

**BULANIK REGRESYON MODELLERİNİN TUTARLILIĐI ÜZERİNE ÇALIŐMALAR**

**SELCEN GÜLSÜM ASLAN ÖZŐAHİN**

**YÜKSEK LİSANS TEZİ**  
**ENDÜSTRİ MÜHENDİSLİĐİ**

**TOBB EKONOMİ VE TEKNOLOĐİ ÜNİVERSİTESİ**  
**FEN BİLİMLERİ ENSTİTÜSÜ**

**ARALIK 2014**

**ANKARA**

Fen Bilimleri Enstitü Onayı

---

Prof. Dr. Osman EROĞUL  
Müdür

Bu tezin Yüksek Lisans derecesinin tüm gereksinimlerini sağladığını onaylarım.

---

Prof. Dr. Tahir HANALIOĞLU  
Anabilim Dalı Başkanı

Selcen Gülsüm ASLAN ÖZŞAHİN tarafından hazırlanan **BULANIK REGRESYON MODELLERİNİN TUTARLILIĞI ÜZERİNE ÇALIŞMALAR** adlı bu tezin Yüksek Lisans tezi olarak uygun olduğunu onaylarım.

---

Prof. Dr. İ Burhan TÜRKŞEN  
Tez Danışmanı

Tez Jüri Üyeleri

Başkan : Prof. Dr. Tahir HANALIOĞLU

Üye : Prof. Dr. İ. Burhan TÜRKŞEN

Üye : Yrd. Doç. Dr. Murat ÖZBAYOĞLU

## **TEZ BİLDİRİMİ**

Tez içindeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edilerek sunulduğunu, ayrıca tez yazım kurallarına uygun olarak hazırlanan bu çalışmada orijinal olmayan her türlü kaynağa eksiksiz atıf yapıldığını bildiririm.

Selcen Glsm ASLAN ZAHİN

**Üniversitesi** : TOBB Ekonomi ve Teknoloji Üniversitesi  
**Enstitüsü** : Fen Bilimleri  
**Anabilim Dalı** : Endüstri Mühendisliği  
**Tez Danışmanı** : Prof. Dr. İ. Burhan TÜRKŞEN  
**Tez Türü ve Tarihi** : Yüksek Lisans – Aralık 2014

**Selcen Gülsüm ASLAN ÖZŞAHİN**

## **BULANIK REGRESYON MODELLERİNİN TUTARLILIĞI ÜZERİNE ÇALIŞMALAR**

### **ÖZET**

Günümüz dünyasında mevcut verilerin kayıt altında tutulması son derece kolay olup bu verilerin gelecekte yapılacak çalışmalara rehberlik edecek nitelikte derlenip işlenmesi; kişi, kuruluş ve devletlere sunduğu proaktif yaklaşım sayesinde önem taşımaktadır. Bu önemin en geçerli sebebi ise geçmişte gerçekleşen bir takım faaliyetlerin rakamsal ölçütü olan gerçek verilerin gelecekte gerçekleşmesi muhtemel faaliyetlere en iyi şekilde rehberlik etmesi ve bu sayede aksiyon planlarında en önemli yol göstericilerden biri olmasıdır.

Birçok otorite tarafından faydası hali hazırda kabul edilmiş bilgi yönetimi ve veri modelleme uygulamalarından olan Bulanık Mantık ile Modelleme yaklaşımının kullanıldığı bu çalışma kapsamında Danimarka devletinin 1993-2013 yılları arasında resmi olarak açıkladığı ithalat tutarı toplamı verileri Döviz Kuru, Üretici Fiyat İndeksi, Altın Rezervi vb. parametreleri girdi olarak kullanılmıştır. Bu veri kümesinde Bulanık Öbekleme Algoritması ile bulanıklık mertebesi ve öbek sayısı birbirinden farklı olan toplamda 221 farklı model için üyelik değerleri hesaplanmış, her bir model için Bulanık Regresyon Analizi yapılmıştır.

Sonuçta elde edilen her bir model için bu modellerin tutarlılık parametreleri olarak dikkate alınan R-kare değeri üzerinden değerlendirmeler ve genel geçer çıkarımlar değerlendirilmiştir. Bu çalışma kapsamında Çok Katmanlı Üyelikler adı altında ilk kez uygulanan ve sonuçları gözlemlenen yaklaşım; Bulanık Regresyon Analizi kapsamında sadece tek bir bulanıklık mertebesi ve öbek sayısı için hesaplanmış üyelik değerlerinin yer aldığı modeller yerine farklı öbek sayısı ve bulanıklık mertebesi için hesaplanmış üyeliklerin ve transformasyonları matrisinin de aynı modelde kullanılarak elde edilen melez modelin tutarlılığının gözlemlenmesidir.

**Anahtar kelimeler:** Finansal Veri Analizi, Bulanık Küme Algoritması, Bulanık Regresyon Uygulaması, Çok Katmanlı Üyelikler

**University** : TOBB Economics and Technology University  
**Institute** : Institute of Natural and Applied Sciences  
**Science Programme** : Industrial Engineering  
**Supervisor** : Prof. Dr. İ. Burhan TÜRKŞEN  
**Degree Awarded and Date** : M.Sc. – December 2014

**Selcen Gülsüm ASLAN ÖZŞAHİN**

## **STUDIES ON THE CONSISTENCY OF FUZZY REGRESSION MODELS**

### **ABSTRACT**

In today's World, collecting data is not a big issue in a proper and processable format with the help of recent technological developments but it is a big deal not only for the government bodies, private sector and also for the individuals. The main idea behind keeping data in a processable format is to discover the existing data patterns to estimate the future with the highest consistency. Enhancements in data collection, processing and estimation technics bring the strong knowledge from the past, ability to control processes and dominate the future based on the historical data. In other words, technical control on data management promises the power of knowledge to the all stakeholders in management and in many areas mainly in economics, engineering, and medical sciences as well as politics and public relations.

One of the leading research area in data management is modeling which are namely stochastic models, statistical models, linear modeling or fuzzy systems. Fuzzy Models are one of the most trendy and consistently estimating approach for modeling. In this study, the data set of Denmark which contains financial indicators as input and import amounts between 1993 and 2013 as output were used to create 221 different models by using Fuzzy C-Means Clustering Algorithm and Fuzzy Regression Analysis with different number of clusters and degree of fuzziness in each model. With the help of high number of experimental models, evaluation of model consistency parameter, R-square, depending on the changes of number of clusters and degree of fuzziness have done and several inferences have been achieved.

Last but not least, outstanding approach has been developed in this study which is named as Multi-Layer Fuzzy. In this approach more than one membership matrix calculated with at least two different number of cluster and degree of fuzziness values have been used in a single model and higher consistency in the estimations have been achieved as expected.

**Key Words:** Financial Data Analysis, Fuzzy C-Means Clustering Algorithm, Fuzzy Regression, Multi-Layer Fuzzy Approach

## TEŞEKKÜR

Lisans eğitimimi tamamlayıp bir müddet ara verdikten sonra başlattığım yüksek lisans eğitimimi de beklentilerimin epey ötesinde sonuçlar elde ettiğim bu tez çalışmasıyla tamamlamaktan çok büyük mutluluk duymaktayım.

Bu tez çalışması boyunca ve yüksek lisans eğitimime başladığım ilk dönem itibari ile verdiği derse katılarak hem bu çalışmaya hem de benim henüz başlayan akademik kariyerime çok büyük etki eden çok değerli hocam Prof. Dr. İ. Burhan TÜRKŞEN'e en derin teşekkürlerimi sunmayı borç bilirim. Hem geniş dünya görüşüne hem de akademik çalışmalarındaki hâkimiyetine çok büyük saygı duyduğum değerli hocam ile önümüzdeki yıllarda da farklı platformlarda yeniden birlikte çalışmayı gönülden isterim.

Yüksek lisans çalışmalarına dahil olduğu andan itibaren her daim desteğini hissettiğim çok değerli eşim Arif ÖZŞAHİN'e yoğun hayat temposunda akademik kariyerime vakit ayırmama olanak sağlayacak şekilde destek olduğu için ayrıca çok teşekkür etmek isterim.

Sadece yüksek lisans eğitimim süresince değil hayatım boyunca tamamladığım tüm güzel işlerde, başarılarında bana en büyük desteği sunan sevgili ailem, babam Sadık ASLAN, annem Kurtuluş ASLAN ve kardeşim Yavuzalp ASLAN'a değerli destekleri ve hayatımdaki her adımda bu denli güçlü şekilde desteklerini hissettirerek yanımda oldukları için sonsuz teşekkürlerimi sunmak isterim.

Son olarak yüksek lisans eğitimimi yoğun çalışma hayatım ile birlikte devam ettirdiğim süre boyunca esnek çalışma saatleri imkanı sunarak çalışanlarının akademik kariyerini geliştirmesine olanak sağlayan TÜBİTAK yönetimine de teşekkür etmek isterim. Aynı şekilde TOBB Ekonomi ve Teknoloji Üniversitesi yönetimine ARGE niteliği taşıyan kurumlarda çalışan ve belli koşulları sağlayan kişilere öğretim ücretinden muaf şekilde burslu yüksek lisans imkanı tanıdıkları için teşekkürlerimi iletmek isterim. Bu süreçte aldığım dersler kapsamında ya da görüşlerine başvurduğumda her daim desteklerini ve rehberliklerini esirgemeyen TOBB ETÜ Endüstri Mühendisliği Bölümü'nün değerli öğretim üyelerine de teşekkürü borç bilirim.

Bu tez çalışması kapsamında yaptığım çalışmaların ve elde ettiğim bulguların bana gelecekte yapacağım akademik çalışmalarımda motivasyon olmasını umar bu çalışma kapsamında bana destek veren herkese bu vesile ile teşekkür etmek isterim.

## İÇİNDEKİLER

	Sayfa
<b>ÖZET</b> .....	iv
<b>ABSTRACT</b> .....	v
<b>TEŞEKKÜR</b> .....	vi
<b>İÇİNDEKİLER</b> .....	vii
<b>ÇİZELGELERİN LİSTESİ</b> .....	ix
<b>ŞEKİLLERİN LİSTESİ</b> .....	xi
<b>KISALTMALAR</b> .....	xiii
<b>SEMBOL LİSTESİ</b> .....	xiv
<b>1. GİRİŞ</b> .....	1
<b>2. BULANIK SİSTEMLERİN GELİŞİMİ</b> .....	2
2.1. Zadeh Bulanık Kümeler Teorisi.....	3
2.2. Muzimoto Bulanık Kural Tabanlı Algoritması.....	4
2.3. Tagaki & Sugeno & Kang Kural Tabanlı Algoritması.....	6
2.4. Bulanık Öbek Ortalaması Algoritması.....	7
<b>3. BCO ÖBEKLEME ALGORİTMASI</b> .....	10
3.1. Veri Seti ve Analizi.....	10
3.1. 1. Zaman Serileri Analizi.....	10
3.1. 2. Basit Doğrusal Regresyon.....	12
3.2. Öbek Sayısı ve Hesaplama Yöntemleri.....	15
3.2. 1. XB İndeksi.....	17
3.2. 2. Kung-Lin İndeksi.....	19
3.2. 3. Geliştirilmiş Bulanık Öbekleme İndeksi.....	20

3.3. Öbek Merkezleri.....	24
3.4. Bulanıklık Mertebesi.....	25
3.5. Uzaklık Hesaplama Metotları.....	26
3.5. 1. Aralık Ölçekli ve Oransal Ölçekli Değişkenler.....	26
3.6. Bulanık Öbek Ortalaması Algoritması ve Adımları.....	28
<b>4. BULANIK REGRESYON MODELİ KURULMASI.....</b>	<b>33</b>
<b>5. DENEY KÜMESİNİN TANIMLANMASI VE MODELLERİN KURULMASI.....</b>	<b>36</b>
5.1. Veri Setinin Tanıtımı.....	38
5.2. Korelasyon Analizi.....	39
5.3. Regresyon Modelinin Oluşturulması.....	46
5.3.1. Basit Doğrusal Regresyon Analizi.....	46
5.4. Öbek Sayısı ve Bulanıklık Mertebesi Aralığının Belirlenmesi.....	48
5.4.1. Öbek Sayısının Belirlenmesi.....	48
5.4.2. Bulanıklık Mertebesinin Belirlenmesi.....	49
5.5. Üyeliklerin Hesaplanması.....	49
5.6. Üyeliklerin Transformasyonlarının Hesaplanması.....	52
5.7. Bulanık Regresyon Modellerinin Tutarlılığı Üzerine Değerlendirmeler.....	53
<b>6. BULANIK KÜME ALGORTİMASINDA YENİ BİR YAKLAŞIM: ÇOK KATMANLI BULANIK REGRESYON MODELİ.....</b>	<b>65</b>
6.1. Çok Katmanlı Bulanık Regresyon Modellerinin Tanımlanması.....	66
6.2. Çok Katmanlı Üyeliklerin Modelin Tutarlılığı Üzerine Etkisi.....	67
<b>7. GELECEK ÇALIŞMALAR.....</b>	<b>92</b>
<b>KAYNAKLAR.....</b>	<b>93</b>
<b>ÖZGEÇMİŞ.....</b>	<b>95</b>



## ÇİZELGELERİN LİSTESİ

Çizelge	Sayfa
Çizelge 3.1. Farklı öbek geçerlilik indeksi hesaplama yöntemleri	27
Çizelge 5.1. Veri setinde yer alan girdilerin isimleri ve indisleri	39
Çizelge 5.2. Korelasyon katsayısına göre sınıflandırma çizelgesi	40
Çizelge 5.3. Korelasyon Tablosu -1	41
Çizelge 5.4. Korelasyon Tablosu -2	43
Çizelge 5.5. Korelasyon Tablosu -3	44
Çizelge 5.6. Regresyon Analizi Çıktılar	47
Çizelge 5.7. Regresyon Analizi Model Tutarlılık Özeti	47
Çizelge 5.8. Modellerin Öbek Sayısı ve Bulanıklık Mertebesi Değişimleri- Bölüm 1	50
Çizelge 5.9. Modellerin Öbek Sayısı ve Bulanıklık Mertebesi Değişimleri- Bölüm 2	51
Çizelge 5.10. Model Tutarlılığı Deney Sonuçları	55
Çizelge 5.11. Regresyon Modeli Tutarlılık Parametreleri	57
Çizelge 5.12. Bulanıklık mertebesinin 1.4 olduğu modellerde model tutarlılığı değişimleri	61
Çizelge 5.13. Bulanıklık mertebesinin 1.8 olduğu modellerde model tutarlılığı değişimleri	62
Çizelge 5.14. Bulanıklık mertebesinin 2.2 olduğu modellerde model tutarlılığı değişimleri	63
Çizelge 5.15. Bulanıklık mertebesi ve öbek sayısına göre model tutarlılığı değişimi	63
Çizelge 5.16. Deneme 1 model içerikleri ve model tutarlılıkları özet tablosu	75
Çizelge 5.17. Deneme 2 model içerikleri ve model tutarlılıkları özet tablosu	86

Çizelge 5.18. Deneme 3 model içerikleri ve model tutarlılıkları özet tablosu	87
Çizelge 5.19. Deneme 4 model içerikleri ve model tutarlılıkları özet tablosu	89
Çizelge 5.20. Deneme 5 model içerikleri ve model tutarlılıkları özet tablosu	90

## ŞEKİLLERİN LİSTESİ

Şekil	Sayfa
Şekil 2.1. Bulanık sistemlerin genel yapısı	3
Şekil 2.2. Hiper-küre (solda) ve hiper-düzlem (sağda) illüstrasyonu	8
Şekil 3.1. Basit doğrusal regresyon doğrusu	14
Şekil 3.2. Gerçek ve tahmin edilen regresyon doğruları	15
Şekil 3.3. Verilerin saçılım grafiği (solda); verilerin öbekleşmesinin gösterimi (sağda)	16
Şekil 3.4. Oran tipi ÖĞİ indeksinin yoğunluk ve ayrılabilirliklerinin gösterimi	18
Şekil 3.5. Bulanık Fonksiyonlar yaklaşımları ile bulanık sistem modelleri yapısı	29
Şekil 5.1. Bulanık Regresyon Modeli iş akış şeması	37
Şekil 5.2. Model tutarlılığı deney sonuçları grafiksel gösterim	56
Şekil 5.3. Bulanıklık mertebesinin 1.4 olduğu modellerde tutarlılık parametresinin öbek sayısına göre değişimi grafiği	58
Şekil 5.4. Bulanıklık mertebesinin 1.8 olduğu modellerde tutarlılık parametresinin öbek sayısına göre değişimi grafiği	58
Şekil 5.5. Bulanıklık mertebesinin 2.2 olduğu modellerde tutarlılık parametresinin öbek sayısına göre değişimi grafiği	59
Şekil 5.6. Temsili örtüşen öbekleşme gösterimi	60
Şekil 5.7. Bulanıklık mertebesi 1.4,1.8 ve 2.2 olan modellerde model tutarlılığının öbek sayısına göre değişimi	64
Şekil 6.1. Klasik Bulanık Regresyon Modeli temsili gösterimi	65
Şekil 6.2. Çok Katmanlı Bulanık Regresyon Modeli temsili gösterimi	65
Şekil 6.3: Model A'da yer alan bir verinin öbeklere aitliğinin gösterimi	68
Şekil 6.4: Model A'da yer alan bir diğer verinin öbeklere aitliğinin gösterimi	69

Şekil 6.5. Model A'da belirtilen üyelik değerlerinin çıktı değeri ile dağılımı	70
Şekil 6.6. Model A'da belirtilen üyelik değerlerinin gösterimi	71
Şekil 6.7. Model B'de belirtilen üyelik değerlerinin çıktı değeri ile dağılımı	72
Şekil 6.8. Model B'de belirtilen üyelik değerlerinin gösterimi	74
Şekil 6.9 Model A ve B'de belirtilen üyelik değerleri ve çıktı değerlerinin gösterimi	76
Şekil 6.10. Model A+B'de belirtilen üyelik değerlerinin gösterimi	77
Şekil 6.11. Model D'de yer alan iki farklı verinin öbeklere aitliğinin gösterimi	78
Şekil 6.12. Model D'de belirtilen üyelik değerleri ile çıktı değerinin dağılımı	78
Şekil 6.13. Model E'de yer alan iki farklı verinin öbeklere aitliğinin gösterimi	79
Şekil 6.14. Model E'de belirtilen üyelik değerleri ile çıktı değerinin dağılımı	80
Şekil 6.15. Model D+E'de belirtilen üyelik değerlerinin gösterimi	81
Şekil 6.16. Model F'de yer alan iki farklı verinin öbeklere aitliğinin gösterimi	81
Şekil 6.17. Model F'de belirtilen üyelik değerleri ile çıktı değerinin dağılımı	82
Şekil 6.18. Model D+F'de belirtilen üyelik değerlerinin gösterimi	83
Şekil 6.19. Model E+F,'de belirtilen üyelik değerlerinin gösterimi	84
Şekil 6.20. Model D+E+F'de belirtilen üyelik değerlerinin gösterimi	85

## KISALTMALAR

<b>Kısaltmalar</b>	<b>Açıklama</b>
<b>BCO</b>	Bulanık Öbek Ortalaması
<b>BDR</b>	Basit Doğrusal Regresyon
<b>BRM</b>	Bulanık Regresyon Modeli
<b>ÇKBRM</b>	Çok Katmanlı Bulanık Regresyon Modeli
<b>GBÖ</b>	Geliştirilmiş Bulanık Öbekleme
<b>EKK</b>	En Küçük Kareler
<b>FCRM</b>	Bulanık Öbek Ortalaması Regresyon Modeli
<b>ÖGİ</b>	Öbek Geçerlilik İndeksi
<b>SGBÖ</b>	Sınıflandırmalı Geliştirilmiş Bulanık Öbekleme
<b>TSK</b>	Takagi-Sugeno-Kang bulanık kural tabanı sistemleri
<b>XB</b>	Xie & Ben

## SEMBOL LİSTESİ

Bu çalışmada kullanılmış olan simgeler açıklamaları ile birlikte aşağıda sunulmuştur.

<b>Simgeler</b>	<b>Açıklama</b>
$\ \cdot\ $	Öklit normu
$\mu_{ik}$	Üyelik değeri
$\alpha$	Doğrusal fonksiyonun sabitidir
$c^*$	Sistemdeki kural sayısı
$E$	$\varepsilon$ hata teriminin gerçek değerinin tahmini
$\varepsilon_t$	Hata terimi
$isr$	Zadeh tarafından literatüre kazandırılan bir değişkenin 0 ya da 1 gibi kesin olmayıp bulanık olduğunun tanımı
$m$	Bulanıklık derecesi
$r$	Korelasyon katsayısı
$U$	Üyelik değerlerinin depolandığı matris
$V$	Öbek merkezi değeri
$\beta_i$	Her bir $i$ fonksiyonu için $i=1, \dots, c$ 'ye kadar regresyon katsayıları
$E$	Durdurma sabiti
$\hat{\alpha}$	$\alpha$ gerçek kesim noktasının tahmini
$\hat{\beta}$	$\beta$ gerçek parametresinin tahmini
$\Phi_i$	Girdi değişkenleri, ilgili öbekteki üyelik değerleri ve bunların dönüşümlerini içeren matris
$\widehat{\omega}_t$	Girdi uzayının üyelik değerlerini kullanan yeni bir uzaya eşlemeleri olan regresyon katsayıları

## 1. GİRİŞ

Günümüz dünyasında belirli girdi değerleriyle doğrudan ya da kısmi ilişki içinde çıktı veren, geleceğe ait bilgileri tahmin edebilme yeteneği sağlayan, başka bir deyişle kurumların ya da ileri teknoloji cihazların proaktif hareket etmelerine olanak sağlayan modeller üzerine kurulmuş sistemlerin kullanım alanları genişlemiş ve varlıkları çok büyük önem kazanmış olup buna yönelik çok farklı alanlarda çalışmalar yürütülmektedir.

Bu husus sadece dinamik ekonomik aktörlerin ani tepkilerine ayak uydurmaya çalışan günümüz müteşebbislerinin değil, maddenin doğasının daha derinlemesine keşfi ile kimya, biyoloji, genetik gibi araştırma alanlarında hatta var olanların yapısını takliden üretilmiş ileri teknoloji cihazların da işletim sistemlerinin temelinde yer almaktadır.

Bu çalışma temelde bir ya da birden çok farklı veri serisinin yine bir ya da birden çok çıktı üzerindeki etkisini belirlemeye yönelik modeller oluşturulması hedeflenmekte, oluşturulan modellerin geçmiş veriler aracılığı ile iyileştirilmesi ve gelecek çıktıların daha iyi tahmin edilmesini amaçlamaktadır.

Bu kapsamda bilinen ve kullanımı en yaygın uygulama regresyon modelleri olup regresyon modellerinin girdileri ve oluşturulan modelin yapısı üzerinde yapılan detay çalışmalar ile daha iyi modellerin elde edilmesi mümkün olmuştur.

Diğer yandan verilerin sınıflandırılması ve çıktı tahmininde son elli yıl içinde verilerin aldıkları değerlere göre birer öbeğe dâhil ettiği; bu öbek merkezlerini referans alarak farklı algoritmalar aracılığı ile hesaplanan üyelik değerlerinin kullanıldığı Bulanık Sistem uygulamaları ile daha hassas sınıflandırılabilmiş, daha tutarlı çıktı tahminleri veren modeller oluşturulabilmiştir.

Bulanık Mantık kapsamında yer alan birden çok yaklaşımdan gözlemlendiği üzere verilerin öncesinde farklı yöntemler ile sayısı belirlenen öbeklere dağılımı ve bu öbeklerin merkezlerine uzaklıkları aracılığı ile hesaplanan üyelik değerlerinin hesaplanması; bu üyelik değerleri ve transformasyonlarının tahmin modeline eklenmesi ile elde edilen Bulanık Regresyon Modelinin [1] çıktı tahmininde klasik regresyon yaklaşımlarına kıyasla daha iyi sonuç verildiği tespit edilmiştir.

## 2. BULANIK SİSTEMLERİN GELİŞİMİ

Gelişen teknoloji ve maddenin doğasının derinlemesine keşfi sebebi ile hali hazırda var olan ya hep ya hiç başka bir deyişle sıfır- bir karar mekanizmaları mevcut sistemleri modelleme de yetersiz kalmaya başlamış ve gerçek hayat problemlerinin modellenmesini gerçek dışılığa itmiştir.

Bu duruma cevaben gelişen Bulanık Sistemler verileri sıfır- bir gibi ya da siyah –beyaz gibi kesin sınıflara ayırmak yerine verilerin belirli sınıflara belli bir dereceye kadar aitliği hususunu gündeme getirmiş ve temel kabul etmiştir.

Bu temel verilerin sınıflara aitliklerini ölçülebilir ve dolayısıyla iyileştirilebilir hale getirmek amacı ile rakamsal olarak ifade edilmesini gerektirmiş ve bu sebeple her verinin bir sınıfa aitliğini temsil eden “üyelik değerini” hesaplamak üzere çok sayıda yöntem ve teknik geliştirilmiştir.

İlk olarak çalışmalara konu olan Bulanık Sistemlerde sınıfların ve verilerin bu sınıflara aitlik derecelerinin belirlenmesinde uzman görüşü esas alınmış, klasik modellere kıyasla daha iyi modeller oluşturulmasına karşın modelin performansında uzmanın görüşünü belirtmekteki nesnelliği ve başarısı da modelin performansını doğrudan etkileyen bir parametre olarak karşımıza çıkmıştır. Uzman görüşü birçok başka alanda da kural tabanı olarak karşımıza çıkmakta olup bu kurallar sistemi tanıyan bilen uzmanlarca oluşturulup modele dâhil edilir.

Bulanık sistemlerin gelişimine baktığımızda ise zaman içinde bu tip modellerin yerini uzman görüşünün ve etkisinin en aza indirildiği ve hatta yok edildiği sistemlere bıraktığı görülmektedir. Bu yeni tip bulanık sistemlerde girdi ve çıktı arasındaki ilişkileri belirlemede kural tabanı yerine bulanık fonksiyonlar kullanıldığından uzman görüşünün sebep olduğu öznellikten uzaklaşmayı başarmış modeller elde etmek mümkün olup, daha nesnel modeller oluşturulabilmiştir.

Başka bir deyişle literatürde bu yeni tip “Bulanık Sistemler” genellikle öncüllerin analizi sonucu elde edilen bulanık kuralların; ardılların (yani çıktılar) tahmin edilmesinde kural tabanlı, fonksiyonlar kullanılarak ya da tamamen kendi kendine öğrenen sistemler aracılığı ile ifade edilebildiği sistemler olarak tanımlanmıştır.

Farklı yaklaşımlar kapsamında ele alınan modellerde fonksiyon doğrusal veya doğrusal olmayan bir yapıda olabilir. Örneğin ilerleyen bölümlerde tanımlanacak olan Zadeh [2] yaklaşımında hem girdiler hem çıktılar kural tabanlı iken sonradan geliştirilen TSK [3] tipi yaklaşımda kural tabanlılık yumuşatılmış olup ardıl kısımları girdi ve çıktı değişkenleri arasındaki bir fonksiyon ile gösterilir.



Bir bulanık kural tabanı yapısında her öncülün etkileşimsiz olduğu varsayılır, ya da öncüller birbirleri ile etkileşimsiz olacak şekilde seçilmeye çalışılır, dolayısıyla her bir girdi değişkeni için birbirinden farklı bulanık kümeler tanımlanır. Tüm yaklaşımları genel olarak özetlemek için ise Şekil 2.1 açıklayıcı olacaktır.



Şekil: 2.1. Bulanık sistemlerin genel yapısı

Bulanık sistem model yapılarının arasında Zadeh [2], Takagi-Sugeno [3] ve Mizumoto [4] tarafından önerilen yaklaşımlar en ünlü ve en çok kullanılanlarıdır. Yıllardır, bu temel bulanık sistemler bir takım değişiklikler ile yeniden önerilmekte ve kullanılmaktadır.

## 2.1. Zadeh [2] Bulanık Kümeler Teorisi

Zadeh [2] Kural Tabanlı yaklaşımı Bulanık Sistemlerin temelini oluşturmaktadır, bu yaklaşımda temel olarak verilerin belirli kümelere ait ya da ait değil şeklinde sınıflandırılması yerine bir dereceye kadar ait olması söz konusudur. Bulanık Kural Tabanlı yapıya sahip olan bu yaklaşımda her bir modelde  $c$  adet kural vardır. Her bir verinin kural tabanlı bu kümelere aitliği söz konusu olup üyelik değeri olarak ifade edilir. Bu yaklaşımda girdiler de çıktılar da kural tabanlıdır. Sistemde uzman görüşünün var olması hasebiyle model öznelik barındırır.

Zadeh'in [2] bulanık kural tabanlı bulanık sistemi aşağıdaki formül aracılığı ile özetlenmektedir.

$$R: \text{KURAL}_{i=1}^c [\text{EĞER } \text{VE}_{j=1}^{nv} (x_j \in X_j \text{ } \alpha_i A_{ji}) \text{ İSE } y \in Y, \alpha_i B_i ] \quad (2.1)$$

Sistemi genel olarak tasvir eden aşağıdaki açıklamalar faydalı olacaktır:

- c sistem modelindeki kuralların sayısını
- $x_{ij}$  inci girdi değişkeni;  $j=1 \dots, nv$ ; nv toplam girdi sayısını
- $X_j$ ;  $x_j$  lerin tanım kümesi
- $X_j$  üyelik fonksiyonundaki i kuralına göre  $A_{ji}$  girdiyle alakalı dilsel etiketi
- y çıktı değerini
- Y çıktı olan y'lerin tanım kümesini
- i inci kuralın üyelik fonksiyonuna göre  $B_i$  çıktı değeri ile ilgili dilsel etiketi
- VE üyelik değerlerinin kural tabanları aracılığıyla girdiler ile ilişkilerinin derecelerini ifade eden mantıksal bağlacı
- İSE mantıksal çıkarım bağlacı
- KURAL bulanık sistemdeki kuralları birbirine bağlayan mantıksal bağlacı
- isr: Zadeh tarafından literatüre kazandırılan bir değişkenin 0 ya da 1 gibi kesin değil bulanık olduğunun tanımını ifade etmektedir.

Model kullanılmadan önce sistemdeki kural sayısının ( $c^*$ ) ve her bir üyelik değeri için üyelik fonksiyonun çeşidinin en başında belirlenmesi uygun olacaktır.

## 2.2. Mizumoto [4] Bulanık Kural Tabanlı Algoritma

Mizumoto [4] bulanık kural tabanı yapıları Zadeh tarafından öne sürülen bulanık kural tabanı yapılarının basitleştirilmiş bir hali olarak tasvir edilmektedir. Burada her kural bir  $B_i$  bulanık kümesi yerine, bir skalar  $b_i$  değeri ile gösterilir. Bu tip çıkarsama yapıları 2.2'de tanımlanmıştır.

$$R: \text{KURAL}_{i=1}^c [\text{EĞER } \text{VE}_{j=1}^{nv} (x_j \in X_j \text{ } \alpha_i A_{ji}) \text{ İSE } y_i = b_i ] \quad (2.2)$$

Bu formülde yer alan  $b_i$ , i inci kural ile ilişkili bir değerdir. Mizumoto [4] bulanık sistemini kullanırken tanımlanacak çıkarsama parametreleri c ile her kuralda her bir girdi değişkeni için oluşturulan üyelik fonksiyonları kapsamında her kural ile ilişkili skalar değerler;  $b_i$ ,  $i=1, \dots, c$  ile ifade edilir.

Genelleştirilmiş bulanık sistem yapısı aşağıdaki gibi açıklanabilir. Bu kapsamda;

- $x' = (x_1', x_2', \dots, x_{nv}')$  çıktı değerleri bilinmeyen bir girdi vektörünü
- $X_j$ ;  $x_j$ 'lerin  $j=1, \dots, nv$ 'ye kadar oluşturduğu girdi veri kümesini
- $A_{ij}$ 'de  $x_j$ 'lerin  $j$  kuralında üyelik fonksiyonu kapsamında oluşturduğu üyelikler kümesini
- $C$ : kural tabanlı bulanık sistemde var olan toplam kural sayısını
- $nv$ : girdi çeşidi sayısını (girdi veri boyutunu) ifade etmektedir.

$(x_k, y_k) = (x_{k1}, x_{k2}, \dots, x_{k nv}, y_k)$ ;  $k=1, \dots, n$ , bulanık kural taban yapısı ve deney verisi kullanılarak yaklaşık  $x'$  değerleri tayin edilmektedir.

Bu tanımlamalar eşliğinde takip edilmesi gerekli adımlar sırası ile şu şekildedir:

**Bulanıklaştırma:** Her bir girdi değerine  $\mu_i(x_j')$ , her  $i=1, \dots, c$ ; her  $j=1, \dots, nv$  formülüne göre üyelik değerlerinin atanması demektir. Her bir girdiye karşılık üyelik değerleri oluşturulur ve sistematik olarak bir  $A_{ij}$  matrisi içinde bu üyelik değerleri depolanır.

**Öncüllerin Derlenmesi:** Bulanıklaştırma adımında hesaplanan üyelik değerleri tek bir ateşleme derecesi dikkate alınarak hesaplanmıştır. Ateşleme derecesi modern ve son yıllarda geliştirilen bulanık sistemlerdeki bulanıklık mertebesine yakın bir parametre olarak belirtilebilir. Derleme işlemi bulanık operatörler kullanılarak yapılır ki bu bulanık operatörlerden T normunda yer alan MIN operatörü:  $\wedge$  ile ifade edilir. Herhangi bir girdi vektörü için kullanılması gereken bulanık ateşleme derecesi şu formül ile hesaplanır:

$$\tau_i(x') = VE_{j=1}^{nv} \mu_i(x'_j), \forall i = 1 \dots c \quad (2.3)$$

**Çıkarım:** (Model Çıktısı Olarak Bulanık Kümelerin Tanımlanması): Ateşleme derecesi öncüller ve girdi değişkenleri arasındaki uyumluluğun bir ölçüsünü gösterse de, aynı zamanda son model çıktısı olan bulanık kümeyi tanımlayan her bir kuralın çıktı bulanık kümesinin katkı seviyesini de temsil etmektedir. Mamdani [5] tipi çıkarsama metodunda MIN işlemi modelin bulanık çıktı kümesinin üyelik fonksiyonlarının belirlenmesinde kullanılmakta olup bu durumda; VE bağlacı ÇIKARIM bağlacı yerine kullanılabilir. Bu durumda üyelik değerlerinin hesaplanmasında kullanılan formül ise şu şekildedir:

$$\mu_i'(y) = \tau_i(x') \vee \mu_i(y) = \tau_i(x') \wedge \mu_i(y), \forall i = 1 \dots c \quad (2.4)$$

**Ardılların Derlenmesi:** Her bir kural için oluşturulan üyelik değerleri kümesi ÇIKARIM adımında oluşturulmaktadır. Bu adımda ise her bir ayrı küme altında toplanılan üyelik derleri KURAL operatörü aracılığı ile tek bir bulanık bütünleşik çıktı kümesi içine yerleştirilir. Mamdani [5] tipi çıkarsama metodunda KURAL aynı zamanda VEYA operatörü yerine geçebilmekte olup aşağıdaki şekilde hesaplanır.

$$\mu'(y) = \bigvee_{i=1}^c \mu_i'(y) \quad (2.5)$$

**Berraklaştırma:** Bulanık çıkarsama metodunun son adımı olan bu adımda Tip 1 bulanık çıkarsamadan kesin bir çıktı elde etmek amaçlanmaktadır. Bu adımda sıklıkla kullanılan berraklaştırma metodu Ağırlık Merkezi Yöntemi olup şu şekilde tanımlanır:

$$y' = \frac{\int_y y \mu'(y)}{\int_y \mu'(y)} \quad (2.6)$$

Sonuç olarak Mizumoto [4] yaklaşımının da kural tabanlı olması hasebiyle içinde öznellik barındırdığı aşikârdır.

### 2.3. Takagi & Sugeno & Kang [3] Kural Tabanlı Algoritması

Takagi-Sugeno-Kang (TSK) [3] bulanık kural tabanı sistemleri yaygın olarak kullanılan ve araştırılan bulanık çıkarım sistemlerinden biri olup, Zadeh'in [2] bulanık kural tabanı ile karşılaştırıldığında kısmen farklıdır. Tek fark kural tabanı yapısının çıktı kısmıdır. Zadeh [2] bulanık kuralların hem girdileri hem de çıktıları göstermek için bulanık kural kümeleri kullanmışken, TSK [3] bulanık kural tabanında, girdi kısım toplama operatörü ile karakterize edilirken çıktı ise regresyon doğrusu ile karakterize edilir. TSK [4] için bulanık kural tabanı yapısı aşağıda tasvir edilmektedir.

$$R: \text{KURAL}_{i=1}^c [\text{EĞER } \bigvee_{j=1}^{nv} (x_j \in X_j \text{ isr } A_{ji}) \text{ İSE } y_i = a_i x^T + b_i] \quad (2.7)$$

Bu formülasyonda;

- $a_i$  ve  $b_i$ ;  $i$  inci kuralla ilgili regresyon katsayılarını
- $y_i$ ;  $i$  inci kurala ait çıktı değerlerini,
- EĞER İSE  $y_i$  değerlerini her bir kurala göre ve bulanıklık ateşleme derecesine göre ağırlıklandırarak her bir kurala göre çıktı değerlerini bulmayı sağlayan bağlacı
- KURAL ise modelin çıktılarının her bir kurala göre ağırlıklı ortalamasını alarak model kapsamında oluşan bulanık çıktıları derleyen bağlacı temsil etmektedir.

TSK [3] bulanık kural tabanı metodunun çıkarsama parametreleri;

- $c$ ; her kuraldaki her girdi değişkeninin üyelik fonksiyonu,
- $a_i$ ; her kuralın regresyon doğrusu katsayıları
- çıkarsama metodunda kullanılan VE, EĞER İSE ve KURAL bağlaçlarıdır.

TSK [3] kural tabanlı bulanık sistemi doğrusal regresyona bağlı olmasına karşılık doğrusal olmayan polinomal fonksiyonlarda da başarı ile uygulanmış ve tutarlı sonuç vermiştir. TSK tipi Kural Tabanlı Bulanık sistemlerde girdi (öncül) ve çıktı (ardıllar) arasındaki ilişki bir fonksiyon olarak gösterilmektedir.

Özetle TSK [3] bulanık kural tabanlı sistemlerde;

- Kural tabanı vardır.
- Çıktılar regresyon modeli şeklinde, girdiler ise yine kural tabanlı olarak elde edilmektedir.
- Uzman görüşünün etkisi azaltılmış, kendi kendine öğrenebilen modellemeye ilk adım atılmıştır.

#### **2.4. Bulanık Öbek Ortalaması Yaklaşımı [6]**

Bezdek tarafından 1981 yılında literatüre kazandırılan Bulanık Öbek Ortalaması Yaklaşımı (BCO Öbekleme Algoritması) [6] bu zamana kadar yapılan çalışmalarda en çok kullanılan yöntem olup birçok başka yaklaşıma da temel teşkil etmektedir. BCO Öbekleme Algoritması kapsamında çıktıları tahmin etmede kullanılacak olan girdilerin toplamda kaç adet öbekte sınıflandırılacağı en başında belirlenmiş ya da en azından sabit bir sayıda sabitlenmiş kabul edilmektedir. Böylelikle de verilerin öbekler arasında dağılımı  $X = \{x_1, x_2, \dots, x_n\}$  olacak şekilde  $c$  sayıdaki öbeğe gerçekleşir. Gerçek dünyada bu tip önceden belirlenmiş parametrelerin algoritmada kullanımı modelin etkinliğini ve çıktıların tutarlılığını olumsuz etkilediğinden başlangıçta öbek sayısını belirlemede kullanılmak üzere oluşturulmuş öbek

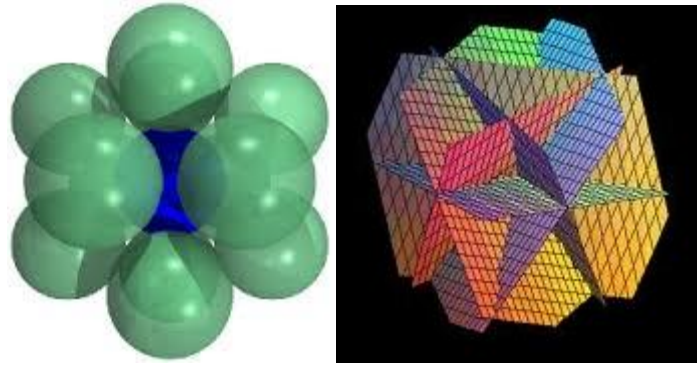
sayısı belirleme indekslerinin oluşumuna BCO Öbikleme Algoritması sebep olmuştur denilebilmektedir.

BCO Öbikleme Algoritması [6] yaklaşımı ilerleyen bölümlerde algoritmanın adımlarını içerecek şekilde açıklanacağından bu bölümde BCO Öbikleme Algoritması'nın temel teşkil ettiği Bulanık Öbek Ortalaması Regresyon Modeli (FCRM (Fuzzy C- Regression Model Clustering Algorithm)'den [7] ve iki yaklaşımın benzer ve farklı yönleri belirtilecektir.

Hathaway ve Bezdek [7] tarafından 1993 yılında ortaya atılan FCRM yaklaşımın amacı birçok başka öbikleme algoritmasında da görüldüğü üzere benzer özellik gösteren verilen aynı öbekler içinde sınıflandırılmasını ya da başka bir deyişle toplanmasını sağlayacak bir algoritma oluşturmaktır.

BCO Öbikleme Algoritması ve FCRM [6,7] yaklaşımları arasındaki farklılıklar şu şekilde özetlenebilmektedir:

- BCO Öbikleme Algoritmasında öbekler hiper-küre olarak tanımlanmakta olup FCRM'de öbekler hiper-düzlem olarak tanımlanmaktadır.



Şekil 2.2. Hiper-küre (solda) ve hiper-düzlem (sağda) illüstrasyonu [8]

- BCO Öbikleme Algoritmasında öbek merkezleri  $v_i$  noktaları iken, FCRM 'de öbekler  $nv$  boyutlu girdi ve tek bir çıktı aracılığı ile temsil eden hiper-düzlemelerdir. FCRM'e temel teşkil eden ilgili model 2.8'de belirtilmiştir.

$$y_i = \beta_i^0 + \beta_i^1 x_1 + \dots + \beta_i^{nv} x_{nv} \quad (2.8)$$

Bu model kapsamında  $\beta_i$  her bir  $i$  fonksiyonu için  $i=1, \dots, c$ 'ye kadar regresyon katsayılarını temsil etmektedir. BCO Öbekleme Algoritması öbek merkezlerini her bir veri vektörünün ortalamasını üyelik değerleri ile birlikte ortalama değer olarak bulurken; FCRM, öbekleri ağırlıklı en küçük kareler regresyon algoritması ile tayin etmektedir.

Sonuç olarak BCO Öbekleme Algoritması ve bu algoritmayı temel alarak geliştirilmiş olan FCRM yaklaşımları için aşağıdaki saptamalar yapılmıştır:

- Girdiler öbeklere; öbek merkezlerine uzaklıklarını esas alan üyelik derecelerine göre dağılmaktadır.
- Üyeliklerin hesaplanmasında bulanıklık mertebesi yer almaktadır.
- Başlangıçta üyeliklere tesadüfi bir merkez atanmakta olup sonrasında her iterasyonda bu merkezler/üyelikler iyileşmektedir.
- Üyeliklerin öbek merkezlerine uzaklıklarının hesaplanması esasına dayanmaktadır.

### **3. BCO ÖBEKLEME ALGORİTMASI**

Üyelik değerlerinin hesaplanması için bu tez çalışması kapsamında güncel yöntemlerden olan BCO Öbekleme Algoritması [6] seçilmiştir. Bu algoritmanın kullanımında öncelikle birden fazla girdisi olan ve en az bir çıktısı olan veri setinin belirlenmesi gerekmektedir. Ardından bu veri setinin kaç adet öbeğe dağıtılacağı belirlenmelidir. Diğer yandan algoritmanın sonucunu önemle etkileyen bir diğer girdi de bulanıklık mertebesi değeri olup, bir değerde karar kılınmalıdır. Bu belirlemelerin ardından üyelik değerlerinin hesaplanması için algoritmanın çalıştırılması mümkün olabilecektir.

Elde edilen üyelik değerleri her bir bulanıklık mertebesi ve öbek sayısı için ayrı bir modele temel teşkil edecek olup bir matriste depolanmalı, verinin kendisi, hesaplanan üyelikler ve bu üyeliklerin karesi, küpü, logaritması gibi farklı transformasyonları bir sonraki adım olan bulanık regresyon modellemesine hazır edilmelidir.

Algoritmanın uygulamasında kullanılacak her bir parametre ve bu parametrelerin belirlenmesinde dikkat edilmesi gereken hususlar bu bölüm kapsamında açıklanmıştır.

#### **3.1. Veri Seti ve Analizi**

Gelişen teknoloji ile biyolojik, genomik, tıp, iklim, sosyal medya ve çevre bilimleri alanlarında çok büyük veri setleri kolayca toplanabilmektedir. İstatistiksel öğrenme, bu veri setlerini anlamak ve anlamlandırmak için kullanılan istatistiksel metotlar bütünüdür. Karar destek mekanizmaları, yapay sinir ağları, karar ağaçları, genelleştirilmiş doğrusal modeller bu kapsamda kullanılan yöntemlerden bazılarıdır. Büyük verileri anlamlandırmak için minimum varsayım gerektiren ve verinin yapısına uygun olarak modelleme yapan parametrik olmayan farklı yöntemler de çalışılmaktadır.

##### **3.1.1. Zaman Serileri Analizi**

Risk yönetimi günümüzde oldukça önem kazanmıştır. Karşılaşılacak herhangi bir risk durumuna karşı tedbir almada, bu durumun oluşumunu öngörmeye kişi ve kurumlara riskin fırsata çevrilmesi olanağı sağladığından son derece önemlidir. Kantil ve ekspektil [9] regresyon modelleri ile bu tahminlerin yapılmasında kullanılan farklı yaklaşımların başında gelmektedir.

Ekonomik değişkenlerin öngörülmesi herhangi bir risk karşısında doğru pozisyonu almak için önemlidir. Zaman serileri modellemeleri ile gösterge niteliğindeki ekonomik verilerin öngörü-



tahmin modellemeleri yapılmaktadır. Örneğin; konut fiyat endeksinin öngörülmesi ekonomik olarak yaygın ve piyasalara yön veren önemli bir uygulamadır.

Bilindiği üzere modellemede esas olarak kullanılan veri setleri zaman serileri olup; zaman değişkeniyle ilişkili bir değişken hakkında elde edilen gözlem değerlerini zamana göre sıralanmış olarak gösteren serilere zaman serisi denilmektedir. Zaman serileri temel alınarak kurulan modellerin gelecek zamana ait tahmin edilmesi istenilen çıktı değerlerini tahmin etmede sıklıkla kullanıldığı gözlemlenmektedir. [10]

İlgili modellerde kullanılan zaman serileri genel olarak üç grupta incelenmektedir [11]:

**Gerçek Zaman Serileri:** Bu zaman serilerinde veri seti geçmişte gerçekleşen değerlerin gözlemlenmesi ve kayıt altına alınması ile oluşturulur. Modellemede sıklıkla kullanılan ve modelin tutarlılığını olumlu olarak etkilediği bilinen veri tipi gerçek zaman serileridir. Endeksin günlük açılış değeri veya Türkiye'deki Yeşilirmak nehrinin yıllık akış debisi verisinin 2 yıl boyunca kayıt altına alınmasıyla oluşturulan veri seti gerçek zaman serilerine örnek olarak verilebilir.

**Yapay Zaman Serileri:** Dağılımı ya da gerçekleşme durumu ön görülen verilerin çeşitli yöntemler ya da istatistik programları aracılığı ile türetildiği veri setleridir. Gerçek veri toplanmanın mümkün olmadığı, çok zor olduğu durumlarda ya da tercih edilmediği durumlarda oluşturulan tahmin modellerinde kullanıldığı gözlemlenmektedir.

**Karma Zaman Serileri:** Hem gerçek hem de yapay zaman serilerinin birlikte kullanıldığı melez yapılar olup çıktıyı istenilen tutarlılıkta tahmin etmede sadece gerçek ya da yapay verilerin yeterli olmadığı durumlarda kullanılması uygun kabul edilen veri setleridir. Verinin yapısı ve mevcut trendine göre tahmin modellerinde iyi sonuç verdiği zaman zaman gözlenebilmektedir.

Tüm tahmin modellerinde olduğu gibi BCO Öbekleme Algoritması ve Bulanık Regresyon Modellemesinde [1,6] de kullanılan zaman serilerinin birbirlerine bağımlı olması ya da yüksek korelasyon içinde olması mümkün olduğunca kaçınılmak istenilen bir durumdur. Yukarıda belirtilen kategorilerden birinde yer alan zaman serileri 3.1'de yer aldığı üzere her bir zaman serisi bir kolonda yer alacak şekilde bir matriste depolanır.

Örneğin; birinci zaman serisi:  $x_1$  ilk kolonda; ikinci zaman serisi  $x_2$  ikinci kolonda olacak şekilde belirtilmelidir.  $X = \{x_1, x_2, \dots, x_n\}$  olarak tanımlı n tane gözlemi gösterebilir. Her k

gözlemi,  $k=1, \dots, n$ 'e kadar devam eden  $nv$  boyutlu bir vektör  $x_k = [x_{1,k}, x_{2,k}, \dots, x_{nv,k}]^T \in R^{nv}$ .  $n \times nv$  boyutlu vektörlerin gösterimi aşağıdaki gibidir:

$$X = \begin{bmatrix} x_{1 \times 1} & x_{1 \times 2} & \dots & x_{1 \times nv} \\ \dots & \dots & \dots & \dots \\ x_{n \times 1} & x_{n \times 2} & \dots & x_{n \times nv} \end{bmatrix} \quad (3.1)$$

### 3.1.2. Basit Doğrusal Regresyon (BDR) [12]

Değişkenler arasındaki ilişkiyi incelemek bilimin uğraşlarından birisi olagelmıştır. Bunu doğal karşılamak gerekir. Çünkü gerek günlük hayatımızda gerekse bilimsel araştırmalarda karşılaştığımız sorunların çoğunluğu iki (veya daha çok) değişken arasında bir ilişki olup olmadığının saptanması ile ilgilidir. İki değişken arasında bir ilişki bulunup bulunmadığı, eğer varsa bu ilişkinin derecesinin saptanması da istatistiksel çözümlenelerde sık sık karşılaşılan bir sorundur. Değişkenler arasındaki ilişkinin incelenmesinde regresyon ilk aklı gelen tekniktir. İstatistiksel anlamda iki değişken arasındaki ilişki, bunların değerlerinin karşılıklı değişimleri arasında bir bağıllık şeklinde anlaşılır. Gerçekten  $X$  değişkeninin değerleri değişirken buna bağlı olarak  $Y$  değişkeninin değerleri de değişiyorsa, bu ikisi arasında bir ilişki bulunduğu söylenebilir. Regresyonda değişkenlerin bağımlı değişken ve bağımsız değişken ya da değişkenler olarak iki gruba ayrılması bir zorunluluktur. Bağımlı değişken, bağımsız değişken ya da değişkenler tarafından açıklanmaya çalışılan değişkendir. Regresyonda bağımlı değişken  $Y$  ve bağımsız değişken ya da değişkenlerde  $X$  ile gösterilir. Regresyonda, amaçlardan biri bağımlı değişkenle bağımsız değişken ya da değişkenler arasındaki ilişkilerin ortaya çıkarılmasıdır.

Örneğin  $Y$  ile  $X$  arasında  $Y_i = \alpha + \beta X_i + \varepsilon_i$  ( $i=1,2,3,\dots$ ) gibi doğrusal bir ilişki öngörülüyorsa ilk adım modelin bilinmeyen  $\alpha$  ve  $\beta$  parametrelerinin tahmin edilmesidir. Modelin bilinmeyen parametreleri tahmin edildiğinde bağımsız değişken ya da değişkenlerin farklı değerleri için bağımlı değişkenin alacağı değeri tahmin etmek regresyonda bir diğer amaçtır.

Regresyon analizi, bağımsız değişken sayısına göre;

- Basit regresyon analizi (Tek bağımsız değişken),
- Çoklu regresyon analizi (Birden çok bağımsız değişken),

Fonksiyon tipine göre;

- Doğrusal regresyon analizi,

- Doğrusal olmayan regresyon analizi (Eğrisel),

Verilerin kaynağına göre;

- Ana kütle verileriyle regresyon analizi,
- Örnek verileri ile regresyon analizi,
- Zaman serilerinde regresyon analizi (Eşleştirilmiş zaman serileri) şeklinde gruplandırılır.

Birçok istatistiksel çalışmada olduğu gibi regresyon analizinde de ana kütle verilerinin tümü yerine bu ana kütlede seçilen örnek verileri ile analiz yapılır. Daha sonra elde edilen sonuçlar ana kütledeki ilişkinin tahmininde kullanılır. Bilindiği gibi, ana kütle birimi sayısı çok fazla olduğundan, zamandan ve araştırma masraflarından tasarruf amacıyla tüm ana kütle birimleri yerine, bu ana kütlelerden tesadüfi olarak belirli sayıda birim (n) seçilerek istatistik analizler yapılır. Ana kütle ve örnek verileriyle yapılan istatistik araştırmalarda tekniklerinin uygulanmasında farklılık yoktur. Ancak teknikler uygulandıktan sonra örnekleme teorisinden yararlanılarak ana kütle parametrelerinin testleri ve tahminleri yapılır. Regresyon analizinde de uygulama aynı şekilde olmaktadır. Büyük harfler ana kütleyle, küçük harfler ise örneğe ait verileri ve istatistik ölçüleri göstermekte kullanılmaktadır. Basit doğrusal regresyon analizi, Y bağımlı değişkeninin tek bir bağımsız (açıklayıcı) değişken X ile arasındaki ilişkinin doğrusal fonksiyonla ifade edilmesine dayanmaktadır. Basit doğrusal regresyon modeli, tek bir serbest değişken içeren 3.2’de belirtilen modeldir.

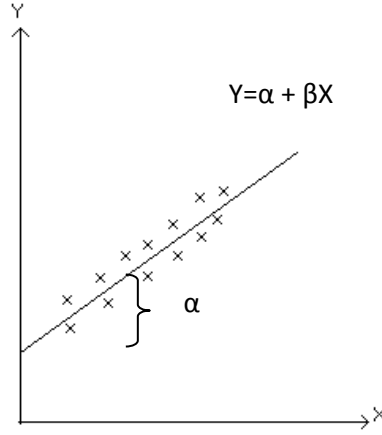
$$Y_i = \alpha + \beta X_i + \varepsilon_i \quad (3.2)$$

Bu modelin  $\alpha$  ve  $\beta$  parametrelerini bulmak için X serbest değişkeni, Y bağlı değişkeni ve  $\varepsilon$  hata terimi ile ilgili gözlemlere gerek duyulur.

Ana kütle içinde birer  $\alpha$  ve  $\beta$  değeri varken, bu ana kütlede çekilen her bir örneklem için ayrı birer  $\hat{\alpha}$  ve  $\hat{\beta}$  elde edilmektedir. İşte bu  $\hat{\alpha}$  ve  $\hat{\beta}$  normal bölünmeye sahip olup beklenen değerleri sırasıyla  $\alpha$  ve  $\beta$ ’dir. Uygulamada tek bir örneklem alınmakta ve bu örneklem yardımıyla ana kütle parametreleri tahmin edilmektedir.

$\alpha$  doğrusal fonksiyonun sabitidir. X= 0 olduğunda regresyon doğrusunun dikey eksen olan Y ile kesiştiği noktayı göstermektedir.  $\beta$  ( $\beta_{yx}$  ile de gösterilebilir) ise doğrusal fonksiyonun eğimidir. Regresyon analizinde bağımsız değişken X deki bir birimlik değişimin bağımlı değişken Y’de (Y cinsinden) ne kadarlık bir değişime yarattığını gösteren regresyon katsayısıdır. Fonksiyon tipinin belirlenmesi için regresyon analizine serpilme (saçılım)

diyagramı çizilerek başlanır. Aşağıdaki serpilme diyagramında gözlem noktalarının dağılımının doğrusal bir eğilimde olduğu açıkça görülmektedir.



Şekil 3.1. Basit doğrusal regresyon doğrusu

$\alpha$  ve  $\beta$  parametrelerinin gösterdiği grafikte regresyon doğrusunun eğiminin pozitif olduğu anlaşılmaktadır.  $\beta$ 'nin işareti iki değişken arasındaki ilişkinin yönünü göstermektedir. Her iki değişken birlikte artıyor veya azalıyor  $\beta$ 'nin işareti pozitif (+), değişkenlerden biri artarken diğeri azalıyor  $\beta$ 'nin işareti negatif (-) olacaktır.  $\beta$ 'nin sıfır (0) olması ise iki değişkenin arasında bir ilişki olmadığını göstermektedir. Sıfır (0)'dan farklılık ise iki değişken arasında belirli bir ilişkinin varlığını ifade etmektedir. Regresyon katsayısının alt sınırı (0) vardır, ancak belirli bir üst sınırı yoktur. Bu nedenle sadece regresyon doğrusuna bakarak ilişkinin gücü hakkında kesin bir şey söylemek mümkün değildir.

Regresyon modeline açıkça dâhil edilemeyen diğer değişkenleri temsil etmek üzere  $Y_i = \alpha + \beta X_i + \varepsilon_i$  modelinde yer verilen  $\varepsilon$  hata terimini gözlemek hiçbir zaman mümkün olmaz. Dolayısıyla  $\varepsilon$  hata terimi hakkında aşağıda değineceğimiz bazı varsayımları yapmak zorunlu hale gelir.

“Y ve X arasındaki gerçek ilişki” ;  $Y = \alpha + \beta X_i + \varepsilon$  iken “gerçek regresyon doğrusu” :  $E(Y_i) = \alpha + \beta X_i$  eşitliğidir.

Öte yandan, “Tahmin edilen ilişki” :  $\hat{Y}_i = \hat{\alpha} + \hat{\beta} X_i + e_i$  şeklinde gösterilmektedir. Tahmin edilen regresyon doğrusu ise 3.3'te belirtilmektedir.

$$\hat{Y}_i = \hat{\alpha} + \hat{\beta}x_i \quad (3.3)$$

Bu eşitliklerde:

$Y_i$  Y değişkeninin gözlenen değerini,

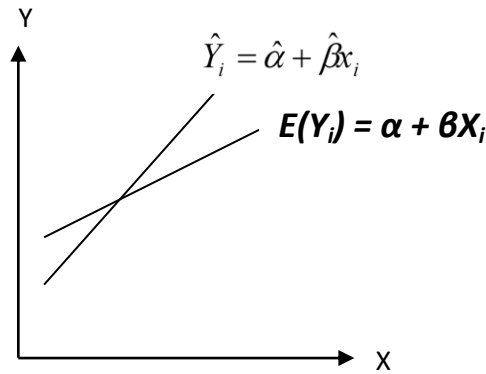
$\hat{Y}_i$  X değişkeninin belli bir değeri veri iken Y değişkeninin tahmin edilen değerini,

$\hat{\alpha}$   $\alpha$  gerçek kesim noktasının tahminini,

$\hat{\beta}$   $\beta$  gerçek parametresinin tahminini,

E  $\varepsilon$  hata teriminin gerçek değerinin tahminini ifade eder.

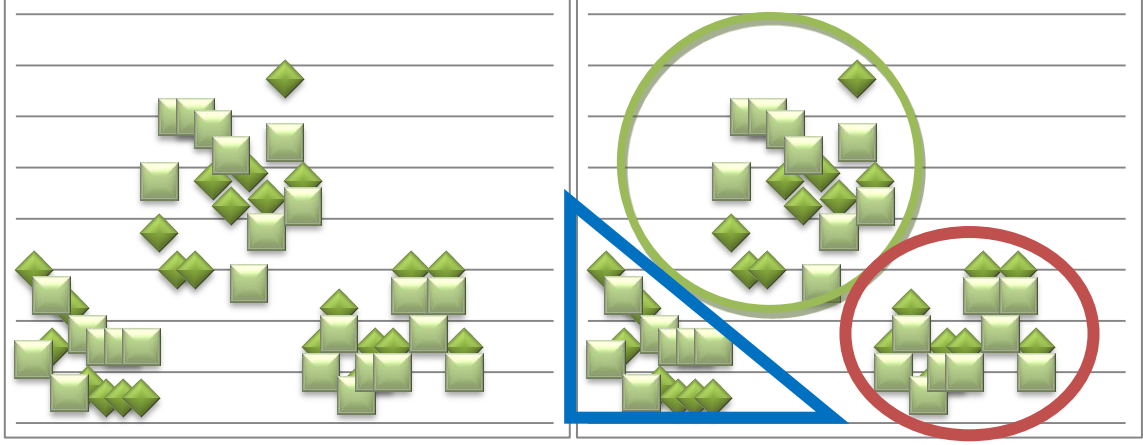
Gerçek ve tahmin edilen regresyon doğruları Şekil 3.2’de yer alan grafikte gösterilmiştir:



Şekil 3.2. Gerçek ve tahmin edilen regresyon doğruları örneği

### 3.2. Öbek Sayısı ve Hesaplama Yöntemleri

Öbek sayısı BCO Öbekleme Algoritması [6] ile üyeliklerin hesaplanmasında ve dolayısı ile üyelikler ve transformasyonlarının yer aldığı bulanık regresyon modelinde modelin tutarlılığını doğrudan etkileyen önemli bir parametredir.



Şekil 3.3. Verilerin saçılım grafiği (solda); verilerin öbeksizmesinin gösterimi (sağda)

Çıktı tahmininin yapılacağı modelde girdi olarak yer alan zaman serilerinin her birinin zamana göre değişimlerine göre birbirlerine ne denli benzer ya da ayrık olduğunun temsilinde üyeliklerin hesaplanmasından önce belirlenen öbek sayısının belirleyici rolü vardır. Model oluşturulmadan önce üyeliklerin hesaplanması esnasında elde olan verilerin tam olarak kaç öbek içerisinde dağılacığı belirlenirken aşağıda sıralanan hususlar dikkate alınmalıdır [13]:

- Problemin ve tahmin modelinin yapısı gereği verilerin önceden belirlenen sayıda öbek içerisinde dağılması istenebilir ya da uygun görülebilir.
- Veriler mantığa uygun bir aralık içerisinde kalan birden çok sayıda öbek içine dağıtılarak her bir öbek sayısı için üyelikler hesaplanabilir ve bu üyelikler ile ayrı ayrı oluşturulan modellerin tutarlılığı en yüksek olan model kabul edilebilir. Öbek sayısı aralığı belirlenirken işlem yapma süresi ve kolaylığı da göz önünde bulundurularak mümkün olduğunca az öbek ile işlem yapılmaya çalışılmalıdır.
- Veriler öbek sayısı doğrulama indeksleri kullanılarak sezgisel olarak modelde tutarlı sonuç vermesi umulan bir öbek sayısı belirlenebilir.

Öbek sayısı belirlenmesine rehberlik eden bazı indeksler de literatürde yer almaktadır. Bu indeksler belli bir aralık içerisinde tarama yaparak o aralıktaki yerel optimum öbek sayısı değerini belirlemeye yaramakta olup kesin olarak modeli en yüksek tutarlılığa taşıyacak en iyi öbek sayısını belirlemede yeterli değildir. Bu durum geçmişte yapılan deneysel çalışmalarda tespit edilmiştir.

Birçok öbek geçerlilik fonksiyonu, iki farklı öbeksizme kavramının birleştirilmesiyle oluşturulur [14]:

- **Yoğunluk:** Her öbeğin öbek elemanları arasındaki benzerliği ölçer. Çoğu geçerlilik indeksi, yoğunluğu ölçmek için öbek-içi uzaklıkları kullanır.
- **Ayrılabilirlik:** Her öbeğin tek tek benzemezliğini ölçer. Çoğu geçerlilik indeksi, ayrılabilirliği ölçmek için öbekler-arası uzaklıkları kullanır.

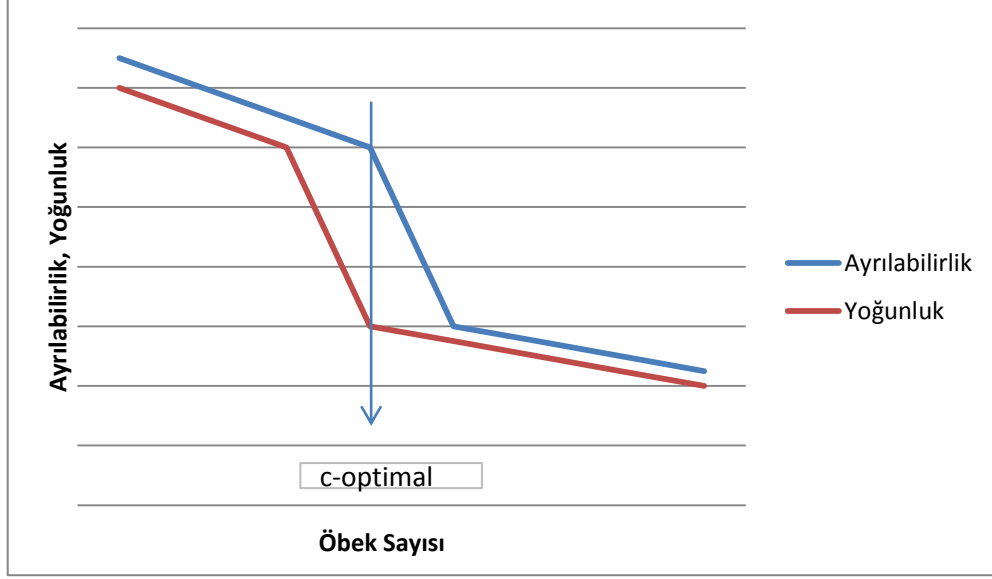
Bir öbekleme algoritmasının yoğunluğu küçük, ayrılabilirliği büyük olduğu zaman etkili olduğu gösterilmiştir [14]. Bu iki kavramın birbirine bağlı olmasına dayanarak, öbek geçerlilik indeksleri oran tipi ve toplam tipi olmak üzere iki farklı tipe ayrılmaktadır. Geçmiş çalışmalar kapsamında kullanılan bazı öbek sayısı doğrulama indeksleri aşağıda detaylı olarak açıklanmıştır.

### 3.2.1. XB (Xie & Ben) İndeksi

*Oran tipi* geçerlilik indeksleri, Xie-Beni [15] indeksi gibi, yoğunluk-ayrılabilirlik oranının ölçülmesi ile oluşturulur. İyi bilinen *oran tipi* (yoğunluk/ayrılabilirlik) öbek geçerlilik indekslerinden biri olan XB öbek geçerlilik indeksi [15] şu şekilde ifade edilir:

$$XB(c) = \frac{(\sum_{i=1}^c \sum_{k=1}^n \mu_{ik}^2 d(x_k, v_i)^2)}{\min_{i,j \neq i} d(v_i, v_j)^2}, d(x_k, v_i) = \|x_k - v_i\|^2 \quad (3.4)$$

Burada  $x_k \in \mathfrak{R}^{nv}$ ,  $k$ 'inci girdi vektörünü  $k=1, \dots, n$  ve  $v_i \in \mathfrak{R}^{nv}$ ,  $i, j=1, \dots, c$ ,  $nv$  boyutlu bir vektör olarak öbek merkezini temsil eder. XB,  $c$  sayısı toplam veri örneklerinin sayısı olan  $n$ 'ye yakın olursa monoton olarak azalır. Kim ve Ramakrishna, 2005 yılında yaptığı bir çalışmada [14] öbek sayısının değişen değerleri için BCO algoritmasından elde edilen öbekler arasındaki yoğunluk ve ayrılabilirliğin davranışını ve ilişkisini incelemişlerdir. [14]'de, yoğunluk ve ayrılabilirlik arasındaki ilişki grafikler ile gösterilir. Kim ve Ramakrishna [14]'ün genellemesine göre, yoğunluk hızlı şekilde arttıkça  $c$  değeri  $c_{optimal}$ 'den  $c_{optimal} - 1$ 'e düşmektedir. Bu da,  $c < c_{optimal}$  için yoğunluk büyük ve  $c > c_{optimal}$  için yoğunluk küçük olacak anlamına gelir. Her nesne kendi öbeğine eşit olduğunda yoğunluk sıfır olur. Yoğunluktaki ani bir düşüş  $c_{optimal}$ 'in göstergesidir.



Şekil 3.4. Oran tipi ÖGİ indeksinin yoğunluk ve ayrılabilirliklerinin gösterimi

BCO Öbekleme Algoritmasının sonuç analizinde, bazı öbeklerin diğerlerinden daha az/çok yoğunluğa sahip olmasından dolayı her öbeğin farklı yoğunluk değerleri olduğu gözlemlenebilir. Öbeklerin sayısı arttıkça/azaldıkça, bu öbeklerin yoğunluklarındaki değişiklik diğer öbeklerden farklı (daha büyük ya da küçük) olacaktır. XB geçerlilik indeksinde, bütün öbek yapısının yoğunluğu her öbeğin yoğunluğunun ortalaması Eşitlik 3.4'teki pay formülü aracılığı ile belirlenir. Fakat, ortalama almak öbeklerin bazılarının yoğunluk değerlerindeki büyük değişikliklerin etkisini bastırabilir. Bu değişikliklere genelde normalden az (ya da çok) öbek sayısına sahip BCO modelleri sebep olur. Bu nedenle, en iyi öbek sayısının belirlenmesi sırasında meydana gelen bu büyük yoğunluk değeri kaymalarını kullanmak için en iyi yol, öbeklerin en büyük yoğunluklarını ölçerek modelin yoğunluğunu yorumlamaktır.

Diğer yandan, Şekil 3.4'den  $c \neq c_{optimal}$  ifadesi sağlandığında yoğunluk ve ayrılabilirliğin bağımlı değişimlerinin bir şekilde benzer olduğu gözlemlenir. Bu nedenle, bunların *oran tipi* geçerlilik indekslerine etkisi de bir şekilde benzer olmalıdır. Kim tarafından yapılan[16]'daki tartışmayı da düşünerek XB geçerlilik indeksinin geliştirilmiş bir versiyonu şu şekilde önerilir:



$$XB^*(c) = \frac{\max_{i=1,\dots,c} \left\{ \frac{\sum_{k=1}^n \mu_{ik}^2 \|x_k - v_i\|^2}{n} \right\}}{\min_{i,j \neq i} \|v_i - v_j\|^2} \quad (3.5)$$

[14]'de  $XB^*$  indeksinin  $XB$  indeksinden daha etkili olduğu ispatlanmıştır, çünkü  $XB^*$  ile öbekler büyük yoğunluk değerlerine sahip olduğunda belirlenir. Bu bilgi ile öbekleme yapısındaki belirsizlikler gözlemlenerek en iyi öbek sayısı tespit edilir. Böylece,  $XB^*$ , FCRM [7] algoritması için önerilen yeni öbek geçerlilik ifadesinin başlangıç noktası olur diyebiliriz.

### 3.2.2. Kung-Lin İndeksi

Kung ve Lin [17] anahtarlamalı regresyon problemlerinin bulanık uyarlaması gibi Bulanık Öbek Ortalaması Regresyon Modeli (FCRM) [7] tipi öbekleme yaklaşımlarını geçerli kılmak için bir geçerlilik indeksi önermişlerdir. FCRM algoritmaları modelleri, gerçek çıktı ile her regresyon model çıktısı arasındaki hata indeksine bağlı önem ya da ağırlık olarak isimlendirilen bir dağılım matrisini (üyelik değerleri matrisi) ve  $c$  regresyon model parametrelerini belirler. Kung ve Lin [17] tarafından ileri sürülen geçerlilik indeksi  $XB$  indeksine bağlıdır, fakat yoğunluk gözlenen çıktı ile her öbeğin doğrusal ya da polinomal fonksiyonlarından elde edilen çıktı arasındaki hata ile ölçülür. Ayrılabilirlik, birim normal vektörlerin iç çarpımının kesin değeri ile tanımlanan öbekler arasındaki benzemezliğin tersi alınarak ölçülür. Kung-Lin'in öbek geçerlilik indeksi şu şekildedir:

$$Kung - Lin(c) = \frac{\left( \frac{\sum_{i=1}^c \sum_{k=1}^n \mu_{ik}^2 \|x^T \beta_i - y_k\|^2}{n} \right)}{\left[ \frac{1}{\max_{i \neq j} \langle u_i, u_j \rangle} + \kappa \right]}, \text{ burada } \beta_i = [x^T \mu_i x]^{-1} [x^T \mu_i] y \quad (3.6)$$

Eşitlik 3.6'daki pay yoğunluk ölçüsünü, payda ise ayrılabilirliği temsil etmektedir.  $u_i$ 'ler her  $c$ -regresyon fonksiyonunun birim normal vektörünü temsil eder. FCRM modelleri [7] regresyon eşitlikleri ile gösterilir ve bu nedenle bunlara ilişkin birim vektörler şu şekilde tanımlanır:

$$u_i = \frac{[n_i]}{\|n_i\|}, n_i = [\beta_{i1} \dots \beta_{i,nv} - 1] \in \mathfrak{R}^{nv+1}$$

$n_i$ : regresyon fonksiyon parametrelerini,

$\beta_{i,nv}$ : bir vektör formunda ifade eder ve  $\|\cdot\|$  Öklit normudur,

$nv$ : girdi veri kümesindeki  $x = [x_1, \dots, x_{nv}]$  değişken sayısını ifade eder.

İki öbeğin iki birim vektörünün iç çarpımı, bunlar arasındaki açının kosinüs değerine eşittir. Bu değer, Kung ve Lin [17]'in geçerlilik formülünün  $c$ -regresyon fonksiyonlarının ayrılabilirliğini ölçer. Fonksiyonlar birbirine dik olduğunda ayrılabilirlik maksimumdur. Kung ve Lin [17], geçerlilik fonksiyonunun FCRM modelleri için iyi bir bulgu olduğunu göstermişlerdir.

### 3.2.3. Geliştirilmiş Bulanık Öbekleme İndeksi

Geliştirilmiş Bulanık Öbeklemeyi (GBÖ) [18] ve sınıflandırma problemleri için GBÖ, Sınıflandırılmalı Geliştirilmiş Bulanık Öbekleme (SGBÖ) [19]'yi geçerli kılmak için Geliştirilmiş Bulanık Öbekleme için Öbek Geçerlilik İndeksi (ÖGİ-GBÖ) ve Sınıflandırılmalı Geliştirilmiş Bulanık Öbekleme için Öbek Geçerlilik İndeksi (ÖGİ-SGBÖ) olmak üzere iki yeni *oran tipi* öbek geçerlilik indeksi sunulmuştur [20]. İlk olarak, regresyon problemleri için GBÖ üzerine odaklanılmıştır. GBÖ algoritmasında öbekler, öbek ön modelleri (merkezleri) ve bunların ilgili regresyon fonksiyonları ile belirlenir. GBÖ algoritmasında hesaplanan üyelik değerleri, her nesnenin her öbeğe ne derecede ait olduğunu gösterir. Ayrıca, girdi ve çıktı değişkenleri arasındaki ilişkiyi regresyon fonksiyonları ile belirlemeye yardımcı olan aday girdi değişkenleridir. Her girdi vektörü, üyelik değerlerinin ve/veya bunların dönüşümlerinin yeni ek girdi olarak kullanıldığı özellik uzayına taşınır. Özellik uzayında her öbek için yeni bir veri kümesi oluşturulur ve bu veri kümesi kullanılarak bir “Bulanık Regresyon Fonksiyonu” tahmin edilir. Bu nedenle, GBÖ [18] algoritmasından elde edilen üyelik değerlerinin çıktı değişkenini bir başka ifadeyle regresyon fonksiyonlarının “iyi” tahmin edicisini açıklaması, özellik uzayında daha iyi bir bulanık dağılımı göstermesi beklenir.

GBÖ algoritmasını geçerli kılarken, öbeklerin temsili regresyon fonksiyonları ölçüldüğü gibi öbekleme yapısı da ölçülerek öbeklerin yoğunluğu ve ayrılabilirliği arasındaki ilişki bulunur. Yeni geçerlilik indeksi, GBÖ [18] modellerinin öbek sayısı geçerli kılınacağı zaman bu kavramların ikisini de içermelidir. Yeni geçerlilik indeksi iki terimi birleştirir:

- $XB^*$  yoğunluğu Eşitlik 3.5'de yer alan formülün payında belirtildiği üzere yoğunluğun ilk terimi olarak kullanılacaktır
- İkinci terim olarak Kung-Lin indeksi Eşitlik 3.6'da yer alan formülün payında belirtildiği üzere yoğunluğun değiştirilmiş bir tipidir. ÖGİ-GBÖ 'nin yoğunluğunun ikinci terimi, gerçek çıktı ile regresyon modeli arasındaki hatayı temsil eder.

Bu regresyon modelleri Bulanık Regresyon Fonksiyonlarıdır,  $f(\Phi_i, \widehat{W}_i) \rightarrow y$ , burada  $\Phi_i$  girdi değişkenleri ve ilgili öbekteki üyelik değerleri ve bunların dönüşümlerini içeren bir matristir ve  $\widehat{W}_i$  girdi uzayının üyelik değerlerini kullanan yeni bir uzaya eşlemeleri olan regresyon katsayılarıdır. Diğer yandan, ÖGİ-GBÖ'nin ayrılabilirliği, regresyon fonksiyonları arasındaki açı ile öbek merkez ön modelleri arasındaki uzaklıkları eşleştirir.

Yeni geçerlilik fonksiyonunda, yoğunluk ölçülürken orijinal sayısal girdiler de fonksiyonlara tahmin edici olarak girerler. Bu, şu şekilde açıklanabilir: GBÖ [18] algoritması, verilen sistemin kısmi modellerini tahmin edebilecek üyelik değerlerini,  $(\mu_i^{imp}), i=1, \dots, c$ , bulur.  $(^{imp})$  simgesi, üyelik değerlerinin Geliştirilmiş Bulanık Öbekleme (GBÖ) yönteminden elde edildiğini belirtir.

Bu üyelik değerleri ve/veya bunların dönüşümleri, her öbek için “En Küçük Kareler (EKK), Destek Vektör Makineleri, Ridge Regresyonu” gibi fonksiyon yaklaşım yöntemlerini kullanarak “Bulanık Regresyon Fonksiyonlar”ını belirlemek için girdi değişkenleri ile birlikte kullanılır. Yeni GBÖ algoritması, her öbeğin kısmi modellerinin hatasını minimize etmek için orijinal girdilerin yanında üyelik değerleri ve bunların dönüşümlerini ek tahmin edici olarak tanıtır.

Bu nedenle, orijinal girdi değişkenlerine ek olarak yeni üyelik değerlerinin davranışını analiz ederek yeni GBÖ algoritmasının en iyi öbek sayısı geçerli kılınır. Bundan sonraki adımlar yeni geçerlilik analizinin düzenini tanımlar:

(i) Herhangi bir  $i$  öbeği için üyelik değerleri  $(\mu_i^{imp})$  ve/veya bunların dönüşümünün ek tahmin edici olarak kullanılmasıyla farklı bir veri kümesi oluşturulur. Bu,  $nv$  girdi değişkenli orijinal girdi uzayını  $(x \in \mathfrak{R}^{nv})$  daha yüksek boyutlu özellik uzayına  $(\mathfrak{R}^{nv+nm})$  taşır. Böylece, her veri vektörü  $(nv+nm)$  özellik uzayında gösterilir.  $nm$ , orijinal girdi uzayına ek tahmin ediciler olan eklemeli üyelik değerleri ve bunların olası dönüşümlerinin sayısıdır. Eşitlik 3.7’de bir boyutlu girdilerin orijinal girdi matrisinin  $(nv+1+1)$  boyutlu yeni bir uzaya taşınmasıyla ve üyelik değerlerini kullanan tek çıktı ile oluşturulan örnek bir özellik uzayı gösterilir  $(nv=1, nm=1)$ :

$$\phi_i(x, \mu_i^{imp}) = \begin{bmatrix} \mu_{i \times 1}^{imp} & x_1 \\ \vdots & \vdots \\ \mu_{i \times n}^{imp} & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \in \mathfrak{R}^{nv+nm} \quad (3.7)$$

Eğer gerekirse, orijinal girdi değişkenlerine ek olarak üyelik değerlerinin  $(\mu_i^{imp})^2$ ,  $(\mu_i^{imp})^m$ ,  $\exp(\mu_i^{imp})$ ,  $\ln((1-\mu_i^{imp})/\mu_i^{imp})$ , vb. şeklindeki matematiksel dönüşümleri kullanılabilir. Burada  $m$ , GBÖ'nün bulanık derecesini temsil eder.

(ii) Daha sonra her öbek için ilgili girdi veri kümeleri kullanılarak regresyon fonksiyonu  $(\Phi_i(x, \mu_i^{imp}))$  belirlenir. Bu nedenle, GBÖ [18] öbeklemesi için tasarlanan yeni geçerlilik indeksi Eşitlik 3.8'de formüle edilmektedir.

$$vc^* = \max_{i=1,\dots,c} \frac{1}{n} \sum_{k=1}^n \mu_{ik}^m \left( \|x_k y_k - v_i\|^2 + (y_k - f_i(\Phi_i, \hat{w}_i))^2 \right) \quad (3.8)$$

$$vs^* = \begin{cases} \min_{i,j \neq i} \left( \|v_i - v_j\|^2 / |\langle \alpha_i, \alpha_j \rangle| \right), & \text{if } |\langle \alpha_i, \alpha_{j_{i=1,\dots,c,j \neq 1}} \rangle| \neq 0 \\ \min_{i,j \neq i} \|v_i - v_j\|^2, & \text{diğer durumlar} \end{cases}$$

$$cvlIFF = \frac{vc^*}{(c \cdot vs^*) + 1}$$

Eşitlik 3.8'de:

$vc^*$  : yeni geçerlilik indeksinin yoğunluğunu,

$vs^*$  : yeni geçerlilik indeksinin ayrılabilirliğini,

$n_i^\Phi = \langle \hat{W}_{i1}, \hat{W}_{i2}, \dots, \hat{W}_{im}, \hat{W}_{i(nm+1)}, \dots, \hat{W}_{i(nm+nv)} \rangle \in \mathfrak{R}^{nv+nm}$  :  $\Phi_i(x, \mu_i^{imp}) \in \mathfrak{R}^{nv+nm}$  özellik uzayındaki veri kümelerinden elde edilen bulanık fonksiyonların yüzeyine dik olan normal vektörü,

$\alpha_i$  :  $|\langle \alpha_i, \alpha_j \rangle| \in [0,1]$ 'deki her "Bulanık Fonksiyon" un normal vektörünü,

temsil eder. Burada,  $i$ . öbek için, birim vektör olan  $\alpha_i = [n_i^\Phi] / \|n_i^\Phi\|$  olarak ifade edilir ve  $\|n_i^\Phi\|$  vektörün uzunluğunu gösterir. İki öbeğin bulanık fonksiyonlarının iki birim vektörünün iç çarpımının değeri, bu iki öbek arasındaki açının kosinüsüne eşittir,  $i, j = 1, \dots, c$ ;  $i \neq j$  :

$$\cos \theta_{i,j} = \langle \alpha_i, \alpha_j \rangle = \frac{\langle n_i^\Phi, n_j^\Phi \rangle}{\|n_i^\Phi\| \|n_j^\Phi\|} = \frac{[\hat{W}_{i1} \hat{W}_{j1} + \dots + \hat{W}_{i(nm+nv)} \hat{W}_{j(nm+nv)}]}{\sqrt{(\hat{W}_{i1})^2 + \dots + (\hat{W}_{i(nm+nv)})^2} \cdot \sqrt{(\hat{W}_{j1})^2 + \dots + (\hat{W}_{j(nm+nv)})^2}} \quad (3.9)$$

Öbek sayısı fazla olduğunda, iki öbek merkezi birbirine çok yakın olacağından aralarındaki uzaklık neredeyse yok olur ( $\cong 0$ ), bu durumda geçerlilik indeksi sonsuza gider. Bunu önlemek için, Eşitlik (3.5)'deki ÖGİ-GBÖ'nin paydası bir artırılır. SGBÖ modelleri için ÖGİ-SGBÖ 'de [19],  $vc^*$  'deki “Bulanık Fonksiyonlar”,  $f_i(\Phi_i(x, \mu_i^{imp}), \widehat{W}_i)$ , sınıflandırıcı modellerden elde edilen sonsal olasılıklar ile yer değiştirir:

$$vc^* = \max_{i=1, \dots, c} \left\{ \frac{1}{n} \sum_{k=1}^n (\mu_{ik}^{imp})^m \left( \|x_k y_k - v_i\|^2 + \left( y_k - \hat{P}_i(y=1 | f(\Phi_i(x, \mu_i^{imp}), \widehat{W}_i)) \right)^2 \right) \right\} \quad (3.10)$$

Yoğunluk ölçüsü, yeni geçerlilik fonksiyonunun  $vc^*$  değeri (ÖGİ-GBÖ'nün payı), ki o da  $XB^*$  indeksinin payındaki yoğunluk ( $vc^*$ ) teriminin Bulanık Fonksiyonların hata karesini içeren farklı bir şeklidir. Bununla birlikte, ÖGİ-GBÖ'nün yoğunluk değeri Şekil 3.6'da gösterilen  $XB^*$  indeksi ile çıktı üzerine benzer bir etki yapar.  $c > c_{optimum}$  olduğunda, öbeklerin yoğunluğu küçük olacaktır. Bu da, öbek sayısı arttıkça öbekler daha benzer nesnelere içereceğinden öbek-içi uzaklıkların azalacağı gerçeğinden ortaya çıkar. Ek olarak, her öbek için bir regresyon fonksiyonu vardır. Benzer şekilde, fonksiyonların sayısı arttıkça, “Bulanık Fonksiyonlar”ın hatası azalacaktır, çünkü regresyon model çıktısı gerçek çıktıya yaklaşacaktır.  $c = n$  olduğunda yoğunluk değeri sıfır olur, ki bu durumda her nesne kendi öbek merkezi ve fonksiyonu olur. Diğer yandan,  $c < c_{optimum}$  olduğunda, öbekler ÖGİ-GBÖ 'daki yoğunluğun ilk terimini artıracak benzemez nesnelere de birlikte içerecektir. Gerçek model sayısından daha az fonksiyon olacağından, gerçek ve beklenen sapma arasındaki hata yüksek olacaktır. Bu nedenle, eğer öbek sayısı en iyi öbek sayısından az olursa, yoğunluk yüksek olacaktır. Öbeklerin sayısı  $c_{optimum}$  'a yaklaştığında yoğunlukta bir düşüş olacaktır. Bu nedenle,  $c$  değeri  $c = c_{optimum}$  'den  $c = n$  'ye arttığında yoğunluk yavaş yavaş sıfıra yaklaşacaktır.

Her öbek farklı yoğunluk değerlerine sahip olabilir. Yeni Öbek Geçerlilik İndeksi (ÖGİ-GBÖ) Eşitlik (3.5)'te gösterildiği üzere  $XB^*$  [14] ile öbeklerin ayrı ayrı yoğunluk değerlerinin ortalamasını alarak öbek sayısı az olduğunda yüksek yoğunluklu öbeklerin etkisini bastırma sıkıntısını paylaşır.  $c < c_{optimum}$  olduğunda, öbeklerin en yüksek ve en düşük yoğunlukları arasındaki fark fazla olacaktır. Oysa  $c \geq c_{optimum}$  iken, her bir ögenin ayrı ayrı yoğunlukları çok düşük olacaktır, böylece öbeklerin en yüksek ve en düşük yoğunlukları arasındaki fark ihmal edilebilir olacaktır. Bu sayede, ortalama yoğunluk değerleri yerine en yüksek yoğunluk

değerlerini analiz ederek ani değişiklikler belirlenebilir. Bunlar ideal koşullar altındaki varsayımlardır ve yeni öbek geçerlilik indeksi, bu öbeklerin ayrılabilirliklerinin ölçüsünü en iyi öbek sayısına yakınsama hassasiyetinin artması için birleştirir.

Yeni geçerlilik indeksinin ayrılabilirliği, ayrıca regresyon ve öbekleme gibi iki farklı yapıdan elde edilen ayrılabilirlik ölçüsünü de birleştirir. Eşitlik (3.5)'te, öbek merkezleri arasındaki uzaklık Öklid uzaklıkları olarak alınır. Ek olarak, her “Bulanık Fonksiyonun  $\langle \alpha_i, \alpha_j \rangle \in [0,1]$  arasındaki açının kosinüsünün değeri, ek ayrılabilirlik kriteri olarak kullanılır. Eğer fonksiyonlar dik ise, bu fonksiyonlar benzemez fonksiyonlardır, ki bu da en iyi modelin bir özelliğidir. Ayrılabilirlik, Eşitlik (3.5)'de verilen ÖGİ-GBÖ'nin [18] paydası, koşullu olarak öbekler arası uzaklıklar ile açılarının oranını alarak birleştirir. Eğer herhangi iki fonksiyon arasındaki açı sıfır ise, bu fonksiyonlar birbirine paraleldir. Bu nedenle, ayrılabilirliği göstermek için bunların öbek merkezleri arasındaki minimum uzaklık kullanılır. Daha iyi ayrılabilirlik daha iyi öbekleme sonuçları sağlar.

### 3.3. Öbek Merkezleri

Öbek sayısına karar verilmesinin ardından o öbek içinde yer alan verilerin ağırlık merkezleri esas alınarak belirlenen öbek merkezleri de her bir verinin bu öbek merkezine uzaklığı üzerinden hesaplanan üyelik değerlerinin hesaplanmasına olanak sağlayacağından algoritma içerisinde önemli bir parametre olarak yer almaktadır.

BCO Öbekleme Algoritmasının [6] ilk iterasyonda belirlediği öbek merkezi yaklaşık bir değer olup her iterasyonda iyileşmektedir. İki iterasyonda arasındaki öbek merkezi değeri öncesinde belirlenen hata payının ne zaman altında kalır ise o zaman algoritma sonlanır ve öbek merkezleri nihai değerine ulaşır. Öbek merkezleri belirlendikten sonra aşağıda belirtildiği şekilde ifade edilir.

$$V = \{v_1, v_2, \dots, v_c\} \in \mathcal{R}^{c \times nv} \quad (3.11)$$

Başlangıçta veri setindeki girdi sayısı kadar ( $nv$ ) öbek merkezi oluşturulması sık kullanılan bir yöntemdir.

$nv$  boyutlu  $X$  veri kümesi vektörleri, bulanık öbekleme algoritmaları ile  $c$  adet öbeğe ayrıştırılır. Ayrıştırılan öbeklerin merkezleri  $V = \{v_1, v_2, \dots, v_c\} \in \mathcal{R}^{nv}$   $nv$  boyutlu olan vektörlere ayrıştırılırlar. Her öbek merkezi ( $v_i$ ),  $nv$  adet gözlemin merkezidir.

BCO Öbekleme Algoritması, amaç fonksiyonunu en iyilemeye çalışır ve bunu yaparken de uzaklık ölçütünü esas alır. Algoritma çalıştırılmadan önce, veri kümesinin kaç tane öbeğe ayrılacağı bilindiği varsayılır. Gerçek hayatta pek mümkün olmayan bu yaklaşıma öneri niteliğinde bir çözüm olarak daha önce bahsedilen öbek geçerlilik indeksleri (ÖGİ) geliştirilmiştir.

### 3.4. Bulanıklık Mertebesi

Belirsizlik teorisinde [2] bulanıklığın bir tür belirsizlik (kesin olmamak) olduğu kabul edilir. Bulanıklık ölçü derecesini belirlemek için çeşitli fonksiyonlar önerilmiştir. Bulanık öbekleme algoritmalarında örtüşme sabiti ( $m$ ), bulanıklık derecesi olarak kullanılır. Sonraki bölümlerde “ $m$ ” diğer kısıtlar ile birlikte önerilen bulanık fonksiyonlar yaklaşımı belirsizliğini tanımlamak için bir parametre olarak kullanılacaktır.

$m \in (1, \infty)$ , bulanık öbekleme algoritmasının “Bulanıklık Derecesi” ya da “Bulanıklaştırıcı”yı göstermekte ve öbeklerin örtüşme derecesini belirlemektedir. “ $m=1$ ” kesin kümeleme yapısını gösterir ve örtüşme yok anlamına gelir.

Bulanıklık mertebesi BCO Öbekleme Algoritması [6] ile üyeliklerin hesaplanmasında ve dolayısı ile üyelikler ve transformasyonlarının yer aldığı bulanık regresyon modelinde modelin tutarlılığını doğrudan etkileyen önemli bir parametredir. Üyeliklerin hesaplandığı algoritmada ve algoritmanın en küçükleme çalıştığı amaç fonksiyonunda üstsel fonksiyon olarak yer almaktadır.

Bulanıklık değerinin ( $m$ ) etkisi, sınırlarda Eşitlik 3.12 'de verilen üyelik değeri hesaplama eşitliğini kullanarak limitinin alınmasıyla aşağıdaki gibi analiz edilebilir:

$$\lim_{m \rightarrow \infty} \mu_{ik}(x) = \lim_{m \rightarrow \infty} \left[ \sum_{j=1}^c \left( \frac{d^2(x_k, v_i)}{d^2(x_k, v_j)} \right)^{1/(m-1)} \right]^{-1} = \frac{1}{c}, \forall i, j = 1, \dots, c \quad (3.12)$$

ve hiçbir öbek merkezinin birbirine benzemediği varsayımı altında Eşitlik (3.13) elde edilir.

$$\lim_{m \rightarrow \infty} \mu_{ik}(x) = \begin{cases} 1, & \text{if } d^2(x_k, v_i) < d^2(x_k, v_i), \forall i, j = 1, \dots, c, i \neq j \\ 0, & \text{diğer durum} \end{cases} \quad (3.13)$$

$m$  değeri artarken  $\mu_{ik}$  değeri 0'a yakınsayacaktır. Örneğin,  $m=6$  olduğunda herhangi bir güçlü üyelik,  $\mu_{ik} = 0.85$  gibi,  $1/c$ 'ye yakın bir değere yaklaşacaktır.  $m$  parametresi öbeklerin örtüşme derecesini temsil ettiğinden,  $m$  büyüdükçe bulanıklık sonuçları ve örtüşme daha geniş

olacaktır.  $m$  değeri küçüldükçe, bulanık öbeleme sonucu, kesin kümeleme modeline daha yakın olacaktır.  $m$  değerinin 1 olması kesin kümeleme ile aynıdır. Bu da, öbekler arasında örtüşme olmaması ve tüm üyelik değerlerinin  $\mu_{ik} \in \{0,1\}$  olması anlamına gelmektedir.

Önceden yapılan araştırmalar [21] genel olarak sistem modelleme analizinde  $m=2$  olması gerektiğini göstermektedir. Daha yeni bir çalışmada [22] üyelik değeri hesaplama fonksiyonu için Taylor açılımı analizine dayanarak,  $m$ 'in minimum ve maksimum değerlerinin [1.4, 2.6] olduğu kanıtlanmıştır.

### 3.5. Uzaklık Hesaplama Metodları

Bir veri setinde yer alan birimlerin kümelmesi işlemi bu birimlerin birbirleriyle olan benzerlikleri ya da birbirlerine olan uzaklıkları kullanılarak gerçekleştirilmektedir. Değişkenlerin kesikli ya da sürekli olmalarına ya da değişkenlerin nominal, ordinal, küresel ya da eliptik dağılıma sahip olmalarına göre hangi uzaklık ölçüsünün ya da hangi benzerlik ölçüsünün kullanılacağına karar verilebilmektedir.

#### 3.5.1. Aralık Ölçekli ve Oransal Ölçekli Değişkenler

Aralık ölçekli değişkenler arasındaki karşılaştırma oranıyla değil fark ile yapılır. Örneğin, sıcaklık değişkeni inceleniyorsa, bir ev diğer evden 2 kat daha sıcak demek yerine, 5 C daha sıcak gibi bir tanımlama yapılır.

Oransal ölçekli değişkenlerse; alan, uzunluk gibi aralarında kıyaslama yapılırken bir oran belirtebileceğimiz değişkenlerdir. Örneğin, oraya ulaşmak için A yolunu kullanırsak B yoluna göre iki kat daha fazla yol yürümüş oluruz, diyebiliriz.

Aralık ve oransal ölçekteki değişkenleri bulunan birimleri kümelerken, değişik uzaklık ölçüleri kullanılır. Bu ölçülerden bazıları: Öklidyen, Mahalanobis, Maksimum Mesafe ve Minkowski uzaklık ölçüleridir.

Uzaklık ölçüsünün genel formülü Eşitlik 3.14'te ifade edildiği şekildedir.

$$d^2(x_k - v_i) = (x_k - v_i)^T A_i (x_k - v_i) \geq 0 \quad (3.14)$$

Bu formül ile ifade edilen uzaklık fonksiyonunun özellikleri ise aşağıdaki şekilde sıralanabilir:

1.  $d(i, j) \geq 0$ ; Uzaklık negatif değil



2.  $d(i, i) = 0$  ; Her birim kendisine olan uzaklığı sıfırlar.
3.  $d(i, j) = d(j, i)$  ; Uzaklık fonksiyonu simetriktir.
4.  $d(i, j) \leq d(i, h) + d(h, j)$  ; iki birimin arasındaki uzaklık bu iki birimin üçüncü bir birime olan uzaklıkları toplamından küçük olamaz (üçgen eşitsizliği)

BCO Öbekleme Algoritması [6] kapsamında her bir verinin öbek merkezlerine uzaklıkları esas alınarak üyelik değerleri hesaplanmakta ve bu kapsamda kullanılan birden fazla yöntem bulunmaktadır.

Modelin iyileşmesi açısından verilerin karakteristik özelliklerine göre modelin tutarlılığı üzerinde daha iyi sonuç verdiği gözlemlenen bazı yöntemler bulunmakta olup bu yöntemlerden öne çıkanlar aşağıda verilmiştir.

Çizelge 3.1. Farklı öbek geçerlilik indeksi hesaplama yöntemleri [23]

Uzaklık Ölçütü	Fonksiyon
Öklid Mesafesi	$d_2(a, b) = \left[ \sum_{i=1}^n (a_i - b_i)^2 \right]^{\frac{1}{2}}$
Minkowski Mesafesi	$d_p(a, b) = \left[ \sum_{i=1}^n  a_i - b_i ^p \right]^{\frac{1}{p}}, \quad p > 0$
Maksimum Mesafe	$d_\infty = (a, b) = \max_{i=1}^{nv}  a_i - b_i $
Mahalanobis Mesafesi	$d_A(a, b) = \sqrt{(a - b)^T A (a - b)}$

Minkowski uzaklık ölçüsündeki m değeri büyük ve küçük farklara verilen ağırlığı değiştirir. İlgili uzaklık hesaplama yönteminde m = 2 değerini alırsa, formül Öklidyen uzaklık ölçüsü formülüne dönüşür.

Eşitlik 3.6'daki norm matrisi  $A_i, (i=1, \dots, c)$ , pozitif tanımlı simetrik bir matristir. Diğer uzaklık ölçüleri de bulanık öbeleme algoritmalarında kullanılabilir. Farklı uzaklık ölçülerine ait kısa bir liste Çizelge 3.1'de verilmiştir. BCO Öbeleme algoritması, Öklid uzaklığını kullanır. Bu nedenle norm matrisi  $A_i$ , girdi matrisi standart sapması 1 ve ortalaması 0 olarak ölçeklendirildiğinden birim matrise eşittir ( $A = I$ ). Diğer taraftan Gustafson ve Kestel [24] "Mahalanobis" uzaklığını kullanır. Bu durumda her bir öbeğin matris modeli, o öbeğin kovaryansının tersine eşittir ( $A_i = C^{-1}$ ).

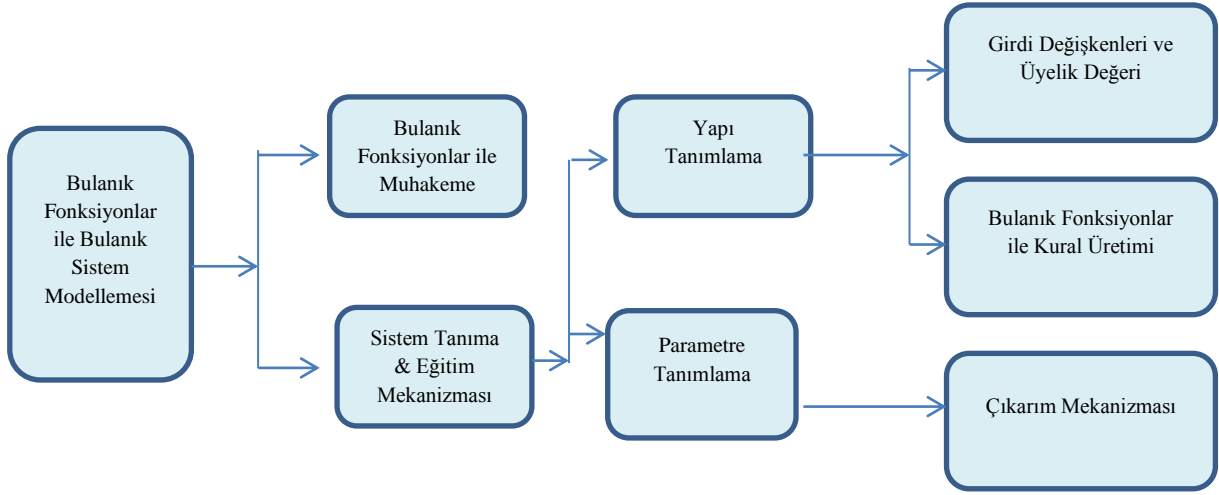
Bu uzaklık hesaplama yöntemlerinde Euclidian in sıklıkla dairesel dağılıma sahip veri setlerinde; Minkowski ve Mahalonobis ise eliptik dağılıma sahip veri setlerinde daha iyi sonuç vermektedir. [25]

### **3.6. Bulanık Öbek Ortalaması (BCO) Algoritması ve Adımları**

Daha önce de bahsedildiği üzere Bulanık Mantık, 1965 yılında Lütü Askerizade'nin yayınladığı bir makalenin sonucu oluşmuş bir mantık yapısıdır.

Bulanık mantığın temeli bulanık küme ve alt kümelerine dayanır. Klasik yaklaşımda bir varlık ya kümenin elemanıdır ya da değildir. Matematiksel olarak ifade edildiğinde varlık küme ile olan üyelik ilişkisi bakımından kümenin elemanı olduğunda "1", kümenin elemanı olmadığı zaman "0" değerini alır. Bulanık mantık klasik küme gösteriminin genişletilmesidir. Bulanık varlık kümesinde her bir varlığın üyelik derecesi vardır. Varlıkların üyelik derecesi, (0, 1) aralığında herhangi bir değer olabilir ve üyelik fonksiyonu  $M(x)$  ile gösterilir.

Klasik kümelerin aksine, bulanık kümelerde elemanların üyelik dereceleri [0, 1] aralığında sonsuz sayıda değişebilir. Bunlar üyeliğin derecelerinin devamlı ve aralıksız bütünüyle ilgili bir kümedir. Keskin (klasik) kümelerdeki soğuk-sıcak, hızlı-yavaş, aydınlık-karanlık gibi birbirinden tamamen farklı ikili değişkenler bulunurken bulanık mantıkta biraz soğuk, biraz sıcak, biraz karanlık gibi esnek niteleyicilerle yumuşatılarak gerçek dünyaya uygun belirteçlerin tanımlanması mümkün olabilmektedir.



Şekil 3.5. Bulanık Fonksiyonlar yaklaşımları ile bulanık sistem modelleri yapısı [26]

BCO Algoritmasının, en iyi sonuçları elde etmek için minimize edilmesi gereken bir kısıtlı en iyileme (optimizasyon) problemi olduğu görülmektedir. Bu nedenle, BCO Algoritması aşağıdaki gibi tek bir en iyileme (optimizasyon) yapısı şeklinde yazılabilir:

$$MinJ(X; U, V) = \sum_{i=1}^c \sum_{k=1}^n (\mu_{ik})^m d^2(x_k - v_i)_A$$

$$0 \leq \mu_{ik} \leq 1, \forall i, k$$

$$\sum_{i=1}^c \mu_{ik} = 1, \forall k > 0 \quad (3.15)$$

$$0 < \sum_{i=1}^c \mu_{ik} < n, \forall i > 0$$

Eşitlik 3.15'te verilen [6] kısıtlı en iyileme modeli Lagrange Çarpanı yöntemi [27] kullanılarak çözülebilir ve bu model, tek bir amaç fonksiyonu ile kısıtsız en iyileme problemine dönüştürülür. Bir eşitlik kısıtlı problem elde etmek için, birincil (primal) kısıtlı en iyileme problemi öncelikle, Lagrange Çarpanları olarak bilinen belirlenmemiş  $\lambda$  parametrelerinin yardımıyla Eşitlik 3.16'da yer alan eşdeğer bir kısıtsız problem haline dönüştürülür;

$$Max W(U, V) = \sum_{i=1}^c \sum_{k=1}^n (\mu_{ik})^m d^2(x_k, v_i)_A - \sum_{i=1}^c \lambda_i (\mu_{ik} - 1) \quad (3.16)$$

Lagrange yöntemine göre, Lagrange fonksiyonu birincil (primal) parametrelere göre minimize edilmeli ve ikincil (dual) parametrelere göre maksimize edilmelidir. Eşitlik

3.16'daki amaç fonksiyonunun model parametreleri olan küme merkezleri ( $V$ ) ve üyelik değerleri ( $U$ )'ne göre türevi alınarak en iyi üyelik değeri ve öbek merkezi hesaplama eşitlikleri aşağıdaki gibi formüle edilirler:

$$\mu_{ik}^{(t)} = \left[ \sum_{j=1}^c \left( \frac{d(x_k, v_i^{(t-1)})}{d(x_k, v_j^{(t-1)})} \right)^{2/(m-1)} \right]^{-1} \quad (3.17)$$

$$v_i^{(t)} = \frac{\sum_{k=1}^n (\mu_{ik}^{(t)})^m x_k}{\sum_{k=1}^n (\mu_{ik}^{(t)})^m}, \forall i = 1, 2, \dots, c \quad (3.18)$$

Eşitlik 3.17'deki  $v_i^{(t-1)}$  değeri  $i$ . öbeğe ait  $(t-1)$ . iterasyondan elde edilen öbek merkezi vektörünü göstermektedir. Benzer şekilde, Eşitlik 3.17 ve Eşitlik 3.18'de  $\mu_{ik}^{(t)}$ ,  $t$ . iterasyonda hesaplanan en iyi üyelik değerlerini ifade etmektedir. Bu işlemin sonucu, üyelik değerlerinin ve öbek merkezlerinin birbirine bağlı olduğunu göstermektedir. Bu nedenle, Bezdek [6] üyelik değerlerini ve öbek merkezlerini hesaplamak için tekrarlı bir algoritma önermiştir. Her iterasyonda  $(t)$ , amaç fonksiyonu  $J^{(t)}$  aşağıdaki gibi hesaplanır;

$$J^{(t)} = \sum_{i=1}^c \sum_{k=1}^n (\mu_{ik}^{(t)})^m d^2(x_k, v_i^{(t)})_A > 0 \quad (3.19)$$

Bu yöntem kullanılırken hesaplanması gereken değerlerin başında üyelikler gelmektedir. Üyelikler ise aşağıdaki adımlar takip edilerek hesaplanabilmektedir.

#### **Algoritmanın Adımları [26]:**

Veri vektörleri  $X = (x_1, x_2, \dots, x_n)$ , öbek sayısı  $c$ , bulanıklık derecesi  $m$  ve durdurma sabiti  $\epsilon$ , verilsin.

Bölünme matrisi  $U$ , rastgele olarak başlatılır.

Hazırlık:

- ▶ *Öbek sayısının ve bulanıklık mertebesinin değerinin ya da aralığının kararlaştırılması:* Üyelik değerlerinin hesaplanmasında önemli rol oynayan öbek sayısı tek bir değer olarak belirlenebileceği gibi belli bir aralıkta yer alan çok sayıda öbek sayısı için çok sayıda üyelik değeri matrisi oluşturmak da bir yöntemdir. Modelin

tutarlılığında en iyi sonucu verecek öbek sayısını en başta kesin doğru olacak şekilde belirlemek mümkün olmadığından makul sonuç vermesi beklenen bir aralıkta birden çok deney yaparak birden çok öbek sayısı için oluşturulacak birden çok üyelik matrisi ile oluşturulan modelin tutarlılığı gözlemlenmesi daha iyi modellerin bulunmasına olanak sağlayacaktır.

- ▶ *[X] Veri setinin belirlenmesi:* Veri seri birden çok girdili bir ya da birden çok çıktı değerine sahip olan sürekli ya da kesikli nitelikte olabilmektedir.
- ▶  $\varepsilon$  belirlenmesi

**Birinci Adım:** Başlangıç merkezleri atanması

Girdi olarak başlangıç bölünme matrisinin üyelik değerleri alınarak başlangıç öbek merkezleri Eşitlik 3.20 kullanılarak hesaplanır. İlk adımda her vektör için ağırlık merkezi başlangıç öbek merkezi olarak atanır.

$$v_i^{(t)} = \left( \sum_{k=1}^n (\mu_{ik}^{(t)})^m x_k \right) / \sum_{k=1}^n (\mu_{ik}^{(t)})^m, \forall i = 1, \dots, c \quad (3.20)$$

**İkinci Adım:** İlk iterasyonun başlatılması

( $t=1$ , maksimum-tekrar) Bütün  $t$  değerleri için tekrarlanır.

- ▶ 3.21 denkleminde göre üyelikler hesaplanır

Öbekteki her bir veri girdisi  $k$  için üyelik değerleri Eşitlik 3.21’de yer alan üyelik değeri hesaplama formülasyonu ile bulunur.

$$\mu_{ik}^{(t)} = \left[ \sum_{j=1}^c \left( \frac{d(x_k, v_i^{(t-1)})}{d(x_k, v_j^{(t-1)})} \right)^{\frac{2}{m-1}} \right]^{-1} \quad (3.21)$$

- ▶ Bu üyelikler Eşitlik 3.22 ‘de yer alan üyelik değerleri matrisinde depolanır.

$$U = \begin{bmatrix} \mu_{1,1} & \mu_{2,1} & \cdots & \mu_{c,1} \\ \cdots & \cdots & \cdots & \cdots \\ \mu_{1,n} & \mu_{2,n} & \cdots & \mu_{c,n} \end{bmatrix} \quad (3.22)$$

**Üçüncü Adım:** Yeni öbek merkezi hesaplanır.

$t$ . iterasyondaki her bir  $i$ . öbek merkezi Eşitlik 3.23 kullanılarak hesaplanır.

$$v_i^{(t)} = \frac{\left(\sum_{k=1}^n (\mu_{ik}^{(t)})^m x_k\right)}{\sum_{k=1}^n (\mu_{ik}^{(t)})^m}, \forall i = 1, \dots, c \quad (3.23)$$

**Dördüncü Adım:** Durdurma koşulu olan iki öbek merkezi arasındaki fark en başta belirlenen  $\varepsilon$  değerinden küçük ise iterasyonlara son verilir,  $|v_i^{(t)} - v_i^{(t-1)}| \leq \varepsilon$ ; aksi halde birinci adıma geri dönülür.

BCO Öbekleme Algoritması [6], standart bulanık öbekleme algoritması olarak kabul edilebilir ve bu algoritmanın uzantıları birçok farklı amaç için öne sürülmüştür. BCO Öbekleme algoritması aynı zamanda bu çalışmalarda önerilen geliştirilmiş öbekleme algoritmasının (GBÖ-FCRM) çekirdek yapısıdır.

#### 4. BULANIK REGRESYON MODELİNİN KURULMASI [28]

1999 yılında Demirci tarafından önerilen bulanık fonksiyon yaklaşımı, bulanık eşitlikler temeline dayanır. Üyelik değerleri, bulanık işlemciler (üçgensel norm, tümleyen norm, vb.) kullanılarak çıkarsama yapılır ve bulanık eşitlikler temelinde bulanık fonksiyonlar tanımlanır. Bulanık kontrol ve teori çerçevesinde bulanık eşitlikler ve fonksiyonların başarılı bir şekilde uygulandığı Demirci'nin çalışmasında söylenir. Bu çalışma ile ilgili detaylı bilgiler Demirci'nin 1999 yılında yaptığı çalışmalardan edinilebilir [28].

Alışılmışın dışındaki bulanık fonksiyonlar yaklaşımı ilk olarak Türkşen tarafından ortaya atılmıştır. Bulanık Regresyon Modeli (BRM) yaklaşımı, bulanık sistem çıkarsaması yapılırken ayrı ayrı tanımlanması gereken bulanık kural tabanlarını 'Bulanık Fonksiyonlar' ile tanımlama ihtiyacından doğmuştur. [28]

Bulanık fonksiyonlar yaklaşımı girdiler arasındaki ilişkiyi, öbeleme algoritmalarından elde edilen üyelik değerleri ve kullanıcı tarafından tanımlı dönüştürülmüş hallerinin gerçek veri kümesi içerisine yeni açıklayıcı (bağımsız) değişkenler olarak katılmasıyla, girdi-çıkı arasında ilişkiyi yerel fonksiyonlarla tahmin eder. Böylece veri kümelerinin her öbeğe göre yapısı belirlenir. [28]

Bulanık fonksiyonlar, farklı bulanık sistem analizleri için oluşturulabilmektedir. Eğer amaç gözlem değerlerini sınıflandırmak ise, sınıflandırma problemlerinde olduğu gibi, sistem modellemesinin amacı sınıflandırılmamış durumları en küçükmektir. Eğer amaç, verilen bağımsız değişkenler ile bağımlı değişken arasındaki ilişkiyi fonksiyonlarla tahmin etmekse, sistemin amacı tahmin hatasını en küçükleyen fonksiyonu bulmaktır.

Bulanık fonksiyonların temelini bulanık öbeleme algoritmaları oluşturur. Genel olarak BCO Öbeleme Algoritması, bulanık fonksiyonlar yaklaşımında üyelik değerlerinin bulunmasında kullanılır.

Bulanık sistem modellemek isteyen uzmanlar bulanık fonksiyonları kullanarak, bulanık kural tabanlarına kıyasla birkaç adım daha kısa işlem yaparak daha kolay bir sistem analizi gerçekleştirebileceklerdir. Uzmanlar bulanık küme ve mantık teorisindeki bulanık işlemcileri örneğin üçgensel norm (min) veya tümleyen (max) normu bilmek zorunda olmadan bulanık sistem modellemesi yapabilmektedir.[28]

Bulanık Regresyon Modeli üyeliklerin BCO Öbekleme Algoritması ya da herhangi bir başka yöntemle hesaplanmasının ardından üyeliklerin kullanıldığı modellerde modelin tutarlılığı en iyileştiren uygulamalardan biri olarak karşımıza çıkmaktadır. Bu yaklaşım kural tabanlı değildir ve veri üzerinden modelin eğitilmesi mantığı üzerine çalışır, modelin her iki tarafı da regresyon denklemlidir. Üyeliklerin, üyeliklerin transformasyonlarının ve verinin kendisinin modele girmesi mümkündür. Ayrıca asıl modelde R-kare değerini iyileştiren adaylar girer.

Genel olarak BRM yaklaşımı aşağıdaki şekilde tanımlanmaktadır.

A modeli bulanıklık mertebesi n, öbek sayısı a olan bir model olsun.

Eşitlik 4.1’de yer alan  $U_A$  matrisi A modeli kapsamında BCO Algoritması aracılığı ile hesaplanan üyelik değerlerinin depolandığı matris olsun.

$$U_A = \begin{bmatrix} \mu_{1,1} & \mu_{2,1} & \cdots & \mu_{c,1} \\ \cdots & \cdots & \cdots & \cdots \\ \mu_{1,n} & \mu_{2,n} & \cdots & \mu_{c,n} \end{bmatrix} \quad (4.1)$$

$\mu_{1,1}$  : A Modeli kapsamında birinci girdi değerinin birinci öbeğe aitlik derecesi, üyelik değeri olsun.

Eşitlik 4.2’de belirtilen  $X^A$  matrisi BRM kapsamında girdi olarak kabul edilen A modeline göre hesaplanmış üyelik değerleri, bu üyeliklerin transformasyonları ve girdi verilerinin depolandığı genel girdi matrisidir.

$$X^A = \begin{bmatrix} 1 & \mu_{1,1} & \mu_{1,1}^2 & x_{1x1} & \cdots & x_{1xnv} \\ 1 & \mu_{1,2} & \mu_{1,2}^2 & x_{2x1} & \cdots & x_{2xnv} \\ \vdots & \vdots & \vdots & \vdots & \cdots & \cdots \\ 1 & \mu_{1,n} & \mu_{1,n}^2 & x_{nx1} & \cdots & x_{nxnv} \end{bmatrix} \quad (4.2)$$

$x_i$  : girdi verilerini temsil etmektedir.

$\mu_{1,n}^2$ : A Modeli kapsamında hesaplanan üyeliklerin transformasyonlarını temsil etmektedir.

BRM girdi değerleri Eşitlik 4.1’de yer aldığı üzere; tek bir bulanıklık mertebesi ve öbek sayısı için hesaplanmış üyelik değerlerinin, bu üyelik değerlerinin transformasyonlarının ve girdi verilerinin özetle Model A’nın verilerinin modelde çıktı tahmini amaçlı kullanımını esas



almaktadır. Bu tip regresyon modellemesinin Basit Doğrusal Regresyon (BDR)'ye kıyasla her durumda daha iyi sonuç verdiği geçmişte yapılan birçok çalışmada olduğu gibi bu çalışmada da ispatlanmıştır.

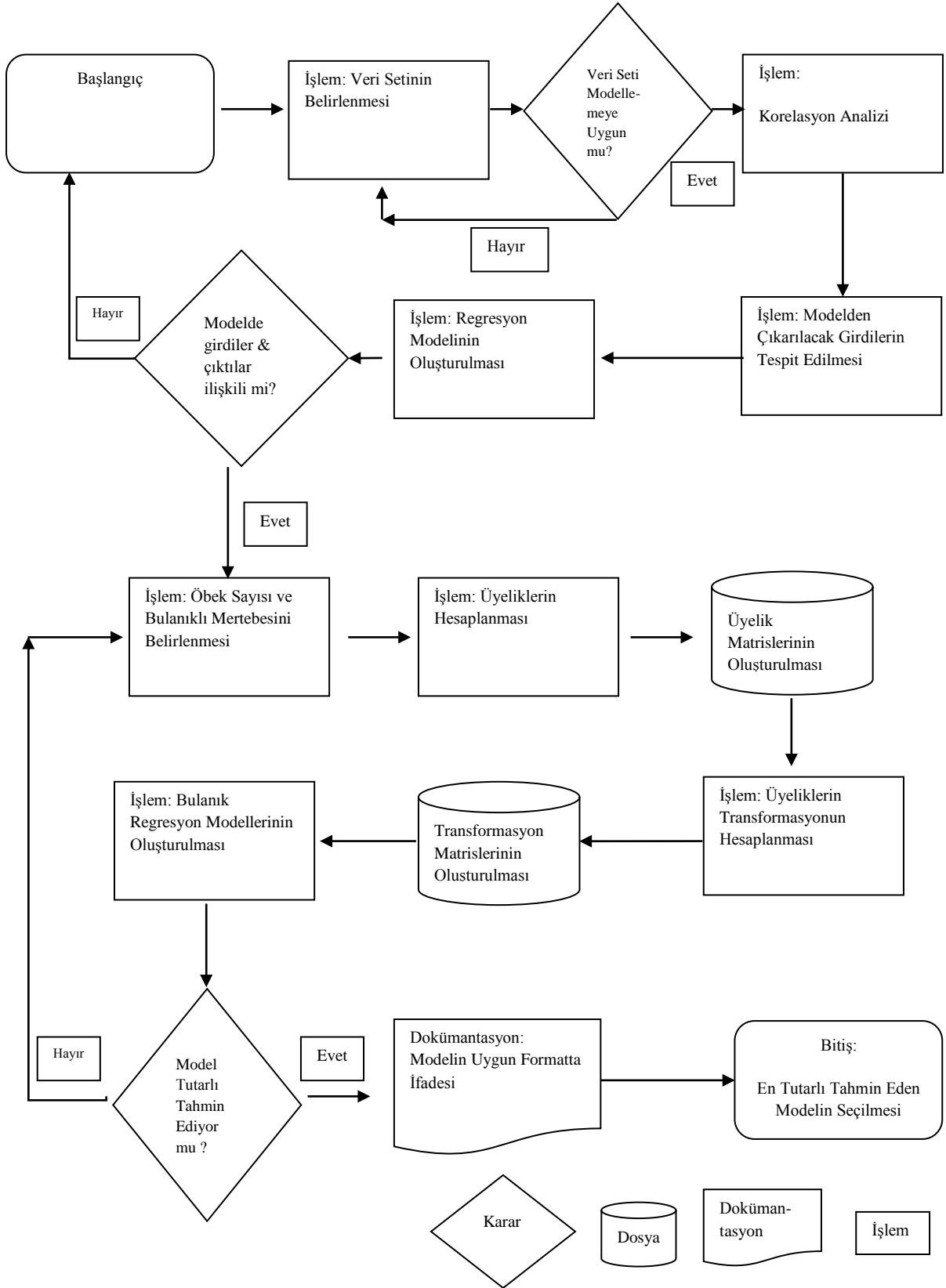
## 5. DENEY KÜMESİNİN TANIMLANMASI VE MODELLERİN KURULMASI

Bu çalışma kapsamında; BCO Öbekleme Algoritması [6] ve BRM [28] bir veri seti üzerinde uygulanmış, çıktılar farklı parametreler eşliğinde denenmiş, sonuçlar değerlendirilmiş ve ilk kez elde edilen bazı bulgulara ulaşılmıştır.

Bu tez çalışması kapsamında öncelikle tek çıktılı birden fazla girdili veri seti analiz edilmiştir. Bu analiz kapsamında girdilerin birbirleri ile ilişkilerini ve bir değişkenin diğer bir girdi değişkenine bağımlı olup olmadığını tespit etmek için korelasyon analizi yapılmış olup ilişkiye sahip olduğu tespit edilen çıktılardan sadece biri modelde kalacak şekilde girdi sayısında azaltma yapılmıştır. Sonrasında modelde kalmasına karar verilen girdiler ve çıktı veri seti için doğrusal regresyon modeli oluşturulmuş ve bu modelin tahmin tutarlılığı değerlendirilmiştir.

Bu değerlendirme sonucunda elde edilen veriler kaydedilmiştir. Sonrasında ise BCO Öbekleme Algoritması kullanılarak üyelik değerleri hesaplanmıştır. Üyelik değerleri sadece en iyi sonucu vermesi beklenen tek bir öbek sayısı ve bulanıklık mertebesi için değil; önceki çalışmaları da temel alacak şekilde mümkün olduğunca geniş bir aralık boyunca seçilmiş ve sonuç olarak da 221 ayrı modele temel oluşturacak sayıda farklı kombinasyon kullanılmıştır. Bu kapsamda hesaplanan 221 farklı model için hesaplanan 221 farklı üyelik değeri matrisi kullanılarak bu değerlerin transformasyonları hesaplanmıştır.

İlerleyen adımda ise hem girdi verileri, hem üyelikler hem de üyeliklerin transformasyonlarının çıktı değerlerini tahmin etmesi için oluşturulan BRM her bir deney kapsamında gerçekleştirilmiş olup bu modellerin tutarlılık parametreleri olarak dikkate alınan R-kare değeri üzerinden değerlendirmeler ve genel geçer çıkarımlar elde edilmiştir. Genel olarak bu bölümde yapılan çalışmalar aşağıdaki iş akışı şemasında özetlenmiştir.



Şekil 5.1. Bulanık Regresyon Modeli iş akış şeması

## 5.1. Veri Setinin Tanıtımı

Günümüz dünyasında mevcut verilerin kayıt altında tutulması son derece kolay olup bu verilerin gelecekte yapılacak çalışmalara rehberlik edecek nitelikte derlenip işlenmesi; kişi, kuruluş ve devletlere sunduğu proaktif yaklaşım sayesinde son derece büyük önem teşkil etmektedir. Bu önemin en geçerli sebebi ise geçmişte gerçekleşen bir takım faaliyetlerin rakamsal ölçütü olan gerçek verilerin gelecekte gerçekleşmesi muhtemel faaliyetlere en iyi şekilde rehberlik etmesi ve bu sayede kişi, kurum ve devletlerin aksiyon planlarında en önemli yol göstericilerden biri olmasıdır.

Bilginin; tasnifi, sistematik olarak kayıt altında tutulması, elde tutulan verilerin işlenmesi, modellenmesi ve nihayetinde gelecekte yapılması planlanan aksiyonlara rehberlik edecek şekilde tahminler elde etmeye uygun hale getirilmesi; kişi, kuruluş ve devletlere sadece maddi kazanç sağlamakla kalmayıp aynı zamanda bilginin temsil ettiği güç olgusunu da en iyi şekilde bahşetmektedir. Söz konusu yaklaşım elde var olan bilginin yönetimin ihtiyaçlarına yönelik olarak kullanımının önünü açmakta olup başta ekonomi alanında olmak üzere mühendislik, tıp, siyaset, kamu bilimi ve daha farklı alanlarda da başarılı uygulamalar ile günümüze kadar gelmiştir.

Bir çok otorite tarafından faydası hali hazırda kabul edilmiş bilgi yönetimi ve veri modelleme uygulamalarının güncel yöntemler ile harmanlanarak sonuçlarının paylaşıldığı bu çalışma kapsamında Danimarka Devleti'nin 1993-2013 yılları arasında resmi olarak açıkladığı ithalat tutarı toplamı verileri; Döviz Kuru, Reel Efektif Döviz Kuru, Tüketici Fiyat İndeksi (%değişim), Tüketici Fiyat İndeksi (ortalama), Üretici Fiyat İndeksi, Üretici Fiyat İndeksi-(ortalama), Üretici Fiyatları, Uzun Vadeli Tahvil Kazancı, Tüketici Fiyatları(%değişim), Döviz Kuru (Ort), İkraz Faizi (%), Para Rezervi, Para Stok Parametresi, Finansal Stok, Döviz Stoğu, Altın Rezervi ve Uluslararası Rezervler parametrelerinin oluşturduğu veri seti girdi olarak kullanılmıştır. Çıktı olarak ise Danimarka'nın 1993-2013 yılları arasında gerçekleştirdiği İthalat Tutarı (ABD Doları cinsinden-milyar) kabul edilmiş, gerçekleşen değerler girdiler kullanılarak çıktı tahmininde ve modelin tutarlılığının belirlemede kullanılmıştır.

Veri setinde girdi olarak kullanılan tüm parametreler Çizelge 5.1'de listelenmiş olup, girdileri temsil eden "x" lerin indis numaraları da karşılıklarına belirtilmiştir.

Çizelge 5.1. Veri setinde yer alan girdilerin isimleri ve indisleri

Parametre Adı	Girdi İndisi	Parametre Adı	Girdi İndisi
Döviz Kuru	3	M1 (%)	14
Reel Efektif Döviz Kuru	4	M2 (%)	15
Tüketici Fiyat İndeksi (% değişim)	5	Para Rezervi	16
Tüketici Fiyat İndeksi	6	Para Stok Parametresi-I	17
Üretici Fiyat İndeksi (% değişim)	7	Para Stok Parametresi-II	18
Üretici Fiyat İndeksi	8	Finansal Stok-I	19
Üretici Fiyatları	9	Finansal Stok-II	20
Uzun Vadeli Tahvil Kazancı	10	Finansal Stok-III	21
Tüketici Fiyatları (% değişim)	11	Döviz Stokları	22
Döviz Kuru (ortalama)	12	Altın Rezervi	23
İkrar Faizi (%)	13	Uluslararası Rezervler	24

Regresyon analizini yapmadan önce, korelasyon analizine veriler tabi tutulmuş ve en düşük korelasyona sahip indeksler modele sokulmaya çalışılmıştır. Yapılan gözlemlerin sonucunda çok düşük korelasyona sahip girdileri bulmanın oldukça zorlu olduğu ve sonuç olarak istenilen tahmin tutarlılığını elde edebilmek için, çeşitli duyarlılık analizleri ile model oluşturulmuştur. Kurulan regresyon modelinden sonra, verilerin birden fazla öbeğe de üye olabileceğini düşündüren Bulanık Regresyon analizi ile kapanış fiyatlarını tahmin etme modeli kurulmuştur.

## 5.2. Korelasyon Analizi

Regresyon analizine başlanmadan önce teknik indeksler arasında herhangi bir ilişkinin-bağımlılık durumunun (korelasyon) olup olmadığı ve teknik indeksler (girdilerin), çıktı değeri ile ne şekilde ilişkisi olduğu yapılan korelasyon çalışması ile incelenmiştir.

Korelasyon analizi, serbest ve bağımlı değişkenler arasındaki ilişki düzeyini veya derecesini ölçen analizdir. Şayet analizde tek bir bağımsız değişken varsa, bu tür analize basit korelasyon analizi, birden çok bağımsız değişken söz konusu ise buna da, çoklu korelasyon adı verilir. Bilimsel araştırmalarda en çok kullanılan basit korelasyon analizidir. Ortalamadan sapmaların esas alınarak ilişki düzeyinin saptanması için uygulanan testlere korelasyon analizi denir. [26]

Korelasyon katsayısı, değişkenlerin yönü, etkileşimlerin ne şekilde olduğu hususlarında bilgi veren bir parametredir. Değişkenlerin birbiri arasındaki etkileşimin varlığı; derecesi korelasyon analizi ile tespit edilebilmektedir.

Korelasyon katsayısı -1 ile +1 arasında değişen değerler almakta olup; katsayı, etkileşimin olmadığı durumda 0, tam ve kuvvetli bir etkileşim varsa 1, ters yönlü ve tam bir etkileşim varsa -1 değerini almaktadır. Korelasyon katsayısı genellikle r harfiyle gösterilmekte olup; buna göre korelasyon katsayısı  $-1 \leq r \leq +1$  eşitsizliğiyle temsil edilecektir [30].

Korelasyon katsayısının yorumunu, tam değerler dışında ara değerler için yapmak oldukça güçtür. Ara değerler için katsayı değerlendirirken, örnek gözlem sayısı (n) oldukça önemlidir. Çok fazla gözleme dayanan değerlendirmelerde 0.25'e kadar düşmüş bir korelasyon katsayısı bile anlamlı sayılabilmektedir. Fakat az sayıda, 10-15 gözleme dayanan değerlendirmelerde korelasyon katsayısının 0.71 üstünde olması beklenmektedir. Korelasyon katsayısının (r) değerine göre korelasyonun derecesi hakkında değerlendirmeler Çizelge 5.2'de genel geçer olarak yapılabilmektedir.

Çizelge 5.2. Korelasyon katsayısına göre sınıflandırma çizelgesi

<b>Kuvvetli (-)</b>	<b>Orta (-)</b>	<b>Zayıf (-)</b>	<b>Zayıf (+)</b>	<b>Orta (+)</b>	<b>Kuvvetli(+)</b>
$-1 \leq r < -0.9$	$-0.9 \leq r < -0.5$	$-0.5 \leq r < 0$	$0 < r \leq 0.5$	$0.5 < r \leq 0.9$	$0.9 < r \leq 1$

Özetle korelasyon, olasılık kuramı ve istatistikte iki rassal değişken arasındaki doğrusal ilişkinin yönünü ve gücünü belirtir. Genel istatistiksel kullanımda korelasyon, bağımsızlık durumundan ne kadar uzaklaşıldığını gösterir.

Bu değerlendirmeler ışığında bir önceki bölümde tanımlanan veri setinin her bir girdi değerinin bir diğeri ile etkileşimini özetleyen korelasyon ilişkileri Çizelge 5.3, 5.4 ve 5.5'te verilmiştir.

Çizelge 5.3. Korelasyon Tablosu -1 ( $x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}, x_{13}$  için)

	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$	$x_8$	$x_9$	$x_{10}$	$x_{11}$	$x_{12}$	$x_{13}$
$x_4$	-0,859										
	0										
$x_5$	-0,438	0,578									
	0	0									
$x_6$	-0,438	0,578	1								
	0	0	0								
$x_7$	-0,494	0,549	0,978	0,978							
	0	0	0	0							
$x_8$	-0,494	0,549	0,978	0,978	1						
	0	0	0	0	0						
$x_9$	-0,1	-0,079	0,199	0,199	0,298	0,298					
	0,114	0,216	0,002	0,002	0	0					
$x_{10}$	0,254	-0,417	-0,913	-0,913	-0,858	-0,858	-0,099				
	0	0	0	0	0	0	0,118				
$x_{11}$	0,188	-0,188	0,109	0,109	0,17	0,17	0,614	-0,004			
	0,003	0,003	0,084	0,084	0,007	0,007	0	0,947			
$x_{12}$	0,994	-0,863	-0,436	-0,436	-0,493	-0,493	-0,104	0,256	0,187		
	0	0	0	0	0	0	0,102	0	0,003		
$x_{13}$	0,284	-0,443	-0,764	-0,764	-0,685	-0,685	-0,048	0,796	0,015	0,287	
	0	0	0	0	0	0	0,449	0	0,816	0	

Çizelge 5.3'te belirtilen korelasyon katsayıları; ilk hücre içinde Pearson Korelasyon Katsayısı ve ikinci hücrede P- değeri olacak şekilde ifade edilmiştir.

Çizelge 5.3'te özetlendiği üzere  $x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}, x_{13}$  girdi değişkenleri arasındaki ilişkiler yapılan analiz çalışması sonrasında şu şekilde gözlenmiştir:

- $x_3$  girdi değişkeni;
  - $x_4$  değişkenleri ile orta şiddette ve ters yönlü korelasyona sahiptir
  - $x_5, x_6, x_7, x_8, x_9$  ile zayıf ters yönlü korelasyona sahiptir
  - $x_{10}, x_{11}, x_{13}$  zayıf aynı yönlü korelasyona sahiptir
  - $x_{12}$  ile güçlü ve aynı yönlü korelasyona sahiptir

Her bir deęişken için yukarıda belirtilen deęerlendirmeler korelasyon tablosu aracılığı ile yapılabilmektedir. Bu tabloda yer alan veriler yardımı ile birbiri ile yüksek korelasyon ilişkisi içinde olan girdi deęişkenlerinin modelden çıkarılması böylece birbirini tekrar eden davranış gösteren verilerden kurtulunması hedeflenmektedir.

Çizelge 5.3'te yer alan verilerden bu karara etki edecek 0,4'den daha yüksek ilişki katsayısına sahip olan gözlemler aşağıda özetlenmiştir:

- $x_3$  girdi deęişkeni  $x_4, x_5, x_6, x_7, x_8$  ile ters yönlü korelasyona sahiptir
- $x_3$  girdi deęişkeni  $x_{12}$  ile aynı yönlü korelasyona sahiptir
- $x_4$  girdi deęişkeni  $x_5, x_6, x_7, x_8$  ile aynı yönlü;  $x_{10}, x_{12}, x_{13}$  ile ters yönlü korelasyona sahiptir
- $x_5$  ile  $x_6$  deęişkenleri aynı yönlü ve tam korelasyona sahiptir
- $x_5$  girdi deęişkeni  $x_7, x_8$  ile aynı yönlü,  $x_{10}, x_{12}, x_{13}$  ile ters yönlü korelasyona sahiptir
- $x_6$  girdi deęişkeni  $x_7, x_8$  ile aynı yönlü;  $x_{10}, x_{12}, x_{13}$  ile ters yönlü korelasyona sahiptir
- $x_7$  ile  $x_8$  deęişkenleri aynı yönlü ve tam korelasyona sahiptir
- $x_7$  girdi deęişkeni  $x_4, x_5, x_6$  ile aynı yönlü;  $x_3, x_{10}, x_{12}, x_{13}$  ile ters yönlü korelasyona sahiptir
- $x_8$  girdi deęişkeni  $x_4, x_5, x_6$  ile aynı yönlü;  $x_3, x_{10}, x_{12}, x_{13}$  ile ters yönlü korelasyona sahiptir
- $x_9$  girdi deęişkeni  $x_{11}$  ile aynı yönlü korelasyona sahiptir
- $x_{10}$  ile  $x_5, x_6, x_7, x_8$  ters yönlü  $x_{13}$  ile aynı yönlü korelasyona sahiptir
- $x_{12}$  ile  $x_4, x_5, x_6, x_7, x_8$  ters yönlü;  $x_3$  ile aynı yönlü korelasyona sahiptir
- $x_{13}$  ile  $x_4, x_5, x_6, x_7, x_8$  ters yönlü korelasyona sahiptir



Çizelge 5.4. Korelasyon Tablosu -2 ( $x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}, x_{13}$  ile  $x_{14}, x_{15}, x_{16}, x_{17}, x_{18}, x_{19}, x_{20}, x_{21}, x_{22}, x_{23}, x_{24}$  için)

	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$	$x_8$	$x_9$	$x_{10}$	$x_{11}$	$x_{12}$	$x_{13}$
$x_{14}$	0,021	-0,06	-0,108	-0,108	-0,156	-0,156	-0,045	-0,106	-0,314	0,017	-0,09
	0,746	0,342	0,088	0,088	0,013	0,013	0,477	0,093	0	0,793	0,157
$x_{15}$	-0,064	0,037	-0,118	-0,118	-0,153	-0,153	-0,167	-0,035	-0,261	-0,059	0,047
	0,317	0,562	0,062	0,062	0,016	0,016	0,008	0,586	0	0,349	0,458
$x_{16}$	-0,454	0,621	0,969	0,969	0,944	0,944	0,197	-0,849	0,098	-0,45	-0,653
	0	0	0	0	0	0	0,002	0	0,124	0	0
$x_{17}$	-0,556	0,652	0,971	0,971	0,965	0,965	0,205	-0,864	0,064	-0,556	-0,694
	0	0	0	0	0	0	0,001	0	0,315	0	0
$x_{18}$	-0,587	0,7	0,944	0,944	0,936	0,936	0,158	-0,823	0,066	-0,588	-0,649
	0	0	0	0	0	0	0,013	0	0,3	0	0
$x_{19}$	-0,648	0,723	0,893	0,893	0,911	0,911	0,234	-0,746	0,132	-0,648	-0,556
	0	0	0	0	0	0	0	0	0,037	0	0
$x_{20}$	-0,601	0,699	0,933	0,933	0,94	0,94	0,257	-0,802	0,151	-0,601	-0,611
	0	0	0	0	0	0	0	0	0,017	0	0
$x_{21}$	-0,078	-0,099	-0,59	-0,59	-0,522	-0,522	-0,233	0,636	-0,177	-0,081	0,565
	0,221	0,118	0	0	0	0	0	0	0,005	0,205	0
$x_{22}$	-0,497	0,565	0,906	0,906	0,892	0,892	0,105	-0,837	-0,024	-0,495	-0,777
	0	0	0	0	0	0	0,098	0	0,702	0	0
$x_{23}$	-0,569	0,546	0,885	0,885	0,928	0,928	0,206	-0,761	0,094	-0,567	-0,642
	0	0	0	0	0	0	0,001	0	0,139	0	0
$x_{24}$	-0,501	0,566	0,907	0,907	0,895	0,895	0,109	-0,836	-0,02	-0,499	-0,773
	0	0	0	0	0	0	0,086	0	0,754	0	0

Çizelge 5.4’te belirtilen korelasyon katsayıları; ilk hücre içinde Pearson Korelasyon Katsayısı ve ikinci hücrede P- değeri olacak şekilde ifade edilmiştir. Çizelge 5.4’te de yer alan veriler incelendiğinde şu saptamalar yapılmaktadır:

- $x_3$  girdi değişkeni  $x_{16}, x_{17}, x_{18}, x_{19}, x_{20}, x_{22}, x_{23}, x_{24}$  girdi değişkenleri ile ters yönlü korelasyon ilişkisi içindedir.
- $x_4$  girdi değişkeni  $x_{16}, x_{17}, x_{18}, x_{19}, x_{20}, x_{22}, x_{23}, x_{24}$  girdi değişkenleri ile aynı yönlü korelasyon ilişkisi içindedir.
- $x_5$  girdi değişkeni  $x_{16}, x_{17}, x_{18}, x_{19}, x_{20}, x_{22}, x_{23}, x_{24}$  girdi değişkenleri ile aynı yönlü  $x_{21}$  ile ters yönlü korelasyon ilişkisi içindedir.

- $x_6$  girdi deęişkeni  $x_{16}, x_{17}, x_{18}, x_{19}, x_{20}, x_{22}, x_{23}, x_{24}$  girdi deęişkenleri ile aynı yönlü  $x_{21}$  ile ters yönlü korelasyon ilişkisi içindedir.
- $x_7$  girdi deęişkeni  $x_{16}, x_{17}, x_{18}, x_{19}, x_{20}, x_{22}, x_{23}, x_{24}$  girdi deęişkenleri ile aynı yönlü  $x_{21}$  ile ters yönlü korelasyon ilişkisi içindedir.
- $x_8$  girdi deęişkeni  $x_{16}, x_{17}, x_{18}, x_{19}, x_{20}, x_{22}, x_{23}, x_{24}$  girdi deęişkenleri ile aynı yönlü  $x_{21}$  ile ters yönlü korelasyon ilişkisi içindedir.
- $x_{10}$  girdi deęişkeni  $x_{16}, x_{17}, x_{18}, x_{19}, x_{20}, x_{22}, x_{23}, x_{24}$  girdi deęişkenleri ile ters yönlü  $x_{21}$  ile aynı yönlü korelasyon ilişkisi içindedir.
- $x_{12}$  girdi deęişkeni  $x_{16}, x_{17}, x_{18}, x_{19}, x_{20}, x_{22}, x_{23}, x_{24}$  girdi deęişkenleri ile ters yönlü korelasyon ilişkisi içindedir.
- $x_{13}$  girdi deęişkeni  $x_{16}, x_{17}, x_{18}, x_{19}, x_{20}, x_{22}, x_{23}, x_{24}$  girdi deęişkenleri ile ters yönlü  $x_{21}$  ile aynı yönlü korelasyon ilişkisi içindedir.

Çizelge 5.5. Korelasyon Tablosu -3 ( $x_{14}, x_{15}, x_{16}, x_{17}, x_{18}, x_{19}, x_{20}, x_{21}, x_{22}, x_{23}, x_{24}$  için)

	$x_{14}$	$x_{15}$	$x_{16}$	$x_{17}$	$x_{18}$	$x_{19}$	$x_{20}$	$x_{21}$	$x_{22}$	$x_{23}$
$x_{15}$	0,797									
	0									
$x_{16}$	-0,078	-0,065								
	0,22	0,306								
$x_{17}$	-0,04	-0,058	0,965							
	0,526	0,362	0							
$x_{18}$	-0,05	-0,019	0,955	0,989						
	0,431	0,76	0	0						
$x_{19}$	-0,098	-0,044	0,918	0,956	0,977					
	0,123	0,493	0	0	0					
$x_{20}$	-0,088	-0,051	0,947	0,971	0,98	0,99				
	0,165	0,424	0	0	0	0				
$x_{21}$	-0,029	0,063	-0,52	-0,443	-0,368	-0,289	-0,424			
	0,653	0,318	0	0	0	0	0			
$x_{22}$	-0,166	-0,289	0,824	0,871	0,824	0,77	0,81	-0,542		
	0,009	0	0	0	0	0	0	0		
$x_{23}$	-0,264	-0,31	0,817	0,876	0,838	0,819	0,838	-0,42	0,942	
	0	0	0	0	0	0	0	0	0	
$x_{24}$	-0,17	-0,291	0,826	0,873	0,827	0,773	0,813	-0,539	1	0,946
	0,007	0	0	0	0	0	0	0	0	0

Çizelge 5.5'te belirtilen korelasyon katsayıları; ilk hücre içinde Pearson Korelasyon Katsayısı ve ikinci hücrede P- değeri olacak şekilde ifade edilmiştir. Çizelge 5.5'te yer alan veriler incelendiğinde şu saptamalar yapılmaktadır:

- $x_{14}$  girdi değişkeni ile  $x_{15}$  aynı yönde korelasyon ilişkisi içindedir.
- $x_{15}$  girdi değişkeni  $x_{14}$  ile aynı yönde korelasyon ilişkisi içindedir.
- $x_{16}$  girdi değişkeni  $x_{17}, x_{18}, x_{19}, x_{20}, x_{22}, x_{23}, x_{24}$  ile aynı yönde;  $x_{21}$  ile ters yönde korelasyon ilişkisi içindedir.
- $x_{17}$  girdi değişkeni  $x_{16}, x_{18}, x_{19}, x_{20}, x_{22}, x_{23}, x_{24}$  ile aynı yönde;  $x_{21}$  ile ters yönde korelasyon ilişkisi içindedir.
- $x_{18}$  girdi değişkeni  $x_{16}, x_{17}, x_{19}, x_{20}, x_{22}, x_{23}, x_{24}$  ile aynı yönde;  $x_{21}$  ile ters yönde korelasyon ilişkisi içindedir.
- $x_{19}$  girdi değişkeni  $x_{16}, x_{18}, x_{17}, x_{20}, x_{22}, x_{23}, x_{24}$  ile aynı yönde;  $x_{21}$  ile ters yönde korelasyon ilişkisi içindedir.
- $x_{20}$  girdi değişkeni  $x_{16}, x_{18}, x_{19}, x_{17}, x_{22}, x_{23}, x_{24}$  ile aynı yönde;  $x_{21}$  ile ters yönde korelasyon ilişkisi içindedir.
- $x_{21}$  girdi değişkeni  $x_{16}, x_{17}, x_{18}, x_{19}, x_{20}, x_{22}, x_{23}, x_{24}$  ile ters yönde korelasyon ilişkisi içindedir.
- $x_{22}$  girdi değişkeni  $x_{16}, x_{18}, x_{19}, x_{17}, x_{23}, x_{24}$  ile aynı yönde;  $x_{21}$  ile ters yönde korelasyon ilişkisi içindedir.
- $x_{23}$  girdi değişkeni  $x_{16}, x_{18}, x_{19}, x_{17}, x_{22}, x_{20}, x_{24}$  ile aynı yönde;  $x_{21}$  ile ters yönde korelasyon ilişkisi içindedir.
- $x_{24}$  girdi değişkeni  $x_{16}, x_{18}, x_{19}, x_{17}, x_{22}, x_{20}, x_{23}$  ile aynı yönde;  $x_{21}$  ile ters yönde korelasyon ilişkisi içindedir.

Tüm bu değerlendirmeler bir arada ele alındıktan sonra korelasyon analiz çalışmaları sonucunda birbiri ile yüksek korelasyona sahip olması sebebi ile modelden çıkarılmasına karar verilen girdi verileri şu şekildedir:

**$x_5, x_6, x_7, x_8, x_{10}, x_{13}, x_{16}, x_{17}, x_{18}, x_{19}, x_{20}, x_{21}, x_{22}, x_{23}, x_{24}$**

Korelasyon analizinin ardından modelde kalmasına karar verilen girdi değişkenleri şu şekildedir:

**$x_3, x_4, x_9, x_{11}, x_{12}, x_{14}, x_{15}$**

### 5.3. Regresyon Modelinin Oluşturulması

#### 5.3.1. Basit Doğrusal Regresyon Analizi

Regresyon analizi, iki ya da daha çok değişken arasındaki ilişkiyi ölçmek için kullanılan analiz metodudur. Eğer tek bir değişken kullanılarak analiz yapılıyorsa buna tek değişkenli regresyon, birden çok değişken kullanılıyorsa çok değişkenli regresyon analizi olarak isimlendirilir. [12]

İncelenen değişkenlerin birinin bağımlı (Y), diğerlerinin bağımsız ( $x_1, x_2, x_3, \dots, x_n$ ) olması halinde; bağımlı değişkenin bağımsız değişkenlerin fonksiyonu olarak:

$Y=f(x_1, x_2, x_3, \dots, x_n)$  şeklinde ifade edilmesi regresyon analizinin konusunu oluşturur.

Yapılan bu analiz çeşitli durumlar için şu şekilde açıklanabilir:

Eğer sadece iki değişkenli (y ve  $x_1$ ) regresyon analizi yapılıyorsa, geometrik olarak X-Y düzeyinde doğrusal regresyon tahmini bu düzeyde olan noktalara en-küçük-kareler prensibine göre en iyi uyan bir doğru uydurmaktır. Eğer bütün noktalar bu tahmin edilen doğru üzerinde ise tahmin edilen doğru uyumu bu veri noktaları için hiç hatasızdır ve veri noktaları doğrudan hiç ayrılık göstermez. Hatta varyasyonu bu halde iken sıfır olur. Veri noktaları tahmin edilen doğrudan uzaklaştıkça hatalar varyasyonu çoğalmaktadır.

Üç değişkenli (y ve  $x_1$  ve  $x_2$  değişkenli) çoklu regresyonda ise tahmin edilen bir üç boyutlu düzeydir. Eğer bu düzey veri noktalarına tam olarak uyarsa bütün veri noktaları tahmin edilen düzey üstüne düşerler. Veri noktaları tahmin edilen düzey etrafında dağılmaya başlarsa, hatalar varyasyonu artmaya başlar.

Üç değişkenden daha çoklu değişkenli regresyon analizi ise grafik olarak verilemez; çok boyutlu uzayı gösterir ve tahmin edilen regresyon katsayıları ise bu çok boyutlu uzayda bir hiper-düzlem ortaya çıkarır ve bu hiper-düzeyden ayrılmalar hata varyasyonudur.

İncelenen olayda; bir bağımlı, bir bağımsız değişken varsa oluşturulacak model tek değişkenli regresyon modeli :  $Y=f(x)$ , incelenen olayda bir bağımlı, birden fazla bağımsız değişken varsa oluşturulacak model çok değişkenli regresyon modeli;  $Y=f(x_1, x_2, x_3, \dots, x_n)$  olarak adlandırılır. Ayrıca, regresyon denklemleri doğrusal ve eğrisel olmak üzere iki ayrı şekilde de sınıflandırılmaktadır.

Bu araştırma çalışması kapsamında korelasyon analizi sonrasında modelde kalmasına karar verilen  $x_3$ ,  $x_4$ ,  $x_9$ ,  $x_{11}$ ,  $x_{12}$ ,  $x_{14}$ ,  $x_{15}$  çıktı değişkenini ne şekilde etkilediğini analiz etmek amacı ile regresyon analizi çalışması yapılmıştır.

Regresyon analizinin sonuçları özet olarak Çizelge 5.6 ve 5.7’de belirtilmiştir.

Regresyon analizi kapsamında eldeki mevcut girdi değişkenleri ile oluşturulan doğrusal model Eşitlik 5.1’de belirtildiği şekildedir:

$$y = -39,14 - 0,517 *x_3 + 0,2564 *x_9 + 0,031 *x_{11} + 0,0522 *x_{15} + 0,4408 *x_4 + 0,578 *x_{12} - 0,0353 *x_{14} \quad (5.1)$$

Çizelge 5.6. Regresyon analizi çıktıları

Term	Coef	SE	Coef	T-Value	P-Value
Constant	-39,14	6,00	-6,53	0,000	
$x_3$	-0,517	0,755	-0,68	0,494	83,46
$x_9$	<b>0,2564</b>	<b>0,0361</b>	<b>7,10</b>	<b>0,000</b>	<b>2,40</b>
$x_{11}$	0,031	0,170	0,18	0,856	2,33
$x_{15}$	<b>0,0522</b>	<b>0,0224</b>	<b>2,33</b>	<b>0,021</b>	<b>3,16</b>
$x_4$	0,4408	0,0503	8,76	0,000	4,68
$x_{12}$	<b>0,578</b>	<b>0,777</b>	<b>0,74</b>	<b>0,458</b>	<b>86,13</b>
$x_{14}$	-0,0353	0,0244	-1,44	0,150	3,47

Çizelge 5.7. Regresyon analizi model tutarlılık özeti

R-kare	R-kare (adj)	R-kare (pred)
64,12%	63,08%	62,21%

Bu adımda elde edilen tutarlılık parametreleri çalışmanın ilerleyen bölümlerinde üyeliklerin ve üyeliklerin transformasyonlarının modele eklenmesi ile elde edilecek olan Bulanık Regresyon Modellerinin her birindeki tutarlılık parametreleri ile karşılaştırılacak ve modelin tahmin tutarlılığında herhangi bir iyileştirme olup olmadığı gözlemlenecektir.

## 5.4. Öbek Sayısı ve Bulanıklık Mertebesi Aralığının Belirlenmesi

Üyeliklerin hesaplanmasından önce belirlenmesi gereken parametrelerin başında öbek sayısı ve bulanıklık mertebesi gelmektedir. Modelin tutarlılığı önemli derecede etkilediği görülen bu parametrelerin problemin gerçek hayattaki yapısına uygun ve işlem hacmini çözüm sürecinin performansına olumsuz etmeyecek düzeyde etkileyecek şekilde seçilmesi son derece önemlidir. Bu seçimde kesin doğru değerleri tek bir işlem ya da algoritma aracılığı ile tespit etmek maalesef mümkün olmadığından en doğru yaklaşım belirli bir aralıkta yapılacak taramanın ardından en yüksek derecede tutarlılığı olanak sağlayan modelin seçilmesidir.

Ancak bu çalışma kapsamında Bulanık Küme Mantığı'nın başka bir deyişle Bulanık Öbek Ortalaması (BCO) Algoritması aracılığı ile elde edilen üyeliklerin, bunların transformasyonlarının Bulanık Regresyon aracılığı ile modellenmesi ve elde edilen modellerin Basit Doğrusal Regresyon (BDR) Modellemesi ile karşılaştırılmasının yanı sıra Bulanık Regresyon Modellemesi (BRM) yaklaşımında yer alan üyeliklerin ve transformasyonlarının farklı öbek sayısı ve bulanıklık mertebesi seçimine bağlı olarak gösterdiği değişimin gözlemlenmesi ve var ise bu parametrelerin modelin tutarlılığı ile sistematik olarak ilişkilendirilmesi amaçlanmıştır.

Bu kapsamda bulanıklık mertebesi ve öbek sayısını bir ya da birkaç farklı değişken olarak belirleyip sınırlı sayıda deney yapmak yerine her iki parametre içinde birer aralık belirleyip bu aralıkta her iki parametrenin aldığı değerlere göre hesaplanan üyelikler ve transformasyonlarının modelin tutarlılığı üzerine etkileri değerlendirilmiştir.

### 5.4.1. Öbek Sayısının Belirlenmesi

Öbek sayısı belirlenirken gerçek hayat problemlerinde en iyi sonucu alabilmek ve elde edilen sonuçların gerçek hayat problemine uygulanabilir olması için öbek sayısının bu kriterlere uygun olarak belirlenmesi önem teşkil etmektedir. Ancak her öbek sayısı her problemde en tutarlı modelin oluşturulması için uygun olmayabilir, bunun yerine verilerin niteliğine, birbirleri arasındaki ilişkilerine ve benzer hareket edişlerine göre verilerin öbeklenmesi istenilen öbek sayısından farklı sayıda öbek sayısı ile modelleme yapılarak daha tutarlı sonuç veren modeller elde edilebilmektedir.

Bu çalışma kapsamında ise öbek sayısının modelin tutarlılığı üzerindeki etkisi; bulanıklık mertebesi ile varsa ilişkisi farklı öbek sayısı ve bulanıklık mertebeleri kullanılarak yapılan deneyler aracılığı ile gözlemlenecektir. Yapılan deneylerde öbek sayısı  $c: 2 < c < 20$  olacak şekilde aşağıdaki kümede belirtilen şekilde olacaktır.

Öbek Sayısı: {2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20}

Modellemede sürekli 250 adet veri kullanılmış olup başka bir deyişle veri seti, üyelikler ve üyeliklerin transformasyonlarına ait matrislerin satır sayısı 250'dir.

#### **5.4.2. Bulanıklık Mertebesinin Belirlenmesi**

Daha önceden de bahsedildiği üzere yapılan araştırmalar [21] genel olarak sistem modelleme analizinde bulanıklık mertebesinin 2 olması gerektiğini göstermektedir. Daha yeni bir çalışmada [21] üyelik değeri hesaplama fonksiyonu için Taylor açılımı analizine dayanarak, bulanıklık mertebesinin minimum ve maksimum değerlerinin [1.4, 2.6] olduğu kanıtlanmıştır.

Bu çalışmanın devamı olarak da bulanıklık mertebesinin önceki çalışmalarda belirlenen ve iyi sonuç verildiği tespit edilen bu aralık içerisinde yapılmasına karar verilmiştir.

Bulanıklık mertebesi değerleri 0.1 artışla  $1.4 \leq m \leq 2.6$  olacak şekilde aşağıdaki şekilde belirtilmiştir.

$m = \{1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6\}$

#### **5.5. Üyeliklerin Hesaplanması**

Üyelikler her bir bulanıklık mertebesi ve öbek sayısı için ayrı ayrı hesaplanmıştır. Toplamda 247 farklı model oluşturularak deney yapılmış olup her bir deney için üyelik derecesi ve bulanıklık mertebesi Çizelge 5.8 ve 5.9'da yer aldığı şekilde kullanılmıştır.

Tüm modellerde 250 adet veri içeren zaman serileri aracılığı ile üyelikler hesaplanmıştır.

Çizelge 5.8. Modellerin Öbek Sayısı ve Bulanıklık Mertebesi Değişimleri- Bölüm 1

Tutarlılık (R-kare)		Öbek Sayısı Değerleri								
		2	3	4	5	6	7	8	9	10
Bulanıklık Mertebesi Değerleri	1,4	Model 2	Model 15	Model 28	Model 41	Model 54	Model 67	Model 80	Model 93	Model 106
	1,5	Model 3	Model 16	Model 29	Model 42	Model 55	Model 68	Model 81	Model 94	Model 107
	1,6	Model 4	Model 17	Model 30	Model 43	Model 56	Model 69	Model 82	Model 95	Model 108
	1,7	Model 5	Model 18	Model 31	Model 44	Model 57	Model 70	Model 83	Model 96	Model 109
	1,8	Model 6	Model 19	Model 32	Model 45	Model 58	Model 71	Model 84	Model 97	Model 110
	1,9	Model 7	Model 20	Model 33	Model 46	Model 59	Model 72	Model 85	Model 98	Model 111
	2,0	Model 8	Model 21	Model 34	Model 47	Model 60	Model 73	Model 86	Model 99	Model 112
	2,1	Model 9	Model 22	Model 35	Model 48	Model 61	Model 74	Model 87	Model 100	Model 113
	2,2	Model 10	Model 23	Model 36	Model 49	Model 62	Model 75	Model 88	Model 101	Model 114
	2,3	Model 11	Model 24	Model 37	Model 50	Model 63	Model 76	Model 89	Model 102	Model 115
	2,4	Model 12	Model 25	Model 38	Model 51	Model 64	Model 77	Model 90	Model 103	Model 116
	2,5	Model 13	Model 26	Model 39	Model 52	Model 65	Model 78	Model 91	Model 104	Model 117
2,6	Model 14	Model 27	Model 40	Model 53	Model 66	Model 79	Model 92	Model 105	Model 118	



Çizelge 5.9. Modellerin Öbek Sayısı ve Bulanıklık Mertebesi Değişimleri- Bölüm 2

Tutarlılık (R-kare)		Öbek Sayısı Değerleri									
		11	12	13	14	15	16	17	18	19	20
Bulanıklık Mertebesi Değerleri	1,4	Model 119	Model 132	Model 145	Model 158	Model 171	Model 184	Model 197	Model 210	Model 223	Model 236
	1,5	Model 120	Model 133	Model 146	Model 159	Model 172	Model 185	Model 198	Model 211	Model 224	Model 237
	1,6	Model 121	Model 134	Model 147	Model 160	Model 173	Model 186	Model 199	Model 212	Model 225	Model 238
	1,7	Model 122	Model 135	Model 148	Model 161	Model 174	Model 187	Model 200	Model 213	Model 226	Model 239
	1,8	Model 123	Model 136	Model 149	Model 162	Model 175	Model 188	Model 201	Model 214	Model 227	Model 240
	1,9	Model 124	Model 137	Model 150	Model 163	Model 176	Model 189	Model 202	Model 215	Model 228	Model 241
	2,0	Model 125	Model 138	Model 151	Model 164	Model 177	Model 190	Model 203	Model 216	Model 229	Model 242
	2,1	Model 126	Model 139	Model 152	Model 165	Model 178	Model 191	Model 204	Model 217	Model 230	Model 243
	2,2	Model 127	Model 140	Model 153	Model 166	Model 179	Model 192	Model 205	Model 218	Model 231	Model 244
	2,3	Model 128	Model 141	Model 154	Model 167	Model 180	Model 193	Model 206	Model 219	Model 232	Model 245
	2,4	Model 129	Model 142	Model 155	Model 168	Model 181	Model 194	Model 207	Model 220	Model 233	Model 246
	2,5	Model 130	Model 143	Model 156	Model 169	Model 182	Model 195	Model 208	Model 221	Model 234	Model 247
2,6	Model 131	Model 144	Model 157	Model 170	Model 183	Model 196	Model 209	Model 222	Model 235	Model 248	

Çizelge 5.8’de numaralandırılan modeller dikkat edildiğinde fark edileceği üzere Model 2’den başlanmaktadır, Model 1, bir önceki adımda yapılan regresyon analizi ardından elde edilen model olarak belirlenmiştir.

Her bir modelde öbek sayısı ve bulanıklık mertebesi değerleri dışında belirlenmesi gereken bazı diğer parametreler de bulunmaktadır. Bu parametreler öbek merkezleri ve uzaklık hesaplama ölçüsü olup her bir deney kapsamında ne şekilde belirlendiği aşağıda açıklanmıştır.

**Öbek Merkezleri:** BCO Algoritmasında [6] öbek sayısının belirlenmesinin ardından o öbekte yer alan veriler aracılığı ile öbek merkezleri belirlenmelidir. Öbek merkezleri o verileri en iyi şekilde açıklayacak ağırlık merkezi noktası olarak belirlenmektedir. Her bir veri seti için öbek merkezi değerinin belirlenmesinde verilerin öbekseneceği öbek sayısı belirleyici olmakta ve her modelde öbek sayısına göre değişmektedir.

**Uzaklık Ölçüsü:** BCO Algoritması kapsamında [6] üyelikler hesaplanırken her bir datanın ait olduğu öbeğe ve diğer öbeksene uzaklıkları dikkate alınmaktadır. Bu uzaklıklar hesaplanırken kullanılacak farklı yöntemler olmakla birlikte en yaygın kullanılan ve pratik olan Öklid Uzaklık Ölçüsü kullanılmıştır. Bu çalışmada kullanılmayan diğer ölçütler önceki bölümlerde belirtildiği üzere Minkowski, Maksimum ve Mahanolabis Uzaklık Ölçüsüdür.

Belirlenen parametreler, öbek sayısı ve bulanıklık mertebeleri değerlerine göre her bir model için BCO Öbeksene Algoritması [6] ile üyelik değerleri hesaplanmıştır.

## 5.6. Üyeliklerin Transformasyonlarının Hesaplanması

Yapılan deneylerde elde edilen üyelik değerlerine ek olarak üyeliklerin transformasyonları da her bir model için elde edilen üyelikler için hesaplanmış ve modeller için yapılan deneyler kapsamında kullanılmıştır.

Deneylerde kullanılan üyelik transformasyonları şu şekildedir:

- Üyelik Değerleri: U matrisi: [U]

$$[U] = \begin{bmatrix} \mu_1 & \cdots & \mu_n \\ \vdots & \ddots & \vdots \\ \mu_m & \cdots & \mu_{m \times n} \end{bmatrix} \quad (5.2)$$

- Üyelik Değerlerinin Karesi:  $U^2$  matrisi:  $[U^2]$

$$[U^2] = \begin{bmatrix} \mu_1^2 & \cdots & \mu_n^2 \\ \vdots & \ddots & \vdots \\ \mu_m^2 & \cdots & \mu_{mxn}^2 \end{bmatrix} \quad (5.3)$$

- Üyelik Değerlerinin Küpü:  $U^3$  matrisi:  $[U^3]$

$$[U^3] = \begin{bmatrix} \mu_1^3 & \cdots & \mu_n^3 \\ \vdots & \ddots & \vdots \\ \mu_m^3 & \cdots & \mu_{mxn}^3 \end{bmatrix} \quad (5.4)$$

- Üyelik Değerlerinin Karekökü:  $\sqrt{U}$  matrisi:  $[\sqrt{U}]$

$$[\sqrt{U}] = \begin{bmatrix} \sqrt{\mu_1} & \cdots & \sqrt{\mu_n} \\ \vdots & \ddots & \vdots \\ \sqrt{\mu_m} & \cdots & \sqrt{\mu_{mxn}} \end{bmatrix} \quad (5.5)$$

- Üyelik Değerlerinin Logaritması (10 tabanında):  $\log(U)$  matrisi:  $[\log(U)]$

$$[\log(U)] = \begin{bmatrix} \log \mu_1 & \cdots & \log \mu_n \\ \vdots & \ddots & \vdots \\ \log \mu_m & \cdots & \log \mu_{mxn} \end{bmatrix} \quad (5.6)$$

Eşitlik 5.2-5.6 kapsamında belirtilen üyeliklerin tranformasyon matrisleri oluşturulan Bulanık Regresyon Modeli (BRM) kapsamında girdi olarak kullanılmıştır

### 5.7. Bulanık Regresyon Modellerinin Tutarlılığı Üzerine Değerlendirmeler

Modellerde kullanılan girdi değerleri; girdi değişkenleri  $\{x_3, x_4, x_9, x_{11}, x_{12}, x_{14}, x_{15}\}$ , üyelikler  $[U]$ , üyeliklerin transformasyonları  $\{[U^2], [U^3], [\sqrt{U}], [\log(U)]\}$  matrisleri olarak belirlenmiş, çıktı da Danimarka'nın 1993-2013 yılları arasında ABD Doları cinsinden (milyar) İthalat Tutarı  $\{x_2\}$  olarak ifade edilmiştir.

Yapılan deneyler sonucunda elde edilen model tutarlılık parametresi olan R-kare değerleri her bir model için Çizelge 5.10'da verilmiştir. Bilindiği üzere R-kare değeri sifıra eşitse ya da sifıra yakınsa, modelin tahmin tutarlılığının iyi olmadığı kabul edilir. Bu demektir ki model ile açıklanan veri hareketleri yani varyasyon tüm varyasyonun %0'ını açıklamakta ve geometrik olarak örneklem verileri regresyon ile elde edilen hiper-düzeyin etrafında dağınık olarak bulunmaktadır. Bu sonuca varılması halinde modelin ya da çıktıların değiştirilmesi

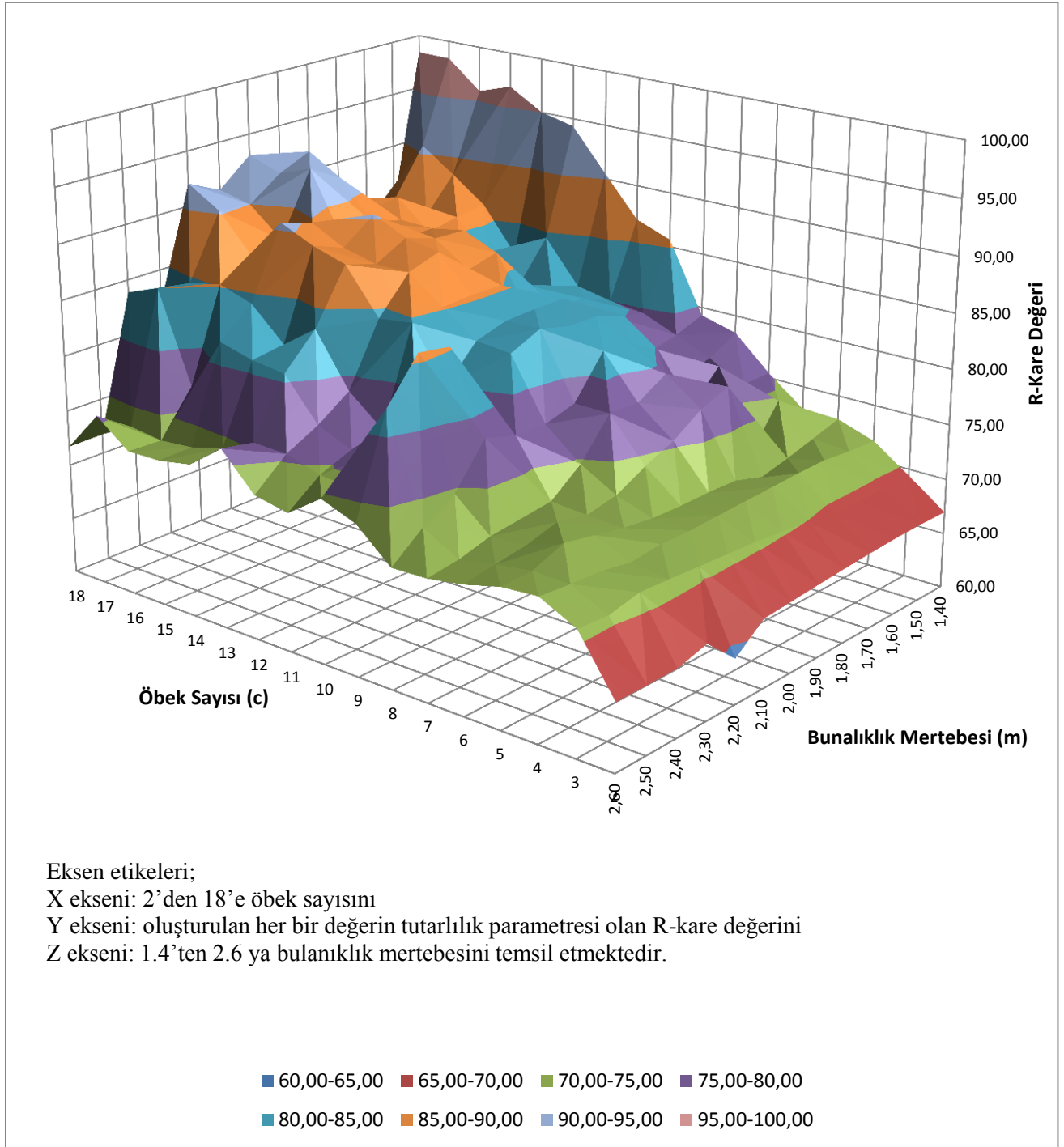
önerilir. Diğer yandan eğer R-kare değeri 100'e yakınsa, modelin tahmin tutarlılığının uygun olduğu sonucu çıkarılır. Geometrik olarak çok değişken boyutlu uzayda, model ile açıklanan veri hareketleri yani varyasyon tüm varyasyonun hepsi uzayda dağılma göstermeyip R-kare kadarı regresyon ile elde edilen hiper-düzey üzerinde bulunmaktadır.

Ayrıca üyelikler hesaplandığında öbek sayısı 18'den büyük olan modellerde üyelik değerlerinde kayda değer bir değişim tespit edilmediğinden öbeklerin örtüşmesi durumu tespit edilmiş olup bu nedenle öbek sayısı 18'den büyük olan modeller değerlendirme kapsamına alınmamıştır.

Çizelge 5.10. Model Tutarlılığı Deney Sonuçları

Tutarlılık (R-kare)		Öbek Sayısı Değerleri																
		2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Bulanıklık Mertebe Değerleri	1,40	66,97	69,42	71,72	72,84	72,94	75,43	78,40	79,26	85,64	86,82	90,51	94,29	95,02	96,83	95,82	98,37	98,41
	1,50	67,03	69,35	71,86	72,12	74,75	75,40	77,05	78,37	79,65	80,03	80,67	84,54	82,26	86,31	88,55	90,69	84,37
	1,60	67,00	69,38	72,16	71,42	71,07	78,12	76,20	77,77	80,17	81,20	81,65	80,96	83,32	85,96	85,40	86,55	83,36
	1,70	66,89	69,41	72,14	70,82	78,15	75,45	79,82	82,10	80,33	81,23	82,65	81,33	84,51	85,40	87,97	86,91	85,63
	1,80	66,75	69,94	72,05	71,66	77,16	79,50	81,36	83,07	83,83	81,87	86,55	87,86	87,32	87,41	87,65	88,10	90,86
	1,90	66,58	70,25	71,95	71,77	75,42	77,94	81,73	83,11	83,85	85,24	85,84	88,00	87,28	88,09	91,47	93,27	92,20
	2,00	66,38	70,55	71,86	71,75	74,30	78,08	78,08	83,86	83,75	85,27	88,38	88,92	90,31	87,35	93,12	92,02	93,06
	2,10	66,20	70,74	71,76	71,85	72,90	74,69	77,31	82,67	82,96	85,14	87,48	87,45	88,19	89,74	87,47	89,98	90,64
	2,20	64,05	70,92	71,83	72,22	72,34	72,57	77,39	78,76	80,02	84,73	84,31	88,33	89,25	90,60	86,98	90,26	92,05
	2,30	66,70	71,12	71,52	72,80	72,94	72,23	78,31	85,28	77,77	76,31	78,86	83,55	81,37	84,24	85,20	84,56	81,99
	2,40	65,96	71,28	72,00	73,36	72,18	71,92	74,61	85,76	76,50	75,64	76,91	80,59	80,95	81,75	76,90	85,00	83,87
	2,50	65,95	70,66	72,18	72,28	71,76	71,16	72,92	82,10	74,31	74,55	75,46	75,55	74,14	73,30	73,00	73,38	71,74
	2,60	65,96	70,86	72,45	72,19	71,32	70,78	70,66	73,43	74,59	72,37	72,97	76,15	73,78	73,21	73,12	75,43	71,71

Bu tabloda elde edilen veriler incelendiğinde Şekil 5.2'deki grafik elde edilmiştir.



Şekil 5.2. Model tutarlılığı deney sonuçları grafiksel gösterim

Bu deneyler ve ilgili grafik detaylı olarak incelendiğinde temel olarak üç farklı çıkarım elde edilmektedir.

**Çıkarım 1:** Üyeliklerin ve üyeliklerin transformasyonlarının modelde yer alması her durumda modelin tutarlılığını artırmıştır.

Model 1 'de elde edilen modelin R-kare değeri Çizelge 5.11'de özetlenmiştir.

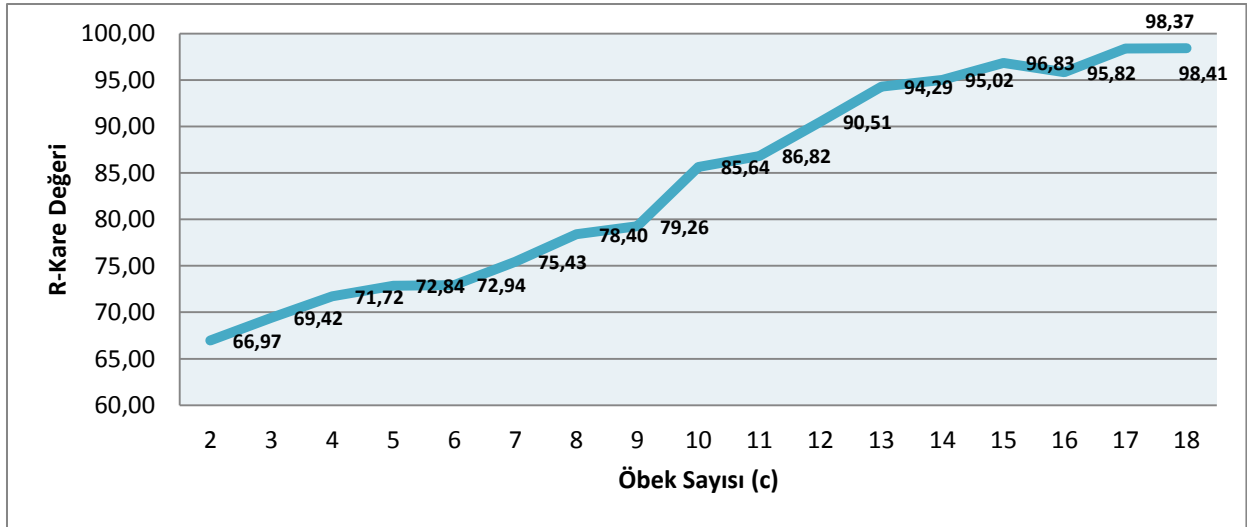
Çizelge 5.11. Regresyon Modeli Tutarlılık Parametreleri

R-kare	R-kare (adj)	R-kare (pred)
64,12%	63,08%	62,21%

Model 1 'in R-kare değeri olan 64,12 değerinden daha düşük bir tutarlılık değeri Çizelge 5.10'da gözlemlenmediğinden herhangi bir öbek sayısı ve bulanıklık mertebesi için hesaplanan üyelikler ve bu üyeliklerin transformasyonlarının modelin tutarlılığını her durumda artırdığını söylemek mümkündür.

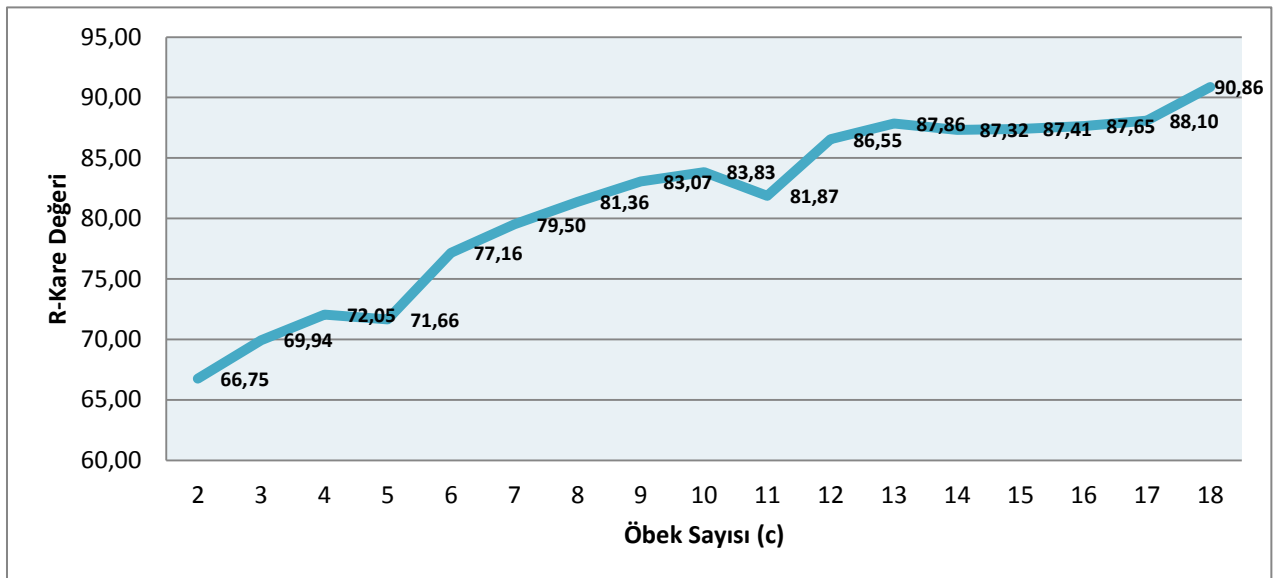
Bu bölüm kapsamında açıklanan çıkarım sadece bulanıklık mertebesinin 1.4,1.8, 2.2 olduğu ve 2'den 18'e birer artışla tüm öbek sayıları için hesaplanan modeller özelinde yapılmıştır ancak bu çalışma kapsamında gerçekleştirilen 221 deneyde de bu çıkarıma aykırı bir duruma rastlanmamış olup tüm çalışmayı kapsamaktadır.

**Çıkarım 2:** Üyelikler hesaplanırken öngörülen öbek sayısının artışı modelin tahmin tutarlılığında iyileşmeye olanak sağlamıştır. Bulanıklık mertebesi 1.4,1.8 ve 2.2 olan modellerin farklı öbek sayıları için tutarlılık değerleri Çizelge 5.10'da belirtilmiştir.



Şekil 5.3. Bulanıklık mertebesinin 1.4 olduğu modellerde tutarlılık parametresinin öbek sayısına göre değişimi grafiği

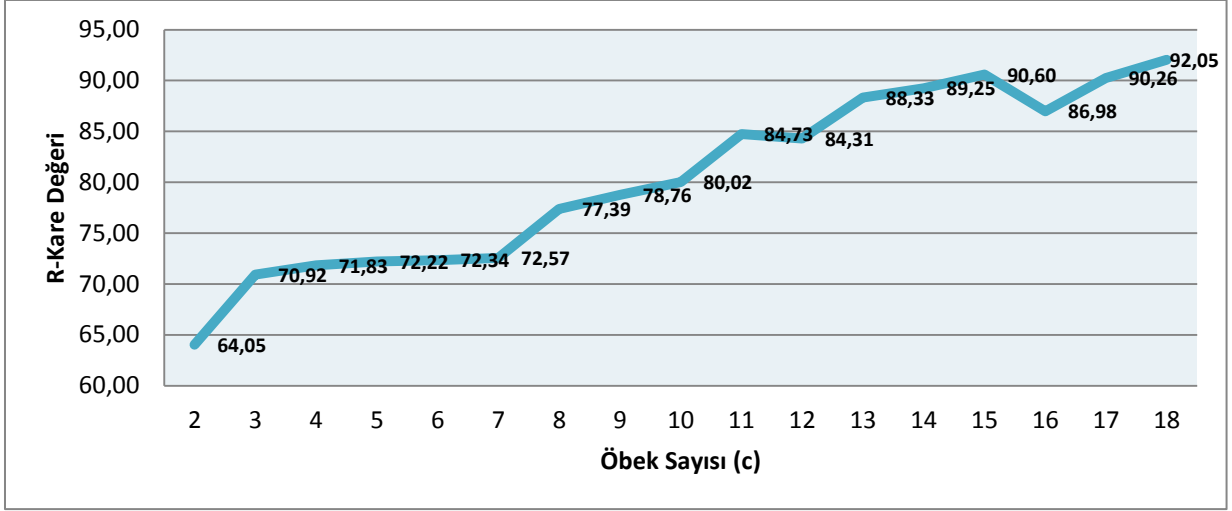
Bulanıklık mertebesinin 1.4 değerini aldığı modellerin farklı öbek sayılarına göre tutarlılık parametresi olan R-kare değerlerinin değişim grafiği Şekil 5.3'te de yer almaktadır. Görüldüğü üzere öbek sayısının 2 olduğu modelde R-kare değeri 66,97 değerini alırken; bulanıklık mertebesi sabit tutulup modeldeki öbek sayısı artırıldığında elde edilen R-kare değerinin artış gösterdiği ve öbek sayısının 18 e ulaştığı noktada 98,14 değerine kadar ulaştığı gözlemlenmiştir.



Şekil 5.4. Bulanıklık mertebesinin 1.8 olduğu modellerde tutarlılık parametresinin öbek sayısına göre değişimi grafiği



Bulanıklık mertebesinin 1.8 değerini aldığı modellerin farklı öbek sayılarına göre tutarlılık parametresi olan R-kare değerlerinin değişim grafiği Şekil 5.4'te yer almaktadır. Görüldüğü üzere öbek sayısının 2 olduğu modelde R-kare değeri 66,75 değerini alırken; bulanıklık mertebesi sabit tutulup modeldeki öbek sayısı artırıldığında elde edilen R-kare değerinin artış gösterdiği ve öbek sayısının 18 e ulaştığı noktada 90,86 değerine kadar ulaştığı gözlemlenmiştir.



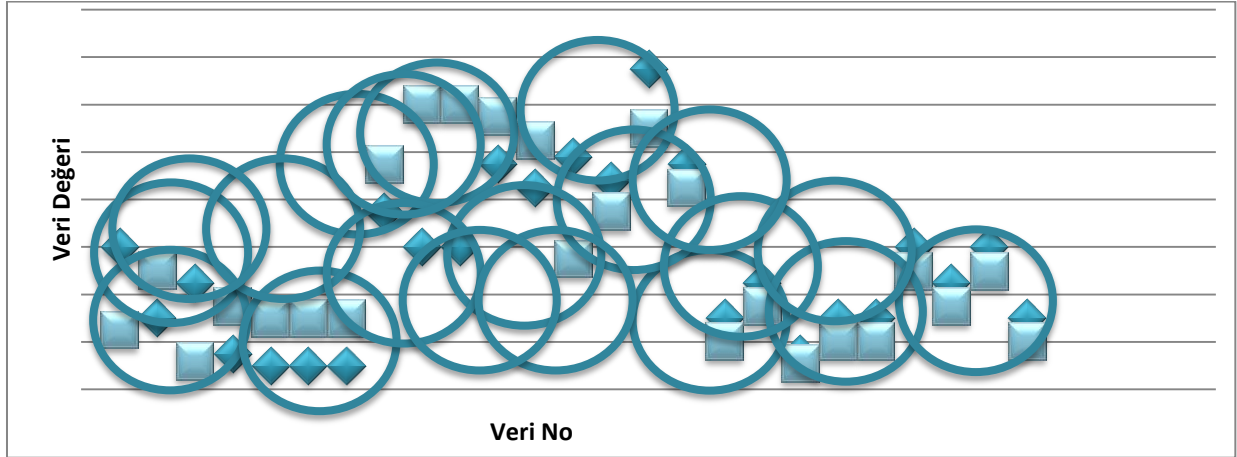
Şekil 5.5. Bulanıklık mertebesinin 2.2 olduğu modellerde tutarlılık parametresinin öbek sayısına göre değişimi grafiği

Bulanıklık mertebesinin 2.2 değerini aldığı modellerin farklı öbek sayılarına göre tutarlılık parametresi olan R-kare değerlerinin değişim grafiği Şekil 5.5'te yer almaktadır. Görüldüğü üzere öbek sayısının 2 olduğu modelde R-kare değeri 64,05 değerini alırken; bulanıklık mertebesi sabit tutulup modeldeki öbek sayısı artırıldığında elde edilen R-kare değerinin artış gösterdiği ve öbek sayısının 18 e ulaştığı noktada 92,05 değerine kadar ulaştığı gözlemlenmiştir.

Bu bölüm kapsamında açıklanan çıkarım sadece bulanıklık mertebesinin 1.4,1.8, 2.2 olduğu ve 2'den 18'e birer artışla tüm öbek sayıları için hesaplanan modeller özelinde yapılmıştır ancak bu çalışma kapsamında gerçekleştirilen 221 deneyde de bu çıkarıma aykırı bir duruma rastlanmamış olup tüm çalışmayı kapsamaktadır.

**Çıkarım 3:** Öbek sayısındaki artışın modelin tutarlılığını artırmada olumlu rol oynadığı gözlemler sonucunda belirlenmiş olmakla birlikte bir modelde çok sayıda öbekleşme olması modelde kullanılan üyelik değeri matrisindeki kolon sayısını, aynı şekilde üyeliklerin

transformasyonları matrisindeki kolon sayısını artırdığından tercih edilen bir durum değildir. Öbek sayısı artışının, modelin iyileşme ivmesi ile denge içinde belirli bir noktada tutulması yapılan deneyler sonucunda elde edilen gözlemlerin bir sonucu olarak önerilmektedir. Bu durum temsili olarak Şekil 5.6’da resmedilmiştir.



Şekil 5.6. Temsili örtüşen öbekleşme gösterimi

Başka bir deyişle modelde üyeliklerin hesaplanmasında kullanılan öbek sayısı değerlerinin çok yüksek seçilmesi; modelin oluşturulması ve tahmin yapılmasını zorlaştırıcı etki oluşturabileceğinden, kontrol edilebilir ve kullanım kolaylığının sürdürülebilir büyüklükte kalması aynı zamanda tahmin tutarlılığını istenen seviyeye yakın olmasını sağlayacak şekilde bir dengede seçilmelidir.

Örneğin Çizelge 5.12’de görülebileceği gibi öbek sayısının 2’den 3’e çıkarılması modelin tutarlılığı üzerinde %3.66 iyileşme sağlamıştır ve bu şekilde elde edilen %69,42 lik R-kare değeri modelin tutarlılığı açısından yeterli değildir. Yine aynı şekilde öbek sayısının 13’ten 14’e çıkarılması modelin tutarlılığı üzerinde %0.77 iyileşme sağlamıştır, ancak bu şekilde elde edilen %95.02 lik R-kare değeri modelin tutarlılığı açısından yeterli seviyededir.

Çizelge 5.12. Bulanıklık mertebesinin 1.4 olduğu modellerde model tutarlılığı değişimleri

Öbek Sayısı	R-kare Değeri	% Değişim	Öbek Sayısı	R-kare Değeri	% Değişim
2	66,97	x	11	86,82	1,38%
3	69,42	3,66%	12	90,51	4,25%
4	71,72	3,31%	13	94,29	4,18%
5	72,84	1,56%	14	95,02	0,77%
6	72,94	0,14%	15	96,83	1,90%
7	75,43	3,41%	16	95,82	-1,04%
8	78,4	3,94%	17	98,37	2,66%
9	79,26	1,10%	18	98,41	0,04%
10	85,64	8,05%			

Sabit bulanıklık mertebesi, bulanıklık mertebesinin 1.4 olduğu durumda öbek sayısının bir artırılması ile model tutarlılığında, başka bir deyişle R-kare değerindeki artış Çizelge 5.12’de gösterilmektedir.

Yine Çizelge 5.10’da yer alan veriler temel alınarak hesaplandığında görülmektedir ki öbek sayısının 13’ten 18’e çıkarılması modelin tutarlılığı üzerinde %4.36 iyileşme sağlamıştır, bu öbek sayısındaki artış modelin işlem hacmini artırışı, işlem zorluğu ve süresi göz önünde bulundurulduğunda tercih edilmeyebilir.

Bir başka örnek de Çizelge 5.10’da yer alan veriler olup, tabloda görülebileceği gibi öbek sayısının 2’den 3’e çıkarılması modelin tutarlılığı üzerinde %4.8 iyileşme sağlamıştır ve bu şekilde elde edilen %69,94’lük R-kare değeri modelin tutarlılığı açısından yeterli değildir. Yine aynı şekilde öbek sayısının 15’ten 16’ya çıkarılması modelin tutarlılığı üzerinde %0.1 iyileşme sağlamıştır, ancak bu şekilde elde edilen %87,41 lik R-kare değeri modelin 1.8 bulanıklık mertebesi tutarlılığı açısından yeterli seviyededir.

Çizelge 5.13. Bulanıklık mertebesinin 1.8 olduğu modellerde model tutarlılığı değişimleri

Öbek Sayısı	R-kare Değeri	% Değişim	Öbek Sayısı	R-kare Değeri	% Değişim
2	66,75	x	11	81,87	-2,3%
3	69,94	4,8%	12	86,55	5,7%
4	72,05	3,0%	13	87,86	1,5%
5	71,66	-0,5%	14	87,32	-0,6%
6	77,16	7,7%	15	87,41	0,1%
7	79,5	3,0%	16	87,65	0,3%
8	81,36	2,3%	17	88,1	0,5%
9	83,07	2,1%	18	90,86	3,1%
10	83,83	0,9%			

Sabit bulanıklık mertebesi, bulanıklık mertebesinin 1.8 olduğu durumda öbek sayısının bir artırılması ile model tutarlılığında, başka bir deyişle R-kare değerindeki artış Çizelge 5.13'te gösterilmektedir.

Yine Çizelge 5.10'da yer alan veriler temel alınarak hesaplandığında görülmektedir ki öbek sayısının 13'ten 18'e çıkarılması modelin tutarlılığı üzerinde %3,41 iyileşme sağlamıştır, bu öbek sayısındaki artış modelin işlem hacmini artırışı, işlem zorluğu ve süresi göz önünde bulundurulduğunda tercih edilmeyebilir.

Aynı şekilde bulanıklık mertebesinin 2.2 olduğu durumda öbek sayısının bir artırılması ile model tutarlılığında, başka bir deyişle R-kare değerindeki artış Çizelge 5.14'de gösterilmektedir.

Çizelge 5.14. Bulanıklık mertebesinin 2.2 olduğu modellerde model tutarlılığı değişimleri

Öbek Sayısı	R-kare Değeri	% Değişim	Öbek Sayısı	R-kare Değeri	% Değişim
2	64,05	x	11	84,73	5,9%
3	70,92	10,7%	12	84,31	-0,5%
4	71,83	1,3%	13	88,33	4,8%
5	72,22	0,5%	14	89,25	1,0%
6	72,34	0,2%	15	90,6	1,5%
7	72,57	0,3%	16	86,98	-4,0%
8	77,39	6,6%	17	90,26	3,8%
9	78,76	1,8%	18	92,05	2,0%
10	80,02	1,6%			

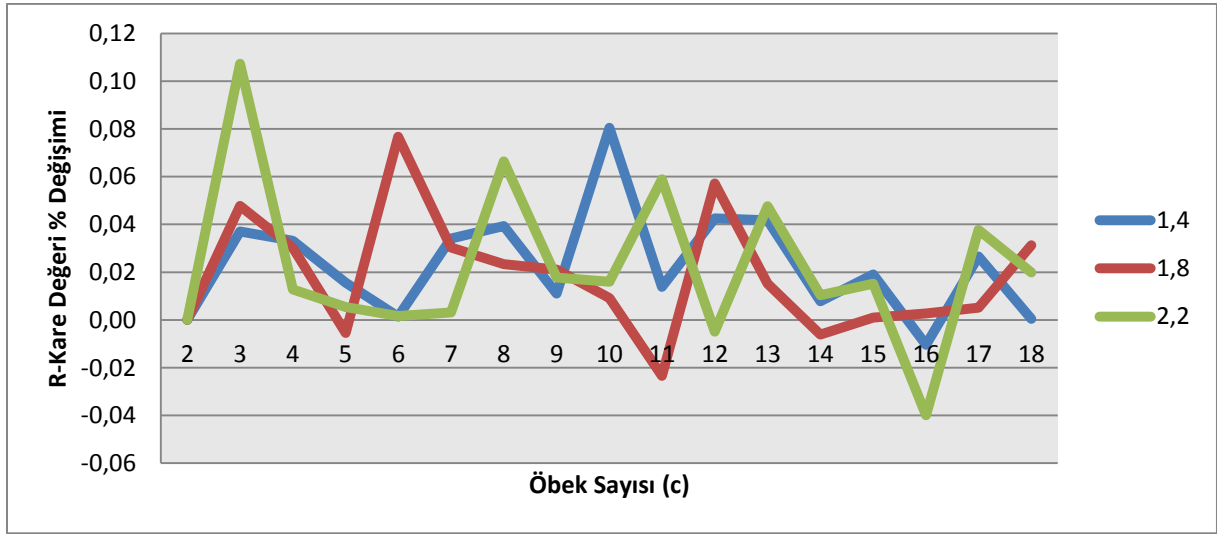
Çizelge 5.15. Bulanıklık mertebesi ve öbek sayısına göre model tutarlılığı değişimi

Öbek Sayısı / Bulanıklık Mertebesi	M=1,4	M=1,8	M=2,2	Öbek Sayısı / Bulanıklık Mertebesi	M=1,4	M=1,8	M=2,2
C=2	x	x	x	C=11	0,01	-2,3%	5,9%
C=3	3,7%	4,8%	10,7%	C=12	4,3%	5,7%	-0,5%
C=4	0,03	3,0%	1,3%	C=13	0,04	1,5%	4,8%
C=5	1,6%	-0,5%	0,5%	C=14	0,8%	-0,6%	1,0%
C=6	0,00	7,7%	0,2%	C=15	0,02	0,1%	1,5%
C=7	3,4%	3,0%	0,3%	C=16	-1,0%	0,3%	-4,0%
C=8	0,04	2,3%	6,6%	C=17	0,03	0,5%	3,8%
C=9	1,1%	2,1%	1,8%	C=18	0,0%	3,1%	2,0%
C=10	0,08	0,9%	1,6%				

Son olarak Çizelge 5.15'te de benzer yorumlar yapmak mümkündür ancak şimdiye dek bahsi geçen hususlara ek olarak dikkat edilmesi gereken bir hususta sabit bulanıklık mertebesi değeri için öbek sayısının kademeli artışında R-kare değerinin artış yerine azalma göstermesinin mümkün olabileceğidir.

Örneğin Çizelge 5.15'te görülebileceği gibi bulanıklık mertebesi 1.4 olduğu durumda öbek sayısı 15'ten 16'ya artarken ve R-kare değerinin artış göstermesi beklenirken aynı zamanda da tüm bu öbek sayısı artışlarında R-kare değeri pozitif bir artış ivmesine sahipken sadece bir değer özelinde, %1 lik bir azalma tespit edilmiştir.

Modelin tutarlılık parametresi olan R-kare değerindeki bu azalmanın 221 deney kapsamında elde edilen çıkarımı olumsuz etkileyecek sıklıkta ve önem seviyesinde olmadığı değerlendirilmiştir.

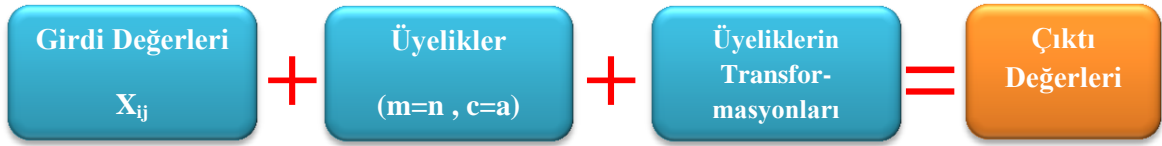


Şekil 5.7. Bulanıklık mertebesi 1.4,1.8 ve 2.2 olan modellerde model tutarlılığının öbek sayısına göre değişimi

Bu bölüm kapsamında bahsedilen çıkarımlar sadece bulanıklık mertebesinin 1.4,1.8, 2.2 olduğu ve 2'den 18'e birer artışla tüm öbek sayıları için hesaplanan modeller özelinde yapılmıştır ancak bu çalışma kapsamında gerçekleştirilen 221 deneyde de bu çıkarıma aykırı bir duruma rastlanmamış olup tüm çalışmayı kapsamaktadır.

## 6. BULANIK KÜME ALGORİTMASINDA YENİ BİR YAKLAŞIM: ÇOK KATMANLI BULANIK REGRESYON MODELİ

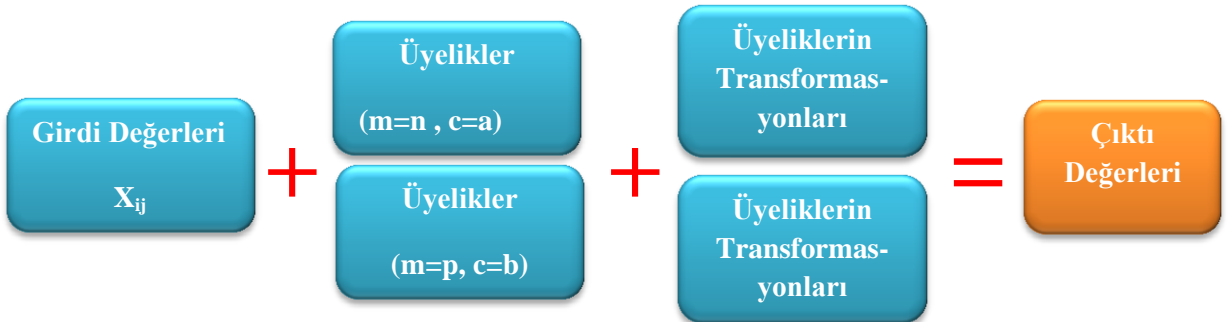
Çok katmanlı üyeliklerin modelde kullanımı ilk kez bu çalışma kapsamında yapılan deney verileri aracılığı ile gerçekleştirilmiştir. Bilindiği üzere yaygın ve hali hazırda kullanımda olan yaklaşım gereği belirlenen birer bulanıklık mertebesi ve öbek sayısı değerleri için üyelik değerleri ve transformasyonları hesaplanır doğrusal regresyon modelinde kullanılarak modelin tutarlılığındaki değişimler gözlemlenir, şimdiye kadar yapılan çalışmalar ve bu tez çalışması kapsamında yapılan deneyler de göstermiştir ki üyelikler ve üyeliklerin transformasyonları modelin tutarlılığının artırılmasında önemli rol oynamıştır.



Şekil 6.1. Bulanık Regresyon Modeli temsili gösterimi [26]

Şekil 6.1'de görülebileceği üzere doğrusal regresyon modeline girdi olarak eklenen üyelikler ve transformasyonları bu çalışma dahil bir çok çalışmada model tutarlılığını artıran genel geçer bir model ve yaklaşımdır.

Bu çalışma kapsamında ilk kez uygulanan ve sonuçları gözlemlenen yaklaşım ise Bulanık Regresyon Modeli (BRM) [28] kapsamında sadece girdi değerlerinin, tek bir bulanıklık mertebesi ve öbek sayısı için hesaplanmış üyelik değerlerinin, bu üyelik değerlerinin transformasyonlarının modelde çıktı tahmini amaçlı kullanımının yanı sıra farklı bir üyelik matrisi ve üyeliklerin transformasyonları martisinin de modelde kullanılarak elde edilen melez modelin tutarlılığının gözlemlenmesidir.



Şekil 6.2. Çok Katmanlı Bulanık Regresyon Modeli temsili gösterimi

Şekil 6.2’de temsil edilen ve ilk kez bu çalışma kapsamında uygulanan Çok Katmanlı Bulanık Regresyon Modeli (ÇKBRM) yaklaşımı ise mevcut Bulanık Regresyon modeline farklı öbek sayısı ve/veya bulanıklık mertebesi için hesaplanmış üyelikler ve transformasyonlarının eklenmesi ve öngörüldüğü üzere Klasik Bulanık Regresyon Modeline kıyasla her deneyde daha yüksek R-kare değeri ile daha tutarlı sonuçlar veren model oluşturulmasına olanak sağlamaktadır.

### 6.1. Çok Katmanlı Bulanık Regresyon Modellerinin Tanımlanması

Bu çalışmada tek bir modelde birden fazla öbek sayısı ve bulanıklık mertebesi için hesaplanan birden fazla üyelik ve üyeliklerin transformasyon matrislerinin modelde girdi olarak kullanılması denenecek olup modelin tutarlılığı üzerinde bu uygulamanın etkisi gözlemlenecektir.

A modeli bulanıklık mertebesi n, öbek sayısı a olan bir model olsun.

6.1’de yer alan  $U_A$  matrisi A modeli kapsamında BCO Algoritması aracılığı ile hesaplanan üyelik değerlerinin depolandığı matris olsun.

$\mu_{A,1,1}$  : A Modeli kapsamında birinci girdi değerinin birinci öbeğe aitlik derecesi, üyelik değeri olsun.

$$U_A = \begin{bmatrix} \mu_{A,1,1} & \mu_{A,2,1} & \cdots & \mu_{A,c,1} \\ \cdots & \cdots & \cdots & \cdots \\ \mu_{A,1,n} & \mu_{A,2,n} & \cdots & \mu_{A,c,n} \end{bmatrix} \quad (6.1)$$

Aynı şekilde B modelini tanımlayalım.

B modeli bulanıklık mertebesi p, öbek sayısı b olan bir model olsun.

6.2’de yer alan  $U_B$  matrisi B modeli kapsamında BCO Algoritması aracılığı ile hesaplanan üyelik değerlerinin depolandığı matris olsun.

$\mu_{B,1,1}$  : B Modeli kapsamında birinci girdi değerinin birinci öbeğe aitlik derecesi, üyelik değeri olsun.

$$U_B = \begin{bmatrix} \mu_{B,1,1} & \mu_{B,2,1} & \cdots & \mu_{B,c,1} \\ \cdots & \cdots & \cdots & \cdots \\ \mu_{B,1,n} & \mu_{B,2,n} & \cdots & \mu_{B,c,n} \end{bmatrix} \quad (6.2)$$



$X^A$  matrisi BRM kapsamında girdi olarak kabul edilecek A modeline göre hesaplanmış üyelik değerleri, bu üyeliklerin transformasyonları ve girdi verilerinin depolandığı genel girdi matrisidir.

$x_i$  : girdi verilerini temsil etmektedir.

$\mu^2_{1,n}$ : A Modeli kapsamında hesaplanan üyeliklerin transformasyonlarını temsil etmektedir.

$$X^A = \begin{bmatrix} 1 & \mu_{A,1,1} & \mu^2_{A,1,1} & x_{1x1} & \cdots & x_{1xnv} \\ 1 & \mu_{A,1,2} & \mu^2_{A,1,2} & x_{2x1} & \cdots & x_{2xnv} \\ \vdots & \vdots & \vdots & \vdots & \cdots & \cdots \\ 1 & \mu_{A,1,n} & \mu^2_{A,1,n} & x_{nx1} & \cdots & x_{nxnv} \end{bmatrix} \quad (6.3)$$

BRM kapsamında girdi değerlerinin 6.3'te yer aldığı üzere; tek bir bulanıklık mertebesi ve öbek sayısı için hesaplanmış üyelik değerlerinin, bu üyelik değerlerinin transformasyonlarının özete Model A'nın verilerinin modelde çıktı tahmini amaçlı kullanımı mümkün olup bu tip regresyon modellemesinin BDR'ye kıyasla her durumda daha iyi sonuç verdiği önceki bölümlerde gösterilmiştir.

$$X^{A+B} = \begin{bmatrix} 1 & \mu_{A,1,1} & \mu^2_{A,1,1} & \mu_{B,1,1} & \mu^2_{B,1,1} & x_{1x1} & \cdots & x_{1xnv} \\ 1 & \mu_{A,1,2} & \mu^2_{A,1,2} & \mu_{B,1,2} & \mu^2_{B,1,2} & x_{2x1} & \cdots & x_{2xnv} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \cdots \\ 1 & \mu_{A,1,n} & \mu^2_{A,1,n} & \mu_{B,1,n} & \mu^2_{B,1,n} & x_{nx1} & \cdots & x_{nxnv} \end{bmatrix} \quad (6.4)$$

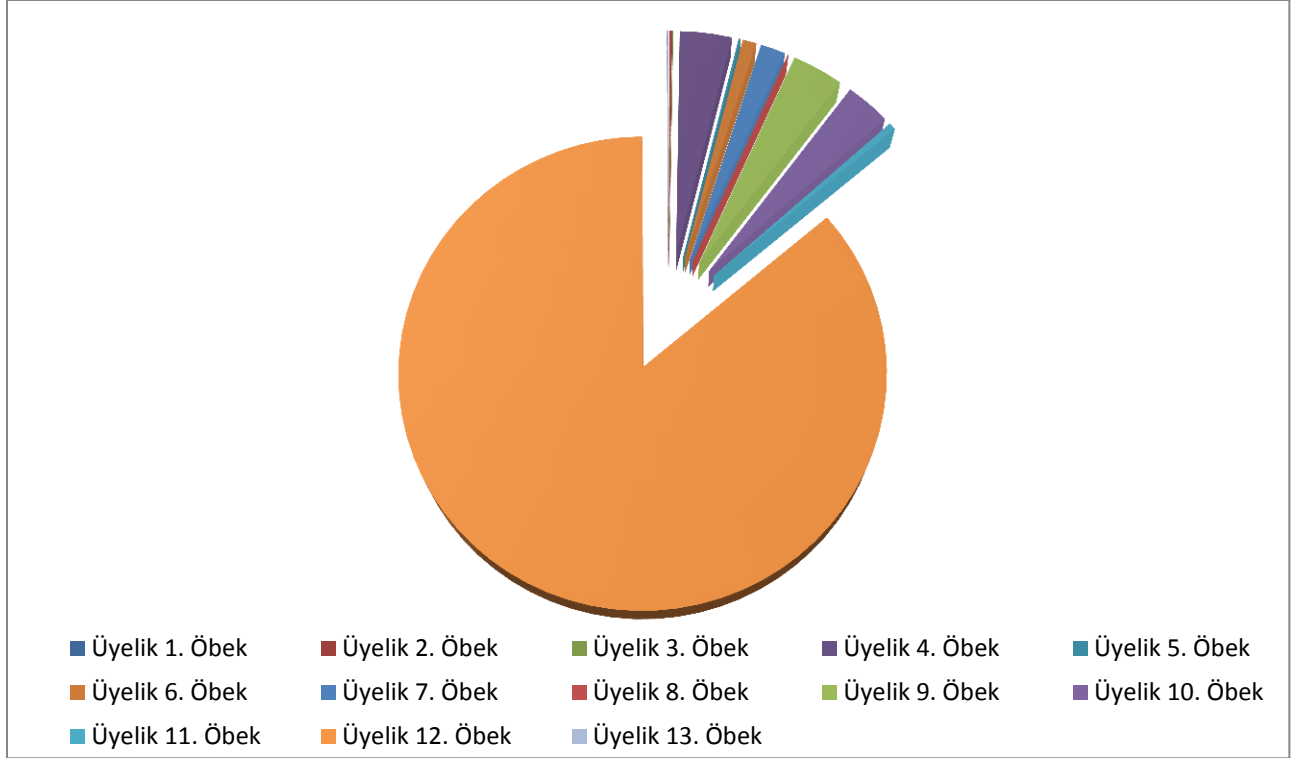
İlk kez bu çalışma kapsamında uygulanan Çok Katmanlı Bulanık Regresyon Modeli (ÇKBRM) yaklaşımı ise mevcut Bulanık Regresyon modeline farklı öbek sayısı ve/veya bulanıklık mertebesi için hesaplanmış üyelikler ve transformasyonlarının eklenmesi aracılığı ile oluşan A+B gibi melez modeller bu tez çalışması kapsamında oluşturulmuş ve modellerin tutarlılık parametresi olan R-kare değerlerindeki değişiklikler gözlemlenmiştir. İlgili ÇKBRM Modelinin girdi matrisi 6.4'te yer almaktadır.

## 6.2. Çok Katmanlı Üyeliklerin Modelin Tutarlılığı Üzerine Etkisi

Çok Katmanlı Üyeliklerin Bulanık Regresyon Modeline eklenmesinin ardından modeldeki etkisi iyileştirmeye olanak sağlayıp sağlamadığı tespit edilmeye çalışılmıştır. Hali hazırda oluşturulan 221 farklı model üzerinden aşağıdaki modeller seçilmiş bu seçilen modeller Çok Katmanlı Bulanık Regresyon Modeline uygun olacak şekilde melez modeller oluşturulmuş ve yeni modellerin tutarlılık parametresi olan R-kare değerleri incelenmiştir.

## Deneme 1:

Bu deneme kapsamında ele alınan ilk model öbek sayısı 13, bulanıklık mertebesi 1.4 olan, R-kare değeri %94.29 olan ve Model A olarak isimlendirilen modeldir.

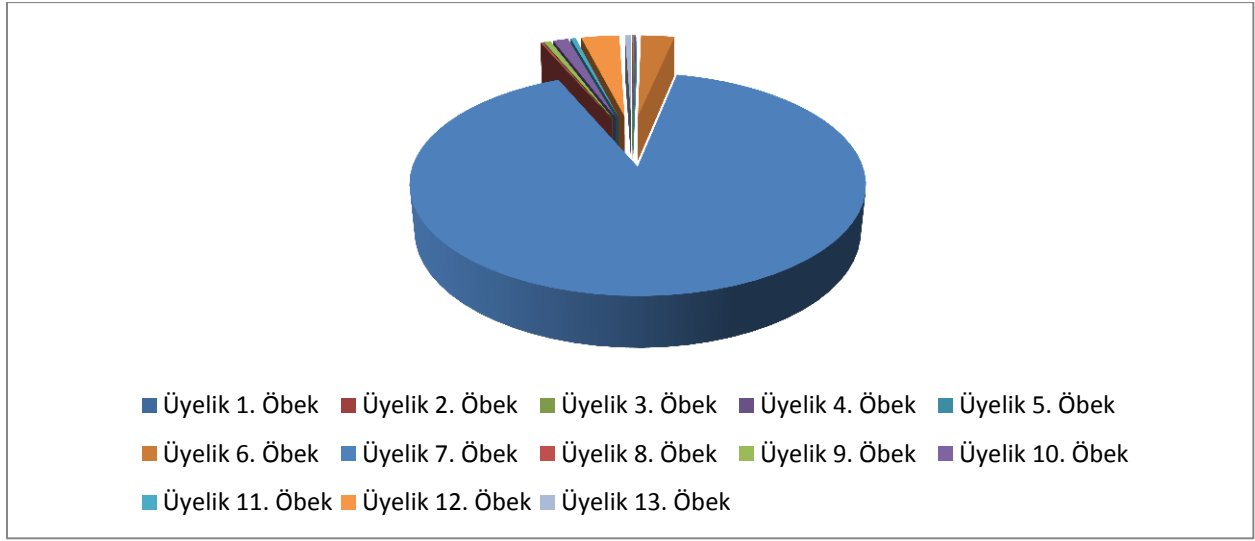


Şekil 6.3: Model A'da yer alan bir verinin öbeklere aitliğinin gösterimi

Model A'da yer alan tek bir verinin her bir öbeğe üyelik dereceleri Şekil 6.3'te gösterildiği gibi olup bu verinin 12. Öbeğin karakteristik özelliklerine daha uygun nitelikte olduğu üyelik değerinin diğer öbeklerin üyelik değerlerine kıyasla çok daha yüksek olması sebebi ile açıktır. Başka bir deyişle bu verinin yüksek tutarlılıkla modellenmesi amacı ile oluşturulacak Bulanık Regresyon Modelinde en büyük ağırlığa sahip olacak; dolayısı ile sonuç tahminine en etkin şekilde etki edecek öbek 12. Öbeğdir.

Bu verinin diğer öbeklere de belli derecede üye olup her bir verinin tüm öbeklere üyelikleri toplamı 1'dir. Yani bu verinin modellenmesi ve karakteristik özelliklerinin tespitinde diğer öbeklerin de az da olsa etkisi vardır, doğrusal regresyon modelleri üyelik değerlerini ihtiva etmediğinden her bir verinin karakteristik özelliğini tek bir öbekmişçesine modelleme ve tahmin etmeye çalışır bu nedenle de bu tez çalışmasındaki gibi birden çok öbeğin karakteristik özelliklerini gösterir nitelikte veriler için yüksek tutarlılıkta modellerin elde edilmesi zor hatta

imkansızdır. Az da olsa mevcut olan farklı öbeklere üyelikler sayesinde verinin daha yüksek tutarlıkta tahmini mümkün olabilmektedir.



Şekil 6.4: Model A’da yer alan bir diğer verinin öbeklere aitliğinin gösterimi

Aynı şekilde başka bir veri dikkate alındığında ise yukarıda yer alan Şekil 6.4 elde edilmiştir. Bu verinin ise bir öncekinin aksine 12. Öbek yerine 7. Öbeğe üyeliğinin en yüksek olacak şekilde vuku bulduğu gözlemlenmiştir.

Model A özetle;

**Model A: c=13, m=1.4 (%94.29)** şeklinde ifade edilmektedir.

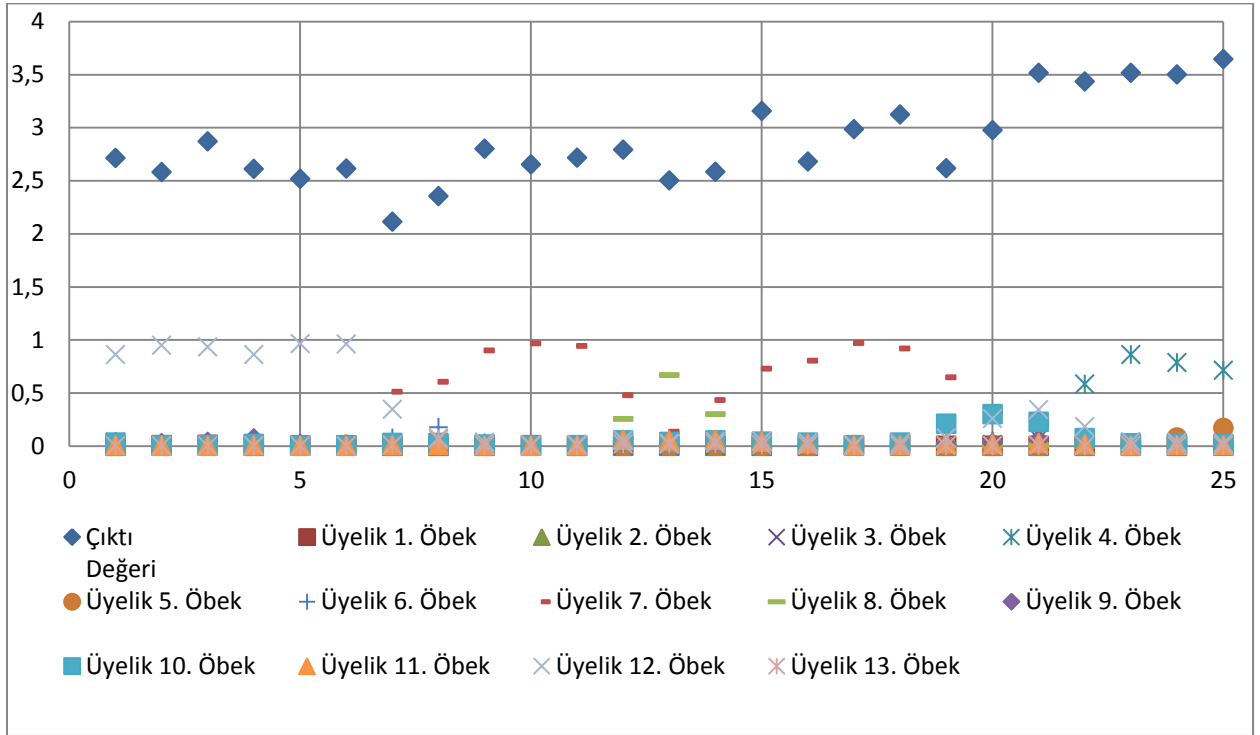
Bu model kapsamında elde edilen üyelik değerleri 6.5’te yer alan  $U_A$  matrisinde depolanmaktadır.

$$U_A = \begin{bmatrix} \mu_{A,1,1} & \mu_{A,2,1} & \cdots & \mu_{A,13,1} \\ \cdots & \cdots & \cdots & \cdots \\ \mu_{A,1,250} & \mu_{A,2,250} & \cdots & \mu_{A,13,250} \end{bmatrix} \quad (6.5)$$

Bu modelde yer alan verilerin tek tek incelenmesinin ardından üyeliklerin modeldeki etkisini gözlemlenmek adına ilk 25 verinin oluşturduğu öncül veri seti ve bu verilere ait üyelik değerleri ile birlikte grafiksel gösterimi Şekil 6.5’te yer almaktadır. Bu şekilde görüleceği üzere çıktı değerlerinin mevcut yönelimi ile bazı öbekler bazı alt veri setleri özelinde benzer şekilde hareket etmiş; veri setinin yönelimi öbekler sayesinde tespit edilebilmiştir.

Örnek verecek olursak;

- 0-5 indisli verilerin yönelimi 12. Öbek'e aitlikleri ile paralellik göstermiş,
- 5-10 indisli verilerin yönelimi 12 ve 7.Öbeklerin üyelik değerlerinin yönelimi ile paralele hareket etmiş,
- 14-20 indisli verilerin yönelimi ise yine 7. Öbek ile benzerlik göstermiş,
- son olarak da 20-25 indisli verilerin yönelimi 4. Öbek'e aitlik verileri ile paralel hareket etmiştir.



Şekil 6.5. Model A'da belirtilen üyelik değerlerinin çıktı değeri ile dağılımı (ilk 25 veri)

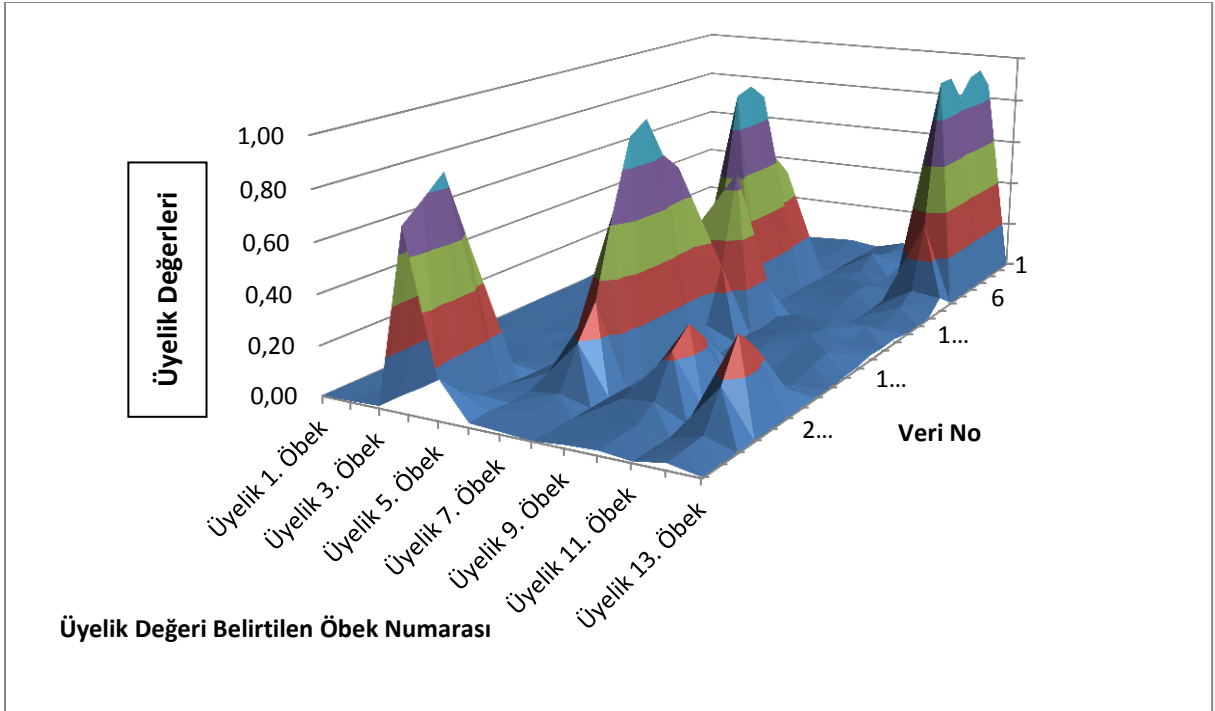
Model A'da verilen ilk 25 verinin her bir öbeğe üyelik değerlerinin gösterildiği Şekil 6.6'da; verilerin ağırlıklı olarak 5 farklı öbekte yüksek değer aldığı gözlemlenmektedir. Bu aslında örneklem olarak alınan 25 verinin karakteristik özelliklerinin farklı veriler özelinde farklı öbeklere benzeştiğini göstermekte olup Bulanık Regresyon Modelinin etki yöntemini özetlemektedir.

$X^A$  matrisi Bulanık Regresyon Modeli (BRM) kapsamında girdi olarak kabul edilecek A modeline göre hesaplanmış üyelik değerleri, bu üyeliklerin transformasyonları ve girdi verilerinin depolandığı genel girdi matrisidir.

$x_i$  : girdi verilerini temsil etmektedir.

$\mu^2_{A,1,n}$ : A Modeli kapsamında hesaplanan üyeliklerin transformasyonlarını temsil etmektedir.

$$X^A = \begin{bmatrix} 1 & \mu_{A,1,1} & \mu^2_{A,1,1} & x_{1x1} & \cdots & x_{1xnv} \\ 1 & \mu_{A,1,2} & \mu^2_{A,1,2} & x_{2x1} & \cdots & x_{2xnv} \\ \vdots & \vdots & \vdots & \vdots & \cdots & \cdots \\ 1 & \mu_{A,1,n} & \mu^2_{A,1,n} & x_{nx1} & \cdots & x_{nxnv} \end{bmatrix} \quad (6.6)$$



Şekil 6.6. Model A’da belirtilen üyelik değerlerinin gösterimi (İlk 25 Veri)

Verilerin öbeklere üyeliklerinin şiddetine göre verilerin öbeklere aitlikleri, şiddetli, orta derecede ve zayıf şeklinde sınıflandırılabilir [31]:

- **Şiddetli Üyelik:** Bir verinin bir öbeğe aitliği yani üyelik değeri 0,99 ile 0,65 arasında ise üyeliği şiddetli üyelik olarak nitelendirilir ve ilgili veri ağırlıklı olarak bu öbeğin karakteristik özelliklerini taşır denilebilmektedir.
- **Orta Dereceli Üyelik:** Bir verinin bir öbeğe aitliği yani üyelik değeri 0,65 ile 0,40 arasında ise üyeliği orta dereceli üyelik olarak nitelendirilir ve ilgili veri genel olarak bu öbeğin karakteristik özelliklerini taşır denilebilmektedir.

- **Zayıf Üyelik:** Bir verinin bir öbeğe aitliği yani üyelik değeri 0, 40 tan az ise üyeliği zayıf üyelik olarak nitelendirilir ve ilgili veri genel olarak bu öbeğin karakteristik özelliklerini taşımamaktadır sadece sınırlı etki söz konusudur denilebilmektedir.

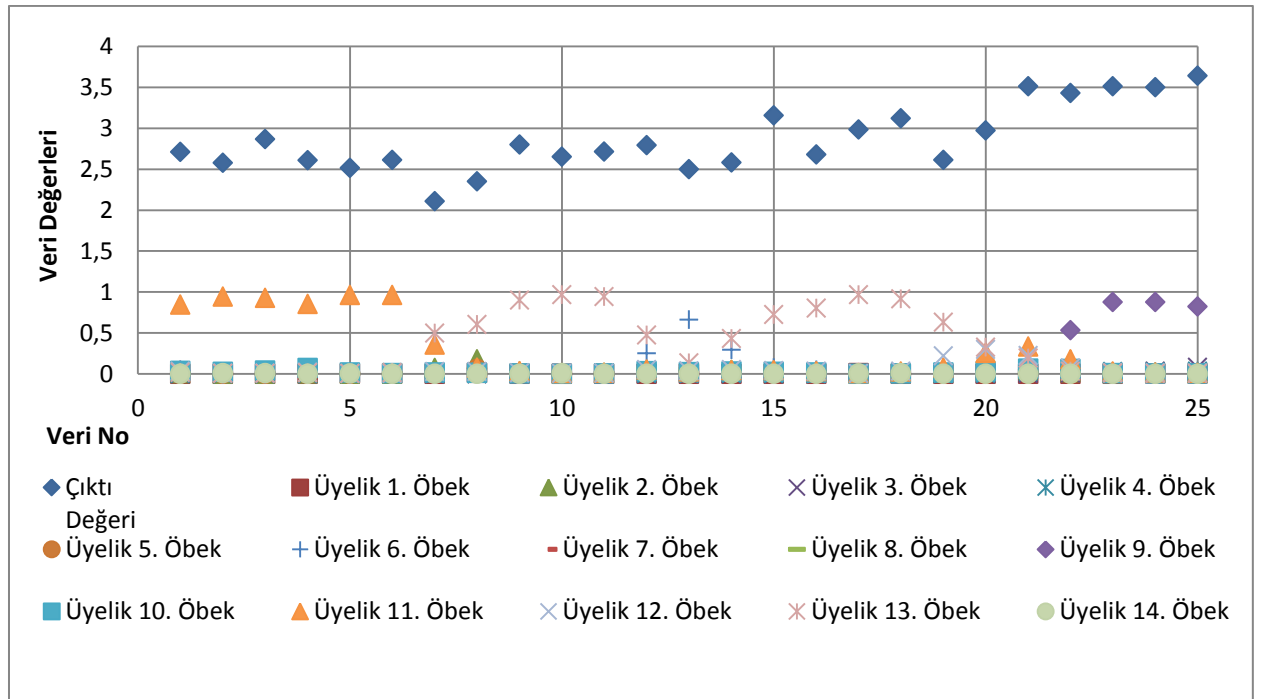
Bu sınıflandırma kapsamında Şekil 6.6'da beş şiddetli üyelik öbeği, bir orta dereceli üyelik öbeği, iki adet de zayıf üyelik öbeği tespit edilmiştir.

Bu deney kapsamında kullanılan diğer model ise öbek sayısı 14, bulanıklık mertebesi 1.4 olan, R-kare değeri %95.02 olan ve Model B olarak isimlendirilen modeldir. Özetle;

**B: c=14, m=1.4 (%95.02)** şeklinde ifade edilmektedir.

Bu model kapsamında elde edilen üyelik değerleri 6.7'de yer alan  $U_B$  matrisinde depolanmaktadır.

$$U_B = \begin{bmatrix} \mu_{B,1,1} & \mu_{B,2,1} & \cdots & \mu_{B,14,1} \\ \cdots & \cdots & \cdots & \cdots \\ \mu_{B,1,n} & \mu_{B,2,n} & \cdots & \mu_{B,14,n} \end{bmatrix} \quad (6.7)$$



Şekil 6.7. Model B'de belirtilen üyelik değerlerinin çıktı değeri ile dağılımı (İlk 25 Veri)

Model B'de hesaplanan üyeliklerin modeldeki etkisini gözlemlemek adına ilk 25 verinin oluşturduğu öncül veri seti ve bu verilere ait üyelik değerleri ile birlikte grafiksel gösterimi Şekil 6.7'de yer almaktadır. Bu şekilde yeniden görüleceği üzere çıktı değerlerinin mevcut

yönelimi ile bazı öbekler bazı alt veri setleri özelinde benzer şekilde hareket etmiş; veri setinin yönelimi öbekler sayesinde tespit edilebilmiştir.

Özetleyecek olursak;

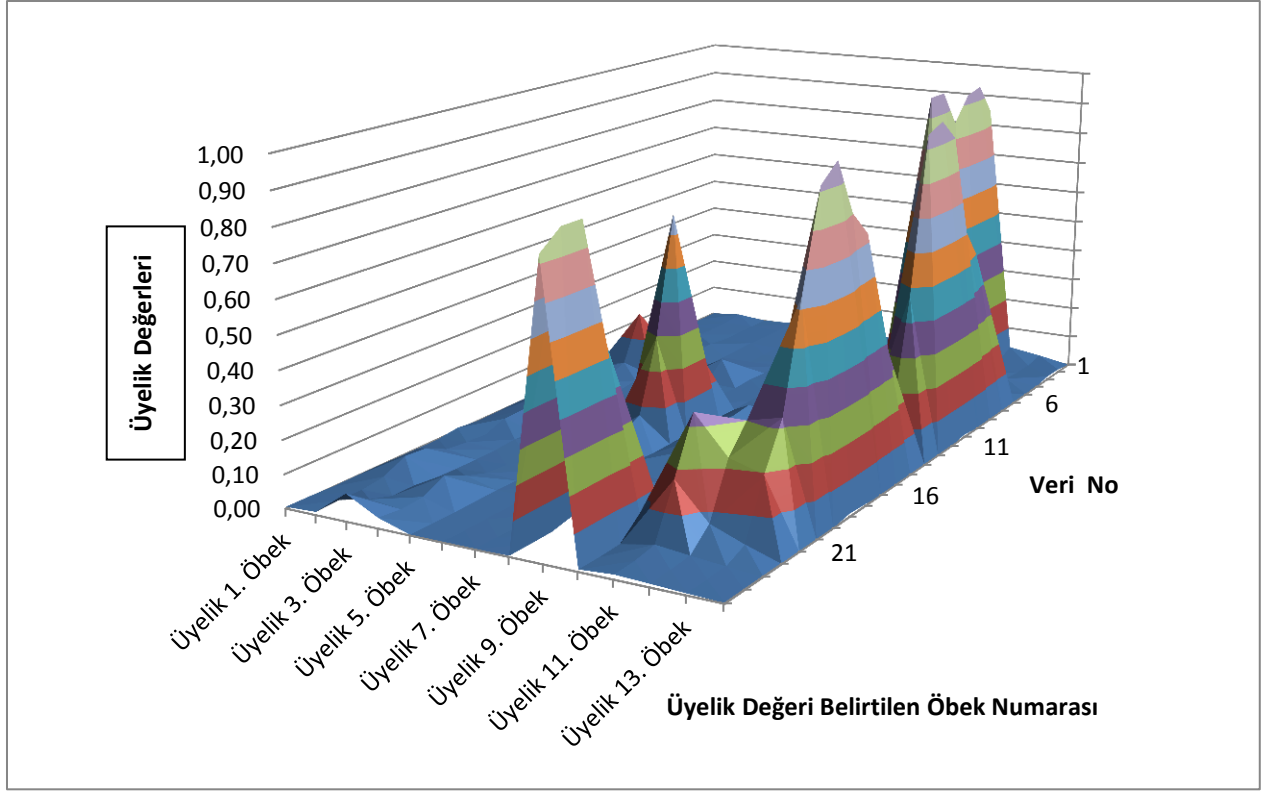
- 0-7 indisli verilerin yönelimi 11. Öbek'e aitlikleri ile paralellik göstermiş,
- 7-12 indisli verilerin yönelimi 13. Öbek'e aitlikleri ile paralellik göstermiş
- 15-20 indisli verilerin yönelimi ise yine 3. Öbek ile benzerlik göstermiş,
- Son olarak da 22-25 indisli verilerin yönelimi 9. Öbek'e aitlik verileri ile paralel hareket etmiştir.

$X^B$  matrisi Bulanık Regresyon Modeli (BRM) kapsamında girdi olarak kabul edilecek B modeline göre hesaplanmış üyelik değerleri, bu üyeliklerin transformasyonları ve girdi verilerinin depolandığı genel girdi matrisi olup 6.8'de gösterilmiştir.

$x_i$  : girdi verilerini temsil etmektedir.

$\mu^2_{B,1,n}$ : A Modeli kapsamında hesaplanan üyeliklerin transformasyonlarını temsil etmektedir.

$$X^B = \begin{bmatrix} 1 & \mu_{B,1,1} & \mu^2_{B,1,1} & x_{1x1} & \cdots & x_{1xnv} \\ 1 & \mu_{B,1,2} & \mu^2_{B,1,2} & x_{2x1} & \cdots & x_{2xnv} \\ \vdots & \vdots & \vdots & \vdots & \cdots & \cdots \\ 1 & \mu_{B,1,n} & \mu^2_{B,1,n} & x_{nx1} & \cdots & x_{nxnv} \end{bmatrix} \quad (6.8)$$



Şekil 6.8. Model B’de belirtilen üyelik değerlerinin gösterimi (İlk 25 Veri)

Model B’da verilen ilk 25 verinin her bir öbeğe üyelik değerlerinin gösterildiği Şekil 6.8’de; verilerin ağırlıklı olarak 5 farklı öbekte yüksek değer aldığı; daha detaylı incelersek beş şiddetli üyelik öbeği, bir orta dereceli üyelik öbeği, iki adet de zayıf üyelik öbeği tespit edilmiştir. Bu aslında örneklem olarak alınan 25 verinin karakteristik özelliklerinin farklı veriler özelinde farklı öbeklere benzeştiğini göstermekte olup Bulanık Regresyon Modelinin etki yöntemini özetlemektedir.

Bu iki model Çok Katmanlı Bulanık Regresyon Modeline girdi olarak sunulduğunda oluşan melez modelin tutarlılık parametresi R-kare değerinin %99.89 olduğu gözlemlenmiş olup bu melez modelin A ve B modellerinin tutarlılığından daha yüksek bir R-kare değerine ulaştığı gözlemlenmiştir. Özetle;

**A+B = %99.89** şeklinde ifade edilebilmektedir. İlgili ÇKBRM Modelinin girdi matrisi 6.9’da yer almaktadır.



$$X^{A+B} = \begin{bmatrix} 1 & \mu_{A,1,1} & \mu^2_{A,1,1} & \mu_{B,1,1} & \mu^2_{B,1,1} & x_{1x1} & \cdots & x_{1xnv} \\ 1 & \mu_{A,1,2} & \mu^2_{A,1,2} & \mu_{B,1,2} & \mu^2_{B,1,2} & x_{2x1} & \cdots & x_{2xnv} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & \mu_{A,1,n} & \mu^2_{A,1,n} & \mu_{B,1,n} & \mu^2_{B,1,n} & x_{nx1} & \cdots & x_{nxnv} \end{bmatrix} \quad (6.9)$$

Ayrıca bu tutarlılık düzeyine Bulanık Regresyon Analizi kapsamında oluşturulan 221 modelin hiç birinde ulaşamadığı; BRM kapsamında ulaşılabilen en yüksek R-kare değerinin %98,41 olduğu dikkat çekmektedir.

Çizelge 5.16. Deneme 1 model içerikleri ve model tutarlılıkları özet tablosu

Deney Numarası	Girdi Model Kodu	Model Tutarlılığı (R-kare)
Deney 1	<b>A: c=13, m=1.4</b>	<b>%94.29</b>
	<b>B: c=14, m=1.4</b>	<b>%95.02</b>
	<b>A+B</b>	<b>%99.89</b>

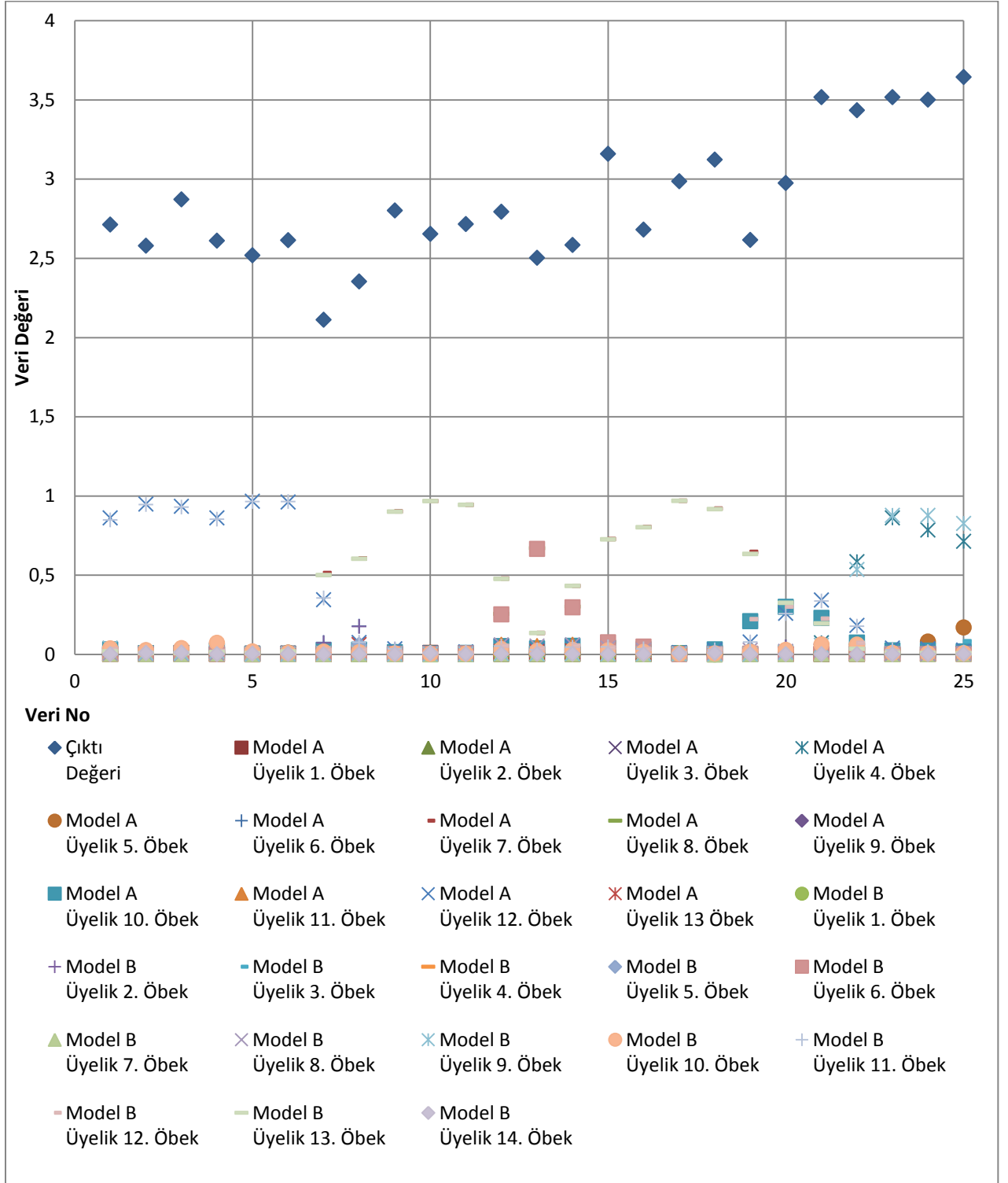
Bulanık Regresyon Analizine farklı bir boyut katan birden fazla öbek sayısı ve bulanıklık mertebesi için hesaplanan üyelik değerlerinin tek bir modelde kullanımı ve yaklaşımın modelin tutarlılık parametresi üzerine etkisi A ve B Modelleri temel alınarak oluşturulan Çok Katmanlı Bulanık Regresyon Modelinde artış göstermiştir. İlgili modellerin genel yapısı ve model tutarlılıkları Çizelge 5.16’da belirtilmiştir.

Model A+B’de hesaplanan üyeliklerin modeldeki etkisini gözlemlemek adına ilk 25 verinin oluşturduğu öncül veri seti ve bu verilere ait üyelik değerleri ile birlikte grafiksel gösterimi Şekil 6.9’da yer almaktadır. Bu şekilde görüldüğü üzere çıktı değerlerinin mevcut yönelimi ile bazı öbekler bazı alt veri setleri özelinde benzer şekilde hareket etmiş; veri setinin yönelimi öbekler sayesinde tespit edilebilmiştir.

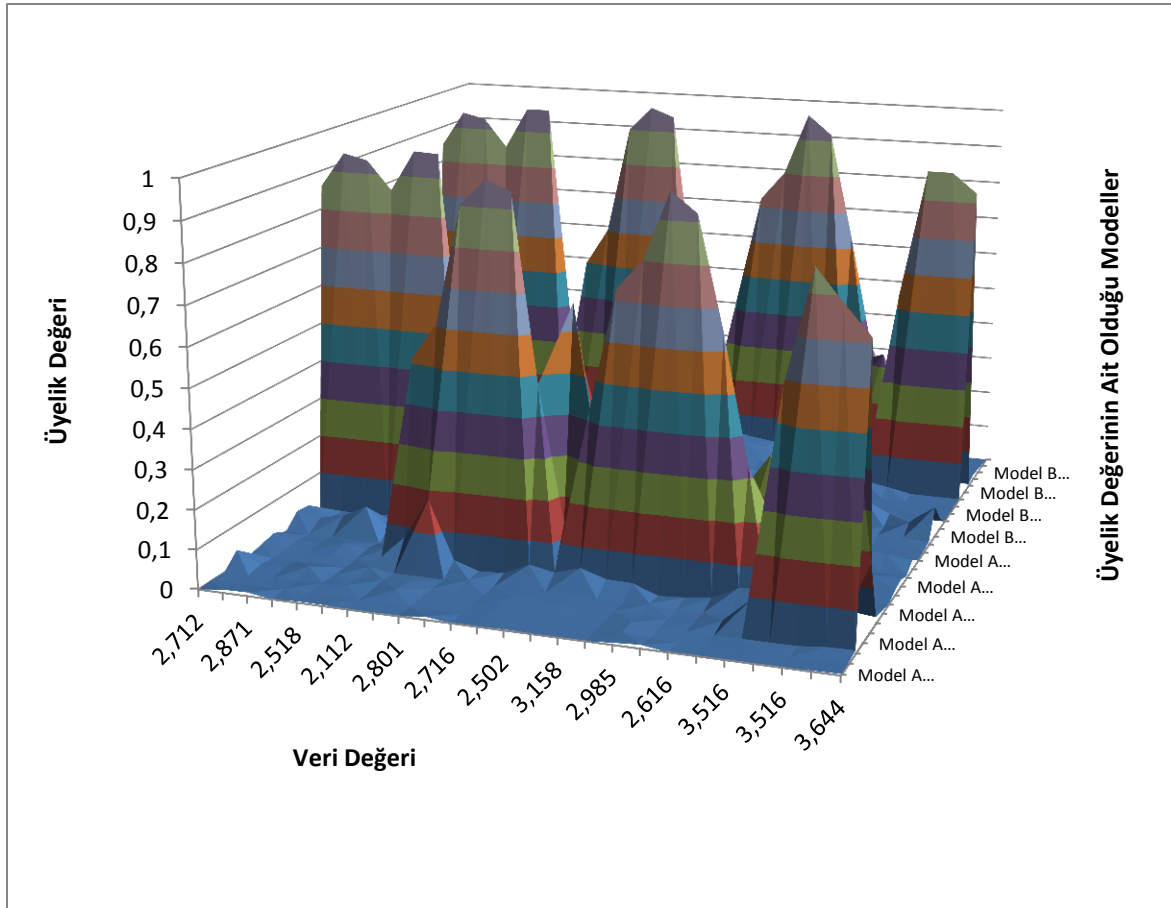
Özetleyecek olursak;

- 0-7 indisli verilerin yönelimi Model B-9.Öbek’e aitlikleri ile paralellik göstermiş,
- 8-19 indisli verilerin yönelimi Model A-8. Öbek’e aitlikleri ile paralellik göstermiş
- 19-25 indisli verilerin yönelimi ise eş zamanlı olarak Model A – 12. Model B- 8. Öbekler ile benzerlik göstermiştir.

Çok Katmanlı Bulanık Regresyon Modeli, hali hazırda kullanımda olan Bulanık Regresyon Analizi'ne ek olarak modelde yer alan verileri daha çok kapsadığı ve yönelimlerini daha çok yansıtan öbek ve üyelik değerlerine sahip olması sebebi ile oluşturulan melez modelin tutarlılığının arttığı gözlemlenmiştir.



Şekil 6.9 Model A ve B’de belirtilen üyelik değerleri ve çıktı değerlerinin gösterimi (İlk 25 Veri)

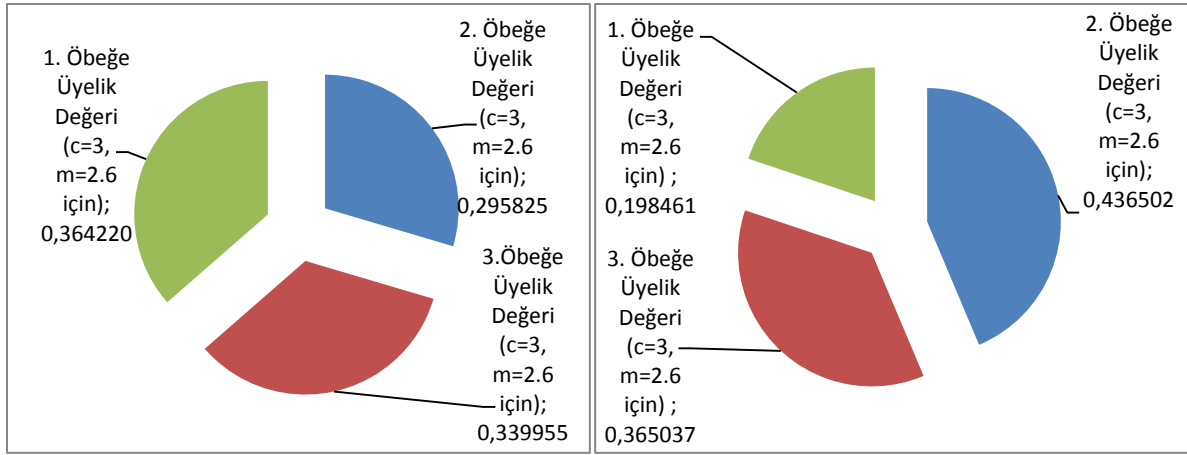


Şekil 6.10. Model A+B’de belirtilen üyelik değerlerinin gösterimi (İlk 25 Veri)

Model A+B’de verilen ilk 25 verinin her bir öbeğe üyelik değerlerinin gösterildiği Şekil 6.9 ve Şekil 6.10’da verilerin ağırlıklı olarak 10 farklı öbekte yüksek değer aldığı gözlemlenmektedir. Bu aslında örneklem olarak alınan 25 verinin karakteristik özelliklerinin farklı veriler özelinde farklı öbeklere benzeştiğini göstermekte olup Bulanık Regresyon Modelinin etki yönteminin Çok Katmanlı Bulanık Regresyon Analizinde aynı Bulanık Regresyon Modelinin doğrusal regresyon modelini güçlendirdiği şekilde Bulanık Regresyon Modelini güçlendirdiğini ve modelin tutarlılığını özetlediğini göstermektedir.

### **Deneme 2:**

Bu deneme kapsamında ele alınan ilk model öbek sayısı 3, bulanıklık mertebesi 2.6 olan, R-kare değeri %66.29 olan ve Model D olarak isimlendirilen modeldir.

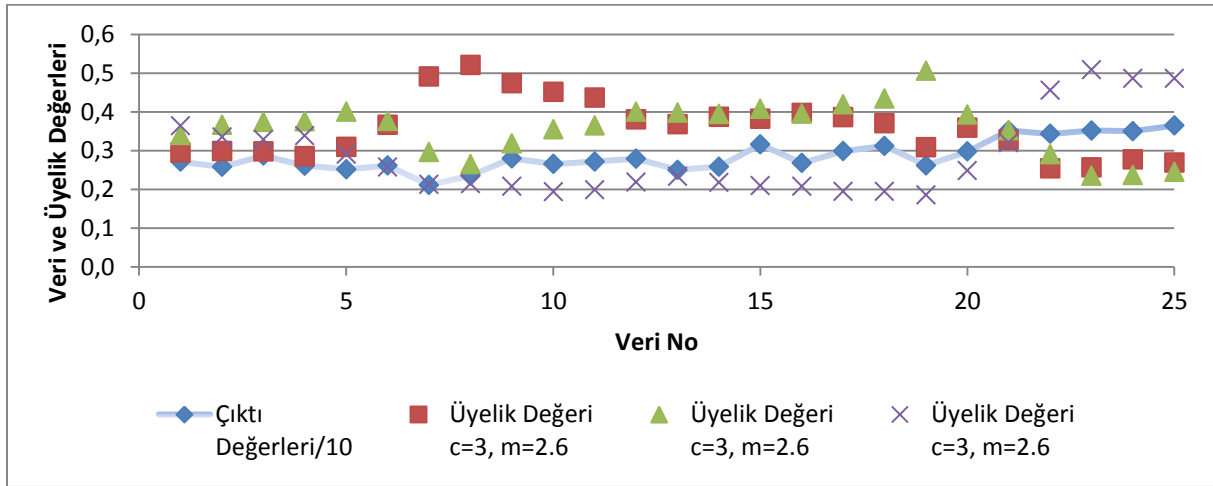


Şekil 6.11. Model D’de yer alan iki farklı verinin öbeğe aitliğinin gösterimi

Bu verinin modellenmesi ve karakteristik özelliklerinin tespitinde üyelik değerlerinin etkisi bulunmaktadır, sadece en yüksek derece ile ait oldukları öbeğin değil daha az etki ile üye olunana öbeğin de verinin yönelimini belirlemede etkisi aşikârdır. Ayrıca bu değerlendirme her bir verinin Şekil 6.11’de belirtildiği üzere farklı üyelik değerlerine sahip olması hasebiyle her bir veri için farklı şekilde geçerlidir.

Özetle;

**D:  $m=2.6$ ,  $c=3$  (%66.29)** şeklinde ifade edilmektedir.



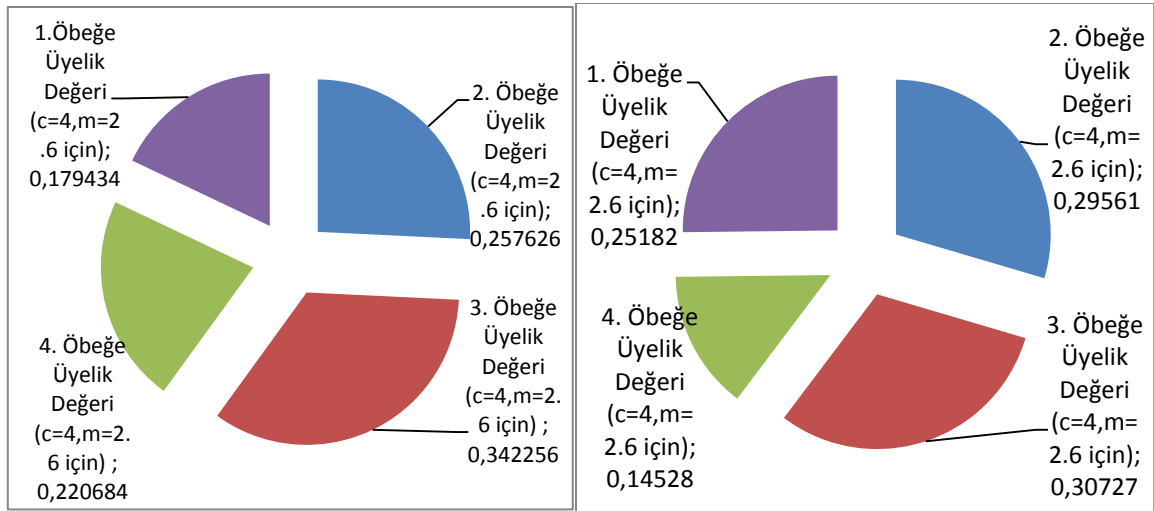
Şekil 6.12. Model D’de belirtilen üyelik değerleri ile çıktı değerinin dağılımı (İlk 25 Veri)

Model D kapsamında hesaplanan üyelik değerlerinden ilk 25’i ve tahmin edilmek üzere model oluşturulmaya çalışılan çıktı değeri yönelimlerini gözlemlemek amacı ile Şekil 6.12 oluşturulmuştur. Şekil 6.12’de görülebileceği ve bir önceki deneyde açıkça ifade edildiği

üzere 25 çıktı değerinin yönelimleri farklı öbeklerce taklit edilmekte olduğundan modelde üyelik değerlerinin varlığı modelin tutarlılığını artırmaktadır.

Bu deney kapsamında kullanılan diğer model ise öbek sayısı 4, bulanıklık mertebesi 2.6 olan, R-kare değeri % 66.97 olan ve Model E olarak isimlendirilen modeldir.

Model D’de hesaplanan üyeliklerin modeldeki etkisini gözlemlemek adına ilk 25 verinin oluşturduğu öncül veri seti ve be bu verilere ait üyelik değerleri ile birlikte grafiksel gösterimi Şekil 6.12’de yer almaktadır. Bu şekilde yeniden görüleceği üzere çıktı değerlerinin mevcut yönelimi ile bazı öbekler bazı alt veri setleri özelinde benzer şekilde hareket etmiş; veri setinin yönelimi öbekler sayesinde tespit edilebilmiştir.

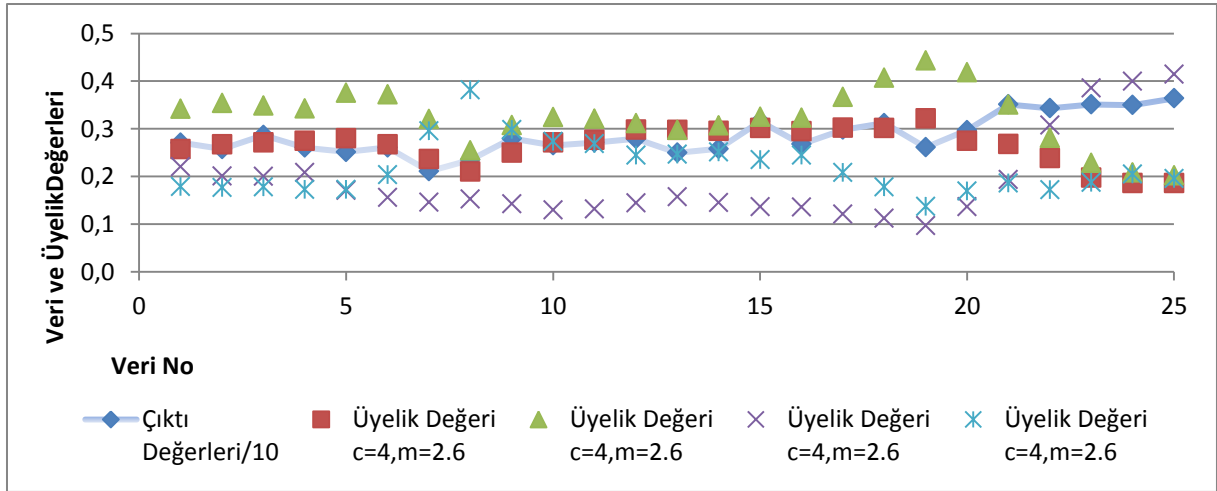


Şekil 6.13. Model E’de yer alan iki farklı verinin öbeklere aitliğinin gösterimi

Şekil 6.13’te ise Model E kapsamında yer alan iki farklı verinin üyelik değerleri ve birbirlerinden farklı oluşları gösterilmiştir. Bu veri kümesinin modellenmesi ve karakteristik özelliklerinin tespitinde üyelik değerlerinin etkisi bulunduğu aşikâr olup sadece en yüksek derece ile ait oldukları öbeğin değil daha az etki ile üye olunana öbeğin de etkisi söz konusudur.

Özetle; **E: m=2.6, c= 4 (%66.97)** şeklinde ifade edilmektedir.

Model E kapsamında hesaplanan üyelik değerlerinden ilk 25’i ve tahmin edilmek üzere model oluşturulmaya çalışılan çıktı değeri yönelimlerini gözlemlemek amacı ile Şekil 6.12 oluşturulmuştur. Şekil 6.12’de görülebileceği ve bir önceki deneyde açıkça ifade edildiği üzere 25 çıktı değerinin yönelimleri farklı öbeklerce taklit edilmekte olduğundan modelde üyelik değerlerinin varlığı modelin tutarlılığını artırmaktadır.



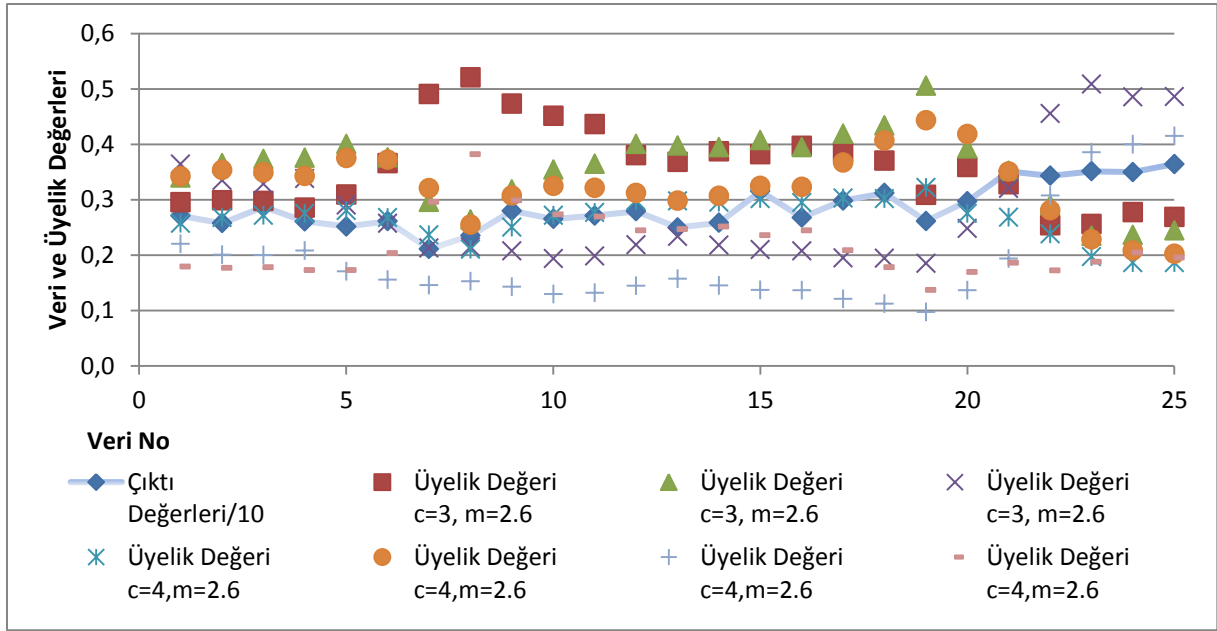
Şekil 6.14. Model E’de belirtilen üyelik değerleri ile çıktı değerinin dağılımı (İlk 25 Veri)

Model E’de hesaplanan üyeliklerin modeldeki etkisini gözlemlemek adına ilk 25 verinin oluşturduğu öncül veri seti ve be bu verilere ait üyelik değerleri ile birlikte grafiksel gösterimi Şekil 6.14’te yer almaktadır. Bu şekilde yeniden görüleceği üzere çıktı değerlerinin mevcut yönelimi ile bazı öbekler bazı alt veri setleri özelinde benzer şekilde hareket etmiş; veri setinin yönelimi öbekler sayesinde tespit edilebilmiştir.

Bu iki model Çok Katmanlı Bulanık Regresyon Modeline girdi olarak sunulduğunda oluşan melez modelin tutarlılık parametresi R-kare değerinin %81.25 olduğu gözlemlenmiş olup bu melez modelin D ve E modellerinin tutarlılığından daha yüksek bir R-kare değerine ulaştığı gözlemlenmiştir. Özetle;

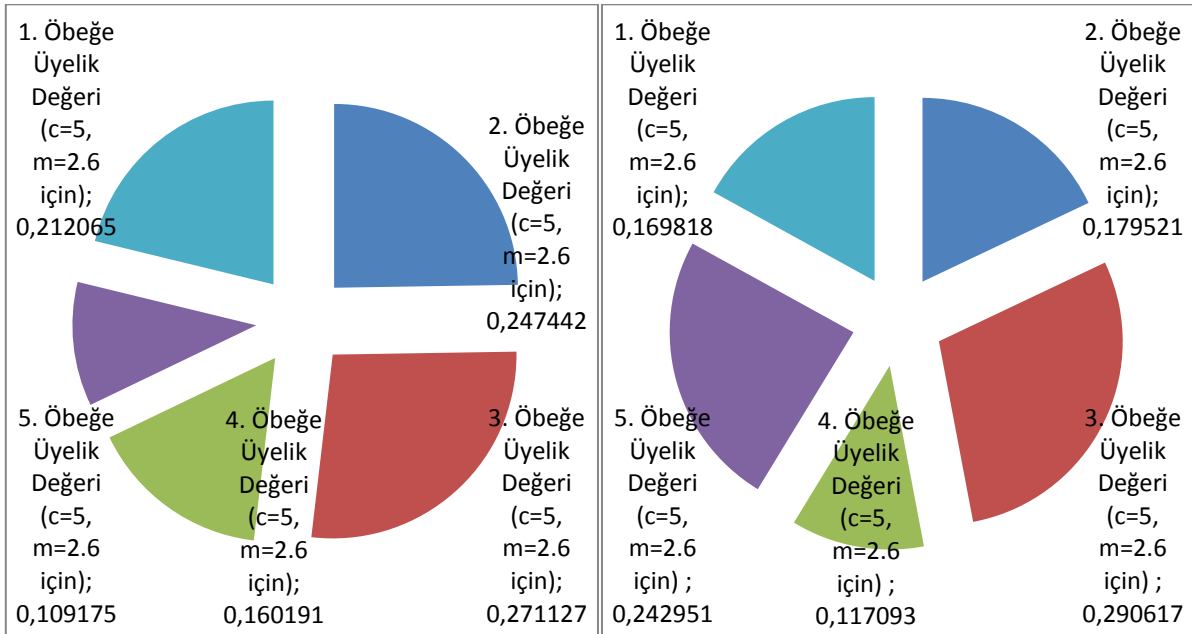
**D+E= %81.25** şeklinde ifade edilebilmektedir.

Model D+E’de hesaplanan üyeliklerin modeldeki etkisini gözlemlemek adına ilk 25 verinin oluşturduğu öncül veri seti ve be bu verilere ait üyelik değerleri ile birlikte grafiksel gösterimi Şekil 6.15’te yer almaktadır.



Şekil 6.15. Model D+E’de belirtilen üyelik değerlerinin gösterimi (İlk 25 Veri)

Bu deneyde bir adım daha atılarak Model F tanımlanmıştır, öbek sayısı 5, bulanıklık mertebesi 2.6 olan, R-kare değeri %66.97 olan ve Model E olarak isimlendirilen modeldir.

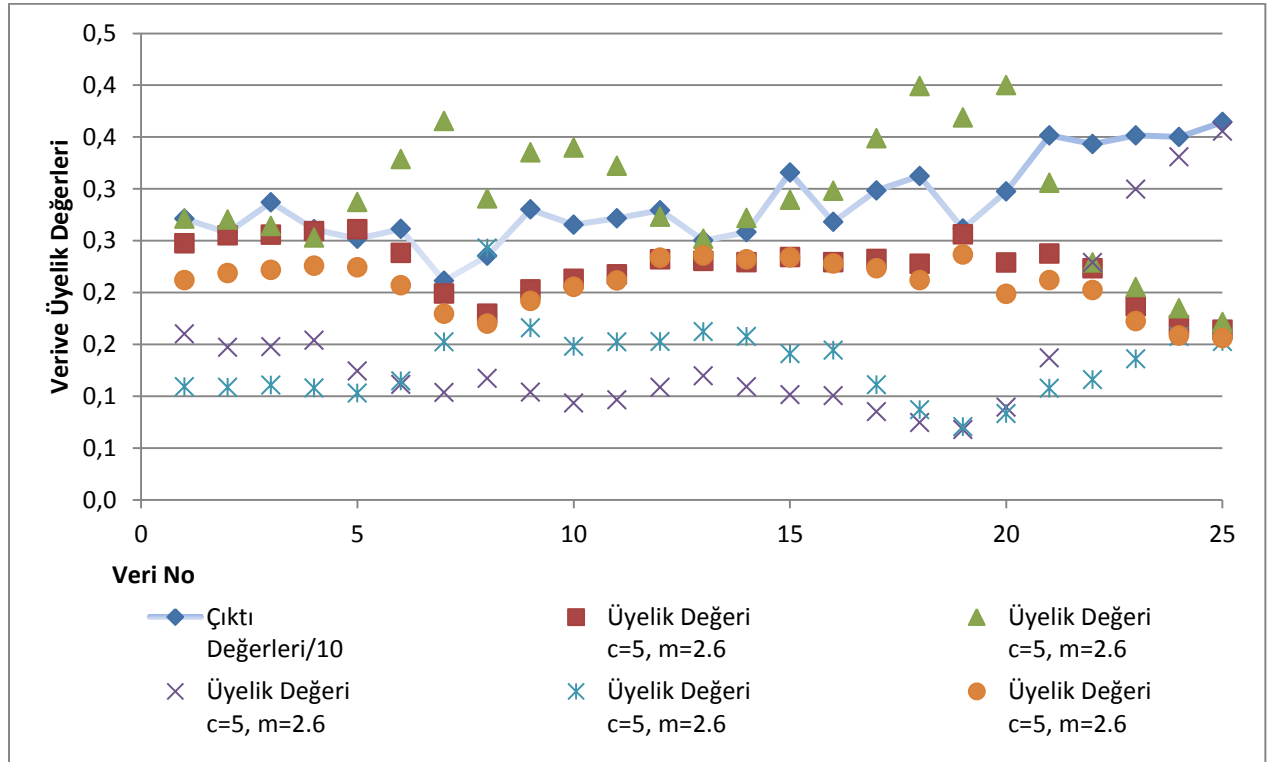


Şekil 6.16. Model F’de yer alan iki farklı verinin öbeklere aitliğinin gösterimi

Şekil 6.16’da Model F’de yer alan iki farklı verinin üyelik değerleri ve birbirlerinden farklı oluşları gösterilmiş olup, bu veri kümesinin modellenmesi ve karakteristik özelliklerinin tespitinde üyelik değerlerinin etkisi bulunduğu teyidi yapılmıştır.

Özetle;

**F: m=2.6, c=5 (%66.29)** şeklinde ifade edilmektedir.



Şekil 6.17. Model F’de belirtilen üyelik değerleri ile çıktı değerinin dağılımı (İlk 25 Veri)

Model F’de hesaplanan üyeliklerin modeldeki etkisini gözlemlemek adına ilk 25 verinin oluşturduğu öncül veri seti ve bu verilere ait üyelik değerleri ile birlikte grafiksel gösterimi Şekil 6.17’de yer almaktadır. Bu şekilde yeniden görüleceği üzere çıktı değerlerinin mevcut yönelimi ile bazı öbekler bazı alt veri setleri özelinde benzer şekilde hareket etmiş; veri setinin yönelimi öbekler sayesinde tespit edilebilmiştir.

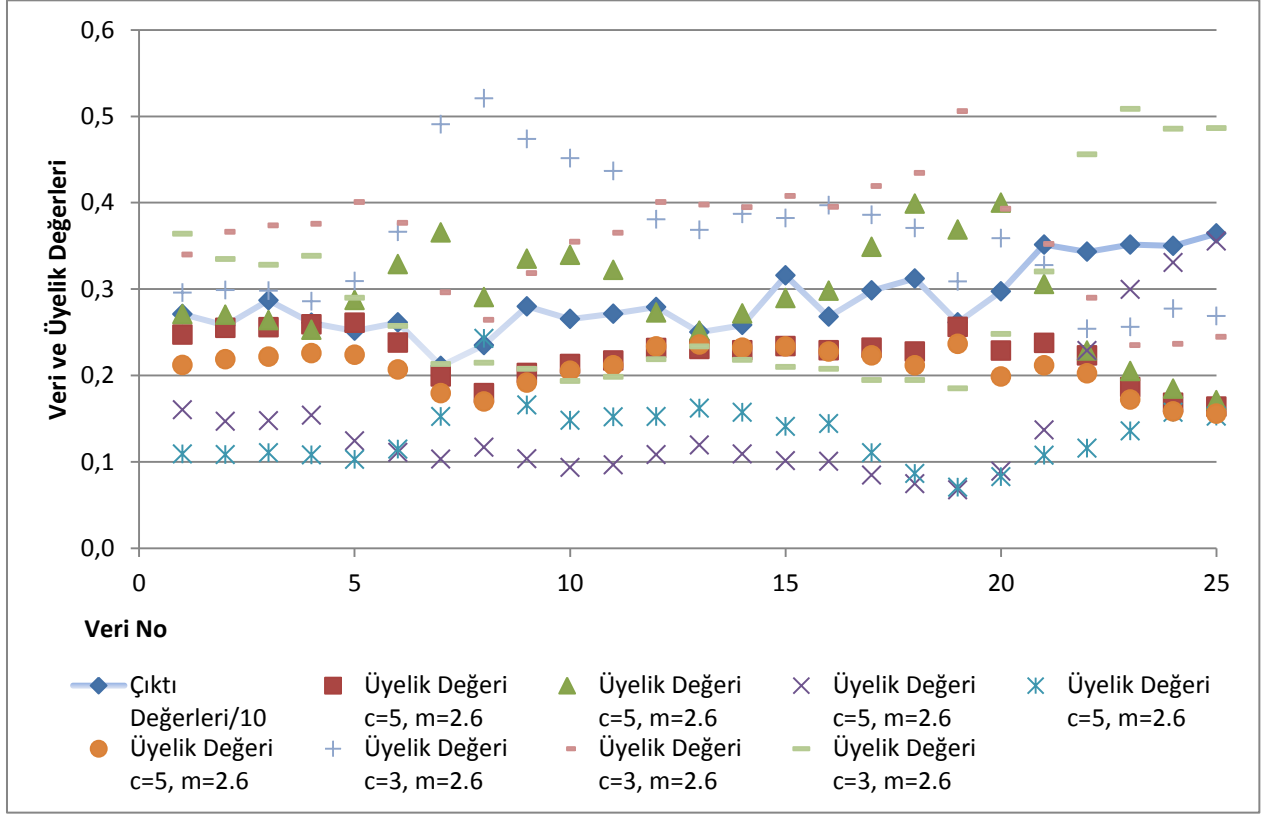
D ve F modellerinin Çok Katmanlı Bulanık Regresyon Modeline girdi olarak sunulduğunda oluşan melez modelin tutarlılık parametresi R-kare değerinin %81.95 olduğu gözlemlenmiş olup bu melez modelin D ve F modellerinin tutarlılığından daha yüksek bir R-kare değerine ulaştığı gözlemlenmiştir.

Özetle;

**D+F= %81.95**



Model D+F’de hesaplanan üyeliklerin modeldeki etkisini gözlemlemek adına ilk 25 verinin oluşturduğu öncül veri seti ve bu verilere ait üyelik değerleri ile birlikte grafiksel gösterimi Şekil 6.18’de yer almaktadır.



Şekil 6.18. Model D+F’de belirtilen üyelik değerlerinin gösterimi (İlk 25 Veri)

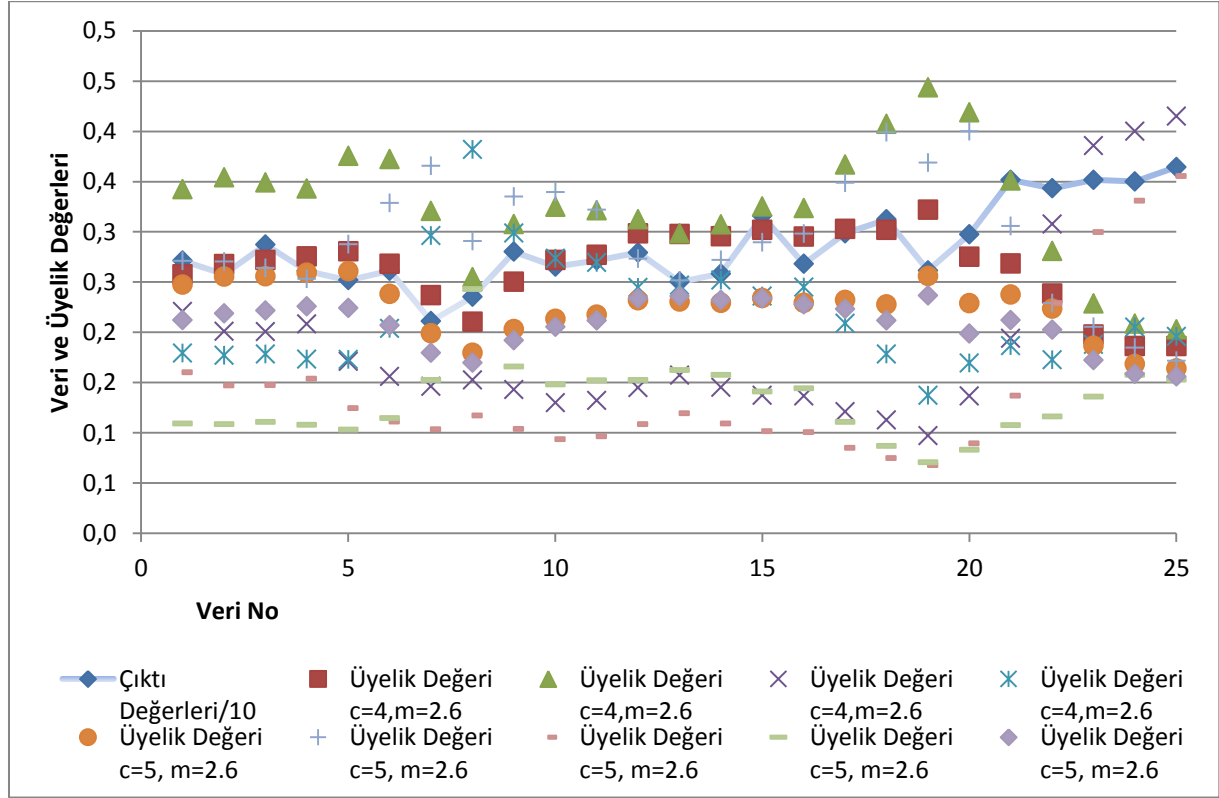
Aynı şekilde E ve F modellerinin Çok Katmanlı Bulanık Regresyon Modeline girdi olarak sunulduğunda oluşan melez modelin tutarlılık parametresi R-kare değerinin %80.48 olduğu gözlemlenmiş olup bu melez modelin D ve F modellerinin BRM Modeli tutarlılığından daha yüksek bir R-kare değerine ulaştığı gözlemlenmiştir.

Özetle;

**E+F=% 80.48**

Model E+F’de hesaplanan üyeliklerin modeldeki etkisini gözlemlemek adına ilk 25 verinin oluşturduğu öncül veri seti ve bu verilere ait üyelik değerleri ile birlikte grafiksel gösterimi Şekil 6.18’de yer almaktadır.

Model E+F’de hesaplanan üyeliklerin modeldeki etkisini gözlemlemek adına ilk 25 verinin oluşturduğu öncül veri seti ve be bu verilere ait üyelik değerleri ile birlikte grafiksel gösterimi Şekil 6.19’da yer almaktadır.



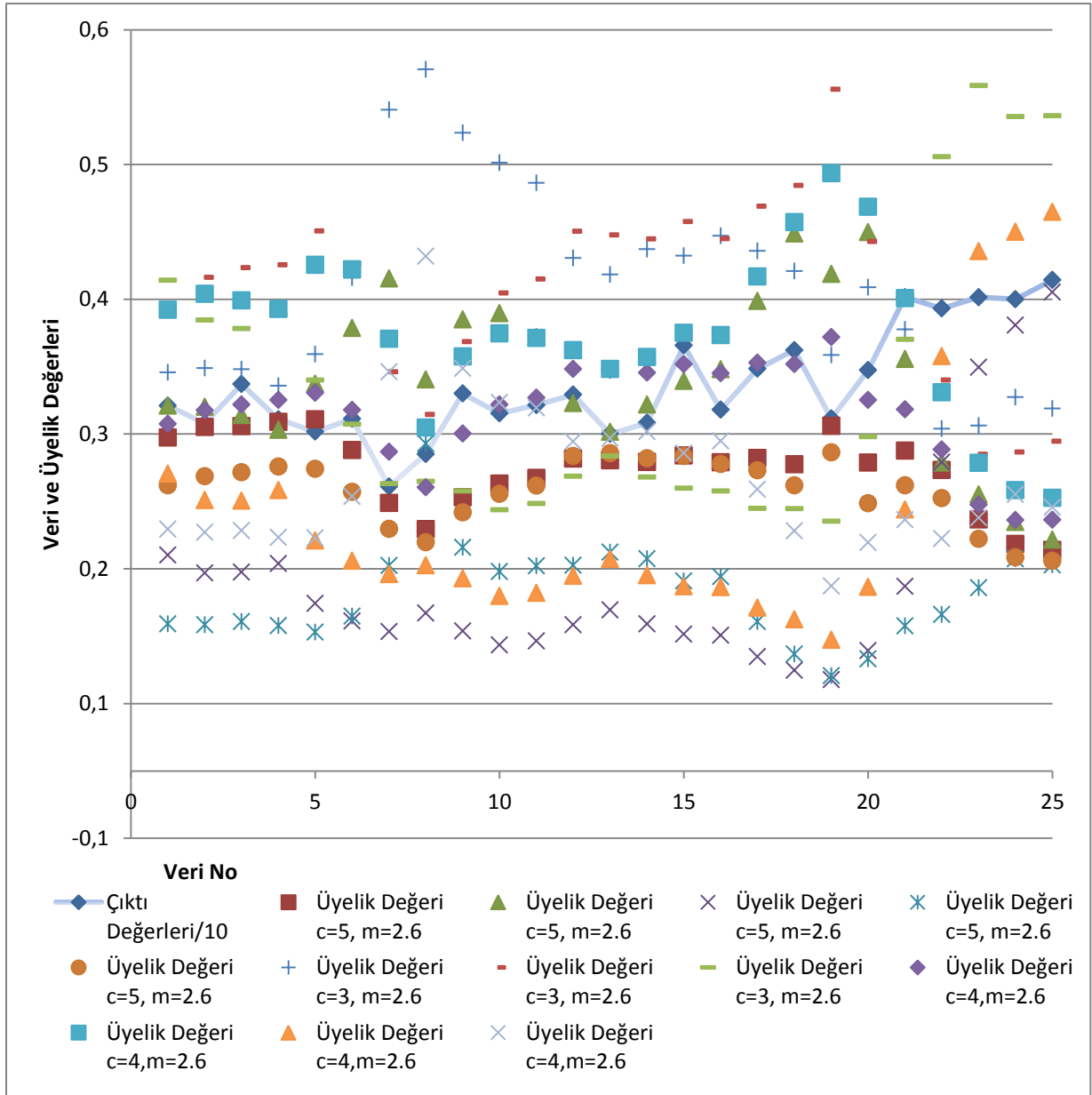
Şekil 6.19. Model E+F,’de belirtilen üyelik değerlerinin gösterimi (İlk 25 Veri)

Son olarak D, E ve F modellerinin Çok Katmanlı Bulanık Regresyon Modeline girdi olarak sunulduğunda oluşan melez modelin tutarlılık parametresi R-kare değerinin %73.79 olduğu gözlemlenmiş olup bu melez modelin D, E ve F modellerinin tutarlılığından daha yüksek bir R-kare değerine ulaştığı gözlemlenmiştir.

Özetle;

**D+E+F = %73.79** şeklinde ifade edilmektedir.

Son olarak Model D+E+F’de hesaplanan üyeliklerin modeldeki etkisini gözlemlemek adına ilk 25 verinin oluşturduğu öncül veri seti ve be bu verilere ait üyelik değerleri ile birlikte grafiksel gösterimi Şekil 6.20’de yer almaktadır.



Şekil 6.20. Model D+E+F’de belirtilen üyelik değerlerinin gösterimi (İlk 25 Veri)

Çok katmanlı üyelik değerlerinin modelde kullanımı çıktı değerini ve bu değer zamana bağlı eğilimlerini BRM’ye kıyasla daha çok sayıda temsil ve taklit edebildiğinden modeldeki varlıklarının modeli iyileştirdiği gözlemlenmiştir.

Çizelge 5.17. Deneme 2 model içerikleri ve model tutarlılıkları özet tablosu

Deney Numarası	Girdi Model Kodu	Model Tutarlılığı (R-kare)
Deney 2	<b>D: m=2.6, c=3</b>	<b>%66.29</b>
	<b>E: m=2.6, c=4</b>	<b>%66.97</b>
	<b>F: m=2.6, c=5</b>	<b>%66.29</b>
	<b>D+E</b>	<b>%81.25</b>
	<b>D+F</b>	<b>%81.95</b>
	<b>E+F</b>	<b>%80.48</b>
	<b>D+E+F</b>	<b>%73.79</b>

Bu deneyde de görüldüğü üzere, Çok Katmanlı Bulanık Regresyon Modeli, Bulanık Regresyon Modelinde oluşturulan modellere kıyasla R-kare değeri daha yüksek melez modellerin oluşturulmasına olanak sağlamaktadır.

### **Deneme 3:**

Bu deneme kapsamında ele alınan ilk model öbek sayısı 7, bulanıklık mertebesi 1.7 olan, R-kare değeri %75,78 olan ve Model G olarak isimlendirilen modeldir. Özetle;

**G: m=1.7, c= 7 (%75,78)** şeklinde ifade edilmektedir.

Bu deney kapsamında kullanılan diğer model ise öbek sayısı 8, bulanıklık mertebesi 1.7 olan, R-kare değeri %80,54 olan ve Model H olarak isimlendirilen modeldir. Özetle;

**H: m=1.7, c= 8 (%80,54)** şeklinde ifade edilmektedir.

Bu iki model Çok Katmanlı Bulanık Regresyon Modeline girdi olarak sunulduğunda oluşan melez modelin tutarlılık parametresi R-kare değerinin %88.69 olduğu gözlemlenmiş olup bu melez modelin G ve H modellerinin tutarlılığından daha yüksek bir R-kare değerine ulaştığı gözlemlenmiştir. Özetle;

**G+H= %88,69** şeklinde ifade edilmektedir.

Bu deneyde fazladan bir adım daha atılarak Model I tanımlanmıştır, öbek sayısı 9, bulanıklık mertebesi 1.7 olan, R-kare değeri %82,89 olan ve Model I olarak isimlendirilen modeldir. Özetle;

**I: m=1.7, c= 9 (%82,89)** şeklinde ifade edilmektedir.

Aynı şekilde G ve I modellerinin Çok Katmanlı Bulanık Regresyon Modeline girdi olarak sunulduğunda oluşan melez modelin tutarlılık parametresi R-kare değerinin %88.54 olduğu gözlemlenmiş olup bu melez modelin G ve I modellerinin tutarlılığından daha yüksek bir R-kare değerine ulaştığı gözlemlenmiştir. Özetle;

**G+I= %88,54** şeklinde ifade edilmektedir.

Yine benzer şekilde H ve I modellerinin Çok Katmanlı Bulanık Regresyon Modeline girdi olarak sunulduğunda oluşan melez modelin tutarlılık parametresi R-kare değerinin %99,21 olduğu gözlemlenmiş olup bu melez modelin H ve I modellerinin tutarlılığından daha yüksek bir R-kare değerine ulaştığı gözlemlenmiştir. Özetle;

**H+I= %92,21** şeklinde ifade edilmektedir.

Son olarak G, H ve I modellerinin Çok Katmanlı Bulanık Regresyon Modeline girdi olarak sunulduğunda oluşan melez modelin tutarlılık parametresi R-kare değerinin %94.63 olduğu gözlemlenmiş olup bu melez modelin G, H ve I modellerinin tutarlılığından daha yüksek bir R-kare değerine ulaştığı gözlemlenmiştir. Özetle;

**G+H+I= %94,63** şeklinde ifade edilmektedir.

Çizelge 5.18. Deneme 3 model içerikleri ve model tutarlılıkları özet tablosu

Deney Numarası	Girdi Model Kodu	Model Tutarlılığı (R-kare)
Deney 3	<b>G: m=1.7, c= 7</b>	<b>%75,78</b>
	<b>H: m=1.7, c= 8</b>	<b>%80,54</b>
	<b>I: m=1.7, c= 9</b>	<b>%82,89</b>
	<b>G+I</b>	<b>%88,54</b>
	<b>H+I</b>	<b>%92,21</b>
	<b>G+H</b>	<b>%88,69</b>
	<b>G+H+I</b>	<b>%94,63</b>

Bu deneyde de görüldüğü üzere, Çok Katmanlı Bulanık Regresyon Modeli, Bulanık Regresyon Modelinde oluşturulan modellere kıyasla R-kare değeri daha yüksek melez modellerin oluşturulmasına olanak sağlamaktadır.

#### **Deneme 4:**

Bu deneme kapsamında ele alınan ilk model öbek sayısı 15, bulanıklık mertebesi 1.5 olan, R-kare değeri %87,62 olan ve Model J olarak isimlendirilen modeldir. Özetle;

**J: m=1.5, c= 15 (%87,62)** şeklinde ifade edilmektedir.

Bu deney kapsamında kullanılan diğer model ise öbek sayısı 16, bulanıklık mertebesi 1.5 olan, R-kare değeri %89,62 olan ve Model K olarak isimlendirilen modeldir. Özetle;

**K: m=1.5, c= 16 (%89,62)** şeklinde ifade edilmektedir.

Bu iki model Çok Katmanlı Bulanık Regresyon Modeline girdi olarak sunulduğunda oluşan melez modelin tutarlılık parametresi R-kare değerinin %98,77 olduğu gözlemlenmiş olup bu melez modelin G ve H modellerinin tutarlılığından daha yüksek bir R-kare değerine ulaştığı gözlemlenmiştir. Özetle;

**J+K= %98,77** şeklinde ifade edilmektedir.

Bu deney kapsamında kullanılan diğer model ise öbek sayısı 16, bulanıklık mertebesi 1.5 olan, R-kare değeri %89,62 olan ve Model L olarak isimlendirilen modeldir. Özetle;

**L: m=1.5,c=17 (%91.33)** şeklinde ifade edilmektedir.

Aynı şekilde J ve L modellerinin Çok Katmanlı Bulanık Regresyon Modeline girdi olarak sunulduğunda oluşan melez modelin tutarlılık parametresi R-kare değerinin %98,95 olduğu gözlemlenmiş olup bu melez modelin G ve I modellerinin tutarlılığından daha yüksek bir R-kare değerine ulaştığı gözlemlenmiştir. Özetle;

**J+L= %98,95** şeklinde ifade edilmektedir.

Yine benzer şekilde K ve L modellerinin Çok Katmanlı Bulanık Regresyon Modeline girdi olarak sunulduğunda oluşan melez modelin tutarlılık parametresi R-kare değerinin %99,37 olduğu gözlemlenmiş olup bu melez modelin K ve L modellerinin tutarlılığından daha yüksek bir R-kare değerine ulaştığı gözlemlenmiştir. Özetle;

**K+L=%99,37** şeklinde ifade edilmektedir.

Son olarak J, K ve L modellerinin Çok Katmanlı Bulanık Regresyon Modeline girdi olarak sunulduğunda oluşan melez modelin tutarlılık parametresi R-kare değerinin %99,87 olduğu

gözlemlenmiş olup bu melez modelin G, H ve I modellerinin tutarlılığından daha yüksek bir R-kare değerine ulaştığı gözlemlenmiştir. Özetle;

**J+K+L=%99,87** şeklinde ifade edilmektedir.

Çizelge 5.19. Deneme 4 model içerikleri ve model tutarlılıkları özet tablosu

Deney Numarası	Girdi Model Kodu	Model Tutarlılığı (R-kare)
Deney 4	<b>J: m=1.5, c= 15</b>	<b>%87,62</b>
	<b>K: m=1.5, c= 16</b>	<b>%89,62</b>
	<b>L: m=1.5,c=17</b>	<b>%91.33</b>
	<b>J+K</b>	<b>%98,77</b>
	<b>J+L</b>	<b>%98,95</b>
	<b>K+L</b>	<b>%99,37</b>
	<b>J+K+L</b>	<b>%99,87</b>

Bu deneyde de görüldüğü üzere, Çok Katmanlı Bulanık Regresyon Modeli, Bulanık Regresyon Modelinde oluşturulan modellere kıyasla R-kare değeri daha yüksek melez modellerin oluşturulmasına olanak sağlamaktadır.

#### **Deneme 5:**

Bu deneme kapsamında ele alınan ilk model öbek sayısı 16, bulanıklık mertebesi 1.9 olan, R-kare değeri %91,47 olan ve Model M olarak isimlendirilen modeldir. Özetle;

**M: m=1.9, c= 16 (%91,47)** şeklinde ifade edilmektedir.

Bu deney kapsamında kullanılan diğer model ise öbek sayısı 17, bulanıklık mertebesi 1.9 olan, R-kare değeri (%93,66) olan ve Model N olarak isimlendirilen modeldir. Özetle;

**N: m=1.9, c= 17 (%93,66)** şeklinde ifade edilmektedir.

Bu iki model Çok Katmanlı Bulanık Regresyon Modeline girdi olarak sunulduğunda oluşan melez modelin tutarlılık parametresi R-kare değerinin %98,41 olduğu gözlemlenmiş olup bu melez modelin M ve N modellerinin tutarlılığından daha yüksek bir R-kare değerine ulaştığı gözlemlenmiştir. Özetle;

**M+N= %98,41** şeklinde ifade edilmektedir.

Bu deney kapsamında kullanılan diğer model ise öbek sayısı 18, bulanıklık mertebesi 1.9 olan, R-kare değeri %92.14 olan ve Model O olarak isimlendirilen modeldir. Özetle;

**O: m=1.9, c= 18 (%92.14)** şeklinde ifade edilmektedir.

Aynı şekilde M ve O modellerinin Çok Katmanlı Bulanık Regresyon Modeline girdi olarak sunulduğunda oluşan melez modelin tutarlılık parametresi R-kare değerinin %98,02 olduğu gözlemlenmiş olup bu melez modelin M ve O modellerinin tutarlılığından daha yüksek bir R-kare değerine ulaştığı gözlemlenmiştir. Özetle;

**M+O= %98,02** şeklinde ifade edilmektedir.

Yine benzer şekilde N ve O modellerinin Çok Katmanlı Bulanık Regresyon Modeline girdi olarak sunulduğunda oluşan melez modelin tutarlılık parametresi R-kare değerinin %98,60 olduğu gözlemlenmiş olup bu melez modelin N ve O modellerinin tutarlılığından daha yüksek bir R-kare değerine ulaştığı gözlemlenmiştir. Özetle;

**N+O= %98,60** şeklinde ifade edilmektedir.

Son olarak M, N ve O modellerinin Çok Katmanlı Bulanık Regresyon Modeline girdi olarak sunulduğunda oluşan melez modelin tutarlılık parametresi R-kare değerinin %98,88 olduğu gözlemlenmiş olup bu melez modelin M, N ve O modellerinin tutarlılığından daha yüksek bir R-kare değerine ulaştığı gözlemlenmiştir. Özetle;

**M+N+O=%98,88** şeklinde ifade edilmektedir.

Çizelge 5.20. Deneme 5 model içerikleri ve model tutarlılıkları özet tablosu

Deney Numarası	Girdi Model Kodu	Model Tutarlılığı (R-kare)
Deney 5	<b>M: m=1.9, c= 16</b>	<b>%91,47</b>
	<b>N: m=1.9, c= 17</b>	<b>%93,66</b>
	<b>O: m=1.9, c= 18</b>	<b>%92.14</b>
	<b>M+N</b>	<b>%98,41</b>
	<b>M+O</b>	<b>%98,02</b>
	<b>N+O</b>	<b>%98,60</b>
	<b>M+N+O</b>	<b>%98,88</b>



Bu deneyde de görüldüğü üzere, Çok Katmanlı Bulanık Regresyon Modeli, Bulanık Regresyon Modelinde oluşturulan modellere kıyasla R-kare değeri daha yüksek melez modellerin oluşturulmasına olanak sağlamaktadır

Bu değerlendirmelere ek olarak Çok Katmanlı Üyeliklerin modele eklenmesi göstermiştir ki işlem hacmini ve işlem süresini ciddi miktarda artıran öbek sayısı artırımına gitmek yerine birden fazla küçük öbek sayısı değerlerine sahip dolayısı ile daha küçük hacimli üyelik verilerinin modele dâhil edilmesi modelin tutarlılığının artırılmasını sağlamaktadır.

## 7. GELECEK ÇALIŞMALAR

Bu çalışma araştırma alanı özelinde daha önde hiç denenmemiş model yaklaşımlarını barındırmakla birlikte yeni sorular ile bir takım yaklaşımların denenmesinin ve modellenmesinin önünü de açmıştır. Bu soruların temelini Çizelge 5.15'te de belirtilen model iyileşme yüzdeleri oluşturmaktadır. Bu tabloda sabit bulanıklık mertebesi değerleri için öbek sayısının bir artırılması ile elde edilen parametreler eşliğinde hesaplanan üyelikler ve transformasyonlarının modelde yer almasının modelin tutarlılığı üzerine etkisi özetlenmiştir. Bu değişim 1.4,1.8 ve 2.2 bulanıklık mertebeleri için bu tabloda özetlenmiş olup 1.4 ile 2.6 arasında tüm bulanıklık mertebeleri için de bu çalışmada yapılan deneyler kapsamında hesaplanmıştır.

Yeni çalışmalara gebe olan ise modeldeki öbek sayısının bir artırma ile elde edilen model tutarlılığı iyileşmesinin güven aralığı kavramı ile ilişkilendirilip ilişkilendirilemeyeceği hususudur. Bu ilişkinin kurulması halinde modelde bir fazla öbek sayısına yer vererek oluşturulan modelin kurulmasının öncesinde güven aralığı yaklaşımı ile yaklaşık model iyileşmesi değerleri model kurulmadan tahmin edilebilecek olup yine bu araştırma alanında bugüne kadar mevzu bahis olmayan başka bir husus daha aydınlatılmış olacaktır.

Bir diğer husus ise farklı tip ve nitelikte veri setleri için bu çalışmada yer alana benzer kapsamda öbek sayısı aralığı ve bulanıklık mertebesi aralığı değerleri belirlenerek deneylerin tekrarlanıp önerilen çıkarım ve sonuçların teyit edilmesidir. Bu çıkarımların farklı veri setlerinde de bu çıkarımları çürütücü sonuçlar ortaya koymaması halinde güvenilirliği artacak olup gelecek çalışmalara rehberlik edebilecek nitelik kazanması söz konusu olacaktır.

Diğer yandan bu çalışmanın ilk bölümünde öbek sayısındaki değişikliklerin modelin tutarlılığı üzerine etkisi çalışılmış olup aynı şekilde bulanıklık mertebesindeki değişikliklerin modelin tutarlılığı üzerine etkisi hem teorik hem de numerik şekilde incelemeye açıktır.

Son olarak bulanık küme yaklaşımın çıktı tahmini yapmak üzere kurulan modellerde kullanım alanlarını genişletmek adına finansal verilerin dışında farklı disiplinlerde ne şekilde kullanılabileceği hususunda yapılacak disiplinler arası bir ön çalışma anlamlı bulunmaktadır.

Bu yaklaşım ve modellemede kullanılan algoritmaların, farklı alanlarda kullanımının önünü açmakla kalmayıp farklı disiplinlerde veriler arasında sebep sonuç ilişkisine bağlı olarak var olan doğrudan ilişkilerin olduğu veri setlerinde kullanımı aracılığı ile oluşan model tutarlılığı iyileşmesi hakkında da fikir verecektir.

## KAYNAKLAR

- [1] Çelikyılmaz, A., Türkşen, İ.B., Modeling Uncertainty with Fuzzy Logic with Recent Theory and Applications, *Springer*, Berlin Heidelberg, 2009.
- [2] Zadeh, L.A., Fuzzy Sets, *Information and Control*, 8, 338-353, 1965.
- [3] Tagaki, T., Sugeno, M., Fuzzy Identification of Systems and Its Applications to Modeling and Control, *IEEE Transactions on Systems, Man and Cybernetics SMC*, 15(1), 116-132, 1985.
- [4] Mizumoto, M., Method of Fuzzy Inference Suitable for Fuzzy Control, *J. Soc. Instrument Control Engineering*, 58, 959-963, 1989.
- [5] Mamdani, E.H., Assilian, S., An Experiment in Linguistic Synthesis with a Fuzzy Logic Controller, *Int. J. Man-Machine Studies*, 7, 1-13, 1974.
- [6] Bezdek, J.C., Pattern Recognition with Fuzzy Objective Function. *Plenum Press*, New York, 1981a.
- [7] Hathaway, R.J., Bezdek, J.C., Switching Regression Models and Fuzzy Clustering, *IEEE Transactions on Fuzzy Systems*, 1(3), 195-204, 1993.
- [8] “Hiper-küre ve hiper-düzlem illüstrasyonu” erişim adresi:  
<http://web.cs.sunyit.edu/~sengupta/>,  
[http://www.math.umd.edu/~jmr/241/lines\\_planes.html](http://www.math.umd.edu/~jmr/241/lines_planes.html), erişim tarihi: 11 Mayıs 2014.
- [9] “Kantil ve ekspektıl” erişim adresi:  
[http://www.medeniyet.edu.tr/istatistik\\_arastirma\\_alanlari.html](http://www.medeniyet.edu.tr/istatistik_arastirma_alanlari.html), erişim tarihi: 18 Nisan 2014.
- [10] “Zaman Serileri Analizi”, erişim adresi:  
[http://www.medeniyet.edu.tr/istatistik\\_arastirma\\_alanlari.html](http://www.medeniyet.edu.tr/istatistik_arastirma_alanlari.html), erişim tarihi: 16 Nisan 2014.
- [11] Yrd. Doç. Dr. Murat Özbayođlu ile 4 Mart 2013’teki görüşme.
- [12] Montgomery, D.C., Runger, G.C., Applied Statistics and Probability for Engineers, *Wiley*, New Jersey, 2006.
- [13] Prof. Dr. İsmail Burhan Türkşen ile 11 Aralık 2013’teki görüşme.
- [14] Kim, E., Ramakrishna, R.S., New Indices for Cluster Validity Assessment, *Pattern Recognition Letters*, 26, 2353-2363, 2005.
- [15] Xie, X.L., Beni G.A., Validity Measure for Fuzzy Clustering, *IEEE Trans. Pattern and Machine Intelligence* 3(8), 841-846, 1991.

- [16] Kim, D.-W., Lee, K.H., Lee, D., Fuzzy Cluster Validation Index Based on Inter-cluster Proximity, *Pattern Recognition Letters*, 24, 2561-2574, 2003.
- [17] Kung, C.-C., Lin, C.-C., A New Cluster Validity Criterion for Fuzzy C-Regression Model and its Application to T-S Fuzzy Model Identification, 13<sup>th</sup> IEEE International Conference on Fuzzy Systems, 3, 1673-1678, Budapest, Hungary, Temmuz 2004.
- [18] Çelikyılmaz, A., Türkşen, İ.B., A New Cluster Validity Index with Fuzzy Functions, 12<sup>th</sup> International Fuzzy Systems Association World Congress, IFSA 2007, Cancun, Mexico, Haziran 2007.
- [19] Çelikyılmaz, A., Türkşen, İ.B., Increasing Accuracy of Two-Class Pattern Recognition with Enhanced Fuzzy Functions, *Expert Systems with Applications*, 2007.
- [20] Türkşen, İ.B., Type-1 and Type-2 Fuzzy System Modeling, *Fuzzy Sets and Systems*, 106, 11-34, 1999.
- [21] Özkan, I., Türkşen, İ.B., Upper and Lower Level of Fuzziness of FCM, *Information Sciences Special Issue*, 177(23), 5143-5152, 2007.
- [22] Türkşen, İ.B., *An Ontological and Epistemological Perspective of Fuzzy Theory*, Elsevier, Amsterdam, 2006.
- [23] Gustafson, D., Kessel, W., Fuzzy Clustering with a Fuzzy Covariance Matrix, *Proc. 18<sup>th</sup> IEEE CDC Conference*, San Diego, USA, 761-766, Aralık 1979.
- [24] Prof. Dr. İsmail Burhan Türkşen ile 25 Nisan 2014'teki görüşme.
- [25] Çelikyılmaz, A., Türkşen, İ.B.: Uncertainty modeling with evolutionary improved fuzzy functions approach. *IEEE Transactions on System, Man, and Cybernetics-SMCB* (2007)
- [26] Khuri, A.I., *Advanced Calculus with Applications in Statistics*, Wiley Interscience, Hoboken, 2003.
- [27] Çelikyılmaz, A., Türkşen, İ.B., Uncertainty Bounds of Fuzzy C-Regression Method, *IEEE World Congress on Computational Intelligence*, Hong Kong, Haziran 2008.
- [28] Demirci, M., Fuzzy Functions and Their Fundamental Properties, *Fuzzy Sets and Systems*, 106(2), 239-246, 1999.
- [29] Türkbâl, A., *Bilimsel Araştırma Yöntemleri ve Yazma Teknikleri*, Aktif Yayınları, Erzurum, 1981.
- [30] Yrd. Doç. Dr. Mükerrerem Bahar Başkır ile 16 Eylül 2013'teki görüşme.

## ÖZGEÇMİŞ

### Kişisel Bilgiler

Soyadı, adı : ASLAN ÖZŞAHİN, Selcen Gülsüm  
Uyruğu : T.C.  
Doğum tarihi ve yeri : 10.03.1988, Ankara  
Medeni hali : Evli  
Telefon : 0 (542 418 27 77)  
E-posta : sgaslan@etu.edu.tr

### Eğitim

Derece	Eğitim Birimi	Mezuniyet tarihi
Y. Lisans	TOBB ETÜ /Endüstri Müh.	2014
Lisans	Atılım Üniversitesi /Endüstri Müh. (ÖSYM Bursu ve Bölüm 3.lüğü ile)	2009

### İş Deneyimi

Yıl	Yer	Görev
2011-...	TÜBİTAK Uluslararası İşbirliği Daire Başkanlığı	Bilimsel Prog. Uzm. Yard.
2009-2010	Atılım Üniversitesi Endüstri Mühendisliği Bölümü	Araşt.Görevlisi

### Yabancı Dil

İngilizce (İleri Derecede)

### Yayımlar

ASLAN ÖZŞAHİN, S.G., TÜRKŞEN İ.B., Usage Of Newly Established Methods in Modelling of Economical Data: Fuzzy Systems, International Conference on Knowledge, Economy and Management,ICKEM14,21-24, Antalya, Turkey, Kasım 2014.