2 (2013) 1362–1365. <https://doi.org/10.1109/ICCAS.2013.6704169>.
- [71] C. Xiong, H. Zhou, D. Lu, Z. Zeng, L. Lian, C. Yu, Rapidly-exploring adaptive sampling tree*: A sample-based path-planning algorithm for unmanned marine vehicles information gathering in variable ocean environments, *Sensors (Switzerland)* 20 (2020). <https://doi.org/10.3390/s20092515>.
- [72] G. Hitz, A. Gotovos, F. Pomerleau, M.É. Garneau, C. Pradalier, A. Krause, R.Y. Siegwart, Fully autonomous focused exploration for robotic environmental monitoring, *Proc. - IEEE Int. Conf. Robot. Autom.* (2014) 2658–2664. <https://doi.org/10.1109/ICRA.2014.6907240>.
- [73] Z.W. Lim, D. Hsu, W.S. Lee, Adaptive informative path planning in metric spaces, *Int. J. Rob. Res.* 35 (2016) 585–598. <https://doi.org/10.1177/0278364915596378>.

- [74] A. Jones, M. Schwager, C. Belta, A receding horizon algorithm for informative path planning with temporal logic constraints, *Proc. - IEEE Int. Conf. Robot. Autom.* (2013) 5019–5024. <https://doi.org/10.1109/ICRA.2013.6631294>.
- [75] J.L. Nguyen, N.R.J. Lawrance, S. Sukkarieh, Nonmyopic planning for long-Term information gathering with an aerial glider, *Proc. - IEEE Int. Conf. Robot. Autom.* (2014) 6573–6578. <https://doi.org/10.1109/ICRA.2014.6907829>.
- [76] K. Tiwari, V. Honore, S. Jeong, N.Y. Chong, M.P. Deisenroth, Resource-constrained decentralized active sensing for multi-robot systems using distributed Gaussian processes, *Int. Conf. Control. Autom. Syst.* 0 (2016) 13–18. <https://doi.org/10.1109/ICCAS.2016.7832293>.
- [77] Q. V. Nguyen, F. Colas, E. Vincent, F. Charpillet, Long-term robot motion planning for active sound source localization with Monte Carlo tree search, 2017 Hands-Free Speech Commun. Microphone Arrays, HSCMA 2017 - Proc. (2017) 61–65. <https://doi.org/10.1109/HSCMA.2017.7895562>.
- [78] D. Jang, J. Yoo, C.Y. Son, H.J. Kim, Fully Distributed Informative Planning for Environmental Learning with Multi-Robot Systems, (2021) 1–9. <http://arxiv.org/abs/2112.14433>.
- [79] J. McMahan, H. Yetkin, A. Wolek, Z.J. Waters, D.J. Stilwell, Towards real-time search planning in subsea environments, *IEEE Int. Conf. Intell. Robot. Syst.* 2017-Sept (2017) 87–94. <https://doi.org/10.1109/IROS.2017.8202142>.
- [80] G. Best, R. Fitch, Probabilistic maximum set cover with path constraints for informative path planning, *Australas. Conf. Robot. Autom. ACRA 2016-Decem* (2016) 97–106.
- [81] W. Chen, L. Liu, Pareto Monte Carlo Tree Search for Multi-Objective Informative Planning, *Robot. Sci. Syst.* (2019). <https://doi.org/10.15607/RSS.2019.XV.072>.
- [82] K. Kurzer, M. Fechner, J.M. Zollner, Accelerating Cooperative Planning for Automated Vehicles with Learned Heuristics and Monte Carlo Tree Search, *IEEE Intell. Veh. Symp. Proc.* (2020) 1726–1733. <https://doi.org/10.1109/IV47402.2020.9304597>.
- [83] B. Kartal, E. Nunes, J. Godoy, M. Gini, Monte carlo tree search for multi-robot task allocation, 30th AAAI Conf. Artif. Intell. AAAI 2016 (2016) 4222–4223. <https://doi.org/10.1609/aaai.v30i1.9945>.
- [84] M. Jara Ten Kathen, I. Jurado Flores, D. Gutiérrez Reina, An informative path planner for a swarm of asvs based on an enhanced pso with gaussian surrogate model components intended for water monitoring applications, *Electron.* 10 (2021). <https://doi.org/10.3390/electronics10131605>.
- [85] M. Carolina Jara Ten Kathen, I. Jurado Flores, D. Gutierrez Reina, A Comparison of PSO-Based Informative Path Planners for Autonomous Surface Vehicles for Water Resource Monitoring, *ACM Int. Conf. Proceeding Ser.* (2022) 271–276. <https://doi.org/10.1145/3529399.3529442>.
- [86] M. Budd, B. Lacerda, P. Duckworth, A. West, B. Lennox, N. Hawes, Markov decision processes with unknown state feature values for safe exploration using gaussian processes, *IEEE Int. Conf. Intell. Robot. Syst.* (2020) 7344–7350. <https://doi.org/10.1109/IROS45743.2020.9341589>.
- [87] T.A. Frewen, H. Sane, M. Kobilarov, S. Bajekal, K.R. Chevva, Adaptive path planning in a dynamic environment using a receding horizon probabilistic roadmap: Experimental demonstration, *Am. Helicopter Soc. Int. - AHS Spec. Meet. Unmanned Rotorcr. Netw. Centric Oper.* 2011 (2011) 176–184.
- [88] S. Anderson, T.D. Barfoot, C.H. Tong, S. Särkkä, Batch nonlinear continuous-time trajectory estimation as exactly sparse Gaussian process regression, *Auton. Robots* 39 (2015) 221–238. <https://doi.org/10.1007/s10514-015-9455-y>.
- [89] S. Anderson, T.D. Barfoot, Full STEAM ahead: Exactly sparse Gaussian process regression for batch continuous-time trajectory estimation on SE(3), *IEEE Int. Conf. Intell. Robot. Syst.* 2015-Decem (2015) 157–164. <https://doi.org/10.1109/IROS.2015.7353368>.
- [90] G.Z. Gandler, C.H. Ek, M. Björkman, R. Stolkin, Y. Bekiroglu, Object shape estimation and modeling, based on sparse Gaussian process implicit surfaces, combining visual data and tactile exploration, *Rob. Auton. Syst.* 126 (2020) 103433. <https://doi.org/10.1016/j.robot.2020.103433>.
- [91] A. Dong, Z. Du, Z. Yan, A sensorless interaction forces estimator for bilateral teleoperation system based on online sparse Gaussian process regression, *Mech. Mach. Theory* 143 (2020).

- <https://doi.org/10.1016/j.mechmachtheory.2019.103620>.
- [92] T. Chen, B. Dai, D. Liu, J. Song, Sparse Gaussian process regression based ground segmentation for autonomous land vehicles, *Proc. 2015 27th Chinese Control Decis. Conf. CCDC 2015* (2015) 3993–3998. <https://doi.org/10.1109/CCDC.2015.7162621>.
- [93] K. Jakkala, S. Akella, Multi-Robot Informative Path Planning from Regression with Sparse Gaussian Processes, (n.d.). <https://github.com/itskalvik/SGP-Tools>.
- [94] H.-H. Bock, Clustering Methods: A History of k-Means Algorithms, (2007) 161–172. https://doi.org/10.1007/978-3-540-73560-1_15.
- [95] Y.A. Prabowo, B.R. Trilaksono, E.M.I. Hidayat, B. Yulianto, Utilizing a Rapidly Exploring Random Tree for Hazardous Gas Exploration in a Large Unknown Area, *IEEE Access* 10 (2022) 15336–15347. <https://doi.org/10.1109/ACCESS.2022.3147720>.
- [96] H. Liu, J. Cai, Y.-S. Ong, Remarks on multi-output Gaussian process regression, *Knowledge-Based Syst.* 144 (2018) 102–121.
- [97] F. Leibfried, V. Dutoir, S. John, N. Durrande, A Tutorial on Sparse Gaussian Processes and Variational Inference, (2020) 1–37. <http://arxiv.org/abs/2012.13962>.
- [98] GHRSSST Level 4 MUR Global Foundation Sea Surface Temperature Analysis (v4.1), No Title, (n.d.) <https://podaac.jpl.nasa.gov/dataset/MUR-JPL-L4-GLO>.
- [99] A. Gotovos, N. Casati, E.T.H. Zurich, G. Hitz, A. Krause, *IJCAI13-202.pdf*, (2009) 1344–1350. <papers3://publication/uuid/07F3A6A8-2905-4B9E-874A-556175725FAF>.
- [100] H. Mohammadi, P. Challenor, D. Williamson, M. Goodfellow, Cross-Validation--based Adaptive Sampling for Gaussian Process Models, *SIAM/ASA J. Uncertain. Quantif.* 10 (2022) 294–316.
- [101] R. Coulom, Efficient selectivity and backup operators in Monte-Carlo tree search, *Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* 4630 LNCS (2007) 72–83. https://doi.org/10.1007/978-3-540-75538-8_7.
- [102] M.C. Fu, A Tutorial Introduction to Monte Carlo Tree Search, *Proc. - Winter Simul. Conf. 2020-Decem* (2020) 1178–1193. <https://doi.org/10.1109/WSC48552.2020.9384090>.
- [103] H. Yetkin, *Toward Real-Time Planning for Robotic Search*, (2018).
- [104] N. Shi, X. Liu, Y. Guan, Research on k-means clustering algorithm: An improved k-means clustering algorithm, *3rd Int. Symp. Intell. Inf. Technol. Secur. Informatics, IITSI 2010* (2010) 63–67. <https://doi.org/10.1109/IITSI.2010.74>.
- [105] M. Daszykowski, B. Walczak, 2.26 - Density-Based Clustering Methods, *Compr. Chemom. Chem. Biochem. Data Anal. Second Ed. Four Vol. Set 2* (2020) 565–580. <https://doi.org/10.1016/B978-0-444-64165-6.03005-6>.
- [106] A. Lukasová, Hierarchical agglomerative clustering procedure, *Pattern Recognit.* 11 (1979) 365–381.
- [107] D. Müllner, Modern hierarchical, agglomerative clustering algorithms, (2011) 1–29. <http://arxiv.org/abs/1109.2378>.
- [108] T. Zhang, R. Ramakrishnan, M. Livny, BIRCH: A new data clustering algorithm and its applications, *Data Min. Knowl. Discov.* 1 (1997) 141–182. <https://doi.org/10.1023/A:1009783824328>.
- [109] T. Zhang, R. Ramakrishnan, M. Livny, BIRCH: An Efficient Data Clustering Method for Very Large Databases, *SIGMOD Rec. (ACM Spec. Interes. Gr. Manag. Data)* 25 (1996) 103–114. <https://doi.org/10.1145/235968.233324>.
- [110] D. Sculley, Web-scale k-means clustering, in: *Proc. 19th Int. Conf. World Wide Web, 2010*: pp. 1177–1178.
- [111] A. Ng, M. Jordan, Y. Weiss, On spectral clustering: Analysis and an algorithm, *Adv. Neural Inf. Process. Syst.* 14 (2001).
- [112] U. Von Luxburg, A tutorial on spectral clustering, *Stat. Comput.* 17 (2007) 395–416. <https://doi.org/10.1007/s11222-007-9033-z>.
- [113] G. Ogbuabor, U. F. N, Clustering Algorithm for a Healthcare Dataset Using Silhouette Score Value,

Int. J. Comput. Sci. Inf. Technol. 10 (2018) 27–37. <https://doi.org/10.5121/ijcsit.2018.10203>.

- [114] A. Naghizadeh, D.N. Metaxas, Condensed silhouette: An optimized filtering process for cluster selection in K-means, *Procedia Comput. Sci.* 176 (2020) 205–214. <https://doi.org/10.1016/j.procs.2020.08.022>.
- [115] A.M. Chiodi, C. Zhang, E.D. Cokelet, Q. Yang, C.W. Mordy, C.L. Gentemann, J.N. Cross, N. Lawrence-Slavas, C. Meinig, M. Steele, others, Exploring the Pacific Arctic seasonal ice zone with saildrone USVs, *Front. Mar. Sci.* 8 (2021) 640690.



ÖZGEÇMİŞ

Perihan KARAKÖSE

[REDACTED]

[REDACTED] [REDACTED]
[REDACTED] [REDACTED]
[REDACTED] [REDACTED]
[REDACTED] [REDACTED]
[REDACTED] [REDACTED]
[REDACTED] [REDACTED]

[REDACTED]

[REDACTED] [REDACTED]
[REDACTED] [REDACTED]

[REDACTED]

[REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED] [REDACTED]
[REDACTED]
[REDACTED] [REDACTED]
[REDACTED] [REDACTED]

[REDACTED]

- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]

[REDACTED]

[REDACTED] [REDACTED]

AKADEMİK FAALİYETLER

Makaleler:

- Karaköse, P., & Koca, A. (2024). Turbine Type Rotary Wave Energy Converter Performance. *Gazi University Journal of Science*, 37(1), 264-282
- Karaköse, P., & Koca, A. (2022). A Novel Rotor Type Wave Energy Converter Design for Maximum Energy Captured in Low Wave Heights. *Batman Üniversitesi Yaşam Bilimleri Dergisi*, 12(2), 136-153.

Bildiriler:

- Ay, M., Ozmen Koca, G., Korkmaz, D., Bingol, M. C., Akpolat, Z. H., Bal, C., & Karakose, P. (2017, October). Prototype design and waterproof production of intelligent robotic fish (i-rof). In *8th International Advanced Technologies Symposium* (pp. 3960-3968)

2. Yamaç, H. İ., Koca, A., Karakose, P., “Numerical Analysis of Wave Amplitude Effect on Power Output of Oscillating Water Column with Coastal Elevation,” presented at the 1st International Engineering And Thecnology Symposium .
3. Karakose, P., Yetkin, H., Bal, C., ““Balanced Distribution Method For Prior Data Selection In Hyperparameter Estimation With Limited Budget”, 4th Nternational Azerbaijan Congress On Life, Engineering, And Applied Sciences,2023
4. Karakose, P., Bal, C., “Determining Water Quality Using Gaussion Regression”, 9th International Asian Congress On Contemporary Sciences, 2023

Projeler:

1. Karakose, P., Koca, A., “Kıyıda Uzak Tıp Fdam' Lı Dalga Enerjisi Dönüştürücüsünün Geliştirilmesi”, Fırat Üniversitesi Bilimsel Araştırma Projesi, Araştırmacı
2. Bal, C., Akpolat, Z. H., Ay, M., Özmen Koca, G., Bingöl, M.C., Korkmaz, D., Karaköse, P., “Üç Boyutlu Hareket Kabiliyetine Sahip Carangiform Türünde Bir Biyomimetik Robot Balığın Tasarımı ve Gerçeklemesi”, Tübitak 1001, Bursiyer
3. İnce, O., Karakose, P., Demir, F., İnce, E.G., Atar, M., “Betonarme binalarda tasarım hatalarının tespitinin yapılması için görüntü işleme teknikleri tabanlı karar destek sistemi geliştirilmesi”, Tübitak 1001, Araştırmacı