



T.C.
VAN YÜZÜNCÜ YIL ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ



**KISITLI ORDİNASYON YÖNTEMLERİNDEN KANONİK UYUM
ANALİZİ VE GEREKSİZLİK ANALİZİ YÖNTEMLERİNİN
KARŞILAŞTIRILMASI: KARDİYOLOJİK VERİ SETİ ÜZERİNDE
BİR UYGULAMA**

Mehmet Tahir HUYUT
BİYOİSTATİSTİK ANABİLİM DALI
(TIP PROGRAMI)
DOKTORA TEZİ

DANIŞMAN
Prof. Dr. Sıddık KESKİN

VAN-2019

T.C.
VAN YÜZÜNCÜ YIL ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ

**KISITLI ORDİNASYON YÖNTEMLERİNDEN KANONİK UYUM
ANALİZİ VE GEREKSİZLİK ANALİZİ YÖNTEMLERİNİN
KARŞILAŞTIRILMASI: KARDİYOLOJİK VERİ SETİ ÜZERİNDE
BİR UYGULAMA**

Mehmet Tahir HUYUT
BİYOİSTATİSTİK ANABİLİM DALI
(TIP PROGRAMI)
DOKTORA TEZİ


DANIŞMAN
Prof. Dr. Sıddık KESKİN

VAN-2019

KABUL VE ONAY

Van Yüzüncü Yıl Üniversitesi Sağlık Bilimleri Enstitüsü, Biyoistatistik Anabilim Dalında Mehmet Tahir HUYUT tarafından hazırlanan "*Kısıtlı Ordınasyon Yöntemlerinden Kanonik Uyum Analizi ve Gereksizlik Analizi Yöntemlerinin Karşılaştırılması: Kardiyolojik Veri Seti Üzerinde Bir Uygulama*" adlı tez çalışması aşağıdaki jüri tarafından DOKTORA TEZİ olarak OY BİRLİĞİ/OY ÇOKLUĞU ile kabul/ret edilmiştir.

Tez Savunma Tarihi: 10/04/2019


Prof. Dr. Siddık KESKİN
Yüzüncü Yıl Üniversitesi
Jüri Başkanı

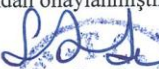

Prof. Dr. Abdullah YEŞİLOVA
Yüzüncü Yıl Üniversitesi
Jüri Üyesi


Doç. Dr. Mahmut KARA
Yüzüncü Yıl Üniversitesi
Jüri Üyesi


Doç. Dr. Serdal Kenan KÖSE
Ankara Üniversitesi
Jüri Üyesi


Doç. Dr. Mustafa Ağah TEKİNDAL
Selçuk Üniversitesi
Jüri Üyesi

Tez hakkında alınan jüri kararı, Van Yüzüncü Yıl Üniversitesi Sağlık Bilimleri Enstitüsü Yönetim Kurulu tarafından onaylanmıştır.


Prof. Dr. Semiha DEDE
Sağlık Bilimleri Enstitüsü Müdürü



ETİK BEYAN

T.C.
VAN YÜZÜNCÜ YIL ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ MÜDÜRLÜĞÜ'NE

Doktora tezi olarak hazırlayıp sunduğum "*Kısıtlı Ordinasyon Yöntemlerinden Kanonik Uyum Analizi ve Gereksizlik Analizi Yöntemlerinin Karşılaştırılması: Kardiyolojik Veri Seti Üzerinde Bir Uygulama*" başlıklı tezim; bilimsel ahlak ve değerlere uygun olarak tarafımdan yazılmıştır. Tezimin fikir/hipotezi tümüyle tez danışmanım ve bana aittir. Tezde yer alan deneysel çalışma/araştırma tarafımdan yapılmış olup, tüm cümleler, yorumlar bana aittir. Bu tezdeki bütün bilgiler akademik kurallara ve etik ilkelere uygun olarak hazırlanıp, bu kural ve ilkeler gereği, çalışmada bana ait olmayan tüm veri, düşünce ve sonuçlara atıf yapılmış ve kaynak gösterilmiştir.

Yukarıda belirtilen hususların doğruluğunu beyan ederim.

Mehmet Tahir HUYUT

25/04/2019

TEŐEKKÜR

Tez alıőmamın gerekleőmesinde yardım ve ilgilerini esirgemeyen, gürüő ve önerileri ile bana yardımcı olan, Tıp Fakóltesi Biyoistatistik Anabilim Dalı Baőkanı danıőman hocam Sayın Prof. Dr. Sıddık KESKİN'e, bana farklı bir bakıő aısı kazandıran ve tezimin őekillenmesinde bana yardımcı olan sayın Prof. Dr. Abdullah YEŐİLOVA'ya ve sayın Do. Dr. Mahmut KARA'ya, alıőmalarım esnasında maddi ve manevi desteklerini esirgemeyen deėerli aėabeyim Tıp Fakóltesi Biyokimya Anabilim Dalı Öğretim Üyesi Do. Dr. Zübeyir HUYUT'a alıőmalarım sırasında gösterdikleri anlayıőtan dolayı eőim Sedanur'a ve kızım Belinay'a ve oėlum Yusuf Eymen'e teőekkürü bir bor bilirim.

ÖZET

Huyut MT, Kısıtlı Ordinasyon Yöntemlerinden Kanonik Uyum Analizi ve Gereksizlik Analizi Yöntemlerinin Karşılaştırılması: Kardiyolojik Veri Seti Üzerinde Bir Uygulama, Van Yüzüncü Yıl Üniversitesi, Sağlık Bilimleri Enstitüsü, Biyoistatistik Anabilim Dalı, Doktora Tezi, Van, 2019.

Doğadaki olaylar, birbirleri ile ilişkili olan çok sayıda değişkenin veya faktörün ayrı ayrı ya da birlikte etkisi ile ortaya çıkmaktadır. Gereksizlik Analizi (Redundancy Analysis, RDA) ve Kanonik Uyum Analizi (Canonical Correspondence Analysis, KUA) bu ilişkileri belirlemede kullanılan analiz yöntemlerindedir. Doğrudan Gradient Analizleri (DGA) olarak da bilinen kısıtlı kanonik analiz yöntemleri ile ilgili yapılan çalışmaların, genellikle çevrebilimciler tarafından yapıldığı ve sağlık alanında sınırlı sayıda çalışmanın olduğu görülmüştür. Bu nedenle tez çalışmasında, KUA ve RDA'nın incelenmesi amaçlanmıştır. Ayrıca, farklı ölçek seviyelerindeki değişkenlere ait verileri içeren Kardiyoloji alanındaki bir veri seti ile yöntemlerin uygulaması yapılarak performansları karşılaştırılmıştır. Performans ölçütleri olarak Belirleme Katsayısı (R^2) ve MAPE (Mean Absolute Percentage Error) değeri hesaplanmıştır. Elde edilen sonuçlara göre farklı ölçek türlerinde olan değişkenler arasındaki ilişki yapılarını analiz eden kısıtlı ordinasyon yöntemleri olan KUA ve RDA'nın varyasyonu yeterli düzeyde açıkladığı görülmüştür. Ayrıca, her iki yöntemin düşük MAPE değeri ile (% 10'dan düşük) iyi sınıflama yaptığı ve ordinasyon diyagramı gerçekleştirdiği böylece sağlık alanında da kullanılabileceği vurgulanmıştır.

Anahtar Sözcükler: Kanonik form, ölçekleme, biplot, boyut indirgeme, MAPE, angina

ABSTRACT

Huyut MT, Comparison of Canonical Correspondence Analysis and Redundant Analysis based on Constrained Ordination Methods: An Application on a Cardiological Data Set, Yuzuncu Yil University, Institute of Health Sciences, Department of Biostatistics, PhD Thesis, Van, 2019. The events in nature arise with the effect of a large number of variables or factors that are related to each other, either individually or together. Redundancy Analysis (RDA) and Canonical Correspondence Analysis (CCA) are the methods of analysis used to determine these relationships. Studies on constrained canonical analysis methods, also known as Direct Gradient Analyzes (DGA), have generally been carried out by ecologists and there are limited number of studies in the field of health. Therefore, it is aimed to examine the CCA and RDA in this study. For this purpose, a data set including variables at different scale levels obtained from cardiology field was used to compare RDA and CCA methods performance. In addition, the coefficients of determination (R^2) and MAPE values were calculated as performance measures of the constrained ordination methods used in the study. According to the findings, it was found that the constrained ordination methods (CCA and RDA), which analyze the relationship structures between the variables found in different scale types, adequately explain the variation. In addition, it is emphasized that both methods have a good classification with a low MAPE value (less than 10%) and produce good ordination diagram thus, they can be used in the health field.

Key words: Canonical form, scaling, biplot, dimension reduction, MAPE, angina

İÇİNDEKİLER

KABUL VE ONAY.....	II
ETİK BEYAN.....	III
TEŞEKKÜR.....	IV
ÖZET.....	V
ABSTRACT.....	VI
İÇİNDEKİLER.....	VII
SİMGELER VE KISALTMALAR.....	IX
ŞEKİLLER LİSTESİ.....	X
TABLolar LİSTESİ.....	XI
1. GİRİŞ.....	1
2. GENEL BİLGİLER.....	3
2. 1. Sınıflandırma ve Ordinasyonu.....	3
2. 2. Ordinasyon Diyagramları.....	3
2. 2. 1. Temel bileşenler diyagramı (Biplotlar).....	5
2. 2. 2. Gradient analizleri (Ordinasyon teknikleri).....	5
2. 3. Gereksizlik Analizi (Redundancy Analysis/RDA).....	7
2. 3. 1. Gereksizlik analizi (RDA) ile kanonik korelasyon analizi (CCorA) arasındaki ilişki.....	11
2. 3. 2. Gereksizlik analizi (RDA) ile temel bileşenler analizi (PCA) arasındaki ilişki.....	11
2. 4. Kanonik Uyum Analizi/KUA (Canonical Correspondence Analysis/CCA).....	12
2. 5. Ağırlıklı Ortalama Ordinasyonu.....	18
3. GEREÇ VE YÖNTEM.....	20
3. 1. Gereç.....	20
3. 2. Yöntem.....	21
3. 2. 1. Kanonik (kısıtlı/sınırlandırılmış) ordinasyon.....	21
3. 2. 2. Gereksizlik analizi (Redundancy Analysis/RDA).....	23
3. 2. 3. Kanonik uyum analizi (canonical correspondence analysis).....	36
3. 2. 4. Kanonik ordinasyonda anlamlılık testleri.....	44

3. 3. Yöntemlerin Performans Değerlendirmesi (Model Performance Evolution).....	45
4. BULGULAR.....	48
4. 1. Kanonik Uyum Analizi (KUA) Uygulaması	48
4. 2. Gereksizlik Analizi (RDA) Uygulaması	52
5. TARTIŞMA VE SONUÇ.....	57
KAYNAKLAR.....	62
ÖZGEÇMİŞ.....	67
EKLER.....	68
EK 1. Terminolojik Sözlük.....	68
EK 2. Etik Kurul Raporu.....	70
EK 3. Tez Orjinallik Raporu.....	71

SİMGELER VE KISALTMALAR

χ^2_L	: Olabilirlik oran istatistiđi
ALS	: Dalgalı En Küçük Kareler Algoritması
CA	: Uyum Analizi
CCorA	: Kanonik Korelasyon Analizi
DCA	: Detrended Uyum Analizi
DGA	: Doğrudan Gradient Analizleri
KUA	: Kanonik Uyum Analizi
MDS	: Çok Boyutlu Ölçekleme
NMDS	: Metrik Olmayan Çok Boyutlu Ölçekleme
RDA	: Gereksizlik Analizi / Redundancy Analysis
TBA	: Temel Bileşenler Analizi
DUA	: Detrended Uyum Analizi

ŞEKİLLER LİSTESİ

Şekil 1. Topluluk verilerinde KUA ve RDA'nın kullanımı için karar ağacı.....	17
Şekil 2. Bir çevresel değişkene karşı üç cevap değişkeninin cevabının ya da beklenen değerinin unimodal eğrileri.....	19
Şekil 3. (a) Ordinasyon, (b) Regresyon ve (c) Kanonik analizin asimetric formları (RDA ve KUA) arasındaki ilişkiler.....	22
Şekil 4. Doğrudan algoritma kullanılarak Gereksizlik Analizi hesaplama adımları.....	26
Şekil 5. RDA biplot'u ve triplotunun şematik gösterimi.....	33
Şekil 6. Rda'DA (a) Tahmin edilen (fitted) nesnelere bir vektör üzerine iz düşümü (b) Vektörler arasındaki açıları vurgulayan şemalar.....	35
Şekil 7. KUA triplotunun şematik gösterimi.....	41
Şekil 8. KUA triplotunda Tip I ve Tip II ölçeklendirme.....	43
Şekil 9. KUA uygulamasının triplot şeması.....	50
Şekil 10. RDA uygulamasının triplot şeması.....	54

TABLULAR LİSTESİ

Tablo 1. Cevap değişkeninin cevap (tepki) çeşitlerine dayanan dolaylı ve doğrudan ordinasyon yöntemleri.....	5
Tablo 2. Ordinasyon ve regresyon arasındaki ilişki.....	21
Tablo 3. RDA ve KUA kullanılarak cevap değişkenleri matrisi ve açıklayıcı değişkenler matrisinin kanonik analizinden elde edilen özdeğerler ve bunlara karşılık gelen özvektörler.....	31
Tablo 4. Cevap değişkeni (site), nesne ve açıklayıcı değişkenlere göre KUA uygulaması.....	46
Tablo 5. Site, cevap değişkeni ve açıklayıcı değişkenlere göre RDA uygulaması	47
Tablo 6. Cevap değişkeni (site), nesne ve açıklayıcı değişkenlere göre KUA uygulama sonucu.....	48
Tablo 7. KUA uygulamasında permütasyon testi sonucu.....	49
Tablo 8. KUA uygulamasında temel eksenlerin özdeğerleri ve varyans açıklama oranları (%).....	49
Tablo 9. Site sayısı, cevap ve açıklayıcı değişkenlere göre RDA uygulama sonucu.....	52
Tablo 10. RDA uygulamasında permütasyon testi sonucu.....	53
Tablo 11. RDA uygulamasında temel eksenlerin özdeğer ve varyans açıklama oranları (%).....	53
Tablo 12. Kısıtlı ordinasyon yöntemleri için değerlendirme ölçütleri	56

1. GİRİŞ

Doğadaki olaylar, birbirleri ile ilişkili olan çok sayıdaki değişkenin veya faktörün ayrı ayrı veya birlikte etkileri ile ortaya çıkmaktadır. Bu ilişkileri doğru bir şekilde belirleyebilmek, olayların akışına istenilen şekilde yön verebilmek açısından oldukça önemlidir. Bu nedenle de uygun istatistik yöntemlerin kullanılması gerekmektedir.

Bilimsel çalışmalarda, genel olarak değişkenler arasında ilişki olup olmadığı ya da ilgilenilen değişken bakımından, uygulanan faktörün seviyeleri arasında fark olup olmadığı incelenmektedir. İlgilenilen değişken ile buna etki edebileceği düşünülen diğer değişken veya değişkenler arasındaki ilişki yapısı; doğrusal ya da doğrusal olmayan ilişkiler olmak üzere iki başlık altında incelenebilir. Doğrusal ilişkilerin belirlenmesinde kullanılan ilişki ölçüleri, değişken sayısına ve ölçek türüne göre farklılık göstermektedir. Değişkenler *sürekli olduğunda* Pearson korelasyon katsayısı (çarpım moment korelasyon) kullanılabilirken, *sıralı* olduğunda Sperman ya da Kendall korelasyon katsayıları, *kategorik* olduğunda ise phi katsayısı kullanılabilir.

Sürekli değişkenler normal dağılım gösterdiğinde, bunlar arasındaki doğrusal ilişkileri belirleme ve boyut indirgeme temeline dayalı, Temel Bileşenler Analizi (PCA), Faktör Analizi ve Çok Değişkenli Regresyon Analizi gibi çok değişkenli analiz yöntemleri ile kullanılabilir. Ancak, aralarındaki ilişki yapısının incelenmek istendiği değişkenler *kategorik* olduğu durumda, bu ilişki yapısını inceleyen ve görsel olarak (genellikle) iki boyutlu uzayda gösteren analiz yöntemi Çoklu Uyum Analizidir.

Bilimsel araştırmalarda yaygın problemlerden birisi de çok sayıda cevap değişkeni ile açıklayıcı (çevresel) değişkenler arasındaki ilişki yapısının yüksek doğrulukla belirlenmesidir. Herhangi bir cevap değişkeni (nesnelerin varlığı, tür kompozisyonu vb) açıklayıcı değişkenlere bağlı olarak değişebilmektedir. Dolayısıyla, farklı lokasyonlarda (yerleşim yerlerinde, sitelerde veya bölgelerde) cevap değişkenleri (hastalıkların, türlerin veya nesnelerin varlığı veya bulunma sıklığı) ile açıklayıcı (çevresel) değişkenler arası ilişkinin belirlenmesi oldukça önemlidir. Bu yapıdaki verilerin analiz edilmesinde; önceleri daha çok değişkenler arası ilişkilerin doğrusal olduğu varsayılmış ya da cevap değişkenleri için ayrı ayrı çoklu regresyon analizlerinin

yapılması tercih edilmiştir. Ancak daha sonra bu yaklaşımların yerine ilişkileri birlikte ele alan yeni yöntemler önerilmiştir. Regresyon ve ordinasyon yöntemleri, çok değişkenli doğrudan gradient analizi algoritmalarına eklenmiş ve bunlar kanonik (ya da kısıtlı) ordinasyon yöntemleri olarak adlandırılmıştır. Kanonik ordinasyon yöntemlerinden en yaygın olanları Gereksizlik Analizi (Redundancy Analysis, RDA) ve Kanonik Uyum Analizidir (Canonical Correspondence Analysis, CCA). Bu yöntemler dışındaki ordinasyon yöntemleri, dolaylı gradient analizi (Indirect gradient analysis) olarak bilinir.

Gereksizlik Analizi (RDA), birden fazla cevap (tepki, yanıt) değişkenindeki varyasyonu, birden fazla açıklayıcı değişken ile açıklamak ve özetlemek için kullanılan bir analiz yöntemidir. Diğer bir ifade ile RDA, bir dizi açıklayıcı değişkenle cevap değişkenlerinin bileşenleri arasındaki doğrusal ilişkileri özetleyen doğrudan bir gradient analiz yöntemidir.

Kanonik Uyum Analizi (KUA/CCA) bazı lokasyonlarda, cevap değişkeni olarak alınan; hastalıkların, nesnelerin veya türlerin varlığı ile bunlara etkili olabileceği düşünülen açıklayıcı değişkenler arası ilişkileri belirlemek için geliştirilen çok değişkenli analiz yöntemlerinden birisidir (Ter Braak, 1986).

Doğrudan Gradient Analizleri (DGA) ile ilgili yapılan çalışmaların, genellikle ekolojistler ve çevre bilimciler tarafından yapıldığı ve bu analizlerin sağlık alanında uygulamalarıyla ilgili sınırlı sayıda çalışmanın olduğu görülmüştür. Böylece tez çalışmasında, KUA ve RDA'nın incelenmesi amaçlanmıştır. Ayrıca farklı ölçek seviyelerindeki değişkenlere ait verileri içeren kardiyoloji alanındaki bir veri seti ile yöntemlerin uygulaması yapılarak performansları değerlendirilmiştir.

2. GENEL BİLGİLER

Bilimsel çalışmalarda, ele alınan özellikler genellikle çok sayıda faktörün etkisi altındadır. Bu nedenle geçerli ve güvenilir sonuçların elde edilebilmesi bakımından, değişkenlerin birlikte ele alınıp incelenmesi önemlidir. Bu kapsamda çok değişkenli analiz yöntemleri ön plana çıkmaktadır. Araştırmalarda ele alınan değişkenlerin; sürekli, kategorik ya da sıralı değişken olması durumuna göre ve açıklayıcı değişkenlerle doğrudan ilişkilendirilecek modellere göre uygun analiz yöntemleri bulunmaktadır. Bu yöntemler arasında; kısıtlı ordinasyon yöntemlerinden, Gereksizlik Analizi (RDA) ve Kanonik Uyum Analizi (KUA) sayılabilir.

2. 1. Sınıflandırma ve Ordinasyon

Çok değişkenli istatistik analizlerde, ordinasyon ve sınıflandırma (kümeleme) yöntemleri, topluluk ekolojisi üzerinde çalışan araştırmacıların sık tercih ettikleri yöntemler olup, bunlar kısmen birbirlerini tamamlayıcı niteliktedir. Ordinasyon yöntemleri, bir veri matrisinin daha düşük boyutta varyasyonunu göstermede ya da temsil etmede kullanılmaktadır (Palmer, 2017). Bu yöntemler, hesaplamalardan sonra dahil olan dışsal değişkenlerden etkilenmediği gibi, kendisini sınırlamadan veri matrisini açıklamaya da olanak tanır. Sınırlandırılmış (kısıtlı) ordinasyon (RDA ve KUA), cevap değişkenlerini içeren matris ve açıklayıcı değişkenleri içeren matrisle doğrudan ilişkilidir. Her iki matriste ordinasyonun aşamalarını gösterir. Bu yaklaşıma Doğrudan Gradient Analizi denir. Bu yaklaşımda, ordinasyon ve çoklu regresyon yöntemleri birleştirilir (Borcard, 2006).

2. 2. Ordinasyon Diyagramları

Ekolojide en yaygın kullanılan veri matrisleri, nesnelere tarafından indekslenen cevap değişkeninin bolluğu ya da öneminden oluşur. Gradientler boyunca cevap değişkeninin dağılımı hakkında bilgi edinmek için ordinasyon tekniklerinin kullanımı gittikçe artmaktadır (Hill ve Gauch, 1980).

Doğrusal (TBA ve RDA) ve ağırlıklı ortalama'ya (UA/DUA ve KUA) dayalı ordinasyon teknikleri, topluluk verilerini büyük ölçüde farklı şekillerde temsil

etmektedir. İki boyutlu ordinasyon şemasının oluşturulması ve yorumlamasının kolaylığı nedeniyle diyagramlar iki boyutta oluşturulur (Ter Braak ve Prentice, 1988).

İki boyutlu ordinasyon çözümde, cevap değişkeni verisinin iki değişkenli Gaussian cevap yüzeylerine uyması amaçlanmaktadır (Ter Braak, 1985b). Ancak, Ter Braak (1986), iki boyutlu çözümde, ilk iki eksenin puanları arasında yaklaşık dörtte bir bağımlılık olduğunda, kemer etkisinin oluşacağını belirtmiştir. Gauch (1982a), uzun bir gradient'in kısa bir gradient'e baskın olduğunu ve her zaman bu etkinin ortaya çıkabileceğini iddia etmiştir. Ter Braak (1986), Detrended (Eğilimsiz) Uyum Analizine olanak sağlayan bazı modifikasyonların, bu etkinin giderilmesi için kanonik uyum analizine dâhil edilebileceğini ifade etmiştir. Detrending işleminin, kemer etkisini ortadan kaldırdığını ve konum ya da lokasyon değerlerinin Gaussian modeline uyumunu artırdığını belirtmiştir.

Ekolojik verilerde, genel olarak çevresel ya da açıklayıcı değişkenlerle cevap değişkenleri doğrusal olmayan ilişkili olabilir. Bunun için iki adımlı bir yaklaşım önerilmiştir: Birinci adımda, cevap değişkeni verilerinden, toplum kompozisyonundaki varyasyon modeli, (detrended) Uyum Analizi gibi bir ordinasyon tekniği ile oluşturulur. İkinci adımda bu model, ilk birkaç koordinasyon eksenini kullanarak, çevresel değişkenlerle ilişkilendirilir (Gauch 1982a; Ter Braak, 1986).

Ter Braak (1986), birden fazla cevap değişkeni içeren çalışmalarda, her cevap değişkeni için ayrı ayrı yapılacak olan regresyon analizlerinin kullanışsız olabileceğini ifade etmiştir. Ayrıca, topluluk kompozisyonunun, çevresel değişkenlere göre nasıl değiştiğini (özellikle çevresel değişkenlerin sayısı iki ya da üçü aştığında) belirlemede, çok değişkenli bir istatistik analiz yönteminin (ortak bir yanıt modeline dayanan) gerekli olduğunu belirtmiştir.

Ter Braak ve Prentice (1988), Regresyon ve kalibrasyonun tersine, ordinasyon probleminin çok sayıda parametrenin eş zamanlı tahminini gerektirdiğini ve uygulamada bazı kısıtlamaları olduğunu belirtmişlerdir. Çıkartılacak ordinasyon eksenlerinin sayısının az olması gerektiğini ve bir çözüme olanak tanınması için yanıt modelinin kısıtlanması gerekliliğini ifade etmişlerdir. Ayrıca, ordinasyon tekniklerinin, basit modellerden sapmalara karşı sağlam (robust) olduğunu vurgulamışlardır.

2. 2. 1. Temel bileşenler diyagramı: Biplotlar

TBA, her bir cevap değişkeninin, koordinasyon eksenleri tarafından tanımlanan alandaki varlığını düzleme uyarlar. Noktaların ve cevap değişkenlerinin oklarla işaretlendiği şemaya "biplot" denir (Gabriel, 1971). Oklar, cevap değişkeninin varlığındaki maksimum varyasyon yönünü gösterir ve uzunluğu, bu maksimum değişim oranıyla orantılı olarak değişir. Ter Braak ve Prentice (1988), diyagramda orijinden uzaklaştıkça katkının veya ilişkinin arttığını, orijine yakın yerleşen değişkenlerin ise varyasyona katkısının daha az olduğunu belirtmişlerdir. Bunun yanısıra, genellikle diyagramın kenarlarında yerleşen cevap değişkenlerini göstermemenin daha uygun olacağını belirtmişlerdir.

2. 2. 2. Gradient analizleri (Ordinasyon teknikleri)

Gradient Analizleri, heterojen bir ortamda çalışan biyotik süreçlerin neden olduğu biyotik toplulukların yapısını incelemek için ordinasyon yöntemlerini kullanır.

Ter Braak ve Verdonschot (1995) gradientlerin, bir ordinasyon diyagramı yoluyla değişkenler arası ilişkileri basit ve görsel olarak sunmak için önemli olduğunu belirtmişlerdir.

Gauch (1982b); Jongman ve ark. (1995); Legendre ve Legendre (1998), TBA ve RDA'da değişkenler arasında doğrusal ilişkinin varsayıldığını ifade etmişler ve bu varsayımın kısa gradientlerde uygulanabilirliğini belirtmişlerdir (Tablo 1).

Tablo 1. Cevap değişkeninin cevap (tepki) çeşitlerine dayanan dolaylı ve doğrudan ordinasyon yöntemleri (Ter Braak, 1988).

Yöntem	Kısa Gradient (doğrusal cevap modeli)	Uzun Gradient (unimodal cevap modeli)
Dolaylı ordinasyon	TBA	CA
Doğrudan ordinasyon	RDA	CCA
Mesafe ölçütü	Öklit uzaklığı	Ki-kare uzaklığı

Hejemanova-Nezerkova ve Hejeman (2006), çevresel gradientin uzunluğunu tespit etmek için Detrended Uyum Analizi (DUA) kullanılabileceğini belirterek, DUA'da Monte Carlo permütasyon testinin kullanılma amacının, elde edilen açıklayıcı

değişkenlerin (çevresel değişkenler) bitki türlerinin (cevap değişkenlerinin) kompozisyonu üzerinde yaptığı etkiyi ortaya koymak olduğunu ifade etmişlerdir.

Dolaylı gradient analizleri

Dolaylı gradient analizlerinde çevresel gradientler doğrudan incelenmeyip bunlar, cevap değişkeni ile birlikte ortaya çıkan değişkenlerin verilerinden elde edilmeye çalışılır (Palmer, 1993). Bu analizlerde, gradientleri ya da veri yapısını açıklamak amacıyla, veri seti bir takım matematiksel işleme tabi tutulur. Ardından bu gradient ya da veri yapısı açıklanmak üzere, aynı örneklemelerin geçerli olduğu başka bir veri seti ile karşılaştırılır.

Gauch (1982a), Dolaylı gradient analizlerinde, özellikle ilgilenilen çevresel değişkenlerin bir alt kümesinin nesnelerin kompozisyonu üzerindeki etkilerini saptamanın anlamlı olmayacağını belirtmiştir. Bu çeşit problemlerin, ancak cevap değişkenlerinin oluşumlarını doğrudan çevresel değişkenlerle ilişkilendiren Doğrudan Gradient Analizi Yöntemleriyle aşılabileceğini savunmuştur.

Temel Bileşenler, Faktör ve Çoklu Regresyon Analizleri gibi ordinasyon yöntemleri, bir veri matrisinin varyasyonunu daha az sayıda boyutta temsil eder. Bu analizlerde yapıların yorumlanması için bir arka plan tasarlanır. Bu yaklaşım *dolaylı gradient analizi* olarak ifade edilir (Borcard, 2006).

Dolaylı gradient analizlerinde, açıklayıcı değişkenler matrisi (X), cevap değişkenlerinin hesaplanmasından bağımsızdır. Açıklayıcı değişkenlerin matrisi (X) üzerinde, ordinasyon vektörlerinin korelasyonu veya regresyonu (ön bilgi için) hesaplanır. Doğrudan gradient analizlerinde ise X matrisi, X'deki değişkenlerin kombinasyonlarıyla maksimum düzeyde ilişkili ordinasyon vektörlerinin sıkıştırmasıyla Y'nin ordinasyonunu üreten hesaplamaya müdahale eder (Legendre ve Legendre, 1998).

Ter Braak ve Prentice (1988), ordinasyon eksenlerinin, cevap değişkenlerinin bolluğunun gradientler boyunca nasıl değiştiğine dair belirli bir istatistik modele göre (doğrusal ya da unimodal) cevap değişkeni verilerinin uyumunu optimize edecek

şekilde inşa edilen gizli değişkenler ya da varsayımsal çevresel değişkenler olarak düşünülebileceğini iddia etmiştir.

Doğrudan gradient analizleri (Kısıtlı Ordinasyon)

Cevap değişkenlerinin, açıklayıcı değişkenlere nasıl cevap verdiğinin belirlenmesi için regresyon ve ordinasyon yöntemleri kullanılabilir. Bunun için çok değişkenli doğrudan gradient analizi yöntemlerine entegre edilmiş ve kanonik (kısıtlı) ordinasyon olarak adlandırılan analiz yöntemleri önerilmiştir. Kısıtlı ordinasyon (çok değişkenli doğrudan gradient analizi) aynı anda tüm cevap değişkenlerini analize dâhil eder (Legendre ve Legendre, 1998).

Doğrudan Gradient Analizinde (DGA) cevap değişkenleri, doğrudan doğruya ölçülebilen açıklayıcı değişkenlerle ilişkilidir. DGA, ölçülen bir gradient boyunca bir konum fonksiyonu olarak, cevap değişkenlerinin dağılımını gösteren bir grafik oluşturabilir. Ancak, ekolojik veriler tipik olarak birçok cevap değişkenine ve çoklu gradientlere sahip olduğundan, DGA boyut indirgeme yöntemiyle, yani ordinasyonla en iyi şekilde eşleştirilebilir (Palmer, 2017).

DGA bir regresyon analizi yöntemi olup, değişkenler arasında herhangi bir ilişki olup olmadığını belirler ve belirli açıklayıcı değişken kümesi tarafından açıklanabilecek olan nesnelerin ya da topluluk kompozisyonundaki varyasyonu incelemeye olanak tanır (Ter Braak ve Prentice, 1988).

2. 3. Gereksizlik Analizi (Redundancy Analysis/RDA)

Gereksizlik (artıklık) Analizi (RDA), iki değişken kümesi arasındaki ilişkileri analiz etmek için kullanılan çok değişkenli istatistik analiz yöntemlerinden birisidir. Israels (1992), Y değişken kümesi için gereksizlik analizi hesaplamasını, Y kümesindeki değişkenler için bir dizi eşzamanlı regresyon eşitliklerinin hesaplanması süreci olarak kabul etmiştir. Ayrıca, bu analizin; açıklayıcı değişkenlerin (X) sadece doğrusal kombinasyon fonksiyonlarını kullandığını ifade etmiş ve bu nedenle yöntemin, azaltılmış sıralama regresyonu olarak da değerlendirilebileceğine dikkat çekmiştir.

Gittins (1985), gereksizlik kavramının açıklanan varyans ile eşanlımlı olduğunu iddia etmiştir.

RDA'da Y veri tablosu, cevap değişkenlerini içerirken, X veri tablosu ise açıklayıcı değişkenleri göstermektedir (Makarenkov ve Legendre, 2002).

Legendre ve Legendre (1998); Makarenkov ve Legendre (2002), RDA'nın, X ve Y değişkenleri arasındaki doğrusal bir ilişki modelini kullandığından, çok değişkenli cevap (tepki) verilerinin modellenmesi için çoklu doğrusal regresyonun bir uzantısı olarak da görülebileceğini belirtmişlerdir. Ayrıca, RDA'nın tahmin edilen (fitted) değerlerin bir özdeğer ayrışımından oluştuğunu göstermişlerdir. Benzer şekilde, Makarenkov ve Legendre (2002), RDA'nın veri noktalarının dağınıklıklarındaki eğilimleri en yüksek düzeyde gösteren Temel Bileşenler Analizi'nin (TBA) kısıtlı bir uzantısı olarak da düşünülebileceğini belirtmişlerdir.

Gereksizlik Analizi (RDA), açıklayıcı değişken kümesinden yararlanarak, cevap (yanıt) değişkeni kümesindeki varyasyonu açıklamak ve özetlemek için kullanılan bir yöntemdir. Diğer bir ifade ile RDA, açıklayıcı değişken kümesi ile "gereksiz (artık)" olan cevap değişkenlerinin bileşenleri arasındaki doğrusal ilişkileri özetleyen bir doğrudan gradient analizi yöntemidir. RDA, çoklu açıklayıcı değişkenlerle çoklu cevap değişkenlerinin regresyonuna izin vererek Çoklu doğrusal regresyonu genişletir. Çoklu doğrusal regresyon yoluyla üretilen tüm cevap değişkenlerinin düzenlenmiş değerlerinin bir matrisi daha sonra Temel Bileşenler Analizine (TBA) tabi tutulur.

Borcard (2006), RDA'nın Temel Bileşenler Analizinin (TBA) kısıtlı bir formu olarak da düşünülebileceğini ifade etmiştir. Ayrıca, Gereksizlik (artıklık) Analizinde, cevap değişkenlerinin doğrusal kombinasyonlarından oluşturulan kanonik eksenlerin, açıklayıcı değişkenlerin doğrusal kombinasyonları olması gerektiğine dikkat çekmiştir.

Legendre ve Legendre (1998), RDA analizinde X'deki değişkenlerin doğrusal kombinasyonları olan ordinasyon vektörleri ile Y'nin ordinasyonunun sınırlandırıldığını vurgulamışlardır.

Wollenberg (1977), RDA ile Kanonik Korelasyon Analizi (CCorA) arasındaki ilişkinin, basit doğrusal regresyon ile doğrusal korelasyon analizi arasındaki ilişki ile

aynı olduğunu göstermiştir. Ayrıca, Gittins (1985) ve Wollenberg (1977), CCorA'dan farklı olarak, RDA'nın asimetrik bir analiz olduğunu ifade etmişlerdir.

Legendre ve Legendre (1998), genelde bir kanonik korelasyon katsayısı verildiğinde, Y değişkenlerindeki fazlalığın X değişkenlerinden farklı olacağını belirtmişlerdir. Bunun yanında RDA, veri noktalarının dağılımlarındaki eğilimleri en yüksek düzeyde gösteren Temel Bileşen Analizi'nin (TBA) kısıtlı bir uzantısı olarak da düşünülebileceğini ifade etmişlerdir. Ayrıca, kanonik ordinasyon vektörlerinin; Y değişkenlerinin doğrusal kombinasyonları olduğunu belirtmişler ve bir dizi kısıtlayıcı (açıklayıcı) değişkenle doğrusal ilişkili olduğunu göstermişlerdir. Ayrıca, RDA'da cevap değişkeninin varlığına ya da frekansına ki-kare dönüşümü yapılır, ancak dönüştürülmüş cevap verisi (Y) ile açıklayıcı değişkenler (X) arasındaki ilişkinin doğrusal olduğu varsayılmaktadır (Makarenkov ve Legendre, 2002).

Ter Braak (1986) RDA'nın, tüm değişken çeşitlerini içeren açıklayıcı (X) değişkenle cevap değişkenlerinin eşzamanlı analizini iki ya da üç boyutlu bir uzayda sunduğunu belirtmiştir. Bunun yanında, RDA'da kullanılan değişken kümelerinin, eş zamanlı analiz yaklaşımının bir kovaryans kriteri için uygun olduğuna dikkat çekmiştir. Ayrıca, "Skor" setlerinin, RDA sonucunun tipik bir özelliği olduğunu ve kullanılan ölçeklemeye bağlı olarak değişebileceğini ifade etmiştir.

Ter Braak (1986), nesne ve cevap değişkeni puanlarının, genellikle "site" ve "tür" puanları olarak rapor edildiğini ifade etmiştir. Elde edilen bu puanlar, noktaları ve vektörleri düzenlemek için kullanılan koordinatlardır. Vektörün yönü, ilgili değişken için artış yönünü göstermektedir. Ayrıca, açıklayıcı değişken puanlarının, kısıtlayıcı değişken puanları olarak da ifade edilebileceğini belirtmiştir.

Makarenkov ve Legendre (2002), RDA'nın temel yaklaşım mantığı olan, cevap değişkenlerinin değişimlerini açıklayıcı değişkenlerdeki değişimlere göre doğrusal olarak ilişkilendirmesinin özel bir nedeninin olmadığını iddia etmişlerdir. Zira, ekolojik süreçler modellendiğinde, doğrusallık varsayımının çoğu durumda gerçekçi olmadığını ve sadece daha uygun analiz yöntemleri mevcut olmadığı için bu yaklaşımın uygulandığını düşünmüşlerdir. Bununla beraber, X ve Y'deki değişkenler arasındaki ilişkiyi modellemede doğrusallık varsayımını ortadan kaldırmak için polinom

regresyona dayanan iki ampirik kanonik analiz yöntemini önermişlerdir. Polinom RDA ve polinom CCA olan bu yöntemlerin, klasik RDA ve CCA'ya alternatif olabileceğini savunmuşlardır.

Van der Burg ve De Leeuw (1983), CCorA'da ayırık verilerin optimal doğrusal olmayan dönüşümlerini bulmak için alternatif en küçük kareler yöntemini kullanmışlardır. Durand (1993) ve Donovan (1998), RDA ve CCA'da doğrusal olmayan değişkenleri ifade etmek için spline dönüşümleri kullanmışlar ve elde edilen dönüşümlerin şekillerinin genellikle yorumlanamayacağını belirtmişlerdir. Benzer şekilde Makarenkov ve Legendre (2002), bilinmeyen fonksiyonel formların doğrusal olmayan ilişkilerini tahmin etmek için polinom regresyonların zarif ve kolay bir yol sunduğunu savunmuşlardır.

Makarenkov ve Legendre (2002), polinom denklemlerinden elde edilen terimlerin, cevap verisi üzerinde önemli etkiye sahip değişkenlerin kombinasyonlarını temsil ettiğini belirtmişler ve ikinci derece önemli terimlerin, açıklayıcı ve cevap değişkenleri arasındaki doğrusal olmayan ilişkileri temsil ettiğini iddia etmişlerdir. Ayrıca, Y veri tablosu büyük veri içerdiğinde, RDA eksenlerinin cevap değişkenlerinin miktarlarındaki farklılıklar açısından yorumlanabileceğini iddia etmişlerdir. Böylece RDA biplotlarındaki bileşen eksenlerinin, açıklayıcı değişkenler tarafından sınırlanan cevap değişkeni miktarlarındaki gradientleri temsil ettiğini savunmuşlardır.

Borcard (2006), kısıtlı bir ordinasyonun açıklayıcı değişkenlerin yanında çok sayıda kanonik eksen ürettiğini ve bu eksenlerin her birinin, tüm açıklayıcı değişkenlerin doğrusal bir kombinasyonu (çoklu regresyon modeli) olduğunu belirtmiştir. Ayrıca, her bir eksen üzerindeki açıklayıcı değişkenlerin kanonik katsayılarının (yani, modellerin regresyon katsayılarının) incelenerek; hangi değişkenin (değişkenlerin) en önemli olduğu hakkında bilgi sahibi olunacağını ifade etmiştir.

Borcard (2006), açıklayıcı değişkenler tarafından açıklanamayan veri matrisinin varyasyonunun, bir dizi kısıtlanmamış eksenle ifade edildiğini belirtmiştir. Birçok durumda, açıklayıcı değişkenler boyutsal olarak homojen olmadığından, genellikle standardize edilmiş açıklayıcı değişkenler ile kanonik düzenlemeler yapıldığını iddia etmiştir. Ayrıca RDA'da, kovaryans veya korelasyon matrisi kullanılarak yapılan

analizler arasında bir farkın olmadığını ve bu seçimin cevap (Y) değişkenleri ile ilgili olduğunu ifade etmiştir.

2. 3. 1. Gereksizlik analizi (RDA) ile kanonik korelasyon analizi (CCorA) arasındaki ilişki

Doğrusal Korelasyon Analizi ile Doğrusal (lineer) Regresyon Analizi arasındaki ilişkiye benzer şekilde, Kanonik Korelasyon Analizi (CCorA) ile Gereksizlik Analizi (RDA) arasında benzer bir ilişki yapısı vardır. CCorA'da, incelenen iki matris simetrik bir şekilde çözümlenirken, RDA'da Y matrisinin açıklayıcı bir X matrisine bağımlı olduğu düşünülmektedir. RDA ve CCA, iki matrisli problemlerin çözümünde, CCorA'dan daha fazla uyum gösterir (Ter Braak, 1988).

CCorA'da kanonik eksenler, Y ve X değişkenlerinin iki setinin doğrusal kombinasyonları arasındaki korelasyonu en üst düzeye çıkarır. Bunun için değişkenler içi grup kovaryansı (ya da korelasyon) yerine, değişkenler arası grup kovaryansı (ya da korelasyonunu) maksimize edilir. Ancak RDA'da, X değişken kümesinden elde edilen doğrusal fonksiyonlarla Y değişken kümesindeki varyasyon maksimum düzeyde açıklanmaya çalışılmaktadır. Kanonik korelasyon analizinde, elde edilen korelasyon katsayıları, değişken kümelerine ait varyasyon hakkında bilgi vermemektedir. Bu durum, Kanonik korelasyon analizi için bir dezavantaj oluşturmaktadır. Bunun giderilebilmesi için gereksizliği belirleme indeksi geliştirilmiştir. Gereksizlik analizi, literatürde gereksizliği belirleme indeksi olarak da bilinmektedir. Bu indeks değeri, değişken kümesindeki varyansın, diğer değişken kümesi tarafından ne kadarlık kısmının açıklandığını belirlemek için kullanılmaktadır (Kalkan ve Özden, 2017). Böylece her bir kanonik bileşenin varyansa olan katkısı elde edilebilmektedir. 1977'de Wallenberg gereksizliği belirleme indeksini maksimize edecek yeni bileşenler oluşturarak, Gereksizliği belirleme analizini türetmiş ve Kanonik Korelasyon Analizine alternatif olarak göstermiştir (Ural, 1992).

2. 3. 2. Gereksizlik analizi (RDA) ile temel bileşenler analizi (PCA) arasındaki ilişki

Temel Bileşenler Analizinde, değişken kümesindeki varyasyona olan katkılarına göre elde edilen bileşenler, boyut sayısını göstermektedir. Böylece birbirinden bağımsız bileşenler yardımı ile ilgilenilen değişken kümesindeki değişken sayısı azaltılabilmektedir. Gereksizlik Analizinde ise oluşturulan bileşenler yardımı ile gereksizliği belirleme indeksi düzenlendiğinden, bu indeksin maksimum değer aldığı bileşenler boyut sayısını göstermektedir. Temel bileşenler ile Gereksizlik analizinin karşılaştırmasının yapıldığı bazı çalışmalarda, Gereksizlik analizi sonucunda daha az sayıda boyut elde edildiği görülmüştür. Bu durumun, gereksizliği belirleme indeks değerinden kaynaklandığı düşünülmekte ve analizde elde edilen bileşenlerin varyasyona olan katkısının daha az sayıda bileşen üzerinde yoğunlaşmasından kaynaklandığı düşünülmektedir (Van der Burg ve De Leeuw, 1983).

2. 4. Kanonik Uyum Analizi/KUA (Canonical Correspondence Analysis/CCA)

Kanonik Uyum Analizi (KUA), ilk olarak Ter Braak (1986) tarafından Uyum Analizinin genişletilmiş hali olarak ortaya atılmış ve yöntem ekolojik verilerde kullanılmıştır. Bölgelerde cevap değişkenlerinin bulunma sıklığı ile çevresel değişkenler arası ilişkiyi incelemek üzere geliştirilmiştir.

Cevap eğrisi modellemesi bağlamında, ağırlıklandırılmış ortalama temel alınarak cevap değişkenlerinin cevap eğrilerinin şeklini belirlemek ve hangi koşullar altında, ağırlıklandırılmış ortalamanın maksimum olasılıkla yakınsayacağı sorusu Ter Braak'ın KUA'yı keşfetmesine neden olmuştur (Ter Braak ve Looman, 1986).

Ekolojik verilerin analizinde yoğun olarak başvurulan Kanonik Uyum Analizi, bir Doğrudan Gradient Analizi Yöntemi olarak diğer atama yöntemlerinin tersine sadece nesne kompozisyonunu temel almaz. Nesne kompozisyonunun yanında çevresel değişkenlerin de analize dâhil edildiği bir atama yöntemi olan KUA, Kısıtlanmış Uyum Analizi olarak da bilinir. Cevap değişkeni ile çevresel ya da açıklayıcı değişkenler arası ilişki konusunda oldukça açıklayıcı bir güce sahiptir (Oksanen, 2004).

Yöntem; cevap değişkeni olarak alınabilen hastalık veya bitki çeşitleri, ürün markaları, meslek tercihleri veya hayvan ırkları ile konumlar (lokasyon site, bölge) arasındaki ilişkileri incelemede de kullanılabilir.

Birçok durumda gözlemsel çalışmalar, değişken kümeleri arasındaki farklı ilişki yapılarını değerlendirmeyi amaçlamaktadır. Bunlar, doğrudan etki eden çoklu bağımlı değişkenlerin tespiti (hastalık oluşumları gibi) olabileceği gibi ilk grubun değişkenlerini etkilemesi beklenen diğer değişkenlerin bulunması da olabilir.

Değişken kümeleri arasındaki ilişkiler genelde asimetriktir ve bu iki değişken kümesini analiz etmede iki temel hedef gözetilir: Birincisi, hangi bağımlı değişkenin/lerin ya da bunların kombinasyonlarının en çok olduğunu belirlemektir. İkincisi ise açıklayıcı değişkenleri ve kombinasyonlarını en fazla etkileyen sonuç değişkenlerini tahmin etmektir. Çok Değişkenli Doğrusal Yöntemler, (çok değişkenli regresyon ya da çok değişkenli varyans analizi (MANOVA) gibi yöntemler) ancak ikinci amacı elde edebilmektedir. Diğer yandan açıklayıcı ve sonuç değişken grupları arasındaki asimetrik ilişkiyi hesaba katmayan, çok değişkenli yöntemler (kanonik korelasyon analizi gibi) ikinci amaca ulaşamaz (Mardia ve ark., 1979). Ancak, yukarıda anlatılan iki hedefi başaran yöntemlerden biri Kanonik Uyum Analizi'dir. (Ter Braak, 1986; Saporta, 1990). KUA, Çok Değişkenli Regresyon, MANOVA ve TBA yöntemlerini birleştirir.

Uyum Analizi (UA) 'nin bir uzantısı olarak Ter Braak (1986, 1987) tarafından geliştirilen Kanonik Uyum Analizi (KUA), cevap değişkenlerinin çevresel gradientlere unimodal cevaplarını tahmin eder, ancak yine de X ve Y grubu değişkenler arasındaki ilişkilerin doğrusal olduğunu varsayar. KUA'dan kaynaklanan bileşen eksenleri temel olarak açıklayıcı değişkenler tarafından sınırlanan gradientleri temsil eder (Makarenkov ve Legendre, 2002).

Ter Braak (1986) KUA'da, koordinasyon eksenlerinin, harici bilgi yardımıyla ya da çevresel değişken verileriyle yorumlandığını belirtmiştir. Ayrıca, Kanonik Uyum Analizinde koordinasyon eksenlerinin, açıklayıcı değişkenlerin doğrusal kombinasyonları olacak şekilde (açıklayıcı değişkenlerin kısıtlaması altında) seçileceğini, bu yolla nesnelere varlığının, doğrudan açıklayıcı değişkenlerdeki değişim

ile ilişkilendirilebileceğini ve açıklayıcı değişkenlerin sürekli ya da kategorik olabileceğini ifade etmiştir. Bunun yanısıra, KUA'nın, nesne kompozisyonunun çevreyle nasıl değiştiği konusunda hızlı bir değerlendirme yapabildiğini ve bu yöntemin, düzenli ordinasyon yönlerini doğrudan gradient analizinin yönleriyle birleştirdiğini belirtmiştir. Diğer yandan, KUA'nın mantığının, cevap değişkenlerinin bileşik çevresel gradientlere göre Gaussian (çan eğrisi şeklinde) cevap yüzeylerine sahip olduğunu varsaymıştır.

Kanonik Uyum Analizi koordinasyon diyagramı, cevap değişkenleri (türler) veya açıklayıcı (çevresel) değişkenler arasındaki yüksek korelasyonlardan etkilenmemektedir. Kanonik Uyum Analizi, cevap değişkenlerinin toplam bolluğu ile doğrusal orantılı, çevresel değişkenlerin kovaryans matrisi ile ters orantılı olarak ağırlıklandırıldığı bir analiz yöntemidir (Ter Braak, 1986).

KUA ile açıklayıcı değişkenlerin temelinde cevap değişkeninin oransal çoğalma eğilimleri bileşen eksenlerinde gösterilmeye çalışılır. İlişkilerin doğrusallık varsayımı için herhangi bir özel neden olmamakla birlikte çevresel değişkenlerdeki değişiklikler eksenlerde birleştirilmeye çalışılır. Ekolojik süreçteki modellemede doğrusallık varsayımı birçok örnekte gerçekçi olmamakla birlikte, daha uygun yöntemin bulunmaması nedeniyle bu şekilde varsayılmaktadır (Anonymous 1, 2018).

"Kanonik" terimi, "bir kurala göre" ya da "en temel biçimine indirgenmiş" anlamına gelir. Bu nedenle KUA, değişken grupları arasında bir ilişki ararken belirli bir ilkeyi izleyen bir UA olarak tanımlanmıştır. Genellikle, KUA uygulamasında iki matrisin belirtilmesi gerekir. Bunlardan birisi gözlemleri (incelenen her bir lokasyonda cevap değişkenlerinin varlığı) içeren matris, diğeri ise gradientleri (incelenen her bir lokasyonda bulunan mineraller gibi çevresel veya açıklayıcı değişkenleri) içeren matristir (Anonymous 1, 2018).

Gözlemler, doğrusal olarak gradientlerle ilişkilirse, KUA yerine bir Kanonik Korelasyon Analizi ya da Gereksizlik Analizi (RDA) kullanılması gerekmektedir. Kanonik Uyum Analizi (KUA), Uyum Analizinin (UA) kanonik şeklidir. Açıklayıcı değişkenler matrisinin UA çözümünün hesaplanmasına müdahale ettiği, Doğrudan

Gradient Analizinin bir formudur. KUA sonucunda, sadece açıklayıcı değişkenlerin matrisi tarafından açıklanabilen uyum yapısı gösterilmiş olur (Anonymous 1, 2018).

Ter Braak (1986), detrending yönteminin, kemer etkisinin giderilmesi için kullanılabileceğini belirtmiştir. Detrending KUA'nın, cevap değişkenlerinin çevresel gradientlere göre çan şeklinde cevap eğrilerine sahip olması durumunda etkili bir ordinasyon yöntemi olduğunu ifade etmiştir. Bu nedenle KUA, topluluk kompozisyonu ve çevresel değişkenlerle ilgili verilerin analizinde kanonik korelasyon analizine göre daha uygun bir yöntemdir. Yöntemde, noktaların cevap değişkeni (tür) ve siteleri temsil ettiğini ve vektörlerin çevresel değişkenleri temsil ettiği bir ordinasyon diyagramına yol açtığını saptamıştır. Böyle bir diyagramın, çevresel değişkenler tarafından en iyi açıklanabilen ve aynı zamanda çevresel değişkenlerin her biri boyunca cevap değişkeni dağılımlarının yaklaşık merkezlerini temsil eden varyasyon modelini gösterdiğini belirtmiştir.

Ter Braak (1987), KUA'nın, cevap değişkeni (türlerin) dağılımlarının maksimum düzeyde ayrıldığı çevresel değişkenlerin doğrusal kombinasyonlarını oluşturduğunu ifade etmiştir. Bunun yanında, KUA tarafından üretilen özdeğerlerin bu ayrımı ölçtüğünü tespit etmiştir. Uygulamalarda KUA'nın, cevap değişkeninin çevreyle olan ilişkilerini tespit etmede ve cevap değişkeninin çevresel değişkenlere cevap (tepki) modeli ile ilgili özel soruları araştırmak için kullanılabileceğini göstermiştir. Yer ve cevap değişkeni noktalarının, Uyum Analizindeki ile aynı yorumu taşıdığını ifade etmiştir. Cevap değişkeninin çevresel değişkenlere karşı unimodal cevap fonksiyonlarına sahip olma eğiliminde olduğunda KUA'nın daha iyi performans gösterdiğini vurgulamıştır. Ayrıca bu durumda, Kanonik Korelasyon Analizi ve Gereksizlik Analizi (RDA)'nin daha başarısız olduğuna dikkat çekmiştir.

Ter Braak (1988a), KUA 'da bağımsız değişkenlerin doğrusal kombinasyonları ile doğrudan ilişkili olan ve açıklanan varyans sırasına göre hesaplanan koordinat eksenlerinin hesaplanması işlemiyle ilgili bir regresyon modelinin yerleştirildiğini göstermiştir. Çoklu Regresyon Analizi ile ordinasyon puanlarının kısıtlandığını belirtmiştir.

Uyum Analizi, KUA gibi cevap deęişkeni sayıları ile örnek puanları arasındaki korelasyonu maksimum yapmaya çalışır. Ancak, KUA'da örneklem puanları açıklayıcı deęişkenlerin doğrusal kombinasyonları olarak sınırlandırılmıştır. Kısıtlanma nedeniyle KUA 'daki özdeęerler, UA'dan daha düşük olacaktır. KUA 'da maksimum yapma, “cevap deęişkeni sayılarının en iyi dağılımını bulma” olarak açıklanabilir (Ter Braak, 1988). Ter Braak (1988), cevap deęişkenlerinin çevresel deęişkenlerin doğrusal kombinasyonlarına göre tek yönlü cevap yüzeylerine sahip olduęu varsayımından hareketle KUA'da cevap deęişkenlerinin noktalarla (iki boyutlu alt uzayda yaklaşık optimumlarına karşılık gelen) ve çevresel deęişkenlerin yönleri ile deęişim oranlarını alt uzayda gösteren oklarla temsil edildiğini ifade etmiştir.

Ter Braak (1988a) KUA ile aralarında doğrusal ya da monotonik bir ilişki olduğunu kabul eden ordinasyon yöntemlerinin aksine, bağımlı ve bağımsız deęişkenler arasında tek modlu (unimodal) bir ilişki olduğunu göstermiştir.

Israels (1984) KUA'yı, enstrümantal deęişkenlere (açıklanan deęişkenlere) sahip Faktöriyel Uyum Analizi (FUA) olarak ifade etmiştir. Sabatier ve ark. (1989), bu yöntemin Kısıtlama Analizi olarak da bilindiğini belirtmişlerdir.

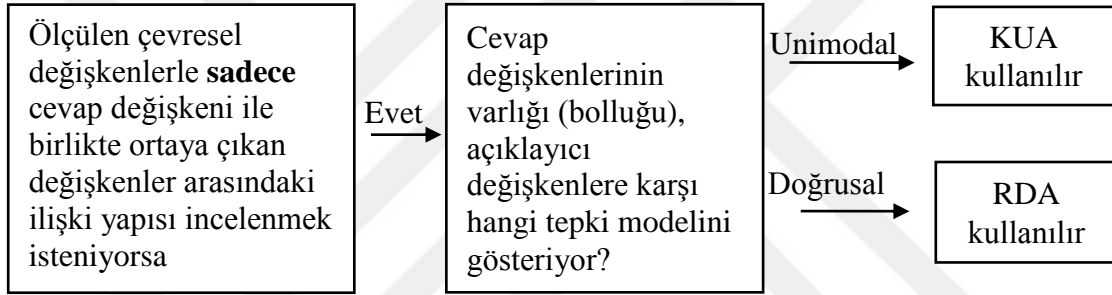
Angers ve ark. (1999), KUA 'da uygulanan testlerin, permütasyon yöntemlerini temel aldığından, veri dağılımı ile ilgili özel bir varsayımının olmadığını ifade etmişlerdir.

Graffelman (2001), alternatif bir yaklaşım olarak KUA'nın, çevresel deęişkenlere göre cevap deęişkenlerinin ağırlıklı ortalamaları için ağırlıklı en küçük kareler yaklaşımı olarak görülebildiğini ifade etmiştir. Ancak Hill (1974), KUA'nın karşılıklı ortalama olarak deęerlendirilebileceğini ifade etmiştir. Ter Braak (1987), Uyum Analizinin karşılıklı ortalama algoritmasını, çevresel deęişkenler üzerinde site puanlarının regresyonu ile genişleterek karşılıklı ortalamaları elde etmiştir. Daha sonra KUA'yı ağırlıklandırılmış ortalamalar matrisinin Ağırlıklı Temel Bileşenler Analizi olarak formüle etmiştir.

Graffelman (2001), çevresel deęişkenlere göre cevap deęişkenlerinin ağırlıklandırılmış ortalamalarının, genellikle optimum deęerlerini tahmin etmek için

KUA'nın kullanıldığını savunmuştur. Ayrıca, KUA'da optimum değerlerin görüntü kalitesinin, tüm özdeğerlerin toplamına göre en büyük iki özdeğer oranının ele alınarak rapor edildiğini belirtmiştir. Bunun yanında, KUA'da elde edilen özdeğerlerin sadece bolluk matrisinin görüntü kalitesini gösterdiğini, ancak cevap değişkenlerinin optimum değerini göstermediğini iddia etmiştir.

KUA'nın uygulanmasının uygun olup olmadığına karar vermede; ölçülen çevresel değişkenlerle sadece cevap değişkeni ile birlikte ortaya çıkan değişkenler arasındaki ilişki yapısının veya çevrenin makul koşullarında türlerin cevaplarının (tepkilerinin) unimodal olup olmamasının sorgulanması gerekmektedir (Şekil 1) (Mc Cune ve Grace, 2002).



Şekil 1. Topluluk verilerinde (cevap değişkeni ile birlikte ortaya çıkan değişkenler ile açıklayıcı değişkenlere ait verilerde) KUA ve RDA'nın kullanımı için karar ağacı (Mc Cune ve Grace, 2002).

Mc Cune ve Grace (2002), KUA'nın kullanımı için uygun olan temel soruların, Dolaylı Ordinasyon Yöntemlerinden farklı olduğunu ve KUA'nın topluluk veri kümelerine en uygun olan yöntem olduğunu belirtmişlerdir. Bu temel sorunların; çevreye karşı cevap değişkenlerinin cevabının (tepkisinin) unimodal (çan şeklinde) olması ve altta yatan (gizli) önemli çevresel değişkenlerin ölçülmesi olduğunu belirtmişlerdir. Birinci sorunun, doğrusal cevap eğrilerini (RDA) kabul eden yöntemlerde sorunlara neden olduğunu ancak KUA için sorun oluşturmayacağını vurgulamışlardır. İkinci sorunun ise diğer ordinasyon yöntemlerinin aksine (kanonik korelasyon ve gereksizlik analizleri hariç) sınırlandırılmış ordinasyon sonuçlarının çevresel matristen elde edileceğini belirtmişlerdir.

KUA'da her diyagramdaki iki koordinasyon eksenini, çevresel değişkenlerin doğrusal kombinasyonlarını temsil eder. Çevresel verilerin en yüksek varyasyonunu

açıklayan iki dik eksen, (Eksen, I Eksen II'den daha fazlasını açıklayacak şekilde) inşa edilir. Elde edilen diyagramda, veriler arasındaki ilişkileri tanımlamak için üç öge kullanılır. Bunlar; örnek noktalar, cevap değişkenlerinin noktaları ve çevresel vektörlerdir. Örnek bir noktanın diyagramdaki konumu, hem türlerin toplam kompozisyonu (yani, numunede toplanan her bir türün sayısı) hem de örnekleme sırasındaki çevresel koşullarla belirlenir. İki nesne, tür kompozisyonu ya da çevresel şartlar açısından farklılık gösteriyorsa, bunlar ordınasyon diyagramında farklı yerlerde görünürken, benzer olması durumunda birbirlerine yakın görünecektir (O'Connell ve ark., 2004).

Wartenberg ve ark. (1987), KUA'nın, Temel Bileşenler Analizinde kullanılan, korelasyon (ya da kovaryans) yerine ki-kare uzaklık matrisini kullandığını ifade etmiştir. KUA'nın, ikinci ordınasyon eksenini sonlandırmak ve ikinci boyuta göre kemer ya da yay etkisini ortadan kaldırmak için aynı algoritmayı kullanan Detrended (eğilimsiz) Kanonik Uyum Analizi olduğunu belirtmiştir.

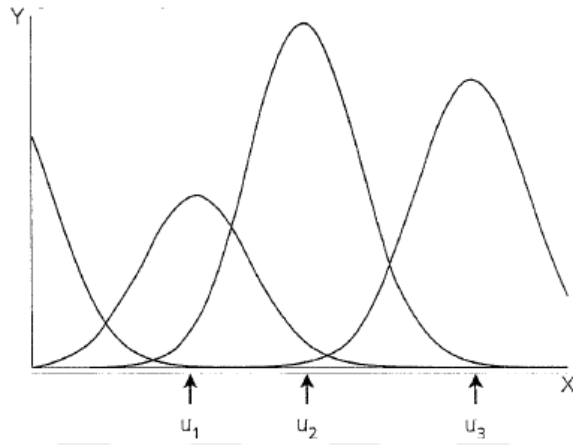
2. 5. Ağırlıklı Ortalama Ordınasyonu

Değişkenlerin koordinasyon şemasında zorunlu olarak ortogonal yönler olarak gösterilmesi gerekli olmamasına rağmen KUA, iki nicel çevresel değişkenle, bu değişkenlere göre cevap değişkeninin ağırlıklı ortalama alma ordınasyonunu kullanarak iki boyutlu bir diyagramda değişkenler arasındaki ilişki yapısı hakkında bilgi verir (Ter Braak, 1986).

Ağırlıklı ortalama ordınasyonunun avantajı; sitelerin ve türlerin eş zamanlı düzenlenmesidir (Escoufier, 1987). Ayrıca; türler, TBA için sorunlar oluşturabilen çevresel gradientlerle doğrusal olmayan ve unimodal ilişkiler içerisinde olduğunda, ağırlıklı ortalama daha hızlı hesaplama (NMDS'ye göre) yapar ve performansı yüksektir (Gauch, 1982a; Ter Braak ve Looman, 1986; Ter Braak, 1987; Ter Braak ve Prentice, 1988).

Cevap değişkenlerinin cevap eğrisi, Şekil (2)' deki gibi simetrik ise türün merkezi yada ağırlıklandırılmış ortalaması türün optimum tahmin değeridir. İlk yapay gradient, ilk koordinasyon eksenini olarak adlandırılır. Koordinasyon ekseninin özdeğeri

ile cevap deęişkeninin yoğunlaştığı bölge (niş) ayırımının maksimum deęeri elde edilir. Ayrıca, sonraki koordinasyon eksenleri nişleri maksimum olarak ayıran çevresel deęişkenlerin doğrusal kombinasyonlarıdır ve daha önce oluşturulan eksen ya da eksenler ile ilişkisiz olması gerektięi sınırlamasına tabidir. Çevresel deęişkenler için birçok koordinasyon eksen çıkarıldığından, artan eksen sayısı ile niş ayırımı (özdeęer) azalır. Bu nedenle prensip olarak sadece ilk birkaç eksenin incelenmesi yeterli olacaktır (Ter Braak ve Verdonschot, 1995).



Şekil 2. Bir çevresel deęişkene (x) karşı üç cevap deęişkeninin (türün) cevabının ya da beklenen deęerinin unimodal eęrileri. (u_k)[$k=1, 2, 3$] ile gösterilen üç türün aęırlıklandırılmış ortalamaları ile optimum deęerleri tahmin edilmiştir. Türlerin optimum deęeri, örneklem aralığının dışında, aęırlıklı ortalamalar ise aralığın içindedir. Çizilen eęriler simetriktir, ancak bu KUA için zorunlu deęildir (Ter Braak ve Verdonschot, 1995).

Karşılıklı ortalama algoritmasında, ilk eksen, site ve tür puanları arasındaki muhtemel maksimum korelasyonu üreten türlerin ve sitelerin sıralamasından oluşur (Gauch, 1982a). İkinci ve daha yüksek eksenler, eksenlerin ortogonal olduęu kısıtlamaya tabi olan maksimum alan-tür korelasyonuna sahiptir. Her bir eksenle ilişkili özdeęerler ise tür puanları ile site puanları arasındaki korelasyon katsayısına eşittir (Gauch, 1982a; Pielou, 1984). Bu nedenle, 1'e yakın özdeęer, türler ve siteler arasında yüksek derecede uyumu gösterirken, 0'a yakın özdeęer uyumun az olduęunu ifade eder.

3. GEREÇ VE YÖNTEM

3. 1. Gereç

Çalışmada uygulama materyali olarak, serbest erişimli “<https://archive.ics.uci.edu/ml/datasets.html>” (Erişim tarihi: 05/09/2018) sitesinden elde edilen veriler kullanılmıştır.

Kardiyoloji alanı ile ilgili olan veri seti, 270 birey ve farklı değişken tipinde (kategorik, nümerik ve ordinal) 13 adet özelliği (değişkeni) içermektedir. Veri setinde yer alan değişkenlerin özellikleri aşağıda özetlenmiştir.

1. Yaş
2. Cinsiyet
3. Göğüsteki ağrı türü (1: Tipik angina 2: Atipik angina 3: Non-anginal pain 4: Asimptomatik)
4. Dinlenme sırasında kan basıncı
5. Serum kolesterol miktarı (mg/dl)
6. Açlık kan şekeri düzeyi > 120 mg/dl (diyabet var, diyabet yok)
7. Dinlenme sırasında elektrokardiyografik (EKG) sonuçları (1: ST-T segmentinde anormallik 2: Sol ventriküler hypertrophy belirtisi 3: EKG sonucu normal)
8. Maximum kalp atış sayısı
9. Egzersiz sırasında ağrı oluşumu (ağrı var, ağrı yok)
10. ST segmentinde dinlenmeye göre egzersiz sırasında çökme olup olmadığı (çökmedi, çöktü)
11. Egzersiz kaynaklı ST segmentinin şekli (1: Yukarı yönlü 2: Düz 3: Aşağı yönlü)

12. Kalpteki açık ana damar sayısı (1: Bir damar kapalı 2: İki damar kapalı 3: Üç damar kapalı 4: Üç damar açık)

13. Talyum sintigrafi sonucu (1: Normal 2: Sabit kusur 3: Geri dönüştürülebilir)

Çalışmaya dâhil edilen değişkenlerden; 1, 4, 5 ve 8 no'lu değişkenler sürekli, 11 no'lu değişken sıralı, 2, 6, 9 ve 10 no'lu değişkenleri ikili (binary), 3, 7, 12 ve 13 no'lu değişkenler ise kategorik yapıdadır.

3. 2. Yöntem

3. 2. 1. Kanonik (kısıtlı/sınırlandırılmış) ordinasyon

Gaussian benzeri cevap modelleri için kısıtlı ordinasyon problemi, açıklayıcı (çevresel) değişkenlerin doğrusal kombinasyonları olan koordinasyon eksenlerini oluşturmaktır. Bu eksenlere göre Gaussian (logit) yüzeyleri, verilere en iyi şekilde uymaktadır. RDA'da olduğu gibi, çevresel değişkenlerin cevap değişkeni üzerindeki ortak etkileri, cevap değişkenlerini etkileyen ve bileşik çevresel gradientler olarak düşünülebilen birkaç koordinasyon eksenini ile oluşturulur. Ter Braak (1986), bu yaklaşımı Gaussian kanonik ordinasyon olarak ifade etmiştir.

Kanonik Ordinasyon Yöntemleri, ilgilenilen değişkenlerin belirli etkilerini tespit etme gücünü büyük ölçüde geliştirmiştir. Bu Yöntemler, birkaç veri tablosunu aynı anda analiz edebilmektedir. *Doğrudan gradient analizlerinde* iki veri matrisinin birlikte incelenme imkânı vardır ve Ordinasyon ile Çoklu Regresyon Yöntemlerini birleştirir (Tablo 2). Çoklu Regresyon Analizine iteratif algoritmalar eklenerek Kanonik Ordinasyon elde edilir (Ter Braak, 1988).

Tablo 2. Ordinasyon ve regresyon arasındaki ilişki (Legendre ve Legendre, 1998).

Cevap değişkeni	Açıklayıcı değişken	Analiz
1 değişken	1 değişken	Basit regresyon
1 değişken	m değişken	Çoklu regresyon
p değişken	-	Basit ordinasyon
p değişken	m değişken	Kanonik ordinasyon

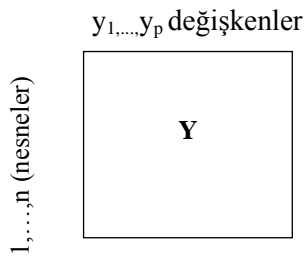
Kanonik Form

Matematikte kanonik form; belli işlevlerin, ilişkilerin veya ifadelerin en az bilgi kaybıyla ifade edilebileceği en basit ancak en kapsamlı bir form olarak görülebilir. Bir kovaryans matrisinin kanonik formu özdeğer matrisidir. Çok boyutlu ölçekleme (MDS) algoritmalarını kullanan bazı kanonik analiz uzantıları tanımlanmış olmasına rağmen, kanonik analiz yöntemleri genel olarak özdeğer analiz (yani özdeğerlerin ve özvektörlerin hesaplanmasını) yöntemini kullanır.

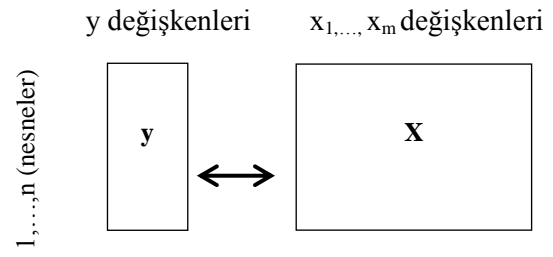
Kanonik Analiz, Ordinasyon ve Regresyon kavramlarını birleştirir. Bir cevap (tepki) matrisini (Y) ve bir açıklayıcı matris (X)'i içerir. Diğer ordinasyon yöntemleri gibi Kanonik Analizler, dağılım şemalarının çizilebileceği (genellikle) ortogonal eksenler üretir. Kanonik Uyum Analizleri, Çoklu Regresyon Analizi ile ilgilidir. Çoklu Regresyon, açıklayıcı değişken seti olarak alınan X değişkenleri ile Y cevap değişkeninin modellenmesi uygun yöntemdir (Legendre ve Legendre, 1998).

(a) Y matrisinin basit ordinasyonu:

Tekil eksenler: Temel Bileşenler Analizi (TBA),
Uyum Analizi (CA)



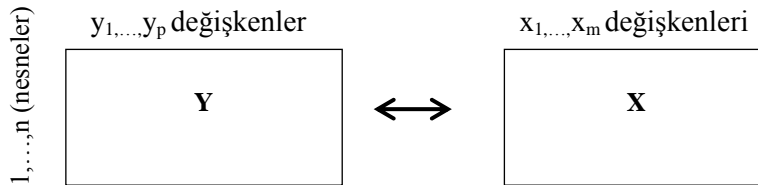
(b) X'in kısıtlaması altında y'nin ordinasyonu: Çoklu Regresyon



$$\text{Model: } \hat{y} = b_0 + b_1x_1 + \dots + b_mx_m$$

(c) X'in kısıtlaması altında Y'nin ordinasyonu:

Gereksizlik Analizi, Kanonik Uyum Analizi



Şekil 3. (a) Ordinasyon, (b) Regresyon ve (c) Kanonik analizin asimetric formları (RDA ve CCA) arasındaki ilişkiler. (c) 'de, Y'nin her bir kanonik eksenini, açıklayıcı değişkenlerin (X'in) doğrusal bir kombinasyonu olarak sınırlandırılmıştır (Legendre ve Legendre, 1998).

Düzenlenmiş değerlerin vektörleri,

$$\hat{y} = b_0 + b_1 x b_{i1} + b_2 b_{i2} + \dots + b_p b_{ip}$$

eşitliği ile bazı nesnelerin (Şekil 3 (b)) yeni bir boyutlu ordinasyonunu gerçekleştirir. Bu nedenle, Çoklu Regresyonda \hat{y} nın ordinasyonu, X'deki değişkenler ile doğrusal ilişkili ve optimal olarak kısıtlanmış olduğundan, y ve \hat{y} ordinatları arasında bir uyum oluşur. Kısıtlama, en küçük kareler yönteminde optimal düzeydedir; yani Çoklu Doğrusal Regresyon, R^2 'yi maksimize eder (Legendre ve Legendre, 1998).

Kanonik Analizler, bu iki yöntem ailesinin özelliklerini birleştirir (Şekil 3(c)) (ordinasyon ve regresyon). Şekil 3(c)'de ikinci bir değişken, X değişken kümesine bağlı olacak şekilde sınırlanan Y'nin ordinasyonlarını üretir. X ve Y arasındaki ilişki yapısı, Kanonik Analiz Yöntemleri arasında farklılık gösterir. RDA, her bir kanonik koordinasyon ekseninde (açıklayıcı değişkenlerin (X) doğrusal bir kombinasyonu ile ilişkili olan nesnelerin (Y matris) çok değişkenli saçılım grafiklerinde) bir yönü gösterir. Bu işlevi ile herhangi bir kanonik eksen, temel bileşenlerle benzerlik gösterir (Legendre ve Legendre, 1998).

Kanonik Diskriminant Analizinde, nesneler niteliksel bir tanımlayıcı tarafından tanımlanan k grub'a ayrılır. Yöntem, k gruplarının merkezlerinin dağılımını maksimuma çıkarır. Bu, nesne-grup dağılımı içine toplanmış nesne-grup dağılımı arasındaki oranın maksimize edilmesiyle elde edilir. Hedef veri seti, nesnelerin bir sınıflandırmasını temsil eden tek bir kalitatif Y değişkeni içerdiğinde, Diskriminant Analizi kullanılabilir (Legendre ve Legendre, 1998).

3. 2. 2. Gereksizlik analizi (Redundancy Analysis/RDA)

Gereksizlik Analizi (RDA) ile ilgili ilk düşünceler iki doğrusal fonksiyon arasında güçlü kanonik korelasyon katsayıları elde edilmesine rağmen bu doğrusal fonksiyonların, değişken kümelerindeki topluluğun varyansı hakkında herhangi bir bilgi vermemesi üzerine gelişmiştir. Rozeboom (1965), tarafından geliştirilen teorik bilgi yardımı ile Stewart and Love (1968), yeni bir ölçüt bulmuşlar ve bu ölçütü, gereksizliği belirleme indeksi olarak adlandırılmışlardır. Bu indeks değeri, bir değişken kümesindeki

ortalama varyans deęerinin, Kanonik Varyans Analizi Yöntemi kullanılarak dięer deęişken kümeleri ile ne kadarının açıklandığının bir ölçüsü olarak tanımlanmıştır (Ural, 1992). Wollenberg (1977), Gereksizlik Analizi (Redundancy Analysis)'ni Gereksizlik İndeksi olarak tanımlamış ve bu indeksin maksimizasyonu Muller (1981), tarafından geliştirilmiştir.

Bu indeks deęeri, bir deęişken kümesindeki ortalama varyans deęerinin dięer deęişken kümelerindeki kanonik bir varyans ile ne kadarının açıklandığının bir ölçüsü olarak tanımlanmıştır (Ural, 1992; Kalkan ve Özden, 2017). Wollenberg (1977), RDA'nın çözümünde sınırlayıcı olan diklik şartını kullanmıştır. Gereksizlik İndeks Deęeri (Eşitlik 8) ile her bir doğrusal fonksiyonun, dięer deęişken kümesinin varyansına olan katkısını göstermektedir. Yani Y, deęişken kümesindeki varyasyon, X deęişken kümesinden elde edilen doğrusal fonksiyonlar ile açıklanmaya çalışılmaktadır. RDA, X açıklayıcı deęişkenlerin regresyonu ile düzenlenmiş Y deęerlerini içeren \hat{Y} matrisindeki nesnelere arası Öklid uzaklığını korur. Bu nedenle, elde edilen \hat{Y} 'deki deęişkenler X deęişkenlerinin doğrusal kombinasyonlarıdır (Lambert ve ark., 1988).

RDA, çoklu açıklayıcı deęişkenlerde çoklu cevap deęişkenlerinin regresyonuna izin vererek Çoklu Doğrusal Regresyon Yöntemini genişletir. Ayrıca bu analizde, MLR (Çoklu Regresyon) yoluyla üretilen tüm cevap deęişkenlerinin düzenlenmiş deęerlerinin matrisi, TBA'ya tabi tutulur. Gereksizlik Analizi hesaplama adımlarına geçmeden veri için aşağıdaki ön hazırlık yapılır.

Ön analiz

1) Cevap deęişkenleri boyutsal olarak homojen deęilse (dięer bir ifade ile veride farklı temel ölçüm birimleri varsa), bu deęişkenler ortalamaları üzerinde merkezileştirilebilir ya da z skoru kullanılarak standartlaştırılabilir. (Standartlaştırma, farklı ölçeklere sahip olabilen regresyon katsayılarının doğrudan karşılaştırılmasına izin verir) (Legendre ve Legendre, 1998).

2) Sistem aşırı-kararlı (over-determined) deęilse, açıklayıcı deęişken sayısının, veri matrisindeki nesne sayısından (siteler, örnekler, gözlemler vb.) az olması gerekir (Legendre ve Legendre, 1998).

Aşırı-kararlılık (overdetermined), cevap (tepki) verisinde belirli bir etki hesaplanırken, gerekli olandan daha fazla açıklayıcı değişken kullanıldığında ortaya çıkan bir durumdur. Diğer bir ifade ile RDA tek bir cevap için birden fazla olası nedeni tespit eder (Legendre ve Legendre, 1998).

Ayrıca, aşırı-kararlılık (over-determined) durumunda, mevcut açıklayıcı değişkenlerin sayısına göre bir veri setindeki nesnelere sayısının artırılabilmesi ifade edilmiştir (Legendre ve Legendre, 1998).

3) Legendre ve Legendre (1998), kalitatif açıklayıcı bir değişkeni doğrusal yanıt verisiyle ilişkilendirmek için RDA'nın kullanılabilmesini belirtmişlerdir. Bu durumda, niteliksel değişken, bir yapay (dummy) değişken olarak yeniden kodlanır ve RDA çalıştırılır. Tahmin edilen (fitted) site puanları nitel açıklayıcı değişkenin, nicel olarak yeniden ölçeklendirilmesini sağlar.

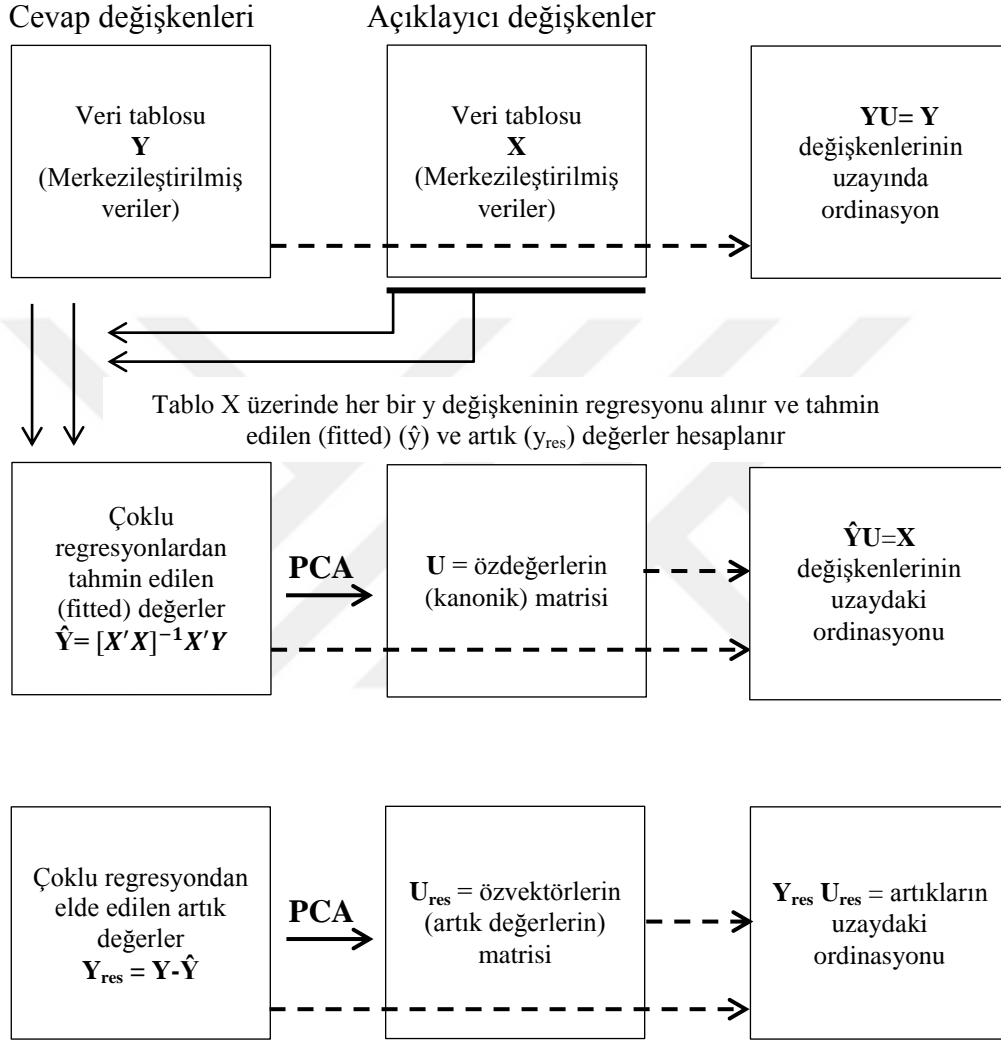
4) Her bir değişkenin, açıklayıcı ve cevap matrisindeki dağılımı incelenebilir. İlişkiler doğrusal olarak belirgin değilse, ilişkileri doğrusallaştırmak ve aykırı (outlier) değerlerin etkisini azaltmak için dönüşümler (transformasyon) uygulanabilir.

A) Gereksizlik (redundancy) analizinin hesaplama adımları

Aşağıdaki cebirsel adımlar, Çoklu Regresyon ve Temel Bileşenler Analizi yoluyla Gereksizlik Analizi eşitliğine (Eşitlik 8)'e nasıl ulaşılabileceğini açıklamaktadır. Şekil (4)'teki adımlar: **i)** X'deki tüm değişkenler üzerinde Y'nin her bir değişkeninin regresyonu alınır, değerleri düzenlenir ve artık değerler hesaplanır. **ii)** Özdeğerler ve özvektörlerin hesaplanması için tahmin edilen (fitted) değerlerin matrisi üzerinde Temel Bileşenler Analizi gerçekleştirilir. **iii)** Sonuçta aşağıda belirtilen iki ordinasyon gerçekleşmiş olur.

Birinci ordinasyonda, açıklayıcı değişkenlerin iki boyutlu uzaydaki ordinasyonu elde edilir. Bu ordinasyonla, “çevresel değişkenlerin doğrusal kombinasyonları” olarak adlandırılan kanonik eksenler elde edilir ve bu eksenlere bağlı olarak site puanları hesaplanır. Ordinasyonla elde edilen kanonik eksenler ortogondur.

İkinci ordinasyonda, bağımlı değişkenlerin (tür alanlarının) alanlardaki ordinasyonu elde edilir. Bu ordinasyon, “örnek puanları” ya da “tür puanlarını” verir. Ancak, buradan elde edilen koordinasyon eksenleri ortogonal değildir (Legendre ve Legendre, 1998; Borcard, 2006).



Şekil 4. Doğrudan algoritma kullanılarak Gereksizlik Analizi hesaplama adımları. Gereksizlik Analizi, iki aşamalı bir süreç olarak anlaşılabilir: (1) X 'deki tüm değişkenler üzerinde, Y 'nin her bir değişkeninin regresyonu alınır ve değerler düzenlenir (2) Özdeğerler ve özvektörlerin hesaplanması için tahmin edilen (fitted) değerlerin matrisi üzerinde Temel Bileşenler Analizi gerçekleştirilir. Sonuçta iki ordinasyon elde edilir. Biri (Y) cevap değişkenlerinin uzayında (YU) ordinasyonu, diğeri (X) açıklayıcı değişkenlerin uzayında ($\hat{Y}U$) ordinasyonu elde edilir (Legendre ve Legendre, 1998).

Bir başka TBA ordinasyonu ise artıkların matrisi kullanılarak elde edilebilir (Legendre ve Legendre, 1998).

1) Şekil (4)' te Y'deki her bir cevap değişkeni için, X'teki tüm değişkenlerle çoklu doğrusal regresyon hesaplanır. Her bir Y değişkeni için regresyon katsayılarına ait eşitlik:

$$b = [X'X]^{-1}X'Y$$

olarak hesaplanır. Bu eşitlik, tüm Y değişkenleri için matris notasyonu ile

$$B = [X'X]^{-1}X'Y \quad (1)$$

olarak ifade edilir. Eşitlik (1)'de B, tüm cevap (Y) değişkenleri için açıklayıcı değişkenlere ait regresyon katsayıları matrisidir.

Çoklu regresyonda, tahmin edilen (fitted) \hat{Y} değerleri;

$$\hat{Y} = XB \quad (2)$$

eşitliği ile elde edilir. Tüm tahmin edilen (fitted) değerler (\hat{Y}) kümesi, Eşitlik (2)'de gösterildiği gibi tek bir matris işlemiyle hesaplanabilir. Eşitlik (1), Eşitlik (2)'de yerine yazıldığında tahmin edilen (fitted) değerler (\hat{Y}) kümesi;

$$\hat{Y} = X[X'X]^{-1}X'Y \quad (3)$$

olarak hesaplanır (X ve Y değişkenleri merkezleştirildiğinden, B vektörlerinde regresyon sabiti parametresi yoktur). Standart çoklu regresyonda olduğu gibi \hat{Y} vektörleri genellikle merkezleştirilir ve $m = n$ olduğunda X karesel matris olur. Bu durumda, çoklu regresyonlar Y matrisindeki değişkenleri tamamen açıklar ve $\hat{Y} = Y$ olur (Legendre ve Legendre, 1998).

2) Tahmin edilen (fitted) \hat{Y} değerlerine karşılık gelen uygun bir kovaryans matrisi;

$$S_{\hat{Y}\hat{Y}} = [1/(n - 1)]\hat{Y}'\hat{Y} \quad (4)$$

eşitliği ile elde edilir. Eşitlik (3)'teki \hat{Y} değeri Eşitlik (4)'te yerine yazıldığında kovaryans-matris;

$$S_{\hat{Y}\hat{Y}} = \left[\frac{1}{n - 1} \right] Y'X[X'X]^{-1}X'X[X'X]^{-1}X'Y \quad (5)$$

olarak hesaplanır. Eşitlik (5) düzenlenerek kovaryans matrisi;

$$S_{\hat{Y}'\hat{Y}} = S_{YX}S_{XX}^{-1}S'_{XX} \quad (6)$$

olarak elde edilir. Eşitlik (6)'da, S_{YY} cevap değişkenleri arasındaki (pxp) boyutlu kovaryans matrisi, S_{XX} çok değişkenli regresyonlar arası mxm boyutlu kovaryans matrisi ve S_{YX} iki değişken seti arasındaki pxm boyutlu bir kovaryans matrisidir ($S'_{YX} = S_{XY}$ (m xp) boyutlu bir matristir). Y değişkenlerinin standardize edilmiş olması durumunda, Eşitlik (6) için $R_{YX}R_{XX}^{-1}R'_{YX}$, çoklu belirleme katsayısı olur (Legendre ve Legendre, 1998).

3) Tahmin edilen (fitted) \hat{Y} değerleri, çözümün boyutunu azaltmak için Temel Bileşenler Analizine tabi tutulur. Bu analizde özdeğer problemini çözmek için eşitlik;

$$(S_{\hat{Y}'\hat{Y}} - \lambda_k \mathbf{I})u_k = 0 \quad (7)$$

olarak ifade edilir. Eşitlik (6) kullanılarak Eşitlik (7);

$$(S_{YX}S_{XX}^{-1}S'_{YX} - \lambda_k \mathbf{I})u_k = 0 \quad (8)$$

eşitliğine dönüştürülür. Elde edilen Eşitlik (8), Gereksizlik Analizi eşitliğini vermektedir.

Normalize edilmiş kanonik özvektörleri (u_k) içeren matris, U olarak adlandırılır. Özvektörler, tahmin edilen \hat{Y} değerlerini oluşturan değişkenlerin kanonik eksenlere katkılarını gösterir. Bu katkılar TBA'daki gibi yorumlanır. pxp boyutlu U matrisi, sadece sıfır olmayan özdeğerlere sahip, min [p, m, n-1] boyutlu özvektörleri içerir (kanonik özvektör sayısının min [p, m, n-1] değerini aşmaması gerekmektedir). Aşağıda anlatılacak biplotlar için kullanılan U matrisi, her bir özvektör uzunluğu $\sqrt{\lambda_k}$ olacak şekilde yeniden ölçeklendirilir.

Çoklu regresyonda, Y matrisi uzayının referans büyüklüğünün p'yi aşmaması gerektiği belirtilmiştir. Burada Y matrisi tek bir değişken içerir ve tahmin edilen (fitted) \hat{Y} değerleri tarafından yapılan ordinasyon tek boyutlu olur. Ayrıca, X'deki değişkenlerin sayısı m'yi, Öklit uzayında n noktası için gerekli boyut sayısı da max (n-1)'i aşmamalıdır (Legendre ve Legendre, 1998).

4) Eşitlik (8)'de bulunan u_k özvektörlerin (U) matrisi ve Temel Bileşenler eşitliği kullanılarak merkezleştirilmiş Y matrisiyle;

$$F = YU \quad (9)$$

eşitliği elde edilir. Y cevap değişkenleri uzayındaki nesnelere ordinasyonu Eşitlik (9) ile hesaplanır.

Site (konum, lokasyon) puanları

Eşitlik (9)'da tanımlanan ordinasyon vektörleri (F'nin sütunları) “site puanları”nın vektörleri olarak adlandırılır. Bu vektörler, özdeğerleri eşit olmayan ancak birbirlerine yakın varyans değerlerine sahiptir. Biplot diyagramlarında F matrisi'nin nasıl temsil edileceği 8. adımda anlatılacaktır.

5) Benzer şekilde, X uzayında nesnelere ordinasyonu aşağıdaki gibi elde edilir.

$$Z = \hat{Y}U = XB\hat{U} \quad (10)$$

Eşitlik (10)'da görüldüğü üzere \hat{Y} matrisindeki vektörler sırasıyla merkezleştirilir. Bu ordinasyon vektörleri (Z matrisinin sütunları) “düzenlenmiş site puanları” veya “bazı paket programlarında çevresel değişkenlerin doğrusal kombinasyonları olan örnek puanlar” olarak da adlandırılır.

Eşitlik (6)'daki “site puanları”; k eksenleri üzerinde orijinal verilerin (Y matrisi) iz düşümü ile elde edilir. İz düşümler yaklaşık olarak gözlenen verilere karşılık gelir. Diğer yandan, Eşitlik (10)'daki “tahmin edilen (fitted) site puanları”, çoklu regresyonların (\hat{Y} matrisi) tahmin ettiği (fitted) değerlerin, k eksenleri üzerindeki iz düşümünden elde edilir (Legendre ve Legendre, 1998).

6) r_k , k boyutludur ve Eşitlik (9)'da hesaplanan Y ve Eşitlik (10)'da hesaplanan X uzayındaki ordinasyon vektörleri arasındaki korelasyonu göstermektedir.

r_k , her bir kanonik (k) eksen tarafından açıklanan iki veri kümesi arasındaki ilişkinin miktarını ölçer. Yüksek oranda cevap değişkeni-açıklayıcı değişken ilişkisine

sahip olan kanonik eksen, Y'deki varyasyonun küçük bir kısmını açıklayabilir. Bu nedenle r_k , dikkatli yorumlanmalıdır, (Legendre ve Legendre, 1998).

7) Açıklayıcı değişkenlerin (X) kanonik ordınasyon eksenlerine katkısının belirlenmesinde, regresyon ya da korelasyon katsayıları incelenir.

Kanonik katsayılar matrisi (C),

$$C = BU \quad (11)$$

olarak düzenlenmiş site puanlarının matrisinde açıklayıcı değişkenlerin (X) ağırlıklarını verir. Açıklayıcı değişkenlerin uzaydaki nesnelere ordınasyonu, XC olarak doğrudan hesaplanabilir. Site puanlarının vektörleri, Eşitlik (10)'daki ile aynıdır. C matrisinin sütunlarındaki katsayılar ise Eşitlik (11)'deki standardize edilmiş açıklayıcı değişkenler (X) matrisinin ordınasyon puanlarının regresyon katsayıları ile aynıdır.

Korelasyonlar, X'deki değişkenler arasında ya da Eşitlik (9)'daki Y uzayı ile Eşitlik (10)'daki X uzayının ordınasyon vektörleri arasında hesaplanabilir. X ile Y uzayındaki ordınasyon vektörleri arasındaki korelasyonlar, biplotlardaki açıklayıcı değişkenleri temsil etmek için kullanılır (Legendre ve Legendre, 1998).

8) RDA'daki biplotlar, iki set noktadan ya da TBA'da olduğu gibi üç set noktadan oluşmaktadır. Bu noktalar; site puanlarını (Eşitlik 9 ve Eşitlik 10'daki F ve Z matrisleri), cevap değişkenlerini (Y) ve açıklayıcı değişkenleri (X) temsil etmektedir. Biplotlar, nesnelere Y ve X değişkenleri açısından ordınasyonunun yorumlanmasını sağlar. Y'de ya da X'de, çok fazla nesne ya da çok fazla değişken olduğunda, ayrı ordınasyon şemaları çizilebilir ve yan yana sunulabilir. Ordınasyonda, TBA'da olduğu gibi, iki temel ölçekleme çeşidi kullanılır (Legendre ve Legendre, 1998).

9) Gereksizlik Analizi genellikle cevap değişkenleri (Y) matrisindeki varyasyonu tamamen açıklamamaktadır. Regresyon adımı sırasında (Şekil 4), her bir y değişkeni için regresyon artıkları hesaplanabilir; artıklar, \hat{Y} matrisindeki düzenlenmiş \hat{y}_{ij} değerleri ve gözlenen y_{ij} değerleri arasındaki farktan elde edilir. Artıklar matrisi (Şekil 4'deki Y_{res}), $n \times p$ boyutludur. Artıklar, Temel Bileşenler Analizi ile incelenebilir, bu da $\min [p, n-1]$ olan özvektörler ve kanonik olmayan özdeğerleri verir (Şekil 4'ün

sonunda gösterilmiştir). Böylece, Y matrisinin tam analizi (yani, düzenlenmiş değerlerin ve artıkların analizi) yapıldığında, Y matrisinin Temel Bileşen Analizi sonucundan daha fazla özvektör elde edilir. Çeşitli kısıtlamalar nedeniyle, maksimum kanonik ve kanonik olmayan eksen sayısı farklılık gösterebilir (Tablo 3). Artıklar matrisinde en fazla min $[p, m, n - 1]$ sıfır olmayan kanonik özdeğerler ve özvektörler ile en fazla min $[p, n - 1]$ kanonik olmayan özdeğerler ve özvektörler hesaplanır (Tablo 3).

Tablo 3. RDA ve KUA kullanılarak cevap değişkenleri matrisi ve açıklayıcı değişkenler matrisinin kanonik analizinden elde edilen özdeğerler ve bunlara karşılık gelen özvektörler (Legendre ve Legendre, 1998).

	Kanonik özdeğerler ve özvektörler	Kanonik olmayan özdeğerler ve özvektörler
RDA	min $[p, m, n-1]$	min $[p, n-1]$
CCA	min $[p-1, m, n-1]$	min $[p-1, n-1]$

X'deki değişkenler Y'deki değişkenleri iyi tanımladığında, bazı durumlarda kanonik özdeğerler ilk kanonik olmayan özdeğerlerden daha büyük çıkabilir. X'deki değişkenler Y'yi büyük oranda açıklayamadığında, artıklar üzerinde hesaplanan ilk kanonik olmayan özdeğerler, kanonik özdeğerlerden daha büyük çıkabilir. Y'nin tek bir cevap değişkeni içerdiği durumlarda Gereksizlik Analizi, Çoklu Doğrusal Regresyon Analizi ile aynı sonuçları verir (Legendre ve Legendre, 1998).

B) RDA biplotlarının ve triplotlarının okunması ve yorumu

RDA ordnasyonu, biplot (Şekil 5 (a)) ya da triplot (Şekil 5 (b)) olarak sunulabilir. Aynı diyagramda nesnelere, cevap değişkenleri ve açıklayıcı değişkenler birlikte sunulur. Açıklayıcı değişkenler, nitel (çok sınıflı olanlar ikili değişken olarak kodlanır) veya nicel olabilir. Triplotlarda/Biplotlarda, nitel açıklayıcı değişkenler koordinat merkezinde, nicel açıklayıcı değişkenler ise vektörler olarak temsil edilir (Borcard, 2006).

Grafiklerin yorumu, ölçekleme seçimine göre farklılık gösterir. Hangi ölçeklemenin kullanılacağı, biplot'ta belirtmek istenen vurguya göre değişir. Nesnelere arasındaki uzaklıklara vurgu yapılmak isteniyorsa Tip 1, değişkenler arasındaki ilişki yapılarına vurgu yapılmak isteniyorsa Tip 2 ölçeklemesi kullanılmalıdır. Ayrıca,

açıklayıcı değişkenlerin çoğunluğu binary olduğunda, Tip 1 ölçeklemesinin daha yaygın kullanıldığı görülmüştür. Bunun yanısıra, X veri kümesindeki değişkenlerin çoğunluğu kantitatif olduğunda Tip 2 ölçeklemesinin daha çok tercih edildiği ve ilk iki özdeğer yaklaşık olarak eşit çıktığında, her iki ölçeğin biplot görüntüsünün de yaklaşık olarak aynı sonucu vereceği belirtilmiştir (Legendre ve Legendre, 1998).

C) RDA'da ölçeklendirme tipleri

Tip 1 Ölçeklemesi- Uzaklık Biplotu (Nesne Odaklı)

Kanonik eksenler boyunca cevap değişkenlerinin puanlarını temsil eden U matrisindeki özvektörler aşağıda belirtildiği şekilde ölçeklenir (yani, özvektörler birim uzunluğuna göre ölçeklenir). X uzayındaki site puanları (Z matrisi) ile Y uzayındaki site puanları (F matrisi), $\sqrt{n\lambda_k}$ uzunluğuna ölçeklendirilir. RDA'daki özdeğerler Y'deki toplam varyansı gösterir. Z ve U ya da F ve U matrisleri biplotlarda birlikte kullanılabilir. Zira, site puanları matrislerine sahip özvektörlerin sonuçları orijinal matrisleri yeniden oluşturabilir (TBA'da olduğu gibi $ZU' = \hat{Y}$ ve $FU' = Y$ orijinal matrisler elde edilebilir).

Bir kantitatif açıklayıcı değişken, düzenlenmiş site puanları ile X'in korelasyonları kullanılarak biplotta temsil edilebilir. Her korelasyon, k eksenine karşılık gelen özdeğeri ile çarpılır. Ölçeklemede yapılan bu düzeltme, site puanları varyanslarının eksenler arasında farklılık gösterdiğini açıklamaktadır. Nesnelere arası uzaklıklar cevap değişkenleri uzayında yaklaşık Öklid uzaklıklarına karşılık geldiğinden, Uzaklık biplotu olarak adlandırılan bu ölçekleme türü nesnelere ordinasyonu üzerine odaklanır (Legendre ve Legendre, 1998).

Bir uzaklık biplotunun (Şekil 5) temel özellikleri aşağıdadır:

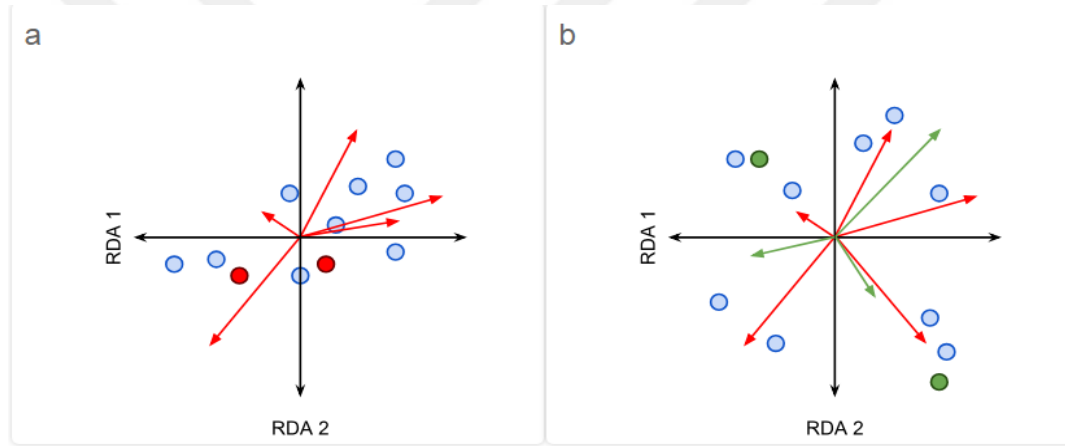
(1) Bir biplottaki nesnelere arasındaki uzaklıklar, yaklaşık Öklid uzaklığına karşılık gelmektedir. Böylece, birbirine yakınlaştırılan nesnelere benzer değişken değerlerine sahip olması beklenir. Bu ise her zaman gerçekleşmeyebilir. Zira RDA, veri kümesindeki varyasyonun her zaman tamamını açıklayamamaktadır.

(2) Bir cevap deęişkeni üzerinde bir nesnenin iz düşümü, bu deęişken boyunca nesnenin yaklaşık deęerini gösterir.

(3) Y deęişkenleri (cevap vektörleri) arasındaki açı anlamsızdır.

(4) Biplottaki X ve Y deęişkenleri arasındaki açı, bu deęişkenlerin arasındaki korelasyonu ifade etmektedir.

(5) Binary açıklayıcı deęişkenlerden, “1” olarak kodlanan deęişkenler nesnelere merkezi olarak temsil edilir. Kalitatif açıklayıcı bir deęişkenin merkezi ile bir cevap deęişkeni arasındaki ilişki, merkezi noktanın deęişken üzerindeki iz düşümü ile (nesnelere için olduğu gibi) bulunur. (Legendre ve Legendre, 1998; Borcard, 2006).



Şekil 5. a) RDA biplot'u b) RDA triplotunun şematik gösterimi. a) Bir RDA biplot ordinasyonunda, nesnelere noktalarla ve cevaplar ya da açıklayıcı deęişkenler ise vektörlerle (kırmızı oklar) temsil edilir. Nominal deęişkenlerin seviyeleri kırmızı noktalarla temsil edilir. b) Bir RDA triplot ordinasyonunda, nesnelere mavi noktalarla temsil edilirken, cevap ve açıklayıcı deęişkenler vektörler (kırmızı ve yeşil oklar) olarak gösterilir. Nominal deęişkenlerin seviyeleri yeşil noktalarla temsil edilir. Varsayılan görsel sonuçların sunumu paket programlarına göre deęişebilir. Metinlerde ölçeklendirmeye baęlı olarak parsellerin yorumlanması tartışılmaktadır (Anonymous 2, 2018).

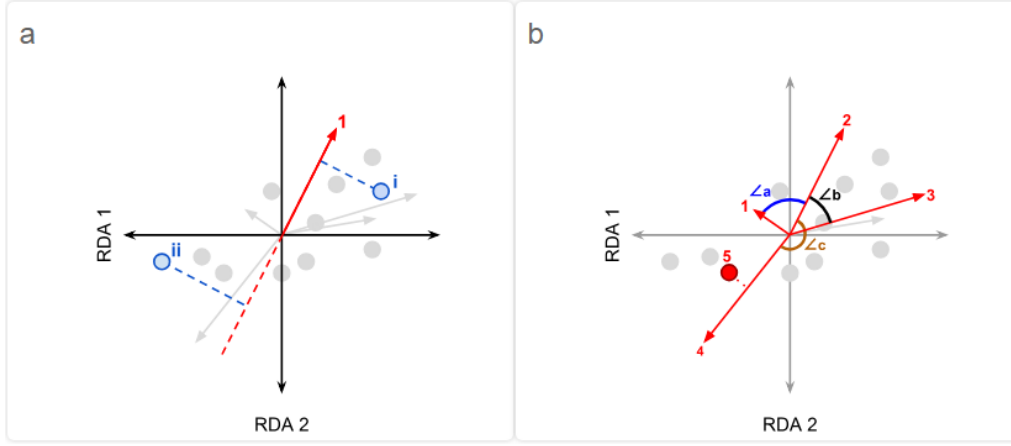
Tip 2 Ölçeklemesi - Korelasyon Biplotu (cevap deęişken odaklı)

Alternatif bir yöntem olarak, U matrisindeki özvektörler $\sqrt{(\lambda k)}$ uzunluęuna yeniden ölçeklendirilerek, cevap deęişken puanları elde edilir (özvektörler, özdeęerlerin kareköküne göre ölçeklenir). Tip 1 ölçekleme için elde edilen X uzayındaki site puanları, $\mathbf{Z}\Lambda^{-1/2}$ transformasyonu ile birim varyansa göre yeniden ölçeklenir. Tip 1

ölçeklenme için elde edilen Y uzayındaki site puanları ise $F\Lambda^{-1/2}$ dönüşümü kullanılarak yeniden ölçeklenir. Bu vektörlerin varyansları, Tip 1 ölçeklemedeki açıklanan varyansa göre genellikle 1 değerinden biraz daha büyük çıkmaktadır (Legendre ve Legendre, 1998).

Z ve U ya da F ve U matrisleri, yeniden ölçeklendirilmiş olarak, biplotların içinde birlikte kullanılabilir. Zira, site puanları matrislerine sahip özvektörlerin sonuçları ile orijinal matrisler yeniden oluşturulabilir (TBA'da olduğu gibi $ZU'=\hat{Y}$ ve $FU'=Y$ orijinal matrisler elde edilebilir). Kantitatif açıklayıcı değişken (X), yukarıda hesaplama adımı 7'de elde edilen yerleştirilmiş site puanları ile X 'in korelasyonları kullanılarak biplotta temsil edilir (Legendre ve Legendre, 1998).

Bir korelasyon biplotunun (Şekil 6) temel özellikleri şunlardır: (1) Biplottaki nesnelere arasındaki uzaklıklar, yaklaşık öklid uzaklığı olarak düşünülmemelidir. (2) Bir cevap ya da açıklayıcı değişken üzerinde bir nesnenin iz düşümü, bu değişken boyunca nesnenin yaklaşık değerini gösterir. (3) Biplottaki değişkenler (X ve Y değişkenleri) arasındaki açı, bu değişkenler arasındaki korelasyonu belirtmektedir. Korelasyon, vektörler arasındaki açının kosinüsüne eşittir (örneğin, 90° açısını tanımlayan bir vektör çifti, $\cos(90) = 0$ ile ilişkisizdir; 20° 'lik bir açıyı tanımlayan bir başka vektör çifti ise $\cos(20) = 0.94$ ile güçlü bir pozitif korelasyona sahiptir). (4) Biplottaki tanımlayıcılar arasındaki açılar, bunların kovaryanslarını ya da korelasyonlarını gösterir. (5) Kalitatif açıklayıcı bir değişkenin merkezi ile bir cevap değişkeni (tür) arasındaki ilişki yapısı, merkezi noktanın değişkene iz düşümü ile (nesnelere için olduğu gibi) bulunur. Binary açıklayıcı değişkenlerden, "1" olarak kodlanan değişkenler nesnelere merkezi olarak temsil edilir (Legendre ve Legendre, 1998; Borcard, 2006).



Şekil 6. a) Tahmin edilen (fitted) nesnelere bir vektör üzerine iz düşümü b) Vektörler arasındaki açıları vurgulayan şemalar. (a) Bir değişkenin vektör üzerine iz düşümü (i noktasının şekil (a)'da iz düşümünün gösterildiği gibi) bu nesnenin yaklaşık değeridir. Dolayısıyla, i değişkeni, diğer birçok değişkene göre değişken 1'in daha yüksek değerlerinde bulunabileceğini göstermektedir. Bunun yanı sıra, ii değişkeni, diğer değişkenlere göre değişken 1'in daha düşük değerlerinde bulunabileceğini göstermektedir. b) Z_a değeri 90° 'ye yaklaşmaktadır. Dolayısıyla "1" ve "2" değişkenleri arasında çok düşük bir ilişki olduğu söylenebilir (yani, bağımsız eksenler gibi neredeyse ortogondur). Z_b değeri 90° 'den küçüktür. O halde "2" ve "3" değişkenleri arasında pozitif korelasyon olduğu söylenebilir. Z_c 180° 'ye yaklaşırken, "2" ve "4" değişkenleri arasında güçlü negatif korelasyon olduğu görülmektedir (değişkenlerin artış yönlerine bakılarak). Değişken 5 nicel değildir ve bir merkezi nokta ile temsil edilmektedir (Anonymous 2, 2018).

D) Gereksizlik analizi diyagramı yorumu

Gereksizlik Analizi diyagramında siteler noktalarla gösterilir. Hem cevap değişkeni (türler veya nesnelere) hem de açıklayıcı (çevresel) değişkenler oklarla gösterilir ve yorumları TBA biplotundaki okların yorumuna benzer şekilde yapılır.

Siteler arasında her bir cevap değişkeni miktarının şablonu tam bir TBA biplotunda olduğu gibi görünür. Ayrıca RDA, cevap değişkenlerinin bolluğu ve çevresel değişkenler arasındaki korelasyonu da gösterir. Aynı yöne işaret eden okların yüksek pozitif korelasyon gösterdiği, dik kesişen okların sıfıra yakın korelasyon gösterdiği ve zıt yönlerde işaret eden okların ise yüksek negatif korelasyon gösterdiği belirtilmiştir. Uzun oklarla gösterilen açıklayıcı değişkenlerin, cevap değişkenleri için önemli değişkenler olduğu ifade edilmiş ve okların uzunluğu ne kadar fazla olursa, cevap değişkenlerinin varlığı ve çevresel değişkenler arasındaki korelasyonuna yüksek

olacağı vurgulanmıştır. Bu nedenle, Ter Braak ve Prentice (1988), RDA ordinasyon şemasındaki okların, cevap değişkenleri ile çevresel değişkenler arasındaki yaklaşık (doğrusal) korelasyon katsayılarını gösterdiğini ifade etmişlerdir (Ordinasyon şemasının amacına uygun olması için açıklayıcı değişkenler, karşılaştırılabilir ok uzunluklarına sahip olacak şekilde sıfır ortalama ve birim varyansa standartlaştırıldığı varsayılarak bu yorum yapılmıştır) (Ter Braak ve Prentice, 1988).

3. 2. 3. Kanonik uyum analizi (canonical correspondence analysis)

KUA, Çok Değişkenli Doğrudan Gradient Analizi üreten bir özvektör ordinasyon yöntemidir (Ter Braak, 1986). Standart ordinasyonda olduğu gibi bir topluluk varyasyonu modelini ve cevap değişkenlerinin, çevresel değişkenler boyunca dağılımlarının temel özelliklerini görselleştirmeyi amaçlamaktadır. (Ter Braak, 1987). Analizde tahminleri elde etmek için Maksimum Olabilirlik (ML) Yöntemi kullanılır. KUA'nın çözümü, genellikle ağırlıklı ortalama algoritması ile elde edilmesine rağmen, yöntem temelde bir özdeğer analizidir ve dolayısıyla herhangi bir özdeğer algoritmasıyla da çözüm sağlanabilir (Ter Braak 1986, 1987).

KUA 'nın temel yaklaşımındaki istatistik model, bir cevap değişkeninin varlığını veya frekansını, çevresel gradientler boyunca unimodal bir konum fonksiyonu olarak ifade eder. KUA; belirli bazı varsayımlar altında bir Gaussian regresyon yaklaşımıdır ve bu varsayımların sağlanmamasına karşı güçlüdür (Ter Braak ve Prentice, 1988).

Cevap değişkenlerinin ağırlıklı ortalamalarından elde edilen site yada lokasyon puanlarının hesaplanmasında; site puanları bağımlı değişkenleri (cevap değişkenleri), çevresel değişkenler ise bağımsız değişkenleri (açıklayıcı değişkenleri) göstermektedir. Yeni site puanları, regresyon eşitliği tarafından tahmin edilen değerlere karşılık gelir. Regresyon eşitliği, değişkenlerin bir doğrusal kombinasyonunu gösterdiğinden, yöntemde ağırlıklı ortalama (WA) yerine doğrusal kombinasyon puanları da kullanılabilir (Palmer, 1993).

A) Model

Bir n bölgesinde m; cevap değişkeninin miktarını ya da oluşumlarını, q_j ($q_j < n$) çevresel değişkenlerin değerlerini, y_{ik} ($y_{ik} \geq 0$) i. sitedeki k cevap değişkeninin varlığını/yokluğunu (1/0) ve z_{ij} ise i.sitedeki j çevresel değişkenlerin değerini göstermiş olsun.

Dolaylı Gradient Analizinin ilk adımı, cevap değişkeni verisindeki temel varyasyonu ordinasyon yoluyla özetlemektir. Gaussian Ordinasyon Yöntemi (Gauch ve ark., 1974); cevap değişkenlerinin, eksen boyunca Gauss cevap eğrilerine en iyi şekilde uyacağı bir eksen oluşturur. Cevap modeli (Gaussian) aşağıdaki gibi yazılır.

$$E(Y_{ik}) = c_k \exp[1/2(x_i - u_k)^2/t_k^2] \quad (12)$$

Eşitlik (12)'de $E(y_{ik})$, ordinasyon ekseninde x puanlarına sahip i. sitedeki y_{ik} 'nin ortalama beklenen değerini gösterir. k cevap değişkeni için parametreler c_k dir. Cevap değişkenlerinin cevap eğrilerinin maksimum değeri olan u_k optimumdur ve t_k (tolerans), ekolojik bolluğun bir ölçüsüdür (Ter Braak, 1986).

Ter Braak (1985b), cevap değişkeni bolluğunun örnekleme dağılımı Poisson dağılım gösterdiğinde; UA'nın Gauss Ordinasyonunun, en çok olabilirlik çözümüne yaklaşacağını göstermiştir. Ordinasyon yöntemlerinde cevap değişkeninin (türün) aşağıdaki kriterleri taşıması gerekir.

C1) Cevap değişkeninin tolerans değeri ($t_k=t, k=1, \dots, m$)

C2) Cevap değişkeninin cevap eğrilerinin maksimum değeri ($c_k=c, k=1, \dots, m$)

C3) Cevap değişkeninin optimum (u_k) değeri, t'ye göre büyük bir A aralığında homojen dağılır.

C4) Site puanları (x_i), A sitesinde ve geniş bir B aralığında homojen bir şekilde dağılır.

“Homojen olarak dağılır” ifadesi, puanların eşit olduğunu veya puanların eşit bir dağılımdan rastgele alındığını ifade eder. C1 ve C2 koşulları, birçok doğal toplulukta

sağlanmayabilir. Ancak uygulamada Uyum Analizinin avantajı; bu koşulların ihlallerine karşı sağlamlığına bağlıdır (Hill ve Gauch, 1980).

Dolaylı Gradient Analizinin ikinci adımında, çevresel değişkenler üzerinde site puanlarının çoklu regresyonları ya da korelasyon katsayıları aşağıdaki gibi hesaplanır (Ter Braak, 1986).

$$x_i = b_0 + \sum_{j=1}^q b_j z_{ij} \quad (13)$$

Eşitlik (13)' ile, ordinasyon eksenini açıklayıcı (çevresel) değişkenlerle grafiksel olarak ilişkilendirilir ve Eşitlik (13)'te b_0 sabit terim ve b_j ise j. çevresel değişken için regresyon katsayısıdır. Cevap değişkeninin optimum değerlerinin (u_k) ve site puanlarının (x_i), cevap değişkeni ilk verilerinden tahmin edildiği unutulmamalıdır. x_i ve u_k sabit tutularak, b_j regresyon katsayısı tahmin edilir. Bu nedenle cevap değişkeni verileri, koordinasyon eksenini aracılığıyla çevresel değişkenlerle dolaylı olarak ilişkilidir.

Burada önerilen yöntemle, aynı zamanda cevap değişkeninin optimum değeri, regresyon katsayıları (Eşitlik (12) ve Eşitlik (13)) birleştirilerek açıklanan model sayesinde site puanları tahmin edilir ve tahminlerde ML yöntemi kullanılır. Gaussian Kanonik Ordinasyon olarak da bilinen bu yöntem, C1-C4 koşulları altında hesaplama yapar. Gaussian Ordinasyonundan, UA'ya geçişi aşağıdaki gibi özetlenebilir.

$$\lambda u_k = \sum_{i=1}^n \frac{y_{ik} x_i}{y_{+k}} \quad (14)$$

$$x_i^* = \sum_{k=1}^m \frac{y_{ik} u_k}{y_{i+}} \quad (15)$$

$$\mathbf{b} = (\mathbf{Z}'\mathbf{R}\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{R}\mathbf{x}^* \quad (16)$$

$$\mathbf{x} = \mathbf{Z}\mathbf{b} \quad (17)$$

Eşitlik (14 ve 15)'de; u_k ; k'nci (m dışında) cevap değişkenlerinin ağırlıklı ortalamasını, x_i ; i'nci (n dışında) sitenin (örn; nem değeri) değerini, y_{ik} ; i sitesinde bulunan k cevap değişkeninin bolluğunu (frekansını), y_{+k} ; k cevap değişkeninin toplam

bolluğunu ve y_{i+} ; toplam siteleri göstermektedir. R ; y_{i+} 'den oluşan $n \times n$ diagonal matristir. $Z = \{Z_{ij}\}$, çevresel verileri içeren $n \times (q+1)$ boyutlu matristir. \mathbf{b} , \mathbf{x} ve \mathbf{x}^* sütun vektörleridir: $\mathbf{b} = (b_0, b_1, \dots, b_q)'$, $\mathbf{x} = (x_0, x_1, \dots, x_n)'$ ve $\mathbf{x}^* = (x_1^*, \dots, x_n^*)'$.

Geçiş eşitlikleri, Eşitlik (14)'teki λ özdeğeri problemini oluşturur. UA'da olduğu gibi eşitliklerde, tüm site ve cevap değişkeni puanları birbirine eşit ve $\lambda = 1$ olduğunda, elde edilen sonuç önemsiz bir çözüme sahip olur. Bu önemsiz çözüm göz ardı edilebilir ya da sıfır ortalama için site puanları merkezileştirilerek önemsiz kısım çıkartılabilir (yani $\sum_i y_{i+} x_i = 0$).

B) CCA algoritması: Karşılıklı ortalama ve regresyon

Geçiş eşitlikleri, karşılıklı ortalama alma ve çoklu regresyona ilişkin aşağıdaki adımlar (iterasyon algoritması) izlenerek çözülebilir.

- 1) Rasgele, ancak eşit olmayan, ilk site puanları ile başlanır (Eşitlik (13)'teki x_i değerleri).
- 2) Site puanlarının ağırlıklı ortalamalarına göre cevap değişkeni puanları hesaplanır (Eşitlik (14)'te $\lambda=1$ alınarak u_k değerleri hesaplanır).
- 3) Cevap değişkeni puanlarının ağırlıklı ortalaması alınarak yeni site puanları hesaplanır (Eşitlik (15)'teki, x_i^* ; ağırlıklandırılmış ortalama ile yeni site puanları hesaplanır).
- 4) Çevresel değişkenler üzerinde site puanlarının ağırlıklı çoklu regresyonu ile regresyon katsayıları elde edilir (Eşitlik (16)'da ki \mathbf{b} değerleri elde edilir).
- 5) Yeni site puanları, Eşitlik (17) ya da ona eşdeğer olan Eşitlik (13) ile hesaplanır. Yeni site puanları, aslında önceki adımın düzeltilmiş regresyon katsayılarıdır.
- 6) Site değerleri;

$$\sum_{i=1}^n y_{i+} x_i = 0 \text{ ve } \sum_{i=1}^n y_{i+} x_i^2 = 1 \quad (18)$$

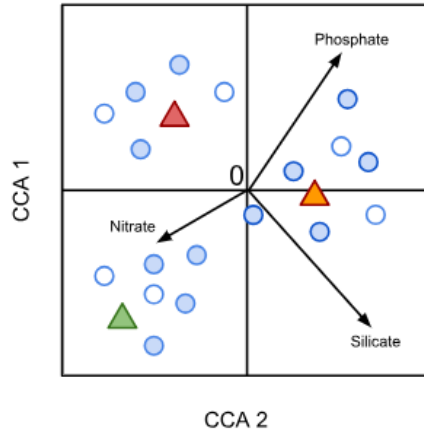
eşitliği ile merkezileştirilir ve standardize edilir.

7) Yeni site puanları, önceki iterasyon site puanlarına yeterince yakınsama sağladığında iterasyon süreci durur. Aksi halde iterasyon, ikinci adımdan itibaren tekrar eder (Ter Braak, 1986).

KUA'nın algortima adımları, UA'nın karşılıklı ortalama algoritmasına benzer olmakla birlikte, KUA'da ilave olarak 4 ve 5 adımları eklenmiştir. Yeni bir yöntem olan KUA, aslında site puanlarındaki kısıtlamalar (4 ve 5 adımları) ile bir UA yöntemine benzemektedir (De Leeuw, 1984). Ter Braak (1986), son regresyon katsayılarının; kanonik katsayılar olarak ve bu regresyondaki çoklu korelasyon katsayısının ise cevap değişkeni–açıklayıcı değişken korelasyonu olarak adlandırıldığını belirtmiştir. Cevap değişkeni–açıklayıcı değişken korelasyonu, nesne kompozisyonundan çıkarılan varyasyonun açıklayıcı değişkenler tarafından ne kadar iyi açıklanabildiğinin bir ölçüsüdür ve site puanları (x_i^* : Eşitlik (14) ile hesaplanan ağırlıklandırılmış ortalama site puanları) arasındaki korelasyona eşittir. Site puanları $\{x_i\}$, Eşitlik (13) ya da Eşitlik (17) ile hesaplanan çevresel değişkenlerin doğrusal kombinasyonudur.

C) Kanonik uyum analizi triplotları ve yorumu

KUA, cevap değişkeninin ve sitelerin (lokasyonların) noktalarla, çevresel değişkenlerin ise oklarla temsil edildiği triplot olarak bilinen bir ordinasyon diyagramı oluşturur (Şekil 7). Site ve cevap değişkeni noktaları UA'dakine benzer şekilde yorumlanır. Noktalar sitelerde bulunan cevap değişkeni kompozisyonundaki varyasyonları gösterirken, çevresel değişkenler oklarla temsil edilir (Ter Braak, 1990). (Ter Braak, 1986; Palmer, 1993). Diyagramda cevap değişkenlerinin, çevresel değişkenlerin doğrusal kombinasyonlarına göre unimodal cevap yüzeylerine sahip olduğu varsayıldığından cevap değişkenleri, iki boyutlu çevresel alt uzaydaki yaklaşık optimumlarına karşılık gelen noktalarla temsil edilir (Şekil 7) (Ter Braak, 1986).



Şekil 7. KUA triplotunun şematik gösterimi. Dolu daireler, nesnelere (örneklem sitelerini); boş daireler, cevap değişkenlerini (bollukları); oklar, kantitatif açıklayıcı değişkenleri (besin konsantrasyonları) ve dolu üçgenler ise kategorik açıklayıcı değişkenleri (kum, kil, tortu tipi vb.) göstermektedir (Anonymous 3, 2018).

UA'da olduğu gibi, KUA'nın unimodal dağılımları gösteren cevap değişkenlerinin analizi için kullanılması önerilmekte ve KUA'ya ait diyagramlarda nesnelere arasındaki ki-kare uzaklıkları korunmaktadır. Bu diyagramlarda, açıklayıcı değişkenlerin doğrusal kombinasyonları ile en çok ilişkili olan cevap değişkenleri bir Öklid uzayında düzenlenir (Legendre ve Legendre, 1998).

Her çevresel değişken için ters yönlü ancak aynı uzunlukta bir vektör (ok) daha bulunmakta olup, bu ters yönlü vektör ilgili çevresel değişkenin azalan etkisini göstermektedir (Ter Braak, 1986).

Cevap değişkeni ve site noktaları, çevresel değişkenler tarafından açıklanabiliyorsa ve çevresel değişkenlerin okları, cevap değişkenlerinin dağılımlarını ortak olarak yansıtıyorsa, bu noktalar, nesne kompozisyonundaki baskın olan modeli ortak olarak temsil eder (Ter Braak, 1986).

Bir KUA diyagramı, açıklayıcı (çevresel) değişkenler tarafından açıklanan topluluk kompozisyonundaki varyasyonun temel modeli hakkında bilgi verir ve aynı zamanda her bir çevresel değişken boyunca cevap değişkeninin dağılımını gösterir (Ter Braak, 1986).

Herhangi bir diyagramın, sitelerin yaklaşık topluluk kompozisyonunu temsil edebilmesi için bu diyagramın, Hill'in (1979) ölçeklemesinde olduğu gibi, ağırlıklı

ortalama cevap deęişkenine ve site puanlarına göre çizilmesi gerekir. Her bir site noktası, o bölgede meydana gelen cevap deęişkeni, noktalarının merkezinde yer aldığından, cevap deęişkeninin belirli bir alanda bulunma olasılığı diyagramdan çıkarılabilir. Ayrıca KUA, Gauss cevap yüzeylerinin uyumu için iyi bir yaklaşım olduğunda, cevap deęişkeni noktaları bu yüzeylerin yaklaşık olarak optimum değerlerini gösterir (Ter Braak, 1986).

Bir cevap deęişkeninin çevresel deęişkenin hangi değerinde meydana geldięi, ağırlıklı ortalama yöntemi ile uygun bir şekilde özetlenebilir. Bir çevresel deęişken (j)'e göre bir cevap deęişkeninin dağılımının ağırlıklı ortalaması (k), o cevap deęişkeninin meydana geldięi bölgelerdeki çevresel deęişkenlerin değerlerinin ortalaması olarak tanımlanır (Eşitlik (19)).

$$\bar{z}_{kj} = \sum_{i=1}^n y_{ik} z_{ij} / y_{+k} \quad (19)$$

Ağırlıklı Ortalama, çevresel deęişken boyunca bir cevap deęişkeninin dağılımının "merkezini" gösterir (Ter Braak ve Looman, 1986). Cevap deęişkenlerinin ağırlıklı ortalamaları arasındaki farklılıklar, çevresel deęişken boyunca cevap deęişkeninin dağılımındaki farklılıkları göstermektedir.

D) KUA'da ölçeklendirme tipleri

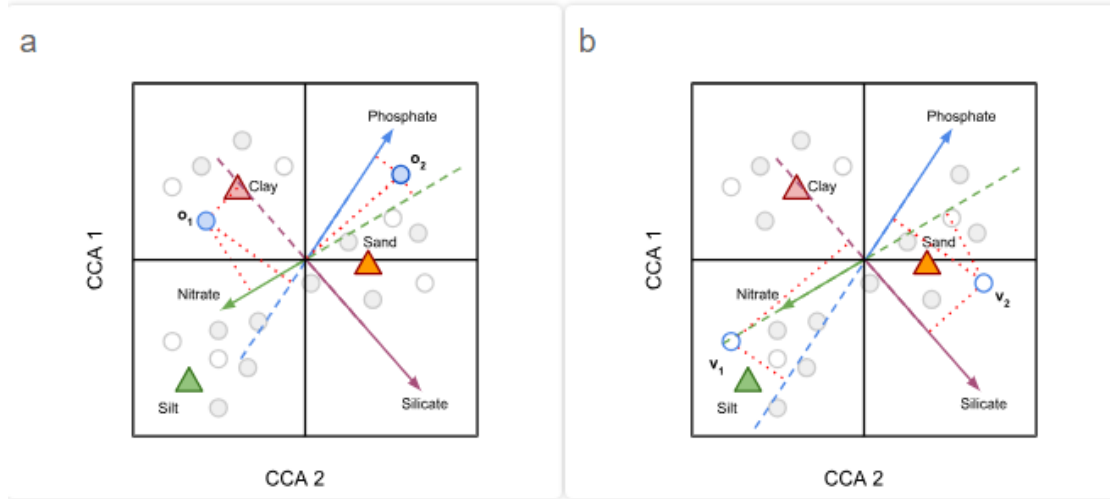
Tip 1 Ölçeklemesi:

Tip 1 ölçeklendirmede, *nesnelere arasındaki ilişkiler vurgulanmaktadır* (Şekil 8 (a)). Bu ölçeklendirmede nesnelere, cevap deęişkenlerinin merkezleri olarak temsil edilir ve nesne noktaları arasındaki uzaklıklar χ^2 uzaklıklarını gösterir.

Tip 2 Ölçeklemesi:

Tip 2 ölçeklendirmede, *cevap deęişkenleri arasındaki ilişkiler vurgulanmaktadır* (Şekil 8 (b)). Bu ölçeklendirmede cevap deęişkenleri, nesnelere merkezleri olarak temsil edilir ve cevap deęişkeni noktaları arasındaki uzaklıklar χ^2 uzaklıklarını gösterir.

Açıklayıcı değişkeni temsil eden bir vektör (ok) üzerinde, bir cevap değişkenini temsil eden noktanın izdüşümü, bu açıklayıcı değişken boyunca cevap değişkeninin maksimum değerinin (optimum değerinin) konumunu gösterir.



Şekil 8. KUA triplotunda Tip I (a) ve Tip II (b) ölçeklendirme. **a)** "O₁", "O₂" iki nesneyi; oklar (oklar çift yönlü olarak, ok'un başı yönünde ve ters yönde kesikli çizgiler boyunca uzatılabilir) üç kantitatif açıklayıcı değişkeni ("Nitrat", "Fosfat", "Silikat"); üçgenler ise üç nominal (kalitatif) değişkeni (Sand", "Silt", "Clay") temsil etmektedir. İzdüşümler noktali kırmızı çizgilerle gösterilmiştir. "O₁" nesnesinin kil çökeltilerinde (Clay) bulunma olasılığı daha yüksektir. "O₂" nesnesinin ise kum çökeltilerinde bulunma olasılığı daha yüksektir. Ayrıca, "O₁" nesnesinin kantitatif açıklayıcı değişkenler üzerine (oklar üzerine) iz düşümleri: nitrat konsantrasyonunun yüksek değerlerinde, fosfat konsantrasyonunun orta ila düşük değerlerinde ve silikat konsantrasyonunun düşük değerlerinde "O₁" nesnesinin var olması daha muhtemeldir yorumları yapılabilir. Benzer şekilde "O₂" nesnesi: fosfat konsantrasyonunun yüksek değerlerinde, silikat konsantrasyonunun orta değerlerinde ve nitrat konsantrasyonunun düşük değerlerinde daha çok bulunmaktadır. **b)** (a) 'dakiye benzer şekilde yorumlanır. Ancak, burada cevap değişkenlerini ("V₁", "V₂") temsil eden noktalar yorumlanır (Anonymous 3, 2018).

E) KUA'nın varsayımları

KUA, doğrusal dış kısıtlamalara sahip maksimum olabilirlik Gaussian Ordinasyon yaklaşımı olarak türetilmiştir (Ter Braak, 1986, 1988a). Bu yöntemde, uygulamalarda doğruluğu test edilemeyen güçlü varsayımlar kullanılmıştır. Ancak, KUA bu varsayımlardan sapmaya karşı sağlam (robust) bir yöntemdir (Ter Braak ve Verdonschot, 1995).

3. 2. 4. Kanonik ordinasyonda anlamlılık testleri

Standart Ordinasyon Yöntemlerinin (Temel Bileşenler Analizi, Temel Koordinatlar Analizi, Faktör Analizi, Uyum Analizi) aksine, Y ve X matrisleri arasındaki ilişki hakkında sıfır hipotezi formüle edilebildiği için, permütasyon yöntemi kullanılarak kanonik analizde özdeğerler test edilebilir. Yöntem, cevap (Y) ve açıklayıcı (X) değişken kümeleri arasında ilişki yapısı hakkında açık hipotezler olduğunda kullanılabilir. Hipotezler, belirli nicel açıklayıcı değişkenleri ya da nicel ve binary değişkenlerin karışımlarını X matrisine dönüştürür (Legendre ve Legendre, 1998). Hipotezler, tek bir sınıflandırma kriteri (faktör) ya da her biri yapay (dummy) değişkenler olarak kaydedilen çeşitli faktörler ve bunların interaksiyonlarını içeren varyans analizi türünde olabilir. Çok değişkenli hipotezleri test etmek için birçok RDA örnekleri uygulanmıştır (Sabatier ve ark., 1989; Ter Braak ve Wiertz, 1994; Verdonschot ve Ter Braak, 1994; Legendre ve Anderson, 1999).

RDA ve KUA'da anlamlılık testleri iki aşamada gerçekleştirilir:

1. Aşama: Birinci aşama, Y ve X değişken kümeleri arasındaki bağımsızlığın (H_0) testidir.

Tüm kanonik özdeğerlerin toplamı, F oranı kullanılarak test edilir. Tüm kanonik özdeğerlerin toplamı olan alternatif hipotez (H_1); ham verilerin permütasyonu ya da artıkların permütasyonu (aşağıda ifade edilmiştir) kullanılarak, Y'deki varyasyonun açıklanmasını test eder. Bu yaklaşım, Y ve X arasındaki ilişkinin genel bir testidir. Y matrisindeki toplam varyasyona bölünen tüm kanonik özdeğerlerin toplamı, çoklu regresyondaki R^2 'ye benzer şekilde, Y'deki varyasyonun X ile açıklanan kısmını belirtir (Legendre ve Legendre, 1998).

2. Aşama: İkinci aşama kanonik özdeğerlerin teker teker anlamlılığının testidir.

Bu aşamada, belirli matrislerle elde edilebilecek olandan daha fazla kanonik varyasyonu açıklayan (bireysel olarak açıklayan) eksenler tespit edilmeye çalışılır. Bu testin ardından, sadece ordinasyon diyagramlarındaki önemli eksenlerin kullanılmasına karar verilebilir. Sıfır hipotezi (H_0), Y ile X arasındaki bağımsızlığı ifade ederken; alternatif hipotez (H_1), ham verilerin permütasyonu ya da artıkların permütasyonu

kullanılarak, Y'deki varyasyonun açıklandığını ifade eder (Legendre ve Legendre, 1998).

Birçok istatistik paket programında, yalnızca ilk kanonik özdeğer anlamlılık açısından test edilebilir. İkinci kanonik özdeğeri test etmek için, birinci özdeğerin etkisi kontrol altında tutulmalıdır. Bu işlem için ilk kanonik eksen, permütasyon testi ve hesaplamaları tekrarlamadan önce analizden kaldırılır ve bir kovaryant'a dönüştürülerek gerçekleştirilir. Bu işlem, her kanonik özdeğer test edilmek istendiğinde tekrarlanabilir (Legendre ve Legendre, 1998).

3. 3. Yöntemlerin Performans Değerlendirmesi (Model Performance Evaluation)

Çalışmada, yöntemlerin tahmin doğruluğunu veya performansını belirlemek üzere kullanılacak ölçütlerden, R^2 (belirleme veya determinasyon katsayısı) ve MAPE (Ortalama Mutlak Yüzde Hata, Mean Absolute Percent Error, MAPE) kullanılmıştır (Anonymous 4, 2019). MAPE,

$$MAPE = \frac{100 (\%)}{n} \sum \left| \frac{y - \hat{y}}{y} \right| \quad (20)$$

eşitliği kullanılarak hesaplanmıştır. Eşitlik (20)'de; y gerçek değeri, \hat{y} tahmin değerini ve n gözlem sayısını ifade etmektedir. MAPE, bir tahmin sisteminin ne kadar doğru olduğunun istatistik ölçüsüdür ve bu doğruluğu yüzde olarak ölçer. Bu ölçüt, mutlak değer kullanımı sayesinde aykırı değerlerin etkilerine karşı dayanıklıdır (Ünver, 1996). Verilerde aşırı değerler yoksa (ayrıca sıfır değerler yoksa) bu ölçütün iyi sonucu veren hata kriteri olduğu belirtilmektedir (Anonymous 5, 2019).

Witt ve Witt (1992), MAPE değerleri %10'un altında olan tahmin modellerini "yüksek doğruluk" derecesine sahip, %10 ile %20 arasında olan modelleri ise doğru tahminler olarak sınıflandırmıştır. Benzer şekilde Lewis (1982), MAPE değeri %10'un altında olan modelleri "çok iyi", %10 ile %20 arasında olan modelleri "iyi", %20 ile %50 arasında olan modelleri "kabul edilebilir" ve %50'nin üzerinde olan modelleri ise "yanlış ve hatalı" olarak sınıflandırmıştır.

Yöntemlere ait belirleme (determinasyon) katsayısı (R^2) aşağıdaki eşitlikle hesaplanmıştır (Anonymous 5, 2019).

$$R^2 = 1 - \frac{SS \text{ Error}}{SS \text{ Total}} = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2} \quad (21)$$

Eşitlik (21)'de SS Error, hatalar kareleri toplamını, SS Total, genel kareler toplamını, y_i , i. cevap değişkeninin gerçek değerini, \hat{y}_i , i. cevap değişkeninin tahmin değerini ve \bar{y} , gerçek verilerin ortalamasını göstermektedir.

R^2 , 0 ve 1 arasında değer alır. Bu değer 1'e (% 100'e) yakın olması cevap değişkenindeki varyansın büyük bir kısmının modeldeki açıklayıcı değişkenler tarafından açıklandığını gösterir (Ünver, 1996).

Çalışmada ele alınan değişkenler arasındaki ilişki yapısı önce Kanonik Uyum Analizi ile sonra Gereksizlik Analizi ile incelenmiştir.

KUA yöntemiyle, cevap değişkeni (lokasyon, site sayısı), nesne (kategorik değişkenler) ve açıklayıcı değişkenlere (sürekli değişken) göre bir uygulama yapılmıştır (Tablo 4).

Tablo 4. Cevap değişkeni (site), nesne ve açıklayıcı değişkenlere göre KUA uygulaması

Cevap Değişkeni (Lokasyon Sayısı)	Nesneler (Kategorik değişkenler)	Açıklayıcı değişkenler (Sürekli değişkenler)
ADS (4 kategorili)	8 Kategorik Değişken (GAT, S-S, D-EKG, Cinsiyet, ESA, Diyabet, ST-D, ST-SE)	2 Sürekli Değişken (MKAS ve Kolesterol)

ST-D: ST segmentinin çökme durumu, GAT: Göğüsteki ağrı türü, ST-SE: ST segmentinin eğimi, ESA: Egzersiz sırasında angina varlığı, MKAS: Maksimum kalp atış sayısı, D-EKG: Dinlenme sırasında EKG sonucu, ADS: Açık ana damar sayısı, S-S: Sintigrafi sonucu.

RDA yöntemiyle, site sayısı (kategorik değişken), cevap değişkeni (sürekli değişken) sayısı ve açıklayıcı değişkenlere göre bir uygulama yapılmıştır (Tablo 5).

Tablo 5. Site sayısı, cevap deęiřkeni ve aıklayıcı deęiřkenlere gre RDA uygulaması

Site sayısı (cevap deęiřkeninin ortaya ıktıęı blgeler)	Aıklayıcı deęiřkenler	Cevap deęiřkenleri (srekli deęiřkenler)
ADS (4 kategorili)	8 Kategorik Deęiřken (GAT, S-S, D-EKG, Cinsiyet, ESA, Diyabet, ST-D, ST-SE) ve 2 Srekli Deęiřken (DSKB, Yař)	2 Srekli Deęiřken (MKAS, Kolesterol)

ST-D: ST segmentinin okme durumu, GAT: Ggsteki aęrı tr, ST-SE: ST segmentinin eęimi, ESA: Egzersiz sırasında angina varlıęı, MKAS: Maksimum kalp atıř sayısı, D-EKG: Dinlenme sırasında EKG sonucu, ADS: Aık ana damar sayısı, S-S: Sintigrafi sonucu, DSKB: Dinlenme sırasında kan basıncı.

Elde edilen verinin analizinde kullanılan, Kanonik Uyum Analizi ve Gereksizlik Analizi Yntemlerinin uygulamaları iin XLSTAT (version: 2018. 5) paket programı kullanılmıřtır.

4. BULGULAR

4.1. Kanonik Uyum Analizi (KUA) Uygulaması

Cevap değişkeni, açıklayıcı değişkenler ve nesnelere göre Kanonik Uyum Analizi yapılmış ve sonuçları Tablo 6'da özetlenmiştir.

Tablo 6. Cevap değişkeni (site), nesne ve açıklayıcı değişkenlere göre KUA uygulama sonucu

Cevap Değişkeni / site sayısı	Nesneler (Kategorik değişkenler)	Açıklayıcı değişkenler (Sürekli değişkenler)	1. Boyutun özdeğeri / açıkladığı varyans oranı (%)	2. Boyutun özdeğeri / açıkladığı varyans oranı (%)	Kısıtlanmış eksenlerin açıkladığı toplam varyans (%)
ADS (4 kategorili)	8 Kategorik Değişken (GAT, S-S, D-EKG, Cinsiyet, ESA, Diyabet, ST-D, ST-SE)	2 Sürekli Değişken (MKAS, Kolesterol)	0.036 / 86.934	0.005 / 13.066	100.00

ST-D: ST segmentinin çökme durumu, GAT: Göğüsteki ağrı türü, ST-SE: ST segmentinin eğimi, ESA: Egzersiz sırasında angina varlığı, MKAS: Maksimum kalp atış sayısı, D-EKG: Dinlenme sırasında EKG sonucu, ADS: Açık ana damar sayısı, S-S: Sintigrafi sonucu.

KUA'nın ordinasyon şeması lokasyonları (siteleri), nesnelere (türleri) ve açıklayıcı (çevresel) değişkenleri göstermektedir.

Cevap değişkeni ile açıklayıcı değişkenler arasındaki ilişki modeli permütasyon testi yardımıyla test edilmiştir. Bu amaçla, pseudo (sahte) F değeri elde edilmiş ve 500 tekrarlı permütasyon testi sonucu elde edilen kritik değere göre:

H_0 : Cevap ve açıklayıcı değişkenler arasında doğrusal bir ilişki yoktur.

H_1 : Cevap ve açıklayıcı değişkenler arasında doğrusal bir ilişki vardır.

hipotezi test edilmiştir. Buna göre elde edilen sonuç Tablo 7'de verilmiştir.

Tablo 7. KUA uygulamasında permütasyon testi sonucu

Permütasyon testi sonucu	
Permütasyon sayısı	500
Pseudo F (sahte F değeri)	1168.000
p-değeri	0.748
alpha değeri	0.050

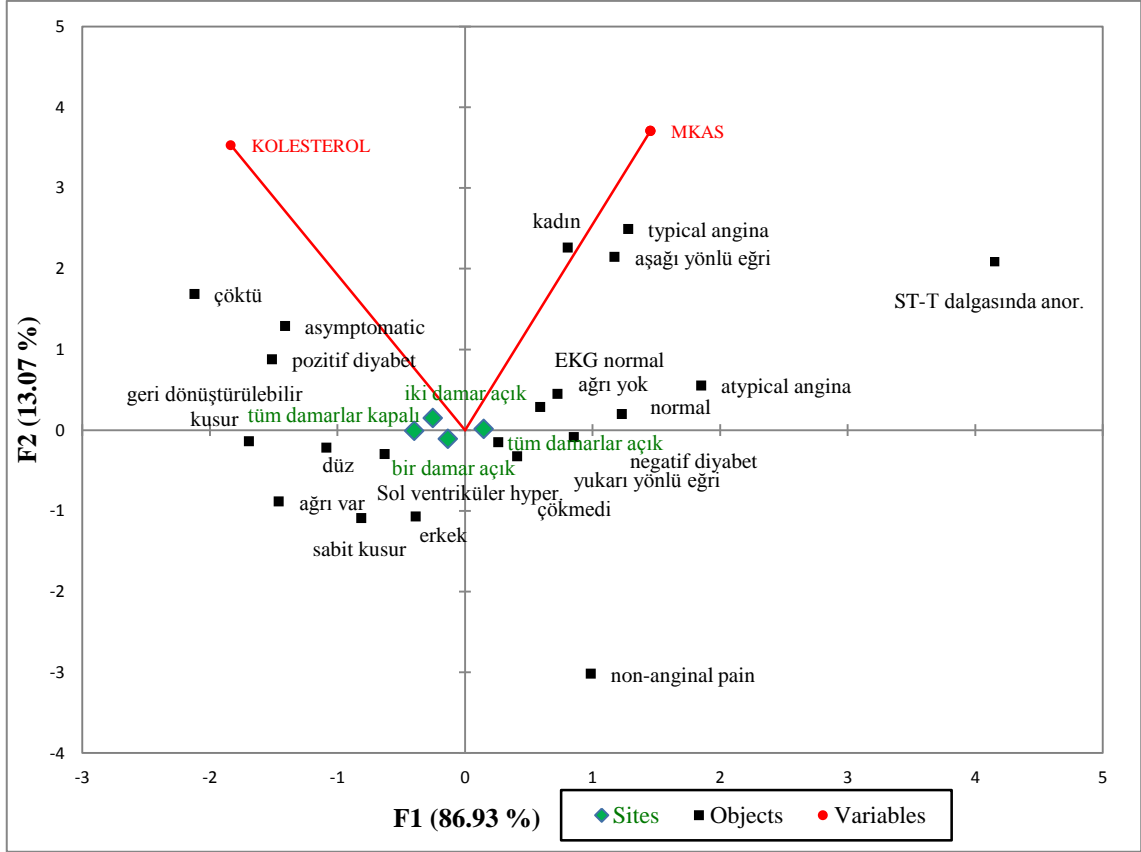
Permütasyon testi sonucunda elde edilen değer $0.748 > 0.05$ olduğundan H_0 hipotezi kabul edilir. Buna göre, KUA uygulaması sonucunda cevap değişkeni ile açıklayıcı değişkenler arasında doğrusal bir ilişki yoktur (unimodal ilişki vardır).

Analizde, eksenler açıklayıcı değişkenlerin doğrusal kombinasyonu olarak kısıtlanmış olduğundan, değişkenlere ait kısıtlı varyansın büyük bir kısmının (% 93.10) birinci eksen tarafından açıklandığı görülmektedir (Tablo 8). İkinci eksenle birlikte, toplam varyansın % 100'ü açıklanabilmektedir. Eksenlerin toplam varyasyonu açıklama oranına göre, iki boyutlu Kanonik Uyum Analizi diyagramının, cevap değişkeni, nesnelere ve açıklayıcı değişkenler arasındaki ilişkileri analiz etmede yeterli olabileceği söylenebilir. Özdeğerlerin, eksenlere göre varyansı açıklama oranları Tablo 8'de verilmiştir.

Tablo 8. KUA uygulamasında temel eksenlerin özdeğerleri ve varyans açıklama oranları (%)

	F1	F2
Özdeğer	0.036	0.005
Kısıtlanmış inertia (%)	86.934	13.066
Kümülatif (%)	86.934	100.000
Toplam inertia (kısıtlanmamış varyans) (%)	82.716	12.432
Kümülatif (%)	82.716	95.148

Cevap değişkeni, nesnelere ve açıklayıcı değişkenlerin, kısıtlı KUA ordinasyonunda elde edilen iki boyutlu uzaydaki triplot gösterimi Şekil 9'da verilmiştir.



Şekil 9. KUA uygulamasının triplot şeması

Şekil 9 incelendiğinde; açıklayıcı değişkenlerin uzunluğuna bakılarak, kolesterol değişkeninin, maksimum kalp atış sayısı açıklayıcı değişkeninden daha önemli olduğu söylenebilir. “Üç damarı açık” bireylerin düşük kolesterol değerine, “iki ve bir damarı açık” bireyler ise daha yüksek kolesterol değerine sahip oldukları görülmektedir. Tipik göğüs ağrısı, yüksek maksimum kalp atış sayısında, asimptomatik göğüs ağrısı ise orta ve yüksek kolesterol seviyesinde görülmüştür. Ayrıca, “atipik angina” göğüs ağrısı değişkeninin, ST segmenti eğrisinin “yukarı yönlü” olması ve EKG sonuçlarının “normal” olması değişkenleriyle yüksek ilişkili olduğu bulunmuştur. Bunun yanında, egzersiz sırasında “ağrısı olan” bireylerin, orta seviyede kolesterol değerine sahip olduğu söylenebilir. ST segment eğrisi “yukarı yönlü” olan bireylerle ve egzersiz sırasında “ağrısı olmayan” bireylerin orta seviyelerde kalp atış sayısına sahip olduğu görülmüştür.

Ayrıca; “Sol ventriküler hypertrophy belirtisi”, ST segmentinin eğiminin “düz” olması ve sintigrafi sonucunun “sabit kusurlu” olması değişkenleri arasında yüksek

korelasyon olduđu söylenebilir. ST segmentinin “çökme” durumu, orta ve yüksek kolesterol değerlerinde daha çok görülmüştür. Bunun yanısıra, ST segmentinde “anormallik”, kategorisi yüksek maksimum kalp atış seviyesinde daha yoğun görülmüştür. “Non-anginal ağrı” değişkeninin kolesterol ve maksimum kalp atış sayısı açıklayıcı değişkenlerinin düşük değerlerinde daha yoğun ortaya çıktığı saptanmıştır. Bununla beraber, bu değişkenin varyasyona en düşük katkı yapan değişken olduğu söylenebilir. Sintigrafi sonucunun “geri dönderilebilir kusur” olması kategorisi ile “sabit kusur” kategorisi arasında yüksek korelasyon bulunmuştur.

ST segmentinin eğimi “aşağı yönlü” olan bireylerde genellikle, “sol ventriküler hypertrophy belirtisi” olduğu, “asimptomatik” göğüs ağrısı yaşadıkları, ST segmentinde “anormallik” olduğu ancak, egzersiz sırasında “ağrı yaşamadıkları” dikkat çekici bulunmuştur. Ayrıca, “kadın” bireylerde ST segmenti eğiminin daha çok “aşağı yönlü” olduğu ve bu bireylerin daha çok “tipik angina” göğüs ağrısı yaşadığı söylenebilir. Bunun yanısıra, sintigrafi sonucu “sabit kusur” ve “geri döndürülebilir kusur” olan, “asimptomatik” göğüs ağrısı yaşayan, “diyabet hastası olan” ve egzersiz sırasında “ağrı yaşayan” bireylerin orta ve düşük düzeyde kolesterol değerine sahip olduğu saptanmıştır. Ayrıca, “tüm damarları açık” bireylerde, diyabet az görülmüş ve ST segmentinde çökme saptanmamıştır.

Orta ve düşük seviyede maksimum kalp atış sayısı ve kolesterol düzeyine sahip olan bireylerde, “sol ventriküler hypertrophy belirtisi” görülmüş ve bu düzeylerde “diyabet varlığına” daha sık rastlanmıştır. Orta ve düşük seviyede kalp atış sayısına sahip bireylerde daha az diyabet görülmüştür. Bunun yanısıra, kolesterol açıklayıcı değişkeninin; ST segmentinin “çökmesi”, “pozitif diyabet” ve “asimptomatik” göğüs ağrısı kategorileri ile yüksek pozitif ilişkisi bulunmuştur. Ayrıca, maksimum kalp atış sayısı açıklayıcı değişkeninin; “kadın”, “tipik angina” göğüs ağrısı ve ST segmentinin eğimi “aşağı yönlü” kategorileri ile yüksek pozitif ilişkisi görülmüştür. “yukarı yönlü” ST segmentine sahip bireylerde ST segmentinin daha az “çöktüğü” söylenebilir. Bunun yanısıra, ST segmenti “çökmeyen” bireylerin düşük kalp atış sayısına sahip olduğu görülmüştür. Ayrıca, ST segmenti “çökmeyen” bireylerin, ST segmentinin “yukarı yönlü” ve sintigrafi sonucunun ise “normal” olduğu bulunmuştur.

“Aşağı yönlü” ST segmentine sahip bireylerin genellikle yüksek kalp atış sayısına sahip olduğu görülmektedir. EKG sonuçlarının, maksimum kalp atış sayısı açıklayıcı değişkeni ile yüksek ilişkisi görülmüştür. “Tipik-angina” ve “atipik angina” göğüs ağrısı türlerinin ise genellikle maksimum kalp atış sayısı ile daha yüksek ilişkili olduğu belirlenmiştir.

4. 2. Gereksizlik Analizi (RDA) Uygulaması

Site sayısı, cevap değişkenleri (sürekli değişken) ve açıklayıcı değişkenlere göre RDA uygulaması yapılmış ve uygulama sonucu Tablo 9 ‘da özetlenmiştir.

Tablo 9. Site sayısı, cevap ve açıklayıcı değişkenlere göre RDA uygulama sonucu

Site sayısı (cevap değişkeninin ortaya çıktığı bölgeler)	Açıklayıcı değişkenler	Cevap değişkenleri (sürekli değişkenler)	1. Boyutun özdeğeri / varyans açıklama oranı (%)	2. Boyutun özdeğeri / varyans açıklama oranı (%)	Kısıtlanmış eksenlerin açıkladığı toplam varyans (%)
ADS (4 kategorili)	8 Kategorik Değişken (GAT, S-S, D-EKG, Cinsiyet, ESA, Diyabet, ST-D, ST-SE) ve 2 Sürekli Değişken (DSKB, Yaş)	2 Sürekli Değişken (MKAS, Kolesterol)	119.993 / 77.493	34.851 / 22.507	100.00

ST-D: ST segmentinin çökme durumu, GAT: Göğüsteki ağrı türü, DSKB: Dinlenme sırasında kan basıncı, ST-SE: ST segmentinin eğimi, ESA: Egzersiz sırasında angina varlığı, MKAS: Maksimum kalp atış sayısı, D-EKG: Dinlenme sırasında EKG sonucu, ADS: Açık ana damar sayısı, S-S: Sintigrafi sonucu.

Cevap değişkeni ile açıklayıcı değişkenler arasındaki ilişki modeli permütasyon testi yardımıyla test edilmiştir. Bu amaçla, pseudo (sahte) F değeri elde edilmiş ve 500 tekrarlı permütasyon testi sonucu elde edilen kritik değere göre:

H_0 : Cevap ve açıklayıcı değişkenler arasında doğrusal bir ilişki yoktur.

H_1 : Cevap ve açıklayıcı değişkenler arasında doğrusal bir ilişki vardır.

hipotezi test edilmiştir. Permütasyon sayısı, mümkün olan minimum önem değerini belirlemektedir. Buna göre elde edilen sonuç Tablo 10’da verilmiştir.

Tablo 10. RDA uygulamasında permütasyon testi sonucu

Permütasyon testi sonucu	
Permütasyon sayısı	500
Pseudo F (sahte F değeri)	9.805
p-değeri	0.0001
alpha (α) değeri	0.050

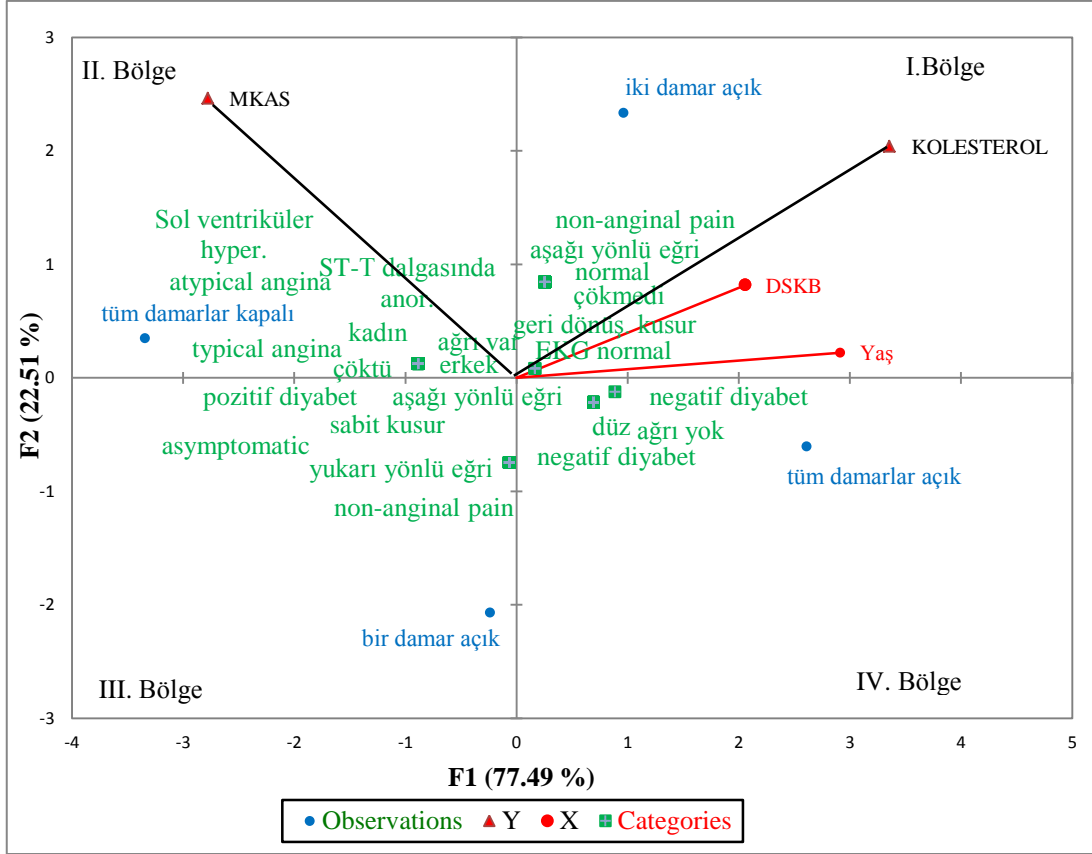
Permütasyon testi sonucunda elde edilen değer $0.0001 < 0.05$ olduğundan H_0 hipotezi ret edilir. Buna göre, RDA sonucunda cevap değişkeni ile açıklayıcı değişkenler arasında doğrusal (lineer) bir ilişki vardır.

Analizde, özdeğerlerin, eksenlere göre varyansı açıklama oranları Tablo 11’de verilmiştir. Eksenler, açıklayıcı değişkenlerin doğrusal kombinasyonu olarak kısıtlanmış olduğundan, cevap değişkenlerine ait açıklanan varyansın % 77.493’ünün birinci eksenle açıklandığı görülmektedir (Tablo 11). İkinci eksenle birlikte, toplam varyansın % 100’ü açıklanmıştır. Eksenlerin toplam varyasyonu açıkladığı bu değerden; iki boyutlu Gereksizlik Analizi diyagramının konumlar (cevap değişkeni ile birlikte değişkenlerin ortaya çıktığı bölgeler), cevap değişkenleri ve açıklayıcı değişkenler arasındaki ilişkileri belirtmek için yeterli olacağı söylenebilir.

Tablo 11. RDA uygulamasında temel eksenlerin özdeğer ve varyans açıklama oranları (%)

	F1	F2
Özdeğer	0.355	0.187
Kısıtlanmış inertia (%)	77.493	22.507
Kümülatif (%)	94.997	100.000
Toplam inertia (kısıtlanmamış varyans) (%)	94.997	5.003
Kümülatif (%)	94.997	100.000

Konumların (lokasyonların), cevap ve açıklayıcı değişkenlerin, kısıtlı RDA ordinasyonunda elde edilen iki boyutlu uzaydaki triplot gösterimi Şekil 10’da verilmiştir.



Şekil 10. RDA uygulamasının triplot şeması

Şekil 10 incelendiğinde; açıklayıcı değişkenlerin uzunluğuna bakılarak, “dinlenme sırasında kan basıncı” değişkeninin, “yaş” açıklayıcı değişkeninden daha önemli olduğu söylenebilir. “Kolesterol” cevap değişkeninin, “yaş” açıklayıcı değişkenine göre “dinlenme sırasında kan basıncı” açıklayıcı değişkeni ile daha yüksek pozitif ilişkili olduğu görülmüştür. “Kalp atış sayısı” cevap değişkeninin; “dinlenme sırasında kan basıncı” açıklayıcı değişkenine göre “yaş” değişkeni ile daha yüksek negatif ilişkili olduğu görülmüştür.

“Non-angina” ile “asimptomatik” göğüs ağrısı yaşayan bireylerin, orta ve düşük seviyede kolesterol ve kalp atışında daha yoğun olduğu görülürken “tipik ve atipik” angina göğüs ağrısı yaşayan bireylerin orta seviyede kalp atışında daha yoğun olduğu görülmüştür.

Ayrıca, “non-angina” göğüs ağrısı yaşayan bireylerde, “dinlenme sırasındaki kan basıncı” değişkeni cevap değişkeni için etkili bulunurken bu bireylerin, “dinlenme

sirasında kan basıncı” ve “yaş” açıklayıcı değişkenleri ile pozitif ilişkili olduğu bulunmuştur. Ayrıca, “dinlenme sırasında kan basıncı” açıklayıcı değişkeninin, birinci bölgedeki kategorik açıklayıcı değişkenlerle pozitif ilişkili olduğu söylenebilir.

“Maksimum kalp atış sayısı” cevap değişkeni ile “kolesterol” cevap değişkeni arasında düşük ilişki belirlenmiştir. Ayrıca, “dinlenme sırasında kan basıncı” ve “yaş” sürekli açıklayıcı değişkenlerinin birinci bölgedeki kategorik açıklayıcı değişkenlerle pozitif ilişkisi görülürken, “kolesterol” cevap değişkeninin birinci bölgedeki açıklayıcı değişkenlerle yüksek ilişkili olduğu söylenebilir. “Yaş” açıklayıcı değişkeni ile “dinlenme sırasında kan basıncı” açıklayıcı değişkenleri arasında yüksek oranda pozitif korelasyon görülmüştür. “Asimptotik” göğüs ağrı kategorisinin, orta ve düşük yaşa sahip bireylerde daha yoğun görüldüğü ve üçüncü bölgedeki kategorik açıklayıcı değişkenlerle yüksek ilişkili olduğu saptanmıştır. Üçüncü bölgedeki kategorik açıklayıcı değişkenlerin “bir damarı açık” bireylerde daha yoğun görüldüğü belirlenmiştir.

Bunun yanısıra, “iki damarı açık bireylerin” orta ve yüksek “kolesterol” ve “kan basıncı” seviyelerinde daha yoğun olduğu saptanmıştır. Ayrıca, “iki damarı açık” bireylerin “kolesterol, kan basıncı ve yaş” değişkenleri ile yüksek oranda pozitif ilişkili olduğu görülmüştür. Bunun yanısıra, “tüm damarları kapalı” bireylerin “maksimum kalp atış sayısı” cevap değişkeni ile ve ikinci bölgedeki kategorik açıklayıcı değişkenlerle yüksek ilişkisi bulunmuştur. “Tüm damarları kapalı” bireylerin ikinci bölgedeki kategorik açıklayıcı değişkenlerle yüksek oranda pozitif ilişkili olduğu bulunmuştur. Bunun yanısıra, “dinlenme sırasında kan basıncı” ve “yaş” açıklayıcı değişkenleri ile “Tüm damarları kapalı” bireylerin negatif ilişkili olduğu söylenebilir.

“Tüm damarları açık” bireylerle dördüncü bölgedeki kategorik açıklayıcı değişkenler arasında pozitif korelasyon bulunmuş ve bu bireylerin genellikle düşük kalp atışına sahip olduğu, ancak orta seviyede “kolesterol” değerlerinde daha yoğun olduğu söylenebilir. Ayrıca, “tüm damarları açık” bireylerin sürekli açıklayıcı değişkenlerle yüksek ilişkisi görülmemiştir. “Maksimum kalp atış sayısı” cevap değişkeninin orta seviyelerinde “sol ventriküler hipertrofi” belirtisi olan ve “ST dalgasında anormallik” olan bireylerin daha yoğun bulunduğu görülmüştür. Ayrıca, ST segmenti eğimi “aşağı yönlü” olan bireylerin, sintigrafi sonucunun “geri dönüştürülebilir kusur” değişkeni ile yüksek oranda pozitif ilişkili olduğu söylenebilir.

Bunun yanısıra, “bir damarı açık” bireylerin dördüncü bölgedeki kategorik açıklayıcı değişkenlerle yüksek ilişkili olduğu görülmüştür. Ancak, “bir damarı açık” bireylerin genellikle kolesterol’un ve kalp atış sayısının düşük seviyelerinde daha yoğun olduğu görülmüştür. Ayrıca, “bir damarı açık” bireylerin “dinlenme sırasında kan basıncı” ve “yaş” sürekli açıklayıcı değişkenleri ile negatif ilişkili olduğu söylenebilir. “Bir damarı açık” kategorinin varyasyona en az katkı sağladığı görülmüş ve mevcut değişkenlerle bu kategorinin çok iyi açıklanamadığı görülmüştür.

Yöntemler için Ortalama Mutlak Yüzde Hata (Mean Absolute Percent Error-MAPE) ve R^2 değerleri Tablo12’ de verilmiştir.

Tablo 12. Kısıtlı ordinasyon yöntemleri için değerlendirme ölçütleri

Yöntem	R^2 (Determinasyon/Belirleme Katsayısı)	MAPE (Ortalama Mutlak Yüzde Hata)	n (Gözlem Sayısı)
CCA	1	0.04	270
RDA	0.94	0.06	270

Tablo 12’de görüldüğü üzere yöntemlerinin varyans açıklama oranlarının yüksek olduğu söylenebilir. KUA ordinasyon diyagramında, açıklayıcı değişkenlerin toplam varyasyonu açıklama oranı % 100 olarak bulunurken, RDA ordinasyonunda bu oran % 94 olarak bulunmuştur. Belirleme katsayısına göre cevap değişkeni, nesnelere ve açıklayıcı değişkenler arasındaki farklı ilişki yapılarını analiz etmek için kısıtlı ordinasyon yöntemlerinin varyasyonu açıklamada yeterli olduğu söylenebilir.

Çalışmada, KUA’nın MAPE değeri 0.04 (% 4) olarak hesaplanırken RDA’nın MAPE değeri ise 0.06 (% 6) olarak hesaplanmıştır. Buna göre, KUA yönteminde elde edilen kısıtlı ordinasyon diyagramında, değişkenler arası farklı ilişki yapılarını belirlemede yapılan hata % 4 iken RDA yönteminde bu hata % 6 olmuştur. Buna göre, iki kısıtlı ordinasyon yönteminde de hatanın birbirine yakın ve % 10’un altında olduğu görülmüştür.

5. TARTIŞMA VE SONUÇ

Çalışmada, RDA ve KUA uygulamalarında, yöntemlerin pseudo F değerleri hesaplanmış ve 500 tekrarlı permütasyon testi yapılarak kritik değer elde edilmiştir. Buna göre, pseudo F değerlerine göre hipotezler test edilmiştir. KUA uygulamasında, cevap ve açıklayıcı değişkenler arasında doğrusal bir ilişkinin olmadığı bulunurken, (unimodal ilişkinin varlığı bulunurken) RDA uygulamasında, cevap ve açıklayıcı değişkenler arasında doğrusal (lineer) bir ilişki bulunmuştur. Ayrıca, bu hipotez kontrolüne göre kısıtlı ordinasyon eksenlerinin, açıklayıcı değişkenlerin doğrusal kombinasyonlarından oluştuğu görülmüştür. Benzer şekilde, Muller (1981); Ramette (2007); Buttigieg ve Ramette (2014), bir RDA/KUA çözümünde, RDA/KUA eksenleri için anlamlılık değerinin permütasyon testi ile belirlenebileceğini ifade etmişlerdir. Genel çözüm anlamlı olduğunda, eksenlerin veya açıklayıcı değişkenlerin önem kontrolünün permütasyon testi ile yapılabileceğini belirtmişlerdir.

KUA ve RDA, determinasyon (belirleme) katsayıları bakımından değerlendirildiğinde; cevap değişkeni, nesnelere ve açıklayıcı değişkenler arasındaki farklı ilişki yapılarını analiz eden kısıtlı ordinasyon yöntemlerinin varyasyonu yeterli oranda (en az % 94 oranında) açıkladığı görülmüştür. Buna göre, kısıtlı ordinasyon yöntemlerinin (KUA ve RDA) ürettiği ordinasyon diyagram modellerinden yapılan ilişki tahminlerinin güvenilir, tutarlı ve toplam varyasyonu büyük oranda açıklayıcı olabileceği söylenebilir.

RDA uygulamalarında, bir kısım açıklayıcı değişkenlerin doğrusal kombinasyonları ile bir dizi cevap değişkeni yüksek oranda açıklanabilmiştir. Benzer şekilde Ter Braak (1988), Kısıtlı Ordinasyon Yöntemlerinin (RDA/KUA gibi) çevre ile ilgili olan topluluk varyasyonunun analizi için ordinasyona eşdeğer bir araç olarak kullanıldığını ifade etmiştir.

Muller (1981), Ter Braak (1994) ve Ramette (2007), RDA'nın, açıklayıcı değişkenlerin bir matrisi ile açıklanabilen cevap matrisindeki ana değişim modellerini özetleyen bir koordinasyon ürettiğini belirtmişlerdir. Ayrıca, veri kümesinin, kısıtlanmış ve kısıtlanmamış varyanslara bölünerek toplam varyansının hesaplandığını vurgulamışlar ve bu işlemin, açıklayıcı değişkenlerdeki varyasyonun cevap

değişkenlerindeki varyasyona göre ne kadarının gereksiz olduğunu gösterdiğini ifade etmişlerdir. Bunun yanında, kısıtlanmış varyansın, kısıtlanmamış varyansdan çok daha yüksek değerler aldığı durumda analizin, cevap (tepki) verisindeki değişimin çoğunun açıklayıcı değişkenler tarafından açıklanabileceğini gösterdiğine vurgu yapmışlardır.

Bunun yanısıra, Muller (1981); Ramette (2007); Buttigieg ve Ramette (2014), bir RDA çözümünün kanonik olmayan (kısıtlanmamış) vektörlerinin koordinasyon ve korelasyon ile incelenmesinin, gereksiz değişkenlerin davranışları hakkında bilgi verdiğini belirtmişlerdir. Ter Braak (1988) ise Kısmi Ordınasyon ve Kısmi Kısıtlı Ordınasyon Yöntemlerinin; belirli çevresel, mekânsal veya zamansal “kovaryantların” etkilerinin ortadan kaldırılmasından sonra kalan (artık) varyasyonu açıklamaya çalıştığını belirtmiştir.

KUA’da cevap değişkenlerinin, açıklayıcı değişkene göre cevap modeli incelendiğinden, eksenlerin kısıtlanma işleminin varyans açıklama oranını doğrudan etkilediği düşünülmektedir. Diğer yandan, RDA uygulamasında, kısıtlı ve kısıtlanmamış eksenlerin cevap verisindeki toplam varyansı açıklama oranı birbirine yakın bulunmuştur. Bu durumun, RDA’da, bir dizi açıklayıcı değişkenin doğrusal kombinasyonları kullanılarak, bir dizi cevap değişkenindeki varyasyonun açıklanmaya çalışılmasından kaynaklandığı düşünülmektedir.

KUA ve RDA yöntemleri için MAPE değeri % 10’un altında bulunmuştur. Buna göre, adı geçen yöntemlerin MAPE kriterine göre iyi sınıflama yaptıkları ve düşük oranda bir hata değeri ile ordınasyon diyagramını gerçekleştirdikleri söylenebilir. Bu nedenle, elde edilen kısıtlı ordınasyon diyagram modellerinin, açıklayıcı ve cevap değişkenlerinin birlikte veya ayrı ayrı diğer nesnelere olan farklı ilişki yapılarının tahmininde, güvenilir ve tutarlı sonuçlar vererek yüksek performans gösterdiği söylenebilir.

Ural (1992), RDA sonucunda, ilgilenilen değişken kümesinin varyasyona olan katkısını maksimum yapan bileşenler bulunduğunu ve bu bileşenler (genellikle ilk iki bileşen) ile boyut indirgeme işleminin mümkün olabileceğini vurgulamıştır. Buna paralel olarak, Ünlükaplan (2008), Temel Bileşenler Analizinde, ilgilenilen değişken kümesine ait varyans’a katkısı en büyük olan bileşenlerin, boyut sayısını gösterdiğini

belirtmiş ve böylece birbirinden bağımsız bileşenler yardımı ile ilgilenilen değişken kümesindeki değişken sayısının azaltılabileceğini ifade etmiştir.

Çalışmada, orijinal değişkenler arasında yüksek korelasyonlu değişkenler olmasına rağmen uygulamalarda anlamlı ikinci bir eksen elde edilmiştir. Bu sonuç Ter Braak'ın (1987), “KUA koordinasyon diyagramı, türler arasında ya da çevresel değişkenler arasındaki yüksek korelasyondan etkilenmez” ifadesini desteklemektedir. Benzer şekilde Palmer (1993), KUA’da yüksek ilişkili değişkenlerin analize dahil edilmesi halinde, anlamlı bir ikinci eksenin elde edilmesine mani olamayacağını belirtmiştir. Ancak, Ter Braak ve Prentice (1988), açıklayıcı değişkenlerin birbirleriyle güçlü bir şekilde ilişkili olduğu durumda (örn.; açıklayıcı değişkenlerin sayısı cevap değişkenlerinin sayısına yaklaştıkça), farklı açıklayıcı değişkenlerin topluluk kompozisyonu üzerindeki etkilerinin ayrıştırılmasının mümkün olmadığını belirtmişler ve bu durumda kanonik katsayıların kararsız hale geldiğine dikkat çekmişlerdir.

Ter Braak (1987), türlerin unimodal cevap fonksiyonları gösterdiği durumda, KUA’ya göre CCorA ve RDA’nın daha başarısız olduğunu belirtmiştir. Benzer şekilde Ter Braak (1988), Gaussian yönteminin, Kanonik ordinasyondaki “sınırlama (kısıtlama)” adımını daha istikrarlı hale getirdiğini belirtmiş, ancak KUA’nın çok daha pratik bir yaklaşım olduğunu belirtmiştir.

Çalışmada, RDA ve KUA uygulama sonuçlarından elde edilen skor değerlerinin I. ve II. Tip ölçeklemeye göre değişiklik gösterdiği tespit edilmiştir. Ancak, çalışmanın diyagram sonuçları I. Tip ölçeklemeye göre elde edilmiştir. Buna paralel olarak, Ter Braak (1994); Buttigieg ve Ramette (2014), değişkenlerin eksenlere göre "skor-puan" değerlerinin, RDA ve KUA sonucunun tipik bir özelliği olduğunu ve kullanılan ölçeklemeye bağlı olarak değişebildiğini ifade etmişlerdir.

Çalışmada, RDA ve KUA uygulamaları yapılırken orijinal veriler kullanılmış ve sonuçlar buna göre değerlendirilmiştir. Palmer (1993), KUA 'daki anlamlılık testlerinin, parametrik dağılım varsayımlarına bağlı olmadığından, değişkenleri herhangi bir dağılıma uygun hale getirmek için orijinal verilerin dönüştürülmesinin gerekli olmadığını belirtmiştir. Ayrıca Palmer (1993), KUA’nın çarpık tür dağılımlarında iyi

performans gösterdiğini bu nedenle, çevresel verilerin dönüşümlerinden minimum seviyede etkileneceğini ifade etmiştir.

Tso (1981) ile Jongman ve ark. (1987), CCorA ile RDA arasında benzerlikler olmasına rağmen hata bileşenine ilişkin varsayımların farklılık gösterebildiğine dikkat çekmişlerdir. RDA'da hataların, korelasyonsuz ve eşit varyansa sahip olduğunu ancak, CCorA'da hataların, korelasyonlu ve normal dağılım gösterdiğini vurgulamışlardır.

Graffelman (2001) çalışmasında, KUA'da aynı anda üç matrisin analize dâhil edilebildiğini belirtmiştir. Bu matrislerin, bolluk matrisi (sitelerdeki türlerin sayısı), çevresel değişkenlere göre türlerin optimal matrisi (ağırlıklandırılmış ortalama matrisinden tahmin edilen) ve çevresel verilere (sitelerdeki ölçülen kimyasal değişkenler) ait matrisler olduğunu ifade etmiştir. Ayrıca, KUA ile elde edilen özdeğerlerin, ağırlıklandırılmış ortalamaların varyansını göstermediğine dikkat çekmiş; ancak, bu özdeğerlerin bolluk matrisinin varyansı için bir uyum yöntemi olduğuna vurgu yapmıştır.

Çalışmadan elde edilen sonuç ve öneriler aşağıda özetlenmiştir.

RDA'da, sürekli açıklayıcı değişkenlerin varyasyona daha büyük katkı yaptığı söylenebilirken açıklayıcı değişkenler, farklı ölçekli değişkenlerden oluştuğunda, birinci eksenin açıkladığı varyans oranında azalma görülmüştür. Bununla beraber RDA'da, cevap değişkenindeki artışın birinci kısıtlı eksene ait açıklanan varyans oranını genellikle düşürdüğü söylenebilir. Diğer yandan, RDA'da, açıklayıcı değişkenlerin 1'den fazla olmasının, kısıtlama işlemi daha anlamlı yapacağı düşünülmekte, iki boyutlu uzayda değişkenlerin iki kısıtlı eksenle daha doğru açıklanabileceği beklenmektedir. Ayrıca, daha doğru RDA sonuçlarının elde edilebilmesi için site sayısının cevap değişkeni sayısından en az 2 fazla olması gerekmekte ve cevap değişkeni sayısının en az 2 olması önerilmektedir.

RDA uygulamalarında, bir kısım açıklayıcı değişkenlerin doğrusal kombinasyonları (TBA'dakine benzer şekilde) ile bir dizi cevap değişkeni yüksek oranda açıklanabilmiştir.

Açıklayıcı değişkenlerin sayısı cevap değişkenlerinin sayısına göre küçük olduğunda, kısıtlı ordinasyon yöntemlerinin (RDA/KUA) varyans açıklama oranının % 100' e yakın değerler verdiği görülmüştür. KUA'da, birinci kısıtlı eksenin açıkladığı varyans miktarının, açıklayıcı değişkenlerle cevap değişkenleri sayısının arasındaki farkın, cevap değişkeni lehine artmasıyla yükseldiği görülürken; RDA'da ise bu farkın açıklayıcı değişkenler lehine artmasıyla yükseldiği dikkat çekici bulunmuştur.

KUA'da, cevap değişkenlerinin açıklayıcı değişkene göre cevap modeli incelendiğinden, eksenlerin kısıtlanma işleminin varyans açıklama oranını doğrudan etkilediği görülmüştür. Ancak, RDA uygulamalarında, kısıtlı ve kısıtlanmamış eksenlerin cevap verisindeki toplam varyansı açıklama oranı birbirine yakın bulunmuştur. Bu durumun, RDA'da, bir dizi açıklayıcı değişkenin doğrusal kombinasyonları kullanılarak, bir dizi cevap değişkenindeki varyasyonun açıklanmaya çalışılmasından kaynaklandığı düşünülmektedir.

Cevap değişkeni, nesnelere ve açıklayıcı değişkenler arasındaki ilişki yapılarını analiz eden kısıtlı ordinasyon yöntemlerinin (KUA ve RDA'nın), varyasyonu yeterli oranda açıkladığı ve elde edilen ordinasyon modelinin yeterli olduğu görülmüştür.

KUA ve RDA yöntemlerinin MAPE kriterine göre iyi sınıflama yaptıkları ve çok düşük oranda (% 10'un altında) bir hata değeri ile ordinasyon diyagramını gerçekleştirdikleri görülmüştür.

Sonuç olarak, farklı ölçek türlerinde olan değişkenler arasındaki ilişki yapılarını analiz eden kısıtlı ordinasyon yöntemleri olan KUA ve RDA'nın varyasyonu yeterli düzeyde açıkladığı görülmüştür. Ayrıca, her iki yöntemin düşük MAPE değeri ile (% 10'dan düşük) iyi sınıflama yaptığı ve ordinasyon diyagramını gerçekleştirdiği böylece sağlık alanında da kullanılabileceği söylenebilir.

KAYNAKLAR

- Angers B, Magnan P, Plantes M, Bernatchez L. Canonical correspondence analysis for estimating spatial and environmental effects on microsatellite gene diversity in brook charr (*Salvelinus fontinalis*). *Mol Ecol*. 1999; 8(1): 1043-53.
- Anonymous 1 (2018). Eriřim adresi: <http://rfunctions.blogspot.com.tr/2016/11/canonical-correspondence-analysis-cca.html>. Eriřim Tarihi: 15/07/2018.
- Anonymous 2 (2018). Eriřim adresi: <https://mb3is.megx.net/gustame/constrained-analyses/rda>. Eriřim Tarihi: 15/07/2018.
- Anonymous 3 (2018). Eriřim adresi: <https://mb3is.megx.net/gustame/constrained-analyses/cca>. Eriřim Tarihi: 15/07/2018.
- Anonymous 4 (2019). Eriřim adresi: https://www.researchgate.net/publication/262980567_Root_mean_square_error_RMSE_or_mean_absolute_error_MAE. Eriřim Tarihi: 12. 04. 2019
- Anonymous 5 (2019). Eriřim adresi: <https://www.statisticshowto.datasciencecentral.com/mean-absolute-percentage-error-mape/>. Eriřim Tarihi:12. 04. 2019
- Austin MP, Gaywood MJ. Current problems of environmental gradients and species response curves in relation to continuum theory. *J Veg Sci*. 1994; 5(1): 473-82.
- Austin MR, Nicholls AO, Doherty MD, Meyers JA. Determining species response functions to an environmental gradient by means of a β function. *J Veg Sci*. 1994; 5(1): 215-28.
- Borcard D. Université Laval Multivariate analysis-February 2006. Eriřim adresi: http://ubio.bioinfo.cnio.es/Cursos/CEU_MDA_07_practicals/Further%20reading/Multivariate%20analysis%20Borcard%202006/Chap_4b.pdf. Eriřim tarihi:18/07/2018.
- Buttigieg PL, Ramette A. A Guide to Statistical Analysis in Microbial Ecology: a community-focused. living review of multivariate data analyses. *FEMS Microbiol Ecol*. 2014; 90(1): 543–50.
- De Leeuw J. The GIFI system of nonlinear multivariate analysis. M. Jambu. L. Lebart. J. Pages. and R. Tomassone. editors. *Data analysis and informatics III*. North-Holland Publishing. Amsterdam. The Netherlands. 1984. 415-24.
- Donovan C. Optimal transformations in multivariate analysis of trawl data. [Thesis]. New Zealand, Auckland: University of Auckland; 1998.
- Durand JF. Generalized principal component analysis with respect to instrumental variables via univariate spline transformations. *CSDA*. 1993; 16(1): 423–40.

- Escoufier Y. The duality diagram: a means for better practical applications. P. Legendre and L. Legendre. editors. *Developments in numerical ecology*. Springer-Verlag. Berlin. 1987. 139-56.
- Gabriel KR. The biplot graphic display of matrices with application to principal component analysis. *Biometrika*. 1971; 58(1): 453–67.
- Gauch HG. *Multivariate analysis in community ecology*. Cambridge University Press. Cambridge. England. 1982a.
- Gauch HG. Noise reduction by eigenvector ordinations. *Ecol*. 1982b; 63(1): 1643-49.
- Gittins R. *Canonical analysis – A review with applications in ecology*. Springer-Verlag. Berlin. 351. 1985; 579-631.
- Graffelman J. Quality statistics in canonical correspondence analysis. *Environmetrics*. 2001; 12(1): 485-97.
- Hejemanova-Nezerkova P, Hejeman M. A canonical correspondence analysis (CCA) of the vegetation- environment relationships in Sudanese savannah. Senegal. *S Afr J Bot*. 2006; 72(1): 256-62.
- Hill MO. Correspondence analysis: a neglected multivariate method. *Appl Statist*. 1974; 23: 340–54.
- Hill MO. DECORANA - A FORTRAN program for detrended correspondence analysis and reciprocal averaging. *Ecology and Systematics*. Cornell University. Ithaca. New York. 1979.
- Hill MO, Gauch HG. Detrended correspondence analysis: an improved ordination technique. *Vegetatio*. 1980; 42: 47–58.
- Israels AZ. Redundancy analysis for qualitative variables. *Psychometrika*. 1984; 49: 331-46.
- Israels AZ. Redundancy Analysis for Various Types of Variables. *Statistica Appl*. 1992; 4(4): 531-42.
- Iwatsubo S. The analytical solutions of eigenvalue problem in the case of applying optimal scoring method to some types in data. In: *Data Analysis and Informatics 3* (Ed. by E. Diday). North-Holland. Amsterdam. 1984. 31–40.
- Jongman RHG, Ter Braak CJF, Van Tongeren OR. *Data Analysis in Community and Landscape Ecology*. Pudoc. Wageningen. 1987.
- Jongman RHG, Ter Braak CJF, Van Tongeren OR. *Data analysis in community and landscape ecology*. Cambridge University Press. Cambridge. UK. 1995.
- Kalkan SB, Özden ÜH. Dünya Üniversitelerinin İtibarını Etkileyen Değişkenlerin Kanonik Korelasyon Analizi ile Belirlenmesi. *Soc Sci Res J*. 2017; 6(2): 11-19.

Karahan M. Turizm Talebinin Yapay Sinir Ağaları Yöntemiyle Tahmin Edilmesi. *J Fac Econ Admin Sci.* 2015; 20 (2): 205.

Lambert Z, Wildt AR, Durand RM. Redundancy analysis: An alternative to canonical correlation and multivariate multiple regression in exploring interspecies associations. *Psychological Bull.* 1988; 104(2): 282-9.

Lebreton JD, Sabatier R, Banco G, Bacou AM. Principal components and correspondence analyses with respect to instrumental variables: an overview of their role in studies of structure-activity and species-environment relationships. In: Devillers. J. Karcher. W. (Eds.). *Applied Multivariate Analysis in SAR and Environmental Studies.* EEC. Brussels. Belgium. 1991. 85-114.

Legendre P, Anderson MJ. Distance-based redundancy analysis: testing multi-species responses in multi-factorial ecological experiments. *Ecol Monogr.* 1999; 189- 96.

Legendre P, Legendre L. *Numerical Ecology.* Second English Edition. Elsevier Science B.V. Amsterdam. Netherlands. 1998. ISBN 978-0444892508.

Lewis CD. *Industrial and Business Forecasting Methods.* Londra: Butterworths Publishing, 1982.

Makarenkov V, Legendre P. Nonlinear Redundancy Analysis and Canonical Correspondence Analysis Based on Polynomial Regression. *Ecol.* 2002; 83(4): 1146-61.

Mardia KV, Kent JT, Bibby JM. *Multivariate analysis.* Academic Press. London. England. 1979.

McCune B, Grace JB. *Analysis of Ecological Communities.* MjM Software Design. Gleneden Beach. Oregon. 2002.

Muller KE. Relationships Between Redundancy Analysis, Canonical Correlation and Multivariate Regression. *Psychometrika.* 1981; 46(2): 139-42.

O'Connell MT, Robert CC, Christopher SS. Fish Assemblage Stability Over Fifty Years in the Lake Pontchartrain Estuary; Comparisons Among Habitats Using Canonical Correspondence Analysis. *Estuaries.* 2004; 5(27): 807-17.

Oksanen J. *Multivariate Analysis in Ecology (Lecture notes).* Department of Biology. University of Oulu. 2004.

Palmer MW. Putting things in even better order: the advantages of canonical correspondence analysis. *Ecol.* 1993; 8(74): 2215-30.

Palmer MW (2017). Erişim adresi: http://ordination.okstate.edu/overview.htm#Redundancy_Analysis. Erişim Tarihi: 18/05/2018.

Pielou EC. *The interpretation of ecological data: a primer on classification and ordination.* Wiley. New York. USA. 1984.

- Ramette A. Multivariate analyses in microbial ecology. *FEMS Microbiol Ecol.* 2007; 62(2): 142–60.
- Rao CR. The use and interpretation of principal component analysis in applied research. *Sankhyá. Ser.* 1964; A26: 329-58.
- Rozeboom W (1965). Linear correlations between sets of variables. *Psychometrika.* 30. 1. 57-71.
- Sabatier RJ, Lebreton D, Chessel R. Principal component analysis with instrumental variables as a tool for modelling composition data. in: R. Coppi & S. Bolasso [eds.] *Multivariate data analysis.* Elsevier Science Publishers B. V. North Holland. the Netherlands. 1989. 341-52.
- Saporta G. *Probabilités. analyses de données et statistiques.* Editions Technip. Paris. 1990. 492-93.
- Stewart D, Love W. A general canonical correlation index. *Psychological Bull.* 1968; 70(1): 160-3.
- Ter Braak CJF. Correspondence analysis of incidence and abundance data: properties in terms of a unimodal response model. *Biometrics.* 1985b; 41: 859-73.
- Ter Braak CJF. Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. *ESA.* 1986; 5(67): 1167-79.
- Ter Braak CJF, Looman CWN. Weighted averaging, logistic regression and the Gaussian response model. *Vegetatio.* 1986; 65: 3-11.
- Ter Braak CJF. The Analysis of Vegetation-Environment Relationships by Canonical Correspondence Analysis. *Vegetatio.* 1987; 64: 69-77.
- Ter Braak CJF. Groep Landbouwwiskunde. Agricultural Mathematics Group. P.O.Box 100. 6700 AC Wageningen. The Netherlands. 1988.
- Ter Braak CJF, Prentice IC. A theory of gradient analysis. *Adv Ecol Res.* 1988;18: 271–317.
- Ter Braak CJF. Canoco—a fortran program for canonical community ordination by [partial] [detrended] [canonical] correspondence analysis, principal component analysis and redundancy analysis (Version 2.1). Technical report LWA-8802. Agricultural Mathematics Group. AC Wageningen. The Netherlands. 1988a.
- Ter Braak CJF. Update notes: CANOCO version 3.10. Agricultural Mathematics Group. Wageningen. 1990.
- Ter Braak CJF. Canonical community ordination. Part I: basic theory and linear methods. *Ecoscience.* 1994; 1: 127–40.

Ter Braak CJF, Looman CWN. Biplots in reduced-rank regression. *Biom J.* 1994; 36: 983-1003.

Ter Braak CJF, Wiertz J. On the statistical analysis of vegetation change: a wetland affected by water extraction and soil acidification. *J Veg Sci.* 1994;5: 361-72.

Ter Braak CJF, Verdonschot PFM. Canonical correspondence analysis and related multivariate methods in aquatic ecology. *Aquat Sci.* 1995; 57-3.

Tso MKS. Reduced-rank regression and canonical analysis. *J Roy Statist Soc.* 1981; 43(B): 183-89.

Ural Y. Gereksizliđi belirleme analizi ve diđer çok deđişkenli analizlerle iliřkisi. [Yüksek Lisans Tezi]. Ankara: Hacettepe Üniversitesi; 1992.

Ünlükaplan Y. Çok deđişkenli istatistiksel yöntemlerin peyzaj ekolojisi arařtırmalarında kullanımı. [Doktora Tezi]. Adana: Çukurova Üniversitesi; 2008.

Ünver Ö. Uygulamalı İstatistik Yöntemler. (İkinci Baskı). Ankara: Siyasal Kitabevi, 1996.

Van der Burg E, De Leeuw J. Non-linear canonical correlation. *Brit J Math Stat Psy.* 1983; 36(1): 54-80.

Verdonschot PFM, Ter Braak CJF. An experimental manipulation of oligochaete communities in mesocosms treated with chlorpyrifos or nutrient additions: multivariate analyses with Monte Carlo permutation tests. *Hydrobiologia.* 1994; 278. 251-66.

Wartenberg D, Ferson S, Rohlf FJ. Putting things in order: a critique of detrended correspondence analysis. *Am Natur.* 1987; 129: 434-48.

Wollenberg ALV. Redundancy Analysis an Alternative for Canonical Correlation Analysis. *Psychometric.* 1977; 42(2): 207-19.

ÖZGEÇMİŞ

Mehmet Tahir Huyut, 1983 yılında Van'da doğdu. Orta öğrenimini, 2000 yılında Van Kazım Karabekir Lisesinde tamamladı. Yükseköğrenimine 2001 yılında Hacettepe Üniversitesi Gıda Mühendisliği Bölümü'nde başladı. Bu bölümü bıraktıktan sonra, 2003 yılında Karadeniz Teknik Üniversitesi Bilgisayar Teknolojileri Öğretmenliği Bölümü'ne kayıt yaptırdı. Ancak bu bölümü de bıraktıktan sonra, 2005 yılında Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Matematik Öğretmenliği Bölümü'ne başladı ve 2010 yılında bu bölümde lisans öğrenimini tamamladı. 2009 yılında Antalya Milli Eğitim Müdürlüğü'nde memur olarak göreve başladıktan sonra, 2011 yılında Matematik öğretmeni olarak çalışmaya başladı. Aynı yıl bahar döneminde Yüzüncü Yıl Üniversitesi Sağlık Bilimleri Enstitüsü, Biyoistatistik Anabilim Dalı'nda Yüksek Lisans öğrenimine başladı ve 2014 yılında bitirdi. 2015 yılında aynı ana bilim dalında Doktora öğrenimine başladı ve 2019 yılında bitirdi. Evli ve iki çocuk babasıdır.

EKLER

EK 1

Terminolojik Sözlük

Topluluk (Community): Bir örnekte birlikte ortaya çıkan bir grup, tür

Örnek (Sample): Örnekleme birimi, birey, nesne, site

Türler (Species): Cevap (tepki) değişkeni, bir regresyon denklemindeki bağımlı değişken, iç değişken

Bolluk/cevap (abundance/response): Bir cevap değişkeninin değeri (genellikle pozitif ya da 0 değer alır)

Çevresel değişken (environmental variable): Açıklayıcı değişken (asıl ilgilenilen değişken), bir regresyon denkleminde bağımsız değişken, dış değişken, uyarıcı değişken, iyileştirici değişken

Kovaryant (covariable): Doğal sonuç değişkeni (eşlik eden değişken), arka plan değişkeni, önemsiz ya da sorun parametrelerine uygun açıklayıcı değişkenler, deney tasarımında blok faktörü

Dolaylı gradient analizi (indirect gradient analysis): İçsel analiz, "faktör analizi", sınırlandırılmamış ordinasyon, sınırlandırılmamış çok boyutlu ölçekleme, dışsal değişkenler üzerinde bir regresyon analizi ile post-hoc olarak izlenebilen

Doğrudan gradient analizi (direct gradient analysis): Dış analiz, kanonik ordinasyon, dış değişkenler tarafından sınırlandırılmış ordinasyon, sınırlı çok değişkenli regresyon, indirgenmiş sıra regresyonu

Ordinasyon (ordination): Dolaylı gradient analizine bakınız

Ordinasyon eksenleri (ordination axis): Özvektör, latent (gizli) değişken, teorik açıklayıcı değişken

Ordinasyon diyagramı (ordination diagram): Özvektör puanlarının dağılım grafiği; biplotlar ve ortak alanlar için kullanılır

Kanonik ordinasyon (canonical ordination): Bu ordinasyon da eksenler, çevresel değişkenlerin doğrusal kombinasyonları olarak kısıtlanır

Kanonik eksenler (canonical axis): Çevresel değişkenlerin doğrusal bir kombinasyonu olarak sınırlanan ordinasyon eksenleri

Özdeğerler (eigenvalue): Bir ordinasyon ekseninin önem ölçüsü

Tür skoru (species score): Özvektör katsayısı; PCA'da yükler, CA ve DCA'da tür eğrisinin merkezi

Örnek skoru (sample score): Bir örnekte özvektör değeri

Biplot (biplot): İki tür varlığın bir koordinasyon diyagramı, örn. belirli bir yorumlama kuralına sahip olan türler ve çevresel değişkenler, yorumlama biplottaki oklarla tanımlanan yönler üzerinde devam eder

Ortak alan (joint plot): Ağırlıklı ortalama alma yöntemine dayanan iki tür varlığın bir ordinasyon diyagramı

Doğrusal yöntem (linear method): Doğrusal bir modele dayalı yöntem, örneğin: doğrusal regresyon, çoklu regresyon, temel bileşenler analizi, gereksizlik analizi

Ağırlıklı ortalama yöntemi (weighted averaging method): Optimum (mod, ideal nokta) ağırlıklı ortalamalama ile tahmin edilen unimodal cevap (tepki) modeline dayanan bir yöntem, örn; uyum analizi

EK 2

Etik Kurul Raporu



**T.C.
YÜZÜNCÜ YIL ÜNİVERSİTESİ
Girişimsel Olmayan Klinik Araştırmalar
Etik Kurul Başkanlığı**

Sayı : B.30.2.YYU.0.01.00.00/22

Tarih: 21/02/2018


Konu : Etik Kurul Başvurunuz

Sn. Prof.Dr. Sıddık KESKİN

İlgi: 08.02.2018 tarih ve bila sayılı yazınız.

İlgi yazı ile Etik Kurulumuza sunulan “**Farklı Değişken Türlerini İçeren Veri Setlerinin Analizinde Kullanılan Bazı Çok Değişkenli İstatistik Analiz Yöntemlerinin İncelenmesi**” isimli proje özetiniz 16.02.2018 tarihinde yapılan Etik Kurul toplantımızda görüşüldü. Yapılan görüşmede çalışmanın proje özetinin içeriği tamamiyle istatistiksel analiz değerlendirmesi olduğundan adı geçen çalışmanın etik kurulla doğrudan ilişkisi saptanmamıştır.

Bilgilerinize rica ederim.


Prof. Dr. Oğuz TUNCER
Girişimsel Olmayan
Klinik Araştırmalar
Etik Kurul Başkanı

Adres: 65080 Zeve Kampüsü - VAN Tel : (0432) 2150470 E-Posta: etikkurull@gmail.com
Fax : (0432) 2168352 Web: http://www.yyu.edu.tr

EK 3

Tez Orjinallik Raporu

	<p style="text-align: center;">T.C. VAN YÜZÜNCÜ YIL ÜNİVERSİTESİ Sağlık Bilimleri Enstitüsü</p>	
DOKTORA TEZİ ORJİNALLİK RAPORU		

Tarih: 06/05/2019

Tez Başlığı / Konusu: KISITLI ORDİNASYON YÖNTEMLERİNDEN KANONİK UYUM ANALİZİ VE GEREKSİZLİK ANALİZİ YÖNTEMLERİNİN KARŞILAŞTIRILMASI: KARDİYOLOJİK VERİ SETİ ÜZERİNDE BİR UYGULAMA


Yukarıda başlığı/konusu belirlenen tez çalışmamın Kapak sayfası, Giriş, Ana bölümler ve Sonuç bölümlerinden oluşan toplam 83 sayfalık kısmına ilişkin, 06/05/2019 tarihinde şahsım/tez danışmanım tarafından TURNİTİN intihal tespit programından aşağıda belirtilen filtreleme uygulanarak alınmış olan orjinallik raporuna göre, tezimin benzerlik oranı % 4 (iki) tür.

Uygulanan filtreler aşağıda verilmiştir:

- Kabul ve onay sayfası hariç,
- Teşekkür hariç,
- İçindekiler hariç,
- Simge ve kısaltmalar hariç,
- Gereç ve yöntemler hariç,
- Kaynakça hariç,
- Alıntılar hariç,
- Tezden çıkan yayınlar hariç,
- 7 kelimedenden daha az örtüşme içeren metin kısımları hariç (Limit match size to 7 words)

Van Yüzüncü Yıl Üniversitesi Lisansüstü Tez Orjinallik Raporu Alınması ve Kullanılmasına İlişkin Yönergeyi inceledim ve bu yönergede belirtilen azami benzerlik oranlarına göre tez çalışmamın herhangi bir intihal içermediğini; aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.

Gereğini bilgilerinize arz ederim.


Mehmet Tahir HUYUT

Öğrencinin Adı Soyadı	Mehmet Tahir HUYUT
Anabilim Dalı	: Biyoistatistik
Öğrenci No	149302052
Program	: <input type="checkbox"/> Yüksek Lisans <input checked="" type="checkbox"/> Doktora
DANIŞMAN ONAYI UYGUNDUR (Prof. Dr. Sıddık KESKİN)	ENSTİTÜ ONAYI UYGUNDUR (Prof. Dr. Semiha DEDE)