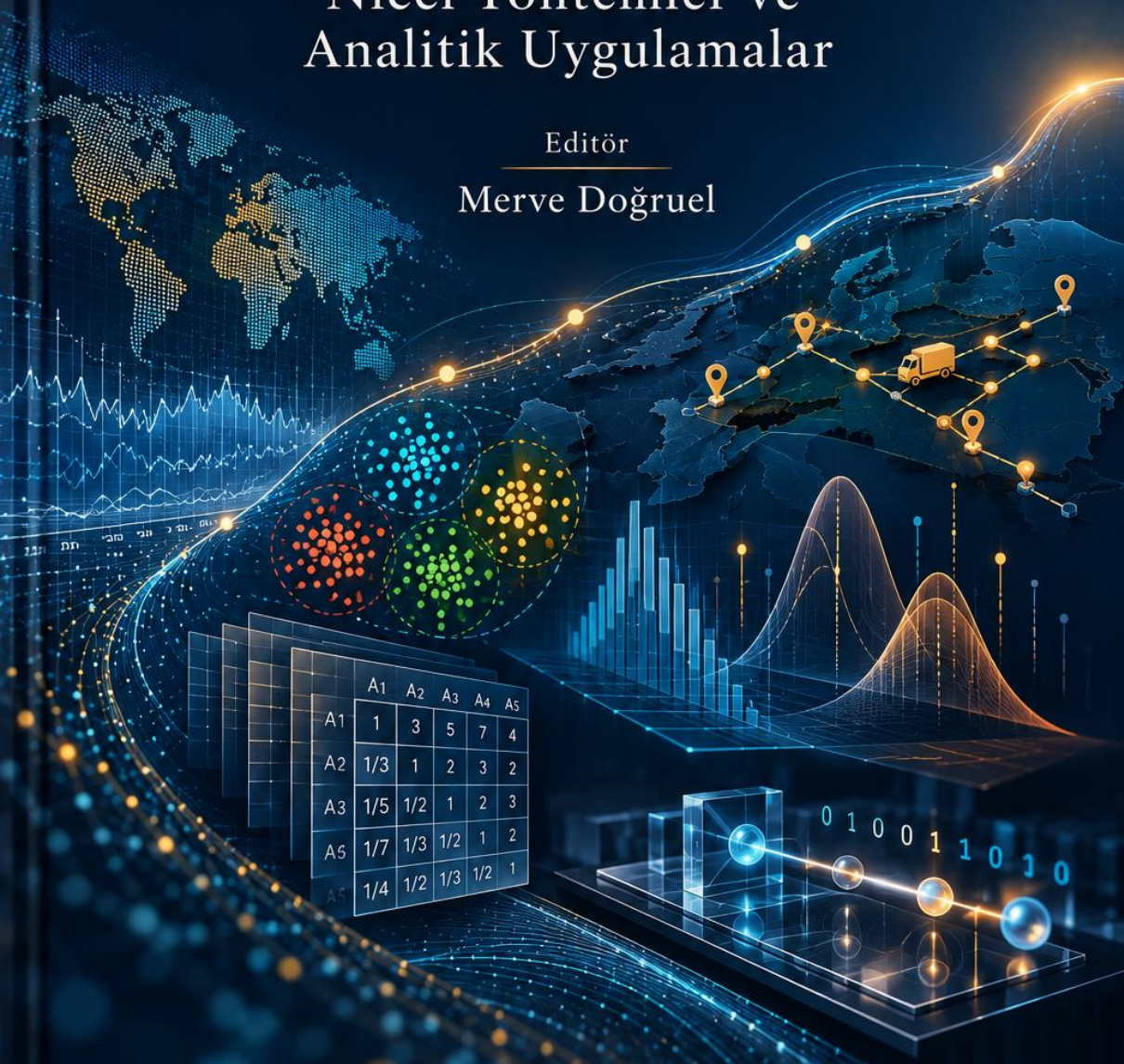
VERİDEN KARARA

Nicel Yöntemler ve Analitik Uygulamalar

Editör

Merve Doğruel



yaz
yayınları

VERİDEN KARARA

Nicel Yöntemler ve Analitik Uygulamalar

Editör

Dr. Öğr. Üyesi Merve DOĞRUEL

İstanbul Esenyurt Üniversitesi
İşletme ve Yönetim Bilimleri Fakültesi
Yönetim Bilişim Sistemleri Bölümü

yaz yayınları

2026

yaz
yayınları

VERİDEN KARARA

Nicel Yöntemler ve Analitik Uygulamalar

Editör: Dr. Öğr. Üyesi Merve DOĞRUEL
İstanbul Esenyurt Üniversitesi
İşletme ve Yönetim Bilimleri Fakültesi
Yönetim Bilişim Sistemleri Bölümü

ORCID: 0000-0003-2299-7182

© YAZ Yayınları

Bu kitabın her türlü yayın hakkı Yaz Yayınları'na aittir, tüm hakları saklıdır. Kitabın tamamı ya da bir kısmı 5846 sayılı Kanun'un hükümlerine göre, kitabı yayınlayan firmanın önceden izni alınmaksızın elektronik, mekanik, fotokopi ya da herhangi bir kayıt sistemiyle çoğaltılamaz, yayınlanamaz, depolanamaz.

E_ISBN 978-625-8926-58-3

Ekim 2026 – Afyonkarahisar

Dizgi/Mizanpaj: YAZ Yayınları

Kapak Tasarım: YAZ Yayınları

YAZ Yayınları. Yayıncı Sertifika No: 73086

M.İhtisas OSB Mah. 4A Cad. No:3/3
İscehisar/AFYONKARAHİSAR

www.yazyayinlari.com

yazyayinlari@gmail.com

ÖNSÖZ

Günümüzde karar problemleri; artan veri hacmi, farklılaşan veri yapıları, belirsizlik ve çok sayıda kısıtın eş zamanlı olarak değerlendirilmesi gerekliliği nedeniyle giderek daha karmaşık bir nitelik kazanmaktadır. Bu değişim, karar süreçlerinde yalnızca veriye ulaşmayı değil, verinin uygun nicel yöntemlerle analiz edilmesini, anlamlı örüntülerin ortaya çıkarılmasını, alternatiflerin değerlendirilmesini ve elde edilen bulguların karar süreçlerini destekleyecek bilgiye dönüştürülmesini gerekli kılmaktadır. Bu bağlamda Nicel Karar Verme Yöntemleri, farklı problem yapılarını sistematik ve analitik bir çerçevede ele almaya olanak sağlayan önemli bir çalışma alanı olarak öne çıkmaktadır.

Veriden Karara: Nicel Yöntemler ve Analitik Uygulamalar başlıklı bu kitap, Nicel Karar Verme Yöntemleri odağında farklı yöntemsel yaklaşımları ve uygulama alanlarını bir araya getirmektedir. Kitabın temel yaklaşımı, nicel yöntemleri yalnızca matematiksel, istatistiksel veya hesaplamalı teknikler olarak ele almak yerine; verinin analizinden kararın desteklenmesine uzanan bütüncül bir sürecin bileşenleri olarak değerlendirmektir. Bu doğrultuda kuramsal ve yöntembilimsel tartışmalar ile gerçek veri ve problem temelli uygulamalar aynı çatı altında sunulmaktadır.

Kitabın ilk bölümünde, K-ortalamalar kümeleme yönteminin temel sorunları ve yöntembilimsel sınırlılıkları, 2020–2026 dönemi literatürü üzerinden bütüncül bir yaklaşımla değerlendirilmektedir. İkinci bölümde, ülkelerin sürdürülebilir kalkınma performanslarının zamansal ve mekânsal yapısı, 2000–2025 dönemindeki gelişim örüntüleri temelinde Dinamik Zaman Bükme (DTW) tabanlı kümeleme, istatistiksel karşılaştırmalar ve coğrafi haritalama aracılığıyla incelenmektedir. Üçüncü bölümde, uluslararası lojistik operasyonlarında araç rotalama ve

maliyet optimizasyonu, kapasite ve zaman penceresi kısıtları ile yakıt ikmali kararlarını içeren matematiksel bir model çerçevesinde ele alınmaktadır. Dördüncü bölümde, metasezgisel algoritmaların zaman serisi hiperparametre optimizasyonundaki istatistiksel güvenilirliği, farklı çapraz doğrulama stratejileri üzerinden karşılaştırmalı olarak değerlendirilmektedir. Beşinci bölümde, lise öğrencilerinin alan seçimini etkileyen ölçütlerin öncelik yapısı, bireysel Analitik Hiyerarşi Süreci (AHP), bootstrap ve Bayesçi yaklaşımlar kullanılarak; belirsizlik, tutarsızlık, öncelik kararlılığı ve toplumsal farklılaşma boyutlarıyla incelenmektedir. Son bölümde ise kuantum tabanlı rastgele sayı üretimi, tek ve çoklu foton ölçümlerine dayalı kuantum devreleri aracılığıyla oluşturulan ikili sayı dizilerinin istatistiksel rastgelelik özellikleri üzerinden ele alınmaktadır.

Böylece kitap, Nicel Karar Verme Yöntemleri odağında, yöntembilimsel değerlendirmelerden veri temelli uygulamalara; kümeleme ve zaman serisi analizlerinden matematiksel modelleme ve optimizasyona; metasezgisel algoritmalardan çok ölçütlü karar verme ve Bayesçi yaklaşımlara; kuantum tabanlı hesaplama ve rastgele sayı üretimine kadar uzanan çok yönlü bir analitik çerçeve sunmaktadır. Bu yöntemsel çeşitlilik, farklı nitelikteki problemlerin uygun nicel araçlarla nasıl yapılandırılabileceğini, analiz edilebileceğini ve elde edilen sonuçların karar süreçlerinin desteklenmesinde nasıl kullanılabileceğini ortaya koymaktadır.

Kitabın bölümleri farklı veri yapıları, yöntemler ve uygulama alanları üzerine kurulmuş olmakla birlikte ortak bir düşünsel eksende birleşmektedir: veriden anlamlı bilgi üretmek, bu bilgiyi nicel yöntemlerle değerlendirmek ve karar süreçlerini destekleyen analitik çıktılara dönüştürmek. Kuramsal tartışmaların uygulamalı çalışmalarla birlikte sunulması; yöntemlerin yalnızca teknik özelliklerinin değil, varsayımlarının, sınırlılıklarının, güvenilirliklerinin ve farklı problem

alanlarındaki kullanım biçimlerinin de değerlendirilmesine olanak sağlamaktadır.

Bu yönüyle eserin, başta Nicel Karar Verme Yöntemleri olmak üzere karar bilimleri, veri analitiği, yöneylem araştırması, optimizasyon ve ilişkili alanlarda çalışan araştırmacılar ile lisansüstü öğrenciler için yöntemsel ve uygulamalı bir kaynak oluşturması amaçlanmaktadır.

Kitabın ortaya çıkmasına bilimsel çalışmalarıyla katkı sağlayan tüm bölüm yazarlarına teşekkür eder; eserin araştırmacılara, öğrencilere ve nicel karar verme yöntemleriyle ilgilenen tüm okuyuculara yararlı olmasını dilerim.

Dr. Merve Doğruel

İstanbul Esenyurt Üniversitesi
İşletme ve Yönetim Bilimleri Fakültesi
Yönetim Bilişim Sistemleri Bölümü

Editör

İstanbul, 2026

İÇİNDEKİLER

K-Ortalamalar Kümeleme Yönteminin Sorunsallarına Bütünsel Bir Bakış: 2020–2026 Literatürü Özelinde Yöntembilimsel Olarak Değerlendirilmesi 1
Özge EREN & Çiğdem ÖZARI

Sürdürülebilir Kalkınma Performansının Zamansal ve Mekânsal Analizi: Dinamik Zaman Bükme Tabanlı Kümeleme ve Coğrafi Haritalama 33
Merve DOĞRUEL

Kapasite Kısıtlı ve Zaman Pencereli Araç Rotalama Problemi ile Maliyet Optimizasyonu: Lojistik Sektöründe Bir Uygulama 59
Emre EKİN & Bahar IŞILAR

Metasezgisel Algoritmaların Zaman Serisi Hiperparametre Optimizasyonundaki İstatistiksel Güvenilirliği: Çapraz Doğrulama Stratejilerinin Karşılaştırmalı Analizi 89
Burcu KARTAL

Liselerde Alan Seçimi Ölçütlerinin Bireysel AHP, Bootstrap ve Bayesçi Yaklaşımlarla Analizi: Öncelik Kararlılığı, Tutarsızlık ve Toplumsal Farklılaşma 116
Ramazan PEHLİVAN

Tek ve Çoklu Foton Ölçümleri ile Kuantum Rastgele Sayı Üretimi 138
Engin KARAMAN

"Bu kitapta yer alan bölümlerde kullanılan kaynakların, görüşlerin, bulguların, sonuçların, tablo, şekil, resim ve her türlü içeriğin sorumluluğu yazar veya yazarlarına ait olup ulusal ve uluslararası telif haklarına konu olabilecek mali ve hukuki sorumluluk da yazarlara aittir."

K-ORTALAMALAR KÜMELEME YÖNTEMİNİN SORUNSALLARINA BÜTÜNSEL BİR BAKIŞ: 2020–2026 LİTERATÜRÜ ÖZELİNDE YÖNTEMBİLİMSEL OLARAK DEĞERLENDİRİLMESİ

Özge EREN[1]

Çiğdem ÖZARI[2]

1. GİRİŞ

K-ortalamalar kümeleme yöntemi, MacQueen (1967) tarafından önerilen etiketlenmemiş gözlemleri (birimleri) önceden belirlenmiş sayıda kümeye (gruba) ayıran denetimsiz öğrenme yöntemidir. Hızı, yorumlanabilirliği ve hemen her yazılım ortamında hazır bulunması, karmaşık veri yapılarını her zaman temsil edememesine rağmen onu çoğu uygulamada ilk başvurulan araç hâline getirmektedir. Buna karşılık yöntemin başlangıç merkezlerine olan duyarlılığı, küme sayısının dışarıdan verilmesi zorunluluğu, aykırı (uç) değerlere olan kırılganlığı, küresel olmayan yapılar dışındaki ayırımda zorlanması, yüksek boyutta uzaklıkların ayırt ediciliğini yitirmesi ve sonuçların geçerliliğine ilişkin belirsizlikler günümüzde de güncelliğini koruyan başlıklardır. Jain'in (2010) "k-ortalamalardan sonraki elli yıl" değerlendirmesinde sıraladığı sorun başlıklarının 2020 tarihli kapsamlı tarama (Ahmed, Seraj ve Islam, 2020) ile 2023 tarihli geniş derlemede (Ikotun, Ezugwu, Abualigah, Abuhaija ve

[1] Profesör Doktor, İstanbul Esenyurt Üniversitesi, Uygulamalı Bilimler Fakültesi, ORCID: 0000-0002-3850-818X.

[2] Doçent Doktor, İstanbul Aydın Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, ORCID: 0000-0002-2948-8957.

Heming, 2023) büyük ölçüde aynı kalması, alanın durağanlığına değil, bu sorunların matematiksel olarak çözülemeyen ancak iyileştirilebilen durumlar olmasına bağlıdır.

Bu kitap bölümün amacı, k-ortalamalar kümeleme sorunsallarını tek tek yeniden anlatmak değil, 2020–2026 dönemi literatürünün bu sorunsallara nasıl baktığını yöntembilimsel olarak değerlendirmektir. Merkezî soru şudur: Dönemin belli başlı araştırmaları sorunsalları hâlâ birbirinden bağımsız başlıklar olarak mı ele almaktadır, yoksa bunlar arasındaki etkileşimi gören bütünsel bir bakış geliştirmiş midir? Bu soruya yanıt verebilmek için her sorunsal üç düzeyde okunmuştur: (1) sorunun matematiksel kaynağı, yani hangi tanımın veya amaç fonksiyonu teriminin sorunu ürettiği; (2) önerilen çözüm ailelerinin biçimsel içeriği; (3) bu çözümlerin beraberinde getirdiği yeni varsayımlar, parametreler ve normatif kararlar. Bu üçlü okuma, "şu yöntem şu sorunu çözdü" biçimindeki doğrusal anlatıyı, "şu yöntem şu koşullarda şu ölçütü iyileştirdi ve karşılığında şu bağımlılığı getirdi" biçimindeki koşullu bir değerlendirmeye dönüştürmeyi amaçlamaktadır.

Gerçekleştirilen literatür incelemesi de dört kanıt tabakasına dayandırılmaktadır. Birinci tabakada amacı doğrudan k-ortalamalar veya kümeleme literatürünü taramak olan çalışmalar yer almaktadır (Ahmed vd., 2020; Ashabi, Sahibuddin ve Salkhordeh Haghighi, 2020; Ezugwu vd., 2022; Ikotun, Ezugwu, Abualigah, Abuhaija ve Heming, 2023; Mussabayev ve Mussabayev, 2024). İkinci tabaka, belirli bir sorunsala yöntem öneren araştırmalardan oluşur. Üçüncü tabaka, kümeleme araştırmasının değerlendirme kültürünü inceleyen yöntembilimsel ve meta-bilimsel çalışmalardır (Hennig, 2015; Akhanli ve Hennig, 2020; Gagolewski, Bartoszuk ve Cena, 2021; Hennig, 2022; Ullmann, Beer, Hünemörder, Seidl ve Boulesteix, 2023). Dördüncü yani son tabaka ise parametresiz ve dağılımsal sağlam k-ortalamalar gibi güncel ön baskılardır; bunlar yerleşmiş

sonuçlar olarak değil, gelişmekte olan araştırma yönleri olarak yorumlanmıştır (Mahon ve Lapata, 2025; Malik, Kargin ve Hassibi, 2026).

Kitap bölümünün anlatış sırası ise şu şekilde düzenlenmiştir. İlk olarak, sorunsalların tamamının kaynağı olan amaç fonksiyonunu ve yapısal sınırlar anlatılmaktadır. Sonrasında yaklaşık on sorunsal, her birinde yalnızca matematiksel kaynak ile çözümün getirdiği yeni bağımlılığı gösterecek kadar kısa biçimde ele alınmış ve nihayetinde son kısımda, bölümün asıl katkısı olan bütünsellik çözümlemesini sunulmaktadır.

2. MATEMATİKSEL ÇERÇEVE: AMAÇ FONKSİYONU VE YAPISAL SINIRLAR

Klasik k-ortalamalar problemi, n gözlemden oluşan $X = \{x_1, \dots, x_n\} \subset \mathrm{R}^d$ veri kümesini k adet kümeye ayırarak kendi oluşturduğu küme içi kareler toplamını (within-cluster sum of squares, WCSS) en aza indirmeyi amaçlar ve Formül (1) ile ana kurgusunu kurar.

$$J(C, M) = \sum_{i=1}^{k} \sum_{x \in C_i} \| x - \mu_i \|_2^2, \qquad C = \{C_1, \dots, C_k\},\ M = \{\mu_1, \dots, \mu_k\} \qquad (1)$$

K-ortalamalar problemi, Formül (1)'de tanımlanan amaç fonksiyonunun en küçüklenmesine dayanır. Ancak tüm olası küme atamalarının doğrudan incelenmesi hesaplama açısından pratik değildir. Bu nedenle uygulamada genellikle Lloyd algoritması kullanılır. Lloyd algoritması, gözlemlerin kümelere atanması ve küme merkezlerinin güncellenmesi adımlarını dönüşümlü olarak tekrar ederek küme içi kareler toplamını azaltır. Bu süreç, küme üyelikleri artık değişmediğinde veya amaç

fonksiyonundaki iyileşme önceden belirlenen bir eşik değerin altına düştüğünde sonlandırılır.

$$C_i^{(t)} = \left\{ x \in X:\ i = \arg\min_{1 \le j \le k} \| x - \mu_j^{(t)} \|^2 \right\}, \qquad \mu_i^{(t+1)} = \frac{1}{\left| C_i^{(t)} \right|} \sum_{x \in C_i^{(t)}} x \qquad (2)$$

Güncelleme adımındaki "ortalama" seçimi keyfî bir tasarım kararı değildir: sabit bir küme için kareli Öklidyen kaybı en aza indiren tek nokta aritmetik ortalamadır. Bu bağ, sağlamlık tartışmasının neden kayıp (uzaklık) fonksiyonunun değiştirilmesini gerektirdiğini de açıklar. Her iki adım J değerini artırmadığından ve olası bölümlemelerin sayısı sonlu olduğundan algoritma sonlu adımda durur; ancak bu yalnızca monoton azalmayı garanti eder, küresel en iyi çözüme ulaşıldığını göstermez.

Amaç fonksiyonu, küme merkezleri ve gözlem üyelikleri birlikte değerlendirildiğinde dışbükey değildir. Bu nedenle algoritmanın küresel optimumu bulması garanti edilemez. Ayrıca k-ortalamalar problemi, en iyi çözümün tüm olası küme atamaları arasında aranmasını gerektirdiğinden hesaplama kuramında NP-zor problemler sınıfında yer almaktadır; başka bir ifadeyle, problem büyüdükçe kesin en iyi çözümü makul sürede bulmak hızla güçleşmektedir (Aloise, Deshpande, Hansen ve Popat, 2009). Bu nedenle başlangıç merkezlerine duyarlılık yalnızca bir uygulama sıkıntısı değil, yöntemin matematiksel yapısından kaynaklanan temel bir zorluktur.

İkinci yapısal sınır veya zorluk, toplam kareler toplamının ayrışmasından doğar. $\bar{x}$ bir veri setinin genel ortalaması olmak üzere Formül (3)'teki gibi ifade edilmektedir.

$$\sum_{x \in X} \| x - \bar{x} \|^2 = \sum_{i=1}^{k} \sum_{x \in C_i} \| x - \mu_i \|^2 + \sum_{i=1}^{k} |C_i| \| \mu_i - \bar{x} \|^2 \qquad (3)$$

Kümeleme literatüründe toplam değişkenlik (Total Sum of Squares, TSS) iki bileşene ayrılmaktadır. İlki, kümelerin kendi içindeki değişkenliği gösteren Within-Cluster Sum of Squares (WCSS) ve diğeri de kümeler arasındaki değişkenliği gösteren Between-Cluster Sum of Squares (BCSS)'tir. Bu ilişki aşağıdaki gibi ifade edilir:

TSS = WCSS + BCSS

WCSS'yi küçültmek BCSS'yi büyütmeye eşdeğerdir ve kümeler arası terim küme büyüklüğüyle ağırlıklandırıldığından, amaç fonksiyonu büyük ve dengeli kümeleri yapısal olarak tercih eder. Küçük fakat bilimsel olarak anlamlı bir alt grubun büyük bir küme içinde eritilmesi, literatürde "tekdüzeleştirme etkisi" olarak adlandırılan bu eğilimin doğrudan sonucudur ve dengesiz küme büyüklüklerinde geçerlilik ölçütlerinin de yanıltıcı olabileceği gösterilmiştir (Xiong, Wu ve Chen, 2009). Dolayısıyla düşük WCSS, bulunan bölümlemenin alan bakımından "doğru" olduğu anlamına gelmez; optimizasyon başarısı ile yapısal doğruluk bu noktada ayrılmalıdır.

3. K-ORTALAMALAR SORUNSALLARIN KISA ENVANTERİ: SORUNLARIN KAYNAĞI , ÇÖZÜM VE ORTAYA ÇIKAN YENİ BAĞIMLILIKLAR

Bu kısımda derlemeye çalıştığımız yaklaşık on sorunsal, bütünsellik çözümlemesine zemin oluşturacak biçimde özetlenmiştir. Her başlıkta sorunun hangi neden ya nedenlerden

kaynaklandığı, öne sürülen çözüm ve varsa bu çözümün getirdiği yeni karar veya kararlar belirtilmiştir.

3.1. Uzaklık Ölçüsü ve Merkez Tanımı

Uygulamada sık rastlanan bir yanılgı, k-ortalamalar yönteminde kullanılan Öklidyen uzaklığın herhangi bir uzaklıkla serbestçe değiştirilebileceği varsayımıdır. Oysa güncelleme adımının geçerliliği uzaklık (kayıp) fonksiyonunun biçimine bağlıdır. Ortalamanın en iyi temsilci olduğu ölçü ailesi Bregman ıraksamaları olarak kabul edilir; kesin dışbükey ve türevlenebilir bir ϕ için

$$d_\phi(x,y) = \phi(x) - \phi(y) - \langle \nabla\phi(y),\, x - y \rangle, \qquad \underset{c}{\operatorname{argmin}}\, \mathrm{E}\left[d_\phi(x,c)\right] = \mathrm{E}[x] \qquad (4)$$

Formül (4), aritmetik ortalamanın kareli Öklidyen uzaklığın yanında Bregman ıraksamaları için de uygun bir merkez olduğunu göstermektedir. Ancak Manhattan uzaklığı kullanıldığında uygun merkez aritmetik ortalama değil bazen medyan, yalnızca gözlemler arasındaki uzaklıkların kullanıldığı durumlarda ise çoğunlukla medoid olur. Kosinüs benzerliği gibi farklı ölçülerde de standart k-ortalamalar güncellemesi ile doğrudan kullanılamaz; merkezin tanımı ve güncelleme adımı kullanılan ölçüye göre yeniden düzenlenmelidir. Bu nedenle yalnızca uzaklık ölçüsünü değiştirip aritmetik ortalamayı merkez olarak korumak, amaç fonksiyonu ile güncelleme adımı arasındaki uyumu bozabilir (Banerjee, Merugu, Dhillon ve Ghosh, 2005; Manthey ve Röglin, 2009).

Ayrıca uzaklık ölçüsünün seçimi, veriye ilişkin örtük bir geometrik varsayım taşır. Örneğin kareli Öklidyen uzaklık, değişkenlerin sürekli ve kümelerin merkez çevresinde benzer biçimde dağıldığını varsayarken; olasılık dağılımları veya metin verileri için göreli entropi gibi farklı Bregman ıraksamaları daha uygun olabilir. Bu nedenle uzaklık ölçüsü yalnızca teknik bir

tercih değil, verinin yapısına ilişkin modelleme kararıdır (Manthey ve Röglin, 2009).

3.2. Başlangıç Merkezleri ve Yerel En İyi Çözüm

K-ortalamalar algoritması, başlangıçta seçilen merkezlerden hareket ederek küme içi kareler toplamını azaltır. Ancak algoritma, tüm olası çözümleri araştırmak yerine yalnızca mevcut merkezlerin çevresinde iyileştirme yaptığı için ulaştığı sonuç yerel en iyi çözüm olabilir. Yerel en iyi çözüm, yakınındaki seçeneklerden daha iyi olmasına karşın veri seti için ulaşılabilecek en düşük WCSS değerini, yani küresel en iyi çözümü, vermeyebilir. Bu nedenle aynı veri seti ve aynı k değeriyle yapılan farklı çalıştırmalar, başlangıç merkezlerine bağlı olarak farklı kümeler üretebilir. Başlangıç merkezlerine duyarlılığın k-ortalamaların temel sorunlarından biri olduğu, farklı başlatma yöntemlerini geniş bir veri grubu üzerinde karşılaştıran çalışmalarla da gösterilmiştir (Celebi, Kingravi ve Vela, 2013). Bu sorunu azaltmak için geliştirilen k-means++ algoritması, ilk merkezi rastgele seçer; sonraki merkezlerin seçiminde ise hâlihazırda seçilmiş merkezlerden uzak olan gözlemlere daha yüksek olasılık verir. Seçim olasılığı şöyledir:

$$p(x) = \frac{D(x)^2}{\sum_{x' \in X} D\,(x')^2}, \qquad \mathrm{E}[J_{k\text{-means++}}] \leq 8(\ln k + 2)\, J_{\mathrm{OPT}} \qquad (5)$$

Bu güvence belirli bir çalıştırmanın iyi olacağını söylemez ve k büyüdükçe gevşer; bu nedenle çoklu başlatma ve en iyi çözümün seçilmesi yerleşik uygulama olmuştur. Fränti ve Sieranoja (2019), yeterli tekrar yapıldığında gelişmiş başlatmanın üstünlüğünün büyük ölçüde eridiğini, kümelerin iyi ayrışmadığı veride ise ne tekrarın ne de akıllı başlatmanın yeterli olduğunu göstermiştir; aynı ekibin kıyaslama çalışması başarının veri ayrışma düzeyine güçlü bağımlılığını belgeler (Fränti ve Sieranoja, 2018). Merkezleri aşamalı ekleyen global k-means

(Likas, Vlassis ve Verbeek, 2003) rastgeleliği azaltır, ancak bunu hesaplama bütçesiyle öder. Başlatma sorunu böylece bir "çözüm" değil, tekrar sayısı ile maliyet arasında bir ödünleşim olarak kalmaktadır. Arthur ve Vassilvitskii (2007), k-means algoritmasında başlangıç merkezlerinin daha etkin seçilmesini sağlayan k-means++ yöntemini önermiştir.

Özarı, Eren ve Alıcı (2019) k-ortalamalar yönteminde başlangıç merkezlerinin elde edilen küme yapıları üzerindeki etkisini inceleyerek, başlangıç merkezlerinin yalnızca veri setindeki gözlemlerden seçilmesi yerine, veri setinin yoğunluk yapısını dikkate alarak veri seti dışında yeni başlangıç merkezleri oluşturulmasına yönelik kolay uygulanabilir bir yöntem geliştirmişlerdir. Takip eden yıllarda Liao vd. (2024) başlangıç merkezlerinin seçimini makine öğrenme yöntemlerinden k-en yakın komşu yardımıyla yoğunluk tepelerine göre otomatik olarak belirleyen k-NNDP algoritmasını geliştirmiş ve uç değerlerin etkisini azaltmak amacıyla merkez güncellemelerinde ortalama yerine medyan kullanımını önermişlerdir. Nguyen Da, Cho ve Nguyen Thanh (2025), küme merkezlerinin seçimini yoğunluk temelli bir yaklaşımla iyileştirerek RBF sinir ağı içerisinde kullanmış ve elektrik yük tahmin performansını geliştirmeyi amaçlamıştır.

3.3. Küme Sayısının Belirlenmesi

K-ortalamalar kümeleme algoritması veride kaç küme olduğuna kendiliğinden karar vermez ve verinin kaç kümeden oluşacağına karşılık gelen k değeri arttıkça WCSS tekdüze azaldığından amaç fonksiyonu bu kararı veremez. Literatürde Dirsek yöntemi, Calinski–Harabasz (1974) ve Davies–Bouldin (1979) indeksleri ile Silhouette katsayısı (Rousseeuw, 1987) gibi ölçümlemeler en sık kullanılan küme sayısı belirleme ölçütlerdir. Her birinin tek tek anlatılması çok yer tutabilir. Bu sebeple birkaçına matematiksel olarak değinilecektir. Örneğin Silhouette,

her gözlemin kendi kümesine uyumunu $a(x)$ kendi kümesine ortalama uzaklık, $b(x)$ en yakın komşu kümeye ortalama uzaklık ilekarşılaştırır Formül 6 ile aşağıda gösterilmektedir:

$$s(x) = \frac{b(x) - a(x)}{\max\{a(x), b(x)\}} \in [-1,1], \qquad S(k) = \frac{1}{n}\sum_{x \in X} s\,(x) \qquad (6)$$

Literatürde bir de boşluk istatistiği olarak tanımlanan küme sayısı belirleme algoritması ise gözlenen küme içi dağılımı, kümesiz bir referans dağılımdan üretilen B örneklemle karşılaştırır (Tibshirani, Walther ve Hastie, 2001):

$$\mathrm{Gap}(k) = \frac{1}{B}\sum_{b=1}^{B} \log W_{kb}^{*} - \log W_k, \qquad \hat{k} = \min\{k:\ \mathrm{Gap}(k) \geq \mathrm{Gap}(k+1) - s_{k+1}\} \qquad (7)$$

Eşitlik (7) ile belirlenen uygun küme sayısı, yalnızca gözlenen veriye değil, karşılaştırmada kullanılan kümesiz referans dağılımın nasıl oluşturulduğuna da bağlıdır.

Küme sayısının kullanıcı tarafından önceden belirlenmesi sorununa karşı geliştirilen yöntemlerden biri X-means algoritmasıdır. X-means, belirli bir aralıktaki farklı k değerlerini değerlendirir ve gerekli gördüğü kümeleri ikiye bölerek uygun küme sayısını kendisi belirlemeye çalışır. Bölme kararında kullanılan Bayesçi Bilgi Ölçütü (BIC), modelin veriye uyumunu değerlendirirken artan küme sayısının yol açtığı model karmaşıklığını cezalandırır; böylece verinin gereğinden fazla sayıda kümeye ayrılmasını önlemeyi amaçlar (Pelleg ve Moore, 2000). Sinaga ve Yang (2020) ile Mahon ve Lapata (2025) tarafından geliştirilen daha yeni yaklaşımlar da kkk değerinin doğrudan kullanıcı tarafından verilmesi gereğini ortadan kaldırmaktadır. Ancak bu yöntemler seçimi tamamen varsayımsız hâle getirmez; açıkça belirlenen k değerinin yerini ceza katsayısı,

model karmaşıklığı veya tanım uzunluğu gibi başka karar ölçütleri alır. Nitekim Kleinberg'in (2002) imkânsızlık teoremi, ölçek değişmezliği, zenginlik ve tutarlılık gibi makul görünen üç özelliğin aynı kümeleme yöntemi tarafından birlikte sağlanamayacağını göstererek, her koşulda geçerli tek bir "doğru kkk" bulunamayacağını kuramsal olarak ortaya koymaktadır. Benzer biçimde Ikotun, Habyarimana ve Ezugwu (2025), on beş iç geçerlilik indeksinin başarısının veri yapısına göre değiştiğini bildirmiştir. Bu nedenle asıl soru yalnızca "en iyi k kaçtır?" değil, "hangi veri işleme adımları ve değerlendirme ölçütü altında elde edilen k kararlı, yorumlanabilir ve araştırmanın amacıyla uyumludur?" olmalıdır. Ju, Shi ve Wang (2024) ise k-ortalamalar algoritmasında uygun küme sayısının belirlenmesinde eski ölçütlerden olan Calinski–Harabasz ölçütünden yararlanarak risk düzeylerinin daha uygun biçimde sınıflandırılmasını sağlamıştır.

3.4. Aykırı Değerler, Gürültü ve Küçük Anlamlı Kümeler

Karesel kayıp (hesaplanan uzaklık) nedeniyle aritmetik ortalamanın kırılma noktası $1/n$'dir. Tek bir gözlemin sonsuza gönderilmesi merkezi de sonsuza taşımaktadır (Formül 8). Çözümler kaybın veya merkez tanımının değiştirilmesi üzerine kuruludur.

Budanmış k-ortalamalar olarak literatürde de görülen alternatif yöntem ise, verilerin (gözlemlerin) belirli bir α oranlık bölümünü amaç dışında bırakır:

$$J_\alpha = \min_{Z \subseteq X,\ |Z| = \lceil (1-\alpha)n \rceil} \sum_{x \in Z} \min_{1 \le i \le k} \| x - \mu_i \|^2 \qquad (8)$$

Budama oran olarak da ifade edilen α'nın kendisi bir model seçimi kararıdır. Eğer yüksek seçilirse anlamlı gözlemler kaybedilir, düşük seçilirse aykırı etki sürer (García-Escudero,

Gordaliza, Matrán ve Mayo-Iscar, 2010; García-Escudero ve Mayo-Iscar, 2024).

K-medoidler ve k-medyanlar ise başlangıç merkezi gerçek bir gözleme veya bileşen medyanına sabitleyerek kırılma noktasını yükseltir. Benzer dönemde dikkat çeken iki gelişme, birden fazla sınırlılığı birlikte ele alan düzenlemelerdir. Heidari, Daneshpour ve Zangeneh (2024) küresel olmayan kümeler ile aykırı değer duyarsızlığını, Khan vd. (2024) winsorlaştırma, Manhattan uzaklıklı Gauss çekirdeği ve değiştirilmiş boşluk istatistiğiyle üç sorunu tek çerçevede birleştirmektedir. Malik, Kargin, ve Hassibi (2026) ise gerçek dağılımın ampirik dağılım çevresindeki Wasserstein-2 belirsizlik kümesinde bulunduğunu varsayarak klasik amacı bir minimaks probleme dönüştürür:

$$\min_{\mu_1,\ldots,\mu_k} \sup_{Q\in \mathrm{B}_{\varepsilon}(\hat{P})} \mathrm{E}_{Q}\left[\min_{1\le i\le k} \| x-\mu_i \|^2\right] \qquad (9)$$

Bu yaklaşım aykırı değer, sınırlı örneklem ve dağılım kaymasını tek çerçevede ele alması bakımından önemlidir; ancak ε yarıçapı yeni bir model seçimi kararıdır ve çalışma 2026 tarihli bir ön baskıdır. Bu başlığın eleştirel öğesi bütün düzenlemelerin uzak gözlemin istenmeyen olduğu varsayımına dayanmasıdır. Oysa aykırı gözlem ile küçük fakat anlamlı bir küme arasındaki ayrım yalnızca geometrik özelliklere dayanılarak yapılamaz. Özellikle sağlık, eğitim ve bölgesel eşitsizlik çalışmalarında az sayıdaki gözlem, temizlenmesi gereken bir gürültüden ziyade korunması gereken anlamlı bir alt grubu temsil edebilir. Bu tür karmaşık ve doğrusal olmayan veri yapıları, klasik k-means yönteminin sınırlılıklarını daha belirgin hâle getirmektedir. Bu kapsamda Berlinski ve Maitra (2025), kernel k-means yaklaşımının hesaplama verimliliğini artırmaya yönelik iyileştirmeler önermiştir.

3.5. Küresel Olmayan, Örtüşen ve Farklı Yoğunluklu Yapılar

K-ortalamalar, her gözlemi kendisine en yakın küme merkezine atar. Bu basit kural, algoritmanın hangi biçimdeki kümeleri bulabileceğini de sınırlar. İki merkeze eşit uzaklıktaki noktaların oluşturduğu küme sınırı doğrusal olduğundan, k-ortalamalar veri uzayını dışbükey Voronoi hücrelerine ayırır. Aşağıdaki Geometrik kısıt, atama kuralının doğrudan sonucudur. İki merkez arasındaki sınır $\| x - \mu_i \|^2 = \| x - \mu_j \|^2$ eşitliğinden elde edilir ve açıldığında doğrusal bir ayırıcıya indirgenir:

$$2(\mu_j - \mu_i)^\top x = \| \mu_j \|^2 - \| \mu_i \|^2 \qquad (10)$$

K-ortalamalar bu nedenle veri uzayını dışbükey Voronoi hücrelerine böler ve iç içe, hilal biçimli veya farklı yoğunluklu yapıları temsil edemez. Çekirdek dönüşümleri, spektral yöntemler ve bulanık üyelik bu kısıtı gevşetir; ancak her birinin bedeli yeni parametrelerdir (bant genişliği, komşuluk yapısı, bulanıklık üsteli). Hennig çalışmasında (2022) dokuz yaygın yöntemi karşılaştırarak her bir yöntemin farklı küme kavramlarını işlemselleştirdiğinden aralarındaki seçim bir performans sorunu değil, araştırma sorusuyla uyum sorunudur.

3.6. Yüksek Boyut ve Temsil Seçimi

K-ortalamalar, gözlemleri merkezlere olan uzaklıklarına göre kümelere ayırdığı için uzaklıkların anlamlı ve ayırt edici olması gerekir. Ancak değişken sayısı arttıkça veri noktaları arasındaki uzaklıklar birbirine benzemeye başlar. Bağımsız ve aynı dağılımlı bileşenler gibi geniş bir koşullar sınıfı altında, bir gözlemin en yakın ve en uzak komşusuna olan uzaklıkları arasındaki göreli fark, boyut arttıkça olasılıkta sıfıra yakınsar (Formül11):

$$\frac{\mathrm{dist}_{\max}(d) - \mathrm{dist}_{\min}(d)}{\mathrm{dist}_{\min}(d)} \xrightarrow[d\to\infty]{p} 0 \qquad (11)$$

Yüksek boyutlu verilerde, kümeleme yapılmadan önce değişken sayısını azaltmak için sıklıkla temel bileşenler analizi (PCA) kullanılır. Ding ve He (2004), PCA ile k-ortalamalar arasında kuramsal bir ilişki bulunduğunu göstermiştir. Ancak PCA, kümeleri en iyi ayıran değişkenleri değil, verideki toplam değişimin büyük bölümünü koruyan bileşenleri seçer. Bu nedenle PCA sonucunda elde edilen bileşenler, küme yapısını her zaman doğru biçimde yansıtmayabilir. Seyrek k-ortalamalar, kümeleme açısından daha önemli değişkenlere daha fazla ağırlık vererek bu sorunu azaltmaya çalışır (Witten ve Tibshirani, 2010). Daha yeni yöntemler ise uygun veri gösterimini ve kümeleri aynı süreç içinde birlikte öğrenmektedir (Fard, Thonet ve Gaussier, 2020). Bununla birlikte, yüksek boyutlu veriyi iki boyuta indirgemek gözlemler arasındaki yakınlık ilişkilerini bozabilir (Chari ve Pachter, 2023). Dolayısıyla hangi değişkenlerin veya bileşenlerin kullanılacağı, yalnızca teknik bir hazırlık değil, elde edilecek kümeleri doğrudan etkileyen önemli bir karardır.

3.7. Ölçeklenebilirlik ve Hız–Kalite Ödünleşimi

Lloyd yinelemesinin adım başına maliyeti $O(nkd)$'dir. Dönemde üç strateji ayırt edilmektedir. Mini-yığın k-ortalamalar (Sculley, 2010) her adımda küçük bir rastgele alt kümeyle merkezleri kademeli günceller; çekirdek küme (coreset) yaklaşımı (Bachem, Lucic ve Krause, 2018) veriyi, her merkez seçimi için amaç değerini göreli olarak koruyan ağırlıklı bir özetle değiştirir:

$$\forall M: \quad |J_S(M) - J_X(M)| \leq \varepsilon J_X(M) \qquad (12)$$

Üçgen eşitsizliğine dayalı hızlandırma (Elkan, 2003) ise gereksiz uzaklık hesaplarını eleyerek aynı başlangıç altında Lloyd'un üreteceği sonucu korur. İlk iki strateji çözüm kalitesinden ödün vererek hız kazanırken üçüncüsü çözümü değiştirmez; ancak yüksek boyutta sınırların ayırt ediciliğini yitirmesi bu kazancı sınırlar. Mussabayev'in (2024) taraması, hız,

kalite ve sadelik arasında tek bir yöntemin bütün veri kümelerinde üstün olmadığını vurgular. Yöntembilimsel uyarı şudur: süre kıyasları ancak aynı donanım, aynı yakınsama eşiği ve aynı çözüm kalitesi düzeyinde anlamlıdır; farklı durdurma ölçütleriyle elde edilen üstünlükler erken durmanın sonucu olabilir. Bu ve benzeri çalışmalar sorunsallara getirdikleri farklı bakış açıları sebebiyle oldukça önemlidir.

3.8. Karışık Türde Verinin Varlığı ve Eksik Gözlem Sorunu

Elimizdeki veri ölçüm verilerinden oluştuğunda ortalama veya medyan hesabı anlamlı iken, kategorik değişkenler için ortalama anlamsız olmakatdır. k-prototipler (Huang, 1998) olarak literatürde yer alan yöntem , sayısal bileşenleri Öklidyen, kategorik bileşenleri uyuşmazlık sayısıyla ölçer ve ikisini bir γ ağırlığıyla birleştirir:

$$d(x,y) = \sum_{j\in \text{num}} \left(x_j - y_j\right)^2 + \gamma \sum_{j\in \text{cat}} \delta\left(x_j, y_j\right), \qquad \delta(a,b) = \begin{cases} 0, & a = b \\ 1, & a \neq b \end{cases} \qquad (13)$$

γ katsayısı iki ölçme düzeyi arasındaki değiş tokuşu belirler ve çoğu uygulamada gerekçelendirilmeden varsayılan değerinde bırakılır. Benzer sessiz kararlar sayısal değişkenlerin standartlaştırılması (her değişkene eşit ağırlık verme yönünde örtük bir tercih) ve eksik gözlemlerin ortalama ile doldurulmasıdır; ikincisi doldurulan noktaları merkeze yaklaştırarak küme içi varyansı yapay olarak düşürür ve WCSS temelli her ölçütü iyimser kılar. Bu üç kararın nadiren raporlanması, aynı veriden hareket eden iki araştırmacının farklı sonuçlara ulaşmasını açıklayan başlıca nedendir (Ahmad ve Dey, 2007; Foss, Markatou, Ray ve Heching, 2016).

3.9. Kümeleme Kararlılığı ve Yeniden Üretilebilirlik

Tek bir çalıştırmanın tek bir sonuç olarak raporlanması, k-ortalamaların temel belirsizliğini görünmez kılar. Kararlılık yaklaşımı ile , veriden bağımsız üretilen iki örneklem üzerindeki bölümlemeler arasındaki beklenen uyuşmazlığı ölçer:

$$\eta(k) = \mathrm{E}\left[d\left(C_k^{(1)}, C_k^{(2)}\right)\right], \qquad \hat{k} = \underset{k}{\operatorname{argmin}} \eta(k) \qquad (14)$$

Yeniden örneklemeye dayalı ortak üyelik matrisi ve küme düzeyinde Jaccard benzerliği aynı mantığın araçlarıdır; Ünlü ve Xanthopoulos (2019) uzlaşmalı kümeleme ile k tahminini bu yolla gerçekleştirmiştir. Bununla birlikte kararlılık tek başına bir doğruluk kanıtı değildir: Ben-David, von Luxburg ve Pál (2006) ile von Luxburg (2010), hatalı bir bölümlemenin de yüksek kararlılık sergileyebileceğini ve simetrik veride kararlılığın gerçek küme sayısını değil, simetrinin bozulduğu eşiği işaret ettiğini göstermiştir. Kararlılık ölçüleri bu nedenle doğrulama kanıtı olarak değil, belirsizliği açığa çıkaran tanılayıcı araçlar olarak okunmalıdır.

3.10. Adalet, Mahremiyet ve Yorumlanabilirlik

2020 sonrası literatürün dikkat çeken yönlerinden biri incelenen dönemin sonuna doğru , k-ortalamaların yalnızca geometrik başarısına değil, elde edilen kümelerin toplumsal ve etik sonuçlarına da odaklanmasıdır. Dönemin en belirgin yeniliği, geometrik bölümlemenin ötesinde normatif hedeflerin amaç fonksiyonuna taşınmasıdır. Sosyal olarak adilce bölümlenmiş k-ortalamalar (Ghadiri, Samadi ve Vempala, 2021), toplam maliyet yerine korunan gruplar üzerindeki en kötü ortalama maliyeti en aza indirir:

$$\Phi_{\text{adil}}(M) = \max_{g \in G} \frac{1}{|X_g|} \sum_{x \in X_g} \min_{1 \le i \le k} \| x - \mu_i \|^2 \qquad (15)$$

Eşitlik (15) Toplam uzaklık yerine yerine gruplar üzerinde maksimum alınmaktadır. Her kümenin demografik bileşimini kısıtlayan "fairlet" tanımı (Chierichetti, Kumar, Lattanzi ve Vassilvitskii, 2017) ise aynı kavramın farklı bir işlemselleştirmesidir ve genel olarak (15)'i gerektirmez; "adil kümeleme" ifadesi dolayısıyla tek başına belirleyici değildir. Alluri, Chen, Thesen, Jacobson ve Marrero (2026), yükseköğretim öğrencilerinin dijital ruh sağlığı verisinde grup eşitsizliklerinin önemli ölçüde azaltılabildiğini bildirmiştir; ancak tek bir uygulama alanından bütün adalet tanımlarına genelleme yapılamaz.

Mahremiyet boyutunda ise diferansiyel mahremiyetli k-ortalamalar, her yinelemede küme toplamlarına ve sayımlarına gürültü ekler (Stemmer ve Kaplan, 2018); mahremiyet bütçesi küçüldükçe gürültü büyür ve küçük kümelerin merkezleri orantısız biçimde bozulur, bu da mahremiyet ile azınlık gruplarının temsili arasında yapısal bir gerilim yaratır.

Yorumlanabilirlik boyutunda açıklanabilir k-ortalamalar (Dasgupta, Frost, Moshkovitz ve Rashtchian, 2020), bölümlemeyi k yapraklı bir karar ağacıyla sınırlandırır ve bunun bedelini biçimsel olarak ölçer:

$$\frac{J_{\text{ağaç}}}{J_{\text{OPT}}} = O(k^2) \qquad (16)$$

Fleissner, Vankadara ve Ghoshdastidar (2024) bu çerçeveyi çekirdek k-ortalamalara genişletmiştir. Üç başlık birlikte değerlendirildiğinde kalite, eşitlik, mahremiyet, açıklanabilirlik ve hesaplama maliyeti arasında bağlama bağlı ödünleşimler bulunduğu görülür; "ek kısıtlar verimliliği mutlaka düşürür" biçimindeki genel yargı ise savunulabilir değildir, çünkü "verimlilik" amaç değeri, çalışma süresi ve toplumsal fayda gibi ayrıştırılması gereken farklı anlamlara da gelmektedir.

4. 2020–2026 LİTERATÜRÜNDE BÜTÜNSEL BAKIŞ: YÖNTEMBİLİMSEL BİR ÇÖZÜMLEME

Bu ve benzeri algoritmalarda belirli bir veri grubunda belirli bir ölçütü iyileştirme karşılığında yeni bir parametre (α, ε, γ, referans dağılım, ceza katsayısı) gelmektedir. Bu parametreler birbirinden bağımsız değildir; budama oranı küme sayısı kararını, temsil seçimi geometri kararını, standartlaştırma aykırı değer tanımını değiştirir. Şimdi bu bölüm itibariyle araştırılan ise ilgili dönemin literatürünün bu bağımlılıkları görüp görmediğidir.

4.1. Bütünselliğin İşlemselleştirilmesi

Çalışmaların sorunlar arasındaki ilişkileri ne ölçüde dikkate aldığını değerlendirebilmek için "bütünsellik" kavramının açık ve gözlenebilir ölçütlerle tanımlanması gerekir. Bu araştırmada bütünsellik beş ölçüt üzerinden değerlendirilmiştir: bir çalışmada birden fazla sorunsalın birlikte ele alınması, sorunsallar arasındaki etkileşimlerin ve ödünleşimlerin açıklanması, küme sayısı ve model seçimi gibi kararların algoritmaya dâhil edilmesi, kullanılan değerlendirme ölçütlerinin ayrıca sorgulanması ve sonuçların yeniden üretilebilmesine yönelik raporlama standartlarının önerilmesi. Bu ölçütler aracılığıyla çalışmaların k-ortalamaların sorunlarını birbirinden bağımsız teknik eksiklikler olarak mı, yoksa birbirini etkileyen bir yöntemsel yapı içinde mi değerlendirdiği belirlenmiştir. İncelenen çalışmalar, söz konusu ölçütleri

karşılama düzeylerine göre parçalı yaklaşımdan bütüncül yaklaşıma uzanan beş düzeyde sınıflandırılmıştır.

Tablo 1. Bütünsellik olgunluk düzeyleri ve 2020–2026 döneminde temsil eden yaklaşımlar

Düzey	Tanım	Dönemde temsil eden yaklaşımlar
D0	Tek sorunsala tek teknik düzeltme; diğer sorunsallar sabit varsayılır	Gul ve Rehman (2023): ölçeklenebilir başlangıç merkezi seçimi; Hussayn ve Shah (2023): entropi temelli başlangıç seçimi
D1	Birden fazla sorunsalın tek yöntemde birlikte ele alınması	Khan vd. (2024): aykırı değer + geometri + k seçimi; Heidari vd. (2024): geometri + sağlamlık
D2	Model seçiminin yeni bir algoritma ile birleştirilmesi	Sinaga ve Yang (2020); Mahon ve Lapata (2025)
D3	Değerlendirme çerçevesinin ve ölçütlerin kendisinin sınanması	Akhanli ve Hennig (2020); Gagolewski vd. (2021); Hennig (2022); Ullmann vd. (2023); Ikotun vd. (2025)
D4	Çeşitli hedeflerin ve güvencelerin amaç fonksiyonuna taşınması	Malik vd. (2026); Alluri vd. (2026); açıklanabilir kümeleme çalışmaları

4.2. Dönemlendirme

Dönemin başında (2020–2021) baskın biçim envanter çıkarmaktır: kapsamlı taramalar sorunsalları sınıflandırmakta, ancak bunları birbirinden bağımsız başlıklar olarak listelemektedir (Ahmed vd., 2020; Ashabi vd., 2020). 2022–2023 arasında iki eğilim aynı anda güçlenir: bir yanda varyant üretimi hızlanır (Ezugwu vd., 2022; Ikotun vd., 2023), diğer yanda değerlendirme kültürüne yönelen meta-bilimsel eleştiri görünür hâle gelir (Gagolewski vd., 2021; Hennig, 2022; Ullmann vd., 2023). 2024–2025 döneminde çok sorunlu birleşik yöntemler (Heidari vd., 2024; Khan vd., 2024) ile model seçimini içselleştiren algoritmalar (Mahon ve Lapata, 2025) öne çıkar. 2026'ya gelindiğinde gündeme normatif ve garanti eksenleri eklenir: dağılımsal sağlamlık ve adalet yani bazı önemli gözlemlerin yok olmayacağının garantisidir (Malik vd., 2026;

Alluri vd., 2026). Ancak bu değerlendirmeler yapılırken çoğunluğun bu şekilde hareket ettiğini söyleyebiliriz.

Bu sıralama bir ilerleme öyküsü gibi okunabilir; ancak böyle bir okuma yanıltıcı olur. Erken dönem derlemeleri başlangıç, k seçimi, veri türü, aykırı değer, geometri, büyük veri ve değerlendirme sorunlarını zaten birlikte ele almaktadır. Son yıllarda otomatik k, dağılımsal sağlamlık ve adalet daha görünür hâle gelmiştir; fakat bu, eski problemlerin yerini yenilerinin aldığı anlamına gelmez. Daha ihtiyatlı ifade, klasik optimizasyon ve ölçeklenebilirlik gündemi sürerken modele ilişkin güvence, belirsizlik ve hakkaniyet tartışmalarının literatüre eklenmiş olmasıdır.

4.3. Dönem Çalışmalarının Eleştirel Kanıt Matrisi

Tablo 2 ile, bölümde tartışılan 2020–2026 dönemi çalışmalarını kanıt türü, katkı, eleştirel sınırlılık ve Tablo 1'deki olgunluk düzeyi bakımından özetlemektedir. Matrisin amacı bir sıralama yapmak değil, her çalışmanın hangi sorunsalı nasıl ele aldığını ve diğer eksenlerle bağ kurup kurmadığını görünür kılmaktır.

Tablo 2. 2020–2026 dönemi bazı çalışmaların katkı, sınırlılık ve olgunluk düzeyi matrisi

Çalışma (anıt türü)	Katkı	Eleştirel sınırlılık	Düzey
Ahmed vd. (2020), tarama	Sorunsalların kapsamlı envanteri ve deneysel kıyas	Sorunsallar bağımsız başlıklar olarak listelenir	D0
Ashabi vd. (2020), sistematik tarama	Varyantların sistematik sınıflaması	Değerlendirme çerçevesi sorgulanmaz	D0
Akhanli ve Hennig (2020), hakemli	Kalibre edilmiş ölçütlerin amaca göre ağırlıklı birleştirilmesi	Ağırlık seçimi açık normatif karar gerektirir	D3
Fard vd. (2020), hakemli	Temsil öğrenme ile k-ortalamaların uçtan uca birleştirilmesi	Temsil, varsayılan geometriye uyacak biçimde şekillenebilir	D1

Çalışma (anıt türü)	Katkı	Eleştirel sınırlılık	Düzey
Sinaga ve Yang (2020), hakemli	Açık k girdisini kaldıran denetimsiz formülasyon	Ceza katsayısı ve sönümleme kuralı örtük karar olarak kalır	D2
Dasgupta vd. (2020), hakemli	Karar ağacıyla açıklanabilir kümeleme ve maliyet sınırı	Yorumlanabilirlik $O(k^2)$ bedelle gelir	D4
Gagolewski vd. (2021), hakemli	İç geçerlilik ölçütlerinin uzman bölümlemeleriyle sınanması	Sonuçlar referans havuzuna bağlıdır	D3
Ghadiri vd. (2021), hakemli	En kötü grup maliyetini en aza indiren adil k-ortalamalar	Başlangıç ve k seçimi klasik hâliyle devralınır	D4
Ezugwu vd. (2022), tarama	Kümeleme algoritmalarının geniş taksonomisi	Kapsam genişliği standart kıyası zorlaştırır	D0
Hennig (2022), hakemli	Dokuz yöntemin ampirik karakterizasyonu	Yöntem seçimini araştırma sorusuyla uyum sorunu olarak kurar	D3
Gagolewski (2022), hakemli	Kümeleme için kıyaslama altyapısı	Benimsenmesi henüz sınırlıdır	D3
Ullmann vd. (2023), hakemli	Yeni algoritmaların değerlendirilmesinde iyimserlik yanlılığının belgelenmesi	Yöntem iddialarının okunma biçimini değiştirir	D3
Ikotun vd. (2023), tarama	Varyantlar ve büyük veri açısından kapsamlı derleme	Ortak deney protokolü sınırlıdır	D0
Heidari vd. (2024), hakemli	Küresel olmayan kümeler ve aykırı değer duyarsızlığının birlikte ele alınması	Genellenebilirlik ve maliyet ayrıca sınanmalıdır	D1
Khan vd. (2024), hakemli	Winsorlaştırma, çekirdek dönüşümü ve değiştirilmiş boşluk istatistiği	Eşik, çekirdek ve referans dağılım yeni bağımlılıklardır	D1
Fleissner vd. (2024), hakemli	Açıklanabilir kümelemenin çekirdek k-ortalamalara genişletilmesi	Çekirdek seçimi yorumlanabilirliği sınırlar	D4
Ikotun vd. (2025), hakemli	On beş geçerlilik indeksinin evrimsel çerçevede kıyası	Sonuçlar algoritmik çerçeveye bağlıdır	D3
Mahon ve Lapata (2025), ön baskı	MDL ile bölme-birleştirme, açık k girdisi kaldırılır	Parametresizlik varsayımsızlık değildir; bağımsız doğrulama sınırlı	D2
Malik vd. (2026), ön baskı	Wasserstein-2 belirsizlik kümesiyle dağılımsal sağlam k-ortalamalar	Yarıçap seçimi ve yeniliği nedeniyle doğrulama gerektirir	D4
Alluri vd. (2026), hakemli	Demografik gruplar arasında maliyet eşitliği hedefi	Tek uygulama alanından genelleme yapılamaz	D4

4.4. Yatay Bütünsellik, Dikey Kopukluk

Tablo 1 ve Tablo 2'nin birlikte gösterdiği görünüm iki yönlüdür. Bir yandan D1–D4 düzeylerinde 2020 öncesine göre belirgin bir yoğunlaşma vardır; bu, bütünsel bakışın gerçekten geliştiğinin kanıtıdır. Öte yandan bu düzeyler birbirini beslemeyen ayrı kümeler hâlindedir. D1'deki birleşik yöntemler değerlendirme sorununu (D3) çoğunlukla tek bir indekse bırakmakta; D2'deki parametresiz algoritmalar sağlamlık ve geometri sorunlarını sabit varsaymakta; D4'teki normatif çalışmalar ise başlangıç ve k seçimi sorunlarını klasik hâliyle devralmaktadır. Başka bir deyişle bütünsellik dikey değil, yatay gerçekleşmiştir: farklı araştırma toplulukları kendi eksenlerinde derinleşmiş, fakat eksenler arası bağ kurulmamıştır.

Bu kopukluğun en somut göstergesi, D3 düzeyindeki bulguların D1 ve D4 düzeyindeki yöntem önerilerine yansımamasıdır. Gagolewski vd. (2021) iç geçerlilik ölçütlerinin uzman bölümlemeleriyle uyumunu sınayarak ölçütlerin kendisini araştırma nesnesi hâline getirmiş; Akhanli ve Hennig (2020) tek bir evrensel indeks yerine, araştırma amacına göre ağırlıklandırılan kalibre edilmiş ölçütlerin birleştirilmesini önermiştir. Ağırlıkların belirlenmesi araştırmacının açık normatif tercihini gerektirir; bu bir kusur değil, şeffaflık koşuludur. Hennig'in (2015) "gerçek küme" kavramının araştırma amacına göreli olduğu savı da aynı yöndedir. Buna karşılık D1 düzeyindeki çok sorunlu yöntemler başarılarını hâlâ tek bir silhouette veya doğruluk değeriyle raporlamakta, böylece D3'ün temel uyarısını görmezden gelmektedir.

Bütünselleşme iddiasını sınırlayan üç gözlem vardır. Birincisi, dönem boyunca yayınların sayısal çoğunluğu hâlâ D0 düzeyindedir; tek bir zayıflığı hedefleyen varyant önerileri baskın biçim olmayı sürdürmektedir. İkincisi, yöntem aileleri arasında ortak ve standart bir deney protokolü yerleşmemiştir; veri

kümeleri, ön işleme zincirleri, başlangıç sayıları ve durdurma ölçütleri çalışmadan çalışmaya değişmektedir. Bu durum, Ullmann vd.'nin (2023) belgelediği iyimserlik yanlılığının yapısal zeminidir: karşılaştırma koşullarını seçen taraf yöntemi öneren taraf olduğunda, yeni yöntem lehine sistematik bir avantaj doğar. Bu bulgu, dönemin tek tek yöntem iddialarının okunma biçimini değiştiren en önemli meta-bilimsel katkıdır. Üçüncüsü, kıyaslama altyapısı (ortak veri havuzları, kalibre edilmiş ölçütler, açık kod) ancak dönemin sonlarına doğru kurumsallaşmaya başlamıştır (Gagolewski, 2022) ve benimsenmesi henüz sınırlıdır.

Bu üç gözlem birlikte değerlendirildiğinde, 2020–2026 döneminde bütünsel bakışın kısmi ve asimetrik biçimde geliştiği sonucuna varılmaktadır. Çok sorunu birlikte ele alan yöntemler, model seçimini içselleştiren algoritmalar ve değerlendirme kültürünü sınayan meta-bilimsel çalışmalar çoğalmıştır; buna karşılık başlangıç, k seçimi, sağlamlık, geometri ve geçerliliği tek bir amaç fonksiyonu altında güvenceye bağlayan birleşik bir çerçeve oluşmamıştır. Üstelik Kleinberg'in (2002) imkânsızlık sonucu ile Hennig'in (2015, 2022) görelilik savı birlikte okunduğunda, böyle bir çerçevenin saf teknik araçlarla kurulamayacağı da anlaşılmaktadır: her bütünleştirme, hangi küme kavramının ve hangi normatif hedefin öncelikli olduğuna ilişkin açık bir kararı gerektirir.

5. SONUÇ

2020–2026 literatürü, k-ortalamaların temel sorunlarının ortadan kalkmadığını, soruna özgü yöntemlerle yönetilmeye çalışıldığını göstermektedir. Bu bölümde gösterildiği üzere sınırlılıkların çoğu algoritmanın uygulanışında değil, amaç fonksiyonunun kuruluşundadır. Daha iyi başlatma yerel çözüm riskini azaltabilir, otomatik k yöntemleri açık kullanıcı girdisini sınırlayabilir, sağlam yöntemler aykırı değer ve dağılım

kaymasına karşı koruma sağlayabilir, adalet duyarlı hedefler grup maliyetlerini (uzaklık minimumluğu) dengeleyebilir. Ancak her iyileştirme yeni varsayımlar, eşikler, hesaplama maliyetleri veya normatif tercihler getirmektedir.

Bölümün asıl bulgusu, bu iyileştirmelerin dönem içinde nasıl örgütlendiğine ilişkindir. Bütünsel bakış gelişmiştir, fakat yatay biçimde: birleşik yöntemler, parametresiz algoritmalar, değerlendirme eleştirisi ve normatif hedefler ayrı araştırma topluluklarında olgunlaşmış, birbirinin bulgularını içselleştirmemiştir. Çok sorunlu yöntemlerin başarılarını hâlâ tek bir indeksle raporlaması ve değerlendirme koşullarının yöntemi öneren tarafça belirlenmesi, bu kopukluğun somut göstergeleridir.

Bu tablo, k-ortalamalar sonuçlarının nasıl yorumlanması gerektiğine ilişkin açık bir yöntembilimsel sonuç doğurmaktadır. Kümeleme çıktıları "algoritmanın bulduğu doğal kümeler" olarak değil, belirli bir veri temsili, ölçek, uzaklık, amaç fonksiyonu ve değerlendirme ölçütü altında üretilmiş analitik bir bölümleme olarak okunmalıdır. Tüm çalışmaları taradığımızda özellikle Hennig'in (2022) ilkesi burada belirleyicidir: yöntemler farklı küme kavramlarını işlemselleştirdiğinden, aralarındaki seçim bir performans sorunu değil, araştırma sorusuyla uyum sorunudur. Güvenilir kullanım, yöntemi eleştirmeden reddetmek veya sorgulamadan kabul etmek yerine, varsayımlarını görünür kılmayı, alternatifleri aynı protokol altında karşılaştırmayı ve belirsizliği raporlamayı gerektirir. Gelecek araştırmalar için en verimli yön, yeni bir varyant eklemekten çok, tarafımızca oluşturulmuş düzeyler arasında bağ kuran, yani çok sorunlu yöntemleri kalibre edilmiş ve bağımsız değerlendirme çerçeveleri altında sınayan çalışmalar olmalıdır.

KAYNAKÇA

Ahmad, A. ve Dey, L. (2007). A k-mean clustering algorithm for mixed numeric and categorical data. *Data & Knowledge Engineering*, *63*(2), 503–527. http://dx.doi.org/10.1016/j.datak.2007.03.016

Ahmed, M., Seraj, R. ve Islam, S. M. S. (2020). The k-means algorithm: A comprehensive survey and performance evaluation. *Electronics*, *9*(8), 1295. https://doi.org/10.3390/electronics9081295

Akhanli, S. E. ve Hennig, C. (2020). Comparing clusterings and numbers of clusters by aggregation of calibrated clustering validity indexes. *Statistics and Computing*, *30*(5), 1523–1544. https://doi.org/10.1007/s11222-020-09958-2

Alluri, P., Chen, Z., Thesen, T., Jacobson, N. C. ve Marrero, W. J. (2026). Fairness-aware K-means clustering in digital mental health for higher education students: A generalizable framework for equitable clustering. *JAMIA Open*, *9*(1), ooaf174. https://doi.org/10.1093/jamiaopen/ooaf174

Aloise, D., Deshpande, A., Hansen, P. ve Popat, P. (2009). NP-hardness of Euclidean sum-of-squares clustering. *Machine Learning*, *75*(2), 245–248. https://doi.org/10.1093/jamiaopen/ooaf174

Arthur, D. ve Vassilvitskii, S. (2007). k-means++: The advantages of careful seeding. *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms* (s. 1027–1035). Philadelphia: SIAM.

Ashabi, A., Sahibuddin, S. B. ve Salkhordeh Haghighi, M. (2020). The systematic review of K-means clustering algorithm. *Proceedings of the 2020 9th International*

Conference on Networks, Communication and Computing (ICNCC) (s. 13–18). New York: ACM. https://doi.org/10.1145/3447654.3447657

Bachem, O., Lucic, M. ve Krause, A. (2018). Scalable k-means clustering via lightweight coresets. *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (s. 1119–1127). New York: ACM. https://doi.org/10.1145/3219819.3219973

Banerjee, A., Merugu, S., Dhillon, I. S. ve Ghosh, J. (2005). Clustering with Bregman divergences. *Journal of Machine Learning Research, 6*, 1705–1749. https://dl.acm.org/doi/10.5555/1046920.1194902

Ben-David, S., von Luxburg, U. ve Pál, D. (2006). A sober look at clustering stability. G. Lugosi ve H. U. Simon (Ed.), *Learning Theory: COLT 2006* içinde (s. 5–19). Berlin: Springer.

Berlinski, J. D., & Maitra, R. (2025). Computational Improvements to the Kernel k k-Means Clustering Algorithm. *Statistical Analysis and Data Mining: An ASA Data Science Journal*, *18*(4), e70032. https://doi.org/10.1002/sam.70032Digital Object Identifier (DOI)

Caliński, T. ve Harabasz, J. (1974). A dendrite method for cluster analysis. *Communications in Statistics, 3*(1), 1–27.

Celebi, M. E., Kingravi, H. A., & Vela, P. A. (2013). A comparative study of efficient initialization methods for the k-means clustering algorithm. *Expert systems with applications*, *40*(1), 200-210. https://doi.org/10.1016/j.eswa.2012.07.021

Chari, T. ve Pachter, L. (2023). The specious art of single-cell genomics. *PLoS Computational Biology, 19*(8), e1011288. https://doi.org/10.1371/journal.pcbi.1011288

Chierichetti, F., Kumar, R., Lattanzi, S. ve Vassilvitskii, S. (2017). Fair clustering through fairlets. *Advances in Neural Information Processing Systems, 30*, 5029–5037.

Dasgupta, S., Frost, N., Moshkovitz, M. ve Rashtchian, C. (2020). Explainable k-means and k-medians clustering. *Proceedings of the 37th International Conference on Machine Learning, PMLR 119*, 7055–7065.

Davies, D. L. ve Bouldin, D. W. (1979). A cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-1*(2), 224–227.

Ding, C. ve He, X. (2004). K-means clustering via principal component analysis. *Proceedings of the 21st International Conference on Machine Learning* (s. 225–232). New York: ACM. https://doi.org/10.1145/1015330.1015408

Elkan, C. (2003). Using the triangle inequality to accelerate k-means. *Proceedings of the 20th International Conference on Machine Learning* (s. 147–153). Menlo Park: AAAI Press.

Ezugwu, A. E., Ikotun, A. M., Oyelade, O. N., Abualigah, L., Agushaka, J. O., Eke, C. I. ve Akinyelu, A. A. (2022). A comprehensive survey of clustering algorithms: State-of-the-art machine learning applications, taxonomy, challenges, and future research prospects. *Engineering Applications of Artificial Intelligence, 110*, 104743. https://doi.org/10.1016/j.engappai.2022.104743

Fard, M. M., Thonet, T. ve Gaussier, E. (2020). Deep k-means: Jointly clustering with k-means and learning

representations. *Pattern Recognition Letters, 138*, 185–192. https://doi.org/10.1016/j.patrec.2020.07.028

Fleissner, M., Vankadara, L. C. ve Ghoshdastidar, D. (2024). Explaining kernel clustering via decision trees. *Proceedings of the Twelfth International Conference on Learning Representations (ICLR 2024).* https://arxiv.org/abs/2402.09881

Foss, A., Markatou, M., Ray, B. ve Heching, A. (2016). A semiparametric method for clustering mixed data. *Machine Learning, 105*(3), 419–458.

Fränti, P. ve Sieranoja, S. (2018). K-means properties on six clustering benchmark datasets. *Applied Intelligence, 48*(12), 4743–4759. https://doi.org/10.1007/s10489-018-1238-7

Fränti, P. ve Sieranoja, S. (2019). How much can k-means be improved by using better initialization and repeats? *Pattern Recognition, 93*, 95–112. https://doi.org/10.1016/j.patcog.2019.04.014

Gagolewski, M. (2022). A framework for benchmarking clustering algorithms. *SoftwareX, 20*, 101270. https://doi.org/10.1016/j.softx.2022.101270

Gagolewski, M., Bartoszuk, M. ve Cena, A. (2021). Are cluster validity measures (in)valid? *Information Sciences, 581*, 620–636. https://doi.org/10.1016/j.ins.2021.10.004

García-Escudero, L. A., Gordaliza, A., Matrán, C. ve Mayo-Iscar, A. (2010). A review of robust clustering methods. *Advances in Data Analysis and Classification, 4*(2–3), 89–109. https://doi.org/10.1007/s11634-010-0064-5

García-Escudero, L. A. ve Mayo-Iscar, A. (2024). Robust clustering based on trimming. *WIREs Computational*

Statistics, 16(4), e1658. https://doi.org/10.1002/wics.1658

Ghadiri, M., Samadi, S. ve Vempala, S. (2021). Socially fair k-means clustering. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21)* (s. 438–448). New York: ACM. https://doi.org/10.1145/3442188.3445906

Gul, M., & Rehman, M. A. (2023). Big data: An optimized approach for cluster initialization. Journal of Big Data, 10, Article 120. https://doi.org/10.1186/s40537-023-00798-1

Hussayn, F., & Shah, S. M. (2023). Parametric entropy based cluster centroid initialization for k-means clustering of various image datasets. In Proceedings of the IEEE International Conference on Signal Processing and Computer Vision. https://doi.org/10.2991/978-94-6463-529-4_5

Heidari, J., Daneshpour, N. ve Zangeneh, A. (2024). A novel K-means and K-medoids algorithms for clustering non-spherical-shape clusters non-sensitive to outliers. *Pattern Recognition, 155*, 110639. https://doi.org/10.1016/j.patcog.2024.110639

Hennig, C. (2015). What are the true clusters? *Pattern Recognition Letters, 64*, 53–62.

Hennig, C. (2022). An empirical comparison and characterisation of nine popular clustering methods. *Advances in Data Analysis and Classification, 16*(1), 201–229. https://doi.org/10.1007/s11634-021-00478-z

Huang, Z. (1998). Extensions to the k-means algorithm for clustering large data sets with categorical values. *Data Mining and Knowledge Discovery, 2*(3), 283–304.

Ikotun, A. M., Ezugwu, A. E., Abualigah, L., Abuhaija, B. ve Heming, J. (2023). K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. *Information Sciences, 622*, 178–210. https://doi.org/10.1016/j.ins.2022.11.139

Ikotun, A. M., Habyarimana, F. ve Ezugwu, A. E. (2025). Benchmarking validity indices for evolutionary K-means clustering performance. *Scientific Reports, 15*, 21842. https://doi.org/10.1038/s41598-025-08473-6

Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters, 31*(8), 651–666. https://doi.org/10.1016/j.patrec.2009.09.011

Junjie, J., Wenhao, S., & Yuan, W. (2024). A risk assessment approach for road collapse along tunnels based on an improved entropy weight method and K-means cluster algorithm. *Ain Shams Engineering Journal, 15*(7), 102805. https://doi.org/10.1016/j.asej.2024.102805

Khan, I. K., Daud, H. B., Zainuddin, N. B., Sokkalingam, R., Abdussamad, Museeb, A. ve Inayat, A. (2024). Addressing limitations of the K-means clustering algorithm: Outliers, non-spherical data, and optimal cluster selection. *AIMS Mathematics, 9*(9), 25070–25097. https://doi.org/10.3934/math.20241222

Kleinberg, J. (2002). An impossibility theorem for clustering. *Advances in Neural Information Processing Systems, 15*, 463–470.

Liao, J., Wu, X., Wu, Y., & Shu, J. (2024). K-NNDP: K-means algorithm based on nearest neighbor density peak optimization and outlier removal. *Knowledge-based systems, 294*, 111742. https://doi.org/10.1016/j.knosys.2024.111742

Likas, A., Vlassis, N. ve Verbeek, J. J. (2003). The global k-means clustering algorithm. *Pattern Recognition, 36*(2), 451–461. https://doi.org/10.1016/S0031-3203(02)00060-2

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. L. M. Le Cam ve J. Neyman (Ed.), *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability* içinde (Cilt 1, s. 281–297). Berkeley: University of California Press.

Mahon, L. ve Lapata, M. (2025). K*-Means: A parameter-free clustering algorithm. https://arxiv.org/abs/2505.11904

Malik, V., Kargin, T. ve Hassibi, B. (2026). Distributionally robust K-means clustering. https://arxiv.org/abs/2604.11118

Manthey, B., & Röglin, H. (2009, December). Worst-case and smoothed analysis of k-means clustering with Bregman divergences. In *International Symposium on Algorithms and Computation* (ss. 1024-1033). Berlin, Heidelberg: Springer Berlin Heidelberg.

Mussabayev, R. ve Mussabayev, R. (2024). Comparative analysis of optimization strategies for K-means clustering in big data contexts: A review. *arXiv:2310.09819v3* [ön baskı]. https://arxiv.org/abs/2310.09819

Nguyen Da, T., Cho, M. Y., & Nguyen Thanh, P. (2025). Optimizing K-means clustering center selection with density-based spatial cluster in radial basis function neural network for load forecasting of smart solar microgrid. *Electrical Engineering*, *107*(7), 8963-8978. https://doi.org/10.1007/s00202-024-02599-y

Özarı, Ç., Eren, Ö., & Alıcı, A. (2019). K-ortalamalar yönteminin başlangıç merkez seçim sorunsalı üzerine bir çalışma. *Business & Management Studies: An International Journal, 7*(2), 1117–1135.

Pelleg, D. ve Moore, A. W. (2000). X-means: Extending K-means with efficient estimation of the number of clusters. *Proceedings of the 17th International Conference on Machine Learning* (s. 727–734). San Francisco: Morgan Kaufmann.

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics, 20*, 53–65.

Sculley, D. (2010). Web-scale k-means clustering. *Proceedings of the 19th International Conference on World Wide Web* (s. 1177–1178). New York: ACM.

Sinaga, K. P. ve Yang, M.-S. (2020). Unsupervised K-means clustering algorithm. *IEEE Access, 8*, 80716–80727. https://doi.org/10.1109/ACCESS.2020.2988796

Stemmer, U. ve Kaplan, H. (2018). Differentially private k-means with constant multiplicative error. *Advances in Neural Information Processing Systems, 31*, 5431–5441.

Tibshirani, R., Walther, G. ve Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B, 63*(2), 411–423.

Ullmann, T., Beer, A., Hünemörder, M., Seidl, T. ve Boulesteix, A.-L. (2023). Over-optimistic evaluation and reporting of novel cluster algorithms: An empirical study. *Machine Learning, 112*(5), 1791–1832. https://doi.org/10.1007/s11634-022-00496-5

Ünlü, R. ve Xanthopoulos, P. (2019). Estimating the number of clusters in a dataset via consensus clustering. *Expert Systems with Applications, 125*, 33–39. https://doi.org/10.1016/j.eswa.2019.01.074

Von Luxburg, U. (2010). Clustering stability: An overview. *Foundations and Trends in Machine Learning, 2*(3), 235–274.

Witten, D. M. ve Tibshirani, R. (2010). A framework for feature selection in clustering. *Journal of the American Statistical Association, 105*(490), 713–726. https://doi.org/10.1198/jasa.2010.tm09415

Xiong, H., Wu, J. ve Chen, J. (2009). K-means clustering versus validation measures: A data-distribution perspective. *IEEE Transactions on Systems, Man, and Cybernetics, Part B, 39*(2), 318–331.

SÜRDÜRÜLEBİLİR KALKINMA PERFORMANSININ ZAMANSAL VE MEKÂNSAL ANALİZİ: DİNAMİK ZAMAN BÜKME TABANLI KÜMELEME VE COĞRAFİ HARİTALAMA

Merve DOĞRUEL[1]

1. GİRİŞ

Sürdürülebilir kalkınma, ekonomik ve sosyal gelişmenin çevresel sınırlar gözetilerek sürdürülmesini ve kalkınmanın farklı boyutlarının bütüncül bir çerçevede değerlendirilmesini gerektirmektedir. Bu doğrultuda Sürdürülebilir Kalkınma Amaçları (SKA), ülkelerin ekonomik, sosyal ve çevresel alanlardaki ilerlemelerinin izlenmesi ve karşılaştırılması için ortak bir değerlendirme çerçevesi sunmaktadır. Bununla birlikte ülkelerin sürdürülebilir kalkınma performansları gerek ulaştıkları düzeyler gerekse zaman içerisindeki gelişim örüntüleri bakımından önemli ölçüde farklılaşmaktadır. Güncel çalışmalar, ülkeler arasındaki bu heterojenliğin performans düzeylerinin yanı sıra gelişim yörüngeleri, ilerleme hızları ve yakınsama süreçleri açısından da gözlendiğini ortaya koymaktadır (Ma et al., 2025; Xu et al., 2025; Bigerna & Micheli, 2025).

Sürdürülebilir kalkınma performansının zaman boyutuyla incelenmesi, belirli bir yıldaki durumun karşılaştırılmasının ötesinde uzun dönemli gelişim örüntülerinin ortaya çıkarılmasına olanak sağlamaktadır. Ma et al. (2025), ülkelerin sürdürülebilir

[1] Dr. Öğretim Üyesi, İstanbul Esenyurt Üniversitesi, İşletme ve Yönetim Bilimleri Fakültesi, Yönetim Bilişim Sistemleri Bölümü, ORCID: 0000-0003-2299-7182.

kalkınma yörüngelerinin homojen olmadığını; Xu et al. (2025) ise performans ve ilerleme hızlarının ülkelerin gelir düzeylerine göre farklılaştığını göstermiştir. Bigerna ve Micheli (2025) tarafından ülkeler arası yakınsama süreçlerine ilişkin elde edilen bulgular da sürdürülebilir kalkınma performansının statik bir ülke karşılaştırmasından ziyade zamansal bir süreç olarak ele alınmasının önemini desteklemektedir. Ayrıca sürdürülebilir kalkınmadaki ilerleme doğrusal ve kesintisiz değildir. COVID-19 pandemisinin SKA ilerlemesi üzerindeki farklılaşan etkileri (Yuan et al., 2023) ve ülkeler arasında belirlenen zamansal ve mekânsal heterojenlik (Fang et al., 2023), uzun dönemli performansın dönem içindeki değişimler ve gerilemelerle birlikte değerlendirilmesi gerektiğini göstermektedir.

Ülke düzeyindeki sürdürülebilir kalkınma performansının küresel endeksler ve nicel yöntemler aracılığıyla değerlendirilmesi önceki çalışmalarda da ele alınmıştır. Doğruel Anuşlu ve Fırat (2020), 116 ülkenin verilerini kullanarak inovasyon göstergeleri ile sürdürülebilir kalkınma ve çevresel performans arasındaki ilişkileri incelemiş ve ülke düzeyindeki sürdürülebilirlik performansının çok boyutlu göstergeler aracılığıyla nicel olarak değerlendirilebileceğini göstermiştir. Kümeleme temelli güncel çalışmalar da ülkelerin sürdürülebilir kalkınma profilleri bakımından homojen olmadığını ve farklı ülke gruplarının veri temelli yöntemlerle belirlenebildiğini ortaya koymaktadır (Marti & Cervelló-Royo, 2023; Özekin & Çelik, 2024; Jena & Basel, 2025).

Bununla birlikte ülkelerin uzun dönemli sürdürülebilir kalkınma performanslarının karşılaştırılmasında yalnızca belirli yıllardaki düzeylere veya geleneksel uzaklık ölçülerine dayalı sınıflandırmalar, zaman serilerinin gelişim biçimindeki benzerlikleri yeterince yansıtmayabilir. Dinamik Zaman Bükme (Dynamic Time Warping-DTW), zaman serileri arasındaki benzerliği zaman eksenindeki olası kaymaları dikkate alarak

değerlendiren elastik bir uzaklık yaklaşımıdır. Maharaj et al. (2024), ülke düzeyindeki yenilenebilir enerji zaman serilerini DTW tabanlı yaklaşımla kümelemiş; Kartal (2025) ise ülkelerin Küresel Cinsiyet Uçurumu Endeksi serilerinin sınıflandırılmasında zaman serisi kümelemesinden yararlanmıştır. DTW'nin zaman serisi kümelemesindeki kullanımı ve diğer elastik uzaklık ölçüleri arasındaki konumu güncel metodolojik çalışmalarda da kapsamlı biçimde ele alınmaktadır (Holder et al., 2024; Vera & Angulo, 2023).

Bu çerçevede çalışmanın temel amacı, ülkelerin 2000–2025 dönemindeki Sürdürülebilir Kalkınma Amaçları Endeksi (SKAE) değerlerini zamansal bir bütünlük içerisinde değerlendirerek benzer gelişim örüntülerine sahip ülke gruplarını belirlemek ve elde edilen kümelerin zamansal, mekânsal ve amaç bazlı özelliklerini karşılaştırmaktır. Çalışmada 169 ülkeye ait 26 yıllık SKAE serileri DTW uzaklığı kullanılarak karşılaştırılmakta ve hiyerarşik kümeleme yöntemiyle ülke grupları oluşturulmaktadır. Ardından kümelerin uzun dönemli değişim ve istikrar özellikleri değerlendirilmekte, mekânsal dağılımları coğrafi olarak incelenmekte ve 2025 yılına ait 17 SKA göstergesi bakımından kümeler arasındaki farklılıklar istatistiksel olarak test edilmektedir.

Çalışmanın yöntemsel katkısı, sürdürülebilir kalkınma performansını yalnızca ülkelerin ulaştıkları düzeylere göre değil, 26 yıllık gelişim örüntülerinin benzerliğini esas alarak sınıflandırmasıdır. DTW tabanlı kümeleme sonuçlarının zamansal değişim ölçütleri, coğrafi haritalama ve amaç bazlı istatistiksel karşılaştırmalarla birlikte değerlendirilmesiyle ülkelerin sürdürülebilir kalkınma performansındaki heterojenlik düzey, zaman, mekân ve amaç boyutlarını içeren bütünleşik bir analitik çerçevede ele alınmaktadır.

2. YÖNTEM

Bu bölümde araştırmanın veri yapısı ve analiz süreci açıklanmakta; DTW tabanlı kümeleme yaklaşımı, zamansal ve mekânsal profil analizleri ile kümelerin karşılaştırılmasında kullanılan istatistiksel yöntemler sırasıyla ele alınmaktadır.

2.1. Araştırma Tasarımı ve Veri Seti

Bu çalışma, ülkelerin sürdürülebilir kalkınma performanslarının uzun dönemli değişimini zamansal ve mekânsal boyutlarıyla incelemek amacıyla nicel ve boylamsal bir araştırma tasarımına dayanmaktadır. Araştırmanın temel analiz birimini ülkeler oluşturmakta; ülkelerin sürdürülebilir kalkınma performansları 2000–2025 dönemini kapsayan yıllık gözlemler üzerinden değerlendirilmektedir. Bu yaklaşım, ülkelerin yalnızca belirli bir yıldaki performans düzeylerinin veya başlangıç ve bitiş değerleri arasındaki değişimin değil, dönem boyunca izledikleri gelişim örüntülerinin de dikkate alınmasına olanak sağlamaktadır.

Araştırmada Sürdürülebilir Kalkınma Raporu 2026 kapsamında yayımlanan veri tabanı kullanılmıştır (Sachs et al., 2026). Sürdürülebilir kalkınma performansının temel göstergesi olarak Sürdürülebilir Kalkınma Amaçları Endeksi (SKAE; Sustainable Development Goals Index [SDGI]) esas alınmıştır. SKAE, ülkelerin 17 Sürdürülebilir Kalkınma Amacına ilişkin genel performansını 0 ile 100 arasında değişen ortak bir ölçekte ifade etmekte ve daha yüksek değerler amaçlara ulaşma yönünde daha yüksek performansı göstermektedir. Çalışmada genel SKAE değerlerinin yanı sıra 17 Sürdürülebilir Kalkınma Amacına ilişkin amaç bazlı performans skorlarından da yararlanılmıştır. Genel SKAE serileri ülkelerin uzun dönemli gelişim örüntülerinin belirlenmesinde, amaç bazlı skorlar ise elde edilen ülke kümelerinin sürdürülebilir kalkınmanın farklı boyutları açısından karşılaştırılmasında kullanılmıştır.

Ham veri seti 2000–2025 dönemine ilişkin 26 yıllık gözlemleri içermektedir. Analiz öncesinde ülke düzeyindeki gözlemler ile bölgesel, ekonomik ve küresel toplulaştırılmış göstergeler ayrıştırılmış; ülke niteliği taşımayan 14 toplulaştırılmış kayıt analiz kapsamından çıkarılmıştır. Veri temizleme işlemleri sonucunda nihai analiz seti 169 ülke ve 26 yıllık dönemden oluşmuş, böylece toplam 4.394 ülke-yıl gözlemi elde edilmiştir.

Ülkelerin zamansal benzerliklerinin analiz edilebilmesi amacıyla veri seti, satırlarda ülkelerin ve sütunlarda yılların yer aldığı 169×26 boyutlu bir zaman serisi matrisine dönüştürülmüştür. Buna göre her ülke, 2000–2025 dönemindeki yıllık SKAE değerlerinden oluşan bir zaman serisi ile temsil edilmiştir. Genel SKAE değişkeninde incelenen dönem boyunca eksik gözlem bulunmadığından zaman serisi benzerliklerinin hesaplanmasında herhangi bir veri tamamlama yöntemine başvurulmamıştır.

Analizin ilk aşamasında ülkelerin SKAE zaman serileri arasındaki benzerlikler Dinamik Zaman Bükme yöntemiyle hesaplanmış ve elde edilen uzaklık matrisi hiyerarşik kümeleme analizinde kullanılmıştır. Daha sonra ülke kümelerinin zamansal değişim ve istikrar özellikleri incelenmiş, mekânsal dağılımları ise coğrafi görselleştirme teknikleriyle değerlendirilmiştir. Son aşamada ise kümeler, 17 Sürdürülebilir Kalkınma Amacına ilişkin performans skorları açısından istatistiksel olarak karşılaştırılmıştır.

2.2. Dinamik Zaman Bükme ve Hiyerarşik Kümeleme

Ülkelerin 2000–2025 dönemindeki sürdürülebilir kalkınma performanslarını yalnızca başlangıç ve bitiş düzeylerine göre değil, zaman içindeki gelişim örüntülerine göre karşılaştırmak amacıyla Dinamik Zaman Bükme (Dynamic Time Warping–DTW) yöntemi kullanılmıştır. DTW, iki zaman serisini

zaman eksenindeki yerel kaymalara izin vererek hizalayan esnek bir uzaklık ölçüsüdür. Bu nedenle benzer değişimlerin farklı zamanlarda veya farklı hızlarda gerçekleştiği serilerin karşılaştırılmasına olanak sağlamaktadır. Güncel zaman serisi kümeleme çalışmalarında DTW ve diğer esnek uzaklık ölçülerinin özellikle örüntü benzerliğini yakalamadaki rolü vurgulanmaktadır (Vera & Angulo, 2023; Holder et al., 2024). Kartal (2025) da ülke düzeyindeki uzun dönemli endeks serilerinin sınıflandırılmasında DTW tabanlı zaman serisi kümelemesini kullanmıştır.

Her bir ülke, 2000–2025 dönemine ait yıllık SKAE değerlerinden oluşan tek değişkenli bir zaman serisi ile temsil edilmiştir. DTW uzaklığının hesaplanmasını göstermek üzere, karşılaştırılan herhangi iki ülkenin SKAE serileri $X = (x_1, x_2, \dots, x_n)$ve $Y = (y_1, y_2, \dots, y_m)$olarak tanımlanmıştır. Bu iki ülkenin $x_i$ve $y_j$gözlemleri arasındaki yerel maliyet aşağıdaki biçimde hesaplanmıştır:

$$c(i,j) = \left(x_i - y_j\right)^2 \quad (1)$$

Yerel maliyetlerden oluşturulan matris üzerinde iki seriyi başlangıçtan sona bağlayan ve toplam maliyeti en aza indiren bükme yolu belirlenir. Her hücreye ulaşan en düşük kümülatif maliyet, mevcut hücrenin maliyetine önceki üç olası geçişten en küçüğünün eklenmesiyle hesaplanmaktadır:

$$D(i,j) = c(i,j) + \min\{D(i-1,j), D(i,j-1), D(i-1,j-1)\} \quad (2)$$

Buna göre DTW uzaklığı, izin verilen bükme yolları arasındaki minimum toplam eşleştirme maliyetine dayanmaktadır. Bu çalışmada kullanılan hesaplama yaklaşımı karesel yerel maliyetlerin optimal yol boyunca toplanıp karekökünün alınması biçiminde ifade edilebilir:

$$DTW(X,Y) = \min_P \sqrt{\sum_{(i,j)\in P} \left(x_i - y_j\right)^2} \quad (3)$$

Burada P, iki seriyi eşleştiren geçerli bükme yolunu göstermektedir. DTW'nin bu yapısı, aynı yöndeki değişimlerin tam olarak aynı yılda gerçekleşmesi zorunluluğunu ortadan kaldırmakta ve serilerin bütünsel gelişim örüntüsünü karşılaştırmaya imkân vermektedir (Holder et al., 2024). Kartal'ın (2025) çalışmasında da DTW, ülkelerin uzun dönemli endeks hareketlerini benzerliklerine göre gruplandırmak amacıyla kullanılmıştır.

DTW hesaplamaları ülkelerin 2000–2025 dönemine ait ham SKAE serileri üzerinden gerçekleştirilmiştir. SKAE değerleri bütün ülkeler için aynı 0–100 ölçeğinde tanımlandığından ve araştırmada yalnızca değişimin biçimi değil, sürdürülebilir kalkınma performansının düzeyi ile zamansal gelişiminin birlikte dikkate alınması amaçlandığından ülke serilerine ayrıca standartlaştırma uygulanmamıştır. Her ülkenin 26 yıllık serisi diğer ülkelerin serileriyle karşılaştırılarak 169 ülke için (169) boyutunda simetrik bir DTW uzaklık matrisi elde edilmiştir.

Elde edilen uzaklık matrisi, birleştirici (agglomerative) hiyerarşik kümeleme analizinin girdisi olarak kullanılmıştır. Bu yaklaşımda başlangıçta her ülke ayrı bir küme olarak kabul edilmekte, daha sonra birbirine en yakın kümeler aşamalı biçimde birleştirilmektedir. Kümeler arasındaki uzaklığın hesaplanmasında ortalama bağlantı (average linkage) yöntemi tercih edilmiştir. İki küme arasındaki ortalama bağlantı uzaklığı:

$$\mathrm{d_{avg}(A,B)} = \frac{1}{|\mathrm{A}||\mathrm{B}|}\sum_{\mathrm{i}\in\mathrm{A}}\sum_{\mathrm{j}\in\mathrm{B}}\mathrm{d_{DTW}(i,j)} \qquad (4)$$

şeklinde tanımlanmaktadır. Burada A ve B karşılaştırılan kümeleri, |A| ve |B| küme büyüklüklerini,$d_{DTW}(i,j)$ ise iki ülke arasındaki DTW uzaklığını göstermektedir. Ortalama bağlantı, kümelerdeki tüm gözlem çiftlerinin uzaklıklarını dikkate aldığı ve önceden hesaplanmış DTW uzaklık matrisiyle doğrudan kullanılabildiği için araştırmanın analiz yapısıyla uyumludur.

Güncel uygulamalarda DTW'nin farklı kümeleme algoritmaları ve bağlantı ölçütleriyle birlikte kullanıldığı görülmektedir (Kartal, 2025).

Nihai küme sayısının belirlenmesinde ortalama silhouette katsayısı esas alınmıştır. Silhouette katsayısı, bir gözlemin kendi kümesi içindeki uyumu ile en yakın alternatif kümeye olan ayrışmasını birlikte değerlendirmektedir. Bir i gözlemi için katsayı:

$$s(i) = \frac{b(i)-a(i)}{\max\{a(i),b(i)\}}, \qquad -1 \leq s(i) \leq 1 \qquad (5)$$

biçimindedir. Burada a(i), gözlemin kendi kümesindeki diğer gözlemlere ortalama uzaklığını; b(i) ise diğer kümelere olan ortalama uzaklıkların en küçüğünü göstermektedir. Değerin 1'e yaklaşması daha güçlü küme uyumuna ve ayrışmaya, 0'a yaklaşması sınırda bir konuma, negatif değerler ise olası uyumsuz kümelenmeye işaret etmektedir.

Bu doğrultuda alternatif küme sayıları için hiyerarşik kümeleme çözümleri oluşturulmuş ve her çözümün ortalama silhouette katsayısı hesaplanmıştır. Nihai yapı belirlenirken silhouette değerleri, küme büyüklükleri ve elde edilen grupların analitik yorumlanabilirliği birlikte değerlendirilmiştir. Böylece ülkeler, tek bir yılın performansına veya yalnızca dönem başı–dönem sonu değişimine göre değil, 2000–2025 dönemindeki performans düzeyi ve zamansal gelişim örüntülerinin bütünü dikkate alınarak sınıflandırılmıştır.

2.3. Zamansal Profil Analizi ve Mekânsal Haritalama

DTW tabanlı kümeleme sonucunda elde edilen ülke gruplarının dönem içerisindeki değişim ve istikrar özelliklerini incelemek amacıyla tamamlayıcı bir zamansal profil analizi gerçekleştirilmiştir. Bu kapsamda her ülke için yıllık SKAE değişimi, yıllık değişimlerin ortalaması ve dalgalanma düzeyi, gerileme yaşanan yıl sayısı ve dönem içerisindeki en büyük tek

yıllık düşüş hesaplanmıştır. Böylece kümelerin yalnızca sürdürülebilir kalkınma performans düzeyleri açısından değil, gelişim hızı ve kısa dönemli istikrar özellikleri bakımından da karşılaştırılması amaçlanmıştır.

Ülke i'nin t yılındaki yıllık SKAE değişimi,

$$\Delta_{i,t} = SDGI_{i,t} - SDGI_{i,t-1} \quad (6)$$

olarak hesaplanmıştır. Pozitif değerler bir önceki yıla göre performans artışını, negatif değerler ise gerilemeyi göstermektedir. Ülkelerin uzun dönemli gelişim hızının belirlenmesinde yıllık değişimlerin ortalamasından, yıllar arasındaki değişkenliğin değerlendirilmesinde ise yıllık değişimlerin standart sapmasından yararlanılmıştır. Dalgalanma ölçütü,

$$\sigma_{\Delta,i} = \sqrt{\frac{1}{T-2}\sum_{t=2}^{T}\left(\Delta_{i,t} - \bar{\Delta}_i\right)^2} \quad (7)$$

şeklinde hesaplanmıştır. Burada T toplam dönem sayısını, $\bar{\Delta}_i$ ise ülkenin ortalama yıllık SKAE değişimini göstermektedir. Daha yüksek standart sapma değerleri yıllık değişimlerin daha dalgalı olduğunu ifade etmektedir. Ayrıca her ülke için yıllık değişimin negatif olduğu yılların sayısı ile dönem içerisindeki en düşük yıllık değişim belirlenmiş; bu göstergeler kümelerin gelişim ve istikrar profillerinin karşılaştırılmasında kullanılmıştır.

Ülke kümelerinin mekânsal dağılımını incelemek amacıyla küme üyelikleri coğrafi sınır verileriyle eşleştirilerek dünya haritası üzerinde görselleştirilmiştir.

2.4. İstatistiksel Analiz

Sürdürülebilir Kalkınma Amaçları bakımından DTW tabanlı kümelerin dönem sonundaki yapısal farklılıklarını değerlendirmek amacıyla karşılaştırmalı istatistiksel analizler gerçekleştirilmiştir. Ülke kümeleri 2000–2025 dönemine ait

SKAE zaman serilerinin tamamı kullanılarak oluşturulurken, kümelerin dönem sonundaki sürdürülebilir kalkınma profillerinin karşılaştırılmasında veri setindeki son gözlem yılı olan 2025 yılına ait SKA 1–SKA 17 göstergeleri kullanılmıştır. Böylece zamansal gelişim örüntülerine göre belirlenen ülke gruplarının inceleme dönemi sonunda hangi sürdürülebilir kalkınma boyutlarında farklılaştığının ortaya konulması amaçlanmıştır.

Küme üyelikleri bağımsız grupları temsil ettiğinden, karşılaştırmalı analizler öncesinde dağılım ve varyans varsayımları değerlendirilmiştir. Bu kapsamda grup bazında Shapiro–Wilk testi ve dağılım özellikleri incelenmiş; varyansların homojenliği Levene ve Brown–Forsythe testleriyle kontrol edilmiştir. Özellikle küçük örneklemli kümelerde normallik testlerinin sınırlılıkları dikkate alınarak istatistiksel test seçimi yalnızca tek bir varsayım testinin sonucuna dayandırılmamıştır.

Küme büyüklüklerinin eşit olmaması ve bazı SKA göstergelerinde varyans homojenliğinin sağlanmaması nedeniyle kümeler arası ortalama farklılıklarının değerlendirilmesinde klasik tek yönlü varyans analizi yerine Welch ANOVA kullanılmıştır. On yedi SKA göstergesinin ayrı ayrı test edilmesinden kaynaklanabilecek çoklu test sorununu kontrol etmek amacıyla omnibus testlerden elde edilen p değerlerine Benjamini–Hochberg yanlış keşif oranı (False Discovery Rate [FDR]) düzeltmesi uygulanmıştır. FDR yaklaşımı, çoklu hipotez testlerinde yanlış keşiflerin kontrol edilmesine yönelik temel yaklaşımlardan biridir (Dai & Zheng, 2023).

FDR düzeltmesi sonrasında anlamlı bulunan Welch ANOVA sonuçlarında, farklılığın hangi kümeler arasında ortaya çıktığını belirlemek amacıyla Games–Howell çoklu karşılaştırma testi uygulanmıştır. Bu yöntemin tercih edilmesinde eşit varyans ve eşit grup büyüklüğü varsayımlarını gerektirmemesi dikkate

alınmıştır. Böylece omnibus testlerle belirlenen genel küme farklılıklarının hangi ikili küme karşılaştırmalarından kaynaklandığı değerlendirilmiştir.

Parametrik bulguların dağılım varsayımlarına duyarlılığını değerlendirmek amacıyla analizler ayrıca Kruskal–Wallis testi ile tekrarlanmıştır. Kruskal–Wallis testlerinden elde edilen p değerlerine de Benjamini–Hochberg FDR düzeltmesi uygulanmış ve sonuçlar Welch ANOVA bulgularıyla karşılaştırılarak sağlamlık kontrolü gerçekleştirilmiştir. Kruskal–Wallis analizlerinde farklılıkların büyüklüğünü değerlendirmek amacıyla epsilon-kare (ε^2) etki büyüklüğü raporlanmıştır. İstatistiksel anlamlılığın etki büyüklüğü ölçütleriyle birlikte değerlendirilmesi, gözlenen farklılıkların büyüklüğünün yorumlanmasına katkı sağlamaktadır (Iacobucci et al., 2023). Tüm istatistiksel analizlerde anlamlılık düzeyi .05 olarak kabul edilmiştir.

3. BULGULAR

Bu bölümde analizlerden elde edilen bulgular; sürdürülebilir kalkınma performansının zamansal değişimi, DTW tabanlı ülke kümeleri, kümelerin zamansal ve mekânsal profilleri ve SKA boyutlarındaki farklılıklar çerçevesinde sunulmaktadır.

3.1. Tanımlayıcı Bulgular ve Sürdürülebilir Kalkınma Performansının Zamansal Değişimi

Araştırmada 169 ülkenin 2000 ile 2025 yılları arasındaki sürdürülebilir kalkınma performansı incelenmiştir. Ülkelerin dönem başındaki Sürdürülebilir Kalkınma Amaçları Endeksi (SKAE) düzeyleri ile dönem boyunca gerçekleşen toplam değişim arasındaki ilişki değerlendirildiğinde, başlangıç düzeyi ile SKAE artışı arasında orta düzeyde ve negatif yönlü bir ilişki belirlenmiştir. Pearson korelasyon katsayısı $r = -,463$ ($p < ,001$),

Spearman sıra korelasyon katsayısı ise ρ = −,484 (p < ,001) olarak hesaplanmıştır. $r = -0.463,\ p < 0.001 \rho = -0.484,\ p < 0.001$Her iki katsayının benzer büyüklük ve yönde olması, başlangıç SKAE düzeyi daha düşük olan ülkelerin 2000–2025 döneminde genel olarak daha yüksek performans artışı gösterme eğiliminde olduklarını ortaya koymaktadır. Bununla birlikte bu bulgu, ülkeler arasında doğrudan bir yakınsama kanıtı olarak değil, başlangıç performansı ile uzun dönemli değişim arasındaki negatif yönlü ilişki olarak değerlendirilmiştir.

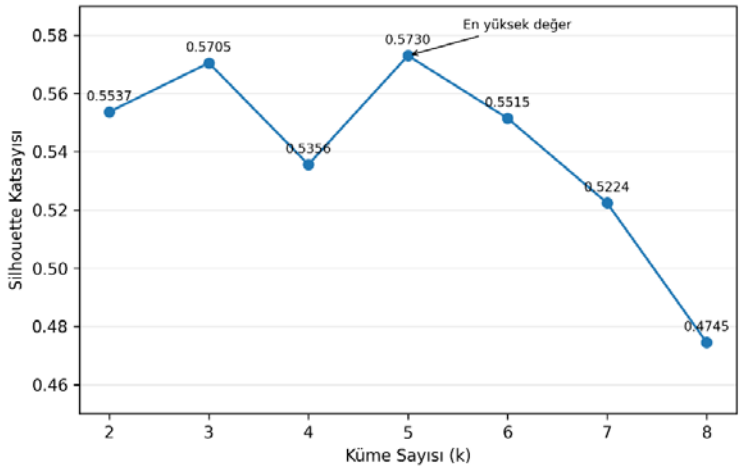
Ülkelerin uzun dönemli performans değişimini değerlendirmek amacıyla 2000 ve 2025 yılı SKAE değerleri arasındaki fark ülke bazında hesaplanmıştır. Bu karşılaştırmada 169 ülkenin tamamında net değişimin pozitif olduğu, diğer bir ifadeyle tüm ülkelerin 2025 yılı SKAE değerlerinin 2000 yılı değerlerinin üzerinde bulunduğu belirlenmiştir. Ancak başlangıç ve bitiş değerleri arasındaki net değişim, ülkelerin 25 yıllık süreç içerisinde izledikleri gelişim örüntüsünü tek başına yansıtmamaktadır.

Bu nedenle dönem içerisindeki değişimin sürekliliğini değerlendirmek üzere her ülke için ardışık yıllara ait SKAE değerleri arasındaki yıllık değişimler hesaplanmıştır. Böylece her ülke için 2001 yılı değerinden 2000 yılı değerinin çıkarılmasıyla başlayan ve 2025 yılı değerinden 2024 yılı değerinin çıkarılmasıyla sona eren 25 yıllık değişim değeri elde edilmiş ve negatif değişimlerin varlığı incelenmiştir. Ülke bazındaki bu hesaplamalar sonucunda, 169 ülkenin 167'sinde (%98,8) 2000 ile 2025 yılları arasında en az bir kez negatif yıllık SKAE değişimi gözlenirken, yalnızca iki ülkede dönem boyunca yıllık gerileme belirlenmemiştir. Dolayısıyla ülkelerin tamamı uzun dönemde başlangıç düzeylerinin üzerinde bir performansa ulaşmış olmakla birlikte, bu gelişimin büyük ölçüde kesintisiz ve doğrusal bir seyir izlemediği görülmüştür.

Başlangıç düzeyi ile toplam değişim arasındaki negatif ilişki ve ülkelerin büyük çoğunluğunda gözlenen dönem içi gerilemeler birlikte değerlendirildiğinde, sürdürülebilir kalkınma performansının yalnızca başlangıç ve bitiş değerleri üzerinden incelenmesinin ülkelerin zamansal gelişim dinamiklerini tam olarak yansıtmadığı anlaşılmaktadır. Bu nedenle izleyen analizlerde ülkeler, 2000 ile 2025 yılları arasındaki SKAE zaman serilerinin tamamı dikkate alınarak DTW tabanlı hiyerarşik kümeleme yöntemiyle gruplandırılmış ve benzer zamansal gelişim örüntülerine sahip ülke kümeleri belirlenmiştir.

3.2. DTW Tabanlı Kümeleme Bulguları

Ülkelerin 2000 ile 2025 yılları arasındaki SKAE zaman serileri arasındaki benzerlikler DTW uzaklıkları kullanılarak belirlenmiş ve elde edilen uzaklık matrisi birleştirici hiyerarşik kümeleme analizine girdi olarak kullanılmıştır. Uygun küme sayısının belirlenmesi amacıyla iki ile sekiz küme arasındaki alternatif çözümler Silhouette katsayısı kullanılarak karşılaştırılmıştır. Elde edilen Silhouette katsayıları Şekil 1'de gösterilmektedir.



Şekil 1. Farklı Küme Sayıları İçin Silhouette Katsayıları

Alternatif çözümler içerisinde en yüksek Silhouette katsayısı beş kümeli çözümde (0,5730) elde edilmiştir. Bununla birlikte üç kümeli çözümün Silhouette katsayısının da 0,5705 ile oldukça yakın olduğu görülmüştür. Beş kümeli çözüm, en yüksek Silhouette değerine sahip olmasının yanı sıra üç kümeli yapının daha ayrıntılı bir biçimde ayrışmasını sağlaması nedeniyle nihai küme yapısı olarak değerlendirilmiştir.

Üç ve beş kümeli çözümler karşılaştırıldığında, üç kümeli çözümde yer alan 40 ülkeden oluşan kümenin beş kümeli çözümde aynı biçimde korunduğu belirlenmiştir. Üç kümeli çözümde 39 ülkeden oluşan grup, beş kümeli çözümde 35 ve 4 ülkeden oluşan iki kümeye; 90 ülkeden oluşan grup ise 58 ve 32 ülkeden oluşan iki kümeye ayrılmıştır. Beş kümeli çözümde farklı üçlü kümelerden gelen ülkelerin aynı kümede birleşmemesi, beş kümeli yapının üç kümeli çözümün temel ayrımını değiştirmekten ziyade bu yapıyı daha ayrıntılı biçimde ortaya koyduğunu göstermektedir.

Nihai çözümde kümelerin büyüklükleri sırasıyla 40, 35, 58, 4 ve 32 ülke olarak belirlenmiştir. Kümelerin 2000 ve 2025 yıllarındaki ortalama SKAE düzeyleri, dönem boyunca gerçekleşen toplam değişimleri ve 26 yıllık ortalama SKAE değerleri Tablo 1'de sunulmuştur.

Tablo 1. DTW Tabanlı Kümelerin Temel SKAE Profilleri

Küme	Ülke sayısı	2000 SKAE	2025 SKAE	2000–2025 Değişimi	2000–2025 Ortalaması
Küme 1	40	75,12	80,79	5,66	78,35
Küme 2	35	45,01	54,02	9,00	49,58
Küme 3	58	63,35	71,51	8,15	67,65
Küme 4	4	36,69	43,97	7,28	39,12
Küme 5	32	53,54	64,37	10,83	58,80

Tablo 1 incelendiğinde, kümelerin yalnızca büyüklükleri açısından değil, SKAE düzeyleri ve uzun dönemli değişimleri

bakımından da farklılaştığı görülmektedir. Küme 1, hem 2000 hem de 2025 yılında en yüksek ortalama SKAE düzeyine sahipken, toplam artışın en düşük olduğu kümedir. Küme 4 en düşük ortalama SKAE düzeyine sahip ülke grubunu oluştururken, Küme 5 dönem boyunca en yüksek ortalama artışı göstermiştir. Küme 2 düşük başlangıç düzeyine karşın belirgin bir artış sergilerken, Küme 3 orta-yüksek performans düzeyindeki ülkeleri içermektedir.

3.3. Kümelerin Zamansal Profilleri ve Mekânsal Dağılımı

DTW tabanlı kümeleme sonucunda belirlenen ülke gruplarının zamansal özelliklerini daha ayrıntılı değerlendirmek amacıyla kümeler; ortalama yıllık SKAE değişimi, yıllık değişimlerin standart sapması, gerileme yaşanan yıl sayısı ve en büyük tek yıllık düşüş açısından incelenmiştir. Bu göstergeler, kümelerin yalnızca SKAE düzeyleri ve uzun dönemli değişimleri açısından değil, gelişim süreçlerinin hızı ve istikrarı bakımından da karşılaştırılmasına olanak sağlamaktadır. Küme bazında elde edilen zamansal profil göstergeleri Tablo 2'de sunulmuştur.

Tablo 2. DTW Tabanlı Kümelerin Zamansal Değişim ve İstikrar Profilleri

Küme	n	Ort. yıllık değişim	Yıllık değişim St Sapması	Ort. gerileme yılı	Medyan gerileme yılı	En büyük tek yıllık düşüş	Ort. en büyük düşüş
Küme 1	40	0,227	0,428	6,875	7,0	-2,158	-0,699
Küme 2	35	0,360	0,570	5,200	5,0	-3,378	-0,815
Küme 3	58	0,326	0,522	5,500	5,5	-2,699	-0,779
Küme 4	4	0,291	0,679	8,250	8,0	-1,133	-0,956
Küme 5	32	0,433	0,667	5,125	4,0	-5,522	-0,896

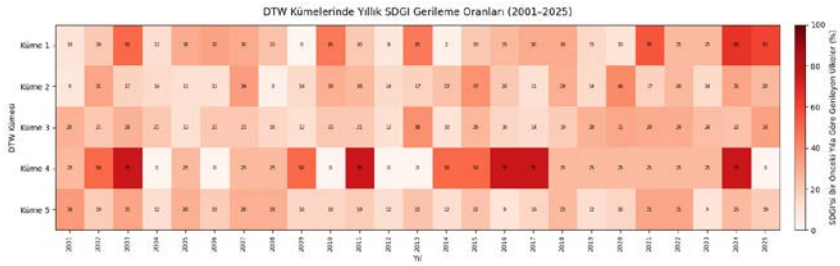
Zamansal profil göstergeleri kümeler arasında belirgin farklılıklar bulunduğunu göstermektedir. ***Küme 1***, kümeler

arasında en yüksek ortalama SKAE düzeyine sahip olmakla birlikte, 0,227 ile en düşük ortalama yıllık artışı ve 0,428 ile en düşük yıllık değişim standart sapmasını göstermektedir. Dolayısıyla bu küme, yüksek sürdürülebilir kalkınma performansının daha yavaş ancak görece istikrarlı bir gelişim örüntüsüyle birlikte görüldüğü ülke grubunu temsil etmektedir. Kümede ülke başına ortalama 6,875 gerileme yılı bulunması, burada kullanılan istikrar kavramının yıllık düşüşlerin hiç yaşanmamasını değil, yıllık değişimlerin görece daha düşük dalgalanma göstermesini ifade ettiğini ortaya koymaktadır. ***Küme 2***, 0,360 ortalama yıllık SKAE artışı ile düşük başlangıç düzeyine karşın görece hızlı bir gelişim göstermektedir. Yıllık SKAE değişimlerinin standart sapmasının 0,570 olması, bu gelişime yıllar arasında belirli ölçüde dalgalanmanın eşlik ettiğini göstermektedir. Kümede ülke başına ortalama 5,2 yılda SKAE gerilemesi yaşanmıştır. ***Küme 3***'te ise ortalama yıllık SKAE artışı 0,326, yıllık değişimlerin standart sapması 0,522 ve ortalama gerileme yılı sayısı 5,5 olarak hesaplanmıştır. Buna göre Küme 3, Küme 2'ye kıyasla biraz daha düşük gelişim hızına, ancak yıllık değişimler açısından daha düşük dalgalanmaya sahip bir yapı sergilemektedir. Her iki kümede de dönem içerisinde geçici gerilemeler görülmekle birlikte, Küme 2 daha hızlı ve nispeten daha dalgalı; Küme 3 ise daha ılımlı ve görece dengeli bir zamansal gelişim örüntüsü göstermektedir. En düşük SKAE düzeyine sahip ***Küme 4***, 0,679 ile yıllık değişimlerin en yüksek standart sapmasına ve ülke başına ortalama 8,25 ile en fazla gerileme yılına sahiptir. Bu bulgular kümenin düşük performans düzeyine daha kesintili ve dalgalı bir gelişim örüntüsünün eşlik ettiğini göstermektedir. Bununla birlikte Küme 4'ün yalnızca dört ülkeden oluşması nedeniyle bu kümeye ilişkin göstergelerin yorumlanmasında örneklem büyüklüğü dikkate alınmalıdır. ***Küme 5*** ise 0,433 ile kümeler arasında en yüksek ortalama yıllık SKAE artışını göstermiştir. Buna karşılık yıllık değişimlerin standart sapmasının 0,667 olması ve gözlenen en büyük tek yıllık

düşüşün -5,522'ye ulaşması, yüksek gelişim hızının aynı zamanda belirgin dönemsel dalgalanmalarla birlikte gerçekleştiğini göstermektedir. Bu nedenle Küme 5, hızlı ancak dalgalı bir gelişim örüntüsü ile diğer kümelerden ayrılmaktadır.

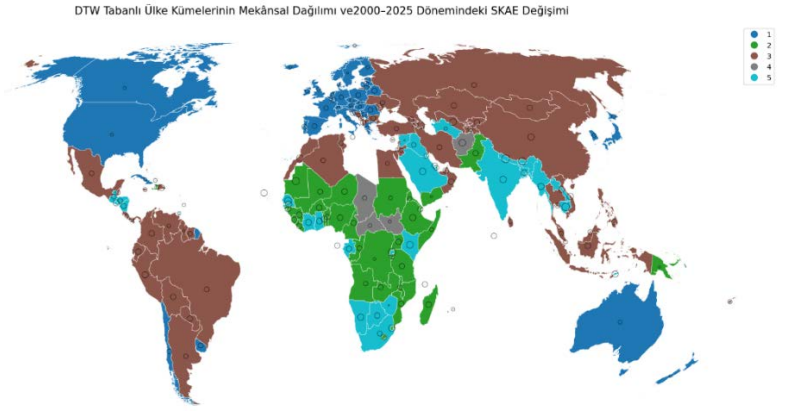
Kümelerin zamansal profilleri birlikte değerlendirildiğinde, benzer uzun dönemli gelişim eğilimlerinin farklı istikrar ve dalgalanma özellikleriyle gerçekleşebildiği görülmektedir. Özellikle düşük SKAE düzeyindeki Küme 2 ile Küme 4'ün; ayrıca üç kümeli çözümde aynı üst yapı içerisinde yer alan Küme 3 ile Küme 5'in zamansal göstergeler bakımından ayrışması, beş kümeli çözümün ülkelerin gelişim örüntülerindeki farklılıkları daha ayrıntılı biçimde ortaya koyduğunu desteklemektedir.

Zamansal değişimin dönem içerisindeki dağılımını daha ayrıntılı incelemek amacıyla, her yıl SKAE değerinde bir önceki yıla göre gerileme yaşayan ülkelerin küme içindeki oranları ayrıca hesaplanmıştır. Şekil 2'de sunulan 2001–2025 dönemine ait ısı haritası, gerileme oranlarının kümelere ve yıllara göre farklılaştığını göstermektedir. Özellikle Küme 4'te bazı yıllarda gerileme oranlarının %75'e ulaşması, bu kümenin Tablo 2'de belirlenen yüksek dalgalanma ve gerileme sıklığıyla uyumludur. Küme 1'de ise yıllık değişimlerin standart sapması görece düşük olmasına rağmen bazı yıllarda gerileme yaşayan ülkelerin oranının yükselmesi, düşük dalgalanmanın yıllık gerilemelerin tamamen ortadan kalktığı anlamına gelmediğini göstermektedir. Küme 2, Küme 3 ve Küme 5'te gerileme oranlarının yıllar arasında değişmesi de uzun dönemli SKAE artışının kesintisiz ve doğrusal bir süreç olmadığını desteklemektedir.



Şekil 2. DTW Tabanlı Kümelerde Yıllara Göre SKAE Gerilemesi Yaşayan Ülkelerin Oranı (%)

Kümelerin zamansal özelliklerinin yanı sıra mekânsal dağılımları da coğrafi haritalama yoluyla incelenmiştir (Şekil 3). Haritalama sonuçları, DTW ile belirlenen kümelerde belirli bölgesel yoğunlaşmalar bulunduğunu göstermektedir. Özellikle yüksek SKAE düzeyi ve görece istikrarlı gelişim örüntüsüyle tanımlanan Küme 1'in ağırlıklı olarak Avrupa ile diğer yüksek performanslı ülkelerde; düşük SKAE düzeyindeki Küme 2'nin ise büyük ölçüde Sahra Altı Afrika ülkelerinde yoğunlaştığı görülmektedir. En düşük ve dalgalı profile sahip dört ülkeden oluşan Küme 4 de kırılgan sürdürülebilir kalkınma örüntüsünün mekânsal olarak belirli ülkelerde yoğunlaştığını göstermektedir. Buna karşılık Küme 3 ve Küme 5'in coğrafi dağılımı daha geniş ve heterojen bir yapı sergilemektedir.



Şekil 3. DTW Tabanlı Ülke Kümelerinin Mekânsal Dağılımı

Bu bulgular, DTW tabanlı kümelerin yalnızca SKAE düzeyleri bakımından değil, gelişim hızı, yıllık dalgalanma ve mekânsal dağılım açısından da farklı profiller oluşturduğunu göstermektedir. Kümelerin sürdürülebilir kalkınmanın hangi boyutlarında farklılaştığının belirlenmesi amacıyla izleyen bölümde 2025 yılına ait SKA 1–SKA 17 göstergeleri karşılaştırılmıştır.

3.4. Kümelerin Sürdürülebilir Kalkınma Amaçları Açısından Karşılaştırılması

DTW tabanlı kümeleme sonucunda belirlenen ülke gruplarının sürdürülebilir kalkınmanın farklı boyutları bakımından nasıl ayrıştığını değerlendirmek amacıyla kümelerin 2025 yılı SKA 1–SKA 17 gösterge düzeyleri karşılaştırılmıştır. Bu analiz, ülkelerin 2000 ile 2025 yılları arasındaki SKAE gelişim örüntülerine göre oluşturulan kümelerin, dönem sonundaki sürdürülebilir kalkınma yapıları bakımından da farklılaşıp farklılaşmadığının belirlenmesini amaçlamaktadır.

Her bir SKA için kümeler arasındaki farklılıklar Welch ANOVA ile incelenmiş ve 17 ayrı karşılaştırmadan kaynaklanan çoklu test sorununu kontrol etmek amacıyla p değerlerine Benjamini–Hochberg FDR düzeltmesi uygulanmıştır. FDR düzeltmesi sonrasında 17 SKA'nın 16'sında kümeler arasında istatistiksel olarak anlamlı farklılık belirlenmiştir ($pFDR < ,05$). Yalnızca SKA 14 bakımından kümeler arasındaki farklılık istatistiksel olarak anlamlı bulunmamıştır ($F = 1,312$; $pFDR = ,278$). SKA 14 için Küme 4'te geçerli gözlem bulunmadığından bu karşılaştırma dört küme üzerinden gerçekleştirilmiştir.

Welch ANOVA sonuçlarının sağlamlığını değerlendirmek amacıyla gerçekleştirilen Kruskal–Wallis analizleri de aynı anlamlılık örüntüsünü ortaya koymuştur. FDR düzeltmesi sonrasında yine 17 SKA'nın 16'sında kümeler arasında anlamlı farklılık belirlenirken, yalnızca SKA 14 için farklılık anlamlı

bulunmamıştır (H = 3,551; pFDR = ,314; ε^2 = ,004). Böylece Welch ANOVA ve Kruskal–Wallis analizlerinin 17 SKA'nın tamamında aynı anlamlılık kararına ulaşması, kümeler arasındaki farklılıklara ilişkin bulguların farklı test yaklaşımları karşısında tutarlı olduğunu göstermektedir.

Kruskal–Wallis analizlerine ilişkin epsilon-kare (ε^2) etki büyüklükleri, kümeler arasındaki farklılaşmanın SKA boyutlarına göre değiştiğini göstermektedir. En yüksek etki büyüklükleri SKA 1 (ε^2 = ,790), SKA 11 (ε^2 = ,779), SKA 6 (ε^2 = ,768), SKA 3 (ε^2 = ,761) ve SKA 9'da (ε^2 = ,740) elde edilmiş; bu amaçların DTW tabanlı ülke kümelerini en belirgin biçimde ayrıştıran boyutlar arasında olduğu görülmüştür. Buna karşılık SKA 14 için etki büyüklüğünün oldukça düşük olması (ε^2 = ,004), bu amaç bakımından kümeler arasındaki ayrışmanın son derece sınırlı olduğunu göstermektedir.

Anlamlı farklılığın hangi kümelerden kaynaklandığını belirlemek amacıyla uygulanan Games–Howell çoklu karşılaştırma testi, kümeler arasındaki ayrışmanın SKA'lara göre farklı bir yapı sergilediğini ortaya koymuştur. SKA 1 ve SKA 2'de tüm ikili küme karşılaştırmaları anlamlı bulunurken, diğer amaçlarda bazı küme çiftleri arasında anlamlı farklılık belirlenmemiştir. Özellikle SKA 10, SKA 15 ve SKA 17'de genel küme farklılığı anlamlı olmasına rağmen bazı ikili karşılaştırmaların anlamlı olmaması, sürdürülebilir kalkınmanın farklı boyutlarında kümeler arasındaki ayrışmanın aynı ölçüde gerçekleşmediğini göstermektedir.

Genel olarak bulgular, ülkelerin uzun dönemli SKAE gelişim örüntülerine göre oluşturulan kümelerin 2025 yılındaki SKA profilleri bakımından da önemli ölçüde farklılaştığını, ancak bu farklılaşmanın bütün amaçlarda aynı büyüklükte olmadığını ortaya koymaktadır. Bu sonuç, ülkelerin sürdürülebilir kalkınma performanslarının yalnızca genel SKAE düzeyi ve zamansal

gelişim örüntüsüyle değil, endeksi oluşturan farklı amaçlar bakımından da ayrıştığını göstermektedir.

4. SONUÇ

Bu çalışmada, 169 ülkenin 2000–2025 dönemindeki sürdürülebilir kalkınma performansı, Sürdürülebilir Kalkınma Amaçları Endeksi (SKAE) zaman serileri üzerinden zamansal ve mekânsal bir yaklaşımla incelenmiştir. Ülkelerin uzun dönemli gelişim örüntülerinin benzerliğini belirlemek amacıyla Dinamik Zaman Bükme (DTW) uzaklığı ile hiyerarşik kümeleme birlikte kullanılmış; elde edilen kümeler zamansal değişim ve istikrar özellikleri, mekânsal dağılımları ve 2025 yılı SKA performansları açısından karşılaştırılmıştır.

Araştırma bulguları, incelenen ülkelerin tamamında 2025 yılı SKAE düzeyinin 2000 yılındaki başlangıç düzeyinin üzerinde olduğunu, ancak bu gelişimin doğrusal ve kesintisiz gerçekleşmediğini göstermiştir. Ülkelerin 167'sinde dönem içerisinde en az bir yıllık gerileme gözlenmesi, uzun dönemli sürdürülebilir kalkınma performansının yalnızca başlangıç ve bitiş değerleriyle değerlendirilmesinin ülkelerin gelişim dinamiklerini tam olarak yansıtmadığını ortaya koymaktadır. Ayrıca başlangıç SKAE düzeyi ile toplam değişim arasındaki negatif yönlü ilişki, başlangıç performansı daha düşük ülkelerin genel olarak daha yüksek artış gösterme eğiliminde olduklarına işaret etmektedir. Bu bulgular, ülkelerin sürdürülebilir kalkınma performanslarının farklı gelişim yörüngeleri izlediğini ortaya koyan güncel çalışmalarla uyumludur (Ma et al., 2025).

DTW tabanlı analiz sonucunda ülkeler beş kümede sınıflandırılmış ve bu kümelerin yalnızca SKAE düzeyleri bakımından değil, gelişim hızları ve yıllık dalgalanma özellikleri açısından da farklılaştığı belirlenmiştir. Bazı kümelerde yüksek performans düzeylerine daha yavaş ve istikrarlı bir gelişim eşlik

ederken, bazı kümelerde daha yüksek gelişim hızının daha belirgin dönemsel dalgalanmalarla birlikte gerçekleştiği görülmüştür. Kümelerin coğrafi dağılımları da belirli bölgesel yoğunlaşmalar ortaya koymuş, ancak tüm kümeler bütünüyle coğrafi sınırlarla açıklanabilecek bir yapı sergilememiştir. Bu sonuç, sürdürülebilir kalkınma performansının zaman ve mekân boyutlarında ülkeler arasında farklılaştığını gösteren Xu et al. (2025) bulgularıyla da örtüşmektedir.

Kümelerin 2025 yılı SKA profilleri açısından karşılaştırılması, 17 amacın 16'sında istatistiksel olarak anlamlı farklılık bulunduğunu göstermiştir. Parametrik ve parametrik olmayan analizlerin aynı anlamlılık örüntüsüne ulaşması bulguların sağlamlığını desteklerken, etki büyüklükleri kümeler arasındaki ayrışmanın amaçlara göre farklı düzeylerde gerçekleştiğini ortaya koymuştur. Bu sonuç, uzun dönemli genel sürdürülebilir kalkınma örüntülerine göre belirlenen ülke gruplarının sürdürülebilir kalkınmanın alt boyutları bakımından da önemli ölçüde ayrışabildiğini göstermektedir.

Çalışmanın temel yöntemsel katkısı, ülke performanslarını tek bir yıl veya başlangıç-bitiş değişimi üzerinden sınıflandırmak yerine, 26 yıllık zaman serilerinin bütünsel gelişim örüntülerini dikkate alan DTW tabanlı bir yaklaşım kullanmasıdır. Bu yaklaşımın zamansal profil göstergeleri, coğrafi haritalama ve amaç bazlı istatistiksel karşılaştırmalarla bütünleştirilmesi, ülkeler arasındaki sürdürülebilir kalkınma farklılıklarının çok boyutlu biçimde değerlendirilmesine olanak sağlamaktadır.

Çalışmanın bazı sınırlılıkları bulunmaktadır. Analizler SKAE ve bu endeksi oluşturan SKA skorlarına dayandığından, ülkelerin sürdürülebilir kalkınma süreçlerini etkileyebilecek ekonomik, kurumsal ve çevresel belirleyiciler doğrudan modellenmemiştir. Ayrıca amaç bazlı karşılaştırmalar dönem

sonunu temsil eden 2025 yılı üzerinden gerçekleştirilmiştir. Gelecek çalışmalarda ülke kümelerinin zaman içerisindeki yapısal değişimleri farklı dönemler için incelenebilir ve kümeler arasındaki farklılaşmanın ekonomik, sosyal, kurumsal ve çevresel belirleyicileri çok değişkenli modeller aracılığıyla araştırılabilir.

KAYNAKÇA

Bigerna, S., & Micheli, S. (2025). Is there worldwide convergence toward the SDGs? *Journal of Policy Modeling, 47*(1), 97–117. https://doi.org/10.1016/j.jpolmod.2024.12.002

Dai, R., & Zheng, C. (2023). False discovery rate-controlled multiple testing for union null hypotheses: A knockoff-based approach. *Biometrics, 79*(4), 3497–3509. https://doi.org/10.1111/biom.13848

Doğruel Anuşlu, M., & Fırat, S. Ü. (2020). Ülkelerin Endüstri 4.0 seviyesinin sürdürülebilir kalkınma düzeylerine etkisinin analizi. *Endüstri Mühendisliği, 31*, 44–58.

Fang, K., Xu, A., Wang, S., Jia, X., Liao, Z., Tan, R. R., Sun, H., & Su, F. (2023). Progress towards Sustainable Development Goals in the Belt and Road Initiative countries. *Journal of Cleaner Production, 424*, 138808. https://doi.org/10.1016/j.jclepro.2023.138808

Holder, C., Middlehurst, M., & Bagnall, A. (2024). A review and evaluation of elastic distance functions for time series clustering. *Knowledge and Information Systems, 66*(2), 765–809. https://doi.org/10.1007/s10115-023-01952-0

Iacobucci, D., Popovich, D. L., Moon, S., & Román, S. (2023). How to calculate, use, and report variance explained effect size indices and not die trying. *Journal of Consumer Psychology, 33*(1), 45–61. https://doi.org/10.1002/jcpy.1292

Jena, S., & Basel, S. (2025). Classifying global economies based on Sustainable Development Goals: A data-driven clustering approach. *Sustainable Development, 33*(3), 4543–4556. https://doi.org/10.1002/sd.3362

Kartal, B. (2025). Zaman serisi kümeleme yöntemi ile Küresel Cinsiyet Uçurumu Endeksinin analizi. *Uluslararası Ekonomi, İşletme ve Politika Dergisi, 9*(1), 420–437. https://doi.org/10.29216/ueip.1633127

Ma, F., Wang, H., Tzachor, A., Hidalgo, C. A., Schandl, H., Zhang, Y., ... Fu, B. (2025). The disparities and development trajectories of nations in achieving the sustainable development goals. *Nature Communications, 16*, 1107. https://doi.org/10.1038/s41467-025-56076-6

Maharaj, E. A., De Giovanni, L., D'Urso, P., & Bhattacharya, M. (2024). Deployment of renewable energy sources: Empirical evidence in identifying clusters with dynamic time warping. *Social Indicators Research, 175*, 741–762. https://doi.org/10.1007/s11205-022-03050-0

Marti, L., & Cervelló-Royo, R. (2023). Disparities in sustainable development goals compliance and their association with country risk. *Sustainable Development, 31*(4), 3038–3051. https://doi.org/10.1002/sd.2568

Özekin, A., & Çelik, F. (2024). Evaluation of sustainable development indices of countries: Global sustainability map. *Journal of Management and Economics Research, 22*(1), 171–199.

Sachs, J. D., Lafortune, G., Fuller, G., & Iablonovski, G. (2026). *Implementing sustainable development: 2030 and beyond. Sustainable Development Report 2026*. Paris: Sustainable Development Solutions Network; Dublin: Dublin University Press. https://doi.org/10.82163/225

Vera, J. F., & Angulo, J. M. (2023). An MDS-based unifying approach to time series K-means clustering: Application in the dynamic time warping framework. *Stochastic*

Environmental Research and Risk Assessment, 37, 4555–4566. https://doi.org/10.1007/s00477-023-02470-9

Xu, Z., Chen, X., Jiang, Q., Wu, X., Bhattarai, N., Mullen, J., ... Liu, J. (2025). Assessing global sustainability performance, imbalance, and coordination over space and time. *Nature Communications, 16*, 9186. https://doi.org/10.1038/s41467-025-64219-y

Yuan, H., Wang, X., Gao, L., Wang, T., Liu, B., Fang, D., & Gao, Y. (2023). Progress towards the Sustainable Development Goals has been slowed by indirect effects of the COVID-19 pandemic. *Communications Earth & Environment, 4*, 184. https://doi.org/10.1038/s43247-023-00846-x

KAPASİTE KISITLI VE ZAMAN PENCERELİ ARAÇ ROTALAMA PROBLEMİ İLE MALİYET OPTİMİZASYONU: LOJİSTİK SEKTÖRÜNDE BİR UYGULAMA

Emre EKİN[1]

Bahar IŞILAR[2]

1. GİRİŞ

Lojistik sektörü günümüzde hızla değişen ve pek çok farklı koşula göre şekillenen bir alan hâline gelmiştir. Uluslararası taşımacılık yapan firmalar; teslimat beklentileri, yasal sürüş kuralları, farklı ülkelerdeki çalışma saatleri ve trafik şartları gibi birçok faktörü aynı anda yönetmek zorunda kalmaktadır. Ayrıca bu faktörlere yönelik doğru planlama yapılmadığında ise gecikme ve maliyet artışlarıyla karşılaşmaktadır. Son yıllarda zaman pencereleri içinde teslimatın daha hassas takip edilmesi, zamanında teslimatın lojistik sektöründe kritik öneme sahip olması gibi durumlar firmaların rekabet gücünü doğrudan etkilemeye başlamıştır. Lojistik sektöründe firmaların rekabet gücüne sahip olabilmesi için zamanında teslimat, doğru rota ve minimum maliyetli dağıtımın gerçekleşmesi gerekmektedir.

Türkiye'de karayolu taşımacılığı lojistik sektörünün önemli bir bölümünü oluşturmaktadır. Nitekim 2025 yılında ülkemiz ihracatının yaklaşık %32'si karayoluyla gerçekleştirilmiştir (T.C. Ticaret Bakanlığı, 2026). Avrupa'ya yapılan ihracat taşımalarının çok büyük bir kısmı hâlâ tırlarla

[1] Doç. Dr., Milli Savunma Üniversitesi, ORCID: 0000-0002-4043-9750.

[2] Endüstri Mühendisi, ORCID: 0009-0007-6758-2411.

gerçekleştirilmekte; İngiltere'ye yapılan taşımalarda feribot kullanılması, gece teslimat zorunlulukları ve farklı ülkelerdeki pazar/hal çalışma saatleri gibi unsurlar planlamayı daha da karmaşık hâle getirmektedir. Çok sayıda bileşeni içerisinde barındıran uluslararası taşımacılıkta verimliliğin sağlanmasında; rotaların kısalması, boşta geçen sürelerin azalması ve düzenli teslimat durumları maliyetleri düşürmekte, yakıt tasarrufu ve güvenilirlik artışı gibi katkılarla genel ekonomik etkinliği de desteklemektedir.

Bu çalışmada, uluslararası karayolu taşımacılığı yapan bir firmanın Antalya'daki ana deposundan Almanya ve İngiltere'deki müşterilere gerçekleştirdiği taşıma operasyonlarında rota planlama problemi ele alınmaktadır. Firmanın mevcut rota planlarının zaman, mesafe ve operasyonel maliyet açısından uygunluğu değerlendirilerek; araç kapasitesi, teslimat zaman pencereleri, sürüş ve dinlenme süreleri ile İngiltere taşımalarında kullanılan feribot geçişleri gibi kısıtlar altında toplam taşıma maliyetini minimize eden alternatif rota ve teslimat planının belirlenmesi amaçlanmaktadır. Bu doğrultuda, firmanın mevcut uygulamasında kullanılan rota planları ile matematiksel optimizasyon modeli sonucunda elde edilen rota planlarının karşılaştırılması ve mevcut planlamaya kıyasla sağlanabilecek maliyet avantajının ortaya konulması hedeflenmektedir.

Bu çalışmanın amacı, uluslararası karayolu taşımacılığı yapan bir firmanın mevcut rotalarını analiz ederek, operasyonel ve finansal açıdan daha uygun bir rota planlamasının mümkün olup olmadığını belirlemektir. Mevcut rotaların süre, mesafe ve toplam operasyonel maliyet açısından uygunluğu sayısal olarak değerlendirilecektir.

Çalışmada; teslimat noktaları, zaman aralıkları, sürüş ve dinlenme kuralları ile feribot kullanımı dikkate alınarak problem matematiksel olarak modellenmektedir. Mevcut uygulama ile

önerilen model karşılaştırılarak zaman, mesafe ve maliyet açısından sağlanabilecek iyileştirmeler ortaya konulacaktır. Bu kapsamda problem, araçların taşıma kapasiteleri ve müşterilere ait teslimat zaman pencereleri başta olmak üzere sürüş-dinlenme süreleri ve feribot geçişleri gibi operasyonel kısıtlar dikkate alınarak Kapasite Kısıtlı ve Zaman Pencereli Araç Rotalama Problemi (Capacitated Vehicle Routing Problem with Time Windows, CVRPTW) şeklinde modellenmiştir. Modelin amaç fonksiyonu, taşıma operasyonuna ilişkin toplam maliyetin minimize edilmesine yöneliktir. Bu kapsamda rota mesafeleri, otoyol geçiş ücretleri, akaryakıt tüketimi ve personel maliyetleri gibi çalışmada kullanılan maliyet unsurları dikkate alınarak optimum rota planı belirlenmiştir. Model GAMS optimizasyon yazılımı kullanılarak çözülmüş ve elde edilen sonuçlar firmanın mevcut rota planları ile karşılaştırılmıştır. Böylece, uluslararası karayolu taşımacılığında belirlenen operasyonel kısıtlar altında toplam taşıma maliyetinin matematiksel optimizasyon yoluyla azaltılabilirliği ortaya konulmuştur.

2. LİTERATÜR ARAŞTIRMASI

Karagül, Tokat ve Aydemir (2016), kapasite kısıtlı araç rotalama probleminde başlangıç rotalarının oluşturulması için ağırlık merkezi esaslı yeni bir sezgisel algoritma önermiş ve yöntemin klasik kurucu sezgisellere göre daha kaliteli başlangıç çözümleri ürettiğini göstermiştir.

Göçken, Yaktubay ve Kılıç (2018), Zaman Pencereli Araç Rotalama Problemi için süpürme ve en yakın komşu sezgisellerinin genetik algoritma ile melezleştirildiği çok amaçlı bir yaklaşım geliştirerek toplam yolu ve araç sayısını birlikte azaltmıştır.

Atasagun ve Karaoğlan (2019), şehir içi trafiğe bağlı hız değişkenliğini Eş Zamanlı Dağıtım ve Toplamalı Araç Rotalama

Problemine (Vehicle Routing Problem with Simultaneous Pickup and Delivery-VRPSPD) zaman bağımlılığı boyutuyla entegre eden bir MILP modelini GAMS ile çözmüştür.

Kay, Karagül, Şahin ve Gündüz (2021), uzun mesafeli çok duraklı kamyon taşımacılığında toplam lojistik maliyetini azaltan tasarruf tabanlı bir rota birleştirme prosedürünün, gerçek veri setlerinde Clarke-Wright algoritmasından daha başarılı olduğunu göstermiştir.

Mor vd. (2022), uzun mesafeli kamyon taşımacılığında yakıt ikmali ile sürücü çalışma-dinlenme kısıtlarını birlikte ele alan iki amaçlı bir rota planlama modeliyle gerçek yol ağında belirgin yakıt tasarrufu sağlamıştır.

Boz vd. (2023), çevresel sürdürülebilirlik, zaman pencereleri ve eş zamanlı toplama-dağıtım kısıtlarını birlikte ele alan Yeşil Zaman Pencereli Eş Zamanlı Topla-Dağıt Araç Rotalama Problemi için bir matematiksel model geliştirmiştir. Problemin çözümünde Genetik Algoritma (GA) ve Ağırlıklı Süperpozisyon Çekim Algoritması (ASÇA) kullanılmış; sonuçlar, GA'nın çözüm kalitesi ve çözüm süresi açısından daha başarılı olduğunu göstermiştir.

Tutumlu vd. (2025), uluslararası taşımacılık yapan bir lojistik firması için araç rotaları ve yakıt ikmal noktalarını birlikte ele alan iki amaçlı bir model geliştirmiştir. Gerçek işletme verileriyle yapılan uygulamada, yakıt maliyetinde %57, sınır bekleme sürelerinde ise %25 iyileşme elde edilmiştir.

Cavecchia vd. (2025), bir ilaç dağıtım firmasının çok depolu ve çok seferli Zaman Pencereli Araç Rotalama Problemini (Vehicle Routing Problem with Time Windows – VRPTW) ayrıştırma tabanlı bir meta-sezgiselle çözerek katı teslimat pencerelerine uyumla filo maliyetini minimize etmiştir.Chen, Gao ve Chen (2025), VRPTW için önerdikleri IGA-ACO

yöntemiyle Solomon test problemlerinde, özellikle R sınıfı örneklerde araç kullanımını %24,45 azaltmıştır.

Çetintaş ve Aslan (2026), bir kargo firmasının eş zamanlı dağıtım-toplamalı, kapasite ve zaman kısıtlı rotalama problemini GAMS ile çözerek maliyette %18,2, mesafede %20,5'e varan iyileştirme sağlamıştır.

3. ARAÇ ROTALAMA PROBLEMİ (ARP)

İşletmeler ürünlerini pazara sunarken yoğun bir rekabet ortamıyla karşılaşmakta; bu rekabet koşullarında rakiplerinden farklılaşmak ve operasyonel maliyetleri düşürmek amacıyla lojistik süreçlerini optimize etmeyi hedeflemektedir. Teknolojik gelişmeler doğrultusunda tüketiciler kolay erişilebilir ürün ve hizmetleri tercih etmeye başlamış; bu tercihlerini de minimum maliyet ve maksimum teslimat hızı beklentisiyle şekillendirmiştir. Tüketici davranışlarındaki bu dönüşümün bir sonucu olarak e-ticaret platformlarının işlem hacimlerinde dikkate değer bir artış gözlenmiştir. İşletmelerin artan bu talepleri en hızlı ve en düşük maliyetle karşılama zorunluluğu, literatürde Araç Rotalama Problemini (ARP) kritik bir araştırma konusu haline getirmiştir. Temel olarak ARP, araçların en kısa mesafe üzerinden bir turu tamamlamasını hedefleyen bir optimizasyon problemidir (Tutumlu vd., 2025).

Taşımacılık, nakliye ve kargo şirketleri gibi lojistik sektörünü doğrudan etkileyen alanlarda etkin dağıtımın sağlanabilmesi; en kısa güzergâhın belirlenmesi, toplam maliyetlerin düşürülmesi ve kullanılan araç sayısının azaltılması ile doğrudan ilişkilidir. Bu doğrultuda ARP modelleri, belirli bir merkezde kümelenmiş ve belirli kapasite kısıtlarına sahip araçların kat ettiği mesafeyi, seyahat süresini, maliyetleri ve araç sayısını minimize ederken; sistem verimliliğini, müşteri memnuniyetini ve kapasite kullanım oranlarını maksimize etmeyi

amaçlamaktadır. Dolayısıyla ARP, belirli kapasiteye sahip araçların minimum maliyet, süre ve mesafe harcayarak maksimum iş çıktısı üretmesi ve başlangıç noktasına geri dönmesi esasına dayanan, farklı sektörel uygulamalarda çok çeşitli varyasyonları geliştirilmiş bir kombinatoryal optimizasyon problemidir (Atasagun & Karaoğlan, 2019).

Son yıllarda literatürde analizin ana hedefi olan en önemli ARP dallarından biri ise ZARP (Zaman Pencereli Araç Rotalama Problemi) adı verilen, belirli bir zaman aralığında hizmetin herhangi bir müşteride başladığı ARP'nin bir genelleştirilmesidir (Şahin Macit, 2024; Ercan Cömert vd., 2015).

3.1. Kapasite Kısıtlı ve Zaman Pencereli Araç Rotalama (CVRPTW)

Lojistik ve tedarik zinciri yönetiminde temel amaç, coğrafi olarak farklı noktalarda bulunan müşteri taleplerinin karşılanması sırasında toplam seyahat mesafesi, seyahat süresi, yakıt tüketimi ve filo maliyetlerinin mümkün olduğunca azaltılmasıdır. Bu operasyonel yapının yöneylem araştırmasındaki temel matematiksel karşılığı Araç Rotalama Problemi (ARP)'dir. Klasik ARP, bir depodan hareket eden araçların belirli müşterileri ziyaret ederek dağıtım veya toplama faaliyetlerini gerçekleştirmesi ve tur sonunda başlangıç noktasına dönmesini sağlayan NP-zor bir kombinatoryal optimizasyon problemidir.

Gerçek saha operasyonlarında araçların taşıma kapasiteleri sınırlı olmakta ve müşterilerin hizmet kabul ettiği zaman aralıkları farklılık göstermektedir. Bu nedenle klasik ARP'ye kapasite ve zaman penceresi kısıtlarının eklenmesiyle Kapasite Kısıtlı ve Zaman Pencereli Araç Rotalama Problemi (Capacitated Vehicle Routing Problem with Time Windows – CVRPTW) ortaya çıkmaktadır. CVRPTW'nin temel unsurları aşağıdaki şekilde özetlenebilir:

Kapasite Kısıtı: Her aracın taşıyabileceği maksimum yük miktarı belirlenmiştir. Bir rotadaki müşterilerin toplam talebi, ilgili aracın taşıma kapasitesini aşamaz.

Zaman Penceresi Kısıtı: Her müşterinin hizmet alabileceği belirli bir zaman aralığı bulunmaktadır. Araç müşteriye bu aralıktan önce ulaşırsa hizmet başlangıç zamanına kadar bekler; zaman penceresinin sona ermesinden sonra gerçekleşen varış ise kısıt ihlali oluşturur.

Akış Dengesi ve Tur Bütünlüğü: Her müşteri tam olarak bir kez ziyaret edilmeli ve müşteriye giriş yapan araç aynı noktadan çıkış yapmalıdır. Bu yapı, geçersiz alt turların oluşmasını önleyerek araçların rotalarını başlangıç deposunda tamamlamasını sağlar.

Bu özellikleriyle CVRPTW, araç kapasitesi ile müşterilerin hizmet zamanlarını eş zamanlı olarak dikkate alan temel rota optimizasyon modellerinden biridir. Problem, farklı lojistik ve dağıtım uygulamalarında karma tamsayılı programlama ve sezgisel ya da meta-sezgisel yöntemler kullanılarak çözülmektedir (Al-Nabulsi vd., 2025; Lutfizein ve Nuha, 2026).

4. MEVCUT SİSTEM ANALİZİ

4.1. Firma Bilgisi

Uygulamada ele alınan firma; 1996 yılında kurulmuş, merkezi Samandağ/Hatay'da bulunan ve uluslararası karayolu taşımacılığı alanında faaliyet gösteren bir lojistik firmasıdır. Bünyesinde 26 şoför, 1 muhasebeci ve 1 destek elemanı olmak üzere 28 çalışan bulunmaktadır. Ayrı bir planlama birimi bulunmamakta, planlama manuel ve deneyime dayalı olarak yürütülmektedir. Sektörü ve firmayı en çok zorlayan unsurlar yakıt ve Avrupa'daki otoyol ücretleridir. Bu nedenle rota seçimi

ve seyahat süresi lojistik firması açısından kritik olup mevcut planlama sürecinin daha sistematik yöntemlerle değerlendirilmesi önem taşımaktadır.

4.2. Firma Mevcut Çalışma Durumu

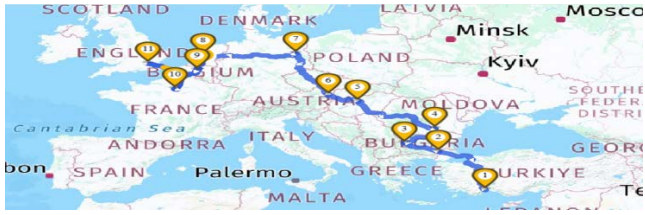
İncelenen firmada uluslararası taşımalar için kullanılan rotalar, büyük ölçüde sevkiyat sırasında gerekli olan belgelerin durumuna göre belirlenmektedir. Firma, bazı belgelerin geçerliliğine ve ülkeler arası geçiş koşullarına bağlı olarak önceden belirlenmiş iki ana rota arasından birini tercih etmektedir. Firma, mevcut rotaları belirlerken teslimat süreleri ve operasyonel koşulları dikkate almaktadır. Ancak bu değerlendirme yapılırken sayısal bir karşılaştırma yerine geçmiş tecrübelere dayanarak yapılmaktadır. Bu çalışma, mevcut uygulamayı temel alarak daha sistematik bir değerlendirme sunmayı hedeflemektedir.

4.3. Mevcut Taşıma Güzergahları ve Otoyol Detayları

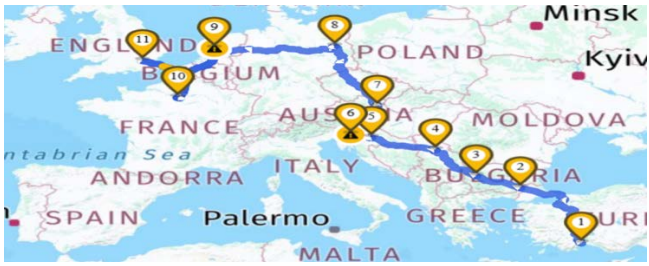
Firma yetkililerinden alınan bilgiler doğrultusunda, operasyonel süreçlerde kullanılan iki temel güzergahın teknik detayları aşağıda analiz edilmiştir:

Güzergah1: Antalya- Kapıkule- Bulgaristan- Romanya- Macaristan- Avusturya- Almanya- Hollanda- Belçika- Fransa- İngiltere

Güzergah2: Antalya- Kapıkule- Bulgaristan- Sırbistan- Hırvatistan- Slovenya-Avusturya- Almanya- Hollanda- Belçika- Fransa- İngiltere



Şekil 1. Güzergah1 Kuzey Hattı



Şekil 2. Güzergah2 Batı Balkan Hattı

4.4. Güzergah Ülkeleri Akaryakıt Fiyatları

Çalışmada kurulan matematiksel modelin maliyet analizlerinde kullanılmak üzere, güzergah üzerinde yer alan ülkelerin güncel motorin (Dizel) birim fiyatları derlenmiştir. Nisan 2026 dönemi itibarıyla Tolls.eu ve ilgili enerji piyasası endekslerinden alınan ortalama pompa fiyatları Tablo 1'de sunulmuştur. Bu veriler, çalışma kapsamında geliştirilen "En Düşük Maliyetli Yakıt İkmal Stratejisi"nin belirlenmesinde temel girdi parametresi olarak kullanılmış ve modelin yakıt arbitrajı kısıtlarının hesaplanmasında ana veri seti olarak modele dahil edilmiştir.

Tablo 1. Birim Yakıt Fiyatları Tablosu

Ülke	Birim Fiyat (€/Litre)
Slovenya	1,87 €
Hırvatistan	1,96 €
Macaristan	1,72 €
Romanya	1,81 €
Avusturya	1,95 €
Belçika	2,19 €
Almanya	2,13 €
Fransa	2,24 €
Sırbistan	1,85 €
Hollanda	2,30 €

4.5. Otoyol Geçiş Ücretleri İçin Varsayımlar ve Kısıtlar

Tüm maliyet hesaplamaları, uluslararası taşımacılık standartlarına uygun 5 akslı ve Euro 6 emisyon sınıfına sahip 40 tonluk (Brüt) araç baz alınarak yapılmıştır. Yerel para birimi kullanan ülkeler (Sırbistan, Bulgaristan, Macaristan, İngiltere) için Ocak 2026 Merkez Bankası çapraz kurları baz alınarak maliyetler Euro cinsine çevrilmiştir. Bu tarihten sonraki dönemde, rotadaki ülkelerin otoyol ücret tarifelerinde majör bir değişiklik (enflasyon düzeltmesi dışında) gözlemlenmediği için, söz konusu veriler çalışmanın temel maliyet hesaplamaları için geçerli kabul edilmiştir.

Avusturya gibi gece/gündüz tarifesi farklılık gösteren ülkelerde, operasyonel gerçekliğe uygun olarak ağırlıklı ortalama birim maliyet kullanılmıştır. Çalışmamızda, Avrupa rotası üzerinde yer alan ülkelerin ortalama kilometre başı otoyol kullanım bedelleri, Tolls.eu veritabanından alınan güncel veriler ışığında hesaplanmış ve ülke bazlı maliyetler Tablo 2'de sunulmuştur (Tolls.eu, 2026).

Tablo 2. Otoyol Geçiş Ücretleri Tablosu

Ülke	Birim Maliyet (€)	Birim
Almanya	0,348	Km Başına
Avusturya	0,56	Km Başına
Macaristan	0,43	Km Başına
Slovenya	0,35	Km Başına
Fransa	0,32	Km Başına
Sırbistan	0,26	Km Başına
Belçika	0,204	Km Başına
Hırvatistan	0,18	Km Başına
Bulgaristan	0,13	Km Başına
Manş Denizi	200	Sabit (Sefer)
Türkiye	90	Sabit (Toplam)
Romanya	29	Sabit (Toplam)
İngiltere	15	Günlük
Hollanda	13	Günlük

4.6. Operasyonel Ek Maliyetler (Personel ve Bakım):

Modelin maliyet fonksiyonunu oluşturan diğer kritik kalemler; sürücü ödemeleri ve araç bakım giderleridir. Firma yetkililerinden alınan güncel veriler doğrultusunda şu maliyetler modele "Sefer Başına Sabit Maliyet" (C_{fixed}) olarak dahil edilmiştir.

Sürücü Maliyeti: Sürücülere, tamamlanan her uluslararası sefer (gidiş-dönüş) başına sabit 1.200 € ödeme yapılmaktadır.

Araç Bakım Maliyeti: Araçlar her sefer dönüşünde periyodik bakıma girmektedir. Bu bakımın maliyeti ortalama 20.000- 25.000 TL'dir.

5. PROBLEM TANIMI VE MODEL KURULUMU

5.1. Problem Bilgisi

Uluslararası karayolu taşımacılığında teslimat zamanları, sürüş dinlenme kuralları, yasal düzenlemeler ve coğrafi koşullar planlama sürecini karmaşıklaştırmakta; hatalı planlama ise gecikme ve ek maliyetlere yol açmaktadır. Bu çalışmada, Türkiye'deki depolardan Almanya ve İngiltere'deki müşterilere yapılan taşımalar için firmanın kullandığı mevcut rotaların zaman, mesafe ve operasyonel maliyet açısından uygunluğu analiz edilerek minimum maliyetli alternatif rotaların belirlenmesi amaçlanmaktadır. Firma, belgelerin durumuna göre iki mevcut rotadan birini kullanmaktadır; ancak bu rotaların en uygun seçenekler olup olmadığı bilinmemektedir. Problem kapsamında, teslimat zaman pencereleri, sürüş-dinlenme kuralları ve İngiltere taşımalarında zorunlu feribot kullanımı dikkate alınmaktadır. Almanya teslimatları 23.00–06.00, İngiltere teslimatları ise 20.00–22.00 saatleri arasında gerçekleştirilmektedir. Bu doğrultuda, yasal ve operasyonel

kısıtlar altında toplam maliyeti en aza indiren alternatif rota ve teslimat planının mümkün olup olmadığı araştırılmaktadır.

Belirtilen bu operasyonel hedeflerin sağlıklı bir şekilde modellenebilmesi ve en uygun rotaların güvenilir bir altyapıyla hesaplanabilmesi için düğüm noktaları arasındaki mesafelerin gerçeğe en yakın şekilde tespiti büyük önem taşımaktadır. Bu doğrultuda çalışmada kullanılan mesafe ve süre hesaplamaları, şirketten temin edilen lokasyon verileri temel alınarak gerçekleştirilmiştir.

Dünya'nın küresel yapısı nedeniyle, iki nokta arasındaki mesafe öklid geometrisiyle hesaplandığında yüksek hata payı oluşmaktadır. Bu çalışmada, iki nokta arasındaki büyük çember mesafesini hesaplamak için Haversine formülü kullanılmıştır. Haversine formülü; enlem (φ) ve boylam (λ) koordinatlarını kullanarak, küresel bir yüzey üzerindeki iki nokta arasındaki en kısa mesafeyi aşağıdaki denklemle hesaplar:

Burada;

- d: İki nokta arasındaki mesafe,
- r: Dünya'nın yaklaşık yarıçapı (6.371 km),
- $\varphi 1, \varphi 2$: Noktaların enlem değerleri (radyan cinsinden),
- $\lambda 1$, $\lambda 2$: Noktaların boylam değerleri (radyan cinsinden) temsil etmektedir.

Bu yöntem ile Türkiye ve Avrupa arasındaki güzergahlarda oluşan coğrafi sapmalar minimize edilmiş ve model girdileri için yüksek doğrulukta bir mesafe matrisi oluşturulmuştur.

Formül:

$$d = 2r \arcsin\left(\sqrt{sin^2(\frac{\phi_2 - \phi_1}{2} + \cos(\phi_1)cos(\phi_2)sin^2(\frac{\lambda_2 - \lambda_1}{2})}\right)$$

5.2. Teslimat Noktaları ve Düğümler

Bu çalışmada her şehir ve teslimat noktası, ağ üzerinde bir düğüm olarak ele alınmıştır. Fiziksel olarak toplam 6 farklı şehir ve 8 farklı lokasyon bulunmasına rağmen, matematiksel modelin akış ve tur dengesini sağlayabilmek amacıyla sistemde toplam 9 düğüm tanımlanmıştır. Bu düğümlerden 1 numaralı düğüm Antalya çıkış deposunu; 2'den 8'e kadar olan düğümler Münih, Bielefeld, Londra, Dortmund ve Stuttgart'taki teslimat noktalarını; 9 numaralı düğüm ise araçların rotalarını tamamlayıp geri döndükleri Antalya varış deposunu (kukla düğüm) temsil etmektedir.

5.3. Varsayımlar ve Kısıtlar

- Çalışma kapsamında tüm araçların aynı teknik özelliklere sahip olduğu varsayılmıştır.
- Modelde kullanılan araçların tamamının taşıma kapasitesi eşittir (22 ton)
- Araçların sabit bir ortalama hız ile hareket ettiği varsayılmıştır.

5.3.1. Zaman Penceresi Kısıtı

Çalışmada ele alınan problemde, teslimat noktalarında teslimatlar yalnızca belirli zaman aralıklarında yapılmaktadır. Örneğin Almanya'daki teslimatlar gece 23.00 sabah 06.00 arası

yapılırken İngiltere'deki teslimatlar akşam 20.00-22.00 arası yapılmaktadır.

5.3.2. Sürüş ve Dinlenme Kısıtları

Uluslararası karayolu taşımacılığında sürücülerin uyması gereken yasal sürüş ve dinlenme kuralları bulunmaktadır. Günlük maksimum sürüş süresi 8 saattir ve şu şekilde gerçekleşir:

4 saat sürüş, 30-45 dk mola; 2 saat sürüş 45-60dk mola; 2 saat sürüş 45-60 dk mola.

Daha sonra araç ve sürücüler 8-9 saat dinlenir. Planlama yapılırken bu sürelerin dikkate alınması gerekir.

5.3.3. Feribot Kullanımı Kısıtı

İngiltere'ye yapılan taşımalar sırasında karayolu taşımacılığına ek olarak feribot kullanımı gerekmektedir. Feribot kullanımı ile (bekleme süresi +karşıya geçiş) yaklaşık 6 saatlik bir ek süre oluşmaktadır ve bu süre toplam seyahat süresine eklenir.

6. UYGULAMA

Bu bölümde, uluslararası karayolu taşımacılığı yapan bir lojistik işletmesinin Hatay/Samandağ merkezli ve Antalya çıkışlı operasyonel verileri üzerinde yürütülen araç rotalama ve yakıt optimizasyonu analizi ele alınmıştır. Problem, Kapasite Kısıtlı ve Zaman Pencereli Araç Rotalama Problemi tabanına dinamik yakıt ikmali ve arbitraj kısıtları entegre edilerek kurgulanmıştır. Model, GAMS/CPLEX ortamında çözülerek firmanın mevcut manuel planlama sonuçları ile karşılaştırılmıştır. Çalışmada amaç; lojistik ağındaki teslimat rotalarının belirlenmesi ve operasyonel maliyetlerin en aza indirilmesidir. Kurulacak olan matematiksel modelde filodaki araçların kat ettiği toplam seyahat ve otoyol

maliyetinin en küçüklenmesini hedefleyen amaç fonksiyonu oluşturulmuştur.

6.1. Veri Seti ve Parametrelerin Tanımlanması

Uygulama çalışmasında Antalya ana deposundan çıkan 5 adet aracın (Euro 6 emisyon sınıfında, 22 ton net kapasiteli) Almanya (Münih, Stuttgart, Dortmund, Bielefeld) ve İngiltere (Londra) bölgelerindeki 7 teslimat adresine sevkiyatı modellenmiştir. Ağ dengesinin sağlanması amacıyla 1 başlangıç deposu (Antalya), 7 müşteri lokasyonu ve araçların geri döneceği 1 bitiş deposu (Antalya - kukla düğüm) olmak üzere toplam 9 düğüm tanımlanmıştır. Düğümler arası mesafelerin hesaplanmasında yerkürenin eğriliğini dikkate alan Haversine formülü kullanılmış, ortalama 75 km/saat seyir hızı üzerinden seyahat süreleri elde edilmiştir. İngiltere hatlarına ise Manş Denizi geçişi için +6 saat feribot süresi eklenmiştir.

Tablo 3. Lojistik Ağı Düğümleri, Talep Miktarları ve Zaman Pencereleri

Düğüm (i)	Şehir/Lokasyon	Talep q_i (Ton)	Hizmet Süresi s_i (Saat)	Zaman Penceresi $[e_i, l_i]$ (Saat)
1	Antalya (Çıkış Deposu)	0.000	0.0	[0, 999]
2	Munich	8.565	2.0	[10, 50]
3	Stuttgart-1	7.513	2.5	[20, 70]
4	Dortmund	16.692	3.0	[40, 90]
5	Bielefeld	9.300	1.5	[50, 100]
6	Stuttgart-2	6.374	2.0	[60, 120]
7	Londra (Leyton)	20.566	3.0	[70, 150]
8	Londra (Enfield)	21.780	2.0	[80, 160]
9	Antalya (Dönüş Deposu)	0.000	0.0	[0, 999]

Tablo 3'te, lojistik ağında yer alan başlangıç ve varış depoları ile Almanya ve İngiltere'deki 7 farklı teslimat noktasına ait temel operasyonel parametreler gösterilmektedir. Müşterilerin CMR taşıma belgelerinden elde edilen gerçek talep miktarları 22 tonluk araç kapasitesi kısıtının temel girdisini oluştururken; her bir düğüm için tanımlanan $[e_i, l_i]$ zaman pencereleri, taze meyve-sebze taşımacılığındaki katı gece teslimat saatlerini ve toptancı hallerinin kabul aralıklarını operasyonel saat cinsinden tam olarak sınırlandırmaktadır.

Tablo 4. Düğümler Arası Haversine Mesafe (d_{ij} - km) ve Seyahat Süre (t_{ij} - Saat) Matrisi

Düğüm	1 (Antalya)	2 (Munich)	3 (Stuttgart-1)	4 (Dortmund)	5 (Bielefeld)	6 (Stuttgart-2)	7 (Londra-Leyton)	8 (Londra-Enfield)
1	0 km (0 s)	1997km (27 s)	2412km (29 s)	2909 km (33 s)	2445 km (32 s)	2182 km (29 s)	2916 km (45 s)	2182 km (45 s)
2	1997 km (27 s)	0 km (0 s)	482 km (2 s)	913 km (6 s)	477 km (6 s)	187 km (2 s)	919 km (18 s)	187 km (18 s)
3	2412 km (29 s)	482 km (2 s)	0 km (0 s)	595 km (4 s)	96 km (5 s)	362 km (0 s)	596 km (16 s)	362 km (16 s)
4	2909 km (33 s)	913 km (6 s)	595 km (4 s)	0 km (0 s)	519 km (1 s)	727 km (4 s)	12 km (13 s)	727 km (13 s)
5	2445 km (32 s)	477 km (6 s)	96 km (5 s)	519 km (1 s)	0 km (0 s)	328 km (5 s)	521 km (14 s)	328 km (14 s)
6	2182 km (29 s)	187 km (2 s)	362 km (0 s)	727 km (4 s)	328 km (5 s)	0 km (0 s)	734 km (16 s)	0 km (0 s)
7	2916 km (45 s)	919 km (18 s)	596 km (16 s)	12 km (13 s)	521 km (14 s)	734 km (16 s)	0 km (0 s)	734 km (0 s)
8	2182 km (45 s)	187 km (18 s)	362 km (16 s)	727 km (13 s)	328 km (14 s)	0 km (16 s)	734 km (0 s)	0 km (0 s)

Tablo 4'te, küresel koordinat sapmalarını gideren Haversine formülasyonu ile hesaplanmış gerçek karayolu mesafeleri ve bu mesafelere bağlı seyahat süreleri sunulmaktadır. Seyahat süreleri ortalama 75 km/saat çekici seyir hızı üzerinden türetilmiş olup, Manş Denizi geçişini gerektiren İngiltere bağlantılarına (4 ve 7 numaralı satır/sütunlar) 6 saatlik Ro-Ro feribot geçiş ve gümrük bekleme süreleri yansıtılmıştır. Bu matris, rotalama kararlarında hem yol maliyetinin hem de zaman sürekliliği kısıtlarının $(t_{jk} \geq t_{ik} + t_{ij} + s_i)$ hesaplanmasına doğrudan girdi sağlamaktadır.

Mesafe ve süre matrislerinin yanı sıra, optimizasyon modelinin amaç fonksiyonunu oluşturan temel maliyet parametreleri de sisteme tanımlanmıştır. Bu doğrultuda; sürücü harcırahı, bakım ve çıkış harçlarını içeren araç başına seferlik sabit maliyet $C_{fixed} = 1714.53$ €, kilometre başı ortalama otoyol maliyeti 0.35 €/km ve güzergah üzerindeki ülkelerin Nisan 2026 dönemi pompa motorin fiyatları (p_i: Türkiye 1.20 €, Romanya 1.81 €, Macaristan 1.72 €, Almanya 2.13 €, Belçika 2.19 €, Fransa 2.24 €) olarak modele aktarılmıştır.

6.2. Matematiksel Modelin Yapısı

Lojistik ağındaki teslimat rotalarının belirlenmesi ve operasyonel maliyetlerin en aza indirilmesi amacıyla Kapasite Kısıtlı ve Zaman Pencereli Araç Rotalama Problemi (CVRPTW) formülasyonu geliştirilmiştir. Modelin kümeleri, parametreleri,

karar değişkenleri ve kısıt denklemleri aşağıda detaylandırılmıştır:

Kümeler (Sets):

- N: Tüm düğümler kümesi (1: Başlangıç Deposu, C: Müşteriler kümesi, 9: Varış Deposu)
- V: Filodaki araçlar kümesi ($v_1, \dots, v_5$)

Parametreler (Parameters):

- c_{ij}: i noktasından j noktasına seyahat mesafesi
- t_{ij}: i noktasından j noktasına seyahat süresi (feribot geçişleri dahil)
- q_i: i müşterisinin talep miktarı
- Q: Maksimum araç kapasitesi (22 Ton)
- e_i: i noktası için en erken hizmete başlama zamanı
- l_i: i noktası için en geç hizmete başlama zamanı
- M: Yeterince büyük bir pozitif sayı (Big-M)
- d_{ij}: Müşteri noktaları (i, j) arasındaki gerçek otoyol mesafelerini (km) temsil eden 9×9 boyutlu mesafe matrisi
- c_d: Kilometre başına düşen toplam operasyonel maliyet katsayısı

Karar Değişkenleri (Decision Variables):

- x_{ijk}: k aracı i noktasından j noktasına gidiyorsa 1, aksi halde 0 (ikili karar değişkeni)
- t_{ik}: k aracının i noktasına varış zamanı (sürekli karar değişkeni)

Amaç Fonksiyonu ve Model Denklemleri:

Modelin amaç fonksiyonu, filodaki araçların kat ettiği toplam seyahat ve otoyol maliyetinin en küçüklenmesini hedeflemektedir:

$$\min Z = \sum_{k \in V} \sum_{i \in N} \sum_{j \in N} \left(d_{ij} \cdot c_d \cdot x_{ijk} \right) \tag{1}$$

Müşteri Ziyareti Kısıtı*:* Her müşteriye yalnızca tek bir araç tarafından ve tam olarak bir kez hizmet verilmesini zorunlu kılar:

$$\sum_{k \in V} \sum_{j \in N} x_{ijk} = 1 \quad \forall i \in C \tag{2}$$

Rota Akış Dengesi: Bir müşteri düğümüne giriş yapan aracın, teslimatı tamamladıktan sonra mutlaka o düğümden çıkış yapmasını garanti altına alır:

$$\sum_{i \in N} x_{ijk} - \sum_{i \in N} x_{jik} = 0 \quad \forall j \in C, \forall k \in V \tag{3}$$

Başlangıç ve Bitiş Deposu Kısıtları: Her aracın rotasına mutlaka Antalya başlangıç deposundan (Düğüm 1) çıkış yaparak başlamasını ve tüm dağıtımlarını bitirdikten sonra Antalya varış deposuna (Düğüm 9) dönerek turunu tamamlamasını sağlar:

$$\sum_{j \in C} x_{1,j,k} = 1 \quad \forall k \in V \tag{4}$$

$$\sum_{i \in C} x_{i,9,k} = 1 \quad \forall k \in V \tag{5}$$

Araç Kapasitesi Kısıtı: Bir araca atanan müşterilerin toplam talep miktarının, o çekicinin maksimum taşıma kapasitesi olan 22 tonu aşmasını engeller:

$$\sum_{i \in C} q_i \sum_{j \in N} x_{ijk} \leq Q \quad \forall k \in V \tag{6}$$

Zaman Sürekliliği ve Alt Tur Eleme Kısıtı: Bir aracın ardışık iki durak arasındaki varış zamanlarını bağlar; seyahat süresi (t_{ij}) ve boşaltma/hizmet süresi (s_i) toplamını ekleyerek mantıksal zaman sürekliliğini kurar ve yasadışı alt turların (sub-tours) oluşmasını engeller:

$$t_{jk} \geq t_{ik} + t_{ij} + s_i - M(1 - x_{ijk}) \quad \forall i, j \in N, \forall k \in V \tag{7}$$

Zaman Penceresi Kısıtları: Araçların müşteri noktalarına varış saatlerinin (t_{ik}), müşteriler tarafından şart koşulan en erken (e_i) ve en geç (l_i) teslimat kabul aralığında gerçekleşmesini temin eder:

$$e_i \leq t_{ik} \leq l_i \quad \forall i \in N, \forall k \in V \tag{8}$$

Yukarıda matematiksel olarak tanımlanan model, operasyonel gerçekliği artırmak adına istasyon bazlı akaryakıt alımlarını ve depo seviyelerini takip eden yakıt ikmali denklemleriyle genişletilerek GAMS ortamına aktarılmıştır.

6.3. Matematiksel Modelin GAMS Kodu

Geliştirilen model Karma Tam Sayılı Doğrusal Programlama (Mixed Integer Programming) yapısındadır. Modelin GAMS ortamındaki küme, parametre, skaler, değişken ve denklem tanımlamaları aşağıda sunulmuştur:



```
Title Uluslararası Kentiçi Dinamik Yakıt Zamanlı CVRPTW Modeli

* ---------------------------------------------------
* 1. KUMELER VE INDISLER (SETS)
* ---------------------------------------------------
Sets
    i   Tum Dugumler Antalya Cikis ve Donus / 1*9 /
    k   Arac Filosu 5 Adet Tir  / v1*v5 /
    c(i) Musteriler Avrupa Noktaları / 2*8 / ;

Alias (i,j);

* ---------------------------------------------------
* 2. PARAMETRELER VE SKALERLER
* ---------------------------------------------------
Scalars
    Q               Araclarin maksimum tonaj kapasitesi / 22 /
    BigM            Alt tur ve yakit akis kisitlari icin buyuk sayi / 100000 /
    hiz             Araclarin ortalama seyir hizi km-saat / 75 /
    TankKap         Tirin maksimum yakit depo kapasitesi Litre / 1200 /
    YakitKritik     Yolda kalmamak icin gereken emniyet tamponu Litre / 50 /
    yakit_tuketimi  1 kmde harcanan yakit litre / 0.335 /
    otoyol_ort      Avrupa rotasi ortalama km basi otoyol Euro / 0.35 /
    sofor_sefer     Sefere cikan soforun sabit ucreti Euro / 1200 /
    bakim_tl        Sefer basi ortalama bakim maliyeti TL / 22500 /
    euro_kuru       Guncel Euro Kuru / 53 /
    tr_cikis        Turkiye cikisi sabit ucret Euro / 90 / ;

Parameters
    sabit_arac      Sefere cikan her arac icin sabit maliyet Euro ;

* Sabit ucret hesabi
sabit_arac = sofor_sefer + (bakim_tl / euro_kuru) + tr_cikis ;

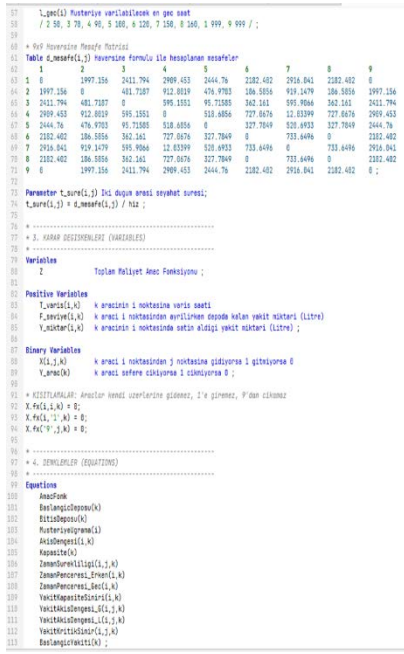
* --- YAKIT FIYATI OPTIMIZASYONU MATRISI ---
Parameter p_yakit(i) Dugum bazli degisen gercek yakit fiyatlari Euro-Litre
    / 1 1.20,
      2 1.01,
      3 1.72,
      4 2.13,
      5 2.13,
      6 2.18,
      7 2.19,
      8 2.24,
      9 1.20 / ;

Parameters
    q_talep(i) Musterilerin CMR Belgelerinden Alinan Gercek Tonaj Talepleri
    / 2 8.565, 3 7.513, 4 16.692, 5 9.888, 6 6.574, 7 20.566, 8 21.780, 1 0, 9 0 /

    s_hizmet(i) Musterideki bosaltma bekleme suresi Saat
    / 2 2, 3 2.5, 4 3, 5 1.5, 6 2, 7 3, 8 2, 1 0, 9 0 /

    e_erken(i) Musteriye varilabilecek en erken saat
    / 2 10, 3 20, 4 40, 5 50, 6 60, 7 70, 8 80, 1 0, 9 0 /
```

Şekil 3. GAMS Modeli Küme, Skaler ve Parametre Tanımlamaları



```
    l_gec(i) Musteriye varilabilecek en gec saat
    / 2 50, 3 70, 4 90, 5 100, 6 120, 7 150, 8 160, 1 999, 9 999 / ;

* 9x9 Haversine Mesafe Matrisi
Table d_mesafe(i,j) Haversine formulu ile hesaplanan mesafeler
          1         2         3         4         5         6         7         8         9
1         0      1997.156  2411.794  2909.453  2444.76   2182.402  2916.041  2182.402  0
2      1997.156     0       481.7187  912.8019  476.9703  186.5856  919.1479  186.5856  1997.156
3      2411.794  481.7187     0       595.1551  95.71585  362.161   595.9066  362.161   2411.794
4      2909.453  912.8019  595.1551     0      518.6856  727.0676  12.83399  727.0676  2909.453
5      2444.76   476.9703  95.71585  518.6856     0      327.7849  520.6933  327.7849  2444.76
6      2182.402  186.5856  362.161   727.0676  327.7849     0      733.6496  0         2182.402
7      2916.041  919.1479  595.9066  12.83399  520.6933  733.6496     0      733.6496  2916.041
8      2182.402  186.5856  362.161   727.0676  327.7849     0      733.6496  0         2182.402
9      0      1997.156  2411.794  2909.453  2444.76   2182.402  2916.041  2182.402  0 ;

Parameter t_sure(i,j) Iki dugum arasi seyahat suresi;
t_sure(i,j) = d_mesafe(i,j) / hiz ;

* ---------------------------------------------------
* 3. KARAR DEGISKENLERI (VARIABLES)
* ---------------------------------------------------
Variables
    Z               Toplam Maliyet Amac Fonksiyonu ;

Positive Variables
    T_varis(i,k)    k aracinin i noktesina varis saati
    F_servis(i,k)   k araci i noktasinda ayrilirken depoda kalan yakit miktari (Litre)
    Y_miktar(i,k)   k aracinin i noktesinda satin aldigi yakit miktari (litre) ;

Binary Variables
    X(i,j,k)        k araci i noktasından j noktasına gidiyorsa 1 gitmiyorsa 0
    Y_arac(k)       k araci sefere cikiyorsa 1 cikmiyorsa 0 ;

* KISITLAMALAR: Araclar kendi uzerlerine gidemez, 1'e giremez, 9'dan cikamaz
X.fx(i,i,k) = 0;
X.fx(i,'1',k) = 0;
X.fx('9',j,k) = 0;

* ---------------------------------------------------
* 4. DENKLEMLER (EQUATIONS)
* ---------------------------------------------------
Equations
    AmacFonk
    BaslangicDeposu(k)
    BitisDeposu(k)
    MusteriyeUgrama(i)
    AkisDengesi(i,k)
    Kapasite(k)
    ZamanSurekliligi(i,j,k)
    ZamanPenceresi_Erken(i,k)
    ZamanPenceresi_Gec(i,k)
    YakitKapasiteSiniri(i,k)
    YakitAkisDengesi_Ic(i,j,k)
    YakitAkisDengesi_Ic(i,j,k)
    YakitKritikSinir(i,j,k)
    BaslangicYakiti(k) ;
```

Şekil 4. GAMS Modeli Mesafe Matrisi, Değişkenler ve Kısıt Denklemleri Bildirimi

Şekil 3'te; lojistik ağ düğümleri (1-9), 5 araçlık filo kümesi, araç taşıma kapasitesi (22 ton), ortalama hız, yakıt tüketim katsayısı, müşteri talep miktarları ile zaman pencerelerini içeren temel skaler ve parametre blokları yer almaktadır.

Şekil 4'te; Haversine formülü ile hesaplanan 9×9 boyutlu mesafe matrisi, seyahat süresi fonksiyonu, sürekli ve ikili (binary) karar değişkenleri ile modelde çözülecek denklemlerin bildirimleri sunulmaktadır.

```
AmacFonk..
    Z =e= sum(k, Y_arac(k) * sabit_arac)
        + sum(k, sum(i, sum(j, d_mesafe(i,j) * otoyol_ort * X(i,j,k))))
        + sum(k, sum(i, Y_miktar(i,k) * p_yakit(i))) ;

BaslangicDeposu(k)..
    sum(j$c(j), X('1',j,k)) =e= Y_arac(k) ;

BitisDeposu(k)..
    sum(i$c(i), X(i,'9',k)) =e= Y_arac(k) ;

MusteriyeUgrama(c)..
    sum(k, sum(i, X(i,c,k))) =e= 1 ;

AkisDengesi(c,k)..
    sum(i, X(i,c,k)) - sum(j, X(c,j,k)) =e= 0 ;

Kapasite(k)..
    sum(i$c(i), q_talep(i) * sum(j, X(i,j,k))) =l= Q ;

ZamanSurekliligi(i,j,k)$(ord(i)<>ord(j))..
    T_varis(j,k) =g= T_varis(i,k) + t_sure(i,j) + s_hizmet(i) - BigM*(1 - X(i,j,k)) ;

ZamanPenceresi_Erken(i,k)..
    T_varis(i,k) =g= e_erken(i) * sum(j, X(i,j,k)) ;

ZamanPenceresi_Gec(i,k)..
    T_varis(i,k) =l= l_gec(i) + BigM*(1 - sum(j, X(i,j,k))) ;

* --- REFUELING (YAKIT IKMAL) KISITLARI ---

* Herhangi bir istasyondan ayrilirken yakit miktari tirin depo kapasitesini (1200 L) asamaz
YakitKapasiteSiniri(i,k)..
    F_seviye(i,k) =l= TankCap * Y_arac(k) ;

* Rotalar arasi yakit degisim dengesi (i'den j'ye gidis durumu)
YakitAkisDengesi_G(i,j,k)$(ord(i)<>ord(j))..
    F_seviye(j,k) =g= F_seviye(i,k) - (yakit_tuketimi * d_mesafe(i,j)) + Y_miktar(j,k) - BigM*(1 - X(i,j,k)) ;

YakitAkisDengesi_L(i,j,k)$(ord(i)<>ord(j))..
    F_seviye(j,k) =l= F_seviye(i,k) - (yakit_tuketimi * d_mesafe(i,j)) + Y_miktar(j,k) + BigM*(1 - X(i,j,k)) ;

* Tirin yolda yakitinin tamamen bitmesini engelleyen kritik emniyet kisiti (En az 50 Litre kalmali)
YakitKritikSinir(i,j,k)$(ord(i)<>ord(j))..
    F_seviye(i,k) - (yakit_tuketimi * d_mesafe(i,j)) =g= YakitKritik - BigM*(1 - X(i,j,k)) ;

* Araclarin sefere ilk baslangic noktasindaki (Antalya) depoyu doldurma durumu
BaslangicYakiti(k)..
    F_seviye('1',k) =e= Y_miktar('1',k) ;

* ---------------------------------------------------
* 5. MODELIN COZULMESI (SOLVE)
* ---------------------------------------------------
Model CVRPTW_Refuel /all/ ;
Solve CVRPTW_Refuel using mip minimizing Z ;

Display X.l, T_varis.l, Y_arac.l, F_seviye.l, Y_miktar.l, Z.l, sabit_arac ;
```

Şekil 5. GAMS Modeli Amaç Fonksiyonu, Akış, Zaman ve Yakıt İkmal Kısıtları

Şekil 5'te; toplam operasyonel maliyeti minimize eden amaç fonksiyonu denklemi, müşteri ziyaret ve rota akış dengesi, araç kapasitesi, Big-M tabanlı zaman sürekliliği ile depo kapasitesi ve kritik emniyet sınırını (50 L) güvenceye alan yakıt ikmal (arbitraj) kısıtları yer almaktadır. Modelin son satırında ise problemin CPLEX çözücüsü ile MIP formatında en küçükleme (minimum Z) yöntemiyle çözüleceği belirtilmektedir.

Kurulan model genelinde; $X(i,j,k)$ ikili karar değişkeni araçların rotalama akışını, $T_varis(i,k)$ varış zamanlarını, $F_seviye(i,k)$ düğüm çıkışlarındaki depo seviyelerini ve

$Y_miktar(i,k)$ ise istasyon bazlı satın alınan yakıt miktarlarını belirlemektedir.

6.4. Model Çözüm Çıktıları ve Karar Değişkenlerinin Analizi

Model, GAMS/CPLEX çözücüsü ile 0,328 saniyede global optimal sonuca ulaşmıştır. Çözüm sonucunda karar değişkenlerinin aldığı değerler çözücü pencere çıktıları halinde aşağıda gösterilmiştir:

```
----    172 VARIABLE X.L  k araci i noktasindan j noktasina gidiyorsa 1 gitmiyorsa 0

            v1          v2          v3          v4          v5

1.2                                 1.000
1.3                     1.000
1.4         1.000
1.7                                             1.000
1.8                                                         1.000
2.6                                 1.000
3.5                     1.000
4.9         1.000
5.9                     1.000
6.9                                 1.000
7.9                                             1.000
8.9                                                         1.000
```

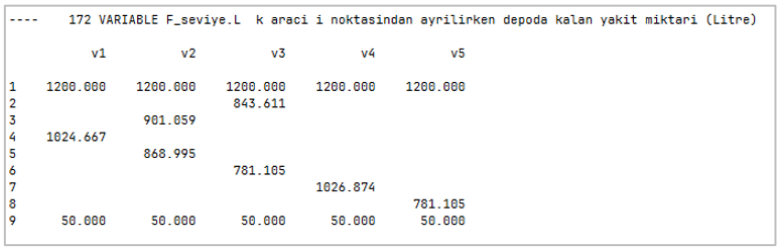
Şekil 6. GAMS Çözüm Çıktısı: Rota Atama Matrisi (X_{ijk})

Şekil 6'da verilen atama matrisine göre, 5 araçlık filonun tamamı sefere atanmış ve aşağıdaki bağımsız rotalar oluşmuştur:

- **Araç 1 (v_1):** Antalya (1) → Dortmund (4) → Antalya (9)
- **Araç 2 (v_2):** Antalya (1) → Stuttgart-1 (3) → Bielefeld (5) → Antalya (9)
- **Araç 3 (v_3):** Antalya (1) → Münih (2) → Stuttgart-2 (6) → Antalya (9)
- **Araç 4 (v_4):** Antalya (1) → Londra-Leyton (7) → Antalya (9)
- **Araç 5 (v_5):** Antalya (1) → Londra-Enfield (8) → Antalya (9)

Oluşturulan bu optimum rotalar boyunca araçların yakıt tüketimleri ve istasyon bazlı dolum stratejileri model tarafından eş zamanlı olarak çözümlenmiştir. Bu doğrultuda araçların uğradıkları düğümlerden ayrılırken depolarında kalan yakıt miktarları (F_seviye) Şekil 7'de, güzergah üzerindeki

noktalardan satın aldıkları optimal yakıt miktarları (Y_miktar) ise Şekil 8'de sunulmuştur.

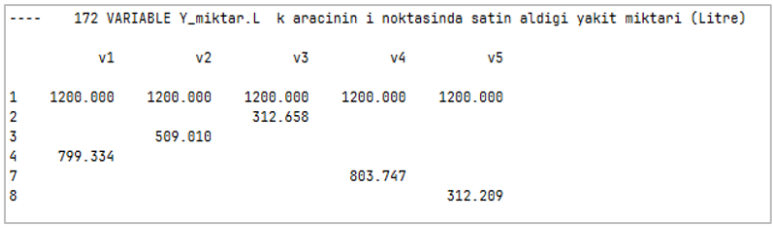


```
----    172 VARIABLE F_seviye.L  k araci i noktasindan ayrilirken depoda kalan yakit miktari (Litre)

          v1          v2          v3          v4          v5

1    1200.000    1200.000    1200.000    1200.000    1200.000
2                            843.611
3                901.059
4    1024.667
5                868.995
6                            781.105
7                                        1026.874
8                                                    781.105
9      50.000      50.000      50.000      50.000      50.000
```

Şekil 7. GAMS Çözüm Çıktısı: Düğümlerden Ayrılıştaki Depo Seviyeleri (F_seviye - Litre)



```
----    172 VARIABLE Y_miktar.L  k aracinin i noktasinda satin aldigi yakit miktari (Litre)

          v1          v2          v3          v4          v5

1    1200.000    1200.000    1200.000    1200.000    1200.000
2                            312.658
3                509.010
4      799.334
7                                        803.747
8                                                    312.209
```

Şekil 8. GAMS Çözüm Çıktısı: Noktalarda Satın Alınan Yakıt Miktarları (Y_miktar - Litre)

Şekil 7 ve Şekil 8 çıktıları birlikte incelendiğinde, modelin dinamik yakıt arbitrajı mantığı belirgin şekilde görülmektedir. Tüm araçlar sefere çıkarken yakıt birim fiyatının en düşük olduğu Türkiye çıkışında (1.20 €/L) depolarını tam (1200 Litre) doldurmuştur. Avrupa güzergahında ise araçlar yalnızca depoya dönüşü sağlayacak ve 50 Litre kritik emniyet sınırını koruyacak miktarlarda ikmal yapmıştır. Araçların tamamının Antalya dönüş deposuna (9 numaralı düğüm) tam 50.00 L seviyesinde girmesi, modelin yüksek fiyatlı ülkelerden gereksiz yakıt alımını engelleyerek maliyet minimizasyonu sağladığını kanıtlamaktadır.

6.5. Mevcut Durum ile Önerilen Modelin Karşılaştırılması

Firmanın geçmiş tecrübelere dayalı yürüttüğü mevcut manuel planlama sonuçları ile geliştirilen GAMS modelinin ürettiği optimal sonuçlar Tablo 5'te karşılaştırılmıştır.

Tablo 5. Mevcut Manuel Planlama ile Geliştirilen GAMS Modelinin Karşılaştırılması

Kriter	Mevcut Manuel Planlama	Geliştirilen GAMS Modeli	İyileştirme / Fark
Toplam Sefer Maliyeti	36.572,65 €	30.243,16 €	%17 Tasarruf
Rota Planlama Süresi	Saatler	0,328 sn	Süre Tasarrufu
Yakıt İkmal Stratejisi	Tahmin Odaklı / Sezgisel	Arbitraj Odaklı / Matematiksel	Maliyet Minimizasyonu
Kısıt Yönetimi	Gözden Kaçabilir	%100 Uyum (Kısıt Garantili)	Hata Payı Sıfır

Tablo 5'te, işletmenin deneyime dayalı yürüttüğü geleneksel planlama yaklaşımı ile önerilen matematiksel optimizasyon modelinin temel performans çıktıları karşılaştırılmaktadır. Elde edilen bulgular, geliştirilen modelin sefer başına maliyeti 36.572,65 € seviyesinden 30.243,16 € seviyesine indirerek %17 oranında (6.329,49 €) doğrudan tasarruf sağladığını göstermektedir. Bu maliyet düşüşü; modelin bölgesel akaryakıt fiyat farklarını değerlendirerek yakıt ikmalini en ucuz istasyonlarda optimize etmesinden ve boş kilometreleri asgariye indiren rota atamalarından kaynaklanmaktadır. Ayrıca saatler süren manuel çizelgeleme süresi 0,328 saniyeye çekilmiş ve katı teslimat pencerelerine %100 uyum garanti edilmiştir. Optimizasyon modeli ile ulaşılan sonuçlar mevcut durumdan daha optimum sonuçlar vermiştir.

6.6. Duyarlılık Analizi

Uluslararası taşımacılıkta akaryakıt fiyatlarının yüksek oynaklığa sahip olması nedeniyle, modelin fiyat dalgalanmalarına karşı dayanıklılığını ölçmek amacıyla duyarlılık analizi

uygulanmıştır. Avrupa yakıt fiyatlarının %10, %20 ve %30 arttığı 3 farklı senaryo GAMS ortamında test edilmiştir.

Tablo 6. Yakıt Fiyat Artış Senaryoları Duyarlılık Analizi Sonuçları

Yakıt Fiyat Artış Oranı	Toplam Sefer Maliyeti (€)	Maliyet Artış Oranı (%)
Baz Senaryo (%0)	**30.243 €**	**-**
+ (%10) Artış	**33.282 €**	**%10,04**
+ (%20) Artış	**36.321 €**	**%20,09**
+ (%30) Artış	**39.360 €**	**%30,14**

Tablo 6'da uluslararası taşımacılık operasyonlarında en yüksek belirsizlik kaynağı olan akaryakıt fiyatlarındaki olası artışların toplam sefer maliyetine etkisi üç farklı stres senaryosu altında incelenmektedir. Analiz sonuçlarına göre; yakıt fiyatlarındaki %10, %20 ve %30'luk artışlar toplam sefer bütçesine sırasıyla %10,04, %20,09 ve %30,14 oranında yansımaktadır. Maliyet artışının yakıt fiyat artışıyla neredeyse birebir doğrusal bir korelasyon sergilemesi, modelin piyasa fiyat şokları karşısında optimal rota yapısından sapmadığını, çözüm kararlılığını koruduğunu ve işletme yönetimine öngörülebilir bir risk kontrol mekanizması sunduğunu doğrulamaktadır.

7. SONUÇ VE DEĞERLENDİRME

Çalışma sonucunda; uluslararası taşımacılıktaki karmaşık zaman ve maliyet kısıtlarının, GAMS tabanlı matematiksel modelleme ile optimize edilerek operasyonel süreçlerde %17 oranında verimlilik artışı sağlanabileceği kanıtlanmıştır. Geliştirilen "Zaman Pencereli Araç Rotalama ve Yakıt İkmal Optimizasyonu" (CVRPTW-F) modeli, GAMS/CPLEX çözücüsü ile global optimum değerlere ulaştırılmıştır. 5 araçlık filo ile yapılan modelleme sonucunda; zaman pencereleri, kapasite ve sürücü dinlenme kısıtları %100 uyumla yönetilmiştir. Yakıt arbitrajı kısıtlarının modele dahil edilmesiyle, toplam sefer maliyeti 30.243 € olarak optimize edilmiştir. Şirketin geleneksel (sezgisel) planlama yöntemi ile modelin ürettiği çözüm

kıyaslandığında, yakıt ikmal stratejisi ve rota optimizasyonu sayesinde %17 oranında somut bir maliyet tasarrufu elde edilmiştir. Gerçekleştirilen duyarlılık analizi ile yakıt fiyatlarındaki %30'a varan dalgalanmaların, operasyonel maliyet üzerindeki etkisi doğrusal bir şekilde modellenmiş ve sistemin yüksek dayanıklılık gösterdiği kanıtlanmıştır. Model, fiyat artışlarına rağmen global optimum çözümden sapmadan, rotayı en ekonomik şekilde yeniden optimize edebilmekte ve yakıt maliyetindeki artışın toplam sefer maliyeti üzerindeki yansıması, geliştirilen model ile şeffaf bir şekilde izlenebilir hale getirilmektedir.

Modelin geliştirilmesine yönelik öneriler kapsamında; operasyonel esnekliği artırmak adına, yakıt ve otoyol ücretlerinin güncel API verileriyle (Tolls.eu) anlık beslenmesi sağlanarak sistem statik modelden canlı karar destek mekanizmasına dönüştürülebilir. Modelde sabit kabul edilen gümrük ve feribot bekleme süreleri, gerçek saha verileriyle istatistiksel olarak dağılımlara (olasılıksal kısıtlar) dönüştürülerek modelin gerçek dünya başarımı artırılabilir. Gelecek çalışmalarda, karbon ayak izi kısıtları ve alternatif yakıtlı araç (EV) senaryoları entegre edilerek modelin "Yeşil Lojistik" vizyonuyla ölçeklendirilmesi hedeflenmektedir.

İşletmenin deneyime dayalı yürüttüğü geleneksel planlama yaklaşımı ile önerilen matematiksel optimizasyon modelinin temel performans çıktıları karşılaştırılmaktadır. Elde edilen bulgular, geliştirilen modelin sefer başına maliyeti 36.572,65 € seviyesinden 30.243,16 € seviyesine indirerek %17 oranında (6.329,49 €) doğrudan tasarruf sağladığını göstermektedir. Bu maliyet düşüşü; modelin bölgesel akaryakıt fiyat farklarını değerlendirerek yakıt ikmalini en ucuz istasyonlarda optimize etmesinden ve boş kilometreleri asgariye indiren rota atamalarından kaynaklanmaktadır. Ayrıca saatler süren manuel çizelgeleme süresi 0,328 saniyeye çekilmiş ve katı

teslimat pencerelerine %100 uyum garanti edilmiştir. Bu doğrultuda çalışma, firmanın mevcut planlama yaklaşımının daha sistematik ve sayısal bir yöntemle değerlendirilmesine katkı sağlamayı amaçlamıştır. Elde edilen sonuçlar, firmanın taşıma operasyonlarını optimize etmiş, daha iyi sonuçlara ulaşılmış ve firmanın gelecekteki taşıma planlamalarında karar destek unsuru olarak kullanılabileceğini göstermiştir.

Çalışmanın farklı araç rotalama problemleri ve kısıtlar kullanılarak geliştirebileceği ve lojistik maliyetlerin minimize edilmesine yönelik matematiksel modelleme literatürüne katkı sunması beklenmektedir.

KAYNAKLAR

Al-Nabulsi, D., Kwigizile, V., Kim, H., & Jin, H.-M. (2025). Analyzing fleet efficiency and passenger delay in demand-responsive transit: A dual-model approach with CVRPTW and TAMOS. *Journal of Advanced Transportation, 2025*. doi:10.1155/atr/6761411

Atasagun, G. C., & Karaoğlan, İ. (2019). A mathematical model for the time dependent vehicle routing problem with simultaneous pick-up and delivery. *Journal of the Faculty of Engineering and Architecture of Gazi University, 34*(3), 1743–1758.

Boz, E., Çalık, A., & Şahin, Y. (2023). Yeşil zaman pencereli ve eş zamanlı topla dağıt araç rotalama problemlerinin metasezgisel yöntemlerle çözümü. *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi, 39*(2), 757-770.

Cavecchia, M., Alves De Queiroz, T., Lancellotti, R., Zucchi, G., & Iori, M. (2025). A real-world multi-depot, multi-period, and multi-trip vehicle routing problem with time windows. *Proceedings of the 14th International Conference on Operations Research and Enterprise Systems (ICORES 2025), 1*, 112-122.

Chen, G., Gao, J., & Chen, D. (2025). Research on vehicle routing problem with time windows based on improved genetic algorithm and ant colony algorithm. *Electronics, 14*(4), 647.

Çetintaş, F., & Aslan, E. (2026). Bir kargo firmasında zaman ve kapasite kısıtlı araç rotalama problemi. *Yönetim ve Ekonomi Dergisi, 33*(2), 377-399.

Ercan Cömert, S., Gül, S., Yazgan, H. R., & Kır, S. (2015). Zaman pencereli araç rotalama problemi için iki aşamalı bir çözüm yöntemi önerisi. *Yöneylem Araştırması ve Endüstri*

Mühendisliği (YAEM) 35. Ulusal Kongresi, Orta Doğu Teknik Üniversitesi, Ankara.

Göçken, T., Yaktubay, M., & Kılıç, F. (2018). Zaman pencereli araç rotalama problemi çözümü için çok amaçlı genetik algoritma yaklaşımı. *Gazi Üniversitesi Fen Bilimleri Dergisi Part C: Tasarım ve Teknoloji, 6*(4), 774-786.

Karagül, K., Tokat, S., & Aydemir, E. (2016). Kapasite kısıtlı araç rotalama problemlerinde başlangıç rotalarının kurulması için yeni bir algoritma. *Mühendislik Bilimleri ve Tasarım Dergisi, 4*(3), 215-226.

Kay, M. G., Karagül, K., Şahin, Y., & Gündüz, G. (2021). Minimizing total logistics cost for long-haul multi-stop truck transportation. *Transportation Research Record, 2676*(2), 367–378.

Lutfizein, F., & Nuha, H. (2026). Penentuan rute distribusi produk tekstil dengan pendekatan capacitated vehicle routing problem with time windows (CVRPTW). *Jurnal Teknik Industri Terintegrasi, 9*, 3185-3195. doi:10.31004/jutin.v9i3.60689

Mor, A., Archetti, C., Jabali, O., Simonetto, A., & Speranza, M. G. (2022). The bi-objective long-haul transportation problem on a road network. *Omega, 106,* 102522.

Şahin Macit, N. (2024). Zaman pencereli araç rotalama problemi (ZPARP): Bibliyometrik bir bakış (2000-2022).

T.C. Ticaret Bakanlığı. (2026). *Dış ticaret lojistiği 2026*. Uluslararası Hizmet Ticareti Genel Müdürlüğü, Ankara.

Tutumlu, B., Ertan, E. D., İnci, G., & Saraç, T. (2025). Uluslararası taşıma yapan bir lojistik firması için araç rotalarının ve yakıt ikmal noktalarının belirlenmesi. *Uludağ Üniversitesi Mühendislik Fakültesi Dergisi, 30*(3), 977–996. doi:10.17482/uumfd.1647330

Tolls.eu. (2026). *European motorway tolls and vignettes*. https://www.tolls.eu/

EKLER

```
$title Ulaslararasi Navlugr Duyarlilik Analizi Modeli

Sets
    i / 1*9 /
    k / v1*v5 /
    c(i) / 1*8 /
    s / Baz, Artis10, Artis20, Artis30 / ;

Alias (i,j);

Scalars
    Q / 22 /, BigM / 100000 /, hiz / 75 /, TankCap / 1200 /, YakitKritik / 50 /
    yakit_tuketimi / 0.335 /, otoyol_ort / 0.35 /, sofor_sefer / 1200 /
    bakim_tl / 22000 /, euro_kuru / 53 /, tn_cikis / 90 / ;

Parameter sabit_arac;
sabit_arac = sofor_sefer + (bakim_tl / euro_kuru) + tn_cikis ;

Parameter p_yakit_baz(i) / 1 1.20, 2 1.80, 3 1.72, 4 2.13, 5 2.13, 6 2.13, 7 2.19, 8 2.24, 9 1.20 /;

Parameter p_yakit(i), q_talep(c), s_hizmet(i), s_erken(i), t_gec(c), d_mesafe(i,j);
q_talep('2')=0.565; q_talep('3')=7.513; q_talep('4')=16.692; q_talep('5')=9.300; q_talep('6')=4.374; q_talep('7')=26.566; q_talep('8')=21.780;
s_hizmet('2')=2; s_hizmet('3')=2.9; s_hizmet('4')=3; s_hizmet('5')=1.5; s_hizmet('6')=2; s_hizmet('7')=3; s_hizmet('8')=2;
s_erken('2')=30; s_erken('3')=20; s_erken('4')=40; s_erken('5')=50; s_erken('6')=60; s_erken('7')=70; s_erken('8')=80;
t_gec('1')=999; t_gec('2')=59; t_gec('3')=70; t_gec('4')=90; t_gec('5')=100; t_gec('6')=125; t_gec('7')=150; t_gec('8')=160;

* Mesafe matrisi (Sadece ust yapi sifir olan tek tek tanimlama)
d_mesafe(i,j) = 0;
d_mesafe('1','2')=1997.156; d_mesafe('1','3')=2411.794; d_mesafe('1','4')=2909.453; d_mesafe('1','5')=2444.76; d_mesafe('1','6')=2182.482; d_mesafe('1','7')=2182.482;
d_mesafe('2','3')=481.7197; d_mesafe('2','4')=912.8319; d_mesafe('2','5')=476.9793; d_mesafe('2','6')=186.5356; d_mesafe('2','7')=919.1479; d_mesafe('2','8')=186.5356; d_mesafe('2','9')=1997.156;
d_mesafe('3','4')=595.1551; d_mesafe('3','5')=95.71565; d_mesafe('3','6')=362.161; d_mesafe('3','7')=595.9066; d_mesafe('3','8')=362.161; d_mesafe('3','9')=2411.794;
d_mesafe('4','5')=518.6856; d_mesafe('4','6')=727.0675; d_mesafe('4','7')=12.03399; d_mesafe('4','8')=727.0675; d_mesafe('4','9')=2909.453;
d_mesafe('5','6')=327.7849; d_mesafe('5','7')=520.6833; d_mesafe('5','8')=327.7849; d_mesafe('5','9')=2444.76;
d_mesafe('6','7')=733.6496; d_mesafe('6','8')=2182.482;
d_mesafe('7','8')=733.6496; d_mesafe('7','9')=2913.341;
d_mesafe('8','9')=2182.482;

Parameter t_sure(i,j); t_sure(i,j) = d_mesafe(i,j) / hiz ;

Variables Z, T_varis(i,k), F_seviye(i,k), Y_miktar(i,k);
Binary Variables X(i,j,k), Y_arac(k);

Equations AmacFonk, BaslangicDeposu(k), BitisDeposu(k), MusteriyeUgrama(i), AkisDengesi(i,k), Kapasite(k),
          ZamanSurekliligi(i,j,k), ZamanPenceresi_Erken(i,k), ZamanPenceresi_Gec(i,k), YakitKapasiteSiniri(i,k),
          YakitAkisDengesi_G(i,j,k), YakitAkisDengesi_L(i,j,k), YakitKritikSinir(i,j,k), BaslangicYakiti(k);

AmacFonk.. Z =e= sum(k, Y_arac(k)*sabit_arac) + sum(k,sum(i,sum(j, d_mesafe(i,j)*otoyol_ort*X(i,j,k)))) + sum(k,sum(i, Y_miktar(i,k)*p_yakit(i)));
BaslangicDeposu(k).. sum(j$c(j), X('1',j,k)) =e= Y_arac(k);
BitisDeposu(k).. sum(i$c(i), X(i,'9',k)) =e= Y_arac(k);
MusteriyeUgrama(c).. sum(k, sum(i, X(i,c,k))) =e= 1;
AkisDengesi(c,k).. sum(i, X(i,c,k)) - sum(j, X(c,j,k)) =e= 0;
Kapasite(k).. sum(i$c(i), q_talep(i) * sum(j, X(i,j,k))) =l= Q;
ZamanSurekliligi(i,j,k)$(ord(i)<>ord(j)).. T_varis(j,k) =g= T_varis(i,k) + t_sure(i,j) + s_hizmet(i) - BigM*(1 - X(i,j,k));
ZamanPenceresi_Erken(i,k).. T_varis(i,k) =g= e_erken(i) * sum(j, X(i,j,k));
ZamanPenceresi_Gec(i,k).. T_varis(i,k) =l= t_gec(i) + BigM*(1 - sum(j, X(i,j,k)));
YakitKapasiteSiniri(i,k).. F_seviye(i,k) =l= TankCap * Y_arac(k);
```

Ek-1 Duyarlılık analizi GAMS kodu

```
YakitAkisDengesi_G(i,j,k)$(ord(i)<>ord(j)).. F_seviye(j,k) =g= F_seviye(i,k) - (yakit_tuketimi*d_mesafe(i,j)) + Y_miktar(j,k) - BigM*(1-X(i,j,k));
YakitAkisDengesi_L(i,j,k)$(ord(i)<>ord(j)).. F_seviye(j,k) =l= F_seviye(i,k) - (yakit_tuketimi*d_mesafe(i,j)) + Y_miktar(j,k) + BigM*(1-X(i,j,k));
YakitKritikSinir(i,j,k)$(ord(i)<>ord(j)).. F_seviye(i,k) - (yakit_tuketimi*d_mesafe(i,j)) =g= YakitKritik - BigM*(2-X(i,j,k));
BaslangicYakiti(k).. F_seviye('1',k) =e= Y_miktar('1',k);
X.fx(i,i,k) = 0; X.fx('1','1',k) = 0; X.fx('9',j,k) = 0;

Model CVRPTW /all/ ;
Parameter Sonuc(s);

Loop(s,
    if(sameas(s,'Baz'), p_yakit(i) = p_yakit_baz(i));
    if(sameas(s,'Artis10'), p_yakit(i)$(not sameas(i,'1')) = p_yakit_baz(i) * 1.10; p_yakit('1') = 1.20);
    if(sameas(s,'Artis20'), p_yakit(i)$(not sameas(i,'1')) = p_yakit_baz(i) * 1.20; p_yakit('1') = 1.20);
    if(sameas(s,'Artis30'), p_yakit(i)$(not sameas(i,'1')) = p_yakit_baz(i) * 1.30; p_yakit('1') = 1.20);
    Solve CVRPTW using mip minimizing Z;
    Sonuc(s) = Z.l;
);
Display Sonuc ;
```

Ek-2 Duyarlılık analizi GAMS kodu devamı

METASEZGİSEL ALGORİTMALARIN ZAMAN SERİSİ HİPERPARAMETRE OPTİMİZASYONUNDAKİ İSTATİSTİKSEL GÜVENİLİRLİĞİ: ÇAPRAZ DOĞRULAMA STRATEJİLERİNİN KARŞILAŞTIRMALI ANALİZİ[1]

Burcu KARTAL[2]

1. GİRİŞ

Makine öğrenmesi modellerinin başarımı, eğitim sürecinden önce belirlenen ve veriden doğrudan öğrenilmeyen hiperparametrelerin değerlerine duyarlıdır. Bu nedenle hiperparametre optimizasyonu, model geliştirme sürecinin ayrılmaz bir adımı hâline gelmiştir. Gradyan bilgisine gereksinim duymadan doğrusal olmayan ve çok tepeli arama uzaylarını tarayabilen metasezgisel algoritmalar bu amaçla yaygın biçimde kullanılmaktadır (Holland, 1992; Kennedy ve Eberhart, 1995; Storn ve Price, 1997; Karaboga ve Basturk, 2007). Model tabanlı yaklaşımlar arasında ise Ağaç Yapılı Parzen Kestiricisi (TPE) gibi Bayesyen optimizasyon yöntemleri öne çıkmaktadır (Bergstra vd., 2011).

---

[1] Bu bölüm, 11–13 Mayıs 2026 tarihlerinde İstanbul'da düzenlenen VII. Uluslararası Uygulamalı İstatistik Kongresi'nde (UYIK-2026) sunulan "Zaman Serisi Verilerinde Metasezgisel Algoritmaların Hiperparametre Optimizasyonundaki İstatistiksel Güvenilirliği: Farklı Çapraz Doğrulama Stratejilerinin Karşılaştırmalı Analizi" başlıklı bildirinin genişletilmiş ve revize edilmiş hâlidir. Bu bölümde deney tasarımı genişletilmiş ve tüm analizler yeniden yapılmıştır; bu nedenle sayısal bulgular bildiriden farklılık göstermektedir.

[2] Dr. Öğr. Üyesi, Recep Tayyip Erdoğan Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, İşletme Bölümü; ORCID: 0000-0001-8340-0234.

Bu algoritmaların ortak özelliği stokastik olmalarıdır: aynı veri, aynı model ve aynı ayarlar altında farklı rastgele tohumlarla yapılan çalıştırmalar farklı hiperparametre değerleri ve farklı doğrulama hataları üretebilir. Aynı durum makine öğrenmesi süreçlerinin geneli için de geçerlidir. Liem ve Panichella (2020), yazılım mühendisliği alanındaki 26 veri seti ve 8 sınıflandırıcı üzerinde rastgele tohum seçiminin başarım sonuçlarında ihmal edilemeyecek düzeyde değişkenlik yarattığını göstermiştir. Aynı çalışmada, incelenen makalelerin yarısından azının rastgele veri örneklemeden kaynaklanan tehditleri ele aldığı, yalnızca %8'inin ise makine öğrenmesinin rastgele doğasını dikkate aldığı saptanmıştır. Picard (2021), CIFAR-10 ve ImageNet veri setlerinde 10^4'e varan sayıda tohumu taramıştır. Varyans çok büyük olmasa bile, ortalamadan belirgin biçimde iyi ya da kötü sonuç veren aykırı tohumların kolayca bulunabildiğini göstermiştir. Eryilmaz vd., (2024) ise CUDA tabanlı hesaplamalardaki denetlenmeyen rastgeleliğin başarım skorlarında %4,77'ye varan farklara yol açabildiğini bildirmiştir.

Bu bulgular, bir algoritmanın başarımını tek bir çalıştırmaya dayandırmanın yanıltıcı çıkarımlara zemin hazırladığını göstermektedir. Bouthillier vd. (2021), veri örnekleme, parametre başlatma ve hiperparametre seçiminden kaynaklanan varyansın algoritma karşılaştırmalarını doğrudan etkilediğini vurgulamıştır. Yazarlara göre tekil ölçümlerle başarım iyileşmelerini saptamada hata payı yüksektir. Bu nedenle farklı varyans kaynaklarının modellenmesine dayalı öneriler geliştirmişlerdir. Boquet vd. (2019), modellerin başarımları arasındaki farkları belirlemek için çok sayıda tekrara dayanan istatistiksel test yöntemleri önermiştir. Vos vd. (2025) ise rastgele tohumu değiştirerek 400'e varan tekrarla sonuçları kararlı hâle getiren bir doğrulama yaklaşımı geliştirmiştir. Bu çalışmalar, Demšar'ın (2006) algoritmaların istatistiksel karşılaştırılmasına ilişkin çerçevesiyle örtüşmektedir. Demšar, makine öğrenmesi

karşılaştırmalarında parametrik testlerin normallik ve varyans homojenliği varsayımlarının çoğunlukla sağlanmadığını belirterek Wilcoxon işaretli sıralar testi ve Friedman testi gibi parametrik olmayan testlerin kullanılmasını önermiştir. Böylece tekil sonuçlar yerine tekrarlı çalıştırmalardan elde edilen dağılımlara dayanan bir değerlendirme kültürü öne çıkmaktadır. Bu bölümde de algoritmalar, 30 bağımsız çalıştırmanın dağılımları üzerinden parametrik olmayan testlerle karşılaştırılmıştır.

Sorun zaman serisi verilerinde daha da karmaşık bir hâl alır. Gözlemlerin zamansal sırası, modelin yalnızca geçmiş bilgiyle eğitilip gelecekte sınanmasını gerektirir. Gözlemleri rastgele karıştıran standart k-katlı çapraz doğrulama, zamansal bağımlılık nedeniyle iyimser hata tahminleri üretebilir (Bergmeir ve Benítez, 2012). Bu nedenle kayan pencere (walk-forward), genişleyen pencere ve bloklu k-katlı çapraz doğrulama gibi zaman serisine özgü stratejiler önerilmiştir (Tashman, 2000; Bergmeir ve Benítez, 2012; Roberts vd., 2017). Hiperparametre optimizasyonunda çapraz doğrulama hatası amaç fonksiyonunun kendisi olduğundan, doğrulama stratejisinin seçimi optimizasyonun çözdüğü problemi de doğrudan değiştirir.

Metasezgisel algoritmalar, zaman serisi tahmininde makine öğrenmesi modellerinin hiperparametrelerini ayarlamak amacıyla enerji, finans ve emtia piyasaları gibi farklı alanlarda yaygın biçimde kullanılmaktadır. Wu, Chai, Zhang ve Zhao (2022), elektrik fiyatı tahmininde XGBoost modelinin sekiz temel parametresini PSO ile optimize etmiştir. Yihan, Liang ve Ruiqi (2023), kısa dönem elektrik yükü tahmininde ampirik mod ayrışımı (EMD) ile birleştirilen bir LSTM modelini GA ile optimize etmiştir. Zulfiqar ve Rasheed (2022) ve Zulfiqar vd. (2024) de kısa dönem elektrik yükü tahmininde LSTM ağının nöron sayısı, epoch sayısı ve yığın büyüklüğü gibi hiperparametrelerini Modifiye edilmiş PSO ve GA ile

belirlemiştir. Yusoff vd. (2023) ise ham petrol fiyatı tahmininde LSTM hiperparametrelerini PSO ile ayarlamıştır. Finansal serilerde Chou vd. (2020), çok çıktılı en küçük kareler destek vektör regresyonu (MLSSVR) modelinin hiperparametrelerini hızlandırılmış PSO ile ayarlayarak hisse senedi ve döviz fiyatları için aralık tahminleri üretmiştir. Zhao (2025) ise GA'yı, rastgele orman modelinden önce uygulanan varyasyonel mod ayrışımının (VMD) parametrelerini belirlemek için kullanmıştır.

Bu çalışmaların ortak noktası, önerilen hibrit modellerin (ör. PSO-XGBoost, GA-LSTM, GA-VMD-RF) geleneksel modellere göre daha düşük RMSE, MAE veya MAPE değerleri verdiğini raporlamalarıdır. Ancak bu çalışmalarda sonuçların birden çok bağımsız çalıştırmanın ortalama ve standart sapmasıyla sunulduğuna dair açık bir ifade bulunmamaktadır. Dolayısıyla raporlanan üstünlüğün tek bir çalıştırmaya mı yoksa tekrarlı deneylere mi dayandığı anlaşılamamaktadır. Meta sezgisel algoritmaların stokastik doğası dikkate alındığında bu durum, önemli bir yöntemsel boşluğa işaret etmektedir.

Bu bölüm, söz konusu iki boyutu, yani optimizasyon algoritmasının stokastikliğini ve çapraz doğrulama stratejisinin etkisini, birlikte ele almaktadır. Altı arama yöntemi (Genetik Algoritma, Diferansiyel Gelişim, Parçacık Sürü Optimizasyonu, Yapay Arı Kolonisi, TPE tabanlı Bayesyen Optimizasyon ve taban çizgisi olarak Rastgele Arama), üç zaman serisi çapraz doğrulama stratejisi ve yapısal karmaşıklığı farklı iki sentetik zaman serisi üzerinde, eşit değerlendirme bütçesiyle ve her koşulda 30 bağımsız çalıştırmayla karşılaştırılmıştır. Bölümün katkıları şöyle özetlenebilir:

- Güvenilirlik, yalnızca varyasyon katsayısıyla değil, bilinen global optimuma göre hesaplanan göreli pişmanlık, başarı oranı ve seçilen hiperparametrenin çalıştırmalar arası dağılımıyla birlikte ölçülmüştür.

- İstatistiksel anlamlılık ile pratik önem birbirinden ayrılmış; Friedman testine etki büyüklüğü (Kendall W) ve çoklu karşılaştırma düzeltmesi (Holm) eklenmiştir.
- Çapraz doğrulama stratejisinin hem seçilen hiperparametreyi hem de genelleme hatası tahminindeki yanlılığı nasıl değiştirdiği bağımsız bir test kümesiyle gösterilmiştir.
- Özellik mühendisliğinde veri sızıntısının bulgular üzerindeki etkisi tartışılmış ve bu riske karşı basit doğrulama kontrolleri önerilmiştir.

Bölümün geri kalanı şöyle düzenlenmiştir: İkinci kısımda kavramsal çerçeve ve kullanılan algoritmalar, üçüncü kısımda veri setleri ve deney tasarımı, dördüncü kısımda bulgular, beşinci kısımda tartışma ve altıncı kısımda sonuçlar sunulmaktadır.

2. KAVRAMSAL ÇERÇEVE

2.1. Hiperparametre Optimizasyonu Problemi ve Ridge Regresyon

Hiperparametre optimizasyonu, λ hiperparametre vektörü ve D veri seti için doğrulama hatasını en küçükleyen $\lambda^* = \text{argmin}_{\lambda \in \Lambda} L_{ÇD}(\lambda; D)$ değerinin aranmasıdır. Burada $L_{ÇD}$, seçilen çapraz doğrulama stratejisiyle hesaplanan ortalama hatadır. $L_{ÇD}$ fonksiyonunun analitik biçimi bilinmediğinden ve her değerlendirmesi modelin yeniden eğitilmesini gerektirdiğinden, problem "kara kutu" bir optimizasyon problemi olarak ele alınır.

Bu çalışmada model olarak Ridge regresyon kullanılmıştır. Ridge regresyon, doğrusal regresyonun katsayılarını L2 cezasıyla küçülten bir düzenlileştirme yöntemidir (Hoerl ve Kennard, 1970):

$$\hat{\beta}(\alpha) = \text{argmin}\ \Sigma_t\ (y_t - \beta_0 - x_t'\beta)^2 + \alpha\|\beta\|^2_2 \quad (1)$$

Ceza parametresi $\alpha > 0$ tek hiperparametredir. α büyüdükçe katsayılar sıfıra doğru büzülür; $\alpha \to 0$ durumunda çözüm sıradan en küçük kareler kestiricisine yaklaşır. Arama doğal logaritma ölçeğinde, $\ln \alpha \in [-4, 6]$ (yani $\alpha \in [0{,}018;\ 403{,}4]$) aralığında yapılmıştır. Tek boyutlu ve düzgün bir amaç fonksiyonu seçilmesi bilinçli bir tasarım kararıdır. Böylece global optimum yoğun bir ızgara ile kesin olarak belirlenebilmekte ve her algoritmanın bu optimuma ne kadar yaklaştığı doğrudan ölçülebilmektedir.

2.2. Arama Algoritmaları

2.2.1. Genetik Algoritma (GA)

GA, doğal seçilim ve kalıtımdan esinlenen, seçilim, çaprazlama ve mutasyon işlemcilerini bir aday çözüm popülasyonu üzerinde yineleyen evrimsel bir algoritmadır (Holland, 1992). Bu çalışmada sürekli arama uzayına uygun, sayı kodlu bir GA uygulanmıştır: ebeveynler ikili turnuva ile seçilmiş, çaprazlamada BLX-0,5 işlemcisi, mutasyonda ise Gauss gürültüsü kullanılmış ve ebeveyn ile yavrular birleştirilerek en iyi bireyler bir sonraki kuşağa aktarılmıştır (elitist $(\mu+\lambda)$ seçilimi). BLX-α çaprazlama işlemcisi, iki ebeveynin tanımladığı aralığın her iki yönde α oranında genişletilmiş hâlinden tekdüze olarak yavru üretir (Eshelman ve Schaffer, 1993).

2.2.2. Diferansiyel Gelişim (DE)

DE, popülasyondaki bireyler arasındaki farkları ölçekleyerek yeni aday çözümler üreten ve sürekli uzaylarda etkin olan bir evrimsel algoritmadır (Storn ve Price, 1997). Uygulamada SciPy kütüphanesinin *differential_evolution* fonksiyonu kullanılmıştır.

2.2.3. Parçacık Sürü Optimizasyonu (PSO)

PSO'da her parçacığın hızı, kendi en iyi konumuna (bilişsel bileşen, c_1), sürünün en iyi konumuna (sosyal bileşen, c_2)

ve eylemsizlik ağırlığına (w) göre güncellenir (Kennedy ve Eberhart, 1995). Çalışmada küresel en iyi (global-best) topolojisi kullanılmış ve PySwarms kütüphanesinden yararlanılmıştır (Miranda, 2018).

2.2.4. Yapay Arı Kolonisi (ABC)

ABC, bal arısı kolonilerinin yiyecek arama davranışını modeller. İşçi arılar kendi besin kaynaklarının komşuluğunu arar, gözcü arılar kaynakları uygunluklarıyla orantılı olasılıkla seçer, belirli bir deneme sınırını (*limit*) aşarak gelişmeyen kaynaklar ise kâşif arılarca terk edilip rastgele yeni kaynaklarla değiştirilir (Karaboga ve Basturk, 2007). Kanonik biçimde uygunluk 1/(1+f) ile hesaplanır, her çevrimde en fazla bir kaynak terk edilir ve o ana kadar bulunan en iyi çözüm ayrıca bellekte tutulur.

2.2.5. Bayesyen Optimizasyon (BO-TPE)

Bayesyen optimizasyon, amaç fonksiyonu için bir vekil (surrogate) model kurar ve bir sonraki değerlendirme noktasını bu modele dayanan bir kazanım ölçütüyle seçer. TPE yaklaşımı, iyi ve kötü gözlemlerin yoğunluklarını ayrı ayrı modelleyerek beklenen iyileşmeyi en büyükleyen noktayı önerir (Bergstra vd., 2011). Uygulamada Optuna kütüphanesinin TPE örnekleyicisi kullanılmıştır (Akiba vd., 2019).

2.2.6. Rastgele Arama

Arama uzayından tekdüze dağılımla bağımsız noktalar çeken rastgele arama, hiçbir öğrenme mekanizması içermediğinden diğer algoritmaların sağladığı katma değeri ölçmek için taban çizgisi olarak kullanılmıştır. Rastgele aramanın hiperparametre optimizasyonunda ızgara aramasına göre daha verimli olabildiği gösterilmiştir (Bergstra ve Bengio, 2012).

2.3. Zaman Serilerinde Çapraz Doğrulama Stratejileri

Eğitim kümesinin uzunluğu n olmak üzere üç strateji kullanılmış ve hepsinde beş kat oluşturulmuştur:

- **Kayan pencere (walk-forward):** Sabit genişlikte bir eğitim penceresi zaman ekseninde ileri kaydırılır; her katta model yalnızca hemen önceki gözlemlerle eğitilip izleyen blokta sınanır (Tashman, 2000). Bu çalışmada $h = \lfloor n/7 \rfloor$ uzunluğunda test blokları ve $2h$ uzunluğunda eğitim pencereleri kullanılmıştır.
- **Genişleyen pencere:** Eğitim kümesi her katta serinin başından itibaren büyür; test bloğu uzunluğu $h = \lfloor n/6 \rfloor$'dır (Hyndman ve Athanasopoulos, 2018).
- **Bloklu k-katlı:** Seri ardışık beş bloğa bölünür; her blok sırayla test kümesi, geri kalan bloklar eğitim kümesi olur. Bu yaklaşım, geçmişin yanı sıra gelecekteki gözlemleri de eğitimde kullandığı hâlde, belirli koşullar altında zaman serilerinde güvenilir sonuçlar verebilmektedir (Bergmeir ve Benítez, 2012).

Kayan pencere ile genişleyen pencere arasındaki temel fark, eğitim penceresinin genişliğidir. Bu çalışmada kayan pencere sabit genişlikte tanımlanarak üç stratejinin birbirinden açıkça ayrışması sağlanmıştır.

2.4. Stokastik Optimizasyonda Güvenilirlik Ölçütleri

Bir algoritmanın $r = 1, \ldots, R$ bağımsız çalıştırmada bulduğu en iyi ÇD-RMSE değerleri f_r, yoğun ızgara ile belirlenen global optimum f^* olmak üzere aşağıdaki ölçütler kullanılmıştır. Varyasyon katsayısı (VK), sonuçların göreli dağılımını ölçer:

$$VK\% = 100 \cdot s_f / \bar{f} \quad (2)$$

Ancak VK tek başına yanıltıcı olabilir: bütün çalıştırmalar aynı yanlış noktaya yakınsayan bir algoritma da VK = 0 verir. Bu nedenle global optimuma uzaklığı ölçen göreli pişmanlık (regret) ve başarı oranı da hesaplanmıştır:

$$\text{Pişmanlık}_r\ (\%) = 100 \cdot (f_r - f^*) / f^* \quad (3)$$

Pişmanlığı %0,1'in altında kalan çalıştırmaların oranı "başarı oranı" olarak tanımlanmıştır. Parametre düzeyindeki kararlılık, seçilen ln α değerlerinin çalıştırmalar arası standart sapmasıyla ($SS_{\ln \alpha}$) ölçülmüştür. Son olarak çapraz doğrulamanın genelleme hatasını ne ölçüde doğru tahmin ettiği, bağımsız test kümesi hatası ile ÇD hatası arasındaki göreli farkla değerlendirilmiştir:

$$\text{Yanlılık}\ (\%) = 100 \cdot (RMSE_{test} - RMSE_{ÇD}) / RMSE_{ÇD} \quad (4)$$

Negatif yanlılık, çapraz doğrulamanın test hatasını olduğundan yüksek (kötümser), pozitif yanlılık ise olduğundan düşük (iyimser) tahmin ettiğini gösterir. Bu ölçüt modelin aşırı öğrenmesini değil, doğrulama tahmininin yanlılığını gösterdiği için "yanlılık" olarak adlandırılmıştır.

3. YÖNTEM

3.1. Veri Setleri

Çalışmada gerçek bir veri seti kullanılmamış; yapısal özellikleri tam olarak bilinen ve kontrol edilebilen iki sentetik zaman serisi, sabit rastgele tohumla (tohum = 42) n = 600 gözlem uzunluğunda üretilmiştir. **Kolay-GBM** serisi, finansal fiyat süreçlerinin basitleştirilmiş bir modeli olan Geometrik Brown Hareketi ile üretilmiştir (Black ve Scholes, 1973):

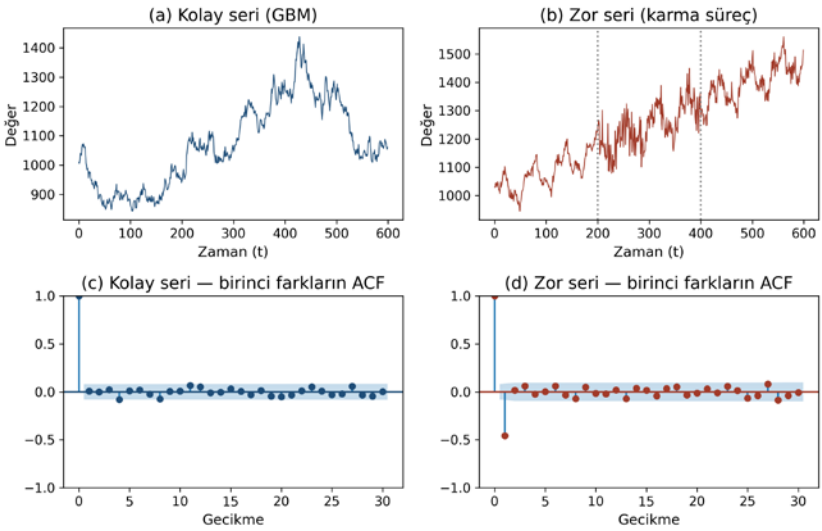
$$P_t = P_{t-1} \cdot \exp(\mu + \sigma\varepsilon_t), \quad \varepsilon_t \sim N(0, 1) \quad (5)$$

Burada $\mu = 0{,}0003$ sürüklenme, $\sigma = 0{,}015$ oynaklık parametresi ve $P_0 = 1000$ başlangıç değeridir.

Bu serinin düzey değerleri güçlü biçimde otokorelasyonludur (rastgele yürüyüş), birinci farkları ise yaklaşık olarak ilişkisizdir (Şekil 1c). Dolayısıyla geçmiş değerlerden öngörülebilir bir bileşen içermez ve kuramsal olarak en iyi tahmin "yarın = bugün" biçimindeki naif tahmine çok yakındır. **Zor-Karma** serisi ise gerçek serilerde sıkça görülen birden çok yapısal bileşeni bir araya getirir:

$$Y_t = 1000 + 0{,}8t + 50\sin(2\pi t/60) + 25\cos(2\pi t/20) + \eta_t + a_t \quad (6)$$

Burada $a_t = 0{,}6a_{t-1} - 0{,}3a_{t-2} + u_t$, $u_t \sim N(0, 5^2)$ bir AR(2) bileşenidir; η_t ise üç rejimli gürültüdür: $t < 200$ için $N(0, 10^2)$, $200 \leq t < 400$ için $N(0, 40^2)$ ve $t \geq 400$ için $N(5, 20^2)$. Böylece seri doğrusal trend, iki frekanslı mevsimsellik, oynaklık rejimi değişimleri ve otoregresif yapıyı birlikte içerir (Şekil 1b).



Şekil 1. Sentetik veri setleri. (a) Kolay-GBM serisi; (b) Zor-Karma serisi (kesikli dikey çizgiler oynaklık rejimi geçişlerini gösterir); (c)–(d) birinci farkların otokorelasyon fonksiyonları (%95 güven bantlarıyla).

3.2. Özellik Mühendisliği ve Veri Sızıntısının Önlenmesi

Her iki seride hedef değişken y_t serinin t anındaki düzey değeridir. Özellik vektörü yalnızca geçmiş bilgiden oluşturulmuştur: beş gecikmeli değer (y_{t-1}, …, y_{t-5}) ile y_{t-1}, …, y_{t-5} değerlerinin ortalaması ve standart sapması. Başlangıçtaki eksik değerler atıldıktan sonra her seride 595 gözlem kalmış; bunların ilk %80'i (476 gözlem) eğitim ve çapraz doğrulama için, son %20'si (119 gözlem) yalnızca son değerlendirme için ayrılan bağımsız test kümesi olarak kullanılmıştır. Standartlaştırma parametreleri her katta yalnızca eğitim bölümünden kestirilmiştir.

Pencere istatistiklerinin hesaplanmasında dikkat edilmesi gereken önemli bir nokta vardır. Hareketli ortalama ve standart sapma t anındaki değerin kendisini de içeren bir pencere üzerinden hesaplanırsa, $y_t = 5 \cdot$ (hareketli ortalama) - ($y_{t-1} + \ldots + y_{t-4}$) özdeşliği gereği hedef değişken, özelliklerin doğrusal bir bileşimi olarak birebir elde edilebilir. Bu durumda doğrusal bir model özdeşliği $\alpha \rightarrow 0$ durumunda tam olarak öğrenir; tüm optimizasyon algoritmaları arama uzayının alt sınırına yakınsar ve gerçekçi olmayan ölçüde düşük hatalar elde edilir. Bu tür hatalar literatürde "veri sızıntısı" (data leakage) olarak adlandırılmaktadır (Kaufman vd., 2012). Bu bölümdeki tüm analizlerde pencere istatistikleri yalnızca t-1 ve öncesindeki değerlerden hesaplanmış, ayrıca hedef değişkenin özelliklerden doğrusal olarak yeniden üretilemediği sayısal olarak doğrulanmıştır.

3.3. Deney Tasarımı

Deney 2 veri seti × 3 çapraz doğrulama stratejisi × 6 algoritma = 36 koşuldan oluşan tam faktöriyel bir tasarımdır. Her koşul, r = 1, …, 30 için 7r + 6 tohum değerleriyle (13, 20, …, 216) 30 kez bağımsız olarak çalıştırılmıştır. Karşılaştırmanın adil olması için her algoritmaya çalıştırma başına aynı bütçe, yani 50

amaç fonksiyonu değerlendirmesi verilmiştir. Eşit bütçe, güvenilirlik karşılaştırmasının algoritmadan çok bütçe farkıyla belirlenmesini önler. Bütçenin sonuçlara etkisini görmek için deney 20 değerlendirmelik düşük bütçeyle de tekrarlanmıştır. Algoritma ayarları Tablo 1'de verilmiştir.

Tablo 1. Algoritma ayarları (bütçe = 50 değerlendirme)

Alg.	Uygulama	Temel ayarlar
GA	Özgün kod	Popülasyon 10; ikili turnuva; BLX-0,5 (p_c = 0,9); Gauss mutasyonu (p_m = 0,2; σ = 1,0); ($\mu+\lambda$) seçilim
DE	SciPy	10 birey; best1bin; F ∈ [0,5; 1]; CR = 0,7; 4 kuşak; yerel iyileştirme kapalı
PSO	PySwarms	10 parçacık; 5 iterasyon; $c_1 = c_2$ = 1,5; w = 0,7
ABC	Özgün kod	5 besin kaynağı; limit = 5; uygunluk 1/(1+f); çevrim başına ≤ 1 kâşif; en iyi çözüm belleği
BO-TPE	Optuna	TPE örnekleyici; ilk 10 deneme rastgele; 50 deneme
Rastgele	NumPy	50 tekdüze örnek

3.4. İstatistiksel Analiz

Her veri seti–çapraz doğrulama koşulunda altı algoritmanın 30 çalıştırmadaki ÇD-RMSE değerleri Friedman testiyle karşılaştırılmıştır (Demšar, 2006). Çalıştırmalar aynı tohum dizisiyle eşleştirildiği için test, eşleştirilmiş bloklar olarak uygulanmıştır. Eşleştirilmiş tohum kullanımının, bağımsız değerlendirmelerde istatistiksel olarak sonuçsuz kalan karşılaştırmaları netleştirebildiği gösterilmiştir (Sharma, 2025). Etki büyüklüğü olarak Kendall uyum katsayısı $W = \chi^2_F / [N(k - 1)]$ hesaplanmıştır; burada N = 30 çalıştırma ve k = 6 algoritma sayısıdır (Kendall ve Smith, 1939). Friedman testinin anlamlı olduğu durumlarda 15 algoritma çifti için Wilcoxon işaretli sıralar testi uygulanmış ve p değerleri çoklu karşılaştırma için Holm yöntemiyle düzeltilmiştir (Demšar, 2006). Tüm testlerde anlamlılık düzeyi %5'tir.

3.5. Yazılım ve Tekrarlanabilirlik

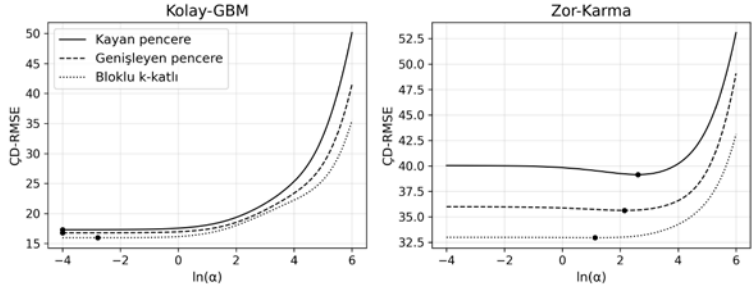
Analizler Python 3.11 ortamında NumPy 2.4, pandas 3.0, scikit-learn 1.8, SciPy 1.17, Optuna 5.0, PySwarms 1.3 ve statsmodels 0.15 kütüphaneleriyle yapılmıştır. Kod, sonuçları

yeniden üretebilecek biçimde tüm tohum değerlerini sabitlemektedir. Analizlerde kullanılan kod, talep edilmesi hâlinde yazardan temin edilebilir.

4. BULGULAR

4.1. Amaç Fonksiyonunun Yapısı

Algoritmaların davranışını yorumlayabilmek için öncelikle amaç fonksiyonunun yapısını incelemek gerekir. Şekil 2'de her koşul için ÇD-RMSE'nin ln α'ya göre değişimi verilmiştir. Kolay-GBM serisinde kayan ve genişleyen pencere stratejilerinde global optimum arama uzayının alt sınırındadır (ln α^* = -4,00); bloklu k-katlı stratejide ise ln α^* = -2,78'dir ve üç eğri de ln α < 0 bölgesinde neredeyse düzdür. Zor-Karma serisinde ise optimum her üç stratejide de iç noktadadır: ln α^* kayan pencerede 2,61 (α ≈ 13,5), genişleyen pencerede 2,15 (α ≈ 8,5) ve bloklu k-katlı stratejide 1,13 (α ≈ 3,1) olarak bulunmuştur. Dolayısıyla aynı veri ve aynı model için "en iyi" hiperparametre, doğrulama stratejisine bağlı olarak dört kattan fazla değişmektedir.



Şekil 2. Çapraz doğrulama RMSE'sinin ln(α)'ya göre değişimi (2.001 noktalı ızgara). Noktalar her stratejinin global optimumunu gösterir.

4.2. Çözüm Kalitesi ve Güvenilirlik

Amaç fonksiyonu altında algoritmaların global optimuma ne ölçüde ve ne kadar tutarlı biçimde ulaştığı Tablo 2'de

özetlenmiştir. En dikkat çekici bulgu, farkların mutlak büyüklüğünün son derece küçük olmasıdır. 36 koşulun tamamında ortalama pişmanlık %0,024'ün altında, tek bir çalıştırmada gözlenen en büyük pişmanlık ise %0,39'dur (ABC, Zor-Karma, kayan pencere). Başarı oranı bütün koşullarda %93,3 veya üzerindedir ve taban çizgisi olan rastgele arama bile koşulların beşinde %100 başarıya ulaşmıştır.

Tablo 2. Güvenilirlik ölçütleri (30 çalıştırma, bütçe = 50)

ÇD stratejisi	Alg.	K: Ort. pişm. (%)	K: VK (%)	K: Başarı (%)	Z: Ort. pişm. (%)	Z: VK (%)	Z: Başarı (%)
Kayan pencere	GA	0,0128	0,0406	96,7	0,0039	0,0078	100,0
	DE	0,0010	0,0007	100,0	0,0017	0,0036	100,0
	PSO	0,0026	0,0032	100,0	0,0015	0,0025	100,0
	ABC	0,0066	0,0361	96,7	0,0169	0,0713	96,7
	BO-TPE	0,0001	0,0001	100,0	0,0003	0,0005	100,0
	Rastgele	0,0039	0,0057	100,0	0,0231	0,0423	93,3
Genişleyen pencere	GA	0,0061	0,0201	96,7	0,0019	0,0036	100,0
	DE	0,0004	0,0003	100,0	0,0013	0,0023	100,0
	PSO	0,0011	0,0014	100,0	0,0010	0,0019	100,0
	ABC	0,0033	0,0181	100,0	0,0060	0,0135	100,0
	BO-TPE	0,0001	0,0001	100,0	0,0004	0,0006	100,0
	Rastgele	0,0017	0,0025	100,0	0,0059	0,0094	100,0
Bloklu k-katlı	GA	0,0012	0,0063	100,0	0,0035	0,0107	100,0
	DE	0,0000	0,0001	100,0	0,0002	0,0005	100,0
	PSO	0,0001	0,0002	100,0	0,0001	0,0003	100,0
	ABC	0,0012	0,0058	100,0	0,0008	0,0031	100,0
	BO-TPE	0,0000	0,0000	100,0	0,0001	0,0001	100,0
	Rastgele	0,0001	0,0001	100,0	0,0016	0,0032	100,0

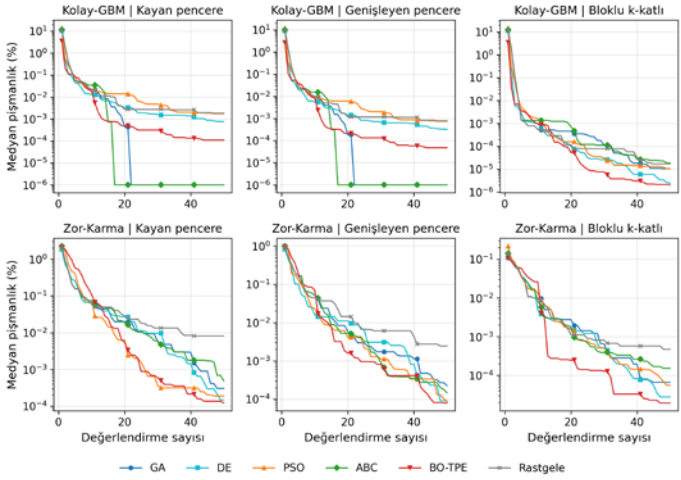
Not: K: Kolay-GBM; Z: Zor-Karma; Pişm.: global optimuma göre göreli pişmanlık (Eşitlik 3); VK: ÇD-RMSE'nin varyasyon katsayısı; Başarı: pişmanlığı %0,1'in altında kalan çalıştırmaların oranı.

Göreli sıralamada ise BO-TPE, altı koşulun tamamında en düşük ortalama pişmanlığı, en düşük VK değerini (≤ %0,0006) ve en küçük $SS_{\ln \alpha}$ değerini vermiş DE ve PSO onu yakından izlemiştir. GA ve ABC ise iki uçlu bir davranış sergilemiştir.

Kolay-GBM serisinde optimum arama sınırında olduğundan, bu iki algoritma çalıştırmaların çoğunda optimumu tam olarak bulmuş, ancak birkaç çalıştırmada görece uzak noktalarda kalmıştır. Bu nedenle VK ve $SS_{\ln \alpha}$ değerleri bu iki algoritmada en yüksektir (ör. GA için $SS_{\ln \alpha} = 0{,}553$).

4.3. Yakınsama Davranışı

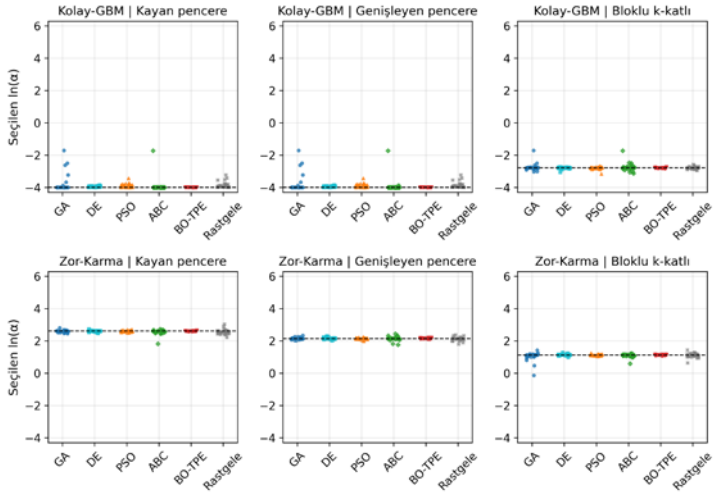
Nihai sonuçların yanı sıra algoritmaların optimuma hangi hızla yaklaştığı da güvenilirlik açısından önemlidir. Şekil 3'te değerlendirme sayısına göre medyan pişmanlık gösterilmiştir. Tüm algoritmalar ilk 10-15 değerlendirmede pişmanlığı %1-10 düzeyinden %0,1 civarına veya altına indirmekte, sonraki kazanımlar ise RMSE ölçeğinde ihmal edilebilir kalmaktadır. BO-TPE, ilk 10 rastgele denemenin ardından vekil modeli devreye girdiğinde hızla iyileşmekte ve optimumun iç noktada olduğu dört koşulda en düşük medyan pişmanlığa ulaşmaktadır. Rastgele aramanın Zor-Karma serisinde diğer yöntemlerin gerisinde kalması, öğrenme mekanizmasına sahip yöntemlerin katma değerinin ancak optimum iç noktada olduğunda görünür hâle geldiğini göstermektedir.



Şekil 3. Değerlendirme sayısına göre medyan göreli pişmanlık (30 çalıştırma, logaritmik ölçek).

4.4. Seçilen Hiperparametrenin Kararlılığı

Hata düzeyindeki yakınlık, seçilen hiper parametrelerin de birbirine yakın olduğu anlamına gelmeyebilir. Şekil 4'te 30 çalıştırmada seçilen ln α değerleri global optimumla (kesikli çizgi) birlikte verilmiştir. Zor-Karma serisinde tüm algoritmaların seçimleri optimum çevresinde dar bir bantta toplanmıştır. Kolay-GBM serisinde ise GA ve ABC'nin birkaç çalıştırması ln α ≈ -1,7 değerine kadar uzaklaşmıştır. Amaç fonksiyonu bu bölgede çok düz olduğundan (Şekil 2), parametre düzeyindeki bu büyük sapma RMSE düzeyinde yalnızca %0,2'lik bir pişmanlığa karşılık gelmektedir. Bu bulgu, güvenilirliğin "hata düzeyinde" ve "parametre düzeyinde" ayrı ayrı değerlendirilmesi gerektiğini göstermektedir.



Şekil 4. Her algoritmanın 30 çalıştırmada seçtiği ln(α) değerleri. Kesikli çizgi global optimumu gösterir.

4.5. İstatistiksel Testler

Gözlenen farkların istatistiksel olarak anlamlı olup olmadığı, Friedman testi ve Holm düzeltmeli ikili Wilcoxon karşılaştırmalarıyla sınanmıştır. Tablo 3'te görüldüğü gibi Friedman testi altı koşulun tamamında anlamlıdır ($p < 0,01$).

Ancak etki büyüklükleri, Kolay-GBM serisinin kayan ve genişleyen pencere koşulları (W = 0,60 ve 0,57) dışında küçüktür (W = 0,12–0,22) ve Holm düzeltmesinden sonra 15 çiftten 2 ile 9'u arasında anlamlı fark bulunmuştur. Sıra tabanlı testler 10^{-4} düzeyindeki farkları da tam bir sıra farkı olarak değerlendirir. Örneğin Kolay-GBM-kayan pencere koşulunda ABC ile BO-TPE'nin medyan ÇD-RMSE değerleri 17,26762 ve 17,26764'tür. Fark istatistiksel olarak anlamlı (Holm düzeltmeli $p < 0,001$) olmasına karşın pratikte önemsizdir.

Tablo 3. Friedman testi, etki büyüklüğü ve Holm düzeltmeli ikili Wilcoxon sonuçları (bütçe = 50)

Veri seti	ÇD stratejisi	χ^2_F	p	W	En iyi ort. sıra	Anlamlı çift
Kolay-GBM	Kayan pencere	90,27	< 0,001	0,60	ABC (1,60)	9/15
Kolay-GBM	Genişleyen pencere	86,15	< 0,001	0,57	ABC (1,70)	9/15
Kolay-GBM	Bloklu k-katlı	19,92	0,001	0,13	BO-TPE (2,50)	2/15
Zor-Karma	Kayan pencere	30,49	< 0,001	0,20	BO-TPE (2,57)	6/15
Zor-Karma	Genişleyen pencere	18,45	0,002	0,12	BO-TPE (2,87)	2/15
Zor-Karma	Bloklu k-katlı	32,58	< 0,001	0,22	BO-TPE (2,67)	4/15

Not: Serbestlik derecesi = 5. "En iyi ort. sıra": 30 çalıştırma üzerinden ortalama sırası en küçük algoritma. W: Kendall uyum katsayısı. "Anlamlı çift": Holm düzeltmesinden sonra $p < 0,05$ olan algoritma çifti sayısı.

4.6. Çapraz Doğrulama Stratejisinin Etkisi ve Genelleme

Algoritmalar arasındaki farkların bu denli küçük olması, belirleyici etkenin başka bir yerde aranması gerektiğini göstermektedir. Tablo 4'te görüldüğü gibi çapraz doğrulama stratejisinin etkisi, algoritma seçiminin etkisinden birkaç mertebe büyüktür. Aynı veri setinde strateji değiştiğinde optimum ÇD-RMSE Kolay-GBM serisinde %8,3, Zor-Karma serisinde %18,8 oranında değişmektedir. Buna karşılık, belirli bir koşulda algoritmalar arasındaki ortalama test RMSE farkı %0,1'in

altındadır. Tüm stratejiler test hatasını olduğundan yüksek tahmin etmiştir (negatif yanlılık). Yanlılık her iki seride de en küçük değerini bloklu k-katlı stratejide (Kolay-GBM: -%0,8; Zor-Karma: -%20,5), en büyük değerini ise kayan pencerede (-%8,5 ve -%31,9) almıştır. Zor-Karma serisindeki büyük kötümser yanlılığın temel nedeni, test kümesinin (t = 481-599) orta oynaklıklı üçüncü rejime, ÇD katlarının bir kısmının ise yüksek oynaklıklı ikinci rejime denk gelmesidir.

Tablo 4. Çapraz doğrulama stratejisine göre optimum, naif tahmin ve test kümesi sonuçları

Veri seti	ÇD stratejisi	ln α^*	f^* (ÇD)	Naif (ÇD)	Test (α^*)	Naif (test)	Yanlılık (%)
Kolay-GBM	Kayan pencere	-4,00	17,27	16,68	15,80	15,91	-8,5
Kolay-GBM	Genişleyen pencere	-4,00	16,76	16,19	15,80	15,91	-5,7
Kolay-GBM	Bloklu k-katlı	-2,78	15,94	15,68	15,81	15,91	-0,8
Zor-Karma	Kayan pencere	2,60	39,16	43,30	26,68	28,59	-31,9
Zor-Karma	Genişleyen pencere	2,15	35,64	40,59	26,45	28,59	-25,8
Zor-Karma	Bloklu k-katlı	1,13	32,95	35,69	26,18	28,59	-20,5

Not: f^*: global optimumdaki ÇD-RMSE; Naif: $\hat{y}_t = y_{t-1}$ tahmininin RMSE'si; Test (α^*): α^* ile tüm eğitim kümesinde yeniden eğitilen modelin test RMSE'si; Yanlılık: Eşitlik (4).

Naif tahminle karşılaştırma da önemli bir bulguyu ortaya koymaktadır. Kolay-GBM serisinde Ridge modeli, üç stratejinin tamamında ÇD aşamasında naif tahminden daha kötü sonuç vermiştir (ör. kayan pencere: 17,27'ye karşı 16,68). Test kümesinde ise naif tahmine göre yalnızca %0,6-0,7 oranında daha iyidir. Bu sonuç, rastgele yürüyüş yapısındaki bir seride geçmiş değerlerin öngörü gücü taşımamasıyla tutarlıdır. Zor-Karma serisinde ise model naif tahmini ÇD aşamasında %7,7-12,2, test kümesinde %6,7-8,4 oranında geride bırakmaktadır.

4.7. Düşük Bütçe Duyarlılık Analizi

Ana deneydeki 50 değerlendirmelik bütçe tek boyutlu bu problem için görece geniş olduğundan, analiz 20 değerlendirmelik düşük bütçeyle de tekrarlanmıştır. Bu durumda algoritmalar arasındaki farklar belirginleşmektedir. BO-TPE altı koşulun tamamında %100 başarıyı korurken, başarı oranı ABC'de %76,7'ye, GA ve rastgele aramada Zor-Karma–kayan pencere koşulunda %76,7'ye, PSO'da ise Kolay-GBM–kayan pencere koşulunda %83,3'e düşmüştür. DE yalnızca Zor-Karma–kayan pencere koşulunda %86,7'ye gerilemiştir. Başarının en çok düştüğü strateji, her iki seride de kayan pencere stratejisidir.

5. TARTIŞMA

Bulgular üç ana noktada toplanabilir. Birincisi, çapraz doğrulama stratejisinin seçimi, optimizasyon algoritmasının seçiminden çok daha belirleyicidir. Strateji değiştiğinde optimum hiperparametre dört kattan fazla, optimum ÇD hatası %19'a kadar, genelleme hatası tahminindeki yanlılık ise Zor-Karma serisinde -%20,5 ile -%31,9, Kolay-GBM serisinde -%0,8 ile -%8,5 arasında değişmiştir. Algoritma seçimi ise 50 değerlendirmelik bütçede test hatasını %0,1'den daha az etkilemiştir. Bu sonuç, doğrulama stratejisinin serinin yapısal özelliklerine göre seçilmesi gerektiğini vurgulayan önceki çalışmalarla (Bergmeir ve Benítez, 2012; Roberts vd., 2017) uyumludur ve bu tartışmayı hiperparametre optimizasyonunun güvenilirliği bağlamına taşımaktadır. Bloklu k-katlı stratejinin her iki seride de en az yanlı tahmini vermesi, Bergmeir ve Benítez (2012) tarafından bildirilen bulgularla uyumludur. Ancak bu sonuç, burada kullanılan iki sentetik seriyle sınırlıdır.

İkincisi, istatistiksel anlamlılık ile pratik önem birbirinden ayrılmalıdır. Friedman testi tüm koşullarda anlamlı olmasına karşın, farklar çoğunlukla ÇD-RMSE'nin dördüncü ya da beşinci

ondalık basamağındadır. Sıra tabanlı testler bu denli küçük ancak tutarlı farkları kolayca yakalar. Bu nedenle algoritma karşılaştırmalarında p değerlerinin yanında etki büyüklüklerinin ve problemin ölçeğinde yorumlanabilir ölçütlerin (pişmanlık, test hatası farkı) raporlanması önerilmektedir. VK ölçütü de tek başına güvenilirlik göstergesi olarak yetersizdir. Örneğin tüm çalıştırmaların arama uzayının sınırındaki aynı noktaya yakınsadığı bir problemde VK = 0 elde edilir; ancak bu değer algoritmanın güvenilirliğini değil, problemin yapısını yansıtır.

Üçüncüsü, algoritmalar arasındaki farklar, arama bütçesi kısıtlı olduğunda ve optimum iç noktada bulunduğunda belirginleşmektedir. 20 değerlendirmelik bütçede BO-TPE tüm koşullarda %100 başarıyı korumuş, ABC ve GA ise kayan pencere koşulunda en fazla gerileyen algoritmalar olmuştur. BO-TPE'nin örneklem verimliliği, her değerlendirmesi maliyetli olan gerçek uygulamalarda (ör. derin öğrenme modelleri) daha da önemli olabilir (Bergstra vd., 2011). Öte yandan tek boyutlu ve düzgün bir problemde rastgele aramanın da yüksek başarı göstermesi, metasezgisel yöntemlerin katma değerinin problemin boyutuna ve karmaşıklığına bağlı olduğunu hatırlatmaktadır. Bu durum, hiçbir optimizasyon algoritmasının tüm problem sınıflarında diğerlerinden üstün olamayacağını ifade eden "bedava öğle yemeği yoktur" teoremiyle de uyumludur (Wolpert ve Macready, 1997).

Son olarak, güvenilir yapay zekâ açısından algoritmik güvenilirliğin ölçülmesi, uygulamanın doğruluğunun bağımsız olarak denetlenmesiyle birlikte ele alınmalıdır. Bölüm 3.2'de açıklanan türden bir özellik mühendisliği hatası, istatistiksel testlerin tamamı "anlamlı" sonuç verse bile bütün sayısal bulguları geçersiz kılabilir. Naif tahminle karşılaştırma, hedefin özelliklerden yeniden üretilip üretilemediğinin sınanması ve algoritmaların bilinen optimumlu problemlerde doğrulanması, bu

tür hataları erken aşamada ortaya çıkarabilecek basit kontrollerdir.

5.1. Kısıtlar

Çalışmanın kısıtları şunlardır: (i) Arama uzayı tek boyutludur. Çok boyutlu ve karma (sürekli–kesikli) hiperparametre uzaylarında algoritmalar arasındaki farkların daha büyük olması beklenebilir. (ii) Model olarak yalnızca Ridge regresyon kullanılmıştır. (iii) Yalnızca iki sentetik seri incelenmiş ve her seri tek bir gerçekleşmeyle (tohum = 42) temsil edilmiştir; bulgular farklı gerçekleşmeler için doğrulanmamıştır. (iv) Algoritmaların parametreleri ayarlanmamış, yaygın varsayılan değerler kullanılmıştır. (v) Gerçek bir veri seti kullanılmamıştır.

6. SONUÇ

Bu bölümde, zaman serisi hiperparametre optimizasyonunda altı arama yönteminin istatistiksel güvenilirliği, üç çapraz doğrulama stratejisi ve iki sentetik seri üzerinde eşit bütçeyle ve 30 bağımsız çalıştırmayla incelenmiştir. Tek boyutlu bu problemde 50 değerlendirme ile tüm yöntemler global optimumun %0,4 yakınına ulaşmış, farklar istatistiksel olarak anlamlı ancak pratikte önemsiz kalmıştır. BO-TPE, en düşük pişmanlık, en düşük dağılım ve düşük bütçede en yüksek başarı oranıyla öne çıkan yöntem olmuştur. Asıl belirleyici etken ise çapraz doğrulama stratejisidir. Strateji, seçilen hiperparametreyi, ulaşılan doğrulama hatasını ve genelleme hatası tahminindeki yanlılığı algoritma seçiminden çok daha fazla etkilemektedir.

Uygulayıcılar için şu öneriler çıkarılabilir: (1) Metasezgisel algoritmalar tek çalıştırmayla değil, çok sayıda bağımsız çalıştırma, eşit bütçe ve etki büyüklüğü içeren testlerle karşılaştırılmalıdır. (2) Güvenilirlik yalnızca varyasyon

katsayısıyla değil, bilinen ya da iyi kestirilmiş bir optimuma göre pişmanlık ve seçilen parametrenin dağılımıyla birlikte raporlanmalıdır. (3) Çapraz doğrulama stratejisi serinin yapısına göre seçilmeli ve ÇD tahmini mutlaka bağımsız bir test kümesiyle ve naif bir taban çizgisiyle karşılaştırılmalıdır. (4) Özellik mühendisliği adımı, hedef değişkenin özelliklerden yeniden üretilip üretilemediği sınanarak veri sızıntısına karşı denetlenmelidir. Gelecek çalışmalarda analizlerin çok boyutlu hiperparametre uzaylarına, doğrusal olmayan modellere ve gerçek zaman serilerine genişletilmesi planlanmaktadır.

KAYNAKÇA

Akiba, T., Sano, S., Yanase, T., Ohta, T., & Koyama, M. (2019). Optuna: A next-generation hyperparameter optimization framework. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* içinde (ss. 2623–2631). https://doi.org/10.1145/3292500.3330701

Bergmeir, C., & Benítez, J. M. (2012). On the use of cross-validation for time series predictor evaluation. *Information Sciences, 191*, 192–213. https://doi.org/10.1016/j.ins.2011.12.028

Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of Machine Learning Research, 13*(2), 281–305.

Bergstra, J., Bardenet, R., Bengio, Y., & Kégl, B. (2011). Algorithms for hyper-parameter optimization. *Advances in Neural Information Processing Systems, 24*, 2546–2554.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy, 81*(3), 637–654. https://www.jstor.org/stable/1831029

Boquet, T., Delisle, L., Kochetkov, D., Schucher, N., Oreshkin, B. N., & Cornebise, J. (2019). Reproducibility and stability analysis in metric-based few-shot learning. *ICLR 2019 Workshop on Reproducibility in Machine Learning*. https://openreview.net/forum?id=B1g-SnUaUN adresinden erişildi.

Bouthillier, X., Delaunay, P., Bronzi, M., Trofimov, A., Nichyporuk, B., Szeto, J., … Vincent, P. (2021). Accounting for variance in machine learning benchmarks.

Proceedings of Machine Learning and Systems, 3, 747–769.

Chou, J., Truong, D. N., & Le, T. L. (2020). Interval forecasting of financial time series by accelerated particle swarm-optimized multi-output machine learning system. *IEEE Access, 8.* 14798-14808. https://doi.org/10.1109/ACCESS.2020.2965598

Demšar, J. (2006). Statistical comparisons of classifiers over multiple data sets. *Journal of Machine Learning Research, 7*, 1-30.

Eryılmaz, B., Koraş, O. A., Schlötterer, J., & Seifert, C. (2024). Investigating the impact of randomness on reproducibility in computer vision: A study on applications in civil engineering and medicine. In *2024 IEEE 6th International Conference on Cognitive Machine Intelligence (CogMI)* (ss. 265-274). IEEE.

Eshelman, L. J., & Schaffer, J. D. (1993). Real-coded genetic algorithms and interval-schemata. *Foundations of genetic algorithms* (Cilt 2, ss. 187–202) içinde. Elsevier.

Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics, 12*(1), 55–67. https://www.jstor.org/stable/1267351

Holland, J. H. (1992). *Adaptation in natural and artificial systems*. Cambridge, MA: MIT Press.

Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and practice* (2nd ed.). Melbourne, Australia: OTexts.

Karaboga, D., & Basturk, B. (2007). A powerful and efficient algorithm for numerical function optimization: Artificial bee colony (ABC) algorithm. *Journal of Global*

Optimization, 39(3), 459–471. https://doi.org/10.1007/s10898-007-9149-x

Kaufman, S., Rosset, S., Perlich, C., & Stitelman, O. (2012). Leakage in data mining: Formulation, detection, and avoidance. *ACM Transactions on Knowledge Discovery from Data, 6*(4), 1–21. http://doi.acm.org/10.1145/2382577.2382579

Kendall, M. G., & Smith, B. B. (1939). The problem of m rankings. *The Annals of Mathematical Statistics, 10*(3), 275–287. https://www.jstor.org/stable/2235668

Kennedy, J., & Eberhart, R. (1995). Particle swarm optimization. *Proceedings of the IEEE International Conference on Neural Networks* içinde (ss. 1942–1948). Perth, Australia. 10.1109/ICNN.1995.488968

Liem, C. C. S., & Panichella, A. (2020). *Run, Forest, Run? On randomization and reproducibility in predictive software engineering* (arXiv:2012.08387). https://arxiv.org/abs/2012.08387 adresinden erişildi.

Miranda, L. J. (2018). PySwarms: A research toolkit for particle swarm optimization in Python. *Journal of Open Source Software, 3*(21), 433. 10.21105/joss.00433

Picard, D. (2021). Torch.manual_seed(3407) is all you need: On the influence of random seeds in deep learning architectures for computer vision (arXiv:2109.08203). https://arxiv.org/abs/2109.08203 adresinden erişildi.

Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., … Dormann, C. F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography, 40*(8), 913–929. https://www.jstor.org/stable/10.2307/90011350

Sharma, U. (2025). When does pairing seeds reduce variance? Evidence from a multi-agent economic simulation https://doi.org/10.48550/arXiv.2512.24145

Storn, R., & Price, K. (1997). Differential evolution - A simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization, 11*(4), 341–359.

Tashman, L. J. (2000). Out-of-sample tests of forecasting accuracy: An analysis and review. *International Journal of Forecasting, 16*(4), 437-450.

Vos, G., van Eijk, L., Sarnyai, Z., & Azghadi, M. R. (2025). Stabilizing machine learning for reproducible and explainable results: A novel validation approach to subject-specific insights. *Computer Methods and Programs in Biomedicine, 269*, 108899. https://doi.org/10.1016/j.cmpb.2025.108899

Wolpert, D. H., & Macready, W. G. (1997). No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation, 1*(1), 67–82. 10.1109/4235.585893

Wu, K., Chai, Y., Zhang, X., & Zhao, X. (2022). Research on power price forecasting based on PSO-XGBoost. *Electronics, 11*(22), 3763. https://doi.org/10.3390/electronics11223763

Yihan, L., Liang, H., & Ruiqi, M. (2023, November). Enhancing Short-term Electric Load Forecasting with EMD-GA Optimized LSTM. In *2023 6th International Conference on Power and Energy Applications (ICPEA)* (ss. 147-152). IEEE. https://doi.org/10.1109/ICPEA59834.2023.10398683

Yusoff, M., Sharif, M. Y., & Sallehud-Din, M. T. M. (2023, November). Long term short memory with particle swarm optimization for crude oil price prediction. In *2023 7th International Symposium on Innovative Approaches in Smart Technologies (ISAS)* (pp. 1-4). IEEE. 10.1109/ISAS60782.2023.10391535.

Zhao, Y. (2025). Real-time multi-step time series forecasting using a GA-optimized VMD-RF framework. *Research Square*. https://doi.org/10.21203/rs.3.rs-7493076/v1

Zulfiqar, M., & Rasheed, M. B. (2022). Short-term load forecasting using long short term memory optimized by genetic algorithm. In 2022 IEEE Sustainable Power and Energy Conference (iSPEC) (ss. 1-5). IEEE. https://doi.org/10.1109/iSPEC54162.2022.10033074.

Zulfiqar, M., Gamage, K. A., Rasheed, M. B., & Gould, C. (2024). Optimised deep learning for time-critical load forecasting using LSTM and modified particle swarm optimisation. Energies, 17(22), 5524. https://doi.org/10.3390/en17225524

LİSELERDE ALAN SEÇİMİ ÖLÇÜTLERİNİN BİREYSEL AHP, BOOTSTRAP VE BAYESÇİ YAKLAŞIMLARLA ANALİZİ: ÖNCELİK KARARLILIĞI, TUTARSIZLIK VE TOPLUMSAL FARKLILAŞMA

Ramazan PEHLİVAN[1]

1. GİRİŞ

Lise döneminde alan seçimi, öğrencinin yalnızca sonraki iki yılda izleyeceği dersleri belirleyen bir tercih değildir. Seçilen alan, yükseköğretime geçişte karşılaşılabilecek seçenekleri ve öğrencinin kendisi için öngördüğü eğitim ve meslek yolunu da etkileyebilmektedir. Bu nedenle alan seçimi, kişisel özellikler ile akademik başarı, aile beklentileri, ekonomik koşullar, sosyal çevre ve geleceğe ilişkin beklentilerin birlikte değerlendirildiği çok ölçütlü bir karar problemi olarak ele alınabilir. Lise öğrencilerinin kariyer tercihleri üzerine yapılan çalışmalar da kararların tek bir ilgi veya yetenek değişkeniyle açıklanamayacağını; aile, okul, akademik deneyim, ekonomik beklentiler ve sosyal çevrenin farklı ağırlıklarda rol oynayabildiğini göstermektedir (Akosah-Twumasi vd., 2018; Koçak vd., 2021; Siddiky & Akter, 2021). Türkiye'deki lise öğrencilerine ilişkin araştırmalar da alan ve kariyer tercihinin bireysel özelliklerle birlikte eğitim ve çevre koşullarıyla ilişkili olduğunu ortaya koymaktadır (Bayraktar, 2018; Karacan-Ozdemir, 2019).

[1] Dr. Öğr. Üyesi, İstanbul Esenyurt Üniversitesi, Uygulamalı Bilimler Fakültesi, Veri Bilimi ve Analitiği Bölümü, ORCID: 0000-0002-7805-3554.

Literatürde öne çıkan ölçütler arasında kişisel ilgi ve yetenek, akademik başarı, aile etkisi, ekonomik beklentiler, iş güvencesi ve sosyal statü bulunmaktadır. Bununla birlikte bu ölçütlerin göreli önemi her bağlamda aynı değildir; kültürel ve toplumsal çevrenin yanı sıra öğrencinin içinde bulunduğu eğitim ortamı da tercihlerin oluşumunda farklılaşmaya yol açabilmektedir (Akosah-Twumasi vd., 2018; Hartung vd., 2022). Dolayısıyla alan seçimini yalnızca hangi faktörlerin etkili olduğu sorusuyla değil, öğrencilerin bu faktörlere hangi göreli önemi verdikleri sorusuyla incelemek daha açıklayıcı bir çerçeve sunmaktadır.

Bu ölçütlerin göreli önemlerinin birlikte değerlendirilmesi için Analitik Hiyerarşi Süreci (AHP) kullanılabilir. AHP, ölçütlerin ikili karşılaştırmalar yoluyla göreli önceliklere dönüştürülmesini ve karar vericilerin öznel yargılarının sistematik biçimde değerlendirilmesini sağlar. (Saaty, 1980; Tavana vd., 2023). Eğitim ve kariyer seçiminde AHP'nin kullanıldığı çalışmalar, öğrencilerin farklı ölçütlere ilişkin tercihlerini ağırlıklar biçiminde ortaya koymanın mümkün olduğunu göstermektedir (Önder vd., 2014; Sumo vd., 2023). Bununla birlikte, AHP uygulamalarında elde edilen sonuçların nasıl toplulaştırıldığı ve bireysel yargıların ne ölçüde korunabildiği önemli bir yöntemsel sorundur.

Grup AHP uygulamalarında katılımcıların ikili karşılaştırmaları çoğu zaman tek bir grup kararını temsil edecek şekilde birleştirilmektedir. Geometrik ortalama gibi toplulaştırma teknikleri grup önceliklerini elde etmek açısından işlevsel olmakla birlikte, bireyler arasındaki farklılıkların önemli bir bölümünü görünmez hâle getirebilir (Forman & Peniwati, 1998; Ossadnik vd., 2016). Bu nedenle, grup düzeyinde elde edilen düşük bir tutarlılık oranının bireysel kararların da tutarlı olduğu şeklinde yorumlanması yanıltıcı olabilir. AHP'de ikili karşılaştırmaların öznel olması nedeniyle tutarsızlık beklenebilir;

bu nedenle tutarlılığın yalnızca toplulaştırılmış matris düzeyinde değil, mümkün olduğunda karar verici düzeyinde de incelenmesi önem taşımaktadır (Grošelj & Dolinar, 2023).

AHP çalışmalarında tek bir kesin yargı değerinin, karar vericinin sahip olduğu belirsizliği her zaman yeterince yansıtmayabileceği görülmektedir. Bu soruna yönelik olarak bulanık, aralıklı ve Bayesçi yaklaşımlar geliştirilmiş ve karar yargılarındaki belirsizliğin modele dâhil edilmesine çalışılmıştır (Altuzarra vd., 2007; Liu vd., 2020; Grošelj ve Dolinar, 2023). Buna karşın, lise öğrencilerinin alan veya kariyer seçimi bağlamında bireysel AHP sonuçlarındaki belirsizliği hem bootstrap hem de Bayesçi yaklaşımla inceleyen çalışmaların sınırlı olduğu görülmektedir. Ayrıca AHP sonuçlarının değerlendirilmesinde yalnızca tutarlılık düzeyine bakmak yeterli değildir. Elde edilen öncelik sıralamasının küçük değişiklikler karşısında korunup korunmadığının da incelenmesi gerekir (Tavana vd., 2023).

Toplumsal farklılaşma da alan seçiminin değerlendirilmesinde dikkate alınması gereken bir başka boyuttur. Kariyer tercihleri; aile arka planı, okul ve sosyal çevre, cinsiyet ve iş gücü piyasasına ilişkin beklentiler gibi değişkenlerle farklılaşabilmektedir (Hartung vd., 2022). Bununla birlikte bu farklılıklar çoğu araştırmada klasik anket değişkenleri üzerinden incelenmekte, öğrencilerin çok ölçütlü karar yapısındaki farklılıklar ise daha sınırlı ele alınmaktadır. Lise öğrencilerinin bireysel ikili karşılaştırmalarını esas alarak tutarsızlık, belirsizlik, öncelik kararlılığı ve toplumsal farklılaşmayı aynı analitik çerçevede ele alan çalışmalar sınırlıdır (Ossadnik vd., 2016; Sumo vd., 2023). Bu çalışmada bu eksiklikten hareketle, öğrencilerin öncelikleri bireysel düzeyde hesaplanmış ve önce bireysel yargıların tutarlılığı incelenmiştir. Daha sonra ağırlıklardaki belirsizlik bootstrap ve hiyerarşik Bayesçi modelleme ile değerlendirilmiş, farklı analiz

koşullarında elde edilen sıralamanın ne ölçüde korunduğu test edilmiştir. Son olarak öncelik yapısının okul, alan ve diğer sosyodemografik özelliklere göre farklılaşıp farklılaşmadığı incelenmiştir.

Araştırmanın verileri 2018–2019 eğitim-öğretim yılında İstanbul'un Başakşehir ilçesindeki iki liseden elde edilmiştir; bu nedenle çalışma güncel bir öğrenci evrenini temsil etme iddiasında değildir. Veriler, sosyal bilimlerde etik kurul izni uygulamasının yaygınlaşmasından önceki döneme aittir. Uygulama okul yönetimlerinin bilgisi ve onayı dâhilinde, ders saati içinde ve gönüllülük esasına göre gerçekleştirilmiş; katılımcılardan ad, soyad, öğrenci numarası veya benzeri tanımlayıcı hiçbir bilgi istenmemiştir. Bu çalışmada yeni katılımcı toplanmamış; daha önce yayımlanmamış veri setinde yer alan bireysel kararlar yöntemsel açıdan yeniden analiz edilmiştir. Bu çerçevede araştırmanın temel soruları şunlardır: (1) Öğrenciler alan seçimini etkileyen ölçütleri hangi göreli önceliklerle değerlendirmektedir? (2) Bireysel yargı tutarsızlığı ve örnekleme bağlı belirsizlik, elde edilen öncelikleri ne ölçüde etkilemektedir? (3) Öncelik sıralaması farklı analiz koşullarında ne ölçüde kararlıdır? (4) Öncelik yapısı okul, alan ve diğer sosyodemografik özelliklere göre farklılaşmakta mıdır?

2. YÖNTEM

2.1. Katılımcılar

Araştırma kesitsel bir desende ve kolayda örnekleme yoluyla yürütülmüştür. Veriler, biri özel karma lise diğeri devlet lisesi olmak üzere iki okulun on beş şubesinden toplanmış; eksiksiz yanıt veren öğrenci sayısı 218 olmuştur. Katılımcıların tamamı alan seçimini yapmış durumdadır. Bu nedenle veriler, öğrencilerin gelecekte yapacakları bir tercihi değil, gerçekleşmiş alan seçiminin ardındaki ölçüt önceliklerini yansıtmaktadır.

Tablo 1. Katılımcıların Dağılımı (N = 218)

Değişken	n	Değişken	n
Kız	138	Özel lise	173
Erkek	80	Devlet lisesi	45
Sayısal (MF)	149	Baba: üniversite	107
Eşit ağırlık (TM)	65	Baba: lise ve altı	111
Yabancı dil	4	Anne: üniversite	60
16 yaş	147	Anne: lise ve altı	158
17 yaş	54	Mesleğe karar verdi	113
18 yaş	17	Henüz kararsız	105

Not: Annelerin 49'u bir işte çalışmaktadır. Yabancı dil alanındaki öğrenci sayısı (n = 4) yetersiz olduğundan alt grup analizlerinde bu grup dışarıda bırakılmıştır.

2.2. Ölçütler ve Verinin Hazırlanması

Alan seçiminde etkili olabilecek ölçütler, uygulama öncesinde yapılan görüşmeler ve öğretmen görüşleri doğrultusunda sekiz başlıkta belirlenmiştir: yetenek ve kişilik, mesleğin manevi tatmini, mesleğin maddi imkânları, ders notları, iş bulma imkânı, toplumsal saygınlık, çevrenin etkisi (rol model) ve ailenin isteği. Ölçütler ikişerli karşılaştırılmış, bu da her katılımcı için yirmi sekiz karşılaştırma anlamına gelmiştir. Karşılaştırmalar Saaty'nin (1980) dokuzlu oran ölçeğiyle kaydedilmiştir. Ölçüt kümesi tek düzeyli olduğu için AHP yalnızca ölçüt ağırlıklarını belirlemek amacıyla kullanılmıştır. Alternatiflerin ölçütler açısından değerlendirildiği bir performans matrisi bulunmadığından, alternatif sıralaması gerektiren yöntemlere bu çalışmada yer verilmemiştir.

Analiz öncesinde veri seti üzerinde üç işlem gerçekleştirilmiştir. Demografik bilgiler iki ayrı çalışma sayfasında ve farklı sınıf sırasıyla saklandığından, demografik bilgiler ile ikili karşılaştırmaların eşleştirilmesi sütun sırasına göre değil, sınıf adı ve sınıf içi sıra numarası kullanılarak yapılmıştır. Karşılıklı değerler veri girişinde iki ondalık haneye

yuvarlandığından (1/9 yerine 0,11) tam ölçek değerlerine geri eşlenmiştir; bu yuvarlama, matematiksel olarak var olmayan bir tutarsızlık üretmektedir. Son olarak on beş hücredeki veri girişi hatası düzeltilmiştir. Hazırlık sonrasında 218 öğrenciye ait 6.104 karşılaştırma hücresinin tamamı doludur; eksik veri bulunmamaktadır.

2.3. Öncelik Türetme ve Tutarlılık

Her katılımcı için oluşturulan sekize sekiz karşılaştırma matrisinden ölçüt ağırlıkları satır geometrik ortalaması yöntemiyle türetilmiştir (Crawford ve Williams, 1985). Sonuçların karşılaştırılması amacıyla özvektör yöntemi de uygulanmıştır. Grup ağırlıkları, yargıların geometrik ortalamasıyla (AIJ) ve bireysel ağırlıkların ortalamasıyla (AIP) hesaplanmıştır. Forman ve Peniwati'nin (1998) ayrımına dayanan bu iki yaklaşımdan AIJ temel yöntem olarak alınmış, AIP ise sağlamlık kontrolü amacıyla kullanılmıştır. Bireysel yargı tutarlılığı üç ölçütle değerlendirilmiştir: Saaty'nin (1980) tutarlılık oranı (n = 8 için rastgele indeks 1,41), satır geometrik ortalamasıyla doğrudan uyumlu geometrik tutarlılık indeksi (Aguarón ve Moreno-Jiménez, 2003; eşik 0,37) ve matristeki en kötü üçlü ilişkiyi yakalayan Koczkodaj (1993) indeksi.

2.4. Belirsizlik, Kararlılık ve Farklılaşma

Grup ağırlıkları nokta tahmini olarak değil, belirsizlik aralıklarıyla raporlanmıştır. İlk olarak, öğrenci düzeyinde 5.000 tekrarlı bootstrap yeniden örneklemesi yapılmıştır. Elde edilen dağılımdan yüzde 95 güven aralıkları ve her ölçütün belirli bir sırada yer alma olasılıkları hesaplanmıştır. Sıra alma olasılığı, stokastik çok ölçütlü kabul edilebilirlik analizinde tanımlanan sıra kabul edilebilirlik indeksiyle ilişkilidir (Lahdelma, Hokkanen & Salminen, 1998). İkinci olarak, öğrenci düzeyinde hiyerarşik bir Bayesçi çarpımsal model kurulmuştur (Altuzarra, Moreno-Jiménez & Salvador, 2007). Modelde her öğrencinin logaritmik

yargılarının kendi öncelik vektörü etrafında normal dağılımlı bir hata ile oluştuğu kabul edilmiştir. Bireysel öncelik vektörlerinin ise ortak bir çok değişkenli normal dağılımdan geldiği varsayılmıştır. Model dört zincirli Gibbs örneklemesiyle kestirilmiş ve yakınsama zincirler arası ölçek indirgeme istatistiği kullanılarak kontrol edilmiştir. Bu iki yaklaşımın birlikte kullanılması, elde edilen sonuçların yönteme ne ölçüde bağlı olduğunu karşılaştırmaya ve Bayesçi modeldeki büzüşme etkisinin tutarsız yargıların sonuçlar üzerindeki etkisini ne ölçüde azalttığını değerlendirmeye olanak sağlamıştır.

Sıralamanın hesaplama tercihlerine duyarlılığı on farklı senaryo üzerinden incelenmiştir. Bu senaryolar temel çözüm, özvektör yöntemi, AIP toplulaştırması, tutarlılık oranı için 0,30, 0,20 ve 0,10 eşiklerinin kullanılması, geometrik indeks için 0,37 eşiğinin uygulanması, karşılaştırmaların bir basamak rastgele kaydırılmasıyla oluşturulan yargı gürültüsü ve her sınıfın sırayla dışarıda bırakıldığı çözümleri kapsamaktadır. Senaryolar arası uyum Spearman sıra korelasyonuyla ölçülmüştür. Ölçüt ağırlıkları toplamı bire eşit bileşenlerden oluştuğu ve bir simpleks üzerinde yer aldığı için, çok değişkenli grup karşılaştırmalarında doğrudan ham ağırlıklar yerine merkezî logaritmik oran (clr) dönüşümü kullanılmıştır (Aitchison, 1986). Gruplar arası farklar permütasyon temelli çok değişkenli varyans analiziyle sınanmıştır. Ölçüt bazındaki karşılaştırmalarda Mann-Whitney U testi kullanılmış, etki büyüklüğü Cliff (1993) deltasıyla hesaplanmış ve çoklu karşılaştırmalar Benjamini ve Hochberg (1995) yöntemiyle düzeltilmiştir. Bütün alt grup analizleri Bayesçi modelin büzüştürülmüş ağırlıklarıyla tekrarlanmıştır.

2.5. Entropinin Rolü

Shannon (1948) entropisi çok ölçütlü karar verme uygulamalarında nesnel bir ağırlıklandırma yöntemi olarak kullanılabilmektedir. Ancak yöntemin temelinde alternatiflerin

ölçütler açısından ne ölçüde farklılaştığının değerlendirilmesi vardır. Bu çalışmada bir alternatifler kümesi bulunmadığından entropi ancak katılımcılar satırlara yerleştirilerek hesaplanabilir; bu durumda ölçülen şey ölçütlerin önemi değil, katılımcıların o ölçüt üzerinde ne ölçüde anlaşamadığıdır. Bu nedenle entropi çalışmada doğrudan ağırlıklandırma aracı olarak kullanılmamış, uzlaşı düzeyini değerlendirmek amacıyla bir tanı ölçütü olarak ele alınmıştır. Bununla birlikte, ağırlıklandırma yöntemi olarak kullanılması durumunda elde edilecek sonuçlar karşılaştırma amacıyla ayrıca sunulmuştur. Uzlaşı ayrıca değişim katsayısı ve Kendall uyuşum katsayısıyla betimlenmiştir. Bütün hesaplamalar Python ortamında, rastgeleliğe dayalı işlemlerde sabit bir çekirdek değeri kullanılarak gerçekleştirilmiştir.

3. BULGULAR

3.1. Ölçüt Ağırlıkları

Toplulaştırılmış ölçüt ağırlıkları Tablo 2'dedir. Öğrenciler en yüksek ağırlığı yetenek ve kişiliğe (%23,4), ikinci sırada mesleğin manevi tatminine (%21,8) vermektedir; bu iki ölçüt birlikte toplam ağırlığın yaklaşık yüzde kırk beşini oluşturmaktadır. Üçüncü sırada maddi imkânlar (%15,5), dördüncü sırada ders notları (%12,8) yer almakta; iş bulma imkânı ve toplumsal saygınlık birbirine yakın değerlerle (%8,4 ve %8,2) izlemektedir. En düşük iki ağırlık çevrenin etkisine (%6,5) ve ailenin isteğine (%3,4) aittir. Toplulaştırma biçiminin seçimi sıralamayı değiştirmemektedir; en belirgin fark, bireysel ağırlıkların ortalaması alındığında ders notlarının yaklaşık bir buçuk puan yükselmesidir. Bu fark, ders notlarına daha yüksek ağırlık veren öğrencilerin bulunduğunu düşündürmektedir.

Tablo 2. Ölçüt Ağırlıkları ve Toplulaştırma Karşılaştırması

Ölçüt	Sıra	AIJ (%)	AIP ort. (%)	AIP ortanca
Yetenek ve kişilik	1	23,36	21,66	23,68
Manevi tatmin	2	21,80	21,06	24,85
Maddi imkânlar	3	15,50	15,81	15,36
Ders notları	4	12,81	14,30	12,29
İş bulma imkânı	5	8,40	8,58	7,59
Toplumsal saygınlık	6	8,23	8,54	7,31
Çevre / rol model	7	6,53	6,59	6,13
Aile isteği	8	3,37	3,46	2,80

Not: AIJ = yargıların geometrik ortalamasıyla toplulaştırma; AIP = bireysel ağırlıkların toplulaştırılması. Sıra sütunu her iki yöntem için aynıdır.

3.2. Bireysel Yargı Tutarsızlığı

Toplulaştırılmış matris üzerinden hesaplanan tutarlılık oranı 0,011'dir; Saaty'nin 0,10 eşiğinin çok altında kalan bu değer, ilk bakışta katılımcıların son derece tutarlı yargılar verdiği izlenimini uyandırmaktadır. Bireysel düzeydeki tablo ise tamamen farklıdır. Tablo 3'te görüldüğü gibi bireysel tutarlılık oranının ortancası 0,269, ortalaması 0,351'dir ve katılımcıların yalnızca 23'ü (%10,6) eşiği sağlamaktadır. Geometrik tutarlılık indeksi benzer bir sonuç vermekte, 0,37 eşiğini sağlayan katılımcı sayısı 28'de (%12,8) kalmaktadır. Koczkodaj indeksinin ortancasının 0,968 olması ise neredeyse her matriste en az bir üçlü ilişkinin ciddi biçimde tutarsız olduğunu göstermektedir.

Bu karşıtlık yöntemsel olarak önemlidir: birleşik matrisin düşük tutarsızlık değeri, katılımcıların bireysel düzeydeki tutarlılığını doğrudan yansıtmamakta; geometrik ortalama ile yapılan toplulaştırmanın düzleştirici etkisini de içermektedir. Bireysel yargılardaki karşıt yönlü sapmalar toplulaştırma sırasında birbirini dengeleyebilmekte ve birleşik matriste daha düşük bir tutarsızlık düzeyi ortaya çıkabilmektedir. Grup AHP uygulamalarında yalnızca birleşik matrisin tutarlılık oranının

raporlanması bu nedenle bireysel yargıların tutarlılığı hakkında yeterli bilgi vermemektedir.

Tablo 3. Bireysel Tutarlılık Ölçütleri (N = 218)

Ölçüt	Ortanca	Ortalama	En az	En çok	Eşiği sağlayan
Tutarlılık oranı (CR)	0,269	0,351	0,006	1,654	23 (%10,6)
Geometrik indeks (GCI)	0,856	1,000	0,020	4,014	28 (%12,8)
Koczkodaj indeksi	0,968	0,950	0,500	0,999	—

Not: Eşikler CR için 0,10; GCI için 0,37'dir; Koczkodaj indeksi için yerleşik bir eşik yoktur. Örneklemin %23,4'ü 0,15; %33,9'u 0,20; %56,4'ü 0,30 eşiğinin altındadır.

3.3. Öncelik Kararlılığı

Yargı tutarsızlığının yüksekliği, elde edilen sıralamanın güvenilmez olduğu anlamına gelmemektedir. On senaryo altında hesaplanan ağırlıklar Tablo 4'tedir. Sekiz senaryoda sıralama temel çözümle birebir aynıdır (Spearman $\rho = 1{,}00$); yalnızca en sıkı iki filtrelemede ders notları ile iş bulma imkânı yer değiştirmekte ve korelasyon 0,976'ya inmektedir. En üstteki üç sıra hiçbir senaryoda değişmemektedir. Tutarlılık eşiği sıkılaştıkça ağırlıklarda sistematik bir kayma gözlenmekte, yetenek ve kişiliğin ağırlığı %28,1'e yükselirken ders notlarının ağırlığı %8,2'ye gerilemektedir; bu, tutarlı yargı veren öğrencilerin içsel ölçütlere daha net bir öncelik verdiğine işaret etmekle birlikte söz konusu alt örneklemler küçüktür ($n = 23$ ve $n = 28$).

Tablo 4. Senaryolara Göre Ölçüt Ağırlıkları (%)

Ölçüt	Temel	AIP	CR<0,30	CR<0,20	CR<0,10	LOO aralığı
Yetenek	23,4	21,7	26,5	28,1	26,9	22,9–23,9
Manevi	21,8	21,1	23,0	23,7	23,4	21,3–22,3
Maddi	15,5	15,8	14,3	12,9	12,7	14,8–16,1
Notlar	12,8	14,3	11,0	9,8	8,2	12,4–13,3
İş bulma	8,4	8,6	8,9	9,4	11,3	8,1–8,7
Saygınlık	8,2	8,5	7,5	7,0	7,6	8,0–8,5
Çevre	6,5	6,6	5,6	5,6	6,2	6,3–6,8
Aile	3,4	3,5	3,4	3,5	3,6	3,3–3,4
Spearman	1,00	1,00	1,00	1,00	0,98	1,00

Not: AIP = bireysel ağırlıkların toplulaştırılması; LOO aralığı = her sınıfın sırayla dışarıda bırakıldığı on beş çözümün uç değerleri. Son satır temel çözümle Spearman sıra korelasyonudur. Tabloya alınmayan iki senaryo (özvektör yöntemi ve yargı gürültüsü) temel çözümden hiçbir ölçütte 0,2 puandan fazla sapmamış, geometrik indekse göre daraltılmış çözümde (n = 28) korelasyon 0,976 çıkmıştır.

3.4. Belirsizliğin İki Yolla Nicelendirilmesi

Bootstrap ve Bayesçi yaklaşımların sonuçları Tablo 5'te karşılaştırılmıştır. İki yaklaşımın ürettiği ağırlıklar arasındaki ortalama mutlak fark 0,01 yüzde puanı, sıra korelasyonu birdir; aralık genişlikleri sırasıyla 2,23 ve 2,21 yüzde puanıdır. Bayesçi modelin yakınsama tanıları sorunsuzdur, sekiz ölçütün tamamında ölçek indirgeme istatistiği 1,000 düzeyindedir. Sıra kabul edilebilirlik indeksleri sıralamanın olasılık temelli bir okumasını sunmaktadır: yetenek ve kişilik yaklaşık yüzde doksan olasılıkla birinci sıradadır, kalan olasılığın tamamı manevi tatmine aittir ve bu iki ölçütün ilk üçte yer alma olasılığı yüzde yüzdür. Maddi imkânların ilk üçte olma olasılığı yüzde doksan dokuza yakın, ders notlarının ise yüzde iki dolayındadır. Sıralamanın üst ucu sabittir; belirsizlik yalnızca ilk iki sıranın kendi içindeki düzenine ilişkindir.

Tablo 5. Bootstrap ve Hiyerarşik Bayesçi Sonuçların Karşılaştırılması

Ölçüt	Bootstrap	%95 AA	Bayesçi	%95 GA	1. sıra
Yetenek ve kişilik	23,36	21,72–24,99	23,35	21,69–25,02	%90
Manevi tatmin	21,80	20,01–23,65	21,77	20,05–23,67	%10
Maddi imkânlar	15,50	14,01–17,06	15,50	14,04–17,09	—
Ders notları	12,81	11,42–14,28	12,79	11,40–14,37	—
İş bulma imkânı	8,40	7,64–9,25	8,40	7,64–9,23	—
Toplumsal saygınlık	8,23	7,45–9,06	8,22	7,47–9,03	—
Çevre / rol model	6,53	5,95–7,14	6,52	5,96–7,12	—
Aile isteği	3,37	3,09–3,68	3,37	3,08–3,67	—

Not: AA = bootstrap güven aralığı (5.000 tekrar); GA = Bayesçi güvenilirlik aralığı (4 zincir, 8.000 yineleme). Son sütun Bayesçi modelden elde edilen birinci sırada yer alma olasılığıdır.

3.5. Uzlaşı, Kutuplaşma ve Entropinin Yeri

Katılımcıların bireysel sıralamaları arasındaki uyumu ölçen Kendall uyuşum katsayısı 0,401'dir; bu değer öğrencilerin ortak bir yönelim paylaştığını, ancak bunun bir uzlaşıdan çok genel bir eğilim düzeyinde kaldığını göstermektedir. Tablo 6'da görüldüğü gibi öğrenciler yetenek ve kişilik ile manevi tatminde büyük ölçüde hemfikirdir (değişim katsayıları 0,52 ve 0,54); buna karşılık aile isteği (1,02) ve toplumsal saygınlıkta (0,90) dağılım belirgin biçimde kutuplaşmıştır. Aile isteği ortalamada en düşük ağırlığı alan ölçüt olmakla birlikte bir kısım öğrenci için hayli yüksek bir ağırlık taşımaktadır.

Aynı tablo, entropinin ağırlıklandırma aracı olarak kullanılması hâlinde ne olacağını da göstermektedir. Dağılımın en geniş olduğu ölçüte en yüksek ağırlığı veren entropi yöntemi, aile isteğini birinci sıraya (%16,6), yetenek ve kişiliği son sıraya (%7,1) yerleştirmektedir; iki ağırlık kümesi arasındaki Spearman korelasyonu –0,83'tür. Yani alternatifler kümesi bulunmayan bir tasarımda entropi, öğrencilerin en önemsiz saydığı ölçütü en önemli ölçüt olarak raporlamaktadır. Bu, yöntemin bu veri yapısına uygun olmadığının somut kanıtıdır.

Tablo 6. Uzlaşı Ölçütleri ve Entropi Yönteminin Vereceği Ağırlıklar

Ölçüt	AHP ağırlığı	Entropi (e)	Değişim kats.	Entropi ağırlığı
Yetenek ve kişilik	%23,4	0,975	0,52	%7,1
Manevi tatmin	%21,8	0,972	0,54	%8,1
Maddi imkânlar	%15,5	0,959	0,69	%11,8
Ders notları	%12,8	0,945	0,80	%15,6
İş bulma imkânı	%8,4	0,955	0,76	%12,8
Toplumsal saygınlık	%8,2	0,947	0,90	%15,3
Çevre / rol model	%6,5	0,955	0,79	%12,9
Aile isteği	%3,4	0,942	1,02	%16,6

Not: Entropi ağırlıkları yalnızca karşılaştırma amacıyla sunulmuştur; çalışmanın bulguları bu ağırlıklara dayanmamaktadır.

3.6. Toplumsal Farklılaşma

Öncelik yapısı okul, öğrencinin alanına ve cinsiyetine göre anlamlı biçimde farklılaşmakta; babanın ve annenin eğitim durumu, annenin çalışma durumu ve mesleğe karar verilip verilmemiş olması bakımından ise farklılaşmamaktadır (Tablo 7). Açıklanan varyans oranları bütün değişkenler için düşüktür; en yüksek değer yüzde 1,5 ile okul bağlamına aittir. Analiz Bayesçi modelin büzüştürülmüş ağırlıklarıyla tekrarlandığında okul bağlamı ve alan farkları korunmakta, hatta okul bağlamının etkisi güçlenmekte; cinsiyet farkı ise anlamlılığını yitirmektedir. Bu, cinsiyete bağlı farkın önemli ölçüde tutarsız yargı veren öğrencilerin uç ağırlıklarınca taşındığına, okul bağlamına bağlı farkın ise daha sağlam bir temele dayandığına işaret etmektedir.

Tablo 7. Öncelik Yapısının Alt Gruplara Göre Farklılaşması

Değişken	Gruplar (n)	F	R^2	p	p (Bayesçi ağırlıklar)
Okul bağlamı	173 / 45	3,34	0,015	0,005	0,002
Alan (MF / TM)	149 / 65	2,25	0,011	0,040	0,045
Cinsiyet	138 / 80	2,28	0,010	0,040	0,098
Meslek kararı	113 / 105	1,62	0,007	0,145	0,214
Baba eğitimi	107 / 111	1,41	0,007	0,205	0,241
Anne çalışma durumu	49 / 169	0,61	0,003	0,710	0,611
Anne eğitimi	60 / 158	0,53	0,003	0,778	0,796

Not: F = permütasyon temelli pseudo-F istatistiği; p değerleri 9.999 permütasyondan elde edilmiştir. Son sütun aynı analizin Bayesçi büzüştürülmüş ağırlıklarla tekrarından elde edilen p değeridir.

Farkların hangi ölçütlerden kaynaklandığı Tablo 8'dedir. Erkek öğrenciler maddi imkânlara kız öğrencilerden yaklaşık üç puan daha yüksek ağırlık vermekte; devlet lisesindeki öğrenciler yetenek ve kişilik ile manevi tatmine belirgin biçimde daha yüksek, maddi imkânlar ile toplumsal saygınlığa daha düşük ağırlık vermekte; sayısal alandaki öğrenciler ise eşit ağırlık alanındakilere kıyasla maddi imkânlar ve iş bulma imkânını öne çıkarmaktadır. Bu farkların istatistiksel anlamlılığı dikkatle değerlendirilmelidir: hiçbiri çoklu karşılaştırma düzeltmesinden sonra geleneksel anlamlılık düzeyine ulaşmamakta, en düşük düzeltilmiş değer 0,095'te kalmakta ve Cliff deltası hiçbir karşılaştırmada 0,25'i aşmamaktadır. Buna karşılık grup düzeyindeki ağırlık farklarının bootstrap dağılımları tutarlı bir yön göstermektedir; örneğin özel lise öğrencilerinin toplumsal saygınlığa daha yüksek ağırlık verme olasılığı yüzde 99,8'dir. Bulguyu "gruplar arasında anlamlı farklar vardır" biçiminde değil, "öncelik yapısı gruplar arasında sistematik ancak küçük ölçekli bir kayma göstermektedir" biçiminde ifade etmek bu nedenle daha doğrudur. Nitekim sıralama alt gruplarda büyük ölçüde korunmakta, on dört alt grubun sekizinde genel çözümle birebir aynı çıkmakta ve hiçbirinde yetenek ile manevi tatminin ilk iki sırayı paylaşması bozulmamaktadır.

Tablo 8. Anlamlı Farklılaşma Gösteren Değişkenlerde Ölçüt Bazlı Karşılaştırma

Değişken / Ölçüt	Karşılaştırma	Fark	p	q	delta
Cinsiyet: Maddi	Erkek – Kız	+3,19	0,027	0,217	0,18
Okul: Yetenek	Özel – Devlet	–4,75	0,012	0,095	–0,24
Okul: Maddi	Özel – Devlet	+3,44	0,050	0,113	0,19
Okul: Manevi	Özel – Devlet	–3,04	0,056	0,113	–0,19
Okul: Saygınlık	Özel – Devlet	+2,49	0,029	0,113	0,21
Alan: Maddi	MF – TM	+3,43	0,027	0,112	0,19
Alan: İş bulma	MF – TM	+1,97	0,028	0,112	0,19

Not: Fark sütunu grup ortalamaları arasındaki yüzde puanı farkıdır. p = Mann-Whitney U testi; q = Benjamini-Hochberg düzeltilmiş değer; delta = Cliff etki büyüklüğü. Yalnızca düzeltme öncesi $p < 0,06$ olan karşılaştırmalar gösterilmiştir.

4. TARTIŞMA

Bulguların öne çıkan yönlerinden biri, öğrencilerin alan seçimi ölçütleri arasında yetenek ve kişilik ile manevi tatmini daha yüksek önceliklerle değerlendirmesidir. Yetenek ve kişilik ile manevi tatmin birlikte ağırlığın neredeyse yarısını almakta, ailenin isteği ise en düşük ağırlıklı ölçüt olarak kalmaktadır. Aile isteğinin ağırlığı yalnızca düşük değil, aynı zamanda dar bir aralıkta kestirilmektedir (yüzde 95 aralığı 3,1–3,7); yani öğrenciler bu ölçüt konusunda belirgin bir konum almaktadır. Ancak bu sonucun ailenin gerçekten etkisiz olduğu biçiminde okunması için veride bir dayanak yoktur. Ölçülen şey ailenin nesnel etkisi değil, öğrencinin kendi kararını nasıl gerekçelendirdiğidir. Aile etkisi doğrudan bir talep biçiminde değil, hangi mesleklerin görünür ve arzu edilir kılındığı, hangi okulun seçildiği, hangi derslere destek verildiği gibi dolaylı kanallardan işliyor olabilir; böyle bir dolaylılık, sorulduğunda düşük ağırlık alan ancak fiilen belirleyici olan bir etki üretebilir. Bununla bağlantılı dolaylı bir bulgu, mesleğe henüz karar vermemiş öğrencilerin aile isteğine karar vermiş olanlardan daha

yüksek ağırlık vermesidir; fark küçük olmakla birlikte bootstrap dağılımı farkın bu yönde olma olasılığını yüzde 99,8 vermektedir. Kararsızlık hâlinde dışsal ölçütlerin görünürlüğünün arttığı düşünülebilir; ancak bu, kesitsel veriyle sınanamayacak bir süreç iddiası olarak hipotez düzeyinde bırakılmaktadır.

Öncelik yapısında belirgin farklılaşmanın gözlendiği değişkenlerden biri okul bağlamıdır ve bu fark Bayesçi büzüşme uygulandığında zayıflamak yerine korunmaktadır. Devlet lisesindeki öğrenciler yetenek, kişilik ve manevi tatmini; özel lisedeki öğrenciler ise maddi imkânlar, toplumsal saygınlık ve iş bulma imkânını görece öne çıkarmaktadır. Bu bulgunun yorumunda üç sınırlılığa dikkat edilmelidir. Birincisi, karşılaştırma iki okul arasındadır; okul düzeyindeki gözlem sayısı ikidir ve bu, "devlet okulu" ile "özel okul" arasında genel bir karşılaştırma yapılmasını imkânsız kılar. İkincisi, iki okul yalnızca türleri bakımından değil cinsiyet bileşimi, yaş ve alan dağılımı bakımından da farklıdır; devlet lisesinde sınıflar cinsiyete göre ayrıdır ve öğrenciler daha yaşlıdır. Gözlenen fark bu nedenle okul, cinsiyet ve yaşın birlikte ürettiği bileşik bir etkidir ve mevcut desenle ayrıştırılamaz. Üçüncüsü, örneklem büyüklükleri dengesizdir. Bulgu bu nedenle okul bağlamının öncelik yapısıyla ilişkili olduğu biçiminde ifade edilmeli, nedensel bir iddiaya dönüştürülmemelidir.

Katılımcıların yalnızca onda birinin klasik tutarlılık eşiğini sağlaması ilk bakışta veri kalitesine ilişkin bir sorun gibi görünebilir. Ancak yirmi sekiz ikili karşılaştırmanın on altı yaşındaki katılımcılar tarafından sınıf ortamında yanıtlandığı düşünüldüğünde bu sonuç şaşırtıcı değildir; tutarlılık oranı, katılımcıdan yargılarını zihninde tutarlı bir oran yapısına oturtmasını bekler ve bu, günlük düşünme biçiminden hayli farklı bir bilişsel taleptir. Asıl bulgu, tutarsızlığın yüksekliği ile sonucun kararlılığı arasındaki karşıtlıktır. Tutarlılık eşiği kademeli sıkılaştırıldığında örneklem onda birine indiğinde bile

sıralama neredeyse değişmemekte; aynı durum yargılara yapay gürültü eklendiğinde, kestirim yöntemi değiştirildiğinde ve sınıflar sırayla dışarıda bırakıldığında da geçerliliğini korumaktadır. Bu, tutarsızlığın rastgele bir dalgalanma niteliğinde olduğuna ve sistematik bir yanlılık üretmediğine işaret etmektedir. Uygulama açısından çıkarılacak sonuç, tutarsız katılımcıların topluca elenmesinin gereksiz bir bilgi kaybı olabileceği; daha verimli yolun tutarsızlığı bireysel düzeyde raporlamak ve sonucun buna duyarlılığını göstermek olduğudur.

Çalışmanın sınırlılıkları şunlardır. Örneklem iki okuldan ve tek bir ilçeden oluşmakta, kolayda örnekleme kullanılmaktadır; bulgular Türkiye'deki lise öğrencilerine genellenemez. Veriler 2018 yılına ait olduğundan, bulgular salgın öncesi bir kesit olarak okunmalıdır. Ölçüt ağırlıkları öz bildirime dayanmakta ve toplumsal beğenilirlik etkisine açık bulunmaktadır. Alt gruplar arası etki büyüklükleri küçüktür ve ölçüt bazındaki farklar düzeltme sonrasında anlamlılığa ulaşmamaktadır; bu farklar kesin sonuçlar değil, daha büyük örneklemlerde sınanmayı bekleyen yönelimlerdir. Desen kesitseldir ve hiçbir bulgu nedensel biçimde yorumlanamaz.

5. SONUÇ

Bu bölümde, iki lisede öğrenim gören 218 öğrencinin alan seçimi ölçütlerine verdikleri ağırlıklar, Analitik Hiyerarşi Süreci bireysel düzeyde uygulanarak çözümlenmiştir. Öğrenciler en yüksek ağırlığı yetenek ve kişiliğe, ikinci sırada mesleğin manevi tatminine vermekte, ailenin isteği ise en düşük ağırlıklı ölçüt olarak kalmaktadır. Bu sıralama denenen on senaryonun sekizinde birebir korunmuş, kalan ikisinde yalnızca alt sıralarda tek bir yer değiştirme gözlenmiştir.

Çalışmanın yöntemsel açıdan öne çıkan bulgusu, toplulaştırılmış matris üzerinden hesaplanan tutarlılık oranının

bireysel tutarlılığın bir göstergesi olmadığıdır. Birleşik matrisin tutarlılık oranı 0,011 iken bireysel ortanca 0,269'dur ve katılımcıların yalnızca yüzde 10,6'sı klasik eşiği sağlamaktadır. Grup AHP uygulamalarında bireysel tutarlılık dağılımının raporlanması, bu nedenle bir ayrıntı değil, sonucun doğru yorumlanması için gerekli bir bilgidir. Buna karşılık yüksek tutarsızlık, öncelik sıralamasının geçersiz olduğu anlamına gelmemektedir; bu çalışmada bireysel yargılardaki değişkenliğe rağmen genel öncelik yapısı büyük ölçüde korunmuştur.

Belirsizliğin iki bağımsız yolla nicelendirilmesi sonuçların yöntem tercihine bağlı olmadığını göstermiş; bootstrap ile Bayesçi modelin ürettiği ağırlıklar arasındaki ortalama fark bir yüzde puanın yüzde biri düzeyinde kalmıştır. Bayesçi modelin sağladığı ek bilgi, okul bağlamına bağlı farklılaşmanın büzüşme sonrasında güçlenmesine karşılık cinsiyete bağlı farklılaşmanın anlamlılığını yitirmesidir. Öncelik yapısı okul bağlamı ve alan ile ilişkili, ebeveyn eğitimi ve annenin çalışma durumuyla ilişkisizdir; ancak etki büyüklükleri küçüktür ve farklılaşma sıralamanın yeniden kurulması biçiminde değil, ölçütler arasındaki mesafelerin daralıp genişlemesi biçiminde ortaya çıkmaktadır. Öğrenciler seçimlerini bireysel yatkınlık diliyle anlatmakta, bu anlatının tonu ise içinde bulundukları okul bağlamına göre değişmektedir. Alternatifler kümesi bulunmayan bir tasarımda Shannon entropisinin ağırlıklandırma aracı olarak kullanılmasının önemi değil uzlaşmazlığı ölçtüğü ve bu veride sıralamayı büyük ölçüde tersine çevirdiği de ayrıca gösterilmiştir.

İleride yapılacak çalışmalar için üç öneri getirilebilir. Ölçüt ağırlıklarının yanı sıra alternatiflerin ölçütler açısından değerlendirilmesi de toplanırsa, öğrencilerin kendi ağırlıklarıyla hangi alana yönelmeleri gerektiği ile fiilen hangi alanı seçtikleri karşılaştırılabilir. Aynı ölçüt kümesinin salgın sonrası bir örneklemde uygulanması, bu çalışmanın 2018 kesitiyle doğrudan karşılaştırılabilir bir zemin sunacaktır. Okul sayısının artırılması

ve okulların çok düzeyli bir modelde ele alınması ise bu çalışmada ayrıştırılamayan okul, cinsiyet ve yaş etkilerinin birbirinden ayrılmasına imkân verecektir.

KAYNAKÇA

Aguarón, J., & Moreno-Jiménez, J. M. (2003). The geometric consistency index: Approximated thresholds. *European Journal of Operational Research, 147*(1), 137–145.

Aitchison, J. (1986). *The statistical analysis of compositional data*. London: Chapman and Hall.

Akosah-Twumasi, P., Emeto, T., Lindsay, D., Tsey, K., & Malau-Aduli, B. (2018). A systematic review of factors that influence youths career choices: The role of culture. *Frontiers in Education, 3*, 58.

Altuzarra, A., Moreno-Jiménez, J. M., & Salvador, M. (2007). A Bayesian priorization procedure for AHP-group decision making. *European Journal of Operational Research, 182*(1), 367–382.

Bayraktar, Y. (2018). Investigation of factors affecting field selection of 11th grade in high school students. *International Journal of Educational Research Review, 3*, 23–31.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B, 57*(1), 289–300.

Cliff, N. (1993). Dominance statistics: Ordinal analyses to answer ordinal questions. *Psychological Bulletin, 114*(3), 494–509.

Crawford, G., & Williams, C. (1985). A note on the analysis of subjective judgment matrices. *Journal of Mathematical Psychology, 29*(4), 387–405.

Forman, E., & Peniwati, K. (1998). Aggregating individual judgments and priorities with the analytic hierarchy

process. *European Journal of Operational Research, 108*(1), 165–169.

Grošelj, P., & Dolinar, G. (2023). Group AHP framework based on geometric standard deviation and interval group pairwise comparisons. *Information Sciences, 626*, 370–389.

Hartung, A., Wessling, K., & Hillmert, S. (2022). Interplay between family background and labour-market conditions in shaping students' occupational status expectations. *Journal of Education and Work, 35*, 405–421.

Karacan-Ozdemir, N. (2019). Associations between career adaptability and career decision-making difficulties among Turkish high school students. *International Journal for Educational and Vocational Guidance, 19*, 1–21.

Koczkodaj, W. W. (1993). A new definition of consistency of pairwise comparisons. *Mathematical and Computer Modelling, 18*(7), 79–84.

Koçak, O., Ak, N., Erdem, S. S., Sinan, M., Younis, M., & Erdoğan, A. (2021). The role of family influence and academic satisfaction on career decision-making self-efficacy and happiness. *International Journal of Environmental Research and Public Health, 18*(11), 5919.

Lahdelma, R., Hokkanen, J., & Salminen, P. (1998). SMAA - Stochastic multiobjective acceptability analysis. *European Journal of Operational Research, 106*(1), 137–143.

Liu, Y., Eckert, C., & Earl, C. (2020). A review of fuzzy AHP methods for decision-making with subjective judgements. *Expert Systems with Applications, 161*, 113738.

Ossadnik, W., Schinke, S., & Kaspar, R. H. (2016). Group aggregation techniques for Analytic Hierarchy Process and Analytic Network Process: A comparative analysis. *Group Decision and Negotiation, 25*, 421–457.

Önder, E., Önder, G., & Taş, N. (2014). Identifying the importance level of factors influencing the selection of nursing as a career choice using AHP. *Procedia - Social and Behavioral Sciences, 122*, 398–404.

Saaty, T. L. (1980). *The analytic hierarchy process: Planning, priority setting, resource allocation*. New York: McGraw-Hill.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal, 27*(3), 379–423.

Siddiky, M. R., & Akter, S. (2021). The students' career choice and job preparedness strategies: A social environmental perspective. *International Journal of Evaluation and Research in Education, 10*(2), 421–431.

Sumo, D. Z., Zhang, L., & Sumo, P. D. (2023). Career choice for ICT among Liberian students: A multi-criteria decision-making study using analytical hierarchy process. *Heliyon, 9*, e16445.

Tavana, M., Soltanifar, M., & Santos-Arteaga, F. (2023). Analytical hierarchy process: Revolution and evolution. *Annals of Operations Research, 326(2)*, 879–907.

TEK VE ÇOKLU FOTON ÖLÇÜMLERİ İLE KUANTUM RASTGELE SAYI ÜRETİMİ

Engin KARAMAN[1]

1. GİRİŞ

Günümüz bilgisayar sistemlerinde, şifreleme anahtarları genellikle dijital sayılar şeklinde temsil edilmekte ve bu anahtarların güvenliği, çeşitli fiziksel süreçler veya algoritmalar aracılığıyla sayıların tahmin edilemezlik düzeylerinin artırılmasıyla sağlanmaktadır. Bu amaçla kullanılan rastgele sayılar, genellikle rastgele sayı üreteçleri tarafından oluşturulmaktadır (Jacak vd., 2021). Rastgelelik, bir modeli olmayan ve tahmin edilemeyen olaylar bütününü kapsar. Bir dizinin rastgele bir dağılım özelliği göstermesinin ön koşulu, tüm öğelerin birbirinden bağımsız olarak üretilmesi ve dizideki bir sonraki elemanın değerinin önceki değerler baz alınarak tahmin edilemez olmasıdır (Bassham vd., 2010) . Örneğin, çeşitli amaçlarla alınan tek kullanımlık şifreler (OTP'ler), bankanın sağladığı pinler, bir siteye giriş esnasında ekranda beliren CAPTCHA'ların hepsinin bir rastgele sayı üretecinin çıktısı olması beklenir (Mannalath vd., 2023). Bu yöntemler, bilgisayar sistemlerindeki kriptografik güvenlik sağlama süreçlerinin temelini oluşturmakta ve kullanıcı verilerinin korunmasına yönelik önemli bir rol oynamaktadır (Murry, 1970).

Bu çalışmada, fotonik tabanlı bir rastgele sayı üreteci olan kuantum rastgele sayı üretecinin (QRNG) tasarımı oluşturmak

[1] Dr. Öğr. Üyesi, Yeditepe Üniversitesi, Bilgisayar ve Bilişim Fakültesi, Bilişim Sistemleri ve Teknolojileri Bölümü, İstanbul, Türkiye, ORCID: 0000-0002-2336-6289.

hedeflenmiştir. Bunun için Bulut temelli çalışma yapabilen Cirq kütüphanesi kullanılarak birçok operatörden oluşan kuantum devresi üzerinden 10 bitlik rastsal sayı kümesi oluşturulmuştur. Yapılan bağımsız testler sonucunda sayı setinin rastsallığı ispat edilecektir.

2. RASTGELE SAYI ÜRETEÇLERİ VE RASTGELELİK TESTLERİ

Her e-posta gönderme işleminde, online banka hesabına bağlanıldığında veya tıbbi muayene kontrol işlemleri gibi sayısız bir çok çevrimiçi etkinliğin güvenliğini korumak için rastgele sayılara ihtiyaç bulunur. Bunu gerçekleştiren ve matematik biliminin bir alt disiplini olan Kriptografi, çevrimiçi ortamda güvenliğimizi sağlamak için kullandığımız araçlar dizisidir ve rastgele sayılar, kriptografinin üzerine inşa edildiği en önemli temeldir. Başka bir deyişle, öngörülemeyen rastgele rakamlar üretemeseydik, güvenli çevrimiçi iletişim mümkün olmazdı.

Önyargılı üreteçler tarafından gerçekten rastgele bit dizilerinin üretilmesi, bilgisayar bilimi, kriptografi ve istatistiksel uygulamalarda önemli bir zorluktur. Baskın görüş, klasik fizik yasalarının determinizmi nedeniyle, hiçbir klasik gerçekleştirmenin gerçekten rastgele bit dizileri üretemeyeceği yönündedir (Landau & Lifshitz, 2007).

2.1. Rastgele Sayı Üreteçleri

Rastgele sayı üreteçleri, zor tahmin edilebilen sayı dizileri oluşturmak için kullanılan mekanizmalar olarak ikiye ayrılır (Mannalath vd., 2023).

Sözde Rastgele Sayı Üreteçleri (Pseudo-Random Number Generators-SRSÜ), deterministik veya matematiksel olarak belirlenmiş algoritmalar kullanarak tahmin edilmesi zor rastgele sayılar üretirler (Boyar, 2023). Bu algoritmalar, önceden

tanımlanmış bir başlangıç noktasından (seed) başlayarak deterministik bir şekilde ilerler ve belirli bir kalıpta sayılar üretirler (James, 1990). Çıkış dizileri genellikle 0 ile 1 arasında mükemmel bir şekilde dengelenmiş olmasına rağmen, kriptografik güvenliği zayıflatabilecek, bilimsel simülasyonlarda beklenmeyen hatalara neden olabilecek veya temel fizik testlerinde açık boşluklara neden olabilecek güçlü bir uzun menzilli korelasyon mevcuttur (Ma vd., 2016). Sözde rastgele sayı üreteçlerinin amacı, hızlı ve rastgele nitelikte sayılar üretmektir. Günümüz programları ve yazılımlarının yüksek hızda çalışması nedeniyle bu hızın sağlanması büyük önem taşımaktadır. Ayrıca, bu sistemler yoğun miktarda rastgele sayı gereksinimi duymaktadırlar. Güvenliğin kritik olmadığı sistemlerde, rastgele sayı üretim maliyetinin düşürülmesi ve hızlı çalışmanın sağlanabilmesi adına, sözde rastgele sayı üreteçleri tercih edilmektedir.

Gerçek Rastgele Sayı Üreteçleri (True-Random Number Generators -GRSÜ) deterministik değildir; yani algoritmalar yerine tahmin edilmesi zor fiziksel süreçleri kullanarak rastgele sayı dizileri oluştururlar. Bu süreçler genellikle radyoaktif bozunma, termal gürültü, Johnson-Nyquist gürültüsü, Zener diyot gürültüsü, kaotik gürültü veya diğer rastgelelik kaynaklarıyla gerçekleşirler (Lee & Cleaver, 2017; Stojanovski vd., 2001; Zhun & Hongyi, 2001). Bu fiziksel süreçler kullanılarak entropi kaynakları oluşturulur ve bu düzensizlik kaynağı aracılığıyla rastgele sayılar üretirler. Entropi kaynağı, üretilen sayıların tahmin edilemez ve tekrar üretilemez olmasını sağlamak açısından kritiktir. Bu düzensizlikler, bilgisayar sistemlerindeki yazılım kaynaklı fiziksel verilerden elde edilebileceği gibi ayrı bir cihazla da sağlanabilir. GRSÜ'lerin genel çalışma mantığı, ölçülen bir değerin belirli sınırların dışında olup olmamasına göre 0 veya 1 değeri atamasına dayanır (Hanköylü, 2019). Bu

kaynaklar, GRSÜ olarak kullanılarak bilgisayar sistemlerinde güvenli rastgele sayı üretimine katkıda bulunurlar.

Gerçek Rastgele Sayı Üreteçleri, tahmin edilemezlik konusunda Sözde Rastgele Sayı Üreteçlerinden daha iddialı olsalar da bazı kısıtlamalarla karşılaşırlar. Öncelikle, GRSÜ'lerde kullanılan düzensizlik kaynakları, 0 veya 1 sayılarının elde edilme olasılıklarının eşit olmayabileceği durumlar üretebilmektedir. Bu nedenle, bazı uygulamalarda bu eğilimleri düzelten ek süreçler kullanma zorunluluğu getirmektedir. Ayrıca, GRSÜ'ler genellikle daha düşük hızlarda ve daha küçük bit aralıklarında sayı ürettikleri için SRSÜ'ler kadar hızlı değildirler (Rukhin vd., 2010).

Bir diğer önemli husus, GRSÜ'lerin kullanımında, elde edilen verilerin tam olarak ne kadar rastgele olduğu konusunda kesin bir sonuca varmanın zorluğudur, çünkü entropi süreçleri bazı durumları göz ardı ederek veri üretirler. Bu süreçlerde meydana gelen hatalar veya yanlışlıkların tespit edilmesi de oldukça zordur. Ayrıca, GRSÜ'lerin sistemlere entegrasyonu zahmetli olabilir ve bu nedenle kullanımları pratikte bazı zorluklar ortaya çıkabilmektedir (Stipčević & Koç, 2014).

Bu iki tür rastgele sayı üreteci, bilgisayar sistemlerinde kriptografik uygulamalardan bilimsel araştırmalara kadar geniş bir yelpazede kullanılmaktadır, veri güvenliği ve güvenilirlik açısından önemli roller üstlenmektedirler. Fiziksel rasgele sayı üreteçleri için olası kendiliğinden kesinti gibi sorunları da problemler ortaya çıkartabilmektedir. Kofler'in de belirttiği gibi, klasik rasgele sayı üreteçleri, (klasik fiziğin deterministik yasaları veya deterministik matematiksel bilgi algoritmaları tarafından belirlenen) deterministik üretim süreci nedeniyle, 0 ve 1 rakamları arasındaki mükemmel dengeye rağmen kaçınılmaz olarak her zaman aşağıdakilerin varlığıyla karakterize edilecek diziler üretir: Bilişim Teknolojileri güvenliği için potansiyel bir

risk oluşturabilecek belirli uzun vadeli deterministik modeller, bilimsel simülasyonlardaki beklenmeyen hatalar veya fiziksel süreç testlerindeki boşluklar (Kofler vd., 2006).

Bununla birlikte bir cihazın bileşenleri çalışmayı durdurursa veya bozulursa çıkış bitlerinin kalitesi değişebilir. Cihaz değerler üretiyorsa, cihazın gizli kusurlarının tespit edilmesi özellikle zordur. Bu nedenle güvenlik önerileri, gerçek kuantum sayı üreteçlerinde bir tür kendi kendini test etmeyi gerektirir. Cihazın durumu her zaman alt sistem tarafından denetlenmelidir (Iavich vd., 2021).

2.2. Kuantum Rastgele Sayı Üreteçleri

Gerçek rastgele sayı üreteçlerinin kullanımı kriptografide çok önemlidir. Kuantum fiziği perspektifinden, kuantum objeleri bir ölçüm yapılıncaya kadar süperpozisyon halinde bulunurlar, yani birçok farklı durum aynı anda var olabilmektedir. Özellikle, birbirine bağlı kuantum sistemleri tek bir dalga fonksiyonu ile tanımlanabilirler, bu duruma dolaşıklık denir (Ananthaswamy, 2019). Bu nedenle, klasik bilgisayar sistemlerinde kullanılan bitler sadece 0 veya 1 değerini alabilirken, kuantum bitleri (kübitler) hem 0 hem de 1 değerlerini aynı anda alabilirler. Bu özellik, başlangıç koşulları aynı olan durumların farklı sonuçlar verebileceği anlamını taşımaktadır. Kuantum sistemleri süperpozisyon durumlarının varlığı, ölçümde çökme, dolanık ama uzay benzeri ayrılmış parçacıkların varlığı, yerel olmayan korelasyonların varlığı, ayırt edilemeyen parçacıkların varlığı gibi büyüleyici yönlere sahiptir.

Kuantum mekaniğinin bu özellikleri, Kuantum Rastgele Sayı Üreteçleri (KRSÜ) tarafından üretilen sayıların rastgeleliğini artırır ve bu sayede iletişimin güvenliğini de artırır. KRSÜ'ler, kuantum süperpozisyon ve dolaşıklık özelliklerini kullanarak, deterministik olmayan, yüksek entropili rastgele sayı dizileri oluşturabilirler (Guillan-Lorenzo vd., 2023). Bu

özellikleri sayesinde, kuantum rastgele sayı üreteçleri kriptografik uygulamalarda güvenilirlik ve güvenlik sağlamak için önemli bir araç haline gelmektedirler (Ergen, 2020).

Mevcut kuantum teknolojisi oldukça olgunlaşmış olup, artık koşulsuz güvenli kriptografi uygulamalarında kullanılabilecek yüksek kaliteli rastgele sayılar üretmek mümkün olmaktadır. Kuantum kriptografisi sayesinde güvenli şifreleme sistemini tasarlanarak, saldırganların onu kırmaya karşı dirençli olan rastgele anahtar üretilmesine olanak sağlanmaktadır (Siswanto & Rudiyanto, 2017).

Kuantum Rastgele Sayı Üreteçleri (KRSÜ), Sözde Rastgele Sayı Üreteçleri (SRSÜ) ve klasik fizik tabanlı Gerçek Rastgele Sayı Üreteçleri (GRSÜ) ile karşılaştırıldığında, belirgin avantajlara sahiptirler. SRSÜ'ler matematiksel temele dayandıkları için deterministik olarak hesaplanabilirler ve bu özellikleri nedeniyle kriptografik uygulamalarda tercih edilmezler. Öte yandan, GRSÜ'ler klasik fizik yasalarına göre çalıştıkları için hızlı ve büyük boyutlarda rastgele bit dizileri üretmede sınırlamalar yaşarlar. Ayrıca, çevre koşulları gibi faktörlerden etkilenebilirler ve üretim sürecinde bazı zorluklarla karşılaşabilirler.

KRSÜ'ler ise, üretim sürecinde kullanılan fiziksel süreçlerden bağımsız olarak rastgele sayılar üretebilirler. Bu özellikleriyle kriptoloji uygulamaları için ideal üreteçler olarak kabul edilirler. Bununla birlikte, günümüzde hala karmaşık cihazlarla üretilmeleri gerektiğinden, iletişim hızının önemli olduğu durumlarda dezavantajları olabilir.

Genel olarak, KRSÜ'lerin hız ve rastgelelik açısından sağladığı avantajlara rağmen, üretim süreçlerinin karmaşıklığı günümüzdeki iletişim ihtiyaçlarını tam olarak karşılamakta yetersiz kalabilir. Bu nedenle, teknolojinin ilerlemesiyle birlikte KRSÜ'lerin pratik kullanımı ve performansı üzerindeki

çalışmalar önem kazanmaktadır. Rastgele sayı üreteçleri tarafından üretilen sayıların gerçekten rastgele kabul edilebilmesi için, üretilen sayıların birbirleriyle ilişkisiz olması ve bir sonraki sayının tahmin edilemez olması gerekmektedir.

KRSÜ'lerin fiziksel süreçlere dayandığı düşünüldüğünde, diğer gerçek rastgele sayı üreteçlerinin hız sorununa çözüm sunabildiği görülmektedir. KRSÜ'ler, kuantum mekanik prensiplerini kullanarak hızlı ve güvenilir bir şekilde gerçek rastgele sayılar üretebilirler, bu da özellikle bilgi güvenliği ve kriptografik uygulamalarda önemli bir avantaj sağlar.

Elde edilen entropi, olayların gerçekleşme olasılıkları açısından bir yanlılık içerdiğinde, bilginin sıkıştırılabilirliği artar; bunun aksi durumda ise sıkıştırılabilirlik azalır. Bilginin daha fazla sıkıştırılabilmesini mümkün kılan olaylar, ilgili entropi kaynağının yeterli rastgelelik içermediğini göstermektedir (Truong vd., 2019).

2.3. Rastgele Sayı Üreteçleri İçin İstatistiksel Testler

Bir sayı dizisinin rastsal bir karaktere sahip olup olmadığının belirlenebilmesi için literatürde pek çok test geliştirilmiştir.

İlk olarak, 1950 yılında algoritmik karmaşıklık teorisinin kurucusu Kolmogorov sayıların rastgeleliği üzerine çalışmalar gerçekleştirmiştir (Jacak vd., 2021). Ancak, sonraki yıllarda rastgeleliğin belirlenmesi ve doğrulanması için bazı istatistiksel testler kullanılmaya başlanmıştır. Bu testler, bir sayı dizisinin rastgeleliğini kesin olarak kanıtlamamakla birlikte (sonlu bir dizinin gerçek bir rastgele sayılar dizisi olup olmadığının tespiti mümkün değildir), sayılar arasında örüntüler olup olmadığı tespiti önem arz etmiştir. Bu şekilde, üretilen sayı dizisinin tahmin edilebilirliği ya da edilemezliği, çeşitli istatistiksel algoritmalar kullanılarak ölçülmektedir.

Literatürde bunun için kompozit (birden çok teknik içeren) test paketleri mevcuttur. Bu çalışmada, on beş istatistiksel rastgelelik parametresine sahip NIST istatistiksel test paketi kullanılacaktır. NIST test paketi (Computer Security Division, 2016), kaotik olaylar, ışık kaynağı, gürültü, kuantum veya sözde rastgele sayı üreteçleri gibi mevcut bir olaya dayalı olarak yazılım veya donanım tarafından oluşturulan ikili veri dizisinin rastgelelik düzeyini test etmek için oluşturulmuş bir çok test kriterinden oluşur. Bu testte kullanılan istatistiksel parametreler, ikili bir veri dizisindeki mevcut rastgele olmayan yönleri tanımlayacaktır. Pakette yer alan testlerin kısa bir tanımlaması aşağıda sunulmuştur (Bassham vd., 2010; Rukhin vd., 2010):

2.3.1. Frequency (Monobits) Test

Bu testin odağı, tüm dizi için sıfır ve birlerin oranıdır. Test, sıfır ve birlerin sayısının 1/2'ye yakın olup olmadığını, yani bir sekanstaki sıfır ve birlerin sayısının yaklaşık olarak eşit olup olmadığını ölçer.

2.3.2. Test For Frequency Within A Block

Bu test, M bitlik bloklar içindeki sıfır ve birlerin oranına odaklanır. Amacı, her bir M-bit blokta sıfır ve birlerin frekansının yaklaşık olarak M/2 olup olmadığını belirlemektir.

2.3.3. Runs Test

Bu testin amacı, dizideki sıfır ve birlerin akışlarının rastgele bir dizi için beklenen dağılımla uyumlu olup olmadığını belirlemektir. Bu, özellikle dizideki sıfır ve birler arasındaki dalgalanmaların hızını kontrol eder.

2.3.4. Test For The Longest Run Of Ones In A Block

Bu testin amacı, test edilen dizi içindeki en uzun dizilerin uzunluğunun, rastgele bir dizide beklenen en uzun dizilerin uzunluğuyla tutarlı olup olmadığını belirlemektir.

2.3.5. Random Binary Matrix Rank Test

Bu test, dizinin ayrık alt matrislerinin sırasını incelemeye odaklanır. Amacı, orijinal dizinin sabit uzunluklu alt dizileri arasında doğrusal bağımlılığı kontrol etmektir.

2.3.6. Discrete Fourier Transform (Spectral) Test

Bu test, ayrık Hızlı Fourier Dönüşümü'nün tepe yüksekliğine odaklanır. Amacı, test edilen sırada rastgele özellik varsayımından sapmayı gösterecek periyodik özellikleri, yani tekrarlayan desenleri tespit etmektir.

2.3.7. Non-Overlapping (Aperiodic) Template Matching Test

Bu test, önceden tanımlanmış hedef alt dizilerin gerçekleşme sayısını değerlendirir. Amacı, belirli bir periyodik olmayan (aperiodic) paternin çok fazla oluşumunu gösteren sekansları reddetmektir. Bu sınama, belirli bir m-bit kalıbını aramak için bir m-bit penceresi kullanır. Desen bulunamazsa, pencere bir bit konumu kayar; desen bulunduğunda ise pencere, bulunan kalıptan sonraki bitten itibaren yeniden başlar ve arama kaldığı yerden devam eder.

2.3.8. Overlapping (Periodic) Template Matching Test

Bu test, önceden tanımlanmış hedef alt dizilerin sayısını değerlendirir.

2.3.9. Maurer's Universal Statistical Test

Bu test, eşleşen modeller arasındaki bit sayısını değerlendirir. Amacı, dizinin bilgi kaybı olmadan önemli ölçüde sıkıştırılıp sıkıştırılamayacağını belirlemektir. Aşırı sıkıştırılabilir bir dizinin rastgele olmadığı kabul edilir.

2.3.10. Linear Complexity Test

Bu test, üreten geri besleme kaydının uzunluğunu inceler. Amacı, dizinin rastgele kabul edilecek kadar karmaşık olup olmadığını belirlemektir. Rastgele diziler daha uzun bir geri besleme kaydı ile karakterize edilirken, kısa bir geri bildirim kaydı rastgele olmama anlamına gelir.

2.3.11. Serial Test

Bu test, tüm sekans boyunca çakışan her bir m-bit paterninin frekansını değerlendirir. Amacı, 2^m m-bit çakışan örüntülerin meydana gelme sayısının rastgele bir dizi için beklenenden yaklaşık olarak aynı olup olmadığını belirlemektir.

2.3.12. Approximate Entropy Test

Bu test, çakışan her bir m-bit paterninin frekansına odaklanır. Amacı, ardışık iki uzunluktaki (m ve m+1) üst üste binen blokların sıklığını rastgele bir dizi için beklenen sonuçlarla karşılaştırmaktır.

2.3.13. Cumulative Sum (Cusum) Test

Bu test, test edilen sekansta meydana gelen kısmi sekansların kümülatif toplamının rastgele sekanslar için bu kümülatif toplamın beklenen davranışına göre çok büyük veya çok küçük olup olmadığını belirler. Rastgele bir dizi için, rastgele yürüyüş sıfıra yakın olmalıdır. Rastgele olmayan diziler için, bu rastgele yürüyüşün sıfırdan uzaklaşmaları büyük olacaktır.

2.3.14. Random Excursions Test

Bu test, rastgele bir yürüyüş içindeki belirli bir duruma yapılan ziyaret sayısının rastgele bir sekans için beklenenden fazla olup olmadığını belirler.

2.3.15. Random Excursions Variant Test

Bu test, belirli bir durumun kümülatif toplam rasgele yürüyüşte gerçekleşme sayısına odaklanır. Amacı, rastgele yürüyüşte çeşitli durumların beklenen oluşum sayısından sapmalarını tespit etmektir.

Testlerin başarılı olması için iki önemli parametre vardır. Bunlar P puanı ve önem seviyesi (α) dir. Önem seviyesinin 0.01 olarak belirlenmesi, test edilecek sayıların rastgeleliğinin %99 güven düzeyine sahip olduğunu gösterir. Rastgelelik ölçütü olarak hesaplanması gereken P-değeri 1'e eşit olduğunda, sayılar mükemmel rastgelelik göstermektedir. Bu değerden sapma durumunda ise, sayıların rastgele olmadığı kabul edilir.

3. DENEYSEL METOT

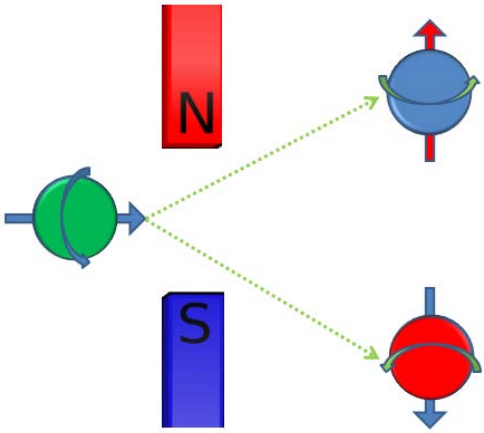
3.1. Kuantum Ölçümüne Dayalı Rastgelelik

Bu bölümde mevcut kuantum işlemcilerde rastgele sayılar üretmek için basit bir algoritma ele alınacaktır. Algoritma, temel durumdaki bir kübite Hadamard kapısı uygular ve ardından hesaplama temelinde kübitin durumunu ölçer. Eş. 1'de görüldüğü gibi, $|0\rangle$ 'a etki eden Hadamard kapısı, eşit genliklere sahip hesaplamalı temel durumların bir süperpozisyon durumunu üretir.

$$H|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \quad (1)$$

Bu durum ölçüldüğünde temel ve uyarılmış durumları elde etme olasılığı eşit olur. Hidary bu çalışma prensibinin, rastgele bir bit üreteci olarak kullanılabileceğini önermiştir (Hidary, 2019). Aynı şekilde Ma, gerçek bir rastgelelik, durumunun tutarlı süperpozisyonunu bozan herhangi bir kuantum sürecinden üretilebileceğini vurgulamıştır (Ma vd., 2016). Bir kuantum durumunun ölçüm sonucu doğası gereği olasılıksal (Sakurai, 1994) olması da bu teorileri desteklemektedir. Şekil

1'de gösterilen Stern-Gerlach deneyi foton davranışlarını gösteren bir simülasyon sunmaktadır.



Şekil 1. Stern-Gerlach deneyinde elektron spin tespiti

Stern-Gerlach deneyinde spin elektronun dikey eksen boyunca $|\uparrow\rangle$ ve $|\downarrow\rangle$ ile gösterilen iki yönde ilerlediği varsayılır. Eğer elektron başlangıçta iki dönüş yönünün süperpozisyonunda ise ($|\rightarrow\rangle = (|\uparrow\rangle + |\downarrow\rangle) / \sqrt{2}$) elektronun konumunun tespit edilmesi tutarlılığı bozacaktır ve sonuç (↑ veya ↓) doğası gereği rastgele olacaktır (Arteaga vd., 2019).

Her iki koldaki fotonların ölçülmesi, üst üste binmiş durumlardan yalnızca birini rastgele olarak verecektir. Yalnızca iki olası sonuç olduğundan, tespit edilen her foton en fazla bir rastgele bit üretecektir. Dedektör olarak ışın bölücüleri ve foto-çoğaltıcı tüpleri içeren, bu prensibe dayanan çeşitli uygulamalar mevcuttur (Weihs vd., 1998).

Rastgelelik yalnızca tek bir fotonun ölçülmesiyle değil aynı zamanda birden fazla foton içeren kuantum durumlarından da üretilebilir. Örneğin tutarlı bir durum farklı foton sayısı (Fock) durumlarının $|n\rangle$ üst üste binmesidir; burada n, foton sayısıdır ve α^2 tutarlı durumun ortalama foton sayısıdır.

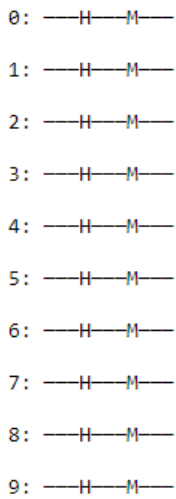
$$|\alpha\rangle = e^{-\frac{|\alpha|^2}{2}} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle (2)$$

Böylece, tutarlı bir lazer darbesinin foton sayısını, foton sayısı çözümleyici bir SPD (tek foton dedektörleri) ile ölçerek, Poisson dağılımını takip eden rastgele sayılar elde edebiliriz. Foton sayısının ölçülmesine dayanan QRNG'ler deneylerde başarıyla gösterilmiştir (Fürst vd., 2010; Ren vd., 2011).

3.2. Cirq ile Kuantum Rastgele Sayı Üreteci Tasarımı

Çalışmanın bu aşamasında fotonik tabanlı bir rastgele sayı üreteci olan kuantum rastgele sayı üretecinin (QRNG) tasarımı oluşturulacaktır. Bunun için Bulut temelli çalışma yapabilen Cirq uygulamasından yararlanılacaktır. **Cirq**, Google'ın kuantum devrelerini yazmak, değiştirmek ve optimize etmek ve ardından bunları kuantum bilgisayarlarda ve kuantum simülatörlerinde çalıştırmak için kullanılan bir Python yazılım kütüphanesidir. Bu sayede, geliştiricinin tekli, ikili ve üçlü operatörlerden oluşan kuantum devreleri oluşturmasına ve yürütmesine olanak tanır (*Cirq,* t.y.).

Gerekli yükleme ve güncellemeler yapıldıktan sonra Cirq üzerinden, bir kübit ile bir devre oluşturulmuştur. İkinci aşamada süperpozisyon oluşturmak için bite Hadamard kapısı uygulanır ve ardından bir ölçüm gerçekleştirilir. Program daha sonra devreyi simülatörde istenilen sayıda tekrarlayabilmektedir. Bu anlamda program, [0, N) tamsayı aralığındaki tek tip rastgele sayı üreteci olarak yorumlanabilir; burada N, devrenin tekrar sayısıdır. Bu çalışmada 10 bitlik rastgele sayılar oluşturulması amaçlanmıştır (N=10).



Şekil 2. 10 kübitli devre tasarımı

Bu devre aracılığıyla 2^{10} (1024) adet farklı ikili sayı dizisi elde edilebilecektir. İlk simülasyon sonucunda elde edilen bit değeri 1111010011 saptanmıştır. Bu değerin 10'luk sistemdeki karşılığı ise 979 olarak hesaplanmıştır.

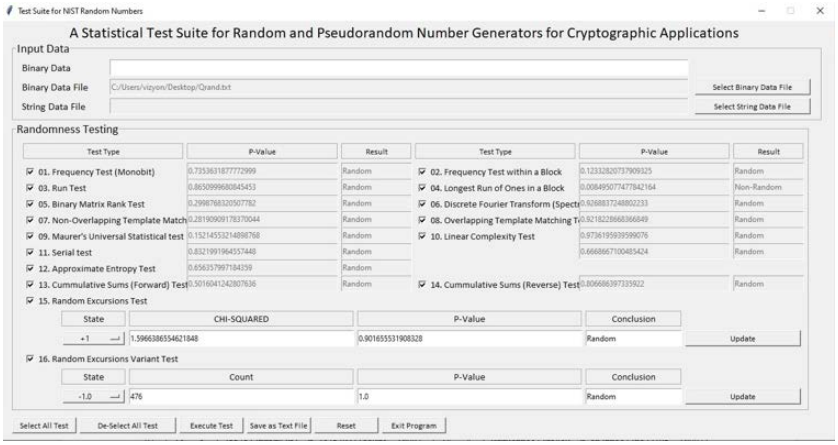
Son aşamada bu devre yapısı kullanılarak 10 bitlik 100 adet sayı oluşturulacaktır. Bu sayılar bir txt uzantılı dosyada (Qrand.txt) konumlandırıp final aşamasında sayıların rastgelelik koşulunun sağlanıp sağlanmadığı test edilecektir.

4. BULGULAR VE TARTIŞMA

Bir önceki bölümde de belirtildiği gibi 100 adet binary sayı seti sanal devre üzerinden oluşturulup, işlenmiştir. Bu sayı kümesinin rastsal bir dağılıma sahip olup-olmadığını saptamak için birçok istatistiksel rastgelelik parametresine sahip NIST istatistiksel test paketi kullanılacaktır. Bunun için GitHub üzerinde bulunan (stevenang/randomness_testsuite) araç kiti ile rastgele sayılar üzerinde NIST testlerini uygulayacağız. Bu kit

ilgili testlerin hepsini aynı anda hesaplama özelliğiyle kullanıcılara kolaylık sağlayan bir python arayüz uygulamasıdır.

Son olarak ürettiğimiz sayı kümesini içeren dosyayı uygulamaya yükleyerek mevcut testler ile analize tabi tutulmuştur. Tüm testlerin sonuçları p value değeri üzerinden hesaplanmış ve rastgelelik durumu Şekil 3'te gösterilmiştir.



Şekil 3. Kriptografik uygulamalara yönelik rastgele ve sözde rastgele sayı üreteçleri için NIST istatistiksel test paketi sonuçları

Şekil 3'te de gösterildiği gibi NIST istatistik testi kullanılarak QRNG çıktısına rastgelelik testi yapıldı. İstatistik testi, boş hipotezin aksine kanıtın gücünü özetleyen P değerinin değerlerini belirlemek için kullanılır. P değeri 1'e eşitse ikili dizinin mükemmel bir rastgeleliğe sahip olduğu kabul edilir. Testin sonuçlarına göre, P değerinin bir test haricinde (*Test For The Longest Run Of Ones In A Block*) tüm değerlerinin 0,01'den büyük olduğunu göstermektedir; bu, QRNG tasarımının çıktı dizilerinin %99 güven düzeyiyle rastgele olduğunu göstermektedir.

5. SONUÇ

Sonuç olarak rastgele sayı üretimi, bilgisayar bilimleri açısından büyük bir öneme sahiptir. Özellikle kriptoloji alanında hayati bir rol oynamaktadır. Şifreleme anahtarlarının üretimi gibi kritik işlemler, büyük ölçüde rastgele sayıların kullanımına dayanır. Bu nedenle, rastgele sayı üretimi titizlikle ele alınmalı ve kesinlikle rastlantıya bırakılmamalıdır. Rastgele sayı üretiminin güvenilir ve etkili bir şekilde gerçekleştirilmesi, hem veri güvenliği hem de sistem bütünlüğü açısından vazgeçilmezdir. Bu bağlamda, rastgele sayı üreteçlerinin tasarımı ve uygulanması dikkatle incelenmeli ve optimize edilmelidir.

Bu çalışmada fotonik tabanlı bir rastgele sayı üreteci olan kuantum rastgele sayı üretecinin (QRNG) tasarımı dizayn edilmiştir. Bunun için Bulut temelli çalışma yapabilen Cirq kütüphanesi kullanılarak birçok operatörden oluşan kuantum devresi üzerinden 10 bitlik rastsal sayı kümesi oluşturulmuştur. Yapılan bağımsız testler sonucunda sayı setinin rastsallığı ispat edilmiştir.

Gerçekleştirilen simülasyonun kapsamı ve genişliği daha da artırılabilir. Rastsal sayı üretimi için Kuantum teknolojilerinin kullanımı giderek farklı tür ve yöntemlerle alanda kendine yer bulmaktadır. Bu nedenle alanda çalışmak isteyen araştırmacılar için bu çalışmanın başlangıç noktası olması çalışmanın temel amacı olarak belirlenmiştir. Her şeyin giderek dijitalleştiği günümüz göz önüne alındığında bu konuda daha fazla eser üretimi gerek ülkemiz gerekse de dünyamız için önem arz etmektedir.

KAYNAKÇA

Ananthaswamy, A. (2019, 19 Haziran). How to turn a quantum computer into the ultimate randomness generator. *Quanta Magazine*. https://www.quantamagazine.org/how-to-turn-a-quantum-computer-into-the-ultimate-randomness-generator-20190619 adresinden erişildi.

Arteaga, O., Garcia-Caurel, E., & Ossikovski, R. (2019). Stern-Gerlach experiment with light: Separating photons by spin with the method of A. Fresnel. *Optics Express, 27*(4), 4758–4768. https://doi.org/10.1364/OE.27.004758

Bassham, L. E., Rukhin, A. L., Soto, J., Nechvatal, J. R., Smid, M. E., Leigh, S. D., . . . Banks, D. L. (2010). *A statistical test suite for random and pseudorandom number generators for cryptographic applications* (NIST Special Publication 800-22 Rev. 1a). Gaithersburg, MD: National Institute of Standards and Technology. https://www.nist.gov/publications/statistical-test-suite-random-and-pseudorandom-number-generators-cryptographic adresinden erişildi.

Boyar, H. (2023, 19 Temmuz). *Kuantum rastsal sayı üreteçlerine giriş*. QTurkey. https://kuantumturkiye.org/kuantum-rastsal-sayi-ureteclerine-giris/ adresinden erişildi.

Cirq. (t.y.). Google Quantum AI. 21 Haziran 2024 tarihinde https://quantumai.google/cirq adresinden erişildi.

Computer Security Division, Information Technology Laboratory. (2016, 24 Mayıs). *Guide to the statistical tests: Random bit generation*. CSRC NIST. https://csrc.nist.gov/Projects/random-bit-generation/Documentation-and-Software/Guide-to-the-Statistical-Tests adresinden erişildi.

Ergen, S. (2020, 4 Haziran). *Rastgele sayılar üzerine*. HATSAT. https://suleymanergen.com/genel/rastgele-sayilar-uzerine adresinden erişildi.

Fürst, H., Weier, H., Nauerth, S., Marangon, D. G., Kurtsiefer, C., & Weinfurter, H. (2010). High speed optical quantum random number generation. *Optics Express, 18*(12), 13029–13037. https://doi.org/10.1364/OE.18.013029

Guillan-Lorenzo, O., Troncoso-Costas, M., Alvarez-Outarelo, D., Diaz-Otero, F. J., & Garcia-Escartin, J. C. (2023). Optical quantum random number generators: A comparative study. *Optical and Quantum Electronics, 55*(2). https://doi.org/10.1007/s11082-022-04396-y

Hanköylü, S. (2019). *Kuantum rastgele sayı üreteci tasarımı ve uygulaması* (Tez). Hacettepe Üniversitesi, Sağlık Bilimleri Enstitüsü, Ankara.

Hidary, J. D. (2019). *Quantum computing: An applied approach.* Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-030-23922-0

Iavich, M., Kuchukhidze, T., Gnatyuk, S., & Fesenko, A. (2021). Novel certification method for quantum random number generators. *International Journal of Computer Network and Information Security, 13*(3), 28–38. https://doi.org/10.5815/ijcnis.2021.03.03

Jacak, M. M., Jóźwiak, P., Niemczuk, J., & Jacak, J. E. (2021). Quantum generators of random numbers. *Scientific Reports, 11*(1), 16108. https://doi.org/10.1038/s41598-021-95388-7

James, F. (1990). A review of pseudorandom number generators. *Computer Physics Communications, 60*(3), 329–344. https://doi.org/10.1016/0010-4655(90)90032-V

Kofler, J., Paterek, T., & Brukner, Č. (2006). Experimenter's freedom in Bell's theorem and quantum cryptography. *Physical Review A, 73*(2), 022104. https://doi.org/10.1103/PhysRevA.73.022104

Landau, L. D., & Lifshitz, E. M. (2007). *Quantum mechanics: Non-relativistic theory* (3rd ed., rev. and enl.). Amsterdam: Elsevier.

Lee, J. S., & Cleaver, G. B. (2017). The cosmic microwave background radiation power spectrum as a random bit generator for symmetric- and asymmetric-key cryptography. *Heliyon, 3*(10), e00422. https://doi.org/10.1016/j.heliyon.2017.e00422

Ma, X., Yuan, X., Cao, Z., Qi, B., & Zhang, Z. (2016). Quantum random number generation. *npj Quantum Information, 2*, 16021. https://doi.org/10.1038/npjqi.2016.21

Mannalath, V., Mishra, S., & Pathak, A. (2023). A comprehensive review of quantum random number generators: Concepts, classification and the origin of randomness. *Quantum Information Processing, 22*(12), 439. https://doi.org/10.1007/s11128-023-04175-y

Murry, H. F. (1970). A general approach for generating natural random variables. *IEEE Transactions on Computers, C-19*(12), 1210–1213. https://doi.org/10.1109/T-C.1970.222860

Ren, M., Wu, E., Liang, Y., Jian, Y., Wu, G., & Zeng, H. (2011). Quantum random-number generator based on a photon-number-resolving detector. *Physical Review A, 83*(2), 023820. https://doi.org/10.1103/PhysRevA.83.023820

Rukhin, A., Soto, J., Nechvatal, J., Barker, E., Leigh, S., Levenson, M., . . . Dray, J. (2010). *A statistical test suite for random and pseudorandom number generators for*

cryptographic applications (NIST Special Publication 800-22 Revision 1a). Gaithersburg, MD: National Institute of Standards and Technology.

Sakurai, J. J. (1994). *Modern quantum mechanics* (Rev. ed.). Reading, MA: Addison-Wesley.

Siswanto, M., & Rudiyanto, B. (2017). Designing of quantum random number generator (QRNG) for security application. *2017 3rd International Conference on Science in Information Technology (ICSITech)* içinde (ss. 273–277). IEEE. https://doi.org/10.1109/ICSITech.2017.8257124

Stipčević, M., & Koç, Ç. K. (2014). True random number generators. Ç. K. Koç (Ed.), *Open problems in mathematics and computational science* içinde (ss. 275–315). Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-319-10683-0_12

Stojanovski, T., Pihl, J., & Kocarev, L. (2001). Chaos-based random number generators. Part II: Practical realization. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications, 48*(3), 382–385. https://doi.org/10.1109/81.915396

Truong, N. D., Haw, J. Y., Assad, S. M., Lam, P. K., & Kavehei, O. (2019). Machine learning cryptanalysis of a quantum random number generator. *IEEE Transactions on Information Forensics and Security, 14*(2), 403–414. https://doi.org/10.1109/TIFS.2018.2850770

Weihs, G., Jennewein, T., Simon, C., Weinfurter, H., & Zeilinger, A. (1998). Violation of Bell's inequality under strict Einstein locality conditions. *Physical Review Letters, 81*(23), 5039–5043. https://doi.org/10.1103/PhysRevLett.81.5039

Zhun, H., & Hongyi, C. (2001). A truly random number generator based on thermal noise. *ASICON 2001: 4th International Conference on ASIC Proceedings* içinde (ss. 862–864). IEEE. https://doi.org/10.1109/ICASIC.2001.982700

VERİDEN KARARA

Nicel Yöntemler ve Analitik Uygulamalar

yaz
yayınları

YAZ Yayınları
M.İhtisas OSB Mah. 4A Cad. No:3/3
İscehisar / AFYONKARAHİSAR
Tel : (0 531) 880 92 99
yazyayinlari@gmail.com • www.yazyayinlari.com

ISBN: 978-625-8926-58-3

