T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

ADRESE DAYALI COĞRAFİ KODLAMA İŞLEMİNDE KALİTE ARAŞTIRMASI

BATUHAN KILIÇ

YÜKSEK LİSANS TEZİ
HARİTA MÜHENDİSLİĞİ ANABİLİM DALI
UZAKTAN ALGILAMA VE CBS PROGRAMI

DANIŞMAN
DOÇ. DR. FATİH GÜLGEN

İSTANBUL, 2017

T.C.

YILDIZ TEKNİK ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

ADRESE DAYALI COĞRAFİ KODLAMA İŞLEMİNDE KALİTE ARAŞTIRMASI

Batuhan KILIÇ tarafından hazırlanan tez çalışması 07.06.2017 tarihinde aşağıdaki jüri tarafından Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü Harita Mühendisliği Anabilim Dalı'nda **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Tez Danışmanı

Doç. Dr. Fatih GÜLGEN

Yıldız Teknik Üniversitesi

Jüri Üyeleri

Doç. Dr. Fatih GÜLGEN

Yıldız Teknik Üniversitesi _____

Doç. Dr. Melih BAŞARANER

Yıldız Teknik Üniversitesi _____

Prof. Dr. Necla ULUĞTEKİN

İstanbul Teknik Üniversitesi _____

ÖNSÖZ

Tez çalışmam boyunca bana karşı hep anlayışlı olan ve manevi desteğini eksik etmeyen biricik eşim Seda KILIÇ'a;

Hem eğitim hemde çalışma hayatım boyunca desteklerini bir an olsun esirgemeyen ve bana her daim göstermiş oldukları anlayış için sevgili aileme;

Tez çalışması ve Yüksek Lisans eğitimim süresince bilgi ve tecrübelerinden faydalandığım, tezin her aşamasında sorunlarımı sabırla dinleyen ve beni yönlendiren, maddi ve manevi her konuda yardımlarını esirgemeyen tez danışmanım ve saygıdeğer hocam Doç.Dr. Fatih GÜLGEN'e;

Derslerinde edindiğim bilgi ve tecrübelerinden dolayı değerli hocalarım Prof.Dr. Türkay GÖKGÖZ, Doç.Dr. Melih BAŞARANER ve Yrd.Doç.Dr. Alper ŞEN ve Yrd.Doç.Dr. Sinan ÇETİNKAYA'ya;

Tez kapsamındaki uygulamada; sorularımı içtenlikle yanıtlayan ve bilgi ve tecrübelerini benimle paylaşan Doç.Dr. Bahattin ERDOĞAN'a;

Tez çalışmam süresince bana vermiş oldukları destek ve anlayışlarından ötürü mesai arkadaşlarım Kadir ŞAHBAZ, Müslüm HACAR, Abdulkadir MEMDUHOĞLU, Ali Hasan DOĞAN, M. Fahri KARABULUT, Yalçın YILMAZ, Barış SÜLEYMANOĞLU, Hüseyin DUMAN ve Mert GÜRTÜRK'e;

Gerek lisans eğitimim gerekse sonrasındaki kariyer planlamalarımda bana hep yardımcı olan, maddi ve manevi desteklerini hiçbir konuda esirgemeyen Prof.Dr. Abdullah ÖZEN'e ve Muhammed Can ÇAĞLAR abime;

Teşekkürü bir borç bilirim.

Haziran, 2017

Batuhan KILIÇ

İÇİNDEKİLER

Sayfa

SİMGE LİSTESİ vi

KISALTMA LİSTESİ vii

ŞEKİL LİSTESİ viii

ÇİZELGE LİSTESİ ix

ÖZET x

ABSTRACT xii

BÖLÜM 1

GİRİŞ 1

1.1 Literatür Özeti 2
1.2 Tezin Amacı 7
1.3 Hipotez 7

BÖLÜM 2

ADRESE DAYALI COĞRAFİ KODLAMA 8

2.1 Dünyada ve Ülkemizde Adres Yapıları 8
2.2 Coğrafi Kodlama 11
2.3 Coğrafi Kodlama Ölçütleri 15
2.3.1 Coğrafi Kodlama Kalitesi 16
2.3.2 Sistemsel Entegrasyon 17
2.3.3 Arayüzün Yapısal Durumu 18
2.3.4 Maliyet Hesabı 18
2.3.5 Kapasite Hesabı 18

BÖLÜM 3

ÇEVRİMİÇİ COĞRAFİ KODLAMA HİZMETLERİNDE KALİTE ARAŞTIRMASI 20

3.1 Referans Verilerin Oluşturulması 21
3.2 Adres Verilerinden Coğrafi Kodlama 25
3.2.1 Konumsal Uyum 26
3.2.2 Adres Uyumu 34
3.2.3 Tersine Coğrafi Kodlamanın Uyumu 38

BÖLÜM 4

SONUÇ VE ÖNERİLER 42

KAYNAKLAR 44

ÖZGEÇMİŞ 49

SİMGE LİSTESİ

E_b	Bing maps sağa değerler
E_g	Google maps sağa değerler
E_i	Servislerden elde edilen sağa değerler
E_{ref}	Referans sağa değerler
n	Ölçü sayısı
N_b	Bing maps yukarı değerler
N_g	Google maps yukarı değerler
N_i	Servislerden elde edilen yukarı değerler
S_b	Bing maps sapma değer
S_g	Google maps sapma değer
S_{ref}	Referans sapma değer
x_i	Her bir ölçü için fark değerleri
ΔE	Sağa değer farkları
ΔN	Yukarı değer farkları
ΔS	Konumsal farklar

KISALTMA LİSTESİ

ABD	Amerika Birleşik Devletleri
AKS	Adres Kayıt Sistemi
API	Uygulama Programlama Arayüzü
BOTAŞ	Boru Hatları İle Petrol Taşıma Anonim Şirketi
CARIS	Computer Aided Resource Information System
CASS	Coding Accuracy Support System
CBS	Coğrafi Bilgi Sistemi
CBSGM	Coğrafi Bilgi Sistemleri Genel Müdürlüğü
DIME	Dual Independent Map Encoding
ERDAS	Earth Resources Data Analysis System
ESRI	Environmental Systems Research Institute
GPS	Global Positioning System
GRASS	Geographic Resources Analysis Support System
HTML	Hypertext Markup Language
INSPIRE	Infrastructure for Spatial Information in Europe
ISO	International Organization for Standardization
ITRF	International Terrestrial Reference Frame
KBS	Kent Bilgi Sistemi
MAKS	Mekânsal Adres Kayıt Sistemi
MERNİS	Merkezi Nüfus İdare Sistemi
MOSS	Map Overlay and Statistical System
NAD	North American Datum
OASIS	Organization for the Advancement of Structured Information Standards
POI	Point of Interest
SQL	Structured Query Language
TIGER	Topologically Integrated Geographic Encoding and Referencing
TUCBS	Türkiye Ulusal Coğrafi Bilgi Sistemleri
UAVT	Ulusal Adres Veri Tabanı
UPU	Universal Postal Union
WGS	World Geodetic System
XML	Extensible Markup Language

ŞEKİL LİSTESİ

Sayfa

Şekil 2.1 *INSPIRE* Ek-1 temaları [49] 9
Şekil 2.2 Klasik coğrafi kodlama işlem akışı [22] 12
Şekil 2.3 Adres standardizasyonu algoritması 13
Şekil 2.4 Adres lineer enterpolasyon algoritması örneği 14
Şekil 2.5 Çevrimiçi coğrafi kodlama işlem akışı [22] 14
Şekil 2.6 Eşleşme oranı ve Eşleşme puanı bilgileri 17
Şekil 3.1 Fatih bölgesi ve Tarihi yarımada 20
Şekil 3.2 Miami Beach bölgesi 21
Şekil 3.3 *Web Content Extractor V.8.4* genel görünüm 23
Şekil 3.4 Fatih'de bir referans otelin (a) harita ve (b) perspektif görüntüsü 24
Şekil 3.5 Miami Beach'de bir referans otelin (a) harita ve (b) perspektif görüntüsü 25
Şekil 3.6 *Python* ile çevrimiçi coğrafi kodlama 25
Şekil 3.7 Çokgen cephesine olan en yakın mesafeler 27
Şekil 3.8 Fatih bölgesi için her bir ölçünün ΔE, ΔN ve ΔS sapma miktarları 30
Şekil 3.9 Miami Beach için her bir ölçünün ΔE, ΔN ve ΔS sapma miktarları 31
Şekil 3.10 Levenshtein mesafesi için örnek matrisler [67] 35
Şekil 3.11 Python ile Levenshtein mesafesi'nin hesaplanması 35
Şekil 3.12 Coğrafi kodlamayla elde edilen adres verilerinin uyum yüzdeleri 36
Şekil 3.13 Tersine coğrafi kodlamayla elde edilen adres verilerinin uyum yüzdeleri 39
Şekil 3.14 Tersine coğrafi kodlama kalitesinin etkilenmesi 41

ÇİZELGE LİSTESİ

Sayfa

Çizelge 2.1 Çevrimiçi coğrafi kodlama servislerinin avantaj ve dezavantajları..... 15
Çizelge 2.2 Servisler arasındaki kapasite farklılıkları..... 19
Çizelge 3.1 Tespit edilen konaklama tesislerine ait sayılar..... 22
Çizelge 3.2 Otel çokgenlerinin içinde ve dışında yer alan ilgi noktaları..... 26
Çizelge 3.3 Fatih bölgesi..... 27
Çizelge 3.4 Miami Beach bölgesi..... 28
Çizelge 3.5 Fatih bölgesi için istatistiksel değerler..... 32
Çizelge 3.6 Miami Beach bölgesi için istatistiksel değerler..... 32
Çizelge 3.7 Standart adres formatlarının birbirleriyle olan ortalama uyum değerleri ... 37
Çizelge 3.8 Tersine coğrafi kodlama sonucu elde edilen adres formatlarının uyum değerleri..... 40

ÖZET

ADRESE DAYALI COĞRAFİ KODLAMA İŞLEMİNDE KALİTE ARAŞTIRMASI

Batuhan KILIÇ

Harita Mühendisliği Anabilim Dalı

Yüksek Lisans Tezi

Tez Danışmanı: Doç. Dr. Fatih GÜLGEN

Coğrafi Bilgi Sistemleri ve bilgi teknolojilerindeki gelişmelerle birlikte adresler, konum bulma ve posta hizmetlerinin yanında, araç navigasyonu ve ulaştırma hizmetlerine, coğrafi ve demografik analizlerin gerçekleştirilmesine altlık oluşturmaktadır. Adres bilgilerinden coğrafi kodlama teknolojisi, adresi oluşturan metinsel tanımlamalar ile mekânsal veri arasındaki köprüyü kurmada önemli rol oynamaktadır.

Coğrafi kodlama işlemi, adres bilgisini enlem ve boylam değerleriyle ifade edilen coğrafi koordinatlara dönüştürmekte kullanılan bir yöntemdir. Coğrafi kodlama çalışmalarında mekânsal analizlerin kalitesinin artırılması, adreslerin güncelliğine, adres bileşenlerinin tutarlılığına ve en önemlisi adreslerin standart bir formatta kullanılmasına bağlıdır. Bu çalışmada, günümüzde sık olarak kullanılan coğrafi kodlama sistemlerinden *Google* ve *Bing Maps* çevrimiçi coğrafi kodlama hizmetlerinin kaliteleri karşılaştırılmaktadır. Adres verileri üzerinden coğrafi kodlama işlemi gerçekleştirilerek elde edilen ilgi noktalarının konumsal sapmaları irdelenmekte ve adresi oluşturan metinleri *Levenshtein mesafe algoritması* kullanılarak değerlendirilmektedir.

Test işlemleri, İstanbul ili Fatih ilçesi ve Amerika Birleşik Devletleri Florida eyaletine bağlı Miami Beach bölgesi sınırları içinde kalan konaklama tesislerinin adres bilgileri kullanılarak gerçekleştirilmiştir. Tesis adresleri *Web Content Extractor V.8.4* yazılımı kullanılarak web sayfaları üzerinden toplanmıştır. Çevrimiçi coğrafi kodlama işlemi, *Python* programlama dilinde geliştirilmiş olan bir coğrafi kodlama kütüphanesi

kullanılarak *Google* ve *Bing Maps* servisleri üzerinde gerçekleştirilmiştir. Adreslerden coğrafi kodlamayla elde edilen ilgi noktalarının koordinatları istatiksel olarak incelenmiş ve referans değerleriyle olan farklılıkları ortaya konulmuştur. Servislerden elde edilen standart adres formatlarının birbirleriyle ve referans verisi ile uyumu Levenshtein mesafesi kullanılarak karşılaştırılmıştır. Ayrıca, servisler üzerinden yapılan tersine coğrafi kodlama işlemlerinin kalitesi irdelenmiştir. Sonuçlar, çevrimiçi coğrafi kodlama servislerden elde edilen noktaların konumsal uyumunun adres bilgisi uyumunun bölgelere göre değişken olduğunu göstermektedir.

Anahtar Kelimeler: CBS, adres, coğrafi kodlama, tersine coğrafi kodlama, ilgi noktası, Levenshtein mesafesi

ABSTRACT

QUALITY INVESTIGATION IN ADDRESS-BASED GEOCODING PROCESS

Batuhan KILIÇ

Department of Geomatic Engineering

MSc. Thesis

Adviser: Assoc. Prof. Fatih GÜLGEN

Nowadays, along with the developments of Geographical Information Systems (GIS) and information technologies, postal addresses are used in many areas such as postal services, vehicle navigation, transportation routing, geographical and demographic analysis etc. Geocoding technology from address information plays an important role in bridging between non-spatial and spatial data.

Geocoding is a method used to convert address information into geographical coordinates expressed in terms of latitude and longitude. The increase in the quality of spatial analysis in geocoding studies depends on the fact that address structures are kept up-to-date in databases, address components are consistent and most importantly, addresses are used in a standard format. In this study, the qualities of the most commonly used online geocoding services, *Google and Bing Maps*, were compared. The positional deviations of the points of interest obtained data are examined by performing the geocoding process on the address and verbal address information is evaluated using the *Levenshtein distance algorithm*.

Two case studies are implemented in Fatih District in İstanbul city of Turkey and Miami Beach Region in Florida States of USA using the addresses of accommodation facilities. The address data of accommodation facilities in those areas are collected via web pages using *Web Content Extractor V.8.4* software. Online geocoding was done on *Google* and *Bing Maps* services using a geocoding library developed in the *Python programming language*. The coordinates of the points of interest obtained from the

services were statistically analyzed and their differences from the reference values were revealed. Then the compatibility of the standard address formats obtained from the services with each other and with the reference data was compared using the *Levenshtein distance*. Finally, the quality of reverse geocoding processes through services was examined. The results show that the positional and address consistencies obtained from online geocoding services vary according to different regions.

Keywords: GIS, address, geocoding, reverse geocoding, points of interest, Levenshtein distance

BÖLÜM 1

GİRİŞ

Harita, mekân ile ilgili bilgilerin bir düzlem üzerinde, ölçeğe bağlı olarak genelleştirilmiş ve içerik olarak sınırlandırılmış bir modeli olmasının yanısıra görsel bir iletişim aracıdır [1]. Bilgisayar teknolojilerindeki gelişmeler coğrafi bilgi sistemlerinin oluşturulmasını zorunlu hale getirmiş ve haritalar basılı ortamdan sayısal ortama taşınmıştır. Günümüzde mobil cihazların ve sosyal ağların gelişmesi coğrafi mekâna ilişkin bilgilere hızlı ve ayrıntılı şekilde ulaşabilmeyi olanaklı kılmıştır. Coğrafi verinin temelinde nokta detaylar önemli bir yer kaplamaktadır. İnsanları uğrak yerlerine sevk eden ve sayısal haritalar üzerinde konum belirten nokta detaylar *ilgi noktaları* (*Points of Interest* – POIs) olarak adlandırılmaktadır.

Çeşitli sistemler ve cihazlar kullanılarak araziden toplanan ya da gönüllü katılımcılar tarafından sosyal ağlar üzerinden elde edilen ilgi noktaları, harita servisleri aracılığıyla kullanıcılara sunulmaktadır. Bir harita servisinde ilgi noktalarının konumları coğrafi koordinat sistemine bağlı enlem ve boylam değerleriyle tanımlıdır. Buna karşın her ilgi noktası gerçek dünyada bir adres bilgisine sahiptir ve coğrafi koordinatlar ile adres arasında bir köprünün kurulmuş olması gerekir. Bu köprü *coğrafi kodlama* (‘*geocoding*’) ile sağlanmaktadır. Coğrafi kodlamanın sağlıklı olarak gerçekleştirilmesi adres bilgilerinin tutarlılığına bağlıdır.

Adres, belirli bir konuma ulaşmanın yanında çeşitli kamu kurumlarındaki bilgilerin birbirleriyle ilişkilendirilmesinde ve geleceğe yönelik stratejik planların oluşturulmasında önemli bir rol oynar. Yerleşim alanlarında karmaşık şekilde inşa edilmiş yol ağının ve bina yapılarının düzenli bir adres sistemine dayandırılması oldukça zordur. Ülkemizde kamu kurum ve kuruluşları tarafından farklı adres formatlarının

kullanılması coğrafi kodlamanın sağlıklı yapılabilmesinin önündeki en büyük engeldir. Bunun yanında; (1) yol başlangıç ve bitişlerinin tespit edilmemiş olması, (2) numarataj çalışmalarındaki yetersizlikler, (3) adres yapılarındaki bileşen tutarsızlıkları, (4) kısaltma tanımlamalarındaki farklılıklar, adres verileri üzerinden yapılan coğrafi kodlamayı zorlaştırmaktadır. Özetle standart bir adres formatının olmayışı coğrafi kodlama işlemlerinde olumsuzluklar yaratmakta ve doğru sonuçlara ulaşmakta problemler çıkarmaktadır.

Bu kapsamda, tez çalışmasının birinci bölümünde coğrafi kodlama ile ilgili literatür aktarılmış, tezin kapsamı, amacı ve hipotezi verilmiştir. İkinci bölümde coğrafi kodlama kavramı irdelenmiş, coğrafi kodlama yöntemleri ve coğrafi kodlama ölçütleri aktarılmıştır. Üçüncü bölümde tez çalışmasının metodolojisi ortaya koyulmuş ve uygulama sonuçları değerlendirilmiştir. Dördüncü bölüm ise sonuç ve önerilerin aktarıldığı kısımdır.

1.1 Literatür Özeti

Coğrafi kodlama, Zandbergen [2] tarafından *"bir referans veritabanındaki adres tanımlamalarının coğrafi koordinatlar ile eşleştirilmesi*" şeklinde tanımlanmıştır. Coğrafi kodlama çalışmalarının başlangıcı 1 Temmuz 1963 yılında Amerika Birleşik Devletleri (ABD) Posta Hizmet Bakanlığı (*US Post Office Department*) tarafından uygulamaya koyulan, beş haneli posta kodlarının kullanılmasına dayanmaktadır [3]. ABD Nüfus Sayım İdaresi (US *Census Bureau*) çalışanı James Corbett ve Donald Cooke 1967 yılında, coğrafi kodlama için ilk modern vektör harita modelini ('*Dual Independent Map Encoding*' – DIME) oluşturmuştur. Bu model, bir cadde ağı boyunca adres dizilerine kod atayan '*percent along*' isimli coğrafi bir kodlama algoritmasıyla gerçekleştirilmiştir. Günümüzde bu algoritma, '*Google Maps'* ve '*Mapquest'* gibi platformlar tarafından halen kullanılmaktadır. *ABD Nüfus Sayım İdaresi*'nin kullanımı için planlanan *DIME*, yol bağlantılarını oluşturan kavşakların düğüm noktaları şeklinde sayısallaştırılmasına ve yollar arasındaki mekânsal ilişkilerin ortaya çıkarılmasına dayanmaktadır. ABD'nin Connecticut eyaletine bağlı olan New Haven, coğrafi olarak kodlanmış yol ağı veritabanına sahip, dünya üzerindeki ilk şehirdir [4],[5].

Yetmişli yılların sonlarında, '*Geographic Resources Analysis Support System'* (GRASS GIS) ve *'Map Overlay and Statistical System'* (MOSS) yazılımları üzerinde kamuya açık iki coğrafi kodlama platformu geliştirilmiştir. Seksenli yılların başında '*Intergraph'*, *'Environmental Systems Research Institute'* (ESRI), *'Computer Aided Resource Information System'* (CARIS), *'Earth Resources Data Analysis System'* (ERDAS) ve *'MapInfo Corporation'* (günümüzde Pitney Bowes) olarak adlandırılan ticari firmalar, coğrafi kodlama araçlarını geliştirmeye başlamıştır [3].

Yirminci yüzyılın sonlarında, coğrafi verinin internet üzerinden dağıtımının yaygınlaşması ve CBS yazılımlarının buna bağlı olarak gelişmesiyle, coğrafi kodlama sistemleri kullanıcı odaklı hale gelmiştir. ABD Nüfus Sayım İdaresi, katılımcıların adres, isim, cinsiyet, ırk, yaş, medeni durum gibi temel bilgilerini içeren geniş kapsamlı bir mekânsal veritabanı oluşturmak için, enterpole edilmiş coğrafi kodlama yöntemine dayalı '*mail-out/mail-back'* isimli bir teknik geliştirmiştir [6]. Bu teknikle elde edilen veritabanı, yollar, binalar, nehirler, göller gibi arazi detaylarını tanımlayan *'Topologically Integrated Geographic Encoding and Referencing'* (TIGER) formatını kullanmaktadır. Adres dizilerini kapsayan *TIGER*, bugünlerde neredeyse tüm coğrafi kodlama yazılım platformlarında kullanılmaktadır [3].

İki binli yılların başlarında '*Coding Accuracy Support System'* (CASS) olarak bilinen adres standardizasyonu ortaya çıkmıştır. Yıllık olarak güncellenen CASS sertifikasyonu, teslimat noktası kodları, posta kodları ve ABD'de kullanılan *ZIP+4* kodlarına odaklıdır [3]. İzleyen yıllarda coğrafi kodlama platformlarında tersine coğrafi kodlama işlemi kullanılmaya başlamıştır [7],[8],[9],[10]. Bu işlem enlem ve boylam değerleriyle ifade edilen coğrafi koordinatlardan metinsel, okunabilir adres formatlarına dönüşüm sağlamaktadır. 2008 ve 2009 yılları arasında '*Google Maps'*, '*MapQuest'* ve '*Bing Maps'* olarak adlandırılan coğrafi kodlamaya olanak sağlayan kullanıcı odaklı platformlar ortaya çıkmıştır [11],[12],[13]. Bu platformlar özellikle akıllı telefonlar ve mobil endüstrinin eş zamanlı gelişimiyle coğrafi kodlama ve ilgi noktalarına erişimi daha kolay hale getirmiştir. Son yıllarda coğrafi kodlama ve tersine coğrafi kodlamaya destek veren hizmet sağlayıcılar, bulut odaklı coğrafi kodlama yapabilen programlama arayüzü (*API*) ve kurum içi coğrafi kodlama platformlarındaki gelişmeler sayesinde daha iyi,

daha doğru ve daha hızlı şekilde adres eşleme ve coğrafi kodlamaya imkân tanımaktadır [3], [14].

Literatürde coğrafi kodlama işleminin kullanıldığı akademik çalışmalar, genellikle adres eşleme, kamu sağlığı ve epidemiyoloji üzerine yoğunlaşmaktadır [15],[16], [17],[18],[19],[20],[21],[22],[23],[24],[25],[26],[27],[28]. McElroy vd. [15] sağlık üzerine yapılan araştırmalarda coğrafi kodlama işlemini gerçekleştirebilmek amacıyla 1988 ve 1994 yılları arasında ABD'nin Wisconsin eyaletinde göğüs kanserine yakalanan kadınlar ile yapılan telefon konuşmalarından alınmış posta adreslerini kurumsal bir yazılım kullanarak kodlamıştır. Farklı tarihlerde yapılan kodlamaların eşleştirilmesi adres bilgisinin kalitesinin artırılmasına ve bu adreslere ait enlem-boylam koordinatlarını elde etmeye dayandırılmıştır. Cayo ve Talbot [16] ikamet adreslerini kullanarak '*MapMarker*' yazılımı ile coğrafi kodlama gerçekleştirmiş ve elde edilen noktaların konumsal doğruluğunu ortofoto harita verileriyle karşılaştırarak değerlendirmiştir. Kırsal alanlardaki konumsal hata oranının diğer bölgelere göre daha fazla olduğunu tespit etmiştir. Karimi vd. [17] coğrafi kodlama teknikleri ile ortaya çıkan eşleşme durumu ve konumsal doğruluk belirsizliklerini incelemiştir. Yerleşim bölgelerine ait adres bilgilerini üç farklı coğrafi kodlama algoritması ile test edip elde ettikleri sonuçları GPS ile toplanmış veriler ile karşılaştırmıştır. Yang vd. [18] eşleme oranı ve konumsal doğruluklar ile ilgili çalışmasında üç ayrı ticari coğrafi kodlama sisteminin ('*ArcView*', '*Automatch*' ve '*ZIP/Geolytics*') artılarını ve eksilerini ortaya koyarak kodlama sonucu meydana gelen hata kaynaklarını belirlemiştir. Bir veritabanından rastgele alınan adres verileri üzerinden yapılan çalışmada üç yazılım içerisinden en iyi sonuçların *ArcView* tarafından sağlandığı tespit edilmiştir. Ward vd. [19] ABD'nin eyaletlerinden biri olan Iowa'da nüfusa odaklı Hodgkin dışı lenfoma hastalığına yakalanan insanların adres bilgileri ile yapılan coğrafi kodlama çalışmaları sonucunda elde edilen konumsal hataları incelemiştir. *ArcView* ve ek bir ticari yazılım kullanılarak elde edilen konumlar, ikamet edilen evlere ait GPS ölçüleriyle karşılaştırılmıştır. Karşılaştırma sonucu konumsal hataların kentsel alanlardan ziyade kırsal alanlarda daha fazla olduğunu saptamıştır. Rushton vd. [20] kanser araştırmalarında coğrafi kodlama işleminin nasıl gerçekleştirilmesi gerektiği üzerine odaklanmıştır. Coğrafi olarak kodlanan kanser kayıtlarının doğruluğunu belirlemek için kanser hastalığı çalışmalarıyla bağlantılı olarak

(1) ekolojik analizler, (2) enterpolasyon ve (3) parsel eşleme olarak adlandırılan üç farklı yöntemin kullanıldığından söz etmiştir. Goldberg vd. [21] coğrafi kodlama ile ilgili gerçekleştirilmiş akademik çalışmaları incelemiş ve coğrafi kodlama kavramı, bileşenleri ve işlem sonrasında ortaya çıkan hatalar hakkında genel bilgiler vermiştir. Roongpiboonsopit ve Karimi [22], '*Geocoder.us*', '*Google*', '*MapPoint*', '*MapQuest*', ve '*Yahoo*' isimli beş farklı çevrimiçi coğrafi kodlama servisinin eşleme kalitesini incelemiştir. Çalışmada, coğrafi kodlama sonuçlarının kalitesine odaklı olarak çevrimiçi coğrafi kodlama servisleri arasındaki farklılıkları ortaya çıkarmayı ve uygulamanın ihtiyaçlarına göre uygun servisin seçimi için genel bir bakış açısı oluşturulması hedeflenmiştir. Her bir servisin kalitesi (1) eşleme oranı, (2) konumsal doğruluk ve (3) eşleşme sonucu servisler arasındaki benzerlik olarak tanımlanmış üç farklı ölçüte bağlı olarak değerlendirilmiştir. Li vd. [23] Çin'deki adres yapısının düzensiz olması nedeniyle coğrafi kodlama çalışmaları için uygun olmadığını belirterek makine öğrenmeye dayalı birliktelik kural çıkarımı (*association rule mining*) yöntemini adres modeli oluşturmak amacıyla kullanmıştır. Oluşturulan adres modeli, coğrafi kodlama çalışmaları için açık ve analiz edilebilir bir yöntem olarak değerlendirilmiştir. Zandbergen [24] yol verilerinin coğrafi kodlama çalışmalarına olan etkisini belirlemek için ABD'nin Florida eyaletine bağlı Volusia bölgesindeki kentsel ve kırsal alanlarda yaşayan tek aileye ait konutların, çoklu ailelere ait konutların, ticari, kurumsal ve endüstriyel alanların adres bilgilerininin coğrafi kodlama sonuçlarının kalitesini üç farklı veri seti üzerinden karşılaştırıp incelemiştir. Goldberg vd. [25] coğrafi kodlama sistemlerinde amaçlar doğrultusunda karşılaşılabilen ihtiyaçları belirlemede büyük kurum ve kuruluşlara yardım etmek amaçlı bir değerlendirme çerçevesi geliştirmiş ve anlaşmalar gereği isimleri belirtilmeyen beş farklı coğrafi kodlama sistemini kullanarak test etmiştir. Test için coğrafi kodlama çalışmalarının kalite bazında değerlendirilmesi, kullanıcıların sistemlere erişebilme olanakları, sistemin sahip olduğu yazılım bilgileri, sisteme ait maliyet giderleri ölçütler olarak kullanılmıştır. Satanta vd. [26] Endonezya'daki adres verileri üzerinden yapılacak coğrafi kodlama çalışmalarındaki sıkıntıları ortaya koymuştur. Kentsel ve kırsal bir alanda *ESRI* yazılımı ile yaptığı çalışmada en uygun sonuçların '*General Single Field*' ve '*US Address Single House*' isimli adres konumlama araçları ile yapılabildiğini saptamıştır. Tian vd. [27] adres yapılarındaki sorunlara

değinerek, Çin'deki adres veri yapısı ile batı ülkelerinin adres yapıları arasındaki farklılıkları ortaya koymuştur. Çin'e özel adres modeli ve adres standardizasyonunu ele alan optimize edilmiş adres eşleme yöntemine odaklı bir coğrafi kodlama servisi ortaya koymuştur. Chow vd. [28] farklı coğrafi kodlama servislerinin adresler üzerindeki eşleme oranı ve konumsal doğrulukları üzerine bir çalışma gerçekleştirmiştir. ABD'nin Teksas eyaletindeki adresler internet üzerinden toplanarak *ESRI* ve *CoreLogic PxPoint* masaüstü yazılımları ile *Google*, *Yahoo*, *Bing Maps*, *Geocoder.us*, *Texas A&M University Geocoder* ve *OpenStreetMap* çevrimiçi yazılımlarından oluşan toplam sekiz farklı coğrafi kodlama servisi tarafından test edilmiş ve GPS ölçüleri ile karşılaştırılmıştır. Karşılaştırma sonucu çevrimiçi yazılımlar içerisinde en iyi sonuçların Google tarafından sağlandığı tespit edilmiştir. Ayrıca coğrafi kodlama, suç analizi [29],[30],[31],[32],[33], trafik kazası tespiti [34],[35] ve ulaştırma problemlerinin çözümü [36] gibi farklı çalışma alanlarında temel bir işlem adımı olmuştur.

Ülkemizde coğrafi kodlamayla ilgili çalışmalar sınırlı sayıdadır [37],[38],[39]. Ersoy ve Diri [37] internet kullanıcılarının mevcut metin tabanlı arama motorlarını kullanarak adres bilgisine ulaşma sürecinin kolaylaştırılmasına yönelik bir çalışma gerçekleştirmiştir. Arama motorlarından elde edilen sonuçların coğrafi koordinatlar ile harita üzerinde sunumunu sağlamıştır. Sistemin çözümleyebildiği adres sayısı, arama yapılan konu, mekân, vb. adres bilgilerinin coğrafi koordinatlara dönüştürme oranını belirlemişlerdir. Yıldırım vd. [38] Türkiye'de var olan adres yapısının coğrafi kodlama çalışmalarına uyumunu belirlemek için *BOTAŞ*, *Nüfus ve Vatandaşlık İşleri Genel Müdürlüğü* ve kargo şirketlerinden topladığı verileri istatiksel olarak incelemiştir. Analiz sonucu Türkiye sokak adres sisteminin coğrafi kodlamaya elverişli olmamasının temel sebebinin numarataj çalışmalarındaki eksiklik ve sokak isimlerindeki uyuşmazlıkların olduğunu saptamıştır. *Coğrafi Bilgi Sistemleri Genel Müdürlüğü* (CBSGM) tarafından 2015 yılında "Ortofoto Haritaları Kullanılarak Coğrafi Veri Altyapısı Üretiminin Gerçekleştirilmesi (*GeoKodlama*) Projesi" gerçekleştirilmiştir. Proje, Adıyaman, Batman, Çorum, Gümüşhane, Malatya, Nevşehir, Uşak ve Yozgat il ve ilçe merkezlerini kapsamaktadır. Pilot proje niteliğinde olan proje kapsamında, ortofotolar üzerinden elde edilen bina ve yol detaylarının, web servisleri kullanılarak bina ve yol verilerine

UAVT kodlarının atanması ve adreslerin belirlenerek veri tabanına işlenmesi sağlanmıştır [39].

1.2 Tezin Amacı

Geleneksel olarak, coğrafi kodlama işlemi, Coğrafi Bilgi Sistemleri yazılım paketlerinde mevcut bulunan coğrafi kodlama araçları ile yapılmaktayken, son yıllarda web servisleri ve teknolojinin ilerlemesiyle geliştirilen çevrimiçi coğrafi kodlama sistemleri, CBS uzmanı olmayan kişiler tarafından da kullanılmaya başlanmıştır. Bu tez çalışması kapsamında, Çevrimiçi çalışan servislerde adres verilerinden gerçekleştirilen coğrafi kodlama kalitesinin ölçülmesi için karşılaştırmalı bir model önerisinin geliştirilmesi ve bunun Türkiye verilerine uyarlanması amaçlanmaktadır.

1.3 Hipotez

Çalışma aşağıdaki hipotezlere dayandırılmaktadır:

- Coğrafi kodlamanın kalitesini, adres verilerinin yapısal olarak belirli bir standartta ifade edilmesi, doğrudan etkiler.
- Bina bazında coğrafi kodlamanın kalitesi sistematik ve güncel numarataj sistemine dayanır.
- Ülkelere ve bölgelere göre farklı çevrimiçi coğrafi kodlama servisleri farklı kalitede sonuçlar üretir.
- Tersine coğrafi kodlamada, koordinat sistemi tanımlamalarında yapılacak hatalar, sonuçların kalitesini doğrudan etkiler.

BÖLÜM 2

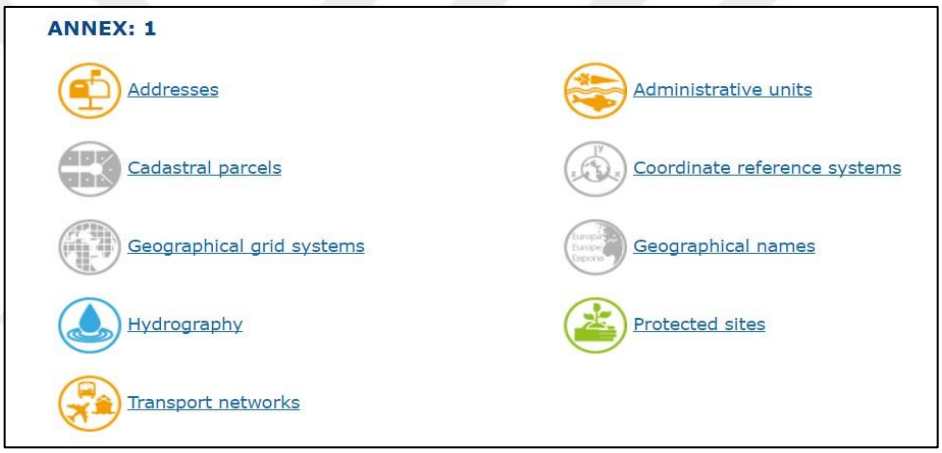
ADRESE DAYALI COĞRAFİ KODLAMA

Adres coğrafi mekânda bir konumu tanımlamak için kullanılan söz dizimidir. Dünya üzerinde yerel yönetimler ve kamu kurum ve kuruluşları ihtiyaçlar doğrultusunda birçok çalışmada adres verilerine ihtiyaç duyar. Kamu hizmetlerinin doğru planlanması ve tüm kamu kurum ve kuruluşları tarafından eşgüdümlü yürütülmesi, devletin tüm vatandaşlarının temel bilgilerine anında ulaşabilmesi, kişilerin ikamet adresleriyle kimlik numaralarını içeren güvenilir nüfus kayıt sistemlerinin kurulması sonucunda güncel olarak elde edilebilir. Adres, konum verisinin tüm varyasyonları arasındaki en temel birim olarak düşünülmektedir. Bir evin, binanın, parselin tek anlamlı bir konumu için bir kod veya bir tanım olarak belirtilmektedir. Kentsel, sağlık, ticari, endüstriyel, dinlenme ya da geleneksel amaçlara bakılmaksızın herhangi bir toprak parçasının temel tanımıdır ve harita üzerindeki coğrafi objelerin konum bilgisine karşılık gelmektedir [40]. Ayrıca, bireyler ile konumları arasında ilişki kurmada kullanılan adres verileri, coğrafi detayların konumlarının belirlenmesinde ve ulusal ya da uluslararası bilgi altyapısı kurulmasında büyük bir öneme sahiptir [2],[23],[24],[38],[41],[42],[43],[44], [45],[46],[47].

2.1 Dünyada ve Ülkemizde Adres Yapıları

Adres verilerinin mekânsal bileşenlerle bütünleştirilmesi ve standart bir yapıya uyarlanması adresi oluşturan bileşenlerin net olarak ifade edilmesiyle sağlanır. Standart adres yapılarının kullanılması yapılacak analiz ve çalışmaların kalitesini yükseltecek, ayrıca ekonomik ve zamansal kayıpları ortadan kaldıracaktır. Uluslararası

kapsamda adres yapıları incelendiğinde, adres bileşenlerinin standartlaşması konusunda Güney Afrika, Avustralya, Yeni Zelanda, Britanya, Danimarka ve Amerika Birleşik Devletleri gibi birçok ülkede çalışmaların devam ettiği görülmektedir. *Uluslararası Posta Birliği* (UPU), *Uluslararası Standartlar Örgütü* (ISO) ve *Yapılandırılmış Bilgi Standartları Gelişim Örgütü* (OASIS) vb. organizasyonlar adres standartlarının oluşturulmasına katkı sağlamaktadır [42]. Avrupa'da mekânsal bilgi altyapılarının oluşturulması ve kurum ve kuruluşlar tarafından kolay erişimin sağlanması sebebiyle 2007 yılında INSPIRE (*Infrastructure for Spatial Information in Europe*) Direktifi yürürlüğe girmiştir [48]. *INSPIRE* Direktifi tarafından önerilen temalar Şekil 2.1'de verilmiştir [49]. Adres teması kapsamında tüm Avrupa adres tanımlamaları incelenmiş ve ortak bir adres veri modeli oluşturulmaya çalışılmıştır.



Şekil 2.1 *INSPIRE* Ek-1 temaları [49]

INSPIRE, adres standartlarını geliştirirken *ISO/TC211* standartlarından faydalanarak ve nesneye yönelik veri modellemesini kullanarak, tüm Avrupa için ortak bir standart oluşturmaya çalışmaktadır. *INSPIRE* direktiflerine göre bir adresin oluşması için gereken bileşenler aşağıda verilmiştir [50]:

- İdari birim ismi (örneğin; kendi yerel yönetim teşkilatı olan belediye isimleri)
- Adresin bulunduğu bölge (örneğin; şehir isimleri)
- Yol adı ve türü (örneğin; cadde isimleri)
- Adres konumlayıcı (örneğin; kapı numarası)
- Posta tanımlayıcı kodları

Ülkemizde kullanılan adres yapıları incelendiğinde, adresi tanımlayan birden fazla niteleyicinin var olduğu görülmektedir. Yerel yönetimler tarafından yapılan numarataj çalışmaları ile ortaya çıkan sokak ismi ve kapı numarası, posta teşkilatının yaptığı çalışmaya bağlı olarak ortaya çıkan posta kodları ve İçişleri Bakanlığı'nın onayı ile verilen mahalle isimleri birer niteleyici adres bileşeni olarak tanımlanmaktadır [51]. Bu bileşenlerin bir arada tutulması kapsamında İçişleri Bakanlığı kontrolünde yürütülen *Adres Kayıt Sistemi* (AKS)'nin verimli bir şekilde çalışabilmesi için Türkiye Cumhuriyeti vatandaşlarının ve Türkiye'de ikamet eden diğer yabancı vatandaşların ikametgâh adreslerinin standartlaştırılması ve kimlik bilgilerinin adres bilgileriyle eşleştirilmesi gerekmektedir. Böylelikle AKS çalışmalarında standartlaştırılmış bir *Ulusal Adres Veri Tabanı* (UAVT) oluşturulmuştur [52].

Adres bilgisine bağlı olarak "25.04.2006 tarih ve 5490 sayılı" Nüfus Hizmetleri Kanunu'nun 69. maddesine dayanarak, adres ve numaralamaya ilişkin bir yönetmelik hazırlanmıştır. Ulusal Adres Veri Tabanındaki adreslerin oluşumu ve güncel olarak takibinden "31.07.2006 tarih ve 26245 sayılı" Resmi Gazete'de yayımlanan "Adres ve Numaralamaya İlişkin Yönetmelik" gereği yetkili idare sorumlu tutulmaktadır. Yetkili idare; belediye teşkilatı olan yerlerde belediye, belediye teşkilatı olmayan yerlerde özel idare ve büyükşehir sınırları içinde büyükşehir belediyesi olarak tanımlanmıştır [53].

Avrupa'da olduğu gibi ülkemizde de adrese odaklı olarak mekânsal bilgi altyapılarının oluşturulması, coğrafi bilginin yönetimi açısından önemlidir. Bu kapsamda teknolojik gelişmelere uygun bir Coğrafi Bilgi Sistemi altyapısının kurulması, kamu kurum ve kuruluşlarının sorumlu oldukları coğrafi bilgileri ortak altyapı üzerinden kullanıcılara sunmaları için bir web portalı oluşturulması ve coğrafi veri değişim standartlarının belirlenmesi amacıyla bir e-devlet projesi olan *Türkiye Ulusal Coğrafi Bilgi Sistemleri* ('TUCBS') oluşturulmuştur [54]. TUCBS kapsamında oluşturulan temel coğrafi veri temaları içerisinde adres ile ilgili olarak *'TUCBS.AD'* Adres Veri Teması, Çevre ve Şehircilik Bakanlığı Coğrafi Bilgi Sistemleri Genel Müdürlüğü'nün yetki ve sorumluluğunda "TUCBS Standartlarının Belirlenmesi Projesi" kapsamında Aralık 2012'de hazırlanmıştır [55]. TUCBS kapsamında yapılan incelemelere göre; Adres Veri Temasının, 10 bakanlığın; 15 genel müdürlüğünün 19 farklı daire başkanlığı tarafından; 27 farklı işte kullanıldığı belirlenmiştir. Adres veri teması, Bina, Parsel, İdari Birim ve

Ulaşım gibi diğer temel coğrafi temalar ile de ilişkilendirilmektedir. Bir binanın birden fazla dış kapısı olabilme durumu ve birden fazla adrese sahip olma durumundan dolayı binalara ait kapı numaraları bina ile ilişkilendirilmelidir. Benzer şekilde bir parselde birden fazla dış kapı olabileceği durumlarda parsel ile kapı numaraları ilişkisi de belirtilmelidir. TUCBS kapsamında adres yapısının oluşturulmasında aşağıdaki durumların göz önünde bulundurulması gerekir [55]:

• *TUCBS.BI* (Bina) temasında, binanın mutlaka en az bir tane iç kapı ve dış kapı numarası bulunmalıdır.

• *TUCBS.TK* (Parsel) temasında tanımlanan parseller bir veya birden fazla adres dış kapı numarasını içerebilir.

• *TUCBS.ID* (İdari Birim) temasında tanımlanan idari birimlerin adı ve özellikleri, adres temasında tanımlanan il, ilçe, mahalle ve köy gibi idari birim isimleriyle ilişkilidir.

• *TUCBS.UL* (Ulaşım) temasında tanımlanan yol bilgisi adres bileşenlerinin oluşturduğu CSBM (Cadde, Sokak, Bulvar, Mahalle) ile ilişkilendirilebilir şeklinde tanımlanmaktadır.

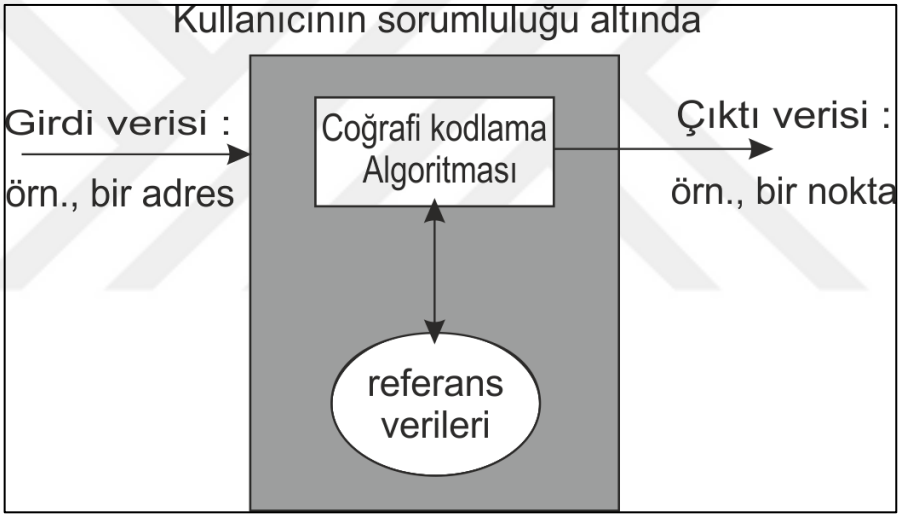
Ayrıca, adres bilgilerine mekânsal boyut kazandırmak amacıyla Nüfus ve Vatandaşlık İşleri Genel Müdürlüğü'nce Adres Kayıt Sistemi'nde metinsel nitelikte tutulan adres bilgilerinin coğrafi koordinatlarla birleştirilmesi ve oluşturulan altyapının diğer sistemlere entegre edilebilmesi amacıyla "*Mekânsal Adres Kayıt Sistemi Projesi* (MAKS)" başlatılmıştır. Bu projenin amaçları doğrultusunda ülke genelindeki tüm adres bileşenlerinin metinsel olarak üzerinde yönetildiği Ulusal Adres Veri Tabanı (*UAVT*)'na mekânsal boyut kazandırılarak daha etkin ve fonksiyonel bir Ulusal Adres Veri Tabanının oluşturulması, Coğrafi Bilgi Sistemi (*CBS*) tabanlı *Kent Bilgi Sistemlerine* (*KBS*) sahip yetkili idarelerin, ikinci bir manuel veri girişine ihtiyaç duymaksızın, sistemlerin entegrasyonu ile kurulacak Mekânsal UAVT'yi güncel tutmalarına olanak sağlayacak servislerin hazırlanması gibi durumlar yer almaktadır[52],[56].

2.2 Coğrafi Kodlama

Mekânsal veriyi elde etmek için kullanılan en basit yol adreslerin yorumlanmasıdır [22]. Metin odaklı adres tanımları haritalar ile doğrudan bağlantılı değildir. Bu ilişki adres verilerini coğrafi koordinatlara dönüştürmeyi hedefleyen coğrafi kodlama işlemiyle

sağlanır [57],[58],[59]. Coğrafi kodlamanın yapılacağı alanının büyüklüğü ve bölgedeki adres sayısı arasında doğru bir orantının olduğu düşünüldüğünde, adres üzerinden coğrafi kodlamanın el ile tek tek yapılması, işlemi zor ve zaman alıcı hale getirir. Günümüzde mekânsal analizlere altyapı oluşturan temel işlemlerden biri olan coğrafi kodlama işlemi, CBS ortamında otomatik olarak; klasik ve (2) çevrimiçi (*'online'*) coğrafi kodlama adları altında, iki şekilde ele alınmaktadır.

Klasik coğrafi kodlama işlemi, coğrafi kodlayıcı (*'geocoder'*) olarak bilinen yazılımlarla gerçekleştirilmektedir. Bir coğrafi kodlayıcı; (1) girdi veriler, (2) coğrafi kodlama algoritmaları, (3) referans veritabanı ve (4) çıktı veriler olmak üzere dört temel bileşene sahiptir (Şekil 2.2) [22].

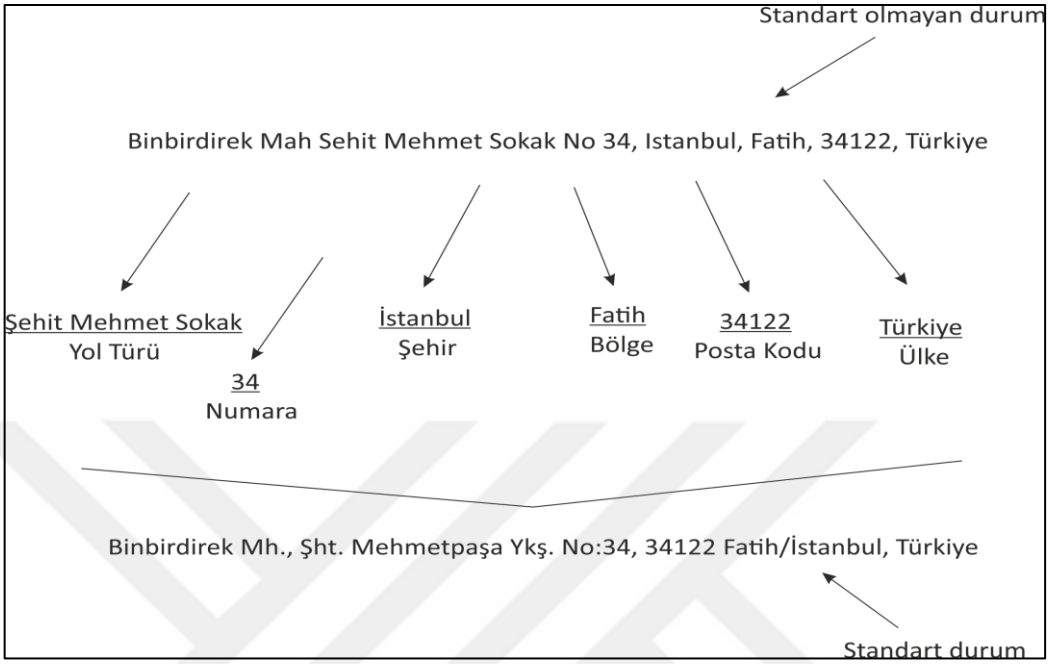


Şekil 2.2 Klasik coğrafi kodlama işlem akışı [22]

Girdi veri, kullanıcıların coğrafi koordinata ulaşabilmesi için mekânsal bilgiyi kapsayan bir coğrafi objenin ya da bir konumun tanımlaması olarak ifade edilebilir. Bu veriler; posta adresleri, otel, okul vb. adlandırılmış mekânlar ve telefon, sayaç numarası vb. diğer ilişkili konumsal tanımlamalardan oluşmaktadır. Çıktı veri ise coğrafi olarak girdi veriye kod bilgilerinin atanmasıyla elde edilen sonuç verisidir.

Klasik coğrafi kodlama işleminde adres standardizasyonunu sağlayan algoritma en önemli işlem araçlarından biridir. Adres standardizasyonu, adres ayrıştırma ('address parsing') ve adres normalizasyonu ('address normalization') olarak adlandırılan iki bölümden oluşmaktadır. Bir coğrafi kodlama sisteminde adres ayrıştırma işleminin rolü

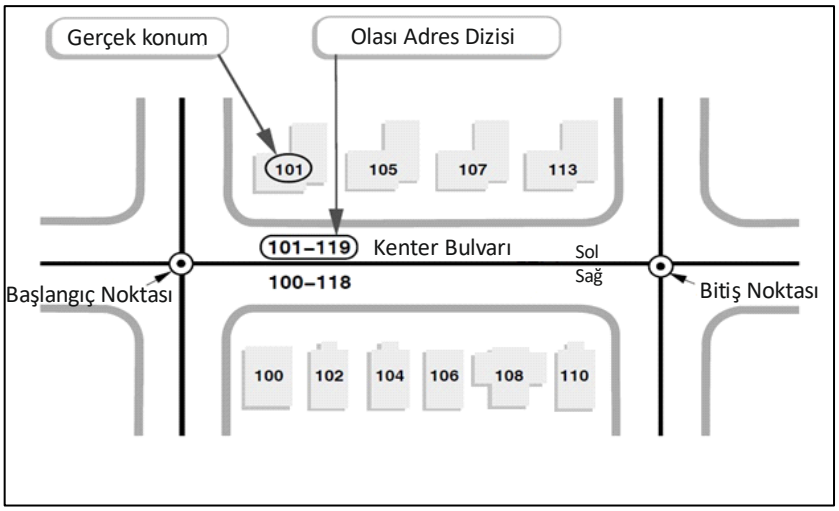
adres tanımlamasını oluşturan bileşenlerin ortaya çıkarılmasıdır [21]. Adres normalizasyonu ise düzensiz veya kısaltılmış olan adres bileşenlerinin standart bir format haline getirilmesini sağlar (Şekil 2.3).



Şekil 2.3 Adres standardizasyonu algoritması

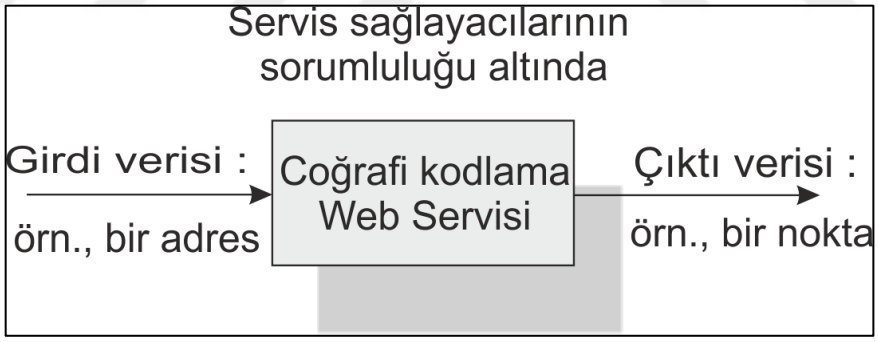
Adres standardizasyonu için kullanılan algoritma referans veritabanı ile uyumlu olarak çalışmaktadır. Referans veritabanı çıktı verilerini elde etmek için algoritmalar tarafından kullanılan coğrafi referans bileşenlerini içermektedir. Referans veritabanını oluşturmak için yol ağı veri modeli, coğrafi birim modeli veya parsel sınırına ait veri modeli (posta kodları, bölgeler, şehirler) ve adreslere ait nokta veri modeli gibi birçok adres modeli bulunmaktadır [22]. Örneğin Şekil 2.3'de girdi adres verisi içindeki her bir adres bileşenine coğrafi kodlama sonucunda servis tarafından bir anahtar kelime atanır. Kullanılan servisin referans veri tabanından bu bileşenlere en yakın değer tespit edilerek, standart adres formatı oluşturulur.

Klasik coğrafi kodlamada adres standardizasyonunu sağlayan algoritma dışında adres dizilerini oluşturan bir enterpolasyon algoritması kullanılmaktadır. Bu algoritma başlangıç ve bitiş noktaları belirli bir isimli yol boyunca adres bilgisini referans veritabanındaki detaylardan biri ile eşleştirerek doğru konumu belirlemektedir (Şekil 2.4) [6].



Şekil 2.4 Adres lineer enterpolasyon algoritması örneği

Günümüzde internetteki ve web servislerindeki gelişmelerle birlikte klasik coğrafi kodlama işlemlerine ilave olarak çevrimiçi coğrafi kodlama sistemleri ortaya çıkmıştır. Çevrimiçi coğrafi kodlama internet tabanlı harita servisleri üzerinden geliştiğinden dolayı coğrafi kodlama işlemlerini otomatik olarak sağlayabilmektedir. Genellikle harita servisleri üzerinde çalışan algoritmalar ve geçerli referans veritabanı kullanıcıların erişimine kapalıdır (Şekil 2.5) [22].



Şekil 2.5 Çevrimiçi coğrafi kodlama işlem akışı [22]

Çevrimiçi coğrafi kodlama işlemi, klasik coğrafi kodlamaya göre farklılık göstermekle birlikte kendi içinde bazı avantaj ve dezavantajlara sahip olmaktadır (Çizelge 2.1).

Çizelge 2.1 Çevrimiçi coğrafi kodlama servislerinin avantaj ve dezavantajları

Avantajlar	**Dezavantajlar**
Kullanım kolaylığı sağlama	Referans veritabanı üzerinde bir kontrol sağlanamaması
Anında koordinat sonuçlarına erişim	Coğrafi kodlama işlem parametreleri üzerinde kontrol sağlanamaması
Kullanıcının referans veri tabanı üzerinde güncelleme yapmaya ihtiyaç duymaması	İnternet altyapısına bağlı kalması
CBS uygulamaları ile entegre olabilmesi	
Çok sayıda adresin aynı anda coğrafi olarak kodlanması	

2.3 Coğrafi Kodlama Ölçütleri

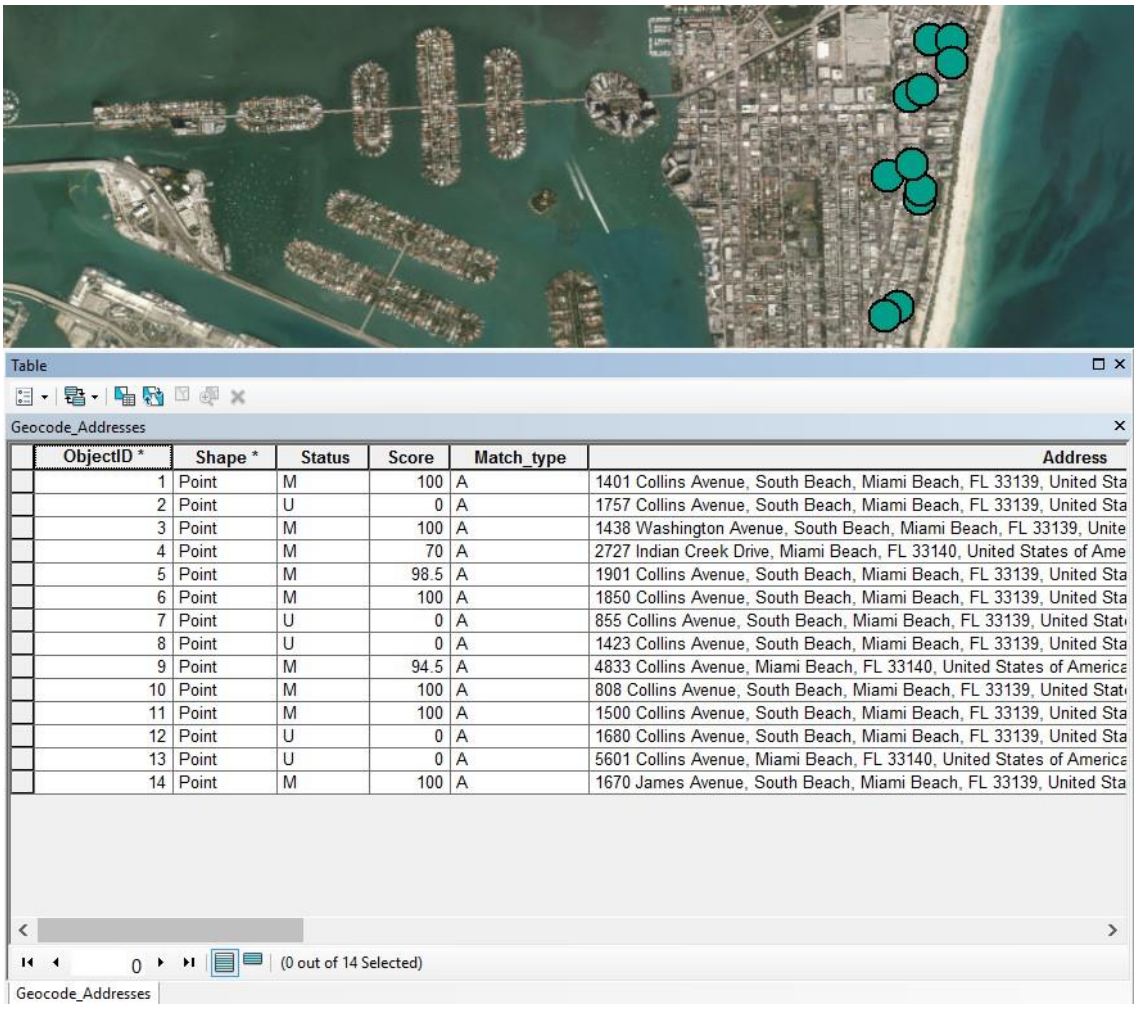
Kar amacı olmayan ya da ticari oluşumlardan yerel, bölgesel ve ulusal seviyedeki kurumlara kadar çeşitlenen araştırma gruplarının ve organizasyonların neredeyse tamamı birçok önemli çalışmada coğrafi kodlamadan yararlanmaktadır [60],[61]. Coğrafi kodlama ile elde edilen konumlar birçok uygulama için temel girdi verileri oluşturmaktadır. Coğrafi kodlama aşamasında ortaya çıkan sorunlar ya da belirsizlikler sonraki işlem adımlarına yayılmakta ve kullanıcıların yanlış ya da güvenilir olmayan yargılara varmasına sebebiyet vermektedir. Coğrafi kodlama sistemlerinin sürekli olarak değişen yeni ürünlerle çeşitlenmesi, kullanıcıların ihtiyaç duyduğu coğrafi kodlama sistemini belirleme sürecinde bazı temel ölçütleri dikkate almasını gerektirmektedir. Bu bağlamda, coğrafi kodlama çalışmalarının gerçekleşmesi esnasında dikkat edilmesi gereken ölçütler (Goldberg vd. [25])'e göre aşağıdaki gibi sıralanabilir:

- Coğrafi kodlama kalitesi
- Sistemsel entegrasyon

- Arayüzün yapısal uyumu
- Maliyet hesabı
- Kapasite hesabı

2.3.1 Coğrafi Kodlama Kalitesi

Coğrafi kodlama sonuçlarının kalitesi, verileri kullanan bireyler ve araştırma grupları için en önemli ölçütlerden biridir. Literatürde bahsi geçen kalite ölçütleri arasında eşleşme oranı, eşleme türü ve mekânsal doğruluk faktörleri yer almaktadır. Eşleşme oranı referans veritabanındaki veriyle coğrafi kodlama sonucu elde edilen veri arasındaki uyum olarak nitelendirilebilir. Eşleme türü, referans veri üzerinden yapılacak coğrafi kodlamanın yöntemi hakkında bilgi veren bir kalite ölçütüdür. Mekânsal doğruluk ise referans verideki bir detayın gerçek konumu ile coğrafi kodlama sonucu elde edilen konum arasındaki mesafe olarak adlandırılabilir. Bu aşamada referans veri tabanı ile coğrafi kodlama sonucu elde edilen veri arasında eşleşmenin mekânsal doğruluğuna dair bir puan elde edilir. Buna eşleşme puanı denilmektedir. Örneğin; ESRI firmasının *ArcGIS* yazılımında yer alan '*Geocode Addresses*' aracı ile yapılan coğrafi kodlama sonucunda eşleşme puan değerleri 0 ile 100 arasında değişmektedir (Şekil 2.6). Veri eşleşmesi tam olarak sağlandığında eşleşme puanı değeri 100 olacaktır. Eğer ki eşleşme puanı 85 ile 99 değerleri arasında ise iyi bir eşleşmenin gerçekleşmiş olduğu fakat bu değerlerin altında bir eşleşme puanı oluşmuş ise servisin bu eşleşmeyi tam olarak sağlayamadığı ortaya çıkmaktadır [62].



ObjectID *	Shape *	Status	Score	Match_type	Address
1	Point	M	100	A	1401 Collins Avenue, South Beach, Miami Beach, FL 33139, United Sta
2	Point	U	0	A	1757 Collins Avenue, South Beach, Miami Beach, FL 33139, United Sta
3	Point	M	100	A	1438 Washington Avenue, South Beach, Miami Beach, FL 33139, Unite
4	Point	M	70	A	2727 Indian Creek Drive, Miami Beach, FL 33140, United States of Ame
5	Point	M	98.5	A	1901 Collins Avenue, South Beach, Miami Beach, FL 33139, United Sta
6	Point	M	100	A	1850 Collins Avenue, South Beach, Miami Beach, FL 33139, United Sta
7	Point	U	0	A	855 Collins Avenue, South Beach, Miami Beach, FL 33139, United Stat
8	Point	U	0	A	1423 Collins Avenue, South Beach, Miami Beach, FL 33139, United Sta
9	Point	M	94.5	A	4833 Collins Avenue, Miami Beach, FL 33140, United States of America
10	Point	M	100	A	808 Collins Avenue, South Beach, Miami Beach, FL 33139, United Stat
11	Point	M	100	A	1500 Collins Avenue, South Beach, Miami Beach, FL 33139, United Sta
12	Point	U	0	A	1680 Collins Avenue, South Beach, Miami Beach, FL 33139, United Sta
13	Point	U	0	A	5601 Collins Avenue, Miami Beach, FL 33140, United States of America
14	Point	M	100	A	1670 James Avenue, South Beach, Miami Beach, FL 33139, United Sta

Şekil 2.6 Eşleşme oranı ve Eşleşme puanı bilgileri

2.3.2 Sistemsel Entegrasyon

Coğrafi kodlama işlemlerini uygulamak isteyen kullanıcıların ve büyük kurumların dikkat edeceği bir diğer ölçüt sistem entegrasyonudur. Bu başlık altındaki ölçütler işletim sistemi desteği, toplu coğrafi kodlama imkânı ve iş akış uyumu olarak üç kısımda incelenmektedir. İşletim sistemi desteği coğrafi kodlama yazılımının ilgili kurum/kurumlar tarafından kullanılan işletim sistemi (*Windows, Unix/Linux/Solaris* vd.) üzerinde uygulanabilir olmasıyla ilişkilidir. Toplu coğrafi kodlama imkânı bir veritabanından sorgulanmış çok sayıda adresin aynı anda toplu olarak coğrafi kodlanmasını (*'batch geocoding'*) sağlayan bir ölçüttür. İş akış uyumu ise coğrafi kodlama sisteminin büyük kurumlar tarafından kullanılan araçlar, sistemler (bağlı kütüphaneler) ile ilişkisini tanımlamaktadır.

2.3.3 Arayüzün Yapısal Durumu

Coğrafi kodlama çalışmaları, kişisel hakların mahremiyeti açısından gizli kalması gereken durumlar ya da veri güvenliği nedeniyle (örneğin, sağlıkla ilişkili kayıtların konumlandırılmasında vb.), masaüstü odaklı (*ArcGIS, MapInfo* vb.) veya kurum içi platformlar (*Geocoder.us, 'Alteryx Community'* vb.) üzerinden gerçekleştirilmektedir. Bunların dışında, internet üzerinden (*Uygulama Programlama Arayüzü* – API) kurum içi diğer yazılımlar arasında birlikte çalışabilen coğrafi kodlama platformları oluşturulabilmektedir.

2.3.4 Maliyet Hesabı

Bir coğrafi kodlama sisteminin gerçek maliyeti için miktar belirlemek oldukça zordur. Ancak, coğrafi kodlama sistemlerinin bazı kıstaslarına göre bir maliyet hesabı oluşturmak mümkündür. Özellikle yazılım için gerekli olan lisans ücreti, istenen referans verileri oluşturmak için gerekli lisans ücreti ve yazılım üzerine eklenecek olan destek sözleşmeler için gerekli maliyet, yazılım satan firmalar tarafından yıllık olarak sabitlenmiş fiyatlar üzerinden belirlenmektedir.

2.3.5 Kapasite Hesabı

Bir coğrafi kodlama sistemine ait temel kapasiteler veya kabiliyetler belirli bir kullanım planlaması içerisinde bir coğrafi kodlama sisteminin uygunluğunu değerlendirmede önemli bir role sahiptir. Bu tür ölçütler, coğrafi kodlama sistemleri içerisinde kaç adet adres verisi ile çalışabilme (sisteme yüklenebilen adres sayısı) ve adres verilerini zamansal olarak hangi aralıklarda kullanabilme (saat, gün, ay ve yıl bazında) gibi durumlar olarak düşünülebilir. Örnek olarak, *ESRI, Locative Technology, Microsoft, Google* ve *America Online* gibi farklı ticari şirketlerin kullanıcılarına sunduğu coğrafi kodlama sistemlerinin erişim imkanları Çizelge 2.2'de gösterilmiştir.

Çizelge 2.2 Servisler arasındaki kapasite farklılıkları

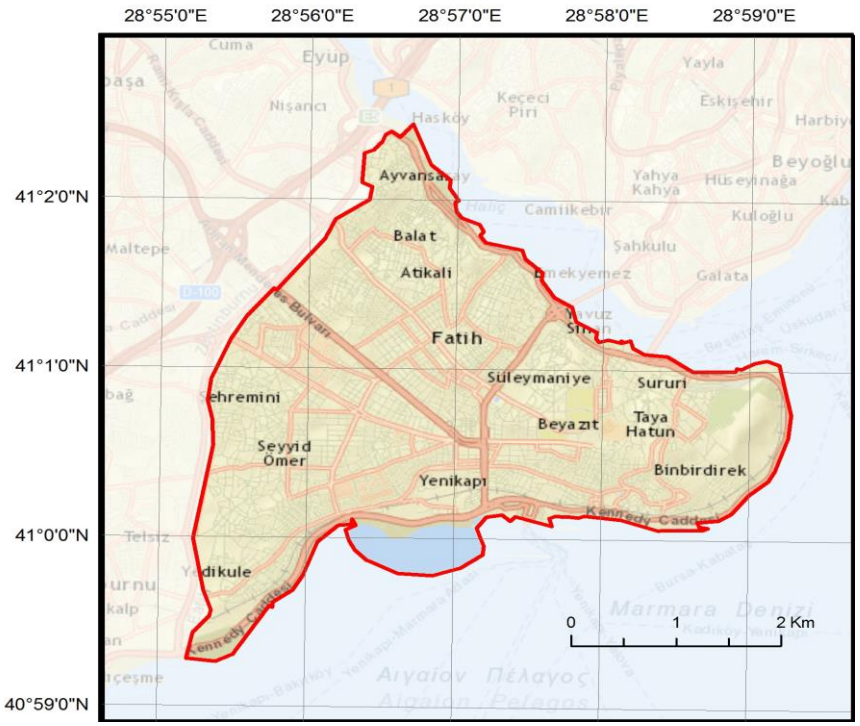
Servis İsmi	**Şirket**	**Kapasite**
ArcGIS	*ESRI*	Lisans zorunluluğuna bağlı olarak limitsiz erişim
Geocoder.us	*Locative Tech.*	En yüksek 1 000 adres verisi
BingMaps	*Microsoft*	Yıllık 1 000 000 adres verisi
Google Maps	*Google*	Günlük 2 500 adres verisi
MapQuest	*America Online, Inc. (AOL)*	Aylık 15 000 adres verisi

Her coğrafi kodlama servisinin yukarıda bahsedilen ölçütler doğrultusunda tasarlanmış kendine özgü bir hizmet yapısı mevcuttur. Ölçütlere göre yapılan karşılaştırmalardan önce adres verilerinin coğrafi kodlama servislerine olan uygunluklarının tespit edilmesi önemli işlem adımlarından biridir.

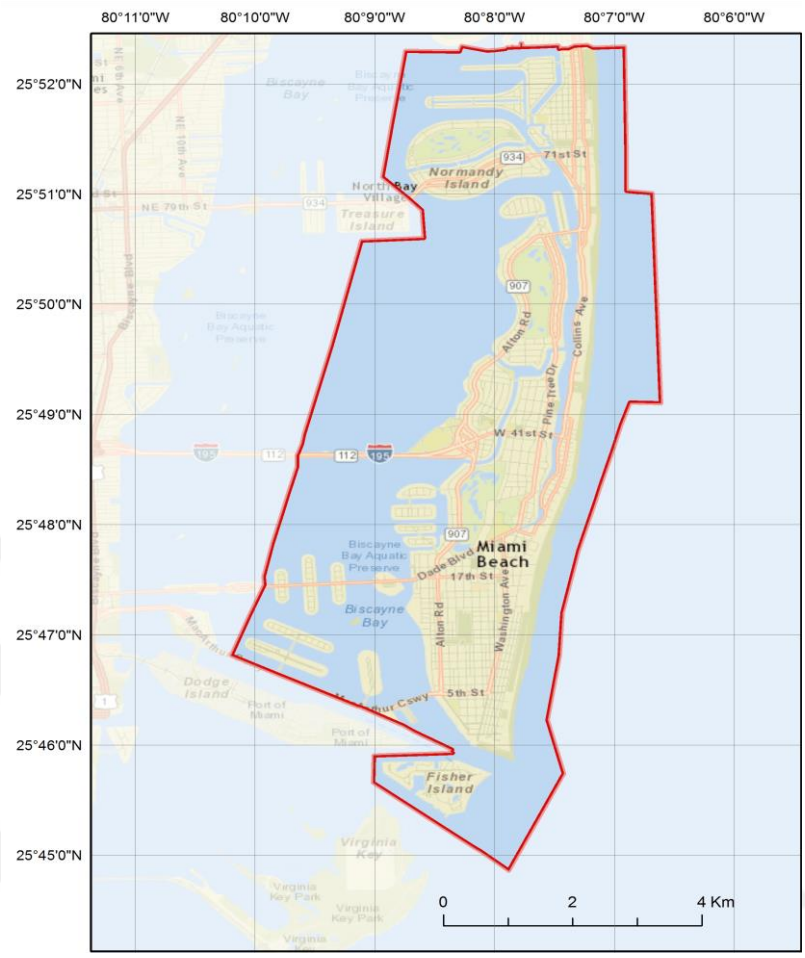
BÖLÜM 3

ÇEVRİMİÇİ COĞRAFİ KODLAMA HİZMETLERİNDE KALİTE ARAŞTIRMASI

Bu çalışma kapsamında çevrimiçi coğrafi kodlama işlemlerinin kalitesi ulusal ve uluslararası adres yapıları üzerinden karşılaştırmalı olarak incelenmiştir. Test için kullanılan adres verileri Türkiye ve Amerika Birleşik Devletleri'ndeki otel benzeri konaklama tesislerine aittir. Uygulama bölgesi olarak otellerin yoğun olduğu iki bölge seçilmiştir. Birinci çalışma bölgesi İstanbul ili Fatih ilçesi sınırları içinde kalan otelleri kapsamaktadır (Şekil 3.1). İkinci bölge Florida eyaletine bağlı Miami Beach sınırları içinde kalan otelleri içermektedir (Şekil 3.2).



Şekil 3.1 Fatih bölgesi ve Tarihi yarımada



Şekil 3.2 Miami Beach bölgesi

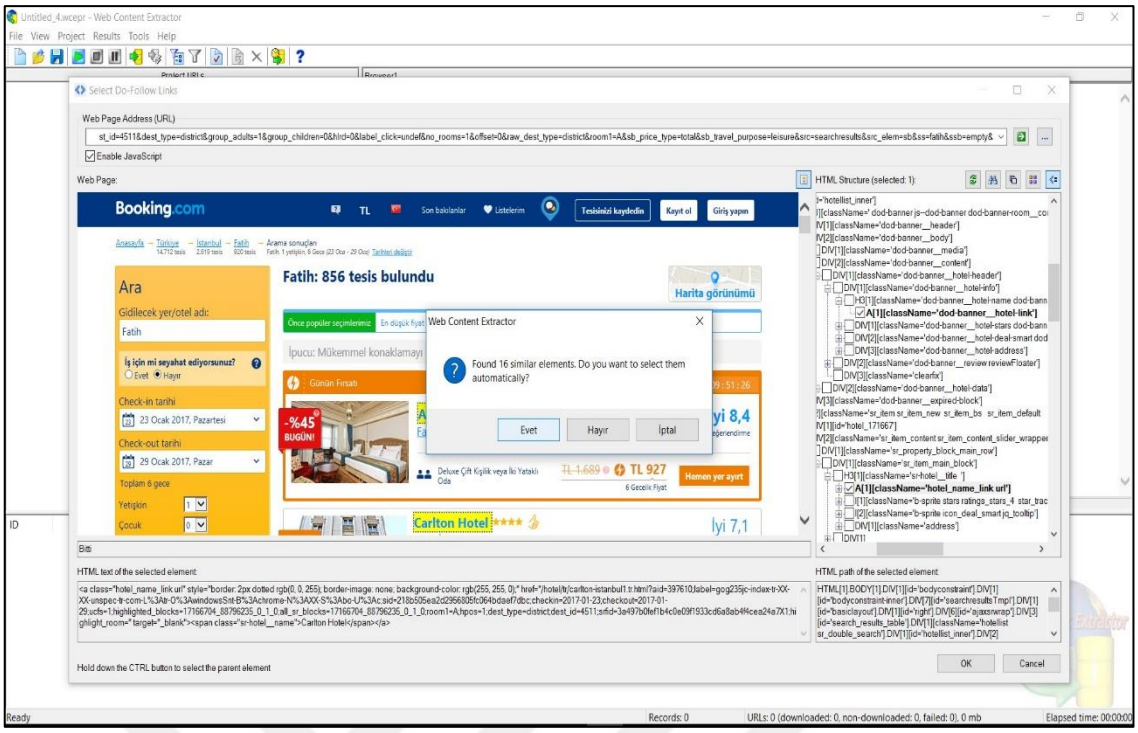
3.1 Referans Verilerin Oluşturulması

Hazırlık aşamasında çalışmaya altlık oluşturan konaklama tesisleri ve bunların adres verileri *www.booking.com* ve *www.hotels.com* olarak bilinen iki web sitesi üzerinden elde edilmiştir. Her iki site üzerinden çalışma bölgelerinde bulunan oteller arandığında çok sayıda farklı otele rastlanmaktadır (Çizelge 3.1). Belirlenen otellerin büyük bir kısmı her iki web sitesinde de bulunan ortak oteller olmasına karşın bazı otel isimlerinin yazımında ve adres tanımlamalarında farklılıklar olduğu görülmüştür. Özellikle Fatih bölgesindeki otellerin adres tanımlamalarındaki farklılıklar, ülkemizde standart bir adres yapısının oluşturulamamış olmasından dolayı, oldukça belirgindir [63].

Çizelge 3.1 Tespit edilen konaklama tesislerine ait sayılar

	Booking.com	**Hotels.com**
Fatih	250	270
Miami Beach	271	382

Her iki web sitesi üzerinden otel bilgilerinin toplanması için '*web scraping*' işlemi gerçekleştirilmiştir. *Web scraping*, internet siteleri üzerinde belirli konumlarda yer alan standart bilgileri toplamak için kullanılan bir veri kazıma işlemidir. *Web scraping* uzun süredir kullanılan bir terim olmasına karşın veri madenciliği alanında '*screen scraping*' ve '*web harvesting*' terimleri de bu işlemi tanımlamaktadır [64]. *Visual Web Ripper, Web Content Extractor, Web Scraper Plus+* vb. çok sayıda ticari *web scraping* yazılımı mevcuttur [65]. Bu yazılım araçlarının ücretsiz olarak kullanımları sınırlıdır ve çeşitli özelliklerine göre yazılım ücretleri değişmektedir. Bu çalışma kapsamında verilerin toplanabilmesi için *Newprosoft* şirketi tarafından geliştirilen sınırlı sayıda ücretsiz işlem yapma yetisine sahip *Web Content Extractor V.8.4* yazılımı kullanılmıştır (Şekil 3.3).



Şekil 3.3 *Web Content Extractor V.8.4* genel görünüm

Web Content Extractor yazılımı:

- Web sitelerinden istenilen verileri otomatik olarak ayıklayabilme,
- Verileri farklı formatlarda saklama (*Microsoft Access* veritabanı, *Microsoft Excel*, metin (.txt), *HTML, XML* ve *SQL* komut dosyaları),
- Programı otomatikleştirmek için kullanılabilecek bir dizi komut satırı seçeneğini destekleme,
- Şifre ile korumalı web sitelerinden verileri ayıklayabilme,
- Birden çok *Proxy* sunucusu kullanabilme,
- Kullanımı kolay yapılandırma sihirbazına sahip olma

gibi çeşitli özelliklere sahiptir.

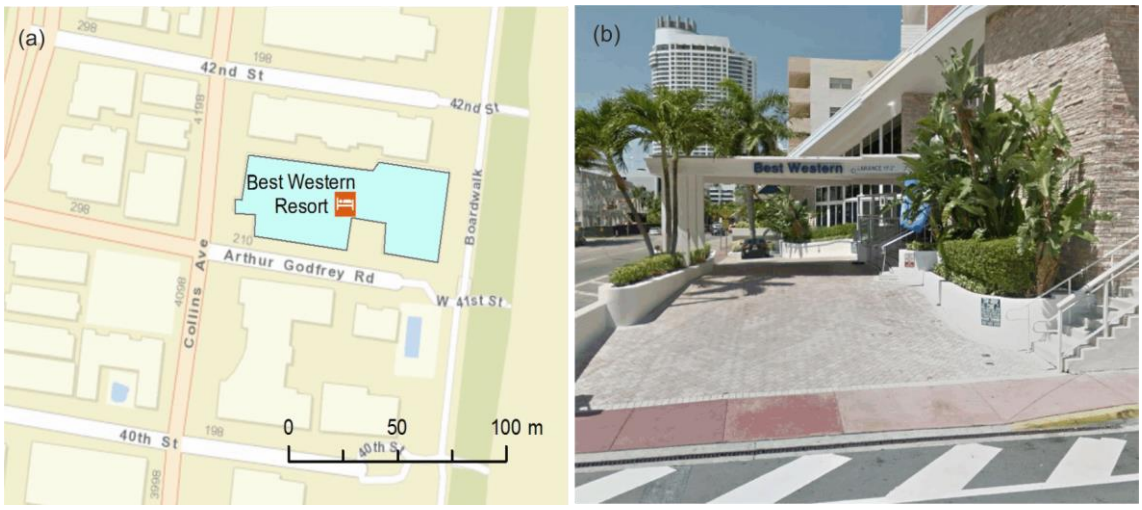
Web Content Extractor yazılımı ile gerçekleştirilen işlemler sonucunda her iki bölgedeki konaklama tesislerinin isim ve adres verileri iki ayrı web sitesi üzerinden otomatik olarak toplanmıştır. Aynı otellere ait adres yazımlarındaki farklılıklar incelendiğinde adreslerin *www.booking.com* sitesinde daha açık ve anlaşılır ifade edildiği

belirlendiğinden referans otel ve adres bilgileri olarak bu siteden toplanan verilerin kullanılmasına karar verilmiştir.

Fatih ve Miami Beach bölgelerinden sırasıyla toplam 250 ve 271 adet toplanmış otele ait adresler incelendiğinde tekrarlı adreslerin olduğu anlaşılmıştır. Gerekli ayıklama işlemi yapıldıktan sonra ilgili otellerin *Google* ve *Bing maps* servisleri üzerinden otomatik çevrimiçi coğrafi kodlama yöntemiyle konumları belirlenmiştir. Bu aşamada gerçekleştirilen coğrafi kodlama yönteminin ayrıntıları adres verileri üzerinden yapılan coğrafi kodlama ile benzerdir. Yöntemin ayrıntıları Bölüm 3.2'de açıklanacaktır. Her iki harita servisinde aynı isimle tanımlı ve coğrafi olarak kodlanan otellerin sokak görüntüleri incelenerek Fatih'de 74, Miami Beach'de ise 82 otelin konumu kesinleştirilmiştir (Şekil 3.4 ve Şekil 3.5). Her iki bölgeden belirlenen otellerin konumları bina çokgeni oluşturacak şekilde, *WGS84* datumu temel alınarak, *ArcGIS World Street Map* servisi üzerinden sayısallaştırılmıştır. Coğrafi kodlamaya referans oluşturacak konumsal veriler noktasal olması gerektiğinden, bina detaylarının ağırlık merkezleri türetilmiştir.



Şekil 3.4 Fatih'de bir referans otelin (a) harita ve (b) perspektif görüntüsü



Şekil 3.5 Miami Beach'de bir referans otelin (a) harita ve (b) perspektif görüntüsü

3.2 Adres Verilerinden Coğrafi Kodlama

Fatih ve Miami Beach bölgelerinde ağırlık merkezleri temel alınarak konumlandırılan oteller ve adresleri, *WGS84* datumlu iki ayrı coğrafi veri setinde referans nokta detaylar olarak depolanmıştır. Detaylara ait adresler *Google maps* ve *Bing maps* servisleri kullanılarak coğrafi kodlanmıştır. Bu işlem *Google maps Geocoding API* ve *Bing maps Locations API* servisleri kullanılarak *Python* programla dilinde hazırlanan coğrafi kodlama kütüphanesi kullanılarak çevrimiçi gerçekleştirilmiştir (Şekil 3.6).

```
#-*- coding: utf-8 -*-

import (1)        ### (1) Kullanılacak kütüphaneyi  giriniz. ###
dosya = open(2)   ### (2) Girdi dosya yolunu giriniz. ###
dosya2 = open(3)  ### (3) Çıktı dosya yolunu giriniz. ###
liste= dosya.read().splitlines() (4) ### (4) Girdi dosyasını liste haline dönüştürme ###
for adres in liste:
    g = (1).(5)(6) ### (5) Kullanılacak servisi seçiniz. ### ### (6) Servise ait parametreleri giriniz.###
    txt= str(g.(7)) + "\n" ### (7) İstenilen sonuç çıktısını giriniz. ###
    dosya2.write(txt)
dosya2.flush()
dosya2.close()
dosya.close()
```

Şekil 3.6 *Python* ile çevrimiçi coğrafi kodlama

Google maps Geocoding API ve *Bing maps Locations API* servisleri, coğrafi kodlama işlemini gerçekleştirirken ilk olarak adres verilerini kendi içinde düzenli bileşenlerine ayırmaktadır. Servisler bileşenlere birer anahtar isim tanımladıktan sonra, parçalanmış adres bileşenlerini tek tek kendi referans veritabanı ile karşılaştırarak en iyi aday

bilgisini göstermektedir. Adres verisi okutulan servis, kendi referans veritabanından en uygun standart adresi belirledikten sonra eşleşmiş kendi standart adres değerlerini (örneğin; şehir adı, ilçe adı, mahalle adı, yol adı, kapı numarası, posta kodu, eşleşme durumu vb.) ve *WGS84* datumundaki enlem-boylam değerlerini sunmaktadır.

3.2.1 Konumsal Uyum

Bu çalışmada konumsal uyum iki aşamada değerlendirilmektedir:

Birinci aşamada, her iki bölgedeki referans adresler kullanılarak *Google maps* ve *Bing maps* servisleri üzerinden elde edilen ilgi noktalarının *ArcGIS World Street Map* servisi üzerinden sayısallaştırılan çokgen detaylar ile mekânsal ilişkisi incelenmiştir (Çizelge 3.2). Fatih'de *Google maps* üzerinden coğrafi olarak kodlanan otellere ait 61 adet ilgi noktası sayısallaştırılan çokgen detayların içinde yer almıştır. *Bing maps* üzerinden elde edilen otellerde çokgen detayın içinde yer alan nokta sayısı 44 adettir. Buna karşın, sayısallaştırılan referans çokgenlerin dışında kalan yani hatalı kodlanan ilgi noktası sayıları; *Google maps* servisi üzerinden 13, *Bing maps* servisi üzerinden 30 adettir. Miami Beach'de *Google maps*'den elde edilen 72 adet ilgi noktası sayısallaştırılan çokgen detaların içinde yer almıştır. *Bing maps*'den çıkartılan otellerde çokgen detayların içinde yer alan nokta sayısı 61 adettir. Referans çokgenlerin dışında kalan ilgi noktası sayıları; *Google maps* üzerinden 10 adet, *Bing maps* üzerinden 21 adet olarak belirlenmiştir.

Çizelge 3.2 Otel çokgenlerinin içinde ve dışında yer alan ilgi noktaları

	Google maps		*Bing maps*	
	İçinde	*Dışında*	*İçinde*	*Dışında*
Fatih	61	13	44	30
Miami Beach	72	10	61	21

Her iki çalışma bölgesinde, *Google* ve *Bing Maps*'de hatalı olarak coğrafi kodlanan (referans çokgenlerin dışında kalan) otel noktalarının ait oldukları en yakın çokgene olan uzaklıkları tespit edilmiştir (Şekil 3.7).



Şekil 3.7 Çokgen cephesine olan en yakın mesafeler

Fatih ve Miami Beach bölgeleri için belirlenen uzaklıklara göre hesaplanan toplam, ortalama, en uzak ve en yakın mesafe değerleri Çizelge 3.3 ve Çizelge 3.4'de sırasıyla verilmektedir. Burada *n*, referans çokgen detaylar dışında kalan hatalı kodlanmış otel noktalarının sayısını göstermektedir.

Çizelge 3.3 Fatih bölgesi için cepheye olan mesafeler

	Toplam uzaklık	Ortalama uzaklık	En uzak mesafe	En yakın mesafe
Google Maps (*n*=13)	121.3 m	9.3 m	28.3 m	0.6 m
Bing Maps (*n*=30)	332.8 m	11.1 m	60.5 m	0.004 m

Çizelge 3.4 Miami Beach bölgesi için cepheye olan mesafeler

	Toplam uzaklık	Ortalama uzaklık	En uzak mesafe	En yakın mesafe
Google Maps (*n*=10)	21.5 m	2.1 m	6.3 m	0.3 m
Bing Maps (*n*=21)	62.2 m	2.9 m	8.5 m	0.7 m

Her iki bölgedeki sonuçlar incelendiğinde, *Bing Maps*'e ait otel noktalarının sayısı *Google Maps*'e ait nokta sayılarından fazla olduğu görülmektedir. Cephelere olan toplam, ortalama ve en uzak mesafe değerleri *Bing Maps*'de daha fazladır. *Bing Maps*'in en yakın mesafe değeri sadece Fatih'de daha düşüktür. Bu değerler, coğrafi kodlamanın, her iki çalışma bölgesinde, *Google maps* servisi tarafından daha başarılı gerçekleştirildiğini göstermektedir.

İkinci aşamada, her iki bölgede *Google maps* ve *Bing maps* üzerinden elde edilen coğrafi koordinatlar ve çokgenlerin ağırlık merkezlerinden türetilmiş, referans kabul edilen, otel koordinatları karşılaştırılmıştır. Servislerden elde edilen coğrafi koordinatlar enlem-boylam değerleriyle ifade edildiğinden, karşılaştırma öncesinde mesafelerin hesaplanabilmesi için projeksiyon koordinatlarına dönüşüm yapılmıştır. Dönüşüm işlemlerinden sonra, Fatih için *ITRF96* datumunda 3°'lik Gauss Kruger projeksiyon koordinatları, Miami Beach için *NAD83* datumunda StatePlane Florida East projeksiyon koordinatları kullanılmıştır.

Projeksiyon koordinatları arasındaki farklar referans detayların koordinatları (Sağa değer: E_{ref} ve Yukarı değer: N_{ref}) ile *Google maps* servisi koordinatları (Sağa değer: E_g ve Yukarı değer: N_g) ve *Bing maps* servisi koordinatları (Sağa değer: E_b ve Yukarı değer: N_b) üzerinden Eşitlik 3.1 yardımı ile çıkartılmıştır.

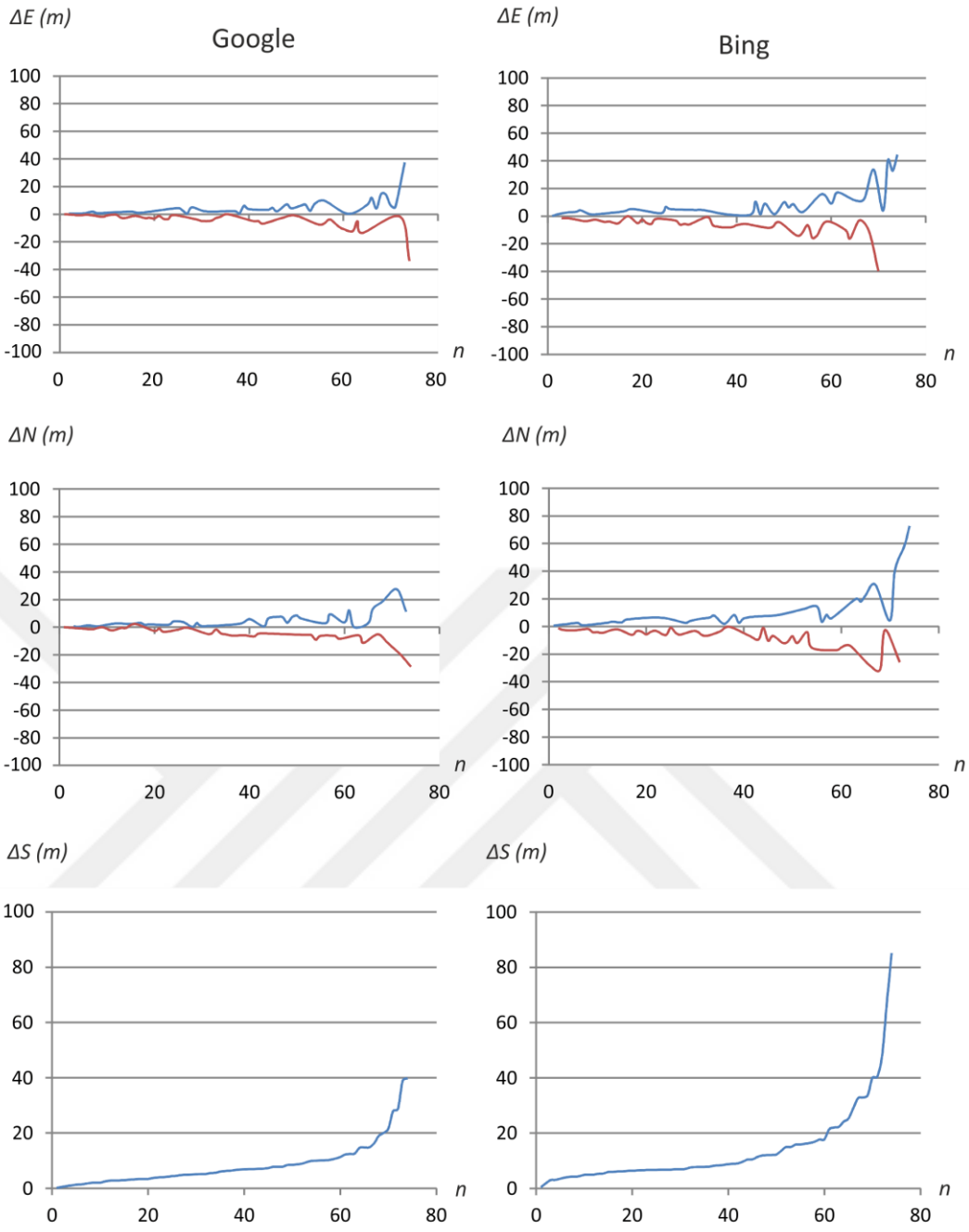
$$\Delta E = E_{ref} - E_i \qquad \Delta N = N_{ref} - N_i \tag{3.1}$$

Eşitlikte geçen E_i ve N_i ilgili servislerden elde edilen koordinat değerlerini göstermektedir. Referans veri ve servisler arasındaki konumsal farklar (ΔS) ise; koordinat fark değerleri kullanılarak Eşitlik 3.2 ile hesaplanmıştır.

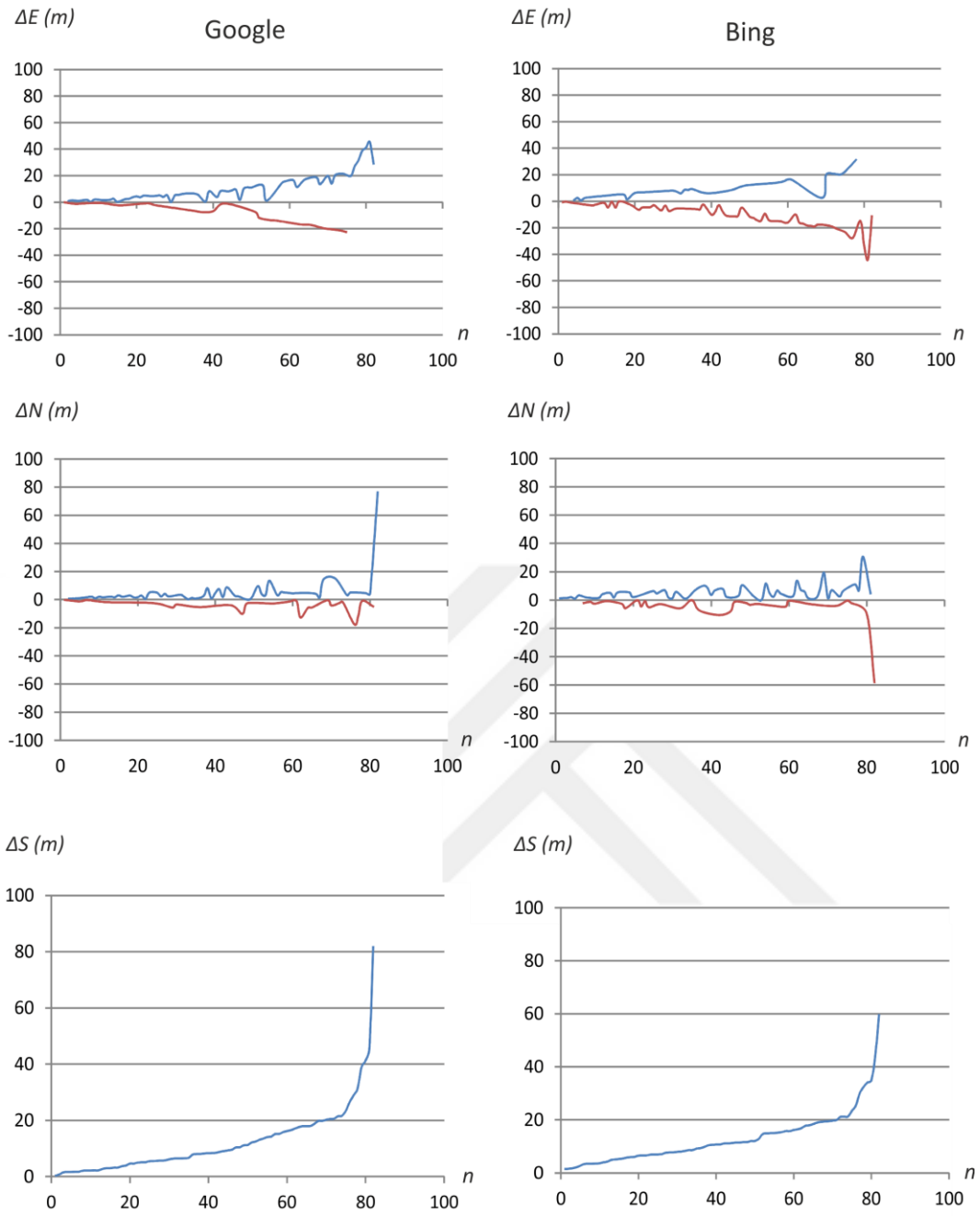
$$\Delta S = \sqrt{(E_{ref} - E_i)^2 + (N_{ref} - N_i)^2} \quad \text{veya} \quad \Delta S = \sqrt{\Delta E^2 + \Delta N^2} \tag{3.2}$$

Fatih ve Miami Beach bölgeleri için her bir ölçünün ΔE, ΔN ve ΔS sapma miktarı sırasıyla Şekil 3.8 ve Şekil 3.9'daki grafiklerle gösterilmiştir. Tüm ölçü farkları ΔS sapma miktarlarına göre küçükten büyüğe sıralanarak belirlenmiştir. Grafiklerin yatay ekseni sıfır ölçü farkını ifade etmektedir. Yatay eksenin üstünde kalan değerler referans değer ile olan pozitif sapma miktarlarını, altında kalan değerler ise negatif sapma miktarlarını göstermektedir. Fatih bölgesindeki tüm grafiklerde, *Bing maps* servisinin verdiği değerlerin yatay eksenden daha fazla sapma gösterdiği görünmektedir. Bu durumda, *Google*'ın *Bing*'e göre daha iyi coğrafi kodlama yaptığı söylenebilir.

Miami beach bölgesine ait grafiklerde, sağa ve yukarı değer farklarının yatay eksenden sapmaları her iki servis için benzerdir. Ancak yukarı değer grafiğine bakıldığında her iki servis için son ölçü değerlerinde büyük sapmalar olduğu görülmektedir. İlgili otellerin alanlarının çok büyük olması, servislerden hesaplanan yukarı değer ve ağırlık merkezi arasında, büyük miktarda sapmaya neden olmuştur. Genel olarak grafiklerde benzerliklerin olmasına karşın *Bing maps* servisinin *Google maps*'e göre daha iyi coğrafi kodlama yaptığı söylenebilir.



Şekil 3.8 Fatih bölgesi için her bir ölçünün ΔE, ΔN ve ΔS sapma miktarları



Şekil 3.9 Miami Beach için her bir ölçünün ΔE, ΔN ve ΔS sapma miktarları

Burada *n*, coğrafi kodlama işleminin gerçekleştirildiği çalışma bölgesindeki toplam otel sayısını göstermektedir. Çizelge 3.5 ve Çizelge 3.6, her iki çalışma alanı için *Min, Max, Medyan, Ort* ve *Norm* ile ifade edilen istatistiki bilgileri içermektedir.

Çizelge 3.5 Fatih bölgesi için istatistiksel değerler

	ΔE		ΔN		ΔS	
	$E_{ref} - E_g$	$E_{ref} - E_b$	$N_{ref} - N_g$	$N_{ref} - N_b$	$S_{ref} - S_g$	$S_{ref} - S_b$
Min	-33.80	-39.78	-28.62	-31.02	0.13	0.42
Max	37.28	44.30	27.47	72.60	39.72	85.05
Medyan	-0.43	-1.15	-0.06	-1.67	6.44	8.24
Ort	0.15	0.83	-0.62	0.92	8.30	13.89
Norm	67.82	105.40	70.10	134.88	97.54	171.17

Çizelge 3.6 Miami Beach bölgesi için istatistiksel değerler

	ΔE		ΔN		ΔS	
	$E_{ref} - E_g$	$E_{ref} - E_b$	$N_{ref} - N_g$	$N_{ref} - N_b$	$S_{ref} - S_g$	$S_{ref} - S_b$
Min	-22.77	-43.67	-17.75	-58.72	0.26	1.44
Max	45.25	31.66	76.88	30.61	81.95	59.71
Medyan	3.61	-4.66	1.63	1.79	8.39	10.89
Ort	4.50	-4.02	1.98	1.62	12.15	12.93
Norm	126.18	120.56	90.50	83.27	155.28	146.52

Min, Max, Medyan, Ort ve *Norm* sırasıyla detaylar arasındaki farkların; en küçük, en büyük, ortanca, ortalama ve norm değerlerini göstermektedir. En küçük değer (*Min*), bir veri örneklemindeki en küçük gözlem değerine karşılık gelir. En büyük değer (*Max*) ise, bir veri örneklemindeki en büyük gözlem değeridir. Ortanca (*Medyan*), bir

örneklem veri serisini küçükten büyüğe doğru sıraladığımızda, seriyi ortadan ikiye ayıran değer olarak bilinmektedir. *Medyan* değeri, Eşitlik 3.3 ile hesaplanmaktadır.

$$Medyan = \frac{(n+1)}{2} \tag{3.3}$$

Eşitlikte n, ölçü sayısına karşılık gelmektedir. *Medyan* ise gösterilecek değerin sıra numarasını vermektedir. Eğer ölçü sayısı tek ise medyanın sıra numarası bir tamsayı olacaktır ve doğrudan *medyan* değeri bulunur. Eğer ölçü sayısı çift ise medyanın sıra numarası ondalıklı bir sayı çıkar. Bu durumda bu sayının etrafındaki iki değerin aritmetik ortalaması *medyan* değerine karşılık gelmektedir.

Ortalama (*Ort*), bir sayı dizisindeki elemanların toplamının eleman sayısına bölünmesi ile elde edilir. İstatistik bilim dalında en çok kullanan merkezi eğilim ölçüsüne karşılık gelmektedir. Ortalama değer Eşitlik 3.4 ile hesaplanır.

$$Ort = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{3.4}$$

Eşitlikte n ölçü sayısını, x_i ise her bir ölçü için Eşitlik 3.1 ve Eşitlik 3.2'den hesaplanan fark değerini göstermektedir.

Norm (*Norm*), her bir ölçü değerinin karşılık geldiği referans değer ile farklarının hesaplanmasında kullanılan matematiksel bir büyüklük olarak tanımlanmaktadır. Ölçü değerleri ile referans değer arasındaki uzunluk vektör olarak kabul edilir ve bu vektörün uzunluk değeri hesaplanarak *norm* değeri belirlenir. *Norm* değeri Eşitlik 3.5 ile hesaplanır.

$$Norm = \sqrt{\sum_{i}^{n} {x_i}^2} \tag{3.5}$$

Her iki test bölgesinde ΔE ve ΔN'in en küçük değerleri negative, en büyük değerleriyse pozitiftirtir. Çizelge 3.5 ve Çizelge 3.6'daki, ΔE ve ΔN en küçük değer değişimleri Fatih'de -28 m ile -40 m arasındayken, Miami Beach'de -17 m ile -59 m arasındadır. Çalışma alanlarındaki en büyük değer değişimleri ise Fatih'de 27 m ile 73 m, Miami Beach'de 30 m ile 77 m arasındadır. ΔS için belirlenen en büyük değer, servislerin bulduğu en başarısız konumu göstermektedir. *Google maps* servisi Fatih'de yaklaşık en

fazla 40 m konumsal sapma göstermişken, *Bing maps* servisinde bu değer 85 m civarındadır. Miami Beach bölgesinde ise en büyük konumsal sapma değerleri *Google maps* için 82 m ve *Bing maps* için 60 m civarındadır. ΔS'in küçük değeri ise servisler tarafından bulunan en başarılı konumsal sonuçları göstermektedir. En küçük ΔS değerleri her iki çalışma bölgesindeki iki servis için yaklaşık bir metre düzeyindedir.

Medyan değerleri, ΔE ve ΔN yönünde Fatih'de yaklaşık 1 m civarındayken, Miami Beach'de 3 m civarında sapma göstermektedir. Medyan değerinin ΔS konumsal sapma miktarı, her iki bölge için *Google maps* servisinin *Bing maps* servisine göre yaklaşık 2 m daha iyidir. ΔE ve ΔN yönünde medyan değerleri için söylenenler ortalama değerler için de söylenebilir. Ortalama değerin ΔS konumsal sapma miktarı, Fatih için *Google maps* servisinin *Bing maps* servisine göre yaklaşık 5 m daha iyi iken, Miami Beach için iki servisin sapma miktarları yaklaşık aynıdır.

Norm değerleri, Fatih'de ΔE ve ΔN yönünde ortalama *Google maps* servisinde 69 m, *Bing maps* servisinde 120 m civarındadır. ΔS konumsal sapma miktarı, *Google maps* servisinde 98 m, *Bing maps* servisinde ise 171 m düzeyindedir. Miami Beach'de ΔE ve ΔN yönünde ortalama *Google maps* servisinde 108 m, *Bing maps* servisinde 102 m civarındadır. ΔS konumsal sapma miktarı, *Google maps* servisinde 155 m, *Bing maps* servisinde ise 146 m düzeyindedir.

Bu aşamada elde edilen değerlere göre; Fatih'de *Google maps* servisi tarafından gerçekleştirilen coğrafi kodlama *Bing maps* servisine göre daha iyi sonuçlar vermektedir. Buna karşın, Miami Beach'de her iki servisin sonuçları birbirine oldukça yakın olduğu söylenebilir.

3.2.2 Adres Uyumu

Her iki çalışma bölgesindeki referans adresler ile *Google* ve *Bing Maps* üzerinden elde edilen standart adres değerleri *'Levenshtein mesafesi'* olarak bilinen bir algoritma kullanılarak karşılaştırılmıştır. Adresler arasındaki uyumu belirlemek için kullanılan *Levenshtein mesafesi* arama motorları başta olmak üzere bilgi teknolojilerinde sıkça kullanılan bir algoritmadır. Pratikte arama sonuçlarında kelimeler arasındaki benzerliği derecelendirmek için kullanılan algoritma, basitçe iki dizi, iki kelime ve iki cümle

arasındaki benzerliği göstermekte ve bunlar arasındaki değiştirme ve ekleme işlemlerini hesaplamaktadır [66]. *Levenshtein mesafesi*'nin hesaplanması için kullanılan örnek matrisler Şekil 3.10'da gösterilmiştir. Buna göre, A matrisinde "Kitap" isminin "Kalem" ismine dönüşmesi için 4 işlemin, B matrisinde ise "Tatil" ve "Sahil" isimlerinin birbirine dönüşmesi için ise 2 işlemin yapılması gerektiği anlaşılmaktadır [67].

A matrisi

	K	İ	T	A	P
K	0	1	2	3	4
A	1	1	2	3	4
L	2	2	2	3	4
E	3	3	3	3	4
M	4	4	4	4	**4**

B matrisi

	T	A	T	İ	L
S	1	2	3	4	5
A	2	1	2	3	4
H	3	2	2	3	4
İ	4	3	3	2	3
L	5	4	4	3	**2**

Şekil 3.10 Levenshtein mesafesi için örnek matrisler [67]

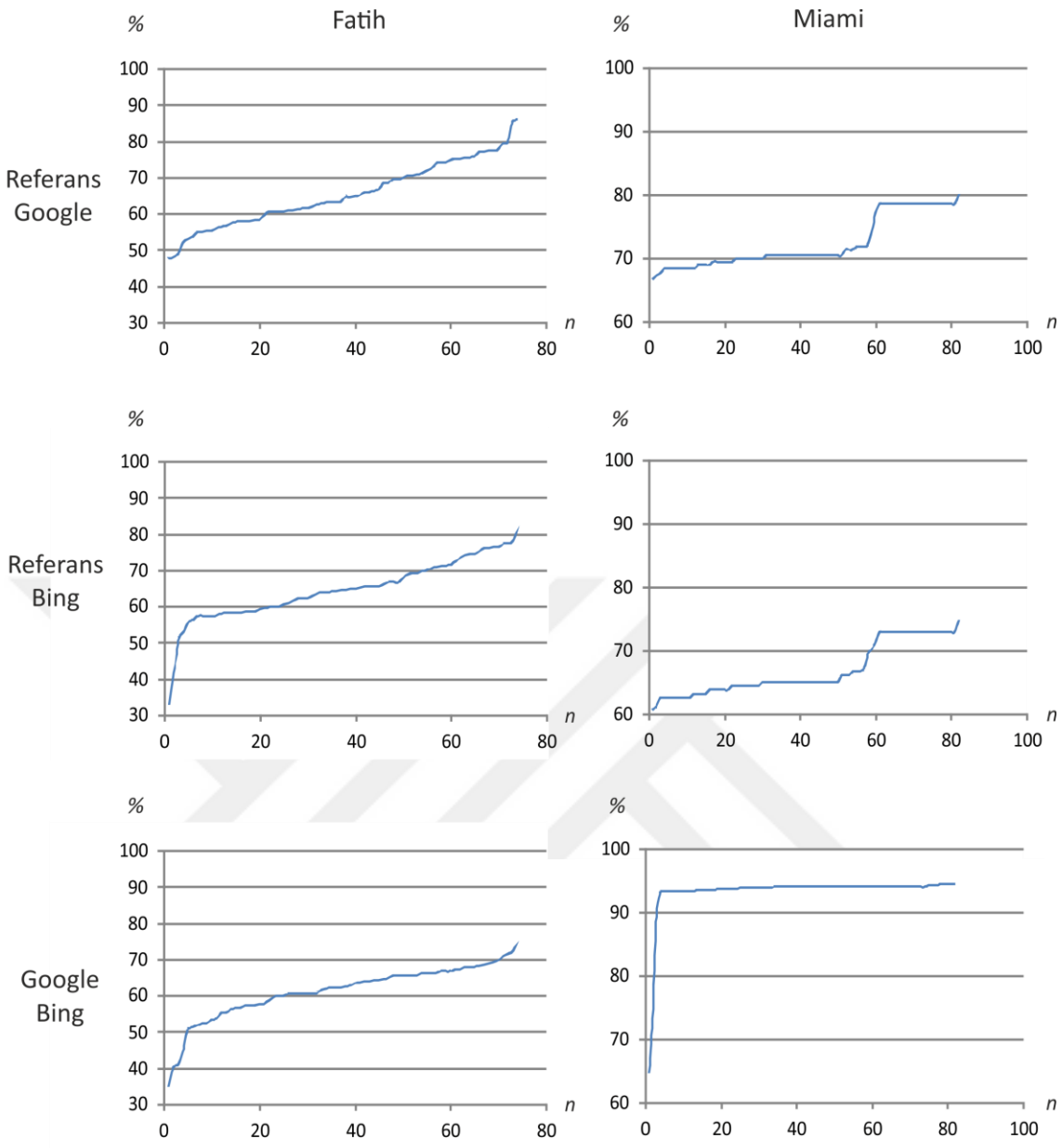
Servislerin birbirleriyle ve referans veriyle olan uyumları, *Python* dilinde hazırlanan *Levenshtein mesafesi* değerlerini hesaplayan bir kütüphane kullanılarak belirlenmiştir (Şekil 3.11).

```
#-*- coding: utf-8 -*-

import (1)    ### (1) Kullanılacak kütüphaneyi giriniz. ###
dosya = open(2)   ### (2) Girdi dosya yolunu giriniz. ###
dosya2 = open(3)  ### (3) 2. Girdi dosya yolunu giriniz. ###
dosya3 = open(4)  ### (4) Çıktı dosya yolunu giriniz. ###
liste= dosya.read().splitlines()    (5)  ### (5) Girdi dosyasını liste haline dönüştürme ###
liste2= dosya2.read().splitlines()  (6)  ### (6) 2. Girdi dosyasını liste haline dönüştürme ###
for i in range(len(liste)):
    g = (1). (7) (liste[i], liste2[i])    ### (7) İstenilen sonuç çıktısını giriniz. ###
    txt = str(g) + "\n"
    dosya3.write(txt)
dosya3.flush()
dosya3.close()
dosya2.close()
dosya.close()
```

Şekil 3.11 Python ile Levenshtein mesafesi'nin hesaplanması

Servisler tarafından üretilen standart adres formatlarının, referans adres formatı ve birbirleriyle olan uyumu Fatih ve Miami Beach için Şekil 3.12'de grafik olarak gösterilmiştir.



Şekil 3.12 Coğrafi kodlamayla elde edilen adres verilerinin uyum yüzdeleri

Şekil 3.12'de, Fatih bölgesinin değerleri incelendiğinde;

- *Google maps* uyum değerlerinin yaklaşık %50 ile %85 arasında,
- *Bing maps* değerlerinin yaklaşık %35 ile %82 arasında,
- İki servisin birbiriyle uyum yüzdesinin yaklaşık %37 - %73 arasında değiştiği görülmektedir.

Miami Beach bölgesinin değerleri incelendiğinde;

- *Google maps* uyum değerlerinin yaklaşık %67 ile %80 arasında,
- *Bing maps* değerlerinin yaklaşık %60 ile %75 arasında,
- İki servisin birbiriyle uyum yüzdesinin yaklaşık %66 - %95 arasında değiştiği görülmektedir.

Levenshtein mesafesi kullanılarak elde edilen tüm değerlerin aritmetik ortalamaları Çizelge 3.7'de verilmektedir.

Çizelge 3.7 Standart adres formatlarının birbirleriyle olan ortalama uyum değerleri

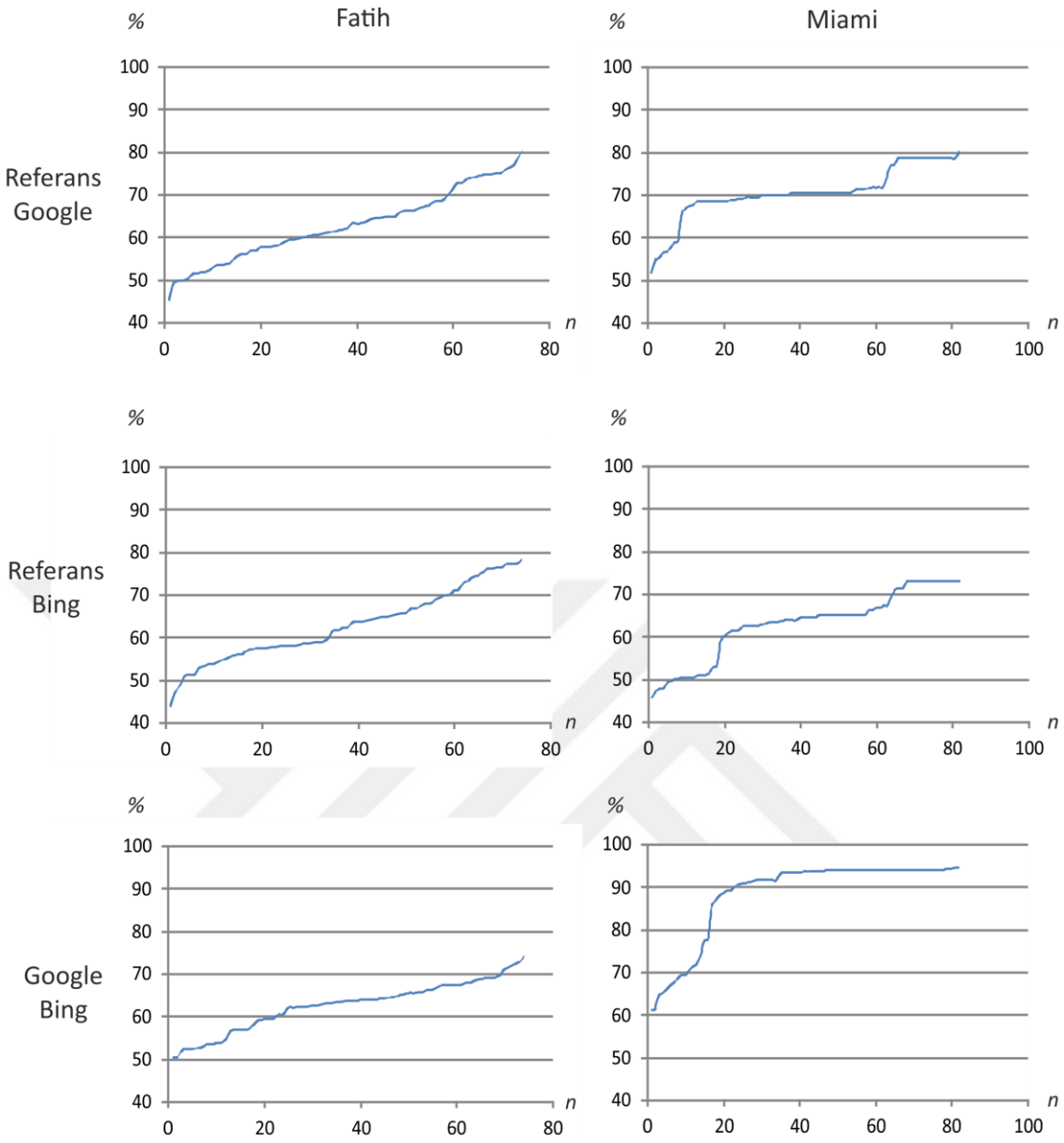
		Referans		*Google maps*
		Google maps	*Bing maps*	*Bing maps*
Uyum (%)	Fatih	65.3	64.6	61.2
	Miami Beach	72.3	66.9	93.2

Referans olarak kullanılan adres verileri Bölüm 3.1'de belirtildiği üzere *www.booking.com* sitesinden elde edilmiştir. Bu site, çeşitli nedenlerle otellerde konaklama yapmak zorunda olan ziyaretçilerin sıklıkla başvurduğu bir rehber olmasına karşın, adres bilgilerini kullanıcılarına açık ve anlaşılır bir standartta sunamamaktadır. Bunun temel nedeni, adres bilgilerinin otel işletmelerinden toplanmış olması ve verilerdeki uyumsuzlukların giderilmeden yayımlanmasıdır. Fatih'de referans adresler ve her iki servis arasındaki *Levenshtein mesafe* değerleri yaklaşık %65 civarındadır. Miami Beach'de referans veri ile *Google* ve Bi*ng maps* servislerinin ürettiği standart adresler arasındaki uyum sırasıyla %72.3 ve %66.9'dur. Bu bölgede, *Google*'ın *Bing*'e göre referans veriyle çok az da olsa daha iyi eşleştiği söylenebilir. *Google* ve *Bing maps* servislerinin birbirleriyle olan uyum değeri Fatih'de %61.2 olarak hesaplanmıştır. Bu değer, her iki servisin Fatih'de adres standardizasyonu yapmakta zorlandığını

göstermektedir. Düşük uyum değerinin elde edilmesinin temel nedeni, özellikle Türkiyede kullanılan adres yapısının karmaşık halidir. Miami Beach'de *Google* ve *Bing maps* servislerinin ürettiği standart adresler %93.2 uyum göstermektedir. Değerin 85 ile 99 değerleri arasında elde edilmesi eşleşmenin oldukça başarılı olduğunu göstermektedir. Miami Beach'de her iki servisin tam eşleşmeye yakın değerler üretmesi ABD'de kullanılan adres yapısına servislerin uyum sağlamış olduğunun önemli bir göstergesidir.

3.2.3 Tersine Coğrafi Kodlamanın Uyumu

Her iki çalışma bölgesindeki tüm binaların ağırlık merkezlerinin *WGS84* datumundaki coğrafi koordinatları Bölüm 3.1'de bahsedildiği şekilde elde edildikten sonra, enlem ve boylam değerleri kullanılarak tersine coğrafi kodlama ile *Google* ve *Bing maps* servislerinden adres bilgileri türetilmiştir. Elde edilen adreslerin referans olarak kullanılan adres verisiyle uyumu yine *Levenshtein mesafe algoritması* kullanılarak incelenmiştir (Şekil 3.13).



Şekil 3.13 Tersine coğrafi kodlamayla elde edilen adres verilerinin uyum yüzdeleri

Şekil 3.13'de, Fatih bölgesinin değerleri incelendiğinde;

- *Google maps* uyum değerlerinin yaklaşık %48 ile %80 arasında,
- *Bing maps* değerlerinin yaklaşık %45 ile %79 arasında,
- İki servisin birbiriyle uyum yüzdesinin yaklaşık %50 - %73 arasında değiştiği görülmektedir.

Miami Beach bölgesinin değerleri incelendiğinde;

- *Google maps* uyum değerlerinin yaklaşık %52 ile %80 arasında,
- *Bing maps* değerlerinin yaklaşık %48 ile %73 arasında,

İki servisin birbirleriyle uyum yüzdesinin yaklaşık %61 - %95 arasında değiştiği görülmektedir.

Fatih'de referans adres verisi ile Google ve Bing maps servislerinin ve test edilen iki servisin uyumları %62 - %63 arasındadır (Çizelge 3.8). Miami Beach'de referans veri ile Google ve Bing maps servislerinin sonuçları arasındaki uyum sırasıyla %70.5 ve %62.9'dur. Bu bölgede, Google'ın Bing'e göre referans veriyle yine az da olsa daha iyi eşleştiği söylenebilir. İki servis arasındaki tersine coğrafi kodlama uyumu ise %88.5'dir. Bu değer de iki servisin oldukça uyumlu olduğunu göstermektedir.

Çizelge 3.8 Tersine coğrafi kodlama sonucu elde edilen adres formatlarının uyum değerleri

		Referans		*Google maps*
		Google maps	*Bing maps*	*Bing maps*
Uyum (%)	Fatih	62.7	62.8	62.5
	Miami Beach	70.5	62.9	88.2

Bu çalışmada, tersine coğrafi kodlama işlemiyle elde edilen adres bilgilerinin uyumları her bir nokta için ikili şekilde teker teker incelendiğinde, her iki çalışma bölgesinde benzer sorunlara rastlanmıştır. Özellikle iki yolun kesişiminde yer alan otellerin hangi yol ile ilişkilendirileceği, çalışmada kullanılan referans noktalar binaların ağırlık merkezlerinden türetildiğinden, servisler tarafından tam olarak tespit edilememektedir (Şekil 3.14).



Şekil 3.14 Tersine coğrafi kodlama kalitesinin etkilenmesi

Eğer ağırlık merkezi otelin giriş kapısının bulunduğu yola değil de diğer yola yakın ise otel adresi yanlış yola bağlanmaktadır. Bu durum tersine coğrafi kodlamanın kalitesini etkileyen en büyük sorundur. Bu problemin çözümüne ilişkin olarak, tersine coğrafi kodlamada kullanıcı tarafından girilen noktaya en yakın bina detayı tespit edilerek, ilgili binanın adres bilgisi/bilgileri geri döndürülebilir. Yakında bir bina detayı belirlenemediği durumda adres, sadece yol bilgisine dayalı olarak geri döndürülebilir.

BÖLÜM 4

SONUÇ VE ÖNERİLER

Bu çalışmada adres bilgisini enlem ve boylam değerleriyle ifade edilen coğrafi koordinatlara dönüştüren coğrafi kodlama ve bu işlemin tersi şekilde işleten yani coğrafi koordinatları adres bilgisine dönüştüren tersine coğrafi kodlama işlemleri, *Google Maps* ve *Bing Maps* servisleri üzerinden karşılaştırmalı olarak incelenmiştir. İki ayrı çalışma bölgesinde gerçekleştirilen uygulamalar, servislerin farklı ülkelerde farklı coğrafi kodlama yeteneklerine sahip olduklarını göstermektedir. ABD'de coğrafi kodlamanın konumsal uyumu çevrimiçi çalışan Google ve Bing gibi iki küresel servis için yaklaşık eşittir. Buna karşın ülkemizde *Google Maps* servisinin *Bing Maps*'e göre daha iyi sonuçlar ortaya koyduğu söylenebilir.

Coğrafi kodlama sonucu bir koordinat çiftiyle temsil edilen herhangi bir olgu haritada nokta ile gösterilmektedir. Yüksek çözünürlüklerde bu olgular belirli bir alan kaplayan çokgenlere karşılık gelir. Bu çalışma kapsamında iki ayrı bölgede gerçekleştirilen uygulamalar detayların büyüklükleri ve coğrafi kodlamanın kalitesi arasında aşağıda bahsedilen iki sonucun elde edilmesini sağlamıştır:

- Özellikle yapılaşmanın birbirine bitişik ya da yakın ve dikey olarak geliştiği Fatih gibi bölgelerde coğrafi kodlamanın konumsal tutarlılığı zayıflamaktadır. Alansal olarak küçük ve birbirine yakın yapılardan oluşan bir bölgede 20 - 30 m'lik sapmalar yanlış yapıların kodlanması anlamına gelmektedir.

- Diğer taraftan, Miami Beach gibi yapılaşmanın yatay geliştiği, büyük alan kaplayan tesisler için coğrafi kodlama işlemi yapı çokgeninin içinde gerçekleşmektedir. Her ne kadar servislerden döndürülen coğrafi kodlanmış noktalar, çokgen detayların ağırlık merkezlerine göre büyük sapmalar gösterse de, yapıların içinde yer aldığından coğrafi kodlamanın konumsal tutarlılığının daha yüksek olduğu söylenebilir. Bu durumda 20 - 30 m'lik sapmalar tolere edilebilir.

Coğrafi kodlamanın kalitesini arttıran en önemli unsurlardan biri kodlamayı gerçekleştirmek için kullanılan servislerin standart adres altyapılarıdır. Eğer servislerin mekânsal veritabanlarında adresler bir sistematik içinde standart olarak tanımlanmışsa, kullanıcıların girdiği adres tanımlamalarıyla uyum yüksek oranda sağlanabilir. Bu noktada, veritabanının güncelliği ve numarataj sistematiğinin sağlanmış olması oldukça önemlidir. Ülkemizde kamu kurum ve kuruluşlarının kullandığı adres sistemleri bu çalışma kapsamında incelendiğinde standart adres yapısı hususunda önemli eksikliklerin olduğu görülmüştür. Adres yapılarının dünyada olduğu gibi yol ağına bağlı, güncel numarataj sistemine oturtulması ve bunun tüm ülke genelinde kullanımının sağlanması için gerekli çalışmaların yapılması gerekmektedir. Peşi sıra gelişmiş ülkelerde olduğu gibi, ülke adres sistematiğine uygun bir coğrafi kodlama servisi geliştirilebilir.

KAYNAKLAR

[1] Uluğtekin, N. ve Doğru, A.Ö., (2005). "Coğrafi Bilgi Sistemi ve Harita: Kartografya", Ege Üniversitesi CBS Sempozyumu, 27-29 Nisan 2005, İzmir.

[2] Zandbergen, P.A., (2008). "A comparison of address point, parcel and street geocoding techniques", Computers, Environment and Urban Systems, 32: 214-232.

[3] Geocoding, Wikipedia - The Free Encyclopedia, https://en.wikipedia.org/wiki/Geocoding, 09.03.2017.

[4] Corbett, J.P., (1979). Topological Principles in Cartography, Technical Paper 48, US Department of Commerce, Bureau of the Census, Washington D.C.

[5] Geographic Information Systems at Yale: Geocoding Resources, Yale University Library, http://guides.library.yale.edu/GIS/geocoding, 09.03.2017.

[6] Spatially enabling the data: What is geocoding?, National Criminal Justice Reference Service, https://www.ncjrs.gov/html/nij/mapping/ch4_3.html, 09.03.2017.

[7] Brownstein, J.S., Cassa, C., Kohane, I.S. ve Mandl, K.D., (2005). "Reverse geocoding: concerns about patient confidentiality in the display of geospatial health data", AMIA Annual Symposium Proceedings: American Medical Informatics Association.

[8] Curtis, A.J., Mills, J.W. ve Leitner, M., (2006). "Spatial confidentiality and GIS: re-engineering mortality locations from published maps about Hurricane Katrina", International Journal of Health Geographics, 5: 44.

[9] Brownstein, J.S., Cassa, C.A., Kohane, I.S. ve Mandl, K.D., (2006). "An unsupervised classification method for inferring original case locations from low-resolution disease maps", International Journal of Health Geographics, 5: 56.

[10] Kounadi, O., Lampoltshammer, T.J., Leitner, M. ve Heistracher, T., (2013). "Accuracy and privacy aspects in free online reverse geocoding services", Cartography and Geographic Information Science, 40: 140-153.

[11] Google geocoding API, Google Maps Developer, https://developers.google.com/maps/documentation/geocoding/start, 09.03.2017.

[12] MapQuest geocoding API, MapQuest Map Developer Network, https://developer.mapquest.com/documentation/geocoding-api/, 09.03.2017.

[13] Bing Maps Geocoding API, Bing Maps Developer Resources, https://www.microsoft.com/maps/choose-your-bing-maps-API.aspx, 09.03.2017.

[14] Roongpiboonsopit, D. ve Karimi, H.A., (2010). "Quality assessment of online street and rooftop geocoding services", Cartography and Geographic Information Science, 37: 301-318.

[15] McElroy, J.A., Remington, P.L., Trentham-Dietz, A., Robert, S.A. ve Newcomb, P.A., (2003). "Geocoding addresses from a large population-based study: lessons learned", Epidemiology, 14: 399-407.

[16] Cayo, M.R. ve Talbot, T.O., (2003). "Positional error in automated geocoding of residential addresses", International Journal of Health Geographics, 2: 10.

[17] Karimi, H.A., Durcik, M. ve Rasdorf, W., (2004). "Evaluation of uncertainties associated with geocoding techniques", Computer - Aided Civil and Infrastructure Engineering, 19: 170-185.

[18] Yang, D.H., Bilaver, L.M., Hayes, O. ve Goerge, R., (2004). "Improving geocoding practices: evaluation of geocoding tools", Journal of Medical Systems, 28: 361-370.

[19] Ward, M.H., Nuckols, J.R., Giglierano, J., Bonner, M.R., Wolter, C., Airola, M., Mix, W., Colt, J.S. ve Hartge, P., (2005). "Positional accuracy of two methods of geocoding", Epidemiology, 16: 542-547.

[20] Rushton, G., Armstrong, M.P., Gittler, J., Greene, B.R., Pavlik, C.E., West, M.M. ve Zimmerman, D.L., (2006). "Geocoding in cancer research: a review", American Journal of Preventive Medicine, 30: S16-S24.

[21] Goldberg, D.W., Wilson, J.P. ve Knoblock, C.A., (2007). "From text to geographic coordinates: The current state of geocoding", Urisa-Washington D.C., 19: 33.

[22] Roongpiboonsopit, D. ve Karimi, H.A., (2010). "Comparative evaluation and analysis of online geocoding services", International Journal of Geographical Information Science, 24: 1081-1100.

[23] Li, B. Zhang, X. ve Chen, Y., (2010). "Automatic construction and visualization of address models", 2010 Sixth International Conference on Natural Computation, 10-12 Ağustos 2010, Yantai, China.

[24] Zandbergen, P.A., (2011). "Influence of street reference data on geocoding quality", Geocarto International, 26: 35-47.

[25] Goldberg, D.W., Ballard, M., Boyd, J.H., Mullan, N., Garfield, C., Rosman, D., Ferrante, A.M. ve Semmens, J.B., (2013). "An evaluation framework for comparing geocoding systems", International Journal of Health Geographics, 12: 50.

[26] Sutanta, H., Chintya, N.P.P., Syarafina, Z., Nuringtyas, T.R., Roto, R., Widyaparaga, A., Mahardika, M., Kusumaadmaja, A. ve Hadi, N., (2015). "Issues and challenges in developing geocoded address in Indonesia", Proceedings of the 1st International Conference on Science and Technology, 11-13 Kasım 2015.

[27] Tian, Q., Ren, F., Hu, T., Liu, J., Li, R. ve Du, Q., (2016). "Using an Optimized Chinese Address Matching Method to Develop a Geocoding Service: A Case Study of Shenzhen, China", ISPRS International Journal of Geo-Information, 5: 65.

[28] Chow, T.E., Dede-Bamfo, N. ve Dahal, K.R., (2016). "Geographic disparity of positional errors and matching rate of residential addresses among geocoding solutions", Annals of GIS, 22: 29-42.

[29] Ratcliffe, J.H., (2004). "Geocoding crime and a first estimate of a minimum acceptable hit rate", International Journal of Geographical Information Science, 18: 61-72.

[30] Nicoara, G., (2005). "Exploring the geocoding process: a municipal case study using crime data", Yüksek Lisans Tezi, University of Texas at Dallas, Dallas, TX.

[31] Harada, Y. ve Shimada, T., (2006). "Examining the impact of the precision of address geocoding on estimated density of crime locations", Computers & Geosciences, 32: 1096-1107.

[32] Bichler, G. ve Balchak, S., (2007). "Address matching bias: Ignorance is not bliss", Policing: An International Journal of Police Strategies & Management, 30: 32-60.

[33] Hart, T.C. ve Zandbergen, P.A., (2013). "Reference data and geocoding quality: Examining completeness and positional accuracy of street geocoded crime incidents", Policing: An International Journal of Police Strategies & Management, 36: 263-294.

[34] Qin, X., Parker, S., Liu, Y., Graettinger, A.J. ve Forde, S., (2013). "Intelligent geocoding system to locate traffic crashes", Accident Analysis & Prevention, 50: 1034-1041.

[35] Mammadrahimli, A., (2015). Assessment of crash location improvements in map-based geocoding systems and subsequent benefits to geospatial crash analysis, Yüksek Lisans Tezi, the Graduate School of Clemson University, Güney Karolina.

[36] Goodchild, M.F., (2000). "GIS and transportation: status and challenges", GeoInformatica, 4: 127-139.

[37] Ersoy, K. ve Diri, B., (2011). "İnternet Sitelerinden Derlenen Adres Bilgilerinin Coğrafi Kodlanarak Harita Üzerinde Gösterimi", 13. Türkiye Harita Bilimsel ve Teknik Kurultayı, 18-22 Nisan 2011, Ankara.

[38] Yildirim, V., Yomralioglu, T., Nisanci, R. ve Inan, H., (2014). "Turkish street addressing system and geocoding challenges", Proceedings of the Institution of Civil Engineers, 167: 99.

[39] T.C. Çevre ve Şehircilik Bakanlığı Coğrafi Bilgi Sistemleri Genel Müdürlüğü, GeoKodlama Projesi, http://www.csb.gov.tr/db/cbs/webmenu/webmenu36904.pdf, 09.03.2017.

[40] Kellison, M.T., (2012). Address Points and A Master Address File: Improving Efficiency in the City of Chino, Doktora Tezi, University of Southern California, California.

[41] Coetzee, S. ve Bishop, J., (2009). "Address databases for national SDI: Comparing the novel data grid approach to data harvesting and federated databases", International Journal of Geographical Information Science, 23: 1179-1209.

[42] Coetzee, S., Cooper, A.K., Lind, M., Wells, M.M., Yurman, S.W., Wells, E., Griffiths, N. ve Nicholson, M.J., (2008). "Towards an international address standard", 10th International Conference for Spatial Data Infrastructure, St. Augustine, Trinidad & Tobago.

[43] Lind, M., (2008). "Addresses as an infrastructure component: Danish experiences and perspectives", Proceedings of the ISO Workshop on Address Standards: Considering the Issues related to an International Address Standard: Copenhagen, Denmark.

[44] Coetzee, S. ve Cooper, A.K., (2007). "What is an address in South Africa?", South African Journal of Science, 103: 449-458.

[45] Geymen, A., Yomralioglu, T. ve Baz, I., (2008). "Developing an urban information system for local governments", Proceedings of the Institution of Civil Engineers-Municipal Engineer: Published for the Institution of Civil Engineers by Thomas Telford Services, London., Vol. 161.

[46] Henry, K.A. ve Boscoe, F.P., (2008). "Estimating the accuracy of geographical imputation", International Journal of Health Geographics, 7: 3.

[47] Farvacque-Vitkovic, C., Godin, L., Leroux, H., Verdet, F. ve Chavez, R., (2005). Street addressing and the management of cities, DC: World Bank, Washington.

[48] INSPIRE Directive (2007). Directive 2007/2/EC of the EU Parliament and of the Council (14 March 2007) Establishing An Infrastructure for Spatial Information in the EU Community (INSPIRE), Official Journal of the European Union, 50.

[49] INSPIRE, Data Specifications, http://inspire.ec.europa.eu/Themes/Data-Specifications/2892, 22.03.2017.

[50] Data Specification on Addresses – Technical Guidelines, Infrastructure for Spatial Information in Europe (INSPIRE), http://inspire.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_AD_v3.0.pdf, 15.03.2017.

[51] Yıldırım, V., (2003). "Adres Bilgi Sistemi Tasarımı ve Uygulaması: Trabzon kent örneği", Yüksek Lisans Tezi, Karadeniz Teknik Üniversitesi, Trabzon.

[52] Ural H., Bediroğlu Ş., Yıldırım V., Nişancı R., Çolak H.E., Erbaş Y.S. ve Memişoğlu T., (2015). "Mekansal Adres Kayıt Sistemine Geçişte Yaşanabilecek Numarataj

ve Geokodlama Sorunları ve Çözüm Önerileri", 7. Kentsel Altyapı Sempozyumu: TMMOB, 13-14 Kasım 2015, Trabzon.

[53] T.C. Resmi Gazete, Adres ve Numaralamaya İlişkin Yönetmelik. (26245), 31.07.2006.

[54] Tapu ve Kadastro Genel Müdürlüğü, Türkiye Ulusal Coğrafi Bilgi Sistemleri (TUCBS) Projesi, https://www.tkgm.gov.tr/tr/icerik/turkiye-ulusal-cografi-bilgi-sistemi-tucbs-projesi, 16.03.2017.

[55] T.C. Çevre ve Şehircilik Bakanlığı Coğrafi Bilgi Sistemleri Genel Müdürlüğü, Türkiye Ulusal Coğrafi Bilgi Sistemi Standartlarının Belirlenmesi Projesi: TUCBS Adres Veri Teması, http://www.csb.gov.tr/db/cbs/editordosya/TUCBS_AD1_1.pdf, 01.02.2017.

[56] Nüfus ve Vatandaşlık İşleri Genel Müdürlüğü, MAKS NEDİR?, https://maks.nvi.gov.tr /, 15.03.2017.

[57] Antenucci, J.C., Brown, K., Croswell, P.L., Kevany, M.J. ve Archer, H., (1991). Geographic Information Systems: a guide to the technology, Springer US, New York: Van Nostrand Reinhold.

[58] Huxhold, W.E., (1991). An Introduction to Urban Geographic Information Systems, Oxford University Press, Oxford.

[59] Longley, P.A., Goodchild, M.F., Maguire, D.J. ve Rhind, D.W., (2001). "Geographic information system and Science", England: John Wiley & Sons, Ltd: 327-329.

[60] Zandbergen, P.A., (2009). "Geocoding quality and implications for spatial analysis", Geography Compass, 3: 647-680.

[61] Goldberg, D., (2008). A geocoding best practices guide, North American Association of Central Cancer Registries (NAACCR), Springfield, IL.

[62] The Geocoding Process, ESRI, http://desktop.arcgis.com/en/arcmap/10.3/guide-books/geocoding/the-geocoding-process.htm, 15.03.2017.

[63] Kilic, B. ve Gülgen, F., (2016). "Coğrafi Kodlama İçin Adres Verilerinin Standardizasyonuna Yönelik Bir Model Önerisi", Dünya CBS Günü 2016: Haliç Kongre Merkezi, 24-25 Kasım 2016, İstanbul.

[64] Mitchell, R., (2015). Web Scraping with Python: Collecting Data from the Modern Web, " O'Reilly Media, Inc.", United States of America.

[65] Rodrigues, F., (2010). POI Mining and Generation, Doktora Tezi, Faculty of Sciences and Technology (FCTUC) Department of Informatics Engineering - University of Coimbra, Portugal.

[66] Bilgisayar Kavramları, Levenshtein Mesafe Algoritması (Levenshtein Distance), http://bilgisayarkavramlari.sadievrenseker.com/2010/12/30/levenshtein_mesafe-algoritmasi-levenshtein-distance/, 22.03.2017.

[67] Hacar, M., (2015). Mekânsal Veri Altyapılarında Geometrik Entegrasyon, Yüksek Lisans Tezi, Yıldız Teknik Üniversitesi, İstanbul.

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı :Batuhan KILIÇ

Doğum Tarihi ve Yeri :10.09.1990, Elazığ

Yabancı Dili :İngilizce

E-posta :batuhank@yildiz.edu.tr

ÖĞRENİM DURUMU

Derece	**Alan**	**Okul/Üniversite**	**Mezuniyet Yılı**
Lisans	Geomatik Mühendisliği	Bülent Ecevit Üniversitesi	2013
Lise	Fen Bilimleri	Elazığ Ahmet Kabaklı Anadolu Öğretmen Lisesi	2008

İŞ TECRÜBESİ

Yıl	**Firma/Kurum**	**Görevi**
2014-	Yıldız Teknik Üniversitesi/ Fen Bilimleri Enstitüsü	Araştırma Görevlisi

YAYINLARI

Bildiri

[1] Karabulut, M. F., Kilic, B., Yilmaz, Y., ve Gülgen, F. (2016). “Proposal of a Model for Tracking the Amount and the Yield of Agricultural Product using GIS”. Geomed 2016: The 4th International Geography Symposium, 23-26 Mayıs 2016, Antalya.

[2] Yilmaz, Y., Saracoglu, A., Karabulut, M. F., Kilic, B., ve Dogan, A. H. (2016). “Performing Local Accuracy Functions Using Web-Based Online GPS Process Services: Application To Spain”. ISCAS - Selçuk International Scientific Conference On Applied Sciences (Poster), 27-30 Eylül 2016, Antalya.

[3] Kilic, B. ve Gülgen, F. (2016). "Coğrafi Kodlama İçin Adres Verilerinin Standardizasyonuna Yönelik Bir Model Önerisi", Dünya CBS Günü 2016: Haliç Kongre Merkezi, 24-25 Kasım 2016, İstanbul.

[4] Kilic, B. ve Gülgen, F. (2017). “Coğrafi Kodlama İçin Adres Standardizasyonu”. 16. Türkiye Harita Bilimsel ve Teknik Kurultayı: Beytepe Kültür ve Kongre Merkezi, 3-6 Mayıs 2017, Ankara.

Kitap Bölümü

[1] Karabulut, M. F., Kilic, B., Yilmaz, Y., ve Gülgen, F. (2017). “A Model for Tracking the Amount and the Yield of Agricultural Product Using GIS”. Contemporary Studies in Environment and Tourism, Efe R., Öztürk M., Eds., Cambridge Scholars Publishing, Newcastle, pp.499-506.

Proje

[1] İstanbul Lojistik Ana Planı Projesi (İSTLAP), İstanbul Büyükşehir Belediyesi Danışmanlık, 2015 – Devam ediyor.

ÖDÜLLERİ

[1] Yilmaz, Y., Saracoglu, A., Karabulut, M. F., Kilic, B., ve Dogan, A. H. (2016). “Performing Local Accuracy Functions Using Web-Based Online GPS Process Services: Application To Spain”. ISCAS - Selçuk International Scientific Conference On Applied Sciences (Best Poster Award), 27-30 Eylül 2016, Antalya.