

## Araştırma Makalesi

### OpenStreetMap Binalarının Mekânsal Doğruluğunun Analizi

Kadir Can Küçük<sup>1</sup>, Berk Anbaroğlu\*<sup>1</sup>

<sup>1</sup>Hacettepe Üniversitesi, Mühendislik Fakültesi, Geomatik Mühendisliği Bölümü, Ankara, Türkiye

#### ÖZ

**Anahtar Kelimeler:**  
Hausdorff mesafesi  
Gönüllü coğrafi bilgi  
TOPOVT  
OSM  
ArcGIS

Bu makalenin amacı, OpenStreetMap(OSM)'nin mekânsal doğruluğunu Türkiye Topoğrafik Vektör Veri Tabanı (TOPOVT)'ye göre 'büyük bina' katmanında analiz etmektir. Açık kaynaklı bir haritalama platformu olan OSM'ye isteyen herkes erişebilir. Bağlayıcı bir standardı olmadığından, OSM'nin mekânsal doğruluğunun değerlendirilmesi açık bir araştırma konusudur. Diğer taraftan TOPOVT, Harita Genel Müdürlüğü tarafından standart bir prosedür izlenerek üretilmektedir. TOPOVT'yi güncellemek masraflı bir işlemdir ve ancak belirli zaman aralıklarında gerçekleştirilebilmektedir. TOPOVT'de güncellenecek yerlerin otomatik tespiti toplam maliyeti düşürecek etkin bir işlem olacaktır. Bu işlemin ön koşulu da OSM verisinin mekânsal doğruluğunu analiz etmektir. Dolayısıyla, bu makalede OSM ve TOPOVT'deki eşleşen poligonlar arasındaki Hausdorff mesafesini hesaplayarak 'bina' katmanının mekânsal doğruluğunun belirlenmesi hedeflenmiştir. Önerilen metodoloji kapsamında eşleşen poligonları tespit etmek için iki farklı yöntem değerlendirilmiştir: 'örtüşme yöntemi' ve 'merkez tabanlı yöntem'. Açık kaynak olarak geliştirilen yazılım iki farklı sahada test edilmiştir. Sonuçlar OSM kalitesinin TOPOVT ile iyi bir şekilde eşleşebileceğini göstermiştir. Tüm süreç bir ArcMap uzantısına entegre edilmiştir ve geliştirilen kod GitHub'da paylaşılmıştır.

### Spatial Accuracy Analysis of Buildings in OpenStreetMap

#### ABSTRACT

**Keywords:**  
Hausdorff Distance  
Volunteered Geographic  
information  
TOPOVT  
OSM  
ArcGIS

The aim of this paper is to assess the spatial accuracy of OpenStreetMap (OSM) with respect to the Turkey Topographic Vector Database (TOPOVT) within the context of 'building' layer. Being an open platform, anyone can access to OSM. Since there are no stringent standards, spatial accuracy assessment of OSM is an open research area. TOPOVT, on the other hand, is produced by the General Directorate of Mapping. Updating this database is a costly process. Therefore, automatic detection of the locations requiring update in TOPOVT would be an effective operation, which would eventually reduce the overall cost of update. However, the spatial accuracy of the geographical features must be analyzed in order to support such a motivation. Therefore, the aim of this paper is to assess the spatial accuracy of 'building' layer by calculating the Hausdorff distance between the matching polygons in OSM and TOPOVT. The proposed methodology consists of two methods to detect the matching polygons: 'overlap method' and 'centroid method'. The results indicate that the quality of OSM could well match with TOPOVT. The whole process is integrated into an ArcMap extension and the code is available on GitHub.

#### \* Sorumlu Yazar

(kadir.kucuk@hacettepe.edu.tr) ORCID ID 0000-0003-1788-2540  
\*(banbar@hacettepe.edu.tr) ORCID ID 0000-0003-2331-6190

## 1. GİRİŞ

Gönüllü Coğrafi Bilgi (GCB), mekânsal bilişim alanında son birkaç yılda önemi hızla beliren konulardan biridir. İnternete erişim kolaylığı ve mobil cihazların geniş kullanımı, milyonlarca insanın coğrafi etiketli verilerini tüm dünyayla paylaşmasına ve başkalarının paylaştıklarından faydalanmasına vesile olmaktadır (Anbaroğlu, 2017). Böylece, internet mekânsal verinin hem üretildiği hem de paylaşıldığı bir merkez durumuna gelmiştir (Goodchild, 2007). GCB'deki ilerleme, 'Açık Bilim', 'Açık Veri' ve 'Sivil Bilim'deki ilerlemeyle etkin bir uyum içindedir; çünkü burada adı geçen tüm farklı yaklaşımlardaki ortak payda; üretilen veriden, bilimsel yayınlara erişime kadar bilimsel süreçteki tüm adımları şeffaf ve erişilebilir kılmaktır (Haklay, 2013; Sevinç ve Karas, 2018).

GCB'nin pratik ve bilimsel amaçlarla kullanılmasının çeşitli faydaları vardır (Feick ve Roche, 2013). İlk olarak, veri toplamanın ekonomik maliyeti önemli ölçüde azaltılabilir. Yüzlerce hatta binlerce insanın katkısı, veri toplamaya harcanan zamanı ve parayı ciddi oranlarda azaltabilir. Büyük şirketler zaten bir süredir GCB'yi etkin bir şekilde kullanmaktadırlar. Örneğin, Google'da ziyaret ettiğiniz lokantalar veya müzeler için bir değerlendirme yapma imkanının kolaylıkla sunulması sayesinde binlerce gönüllü vatandaş tarafından farklı mekânların popülerliği değerlendirilebilmektedir. Benzer şekilde devlet kurumları da GCB'den faydalanarak haritalandırma projelerini daha kısa sürelerde ve daha az maliyetle gerçekleştirilebilmektedir (Çabuk, 2015; Taşkanat vd., 2018).

Operasyonel amaçlar için GCB'ye güvenmenin ikinci avantajı da GCB ile üretilen verinin güncel olmasıdır. Bir mahallede yaşayan insanlar, o mahalleyle ilgili en güncel durumu zamanında alabilirler. Örneğin, OSM'de bulunan; ancak Münih'in referans veri seti olan ATKIS'te bulunmayan 1200'den fazla yeni bina inşa edildiği tespit edilmiştir (Fan vd., 2014). Bunun yanında, gecekondu mahallelerinin ve gayri resmi yerleşimlerin izlenmesi yoluyla ve kentsel dönüşüm amacıyla da GCB ile üretilen veri değerlendirilmiştir (Hachmann vd., 2018). Günümüzde karşılaşılan birçok durumda, resmî kurumlar tarafından tam anlamıyla güncel verilerin elde edilmesinin mümkün olmayacağı ve ancak GCB'ye katkı veren gönüllü vatandaşların katkılarıyla daha güncel veriyi elde etmenin mümkün olacağı anlaşılmaktadır. Neticede, araştırmacılar hava kalitesini izleyen sensörlerin konumlarının tespitinden (Gupta vd., 2018), salgın hastalıkların önlenmesine kadar (Qi vd., 2018) birçok alanda GCB'nin güncel olmasından fayda sağlamaktadır.

Birçok farklı araştırma alanında GCB yoğun olarak kullanılmaktadır. Ancak, veri kalitesi GCB'nin resmî olarak kullanılmasının önündeki temel endişe kaynağı olmaya devam etmektedir. Öncelikle, veri gönüllü

kişilerce toplanmaktadır ve hemen herkes GCB'ye katkıda bulunabilmektedir. Kapsayıcı bir süreç olmanın sonucu olarak, GCB'de veri toplama süreci ve dolayısıyla veri kalitesinin temin edilmesi yönünde belirli bir standart bulunmamaktadır. Anket, metin, görüntü veya harita tabanlı GCB'nin veri kalitesinin değerlendirilmesi için çeşitli yöntemler önerilmiştir (Senaratne vd., 2017). Buradan da anlaşılacağı üzere, GCB'nin veri kalitesinin değerlendirilmesi üzerine bilimsel çalışmalar olgunlaşmış durumdadır.

Mevcut araştırmalar, gönüllüler tarafından toplanan verilerin yetkili kurumlarca üretilmiş veri ile eşleşebileceğini ve resmi amaçlar için kullanılabilirliğini göstermektedir (Haklay, 2010). Milyonlarca gönüllünün açık bir şekilde dünyayı haritalamak için işbirliği yaptığı OpenStreetMap (OSM) için bu durum özellikle geçerlidir (Brovelli ve Zamboni, 2018). GCB'yi kullanmanın faydaları ve temel endişe kaynağı olan veri kalitesinin de gerçek bir sınırlama olmadığı görülmektedir. Dolayısıyla, Amerika Birleşik Devletleri Jeolojik Araştırma (USGS) Coğrafi Bilgi Bilimi için Mükemmeliyet Merkezi'nin de GCB'ye güvendiğini görmek şaşırtıcı değildir (2006 yılından beri devam eden CEGIS programı (USGS, 2019).

Bu makale, OpenStreetMap'te bina ayak izlerinin mekânsal doğruluğunu, Harita Genel Müdürlüğü tarafından üretilen referans veri seti olan Türkiye Topografik Vektör Veri tabanı (TOPOVT) ile karşılaştırarak analiz edilmesini amaçlamaktadır. Eşleşen (homolog) poligonlar, 'Örtüşme Yöntemi' ve 'Merkez Tabanlı Yöntem' olarak adlandırılan iki farklı yaklaşım ile tespit edilmektedir. Sonrasında ise, eşleşen OSM binası ile TOPOVT binası arasındaki Hausdorff Mesafesi tespit edilerek, OSM binasının doğru olarak kabul edilen TOPOVT'ye uzaklığı belirlenmektedir (Avbelj vd., 2015). Geliştirilen metodolojiyi gerçek hayatta da uygulayabilmek için bir ArcGIS eklentisi geliştirilmiştir ve açık olarak paylaşılmaktadır (Küçük, 2019).

Bu makale beş bölümden oluşmaktadır. Yazın araştırması ikinci bölümde sunulmaktadır. Eşleşen binaların tespiti için incelenen iki yöntem ise üçüncü bölümde anlatılmaktadır. Makalede kullanılan veri setlerinin tanıtımı ve elde edilen sonuçlar dördüncü bölümde sunulmaktadır. Son olarak, bu makale kapsamında elde edilen kazanımlar ve gelecek araştırma alanları beşinci bölümde takdim edilmiştir.

## 2. LİTERATÜR TARAMASI

Mekânsal veri kalitesi, bu kapsamda geliştirilmiş olan uluslararası standartlar nedeniyle (örneğin ISO 19157:2013) iyi çalışılmış bir konudur. Bütünlük, komisyon/ihmal hataları, mantıksal tutarlılık ve mekânsal/zamansal/tematik doğruluk gibi farklı yönleri kapsar. Bununla birlikte, 'kullanıma uygunluk' 20 yılı aşkın bir süredir geçerli bir veri kalitesi parametresi olarak kabul edilmiştir. Özne olmakla da

birçok uygulama için en önemli parametre olarak değerlendirilmektedir (Veregin, 1999).

OSM'nin veri kalitesi değerlendirmesi iki ana nedenden dolayı devam eden bir araştırma alanıdır. Birincisi, OSM'ye yapılan katkılar artmakta ve bazı Avrupa kentleri kapsamlı bir şekilde haritalanmıştır. İkinci olarak, OSM'ye herhangi biri veri üretim sürecine doğrudan katkıda bulunabilmektedir. Sonuç olarak, GCB yoluyla oluşturulan haritaların gerçeğe ne ölçüde karşılık geldiğini değerlendirmek faydalı olacaktır. ISO 19113 veya ISO 19157 gibi uluslararası standartlar olsa da araştırmacılar OSM'nin veri kalitesini değerlendirmek için iki ana stratejiye güvenmektedirler: içsel ve dışsal yöntemler (Barron vd., 2014).

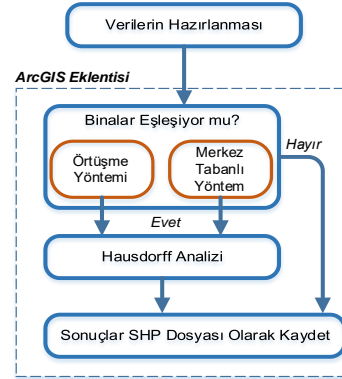
İçsel yöntemler bir coğrafi nesnenin zaman içinde nasıl geliştiğini inceleyerek, veri kalitesini belirlemek için yöntemler geliştirir. Bir başka ifade ile, içsel yöntemlerde doğru kabul edilen bir referans veri seti bulunmamaktadır. İçsel yöntemler üzerinden veri kalitesinin belirlenmesinin ardındaki temel mantık, referans veri kümelerinin genellikle yüksek maliyetlerle elde edilebilmesi veya kısıtlayıcı lisanslara sahip olmasıdır. Bu nedenle, araştırmacılar veri kalitesini tahmin etmek için verilerin kendisine ve geçmiş kayıtlara güvenirlere. Örneğin, Haklay vd. (2010) 'Linus' Yasası'nın mekânsal doğruluk değerlendirmesi için geçerli olduğunu doğrulamıştır. Bir başka ifadeyle, belirli bir mekânsal nesneyi veya bölgeyi haritalamak için gönüllülerin sayısı arttıkça, mekânsal doğruluk da artacaktır. Başka bir araştırma, öznitelik bütünlüğü için bir gösterge sağlamak amacıyla ev sayısı / adı ile binaların toplam sayısına oranı araştırılmıştır. Ancak, Barron vd. (2014)'nin belirttiği gibi veri kalitesinin nesnel kriterler çerçevesinde belirlenebilmesi için yüksek kaliteli referans veri setine ihtiyaç duyulmaktadır.

Dışsal yöntemler, referans veri setlerinin varlığını kabul eder. Sonuç olarak, GCB ile üretilen coğrafi verinin kalitesinin analiz edilmesi halen araştırmacıların ilgisini çeken bir araştırma alanıdır. En büyük GCB projelerinden biri olan OSM'nin de veri kalitesinin doğru bir tespit edilebilmesi için referans veri setine ihtiyaç duyulmaktadır. Bu tür referans veri setleri de genellikle veri üretim konusunda yetkili kurumlar tarafından oluşturulur. Güncel araştırmalar hem içsel hem de dışsal yöntemlerin birlikte değerlendirilmesinin, her iki stratejinin de sınırlamalarını aşabileceğini belirtmektedir (Touya vd.,2017).

### 3. METODOLOJİ

Bu makalenin amacı, OSM'deki binaları Türkiye'deki resmî verilerle karşılaştırarak mekânsal doğruluğunu incelemektir. Bu kapsamda resmî veri olarak Harita Genel Müdürlüğü (HGM) tarafından üretilen TOPOVT referans veri seti olarak kabul edilmiştir. Bunu yaparken de, çalışmanın gerçek

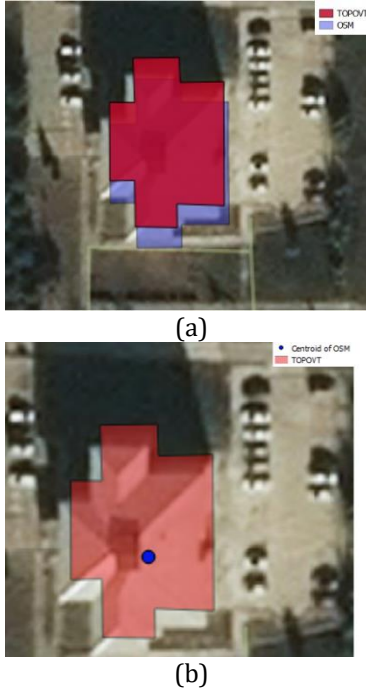
hayatta uygulanabilir olmasını sağlamak da hedeflenmiştir ve geliştirilen metodoloji bir ArcGIS eklentisi olarak geliştirilmiştir. ArcGIS eklentisi geliştirmenin temel nedeni de TOPOVT'nin ArcGIS üzerine inşa edilmiş olmasıdır. Makalenin metodolojisi Şekil 1'de gösterilmiştir. Önerilen metodolojide, önerilen bina eşleşme yöntemlerinin ve aynı zamanda veri setlerinin daha iyi anlaşılması için OSM'de bir veri temizleme işlemi gerçekleştirilmemiştir.



Şekil 1. Araştırmanın Metodolojisi

İlk olarak, 'Veri Hazırlama' adımıyla, TOPOVT verileri, test bölgesine karşılık gelen 128 özellik veri katmanını içeren bir geodatabase (.mdb) dosyası olarak sağlanmaktadır. Bu katmanlar içinde 'büyük bina' katmanı ile OSM'deki binalar karşılaştırılarak, bu binaların mekânsal doğruluğunun tespit edilmesi sağlanmıştır. OSM'deki bina poligonlarının elde edilmesinin en kolay yolu QGIS'teki ilgili işlevselliği kullanmaktır. Özellikle, belirlenen araştırma alanının OSM verilerini indirmek için 'OpenStreetMap Verilerini Arama ve İndirme' başlıklı eğitim içeriği takip edilebilir (QGIS, 2019). Son olarak, her iki veri setinin de aynı coğrafi referans sistemine sahip olması için gerekli koordinat sistemi dönüşümü gerçekleştirilmelidir. Her iki veri seti de başlangıçta WGS 84 datum ve coğrafi koordinatlarda 4326 SRID'ye sahip olarak kaydedilmiştir. Ancak, Hausdorff mesafesi bir metrik sistem gerektirdiğinden, her iki veri seti de SRID: 32636'ya dönüştürülür. Böylece, veri setleri 36N UTM bölgesinde kalmakta olup, test sahalarını uygun bir şekilde temsil etmektedir.

Metodolojinin ana bileşeni OSM ve TOPOVT'de eşleşen binaları otomatik bir şekilde tespit etmektir. Eşleşen bir bina, OSM ve TOPOVT'deki aynı uzamsal nesneyi temsil ettiği kabul edilmektedir. Eşleşen binaları tespit etmek için iki farklı yaklaşım önerilmiştir. 'Örtüşme Yöntemi'nde, iki poligon kesişiyorsa, o iki poligonun eşleştiği kabul edilir. 'Merkez Tabanlı Yöntem'de ise, OSM poligonunun merkezi TOPOVT poligonunun içinde kalıyorsa, iki poligonun eşleştiği kabul edilmektedir. Bu iki yaklaşımın bir örneği Şekil 2'de gösterilmiştir. Hecht vd. (2013) OSM verilerinin bütünlüğünü değerlendirmek için benzer bir metodolojiyi uygulamışlardır.



**Şekil 2.** Eşleşen poligonların, (a) örtüşme yöntemi ve (b) merkez tabanlı yöntem kullanarak tespiti

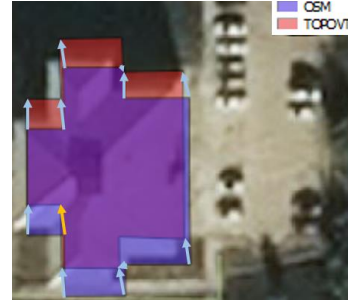
Eşleşen poligonlar belirlendikten sonra, Hausdorff mesafesi ile aralarındaki mesafe ölçülebilmektedir. Hausdorff mesafesi, eşleşen poligonların köşe koordinatlarını temsil eden iki nokta kümesi arasındaki ölçüdür. En yakın köşe noktaları arasındaki maksimum mesafeyi belirler. Poligonlar (ve ayrıca çizgiler) bir nokta kümesi olarak kabul edilebilir, bu tür coğrafi nesnelerin birbirine yakınlığını ölçmek için kullanılacak bir yöntemdir. Hausdorff mesafesi ne kadar düşükse, eşleşen poligonlar da birbirine o kadar yakın olduğu söylenebilir. Hausdorff mesafesi Denklem 1’de gösterildiği gibi hesaplanır.

$$H(\mathbf{A}, \mathbf{B}) = \max_{a \in \mathbf{A}} \left( \min_{b \in \mathbf{B}} (d(a, b)) \right) \quad (1)$$

**Denklem 1.** Hausdorff mesafesi

Denklem 1’de aynı coğrafi nesneyi temsil eden ve eşleşen iki poligon A ve B olarak gösterilmektedir. Her iki poligon da sırasıyla m ve n noktadan oluşan noktalar kümesidir. Poligon A  $\{a_1, a_2, \dots, a_m\}$  noktalarından, poligon B ise  $\{b_1, b_2, \dots, b_n\}$  noktalarından oluşmaktadır. Her iki veri kümesi de aynı binayı temsil ettiği kabul edildiğinden, köşe noktaları arasındaki Öklid uzaklığı hesaplanır ve bu da  $d(a, b)$  olarak gösterilmektedir. Her poligonun içerdiği nokta sayısı arasında matematiksel bir ilişki olmadığı unutulmamalıdır. Bir başka ifadeyle  $m \geq n$  veya  $m \leq n$  olabilir. Hausdorff mesafesi tek yönlü bir mesafe ölçüsü olsa da kolayca çift yönlü bir ölçüye dönüştürülebilir (Schlesinger vd., 2014).

Bu araştırmanın amacı OSM binalarının mekânsal doğruluğunu değerlendirmek olduğundan, A ve B poligonları sırasıyla OSM ve TOPOVT’deki poligonları temsil etmektedir. Mesafe ölçüsünün görsel tasviri Şekil 3’de gösterilmiştir. OSM poligonunun her tepe noktası için, TOPOVT’deki en yakın komşusu başlangıçta ve daha sonra bu mesafelerin maksimum kısmında Hausdorff mesafesi olarak kabul edilir.



**Şekil 3.** Hausdorff mesafesinin görsel tasviri

Örtüşme yöntemi kullanıldığında, OSM’deki bir poligon ile TOPOVT’de bir poligonunun eşleşeceği kabul edilmiştir. Ancak, kimi durumlarda birden çok OSM poligonuyla tek bir TOPOVT poligonu da eşleşebilmektedir. Böyle bir durumda, örtüşme yöntemi sadece en düşük Hausdorff mesafesine sahip olanların eşleşen olduğunu varsaymaktadır. Başka bir deyişle, sadece bire bir (1:1) ve en düşük Hausdorff mesafesine sahip olanların bu çalışmada eşleştiği kabul edilmektedir. Elde edilen analiz sonuçlarının tekrarlanabilirliğini sağlamak ve çalışmanın sürdürülebilir bir yapıda gelişmesini sağlamak için, geliştirilen tüm kod ve test verileri projenin GitHub sayfasında mevcuttur (Küçük, 2019).

Metodolojinin son adımı olarak da geliştirilen metodolojinin pratikte uygulanmasını desteklemek için bir adım atılmıştır ve bir ArcGIS eklentisi geliştirilmiştir. ArcGIS’in yeteneklerinin daha etkin bir şekilde kullanılabilmesi ve geliştirilebilmesi için ArcPy adlı Python kütüphanesi kullanılmaktadır. Her ne kadar birçok fonksiyon ArcMap’ın grafik arayüzünden erişilebilir olsa da, iş süreçlerini etkin bir şekilde modellemek ve yönetebilmek için ArcMap üstüne eklenti geliştirilebilmektedir. Bir poligonun merkezini bulmak veya iki poligonun kesişip kesişmediğini saptamak da dahil olmak üzere tüm konumsal sorgular ArcPy’da yapılabilmektedir. Bu kapsamda geliştirilen metodoloji ArcPy’da uygulanarak, bir ArcGIS eklentisi geliştirilmiştir. Eklenti çalıştırdıktan sonra, OSM’de bulunan ancak TOPOVT’de bulunmayan poligonları, bir başka deyişle güncelleştirme gerektirebilecek alanları belirtmek için ayrı bir SHP dosyası da kaydedilmektedir.

OSM’nin lisansı, ‘Open Data Commons Open Database License (ODbL)’ bu kapsamda vurgulanmalıdır. Bu lisans, OSM ve katkıda bulunanlar referans verildiği sürece herkesin OSM verilerini kendi amaçları doğrultusunda kullanmasına, dağıtmasına ve

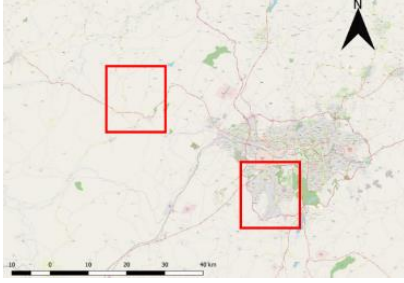


uyarlamasına olanak tanımaktadır. Buna ek olarak, OSM verilerini kullananlar da üretilen haritaları aynı lisans altında dağıtmak zorundadır (Brovelli ve Zamboni 2018; OSM, 2019).

#### 4. ANALİZ VE SONUÇLAR

##### 4.1 Veri Setleri

OSM'deki binaların mekânsal doğruluğunu analiz etmek için geliştirilen metodoloji iki farklı bölgede incelenmiştir. Her iki bölge de Şekil 4'de gösterilmiştir ve bu bölgeler Ankara il sınırları içinde kalmaktadır. Bu bölgelerden biri kırsal çevreye, diğeri ise kentsel çevreye karşılık gelmektedir. Her iki bölge de yaklaşık 150 km<sup>2</sup>'lik bir alana karşılık gelmektedir. Kuzeyde yer alan bölge kırsal çevreyi, diğeri bölge ise şehri temsil etmektedir.



Şekil 4. Çalışma alanı

OSM'deki binaların kalitesi, Harita Genel Müdürlüğü (HGM) tarafından üretilen referans veri seti ile karşılaştırılarak değerlendirilmektedir. Türkiye'nin 1:25000 ölçekli topoğrafik haritalarının üretiminden sorumlu ulusal haritalama kurumu HGM'dir. HGM tarafından üretilen referans veri seti, Türkçe'de 'Türkiye Topoğrafik Vektör Veri Tabanı'nın kısaltması olan 'TOPOVT' olarak adlandırılmaktadır. Bölgelerin hava fotoğrafları 2011 yılında çekilmiş ve fotogrammetrik değerlendirme 2012 yılında gerçekleştirilmiştir.

TOPOVT stereo hava fotoğraflarının derlenmesi yoluyla üretilir. Veri tabanında yollardan binalara ve mezarlıktan parklara kadar 128 farklı özellikte veri seti bulunmaktadır. Bu çalışma kapsamında TOPOVT'deki 'Büyük Bina' katmanı incelenmiştir. Geometrik doğruluk TOPOVT için yatay ve dikey bileşenlerde  $\pm 3$  m'dir. Böyle bir sistemi canlı ve güncel tutmak, hükümet organları ve vatandaşlar da dahil olmak üzere tüm paydaşların ortak çalışmasını gerektirmektedir (Yılmaz ve Canıberk, 2018). Sonuç olarak, OSM'nin TOPOVT'nin güncelleme sürecine entegre edilebilmesi pratik açıdan önem arz eden, ancak zorlu bir araştırma hedefi olarak belirmektedir. Veri kümeleri ve çalışma alanları arasındaki genel karşılaştırma Tablo 1'de gösterilmiştir.

Genel karşılaştırma birkaç önemli hususu belirtmektedir. Birincisi, OSM'de şehir bölgesinde TOPOVT'ye göre daha fazla poligon vardır. Bunun iki

ana nedeni var. İlk olarak, TOPOVT'nin temel verileri, yani havadan çekilen görüntüler 2011 yılında yakalanırken, OSM'deki verilerin güncel hali kullanılmıştır. İkinci olarak, OSM kayıtları sadece tekil binalar değil, aynı zamanda bina bloklarını da içermektedir.

**Tablo 1.** Veri setlerinin ve çalışma alanlarının karşılaştırılması

	Şehir		Kırsal	
	OSM	TOPOVT	OSM	TOPOVT
Bina sayısı	6404	1123	20	123
Min alan (m <sup>2</sup> )	9.56	2.01	93.66	73.14
Maks alan (m <sup>2</sup> )	84740	16496	92383	6941
Ort alan (m <sup>2</sup> )	1031	778	13065	606
Std Sap (m <sup>2</sup> )	4102	885	27110	828

Kırsal alanda ise, TOPOVT'de OSM'ye göre daha fazla bina bulunduğu gözlemlenmiştir. Kentsel alanların kırsal alanlara göre OSM'de daha iyi haritalanmış olduğu gerçeği önceki çalışmaları destekler niteliktedir (Hecht vd., 2013). Ayrıca, OSM'de mevcut tüm poligon verisi değerlendirildiği ve herhangi bir veri temizleme işlemi uygulanmadığından, incelenen poligonlar arasındaki büyüklük farklılık dikkati çekmektedir. Bu heterojen durumun analiz sonuçlarını nasıl etkilediği bölüm 4.3'de detaylı bir şekilde tartışılmıştır.

##### 4.2 Sonuçlar

Önerilen metodoloji hem kentsel hem de kırsal çalışma bölgesinde uygulanmıştır. İlk olarak, eşleşen binalar merkez tabanlı ve örtüşme yöntemleri kullanılarak tespit edilmiştir. Eşleşen binalar için, Hausdorff uzaklığı (H) hesaplanmıştır. Genel sonuçlar Tablo 2'de gösterilmiştir.

**Tablo 2.** Genel sonuçlar

	Şehir		Kırsal	
	Merkez Örtüşme Tabanlı	Örtüşme	Merkez Örtüşme Tabanlı	Örtüşme
Eşleşen bina sayısı	595	608	6	4
Ort H (m)	9.65	9.68	61.59	5.51
Std. sap. H (m)	13.15	17.99	98.58	2.29
Maks H (m)	93.21	247.90	274.50	8.80
Min H (m)	0.53	0.53	2.61	2.61
İşlem Süresi (dak.)	11.00	14.00	0.03	0.03

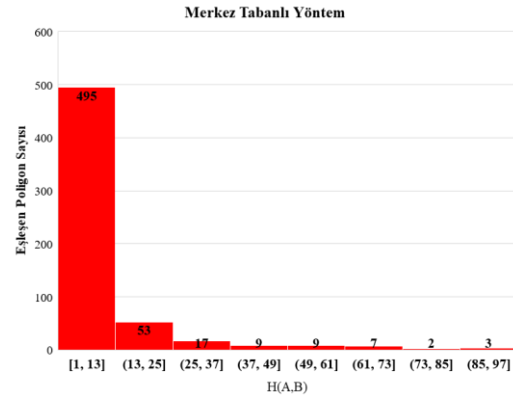
Analiz sonuçları incelendiğinde çeşitli tespitler yapılabilir. İlk olarak, beklendiği gibi, kırsal alanda çok az sayıda eşleşen bina bulunmaktadır. İkinci olarak, 'merkez tabanlı yöntem'de Hausdorff uzaklık değerlerinin standart sapması, 'örtüşme yöntemin'den daha düşüktür. Dolayısıyla, 'merkez tabanlı yöntem'in daha güvenilir olduğu söylenebilir.

Hata miktarlarının anlamlı OSM poligonlarının analiz edilmesiyle daha düşürülebileceği vurgulanmalıdır. Daha önceden de belirtildiği gibi, gerek önerilen bina eşleşme yöntemlerinin gerekse OSM verisinin daha iyi anlaşılması için, analizden önce OSM’de bir veri temizleme işlemi gerçekleştirilmemiştir. OSM’de tekil bir bina yapısına ek olarak, bir parseli veya siteyi temsil eden büyük poligonları da gözlemlemek mümkündür. Böyle büyük bir poligonun merkezinin bir TOPOVT binası içinde kalması da muhtemeldir. Böyle durumlar da elde edilen Hausdorff uzaklıklarını büyük ölçüde arttırmaktadır.

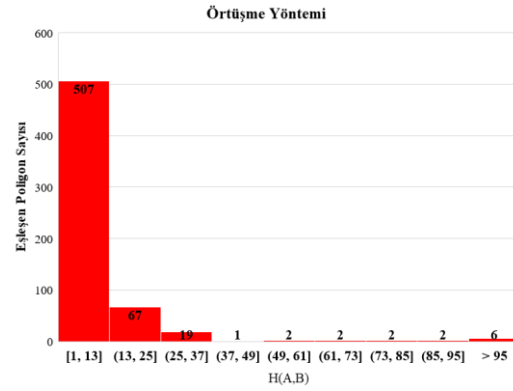
Şehir bölgesinde, kırsal çevreye göre, çok daha fazla bina eşleşmesi tespit edildiği için ‘örtüşme’ ve ‘merkez tabanlı’ yöntemlerin daha anlamlı bir şekilde karşılaştırılması sağlanabilir. Merkez tabanlı yöntemde 595 bina eşleşmişken, örtüşme yönteminde bu sayı biraz daha fazladır. Ortalama Hausdorff uzaklıkları incelendiğinde ise, merkez tabanlı yöntemin 9,65 metre ile örtüşme yöntemine göre çok az bir farkla daha iyi olduğu görülmüştür. Maksimum Hausdorff mesafesi ve mesafelerin standart sapması incelendiğinde, yöntemler arasındaki temel fark belirginleşmektedir. Her iki analiz sonucunda da ‘merkez tabanlı’ yöntemin daha başarılı olduğu gözlemlenmektedir. Ayrıca, ‘merkez tabanlı’ yöntem daha hızlı da çalıştığı için örtüşme yönteminden daha başarılı olduğu belirtilebilir. Daha çok sayıda bina içeren kentsel bir ortamda, burada belirtilen avantajlarından ötürü merkez tabanlı yönteminin eşleşen binaların mekânsal doğruluğunu belirlemek için tercih edilebileceği düşünülmektedir. Merkez tabanlı ve örtüşme yöntemleri için Hausdorff mesafelerinin histogramları Şekil 5’de gösterilmiştir.

Histogramların analiz edilmesiyle ilgi çekici başka gözlem yapılabilir. Örtüşme yöntemi ile düşük Hausdorff mesafesine sahip daha çok eşleşen bina olduğu görülmektedir. Bir başka deyişle, 507 eşleşen binanın Hausdorff uzaklığı 1 ve 13 metre arasındadır. Merkez tabanlı yöntemde ise, aynı Hausdorff uzaklık aralığı için bu sayı 495’e düşmektedir. Bu nedenle, örtüşme yönteminin birkaç yüksek hatalı eşleşmesi sonucunda yukarıda belirtilen çıkarım yapılmaktadır.

Her iki yöntemde de gözlemlenen yüksek hatalı eşleşmelerin temel nedeni de OSM veri setinde herhangi bir ön temizleme işleminin gerçekleştirilmemiş olmasıdır. Bu kapsamda yapılacak çalışmalar ile yüksek hatalı eşleşmelerin önüne kolaylıkla geçilebileceği ön görülmektedir.



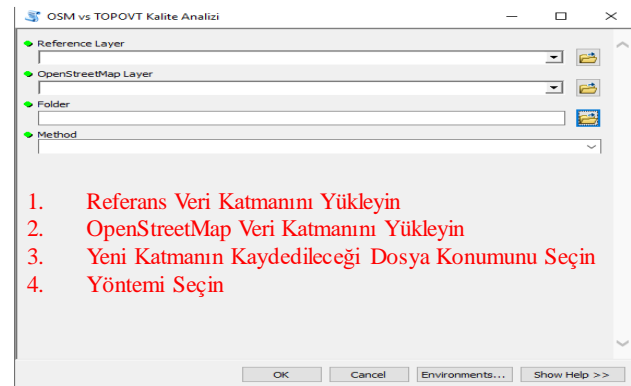
(a)



(b)

Şekil 5. Hausdorff mesafelerinin (a) merkez tabanlı yöntem ve (b) örtüşme yöntemine göre histogramları

Son olarak, geliştirilen kod dört girdisi olan ve Şekil 6’da gösterilen bir ArcGIS uzantısına entegre edilmiştir. İlk iki girdi sırayla TOPOVT ve OSM katmanlarını belirtmektedir. Üçüncü girdi de oluşturulan çıktı dosyalarının kaydedileceği dizin belirlenmektedir. Dördüncü girdi ise, hangi eşleşme yönteminin kullanılacağını belirtmektedir.



Şekil 6. ArcGIS eklentisi

Geliştirilen eklentinin kullanımı kolaydır. Üçüncü adımda seçilen dizine iki farklı dosya kaydedilecektir: OSM’de bulunan; ancak TOPOVT’de bulunmayan binalar bulunmaktadır. Bunun yanında eşleşen binalar ve aralarındaki Hausdorff uzaklığını gösteren katman da çıktı olarak kullanıcıya yansıtılacaktır.

### 4.3 Durum Analizi

Yöntemlerin etkinliğini daha iyi anlamak için bu bölümde maksimum Hausdorff mesafesinin hangi durumlarda oluştuğu incelenecektir. Kentsel çevrede maksimum Hausdorff mesafesinin 247.90 olmasına neden olan durum Şekil 7’de gösterilmiştir. Büyük OSM poligonu batı sınırında tek bir TOPOVT binası ile örtüşmektedir. Diğer poligonlar ise OSM’nin doğrudan içinde olduğu için örtüşme yoktur. Bu tek kesişen poligon nedeniyle, hesaplanan Hausdorff mesafesi yüksektir.



Şekil 7. Şehir bölgesindeki maksimum Hausdorff mesafesi

Bu örnek aynı zamanda OSM’nin heterojen doğasını da göstermektedir, çünkü poligonlar gönüllüler tarafından kaydedilmektedir ve bir site alanı veya parsel gibi büyük bir alan ‘bina’ olarak etiketlenebilir. Benzer bir durum, merkez tabanlı yöntem için de geçerlidir. Şekil 8’de gösterildiği gibi bir durumda örtüşme yönteminde büyük hatalar gözlemlenebilir.



Şekil 8. Kırsal bölgedeki maksimum Hausdorff mesafesi

Bu bağlamda üç OSM poligonu vardır. Diğer iki poligonu içeren en büyük poligonun merkezi bir TOPOVT binasının içindedir. Bu nedenle, merkez tabanlı yöntemde bu iki poligonun eşleşip Hausdorff mesafesinin buna göre hesaplandığı varsayılmıştır. Benzer şekilde, ikinci büyük OSM poligonunun merkezi de haritanın güney kısmında gösterildiği gibi bir

TOPOVT binası içinde bulunmuştur. Sadece birkaç eşleşen poligon olduğundan, bu tür büyük uzaklıklar genel sonuçları da etkilemektedir.

### 5. SONUÇ

Günümüzde GCB ile birlikte gönüllü kişiler bir haritanın üreticisi konumuna da gelmişlerdir. Her ne kadar bu gelişme ile birlikte bir haritanın güncelliği daha az maliyetle sağlanabiliyorken, veri kalitesi ayrı bir araştırma alanı olarak belirmektedir. Nitekim, güncel verinin ne kadar kullanışlı olabileceği ancak veri kalitesi üzerindeki şüpheler giderildikten sonra belirlenebilir. Bu makalenin amacı da önemli GCB projelerinden biri olan OSM’nin bina poligonları bazında mekânsal doğruluğunu değerlendirmektedir. Bunun için, OSM poligonları ile referans veri seti olan TOPOVT’nin ‘büyük bina’ poligonları karşılaştırılarak, eşleşen poligonlar arasındaki Hausdorff uzaklığı hesaplanmıştır. Araştırma kapsamında merkez tabanlı yöntem ve örtüşme yöntemi olarak adlandırılan iki yöntemin eşleşen bina bulma kapsamında başarıları analiz edilmiştir. Merkez tabanlı yöntemde bir OSM poligonunun merkezi TOPOVT poligonunun içindeyse, iki binanın eşleştiği kabul edilmiştir. Örtüşme yönteminde iki binanın eşleşmesi içinse, ilgili poligonların kesişmesi gerekmektedir.

Deneyler Ankara il sınırları içinde kalan kentsel ve kırsal olmak üzere iki farklı bölgede gerçekleştirilmiştir. Kentsel bölge ile ilgili sonuçlar, ortalama olarak, her iki yöntemin eşleşen binaları benzer sayıda algıladığını göstermektedir. Referans veri setinde bulunan 1123 binadan yaklaşık 600’ü her iki yöntemle de eşleştiği kabul edilmektedir. Benzer şekilde, eşleşen binalar için Hausdorff mesafesi hesaplanmaktadır. Bu mesafenin ortalama 9,5 metre olduğu tespit edilmiştir. Merkez tabanlı yöntemin daha hızlı olduğu ve Hausdorff mesafelerinin daha düşük olduğu düşünülürse, ileride yapılacak çalışmalarda merkez tabanlı yöntemin kullanılması daha etkili olabilir. Ancak, yöntemler arasındaki fark yalnızca sınırdan olduğundan, yapılan çıkarım bir gözlem niteliğindedir ve farklı test sahalarının da analiz edilmesiyle daha doğru bir sonuç elde edilecektir. Ayrıca merkez tabanlı yöntemin kırsal alanda başarısız olduğu da unutulmamalıdır.

Bu araştırma aynı zamanda kentsel-kırsal ayrımı ve kırsal alanların GCB açısından gelişmemiş olduğu ile ilgili mevcut araştırma sonuçlarını desteklemektedir. Her ne kadar OSM’de güncel veriler kullanılmış olsa da ve referans veri setinin 2011 yılındaki duruma göre sayısallaştırıldığı bilirse de poligon sayıları arasındaki büyük farklılık OSM’nin kırsal alanlarda aslında pek de güncel olamayabileceğinin en büyük göstergesidir. Dolayısıyla, kırsal alandaki çok düşük eşleşen bina poligonu nedeniyle, anlamlı bir sonuca varmak zor olsa da mevcut sonuçlar örtüşme yönteminin bu bölgede daha başarılı olduğunu göstermektedir. Daha güvenilir sonuçlar elde etmek için kırsal alanlarda GCB’nin

içeriğinin zenginleştirilmesiyle birlikte, daha doğru çıkarımlar yapılabilir.

Bu makalede sadece bire-bir bina eşleşmeleri analiz edilmiştir. Ancak, özellikle örtüşme yönteminde, bir veri kümesindeki tek bir poligonun diğer veri kümesinde birden çok poligonla eşleşebileceği çeşitli durumlar gözlemek mümkündür. Bu gibi durumlarda, eşleşen poligonlardan Hausdorff mesafesi en düşük poligonların eşleştiği kabul edilmiştir. Sonuç olarak, gelecek çalışmalarda, bire-çoklu eşleşmelerin daha ayrıntılı olarak analiz edilmesi önem arz etmektedir. Buna ek olarak, veri kalitesinin daha doğru hesaplanabilmesi için, OSM’de veri temizliği işlemi gerçekleştirilebilir. Bina olmayan poligonların analizden çıkarılması ile hesaplanan Hausdorff uzaklıklarının düşeceği ön görülmektedir. Nitekim, OSM’de gerçekleştirilecek bir veri temizleme adımı ile bölüm 4.3’te incelenen durumlar önlenebilecektir. Farklı veri temizleme ve ön işleme stratejilerini değerlendirmek ilgi çekici bir araştırma alanı olacaktır. Son olarak, yollar gibi diğer coğrafi olguların da mekânsal doğruluğunun da incelenmesi ile daha bütüncül bir yaklaşımla OSM veri kalitesinin incelenmesi sağlanabilir.

## TEŞEKKÜRLER

TOPOVT verilerini sağladıklarından dolayı Harita Genel Müdürlüğü’ne teşekkür ederiz. Ayrıca, Dr. Alb. Altan Yılmaz ve Bnb. Mustafa Camberk’in geri bildirimleri için de müteşekkirimiz.

## REFERANSLAR

- Anbaroğlu, B. (2017). Gönüllü Coğrafi Bilgi: Mekânsal Bilişim Çalışmalarına Web 2.0 Devrinde Yeni Bir Yaklaşım. 9.
- Avbelj, J., Müller, R., & Bamler, R. (2015). A Metric for Polygon Comparison and Building Extraction Evaluation. *IEEE Geoscience and Remote Sensing Letters*, 12(1), 170–174. <https://doi.org/10.1109/LGRS.2014.2330695>
- Barron, C., Neis, P., & Zipf, A. (2014). A Comprehensive Framework for Intrinsic OpenStreetMap Quality Analysis. *Transactions in GIS*, 18(6), 877–895. <https://doi.org/10.1111/tgis.12073>
- Brovelli, M. A., & Zamboni, G. (2018). A New Method for the Assessment of Spatial Accuracy and Completeness of OpenStreetMap Building Footprints. *ISPRS International Journal of Geo-Information*, 7(8), 289. <https://doi.org/10.3390/ijgi7080289>
- Brovelli, M., & Zamboni, G. (2018). A New Method for the Assessment of Spatial Accuracy and Completeness of OpenStreetMap Building Footprints. *ISPRS International Journal of Geo-Information*, 7(8), 289. <https://doi.org/10.3390/ijgi7080289>

- Çabuk, S. (2015). Open Street Map Verilerinden Yararlanılarak 1/50K Ölçekli Harita Üretilebilirliğinin Araştırılması. 9.
- Fan, H., Zipf, A., Fu, Q., & Neis, P. (2014). Quality assessment for building footprints data on OpenStreetMap. *International Journal of Geographical Information Science*, 28(4), 700–719. <https://doi.org/10.1080/13658816.2013.867495>
- Feick, R., & Roche, S. (2013). Understanding the Value of VGI. İçinde D. Sui, S. Elwood, & M. Goodchild (Ed.), *Crowdsourcing Geographic Knowledge* (ss. 15-29). [https://doi.org/10.1007/978-94-007-4587-2\\_2](https://doi.org/10.1007/978-94-007-4587-2_2)
- Goodchild, M. F. (2007). Citizens as sensors: The world of volunteered geography. *GeoJournal*, 69(4), 211–221. <https://doi.org/10.1007/s10708-007-9111-y>
- Gupta, S., Pebesma, E., Degbelo, A., & Costa, A. C. (2018). Optimising Citizen-Driven Air Quality Monitoring Networks for Cities. *ISPRS International Journal of Geo-Information*, 7(12), 468. <https://doi.org/10.3390/ijgi7120468>
- Hachmann, S., Jokar Arsanjani, J., & Vaz, E. (2018). Spatial data for slum upgrading: Volunteered Geographic Information and the role of citizen science. *Habitat International*, 72, 18–26. <https://doi.org/10.1016/j.habitatint.2017.04.011>
- Haklay, M. (Muki), Basiouka, S., Antoniou, V., & Ather, A. (2010). How Many Volunteers Does it Take to Map an Area Well? The Validity of Linus’ Law to Volunteered Geographic Information. *The Cartographic Journal*, 47(4), 315–322. <https://doi.org/10.1179/000870410X12911304958827>
- Haklay, Mordechai. (2010). How Good is Volunteered Geographical Information? A Comparative Study of OpenStreetMap and Ordnance Survey Datasets. *Environ Plann B Plann Des*, 37(4), 682–703. <https://doi.org/10.1068/b35097>
- Haklay, Muki. (2013). Citizen Science and Volunteered Geographic Information: Overview and Typology of Participation. İçinde D. Sui, S. Elwood, & M. Goodchild (Ed.), *Crowdsourcing Geographic Knowledge: Volunteered Geographic Information (VGI) in Theory and Practice* (ss. 105–122). [https://doi.org/10.1007/978-94-007-4587-2\\_7](https://doi.org/10.1007/978-94-007-4587-2_7)
- Hecht, R., Kunze, C., & Hahmann, S. (2013). Measuring Completeness of Building Footprints in OpenStreetMap over Space and Time. *ISPRS International Journal of Geo-Information*, 2(4), 1066–1091. <https://doi.org/10.3390/ijgi2041066>
- Küçük. (2019). <https://github.com/kadirkuçuk/Proje>. Son erişim tarihi: 10 Aralık 2019
- OSM. (2019). OpenStreetMap Copyright and Licence. Gelişim tarihi gönderen <https://www.openstreetmap.org/copyright/en>



- QGIS. (2019). OpenStreetMap Verisinde Arama Yapma ve Veriyi İndirme—QGIS Tutorials and Tips. Geliş tarihi gönderen [http://www.qgistutorials.com/tr/docs/downloading\\_osm\\_data.html](http://www.qgistutorials.com/tr/docs/downloading_osm_data.html)
- Qi, Y., Zhang, C., Zhi, Z., Guo, K., & Guo, D. (2018). A VGI-based Foodborn Disease Report and Forecast System. Proceedings of the 4th ACM SIGSPATIAL International Workshop on Safety and Resilience, 18:1–18:7. <https://doi.org/10.1145/3284103.3284124>
- Schlesinger, M. I., Vodolazskii, Y. V., & Yakovenko, V. M. (2014). Recognizing the Similarity of Polygons in a Strengthened Hausdorff Metric. *Cybern Syst Anal*, 50(3), 476–486. <https://doi.org/10.1007/s10559-014-9636-2>
- Senaratne, H., Mobasher, A., Ali, A. L., Capineri, C., & Haklay, M. (Muki). (2017). A review of volunteered geographic information quality assessment methods. *International Journal of Geographical Information Science*, 31(1), 139-167. <https://doi.org/10.1080/13658816.2016.1189556>
- Sevinç, H. K., & Karaş, İ. R. (2018). Gönüllü Coğrafi Bilgi, Sivil Bilim ve Katılımcı Coğrafi Bilgi Sistemleri Arasındaki Benzerlikler ve Farklılıklar. <http://dx.doi.org/10.15659/uzalcbs2018.6661>
- Taşkanat, T., Karaağaç, A., Beşdok, E., & Bostancı, B. (2018). Kentsel Sorunların Yönetimi için Bir Gönüllü Coğrafi Bilgi Mobil Uygulaması Geliştirilmesi. *Geomatik*, 3(1), 84–91. <https://doi.org/10.29128/geomatik.371144>
- Touya, G., Antoniou, V., Olteanu-Raimond, A.-M., & Van Damme, M.-D. (2017). Assessing Crowdsourced POI Quality: Combining Methods Based on Reference Data, History, and Spatial Relations. *ISPRS International Journal of Geo-Information*, 6(3), 80. <https://doi.org/10.3390/ijgi6030080>
- USGS. (2019). Volunteered Geographic Information (VGI). Geliş tarihi gönderen <https://www.usgs.gov/core-science-systems/ngp/cegis/vgi>
- Veregin, H. (1999). Data Quality Parameters. *İçinde Geographical Information Systems, Principles and Applications , Principles and Technical Issues* (2. bs, C. 1, ss. 177-189). New York: John Wiley & Sons.
- Yılmaz, A., & Canıberk, M. (2018). Real Time Vector Database Updating System: A Case Study for Turkish Topographic Vector Database (TOPOVT). *International Journal of Engineering and Geosciences*, 3(2), 73–79. <https://doi.org/10.26833/ijeg.383054>