

**T.C.
MİMAR SİNAN GÜZEL SANATLAR ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

UZAMSAL REGRESYON ANALİZİ

DOKTORA TEZİ

Özlem DENİZ BAŞAR

İstatistik Anabilim Dalı

İstatistik Programı

Tez Danışmanı: Prof. Dr. Aydın ERAR


TEMMUZ 2009


Özlem DENİZ BAŞAR tarafından hazırlanan UZAMSAL REGRESYON ANALİZİ adlı
bu tezin DOK.TÖ.RA... tezi olarak uygun olduğunu onaylarım.


Prof. Dr. M. Aydın ERAR


Tez Yöneticisi


Bu çalışma, jürimiz tarafından İSTATİSTİK..... Anabilim Dalında
Doktora..... tezi olarak kabul edilmiştir.

Başkan : Prof. Dr. M. Aydın ERAR (Dahizmen) 

Üye : Prof. Dr. Münevver Turan 

Üye : Prof. Dr. Fazıl Güler 

Üye : Prof. Dr. Gözay Kuroğlu 

Üye : Prof. Dr. Sahamet Rabbil 

Bu tez, Mimar Sinan Güzel Sanatlar Üniversitesi Fen Bilimleri Enstitüsü tez yazım kurallarına uygundur.

ÖZET

Bu çalışmada amaç, uzamsal regresyon analizinde kullanılacak ağırlık matrislerini düzenlemek üzere farklı uzamsal ağırlık fonksiyonlarının oluşturulmasıdır. Bu amaç doğrultusunda, incelenen bölgelerin coğrafi konumlarının yanı sıra farklı etkiler de fonksiyonlara dahil edilerek uzamsal ağırlık matrisleri düzenlenmiş ve uzamsal modeller oluşturulmuştur.

Birinci bölüm giriş niteliğinde olup tezi oluşturan konu başlıklarından kısaca bahsedilmiştir.

İkinci bölümde; genel olarak uzamsal verilerin yapısı ele alınmıştır. İlk adımda uzamsal verilerin gerekliliği incelenmiş ve uzamsal veri çeşitleri açıklanmıştır. Daha sonra, bu veriler için örnekleme yöntemlerinin kullanımları ayrıntıları ile verilmiştir. Bunların yanı sıra uzamsal otokorelasyon fonksiyonları ve uzamsal otokorelasyon sorununun belirlenmesine kullanılan testler de incelenmiştir.

Üçüncü bölümde; uzamsal analizde kullanılacak modeller ayrıntılı bir biçimde ele alınmıştır. Modeller, ilişkili ve ilişkisiz hataların bulunması durumunda kullanılan doğrusal modeller olmak üzere iki ayrı grupta incelenmiştir. Ayrıca bu modellerin oluşturulmasında kullanılacak ağırlık matrislerinin elde edilme yöntemlerine de yer verilmiştir.

Dördüncü bölümde; uzamsal regresyon analizinin uygulanmasında izlenecek süreç ve kullanılacak istatistiksel testler açıklanmıştır. Ağırlık fonksiyonlarını oluşturmak için dikkate alınması gereken kurallar bu bölümde gerekçeleri ile verilmiştir.

Beşinci bölümde; oluşturulan farklı ağırlık fonksiyonları yardımıyla ağırlık matrisleri düzenlenmiş ve uzamsal regresyon analizi uygulanmıştır. Analiz sonucunda elde edilen modellerin geçerlilikleri sınanmış ve modeller yardımıyla tahminler yapılmıştır.

Son bölüm olan altıncı bölümde ise çalışmanın sonuçları tartışılmış ve ileriye yönelik yapılabilecekler üzerinde durulmuştur.

SPATIAL REGRESSION ANALYSIS

SUMMARY

The aim of this study is to form different spatial weight functions in order to create spatial weight matrix which is used for spatial regression analysis. For this purpose, other impacts are included along with the geographical coordinates of the locations.

The first chapter is an introduction chapter which covers the main titles of the thesis.

In the second chapter, structures of spatial data are discussed generally. First of all necessity and types of spatial data are studied. Then application of sampling methods to this data are explained in detail. Apart from these spatial autocorrelation functions and tests for determining of spatial autocorrelation problem are also explained.

In the third chapter, models that can be used in spatial analysis are examined in details. These models are studied in two groups: models with correlated errors and uncorrelated errors. Spatial weight matrix which is used to build these models are also explained.

In the fourth chapter, the process of spatial regression analysis and the statistical tests are examined. Rules for building spatial weight functions are also given in this chapter.

In the fifth chapter, spatial weight matrix are built by different spatial weight functions and spatial regression analysis applied. Validity of these models are checked and they are used for prediction.

In the sixth chapter which is the final chapter, the results of the study are discussed and recommendations for future studies are given.

ÖNSÖZ

Çok uzun süren ve oldukça yorucu olan bir sürecin ardından sonucuna ulaşabildiğim tezim için teşekkür etmem gereken o kadar çok kişi var ki...

En başta, gece gündüz demeden, bana destek olmak için en az benim kadar çalışan ve kıymetli vaktini hiçbir gün geri çevirmeden ayıran, bana bilimsel anlamda kattıklarının yanı sıra sabretmeyi de öğreten danışmanım Prof.Dr. Aydın ERAR'a,

Uzun yıllardır birlikte çalıştığım, manevi desteğini her zaman arkamda hissettiğim, bu sürecin hiçbir aşamasında beni yalnız bırakmayan Prof.Dr. Münevver TURANLI'ya,

Kararlarım ile ilgili en önemli sayfalarda imzası olan ve tanımaktan büyük onur duyduğum Prof. Dr. Gülay KIROĞLU'na,

Hayata gözlerimi açtığım ilk günden beri, beni her konuda destekleyen, katlanan, cesaretlendiren, umutlandıran, sevindiren kısacası beni ben yapan canım aileme,

Başından beri yanımda olan, her adımda desteğini esirgemeyen, kendisinden çaldığım zamanlar için bana sabır gösteren biricik eşime,

Üzerimde emeği olan ve adlarını buraya sığdıramadığım herkese sonsuz teşekkürü bir borç bilirim.

Temmuz, 2009

Özlem DENİZ BAŞAR

İÇİNDEKİLER DİZİNİ

ÖZET.....	i
SUMMARY	ii
ÖNSÖZ.....	iii
İÇİNDEKİLER DİZİNİ	iv
TABLO DİZİNİ.....	vi
ŞEKİL DİZİNİ.....	ix
1. GİRİŞ.....	1
2. UZAMSAL VERİLER.....	4
2.1. UZAMSAL VERİLERİN GEREKLİLİĞİ	5
2.2. UZAMSAL VERİ ÇEŞİTLERİ	7
2.2.1. Jeostatistik Veriler	7
2.2.2. Kafes Veriler	8
2.2.3. Nokta Düzenleri.....	9
2.3. UZAMSAL ORTALAMA	11
2.4. UZAMSAL OTOKORELASYON FONKSİYONLARI	12
2.4.1. Moran Katsayısı	14
2.4.2. Geary Oranı	15
2.4.3. Zaman Serilerinin Otokorelasyon Fonksiyonu	16
2.4.4. Uzamsal Otokorelasyon Fonksiyonu	16
2.5. UZAMSAL ÖRNEKLEME.....	17
2.5.1. Uzamsal Kitle ve Örneklemeye	17
2.5.2. Uzamsal Örneklemeye Yöntemleri.....	19
2.5.2.1. Tekdüze (Uniform) Rasgele Örneklemeye.....	19
2.5.2.2. Tabakalı Rasgele Örneklemeye	20
2.5.2.3. Sistematik Örneklemeye	20
3. UZAMSAL VERİLERDEKİ DEĞİŞİMLERİN MODELLENMESİ.....	22
3.1. İLİŞKİSİZ HATALAR İÇİN DOĞRUSAL MODELLER	24
3.1.1. Sürekli Değişkenler için Tanımlayıcı Modeller	24
3.1.1.1. Bağımsız Hatalarla Trend Yüzey Modelleri.....	24
3.1.1.2. Semi-Variogram ve Kovaryans Modelleri.....	24
3.1.2. Kesikli Değişkenler için Tanımlayıcı Modeller	25
3.1.3. Klasik Açıklayıcı Modeller.....	26
3.1.4. Genelleştirilmiş Doğrusal Modeller	32
3.1.5. Coğrafi Olarak Ağırlıklandırılmış Regresyon Yöntemi.....	33
3.1.6. Coğrafi Olarak Ağırlıklandırılmış Regresyon Yönteminde Uzamsal Ağırlık Fonksiyonunun Seçimi	36

3.1.7. Karma Coğrafi Olarak Ağırlıklandırılmış Regresyon Modeli	37
3.1.8. Coğrafi Olarak Ağırlıklandırılmış Regresyon Modelinde Aykırı Değerler	40
3.1.9. Uzamsal Yapıda Oluşturulan Modeller	42
3.1.9.1. Komşuluk Düzeltmeli Modeller – Papadakis Analizi	42
3.1.9.2. Birinci Farklar Modeli	43
3.2. İLİŞKİLİ HATALAR İÇİN DOĞRUSAL MODELLER.....	44
3.2.1. Geniş Alanlı Uzamsal Değişimler için Modeller.....	46
3.2.2. Dar Alanlı Uzamsal Değişimler için Modeller	47
3.2.2.1. Yüzey Veriler için Modeller	47
3.2.2.2. Sürekli Değerli Alan Verileri için Modeller	48
3.2.2.3. Kesikli Değerli Alan Verileri için Modeller.....	52
3.2.3. Karma Modeller	53
3.2.4. Uzamsal Genişleme Modeli.....	54
3.2.5. Uzamsal Ayarlanabilir Filtreleme Modeli.....	54
3.2.6. Çok Düzeyli Modelleme.....	55
3.2.7. Rasgele Katsayılar Modelleri.....	55
4. UZAMSAL REGRESYON ANALİZİ.....	57
4.1. UZAMSAL BAĞIMLILIĞIN SAYISAL OLARAK GÖSTERİLMESİ.....	57
4.1.1. Arazi: İki Boyutlu Sürekli Uzaydan Veriler.....	58
4.1.2. Nesnelere: İki Boyutlu Kesikli Uzaydan Veriler	58
4.2. UZAMSAL AĞIRLIK FONKSİYONLARI	59
4.3. UZAMSAL REGRESYON ANALİZİNDE KULLANILAN MODELLER.....	63
4.3.1. Uzamsal Otoregresif Modeller	63
4.3.1.1. Birinci Dereceden Uzamsal Otoregresif Model	64
4.3.1.2. Karma Otoregresif Model	65
4.3.1.3. Uzamsal Hata Modeli	66
4.3.2. Tanımlayıcı İstatistikler ve Uzamsal Etkiler.....	67
5. UZAMSAL REGRESYON ANALİZİNDE AĞIRLIK MATRİSLERİNE İLİŞKİN UYGULAMA.....	70
5.1. DEĞİŞKENLERDE EKSİK GÖZLEMLERİN OLMASI DURUMUNDA VERİ ANALİZLERİ	71
5.1.1. Klasik Regresyon Analizinin Uygulanması.....	74
5.1.2. Farklı Ağırlık Matrisleri Kullanılarak Uzamsal Gecikme Modelinin Elde Edilmesi	76
5.1.2.1. En Yakın Komşuluk Matrisi ile Model.....	77
5.1.2.2. Eşik Uzaklık Matrisi ile Model	78
5.1.2.3. Oluşturulan Modellerin Birbirleri ile Karşılaştırılması.....	78
5.2. DEĞİŞKENLERDE EKSİK GÖZLEMLERİN GİDERİLDİĞİ DURUMUNDA VERİ ANALİZLERİ	79
5.2.1. Klasik Regresyon Analizinin Uygulanması.....	82
5.2.2. Farklı Ağırlık Matrisleri Kullanılarak Uzamsal Gecikme Modelinin Elde Edilmesi	84
5.2.2.1. En Yakın Komşuluk Matrisi ile Model.....	84
5.2.2.2. Eşik Uzaklık Matrisi ile Model	84
5.2.2.3. Oluşturulan Modellerin Birbirleri ile Karşılaştırılması.....	85
5.3. AĞIRLIK FONKSİYONLARI YARDIMIYLA AĞIRLIK MATRİSLERİNİN OLUŞTURULMASI	85
5.3.1. w_1 Ağırlık Fonksiyonunun Kullanılması ile Oluşturulan Modeller.....	87
5.3.2. w_2 Ağırlık Fonksiyonunun Kullanılması ile Oluşturulan Modeller.....	91
5.3.3. w_3 Ağırlık Fonksiyonunun Kullanılması ile Oluşturulan Modeller.....	94
5.3.4. Farklı Ağırlık Fonksiyonunun Kullanılması ile Oluşturulan Modellerin Birbirleri ile Karşılaştırılması.....	98
5.4. MODELİN GEÇERLİLİĞİNİN İNCELENMESİ VE TAHMİNLER	100

6. SONUÇ VE ÖNERİLER.....	104
KAYNAKLAR.....	106
EKLER	112
EK – 1. İnternet Kullanımına Göre Ülkelerin Kümelenmesi	113
EK – 2. En Yakın Komşuluk Matrisi	114
EK – 3. Eşik Uzaklık Matrisi	115
EK – 4. w_1 Ağırlık Matrisi.....	116
EK – 5. w_2 Ağırlık Matrisi.....	117
EK – 6. w_3 Ağırlık Matrisi.....	118
EK – 7. R Programı Kodları	119
ÖZGEÇMİŞ	121

TABLO DİZİNİ

	Sayfa No
Tablo 2.1 Uzamsal Verilerin Sınıflandırılması	10
Tablo 3.1. Genel ve Bölgesel İstatistikler Arasındaki Farklar.....	23
Tablo 3.2. Klasik Regresyon Modeli ile GWR Modelin Karşılaştırılması	34
Tablo 3.3. Kullanılan Semi-Variogram Örnekleri	48
Tablo 4.1. Regresyon Modelleri için Uzamsal Bağımlılık Testleri	68
Tablo 5.1. Uygulamada Kullanılan Değişkenler.....	70
Tablo 5.2. Eksik Gözlemler İçeren Değişkenler.....	71
Tablo 5.3. Tüm Verilerle Klasik Regresyon Analizi Sonuçları.....	75
Tablo 5.4. Varsayımlardan Sapmalara İlişkin Ölçümler	75
Tablo 5.5. Uzamsal Bağımlılık İstatistikleri	76
Tablo 5.6. Farklı Modellere İlişkin Sonuçlar.....	77
Tablo 5.7. Değişkenler ve Normale Yaklaştırmak için Yapılan Dönüşümler (xp)	79
Tablo 5.8. Eksik Gözlemlerin Giderildiği Klasik Regresyon Analizi Sonuçları	82
Tablo 5.9. Eksik Gözlem Olmasızın Varsayımlardan Sapmaların İncelenmesi.....	83
Tablo 5.10. Eksik Gözlem Olmaması Durumunda Uzamsal Bağımlılık İstatistikleri	83
Tablo 5.11. Veriler ve Kullanılan Fonksiyonların İsimleri	86
Tablo 5.12. Uzamsal Bağımlılık İstatistikleri	88
Tablo 5.13. w1 Ağırlık Matrisli Uzamsal Gecikme Modeline İlişkin Sonuçlar	89
Tablo 5.14. Etkin Gözlemler.....	90
Tablo 5.15. Uzamsal Bağımlılık İstatistikleri	91
Tablo 5.16. w2 Ağırlık Matrisli Uzamsal Gecikme Modeline İlişkin Sonuçlar	92
Tablo 5.17. Etkin Gözlemler.....	93
Tablo 5.18. Uzamsal Bağımlılık İstatistikleri	95
Tablo 5.19. w3 Ağırlık Matrisli Uzamsal Gecikme Modeline İlişkin Sonuçlar	96
Tablo 5. 20. Etkin Gözlemler.....	97
Tablo 5.21. Farklı Ağırlık Matrisleri ile Oluşturulan Modellere İlişkin Sonuçlar	98
Tablo 5.22. Almanya'nın Çıkarıldığı Veri Setinin Uzamsal Bağımlılık İstatistikleri.....	99
Tablo 5.23. Kukla Değişken Eklenmesi ile Elde Edilen Uzamsal Bağımlılık İstatistikleri.....	100

Tablo 5.24. Anlamsız Bulunan Değişkenlerin Çıkartılması ile Hesaplanan Uzamsal Bağımlılık İstatistikleri	100
Tablo 5.25. 28-75-103 Gözlemlili Verilerle Geçerliliğin İncelenmesi için Hesaplanan Değerler	102
Tablo 5.26. Farklı Ağırlık Matrisleri ile Türkiye için Yapılan Tahminler.....	103

ŞEKİL DİZİNİ

	Sayfa No
Şekil 2.1. USA National Climatic Data Center Raporları Sonucunda Hava İstasyonları.....	8
Şekil 2.2. Çocukların Kurşun Zehirlenmelerine İlişkin Oranlar.....	8
Şekil 2.3. 17/20 Nisan 2003 Tarihinde Gözlenen Şimşeklerin Yeri.....	9
Şekil 2.4. Ortalama Hesaplamasında Kullanılan Bölgelere İlişkin Bilgiler ve Ortalama Değerleri	11
Şekil 2.5. Uzamsal Ortalamalara İlişkin Frekans Dağılımı	12
Şekil 2.6. Buffalo ve Alt Bölgelerine İlişkin Suç Oranları Yardımı ile Otokorelasyon İncelemesi.....	13
Şekil 2.7. Matrisin Oluşturulması.....	14
Şekil 2.8. Üç Bölge Yardımıyla Otokorelasyonun Gösterilmesi	15
Şekil 2.9. Örneklem Planlarına İlişkin Şablonlar	18
Şekil 2.10. Örneklemde Çerçevelerin Oluşturulması	19
Şekil 3.1. Verilerin Toplanmasına İlişkin Sonuçlar.....	23
Şekil 3.2. Birinci Dereceden Komşulukların Kodlanması	26
Şekil 3.3. Birimlerin Birbirlerine göre Komşuluklarının Gösterilmesi	50
Şekil 4.1. Otuz Derecelik Bölmelerle Ayrılmış Uzaklık Bantlarıyla Gösterilmiş Komşular.....	58
Şekil 4.2. Delaunay Üçgenleştirme Yönteminin Gösterilmesi	59
Şekil 4.3. Farklı Komşuluk Yapıları ve Komşuluk Matrislerinin Gösterimi	60
Şekil 4.4. Uzamsal Regresyon Analizi Süreci.....	69
Şekil 5.1. Ülkelerin Merkez Noktalarının ve İnternet Kullanım Oranları Büyüklüklerinin Koordinat Sisteminde Temsili Gösterimleri	72
Şekil 5.2. Değişkenlere İlişkin Histogramlar	73
Şekil 5.3. İnternet Kullanımı, Telefon Hatları ve Eğitim Harcamaları Değişkenlerinin Bir Arada Gösterilmesi	74
Şekil 5.4. Kullanılan Değişkenlerin Dağılımlarının Gösterimi.....	80
Şekil 5.5. Bağımlı Değişken ile Dönüşüm Uygulanmış Değişkenlerin Düzleştirme Eğrileri.....	81
Şekil 5.6. Ülkelerin Başkentleri ve Merkez Noktalarının Birbirlerine Uzaklıkları	86
Şekil 5.7. w1 Ağırlığı için Elde Edilen Moran I Saçılım Grafiği.....	90
Şekil 5.8. w2 Ağırlığı için Elde Edilen Moran I Saçılım Grafiği.....	93

Şekil 5.9. Optimum Delta Değerinin Belirlenmesi için Kırılma Noktasının Tespit Edilmesi	95
Şekil 5.10. w3 Ağırlığı için Elde Edilen Moran I Saçılım Grafiği	97
Şekil 5.11. (a) w1 Ağırlık Matrisi, (b) w2 Ağırlık Matrisi, (c) w3 Ağırlık Matrisi ile Oluşturulmuş Modelde ($y - gecikmis\hat{y}$) Çizimleri (d) a-b-c Şekillerinin Üst Üste Gösterilmesi	101

1. GİRİŞ

Uzamsal regresyon analizi özellikle son yirmi yılda istatistiksel analizlerde sık kullanılan yöntemlerden biri olmuştur. Konumun önemli olduğu durumlarda, değişimi açıklamak ve öngörülerde kullanılmak üzere kullanılan klasik modellerin yerini, coğrafi anlamdaki bilgileri içeren ve konumların gözlemler üzerindeki etkilerini de dikkate alarak oluşturulan uzamsal modeller almıştır.

Günümüzde uzamsal analizler birçok alanda uygulanmaktadır. Bu analizler, özellikle son dönemde tıpta, kanserle ilgili yapılan çalışmalarda en sık kullanılan yöntem olarak karşımıza çıkmaktadır. Bunun yanı sıra jeoloji, toprak bilimleri, imaj analizleri, epidemiyoloji, tarım bilimi, ekoloji, ormancılık, astronomi ve atmosferik bilimlerde kullanımı her geçen gün artmaktadır.

Uzamsal regresyon analizinin uygulanmasına karar verdikten sonra en önemli nokta, kullanılacak ağırlık matrisinin oluşturulmasıdır. Literatürde ağırlık matrislerinin oluşturulmasında en sık kullanılan yöntem komşuluk esasına dayanan, sınırdaş bölgelerin bir, diğerlerinin sıfır olarak kodlandığı matrislerin kullanılmasıdır. Ancak; analizlerde sadece sınırdaş olmanın sınama için yeterli bir kriter olmamaya başlaması ve farklı etkilerin de modele ağırlık olarak dahil edilmek istenmesi nedeniyle uzamsal süreçte yeni arayışlara gidilmiştir. Bilgisayar programlama teknolojisinin de her geçen gün gelişmesi ve karmaşık istatistiksel analizlerin uygulanmasını kolaylaştırması sayesinde, uzamsal ağırlık matrislerinin oluşturulmasında birçok etkenin barındırılabilirdiği fonksiyonlar kullanılmaya başlanmıştır. Uzamsal süreçte kullanılan bu fonksiyonlar “uzamsal ağırlık fonksiyonları” adını almışlardır.

Uzamsal ağırlık fonksiyonlarının doğru oluşturulması, sadece incelenen bölgelerin birbirlerine olan coğrafi uzaklıklarının değil, farklı etkilerin de dahil edilebilmesi açısından, uzamsal ağırlık matrislerinin düzenlenmesinde büyük önem taşımaktadır. Uzamsal ağırlıklar oluşturulurken dikkat edilmesi gereken en önemli nokta, incelenen bölgeler arasındaki mesafenin artmasına bağlı olarak uzamsal ağırlık değerinin daha küçük çıkmasının gerekliliğidir. Diğer bir ifade ile bölgeler arasındaki

mesafe arttıkça etkileşimin azalması, dolayısıyla ağırlık değerinin küçülmesi beklenmektedir. Bu nedenden dolayı oluşturulacak her fonksiyona, uzamsal ağırlık fonksiyonuymuş gibi yaklaşmak yanlış olacaktır.

Bu çalışmada amaç, farklı etkilerin barındırıldığı çeşitli uzamsal ağırlık fonksiyonları yardımıyla uzamsal ağırlık matrislerinin oluşturulması, bu matrisleri kullanarak uzamsal regresyon analizi sürecinin uygulanmasıdır. Bu amaç doğrultusunda çalışmanın ikinci bölümünde uzamsal veriler hakkında genel bilgiler verilmiştir. Düşünülmesi gereken ilk nokta, uzamsal verileri kullanmanın gerekli olup olmadığıdır. Eğer uzamsal verilerin kullanımı gerekli görülürse, hangi çeşit verinin kullanılması gerektiğine karar verilmelidir. Bölgelere ilişkin bu verilerin kullanılması durumunda, yakın konumlanmış bölgelerin etkileşim içerisinde olması kaçınılmazdır. Bu nedenden dolayı uzamsal otokorelasyonun yönünün ve büyüklüğünün belirlenmesi açısından en sık kullanılan testler, bu bölümde ayrıntıları ile açıklanmıştır. Bunların yanı sıra, bir analizin başlangıcında izlenmesi gereken en önemli süreç olan kitlenin tanımlanması ve gerekli örnekleme yönteminin belirlenerek verilerin elde edilmesi, uzamsal verilerde kullanılan teknikleri ile ikinci bölümde gösterilmiştir.

Üçüncü bölümde, uzamsal verilerdeki değişimlerin gösterilebilmesi için kullanılan modeller açıklanmıştır. Modeller, ilişkisiz ve ilişkili hataların olması durumunda kullanılan doğrusal modeller olmak üzere iki grupta incelenmiştir. İlişkisiz hataların olması durumunda kullanılan tanımlayıcı modeller, klasik açıklayıcı modeller, genelleştirilmiş doğrusal modeller, coğrafi olarak ağırlıklandırılmış regresyon modelleri ve uzamsal yapıda oluşturulan modeller ayrıntıları ile gösterilmiştir. İlişkili hataların söz konusu olduğu durumlarda kullanılan geniş alanlı ve dar alanlı uzamsal değişim modelleri, karma modeller, uzamsal genişleme modelleri, uzamsal ayarlanabilir filtreleme modelleri, çok düzeyli modeller ve rasgele katsayılar modelleri de bu bölümde gösterilmiş ve kestirilmelerine ilişkin bilgiler gerekçeleri ile sunulmuştur.

Uzamsal regresyon analizi sürecinin incelendiği dördüncü bölümde ise, ilk olarak uzamsal modellerin kullanılmasının nedeni olan uzamsal bağımlılık, sayısal olarak gösterilmiştir. Bu çalışmanın asıl konusunu oluşturan uzamsal ağırlık fonksiyonlarına ilişkin bilgiler yine bu bölümde verilmiştir. Uzamsal regresyon analizinde kullanılacak ve uzamsal ağırlık matrisleri ile oluşturulan uzamsal modeller bu bölümde ayrıntıları ile açıklanmıştır.

Beşinci bölümde sınırdaşlık esas alınan ağırlık matrislerinin uygulandığı modellerin yanı sıra uzamsal ağırlık fonksiyonlarının da dahil edildiği uzamsal modeller oluşturulmuştur. Modellerin elde edilmesinin ardından modellere ilişkin geçerlilik analizleri yapılmış ve farklı ağırlık fonksiyonları ile oluşturulmuş modeller için Türkiye'nin coğrafi konumunun esas alındığı tahminler yapılmıştır. Dolayısıyla tahminlerde Türkiye'deki değişimlerin komşularını, incelenen kriterler ile birlikte ne kadar etkilediği belirlenmiştir.

Son bölüm olan altıncı bölümde ise çalışmanın sonuçları tartışılmış ve ileriye yönelik olarak neler yapılabileceğinin üzerinde durulmuştur.

2. UZAMSAL VERİLER

Coğrafi sistemlere nicel yaklaşımlar uzamsal yöntemlerin uygulanması ile sağlanabilmektedir. Çok uzun bir geçmişe sahip olmayan yöntemlerin literatürde yer alması 1973 yılında Cliff ve Ord tarafından yapılan çalışmalar ile başlamıştır. Odland (1988) tarafından gerçekleştirilen “uzamsal otokorelasyon” çalışmaları, Boots ve Getis (1988) tarafından geliştirilen “nokta düzen analizleri”, Anselin (1988) ve Ripley (1988) tarafından yapılan “uzamsal ekonometri ve uzamsal istatistik” çalışmaları, uzamsal analizlerin geliştirilmesinde önemli adımlar olarak gösterilmiştir. Gelişen bilgisayar teknolojileri ile günümüzde daha fazla kullanım alanı bulunan uzamsal veriler ve analizlerinin literatürdeki örnekleri her geçen gün artmaktadır.

Bilindiği gibi gözlenen tüm olaylar için uzay ve zaman koordinatları var olmaktadır. Çünkü tüm olaylar herhangi bir yerde herhangi bir zamanda meydana gelmektedirler. Deneysel bilimlerde deneylerin gerçekleştiği uzamsal koordinatların tam değerleri veri kümesine dahil edilmemektedir. Bu tarz verilerin kendilerini, açıklayıcı değişkenler içinde belli ettiği ve analizlerin sonucuna etki edecek nitelikte olmadığı düşünülür. Ancak sosyal ve çevresel bilimlerde gözlemler, deneysel bilimlerde olduğu gibi elde edilemezler. Araştırılan konuya ait veriler olduğu gibi alınır ve araştırmacı açıklayıcı değişkenler için deneysel değerler elde edilemez. Koordinat değerlerinin dahil edilmemesi durumunda bir sonraki adım olan eldeki verilerle model elde etme aşamasında, açıklayıcı değişkenlerin tasarım matrisinin oluşturulmasında değişkenlerin ölçüm değerleri ile düzeyleri birbirine karışacaktır. Bu durum ilerleyen adımlarda model hatalarının, hata ölçümlerinin etkilerinin ve örnekleme hatalarının ortaya çıkmasına neden olur.

Gözlemsel bilimlerin birçok alanında, veri kümesindeki her bir birey için yer ve zaman değerlerinin kaydedilmesi oldukça önemlidir. Bu önem aşağıdaki nedenlerle açıklanabilir:

1. Sosyal bilimlerde farklı türde mekanlar ve uzaylar içeren süreçler üzerinde çalışır (sosyal ve ekonomik süreçlerin mekan ve uzayların yapılarına göre farklı durumlar göstermesi gibi).

2. Olayın meydana geldiği yerin kaydedilmesi demek, bu verinin farklı veri kümelerindeki verilerle birleştirilebilir hale getirilebilmesi anlamına gelmektedir (adres tabanlı sağlık verileri ile nüfus sayımından elde edilen sosyoekonomik verilerin birleştirilmesi gibi).

Bu açıklamalar doğrultusunda uzamsal veri analizinin bilimsel araştırmaları destekleyen bir rolü olduğu söylenebilir. Ayrıca coğrafi uzaydaki gözlemlerin bağımlı olması nedeniyle birçok genel problemin çözümünde önemli bir role sahiptir. Coğrafi olarak birbirine yakın verilerin uzak olan verilere göre daha fazla benzerlik göstermesi bağımlılık yapısını ortaya çıkarır. Coğrafi uzayın en büyük ve en belirgin özelliği ve belki de problemi budur. Bu tarz verilere klasik istatistik teorisinin uygulanması da sorunlara neden olur (Haining, 2003, syf.15-16).

Uzamsal veriler jeoloji, toprak bilimleri, imaj analizleri, epidemiyoloji, tarım bilimi, ekoloji, ormancılık, astronomi, atmosferik bilimler veya farklı uzamsal noktalardan elde edilen bilgileri içeren verilerdir. Daha önce de belirtildiği gibi, bu tarz verilerde farklı bölgelerdeki ölçümler arasında bağımlılık görülebilir (Cressie, 1993, syf.7).

Uzamsal istatistik ise uzayda bazı süreçlerde elde edilen gözlenebilir veriler yardımıyla bulunan sonuçları içerir. Ayrıca bu süreçlerin davranışlarını tanıtmaya veya açıklamaya yardımcı yöntemler ve bunların diğer uzamsal yaklaşımlarla olası ilişkilerini ortaya çıkaran sistemin tamamı da uzamsal istatistikleri oluşturur.

Coğrafi yer ve şekillere ilişkin yapılan çalışmalar ve bunların arasındaki ilişkilerin incelenmesine de “uzamsal analiz” adı verilir (Albert, 2000, syf.5).

2.1. UZAMSAL VERİLERİN GEREKLİLİĞİ

Uzamsal verilerle ve istatistiklerle çalışılmasının en büyük nedeni olarak her zaman “ne kadar” sorusuyla değil, zaman zaman “ne kadarı nerede” sorusuna cevap bulunmasının gerekliliğidir. Birçok veri kümesinde gözlemlerin hangi coğrafi konumdan elde edildiğine ilişkin bilgiler bulmak mümkündür.

Uzamsal istatistiklerle ilgili çalışan birçok bilim adamının ilgilendiği en önemli nokta, uzaydaki gözlemlerden oluşan uzamsal verilerin ölçümleri arasındaki bağımlılık anlamına gelen otokorelasyonlarıdır. Bu bağımlılık nedeniyle “klasik” istatistiksel modelleme ile “uzamsal” modelleme birbirlerinden ayrılmaktadır. Uzamsal verilere klasik istatistiksel yöntemler uygulanması, sonuçlarda yeterlilik ve tutarlılık varsayımlarının bozulmasına neden olur (Schabernberger, 2005, syf.1).

Özel bir durum olarak dikey kesit veriler: ilişkili gözlemler içerdiğinden uzamsal verilerle karıştırılabilirler; ancak aralarında önemli farklılıkların olduğu bilinir.

Dikey kesit çalışmada oluşturulan model;

$$\mathbf{Y}_i = \mathbf{X}_i\boldsymbol{\beta} + e_i \quad (2.1)$$

şeklindedir ve $e_i \sim N(0, V_i(\theta_i))$ olur. $V_i(\theta_i)$; aynı hataya ait gözlemler arasındaki varyans ve kovaryansları gösterir. Burada,

1. β ve θ_i parametreleri kestirilmelidir.
2. Bu parametrelere ait hipotez testleri gerçekleştirilmelidir.
3. $E[Y_i] = \mathbf{X}_i\boldsymbol{\beta}$ ortalama vektörleri kestirilmelidir.
4. Model yardımıyla t anındaki istenen değer kestirilmelidir.

Bu açıklamanın uzamsal bir süreçle karşılaştırılması için bir bölgede n gözlemin bulunduğu varsayalım. Dikey kesit sürece benzer olarak, akışı belirleyecek yeni sorulara cevap aranması gerekmektedir. Bu sorular,

1. Verileri açıklayacak parametreler hakkında
2. Bölgenin herhangi iki yarısındaki ortalama değerlerin karşılaştırılması hakkında
3. Bölgenin tamamındaki ortalama değer hakkında
4. Ölçülmemiş bölgedeki ortalama değer hakkında

olmalıdır. İki çalışma hakkındaki araştırılacaklar arasında benzerlik görülse de, ikinci çalışmada incelenen Z vektörünün $s = [x, y]'$ uzamsal koordinatı da göz önüne alınmalıdır. Uzamsal koordinatların dikkate alındığı model aşağıda gösterildiği gibi olur.

$$\mathbf{Z}(s) = \mathbf{X}'_s\boldsymbol{\alpha} + \gamma \quad (2.2)$$

Modeldeki $\boldsymbol{\alpha}$, katsayılar vektörünü; s noktası ise incelenen bölgede x ve y koordinatlarıyla belirlenmiş yeri gösterir. Eğer n gözlem alınırsa,

$$\mathbf{Z}(s) = [Z(s_1), Z(s_2), \dots, Z(s_n)]' \quad (2.3)$$

olur ve böylece model aşağıdaki gibi yorumlanır.

$$\mathbf{Z}(s) = \mathbf{X}_s \boldsymbol{\alpha} + \gamma \quad (2.4)$$

$\gamma \sim (0, \boldsymbol{\Sigma}(\theta_s))$ parametreleri ile dağılmaktadır. $\boldsymbol{\Sigma}$, $\mathbf{Z}(s)$ vektörünün kovaryans matrisi olmak üzere yukarıdaki modeller,

$$\text{Dikey kesit için } \hat{\boldsymbol{\beta}} = (\mathbf{X}'_1 \mathbf{V}(\theta_1)^{-1} \mathbf{X}_1)^{-1} \mathbf{X}'_1 \mathbf{V}(\theta_1)^{-1} \mathbf{Y} \quad (2.5)$$

$$\text{Uzamsal için } \hat{\boldsymbol{\alpha}} = (\mathbf{X}'_s \boldsymbol{\Sigma}(\theta_s)^{-1} \mathbf{X}_s)^{-1} \mathbf{X}'_s \boldsymbol{\Sigma}(\theta_s)^{-1} \mathbf{Z}(s) \quad (2.6)$$

olarak kestirilir. İki model arasındaki farklar teknik, istatistiksel ve hesapsal farklılıklar olarak açıklanır. Örneklemeden kaynaklanan farkların yanı sıra teknik farklılık olarak adlandırılan varyans-kovaryans matrisinden kaynaklanan farklar da olur (Schabenberger, 2005, syf.3-5).

2.2. UZAMSAL VERİ ÇEŞİTLERİ

Uzamsal istatistikte kullanılan yöntemler, analiz edilen uzamsal verinin yapısına göre farklılıklar göstermektedir. Uzamsal veriler yapılarına göre jeostatistik (geostatistical) veriler, kafes (lattice) veriler ve nokta düzen verileri olmak üzere üç gruba ayrılırlar.

2.2.1. Jeostatistik Veriler

1980'lerin başından itibaren sıkça kullanılan jeostatistik veriler, maden mühendisliği, jeoloji, matematik ve istatistik disiplinlerinin karışımı sonucunda elde edilmiştir. Jeostatistik verilerin modelleri uzamsal eğilim ve uzamsal korelasyon içerebilirler. Bu tarz verilerde eğilim – alan yöntemleri adı verilen özel yöntemler de kullanılmaktadır. Bu yöntemler ancak geniş ölçekli değişimler ve bağımsız hataların varlığının kabul edildiği durumlarda kullanılmalıdır. Jeostatistik veriler, uzamsal sürekli veriler olarak da adlandırılabilirler.

Jeostatistiğin literatürde en sık kullanıldığı konu madenciliktir. Bu konunun en önemli problemlerden biri gözlenmiş maden blok verilerinde cevherin bulunduğu yerin saptanmasıdır. Matheron (1963) çeşitli uyarlamalar sonucunda elde ettiği bu saptama sürecine “Kriging” adını vermiştir. Bunun yanı sıra toprak bilimleri ve erozyon öngörülerinde, yağmur ormanları incelemelerinde, atmosferik bilimlerde, toprak ölçümlerinde ve yeraltı sularına ilişkin ölçümlerde jeostatistik veriler sıkça kullanılmaktadır.

Jeoistatistik verilerin kullanımına örnek olarak meteorolojide hava sıcaklığının ölçülmesi gösterilebilir. Bölgeler ve yerleşim yerlerindeki hava sıcaklığını tam olarak ölçmek mümkün olmayacaktır. Çünkü çok kısa mesafelerde dahi hava sıcaklık değerinin değiştiği durumlarla karşılaşılabilir. Bu nedenle belli bölgelere kurulan hava istasyonları esas alınabilir. Şekil 2.1'de USA National Climatic Data Center raporları yardımıyla oluşturulmuş ve jeoistatistik yapıdaki verileri içeren şekil görülmektedir.

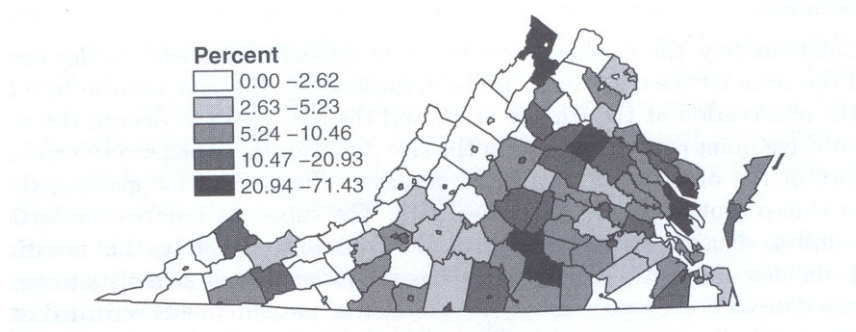


Şekil 2.1. USA National Climatic Data Center Raporları Sonucunda Hava İstasyonları (Kaynak: Schabenberger, 2005, syf.8)

Şekilde bulunan tüm noktalar Amerika'da bulunan hava istasyonlarını göstermektedir. Bu tür bir çalışma için mevcut istasyonlardaki hava sıcaklıkları dikkate alınacaktır.

2.2.2. Kafes Veriler

Kafes veriler rasgele olmayan, sayılabilir ve kesikli yapıya sahip olan veriler olarak adlandırılabilir. Bu tarz veriler uzayda bir noktadan çok bölgesel bir alanı ifade etmektedirler. Uydu yardımıyla elde edilen elektromanyetik dalgalar, arazideki mineral dağılımı bu tarz verilere örnek olarak gösterilebilir (Cressie, 1993, syf.10-13).

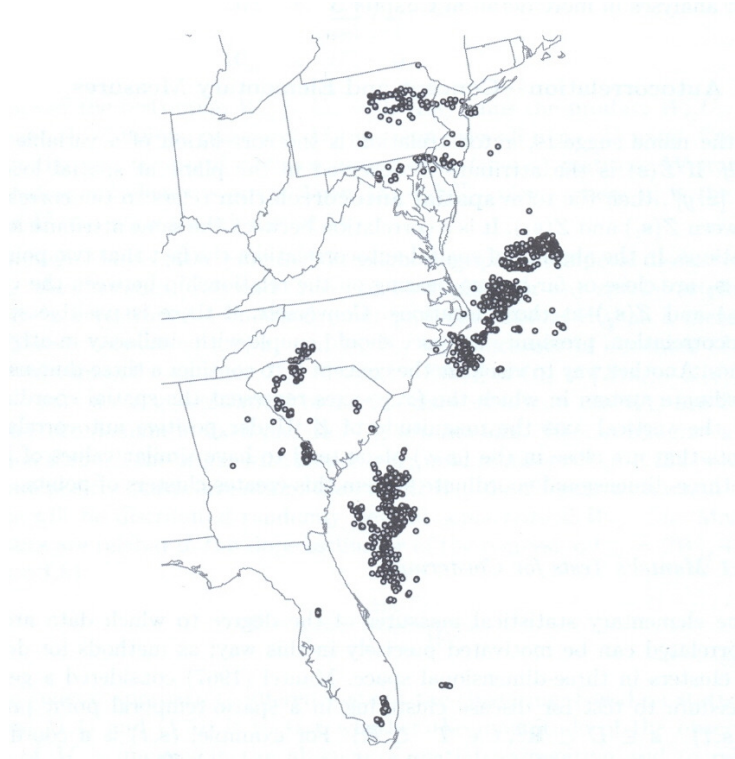


Şekil 2.2. Çocukların Kurşun Zehirlenmelerine İlişkin Oranlar (Kaynak: Schabenberger, 2005, syf.10)

Şekil 2.2'de Virginia Sağlık Departmanından alınan ve çocukların kurşun zehirlenmeleri oranını gösteren bir şekil bulunmaktadır. Departmanın çocuklarda kurşun zehirlenmesi olayını ortadan kaldırmak için uyguladığı program çerçevesinde 72 aylıktan daha küçük çocukların kan düzeyleri incelenmiştir. Test edilen çocuk sayısı ve onlara ait kandaki kurşun seviyelerine ilişkin değerler Virginia'nın tüm şehirleri için her yıl kaydedilmiştir. 2000 yılı için çocuklara ait kandaki ortalama kurşun oranı, tüm şehirler için yukarıdaki şekilde görüldüğü gibi elde edilmiştir. Bu gösterim şekli, kafes verilerin varlığı durumunda sıkça kullanılır.

2.2.3. Nokta Düzenleri

Çalışmada "olayın yeri"nin önemli olduğu durumlarda bu tarz veriler kullanılır. Ege bölgesindeki uzun iğneli çamların konumunun inceleneceği örnekte ilk aranacak cevap, bu ağaçların kümelenmesindeki biyolojik anlamlılığın ne olduğu sorusudur. Ağaçların gövde yükseklikleri değişken olarak alınabilir ve ağaçların konumları boyunca ölçümler yapılabilir. Büyük veya küçük ağaçların kümelenip kümelenmediği incelenebilir ve bu ağaçların birbirlerini nasıl etkilediği araştırılabilir. Buradaki boyut değişkeni "işaret değişkeni" olarak adlandırılır ve bu süreç "işaretlenmiş uzamsal nokta süreci" olarak adlandırılır.



Şekil 2.3. 17/20 Nisan 2003 Tarihinde Gözlenen Şimşeklerin Yeri (Kaynak: Schabenberger, 2005, syf.13)

Şekil 2.3'de Amerika'nın doğu sahillerinde 17 Nisan ve 20 Nisan 2003 tarihleri arasında gözlenen şimşekler nokta düzende gösterilmiştir. Bu tarz veriler hava yolu trafiğinin düzenlenmesinde hava trafik kontrolörlerince, deniz yolu ile gerçekleştirilen tüm olayların planlanmasında ve hatta olası bir fırtınanın gelişinin hesaplanması konusunda oldukça sık kullanılmaktadır. Bu veriler şimşeğin yüzeye çarptığı anda yaydığı elektromanyetik sinyallerin sayılması ile elde edilmiştir (Schabenberger, 2005, syf.13).

Tüm bu sınıflandırmaların yanı sıra, değişkenlerin ölçüm düzeylerine göre düzenlenmesi de farklı ölçülmüş değişkenlere uygulanacak istatistiksel analizlerin birbirlerinden farklılık göstermesi nedeniyle büyük önem taşımaktadır. Uzamsal veriler de klasik verilerde olduğu gibi sınıflayıcı (nominal), sıralayıcı (ordinal), aralıklı (interval) ve oranlı (ratio) ölçek olmak üzere 4 ölçekte toplanırlar. Tablo 2.1'de ölçek türlerine göre uzamsal veriler ve bunlara ilişkin örnekler yer almaktadır.

Tablo 2.1 Uzamsal Verilerin Sınıflandırılması

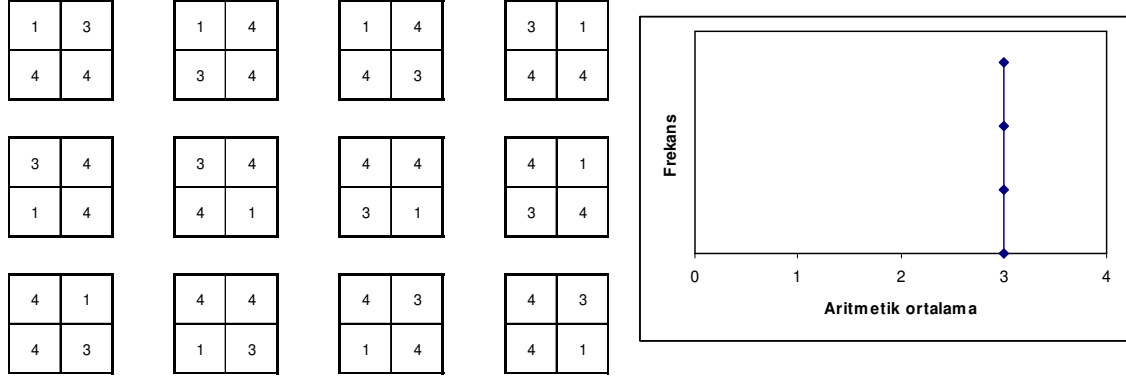
UZAMSAL GÖSTERİM				
Ölçüm Düzeyi	Nokta (P)	Çizgi (L)	Alan (A)	Yüzey (S)
Sınıflayıcı	Ev (soyulmuş/soyulmamış)	Yol (tamirde/değil)	Hayat tarzlarına göre nüfus alanları	Arazi kullanım türleri
Sıralayıcı	Yaşam kalitesine göre şehirlerin sıralanması	Yolların sıralandırılması (otoban, çevreyolu, vb.)	Gelir sınıflarına göre nüfus alanlarının sıralanması	Toprağın yapısı
Aralıklı	Şehirlerin gelişmişlik indeks değerleri	Greenwich meridyenine göre olan uzaklıklar	Alanlara göre gelişmişlik indeks değeri	Yeryüzü sıcaklığı
Oranlı	Fabrikadan elde edilen üretim oranları	Fabrika çıkışında gemiyle taşınan yük oranı	Bölgesel gelir oranları	Alana düşen yağış oranı

(Kaynak: Haining, 2003, syf.52)

Sıralayıcı ya da daha yüksek düzeyde ölçülmüş veriler daha düşük düzeylere çekilebilirler; ancak bu durumda bilgi kaybedileceği bilinir. Sınıflayıcı veriler mod ve frekans dağılımları kullanılarak karşılaştırılabilirler. Sıralayıcı verilerin karşılaştırılmasında medyan değerleri ve kutu çizimleri kullanılabilir. Aralıklı ve oranlı ölçek türüne sahip veriler de standart sapma ve ortalamalar yardımıyla karşılaştırılabilirler (Haining, 2003, syf.50-51).

2.3. UZAMSAL ORTALAMA

Uzamsal verilerden hesaplanan ortalama ile klasik yöntemle hesaplanan ortalama birbirinden farklılık göstermektedir. Bunu açıklamak için aşağıdaki şekilde verilen örnek kullanılabilir.



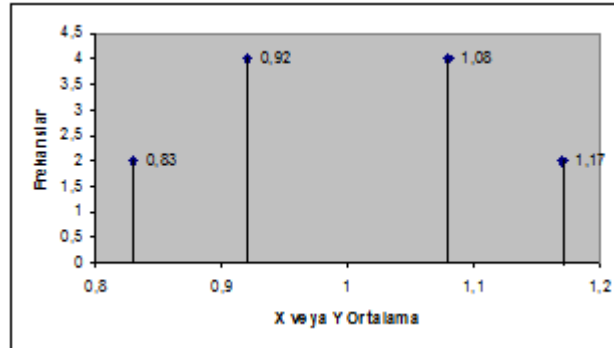
Şekil 2.4. Ortalama Hesaplamasında Kullanılan Bölgelere İlişkin Bilgiler ve Ortalama Değerleri
(Kaynak: Arlinghaus, 1995, syf.19)

Yukarıda gösterilmiş olan 12 karenin bir bölgenin haritası olduğu ve üzerindeki 4 kutunun da 4 farklı şehri gösterdiği varsayalım. Klasik ortalama hesaplandığında bu 12 bölgenin her birinin ortalamasının 3 olduğu görülmektedir. Ortalamasının bu şekilde hesaplandığı durumda uzamsal tüm etkilerin ortadan kalktığı, tüm bölgelerin aynı özelliğe sahip olduklarının düşünülmektedir. Ancak analizlerde bölgelerin ve şehirlerin uzamsal konumlarının da dikkate alınması gerekir. Şekilde birinci bölgeyi temsil eden ilk kutu dikkate alınır, bölgenin uzamsal ortalaması hesaplanırken izlenecek yol aşağıdaki gibidir:

- Kutu içinde yer alan ve her bir şehir için ayrı kodlanan değerler ağırlık olarak alınır.
- Her bir büyük karenin sol alt köşesi, dikdörtgensel koordinat ekseninin orijinine denk gelecek şekilde yerleştirilir.
- Her bir şehrin genişliğini gösterecek şekilde x ve y eksenleri işaretlenir.
- Her bir şehrin merkezi noktaları koordinat ekseninde işaretlenir ve etiketlenir: $(1/2, 1/2)$, $(3/2, 1/2)$, $(1/2, 3/2)$ ve $(3/2, 3/2)$
- Karelerin ortalamalarının hesaplanmasında ilk sütundaki ağırlıklar toplanıp bu sütunun merkez x koordinatı ile çarpılır ve ikinci sütundaki ağırlıklar toplanıp bu sütunun merkez x koordinatı ile çarpılır. Bu değerler toplanıp toplam ağırlığa bölünür. (Birinci kutu için; $(1+4)*1/2 + (3+4)*3/2/12 = 1,08$)

- Aynı işlem tüm karelere uygulanır.

Bu açıklama doğrultusunda şekilde gösterilen örneğin uzamsal ortalamasına ilişkin frekans dağılımı aşağıda gösterildiği gibi olacaktır.



Şekil 2.5. Uzamsal Ortalamalara İlişkin Frekans Dağılımı (Kaynak: Arlinghaus, 1995, syf.20)

Açıklamalarda ortalamanın hesaplanmasında sütun merkezinin x koordinatı ile çarpıldığı söylenmişti. Benzer şekilde satır merkezinin y koordinat değeri kullanılarak da ortalama hesaplanabilir. Bu durumda örneğin sol üst ilk kutusunun uzamsal ortalaması (1,08; 0,83) olur. Bu örnekte (1, 1) noktası için hesaplanan ortalama, verilmiş değerlerle ağırlıklandırılmış olan yüzeyin denge noktası olur.

Uzamsal medyan incelendiğinde ise, medyanın klasik anlamından farklılık göstermediği görülecektir. X değerlerinin yarısının birinci koordinat tarafında, y değerlerinin yarısının da ikinci koordinat tarafında olacağı bilinmektedir. Uzamsal medyan seçilen eksene bağlıdır (Arlinghaus, 1995, syf.18-22).

2.4. UZAMSAL OTOKORELASYON FONKSİYONLARI

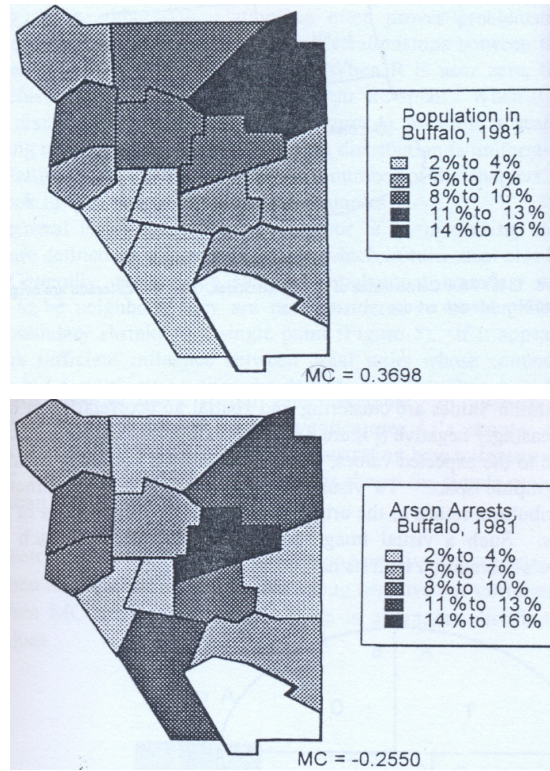
Otokorelasyon, değişkenin kendi içindeki korelasyonu anlamına gelmektedir. Uzamsal $s=[x,y]$ noktasında gözlenen $Z(s)$ değerleri dikkate alındığında uzamsal otokorelasyon, $Z(s_i)$ ve $Z(s_j)$ değerleri arasındaki ilişkiyi göstermektedir. Uzamsal otokorelasyonun bulunmaması, s_i ve s_j noktalarında bulunan $Z(s_i)$ ve $Z(s_j)$ değerleri arasında ilişki olmadığını gösterir (Schabenberger, 2005, syf.14).

Uzamsal otokorelasyonun incelenmesi durumunda karşılaşılan sonuçlarda verilerin genel durumu için aşağıdaki yorumlar yapılabilir:

- Pozitif uzamsal otokorelasyon olması incelenen değişken için benzer birimlerin uzayda bir arada kümelenildiğini gösterir.

- Negatif uzamsal otokorelasyon olması benzer birimlerin birbirinden ayrılarak benzemez birimler arasına karıştığını; ancak benzemez birimlerin uzayda kümelenmiş halde olduğunu gösterir.
- Uzamsal otokorelasyonun olmaması ise birimlerin uzayda rasgele dağıldığını göstermektedir.

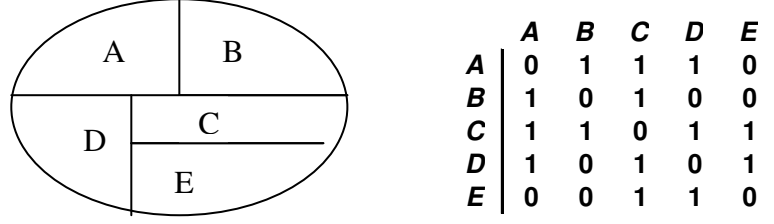
Bu durumu daha iyi açıklayabilmek için Şekil 2.6'da bir yerleşim yeri ve alt bölgelerinde suç oranına ilişkin bilgiler gösterilmiştir.



Şekil 2.6. Buffalo ve Alt Bölgelerine İlişkin Suç Oranları Yardımı ile Otokorelasyon İncelemesi
(Kaynak: Arlinghaus, 1995, syf.24-25)

İlk şekilde pozitif uzamsal otokorelasyonun varlığından söz edilir. Çünkü suç oranı şehrin kuzey kesiminde toplanmış ve güneye doğru indikçe azalmıştır. İkinci şekilde ise negatif uzamsal otokorelasyonun varlığından söz edilir. Çünkü suç oranları birbirinden bağımsız ve karışık şekilde harita üzerinde gösterilmiştir.

Örnekteki durum, matris yardımıyla da gösterilebilir. Bölgelerin i ve j olarak gösterildiği durumda, eğer i alanı j alanına bitişik veya komşuysa 1, değilse 0 değerini alır. Aşağıdaki şekil bu matrisin oluşturulmasına örnek olarak gösterilebilir:



Şekil 2.7. Matrisin Oluşturulması (Kaynak: Arlinghaus, 1995, syf.26)

Uzamsal otokorelasyonu gösteren başka istatistikler de bulunmaktadır. Moran katsayısı (The Moran Coefficient – Moran's I - MC) ve Geary oranı (Geary Ratio – Geary's C - GR) kullanılan istatistiklerden bazılarıdır.

2.4.1. Moran Katsayısı

Uzamsal ağırlıklara dayanan bu katsayının hesaplanmasında aşağıdaki eşitlikler kullanılır.

$$S_0 = \sum_{i=1}^n \sum_{j=1}^n w_{ij} \quad (2.7)$$

$$S_1 = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (w_{ij} + w_{ji})^2 \quad (2.8)$$

$$S_2 = \sum_{i=1}^n \left(\sum_{j=1}^n w_{ij} + \sum_{j=1}^n w_{ji} \right)^2 \quad (2.9)$$

$$k = \frac{m_4}{m_2^2} \quad (2.10)$$

Burada w , ağırlıklandırılmış uzamsal bağlanabilirlik (connectivity) matrisi; n , örneklem birim sayısı ve m , momentler olmak üzere Moran I katsayısı,

$$I = \frac{n}{S_0} \times \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (2.11)$$

ile verilir. Moran I katsayısının ortalama ve varyansı,

$$E(I) = -\frac{1}{(n-1)} \quad (2.12)$$

$$Var(I) = \frac{n}{(n-1)^3 S_0^2} \left\{ [(n^2 - 3n + 3)S_1 - nS_2 + 3S_0^2] - k[(n^2 - n)S_1 - 2nS_2 + 6S_0^2] \right\} - \frac{1}{(n-1)^2} \quad (2.13)$$

olarak hesaplanır. Hesaplanan değerler için:

- MC değeri +1'e yaklaşıyorsa güçlü pozitif otokorelasyon,
- MC değeri -1'e yaklaşıyorsa güçlü negatif otokorelasyon,
- MC değeri -1/(n-1)'e yaklaşıyorsa değerlerin rasgele oldukları söylenebilir.

2.4.2. Geary Oranı

Yukarıda tanımlanmış tüm değerler aynı kalmak kaydıyla Geary oranı, ortalaması ve varyansı,

$$c = \frac{n-1}{2S_0} \times \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - x_j)^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (2.14)$$

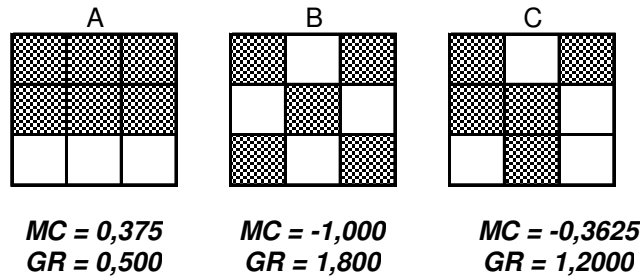
$$E(c) = 1 \quad (2.15)$$

$$Var(c) = \left\{ \begin{array}{l} (n-1)S_1 [n-3n+3-(n-1)k] - \frac{1}{4}(n-1)S_2 [n+3n-6-(n^2-n+2)k] \\ + S_0^2 [n^2-3-(n-1)^2 k] \end{array} \right\} \times \frac{1}{n(n-2)^2 S_0^2} \quad (2.16)$$

olarak hesaplanır. Hesaplanan değer için:

- GR değeri 0'a yaklaşıyorsa güçlü pozitif otokorelasyon,
- GR değeri 2'ye yaklaşıyorsa güçlü negatif otokorelasyon,
- GR değeri 1'e yaklaşıyorsa değerlerin rasgele oldukları söylenir.

Şekil 2.8'de bulunan A, B ve C haritaları bu ölçümlerdeki görsel farklılıkları göstermektedir.



Şekil 2.8. Üç Bölge Yardımıyla Otokorelasyonun Gösterilmesi (Kaynak: Arlinghaus, 1995, syf.27)

Şekilde A haritası pozitif otokorelasyonu, B haritası negatif otokorelasyonu ve C haritası ise rasgele dağılımı göstermektedir (Arlinghaus, 1995, syf.24-28).

2.4.3. Zaman Serilerinin Otokorelasyon Fonksiyonu

Seriler söz konusu olduğunda otokorelasyon, stokastik süreci oluşturacak rasgele değişkenler ailesindeki korelasyon olarak açıklanabilir. Zaman serilerinde korelasyonun bu şekli “seri korelasyon” olarak bilinir.

Durağan bir zaman serisinde $E[Z(t_i)] = 0$ ve $Var[Z(t_i)] = \sigma^2$ olacaktır. Zaman serilerinde otokorelasyon tek bir ölçüm yapmaktansa zaman noktalarının fonksiyonlarının kullanılması ile ölçülür. t_i ve t_j noktalarındaki serilerin kovaryans fonksiyonu,

$$Cov[Z(t_i), Z(t_j)] = E[Z(t_i), Z(t_j)] = C(t_j - t_i) \quad (2.17)$$

ve otokorelasyon fonksiyonu,

$$R(t_j - t_i) = \frac{Cov[Z(t_i), Z(t_j)]}{\sqrt{Var[Z(t_i)]Var[Z(t_j)]}} = \frac{C(t_j - t_i)}{C(0)} \quad (2.18)$$

olur. Zaman serilerinde veya uzamsal verilerde daha sık olarak pozitif otokorelasyona rastlanır.

2.4.4. Uzamsal Otokorelasyon Fonksiyonu

Uzamsal durumlarda otokorelasyon, s_i ve s_j yerlerindeki değerlerin stokastik bağımsızlığını yansıtmaktadır. Eğer korelasyon pozitifse yüksek (alçak) değerlerin yüksek (alçak) değerlerle çevrelediği düşünülür. Eğer negatifse yüksek (düşük) değerlerin düşük (yüksek) değerlerle çevrelenmiş olması beklenir. Buradan gelen çıkarım uzamsal ayrılmanın büyümesiyle korelasyonun düşeceği

Uzamsal sürecin kovaryans fonksiyonu,

$$C(s, h) = Cov[Z(s), Z(s + h)] = E[\{Z(s) - \mu(s)\}\{Z(s + h) - \mu(s + h)\}] \quad (2.19)$$

ve otokorelasyon fonksiyonu (veya korelogramı),

$$R(s, h) = \frac{C(h)}{\sqrt{Var[Z(s)]Var[Z(s + h)]}} \quad (2.20)$$

olur. Verilen bu fonksiyonların zaman serilerinin otokorelasyon fonksiyonları tanımının biraz daha genişletilmiş biçimi olduğu görülmektedir. Bu fonksiyonların birbirine benzer olması iki yöntemin aynı süreçleri içerdiği düşüncesini akla getirmemelidir. Çünkü, (i) uzamsal veri analizi için bulunan bir çok istatistiksel yöntem, zaman serileri analizindeki yöntemlerden bağımsız olarak geliştirilmiştir (ii) zaman serileri modellerine uzamsal uyarlamalar oldukça zor gerçekleşmektedir.

Zaman serileri analizlerinde olduğu gibi uzamsal veri analizinde de çeşitli varsayımların gerçekleştirilmesi gerekmektedir. Kovaryans fonksiyonu $C(s,h)$ 'ın durumu üzerindeki varsayım durağanlık varsayımı olarak bilinir (Schabenberger, 2005, syf.25-27).

2.5. UZAMSAL ÖRNEKLEME

Uzamsal örnekleme, coğrafi olarak belirlenmiş kitleden verilerin nasıl toplanacağını belirlemek için geliştirilmiştir. Birçok olay veya birim, uzamsal konumları ile ilişkilendirilebildiği için yapılacak herhangi bir örnekleme hatası ciddi sorunlara neden olur.

İlk klasik örnekleme çalışmalarında [Cochran 1977, Deming 1950, Jessen 1978, Kish 1965, Yates 1981] uzamsal verilerin varlığının bilinmemesi ve uzamsal analizlerin kullanılmaması yüzünden örnekleme sürecinde verilerin uzamsal durumu göz ardı edilmiştir. Ancak uzamsal analizin kullanılmaya başlanması ile birlikte örnekleme yapısı ile uzamsal istatistik arasındaki etkileşim daha da önem kazanmıştır.

Örneklemenin, kitle parametrelerini (ortalamalar, toplamalar veya oranlar) kestirmede çok önemli olduğu bilinmektedir. Bu parametrelerin örnekleme istatistiklerinden kestirilmesi "tasarım tabanlı" (design based) yöntem olarak adlandırılır. Kitlenin uzayda dağıldığı durumlarda örneğin analiz edilmesinde örnekleme birimlerinin uzamsal koordinatları kullanılabilir (Arlinghaus, 1995, syf.31-32).

2.5.1. Uzamsal Kitle ve Örnekleme

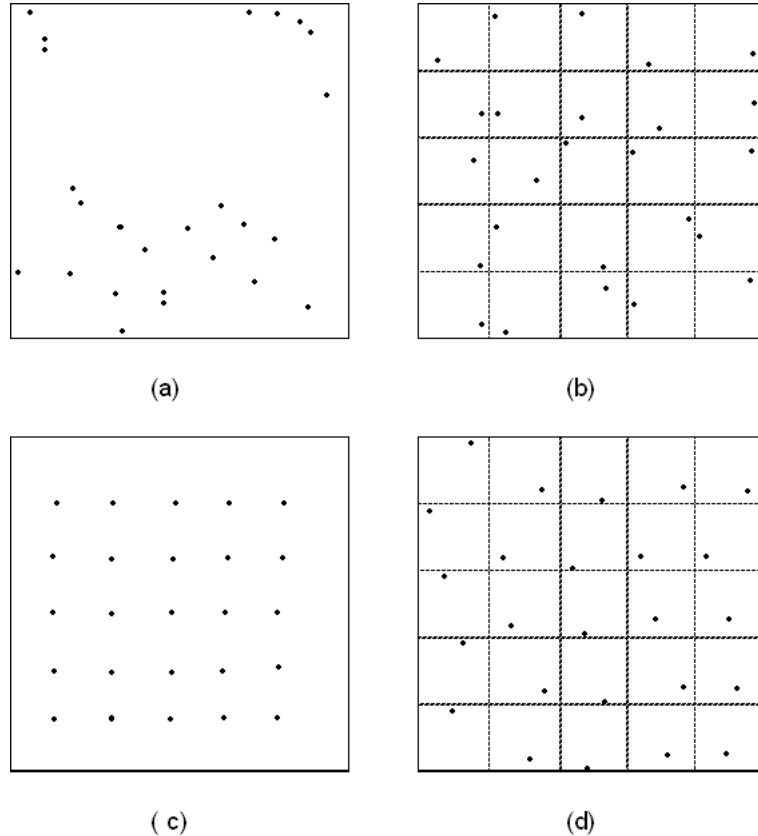
Bu konuyu açıklayabilmek için "U" olarak isimlendirilen kitle, örnekleme probleminde ilgilenilen bölge olsun. Bu kitlenin birimleri de bu bölgede yapılmış olan ölçümler olur. Kesikli bir kitle $U=\{u_1, u_2, \dots, u_N\}$ (göller, ağaçlar gibi) şeklinde sonlu sayıda birim içerir. Çoğu zaman bu birimlere ait tek bir nokta olarak algılanacak uzamsal konumlar olur. Ancak "insan" gibi birimler için, uzayda durağan olmadıkları için sabit ve tek bir konum belirlemek mümkün olmayacaktır. Bu nedenle örneklemenin

yapılacağı anda tam koordinatları yerine birimi içine alacak kareler oluşturmak suretiyle uzamsal konum belirlenebilir. Bu yüzden hareketli kitlenin örnekleme sürecinde sorunlara neden olduğu bilinmektedir.

Sürekli kitleler sürekli alan ya da bölgeleri içerirler (parklar, bataklıklar vb.). Bu durumda bu kitledeki birimler $U=\{u:u\in R^d, d=1,2\}$ şeklinde gösterilebilir. Bu tarz kitleler söz konusu olduğunda alan bölümlere ayrılarak sonlu bir küme haline getirilir. Bu durum kitlenin kendi içinde çerçevelere bölünmesi ile sağlanır. Örnekleme sürecinde de bu çerçeveler kullanılır.

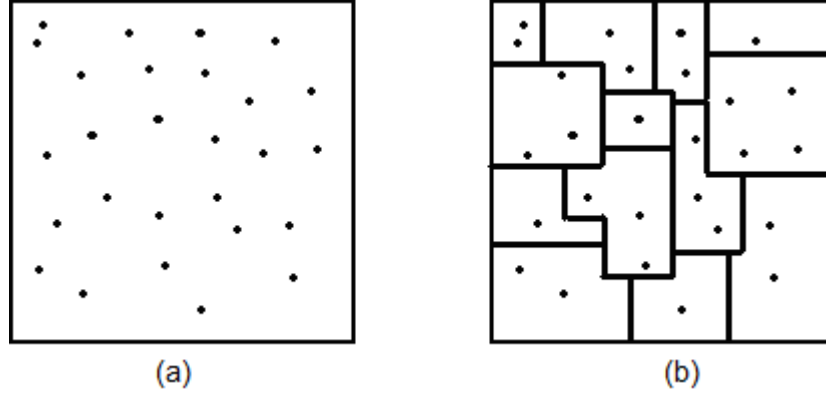
Çerçeveleme işleminin gerçekleşmesinden sonra hedef kitle ile çerçevelenmiş kitle arasındaki farklılıkların gösterilmesi oldukça önemlidir. Hedef kitle çalışmanın konusunu oluştururken, çerçevelenmiş kitle ise sadece örneklemede kullanılan birimleri göstermektedir (Arlinghaus, 1995, syf.33-36).

n birimli örneklem noktasının $\{x_1, \dots, x_n\}$ olarak gösterildiği birçok örnekleme planı vardır. Bunlardan bazıları Şekil 2.9'da gösterilmiştir.



Şekil 2.9. Örneklem Planlarına İlişkin Şablonlar (Kaynak: Ripley, 2004, syf.20-21)

(a) şeklinde tekdüze (uniform) rasgele, (b) şeklinde merkezi (centric) sistematik, (c) şeklinde tabakalı (stratified) rasgele, ve (d) şeklinde ise bağlantısız (nonaligned) sistematik örnekleme düzenekleri görülmektedir (Ripley, 2004, syf.19-21).



Şekil 2.10. Örneklemelerde Çerçevelerin Oluşturulması (Kaynak: Arlinghaus, 1995, syf.34-35)

Şekil 2.10'da gösterilen (a) şeklinde sürekli uzamsal bir kitlenin bireylerinin konumları görülmektedir. Bu tarz yapıya sahip veri kümesine uzamsal örnekleme uygulandığında çerçevelerin oluşturulması gerekmektedir. (b) şeklinde de uzamsal örnekleme sürecinin çerçevenleme kısmı gösterilmiştir (Arlinghaus, 1995, syf.35).

2.5.2. Uzamsal Örnekleme Yöntemleri

Bir A bölgesindeki x ölçümleri $Z = Z(x)$ olarak gösterilsin. Burada $\mu = E(Z(x))$ ve $\sigma^2 = Var(Z(x))$ olarak gösterilir. $\tilde{Z}(A) = \int_A Z(x)dx / a$ ifadesinin kestiriminde

$\bar{Z} = \frac{1}{n} \sum_{i=1}^n Z(x_i)$ eşitliği kullanılacaktır. Bu ifadedeki a , A bölgesinde bir alanı veya

hacmi göstermektedir. Bilinen bu ifadeler yardımı ile örnekleme yöntemleri aşağıdaki gibi açıklanacaktır:

2.5.2.1. Tekdüze (Uniform) Rasgele Örnekleme

x_1, \dots, x_n noktalarının rasgele konumlarının beklenen değerlerinden,

$$E\{\bar{Z} - \tilde{Z}(A) / Z\} = \frac{1}{n} \sum_{i=1}^n \frac{1}{a} \int_A Z(x)dx - \frac{1}{a} \int_A Z(x)dx = 0 \quad (2.21)$$

$$Var\{\bar{Z} - \tilde{Z}(A) / Z\} = \frac{1}{n^2} \sum var\{Z(x_i)\} \text{ ve bağımsızlık altında,}$$

$$\begin{aligned}
&= \frac{1}{n} \text{var}[Z(x_1)] = \frac{1}{na} \int_A [Z(x) - \tilde{Z}(A)]^2 dx \\
&= \frac{1}{na} \int_A [Z(x) - \mu]^2 dx - \frac{1}{na^2} \iint_A [Z(X) - \mu][Z(Y) - \mu] dx dy \quad (2.22)
\end{aligned}$$

$$\text{Var}\{\bar{Z} - \tilde{Z}(A)\} = E[\text{var}\{\bar{Z} - \tilde{Z}(A) / Z\}] = \frac{1}{n} [\sigma^2 - E\{C(X, Y)\}] \quad (2.23)$$

Olur; X ile Y, A noktasında bağımsız tekdüze dağılıma sahiptir (Ripley, 2004, syf.22).

2.5.2.2. Tabakalı Rasgele Örnekleme

A'nın s_1, \dots, s_m olmak üzere m tabakaya ayrıldığı varsayalım. Buradan aşağıdaki eşitlikler yazılabilir:

$$E\{\bar{Z}|Z\} = E\left(\frac{1}{m} \sum \bar{Z}_i | Z\right) = \frac{1}{m} \sum \tilde{Z}(S_i) = \tilde{Z}(A) \quad (2.24)$$

$$\text{Var}\{\bar{Z} - \tilde{Z}(A)|Z\} = \frac{1}{m^2} \sum_i \frac{1}{k s_i} \left[\int_{s_i} (Z(X) - \mu)^2 dx - \iint_{s_i} (Z(X) - \mu)(Z(Y) - \mu) dx dy \right] \quad (2.25)$$

Her bir tabaka için hesaplanan $\bar{Z}_1, \bar{Z}_2, \dots, \bar{Z}_m$ ortalamalarından,

$$\text{Var}\{\bar{Z} - \tilde{Z}(A)\} = E[\text{Var}\{\bar{Z} - \tilde{Z}(A)|Z\}] \quad (2.26)$$

$$= \frac{1}{m^2 k} \sum [\sigma^2 - E\{C(X_i, Y_i)\}] \quad (2.27)$$

$$= \frac{1}{n} [\sigma^2 - E\{C(X_i, Y_i)\}] \quad (2.28)$$

bulunur. Burada X_i, Y_i noktaları i'nci tabakadaki rasgele noktaları göstermektedir. Tabakalı örneklemede noktalar $E\{C(X, Y)\}$ 'yi olabildiğince büyük yapacak şekilde küçük seçilmelidir (Ripley, 2004, syf.23).

2.5.2.3. Sistemik Örnekleme

$u \in A$ olarak etiketlenen noktaların örnek olduğu varsayalım. Buna göre:

$$E(\bar{Z}) = \frac{1}{n} \sum_u E(Z(u)) = \frac{1}{n} \sum_u \mu = \mu \quad (2.29)$$

$$\text{Var}(\bar{Z} - \tilde{Z}(A)) = \frac{1}{n^2} E \left[\sum_u \{Z(u) - \tilde{Z}(A)\} \right]^2 \quad (2.30)$$

$$= \frac{1}{n^2} \sum_{u,v} C(u,v) - \frac{2}{n} \sum_u E \{ (Z(u) - \mu)(\tilde{Z}(A) - \mu) \} + E \{ \tilde{Z}(A) - \mu \}^2 \quad (2.31)$$

$$= \frac{1}{n^2} \sum_{u,v} C(u,v) - 2 \sum_u \frac{1}{an} \int_A C(u,y) dy + \frac{1}{a^2} \iint_A C(x,y) dx dy \quad (2.32)$$

olur (Ripley, 2004, syf.23).

Tüm bu açıklamalardan yola çıkarak A bölgesinde dikkate alınan n grup için,

- Basit rasgele örneklemede, gruplar bağımsız olarak ve A bölgesinde tekdüze dağılıma sahip olarak seçilir.
- Tabakalı rasgele örneklemede A bölgesi birbirinden farklı tabakalara bölünür ve her tabakaya basit rasgele örnekleme uygulanır.
- Sistemik rasgele örnekleme aynı değerlerle sistemik olarak grupların seçildiği örnekleme yöntemidir.
- Kümelere göre rasgele örnekleme uzamsal olarak birbirine yakın olan gruplardan rasgele seçimler yapılması esasına dayanır. Ancak uzamsal literatürde bu örnekleme yöntemi çok sık kullanılmamaktadır.

3. UZAMSAL VERİLERDEKİ DEĞİŞİMLERİN MODELLENMESİ

Uzamsal veriler yer ve nitelik ile ilgili bilgiler içerirken, uzamsal olmayan veriler sadece niteliğe ilişkin bilgiler içermektedir. Örneğin, bir fabrikadaki işçilerin maaşları uzamsal olmayan veriye örnek gösterilebilirken bir hastalığa sahip insanların ülkenin farklı bölgelerine göre dağılım oranları uzamsal verilere örnek gösterilebilir. Farklı zamanlarda bir bölgedeki işsizlik oranı uzamsal olmayan ama farklı bölgelerdeki işsizlik oranı uzamsal veriye örnek gösterilebilir. Uzamsal olmayan ve uzamsal veriler arasındaki farkın, uygulanacak yöntemin belirlenmesi için bilinmesi gerekmektedir.

Bölgesel ilişkilerin söz konusu olduğu ancak uzamsal yapıya sahip olmayan veriler için birçok teknik geliştirilmiştir. Spline fonksiyonları, konumsal ağırlıklı en küçük kareler regresyonu (LOWESS), çekirdek regresyon yöntemi ve ekonometri literatüründeki değişken parametre modelleri bu tekniklerden bazılarıdır (Fotheringham, 2006, syf.1-3).

Uzamsal yapıya sahip verilere ilişkin hesaplanan istatistikler bölgesel (lokal) ve genel (global) uzamsal istatistikler olarak iki grup için yorumlanabilirler. Örneğin, ABD'deki ortalama yağış miktarı genel istatistik olarak adlandırılabilirken; ölçüm istasyonlarından elde edilen yağış ölçüm verileri bölgesel istatistik olarak adlandırılabilirler. Çalışılan alanda eşit ağırlıklı veriler ile ayarlanmış olan model genel model olarak adlandırılır ve bu model yardımıyla genel parametreler kestirilir. Uzamsal limitlenmiş verilerle ayarlanmış olan model de bölgesel parametrelerin kestirildiği bölgesel modeller olarak adlandırılır. Bölgesel ve genel istatistikler arasındaki farklar aşağıdaki Tablo 3.1'de gösterilmiştir.

Genel istatistiklerin tek değerli olması, örneğin tek ortalamanın, tek standart sapmanın hesaplanması demektir. Bölgesel istatistiklerin birden çok değerli olması demek, çalışma dahilindeki her bir bölge için ayrı ayrı değerlerin hesaplanması anlamına gelmektedir.

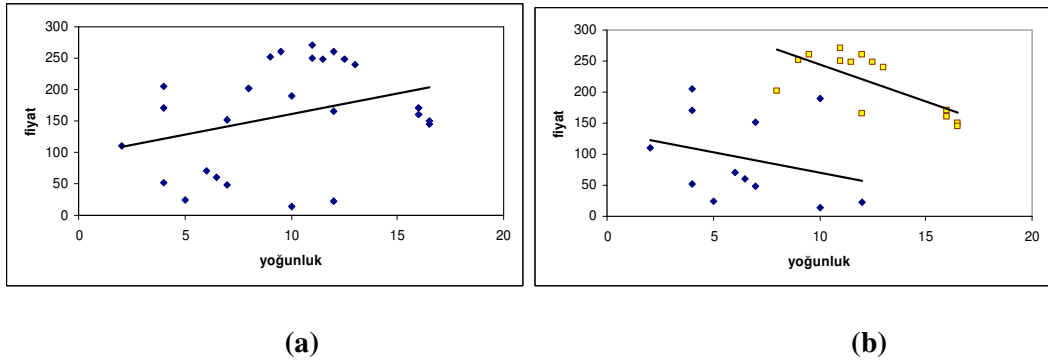
Tablo 3.1. Genel ve Bölgesel İstatistikler Arasındaki Farklar

Genel	Bölgesel
○ Tüm bölge için verileri özetler	○ Genel istatistikleri bölgesel olarak dağıtır
○ Tek bir istatistik değeri verir	○ Birden çok istatistik değeri verir (bölgeler için)
○ Haritalanamaz	○ Haritalanabilir
○ Coğrafi Bilgi Sistemleri kullanılamaz	○ Coğrafi Bilgi Sistemleri kullanılabilir
○ Uzamsal değildir veya limitli uzamsaldır	○ Uzamsaldır
○ Uzaydaki benzerlikleri vurgular	○ Uzaydaki farklılıkları vurgular
○ Düzenlilik veya kurallar arar (klasik regresyon gibi)	○ İstisna veya bölgesel "önemli" noktaları arar (coğrafi ağırlıklandırılmış regresyon gibi)

(Kaynak: Fortheringham, 2006, syf.6)

Genel istatistiklerin haritalanamaz ve coğrafi bilgi sistemleri (GIS) kullanılamaz olması demek, tek bir istatistik değerinin olmasından dolayı coğrafi bilgi sistemlerinin uygulanamaz olması anlamına gelmektedir.

Bölgesel ilişkilerin genel kestiricilerle açıklanması yanlış sonuçların ortaya çıkmasına neden olacaktır. Bu durum Simpson Paradoksu (Simpson's Paradox, Simpson, 1951) ile açıklanabilir. Simpson Paradoksu veri gruplarının ayrı ayrı veya bir arada analiz edilmesinin neden olduğu ters sonuçları ortaya koymaktadır. Bunu açıklamak için aşağıdaki örnek kullanılabilir. Şekildeki noktalar evlerin fiyatları ile bölgedeki nüfus yoğunluğu arasındaki ilişkiyi göstermektedir.



Şekil 3.1. Verilerin Toplanmasına İlişkin Sonuçlar (Fortheringham, 2006, syf.18)

İncelenen (a) şeklinde elde edilen doğrusal regresyon çizgisi iki nitelik arasındaki pozitif ilişkiyi göstermektedir. Yani ev fiyatları arttıkça nüfus yoğunluğu da artacaktır. Ancak (b) şekli incelendiğinde ve veriler iki bölgeye ayrılıp dikkate alındığında her iki bölgede de negatif ilişki olduğu görülecektir. Bu nedenden dolayı Simpson

Paradoksu veri setlerinin toplanmasının sakıncalarını göstermekte ve karşılaşılabilecek hataları minimuma indirmektedir (Fotheringham, 2006, syf.6-8).

3.1. İLİŞKİSİZ HATALAR İÇİN DOĞRUSAL MODELLER

Uzamsal yapıya sahip veriler, konum bilgisi içermeyen ve uzamsal yapıda olmayan veriler gibi modellenebilirler ve böylece değişimler oluşturulan modeller yardımıyla açıklanabilirler.

Bu modeller, “ilişkisiz hatalar için doğrusal modeller” ve “ilişkili hatalar için doğrusal modeller” başlıkları altında incelenecektir. Bunların yanı sıra modeller kendi içlerinde yapılarına göre tanımlayıcı ve açıklayıcı değişkenlerin yer aldığı modeller olarak ikiye ayrılacaktır.

Bu bölümde ilişkisiz hataların söz konusu olduğu durumlarda kullanılabilecek modeller açıklanacaktır. Ancak literatürde model hatalarının korelasyonsuz olduğu zamanlarda standart doğrusal modelin kullanılmasının uygun olduğu belirtilmiştir.

3.1.1. Sürekli Değişkenler için Tanımlayıcı Modeller

Literatürde, sürekli değişkenler için açıklanacak olan modeller, bağımsız hatalarla trend yüzey (trend surface) modelleri, semi-variogram ve kovaryans modelleri olarak bilinir.

3.1.1.1. Bağımsız Hatalarla Trend Yüzey Modelleri

Y bağımlı değişkeni için trend modeli,

$$\mathbf{y} = \mathbf{A}\theta + e \quad (3.1)$$

olacaktır. \mathbf{y} , n örneklem grubuna ilişkin $n \times 1$ boyutlu veri vektörü; \mathbf{A} , konum koordinatları matrisi; θ , trend yüzey parametresini ve e , hata terimlerini göstermektedir. \mathbf{A} ve θ 'nın büyüklüğü trende bağlı olarak belirlenmektedir.

Bu modellerde karşılaşılan en önemli sorun, modelin uyumunun az olması durumunda trend yardımıyla hesaplanan hata terimlerinin uzamsal otokorelasyona sahip olmasıdır. Hataların otokorelasyona sahip olması da modelin anlamlı olmadığı halde istatistiksel olarak anlamlı çıkmasına neden olur.

3.1.1.2. Semi-Variogram ve Kovaryans Modelleri

Semi-variogram ve kovaryans çizimleri, ortalama etrafındaki uzamsal dağılımın gösterilmesinde kullanılır. Buradaki problem, bu dağılımı gösterebilecek bir modelin

tanımlanmasıdır. Örneklem değerlerine modeli uydurabilmenin iki yolu vardır: görsel ve grafiksel yöntemler kullanarak şekle uygun modelin kestirilmesi; istatistiksel yöntemlerle modelin kestirilmesi. İkinci yöntem, semi-variogram veya kovaryans şekline uygunluğuna göre, normal dağıldığı varsayımı ile en çok olabilirlik veya kısıtlanmış en çok olabilirlik gibi çeşitli yöntemlere ayrılmaktadır. Yöntemlerin uygun olanı, “parsimoni” prensipleri de dikkate alınarak, Akaike bilgi kriteri (AIC) yardımıyla belirlenebilmektedir. MSR, hata kareleri ortalaması iken AIC kriteri,

$$AIC = n \ln(MSR) + 2p \quad (3.2)$$

ile verilir. Burada n, semi-variogram çizimindeki noktaları; p, modeldeki parametre sayısını göstermektedir. AIC değeri küçük olan modelin seçilmesi kuralı burada da geçerlidir (Haining, 2003, syf.328-329).

3.1.2. Kesikli Değişkenler için Tanımlayıcı Modeller

Tam uzamsal modeller, genelleştirilmiş doğrusal modeller için geliştirilen standart yöntemlere tam uyum göstermemektedirler. y_i değerlerinin bağımsız olma koşulu uzamsal modellerde sağlanamadığı için uyum bozulmaktadır. Bu nedenle iki kestirim yöntemi önerilmektedir: Yapay (pseudo) olabilirlik kestirimi ve kodlama (coding) yöntemleri.

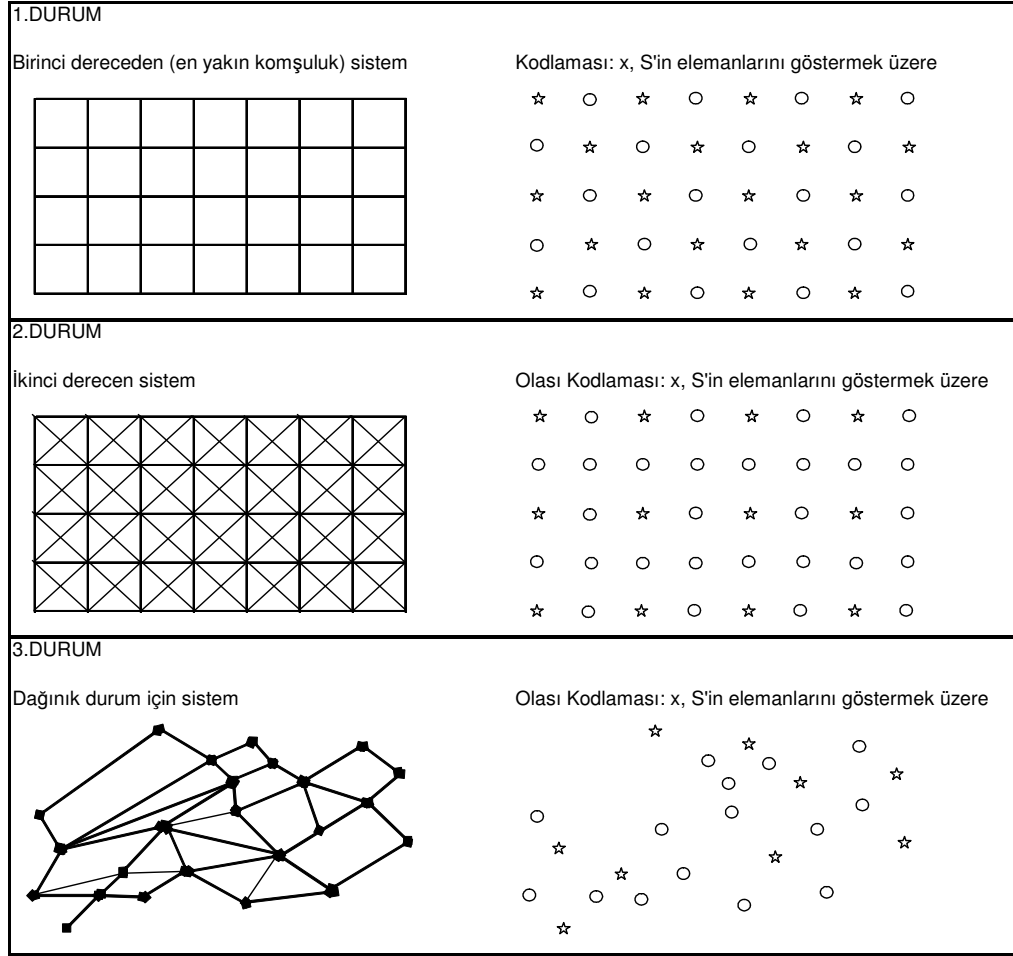
Yapay olabilirlik kestirimi, bilinmeyen parametrelerin olması durumunda koşullu olasılıkların logaritmalarının toplamını en büyükleyen yöntemdir. Kodlama yöntemi de gözlemlerin koşullu bağımsız olduğu yerlerde bir alt küme kullanılmasını gerektirir. Yapay olabilirlik kestirimlerinin maksimizasyonunu hesaplamak için kullanılacak olan koşullu olabilirlik, $\alpha_i = \theta_{0,0} + \theta_{1,0}s_1(i) + \theta_{0,1}s_2(i)$ ve

$\sum_{j \in N(i)} b(i, j)y(j) = \beta \sum_{j \in N(i)} w(i, j)y(j)$ olarak kabul edildiği durumda, i, N(i) birim içerisinde birinci dereceden komşu i birimi, w(i,j) ağırlık matrislerini ve $(\theta_{0,0}, \theta_{1,0}, \theta_{0,1}, \beta)$ parametreleri göstermek üzere,

$$\sum_i \ln \left\{ \exp \left[y(i) \left(\alpha(i) + \beta \sum_{j \in N(i)} w(i, j)y(j) \right) \right] / \left[1 + \exp \left[\left(\alpha(i) + \beta \sum_{j \in N(i)} w(i, j)y(j) \right) \right] \right] \right\} \quad (3.3)$$

şeklinde hesaplanır. Kodlanmış kestirici de yukarıdaki ifadenin maksimizasyonu ile elde edilebilir.

Şekil 3.2'de grupların birbirlerine komşuluklarını ve birinci dereceden kodlamalarını göstermektedir.



Şekil 3.2. Birinci Dereceden Komşulukların Kodlanması (Haining, 2003, syf.335)

Yapay olabilirlik kestirimi yüksek düzeyli durumlarda daha iyi sonuçlar vermektedir. Ancak kodlanmış kestirimler parametreler hakkında sonuçlar çıkarılmasına yardımcı olacaktır (Haining, 2003, syf.334-335).

3.1.3. Klasik Açıklayıcı Modeller

Regresyon teorisi birçok denemenin sonuçlarına uyarlanabilmektedir. Farklı denemelerin düzeylerinin (X_1, \dots, X_k) bağımlı değişkenin düzeyini (Y) etkilediği bilinmektedir. Hata terimleri de deneme sonuçlarına göre elde edilmektedir. Normal dağılımın varlığında doğrusal regresyonda bağımlı değişken $Y(s)$ ve ortalama fonksiyonu veya beklenen değeri de $\mu(s)$ olarak gösterilir. Bu bölümün amacı $\mu(s)$ yardımıyla farklı model çeşitlerini açıklamaktır. Uzamsal değişimin gösterileceği modelde uzamsal birimlerin (alan veya noktalar) bir şekilde gösterilmesi gerekmektedir. $s[x,y]$ uzamsal koordinatları göstermek üzere, bu bölgedeki bağımlı

değişkenin düzeyiyle kısıtlanması gerekmemektedir. Burada normal doğrusal model kullanılacaktır.

Klasik anlamda doğrusal regresyon modeli,

$$Y(s) = \mu(s) + e(s) \quad s = 1, \dots, n \quad (3.4)$$

$$\mu(s) = \beta_0 + \beta_1 X_1(s) + \dots + \beta_k X_k(s) \quad (3.5)$$

ile verilir. Burada β_0 sabit parametreyi ve β_1, \dots, β_k regresyon parametrelerini göstermektedir. n , nokta veya alanların sayısıdır. $\{e(s)\}$ bağımsız ve $N(0, \sigma^2)$ 'dir. Eğer açıklayıcı değişkenlerin Y üzerindeki etkileri "dikey" ise (s noktasındaki cevap değeri s 'deki açıklayıcı değişken düzeyinin fonksiyonuysa) ve modeldeki $\{e(s)\}$ hataları bağımsız varsayılıyorsa bu model "uzamsal değişim modeli" olarak kabul edilir. Eğer açıklayıcı değişkenlerde uzamsal korelasyon görülürse bu durum hata terimlerini de etkileyecektir. Böyle bir durumda sabit varyanslılık varsayımının tekrar gözden geçirilmesi gerekmektedir. Varsayımların sağlanması durumunda iyi modelin yüksek uyuma (R^2) sahip olması beklenmektedir (Haining, 2003, syf.351).

Bunun yanı sıra uzamsal veri analizinde süreç, modeli elde etmek için verilerin kullanılmasıyla başlar. Bir sonraki adımda Moran testi yardımıyla hataların uzamsal otokorelasyona sahip olup olmadıkları incelenir. Bu testte sıfır hipotezi hata terimlerinin bağımsız olduğunu, alternatif hipotez hataların uzamsal otokorelasyona sahip olduğunu gösterir.

Eğer uzamsal otokorelasyonun varlığı tespit edilirse "uzamsal hata modeli" en çok olabilirlik ile oluşturulmalıdır. Uzamsal otokorelasyonun olması durumunda regresyon modelinin kullanılması regresyon parametrelerinin en küçük kareler kestiricilerinin etkisiz ve varyans kestiricisinin olması gerekenden daha düşük çıkmasına neden olur. Bu durum R^2 değerinin olması gerekenden daha yüksek çıkmasına ve hipotezlerin 1. tip hata düzeylerinin büyümesine neden olur. Uzamsal otokorelasyon testinin sonucunda alternatif hipotezin kabul edilmesi (hataların uzamsal otokorelasyona sahip olması) durumunda çalışmayı yapan kişinin aşağıdaki durumlara dikkat etmesi gerekmektedir:

- (i) Modelin, değişimi açıklamada kullanılan doğru bir model olup olmadığı;
- (ii) Hataların, bağımlı değişkenin durumu hakkında bilgi içerip içermediği ve bunun dikkate alınıp alınmadığı;
- (iii) Seçilen anlamlılık düzeyinin çalışma için uygun olup olmadığı.

Yeni açıklayıcı değişkenlerin girişi ile uygulanacak değişik süreçte Lagrange çarpanı (LM) kullanılabilir. LM testleri genel alternatif hipotezlerden biraz daha özel testlerdir.

Açıklayıcı değişkenler ile bağımlı değişken arasındaki etki “dikey” değil de “yatay” bileşenlere sahip olabilir. Bu, $s[x,y]$ noktasındaki cevap değişken değerinin bir veya iki açıklayıcı değişken düzeyinin bir fonksiyonu olabildiği anlamına gelmektedir. $s[x,y]$ 'nin komşularını gösteren bölgeler kümesi bunun için örnek olarak gösterilebilir. Komşuluk etkilerinin de gösterildiği model aşağıdaki gibi düzenlenir:

$$Y(s) = \mu(s) + e(s) \quad (3.6)$$

$$\mu(s) = \beta_0 + \beta_1 X_1(s) + (\beta_{1,N} \sum_{j \in N(i)} w(i, j) X_1(j)) + \beta_2 X_2(i) + \dots + \beta_k X_k(i) \quad (3.7)$$

Burada $\beta_{i,N}$, komşuluk ilişkilerini gösteren parametre olur. Komşuluk etkilerinin uzamsal skalası, ağırlık matrisinin 0'dan farklı değerleri ile ele alınır. Bu durumda ağırlık matrisi standartlaştırılır ve böylece komşuluk etkisi, komşuluk ortalaması olarak gösterilir. X_1 açıklayıcı değişkeninin “yatay” etkisi de diğer açıklayıcı değişkenler gibi $\mu(s)$ ile modellenir. Oluşturulacak model aşağıdaki gibidir:

$$Y(s) = \mu(s) + \rho \sum_{j \in N(i)} w(i, j) Y(j) + e(s) , \quad s=1, \dots, n \quad (3.8)$$

$$\mu(s) = \beta_0 + \beta_1 X_1(s) + \beta_2 X_2(s) + \dots + \beta_k X_k(s) \quad (3.9)$$

Modelin matris notasyonunda gösterimi aşağıdaki gibidir.

$$\mathbf{Y} = \rho \mathbf{WY} + \mathbf{X}\beta + \mathbf{e} \quad (3.10)$$

$$\mathbf{Y} - \rho \mathbf{WY} = \mathbf{X}\beta + \mathbf{e} \quad (3.11)$$

$$(\mathbf{I} - \rho \mathbf{W})\mathbf{Y} = \mathbf{X}\beta + \mathbf{e} \quad (3.12)$$

$$\mathbf{Y} = (\mathbf{I} - \rho \mathbf{W})^{-1} \mathbf{X}\beta + (\mathbf{I} - \rho \mathbf{W})^{-1} \mathbf{e} \quad (3.13)$$

$$\equiv (\mathbf{I} - \rho \mathbf{W})^{-1} \mathbf{X}\beta + \mathbf{u} \quad (3.14)$$

elde edilir. Eğer $e = (e(1), \dots, e(n))^T \approx N(0, \Lambda)$ ise $u \approx MVN(\mathbf{0}, (\mathbf{I} - \rho \mathbf{W})^{-1} \Lambda (\mathbf{I} - \rho \mathbf{W}^T)^{-1})$ olacaktır. $\Lambda = \sigma^2 \mathbf{I}$ olarak da gösterilebilir (Haining, 2005, syf.314,315).

Öte yandan,

$$Z(s) = X(s)\beta + e(s) , \quad e(s) \approx (0, \Sigma(\theta)) \quad (3.15)$$

biçiminde tanımlandığında en küçük kareler kestirimleri,

$$\hat{\beta}_{EKK} = (X(s)'X(s))^{-1} X(s)'Z(s) \quad (3.16)$$

$$\hat{\sigma}^2 = \frac{1}{n - \text{rank}\{X(s)\}} (Z(s) - X(s)\hat{\beta}_{OLS})' (Z(s) - X(s)\hat{\beta}_{OLS}) \quad (3.17)$$

olur. Buradan, kestiricilere ilişkin özellikler,

$$E[\hat{\beta}_{EKK}] = \beta \quad (3.18)$$

$$E[\hat{e}_{EKK}(s)] = 0 \quad (3.19)$$

$$\text{Var}[\hat{\beta}_{EKK}] = \sigma^2 (X(s)'X(s))^{-1} \quad (3.20)$$

$$\text{Var}[\hat{e}_{EKK}(s)] = \sigma^2 \mathbf{M} \quad (3.21)$$

$$I'\hat{e}_{EKK}(s) = 0 \quad (3.22)$$

biçiminde yazılır. M matrisi, şapka matrisi (H) türünden de tanımlanabilir:

$\mathbf{H} = \mathbf{X}(s)(\mathbf{X}(s)'\mathbf{X}(s))^{-1} \mathbf{X}(s)'$ iken \mathbf{M} 'nin, $\mathbf{M} = \mathbf{I} - \mathbf{H}$ olur ve izdüşüm matrisi (simetrik ve idempotent) biçimde tanımlaması yapılabilir (Schabenger, 2005, syf.303-305).

Ayrıca, modelinin kullanıldığı durumlarda oluşturulacak hipotez $H_0 : \mathbf{L}\beta = \ell_0$ olacaktır. Burada \mathbf{L} , $\ell \times p$ boyutlu katsayılar matrisi ve ℓ_0 , $\ell \times 1$ boyutlu vektör olacaktır. Bu durumda ℓ vektöründeki i 'nci satır α_i katsayısını gösterecektir. Bu hipotez doğrultusunda kullanılacak F testi,

$$F = (\mathbf{L}\hat{\beta}_{EKK} - \ell_0)' [\mathbf{L}(\mathbf{X}(s)'\mathbf{X}(s))^{-1} \mathbf{L}']^{-1} (\mathbf{L}\hat{\beta}_{EKK} - \ell_0) / \hat{\sigma}^2 \text{rank}\{\mathbf{L}\} \quad (3.23)$$

olur. $F_{\alpha,1,v} = t_{\alpha/2,v}$ ifadesinden yola çıkarak,

$$t = \text{sign}(\hat{\beta}_j) \times \sqrt{F} \quad (3.24)$$

ve β_j 'ler için güven aralığı, kestirilmiş standart hataların $S(\hat{\beta}_j)$ olarak gösterilmesi ile,

$$\hat{\beta}_j \pm t_{\alpha/2, n - \text{rank}\{X(s)\}} S(\hat{\beta}_j) \quad (3.25)$$

olacaktır. Eğer artık kareler toplamı (SSR) dikkate alınarak F test istatistiği hesaplanacaksa,

$$SSR = \left(\mathbf{Z}(s) - \mathbf{X}(s)' \hat{\beta}_{EKK} \right)' \left(\mathbf{Z}(s) - \mathbf{X}(s)' \hat{\beta}_{EKK} \right) \quad (3.26)$$

eşitliği kullanılacaktır. Burada, SSR_T , $H_0: \mathbf{L}\beta = \ell_0$ doğrusal kısıtının kareler toplamı olmak üzere,

$$F = \frac{(SSR_T - SSR) / \text{rank}\{L\}}{\hat{\sigma}^2} \quad (3.27)$$

olarak hesaplanacaktır.

Klasik regresyon sürecinin bir sonraki aşamasında artıklar farklı yöntemlerle incelenebilirler. Bu bağlamda student türü artıklar için,

$$r_i = \frac{\hat{e}_{EKK}(s_i)}{\sqrt{\text{var}[\hat{e}_{EKK}(s_i)]}} = \frac{\hat{e}_{EKK}(s_i)}{\hat{\sigma} \sqrt{1 - h_{ii}}} \quad (3.28)$$

eşitliği kullanılır. Bununla birlikte gözlem uzaklığı (Leverage) ölçütü (h_{ii}) yardımıyla X düzlemindeki farklı olan gözlemler de belirlenebilir. Bunun için uygun aralık $1/n \leq h_{ii} \leq 1$ olarak belirlenmiştir.

PRESS (prediction sum of squares) ise,

$$\hat{e}(s_i)_{-i} = Z(s_i) - \hat{Z}(s_i)_{-i} = \frac{\hat{e}(s_i)}{1 - h_{ii}} \quad (3.29)$$

eşitliği ile incelenir. Bunların yanı sıra artıkların incelenmesinde kullanılacak olan alışlagelen bir uzaklık da,

$$D = \frac{r_i^2 h_{ii}}{k(1 - h_{ii})} \quad (3.30)$$

ile gösterilen Cook uzaklığıdır. Burada $k = \text{rank}\{X(s)\}$ olarak gösterilecektir. Başka bir istatistik ise,

$$DFFits_i = t_i \sqrt{\frac{h_{ii}}{1-h_{ii}}} \quad (3.31)$$

ile gösterilen DFFits istatistiğidir (Schabenberger, 2005, syf.305-307).

Artıkların incelenmesi modellerdeki varsayımların bozulmalarının belirlenmesinde oldukça önemli bir süreçtir. Örneğin $\hat{e}_{EKK}(s)$ için çizilen QQ çizimleri veya histogramlar normal dağılım varsayımının kontrolünde, saçılım grafikleri de sabit varyans varsayımının incelenmesinde kullanılmaktadır.

EKK'dan elde edilen satır artıkları,

$$\hat{e}_{EKK} = \mathbf{Z}(s) - \hat{\mathbf{Z}}(s) = \mathbf{Z}(s) - \mathbf{HZ}(s) = \mathbf{MZ}(s) \quad (3.32)$$

olarak gösterilir. Burada $\mathbf{H} = \mathbf{X}(s)(\mathbf{X}(s)' \mathbf{X}(s))^{-1} \mathbf{X}(s)'$ olarak tanımlanan şapka matrisi olacaktır. Beklenen durum $\mathbf{e}(s)$ 'lerin sıfır ortalamalı ve eşvaryanslı olmalarıdır. Bu durumların dışında özelliklere sahip olan modellerin kullanılması uygun olmayacaktır.

EKK ile elde edilen artık değerlerine bağlı olarak hesaplanacak Moran istatistik değeri,

$$I_{artik} = \frac{n \hat{e}_{EKK}(s)' \mathbf{W} \hat{e}_{EKK}(s)}{w_{..} \hat{e}_{EKK}(s)' \hat{e}_{EKK}(s)} \quad (3.33)$$

olur. Burada, $w_{..} = \mathbf{1}' \mathbf{W} \mathbf{1} = \sum_{i=1}^n \sum_{j=1}^n w_{ij}$ 'dir; \mathbf{W} da bağlanabilirlik (connectivity) matrisidir.

I_{artik} aynı zamanda, daha önce tanımlanan $\mathbf{M} = \mathbf{I} - \mathbf{H}$ iken, aşağıdaki gibi de yazılabilir:

$$I_{artik} = \frac{n \mathbf{e}(s)' \mathbf{M} \mathbf{W} \mathbf{M} \mathbf{e}(s)}{w_{..} \mathbf{e}(s)' \mathbf{M} \mathbf{e}(s)} \quad (3.34)$$

Ayrıca, $k = \text{rank}\{\mathbf{X}(s)\}$ iken, I_{artik} asimptotik olarak normal dağılıma sahiptir ve ortalaması,

$$E[I_{artik}] = - \left[\frac{n}{(n-k)w_{..}} \right] iz \{ (\mathbf{X}(s)' \mathbf{X}(s))^{-1} \mathbf{X}(s)' \mathbf{W} \mathbf{X}(s) \} \quad (3.35)$$

$$= \frac{n}{(n-k)w..} iz\{\mathbf{MW}\} \quad (3.36)$$

ve varyansı,

$$Var[I_{artik}] = \left\{ \frac{n^2}{w..^2 (n-k)(n-k+2)} \right\} \times \left\{ S_1 + 2iz\{\sigma^2\} - iz\{F\} - \frac{2[iz\{G\}]^2}{(n-k)} \right\} \quad (3.37)$$

biçiminde verilir. Burada,

$$S_1 = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (w_{ij} + w_{ji})^2 \quad (3.38)$$

$$F = (\mathbf{X}(s)' \mathbf{X}(s))^{-1} \mathbf{X}(s)' (\mathbf{W} + \mathbf{W}')^2 \mathbf{X}(s) \quad (3.39)$$

$$G = (\mathbf{X}(s)' \mathbf{X}(s))^{-1} \mathbf{X}(s)' \mathbf{W} \mathbf{X}(s) \quad (3.40)$$

olarak tanımlanmıştır. Normal dağılıma sahip olduğu varsayımı dikkate alındığında, uzamsal otokorelasyonun olmadığını gösteren sıfır hipotezi,

$$Z = \frac{I_{artik} - E[I_{artik}]}{\sqrt{Var[I_{artik}]}} \quad (3.41)$$

eşitliği ile test edilir (Schabenberger, 2005, syf.314-315).

3.1.4. Genelleştirilmiş Doğrusal Modeller

Lojistik, Binom ve Poisson modeller bağımsız hatalarla birlikte genelleştirilmiş doğrusal modeller sınıfının birer elemanlarıdır.

$\mu(s)$, $Y(s)$ 'nin beklenen değeri iken;

$$\text{Lojistik için } \mu(s) = prob\{Y(s) = 1\} = \theta(s) = \exp(\alpha(s)) / (1 + \exp(\alpha(s))) \quad (3.42)$$

$$\text{Binom için } \mu(s)/c(s) = \theta(s) = \exp(\alpha(s)) / (1 + \exp(\alpha(s))) \quad (3.43)$$

$$\text{Poisson için } \mu(s) = \lambda(s) = \exp(\alpha(s)) \quad (3.44)$$

olur. Binom modeldeki $c(s)$ ifadesi bilinen deneme sayısıdır. Poisson model için link fonksiyonu logaritmik bir fonksiyon, Binom ve lojistik modeller içinse lojit fonksiyondur ve $\alpha(s)$ 'ye "log olabilirlik" denilir.

Bir örnekle açıklanır: $O(s)$, $s=1, \dots, n$ bölgelerindeki hasta sayısını gösterebilir. Bu durumda $O(s)$ 'nin, $\lambda(s) = E(s)r(s)$ olmak üzere Poisson dağıldığı söylenebilir. $E(s)$, beklenen hasta sayısını; $r(s)$ 'nin ise göreceli riski belirtir. Bu durumda Poisson modeli,

$$\log(E[O(s)]) = \log(\lambda(s)) = \log(E(s)) + \log[r(s)], \quad s=1, \dots, n \quad (3.45)$$

$$= \log(E(s)) + \beta_0 + \beta_1 X_1(s) + \beta_2 X_2(s) + \dots + \beta_k X_k(s) \quad (3.46)$$

olur. $\log(E(s))$, "offset" değişken adını alacaktır. Eğer $\{\log(E(s))\}$ merkezleştirilmiş ve $\sum_j O(i) \neq \sum_i E(i)$ ise ya da ortak değişkenler (covariates) merkezleştirilmemişse β_0 yerini korur. Poisson modelde ortalama ve varyansın birbirine eşit olması gerekmektedir. Ancak uzamsal heterojenlik yüzünden bu değerler farklı çıkmaktadır. Bu nedenle aşırı yayılımın (overdispersion) bulunması istatistiksel anlamlılığın kestirilmesine yardımcı olur. Aşırı yayılımın olması demek gerçek I. tip hata değerlerinin, genelleştirilmiş doğrusal model yardımıyla hesaplanan değerden daha büyük çıkması anlamına gelecektir. Bayesyen modeller aşırı yayılım etkisini gidererek modeli uzamsal ve uzamsal olmayan bileşenlerine ayırabilir. Aşırı yayılımın olması durumunda, uygun paket programın kullanımı ile, iteratif yeniden ağırlıklandırılmış en küçük kareler yöntemiyle genelleştirilmiş doğrusal modeller kestirilebilirler.

Lojistik ve binom modeller söz konusu olduğunda, $\exp(\beta_j)$, X_j 'ye uygun olabilirlik oranını (odds ratio) göstermektedir. X_j 'deki 1 birim artış için olabilirlik oranının hesaplanması,

$$\exp(\beta_j) = \frac{[\theta(s)/(1-\theta(s))]}{[\theta^*(s)/(1-\theta^*(s))]} \quad (3.47)$$

şeklinde olur. $\theta(s)$, $X_j = x_j + 1$ iken $Y(s) = 1$ olması olasılığını ve $\theta^*(s)$, $X_j = x_j$ iken $Y(s) = 1$ olması olasılığını gösterir. Örneğin, $\beta_j = 0,2231$ ise $\exp(\beta_j) = 1,25$ olur. Bu da $Y=0$ 'dan $Y=1$ olma olasılığının X_j 'nin her birimin artmasında %25 artacağı anlamına gelmektedir (Haining, 2003. syf.312-323).

3.1.5. Coğrafi Olarak Ağırlıklandırılmış Regresyon Yöntemi

Coğrafi olarak ağırlıklandırılmış regresyon yönteminin (Geographically Weighted Regression - GWR) açıklanması için, ilk adımda genel (global) regresyon modelinin incelenmesi gerekmektedir. Genel regresyon modeli,

$$y_i = \beta_0 + \sum_k \beta_k x_{ik} + \varepsilon_i \quad (3.48)$$

şeklinde olur. Model bölgesel (local) şekle dönüştürüldüğünde,

$$y_i = \beta_0(u_i, v_i) + \sum_k \beta_k(u_i, v_i) x_{ik} + \varepsilon_i \quad (3.49)$$

dur. Burada (u_i, v_i) , uzayda i noktasının koordinatlarıdır ve $\beta_k(u_i, v_i)$, i noktasında $\beta_k(u, v)$ sürekli fonksiyonunun uyarlanmış biçimidir. Böylece parametrelerin sürekli değerler almasını ve belli noktalardaki uzamsal değişimin gösterilmesini sağlamaktadır. Katsayı kestirimi,

$$\hat{\beta}(u_i, v_i) = (\mathbf{X}^T \mathbf{W}(u_i, v_i) \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}(u_i, v_i) \mathbf{y} \quad (3.50)$$

olarak verilir. Burada $\mathbf{W}(u_i, v_i)$, köşegen elemanları dışındakiler 0 olan, köşegen elemanları ise i regresyon noktasına ait coğrafi ağırlıkları temsil eden $n \times n$ boyutlu matris olur.

Tablo 3.2. Klasik Regresyon Modeli ile GWR Modelin Karşılaştırılması

Klasik regresyon modeli	GWR modeli
$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$	$\mathbf{Y} = (\boldsymbol{\beta} \otimes \mathbf{X})\mathbf{1} + \boldsymbol{\epsilon}$
$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}$	$\hat{\boldsymbol{\beta}}(i) = (\mathbf{X}^T \mathbf{W}(i) \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}(i) \mathbf{Y}$

Klasik regresyon modeli ile coğrafi olarak ağırlıklandırılmış regresyon modellerinin yapısal anlamda birbirleri ile karşılaştırılmaları Tablo 3.2'de verilmiştir. \otimes simgesi (logical multiplication operator), $\boldsymbol{\beta}$ 'nin her bir elemanının bağlantılı \mathbf{X} elemanı ile çarpılması anlamına gelmektedir. n veri noktası ve k açıklayıcı değişken varsa, $\boldsymbol{\beta}$ ve \mathbf{X} $n \times (k+1)$ boyutlu ve $\mathbf{1}$, $(k+1) \times 1$ boyutlu $\mathbf{1}$ 'ler vektörü olur.

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0(u_1, v_1) & \beta_1(u_1, v_1) & \dots & \beta_k(u_1, v_1) \\ \beta_0(u_2, v_2) & \beta_1(u_2, v_2) & \dots & \beta_k(u_2, v_2) \\ \dots & \dots & \dots & \dots \\ \beta_0(u_n, v_n) & \beta_1(u_n, v_n) & \dots & \beta_k(u_n, v_n) \end{bmatrix} \quad (3.51)$$

olarak bulunur ve ağırlık matrisi,

$$\mathbf{W}(i) = \begin{bmatrix} w_{i1} & 0 & \dots & 0 \\ 0 & w_{i2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & w_{in} \end{bmatrix} \quad (3.52)$$

şeklinde oluşturulur (Fortheringham, 2006, syf.52-54).

Bölgesel model varyansları da aşağıdaki gibi verilir.

$$\hat{\beta}(u_i, v_i) = \mathbf{C}\mathbf{y} \quad (3.53)$$

biçimine getirilebilir. Bu eşitlikte,

$$\mathbf{C} = (\mathbf{X}^T \mathbf{W}(u_i, v_i) \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}(u_i, v_i) \quad (3.54)$$

olacaktır. Kestirime ilişkin varyans,

$$\text{Var}[\hat{\beta}(u_i, v_i)] = \mathbf{C}\mathbf{C}^T \sigma^2 \quad (3.55)$$

ve

$$\sigma^2 = \sum_i (y_i - \hat{y}_i) / (n - 2\gamma_1 + \gamma_2) \quad (3.56)$$

şeklinde hesaplanır. Burada $\gamma_1 = \text{iz}(\mathbf{H})$ ve $\gamma_2 = \text{iz}(\mathbf{H}^T \mathbf{H})$ olduğu bilinmektedir. \mathbf{H} , şapka matris olarak bilinir ve $\hat{\mathbf{y}} = \mathbf{H}\mathbf{y}$ olur. \mathbf{H} 'nin her bir satırı,

$$\mathbf{r}_i = \mathbf{X}_i (\mathbf{X}^T \mathbf{W}(u_i, v_i) \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}(u_i, v_i) \quad (3.57)$$

olarak hesaplanır. Artıkların serbestlik derecesi $n - 2\gamma_1 + \gamma_2$ olur. $2\gamma_1 - \gamma_2$ değeri de genel doğrusal regresyon modelindeki parametre sayısını verir (Fortheringham, 2006, syf.54-56).

Coğrafi olarak ağırlıklandırılmış regresyon modellerinde Akaike bilgi kriteri (AIC) aşağıdaki gibi hesaplanır.

$$AIC_C = 2n \log(\hat{\sigma}) + n \log_e(2\pi) + n \left\{ \frac{n + \text{iz}(\mathbf{H})}{n - 2 - \text{iz}(\mathbf{H})} \right\} \quad (3.58)$$

İndis olarak gösterilen C, düzeltilmiş (corrected) AIC kestirimi olduğunu gösterir. N, örneklem büyüklüğünü; $\hat{\sigma}$, hata terimlerinin kestirilmiş standart sapmasını; $iz(\mathbf{H})$ de şapka matrisinin izini göstermektedir. Matris izi bant genişliği yardımıyla hesaplanmıştır. AIC değeri doğrusal modellerin yanı sıra Poisson ve lojistik GWR'de de kullanılabilir. Daha basit olarak,

$$AIC = 2n \log_e(\hat{\sigma}) + n \log_e(2\pi) + n + iz(\mathbf{H}) \quad (3.59)$$

biçiminde gösterilir. Ancak modellerin karşılaştırılmaları AIC_c ile AIC değerleri arasında olmamalıdır. Her iki eşitlikte de yer alan $\hat{\sigma}$ değeri,

$$\hat{\sigma}^2 = \frac{AKT}{n} \quad (3.60)$$

ile elde edilmektedir. Buradaki AKT ifadesi (artık kareler toplamı),

$$AKT = \mathbf{y}^T (\mathbf{I} - \mathbf{S})^T (\mathbf{I} - \mathbf{S}) \mathbf{y} \quad (3.61)$$

ile hesaplanır.

Bunun yanı sıra bant genişliği yardımıyla Bayesçi bilgi kriteri (BIC-veya Schwarz kriteri SIC),

$$BIC = -2\log_e(L) + k\log_e(n) \quad (3.62)$$

biçiminde verilir. L, model olabirliğini; k, parametre sayısını ve n örneklem sayısını göstermektedir (Fortheringham, 2006, syf.59-62).

3.1.6. Coğrafi Olarak Ağırlıklandırılmış Regresyon Yönteminde Uzamsal Ağırlık Fonksiyonunun Seçimi

Birinci bölümde açıklanmış olan komşuluk matrisi ve bölgelerin ortak sınırlarının olması mantığı dikkate alındığında ağırlık değerleri,

$$w_{ij} = 1, \quad \text{eğer } d_{ij} < d$$

$$w_{ij} = 0, \quad \text{diğer durumlarda}$$

olur. Ancak uzamsal ağırlık fonksiyonunda süreksizlik (kesiklilik) problemi karşımıza çıkar. Regresyon noktasının değişmesiyle kestirilmiş katsayılar da değişir.

Ağırlıkların kesiklilik sorununu ortadan kaldırabilmek için kullanılan bir yol, d_{ij} i ve j arasındaki uzaklığı göstermek üzere, w_{ij} 'yi d_{ij} 'nin sürekli bir fonksiyonu haline getirmektir. Bunun yöntemlerinden biri,

$$w_{ij} = \exp\left[-1/2(d_{ij}/h)^2\right] \quad (3.63)$$

olarak ağırlık fonksiyonunu düzenlemektir. h bant genişliğini (bandwidth) gösterir.

Alternatif olarak çekirdek (kernel) uygulaması,

$$w_{ij} = \left[1 - (d_{ij}/h)^2\right]^2, \quad \text{eğer } d_{ij} < b \quad (3.64)$$

$$= 0, \quad \text{diğer durumlarda}$$

Farklı bir yol olarak da her bir j noktasının i noktasına uzaklıkları yardımıyla elde edilen rank değerleri R_{ij} ile gösterilirse ağırlıklar,

$$w_{ij} = \exp(-R_{ij}/h) \quad (3.65)$$

ile hesaplanabilir.

Diğer bir yol olarak da herhangi i noktası için ağırlıkların toplamını C gibi bir sabite eşitlemek olabilir:

$$\sum_j w_{ij} = C, \quad \text{tüm i'ler için} \quad (3.66)$$

C'nin optimum değerini belirlerken uygulanacak adımlar aşağıdaki gibi özetlenebilir:

1. Adım: Herhangi bir C değeri belirlenir.
2. Adım: Ağırlık fonksiyonu, seçilen C değeri için kalibre edilir.
3. Adım: Farklı bir C değeri belirlenir.
4. Adım: 2. adım tekrarlanır.

Bu süreç optimal uyum sağlanana kadar tekrarlanır (Fortheringham, 2006, syf.56-59).

3.1.7. Karma Coğrafi Olarak Ağırlıklandırılmış Regresyon Modeli

Bazı durumlarda analiz edilecek tüm değişkenler uzamsal yapı gösterebildiği gibi, değişkenlerin bir kısmı farklı yapı gösterebilirler. Örneğin ev fiyatlarına ilişkin yapılan bir çalışmada yatak odası sayısı coğrafi bir etki gösterirken; bölgedeki işsizlik düzeyi

gibi sosyoekonomik deęişkenlerin her yerde etkileri aynı olabilir. Bu durumda karma modeller oluşturulur. Karma modeller ařaęıdaki gibi yazılabilir:

$$y_i = \sum_{j=1, k_a} a_j x_{ij}(a) + \sum_{l=1, k_b} b_l(u_i, v_i) x_{il}(b) + \epsilon_i \quad (3.67)$$

Burada i gözlemi için, y_i , baęımlı deęişken deęeri ve (u_i, v_i) , konumu (lokasyon) gösterirken; $\{a_1, \dots, a_{k_a}\}$, k_a tane genel (global) katsayı ve $\{b_1(u,v), \dots, b_{k_b}(u,v)\}$ ise k_b tane bölgesel (lokal) katsayılar fonksiyonudur. $\{x_{i1}(a), \dots, x_{ik_b}(a)\}$ bölgesel katsayılara ait baęımsız deęişkenlerdir. İki grup baęımsız deęişkenler a-grubu ve b-grubu olarak adlandırılabilir. Ancak her iki grup için ayrı ayrı sabitler (constant) olacağı düşünülmemelidir. Eęer a-grubu olmazsa basit GWR, b-grubu olmazsa standart doğrusal regresyon elde edilir. Karma modelin çözümlü için model ilk adımda vektör-matris gösterimiyle yazılır:

$$\mathbf{y} = \mathbf{X}_a \mathbf{a} + \mathbf{m} + \boldsymbol{\epsilon} \quad (3.68)$$

Burada \mathbf{y} , baęımlı deęişken vektörü; \mathbf{X}_a , a-grubu deęişkenlerinin matrisi; \mathbf{a} , a-grubu katsayılar vektörü; $\boldsymbol{\epsilon}$, hata terimleri vektörü ve \mathbf{m} , i 'nci elemanı $\sum_{l=1, k_b} b_l(u_i, v_i) x_{il}(b)$ olan bir vektördür. \mathbf{m} 'nin, coęrafi olarak aęırlıklandırılmış bir terim olduęu belirtilir. (3.68) eřitlięinden,

$$\mathbf{y} - \mathbf{X}_a \mathbf{a} = \mathbf{m} + \boldsymbol{\epsilon} \quad (3.69)$$

yazılır ve \mathbf{a} 'nın bilindięi varsayımı altında \mathbf{m} 'yi kestirmek için basit GWR kullanılır. \hat{y} 'nin kestirilmesinde daha önce, $\hat{y} = \mathbf{H}\mathbf{y}$ kullanılmış ve \mathbf{H} , řapka (hat) matrisi olarak tanımlanmıştı. Buradan,

$$\hat{\mathbf{m}} = \mathbf{H}(\mathbf{y} - \mathbf{X}_a \mathbf{a}) \quad (3.70)$$

olarak kestirilir. Bu eřitlięe baęlı olarak gerekli düzenlemeler yapılırsa,

$$(\mathbf{I}-\mathbf{H}) \mathbf{y} = (\mathbf{I}-\mathbf{H}) \mathbf{X}_a \mathbf{a} + \boldsymbol{\epsilon} \quad (3.71)$$

olur. Eęer $\mathbf{z} = (\mathbf{I}-\mathbf{H}) \mathbf{y}$ ve $\mathbf{Q} = (\mathbf{I}-\mathbf{H}) \mathbf{X}_a$ olursa,

$$\mathbf{z} = \mathbf{Q}\mathbf{a} + \boldsymbol{\epsilon} \quad (3.72)$$

olur. Bu sıradan bir regresyon problemidir (EKK) ve standart yöntemler kullanılarak \mathbf{a} kestirilir:

$$\hat{a} = (\mathbf{Q}^T \mathbf{Q})^{-1} \mathbf{Q}^T \mathbf{z} \quad (3.73)$$

$$= (\mathbf{X}_a^T (\mathbf{I}-\mathbf{H})^T (\mathbf{I}-\mathbf{H}) \mathbf{X}_a)^{-1} \mathbf{X}_a^T (\mathbf{I}-\mathbf{H})^T (\mathbf{I}-\mathbf{H}) \mathbf{y} \quad (3.74)$$

\hat{a} 'nın elde edilmesinden sonra (3.67) numaralı eşitlikte bu terim sol taraftaki ifadelerden çıkarılırsa b-grubu katsayılar fonksiyonu bulunur ve bunların kestirilmesinde de basit GWR kullanılabilir. Bunu elde etmede kullanılacak algoritma aşağıdaki gibi olur:

1. Adım: \mathbf{X}_a 'nın her bir sütunu için,
 - (i) Basit GWR kullanılarak sütun, \mathbf{X}_b ile regresyon uygulanır.
 - (ii) Bu regresyon ile artıklar hesaplanır.
2. Adım: Basit GWR kullanılarak \mathbf{y} ile \mathbf{X}_b için regresyon uygulanır.
3. Adım: Bu regresyonun artıkları hesaplanır.
4. Adım: \mathbf{y} -artıkları ve \mathbf{X}_a -artıkları için EKK kullanılarak regresyon uygulanır. Bu işlem \hat{a} kestirimini verir.
5. Adım: $\mathbf{X}_a \hat{a}$, \mathbf{y} 'den çıkarılır. Elde edilen değer yine basit GWR kullanılarak \mathbf{X}_b ile regresyona alınır ve coğrafi dağılım (b-grubu) katsayıları kestirilir.

Bu modellerde bilinen $\mathbf{W} = (\mathbf{I}-\mathbf{H})^T (\mathbf{I}-\mathbf{H})$ eşitliğinin,

$$\hat{a} = (\mathbf{X}_a \mathbf{W} \mathbf{X}_a)^{-1} \mathbf{X}_a^T \mathbf{W} \mathbf{y} \quad (3.75)$$

şekline geleceği bilinmektedir. \mathbf{W} , pozitif yarı tanımlı simetrik matris olacaktır. Diğer bir durum da karma GWR için şapka matrisinin belirlenmesidir. (3.68) numaralı modelde \hat{m} ve \hat{a} yerine konulursa,

$$\hat{\mathbf{y}} = \mathbf{X}_a (\mathbf{X}_a \mathbf{W} \mathbf{X}_a)^{-1} \mathbf{X}_a^T \mathbf{W} \mathbf{y} + \mathbf{H} \mathbf{y} = \mathbf{H}^* \mathbf{y} \quad (3.76)$$

ve

$$\mathbf{H}^* = \mathbf{X}_a (\mathbf{X}_a \mathbf{W} \mathbf{X}_a)^{-1} \mathbf{X}_a^T \mathbf{W} + \mathbf{H} \quad (3.77)$$

olur. Eğer k_a tane a-grubu değişken varsa, her k_b tane değişken için k_a+2 tane basit GWR kalibrasyonu yapılacaktır (Fortheringham, 2006, syf.65-68).

3.1.8. Coğrafi Olarak Ağırlıklandırılmış Regresyon Modelinde Aykırı Değerler

GWR'nin kullanılması planlanan çalışmalarda, aykırı değerler, bölgesel parametre kestirimlerinin gerçek değerlerinden uzaklaşmalarına neden olacaktır. Aykırı değerlerin belirlenmesinde $e_i = y_i - \hat{y}_i$ değerlerinin önemi büyük olacaktır. GWR için oluşturulan modeller doğrultusunda,

$$\hat{y} = \mathbf{Hy} \quad (3.78)$$

ve böylece,

$$\mathbf{e} = \mathbf{y} - \mathbf{Hy} = (\mathbf{I} - \mathbf{H})\mathbf{y} \quad (3.79)$$

olur. Burada \mathbf{e} , e_i 'lerin vektörü olacaktır.

$$\text{Var}(\mathbf{e}) = (\mathbf{I} - \mathbf{H})(\mathbf{I} - \mathbf{H})^T \text{var}(\mathbf{y}) = (\mathbf{I} - \mathbf{H})(\mathbf{I} - \mathbf{H})^T \sigma^2 \quad (3.80)$$

ile elde edilir. Student türü hatalar da,

$$r_i = \frac{e_i}{\hat{\sigma} \sqrt{q_{ii}}} \quad (3.81)$$

ile verilir. Burada q_{ii} , $\mathbf{Q} = (\mathbf{I} - \mathbf{H})(\mathbf{I} - \mathbf{H})^T$ ifadesinin i 'nci köşegen ögesidir. $\hat{\sigma}^2$ yerine $\hat{\sigma}_{-i}^2$ (kalibre edilmiş verilerden i 'nci gözlem için elde edilen σ^2 kestirimi) kullanılırsa,

$$\hat{r}_i^* = \frac{e_i}{\hat{\sigma}_{-i} \sqrt{q_{ii}}} \quad (3.82)$$

bulunur. Burada amaç, i gözlemi aykırı ise, i 'nin çıkarılması ile elde edilen değerlerin incelenmesidir. r_i terimi "içsel student türü hatalar" ve r_i^* terimi "dışsal student türü hatalar" adını alır. $|r_i^*|$ değerinin 3'den büyük çıkması gözlemin aykırı olması anlamına gelebilir (Fortheringham, 2006, syf.73-80).

Birim parametrelerinin durağanlığının incelenmesinde, q parametresinin varyansı V_q olarak gösterilirse,

$$V_q = \frac{1}{n} \sum_{i=1}^n \left(\hat{\beta}_{iq} - \frac{1}{n} \sum_{i=1}^n \hat{\beta}_{iq} \right)^2 \quad (3.83)$$

olarak hesaplanır. Analitik olarak dağılımın hesaplanması zor olduğunda durağanlığın incelenmesinde Monte Carlo yöntemi uygulanır. Bu uygulama birimlerin katsayılarının anlamlılık testlerinin yapılmasına olanak sağlamaktadır. Burada yapılacak uygulamalarda sıfır hipotezi katsayıların durağan olduğunu gösterecek şekilde düzenlenmiştir. Bunun için ilk önce $\hat{\beta}_q$ sütun vektörünün oluşturulması gerekmektedir. Bu vektörün elemanları $[\hat{\beta}_{1q}, \dots, \hat{\beta}_{nq}]$ olur ve

$$V_q = \frac{1}{n} \hat{\beta}_q^T [\mathbf{I} - (1/n)\mathbf{J}]^T [\mathbf{I} - (1/n)\mathbf{J}] \hat{\beta}_q \quad (3.84)$$

elde edilir. Burada \mathbf{I} , nxn boyutlu kimlik matris; \mathbf{J} , tüm elemanları 1'e eşit olan nxn boyutlu matristir. $(1/n)\mathbf{J} \hat{\beta}_q$, $\hat{\beta}_q$ 'nin ortalama vektörünü göstermektedir. Ayrıca idempotentlikten ötürü $(\mathbf{I} - (1/n)\mathbf{J})^T (\mathbf{I} - (1/n)\mathbf{J}) = (\mathbf{I} - (1/n)\mathbf{J})$ olacağından eşitlik,

$$V_q = \frac{1}{n} \hat{\beta}_q^T [\mathbf{I} - (1/n)\mathbf{J}]^T \hat{\beta}_q \quad (3.85)$$

Biçiminde yazılabilir. $\hat{\beta}_q$, bağımlı değişkenin doğrusal dönüşümü iken, $\hat{\beta}_q = \mathbf{B}\mathbf{y}$ olur ve

$$\mathbf{B} = \begin{bmatrix} e_q^T [X^T W_1 X]^{-1} X^T W_1 \\ \vdots \\ e_q^T [X^T W_n X]^{-1} X^T W_n \end{bmatrix} \quad (3.86)$$

yazılır. Burada, $\mathbf{W}_1, \dots, \mathbf{W}_n$, göze konumları etrafında "çekirdek merkezleştirilmiş köşegen ağırlık matrisleri" olarak bilinir. Ayrıca \mathbf{e}_q , q'ncü elemanı 1, gerisi 0 olan bir sütun vektördür. Böylece,

$$V_q = \mathbf{y}^T (1/n) \mathbf{B}^T (\mathbf{I} - (1/n) \mathbf{J}) \mathbf{B} \mathbf{y} \quad (3.87)$$

olur. Buradan yola çıkarak,

$$E(V_q) = \text{iz} ((1/n) \mathbf{B}^T (\mathbf{I} - (1/n) \mathbf{J}) \mathbf{B}) \sigma^2 \quad (3.88)$$

elde edilir. $E(V_q) / \sigma^2$, yaklaşık olarak, $\text{iz}((1/n) \mathbf{B}^T (\mathbf{I} - (1/n) \mathbf{J}) \mathbf{B})$ serbestlik derecesiyle ki-kare dağılır. Yaklaşık ki-kare değeri, σ^2 'nin kestirimine bölünürse, $n - 2\text{iz}(\mathbf{H}) - \text{iz}(\mathbf{H}^T \mathbf{H})$ serbestlik dereceli bir ki-kare değeri daha elde edilir. Bu iki ki-kare değerinin oranı $\text{iz}((1/n) \mathbf{B}^T (\mathbf{I} - (1/n) \mathbf{J}) \mathbf{B})$ ve $n - 2\text{iz}(\mathbf{H}) - \text{iz}(\mathbf{H}^T \mathbf{H})$ serbestlik dereceli F

değerini verecektir. Bu F testi q'uncu parametrenin durağan olup olmadığını gösterecektir (Fortheringham, 2006, syf.92-94).

3.1.9. Uzamsal Yapıda Oluşturulan Modeller

Bu bölümde artıklarının korelasyonsuz olduğu belirlenen ancak uzamsal bir yapıya sahip modeller açıklanmıştır.

3.1.9.1. Komşuluk Düzeltmeli Modeller – Papadakis Analizi

Literatürde bu tarz modellere, tarım ile ilgili olarak araziyi farklı parsellere bölerek çalışılan verilerin incelendiği durumlarda karşılaşılmaktadır. Burada kullanılan değişkenler komşu bölgelerdeki artıkların fonksiyonları olarak oluşturulur. Bu süreç Papadakis modeli (1937) adı verilen bir süreç ile açıklanabilir.

İki boyutlu bir düzende r satırları ve c sütunları gösterebiliriz. rc sayıda deneysel birimin b sayıda blokta gruplanmış oldukları varsayalım. Bu bloklarda gerçekleşen deneyler T_1, \dots, T_t olarak gösterilsin. Bu durumda her bir blokta n_i kadar T_i deney gerçekleştirilecektir. Bu deneysel grubun modeli,

$$Z(i, j)_{kl} = \mu + \rho_k + \tau_l + e_{kl} \quad (3.89)$$

$$i = 1, \dots, r, \quad j = 1, \dots, c$$

$$k = 1, \dots, b, \quad l = 1, \dots, t$$

biçiminde gösterilir. ρ_k , k'inci bloktaki etkiyi; τ_l , l'inci denemedeki etkiyi gösterecektir.

Papadakis analizinde yukarıdaki modellerdeki blok etkileri yine bu modeldeki EKK artıklarının fonksiyonları ile yer değiştirilir. Bu süreç,

$$\hat{e}(i, j) = Z(i, j)_{kl} - \hat{E}[Z(i, j)_{kl}] = Z(i, j)_{kl} - \bar{Z}_k \quad (3.90)$$

ile başlar. Buradaki \bar{Z}_k , i'inci satır ve j'inci sütundaki k denemenin ortalamasını göstermektedir.

Artıklar ve uzamsal komşuluklar dikkate alınarak değişkenler,

$$x_1(i, j) = \frac{1}{2}(\hat{e}(i, j-1) + \hat{e}(i, j+1)) \quad (3.91)$$

$$x_2(i, j) = \frac{1}{2}(\hat{e}(i-1, j) + \hat{e}(i+1, j)) \quad (3.92)$$

şeklinde oluşturulur ve model,

$$Z(i, j) = \beta_0 + \tau_i + \beta_1 x_1(i, j) + \beta_2 x_2(i, j) + e_{kl}^* \quad (3.93)$$

eşitliği ile gösterilir. Bu bilgiler ışığında komşuluk kavramının da dahil olmasındaki süreci açıklayan model, matris formunda, $Z(s) = X\tau + e$, $\hat{e} = MZ(s)$, $M = I - X(X'X)^{-1}X'$ ve $\hat{\tau}_{EKK} = (X'X)^{-1}X'Z(s)$ ifadeleri yardımıyla,

$$Z(s) = X\tau + \beta A \hat{e} + e^* \quad (3.94)$$

$$= X\tau + \beta AMZ(s) + e^* \quad (3.95)$$

$$= X\tau + \beta A(Z(s) - X\hat{\tau}_{OLS}) + e^* \quad (3.96)$$

olarak gösterilir. Modeldeki A ifadesi, artıkların ortalama matrisini gösterir.

3.1.9.2. Birinci Farklar Modeli

Hataların korelasyonsuz olduğu durumlarda Besag ve Kempton (1986) tarafından deney düzenleme ile oluşturulan model de kullanılabilir. Bu modelin satır sayısının sütun sayısından çok daha fazla olduğu durumlarda kullanılması tercih edilmektedir. Tek sütunun olduğu durumda birinci farklar modeli,

$$Z(s) = X\tau + e(s), \quad e(s) \sim (0, \sigma^2 \Sigma) \quad (3.97)$$

olacaktır. Z(i) için $\text{Var}[Z(i) - Z(i+1)] = \sigma^2$ ve $\text{Kov}[Z(i) - Z(i+1), Z(k) - Z(k+1)] = 0$, $i \neq k$, olduğu da varsayılır. Diğer bir ifade ile aynı sütundaki gözlemlerin birinci farkları eş varyanslı ve ilişkisiz olmalıdır.

(3.97) ile gösterilen modelin her iki tarafı aşağıdaki matris ile çarpılırsa birinci farklar modeli elde edilir:

$$\Delta = \begin{bmatrix} 1 & -1 & 0 & \dots & 0 \\ 0 & 1 & -1 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 1 & -1 \end{bmatrix} \quad (3.98)$$

Elde edilen model bir EKK modeli olur.

$$\Delta Z(s) = \Delta X\tau + \Delta e(s) \quad (3.99)$$

$$Z(s)^* = X^* \tau + e(s)^* , \quad e(s)^* \approx (0, \sigma^2 I) \quad (3.100)$$

Buradaki deneme etkileri ve kestirimleri,

$$\hat{\tau} = (X^{*'} X^*)^{-1} X^{*'} Z(s)^* \quad (3.101)$$

$$Var[\hat{\tau}] = \hat{\sigma}^2 (X^{*'} X^*)^{-1} \quad (3.102)$$

$$\hat{\sigma}^2 = \frac{1}{n - rank\{X\}} (Z(s)^* - X^* \hat{\tau})' (Z(s)^* - X^* \hat{\tau}) \quad (3.103)$$

şeklinde elde edilir.

İlişkili artıkların olduğu bilindiğinde genel doğrusal ortalama fonksiyonu,

$$Z(s) = X(s) \beta + e(s) , \quad e(s) \sim (0, \sigma^2 \Sigma) \quad (3.104)$$

olur. Σ bilindiğinde β 'nin genelleştirilmiş en küçük kareler kestirimi, dönüştürülmüş modele klasik en küçük kareler yönteminin uygulanmasıyla elde edilir. Bir L matrisi oluşturulur ve $LL' = \Sigma$ olarak tanımlanırsa,

$$L^{-1}Z(s) = L^{-1}X(s)\beta + L^{-1}e(s) \quad (3.105)$$

$$Z(s)^* = X(s)^* \beta + e(s)^* , \quad e(s)^* \approx (0, \sigma^2 I) \quad (3.106)$$

olarak hesaplanır. Dönüştürülmüş modeldeki klasik en küçük kareler kestirimi, korelasyonlu hata modelindeki genelleştirilmiş en küçük kareler kestirimi olur ve

$$\hat{\beta}_{GLS} = (X(s)^{*'} X(s)^*)^{-1} X(s)^{*'} Z(s)^* \quad (3.107)$$

$$= (X(s)' L^{-1'} L^{-1} X(s))^{-1} X(s)' L^{-1'} L^{-1} Z(s) \quad (3.108)$$

$$= (X(s)' \Sigma^{-1} X(s))^{-1} X(s)' \Sigma^{-1} Z(s) \quad (3.109)$$

ile gösterilir (Schabenberger, 2005, syf.320-321).

3.2. İLİŞKİLİ HATALAR İÇİN DOĞRUSAL MODELLER

Bu bölümde, genel anlamda, iki yaklaşım incelenecektir: Birincisi, uzamsal sürekli (coğrafi istatistik) verilere uygulanan kovaryans fonksiyonuna dayalı modellerdir. İkincisi, bölgesel verilere uygulanan uzamsal yakınsaklığı dikkate alan modellerdir.

Genelleştirilmiş en küçük kareler kestirimi, (3.109) numaralı eşitlik ile elde edilir. Bu ifadenin kullanılabilmesi için dağılım özelliklerinin bilinmesi gerekmektedir. Eğer $e(s)$ normal dağılıma sahipse β 'nin en çok olabilirlik kestirimi, genelleştirilmiş en küçük kareler kestirimine eşit olacaktır. Bu durumda $\hat{\beta}_{GLS} \sim G(\beta, (X(s)\Sigma^{-1}X(s))^{-1})$ olacaktır. Hatalar normal dağılıma sahip değilse gerekli olan varsayımların sağlanabildiği söylenemez. $\hat{\beta}_{GLS}$ 'nin tutarlılığı $X(s)$ matrisi ve kovaryans matrisi Σ 'ya bağlıdır.

β için $\hat{\beta}_{GLS}$ 'nin tutarlılığının söylenebilmesi için koşullardan biri,

$$\lim_{n \rightarrow \infty} \frac{(X(s)'\Sigma^{-1}X(s))}{n} = Q \quad (3.110)$$

limitinin sağlanmasıdır. Q sonlu ve tekil olmayan matristir. $\hat{\beta}_{GLS}$ 'nin asimptotik özellikleri,

$$\sqrt{n}(\hat{\beta}_{GLS} - \beta) = \left(\frac{X(s)'\Sigma^{-1}X(s)}{n} \right)^{-1} \frac{X(s)'\Sigma^{-1}e(s)}{\sqrt{n}} \quad (3.111)$$

ifadesinin asimptotik özelliklerinden türemiştir.

Burada uygulanan diğer bir koşul Lindberg-Feller merkezi teoremidir ve

$$\sqrt{n}(\hat{\beta}_{GLS} - \beta) \xrightarrow{d} G(0, Q^{-1}) \quad (3.112)$$

ile gösterilir. Merkezi limit teoremlerinde bazı düzenlemeler yapılması sonucunda,

$$\hat{\beta}_{GLS} \sim G(\beta, (X(s)'\Sigma^{-1}X(s))^{-1}) \quad (3.113)$$

elde edilir.

Eğer Σ bilinmiyorsa β ve Σ 'nin kestirilmesinde iteratif yeniden ağırlıklandırılmış genelleştirilmiş en küçük kareler kullanılır ve böylece β 'nin kestirilmiş genelleştirilmiş en küçük kareler kestiricisi,

$$\tilde{\beta}_{EGLS} = (X(s)'\tilde{\Sigma}^{-1}X(s))^{-1} X(s)'\tilde{\Sigma}^{-1}Z(s) \quad (3.114)$$

ile hesaplanacaktır. $\tilde{\Sigma}$ kovaryans parametrelerinin kestirimi olur. $\tilde{\beta}_{EGLS}$ 'nin dağılım özelliklerinin çıkarımı oldukça zordur. Çünkü $\tilde{\Sigma}$ ve $e(s)$ korelasyonlu olur. Bu nedenle asimptotik sonuçlara dönülecek olursa, önceki koşullara ek olarak,

$$\lim_{n \rightarrow \infty} n^{-1} X(s) [\tilde{\Sigma} - \Sigma] X(s) \xrightarrow{p} 0 \quad (3.115)$$

$$\lim_{n \rightarrow \infty} n^{-1/2} X(s)' [\tilde{\Sigma} - \Sigma] e(s) \xrightarrow{p} 0 \quad (3.116)$$

ifadelerinin yanı sıra $\tilde{\beta}_{EGLS}$, $\hat{\beta}_{GLS}$ ile aynı limit teoremine sahip olur ve

$$\sqrt{n}(\tilde{\beta}_{EGLS} - \beta) \xrightarrow{d} G(0, Q^{-1}) \quad (3.117)$$

olur. Buradan yola çıkarak $\tilde{\beta}_{EGLS}$ için,

$$\tilde{\beta}_{EGLS} \sim G(\beta, (X(s)' \Sigma^{-1} X(s))^{-1}) \quad (3.118)$$

şeklinde dağıldığı söylenebilir.

3.2.1. Geniş Alanlı Uzamsal Değişimler için Modeller

Geniş alanlı değişimler için kullanılan modeller hem sürekli hem de alan verilerine uygulanabilirler. $\mathbf{s}(i) = (s_1(i), s_2(i))$ yüzey veya alandaki i 'nci gözenin (birimin) koordinatları olsun. Bu durumda yazılacak genel ifade,

$$f(\mathbf{s}(i)) = \sum_{f+s \leq q} \theta_{r,s} s_1(i)^r s_2(i)^s \quad (3.119)$$

olacaktır. Burada q bir tam sayı olur ve yüzeyin (alanın) sayısını ifade eder. $q = 0$ olması durumunda i 'den bağımsız olarak, ortalama sabit olur. Doğrusal trend söz konusu olduğunda ($q=1$), kuadratik trend söz konusu olduğunda ($q=2$) olur. Yüzeylerin sıralanmış hali için θ 'lar,

Yüzey, $\theta_{0,0}$

Doğrusal, $\theta_{0,0} + \theta_{1,0} s_1(i) + \theta_{0,1} s_2(i)$

Kuadratik, $\theta_{0,0} + \theta_{1,0} s_1(i) + \theta_{0,1} s_2(i) + \theta_{2,0} s_2(i)^2 + \theta_{1,1} s_1(i) s_2(i) + \theta_{0,2} s_2(i)^2$

ifadeleri ile yazılır. Buradan yola çıkarak geniş alanda değişimi tanımlamak üzere kurulacak olan Z modeli, belirtilen ifadeye bağımsız hata terimlerinin eklenmesi ile elde edilir:

$$Z(\mathbf{s}(i)) = Z(i) = f(\mathbf{s}(i)) + e(\mathbf{s}(i)) \quad (3.120)$$

Buradaki $e(\mathbf{s}(i))$, 0 ortalama ve σ^2 varyans ile dağılır. $Z(i)$ 'nin ortalaması (beklenen değeri) ise,

$$E[Z(i)] = \mu(i) = f(\mathbf{s}(i)) = \sum_{r+s \leq q} \theta_{r,s} s_1(i)^r s_2(i)^s \quad (3.121)$$

şeklinde olur (Haining, 2003, syf.293).

3.2.2. Dar Alanlı Uzamsal Değişimler için Modeller

Dar alanda uzamsal değişimlerin incelenmesi istendiğinde modeller, yüzey (surface) verileri için modeller, sürekli değer alan veriler için modeller ve kesikli değer alan veriler için modeller olmak üzere üç gruba ayrılır.

3.2.2.1. Yüzey Veriler için Modeller

Bu tarz modeller otokovaryans terimleri ile ifade edilebilirler. Sabit ortalama için zayıf durağanlık olması,

$$E[Z(\mathbf{s}(i))] = \mu \text{ ve } \text{Cov}[Z(\mathbf{s}(i)), Z(\mathbf{s}(j))] = C(\mathbf{s}(i) - \mathbf{s}(j)) \quad (3.122)$$

ile belirtilir. Buradaki $E[.]$ ortalaması konumdan bağımsızdır ve iki değer arasındaki kovaryans ise sadece göreceli konumlarına bağlıdır. Eğer $\text{cov}[.,.]$ kovaryansı iki nokta arasındaki $d(\mathbf{s}(i) - \mathbf{s}(j))$ uzaklığına bağlıysa, bu sürece “izotropik (isotropik) süreç” denir.

Durağanlık, zayıf durağan olmaktan daha çok incelenen bir genel varsayımdır ve uzamsal değişimin modellenmesinde “semi-variogram”ların ($\nu(.)$) kullanılmasına neden olur. Durağanlık olması,

$$E[Z(\mathbf{s}(i))] = \mu \text{ ve } \text{Var}[Z(\mathbf{s}(i)) - Z(\mathbf{s}(j))] = 2\nu(\mathbf{s}(i) - \mathbf{s}(j)) \quad (3.123)$$

biçiminde belirtilir. Semi-variogramlar sadece dar alanlı uzamsal değişim olması durumunda kullanılır. Semi-variogram sadece $(\mathbf{s}(i) - \mathbf{s}(j))$ farkının fonksiyonudur ve konum fonksiyonu değildir. İncelenen doğrultuda veya yönde bağımlılık yok ise, semi-variogram izotropik olduğunu gösterir ve h uzaklığı göstermek üzere, $\nu(h)$ olarak gösterilir. Diğer durumlarda süreç “izotropik olmama (anisotropic)” adını alır.

İzotropik zayıf durağanlık sürecinde semi-variogram, $\nu(h) = (C(0) - C(h))$ biçimini alır. Bu, zayıf durağanlık sürecinde semi-variogram modeli için olası herhangi bir kovaryans fonksiyonunun kullanılabileceğini gösterir. Semi-variogramlar için bazı örnekler Tablo 3.3’de gösterilmiştir (Haining, 2003, syf.294-296).

Tablo 3.3. Kullanılan Semi-Variogram Örnekleri

Üstel	$v(h) = C_0 + C \left[1 + \exp\left(\frac{3h}{2}\right) \right]$	
Gauss	$v(h) = C_0 + C \left[1 - \exp\left(-3\left(\frac{h}{2}\right)^2\right) \right]$	
Hole	$v(h) = C_0 + C \left[1 - a \sin(ah) / h \right]$	
Küresel	$v(0) = 0$ $v(h) = C \left[\frac{3h}{2a} - \frac{1}{2} \left(\frac{h}{a} \right)^3 \right] + C_0$ $v(h) = C + C_0 \quad h > a$	

(Kaynak: Haining, 2003, syf.296)

3.2.2.2. Sürekli Değerli Alan Verileri için Modeller

Durağanlık ve izotropi ile ilgili anlatılanlar bu tarz verilere ilişkin modellerde kullanılmamaktadır. $\{Z(\mathbf{s}(i)) : \mathbf{s}(i) \in D\}$ için oluşturulan olasılık modellerinde izlenen genel yol, bölümlerin uzamsal konumlarındaki bilgiler ve grafik yapısı (her bir bölümün birbirine komşuluğu) varsayımlarının incelenmesidir.

Bu tarz veriler için model oluşturmaya başlarken uzamsal bağımlılık içeren mevcut modelin “Markov özelliği”ne sahip olduğu düşünülür. Uzamsal Markov özelliği, zaman serisi modellerindeki Markov özelliği ile benzerlik göstermektedir. Örneğin zaman serisi gözlemleri kesikli ve $t = 1, 2, 3, \dots$ anında ölçülmüşlerdir. $\text{Prob}\{Z(t) = z(t) | Z(1) = z(1), \dots, Z(t-1) = z(t-1)\}$ notasyonu, Z raslantı değişkeninin t anında z(t)'ye eşit olduğunu gösteren koşullu olasılık fonksiyonunu göstermektedir. Markov özelliğinin tanımından yola çıkarak koşullu olasılık;

$$\text{Prob}\{Z(t) = z(t) \mid Z(1) = z(1), \dots, Z(t-1) = z(t-1)\} \quad (3.124)$$

$$= \text{Prob}\{Z(t) = z(t) \mid Z(t-1) = z(t-1)\} \quad (3.125)$$

biçiminde yazılır.

Şekil 3.3'de tanımlanmış komşuluklar görülmektedir. (a), (b) 1 boyutlu uzamsal sistemi, (c), (d) 2 boyutlu düzgün olmayan uzamsal sistemi, (e), (f) ise 2 boyutlu düzgün uzamsal sistemi göstermektedir.

(a) şeklinde bölgelerin dağılımı $i(N(i))$ bölgesindeki komşuların kümesiyle birlikte gösterilmiştir. Uzamsal Markov özelliği;

$$\begin{aligned} \text{Prob}\{Z(i) = z(i) \mid \{Z(j) = z(j)\}, \quad j \in D, j \neq i\} \\ = \text{Prob}\{Z(i) = z(i) \mid \{Z(j) = z(j)\}, \quad j \in N(i)\} \end{aligned} \quad (3.126)$$

olarak yazılır. (e) şeklindeki gibi dikdörtgensel kafeslerde i bölgesinin komşuluğu 4 bölge (kuzey, güney, doğu, batı) ile gösterilebilir.

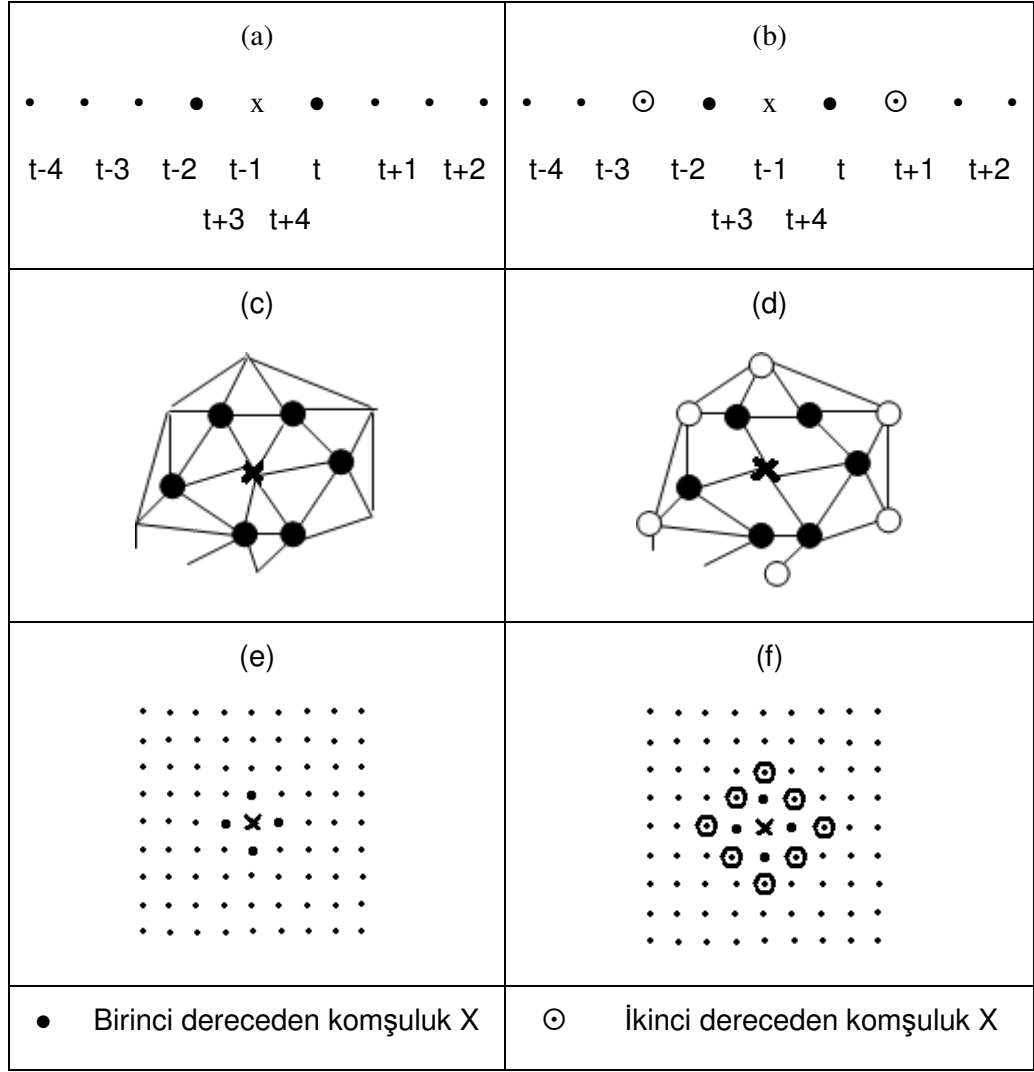
Zaman serisi modellerinde Markov özelliği farklı şekildedir. İkinci boyuttan Markov özelliğini gösteren bir model;

$$\text{Prob}\{Z(t) = z(t) \mid Z(t-1) = z(t-1), Z(t-2) = z(t-2)\} \quad (3.127)$$

olarak verilir. Bu model (t-1) ve (t-2) olmak üzere iki farklı zaman parametreleri için bağımlılık içermektedir. Bu model uzamsal sisteme uyarlandığında;

$$\text{Prob}\{Z(i) = z(i) \mid \{Z(j) = z(j)\}, \{Z(k) = z(k)\}, \quad j \in N_1(i), k \in N_2(i)\} \quad (3.128)$$

şeklinde bulunur. Burada $N_1(i)$, i bölgesine birinci boyuttan komşulukları, $N_2(i)$ ise i bölgesine ikinci boyuttan komşulukları gösterir.



Şekil 3.3. Birimlerin Birbirlerine göre Komşuluklarının Gösterilmesi (Kaynak: Haining, 2003, syf.298)

Birinci boyuttan Markov özelliğine sahip birçok değişkenli normal uzamsal model için;

$$E[Z(i) = z(i) | \{Z(j) = z(j), j \in N(i)\}] \quad (3.129)$$

$$= \mu(i) + \sum_{j \in N(i)} k(i, j) [Z(j) - \mu(j)], \quad i = 1, \dots, n \quad (3.130)$$

ve

$$\text{Var} [Z(i) = z(i) | \{Z(j) = z(j)\}, j \in N(i)] = \sigma(i)^2, \quad i=1, \dots, n \quad (3.131)$$

yazılır. Burada $j \in N(i)$ olmadığı sürece $k(i, i) = 0$ ve $k(i, j) = 0$ olacaktır. Buna “otonormal (autonormal)” veya “koşullu otoresif model (conditional autoregressive model – CAR)” adı verilir. Modelin koşulsuzluk durumu söz konusu olduğunda kullanılacak matris gösterimi aşağıdaki gibidir:

$$E[\mathbf{Z}] = \boldsymbol{\mu} \quad \text{ve} \quad \text{cov}[(\mathbf{Z}-\boldsymbol{\mu}), (\mathbf{Z}-\boldsymbol{\mu})^T] = (\mathbf{I} - \mathbf{K})^{-1}\mathbf{M} \quad (3.132)$$

\mathbf{K} , $n \times n$ boyutlu $\{k(i,j)\}$ ve $\boldsymbol{\mu} = (\mu(1), \dots, \mu(n))^T$ matrisleridir. \mathbf{M} ise $m(i, i) = \sigma(i)^2$ olan köşegen matristir. $(\mathbf{I} - \mathbf{K})^{-1}\mathbf{M}$ 'nin kovaryans matrisi olmasından dolayı simetrik ve pozitif tanımlı olması gerekir. Bundan dolayı $k(i,j)\sigma(j)^2 = k(j,i)\sigma(i)^2$ olması beklenir.

Bu modelde $\mu(i) = \mu$, $k(i,j) = \tau w(i,j)$ olarak ayarlanabilir. Burada τ bir parametreyi ve $\sigma(i)^2 = \sigma^2$ 'yi temsil etmektedir. $(\mathbf{I} - \tau \mathbf{W})$ 'nun tersinebilirliğini gösterebilmek için τ 'nun, w_{\min} ve w_{\max} 'ın \mathbf{W} bağlantı (connectivity) matrisinin en küçük ve en büyük özdeğerleri olmak üzere, $(1 / w_{\min})$ ve $(1 / w_{\max})$ arasında bir değer alması gerekir.

Her bir bölgenin dört komşusu olduğu ($w(i,j) = 0$ veya 1) kare kafeslerde, kafes büyüklüğü arttıkça $w_{\min} \downarrow -4$ ve $w_{\max} \uparrow +4$ olur. Böylece $|\tau| < 0,25$ olur. τ 'nun 0'dan yukarı doğru artması incelenen bölgelerdeki korelasyonun artması anlamına gelir.

Eğer bölgesel veriyle analiz yapan kişi Markov özelliğini dikkate almazsa modeli belirlemek için eşanlı (simultaneous) yaklaşımı kullanması gerekir. $e \sim N(0, \sigma^2 \mathbf{I})$ ve $e(i)$, i ($i=1, \dots, n$) sitesiyle oluşturulmuş bir değişken ise,

$$Z(i) = \mu(i) + \sum_{j \in N(i)} s(i,j)[Z(j) - \mu(j)] + e(i) \quad , \quad i = 1, \dots, n \quad (3.133)$$

biçiminde yazılır. Burada $s(i,i) = 0$ olacaktır. Matris gösteriminde,

$$(\mathbf{I} - \mathbf{S})(\mathbf{Z} - \boldsymbol{\mu}) = \mathbf{e} \quad (3.134)$$

eşitliği ile yazılır. Burada \mathbf{S} , $\{s(i,j)\}$ 'den oluşan $n \times n$ boyutlu matris ve $\boldsymbol{\mu} = (\mu(1), \dots, \mu(n))^T$ olduğu bilinmektedir.

Eğer \mathbf{e} 'nin varyans-kovaryans matrisi köşegen yapıya sahip olan Λ olursa,

$$\text{Cov}[(\mathbf{Z} - \boldsymbol{\mu}), (\mathbf{Z} - \boldsymbol{\mu})^T] = (\mathbf{I} - \mathbf{S})^{-1} \Lambda (\mathbf{I} - \mathbf{S}^T)^{-1} \quad (3.135)$$

bulunur. Bu modele "eşanlı otoregresif model (simultaneous autoregressive model – SAR model)" adı verilir. Burada $\mu(i) = \mu$ ve $s(i,j) = \rho w(i,j)$ alınır ve ρ 'nun parametre, $w(i,j) = w(j,i)$ olduğu bilinirse modelin basit hali elde edilir. ρ , \mathbf{W} 'da en büyük ve en küçük özdeğerler olarak kısıtlanabilir. ρ 'nun artmasıyla gruplanmış bölgeler arasındaki ilişki de artacaktır. Bağlantı matrisi \mathbf{W} 'nun simetrik olması gerekmemektedir. Böylece satır toplamları 1'e standartlanabilir ve bu durumda $|\rho| < 1.0$ olur (Haining, 2003, syf.297-301).

3.2.2.3. Kesikli Değerli Alan Verileri için Modeller

Bu başlık altında ikili (binary) veriler için lojistik modelin uzamsal versiyonu ve sayılabilir (count) veriler için Binom ve Poisson modellerin uzamsal versiyonlarından söz edilebilir.

Lojistik model bir olayın var olup olmaması durumunda kullanılır ($Z(s(i)) = 0$ veya 1). Binom veya Poisson olasılık modelleri sayılabilen durumların olması durumunda kullanılır ($Z(s(i)) = 0, 1, 2, \dots$). Poisson model, olayın gerçekleşme olasılığı p 'nin 0 'a yaklaşması ve n 'in artması durumunda Binom modele yaklaşır. Poisson modeller daha çok nadir, ender görülen olaylar için kullanılırlar. Komşuluğun söz konusu olduğu durumlarda bu modellerin uzamsal (auto) versiyonları koşullu olasılıklara dayanır. Buradan yola çıkarak otolojistik (autologistic) model ($z(i) = 0$ veya 1),

$$\text{Prob}\{Z(i) = z(i) \mid \{Z(j)\}, \quad j \in N(i)\} \quad (3.136)$$

$$= \exp [z(i)(\alpha(i) + \sum_{j \in N(i)} b(i, j)z(j))] / [1 + \exp [(\alpha(i) + \sum_{j \in N(i)} b(i, j)z(j))] \quad (3.137)$$

biçiminde gösterilir. Bu koşullu olasılıkta $Z(i)$ 'nin 0 veya 1 değerini alması, Z 'nin komşu bölgelerdeki değerlerine bağlıdır. Burada $b(i, j) = b(j, i)$ olacaktır. $b(i, j)$ 'nin tüm i ve j değerleri için 0 değerlerini alması modelin uzamsal olmayan bir yapı sergileyeceğini gösterir.

$Z(i) = 0, 1, 2, \dots, c(i)$ için tanımlanan otobinom (autobinomial) model,

$$\text{Prob}\{Z(i) = z(i) \mid \{Z(j)\}, \quad j \in N(i)\} \quad (3.138)$$

$$= (C(i)! - z(i)! (c(i) - z(i))!) \theta(i)^{z(i)} (1 - \theta(i))^{c(i) - z(i)} \quad (3.139)$$

$$\theta(i) = \exp [\alpha(i) + \sum_{j \in N(i)} b(i, j)z(j)] / [1 + \exp [(\alpha(i) + \sum_{j \in N(i)} b(i, j)z(j))] \quad (3.140)$$

ifadesi ile yazılır. Burada $c(i) = 1$ olduğunda model otolojistik modele dönüşür. Bu model için gerekli koşul $b(i, j) = b(j, i)$ olmasıdır. Modelde $c(i)$, i bölgesindeki deneme sayısını, $\theta(i)$ i bölgesindeki başarı oranını gösterir.

Oto-poisson model ise ($Z(i) = 0, 1, 2, \dots$) aşağıdaki gibi gösterilir:

$$\text{Prob}\{Z(i) = z(i) \mid \{Z(j)\}, \quad j \in N(i)\} = \exp [-\lambda(i)] \lambda(i)^{z(i)} / z(i)! \quad (3.141)$$

$$\lambda(i) = \exp [(\alpha(i) + \sum_{j \in N(i)} b(i, j)z(j))] \quad (3.142)$$

Uzamsal bağımlılık $\lambda(i)$ parametresi yardımıyla incelenir. Bir önceki modelde olduğu gibi $b(i,j)=b(j,i)$ koşulu burada da geçerlidir. Bunun yanı sıra $b(i,j) \leq 0$ (tüm i ve j 'ler için) koşulunun da sağlanması gerekir (Haining, 2003, syf.304-306).

3.2.3. Karma Modeller

Bir çok çalışmada $e(s)$ 'lerdeki değişimin sistematik bileşenleri olduğu bilinir. Böyle durumlarda hem sabit hem de rassal etkilerin dikkate alındığı karma modeller (mixed models) kullanılması uygun olur.

Doğrusal karma modelin (LMM) genel formu,

$$Z(s) = X(s)\beta + U(s)\alpha + \varepsilon(s) \quad (3.143)$$

şeklinde yazılır. Burada α , $K \times 1$ boyutlu ve 0 ortalama ile G varyansa sahip rassal etkileri gösteren parametredir. $\varepsilon(s)$ model hataları vektörü α 'dan bağımsız ve 0 ortalama ile R varyansa sahiptir. Buradaki tüm etkileri göz önüne alıp kestirim yapmak oldukça karışık ve zordur.

Karma bir model signal model ile ilişkilendirilebilir.

$$Z(s) = S(s) + \varepsilon(s) \quad (3.144)$$

olarak gösterilen karma modelde,

$$S(s) = X(s) \beta + W(s) + \eta (s) \quad (3.145)$$

olarak yazılırsa belirtilen signal model yardımıyla, $U(s)$, $W(s) + \eta (s)$ haline dönüştürülmüş olur.

G ve R için de gerekli varsayımların incelenmesi gerekmektedir. Karma model denklemleri, $[\alpha, \varepsilon(s)]$ 'nin bileşik olabilirliliğinin belirlenmesi ile başlar ve β ile α 'nın maksimize edilmesi ile devam eder. Normallik varsayımı altında bileşik yoğunluk,

$$f(\alpha, \varepsilon(s)) = \frac{1}{(2\pi)^{(n+k)/2}} \begin{vmatrix} G & 0 \\ 0 & R \end{vmatrix}^{-1/2} \times \exp \left\{ -\frac{1}{2} \begin{bmatrix} \alpha \\ Z(s) - X(s)\beta - U(s)\alpha \end{bmatrix} \begin{bmatrix} G & 0 \\ 0 & R \end{bmatrix}^{-1} \begin{bmatrix} \alpha \\ Z(s) - X(s)\beta - U(s)\alpha \end{bmatrix} \right\} \quad (3.146)$$

biçiminde yazılır. Bu fonksiyonun maksimizasyonu,

$$Q(\beta, \alpha) = (Z(s) - X(s)\beta - U(s)\alpha)' R^{-1} (Z(s) - X(s)\beta - U(s)\alpha) + \alpha' G^{-1} \alpha \quad (3.147)$$

ifadesi minimize edildiği zaman sağlanmış olur. Tüm adımlar sonucunda Henderson'un karma modeli olarak da bilinen eşitlik,

$$\begin{bmatrix} X(s)'R^{-1}X(s) & X(s)'R^{-1}U(s) \\ U(s)'R^{-1}X(s) & U(s)'R^{-1}U(s) + G^{-1} \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{\alpha} \end{bmatrix} = \begin{bmatrix} X(s)'R^{-1}Z(s) \\ U(s)'R^{-1}Z(s) \end{bmatrix} \quad (3.148)$$

olarak belirtilir. Sonuçlar ise,

$$\hat{\beta} = (X(s)'V^{-1}X(s))^{-1} X(s)'V^{-1}Z(s) \quad (3.149)$$

$$\hat{\alpha} = GU(s)'V^{-1}(Z(s) - X(s)\hat{\beta}) \quad (3.150)$$

$$V = U(s)GU(s)' + R \quad (3.151)$$

eşitlikleri ile gösterilir (Schabenberger, 2005, syf.325-326).

3.2.4. Uzamsal Genişleme Modeli

Burada kullanılacak model,

$$y_i = \alpha + \beta x_{i1} + \dots + \tau x_{im} + \epsilon_i \quad (3.152)$$

$$\alpha_i = \alpha_0 + \alpha_1 u_i + \alpha_2 \gamma_i \quad (3.153)$$

$$\beta_i = \beta_0 + \beta_1 u_i + \beta_2 \gamma_i \quad (3.154)$$

$$\tau_i = \tau_0 + \tau_1 u_i + \tau_2 \gamma_i \quad (3.155)$$

şeklinde gösterilir. Burada u_i ve γ_i , i bölgesinin uzamsal koordinatlarını gösterir.

Genişleme modeli,

$$y_i = \alpha_0 + \alpha_1 u_i + \alpha_2 \gamma_i + \beta_0 x_{i1} + \beta_1 u_i x_{i1} + \beta_2 \gamma_i x_{i1} + \dots + \tau_0 x_{im} + \tau_1 u_i x_{im} + \epsilon_i \quad (3.156)$$

şeklinde bulunur (Fotheringham, 2006, syf.16).

3.2.5. Uzamsal Ayarlanabilir Filtreleme Modeli

Çok değişkenli zaman serileri verilerine uygulandığında regresyon parametrelerinin zaman içindeki kaymalarını gösteren bir yöntemdir.

$$y_t = \sum_j x_{tj} \beta_{tj} + \epsilon_t \quad (3.157)$$

modelinde t, kesikli zaman noktasını gösterir. t zamanında yeni bir çok değişkenli gözlem eklenirse var olan $\hat{\beta}_{t-1}$ regresyon katsayısı bağımlı değişkeni kestirmek için kullanılır. Kestirim için,

$$\hat{\beta}_{jt} = \hat{\beta}_{jt-1} + \left| \hat{\beta}_{jt-1} \right| \alpha_j (y_t - \hat{y}_t) / |\hat{y}_t| \quad (3.158)$$

ifadesi kullanılır. Burada $\hat{\beta}_{jt}$, $\hat{\beta}_t$ matrisinin j'nci elemanıdır (Fortheringham, 2006, syf.17).

3.2.6. Çok Düzeyli Modelleme

İncelenen etkileri çok boyutta gösteren modellerdir.

$$y_{ij} = \alpha_j + \beta_j x_{ij} + \epsilon_{ij} \quad (3.159)$$

olarak gösterilen modelde y_{ij} , j bölgesinde i'nin davranışını, x_{ij} , i'nin j bölgesindeki x değerini, α_j ve β_j ise yer-özel parametrelerini göstermektedir.

$$\alpha_j = \alpha + \mu_j^\alpha \quad (3.160)$$

$$\beta_j = \beta + \mu_j^\beta \quad (3.161)$$

Bu parametrelerin elde edilmesi ile çok düzeyli model,

$$y_{ij} = \alpha + \beta x_{ij} + (\epsilon_{ij} + \mu_j^\alpha + \mu_j^\beta x_{ij}) \quad (3.162)$$

ile gösterilir. Bu model μ_j^α ve μ_j^β sıfır olmadıkça en küçük kareler ile çözülemez (Fortheringham, 2006, syf.18-19).

3.2.7. Rasgele Katsayılar Modelleri

Doğrusal regresyon uygulamasında her bir etkinin sabit olduğu varsayılır. Rasgele katsayılar modelinde model parametrelerinin sabit olmadığı, durumdan duruma değişebildiği ve genellikle normal olmak üzere rasgele dağılımlara sahip olduğu bilinir. Bu durumda model,

$$y_i = \sum_j x_{ij} \beta_{ij} + \epsilon_i \quad (3.163)$$

biçiminde gösterilir. β_{ij} rasgele bir deęişken olurr ve $\beta_{ij} \sim N(\beta_j, \sigma_j^2)$ ile daęılır. Çözümler her bir durum için Bayes teoremi yardımıyla elde edilebilir (Fortheringham, 2006, syf.20).

4. UZAMSAL REGRESYON ANALİZİ

Klasik doğrusal regresyon modelleri hata terimlerinin bağımsız ve normal dağılıma sahip olma koşullarının yerine getirilmesi durumunda uygulanması uygun olan yöntemlerdir. Analiz sürecinde doğrusal regresyon modeli oluşturulduktan sonra varsayımların incelenmesi gerekir. Bilindiği gibi modellerin tercih edilmesi çoğunlukla gözlenen ve kestirilen değerlerin farklarından elde edilen artıkların incelenmesi ile sağlanır. Eğer varsayımlardan sapmalar görülürse kestirilen modelin kullanılmasının uygun olmadığı ve hatalı kestirimlerin elde edileceği bilinir.

Artık değerlerinin bağımlılığının söz konusu olduğu ve uzamsal değişkenlerin bulunduğu modellerde uzamsal regresyon analizlerinin uygulanması uygundur (Chi, 2008). Bu nedenle uzamsal bağımlılığın incelenmesinin ardından uzamsal yapıda modellerin kurulması çalışmanın doğruluğu ve güvenilirliği için önem taşımaktadır.

Uzamsal regresyon analizinin ilk adımında açıklayıcı uzamsal veri analizleri (exploratory spatial data analysis – ESDA) uygulanır. ESDA yardımıyla verilerde uzamsal yapının görülmesi, uzamsal kümelerin ve uzamsal aykırı değerlerin ortaya çıkarılması ve regresyon modelindeki uzamsal yapıyı ortaya çıkaracak istatistiklerin elde edilmesi sağlanır. Daha açık bir ifade ile ESDA, uzamsal otokorelasyonun, uzamsal heterojenliğin, uzamsal komşuluk yapısına bağlı olarak uzamsal ağırlık matrisinin belirlenmesini sağlar.

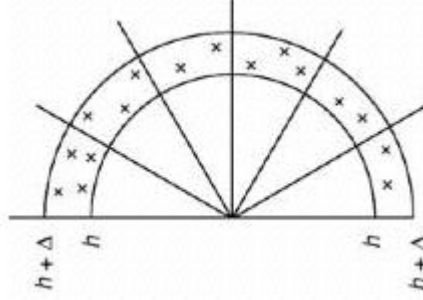
4.1. UZAMSAL BAĞIMLILIĞIN SAYISAL OLARAK GÖSTERİLMESİ

Uzamsal bağımlılığın coğrafi yapılarda çok sık görülen bir sorun olduğu bilinmektedir. Bu yüzden bağımlılığın incelenmesi oldukça büyük önem taşır.

Uzamsal bağımlılığı gösterebilmek için bilinen yollardan biri noktalar kümesindeki veya alanlardaki nesnelere arasında var olan uzamsal ilişkilerin gösterilmesidir. Bunun gösteriminin ardından uzamsal bağımlılık farklı yöntemlerle açıklanabilir.

4.1.1. Arazi: İki Boyutlu Sürekli Uzaydan Veriler

Bu durumdaki verilerde uzamsal ilişki, uzaklık veya uzaklık bantları ile ifade edilirler. $s(i)$, bölgenin içindeki noktalar arası $d(i,j)$ uzaklık bilgisi olarak kaydedilmesi ile $[(i,j)|d(i,j)=h]$ olarak gösterilsin. Buradaki koşul, i 'den h kadar uzaklıkta bir j 'nin seçilmiş olduğunu göstermektedir. Bu koşul, h değerinin sürekli değişebilir olmasından dolayı esnetilebilmekte ve $[(i,j)|d(i,j)=h\pm\Delta]$ olarak da gösterilebilmektedir.



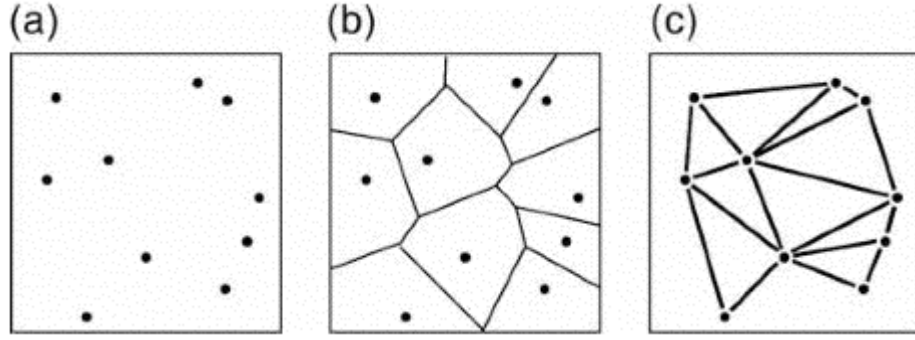
Şekil 4.1. Otuz Derecelik Bölmelerle Ayrılmış Uzaklık Bantlarıyla Gösterilmiş Komşular (Kaynak: Haining, 2003, syf.75)

Şekilde gösterilen her bir bölme, incelemenin daha ayrıntılı yapılabilmesi için b olarak adlandırılırsa, i ve j arasındaki uzaklık $[(i,j) | d_b(i,j) = h \pm \Delta]$ olarak gösterilir (Haining, 2003, syf.75). Elde edilen uzaklık değerleri, uzamsal korelasyonun hesaplanmasında kullanılan Moran I testi ve Geary oranının elde edilmesinde çıkış noktasıdır.

4.1.2. Nesnelere: İki Boyutlu Kesikli Uzaydan Veriler

Kesikli uzayda komşulukların tanımlanmasında aşağıdaki kavramlar kullanılmaktadır:

- Düz (straight) doğru uzaklıkları: Her nokta diğer noktalara ortak sınır varlığı gözlemeksizin belirlenmiş uzaklıklar ile bağlıdır.
- En yakın komşuluk: Her nokta kendine en yakın k nokta ile ($k=1, 2, \dots$) birleştirilir. Ancak A 'nın en yakın komşusunun B olmasının, B 'nin en yakın komşusunun A olduğu anlamına gelmediği bilinmelidir (Haining, 2003, syf.80).
- Delaunay üçgenleştirme yöntemi: Dirichlet hücreleri adı verilen ve alanı, her bir noktayı hücreler içinde alacak şekilde parçalara bölen sistemde, oluşturulan hücrelerin köşelerindeki tüm noktaların birleştirilmesi ile uygulanan yöntemdir (Ripley, 2004, syf.39).



Şekil 4.2. Delaunay Üçgenleştirme Yönteminin Gösterilmesi (Kaynak: Haining, 2003, syf.81)

Delaunay üçgenleştirme yönteminin uygulanmasında, şekil 4.2.(a)'da bir alan içerisinde yer alan noktaların (b)'de Dirichlet hücrelerinin içerisine alınması ve (c)'de hücrelerin köşe noktalarının birleştirilmesi gösterilmiştir.

Uzamsal ilişkiler sınırdaşlık veya bağlanırlık (connectivity) matrisi (C) ile gösterilebilirler. Eğer n nesne (nokta veya alan) varsa, nxn boyutlu bir matris oluşturulabilir. C(i,j) C matrisindeki i satırı ve j sütunundaki değer ise, iki nesnenin komşu olması durumunda,

$$C(i,j) = C(j,i) = 1 \quad (4.1)$$

olur. Ancak ilişkilerin tanımlanmasında 0 ve 1 değerlerinin kullanılması zorunlu değildir (Haining, 2003, syf.82).

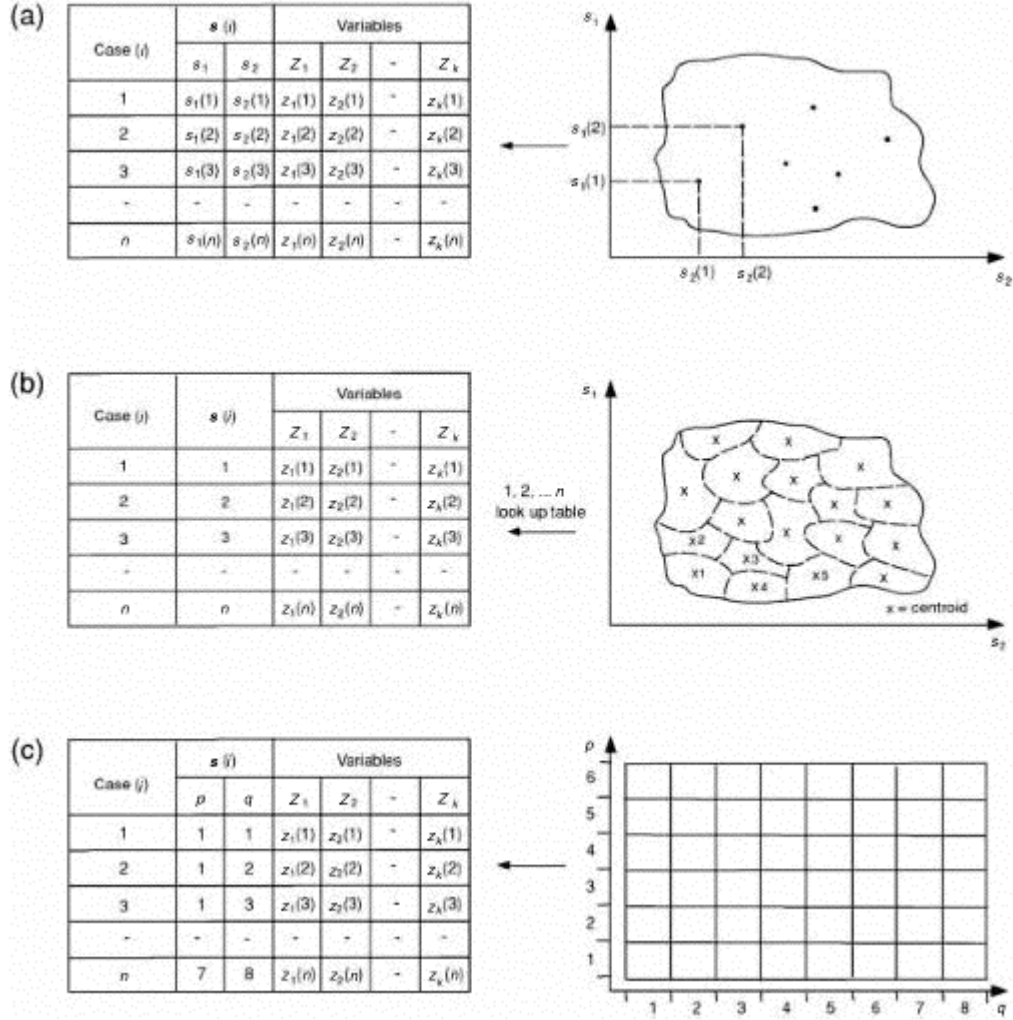
4.2. UZAMSAL AĞIRLIK FONKSİYONLARI

Uzamsal regresyon analizinde belirtilen bölgelerin birbirleri ile etkileşimlerinin büyüklüklerini göstermek için “uzamsal ağırlık matrisleri” (W) oluşturulur. Bu matrislerin kullanılmasındaki amaç, koordinatlar dikkate alındığında birbirine daha yakın olan bölgelerde incelenen olayın, birbirine uzak bölgelerden daha fazla etkileşim halinde olduğunu göstermektir.

Z_1, Z_2, \dots, Z_k , k sayıda değişken değerini, s de noktanın ya da alanın konumunu gösterebilir. Buradan uzamsal veri matrisi,

$$\begin{array}{c}
 \text{Veriler} \qquad \qquad \text{Konum} \\
 \left[\begin{array}{cccccc}
 z_1(1) & z_2(1) & \cdots & z_k(1) & s(1) \\
 z_1(2) & z_2(2) & \cdots & z_k(2) & s(2) \\
 \vdots & \vdots & \vdots & \vdots & \vdots \\
 z_1(n) & z_2(n) & \cdots & z_k(n) & s(n)
 \end{array} \right]
 \end{array} \quad (4.2)$$

biçiminde oluşur. Bunun yanı sıra farklı yapılarda verilerin komşuluk durumları aşağıdaki şekilde ayrıntılı olarak gösterilmiştir:



Şekil 4.3. Farklı Komşuluk Yapıları ve Komşuluk Matrislerinin Gösterimi (Kaynak: Haining, 2003, syf.55)

Verilerin bir bölgede i noktalarının belirtildiği koordinat sistemi, (a) şeklinde gösterilmiştir. Bu tip yapıya sahip verilerde ağırlık matrisleri oluşturulurken her bir gözlem için koordinat eksenlerindeki her iki konum da dikkate alınır. Yapısal anlamda (b) şeklinde gösterilen gözlemler, incelenen konumların merkezleri olarak ele alınırsa uzamsal nesnelerin konumu bölge numaraları olarak gösterilebilirler. Eğer veriler (c) şeklinde olduğu gibi kare pikseller halinde verilmişse, verilen matriste p ve q olarak gösterilen bölge numaraları ile etiketlenebilirler.

Genel anlamda ağırlık matrislerini oluştururken ilişkileri açıklamak için ilişkinin var olması durumunda 1, olmaması durumunda 0 yapısının kullanılması zorunlu değildir. Ağırlık matrisleri oluşturmak için literatürde birbirinden farklı birçok ağırlık

fonksiyonu kullanılmıştır. Bunlardan bilinen ve en sık kullanılan birkaç tanesi aşağıda açıklanmıştır:

(i) Uzaklık: $w_{ij} = d_{ij}^{-\delta}$ olarak gösterilebilir. Burada d_{ij} , i ve j arasındaki uzaklık ve $\delta \geq 0$ 'dır. Gatrell (1983) bu fonksiyonun yanı sıra uzaklığı birçok farklı metrik yöntemle de göstermiştir.

(ii) Uzaklığın üstel fonksiyonu: $w_{ij} = \exp(d_{ij}^{-\delta})$ ifadesinde $\exp()$ üstel fonksiyonu göstermektedir.

(iii) Ortak sınır: $w_{ij} = (l_{i,j} / l_i)^\tau$ ifadesinde l_{ij} , i ve j arasındaki ortak sınırın uzunluğunu; l_i ise i'nin sınır uzunluğunu göstermektedir. $\tau \geq 0$ olarak belirlenir.

(iv) Ortak sınır ve uzaklık ağırlığı: $w_{ij} = (l_{i,j} / l_i)^\tau d_{ij}^{-\delta}$ olarak, (i) ve (iii) fonksiyonlarının birleştirilmesi ile oluşturulur.

(i) ve (ii) durumunda ağırlıklar sıfırdan farklı hesaplanmalıdır. Ancak bilinmesi gereken en önemli nokta uzaklık değerinin artmasıyla ağırlığın daha düşük değerler almasının gerektiğidir. (iii) durumunda sadece ortak sınırı olmayan değerler için ağırlık 0 değerini alır. Bu ağırlık j ile i arasındaki sınır küçüldükçe daha düşük değerler alır ve τ 'nin büyük değerleri için de azalır. (iv) durumunda ise sadece ortak sınırı olan alanlar dışında kalan değerler sıfır olur ve j ile i arasındaki sınır küçüldükçe ağırlık da düşük değerler alır.

Bazı ağırlık matrisleri satır standartlaştırılmış şekilde düzenlenebilir. Satırdaki her bir değerın satır toplamına bölünmesi ve sonucunda satır toplamının 1'e eşit olması gerekmektedir. Standartlaştırılmış matris W^* ile gösterilirse,

$$w_{ij}^* = (w_{ij} / \sum_{j=1}^n w_{ij}) \quad (4.3)$$

olarak ifade edilir (Haining, 2003, syf.84).

Bavaud (1998) genel ağırlık matrislerinin özellikleri üzerine çalışmış ve toplam alanda her bir bölgenin veya yerin önemini açıklamak gibi, ağırlığın uzamsal sistemdeki önemini belirtmiştir. Tüm bunların yanı sıra birçok yazar ağırlık matrislerinin özdeğer ve özvektörlerini incelemiş ve uzamsal yapıların nasıl karakterize edilebileceğini açıklamaya çalışmışlardır.

Uzamsal ağırlıklar için yapılan matris tanımlamalarında Cliff ve Ord'un (1981) önerdiği ikili (binary) veya satır standartlaştırılmış sınırdışlık (contiguity), ortak sınır

uzunluğu, merkez uzaklıklar fonksiyonları gibi yöntemlerin kullanıldığı bilinmektedir. Bunun yanı sıra Stetzer'in (1982) uzamsal ağırlıklar oluşturmak ile ilgili yürüttüğü simülasyon deneyinde ise ulaştığı sonuçlardan bazıları aşağıdaki gibi özetlenebilir:

- Tanımlanamamış uzamsal ağırlık matrisi kullanmak standart hataları arttırmaktadır. Doğru tanımlanmış uzamsal ağırlık matrisi ise her zaman en düşük standart hataların bulunmasına imkan sağlar.
- Uzamsal ağırlık matrisinin gereğinden yüksek sayılarla oluşturulması standart hataların şişmesine neden olacaktır.
- Uzamsal çalışma alanında etkin bölgenin seçilmesi, ağırlıkların fonksiyonel formlarının belirlenmesinden daha önemlidir.
- Küçük örnekleme çalışılması matrisin etkinliğinde sorunlar çıkmasına neden olacaktır.

Florax ve Rey (1995) coğrafi ağırlık matrislerinin tanımlanmasındaki sorunların uzamsal hipotez testlerinin istatistiksel güçlerine etkileri üzerine çalışmışlar ve

- uzamsal ağırlık matrisinin tanımlanamamasının testlerin güçlerinde kayıplara neden olduğunu,
- gereğinden düşük değerli uzamsal ağırlık matrisinin olması, pozitif uzamsal otokorelasyon olması durumunda testlerin güçlerinde artışa, negatif uzamsal otokorelasyon olması durumunda ise azalışa neden olduğunu,
- uzamsal ağırlık matrisinde tanımlanma sorunu olduğu zaman Moran katsayılarının, uzamsal otokorelasyon test istatistiklerinden daha iyi sonuçlar verdiğini,
- yüksek veya düşük tanımlanmış uzamsal ağırlık matrislerinin uzamsal ekonometrik modellerin ortalama hata karelerinde artışa neden olduğunu

belirlemişlerdir.

Upton (1991) merkez (centroid) uzaklıklar tanımlamasını kullanarak,

$$(populasyon_j)^\alpha (alan_j)^\tau / d_{ij}^{-\gamma}$$

ifadesini ağırlık olarak kullanmıştır. Bu ifadedeki α ve τ ikili üsteller; d_{ij} ise i ve j bölgelerinin merkezleri arasındaki uzaklığı göstermektedir. Upton'a göre gerçek modelin,

$$Y = X\beta + V^{1/2} \varepsilon \quad (4.4)$$

olduğu ve düzenlenmiş ağırlık matrisinin etkisiyle,

$$Y = X\beta + \Omega^{1/2} \varepsilon \quad (4.5)$$

şekline dönüştürdüğü belirtilmiştir. Modellerdeki V ve Ω terimleri uzamsal bağımlılığın derecesini ve yapısını belirten kovaryans matrisleridir. Bunlara bağlı olarak oluşturulabilecek teoremler aşağıda verilmiştir:

TEOREM 1: Uzamsal ağırlık matrisinin tanımlanamaması, $b = (X' \Omega X)^{-1} X' \Omega Y$ katsayısının yanlış çıkmasına neden olur.

TEOREM 2: Uzamsal istatistik modelinde uzamsal ağırlık matrisinin tanımlanamaması, modelin etkinliğini ortadan kaldırır ve $Var(\hat{\mu}_\Omega) / Var(\hat{\mu}_V) \leq 1$ 'dir.

TEOREM 3: Uzamsal ağırlığın tanımlanamaması, (4.6) ile gösterilen varyans kestiriminin yanlış çıkmasına neden olur.

$$s^2 = (Y - X\beta)' \Omega^{-1} (Y - X\beta) / (n - k - 1) \quad (4.6)$$

Bu eşitlikteki X matrisinde $(k+1)$ değişken bulunacaktır (Arlinghaus, 1996, syf.68-71).

4.3. UZAMSAL REGRESYON ANALİZİNDE KULLANILAN MODELLER

Verilerde komşulukların önemli olduğu, birbirlerine göre konumlarının etkiye neden olduğu ve hatta bazı çalışmalarda koordinatların dikkate alındığı durumlarda uzamsal regresyon analizinin uygulanması ve buna bağlı olarak uzamsal modellerin kullanılması, uzamsal etkinin gösterilmesinde daha iyi sonuçlar verecektir. Kullanılan modellere ilişkin özellikler ve süreç ayrıntıları ile açıklanmıştır.

4.3.1. Uzamsal Otoregresif Modeller

Zaman serilerinde otoregresif modeller t zamanındaki veriyi geçmiş değerlerle doğrusal birleşimi şeklinde göstermek için kullanılır. Uzamsal yapıda otoregresif modeller ise s noktasındaki veriyi komşu değerlerin doğrusal bileşimi olarak gösterir. Otoregresyon verilerde uzamsal bağımlılığın olduğu varsayılır.

Uzamsal bir örneklemden alınan veriler ile oluşturulan otoregresif modeller grubu,

$$y = \rho W_1 y + X\beta + u \quad (4.7)$$

$$u = \lambda W_2 u + \varepsilon \quad (4.8)$$

$$\varepsilon \approx N(0, \sigma^2 I_n) \quad (4.9)$$

olarak gösterilir. Burada y , $n \times 1$ boyutlu bağımlı değişkenler vektörünü; X , $n \times k$ boyutlu bağımsız değişkenler vektörünü; W_1 ve W_2 , bilinen $n \times n$ boyutlu uzamsal ağırlık matrisini göstermektedir. Kullanılacak diğer birçok uzamsal model, bu modeller grubundan türetilerek oluşturulmuştur.

4.3.1.1. Birinci Dereceden Uzamsal Otoregresif Model

(4.7) ile gösterilen genel modelde, $X = 0$ ve $W_2 = 0$ alınırsa birinci dereceden uzamsal otoregresif model elde edilir ve,

$$y = \rho W y + \varepsilon \quad (4.10)$$

$$\varepsilon \approx N(0, \sigma^2 I_n) \quad (4.11)$$

şeklinde gösterilir. Bu model y 'deki değişimi açıklayıcı değişkenler olmaksızın komşuluk birimlerinin doğrusal kombinasyonu olarak gösterir. (4.10) numaralı modele birinci dereceden otoregresif model denilmesinin nedeni zaman serisi analizlerinde kullanılan ve değişimi önceki dönemlerdeki değişimlerle açıklayan $y_t = \rho y_{t-1} + \varepsilon_t$ modeline uzamsal etkilerin eklenmesi sonucundan oluşturulmuş olmasıdır.

En küçük kareler kestirimi kullanılarak uzamsal otoregresif modelde bulunan ρ parametresi,

$$\hat{\rho} = (y' W' W y)^{-1} y' W' y \quad (4.12)$$

ile kestirilir. Kestirim $E(\varepsilon) = 0$ için yansızdır:

$$E(\hat{\rho}) = E[(y' W' W y)^{-1} y' W' (\rho W y + \varepsilon)] \quad (4.13)$$

$$= \rho + E[(y' W' W y)^{-1} y' W' \varepsilon] \quad (4.14)$$

ρ 'nun en çok olabilirlik kestirimi,

$$L(y|\rho, \sigma^2) = \frac{1}{2\pi\sigma^{2(n/2)}} |\mathbf{I}_n - \rho W| \exp\left\{-\frac{1}{2\sigma^2} (y - \rho W y)' (y - \rho W y)\right\} \quad (4.15)$$

eşitliğinde $\hat{\sigma}^2 = (1/n)(y - \rho W y)' (y - \rho W y)$ kullanılırsa,

$$Ln(L) = -\frac{n}{2} \ln(\pi) - \frac{n}{2} \ln(y - \rho Wy)'(y - \rho Wy) + \ln|I_n - \rho W| \quad (4.16)$$

elde edilir. Bu ifade basit tek deęişkenli optimizasyon süreci uygulanarak en büyüklenir ve $\hat{\rho}$ kestirilir (LeSage, 1998, syf.43-47).

4.3.1.2. Karma Otoregresif Model

Karma otoregresif modeller (uzamsal gecikme modelleri), birinci dereceden uzamsal otoregresif modele klasik regresyon modellerinde olduęu gibi X açıklayıcı deęişkenler matrisinin eklenmesi ile elde edilen modellerdir. Anselin (1989) bu modelin parametrelerinin kestiriminde en çok olabilirlik yönteminin kullanılmasını saęlayan bir çok çalışma yapmış ve kullandığı modeli “karma regresif – uzamsal otoregresif model” (spatial autoregressive model – SAR) olarak adlandırmıştır. SAR model,

$$y = \rho Wy + X\beta + \varepsilon \quad (4.17)$$

$$\varepsilon \approx N(0, \sigma^2 I_n)$$

olarak gösterilmiştir. Burada y, nx1 boyutlu bağımlı deęişken vektörü; X, nxk boyutlu açıklayıcı deęişkenler matrisi; W uzamsal ağırlık matrisidir. ρ parametresi uzamsal gecikmeli bağımlı deęişkenin katsayısı; Wy ve β açıklayıcı deęişkenlerin bağımlı deęişken üzerindeki etkilerini gösteren ifadedir. Bu modele Anselin tarafından uzamsal otoregresif model adının verilmesinin nedeni, zaman serileri analizlerinde kullanılan gecikmeli modeli anımsatmasıdır. Model, standart regresyon modelinin uzamsal gecikmeli bağımlı deęişken ile birleştirmesidir.

Karma otoregresif modelin oluşturulma sürecine ilişkin algoritma aşağıdaki gibi gösterilebilir (LeSage, 1998).

- $y = X\beta_0 + \varepsilon_0$ klasik regresyon modeli kestirilir.
- β_L gecikmeli modelin bağımsız deęişkenini göstermek üzere $Wy = X\beta_L + \varepsilon_L$ modeli klasik sistemde oluşturulur.
- $e_0 = y - X\hat{\beta}_0$ ve $e_L = Wy - X\hat{\beta}_L$ yardımıyla artıklar hesaplanır.
- Elde edilmiş olan e_0 ve e_L yardımıyla ρ deęerini bulmak için kullanılacak olabilirlik fonksiyonu oluşturulur:

$$L = -(n/2)\ln(\pi) - (n/2)\ln(1/n)(e_0 - \rho e_L)'(e_0 - \rho e_L) + \ln|I - \rho W| \quad (4.18)$$

- Elde edilen $\hat{\rho}$ yardımıyla katsayılar kestirilir:

$$\hat{\beta} = (\hat{\beta}_0 - \rho \hat{\beta}_L) \text{ ve } \hat{\sigma}_\varepsilon^2 = (1/n)(e_0 - \rho e_L)'(e_0 - \rho e_L) \quad (4.19)$$

Belirtilen algoritmanın uygulanması sonucunda gecikme modeline ilişkin katsayılar elde edilmiş olacaktır (LeSage, 1998, syf.63-64).

4.3.1.3. Uzamsal Hata Modeli

(4.7) ile gösterilen genel uzamsal model çeşitli düzenlemeler ile (4.20) ile gösterilen şekle getirilir ve oluşturulan model “uzamsal hata modeli” (Spatial Error Model – SEM) adını alır.

$$\begin{aligned} \mathbf{y} &= \mathbf{X}\boldsymbol{\beta} + u \\ u &= \lambda \mathbf{W}u + \varepsilon \\ \varepsilon &\approx N(0, \sigma^2 I_n) \end{aligned} \quad (4.20)$$

Modelde y , $n \times 1$ boyutlu bağımlı değişken matrisini; X matrisi, $n \times k$ boyutlu bağımsız değişkenler matrisini göstermektedir. W , uzamsal ağırlık matrisi ve λ , uzamsal ilişkili hataların katsayısı olarak belirtilmiştir.

Artıklardaki uzamsal otokorelasyonun varlığı birçok istatistiksel test ile sınanabilir. Bunlardan en bilineni Moran istatistiğidir. Bu teste ilişkin özellikler (2.4.1.) numaralı bölümde incelenmiştir. Moran istatistiği uzamsal ağırlık matrisi W 'nun standartlaştırılmış olup olmamasına göre iki farklı yapıda oluşturulabilir:

1. W ağırlık matrisi standartlaştırılmamış ise, $I = (n/s)[e'We]/e'e$
2. W ağırlık matrisi standartlaştırılmış ise, $I = e'We/e'e$

İncelenen istatistiklerdeki e ifadesi regresyon artıklarını göstermektedir. Cliff ve Ord (1972, 1973, 1981) en küçük kareler artıklarına dayanan ve I değerinden bu istatistiğin ortalamasının çıkartılıp yine aynı istatistiğin standart sapmasına bölünmesi ile düzeltilmiş Moran istatistiği için asimptotik dağılımı gösteren çalışmalar yapmışlardır. Bu düzeltme W matrisinin standart olup olmaması benzer şekilde iki yapıda gerçekleştirilmektedir (LeSage, 1998, syf.71-72).

- W matrisinin standartlaştırılmamış olması durumunda ve $M = (I - X(X'X)^{-1}X')$ matrisini göstermek üzere:

$$\begin{aligned}
E(I) &= (n/s)iz(MW)/(n-k) \\
V(I) &= \left[(n/s)^2 \left[iz(MWMMW') + iz(MW)^2 + (iz(MW))^2 \right] / d \right] - E(I)^2 \\
d &= (n-k)(n-k+2) \\
Z_I &= [I - E(I)] / V(I)^{1/2}
\end{aligned} \tag{4.21}$$

- W matrisinin standartlaştırılmış olması durumunda:

$$\begin{aligned}
E(I) &= iz(MW)/(n-k) \\
V(I) &= \left[iz(MWMMW') + iz(MW)^2 + (iz(MW))^2 \right] / d - E(I)^2 \\
d &= (n-k)(n-k+2) \\
Z_I &= [I - E(I)] / V(I)^{1/2}
\end{aligned} \tag{4.22}$$

4.3.2. Tanımlayıcı İstatistikler ve Uzamsal Etkiler

Uzamsal etkilerin varlığı en küçük kareler kestirimlerinin geçersiz kabul edilmesine neden olmaktadır. Bu nedenle literatürde model tanımlamasına ilişkin çok sayıda testler bulunmaktadır. En bilinen olanlarından Lagrange çarpanı (Lagrange multiplier – LM), olabilirlik oranı (likelihood ratio – LR) ve Wald (W) istatistikleri kısıtlı ve kısıtlı olmayan kestirimler arasındaki uzaklık yardımıyla hesaplanmaktadır. Kısıtların kabul edilebilirliği (sıfır hipotezi) uzaklıkların genişliğine bağlıdır. Eğer çok büyükse kısıtlar reddedilecektir. LR testi hem kısıtlı, hem de kısıtlanmamış kestirimler gerektirirken, Wald testi sadece kısıtlanmamış kestirimleri, LM testi de sadece kısıtlı kestirimleri gerektirmektedir.

Uzamsal otokorelasyonu göstermek için kullanılan testlerden biri olan Wald testi,

$$\begin{aligned}
W &= \lambda^2 [t_2 + t_3 - (1/n)t_1^2] \\
t_1 &= iz(W.*B^{-1}) \\
t_2 &= iz(WB^{-1})^2 \\
t_3 &= iz(WB^{-1})'(WB^{-1})
\end{aligned} \tag{4.23}$$

olarak hesaplanır. Burada $B = (I_n - \lambda W)$; λ , uzamsal ilişkili hataların katsayısı ve $*$ ifadesi her bir nesne için matris çarpımını gösterir.

Uzamsal hata bağımlılığı için Burridge (1980) tarafından geliştirilen LM testi,

$$LM-Err = (e'We/s^2)^2 / T \tag{4.24}$$

şeklinde hesaplanır. Burada e , EKK artıkları vektörü; W , uzamsal ağırlık matrisi; $s^2 = e'e / N$ ve $T = iz(W'W + W^2)$ olarak gösterilmiştir. Bu istatistiğin bir serbestlik derecesi ile ki-kare dağılıma uygunluk gösterdiği bilinmektedir.

Uzamsal gecikmenin bağımlılığının testi Anselin (1988) tarafından ve klasik regresyon esaslarına dayanılarak,

$$LM-Lag = (e'Wy / s^2)^2 / (RJ_{\rho-\beta}) \quad (4.25)$$

olarak gösterilmiştir. Burada e , en küçük kareler artıkları vektörü; y , bağımlı değişken ve $RJ_{\rho-\beta} = [T + (WX\beta)'M(WX\beta) / s^2]$ olur. Bu eşitlikteki $WX\beta$, en küçük kareler regresyon modelinden kestirilmiş değerlerin uzamsal gecikmeleri ve $M = I - X(X'X)^{-1}X'$ izdüşüm matrisidir. Bu test istatistiğinin de bir önceki istatistik gibi bir serbestlik dereceli ki-kare dağılıma uygunluk gösterdiği bilinmektedir.

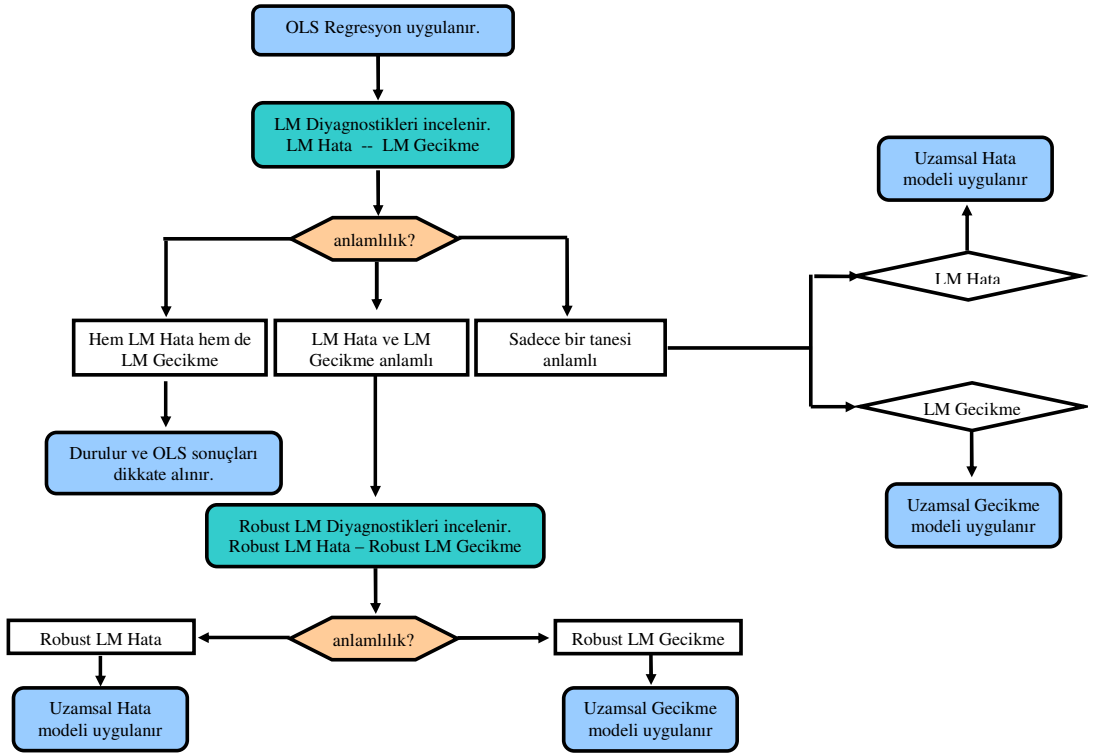
Tablo 4.1. Regresyon Modelleri için Uzamsal Bağımlılık Testleri

Test	Formülasyon	Dağılım	Kaynak
Moran	$e'We / e'e$	N(0,1)	Cliff ve Ord (1981)
LM-Err	$(e'We / s^2)^2 / T$	$\chi^2(1)$	Burridge (1980)
LM-Lag	$(e'Wy / s^2)^2 / (RJ_{\rho-\beta})$	$\chi^2(1)$	Anselin (1988)

(Kaynak: Varga, 1998, syf.39)

Tabloda regresyon modellerinde uzamsal bağımlılığın varlığının tespitinde kullanılan testler gösterilmiştir (Varga, 1998).

Uzamsal regresyon analizinin uygulanma süreci tabloda gösterilmiştir.



Şekil 4.4. Uzamsal Regresyon Analizi Süreci

Uzamsal regresyon analizinde süreç klasik regresyon analizinin uygulanması ile başlar. Modele ilişkin sonuçların elde edilmesinin ardından Lagrange Çarpanları istatistikleri incelenir. Bu adımda uzamsal hata veya uzamsal gecikmenin bağımlılığının incelendiği testlerin anlamlılıklarına bakılır. Eğer her iki test anlamsız bulunursa klasik modelden elde edilen sonuçların kullanılmasının uygun olduğuna karar verilir. Eğer sadece uzamsal hata bağımlılığının incelendiği test anlamlı sonuç verirse çalışmaya uzamsal hata modeli ile devam edilmesine, eğer sadece uzamsal gecikmenin bağımlılığının anlamlı olduğuna karar verilirse çalışmaya uzamsal gecikme modeliyle devam edileceğine karar verilir. Eğer her iki bağımlılığı da inceleyen test anlamlı çıkarsa bu adımda Robust Lagrange Çarpanı istatistikleri incelenir. Sürecin devamında daha önceki adımda olduğu gibi, eğer Robust Lagrange çarpanı hata bağımlılığının incelendiği test anlamlı çıkarsa uzamsal hata modeli, Robust Lagrange çarpanı uzamsal gecikmenin bağımlılığının incelendiği test anlamlı çıkarsa uzamsal gecikme modeli kullanılarak analize devam edilmez.

5. UZAMSAL REGRESYON ANALİZİNDE AĞIRLIK MATRİSLERİNE İLİŞKİN UYGULAMA

Bu uygulamada amaç uzamsal regresyon modellerinde kullanılacak farklı ağırlık matrisleri oluşturmak ve bu matrislerden elde edilen modelleri birbirleriyle karşılaştırmaktır. Bunun için “Global Information Technology Report 2007-2008” (URL-2, 2008) isimli rapor esas alınarak 127 ülkeye ait teknoloji ve teknolojik açıdan gelişebilme alt yapılarına ilişkin veriler derlenmiştir. Elde edilen değişkenler tabloda verildiği gibidir.

Tablo 5.1. Uygulamada Kullanılan Değişkenler

1. Ülkelerin alfabetik olarak sıra numaraları (ID Numaraları)
2. Ülkelerin isimleri
3. İnternet kullanıcıları oranları (her 100 hanede internet kullanılanların oranı)
4. En son teknolojinin ulaşılabilirliği (1=çok yaygın değil ve kullanılmaz, 7=çok yaygın ve kullanılır)
5. Toplam vergi oranı (Bu değer kazanç vergisi, emek vergisi ve diğer vergilerin tamamını içermektedir)
6. Telefon hatları sayısı (Her 100 haneye düşen telefon hattı sayısı)
7. Güvenli internet sunucuları sayısı (Her 1 milyon haneye düşen sunucu sayısı)
8. Elektrik üretimi
9. Eğitim harcamaları (GNI içerisinde eğitime ayrılan oran)
10. Kişisel bilgisayar sayısı (100 haneye düşen kişisel bilgisayar sayısı)
11. Yenilik kapasitesi (Teknolojiye sahip işletmeler, 1=işletmenin yeni lisanslar almaması, 7=işletmelerin düzenli olarak yenilikleri takip ederek yeni lisanslı teknolojik ürünler alması)
12. Ülkenin orta noktasının X koordinatları
13. Ülkenin orta noktasının Y koordinatları
14. Ülkenin yüzölçümü
15. İnsani Gelişim İndeks Değerleri (HDI 2007-2008)

Uygulamada GeoDa 0.9.5-i (URL-3, 2008) ve R (URL-1, 2008) programları kullanılmıştır. R programı istatistiksel hesaplamaları ve grafikleri elde etmek

amacıyla John Chambers tarafından Bell Laboratuvarlarında geliştirilmiş olan ve S dili kullanılarak kod yazma suretiyle çalıştırılan açık kaynaklı bir programdır. GeoDa programı Illinois Üniversitesi Uzamsal Analiz Laboratuvarlarında Luc Anselin tarafından oluşturulmuş olan ekip tarafından hazırlanmış olan ve uzamsal analizlerin uygulanmasını sağlayan bir programdır.

Verilerin analizlerinin başlanması için ilk adımda Excel formatında düzenlenen veriler özel bir program yardımıyla “comma seperated values file-CSV” uzantılı hale getirilmiştir. Buradan “not defteri”ne (notepad) aktarılan veriler her iki programda da kullanılabilir hale getirilmiştir.

Çalışma süresince internet kullanıcıları oranları değişkeni bağımlı değişken olarak ele alınmış ve farklı ağırlık matrisleri kullanılarak değişkenlerin bağımlı değişken üzerindeki etkileri, ülkelerin coğrafi konumları da dikkate alınarak incelenmiştir.

127 ülkeye ait veriler incelenirken bazı ülkelere ilişkin bir kısım değişkenlerin değerleri elde edilememiştir. Bu değişkenler için çalışmada eksik gözlemlerin olduğu söylenebilir. Eksik gözlemlerin bulunduğu değişkenler ve eksik gözlem sayıları aşağıdaki tabloda belirtildiği gibidir.

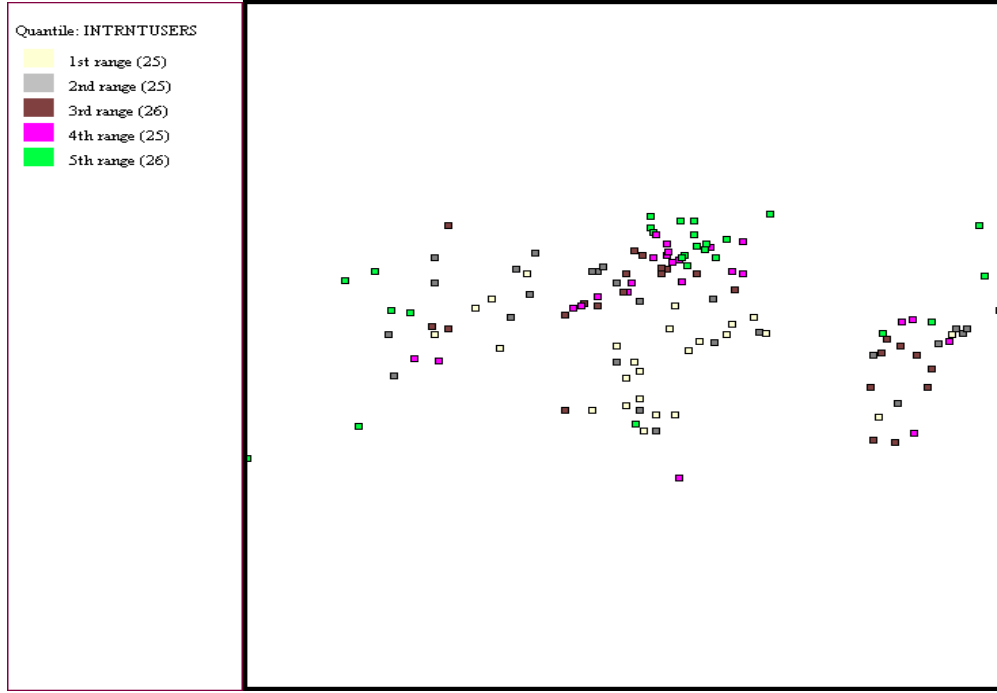
Tablo 5.2. Eksik Gözlemler İçeren Değişkenler

	Gözlem	Eksik	Toplam
Toplam vergi oranı	121	6	127
Güvenli internet sunucuları sayısı	124	3	127
Elektrik üretimi	113	14	127
Eğitim harcamaları	120	7	127
Kişisel bilgisayar sayısı	126	1	127

Çalışmanın ilk bölümünde mevcut eksik gözlemler olduğu gibi bırakılacak ve analizler eksik gözlemlerin varlığı altında yapılacaktır.

5.1. DEĞİŞKENLERDE EKSİK GÖZLEMLERİN OLMASI DURUMUNDA VERİ ANALİZLERİ

Çalışmanın ilk adımında koordinat sistemi kullanılarak elde edilen verilere ilişkin haritalama işlemi yapılmıştır. Bu işlemin yapılabilmesi için koordinat eksen düzeni göz önüne alınarak bölgelere ilişkin koordinatlar belirlenirken, Y eksenini üzerinde kuzey bölgeler +, güney bölgeler -, X eksenini üzerinde doğu bölgeler +, batı bölgeler - olarak kaydedilmiştir. Ülkelerin merkez noktalarının gösterildiği, bunun yanı sıra bağımlı değişken olarak incelenecek olan internet kullanım oranlarının en düşükten en yükseğe doğru sıralandığı temsili harita şeklindeki gibi olacaktır.



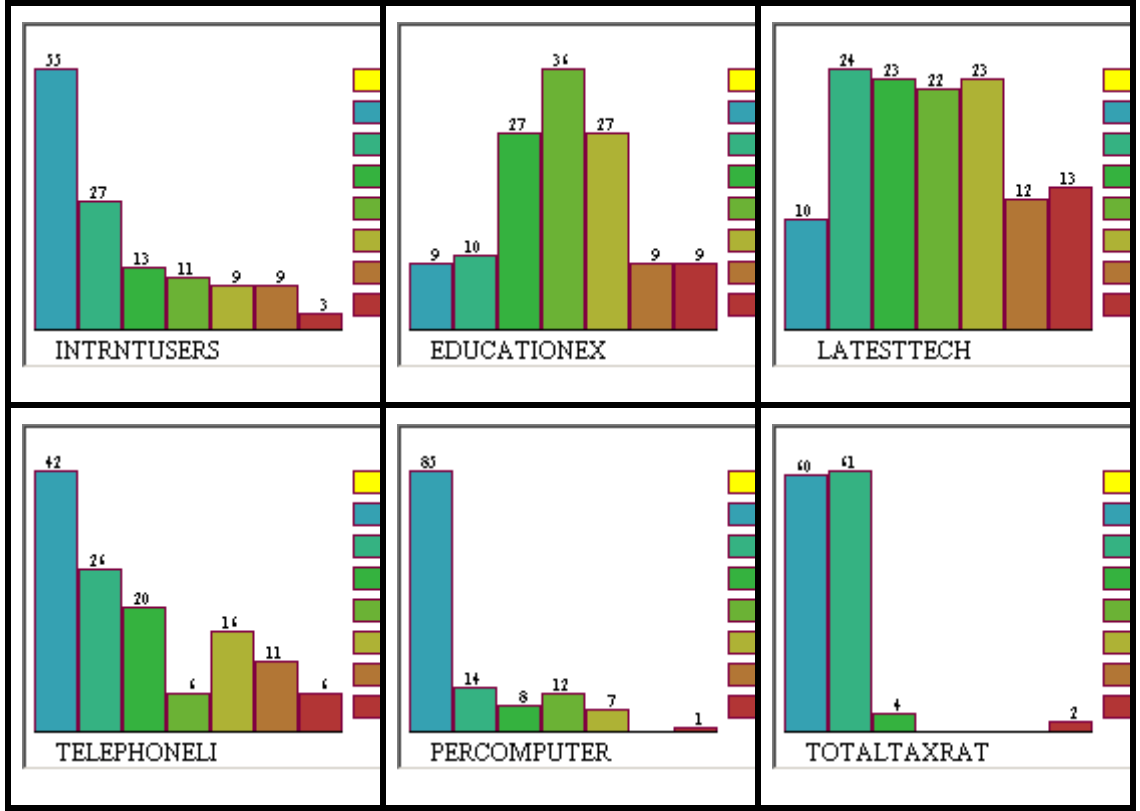
Şekil 5.1. Ülkelerin Merkez Noktalarının ve İnternet Kullanım Oranları Büyüklüklerinin Koordinat Sisteminde Temsili Gösterimleri

Temsili haritada ülkeler internet kullanım oranları, en az yüzdeye sahip olanlar 1, en yüksek yüzdeye sahip olanlar 5 olmak üzere, 5 ayrı gruba bölünmüştür. Mozambik, Tanzania ve Zambiya gibi ülkeleri internet kullanımının en az olduğu 1 numaralı sınıfta, Çin, Hindistan, Pakistan gibi ülkeleri 2 numaralı sınıfta, Türkiye, Arjantin, Uruguay gibi ülkeleri 3 numaralı sınıfta, Singapur, Malezya, Çek Cumhuriyeti, İspanya, Portekiz gibi ülkeleri 4 numaralı sınıfta ve Avustralya, Japonya, Kore gibi ülkeleri de internet kullanımının en yoğun olduğu 5 numaralı sınıfta görülmektedir. Ülkelerin buldukları sınıflara ilişkin bilgiler Ek-1'de gösterilmiştir. Çalışmanın devamında koordinatların dikkate alındığı durumda coğrafi konumların ülkelerin internet kullanımı üzerindeki etkileri incelenecektir.

Uzamsal analizlerde klasik analizlerde olduğu gibi açıklayıcı veri analizleri (EDA) uygulanabilir. Bu sayede değişkenlerin özellikleri ayrıntıları ile gösterilebilir. Açıklayıcı veri analizlerine değişkenlerin histogramlarının oluşturulması ile başlanabilir.

Şekil 5.2'de verilen histogramlar incelendiğinde ülkelerin sahip oldukları özellikler ayrıntıları ile görülebilir. İnternet kullanıcıları dikkate alındığında, 55 ülkenin internet kullanım oranınının 0,21-12,88 arasında değiştiği ve sadece 3 ülkenin (Yeni Zelanda, Hollanda, İsviçre) 76,20-88,87 oranında internet kullandığı söylenebilir. Eğitim harcamalarının yaklaşık simetrik olduğu ve GNI değerlerinin yüzde 3,46-4,63'lük değerini kapsadığı belirlenmiştir. İnternetin en fazla kullanıldığı 3 ülkeden sadece

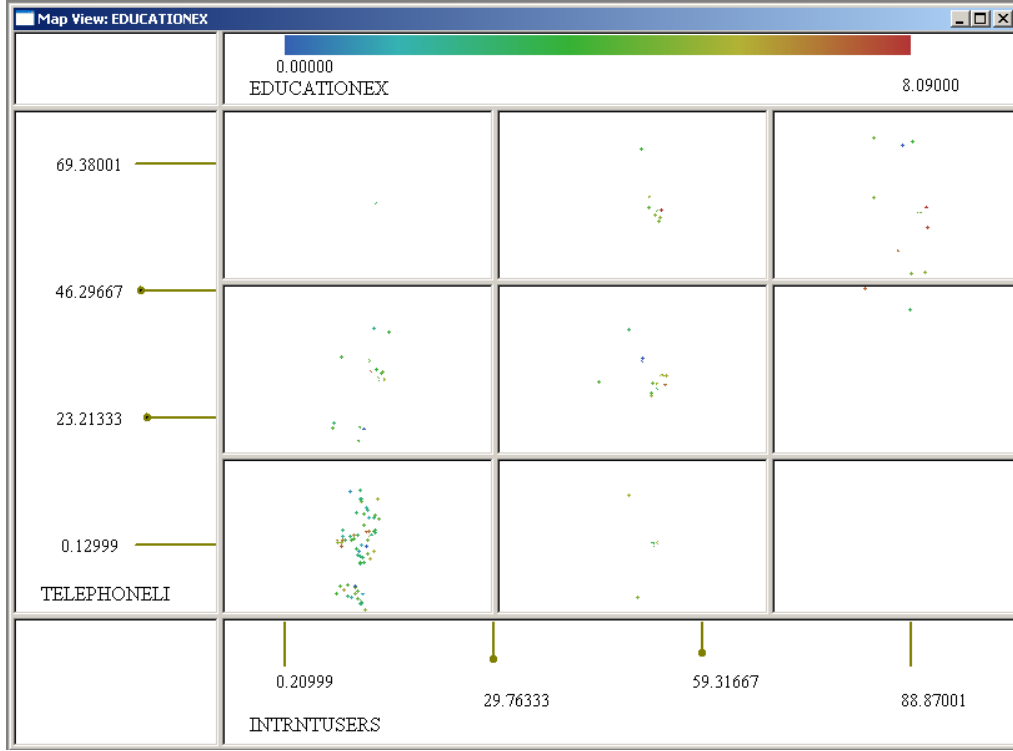
Hollanda'nın kırmızı ile gösterilen ve eğitim indeksinde en büyük oranı temsil eden ülkeler arasında olmadığı söylenebilir. Diğer değişkenlere de ilişkin histogramlar benzer şekillerde yorumlanabilir. Bağımlı değişken olarak kullanılacak olan internet kullanım oranı için Türkiye'nin konumu incelendiğinde 27 ülkenin bir arada yer aldığı yeşil sütunda olduğu söylenir.



Şekil 5.2. Değişkenlere İlişkin Histogramlar

Şekil 5.2. incelendiğinde değişkenlerin dağılımlarının birbirinden farklı oldukları açıkça görülmektedir. Çalışmanın ilerleyen aşamalarında çeşitli dönüşümlerle tüm değişkenler yaklaşık olarak simetrik hale getirilip sonuçlar elde edilecek, dönüştürülmemiş değişkenlerle elde edilecek olan sonuçlarla karşılaştırılacaktır.

İnternet kullanımını en fazla etkileyeceği düşünülen değişkenlerden telefon hatları sayısı ve eğitim harcamalarının etkisi bir arada Şekil 5.3'deki gibi gösterilebilir.



Şekil 5.3. İnternet Kullanımı, Telefon Hatları ve Eğitim Harcamaları Değişkenlerinin Bir Arada Gösterilmesi

Burada ilk kutuda telefon hatları sayısı yüksek olan sadece bir ülkenin internet kullanımının en düşük düzeyde olduğu ve eğitime ayırdığı bütçenin ortalama düzeyde olduğu gösterilmiştir. Ancak yoğun olarak telefon hattı sayısı az olan ülkelerde internet kullanımının da az olduğu görülmektedir.

5.1.1. Klasik Regresyon Analizinin Uygulanması

Uzamsal regresyon analizine ilişkin sürecin ilk adımında verilere klasik regresyon analizi uygulanmıştır. Analizin bu aşamasında varlığı tespit edilmiş olan eksik gözlemler çalışmadan çıkarılmış ve sürece bu şekilde devam edilmiştir. İnternet kullanım oranı bağımlı değişken, en son teknolojinin ulaşılabilirliği, toplam vergi oranı, telefon hatları sayısı, güvenli internet sunucuları sayısı, elektrik üretimi, eğitim harcamaları, kişisel bilgisayar sayısı ve yenilik kapasitesi değişkenleri bağımsız değişkenler olarak alınarak oluşturulan klasik regresyona ilişkin sonuçlar Tablo 5.3'de verildiği gibi hesaplanmıştır.

Tablo 5.3'de verilen sonuçlar incelendiğinde F istatistik değeri yardımıyla modelin anlamlı olduğu söylenebilir ($p < 0,05$). Modelin belirtme katsayısı (R^2) incelendiğinde ilgili bağımsız değişkenlerin bağımlı değişkenleri 0,80 oranında açıkladığı söylenebilir.

Tablo 5.3. Tüm Verilerle Klasik Regresyon Analizi Sonuçları

KLASİK REGRESYON ANALİZİ SONUÇLARI				
Bağımlı Değişkenin Ortalaması	:24.3808	Düzeltilmiş R Kare	:0.804154	
Bağımlı Değişkenin Standart Sapması	:22.4543	Artık Kareler Toplamı	:11744.3	
Gözlem Sayısı	:127	Regresyon Standart Hatası	:9.97637	
Değişken Sayısı	:9	F İstatistiği	:65.6704	
Serbestlik Derecesi	:118	P. (F istatistiği)	:7.25682e-040	
		Log Olabilirlik	:-467.666	
		AIC	:953.331	
		BIC	:978.929	
DEĞİŞKEN	KATSAYI	STD. HATA	T	OLASILIK
Sabit	-10.49744	6.061766	-1.731746	0.0859323
Teknoloji	2.310137	1.588157	1.454603	0.1484340
Vergi	0.000800543	0.0268334	0.02983381	0.9762075
Tel. Hattı	0.5179215	0.09277642	5.582468	0.0000002
Int. Sunucuları	0.006617293	0.00744308	0.8890526	0.3757841
Elek. Üretimi	0.00034767	0.00028049	1.239498	0.2176208
Eğt. Harcaması	0.4588355	0.5275182	0.8698003	0.3861759
Kişisel Bilg.	0.1954028	0.079766783	2.449644	0.0157685
Yenilik Kap.	1.437346	1.913495	0.7511628	0.4540488

Bu süreçte hesaplanan Akaike Bilgi Kriteri (AIC) ve Schwarz Kriteri (BIC) elde edilen bu modeli daha sonraki adımlarda elde edilen modellerle karşılaştırmada kullanılacaktır.

Modele ilişkin değişkenlerin anlamlılıkları incelendiğinde 0,10 anlamlılık düzeyi için sabit terim, telefon hatları sayısı ve kişisel bilgisayar sayısına ilişkin değişkenlerin anlamlı oldukları, diğer bütün değişkenlerin ise modelin açıklanmasında anlamsız oldukları belirlenmiştir.

Analizin bir sonraki adımında istatistiksel varsayımlardan sapmalara ilişkin ölçümler incelenmiş ve sonuçlar Tablo 5.4'deki gibi gösterilmiştir.

Tablo 5.4. Varsayımlardan Sapmalara İlişkin Ölçümler

REGRESYON İSTATİSTİKLERİ			
	Değer	Ser. Derecesi	Olasılık (p)
Çoklu Bağlantı Koşul Değeri	25.65733		
Jarque-Bera Testi	58.13228	2	0.0000000
DEĞİŞEN VARYANSLILIK İSTATİSTİKLERİ			
Breusch-Pagan Testi	145.1746	8	0,0000000
White Testi	95.68099	44	0,0000107

Çoklu bağlantı probleminin incelenmesini sağlayan çoklu bağlantı koşul değeri (multicollinearity condition number) için karşılaştırma kriteri 30 sayıdır. Tablo

5.4'deki kritik deęer incelendięinde $25,66 < 30$ olduęundan modelde oklu baęlantı problemi olmadığı sylenebilir.

Verilerde normallik ve deęişen varyanslılık varsayımlarının incelenmesi iin Jarque-Bera, Breusch-Pagan ve White testleri de yapılmıő, incelenen verilerde normallik ve deęişen varyanslılık varsayımlarından sapmalar olduęu belirlenmiőtir. Modelin uzamsal yapıda kurulması ile bu sapmaların ortadan kalkacaęı dőünölmektedir.

Uygulamada kullanılan verilerin lkelere ait olması nedeniyle lkelerin konumlarının ve komőuluk iliőkilerinin internet kullanımını etkileyebileceęi dőünölmektedir. Bu nedenle alıőmaya uzamsal olarak devam edilip edilemeyeceęini lmek amacıyla uzamsal baęımlılıęın lldüęü istatistiklerin incelenmesine karar verilmiőt ve bulunan sonular tablodaki gibi gsterilmiőtir.

Tablo 5.5. Uzamsal Baęımlılık İstatistikleri

UZAMSAL BAęIMLILIK İSTATİSTİKLERİ			
	Deęer	S.D.	Olasılık (p)
Moran I (hata)	2.2086426		0.0271994
Lagrange arpanı (gecikme-lag)	3.7703163	1	0.0501698
Robust LM (gecikme)	2.0081727	1	0.1564537
Lagrange arpanı (hata-error)	1.9447825	1	0.1631502
Robust LM (hata)	0.1826389	1	0.6691152

Elde edilen istatistikler incelendięinde Lagrange Gecikme arpanı iin elde edilen deęerin anlamlı ıkmasından dolayı alıőmaya farklı aęırlık matrisleri kullanmak suretiyle uzamsal gecikme modeli oluőturularak devam edilmesinin uygun olduęuna karar verilmiőtir.

5.1.2. Farklı Aęırlık Matrisleri Kullanılarak Uzamsal Gecikme Modelinin Elde Edilmesi

Uzamsal gecikme modelini kurarak alıőmaya devam edilebilmesi iin bir aęırlık matrisinin oluőturulması gerekmektedir. Aęırlık matrisleri oluőturmak iin eőttili yollar vardır. Ancak literatürde en sık kullanılanları eksik gözlemlerin söz konusu olduęu durumlarda bile kullanılabilen en yakın komőuluk matrisi (k-nearest neighbor) ve eőtik uzaklık matrisi (threshold distance) aęırlıklarıdır. alıőmanın devamında aęırlık matrisleri oluőturulmuőt ve uzamsal gecikme modeline iliőklin sonular elde edilmiőtir.

Oluőturulan en yakın komőuluk matrisi, eőtik uzaklık matrisi ve aynı veri seti iin elde edilen klasik modele iliőklin sonular Tablo 5.6'da gsterilmiőtir.

Tablo 5.6. Farklı Modellere İlişkin Sonuçlar

	KLASİK MODEL SONUÇLARI		EN YAKIN KOMŞULUK MATRİSLİ MODEL SONUÇLARI		EŞİK UZAKLIK MATRİSLİ MODEL SONUÇLARI		
	Değer	Olasılık	Değer	Olasılık	Değer	Olasılık	
Karşılaştırma Ölçütleri	R kare	0.816589		0.823055		0.828102	
	Standart Hata	9.97637		9.44536		9.30968	
	Log Olabilirlik	-467.666		-465.732		-463.797	
	AIC	953.331		951.464		947.593	
	BIC	978.929		979.906		976.035	
Katsayılar	Ağırlık (W)			0.1277596	0.0680974	0.2958319	0.0028981
	Sabit	-10.49744	0.0859323	-10.17981	0.0761722	-15.59923	0.0080048
	Teknoloji	2.310137	0.1484340	2.448931	0.1033879	2.626784	0.0777942
	Vergi	0.000800543	0.9762075	0.004340	0.8651882	-0.005015	0.8417952
	Tel. Hatları	0.5179215	0.0000002	0.4718737	0.0000002	0.4809921	0.0000000
	Int. Sunucuları	0.006617293	0.3757841	0.0071850	0.3079263	0.0035029	0.6153185
	Elek. Üretimi	0.00034767	0.2176208	0.0002719	0.3086688	0.0002742	0.2998570
	Eğt. Harcaması	0.4588355	0.3861759	0.3673445	0.4652327	0.3650523	0.4594679
	Kişisel Bilg.	0.1954028	0.0157685	0.2085429	0.0058829	0.2057405	0.0057651
	Yenilik Kap.	1.437346	0.4540488	0.673839	0.7145673	1.088569	0.5448376
Test Değeri	Breusch-Pagan Testi	145.1746	0,0000000	126.544	0,0000000	151.4651	0,0000000
	Olabilirlik Oran Testi			3.866888	0.0492478	7.73804	0.0054069

5.1.2.1. En Yakın Komşuluk Matrisi ile Model

Verilere klasik regresyon analizinin uygulanmasının ardından k en yakın komşuluk matrisi oluşturulmuş ve ülkelerin birbirlerine göre konumlarının internet kullanımı üzerindeki etkileri diğer değişkenlerle birlikte incelenmeye çalışılmıştır. En yakın komşuluk matrisinin aldığı değerler Ek-2'de gösterilmiştir.

En yakın komşuluk matrisi oluşturulurken ilk adımda tamamen araştırmacının tercihinin bırakılmış olan ve k ile gösterilen ($k > 0$) komşu sayısı belirlenir. Belirlenen ülkeler arasındaki uzaklıklar, ülke koordinatlarının koordinat eksenine yerleştirilerek Pisagor teoremi kullanılması ile hesaplanmıştır. En yakın komşuluklar belirlenirken veri kümesinin dünyadaki birçok ülkeyi kapsamaması nedeniyle her bir ülkenin sahip olabileceği komşu sayısı 5 olarak kısıtlanmıştır. Burada sınırdaşlık gözetenilmeksizin, koordinatlar yardımıyla her bir ülkeye en yakın 5 ülke komşu olarak belirlenmiştir.

Bu bilgiler ışığında AIC dikkate alındığında değerlerin daha düşük çıkmasından dolayı, en yakın komşuluk matrisiyle oluşturulan gecikme modelinin klasik modelden daha iyi sonuçlar verdiği söylenebilmektedir. Katsayılar incelendiğinde ise 0,10 anlamlılık düzeyinde ağırlık, sabit terim, telefon hatları sayısı ve evlerdeki kişisel bilgisayar sayısı değişkenlerinin anlamlı çıktıkları söylenir.

Analizin bir sonraki adımında modelin geçerliliğinin sınanması açısından regresyona ilişkin istatistikler elde edilmiştir. Breusch-Pagan testi yardımıyla modeldeki değişen varyanslılık sorunu incelenir. Bu test sonucunda elde edilen istatistik değeri dikkate alındığında modelde değişen varyanslılık probleminin olduğu söylenebilir. Bunun yanı sıra otoregresif katsayının anlamlılığını ölçmeye yarayan olabilirlik oran testi sonucuna göre incelenen katsayı anlamlıdır.

5.1.2.2. Eşik Uzaklık Matrisi ile Model

Klasik regresyon uygulanmış verilere uzamsal gecikme modelini uygulamak üzere eşik değeri yardımıyla bir uzaklık (komşuluk) matrisi oluşturulmuştur. Verilerdeki koordinatlar dikkate alındığında eşik değeri çalışma için optimum olarak belirlenen 43,32435 olarak alınmıştır. Bu matrisin aldığı değerler Ek-3'de gösterilmiştir. Ülkeler arasındaki uzaklıklar hesaplanırken, en yakın komşuluk matrisinin hesaplanmasında olduğu gibi, koordinat eksenini üzerindeki değerlerinin Pisagor teorisi ile çözümlenmesi süreci uygulanmıştır. Ülkelerin konumları birbirine göre incelenmiş ve uzaklıklar yardımıyla hesaplanan değerleri eşik değeri olarak belirlenmiş değerden daha düşük çıkan ülkeler komşu olarak belirlenmiş ve komşu sayısında herhangi bir kısıtlamaya gidilmemiştir.

Tablo 5.6'da gösterilen modele ilişkin istatistikler incelendiğinde AIC değerinin elde edilmiş olan diğer modellerden daha düşük çıktığı, dolayısıyla daha iyi sonuç verdiği belirlenmiştir. Bunun yanı sıra model katsayıları incelendiğinde ağırlık katsayısının, sabit terimin, en son teknolojinin ulaşılabilirliği, telefon hattı sayısı ve kişisel bilgisayar oranı değişkenlerinin 0,10 için anlamlı çıktığı söylenebilmektedir.

Varsayımlardan sapmaların incelenmesi için elde edilen istatistiklerden yola çıkarak, Breusch-Pagan test sonucu yardımıyla elde edilen modelde değişen varyanslılık sorununun olduğu belirlenmiştir. Bunun yanı sıra olabilirlik oran testi yardımıyla otoregresif katsayının anlamlı çıktığı söylenebilmektedir.

5.1.2.3. Oluşturulan Modellerin Birbirleri ile Karşılaştırılması

Ağırlık matrisi kullanılarak oluşturulmuş olan iki model de istatistiksel açıdan anlamlı çıkmıştır. Ancak farklı ağırlık matrisleri kullanılarak elde edilen tüm modellerde değişen varyanslılık sorunundan kurtulunamamıştır. Buna karşılık en yakın komşuluk ve eşik uzaklık matrisli modellerde otoregresif katsayı anlamlı çıkmıştır.

Varsayımlardan sapmaların düzeltilemediği modellerin kestirimlerde kullanması problemlerin ortaya çıkmasına neden olacaktır.

5.2. DEĞİŞKENLERDE EKSİK GÖZLEMLERİN GİDERİLDİĞİ DURUMUNDA VERİ ANALİZLERİ

Çalışmanın bu aşamasında uygulamanın ilk kısmında belirtilen eksik gözlemlerin istatistiksel yöntemlerle giderilmesi ve daha önce incelenmiş olan modelleri, tüm veri setiyle oluşturulmuş olanlarla karşılaştırılması amaçlanmaktadır. Ancak bu değerler bulunmadan önce doğrusal regresyon analizinin en önemli varsayımlarından biri olan değişkenlerin normal dağılıma sahip olma varsayımı incelenecektir. Gerektiği durumlarda değişkenleri normal dağılıma yaklaştırmak için dönüşümler yapılacaktır.

Kullanılan değişkenlerin dağılımlarının gözlenebileceği şekiller Şekil 5.4'de gösterilmiştir. Şekilden de görüleceği üzere değişkenlerin birçoğu normal dağılıma yakınlık göstermemektedir.

Değişkenleri normal dağılıma yaklaştırmak için yapılması uygun görülen dönüşümler Tablo 5.7'de gösterildiği gibi belirlenmiştir.

Tablo 5.7. Değişkenler ve Normale Yaklaştırmak için Yapılan Dönüşümler (x^p)

Değişken Adı	Dönüşüm (p)
İnternet kullanıcıları oranları	1/3
En son teknolojinin ulaşılabilirliği	-
Toplam vergi oranı	Log
Telefon hatları sayısı	1/2
Güvenli internet sunucuları sayısı	Log(SecInt+1)
Elektrik üretimi	- *
Eğitim harcamaları	-
Kişisel bilgisayar sayısı	1/3
Yenilik kapasitesi	-

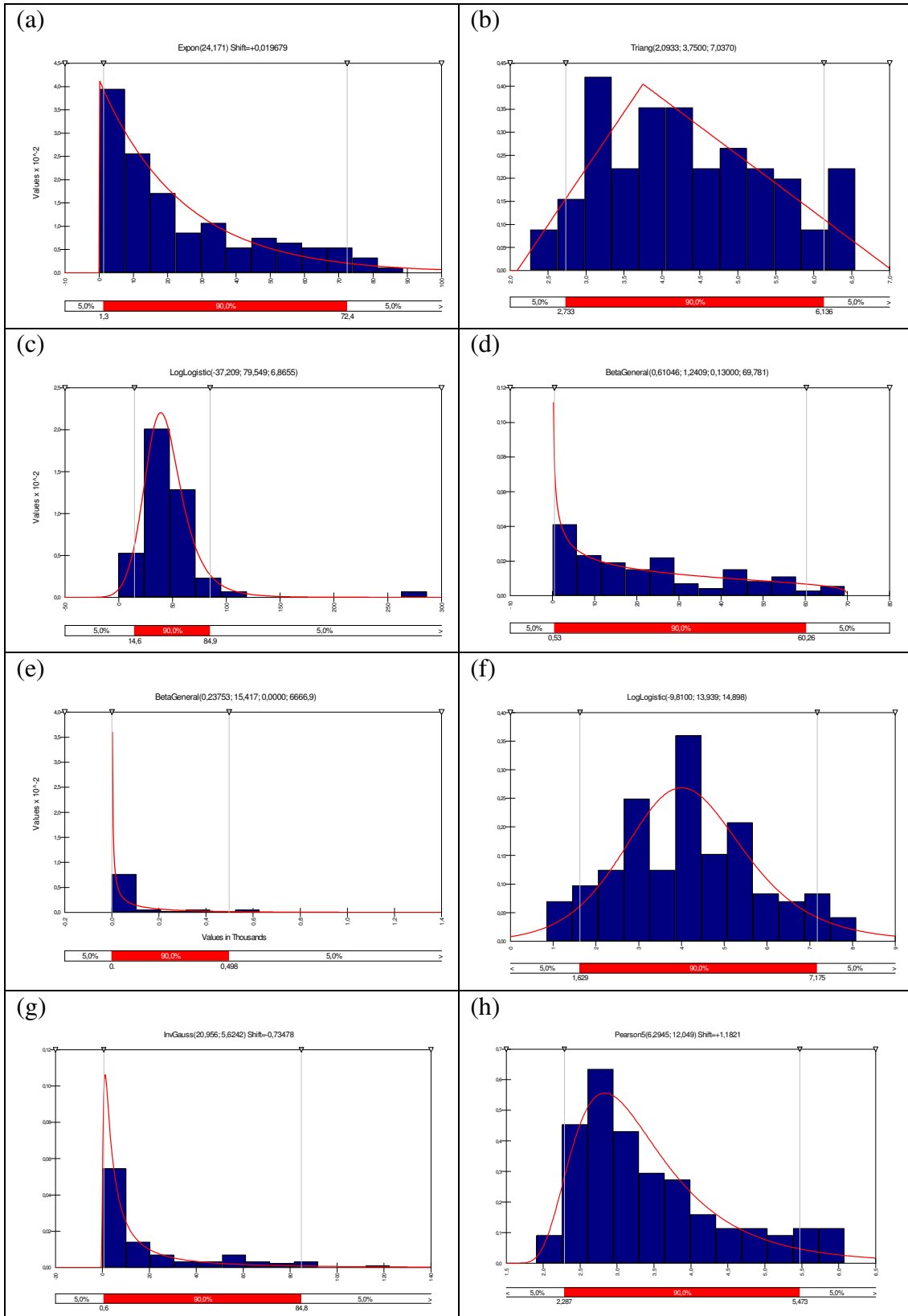
** Kukla değişken kullanılmıştır.*

Tabloda belirtilen dönüşümler yapıldığında değişkenler yaklaşık simetrik duruma gelmiştir. Ancak elektrik üretimi değişkeninde özel bir durum söz konusudur. Bu değişkene birçok dönüşüm uygulanmış ancak simetrik durum sağlanamamıştır. Denemeler sonucunda seri ancak bir kukla değişken eklenmesi sonucunda yaklaşık simetrik hale getirilebilmiştir. Bu kukla değişken oluşturulurken şekil olarak,

$$\text{Elektrik üretimi} < 10000 \text{ ise,} \quad \text{GEP} = 1$$

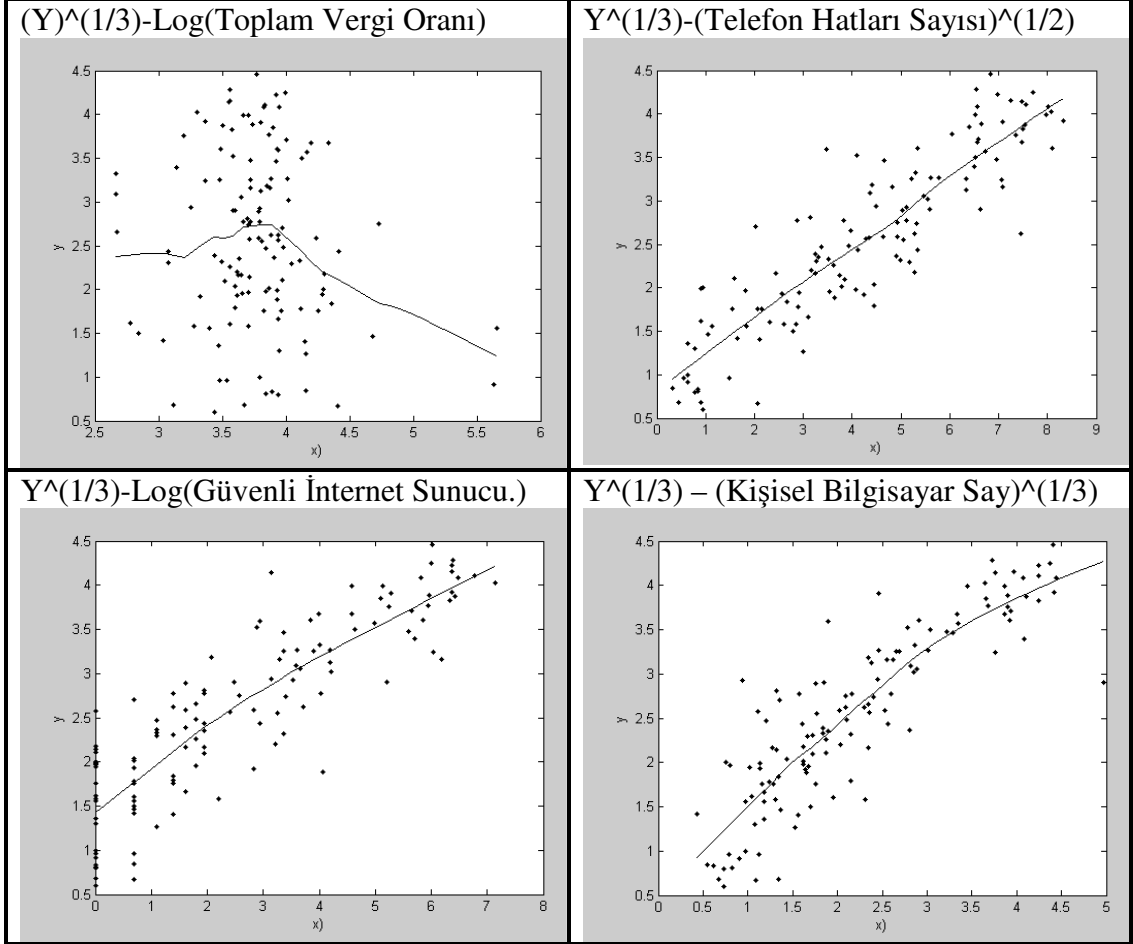
$$\text{Diğer durumlarda,} \quad \text{GEP} = 0$$

kısıtları esas alınmıştır.



Şekil 5.4. Kullanılan Değişkenlerin Dağılımlarının Gösterimi

Değişkenler arasındaki bağıntıları incelemek üzere düzleştirme (smoothing) çizimleri de dikkate alınmıştır. Bağımlı değişken olarak alınan internet kullanıcıları oranı değişkeni ile dönüşüm uygulanmış değişkenlerin düzleştirme eğrileri Şekil 5.5'te verilmiştir.



Şekil 5.5. Bağımlı Değişken ile Dönüşüm Uygulanmış Değişkenlerin Düzleştirme Eğrileri

Şekil 5.5'te de görüleceği gibi dönüştürülmüş bağımlı değişken ile telefon hatları sayısı, güvenli internet sunucuları sayısı ve kişisel bilgisayar sayısı değişkenleri arasında yaklaşık bir doğrusallık sağlanmasına karşın, toplam vergi oranı değişkeninde bu doğrusallık tam olarak sağlanamamıştır. Çalışmanın ilerleyen aşamalarında modele katkısının olmadığı belirlenen bu değişken çalışmadan çıkarılacak ve sonuçlar mevcut değişkenler ile değerlendirilecektir.

Tüm değişkenlerde yaklaşık simetriklik sağlandığından bir sonraki adımda eksik gözlemlerin ortadan kaldırılmasına yönelik işlemler gerçekleştirilecektir. Çalışmalarda eksik gözlemleri veriden çıkarmak bir yöntem olarak gösterilebilir. Ancak bu çalışmada Tablo 5.2'de gösterilen ve 31 eksik gözleme sahip verinin

çalışmadan çıkarılmasının, sonucu etkileyeceği düşünülmüştür. Bu nedenle eksik verileri istatistiksel yöntemlerle tahmin etme yoluna gidilmiştir.

Bu değerleri elde edebilmek için SPSS 13.0 paket programı yardımıyla kullanılan “missing value analysis” modülü uygulanmıştır. Bu modül içinde değişkenlerin ortalaması, kovaryans matrisleri ve nicel değişkenlerin korelasyonlarını dikkate alarak değişkenlerdeki eksik değerlerin hesaplanmasını sağlayan EM algoritmasının kullanılması tercih edilmiştir. Böylece veri kümesindeki eksik gözlem içermeyen değişkenler ile eksik gözlemlili değişkenler regresyona tabi tutulup eksik değerler kestirilmiştir. Burada modele yerleştirilirken yaklaşık simetrik duruma gelmiş değişkenler kullanılmıştır.

Çalışmanın bundan sonraki aşamalarında normal hale getirilmiş ve eksik gözlem içermeyen değişkenler kullanılmıştır. Bu adımda SPSS ile incelenmiş olan veriler “.CSV” uzantılı dosya haline getirilerek GeoDa programına aktarılmış ve analizlere bu program ile devam edilmiştir.

5.2.1. Klasik Regresyon Analizinin Uygulanması

Eksik gözlemlerin giderilmesinin ardından uzamsal modellere geçmeden önce ilk adımda klasik regresyon analizi uygulanmış Tablo 5.8’deki sonuçlar elde edilmiştir.

Tablo 5.8. Eksik Gözlemlerin Giderildiği Klasik Regresyon Analizi Sonuçları

	KLASİK MODEL SONUÇLARI		EN YAKIN KOMŞULUK MATRİSLİ MODEL SONUÇLARI		EŞİK UZAKLIK MATRİSLİ MODEL SONUÇLARI		
	Değer	Olasılık	Değer	Olasılık	Değer	Olasılık	
Karşılaştırma Ölçütleri	R kare	0.843119		0.857130		0.859405	
	Standart Hata	0.393557		0.374091		0.3711	
	Log Olabilirlik	-56.5663		-55.5128		-54.4256	
	AIC	133.133		133.026		130.851	
	BIC	161.574		164.312		162.137	
Katsayılar	Ağırlık (W)			0.0799199	0.1821744	0.2053867	0.0330697
	Sabit	0.3309628	0.4716662	0.1908699	0.6684311	-0.0642284	0.8910308
	Teknoloji	0.115366	0.1031907	0.1243927	0.0624972	0.1308513	0.0493548
	Vergi	0.05470207	0.5303705	0.0644237	0.4382062	0.0080858	0.9235697
	Tel. Hattı	0.2312506	0.0000001	0.2175152	0.0000000	0.2136485	0.0000000
	Int. Sunucuları	0.0599894	0.2152553	0.0581521	0.2038409	0.0494235	0.2767336
	Elek. Üretimi	-9.0133e-006	0.5917353	-1.13e-005	0.4799405	-6.99e-006	0.6583850
	Kukla_Elek.Üret.	-0.1584583	0.4634634	-0.1759663	0.3919935	-0.0700737	0.7343802
	Eğt. Harcaması	0.05224849	0.0537942	0.0483361	0.0605509	0.0522386	0.0388834
	Kişisel Bilg.	0.1644062	0.0774469	0.164928	0.0601069	0.1750004	0.0445616
	Yenilik Kap.	-0.00635404	0.9330925	-0.0155118	0.8293815	-0.0136931	0.8480563
Test Değeri	Breusch-Pagan Testi	15.85505	0.0699709	5.184156	0.8179684	9.463037	0.3956744
	Olabilirlik Oran Testi			2.106958	0.1466305	4.281255	0.0385349

Anlamlı olarak bulunan modelin belirtme katsayısı incelendiğinde, dönüşüm yapılmamış verilerle oluşturulan modelden daha yüksek bir değere sahip olduğu görülmektedir. Değişkenler incelendiğinde ise telefon hatları sayısı, eğitim harcamaları ve kişisel bilgisayar sayısı değişkenlerinin 0,10 anlamlılık ile modeli açıklayan değişkenler olduğu söylenir.

Tablo 5.9. Eksik Gözlem Olmasızın Varsayımlardan Sapmaların İncelenmesi

REGRESYON İSTATİSTİKLERİ			
	Değer	Ser. Derecesi	Olasılık (p)
Çoklu Bağlantı Koşul Değeri	49.40889		
Jarque-Bera Testi	1.275919	2	0.5283695
DEĞİŞEN VARYANSLILIK İSTATİSTİKLERİ			
Breusch-Pagan Testi	15.85505	9	0.0699709

Analizin bir sonraki adımında klasik istatistiksel varsayımlardan sapmalara ilişkin ölçümler incelenmiş ve sonuçlar Tablo 5.9'daki gibi gösterilmiştir. Değerler incelendiğinde çoklu bağlantı koşul sayısının 49,40889 olarak hesaplanması (MCN>30) modelde çoklu bağlantı problemi olmakla birlikte çok güçlü olmadığını göstermektedir. Jarque-Bera test istatistiği yardımıyla modelden elde edilen hataların normal dağılıma sahip oldukları ve Breusch-Pagan test istatistiği ile modelde eş varyanslı bir yapı olduğu söylenir.

Modeldeki varsayımdan sapmaların ve ülkelerin komşuluk durumlarının dikkate alınması ile uzamsal bağımlılığın incelenmesi için istatistikler hesaplanmış ve Tablo 5.10'da gösterilmiştir.

Tablo 5.10. Eksik Gözlem Olmaması Durumunda Uzamsal Bağımlılık İstatistikleri

UZAMSAL BAĞIMLILIK İSTATİSTİKLERİ			
	Değer	S.D.	Olasılık (p)
Moran I (hata)	2.2086426		0.0271994
Lagrange Çarpanı (gecikme-lag)	3.7703163	1	0.0501698
Robust LM (gecikme)	2.0081727	1	0.1564537
Lagrange Çarpanı (hata-error)	1.9447825	1	0.1631502
Robust LM (hata)	0.1826389	1	0.6691152

Uzamsal analizlerin kullanımlarının uzamsal otokorelasyonun varlığı durumunda özellikle gerekli olduğu bilinir. Uzamsal otokorelasyonun varlığının incelenmesi için Moran I değeri kullanılmıştır. Elde edilen Moran I test istatistiğine göre verilerde pozitif otokorelasyon olduğu söylenir. Bu durumda benzer değerlerin uzayda bir arada kümelendikleri de söylenir. Bu nedenden dolayı bu verilere uzamsal regresyon analizinin uygulanması ve kümelenmenin neden olduğu etkiyi, coğrafi etkileri de modele ekleyerek ortadan kaldırılmasının uygun olduğu istatistiksel olarak da gösterilmiştir.

Elde edilen istatistikler incelendiğinde Lagrange Gecikme Çarpanı için elde edilen değer anlamlı çıkmasından dolayı çalışmaya farklı ağırlık matrisleri kullanmak suretiyle uzamsal gecikme modeli oluşturularak devam edilmesinin uygun olduğuna karar verilmiştir.

5.2.2. Farklı Ağırlık Matrisleri Kullanılarak Uzamsal Gecikme Modelinin Elde Edilmesi

Eksik gözlemlerin olması durumunda izlenen süreç, eksik gözlemlerin giderilmesi ile elde edilen verilere aynen uygulanmıştır. Bu nedenle klasik modelin elde edildiği verilere en yakın komşuluk matrisi ve eşik uzaklık matrisleri eklenerek uzamsal gecikme modelleri elde edilmiş ve yorumlanmıştır.

5.2.2.1. En Yakın Komşuluk Matrisi ile Model

Bu aşamada kullanılan matris, daha önce açıklanan ve sınır gözetmeksizin birbirine en yakın olan ülkelerin 4 tanesini komşu kabul ederek oluşturulmuştur. Yapılan analizler sonucunda elde edilen modele ilişkin sonuçlar Tablo 5.8'de gösterilmiştir.

Model için hesaplanmış olan ve yapay R^2 'ye eşdeğer olan ifade 0,86 olarak bulunmuştur. Model için elde edilen sonuçlar klasik model ile elde edilen sonuçlarla karşılaştırıldığında, hesaplanan olabilirlik oran değerinin klasik modelde hesaplanan değerden daha büyük ve AIC değerinin de yine klasik model için hesaplanan değerden daha düşük olduğu belirlenmiştir. Bu da en yakın komşuluk matrisiyle elde edilen modelin klasik modelden daha iyi sonuçlar içerdiğini göstermektedir.

Değişkenlerin anlamlılıkları incelendiğinde ise en son teknolojinin ulaşılabilirliği, telefon hattı sayısı, eğitim harcamaları ve kişisel bilgisayar sayısı değişkenleri model için anlamlı değişkenler olarak belirlenmişlerdir.

İstatistiksel varsayımlardan sapmaların incelenmesi için uygulanan testlerden biri olan Breusch-Pagan testi yardımıyla ($p>0,05$) verilerin eş varyanslı bir yapıya sahip oldukları söylenebilir. Burada gösterilen olabilirlik oran testi yardımıyla da modelde bulunan uzamsal otoregresif katsayının anlamsız çıktığı söylenebilir. Dolayısıyla gecikme modeli olarak kurulan modelin kullanılması uygun bulunmamaktadır.

5.2.2.2. Eşik Uzaklık Matrisi ile Model

Çalışmanın bu bölümünde coğrafi etkileri daha iyi gözlemleyebilmek için eşik değeri 43,32435 olarak kabul edilmiş ve satır standartlaştırılmış ağırlık matrisi modele dahil edilmiştir. Model için hesaplanan katsayılar ve istatistikler Tablo 5.8'de gösterilmiştir.

Model için belirtme katsayısı 0,86 olarak hesaplanmıştır. Oluşturulan model aynı verilere uygulanan klasik model ile karşılaştırıldığında, modelin olabilirlik oran değeri klasik modelden daha yüksek ve AIC değeri de klasik modeldeki AIC değerinden daha düşük çıkmıştır.

Değişkenlerin anlamlılıkları incelendiğinde ise ağırlık, son teknolojinin ulaşılabilirliği, telefon hattı sayısı, eğitim harcamaları ve kişisel bilgisayar sayısı değişkenleri model için anlamlı değişkenler olarak belirlenmişlerdir.

Varsayımlardan sapmalara ilişkin testler incelendiğinde ise, Breusch-Pagan testi yardımıyla ($p > 0,05$) verilerin eş varyanslı bir yapıya sahip oldukları söylenebilir. Hesaplanmış olan olabilirlik oran testi yardımıyla da uzamsal otoregresif katsayılarının anlamlı olduğu söylenebilir.

5.2.2.3. Oluşturulan Modellerin Birbirleri ile Karşılaştırılması

Eksik gözlemlerin giderilmesi sonucunda oluşturulan klasik model, en yakın komşuluk matrisli uzamsal gecikme modeli ve eşik uzaklık matrisli uzamsal gecikme modeli birbirleri ile karşılaştırıldığında gerek log olabilirlik değerleri, gerek yapay R kare değerleri, gerekse AIC değerleri bakımından, eşik uzaklık matrisi kullanılarak oluşturulan uzamsal gecikme modelinin coğrafi etkilerin modele dahil edilmesi nedeniyle tercih edilmesi gereken model olduğuna karar verilmiştir.

5.3. AĞIRLIK FONKSİYONLARI YARDIMIYLA AĞIRLIK MATRİSLERİNİN OLUŞTURULMASI

Uzamsal modellerin oluşturulmasında literatürde çok sık kullanılan eşik uzaklık ağırlık matrisi ve en yakın komşuluk ağırlık matrislerinin yanı sıra özel fonksiyonların kullanılması ile elde edilen değerler yardımıyla oluşturulan matrisler de kullanılmaktadır. Belirlenen komşuluklara ilişkin ağırlıkların oluşturulmasında fonksiyonların kullanılması araştırmacıların çok tercih ettikleri bir yöntem olmamakla birlikte, sonuçların sadece sınırdışlık veya uzaklık durumuna göre değil, incelenen bölgedeki daha özel durumların ortaya çıkarılması için farklı etmenlerin de dahil edilmesi bakımından önem verilmesi gereken bir konudur. Ağırlık fonksiyonları oluşturulurken karşılaşılan en büyük sorun fonksiyonun, incelenen birimler arasındaki uzaklığın artması ile etkileşimin azaldığını gösterecek şekilde düzenlenebilmesidir.

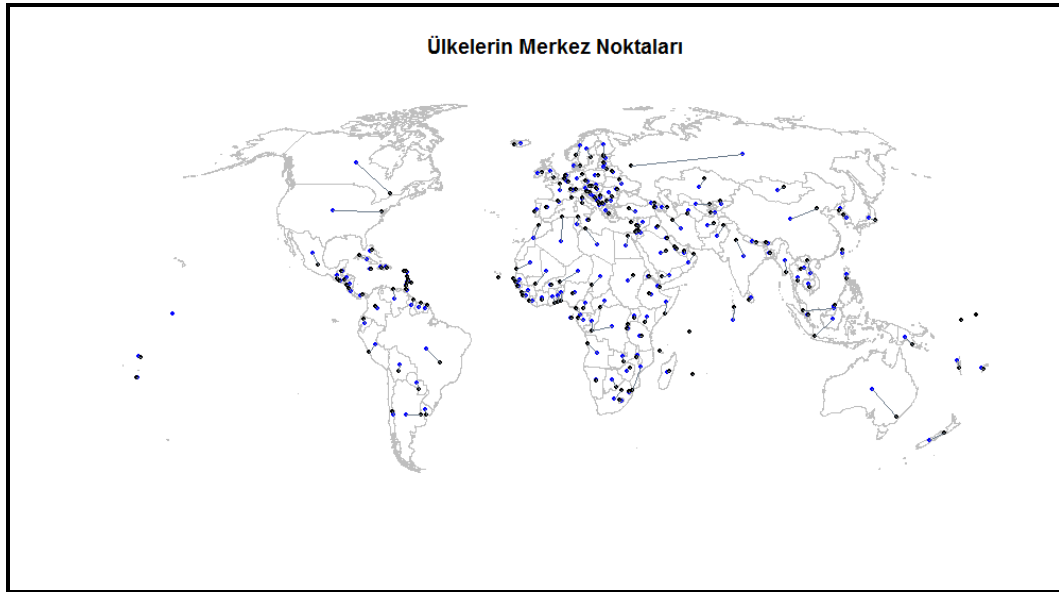
Çalışmanın bu aşamasında fonksiyonlar yardımıyla hesaplanacak olan matrislerin kullanılması ile uzamsal modellerin oluşturulması amaçlanmıştır. Oluşturulan

fonksiyonlar ilk olarak sınırdaş hiç komşusu olmayan ada ülkelerinin çalışmadan çıkarılması ile oluşturulmuş verilere, ikinci olarak da ada ülkeleri dahil her ülke kendisine en yakın veya sınırdaş olan en az bir ülke ile komşu kabul edilerek oluşturulan verilere uygulanmıştır. Ancak ada ülkeleri dahil her ülkenin en az bir komşu ile ilişkilendirildiği verilerle farklı ağırlık fonksiyonları için oluşturulan modellerin tamamı anlamsız çıktığından bu çalışmada açıklanmamıştır. Çalışma boyunca birçok farklı ağırlık fonksiyonu uygulanmış ancak anlamlı bulunan üç tanesi ayrıntıları ile anlatılmıştır. Kullanılan veri setine ilişkin çalışılan ağırlık isimleri Tablo 5.11'de gösterilmiştir.

Tablo 5.11. Veriler ve Kullanılan Fonksiyonların İsimleri

VERİLER	AĞIRLIKLAR		
Hiç komşusu olmayan ada ülkeleri çalışmadan çıkarılmıştır.	w_1	w_2	w_3

Kullanılan tüm ağırlık matrislerinde ortak olarak, incelenen ülkelerin birbirlerine uzaklık değerleri bulunmaktadır. Bu uzaklık değerleri dünya atlası esas alınarak ülkelerin merkez noktalarının (centroids) birbirlerine uzaklıkları hesaplanarak oluşturulmuştur. Merkez noktaların elde edilmesi ve bundan sonra incelenecek olan ağırlık fonksiyonlarına bağlı tüm analizler "R" programı ile gerçekleştirilmiştir. Programının çalıştırılması sürecinde kullanılan kodlardan bazıları Ek-7'de verilmiştir. Uzaklık değerlerinin santimetre cinsinden elde edilmesi için belirlenen, ülkelerin merkez uzaklıklarına ilişkin şekil aşağıda gösterilmiştir.



Şekil 5.6. Ülkelerin Başkentleri ve Merkez Noktalarının Birbirlerine Uzaklıkları

Uzaklık deęerleri kullanılarak yapılan bu tarz alıřmalarda lkelerin bařkentlerinin yerine her bir lkenin merkez noktasının kullanılması, aęırlıęın homojen olarak gsteriminde byk nem tařımaktadır. Őekil 5.6'da gsterilen uzaklık deęerleri dikkate alındıęında Kanada'nın bařkenti ile ABD'nin bařkenti arasındaki uzaklıęı ok az olduęu grlmektedir. Dolayısıyla iki lkenin ok fazla Őekilde etkileřim altında oldukları sylenebilmeliydi. Ancak yzlmlerinin byklę dikkate alındıęında, sadece bařkentleri birbirine yakın diye bu iki lkenin neredeyse birbirinin aynısı olduęunu sylemek byk yanlıřlara neden olacaktır. Bu nedenle merkez noktaların lkenin genelini daha iyi temsil ettięi dřnlr ve bu tarz alıřmalarda uzaklık deęerlerinin belirlenmesi iin bu noktalar kullanılır.

5.3.1. w_1 Aęırlık Fonksiyonun Kullanılması ile Oluřturulan Modeller

alıřmanın bu blmnde ekirdek (kernel) bant geniřlięi (h) kullanılmıř ve fonksiyon yardımıyla daha zel yapıda aęırlık matrisleri oluřturulmuřtur. ekirdek bant geniřlięi hesaplanırken “apraz geerlilik” yntemi uygulanmıř ve bu deęerlerin elde edilmesinde Serdar Demir (2005) tarafından Delphi-6 yazılımı ile desteklenmiř olarak hazırlanan paket program kullanılmıřtır. Bu program bant geniřlięinin hesaplanmasında Gaussian ekirdek fonksiyonlarını veya Epanechnikov ekirdek fonksiyonlarını kullanmaktadır. Yine Serdar Demir'in (2005) hazırladıęı “Regresyon Fonksiyonlarının Uyarlanabilir Nadaraya-Watson ekirdek Kestirimleri” isimli tezinde eřitli benzetim alıřmaları ile ispatladıęı gibi, lkelere iliřkin gzlem sayısının 127 ($n > 100$) olması nedeniyle Gaussian ekirdek fonksiyonunun, Epanechnikov ekirdek fonksiyonundan daha iyi sonular vereceęine karar verilmiřtir. Oluřturulan matrisle iliřkin deęerler Ek-4'de verilmiřtir.

alıřmanın bundan sonraki tm ařamalarında baęımlı deęiřken olarak yaklařık simetrik hale getirilmiř olan internet kullanıcıları deęiřkeni alınmıřtır. Bunun yanı sıra dięer tm deęiřkenler baęımlı deęiřken ile birlikte tek tek incelenmiř ve apraz geerlilik incelemesi sonucunda bant geniřlięi tm deęiřkenler iin yaklařık olarak -0,3 ve 0,3 olarak bulunmuřtur. Ancak ilgilenilen coęrafi blgeler arasındaki uzaklık arttıka aęırlıęın daha dřk deęerler alması gerektięi mantıęından yola ıkarak pozitif bant geniřlięi deęerinin seilmesinin uygun olduęuna karar verilmiřtir.

Elde edilen bant geniřlięi yardımıyla yeni bir aęırlık fonksiyonu oluřturulmuřtur. Bu fonksiyon oluřturulurken daha nce de belirtilen aęırlık deęerinin odak nokta ve komřuları arasındaki farkın artması ile daha dřk deęerler alması mantıęından yola ıkarak, fonksiyonun ssel yapıda oluřturulmasına karar verilmiřtir. Bunun sonucunda fonksiyon,

$$w_1 = \exp\left(-\frac{1}{2} \frac{d_{ij}^2}{h^2}\right) \quad (5.1)$$

şeklinde oluşturulmuştur (Fotheringham, 2006, syf.45). Literatürde oldukça sık kullanılan bu fonksiyon “Gaussian ağırlık fonksiyonu” olarak da adlandırılır. Fonksiyondaki h değeri çekirdek bant genişliği; d değeri dünya atlası üzerinde ülkeler arasındaki uzaklığı göstermektedir.

Uzamsal modellere geçmeden önce ilk adımda klasik regresyon analizinin uygulanmasının gerektiği bilinmektedir. Çalışmaya eksik gözlemlerin giderilmesi ile devam edileceği düşünüldüğünden Tablo 5.8 ile gösterilmiş olan klasik model sonuçları geçerlidir.

Çalışılacak uzamsal modeli belirlemek amacıyla w_1 ağırlık matrisi dikkate alınarak elde edilmiş olan uzamsal bağımlılık istatistikleri tabloda gösterilmiştir.

Tablo 5.12. Uzamsal Bağımlılık İstatistikleri

UZAMSAL BAĞIMLILIK İSTATİSTİKLERİ			
	Değer	S.D.	Olasılık (p)
Lagrange Çarpanı (gecikme-lag)	0.1026	1	0.07469
Robust LM (gecikme)	0.0186	1	0.08406
Lagrange Çarpanı (hata-error)	0.6294	1	0.42760
Robust LM (hata)	0.5454	1	0.46020

Elde edilen istatistikler incelendiğinde Lagrange Gecikme Çarpanı için elde edilen değerlerin anlamlı çıkmasından dolayı w_1 ağırlık matrisi ile uzamsal gecikme modeli oluşturularak devam edilmesinin uygun olduğu belirlenmiştir. Oluşturulan gecikme modeline ilişkin sonuçlar Tablo 5.13'te gösterilmiştir.

Klasik modelde çalışmada bulunan toplam vergi oranı ve en son teknolojinin ulaşılabilirliği değişkenleri tüm modellerde anlamsız bulunmaları ve simetrik hale getirilememesi nedeniyle, yeni ağırlık matrislerinin kullanıldığı uzamsal modellerin elde edilmesi sürecinde çalışmadan çıkarılmıştır. Bu değişkenler çıkarılmadan model oluşturulduğunda tam çoklu bağlantı sorunu ile karşılaşmış ve analizlerden sonuç alınamamıştır. Bu nedenle ilgilenilen model ve daha sonra oluşturulacak tüm modellerde bu iki değişken çalışma dışında bırakılmıştır.

Tablo 5.13. w_1 Ağırlık Matrisli Uzamsal Gecikme Modeline İlişkin Sonuçlar

UZAMSAL GECİKME MODELİ SONUÇLARI				
Gözlem Sayısı	:103			
Değişken Sayısı	:9	Log Olabilirlik		:-43.01999
Serbestlik Derecesi	:94	AIC		:104.04
DEĞİŞKEN	KATSAYI	STD. HATA	T	OLASILIK
Sabit	5.9682e-01	1.9940e-01	2.9930	0.002762
Ağırlık (w_1)	-0.00014521	0.0000720	2.0236	0.043011
Tel. Hattı	2.1872e-01	4.0640e-02	5.3820	7.365e-08
Int. Sunucuları	7.1212e-02	5.1284e-02	1.3886	0.164960
Elek. Üretimi	5.5577e-06	1.2516e-05	0.4441	0.657006
Eğt. Harcaması	6.1151e-02	2.8833e-02	2.1209	0.033933
Kişisel Bilg.	1.5368e-01	1.0771e-01	1.4268	0.153642
Yenilik Kap.	6.0269e-02	6.7255e-02	0.8961	0.370183

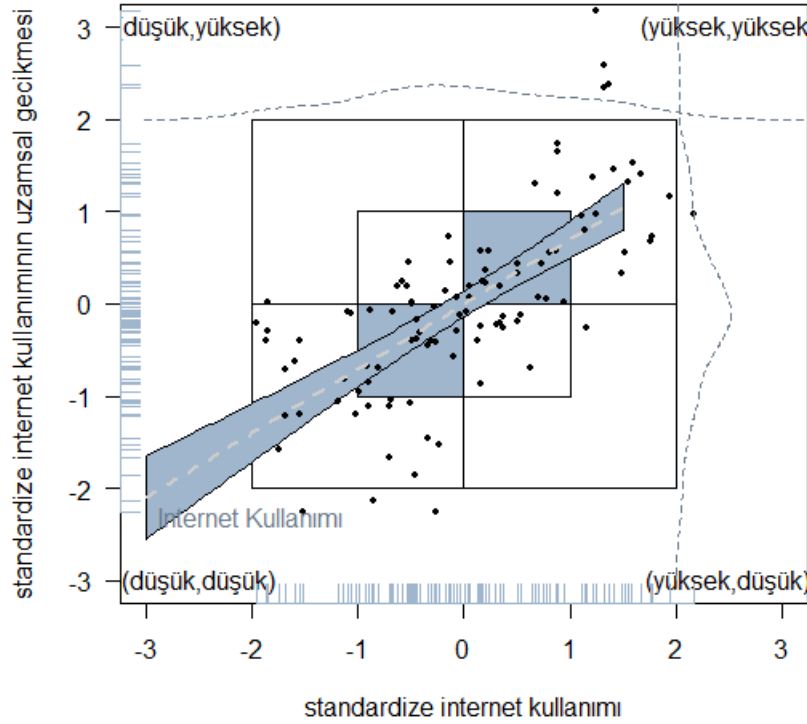
Modeli oluşturan değişkenler incelendiğinde sabit, w_1 ağırlığı, telefon hattı sayısı ve eğitim harcamaları değişkenlerinin anlamlı oldukları belirlenmiştir. Elde edilen sonuçlar Tablo 5.8'de gösterilen klasik modele ilişkin sonuçlar karşılaştırıldığında, uzamsal modelde anlamlı olan değişken sayısının daha fazla olduğu görülmüştür. Bunun yanı sıra AIC değeri (104.04), 103 gözlemlili klasik model için hesaplanan değerden daha düşük çıkmıştır. İki modelden birinin tercih edilmesi durumunda klasik model dikkate alınmalıdır.

Oluşturulan modelden elde edilen artıkların uzamsal otokorelasyona sahip olup olmadıklarının incelenmesi için kullanılan Moran I test değeri,

Moran / istatistik değeri : 0.9349

p-değeri : 0.3499

olarak hesaplanmıştır. Elde edilen Moran I test istatistiğine göre w_1 ağırlığı ile oluşturulmuş olan modelden elde edilen artık değerlerinde uzamsal otokorelasyonun bulunmadığı ($p>0.05$) belirlenmiştir. Uzamsal otokorelasyonun bulunmaması verilerde herhangi bir şekilde kümelenmenin olmadığını, dolayısıyla bu ağırlık ile oluşturulmuş olan modelde coğrafi anlamda verilerin bir arada kümelenmesinin önüne geçildiği söylenebilmektedir.



Şekil 5.7. w_1 Ağırlığı için Elde Edilen Moran I Saçılım Grafiği

Moran I saçılım grafiği incelendiğinde test istatistiğinden elde edilen sonuçlara benzer şekilde verilerin birbirinden bağımsız şekilde dağıldıkları görülmektedir. Oluşturulan grafikte görüldüğü gibi, standartlaştırılmış bağımlı değişken ve bağımlı değişkenin uzamsal gecikmeli haline ilişkin oluşturulan şekilde verilerin $\pm 2\sigma$ arasında dağılımları beklenmektedir. Ancak şekilde de görüleceği gibi [-2, 2] için oluşturulan kutunun dışında bazı gözlemlerin olduğu belirlenmiştir. Bu gözlemler için çalışmada etkin gözlemler oldukları söylenebilir. Bu etkili ülkelerin hangileri oldukları ve bunların belirlenmesinde kullanılan istatistikler Tablo 5.14'de gösterilmiştir.

Tablo 5.14. Etkin Gözlemler

No	Ülke Adı	Dfb.1	dfb.x	dffit	cov.r	cook.d
7	Bangladeş	-0.10	0.09	-0.10	1.06 *	0.00
13	Brezilya	0.07	0.06	0.36	0.81 *	0.06
17	Kamboçya	0.07	-0.06	0.07	1.06 *	0.00
20	Çad	0.03	-0.02	0.03	1.06 *	0.00
22	Çin	0.15	-0.06	0.25	0.91 *	0.03
31	Etiyopya	-0.08	0.07	-0.08	1.07 *	0.00
35	Almanya	-0.27	0.42	0.53 *	0.85 *	0.13
66	Hollanda	0.15	-0.20	-0.22	1.06 *	0.02
78	Rusya	0.05	0.07	0.33	0.85 *	0.05
86	İsveç	0.10	-0.13	-0.15	1.06 *	0.01
89	Tacikistan	0.02	-0.02	0.02	1.07 *	0.00

w_1 ağırlık matrisi ile oluşturulmuş gecikme modelinin sonuçları ve etkin olarak belirlenen bu ülkelerin diğer ülkelerden farklı çıkma nedenleri, diğer ağırlık matrislerinden elde edilen sonuçlarla karşılaştırılarak çalışmanın devamında sunulacaktır.

5.3.2. w_2 Ağırlık Fonksiyonunun Kullanılması ile Oluşturulan Modeller

Ağırlık matrisini oluşturmak için kullanılan fonksiyon,

$$w_2 = d^{-h} \quad (5.2)$$

şeklinde verilmiştir. Bu fonksiyonda kullanılan d uzaklık değerlerini, h ise bir önceki bölümde hesaplanma şekli açıklanan çekirdek bant genişliğini göstermektedir. Gatrell (1983) çalışmalarında bu ağırlık fonksiyonunu, bant genişliğinin yerine $\delta \geq 0$ olacak şekilde bir parametre tanımlayarak kullanmıştır. Bu ağırlık matrisinin de oluşturulmasında, yakın bölgelerin birbirlerini uzak bölgelere oranlara daha fazla etkilediği mantığı esas alınmıştır. Bu matrise ilişkin değerler Ek-5'de gösterilmiştir.

Tablo 5.8'de gösterilen klasik regresyon sonuçları, bu ağırlık matrisinin kullanımdan önce de geçerlidir. Klasik regresyon sonuçlarının elde edilmesinin ardından w_2 için uzamsal bağımlılık istatistikleri incelenmiştir.

Tablo 5.15. Uzamsal Bağımlılık İstatistikleri

UZAMSAL BAĞIMLILIK İSTATİSTİKLERİ			
	Değer	S.D.	Olasılık (p)
Lagrange Çarpanı (gecikme-lag)	0.9652	1	0.03259
Robust LM (gecikme)	0.5821	1	0.04455
Lagrange Çarpanı (hata-error)	0.9208	1	0.33730
Robust LM (hata)	0.5377	1	0.46340

Elde edilen istatistikler incelendiğinde Lagrange Gecikme Çarpanı için elde edilen değer anlamlı çıkmasından dolayı w_2 ağırlık matrisi ile uzamsal gecikme modeli oluşturularak devam edilmesinin uygun olduğu belirlenmiştir. Oluşturulan gecikme modeline ilişkin sonuçlar Tablo 5.16'da gösterilmiştir.

Tablo 5.16. w_2 Ağırlık Matrisli Uzamsal Gecikme Modeline İlişkin Sonuçlar

UZAMSAL GECİKME MODELİ SONUÇLARI				
Gözlem Sayısı	:103			
Değişken Sayısı	:9	Log Olabilirlik		:-56.64018
Serbestlik Derecesi	:94	AIC		:103.79
DEĞİŞKEN	KATSAYI	STD. HATA	T	OLASILIK
Sabit	5.9813e-01	1.9909e-01	3.0043	0.002662
Ağırlık (w_2)	0.0012042	0.0006390	1.8848	0.059463
Tel. Hattı	2.1169e-01	4.0518e-02	5.2247	1.745e-07
Int. Sunucuları	7.2217e-02	5.0707e-02	1.4242	0.254389
Elek. Üretimi	6.4881e-06	1.2613e-05	0.5144	0.606982
Eğt. Harcaması	6.1374e-02	2.8796e-02	2.1313	0.033063
Kişisel Bilg.	1.5586e-01	1.0757e-01	1.4488	0.147385
Yenilik Kap.	5.4666e-02	6.6877e-02	0.8174	0.413695

Modeli oluşturan değişkenler incelendiğinde w_1 ağırlığı ile oluşturulan uzamsal gecikme modeline benzer şekilde sabit, w_2 ağırlığı, telefon hattı sayısı ve eğitim harcamaları değişkenlerinin anlamlı oldukları belirlenmiştir. Bunun yanı sıra AIC değeri 103.79 olarak hesaplanmıştır. Hesaplanan değer ilk ağırlık ile hesaplanan değerden daha düşük çıktığı için, w_2 ağırlık matrisi ile oluşturulan modelin tercih edilebilir bir model olduğu söylenebilir.

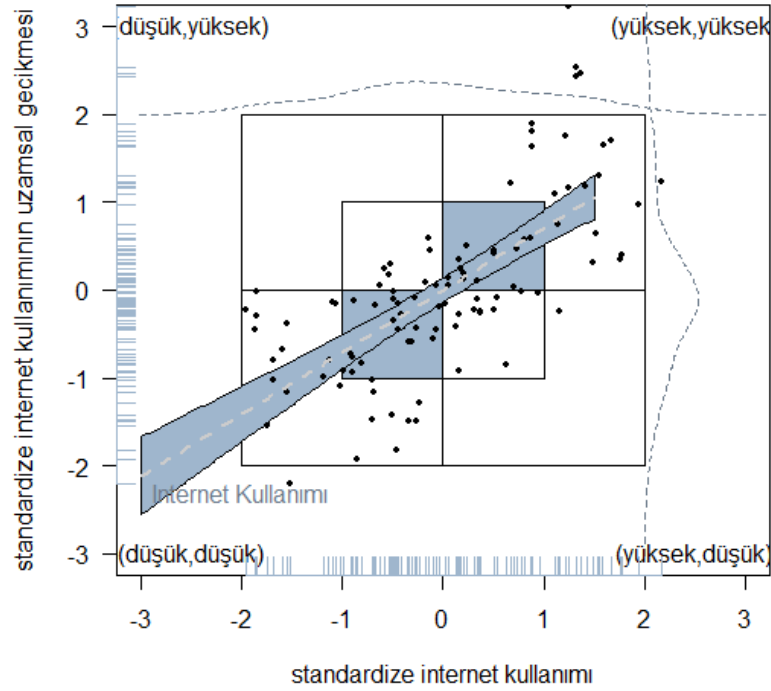
Model için elde edilen artıkların uzamsal otokorelasyona sahip olup olmadığının incelenmesi için Moran I istatistik değeri,

Moran I istatistik değeri : 1.098

p-değeri : 0.2722

olarak hesaplanmıştır. Elde edilen sonuca göre w_2 ağırlığı ile oluşturulmuş olan modelden elde edilen artık değerlerinde uzamsal otokorelasyonun bulunmadığı ($p>0.05$) belirlenmiştir. Test istatistiğinin sonucuna göre verilerin kendi içerisinde kümelenmedikleri, dolayısıyla oluşturulan gecikme modelinde uzamsal etkilerin ortadan kaldırıldığı ve konumların belirleyici etkilerinin azaltıldığı söylenebilir.

Standartlaştırılmış bağımlı değişken ve standartlaştırılmış bağımlı değişkenin gecikmelerine ilişkin Moran I test değeri referans alınarak oluşturulan Moran saçılım grafiği Şekil 5.8'de gösterilmiştir.



Şekil 5.8. w_2 Ağırlığı için Elde Edilen Moran I Saçılım Grafiği

Moran I saçılım grafiği incelendiğinde test istatistiğinden elde edilen sonuçlara benzer şekilde verilerin birbirinden bağımsız şekilde dağıldıkları görülmektedir. Elde edilen şekilde etkin gözlemlerin olduğu görülmektedir. Bu gözlemlerin hangi ülkelere ait oldukları ve tespit edilmesinde kullanılan istatistiklere ilişkin sonuçlar Tablo 5.17'de gösterilmiştir.

Tablo 5.17. Etkin Gözlemler

No	Ülkeler	dfb.1	dfb.x	dffit	cov.r	cook.d	hat
5	Avusturya	-0.24	0.36	0.45_*	0.92_*	0.09	0.03
7	Bangladeş	-0.07	0.06	-0.07	1.06_*	0.00	0.04
13	Brezilya	0.05	0.04	0.23	0.93_*	0.03	0.01
19	Kanada	0.27	-0.38	-0.43_*	0.97	0.09	0.04
20	Çad	0.03	-0.03	0.03	1.06_*	0.00	0.04
31	Etiyopya	-0.04	0.04	-0.04	1.07_*	0.00	0.05
35	Almanya	-0.31	0.49	0.62_*	0.79_*	0.17	0.03
66	Hollanda	0.06	-0.08	-0.09	1.08_*	0.00	0.06
89	Tacikistan	0.05	-0.04	0.05	1.07_*	0.00	0.04

w_1 ağırlık matrisi ile oluşturulan modelden farklı olarak w_2 ağırlık matrisi ile oluşturulan modelde, Kamboçya, Çin, Rusya ve İsveç etkin gözlem olarak hesaplanmazken, diğerlerinden farklı olarak Avusturya ve Kanada etkin olarak

belirlenmiştir. Ayrıca etkin olduğundan şüphelenilen gözlem sayısının azalması, modelin veriler için daha uyumlu olduğunun göstergesi olabilir.

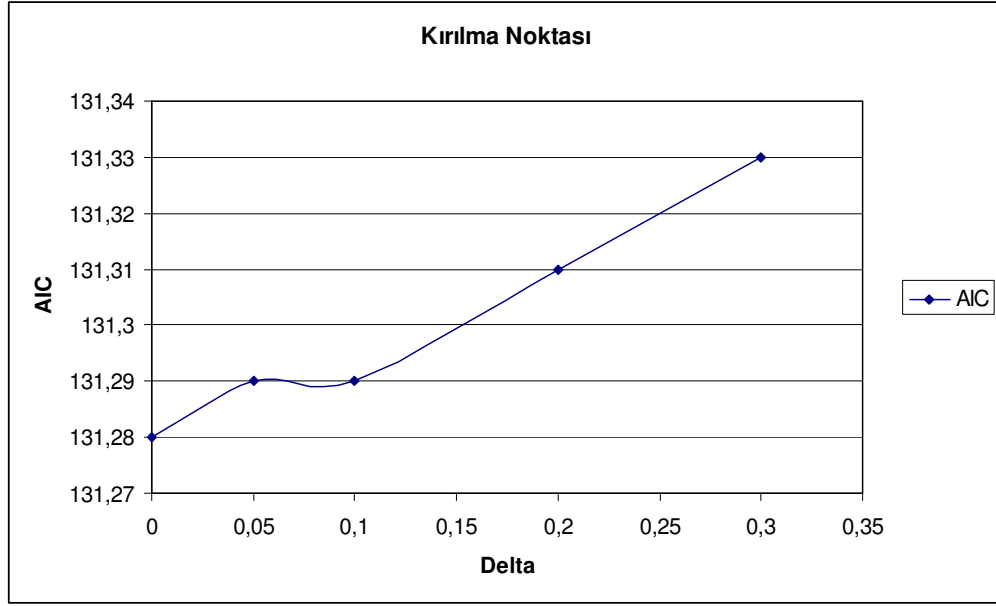
5.3.3. w_3 Ağırlık Fonksiyonunun Kullanılması ile Oluşturulan Modeller

Uzamsal modellerle çalışılırken verilerin konumlarının etkilerini modele dahil edip, kestirimlerin daha tutarlı olmasına imkan sağlayan ağırlık matrisinin değişik yollarla oluşturulabileceği belirtilmişti. Çalışmanın bu kısmında da farklı bir ağırlık fonksiyonu yardımıyla ağırlık matrisi oluşturulmuştur.

Ülkelerin internet kullanımının incelendiği modelde sadece ülkelerin birbirlerine göre uzaklıklarının değil, farklı etkilerin de model sonucunda değişikliklere neden olduğu düşünülmüştür. İncelenebilecek birçok etken vardır. Ancak bu çalışma için Birleşmiş Milletler Kalkınma Programı (UNDP) çerçevesinde oluşturulan “İnsani Gelişme İndeksi” (Human Development Index - HDI) (URL-4, 2008) bir etken olarak alınmıştır. Bu indeksin seçilmesinin nedeni, indeksin hesaplanması için birçok farklı kriteri bünyesinde barındırmasıdır. Bilindiği gibi HDI, yaşam süresi indeksinin, eğitim indeksinin ve GSMH indeksinin bir araya getirilmesi ile ülkeleri gelişmiş, gelişmekte olan ve az gelişmiş olmak üzere üç kategoriye ayırmayı sağlayan değerlerden oluşmaktadır. Çalışmada sadece uzaklıkların değil, tüm bu faktörlerin internet kullanımı üzerindeki etkilerini görmek amacıyla w_3 olarak isimlendirilen ve uzaklığın yanı sıra incelenen ülkelerin HDI değerlerinin birbirlerine oranlarının da dahil edildiği ağırlık matrisinin kullanılması amaçlanmıştır. Bu matrisi için elde etmek için oluşturulan fonksiyon bu çalışmada,

$$w_3 = d^{-h} \left(\frac{HDI_i}{HDI_j} \right)^{-\delta} \quad (5.3)$$

şeklinde gösterilmiştir. Fonksiyonda d ülkelerin merkez noktaları arasındaki uzaklığı, h çekirdek bant genişliğini, $HDI_{i,j}$ i ve j ülkelerinin insani gelişme indeks değerlerini, δ ise bilinmeyen parametreyi göstermektedir. Bu parametrenin, uzaklıklar arttıkça ülkelerin birlerini daha az etkilemeleri mantığından yola çıkarak $(0,1]$ arasında değer alması gerektiği belirlenmiştir. Eğer indeks değerinin etkisi fonksiyona dahil edilecekse 0 değerini alamayacağı çok açıktır. En uygun δ değerinin belirlenmesi için uzamsal modelde belirlenen ağırlık matrisini kullanarak farklı δ değerleri için modeller oluşturulmuş ve bu modellerden elde edilen AIC değerleri karşılaştırılmıştır. Karşılaştırmaya ilişkin sonuçlar Şekil 5.9'daki gibi gösterilmiştir.



Şekil 5.9. Optimum Delta Değerinin Belirlenmesi için Kırılma Noktasının Tespit Edilmesi

δ 'nın alacağı değer 0 olamayacağı belirtilmişti. Dolayısıyla şekil incelendiğinde kırılma noktasının 0,1 olduğu görülmektedir. Buradan yola çıkarak incelenen ağırlık matrisi ile oluşturulan modelde kullanılacak optimum δ değerinin 0,1 olmasına karar verilmiştir. Bu değer belirlenmesi ile 103 gözlemi içeren veri setine uzaklıkların yanı sıra insani gelişmişlik indeks değerinin de dahil edildiği ağırlık matrisini oluşturmak için kullanılacak fonksiyon,

$$w_3 = d^{-h} \left(\frac{HDI_i}{HDI_j} \right)^{-0,1} \quad (5.4)$$

olarak düzenlenmiştir. Matrise ilişkin değerler Ek-6'da gösterilmiştir.

Ağırlık fonksiyonunun belirlenmesinin ardından uzamsal modellerden hangisinin tercih edilmesi gerektiğini gösteren uzamsal bağımlılık istatistikleri hesaplanmıştır.

Tablo 5.18. Uzamsal Bağımlılık İstatistikleri

UZAMSAL BAĞIMLILIK İSTATİSTİKLERİ			
	Değer	S.D.	Olasılık (p)
Lagrange Çarpanı (gecikme-lag)	0.9847	1	0.03210
Robust LM (gecikme)	0.5983	1	0.04392
Lagrange Çarpanı (hata-error)	0.9193	1	0.33770
Robust LM (hata)	0.5329	1	0.46540

Uzamsal bağımlılık istatistikleri yardımıyla çalışmaya uzamsal gecikme modeli ile devam edilmesine karar verilmiştir. Buradan yola çıkarak uzamsal gecikme modeline ilişkin sonuçlar tablodaki gibi bulunmuştur.

Tablo 5.19. w_3 Ağırlık Matrisli Uzamsal Gecikme Modeline İlişkin Sonuçlar

UZAMSAL GECİKME MODELİ SONUÇLARI				
Gözlem Sayısı	:103			
Değişken Sayısı	:9	Log Olabilirlik		:-42.89069
Serbestlik Derecesi	:94	AIC		:103.78
DEĞİŞKEN	KATSAYI	STD. HATA	T	OLASILIK
Sabit	5.9782e-01	1.9907e-01	3.0031	0.002673
Ağırlık (w_3)	0.0012124	0.0006440	1.8815	0.05771
Tel. Hattı	2.1172e-01	4.0467e-02	5.2319	1.678e-07
Int. Sunucuları	7.2201e-02	5.0704e-02	1.4240	0.154456
Elek. Üretimi	6.4930e-06	1.2612e-05	0.5148	0.606686
Eğt. Harcaması	6.1402e-02	2.8796e-02	2.1323	0.032982
Kişisel Bilg.	1.5587e-01	1.0757e-01	1.4490	0.147342
Yenilik Kap.	5.4612e-02	6.6874e-02	0.8167	0.414127

w_3 ağırlığı ile oluşturulan uzamsal gecikme modelinde sabit, w_3 ağırlığı, telefon hattı sayısı ve eğitim harcamaları değişkenlerinin anlamlı oldukları belirlenmiştir. Bunun yanı sıra AIC değeri 103.78 olarak hesaplanmıştır.

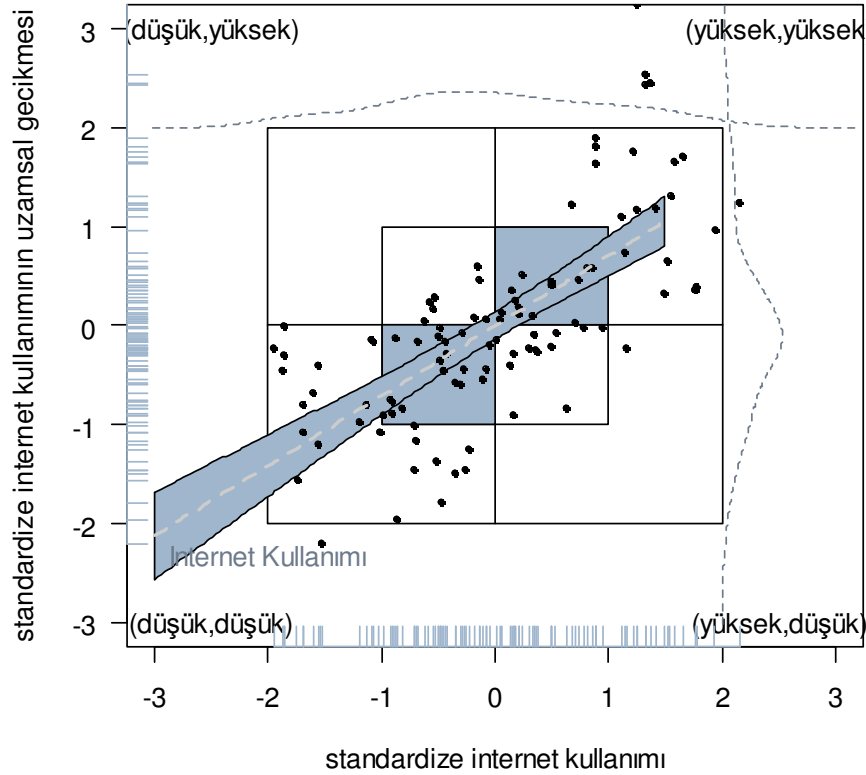
Model için elde edilen artıkların uzamsal otokorelasyona sahip olup olmadığının incelenmesi için hesaplanan Moran I istatistik değeri,

Moran I istatistik değeri : 1.0972

p-değeri : 0.2726

olarak bulunmuştur. Elde edilen sonuca göre w_3 ağırlığı ile oluşturulmuş olan modelden elde edilen artık değerlerinde uzamsal otokorelasyonun bulunmadığı ($p>0.05$) belirlenmiştir.

Standartlaştırılmış bağımlı değişken ve standartlaştırılmış bağımlı değişkenin gecikmelerine ilişkin Moran I test değeri referans alınarak oluşturulan Moran saçılım grafiği, Şekil 5.10'da gösterilmiştir.



Şekil 5.10. w_3 Ağırlığı için Elde Edilen Moran I Saçılım Grafiği

Moran I saçılım grafiğinde Moran I testinde elde edilen sonuçlara uygun olarak verilerin birbirinden bağımsız şekilde dağıldığı, dolayısıyla bir etkileşim içerisinde olmadıkları söylenebilmektedir. Elde edilen şekilde belirtilen kutunun dışında kalan ve etkin özelliğe sahip gözlemlerin olduğu görülmektedir. Bu gözlemlerin kaç numaralı ülkelere ait olduğu ve bunların belirlenmesinde kullanılan test değerleri Tablo 5.20'de gösterilmiştir.

Tablo 5. 20. Etkin Gözlemler

No	Ülkeler	dfb.1_	dfb.x	dffit	cov.r	cook.d	hat
5	Avusturya	-0.24	0.36	0.44_*	0.92_*	0.09	0.03
7	Bangladeş	-0.07	0.07	-0.07	1.06_*	0.00	0.04
13	Brezilya	0.05	0.04	0.23	0.93_*	0.02	0.01
19	Kanada	0.27	-0.37	-0.43_*	0.97	0.09	0.04
31	Etiyopya	-0.04	0.04	-0.04	1.07_*	0.00	0.05
35	Almanya	-0.31	0.49	0.62_*	0.79_*	0.17	0.03
66	Hollanda	0.06	-0.08	-0.09	1.08_*	0.00	0.06
89	Tacikistan	0.04	-0.04	0.04	1.07_*	0.00	0.04

Burada bulunan etkin gözlemler w_2 ağırlığı ile hesaplanan modelde elde edilen etkin gözlemlerle, bir ülkenin durumu hariç, aynı çıkmıştır. Çad, w_2 ağırlık matrisi ile hesaplanan modelde etkin gözlem iken, w_3 ile hesaplanan modelde etkin gözlem olarak hesaplanmamıştır.

5.3.4. Farklı Ağırlık Fonksiyonun Kullanılması ile Oluşturulan Modellerin Birbirleri ile Karşılaştırılması

Çalışmanın ilgilenilen bölümünde üç farklı ağırlık matrisi kullanılmıştır. Bu matrislerden w_1 olarak isimlendirilen literatürde “Gaussian” olarak adlandırılan ağırlık fonksiyonundan türetilmiş bir matrisken, w_2 olarak gösterilen ve “ters fonksiyon” olarak isimlendirilen çekirdek bant genişliğinin fonksiyona dahil edilmesi ile elde edilmiş olan ağırlık matrisidir. w_3 olarak isimlendirilen son matris ise literatürde henüz kullanılmamış ve bu çalışma için özel olarak hazırlanmış olan, insani gelişmişlik değerlerinin fonksiyonlaştırılması yardımıyla elde edilmiş ağırlık matrisidir. Üç farklı ağırlık matrisi ile oluşturulmuş olan modeller birbirleri ile karşılaştırılıp kullanımını en uygun bulunan model belirlenmeye çalışılmıştır. Üç matris yardımıyla oluşturulan modellere ilişkin sonuçlar Tablo 5.21’de gösterilmiştir.

Tablo 5.21. Farklı Ağırlık Matrisleri ile Oluşturulan Modellere İlişkin Sonuçlar

	w_1 AĞIRLIK MATRİSİ İLE OLUŞTURULAN MODEL SONUÇLARI		w_2 AĞIRLIK MATRİSİ İLE OLUŞTURULAN MODEL SONUÇLARI		w_3 AĞIRLIK MATRİSİ İLE OLUŞTURULAN MODEL SONUÇLARI		
	Değer	Olasılık	Değer	Olasılık	Değer	Olasılık	
Karşılaştırma Ölçütleri	Standart Hata	0.36741		0.36695		0.36694	
	Log Olabilirlik	-43.01999		-56.64018		-42.89069	
	AIC	104.04		103.79		103.78	
Katsayılar	Ağırlık (W)	5.9682e-01	0.002762	5.9813e-01	0.002662	5.9782e-01	0.002673
	Sabit	-0.00014521	0.043011	0.0012042	0.059463	0.0012124	0.05771
	Tel. Hattı	2.1872e-01	7.365e-08	2.1169e-01	1.745e-07	2.1172e-01	1.678e-07
	Int. Sunucuları	7.1212e-02	0.164960	7.2217e-02	0.254389	7.2201e-02	0.154456
	Elek. Üretimi	5.5577e-06	0.657006	6.4881e-06	0.606982	6.4930e-06	0.606686
	Eğt. Harcaması	6.1151e-02	0.033933	6.1374e-02	0.033063	6.1402e-02	0.032982
	Kişisel Bilg.	1.5368e-01	0.153642	1.5586e-01	0.147385	1.5587e-01	0.147342
	Yenilik Kap.	6.0269e-02	0.370183	5.4666e-02	0.413695	5.4612e-02	0.414127
Test	Moran I Testi	0.9349	0.3499	1.098	0.2722	1.0972	0.2726

Üç model için hesaplanmış olan test incelendiğinde; Moran / testi tüm modellerin artıklarında uzamsal otokorelasyon sorunun olmadığını göstermekte ve incelenen tüm modellerde otoregresif katsayının anlamlı olduğunu göstermektedir.

Modellerde kullanılan deęişkenlerden aęırlığın, sabit terimin, telefon hattı sayısının ve eęitim harcamalarının her üç modelde anlamlı çıkan deęişkenler oldukları görölmektedir.

103 gözlem için incelenen klasik modelde de aynı deęişkenlerin anlamlı çıktıkları tespit edilmiştir. Buradan yola çıkarak verilere klasik model uygulamak yerine uzamsal konumlarının da dikkate alındığı uzamsal gecikme modellerinin kullanılmasının uygun olduğuna karar verilmiştir. Bu modellerden hangisinin daha tercih edilebilir olduğuna karar verebilmek için karşılaştırma ölçüleri dikkate alınabilir. Bu durumda en büyük log olabilirlik ve en küçük AIC deęerine sahip olan ve w_3 aęırlık matrisi ile oluşturulmuş olan modelin kullanılması önerilmektedir.

Üç modeldeki etkin gözlemler incelendiğinde Almanya'nın dięer tüm ölkelerden çok farklı olarak etkin çıktığı belirlenmiştir. Bu nedenden dolayı analizin devamında bu gözlem çalışmadan çıkarılmış ve uzamsal istatistikler incelenmiştir.

Tablo 5.22. Almanya'nın Çıkarıldığı Veri Setinin Uzamsal Baęımlılık İstatistikleri

UZAMSAL BAęIMLILIK İSTATİSTİKLERİ						
	W_1		W_2		W_3	
	Deęer	Olasılık	Deęer	Olasılık	Deęer	Olasılık
Lagrange Çarpanı (gecikme-lag)	0.4437	0.5054	2.1649	0.1412	2.1992	0.1381
Robust LM (gecikme)	0.2175	0.6409	1.522	0.2173	1.5529	0.2127
Lagrange Çarpanı (hata-error)	0.7346	0.3914	1.0597	0.3033	1.0579	0.3037
Robust LM (hata)	0.5084	0.4758	0.4169	0.5185	0.4117	0.5211

Tablo 5.22'de verilmiş olan sonuçlar incelendiğinde Almanya'nın veriler arasından çıkarıldığı durumda uzamsal analizlerin yapılmasının uygun olmadığı belirlenmiştir. Bu nedenden dolayı Almanya'nın etkin gözlem olarak belirlenmesine rağmen veriler arasında bırakılmasına karar verilmiştir.

Fazla sayıda etkin gözlem çıkmasından dolayı analize yeni bir deęişken eklenmesine karar verilmiştir. Bu deęişken üç model de etkin çıkan gözlemlerin bir olarak kodlandığı kukla deęişken olarak düzenlenmiştir. Bu deęişkenin de modele dahil edilmesi ile uzamsal analizlerin uygulanabilirliğini incelenek amacıyla uzamsal baęımlılık istatistikleri elde edilmiş ve Tablo 5.23'deki gibi gösterilmiştir.

Tablo 5.23. Kukla Değişken Eklenmesi ile Elde Edilen Uzamsal Bağımlılık İstatistikleri

UZAMSAL BAĞIMLILIK İSTATİSTİKLERİ						
	W ₁		W ₂		W ₃	
	Değer	Olasılık	Değer	Olasılık	Değer	Olasılık
Lagrange Çarpanı (gecikme-lag)	0.4443	0.5050	1.1514	0.2833	1.7880	0.1812
Robust LM (gecikme)	0.1860	0.6663	1.7685	0.1836	1.1685	0.2797
Lagrange Çarpanı (hata-error)	1.1247	0.2889	1.3524	0.2449	1.3505	0.2452
Robust LM (hata)	0.8664	0.3520	0.7353	0.3912	0.7310	0.3925

Sonuçlar incelendiğinde etkin gözlemler için oluşturulan kukla değişkenin kullanıldığı verilere uzamsal analizlerin yapılmasının uygun olmadığı belirlenmiştir.

Çalışmada kullanılan üç farklı ağırlık matrisi yardımıyla oluşturulan modellerde aynı değişkenlerin istatistiksel olarak anlamsız çıktıkları belirlenmiştir. Farklı bir yaklaşım olması açısından anlamsız bulunan tüm değişkenler çalışmadan çıkarılmış ve uzamsal analizlerin uygulanabilirliğini sınamak amacıyla uzamsal bağımlılık istatistikleri hesaplanmıştır.

Tablo 5.24. Anlamsız Bulunan Değişkenlerin Çıkarılması ile Hesaplanan Uzamsal Bağımlılık İstatistikleri

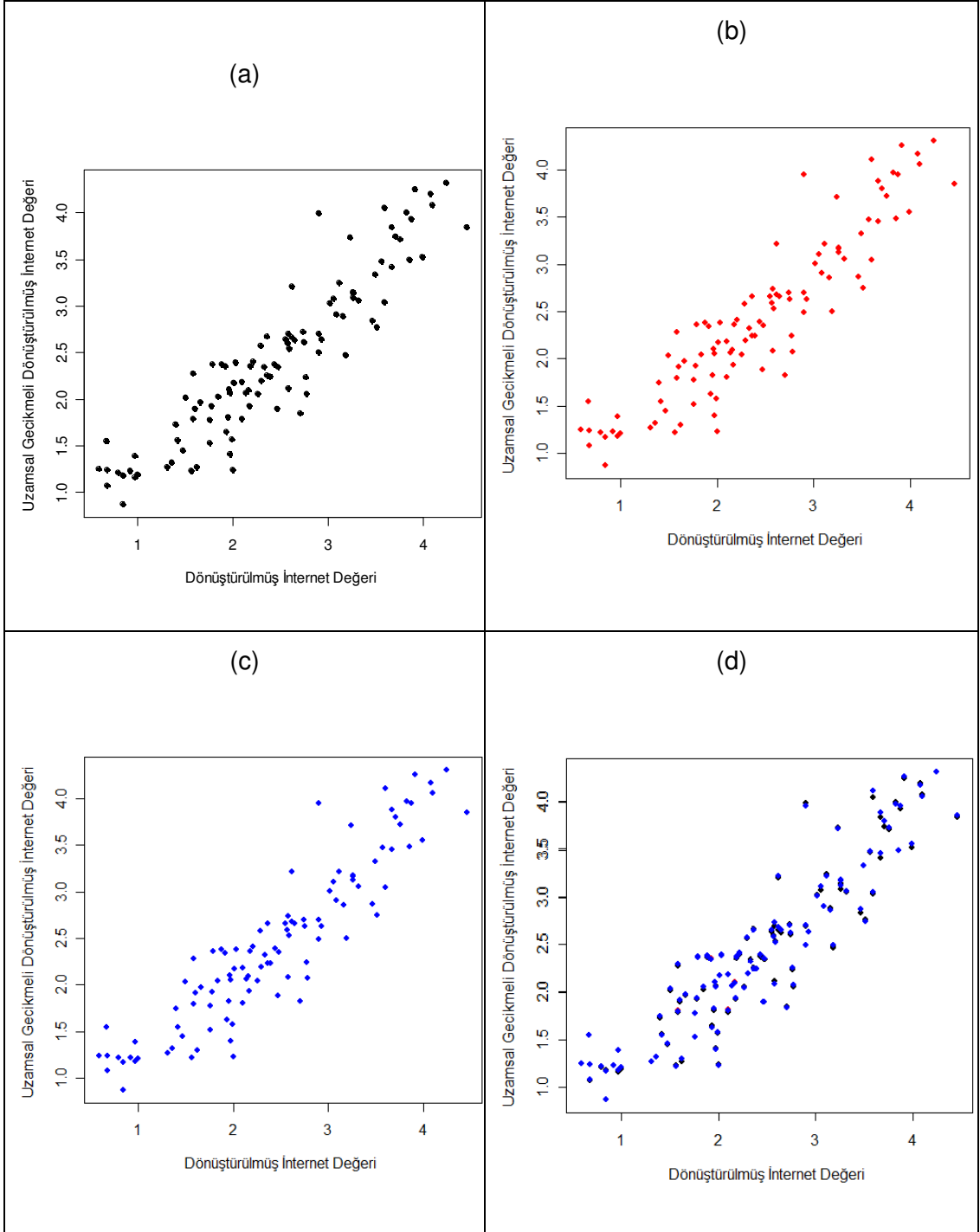
UZAMSAL BAĞIMLILIK İSTATİSTİKLERİ						
	W ₁		W ₂		W ₃	
	Değer	Olasılık	Değer	Olasılık	Değer	Olasılık
Lagrange Çarpanı (gecikme-lag)	0.1026	0.8917	0.9652	0.3259	0.9847	0.3210
Robust LM (gecikme)	0.0186	0.8917	0.5821	0.4455	0.5983	0.4392
Lagrange Çarpanı (hata-error)	0.6294	0.4276	0.9208	0.3373	0.9193	0.3377
Robust LM (hata)	0.5454	0.4602	0.5377	0.4634	0.5329	0.4654

Tablo 5.24’de verilen sonuçlar incelendiğinde, anlamsız değişkenlerin çıkarılması ile oluşturulan verilere uzamsal analizlerin uygulanmasının uygun olmadığına karar verilmiştir.

5.4. MODELİN GEÇERLİLİĞİNİN İNCELENMESİ VE TAHMİNLER

Çalışmanın bu bölümünde oluşturulan ağırlık matrisleri ile elde edilen modellerin geçerliliği sınanacaktır. Bu sınamanın yapılmasından önce modellerde kullanılan “dönüştürülmüş internet kullanımı” değişkeninin, uzamsal gecikmeli modellerinden

elde edilen “uzamsal gecikmeli dönüştürülmüş internet kullanımı” değerleri ile saçılım grafikleri incelenmiştir.



Şekil 5.11. (a) w_1 Ağırlık Matrisi, (b) w_2 Ağırlık Matrisi, (c) w_3 Ağırlık Matrisi ile Oluşturulmuş Modelde ($y - gecikmiş$) Çizimleri (d) a-b-c Şekillerinin Üst Üste Gösterilmesi

(a) saçılım grafiğinde w_1 ağırlık matrisi kullanılarak oluşturulan modeldeki bağımlı değişken ile gecikmeli bağımlı değişken arasındaki ilişki, (b) saçılım grafiğinde w_2 ağırlık matrisi kullanılarak oluşturulan modeldeki bağımlı değişken ile gecikmeli bağımlı değişken arasındaki ilişki ve (c) saçılım grafiğinde w_3 ağırlık matrisi

kullanılarak oluşturulan modeldeki bağımlı değişken ile gecikmeli bağımlı değişken arasındaki ilişki görülmektedir. Aslında her üçü de birbirinin aynısı gibi görünse de aralarında farklılıklar vardır. Hesaplanan değerler arasındaki farklar, (a), (b) ve (c) tablolarının üst üste çizilmesi sonucunda (d) şeklindeki gibi gösterilmiştir. Bu değerlerden kaynaklanan farklılıklar etkin gözlemlerin ağırlıktan ağırlığa çok az da olsa farklı çıkmasına neden olmuştur.

Bu çalışmada modellerin geçerliliğinin incelenmesi için izlenen yöntem şu şekildedir:

- 103 gözlemleri veri seti, ülkelerin birbirleri ile komşulukları dikkate alınıp, 28 ve 75 gözlem içerecek iki gruba ayrılmıştır.
- 28, 75 ve 103 gözlemleri veriler için uzamsal gecikme modelleri ayrı ayrı oluşturulmuş ve bu modellerde $[(y - \hat{y})/n]$ değerleri hesaplanmıştır. Beklenen durum aynı ağırlık için 2 gruba bölünmüş verilerin değerlerinin yaklaşık olarak aynı çıkmasıdır.
- En son adımda 2 gruba ayrılmış veriler ile hesaplanmış olan modellerden elde edilen kestirimler ile orijinal değerler arasındaki korelasyonlar incelenmiştir.

Tüm adımlarda elde edilen sonuçlar Tablo 5.25'de gösterilmiştir.

Tablo 5.25. 28-75-103 Gözlemleri Verilerle Geçerliliğinin İncelenmesi için Hesaplanan Değerler

	28 gözlem $[(y - \hat{y})/n]$	75 gözlem $[(y - \hat{y})/n]$	103 gözlem $[(y - \hat{y})/n]$	orijinal veri ile korelasyonu _28gozlem	orijinal veri ile korelasyonu _75gozlem	orijinal veri ile korelasyonu _103gozlem
w1	0,1262	0,1093	0,1350	0,8838	0,9237	0,9196
w2	0,1271	0,1102	0,1346	0,8829	0,9230	0,9199
w3	0,1271	0,1103	0,1346	0,8829	0,9230	0,9199

Hesaplanan değerler incelendiğinde w_2 ve w_3 ağırlık matrisleri ile oluşturulmuş olan modellerin, w_1 ağırlık matrisi ile oluşturulmuş olan modelden daha geçerli oldukları söylenebilir. Modeller için elde edilen istatistikler de dikkate alındığında, internet kullanımının incelendiği ve 103 ülkenin bilgilerini içeren veriler için, w_3 ağırlık matrisi ile oluşturulan modelin tahminlerde kullanılmasının, diğer modellere göre daha uygun sonuçlar verebileceği söylenebilir.

Çalışmanın en son aşamasında elde edilen modeller yardımıyla tahminler yapılacaktır. Burada amaç incelenen ülkeler arasında bulunan Türkiye'nin internet kullanımına ilişkin bilgileri yardımıyla, sınırdaş komşularının Türkiye'nin değerlerindeki değişimlere nasıl tepki vereceklerinin ortaya çıkarılmasıdır. Buradaki değişimin incelenmesi için iki durum göz önüne alınmıştır. İlk durumda Türkiye'de

internet kullanımının %1 olduğu, ikinci durumda da bu oranın %100 olduğu kabul edilmiştir. Buradan yola çıkarak her iki durumun, komşu ülkelerin internet kullanımları üzerinde pozitif veya negatif artış miktarları incelenmiştir. Elde edilen sonuçlar Tablo 5.26'de gösterilmiştir.

Tablo 5.26. Farklı Ağırlık Matrisleri ile Türkiye için Yapılan Tahminler

AGIRLIK 1			
Türkiye'nin İnternet Kullanımı %1 Olsaydı		Türkiye'nin İnternet Kullanımı %100 Olsaydı	
ÜLKE	ARTIŞ (%)	ÜLKE	ARTIŞ (%)
Suriye	0,597	Suriye	2,769
Yunanistan	0,596	Gürcistan	2,768
Gürcistan	0,596	Bulgaristan	2,768
Bulgaristan	0,596	Ermenistan	2,768
Ermenistan	0,596	Yunanistan	2,767
AGIRLIK 2			
Türkiye'nin İnternet Kullanımı %1 Olsaydı		Türkiye'nin İnternet Kullanımı %100 Olsaydı	
ÜLKE	ARTIŞ (%)	ÜLKE	ARTIŞ (%)
Suriye	2,546	Suriye	11,818
Ermenistan	2,254	Ermenistan	10,464
Gürcistan	2,201	Gürcistan	10,216
Bulgaristan	2,201	Bulgaristan	10,216
Yunanistan	2,108	Yunanistan	9,787
AGIRLIK 3			
Türkiye'nin İnternet Kullanımı %1 Olsaydı		Türkiye'nin İnternet Kullanımı %100 Olsaydı	
ÜLKE	ARTIŞ (%)	ÜLKE	ARTIŞ (%)
Suriye	2,562	Suriye	11,892
Ermenistan	2,253	Ermenistan	10,459
Gürcistan	2,206	Gürcistan	10,239
Bulgaristan	2,186	Bulgaristan	10,148
Yunanistan	2,07	Yunanistan	9,609

Türkiye'de internet kullanımı verilerde %16,56 olarak gösterilmiştir. Tablo 5.23'de gösterilen değerlere göre Türkiye'nin internet kullanım oranı %1'e de düşse, %100'e de çıksa komşularının internet kullanımını arttırıcı yönde bir etki yapmaktadır. Öyle ki; w_1 ile gösterilen birinci ağırlık matrisi kullanılarak oluşturulan modelde, Türkiye'nin internet kullanımı %1 olsaydı, bu durum Suriye'nin internet kullanımında %0,597 oranının da bir artışa neden olmaktadırken, Türkiye'nin internet kullanımı %100 olsaydı, Suriye'nin oranında %2,769 oranında bir artışa neden olacaktı. Diğer değerler de benzer şekilde yorumlanabilir.

Oransal anlamda en büyük artışlarla, w_3 olarak gösterilen üçüncü ağırlık matrisinin kullanılması ile oluşturulmuş modelin tahminlerinde karşılaşılmıştır.

6. SONUÇ VE ÖNERİLER

Farklı uzamsal ağırlık fonksiyonları kullanılarak uzamsal regresyon analizinin uygulandığı bu çalışmada öncelikle veriler incelenmiş ve eksik gözlemlerin varlığı belirlenmiştir. Ülkelerin internet kullanım oranlarının çeşitli değişkenlerle açıklanmasını sağlamak amacıyla ilk bölümde veriler, bu şekilleriyle alınmış ve en yakın komşuluk ve eşit uzaklık ağırlık matrisleri kullanılarak uzamsal regresyon analizi uygulanmıştır. Bu bölümde verilerdeki eksikliklerin ağırlık fonksiyonlarının kullanılmasını engellediği belirlenmiştir. Bu nedenden dolayı sadece komşuluğu esas alan ağırlık fonksiyonları ile işlemler yapılabilmektedir.

Uygulamanın sonraki bölümünde verilerdeki eksiklikler EM algoritması ile giderilmiş ve kullanılan tüm değişkenler yaklaşık simetrik hale getirilmiştir. Burada da en yakın komşuluk ve eşit uzaklık matrisleri kullanılmış, ancak sadece uzaklıkların dikkate alındığı ağırlık matrisleri ile oluşturulan modeller için sonuçlar yazılmıştır.

Çalışmanın asıl amacı olan uzamsal ağırlık fonksiyonlarının kullanımı bir sonraki bölümde gerçekleştirilmiştir. İlk olarak literatürde sıkça kullanılan ve sadece uzaklığın etkilerinin dahil edildiği iki ağırlık fonksiyonuna ilişkin uzamsal modeller oluşturulmuştur. Sonraki adımda daha önce uygulanmamış olan, öncelikle sadece bu veri seti için tasarlanan, ama parametrelerinin doğru tanımlanabilmesi ile farklı veri setleri için de kullanılması uygun görülen ağırlık fonksiyonu yardımıyla yeni uzamsal model oluşturulmuştur. Bu fonksiyona bölgelerin birbirleri ile karşılaştırılması için uzaklık değerlerinin yanı sıra, içerisinde siyasi, ekonomik, teknolojik ve daha birçok etkiyi barındıran insani gelişim indeksi dahil edilmiştir. Burada karşılaşılan en büyük problem, ağırlıklar oluşturulurken uzaklıklar arttıkça ağırlığın etkisinin azalmasını sağlayacak fonksiyonun düzenlenmesi aşamasında karşımıza çıkmıştır. Bu nedenden dolayı tasarlanan onlarca fonksiyon arasından sadece birkaç tanesi belirlenen kriterler doğrultusunda uzamsal modelin oluşturulmasında kullanılabilmektedir. Diğer büyük zorlukla oluşturulan matrislerin istatistiksel programlarda kullanılabilecek şekilde dönüştürülmesinde karşılaşılmıştır. Analizlerin en başında kullanılan GeoDa programı, fonksiyonların oluşturulmasıyla elde edilen ağırlık matrislerinin kullanılmasında yetersiz kalmış, bunun yerine R

programı ile kodlar yazmak suretiyle çalışmaya devam edilmiştir. Bu bölümde ada ülkelerinin tanımlanması başka bir sorun olarak karşımıza çıkmıştır. Çalışmada sınırdaşlık esas alınıp ağırlık fonksiyonları tanımlandığı için ada ülkelerinin konumları probleme neden olmuştur. Bu nedenle ada ülkelerini ilk adımda modelden çıkarmak yerine, sınır komşuluğu olmamasında rağmen kendisine uzaklık bakımından en yakın olan ülke ile komşu kabul edilmesi düşünülmüştür. Analizler bu şekilde uygulandığında elde edilen uzamsal modellerin tamamının anlamsız çıkmasından dolayı, ada ülkelerinin çalışmadan çıkarılmasına karar verilmiştir. Dolayısıyla 127 olan gözlem sayısı 103'e düşmüştür.

Farklı ağırlık fonksiyonları ile oluşturulan modellerle klasik regresyon analizde olduğu gibi etkin gözlemler belirlenmiştir. Burada çarpıcı olarak Almanya, oluşturulan tüm modellerde, diğer ülkelerden çok farklı bir biçimde etkin olarak karşımıza çıkmıştır. Bunun gerekçesi olarak, dünya ülkelerinin yakın tarihteki teknolojik ve siyasi anlamdaki gelişimleri dikkate alındığında, özellikle Berlin Duvarı'nın yıkılmasının ardından, Almanya'nın diğer Avrupa ülkelerinin gelişimine yetişebilmek için uyguladığı disiplinli politika gösterilebilir. Analizlerde Almanya'nın etkisini giderebilmek adına, bu ülke çalışmadan çıkarılmış ve analizler tekrar uygulanmıştır. Ancak; bu gözlemin çıkarılması durumunda elde edilen sonuçlar, kalan verilere uzamsal regresyon analizinin uygulanamayacağını göstermiştir.

Çalışmanın en son bölümünde elde edilen modellerin geçerliliği sınanmış ve farklı ağırlık fonksiyonları ile oluşturulmuş modeller için Türkiye'nin coğrafi konumunun esas alındığı tahminler yapılmıştır. Dolayısıyla tahminlerde Türkiye'deki değişimlerin komşularını, belirlenen kriterler ile birlikte ne kadar etkilediği belirlenmiştir.

Çalışmada oluşturulan modellerin birbirleri ile karşılaştırılmasında Akaike Bilgi Kriteri ve Log olabilirlik değerleri kullanılmıştır. İzleyen çalışmalarda modellerin birbirleri ile karşılaştırılmasında bu kriterlerin yanı sıra ICOMP kriterinin de kullanılması düşünülmektedir.

KAYNAKLAR

- Albert, D.P., Gesler, W.M., Levergood, B.**, 2000. *Spatial Analysis, GIS and Remote Sensing Applications in the Health Sciences*, CRC Press, Boca Raton.
- Andresen, M.**, 2006. Crime Measures and The Spatial Analysis of Criminal Activity, *Advance Access Publication*, **46**, 258-285.
- Anselin, L.**, 1989. What is Special About Spatial Data? Alternative Perspectives on Spatial Data Analysis, *Symposium of Spatial Statistics, Past, Present and Future*, New York.
- Anselin, L.**, 2005. *Exploring Spatial Data with GeoDa: A Workbook*, Center for Spatially Integrated Social Science, Urbana.
- Arlinghaus, S.L.** (ed.), 1995. *Practical Handbook of Spatial Statistics*, CRC Press, Boca Raton.
- Balçı, C., Yüzbaşıoğlu, N., Ünlü, B., Erkaya, M.** (ed.), 2006. *Dünya Atlası 2006, Encyclopedia millenia world atlas*, gün. 3. bas., Boyut Yayıncılık, İstanbul.
- Bavaud, F.**, 1998. Models for Spatial Weights: A Systematic Look, *Journal of Geographical Analysis*, **30**, 153-171.
- Beguieria, S., Vicente-Serrano, S. M.**, 2006. Mapping the Hazard of Extreme Rainfall by Peaks over Threshold Extreme Value Analysis and Spatial Regression Techniques, *Journal of Applied Meteorology and Climatology*, **45**, 108-124.
- Bhattacharjee, A., Jensen-Butler, C.**, 2006. Estimation of Spatial Weights Matrix, with an Application to Diffusion in Housing Demand.
- Bocci, C.**, 1999. Linear Regression With Spatially Correlated Data, *PhD of Philosophy in Mathematics*, University of Ottawa, Canada.
- Buckeridge, D.**, 1998. A Spatial Analysis of the Relationship Between Respiratory Health and Motor Vehicle Emissions Amongst Small Urban Areas, *Master Thesis of Science in Epidemiology*, Community Health, University of Toronto, Canada.
- Buntgen, U., Frank, D. C., Nievergelt, D., Esper, J.**, 2006. Summer Temperature Variations in the European Alps, A.D. 755-2004, *Journal of Climate*, **19**, 5606-5623.

- Cakmak, S., Jerrett, M., Goldberg, M., Pope, A., Ma, R.**, 2003. Spatial Regression Models for Large-Cohort Studies Linking Community Air Pollution and Health, *Journal of Toxicology and Environmental Health, Part A*, **66**, 1811-1823.
- Caleiro, A., Guerreiro, G.**, 2005. Understanding the Election Results in Portugal, *Portuguese Economic Journal*, **4**, 207-228.
- Chi, G., Zhu, J.**, 2008. Spatial Regression Models for Demographic Analysis, *Population Research and Policy Review*, **27**, 17-42.
- Cliff, A.D., Ord, J.K.**, 1981. *Spatial Processes: Models & Applications*, Pion Limited, London.
- Cliff, A.D., Ord, J.K.**, 1973. *Spatial Autocorrelation*, Pion Limited, London.
- Congdon, P.**, 2006. A Model for Non-Parametric Spatially Varying Regression Effects, *Journal of Computational Statistics and Data Analysis*, **50**, 422-445.
- Cooner, F., Banerjee, S.**, 2006. Modelling Geographically Referenced Survival Data with a Cure Fraction, *Statistical Methods in Medical Research*, **15**, 307-324.
- Cressie, N.A.**, 1993. *Statistics for Spatial Data*, Wiley Series in Probability and Mathematical Statistics, New Jersey.
- Csillag, F., Kabos, S.**, 2008. A Spatial Clustering Perspective on Autocorrelation and Regionalization, *Environ. Ecol. Stat.*, **15**, 385-401.
- Cubukcu, K.M.**, 2006. Şehir İçi Otobüs Hizmeti Maliyetinin Modellenmesinde Coğrafi Bilgi Sistemleri Kullanımı, *Coğrafi Bilgi Sistemleri Bilişim Günleri*, 13-16 Eylül 2006, Fatih Üniversitesi, İstanbul.
- Dale-Johnson, D., Brzeski, W.J.**, 2001. Spatial Regression Analysis of Commercial Land Price Gradient, *Presented at the Asian Real Estate Society Sixth Annual Conference at Keio University*, Mita Campus, July 31 to August 2, Tokyo, Japan.
- Demir, S.**, 2005. Regresyon Fonksiyonlarının Uyarlanabilir Nadaraya-Watson Çekirdek Kestirimleri, *Doktora Tezi*, Hacettepe Üniversitesi FenBilimleri Enstitüsü İstatistik Anabilim Dalı.
- De Silva, D.**, 2005. Synergies in Recurring Procurement Auctions: An Empirical Investigation, *Economic Inquiry*, **43**, 1.
- Diggle, P., Ribeiro, P.**, 2007. *Model-Based Geostatistics*, Springer Series in Statistics, New York.
- Diniz-Filho, J., Hawkins, B., Bini, M., De Marco, P., Blackburn, T.**, 2007. Are Spatial Regression Methods a Panacea or a Pandora's Box?, *Ecography*, **30**, 848-851.
- Dutta, S., Meyer, A., Jain, A., Richter, G.** (ed.), 2006. *The Information Society in an Enlarged Europe*, Springer, Berlin.

- Fischer, M.M., Getis, A.** (ed.), 1997. *Recent Developments in Spatial Analysis: Spatial Statistics, Behavioural Modelling and Computational Intelligence*, Springer, New York.
- Florax, R.J.G.M., Rey, S.**, 1995. *The Impact of Misspecified Spatial Structure in Linear Regression Models*, New Directions in Spatial Econometrics, (ed.) L. Anselin and R.J.G.M. Florax, Springer-Verlag, Berlin.
- Fotheringham, A.S., Brunson, C., Charlton, M.**, 2006. *Geographically Weighted Regression, The Analysis of Spatially Varying Relationships*, John Wiley & Sons Ltd, West Sussex.
- Freisthler, B., Bruce, E., Needell, B.**, 2007. Understanding the Geospatial Relationship of Neighborhood Characteristics and Rates of Maltreatment for Black, Hispanic, and White Children, *Journal of Social Work*, **52**, No.1, 7-16.
- Freisthler, B., Lery, B., Gruenewald, P., Chow, J.**, 2006. Methods and Challenges of Analyzing Spatial Data for Social Work Problems: The Case of Examining Child Maltreatment Geographically, *Social Work Research*, **30**, 4.
- Gamerman, D., Moreira, A.**, 2004. Multivariate Spatial Regression Models, *Journal of Multivariate Analysis*, **91**, 262-281.
- Gao, X., Asami, Y., Chung, C.**, 2006. An Empirical Evaluation of Spatial Regression Models, *Journal of Computers and Geosciences*, **32**, 1040-1051.
- Gatrell, A.C.**, 1983. *Distance and Space: A Geographical Perspective*, Clarendon Press, Oxford.
- Gelfand, A., Vounatsou, P.**, 2003. Proper Multivariate Conditional Autoregressive Models for Spatial Data Analysis, *Journal of Biostatistics*, **4**,1, 11-25.
- Geraci, M., Bottai, M.**, 2006. Use of Auxiliary Data in Semi-Parametric Spatial Regression with Nonignorable Missing Responses, *Statistical Modelling*, **6**, 321-336.
- Gerli, M., Marini, G.**, 2005. Spatial and Temporal Time Series Conversation: A Consistent Estimator of The Error Variance – Covariance Matrix, *Journal of Business Cycle Measurement and Analysis*, **2**, No. 3, 373-405.
- Getis, A., Aldstadt, J.**, 2004. Constructing the Spatial Weights Matrix Using a Local Statistic, *Journal of Geographical Analysis*, Vol. **36**, no. 2, 90-104.
- Getis, A., Mur, J., Zoller, H.G.** (ed.), 2004. *Spatial Econometrics and Spatial Statistics*, Palgrave Macmillan, New York.
- Gezici, F., Hewings, G.**, 2003. Spatial Analysis of Regional Inequalities in Turkey, *43rd Congress of European Regional Science*, 27-30 August, Finland.
- Gonzales, M., Soibelman, I., Formoso, C.**, 2005. A New Approach to Spatial Analysis in CAMA, *Property Management*, Vol. **23**, No. 5, 312-327.
- Griffith, D., Wong, D.**, 2007. Modeling Population Density Across Major US Cities: A Polycentric Spatial Regression Approach, *Journal of Geographical Systems*, **9**, 53-75.

- Griffith, D.A.**, 1993. *Spatial Regression Analysis on The PC: Spatial Statistics Using SAS*, Association of American Geographers, Washington.
- Griffith, D.A., Layne, L.J.**, 1999. *A Casebook for Spatial Statistical Data Analysis; A Compilation of Analyses of Different Thematic Data Sets*, Oxford University Press, New York.
- Haining, R.P.**, 2003. *Spatial Data Analysis: Theory and Practice*, Cambridge University Press, New York.
- Hsu, C., Mas, F.S., Hickey, J.M., Miller, J., Lai, D.**, 2006. Surveillance of the Colorectal Cancer Disparities Among Demographic Subgroups: A Spatial Analysis, *Southern Medical Journal*, Vol. **99**, Number 9, 949-956.
- Huang, Y., Leung, Y.**, 2002. Analysing regional industrialisation in Jiangsu province using geographically weighted regression, *Journal of Geographically Systems*, **4**, 233-249.
- Hubbard, K.G., You, J.**, 2005. Sensitivity Analysis of Quality Assurance Using the Spatial Regression Approach – A Case Study of the Maximum/Minimum Air Temperature, *Journal of Atmospheric and Oceanic Technology*, Vol.**22**, 1520-1530.
- Isik, O., Pinarcioglu, M.M.**, 2007. Geographies of a Silent Transition: A Geographically Weighted Regression Approach to Regional Fertility Differences in Turkey, *Eur J Population*, **22**, 399-421.
- Joines, J., Picciotto, I. H., Carey, T., Gesler, W., Suchindra, C.**, 2003. A Spatial Analysis of County-Level Variation in Hospitalization Rates for Low Back Problems in North Carolina, *Journal of Social Science & Medicine*, **56**, 2541-2553.
- Kaimowitz, D., Mendez, P., Puntodewo, A., Vanclay, J.**, 2002. *Spatial Regression Analysis of Deforestation in Santa Cruz, Bolivia*, C.H. Wood and R. Porro (eds) *Land Use and Deforestation in the Amazon*, University Press of Florida, p. 41-65.
- Kosfeld, R., Lauridsen, J.**, 2006. A test strategy for spurious spatial regression, spatial nonstationarity, and spatial cointegration, *Papers in Regional Sciences*, Vol. **85**, No. 3.
- Lauridsen, J.**, 1996. Dependence and Regimes in Applied Spatial Regression Analysis, *36. European Regional Science Association Congress*, Zurich.
- Law, S.**, 2000. Spatial Analysis of Multivariate Demographic Data for Identifying Communities in GIS, *PhD Thesis in Philosophy*, University of New Brunswick.
- Lehmann, A., Overton, J.M., Leathwick, J.R.**, 2003. GRASP: Generalized Regression Analysis and Spatial Prediction, *Journal of Ecological Modelling*, 165-183.

- Leishman, C., Watkins, C.**, 2002. Estimating Local Repeat Sales House Price Indices for British Cities, *Journal of Property Investment and Finance*, Vol. **20**, No. 1, 36-58.
- Leitner, M., Brecht, H.**, 2007. Crime Analysis and Mapping with GeoDa 0.9.5-i, *Social Science Computer Review*, Vol. **25**, Number 2, 265-271.
- LeSage, J.P.**, 1999. *The Theory and Practice of Spatial Econometrics*, Department of Economics, University of Toledo.
- LeSage, J.P.**, 1997. Regression Analysis of Spatial Data, *The Journal of Regional Analysis and Policy*, **27**, 2, 83-94.
- Leyland, A., Langford, I., Rasbash, J., Goldstein, H.**, 2000. Multivariate Spatial Models for Event Data, *Statistics in Medicine*, **19**, 2469-2478.
- Lin, T., Wu, C., Lee, F.**, 2006. "Neighborhood" Influence on the Formation of National Identity in Taiwan: Spatial Regression with Disjoint Neighborhoods, *Journal of Political Research Quarterly*, Vol.**59**, No.1, 35-46.
- Petrucci, A., Salvati, N., Seghieri, C.**, 2003. The Application of a Spatial Regression Model to the Analysis and Mapping of Poverty, *Food and Agriculture Organization of the United Nations*, Rome.
- Ripley, B.D.**, 2004. *Spatial Statistics*, Wiley Series in Probability and Statistics, New Jersey.
- Robertson, D.M., Saad, D.A., Heisey, D.M.**, 2005. A Regional Classification Scheme for Estimating Reference Water Quality in Streams Using Land-Use-Adjusted Spatial Regression-Tree Analysis, *Journal of Environmental Management*, Vol.**37**, No.2, 209-229.
- Royle, J., A., Berliner, M.**, 1999. A Hierarchical Approach to Multivariate Spatial Modeling and Prediction, *Journal of Agricultural, Biological, and Environmental Statistics*, Vol. **4**, No. 1, 29-56.
- Schabenberger, O., Gotway, C.A.**, 2005. *Statistical Methods for Spatial Data Analysis*, Chapman & Hall/CRC, Boca Raton.
- Schnabel, U., Tietje, O.**, 2003. Explorative Data Analysis of Heavy Metal Contaminated Soil Using Multidimensional Spatial Regression, *Environmental Geology*, **44**, 893-904.
- Shearmur, R., Apparicio, P., Lizion, P., Polese, M.**, 2007. Space, Time and Local Employment Growth: An Application of Spatial Regression Analysis, *Journal of Growth and Change*, Vol. **38**, No. 4, 696-722.
- Shearmur, R., Apparicio, P., Lizion, P., Polese, M.**, 2007. Space, Time, and Local Employment Growth: An Application of Spatial Regression Analysis, *Growth and Change*, Vol. **38**, No. 4, 696-722.
- Shi, H., Zhang, L., Liu, J.**, 2006. A New Spatial-Attribute Weighting Function for Geographically Weighted Regression, *Canadian Journal of Forest Research*, **36**, 4.

- Stetzer, F.**, 1982. Specifying Weights in Spatial Forecasting Models: The Results of Some Experiments, *Environment and Planning A*, Vol. **14**, Issue 5, 571–584.
- Tiefelsdorf, M.**, 1988. Modelling Spatial Processes The Identification and Analysis of Spatial Relationships in Regression Residuals by Means of Moran's I, *PhD Thesis in Philosophy*, Free University of Berlin, Germany.
- Upton, G.**, 1991. Rectangular Cartograms, Spatial Autocorrelation and Interpolation, *Papers in Regional Science*, **70**, 287-302.
- URL-1**, <http://www.r-project.org>, 4 Ocak 2008.
- URL-2**, <http://www.insead.edu/v1/gitr/wef/main/analysis/choosedatavariablen.cfm>, 10 Ocak 2008.
- URL-3**, <http://geodacenter.asu.edu>, 4 Ocak 2008.
- URL-4**, <http://hdr.undp.org/en/reports/global/hdr2007-2008>, 9 Şubat 2008.
- Varga, A.**, 1998. *University Research and Regional Innovation: A Spatial Econometric Analysis of Academic Technology Transfers*, 2. Edition, Kluwer Academic Publishers, Springer, New York.
- Voss, P., Long, D., Hammer, R.**, 2006. County Child Poverty Rates in the US: a Spatial Regression Approach, *Popul Res Policy Rev*, **25**, 369-391.
- Vucetic, S., Fiez, T., Obradovic, Z.**, 1999. A Data Partitioning Scheme for Spatial Regression, *INEEL University Research Consortium Project, No. C94-175936*.
- Wackernagel, H.**, 2003. *Multivariate Geostatistics, An Introduction with Applications*, Springer, New York.
- Wakefield, J.**, 2007. Disease Mapping and Spatial Regression with Count Data, *Journal of Biostatistics*, **8**, 2, 158-183.
- Wheeler, D., Tiefelsdorf, M.**, 2005. Multicollinearity and correlation among local regression coefficients in geographically weighted regression, *Journal of Geographical Systems*, **7**, 161-187.
- Zhang, L., Gove, J. H.**, 2005. Spatial Assessment of Model Errors from Four Regression Techniques, *Journal of Forest Science*, **51**, 4, 334-346.
- Zhang, L., Shi, H.**, 2004. Local Modelling of Tree Growth by Geographically Weighted Regression, *Forest Science*, **50**, 2, 225.

EKLER

EK – 1. İnternet Kullanımına Göre Ülkelerin Kümelenmesi

1	Bangladeş Botswana Burkina Faso Burundi Kamboçya Kamerun Çad Etiyopya Nepal	Gambiya Lesoto Libya Madagaskar Mali Moritanya Mozambik Namibya	Nijerya Paraguay Uganda Zambiya Sri Lanka Tacikistan Tanzanya Nikaragua
2	Cezayir Ermenistan Azerbaycan Benin Bolivya Çin El Salvador Gürcistan	Hindistan Endonezya Kazakistan Kenya Kırgızistan Moğolistan Pakistan Guatemala	Suriname Suriye Zimbabve Panama Filipinler Senegal Güney Afrika
3	Arnavutluk Arjantin Bahreyn Brezilya Kolombiya Ekvator Yunanistan Guyana Venezüella	Morityus Meksika Moldova Fas Umman Peru Rusya Sudi Arabistan Uruguay	Tayland TrinidadandTobago Tunus Türkiye Vietnam Ürdün Makedonya Ukrayna
4	Belçika Bosna Hersek Bulgaristan Şili Kosta Rica Hırvatistan Kıbrıs Çek Cumhuriyeti	Dominik Cumhuriyeti Mısır Macaristan İrlanda İsrail Kuveyt Litvanya Malezya Malta	Katar Romanya Slovakya Birleşik Arap Emirlikleri Portekiz Polonya Singapur İspanya Porto Riko
5	Avustralya Avusturya Barbados Kanada Danimarka Estonya Finlandiya Fransa	Almanya Hong Kong İzlanda İtalya Jameika Japonya Kore Litvanya	Luksemburg Norveç Slovenya İsveç Hollanda Yeni Zellanda İsviçre Taiwan

EK – 2. En Yakın Komşuluk Matrisi

0 127 ozlem ID					
1 68	2.06155281	18 71	4.47213595	35 46	2.70416346
1 43	2.82842712	18 12	5.55990108	35 84	3.58503835
1 44	3.60555128	18 85	10.4403065	35 27	5.74042681
1 17	3.60555128	18 100	12.0415946	35 89	9.66190975
1 48	6	18 73	12.2065556	36 63	2.23606798
2 78	8.94427191	19 118	4.74236228	36 66	3.60555128
2 116	12.5299641	19 113	5.68242906	36 38	5
2 71	13.0384048	19 59	9.08240056	36 93	9.21954446
2 72	13.6041354	19 126	11.7	36 108	11.4017543
2 105	13.892444	19 37	13.8452158	37 59	7
3 123	8.06225775	20 125	3.16227766	37 118	9.21954446
3 24	8.06225775	20 114	5.38516481	37 19	13.8452158
3 90	12.5299641	20 101	11.8776597	37 113	14.3178211
3 13	17.0293864	20 47	12.9047472	37 99	18.3847763
3 16	25.6320112	20 70	12.9529919	38 36	5
4 7	2.3194827	21 85	5.65685425	38 63	7.07106781
4 41	2.62488095	21 12	10.3880941	38 66	8.24621125
4 110	8.60232527	21 23	11.4017543	38 108	11.1803399
4 117	10.4048756	21 18	15.6524758	38 93	13.4164079
4 61	10.707474	21 71	19.4164878	39 11	4.9244289
5 51	25.5538647	22 122	22.090722	39 67	5.35840461
5 70	35.8745035	22 75	37.6563408	39 82	7.18279194
5 101	40.8386924	22 44	44.9623454	39 105	8.48528137
5 92	41.484937	22 55	45.4391076	39 42	8.60232527
5 83	43.324358	22 46	45.8332848	40 100	2.44826469
6 103	1.7149344	23 65	10.198039	40 73	7.99962499
6 28	2.96984848	23 21	11.4017543	40 71	12.888522
6 54	4.75184175	23 85	12.083046	40 18	14.3427334
6 42	5.6639209	23 34	16.2788206	40 12	18.9135005
6 17	5.76888204	23 12	17.7879875	41 4	2.62488095
7 4	2.3194827	24 3	8.06225775	41 7	4.34626276
7 41	4.34626276	24 13	14.3178211	41 110	8.78009112
7 110	10.7042048	24 90	14.7648231	41 117	8.82553115
7 61	11.1628849	24 123	15.2970585	41 29	12.4535136
7 117	12.3685084	24 91	20.6155281	42 67	3.28823661
8 96	1.07814656	25 77	11	42 82	3.78054229
8 120	4.17958132	25 47	15.7458725	42 11	5.02493781
8 99	5.42299733	25 9	18.6010752	42 31	5.09901951
8 61	5.93252897	25 125	19.0262976	42 6	5.6639209
8 87	8.33600024	25 111	19.821453	43 68	2.5
9 81	7.21110255	26 124	7.21110255	43 1	2.82842712
9 114	13.453624	26 33	8.00562302	43 14	6.40312424
9 50	13.6014705	26 89	9.43398113	43 17	6.40312424
9 125	17.8885438	26 45	13.0384048	43 97	7.61577311
9 25	18.6010752	26 115	13.0384048	44 35	2.40831892
10 115	2.68931218	27 84	3.16227766	44 46	3.86167062
10 45	8.10631852	27 89	4.12310563	44 84	5.64025709
10 124	8.40430842	27 46	5.50363516	44 27	8.11865137
10 95	8.61527132	27 35	5.74042681	44 89	11.9462337
10 107	9.68671255	27 44	8.11865137	45 107	3.16227766
11 82	2.31138487	28 103	1.26372465	45 115	6.32455532
11 67	2.3478138	28 17	2.91547595	45 124	7.61577311
11 39	4.9244289	28 14	2.91547595	45 10	8.10631852
11 42	5.02493781	28 6	2.96984848	45 26	13.0384048
11 121	6.94622199	28 54	3.82099463	46 84	2.38537209
12 18	5.55990108	29 53	3.97397786	46 35	2.70416346
12 85	5.89173149	29 117	4.47213595	46 44	3.86167062
12 71	9.85456747	29 110	5	46 27	5.50363516
12 21	10.3880941	29 57	5	46 89	8.7
12 100	16.820003	29 34	8.54400375	47 111	6.99517691
13 90	9.21954446	30 104	22.2236923	47 125	10.170177
13 16	12.2065556	30 64	23.9796268	47 92	12.0885276
13 91	13.0384048	30 80	27.5025908	47 20	12.9047472
13 24	14.3178211	30 109	28.1411176	47 25	15.7458725
13 3	17.0293864	30 15	28.7957028	48 102	1.56524758
14 14	0	31 42	5.09901951	48 17	3.60555128
14 28	2.91547595	31 82	5.86451191	48 14	3.60555128
14 48	3.60555128	31 86	6	48 93	5
14 1	3.60555128	31 67	7.62315551	48 28	5.06951674
14 103	4.07492331	31 108	7.81024968	49 52	15.6204994
15 127	6.32455532	32 95	4.18718282	49 121	19.4164878
15 104	7	32 55	6.95215794	49 11	26.3486243
15 80	7	32 124	11.8473626	49 82	26.6681927
15 109	8.47230783	32 115	12.3434193	49 94	27.5769832
15 64	8.47230783	32 10	12.5529439	50 81	10.6301458
16 13	12.2065556	33 26	8.00562302	50 88	12.2065556
16 90	13.3416641	33 91	8.10493677	50 9	13.6014705
16 107	14.0356688	33 89	11.3265176	50 106	13.6014705
16 45	15.5241747	33 27	13.7437258	50 112	19.9248588
16 91	21	33 124	15.0894003	51 70	10.6103723
17 14	0	34 53	6.18809341	51 101	17.6521613
17 28	2.91547595	34 57	7.21110255	51 92	18.1107703
17 48	3.60555128	34 29	8.54400375	51 20	23.430749
17 1	3.60555128	34 110	11.3137085	51 125	25.2388589
17 103	4.07492331	34 117	13	52 121	6.08276253
		35 44	2.40831892	52 39	12.2065556

EK – 3. Eşik Uzaklık Matrisi

0 127 ozlem ID	2 41	42.6625128	4 76	17.4642492
1 87 42.0594817	2 42	23.7697286	4 82	41.4185043
1 117 15.132746	2 43	21.9544984	4 86	41.3400532
1 110 18.973666	2 48	25.4950976	4 93	27.7308492
1 61 27.9651569	2 49	42.5440948	4 97	20.880613
1 57 18.8679623	2 52	27.3130006	4 102	27.0379363
1 53 17.4038071	2 53	31.622658	4 103	31.1079572
1 43 2.82842712	2 54	17.3349358	4 108	37.2021505
1 34 17.2046505	2 57	33.1360831	4 119	15.8113883
1 29 14.3178211	2 61	42.4199246	5 83	43.324358
1 7 27.3089729	2 63	36.4005494	5 51	25.5538647
1 4 25.019992	2 66	35	5 70	35.8745035
1 120 38.0131556	2 67	21.6728517	5 92	41.484937
1 118 41.761226	2 68	23.307724	5 101	40.8386924
1 99 29.6816442	2 72	13.6041354	6 99	38.7824703
1 96 34.8828396	2 76	32.2024844	6 65	22.5228773
1 65 16.2788206	2 78	8.94427191	6 23	32.7181907
1 37 37.5898923	2 82	24.4231959	6 8	42.756016
1 23 26.0192237	2 86	34.7131099	6 117	23.2912001
1 8 33.8365025	2 93	29.4108823	6 110	27.6383791
1 116 11.1803399	2 94	15.7699081	6 97	11.86086
1 105 24.0208243	2 97	28.4253408	6 76	15.8012658
1 94 28.0515597	2 102	26.1122577	6 68	10.4847508
1 78 26.5706605	2 103	21.4136639	6 61	36.840874
1 72 7.88495403	2 105	13.892444	6 57	27.9692688
1 2 21.4009346	2 108	36.0555128	6 53	26.540017
1 85 33.2415403	2 110	35.6931366	6 48	6.80294054
1 73 38.2753184	2 116	12.5299641	6 43	12.0283
1 71 33.9411255	2 117	33.8378486	6 41	30.5458672
1 21 35.9026461	2 119	35.8050276	6 34	26.2731802
1 18 35.6089876	2 121	26.4764046	6 29	23.256827
1 12 36.3801113	3 13	17.0293864	6 17	5.76888204
1 6 9.20217366	3 16	25.6320112	6 14	5.76888204
1 11 18.6077941	3 24	8.06225775	6 7	34.7910908
1 14 3.60555128	3 26	38.8329757	6 4	32.6049076
1 17 3.60555128	3 33	34.6538598	6 1	9.20217366
1 28 6.23698645	3 45	39.3192065	6 116	9.21303425
1 31 18.0277564	3 90	12.5299641	6 105	18.6461792
1 36 18.973666	3 91	26.8328157	6 103	1.7149344
1 38 23.7697286	3 107	38.8329757	6 94	22.6241022
1 39 18.6815417	3 123	8.06225775	6 78	23.712444
1 41 23.3214494	3 124	42.0475921	6 72	11.7563812
1 42 14.8660687	4 1	25.019992	6 54	4.75184175
1 48 6	4 87	22.4722051	6 39	11.2641023
1 52 30.4630924	4 50	37.7359245	6 28	2.96984848
1 54 7.64852927	4 112	26.0192237	6 2	21.7412051
1 63 16.7630546	4 88	26.925824	6 100	42.9194595
1 66 15.5241747	4 81	40.8044115	6 85	37.5616826
1 67 16.2669143	4 117	10.0498756	6 73	37.0793743
1 68 2.06155281	4 110	8.60232527	6 71	34.754568
1 76 10.8166538	4 61	10.707474	6 21	41.217472
1 82 18.4226084	4 57	12.7279221	6 18	37.4256596
1 86 23.2594067	4 53	13.6745201	6 12	39.4780002
1 93 11	4 43	23.0217289	6 11	9.77394496
1 97 7.07106781	4 34	19.8494332	6 31	9.36375993
1 102 7.43303437	4 29	13	6 36	17.409193
1 103 7.48765651	4 120	18.3575598	6 38	21.1206061
1 108 21.5870331	4 118	41.1096096	6 42	5.6639209
1 119 14.4222051	4 99	15	6 49	35.9204677
1 121 25.5538647	4 96	15.9346321	6 52	21.979081
2 1 21.4009346	4 65	31.7647603		
2 118 39.6232255	4 59	39.6232255		
2 99 42.1070065	4 37	32.7566787		
2 65 14.3178211	4 23	36.0693776		
2 37 40.3112887	4 8	14.9802837		
2 23 20.6155281	4 116	36.0138862		
2 19 41.3363037	4 72	30.9785813		
2 34 27.0185122	4 6	32.6049076		
2 100 22.0227155	4 7	2.3194827		
2 85 18.6815417	4 11	42.3231615		
2 73 17	4 14	27.2946881		
2 71 13.0384048	4 17	27.2946881		
2 40 24.3046086	4 28	30.1346976		
2 21 23.7697286	4 31	38.4837628		
2 18 15.8113883	4 36	26.8700577		
2 12 18.7193082	4 38	30.6104557		
2 6 21.7412051	4 41	2.62488095		
2 11 22.5222113	4 42	37.6430604		
2 14 21.9317122	4 48	25.96151		
2 17 21.9317122	4 54	32.596012		
2 28 21.0641876	4 58	24.3515913		
2 29 30.8058436	4 62	30.016662		
2 31 28.8617394	4 63	26.2488095		
2 36 38.6005181	4 66	26.4007576		
2 38 42.7200187	4 67	40.0313939		
2 39 18.0277564	4 68	23.0488611		

EK – 4. w₁ Ağırlık Matrisi

Ulke	Komsu	w1						
1	43	0,999950001	33	26	0,999550101	75	44	0,998578789
1	43	0,999950001	33	91	0,999328004	76	119	0,999861121
2	116	0,999444599	34	53	0,999550101	76	97	0,999861121
2	65	0,998578789	34	65	0,998750781	77	98	0,992423293
2	85	0,998578789	35	46	0,999911115	77	25	0,997780245
2	71	0,998395733	35	44	0,999911115	78	2	0,999061552
2	73	0,997552999	36	98	0,970035876	79	113	0,999644508
2	78	0,999061552	36	63	0,999950001	79	104	0,997314723
3	24	0,999727815	37	59	0,999444599	79	127	0,999644508
3	13	0,995338676	38	108	0,99920032	79	126	0,999444599
3	90	0,997314723	38	86	0,999061552	80	126	0,998201619
3	16	0,9911505	38	98	0,971610767	80	15	0,999550101
3	123	0,998911704	39	11	0,999644508	80	104	0,999061552
4	7	0,999950001	39	42	0,999444599	81	25	0,996251488
4	117	0,99920032	39	109	0,99980002	81	50	0,999444599
6	42	0,999644508	39	54	0,999328004	82	42	0,999861121
6	30	0,999727815	39	105	0,99920032	82	11	0,999977778
6	48	0,99980002	41	98	0,976801108	84	46	0,999911115
6	103	0,99980002	41	7	0,999861121	84	27	0,999911115
6	54	0,99980002	41	117	0,999061552	85	21	0,999727815
7	41	0,999861121	42	93	0,999328004	85	12	0,99980002
7	98	0,978191324	42	30	0,999727815	86	108	0,999861121
7	4	0,999950001	42	6	0,999644508	86	38	0,999061552
9	50	0,998911704	42	109	0,99980002	87	99	0,998750781
11	82	0,999977778	42	39	0,999444599	87	120	0,999861121
11	39	0,999644508	42	82	0,999861121	88	25	0,990704583
12	85	0,99980002	43	1	0,999950001	88	50	0,998750781
12	115	0,999977778	43	68	0,999950001	89	26	0,999444599
13	16	0,998578789	43	17	0,99980002	89	27	0,999861121
13	90	0,99920032	43	117	0,998750781	90	13	0,99920032
13	3	0,995338676	44	75	0,998578789	90	16	0,998201619
13	24	0,995012479	44	46	0,998395733	90	3	0,997314723
13	91	0,998578789	44	35	0,999911115	91	33	0,999328004
14	28	0,999950001	45	107	0,999911115	91	26	0,998578789
15	80	0,999550101	45	16	0,997996454	91	16	0,996251488
15	104	0,999444599	45	124	0,999550101	91	13	0,998578789
15	126	0,999328004	46	84	0,999911115	93	98	0,962450905
15	127	0,999644508	46	35	0,999911115	93	30	0,99980002
16	124	0,996533799	46	44	0,998395733	93	42	0,999328004
16	45	0,997996454	48	102	0,999950001	93	119	0,998750781
16	107	0,998395733	48	28	0,999911115	93	102	0,999727815
16	123	0,995012479	48	6	0,99980002	94	105	0,999861121
16	3	0,991585601	48	97	0,99980002	97	119	0,999550101
16	90	0,998201619	50	88	0,998750781	97	76	0,999861121
16	13	0,998578789	50	25	0,99321755	97	17	0,999861121
16	26	0,995338676	50	81	0,999444599	97	48	0,99980002
16	91	0,996251488	50	9	0,998911704	98	25	0,982728667
17	97	0,999861121	51	70	0,997780245	98	77	0,992423293
17	117	0,999061552	52	121	0,99980002	98	58	0,9911505
17	43	0,99980002	53	34	0,999550101	98	7	0,978191324
18	71	0,99980002	54	39	0,999328004	98	41	0,976801108
19	113	0,999727815	54	109	0,999861121	98	119	0,971610767
20	114	0,999727815	54	6	0,99980002	98	36	0,970035876
20	125	0,99980002	54	103	0,999911115	98	38	0,971610767
21	85	0,999727815	57	110	0,999861121	99	57	0,998750781
21	23	0,998911704	57	99	0,998750781	99	61	0,999644508
22	122	0,987281572	58	98	0,9911505	99	120	0,999328004
23	65	0,998578789	58	25	0,983930514	99	87	0,998750781
23	21	0,998911704	58	62	0,998578789	100	73	0,999550101
23	85	0,998578789	59	37	0,999444599	100	71	0,998911704
24	13	0,995012479	59	113	0,999550101	102	30	0,999861121
24	3	0,999727815	59	118	0,999727815	102	48	0,999950001
25	98	0,982728667	61	99	0,999644508	102	93	0,999727815
25	77	0,997780245	62	58	0,998578789	103	6	0,99980002
25	60	0,996533799	62	25	0,991585601	103	28	0,999977778
25	125	0,996251488	62	112	0,999911115	103	54	0,999911115
25	50	0,99321755	63	36	0,999950001	104	80	0,999061552
25	81	0,996251488	63	98	0,970035876	104	15	0,999444599
25	88	0,990704583	63	66	0,999911115	104	127	0,998750781
25	58	0,983930514	64	104	0,999861121	104	79	0,997314723
25	62	0,991585601	65	34	0,998750781	105	39	0,99920032
25	112	0,990704583	65	23	0,998578789	105	94	0,999861121
26	124	0,999328004	65	2	0,998578789	107	45	0,999911115
26	16	0,995338676	65	116	0,999061552	107	16	0,998395733
26	91	0,998578789	66	98	0,968420496	108	38	0,99920032
26	33	0,999550101	68	43	0,999950001	108	86	0,999061552
26	89	0,999444599	70	51	0,997780245	109	42	0,99980002
27	84	0,999911115	70	114	0,998201619	109	54	0,999861121
27	89	0,999861121	71	73	0,999328004	109	39	0,99980002
28	103	0,999977778	71	2	0,998395733	110	117	0,999644508
28	48	0,999911115	71	18	0,99980002	110	57	0,999861121
28	14	0,999950001	71	100	0,998911704	112	62	0,999911115
30	42	0,999727815	73	2	0,997552999	112	25	0,990704583
30	102	0,999861121	73	71	0,999328004	113	59	0,999550101
30	6	0,999727815	73	100	0,999550101	113	79	0,999644508
30	93	0,99980002	75	122	0,996251488	113	126	0,99920032

EK – 5. w₂ Ağırlık Matrisi

Ulke	Komsu	w ₂						
1	43	5,712992101	33	26	4,108915851	75	44	3,457514461
1	43	5,712992101	33	91	3,868852695	76	119	4,901274189
2	116	3,981071706	34	53	4,108915851	76	97	4,901274189
2	65	3,457514461	34	65	3,525109584	77	98	2,688682267
2	85	3,457514461	35	46	5,240611947	77	25	3,233635033
2	71	3,395199698	35	44	5,240611947	78	2	3,679738819
2	73	3,186648866	36	98	2,183888614	79	113	4,256699613
2	78	3,679738819	36	63	5,712992101	79	104	3,142484873
3	24	4,430682073	37	59	3,981071706	79	127	4,256699613
3	13	2,892548174	38	108	3,769169136	79	126	3,981071706
3	90	3,142484873	38	86	3,679738819	80	126	3,337476746
3	16	2,626527804	38	98	2,201913484	80	15	4,108915851
3	123	3,598832131	39	11	4,256699613	80	104	3,679738819
4	7	5,712992101	39	42	3,981071706	81	25	2,988876673
4	117	3,769169136	39	109	4,640391524	81	50	3,981071706
6	42	4,256699613	39	54	3,868852695	82	42	4,901274189
6	30	4,430682073	39	105	3,769169136	82	11	6,451950121
6	48	4,640391524	41	98	2,270523457	84	46	5,240611947
6	103	4,640391524	41	7	4,901274189	84	27	5,240611947
6	54	4,640391524	41	117	3,679738819	85	21	4,430682073
7	41	4,901274189	42	93	3,868852695	85	12	4,640391524
7	98	2,291911455	42	30	4,430682073	86	108	4,901274189
7	4	5,712992101	42	6	4,256699613	86	38	3,679738819
9	50	3,598832131	42	109	4,640391524	87	99	3,525109584
11	82	6,451950121	42	39	3,981071706	87	120	4,901274189
11	39	4,256699613	42	82	4,901274189	88	25	2,607142942
12	85	4,640391524	43	1	5,712992101	88	50	3,525109584
12	115	6,451950121	43	68	5,712992101	89	26	3,981071706
13	16	3,457514461	43	17	4,640391524	89	27	4,901274189
13	90	3,769169136	43	117	3,525109584	90	13	3,769169136
13	3	2,892548174	44	75	3,457514461	90	16	3,337476746
13	24	2,863278707	44	46	3,395199698	90	3	3,142484873
13	91	3,457514461	44	35	5,240611947	91	33	3,868852695
14	28	5,712992101	45	107	5,240611947	91	26	3,457514461
15	80	4,108915851	45	16	3,283778989	91	16	2,988876673
15	104	3,981071706	45	124	4,108915851	91	13	3,457514461
15	126	3,868852695	46	84	5,240611947	93	98	2,109971418
15	127	4,256699613	46	35	5,240611947	93	30	4,640391524
16	124	3,024252145	46	44	3,395199698	93	42	3,868852695
16	45	3,283778989	48	102	5,712992101	93	119	3,525109584
16	107	3,395199698	48	28	5,240611947	93	102	4,430682073
16	123	2,863278707	48	6	4,640391524	94	105	4,901274189
16	3	2,646553136	48	97	4,640391524	97	119	4,108915851
16	90	3,337476746	50	88	3,525109584	97	76	4,901274189
16	13	3,457514461	50	25	2,733880874	97	17	4,901274189
16	26	2,892548174	50	81	3,981071706	97	48	4,640391524
16	91	2,988876673	50	9	3,598832131	98	25	2,374343733
17	97	4,901274189	51	70	3,233635033	98	77	2,688682267
17	117	3,679738819	52	121	4,640391524	98	58	2,626527804
17	43	4,640391524	53	34	4,108915851	98	7	2,291911455
18	71	4,640391524	54	39	3,868852695	98	41	2,270523457
19	113	4,430682073	54	109	4,901274189	98	119	2,201913484
20	114	4,430682073	54	6	4,640391524	98	36	2,183888614
20	125	4,640391524	54	103	5,240611947	98	38	2,201913484
21	85	4,430682073	57	110	4,901274189	99	57	3,525109584
21	23	3,598832131	57	99	3,525109584	99	61	4,256699613
22	122	2,486724246	58	98	2,626527804	99	120	3,868852695
23	65	3,457514461	58	25	2,400390349	99	87	3,525109584
23	21	3,598832131	58	62	3,457514461	100	73	4,108915851
23	85	3,457514461	59	37	3,981071706	100	71	3,598832131
24	13	2,863278707	59	113	4,108915851	102	30	4,901274189
24	3	4,430682073	59	118	4,430682073	102	48	5,712992101
25	98	2,374343733	61	99	4,256699613	102	93	4,430682073
25	77	3,233635033	62	58	3,457514461	103	6	4,640391524
25	60	3,024252145	62	25	2,646553136	103	28	6,451950121
25	125	2,988876673	62	112	5,240611947	103	54	5,240611947
25	50	2,733880874	63	36	5,712992101	104	80	3,679738819
25	81	2,988876673	63	98	2,183888614	104	15	3,981071706
25	88	2,607142942	63	66	5,240611947	104	127	3,525109584
25	58	2,400390349	64	104	4,901274189	104	79	3,142484873
25	62	2,646553136	65	34	3,525109584	105	39	3,769169136
25	112	2,607142942	65	23	3,457514461	105	94	4,901274189
26	124	3,868852695	65	2	3,457514461	107	45	5,240611947
26	16	2,892548174	65	116	3,679738819	107	16	3,395199698
26	91	3,457514461	66	98	2,166486176	108	38	3,769169136
26	33	4,108915851	68	43	5,712992101	108	86	3,679738819
26	89	3,981071706	70	51	3,233635033	109	42	4,640391524
27	84	5,240611947	70	114	3,337476746	109	54	4,901274189
27	89	4,901274189	71	73	3,868852695	109	39	4,640391524
28	103	6,451950121	71	2	3,395199698	110	117	4,256699613
28	48	5,240611947	71	18	4,640391524	110	57	4,901274189
28	14	5,712992101	71	100	3,598832131	112	62	5,240611947
30	42	4,430682073	73	2	3,186648866	112	25	2,607142942
30	102	4,901274189	73	71	3,868852695	113	59	4,108915851
30	6	4,430682073	73	100	4,108915851	113	79	4,256699613
30	93	4,640391524	75	122	2,988876673	113	126	3,769169136

EK – 6. w₃ Ağırlık Matrisi

Ulke	Komsu	w ₃							
1	43	5,796448917	33	26	4,118918185	75	44	3,394145646	
1	116	3,998641597	33	91	3,86935355	76	119	4,95402615	
2	65	3,495658012	34	53	4,223432063	76	97	4,969523258	
2	85	3,307222345	34	65	3,576387983	77	98	2,725505832	
2	71	3,179313	35	46	5,21510519	77	25	3,267557955	
2	73	3,096421562	35	44	5,20685149	78	2	3,726525866	
2	78	3,63353919	36	98	2,168692991	79	113	4,340817922	
3	24	4,429661295	36	63	5,709661872	79	104	3,324344805	
3	13	2,828636255	37	59	4,081606839	79	127	4,381790586	
3	90	3,098602831	38	108	3,770749834	79	126	4,03010026	
3	16	2,604887812	38	86	3,685876972	80	126	3,205353156	
3	123	3,591729082	38	98	2,164482072	80	15	4,11143744	
4	7	5,691245712	39	11	4,254009179	80	104	3,69310494	
4	117	3,769169136	39	42	3,973904868	81	25	3,103101433	
6	42	4,250826029	39	109	4,641851763	81	50	4,040318689	
6	30	4,40329239	39	54	3,864358967	82	42	4,891937161	
6	48	4,602829993	39	105	3,767979685	82	11	6,447195281	
6	103	4,624989219	41	98	2,284579604	84	46	5,2331836	
6	54	4,636953636	41	7	4,896048901	84	27	5,333265498	
7	41	4,906505054	41	117	3,689861202	85	21	4,485924557	
7	98	2,308561184	42	93	3,841076516	85	12	4,606732373	
7	4	5,734821584	42	30	4,409376644	86	108	4,895164066	
9	50	3,643610304	42	6	4,262581312	86	38	3,674334247	
11	82	6,456708469	42	109	4,650223226	87	99	3,547824711	
11	39	4,259391747	42	39	3,988251468	87	120	4,932857057	
12	85	4,674296606	42	82	4,910629039	88	25	2,69830964	
12	115	6,8660237	43	1	5,63074392	88	50	3,566371063	
13	16	3,506505492	43	68	5,63074392	89	26	3,970654007	
13	90	3,800509661	43	17	4,586551215	89	27	4,921420026	
13	3	2,957904158	43	117	3,462913902	90	13	3,738087057	
13	24	2,927298788	44	75	3,522066374	90	16	3,356854766	
13	91	3,494487362	44	46	3,400581638	90	3	3,186988367	
14	28	5,745581133	44	35	5,274591303	91	33	3,868351905	
15	80	4,106395809	45	107	5,257145202	91	26	3,465482453	
15	104	3,993081872	45	16	3,305040548	91	16	2,999155923	
15	126	3,713414189	45	124	4,131365695	91	13	3,420932746	
15	127	4,154578464	46	84	5,248050839	93	98	2,092869186	
16	124	3,021214197	46	35	5,266243457	93	30	4,651472632	
16	45	3,262654206	46	44	3,389826276	93	42	3,896829733	
16	107	3,384000519	48	102	5,705760786	93	119	3,490384881	
16	123	2,881367043	48	28	5,22604027	93	102	4,427104183	
16	3	2,668539261	48	6	4,678259577	94	105	4,928972279	
16	90	3,318210589	48	97	4,606939789	97	119	4,096102486	
16	13	3,409207906	50	88	3,484325484	97	76	4,833962421	
16	26	2,889277465	50	25	2,796743412	97	17	4,907865646	
16	91	2,978632654	50	81	3,922699687	97	48	4,674086159	
17	97	4,894691585	50	9	3,55460426	98	25	2,366836494	
17	117	3,657248272	51	70	3,268736773	98	77	2,652356215	
17	43	4,69486385	52	121	4,634062403	98	58	2,623895991	
18	71	4,652783151	53	34	3,997504688	98	7	2,275381806	
19	113	4,485462921	54	39	3,873351649	98	41	2,256553792	
20	114	4,550567628	54	109	4,908517837	98	119	2,198039213	
20	125	4,735816911	54	6	4,643831962	98	36	2,199190711	
21	85	4,376119879	54	103	5,227089952	98	38	2,239992216	
21	23	3,487013258	57	110	4,869281763	99	57	3,507801171	
22	122	2,484124406	57	99	3,542503401	99	61	4,296404585	
23	65	3,725256344	58	98	2,629162257	99	120	3,894740835	
23	21	3,714236726	58	25	2,395200776	99	87	3,525976876	
23	85	3,524444032	58	62	3,412266001	100	73	4,149095791	
24	13	2,801257053	59	37	3,883012879	100	71	3,502110898	
24	3	4,431703085	59	113	4,064200855	102	30	4,91694884	
25	98	2,381874785	59	118	4,416883592	102	48	5,720232581	
25	77	3,20006429	61	99	4,217361572	102	93	4,434262855	
25	60	3,076109979	62	58	3,50336294	103	6	4,655845123	
25	125	2,971503808	62	25	2,675850193	103	28	6,403183684	
25	50	2,6724313	62	112	5,223030778	103	54	5,254168923	
25	81	2,878856513	63	36	5,716324272	104	80	3,666421072	
25	88	2,519056457	63	98	2,169957907	104	15	3,969097663	
25	58	2,405591167	63	66	5,244886781	104	127	3,430191521	
25	62	2,617576843	64	104	5,002853136	104	79	2,970573678	
25	112	2,569947477	65	34	3,474566417	105	39	3,770358962	
26	124	3,869341526	65	23	3,209015741	105	94	4,873731748	
26	16	2,895822585	65	2	3,41978712	107	45	5,224130689	
26	91	3,449564789	65	116	3,655649429	107	16	3,40643594	
26	33	4,098937807	66	98	2,150911951	108	38	3,7675891	
26	89	3,991516736	68	43	5,79644168	108	86	3,684331853	
27	84	5,149568045	70	51	3,198910238	109	42	4,630580609	
27	89	4,88121082	70	114	3,324920506	109	54	4,894041231	
28	103	6,501087963	71	73	4,01457988	109	39	4,638931745	
28	48	5,255224255	71	2	3,625745874	110	117	4,28577451	
28	14	5,680587915	71	18	4,6280329	110	57	4,933476814	
30	42	4,452090447	71	100	3,698224609	112	62	5,258252296	
30	102	4,885649508	73	2	3,279505323	112	25	2,644876747	
30	6	4,458242127	73	71	3,728415332	113	59	4,154122808	
30	93	4,629336815	73	100	4,069125015	113	79	4,174211385	
			75	122	3,030195119	113	126	3,741647852	

EK – 7. R Programı Kodları

ÜLKELERİN MERKEZİ NOKTALARININ ÇİZİLMESİ

```
*****
dd <- c("C:\\Users\\odeniz.ITICU0\\Desktop\\spatial kitap data\\srm\\")
setwd(dd)
library(RColorBrewer)
library(maptools)
library(spdep)
library(sp)
library(rgdal)
rob.shp <- read.shape("wg2002worldmap.shp")
rob.map <- Map2poly(rob.shp,region.id =
unique(as.character(rob.shp$att.data$FIPS_CNTRY)))
tr <- readShapePoly("wg2002worldmap",
IDvar="FIPS_CNTRY", proj4string=CRS("+proj=robin +lon 0=0"))
ct <- na.omit(rob.shp$att.data[,c(1, 18:20)])
colnames(ct) <- c("ID","x", "y", "City_POP")
ct$x <- as.numeric(as.character(ct$x))
ct$y <- as.numeric(as.character(ct$y))
coordinates(ct) <- c("x", "y")
proj4string(ct) <- CRS("+proj=longlat +datum=WGS84")
ct_rb <- spTransform(ct, CRS=CRS("+proj=robin +lon0=0"))
plot(rob.map, border="Grey", forcefill=T,
xaxt="n", yaxt="n", bty="n",
lwd=.00000000125, las=1, ylab="",
main="Centroids and Capitals", xlab="")
points(coordinates(tr), pch=19,cex=.5, col="grey")
points(coordinates(ct_rb), pch=19, cex=.5, col="black")
tr_or <- coordinates(tr)
rownames(tr_or) <- as.character(attributes(tr)$data)$FIPS_CNTRY
ct_rb_or <- coordinates(ct_rb)
rownames(ct_rb_or) <- as.character(ct_rb$ID)
coord_dif <- cbind(tr_or[-91,], ct_rb_or[rownames(tr_or),][-91,])
x1 <- coord_dif[,1]
x2 <- coord_dif[,3]
y1 <- coord_dif[,2]
y2 <- coord_dif[,4]
segments(x1, y1, x2, y2,col="slategray4")
```

HER BİR AĞIRLIK DEĞERİ İÇİN ETKİN GÖZLEMLERİN MORAN PLOT İLE GÖSTERİLMESİ

```
*****
moran.plot(DonInt,agirlik1w1)
```

HER BİR AĞIRLIK İÇİN MORAN SCATTER GRAFIKLERİNİN GÖSTERİMİ

```
*****
dem <- (resid(tezverisi.lm))
ds <- (DonInt-mean(DonInt))/sqrt(var(DonInt))
ds.slag <- as.vector(agirlik1 %>% ds)
ds.slag <- (ds.slag-mean(ds.slag))/sqrt(var(ds.slag))
plot(ds, ds.slag, xlim=c(-3,3), ylim=c(-3,3), pch=20, las=1,
xlab="standardize internet kullanımı",
ylab="standardize internet kullanımının uzamsal gecikmesi")
reg1 <- lm(ds.slag~ds)
xgrid <- seq(-3,1.5,length.out=158)
x0 <- list(ds=xgrid)
pred.out <- predict(reg1, x0, interval="confidence")
lines( c(-2,-2,+2,+2,-2), c(-2,+2,+2,-2,-2) )
lines( c(-1,-1,+1,+1,-1), c(-1,+1,+1,-1,-1) )
lines( c(-2,+2), c(0,0) )
lines( c(0,0), c(-2,+2) )
text(-2.5,3,"(düşük,yüksek)"); text(2.5,3,"(yüksek,yüksek)")
text(-2.5,-3,"(düşük,düşük)"); text(2.5,-3,"(yüksek,düşük)")
polygon(x=c(-1,0,0,-1), y=c(-1,-1,0,0), col = "slategray3")
polygon(x=c(0,1,1,0), y=c(0,0,1,1), col = "slategray3")
polygon(x=c(xgrid, rev(xgrid)),
y=c(pred.out[,3], rev(pred.out[,2])), col="slategray3", border=T)
points(ds, ds.slag, pch=20)
sldensity <- density(ds.slag)
lines(sldensity$y+2, sldensity$x, lty=2, col="slategray4")
ddensity <- density(ds)
lines(ddensity$x, ddensity$y+2, lty=2, col="slategray4",xlim=c(-2,2))
```

```

lines(xgrid, pred.out[,1], type="l", lty=2, col="gray80", lwd=2)
rug(jitter(ds, factor=2), col="slategray3")
rug(ds.slag, side=2, col="slategray3")
text(-2., -2.3, "Internet Kullanımı", col="slategray4")

```

KLASİK MODEL SONUÇLARI

```

*****
tezverisi <- read.csv("C:\\Users\\odeniz.ITICU0\\Desktop\\tezverisi\\tezverisi.txt",header=TRUE,skip=1)
summary(tezverisi)
lm(DonInt~ElPro+AvLaTe+EduExp+CapInn+DonTax+DonTelLi+DonSecI+DonPerCo+Dummy_EP,data=tezverisi)
tezverisi.lm<-
lm(DonInt~ElPro+AvLaTe+EduExp+CapInn+DonTax+DonTelLi+DonSecI+DonPerCo+Dummy_EP,data=tezverisi)
summary(tezverisi.lm)
library(spdep)
library(geoR)
library(lmtest)
dwtest(tezverisi.lm)

```

AĞIRLIK 1

```

*****
agirlik1 <-
matrix(scan("C:\\Users\\odeniz.ITICU0\\Desktop\\Orijinal_Veri_Agirlik_Denemeleri\\agirliklar\\103agirlikw1.txt", sep=","), nrow=103, ncol=103, byrow=TRUE)
range(rowSums(agirlik1))
agirlik1[1:103,1:103]
agirlik1w1 <- mat2listw(agirlik1)
print(is.symmetric.nb(agirlik1w1$neighbours))
print(agirlik1w1$neighbours)
w1sonuc.lag<-
lagsarlm(DonInt~ElPro+EduExp+CapInn+DonTelLi+DonSecI+DonPerCo,data=tezverisi103gozlem, agirlik1w1)
summary(w1sonuc.lag)
agirlik1w1$weights
tezverisi103gozlem.e<-resid(tezverisi.lm)
tezverisi.morane<-moran.test(tezverisi103gozlem.e,agirlik1w1, randomisation=FALSE,alternative="two.sided")
print(tezverisi.morane)
tezverisi103gozlem.lagrange<-lm.LMtests(tezverisi.lm,agirlik1w1, test=c("LMerr","RLMerr","LMlag","RLMlag","SARMA"))
print(tezverisi103gozlem.lagrange)

```

ÖZGEÇMİŞ

Adı Soyadı : Özlem DENİZ BAŞAR

Doğum Yeri : Edremit / BALIKESİR

Doğum Tarihi : 21.07.1980

Eğitim Durumu

Lise : Çankaya Milli Piyango Anadolu Lisesi (1998)

Lisans : Mimar Sinan Güzel Sanatlar Üniversitesi Fen Edebiyat
Fakültesi İstatistik Bölümü (2002)

Yüksek Lisans : Mimar Sinan Güzel Sanatlar Üniversitesi Fen Bilimleri
Enstitüsü İstatistik A.B.D. (2005)

İş Tecrübesi

2002- : İstanbul Ticaret Üniversitesi Fen Edebiyat Fakültesi İstatistik
Bölümü Araştırma Görevlisi