


Google Trends Arama Verilerinin Türkiye'ye Yönelik Uluslararası Turizm Talebinin Nowcasting'inde Kullanımı: SARIMAX Yaklaşımı

Using Google Trends Search Data for Nowcasting International Tourism Demand for Türkiye: A SARIMAX Approach

Abdulhalik PINAR^a 

^a Arş. Gör. Dr., Harran Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, Yönetim Bilişim Sistemleri Bölümü, Şanlıurfa, Türkiye. ORCID: 0000-0002-1716-5114

Özet

Turizm istatistiklerinin resmî olarak yayımlanması belirli bir gecikmeyle gerçekleşmekte, bu durum sektördeki karar alıcıların güncel talebi zamanında değerlendirmesini güçleştirmektedir. Bu çalışma, Google Trends arama hacmi verilerinin Türkiye'ye yönelik aylık uluslararası ziyaretçi sayısının resmî istatistikler açıklanmadan tahmin edilmesinde (nowcasting) öncü gösterge olarak kullanılıp kullanılamayacağını incelemektedir. Türkiye İstatistik Kurumu'nun 2012 Ocak ile 2025 Aralık dönemine ait 165 aylık ziyaretçi serisi ile Google Trends üzerinden elde edilen üç turizm temelli arama terimi ("turkey holiday", "Antalya", "İstanbul") birleştirilmiş; mevsimsel otoregresif bütünleşik hareketli ortalama (SARIMA) temel modeli ile arama verisini dışsal değişken olarak içeren SARIMAX modeli karşılaştırılmıştır. Analizler Python ortamında yürütülmüş, modeller MAPE, RMSE ve AIC ölçütleriyle değerlendirilmiştir. Bulgular, "Antalya" arama teriminin ziyaretçi sayısı ile en güçlü ilişkiyi gösterdiğini ($r = 0,63$) ve aramaların ziyaretçi hareketini yaklaşık iki ay önceden sinyallediğini ortaya koymaktadır. Arama verisinin modele eklenmesi test dönemi tahmin hatasını belirgin biçimde düşürmüş, ortalama mutlak yüzde hatası %11,4'ten %6,5'e gerilemiştir. Sonuçlar, Google Trends verisinin Türkiye turizmde gerçek zamanlı talep izleme ve erken uyarı amacıyla pratik bir araç olarak değerlendirilebileceğine işaret etmektedir.

Anahtar Kelimeler: Nowcasting, Google Trends, Turizm Talebi Tahmini, SARIMAX, Türkiye Turizmi

Abstract

The official release of tourism statistics occurs with a considerable delay, which makes it difficult for decision makers in the sector to assess current demand in a timely manner. This study examines whether Google Trends search volume data can serve as a leading indicator for nowcasting the monthly number of international visitors to Türkiye before official statistics are published. A monthly visitor series of 165 observations covering January 2012 to December 2025, obtained from the Turkish Statistical Institute, was combined with three tourism related search terms retrieved from Google Trends ("turkey holiday", "Antalya", "İstanbul"). A baseline seasonal autoregressive integrated moving average (SARIMA) model was compared with a SARIMAX model that incorporates search data as an exogenous variable. The analyses were conducted in Python, and the models were evaluated using MAPE, RMSE, and AIC. The findings show that the search term "Antalya" exhibits the strongest association with visitor numbers ($r = 0.63$) and that searches signal visitor movement approximately two months in advance. Adding search data to the model substantially reduced the test period forecast error, with the mean absolute percentage error falling from 11.4% to 6.5%. The results indicate that Google Trends data can be considered a practical tool for real time demand monitoring and early warning in Turkish tourism.

Keywords: Nowcasting, Google Trends, Tourism Demand Forecasting, SARIMAX, Türkiye Tourism

Makale Bilgisi

Geliş Tarihi 02.06.2026
Kabul Tarihi 14.06.2026

Sayı Editörü

Sağbetullah MERİÇ

Sorumlu Yazar

Abdulhalik PINAR
abdulhalik.pinar@harran.edu.tr

Önerilen Atf:

Pınar, A., (2026). Google Trends Arama Verilerinin Türkiye'ye Yönelik Uluslararası Turizm Talebinin Nowcasting'inde Kullanımı: SARIMAX Yaklaşımı. *Journal of Academic Tourism Studies*, 7(1): 109-120.

1. GİRİŞ

Turizm, Türkiye ekonomisinin önemli gelir kaynaklarından biridir ve döviz girdisi, istihdam ve bölgesel kalkınma açısından belirleyici bir rol oynamaktadır. Türkiye İstatistik Kurumu verilerine (TÜİK, 2026) göre ülkeye gelen uluslararası ziyaretçi sayısı 2012 yılında yaklaşık 36,5 milyon iken 2019 yılında 51,9 milyona ulaşmış, COVID-19 salgınının etkisiyle 2020 yılında 15,8 milyona gerilemiş, ardından hızla toparlanarak 2024 yılında 62,2 milyona çıkmıştır (TÜİK, 2026). Bu ölçekteki bir sektörün etkin biçimde yönetilmesi, talebin doğru ve zamanında tahmin edilmesine bağlıdır.

Turizm talebi tahmini uzun süredir araştırmacıların ilgisini çeken bir alandır; ancak geleneksel istatistiklerin önemli bir kısıtı, gecikmeli olarak yayımlanmasıdır. Resmî veriler genellikle ilgili dönemin bitiminden haftalar sonra erişilebilir hale gelir. Bu gecikme, kapasite planlaması, fiyatlandırma ve tanıtım gibi kararların güncel olmayan bilgilere dayanmasına yol açar. Choi ve Varian (2012), bu soruna yönelik olarak arama motoru verilerinin "şimdi tahmin etme" (nowcasting) amacıyla kullanılabilirliğini öne sürmüştür ve Google arama verilerinin çeşitli ekonomik göstergelerle eş zamanlı biçimde hareket ettiğini göstermiştir.

Turistler bir seyahate çıkmadan önce destinasyon, konaklama ve ulaşım hakkında yoğun biçimde çevrimiçi araştırma yapar. Bu davranış, arama hacmi verilerini turizm talebinin doğal bir öncü göstergesi haline getirir; aramalar, gerçekleşen ziyaretten haftalar hatta aylar önce yoğunlaşır (Siliverstovs ve Wochner, 2018). Bu nedenle Google Trends verisi, turizm talebi tahmininde giderek daha fazla kullanılan bir kaynak olmuştur (Dinis vd., 2019).

Uluslararası literatürde Google Trends verisinin turizm tahminine katkısı çeşitli destinasyonlar için belgelenmiştir. Sun vd. (2025), Singapur'a yönelik ziyaretçi sayısını arama verisi ve havalimanı yolcu istatistikleriyle birlikte modelleyen bir SARIMAX yaklaşımının, SARIMA, Holt-Winters ve Prophet gibi tek değişkenli modellerin yanı sıra LSTM gibi makine öğrenmesi yöntemlerini de geçtiğini göstermiştir. Botha ve Saayman (2024), COVID-19 sonrası toparlanma döneminde Google Trends içeren modellerin başarı oranının belirgin biçimde arttığını ortaya koymuştur. Höpken vd.(2019) ise birden çok arama teriminin dikkatli biçimde seçilip birleştirilmesinin tahmin doğruluğunu artırdığını göstermiştir.

Türkiye bağlamında ise çalışmalar görece sınırlıdır. Öncel vd. (2022), Rusya'dan Antalya'ya gelen ziyaretçileri "Antalya" arama terimiyle modellemiş ve Google Trends verisinin ARIMAX modelinde en güçlü açıklayıcı değişken olduğunu, ancak ham verinin yıllar arası karşılaştırılabilirlik açısından düzeltilmesi gerektiğini belirtmiştir. Çiçekdağı (2021), Türkiye'deki UNESCO Dünya Mirası alanlarına yönelik aramaları trend analizi ile incelemiştir. Önder (2017) ise Google Trends tabanlı tahminlerin ülke ve şehir düzeyinde farklı performans sergilediğini ortaya koymuştur. Bununla birlikte, ulusal düzeyde toplam uluslararası ziyaretçi sayısını güncel veriyle (2020 sonrası dönem dahil) ele alan ve arama terimi seçiminin tahmin başarısına etkisini sistematik biçimde inceleyen bir çalışmaya rastlanmamıştır.

Bu çalışma söz konusu boşluğu kapatmayı amaçlamakta ve şu sorulara yanıt aramaktadır: (1) Google Trends arama hacmi, Türkiye'ye yönelik toplam uluslararası ziyaretçi sayısının öncü göstergesi olarak anlamlı bir ilişki taşımakta mıdır? (2) Hangi arama terimi en güçlü öncü sinyali sağlamaktadır ve bu sinyal kaç ay önceden ortaya çıkmaktadır? (3) Arama verisini dışsal değişken olarak içeren SARIMAX modeli, temel SARIMA modeline kıyasla tahmin başarısını anlamlı biçimde artırmakta mıdır? Çalışmanın katkısı, Türkiye'nin toplam uluslararası turizm talebini 2012-2025 dönemini kapsayan güncel ve kesintisiz bir veriyle ele alması ve arama terimi seçiminin tahmin başarısı üzerindeki belirleyici etkisini ampirik olarak göstermesidir.

2. KAVRAMSAL ÇERÇEVE VE İLGİLİ ÇALIŞMALAR

2.1. Turizm Talebi Tahmini ve Büyük Veri

Turizm talebi tahmini literatürü, klasik zaman serisi modellerinden makine öğrenmesi ve derin öğrenme yöntemlerine uzanan geniş bir yelpazeye sahiptir. Wu vd. (2025), büyük veri tabanlı turizm tahmin çalışmalarını web arama hacmi verileri, sosyal medya istatistikleri, metinsel veriler, fotoğraf verileri ve video verileri olmak üzere beş kategoride toplamıştır. Bu kategoriler arasında web arama verisi, erişilebilirliği ve zamansal sürekliliği nedeniyle en yaygın kullanılan kaynaktır. Geleneksel istatistiksel verinin aksine, arama verisi neredeyse gerçek zamanlı olarak elde edilebilir ve bu özelliği onu nowcasting uygulamaları için elverişli kılar.

Nowcasting kavramı, geleceğe yönelik klasik öngörüden (forecasting) farklı olarak, henüz resmî verisi açıklanmamış olan mevcut dönemin tahminine odaklanır (Choi ve Varian, 2012). Minora vd. (2023), Avrupa'daki on bir ülkede bölgesel düzeyde geceleme sayısını insan hareketliliği verisi kullanarak yüksek doğrulukla tahmin etmiş ve alternatif veri kaynaklarının resmî turizm istatistiklerinin üretim gecikmesini telafi edebileceğini göstermiştir.

2.2. Google Trends ve Turizm Talebi

Google Trends verisinin turizm tahminindeki kullanımı, Choi ve Varian'ın (2012) öncü çalışmasının ardından hızla yaygınlaşmıştır. Bu yaklaşım, arama hacminin destinasyon talebinin bir vekil değişkeni olarak hareket ettiği varsayımına dayanır. Gunter ve Önder (2016), Viyana destinasyonu için Google Trends endekslerinin tahmin doğruluğunu yalnızca otoregresif bir temel modele kıyasla artırdığını göstermiştir. Bangwayo-Skeete ve Skeete (2015), karışık veri örnekleme (MIDAS) yaklaşımıyla arama verisinin ziyaretçi tahminine katkısını ortaya koymuştur.

Arama verisinin kullanımında iki önemli yönetsel mesele öne çıkar. Birincisi arama terimi seçimidir; ilgili ve güçlü sinyal taşıyan terimlerin belirlenmesi tahmin başarısını doğrudan etkiler (Höpken vd., 2019). İkincisi, birden çok terimin nasıl birleştirileceğidir. Siliverstovs ve Wochner (2018), tek bir bileşik endekse dayanan boylamsal yaklaşımın bazı ölçüm sorunları taşıdığını, bunun yerine birden çok seriyi karşılaştıran kesitsel yaklaşımın daha sağlam sonuçlar verebileceğini İsviçre turizmi örneğinde göstermiştir. Song ve diğerlerinin değerlendirmesinde de belirtildiği üzere, arama sorgu verisi zengin bilgi taşısa da dikkatli ve eleştirel biçimde modellenmediğinde geçersiz sonuçlara yol açabilir (Lee, 2025).

Ham Google Trends verisinin önemli bir kısıtı, sıfır ile yüz arasında görel olarak normalize edilmiş olmasıdır. Bu normalizasyon, internet kullanımının zaman içindeki büyümesini gizleyebileceği gibi, bir terimin mutlak arama hacmi artsa bile diğer terimlerin görece popülerliği nedeniyle endeksin düşmesine yol açabilir (Siliverstovs ve Wochner, 2018). Öncel vd. (2022) bu sorunu gidermek için ham veriyi internet kullanıcı sayısı ile çarparak düzeltmeyi önermiştir.

2.3. COVID-19 ve Model Performansı

COVID-19 salgını, turizm zaman serilerinde keskin bir yapısal kırılmaya yol açmış ve geleneksel tahmin modellerinin başarısını zorlaştırmıştır. Botha ve Saayman (2024), Güney Afrika örneğinde Google Trends içeren modellerin salgın öncesi dönemde modellerin yaklaşık %40'ında en iyi performansı sergilediğini, salgın sonrası toparlanma döneminde ise bu oranın %77'ye yükseldiğini bulmuştur. Bu bulgu, kriz ve toparlanma dönemlerinde arama verisinin öncü gücünün arttığına işaret etmektedir. Wu vd. (2023) de çok kaynaklı heterojen verinin salgın döneminde ziyaretçi tahmin başarısını artırabileceğini göstermiştir.

3. VERİ VE YÖNTEM

3.1. Veri Kaynakları

Çalışmanın bağımlı değişkeni, Türkiye'ye gelen aylık toplam uluslararası ziyaretçi sayısıdır. Bu veri, Türkiye İstatistik Kurumu'nun turizm istatistikleri kapsamında yayımladığı ve ziyaretçilerin aylara göre turizm geliri ile kişi sayısını içeren resmî tablodan elde edilmiştir (TÜİK, 2026). Veri seti 2012 Ocak ile 2025 Aralık dönemini kapsamakta olup toplam 165 aylık gözlemden oluşmaktadır. Seride hiçbir eksik gözlem bulunmamaktadır.

Çalışmanın bağımsız değişkeni, Google Trends platformundan elde edilen arama hacmi verisidir (Google, 2026). Türkiye'ye yönelik turizm talebini yansıtabileceği değerlendirilen üç arama terimi seçilmiştir: "turkey holiday", "Antalya" ve "Istanbul". Bu terimler, sırasıyla genel tatil niyetini, kıyı turizmi talebini ve kültür-şehir turizmi talebini temsil etmektedir. Bunun terimleri seçmemizin gerekçesi vardır. Birincisi, "Antalya" ve "Istanbul" gibi özel adlar dilden bağımsız olup farklı kaynak ülkelerdeki kullanıcılar tarafından büyük ölçüde aynı biçimde aranmaktadır; nitekim Öncel Çekim ve Koyuncu (2022) da dil yanlılığını azaltmak amacıyla bilinçli olarak tek bir özel ad ("Antalya") tercih etmiştir. İkincisi, çalışma ulusal düzeyde toplam talebi modellediğinden, tek bir kaynak ülkeye özgü dildeki aramalar yerine, dünya geneli kapsamda toplulaştırılmış arama hacmi kullanılmıştır. Veriler, Google Trends'in resmi arayüzü üzerinden, dünya geneli (worldwide) coğrafi kapsam ve 2012 Ocak ile 2025 Aralık zaman aralığı seçilerek aylık frekansta indirilmiştir. Google Trends verisinin dünya geneli kapsamda alınmasının nedeni, Türkiye'ye gelen ziyaretçilerin yurt dışında ikamet eden bireyler olması ve aramalarını kendi ülkelerinde gerçekleştirmesidir.

3.2. Verinin İşlenmesi

Tüm veri işleme ve modelleme adımları Python programlama dili kullanılarak gerçekleştirilmiştir. TÜİK'in Excel biçimindeki tablosu xlr d kütüphanesi ile okunmuş; yıl ve ay bilgileri ayrıştırılarak tarih indeksli bir zaman serisine dönüştürülmüştür. Google Trends'ten indirilen üç ayrı CSV dosyası pandas kütüphanesiyle okunmuş, başlık satırları temizlenmiş ve ortak tarih aralığında hizalanmıştır.

Ham Google Trends serileri sıfır ile yüz arasında farklı ölçeklerde olduğundan, üç terim scikit-learn kütüphanesinin MinMaxScaler işleviyle ortak bir ölçeğe getirilmiş ve aritmetik ortalamaları alınarak bir bileşik arama endeksi oluşturulmuştur. Ancak ön analizlerde bileşik endeksin tek başına "Antalya" teriminden daha zayıf bir ilişki gösterdiği belirlenmiş; bu nedenle Siliverstovs ve Wochner'in (2018) tek terim yaklaşımına ilişkin değerlendirmeleri doğrultusunda, nihai modelde en güçlü sinyali taşıyan "Antalya" terimi dışsal değişken olarak kullanılmıştır.

Öncel Çekim ve Koyuncu (2022), ham Google Trends verisinin yıllar arası karşılaştırılabilirliğini artırmak için internet kullanıcı sayısı ile düzeltilmesini önermiştir. Bu çalışmada söz konusu düzeltme uygulanmamıştır; bunun nedeni, analiz döneminin (2012-2025) büyük bölümünde Türkiye'ye turist gönderen başlıca ülkelerde internet penetrasyonunun zaten yüksek düzeyde ve görece istikrarlı olması, dolayısıyla düzeltmenin sağlayacağı katkının sınırlı kalacağını değerlendirilmesidir. Ayrıca SARIMAX modelinde uygulanan birinci dereceden fark alma işlemi, serideki yavaş seyreden trend bileşenlerini (internet kullanımının kademeli artışı dahil) önemli ölçüde gidermektedir. Yine de bu düzeltmenin uygulanması, özellikle daha uzun veya internet kullanımının hızla değiştiği dönemleri kapsayan çalışmalarda yararlı olabilir.

Ziyaretçi serisi, varyansı dengelemek amacıyla logaritmik dönüşüme tabi tutulmuştur. Serinin durağanlığı genişletilmiş Dickey-Fuller (ADF) testiyle sınanmış; mevsimsel ve birinci dereceden fark alındıktan sonra serinin durağan hale geldiği belirlenmiştir (ADF $p < 0,001$). COVID-19 salgınının yarattığı yapısal kırılmayı kontrol altına almak için 2020 Mart ile 2021 Haziran arası kapsayan bir kukla (dummy) değişken modele dahil edilmiştir.

3.3. Gecikme Analizi

Google Trends aramalarının ziyaretçi hareketini kaç ay önceden sinyallediğini belirlemek için, arama serisi ile logaritmik ziyaretçi serisi arasında çapraz korelasyon fonksiyonu (CCF) hesaplanmıştır. Bir ile altı ay arasındaki gecikmeler için korelasyon katsayıları karşılaştırılmış ve en yüksek korelasyona karşılık gelen gecikme, optimal öncü süre olarak belirlenmiştir. Bu analiz, statsmodels kütüphanesinin ccf işlevi kullanılarak yapılmıştır.

3.4. Modelleme ve Değerlendirme

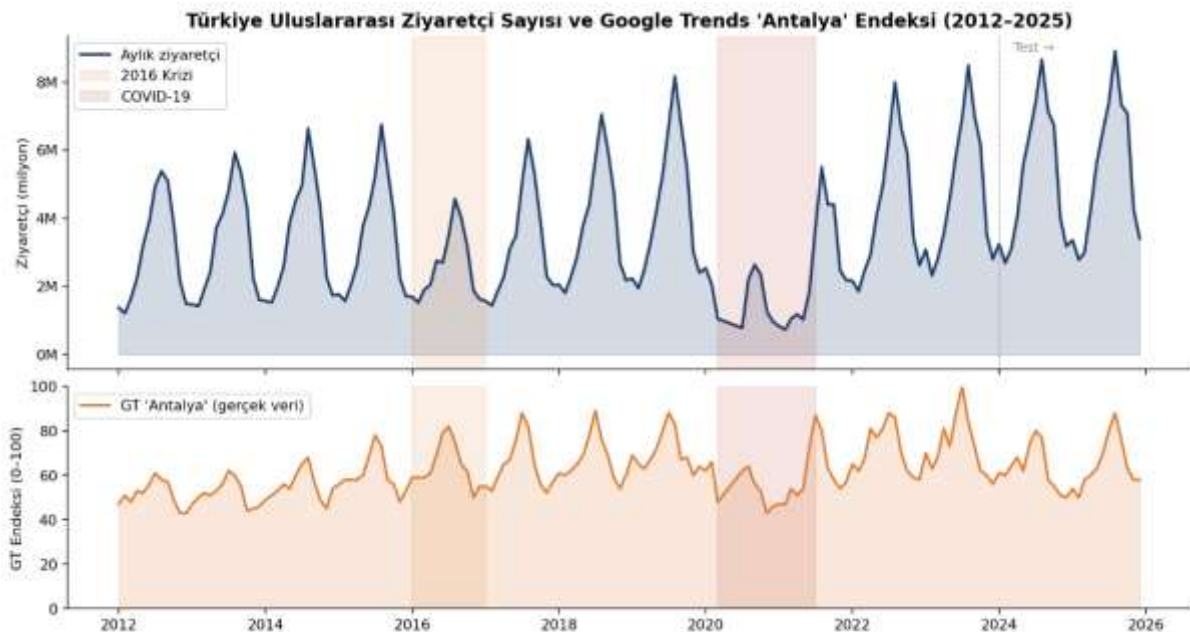
Çalışmada iki model kurulmuştur. Temel model, dışsal arama değişkeni içermeyen mevsimsel SARIMA modelidir; yalnızca geçmiş ziyaretçi serisini ve COVID-19 kukla değişkenini kullanır. İkinci model olan SARIMAX, bu yapıya iki ay gecikmeli "Antalya" arama endeksini dışsal değişken olarak ekler. Her iki model de (1,1,1) sıradan ve (1,1,1,12) mevsimsel parametre yapısıyla, statsmodels kütüphanesinin SARIMAX sınıfı kullanılarak tahmin edilmiştir. Parametre yapısının belirlenmesinde otokorelasyon (ACF) ve kısmi otokorelasyon (PACF) grafiklerinden yararlanılmıştır.

Veri seti, 2012 Ocak ile 2023 Aralık arasını kapsayan 141 aylık eğitim dönemi ile 2024 Ocak ile 2025 Aralık arasını kapsayan 24 aylık test dönemine ayrılmıştır. Modellerin test dönemindeki tahmin başarısı üç ölçütle değerlendirilmiştir: ortalama mutlak yüzde hatası (MAPE), kök ortalama kare hatası (RMSE) ve Akaike bilgi ölçütü (AIC). MAPE ve RMSE hesaplamaları scikit-learn kütüphanesinin ilgili işlevleriyle yapılmıştır. Logaritmik ölçekte üretilen tahminler, değerlendirme öncesinde üstel dönüşümle özgün ölçeğe geri çevrilmiştir.

4. BULGULAR

4.1. Türkiye Turizminin Genel Görünümü

Şekil 1, 2012-2025 döneminde Türkiye'ye gelen aylık uluslararası ziyaretçi sayısı ile bileşik Google Trends endeksinin birlikte seyrini göstermektedir. Ziyaretçi serisi belirgin bir mevsimsel örüntü sergilemekte; her yıl yaz aylarında doruğa ulaşmaktadır. Tüm dönem ortalamasında en yüksek ziyaretçi sayısı Ağustos ayında (6,61 milyon), en düşük sayı ise Şubat ayında (1,77 milyon) gerçekleşmektedir. Seride iki büyük kırılma dikkat çekmektedir: 2016 yılındaki siyasi gerilimlerin yol açtığı düşüş ve 2020 yılında COVID-19 salgınının neden olduğu keskin çöküş. Salgının en derin noktası olan 2021 Şubat ayında ziyaretçi sayısı yalnızca 727 bine kadar gerilemiştir.



Şekil 1. Türkiye'ye gelen aylık uluslararası ziyaretçi sayısı ve bileşik Google Trends endeksi (2012-2025).

Şekil 1, aylık ziyaretçi sayısını milyon kişi cinsinden, alt panel beş arama teriminden oluşturulan bileşik Google Trends endeksini göstermektedir. Turuncu alan 2016 krizini, kırmızı alan COVID-19 salgın dönemini işaret etmektedir. Her iki serinin de mevsimsel doruklarının yaz aylarında örtüştüğü, salgın döneminde ise eş zamanlı biçimde çöktüğü görülmektedir.

Yıllık toplamlar incelendiğinde, sektörün COVID-19 sonrası güçlü bir toparlanma sergilediği görülmektedir. 2020 yılında 15,8 milyona kadar gerileyen toplam ziyaretçi sayısı, 2022 yılında salgın öncesi düzeyi aşmış, 2024 yılında 62,2 milyona ve 2025 yılında 63,9 milyona ulaşmıştır.

4.2. Arama Terimlerinin Karşılaştırması

Şekil 2, üç arama teriminin ziyaretçi serisiyle ilişkisini ayrı ayrı göstermektedir. "Antalya" terimi, ziyaretçi sayısı ile en güçlü eş zamanlı korelasyonu sergilemektedir ($r = 0,63$). "Turkey holiday" terimi orta düzeyde bir ilişki gösterirken ($r = 0,31$), "İstanbul" terimi en zayıf ilişkiyi taşımaktadır ($r = 0,28$). Bu sonuç, Antalya'nın Türkiye'nin kıyı turizminin merkezi olması ve mevsimsel talebi en doğrudan yansıtan destinasyon olmasıyla tutarlıdır. İstanbul'a yönelik aramaların daha zayıf ilişki göstermesi, bu kentin turizm dışı amaçlarla da yoğun biçimde aranmasından kaynaklanıyor olabilir.



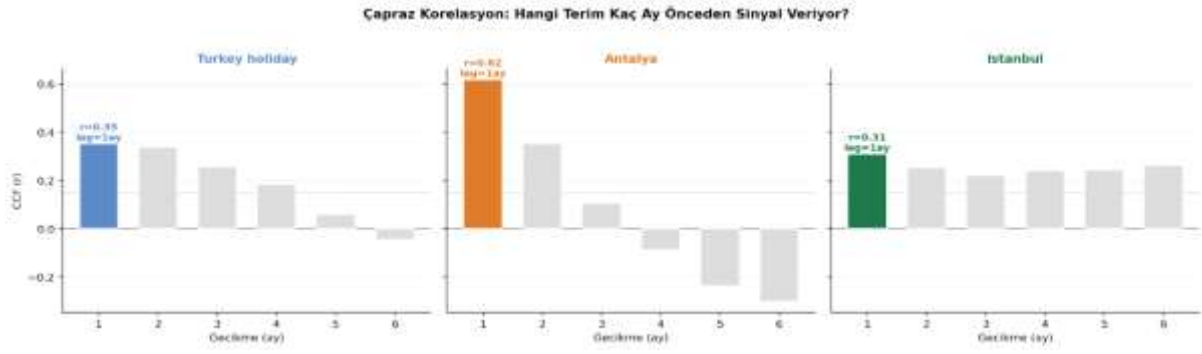
Şekil 2. Üç arama teriminin ziyaretçi sayısı ile karşılaştırması ve korelasyon katsayıları.

Açıklama. Her panelde gri çizgi aylık ziyaretçi sayısını, renkli çizgi ilgili Google Trends arama endeksini göstermektedir. Sağ alt köşedeki r değeri, iki seri arasındaki eş zamanlı Pearson korelasyon katsayısıdır. "Antalya" teriminin ziyaretçi serisiyle açık biçimde en güçlü ilişkiyi taşıdığı görülmektedir.

Üç terimin belirgin biçimde farklı korelasyon değerleri üretmesi, arama terimi seçiminin tahmin başarısı için tesadüfi bir ayrıntı değil, belirleyici bir karar olduğunu ortaya koymaktadır. "Antalya" teriminin öne çıkması ($r = 0,63$), Türkiye'ye gelen uluslararası talebin ağırlıklı olarak deniz-kum-güneş eksenli kıyı turizmine dayandığını ve bu talebin tek bir destinasyon adıyla şaşırtıcı ölçüde iyi temsil edilebildiğini göstermektedir. Buna karşılık "İstanbul" teriminin zayıf kalması ($r = 0,28$), bir arama teriminin popülerliğinin tek başına yeterli olmadığını gösteren öğretici bir bulgudur: İstanbul yoğun biçimde aranan bir terim olsa da bu aramaların önemli bölümü iş, eğitim, haber veya transit gibi turizm dışı amaçlardan kaynaklandığı için tatil talebinin gürültülü bir vekili haline gelmektedir. Bu karşılaştırma, Höpken vd.'nin (2019) terim seçiminin titizlikle yapılması gerektiği yönündeki vurgusunu ampirik olarak doğrulamaktadır.

4.3. Gecikme Analizi

Şekil 3, üç arama terimi için çapraz korelasyon analizinin sonuçlarını sunmaktadır. Tüm terimler için en yüksek korelasyon belirli bir gecikmede ortaya çıkmakta; bu da aramaların ziyaretçi hareketini önceden sinyallemediğini doğrulamaktadır. "Antalya" terimi iki aylık gecikmede en güçlü ilişkiyi göstermektedir. Bu bulgu, turistlerin Türkiye'ye gelmeden yaklaşık iki ay önce destinasyon araştırması yaptığı yönündeki seyahat planlama davranışıyla uyumludur ve Siliverstovs ve Wochner'in (2018) tatil planlamasının haftalar veya aylar öncesinden başladığı gözlemini desteklemektedir.



Şekil 3. Arama terimleri ile ziyaretçi sayısı arasındaki çapraz korelasyon analizi.

Açıklama. Her panel, bir arama terimi ile logaritmik ziyaretçi serisi arasındaki çapraz korelasyon katsayılarını bir ile altı ay arası gecikmeler için göstermektedir. Koyu renkli çubuk en yüksek korelasyona karşılık gelen optimal gecikmeyi, kesikli çizgi ise istatistiksel anlamlılık sınırını temsil etmektedir. "Antalya" terimi için optimal öncü süre iki aydır.

Korelasyonun sıfır gecikmede değil, belirli bir gecikmede tepe yapması, arama verisinin eş zamanlı bir gösterge olmanın ötesinde gerçek bir öncü gösterge işlevi taşıdığını kanıtlamaktadır; nowcasting iddiasının dayandığı temel bulgu budur. "Antalya" teriminin gecikmeli yapıda da en yüksek korelasyonu koruması, bu terimin yalnızca güçlü değil, aynı zamanda zamanlama açısından da en güvenilir sinyal olduğunu göstermektedir. Bu öncü sürenin varlığı, turistlerin seyahat kararı ile fiili varış arasında geçen planlama dönemini yansıtmakta ve uygulamada kritik bir fırsat penceresi sunmaktadır: arama verisi izlenerek, resmî istatistikler açıklanmadan önce talebin yönü hakkında öngörü elde edilebilmektedir. Diğer iki terimin daha düşük tepe değerleri, Şekil 2'deki korelasyon sıralamasının gecikmeli yapıda da korunduğunu, dolayısıyla terim seçiminin tutarlı biçimde belirleyici olduğunu doğrulamaktadır.

4.4. Model Karşılaştırması

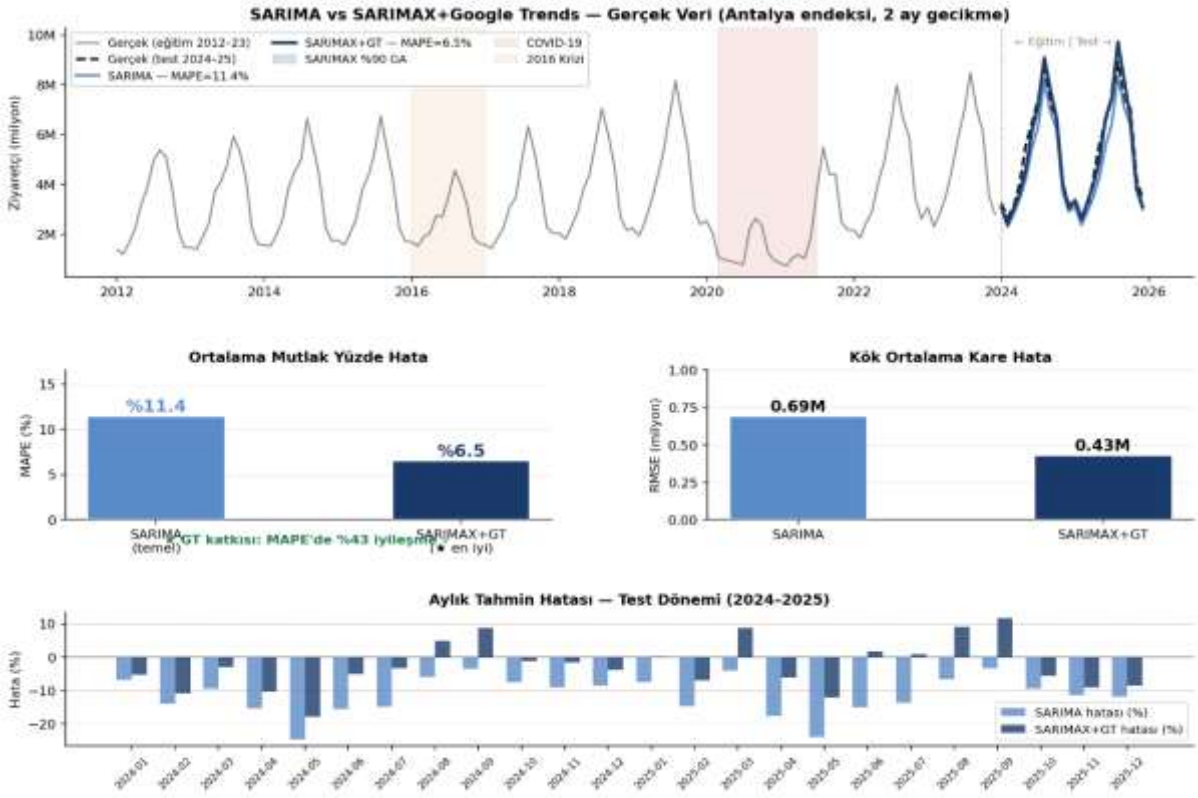
Tablo 1, temel SARIMA modeli ile arama verisini içeren SARIMAX modelinin test dönemi (2024-2025) tahmin başarısını özetlemektedir. SARIMAX modeli her üç ölçütte de temel modelden daha iyi sonuç vermiştir. Ortalama mutlak yüzde hatası %11,4'ten %6,5'e gerilemiş; bu, arama verisinin eklenmesiyle tahmin hatasında yaklaşık %43'lük bir iyileşme anlamına gelmektedir. Kök ortalama kare hatası ise yaklaşık %38 oranında azalmıştır.

Tablo 1. Model karşılaştırması: test dönemi (2024-2025) tahmin başarısı.

Model	MAPE (%)	RMSE	AIC
SARIMA (temel)	11,4	690.395	-50,32
SARIMAX + Google Trends	6,5	430.025	-50,53

MAPE ortalama mutlak yüzde hatasını, RMSE kök ortalama kare hatasını, AIC ise Akaike bilgi ölçütünü göstermektedir. Üç ölçütte de daha düşük değer daha iyi performansa işaret eder. SARIMAX modeli, dışsal değişken olarak iki ay gecikmeli "Antalya" arama endeksini kullanmaktadır. Değerler 2024 Ocak ile 2025 Aralık arasını kapsayan 24 aylık test dönemi üzerinden hesaplanmıştır.

Şekil 4, iki modelin tahminlerini gerçek değerlerle birlikte görselleştirmektedir. Üst panelde SARIMAX modelinin test dönemindeki gerçek ziyaretçi sayısını, özellikle yaz doruklarında, temel modele kıyasla daha yakından izlediği görülmektedir. Alt panellerdeki hata dağılımı, SARIMAX modelinin aylık tahmin hatalarının çoğunlukla daha dar bir bant içinde kaldığını göstermektedir.



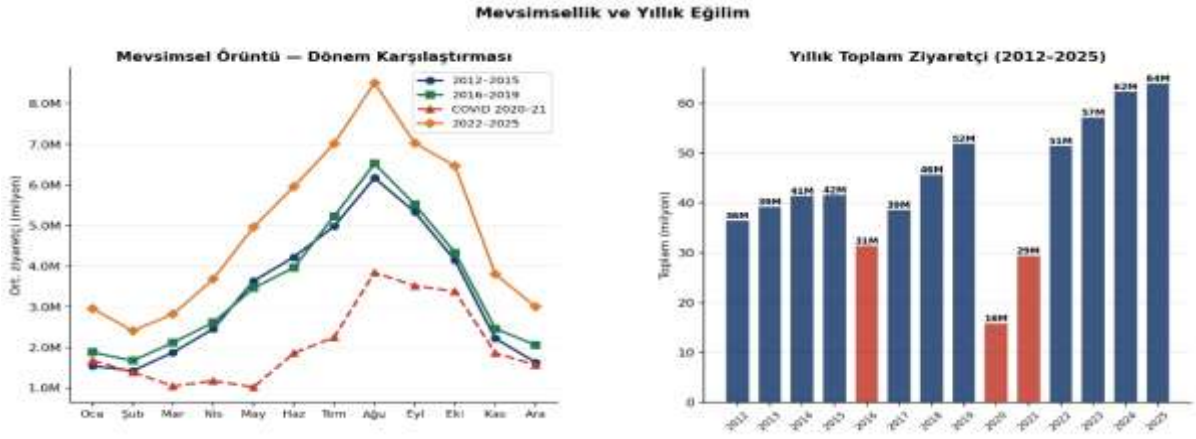
Şekil 4. SARIMA ve SARIMAX modellerinin gerçek değerlerle karşılaştırılması.

Üst panel tüm dönem boyunca gerçek ziyaretçi sayısını ve iki modelin test dönemi tahminlerini göstermektedir; gölgeli bant SARIMAX modelinin %90 güven aralığıdır. Sol alt panel ortalama mutlak yüzde hatasını, sağ alt panel kök ortalama kare hatasını karşılaştırmaktadır. En alttaki panel test dönemindeki aylık tahmin hatalarını yüzde cinsinden sunmaktadır. SARIMAX modelinin her iki hata ölçütünde de temel modelden üstün olduğu görülmektedir.

Tahmin hatasındaki yaklaşık %43'lük düşüş, arama verisinin modele yalnızca marjinal değil, kayda değer bir katkı sağladığını göstermektedir; bu, çalışmanın merkezi bulgusudur. Üst paneldeki tahmin eğrileri, iyileşmenin nerede gerçekleştiğini açığa çıkarmaktadır: temel SARIMA modeli mevsimsel doruklarda gerçek değerlerden belirgin biçimde sapmakta, SARIMAX modeli ise yaz zirvelerini daha doğru yakalamaktadır. Bu durum, arama verisinin özellikle talebin en yüksek ve yönetim açısından en kritik olduğu dönemlerde değer kattığını göstermesi bakımından önemlidir; çünkü kapasite ve fiyatlandırma kararlarının en çok önem taşıdığı dönemler tam da bu doruk aylardır. En alttaki hata panelinde SARIMAX hatalarının daha dar bir bant içinde kalması, modelin yalnızca ortalamada değil, ay bazında da daha istikrarlı tahmin ürettiğini ortaya koymaktadır. Elde edilen %6,5'lik MAPE değeri, Sun ve diğerlerinin (2025) Singapur için raporladığı %7,3 düzeyiyle karşılaştırılabilir olup yaklaşımın uluslararası ölçekte de geçerli olduğunu desteklemektedir.

4.5. Mevsimsellik ve Yıllık Eğilim

Şekil 5, ziyaretçi sayısının mevsimsel örüntüsünü farklı dönemler için ve yıllık toplamı karşılaştırmalı olarak sunmaktadır. Sol paneldeki mevsimsel eğriler, COVID-19 dönemi dışında tüm dönemlerde benzer bir biçim izlemekte; talep ilkbaharda yükselmeye başlayıp yaz aylarında doruğa ulaşmakta ve sonbaharda gerilemektedir. Salgın dönemine ait eğri ise belirgin biçimde basıktır. Sağ paneldeki yıllık toplamlar, 2016 ve 2020-2021 kırılmalarını ve sonraki güçlü toparlanmayı açıkça ortaya koymaktadır.



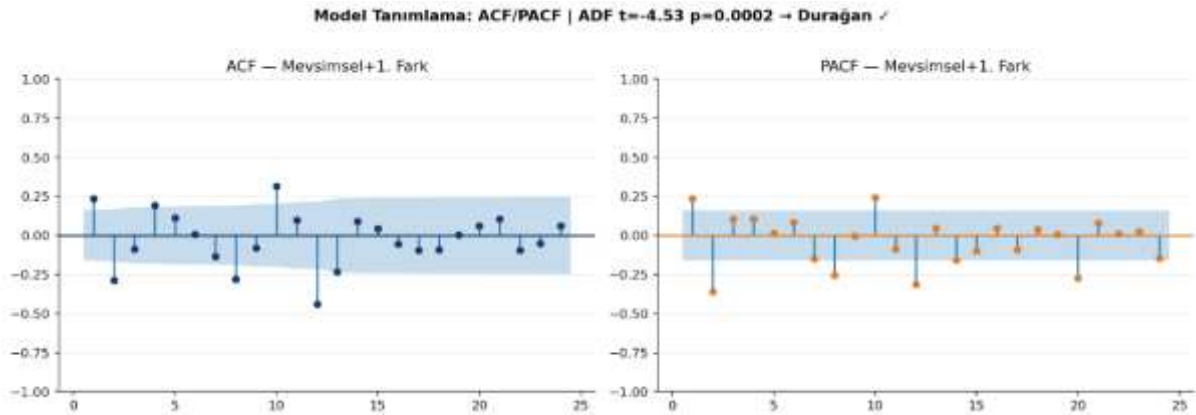
Şekil 5. Mevsimsel örüntü ve yıllık toplam ziyaretçi sayısı.

Sol panel, dört farklı dönem için aylık ortalama ziyaretçi sayısını göstermektedir; her eğri bir dönemi temsil eder. Sağ panel yıllık toplam ziyaretçi sayısını sunmakta, kırmızı çubuklar kriz yıllarını (2016, 2020, 2021) işaret etmektedir. Mevsimsel doruğun istikrarlı biçimde yaz aylarında yoğunlaştığı, COVID-19 döneminin ise bu örüntüyü geçici olarak bozduğu görülmektedir.

Mevsimsel eğrilerin COVID-19 dönemi dışında neredeyse üst üste binmesi, Türkiye turizm talebinin son derece istikrarlı ve öngörülebilir bir mevsimsel yapıya sahip olduğunu göstermektedir; bu istikrar, mevsimsel zaman serisi modellerinin bu veride neden başarılı olduğunu da açıklamaktadır. Salgın eğrisinin belirgin biçimde basıklaşması, COVID-19'un yalnızca talebin düzeyini değil, mevsimsel yapısını da geçici olarak bozduğunu ortaya koymakta ve bu dönemin modelde ayrı bir kukla değişkenle ele alınmasının gerekçesini görsel olarak haklı çıkarmaktadır. Sağ paneldeki yıllık toplamlar ise dikkat çekici bir toparlanma kapasitesine işaret etmektedir: hem 2016 krizinin hem de pandeminin yarattığı keskin düşüşlerin ardından talep yalnızca eski düzeyine dönmekle kalmamış, onu aşarak yeni zirvelere ulaşmıştır. Bu örüntü, Türkiye turizminin dış şoklara karşı kırılğan ancak güçlü bir biçimde dirençli olduğunu düşündürmektedir.

4.6. Model Tanımlama Sınamaları

Şekil 6, mevsimsel ve birinci dereceden fark alınmış serinin otokorelasyon ve kısmi otokorelasyon grafiklerini göstermektedir. Bu grafikler, SARIMA modelinin parametre yapısının belirlenmesinde kullanılmıştır. Genişletilmiş Dickey-Fuller testi, farklanmış serinin durağan olduğunu doğrulamaktadır ($p < 0,001$), bu da seçilen fark alma derecelerinin uygun olduğunu göstermektedir.



Şekil 6. Mevsimsel ve birinci dereceden farklanmış serinin ACF ve PACF grafikleri.

Sol panel otokorelasyon fonksiyonunu (ACF), sağ panel kısmi otokorelasyon fonksiyonunu (PACF) yirmi dört gecikmeye kadar göstermektedir. Gölge alan %95 güven aralığını temsil eder. Grafiklerdeki örüntü ve genişletilmiş Dickey-Fuller test sonucu ($p < 0,001$), serinin uygulanan dönüşümlerin ardından durağan hale geldiğini ve seçilen model yapısının uygun olduğunu doğrulamaktadır.

Bu grafikler ve eşlik eden durağanlık testi, modelin metodolojik geçerliliğini güvence altına almaktadır. Genişletilmiş Dickey-Fuller testinin güçlü sonucu ($p = 0,0002$), uygulanan mevsimsel ve birinci dereceden fark alma işlemlerinin serinin durağan olmama sorununu giderdiğini, dolayısıyla SARIMA/SARIMAX ailesinin ön koşulunun sağlandığını doğrulamaktadır. ACF ve PACF grafiklerinde anlamlılık sınırlarını aşan az sayıda gecikmenin bulunması, seçilen düşük dereceli (1,1,1)(1,1,1,12) parametre yapısının veriye uygun olduğunu ve modelin aşırı parametrelendirilmediğini göstermektedir. Böylece bu görsel, bulgular bölümünde sunulan tahmin sonuçlarının sağlam bir model spesifikasyonuna dayandığını ve elde edilen başarının model kurgusundaki bir hatadan değil, arama verisinin gerçek katkısından kaynaklandığını teyit etmektedir.

5. TARTIŞMA

Bu çalışmanın bulguları, Google Trends arama verisinin Türkiye'ye yönelik uluslararası turizm talebinin nowcasting'inde anlamlı bir öncü gösterge işlevi gördüğünü ortaya koymaktadır. Arama verisinin modele eklenmesiyle elde edilen %43'lük tahmin hatası iyileşmesi, uluslararası literatürdeki bulgularla uyumludur. Sun vd. (2025) Singapur için %7,3 düzeyinde bir MAPE değeri raporlamış; bu çalışmada elde edilen %6,5'lik değer benzer bir başarı düzeyine işaret etmektedir. Arama verisinin tahmin başarısını artırdığı yönündeki bu sonuç, Gunter ve Önder (2016), Bangwayo-Skeete ve Skeete (2015) ve Botha ve Saayman'ın (2024) bulgularıyla da örtüşmektedir.

Çalışmanın dikkat çekici bir bulgusu, arama terimi seçiminin tahmin başarısı üzerindeki belirleyici etkisidir. "Antalya" teriminin tek başına, üç terimin birleştirilmesiyle oluşturulan bileşik endeksten daha güçlü bir ilişki göstermesi, ilk bakışta sezgilere aykırı görünebilir. Ancak bu sonuç, Siliverstovs ve Wochner'in (2018) bileşik endekslerin ölçüm sorunları taşıyabileceği yönündeki uyarısıyla ve Lee'nin (2025) arama verisinin dikkatli modellenmesi gerektiği yönündeki değerlendirmesiyle tutarlıdır. "Turkey holiday" ve "Istanbul" terimlerinin taşıdığı gürültü, bileşik endekste "Antalya" teriminin güçlü sinyalini zayıflatmıştır. Bu bulgu, Öncel vd.'nin (2022) Antalya odaklı tek terim tercihinin isabetli olduğunu da desteklemektedir.

İki aylık optimal gecikme süresi, turistlerin seyahat kararı ile fiili ziyaret arasındaki planlama dönemini yansıtmaktadır. Bu süre, destinasyon araştırması, konaklama ve ulaşım rezervasyonu gibi hazırlık aşamalarının zamanlamasıyla tutarlıdır. Pratik açıdan bu gecikme, resmî istatistikler açıklanmadan yaklaşık iki ay önce talep yönü hakkında sinyal elde edilebileceği anlamına gelir.

Çalışmanın bazı sınırlılıkları bulunmaktadır. Birincisi, analiz ulusal düzeyde toplam ziyaretçi sayısı ile yürütülmüştür; destinasyon veya kaynak ülke bazında ayrıştırılmış bir analiz daha ayrıntılı bilgi sağlayabilirdi. İkincisi, ham Google Trends verisi görece olarak normalize edildiğinden, Google'ın algoritma değişikliklerinden kaynaklanan tutarsızlıklar tam olarak kontrol edilememektedir (Siliverstovs ve Wochner, 2018). Üçüncüsü, çalışmada klasik bir zaman serisi yöntemi olan SARIMAX kullanılmıştır; makine öğrenmesi ve derin öğrenme yöntemleriyle karşılaştırma yapılması gelecek çalışmalar için verimli bir alan oluşturmaktadır (Wu et al., 2023).

6. SONUÇ VE ÖNERİLER

Bu çalışma, Google Trends arama verisinin Türkiye'ye yönelik aylık uluslararası ziyaretçi sayısının nowcasting'inde etkili bir öncü gösterge olduğunu ampirik olarak göstermiştir. "Antalya" arama terimi ziyaretçi sayısı ile en güçlü ilişkiyi sergilemekte, aramalar ziyaretçi hareketini yaklaşık iki ay önceden sinyallemektedir. Arama verisini dışsal değişken olarak içeren SARIMAX modeli, temel SARIMA modeline kıyasla tahmin hatasını belirgin biçimde düşürmüştür.

Bulgular, kamu kurumları ve sektör paydaşları için pratik çıkarımlar taşımaktadır. Kültür ve Turizm Bakanlığı ile sektör temsilcileri, aylık Google Trends endekslerini izleyerek resmî istatistikler açıklanmadan önce talep yönü hakkında erken sinyal elde edebilir; bu sinyal kapasite planlaması, tanıtım bütçesi tahsisi ve fiyatlandırma kararlarında kullanılabilir. Özellikle kriz ve toparlanma dönemlerinde arama verisinin öncü gücünün arttığı dikkate alındığında (Botha ve Saayman, 2024),

bu yaklaşımın erken uyarı mekanizması olarak değerlendirilmesi anlamlıdır. Önerilen yaklaşımın operasyonel bir erken uyarı aracına dönüştürülmesi teknik açıdan basit bir iş akışıyla mümkündür. İlk aşamada, "Antalya" arama endeksi her ayın başında Google Trends'ten otomatik olarak çekilir; bu işlem, açık kaynaklı bir betik (örneğin Python pytrends kütüphanesi) ile zamanlanmış görev biçiminde gerçekleştirilebilir. İkinci aşamada, güncellenen arama verisi mevcut SARIMAX modeline dışsal değişken olarak beslenir ve gelecek iki aya ilişkin ziyaretçi tahmini üretilir. Üçüncü aşamada, üretilen tahmin önceden belirlenen eşik değerlerle karşılaştırılır; belirli bir sapma gözlemlendiğinde ilgili birimlere otomatik bildirim gönderilir. Bu üç adımlı döngü, resmî istatistikler açıklanmadan yaklaşık iki ay önce talep yönü hakkında düzenli ve düşük maliyetli bir sinyal sağlayarak kapasite planlaması ve tanıtım kararlarına girdi oluşturabilir.

Gelecek çalışmalar, analizi destinasyon ve kaynak ülke düzeyinde ayrıştırarak mekansal bir boyut ekleyebilir, makine öğrenmesi yöntemleriyle karşılaştırmalı değerlendirme yapabilir ve sosyal medya verisi gibi ek alternatif kaynakları modele dahil edebilir. İkinci olarak, kaynak ülkeye özgü dillerde (Rusça, Almanca, vb.) arama terimleri kullanılarak ülke bazında ayrıştırılmış bir analiz yürütülebilir; bu yaklaşım, hangi kaynak pazarın arama davranışının daha güçlü öncü sinyal taşıdığını ortaya koyabilir. Bu genişletmeler, arama verisine dayalı turizm talebi tahmininin Türkiye bağlamındaki potansiyelini daha da ortaya çıkaracaktır.

Kaynakça

- Bangwayo-Skeete, P. F., ve Skeete, R. W. (2015). Can Google data improve the forecasting performance of tourist arrivals? Mixed-data sampling approach. *Tourism Management*, 46, 454-464. doi.org/10.1016/j.tourman.2014.07.014
- Botha, I., ve Saayman, A. (2024). Does google analytics improve the prediction of tourism demand recovery? *Forecasting*, 6(4), 908-924. https://doi.org/10.3390/forecast6040045
- Choi, H., ve Varian, H. (2012). Predicting the present with Google Trends. *Economic record*, 88, 2-9. https://doi.org/10.1111/j.1475-4932.2012.00809.x
- Çiçekdağı, M. (2021). Google trends verilerine göre turizmde tanınmışlık ve talep tahmini. *Çatalhöyük Uluslararası Turizm ve Sosyal Araştırmalar Dergisi*, (6), 140-157.
- Dinis, G., Breda, Z., Costa, C., ve Pacheco, O. (2019). Google trends in tourism and hospitality research: A systematic literature review. *Journal of Hospitality and Tourism Technology*, 10(4), 747-763. https://doi.org/10.1108/JHTT-08-2018-0086
- Google. (2026). *Google Trends: "turkey holiday", "Antalya", "Istanbul" arama terimleri için arama ilgisi, dünya geneli, Ocak 2012 - Aralık 2025* [Veri kümesi]. Google. https://trends.google.com/trends
- Höpken, W., Eberle, T., Fuchs, M., ve Lexhagen, M. (2019). Google Trends data for analysing tourists' online search behaviour and improving demand forecasting: The case of Åre, Sweden. *Information Technology ve Tourism*, 21(1), 45-62. https://doi.org/10.1007/s40558-018-0129-4
- Lee, G.-C. (2025). A Data-Driven Approach to Tourism Demand Forecasting: Integrating Web Search Data into a SARIMAX Model. *Data*, 10(5), 73. https://doi.org/10.3390/data10050073
- Minora, U., Iacus, S. M., Batista e Silva, F., Sermi, F., ve Spyrtatos, S. (2023). Nowcasting tourist nights spent using innovative human mobility data. *PloS one*, 18(10), e0287063. https://doi.org/10.1371/journal.pone.0287063
- Öncel Çekim, H. O., ve Koyuncu, A. (2022). The impact of Google Trends on the tourist arrivals: A case of Antalya tourism. *Alphanumeric Journal*, 10(1), 1-18. https://doi.org/10.17093/alphanumeric.931652
- Önder, I. (2017). Forecasting tourism demand with Google trends: Accuracy comparison of countries versus cities. *International Journal of Tourism Research*, 19(6), 648-660. https://doi.org/10.1002/jtr.2137
- Önder, I., ve Gunter, U. (2016). Forecasting tourism demand with Google Trends for a major European city destination. *Tourism Analysis*, 21(2-3), 203-220.

- Silverstovs, B., ve Wochner, D. S. (2018). Google Trends and reality: Do the proportions match?: Appraising the informational value of online search behavior: Evidence from Swiss tourism regions. *Journal of Economic Behavior ve Organization*, 145, 1-23. <https://doi.org/10.1016/j.jebo.2017.10.011>
- Sun, S., Zhong, S., Wu, J., ve Song, H. (2025). A data-driven approach to tourism demand forecasting: Integrating web search data into a SARIMAX model. *Data*, 10(5), 73. <https://doi.org/10.3390/data10050073>
- TÜİK. (2026). *Ziyaretçilerin aylara göre turizm geliri, kişi sayısı ve kişi başı ortalama harcama, 2012–2025* [Veri kümesi]. TÜİK. <https://data.tuik.gov.tr>
- Wu, D. C., Zhong, S., Wu, J., ve Song, H. (2025). Tourism and hospitality forecasting with big data: A systematic review of the literature. *Journal of Hospitality ve Tourism Research*, 49(3), 615-634. <https://doi.org/10.1177/1096348023122>
- Wu, J., Li, M., Zhao, E., Sun, S., ve Wang, S. (2023). Can multi-source heterogeneous data improve the forecasting performance of tourist arrivals amid COVID-19? Mixed-data sampling approach. *Tourism Management*, 98, 104759. <https://doi.org/10.1016/j.tourman.2023.104759>