

## Kısa, Orta ve Uzun Vadeli Trafik Akış Hızı Tahmini ve Görselleştirilme Aracı

*Short, Medium and Long Term Traffic Flow Forecasting and Data Visualization Tool*

Ibrahim TAKAK<sup>1</sup> , Halit GÖRMEZ<sup>1</sup> , H. İrem TÜRKMEN<sup>1</sup> , M. Amaç GÜVENSAN\*<sup>1</sup> 

<sup>1</sup>Yıldız Teknik Üniversitesi, Elektrik-Elektronik Fakültesi, Bilgisayar Mühendisliği Bölümü, 34220,

Esenler / İstanbul

### Öz

İstanbul içerisinde her geçen gün artan insan ve araç sayısı beraberinde ciddi bir trafik yoğunluğu getirmektedir. Oluşan bu araç yoğunluğunun giderilmesi veya kontrol edilebilmesi için eldeki verilerin iyi yorumlanması gerekmektedir. İstanbul il sınırları içerisinde ölçümleri yapılan tüm hız verileri İstanbul Büyükşehir Belediyesi tarafından kayıt altına alınmaktadır. Fakat bu verilerin sayılar üzerinden yorumlanması oldukça güçtür. Görselleştirme, sayısal verilerden anlam çıkarılması ve yorum yapılabilmesi için sıkça başvurulan bir yöntemdir. Bu çalışmada trafik yoğunluğunu analiz ederek görselleştiren, kısa, orta ve uzun vadede hız tahmini yapan bir araç geliştirilmiştir. Analiz kısmında İstanbul sınırları içerisinde birçok noktadaki sensörden alınan hız verileri belli bölge, zaman ve lokasyona göre özelleştirilmiş ve kullanıcıya talepleri doğrultusunda görsel bir içerik sunulmuştur. İçerik hazırlanırken ilgili bilgiler harita üzerinde işaretlenmiş veya sayısal grafiklerden yararlanılmıştır. Trafik tahmini yapabilmek için regresyon ve temel istatistiksel veriye dayalı üç farklı yöntem denenmiş ve elde edilen sonuçlar karşılaştırılmıştır. Geliştirilen sistem günün farklı zaman dilimleri için çıkarılan regresyon modellerini kullanarak %23.99 Ortalama Mutlak Yüzde Hatası ve 8.5 Ortalama Mutlak Hata ile 1 hafta sonraya kadar trafik akışı tahmini yapabilmektedir.

**Anahtar Kelimeler:** Trafik Akış Hızı Tahmini, Trafik Akış Analizi, Veri Görselleştirme, Regresyon.

### Abstract

The number of people and vehicles that are increasing day by day brings a serious traffic density to İstanbul. In order to eliminate or control this vehicle density, historical data should be interpreted well. All speed data that were measured in the provincial borders of İstanbul were recorded numerically by İstanbul Metropolitan Municipality. However, it is very difficult to evaluate these numerical data efficiently. Visualization is a frequently used method to interference from numerical data and evaluation of it. In this study, a tool that analyzes and visualizes traffic density and estimates speed in the short, medium and long term has been developed. In the analysis part, the speed data received from the sensors at many points within the borders of İstanbul are customized according to a specific region, time and location and a visual content is presented to the user on their demands. Relevant information was marked on the map or numerical graphics were used in order to present visual content. In order to perform traffic forecasting, three different methods based on regression and basic statistical data were employed and the obtained results were compared. Developed system can predict up to 1 week with a 23.99% Mean Absolute Percentage Error and 8.5 Mean Absolute Error by exploiting regression models that are generated for different time slots of a whole day.

**Keywords:** Traffic Flow Forecasting, Traffic Flow Analysis, Data Visualization, Regression.

## I. GİRİŞ

Jeopolitik konumu ve imkânlarından dolayı İstanbul'un nüfusu her geçen gün artmaktadır. Bu nüfus artışı beraberinde birçok problemi getirmektedir. Araç sayısındaki artış sonucu oluşan trafik de bu sorunların başında gelmektedir. Trafik sorununun çözümü ve sorunun kaynağının anlaşılması için öncelikle trafik verilerinin analiz edilmesi gerekmektedir. Yapılan analizler sonucunda trafik sorununun çözümü için daha az ve doğru yerlere yapılacak yatırımlarla daha verimli sonuçların alınması sağlanacaktır. Sorunların çözümünün yanında insanlar gidecekleri yer için alternatif ve hızlı yollar aramaktadır. Trafik yoğunluğunun az olduğu bir başka deyişle araçların hızlı ilerleyebildiği noktaların tespit edilmesi gerekmektedir. Bu da her bir lokasyon için bir hız tahmini yapılmasını gerektirmektedir. Çalışmanın amacı hem trafik verilerini analiz eden hem de kısa, orta ve uzun vadede hız tahmini yapabilen bir araç geliştirmektir. Geliştirilen uygulama şehir ve ulaşım planlama alanlarında katkı sağlayacak niteliktedir. Trafik akış hızı tahmini son yıllarda üzerinde sıklıkla çalışılan bir konudur. Literatürde büyük çoğunluğu kısa vadeli trafik akış hızı tahmini yapan pek çok çalışma mevcuttur [18][19]. Zaman serisi modellerinden Özbâğlanımlı Tümlleşik Yürüyen Ortalama (Autoregressive Integrated Moving-Average - ARIMA) [1], çeşitli regresyon modelleri [2,3] ve Yapay Sinir Ağları (Artificial Neural Networks) [4-6] sıklıkla kullanılan yöntemler arasındadır. Derin öğrenmenin karmaşık ve doğrusal olmayan sistemleri modellemedeki başarısı, trafik akışı tahmini için derin öğrenme yaklaşımlarını kullanan çalışmaların son yıllarda artmasını sağlamıştır [7,8].

Trafik akış hızı tahmini yapan çalışmalardan çok azı orta ve uzun vadeli trafik akış hızı tahmini üzerinde yoğunlaşmıştır. Zhao ve arkadaşları 15, 30, 45 ve 60 dakika sonrası trafik akışını tahmin edebilmek için Uzun Kısa-Vadeli Bellek Ağlarını önermişlerdir [9]. Chen ve arkadaşları GPS verisi üzerinde Bulanık Derin Konvolüsyonel Sinir Ağları (Fuzzy Deep Convolutional Neural Networks) modelini kullanarak 6 saate kadar uzaklık için trafik hızı tahmini yapmışlardır [10]. Peng ve arkadaşları Sezonluk Özbağlanımlı Tümlşik Yürüyen Ortalama (Seasonal ARIMA), Üstel Yumuşatma (Exponential Smoothing) ve İleri Beslemeli Yapay Sinir Ağları (Feed Forward Neural Networks) yöntemleri ile yağış verisinin kullanıldığı ve kullanılmadığı senaryolar için 24 saat uzağa kadar tahminleme yapmışlardır. 24 saat sonrası tahmin etmek için yapılan testlerde yaklaşık %17'lik bir ortalama mutlak yüzde hatası oranı ile en başarılı sonucu Sezonluk Özbağlanımlı Tümlşik Yürüyen Ortalama yöntemi vermiştir [11].

Literatürde trafiğe dair farklı problemlere yönelik görselleştirme çalışmaları yapılmıştır. Bu problemler şehir trafik akışının, trafik kazalarının, yaya hareketliliğinin ve hava kirliliğinin analizi ve görselleştirilmesi alt başlıkları altında değerlendirilebilir [25]. Bu çalışmalar arasında trafik akışını görselleştirmeye yönelik olanlar önemli bir yer tutmaktadır [20]. Trafik akış özelliklerini görselleştiren çalışmalardan bir kısmı kameralardan gelen görüntüleri analiz ederek akış görselleştirmesi yapmakta [21, 22, 24], bazıları ise rota bilgisinden yararlanmaktadır [23] ve kendi içinde trafik verisini uzamsal, zamansal veya kategorik açıdan ele alarak görselleştiren çalışmalar olarak ayrılabilir. Trafik sıklığının uzay-zamansal değişiminin harita üzerinde görselleştirilmesi [25], haftanın ya da ayın günlerine göre farklı saatlerdeki trafik sıklığının ısı haritası şeklinde ifade edilmesi [26] ve harita üzerinde farklı noktalara ait giriş ve çıkış yoğunluklarının görselleştirilmesi [27] farklı uygulama örnekleri olarak karşımıza çıkmaktadır.

Literatürdeki çalışmalar incelendiğinde, trafik görselleştirmesine dair farklı bakış açısına sahip çalışmalar olmakla birlikte trafik analizi ve ileriye yönelik tahminin birlikte yapıldığı bir araç ya da uygulamaya rastlanmamıştır. Yandex'in tahmin ve analiz için çalışmaları olmakla birlikte bu çalışmalar tek bir araçta toplanmamış durumdadır. Proje kapsamında geliştirilen araç hem geçmiş trafik verilerini analiz etmesi hem de hız tahmini yapması bakımından piyasadaki diğer örneklerden ayrılmaktadır. GoogleMaps anlık olarak trafik yoğunluğunu gösterebilmekte ve seçilen iki konum arası gidiş süresini yine bu anlık trafik yoğunluğuna göre hesaplamaktadır. Ancak İstanbul için düşük hassasiyet oranı ile bir trafik tahmini yapmaktadır [12]. YandexMaps ise anlık trafik durumunu gösterebilmekte, bir sonraki gün için ve yaklaşık bir ay ile sınırlandırılmış uzun vadede tahmin

yapabilmektedir [13]. Literatür, trafiğin analiz edilmesi açısından incelendiğinde YandexMaps'in geliştirdiği araç genel erişime açtığı görülmektedir [14]. YandexMaps'in analiz aracı incelendiğinde 2015, 2016, 2017 yılları üzerinde çalışmalar yapıldığı görülmektedir. Analiz yöntemlerine bakıldığında genel olarak bir yılın ortalamalarının alındığı ya da belirli ayların analizinin yapıldığı ve bunların kullanıcıya seçim hakkı tanımayan analizler olduğu görülmektedir. Geliştirilen aracın, YandexMaps analiz aracından farkı, kullanıcıya çok çeşitli ve esnek analiz yöntemleri sunmasıdır. Kullanıcı istediği tarih ve dönem aralıklarını kendi seçimi ile inceleyebilmektedir.

## II. VERİSETİ

Araç için gerekli olan hız verileri İstanbul Büyükşehir Belediyesi (İBB) tarafından temin edilmiştir. Temin edilen veriler MS SQL veri tabanı üzerinde tutulmuştur. Bu veriler 2017 yılına ait 7000 algılayıcıdan dakikada bir kez alınan ölçümleri içermektedir. Trafik akış hızı verileri MS SQL veri tabanı üzerinde ilgili algılayıcının numarası, yönü ve zaman bilgisi ile birlikte tutulmaktadır. Hava durumu etkisini görmek için kullanılacak olan hava durumu verileri Meteoroloji Genel Müdürlüğü (MGM)'den temin edilmiştir. Temin edilen bilgiler Excel dosyası şeklinde tutulmaktadır. Veriler istasyon numarası, istasyon ismi, ay, gün, saat ve yağış miktarı bilgilerinden oluşmaktadır. Daha hızlı sorgu sonuçları elde etmek için 2017 yılı hava durumu verileri her bir bölge için alt parçalara filtrelenmiştir.

## III. TRAFİK ANALİZ VE RAPORLAMA ARACI

Çalışma kapsamında gerçekleştirilen araç, Analiz ve Tahmin olmak üzere iki ana modülden oluşmaktadır. Analiz modülü Günlük Analiz, Aylık Hız Verileri, Aylar-Günler, Kritik Lokasyonlar, Renkli Yollar ve Hava Durumu Etkisi isimlerine sahip altı alt modülden oluşmaktadır.

Uygulamada önyüz tasarımı için HTML/CSS, JS dilleri kullanılmıştır. Arka yüz işlemleri için Python, SQL dilleri ve Django kütüphanesi kullanılmıştır. Harita işlemleri için ise GoogleMaps API kullanılmıştır.

### 3.1. Trafik Hızının Analiz Edilmesi

Bu modül geçmiş trafik hız verilerini belirli kriterler doğrultusunda görselleştirerek yorum yapmamızı sağlamaktadır.

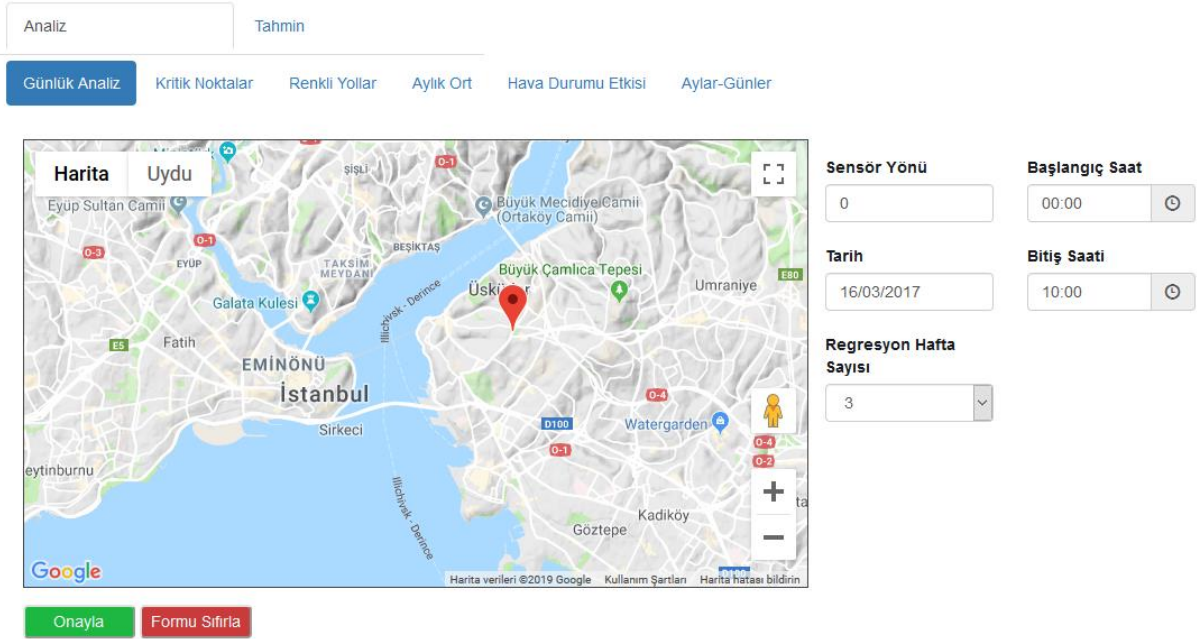
#### 3.1.1. Günlük Analiz

Günlük analiz kısmında algılayıcı yeri, algılayıcı yönü, tarih, başlangıç-bitiş saati ve oluşturulacak regresyon modeli için kaç hafta öncesine kadar verilerin değerlendirileceği bilgileri kullanıcıdan alınmıştır. Giriş ekranı Şekil 1'de gösterilmiştir. Alınan bilgiler doğrultusunda belirlenen kriterlere göre veri tabanından filtreleme yapılmıştır. İşlem hızını arttırmak

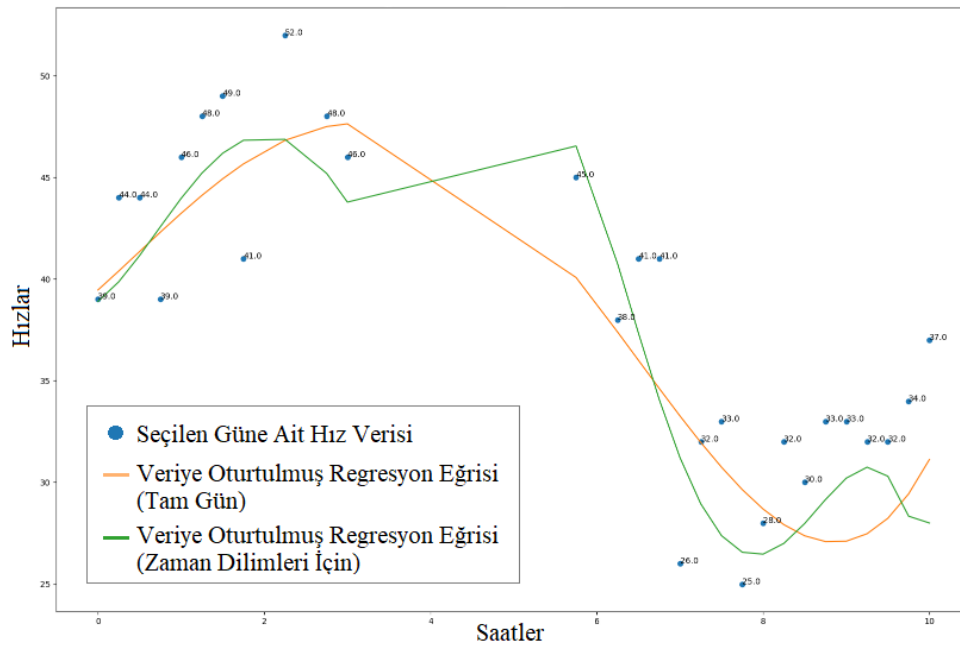
çin dakikalık veriler 15 dakikalık gruplara toplanmıştır.

Çıkarılan 15 dakikalık hız verileri grafik üzerinde üç farklı şekilde sunulmuştur. İlkinde, seçilen günün verileri noktasal olarak grafik üzerinde gösterilmiştir. Verilerin noktasal olarak gösterilmesinin yanı sıra hız ve saatler arasındaki ilişkinin formülize edilebilmesi için bu noktalara iki farklı regresyon eğrisi oturtulmuştur [15]. İlk regresyon eğrisi günü sabit belirli parçalara ayırarak (iş giriş, öğle iş çıkış gibi) her bir parça için kendi içerisinde regresyon grafiği oluşturmuş ve sonra eğriler birleştirilmiştir. İkincisinde

ise günün tamamı için tek bir regresyon eğrisi çıkarılmıştır. Oturtulan regresyon eğrileri ve gerçek veri kullanıcıya çıktı olarak gösterilmektedir. Çıktı ekranı Şekil 2'de verilmiştir. Dinamik olarak oluşturulan grafik çıktıları, doğrudan tahminleme yapmamakta, kullanıcıya trafik karakteristiği hakkında bilgi vermektedir. Tahmin modülünde kullanılacak en başarılı regresyon modelinin seçilmesi için yapılması gereken denemelerin sistematik olarak gerçekleştirilebilmesine yardımcı olmak adına gerçek hız verilerine oturtulmuş regresyon eğrileri de analiz aracında yer almaktadır.



Şekil 1. Günlük analiz bilgi girişi ekranı örnek görüntüsü



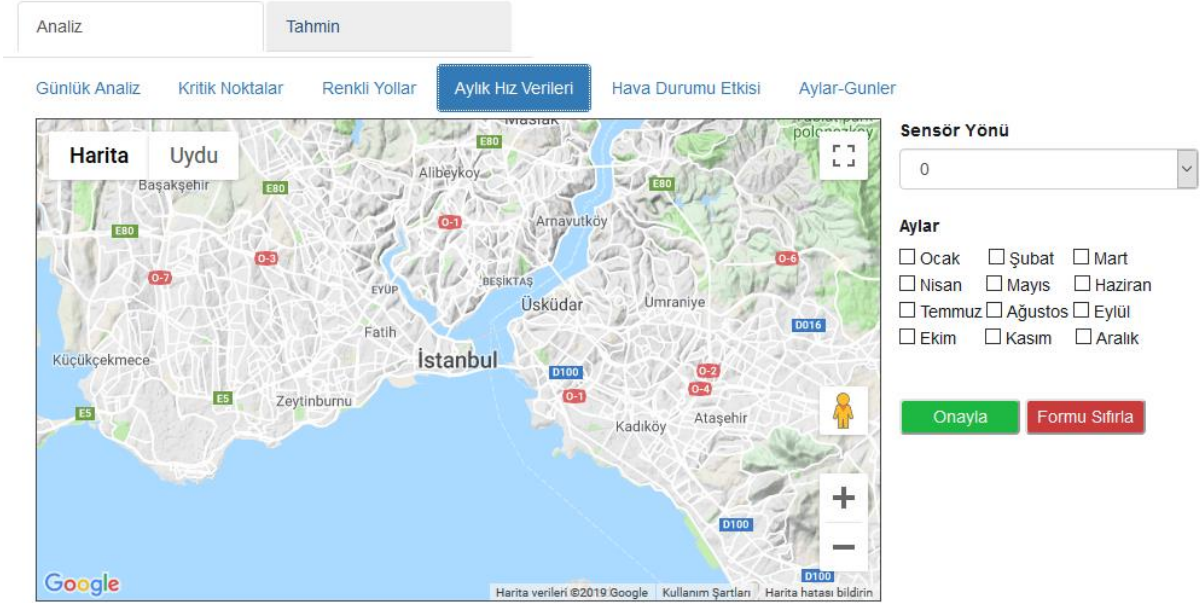
Şekil 2. Günlük analiz çıktı ekranı örnek görüntüsü

### 3.1.2. Aylık Hız Verileri

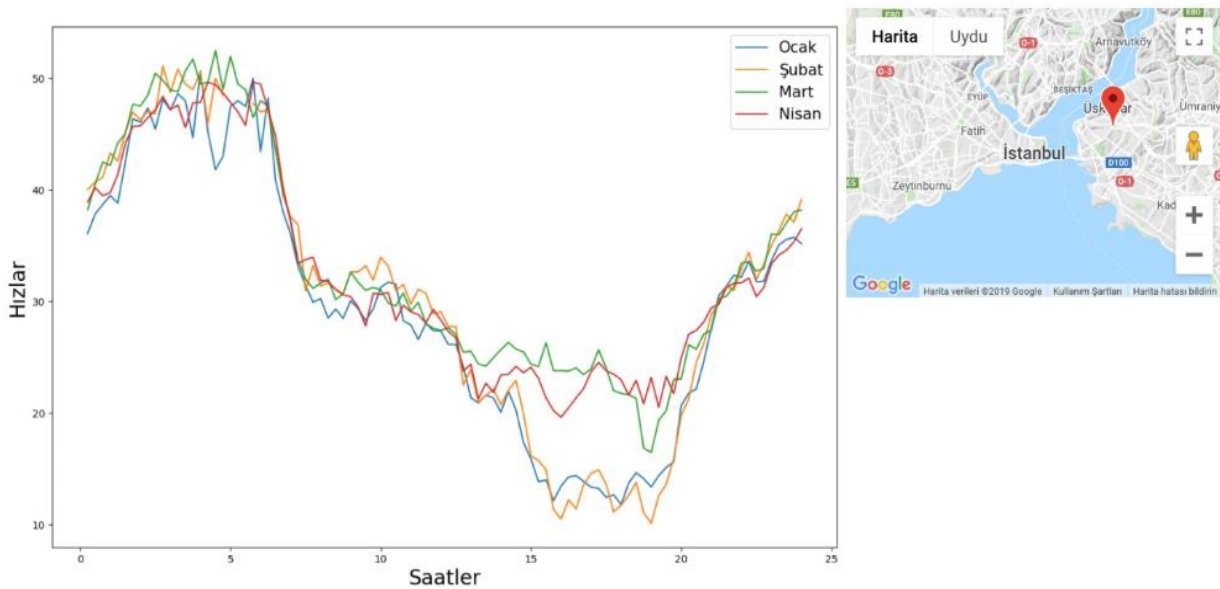
Aylık Hız Verileri modülü, trafik akış özelliklerinin farklı aylarda nasıl değiştiğinin gözlemlenebilmesi için geliştirilmiştir. Aylık Hız Verileri sekmesinde harita üzerinden istenen lokasyon seçimi yapılmıştır. Yapılan bu seçime göre en yakın algılayıcı tespit edilmiştir. Seçilen lokasyon için algılayıcı yönü seçimi yapılmıştır. Ayların karakteristiğini belirlemek için yapılan bu analizde seçim kutucuklarından istenilen aylar seçilmiştir. Veri giriş ekranı Şekil 3'te gösterilmiştir. Kullanıcıya çıktı olarak seçimi yapılmış aylara ait gün içinde 15 dakikalık hız değerlerini

gösteren grafik çıkarılmıştır. Çıkarılan bu grafiğin yanında bilgilendirmek amacıyla seçimi yapılan lokasyon harita üzerinde gösterilmiştir. Çıktı ekranı Şekil 4'te gösterilmiştir.

Kullanıcı, Aylık Hız Verileri modülünde seçilen ay(lar)ın günlük saatlerinin karakteristiğini görebilmektedir. Ayrıca birden fazla ay seçmesi durumunda, seçilen ayların günlük aynı saatlerine denk gelen hız ortalamalarını karşılaştırabilmektedir. Bu karşılaştırma sonucunda İstanbul trafiğinin aylara göre nasıl değişim gösterdiği yorumlanabilmektedir.



Şekil 3. Aylık hız verileri veri girişi ekranı örnek görüntüsü

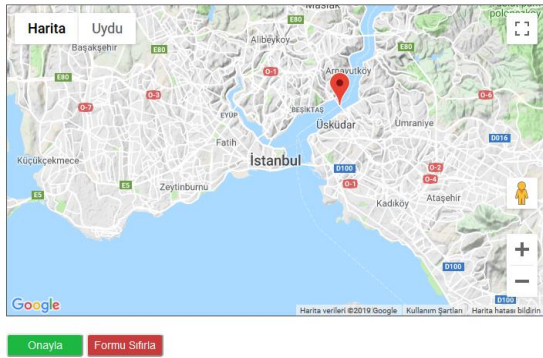


Şekil 4. Aylık hız verileri çıktı ekranı örnek görüntüsü



### 3.1.3. Aylar Günler

Aylar-Günler sekmesinde kullanıcıdan algılayıcı seçmesi, bir analiz türü seçmesi, Aylar ve Günler seçim kutucuklarından istediği gün(ler) ve ay(lar)ı seçmesi ve algılayıcı yönü seçmesi istenmektedir. Ayın Günlük Hızları ve Günün Aylık Hızları olmak üzere iki analiz türü vardır. Yapılan seçim sonucuna göre veritabanından filtreleme yapılmıştır. Seçilen analiz tipine göre seçilen aylardaki seçilen günlerin karakteristikleri veya seçilen günlerdeki seçilen ayların karakteristikleri çizgi grafiği ile gösterilmiştir. Seçim türlerinden Ayın Günlük Hızları'nda seçilen bir aydaki seçilen günlere ait akış karakteristikleri çizgi grafiğiyle gösterilmiş ve seçilen günlerin seçilen ay içerisindeki karakteristikleri incelenmiştir. Bu seçim için giriş ekranı

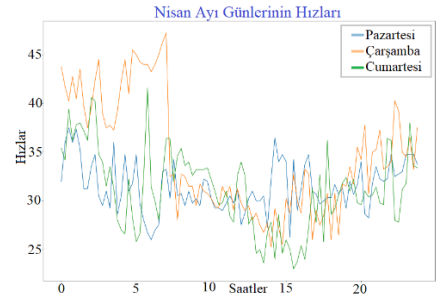


(a)

örnek görüntüsü Şekil 5a'da, çıkış ekran görüntüsü ise Şekil 5b'de verilmiştir.

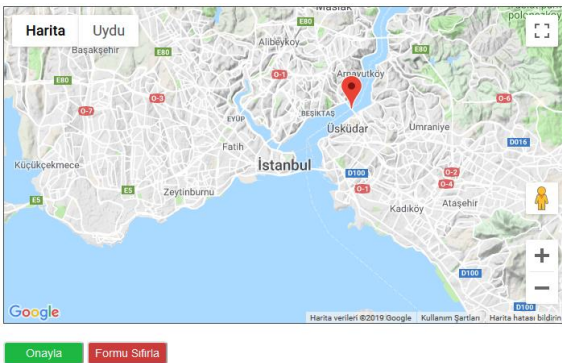
Diğer bir seçim olan Günün Aylık Hızları'nda ise seçilen günün, seçilen farklı aylardaki trafik akış karakteristikleri çizgi grafiğiyle gösterilmiştir. Bu seçim için giriş ekranı örnek görüntüsü Şekil 6a'da, çıkış ekran görüntüsü ise Şekil 6b'de gösterilmiştir.

Bu modülde yapılan karşılaştırmalar sonucunda günlük hız verilerinin aylardaki değişimleri incelenerek mevsimsel etkiler görülebilmektedir. Ayrıca aylık hız verilerinin günlerdeki değişimleri incelenerek hafta içi ve hafta sonu trafik karakteristikleri incelenebilmektedir.

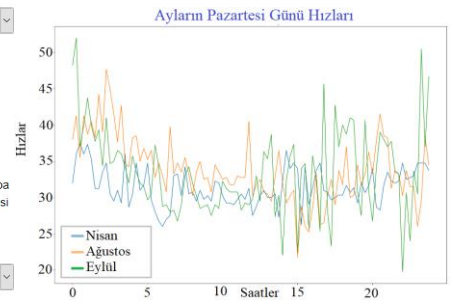


(b)

Şekil 5. Ayın günlük hızları analizi modülü a) Veri giriş ekranı ekran görüntüsü b) Çıkış ekran görüntüsü



(a)



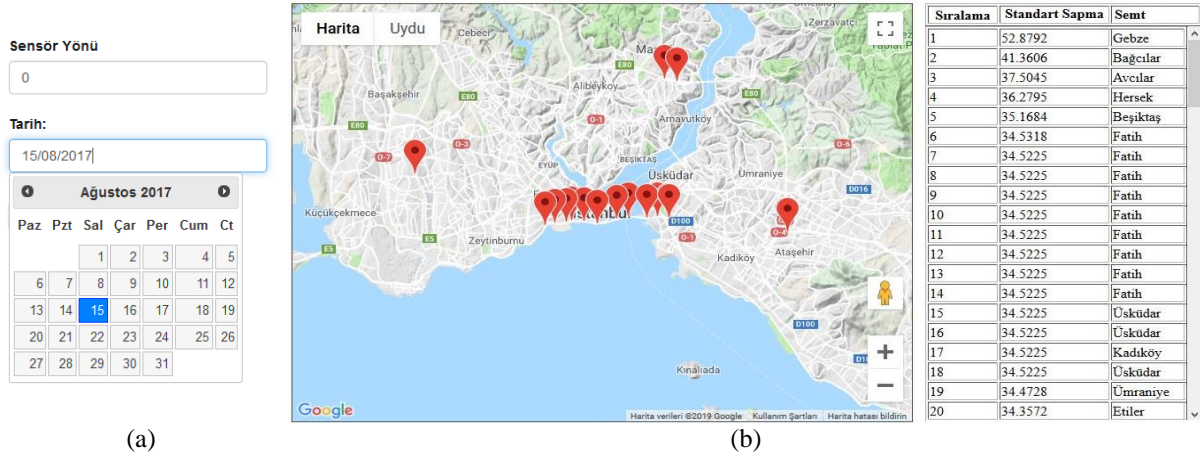
(b)

Şekil 6. Günün aylık hızları analizi modülü a) Veri giriş ekranı ekran görüntüsü b) Çıkış ekran görüntüsü

### 3.1.4. Trafikte Kritik Lokasyonlar

Trafik analizi ve hız tahmini aracının ikinci kısmı olan hız tahmininde İstanbul il sınırları içerisinde seçilen bir noktadaki ve belirli tarihteki hız tahmin edilmiştir. Bu kısım için başlangıç olarak İstanbul'daki hız değerlerinin en değişken olduğu bölgelerin pilot bölge olarak seçilmesine karar verilmiştir. Hızların en değişken olduğu bölgeleri tespit etmek için ise hızların standart sapmasına bakılmıştır. Bu şekilde belirlenen noktalara ise Kritik Lokasyonlar ismi verilmiştir. Kritik Lokasyonlar sekmesinde GoogleMaps üzerinden işaretleme yapılmıştır. Kullanıcıdan alınan tarih ve algılayıcı yönü bilgilerine göre veriler veritabanından filtrelenmiştir. Giriş ekranı Şekil 7a'da gösterilmiştir. Filtrelenen verilerdeki her bir algılayıcının hız değerleri

listeye kaydedilmiştir. Her bir algılayıcıya ait listelenen hız değerlerinin standart sapmaları hesaplanmıştır. Her bir algılayıcının standart sapma değeri ve algılayıcı numarası liste olarak dosyaya kaydedilmiştir. Kaydedilen standart sapmalar büyükten küçüğe sıralanmıştır. Algılayıcılardan en yüksek standart sapmaya sahip ilk 20 sensörün adres bilgisi çıkarılmıştır. Çıkarılan 20 algılayıcı GoogleMaps üzerinde pinlerle gösterilmiştir. Çıkarılmış olan adres bilgileri pinlere açıklama olarak eklenmiştir. Pinlerin işaretlendiği haritanın yanısıra standart sapması en yüksek 100 algılayıcının algılayıcı numarası, standart sapma değeri ve konumu da ekranda yer almaktadır. Çıkış ekranı Şekil 7b'de gösterilmiştir.



(a) (b)  
Şekil 7. Trafikte kritik noktalar a) Veri giriş ekranı ekran görüntüsü b) Çıkış ekran görüntüsü

### 3.1.5. Hız Bazlı Renkli Yollar

Bu modülde geçmiş hız verileri belirli lokasyon ve tarihe göre filtrelenmiştir. Filtrelenen veriler araçların hızlarına göre GoogleMaps üzerinde uygun renklerde gösterilmiştir. Renkli Yollar kısmında kullanıcıdan bir saat bilgisi, algılayıcı yönü ve tarih girdisi alınmıştır. Giriş ekranı Şekil 8a'da gösterilmiştir. Alınan bilgiler doğrultusunda veriler veritabanından filtrelenmiştir. Filtreleme sonucunda seçilen tarih ve saatteki tüm algılayıcılardan okunan hız değerleri çıkarılmıştır. Algılayıcıların yerleri belirlenerek harita üzerinde ardışık gelen her iki algılayıcı arası hız değerlerine göre GoogleMaps API'si yardımıyla renklendirilmiştir [16]. Trafik sıkışıklığının daha rahat anlaşılabilmesi için belli hız aralıklarının kendine özgü renkleri bulunmaktadır. Bu hız aralıkları İBB navi uygulamasındaki aralıklarla uyumlu olarak aşağıdaki şekilde seçilmiştir (km):

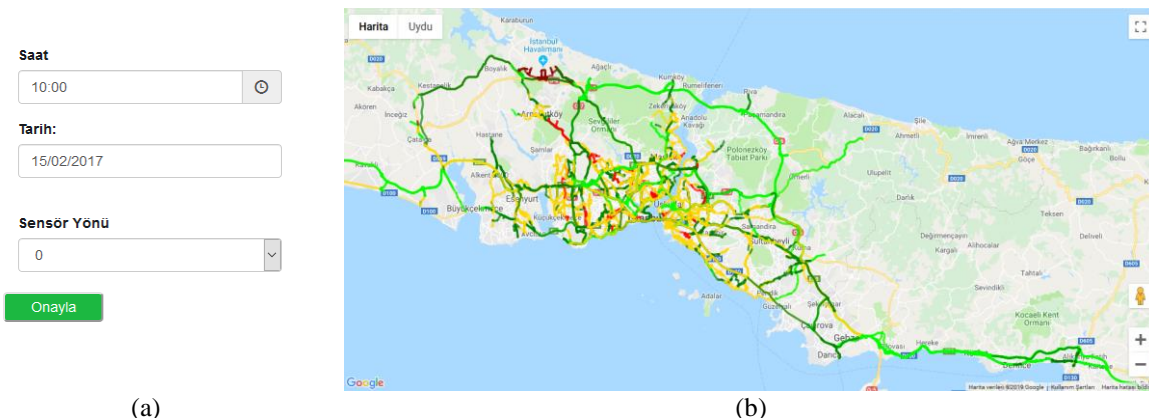
- [0-20] : Koyu kırmızı
- [20-30]: Açık kırmızı
- [30-45]: Turuncu
- [45-60]: Sarı
- [60-75]: Koyu yeşil
- [75-∞]: Açık yeşil

Çıkış ekranı Şekil 8b'de gösterilmiştir. Oluşturulan renklendirilmiş harita ile İstanbul'un genel hız yapısı

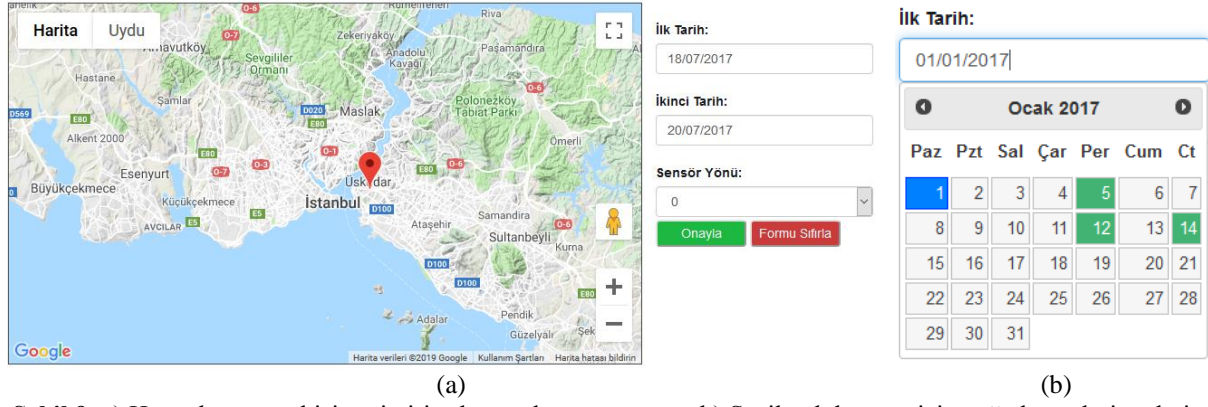
görülmüştür. Renkli Yollar sekmesi aynı zamanda geliştirilen aracın ikinci modülü olan Tahmin kısmının alt yapısını oluşturmuştur.

### 3.1.6. Hava Durumu Etkisi

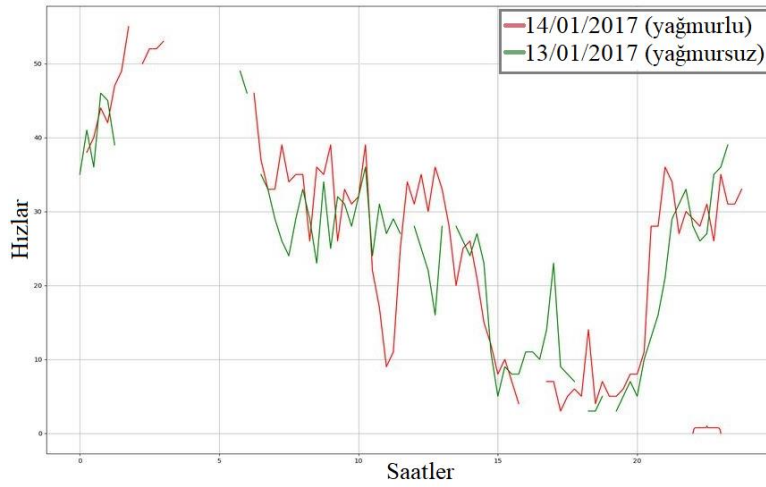
Hız verilerinin analiz edilmesinde hava durumu da önemli bir parametredir. Bu nedenle Meteoroloji Genel Müdürlüğü'nden temin edilen hava durumu verileri ile  $m^2$ 'ye düşen yağış miktarı baz alınarak Hava Durumu Etkisi sekmesi oluşturulmuştur. Bu sekmede kullanıcıdan lokasyon, algılayıcı yönü ve iki tarih bilgisi alınmaktadır. Giriş ekranı Şekil 9a'da gösterilmiştir. Alınan girdiler doğrultusunda veriler veritabanından filtrelenmiştir. Kullanıcı girdi ekranında öncelikle harita üzerinden seçim yapmıştır. Böylece seçilen lokasyonda hangi tarihlerde yağış olduğu takvim üzerinde işaretlenmiştir. Yeşil renk ile işaretli günler yağış alan günlerdir. Takvim üzerinde işaretleme örnek görüntüsü Şekil 9b'de gösterilmiştir. Kullanıcı, takvimler üzerindeki işaretlemeler sayesinde yağışlı ve yağışsız günleri görerek seçim yapıp yağışın etkisini inceleyebilmektedir. Çıktı olarak kullanıcıya seçilen iki tarihteki hız verilerinin çizgi grafiği gösterilmiştir. Çıkış ekranı Şekil 10'da verilmiştir. Algılayıcı arızaları sebebi ile bazı saatlerde ölçüm yapılamadığında ilgili grafikte kesintiler oluşmaktadır.



(a) (b)  
Şekil 8. Hız bazlı renkli yollar a) Veri giriş ekranı ekran görüntüsü b) Çıkış ekran görüntüsü



Şekil 9. a) Hava durumu etkisi veri giriş ekranı ekran görüntüsü b) Seçilen lokasyon için yağışlı günlerin takvim üzerinde gösterilmesi



Şekil 10. Seçilen tarihlerdeki saatlik hızlar ve bu tarihlerdeki yağış durumunu gösteren grafiğin ekran görüntüsü

### 3.2. Uzun Vadeli Trafik Tahmini Üzerine Çözümler

Geliştirilen aracın görevi kısa, orta ve uzun vadeli trafik tahmin ve analizi yapmaktır. Bu amaçla sistemde uzun vadeli tahmin için birden fazla çözüm yöntemi sunulmuştur. Bunlar sırasıyla; önceki haftaların ortalamasının tahmin için kullanılması, tüm bir gün için çıkarılan regresyon eğrisinin kullanılması ve karakteristiği birbirine benzeyen zaman dilimleri için çıkarılan ayrı regresyon eğrilerinin kullanılmasıdır. Regresyon analizi, bir değişkenin değerinin başka bir değişkenin değerine göre tahmin edilmesidir. Tahmin edilmek istenen değişken bağımlı değişken, bağımlı değişkenin değerini tahmin etmek için kullanılan değişken ise bağımsız değişken olarak adlandırılır. Y ve X sırası ile bağımlı ve bağımsız değişkenleri,  $\beta_i$  polinom katsayılarını ve  $\epsilon$  hata katsayısını göstermek üzere n. dereceden bir polinomial regresyon modeli **Eşitlik (1)**'de verilmiştir.

$$Y_i = \beta_0 + \beta_1 X + \beta_2 X^2 + \dots + \beta_n X^n + \epsilon \quad (1)$$

Regresyon modelinin katsayılarının bulunmasında en küçük kareler yöntemi kullanılmaktadır. Geliştirilen araçta kullanılacak olan regresyon eğrisi için kullanıcının belirlediği hafta sayısı kadar geriye gidilerek hız verilerinin ortalaması alınmıştır. Ortalama

hız verilerinden oluşturulan regresyon eğrisinin en uygun denklem derecesini bulabilmek için sırayla denemeler yapılmıştır. Yapılan denemelerde her bir denklem derecesi için gerçek hız değeri ile regresyon sonucunda bulunan değer arasındaki  $\epsilon$  hata miktarı hesaplanmıştır. Daha sonra her bir denklem derecesindeki  $\epsilon$  hata miktarı arası farklar hesaplanmıştır. Farkın daha önceden belirlenen  $\Delta$  değerinden az olduğu derece regresyon derecesi olarak seçilmiştir. Yapılan denemelerde trafik akış hızı tahmini için en uygun regresyon denklem derecesinin 7 olduğu görülmüştür.

#### 3.2.1. Önceki Haftaların Ortalamasına Göre Hız Tahmini

Trafik tahmini için kullanılan yöntemlerden biri, önceki haftaların ortalamasının alınmasıdır [17]. Bu yöntemde kullanıcı, arayüz üzerinden tahmin etmek istediği gün, saat, trafik akış yönü ve tahmin modelinde kullanılacak hafta sayısı belirlemektedir. Kullanıcının girdiği hafta sayısı kadar geriye gidilerek aynı saat ve dakikaya ait hız verilerinin ortalaması alınmıştır. k kullanıcının geriye giderek ortalamaya dahil etmek istediği hafta sayısı,  $V_i$  haftanın seçilen gün ve saati için i hafta kadar geriye gidildiğinde ölçülen hız değeri



olmak üzere, kullanıcının tahmin etmek istediği gün ve saat  $V_0$ , Eşitlik (2) ile hesaplanmıştır.

$$V_0 = \frac{1}{k} \sum_{i=1}^k V_i \quad (2)$$

Ortalama alınan hız verileri Hız Bazlı Renkli Yollar modülü alt yapısı kullanılarak renklendirilmiş olarak ekranda gösterilmektedir.

### 3.2.2. Tam Gün İçin Çıkarılan Regresyon Eğrisi ile Tahmin

Tahmin için kullanılan ikinci yöntem ise günlük regresyondur. Tahmin edilmesi istenilen günün seçilen hafta sayısı kadar geriye gidilip aynı günlerin ortalaması alınmıştır. Hesaplanan ortalama hız verileri kullanılarak bütün bir gün için regresyon eğrisi çıkarılmıştır. Bu eğriden yola çıkılarak istenen saat ve dakikaya ait hız verisi tahmin edilmiştir. Bu işlem tüm algılayıcılar için uygulanmış ve ilgili yollar yoğunluklarına göre renklendirilmiştir.

### 3.2.3. Günün Zaman Dilimleri İçin Çıkarılan Regresyon Eğrisi ile Tahmin

Tahmin için kullanılan son yöntem bir günün belirli zaman dilimleri için birden fazla regresyon eğrisi grafiği çıkarılmasıdır. Hız karakteristikleri farklılık gösteren iş girişi, öğle, iş çıkışı, akşam ve gece saatleri için ayrı ayrı regresyon denklemleri çıkarılmıştır. Böylece gerçek hız verilerine daha iyi oturan bir regresyon eğrisi oluşturulmuştur.

## IV. DENEYSEL SONUÇLAR

Hız tahmini için öne sürülmüş yaklaşımların başarısını ölçmek için Ortalama Mutlak Yüzde Hata (Mean Absolute Percentage Error, MAPE) ve Ortalama Mutlak Hata (Mean Absolute Error, MAE) metriklerinden faydalanılmıştır [17]. MAPE bir tahmin sisteminin ne kadar doğru olduğunu istatistiksel bir ölçüsüdür ve  $A_i$  gerçek değer ve  $P_i$  ise tahmin değeri olmak üzere Eşitlik (3) ile hesaplanmaktadır. Bu metrik doğruluğu yüzde olarak ölçmektedir.

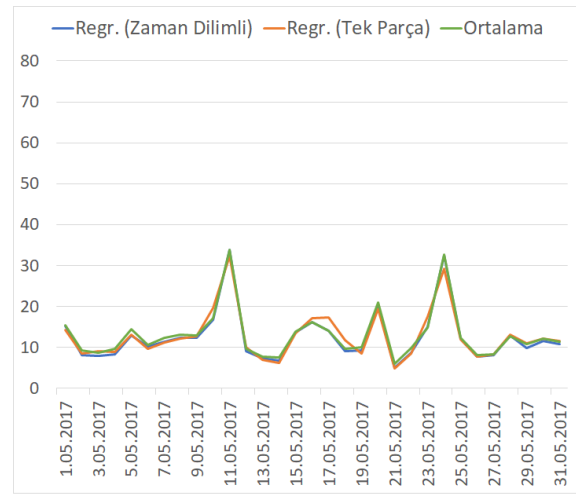
$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{A_i - P_i}{A_i} \right| \quad (3)$$

MAE ise hataların ortalama büyüklüğünü hata yönünü dikkate almadan ölçmektedir ve Eşitlik (4) ile ifade edilmektedir.

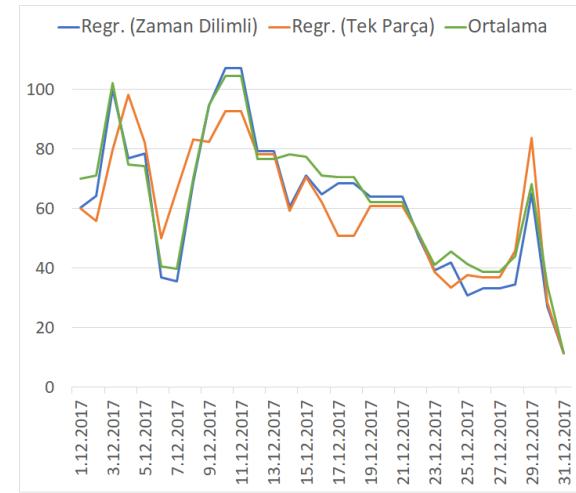
$$MAE = \frac{1}{n} \sum_{i=1}^n |A_i - P_i| \quad (4)$$

Önerilen üç farklı tahmin yönteminin başarılarının karşılaştırılması için, Mayıs ve Aralık aylarına ait beş farklı lokasyondan alınan trafik akış hızı verileri değerlendirilmiştir. Her bir günün ilgili lokasyonundaki trafik akış hızı tahmini için, ilgili tarihten 3 hafta geriye

kadar gidilerek aynı algılayıcının ilgili gün ve saatteki verileri kullanılmıştır. Tahmin edilmek istenen her tarih için, o tarihten önceki haftaların bilgisi kullanıldığından, model oluşturmak için başvuru verileri ve test için kullanılan veriler birbirinden tamamen ayrılmış durumdadır. Böylece 1 hafta sonrasına kadar akış hızı tahmini yapmak mümkün olmaktadır. Her bir lokasyon ve tarih için o günün her bir saatinde yönelik tahminler yapılmış ve yapılan tahminlerin MAPE ortalamaları alınmıştır. Yatay ekseninde tarihler, dikey ekseninde ise elde edilen ortalama MAPE değerleri olmak üzere her bir ay ve lokasyona ait performans grafikleri Şekil 11, 12, 13, 14 ve 15'te verilmiştir.



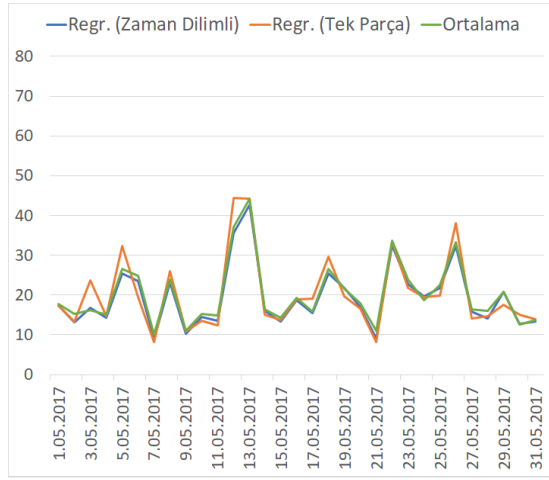
(a)



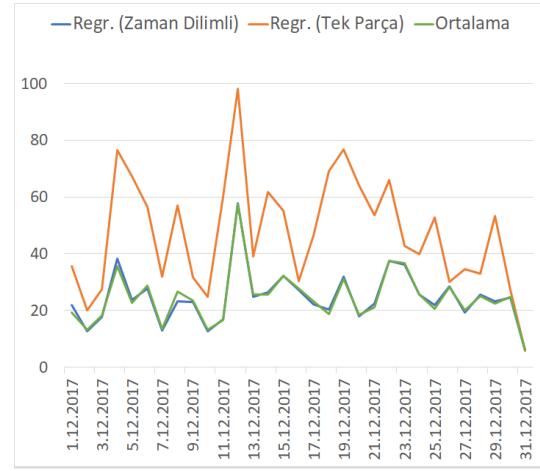
(b)

Şekil 11. Beşiktaş için a) Mayıs ayı b) Aralık ayı ortalama MAPE oranları



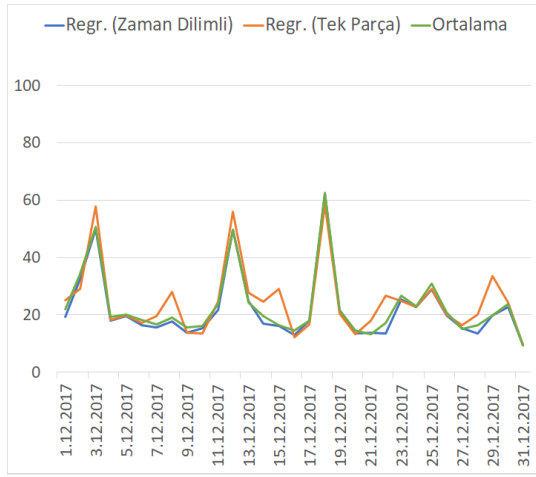


(a)



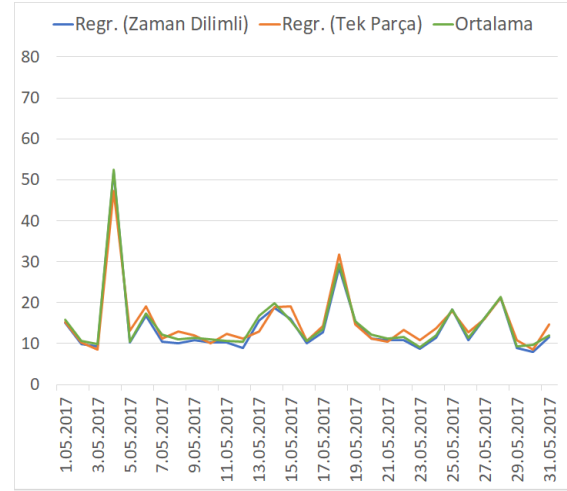
(b)

Şekil 13. Kadıköy için a) Mayıs ayı b) Aralık ayı ortalama MAPE oranları

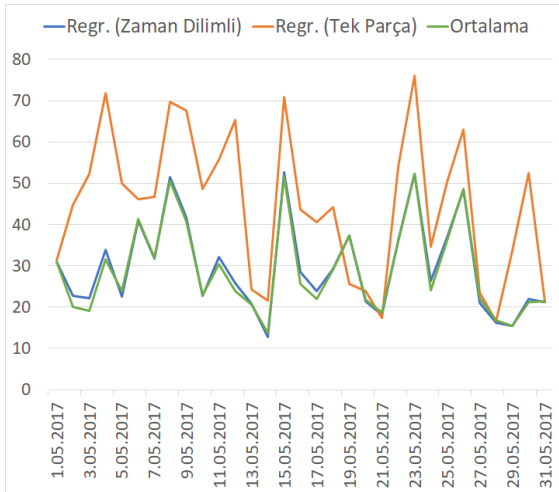


(b)

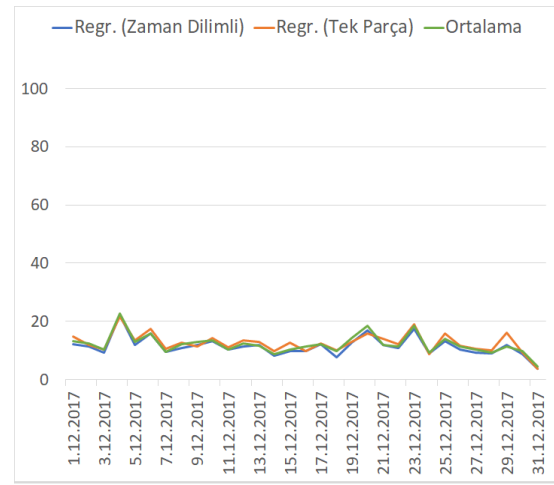
Şekil 12. Üsküdar için a) Mayıs ayı b) Aralık ayı ortalama MAPE oranları



(a)

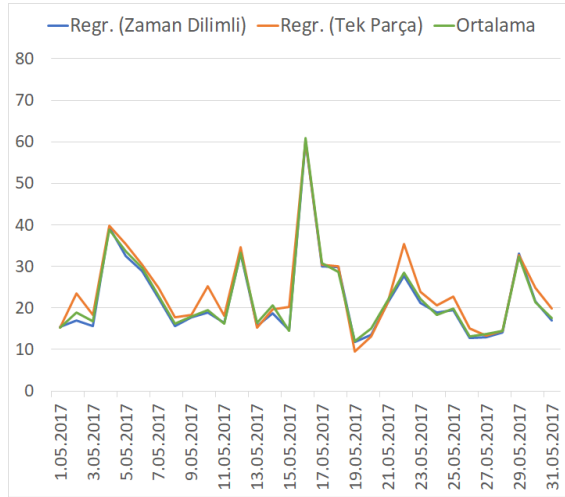


(a)

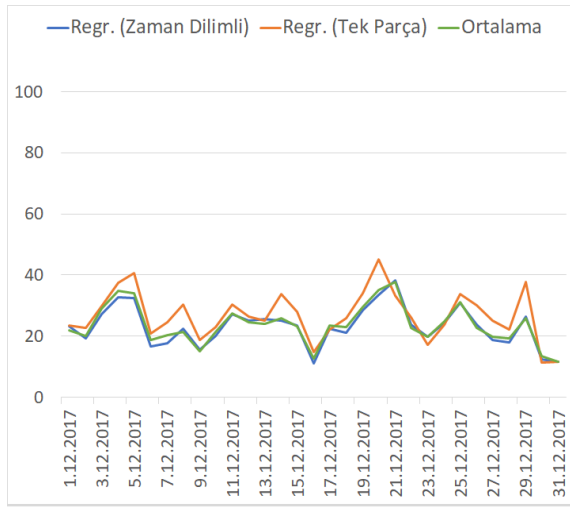


(b)

Şekil 14. FSM Köprüsü için a) Mayıs ayı b) Aralık ayı ortalama MAPE oranları



(a)



(b)

**Şekil 15.** Şişli için a) Mayıs ayı b) Aralık ayı ortalama MAPE oranları

Sonuçlar incelendiğinde genel olarak Mayıs ayındaki günlerin Aralık ayına göre daha yüksek başarı ile tahmin edildiği görülmektedir. Bunun sebebi olarak kış mevsiminin değişken hava koşulları sebebiyle, trafik karakteristiğinin öngürülmesinin daha zor bir örüntü sergilemesi olduğu düşünülmektedir. Bununla birlikte grafiklerde hata oranlarının birden yükseldiği günlerin çoğunlukla Cuma günlerine denk geldiği gözlenmiştir.

Mayıs ayı için beş lokasyondan elde edilen ortalama MAPE oranları Tablo 1'de, Aralık ayı için elde edilen MAPE oranları ise Tablo 2'de verilmiştir.

Sonuçlar incelendiğinde günün farklı zaman dilimleri için çıkarılan regresyon eğrisinin diğer yöntemlere göre az da olsa daha iyi sonuçlar verdiği gözlenmiştir. Bu yöntem ile her iki ay için ortalama %23.99 MAPE hata oranı alınırken, MAE hata oranı iki ay için ortalama 8.52 olmuştur. Öte yandan geçmiş hafta verilerinin ortalamalarını alarak yapılan hesaplamaların da basit

olmasına karşın oldukça tatminkâr sonuçlar verdiği gözlenmiştir.

**Tablo 1.** Farklı yöntemler ile yapılan tahminlerde Mayıs ayı için beş lokasyondan elde edilen ortalama MAPE oranları

Lokasyon	Regresyon (Zaman Dilimli)	Regresyon (Tek Parça)	Ortalama
Beşiktaş	12,59	12,87	13,15
Üsküdar	19,44	20,31	20,23
Kadıköy	29,67	44,74	29,11
FSM Köprüsü	14,19	15,09	14,83
Şişli	22,18	24,02	22,64
<b>Ortalama</b>	<b>19,61</b>	<b>23,41</b>	<b>19,99</b>

**Tablo 2.** Farklı yöntemler ile yapılan tahminlerde Aralık ayı için beş lokasyondan elde edilen ortalama MAPE oranları

Lokasyon	Regresyon (Zaman Dilimli)	Regresyon (Tek Parça)	Ortalama
Beşiktaş	60,70	60,68	63,52
Üsküdar	21,98	24,85	23,04
Kadıköy	24,58	47,38	24,51
FSM Köprüsü	11,43	12,61	12,14
Şişli	23,17	26,76	23,70
<b>Ortalama</b>	<b>28,37</b>	<b>34,46</b>	<b>29,38</b>

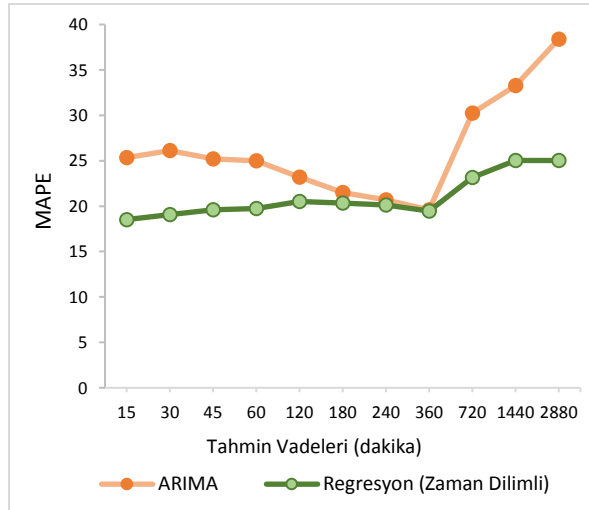
Analiz için İstanbul il sınırları içerisinde konumlandırılmış algılayıcılar içerisinde trafik karakteristiğinin en çok değişiklik gösterdiği (Bölüm 3.1.4) beş farklı lokasyon (Beşiktaş, Kadıköy, Üsküdar, FSM Köprüsü ve Şişli) seçilmiştir. Bu lokasyonlarda belirli zamanlarda yoğun trafik gözlenmekte, gece ve gündüze ait saatlerde trafik akış hızı büyük oranda değişmektedir. Bununla birlikte trafik yoğunluğunun yüksek olduğu bölgeler trafik kazası, yol çalışması, yağış gibi dış etmenlerden daha fazla etkilenmektedir. Tüm bu nedenlerle, hangi yöntem kullanılırsa kullanılsın, bu lokasyonlar için yapılacak tahmin zorlaşmakta ve elde edilen başarı oranları, seyir hızının daha yüksek ve trafik karakteristiğinin daha kararlı olduğu bölgeler için elde edilecek başarı oranlarından daha düşük olabilmektedir. Zaman dilimleri için çıkarılan regresyon modelleri ile Aralık ayına ait tahminlerde Beşiktaş, Mayıs ayına ait tahminlerde ise Kadıköy bölgesinde nispeten daha yüksek hata oranları elde edildiği gözlenmiştir. Bu yüksek hata oranlarının ilgili algılayıcı ve zaman dilimine ait veri eksikliği ve karakteristiği etkileyen harici bir etmene (etkinlik, kaza vb.) bağlı olarak geliştiği düşünülmektedir. Veri eksikliğinin giderilmesi ve harici etmenlerin de

değerlendirmeye katılması sistem başarısını önemli oranda olumlu etkileyecektir.

Bu çalışmada önerilen farklı zaman dilimleri için çıkarılan regresyon eğrisi modelinin uzun ve kısa vadeli tahminlemedeki başarısını gözlemleyebilmek için bu model literatürde sıklıkla kullanılan ARIMA model ile karşılaştırılmıştır. Karşılaştırma için yapılan testler kapsamında 15 dakikadan 2 güne kadar farklı tahmin vadeleri için ARIMA ve farklı zaman dilimleri için elde edilmiş regresyon eğrilerinin verdiği MAPE oranları incelenmiştir.

ARIMA, özellikle kısa dönemli hız tahmini için literatürde sıklıkla kullanılan bir yöntemdir. Bir zaman serisinin geçmiş değerleri, rassal şokların doğrusal bir fonksiyonu ile modeller ve ARIMA(p, d, q) şeklinde gösterilir. Bir ARIMA model d. dereceden fark alınarak durağan hale getirilmiş serilere uygulanan AR (Otoregresif modeli) ve MA (Hareketli ortalama modeli) modellerinin birleşimidir. p ve q sırasıyla AR ve MA modellerinin derecesini göstermektedir [28]. Yapılan testlerde en iyi sonucu verdiği için ARIMA(1, 1, 0) parametreleri ile modelleme yapılmıştır.

Zaman dilimli regresyon ile tahminde Mayıs ve Aralık aylarının her günü için k=3 hafta geriye gidilerek ilgili gün için eğriler oluşturulmuş ve bu eğriler istenen dakika ve saatin tahmininde kullanılmıştır. Sonuçlar Şekil 16'da özetlenmiştir.



**Şekil 16.** ARIMA ve zaman dilimli regresyon modellerine ait farklı tahmin vadeleri için MAPE oranları

Şekil 16 incelendiğinde farklı zaman dilimleri için oluşturulmuş regresyon modelinin tahmin vadesi arttıkça ARIMA modele göre oldukça başarılı sonuçlar verdiği gözlenmiştir. ARIMA modelinin başarısı özellikle 12 saatten daha ileriye yönelik tahminlerde bariz bir şekilde düşüş göstermektedir.

Önerilen farklı zaman dilimleri için oluşturulmuş regresyon modelinin uzun vadeli tahminlemedeki

başarısı ve elde edilen 8.52 MAE oranı sistemin endüstriyel kullanım için umut vadeder nitelikte olduğunu ortaya koymaktadır.

## V. SONUÇ VE TARTIŞMA

Bu çalışmada kısa, orta ve uzun vadeli trafik hız tahmini yapan ve geçmiş verileri analiz eden bir veri görselleştirme aracı geliştirilmiştir. Analiz kısmında verileri görselleştirmek için grafik oluşturma, regresyon eğrisi ile yakınsama ve haritalandırma yöntemleri kullanılmıştır. Geliştirilen araç ile standart sapma kullanılarak trafik karakteristiği en çok değişen lokasyonlar belirlenebilmektedir. Geçmiş hız verilerini kullanarak 1 hafta uzağa kadar trafik akış hız tahmini yapılabilmektedir. Tasarlanan sistem ile, yetkililerin gün özelliği (hafta sonu, hafta içi, resmi tatil vs) ve hava durumu bilgisini de göz önünde bulundurularak trafik planlaması yapmalarına yardımcı olacak kullanıcı dostu bir arayüz sunulmuştur. Ayrıca geliştirilen arayüz ile, sinyalizasyon ve kavşak çalışması gibi yol düzenlemelerinin trafiğe olumlu/olumsuz etkisi gözlemlenebilmektedir.

Öte yandan literatürde yağmurlu trafik karakteristiğine etkilerine işaret eden pek çok çalışma mevcuttur [29]. Önerilen regresyon ve ortalama tabanlı öğrenme modelleri ilgili gün için k hafta kadar geri gidilerek gözlenen hız verilerini tahmin için kullanılmaktadır. Tüm modellerde en başarılı sonuç k=3 için elde edilmiştir. Ancak önceki 3 hafta içerisinde denk gelecek yağmurlu gün sayısının az olması, sistemin karakteristiği öğrenmesinin önüne geçmektedir. Örneğin yağışlı bir pazartesi gününe ait trafik tahmini yapılması isteniyorsa, modelleme için geçmiş verilerden k adet yağışlı pazartesiye ihtiyaç vardır. İlgili veriyi elde edebilecek kadar geriye gitmek mevsim değişikliğine, dolayısı ile yanlış yönlendirilmiş bir modele yol açmaktadır. Bu sebeple yağış bilgisi bu çalışmada tahmin için kullanılmamış, analiz için değerlendirilmiştir.

İleriki çalışmalarda sistem performansını arttırabilmek için, farklı lokasyonlarda veya zaman dilimlerinde algılayıcı arızası sebebi ile eksik veri bulunması durumunda interpolasyon yöntemi ile aradaki az miktardaki eksik verinin elde edilmesi yoluna gidilerek başarıya katkı sağlanması planlanmaktadır. Özellikle İstanbul'da hava durumu dışında trafik akışını etkileyen yol çalışması, sosyal etkinlikler, trafik kazaları gibi gelecek çalışmalarda değerlendirilebilecek pek çok etmen bulunmaktadır. Bu etmenlerin tahmin sürecine dahil edilmesi başarı oranlarını önemli ölçüde arttıracaktır. Sistemi, farklı kaynaklardan besleyerek bu bilgileri de analiz ve tahmin araçlarına dahil edecek şekilde geliştirmek ve ilgili veriler ile birlikte kamusal kullanıma açmak hem konu hakkındaki araştırmacılara hem de karar vericilere önemli bir katkı sağlayacaktır.

## VI. TEŞEKKÜR

Bu çalışma 120E357 numaralı TÜBİTAK 1001 projesi kapsamında TÜBİTAK tarafından desteklenmektedir.

## KAYNAKLAR

- [1] Vasantha Kumar, S., & Vanajakshi, L. (2015). Short-term traffic flow prediction using seasonal ARIMA model with limited input data. *European Transport Research Review*, 7(3), 1-9.
- [2] Jeong, Y. S., Byon, Y. J., Castro-Neto, M. M., & Easa, S. M. (2013). Supervised weighting-online learning algorithm for short-term traffic flow prediction. *IEEE Transactions on Intelligent Transportation Systems*, 14(4), 1700-1707.
- [3] Chang, H., Lee, Y., Yoon, B., & Baek, S. (2012). Dynamic near-term traffic flow prediction: system-oriented approach based on past experiences. *IET intelligent transport systems*, 6(3), 292-305.
- [4] Van Lint, J. W. C., Hoogendoorn, S. P., & van Zuylen, H. J. (2005). Accurate freeway travel time prediction with state-space neural networks under missing data. *Transportation Research Part C: Emerging Technologies*, 13(5-6), 347-369.
- [5] Zhong, M., Sharma, S., & Lingras, P. (2005). Short-term traffic prediction on different types of roads with genetically designed regression and time delay neural network models. *Journal of computing in civil engineering*, 19(1), 94-103.
- [6] Kumar, K., Parida, M., & Katiyar, V. K. (2013). Short term traffic flow prediction for a non urban highway using artificial neural network. *Procedia-Social and Behavioral Sciences*, 104, 755-764.
- [7] Lv, Y., Duan, Y., Kang, W., Li, Z., & Wang, F. Y. (2014). Traffic flow prediction with big data: a deep learning approach. *IEEE Transactions on Intelligent Transportation Systems*, 16(2), 865-873.
- [8] Wu, Y., Tan, H., Qin, L., Ran, B., & Jiang, Z. (2018). A hybrid deep learning based traffic flow prediction method and its understanding. *Transportation Research Part C: Emerging Technologies*, 90, 166-180.
- [9] Zhao, Z., Chen, W., Wu, X., Chen, P. C., & Liu, J. (2017). LSTM network: a deep learning approach for short-term traffic forecast. *IET Intelligent Transport Systems*, 11(2), 68-75.
- [10] Chen, W., An, J., Li, R., Fu, L., Xie, G., Bhuiyan, M. Z. A., & Li, K. (2018). A novel fuzzy deep-learning approach to traffic flow prediction with uncertain spatial-temporal data features. *Future generation computer systems*, 89, 78-88.
- [11] Peng, H., Bobade, S. U., Cotterell, M. E., & Miller, J. A. (2018, June). Forecasting traffic flow: Short term, long term, and when it rains. In *International Conference on Big Data* (pp. 57-71). Springer, Cham.
- [12] Google haritalar. [Online]. Available: <https://www.google.com.tr/maps/@41.0344773,28.9327462,15615m/data=!3m1!1e3?hl=tr>
- [13] Yandex controls, developer's guide. [Online]. Available: <https://yandex.com.tr/dev/maps/jsapi/doc/2.1/dg/concepts/controls.html/>
- [14] Yandex İstanbul için 3 yıllık trafik analizi. [Online]. Available: [https://yandex.com.tr/company/press\\_center/info/graphics/istanbul\\_traffic](https://yandex.com.tr/company/press_center/info/graphics/istanbul_traffic)
- [15] Ostertagová, E. (2012). Modelling using polynomial regression. *Procedia Engineering*, 48, 500-506.
- [16] Google. Googlemaps. [Online]. Available: <https://developers.google.com/maps/documentation/?hl=tr>
- [17] Yasli, F., Turkmen, H. İ., & Guvensan, M. A. Long-term Traffic Speed Estimation via Regression using Weekly Day Patterns. In *2019 Innovations in Intelligent Systems and Applications Conference (ASYU)* (pp. 1-6). IEEE.
- [18] Akin, M., & Sağiroğlu, Ş. (2019, November). Büyük Veri Perspektifinden Trafik Tahmini Traffic Prediction Based on Big Data Perspective. In *2019 1st International Informatics and Software Engineering Conference (UBMYK)* (pp. 1-6). IEEE.
- [19] Ozdag, M. E. (2020). Derin Öğrenme Teknikleri Kullanılarak Anayol Trafik Analizi (Doctoral dissertation).
- [20] Chen, W., Guo, F., & Wang, F. Y. (2015). A survey of traffic data visualization. *IEEE Transactions on Intelligent Transportation Systems*, 16(6), 2970-2984.
- [21] Dereli, S. (2016, November). Trafik Yoğunluğunun Veri Görselleştirme Yöntemi ile Gerçek Zamanlı İzlenmesi İçin Yeni Bir Sistem Yaklaşımı. In *4th International Symposium on Innovative Technologies in Engineering and Science (ISITES2016) 3-5 Nov 2016 Alanya/Antalya-Turkey*.
- [22] Hsieh, C. Y., & Wang, Y. S. (2016). Traffic situation visualization based on video composition. *Computers & Graphics*, 54, 1-7.
- [23] Wang, Z., Lu, M., Yuan, X., Zhang, J., & Van De Wetering, H. (2013). Visual traffic jam analysis based on trajectory data. *IEEE transactions on visualization and computer graphics*, 19(12), 2159-2168.
- [24] Tsuboi, T. (2020). Traffic congestion visualization by traffic parameters in India. In *International Conference on Innovative Computing and Communications* (pp. 371-380). Springer, Singapore.
- [25] Andrienko, G., Andrienko, N., Bak, P., Keim, D., & Wrobel, S. (2013). Visual analytics focusing on spatial events. In *Visual analytics of movement* (pp. 209-251). Springer, Berlin, Heidelberg.



- 
- [26] Song, Y., & Miller, H. J. (2012). Exploring traffic flow databases using space-time plots and data cubes. *Transportation*, 39(2), 215-234.
- [27] Demissie, M. G., de Almeida Correia, G. H., & Bento, C. (2013). Exploring cellular network handover information for urban mobility analysis. *Journal of Transport Geography*, 31, 164-170.
- [28] Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.
- [29] Qi, Y., Zheng, Z., & Jia, D. (2020). Exploring the spatial-temporal relationship between rainfall and traffic flow: a case study of Brisbane, Australia. *Sustainability*, 12(14), 5596.