

Eylül'26

Sayı/Number: 3 / Volume: 16
Yıl/Year: 2026

Yayın Sahibi

TMMOB

Elektrik Mühendisleri Odası Adına
Mahir ULUTAŞ

Sorumlu Yazı İşleri Müdürü

Ender KELLEÇİ

Yayın İdare Merkezi

İhlamur Sokak No:10 Kat:3
Kızılay - Ankara
Tel: (312) 425 32 72
Faks: (312) 417 38 18
<http://bilimseldergi.emo.org.tr>
bilimseldergi@emo.org.tr
EMO Üyelerine parasız dağıtılır

Dergi Koordinatörü

Sevgi Kinacı
sevgi.kinaci@emo.org.tr

Sayfa Düzeni

TMMOB

Elektrik Mühendisleri Odası

Yayın Türü

Yerel Süreli Yayın
4 ayda bir yayınlanır

Basım Adedi

500

Basım Tarihi

Eylül 2026

EMO BİLİMSEL DERGİ

Elektrik, Elektronik, Bilgisayar, Biyomedikal
Mühendisliği Bilimsel Dergisi

The Journal of Electrical, Electronics, Computer and
Biomedical Engineering

YAYIN KURULU

BAŞ EDITÖR/EDITOR IN CHIEF

Prof. Dr. Arif NACAROĞLU
Gaziantep Üniversitesi

EDİTÖRLER/EDITORIAL BOARD

Prof. Dr. M. Timur AYDEMİR
Kadir Has Üniversitesi

Prof. Dr. Erhan AKIN
Fırat Üniversitesi

Prof. Dr. Aydoğan ÖZDEMİR
İstanbul Teknik Üniversitesi

Prof. Dr. Muhittin GÖKMEN
MEF Üniversitesi

Prof. Dr. Tayfun AKGÜL
İstanbul Teknik Üniversitesi



TMMOB

Elektrik Mühendisleri Odası

UCTEA/Chamber of Electrical Engineers

EMO Bilimsel Dergi Danışma Kurulu

Prof. Dr. A. Hamit Serbest	Çukurova Üniversitesi
Prof. Dr. Altay Güvenir	Bilkent Üniversitesi
Prof. Dr. Cengizhan Öztürk	Boğaziçi Üniversitesi
Prof. Dr. Erhan Akın	Fırat Üniversitesi
Prof. Dr. Güven Önbilgin	Ondokuz Mayıs Üniversitesi
Prof. Dr. Murat Eyüboğlu	Orta Doğu Teknik Üniversitesi
Prof. Dr. Timur Aydemir	Gazi Üniversitesi
Prof. Dr. Ali Hikmet Doğru	Orta Doğu Teknik Üniversitesi
Prof. Dr. Arif Nacaroglu	Gaziantep Üniversitesi
Prof. Dr. Aydın Köksal	Bilişim A.Ş.
Prof. Dr. Aydoğan Özdemir	İstanbul Teknik Üniversitesi
Prof. Dr. Aysın Baytan Ertüzün	Boğaziçi Üniversitesi
Prof. Dr. Belgin Turkey	İstanbul Teknik Üniversitesi
Prof. Dr. Bülent Sankur	Boğaziçi Üniversitesi
Prof. Dr. Cüneyt Güzeliş	Yaşar Üniversitesi
Prof. Dr. Erdal Panayırıcı	Kadir Has Üniversitesi
Prof. Dr. Erkan Afacan	Gazi Üniversitesi
Prof. Dr. Ferit Acar Savacı	İzmir İleri Teknoloji Enstitüsü
Prof. Dr. H. Bülent Ertan	Atılım Üniversitesi
Prof. Dr. Haldun Karaca	Dokuz Eylül Üniversitesi
Prof. Dr. İbrahim Eksin	İstanbul Teknik Üniversitesi
Prof. Dr. İnci Çilesiz	İstanbul Teknik Üniversitesi
Prof. Dr. İrfan Karagöz	Gazi Üniversitesi
Prof. Dr. İsmail Hakkı Altaş	Karadeniz Teknik Üniversitesi
Prof. Dr. İsmail Hakkı Çavdar	Karadeniz Teknik Üniversitesi
Prof. Dr. Kemal Leblebicioğlu	Orta Doğu Teknik Üniversitesi
Prof. Dr. Lale Tükenmez Ergene	İstanbul Teknik Üniversitesi
Prof. Dr. Mithat İdemen	İstanbul Teknik Üniversitesi
Prof. Dr. Muhittin Gökmen	MEF Üniversitesi
Prof. Dr. Murat Aşkar	İzmir Ekonomi Üniversitesi
Prof. Dr. Osman Eroğlu	TOBB Ekonomi ve Teknoloji Üniversitesi
Prof. Dr. Oya Kalıpsız	Yıldız Teknik Üniversitesi
Prof. Dr. Özlem Özgün	Hacettepe Üniversitesi
Prof. Dr. Sermin Onaygil	İstanbul Teknik Üniversitesi
Prof. Dr. Tayfun Akgül	İstanbul Teknik Üniversitesi
Doç. Dr. Ahmet Koltuksuz	Yaşar Üniversitesi
Doç. Dr. Osman Abul	TOBB Ekonomi ve Teknoloji Üniversitesi
Prof. Dr. Ece Olcay Güneş	İstanbul Teknik Üniversitesi
Prof. Dr. Hülya Gökarp	Ondokuz Mayıs Üniversitesi
Prof. Dr. Bilal Gümüş	Dicle Üniversitesi
Prof. Dr. Alkan Alkaya	Mersin Üniversitesi
Doç. Dr. Ömer Melih Gül	İstanbul Teknik Üniversitesi
Doç. Dr. Mustafa Berke Yelten	İstanbul Teknik Üniversitesi
Doç. Dr. Oğuzhan Ceylan	Marmara Üniversitesi
Doç. Dr. Mete Çubukçu	Ege Üniversitesi
Bora Güngören	Portakal Teknoloji
Fikret Küçükdeveci	TEPA
Hasan Ali Pazar	Siemens

İÇİNDEKİLER / CONTENTS

Kablosuz Algılayıcı Ağlar (KAA) ile Tasarlanmış Afet Uyarı Sistemleri ve Öncelikli Araçların Afet Sonrası Geçişlerinin Kolaylaştırılması Tasarımı <i>Disaster Warning Systems Designed with Wireless Sensor Networks (WSN) and Design of Facilitating Passage of Priority Vehicles After Disasters</i> Merve Babaoğlu Derleme Makalesi	7
Stub Yapılı Kuplaj Tabanlı Mikrodalga Sensör ile Motor Yağı Kullanım Ömrü Tespiti <i>Detection of Engine Oil Service Life Using a Stub-Structured Coupled-Line-Based Microwave Sensor</i> Berkay Pehlivan, Mustafa Yıldırım, Mahmut Ahmet Gözel Araştırma Makalesi	15
Extracting Sparse Periodic Fault Signatures from Noisy Signals in Rolling Element Bearings <i>Gürültülü Rulman Sinyallerinden Seyrek ve Periyodik Arıza Verilerinin Çıkarılması</i> Aziz Kubilay Ovacıklı Araştırma Makalesi	25
A Hybrid Stochastic-Robust Optimization Framework for Industrial Energy Management under Renewable Generation and Price Uncertainty <i>Yenilenebilir Enerji Üretimi ve Fiyat Belirsizliği Altında Endüstriyel Enerji Yönetimi için Hibrit Stokastik-Sağlam Optimizasyon Çerçevesi</i> Erhan Baran Araştırma Makalesi	37
Adaptive Signal Source Estimation Using Eigenvalue Analysis and Optimized Dynamic Thresholding <i>Uyarlamalı Sinyal Kaynağı Tahmini: Özdeğer Analizi ve Optimize Edilmiş Dinamik Eşikleme Yaklaşımı</i> Gökhan Yüksek, Timur Lale Araştırma Makalesi	49
Machine Learning for Physiological-Data-Based Driver Monitoring Systems: A Comparative Review and Experimental Case Study on Classification and Regression Approaches <i>Fizyolojik Veriye Dayalı Sürücü İzleme Sistemleri İçin Makine Öğrenmesi: Sınıflandırma ve Regresyon Yaklaşımlarına Yönelik Karşılaştırmalı Bir Derleme ve Deneysel Örnek Olay Çalışması</i> Barış Aykent Araştırma Makalesi	61
Giriş Gecikmeli Robot Manipülatörlerinin Görev Uzayında Denetimi: Kendi Kendini Uyarlayan Bulanık Mantık Tabanlı Bir Yaklaşım <i>Task-Space Control of Robot Manipulators Subject to an Input Delay: A Self-Tuning Fuzzy-Logic-Based Approach</i> Serhat Obuz Araştırma Makalesi	71
LLM-as-a-Judge for Multilingual Question Answering Evaluation: An Experimental Study on an English–Turkish Bilingual SQuAD Dataset <i>Çok Dilli Soru-Cevap Değerlendirmesinde Hakem Olarak LLM: İngilizce–Türkçe İki Dilli SQuAD Veri Kümesi Üzerine Deneysel Bir İnceleme</i> İlkay Yelmen Araştırma Makalesi	81

Kablosuz Algılayıcı Ağlar (KAA) ile Tasarlanmış Afet Uyarı Sistemleri ve Öncelikli Araçların Afet Sonrası Geçişlerinin Kolaylaştırılması Tasarımı

Disaster Warning Systems Designed with Wireless Sensor Networks (WSN) and Design of Facilitating Passage of Priority Vehicles After Disasters

Merve Babaoğlu

Bilgisayar Teknolojileri Bölümü, Meslek Yüksek Okulu, Mardin Artuklu Üniversitesi, Mardin

 0000-0003-3030-8690

mervebabaoğlu@artuklu.edu.tr

Özet

Afetler, doğal yollarla veya insan eliyle gerçekleşebilen ve tehlike yaratan olaylardır. Doğal afetler, doğadaki canlıları ve yapıları ciddi şekilde etkileyebilir. Ülkeler, oluşabilecek zararları azaltmak için yatırımlar yaparak afet öncesi önlemler alabilir ve kayıpları en aza indirebilir.

Afet öncesi, sırası veya sonrasında kullanılacak erken uyarı sistemleri, önlemlerin hızlı alınmasını kolaylaştırır. Geçmişte farklı teknolojiler kullanılmış olsa da bağlantı kopuklukları ve bilgi iletimindeki gecikmeler Kablosuz Algılayıcı Ağlar (KAA) kullanımını artırmıştır. Bu çalışmada, farklı afet türleri için tasarlanmış teknolojiler incelenmiş ve kavramsal bir model önerilmiştir. Önerilen modelde, IR sensörlerle acil geçiş araçlarının kavşaklardaki mesafesi takip edilerek trafik ışıkları otomatik yönetilmektedir.

Bu sistem, afet sonrası veya normal durumlarda acil araç geçişlerini kolaylaştırmayı amaçlamaktadır. Çalışma bir fikir ve tasarım niteliğindedir; henüz gerçek ortamda uygulanmamıştır. Gelecekte simülasyon ve prototip çalışmalarıyla test edilebilir.

Anahtar kelimeler: Kablosuz Algılayıcı Ağ, Sensör, Acil Geçiş, Afet Erken Uyarı Sistemleri

Abstract

Disasters are hazardous events that may occur due to natural processes or human activities. Natural disasters can have severe impacts on living organisms and structures. Countries can take preventive measures before disasters occur by making investments to reduce potential damage and minimize losses.

Early warning systems, which can be used before, during, or after a disaster, facilitate the rapid implementation of necessary measures. Although various technologies have been used in the past, communication disruptions and delays in information transmission have increased the use of Wireless Sensor Networks (WSNs). In this study, technologies designed for different types of disasters are examined, and a conceptual model is proposed. In the proposed model, infrared (IR) sensors monitor the distance of emergency response vehicles from intersections, and traffic signals are automatically controlled accordingly. The system aims to facilitate the passage of emergency vehicles both in the aftermath of disasters and under normal conditions.

This study is conceptual and design-oriented; the proposed system has not yet been implemented in a real-world environment. In future studies, the proposed model can be tested and evaluated through simulations and prototype implementations.

Keywords: Wireless Sensor Networks, Sensors, Emergency Vehicle Priority, Disaster Early Warning Systems

1.Giriş

İnsan yaşamını veya doğadaki tüm canlıları etkileyebilecek, doğal ve/veya fiziksel yapıları etkileyen bir olguya sahip, hayatı durma noktasına getirebilecek yapıda, insan eliyle veya kendiliğinden gerçekleşen olayların tümüne afet denmektedir. Afet öncesinde, afet sırasında ve afet sonrasında yaşanabilecek sıkıntılar, alınabilecek önlemler her daim insanoğlunu ilgilendirmiş ve bu yönüyle toplumsal ve akademik çalışmalara konu olmuştur.

Afetler oluşmadan önce alınabilecek önlemler, hiçbir şekilde zarara uğranmamasını veya zararları asgari düzeye düşürmeyi sağlayabilirler. Sıkça afet yaşayan veya risk potansiyeli yüksek olan ülkeler bu manada alacakları kararlar ve yapacakları uygulamalar ile hem bu risk faktörlerini ortadan kaldırırlar hem de kalkınma düzeylerini yükseltmektedirler.

Türkiye, en başta deprem olmak üzere; yangın, sel, çığ ve heyelan gibi birçok insan kaynaklı veya doğal afetin yaşandığı ve canlı ölümü açısından dünyada en riskli bölgeler arasında yer almaktadır [1].

Afetler insani veya doğal sebeplerden dolayı oluşabilmektedirler. İnsani bir etki olmadan, evrenin kendi düzeni içerisinde gerçekleşen ve birden fazla türü bulunan afetler doğal afetlerdir. Bunlar meteorolojik, biyolojik, iklimsel ve jeolojik vb. yapıda görülebilmektedir.

Gerçekleşen doğal afetlerin büyük bir bölümü de meteorolojik kaynaklıdır [2]. İnsan kaynaklı afetler ise dikkatsizlik veya yanlış bir davranış sonucu yine hayatı tehlikeye yol açabilecek durumların ortaya çıkması durumudur. Doğal veya insani afetlerin neler olduğu Tablo1.'de gösterilmiştir.

Tablo 1. Afet Türleri.

Afet Türü	Örnekler
Doğal Afetler – Jeolojik Kökenli	Toprak kayması, deprem, yanardağ patlaması, tsunami vb.
Doğal Afetler – Meteorolojik Kökenli	Dolu, erozyon, yıldırım, taşkın, kuraklık, hortum, çığ, hava kirliliği, sel vb.
İnsan Kaynaklı Afetler	Kazalar, Biyolojik Saldırıları, Terör Olayları vb.

Afetlere müdahale; can ve mal kayıplarının en aza indirilmesi, acil durumda bulunan hasta ve yaralıların sağlık hizmetlerinden hızlı bir şekilde yararlanabilmesi, güvenlik ve barınma ihtiyaçlarının karşılanması ve afet bölgesinin daha etkin bir şekilde tanınabilmesi amacıyla haritalandırma çalışmalarının yapılması gibi çeşitli önlemler alınarak gerçekleştirilmektedir. İyileştirme ise, afet sonrası dönemde, yaşanan aksaklıkların saptanması, hataların giderilmesi ve düzeltme yoluna gidilmesi durumudur. Ayrıca afet bölgesinde gerçekleşmiş yapısal ve psikolojik bozulmanın devlet eliyle düzeltilmeye çalışılmasıdır. Ancak, afet hazırlık aşamaları daha çok kalkınmış ülkelerde gerçekleşmekte, kalkınmamış veya kalkınmakta olan ülkelerde benzer biçimde gerçekleşmediği için zararlar çok daha fazla meydana gelmektedir [2][3]. Hiç şüphesiz afet erken uyarı sistemlerine sahip bir ülkede kalkınma daha hızlı gerçekleşecek ve karşılaşılabilecek riskler bertaraf edilebilecektir.

Bu çalışmanın temel katkısı, afet sonrası merkezi haberleşme altyapısında yaşanabilecek kesintiler dikkate alınarak kavşak bazlı çalışan düşük gecikmeli bir acil araç öncelik sistemi önerilmesidir. Literatürde bulunan GPS veya RF tabanlı sistemlerden farklı olarak önerilen yapıda, yönlü algılama sağlayan IR sensörler ile Kablosuz Algılayıcı Ağ altyapısının birlikte kullanılması hedeflenmiştir.

Bir sonraki bölümde kablosuz algılayıcı ağ (KAA) hakkında bilgi verilecek daha sonra KAA ile tasarlanmış bazı afet erken uyarı sistemleri gösterilecektir.

Üçüncü bölümde ise afet sonrasında acil geçiş araçlarının ulaşımını kolaylaştırıp hızlandıracak Kızılötesi (IR) Teknolojisi ile tasarlanmış ve KAA altyapısı ile araç bilgisinin gerekli merkeze iletimini sağlayan bir yapı önerilecektir.

2. Kablosuz Algılayıcı Ağ (KAA)

KAA kavramı ilk kez 1980’lerin başlarında ortaya çıkmıştır. İlk başta özellikle askeri alanda kullanılan kablosuz algılayıcı (sensör) ağları; maliyetlerinin gün geçtikçe düşmesi ile yaygın bir biçimde kullanılmaya başlamıştır [4]. Kablosuz algılayıcı ağlar, bulundukları ortamın fiziki koşullarını izleyip kaydetmek ve/veya bilgileri merkezi bir ortamda düzenlemek için oluşturulmuş ve dağıtılmış özel sensör manasına gelmektedirler. Isı, gaz, kirlilik, nem, basınç, titreşim, mesafe vb. durumları ölçebilmektedirler. Böylelikle ortamdaki uzakta, ortamla ilgili bilgiler anlık olarak alınabilmekte ve gerekli adımlar daha hızla atılabilmektedir. Kablosuz algılayıcı ağların kullanım yerleri Tablo 2’de gösterilmiştir.

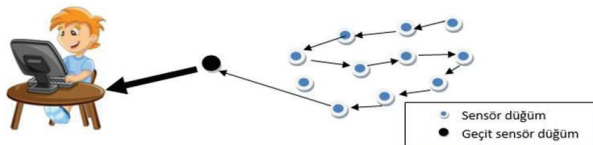
Teknolojide gerçekleşen değişim ve gelişmeler, ayrıca sağlamış olduğu kolaylıklardan dolayı KAA teknolojileri giderek artmaktadır. Habitat izleme, çevre ve sağlık alanında yoğunlaşmıştır. Ayrıca eğitim, endüstri, tarım, trafik gibi alanlarda yaygın olarak kullanılmaya başlamış ve hemen hemen tüm sektörlerde tercih edilmeye başlamıştır [5].

Kablosuz algılayıcı ağlar, profesyonel bilgiye sahip olmadan kolayca kurulabilme, güvenilir sonuçlar verebilme, esnek ve maliyet verimliliği gibi sağladığı birçok fayda ile farklı alanlarda tercih edilmektedir.

Tipik bir KAA, kablosuz bir ortamda birbirine bağlı ve birbiriyle bilgi alışverişi yapan yüzlerce hatta binlerce sensör düğümünden oluşur [6].

Sensör düğümleri bir film teneke kutu içerisinde kendi pili, sensör panosu (board), mikro kontrolörü ve RF adaptörü ile tümleşik bir yapı oluşturur [7].

Sensörler genişçe bir alana dağıtılarak bir adhoc ağ oluşturulur. Bu ağ sayesinde birbirleri ile iş birliği yapan sensörler bilgiye her an, kolayca ulaşılabilmesini sağlarlar. Sonunda ağ, zeki bir ortam oluşturulmuş olur. Bir kablosuz ağ yapısı Şekil 1’te gösterildiği gibidir. Bir kablosuz düğümün genel yapısı ise Şekil 2’de gösterilmiştir.

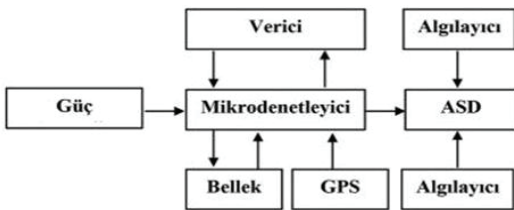


Şekil 1. Kablosuz Algılayıcı Ağ Yapısı

Tablo 2. Kablosuz Algılayıcı Ağ Kullanım Alanları.

Algılayıcı Ağ	
Takip	İzleme
Doğa Takibi (Hayvan takibi, Orman takibi vb.)	Doğa (Hayvan nesli izleme, Jeolojik ölçümler vb.)
Askeri (Düşman takibi vb.)	Askeri (Güvenlik, yasadışı giriş izleme vb.)
Kamu/Endüstriyel (Araç takibi, Eleman takip vb.)	Ticari (Envanter kontrolü vb.)
Ticaret (Müşteri takip vb.)	Kamu/Endüstriyel (Kimyasal atık izleme, Fabrika izleme vb.)
Otomasyon (Akıllı ev uygulamaları vb.)	Çevre (Kimyasal Gaz izleme, Orman yangını izleme vb.)
	Sağlık (Hasta takip vb.)

Ağın kendi yapısında, bütün algılayıcılara komut iletmek ve algılayıcı düğümlerden bilgi almak adına bir toplama istasyonu bulunmaktadır [8]. Algılamaya ek olarak düğümler, diğer bütün düğümlerin elde ettikleri bilgileri toplama istasyonuna iletmek görevini de gerçekleştirmektedirler [9].



Şekil 2. Kablosuz Düğümün Genel Yapısı [10].

2.1. Kablosuz Algılayıcı Ağlar ile Tasarlanmış Afet Erken Uyarı Sistemleri

Hangi tip afet olursa olsun insan odaklı olmak ve her açıdan güvenliği sağlamak ülkelerin tartışmasız ilkelerinden biri olmalıdır. Özellikle doğal afetlerin yaşama olan etkisi ve yarattığı tehlike, bu manada erken uyarı sistemlerin önemini göstermektedir.

Afet erken uyarı sistemlerinin hazırlanması sayesinde insanlar hızlı hareket edip önlem alabilir, cihazlar ve ihtiyaç duyulacak malzemeler korunabilmektedir. Afet yönetiminde acil durumların müdahalesinde hız çok önemli bir etkidir. Bundan dolayı zarar gören iletişim altyapılarının tamir edilmesinden önce, farklı iletişim sistemlerinin tasarlanması ve kullanımının sağlanması daha uygundur. [11] [12].

Afet anında; iletişim sistemleri çıktığında afetzedelerin konumlarının belirlenmesi veya saha personeli ile haberleşmek belki de en büyük sorun olmaktadır. Doğru bilgiye doğru zamanda erişim birçok hayat kurtarabilir. Konum tabanlı ve konum altyapılı sistemler bu sorunun çözülmesinde önemli rol oynayabilir. Uydu sistemleri sayesinde elde edilecek konum bilgisi ve iletişim sayesinde doğru bilgiye en hızlı ve en doğru şekilde ulaşmak mümkün olmaktadır [13].

Türkiye özelinde, jeolojik olarak genç bir ülkedir. Bu nedenle sürekli olarak doğal afet tehlikesiyle karşı karşıyadır. Ülkemizde sel, heyelan, orman yangınları, deprem gibi birçok afet görülmektedir. Afetlerin oluşturacağı can kaybı, ekonomik ve psikolojik sorunlar ciddi sıkıntı yaratmaktadır. Bu türden problemleri ülkemizde yaşamamak, yaşanmadan önce haberdar olmak ve/veya en az zararla atlatabilmek adına kablosuz sensör ağ yapılarına çokça yatırım yapılmalı ve takibine önem verilmelidir.

Sensörler deprem zamanında veya metrolarda da kullanılabilirler. Mesela bir metroda herhangi bir nedenle mahsur kalmış insanların nerede mahsur kaldıkları verisi bilgisi yardım ekiplerine gönderilebilir. Deprem anında, çöken binalarda insanların tahmini olarak nerede veya hangi yükseklikte oldukları bilgisi yine ekiplere bildirilebilir. [14]. Farklı afet türlerine göre geliştirilmiş, kablosuz sensör ağlar kullanılarak oluşturulmuş bazı erken uyarı sistemleri bir sonraki bölümde gösterilecektir.

2.1.1. Orman Yangını Algılama Sistemleri

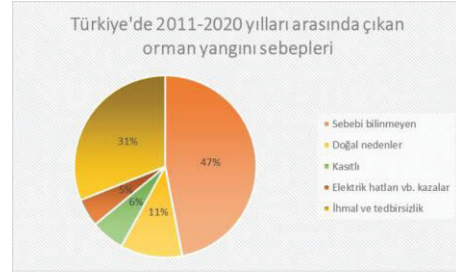
Ormanlar ülkelerin hazineleridirler. İçerisinde yaşayan canlılara besin ve yuva, insanlara da oksijen, gıda, yakıt deposu ve daha sayılamayacak nice faydaları vardır. Böylesine değerli yapıların yangınlarla yok olmasını engellemek adına çok sayıda orman yangını uyarı sistemi geliştirilmiştir. Bu yangınlar meteorolojik kaynaklı olabileceği gibi insan hatasından dolayı da gerçekleşebilmektedir.

Türkiye’de 2011-2020 yılları arasında çıkan orman yangınlarının sadece %11’inin doğal nedenlerden kaynaklandığı gözükmemektedir. %89’luk büyük bir kısmı ise insan kaynaklı çıkmıştır [15].(Şekil 3).

Kablosuz sensör ağları, yangınların tespiti için önemli olan, CO, CO₂, ortam nem miktarı, ortam sıcaklığı gibi değerlerde ölçülmekte ve ilgili merkeze bildirmektedir. Böylelikle beklenmedik değerler ölçüldüğünde hızla hareket edilebilir ve herhangi bir sıkıntı durumunda müdahale edilebilir.

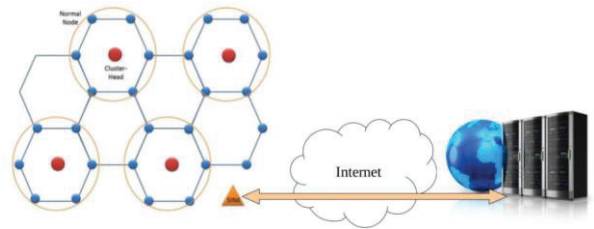
Orman yangınlarının özellikle ilk 5 dakikalık sürede fark edilmesi çok önemlidir. Sonraki 15-20 dakikada gerekli müdahalede bulunulamayan yangınlar çok geniş alanlara yayılarak kontrol edilmesi güçleşebilmektedir. Kurulan sistemler ile orman yangınlarının daha henüz başlangıç aşamasında tespit edilmesi ve yangınla mücadele ekiplerinin bilgilendirilmesinin ardından acil müdahalede bulunulması ve yangının ilerlemesinin önüne geçilmesi mümkün kılınmaktadır [16].

Şekil 4’de gösterilen kablosuz sensör ağında orman yangınlarını erken fark etmek ve müdahalede bulunabilmek adına oluşturulan sistemde örgü ağ topolojisi kullanılmış ve kablosuz cihazlara bağlı sensörler, ormanlık alanlarda ölçüm yaparak olası bir tehdidi algılayıp merkeze iletmektedir. Ayrıca ormanın bulunduğu alanın hava şartları, nem durumu gibi bilgilerde takip edilmektedir.



Şekil 3. Türkiye’de orman yangını çıkış sebepleri [15].

Orman yangınlarını önceden tespit edebilmek için kullanılan bir başka kablosuz sensör ağ sistemi olan FireBugs, yangın ortamlarında real time veri toplamak için ağlar içinde kendisini organize eden TinyOS işletim sistemine dayanan küçük KAA sistemidir. Bu sistemde ortamın davranışlarını izleyen ve bunu internet üzerinden gönderen bir teknoloji ve kundakçı algılayıcı teknolojisi kullanılmaktadır [17].



Şekil 4. Orman yangını erken uyarı sistemi işleyişi [16].

2.1.2. Deprem Algılama Sistemleri

Türkiye deprem afeti açısından dünyanın en tehlikeli yerleri arasında bulunmakta ve “yüksek risk” grubu içinde yer almaktadır. En tehlikeli ve yıkıcı faylar, kalabalık şehirlerimizden geçmektedir. Dolayısıyla ülkemizde oluşturduğu can ve mal kaybı oran olarak da fazla olmaktadır. Bu nedenle özellikle yapıların bu durum gözetilerek oluşturulması, takibinin yapılması, olası iletişim kopukluğu için hazırlıkların yapılması elzemdir.

Afet uyarı sistemlerinin önemi 6 Şubat 2023 Kahramanmaraş-Pazarcık depreminde bir kez daha görülmüştür. Kahramanmaraş’ta gerçekleşen depremin 50 saniye sonra diğer illerde yıkıcı etki yarattığı bilgisi başkanlığını Prof. Dr. Tuncay Taymaz’ın yaptığı ve aralarında Kocaeli Üniversitesi’nden Doç. Dr. Serkan İrmak’ın da bulunduğu ekip tarafından belirlenmiştir [18]. Dolayısıyla, Kahramanmaraş dışındaki şehirlere zamanında bilgi aktarılmasını sağlayan bir uyarı sistemi, olası bir afetin etkilerinin ve meydana gelebilecek can ve mal kayıplarının azaltılmasına katkı sağlayabilirdi.

Depremin büyüklüğü ve gerçekleştiği yönü tespit ve bina sağlamlığının değerlendirilmesi için bir binaya entegre edilebilen sismik algılayıcılar, afet öncesinde uyarı verebilmektedirler [19].

Başka bir sistem ise algılayıcıları konum algılamak için insanlar tarafından giyilen cep algılayıcıları, GPS alıcılarından oluşmaktadır. Kişinin konum hareketi ve konum bilgileri iletilmektedir [20].

Dünyada deprem konusunda en fazla zorluk çeken ülkelerin başında hiç şüphesiz Japonya gelmektedir. Japonya’da çokça kullanılan erken uyarı (Earthquake Early Warnings-EEW) sistemi, deprem bildirimi konusunda Japonya’da kullanılan en iyi uygulamadır. 1 Ekim 2007’den bu yana kullanılan bu uygulama Japonya genelinde birden fazla cihaz aracılığıyla çalışır durumda olup güçlü olma olasılığı yüksek olan sarsıntılar gelmeden önce uyarı sağlamaktadır.[21]. Suzuki ve diğ. [36], afet sonrası ortamlarda özellikle enkaz altında kalanlar için yer altı alanlarında kullanılabilen ve kablosuz algılayıcı ağ aracılığı ile bilgilerin toplandığı mobil robot sistemi geliştirdiler.

Depremde erken uyarı sistemleri, insanlara kendilerini koruyabilecekleri ve güvenli bir alana geçebilecekleri kıymetli birkaç saniye vermektedir.

2.1.3. Sel / Taşkın Algılama Sistemleri

Türkiye’de gerçekleşen depremlerden sonra en çok can ve mal kaybına neden olan doğal afet olan sel/su baskınları nedeniyle her yıl en az 100 milyon dolar ekonomik kayıp yaşanmaktadır [22].

Fazla yağış, kar erimesi, baraj patlaması vb. durumlardan kaynaklanan, su yataklarının kapasitesini aşması ve sonrasında neden olan seller, su basmalarına veya taşkınlara sebep olmaktadır. Bu taşkınlar canlı yaşamına, yerleşim yerlerine ve tarım arazilerine ciddi boyutlarda zarar verebilmektedir.

Akarsu ve derelerdeki taşkını öncesinden haber verecek birtakım sistemler geliştirilmiştir. Bunlardan bir tanesi Devlet Su İşleri Genel Müdürlüğü tarafından geliştirilen, TEFER (Taşkın Tahmini ve Erken Uyarısı) projesidir. Bu proje radar sensörlü gözlem istasyonu ile dere yataklarındaki su seviyesini anlık (real time) olarak izleyip, yaklaşık bir saat öncesinde bölge halkına haber vererek can ve mal kaybını önlemektedir [23]. (Şekil 5).

Afet yönetimi konusunda Fang ve diğerleri [24], önerdikleri sel/taşkın erken uyarı sistemi yönteminde web servisleri, akıllı algılayıcılar ve uzaktan algılama görüntüleri kullanarak oluşturdukları modelin prototipini Çin’in Xinjiang şehrinde yer alan Quergou nehrinde meydana gelen kar erimesi sonucunda oluşan değişimleri izlemeye başarı bir şekilde uygulamışlardır.



Şekil 5. TEFER projesi uygulaması [23].

Elman diye adlandırılan bir başka sistemde su sistemlerinin sebep olduğu taşkın durumlarının önlenmesini sağlamak için kullanılmaktadır. Tehlikeli olabilecek noktalara yerleştirilen bu cihaz sahip olduğu sensörler sayesinde bu noktalardan geçen su miktarının tespitini gerçekleştirir ve anlık olarak bilgilendirme yapar. Acil durumlarda ise halka gerekli bilgilendirmeyi farklı cihazlardan gerçekleştirebilmektedir [25]. (Şekil 6)



Şekil 6. ELMAN Taşkın Bildirim Sistemi [25].

2.1.4. Heyelan Bildirim Sistemleri

Heyelan algılama sistemleri heyelan esnasında olabilecek en küçük toprak hareketlerini algılayabilmek için bir KAA kullanılmaktadır. Gerekli tüm verileri toplayarak uzun heyelan oluşumu gerçekleşmeden önce bu bilgilerin iletilmesini anlık olarak sağlamaktadır.

Bozkurt & Erenoğlu [26], çalışmalarında, heyelanın mühendislik yapılarını etkileyerek oluşabilecek kazaların önüne geçilebilmesi için bulut bilişim ve nesnelerin interneti tercih edilerek bir erken uyarı sistemi tasarlanmıştır. Bunun için ülkemizde heyelan oluşma riski fazla olan Edirne ili Uzunköprü ilçesi Demirtaş Köyü yakınlarında, 41°17'45.0"N, 26°41'53.0"E coğrafi koordinatlarını tercih etmişlerdir. Mühendislik yapısı olarak da demiryolu kullanılmış ve burada oluşan mesafe ölçümleri sürekli izlenmiş, sonuçlar Google Drive üzerinden kaydedilmiştir. Sistem geliştirilerek farklı yapılarda da test edilebilir ve kullanışlı bir sistem geliştirilebilir.

Sofwan ve ark. [27], nesnelerin interneti kullanarak heyelana neden olan ve birçok fiziksel parametreyi ölçebilen bir KAA yapısını başarıyla geliştirilmiştir.

Frigerio ve ark. [35], Rotolon heyelanının (Doğu İtalya Alpleri) otomatik ve sürekli izlenmesi için web tabanlı bir platformun tasarımı gerçekleştirmişlerdir.

2.1.5. Maden Afet Uyarı Sistemleri

İş güvenliği açısından yer altı madencilik ciddi tehlikeler barındırabilen ve iş güvenliği açısından önem verilmesi gereken alanlardan biridir. Madencilikte tehlike yaratabilen ve takip edilmesi gereken durumlar başlıca aşağıdaki gibidir:

- Yangın
- Çökme
- Gaz Sıkışması
- Su Baskını
- Toz

Yukarıda belirtilenler ve bunların dışında ortaya çıkabilecek birçok tehlikenin önceden belirlenmesi belki de dünyanın bu en riskli çalışma ortamlarını güvenli hale getirmekte, can ve mal kaybına engel olmaktadır. Kablosuz yeraltı algılayıcı ağ yapısı son zamanlarda özellikle yurtdışındaki madenlerde çokça tercih edilmeye başlamıştır. Bu teknik kullanılarak yeraltı kömür ocaklarında uygun yerlere algılayıcı sensörler yerleştirilerek çevresel parametrelerin (nem, hava akımı, sıcaklık, maden gazları) izlenmesi ve maden içinde oluşabilecek göçük, yangın ve patlama gibi olası tehlikelerin varlığı önceden tespit edilebilmiştir [28].

Özellikle kömür madenlerinde KAA uygulaması ile izlenebilen çevresel parametreler 8 başlık altında toplanmaktadır [29]. Bunlar;

1. Ocaktaki çeşitli patlamaların izlenmesi
2. Sıcaklık ve nemin izlenmesi
3. Personel ve makine ekipmanın izlenmesi
4. Göçüklerin izlenmesi
5. Ocak yangınlarının izlenmesi
6. Zararlı gazların izlenmesi
7. Tozların (tas ve kömür tozları) izlenmesi
8. Hava basıncı veya hava akım hızının izlenmesi

Gaz sensörleri ise belli bir bölgede mevcut olan farklı gazların varlığını tespit etmek için kullanılır. Sensörler belirlenecek gazın yoğunluğuna bağlı olarak orantılı bir elektriksel tepki verirler. Eğer izlenen gazın yoğunluğu önceden belirlenmiş eşik değerleri aşarsa sensör/sensörlerin kullanıldığı alete, yakındaki personele alarm vermekte veya ekipmanın güç kaynağını kapatma gibi acil eylemleri gerçekleştirebilmektedirler [30].

Özellikle yer altı bildiri sistemlerinde mesafelerin nasıl ölçüldüğüne göre lokalizasyon (yer belirleme) algoritmaları menzile dayalı ve menzile dayalı olmayan şekilde iki kategoriye ayrılır:[33]. Menzile dayalı yaklaşımlar, bilinmeyen düğümlerin konumunu tahmin etmek için komşu düğümler arasındaki mutlak noktadan noktaya mesafe veya yönlendirme bilgilerini kullanır. Bu teknikler ek donanım gerektirir, böylece onları daha pahalı hale getirir. Bu nedenle, büyük ölçekli KAA'da uygulanamayabilirler. Menzile dayalı yaklaşımlardan farklı olarak, menzilsiz yaklaşımlar mesafe veya açı bilgisi gerektirmez. Bu nedenle, ek maliyetli donanım kullanmazlar, bu da onları daha uygun maliyetli ve daha basit alternatifler haline getirir. [34].

3. Materyal ve Yöntem (Material and Method)

3.1. Önerilen Sistemin Farklı Sistemlerle Karşılaştırılması

Çalışmada, bir önceki bölümde anlatılan sistemler dışında, afet sonrası veya acil durumlarda müdahale öncesi ve sonrasında öncelikli araçların geçişlerini kolaylaştırmak adına WSN ağları kullanılması düşünülmüştür. İtfaiye, ambulans, emniyet araçları ve AFAD gibi acil müdahale ekiplerinin özellikle büyük ve kalabalık şehirlerde kavşak noktalarında yoğun trafik ve kırmızı ışık sürelerinden ötürü yaşadıkları gecikmenin en aza indirgenmesi amaçlanmıştır. Bu amaçla birden fazla yöntem kullanılabilir. Literatürde çokça tercih edilen farklı yöntemler ile önerdiğimiz yöntem karşılaştırması Tablo 3'te verilmiştir.

Önerilen IR teknolojisi düşük gecikme süresi, düşük enerji tüketimi ve hızlı tepki avantajları sunmasına rağmen; haberleşme mesafesi ve doğrudan görüş hattı (Line of Sight) gereksinimi gibi bazı teknik sınırlılıklar barındırmaktadır. Özellikle büyük ölçekli yol ve kavşak sistemlerinde bu durum sensör sayısının artmasına ve sistem yönetim karmaşıklığının yükselmesine neden olabilmektedir.

Tablo 3. Farklı Sistemlerle Karşılaştırma.

Sistem	Teknoloji	Avantaj	Dezavantaj
Optik Sistem	IR/Optik	Düşük maliyet, hızlı tepki	Görüş hattı gerekir, hava koşullarından etkilenir
RF Sistem	Radyo Frekansı	Uzun menzil, engel tanımaz	Parazit riski, maliyet yüksek
GPS Tabanlı Sistem	GPS + GSM	Araç takibi, merkezi kontrol	Gecikme olabilir, altyapıya bağımlı
Önerilen Sistem	KKA + IR	Düşük gecikme, düşük enerji tüketimi, gerçek zamanlı haberleşme	Kısmi görüş hattı gerektirir

Tablo 3'te gösterildiği gibi her sistemin kendine göre avantaj ve dezavantajları bulunmaktadır.

Bununla birlikte IR tabanlı yapı, çalışmada önerilen sistem için temel algılama mekanizması olarak değerlendirilmiş olup; kısa mesafeli ve gerçek zamanlı acil araç tespiti açısından avantaj sağlamaktadır. Ancak afet sonrası haberleşme sürekliliği, uzun menzil ihtiyacı ve görüş hattı problemlerinin azaltılması amacıyla sistemin gelecekte hibrit haberleşme altyapıları ile desteklenmesi planlanmaktadır.

Bu kapsamda:

- ZigBee ,
- 433 MHz RF haberleşme sistemleri,
- LoRa,
- NB-IoT ,
- ve GSM/LTE tabanlı hücresel IoT

altyapılarının sisteme entegre edilmesiyle daha geniş kapsama alanı, düşük altyapı bağımlılığı ve daha güvenilir veri iletimi sağlanabileceği değerlendirilmektedir.

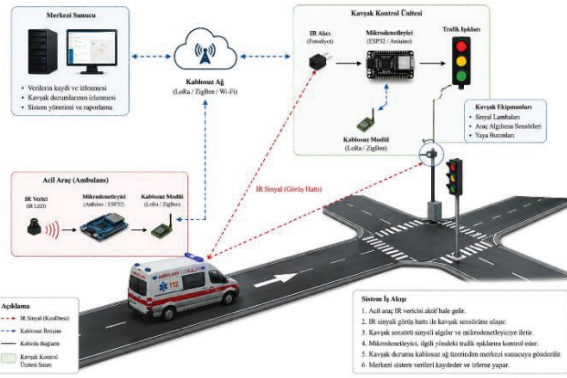
Böylece önerilen sistemin yalnızca IR tabanlı değil, çok katmanlı ve ölçeklenebilir bir akıllı ulaşım altyapısına dönüştürülebileceği öngörülmektedir.

3.2. Önerilen Sistem Yapısı

Tercih ettiğimiz sistemde araçtaki verici ile direklerdeki alıcılar çeşitli teknolojiler kullanılarak birbirleri ile iletişim kurabilirler.

- HGS sistemlerinde kullanılan Proximity veya Mifare teknolojisi,
- Kızılötesi (IR) teknolojisi,
- Bluetooth ve Wifi gibi WSN alıcı ve verici teknolojisi.

Acil durum araçlarında önceden kurulacak olan acil durum bildirim sensörleri ile yaklaşılan kavşaktaki alıcılara sinyal gönderilerek ilgili kavşaktaki trafik ışıklarının gelen acil müdahale aracı yönünde yeşile döndürülerek geçişinin kolaylaştırılması planlanmaktadır. Çalışmanın yapısal özeti anlatan görsel Şekil 7'deki gibidir.



Şekil 7. Önerilen IR Tabanlı acil geçiş öncelik sisteminin genel mimarisi.

Çalışmada, Wifi, Bluetooth ağları, Proximity veya Mifare teknolojileri ile yapılacak sistemlerde kavşaktaki alıcı sensörler arasındaki mesafeler yakın olabileceğinden verilen sinyallerin kavşaktaki diğer algılayıcılar tarafından algılanıp karışıklığa sebep olacağı düşünüldüğünden, tek yönlü iletişim kurabilmek adına IR teknolojisi kullanmanın daha sağlıklı sonuç vereceği düşünülmektedir.

3.3. Kızılötesi (IR) Teknolojisi

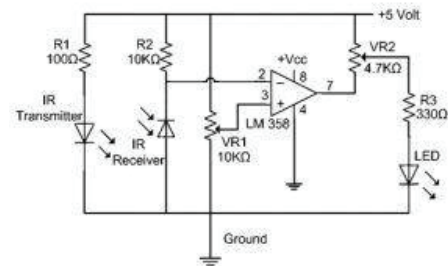
İnfrared (IR) radyasyon ilk kez William Herschel tarafından 1800'lü yıllarda keşfedilmiştir [31]. IR teknolojisi, uzaktan kumanda ve algılamayı içeren çok çeşitli kablosuz uygulamalarda kullanılmaktadır. Bu sensörde verici olarak IR LED, dedektör olarak fotodiyot kullanılır. Fotodiyot üzerine bir kızılötesi ışık düştüğünde, çıkış voltajı ve direnci alınan IR ışık büyüklüğü ile orantılı olarak değişecektir. IR Sensörü çalışma prensibi, bir kızılötesi sensör, verici ve alıcı (verici ve alıcı) olmak üzere iki bölümden oluşur.

Ortaklaşa bir optokuplör veya bir fotokuplör olarak adlandırılır. Burada verici olarak IR LED, alıcı olarak IR fotodiyot kullanılır. Bunda kullanılan fotodiyot, bir kızılötesi LED aracılığıyla üretilen kızılötesi ışığa çok duyarlıdır. Fotodiyot ve çıkış voltajının direnci, elde edilen kızılötesi ışıkla orantılı olarak değiştirilebilir. Bu, temel IR sensörünün çalışma prensibidir [32].



Şekil 8. IR Sensör Örneği

Kızılötesi sensör devresi, bir cihazdaki temel ve popüler sensör modüllerinden biridir. Bu devre Şekil 9 'de gösterildiği gibidir [32].



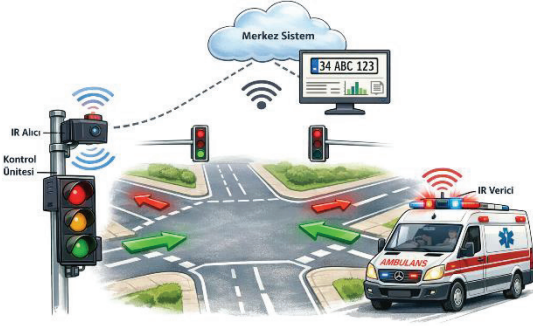
Şekil 9. IR Sensör Devre Şeması.

3.4. Sistemin İşleyişi

Bu çalışmada önerilen sistem, afet sonrası ve acil durum senaryolarında öncelikli araç geçişlerini kolaylaştırmaya yönelik kavramsal bir tasarım niteliğindedir. Öneri kapsamında, acil geçiş önceliğine sahip araçların tepesine monte edilecek IR sensör vericilerinin, görevli personel tarafından aktif edilen tepe lambalarıyla eş zamanlı olarak çalışması ve sürekli sinyal üretmesi öngörülmektedir. Kavşağa yaklaşan aracın gönderdiği sinyal, trafik lambası direklerine konumlandırılmış IR sensör alıcıları tarafından algılanarak işlenecektir.

Bu işleme süreci sonucunda, trafik kontrol kartlarına iletilen veri sayesinde aracın seyir yönündeki trafik ışıklarının uygun şekilde yeşile dönmesi, diğer yönlerde ise trafik akışının kontrollü biçimde durdurulması sağlanabilecektir. Ayrıca sistemin, güvenlik ve denetim amacıyla araç plakası veya tanımlanmış özel kod bilgisini kaydederek kablosuz algılayıcı ağ altyapısı üzerinden merkezi birimlere iletmesi planlanmaktadır. Böylece hem operasyonel kolaylık hem de kayıt altına alınabilirlik sağlanmış olacaktır.

IR sensörlerin hem trafik lambası direklerine hem de acil müdahale araçlarına entegre edilebilmesi, sistemin teknik açıdan uygulanabilirliğini artırmaktadır. Bu çalışma, henüz kavramsal ve tasarım düzeyinde bir öneri olmakla birlikte; geliştirilmesi, prototip üretimi ve saha testleri ile desteklenmesi durumunda özellikle büyük ve yoğun trafikli şehirlerde acil müdahale sürelerini azaltarak önemli ölçüde can ve mal kaybının önlenmesine katkı sağlayabilecek potansiyele sahiptir.



Şekil 9. Acil Durum Trafik Kontrol Sistemi.

Dolayısıyla önerilen model, afet yönetimi ve akıllı ulaşım sistemleri literatürüne uygulanabilir ve geliştirilebilir bir fikir sunmakta olup, gelecekte yapılacak deneysel çalışmalar için temel oluşturabilecek niteliktedir.

IR sensör Şekil 9 da gösterildiği biçimde hem trafik lambası direğine hem de acil geçişli araçlara konumlandırılabilir.

4. Sonuç ve Öneriler

Çalışmada öncelikle Türkiye’de ve dünyada yapılmış olan Kablosuz Algılayıcı Ağ (KAA) uygulamaları incelenmiş ve farklı afetlere göre sınıflandırılarak anlatılmıştır. Amaç, Kablosuz Algılayıcı Ağlar (KAA) afet uygulamaları konusunda temel kaynak oluşturmaktır.

Kablosuz Algılayıcı Ağlar afet uyarı sistemlerinde kullanıldıklarında olası tehlike ve tehditleri engelleyebilme, doğru ve güvenilir sonuca en kısa sürede ulaştırma ve tedbir alınmasını kolaylaştırma konusunda önemli faydalar sağlamaktadır. Ancak KAA’nın en büyük sorunlarından birini maliyet oluşturmaktadır. Türkiye gibi KAA üretimi yapılmayan ve ithal edilmek zorunda kalan bir ülkede bu manada sıkıntı daha büyük olabilmektedir. Doğal veya yapay afet konusunda dezavantajlı bir durumda olan ülkemizde KAA teknolojisi üretim yapılmalı, teşvik edilmeli ve böylelikle afet teknolojileri üretimi kolaylaştırılmalıdır.

Önemli olan bir diğer nokta bu teknolojilerin kurulum ve uygulama aşamalarında görevli olacak çalışanların bu teknolojiler hakkında bilgilendirilmesi, sonraki dönemlerde aksaklık olmaması amacıyla kapasite geliştirme çalışmalarının yapılması durumudur. Ülkemizde gelecekte karşılaşılabilecek durumlar ön görülerek mevcut altyapı güçlendirilmeli ve gelişen teknoloji yakından takip edilip uygulanmasının hızla sağlanması gerekmektedir. Ayrıca yasalarla denetlenebilecek sunucular ile sensörlerden alınan verilerin güvenliği koruma altına alınmalıdır.

Çalışmada, IR teknolojisi kullanılarak öncelikli araçların geçişlerini kolaylaştırmak, oluşabilecek karmaşayı ve gecikmeyi engelleyecek, geliştirilebilir ve gerçek hayatta kullanılabilir ve bu araçlarla ilgili verilerin yine kablosuz algılayıcı ağ sayesinde gerekli mercilere iletim yapısı önerilmektedir. Bu yapı geliştirilip uygulanabilir duruma getirilirse, ülkemizde özellikle büyük ve kalabalık şehirlerde faydalı olacaktır.

5. Kaynaklar

- [1] AFAD, “Türkiye’de Afet Yönetimi ve Doğal Kaynaklı Afet İstatistikleri,” 2018. URL: <https://www.afad.gov.tr/kitaplar> (Erişim tarihi: 2 Kasım 2019).
- [2] M. Kadioğlu, *Afet Yönetimi: Beklenmeyi Beklemek, En Kötüsünü Yönetmek*, Marmara Belediyeler Birliği Kültür Yayınları, İstanbul, 2017.
- [3] F. Şenses, “Sunuş: Afetler ve Kalkınma Konusuna Toplu Bakış,” M. Tiryakioğlu (Ed.), *Afetlerle Kalkınma: Tecrübeler, Politikalar ve Beklentiler*, Efil Yayınevi, Ankara, ss. 17–34, 2016.
- [4] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, “Wireless Sensor Networks – A Survey,” *Computer Networks*, vol. 38, pp. 393–422, 2002.
- [5] T. Soylu, “Kablosuz Algılayıcı Ağların Uygulama Alanları ve Bir Algılayıcı Düğüm Tasarımı,” Yüksek Lisans Tezi, Bilgisayar Mühendisliği Anabilim Dalı, Trakya Üniversitesi, 2012.
- [6] B. Dolay, “Kablosuz Sensör Ağlar,” 2023. URL: <https://e-bergi.com/y/kablosuz-sensor-aglari/> (Erişim tarihi: 11 Mayıs 2023).
- [7] Wikipedia, “Wireless Sensor Network.” URL: https://en.wikipedia.org/wiki/Wireless_sensor_network (Erişim tarihi: 10 Mayıs 2023).
- [8] Y. Shen and H. Ju, “Energy-Efficient Cluster-Head Selection Based on a Fuzzy Expert System in Wireless Sensor Networks,” *2011 IEEE/ACM International Conference on Green Computing and Communications*, Sichuan, pp. 110–113, 2011.
- [9] P. V. S. Reddy, “Fuzzy Conditional Inference and Application to Wireless Sensor Network Fuzzy Control Systems,” *2015 IEEE 12th International Conference on Networking, Sensing and Control*, Taipei, pp. 1–6, 2015.
- [10] A. E. Özdemir, “Orman Yangınlarının Erken Algılanması İçin Kablosuz Algılayıcı Düğüm Teknolojisi,” *TMMOB Elektrik Mühendisleri Odası Dergisi*, Ordu Üniversitesi, Fatsa Deniz Bilimleri Fakültesi, 2021.
- [11] G. Tuna, T. V. Mumcu, and K. Gulez, “Design Strategies of Unmanned Aerial Vehicle-Aided Communication for Disaster Recovery,” *Proceedings of High Capacity Optical Networks and Emerging/Enabling Technologies*, IEEE, Istanbul, Turkey, pp. 115–119, 2012.
- [12] E. Yulianto, P. Utari, and I. A. Satyawan, “Communication Technology Support in Disaster-Prone Areas: Case Study of Earthquake, Tsunami and Liquefaction in Palu, Indonesia,” *International Journal of Disaster Risk Reduction*, vol. 45, 101457, pp. 1–8, 2020.
- [13] Y. Xu, X. Chen, and L. Ma, “LBS Based Disaster and Emergency Management,” *The 18th International Conference on Geoinformatics: GIScience in Change*, Peking University, Beijing, China, June 18–20, 2010.
- [14] S. Doğru Dağın, “Mobil Sensörlerin Afet Anında Kullanım Alanları,” *15. Türkiye Harita Bilimsel ve Teknik Kurultayı*, TMMOB Harita ve Kadastro Mühendisleri Odası, Ankara, 25–28 Mart 2015.
- [15] TRTHABER, “Orman Yangınlarının Yüzde 89’u İnsan Kaynaklı,” 2023. URL: <https://www.trthaber.com/haber/turkiye/orman-yanginlarinin-yuzde-89u-insan-kaynakli-557050.html> (Erişim tarihi: 9 Mayıs 2023).

- [16] Bimetri, "Orman Yangını Erken Uyarı Sistemi," 2023. URL: <https://www.bimetri.com/cozumler/orman-yangini-erken-uyari-sistemi/> (Erişim tarihi: 1 Mayıs 2023).
- [17] K. Sohraby, D. Minoli, and T. Znati, *Wireless Sensor Networks: Technology, Protocols, and Applications*, John Wiley & Sons, Inc., New Jersey, 2007.
- [18] NTV, "Kırılan Fayın Yıkıcı Etkisi Hatay'a 50 Saniye Sonra Ulaşmış," 2023. URL: <https://www.ntv.com.tr/turkiye/kirilan-fayin-yikici-etkisi-hataya-50-saniye-sonra-ulasmis> (Erişim tarihi: 2 Mayıs 2023).
- [19] J. A. Jamalipour, *Wireless Sensor Networks: A Networking Perspective*, Institute of Electrical and Electronics Engineers, New Jersey, 2009.
- [20] J. Yick, B. Mukherjee, and D. Ghosal, "Wireless Sensor Network Survey," *Computer Networks*, vol. 52, no. 12, pp. 2292–2330, 2008.
- [21] HaberTürk, "Japonlar Depremi Erken Uyarı Sistemi ile Önceden Biliyor," 2023. URL: <https://www.haberturk.com/japonlar-depremi-ne-kadar-sure-once-biliyor-2862074> (Erişim tarihi: 13 Mayıs 2023).
- [22] C. Şahin and Ş. Sipahioğlu, *Doğal Afetler ve Türkiye*, 2. Baskı, Gündüz Eğitim ve Yayıncılık, Ankara, 2003.
- [23] Vatan, "Taşkın Erken Uyarı Sistemi Kuruluyor!" 2023. URL: <https://www.gazetevatan.com/gundem/taskin-erken-uyari-sistemi-kuruluyor-hemen-uyarilacaklar-1396338> (Erişim tarihi: 12 Mayıs 2023).
- [24] S. Fang, L. D. Xu, Y. Zhu, J. Ahati, H. Pei, J. Yan, and Z. Liu, "An Integrated System for Regional Environmental Monitoring and Management Based on Internet of Things," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 2, pp. 1596–1605, 2014. URL: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6725615>.
- [25] Elman, "Sel Erken Uyarı Sistemi," 2023. URL: <https://elmantech.com/ehs/> (Erişim tarihi: 15 Mayıs 2023).
- [26] O. Bozkurt and C. Erenoğlu, "Doğal Afetlerin Mühendislik Yapıları Üzerindeki Etkilerinin Yer Bilimleri Tabanlı Disiplinlerarası Bir Yaklaşımla Erken Uyarı Sistemi Tasarımı," *Lâpseki Meslek Yüksekokulu Uygulamalı Araştırmalar Dergisi*, vol. 2, no. 3, p. 77, 2021.
- [27] A. Sofwan, M. Aghus, M. Ridho, and A. Goni, "Wireless Sensor Network Design for Landslide Warning System in IoT Architecture," *2017 4th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE)*, IEEE, 2017.
- [28] J. Zhi-Zhong, D. Simplot-Ryl, and V. C. M. Leung, "Ad Hoc Networks," *Second International Conference, ADHOCNETS 2010*, Victoria, BC, Canada, vol. 49, Springer, 2010.
- [29] L. Muduli, D. P. Mishra, and P. K. Jana, "Application of Wireless Sensor Network for Environmental Monitoring in Underground Coal Mines: A Systematic Review," *Journal of Network and Computer Applications*, vol. 106, pp. 48–67, 2018.
- [30] Y. Aydın and K. Barış, "Yeraltı Kömür Ocaklarında Gaz İzleme Ağları: Sensör Ölçümlerinin ve Konumlarının Doğruluğunun Analizi," *Madencilik*, vol. 54, no. 2, pp. 19–32, 2015.
- [31] C. Skjöldebrand, "Infrared Heating," in R. Richardson (Ed.), *Thermal Technologies in Food Processing*, Boca Raton, Florida, pp. 208–227, 2001.
- [32] FMUSER, "IR Sensör Nedir: Devre Çalışması," 2023. URL: <https://tr.fmuser.net/wap/content/?21028.html> (Erişim tarihi: 8 Mayıs 2023).
- [33] X. Y. Li, L. S. Yan, W. Pan, and B. Luo, "Secure and Robust DV-Hop Safety Localization Based on the Vector Refinement Feedback Method for Wireless Sensor Networks," *The Computer Journal*, vol. 60, no. 6, pp. 810–821, 2017.
- [34] S. Phoemphon, C. So-In, and N. Leelathakul, "Optimized Hop Angle Relativity for DV-Hop Safety Localization in Wireless Sensor Networks," *IEEE Access*, 2018.
- [35] S. Frigerio, L. Schenato, G. Bossi, M. Cavalli, M. Mantovani, G. Marcato, and A. Pasuto, "A Web-Based Platform for Automatic and Continuous Landslide Monitoring: The Rotolon (Eastern Italian Alps) Case Study," *Computers & Geosciences*, vol. 63, pp. 96–105, 2014.
- [36] T. Suzuki, R. Sugizaki, K. Kawabata, Y. Hada, and Y. Tobe, "Autonomous Deployment and Restoration of Sensor Network Using Mobile Robots," *International Journal of Advanced Robotic Systems*, vol. 7, no. 2, pp. 105–114, 2010.

Özgeçmişler



Merve Babaoğlu, lisans eğitimini 2006–2010 yılları arasında Dumlupınar Üniversitesi Bilgisayar Mühendisliği Bölümü'nde tamamlamıştır. Yüksek lisans eğitimini 2019–2021 yılları arasında Hasan Kalyoncu Üniversitesi Elektronik ve Bilgisayar Mühendisliği Bölümü'nde tamamlamış olup, doktora eğitimine Atatürk Üniversitesi Bilgisayar Mühendisliği Bölümü'nde devam etmektedir. Özel sektörde uzun yıllar yazılım geliştirme uzmanı olarak görev yaptıktan sonra, 2018 yılından itibaren Mardin Artuklu Üniversitesi'nde Öğretim Görevlisi olarak akademik hayatını sürdürmektedir.

Akademik ilgi alanları arasında yazılım mühendisliği, web teknolojileri, veri tabanı sistemleri, yapay zekâ, makine öğrenmesi, kablosuz algılayıcı ağlar ve afet yönetimi bilgileri teknolojilerinin kullanımı yer almaktadır. Araştırmalarında bilişim teknolojilerinin gerçek yaşam problemlerine uygulanmasına, veri odaklı yöntemlerin geliştirilmesine ve yenilikçi teknoloji çözümlerinin ortaya konulmasına odaklanmaktadır. Ayrıca öğrencilerin uygulamalı yazılım geliştirme becerilerini desteklemeye yönelik proje ve uygulama temelli çalışmalar gerçekleştirmektedir. Bulanık mantık, makine öğrenmesi, derin öğrenme, meta-sezgisel algoritmalar, optimizasyon algoritmaları ve veri bilimi başlıca araştırma alanları arasında yer almaktadır.

Stub Yapılı Kuplaj Tabanlı Mikrodalga Sensör ile Motor Yağı Kullanım Ömrü Tespiti

Detection of Engine Oil Service Life Using a Stub-Structured Coupled-Line-Based Microwave Sensor

ID 0009-0008-8191-6534

ID 0000-0002-5675-1714

ID 0000-0002-0360-7188

ORCID: B. PEHLİVAN (0009-0008-8191-6534), M. YILDIRIM (0000-0002-5675-1714), M. A. GÖZEL (0000-0002-0360-7188)

Elektrik-Elektronik Mühendisliği Bölümü, Mühendislik ve Doğa Bilimleri Fakültesi

Süleyman Demirel Üniversitesi, Isparta, Türkiye

berkaypehlivanog@gmail.com, mustfayildirimm@gmail.com, *mahmutgozel@sdu.edu.tr

Özet

Bu çalışmada, motor yağı kullanım ömrünün mikrodalga frekans bölgesinde tahribatsız ve yüksek hassasiyetle izlenebilmesi amacıyla kuplaj hattı destekli, stub yapılı kompakt bir sensör tasarlanmış ve deneysel olarak doğrulanmıştır. AD 255C taban malzemesi üzerine inşa edilen 120×30 mm² boyutlarındaki sensör, çeyrek dalga empedans uyumlamasına göre optimize edilerek elektromanyetik alanın numune ile etkileşimi artırılmıştır. Tasarımın performansı, 0 km ile 21.100 km aralığındaki motor yağı numuneleri kullanılarak S₂₁ iletim parametresi üzerinden analiz edilmiştir. Deneysel sonuçlar, yağın kullanım kilometresi arttıkça dielektrik sabitindeki değişime bağlı olarak rezonans frekansında kayma sergilediğini kanıtlamıştır. Hava ortamında 5.48 GHz olan rezonans frekansı; kullanılmamış motor yağında 5.20 GHz'e, 15.500 km'de 5.18 GHz'e ve 21.100 km'de 5.12 GHz seviyesine gerilemiştir. Toplamda 360 MHz'lik bir frekans kayması elde edilmiştir. Önerilen S₂₁ tabanlı yaklaşım, literatürdeki yansıma temelli sistemlerden farklı bir ölçüm altyapısı sunarak, motor yağı kullanım durumunun belirlenmesine yönelik bakım uygulamaları için pratik bir yaklaşım ortaya koymaktadır.

Anahtar kelimeler: Motor yağı, mikrodalga sensör, kuplaj, stub, motor yağı kullanım ömrü.

Abstract

In this study, a coupling line-supported, stub-structured compact microwave sensor was designed and experimentally validated to monitor engine oil service life non-destructively and with high sensitivity in the microwave frequency range. The sensor, fabricated on an AD 255C substrate with dimensions of 120×30 mm², was optimized according to quarter-wave impedance matching conditions to enhance the interaction between the electromagnetic field and the oil sample. The sensor performance was evaluated using the S₂₁ transmission parameter for engine oil samples with usage histories ranging from 0 km to 21,100 km. Experimental results demonstrated that the resonance frequency systematically shifted toward lower frequencies as the oil usage increased, due to changes in the

dielectric constant. The resonance frequency decreased from 5.48 GHz in air to 5.20 GHz for unused oil and to 5.12 GHz for oil used for 21,100 km, resulting in a total frequency shift of 360 MHz. The proposed S₂₁-based approach provides a different measurement infrastructure from the reflection-based systems reported in the literature, offering a practical approach for maintenance applications aimed at determining the condition of engine oil.

Keywords: Engine oil, microwave sensor, coupling, stub, engine oil lifetime.

1. Giriş

Mikrodalga frekanslarında çalışan sensör sistemleri, yüksek hassasiyetleri, düşük üretim maliyetleri ve temassız ölçüm kabiliyetleri sayesinde malzeme karakterizasyonu ve içerik analizi uygulamalarında kullanımı gün geçtikçe artmaktadır [1-4]. Bu sistemler, ölçüm yapılan ortamın dielektrik özelliklerinde meydana gelen değişimleri, S-parametreleri üzerinden rezonans frekansı kaymaları ve genlik değişimleri aracılığıyla hassas bir biçimde algılayabilmektedir [5, 6]. Elektromanyetik dalgaların madde ile etkileşimine dayanan bu yaklaşım, özellikle sıvı veya katı ortamların elektriksel özelliklerindeki değişimlerin izlenmesinde başarılı sonuçlar sunmaktadır. Mikrodalga sensörler hızlı ve tahribatsız ölçüm avantajları sayesinde gıda güvenliği, kimyasal saflık analizi ve biyolojik sıvıların sınıflandırılması gibi farklı disiplinlerde uygulanmaktadır [7]. Mikrodalga sensörlerin algılama hassasiyetini arttırmak için farklı teknikler tercih edilmektedir [8-12]. Mikrodalga sensörlerindeki teknikler arasında metalmalzemeler [13, 14], anten tabanlı sistemler [15], diferansiyel yöntemler [16], iletim hatları [17], RFID sistemli uzaktan sensörler [17-19], rezonatör tasarımları [20] ve kuplör [5] uygulamalarıdır.

Motorlu taşıtlarda motor yağının işlevsel ömrü ve durumu; sürtünme ve aşınmanın azaltılması, ısının uzaklaştırılması ve sızdırmazlığın korunması açısından kritik öneme sahiptir. Uygulamada yağ değişim periyotları çoğunlukla sabit kilometre veya zaman aralıklarına göre belirlenmektedir. Ancak sürüş koşulları, çevresel faktörler, yakıt kalitesi ve motor içi kirlenme gibi çok sayıda değişken nedeniyle yağın gerçek kullanım ömrü bu tahminlerden önemli ölçüde sapabilmektedir. Kurum,

oksidasyon ürünleri, su ve yakıt seyreltmesi ile metal aşınma partiküllerinin birikimi, yağın hem fiziksel hem de dielektrik özelliklerini zamana bağlı olarak değiştirmekte; bu durum erken değişim nedeniyle maliyet ve atık artışına ya da gecikmiş bakım sonucu performans kaybı ve arıza riskinin yükselmesine neden olabilmektedir. Bu bağlamda, motor yağının anlık durumunu doğrudan yansıtan, araç üzerinde uygulanabilir ve ekonomik bir izleme yaklaşımına duyulan ihtiyaç giderek artmaktadır. Motor yağı özelinde yapılan bir çalışmada araç motor yağının ömrü ve seviyesini gerçek zamanlı olarak izlemek amacıyla 2.4 GHz’de çalışan Asimetrik Eş-düzlemsel Şerit beslemeli bir mikroşerit anten tasarlanmış ve antenin motor yağı içerisindeki farklı derinliklerdeki konumuna bağlı olarak rezonans frekansı, bant genişliği ve yansıma genliği değişimlerini analiz edilmiştir. Elde edilen sonuçlar, kullanılmamış motor yağının 2.20 GHz’de, 10 bin km kullanılmış motor yağının 2.18 GHz’de ölçüldüğü belirtilmiştir. kullanım kilometresi arttıkça dielektrik özelliklerde meydana gelen değişimlerin rezonans karakteristiklerine yansıdığını açıkça ortaya koymuştur [15]. Motor yağı özelinde yapılan bir diğer çalışmada ise X-bandında çalışan kiral metamalzeme tabanlı sensör önerilmektedir. Önerilen çalışmada, IS680 alt tabakası üzerine basılmış kare ve dairesel rezonatörlerden oluşan kiral bir metamalzeme sensörü geliştirilmiş, sayısal olarak analiz edilmiş ve deneysel olarak doğrulanmıştır. Sensörün performansı, motor yağının bozulmasına bağlı olarak değişen dielektrik sabitinin rezonans frekansında oluşturduğu kaymalar üzerinden değerlendirilmiştir. Rapor edilen sonuçlara göre, temiz ve kullanılmış motor yağı örnekleri arasında belirgin rezonans frekansı kaymaları gözlemlenmiştir. S_{11} parametresi için 0–10.000 km aralığında yaklaşık 40 MHz, 0–5.000 km aralığında ise 60 MHz’lik kaymalar rapor edilirken; S_{22} parametresi için bu değerlerin sırasıyla 30 MHz ve 120 MHz olduğu belirtilmiştir. Sayısal simülasyon sonuçları ile deneysel ölçümler arasında genel olarak iyi bir uyum elde edilmiş olup, küçük sapmaların kalibrasyon ve laboratuvar koşullarından kaynaklandığı ifade edilmektedir [21].

Mikrodalga sensörlerin yalnızca biyolojik ve gıda numunelerinde değil, petrol türevi sıvıların dielektrik özelliklerinin karakterizasyonunda da yüksek potansiyele sahip olduğu, güncel çalışmalarla ortaya konmuştur. Yapılan çalışmalarda mazot içerisine farklı oranlarda gaz yağı ilave edilmesiyle oluşan dielektrik özellik değişimlerinin belirlenmesi amacıyla yansıma katsayısı (S_{11}) tabanlı, yüksek kalite faktörüne sahip bir mikrodalga rezonatör sensörü geliştirmiştir. Yaklaşık 5.08 GHz rezonans frekansında çalışan sensörün elektromanyetik tasarımı 3.95–5.85 GHz frekans aralığında gerçekleştirilmiş, numunelerin dielektrik özellikleri açık uçlu koaksiyel prob yöntemiyle belirlenerek elektromanyetik benzetim modeline aktarılmıştır. Gaz yağı oranının artmasına bağlı olarak rezonans frekansında düzenli kaymalar meydana geldiği ve önerilen sensörün yakıt numunelerindeki dielektrik değişimlerin karakterizasyonunda yüksek duyarlılık sağladığı gösterilmiştir [22]. Petrol türevi sıvılar üzerinde elde edilen bu sonuçların ardından benzer mikrodalga sensör yaklaşımı, gıda ürünlerinde taşıyıcı analizine de uygulanmıştır. Bu alanda yapılan çalışmada ise 3.95–5.85 GHz frekans aralığında çalışan, S_{11} tabanlı yüksek kalite faktörüne sahip bir mikrodalga rezonatör sensörü kullanarak süt taşıyıcısının belirlenmesini amaçlamıştır. Farklı oranlarda yağsız süt ilave edilen numunelerin dielektrik özellikleri açık uçlu koaksiyel prob yöntemiyle belirlenmiş, yaklaşık 5.127 GHz

rezonans frekansında çalışan sensörün, katkı oranına bağlı olarak belirgin frekans kaymaları oluşturduğu gösterilmiştir. Bu sonuçlar, mikrodalga rezonatör tabanlı sensörlerin süt numunelerindeki dielektrik değişimlerin hassas biçimde karakterize edilmesinde etkili bir yöntem sunduğunu ortaya koymuştur [23]. Benzer yaklaşım bitkisel yağlarda taşıyıcı analizine de uyarlanmıştır. Zeytinyağına farklı oranlarda mısır yağı ilavesinin belirlenmesi amacıyla metamalzeme tabanlı, S_{11} esaslı bir mikrodalga rezonatör sensörü geliştirmiştir. Elektromanyetik analizleri 3.95–5.85 GHz frekans aralığında gerçekleştirilen ve yaklaşık 5.05 GHz rezonans frekansında çalışan sensör ile katkı oranına bağlı olarak rezonans frekansında belirgin değişimler elde edilmiş; kalite faktörü, normalize edilmiş hassasiyet ve başarımlı ölçütü bakımından olumlu sonuçlar rapor edilmiştir. Bu çalışma, metamalzeme tabanlı rezonatör sensörlerin sıvı gıda ürünlerindeki dielektrik değişimlerin belirlenmesinde yüksek duyarlılık sağlayabileceğini göstermiştir [24]. Bu çalışmanın yanı sıra zeytinyağındaki ayçiçek yağı taşıyıcısının belirlenmesine yönelik metamalzeme tabanlı düzlemsel, S_{11} tabanlı mikrodalga rezonatör sensörü geliştirmiştir. FR-4 alt tabaka üzerine tasarlanan ve 8–12 GHz X-band frekans aralığında optimize edilen sensörün numunesiz rezonans frekansı yaklaşık 9.48 GHz olarak belirlenmiştir. Açık uçlu koaksiyel prob yöntemiyle elde edilen dielektrik verileri kullanılarak gerçekleştirilen analizler, ayçiçek yağı oranındaki artışa bağlı olarak rezonans frekansında düzenli kaymalar meydana geldiğini göstermiştir. Bulgular, farklı rezonatör geometrileri ve çalışma frekanslarında geliştirilen mikrodalga sensörlerin sıvı numunelerdeki dielektrik özellik değişimlerinin yüksek hassasiyetle karakterize edilmesinde güvenilir bir yaklaşım sunduğunu ortaya koymaktadır [25]. Bir başka çalışmada ise bal numunelerinde su taşıyıcısının belirlenmesine yönelik yansıma katsayısı tabanlı, düzlemsel bir mikrodalga rezonatör sensörü geliştirmiştir. Yaklaşık 2.45 GHz rezonans frekansında çalışan sensörün elektromanyetik tasarımı gerçekleştirilmiş, farklı oranlarda su ilave edilen bal numunelerinde oluşan dielektrik özellik değişimleri rezonans frekansındaki kaymalar üzerinden değerlendirilmiştir. Su oranındaki artışa bağlı olarak rezonans frekansında düzenli değişimler meydana geldiği ve önerilen sensörün bal numunelerindeki dielektrik özellik değişimlerinin karakterizasyonunda yüksek hassasiyet sunduğu gösterilmiştir [26].

Literatürde, motor yağı kalitesinin gerçek zamanlı ve invaziv olmayan yöntemlerle izlenmesine yönelik olarak X-bandında çalışan elektroformlu iris dalga kılavuzu tabanlı mikrodalga sensörler rapor edilmiştir. Üçüncü dereceden elektroformlu dalga kılavuzu filtresi kullanan bu sensör, motor yağının yaşlanmasına bağlı dielektrik özellik değişimlerini rezonans frekansı kaymaları üzerinden tespit edebilmekte ve geleneksel laboratuvar temelli yağ analizi yöntemlerine kıyasla sürekli ve yerinde izleme avantajı sunmaktadır. Deneysel çalışmalarda, 148, 316 ve 604 saat boyunca yaşlandırılmış motor yağı numuneleri için rezonans frekansında belirgin kaymalar gözlemlenmiş; en ileri yaşlanma durumunda yaklaşık 10 MHz’e kadar frekans kayması ölçülmüştür. Ek olarak, koaksiyel prob tekniği ile gerçekleştirilen dielektrik ölçümler, mikrodalga sensör verileriyle tutarlı sonuçlar vermiştir. Elde edilen bulgular, rezonans frekansının yağın yaşlanma süresine yüksek duyarlılık gösterdiğini ve sensörün öngörücü bakım uygulamalarında kullanılabilirliğini doğrulamaktadır. Elektroformlama yöntemiyle üretilen dalga kılavuzu yapısı,

düşük maliyet, hafiflik ve yüksek hassasiyet avantajları sunarken, sensörün sıvı konsantrasyon analizi gibi farklı endüstriyel uygulamalara da ölçeklenebilir bir çözüm sağladığı literatürde vurgulanmaktadır [27]. Mikrodalga tabanlı sensörler, motor yağının yaşlanma ve bozulma süreçleri sırasında meydana gelen bağıl dielektrik geçirgenlik ve kayıp tanjantı değişimlerini doğrudan algılayabilmeleri sayesinde bu ihtiyaca güçlü bir çözüm sunmaktadır.

Bu çalışmada, motor yağının kullanım ömrünün yüksek duyarlılıkla değerlendirilmesi amacıyla stub yapılı ve kuplaj hattı destekli kompakt bir mikrodalga sensör önerilmektedir. Stub ve kuplaj hattı, çeyrek dalga koşullarına göre boyutlandırılmış; kuplaj bölgesinin optimize edilmesiyle elektromanyetik alanın yağ numunesi ile etkileşiminin artırılması hedeflenmiştir. Bu sayede frekans seçiciliği yükseltilmiş ve dielektrik özelliklerde meydana gelen küçük değişimlere karşı hassas bir sensör yanıtı elde edilmiştir. Bu çalışmanın özgün katkıları aşağıda özetlenmiştir:

(i) Motor yağı kullanım ömrünün belirlenmesine yönelik olarak S_{21} iletim parametresi tabanlı, stub yapılı bir mikrodalga sensör mimarisi önerilmiştir.

(ii) Yansıma parametrelerine dayalı yaklaşımlardan farklı olarak, iletim parametresi kullanılarak bir ölçüm yaklaşımı sunulmuştur.

(iii) Önerilen sensör, farklı kilometrelerde kullanılmış gerçek motor yağı numuneleri ile deneysel olarak doğrulanmıştır.

2. Sensör Tasarımı

Bu çalışmada, sıvı karakterizasyonu ve dielektrik özelliklerin derişimini tespiti amacıyla kuplaj hattı Tabanlı Stub Yapılı bir mikrodalga sensör tasarımı gerçekleştirilmiştir. Sensörün temel geometrisi, test edilecek numunelerin dielektrik katsayılarında meydana gelen değişimlere karşı yüksek duyarlılık sağlayacak şekilde optimize edilmiştir. Sensör tasarımında mikroşerit iletim hattı yapısı esas alınmış olup, etkin dielektrik sabiti ve karakteristik empedans değerleri klasik mikroşerit hat teorisine dayalı olarak hesaplanmıştır. Bu kapsamda, mikroşerit hattın etkin dielektrik sabiti (ϵ_{eff}), alt tabaka dielektrik sabiti (ϵ_r), alt tabaka kalınlığı (h) ve iletim hattı genişliği (W) cinsinden aşağıdaki eşitlik ile ifade edilmektedir:

$$\epsilon_{eff} = \frac{\epsilon_r + 1}{2} + \frac{\epsilon_r - 1}{2} \cdot \left(1 + 12 \cdot \frac{h}{W}\right)^{-1/2} \quad (1)$$

Burada ϵ_{eff} , mikroşerit hattın elektromanyetik alanının bir kısmının hava ortamında yayılmasından kaynaklanan etkin dielektrik sabitini temsil etmektedir. Etkin dielektrik sabiti, iletim hattının faz hızı ve rezonans frekansının belirlenmesinde kritik bir parametredir. Etkin dielektrik sabitinin belirlenmesinin ardından, mikroşerit hattın karakteristik empedansı (Z_0) aşağıdaki eşitlik kullanılarak hesaplanmıştır:

$$Z_0 = \frac{60}{\sqrt{\epsilon_{eff}}} \cdot \ln\left(\frac{8h}{W} + 0.25 \cdot \frac{W}{h}\right) \quad (2)$$

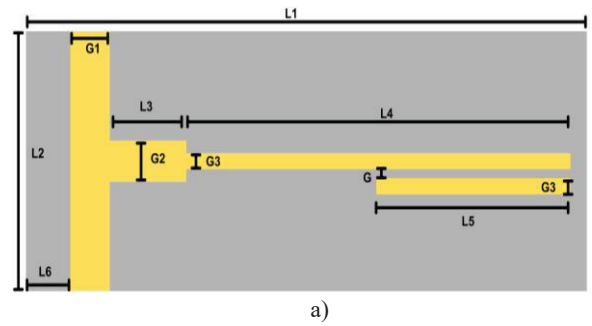
Karakteristik empedans değeri, sensörün giriş-çıkış uyumu ve maksimum güç aktarımının sağlanması açısından önem taşımaktadır. Bu çalışmada, sensörün ölçüm sistemleriyle

uyumlu olacak şekilde 50 Ω empedans koşulu hedeflenmiştir. Stub yapılı rezonatör yapısının rezonans davranışı, eşdeğer devre modeli üzerinden LC rezonans yaklaşımı ile analiz edilmiştir. Bu bağlamda, rezonans frekansı aşağıdaki eşitlik ile ifade edilmektedir:

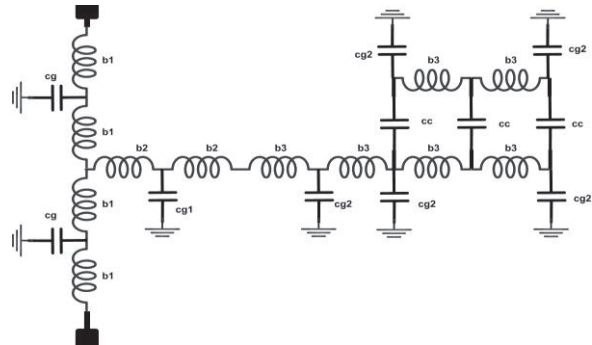
$$f = \frac{1}{2\pi\sqrt{LC}} \quad (3)$$

Burada L ve C parametreleri, rezonatör yapısının eşdeğer endüktans ve kapasitans değerlerini temsil etmektedir. Rezonans frekansı, test edilen motor yağı numunelerinin dielektrik özelliklerine bağlı olarak değişmekte olup, bu değişim sensörün algılama mekanizmasının temelini oluşturmaktadır. Bu teorik ifadeler doğrultusunda, sensör geometrisi çeyrek dalga koşulları dikkate alınarak optimize edilmiş ve elektromanyetik alanın numune ile maksimum etkileşimi sağlayacak şekilde tasarım parametreleri belirlenmiştir.

Sensör tasarımında, yüksek frekans uygulamalarına uygun mikrodalga alt tabaka malzemesi olan Rogers AD 255C tercih edilmiştir. Kullanılan alt tabaka malzemesinin bağıl dielektrik sabiti (ϵ_r) 2.55, kayıp tanjantı ($\tan\delta$) 0.0014 (10 GHz'de) ve kalınlığı 0.762 mm'dir; bakır kaplama kalınlığı ise 35 μ m olarak belirlenmiştir. Sensörün toplam fiziksel boyutları 120 $\times$ 30 mm² olarak belirlenmiş olup, kompakt bir yapı içerisinde yüksek hassasiyet göstermesi hedeflenmiştir. Bu kapsamda tasarlanan stub yapılı mikrodalga sensörün tasarım parametreleri ve eşdeğer devre modeli Şekil 1'de birlikte sunulmuştur.



a)



b)

Şekil 1: Önerilen sensörün tasarım parametreleri (a) ve eşdeğer devre modeli (b)

Şekil 1(a)'da önerilen iki portlu sensör yapısının tasarımına ait parametreler verilirken Şekil 1(b)'de bu parametreler ile tasarlanan sensörün eşdeğer devre modeli yer almaktadır.

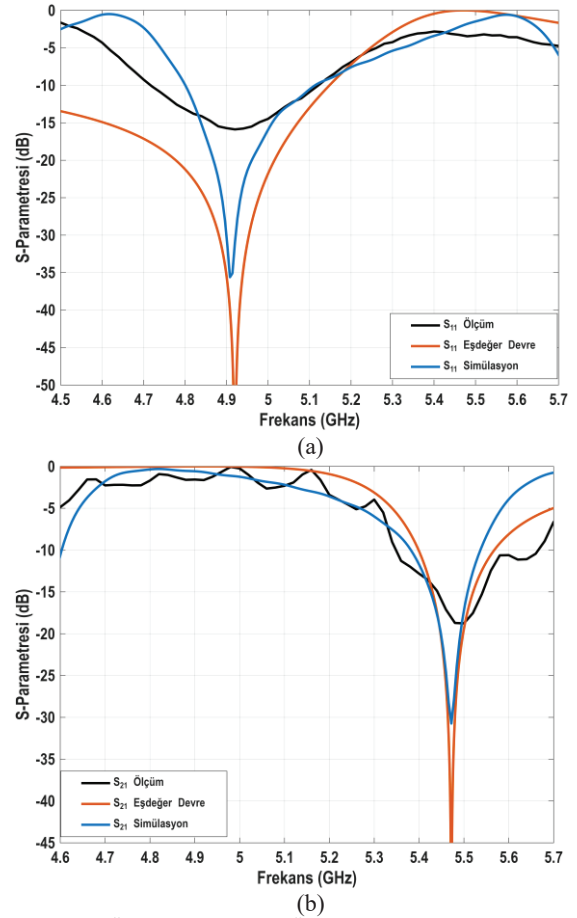
Sensör, ana iletim hattı üzerine yerleştirilmiş T-saplamalı rezonatör ve kuplaj bölgesini içermekte olup, rezonans koşullarının çeyrek dalga empedans uyumlamaya göre tasarlanmıştır. Sensör üzerinde tanımlanan parametreler, elektromanyetik alanın sıvı numune ile etkileşimini maksimize edecek şekilde optimize edilmiştir. Sunulan eşdeğer devre modeli ise iletim hattı segmentlerinin endüktansları (b_1 , b_2 , b_3) ve kuplaj ile saçılma etkilerini temsil eden kapasitif elemanlar (c_g , c_{g1} , c_{g2} , c_c) kullanılarak oluşturulmuştur. Bu model sayesinde sensörün rezonans frekansı yükleme etkilerine karşı gösterdiği davranış analitik olarak incelenebilmektedir. Önerilen sensörün tasarım parametreleri ve eşdeğer devresinde kullanılan elemanların değerleri Tablo 1 ile verilmektedir.

Tablo 1: Sensör tasarımında ve eşdeğer devrede kullanılan parametreler ve değerleri

Parametreler	Değerler (nH)	Parametreler	Değerler (pF)
b_1	0.3838	c_g	0.02653
b_2	2.41302	c_{g1}	0.3269
b_3	3.5089	c_{g2}	0.4310
		c_c	1.2991
Parametreler	Değerleri (mm)	Parametreler	Değerleri (mm)
L1	120	L6	8
L2	30	G	0.5
L3	6.5	G1	1.8
L4	100	G2	2.8
L5	50	G3	1

Tablo 1’de, önerilen stub yapılı mikrodalga sensörün eşdeğer devre modelinde kullanılan elemanların parametreleri ile geometrik boyutları sunulmaktadır. Eşdeğer devrede yer alan endüktans (b_1 – b_3) ve kapasitans (c_g , c_{g1} , c_{g2} , c_c) değerleri, sensörün elektromanyetik davranışını en iyi temsil edecek şekilde optimize edilmiş; buna karşılık L1–L6 uzunlukları ve G, G1–G3 boşluk parametreleri, rezonans frekansı, kuplaj seviyesi ve alan yoğunlaşmasını kontrol edecek biçimde belirlenmiştir. Önerilen sensörün eşdeğer devre modeli, Keysight Advanced Design System (ADS) 2016 yazılımı kullanılarak oluşturulmuştur. Eşdeğer devrede yer alan endüktans (b_1 – b_3) ve kapasitans (c_g , c_{g1} , c_{g2} , c_c) değerleri, sensörün elektromanyetik davranışını en iyi temsil edecek şekilde optimize edilmiş; L1–L6 uzunlukları ile G ve G1–G3 boşluk parametreleri ise rezonans frekansı, kuplaj seviyesi ve elektromanyetik alan yoğunlaşmasını kontrol edecek şekilde belirlenmiştir. Ardından, ADS ortamındaki Tuning ve Optimization araçları kullanılarak devre elemanlarının parametreleri iteratif olarak optimize edilmiş, eşdeğer devre modelinin elektromanyetik benzetim sonuçlarıyla kabul edilebilir düzeyde uyum sağlaması hedeflenmiştir.

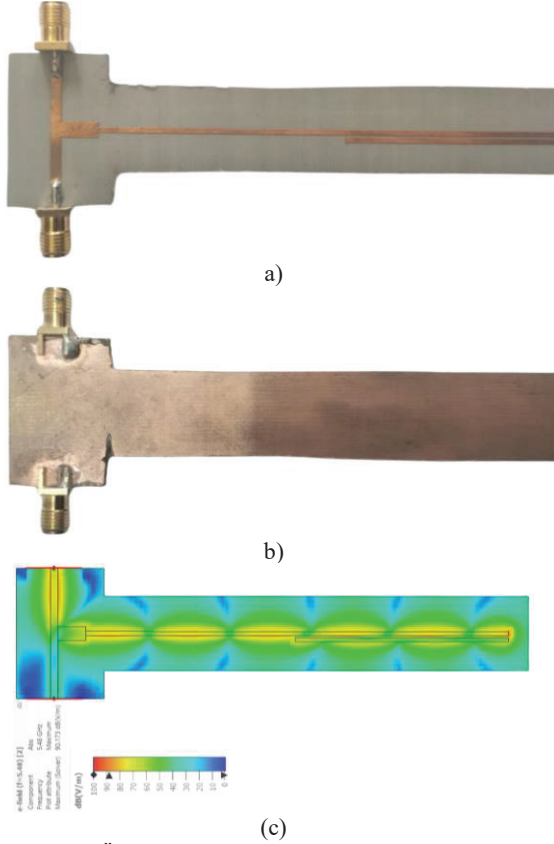
Önerilen sensör yapısının elektromanyetik performansını doğrulamak amacıyla, ölçüm sonuçları ile eşdeğer devre ve elektromanyetik simülasyon çıktıları S-parametreleri üzerinden karşılaştırılmıştır. Önerilen sensörün elektromanyetik benzetimleri CST Studio Suite 2021 yazılımı kullanılarak gerçekleştirilmiştir. Simülasyonlarda Elektromanyetik alanın serbest uzay koşullarında yayılımını modellemek amacıyla tüm yönlerde “Open (Add Space)” sınır koşulları uygulanmıştır.



Şekil 2: Önerilen sensörün Ölçüm, Simülasyon ve Eşdeğer Devre S-Parametreleri

Şekil 2’de sunulan S₂₁ ve S₁₁ parametreleri için ölçüm, eşdeğer devre simülasyonu ve elektromanyetik simülasyon sonuçlarının karşılaştırılması verilmiştir. S₂₁ parametresi incelendiğinde, sensörün 5.48 GHz frekansında iletimin en az olduğu belirgin bir rezonans sergilediği ve ölçüm sonuçlarının simülasyon verileriyle kabul edilebilir düzeyde uyumlu olduğu görülmektedir. Bu rezonans frekansı, motor yağı kullanım ömrünün belirlenmesinde temel gösterge olarak kullanılmış ve yağın dielektrik özelliklerindeki değişimlere bağlı frekans kaymaları S₂₁ üzerinden analiz edilmiştir. S₁₁ parametresi açısından ise eşdeğer devre simülasyonu ve ölçüm sonuçlarının 4.91 GHz aynı rezonans frekansı verdiği fakat farklı genlik oranlarını olduğu gözlemlenmiştir. Bu durum logaritmik ölçekte her ne kadar genlik farkının olduğu gözükse de lineer ölçekte sonuçların kabul edilebilir düzeyde uyumludur. Bazı frekans bantlarında gözlemlenen uyumsuzluklar, sensörün üretim sürecinden kaynaklanan tolerans hatalarına atfedilmektedir. Özellikle bakır iletim hatlarının aşındırma aşamasında oluşan 0.1 mm’den küçük boyutsal sapmalar, elektromanyetik davranış üzerinde sınırlı düzeyde etkiler oluşturarak bu uyumsuzluklara neden olmuştur. Tasarlanan stub yapılı mikrodalga sensörün deneysel olarak doğrulanabilmesi amacıyla, belirlenen geometrik parametrelere uygun olarak baskı devre kartı üzerinde üretimi gerçekleştirilmiştir. Önerilen sensör prototipi, üniversitenin laboratuvar imkânları

kapsamında kimyasal aşındırma yöntemi kullanılarak üretilmiştir. Sensör geometrisi AD 255C alt tabaka üzerine bakır kaplı PCB olarak aktarılmış, iletken desenler kimyasal aşındırma işlemi ile elde edilmiştir. Üretim tamamlandıktan sonra sensör giriş ve çıkış portlarına 50 Ω SMA konnektörler laboratuvar ortamında lehimlenmiş, ardından spektrum analizörün yansıma ve iletim kalibrasyonları ayrı ayrı gerçekleştirilerek S_{11} ve S_{21} parametreleri ölçülmüştür.



Şekil 3: Önerilen sensörün baskısının ön (a), arka (b) yüzeyleri ve elektrik alan dağılımı (c)

Şekil 3'te, üretilen sensör prototipinin üst yüzeyi ve toprak düzlemi olan alt yüzeyi sırasıyla gösterilmektedir. Üst katmanda, ana iletim hattı, T-saplamalı rezonatör yapısı ve kuplaj bölgesi bulunurken; alt yüzeyi ise toprak yapısı oluşturmaktadır. Önerilen sensör yapısında bulunan kuplaj hattı elektrik alanı yoğunlaştırarak sensöre algılama hassasiyetini artırmaktadır. Sensör, ölçüm sistemine entegrasyonun sağlanabilmesi amacıyla SMA konnektörler ile donatılmıştır. Üretilen prototip, simülasyon ve eşdeğer devre çalışmalarında elde edilen tasarım parametreleri ile uyumlu olup, deneysel ölçümlerin gerçekleştirilmesinde kullanılmıştır.

3. Ölçüm Sistemi ve Sonuçları

Bu çalışmada önerilen stub yapılı mikrodalga sensörün motor yağı kullanım ömrünü ayırt edebilme yeteneğini değerlendirmek amacıyla kontrollü bir ölçüm sistemi oluşturulmuştur. Deneysel düzenek Rohde & Schwarz FSH6 spektrum analizör, önerilen sensör, sabitleme mekanizması ve farklı kullanım kilometrelerine sahip motor yağı

numunelerinden oluşmaktadır. Deneysel ölçümlerden önce FSH6 spektrum analizör 4.5–5.7 GHz frekans aralığında S_{11} Açık (Open)-Kısa Devre (Short)-Yük (Load) kalibrasyon yöntemi kullanılarak, S_{21} ise kabloların birbirine doğrudan bağlanarak daha sonra Yük bağlanarak kalibre edilmiştir. Kalibrasyon işleminde referans yük olarak 50 Ω standart yük kullanılmıştır. Ölçüm ve kalibrasyon işlemlerinde ise karakteristik empedansı 50 Ω olan ve 0–18 GHz frekans aralığında çalışabilen düşük kayıplı RF koaksiyel kablolar tercih edilmiştir. Böylece ölçüm sisteminden kaynaklanabilecek sistematik hataların en aza indirilmesi ve referans düzleminin sensör girişine taşınması sağlanarak ölçümlerin doğruluğu ve tekrarlanabilirliği artırılmıştır. Ölçümler sırasında çevresel etkilerin ve mekanik titreşimlerin azaltılması hedeflenmiş; sensör-numune etkileşiminin tekrarlanabilir ve kararlı koşullar altında gerçekleştirilmesine özen gösterilmiştir. Ölçüm sistemi ve kullanılan motor yağı numunelerine ait görseller Şekil 4'te sunulmaktadır.



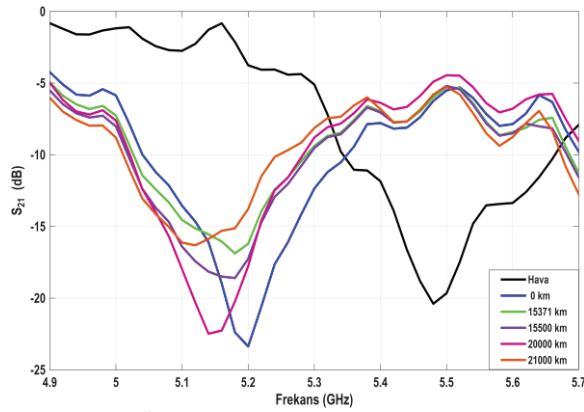
Şekil 4: Önerilen ölçüm sistemi ve farklı kullanım değerlerine sahip motor yağları

Şekil 4'te stub yapılı mikrodalga sensörün deneysel ölçüm düzeneği, motor yağı numuneleri ve ölçüm kurulumu yer almaktadır. Sensör, sabit bir mekanik düzenek yardımıyla konumlandırılmıştır. Ölçümler, hiç kullanılmamış motor yağı (0 km), yaklaşık 15.500 km ve 15.371 km kullanılmış yağlar, 20.000 km ve 21.100 km kullanılmış motor yağlarından oluşmaktadır. Deneysel çalışmada 0 km yeni motor yağı ile kullanım ömrü farklı dört adet ticari motor yağı numunesi kullanılmıştır. Numuneler; 15.371 km ve 15.500 km kullanım ömrüne sahip 5W30 viskozite sınıfındaki benzinli araç motor yağları ile 20.000 km ve 21.000 km kullanım ömrüne sahip

5W30 viskozite sınıfındaki dizel araç motor yağlarından temin edilmiştir. Yağ numunelerinin alındığı araçlar genellikle şehir içi ulaşımda kullanılmıştır.

Sensörün kompakt boyutları ve sade geometrik yapısı yüzeyin hızlı ve kalıntısız temizlenmesine imkân tanımakta; böylece ardışık ölçümler arasındaki çapraz bulaşma riski ortadan kaldırılarak sistem bir sonraki teste hızla hazır hale getirilmektedir. Ölçümler öncesinde tüm numuneler aynı laboratuvar ortamında muhafaza edilmiş ve oda sıcaklığında bekletilerek ölçümler gerçekleştirilmiştir.

Motor yağının kullanım süresi boyunca maruz kaldığı termal, mekanik ve kimyasal etkiler; yağın dielektrik sabiti, kayıp faktörü ve viskozitesi gibi elektromanyetik parametrelerinde ölçülebilir değişimlere neden olmaktadır. Bu değişimler, mikrodalga frekanslarında çalışan rezonans tabanlı sensörler aracılığıyla, temaslı veya temassız olarak izlenebilmektedir. Özellikle stub yapılı ve kuplajlı rezonatör mimarileri, elektromanyetik alanın örneklerle etkileşimini artırarak küçük dielektrik değişimlere karşı belirgin rezonans frekansı kaymaları üretmektedir. Bu çalışmada önerilen stub yapılı mikrodalga sensör kullanılarak, motor yağı numuneleri oda sıcaklığında S_{21} iletim parametresi üzerinden değerlendirilmiştir. Şekil 5'te, hava ortamı referansı ile farklı kullanım kilometrelerine sahip motor yağı numuneleri için elde edilen S_{21} frekans yanıtları sunulmaktadır.



Şekil 5: Önerilen sensör ile gerçekleştirilen S_{21} ölçüm sonuçları

Şekil 5'te görüldüğü üzere, stub yapılı mikrodalga sensörün rezonans frekansı, ortamda bulunan motor yağının kullanım kilometresi arttıkça daha düşük frekanslara doğru kaymaktadır. Referans olarak alınan hava ortamında rezonans frekansı yaklaşık 5.48 GHz civarında gözlemlenirken, kullanılmamış motor yağına karşılık gelen 0 km numunesinde bu değer yaklaşık 5.20 GHz seviyesine gerilemiştir. Benzer şekilde, 15.500 km ve 15.371 km kullanılmış yağ numuneleri için rezonans frekansı yaklaşık 5.18 GHz, 20.000 km için 5.14 GHz ve 21.100 km için ise yaklaşık 5.12 GHz olarak ölçülmüştür. Rezonans frekansındaki bu aşağı yönlü kayma, motor yağının kullanım süresi arttıkça dielektrik sabitinin ve kayıp mekanizmalarının değişmesiyle ilişkilendirilmektedir. Yağ içerisindeki oksidasyon ürünleri, metal aşınma partikülleri ve katkı maddelerinin bozulması, sensörün rezonans frekansını düşük frekanslara taşımaktadır.

Tablo 2: Önerilen sensör ile farklı kullanım oranına sahip numunelerin S_{21} ölçüm sonuçları

Ölçümler	Rezonans frekansı	Rezonans frekansı Değişimi
Hava (Yüksüz)	5.480	-
0 km (Kullanılmamış)	5.200	280 MHz
15.371 km	5.180	300 MHz
15.500 km	5.180	300 MHz
20.000 km	5.140	340 MHz
21.100 km	5.120	360 MHz

Tablo 2 incelendiğinde, yağın kullanım mesafesi arttıkça rezonans frekansının 5.200 GHz seviyesinden 5.120 GHz seviyesine doğru değişim sergilediği görülmektedir. Yapılan ölçümlerde 15.371 km ve 15.500 km numunelerinin her ikisinde de 5.18 GHz rezonans frekansı elde edilmiştir. Bu durum, söz konusu kilometre aralığında yağın dielektrik özelliklerinde meydana gelen değişimin, sensörün spektrum analizörü üzerindeki frekans çözünürlük limitlerinin altında kaldığını ve yağın bu evrede kimyasal olarak benzer karakteristikler sergilediğini göstermektedir.

Tablo 3'te, motor yağı kullanım ömrünün ve kalitesinin belirlenmesine yönelik olarak literatürde rapor edilen seçilmiş çalışmalar özetlenmiş ve kullanılan sensör türleri, çalışma frekansları, elde edilen rezonans frekansı değişimleri, ölçülen S-parametreleri ve test edilen sıvı türleri karşılaştırmalı olarak sunulmuştur.

Tablo 3: Motor yağı kullanım ömrünün tespit edilmesi için yapılan çalışmalar

Referans	Sensör türü	f_0 (GHz)	Δf (MHz)	S-param.	Numune
[15]	Mikroşerit Anten	2.40	130	S_{11}	Motor yağı
[21]	Dalga kılavuzu	X bandı	120	S_{11} - S_{22}	Motor yağı
[28]	CSRR Yapılı Anten	7.32	810	S_{11}	Motor yağı
[29]	Dipol Anten	2.40	510	S_{11}	Motor yağı
Bu çalışma	Stub yapılı İletim Hattı	5.48	360	S_{21}	Motor yağı

Tablo 3'te sunulan literatür karşılaştırması incelendiğinde, bu çalışmada elde edilen rezonans frekansı değişiminin bazı anten ve metamalzeme tabanlı yaklaşımlara kıyasla daha düşük olduğu görülmektedir. Ancak önerilen yaklaşım, literatürdeki çalışmalardan farklı olarak iletim hattı (stub yapılı) tabanlı bir sensör mimarisi kullanmakta ve yansımaya katsayıları (S_{11} , S_{22}) yerine doğrudan iletim parametresi S_{21} üzerinden değerlendirme yapmaktadır. S_{21} parametresinin kullanılması, sensör-numune etkileşiminin iletim karakteristiği üzerinden

değerlendirilmesine ve motor yağının kullanım durumuna bağlı değişimlerin rezonans frekansındaki kaymalar aracılığıyla karakterize edilmesine olanak sağlamaktadır. Ayrıca, önerilen iletim hattı tabanlı sensör yapısı, spektrum analizörü ile gerçekleştirilebilecek ölçümler açısından pratik bir yapı sunmaktadır. S_{11} parametresine dayalı yansıma ölçümlerinde ölçüm sistemine uygun bir yönlü bağlaştırıcı veya benzeri yansıma ölçüm bileşenlerinin gerekebilmesine karşın, iletim hattı tabanlı yapıda sensörün giriş ve çıkış portları arasındaki iletim karakteristiğinin doğrudan değerlendirilmesi mümkün olabilmektedir. Bu yönüyle önerilen yapı, motor yağı kullanımına bağlı dielektrik değişimlerin S_{21} rezonans frekansındaki değişimler üzerinden değerlendirilmesi için pratik bir ölçüm yaklaşımı sunmaktadır. Bu yönüyle önerilen sensör, yüksek frekans kayması hedefleyen anten tabanlı yaklaşımlardan ziyade, kestirimci bakım ve sürekli durum izleme uygulamalarına uygun, kararlı ve güvenilir bir ölçüm altyapısı sağlamaktadır. Dolayısıyla bu çalışma, literatüre motor yağı kullanım ömrünün belirlenmesinde S_{21} tabanlı, iletim hattı destekli mikrodalga sensör yaklaşımını kazandırması bakımından özgün bir katkı sunmaktadır.

4. Sonuçların Tartışılması

Bu çalışmada, motor yağı kullanım durumunun mikrodalga frekans bölgesinde temassız ve yüksek hassasiyetle izlenebilmesi amacıyla stub yapılı, kuplaj hattı destekli kompakt bir mikrodalga sensör tasarlanmış ve deneysel olarak doğrulanmıştır. Ölçüm sonuçları, sensörün rezonans frekansının motor yağının kullanım kilometresi arttıkça sistematik biçimde daha düşük frekanslara kaydığını açıkça göstermiştir. Hava ortamında 5.48 GHz olarak ölçülen rezonans frekansı, kullanılmamış motor yağı (0 km) için 5.20 GHz'e düşerken, 15.371–15.500 km aralığındaki yağlarda 5.18 GHz, 20 000 km'de 5.14 GHz ve 21 100 km'de 5.12 GHz olarak ölçülmüştür. Buna karşılık rezonans frekansı değişimi, hava referansına göre maksimum 360 MHz'e ulaşmıştır. Stub yapısının çeyrek dalga koşullarına göre optimize edilmesi ve kuplaj bölgesinin güçlendirilmesi sayesinde elektromanyetik alanın numune ile etkileşimi artırılmış, böylece küçük dielektrik değişimlere karşı yüksek frekans seçiciliği elde edilmiştir. Elde edilen deneysel bulgular, önerilen sensörün motor yağı durum izleme uygulamalarında anlık, düşük maliyetli ve saha koşullarına uygun bir çözüm sunduğunu ortaya koymaktadır. Gelecek çalışmalarda, daha yüksek hassasiyet sağlayan algılama teknikleri kullanılarak sensörün performansının geliştirilmesi ve farklı kullanım durumlarına sahip çeşitli motor yağı türleri ile kapsamlı ölçümlerin gerçekleştirilmesi hedeflenmektedir.

Teşekkür

Bu çalışma, Türkiye Bilimsel ve Teknolojik Araştırma Kurumu (TÜBİTAK) tarafından 2209-A Üniversite Öğrencileri Araştırma Projeleri Destekleme Programı kapsamında, 1919B012415165 proje numarası ile desteklenmiştir. Yazarlar,

sağlanan maddi ve akademik destekten dolayı TÜBİTAK'a teşekkür eder.

Kaynaklar

- [1] Schueler, M., vd., "Metamaterial inspired microwave sensors", IEEE Microwave Magazine, 13(2), 57-68, 2012.
- [2] Martín, F., vd., Planar Microwave Sensors, John Wiley & Sons, 2022.
- [3] Ebrahimi, A., Beziuk, G. ve Ghorbani, K., "Amplitude variation microwave sensors for high-sensitivity detection of solid and microfluidic dielectrics", IEEE Transactions on Microwave Theory and Techniques, 72(11), 6460-6472, 2024.
- [4] Genç, İ., Gözel, M.A. ve Kahrıman, M., "Microwave sensor design with directional coupler-based CSRR structure for the characterization of the dielectric properties of solid materials", Measurement, 119548, 2025.
- [5] Ebrahimi, A. ve Ghorbani, K., "High-sensitivity detection of solid and liquid dielectrics using a branch line coupler sensor", IEEE Transactions on Microwave Theory and Techniques, 71(12), 5233-5245, 2023.
- [6] Mohammadi, P., Mohammadi, A. ve Kara, A., "A differential and high-sensitivity thickness measurement sensor for dielectric solid materials using a branch line coupler", IEEE Sensors Journal, 24(10), 16077-16084, 2024.
- [7] Yıldırım, M. ve Gözel, M.A., "High-sensitivity microwave sensor for buffalo milk fat rate and adulteration analysis with branch line coupler", Engineering Science and Technology, an International Journal, 71, 102172, 2025.
- [8] Zhang, N., vd., "Dual-mode anti-interference humidity detection: Differential microwave sensor based on microstrip circuit", Sensors and Actuators B: Chemical, 390, 133944, 2023.
- [9] Gulsu, M.S., vd., "Metamaterial-based sensor with a polycarbonate substrate for sensing the permittivity of alcoholic liquids in a WR-229 waveguide", Sensors and Actuators A: Physical, 312, 112139, 2020.
- [10] Araujo, J., vd., "Non-invasive antenna sensor for dielectric material identification at a distance", Measurement, 224, 113842, 2024.
- [11] Dalgac, S., vd., "Investigation of methanol contaminated local spirit using metamaterial based transmission line sensor", Measurement, 178, 109360, 2021.
- [12] Erdoğan, M., vd., "Determination of frying sunflower oil usage time for local potato samples by using microwave transmission line based sensors", Measurement, 163, 108040, 2020.
- [13] Hakim, M.L., vd., "Intensified electric field based highly sensitive backbone-shaped C and X band microwave metamaterial sensor for permittivity characterization applications", Engineering Science and Technology, an International Journal, 60, 101877, 2024.

- [14] Sim, M.S., vd., "Microwave sensors loaded with metamaterial-inspired resonators for dielectric material characterization: A review", *Sensors and Actuators A: Physical*, 372, 115322, 2024.
- [15] Yıldırım, M. ve Gözel, M.A., "Asimetrik eş-düzensel şerit beslemeli anten ile motor yağ seviye ve kullanım ömrü tespiti", *Mühendislik Bilimleri ve Tasarım Dergisi*, 11(3), 904-915, 2023.
- [16] Wang, B.-X., vd., "Sensitivity optimization of differential microwave sensors for microfluidic applications", *Sensors and Actuators A: Physical*, 330, 112866, 2021.
- [17] Palandoken, M., vd., "Novel microwave fluid sensor for complex dielectric parameter measurement of ethanol-water solution", *IEEE Sensors Journal*, 23(13), 14074-14083, 2023.
- [18] Gao, M., vd., "RFID-based microwave biosensor for non-contact detection of glucose solution", *Biosensors*, 11(12), 480, 2021.
- [19] Palandoken, M. ve Gocen, C., "RFID-enabled ML-assisted microwave liquid sensor design for complex dielectric characterization of water-methanol mixture", *Sensors and Actuators A: Physical*, 382, 116142, 2025.
- [20] Islam, M.R., vd., "Metamaterial sensor based on reflected mirror rectangular split ring resonator for the application of microwave sensing", *Measurement*, 198, 111416, 2022.
- [21] Dalgacı, Ş., vd., "Chiral metamaterial-based sensor applications to determine quality of car lubrication oil", *Transactions of the Institute of Measurement and Control*, 43(7), 1640-1649, 2021.
- [22] Korkmaz, H., Hasar, U. ve Delihacıoğlu, K., "Mikrodalga frekanslarda yakıt içeriğindeki değişimlerin karakterizasyonu", *Kahramanmaraş Sütçü İmam Üniversitesi Mühendislik Bilimleri Dergisi*, 29(1), 1-12, 2026.
- [23] Korkmaz, H., "Süt taşıma tespitinde mikrodalga sensör yöntemi", *EMO Bilimsel Dergi*, 15(1), 143-152, 2025.
- [24] Korkmaz, H., "Metamalzeme tabanlı hassas mikrodalga sensör ile zeytinyağındaki taşıma tespiti", *EMO Bilimsel Dergi*, 15(1), 41-52, 2025.
- [25] Korkmaz, H. ve Hasar, U., "Sensitivite microwave sensor for adulteration detection in olive oil", *Natural Sciences and Engineering Bulletin*, 1(2), 27-37, 2024.
- [26] Hasar, H., vd., "Sensitive microwave sensor for detection and quantification of water in adulterated honey", *IEEE Transactions on Instrumentation and Measurement*, 74, 1-13, 2025.
- [27] Zaarour, Y., vd., "Engine oil quality monitoring using an additively manufactured X-band microwave waveguide sensor", *IEEE Access*, 12, 181341-181350, 2024.
- [28] Kumar, K.S. ve Madhan, M.G., "A novel metamaterial-based microwave sensor design for the characterization of lubricating engine oils and alcohols", *Journal of Optoelectronics and Advanced Materials*, 27(9-10), 427-438, 2025.
- [29] Mejri, F. ve Aguil, T., "Design of a passive microwave sensor for the characterization of mobile engine oil", 2017 IEEE-APS Topical Conference on Antennas and Propagation in Wireless Communications (APWC), 2017, 266-269.

Özgeçmişler



Berkay Pehlivan, 2001 yılında Eskişehir’de doğmuştur. 2025 yılında Süleyman Demirel Üniversitesi Elektrik ve Elektronik Mühendisliği Lisans Programı’ndan mezun olmuştur. Araştırma ilgi alanları mikrodalga sensörler ve otomasyon sistemleridir.



Mustafa Yıldırım, 1999 yılında İzmir’de doğmuştur. 2022 yılında Burdur Mehmet Akif Ersoy Üniversitesi Elektrik ve Elektronik Mühendisliği Bölümü’nden lisans derecesini, 2025 yılında ise Süleyman Demirel Üniversitesi Elektrik ve Elektronik Mühendisliği Bölümü’nden yüksek lisans derecesini almıştır. Halen Süleyman Demirel Üniversitesi Elektrik ve Elektronik Mühendisliği Bölümü’nde doktora öğrencisi olarak çalışmalarını sürdürmektedir. Araştırma ilgi alanları arasında materyal karakterizasyonu ve içerik analiz için mikrodalga rezonatör sensörler yer almaktadır.



Doç. Dr. Mahmut Ahmet GÖZEL, 1987 yılında İstanbul’da doğmuştur. Lisans, yüksek lisans ve doktora derecelerini sırasıyla 2011, 2015 ve 2020 yıllarında Süleyman Demirel Üniversitesi Elektronik ve Haberleşme Mühendisliği Anabilim Dalı’nda tamamlamıştır. Halen aynı üniversitenin Elektrik-Elektronik Mühendisliği Bölümü’nde Doçent Dr. ünvanıyla öğretim üyesi olarak görev yapmaktadır. Başlıca araştırma alanları; antenler ve RF enerji hasadı, RFID sistemleri, mikrodalga aktif/pasif devre tasarımı, geniş bant teknolojileri ile materyal karakterizasyonuna yönelik mikrodalga rezonatör sensör tasarımlarıdır.

Extracting Sparse Periodic Fault Signatures from Noisy Signals in Rolling Element Bearings

Gürültülü Rulman Sinyallerinden Seyrek ve Periyodik Arıza Verilerinin Çıkarılması

Aziz Kubilay Ovacıklı

 0000-0001-6687-7794

Software Engineering Department, Istanbul Arel University, Istanbul, Türkiye
azizkubilayovacikli@arel.edu.tr

Öz

Bu çalışmada, rulmanlarda meydana gelen arızalara ait seyrek ve periyodik sinyal bileşenlerinin gürültülü gözlem verilerinden geri kazanımı incelenmektedir. Hedef sinyallerin darbeselliğini ve seyrekliğini artırmak amacıyla, yüksek mertebeden istatistiklerin (çarpıklık ve basıklık) optimizasyonu ile L_1 -normu minimizasyonu yöntemleri karşılaştırılmıştır. İlgilenilen sinyal, harmonik olarak ilişkili sinüs dalgalarının toplamı şeklinde modellenmiştir; bu optimizasyon fonksiyonlarının performansları, beyaz gürültü ve lineer bozulmanın bulunduğu ve bulunmadığı durumlarda değerlendirilmiştir. Elde edilen bulgular, gradyan tabanlı optimizasyonun istenmeyen bileşenleri bastırarak arıza kaynaklı periyodik darbeleri başarıyla geri kazandığını göstermektedir. Ayrıca sayısal deneyler, L_1 -normu minimizasyonunun, çarpıklık ve basıklık maksimizasyonuna kıyasla, seyrekliği artırmada daha üstün bir performans ve daha hızlı yakınsama sergilediğini doğrulamaktadır.

Anahtar kelimeler: Seyrek sinyal geri kazanımı, Gradyan arama optimizasyonu, Arıza teşhisi, Öznelik çıkarımı

Abstract

This study investigates recovering sparse, periodic fault signatures from noisy observations in rolling-element bearings. We compare optimizing higher-order statistics (skewness and kurtosis) with L_1 -norm minimization to enhance the impulsiveness and sparseness of the target signals. By modeling the signal of interest as a sum of harmonically related sinusoids, these objective functions are evaluated in the presence and absence of additive noise and linear distortion. The results demonstrate that gradient-based optimization successfully recovers defect-related periodic impulses by suppressing undesired components. Furthermore, numerical experiments confirm that L_1 -norm minimization provides superior performance and faster convergence in promoting sparsity compared to skewness- and kurtosis-maximizing methods.

Keywords: Sparse signal recovery, Gradient search optimization, Fault diagnosis, Feature extraction

1. Introduction

Periodic and sparse impulse trains, representing either periodic excitation or a response to such an event, carry valuable information about phenomena occurring in the process being monitored. Such impulsive appearances are studied in seismic wave propagation [1], acoustic signals [2], power-line communications [3-4], mechanical vibrations [5-6], and teletraffic and hydrology [7]. For example, such sparse signals could arise from high-energy impacts due to a defect in a rolling bearing as an almost-periodic event [8] or be generated and transferred via a modulator to excite a system, as in ultrasound [9-12].

In an experimental setting where random noise and disturbances corrupt the source signal, extracting impulsive events from noisy observations remains challenging. By exploiting the side information at hand - the hypothesis imposed on the physical or statistical characteristics of the source signal - the undesired components can be suppressed to recover the information-rich signal. After recovery, the resulting output may indicate a fault in the observed system, for instance, in rolling bearings widely used in rotating machinery [13-14]. These fault signatures possess certain characteristics, such as impulsiveness and sparseness, that can be enhanced by optimizing measures of merit (objectives) such as higher-order statistics (HOS) and the L_1 -norm, respectively.

The periodic impulses are modeled and simulated to study and compare recovery methods for vibration analysis of rolling bearings, thereby building a robust predictive maintenance scheme [15]. A periodic impulse train representing an outer race defect will be the signal of interest, which is disturbed by an unknown linear propagation path (transfer function) and additive noise. Such a sparse spike train can be extracted by enhancing impulsiveness via skewness- and kurtosis-maximization [6] or by promoting sparseness via L_1 -norm minimization [16]. In addition to this, modeling the periodic impulses as a sum of harmonically related sinusoids will enable a comparison of the ability of these measures as sparsity-promoting functions on a real Fourier basis consisting of mutually orthogonal components [17].

The main contribution of this study is the comparative evaluation of higher-order statistics (skewness and kurtosis)

and L_1 -norm minimization, conducted strictly within the framework of a real Fourier basis specifically designed for bearing fault diagnosis. While previous studies have independently explored HOS and sparsity for signal recovery, this paper uniquely models the fault-induced periodic impulses as a weighted sum of harmonically related sinusoids to benchmark these objective functions. Specifically, this study contributes by demonstrating that L_1 -norm minimization not only yields a better approximation of the pristine sparse signal by enforcing a more rapid decay of harmonic weights, but it also achieves significantly faster convergence compared to traditional skewness and kurtosis maximization in the presence of noise and linear distortion.

2. Extracting Fault Signatures by Optimizing a Measure of Merit

2.1. Background And Prior Work

In a problem setting without access to the signal of interest or a training signal, impulsive signatures can be extracted using a *blind adaptation* scheme. In this approach, a linear filter is adapted stepwise to optimize an objective that promotes a specific measure of merit of the source signal while suppressing the undesired components that mask it. In condition-based maintenance (CBM) of rolling element bearings, the underlying failures manifest as sparse, periodic impacts with (trivially) asymmetric or heavy-tailed probability density functions (PDFs). These statistical attributes can be quantified by scale-invariant 3rd and 4th order moments, respectively. By maximizing these objectives, periodic impulses can be recovered blindly from noisy vibration signals [5-6, 18]. Because these measures involve multi-model surfaces that can cause local convergence issues, previous studies have investigated their convergence properties for different signal classes [6, 19-21].

In another attempt to recover fault signatures by enhancing sparsity, local convergence issues associated with the above-mentioned objectives can be avoided under certain circumstances. Sparseness in a discrete-time vector $\mathbf{x}_n \in \mathbb{R}^N$ can be quantified by L_0 -norm $\|\mathbf{x}_n\|_0 = \#\{i: x_i \neq 0\}$ that can be approximated (convex relaxation) in a signal recovery framework (from noisy observations) by L_1 -norm $\|\mathbf{x}_n\|_1 = \sum_{i=1}^N |x_i|$ allowing linear programming [22-23]. Then, by minimizing the L_1 -norm of an observed signal that is defined on a proper basis, sparseness, and therefore the temporal resolution of spikes can be enhanced to reveal fault signatures [16].

The methods briefly mentioned so far are all implemented via gradient search: maximizing HOS is steepest ascent [24], whereas L_1 -norm minimization is steepest descent. Sparsity enhancement is generally studied under basis pursuit (BP) [25-27], with sufficient conditions for successful recovery of sparse signals given in [28], where the uniqueness of sparse representation [27] on orthonormal bases was discussed in [29] and objectives promoting sparsity stronger than L_1 were given in [30].

It is worth mentioning recovering sparse spike trains buried in noise [23, 31], for instance, using a wavelet packet dictionary and comparing against matching pursuit (MP) at varying levels of defect severity for inner and outer races [16]. L_1 -norm minimization was utilized to reveal the impulse

response of flaws in non-destructive evaluation (NDE) of materials [32], in deconvolution of seismic traces in the presence of noise [33], and in recovering lamb waves to remove multipath interference in ultrasonic testing [34]. In [35], impulsive signatures were enhanced using a normalized objective L_1/L_2 and compared with kurtosis maximization, and in [36], L_1 -norm minimization was employed on a subset of the spectrum of noisy vibrations from defective structures. Since the geometry of rolling bearings is known a priori [37], defect frequencies (given the rotational speed) can be estimated and used in [38] to detect mechanical faults.

2.2. Notation and Signal Model

This study focuses on outer-race defects to simplify the presentation. The deterministic, non-stationary, and periodic

discrete-time source signal $s_n \triangleq \sum_{m=-\infty}^{\infty} \delta[n - mT]$ with period

T , discrete-time index n , and the Kronecker delta function $\delta[n]$ is a sparse impulse train representing the impacts [38-40] arising from defective rolling bearings. This signal (picket fence) can be represented as a sum of harmonically related sinusoids [41] by a model with an implicit Fourier series expansion. By expanding the Fourier series to separate the DC component, the model then becomes

$$s_n = \frac{1}{T} + \frac{2}{T} \sum_{r=-\infty}^{\infty} \cos[r\omega_0 n], \text{ where } \omega_0 = \frac{2\pi}{T} \text{ is the ball pass}$$

frequency of a defect. In practical vibration analysis, piezoelectric sensors (accelerometers) are inherently AC-coupled and naturally reject the DC component ($1/T$). Furthermore, to ensure algorithmic stability and eliminate any static drift, the observed signals are explicitly pre-processed to be zero-mean. Consequently, the DC term is safely omitted from the analytical dictionary, and the basis matrix $\Psi[n, \omega_0]$ is constructed starting from the first fundamental harmonic ($r=1$). In practice, the physical sensor used for data acquisition has a finite bandwidth, limiting the observable high-frequency components. Consequently, the maximum number of harmonics N that can be accurately captured is strictly bounded by the Nyquist limit associated with the system's sampling frequency ω_s , formulated as $N < \omega_s/2\omega_0$, where ω_0 is the fundamental ball-pass frequency of the defect. This will produce a signal with a sparse representation in the given (orthogonal) real Fourier basis containing a certain number of impacts within an observation window. Through the mechanical structure represented by a linear time-invariant (LTI) system with impulse response h_n , the fault impacts s_n undergo amplitude and phase distortion, producing the distorted signal $\mathbf{x}_n$. The final observed signal $\mathbf{u}_n$ is then formed by the addition of measurement noise $\mathbf{z}_n$, expressed mathematically as $\mathbf{u}_n = \mathbf{x}_n + \mathbf{z}_n$. As established, this can be represented in matrix form as $\mathbf{u}_n = \mathbf{\Theta} + \mathbf{z}_n$. A more detailed analysis of the modeling and construction is available in [6]. Here, the emphasis (and thus novelty) is on the matrix notation that will be given now for the sparse enhancement of fault-related periodic impulses.

Assuming the source signal is sufficiently sparse, a concise representation in the proper basis enables sparse reconstruction of the fault impacts from noisy vibrations. Here, an asymmetric or heavy-tailed distribution (impulsiveness) does not necessarily imply sparseness; however, it aligns with this signal model.

The noisy observations will be synthesized as a weighted sum of harmonics. Let M denote the number of samples in the observation window, and N denote the number of considered harmonics. For a discrete-time index $n=0, 1, \dots, M-1$ spanning the observation window, we define the harmonic basis row vector $\Psi[n, \omega_0] \in \mathbb{R}^{1 \times N}$ evaluated at time n as defined in Eq. 1:

$$\Psi[n, \omega_0] = [\cos(\omega_0 n), \cos(2\omega_0 n), \dots, \cos(N\omega_0 n)] \quad (1)$$

The signal of interest, $\mathbf{s}_n$, representing the equally spaced and scaled periodic impacts, can be expressed in vector-matrix notation as defined in Eq. 2:

$$\mathbf{s}_n = \Psi[n, \omega_0] \mathbf{c} \text{ where } \mathbf{c} = \left[\frac{2}{T}, \frac{2}{T}, \dots, \frac{2}{T} \right]^T \in \mathbb{R}^N \quad (2)$$

As the signal travels through the mechanical structure, it is convolved with the linear time-invariant (LTI) system characterized by the impulse response h_n . The frequency response of this transmission path is mathematically defined by its Discrete-Time Fourier Transform (DTFT) as $H(e^{j\omega}) = \sum_{m=-\infty}^{\infty} h_m e^{-j\omega m}$. This system introduces linear distortion, filtering the harmonic components. Consequently, the parameter vector $\boldsymbol{\theta} = [\alpha_1, \alpha_2, \dots, \alpha_N]^T \in \mathbb{R}^N$ represents the modified weights of the harmonics, such that the initial amplitude of each r -th harmonic is scaled by the magnitude response of the transmission path at that specific harmonic frequency, yielding $\alpha_r = \frac{2}{T} |H(e^{jr\omega_0})|$. The resulting distorted signal $\mathbf{x}_n = \Psi[n, \omega_0] \boldsymbol{\theta}$ and the final noisy observation $\mathbf{u}_n$ are modeled as defined in Eq. 3:

$$\mathbf{u}_n = \mathbf{x}_n + \mathbf{z}_n \quad (3)$$

where the parameter vector $\boldsymbol{\theta} = [\alpha_1, \alpha_2, \dots, \alpha_N]^T \in \mathbb{R}^N$ represents the modified weights of the harmonics such that $\alpha_r = \frac{1}{T} |H(e^{jr\omega_0})|$. The term $\mathbf{z}_n$ represents the additive background noise inherent to the measurement environment, bounded by $\|\mathbf{z}_n\|_2 \leq \epsilon$.

This mathematical model with an implicit dictionary is the analytical approach that will allow us to seek the $\tilde{\boldsymbol{\theta}} = [\tilde{\alpha}_1, \dots, \tilde{\alpha}_N]^T$ one that produces the sparsest signal in both the absence and presence of noise. By doing so, we will assess the optimized functions for their ability to increase sparsity. Then, the aim will be to extract the impulses corrupted by an LTI transfer function and additive noise. The following section presents a simulation experiment to compare the methods for extracting fault signatures.

The proposed analytical approach uses a fixed, predefined dictionary constructed from the bearing's known kinematics. Specifically, the basis matrix $\Psi[n, \omega_0]$ serves as a deterministic dictionary of harmonically related sinusoids. Unlike "dictionary learning" frameworks, where atoms are adaptively learned from data, our dictionary is analytically constructed to reflect the physical periodicity of expected fault signatures, ensuring that the optimization process focuses solely on finding the sparsest weight vector $\boldsymbol{\theta}$ within this established physical basis.

The processing has been performed in the context of blind deconvolution, treating mechanical vibration as a sum of harmonics to analyze the convergence properties of skewness [6, 42]. A similar study on seismic data used a normalized sparsity measure [43]. In addition to extracting impulses from the given signal model, this application can also be seen as a harmonic content enhancement by L_1 -regularization [44], as it was done by skewness maximization in [45].

3. Sparse Enhancement of Fault Signatures

3.1. Comparison of Sparsity-Enhancing Functions in the Absence of Noise and Linear Distortion

This section compares functions optimized to enhance sparse fault signatures. As a sparsity measure, the expectation of the functions is to produce weights that are rapidly decreasing in the parameter vector $\tilde{\boldsymbol{\theta}}$, imposing a quick decay on them [46]. In other words, the more evenly distributed the total energy among the α -coefficients, the less sparse the signal is produced.

The following analyses are primarily fundamental and theoretical. Rather than proposing a competitive algorithm to outperform current state-of-the-art diagnostic tools, the aim is to investigate the mathematical behavior, coefficient decay, and convergence properties of specific sparsity-promoting functions when applied strictly to a real Fourier basis representation.

3.1.1. Maximizing HOS via Gradient Ascent on $\Psi[n, \omega_0] \boldsymbol{\theta}$

The scale-invariant p^{th} order moment of a zero-mean discrete-time signal $\mathbf{u}_n$ is mathematically defined as

$$M_p(\mathbf{u}_n) = \frac{E[\mathbf{u}_n^p]}{(E[\mathbf{u}_n^2])^{p/2}}. \text{ Based on this general definition,}$$

Skewness (S) and Kurtosis (K) correspond to the 3rd-order ($p = 3$) and 4th-order ($p = 4$) moments, respectively. Thus, substituting $p=3$ into Eq. 4 yields Skewness $S(\mathbf{u}_n) = M_3(\mathbf{u}_n)$, whereas substituting $p=4$ into Eq. 5 yields Kurtosis $K(\mathbf{u}_n) = M_4(\mathbf{u}_n)$:

$$S(\mathbf{u}_n) = \frac{E\{\mathbf{u}_n^3\}}{(E\{\mathbf{u}_n^2\})^{3/2}} \quad (4)$$

$$K(\mathbf{u}_n) = \frac{E\{\mathbf{u}_n^4\}}{(E\{\mathbf{u}_n^2\})^2} \quad (5)$$

The complete mathematical derivations of the gradients for Skewness and Kurtosis, including their step-by-step formulations and local convergence proofs, have been extensively established in prior work [6]; therefore, they are omitted here for brevity. However, for the practical numerical implementation of the simulation experiments, it is crucial to explicitly define how the statistical expectation operator $E[\cdot]$ is computed. In our discrete-time finite sample versions, the expectation is calculated as an unbiased time average over the observation window of length M . Specifically, the p^{th} order expectation of a discrete-time signal vector $\mathbf{v}$ is defined as in Eq. 6:

$$E[v^p] \approx \frac{1}{M} \sum_{n=0}^{M-1} v_n^p \quad (6)$$

Algorithmic Presentations

While highly efficient, state-of-the-art algorithms such as FastICA or the Fast Kurtogram exist in the literature for HOS maximization, we deliberately chose standard steepest ascent/descent algorithms with fixed step sizes for all objective functions in this study. The primary goal is not to benchmark the computational sophistication of the solvers, but rather to conduct a controlled, baseline comparison of the inherent mathematical topography (e.g., convexity, gradient steepness, and local optima) of the objective functions themselves. By confining all methods to the same basic gradient-based search mechanism within the fixed Fourier basis, the inherent convergence behavior and coefficient-decay properties dictated strictly by the chosen metrics (HOS vs. L_1 -norm) can be isolated and evaluated fairly.

To maximize the higher-order statistics (HOS) and minimize the L_1 -norm, we use iterative gradient-based approaches. Algorithm 1 and Algorithm 2 summarize the procedures.

Algorithm 1: HOS (Skewness/Kurtosis) Maximization via Gradient Ascent

1. **Initialize:** Initialize: Define starting parameter vector θ_0 as a uniform weight vector $\theta_0 = [1, 1, \dots, 1]^T$. Set step size $\mu > 0$. Set iteration index $k = 0$.
2. **Normalize:** Scale the initial vector to unit norm: $\theta_0 = \frac{\theta_0}{\|\theta_0\|_2}$.
3. **Repeat** until convergence (e.g., $\|\theta_{k+1} - \theta_k\|_2 < \delta$ or maximum iterations reached):
 - a. Compute the target signal estimate: $\hat{u}_k = \Psi[n, \omega_0] \theta_k$.
 - b. Calculate the gradient of the objective function (Skewness or Kurtosis): $\nabla M(\theta_k)$.
 - c. Update the parameter vector via steepest ascent: $\theta_{k+1} = \theta_k + \mu \nabla M(\theta_k)$.
 - d. Enforce unit energy constraint: $\theta_{k+1} = \frac{\theta_{k+1}}{\|\theta_{k+1}\|_2}$.
 - e. Update iteration index: $k = k + 1$.
4. **Output:** Optimal parameter vector $\tilde{\theta} = \theta_k$.

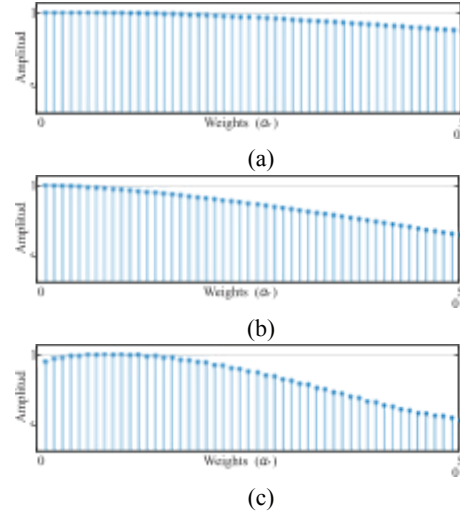


Figure 1. (a) The α -weights (scaled to have unit amplitude) in the vector $\tilde{\theta}$ after kurtosis maximization, (b) after skewness maximization, and (c) L_1 -norm minimization as formulated in (7).

For a numerically tractable algorithm, the norm of the parameter vector is constrained to be unity, which implies $\|\theta\|_2^2 = \theta^T \theta = \sum_{r=1}^N \alpha_r^2 = 1$. Scale-invariant 3rd (skewness) and 4th (kurtosis) order moments will be maximized to produce a sparse signal, both in the absence and presence of noise, to compare their performances against L_1 -norm minimization.

3.1.2. Minimizing L_1 -norm via Gradient Descent on $\Psi[n, \omega_0] \theta$

Sparseness is promoted by minimizing the L_1 -norm, which is defined as the sum of the absolute values of the vector components. For the parameter vector θ :

$$\|\theta\|_1 = \sum_{r=1}^N |\alpha_r| \quad (7)$$

Algorithm 2: L_1 -Norm Minimization via Gradient Descent / Subgradient Method

1. **Initialize:** Define starting parameter vector as a zero vector $\theta_0 = [0, 0, \dots, 0]^T$. Set step size $\mu > 0$. Set regularization parameter λ . Set $k = 0$.
2. **Formulate Lagrangian Objective:** $J(\theta) = \|\theta\|_1 + \frac{1}{\lambda} \sum_{n=0}^{M-1} (\Psi[n, \omega_0] \theta_k - u_n)^2$.
3. **Repeat** until convergence:

Compute the gradient of the quadratic error term:

$$\nabla E(\theta_k) = \frac{2}{\lambda} \sum_{n=0}^{M-1} \Psi[n, \omega_0]^T (\Psi[n, \omega_0] \theta_k - u_n)$$
 - a. Compute the subgradient of the L_1 -norm: $\text{sgn}(\theta_k)$.
 - b. Update the parameter vector via steepest descent: $\theta_{k+1} = \theta_k - \mu (\text{sgn}(\theta_k) + \nabla E(\theta_k))$.
 - c. Update iteration index: $k = k + 1$.
4. **Output:** Optimal sparse parameter vector $\tilde{\theta} = \theta_k$.

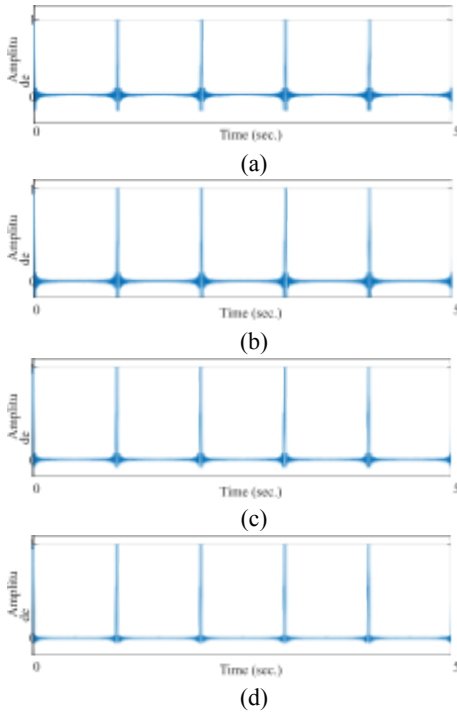


Figure 2. (a) The simulated source signal with $N = 50$ and $f_0 = 1$ Hz, (b) the generated signal after kurtosis maximization, (c) after skewness maximization, and (d) L_1 -norm minimization, where the signals are scaled to have unit amplitude. The relative improvement in time resolution after optimizing each function is evident.

3.1.3. Simulations

The performance and relevance of the functions in promoting sparsity can be compared in the absence of noise ($z_n = 0$) and linear distortion (trivial h_n) by seeking the weights producing the sparsest solution on the given signal model. The Θ that maximizes the skewness and kurtosis, or minimizes the L_1 -norm, should produce a compressible signal (or, with a stricter notion, sparse) on the basis that the atoms (columns) are scaled by $\tilde{\alpha}_r$ with a magnitude that decays according to a power law. In this context, a source signal is simulated as a sum of 50 equally strong harmonics ($N = 50$), resulting in spikes that are 1 sec. long. apart ($f_0 = 1$ Hz) as shown in Fig. 2a. Maximizing the kurtosis and skewness until convergence produced Θ settings as shown in Figure 1a and Figure 1b, respectively, whereas minimizing the L_1 -norm resulted in a setting as depicted in Figure 1c. Here, to solve problem (7) designed on the signal model in (1), CVX was utilized, a package for specifying and solving convex programs [49, 50].

3.1.4 Simulation Parameters and Practical Implications

To ensure the reproducibility of the numerical experiments, the key hyperparameters utilized in the simulations are explicitly defined and justified in Table 1. The selection of these hyperparameters directly impacts the recovery performance, computational complexity, and real-time feasibility of the methods. For instance, increasing the filter length L or the number of harmonics N improves the temporal resolution of the extracted impulses but significantly increases the computational cost per iteration.

Table 1. Summary of Simulation Parameters

Parameter	Symbol	Value	Justification
Number of Harmonics	N	50 (clean) / 15 (noisy)	Chosen based on presumed sensor bandwidth limits and required signal resolution.
Fundamental Frequency	f_0	1 Hz / 50 Hz	Selected to emulate typical outer-race bearing fault frequencies (BPFO).
Step Size (Skewness)	μ_s	10^{-2}	Empirically tuned to balance convergence speed and algorithmic stability.
Step Size (Kurtosis)	μ_k	10^{-3}	Empirically tuned; required a smaller step than skewness to prevent divergence.
Number of Iterations	k_{max}	600	Determined empirically as the threshold where HOS algorithms reached a steady-state optimum.
FIR Filter Length	L	225	Chosen to encompass enough impact periods for accurate statistical estimation.
Regularization Param.	λ	0.5	Tuned to achieve an optimal trade-off between strict sparsity and data fidelity.
Window Length	M	3 seconds (depending on the sampling frequency)	Defined to capture enough mechanical cycles within the observation timeframe.

Furthermore, the step sizes (μ) dictate the stability of the gradient search; a step size that is too large causes the adaptation to diverge, while a conservatively small step size drastically increases the required number of iterations,

delaying diagnosis. Notably, the L_1 -norm minimization demonstrated convergence in substantially fewer iterations (e.g., 9 iterations) compared to HOS maximization (600 iterations), highlighting its better computational efficiency for

potential real-time condition monitoring. Initialization also plays a critical role in these algorithms. For HOS maximization (Algorithm 1), which navigates non-convex multi-modal surfaces, the parameter vector was initialized uniformly as $\theta_0 = [1, 1, \dots, 1]^T / \sqrt{N}$ to provide an unbiased starting point with unit energy. Conversely, for the L_1 -norm minimization (Algorithm 2), the parameter vector was initialized as a zero vector ($\theta_0 = 0$). Because the L_1 -norm formulation poses a convex optimization problem, it inherently guarantees convergence to the global optimum regardless of the starting point. Furthermore, when the adaptation is shifted from the harmonic weights (θ) to a time-domain FIR filter ($\mathbf{f}$) in Section 3.2, the filter is initialized as a center-spike (Dirac delta) sequence (e.g., $\mathbf{f}_0 = [0, \dots, 0, 1, 0, \dots, 0]^T$) to pass the initial signal unaltered before gradient updates commence.

Theoretical Sensitivity and Robustness: While an exhaustive empirical robustness analysis is beyond the scope of this study, the theoretical implications of parameter variations are critical for practical implementation. In gradient-based HOS maximization, the step size (μ) is the most sensitive parameter; slight perturbations can drastically alter the convergence landscape, with excessively large values causing rapid divergence, particularly at lower Signal-to-Noise Ratios (SNR). Conversely, the L_1 -norm minimization relies heavily on the regularization parameter (λ), which strictly dictates the trade-off between absolute sparsity and data fidelity. If λ is set too high in a low-SNR environment, the algorithm may oversmooth the signal, suppressing true physical effects along with noise. Generally, as background noise increases (lower SNR), all discussed methods theoretically require larger observation windows or longer FIR filter lengths to maintain reliable gradient estimates, albeit at the cost of significantly increased computational overhead.

In accordance with these results, it can be concluded that the coefficients α_r . Scaling the harmonics does not cause them to decay rapidly after HOS maximization; rather, the disparity between them decreases with r . On the other hand, seeking the sparsest solution as in (7) produced weights that decay relatively more quickly, making it a more proper measure of sparsity. Synthesizing the signals on the given basis with the resulting parameter vector θ^* shows the improvement in the sparsity in the output signals. The ripples due to the finite number of harmonics (limited bandwidth) are suppressed, resulting in the sparsest signal after L_1 -norm minimization as illustrated in Figure 2d. Skewness-maximizing output, as in Figure 2c, yields a better approximation than kurtosis-maximizing output, as in Figure 2b.

Such a performance assessment was also done on different functions [46], where kurtosis was normalized, satisfying more constraints than its original counterpart. As mentioned before, the purpose is to find a sparsity-enhancing function that yields a parameter setting with a few strong coefficients that characterize the entire signal on a proper basis. The method in (7) appears to better satisfy this expectation than maximizing the HOS in the model of (1). The next section will investigate the case where $z_n \neq 0$.

3.2. Comparison Of Sparsity-Enhancing Functions in the Presence of Noise and Linear Distortion

In case the source signal is subject to amplitude and phase distortion and corrupted by measurement noise, the goal is to approximately invert the disturbing system by promoting sparsity [51, 52] from noisy measurements [48, 53, 54]. To extract sparse periodic impulses arising from defects, the observed signal u_n is iteratively filtered by a finite impulse response (FIR) filter to maximize the higher-order statistics objective functions as defined in Eq. 4 and Eq. 5. Similarly, promoting sparseness is achieved by minimizing the L_1 -norm as defined in Eq. 7. The processing is accomplished as in Sec. III-A1 in an iterative manner (with index k), where the gradient is estimated with respect to the filter f_k^n instead of θ^k .

The output signal $v_n = \mathbf{f}^T \mathbf{u}_n$ is expected to provide conclusive evidence of the existence of a defect, where the adaptation process suppresses the undesired components. Here, the constraint on the filter vector is still valid $\|\mathbf{f}^k\|_2 = 1$, as was the case for the parameter vector of weights θ^k .

For the same purpose of recovering the signal of interest by restoring its structure, sparseness can also be promoted in noisy observations by minimizing the L_1 -norm, as defined in (7). To conduct experiments with these methods, the source signal was simulated as a sum of 15 harmonics ($N = 15$) with a fundamental frequency of $f_0 = 50$ Hz to form a sparse impulsive signal representing an outer race defect as shown in Figure 3a. To simulate the linear distortion introduced by the mechanical transfer function h_n , each r^{th} harmonic component of the source signal was subjected to arbitrary amplitude attenuation and phase shifts. Specifically, the intermediate distorted sequence x_n was numerically generated as

$$x_n = \sum_{r=1}^N \alpha_r \cos(r\omega_0 n + \gamma_r), \text{ where the weights } \alpha_r \in (0, 1]$$

and phase shifts $\gamma_r \in [0, 2\pi)$ represent the magnitude and phase response of the transmission path, respectively. Following this linear degradation, additive white Gaussian noise z_n was introduced to yield the final observed signal $u_n = x_n + z_n$, as illustrated in Figure 3b.

For this illustrative simulation, the additive white Gaussian noise z_n was scaled to yield a Signal-to-Noise Ratio (SNR) of approximately 8.5 dB. This noisy signal is processed by an FIR filter maximizing the kurtosis of its output iteratively with a fixed step size $\mu = 10^{-3}$, resulting in a signal as shown in Figure 6a, whereas maximizing the skewness with a fixed step size $\mu = 10^{-2}$ produced an output as in Figure 6b. In both cases, the algorithms were iterated for 600 iterations at equal filter lengths, allowing the HOS to converge.

In any sparsifying dictionary, the source signal's sparsity depends on the rotation speed, provided we use a constrained local observation window. This creates a relation between sparseness and rotation speed, which can further be analyzed for diagnostics of slowly rotating machinery.

After noting this remark, the result of promoting sparseness in u_n via the objective formulated in (7) is shown in Figure 6c, where the defect-related impulses are recovered after 9 iterations. The parameter vector θ^* forming the sparse output signal $v_n = \Psi[n, \omega_0] \theta^*$ from the real Fourier basis consisted of almost equally strong weights α_r , unlike the case in the absence of noise and linear distortion.

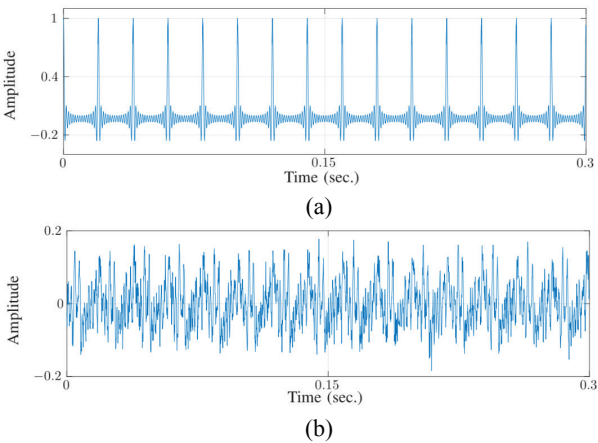


Figure 3. (a) The simulated source signal s_n as a sparse impulse train and (b) the observed signal u_n after linear distortion and additive noise.

To provide an objective, quantitative assessment of recovery performance beyond visual inspection, the Mean Squared Error (MSE) and Pearson correlation coefficient (r) were calculated between the impulsive source signal (s_n) and the final reconstructed outputs (v_n). Table 2 summarizes the test-case results.

Table 2. Quantitative Performance Metrics for Signal Recovery

Optimization Method	MSE	Correlation Coefficient (r)
Kurtosis Maximization	0.00523	0.9248
Skewness Maximization	0.00519	0.9317
L_1 -norm Minimization	0.00483	0.9428

These metrics align with the visual findings, demonstrating that the L_1 -norm minimization not only converges in significantly fewer iterations but also yields a reconstructed signal that mathematically correlates more strongly with the original sparse impulse train, minimizing the residual error compared to the HOS-based approaches.

Noise Robustness and Breakdown Limits: To rigorously evaluate the noise tolerance of the compared objective functions, additional simulations were conducted across a wide range of Signal-to-Noise Ratios (SNR = [5, 0, -5, -10, -15] dB) under purely additive noise conditions. The evaluation metrics, Mean Squared Error (MSE) and Pearson correlation coefficient (r), are visually presented in Figure 4 and Figure 5, respectively. The results reveal a stark contrast in robustness between the methods. As the background noise begins to overpower the defect-induced impacts, higher-order statistics experience catastrophic breakdowns. Specifically, kurtosis maximization collapses below 0 dB, showing a sharp MSE increase (Figure 4a) and dropping to near-zero correlation (Figure 5a) because 4th-order moments exponentially amplify noise variance. Skewness maximization exhibits a similar, albeit slightly softer, degradation (Figure 4b

and Figure 5b). Conversely, the L_1 -norm minimization demonstrates exceptional resilience, maintaining a minimal MSE (Figure 4c) and a correlation coefficient above 0.82 (Figure 5c) even at extreme noise levels (-15 dB). These breakdown limits theoretically confirm that while HOS methods require strict pre-whitening in noisy environments, L_1 -norm minimization provides a vastly superior and more stable approximation of sparse fault signatures under heavy background noise.

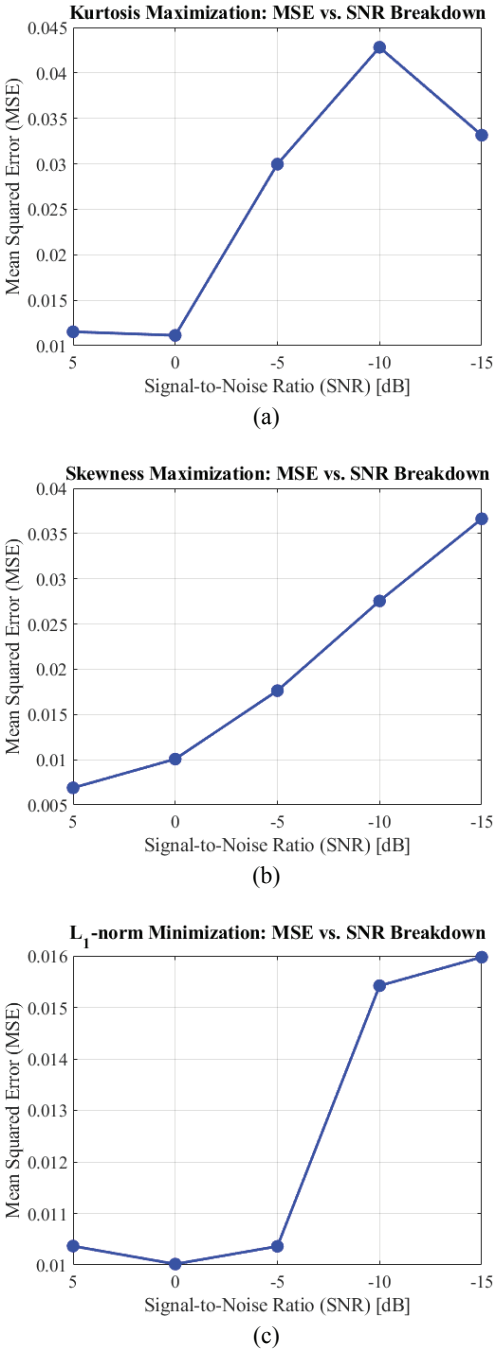


Figure 4. The breakdown limits of the objective functions evaluated by Mean Squared Error (MSE) under varying Signal-to-Noise Ratios (SNR): (a) Kurtosis maximization, (b) Skewness maximization, and (c) L_1 -norm minimization.

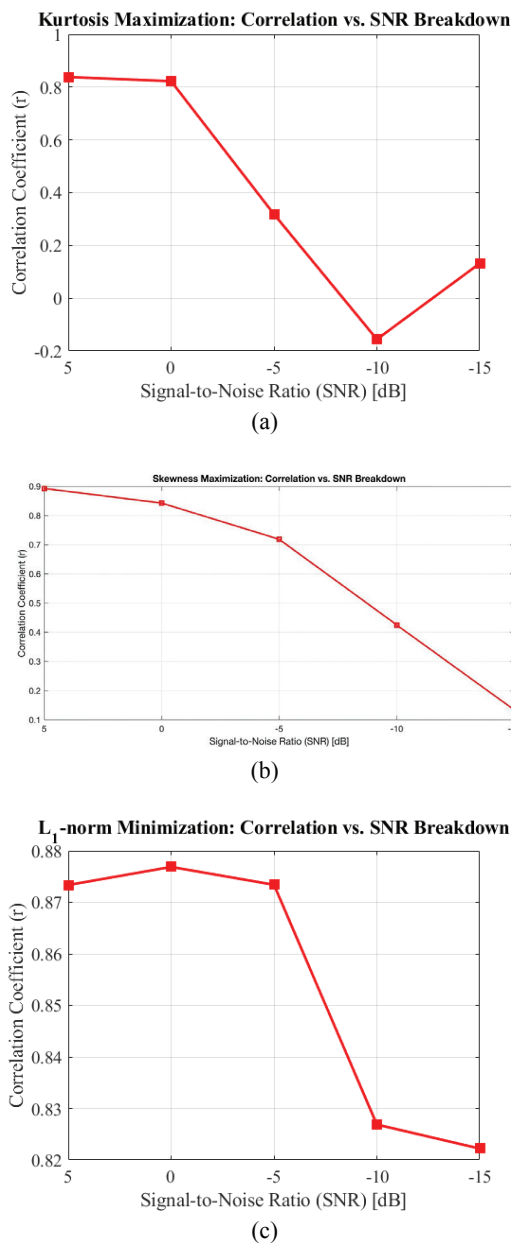


Figure 5. The breakdown limits of the objective functions evaluated by Pearson correlation coefficient (r) under varying Signal-to-Noise Ratios (SNR): (a) Kurtosis maximization, (b) Skewness maximization, and (c) L_1 -norm minimization.

3.3. Discussion on Scope and Applicability

While this study primarily uses an outer race defect model to demonstrate the theoretical framework, the underlying hypothesis of sparse, periodic impulses naturally extends to other localized faults, such as inner race or rolling-element defects. However, practitioners should note that inner race faults often exhibit amplitude modulation due to the rotating load zone. This modulation introduces sidebands that slightly deviate from the strictly uniform harmonic weights modeled in our pristine signal s_n , though the core sparsity assumptions remain valid.

Regarding the selection of optimization methods, L_1 -norm minimization is highly advantageous-and recommended-when the structural kinematics (e.g., shaft speed and bearing geometry) are known a priori. In such cases, the basis matrix $\Psi[n, \omega_0]$ can be accurately constructed, and the algorithm will robustly handle high noise levels and strong linear distortion while guaranteeing global convergence. Conversely, HOS-based blind deconvolution (such as skewness maximization) might still be preferable in purely exploratory or "blind" scenarios. If the exact machinery kinematics are unknown, an adaptive FIR filter that maximizes HOS does not require a predefined frequency basis to extract an impulsive signature, making it a flexible, albeit mathematically riskier, alternative, given the potential for local convergence issues.

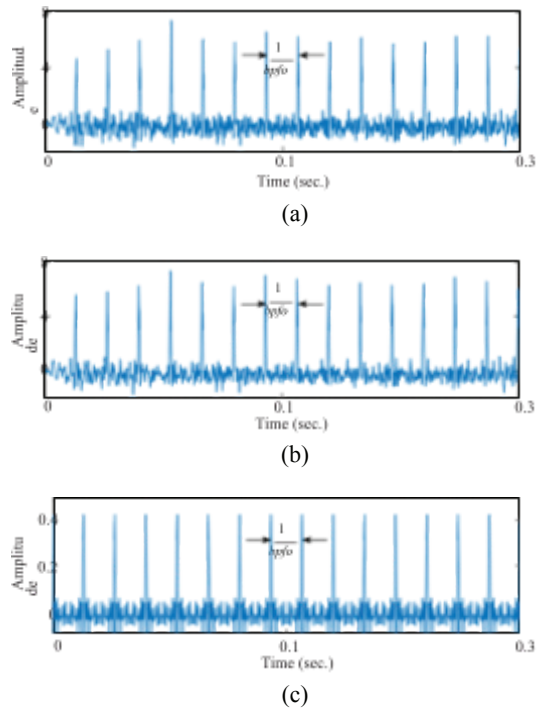


Figure 6. (a) The output signal v_n after kurtosis maximization, (b) after skewness maximization, and (c) after L_1 -norm minimization, zoomed in and marked with outer race ball-pass period, the inverse of the ball-pass frequency outer race ($1/bpfo$).

4. Conclusion

This study provided a theoretical and numerical comparison of Higher-Order Statistics (skewness and kurtosis) and L_1 -norm minimization for recovering sparse periodic fault signatures in rolling element bearings. Using a real Fourier basis as an analytical dictionary, we showed that both methods could suppress noise and compensate for linear distortion, revealing the underlying impact trains.

However, key performance differentiators were identified. Our numerical experiments L_1 -norm minimization achieves a statistically superior approximation of the true sparse signal, with a more rapid decay of harmonic weights compared to HOS-based methods. Crucially, the computational L_1 -norm approach converged in significantly fewer iterations (typically under 10) than the 600 iterations required for HOS

maximization to reach steady state. Furthermore, while HOS objectives are prone to local convergence issues depending on initialization, L_1 -norm minimization offers robust global convergence properties within the established basis. These findings suggest that for predictive maintenance regimes where bearing kinematics are known, L_1 -based sparsity enhancement provides a more efficient and mathematically sound framework for automated fault diagnosis.

A notable limitation of the current study is the absence of experimental validation using real-world measured data. To strictly isolate and mathematically compare the convergence dynamics and coefficient decay of the optimization functions on a real Fourier basis, this study relied on a controlled synthetic framework. Consequently, the performance of the proposed approach on benchmark empirical data remains untested. Validating the proposed L_1 -norm minimization framework on publicly available datasets-such as the SU RF Sensing Lab, CWRU, or Paderborn bearing datasets-constitutes a primary direction for future work. Applying this analytical dictionary approach to such real signals would be highly feasible in subsequent studies, provided that the complex, time-varying structural transfer paths and non-Gaussian mechanical background noises inherent to physical machinery can be adequately decoupled or robustly integrated into the measurement model.

5. References

- [1] K. Wang, H. Jiang, B. Hai and R. Yao, "Rolling bearing fault feature detection using nonconvex wavelet total variation", *Measurement*, vol. 179, pp. 109471, 2021.
- [2] S. Abrar, A. Zerguine and K. A. Meraim, "Adaptive blind equalization in impulsive noise", 2020 54th Asilomar Conference on Signals, Systems, and Computers, pp. 1415-1419, 2020.
- [3] Z. Yu and B. Zhang, "The art of noise modeling and mitigation on power line communication in distributed energy system", *IEEE Transactions on Instrumentation and Measurement*, vol. 74, pp. 1-23, 2025.
- [4] M. Zhan, Z. Pang, D. Dzung, K. Yu and M. Xiao, "Short-packet interleaver against impulse interference in practical industrial environments", *IEEE Transactions on Wireless Communications*, vol. 21, no. 12, pp. 10257-10270, 2022.
- [5] L. He, J. He, Q. Zhou, C. Yi and D. Wang, "Non-iterative optimal blind deconvolution and its application to machine condition monitoring", *Mechanical Systems and Signal Processing*, vol. 243, pp. 113715, 2026.
- [6] A. K. Ovacıklı, P. Pääjärvi, J. P. LeBlanc, and J. E. Carlson, "Recovering periodic impulsive signals through skewness maximization," *IEEE Transactions on Signal Processing*, vol. 64, no. 6, pp. 1586–1596, March 2016.
- [7] S. Li, Y. Han, J. Gaber, S. Yang and Q. Yang, "A multi-antenna spectrum sensing method based on CEEMDAN decomposition combined with wavelet packet analysis", *Electronics*, vol. 12, no. 18, pp. 3823, 2023.
- [8] A. Gabrielli, M. Battarra, E. Mucchi and G. Dalpiaz, "A procedure for the assessment of unknown parameters in modeling defective bearings through multi-objective optimization", *Mechanical Systems and Signal Processing*, vol. 185, pp. 109783, 2023.
- [9] A. K. Ovacıklı, J. E. Carlson, and P. Pääjärvi, "Blind pulse compression through skewness maximization on overlapping echoes from thin layers", in *Ultrasonics Symposium (IUS)*, 2016 IEEE International. IEEE, 2016, pp. 1–4.
- [10] J. E. Carlson, A. K. Ovacıklı, and P. Pääjärvi, "Material impulse response estimation from overlapping ultrasound echoes using a compressed sensing technique", in *Ultrasonics Symposium (IUS)*, 2017 IEEE International. IEEE, 2017, pp. 1–4.
- [11] J. E. Carlson, P. Gupta and N. Kumar, "Estimation of compound layer thickness and porosity in nitrocarburized hardening", 2024 IEEE Ultrasonics, Ferroelectrics, and Frequency Control Joint Symposium (UFFC-JS), pp. 1-4, 2024.
- [12] S. Liang, B.E. Treeby and E. Martin, "Review of the low-temperature acoustic properties of water, aqueous solutions, lipids, and soft biological tissues", *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 21, no. 5, pp. 607-620, 2024.
- [13] N. Kumar and R. K. Satapathy, "Bearings in aerospace, application, distress, and life: a review", *Journal of Failure Analysis and Prevention*, vol. 23, pp. 915–947, 2023.
- [14] E. Park, Y. Jung, I. Kim and U. Lee, "Charlie and the semi-automated factory: data-driven operator behavior and performance modeling for human-machine collaborative systems", *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, pp. 1-16, 2023.
- [15] M. Romanssini, P. C. C. de Aguirre, L. C. Severo and A. G. Girardi, "A review on vibration monitoring techniques for predictive maintenance of rotating machinery", *Eng.*, vol. 4, no. 3, pp. 1797-1817, 2023.
- [16] Y. Bai, W. Cheng, W. Wen and Yang Liu, "Application of time-frequency analysis in rotating machinery fault diagnosis", *Shock and Vibration*, vol. 2023, no. 1, pp. 9878228, 2023.
- [17] R. P. Masini, M. C. Medeiros and E. F. Mendes, "Machine learning advances for time series forecasting", *Journal of Economic Surveys*, vol. 37, no. 1, pp. 76-111, 2023.
- [18] B. Hou, Y. Wang and D. Wang, "Investigations on multiclass classification model-based optimized weights spectrum for rotating machinery condition monitoring", *Journal of Dynamics, Monitoring and Diagnostics*, vol. 4, no. 3, pp. 194-202, 2025.
- [19] M. Eppert, S. Mukherjee and D. Ghoshdastidar, "Recovering imbalanced clusters via gradient-based projection pursuit", vol. 212, pp. 105530, 2026.
- [20] S. Li, Y. Gao, J. Wu and L. Cao, "Blind deblurring of astronomical images using a SCGTV-based single-frame method", vol. 32, no. 20, pp. 35579-35593, 2024.
- [21] S. Wu, Y. Li, J. Wang, H. Zhou, C. Tang, Y. Guan, F. Guo and X. Li, "Numerical MCRT simulation and inter-symbol interference analysis of laser multipath transmission in airborne air-to-air channels", *Photonics*, vol. 11, no. 3, pp. 280, 2024.
- [22] Z. Huang and L. Cao, "Quantitative phase imaging based on holography: trends and new perspectives", *Huang and Cao Light: Science & Applications*, vol 13, no. 145, pp 1-43, 2024.
- [23] L. F. González, M. F. Bande and W. G. Manteiga, "A critical review of lasso and its derivatives for variable

- selection under dependence among covariates", vol. 90, no. 1, pp. 118-145, 2022.
- [24] A. M. Nassef, M. A. Abdelkareem, H. M. Maghrabie and A. Baroutaji, "Review of metaheuristic optimization algorithms for power systems problems", *Sustainability*, vol. 15, no. 12, pp. 9434, 2023.
- [25] R. Heckel, M. Jacob, A. Chaudhari, O. Perlman and E. Shimron, "Deep learning for accelerated and robust MRI reconstruction", *Magnetic Resonance Materials in Physics, Biology and Medicine*, vol. 37, pp. 335-368, 2024.
- [26] E. J. Candès et al., *Compressive sampling*, in *Proceedings of the International Congress of Mathematicians*, vol. 3. Madrid, Spain, 2006, pp. 1433-1452.
- [27] Y. F. Liu, T. H. Chang, M. Hong, Z. Wu, A. M. C. So and E. A. Jorswieck, "A survey of recent advances in optimization methods for wireless communications", *IEEE Journal on Selected Areas in Communications*, vol. 42, no. 11, pp. 2992-3031, 2024.
- [28] Y. Zhang, E. Yang, H. H. Yoon, Q. Cheng, Z. Sun, T. Hasan and W. Cai, "Reconstructive spectrometers: hardware miniaturization and computational reconstruction", *eLight*, vol. 5, no. 23, pp. 1-43, 2025.
- [29] S. Kalra and N. K. Shukla, "Uncertainty principle for convolutional tight frames and its applications in signal recovery", *Sampling Theory, Signal Processing, and Data Analysis*, vol. 23, no. 21, 2025.
- [30] Y. Tian and Y. Zhang, "A comprehensive survey on regularization strategies in machine learning", vol. 80, pp. 146-166, 2022.
- [31] A. Dades and R. Daher, "New uncertainty principles for the free metaplectic wavelet transform", *Integral Transforms and Special Functions*, vol. 37, no. 2, pp. 83-96, 2026.
- [32] R. B. Mhenni, S. Bourguignon and J. Idier, "A greedy sparse approximation algorithm based on L1-norm selection rules", *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 5390-5394, 2020.
- [33] K. N. Filho, R. Lopes, R. D. B. Brotto, T. C. Senna and J. M. T. Romano, "Algorithms for sparse multichannel blind deconvolution", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1-7, 2023.
- [34] Z. Liu, M. Wang, Y. Zhu, X. Liu, Y. Hu and C. He, "A sparse optimization model of 12-11-2 for estimating guided waves dispersion curves", *Measurement*, vol. 253, pp. 117399, 2025.
- [35] Y. Zhao and B. Guo, "A hybrid signal decomposition and adaptive deep deconvolution method for rotating machinery fault diagnosis", *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, vol. 48, no. 11, 2026.
- [36] S. Ereiz, I. Duvnjak and J. F. J. Alonso, "Review of finite element model updating methods for structural applications", *Structures*, vol. 41, pp. 684-723, 2022.
- [37] A. C. Dumitraşcu, D. Cojocaru, V. Cârlescu, A. Zaharia and D. Olaru, "Influence of the cage on the friction torque in miniature angular contact ball bearings", *Lubricants*, vol. 14, no. 4, pp. 1-18, 2026.
- [38] R. N. A. Algburi, H. S. S. Aljibori, Z. Al-Huda, Y. H. Gu and M. A. Al-antari, "Advanced fault diagnosis in industrial robots through hierarchical hyper-laplacian priors and singular spectrum analysis", *Complex & Intelligent Systems*, vol. 11, no. 282, 2025.
- [39] P. P. Yuan, Z. J. Zhao and W. X. Ren, "A combination of improved spline-kernelled chirplet transform and multi-synchrosqueezing technique for analyzing nonstationary signals and structural instantaneous frequency identification", *Mechanical Systems and Signal Processing*, vol. 223, pp. 111873, 2025.
- [40] J. Yang, W. Bao, Y. Liu and X. Li, "Class metric regularized deep belief network with sparse representation for fault diagnosis", *International Journal of Intelligent Systems*, vol. 37, no. 9, pp. 5996-6022, 2022.
- [41] K. K. R. Nareddy, A. J. Kamath and C. S. Seelamantula, "Tight-frame-like analysis-sparse recovery using nontight sensing matrices", *SIAM Journal on Imaging Sciences*, vol. 17, no. 3, pp. 1587-1618, 2024.
- [42] A. K. Ovacıklı, P. Pääjärvi, and J. P. LeBlanc, Skewness as an objective function for vibration analysis of rolling element bearings, in *8th International Symposium on Image and Signal Processing and Analysis (ISPA)*. IEEE, Sep. 2013, pp. 462-466.
- [43] C. Peeters, J. Antoni and J. Helsen, "Blind filters based on envelope spectrum sparsity indicators for bearing and gear vibration-based condition monitoring", *Mechanical Systems and Signal Processing*, vol. 138, pp. 106556, 2020.
- [44] L. C. Kwek, A. W. C. Tan, H. S. Lim, C. H. Tan and K. A. Alaghbari, "Sparse representation and reproduction of speech signals in complex fourier basis", *International Journal of Speech Technology*, vol. 25, pp. 211-217, 2022.
- [45] A. K. Ovacıklı, P. Pääjärvi, J. P. LeBlanc, and J. E. Carlson, Uncovering Harmonic Content via Skewness Maximization - A Fourier Analysis, in *Signal Processing Conference (EUSIPCO), 2014 Proceedings of the 22nd European. IEEE, 2014*, pp. 481-485.
- [46] C. Peeters, J. Antoni and J. Helsen, "A multi-delay extension of the discrete/random separation method", *Mechanical Systems and Signal Processing*, vol. 236, pp. 112959, 2025.
- [47] S. Yoon, H. Yang and W. Seong, "Ray-based blind deconvolution with maximum kurtosis phase correction", *J. Acoust. Soc. Am*, vol. 151, pp. 4237-4251, 2022.
- [48] K. Chen, C. Ding, B. Wang, J. Wang, W. Huang and Zhongkui Zhu, "Interpretable transfer unrolling network for intelligent cross-domain condition monitoring of rotating machinery", *IEEE Transactions on Instrumentation and Measurement*, vol. 74, pp. 1-12, 2025.
- [49] C. Pan, G. Zhou, K. Zhi, S. Hong, T. Wu and Y. Pan, "An overview of signal processing techniques for ris/irs-aided wireless systems", *IEEE Journal of Selected Topics in Signal Processing*, vol. 16, no. 5, pp. 883-917, 2022.
- [50] D. Malyuta, T. P. Reynolds, M. Szmuk, T. Lew, R. Bonalli and M. Pavone, "Convex optimization for trajectory generation: a tutorial on generating dynamically feasible trajectories reliably and efficiently", *IEEE Control Systems Magazine*, vol. 42, no. 5, pp. 40-113, 2022.
- [51] S. Ravishankar, J. C. Ye and J. A. Fessler, "Image reconstruction: from sparsity to data-adaptive methods

- and machine learning”, Proceedings of the IEEE, vol. 108, no. 1, pp. 86-109, 2020.
- [52] B. Chen, L. Liu, Z. Zou and Z. Shi, “Target detection in hyperspectral remote sensing image: current status and challenges”, Remote Sensing, vol. 15, no. 13, pp. 3223, 2023.
- [53] C. Zhang, J. Qin, Q. Ma, Y. Shi and M. Li, “Resilient distributed state estimation for lti systems under time-varying deception attacks”, vol. 10, no. 1, pp. 381-393, 2023.
- [54] S. Grebien, E. Leitinger, K. Witrisal and B. H. Fleury, “Super-resolution estimation of uwb channels including the dense component-an sbl-inspired approach”, IEEE Transactions on Wireless Communications, vol. 23, no. 8, pp. 10301-10318, 2024.
- [55] J.-J. Fuchs, Detection and estimation of superimposed signals, in Acoustics, Speech and Signal Processing, 1998. Proceedings of the 1998 IEEE International Conference on, vol. 3. IEEE, 1998, pp. 1649–1652.

Özgeçmişler



Aziz Kubilay OVACIKLI was born in Erzincan, Turkey, in 1985. He received his B.S. degree in Electrical Engineering from Anadolu University, Eskişehir, Turkey, in 2007, and his M.Sc. and Ph.D. degrees in Signal Processing from Luleå University of Technology, Sweden, in 2010 and 2017, respectively. He has been conducting research on blind adaptive systems applied to vibration analysis and ultrasound. He currently serves as the CTO at Ascribo Applied Sciences Inc. in Chicago, USA, and as an Assistant Professor at Istanbul Arel University, Istanbul, Turkey. His research interests encompass Internet of Things (IoT), predictive maintenance, anomaly and fault detection, industrial artificial intelligence and AI-driven optimization for engineering applications.

A Hybrid Stochastic-Robust Optimization Framework for Industrial Energy Management under Renewable Generation and Price Uncertainty

Yenilenebilir Enerji Üretimi ve Fiyat Belirsizliği Altında Endüstriyel Enerji Yönetimi için Hibrit Stokastik-Sağlam Optimizasyon Çerçevesi

Erhan Baran¹

 0000-0002-5897-8680

¹Department of Electronic and Automation, TUSAŞ-Kazan Vocational School, Gazi University, Ankara, Türkiye
erhanbaran@gazi.edu.tr

Abstract

This study develops a hybrid stochastic-robust optimization framework for industrial energy management under photovoltaic generation and electricity price uncertainty. The proposed two-stage model integrates scenario-based scheduling, robust operational adjustments, battery storage dynamics, and carbon emission costs. Its effectiveness is evaluated using a grid-connected industrial microgrid and compared with deterministic, stochastic, and robust optimization approaches. The hybrid model achieves a daily operating cost of \$1,173.9, with 86.4% photovoltaic utilization and 751.8 kg of carbon emissions. Sensitivity analyses demonstrate that adjustable risk preferences improve protection against unfavorable operating conditions, while increased battery capacity supports renewable energy integration. These findings indicate that the proposed framework offers a flexible approach to balancing economic efficiency, operational resilience, and environmental performance in industrial energy systems.

Keywords: Hybrid optimization, Energy management, PV uncertainty, Battery storage, Industrial microgrids

Özet

Bu çalışma, fotovoltaik üretim ve elektrik fiyatı belirsizliklerine maruz kalan şebekeye bağlı endüstriyel mikro şebekelerde enerji yönetimi için iki aşamalı hibrit stokastik-sağlam optimizasyon çerçevesi önermektedir. Senaryo tabanlı planlama, sağlam gerçek zamanlı düzenlemelerle birleştirilirken batarya işletimi ve karbon maliyetleri karar modeline dahil edilmiştir. Önerilen yaklaşım, 24 saatlik endüstriyel bir vaka çalışmasında deterministik, stokastik ve sağlam optimizasyon modelleriyle karşılaştırılarak değerlendirilmiştir. Hibrit strateji, günlük 1.173,9 ABD doları işletme maliyeti,

%86,4 fotovoltaik enerji kullanım oranı ve 751,8 kg karbon emisyonu sağlamıştır. Duyarlılık analizi, risk korumasının artırılmasının en kötü durum maliyetlerini azalttığını, daha yüksek depolama kapasitesinin ise şebeke bağımlılığını düşürdüğünü göstermiştir. Genel olarak önerilen çerçeve, işletme maliyetleri, belirsizliklere maruz kalma ve çevresel faktörler arasında denge kuran uyarlabilir bir enerji planlama yaklaşımı sunmaktadır.

Anahtar Kelimeler: Hibrit optimizasyon, Enerji Yönetimi, Fotovoltaik üretim belirsizliği, Batarya depolama, Endüstriyel mikro şebekeler

1. Introduction

The rapid decarbonization of industrial energy systems has significantly increased the adoption of on-site renewable energy sources, particularly photovoltaic (PV) systems, within industrial microgrids. While renewable integration reduces operational costs and carbon emissions, it also introduces substantial uncertainty due to intermittency and market price volatility. Industrial energy management systems (EMS) are therefore required to operate under multi-dimensional uncertainty while maintaining economic efficiency, reliability, and environmental compliance [1].

Unlike traditional deterministic energy scheduling approaches, contemporary EMS frameworks must explicitly address uncertainties associated with renewable generation, electricity market prices, and demand fluctuations. In industrial environments—where production continuity, contractual power limits, and penalty costs are critical—suboptimal decisions caused by uncertainty mismanagement can lead to considerable financial losses [2].

To cope with uncertainty, two principal modeling paradigms have been widely adopted in the literature: stochastic programming and robust optimization. Stochastic programming models uncertainty through probabilistic scenarios and optimize expected system performance. Although effective in

capturing average-case behavior, these models may perform poorly under extreme realizations or rare but costly events. On the other hand, robust optimization provides protection against worst-case deviations within predefined uncertainty sets, ensuring solution feasibility under adverse conditions. However, purely robust formulations often lead to overly conservative strategies and inflated operational costs.

Despite extensive research in both domains, most existing studies adopt either stochastic programming or robust optimization independently. The lack of hybrid formulations that jointly balance expectation-based efficiency and worst-case resilience represents a critical methodological gap [3]. Particularly in industrial microgrids with PV generation and battery energy storage systems (BESS), there is limited research integrating both uncertainty treatment mechanisms within a unified EMS framework.

Moreover, recent regulatory developments introducing carbon pricing mechanisms have further complicated energy management decisions. Industrial operators are now required to consider not only electricity procurement and storage costs, but also carbon-related operational impacts [4]. Nevertheless, carbon cost integration into uncertainty-aware EMS formulations remains underexplored.

Motivated by these challenges, this study proposes a two-stage hybrid stochastic–robust optimization framework for industrial energy management under renewable generation and price uncertainty [5]. In the first stage, day-ahead energy scheduling decisions are optimized using scenario-based stochastic programming. In the second stage, real-time adjustments are safeguarded against adverse deviations through a budget-of-uncertainty robust counterpart formulation. A tunable risk-sensitivity parameter is introduced to systematically balance expected performance and worst-case resilience.

The major contributions of this study are summarized as follows:

- Development of a hybrid stochastic–robust EMS model for PV-integrated industrial microgrids.
- Integration of electricity price uncertainty, renewable generation variability, and carbon cost considerations within a unified optimization framework.
- Introduction of a risk-adjustable structure enabling decision-makers to control conservativeness levels.
- Comprehensive numerical analysis comparing deterministic, purely stochastic, purely robust, and proposed hybrid models.
- Managerial insights on cost–risk–emission trade-offs for industrial energy operators.

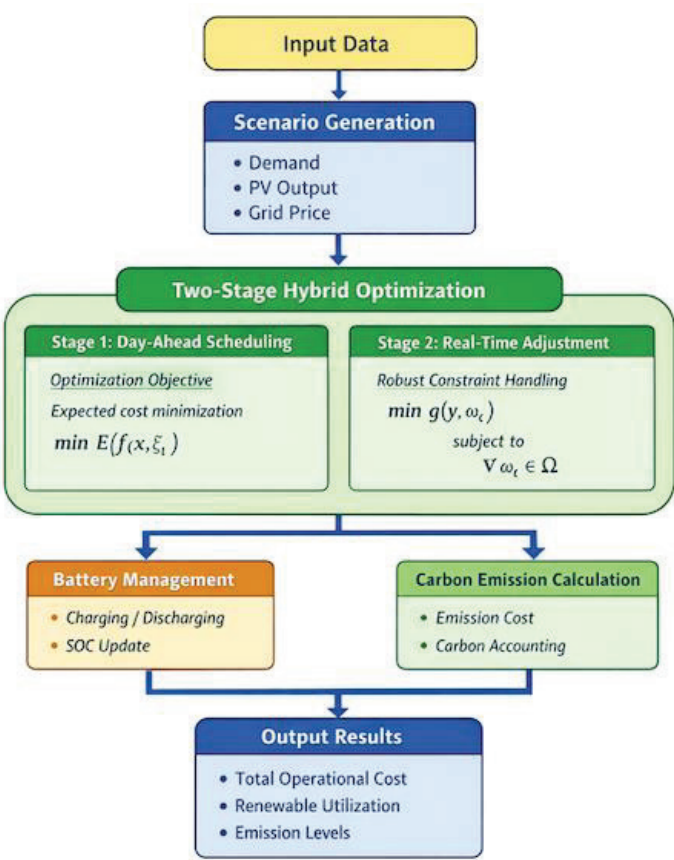


Figure 1: Hybrid stochastic-robust EMS framework for industrial microgrids

Hybrid stochastic-robust energy management systems (EMS) framework for industrial microgrids is presented in Figure 1.

The remainder of this paper is organized as follows: Section 2 reviews related literature on uncertainty-aware energy management. Section 3 presents the mathematical formulation of the hybrid model. Section 4 describes the case study and experimental setup. Section 5 discusses computational results and comparative analyses. Section 6 presents the sensitivity analysis. Finally, Section 7 concludes the study and outlines future research directions.

2. Literature Research

The increasing penetration of renewable energy resources in industrial energy systems has significantly intensified the complexity of energy management problems, particularly under uncertainty. In recent years, extensive research has been conducted on optimizing energy management systems (EMS) for microgrids integrating photovoltaic (PV) generation and battery energy storage systems (BESS). Early studies primarily relied on deterministic optimization techniques due to their computational efficiency and simplicity [6-7]. However, these approaches assume perfect knowledge of system parameters and fail to capture the stochastic nature of renewable generation and electricity market dynamics.

To address uncertainty, stochastic programming has been widely applied in EMS studies. Two-stage and multi-stage stochastic optimization models have been proposed to account for uncertainties in renewable generation, load demand, and electricity prices [8-10]. For instance, stochastic scheduling frameworks for microgrids have demonstrated improved cost efficiency and operational flexibility compared to deterministic models. Despite these advantages, stochastic approaches depend heavily on accurate probability distributions and scenario generation, and their performance may deteriorate under extreme or rare events [11].

Robust optimization has emerged as an alternative approach to address uncertainty without relying on probabilistic information. In robust EMS models, uncertain parameters are represented using bounded uncertainty sets, and solutions are optimized against worst-case realizations [12–15]. This approach ensures system reliability and feasibility under adverse conditions and has been successfully applied in microgrid operation and energy scheduling problems. However, robust optimization often results in overly conservative solutions, leading to increased operational costs and reduced economic efficiency.

In recent years, hybrid stochastic–robust optimization frameworks have gained attention as a promising approach to balance expected performance and worst-case protection. Several studies have proposed hybrid models that combine scenario-based stochastic programming with robust optimization techniques [16–18]. For example, two-stage hybrid frameworks have been applied in microgrid energy management, enabling improved trade-offs between cost efficiency and system reliability [19-20]. Nevertheless, the application of such hybrid models in industrial energy management systems remains relatively limited, particularly in the context of PV-integrated microgrids with battery storage.

At the same time, environmental considerations have become increasingly important in EMS research. With the implementation of carbon pricing mechanisms, several studies have incorporated carbon emission costs into energy management models [21]. These works highlight the importance of integrating economic and environmental objectives in energy system optimization. However, the interaction between carbon pricing and uncertainty modeling has not been sufficiently explored in existing literature.

Moreover, recent studies have explored advanced approaches such as machine learning-based forecasting, real-time optimization, and digital twin technologies to enhance EMS performance [22]. While these methods improve prediction accuracy and operational adaptability, they often lack rigorous uncertainty modeling frameworks capable of handling extreme deviations in renewable generation and electricity prices.

Based on the above discussion, it is evident that although significant progress has been made in uncertainty-aware EMS modeling, several important gaps remain. Existing studies predominantly focus on either stochastic programming or robust optimization, with limited integration of both approaches in a unified framework. Furthermore, few studies simultaneously consider renewable uncertainty, electricity price volatility, battery operation, and carbon emission costs within an industrial microgrid context. To address these

limitations, this study proposes a two-stage hybrid stochastic–robust optimization framework that integrates scenario-based uncertainty modeling with robust protection mechanisms, providing a flexible and risk-aware solution for industrial energy management.

3. Model Structure

3.1. System Description

This study considers an industrial microgrid composed of:

- Photovoltaic (PV) generation units
- Battery Energy Storage System (BESS)
- Utility grid connection
- Industrial electricity demand
- Carbon pricing mechanism

The system operates in a discrete time horizon $t \in T$ (e.g., hourly scheduling over 24 hours). The objective is to determine optimal day-ahead and real-time energy scheduling decisions under renewable generation and electricity price uncertainty.

Uncertainty sources include:

- PV generation variability
- Electricity market price fluctuation
- Load variability

To simultaneously capture expected performance and worst-case protection, a two-stage hybrid stochastic–robust optimization framework is developed.

3.2. Two-Stage Hybrid Decision Structure

3.2.1. First Stage (Here-and-Now Decisions)

These decisions are taken before uncertainty realization:

- Scheduled grid purchase P_t^{DA}
- Battery charging/discharging plan
- Energy procurement commitments

The first stage minimizes the expected operational cost over a discrete scenario set Ω .

3.2.2. Second Stage (Wait-and-See & Robust Adjustment)

After uncertainty realization, recourse decisions adjust:

- Real-time power deviation $\Delta P_{t,\omega}$
- Battery corrective dispatch
- Penalty handling

Robust optimization is applied within each scenario using a budget-of-uncertainty parameter Γ to hedge against worst-case deviations.

3.3. Objective Function

The hybrid objective integrates expected cost and worst-case protection:

$$\min \{ \mathbb{E}_{\omega \in \Omega} [C^{\text{stoch}}(x, \xi_{\omega})] + \lambda \cdot C^{\text{robust}}(x, \Gamma) \}. \quad (1)$$

where (x) denotes the vector of first-stage energy scheduling decisions, Ω is the set of stochastic scenarios, ω represents an individual scenario, ω denotes the vector of uncertain parameters associated with scenario ω , and $(C^{stoch}(\omega))$ represents the scenario-dependent stochastic operating cost. The parameter λ is the risk-sensitivity coefficient that determines the relative importance assigned to protection against adverse realizations, while C^{robust}/ω denotes the robust cost component associated with worst-case deviations. The parameter ω represents the uncertainty budget and controls the degree of conservativeness of the robust solution.

Operational Cost Components:

The total cost consists of:

- Electricity procurement cost
- Battery degradation cost
- Carbon emission cost
- Demand violation penalty

$$C = \sum_t (\pi_{t,\omega} P_{t,\omega} + c^{deg} B_{t,\omega} + C^{CO2} E_{t,\omega} + C^{pen} S_{t,\omega}) \quad (2)$$

where (T) denotes the set of scheduling periods, (t) represents the individual time period, π_t is the electricity price at time (t) under scenario ω , and P_t denotes the electricity purchased from the utility grid. The parameter c^{deg} represents the unit battery degradation cost, while $B_{t,\omega}$ denotes the battery energy throughput. Furthermore, c^{CO2} is the unit carbon emission cost and $E_{t,\omega}$ represents the corresponding carbon emissions. Finally, c^{pen} denotes the unit penalty cost and $S_{t,\omega}$ is the unmet-demand or operational violation variable.

3.4. Constraints

3.4.1. Power Balance Constraint

$$P_{t,\omega}^{grid} + P_{t,\omega}^{PV} + P_{t,\omega}^{dis} = D_{t,\omega} + P_{t,\omega}^{ch} \quad (3)$$

where $P_{t,\omega}$ represents the power purchased from the utility grid, $P_{t,\omega}^{PV}$ is the available photovoltaic generation, and $P_{t,\omega}^{dis}$ denotes the battery discharge power at time (t) under scenario ω . Similarly, $D_{t,\omega}$ represents the industrial electricity demand and $P_{t,\omega}^{ch}$ denotes the battery charging power. Equation (3) ensures that the total electricity supplied to the industrial microgrid is equal to the total electricity demand and battery charging requirement in each period and scenario.

3.4.2. Battery Dynamics

$$SOC_t = SOC_{t-1} + \eta^{ch} P_t^{ch} - \frac{1}{\eta^{dis}} P_t^{dis} \quad (4)$$

where $SOC_{t,\omega}$ denotes the battery state of charge at time (t) under scenario ω , $SOC_{t-1,\omega}$ is the state of charge in the preceding period, η^{ch} and η^{dis} represent the battery charging and discharging efficiencies, respectively, and Δt denotes the duration of each scheduling interval. The charging and

discharging powers are represented by $P_{t,\omega}^{ch}$ and $P_{t,\omega}^{dis}$, respectively.

The allowable battery state-of-charge range is constrained as

$$SOC^{min} \leq SOC_t \leq SOC^{max} \quad (5)$$

where SOC^{min} and SOC^{max} represent the minimum and maximum permissible battery state-of-charge levels, respectively. These limits prevent excessive charging or discharging and ensure technically feasible battery operation.

3.4.3. Grid Capacity Limit

$$0 \leq P_t^{grid} \leq P^{max} \quad (6)$$

Where $P_t^{grid} \leq P^{max}$ denotes the maximum power that can be imported from the utility grid. Equation (6) restricts grid electricity procurement within the technical connection capacity of the industrial microgrid.

3.4.4. Robust Uncertainty Set

PV output deviation:

$$P_t^{PV} = P_t^{PV} + \delta_t \quad (7)$$

$$|\delta_t| \leq \delta_t \quad (8)$$

where P_t^{PV} represents the uncertain realized PV generation, P_t^{PV} denotes the forecasted PV generation, δ_t is the deviation of actual PV output from its forecasted value, and δ_t represents the maximum allowable magnitude of this deviation at time (t) . Equation (8) therefore defines the interval within which PV generation uncertainty is allowed to vary.

Budget constraint:

$$\sum_t \frac{|\delta_t|}{\delta_t} \leq \Gamma \quad (9)$$

This structure ensures protection against simultaneous worst-case deviations up to Γ periods.

3.5. Hybridization Logic

The proposed model integrates stochastic and robust paradigms as follows:

- Scenario-based modeling captures distributional variability.
- Robust set within each scenario prevents extreme realizations.
- Risk parameter λ balances economic optimality and conservativeness.
- Budget parameter Γ controls robustness intensity.

3.6. Solution Approach

The model is formulated as a Mixed-Integer Linear Programming (MILP) problem.

Solution steps:

- Scenario generation via Monte Carlo simulation
- Reformulation of robust counterpart
- Linearization of bilinear terms
- Solution using CPLEX/Gurobi

To evaluate performance, four benchmark models are implemented:

- Deterministic EMS
- Pure stochastic EMS
- Pure robust EMS
- Proposed hybrid EMS

3.7. Managerial Interpretation

The hybrid model enables industrial operators to:

- Quantify risk–cost trade-offs
- Adjust conservativeness through λ and Γ
- Evaluate carbon price impact
- Reduce vulnerability to renewable volatility

4. Case Study and Experimental Setup

4.1. Industrial Microgrid Description

To evaluate the effectiveness of the proposed hybrid stochastic–robust energy management framework, a grid-connected industrial microgrid is considered as the case study. The system represents a medium-scale manufacturing facility equipped with on-site renewable generation and energy storage capabilities.

The microgrid consists of the following components:

- Photovoltaic (PV) generation system
- Battery Energy Storage System (BESS)
- Utility grid connection
- Industrial electricity demand

The energy management system determines optimal day-ahead and real-time operational strategies over a 24-hour scheduling horizon, discretized into hourly intervals.

In this study, industrial microgrid system parameters are shown in Table 1.

Table 1: Industrial microgrid description

Parameter	Symbol	Value	Unit
Scheduling horizon	T	24	h
Time interval	Δt	1	h
Maximum grid purchase power	P_{grid}^{max}	500	kW
Battery energy capacity	E_{BESS}^{max}	600	kWh
Maximum charging power	P_{ch}^{max}	150	kW

Maximum discharging power	P_{dis}^{max}	150	kW
Charging efficiency	η^{ch}	0.95	-
Discharging efficiency	η^{dis}	0.95	-
Minimum state of charge	SOC^{min}	0.20	p.u.
Maximum state of charge	SOC^{minmax}	0.90	p.u.
Initial state of charge	SOC^{rini}	0.50	p.u.
Battery degradation cost	c^{deg}	0.015	USD/kWh
Carbon price	c^{CO2}	40	USD/tCO ₂
Grid emission factor	EF	0.42	kgCO ₂ /kWh
Demand penalty cost	c^{pen}	0.30	USD/kWh

The industrial load profile reflects typical electricity consumption patterns observed in manufacturing plants, where energy demand is relatively stable during production hours and decreases during non-operational periods. Table 2 presents hourly industrial load profile.

Table 2: Hourly industrial load profile

Hour	Load Demand (kW)	Hour	Load Demand (kW)	Hour	Load Demand (kW)
1	210	9	355	17	390
2	205	10	380	18	360
3	200	11	395	19	335
4	198	12	410	20	300
5	202	13	420	21	275
6	220	14	415	22	250
7	260	15	405	23	230
8	310	16	398	24	220

Table 3: Forecasted PV generation profile

Hour	PV Generation (kW)	Hour	PV Generation (kW)	Hour	PV Generation (kW)
1	0	9	130	17	70
2	0	10	175	18	25
3	0	11	210	19	0
4	0	12	225	20	0
5	0	13	220	21	0
6	15	14	205	22	0
7	40	15	170	23	0
8	85	16	120	24	0

4.2. Photovoltaic Generation Modeling

PV generation is inherently uncertain due to weather variability. To capture this uncertainty, a scenario-based approach is adopted. Forecasted PV generation profile is given in Table 3.

Historical solar irradiation data are used to construct multiple PV generation scenarios. Monte Carlo simulation is employed to generate a set of representative scenarios reflecting possible PV output realizations throughout the scheduling horizon.

Ω denote the set of PV generation scenarios and p_ω represent the probability of scenario ω . The expected PV generation in scenario ω at time t is denoted as: $P_{t,\omega}^{PV}$

To reduce computational complexity, scenario reduction techniques are applied to retain the most representative scenarios while preserving statistical characteristics.

4.3. Electricity Price and Uncertainty

Electricity market prices are subject to fluctuations caused by demand variations, fuel prices, and market dynamics. In this study, day-ahead electricity prices are modeled as uncertain parameters. Day-ahead electricity price profile is given in Table 4.

Table 4: Day-ahead electricity price profile

Hour	Price (USD/kWh)	Hour	Price (USD/kWh)	Hour	Price (USD/kWh)
1	0.082	9	0.126	17	0.147
2	0.078	10	0.138	18	0.158
3	0.075	11	0.145	19	0.166
4	0.074	12	0.149	20	0.154
5	0.076	13	0.151	21	0.136
6	0.085	14	0.148	22	0.118
7	0.097	15	0.143	23	0.099
8	0.112	16	0.139	24	0.089

A scenario generation approach similar to PV modeling is used to represent electricity price uncertainty. Historical market data are used to generate price scenarios that capture realistic price volatility.

$\pi_{t,\omega}$ represent the electricity price at time t under scenario ω . These price scenarios are integrated into the stochastic component of the optimization model.

4.4. Battery Energy Storage System Modeling

The battery energy storage system plays a critical role in balancing supply and demand, mitigating renewable intermittency, and enabling cost-effective energy management. Uncertainty modeling parameters is given in Table 5.

Table 5: Uncertainty modeling parameters

Parameter	Symbol	Value
Number of scenarios	Ω	20
PV deviation bound	δ_t^{PV}	15%
Price deviation bound	δ_t^{price}	12%
Load deviation bound	δ_t^{load}	8%
Robustness budget	Γ	6
Risk preference coefficient	λ	0.35

The BESS is characterized by the following parameters:

- Maximum charging power
- Maximum discharging power
- Storage capacity
- Charging and discharging efficiencies

Battery degradation cost is incorporated into the objective function to account for operational wear associated with charging and discharging cycles.

The state-of-charge (SOC) evolution is governed by the battery dynamics described in Section 3.

4.5. Carbon Emission Consideration

To reflect increasing regulatory pressure on industrial emissions, carbon costs are integrated into the energy management model.

Electricity purchased from the grid is associated with an emission factor representing the carbon intensity of grid electricity. The corresponding carbon emission cost is calculated as:

$$C^{CO2} = C^{CO2}.EF.p_{grid}$$
 (10)

where $C_{t,\omega}^{CO2}$ represents the carbon-related cost associated with grid electricity consumption at time (t) under scenario ω , c^{CO2} is the carbon price, (EF) denotes the grid emission factor, $P_{t,\omega}^{grid}$ represents the electricity imported from the utility grid, and Δt is the duration of the scheduling interval. This formulation monetizes the environmental impact of grid electricity consumption and incorporates it directly into the energy management objective.

4.6. Benchmark Models

To assess the performance of the proposed hybrid approach, four different energy management models are implemented for comparison:

- Deterministic EMS
- Stochastic EMS
- Robust EMS
- Proposed Hybrid Stochastic–Robust EMS

Benchmark models is given in Table 6.

Table 6: Benchmark models used in the comparison

Model	Uncertainty Handling	Main Characteristic
Deterministic EMS	None	Uses forecasted average values only
Stochastic EMS	Scenario-based	Minimizes expected cost
Robust EMS	Uncertainty sets	Protects against worst-case deviations
Hybrid Stochastic-Robust EMS	Scenario + robust set	Balances expected performance and resilience

The deterministic model ignores uncertainty and uses expected values of uncertain parameters.

The stochastic model minimizes expected operational cost across scenarios.

The robust model protects against worst-case deviations within uncertainty sets.

The hybrid model simultaneously accounts for expected performance and worst-case protection.

4.7. Solution Environment

The optimization model is formulated as a Mixed-Integer Linear Programming (MILP) problem.

- All simulations are performed using:
- Python / MATLAB optimization environment
 - Gurobi or CPLEX solver

Computational experiments are conducted on a standard workstation to evaluate solution quality and computational performance.

5. Results and Discussion

5.1. Cost Performance Comparison

The performance of the proposed hybrid EMS model is evaluated by comparing its operational cost with those of the benchmark models. Comparative cost results are given in Table 7.

Table 7: Comparative cost results of EMS models

Model	Energy Cost (USD/day)	Battery Deg. Cost (USD/day)	Carbon Cost (USD/day)	Penalty Cost (USD/day)	Total Cost (USD/day)
Det. EMS	1038.4	18.6	79.2	26.7	1162.9
Stoc. EMS	1056.8	20.4	76.9	14.8	1168.9
Robust EMS	1104.7	24.3	74.5	4.9	1208.4
Hybrid EMS	1068.5	21.7	75.1	8.6	1173.9

Results show that the deterministic model yields the lowest nominal cost under expected conditions but reforms poorly when uncertainty is realized. In contrast, the robust model ensures system feasibility under extreme conditions but results in significantly higher operational costs due to overly conservative decisions. Cost comparison graph is given in Figure 2.

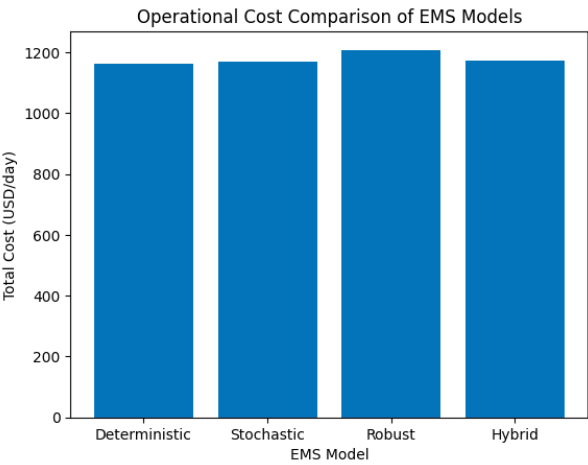


Figure 2: Cost comparison graph

The stochastic model provides a balance between cost efficiency and uncertainty management but lacks protection against extreme scenarios.

The proposed hybrid stochastic-robust model achieves a superior balance by maintaining relatively low expected costs while providing protection against adverse deviations.

5.2. Risk-Cost Trade-off Analysis

The hybrid model introduces a tunable risk parameter λ , which allows decision-makers to control the level of conservativeness in the energy management strategy. Operational performance comparison is given Table 8.

Table 8: Operational performance comparison

Model	Grid Purchase (kWh/day)	PV Utilization Rate (%)	Battery Throughput (kWh/day)	Carbon Emissions (kgCO ₂ /day)
Det. EMS	1885	81.4	620	791.7
Stoc.EMS	1830	84.9	655	768.6
Robust EMS	1775	87.2	710	745.5
Hybrid EMS	1790	86.4	688	751.8

- Sensitivity analysis indicates that:
- Low values of λ prioritize economic efficiency.
 - High values of λ increase system robustness.

As λ increases, operational cost slightly increases, but system resilience improves significantly under worst-case scenarios.

This flexibility enables industrial operators to tailor energy management strategies according to their risk preferences.

5.3. Impact of Renewable Uncertainty

The analysis reveals that renewable generation uncertainty significantly affects energy scheduling decisions. Hourly optimal schedule is given in Table 9.

Table 9: Hourly solution schedule of the proposed hybrid EMS

Hour	Load (kW)	PV (kW)	Grid Purchase (kW)	Battery Charge (kW)	Battery Discharge (kW)
1	210	0	240	30	0
2	205	0	235	30	0
3	200	0	230	30	0
4	198	0	228	30	0
5	202	0	232	30	0
6	220	15	235	30	0
7	260	40	250	30	0
8	310	85	255	30	0
9	355	130	225	0	0
10	380	175	205	0	0
11	395	210	185	0	0
12	410	225	185	0	0
13	420	220	200	0	0
14	415	205	210	0	0
15	405	170	205	0	30
16	398	120	198	0	80
17	390	70	200	0	120
18	360	25	215	0	120
19	335	0	305	0	30
20	300	0	300	0	0
21	275	0	275	0	0
22	250	0	250	0	0
23	230	0	230	0	0
24	220	0	220	0	0

Under high PV variability conditions:

- Battery utilization increases
- Grid reliance becomes more dynamic
- Robust protection becomes more valuable

The hybrid model effectively mitigates these effects by adjusting scheduling strategies based on uncertainty levels. Load-PV-Grid profile is given in Figure 3 and also Battery SOC is given in Figure 4.

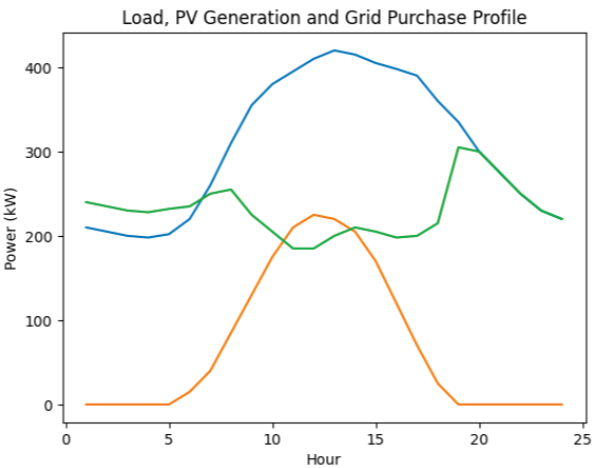


Figure 3: Load-PV-grid

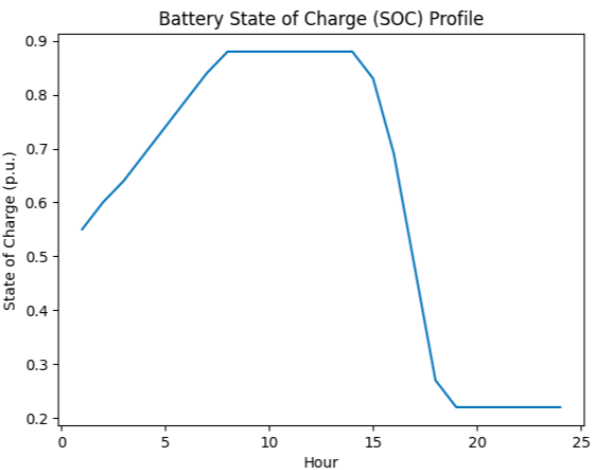


Figure 4: Battery SOC

5.4. Carbon Cost Implications

The integration of carbon pricing into the EMS model influences energy procurement strategies.

Higher carbon prices encourage:

- Greater utilization of PV generation
- Increased reliance on battery storage
- Reduced grid electricity consumption

This demonstrates that carbon pricing mechanisms can significantly influence industrial energy management decisions.

5.5. Managerial Insights

The results provide several important managerial insights for industrial energy operators:

- Hybrid uncertainty modeling provides superior operational flexibility compared to traditional methods.
- Battery storage significantly improves system resilience under renewable uncertainty.
- Carbon pricing policies can effectively promote low-carbon energy strategies.
- Risk-adjustable EMS models allow operators to customize energy management strategies based on operational priorities.

5.6. Discussion

Overall, the proposed hybrid stochastic-robust optimization framework provides a practical and effective solution for industrial energy management under uncertainty.

Compared with existing approaches, the proposed model offers:

- Improved cost stability
- Enhanced resilience to uncertainty
- Flexible risk management capability

These advantages highlight the potential of hybrid optimization approaches in next-generation energy management systems.

6. Sensitivity Analysis

To evaluate the robustness and flexibility of the proposed hybrid energy management model, a sensitivity analysis is conducted on key system parameters. The purpose of this analysis is to investigate how variations in critical parameters influence operational cost, energy scheduling decisions, and system performance.

6.1. Impact of Risk Parameter (λ)

The parameter λ determines the trade-off between expected operational cost and worst-case protection in the hybrid objective function. Sensitivity analysis is shown in figure 5. To analyze its effect, λ is varied across different values while keeping other parameters constant. Sensitivity analysis for risk parameter is given in Table 10.

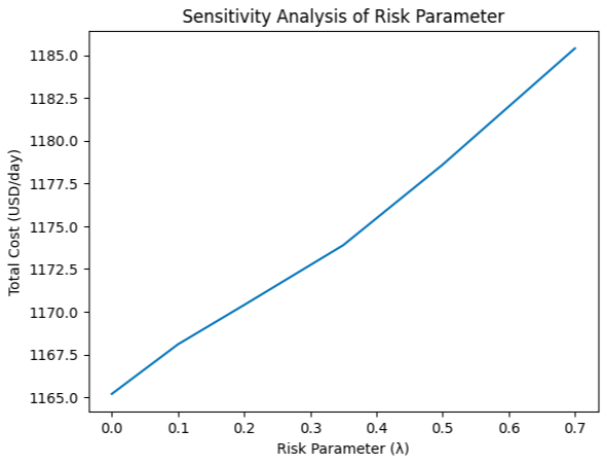


Figure 5: Sensitivity Analysis

Table 10: Sensitivity analysis for risk parameter

λ	Total Cost (USD/day)	Worst-Case (USD/day)	Grid Purchase (kWh/day)	PV Utilization (%)
0.00	1165.2	1298.6	1842	84.7
0.10	1168.1	1275.9	1828	85.1
0.20	1170.4	1259.3	1815	85.8
0.35	1173.9	1236.8	1790	86.4
0.50	1178.6	1220.1	1772	86.9
0.70	1185.4	1208.7	1759	87.3

Results indicate that:

- When λ is small, the model prioritizes expected cost minimization.
- As λ increases, the solution becomes more conservative.

Higher values of λ :

- increase system robustness
- reduce vulnerability to extreme price spikes
- slightly increase operational cost.

This behavior confirms that the hybrid model allows decision-makers to control risk exposure according to operational preferences.

6.2. Impact of PV Generation Variability

The variability of photovoltaic generation significantly affects energy scheduling decisions.

To evaluate this impact, PV uncertainty levels are increased by introducing higher variance in PV generation scenarios. Table 11 presents sensitivity analysis for battery capacity.

Table 11: Sensitivity analysis for battery capacity

Battery Capacity (kWh)	Total Cost (USD/day)	Grid Purchase (kWh/day)	Carbon Emissions (kgCO ₂ /day)	PV Curtailment (%)
200	1214.7	1948	818.2	10.6
400	1190.3	1864	782.9	7.2
600	1173.9	1790	751.8	4.5
800	1168.1	1762	740.0	3.7
1000	1166.8	1754	736.7	3.4

Results show that:

- Battery utilization increases under higher PV uncertainty.
- Grid electricity procurement becomes more dynamic.
- The hybrid model outperforms deterministic and stochastic models in maintaining cost stability.

This demonstrates the advantage of incorporating robust protection against renewable intermittency.

6.3. Impact of Carbon Price

Carbon pricing plays an important role in shaping industrial energy management strategies.

To analyze its influence, carbon price levels are varied across a predefined range. Table 12 presents sensitivity analysis for carbon price.

Table 12: Sensitivity analysis for carbon price

Carbon Price (USD/tCO ₂)	Total Cost (USD/day)	Grid Purchase (kWh/day)	PV Utilization (%)	Carbon Emissions (kgCO ₂ /day)
0	1143.6	1858	83.8	780.4
20	1158.2	1826	84.9	767.0
40	1173.9	1790	86.4	751.8
60	1189.8	1761	87.6	739.6
80	1204.5	1739	88.2	730.4

The results reveal that:

- Higher carbon prices encourage increased use of renewable generation.
- Battery storage utilization increases to reduce grid electricity purchases.
- Total carbon emissions decrease as carbon prices rise.

These findings highlight the potential effectiveness of carbon pricing policies in promoting low-carbon industrial energy systems.

6.4. Impact of Battery Capacity

Battery capacity significantly affects system flexibility.

Increasing battery storage capacity results in:

- reduced reliance on grid electricity
- improved renewable energy utilization
- lower operational cost during peak price periods.

However, beyond a certain threshold, the marginal benefit of additional battery capacity decreases.

This finding suggests that optimal battery sizing is critical for efficient energy management.

7. Conclusion and Future Work

7.1. Conclusion

This study proposed a hybrid stochastic–robust optimization framework for energy management in industrial microgrids with renewable energy integration. The model simultaneously addresses expected system performance and worst-case uncertainty, enabling more reliable and cost-effective operational strategies.

The proposed approach incorporates photovoltaic generation uncertainty, electricity price volatility, and carbon emission costs into a unified optimization framework. A two-stage

decision structure was developed to represent day-ahead scheduling and real-time operational adjustments.

Simulation results from the industrial case study demonstrate that the proposed hybrid model achieves a superior balance between economic efficiency and operational robustness compared with deterministic, stochastic, and purely robust energy management approaches.

Key findings of the study include:

- Hybrid uncertainty modeling significantly improves cost stability under uncertain operating conditions.
- Battery energy storage systems play a critical role in mitigating renewable generation variability.
- Adjustable risk parameters provide flexibility for industrial operators to tailor energy management strategies according to their risk tolerance.
- Carbon pricing mechanisms can effectively promote low-carbon operational decisions.

Overall, the proposed model provides a practical decision-support tool for industrial energy management under uncertainty.

7.2. Future Work

Although the proposed framework provides valuable insights, several extensions can be explored in future research.

First, the model can be extended to incorporate additional distributed energy resources such as combined heat and power (CHP) systems and electric vehicle charging infrastructure.

Second, real-time energy management strategies using machine learning techniques could be integrated with the optimization framework to enhance predictive capabilities.

Third, the proposed model can be expanded to multi-microgrid environments where energy sharing and cooperative energy management are possible.

Finally, future studies may investigate large-scale implementations of hybrid uncertainty models in national energy systems and industrial clusters.

8. References

[1] M’hamed, R., Youcef, H., Bousmaha, B., Mouaadh, Y., “Energy management of photovoltaic plant for smart street lighting system”, *Energy Engineering*, 122(12):4899-4918, (2025).

[2] Ye, W., Huang, S., Herdem, M.S., Sun, W., Nathwani, J., Wen, J., “Transformer-based predictive energy management in hydrogen-integrated renewable systems”, *Energy*, 341:139474, (2025).

[3] Zhang, L., Feng, G., Li, A., Bi, Y., Huang, K., Wang, Z., Chang, S., Liu, Y., “An optimized energy management strategy for wing-PV hybrid heat pump

systems with dual storage: enhancing power-to-load interaction”, *Energy*, 340:139046, (2025).

- [4] Baran, E., “A multi-criteria decision-making approach integrated with machine learning for energy resource supply”, *Systems*, 14(2):200, (2026).
- [5] Chen, G., Wang, H., Zhang, C., Luo, L., Zhang, B., “Research on composite micro-energy systems and adaptive energy management method”, *Energy Conversion and Management: X*, 28:101347, (2025).
- [6] Ikeda, S., Ooka, R., “Metaheuristic optimization methods for a comprehensive operating schedule of battery, thermal energy storage, and heat source in a building energy system”, *Applied Energy*, 151:192-205, (2015).
- [7] Zhang, X., Xu, Y., Zhou, X., Zhang, Y., Li, W., Zuo, Z., Guo, H., Huang, Y., Chen, H., “A near-isothermal expander for isothermal compressed air energy storage system”, *Applied Energy*, 225:955-964, (2018).
- [8] Bermudez, M.F.O., Cabrera, J.D.G., Montoya, O.D., Ramirez, D.A.G., “Energy management system for battery banks in active distribution networks using the arithmetic optimization algorithm”, *Results in Engineering*, 28:108259, (2025).
- [9] Bouaziz, F., Abdelkafi, A., Masmoudi, A., Krichen, L., “Energy management for hybrid building with grid-connected shared DC parking equipment with PV-energy storage system”, *Journal of Building Engineering*, 115:114534, (2025).
- [10] Mazouzi, A., Hadroug, N., Hafaifa, A., Iratni, A., Colak, I., “Particle swarm optimization of fuzzy logic-based energy management system for enhanced efficiency in fuel cell hybrid electric vehicles”, *Sustainable Computing: Informatics and Systems*, 48:101239, (2025).
- [11] Baran, E., “Deep learning-based detection and forecasting of performance losses in solar PV systems using multi-sensor data”, *Applied Sciences*, 16(4):1709, (2026).
- [12] Siluk, J.C.M., Carvalho, P.S., Thomasi, V., Pappis, C.A.O., Schaefer, J.L., “Cloud-based energy management systems: terminologies, concepts and definitions”, *Energy Research & Social Science*, 106:103313, (2023).
- [13] Mezdi, K.E., Magri, A.E., Myasse, I.E., Watil, A., Bahatti, L., Elaadouli, N., Ouabi, H., “Performance improvement through nonlinear control design and power management of a grid-connected wind-battery hybrid energy storage system”, *Results in Engineering*, 20:101491, (2023).
- [14] Modu, B., Abdullah, M.P., Bukar, A.L., Hamza, M.F., Adewolu, M.S., “Energy management and capacity planning of photovoltaic-wind-biomass energy system considering hydrogen-battery storage”, *Journal of Energy Storage*, 73:109294, (2023).
- [15] Quintero, J.F.F., Robles, R.V., Martinez, S.B., Escobar, M.C., Miracle, D.M., “A flexibility management system for behind-the-meter flexibility with distributed energy resources: a sensitivity analysis”, *Sustainable Energy Technologies and Assessments*, 60:103404, (2023).
- [16] Tong, W., Lu, Z., Hunt, J.D., Zhao, H., Han, M., Zhao, G., “Energy management system for modular-gravity energy storage plant”, *Journal of Energy Storage*, 74:109283, (2023).
- [17] Molu, R.J.J., Naoussi, S.R.D., Wira, P., Mbasso, W.F., KEnfack, S.T., Kamel, S., “Optimization-based energy management system for grid-connected photovoltaic/battery microgrids under uncertainty”, *Case Studies in Chemical and Environmental Engineering*, 8:100464, (2023).
- [18] Javid, S.S.M., Derakhshan, G., Hakimi, S.M., “Energy scheduling in a smart energy hub system with hydrogen storage systems and electrical demand management”, *Journal of Building Engineering*, 80:108129, (2023).
- [19] Taye, B.A., Choudhury, N.B.D., “Adaptive filter based method for hybrid energy storage system management in DC microgrid”, *Advances in Electrical Engineering, Electronics and Energy*, 5:100259, (2023).
- [20] Hu, Y., Q., L., Gu, F., Fathi, G., “Comprehensive energy system optimization using developed coyote algorithm for effective management of battery, heat source and thermal storage”, *Energy Reports*, 10:4218-4230, (2023).
- [21] Baran, E., KorkusuzPolat, T., “Classification of industry 4.0 for total quality management: a review”, *Sustainability*, 14(6):3329, (2022).
- [22] KorkusuzPolat, T., Baran, E., “A blockchain-based quality 4.0 application for warehouse management system”, *Applied Sciences*, 14(23):10950, (2024).

Özgeçmişler



Erhan Baran graduated from Kırıkkale University, Department of Industrial Engineering in 2010. He completed his master's degree in Industrial Engineering at Gazi University in 2012. And he completed his doctorate in Industrial Engineering at Sakarya University in 2023. He continues to work as Assistant Professor Dr. in Gazi University, at Electronic and Automation department, TUSAŞ-Kazan Vocational School. He is married. His main research areas are machine learning, deep learning, renewable energy sources, blockchain and artificial intelligence.

Adaptive Signal Source Estimation Using Eigenvalue Analysis and Optimized Dynamic Thresholding

Uyarlamalı Sinyal Kaynağı Tahmini: Özdeğer Analizi ve Optimize Edilmiş Dinamik Eşikleme Yaklaşımı

Gökhan Yüksek¹, Timur Lale²

 0000-0002-6832-8622

 0000-0002-6958-5057

^{1,2}Department of Electrical and Electronics Engineering, Batman University, Batman, Türkiye
gokhan.yuksekk@batman.edu.tr¹, timur.lale@batman.edu.tr²

Öz— Bu çalışma, radar, kablosuz haberleşme ve sensör ağlarında kullanılmak üzere uyarlamalı eşiklemeyle dayalı bir sinyal kaynağı belirleme yöntemi sunmaktadır. Minimum Tanım Uzunluğu (MDL), Akaike Bilgi Kriteri (AIC) ve yeni bir özdeğer tabanlı uyarlamalı eşikleme yaklaşımı olmak üzere üç yöntem; gürültülü, kısa menzilli ve birbirine yakın kaynakların bulunduğu koşullar için tasarlanan ultra geniş bant (UWB) sensör dizileri üzerinde gerçekleştirilen simülasyonlarla karşılaştırılmıştır. MDL ve AIC, model karmaşıklığı ile veri uyumunu dengelemek amacıyla kovaryans matrisi özdeğerlerini kullanırken, önerilen uyarlamalı eşikleme yöntemi özdeğer dağılımına göre gürültü eşiğini dinamik olarak ayarlayarak yüksek gürültü koşullarında kaynak sayısı tahmin performansını artırmaktadır. Sonuçlar, tüm yöntemlerin düşük gürültü koşullarında iyi performans gösterdiğini, ancak önerilen yöntemin yüksek gürültü seviyelerinde veya kaynakların birbirine yakın olduğu durumlarda daha yüksek başarı sağladığını göstermektedir. Ayrıca eşik fonksiyonundaki duyarlılık parametreleri hafif hesaplama yüküne sahip bir arama yöntemiyle optimize edilerek farklı gürültü seviyeleri ve kaynak konfigürasyonları altında kararlı performans elde edilmektedir.

Anahtar Kelimeler— Uyarlamalı Eşikleme, Kaynak Tahmini, Özdeğer Analizi, Optimizasyon

Abstract— This study presents a method for identifying signal sources using adaptive thresholding, crucial in radar, wireless communication, and sensor networks. Three approaches, Minimum Description Length (MDL), Akaike Information Criterion (AIC), and a novel eigenvalue-based adaptive thresholding are compared through simulations on ultra-wideband (UWB) sensor arrays designed for noisy, low-range conditions with closely spaced sources. MDL and AIC employ covariance matrix eigenvalues to balance model complexity and data fit, while the proposed adaptive thresholding dynamically adjusts the noise floor according to the eigenvalue distribution, improving source-enumeration performance under high-noise conditions. Results show that all methods perform well under low-noise conditions, but the proposed method excels when noise is high or sources are closely located. Additionally, sensitivity parameters in the threshold function are

optimized via a lightweight search, providing stable performance across varying noise and source configurations.

Key Words— Adaptive Thresholding, Source Estimation, Eigenvalue Analysis, Optimization

I. Introduction

Estimating the number of active signal sources remains a major challenge in signal processing, which underpins applications in sensor networks, radar systems, wireless communication, and medical imaging. Accurate source estimation improves network efficiency and channel allocation in wireless systems, enables reliable multi-target tracking in radar applications, and supports fault detection and environmental monitoring in sensor arrays. [1], [2], [3], [4], [5], [6], [7].

Eigenvector-based techniques are widely applied to separate signal and noise components by analysing the eigenvalues of the covariance matrix, which represent energy distribution across dimensions. [8] The present work employs this property to estimate the number of sources in real time without explicitly constructing subspaces. Compared with conventional time-domain approaches, eigenvalue-based methods provide source-number information directly from the covariance eigenvalue spectrum without requiring prior knowledge of the exact source waveforms or explicit construction of an array manifold [9], [10].

Several notable studies have contributed to this field. Hyvärinen and Oja advanced independent component analysis algorithms [11]; Huang et al. introduced an MMSE-based MDL estimator [12]. Wiesel and Hero explored decomposable principal component analysis [13];

Ouedraogo et al. developed a non-negative blind source separation approach [14]; and Han and Nehorai enhanced source estimation using nested and linear arrays [15].

This study compares three representative techniques: Minimum Description Length (MDL), Akaike Information Criterion (AIC), and a proposed eigenvalue-driven adaptive thresholding scheme. MDL [16] and AIC [17], [18] are established model-selection criteria balancing model complexity and data fit through eigenvalues. While MDL favours simplicity to prevent overfitting, AIC emphasises prediction accuracy. In contrast, adaptive thresholding dynamically adjusts the noise floor from the eigenvalue distribution, maintaining effectiveness under varying noise and source configurations.

The paper makes several contributions. It offers a systematic comparison of MDL, AIC, and adaptive thresholding under different noise environments, demonstrating the latter's advantage when noise is high or sources are closely spaced. The study also integrates a lightweight optimization framework for threshold parameters (α, β), reducing heuristic dependence and ensuring consistent performance. A brief computational analysis confirms that the proposed approach shares the same cost order as MDL and AIC but avoids extra overheads, making it suitable for practical ultra-wideband (UWB) sensor implementations.

The main contribution of this study is an optimized adaptive source enumeration method that combines spectral curvature information, eigenvalue-based noise-floor estimation, and data-driven threshold parameter selection within a single procedure. Unlike MDL and AIC, which determine the source number through penalized model-order selection, the proposed method directly evaluates the structure of the ordered covariance eigenvalues. First, the second discrete difference of the eigenvalue sequence is used to identify sharp spectral transitions between signal- and noise-dominated components. This information is incorporated into a weighted adaptive threshold rather than treating all eigenvalues equally. Second, an MP-inspired spread term is included to account for the finite-sample spread of noise eigenvalues as a function of the sensor-to-sample ratio. Third, the threshold sensitivity parameters are selected through a lightweight coarse-search and local-refinement procedure and are subsequently fixed for the reported simulations. Thus, the contribution lies in the combined use of discrete spectral-curvature weighting, finite-sample noise correction, and optimized threshold selection for source enumeration, rather than in eigenvalue

thresholding alone.

The remainder of the paper is organized as follows. Section II introduces MDL and AIC and presents the proposed adaptive thresholding method, including the weighting strategy, threshold formulation, parameter optimization, and computational complexity. Section III describes the implementation and simulation setup. Section IV presents the comparative results under different SNR levels and source separations and discusses the main limitations of the method. Finally, Section V summarizes the main findings and conclusions.

II. Methods

A. Minimum Description Length

The MDL [16], [18] criterion is a model selection approach based on information theory, aiming to minimize the total description length of both the model and the observed data. The number of active sources in sensor array systems can be estimated using MDL for signal processing. This is done by analysing the eigenvalues of the sample covariance matrix. The main principle behind MDL is to balance model accuracy with complexity, preventing overfitting by advancing simpler models. Given a covariance matrix R derived from received signals at an array of N sensors, its eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ represent the variance of the signal components along different dimensions. The MDL criterion for estimating the number of sources k is formulated as:

$$MDL(k) = -T \sum_{i=k+1}^N \log \left(\lambda_i^{\frac{1}{N-k}} \right) + k \frac{(2N - k + 1)}{2} \log(T), \quad (1)$$

where T is the number of time samples. The first term in this expression represents the likelihood of the data fitting the assumed signal model, while the second term penalizes complex models to prevent overestimation. The optimal number of sources k is determined by selecting the value that minimizes $MDL(k)$. This method is effective in distinguishing between the signal and noise subspaces, as it identifies the significant eigenvalues corresponding to the signal sources.

Compared with AIC, MDL applies a stronger penalty to model complexity, which generally reduces the tendency toward model-order overestimation. But this stronger penalty

does not imply better performance under high-noise conditions. At low SNR levels, weak signal eigenvalues may approach the noise eigenvalue cluster, increasing the tendency of MDL to underestimate the number of active sources. So, its performance depends on the separability of the signal- and noise-related eigenvalues as well as the applied complexity penalty [16], [17], [18], [19], [20], [21].

B. Akaike Information Criterion

The AIC [16], [17], [18] is a model selection method that seeks to strike a balance between model intricacy and data congruence. Statisticians generally use it in the context of statistical signal processing to estimate the number of active sources for sensor arrays. They do this by analysing the eigenvalues of the sample covariance matrix. AIC is different from the MDL criterion in that it is designed to minimise the relative information loss by selecting the model with the lowest penalised likelihood score. This is in contrast to the MDL criterion, which imposes a stronger penalty on model complexity. Given a covariance matrix R derived from the received signals at an array of N sensors, its eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ represent the variance of the signal components along different dimensions. The AIC criterion for estimating the number of sources k is as follows:

$$AIC(k) = -2T \sum_{i=k+1}^N \log\left(\lambda_i^{\frac{1}{N-k}}\right) + 2k(2N - k + 1), \quad (2)$$

where T is the number of time samples. The first term in this equation corresponds to the likelihood function, measuring how well the model fits the data, while the second term serves as a penalty to prevent overfitting by discouraging excessive model complexity. The optimal number of sources k is determined by selecting the value that minimizes $AIC(k)$. AIC uses a weaker penalty than MDL, so it often chooses more complex models. This makes it useful when underestimating the number of sources is a bigger problem than overestimating. However, in high-noise conditions, AIC may overestimate the number of sources due to its lower penalization of complexity. [16], [17], [18], [19], [20], [21].

C. Adaptive Thresholding

The adaptive thresholding method dynamically determines the number of active signal sources by analysing the eigenvalue spectrum. Unlike traditional model selection criteria such as MDL and AIC, which rely on predefined penalty functions, adaptive thresholding adjusts itself based

on the statistical distribution of eigenvalues, allowing it to perform well across different noise levels and signal conditions. However, conventional adaptive thresholding methods often consider all eigenvalues equally, potentially leading to incorrectness when sources have varying power levels. To address this, a weighted adaptive thresholding approach is introduced, enhancing sensitivity to dominant sources while suppressing noise components.

In an array signal model, a linear array that formed by N sensors receives signals from k active sources. The received signal matrix X can be expressed as;

$$X = AS + W, \quad (3)$$

where A is the mixing matrix including the steering vectors of the sources, S represents the transmitted source signals over T time samples, and W is an additive noise matrix, typically modeled as zero-mean Gaussian noise. From this signal model, the covariance matrix R is computed as;

$$R = \frac{1}{T} XX^H, \quad (4)$$

where X^H is the Hermitian transpose of X . The eigenvalues of R , denoted as $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$, provide insight into the energy distribution of the signal and noise components. The larger eigenvalues correspond to active sources, while the smaller eigenvalues represent noise. The signal model considered in this study assumes additive spatially white Gaussian noise and non-coherent source signals. Under these conditions, the noise covariance is proportional to the identity matrix, while statistically independent source components contribute distinct signal-related eigenvalues to the covariance spectrum. These assumptions provide the basis for separating signal- and noise-dominated eigenvalues. Therefore, the proposed method does not assume an arbitrary signal environment; rather, it avoids requiring prior knowledge of the exact source waveforms or an explicitly constructed array manifold.

A key challenge in adaptive thresholding is the dynamic determination of a threshold τ that effectively distinguishes signal-dominant and noise-dominant eigenvalues. Instead of relying on a fixed threshold, this method estimates the noise floor using statistical properties of eigenvalues. A common approach is to define τ as a function of the mean or median of the eigenvalues;

$$\tau = \alpha \cdot \text{median}(\lambda_i) \text{ or } \tau = \alpha \cdot \text{mean}(\lambda_i), \quad (5)$$

where α is a scaling factor that adjusts the threshold sensitivity. While effective in moderate noise environments, this formulation does not account for variations in eigenvalue importance. To increase the truthfulness of estimation, a new weighted thresholding formulation incorporating eigenvalue characteristics is proposed;

$$\tau = \alpha \cdot \sum_{i=1}^N \omega_i \lambda_i + \beta \cdot \sigma_n^2 (1 + \sqrt{C}). \quad (6)$$

Where ω_i denotes the weight assigned to each eigenvalue according to the local spectral variation described below, while α controls the sensitivity of the eigenvalue-based threshold component. The proposed weighting approach assigns greater importance to regions where the ordered eigenvalue spectrum exhibits sharp local variations. Since the eigenvalue index is discrete, this variation is characterized by the second discrete difference;

$$\Delta^2 \lambda_i = \lambda_{\{i+1\}} - 2\lambda_i + \lambda_{\{i-1\}} \quad \left. \vphantom{\Delta^2 \lambda_i} \right\} \quad i = 2, \dots, M-1. \quad (7)$$

A larger magnitude of $(\Delta^2 \lambda_i)$ represents a stronger local change in the spectral slope and therefore indicates a potential transition between signal- and noise-dominated eigenvalues.

The weighting strategy is based on the local structure of the ordered eigenvalue spectrum. A pronounced change in the second discrete difference indicates a stronger variation in spectral curvature and therefore provides information about a possible transition between signal- and noise-dominated components. This transition information is incorporated into the adaptive threshold so that eigenvalues are not treated equally. In this formulation, regions with stronger local curvature receive greater emphasis, allowing the threshold to respond to the spectral separation observed in the covariance eigenvalues. Since the calculation uses only the ordered eigenvalues already obtained from the covariance matrix, no additional eigenvector-based processing is required.

The subspace separation in the proposed adaptive thresholding method is performed by analysing the ordered eigenvalue spectrum rather than explicitly using the corresponding eigenvectors. The second discrete difference provides local spectral-curvature information, and eigenvalues around stronger transition regions are assigned

greater importance in the adaptive threshold. Thus, the weighting strategy incorporates the local structure of the eigenvalue spectrum instead of treating all eigenvalues equally. The procedure operates on the eigenvalues already obtained from the covariance matrix and does not require additional eigenvector-based processing.

The threshold formulation includes an MP-inspired correction term [10], [22], to account for the finite-sample spread of noise-related eigenvalues. In random matrix theory, this spread depends on the ratio between the number of sensors and the number of temporal observations. Accordingly, the correction term incorporates the ratio $C = N/T$, together with the estimated noise variance σ_n^2 . In the proposed formulation, β controls the contribution of this noise-dependent correction to the adaptive threshold.

$$C = N/T, \quad (8)$$

the ratio of the number of sensors (N) to the number of time samples (T). The estimated number of sources is then determined by;

$$n = \sum_{i=1}^N \mathbb{I}(\lambda_i > \tau). \quad (9)$$

Here, $\mathbb{I}(\cdot)$ is an indicator function that assigns 1 if the condition is met and 0 in other cases. This function forms the output that only eigenvalues significantly above the noise floor are classified as signal components.

Adaptive thresholding is an effective solution for source estimation in array signal processing. This is achieved by leveraging dynamic noise floor estimation, weighted eigenvalue distributions and statistical noise modelling. Incorporating noise-adaptive scaling factors and weighted thresholding enhances its capacity to function in real-world applications such as radar, ultra-wideband (UWB) communication and multi-sensor networks, where signal characteristics fluctuate dynamically. Furthermore, its utility is demonstrated in civil defence scenarios, for example in identifying the number of heartbeats beneath a collapsed structure.

D. Optimization of Threshold Parameters α and β

The sensitivity parameters in α and β in (10) and (11) directly control the threshold's position relative to the eigenvalue spectrum and play a decisive role in

distinguishing signal and noise components. Instead of assigning fixed values, these parameters are optimized as an integral part of the proposed method.

To evaluate each candidate pair (α, β) , a success rate metric is used. For a given scenario d , the success rate $\eta(\alpha, \beta | d)$ is defined as

$$\eta(\alpha, \beta | d) = \frac{\text{Number of correct source est.}}{\text{Total number of trials}}. \quad (10)$$

The success rate $\eta(\alpha, \beta | d)$ corresponds to the fraction of trials in which the estimated number of sources matches the ground truth. Maximizing its average across the validation set $\mathcal{D}$ defines the objective function of the optimization.

The optimization objective is then expressed as

$$J(\alpha, \beta) = \frac{1}{|\mathcal{D}|} \sum_{d \in \mathcal{D}} \eta(\alpha, \beta | d), \quad (11)$$

where $\mathcal{D}$ denotes a validation set including representative conditions such as SNR levels between -10 and $+5$ dB, sensor array sizes of 7 and 9 elements, and minimum source spacings between 0.05 m and 0.20 m. The optimal parameters are determined by

$$(\alpha^*, \beta^*) = \arg \max_{\alpha, \beta} J(\alpha, \beta). \quad (12)$$

In practice, the optimization follows a two-step procedure:

- Coarse grid search [23]: A discrete grid of candidate values is generated for $\alpha \in [0.8, 2.0]$ and $\beta \in [0, 0.3]$. Each pair is evaluated by computing $J(\alpha, \beta)$ over the validation set $\mathcal{D}$.
- Local refinement: The best-performing candidates are further refined using a Nelder–Mead [24] simplex search, which efficiently locates the optimum without requiring derivative information.

This lightweight procedure converges quickly and introduces negligible additional cost, since the dominant computation remains the eigenvalue decomposition of the covariance matrix. The optimized pair (α^*, β^*) is subsequently fixed and applied consistently in all reported simulations.

Incorporating this parameter optimisation makes the adaptive thresholding method statistically adaptive through eigenvalue distribution analysis, as well as parametrically tuned. This ensures stable performance across different SNR and geometry scenarios.

E. Computational Complexity

The dominant cost in all methods is the eigenvalue decomposition of the covariance matrix, which scales as $O(N^3)$ for an array of N sensors. MDL and AIC additionally require scoring candidate orders, resulting in $O(N^3 + KN)$, where K is the maximum number of sources considered. The proposed adaptive thresholding performs only a threshold comparison after the decomposition, maintaining $O(N^3)$ complexity. The lightweight optimization of α and β (Sec. 2.D) involves a coarse grid and a small number of refinement steps; since each step reuses the same decomposition, the additional cost remains negligible compared to the baseline operations.

III. Implementation

A. Pseudo-code

The steps of the proposed method are outlined in the pseudocode below;

Pseudocode:

i. Input: Sensor data matrix X , number of sensors M , number of temporal samples T , and noise variance σ_n^2 .

ii. Compute: Sample covariance matrix:

$$R = (1/T) XX^H$$

iii. Compute: Eigenvalues of R and sort them in descending order:

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$$

iv. Compute: Second discrete differences of the ordered eigenvalues:

$$\Delta^2 \lambda_i = \lambda_{i+1} - 2\lambda_i + \lambda_{i-1}, \quad i = 2, \dots, M-1$$

v. Identify: Dominant local variations in the second discrete-difference sequence and determine the corresponding eigenvalue weights w_i .

vi. Compute: Adaptive threshold τ using the weighted eigenvalue information, optimized sensitivity parameters α and β , and the MP-inspired correction term.

vii. Count: Eigenvalues above the adaptive threshold:

$$\lambda_i > \tau$$

viii. **Output:** Estimated number of active signal sources:
 $\hat{K} = \sum_{i=1}^M I(\lambda_i > \tau)$

B. Simulation

The present simulation example is concerned with the estimation of the number of active signal sources in a linear sensor array configuration by means of a technique known as eigen-value distribution analysis. The example setup is shown in Fig. 1. The setup consists of an array that has been formed by 9 sensors which have been arranged linearly along the x-axis, with a spacing of 0.12 metres to optimise resolution.

Each sensor in the linear array receives signals from 7 sources located at varying distances (ranging from 1.7 m to 4 m) and angles (ranging from -15° to 15°). The signals received at array experience time delays depending on the distances they travel, so interpolation is used to align the signals at each sensor. The sensor outputs are then used to compute the sample covariance matrix, from which eigenvalue decomposition is performed.

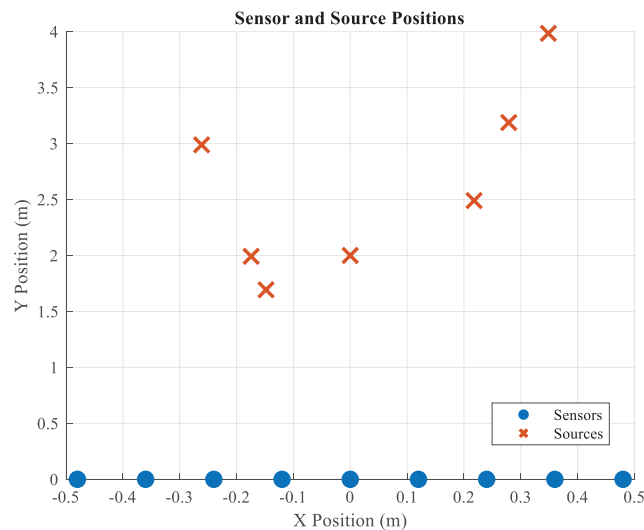


Fig. 1. Distribution of the linear sensor array and signal sources.

The system works at a centre frequency of 2.4 GHz, with a sampling rate of 24 GHz to make sure that time is measured very accurately. The SNR calculation is used to evaluate performance. This is done by simulating realistic signal conditions using Gaussian noise with a proper variance. Broadband characteristics suitable for UWB applications are ensured by the transmitted signals being high-order Morlet

pulses.

The simulations are conducted under a temporally oversampled condition, where the number of temporal observations (T) is greater than the number of array sensors (M), i.e., ($T > M$). Accordingly, the received data matrix contains sufficient temporal observations for estimating the ($M \times M$) sample covariance matrix. This condition reduces the risk of rank deficiency associated with an insufficient number of observations and provides a more suitable eigenvalue spectrum for signal-noise separation. It is also consistent with the ratio ($c = M/T < 1$) used in the MP-inspired correction term. Therefore, the reported performance corresponds to the temporally oversampled operating condition considered in this study.

Fig. 2 presents the Morlet wavelet signal and its corresponding Fourier spectrum at a centre frequency of approximately 2.5 GHz. The top plot illustrates the Morlet wavelet in the time domain, locally showing a sinusoidal oscillation. The bottom plot displays its Fourier spectrum, which exhibits a sharp peak around 2.4 GHz, confirming the dominant frequency of the wavelet. The frequency spectrum also shows minor side lobes, which result from the finite time window of the signal. The presence of low-level noise in both the time and frequency domains suggests realistic signal conditions, which may impact signal resolution but remain within acceptable limits for the UWB applications.

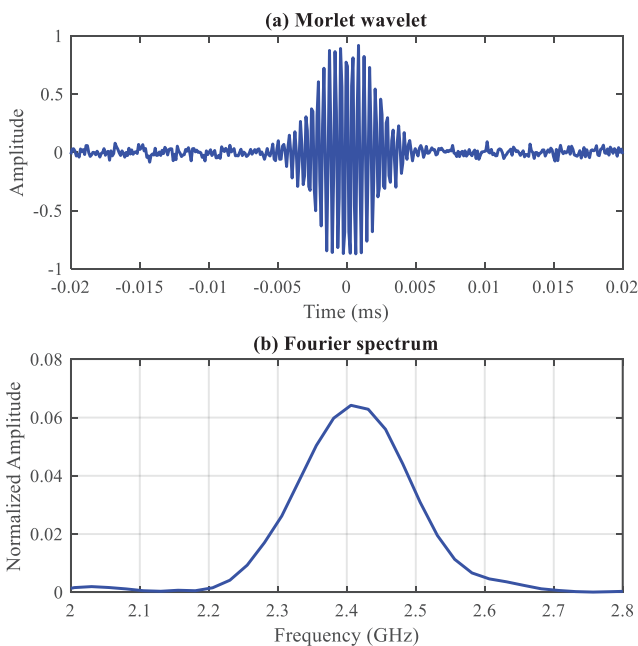


Fig. 2. Ultra-Wideband signal used in implementation.

This signal type and system geometry is well-organized for electromagnetic near-field conditions [25], [26], because in this region the source distance is comparable to or smaller than the wavelength, causing the wavefront to take a spherical form rather than planar propagation. Conventional methods, which assume planar wavefronts, may introduce errors in source estimation under such circumstances. In contrast, the proposed adaptive thresholding method does not depend on explicit array manifold models or steering vectors; instead, it relies solely on the statistical separation of signal- and noise-dominated eigenvalues. This separation continues to hold in near-field propagation, as the covariance matrix still reflects the number of independent sources regardless of the wavefront shape. Moreover, the integration of parameter optimization and MP-based correction stabilizes the threshold even when eigenvalue spreads are affected by near-field curvature or broadband effects. As a result, the method remains effective in indoor radar and short-range communication systems, where high-resolution source detection is required despite the complex wave interactions inherent to near-field geometries.

For each time step, the sorted eigenvalues of the covariance matrix are analysed. MDL and AIC compute likelihood-based penalties, while the adaptive threshold dynamically identifies significant eigenvalues above the noise floor. The estimated number of sources is derived by selecting the minimizing index for MDL and AIC, while adaptive thresholding counts eigenvalues exceeding the computed noise threshold [27], [28].

For each scenario in the validation set $\mathcal{D}$, the optimized parameter pair (α^*, β^*) obtained in Sec. 2.D was applied, ensuring consistent parameter settings across different SNR levels and array geometries.

IV. Results and Discussion

A. Source Enumeration under Noise Variation

The performance of MDL, AIC, and proposed adaptive thresholding is evaluated under different SNR levels and source configurations. Each method is tested multiple times in a loop to determine its ability to accurately estimate the number of sources in different randomized noise environments. The results show how noise affects their performance and highlight the advantages of adaptive thresholding in challenging conditions. Fig. 3 shows the results for noise environments by their SNR in the following.

Formulation of the success rate is determined in (10). In

Fig. 3 one can see the performance metrics of all three algorithms. Measurements are taken by 1000 independent runs for each SNR level between -15dB and 15dB. The result differences on each run are caused by the randomness of the noise figure. Algorithm's truthfulness is one of the key expectations of the evaluation.

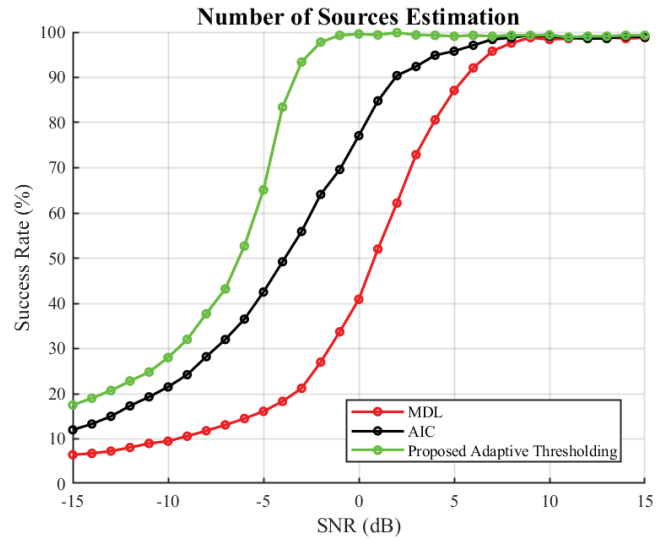


Fig. 3. performance evaluation of the algorithms (for 1000 runs)

The performance evaluation of the number of source estimation methods whose results are given in Fig. 3 which demonstrates the advantages of adaptive thresholding versus the traditional approaches. In low-SNR conditions, MDL and AIC fail to distinguish signal components from noise that results with low success rates.

The proposed adaptive thresholding method outperforms both, maintaining relatively higher accuracy by dynamically adjusting the noise floor. However, this improvement does not necessarily imply acceptable overall performance. As the SNR increases, AIC shows improvement, but adaptive thresholding reaches a better success rate faster. This indicates that the adaptive method is more effective in handling varying noise levels and closely spaced sources.

Table I. presents the summary of the achievements and comments for categorized SNR levels. At high SNR values, all methods eventually achieve nearly 100% accuracy, but the adaptive approach converges more quickly. MDL remains the slowest, often underestimating the number of sources in moderate-SNR conditions. AIC performs better than MDL but still lags behind the adaptive method in complex scenarios.

TABLE I
AVERAGE SUCCESS EVALUATION FOR CATEGORIZED SNR LEVELS

	SNR ranges					
	(-15, -10) dB	(-10, -5) dB	(-5, 0) dB	(0, 5) dB	(5, 10) dB	(10, 15) dB
MDL Avg (%)	7.5	11.9	23.2	61.7	94.2	98.6
AIC Avg (%)	15.3	28.4	56.2	87.9	97.8	98.8
Proposed Avg (%)	20.9	38.7	87.8	99.1	99.2	99.5
Observation	Success rates are low. MDL is the lowest, the proposed method performs better.	Success rates are increasing. MDL remains low, while AIC and the proposed method perform better.	Success rates rise rapidly. The proposed method shows the best performance.	High success levels. The proposed method is nearly perfect	All methods achieve very high success rates.	All methods show near-optimal performance in this range.

One advantage of the proposed method is that it does not directly depend on the geometric distribution of sources. Traditional subspace-based methods often require well-separated sources for accurate estimation, but adaptive thresholding can detect sources even in complex or irregular geometries. This makes it useful in environments where sources are randomly placed or closely spaced. Another strength is its ability to resolve neighbouring sources more effectively. Since the method adapts dynamically to the eigenvalue distribution, it can separate weak and strong sources without assuming a predefined structure.

B. Close-Source Discrimination Performance

In Fig. 4, an additional simulation experiment is conducted to comparatively evaluate the ability of the algorithms to resolve closely spaced active sources. Two separate sources are added to the same experimental scene, positioned at an average distance of 2 meters from the array centre. The distance between these two closely spaced sources varies between 0.05 and 0.25 meters. Each distance configuration is simulated 1000 times, and the following results are obtained regarding the success rates.

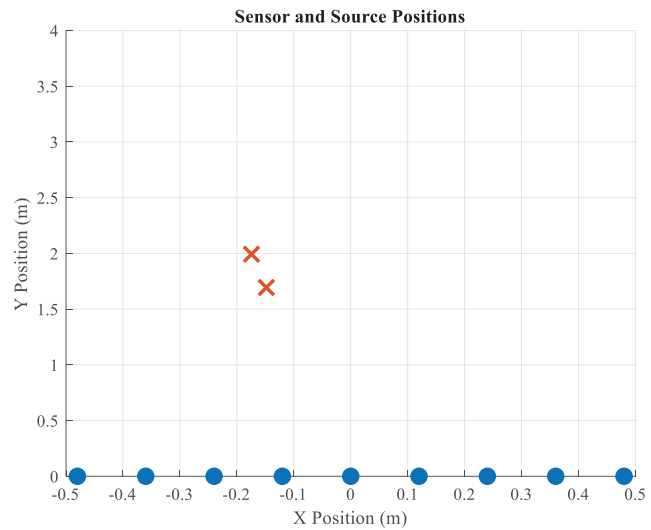


Fig. 4. Simulation scenario for detection of close sources.

Table II presents the success rates of the three methods in detecting closely spaced sources. The adaptive thresholding method significantly outperforms MDL and AIC, especially when the distance between sources is small. At 0.05 meters separation, the adaptive method achieves nearly 60% success, while MDL and AIC struggle below 30%. As the source distance increases, all methods improve, but the adaptive thresholding still provides the highest success rate. At 0.25 meters, all methods approach good accuracy, but the adaptive method converges to this level much faster.

When the distance between sources was very small (0.05–0.10 m), the optimized parameters (α^*, β^*) provided a clearer separation between signal and noise eigenvalues. These results include the optimized parameter setting described in Sec. 2.D, which ensures stable performance even in close-source scenarios.

The close-source results show that the performance differences between the methods are most evident at small source separations. At 0.05 m, the proposed method achieves a success rate of 58.9%, compared with 28.4% for AIC and 19.5% for MDL. At 0.10 m, the corresponding success rates increase to 75.3%, 43.2%, and 32.1%, respectively. These results indicate that the eigenvalue-based adaptive threshold provides a clearer distinction between signal- and noise-related components under the tested close-source conditions. The optimized parameter setting described in Section II-D is used throughout these experiments and contributes to the observed performance across the evaluated source spacings.

TABLE II
 SUCCESS RATES FOR CLOSE SOURCE DISCRIMINATION

Distance Between Sources (m)	MDL Success Rate (%)	AIC Success Rate (%)	Adaptive Thresholding Success Rate (%)
0,05	19,5	28,4	58,9
0,10	32,1	43,2	75,3
0,15	50,8	71,7	84,2
0,20	86,4	88,9	97,6
0,25	96,3	98,2	99,3

As the source spacing increases, the performance difference between the methods gradually decreases. At 0.20 m, the proposed method reaches 97.6% success, while MDL and AIC reach 86.4% and 88.9%, respectively; at 0.25 m, all three methods exceed 96%. The proposed method determines the threshold from the ordered eigenvalue spectrum and uses the second discrete difference to characterize local spectral transitions. Since these operations are performed directly on the eigenvalues obtained from the covariance matrix, the procedure does not require additional eigenvector-based processing. This keeps the source-enumeration procedure computationally simple while preserving the performance advantage observed under the more challenging source-spacing conditions.

C. Statistical Confidence Analysis

To quantify the statistical uncertainty of the reported success rates, a confidence analysis was performed using the 1000 independent trials conducted for each source-spacing condition. Since each trial results in either correct or incorrect source-number estimation, the success rate is treated as a binomial proportion. For a given method and test condition, the estimated success probability is,

$$\hat{p} = \frac{N_{\text{correct}}}{N_{\text{runs}}} , \tag{13}$$

where (N_{correct}) denotes the number of correct source-number estimates and ($N_{\text{runs}} = 1000$). The corresponding standard error is

$$SE(\hat{p}) = \sqrt{\frac{\hat{p}(1 - \hat{p})}{N_{\text{runs}}}} . \tag{14}$$

Based on this formulation, the approximate 95% confidence interval is calculated as,

$$CI_{95\%} = \hat{p} \pm 1.96SE(\hat{p}) . \tag{15}$$

Table III presents the resulting confidence intervals for the close-source experiment. The statistical analysis supports the differences observed in Table II. At the most challenging spacing of 0.05 m, the proposed method achieves a success rate of 58.9% with a 95% confidence interval of 55.9–61.9%, whereas AIC and MDL yield intervals of 25.6–31.2% and 17.0–22.0%, respectively. The confidence intervals of the proposed method and both conventional methods remain non-overlapping between 0.05 m and 0.20 m. This indicates that the observed performance differences within this range are substantially larger than the sampling uncertainty associated with the 1000 independent trials. At 0.25 m, the intervals become closer as all methods reach high source-enumeration success rates.

TABLE III
 95% CONFIDENCE INTERVALS FOR CLOSE-SOURCE ENUMERATION SUCCESS RATES

Source Spacing (m)	MDL 95% CI (%)	AIC 95% CI (%)	Proposed 95% CI (%)
0.05	17.0–22.0	25.6–31.2	55.9–61.9
0.10	29.2–35.0	40.1–46.3	72.6–78.0
0.15	47.7–53.9	68.9–74.5	81.9–86.5
0.20	84.3–88.5	87.0–90.8	96.7–98.5
0.25	95.1–97.5	97.4–99.0	98.8–99.8

To quantify the statistical uncertainty of the reported success rates, 95% confidence intervals were calculated using the binomial proportion model, with ($N=1000$) independent trials for each source-spacing condition. The resulting intervals are presented in Table III. The separation between the methods is especially clear at short source distances. At 0.05 m, the proposed method achieves a 95% confidence interval of 55.9–61.9%, compared with 25.6–31.2% for AIC and 17.0–22.0% for MDL. The confidence intervals remain non-overlapping up to 0.20 m between the proposed method and both conventional methods, indicating that the observed advantage is substantially larger than the Monte Carlo sampling uncertainty. At 0.25 m, the intervals become closer as all three methods approach their high-SNR/high-separation performance region.

D. Practical Considerations and Limitations

The ability to resolve closely spaced sources is influenced not only by the estimation method but also by various physical and signal processing parameters. Line-of-Sight (LOS) conditions play a crucial role, as multipath interference and non-line-of-sight (NLOS) effects can alter the eigenvalue structure, affecting source separation. Operating frequency and bandwidth also impact resolution, where higher frequencies and wider bandwidths generally improve spatial resolution but may introduce additional noise sensitivity.

The present evaluation assumes spatially white Gaussian noise and non-coherent sources. Under this assumption, the noise covariance satisfies $R_n = \sigma_n^2 I$, which supports the formation of a compact noise-eigenvalue region used by the proposed threshold. For colored noise, where $R_n \neq \sigma_n^2 I$, the noise eigenvalues may exhibit a wider and non-uniform distribution, and the present MP-inspired correction may therefore require modification. Similarly, strong source correlation or coherence can reduce the effective rank of the signal covariance and decrease the separation between signal- and noise-related eigenvalues. These cases are outside the present simulation scope and constitute important directions for further evaluation, including covariance prewhitening or spatial-smoothing strategies.

Despite the many advantages of adaptive thresholding, there are challenges when implementing it in real-world scenarios. The approach depends on assumptions about the behaviour of the eigenvalues, which may require modifications for different contexts. While adaptive

thresholding is generally effective, sudden changes in noise can affect its accuracy. Subsequent studies could improve upon this methodology by refining the threshold calculation to accommodate a broader range of scenarios.

The proposed method's ability to function in both near-field and wideband scenarios is another of its key strengths. The eigenvalue-based thresholding strategy is different from traditional space-based approaches, which rely on planar wavefront assumptions. It does not require explicit manifold modelling. Even in situations involving spherical wavefronts, such as in short-range UWB conditions, the method maintains its capacity to effectively differentiate between signal and noise eigenvalues. The broadening of the applicability of the approach to indoor radar and short-range communication systems, where near-field effects are dominant, is a feature of this.

V. Conclusion

This study presents an optimized adaptive thresholding method for estimating the number of active signal sources and compares its performance with MDL and AIC. The results show that the proposed method achieves higher source-enumeration success rates under low-SNR conditions by dynamically adjusting the noise threshold. It also reaches high success rates at lower SNR levels and performs consistently across the tested noise conditions. The method requires only eigenvalue-based operations after covariance estimation, which keeps the computational structure simple. In addition, the threshold parameters are optimized before testing and then fixed for the reported simulations, allowing the method to adapt to different SNR and source-spacing conditions without adding substantial computational cost. Overall, the proposed method provides an effective alternative to MDL and AIC for source enumeration in radar, wireless communication, and sensor-array applications.

References

- [1] D. G. . Manolakis, V. K. . Ingle, and S. M. . Kogon, "Statistical and adaptive signal processing: spectral estimation, signal modeling, adaptive filtering, and array processing," p. 796, 2005.
- [2] D. H. D. H. Johnson and D. E. Dudgeon, "Array signal processing: concepts and techniques," *Book*, p. 533, 1993, Accessed: Mar. 16, 2025. [Online]. Available: <https://search.worldcat.org/title/25546138>
- [3] J. Bin, L. An'Nan, and X. Jie, "An improved signal number estimation method based on information theoretic criteria in array processing," *2019 IEEE 11th International*

- Conference on Communication Software and Networks, ICCSN 2019, pp. 193–197, Jun. 2019, doi: 10.1109/ICCSN.2019.8905347.
- [4] Jacob. Benesty, Israel. Cohen, and J. . Chen, “Fundamentals of signal enhancement and array signal processing,” 2019.
- [5] J. M. Kahn and J. R. Barry, “Wireless infrared communications,” *Proceedings of the IEEE*, vol. 85, no. 2, pp. 265–298, 1997, doi: 10.1109/5.554222.
- [6] R. F. Cinar, A. Demirkol, and Z. Demir, “An ambiguity function aided target density function for radar imaging,” *Optoelectronics and Advanced Materials, Rapid Communications*, vol. 9, no. 5–6, pp. 850–854, 2015.
- [7] R. F. Cinar and A. Demirkol, “An active sensor imaging for dense target environments,” *Optoelectronics and Advanced Materials, Rapid Communications*, vol. 14, no. 11–12, pp. 501–508, Nov. 2020.
- [8] J. Rissanen, “Modeling by shortest data description,” *Automatica*, vol. 14, no. 5, pp. 465–471, Sep. 1978, doi: 10.1016/0005-1098(78)90005-5.
- [9] W. Cheng, Z. Zhang, H. Cao, Z. He, and G. Zhu, “A Comparative Study of Information-Based Source Number Estimation Methods and Experimental Validations on Mechanical Systems,” *Sensors 2014, Vol. 14, Pages 7625–7646*, vol. 14, no. 5, pp. 7625–7646, Apr. 2014, doi: 10.3390/S140507625.
- [10] Y. Lee *et al.*, “Source Enumeration Approaches Using Eigenvalue Gaps and Machine Learning Based Threshold for Direction-of-Arrival Estimation,” *Applied Sciences 2021, Vol. 11, Page 1942*, vol. 11, no. 4, p. 1942, Feb. 2021, doi: 10.3390/APP11041942.
- [11] A. Hyvärinen and E. Oja, “Independent component analysis: algorithms and applications,” *Neural Networks*, vol. 13, no. 4–5, pp. 411–430, Jun. 2000, doi: 10.1016/S0893-6080(00)00026-5.
- [12] L. Huang, T. Long, E. Mao, and H. C. So, “MMSE-based MDL method for robust estimation of number of sources without eigendecomposition,” *IEEE Transactions on Signal Processing*, vol. 57, no. 10, pp. 4135–4142, 2009, doi: 10.1109/TSP.2009.2024043.
- [13] A. Wiesel and A. O. Hero, “Decomposable principal component analysis,” *IEEE Transactions on Signal Processing*, vol. 57, no. 11, pp. 4369–4377, 2009, doi: 10.1109/TSP.2009.2025806.
- [14] W. S. B. Ouedraogo, A. Souloumiac, M. Jaïdane, and C. Jutten, “Non-negative blind source separation algorithm based on minimum aperture simplicial cone,” *IEEE Transactions on Signal Processing*, vol. 62, no. 2, pp. 376–389, Jan. 2014, doi: 10.1109/TSP.2013.2287683.
- [15] K. Han and A. Nehorai, “Improved source number detection and direction estimation with nested arrays and ULAs using jackknifing,” *IEEE Transactions on Signal Processing*, vol. 61, no. 23, pp. 6118–6128, 2013, doi: 10.1109/TSP.2013.2283462.
- [16] T. Salman, A. Badawy, T. M. Elfouly, A. Mohamed, and T. Khattab, “Estimating the number of sources: An efficient maximization approach,” in *2015 International Wireless Communications and Mobile Computing Conference (IWCMC)*, IEEE, Aug. 2015, pp. 199–204. doi: 10.1109/IWCMC.2015.7289082.
- [17] H. Akaike, “Information Theory and an Extension of the Maximum Likelihood Principle,” in *Biogeochemistry*, vol. 1998, Springer Science and Business Media Deutschland GmbH, 1998, pp. 199–213. doi: 10.1007/978-1-4612-1694-0_15.
- [18] H. Akaike, “A New Look at the Statistical Model Identification,” *IEEE Trans. Automat. Contr.*, vol. 19, no. 6, pp. 716–723, 1974, doi: 10.1109/TAC.1974.1100705.
- [19] M. Wax and T. Kailath, “Detection of signals by information theoretic criteria,” *IEEE Trans. Acoust.*, vol. 33, no. 2, pp. 387–392, 1985, doi: 10.1109/TASSP.1985.1164557.
- [20] S. Kritchman and B. Nadler, “Determining the number of components in a factor model from limited noisy data,” *Chemometrics and Intelligent Laboratory Systems*, vol. 94, no. 1, pp. 19–32, Nov. 2008, doi: 10.1016/J.CHEMOLAB.2008.06.002.
- [21] K. P. . Burnham and D. Raymond. Anderson, “Model selection and multi-model inference: a practical information-theoretic approach,” 2011.
- [22] Z. Lu, “Source Enumeration in Sensor Array Processing: A Model Order Selection Problem,” Nov. 2012.
- [23] P. Liashchynskiy and P. Liashchynskiy, “Grid Search, Random Search, Genetic Algorithm: A Big Comparison for NAS,” Dec. 2019, Accessed: Sep. 09, 2025. [Online]. Available: <https://arxiv.org/pdf/1912.06059>
- [24] S. Singer, J. N.- Scholarpedia, and undefined 2009, “Nelder-mead algorithm,” *var.scholarpedia.org*, Accessed: Sep. 09, 2025. [Online]. Available: http://var.scholarpedia.org/article/Nelder-Mead_algorithm
- [25] J. C. Chen, R. E. Hudson, and Kung Yao, “Maximum-likelihood source localization and unknown sensor location estimation for wideband signals in the near-field,” *IEEE Transactions on Signal Processing*, vol. 50, no. 8, pp. 1843–1854, Aug. 2002, doi: 10.1109/TSP.2002.800420.
- [26] R. Schmidt, “Multiple emitter location and signal parameter estimation,” *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, Mar. 1986, doi: 10.1109/TAP.1986.1143830.
- [27] M. Shakeri, K. R. Pattipati, and D. L. Kleinman, “Optimal measurement scheduling for state estimation,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 31, no. 2, pp. 716–729, Apr. 1995, doi: 10.1109/7.381919.
- [28] Q. Liang, W. Wang, X. Liu, Z. Na, M. Jia, and B. Zhang, Eds., *Communications, Signal Processing, and Systems*, vol. 571. Singapore: Springer Singapore, 2020. doi: 10.1007/978-981-13-9409-6.

Özgeçmişler



Gökhan Yüksek is currently working as an Associate Professor in the Department of Electrical and Electronics Engineering at Batman University, Turkey. He received his B.Sc. in Electrical and Electronics Engineering from Nigde University, Nigde, and his M.Sc. and Ph.D. degrees in Electrical and Electronics Engineering from Mersin University, Mersin, Turkey. His research interests include metaheuristic optimization algorithms, intelligent and adaptive control systems, renewable energy, battery management, and hybrid energy storage technologies. He has authored and co-authored several research papers in international journals and conferences indexed in SCI/SCI-Expanded. He also serves as a reviewer for various leading journals in the fields of control engineering, optimization, and energy systems.



Timur LALE received his B.Sc. degree in Electrical and Electronics Engineering from Niğde University in 2012 and his M.Sc. and Ph.D. degrees in Electrical and Electronics Engineering from Dicle University in 2018 and 2024, respectively. His doctoral research focused on developing methods for the detection of demagnetization and stator inter-turn short-circuit faults in permanent magnet synchronous motors. He is currently an Assistant Professor in the Department of Electrical and Electronics Engineering, Division of Electrical Machines, at Batman University, Türkiye. His research interests include fault diagnosis and condition monitoring of electrical machines, artificial intelligence and deep learning methods, metaheuristic optimization, forecasting applications in electrical energy systems, electric vehicles, and battery management systems. He has published several research articles in national and international peer-reviewed journals in these fields.

Machine Learning for Physiological-Data-Based Driver Monitoring Systems: A Comparative Review and Experimental Case Study on Classification and Regression Approaches

Fizyolojik Veriye Dayalı Sürücü İzleme Sistemleri İçin Makine Öğrenmesi: Sınıflandırma ve Regresyon Yaklaşımlarına Yönelik Karşılaştırmalı Bir Derleme ve Deneysel Örnek Olay Çalışması

Bariş Aykent

 0000-0001-6711-9311

Electrical and Electronics Engineering Department
Işık University, İstanbul, Türkiye
b.aykent@gmail.com

Özet

Trafik kazaları, küresel çapta büyük kayıplara yol açan ciddi bir sorundur ve temelinde yorgunluk, dikkat kaybı ile uyku hali gibi insani faktörler yatar. Erken dönem ADAS sistemleri dış tehditleri ve araç dinamiklerini izleyerek daha çok reaktif (tepkisel) önlemlere odaklanırken, günümüzde proaktif (önleyici) Sürücü İzleme Sistemlerine (SİS) doğru temel bir paradigma değişimi yaşanmaktadır. Bu modern sistemler, kritik bir durum henüz ortaya çıkmadan riskleri tahmin edip önleyebilmek adına doğrudan sürücünün vücudundan alınan fizyolojik verileri kullanır. Bu çalışma, makine öğrenmesi (MÖ) modellerinin sürücü fizyolojik verilerini analiz etmede nasıl kullanıldığını incelemekte; sınıflandırma ve regresyon yaklaşımlarını karşılaştırmalı olarak ele alan kapsamlı bir literatür derlemesi sunmaktadır. Çalışmada temel fizyolojik ve davranışsal metrikler incelenmiş, çeşitli MÖ modellerinin uygulama alanları ve etkinlikleri analiz edilmiş; ayrıca veri değişkenliği, gerçek zamanlı işleme zorlukları ve sensör kısıtlamaları gibi kritik sorunlar tartışılmıştır. Derleme sonucunda, derin öğrenme yöntemlerinin artan gücü ve bu sistemlerin gerçek dünya koşullarında sağlam, genellenebilir bir şekilde çalışabilmesi için çok modlu sensör füzyonunun taşıdığı stratejik önem açıkça vurgulanmaktadır. Anahtar kelimeler: Makine Öğrenmesi (MÖ), Derin Öğrenme (DÖ), Sürücü İzleme Sistemi (SİS), Fizyolojik Sinyaller, ADAS

Abstract

Traffic collisions remain a devastating global issue. Human error — specifically fatigue, loss of focus, and drowsiness — plays a massive role in these tragic events. Early Advanced Driver Assistance Systems (ADAS) mainly took a reactive approach, keeping an eye on external threats and vehicle movements. However, the automotive industry is currently undergoing a massive philosophical shift toward proactive Driver Monitoring Systems (DMS). Instead of waiting for a car to swerve, these modern setups look directly at the driver's own physiological data. By doing so, they can predict and neutralize

risks long before an actual emergency happens. This paper takes a deep dive into how researchers are applying machine learning (ML) models to analyze this physiological data, specifically comparing classification and regression techniques. The core behavioral and physical metrics are broken down, to evaluate how well different ML algorithms actually perform, and tackle the messy, real-world challenges like sensor limits and unpredictable data. Ultimately, this review points out that deep learning is rapidly taking over the field, and combining different types of sensors (multimodal fusion) is the smartest strategy for building reliable systems that work in chaotic, everyday driving conditions.

Keywords: Machine Learning (ML), Deep Learning (DL), Driver Monitoring System (DMS), Physiological Signals, ADAS

1. Introduction

Driving is a complex, dynamic task that requires drivers to continuously integrate visual input, cognitive judgment, and physical control [6]. This continuous demand on attention causes mental workload to fluctuate, at times pushing drivers toward inattention or fatigue — factors that are responsible for a substantial share of traffic accidents worldwide [7]. Existing driver-assistance technologies, such as lane-departure and forward-collision warnings, address this problem reactively: they intervene only after a hazardous situation has already begun to develop, rather than addressing its root cause, namely the compromised physiological or cognitive state of the driver [1].

Consequently, the automotive safety field is undergoing a shift from reactive interventions toward proactive Driver Monitoring Systems (DMS). Continuous assessment of physiological data provides a direct and objective measure of a driver's internal state [2], enabling predictive rather than reactive interventions. This need is amplified by the growing automation of driving tasks, which paradoxically increases the risk of driver disengagement and reduced vigilance.

Driver-state monitoring has been studied by human-factors researchers for decades, drawing on foundational concepts such as the Yerkes–Dodson law of arousal [6]. Recent advances in

low-cost, high-fidelity sensing have accelerated interest from both academia and industry in monitoring driver state alongside vehicle state [6]. This interest has translated into regulatory action: the European Union mandated DMS in all new vehicle models beginning in 2024, and Euro NCAP now awards safety-rating credit to vehicles equipped with such systems [2]. Recent systematic reviews confirm that the field has grown rapidly across vision-based, physiological, and multimodal approaches, with an increasing share of studies now employing deep learning rather than classical pipelines [33], and highlight that distraction- and drowsiness-detection research remains dominated by visual modalities even as physiological and multimodal methods gain traction [34]. Complementary surveys of cognitive-state recognition report that deep learning became the dominant methodology in this space by 2023, alongside a broader movement toward multimodal physiological sensing [3], while a systematic review of driver-state detection technologies further catalogues the growing diversity of physiological and vehicular sensing approaches [35], and device-level reviews continue to track new wearable and health-monitoring hardware entering this application space [36].

Modern sensing technologies allow continuous, largely unobtrusive acquisition of physiological signals — including heart rate, respiration, and eye movement — while a person drives [6]. However, raw physiological data has limited value without interpretation, which is where machine learning becomes essential: ML algorithms are able to extract subtle, non-obvious patterns from complex biological signals that correspond to states such as drowsiness or distraction [10].

This paper reviews current machine learning strategies for interpreting driver physiological data, focusing on two complementary paradigms: classification, which assigns discrete driver states (e.g., "awake" versus "drowsy"), and regression, which predicts driver state on a continuous scale. Section 2 outlines the physiological and behavioral indicators used in this literature; Section 3 reviews the ML models applied to these indicators; Section 4 discusses evaluation metrics, data-quality challenges, and multimodal sensor fusion; Section 5 presents an original experimental case study applying several of these models to a physiological dataset; and Section 6 concludes with implications for future work.

2. Methods: physiological and behavioral indicators

2.1. Physiological measures from the human body

2.1.1. Cardiovascular markers

Heart rate (HR) and heart rate variability (HRV) are frequently examined because they provide a direct window into autonomic nervous system activity [6]. Under stress, fatigue, or cognitive overload, heart rate typically increases while HRV decreases, and these physiological shifts are often associated with degraded driving performance [12]. Electrocardiography (ECG) remains the gold-standard acquisition method, although its reliance on skin-contact electrodes limits comfort and practicality in everyday driving. Photoplethysmography (PPG) offers a less intrusive alternative with comparable diagnostic value [13].

2.1.2. Neurophysiological signals

Electroencephalography (EEG) captures the brain's electrical activity and is highly informative regarding alertness state. The slow δ (0.5–4 Hz) and θ (4–8 Hz) bands become dominant during the onset of sleep, whereas α (8–14 Hz) and β (14–25 Hz) bands are associated with relaxed or actively attentive states [14]. A shift in spectral power from β toward α and θ activity is a reliable indicator of increasing drowsiness [14]. Electrooculography (EOG) is similarly effective for drowsiness detection, as it measures blink duration and Slow Eye Movements (SEM) [15]. Electrodermal activity (EDA, also referred to as GSR) quantifies sweat-gland activity via skin electrical resistance and serves as a proxy for mental workload [6].

2.2. Behavioral and vehicular cues

Camera-based facial monitoring provides an effective, non-invasive alternative to physiological sensing [16]. In-cabin cameras can detect heavy eyelids, yawning, and head-drooping behavior. A widely used metric is PERCLOS, defined as the percentage of time a driver's eyelids are at least 80% closed [8]. Vehicle-based indicators, such as irregular steering or lane-position drift, can likewise suggest a compromised driver state, although such measures must be interpreted cautiously, since poor road conditions or individual driving habits can confound the signal [19].

Table 1: Summary of physiological and vehicular signal types for driver state monitoring.

Signal Type	Derived Metric(s)	Indicator(s)	Source	Deployment
Cardiovascular	HR, HRV	Workload, Stress	ECG, PPG	Contact / Non-contact
Neurophysiological	EEG Bands ($\delta, \theta, \alpha, \beta$)	Alertness, Drowsiness	EEG	Invasive
Ocular	Blink Duration, SEM	Drowsiness	EOG, Camera	Contact / Non-intrusive
Autonomic (EDA)	SCL, SCR	Arousal, Workload	GSR Sensor	Contact
Behavioural	PERCLOS, EAR, MAR	Drowsiness, Fatigue	Camera	Non-intrusive
Vehicular	Steering Angle, Lane Dev.	Fatigue, Aggression	In-vehicle Sensors	Non-intrusive

3. Machine learning paradigms for driver state analysis

3.1. Driver state classification: categorical detection

3.1.1. Traditional machine learning

Traditional ML algorithms dominated the field prior to the widespread adoption of deep learning. Support Vector Machines (SVMs), for example, have proven effective at detecting impaired driving; combining physiological and vehicle-movement features, researchers achieved approximately 70% classification accuracy [21]. Random Forest classifiers also perform well when provided with hand-crafted features such as eye aspect ratio (EAR), effectively separating alert from drowsy drivers [23].

3.1.2. The rise of deep learning

Convolutional Neural Networks (CNNs) have substantially advanced the field, particularly for the analysis of video streams and raw physiological signals [3]. Numerous well-established architectures have been applied to the drowsiness-detection problem, including VGG16, MobileNet, ResNet50V2, and AlexNet [16]. Reported results are often striking: one study

achieved 99.71% accuracy with ResNet50V2 on the NITYMED database [26], while another employed an ensemble of CNNs operating solely on in-vehicle sensor data and achieved a 94.2% detection rate [20].

3.2. Driver state regression: continuous prediction

3.2.1. Continuous physiological metrics

Rather than assigning discrete labels such as "awake" or "asleep," regression models predict driver state on a continuous scale, which is well suited to variables such as PERCLOS or seat-pressure distribution [9]. XGBoost has demonstrated particularly strong performance in this context: one study predicting PERCLOS achieved an RMSE of 3.847 and an adjusted R^2 of 0.996 [18].

3.2.2. Predicting subjective states

The Karolinska Sleepiness Scale (KSS), a 9-point self-report instrument, is commonly used and correlates strongly with objective EEG measures [28]. Regression-based estimation of KSS enables multi-stage warning architectures, for example issuing a mild alert at KSS level 5, an intensified alert at level 7, and an autonomous braking or pull-over response at level 9.

Table 2: Comparative Overview of Machine Learning Models for Driver State Detection.

Model	Type	Typical Input	Performance	Key Strength
SVM	Classification	HRV, blink duration	70% accuracy	Effective with small datasets
Random Forest	Classification	EAR, MAR	High accuracy	Robust; diverse data types
CNN (ResNet50V2, MobileNet)	Classification	Image/video streams	Up to 99.71%	Automates feature extraction
Ensemble CNN	Classification	In-vehicle sensors	Up to 94.2%	Handles multiple states
XGBoost	Regression	Physiological/behav.	RMSE=3.847, R^2 =0.996	Efficient, explainable
Functional Lin. Reg.	Regression	Physiological vars.	~90% accuracy	Interpretable models

4. Performance metrics, datasets, and multimodal fusion

4.1. Evaluation metrics for robustness

Accuracy alone can be misleading, particularly when the underlying dataset is imbalanced [20]. For classification tasks, Precision, Recall, and F1-score provide a more complete picture of model behavior [9]. For regression tasks, RMSE and MAE remain the standard metrics for quantifying prediction error [18].

4.2. Challenges in data collection and generalization

Achieving reliable performance in a controlled simulator does not guarantee equivalent performance under real-world driving conditions, and generalization remains a significant challenge [20]. Inter-individual physiological variability — arising from genetic, cultural, and situational factors — means that a stress-induced heart-rate response can differ substantially between

individuals [30]. Consequently, it is not unusual for model accuracy to fluctuate by several percentage points across different test sessions or conditions [30]. Sensor-comfort trade-offs compound this challenge: high-fidelity sensors such as EEG caps are often impractical for continuous everyday use, while low-cost, non-intrusive cameras are highly sensitive to lighting conditions and frequently degrade in performance after dark [19].

4.3. Overcoming limitations with multimodal sensor fusion

Multimodal sensor fusion addresses the limitations of any single sensing modality. By combining physiological signals, facial-camera feeds, and steering-wheel inputs, researchers obtain a substantially more complete representation of driver state [4]. No individual sensor is fully reliable on its own; fused approaches have been reported to push F1-scores as high as 96.10% [9]. Industry implementations, such as those from Rheinmetall and BlackBerry, similarly demonstrate that combining camera and radar data is central to building systems that are robust in real-world deployment.

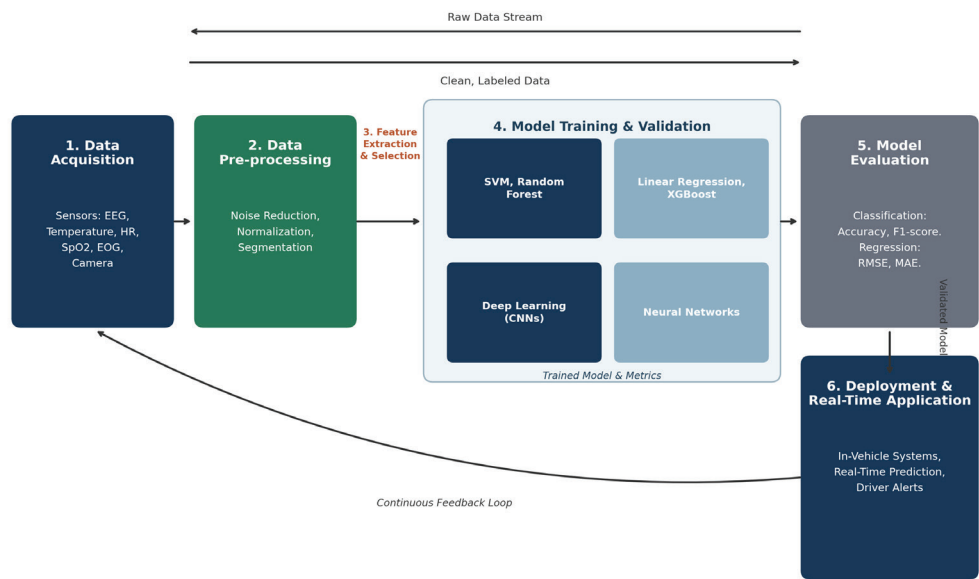


Figure 1: Driver State Analysis Flowchart in This Study.

5. Experimental study: ML models on physiological dataset

To evaluate these concepts empirically, the publicly available Kaggle Driver Physiological Dataset [37] was used, which provides continuous recordings of three physiological channels — body temperature, heart rate, and peripheral oxygen saturation (SpO2) — collected from [authors: please confirm exact sample size, sensor hardware, and sampling protocol] under simulated driving conditions. Records containing missing values or physiologically implausible readings (e.g., SpO2 outside the 70–100% range) were removed prior to analysis, and raw body-temperature readings were converted from Fahrenheit to Celsius. The resulting features were standardized (zero mean, unit variance) and expanded into a five-dimensional feature space [authors: please confirm the specific derived variables used to expand from three raw signals to five features] before Principal Component Analysis (PCA) was applied, reducing the feature space to two principal components that jointly explained over 90% of the total variance. The dataset was then partitioned into training and test subsets using an 80/20 split, with class balance verified between the stressed (Driver State ≥ 7) and non-stressed (Driver State < 7) groups.

5.1. Regression results

Three regression algorithms were evaluated for their ability to predict the driver's continuous state: Linear Regression, a Decision Tree, and a Random Forest.

Table 3: Regression Model Results.

Model	MSE	RMSE	MAE	R ²
Linear Regression	1.59	1.26	0.90	0.70
Decision Tree	0.31	0.56	0.15	0.94
Random Forest	0.23	0.48	0.20	0.96

The Random Forest Regressor substantially outperformed the alternative models ($R^2 = 0.96$, $RMSE = 0.48$), demonstrating strong generalization without evidence of overfitting to noise in the training data. Linear Regression performed considerably worse ($R^2 = 0.70$), consistent with the expectation that physiological relationships are rarely well approximated by a linear model.

5.2. Classification results

The models were then evaluated on a binary classification task: whether the driver was stressed (Driver State ≥ 7 , class 1) or relatively calm (Driver State < 7 , class 0).

Table 4: Classification performance comparison of models for driver state detection.

Model	Accuracy	Prec. 0	Rec. 0	F1 0	Prec. 1	Rec. 1	F1 1
Logistic Reg.	0.88	0.81	0.81	0.81	0.91	0.91	0.91
Decision Tree	0.99	0.97	0.99	0.98	0.99	0.99	0.99
Random Forest	0.98	0.95	1.00	0.98	1.00	0.98	0.99

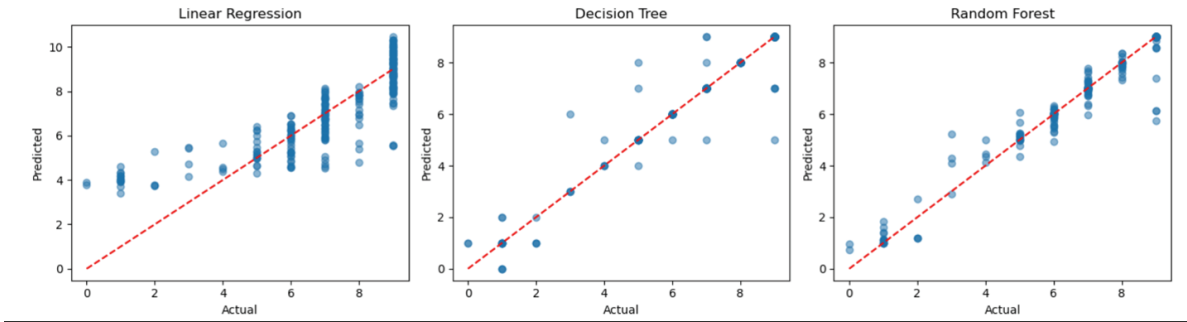


Figure 2: Regression Plots: Actual vs. Predicted Values.

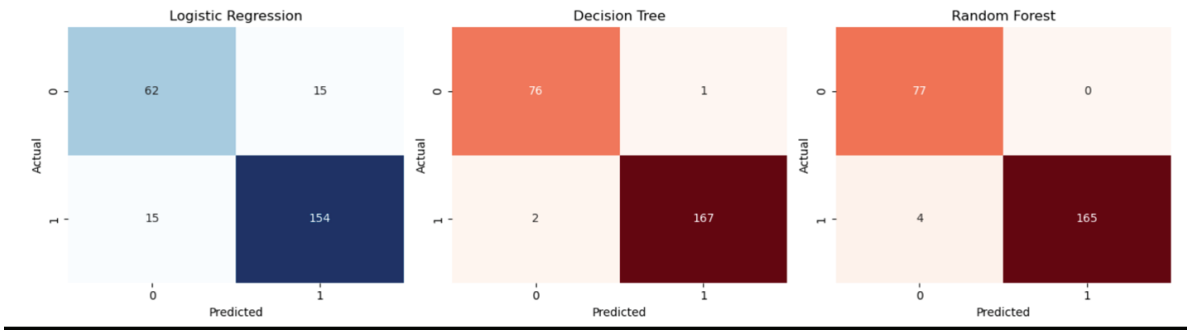


Figure 3: Confusion Matrices for Classification Models.

Both the Decision Tree and Random Forest classifiers achieved high accuracy, exceeding 98%. Random Forest in particular achieved perfect recall for the calm state (class 0) and an AUC approaching 1.00. Logistic Regression, despite its simplicity, still achieved a respectable 87.8% accuracy. These results indicate that tree-based ensemble methods are highly effective for classifying driver state from physiological data.

5.3. Feature importance and permutation analysis

Model interpretability analysis revealed that SpO2 consistently emerged as the dominant predictive variable across all models. Body temperature ranked second in importance, while heart rate contributed comparatively less to model decisions. A permutation importance test confirmed that this ranking was robust rather than incidental (Table 5). Partial Dependence

Plots (PDPs) further revealed a clear trend: higher SpO2 values were strongly associated with classification as the stressed state. This finding is biologically plausible, as variation in blood oxygen saturation is closely linked to overall physiological and cognitive state.

Table 5: Permutation Importance of Features.

Feature	Importance Mean	Importance Std
SpO2	0.38	0.02
Body Temperature	0.09	0.02
Heart Rate	0.05	0.01

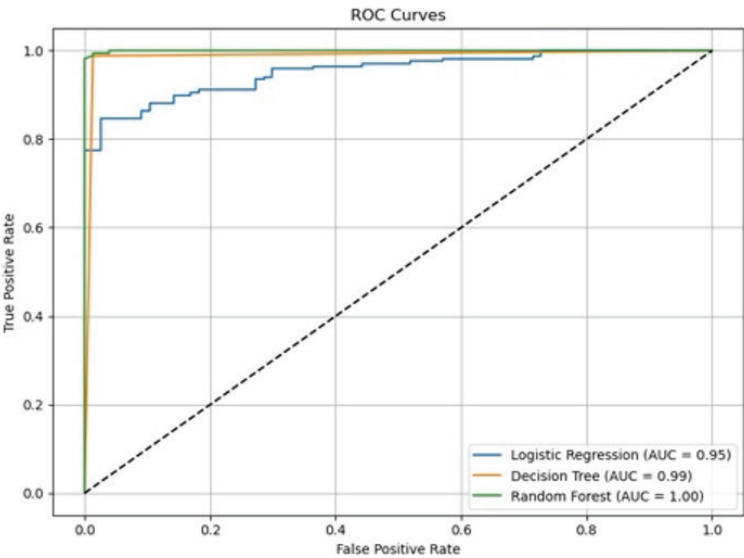


Figure 4: ROC Curves for Classification Models.

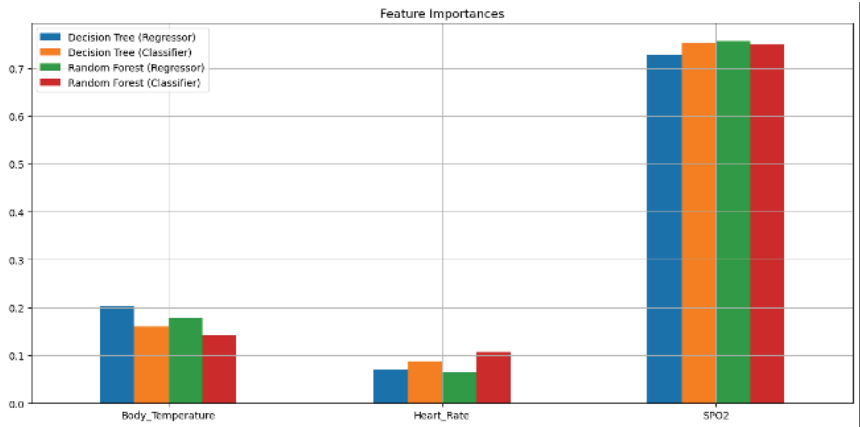


Figure 5: Feature Importance Across Regression and Classification Models.

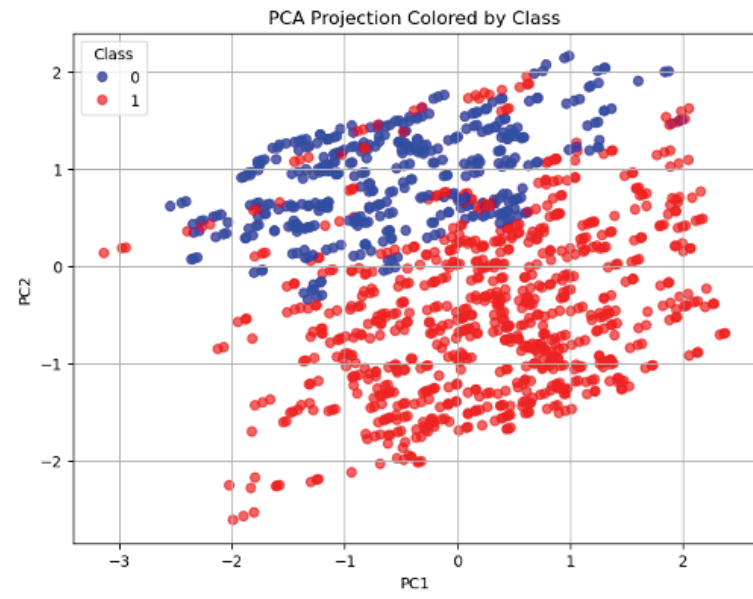


Figure 6: Principal Component Analysis (PCA) Projection Colored by Class.

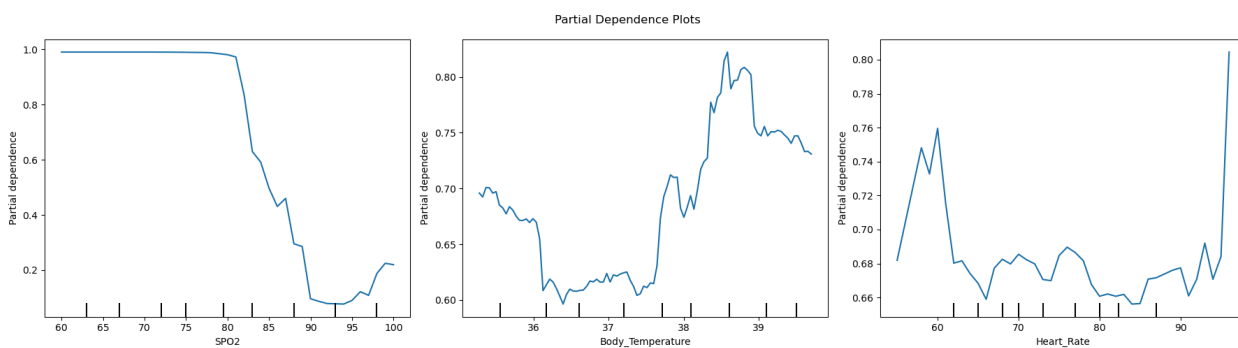


Figure 7: Partial Dependence Plots (PDP) for Top Features.

6. Conclusions

This experimental case study demonstrates that machine learning models can accurately track and predict physiological driver state. Among the features examined, SpO2 emerged as the most influential predictor. For continuous-state prediction, the Random Forest Regressor performed best, achieving an R^2 of 0.96; for discrete-state classification, the Random Forest Classifier again performed best, achieving approximately 98.4% accuracy. These results are consistent with known physiological relationships, as blood-oxygen saturation is closely linked to human stress response and cognitive performance.

The broader field has evolved considerably beyond classical machine learning toward deep learning and complex multimodal sensor fusion, and the present experimental results — obtained from a comparatively simple, single-modality tabular dataset using classical models — should be interpreted within that context. As reviewed in Section 3.1.2 and Section 4.3, CNN-based architectures and multimodal fusion pipelines have achieved substantially higher and more robust performance than the classical models evaluated here when applied to richer, higher-dimensional sensor streams such as video or multi-channel physiological recordings

[3,4,9,16,20,26,33,34]. A natural extension of this work is therefore to apply deep learning architectures (e.g., 1D-CNNs or LSTM networks) to time-resolved physiological signals, and to augment the present single-modality dataset with complementary sensing channels — such as facial video or EEG — to evaluate whether multimodal fusion yields the accuracy and robustness gains reported elsewhere in the literature; this extension was beyond the scope of the tabular, session-level dataset used in the present case study and is identified here as a priority for future work. Beyond this, the industry must focus on several additional priorities. First, more comfortable sensing modalities (e.g., remote photoplethysmography, rPPG) are needed. Second, models must be developed that adapt to inter-individual physiological variability. Third, Explainable AI (XAI) techniques should be prioritized so that regulators and end users can trust these systems. Finally, validated algorithms must be shown to operate reliably in real time within an actual vehicle.

Acknowledgment

The data utilized in this specific experimental portion was retrieved from the Kaggle Driver Physiological Dataset. The

raw body temperature data was converted into Celsius before running the analyses.

7. References

- [1] NHTSA, "Driver Assistance Technologies," accessed Sep. 2025. [Online]. Available: <https://www.nhtsa.gov/vehicle-safety/driver-assistance-technologies>
- [2] KORE Wireless, "What Is a Driver-Monitoring System (DMS)?" accessed Sep. 2025.
- [3] K. Jin, A. Rubio-Solis, R. Naik, D. Leff, J. Kinross, and G. Mylonas, "Human-Centric Cognitive State Recognition Using Physiological Signals: A Systematic Review of Machine Learning Strategies Across Application Domains," *Sensors*, vol. 25, no. 13, Art. 4207, 2025. doi: 10.3390/s25134207
- [4] ResearchGate, "Multimodal Fatigue Detection in Drivers via Physiological and Visual Signals," 2024.
- [5] ResearchGate, "ECG-based fatigue driving detection model," 2024.
- [6] E. T. Solovey, M. Zec, E. A. Garcia Perez, B. Reimer, and B. Mehler, "Classifying Driver Workload Using Physiological and Driving Performance Data: Two Field Studies," *Proc. ACM CHI '14*, 2014.
- [7] Y. Sun, "An In-Vehicle Physiological Signal Monitoring System for Driver Fatigue Detection," *TRB*, 2011.
- [8] T. Abe, "PERCLOS-based technologies for detecting drowsiness: current evidence and future directions," *SLEEP Advances*, vol. 4, no. 1, zpad006, 2023. doi: 10.1093/sleepadvances/zpad006
- [9] ResearchGate, "Driver Fatigue Transition Prediction in Highly Automated Driving," 2020.
- [10] SPIE, "Assessing Deep Learning Techniques for Human Driving Behavior Prediction," 2025.
- [11] BlackBerry QNX, "What Is a Driver Monitoring System?" accessed Sep. 2025.
- [12] S. Narayanan et al., "Estimating Markers of Driving Stress through Multimodal Physiological Monitoring," arXiv:2507.14146, 2025.
- [13] Eventiotic, "Physiological Measurements for Determining Driver's State," 2020.
- [14] PMC, "Driving Drowsiness Detection Using Spectral Signatures of EEG," *Frontiers in Neuroscience*, 2023.
- [15] PMC, "An EOG Sensor's Ability to Detect Driver Drowsiness," 2023.
- [16] MDPI Applied Sciences, "AI- and Deep Learning-Powered Driver Drowsiness Detection," vol. 15, no. 3, 2025.
- [17] IRJET, "Machine Learning Based Driver Monitoring System," vol. 10, no. 12, 2023.
- [18] Univ. of Michigan, "Predicting Driver Fatigue in Automated Driving with Explainability," 2021.
- [19] IRJET, "Driver Drowsiness Detection System," vol. 8, no. 5, 2021.
- [20] PMC, "Ensemble CNN to Detect Drowsy Driving with In-Vehicle Sensor Data," 2021.
- [21] A. Shi et al., "SVM Classification of Drunk Driving Behaviour," *Int. J. Environ. Res. Public Health*, vol. 14, no. 1, 2017.
- [22] ResearchGate, "SVM Classification of Drunk Driving Behaviour," 2017.
- [23] IJARST, "Driver Drowsiness Detection using ML and DL," 2023.
- [24] PMC, "Deep Learning in Physiological Signal Data: A Survey," *IEEE Rev. Biomed. Eng.*, 2020.
- [25] ResearchGate, "DL Model for Driver Drowsiness Detection in Low-Light Conditions," 2023.
- [26] MDPI Applied Sciences, "CNN-Based Approach for Driver Drowsiness Detection," vol. 13, no. 13, 2023.
- [27] TOJDAC, "Regression Model for Driving Fatigue Detection Based on Seat Pressure Distribution," vol. 8, 2018.
- [28] T. Åkerstedt and M. Gillberg, "Subjective and objective sleepiness in the active individual," *Int. J. Neuroscience*, vol. 52, pp. 29–37, 1990.
- [29] V. Putilov and O. Donskaya, "Construction and Validation of EEG Analogues of the KSS," Semantic Scholar, 2014.
- [30] ResearchGate, "Detecting Stress During Real-World Driving Tasks Using Physiological Sensors," 2003.
- [31] Rheinmetall, "Sensor Fusion," 2023.
- [32] insideAI News, "Using Machine Learning with Health Data: Challenges and Pitfalls," 2017.
- [33] O. F. Hassan, A. F. Ibrahim, M. A. Makhoul, and B. Hafiz, "Intelligent driver monitoring systems: a survey of drowsiness detection technologies for road safety," *Artificial Intelligence Review*, Springer, 2026. doi: 10.1007/s10462-026-11505-w
- [34] A. Dontoh, S. Ivey, L. Sirbaugh, A. Danyo, and A. Aboah, "Visual Dominance and Emerging Multimodal Approaches in Distracted Driving Detection: A Review of Machine Learning Techniques," arXiv:2505.01973, 2025.
- [35] M. S. AL-Quraishi, S. S. A. Ali, M. AL-Qurishi, T. B. Tang, and S. Elferik, "Technologies for detecting and monitoring drivers' states: A systematic review," *Heliyon*, vol. 10, no. 20, e39592, 2024. doi: 10.1016/j.heliyon.2024.e39592
- [36] L. Melders, R. Smigins, and A. Birkavs, "Recent Advances in Vehicle Driver Health Monitoring Systems," *Sensors*, vol. 25, no. 6, Art. 1812, 2025. doi: 10.3390/s25061812
- [37] Kaggle (user: gowrishankarachari), "Driver Physiological Dataset," accessed 2025. [Online]. Available: <https://www.kaggle.com/datasets/gowrishankarachari/driver-physiological-dataset>

Özgeçmişler



Dr. Baris Aykent is a dynamics and controls expert specializing in real-time software development for ground and aerial vehicles, including UAVs. He holds a PhD in Control Engineering from Arts et Métiers ParisTech (2013), an MSc in System Dynamics & Control from Istanbul Technical University, and a BSc in Mechanical Engineering from Yıldız Technical University. Over a career spanning Germany (BAST-Daimler-UniBw), France (Renault), and Türkiye (Hexagon, Profen, Togg, Altınay, and Evry), he has led development of driver-assistance and vehicle-control systems such as ESC, TCS, ABS, EPS, BMS, and Hill Assist Control across MIL/SIL/HIL/DIL/VIL environments, most recently serving as Head of Control Software at Evry. His expertise also spans driving simulators and motion/simulator sickness, system identification, Kalman filtering, and machine learning (including FFNN, NARX, and reinforcement learning) applied to autonomous and electric vehicle systems. He currently lectures on Applied AI in Automotive Systems at both MEF University and Işık University, and has authored several Springer book chapters on ESC and vehicle control.

Giriş Gecikmeli Robot Manipülatörlerinin Görev Uzayında Denetimi: Kendi Kendini Uyarlayan Bulanık Mantık Tabanlı Bir Yaklaşım

Task-Space Control of Robot Manipulators Subject to an Input Delay: A Self-Tuning Fuzzy-Logic-Based Approach

Serhat Obuz¹

 0000-0002-3060-9306

¹Elektrik-Elektronik Mühendisliği Bölümü, Tarsus Üniversitesi, Mersin 33400, Türkiye
serhat.obuz@tarsus.edu.tr

Özet

Bu çalışmada, denetim girişi bilinen ve sabit bir gecikmeye uğrayan n -bağlantılı robot manipülatörler için, doğrudan görev uzayında çalışan kendi kendini uyarlayan bulanık mantık tabanlı bir denetleyici geliştirilmektedir. Manipülatör dinamiklerinin bilinmediği varsayılmaktadır. Tüm belirsizlik, çıkış ağırlıkları ile üyelik fonksiyonlarının merkez ve genişlikleri projeksiyon operatörlü e -modifikasyon yasalarıyla çevrimiçi güncellenen bir bulanık ağ tarafından yaklaşılmaktadır. Gecikme bilindiğinden, telafi terimi gerçek geçmiş denetim değerlerinin üstel ağırlıklı integraliyle kurulmaktadır. Böylece gecikmeli giriş, hata dinamiklerinden gecikme uyumsuzluğu terimi doğurmadan çıkarılmaktadır. İzleme hatasının yarı-globel düzgün nihai sınırlı (SGUUB) olduğu, Lyapunov-Krasovskii terimleriyle genişletilmiş bileşik bir fonksiyonel ile kanıtlanmaktadır. Ayrıca kazanç seçimi için açık bir rehber, yapıcı bir doğrusal olmayan sönümleme alternatifi ve 10–175 ms gecikme aralığını kapsayan sayısal bir çalışma sunulmaktadır. Sıfır ağırlıkla ve hiçbir dinamik model bilgisi kullanılmadan başlayan bulanık sistem, ideal ileri beslemeyi çevrimiçi kurmakta ve 10 mm'lik başlangıç hatası 0.31–2.43 mm bandına inmektedir. Gecikme telafi teriminin 50 ms, bulanık mantık tabanlı ileri beslemenin ise 100 ms ve üzerindeki gecikmelerde kararlılık için gerekli olduğu gösterilmektedir.

Anahtar Kelimeler: Doğrusal olmayan denetim, uyarlamalı bulanık denetim, robot manipülatörler, görev uzayı denetimi, giriş gecikmesi.

Abstract

A self-tuning (self-organized) fuzzy-logic controller is developed for an n -link robot manipulator that operates directly in the task space and whose control channel is affected by a known constant input delay. The lumped dynamic uncertainty is reconstructed on-line by a fuzzy-logic network whose output weights and whose membership-function centers and widths are adapted through e -modification update laws equipped with a projection operator. Because the delay is known, a delay-free auxiliary signal is injected into the open-loop error system through an exponentially weighted integral of the true past control values. The delayed input is therefore removed from the error dynamics without a delay-mismatch term. A composite Lyapunov functional augmented with Lyapunov-Krasovskii (LK) terms is used to establish that the task-space tracking error is semi-globally uniformly ultimately bounded (SGUUB). A design procedure and a constructive nonlinear-damping alternative are also provided. Analytical development is emphasized and supported by a numerical study on a two-link revolute manipulator in task space under known input delays of 10–175 ms. Starting from zero weights and without any dynamic model knowledge, the fuzzy system reconstructs the lumped uncertainty online with a few percent relative error and drives a 10 mm initial error to a 0.31–2.43 mm steady state error. The delay-compensation term and the fuzzy feedforward are each shown to be necessary for stability beyond 50 ms and 100 ms, respectively.

Key Words: Nonlinear control, adaptive fuzzy control, robot manipulators, task-space control, input delay.

hesap yükü getirir hem de tekillik yakınlarında güvenilirliğini yitirir. Denetleyiciyi doğrudan görev uzayında tasarlamak bu adımı ortadan kaldırır. Karşılığında belirsiz dinamikler manipülatör Jacobian'ı ile iç içe geçer ve hem tasarımın hem de kararlılık kanıtının bu eşleşmeyi açıkça taşıması gerekir [6–8].

Robot manipülatörler için denetleyici tasarımı esnasında karşılaşılan ikinci güçlük giriş gecikmelerinin bozucu etkilerinden kaynaklanmaktadır. Gerçek bir sistemde komutlanan tork motora anında ulaşmaz. Eyleyici dinamikleri, sinyal işleme ve özellikle ağ üzerinden haberleşme araya bir gecikme koyar [9]. Bu gecikme küçük olsa bile denetleyici başarımını önemli ölçüde düşürür. Denetleyici tasarımında hiç

1. Giriş

Bir robot manipülatörün uç noktasının istenen yörüngede tutmak, dinamiklerin doğrusal olmayan ve genellikle tam bilinmeyen yapısı yüzünden halen açık bir problemidir [1, 2]. Robot manipülatörler için denetleyici tasarımının bahsi geçen bu zorluklarına ek olarak uygulamadan kaynaklanan bir güçlük daha mevcuttur. Hareket denklemleri eklem uzayında yazılır, oysa görev çoğu zaman Kartezyen uzayda tarif edilir [3]. Denetleyici eklem uzayında tasarlanırsa, istenen Kartezyen yörüngeyi her örnekleme anında eklem açılarına çevirmek, yani ters kinematik çözmek gerekir [4, 5]. Bu hem

hesaba katılmazsa iyi ayarlanmış bir kontrol çevrimini kararsızlaştırabilir. Doğrusal olmayan sistemlerde gecikmeyle baş etmenin genel olarak iki temel yolu vardır. Birincisi öngörücü yöntemlerdir. Sistemin gecikme kadar sonraki durumu kestirilir ve problem gecikmesiz bir probleme dönüştürülür. Euler–Lagrange sistemleri için temsili bir örnek literatürde verilmiştir [10]. İkincisi Lyapunov–Krasovskii (LK) yaklaşımıdır. Bu yolda durum öngörüsü yapılmaz, bunun yerine geçmiş denetim değerlerinin bir integrali yardımcı bir sinyal olarak tanımlanır ve açık çevrim hata sistemi gecikmesiz biçimde yazılabilir [11–14]. Bu ikinci yol güncel yazın taramasında, gecikme zamanla değişse bile belirsiz Euler–Lagrange sistemlerini ele alabilecek olgunluktadır [11, 13]. Gecikme bilinmiyorsa, bu tarz yaklaşımın sonucu olarak boyutu gecikme-kestirim hatasıyla büyüyen bir kalıntı nihai sınırdır [12, 14].

Gecikmenin bilinen ve sabit olması, bu iki yolun kesişimindeki en elverişli durumdur. Bu durumda telafi terimi gerçek geçmiş denetim değerleriyle kurulabilir. Sonuçta nihai hata sınırında gecikme kestirimine bağlı hiçbir bileşen kalmaz ve sınır denetim kazancıyla monoton biçimde küçülür. Gecikme bilinmediğinde ise kazanç koşulları, tolere edilebilen gecikme uyumsuzluğuyla ters yönde çalışır [12, 14]. Bu çalışma tam da bu iki durum arasındaki farkı ele almaktadır. Aynı bulanık mantık altyapısının gecikmenin kendisini kestirmek için kullanılması, doğal bir sonraki adımdır. Bu konuda öncü olabilecek bu çalışmada, gecikmenin bilindiğini varsaymak, ağ tabanlı ve örneklenmiş-veri sistemlerinde gerçekçi olabilir, çünkü paketlere zaman damgası konabildiği veya taşıma gecikmesi ölçülebildiği için gecikme bazı uygulamalarda bilinebilir.

Belirsizlik tarafında ise çevrimiçi fonksiyon yaklaşımcıları standart araç olmuştur. Dinamikler bilinmeyen parametrelere göre doğrusal yazılabiliyorsa uyarlamalı yöntemler [15, 16], son yıllarda ise bileşik ve eşzamanlı öğrenme şemaları [8, 17–23] kullanılır. Böyle bir parametrik model yoksa, bulanık mantık ya da sinir ağı yaklaşımcıları belirsizliği doğrudan bir fonksiyon olarak öğrenir [24]. Bulanık sistemlerin ayırt edici yanı, uzman bilgisinin kural tabanına doğrudan yazılabilesidir. Yalnızca çıkış ağırlıkları uyarlanırsa, yaklaşımın kalitesi baştan seçilen ve sonradan değişmeyen üyelik fonksiyonlarıyla sınırlı kalır. Üyelik fonksiyonlarının merkezleri ve genişlikleri de çevrimiçi güncellenirse bu sınır kalkar. Bu tür yapılara kendi kendini uyarlayan bulanık mantık tabanlı sistemler denir [24, 25]. Lyapunov analizinde bu ek serbestliğin gerektirdiği kalan terim sınırları literatürde verilmiş [24], yaklaşım daha sonra görev-uzayı denetimine taşınmıştır [25]. Bu çalışmanın başlıca özgün yanları aşağıda sunulmuştur.

1. Bu çalışmada, bilinen sabit bir giriş gecikmesi altında, doğrudan görev uzayında çalışan kendi kendini uyarlayan bulanık bir denetleyici için Lyapunov tabanlı bir kararlılık analizi verilmektedir. Bilinmeyen dinamikler, ağırlıkları ile üyelik merkez ve genişlikleri projeksiyon operatörlü e-modifikasyon yasalarıyla uyarlanan bir bulanık ağ tarafından yaklaşılmaktadır. Bu yönüyle, güncel yazın dizisindeki [24–26] birçok yaklaşımdan farklıdır.
2. Üstel ağırlıklı bir gecikme telafi terimi ve LK fonksiyonelleri kullanılarak izleme hatasının yarı-global düzgün nihai sınırlı (SGUUB) olduğu kanıtlanmaktadır. Elde edilen sınırdaki gecikme uyumsuzluğuna bağlı hiçbir bileşen bulunmamakta ve sınır denetim kazancıyla monoton azalmaktadır. Bu davranış, gecikme bilinmediğinde kazancın bu iki etkiyi karşı karşıya getiren

ödüleşiminin tersidir [12, 14]. Ayrıca kazanç seçimi için açık bir rehber ile 10–175 ms gecikme aralığını kapsayan sayısal bir çalışma sunulmaktadır.

Gösterim. $\|\cdot\|$ vektörler için Öklid normunu, matrisler için indüklenmiş 2-normunu gösterir. Simge I_n , $n \times n$ birim matristir. Gösterimler $\lambda_{\min}(\cdot)$ ve $\lambda_{\max}(\cdot)$ en küçük ve en büyük özdeğerlerdir. Ayrıca $\text{tr}\{\cdot\}$ matris izidir. Küme $\mathcal{L}_\infty$ sınırlı fonksiyonlar uzayıdır. Bir s niceliği için $\underline{s}$ ve $\bar{s}$ bilinen alt ve üst sınırları, $\hat{s}$ kestirimi, $\tilde{s} \triangleq s - \hat{s}$ ise kestirim hatasını gösterir. Bağlamdan anlaşıldığında fonksiyon argümanları yazılmayacaktır.

2. Dinamik ve Kinematik Model

Denetim girişi gecikmeye uğrayan $n \in \mathbb{N}$ -bağlantılı bir manipülatörün hareket denklemi şudur [27, 28]

$$M(q)\ddot{q} + C(q, \dot{q})\dot{q} + G(q) + F\dot{q} = u(t - \tau), \quad (1)$$

burada $q, \dot{q}, \ddot{q} \in \mathbb{R}^n$ eklem konumu, hızı ve ivmesi, $M(q) \in \mathbb{R}^{n \times n}$ atalet matrisi, $C(q, \dot{q}) \in \mathbb{R}^{n \times n}$ merkezci–Coriolis matrisi, $G(q) \in \mathbb{R}^n$ yerçekimi vektörü, $F \in \mathbb{R}^{n \times n}$ viskoz sürtünme matrisi ve $u \in \mathbb{R}^n$ komutlanan denetim girişidir. Denklemin sağ tarafındaki $u(t - \tau)$ yazımı sisteme etkileyen tork, o an komutlanan değil, $\tau \in \mathbb{R}_+$ kadar önce komutlanan değerdir.

Varsayım 1. $M(q)$, $C(q, \dot{q})$, $G(q)$ ve F terimleri, birinci dereceden kısmi türevleriyle birlikte, $q, \dot{q} \in \mathcal{L}_\infty$ olduğu her durumda düzgün sınırlıdır [28].

Özellik 1. Atalet matrisi simetrik ve pozitif tanımlıdır ve tüm $\xi \in \mathbb{R}^n$ için $m_1 \|\xi\|^2 \leq \xi^T M(q) \xi \leq m_2 \|\xi\|^2$ eşitsizliğini sağlar, burada $0 < m_1 \leq m_2 < \infty$ bilinen sabitlerdir [27].

Özellik 2. Atalet ve merkezci–Coriolis matrisleri, tüm $\xi \in \mathbb{R}^n$ için $\xi^T \left(\frac{1}{2} \dot{M}(q) - C(q, \dot{q}) \right) \xi = 0$ ters simetri bağıntısını sağlar [27].

Manipülatörün görev-uzayı kinematiki ise [29]

$$x = f(q), \quad \dot{x} = J(q)\dot{q}, \quad (2)$$

biçimindedir. Burada $x \in \mathbb{R}^n$ uç işlevcinin konumu, $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ ileri kinematik dönüşümü ve $J \triangleq \partial f / \partial q \in \mathbb{R}^{n \times n}$ manipülatör Jacobian’ıdır. Denetleyici doğrudan x üzerinden tasarlanacağından, dinamikler ile kinematik bu iki denklem aracılığıyla birbirine bağlanacaktır.

Varsayım 2. Manipülatör kinematik tekilliklerden uzakta çalışır. Bu nedenle J^{-1} mevcuttur ve $\|J\| \leq \bar{J}$ ile $\|J^{-1}\| \leq \bar{J}_i$ sağlanır. Ayrıca, (2) içindeki kinematik terimler ve bunların birinci dereceden kısmi türevleri, sınırlı argümanlar için düzgün sınırlıdır [25].

Varsayım 3. Referans yörünge $x_d(t) \in \mathbb{R}^n$, $x_d, \dot{x}_d, \ddot{x}_d$ mevcut ve sınırlı olacak, yani $x_d, \dot{x}_d, \ddot{x}_d \in \mathcal{L}_\infty$ olacak şekilde tasarlanır.

2.1. Bulanık Yaklaşımcı

Manipülatör dinamiklerinin içerdiği belirsizliklerden dolayı, denetleyici doğrudan bu terimleri içermeyecektir. Bilinmeyen tüm etkiler, Bölüm 3’te türetilen tek bir $f_d \in \mathbb{R}^n$ fonksiyonunda toplanacaktır. Bu fonksiyonun önemli özelliği, yalnızca ölçülebilen ve bilinen sinyallere, yani $\zeta \triangleq [x^T, \dot{x}^T, x_d^T, \dot{x}_d^T, \ddot{x}_d^T]^T \in \mathbb{R}^{5n}$ vektörüne bağlı olmasıdır. Bulanık sistemlerin evrensel yaklaşım özelliği [24, 25] sayesinde f_d , herhangi bir sınırlı çalışma bölgesi üzerinde

$$f_d = W^T \phi(p, \Delta, \zeta) + \epsilon, \quad \|\epsilon\| \leq \bar{\epsilon}, \quad (3)$$

biçiminde yazılabilir. Burada $W \in \mathbb{R}^{a \times n}$ sabit ideal ağırlık matrisi, $\phi \in \mathbb{R}^a$ üyelik fonksiyonları vektörü, $p \in \mathbb{R}^{5an}$ ve $\Delta \in \mathbb{R}^a$ bu üyelik fonksiyonlarının merkezleri ve genişlikleri, $\epsilon \in \mathbb{R}^n$ ise sonlu sayıda kuralla kaçınılmaz olan yaklaşım hatasıdır. Her kuralın merkezi, girdi vektörü $\zeta \in \mathbb{R}^{5n}$ ile aynı uzayda bir nokta olup p , bu a merkezin yığılmış hâlidir. Genişlik ise kural başına tek bir sayıdır. Buna göre $\frac{\partial \hat{\phi}}{\partial p} \in \mathbb{R}^{a \times 5an}$ ve $\frac{\partial \hat{\phi}}{\partial \Delta} \in \mathbb{R}^{a \times a}$ olur. Denetleyici gerçek W, p, Δ değerlerine erişemez. Onların yerine $\hat{W}, \hat{p}, \hat{\Delta}$ kestirimlerini kullanır ve kendi üyelik vektörünü $\hat{\phi} \triangleq \phi(\hat{p}, \hat{\Delta}, \zeta)$ olarak hesaplar. Merkez ve genişlikler de uyarlandığından, ϕ bu parametrelere doğrusal olmayan biçimde bağlıdır. Analizde bu bağımlılığı ele almak için $(\hat{p}, \hat{\Delta})$ çevresinde birinci dereceden Taylor açılımı kullanılır:

$$\tilde{\phi} \triangleq \phi - \hat{\phi} = \frac{\partial \hat{\phi}}{\partial p} \tilde{p} + \frac{\partial \hat{\phi}}{\partial \Delta} \tilde{\Delta} + \xi, \quad (4)$$

burada ξ yüksek dereceden kalan terim, $\frac{\partial \hat{\phi}}{\partial p}$ ve $\frac{\partial \hat{\phi}}{\partial \Delta}$ ise $\hat{\phi}$ 'nin merkezlere ve genişliklere göre Jacobian'larıdır. Bu iki Jacobian matrisi, ilerideki uyarlama yasalarında merkez ve genişlik güncellemelerini yönlendiren niceliklerdir.

Varsayım 4. Üyelik değerleri ve bunların parametre gradyanları yapıları gereği sınırlıdır: $\|\hat{\phi}\| \leq \bar{\phi}$, $\left\| \frac{\partial \hat{\phi}}{\partial p} \right\| \leq \bar{\phi}_p$, $\left\| \frac{\partial \hat{\phi}}{\partial \Delta} \right\| \leq \bar{\phi}_\Delta$, ve ideal sabitler $\|W\| \leq \bar{W}$, $\|p\| \leq \bar{p}$, $\|\Delta\| \leq \bar{\Delta}$ eşitsizliklerini sağlar. Ayrıca, [24]'daki yaklaşımı izleyerek, sınırlı çalışma bölgesi üzerinde $a_0, \dots, a_4 \in \mathbb{R}_+$ bilinen sabitler seçilebilir ve kalan terimler aşağıdaki gibi sınırlanır:

$$\|r^T W^T \xi\| \leq (a_0 + a_1 \|\tilde{\Delta}\| + a_2 \|\tilde{p}\|) \|r\|, \quad (5)$$

$$\left| \Delta^T \left(\frac{\partial \hat{\phi}}{\partial \Delta} \right)^T \tilde{W} r \right| \leq a_3 \|\tilde{W}\| \|r\|, \quad (6)$$

$$\left| p^T \left(\frac{\partial \hat{\phi}}{\partial p} \right)^T \tilde{W} r \right| \leq a_4 \|\tilde{W}\| \|r\|, \quad (7)$$

burada $r \in \mathbb{R}^n$, Bölüm 3'te tanıtılan yardımcı hatadır. Projeksiyon operatörü $\tilde{W}, \hat{\Delta}, \hat{p}$ 'yi bilinen sınırlı bölgeler içinde tutar. Bu durum, sabit ideallerle birlikte (5)–(7)'yi iyi tanımlı kılar.

3. Hata Sistemi ve Denetleyici Tasarımı

Bu bölümde hata sinyalleri ve denetleyici tasarlanmaktadır. Tasarımın iki pratik özelliği baştan belirtilmelidir: her örnekleme anında ters kinematik çözülmemekte ve manipülatör dinamiklerinin bilinmesi gerekmemektedir. Denetleyicinin kullandığı bilgiler yalnızca ölçülen eklem konumu ve hızı, bilinen kinematik ve bilinen gecikmedir. Tasarımın üç hedefi vardır.

1. *İzleme hedefi*: uç işlevcinin konumu x , gecikmeye rağmen görev-uzayı referansı x_d 'yi izlemelidir.
2. *Gecikme telafisi hedefi*: gecikmeli denetim girişi, gecikmenin bilinmesinden yararlanılarak açık çevrim hata sisteminde çıkarılmalı. Böylece analiz gecikmesiz bir hata sistemi üzerinde yürütülebilmelidir.
3. *Yaklaşım hedefi*: bilinmeyen dinamikleri toplayan f_d fonksiyonu, kalıcı uyarım koşulu aranmadan, kendi kendini uyarlayan bulanık sistem tarafından çevrimiçi öğrenilmelidir.

3.1. Hata Sinyallerinin Tasarımı

İzleme hedefini niceleyen görev-uzayı hatası $e \in \mathbb{R}^n$

$$e \triangleq x_d - x. \quad (8)$$

biçiminde tanımlanır. Tasarımın çekirdeği ise yardımcı (filtrelenmiş) hata sinyalıdır:

$$r \triangleq J^{-1}(\dot{x}_d + ae) - \dot{q} - e_u - u, \quad (9)$$

Bu tanım şu şekilde yorumlanabilir: $J^{-1}(\dot{x}_d + ae)$ terimi, hatayı a oranında sönümleyecek *istenen* eklem hızıdır. Ondan gerçek hız $\dot{q}$ çıkarılır. Kalan iki terim, e_u ve u , gecikme telafisi için eklenmiştir. Burada e_u (11)'de tanımlanacak, u ise (17) ile üretilen denetim giriştir. Bu tanımlarda $e, r, e_u, u \in \mathbb{R}^n$ olup $\alpha = \text{diag}(\alpha_1, \dots, \alpha_n)$ köşegen elemanları pozitif bir kazanç matrisidir. Yardımcı hata tanımına (9)'da soldan J ile çarpılıp (2) ve (8) kullanılırsa, izleme hatasının kapalı çevrim dinamiği

$$\dot{e} = -ae + Jr + Je_u + Ju. \quad (10)$$

olarak elde edilir. Buradaki ilk terim istenen üstel sönümü sağlar. Kalan üç terim ise r, e_u ve u küçük tutulduğu ölçüde küçük kalır. Analizin geri kalanı, tam olarak bu üç niceliği sınırlamaya ayrılmıştır.

Gecikmenin bilinmesi, telafi teriminin *gerçek* geçmiş denetim değerleriyle kurulmasına izin verir:

$$e_u \triangleq \int_{t-\tau}^t e^{\alpha_u(\xi-t)} u(\xi) d\xi, \quad (11)$$

burada $\alpha_u \in \mathbb{R}_+$ telafi kazancıdır. Durum e_u , son τ saniyede uygulanan girişlerin üstel ağırlıklı bir ortalamasıdır. Üstel ağırlık, uzak geçmişin etkisini azaltır. Leibniz kuralıyla türevi alınır

$$\dot{e}_u = -\alpha_u e_u + u - e^{-\alpha_u \tau} u(t - \tau). \quad (12)$$

elde edilir. Bu ifadenin işlevi şudur: sağ taraftaki $-e^{-\alpha_u \tau} u(t - \tau)$ terimi, sistemin (1)'deki gecikmeli girişiyle birleşecek ve gecikmeli girişin hata sistemindeki etkisini bilinen, sınırlı tek bir terime indirgeyecektir. Başka bir deyişle e_u , gecikmeli giriş hata denklemlerinden temizlemek için özel olarak tasarlanmış bir durumdur.

3.2. Açık Çevrim Hata Dinamiği

Yardımcı hata tanımı (9)'un zamana göre türevi M ile çarpılır ve (1)'deki dinamikler ve (12)'deki telafi türevi yerine konur. Özellik 2'deki ters simetri kullanılırsa r 'nin açık çevrim dinamiği

$$M\dot{r} = f_d - Cr + \Pi - Mu - M\dot{u} - \varepsilon_1 u(t - \tau), \quad (13)$$

olur. Burada üç yeni nicelik tanımlanmıştır. Birincisi, denetleyicinin bilmediği her şeyi tek bir fonksiyonda toplayan

$$f_d \triangleq M \frac{d}{dt} \{J^{-1}(\dot{x}_d + ae)\} + C J^{-1}(\dot{x}_d + ae) + G + F\dot{q}, \quad (14)$$

terimidir. Bu, (3)'te bulanık ağ tarafından yaklaşırlacak olan fonksiyondur ve yalnızca ölçülebilir ζ sinyaline bağlıdır. İkincisi, telafi terimi ile Coriolis matrisinin etkileşiminden doğan

$$\Pi \triangleq M\alpha_u e_u - C e_u - C u, \quad (15)$$

yardımcı terimdir. Üçüncüsü ise gecikmeli girişin önündeki

$$\varepsilon_1 \triangleq I_n - e^{-\alpha_u \tau} M. \quad (16)$$

ağırlık matrisidir. Varsayım 1 ve Özellik 1 gereğince bu terimler sınırlı çalışma bölgesinde sınırlıdır: $\|\Pi\| \leq (m_2 \alpha_u + \bar{C}_D) \|e_u\| + \bar{C}_D \|u\|$, burada $\bar{C}_D$ Coriolis matrisi üzerindeki

norm sınırdır. Matris ε_1 simetrik olduğundan $\|\varepsilon_1\| \leq \bar{\varepsilon}_1 \triangleq \max\{|1 - e^{-a_u \tau} m_1|, |1 - e^{-a_u \tau} m_2|\}$ yazılabilir. Bu özdeğer, $\bar{\varepsilon}_1$ nihai sınıra karesel girdiğinden, daha kaba olan $1 + e^{-a_u \tau} m_2$ sınırına tercih edilir.

Açıklama 1. Açık çevrim hata dinamiği (13)'ün kilit özelliği, içinde gecikme uyumsuzluğuna bağlı hiçbir terim bulunmamasıdır. Telif terimi (11) gerçek gecikmeyi kullanarak kurulduğundan, gecikmeli girişin hata sistemine katkısı tek bir bilinen ağırlıkla temsil edilir ve ε_1 katsayısı tamamen bilinmektedir. Bu durumda (13)'te kalan tek gecikmeli nicelik, Bölüm 4'te LK fonksiyoneliyle ele alınacak olan $\varepsilon_1 u(t - \tau)$ terimidir.

3.3. Denetim Yasası ve Uyarlama Yasalarının Tasarımı

Bölüm 4'teki analize dayanarak, denetim girişi doğrudan bir cebirsel ifadeyle değil, filtrelenmiş bir yasayla üretilir:

$$\dot{u} = -(k_u + I_n)u + kr + e^{-a_u \tau} \hat{W}^T \hat{\phi}, \quad (17)$$

burada $k = \text{diag}(k_1, \dots, k_n)$ ve $k_u = \text{diag}(k_{u,1}, \dots, k_{u,n})$ köşegen elemanları pozitif kazanç matrisleridir. Köşegen yapı her görev-uzay eksenine ayrı bir kazanç atanmasına izin verir. Eksenler ne atalet ne de gecikme payı bakımından eşdeğer olduğundan bu serbestlik tasarımda anlamlıdır. Girdiler yine de birbirinden bağımsız seçilemez: izin verilen ayrışmayı Bölüm 4'teki $\kappa > 0$ koşulu sınırlamaktadır. Sağ taraftaki üç terimin işlevi ayrıdır: $-(k_u + I_n)u$ girişi kendi başına sönümler, kr geri besleme düzeltilmesidir ve $e^{-a_u \tau} \hat{W}^T \hat{\phi}$ ise bulanık ağırlık ürettiği ileri besleme, yani öğrenilen dinamiklerin karşılanmasıdır. Öndeki $e^{-a_u \tau}$ ağırlığı rastgele seçilmemiştir. Bu ağırlık, (16)'daki gecikme ağırlıklandırma matrisiyle uyumlu olacak biçimde belirlenmiştir ve kapalı çevrimde tam da bu uyum sayesinde bulanık terim ile f_d birbirini götürecektir. Bulanık sistemin üç parametre grubu, sönümlenmesi $\|r\|$ ile ölçeklenen e-modifikasyon yasalarıyla güncellenir. Düzgün projeksiyon operatörü $\text{Proj}\{\cdot\}$ [30] kestirimleri önceden bilinen sınırlı bölgelerde tutar:

$$\begin{aligned} \hat{W} &= \text{Proj}\{\Gamma_W \hat{\phi} r^T - \Gamma_W \frac{\partial \hat{\phi}}{\partial \Delta} \hat{\Delta} r^T \\ &\quad - \Gamma_W \frac{\partial \hat{\phi}}{\partial p} \hat{p} r^T - k_W \|r\| \Gamma_W \hat{W}\}, \end{aligned} \quad (18)$$

$$\hat{\Delta} = \text{Proj}\{\Gamma_{\Delta} (\frac{\partial \hat{\phi}}{\partial \Delta})^T \hat{W} r - k_{\Delta} \|r\| \Gamma_{\Delta} \hat{\Delta}\}, \quad (19)$$

$$\hat{p} = \text{Proj}\{\Gamma_p (\frac{\partial \hat{\phi}}{\partial p})^T \hat{W} r - k_p \|r\| \Gamma_p \hat{p}\}, \quad (20)$$

burada $\Gamma_W, \Gamma_{\Delta} \in \mathbb{R}^{a \times a}$ ve $\Gamma_p \in \mathbb{R}^{5an \times 5an}$ simetrik pozitif tanımlı uyarlama kazanç matrisleri, $k_W, k_{\Delta}, k_p > 0$ ise sönümlenme kazançlarıdır. Ağırlık uyarlama yasası (18)'deki ilk terim standart gradyan güncellemesidir. Onu izleyen iki terim, merkez ve genişliklerin de uyarlanmasından kaynaklanan (4)'teki çapraz etkileri dengelemek için eklenmiştir. Sönümlenme terimleri ise parametrelerin zamanla sürüklenmesini engeller.

3.4. Kapalı Çevrim Hata Dinamiği

Denetim yasası (17) ve bulanık gösterim (3), açık çevrim (13)'te yerine konur. Burada $Me^{-a_u \tau} = I_n - \varepsilon_1$ özdeşliği ile $W^T \phi - \hat{W}^T \hat{\phi} = \hat{W}^T \hat{\phi} + W^T \tilde{\phi}$ ayrıştırması kullanıldığında kapalı çevrim

$$\begin{aligned} M\dot{r} &= -Cr - Mkr + \hat{W}^T \hat{\phi} + W^T \tilde{\phi} + \epsilon + Mk_u u \\ &\quad + \varepsilon_1 \hat{W}^T \hat{\phi} + \Pi - \varepsilon_1 u(t - \tau). \end{aligned} \quad (21)$$

olarak bulunur. Bu denklem, nihai hatanın üç kaynağını açıkça göstermektedir: bulanık yaklaşımın kendi hatası $\hat{W}^T \hat{\phi} + W^T \tilde{\phi} + \epsilon$, telafiden artakalan $\varepsilon_1 \hat{W}^T \hat{\phi}$ terimi ve gecikmeli giriş $\varepsilon_1 u(t - \tau)$. Son iki terim, bir sonraki bölümde kurulacak LK yapısıyla sınırlandırılacaktır.

4. Kararlılık Analizi

Kararlılık analizinde üç hata sinyali tek bir vektörde toplanır:

$$z \triangleq [e^T, r^T, u^T]^T \in \mathbb{R}^{3n}. \quad (22)$$

Burada $\underline{a}$, $\underline{k}$ ve $\underline{k}_u$ sırasıyla a , k ve k_u 'nın en küçük özdeğerleridir. Denetim kazancı hata dinamiğine Mk çarpımı üzerinden girdiğinden, aşağıdaki skaler nicelik tanımlanır:

$$\kappa \triangleq \lambda_{\min}\left(\frac{1}{2}(Mk + kM)\right) > 0. \quad (23)$$

Bu tanımın nedeni şudur: analizde $r^T Mkr$ biçiminde bir terim ortaya çıkar ve bir karesel form yalnızca kendi bakışımı parçasını görür,

$$r^T Mkr = r^T \frac{1}{2}(Mk + kM)r \geq \kappa \|r\|^2. \quad (24)$$

Burada M ve k ayrı ayrı simetrik olsa da çarpımları genel olarak simetrik değildir. Bu nedenle gecikmesiz tasarımlarda sıkça kullanılan $\kappa \geq m_1 \underline{k}$ eşitsizliği burada her zaman geçerli değildir.

Kazanc matrisi ölçekli birim seçildiğinde, yani $k = k_0 I_n$ için çarpım simetrik olur ve

$$\kappa = k_0 m_1 \quad (25)$$

bulunur. Kazançla doğrusal ölçekleme korunur. Köşegen ama eşit olmayan girdilerde pozitiflik kendiliğinden gelmez. Kazanc

$$k = \underline{k} I_n + \Delta k, \quad \|\Delta k\| = \lambda_{\max}(k) - \underline{k} \quad (26)$$

biçiminde ayrıştırılırsa

$$\kappa \geq m_1 \underline{k} - m_2 \|\Delta k\| \quad (27)$$

yazılır ve buradan $\kappa > 0$ için yeterli koşul

$$\frac{\lambda_{\max}(k) - \underline{k}}{\underline{k}} < \frac{m_1}{m_2} \quad (28)$$

elde edilir: köşegen girdiler arasındaki bağıl yayılım, atalet matrisinin koşul sayısının tersiyle sınırlıdır. Bu koşul yalnızca yeterlidir.

Analizde şu bileşik Lyapunov–Krasovskii fonksiyoneli kullanılır:

$$V \triangleq V_z + V_b + Q_1 + Q_2, \quad (29)$$

burada hata sinyalleri, bulanık parametre hataları ve gecikme penceresine ilişkin terimler sırasıyla

$$V_z \triangleq \frac{1}{2} e^T e + \frac{\gamma_2}{2} u^T u + \frac{1}{2} r^T M r, \quad (30)$$

$$V_b \triangleq \frac{1}{2} \text{tr}\{\tilde{W}^T \Gamma_W^{-1} \tilde{W}\} + \frac{1}{2} \tilde{\Delta}^T \Gamma_{\Delta}^{-1} \tilde{\Delta} + \frac{1}{2} \tilde{p}^T \Gamma_p^{-1} \tilde{p}, \quad (31)$$

$$Q_1 \triangleq c_1 \int_{t-\tau}^t \|u(\theta)\|^2 d\theta, \quad (32)$$

$$Q_2 \triangleq c_2 \int_{t-\tau}^t \int_s^t \|u(\theta)\|^2 d\theta ds \quad (33)$$

olarak tanımlanır ve $\gamma_2, c_1, c_2 \in \mathbb{R}_+$ sabitlerdir. Fonksiyonelin üç işlevi vardır. Denklem (30) hata vektörü z 'nin karesel büyüklüğünü ölçer. Denklem (31) bulanık parametre hatalarının büyüklüğünü ölçer. Türevi alındığında uyarlama yasaları ortaya çıkar. Denklem (32) ile (33) Lyapunov–Krasovskii terimleridir. Bu iki terim, $[t - \tau, t]$ aralığında

uygulanan girişin karesinin integralini tutar. Amaçları, (21)'de kalan $\varepsilon_1 u(t - \tau)$ gecikmeli terimini karşılamaktır. Lyapunov-Krasovskii terimleri $Q_1, Q_2 \geq 0$ olduğundan ve parametre hataları projeksiyon nedeniyle sınırlı kaldığından

$$\lambda_L \|z\|^2 \leq V \leq \lambda_U \|z\|^2 + \varpi, \quad (34)$$

$$\|z_t\| \triangleq \sup_{\theta \in [t-\tau, t]} \|z(\theta)\|, \quad (35)$$

yazılabilir. Skaler sınır sabitleri

$$\lambda_L \triangleq \frac{1}{2} \min\{1, \gamma_2, m_1\}, \quad (36)$$

$$\lambda_U \triangleq \frac{1}{2} \max\{1, \gamma_2, m_2\} + c_1 \tau + \frac{1}{2} c_2 \tau^2 \quad (37)$$

biçimindedir. Üst sınırdaki anlık norm yerine gecikme penceresi üzerindeki üstünlük normunun kullanılması zorunludur: Q_1 ve Q_2 , anlık $z(t)$ 'nin değil, u 'nun geçmişinin fonksiyonelleridir ve

$$\begin{aligned} Q_1 &\leq c_1 \tau \sup_{\theta \in [t-\tau, t]} \|u(\theta)\|^2, \\ Q_2 &\leq \frac{1}{2} c_2 \tau^2 \sup_{\theta \in [t-\tau, t]} \|u(\theta)\|^2 \end{aligned} \quad (38)$$

sağlanır. Son olarak $\varpi \in \mathbb{R}_+ \cup \{0\}$, projeksiyon kümesinde parametre hatalarını sınırlayan sabittir.

Kendi kendini uyarlayan bulanık mantık tabanlı sistemde uyarlanan üç parametre grubu $\dot{V}$ 'de çok sayıda terim üretir. Aşağıdaki lemma bu terimleri toparlar ve hepsinin, kazançlarla ayarlanabilen tek bir γ sabiti ile $\gamma \|r\|$ biçiminde sınırlanabileceğini gösterir. Bu, standart kendi kendini uyarlayan bulanık mantık tabanlı sisteminin kararlılık analizini kolaylaştırmak için gerekli bir özetir.

Lemma 1. Varsayım 4 altında, uyarlama yasaları (18)–(20) kullanıldığında, $\dot{V}$ içinde (31) ile tanımlanan parametre hatalarının türevi ile bulanık yaklaşım hatasının r^T ile çarpılan katkısı birlikte

$$\dot{V}_b + r^T (\tilde{W}^T \hat{\Phi} + W^T \tilde{\Phi} + \epsilon) \leq \gamma \|r\| \quad (39)$$

eşitsizliğini sağlar; burada

$$\begin{aligned} \gamma &\triangleq a_0 + \bar{\epsilon} + \frac{(a_3 + a_4 + k_W \bar{W})^2}{4k_W} \\ &\quad + \frac{(a_1 + k_\Delta \bar{\Delta})^2}{4k_\Delta} + \frac{(a_2 + k_p \bar{p})^2}{4k_p}. \end{aligned} \quad (40)$$

Skaler sabit $\gamma \in \mathbb{R}_+$, sönümleme kazançları üzerinden $k_W = (a_3 + a_4)/\bar{W}$, $k_\Delta = a_1/\bar{\Delta}$, $k_p = a_2/\bar{p}$ seçiminde en küçük değerini alır. Bu seçim $\gamma_{\min} = a_0 + \bar{\epsilon} + (a_3 + a_4)\bar{W} + a_1\bar{\Delta} + a_2\bar{p}$ 'yi verir.

İspat. İdeal ağırlık ve üyelik parametreleri sabit olduğundan parametre hatalarının türevleri

$$\dot{\tilde{W}} = -\dot{\hat{W}}, \quad \dot{\tilde{\Delta}} = -\dot{\hat{\Delta}}, \quad \dot{\tilde{p}} = -\dot{\hat{p}} \quad (41)$$

biçimindedir. Denklem (31) ile tanımlanan V_b 'nin türevi alınıp (41) kullanıldığında

$$\dot{V}_b = -\text{tr}\{\tilde{W}^T \Gamma_W^{-1} \dot{\tilde{W}}\} - \tilde{\Delta}^T \Gamma_\Delta^{-1} \dot{\tilde{\Delta}} - \tilde{p}^T \Gamma_p^{-1} \dot{\tilde{p}} \quad (42)$$

elde edilir. Kapalı çevrim dinamiği (21), $r^T M \dot{r}$ üzerinden $\dot{V}_z$ 'ye bulanık yaklaşım hatasını taşır; bulanık gösterim (3) ile bu katkı

$$r^T (\tilde{W}^T \hat{\Phi} + W^T \tilde{\Phi} + \epsilon) \quad (43)$$

biçimindedir. Denklem (42) ile (43) toplandığında, $\dot{V}$ içindeki bulanık yaklaşım ve uyarlamadan doğan bütün terimler

$$\begin{aligned} \dot{V}_b + r^T (\tilde{W}^T \hat{\Phi} + W^T \tilde{\Phi} + \epsilon) &= r^T (\tilde{W}^T \hat{\Phi} + W^T \tilde{\Phi} + \epsilon) \\ &\quad - \text{tr}\{\tilde{W}^T \Gamma_W^{-1} \dot{\tilde{W}}\} - \tilde{\Delta}^T \Gamma_\Delta^{-1} \dot{\tilde{\Delta}} - \tilde{p}^T \Gamma_p^{-1} \dot{\tilde{p}} \end{aligned} \quad (44)$$

toplamında birikir. Bu toplam, (39) ile verilen eşitsizliğin sol tarafıdır. Birinci dereceden açılım (4), $W^T \tilde{\Phi}$ yerine konur:

$$W^T \tilde{\Phi} = W^T \left(\frac{\partial \hat{\Phi}}{\partial \Delta} \tilde{\Delta} + \frac{\partial \hat{\Phi}}{\partial p} \tilde{p} \right) + W^T \xi, \quad (45)$$

burada $\xi \in \mathbb{R}^n$ birinci dereceden açılımdan artan yüksek dereceli terimlerdir. Ardından uyarlama yasaları (18)–(20), (44)'e yerleştirilir; projeksiyon operatörü kendi katkısını negatif tuttuğundan yalnızca yasaların sönümlemesiz kısmı hesaba katılır. Gradyan güncellemesinden gelen terim (44)'ün ilk terimiyle birebir sadeleşir:

$$-\text{tr}\{\tilde{W}^T \Gamma_W^{-1} \Gamma_W \hat{\Phi} r^T\} + r^T \tilde{W}^T \hat{\Phi} = 0. \quad (46)$$

Aynı biçimde ağırlık yasasındaki düzeltme terimleri, (45)'in $\tilde{\Delta}$ ve $\tilde{p}$ içeren kısımlarını yok eder:

$$\begin{aligned} -\text{tr}\{\tilde{W}^T \frac{\partial \hat{\Phi}}{\partial \Delta} \hat{\Delta} r^T\} - \text{tr}\{\tilde{W}^T \frac{\partial \hat{\Phi}}{\partial p} \hat{p} r^T\} \\ + r^T W^T \left(\frac{\partial \hat{\Phi}}{\partial \Delta} \tilde{\Delta} + \frac{\partial \hat{\Phi}}{\partial p} \tilde{p} \right) = 0. \end{aligned} \quad (47)$$

Sadeleştirmelerden sonra bu toplam, yalnızca ideal değerleri ve sönümlemeleri taşıyan

$$\begin{aligned} r^T W^T \xi + r^T \epsilon + \Delta^T \left(\frac{\partial \hat{\Phi}}{\partial \Delta} \right)^T \tilde{W} r + p^T \left(\frac{\partial \hat{\Phi}}{\partial p} \right)^T \tilde{W} r \\ + k_W \|r\| \text{tr}\{\tilde{W}^T \tilde{W}\} + k_\Delta \|r\| \tilde{\Delta}^T \tilde{\Delta} + k_p \|r\| \tilde{p}^T \tilde{p} \end{aligned} \quad (48)$$

ifadesine indirgenir. Kalan terimler tek tek sınırlanır. Yüksek dereceli terimlerin sınırı (5) ve $\|\epsilon\| \leq \bar{\epsilon}$ ile

$$\begin{aligned} |r^T W^T \xi + r^T \epsilon| \\ \leq (a_0 + \bar{\epsilon} + a_1 \|\tilde{\Delta}\| + a_2 \|\tilde{p}\|) \|r\|, \end{aligned} \quad (49)$$

türev sınırları (6) ve (7) ile

$$\begin{aligned} \left| \Delta^T \left(\frac{\partial \hat{\Phi}}{\partial \Delta} \right)^T \tilde{W} r + p^T \left(\frac{\partial \hat{\Phi}}{\partial p} \right)^T \tilde{W} r \right| \\ \leq (a_3 + a_4) \|\tilde{W}\| \|r\| \end{aligned} \quad (50)$$

elde edilir. Sönümleme terimleri için $\hat{\bullet} = \bullet - \bar{\bullet}$ konulduğunda ağırlık sönümlemesi

$$\begin{aligned} k_W \|r\| \text{tr}\{\tilde{W}^T \tilde{W}\} &= k_W \|r\| \left(\text{tr}\{\tilde{W}^T W\} - \|\tilde{W}\|^2 \right) \\ &\leq k_W \|r\| \left(\bar{W} \|\tilde{W}\| - \|\tilde{W}\|^2 \right) \end{aligned} \quad (51)$$

biçimini alır. Aynı sönümleme yapısı $\tilde{\Delta}$ ve $\tilde{p}$ sönümlemeleri için de geçerlidir. Eşitsizlikler (49)–(51) bir araya getirilip ortak $\|r\|$ çarpanı dışarı alındığında, $\bullet \in \{W, \Delta, p\}$ için her parametre grubu

$$\begin{aligned} -k_\bullet \|\bullet\|^2 + (c_\bullet + k_\bullet \bar{\bullet}) \|\bullet\|, \\ c_W = a_3 + a_4, \quad c_\Delta = a_1, \quad c_p = a_2 \end{aligned} \quad (52)$$

ifadesine indirgenir. İfade (52), $\|\bullet\|$ değişkeninde aşağı doğru açılan bir parabol olduğundan parametre hatasından bağımsız bir üst sınıra sahiptir:

$$\max_{\|\bullet\| \geq 0} \{-k_\bullet \|\bullet\|^2 + (c_\bullet + k_\bullet \bar{\bullet}) \|\bullet\|\} = \frac{(c_\bullet + k_\bullet \bar{\bullet})^2}{4k_\bullet}. \quad (53)$$

Üç grubun (53) değerleri $a_0 + \bar{\epsilon}$ ile toplandığında (39) ve (40) elde edilir. Son olarak (40) içindeki her parantez $k_\bullet$ üzerinden en küçükleir. Türevin sıfırlanması, yani $\partial[(c_\bullet + k_\bullet \bar{\bullet})^2 / (4k_\bullet)] / \partial k_\bullet = 0$ koşulundan

$$k_W = \frac{a_3 + a_4}{\bar{W}}, \quad k_\Delta = \frac{a_1}{\bar{\Delta}}, \quad k_p = \frac{a_2}{\bar{p}} \quad (54)$$

bulunur ve bu seçim $\gamma_{\min}$ değerini verir. □

Teorem 1. Bilinen sabit giriş gecikmesi altında (1) ile verilen dinamikler ve Varsayımlar 1–4 göz önüne alınsın. Bu durumda (17) denetleyicisi, (11) telafi terimi ve (18)–(20) uyarlama yasaları, tüm kapalı çevrim sinyallerinin sınırlı kalmasını ve görev-uzayı izleme hatasının

$$\limsup_{t \rightarrow \infty} \|e(t)\| \leq \sqrt{\frac{1}{\lambda_L} \left(\frac{\lambda_{UC}}{\lambda} + \varpi \right)} \quad (55)$$

anlamında yarı-global düzgün nihai sınırlı olmasını sağlar; bunun için Lyapunov–Krasovskii kazançlarının

$$c_1 \geq \frac{3\bar{e}_1^2}{\kappa}, \quad c_2 \geq \left[\frac{3\bar{f}^2}{2\alpha} + \frac{3(m_2\alpha_u + \bar{c}_D)^2}{\kappa} \right] \alpha_u^{-1} \quad (56)$$

koşullarını, denetim kazançlarının ise başlangıç koşullarına göre yeterince büyük seçilerek

$$\kappa > \frac{3\bar{f}^2}{\alpha}, \quad (57)$$

$$\frac{3}{4}\gamma_2(\underline{k}_u + 1) > \frac{3\bar{f}^2}{2\alpha} + c_1 + c_2\tau + \frac{3[(\gamma_2\bar{k})^2 + (m_2\bar{k}_u)^2 + \bar{c}_D^2]}{\kappa} \quad (58)$$

yeterli koşullarını sağlaması gerekir. Burada $\bar{k} \triangleq \|k\|$, $\bar{k}_u \triangleq \|k_u\|$, sınır sabiti

$$\varsigma \triangleq \frac{3\Gamma_r^2}{\kappa} + \frac{\gamma_2\bar{W}^2\bar{\phi}^2}{\underline{k}_u + 1}, \Gamma_r \triangleq \gamma + \bar{e}_1\bar{W}\bar{\phi} \quad (59)$$

ve yakınsama oranı

$$\begin{aligned} \lambda &\triangleq \min\{\lambda_e, \lambda_r, \lambda_u\}, & \lambda_e &\triangleq \frac{\alpha}{2}, & \lambda_r &\triangleq \frac{\kappa}{2} - \frac{3\bar{f}^2}{2\alpha}, \\ \lambda_u &\triangleq \frac{3}{4}\gamma_2(\underline{k}_u + 1) - \frac{3\bar{f}^2}{2\alpha} - c_1 - c_2\tau \\ &\quad - \frac{3[(\gamma_2\bar{k})^2 + (m_2\bar{k}_u)^2 + \bar{c}_D^2]}{\kappa} \end{aligned} \quad (60)$$

olarak tanımlıdır; (57) ile (58) koşulları, sırasıyla $\lambda_r > 0$ ve $\lambda_u > 0$ koşullarına denktir.

İspatta kullanılan Young ağırlıkları serbest sabit değildir. Her biri negatif tanımlı terimlerdeki karesel payların sabit bir kesri olarak seçilmiştir: çapraz terimler $\|e\|^2$ ve $\|r\|^2$ paylarının yarısını, ileri besleme terimi Ξ_u ise kendi payının dörtte birini tüketir. Bu nedenle (56)–(58) yalnızca kazançlar, gecikme ve model sınırları ($\bar{f}$, m_2 , $\bar{c}_D$, $\bar{e}_1$) cinsindendir.

Nihai sınırdaki ϖ terimi (31)’deki parametre hatalarından gelir. Projeksiyon operatörü bu hataları sınırlı tutar ancak sıfırlamaz. Bu nedenle (55) kazançlar büyütülerek keyfi biçimde küçültülemez. Buna karşılık kazanç ölçeklemesi ortalama-kare anlamında korunur. Bunu görmek için (74) ifadesi $[0, T]$ aralığında integrali alınır:

$$V(T) - V(0) \leq -\lambda \int_0^T \|z\|^2 dt + \varsigma T. \quad (61)$$

Fonksiyonelin negatif olmaması, yani $V(T) \geq 0$ kullanılarak integral terimi yalnız bırakılır:

$$\lambda \int_0^T \|z\|^2 dt \leq V(0) + \varsigma T. \quad (62)$$

Her iki taraf λT ile bölündüğünde ortalama-kare değeri için

$$\frac{1}{T} \int_0^T \|z\|^2 dt \leq \frac{\varsigma}{\lambda} + \frac{V(0)}{\lambda T} \quad (63)$$

elde edilir. Sağdaki ikinci terim $T \rightarrow \infty$ için sıfıra gittiğinden

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T \|z\|^2 dt \leq \frac{\varsigma}{\lambda} = \mathcal{O}(1/\kappa) \quad (64)$$

bulunur. Ortalama-kare sınırı ϖ içermez. Dolayısıyla kazançlar büyütüldükçe ς ile birlikte küçülür.

İspat. (29) ile tanımlanan fonksiyonelin türevi alınır. Denklem (30)’un türevinde açık çevrim dinamiği (21), hata dinamiği (10) ve Özellik 2 yerine konur; ters simetri sayesinde

$$\frac{1}{2}r^T \dot{M}r - r^T Cr = 0 \quad (65)$$

sağlandığından Coriolis terimi düşer. Bulanık yaklaşım ve uyarlamadan doğan terimler Lemma 1 ile $\gamma\|r\|$ içinde toplandığından, türevin Lyapunov–Krasovskii dışı kısmı

$$\begin{aligned} \dot{V}_1 &\leq -e^T a e - \gamma_2 u^T (k_u + I_n) u - r^T M k r + \Gamma_r \|r\| \\ &\quad + \Xi_x + \Xi_{e_u} + \Xi_d + \Xi_u \end{aligned} \quad (66)$$

biçiminde sınırlanır. İlk üç terim negatif tanımlı terimlerdir. Kalan çapraz terimler

$$\begin{aligned} \Xi_x &\triangleq e^T J r + e^T J u + \gamma_2 u^T k r + r^T M k u, \\ \Xi_{e_u} &\triangleq e^T J e_u + r^T \Pi, & \Xi_d &\triangleq -r^T \varepsilon_1 u(t - \tau), \\ \Xi_u &\triangleq \gamma_2 e^{-\alpha_u \tau} u^T \hat{W}^T \hat{\phi} \end{aligned} \quad (67)$$

olarak tanımlanır; bunlar skaler çapraz terimlerdir, yani $\Xi_x, \Xi_{e_u}, \Xi_d, \Xi_u \in \mathbb{R}$. Skaler Γ_r , bulanık sabit γ ile $r^T \varepsilon_1 \hat{W}^T \hat{\phi}$ teriminden gelen $\bar{e}_1 \bar{W} \bar{\phi}$ sınırını birleştirir. Geçmişe bağlı terimler ele alınır. Leibniz kuralı ile (32) ve (33)’ün türevleri

$$\begin{aligned} \dot{Q}_1 &= c_1 (\|u\|^2 - \|u(t - \tau)\|^2), \\ \dot{Q}_2 &= c_2 \left(\tau \|u\|^2 - \int_{t-\tau}^t \|u(\theta)\|^2 d\theta \right) \end{aligned} \quad (68)$$

biçimindedir. Durum çapraz terimlerine ve gecikmeli terime Young eşitsizliği, ağırlıklar serbest sabitler yerine negatif tanımlı terimlerdeki karesel payların sabit kesirleri olacak biçimde uygulanır: $\|e\|^2$ payından ayrılan her parça $\alpha/6$, $\|r\|^2$ payından ayrılan her parça $\kappa/12$ seçilir ve $r^T \Pi$, $\|\Pi\|$ üzerindeki üst sınır ile iki parçaya ayrılır. Böylece

$$\begin{aligned} e^T J r &\leq \frac{\alpha}{6} \|e\|^2 + \frac{3\bar{f}^2}{2\alpha} \|r\|^2, \\ e^T J u &\leq \frac{\alpha}{6} \|e\|^2 + \frac{3\bar{f}^2}{2\alpha} \|u\|^2, \\ e^T J e_u &\leq \frac{\alpha}{6} \|e\|^2 + \frac{3\bar{f}^2}{2\alpha} \|e_u\|^2, \\ \gamma_2 u^T k r &\leq \frac{\kappa}{12} \|r\|^2 + \frac{3(\gamma_2 \bar{k})^2}{\kappa} \|u\|^2, \\ r^T M k u &\leq \frac{\kappa}{12} \|r\|^2 + \frac{3(m_2 \bar{k}_u)^2}{\kappa} \|u\|^2, \\ r^T \Pi &\leq \frac{\kappa}{6} \|r\|^2 + \frac{3(m_2 \alpha_u + \bar{c}_D)^2}{\kappa} \|e_u\|^2 \\ &\quad + \frac{3\bar{c}_D^2}{\kappa} \|u\|^2, \\ \Xi_d &\leq \frac{\kappa}{12} \|r\|^2 + \frac{3\bar{e}_1^2}{\kappa} \|u(t - \tau)\|^2 \end{aligned} \quad (69)$$

elde edilir. Doğrusal terim ile Ξ_u için aynı kural ve $e^{-\alpha_u \tau} \leq 1$ kullanıldığında

$$\begin{aligned} \Gamma_r \|r\| &\leq \frac{\kappa}{12} \|r\|^2 + \frac{3\Gamma_r^2}{\kappa}, \\ \Xi_u &\leq \frac{1}{4}\gamma_2(\underline{k}_u + 1) \|u\|^2 + \frac{\gamma_2 \bar{W}^2 \bar{\phi}^2}{\underline{k}_u + 1} \end{aligned} \quad (70)$$

elde edilir. Sağdaki iki sabit toplanarak (59)’daki ς oluşur. Geçmişe bağlı terimler için (11) ile verilen ifadeye Cauchy–Schwarz eşitsizliği uygulanır:

$$\|e_u\|^2 \leq \alpha_u^{-1} \int_{t-\tau}^t \|u(\theta)\|^2 d\theta. \quad (71)$$

Bu bağıntı ve (56) seçimi ile, (69)’daki $\|u(t - \tau)\|^2$ ve $\|e_u\|^2$ katsayıları (68)’deki negatif karşılıkları tarafından soğurulur:

$$\begin{aligned} \frac{3\bar{e}_1^2}{\kappa} \|u(t - \tau)\|^2 - c_1 \|u(t - \tau)\|^2 &\leq 0, \\ \left[\frac{3\bar{f}^2}{2\alpha} + \frac{3(m_2 \alpha_u + \bar{c}_D)^2}{\kappa} \right] \|e_u\|^2 &\end{aligned}$$

$$-c_2 \int_{t-\tau}^t \|u(\theta)\|^2 d\theta \leq 0. \quad (72)$$

Böylece geçmişe bağlı bütün terimler $\dot{V}$ 'den giderilir ve geriye yalnızca $c_1 + c_2\tau$ ile ölçeklenen $\|u\|^2$ katsayıları kalır. Negatif tanımlı terimler ise

$$-e^T a e - \gamma_2 u^T (k_u + I_n) u - r^T M k r \\ \leq -\alpha \|e\|^2 - \gamma_2 (k_u + 1) \|u\|^2 - \kappa \|r\|^2 \quad (73)$$

biçiminde alt sınırlanır. Katsayılar terim terim toplanır: $\|e\|^2$ payından ayrılan üç parça toplam $\alpha/2$ 'yi, $\|r\|^2$ payından ayrılan altı parça toplam $\kappa/2$ 'yi tüketir; kalan sönümleme katsayıları (60)'taki $\lambda_e, \lambda_r, \lambda_u$ bileşenleridir. Koşullar (57) ile (58) birlikte son iki bileşenin pozitifliğini sağladığından

$$\dot{V} \leq -\lambda \|z\|^2 + \varsigma \quad (74)$$

yazılabilir. Karşılaştırma lemması (74) ile (34)'teki alt sınıra birlikte uygulandığında

$$V(t) \leq V(0) e^{-\frac{\lambda}{\lambda_U} t} + \frac{\lambda_U}{\lambda} \varsigma \quad (75)$$

elde edilir. Alt sınır $\lambda_L \|e\|^2 \leq V$ kullanılıp (75) üst limiti alındığında (55) elde edilir. Sinyallerin sınırlılığı (75)'ten başlayarak izlenir. Önce $V \in \mathcal{L}_\infty$ ve (34) gereğince $z \in \mathcal{L}_\infty$, dolayısıyla $e, r, u \in \mathcal{L}_\infty$ olur. Referans yörünge ve türevleri Varsayım 3 gereğince sınırlı olduğundan $e \in \mathcal{L}_\infty$ ile $x \in \mathcal{L}_\infty$ ve tekliklikten uzak çalışma bölgesinde $q \in \mathcal{L}_\infty$ elde edilir. Buradan J ile J^{-1} sınırlıdır. Telfa durumu (11), $u \in \mathcal{L}_\infty$ ile sürülen kararlı bir birinci mertbe filtre olduğundan $e_u \in \mathcal{L}_\infty$ sağlanır. Filtrelenmiş hata tanımı (9) çözülerek $\dot{q} \in \mathcal{L}_\infty$, dolayısıyla $\dot{x} = J\dot{q} \in \mathcal{L}_\infty$ bulunur. Projeksiyon operatörü $\hat{W}, \hat{\Delta}, \hat{p}$ kestirimlerini sınırlı bölgelerde tuttuğundan $\hat{p}$ ve $\hat{W}^T \hat{p}$ sınırlıdır. Böylece (17)'den $\dot{u} \in \mathcal{L}_\infty$ ve (11)'den $\dot{e}_u \in \mathcal{L}_\infty$ çıkar. Uyarlama yasaları (18)–(20)'nin sağ tarafları sınırlı bileşenlerden oluştuğundan $\hat{W}, \hat{\Delta}, \hat{p} \in \mathcal{L}_\infty$ olur. Son olarak M^{-1} Özellik 1 gereğince sınırlı olduğundan kapalı çevrim dinamiği (21)'den $\dot{r} \in \mathcal{L}_\infty$ ve $\dot{q} \in \mathcal{L}_\infty$ elde edilir. Böylece kapalı çevrimdeki tüm sinyaller sınırlıdır. $\square$

5. Benzetim Çalışması

Önerilen denetleyici, düzlemde çalışan iki bağlantılı döner bir manipülatör üzerinde sayısal olarak sınanmıştır. Bu örnek, ele alınan sistem sınıfını somutlaştırmakta ve Özellik 1 ile 2'nin hangi parametrelerle sağlandığını göstermektedir. Analiz bu özel seçime bağlı değildir. Kısaltmalar $s_2 \triangleq \sin q_2$ ve $c_2 \triangleq \cos q_2$ ile model terimleri şöyledir [27]

$$M(q) = \begin{bmatrix} p_1 + p_2 + 2p_3 c_2 & p_2 + p_3 c_2 \\ p_2 + p_3 c_2 & p_2 \end{bmatrix}, \\ C(q, \dot{q}) = \begin{bmatrix} -p_3 s_2 \dot{q}_2 & -p_3 s_2 (\dot{q}_1 + \dot{q}_2) \\ p_3 s_2 \dot{q}_1 & 0 \end{bmatrix}, \\ F = \begin{bmatrix} f_1 & 0 \\ 0 & f_2 \end{bmatrix}, \quad (76)$$

burada atalet parametreleri $p_1 = m_1 \ell_{c1}^2 + m_2 \ell_1^2 + I_1$, $p_2 = m_2 \ell_{c2}^2 + I_2$ ve $p_3 = m_2 \ell_1 \ell_{c2}$ olup, $m_i \in \mathbb{R}_+$, $\ell_i \in \mathbb{R}_+$, $\ell_{ci} \in \mathbb{R}_+$ ve $I_i \in \mathbb{R}_+$ sırasıyla bağlantı kütleleri, uzunluğu, kütle-merkezi mesafesi ve atalet momentidir. Benzetimde bu parametreler $p_1 = 3.31$, $p_2 = 0.116$, $p_3 = 0.16 \text{ kg m}^2$ ve $f_1 = f_2 = 0.2 \text{ N m s}$ olarak alınmıştır. Bu değerlerle $M(q)$ her konfigürasyonda pozitif tanımlı kalır ve Özellik 1 sağlanır. Özellik 2 ise modelin yapısı gereği zaten geçerlidir. Kol yatay düzlemde çalıştığı için yerçekimi vektörü sıfırdır.

Görev-uzay1 referansı

$x_d = [0.65 + 0.15(\cos 0.1t - 1), 0.25 + 0.05 \sin 0.2t]^T \text{ m}$ olarak seçilmiştir. Kapalı çevrim, gecikmeli giriş için doğrusal aradeğerlemeli sabit adımlı dördüncü mertbe Runge-Kutta yöntemiyle 1 ms adımla sayısal olarak çözülmüş; giriş gecikmesi için $\tau \in \{10, 25, 50, 75, 100, 125, 150, 175\}$ ms değerleri ele alınmıştır. Başlangıçta uç işlevci referansın 10 mm uzağa yerleştirilmiş, tüm denetleyici durumları sıfır alınmıştır: $u(0) = 0$ ve $e_u(0) = 0$. Bulanık sistem 20 kuraldan oluşur. Merkezler $\hat{p}(0)$ yalnızca bilinen referansın ideal izleme yörüngesi üzerinde yay uzunluğuna göre eşit aralıklarla yerleştirilmiş, genişlikler komşu merkezler arası ortalama uzaklığın iki katı olan $\hat{\Delta}(0) = 2.18$ seçilmiş ve çıkış ağırlıkları $\hat{W}(0) = 0$ ile başlatılmıştır. Böylece denetleyiciye hiçbir dinamik model bilgisi verilmemekte, kural merkezlerinin başlangıç koşulu tasarımı zaten bilinen referans yörüngesinden alınmıştır.

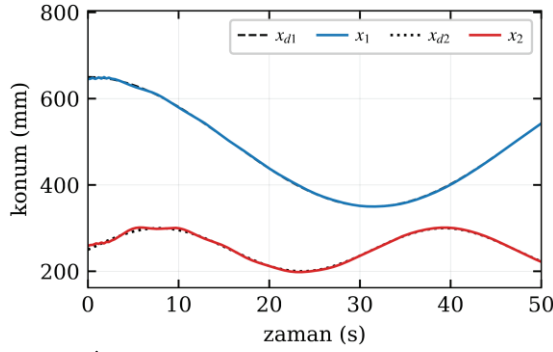
Ortak kazançlar $k_u = \text{diag}(0.02, 0.02)$, $\alpha_u = 0.4$; uyarlama kazançları $\Gamma_W = 300 I_a$, $\Gamma_p = \Gamma_\Delta = 3 \times 10^5 I$; sönümleme kazançları $k_W = 10^{-4}$, $k_p = k_\Delta = 10^{-6}$; projeksiyon kümeleri $\|\hat{W}\|_\infty \leq 2$, $\|\hat{p}\|_\infty \leq 3$, $\hat{\Delta} \in [0.15, 12]^*$ 'dir. Gecikmeye bağlı (k, α) seçimleri Tablo 1'de verilmiştir; her tasarım $|e(0)| \leq 10 \text{ mm}$ başlangıç hatası için hesaplanmıştır.

Tablo 1. Kalıcı rejim izleme hatası (RMS, mm; son referans periyodu, $|e(0)| = 10 \text{ mm}$). Sütun $\Lambda = 0$: aynı kazançlarla bulanık mantık tabanlı ileri besleme kaldırılmış; $e_u \equiv 0$: aynı kazançlarla gecikme telfa filtresi kaldırılmış.

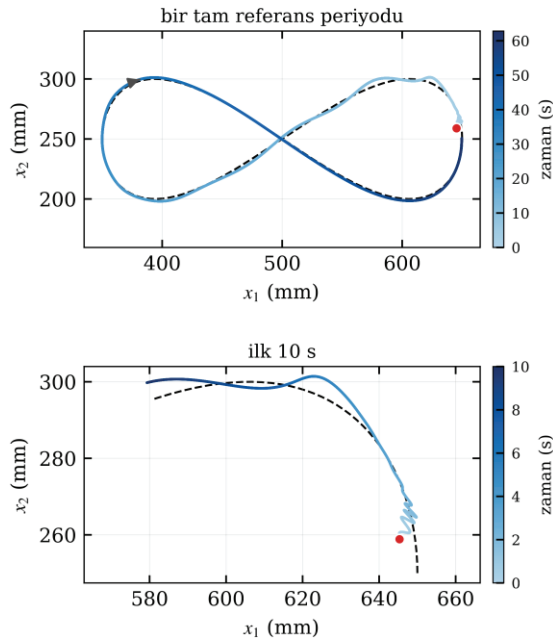
τ (ms)	k	α	önerilen	$\Lambda = 0$	$e_u \equiv 0$
10	(30, 40)	(24, 24)	0.31	0.32	0.31
25	(30, 40)	(16, 16)	0.47	0.49	0.46
50	(30, 30)	(12, 12)	0.64	0.67	ıraksıyor
75	(20, 30)	(8, 8)	1.00	1.05	ıraksıyor
100	(50, 23)	(6, 6)	1.37	ıraksıyor	ıraksıyor
125	(40, 20)	(5, 6.65)	1.38	ıraksıyor	ıraksıyor
150	(30, 20)	(4, 5.32)	1.77	ıraksıyor	ıraksıyor
175	(60, 10)	(3, 3.99)	2.43	ıraksıyor	ıraksıyor

Tablo 1 iki bileşenin rolünü aynı kazançlar altında ayırtmaktadır. Kısa gecikmelerde (10 ve 25 ms) her iki bileşen de kararlılık için gerekli değildir; katkıları gecikme büyüdükçe ortaya çıkar. Gecikme telfa filtresi e_u , 50 ms ve üzerindeki gecikmelerde kararlılığın ön koşuludur: kaldırıldığında çevrim yaklaşık on saniye içinde ıraksamaktadır. Bulanık ileri besleme ise 100 ms ve üzerinde ön koşul haline gelmekte, 50–75 ms'de ise izlemeyi iyileştirmenin yanında ideal ileri beslemenin çevrimiçi kurulmasını sağlamaktadır. Kalıcı rejim giriş enerjisi tüm durumlarda aynıdır (RMS $\approx 0.016 \text{ N m}$); bulanık bileşen ilave bir tork maliyeti getirmemektedir.

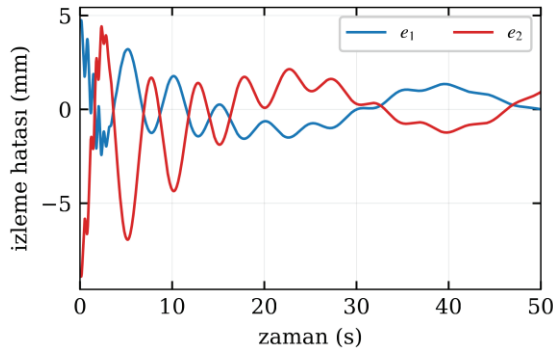
Şekil 1–5, $\tau = 100 \text{ ms}$ için 50 s'lik koşuyu göstermektedir (Şekil 2 bir tam referans periyodunu kapsar). 10 mm'lik başlangıç hatası aşım yapmadan (tepe 10.1 mm) yaklaşık otuz saniyede 1.37 mm etkin değerli kalıcı banda inmekte (Şekil 3), denetim girişi 0.080 N m tepe değerinden düzgün biçimde sönmemektedir (Şekil 4). Sıfır ağırlıkla başlatılan bulanık ağ, ideal ileri beslemeyi ilk birkaç saniyede yakalamakta ve kalıcı rejimde yaklaşık %2 bağıl hatayla yeniden üretmektedir (Şekil 5).



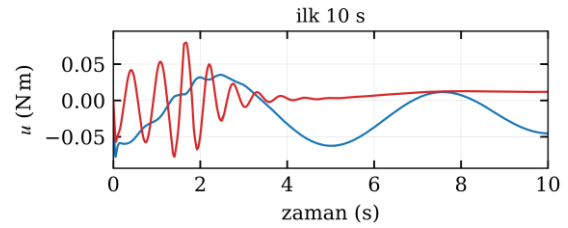
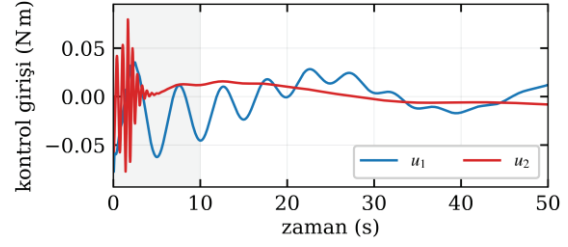
Şekil 1. İstenen ve gerçekleşen uç-işlevci konumu ($\tau = 100$ ms).



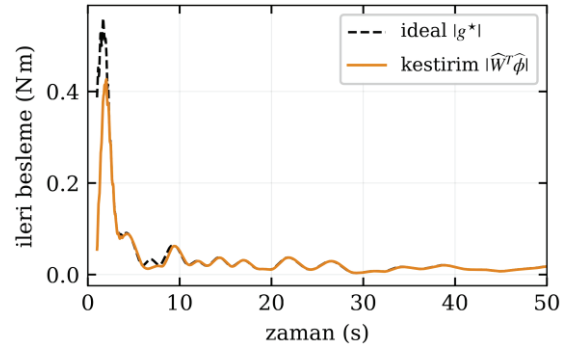
Şekil 2. Görev-uzayı düzleminde yörünge (bir tam referans periyodu); renk zamanı, kırmızı nokta başlangıç konumunu gösterir. Sağ panel, 10 mm'lik başlangıç sapmasının kapandığı ilk 10 s'yi yakınlaştırmaktadır.



Şekil 3. Görev-uzayı izleme hatası bileşenleri.



Şekil 4. Denetim girişleri; alt panel ilk 10 s'yi yakınlaştırmaktadır. Bu penceredeki yüksek frekanslı bileşen uyarlama geçici rejimidir ve kalıcı rejime taşmamaktadır.



Şekil 5. İdeal ileri besleme $\|g^*\|$ ile bulanık ağırlık çevrimiçi kestirimi $\|\hat{W}^T \hat{\phi}\|$ (1 s'lik kayan ortalama).

Son olarak, uyarlamanın değeri bir yük değişimi denemesiyle sınanmıştır: benzetim sırasında atalet parametreleri (p_1, p_2, p_3) sırasıyla 1.15, 1.6 ve 1.6 katına çıkarıldığında, uyarlama etkinken kestirim hatası %2.3 düzeyinde kalmakta, parametreler dondurulduğunda ise %82'ye yükselmektedir. Uyarlama yasalarının işlevi tam olarak bu model bozulmasının çevrimiçi giderilmesidir.

Açıklama 2. Nihai sınırın (59)'daki yapısı, kazançların seçimi için doğrudan bir rehber verir. Gecikme bilindiğinden sınır sabiti ζ , denetim kazancı $\underline{k}$ büyüdükçe monoton azalan bir terim ile kazançtan bağımsız bir tabandan oluşur; dolayısıyla $\underline{k}$ 'i büyütmek sınırı her zaman sıkılaştırır ve iç optimum ortaya çıkmaz. Telifi kazancı α_u sınır sabitini belirleyen başlıca ayar parametresidir: büyüdükçe $e^{-\alpha_u \tau} \rightarrow 0$ olur ve telifi kaynaklı terim küçülür, ancak aynı ağırlık ileri beslemeyi de zayıflattığından ölçülü seçilmelidir. Sönümlenme kazançları $k_W = (a_3 + a_4)/\bar{W}$, $k_\Delta = a_1/\bar{\Delta}$, $k_p = a_2/\bar{p}$ seçimiyle en küçük değerini alır ve aynı terimler parametre sürüklenmesini engeller. Kazançların ölçülen başarıma etkisi Bölüm 5'te sayısal olarak incelenmektedir.

Denetim kazancından δ büyüklüğünde bir pay ayırmak yerine açık bir kazanç formülü tercih ediliyorsa, bulanık ve çapraz terimler, denetim yasasına (17) doğrusal olmayan sönümlenme eklenerek, yani $D_\phi = \|\hat{\phi}\|^2 + \left\| \frac{\partial \hat{\phi}}{\partial \Delta} \right\|^2 + \left\| \frac{\partial \hat{\phi}}{\partial \bar{p}} \right\|^2$ ile $k_n D_\phi r$ eklenip bir kazanç $K_r = \alpha_r + k_n(1 + \bar{W}^2(a_0^2 + a_1^2 +$

$a_2^2))I_n$ kullanılarak sınırlandırılabilir. Bu da bulanık kaynaklı sınır sabitini açık bir $O(1/k_n)$ sabitine dönüştürür [28]. Gecikme bilindiğinden, bu yapıcı varyant, gecikme belirsiz olduğunda ortaya çıkan *gecikme kaynaklı bozulmadan etkilenmez*. Böylece iki tasarım yalnızca doğrusal olmayan sönümlemenin e-modifikasyona göre bilinen tutuculuğu bakımından farklıdır.

6. Sonuç

Bu çalışmada, denetim kanalı bilinen sabit bir giriş gecikmesi içeren n -bağlantılı robot manipülatörler için görev uzayında çalışan, kendi kendini uyarlayan bulanık mantık tabanlı bir denetleyici geliştirilmiştir. Yöntemin özü şudur: gecikme bilindiğinden, gerçek geçmiş denetim değerlerinin üstel ağırlıklı integrali bir durum olarak tanımlanmış ve hata sistemine gecikmesiz bir bileşen eklenmiştir. Bu sayede gecikmeli giriş, hata dinamiklerinden gecikme uyumsuzluğu terimi doğurmadan çıkarılmıştır. Bilinmeyen dinamikler ise ağırlıkları, üyelik merkezleri ve genişlikleri projeksiyonlu e-modifikasyon yasalarıyla uyarlanan bir bulanık ağ tarafından çevrimiçi yaklaşılmıştır. Lyapunov–Krasovskii terimleriyle genişletilmiş bileşik bir fonksiyonel kullanılarak izleme hatasının yarı-global düzgün nihai sınırlı olduğu kanıtlanmıştır. Elde edilen sınır, gecikme belirsizliğine bağlı hiçbir bileşen taşımamakta ve denetim kazancıyla monoton azalmaktadır. Telifi kazancı α_u ise sınır sabitini belirleyen başlıca ayar parametresi olarak öne çıkmaktadır. Sayısal çalışma bu çözümlemeyi desteklemektedir: 10–175 ms gecikme aralığında denetleyici, 10 mm’lik başlangıç hatasını 0.3–2.4 mm düzeyine indirmekte, bulanık ağ ise sıfır ağırlıkla başlatılmasına karşın ideal ileri beslemeyi birkaç yüzde bağlı hatayla yeniden üretmektedir. Gecikme telifi teriminin 50 ms, bulanık mantık tabanlı ileri beslemenin ise 100 ms ve üzerinde kararlılık için gerekli olduğu, uyarlamanın ise model değiştiğinde kestirimi geçerli tuttuğu ayrıca gösterilmiştir. Deneysel doğrulama ve yarı-global sınırlı-bölge gereksiniminin gevşetilmesi, gelecekteki çalışmaların doğal yönlerini oluşturmaktadır.

Beyanlar

Yazar, herhangi bir çıkar çatışması olmadığını beyan eder.

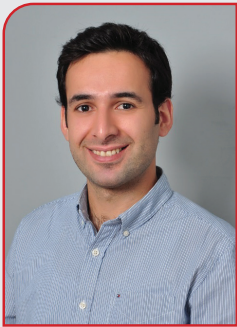
Kaynaklar

- [1] M. W. Spong, “An historical perspective on the control of robotic manipulators”, *Annual Review of Control, Robotics, and Autonomous Systems*, Cilt 5, Sayı 1, ss. 1–31, 2022.
- [2] A. Izadbakhsh, ve S. Khorashadizadeh, “Robust adaptive control of robot manipulators using bernstein polynomials as universal approximator”, *International Journal of Robust and Nonlinear Control*, Cilt 30, Sayı 7, ss. 2719–2735, 2020.
- [3] L. Sciavicco, ve B. Siciliano, *Modelling and control of robot manipulators*, Springer Science & Business Media, 2012.
- [4] B. Xian, M. de Queiroz, D. Dawson, ve I. Walker, “Task-space tracking control of redundant robot manipulators via quaternion feedback”, *Proceedings of the 2001 IEEE International Conference on Control Applications*, IEEE, 2001, ss. 363–368.

- [5] B. Siciliano, “Kinematic control of redundant robot manipulators: A tutorial”, *Journal of intelligent and robotic systems*, Cilt 3, Sayı 3, ss. 201–212, 1990.
- [6] U. Özbay, H. T. Sahin, ve E. Zergeroglu, “Robust tracking control of kinematically redundant robot manipulators subject to multiple self-motion criteria”, *Robotica*, Cilt 26, Sayı 6, s. 711, 2008.
- [7] K. Cetin, E. Tatlicioglu, ve E. Zergeroglu, “An extended jacobian-based formulation for operational space control of kinematically redundant robot manipulators with multiple subtask objectives: An adaptive control approach”, *Journal of Dynamic Systems, Measurement, and Control*, Cilt 141, Sayı 5, 2019.
- [8] S. Obuz, E. Tatlicioglu, ve E. Zergeroglu, “Adaptive cartesian space control of robotic manipulators: A concurrent learning based approach”, *Journal of the Franklin Institute*, Cilt 361, Sayı 6, s. 106701, 2024.
- [9] N. Sharma, S. Bhasin, Q. Wang, ve W. E. Dixon, “Predictor-based control for an uncertain Euler–Lagrange system with input delay”, *Automatica*, Cilt 47, Sayı 11, ss. 2332–2342, 2011.
- [10] M. Krstic, “Lyapunov tools for predictor feedbacks for delay systems: Inverse optimality and robustness to delay mismatch”, *Automatica*, Cilt 44, Sayı 11, ss. 2930–2935, Kasım 2008.
- [11] R. Kamalapurkar, N. Fischer, S. Obuz, ve W. E. Dixon, “Time-varying input and state delay compensation for uncertain nonlinear systems”, *IEEE Transactions on Automatic Control*, Cilt 61, Sayı 3, ss. 834–839, 2016.
- [12] S. Obuz, A. Parikh, I. Chakraborty, ve W. E. Dixon, “Lyapunov-based control of an uncertain Euler–Lagrange system with uncertain time-varying input delays without delay rate constraints”, *IFAC-PapersOnLine (Proc. of the IFAC Workshop)*, Elsevier, 2016.
- [13] I. Chakraborty, S. Obuz, ve W. E. Dixon, “Control of an uncertain nonlinear system with known time-varying input delays with arbitrary delay rates”, *IFAC-PapersOnLine (Proc. of the 10th IFAC Symp. on Nonlinear Control Systems, NOLCOS)*, Elsevier, 2016.
- [14] S. Obuz, J. R. Klotz, R. Kamalapurkar, ve W. Dixon, “Unknown time-varying input delay compensation for uncertain nonlinear systems”, *Automatica*, Cilt 76, ss. 222–229, 2017.
- [15] N. Sadegh, ve R. Horowitz, “An exponentially stable adaptive control law for robot manipulators”, *IEEE Trans. Robot. Autom.*, Cilt 6, Sayı 4, ss. 491–496, Ağustos 1990.
- [16] J. J. Slotine, ve W. Li, “Composite adaptive control of robot manipulators”, *Automatica*, Cilt 25, Sayı 4, ss. 509–519, Temmuz 1989.
- [17] E. Zergeroglu, W. E. Dixon, D. Haste, ve D. M. Dawson, “Composite adaptive output feedback tracking control of robotic manipulators”, *Robotica*, Cilt 17, ss. 591–600, 1999.
- [18] P. Patre, W. Mackunis, M. Johnson, ve W. E. Dixon, “Composite adaptive control for Euler-Lagrange systems with additive disturbances”, *Automatica*, Cilt 46, Sayı 1, ss. 140–147, 2010.
- [19] G. Chowdhary, ve E. Johnson, “Concurrent learning for convergence in adaptive control without persistency of

- excitation”, *Proc. IEEE Conf. Decis. Control*, IEEE, 2010, ss. 3674–3679.
- [20] A. Parikh, R. Kamalapurkar, ve W. E. Dixon, “Integral concurrent learning: Adaptive control with parameter convergence using finite excitation”, *International Journal of Adaptive Control and Signal Processing*, Cilt 33, Sayı 12, ss. 1775–1787, 2019.
- [21] S. Basu Roy, S. Bhasin, ve I. N. Kar, “Composite adaptive control of uncertain euler-lagrange systems with parameter convergence without pe condition”, *Asian Journal of Control*, Cilt 22, Sayı 1, ss. 1–10, 2020.
- [22] K. Guo, Y. Pan, D. Zheng, ve H. Yu, “Composite learning control of robotic systems: A least squares modulated approach”, *Automatica*, Cilt 111, s. 108612, 2020.
- [23] S. Obuz, E. Zergeroglu, ve E. Tatlicioglu, “Adaptive control of a class of nonlinear systems with guaranteed parameter estimation: A concurrent learning based approach”, *IET Control Theory & Applications*, 2024.
- [24] S. Commuri, ve F. L. Lewis, “Adaptive-fuzzy logic control of robot manipulators”, *Proc. IEEE Int. Conf. on Robotics and Automation (ICRA)*, Minneapolis, MN, USA, IEEE, 1996, ss. 2604–2609.
- [25] B. M. Yilmaz, E. Tatlicioglu, A. Savran, ve M. Alci, “Self-adjusting fuzzy logic based control of robot manipulators in task space”, *IEEE Transactions on Industrial Electronics*, Cilt 69, Sayı 2, ss. 1620–1629, 2022.
- [26] V. Kumar, K. Rana, ve D. Kler, “Efficient control of a 3-link planar rigid manipulator using self-regulated fractional-order fuzzy pid controller”, *Applied Soft Computing*, Cilt 82, s. 105531, 2019.
- [27] F. L. Lewis, D. M. Dawson, ve C. Abdallah, *Robot Manipulator Control Theory and Practice*, CRC, 2003.
- [28] W. E. Dixon, A. Behal, D. M. Dawson, ve S. Nagarkatti, *Nonlinear Control of Engineering Systems: A Lyapunov-Based Approach*, Birkhauser, Boston, 2003.
- [29] M. Spong, ve M. Vidyasagar, *Robot Dynamics and Control*, John Wiley & Sons Inc., New York, 1989.
- [30] Z. Cai, M. S. de Queiroz, ve D. M. Dawson, “A sufficiently smooth projection operator”, *IEEE Trans. Autom. Control*, Cilt 51, Sayı 1, ss. 135–139, Ocak 2006.

Özgeçmişler



Serhat Obuz lisans derecesini 2007 yılında İnönü Üniversitesi Elektrik-Elektronik Mühendisliği Bölümü’nden almıştır. Millî Eğitim Bakanlığı’nın yurt dışı lisansüstü öğrenim bursuyla 2012 yılında Clemson Üniversitesi’nde Elektrik ve Bilgisayar Mühendisliği alanında yüksek lisansını, 2016 yılında Florida Üniversitesi’nde Makine ve Havacılık Mühendisliği alanında doktorasını tamamlamıştır. 2021 yılından bu yana Tarsus Üniversitesi Elektrik-Elektronik Mühendisliği Bölümü’nde Dr. Öğretim Üyesi olarak çalışmaktadır. Araştırma alanları Lyapunov tabanlı doğrusal olmayan denetim, giriş ve durum gecikmesi telafisi, uyarlamalı ve gürbüz denetim, akıllı şebekelerde geçici hal kararlılığı ve dağıtık enerji kaynaklarının koordinasyonudur. 2015 yılında IEEE Conference on Control Technology and Applications (CCTA; o dönemki adıyla IEEE Conference on Control Applications) En İyi Öğrenci Makalesi Ödülü’nü almıştır.

LLM-as-a-Judge for Multilingual Question Answering Evaluation: An Experimental Study on an English–Turkish Bilingual SQuAD Dataset

Çok Dilli Soru-Cevap Değerlendirmesinde Hakem Olarak LLM: İngilizce–Türkçe İki Dilli SQuAD Veri Kümesi Üzerine Deneysel Bir İnceleme

Ilkay Yelmen

 0000-0002-1684-9717

Department of Software Engineering, Faculty of Engineering and Natural Sciences, Istinye University, Istanbul, Türkiye

ilkay.yelmen@istinye.edu.tr

Özet

Bu çalışma, büyük dil modellerinin İngilizce ve Türkçe soru-cevap değerlendirmesinde otomatik değerlendirici olarak kullanılabilirliğini karşılaştırmalı olarak incelemektedir. Bu kapsamda, beş farklı büyük dil modelinin kısmi insan anotasyonu ve kural tabanlı vekil skorlardan oluşan karma referans puanlama yapısıyla uyumu, sınıflandırma başarımı, diller arası tutarlılığı ve prompt duyarlılığı analiz edilmiştir. Paralel İngilizce–Türkçe SQuAD veri kümesinden seçilen 100 örnek için doğru, kısmen doğru ve yanlış olmak üzere üç farklı aday yanıt oluşturulmuş, bu yanıtlar basit ve rubrik tabanlı iki prompt stratejisiyle değerlendirilmiştir. Model çıktıları Spearman sıra korelasyonu, doğruluk, makro-F1, Cohen'in kappa katsayısı ve Wilcoxon işaretli sıralar testi kullanılarak incelenmiştir. Bulgular, tüm modellerin karma referans puanlarıyla yüksek düzeyde sıra korelasyonu gösterdiğini ortaya koymuştur. GPT-4o-mini İngilizcede en yüksek korelasyonu üretirken, Türkçe değerlendirmede Claude Sonnet 4.6 ve GPT-5.4-mini en yüksek korelasyonu paylaşmış; birleşik İngilizce–Türkçe sonuçlarda ise Claude Sonnet 4.6 en yüksek korelasyonu göstermiştir. Claude Sonnet 4.6, sınıflandırma başarımı ve prompt kararlılığı açısından en dengeli model olarak öne çıkmıştır. Sonuçlar, LLM-as-a-Judge yaklaşımının Türkçe soru-cevap değerlendirmesi için umut verici olduğunu ancak model seçimi, dil etkisi, prompt tasarımı ve puan kalibrasyonunun birlikte ele alınması gerektiğini göstermektedir.

Anahtar kelimeler: Büyük dil modeli, değerlendirici olarak büyük dil modelleri, soru-cevap değerlendirmesi, çok dilli değerlendirme, doğal dil işleme

Abstract

This study conducts a comparative analysis of the feasibility of using large language models as automatic evaluators in English and Turkish question-answering assessments. In this context, the study analyzes the alignment of five different large language models with a hybrid reference scoring system consisting of partial human annotation and rule-based proxy scores, as well as their classification performance, cross-lingual consistency, and prompt sensitivity. For 100 examples selected from the parallel English–Turkish SQuAD dataset, three different candidate responses, namely correct, partially correct, and incorrect responses, were generated, and these responses were evaluated using two prompt strategies: a simple prompt and a rubric-based prompt. The model outputs

were analyzed using Spearman's rank correlation, accuracy, macro-F1, Cohen's kappa coefficient, and the Wilcoxon signed-rank test. The findings revealed that all models exhibited a high level of rank correlation with the hybrid reference scores. While GPT-4o-mini produced the highest correlation in English, Claude Sonnet 4.6 and GPT-5.4-mini shared the highest correlation in Turkish; in the combined English–Turkish results, Claude Sonnet 4.6 achieved the highest correlation. Claude Sonnet 4.6 emerged as the most balanced model in terms of classification performance and prompt stability. The results indicate that the LLM-as-a-Judge approach shows promise for evaluating Turkish question-answering systems, but that model selection, language effects, prompt design, and score calibration must be addressed collectively.

Keywords: Large language model, LLM-as-a-Judge, question-answering evaluation, multilingual evaluation, natural language processing

1. Introduction

The widespread adoption of large language models in question-answering systems has made the question of how to reliably evaluate the responses generated by these systems even more critical. The generated responses often appear not merely as correct or incorrect, but in context-dependent, incomplete, partially correct, or semantically plausible yet erroneous forms. Therefore, the evaluation of question-answering systems requires more flexible evaluation approaches that go beyond simple superficial similarity metrics and take into account the response's contextual alignment, semantic accuracy, and scope.

Traditional evaluation metrics are primarily based on the similarity of words or phrases between the candidate response and the reference response. While metrics such as BLEU and ROUGE are useful for certain tasks, they do not always adequately capture aspects of question-answering evaluation such as the response's contextual appropriateness, semantic accuracy, and completeness. This issue becomes particularly evident when evaluating responses that are partially correct, incomplete, or contextually limited.

Due to these limitations, the “LLM-as-a-Judge” approach which utilizes large language models as automatic evaluators has gained prominence in recent years. In this approach, large language models can score, rank, or classify candidate responses by evaluating them in conjunction with the question, context, and reference answer. The LLM-as-a-Judge

method is more scalable than human evaluation; it also has the potential to perform more contextual and semantic evaluations compared to traditional metrics [1]. However, previous studies have shown that LLM judges are not equally reliable across all tasks and contexts, and that their alignment with human judgments can vary, particularly in tasks requiring expertise or involving more complex evaluation dimensions such as user satisfaction [2], [3].

However, there is a significant research gap in the context of multilingual evaluation. The majority of LLM-as-a-Judge studies have been conducted primarily in English; the reliability of this approach has been less thoroughly examined in languages such as Turkish, which are morphologically rich and have more limited evaluation resources. Due to its agglutinative structure and distinct syntactic features, Turkish may present different challenges for automatic evaluation compared to English. Therefore, language-specific and systematic analyses are needed to evaluate the outputs of large language models in Turkish. While studies such as TurkBench highlight the importance of specialized benchmarks for Turkish LLM evaluation [4], the alignment of the LLM-as-a-Judge approach with hybrid reference scores, its classification performance, cross-lingual consistency, and prompt sensitivity in Turkish question-answering evaluation have not yet been sufficiently and comprehensively investigated.

This study examines the feasibility and consistency of the LLM-as-a-Judge approach in evaluating English–Turkish question-answering tasks. To this end, 100 examples were selected from the parallel English–Turkish SQuAD dataset; for each example, three different types of candidate answers were generated: correct, partially correct, and incorrect. Five large language models scored these candidate responses on a scale of 1–5 under two different prompt strategies: a simple prompt and a rubric-based prompt. The model outputs were analyzed in terms of agreement with hybrid reference scores, classification accuracy, macro-F1, Cohen’s kappa, cross-language scoring differences, and prompt sensitivity.

The main contributions of the study can be summarized as follows:

- This study provides preliminary empirical evidence on the use of large language models as automatic evaluators in a language like Turkish, which is morphologically complex and underrepresented in the LLM-as-a-Judge literature.
- By evaluating the English and Turkish equivalents of the same question, context, and response levels, this approach directly compares not only the models’ overall performance but also the consistency of their evaluations across languages.
- By considering the “correct,” “partially correct,” and “incorrect” response categories together, it highlights the models’ ability to distinguish “partially correct” responses, which constitute a particularly critical and challenging category for evaluation.
- The study demonstrates that high agreement with reference scores does not always translate to reliable classification performance, highlighting the need to measure score calibration and class discrimination performance separately in LLM-as-a-Judge evaluations.
- By comparing simple and rubric-based prompt strategies, it demonstrates that evaluation outcomes can vary depending on prompt design and language; thus, it positions prompt robustness as a

fundamental evaluation dimension for multilingual LLM-as-a-Judge systems.

- The findings provide a practical evaluation framework for model selection, prompt design, and score calibration in Turkish and multilingual question-answering systems.

The remaining sections of the paper are structured as follows. Section 2 summarizes the literature on the LLM-as-a-Judge approach and multilingual evaluation studies. Section 3 describes the dataset, candidate response generation, judge models, the prompting process, and evaluation metrics. Section 4 presents the experimental setup and reproducibility settings. Section 5 presents results on model performance, agreement with hybrid reference scores, language differences, and prompt sensitivity. Section 6 discusses the findings, while Section 7 presents conclusions and recommendations for future work.

2. Related Work

The LLM-as-a-Judge approach has emerged in recent years as a scalable alternative to human evaluation in automated assessment processes. Li et al. have presented a comprehensive literature review examining this approach across various scenarios, such as evaluation, scoring, ranking, and selection. The study highlights that traditional matching-based or small-model-based evaluation methods have limitations when applied to open-ended and dynamic tasks; the LLM-as-a-Judge approach has been classified based on the dimensions of “what is evaluated,” “how it is evaluated,” and “which benchmarks are used for measurement” [1]. This framework demonstrates that LLM-based evaluators are not merely automated scoring tools, but systems that must be carefully examined in terms of reliability, bias, and assessment design.

Current studies show that LLM judges can produce promising results on various tasks, but that this performance varies depending on the task type and evaluation context. Yu et al. treated the “LLM-as-a-Judge” capability as a general skill of large language models and proposed a two-stage training approach consisting of supervised fine-tuning and direct preference optimization. The study demonstrates that judgment performance can be improved not only through prompt design but also through specialized training and preference optimization processes [5]. In contrast, Huang et al. noted that while judge models fine-tuned based on open-source models may demonstrate high performance on their own test datasets, they still lag behind GPT-4 in terms of generalizability, fairness, and adaptability [6]. These findings indicate that high performance alone is not sufficient in LLM-as-a-Judge systems; reliability analysis is necessary across different data domains, tasks, and usage conditions.

The reliability of the LLM-as-a-Judge approach is closely related to its alignment with human judgments and the transparency of its evaluation criteria. Yamasaki et al. proposed the Function-Based Labels approach to more clearly define the ambiguous relationships between product pairs in complementary product recommendations. In this study, a dataset containing 2,759 human-annotated product pairs was created based on the functional features of the products, independent of user behavior history. ModernBERT demonstrated the best machine learning performance with a Macro-F1 score of 0.911; GPT-4o-mini, on the other hand, achieved a consistency score of 0.989 and a classification accuracy of 0.849 when a detailed FBL definition was used [7]. These results are significant in that they demonstrate that

clear and detailed evaluation definitions can improve alignment with human judgments in LLM-based annotations.

Domain-specific studies highlight both the potential and limitations of LLM judges in different usage contexts. Türkmen and Keleş examined the “LLM-as-a-Judge” approach to evaluating RAG systems using 150,000 real e-commerce interactions on the Trendyol QA Assistant. In the study, retrieval and generation-related errors were distinguished by combining user satisfaction labels, LLM evaluator scores, and factor-based discriminant analyses. The findings have shown that LLM-based judge models can generally reflect trends in user satisfaction, but may overlook certain details of dissatisfaction [2]. Szymanski et al., on the other hand, found that LLM judges demonstrated only limited agreement with human experts in specialized fields such as dietetics and mental health. In the study, LLM judges agreed with experts 68% of the time in the field of dietetics and 64% of the time in the field of mental health [3]. These studies show that LLM evaluators should be viewed as complementary assessment tools used in conjunction with human feedback, rather than as a complete replacement for human evaluation.

Research in various fields—such as healthcare, creative writing, and software engineering—also demonstrates that the “LLM-as-a-Judge” approach must be evaluated in a task-sensitive manner. Khan et al. proposed a framework called “AI on Trial” that uses the “LLM-as-a-Judge” approach in clinical decision-making processes and ensures that responses are evaluated based on criteria of quality, completeness, safety, and clarity [8]. Bellard et al. integrated an LLM-as-a-Judge component into an agent-based RAG system for the management of heart failure and demonstrated that this architecture can improve the robustness of responses [9]. Mahadik et al. used multidimensional evaluation criteria such as factual consistency, emotional tone, emotional expression, readability, and overall response quality in medical dialogue systems [10]. Abir et al. focused on evaluating creative and semantically complex texts by using the ROUGE, BERTScore, and LLM-as-a-Judge metrics together in a poetry summarization task [11]. Crupi et al., on the other hand, examined the “LLM-as-a-Judge” approach in code generation and code summarization tasks; they demonstrated that, although GPT-4-turbo was the most successful judge model, it could occasionally misjudge code accuracy and summary quality [12]. These findings indicate that a single, general performance metric is insufficient in LLM-based evaluation systems; instead, a multidimensional evaluation tailored to the specific task, domain, and output type is necessary.

Another key trend emerging in LLM-as-a-Judge research is the expansion of evaluation processes to include multimodal scenarios and those sensitive to different linguistic contexts. The MLLM-as-a-Judge benchmark, developed in this vein, examines the evaluation capabilities of vision-language models by adapting the LLM-as-a-Judge approach to multimodal large language models. Within the scope of this benchmark, the models were evaluated using rating, pairwise comparison, and aggregate ranking tasks. The findings show that the models are capable of making human-like decisions in pairwise comparisons, but may exhibit significant differences from human preferences in rating and aggregate ranking tasks [13].

In general, the literature indicates that the LLM-as-a-Judge approach holds strong potential for scalable and context-aware evaluation; however, for reliable use, factors such as human collaboration, task type, language effects, prompt design, and evaluation metrics must be considered together. In this

context, the present study contributes to addressing the multilingual reliability gap in the literature by jointly examining the agreement of LLM evaluators with hybrid reference scores, their classification performance, cross-lingual consistency, and prompt sensitivity in English-Turkish question-answering evaluations.

3. Methodology

The methodological design of this study was structured to examine the extent to which large language models can reliably score English and Turkish question-answer pairs using an automatic evaluator. In the experimental process, parallel English-Turkish SQuAD examples with the same content structure were used, and candidate answers representing different levels of accuracy were generated for each example. These candidate responses were categorized into three groups: correct, partially correct, and incorrect. This allowed for an analysis not only of the models’ ability to distinguish clearly correct or incorrect answers but also of their capacity to correctly classify partially correct answers, which are more challenging to evaluate. The study design was structured to evaluate model performance across complementary dimensions such as agreement with hybrid reference scores, classification performance, cross-lingual consistency, and prompt sensitivity rather than reducing it to a single metric.

In the experimental workflow, 100 parallel English-Turkish question-answer pairs were first selected, and then three different candidate responses were generated for each pair using the GPT-4o-mini model. Subsequently, five different large language models were used as evaluators, and each was asked to rate each candidate response on a scale of 1 to 5. The evaluation process was conducted under two different prompt strategies: simple prompt and rubric-based prompt. This structure allowed for an examination of both the models’ overall evaluation performance and whether there were changes in scoring behavior depending on the prompt type. The obtained scores were analyzed using metrics such as Spearman correlation, accuracy, Macro-F1, and Cohen’s kappa; additionally, language differences and prompt sensitivity were evaluated using statistical methods such as the Wilcoxon signed-rank test and the Friedman test. The general experimental process followed in this context is summarized in Figure 1.

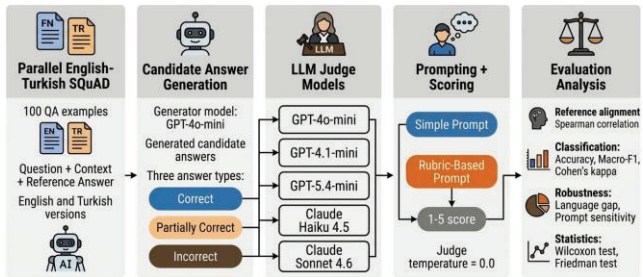


Figure 1. Experimental pipeline for multilingual LLM-as-a-Judge evaluation

3.1. Dataset and Candidate Answer Generation

In this study, English-Turkish parallel examples from the validation split of the Stanford Question Answering Dataset (SQuAD) were used. The experimental dataset consists of 100 question-answer pairs with successfully established English-Turkish alignment. The 100 examples were selected deterministically from the SQuAD v1.1 validation split and its Turkish counterpart, SQuAD-TR. English and Turkish records

were matched using their common SQuAD identifiers, and only examples with non-empty reference answers in both languages were considered eligible. Among the 2,346 aligned English–Turkish examples satisfying these criteria, the first 100 pairs encountered in the original dataset order were included in the study. No random sampling, stratification, or filtering based on question type, answer type, answer length, or context length was applied. When multiple reference answers were available, the first reference answer provided for each example was used consistently in both the candidate-response generation and evaluation procedures. Each example consists of three basic components: a question, a context paragraph, and a reference answer. Using the English and Turkish versions of these components enabled the models to be compared under the same evaluation framework in both languages. Thus, model performance was assessed using parallel English and Turkish versions of the same underlying question-answer content.

Candidate responses were generated separately for each example and each language. At this stage, the GPT-4o-mini model was used as the generative model, and the temperature value was set to 0.7. The model was asked to generate candidate responses at three different levels of accuracy, taking into account the given context and the reference answer. These accuracy levels were defined as “correct,” “partially correct,” and “incorrect.” “Correct” responses represent context-based answers that are fully consistent with the reference answer; “partially correct” responses contain accurate information but are incomplete or limited; and “incorrect” responses are clearly erroneous but may appear superficially plausible.

To construct the reference scoring scheme used in the evaluation process, a hybrid scoring approach was adopted. For 25 of the 100 parallel examples, corresponding to 150 candidate responses in total, the generated answers were manually scored by the author on a 1–5 scale. These manually assigned scores were used as human reference scores for this subset. For the remaining 75 examples, rule-based proxy scores were assigned according to the predefined accuracy levels. Accordingly, correct responses were represented by 5 points, partially correct responses by 3 points, and incorrect responses by 1 point. This proxy scoring scheme was selected to maintain consistency with the score-to-label conversion rule used in the classification analysis, where scores ≤ 2 indicate incorrect responses, a score of 3 indicates partially correct responses, and scores ≥ 4 indicate correct responses. This framework made it possible to examine not only the ability of large language models to distinguish between clearly correct and incorrect responses but also their capacity to identify partially correct responses, which are more ambiguous in terms of evaluation. However, because the human-scored subset was annotated by a single human evaluator and covered only part of the dataset, the resulting reference structure should be interpreted as a partially human-annotated and partially proxy-based scoring framework rather than as a fully human-annotated gold standard. The prompt template used to generate candidate responses was designed to include explicit constraints to ensure that each type of response was formulated in accordance with the specified quality level.

3.2. Judge Models and Prompting Procedure

In this study, five large language models from two different providers were used as evaluator models for the automatic evaluation of candidate responses, as shown in Table 1. The GPT-4o-mini, GPT-4.1-mini, and GPT-5.4-mini

models offered by OpenAI, as well as the Claude Haiku 4.5 and Claude Sonnet 4.6 models offered by Anthropic, were included in the evaluation process. All judge models were run with a temperature value of 0.0 to minimize sampling variability and improve reproducibility. Each model was presented with the question in the relevant language, the reference answer, and the candidate answer; the models were asked to rate the candidate answer on a scale of 1 to 5. The original SQuAD context paragraph was not included in the judge prompt.

Table 1. Judge models used in the evaluation

Model	Provider	Model Family
GPT-4o-mini	OpenAI	GPT-4o
GPT-4.1-mini	OpenAI	GPT-4.1
GPT-5.4-mini	OpenAI	GPT-5
Claude Haiku 4.5	Anthropic	Claude Haiku
Claude Sonnet 4.6	Anthropic	Claude Sonnet

Two different guidance strategies, simple prompting and rubric-based prompting, were used in the prompting process. While the simple prompt required the candidate’s response to be scored based on the criteria of correctness and clarity, the rubric-based prompt explicitly listed correctness, clarity, and completeness as evaluation criteria, providing a more structured evaluation instruction than the simple prompt. The same evaluation workflow was applied to both English and Turkish examples, thereby enabling a comparison of the models’ cross-lingual consistency and sensitivity to different prompt types. To improve reproducibility, representative examples of the prompt templates are provided here. In the simple-prompt condition, the evaluator was instructed: “Score the candidate answer from 1 to 5 based on correctness and clarity.” In the rubric-based condition, the evaluator was instructed: “Evaluate the candidate answer based on correctness, clarity, and completeness.” Both prompt types then presented the question, reference answer, and candidate answer, and required the model to return a JSON-formatted score and short justification. Equivalent Turkish-language templates were used for the Turkish evaluation condition. As part of the experimental design, a total of 6,000 evaluations were conducted using a combination of 5 models, 100 examples, 2 languages, 3 response labels, and 2 prompt strategies.

3.3. Scoring Protocol and Evaluation Metrics

In this study, the judge models were asked to score each candidate response on a scale of 1 to 5. During the scoring process, the candidate response was evaluated alongside the relevant question and reference answer. The original SQuAD context paragraph was not provided to the judge models. A hybrid reference scoring scheme was used. For 25 of the 100 parallel examples, corresponding to 150 candidate responses, scores were manually assigned by the author on a 1–5 scale. For the remaining 75 examples, rule-based proxy scores were used: correct answers were assigned 5 points, partially correct answers 3 points, and incorrect answers 1 point. In this way, the scores generated by the model for each candidate response were made comparable to partially human-annotated and partially proxy-based reference scores. This framework was used to examine whether the models ranked the responses in a manner consistent with the reference scoring scheme.

A predefined threshold rule was applied when converting the model scores into a three-class evaluation framework.

Accordingly, responses with a score ≤ 2 were classified as “incorrect,” those with a score = 3 as “partially correct,” and those with a score ≥ 4 as “correct.” Thanks to this transformation, the models were evaluated not only on their ability to assign scores between 1 and 5, but also on their ability to distinguish between correct, partially correct, and incorrect answers. In particular, maintaining the “partially correct” class as a separate category made it possible to examine the extent to which the models could distinguish answers that were incomplete or only partially correct.

Multiple complementary metrics were used in the evaluation. The agreement between model scores and reference scores was measured using Spearman’s rank correlation. Pearson correlation was additionally reported as a complementary measure to assess the linear association between model scores and reference scores and to compare it with the rank-based agreement captured by Spearman’s correlation. Classification performance was calculated using accuracy and macro-F1 scores following score-to-label conversion. While accuracy provides the ratio of examples correctly classified by the model to the total number of examples, macro-F1 was used to evaluate performance across the three classes in a balanced manner. The agreement between the labels generated by the model and the reference labels was measured using Cohen’s kappa coefficient.

Prompt sensitivity was defined as the absolute difference between the mean scores obtained for the same example under the simple and rubric-based prompt conditions, calculated as $|\text{mean_score_simple} - \text{mean_score_rubric}|$. The absolute difference was used to quantify the magnitude of prompt-induced score variation irrespective of direction; therefore, this metric does not indicate whether the rubric-based prompt produced systematically higher or lower scores. Language difference, on the other hand, was assessed based on the difference between the scores for the same example in English and Turkish. The 95% confidence intervals for the metrics were obtained using the percentile bootstrap method. In the bootstrap process, the number of resamples was set to 1,000, and the random seed was fixed at 42 (BOOTSTRAP_RANDOM_SEED) to ensure reproducibility. In each bootstrap iteration, sampling was performed with replacement to preserve the original sample size. The sampling unit was defined at the response-pair level, depending on the metric in question; for classification metrics, the reference–prediction pairs were used, while for correlation metrics, the model-score–reference-score pairs were used. Grouped bootstrap was not applied; correct, partially correct, and incorrect candidate responses to the same question were treated as independent response pairs. The original sample size of $n = 600$ was maintained for each cell; confidence intervals were reported using the 2.5th and 97.5th percentiles of the bootstrap distribution. The statistical significance of language and prompt effects was assessed using the Wilcoxon signed-rank test; for the prompt-sensitivity analyses, p-values were additionally adjusted using the Benjamini–Hochberg procedure to account for multiple comparisons. Differences between models were analyzed using the Friedman test. Because the human-scored subset was annotated by a single evaluator, the findings should be interpreted as agreement with a hybrid reference structure rather than validation against a fully human-annotated gold standard. Although the experimental design yielded 6,000 model judgments, these judgments were derived from 100 underlying parallel question–answer pairs. Therefore, the 6,000 judgments should be interpreted as repeated evaluation

observations generated from 100 fundamental examples rather than as 6,000 statistically independent samples.

4. Experimental Setup

4.1. Model Configuration

In the experimental evaluation, five large language models from two different API providers were used as evaluation models. The GPT-4o-mini, GPT-4.1-mini, and GPT-5.4-mini models from OpenAI, and the Claude Haiku 4.5 and Claude Sonnet 4.6 models from Anthropic were included in the evaluation process. The evaluator models were used to evaluate candidate responses generated from English and Turkish question-answering examples. All evaluator models were run with a temperature parameter of 0.0 to minimize sampling variability and improve reproducibility. Candidate response generation was carried out as a separate step; in this step, GPT-4o-mini was used as the generative model, and the temperature value was set to 0.7.

The experimental setup is based on 100 parallel English/Turkish examples selected from the SQuAD validation set. For each example, fields for the English and Turkish questions, the context paragraph, and the reference answer were used. For each example and language, three different candidate response labels—correct, partially correct, and incorrect—were evaluated. Additionally, each evaluator model scored the candidate responses using two different prompt strategies: simple and rubric-based. The experiment was designed with 5 models, 100 examples, 2 languages, 3 response labels, 2 prompt strategies, and 1 repetition. As a result of this structure, a total of 6,000 model judgments were obtained. These judgments were generated from 100 underlying parallel question–answer pairs and therefore do not represent 6,000 independent experimental units.

4.2. Evaluation Procedure

The evaluation process was conducted on 100 parallel English/Turkish examples selected from the SQuAD validation set. For each example, the question, reference answer, and candidate answer were provided to the relevant evaluator model. The original context paragraph was not included in the judge-model input. The candidate responses were created at three different levels of accuracy: correct, partially correct, and incorrect; each candidate response was evaluated in both the English and Turkish versions. The evaluator models were instructed to generate a score between 1 and 5 by comparing the given candidate response with the reference response.

Each model was run through the same evaluation workflow. Each of the five evaluators scored candidate responses for 100 examples across two languages, three candidate response labels, and two prompt strategies. Two different types of prompts were used in the prompting process: simple prompts and rubric-based prompts. In the simple prompt, evaluation was based on the criteria of correctness and clarity; in the rubric-based prompt, however, evaluation was structured around the dimensions of correctness, clarity, and completeness. The experimental design was implemented as $5 \text{ models} \times 100 \text{ examples} \times 2 \text{ languages} \times 3 \text{ labels} \times 2 \text{ prompts} \times 1 \text{ repetition}$, resulting in a total of 6,000 judgments.

The model scores obtained were converted to class labels using a predefined score-to-label conversion rule. Accordingly, responses with a score ≤ 2 were considered

incorrect, responses with a score = 3 were considered partially correct, and responses with a score ≥ 4 were considered correct. Following this process, the model scores were compared with the hybrid reference scores, and the transformed model labels were compared with the reference labels. The evaluation results were analyzed using Spearman’s rank correlation, accuracy, macro-F1, and Cohen’s kappa coefficient. Additionally, prompt sensitivity and English–Turkish score differences were calculated in a paired manner using the same examples.

4.3. Reproducibility Settings

To ensure the reproducibility of the experiments, a fixed and predefined experimental framework was followed during the evaluation process. The dataset was limited to 100 parallel English/Turkish examples selected from the SQuAD validation set; for each example, the question, context paragraph, and reference answer fields were used during dataset preparation and candidate-response generation. The candidate responses were generated under three categories correct, partially correct, and incorrect and all judge models evaluated them using the same candidate response structure. Thus, model comparisons were conducted based on the same data, the same language conditions, and the same response categories.

The model settings were also kept constant to ensure comparability. While the GPT-4o-mini model was used with a temperature of 0.7 to generate candidate responses, all evaluator models were run with a temperature of 0.0 during the evaluation phase. Each model was evaluated on the same 100 examples in two languages—English and Turkish—using three candidate response labels and two prompt strategies. The experimental design was implemented as 5 models $\times$ 100 examples $\times$ 2 languages $\times$ 3 labels $\times$ 2 prompts $\times$ 1 repetition, resulting in a total of 6,000 judgments.

During the analysis phase, predefined rules were used for scoring and labeling. A hybrid reference scoring scheme was adopted: 25 examples, corresponding to 150 candidate responses, were manually scored by the author on a 1–5 scale, while the remaining 75 examples were scored using rule-based proxy scores. For the proxy-scored examples, correct, partially correct, and incorrect answers were assigned scores of 5, 3, and 1, respectively. When converting model scores to labels, the following rule was applied: a score of ≤ 2 was considered incorrect, a score of 3 was considered partially correct, and a score of ≥ 4 was considered correct. Confidence intervals were calculated at the 95% level using 1,000 bootstrap resamples; the significance of language and prompt effects was assessed using the Wilcoxon signed-rank test, with Benjamini–Hochberg adjustment applied to the prompt-sensitivity analyses, while differences between models were evaluated using the Friedman test. The human-scored subset was included to support the reference scoring framework, but because it was annotated by a single evaluator, it was not used to estimate inter-annotator agreement.

5. Results

5.1. Model Performance Results

The performance of the evaluator models was analyzed from two complementary perspectives: their agreement with the hybrid reference scores and their ability to classify candidate responses into the correct, partially correct, and

incorrect categories. This distinction is important because a model may preserve the overall ranking of response quality while still showing limited performance in discrete class assignment. Accordingly, Table 2 reports both rank-based agreement and classification-based performance metrics for each model under the English and Turkish evaluation settings.

Table 2. Agreement and classification performance metrics by model and language.

Model	Language	Spearman ρ	Accuracy	Macro-F1	Cohen’s κ
GPT-4o-mini	English	0.925	0.747	0.695	0.620
GPT-4o-mini	Turkish	0.867	0.728	0.687	0.593
Claude Sonnet 4.6	English	0.908	0.772	0.735	0.658
Claude Sonnet 4.6	Turkish	0.901	0.753	0.717	0.630
GPT-5.4-mini	English	0.903	0.670	0.552	0.505
GPT-5.4-mini	Turkish	0.901	0.692	0.611	0.538
Claude Haiku 4.5	English	0.890	0.678	0.577	0.518
Claude Haiku 4.5	Turkish	0.870	0.705	0.642	0.558
GPT-4.1-mini	English	0.882	0.702	0.639	0.553
GPT-4.1-mini	Turkish	0.849	0.705	0.656	0.558

As shown in Table 2, all models produced high Spearman correlation coefficients in both English and Turkish evaluations, indicating that their score rankings were generally consistent with the expected ordering of response quality. In the English evaluation, the highest Spearman correlation was obtained by GPT-4o-mini ($\rho = 0.925$), followed by Claude Sonnet 4.6 ($\rho = 0.908$) and GPT-5.4-mini ($\rho = 0.903$). In the Turkish evaluation, the highest Spearman correlation was observed for Claude Sonnet 4.6 and GPT-5.4-mini, both with $\rho = 0.901$, followed by Claude Haiku 4.5 ($\rho = 0.870$) and GPT-4o-mini ($\rho = 0.867$). However, the classification metrics reveal a different pattern. Claude Sonnet 4.6 achieved the highest accuracy, macro-F1, and Cohen’s κ values in both languages, with an accuracy of 0.772, a macro-F1 score of 0.735, and $\kappa = 0.658$ in English, and an accuracy of 0.753, a macro-F1 score of 0.717, and $\kappa = 0.630$ in Turkish. Although GPT-5.4-mini produced one of the highest Spearman correlation values in both languages, its classification performance was lower, particularly in English, where it achieved an accuracy of 0.670, a macro-F1 score of 0.552, and $\kappa = 0.505$. This finding indicates that high rank correlation does not necessarily imply equally strong classification performance across the three response categories, namely correct, partially correct, and incorrect.

These results show that evaluating differences between models based solely on correlation coefficients is not sufficient. In particular, differences in classification performance indicate that, although the models exhibit similar ranking behavior, they may apply decision boundaries differently. Therefore, a visual comparison of model accuracy helps to highlight these performance differences more clearly. This distinction is particularly important when assessing the practical reliability of judge models, since similar correlation levels may conceal meaningful differences in categorical decision performance. To provide a more direct comparison of classification performance across models and languages, Figure 2 visualizes the accuracy scores reported in Table 2.

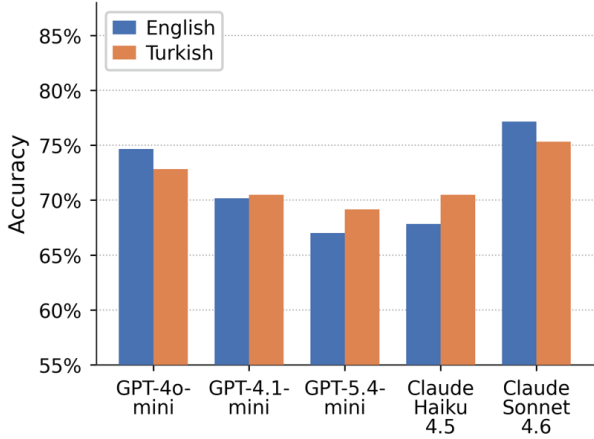


Figure 2. Classification accuracy of LLM-as-a-Judge models in English and Turkish.

According to Figure 2, the highest classification accuracy was achieved by the Claude Sonnet 4.6 model. This model achieved an accuracy of approximately 77% for English and approximately 75% for Turkish. Although GPT-4o-mini demonstrated strong performance in English, it produced lower accuracy in Turkish. In contrast, the GPT-5.4-mini and Claude Haiku 4.5 models showed relatively lower accuracy in English. Overall, the results indicate that the models’ classification performance is at a medium-to-high level, but not all models demonstrate the same level of balanced performance.

Figure 3 presents the confusion matrix of Claude Sonnet 4.6 under the Turkish evaluation setting. This model was examined in greater detail because it achieved the most balanced overall classification performance among the evaluated judge models. The confusion matrix provides a more fine-grained view of model behavior by showing how often each true response category was assigned to the correct or incorrect predicted label.

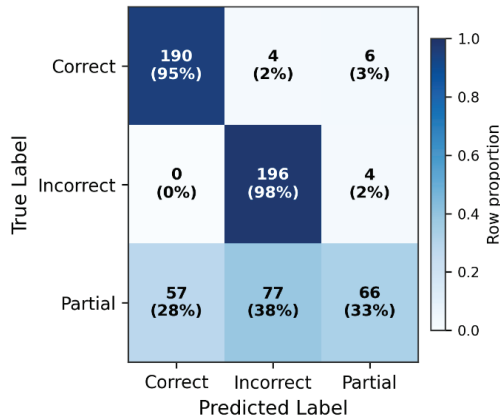


Figure 3. Confusion matrix of Claude Sonnet 4.6 for Turkish evaluation.

As shown in Figure 3, Claude Sonnet 4.6 classified the clearly defined categories with high reliability, correctly identifying 98% of incorrect responses and 95% of correct responses. In contrast, performance on the partially correct category was substantially weaker, with only 33% of partial responses correctly assigned to the partial class. Most errors in this category involved confusion with the incorrect label,

accounting for 38% of partial responses, whereas 28% were misclassified as correct. This pattern indicates that the model was highly effective in separating clearly correct and clearly incorrect answers, but less reliable in identifying responses with intermediate or borderline quality.

Figure 4 presents the confusion matrix of Claude Sonnet 4.6 under the English evaluation setting, allowing a class level comparison of prediction behavior across correct, incorrect, and partially correct responses.

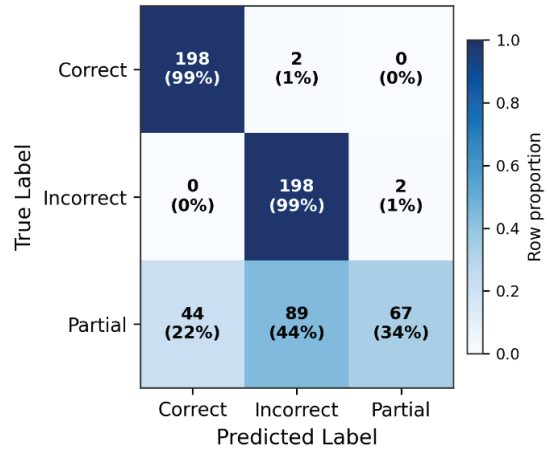


Figure 4. Confusion matrix of Claude Sonnet 4.6 for English evaluation.

As shown in Figure 4, Claude Sonnet 4.6 demonstrated very high classification accuracy for the clearly defined categories in the English evaluation setting, correctly identifying 99% of correct responses and 99% of incorrect responses. In contrast, performance on partially correct responses remained limited, with only 34% assigned to the partial class. Most errors in this category were due to confusion with the incorrect label, which accounted for 44% of partial responses, whereas 22% were misclassified as correct. This pattern confirms that, even under the English evaluation condition, the model was highly reliable for clearly correct and clearly incorrect answers but substantially less effective in identifying responses with intermediate quality.

To examine overall agreement across languages, the English and Turkish evaluation results were combined and analyzed jointly. Table 3 reports the combined correlation results between judge scores and the hybrid reference scores for each model.

Table 3. Combined English-Turkish correlation between judge scores and hybrid reference scores (n = 1,200 per model).

Model	Spearman ρ	Lower CI	Upper CI	Pearson r	p value
GPT-5.4-mini	0.902	0.885	0.916	0.886	< 0.001
Claude Sonnet 4.6	0.905	0.888	0.920	0.900	< 0.001
GPT-4o-mini	0.896	0.881	0.910	0.891	< 0.001
Claude Haiku 4.5	0.880	0.864	0.896	0.873	< 0.001
GPT-4.1-mini	0.865	0.843	0.885	0.857	< 0.001

As shown in Table 3, all models produced high combined Spearman correlation values, indicating strong overall agreement with the hybrid reference scores across the bilingual evaluation setting. Claude Sonnet 4.6 achieved the highest combined Spearman correlation with a value of 0.905, followed closely by GPT-5.4-mini with 0.902 and GPT-4o-mini with 0.896. In terms of linear association, the highest Pearson correlation coefficient was obtained by Claude Sonnet 4.6 with a value of 0.900. Moreover, the p values were below 0.001 for all models, indicating that the observed relationships between model scores and the hybrid reference scores were statistically significant.

To examine scoring behavior beyond classification-based performance, Table 4 reports the mean judge scores assigned to correct, partially correct, and incorrect responses in English and Turkish, averaged across the two prompt strategies. These mean scores provide a more direct view of how each model used the 1 to 5 scale and how distinctly it separated the three response quality categories.

Table 4. Mean judge scores for each candidate answer label and language.

Model	Language	Correct	Partially Correct	Incorrect
GPT-4o-mini	English	4.93	2.94	1.28
GPT-4o-mini	Turkish	4.65	2.82	1.46
Claude Sonnet 4.6	English	4.96	2.65	1.12
Claude Sonnet 4.6	Turkish	4.82	2.78	1.15
Claude Haiku 4.5	English	4.89	2.67	1.19
Claude Haiku 4.5	Turkish	4.80	2.71	1.32
GPT-4.1-mini	English	4.80	2.82	1.21
GPT-4.1-mini	Turkish	4.69	2.96	1.24
GPT-5.4-mini	English	4.92	2.60	1.04
GPT-5.4-mini	Turkish	4.84	3.18	1.10

As shown in Table 4, correct responses are assigned high scores, incorrect responses are assigned low scores, and partially correct responses are assigned intermediate scores across all models. This pattern indicates that the models are able to distinguish between the three response categories in terms of their scoring behavior. While the average scores for correct answers range from 4.65 to 4.96, the average scores for incorrect answers range from 1.04 to 1.46. Partially correct answers, however, show greater variability across models and languages. This suggests that the models are able to distinguish between correct and incorrect answers more consistently, whereas model and language-based differences are more pronounced in the scoring of partially correct answers.

5.2. Statistical Analysis

To further examine the consistency of model scoring behavior, additional statistical analyses were conducted with a particular focus on cross linguistic differences and prompt-related variation. As an initial step, score differences between

the English and Turkish evaluation settings were analyzed to determine whether the models exhibited systematic language dependent scoring patterns. The results of the Wilcoxon signed-rank test for paired English and Turkish scores are presented in Table 5.

Table 5. Wilcoxon signed-rank test results for the difference between English and Turkish scores.

Model	Statistic	p value	Effect Size (r)	Significant?
GPT-5.4-mini	5219.5	0.000117	-0.330	Yes
GPT-4o-mini	12362.5	0.084	+0.123	No
Claude Haiku 4.5	13633.5	0.496	-0.049	No
GPT-4.1-mini	14312.0	0.557	-0.042	No
Claude Sonnet 4.6	9673.5	0.727	-0.028	No

As shown in Table 5, a statistically significant difference between the English and Turkish scores was observed only for GPT-5.4-mini, with p = 0.000117 and an effect size of r = -0.330. For all other models, the p values remained above the significance threshold, indicating no statistically significant language related score difference. These results suggest that the effect of language on scoring behavior was model dependent and was most pronounced in GPT-5.4-mini.

To examine the effect of prompt type on model scores, paired simple-prompt and rubric-based-prompt scores were compared using the Wilcoxon signed-rank test. Prompt sensitivity was summarized separately as the mean absolute difference between the corresponding mean scores. Table 6 presents the mean sensitivity values, 95% confidence intervals, p values, effect sizes, and significance outcomes for each model and language.

Table 6. Wilcoxon signed-rank test results for prompt-type differences, with mean sensitivity, 95% CIs, and BH-adjusted p-values.

Model	Language	Mean Sensitivity	95% CI	p value	p (BH)	Effect Size	Significant?
GPT-4.1-mini	English	0.230	[0.177, 0.287]	< 0.001	< 0.001	0.884	Yes
GPT-4o-mini	Turkish	0.230	[0.177, 0.290]	< 0.001	< 0.001	0.793	Yes
GPT-5.4-mini	Turkish	0.163	[0.113, 0.210]	< 0.001	< 0.001	0.754	Yes
Claude Haiku 4.5	Turkish	0.100	[0.070, 0.133]	0.028	0.0569	-0.400	No
Claude Sonnet 4.6	English	0.033	[0.013, 0.057]	0.002	0.0039	1.000	Yes
GPT-4.1-mini	Turkish	0.173	[0.127, 0.227]	0.050	0.083	0.310	No
GPT-4o-mini	English	0.117	[0.080, 0.157]	0.139	0.1737	0.257	No
GPT-5.4-mini	English	0.133	[0.087, 0.183]	0.137	0.1737	-0.286	No
Claude	Turkish	0.087	[0.053, 0.121]	0.705	0.783	0.077	No

Sonnet 4.6	h		0.120]		8		
Claude Haiku 4.5	English	0.057	[0.033, 0.083]	0.819	0.8185	0.059	No

As shown in Table 6, prompt sensitivity varied across both models and languages. The highest mean sensitivity was observed for GPT-4.1-mini in English and GPT-4o-mini in Turkish, both with a value of 0.230, followed by GPT-4.1-mini in Turkish with 0.173 and GPT-5.4-mini in Turkish with 0.163. After applying the Benjamini–Hochberg correction, statistically significant prompt effects were observed for GPT-4.1-mini in English, GPT-4o-mini in Turkish, GPT-5.4-mini in Turkish, and Claude Sonnet 4.6 in English. Among the Claude models, Claude Sonnet 4.6 exhibited a statistically significant but very small sensitivity in English, whereas Claude Haiku 4.5 in Turkish was not significant after correction. Overall, these results indicate that prompt formulation affected the scoring behavior of some models more strongly than others, and that this effect varied depending on both the model and the evaluation language.

To examine how scoring behavior changed across prompt types, prompt sensitivity was quantified as the mean absolute difference between the corresponding mean scores obtained under the simple and rubric-based prompt conditions for each model in the English and Turkish evaluation settings. Figure 5 presents these prompt sensitivity values and provides a visual comparison of how strongly each model’s scoring behavior varied across the two prompt conditions.

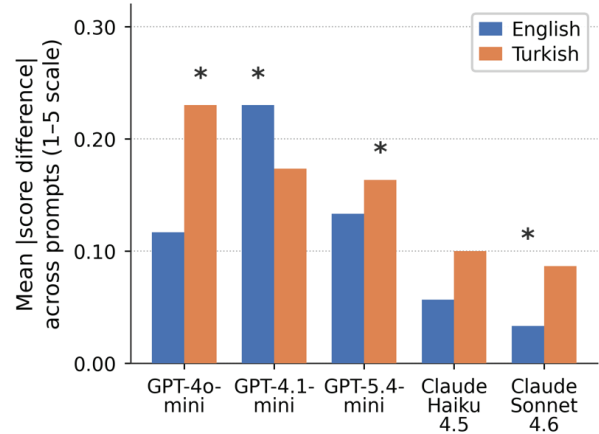


Figure 5. Prompt sensitivity across English and Turkish, measured as the mean absolute score difference between the simple and rubric prompts (1–5 scale). Error bars: 95% bootstrap CIs. Asterisks (*) denote Benjamini–Hochberg-adjusted significance across ten comparisons (Wilcoxon signed-rank test, $p < 0.05$).

As shown in Figure 5, prompt sensitivity differed substantially across models and languages. Among the OpenAI models, GPT-4o-mini exhibited the highest sensitivity in Turkish, whereas GPT-4.1-mini showed the highest sensitivity in English; both effects were statistically significant after Benjamini–Hochberg correction. GPT-5.4-mini also displayed noticeable prompt-related variation, with a statistically significant effect in Turkish but not in English. In contrast, the Claude models were generally less sensitive to prompt type. Claude Haiku 4.5 showed moderate sensitivity in Turkish in terms of mean score difference, but this effect was not statistically significant after correction. Claude Sonnet 4.6

exhibited the lowest sensitivity overall, although its small English prompt effect was statistically significant. These results indicate that prompt formulation influenced the scoring behavior of some models more strongly than others, particularly for GPT-4o-mini in Turkish, GPT-4.1-mini in English, and GPT-5.4-mini in Turkish, while Claude Sonnet 4.6 produced comparatively more stable evaluations across prompt conditions.

Finally, to determine whether the five judge models differed significantly in their scoring behavior, a Friedman test was conducted separately for each language, using the paired scores assigned to the same candidate responses across models. This analysis complements the preceding Wilcoxon tests by shifting the focus from language and prompt effects to differences among the judge models themselves. The results are presented in Table 7.

Table 7. Friedman test results for differences among the five judge models, reported separately for each language

Language	Friedman χ^2	df	p value
English	74.63	4	< 0.001
Turkish	23.68	4	< 0.001

As shown in Table 7, the Friedman test revealed statistically significant differences among the judge models in both English ($\chi^2(4) = 74.63$, $p < 0.001$) and Turkish ($\chi^2(4) = 23.68$, $p < 0.001$). These results confirm that the choice of judge model has a significant effect on the assigned scores in both languages. The larger test statistic obtained for English indicates that inter-model differences were more pronounced in the English evaluation setting, whereas the models produced comparatively more homogeneous scores in Turkish. Taken together with the classification and correlation results, this finding reinforces that judge-model selection is a decisive factor in multilingual LLM-as-a-Judge evaluation.

6. Discussion

The findings from this study provide promising evidence that large language models can be used as automatic evaluators in question-answering assessment tasks. The fact that Spearman’s correlation coefficients were high for all five models examined indicates that the models generally exhibited a ranking behavior consistent with the hybrid reference scoring framework. This result suggests that the LLM-as-a-Judge approach may serve as a viable evaluation method not only in high-resource languages such as English but also in morphologically rich languages such as Turkish, under controlled experimental conditions.

However, the performance of the models should not be evaluated solely based on correlation values. The findings reveal that, despite high rank correlation, there are significant differences among the models in terms of classification performance and macro-F1 scores. This indicates that, although the models can rank responses relatively accurately, they are not equally successful when converting them into discrete categories such as “correct,” “partially correct,” and “incorrect.” Therefore, in LLM-as-a-Judge systems, correlation metrics must be evaluated alongside classification-based metrics such as accuracy, macro-F1, and Cohen’s kappa coefficient. This finding also shows that the thresholding and calibration strategies used to convert model scores into class labels can have a decisive impact on the results. In this study, the proxy score for partially correct responses was set to 3 to ensure consistency with the score-to-label conversion rule. Nevertheless, because the “partially correct” category is more

ambiguous than the clearly correct and incorrect categories, future analyses should examine alternative score-to-label thresholding schemes to test the robustness of classification results.

In terms of model comparisons, Claude Sonnet 4.6 demonstrated the most balanced performance. This model showed high classification accuracy in both English and Turkish results, exhibited only limited differences between the languages, and demonstrated a relatively stable structure in the face of prompt changes. GPT-4o-mini, on the other hand, achieved the highest Spearman correlation on the English dataset but showed a more pronounced decline in Turkish. This result should be interpreted with caution because GPT-4o-mini was used both to generate candidate responses and as one of the judge models, which may introduce same-model or self-preference bias. This finding indicates that, in multilingual evaluation scenarios, model selection should not be based solely on the highest performance in a single language, but should also take into account cross-lingual consistency and robustness. However, it should be noted that these comparisons are limited by the dataset used, the candidate response generation method, the prompt design, and the scoring protocol.

Another notable finding of the study is the weaker class-level discrimination observed for GPT-5.4-mini. Despite its competitive Spearman correlation values, this model showed lower classification performance, particularly for the “partially correct” class. The lower macro-F1 score suggests that a substantial portion of the classification errors was associated with this intermediate category. This suggests that while the model preserves the relative ranking of responses, it may not use the scoring scale with sufficient calibration for fine-grained classification. Therefore, although GPT-5.4-mini appears suitable for ranking-based evaluations, it should be used with caution in three-class classification tasks. In particular, when evaluating the “partially correct” class, it may be necessary to calibrate the score ranges and class thresholds on a per-model basis.

The results regarding language differences and prompt sensitivity also highlight the importance of calibration when using LLM-as-a-Judge. The fact that a significant difference was found only between the English and Turkish scores for GPT-5.4-mini suggests that some models may exhibit different scoring tendencies when evaluating Turkish responses. Furthermore, the fact that prompt sensitivity is more pronounced in OpenAI models suggests that evaluation outputs may be influenced by the way prompts are phrased. For this reason, particularly in Turkish evaluation scenarios, prompt language, rubric details, and scoring criteria should be treated as fundamental components of the experimental design. It should also be noted that the simple and rubric-based prompts differed slightly in the verbal descriptions of the scoring scale, which may have contributed to the observed prompt-sensitivity effects. These results suggest that, in multilingual LLM-as-a-Judge systems, a single general prompt or fixed scoring scale may not be equally suitable for every model and language.

Overall, the findings indicate that the LLM-as-a-Judge approach shows promise for evaluating Turkish question-answering systems; however, for reliable use, model selection, language effects, prompt design, and score calibration must be evaluated together. Since the reference scoring framework was only partially based on manual human scoring, the findings should not be interpreted as direct evidence of full alignment with human judgment, but rather as preliminary evidence of agreement with a partially human-annotated and partially

proxy-based reference structure. Furthermore, although a total of 6,000 model judgments were obtained, the number of fundamental independent examples was 100 question-answer pairs. The remaining observations were repeated evaluations derived from these same underlying examples across models, languages, response categories, and prompt conditions. Consequently, observations originating from the same question may exhibit structural dependence because they share the same question, context, and reference answer. Therefore, the Wilcoxon, Friedman, and observation-level bootstrap results should be interpreted with this dependence structure in mind. Another limitation is the possible same-model bias caused by using GPT-4o-mini both as the candidate-response generator and as one of the judge models. Future studies should separate the response-generation and evaluation roles more strictly by using a generator model that is not included among the judge models or by generating candidate responses from multiple independent models. In future work, the number of examples and the coverage of human annotations will be increased, and multiple independent annotators will be included to provide stronger evidence of reliability. Future studies should also examine model behavior across different task types and develop language-specific calibration strategies aimed specifically at more reliably distinguishing “partially correct” responses. Mixed-effects models or question-level grouped bootstrap approaches could also be used to draw more precise conclusions. The 100 examples were selected through deterministic sequential truncation rather than random or stratified sampling. Consequently, the sample may not fully represent the broader distribution of question types, topics, and difficulty levels in the complete aligned SQuAD-SQuAD-TR pool. Future studies should therefore evaluate the robustness of the findings using larger randomly selected or stratified samples.

7. Conclusion

This study systematically investigated the feasibility of using large language models as automatic evaluators in English and Turkish question-answering assessments. The findings reveal that all models examined exhibited a high degree of rank correlation with the hybrid reference scores and that the LLM-as-a-Judge approach offers a promising evaluation method even in a morphologically complex language such as Turkish, under certain experimental conditions.

The results show that model selection is a determining factor in multilingual evaluation performance. Claude Sonnet 4.6 emerged as the most balanced model in terms of classification performance, cross-language consistency, and relative stability against prompt changes. While GPT-4o-mini produced the highest rank correlation in the English evaluation, Claude Sonnet 4.6 and GPT-5.4-mini shared the highest Spearman correlation in the Turkish evaluation, and Claude Sonnet 4.6 achieved the highest correlation in the combined English-Turkish results. However, the fact that GPT-5.4-mini struggled to distinguish the “partially correct” category clearly demonstrates that high correlation values alone are not a sufficient indicator of reliability.

These findings demonstrate that, for LLM-as-a-Judge systems to be used reliably, correlation-based metrics must be evaluated in conjunction with classification performance, language effects, prompt sensitivity, and score calibration. The study provides an empirical foundation based on preliminary findings regarding the use of multilingual LLM evaluators for Turkish question-answering evaluation and offers guiding

insights for future research with larger datasets and broader human annotation coverage. Since the reference scoring framework was only partially based on manual human scoring, the findings should not be interpreted as direct evidence of full alignment with human judgment, but rather as preliminary evidence of agreement with a partially human-annotated and partially proxy-based reference structure. Future work may extend this analysis by increasing the number of examples, expanding the coverage of human annotations, and including multiple independent annotators to provide stronger evidence of reliability.

8. References

- [1] D. Li, B. Jiang, L. Huang, A. Beigi, C. Zhao, Z. Tan, A. Bhattacharjee, Y. Jiang, C. Chen, T. Wu, K. Shu, L. Cheng, and H. Liu, "From generation to judgment: Opportunities and challenges of LLM-as-a-Judge," in *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, 2025, pp. 2757–2791. doi: 10.18653/v1/2025.emnlp-main.138.
- [2] A. Türkmen and K. E. Keleş, "Whom to trust? Analyzing the divergence between user satisfaction and LLM-as-a-Judge in e-commerce RAG systems," in *Proceedings of the First Workshop on Multilingual Multicultural Evaluation*, 2026. pp. 189–195. doi: 10.18653/v1/2026.mme-main.12.
- [3] A. Szymanski, N. Ziems, H. A. Eicher-Miller, T. J.-J. Li, M. Jiang, and R. A. Metoyer, "Limitations of the LLM-as-a-Judge approach for evaluating LLM outputs in expert knowledge tasks," in *Proceedings of the 30th International Conference on Intelligent User Interfaces (IUI '25)*, 2025. doi: 10.1145/3708359.3712091.
- [4] C. Toraman et al., "TurkBench: A benchmark for evaluating Turkish large language models," in *Proceedings of the Second Workshop on Natural Language Processing for Turkic Languages*, 2026, pp. 126–154. doi: 10.18653/v1/2026.sigturk-1.12.
- [5] J. Yu, S. Sun, X. Hu, J. Yan, K. Yu, and X. Li, "Improve LLM-as-a-Judge ability as a general ability," in *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, 2025, pp. 14099–14115. doi: 10.18653/v1/2025.emnlp-main.712.
- [6] H. Huang, X. Bu, H. Zhou, Y. Qu, J. Liu, M. Yang, B. Xu, and T. Zhao, "An empirical study of LLM-as-a-Judge for LLM evaluation: Fine-tuned judge model is not a general substitute for GPT-4," in *Findings of the Association for Computational Linguistics: ACL 2025*, 2025, pp. 5880–5895. doi: 10.18653/v1/2025.findings-acl.306.
- [7] C. Yamasaki, K. Sugahara, Y. Nagi, and K. Okamoto, "Function-based labels for complementary recommendation: Definition, annotation, and LLM-as-a-Judge," *Pattern Recognition Letters*, vol. 200, pp. 8–15, 2026. doi: 10.1016/j.patrec.2025.11.042.
- [8] T. Khan, M. Rafiq, and A. Michalas, "AI on Trial: LLM-as-a-Judge for private and reliable clinical decision-making," in *2025 Annual Computer Security Applications Conference Workshops (ACSAC Workshops)*, 2025, pp. 256–262. doi: 10.1109/ACSACW69556.2025.00033.
- [9] B. Bellard, X. C. Liu, R. Lohani, S. Raza, P. E. Velmovitsky, Q. Pham, and S. Saha, "Development and evaluation of an agentic RAG with embedded LLM-as-a-Judge for heart failure management: A pilot study," in *2025 IEEE International Conference on Big Data (BigData)*, 2025, pp. 7770–7774. doi: 10.1109/BigData66926.2025.11402164.
- [10] R. Mahadik, A. Mandal, S. Tripathi, and S. Jaiswal, "LLM evaluations for emotional expressiveness and factual consistency in medical dialogue systems using LLM-as-a-Judge," in *2025 International Conference on Responsible, Generative and Explainable AI (ResGenXAI)*, 2025. doi: 10.1109/ResGenXAI64788.2025.11344029.
- [11] M. I. H. Abir, N. Ferdous, A. Tasnim, and N. Mustaqeem, "Assessing large language models on zero-shot poetry summarization using ROUGE, BERT, and LLM-as-a-Judge metrics," in *2026 5th International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE)*, 2026. doi: 10.1109/ICECTE69292.2026.11429378.
- [12] G. Crupi, R. Tufano, A. Velasco, A. Mastropaolo, D. Poshyvanyk, and G. Bavota, "On the effectiveness of LLM-as-a-Judge for code generation and summarization," *IEEE Transactions on Software Engineering*, vol. 51, no. 8, pp. 2329–2345, 2025. doi: 10.1109/TSE.2025.3586082.
- [13] D. Chen, R. Chen, S. Zhang, Y. Wang, Y. Liu, H. Zhou, Q. Zhang, Y. Wan, P. Zhou, and L. Sun, "MLLM-as-a-Judge: Assessing multimodal LLM-as-a-Judge with vision-language benchmark," in *Proceedings of the 41st International Conference on Machine Learning*, vol. 235, 2024, pp. 6562–6595. doi: 10.5555/3692070.3692324.

Özgeçmişler



Dr. Ilkay Yelmen is an Assistant Professor in the Department of Software Engineering, Faculty of Engineering and Natural Sciences, Istinye University, Türkiye. He received his B.Sc. degree in Software Engineering from Istanbul Aydın University in 2013, his M.Sc. degree in Computer Engineering in 2016, and his Ph.D. degree in Computer Engineering in 2023 from the same university. His academic and research interests focus on computer science, including machine learning, deep learning, natural language processing and large language models.