

**T.C.  
YILDIZ TEKNİK ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ**

**KARTOGRAFİK GENELLEŞTİRMEDE SEÇME/ELEME İŞLEMİ İÇİN YAPAY  
ZEKA YÖNTEMLERİNİN AKARSU AĞLARINA UYGULANABİLİRLİĞİ**

**ALPER ŞEN**

**DOKTORA TEZİ  
HARİTA MÜHENDİSLİĞİ ANABİLİM DALI  
UZAKTAN ALGILAMA VE CBS PROGRAMI**

**DANIŞMAN  
DOÇ. DR. TÜRKEY GÖKGÖZ**

**İSTANBUL, 2013**

T.C.  
YILDIZ TEKNİK ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ

**KARTOGRAFİK GENELLEŞTİRMEDE SEÇME/ELEME İŞLEMİ İÇİN YAPAY  
ZEKA YÖNTEMLERİNİN AKARSU AĞLARINA UYGULANABİLİRLİĞİ**

Alper ŞEN tarafından hazırlanan tez çalışması 12.06.2013 tarihinde aşağıdaki jüri tarafından Yıldız Teknik Üniversitesi Fen Bilimleri Harita Mühendisliği Anabilim Dalı'nda **DOKTORA TEZİ** olarak kabul edilmiştir.

**Tez Danışmanı**

Doç. Dr. Türkay GÖKGÖZ  
Yıldız Teknik Üniversitesi

**Jüri Üyeleri**

Doç. Dr. Türkay GÖKGÖZ  
Yıldız Teknik Üniversitesi

\_\_\_\_\_

Prof. Dr. Mehmet SELÇUK  
Yıldız Teknik Üniversitesi

\_\_\_\_\_

Prof. Dr. Necla ULUĞTEKİN  
İstanbul Teknik Üniversitesi

\_\_\_\_\_

Doç. Dr. Melih BAŞARANER  
Yıldız Teknik Üniversitesi

\_\_\_\_\_

Prof. Dr. Cengizhan İPBÜKER  
İstanbul Teknik Üniversitesi

\_\_\_\_\_

Bu alıřma, TBİTAK Bilim İnsanı Destekleme Daire Başkanlıđı'nın 2214 yurt dıřı doktora arařtırma burs programı ile desteklenmiřtir.

## ÖNSÖZ

---

Öncelikle çalışmamın her aşamasında ve her konuda bana destek olan, yardım eden, yönlendiren, bilgi ve deneyimini esirgemeyen danışman hocam Kartografya Anabilim Dalı Başkanı Sayın Doç. Dr. Trkay GKGZ'e sonsuz Őkranlarımı sunarım. Desteklerini her zaman hissettiğim, çalışmamın her ara raporunu titizlikle inceleyen, bilgi ve deneyimlerini esirgemeyen tez izleme komitesindeki hocalarım Sayın Prof. Dr. Mehmet SELÇUK ve Sayın Prof. Dr. Necla ULUĞTEKİN'e çok teşekkür ederim.

Kartografik genelleştirme alanında yıllardan beri sürdürdüğü araştırmalarla haklı ve özel bir yer kazanan, Hannover Üniversitesi Kartografya ve Jeoenformatik Enstitüsü'nde gerçekleştirdiğim araştırmada, bilgi ve deneyimleriyle bana yardımcı olan Enstitü Başkanı Sayın Prof. Dr. -Ing. habil. Monika SESTER'a, tüm Enstitü çalışanlarına ve yurt dışı doktora araştırması desteğinden dolayı TÜBİTAK'a teşekkür ederim.

Tez çalışması süresince öneri ve yardımlarıyla bana destek olan hocalarım Sayın Prof. Dr. Cengizhan İPBÜKER, Sayın Doç. Dr. Ali Melih BAŞARANER, Sayın Yrd. Doç. Dr. Fatih GÜLGEN ve Sayın Yrd. Doç. Dr. Birsan Eygi ERDOĞAN'a teşekkür ederim.

Harita Mühendisliği Bölümü'ndeki diğer hocalarıma ve dostluklarını her zaman hissettiğim tüm Araştırma Görevlisi arkadaşlarıma teşekkür ederim.

Çalışma boyunca beni ilgi ve sabırla destekleyen sevgili eşim Oya ŞEN'e, beni özlemle bekleyen biricik kızım Defne ŞEN'e ve beni yetiştiren anne ve babama teşekkürü bir borç bilirim.

Haziran, 2013

Alper ŞEN

## İÇİNDEKİLER

|                                                                               | Sayfa |
|-------------------------------------------------------------------------------|-------|
| SİMGE LİSTESİ.....                                                            | viii  |
| KISALTMA LİSTESİ.....                                                         | x     |
| ŞEKİL LİSTESİ.....                                                            | xi    |
| ÇİZELGE LİSTESİ .....                                                         | xiv   |
| ÖZET .....                                                                    | xvi   |
| ABSTRACT.....                                                                 | xvii  |
| BÖLÜM 1                                                                       |       |
| GİRİŞ.....                                                                    | 1     |
| 1.1    Literatür Özeti .....                                                  | 1     |
| 1.2    Tezin Amacı .....                                                      | 13    |
| 1.3    Hipotez .....                                                          | 13    |
| BÖLÜM 2                                                                       |       |
| ÇALIŞMANIN HİDROJEOMORFOLOJİK TEMELLERİ .....                                 | 15    |
| 2.1    Tanımlar .....                                                         | 15    |
| 2.2    Akarsu Ağı Desenlerine Göre Akarsu Ağlarının Sınıflandırılması .....   | 18    |
| 2.3    Akarsu Sıralaması .....                                                | 21    |
| 2.4    Akarsu Kolu Desenlerine Göre Akarsu Kollarının Sınıflandırılması ..... | 23    |
| 2.5    Akarsu Ağı ve Havzasına Ait Jeomorfometrik Parametreler.....           | 24    |
| BÖLÜM 3                                                                       |       |
| ÇALIŞMADA KULLANILAN VERİLERİN ÖZELLİKLERİ .....                              | 25    |
| 3.1    Hidrografik Nesne Tipleri ve Özellikleri .....                         | 26    |
| 3.2    Anlamlı Hidrografik Parçalar .....                                     | 28    |
| 3.2.1    Anlamlı Akarsu Parçalarının Kodlanması .....                         | 31    |

|                        |                                                                                  |            |
|------------------------|----------------------------------------------------------------------------------|------------|
| 3.2.2                  | Anlamalı Akarsu ve Kıyı Parçaları Arasındaki Akış İlişkilerinin Kodlanması ..... | 39         |
| 3.3                    | Akarsu Düzeyleri .....                                                           | 45         |
| 3.4                    | Anlamalı Akarsu Parçaları Akış İlişkileri Arasında Akarsu Düzeyleri İzleme ..... | 46         |
| 3.5                    | Coğrafi İsimler .....                                                            | 48         |
| 3.6                    | Hidrolojik Birimlerin Organizasyonu ve Örnekler .....                            | 48         |
| 3.7                    | Veri Toplama Koşulları .....                                                     | 50         |
| <b>BÖLÜM 4</b>         |                                                                                  |            |
| <b>YÖNTEMLER .....</b> |                                                                                  | <b>51</b>  |
| 4.1                    | İstatistiksel Yöntemler .....                                                    | 51         |
| 4.1.1                  | Tanımlayıcı İstatistik .....                                                     | 51         |
| 4.1.2                  | Normal Dağılım .....                                                             | 52         |
| 4.1.3                  | Normalizasyon .....                                                              | 55         |
| 4.1.4                  | Hipotez Testleri .....                                                           | 56         |
| 4.1.5                  | Normal Dağılıma Uygunluğun Araştırılması .....                                   | 59         |
| 4.1.5.1                | Çarpıklık ve Basıklık Katsayısı .....                                            | 60         |
| 4.1.5.2                | Histogram Grafiği, Kantil-Kantil Grafiği ve Kutu Grafiği .....                   | 61         |
| 4.1.5.3                | Kolmogrov-Smirnov Uygunluk Tek Örneklem Testi .....                              | 64         |
| 4.1.6                  | Khi-Kare Bağımsızlık Testi .....                                                 | 65         |
| 4.2                    | Kümeleme ve Sınıflandırma Yöntemleri .....                                       | 70         |
| 4.3                    | Yapay Zeka .....                                                                 | 71         |
| 4.3.1                  | Yapay Sinir Ağları .....                                                         | 73         |
| 4.3.2                  | Kendini Düzenleyen Haritalar .....                                               | 75         |
| 4.3.3                  | İstatistiksel Öğrenme Teorisi .....                                              | 89         |
| 4.3.4                  | Destek Vektör Makineleri .....                                                   | 95         |
| <b>BÖLÜM 5</b>         |                                                                                  |            |
| <b>UYGULAMA .....</b>  |                                                                                  | <b>102</b> |
| 5.1                    | Veriler ve Çalışma Alanı .....                                                   | 104        |
| 5.2                    | Ön İşlemler .....                                                                | 112        |
| 5.2.1                  | Çalışmada Kullanılan Özniteliklerin Belirlenmesi .....                           | 112        |
| 5.2.1.1                | Geometrik Öznitelikler .....                                                     | 113        |
| 5.2.1.2                | Topolojik Öznitelikler .....                                                     | 113        |
| 5.2.1.3                | Semantik Öznitelikler .....                                                      | 116        |
| 5.2.1.4                | Özniteliklerin Normal Dağılıma Uygunluğunun Araştırılması ....                   | 116        |
| 5.2.2                  | Ağırlıkların Belirlenmesi .....                                                  | 123        |
| 5.3                    | Kümelerin Seçimi için Geliştirilen Yöntem .....                                  | 129        |
| 5.4                    | Destek Vektör Makineleri ile Uygulama .....                                      | 132        |
| 5.5                    | Son İşlemler .....                                                               | 134        |
| <b>BÖLÜM 6</b>         |                                                                                  |            |
| <b>BULGULAR .....</b>  |                                                                                  | <b>135</b> |
| <b>BÖLÜM 7</b>         |                                                                                  |            |

|                                                |     |
|------------------------------------------------|-----|
| SONUÇ VE ÖNERİLER .....                        | 151 |
| KAYNAKLAR.....                                 | 154 |
| EK-A                                           |     |
| ALT HAVZALARIN KÜMELEMESİNDE U-MATRİSLERİ..... | 162 |
| ÖZGEÇMİŞ .....                                 | 170 |

## SİMGE LİSTESİ

---

|               |                                                                                     |
|---------------|-------------------------------------------------------------------------------------|
| $A_b$         | Havza alanı                                                                         |
| $C_b$         | Harita işareti için kullanılan abartma sabiti                                       |
| $Cir_b$       | Havzanın daireselliği                                                               |
| $Cr_b$        | Havzanın zirve noktası                                                              |
| $C_z$         | Kartografik işaretin türü                                                           |
| $C_B(p_k)$    | $p_k$ noktasının arasında olma bakımından merkezlik derecesi                        |
| $C_B'(p_k)$   | $p_k$ noktasının normalize edilmiş arasında olma bakımından merkezlik derecesi      |
| $C_C(p_k)$    | $p_k$ noktasının yakınlık bakımından merkezlik derecesi                             |
| $C_C'(p_k)$   | $p_k$ noktasının normalize edilmiş yakınlık bakımından merkezlik derecesi           |
| $C_D(p_k)$    | $p_k$ noktasının bağılılık bakımından merkezlik derecesi                            |
| $C_D'(p_k)$   | $p_k$ noktasının normalize edilmiş bağılılık bakımından merkezlik derecesi          |
| $C$           | Ceza parametresi                                                                    |
| $D_b$         | Akarsu yoğunluğu                                                                    |
| $d(i,j)$      | $i$ ve $j$ vektörleri arasındaki uzunluk                                            |
| $D_s$         | Denk gösterilen nesne sayısı                                                        |
| $d_\alpha$    | Smirnov'un kritik tablo değerleri                                                   |
| $d(p_i, p_k)$ | $p_i, p_k$ düğümleri arasındaki en kısa güzergahda bulunan kenar sayısı             |
| $D_u$         | Denk gösterilen toplam uzunluk değeri                                               |
| $E_{ij}$      | Beklenen frekans değeri                                                             |
| $E_s$         | Eksik gösterilen nesne sayısı                                                       |
| $E_u$         | Eksik gösterilen toplam uzunluk değeri                                              |
| $F_s$         | Fazladan gösterilen nesne sayısı                                                    |
| $F_u$         | Fazladan gösterilen toplam uzunluk değeri                                           |
| $g_{ij}$      | $i$ ve $j$ düğümleri arasındaki en kısa güzergah sayısı                             |
| $g_{ij}(p_k)$ | $p_i$ ve $p_j$ düğümleri arasındaki $p_k$ noktasından geçen en kısa güzergah sayısı |
| $h_{ci}$      | Komşuluk fonksiyonu                                                                 |
| $k$           | Serbestlik derecesi                                                                 |
| $K_n$         | Kantil değeri                                                                       |
| $L_c$         | Havzada toplam akarsu uzunluğu                                                      |
| $M$           | Melton'ın engebe değeri                                                             |
| $m_a$         | Kaynak ölçek sayısı                                                                 |
| $m_f$         | Hedef ölçek sayısı                                                                  |
| $n$           | Gözlem sayısı                                                                       |

|             |                                                                          |
|-------------|--------------------------------------------------------------------------|
| $n_a$       | Kaynak nesne sayısı                                                      |
| $n_f$       | Hedef nesne sayısı                                                       |
| $O_{ij}$    | Gözlem frekans değeri                                                    |
| $O_u$       | Fazla gösterilen toplam uzunluk değerinin uzunluk büyüme faktörüne oranı |
| $P_b$       | Havzanın çevresi                                                         |
| $R_b$       | Rölyef                                                                   |
| $R(w)$      | Beklenen risk                                                            |
| $s$         | Örneklemin standart sapması                                              |
| $S_b$       | Havzanın eğimi                                                           |
| $SH_b$      | Basıklık standart hatası                                                 |
| $SH_ç$      | Çarpıklık standart hatası                                                |
| $X'$        | Normalize edilmiş değer                                                  |
| $\bar{X}$   | Örneklemin ortalama değeri                                               |
| $\chi^2$    | Khi-kare                                                                 |
| $w$         | Ağırlık vektörü                                                          |
| $Z$         | Standart değer                                                           |
| $\sum L_c$  | Havzada 10 m. aralıklı yükseklik eğrilerinin toplam uzunluğu             |
| $\alpha$    | Anlamlılık düzeyi                                                        |
| $\Gamma(n)$ | Gama fonksiyonu                                                          |
| $\phi_c$    | Cramer'in V katsayısı                                                    |
| $\xi_i$     | Gevşek değişkenler                                                       |
| $\emptyset$ | Çekirdek fonksiyon                                                       |
| $\gamma$    | Kernel parametresi                                                       |
| $\sigma$    | Anakütlenin standart sapması                                             |
| $\sigma^2$  | Varyans                                                                  |
| $\mu$       | Anakütlenin ortalama değeri                                              |

## KISALTMA LİSTESİ

---

|       |                                                  |
|-------|--------------------------------------------------|
| AAP   | Anlamlı Akarsu Parçası                           |
| ABD   | Amerika Birleşik Devletleri                      |
| AKP   | Anlamlı Kıyı Parçası                             |
| ASKP  | Anlamlı Su Kütlesi Parçası                       |
| CBS   | Coğrafi Bilgi Sistemleri                         |
| DEM   | Digital Elevation Model                          |
| DLG   | Digital Line Graph                               |
| DVM   | Destek Vektör Makineleri                         |
| GNIS  | Geographic Names Information System              |
| ISO   | International Organization for Standardization   |
| KDH   | Kendini Düzenleyen Haritalar                     |
| K-S   | Kolmogorov-Smirnov                               |
| LEF   | Uzunluk Büyüme Faktörü (Length Expansion Factor) |
| NHD   | National Hydrography Dataset                     |
| NSDI  | National Spatial Data Infrastructure             |
| ÖÇK   | Örtüşen Çizgi Katsayısı                          |
| PT    | Pomme de Terre                                   |
| Q-Q   | Quantile-Quantile                                |
| RF    | Reach File                                       |
| SBP   | South Branch Potomac                             |
| SCM   | Sayısal Coğrafi Model                            |
| SKM   | Sayısal Kartografik Model                        |
| SOM   | Self Organizing Maps                             |
| SVM   | Support Vector Machines                          |
| UHVS  | Ulusal Hidrografi Veri Seti                      |
| USEPA | United States Environmental Protection Agency    |
| USGS  | United States Geological Survey                  |
| UTM   | Universal Transverse Mercator                    |
| VC    | Vapnik-Chervonenkis                              |
| YSA   | Yapay Sinir Ağları                               |
| YRM   | Yapısal Risk Minimizasyonu                       |

## ŞEKİL LİSTESİ

|                                                                                             | Sayfa |
|---------------------------------------------------------------------------------------------|-------|
| Şekil 1.1 Model ve kartografik genelleştirme .....                                          | 3     |
| Şekil 1.2 Hidrografik nesnelerin seçimi .....                                               | 4     |
| Şekil 2.1 Bir akarsu ağının planı.....                                                      | 17    |
| Şekil 2.2 Ağaçsı desen .....                                                                | 18    |
| Şekil 2.3 Paralel desen.....                                                                | 19    |
| Şekil 2.4 Kafes desen.....                                                                  | 19    |
| Şekil 2.5 Radyal desen.....                                                                 | 20    |
| Şekil 2.6 Merkezci desen .....                                                              | 20    |
| Şekil 2.7 Dairesel desen .....                                                              | 20    |
| Şekil 2.8 Dengesiz desen .....                                                              | 21    |
| Şekil 2.9 Dikdörtgen yapıli desen .....                                                     | 21    |
| Şekil 2.10 Akış aşağı akarsu sıralaması ve akış yukarı ana kollar .....                     | 22    |
| Şekil 2.11 Shreve sıralaması .....                                                          | 22    |
| Şekil 2.12 Akarsu kolları .....                                                             | 23    |
| Şekil 3.1 Geçit nesnesinin gösterimi .....                                                  | 27    |
| Şekil 3.2 Nokta, çizgi ve alan hidrografik nesneler.....                                    | 28    |
| Şekil 3.3 UHVS mimarisi .....                                                               | 30    |
| Şekil 3.4 AAP'lerin belirlenmesi (aynı tip nesneler için).....                              | 32    |
| Şekil 3.5 AAP'lerin belirlenmesi (farklı tip nesneler için) .....                           | 32    |
| Şekil 3.6 AAP'lerin belirlenmesi (farklı tip nesneler için) .....                           | 33    |
| Şekil 3.7 Önemli göl/gölcük nesneleri için AAP'lerin belirlenmesi .....                     | 33    |
| Şekil 3.8 Kavşak-kavşak kuralına göre AAP'lerin belirlenmesi .....                          | 34    |
| Şekil 3.9 Kavşak-kavşak kuralına göre AAP'lerin belirlenmesi (aynı tip nesneler için) ..... | 34    |
| Şekil 3.10 Çoklu kanallarda AAP'lerin belirlenmesi .....                                    | 35    |
| Şekil 3.11 AAP'lerin belirlenmesinde önemli ve önemli kavşaklar (5 mil kuralı) .....        | 36    |
| Şekil 3.12 Akarsu/nehir ile kanal/hendek kavşağında AAP'lerin belirlenmesi .....            | 37    |
| Şekil 3.13 Dallanan orta eksenlerle AAP'lerin belirlenmesi.....                             | 37    |
| Şekil 3.14 Dallanan orta eksenlerle AAP'lerin belirlenmesi.....                             | 38    |
| Şekil 3.15 Göllerde AAP'lerin belirlenmesi .....                                            | 39    |
| Şekil 3.16 AAP'ler arasında akış ilişkileri .....                                           | 41    |
| Şekil 3.17 AAP'ler arasında akış ilişkileri .....                                           | 42    |
| Şekil 3.18 Dallanan AAP'ler arasında akış ilişkileri.....                                   | 42    |
| Şekil 3.19 Akış olmayan bağlantı ve akış ilişkileri.....                                    | 43    |

|                                                                                                                                                                                                                 |     |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Şekil 3.20 AAP ve AKP'ler boyunca akış ilişkilerinin sıralanması.....                                                                                                                                           | 44  |
| Şekil 3.21 Akarsu ağında akarsu düzeylerinin atanması .....                                                                                                                                                     | 45  |
| Şekil 3.22 Mississippi nehri boyunca akarsu düzeyleri .....                                                                                                                                                     | 46  |
| Şekil 3.23 Akış ilişkileri kullanılarak kodlanan akarsu düzeyleri .....                                                                                                                                         | 47  |
| Şekil 3.24 AAP ve ASKP'lerin coğrafi isimleri .....                                                                                                                                                             | 48  |
| Şekil 3.25 New England bölgesi hidrolojik birimleri .....                                                                                                                                                       | 49  |
| Şekil 4.1 Aynı ortalamaya sahip farklı varyanslı iki farklı normal dağılım .....                                                                                                                                | 53  |
| Şekil 4.2 Standart normal dağılımın olasılık yoğunluk fonksiyonu ve kümülatif dağılım<br>fonksiyonu .....                                                                                                       | 54  |
| Şekil 4.3 İki yanlı güven düzeyi için test büyüklüğüne ilişkin kabul ve ret bölgeleri.....                                                                                                                      | 57  |
| Şekil 4.4 Tek yanlı güven düzeyi için test büyüklüğüne ilişkin kabul ve ret bölgeleri...                                                                                                                        | 58  |
| Şekil 4.5 Test türünün belirlenmesi .....                                                                                                                                                                       | 59  |
| Şekil 4.6 Normal ve çarpık dağılımlar.....                                                                                                                                                                      | 60  |
| Şekil 4.7 Normal dağılımlı histogram .....                                                                                                                                                                      | 61  |
| Şekil 4.8 Normal dağılımlı verilerin Q-Q grafiği.....                                                                                                                                                           | 62  |
| Şekil 4.9 Normal dağılım ve kutu grafik .....                                                                                                                                                                   | 63  |
| Şekil 4.10 Kutu grafikte aykırı değerlerin gösterimi.....                                                                                                                                                       | 63  |
| Şekil 4.11 K-S testinin kümülatif dağılım fonksiyonunda gösterimi .....                                                                                                                                         | 65  |
| Şekil 4.12 Farklı serbestlik derecelerine göre khi-kare dağılımları .....                                                                                                                                       | 66  |
| Şekil 4.13 Tek yanlı güven düzeyi için kabul ve ret bölgeleri.....                                                                                                                                              | 67  |
| Şekil 4.14 Denetimli ileri beslemeli sinir ağları .....                                                                                                                                                         | 74  |
| Şekil 4.15 KDH'de girdi ve çıktı katmanları .....                                                                                                                                                               | 76  |
| Şekil 4.16 Kare topoloji ve altıgen topoloji .....                                                                                                                                                              | 77  |
| Şekil 4.17 Kazanan nöronun girdi vektörüne göre değişimi .....                                                                                                                                                  | 78  |
| Şekil 4.18 Eğitim boyunca azalan komşuluk yarıçapı.....                                                                                                                                                         | 80  |
| Şekil 4.19 KDH eğitimi .....                                                                                                                                                                                    | 80  |
| Şekil 4.20 U-matrisinde nöronlar arası benzerlik.....                                                                                                                                                           | 81  |
| Şekil 4.21 Küme eleman sayıları .....                                                                                                                                                                           | 81  |
| Şekil 4.22 Toplam 100,000 döngüden oluşan bir KDH ağının eğitiminde her 10,000<br>döngü sonunda iki girdi vektörünün konumları .....                                                                            | 82  |
| Şekil 4.23 Kazanan ve komşu nöronların güncelleme öncesi (devamlı çizgi) ve sonrası<br>(kesikli çizgi) durumları.....                                                                                           | 82  |
| Şekil 4.24 Örnek KDH eğitimi .....                                                                                                                                                                              | 84  |
| Şekil 4.25 Veri setinin eğitim, test ve doğrulama olarak bölünmesi .....                                                                                                                                        | 91  |
| Şekil 4.26 Regresyon problemine bir örnek.....                                                                                                                                                                  | 93  |
| Şekil 4.27 Model karmaşıklığının (VC boyutunun) artması ile genelleme ve deneysel<br>hatanın değişimi .....                                                                                                     | 93  |
| Şekil 4.28 İki boyutlu uzayda doğrusal olarak ayrılabilen veriler.....                                                                                                                                          | 96  |
| Şekil 4.29 Doğrusal olarak ayrılabilen veriler arasındaki en büyük boşluk.....                                                                                                                                  | 96  |
| Şekil 4.30 C'nin alacağı değerlere göre marjin.....                                                                                                                                                             | 99  |
| Şekil 4.31 İki boyutlu uzayda doğrusal olmayan bir sınıflandırıcı ile ayrılabilen verilerin<br>bir çekirdek fonksiyonu ile üç boyutlu uzaya taşınması ve bir doğrusal<br>sınıflandırıcı ile ayrılabilmesi ..... | 100 |
| Şekil 5.1 Metodoloji akış şeması .....                                                                                                                                                                          | 103 |
| Şekil 5.2 USGS TNM 2.0 Viewer .....                                                                                                                                                                             | 104 |
| Şekil 5.3 PT alt havza sınırları .....                                                                                                                                                                          | 105 |

|            |                                                                                                                   |     |
|------------|-------------------------------------------------------------------------------------------------------------------|-----|
| Şekil 5.4  | SBP alt havza sınırları .....                                                                                     | 106 |
| Şekil 5.5  | Orijinal 1:24,000 ölçekli PT'nin 1:400,000 ölçeğinde gösterimi .....                                              | 107 |
| Şekil 5.6  | Orijinal 1:24,000 ölçekli SBP'nin 1:500,000 ölçeğinde gösterimi .....                                             | 108 |
| Şekil 5.7  | Orijinal 1:100,000 ölçekli PT'nin 1:400,000 ölçeğinde gösterimi .....                                             | 109 |
| Şekil 5.8  | Orijinal 1:100,000 ölçekli SBP'nin 1:500,000 ölçeğinde gösterimi .....                                            | 110 |
| Şekil 5.9  | Kıvrımlılık özneteliği .....                                                                                      | 113 |
| Şekil 5.10 | Bir akarsu ağı ve bağlantı çizgesi .....                                                                          | 114 |
| Şekil 5.11 | Pajek yazılımı arayüzü .....                                                                                      | 115 |
| Şekil 5.12 | PT'ye dair histogram ve kutu grafikleri .....                                                                     | 120 |
| Şekil 5.13 | PT'ye dair kantil-kantil grafikleri .....                                                                         | 121 |
| Şekil 5.14 | SBP'ye dair histogram ve kutu grafikleri .....                                                                    | 122 |
| Şekil 5.15 | SBP'ye dair kantil-kantil grafikleri .....                                                                        | 123 |
| Şekil 5.16 | Ağırlıklara dair grafik .....                                                                                     | 125 |
| Şekil 5.17 | Orijinal 1:100,000 ölçekli PT akarsu ağında öznetelik ilişkileri .....                                            | 127 |
| Şekil 5.18 | Orijinal 1:100,000 ölçekli SPB akarsu ağında öznetelik ilişkileri .....                                           | 128 |
| Şekil 5.19 | Brush alt havzasında KDH eğitimi sonunda oluşan düzlemler, U-matrisi ve küme eleman sayıları .....                | 129 |
| Şekil 5.20 | U-matrisinde kümelerin seçimi için geliştirilen yöntem .....                                                      | 131 |
| Şekil 5.21 | U-matrisinde seçilen kümelerin eleman sayıları .....                                                              | 132 |
| Şekil 6.1  | Orijinal 1:24,000, orijinal 1:100,000, KDH 1:100,000 ve DVM 1:100,000 ölçeğinde kafes desenli benzer kollar ..... | 136 |
| Şekil 6.2  | Orijinal 1:24,000, orijinal 1:100,000, türetilen 1:100,000 .....                                                  | 140 |
| Şekil 6.3  | Orijinal PT'de ve KDH'ye göre türetilen haritada daimi ve mevsimsel dere sayısı ile ilgili sütun grafik .....     | 142 |
| Şekil 6.4  | Orijinal PT'de ve DVM'ye göre türetilen haritada daimi ve mevsimsel dere sayısı ile ilgili sütun grafik .....     | 142 |
| Şekil 6.5  | Orijinal SBP'de ve KDH'ye göre türetilen haritada daimi ve mevsimsel dere sayısı ile ilgili sütun grafik .....    | 142 |
| Şekil 6.6  | Orijinal SBP'de ve DVM'ye göre türetilen haritada daimi ve mevsimsel dere sayısı ile ilgili sütun grafik .....    | 143 |
| Şekil 6.7  | Orijinal PT'de ve KDH'ye göre türetilen haritada akarsu düzeyleri ile ilgili sütun grafik .....                   | 145 |
| Şekil 6.8  | Orijinal PT'de ve DVM'ye göre türetilen haritada akarsu düzeyleri ile ilgili sütun grafik .....                   | 146 |
| Şekil 6.9  | Orijinal SBP'de ve KDH'ye göre türetilen haritada akarsu düzeyleri ile ilgili sütun grafik .....                  | 146 |
| Şekil 6.10 | Orijinal SBP'de ve DVM'ye göre türetilen haritada akarsu düzeyleri ile ilgili sütun grafik .....                  | 146 |
| Şekil 6.11 | Kendini düzenleyen haritalar yöntemiyle türetilen 1:100,000 PT akarsu ağı (1:400,000 ölçekli gösterim) .....      | 147 |
| Şekil 6.12 | Destek vektör makineleri yöntemiyle türetilen 1:100,000 PT akarsu ağı (1:400,000 ölçekli gösterim) .....          | 148 |
| Şekil 6.13 | Kendini düzenleyen haritalar yöntemiyle türetilen 1:100,000 SBP akarsu ağı (1:500,000 ölçekli gösterim) .....     | 149 |
| Şekil 6.14 | Destek vektör makineleri yöntemiyle türetilen 1:100,000 SBP akarsu ağı (1:500,000 ölçekli gösterim) .....         | 150 |

## ÇİZELGE LİSTESİ

|                                                                                                       | Sayfa |
|-------------------------------------------------------------------------------------------------------|-------|
| Çizelge 1.1 Bazı hidrografik nesneler için genelleştirme kuralları .....                              | 8     |
| Çizelge 1.2 Detay önem hiyerarşisi .....                                                              | 9     |
| Çizelge 1.3 Bağlılık katsayısı .....                                                                  | 9     |
| Çizelge 2.1 Akarsu ağı ve havzasına ait jeomorfometrik parametreler .....                             | 24    |
| Çizelge 3.1 Nesne tipleri ve tanımları .....                                                          | 26    |
| Çizelge 3.2 Hidrografi kategorisi ile ilgili bilgiler .....                                           | 26    |
| Çizelge 3.3 Baraj/bent ile ilgili kodlama .....                                                       | 27    |
| Çizelge 3.4 AAP'ler arasında akış ilişkileri .....                                                    | 41    |
| Çizelge 3.5 AAP'ler arasında akış ilişkileri .....                                                    | 42    |
| Çizelge 3.6 Dallanan AAP'ler arasında akış ilişkileri .....                                           | 43    |
| Çizelge 3.7 Akış olmayan bağlantı ve akış ilişkileri.....                                             | 43    |
| Çizelge 3.8 AAP ve AKP'ler boyunca akış ilişkilerinin sıralanması .....                               | 44    |
| Çizelge 3.9 Akış ilişkileri kullanarak kodlanan akarsu düzey değerlerinin farkları .....              | 47    |
| Çizelge 4.1 Smirnov'un kritik tablo değerleri .....                                                   | 65    |
| Çizelge 4.2 Farklı anlamlılık düzeylerinde serbestlik derecesine göre kritik tablo<br>değerleri ..... | 66    |
| Çizelge 4.3 Kontenjans tablosu .....                                                                  | 69    |
| Çizelge 4.4 Gözlenen değerler ve beklenen değerler .....                                              | 70    |
| Çizelge 4.5 Akarsulara ait özniteliklerden oluşan girdi vektörleri.....                               | 83    |
| Çizelge 4.6 Rastgele belirlenen 9 adet referans (ağırlık) vektörü .....                               | 83    |
| Çizelge 4.7 Birinci vektörün nöronlara olan Öklid uzaklıkları .....                                   | 84    |
| Çizelge 4.8 Kazanan nöron ve komşularının ağırlıklarının değişimi .....                               | 85    |
| Çizelge 4.9 İkinci vektörün nöronlara olan Öklid uzaklıkları .....                                    | 85    |
| Çizelge 4.10 Kazanan nöron ve komşularının ağırlıklarının değişimi.....                               | 86    |
| Çizelge 4.11 Üçüncü vektörün nöronlara olan Öklid uzaklıkları .....                                   | 86    |
| Çizelge 4.12 Kazanan nöron ve komşularının ağırlıklarının değişimi.....                               | 87    |
| Çizelge 4.13 Dördüncü vektörün nöronlara olan Öklid uzaklıkları .....                                 | 87    |
| Çizelge 4.14 Kazanan nöron ve komşularının ağırlıklarının değişimi.....                               | 88    |
| Çizelge 4.15 Öklid uzaklıkları verilen dört vektör ve oluşan iki ayrı küme.....                       | 89    |
| Çizelge 5.1 Çalışma bölgesine ait eyelet, havza no ve fizyografik bölüm bilgisi .....                 | 111   |
| Çizelge 5.2 Akarsu ağları ve havzalarına ait jeomorfometrik parametreler .....                        | 111   |
| Çizelge 5.3 Kullanılan özniteliklerin tipleri ve değer aralıkları .....                               | 112   |
| Çizelge 5.4 PT'ye ait istatistikler.....                                                              | 117   |

|              |                                                                                                                                                                                          |     |
|--------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Çizelge 5.5  | SBP'ye ait istatistikler.....                                                                                                                                                            | 118 |
| Çizelge 5.6  | Kolmogorov-Smirnov uygunluk tek örneklem testi.....                                                                                                                                      | 119 |
| Çizelge 5.7  | Özniteliklerin hesaplanan ağırlıkları .....                                                                                                                                              | 125 |
| Çizelge 5.8  | Ceza ve kernel parametrelerine göre 10-kat çapraz doğruluk değerleri..                                                                                                                   | 133 |
| Çizelge 6.1  | Önerilen yöntemler ile türetilen akarsu ağlarına ait ÖÇK değerleri .....                                                                                                                 | 136 |
| Çizelge 6.2  | 1:24,000 ve 1:100,000 ölçeğinde orijinal ve 1:100,000 ölçeğinde türetilen PT ve SBP akarsu ağlarına ait nicel bilgiler ve Töpfer'in Radikal Kanununa göre hesaplanan eşik değerler ..... | 138 |
| Çizelge 6.3. | Orijinal 1:24,000 ve 1:100,000 ölçeğinde PT ve SBP akarsu ağlarına ait akarsu tipi ile ilgili nicel bilgiler .....                                                                       | 139 |
| Çizelge 6.4  | KDH yöntemiyle türetilen 1:100,000 ölçeğinde akarsu ağlarının, orijinal akarsu ağları ile akarsu tipi açısından karşılaştırılması .....                                                  | 141 |
| Çizelge 6.5  | DVM yöntemiyle türetilen 1:100,000 ölçeğinde akarsu ağlarının, orijinal akarsu ağları ile akarsu tipi açısından karşılaştırılması .....                                                  | 141 |
| Çizelge 6.6  | Orijinal 1:24,000 ve 1:100,000 ölçeğinde PT ve SBP akarsu ağlarına ait akarsu düzeyi ile ilgili nicel bilgiler .....                                                                     | 144 |
| Çizelge 6.7  | KDH yöntemiyle türetilen 1:100,000 ölçeğinde akarsu ağlarının, orijinal akarsu ağları ile akarsu düzeyi açısından karşılaştırılması .....                                                | 144 |
| Çizelge 6.8  | DVM yöntemiyle türetilen 1:100,000 ölçeğinde akarsu ağlarının, orijinal akarsu ağları ile akarsu düzeyi açısından karşılaştırılması .....                                                | 145 |

---

**KARTOGRAFİK GENELLEŞTİRMEDE SEÇME/ELEME İŞLEMİ İÇİN YAPAY  
ZEKA YÖNTEMLERİNİN AKARSU AĞLARINA UYGULANABİLİRLİĞİ**

Alper ŞEN

Harita Mühendisliği Anabilim Dalı

Doktora Tezi

Tez Danışmanı: Doç. Dr. Trkay GKGZ

Bu alıřmada, akarsu aęlarının genelleřtirmesinde seme/eleme iřlemi iin Kendini Dzenleyen Haritalar (KDH) yntemi kullanılarak yeni bir yntem geliřtirilmiřtir ve Destek Vektr Makineleri (DVM) yntemiyle karřılařtırılmıřtır. KDH ve DVM’de girdi verisi olarak kullanılacak en uygun znitelikler arařtırılmıřtır. znitelikler, khi-kare ( $\chi^2$ ) baęımsızlık testinde belirlenen iliřkilere gre aęrılıklandırılmıřtır. Radikal Kanun, seilecek nesne sayısının belirlenmesinde kullanılmıřtır. KDH’de kmelerin seimi iin kademeli bir yaklařım geliřtirilmiřtir. Uygulamada, 1:100,000 lekli akarsu aęlarını tretmek iin United States Geological Survey Ulusal Hidrografi Veri Seti’nde bulunan 1:24,000 lekli ve farklı akarsu aęı desenlerine (aęası, kafes ve dikdrtgen) sahip iki akarsu aęı kullanılmıřtır. Tretilen akarsu aęları, 1:100,000 lekli orijinal akarsu aęlarına hem nitelik hem de grsel aıdan olduka yakındır. KDH yntemiyle her alt havzada akarsu yoęunluęu ve akarsu aęı deseni byk lde korunmuřtur. Srekli ve semantik olarak doęru aęlar elde edilmiřtir.

**Anahtar Kelimeler:** Kartografik genelleřtirme, akarsu aęı, kendini dzenleyen haritalar, destek vektr makineleri, yapay zeka

## ABSTRACT

---

# THE APPLICABILITY OF ARTIFICIAL INTELLIGENCE METHODS FOR THE SELECTION/ELIMINATION PROCESS TO THE STREAM NETWORKS IN CARTOGRAPHIC GENERALIZATION

Alper ŞEN

Department of Geomatic Engineering

Ph.D. Thesis

Advisor: Assoc. Prof. Dr. Türkay GÖKGÖZ

In this study, a new method was developed using Self Organizing Maps (SOM) method for the selection/elimination process in the generalization of stream networks and it was compared to Support Vector Machines (SVM) method. The most suitable attributes to be used as input to the SOM and SVM were explored. The attributes were weighted in accordance with the associations determined in chi-square ( $\chi^2$ ) independence test. The Radical Law was used in the determination of the number of features should be selected. An incremental approach was developed for the selection of the clusters in the SOM. In the case study, two different stream networks which have different drainage patterns (dendritic, trellis and rectangular) at 1:24,000-scale in the National Hydrography Dataset of United States Geological Survey were used in order to derive the stream networks at 1:100,000-scale. Derived stream networks are quite close to the original stream networks at 1:100,000-scale from both quantity and visual aspects. Stream density and drainage pattern were mostly maintained in each subbasin by SOM method. Continuous and semantically correct networks were obtained.

**Keywords:** Cartographic generalization, stream network, self organizing maps, support vector machines, artificial intelligence

---

YILDIZ TECHNICAL UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

#### 1.1 Literatür Özeti

Coğrafi Bilgi Sistemi (CBS) uygulamalarında, aynı veri tabanına ait veriler, yer bilimleri, doğal kaynak yönetimi, çevre koruma, iklim, orman işleri, savunma, ulaşım, turizm, istatistik, kentsel ve kırsal planlama gibi yeryüzü ile ilgili birçok uygulama ve farklı disiplinlerdeki kullanıcılar tarafından farklı amaç ve ölçekte kullanılmaktadır. Günümüzde tek bir mekansal veri tabanından birçok modelin elde edilmesi ve farklı ayrıntı düzeylerinde analiz, değerlendirme ve sunumu, CBS için büyük önem arz eder.

Yerel, ulusal, bölgesel/çok uluslu ve küresel düzeyde mekansal veri altyapısı projelerine ciddi yatırımlar yapılmaktadır. Bu son derece büyük mekansal veri havuzlarının oluşturulması, bakımı ve kullanımını olanaklı hale getirmek için birçok metodolojinin ortaya konması gerekmektedir. Veri modelleme, veri yönetimi ve veri dağıtımı gibi yöntemlerin yanında otomatik mekansal veri genelleştirme teknikleri bu bağlamda son derece önemlidir [1].

Gelişmiş ülkelerde ulusal harita kurumlarının temel görevi, harita ve coğrafi bilgi üretiminde kullanılan mekansal veri tabanlarının kurulmasıdır. Veri tabanlarının bakımı ve güncellenmesi zorunlu ve henüz standart çözümü olmayan bir problemdir. CBS’de farklı gösterimler, farklı ayrıntı düzeylerinde saklanmaktadır. Çoklu gösterimler, mekansal verinin gösterimi ve depolanması konusunda yaşanan çeşitli problemler sonucu geliştirilen modelin temelini oluşturan ve tek bir mekansal veri tabanının farklı amaç ve ölçek/çözünürlükteki gösterimleridir. Genelleştirme yöntemleri kullanılarak,

çoklu gösterim veri tabanının gösterim düzeylerinden, farklı türde haritalar üretme yeteneği, sisteme esneklik kazandırmaktadır [2].

[3], [4], [5] ve [6]'ya göre bir işlem olarak genelleştirme, hedef ölçeğe ve amaca göre, bazı coğrafi verilere ya da coğrafi verilerin grafik gösterimlerine uygulanan işlemler olarak kabul edilmektedir. Başka bir ifade ile genelleştirme; ayrıntılı mekansal veri kaynağından ya da setinden, semantik ve geometrik dönüşümlerle istenen özelliklere uygun, sayısal ya da işaretli olarak kodlanmış, daha az ayrıntıya sahip bir veri seti türetme işlemi olarak tanımlanabilir [1].

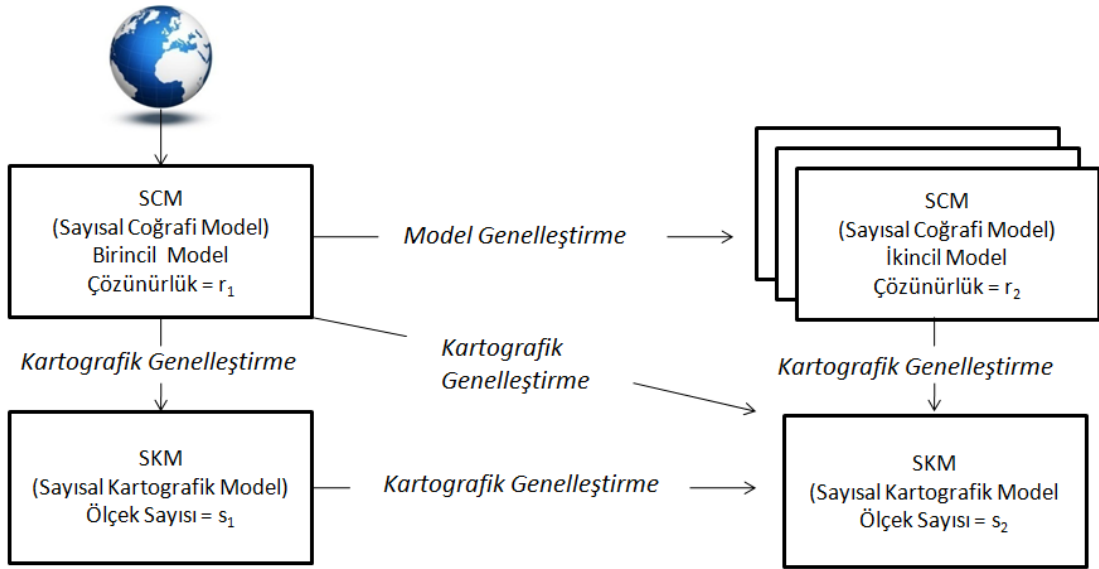
[7]'ye göre genelleştirme, tamamlayıcı terimi olan özet gösterim (abstraction) ile yakından ilgilidir. Özet gösterim, genel ve önemli bileşenleri korumak amacıyla somut gerçeklikten özel ya da gereksiz ayrıntıların elenmesi anlamına gelir [1].

[8]'e göre genelleştirme, mekansal veri seti hacminin önemli yapılar korunarak azaltılması için kullanılan bir işlemdir. Genelleştirme işlemi; hedef ölçeğe, uygulamaya, mekansal duruma ve içeriğe bağlıdır.

Genelleştirme, CBS'nin doğasından kaynaklanan nedenlerle çoğu uzman tarafından model genelleştirmesi ve kartografik genelleştirme olarak ikiye ayrılmaktadır. Model genelleştirmesinde hedeflenen çözünürlükte en uygun veri yapısına ve yoğunluğuna ulaşılması amaçlanırken, kartografik genelleştirmede, model genelleştirmesi yoluyla elde edilen bir sayısal modelden, hedeflenen ölçekte en uygun sayısal görsel modelin elde edilmesine çalışılmaktadır. Kartografik genelleştirmenin sonucu görselleştirilmiş bir kartografik ürün iken, model genelleştirmesinde görselleştirme ile uğraşılmaz [9]. Model genelleştirme, veri analizi için orijinal veri tabanının alt kümesini elde eden bir filtreleme işlemidir.

Model ve kartografik genelleştirme, Sayısal Coğrafi Modelden (SCM) Sayısal Kartografik Model (SKM) elde etmek için kullanılmaktadır. SCM ve SKM mekansal veri tabanının bileşenleridir. SCM, arazinin kapsamlı bir modelini oluşturur ve genellikle temel topografik ölçekli harita formundadır. Model genelleştirme ile birincil SCM'den türetilen veri tabanı, gerçekliğin özel bir amaca yönelik ikincil modelini oluşturur. Hem birincil hem de ikincil modellere uygulanabilen kartografik genelleştirme işlemleri ile

SKM elde edilebilmektedir (Şekil 1.1). CBS, analiz ve görselleştirme işlemlerini yerine getirmek için model ve kartografik genelleştirmeye gereksinim duymaktadır [10].



Şekil 1.1 Model ve kartografik genelleştirme [10]

Literatürde genelleştirme işlemleri farklı biçimlerde sınıflandırılmaktadır. [11] tarafından, genelleştirme işlemini daha anlaşılır hale getiren ilk modellerden biri geliştirilmiştir. İşlem, iki ana aşamaya ayrılmaktadır: (1) Seçim (ön işlem aşaması), (2) Nesneye geometrik ve istatistiksel dönüşümlerin uygulandığı genelleştirme işlemleridir. Seçim aşaması nesnelerin veri tabanında tutulması veya elenmesi işlemlerini içermektedir. Genelleştirme aşaması ise basitleştirme, sınıflandırma ve işaretleştirme işlemlerini içermektedir.

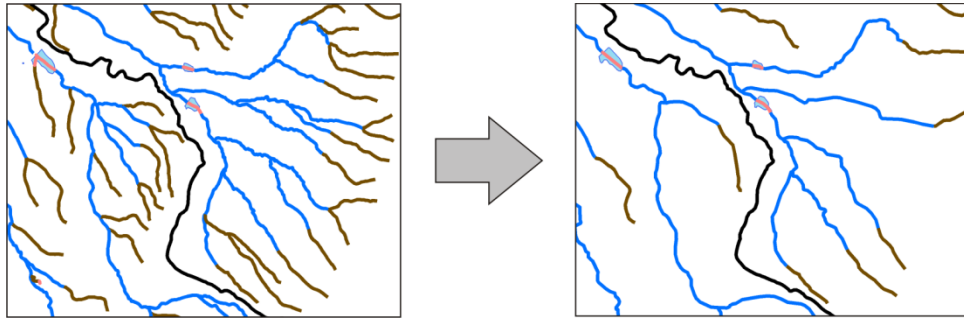
[12] tarafından, genelleştirme işleminin kavramsal modeli oluşturulmuştur. Kavramsal model için üç önemli bileşen tanımlanmaktadır: (1) Teorik amaç veya *niçin genelleştirilir?* (2) Kartometrik değerlendirme veya *ne zaman genelleştirilir?* (3) Mekansal ve öznel dönüşümü veya *nasıl genelleştirilir?*

Genelleştirme işlemleri; seçme/eleme, semantik ve geometrik dönüşüm işlemlerinden oluşmaktadır. Semantik dönüşüm işlemleri; sınıflandırma (classification) ve işaretleştirmedir (symbolization). Geometrik dönüşüm işlemleri ise basitleştirme (simplification), yumuşatma (smoothing), nokta birleştirme (aggregation), çizgi birleştirme (merging), alan birleştirme (amalgamation), dönüşüm (collapse), arıtma ya da seçme (refinement), abartma (exaggeration), iyileştirme (enhancement) ve

ötelemedir (displacement) [13]. Veri tabanının özet gösterimi ile ilgili genelleştirme işlemleri ise model genelleştirme başlığı altında toplanmaktadır.

[14] tarafından, genelleştirme işlemi üç temel sınıfa ayrılmaktadır: (1) Nesnelerin sayısını azaltan işlemler: Seçme, gruplama ve birleştirme, (2) Geometrik indirgeme işlemleri: Birleştirme, dönüştürme ve basitleştirme, (3) Öznitelik basitleştirme (sınıflandırma işlemleri) [15].

Seçme/eleme, hedef harita ölçeği ve amaca göre haritadaki detayları silme veya bırakma işlemidir. Genel prensip, daha küçük ölçekli haritaya giren alanın genel özelliklerini ifade etmeye devam eden detayların kalması, daha önemsiz olanların elenmesidir [16]. Seçme/eleme, nesne sayısında yapılan azaltma işlemidir. Seçme/eleme işlemi, genellikle geometrik ve semantik dönüşümlerin öncesinde, veri tabanındaki nesnelerin sayısında yapılan bir azaltma olarak düşünülmektedir [17]. Hidrografik nesnelerin seçimine dair bir örnek Şekil 1.2’de gösterilmiştir.



Şekil 1.2 Hidrografik nesnelerin seçimi

Seçme/eleme işlemi, kartografik genelleştirmede öncelikli öneme sahiptir. Bir haritadaki tüm işaretlerin okunaklı olabilmesi için belirli alanları işgal etmeleri gerekmektedir. Bilinçli bir seçim işleminde, daha küçük bir ölçekte tüm harita detayından hangi nesnelerin gösterilmesi gerektiğine karar verilmelidir. Harita ölçeğine bağlı olarak harita üzerindeki nesne sayısının değişimine ilişkin basılı harita serilerinin karşılaştırıldığı deneysel çalışmalar, Töpfer’in “Radikal Kanun”u ya da “Seçme İlkesi” olarak adlandırılan bazı ilkeleri ortaya koymuştur [1]. “Radikal Kanun” ya da “Seçme İlkesi”, daha küçük bir ölçek için seçilebilecek nesne sayısını, başka bir deyişle, genelleştirmenin niceliksel yönünü göstermektedir. Ancak, hangi nesnenin seçilmesi gerektiğini göstermemektedir.

Çeşitli ülkelerin atlas çalışmalarında Radikal Kanun'un, kaynak materyalden daha küçük ölçekli haritaların türetilmesi için büyük bir potansiyel değer olduğu kanıtlanmasına rağmen, nicel faktörlerin yanında diğer nitel faktörler; "tecrübe", "coğrafi bilgi" ve "haritanın amacı" seçme eleme işleminde kullanılmalıdır [18]. Halen, "Radikal Kanun" ya da "Seçme İlkesi", nesnelerin seçilmesinde en yaygın kullanılan niceliksel ölçüttür [10].

Seçme ilkesi, en temel haliyle [18];

$$n_f = n_a \sqrt{m_a/m_f} \quad (1.1)$$

eşitliği ile hesaplanabilir.

Burada,  $n_f$  hedef nesne sayısını,  $n_a$  kaynak nesne sayısını,  $m_a$  kaynak ölçek sayısını,  $m_f$  hedef ölçek sayısını göstermektedir.

Bu seçme/eleme ilkesi, kaynak ve hedef ölçekteki detayların benzer geometrik işaretlerle sunulduğu topografik harita genelleştirmesinde kullanılabilir. Buna karşın, kaynak ölçekteki bazı nesnelerin hedef ölçekte değişik işaretlerle gösterilebildiği ve amaca bağlı olarak vurgulanması gereken içeriğin değişebildiği küçük ölçekli haritalarda (atlas haritaları vb.), bu ölçütün kullanılması uygun değildir. Bu durumda, seçme/eleme ölçütü, kartografik işaretin türüne (nokta, çizgi ve alan) ve işaretin abartma miktarına göre (1.2) eşitliğindeki gibi düzenlenebilir [17], [18].

$$n_f = n_a c_b c_z \sqrt{m_a/m_f} \quad (1.2)$$

$c_b$ : Harita işareti için kullanılan abartma sabiti

$c_z$ : Kartografik işaretin türü

[18] tarafından, kaynak ve türetilen haritalarda aynı çizgi kalınlıkları kullanıldığında (1.2) eşitliği, (1.3) eşitliğindeki gibi düzenlenmiştir. Kartografik işaretin türüne göre  $x$  üstel değeri değişmektedir [17], [19].

$$n_f = n_a \sqrt{(m_a/m_f)^x} \quad (1.3)$$

$x$ , işaretin türüne göre aşağıdaki değerleri almaktadır.

- Nokta işaretler için  $x = 1$

- Çizgi işaretler için  $x = 2$
- Alan işaretler için  $x = 3$

Bazı harita nesneleri ağ yapısı göstermekte (yollar, demiryolları, nehirler vb.) ve genelleştirme işlemi süresince özel bir dikkat gerektirmektedir. Akarsular ve yollar, alansal ve çizgisel özellik göstermektedirler. Akarsu genişliği çok fazla değişiklik gösterdiğinden, kartografik model oluşturma işlemi daha karmaşıktır. Göl ve haliç gibi nesneler de modeli daha fazla karmaşıktırır [15].

[20]'ye göre, hidrografik genelleştirme ve yükseklik eğrilerinin genelleştirmesi, model genelleştirme ve kartografik genelleştirme işlemlerinin önemli bir bölümünü oluşturmaktadır. [21]'e göre, topografik haritalar genelleştirilirken nesne sınıflarının işlenmesi şu sırayı izlemelidir [1], [22]:

- Su ağı ve su alanları,
- Demiryolları,
- Yollar,
- Binalar,
- Sınırlar,
- Topografya ve
- Bitki örtüsü

Hidrografya sınıfı, bir harita ya da mekansal veri tabanı için en önemli nesne sınıflardan biridir. Hidrografik genelleştirmenin amacı, topografik haritaların üretimi için gerekli olan hidrografya verilerinin, uygun içerikte ve nitelikte hazırlanmasıdır. Hidrografik genelleştirme, desen çeşitliliği, coğrafi ve hidrografik anlam, bağlantı özelliği, hiyerarşik düzen, akış ilişkisi ve semantik özellikler düşünüldüğünde çok karmaşık, çok araştırılan ve tam olarak çözülememiş bir konudur. Anamlı bir akarsu ağı, topolojik olarak doğru, model akışını sağlayan ve su ile ilgili özelliklerin birbirinden ayırt edilebildiği bir ağıdır [23]. [24]'e göre, akarsu ağlarının genelleştirmesinde seçme/eleme şu iki soruyu yanıtlayabilmelidir: Kaç adet kol seçilmelidir? Hangi kollar daha önemlidir? [18] tarafından birinci soru yanıtlanabilmiştir. Fakat ikinci sorunun yanıtı net olarak

verilememektedir. Akarsu ağına ait hangi kolların seçilmesi gerektiği akarsuya ait geometrik, topolojik ve semantik özniteliklerin bir arada değerlendirilmesi ile ortaya çıkmaktadır. Akarsuların bir havza alanındaki önemi, farklı düzeylerde düşünülmelidir. Bunlar:

- Makro düzeyde akarsu ağları için desen tanıma (pattern recognition),
- Mezo düzeyde dağılım yoğunluğu ve
- Mikro düzeyde geometrik özelliklerdir.

Akarsuların genelleştirmesi ile ilgili mevcut kurallar, çoğunlukla uzman kartograflara yönelik ve klasik üretim için tanımlanmış kurallardır. Harita Genel Komutanlığı (HGK) hidrografik genelleştirme kurallarının bazıları aşağıdadır [25]:

- Mavi kalıp, paftanın iskeleti olup, ana kalıptır ve diğer bütün detayların mavi kalıba uyma zorunluluğu vardır [26].
- Normal olarak ana dereler muhafaza edilir. Ölçek küçüldükçe ana derenin küçük kolları derece derece elemine edilir [27].
- Genelleştirme yapılırken nehirler, isimli dereler, uzun arklar ve arazinin şeklini ifade eden mevsimsel (kuru) derelerin seçimine öncelik verilmelidir [26].
- Pafta dahilinde mevsimsel derelerin hepsi alınmaz, önem derecesine göre bir ayıklama yapmak gerekir. Mevsimsel uzun derelerin hemen hepsi alınır [28].
- Mevsimsel derelerden isimli olanlar ve büyük olanlar tersim edilir [26].
- Küçük olmasına rağmen arazinin topografik yapısını tamamlayan mevsimsel dereler de seçilerek alınır [26].
- Kanaletler yoğunluklarına göre, küçükler ayıklanarak 1:25,000 ölçeğindeki ölçü ve şekline göre tersim edilir [26].
- Büyüklük kriteri genişlik veya uzunluk veya her ikisinin kombinasyonu olabilir [27].

Ülkemizde yapılan genelleştirme ile ilgili doktora çalışmaları, [1], [16], [17], [25], [29], [30], [31], [32], [33] ve [34] tarafından gerçekleştirilmiştir. Hidrografik genelleştirme ile ilgili olan doktora çalışmaları, [16] ve [25] tarafından yapılmıştır.

[16] tarafından, Türkiye’de genelleştirme konusunda ilk doktora çalışması yapılmıştır. 1:25,000 ölçekli haritalardan, klasik genelleştirme yoluyla, 1:50,000 ölçekli harita üretimi prensip ve kurallarından yararlanılarak yarı-otomatik genelleştirme yapılmıştır. Detayların tek anlamlı tanımı için katman-renk bileşiminden oluşturulan detay kodlama kütüğü oluşturulmuştur. Klasik genelleştirme kuralları göz önünde tutularak, 470 adet detay tek tek ele alınmış ve nasıl genelleştirileceği tasarlanmıştır. Böylece uygulanacak olan genelleştirme fonksiyonu belirlenmiştir. [16]’a göre hidrografik nesneler için genelleştirme kurallarının bir kısmı Çizelge 1.1’de verilmiştir (L: Çizgi, A: Alan).

Çizelge 1.1 Bazı hidrografik nesneler için genelleştirme kuralları [16]

| İşaret | Detay Adı                  | Katman | Renk | Tip | Kullanılacak İşlem                     |
|--------|----------------------------|--------|------|-----|----------------------------------------|
| 313    | Mevsimsel dere             | 30     | 15   | L   | Seçim - Seyreletme - Kartografik yorum |
| 506    | Daimi dere (geniş yataklı) | 30     | 16   | A   | Seçim - Kartografik yorum              |
| 312    | Daimi dere (dar yataklı)   | 30     | 17   | L   | Seçim - Kartografik yorum              |
| 568    | Değişken akarsu yatağı     | 30     | 19   | A   | Seçim - Kartografik yorum              |
| 308    | Ark (daimi)                | 30     | 8    | L   | Seçim - Kısmi seçim (>500 m.)          |
| 309    | Ark (mevsimsel)            | 30     | 14   | L   | Seçim - Kısmi seçim (>500 m.)          |
| 310    | Kanal                      | 30     | 21   | L   | Seçim - Kısmi seçim (>500 m.)          |

[25] tarafından, topografik haritaların üretiminde akarsu ve su iletim hatlarının 1:25,000-1:100,000 ölçek aralığında genelleştirilmesi için seçim ağırlıkları hesaplanarak, Kritik Yörünge yöntemine (Critical Path method) dayalı bir seçim algoritması geliştirilmiştir. Akarsuların seçim işleminde, birçok parametre ve ölçüt dikkate alınmıştır. Bu parametre ve ölçütlerin bir kısmı Çizelge 1.2 ve 1.3’te verilmiştir. Her bir

akarsu için, (1.4) eşitliğinde verilen ağırlık hesaplama eşitliği kullanılarak ağırlıklar hesaplanmıştır.

$$M = (M_1 \times M_2 + M_3) \times L \quad (1.4)$$

Burada,  $M$  toplam ağırlık,  $M_1$  birinci ağırlık katsayısı,  $M_2$  ikinci ağırlık katsayısı,  $M_3$  bağıllık ve akarsu ağ yapısını güçlendirici ağırlık katsayısı ve  $L$  detay uzunluğudur.

Çizelge 1.2 Detay önem hiyerarşisi [25]

| Detay Adı                      | Katsayı ( $M_1$ )         | Katsayı ( $M_2$ ) |
|--------------------------------|---------------------------|-------------------|
| Dere önemli (alan)             | Alan/Toplam Uzunluk       | 1.5               |
| Dere önemli                    | 30                        | 1.5               |
| Dere                           | 20                        | 1.5               |
| Dere (mevsimsel) önemli (alan) | Alan/(Toplam uzunluk x 2) | 1.5               |
| Dere (mevsimsel) önemli        | 15                        | 1.5               |
| Dere (mevsimsel)               | 10                        | 1.5               |

Çizelge 1.3 Bağıllık katsayısı [25]

| Durum                                        | İlave Katsayı ( $M_3$ ) |
|----------------------------------------------|-------------------------|
| Durgun sulara (deniz, göl) dökülen akarsular | 4                       |
| Su iletim hatlarına dökülen akarsular        | 5                       |
| Seçilmiş akarsulara dökülen akarsular        | 3                       |

[25]'e ait çalışmada akarsu seçiminde aşağıda belirtilen durumlar dikkate alınmıştır:

- Detay önem hiyerarşisi: Detay önem hiyerarşisi Çizelge 1.2'de verilmektedir. Örneğin daimi dere, mevsimsel dereye göre daha önemlidir.
- Detayın özel isminin olması durumu: Topografik harita üretiminde, ismi olan detay, ismi olmayan detaya göre daha önemlidir. Bu durum, Çizelge 1.2'de belirtilen  $M_2$  katsayısı ile ifade edilmektedir.

- Akarsu genişliği: Geniş olan bir dere, dar olan dereye göre daha önemlidir. Bu durum, akarsu alan detayları için Çizelge 1.2’de gösterilen  $M_1$  katsayısının hesaplanmasında dikkate alınmaktadır.
- Akarsu uzunluğu: Uzun olan bir dere, kısa olan bir dereye göre daha önemlidir. Bu durum,  $M$  ağırlık katsayısının hesaplanmasında  $L$  uzunluğu ile dikkate alınmaktadır.
- Akarsu ağ yapısını koruma: Bir akarsuyun seçiminde, akarsu ağ yapısı içindeki bağlantı görevi de dikkate alınmalıdır. Bu durum akarsuların seçimi algoritmasında Kritik Yörünge yönteminin kullanımı ile dikkate alınmaktadır.
- Akarsu ağ yapısını güçlendirme: Akarsuların iterasyonlu seçim işleminde, seçilmiş olan akarsu ağına doğrudan dökülen akarsulara ilave ağırlık verilerek ve seçim iterasyonuna devam edilerek sağlanır. Bunun için Çizelge 1.3’te belirtilen  $M_3$  ilave ağırlık katsayısı kullanılmaktadır.
- Bağlılık ilişkilerinin dikkate alınması: Topografik harita üretiminde, bir göle dökülen dere daha önemlidir. Su iletim hatları ile bağlantıları bulunan dere daha önemlidir. Dolayısıyla bu durum, durgun sulara ve su iletim hatlarına dökülen akarsulara Çizelge 1.3’te belirtilen  $M_3$  ilave ağırlık katsayısı verilerek dikkate alınmaktadır.

Hidrografik genelleştirmede seçme/eleme ile ilgili yapılan diğer çalışmalar şunlardır. [35] tarafından, akarsu kollarının hiyerarşik olarak sınıflandırıldığı bir yöntem geliştirilmiştir. [36] tarafından, bu yöntemeye dayalı yeni bir yöntem geliştirilmiştir. [37] tarafından, farklı bir hiyerarşik sıralama yöntemi önerilmiştir. [38] tarafından, hidrografik genelleştirme için ağırlıklandırılmış bir akarsu ağı önerilmiştir. [39] tarafından, akarsu seçimi için hiyerarşik filtreler kullanılmıştır. [40] tarafından, Gestalt teorisine dayalı iyi devamlılık (good continuation) prensibi, yol ve akarsu ağlarına uygulanmıştır. Gestalt teorisine dayalı iyi devamlılık prensibi, birbirini takip eden yol veya akarsu parçalarından devamlılık gösteren hat (stroke) dokusu meydana getirenlerin seçilmesi prensibidir. Devamlılık gösteren hat olarak adlandırılan yol veya akarsu dokuları, birbirini izleyen ve keskin yön değişimlerinin olmadığı parçalardan oluşmaktadır. [41] tarafından, 1:100,000 ölçeğinde türetilen akarsu ağları için geometrik, topolojik ve semantik öznitelikler ve kurallar belirlenerek seçme/eleme yapılmıştır. [24] tarafından, havza alanları; akarsu kollarının uzunluğu ve kollar

arasındaki mesafe hakkında bilgi verdiği için, tüm alt havzaların alanları hesaplanıp havza hiyerarşisi ile birlikte kullanılarak eşik değerlere göre akarsular elenmiştir. [42] tarafından, akarsu ağının model genelleştirmesi için iyi devamlılık prensibi kullanılmıştır. [42]'nin çalışmasında, iyi devamlılık prensibine göre belirlenen hatlar hiyerarşik sıra esas alınarak seçilmiştir. [43] tarafından, 1:24,000-1:100,000 ölçek aralığında akarsu seçimi için uzunluk ve akarsu yoğunluk değerleri kullanılmıştır. [44] tarafından yapılan çalışmada ise hidrografik genelleştirmede seçme/eleme için ilk kez Kendini Düzenleyen Haritalar (KDH) (Self Organizing Maps) ve k-ortalamalar yöntemleri karşılaştırılmıştır.

Hidrografik nesneler haricindeki nesnelere uygulanan seçme/eleme işlemi ile ilgili çalışmalar ise şunlardır: [45] tarafından, İsrail haritalarında Radikal Kanun ve ağırlık değerleri kullanılarak şehirler seçilmiştir. [46] tarafından, Radikal Kanun İngiltere'de şehirlerin seçiminde kullanılmıştır. [47] tarafından, tematik haritalar için seçim işlemi kullanılmıştır. [48] tarafından, yol orta eksenleri kullanılarak yollar seçilmiştir. [49] ve [50] tarafından, sokaklarla çevrili kentsel bir blokta binaların sayısını azaltmak için iç kısıtlamalara uygun bir eşik değer kullanılmıştır. Elenecek nesnenin belirlenmesi için bir değer fonksiyonu (cost function), bir dizi kısıtlamayı bir araya getirmek için kullanılmaktadır. Yöntem, yoğun bölgelerde semantik açıdan bir önemi olmayan ve yoldan uzakta olan küçük binaları elemektedir. [51] tarafından, yol ağı seçiminde hatlar kullanılarak iyi devamlılık prensibi kullanılmıştır. [30] tarafından yapılan çalışmada, Hannover Üniversitesi Kartografya ve Jeoenformatik Enstitüsü'nde geliştirilen Change yazılımı kullanılarak, 1:1000-1:25,000 ölçek aralığında yol ağı genelleştirmesinde seçme/eleme konusu ele alınmıştır. [52] tarafından, KDH algoritması önemli yolların seçimi için kullanılmıştır. KDH algoritmasında kullanılan girdi verisi olarak, çizge (graph) temelli bir yaklaşım kullanılarak, bağlılık bakımından merkezlik (degree centrality), yakınlık bakımından merkezlik (closeness centrality) ve arasında olma bakımından merkezlik (betweenness centrality) değerleri tanımlanmıştır. [52] tarafından, yol ağında seçme/eleme için uygulanan KDH yönteminde, kümelerin karşılaştırılması yoluyla en uygun haritanın belirlenmesi kullanıcıya bırakılmıştır. [1] tarafından, 1:25,000-1:50,000 ölçek aralığında, nesne yönelimli LAMPS2 yazılımı kullanılarak, bina ve yerleşim alanları otomatik genelleştirmesi için gerekli yol ağı genelleştirmesi adına hedef ölçekte yeterli

büyükölçeğe sahip olmayan ve yerleşim adalarını çevreleyen bazı yollar, belirlenen uzunluk ölçütüne göre interaktif ve manuel olarak elenmiştir. [31] tarafından, kaynak olarak 1:25,000 ölçekli veri setleri kullanılarak, 1:100,000 ölçekli haritalar için yol verilerinin otomatik olarak iyi devamlılık prensibine göre seçilmesi amaçlanmıştır. Belirlenen güzergahlarda seçme/eleme işlemi; yerleşim dışı bölgeler için örümcek diyagramlar, yerleşim bölgeleri için ise uzunluk ve bağlantı sayıları ölçüt alınarak ve yine kullanıcı tarafından girilen eşik değerlerle karşılaştırılarak gerçekleştirilmiştir. [32] tarafından yapılan çalışmada, araç navigasyon haritalarında kullanıcı tarafından belirlenen güzergaha bağlı olarak seçilecek yolların belirlenmesi ve en uygun şekilde sunulması üzerine çalışılmıştır. Seçilecek yollar, iyi devamlılık prensibine dayalı olarak, yolların sapma açılarına göre belirlenmektedir. [17] ve [53] tarafından, ada temelli bir yol ağı seçme/eleme yöntemi geliştirilmiştir. 1:25,000 orijinal, 1:50,000 ve 1:100,000 ise hedef ölçekler olmak üzere, Java programlama dili kullanılarak Radius Clarity ve nesne yönelimli Gothic veri tabanında, geliştirilen koşulları sağlayacak şekilde adalar birleştirilmiş ve böylece yollar elenmiştir. [33] tarafından, 1:25,000 ölçekli orijinal haritalardan, 1:50,000 ve 1:100,000 ölçekli haritalar üretilirken bina nesnelerinde seçme/eleme işlemi gerçekleştirilmiştir. [34] tarafından, haritaların içerdikleri mekansal bilgi miktarının nicel bir ölçüt kullanılarak belirlenmesine yönelik istatistiksel, topolojik ve metrik yöntemlerin geçerliliği araştırılmıştır.

Çevresel değişim ve ekonomik gelişme ile beraber günümüzde dev boyutlarda coğrafi bilgi üretilmektedir. Teknolojinin gelişmesi ile veri hacmi artmaya devam etmektedir. Veri miktarının artması, çok boyutlu verinin işlenmesini gerektirmektedir. Artan veri miktarı karşısında, geleneksel yöntemler başarısız olmakta veya coğrafi yapıyı, ilişkileri ve anlamı araştırırken bir takım engeller ortaya çıkartmaktadır [54].

Genelleştirme probleminin karmaşık yapısı; coğrafi nesnelerin, özniteliklerin, araçların ve amaçların çeşitliliği sebebiyle coğrafi bilginin formüller ile ifade edilememesinden doğmaktadır. Ayrıntı düzeylerinde nesnelerin geometrik, topolojik ve semantik öznitelikleri arasında bulunan bu karmaşık ilişki ve mekansal verinin çok boyutluluğu, veri tabanı analizlerini, veri madenciliği tekniklerini ve yapay zeka yöntemlerini gerektirmektedir.

Genelleştirme yöntemleri ve işlemleri, bilgi teknolojisindeki gelişmelerden çok fazla etkilenmektedir. Modern veri tabanı tasarımlarında yapay zeka algoritmalarını kullanan teknoloji ve yöntemler, kartografik ve model genelleştirme kapasitesini de artırmaktadır [55]. [16]'ya göre bu durum şöyle ifade edilmiştir: “Önümüzdeki yıllarda yazılım ve donanımlardaki gelişmeye paralel olarak yapay zeka konusundaki gelişmelerin otomatik genelleştirmeye çok büyük oranda çeşitlilik katacağı beklenmelidir.”

Bu bağlamda, yapay zeka yöntemlerinin kartografik genelleştirme açısından değerlendirilmesi çok büyük önem arz etmekte olup, günümüzde bu yöntemleri kartografik genelleştirmede kullanan çok az sayıda çalışma bulunmaktadır.

## **1.2 Tezin Amacı**

1:24,000 ölçeğinden 1:100,000 ölçeğine geçişte akarsu ağının model genelleştirme işleminin tam otomatik gerçekleştirilebilmesi için yapay zeka alanında kullanılan yöntemlere dayanan yeni bir yöntem ve metodolojinin geliştirilmesidir.

## **1.3 Hipotez**

Hidrografik genelleştirmede seçme/eleme, kartografyanın henüz tam olarak çözülememiş problemlerinden biridir. Bu bağlamda, yapay zeka alanında kullanılan yöntemlerden (KDH ve DVM) yararlanılabilir.

KDH yöntemi, nesneleri kümelemektedir ancak kümelerin seçimi [52]'ye göre kullanıcıya bırakılmaktadır. Oysaki akarsu ağının model genelleştirmesinde hangi nesnelerin seçileceği en önemli sorulardan biridir. Bu bağlamda, KDH yöntemi bu sorunun cevabını verecek hale getirilebilir (Bölüm 5.3).

Hidrografik nesneler, yol, bina, yükseklik eğrisi gibi nesnelerden farklıdır. Bu bağlamda, bilinen geometrik, topolojik ve semantik özniteliklerden hangilerinin akarsu ağının model genelleştirmesi için uygun olacağı tespit edilebilir (Bölüm 5.2.1).

Akarsu ağının model genelleştirmesinde kullanılacak özniteliklerin ağırlıkları eşit olmamalıdır. Bu bağlamda, öznitelikler uygun bir yöntemle ( $\chi^2$  bağımsızlık testi) ağırlıklandırılabilir (Bölüm 5.2.2).

Akarsu ađının model genelleřtirmesinin hedeflerinden biri de akarsu ađı deseninin korunmasıdır. Bu bađlamda, akarsu ađı deseni de g z  n ne alındıđında daha iyi sonu lar elde edilebilir.

Geliřtirilen KDH y ntemi ve DVM y ntemi ile elde edilen sonu lar karřılařtırılarak, k meleme ve sınıflandırmanın  st nl kleri ve zayıflıkları belirlenebilir.

### ÇALIŞMANIN HİDROJEOMORFOLOJİK TEMELLERİ

Havza analizi çalışmalarında, yeryüzü biliminin (earth science) jeomorfoloji ve hidroloji olmak üzere iki önemli dalı vardır. Bu ikisi birleşerek hidrojeomorfoloji adını almaktadır. Jeomorfoloji, yeryüzü şekillerini araştıran; hidroloji ise su ile ilgili çalışmaları kapsayan bilimdir. Hidrojeomorfoloji, hidrolojik sürecin yeryüzü üzerindeki etkilerini araştırmaktadır [56]. Bu bölümde, akarsu ağları ve desenleri açısından, bazı tanımlara, havzaların yapısı ile ilgili jeomorfometrik parametrelere ve akarsu sıralamasına yer verilmiştir.

#### 2.1 Tanımlar

**Akarsu:** Karalar üzerindeki yüzeysel sular, yerçekimi tesiri ile en büyük eğim yönünde belirli bir mecrada toplanarak çizgisel bir akım oluşturur. Akarsu, bu şekilde oluşan doğal su yolları içinde hareket eden sular için kullanılan genel terimdir [57]. Akarsular, binlerce yıldan fazla sürede kararlı akış sağlanıncaya kadar doğal erozyonlar vasıtasıyla oluşmaktadır. Akarsular iklim, hidroloji, jeoloji, toprak ve kara parçası gibi çeşitli fiziksel özelliklerin bir sonucu olarak işlev gösterirler [56].

**Akarsu ağı:** Bir ana akarsu kolu (ana akarsu veya gövde) ile yan kolların tümünün meydana getirdiği şebekeye akarsu ağı (drenaj ağı, kanal ağı) denir [57].

**Kaynak deresi:** Bir akarsuyun başlangıç noktası genel olarak bir veya birden fazla kaynaktır. Bununla beraber başlangıç noktası göl, yer altı suyu veya buzul olan

akarsular da mevcuttur. Bir akarsuyun çıktığı yerden ilk dere ile birleştiği yere kadarki kesimine kaynak deresi veya başlangıç deresi (headwater stream) denir [57].

**Akarsu kavşağı:** İki veya daha fazla akarsuyun birleştikleri yere akarsu kavşağı denir [57].

**Ana akarsu kolu:** Bir akarsu ağı, ana akarsu kolu ve buna karışan yan kollardan oluşur. Akarsu kollarından debisi daha büyük veya boyu daha uzun veya kaynağı daha yüksekte olanı, genellikle ana akarsu kolu olarak isimlendirilir [57].

**Akarsu ayağı:** Daha geniş bir akarsudan ya da bir gölden ayrılan akarsu kesimine ayak (akarsu ayağı, göl ayağı) denir [57].

**Akarsu ağız:** Bir havza bölümünden gelen yüzeysel suların toplanarak havzayı terk ettiği akarsu kesimine çıkış noktası; akarsuyun deniz, göl veya hazne ile birleştiği yere ağız denir [57].

**Memba bölgesi:** Bir akarsu kesitinin kaynak tarafında kalan akarsu bölümüne memba bölgesi denir [57].

**Mansap bölgesi:** Bir akarsu kesitinin ağız tarafında kalan bölümüne ise mansap bölgesi denir [57].

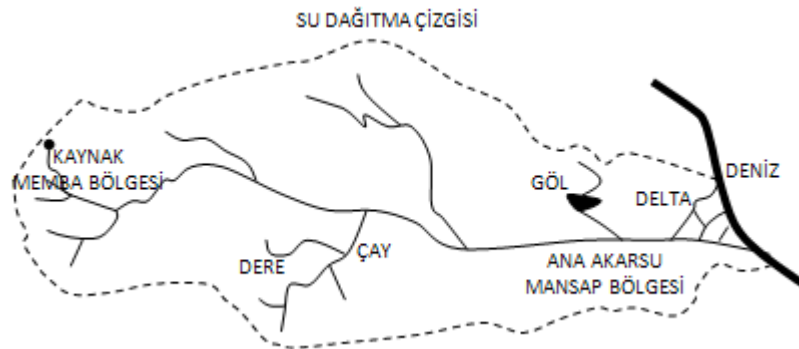
**Karakteristik çizgiler:** Su toplama ve dağıtma çizgileridir. Karakteristik çizgiler, yükseklik eğrileri arasındaki köprüler gibi düşünülebilir. Başka bir ifadeyle yükseklik eğrileri arasındaki jeomorfolojik ilişkinin kartografik işaretleridir. Karakteristik çizgiler, en büyük eğim çizgilerinin birleşerek aldıkları ortak yollar ya da kısaca özel en büyük eğim çizgileri olarak tanımlanabilir. Burada en büyük eğim çizgileri ile kastedilen en büyük eğim doğrultusunda ya da doğrultularında alınan yollardır [29].

**Su dağıtma çizgisi:** İki komşu akarsu havzasını ayıran sırtlara (ridges) su dağıtma çizgisi veya havza sınırı (watershed boundary) denir [56].

**Su toplama çizgisi:** Vadilerin en derin noktalarını birleştiren çizgilere su toplama çizgileri denir. Mevcut olan kaynak veri tabanında, daimi ve mevsimsel (kuru) dere türü detaylardır [25].

**Havza alanı:** Bir akarsuyun sularını toplayan alana akarsu havzası (drenaj havzası, su toplama havzası) denir [57]. Havza alanları büyüklüğüne ve karmaşıklığına göre akarsu havzaları ve alt havza alanları gibi sınıflara ayrılmaktadır. Havza alanı, analiz için daha küçük alt havzalara bölünür. Bu nedenle havza alanları akarsular ile ilişkilidir. Havza alanı, literatürde “basin”, “watershed” ve “catchment” gibi isimlerle adlandırılmaktadır [56].

Akarsu; dere, çay, su, ırmak ve nehir sözcüklerinin hepsinin ortak adıdır. Dilimizde hiç veya önemli bir yan kolu olmayan en küçük akarsulara dere, derelerin birleşmesi ile oluşan akarsulara çay, çay ve derelerin birleşmesi ile oluşan akarsulara ise nehir ismi verilmektedir. Bu sözcükler akarsuyun geçirdiği su miktarına göre tarif edilmediğinden birbirinden kesin sınırlarla ayrılmamıştır. Bu nedenle de birçok dere, çay; birçok çay ise nehir büyüklüğünde olabilmektedir. Dilimizde kullanılan ırmak ve su kelimeleri ise belirsiz bir anlam taşımakta, aralarında açık bir fark olmadığı gibi nehir veya çay büyüklüğünde olabilmektedirler (Kızılırmak, Yeşilirmak, Karasu, Göksu gibi) [57]. Şekil 2.1’de bir akarsu ağının planı gösterilmiştir.



Şekil 2.1 Bir akarsu ağının planı [57]

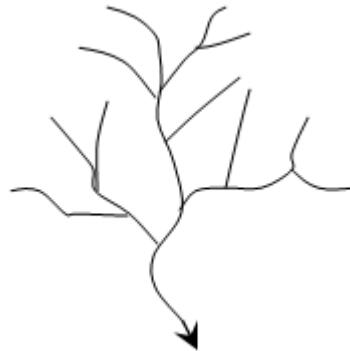
**Sürekli Akışlı Akarsular:** Yatağında yıl boyunca su bulunan akarsular (daimi/sulu dere) [57].

**Periyodik Akışlı Akarsular:** Yatağında yalnız bol yağışlı mevsimlerde su bulunan ve yağışsız zamanlarda su bulunmayan akarsular (mevsimsel/kuru dere) [57].

## 2.2 Akarsu Ağı Desenlerine Göre Akarsu Ağlarının Sınıflandırılması

Doğal akarsu ağı desenleri (drainage pattern), topografik, jeolojik, morfolojik, bitki örtüsü, toprak yapısı ve iklim koşullarına göre çeşitlenmektedir. Akarsular yapı, desen ve ağ bakımından sınıflanabilmektedir. Akarsu ağı deseni daimi ve mevsimsel dereler tarafından meydana getirilen akarsu ağının şeklidir. Akarsu ağı desenleri genel olarak temel (basic) ve değişmiş (modified) olarak ikiye ayrılmaktadır. Temel desen, bir ağın toptan görünümü bakımından diğer desenlerden ayırt edilmesidir. Değişmiş desen ise ağın belli bölgelerinde farklı özellik gösteren ağlardır. Çoğu değişmiş desen, temel desenler arası geçiş teşkil etmektedir. Temel akarsu ağı desenleri; ağaçsı/dal yapılı (dendritic), paralel, kafes (trellis), radyal, merkezci, dairesel, dengesiz ve dikdörtgen yapılı olarak sınıflandırılabilir. Değişmiş desenlere ise, alt ağaçsı (subdendritic), tüysü (pinnate), aşağıya veya dışarıya eğilmiş (recurved) kafes yapılı gibi örnekler verilebilir. Desen değişimleri, genellikle anlaşılması güç iç detaylar ile ilgilidir. İki komşu desen bir bileşik desen de meydana getirebilir. Bir akarsu ağının beklenen deseni, anomali nedeniyle bölgesel, jeolojik ve topografik etmenlerin kontrolünden uzaklaşabilir. Akarsu anomali; deformasyon, diferansiyel çökme veya hidrolojik rejim değişikliği hakkında bilgi sağlayabilir [56], [58].

**Ağaçsı/Dal Yapılı (Dendritic) Desen:** Akarsular, ağacın dallanmasına benzer bir yapı gösterirler. Akarsu kolları değişik yönlerde doğru akan bir düzendedir. Ağaçsı desenler, aşınır (erodible) sediment üzerinde dengelenmiş (adjusted) sistemleri ile yaygın ve tipik bölgelerdir. Ağaçsı havza alanları, tipik olarak armut şekline (pear-shaped) benzer (Şekil 2.2) [56].



Şekil 2.2 Ağaçsı desen [56]

**Paralel Desen:** Akarsu kolları aynı yöne doğru akar ve küçük açılar ile birleşirler. Ağaçsı desenlerin uzatılmış bir biçimidir. Bu desen, sarp bölgelerde görülür (Şekil 2.3) [56].



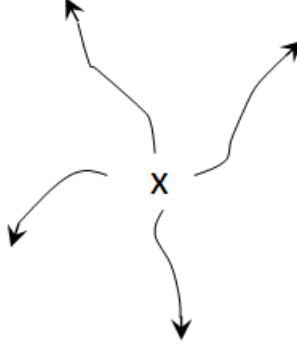
Şekil 2.3 Paralel desen [56]

**Kafes (Trellis) Desen:** Ana akarsu, karşılıklı olarak birbirine kavuşan kolların gövdeye bağlanması ile ana kaya (bedrock) yapısına paralel şekilde akar. Karelere bölünmüş akarsu ağı deseni görünümündedir. Eğimli (tilted), çöküntülü (sedimentary) kayalardan oluşan zayıf düzlemler üzerinde, dirençli (resistant) tabakaları yarıp geçerek aralık (gap) oluştururlar. Kıvrımlı kayalık bölgelerde yaygındırlar (Şekil 2.4) [56].



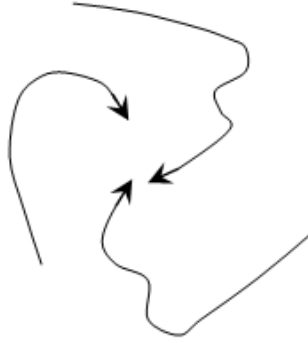
Şekil2.4 Kafes desen [56]

**Radyal Desen:** Merkezi yüksek olan bölgeden dışa doğru ve her yöne doğru akan, dairesel bir yapıda, granit kayalar, tuz domu/tümseği (salt dome), volkanik koniler (volcanic cone) ve diğer lokal arazi çıkıntıları civarında yaygın olarak görülen bir desendir (Şekil 2.5) [56].



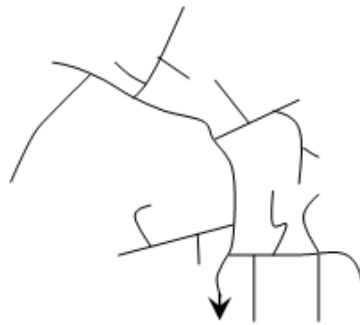
Şekil 2.5 Radyal desen [56]

**Merkezcil (Centripetal) Desen:** Yaygın olmamakla birlikte bazen karstik topografyada görülmektedir. Çöllerde ve tuz gölü veya havzalarında bulunan kesikli akarsulardır. Akarsular bir merkeze doğru akmaktadırlar. Su, her yönden bir alanın merkezine doğru (içe doğru) akar. Dairesel bir yapı göstermektedir (Şekil 2.6) [56].



Şekil 2.6 Merkezcil desen [56]

**Dairesel (Annular) Desen:** Bir kıvrık kafesi andırır. Aşınmış volkanlarda veya çöküntü domları gibi derinden aşınmış (erozyona uğramış) domlarda (tümseklerde) yaygın olarak görülmektedir (Şekil 2.7) [56].



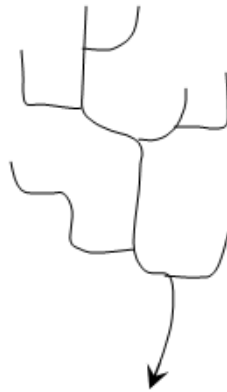
Şekil 2.7 Dairesel desen [56]

**Dengesiz (Deranged) Desen:** Volkanik çöküntülerde ve buzul bölgelerde görülmektedir. Su, en az direnç gösteren yolları takip eder. Akarsu zamanla yönünü ayarlar (Şekil 2.8) [56].



Şekil 2.8 Dengesiz desen [56]

**Dikdörtgen Yapılı (Rectangular) Desen:** Dikdörtgen ya da grid benzeri bir yapı göstermektedir. Akarsular, fisürleri (fissur), tektonik fayları (tectonic fault) veya ana kaya katılımını takip ettiğinden, sivri görünümlü kıvrımlar yaklaşık olarak 90 derecelik dönüşler yapar (Şekil 3.9) [56].



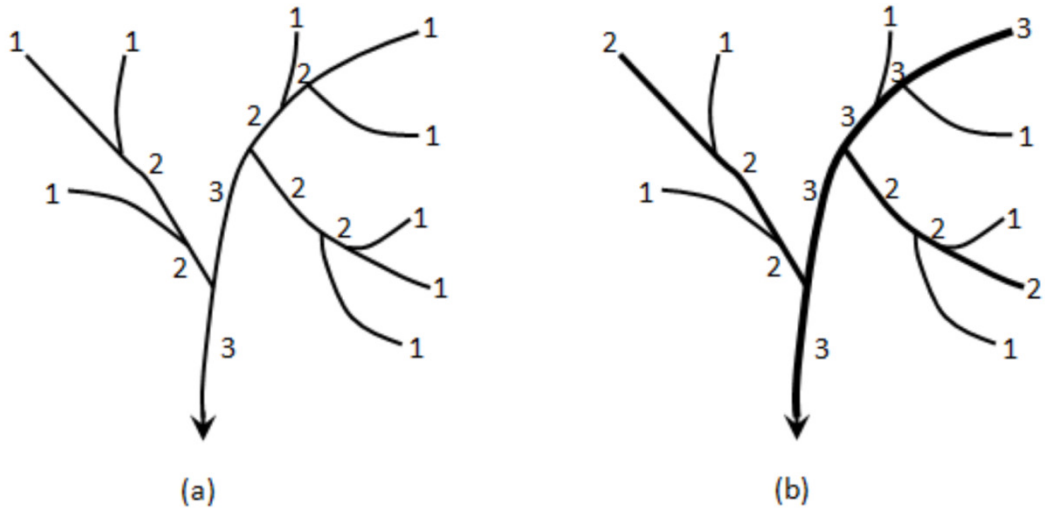
Şekil 2.9 Dikdörtgen yapılı desen [56]

### 2.3 Akarsu Sıralaması

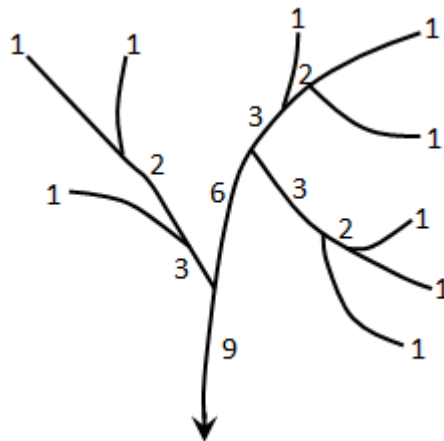
Akarsu; toprak ve ana kayanın (bedrock) en az direnç gösterdiği yolu takip ederek havza alanı genişliğinde akar. Akarsu ağının yapısı, deseni veya dallanması, havza alanının fiziksel özelliklerine göre çeşitlenmektedir.

Akarsu sınıflandırmalarından en önemli olanları, akarsu sıralamasına (stream order) dayanmaktadır. [35] tarafından, akarsu hiyerşisini sınıflandıran bir yöntem geliştirilmiştir. Ardından [36] tarafından bu sıralama iyileştirilmiştir. Şekil 2.10'da

gösterilmiş bu sistem (Horton ve Strahler), iki adımda uygulanmaktadır: (1) En küçük kollar birinci sırada (first-order) olan akarsulardır. Akış aşağıya (downstream) iki adet birinci sıra akarsu birleştiğinde, bir ikinci sıra (second-order) akarsuyu; iki adet ikinci sıra akarsu birleştiğinde, bir üçüncü sıra (third-order) akarsuyu oluştururlar ve sıralama böyle devam eder. İki aynı sırada parça birleştiğinde, akış aşağı olan parçanın sırası bir artar. Sulama kanalları ve diğer yapay sistemler sıralamaya dahil edilmezler. (2) Akış yukarı uzunluk ve sapma açılarına göre ana gövdeler belirlenir [40], [42]. [37] tarafından ise sıra numaraları akış aşağı toplanarak hiyerarşik bir sıralama oluşturulmuştur (Şekil 2.11).



Şekil 2.10 a) Akış aşağı akarsu sıralaması ve b) akış yukarı ana kollar [59]



Şekil 2.11 Shreve sıralaması [59]

Akarsuyun sırası ve büyüklüğü arasında doğrudan bir ilişki bulunmaktadır. Akarsu kolları birleştikçe, sıra numarası da artacaktır. Doğal olarak, akarsuların genişliği, derinliği, akışı ve boşaltması (discharge) da sıra numarası ile beraber artacaktır. Akarsu sıra numarası arttıkça, akarsu yatağının eğimi azalma eğilimindedir.

Birinci, ikinci ve üçüncü sırada olan akarsular başlangıç deresi (headwater stream) olarak sınıflandırılır ve dünya üzerindeki akarsu toplam uzunluğunun %80'den fazlasını oluşturmaktadır. Dört, beş ve altıncı sıradaki akarsular orta (medium) seviyededir. Akarsu (stream), yedinci ve daha büyük sırada, nehir (river) adını alır. Amazon nehri, dünya üzerindeki en büyük nehir olmasına rağmen, 12. sırada yer alır. Şekil 2.10b'de, dere (brook) hiyerarşide birinci sırada, çay ikinci sırada ve nehir üçüncü sırada olarak değerlendirilebilir [56].

#### 2.4 Akarsu Kolu Desenlerine Göre Akarsu Kollarının Sınıflandırılması

Akarsu geometrisi; akış (flow) ve taşınan su miktarı (load), akarsu karakteristiği (genişlik, eğim, derinlik, desen), vadinin yapısı ve akarsu kıyıları gibi hidrolojik faktörlerden etkilenmektedir. Hidroloji literatüründe kıvrımlılık ve güzergaha göre beş adet geometrik ve hidrolojik akarsu kolu vardır: Dolambaçlı (tortuous), düzensiz (irregular), düzenli (regular), geçişli (transitional) ve düz (straight) (Şekil 2.12) [41].



Şekil 2.12 Akarsu kolları [41]

## 2.5 Akarsu Ağı ve Havzasına Ait Jeomorfometrik Parametreler

Yer yüzeyinin geometrik özelliklerinin niceliksel olarak tanımı ve analizi jeomorfometri olarak tanımlanmaktadır. Temel olarak jeomorfometri, genellikle sayısal yükseklik modellerinden (Digital Elevation Model-DEM) yüzey parametreleri ve özelliklerini çıkaran bir dizi sayısal ölçü kullanmaktadır. Engebe (ruggedness) ve eğim (slope) parametrelerine göre yeryüzü düz, tepelik ve dağlık olarak üçe ayrılabilir. Engebe, sıfır ile bir arasında değişen bir değerdir. Engebe değeri sıfır ise düz alanları, bir ise dağlık alanları göstermektedir. Parametreler ve tanımları Çizelge 2.1’de verilmiştir [60].

Çizelge 2.1 Akarsu ağı ve havzasına ait jeomorfometrik parametreler [60]

| Parametreler                                                  | Simge      | Tanım                                                                            |
|---------------------------------------------------------------|------------|----------------------------------------------------------------------------------|
| Havza alanı                                                   | $A_b$      | Havzanın toplam alanı ( $\text{km}^2$ ).                                         |
| Havzanın zirve noktası (Basin Crest)                          | $Cr_b$     | Havzanın en yüksek noktası ( $Z_{max}$ ) (m)                                     |
| Havzanın çevresi                                              | $P_b$      | Havza sınırının toplam uzunluğu (km)                                             |
| Havzadaki toplam akarsu uzunluğu                              | $L_c$      | (km)                                                                             |
| Akarsu yoğunluğu [61]                                         | $D_b$      | Akarsuların toplam uzunluğunun havza alanına oranı                               |
| Havzadaki 10 m aralıklı yükseklik eğrilerinin toplam uzunluğu | $\sum L_c$ | (km)                                                                             |
| Rölyef [62]                                                   | $R_b$      | Havzadaki en büyük ile en küçük yükseklik farkı (m)<br>$R_b = Z_{max} - Z_{min}$ |
| Engebe değeri [63]                                            | $R_n$      | $R_n = D_b R_b$                                                                  |
| Melton’ın engebe değeri [64]                                  | $M$        | $M = R_b / \sqrt{A_b}$                                                           |
| Havzanın eğimi [65]                                           | $S_b$      | $e \sum L_c / A_b$ , (e, eş yükseklik değeri)                                    |

### ÇALIŞMADA KULLANILAN VERİLERİN ÖZELLİKLERİ

United States Geological Survey (USGS) Ulusal Mekansal Veri Programı (National Geospatial Program)'nın temel taşlarından biri olan Ulusal Harita (National Map), topografik bilgiyi sunmak ve dağıtmak için USGS, diğer eyalet ve yerel ortaklar arasında işbirliği ile oluşturulmuştur. Rekreasyon, bilimsel analiz ve acil müdahale gibi birçok faydası vardır. Ulusal Harita'nın ürün, servis ve verisine web üzerinden erişilebilmektedir. Ulusal Harita; ortofoto, yükseklik, coğrafi isimler, hidrografi, sınırlar, ulaşım, yerleşim/meskun alan ve bitki örtüsü ile ilgili mekansal bilgileri içermektedir. Ulusal Harita, Ulusal Mekansal Veri Altyapısı (National Spatial Data Infrastructure (NSDI))'na önemli bir katkı sağlamaktadır [66]. Amerika Birleşik Devletleri (ABD) Ulusal Hidrografi Veri Seti (UHVS) (National Hydrography Dataset (NHD)), United States Environmental Protection Agency (USEPA) ve USGS tarafından geliştirilen, birçok veri kaynağından ve web harita servisinden oluşturulan, ABD'nin hidrografisini gösteren mekansal vektör veri tabakasıdır. ABD Ulusal Harita servislerinden biridir. Ülke genelinde, orta çözünürlük (1:100,000 ölçeğinde) ve yüksek çözünürlük (1:24,000 ölçeğinde) hatta daha yüksek çözünürlükte bulunmaktadır. ABD'nin hidrografik verileri USGS web sitesinde UHVS'de mevcuttur. UHVS, doğal ve yapay su kütleleri ve ilişkili varlıklar ile ilgili bilgiyi kodlayan kapsamlı bir sayısal mekansal veri setidir. Bu nesnelere dair şu bilgiler kodlanmıştır: Meta veri, tip, sınıf, geometri (delineation), coğrafi isim, konum, uzunluk, alan vd. Ayrıca, bu bilgilerin ilişkilendirilebilmesi için “anlamlı hidrografik parça (segment) kodu” (reach code) ve akış yönü bilgisini içermektedir. Veri seti sürdürülebilir (güncellenebilir ve geliştirilebilir) yapıdadır [67].

### 3.1 Hidrografik Nesne Tipleri ve Özellikleri

Bir nesne, varlığın veri tabanındaki temsili olarak tanımlanmaktadır. Hidrografik nesneler, doğal ve yapay su kütlelerini ve ilişkili varlıkları temsil etmektedir. Nesne tiplerine göre sınıflandırılmakta ve “nesne kodu” (feature code), “tanımlama alanı” (description field) ve her nesneyi eşsiz bir kod ile tanımlayan “tanımlayıcı” (common identifier) gibi özniteliklerle tanımlanabilmektedir. Hidrografik nesneler; nokta, çizgi veya alan geometrisine sahip olabilirler. Çizgiler, su akış yönündedir. Ayrıca, veri tabanında orta eksen (artificial path), konektör (connector) ve geçit (underpass) gibi ağın bağlantı özelliğini sağlayan nesneler bulunmaktadır. Nesne tiplerine örnek olarak; “akarsu/nehir”, “kanal/hendek” ve “göl/gölcük” verilebilir. Her tipin bir ismi ve tanımı vardır. Örneğin, en sık karşılaşılan üç nesne tipi ve tanımları Çizelge 3.1’de gösterilmiştir. Her nesne özelliğinin bir ismi, tanımı ve değer listesi vardır. Örneğin, göl/gölcük ve akarsu/nehir nesneleri hidrografi kategorisine girmektedir (Çizelge 3.2).

Çizelge 3.1 Nesne tipleri ve tanımları [67]

| Nesne tipi   | Tanım                                                                                                                                                                |
|--------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Akarsu/Nehir | Akan su kütlesi                                                                                                                                                      |
| Göl/Gölcük   | Kara ile çevrili doğal su kütlesi                                                                                                                                    |
| Kanal/Hendek | Suyu taşımak, toprağı sulamak, iki veya daha fazla su kütlesini bağlamak, deniz taşıtları için bir su yolu olarak hizmet vermek için inşa edilen yapay, açık su yolu |

Çizelge 3.2 Hidrografi kategorisi ile ilgili bilgiler [67]

|          | İsim                     | Tanım                                                                    |
|----------|--------------------------|--------------------------------------------------------------------------|
| Özellik  | Hidrografi kategorisi    | Su içeren nesne                                                          |
| Değerler | Mevsimsel (intermittent) | Daha çok yağmur yağınca ve kar eriyince yılın belli bir kısmı su içerir. |
|          | Daimi (perennial)        | Nadir kuraklık dönemleri hariç yıl boyunca su içerir.                    |

Her nesne tipi beş basamaklı nesne kodu kullanılarak nesne tipi, özelliği ve değerine göre kodlanmaktadır. İlk üç basamak nesne tipi, son iki basamak özellik ve değeridir.

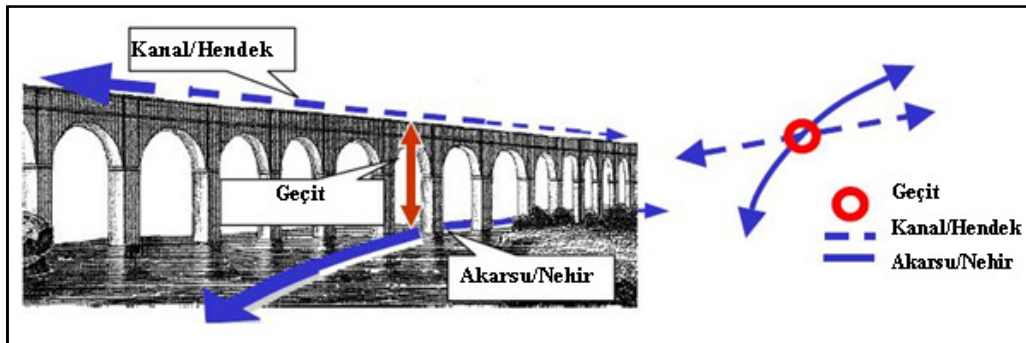
Örneğin, “baraj/bent” nesne tipi kodu “343”tür. Bu nesne tipine ait özellik ve değer bilgisi ise “00” ile “04” arası değişen beş farklı şekilde kodlanmaktadır (FFFCC: FFF nesne tipi, CC özellik) (Çizelge 3.3).

Çizelge 3.3 Baraj/bent ile ilgili kodlama [67]

| Nesne Kodu | Nesne Tipi | Özellik ve Değer (Tanım)                         |
|------------|------------|--------------------------------------------------|
| 34300      | Baraj/Bent | Özellik tipi: Özellik yok                        |
| 34301      | Baraj/Bent | Malzeme: Doğal-Kullanım Durumu: Kullanımda       |
| 34302      | Baraj/Bent | Malzeme: Doğal-Kullanım Durumu: Yapım aşamasında |
| 34303      | Baraj/Bent | Malzeme: Yapay-Kullanım Durumu: Kullanımda       |
| 34304      | Baraj/Bent | Malzeme: Yapay-Kullanım Durumu: Yapım aşamasında |

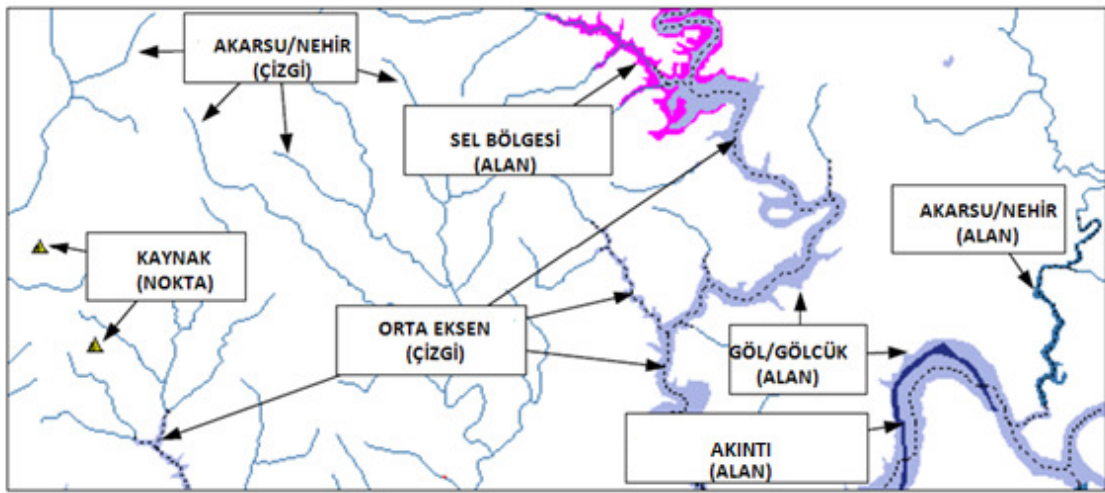
Her nesneye bir tanım alanı, nesne tipi ve nesne kodu atanmaktadır. “FCODE” adlı öznitelik, nesne kodu; “FTYPE” adlı öznitelik, nesne tipi ve “DESCRIPT” adlı öznitelik, tanım alanı (description field) bilgisi içermektedir. Nesne tipi ve tanım alanı “character string” veri tipindedir.

Orta eksen ve konektör, akarsu ağını tamamlayan nesnelerdir. Orta eksen, alan olarak gösterilen nesnenin ortasından geçerek su akışını çizgi olarak temsil etmektedir. Konektörler ise birleşmeyen (kopuk) ağ parçaları (network segment) arasında bağlantı sağlamaktadır. Geçit nesnesi, “üzerinde” ve “altında” olmak üzere iki ilişki tipini göstermektedir. Nesneler farklı yüksekliklerde karşılaştıklarında geçit nesnesi yerleştirilmektedir. Kanalin/hendeğin, akarsuyun üzerinden geçmesi halinde, kanal/hendek “üzerinde”, akarsu/nehir ise “altında” olarak kodlanmaktadır (Şekil 3.1).



Şekil 3.1 Geçit nesnesinin gösterimi [67]

Hidrografik nesneler; nokta, çizgi ve alan geometrileri ile veri tabanlarında saklanmaktadır (Şekil 3.2). Çizgi, bir düğüm noktası ile başlar ve başka bir düğüm noktası ile biter. İki akarsuyun kesişim noktasında bir düğüm noktası vardır. Düğüm noktaları, farklı yükseklikte çizgilerin karşılaştığı yerlerde bulunmamaktadır. Örneğin, bir kanalın akarsuyu geçtiği yerde düğüm noktası bulunmaz. Düğüm noktaları, orta eksenlerin (artificial path) kollara (branch) ayrıldığı yerlerde de bulunur. Çizgilerin (orta eksen, kanal/hendek, konektör, boru hattı ve akarsu/nehir) her zaman bir yönü vardır. Çizgiler akış yönünde çizilmektedir. Belirsiz kaynak olduğunda ve gelgit havzasında akış yönünde çizilmeyebilir [67].



Şekil 3.2 Nokta, çizgi ve alan hidrografik nesneler [67]

### 3.2 Anlamlı Hidrografik Parçalar

1970'li yıllarda USEPA, ABD hidrografik veri tabanlarını oluşturmaya başlamıştır. Anlamlı hidrografik parça (reach) dosya veri tabanlarının yapısı ve içeriği; hiyerarşik sıralama oluşturmak, hidrolojik navigasyon ve anlamlı hidrografik parça kodları diye bilinen tanımlayıcı veriyi sağlamak amacıyla oluşturulmuştur. Anlamlı hidrografik parça dosyaları, akarsu ağını tanımlamak için parçalar arasındaki ilişkileri kullanmaktadır. Bu ilişkiler, -topolojik olarak bağlı olmasa da- parçalar arasında bağlantıyı sağlamaktadır. Su akış yönü bu ilişkiler kullanılarak kodlanmaktadır. Bu bağlanabilirlik, parçaların hiyerarşik sırasını belirler (akarsu ağındaki belirli bir noktada akış yukarı ve akış aşağı ilişkisinin tanımlanması). Akış yukarı veya akış aşağı yönde ağ navigasyonu sağlanmaktadır.

Anlamalı Hidrografik Parça dosyaları ilk kez 1970'lerin başında ortaya çıkmıştır (Reach File Version 1.0 Alpha (RF1 A)). 1975 yılında tamamlanmıştır. İlk tam uygulama (Reach File Version 1.0 (RF1)) 1982 yılında tamamlanmıştır. RF1'de, 1:250,000 ölçekli USGS hidrografik haritalar, fotogrametrik olarak 1:500,000 ölçeğine küçültülmüştür. RF1 yaklaşık olarak 650,000 mil akarsu içeren 68,000 parçadan (segment) oluşmaktadır.

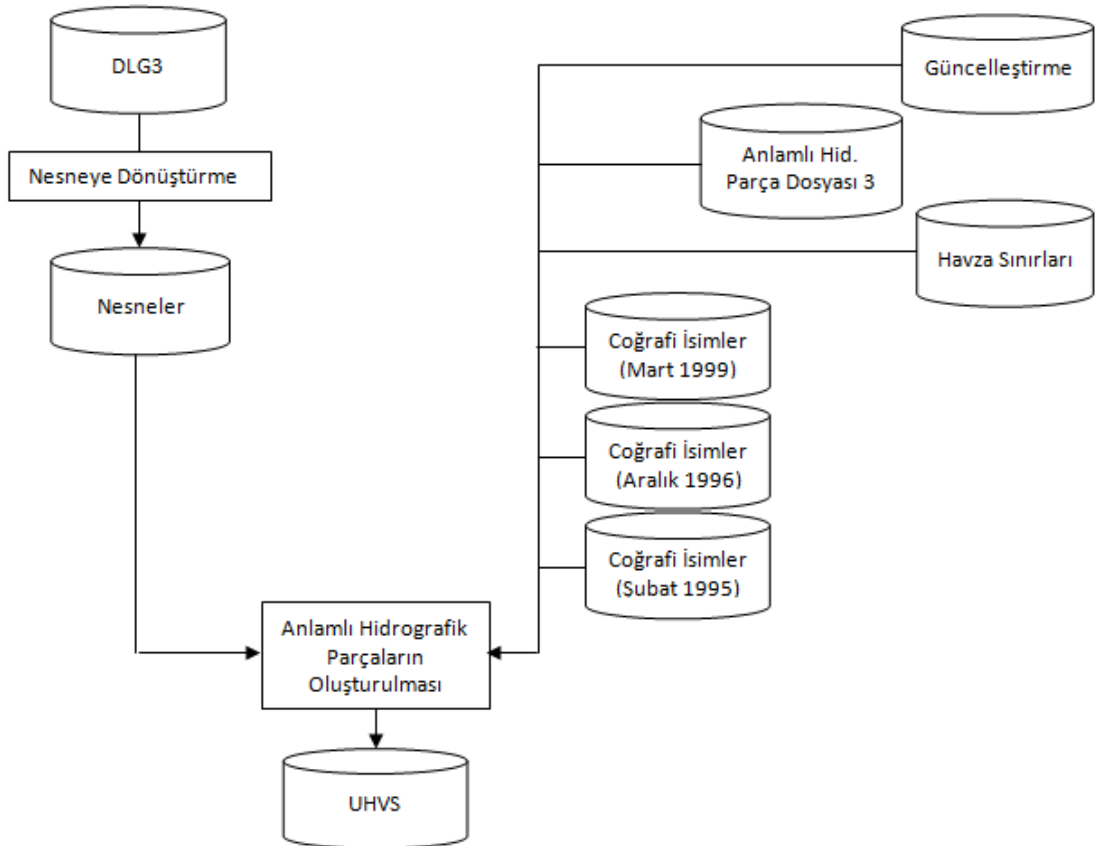
RF1, halen ulusal uygulamaları desteklemesine rağmen, daha detaylı akarsu ağı ihtiyacı Anlamalı Hidrografik Parça dosyası 2.0 sürümü (Reach File Version 2.0 (RF2)) ile 1980'lerin başında gündeme gelmiştir. RF2, şimdiki adı National Geographic Names Database olan USGS Geographic Names Information System (GNIS) nesne dosyası kullanılarak oluşturulmuştur. RF1 parçalarına yeni bir tabaka eklenerek elde edilen RF2, 170,000 parça içermektedir.

Daha kapsamlı ulusal veri tabanı ihtiyacı Anlamalı Hidrografik Parça dosyası üçüncü sürümünün (Reach File Version 3 (RF3)) geliştirilmesini sağlamıştır. Bu veri tabanı, RF1 ve RF2 verisi GNIS ile birlikte 1988 tarihli 1:100,000 ölçekli USGS DLG (Digital Line Graph) verisi içermektedir. Yaklaşık 3,200,000 hidrografik parça ve 93,000,000 nokta içermektedir. UHVS, anlamalı hidrografik parça verisi gelişiminde 4. seridir.

1970 sonrası USGS, dijital teknolojinin kullanımı ile beraber kenar/düğüm veri modelini oluşturmak için kartografik olarak üretilmiş basılı haritalardan elle sayısallaştırma yoluyla DLG (Digital Line Graph)'yi oluşturmuştur. 1990'lı yıllarda ISO (International Organization for Standardization) tarafından kabul edilen Mekansal Veri Transferi Standartları (Spatial Data Transfer Standart) [68] gelişimine paralel olarak USGS, DLG yapısını nesne tabanlı CBS veri modeli DLG-E (DLG 2) ile geliştirmiştir. İleriki yıllarda yapılan iyileştirmeler ile nesnelerin öznitelikleri ve topolojik ilişkilerinin oluşturulduğu DLG-F (DLG 3) ortaya çıkmıştır. Ulusal Hidrografi Veri Seti (UHVS), DLG verisi üzerinde geliştirilen anlamalı hidrografik parça geometrileri, kodları, su akış yönü ve coğrafi isimlerin konumlarını içeren anlamalı hidrografik parça dosyaları, havza sınırları, coğrafi isimler ve güncelleştirmeler ile birleştirilerek oluşturulmuştur. Şekil 3.3'de UHVS mimarisi gösterilmiştir [67].

UHVS'yi oluşturan bilgi kaynakları [67]:

- Digital Line Graph 3 (DLG-3) verisi (orta eksen ve konektör nesneleri hariç), USGS topografik haritalarının sayısallaştırılması ve basılı olmayan kaynak materyallerinden elde edilmiştir. Topolojik olarak yapılandırılmıştır. UTM koordinat sisteminde ve North American Datum 1927 (NAD27) datumundadır.
- Anlamli Hidrografik Parça dosyası üçüncü sürüm (Reach File Version 3 - RF3) verisi, parça geometrileri, kodları, su akış yönleri ve coğrafi isimleri içermektedir. Bu veri seti, 1:100,000 DLG verisi üzerinde geliştirilmiştir.
- Havza sınırları, nesne ve parçaları alt havzalara atamak için farklı kurumlar tarafından geliştirilmiştir.
- Coğrafi isimler, USGS Coğrafi İsimler Bilgi Sistemi'ndeki Ulusal Coğrafi İsimler Veri Tabanı verisidir.



Şekil 3.3 UHVS mimarisi [67]

Anlamalı bir hidrografik parça, hidrolojik özellikler gösteren su nesnesine ait anlamalı bir parçadır. Üç tip anlamalı hidrografik parça bulunmaktadır: Anlamalı Akarsu Parçası (AAP) (transport reach), Anlamalı Kıyı Parçası (AKP) (coastline reach) ve Anlamalı Su Kütlesi Parçası (ASKP) (waterbody reach). AAP ve AKP, çizgi nesnelerden oluşmaktadır. AAP, akış yönünde çizilmektedir. AKP ise su, parçanın sağında kalacak yönde çizilmektedir. ASKP, alan nesnelerdir (göl/gölcük gibi). Anlamalı hidrografik parça kodları, her parçayı etiketlemektedir. Tanımlayıcı ise her parçayı tanımlamaktadır [67].

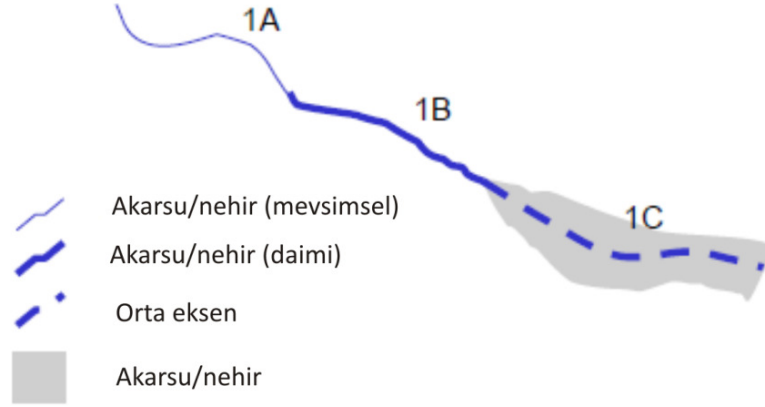
### 3.2.1 Anlamalı Akarsu Parçalarının Kodlanması

Anlamalı akarsu parçası, akarsu ağı içerisinde suyun akış yolunu belirlemektedir. Akış yönünün kodlanmasında kullanılmaktadır. Anlamalı akarsu parçaları arası akış ilişkilerinin kodlanması, akarsu ağı içerisinde navigasyonu sağlamakta ve ağda hiyerarşik sıralama (akarsu düzeylerinin belirlenmesi) için temel oluşturmaktadır. “COM\_ID” isimli bir öznitelik ile her parça 10 basamaklı bir tamsayı ile tek anlamalı (eşsiz) olarak tanımlanmaktadır. Ayrıca, “RCH\_CODE” isimli bir öznitelik ile her bir parça aynı zamanda tek bir sayısal koda sahiptir. 14 basamaklı kod iki bölümden oluşmaktadır. İlk sekiz basamak; parçanın içinde yer aldığı havzanın hidrolojik birim kodu, son altı basamak ise ardışık sıra numarasıdır [67].

Anlamalı akarsu parçaları üç kurala göre belirlenmektedir:

- Temel nesne kuralı (underlying feature rule),
- Kavşak-kavşak kuralı (confluence-to-confluence rule) ve
- Dallanan orta eksen kuralı (branched path rule).

**Temel Nesne Kuralı:** Birbirini izleyen AAP’ler, temel nesne kurallarına göre belirlenmektedir. Aynı tip nesne parçaları birleşirse AAP değeri değişmez. Temel nesne kuralında, bir alan nesne (akarsu/nehir, göl/gölcük gibi) içine çizilen orta eksen, alan nesne ile aynı tipte kabul edilir. Şekil 3.4’de, AAP-1; A ve B akarsularını ve akarsu/nehir nesnesine ait C orta eksenini izlemektedir.



Şekil 3.4 AAP'lerin belirlenmesi (aynı tip nesneler için) [67]

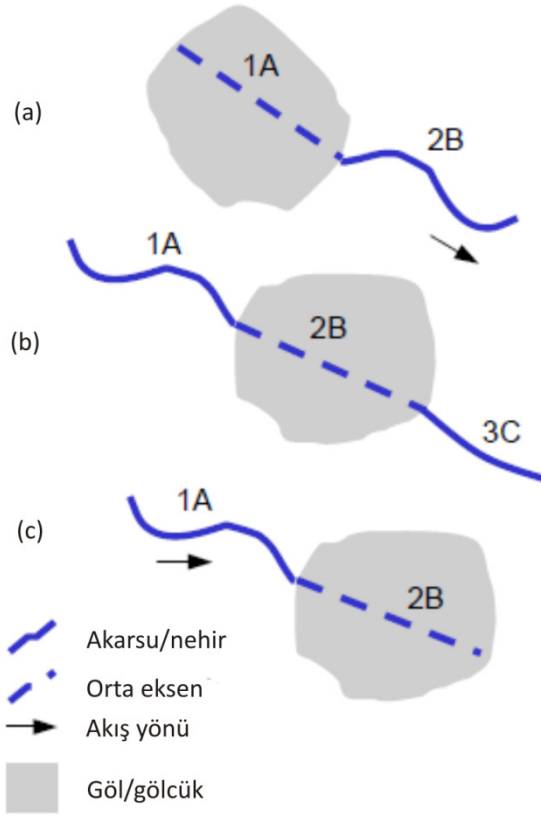
Nesne tipi değiştiğinde farklı bir parçaya geçilmektedir. Şekil 3.5'de, AAP-1, A kanalını ve AAP-2, B boru hattını izlemektedir.



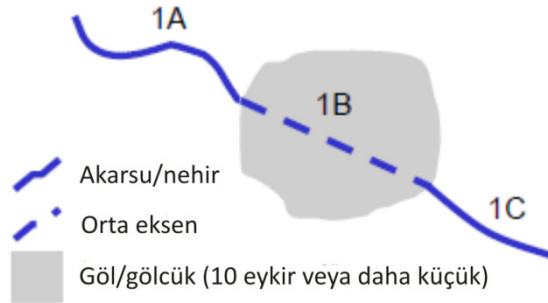
Şekil 3.5 AAP'lerin belirlenmesi (farklı tip nesneler için) [67]

Şekil 3.6a'da, AAP-1, göl/gölcük nesnesine ait A orta eksenini ve AAP-2, B akarsuyunu izlemektedir. Şekil 3.6b'de, AAP-1, A akarsuyunu; AAP-2, göl/gölcük nesnesine ait B orta eksenini ve AAP-3, C akarsuyunu izlemektedir. Şekil 3.6c'de, AAP-1, A akarsuyunu ve AAP-2, göl/gölcük nesnesine ait B orta eksenini izlemektedir.

10 eykir (acre) (40,469 m<sup>2</sup>) veya daha küçük göl/gölcük nesnesine ait orta eksen için ayrı bir AAP belirlenmemektedir. Şekil 3.7'de, AAP-1; A akarsuyunu, B orta eksenini (10 eykir veya daha küçük göl/gölcük nesnesine ait) ve C akarsuyunu izlemektedir.



Şekil 3.6 AAP'lerin belirlenmesi (farklı tip nesneler için) [67]



Şekil 3.7 Önemsiz göl/gölcük nesneleri için AAP'lerin belirlenmesi [67]

**Kavşak-Kavşak Kuralı:** Bir AAP,

- İki kavşak arasında,
- Kaynak ile kavşak arasında,
- Kavşak ile ağız (son nokta) arasında ve
- Kaynak ile ağız (son nokta) arasında

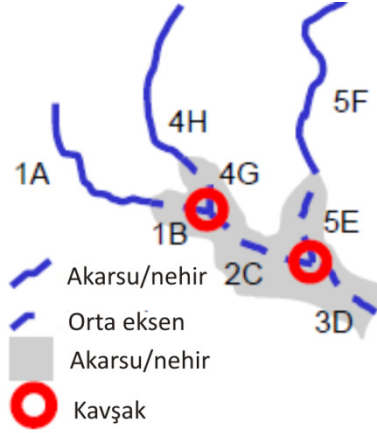
olabilir.

Şekil 3.8’de, AAP-1, A akarsuyunu; AAP-2, B akarsuyunu; AAP-3, C akarsuyunu; AAP-4, D akarsuyunu ve AAP-5, E akarsuyunu izlemektedir.



Şekil 3.8 Kavşak-kavşak kuralına göre AAP’lerin belirlenmesi [67]

Şekil 3.9’da, AAP-1, A akarsuyunu ve akarsu/nehir nesnesine ait B orta eksen parçasını; AAP-2, C orta eksen parçasını; AAP-3, D orta eksen parçasını; AAP-4, H akarsuyunu ve G orta eksen parçasını ve AAP-5, F akarsuyunu ve E orta eksen parçasını izlemektedir.



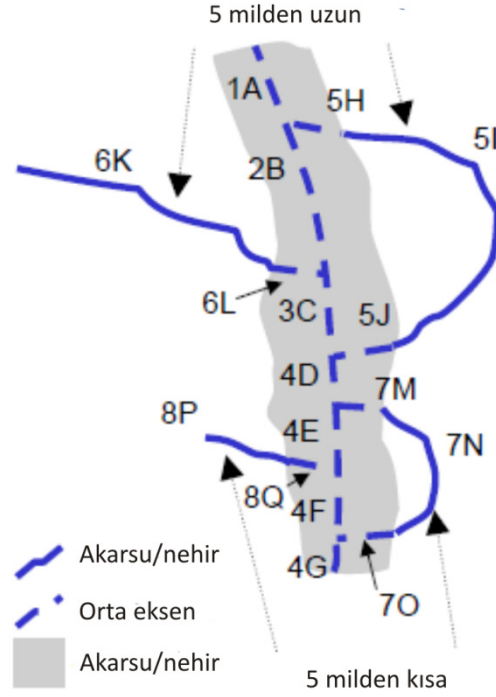
Şekil 3.9 Kavşak-kavşak kuralına göre AAP’lerin belirlenmesi (aynı tip nesneler için) [67]

En çok kolu olan veya en geniş akarsu, ana akarsu olarak seçilmektedir. Akarsu/nehir tipinde bir alan nesne iki kola ayrıldığında AAP belirlenmesinde ikincil kolun izlenip izlenmeyeceğine karar verilmelidir. Kolları ayıran kara parçasının büyüklüğü, şekli ve ikincil kola ait diğer kolların varlığı ikincil kolun önemliliğini göstermektedir. Kolları ayıran kara parçası büyük ve uzun olursa birincil ve ikincil kollar farklı akarsular olmaktadır. Şekil 3.10a’da, AAP-1, akarsu/nehir nesnesine ait A orta eksen parçasını; AAP-2, B orta eksen parçasını; AAP-3, C orta eksen parçasını; AAP-4, D orta eksen

Figure 1 consists of two maps, (a) and (b), showing the Akarsu catchment area. Map (a) is a detailed view of the catchment area, showing the main channel (Akarsu/nehir) and its tributaries (Orta eksen). The main channel is labeled 1A, 2B, 3C, 4D, 5E, 6F, 7G, 7H, 8I, and 8J. The tributaries are labeled 1A, 2B, 3C, 4D, 5E, 6F, 7G, 7H, 8I, and 8J. Map (b) is a simplified view of the catchment area, showing the main channel (Akarsu/nehir) and its tributaries (Orta eksen). The main channel is labeled 1A, 2B, 3C, 4D, 4E, 5F, and 5G. The tributaries are labeled 1A, 2B, 3C, 4D, 4E, 5F, and 5G. A legend indicates that the solid blue line represents the Akarsu/nehir, the dashed blue line represents the Orta eksen, and the grey area represents the Akarsu/nehir catchment area.

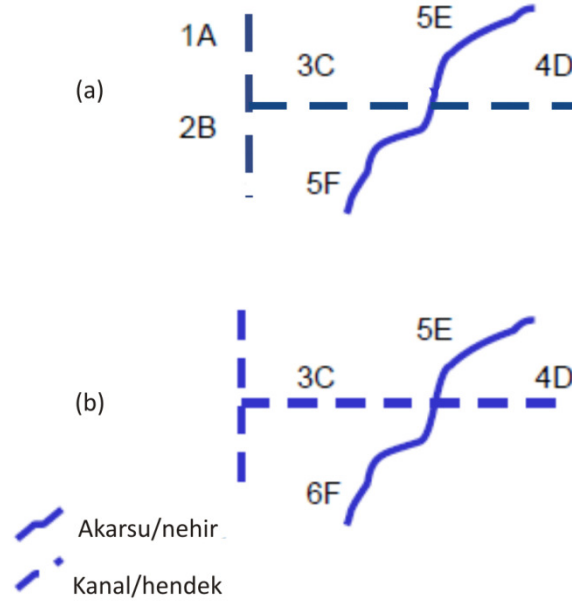
5 milden küçük bir AAP, bir nesne ile kesiştiğinde kavşak-kavşak kuralı uygulanmamaktadır. Bu yaklaşım büyük bir akarsuyu gereksiz yere küçük AAP'lere bölmeyi önlemektedir. Şekil 3.11'de, AAP-1, akarsu/nehir nesnesine ait A orta eksen parçasını; AAP-2, B orta eksen parçasını; AAP-3, C orta eksen parçasını; AAP-4; D, E, F ve G orta eksen parçalarını (önemsiz kolları nedeniyle); AAP-5, H orta eksen parçasını, I akarsuyunu ve J orta eksen parçasını; AAP-6, K akarsuyunu ve L orta eksen parçasını;

AAP-7, M orta eksen parçasını, N akarsuyunu ve O orta eksen parçasını ve AAP-8, P akarsuyunu ve Q orta eksen parçasını izlemektedir.



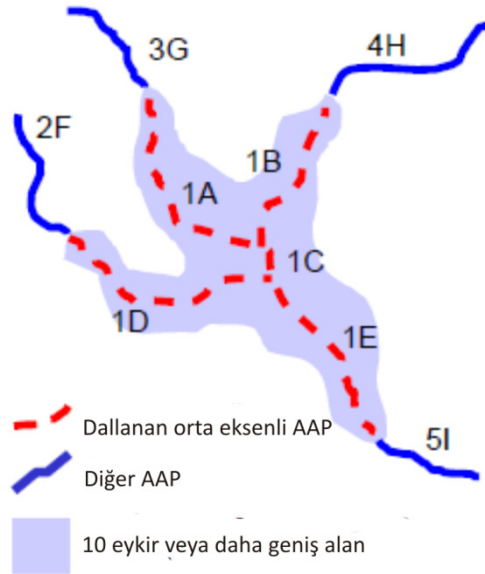
Şekil 3.11 AAP'lerin belirlenmesinde önemli ve önemsiz kavşaklar (5 mil kuralı) [67]

Akarsu/nehir ve kanal/hendek aynı yükseklikte kesişirse, genellikle kanal/hendek için kavşak kuralı uygulanmakta, akarsu/nehir için uygulanmamaktadır. Şekil 3.12a'da, AAP-1, A kanalını; AAP-2, B kanalını; AAP-3, C kanalını; AAP-4, D kanalını ve AAP-5; E ve F akarsularını izlemektedir. Alternatif yaklaşımda ise, her iki nesne tipi için kavşak-kavşak kuralı uygulanmaktadır. Şekil 3.12b'de, AAP-3, C kanalını; AAP-4, D kanalını; AAP-5, E akarsuyunu ve AAP-6, F akarsuyunu izlemektedir. Farklı yükseklikte karşılaşan nesneler için kavşak-kavşak kuralı uygulanmamaktadır. AAP'ler, farklı yüksekliklerde karşılaştıkları noktada kırılma göstermemektedir.



Şekil 3.12 Akarsu/nehir ile kanal/hendek kavşağında AAP'lerin belirlenmesi [67]

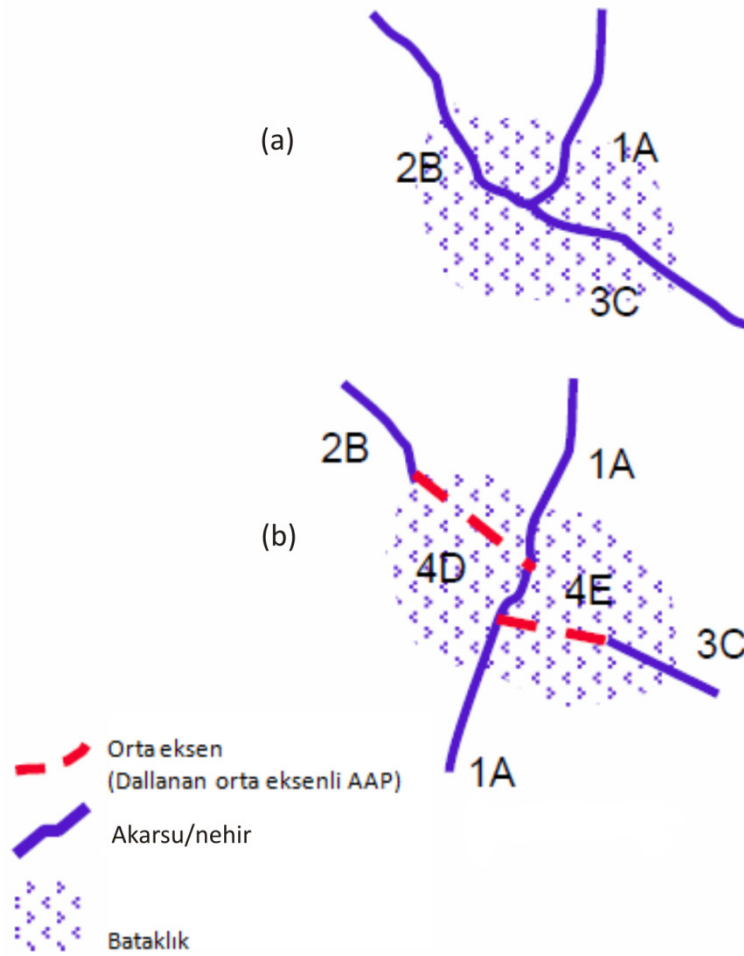
**Dallanan Orta Eksen Kuralı:** Bu kural, 10 eykir veya daha büyük (yaklaşık 4.05 hektar veya daha fazla) göl/gölcük veya bataklık gibi bir alana giren ve çıkan AAP'ler için uygulanmaktadır. Orta eksenler, AAP'lerin belirlenmesinde kullanılan nesnelerdir. Şekil 3.13'de, AAP-1; A, B, C, D ve E orta eksen parçalarını (10 eykir veya daha büyük alana ait); AAP-2, F akarsuyunu; AAP-3, G akarsuyunu; AAP-4, H akarsuyunu ve AAP-5, I akarsuyunu izlemektedir.



Şekil 3.13 Dallanan orta eksenlerle AAP'lerin belirlenmesi [67]

Bu üç kural, anlamlı akarsu parçalarının belirlenmesini yönetmektedir. Bu kurallar için bazı istisnalar vardır. Nesnelerin olağan dışı biçimleri ve bilgi kaynağındaki belirsizlikler kuralların değiştirilmesini gerektirmektedir.

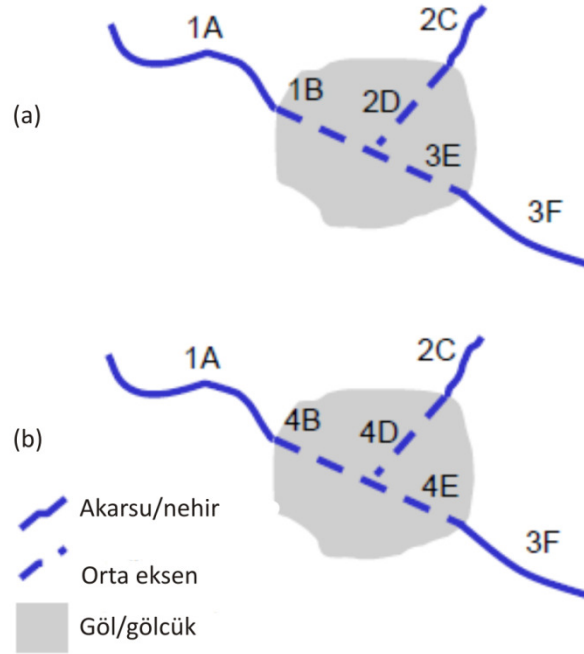
Bir nesne içerisindeki akış kopuksa, dallanan orta eksenli AAP bağlantıyı sağlamaktadır. Şekil 3.14a'da, AAP-1, A akarsuyunu; AAP-2, B akarsuyunu ve AAP-3, C akarsuyunu izlemektedir. Şekil 3.14b'de, AAP-1, A akarsuyunu; AAP-2, B akarsuyunu; AAP-3, C akarsuyunu ve AAP-4, bataklık nesnesine ait D ve E orta eksen parçalarını izlemektedir.



Şekil 3.14 Dallanan orta eksenlerle AAP'lerin belirlenmesi [67]

10 eykirden küçük göl/gölcük nesneleri içindeki orta eksenlere dallanan orta eksen kuralı uygulanmamaktadır. Böyle durumlarda göl içerisindeki orta eksene kavşak-kavşak kuralı uygulanmaktadır. Şekil 3.15a'da, AAP-1, A akarsuyunu ve göl/gölcük nesnesine ait B orta eksen parçasını; AAP-2, C akarsuyunu ve D orta eksen parçasını; AAP-3, F akarsuyunu ve E orta eksen parçasını izlemektedir. 10 eykirden daha büyük

göl/gölcükte dallanan orta eksen kuralı uygulanmaktadır: Şekil 3.15b’de, AAP-1, A akarsuyunu; AAP-2, C akarsuyunu; AAP-3, F akarsuyunu ve AAP-4, göl/gölcük nesnesine ait B, D ve E orta eksen parçalarını izlemektedir.



Şekil 3.15 Göllerde AAP’lerin belirlenmesi [67]

### 3.2.2 Anlamlı Akarsu ve Kıyı Parçaları Arasındaki Akış İlişkilerinin Kodlanması

AAP ve AKP arasındaki akış ilişkileri, parçalar arasında bağlantılı bir akarsu ağı oluşturmaktadır. AAP arasındaki ilişkiler, bağlantılı bir akarsu ağını tanımlamakta ve parçalar arasında su akış yönünü belirlemektedir. Bu bağlantı, AAP’lerin hidrolojik sıralamasına olanak vermektedir (akarsu ağında belirli bir noktada akış yukarı ve akış aşağı ilişkisinin tanımlanması). Akış yukarı veya akış aşağı yönde ağın navigasyonu sağlanmaktadır. AAP ve AKP arasındaki ilişkiler, akarsu ağını kıyı çizgisine bağlamakta ve kıyı çizgisi boyunca ağın ağızları (mansap) sıralanmaktadır. Akış ilişkileri ardışık iki parça arasındaki su akışı tanımlanarak kodlanmaktadır. Su akış yönü her zaman bilinmemektedir. Böyle durumlarda, akış ilişkisi oluşturulamamaktadır [67].

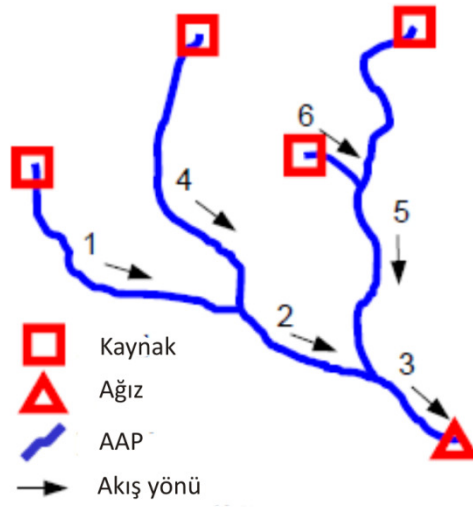
**Akış İlişkileri Kullanılarak Akış Yönlerinin Belirlenmesi:** Ardışık parçaların tanımlayıcıları ve parçalar arası su akış yönleri, akış ilişkilerinin kodlanmasında kullanılmaktadır. Aşağıda verilen ilk beş madde, AAP’ler arası su akış ilişkilerini

tanımlamaktadır. Altıncı madde ise AKP'ler arasındaki ya da AAP'ler ve AKP'ler arasındaki bağlantıyı tanımlamaktadır [67].

1. **İçeri:** Birinci AAP, ikinci AAP'nin içine akar. Parçalar arası çoğu akış ilişkisi bu tiptedir (Şekil 3.16, Çizelge 3.4).
2. **Dışarı:** Birinci AAP, ikinci AAP'den çıkar (Şekil 3.17, Çizelge 3.5).
3. **İki yönlü:** Her iki yöne akış vardır (uygulanmamaktadır).
4. **Ağ başlangıcı/Kaynak:** İkinci AAP, akarsu ağının kaynağından itibaren başlar. Birinci AAP, boş değer (0) almaktadır (Şekil 3.16, Çizelge 3.4).
5. **Ağ sonu/Ağız:** Birinci AAP, akarsu ağının son noktasında sonlanır. İkinci AAP, boş değer almaktadır (Şekil 3.16, Çizelge 3.4).
6. **Akış olmayan bağlantı:**
  - a) AKP'ler arasında akış olmayan bağlantı bulunmaktadır. AKP, su sağ tarafta kalacak biçimde yönlendirilmektedir. Adalar ise saat ibresinin tersi yönünde yönlendirilmektedir. AKP'ler arasındaki akış ilişkisi sıralıdır. Bir AKP (ilişkide birinci parça), diğer parçanın başlangıcına (ilişkide ikinci parça) bağlanmaktadır (Şekil 3.19, Çizelge 3.7).
  - b) AAP'ler ve AKP'ler arasında akış olmayan bağlantı bulunmaktadır. Bir AAP (ilişkide birinci parça), AKP'nin başlangıcına (ilişkide ikinci parça) bağlanmaktadır. AAP ve AKP arasındaki ilişki sayesinde akarsu ile kıyı çizgisinin kesişim noktası belirlenmektedir (Şekil 3.19, Çizelge 3.7).

### Akış ilişkisi ile ilgili örnekler:

#### Örnek 1:

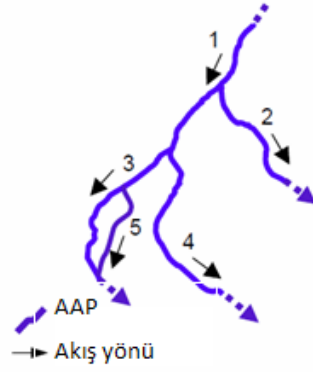


Şekil 3.16 AAP'ler arasında akış ilişkileri [67]

Çizelge 3.4 AAP'ler arasında akış ilişkileri [67]

| Akış İlişkileri |            |                      |
|-----------------|------------|----------------------|
| Birinci AAP     | İkinci AAP | Yön                  |
| 1               | 2          | İçeri                |
| 4               | 2          | İçeri                |
| 6               | 5          | İçeri                |
| 5               | 3          | İçeri                |
| 2               | 3          | İçeri                |
| 0               | 1          | Ağ Başlangıcı/Kaynak |
| 0               | 4          | Ağ Başlangıcı/Kaynak |
| 0               | 6          | Ağ Başlangıcı/Kaynak |
| 0               | 5          | Ağ Başlangıcı/Kaynak |
| 3               | 0          | Ağ Sonu/Ağız         |

### Örnek 2:

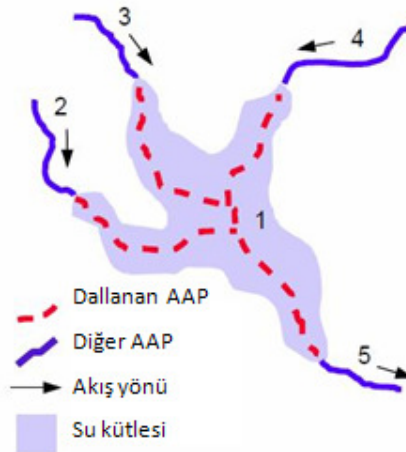


Şekil 3.17 AAP'ler arasında akış ilişkileri [67]

Çizelge 3.5 AAP'ler arasında akış ilişkileri [67]

| Akış ilişkileri |            |        |
|-----------------|------------|--------|
| Birinci AAP     | İkinci AAP | Yön    |
| 1               | 3          | İçeri  |
| 1               | 4          | İçeri  |
| 2               | 1          | Dışarı |
| 5               | 3          | Dışarı |
| 5               | 3          | İçeri  |

### Örnek 3:

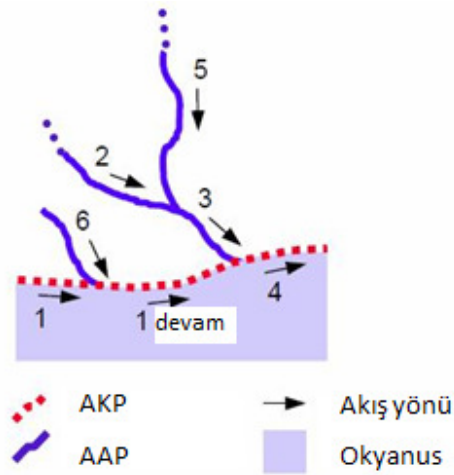


Şekil 3.18 Dallanan AAP'ler arasında akış ilişkileri [67]

Çizelge 3.6 Dallanan AAP'ler arasında akış ilişkileri [67]

| Akış ilişkileri |            |       |
|-----------------|------------|-------|
| Birinci AAP     | İkinci AAP | Yön   |
| 2               | 1          | İçeri |
| 3               | 1          | İçeri |
| 4               | 1          | İçeri |
| 1               | 5          | İçeri |

Örnek 4:



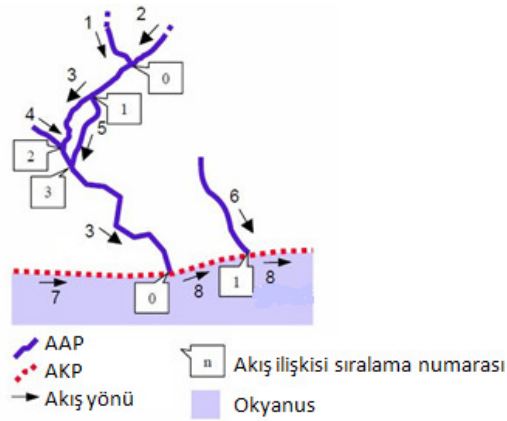
Şekil 3.19 Akış olmayan bağlantı ve akış ilişkileri [67]

Çizelge 3.7 Akış olmayan bağlantı ve akış ilişkileri [67]

| Akış ilişkileri |              |                       |
|-----------------|--------------|-----------------------|
| Birinci parça   | İkinci parça | Yön                   |
| 1               | 4            | Akış olmayan bağlantı |
| 3               | 4            | Akış olmayan bağlantı |
| 2               | 3            | İçeri                 |
| 5               | 3            | İçeri                 |
| 0               | 6            | Ağ başlangıcı/Kaynak  |
| 6               | 1            | Akış olmayan bağlantı |

Akış ilişkileri, 4 adet öznitelik ile verilmektedir. Birinci parça, “COM\_ID\_1”; ikinci parça, “COM\_ID\_2”; yön tanımı, “character string” olarak “DIR\_TEXT”, sayısal olarak ise “DIRECTION” isimli öznitelikler ile verilmektedir [67].

**Anlamlı Hidrografik Parçalar Boyunca Akış Sıralaması:** AAP akış ilişkileri (içeri, dışarı vs.) akış aşağıya doğru, ardışık sıra numaraları ile sıralanmaktadır. AAP ile AKP bağlantıları da kıyı çizgisi boyunca, ardışık sıra numaraları ile sıralanmaktadır (Şekil 3.20). Sıralama, “SEQUENCE” isimli öznitelik ile verilmektedir (Çizelge 3.8).



Şekil 3.20 AAP ve AKP'ler boyunca akış ilişkilerinin sıralanması [67]

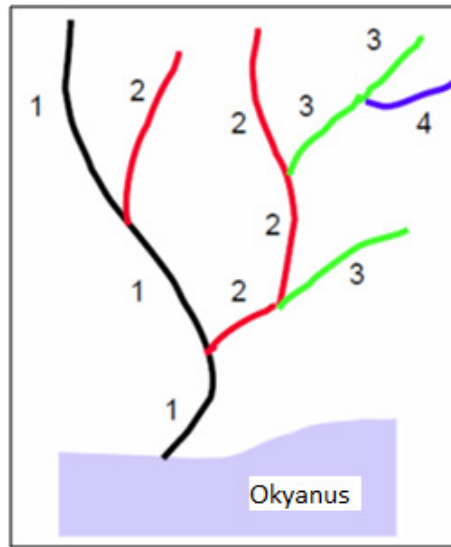
Çizelge 3.8 AAP ve AKP'ler boyunca akış ilişkilerinin sıralanması [67]

| Akış İlişkileri |              |                       |      |
|-----------------|--------------|-----------------------|------|
| Birinci parça   | İkinci parça | Yön                   | Sıra |
| 1               | 3            | İçeri                 | 0    |
| 2               | 3            | İçeri                 | 0    |
| 5               | 3            | Dışarı                | 1    |
| 5               | 3            | İçeri                 | 3    |
| 4               | 3            | İçeri                 | 2    |
| 3               | 8            | Akış olmayan bağlantı | 0    |
| 0               | 6            | Ağ başlangıcı/Kaynak  | 0    |
| 6               | 8            | Akış olmayan bağlantı | 1    |
| 7               | 8            | Akış olmayan bağlantı | 0    |

Düzye güzergahı (level path), kaynak ile ağız arasını veya kaynak ile diğey bir su akışı arasını izler. Düzye farkı (delta level), akış ilişkileri arasında ana su akış yollarını tanımlamaktadır. Akarsu düzyeyi (stream level) de bir AAP'nin ait olduğı ana yolu tanımlamaktadır. Coğrafi isimler genellikle bu yolları izlemektedir [67].

### 3.3 Akarsu Düzyeyleri

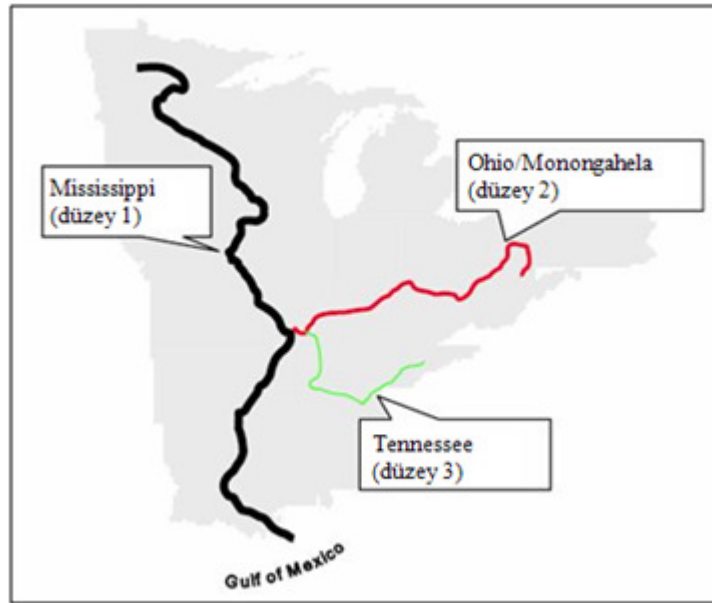
Akarsu düzyeyi, bir akarsu ağında her su akış güzergahını hiyerarşik olarak tanımlayan bir koddur. Akarsu düzye değeryleri, anlamlı akarsu parçaları arasındaki akış-yukarı ilişkilerin hesaplanması ile belirlenmektedir. Şekil 3.21'de görüldüğü gibi, akarsu düzyeyi, akarsu ağı bitiş (son) noktası (ağız) tespit edilerek atanmaktadır. Akışın bitişindeki AAP'den itibaren kaynağıya doğru (akış yukarı) ana akış yolu üzerindeki AAP'lere en düşük değey atanmaktadır. Bu yolda son bulan (bu yola bağlanan) AAP'lere ise kendi kaynaklarına kadar düzye bir artırılarak akarsu düzye değeyi atanmaktadır. Bu işlem, tüm AAP'lere düzye değeryleri atanana kadar devam etmektedir [67]. Bu sistem, Horton ve Strahler akış yukarı sıralaması olarak tanımlanmaktadır [35], [36].



Şekil 3.21 Akarsu ağında akarsu düzyeylerinin atanması [67]

Şekil 3.22'de, Mississippi nehri Meksika Körfezi'nde son bulmaktadır. Ağızdan kaynağıya kadar nehrin ana akış yolunu izleyen AAP'ye 1 değeyi atanmaktadır. Mississippi nehrine bağlanan her kolun (Ohio/Monongahela nehri gibi) ana akış yolunu izleyen AAP'lerine kendi kaynaklarına kadar 2 değeyi atanmaktadır. 2 numaralı kollara bağli her kolun

(Tennessee nehri gibi) ana akış yolunu izleyen AAP'lerine ise kendi kaynaklarına kadar 3 değeri atanmaktadır (Şekil 3.22).



Şekil 3.22 Mississippi nehri boyunca akarsu düzeyleri [67]

İdeal olarak, her ana kol en büyük hacimli su akışını izlemektedir. Anlamlı hidrografik parçalar 1.0 ve 2.0 sürümünde (Reach File Version 1.0 ve 2.0) kodlanan akarsu düzeyleri, akış hacim verisine dayanmaktadır. Bu düzey değerleri mümkün olduğunca korunmaktadır. Akış hacim verisi mevcut olmayan AAP'ler için, aynı coğrafi isimli güzergah, düzeyi belirlemek için kullanılmaktadır. Coğrafi isimler mevcut olmadığında, en uzun ve düz (straightest) güzergah, ana yolu belirlemek için kullanılmaktadır. Bu kural da yeterli olmazsa, akış yukarı bakarak en sağda olan AAP, ana yolun devamı olarak varsayılmaktadır. Akarsu düzeyi "LEVEL" isimli bir öznitelik ile verilmektedir. "-9998" belirsiz değeri ifade etmektedir. Bu değer, akış ilişkilerinin belirlenemediği veya kodlanmadığı durumlarda kullanılmaktadır. Bu değer, AKP'ye de atanmaktadır [67].

### 3.4 Anlamlı Akarsu Parçaları Akış İlişkileri Arasında Akarsu Düzeyleri İzleme

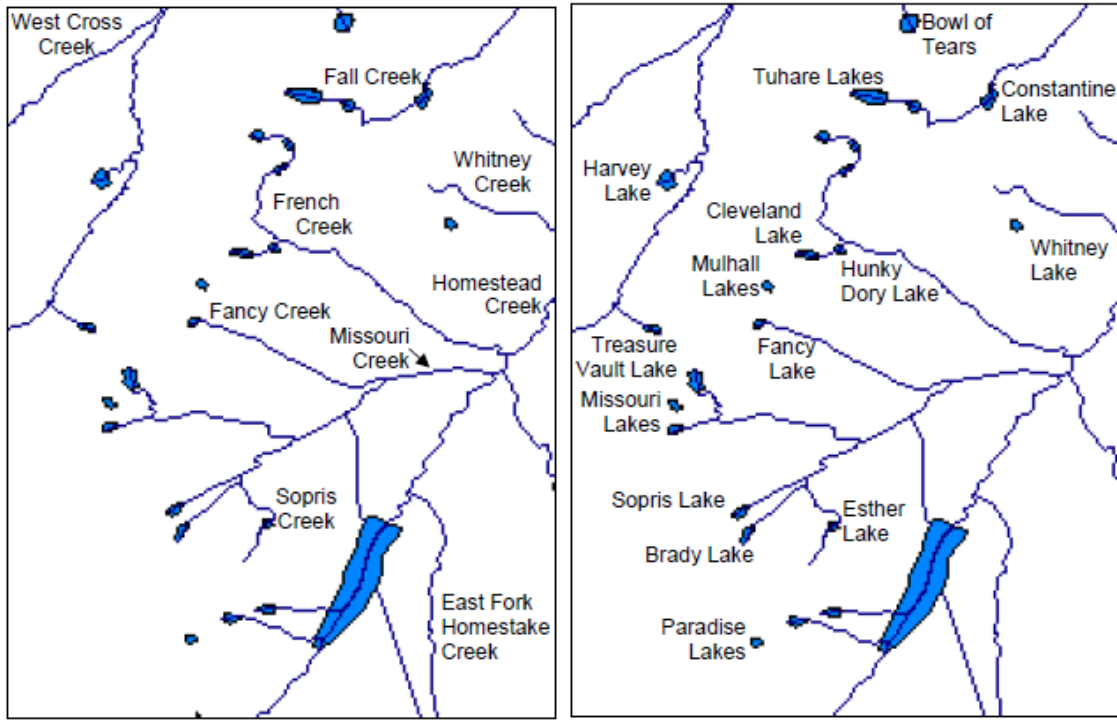
Ana güzergahların tespit edilebilmesi için akarsu düzeyleri arasındaki farktan da faydalanılmaktadır. Bu değer belirlenmesi için ikinci AAP düzey değerinden, birinci AAP düzey değeri çıkarılmaktadır. Farkın sıfır olması iki AAP'nin aynı ana güzergah üzerinde olduğunu göstermektedir. Farkın "-9999" olması; kaynakları, ağızları veya akış

Çizelge 3.9 Akış ilişkileri kullanarak kodlanan akarsu düzey değerlerinin farkları [67]

47

### 3.5 Coğrafi İsimler

AAP'ler ve nesneler coğrafi isim taşıyabilmektedir. İsimler, USGS Coğrafi İsimler Bilgi Sistemi'ndeki Ulusal Coğrafi İsimler Veri Tabanı'ndan alınmıştır. Coğrafi isimler, "NAME"; bir isim için kodlanan sekiz basamaklı tanımlayıcı ise "GNIS\_ID" isimli öznitelikler ile verilmektedir [67].



Şekil 3.24 AAP ve ASKP'lerin coğrafi isimleri [67]

### 3.6 Hidrolojik Birimlerin Organizasyonu ve Örnekler

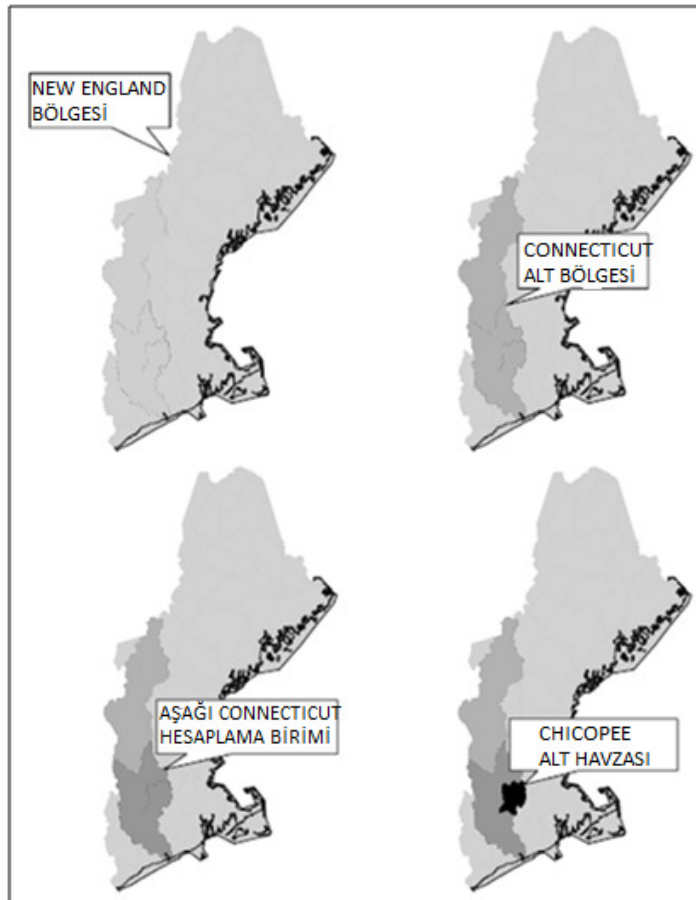
Bir hidrolojik birim, ABD'yi havzalara ayıran ve tanımlayan hiyerarşik sistemin bir üyesidir. Dört adet hidrolojik hiyerarşi düzeyi bulunmaktadır. Bunlar [67]:

- **Birinci düzey:** Ana akarsu drenaj alanını (örneğin, Missouri bölgesi) veya drenaj alanlarının birleşimini (örneğin, Teksas-Gulf bölgesi) içeren bölgeler düzeyidir. 21 bölge bulunmaktadır.
- **İkinci düzey:** Toplam 21 bölgenin 222 alt bölgeye ayrıldığı düzeydir. Her bir alt bölge, bir nehrin ve kollarının bir kısmını, bir veya daha fazla kapalı havzayı içermektedir.

- **Üçüncü düzey:** Toplam 222 alt bölgenin 379 hesaplama birimine (accounting units) ayrıldığı düzeydir.
- **Dördüncü düzey:** Toplam 379 hesaplama biriminin 2,267 alt havzaya ayrıldığı düzeydir. Çoğu alt havza 700 mil kareden ( $1,813 \text{ km}^2$ ) büyüktür.

Sekiz basamaklı hidrolojik birim kodu, dört düzeyi tanımlamaktadır. İlk iki basamak, bölgeyi; ilk dört basamak, alt bölgeyi; ilk altı basamak, hesaplama birimini ve sekiz basamak ise alt havzayı tanımlamaktadır. Şekil 3.25’de, New England bölgesi hidrolojik birimleri gösterilmiştir. Bu bölgeye ait hidrolojik birim kodları aşağıda verilmiştir [67].

- 01: New England bölgesi,
- 0108: Connecticut alt bölgesi,
- 010802: Aşağı Connecticut hesaplama birimi,
- 01080204: Chicopee alt havzası.



Şekil 3.25 New England bölgesi hidrolojik birimleri [67]

### 3.7 Veri Toplama Koşulları

Hiç şüphesiz her ülkede genelleştirme ile ilgili ortaya çıkan kurallar, her ülkenin kendi topografik harita üretim sürecini yansıtmaktadır. Amerika'da 1884 yılında, standart topografik harita üretimine başlanması ile ilgili yapılan ilk kongreden, günümüz USGS topografik veri tabanlarına dek geçen süreçte, genelleştirme ile ilgili standartlar ve yaklaşımlar geliştirilmeye çalışılmıştır.

Verinin tamlığı veri kaynaklarının içeriğine -genellikle USGS topografik haritalarına- bağlıdır. Ölçek ve okunabilirlik kısıtlamalarından dolayı, kaynak haritadaki nesneler eleme veya genelleştirme işlemine tabi tutulmuştur.

USGS UHVS standartları [69]'a göre, 1:24,000 ölçekli akarsu ağları için veri toplama koşulları aşağıdaki gibidir.

- Eğer akarsu/nehir, göl/gölcük veya pınar nesnelerinden çıkıyorsa veri tabanına aktarılmıştır.
- Eğer akarsu/nehir en uzun ekseni  $\geq 0.76$  km ise veri tabanına aktarılmıştır.
- Eğer akarsu/nehir daimi ve kurak bir bölgede ise veri tabanına aktarılmıştır.

USGS UHVS standartları [69]'a göre, 1:100,000 ölçekli akarsu ağları için veri toplama koşulları aşağıdaki gibidir.

- Eğer akarsu/nehir daimi ve göl/gölcük veya pınar nesnelerinden çıkıyorsa veri tabanına aktarılmıştır.
- Eğer akarsu/nehir mevsimsel ve kesin olarak konumlandırılabilirse ve göl/gölcük veya pınar nesnelerinden çıkıyorsa veri tabanına aktarılmıştır.
- Eğer akarsu/nehir daimi ve en uzun ekseni  $\geq 1.6$  km ise veri tabanına aktarılmıştır.
- Eğer akarsu/nehir mevsimsel ve kesin olarak konumlandırılabilirse ve kurak bir bölgede değilse ve en uzun ekseni  $\geq 1.6$  km ise veri tabanına aktarılmıştır.
- Eğer akarsu/nehir mevsimsel ve kesin olarak konumlandırılabilirse ve kurak bir bölgede ise ve en uzun ekseni  $\geq 3.05$  km ise veri tabanına aktarılmıştır.

#### 4.1 İstatistiksel Yöntemler

##### 4.1.1 Tanımlayıcı İstatistik

Temeli olasılık kuramına dayanan matematiksel istatistiğin konusu yığın nesneler ya da ölçülerdir. Rastgele seçme ve ölçme (gözlem) istatistikte rastgele deney olarak adlandırılır. Sınırlı sayıda yinelenen bir rastgele deneyin sonuçları, olanaklı tüm sonuçları kapsayan toplumdaki (deney ortamı) bir örneklem (örnek uzay) oluşturur. Bir örneklem, ortalama değeri ve standart sapması ile tanımlanır. Bir toplumdaki türetilen çok sayıda farklı rastgele örnekleme ilişkin istatistikler, bir örneklemde ötekine değişir. Bu nedenle istatistik, rastgele örnekleme bağlı bir rastgele değişkendir. Bu tezde kısaca değişken olarak ifade edilecektir. Bir örnekleme ilişkin istatistikler, ilgili toplumun beklenen değeri ve standart sapmasının (toplumun parametreleri) kestiricileridir [70].

İstatistiksel özelliklerin tanımlanması, problemin çözümünde kullanılacak yöntemlerin belirlenmesinde önemli rol oynamaktadır. Tanımlayıcı istatistikler kategorisi içinde; değişken türünün tanımlanması, ortalama, medyan (ortanca değer) ve mod (en çok tekrar eden değer) gibi merkezi eğilim ölçütleri, standart sapma ve varyans gibi ortalamadan sapma ölçütleri ile çarpıklık ve basıklık gibi normalden sapma ölçütleri yer almaktadır [71].

Değişkenleri başlıca iki ana grupta toplamak mümkündür.

**Sürekli değişken:** Eğer bir değişken belli değerler arasında sonsuz değer alabiliyorsa, bu tür değişkenlere sürekli değişken denir. Bunlar tam sayı ile ifade edilebileceği gibi ondalıklı sayılarla da ifade edilebilir [72]. Akarsu uzunluğu, sürekli bir değişkendir.

**Kesikli değişken:** Eğer bir değişken belli sınırlar arasında ancak belirli değerleri alabiliyorsa, bu tür değişkenlere kesikli değişken denir. Genellikle tam sayı ile ifade edilirler [72]. Akarsu tipi (mevsimsel/daimi dere), kesikli bir değişkendir.

Ölçek, gözlem sonuçlarının belirli kurallara göre gösterimidir ve temel olarak 4 tip ölçek vardır. Ölçek türleri, aynı zamanda veri türlerini de ifade eder. Bunlar:

**Sınıflama (nominal) ölçeği:** Nesnelerin benzerliklerine göre sayı ya da işaretlerle gruplandırılmasıdır. İşaret ya da sayılar veriyi nitel olarak ayırır; bunun ötesinde hiçbir anlamları yoktur. Akarsu tipine göre ayırmada; daimi derelere “1”, mevsimsel derelere “0” sayıları atanarak iki sınıf oluşturulabilir.

**Sıralama (ordinal) ölçeği:** Sınıflama ölçeğinden farklı olarak nesnelerden birinin diğerine göre nerede olduğu belirtilir. Gözlem değerleri, belli bir ölçüte göre sıraya dizilerek (hiyerarşi oluşturularak) sıralama değeri alırlar. Bir akarsuyun Horton sıralamasında hiyerarşi değeri “5” olabilir.

**Eşit aralıklı (interval) ölçek:** Gözlem değer aralıklarının eşit olduğu kabulü üzerine kurulmuştur. Belli bir başlangıç (sıfır) noktası vardır; ancak bu, yokluk anlamına gelmemektedir. Su sıcaklığı, eşit aralıklı ölçektir.

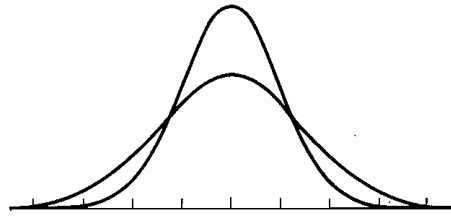
**Oranlı (ratio) ölçek:** Gözlem değerleri arasında oranlamalar yapılabilmektedir ve sıfır, gerçek bir değer olup, yokluk ifade eder. Akarsu uzunluğu, oranlı ölçektir.

Kesikli değişkenler, sınıflama ve sıralama ölçeğiyle; sürekli değişkenler ise eşit aralıklı ve oranlı ölçekle ölçülmüş değişkenlerdir. Sınıflama ve sıralama ölçeğinde kesikli değişkenler, kategorik veriler olarak isimlendirilmektedir.

#### **4.1.2 Normal Dağılım**

Normal dağılımlı değişkenler, ortalama değere göre simetriktir ve dağılımları “çan eğrisi” biçimindedir. Normal dağılım, doğadaki birçok değişkenin dağılımının yansıdığı hipotetik bir dağılımdır [73]. Rastgele deneylere ve ölçmelere ilişkin değişkenler ve

istatistiksel test büyüklükleri, genellikle normal dağılımlı ya da yaklaşık normal dağılımlıdır [70]. Bu nedenle, sürekli değişkenlerin ortaya koyduğu olasılık dağılımlarının en önemlisi normal dağılımdır. Dağılımın üst ve alt limitleri artı ve eksi sonsuzdur. Ortalama ve varyans ( $\sigma^2$ ) değerlerine göre çeşitli normal dağılımlar mevcuttur. Şekil 4.1’de aynı ortalamaya sahip, farklı varyanslı iki farklı normal dağılım gösterilmiştir. Varyansları farklı olmasına rağmen eğrinin biçimi değişmemektedir. Bir normal dağılımın ortalaması ve varyansı biliniyorsa, normal dağılım eğrisinin biçimi tam olarak bilinebilmektedir [73].



Şekil 4.1 Aynı ortalamaya sahip farklı varyanslı iki farklı normal dağılım [73]

Anakütleye ait standart sapma ( $\sigma$ ):

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n}} \quad (4.1)$$

Örnekleme ait standart sapma ( $s$ ):

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} \quad (4.2)$$

$x_i$ : Değişken

$\mu$ : Anakütlenin ortalama değeri

$\bar{x}$ : Örneklemin ortalama değeri

$n$ : Gözlem sayısı

$\sigma^2$ : Varyans

Normal dağılımlı  $X$  değişkeninin beklenen değeri  $\mu = E(X)$  ve varyansı  $\sigma^2 = E((X - \mu)^2)$  ya da standart sapması  $\sigma$  olsun. Normal dağılımlı  $X$  kısaca,  $X \in N(\mu, \sigma)$  biçiminde gösterilir. Normal dağılımın olasılık yoğunluk fonksiyonu (4.3) eşitliğinde verilmiştir.

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\left(-\frac{1}{2}\right)\left[\frac{x-\mu}{\sigma}\right]^2} \quad -\infty < x < +\infty \quad (4.3)$$

$X$  değişkene ilişkin bir değer ya da rasgele deney sonucu  $x$  olduğuna göre normal dağılımın kümülatif dağılım fonksiyonu,  $X < x$  ya da  $-\infty < X < x$  örnekleminin olasılığına eşittir [70]:

$$F(x) = P(X < x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi\sigma^2}} e^{\left(-\frac{1}{2}\right)\left[\frac{x-\mu}{\sigma}\right]^2} dx \quad (4.4)$$

Normal dağılımın özellikleri şöyle sıralanabilir:

- a) Normal dağılım eğrisi altında kalan alan bire eşittir. Bir başka deyişle,  $f(x)$  olasılık yoğunluk fonksiyonunun  $-\infty$ 'dan  $+\infty$ 'a kadar integrali "1" değerini verir.

$$\int_{-\infty}^{+\infty} f(x) dx = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{\left(-\frac{1}{2}\right)\left[\frac{x-\mu}{\sigma}\right]^2} dx = 1 \quad (4.5)$$

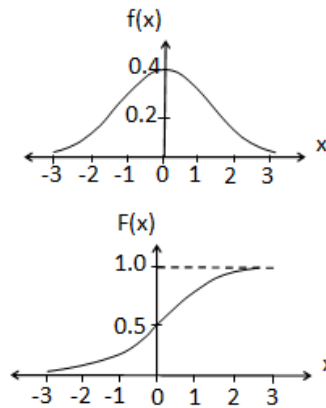
- b) Normal dağılım ortalamaya göre simetriktir. Aritmetik ortalaması, modu ve medyanı eşittir.

$$\int_{-\infty}^{\mu} f(x) dx = \int_{\mu}^{+\infty} f(x) dx = 0.5 \quad (4.6)$$

- c) Gözlemlerin %68.26'sı ortalama ( $\mu$ ) ile  $\pm 1\sigma$  aralığında, %95.44'ü  $\pm 2\sigma$  aralığında ve yaklaşık %100'ü  $\pm 3\sigma$  aralığında bulunur.

- d) Standart normal dağılım, ortalaması 0 ve standart sapması 1 olan dağılımdır.

Standart normal dağılımın olasılık yoğunluk ve kümülatif dağılım fonksiyonları Şekil 4.2'de gösterilmiştir [71], [72].



Şekil 4.2 Standart normal dağılımın olasılık yoğunluk fonksiyonu (üstte) ve kümülatif dağılım fonksiyonu (altta) [70]

#### 4.1.3 Normalizasyon

Kümeleme ve sınıflandırmanın ön işlem adımlarından biri olan normalizasyon yöntemlerinden hangisinin kullanılacağı, verinin istatistiksel tanımı ve uygulamanın amacına göre değişiklik göstermektedir. Değişkenlerin ortalama ve varyansları birbirinden önemli ölçüde farklı olduğu durumda, büyük ortalama ve varyansa sahip değişkenlerin diğerleri üzerindeki etkisi daha fazla olur ve onların rollerini önemli ölçüde azaltır. Ayrıca değişkenlerin sahip olduğu çok büyük ve çok küçük değerler de analizin sağlıklı biçimde gerçekleştirilmesini olumsuz etkiler. Bu nedenle bir dönüşüm yöntemi uygulayarak söz konusu değişkenlerin normalize edilmesi veya standartlaştırılması uygun bir yol olur. Veri dönüşüm yöntemlerinden bazıları şunlardır: Maksimum-minimum normalizasyonu, z-skor standardizasyonu, logaritmik dönüşüm, lojistik dönüşüm ve histogram eşitlemedir. Maksimum-minimum normalizasyon işlemi, değişken değerleri arasındaki oranlar değişmeden, değerlerin 0-1 aralığına getirilmesi işlemidir. Z-skor standardizasyonu; değişken değerlerinin aritmetik ortalaması "0" ve varyansı "1" olacak şekilde yeni değerlere dönüştürülmesi esasına dayanır [74]. Maksimum-minimum ve z-skor yöntemleri doğrusal yöntemler olup, verinin dağılımını değiştirmemektedir. Buna karşın, logaritmik dönüşüm, lojistik dönüşüm ve histogram eşitleme doğrusal olmayan yöntemlerdir ve verinin dağılımını değiştirmektedir. Değişken değerlerinin üssel olarak dağıldığı durumlarda, logaritmik dönüşüm uygun olacaktır. Bu yöntemin dezavantajı, değeri "0" veya negatif olan veriye uygulanamayışıdır. Bu dezavantaj, bazı çalışmalarda veri pozitif hale dönüştürülerek giderilmeye çalışılmaktadır. Lojistik dönüşüm, lojistik fonksiyonları kullanarak veriyi 0-1 aralığına dönüştürmektedir. Bu dönüşüm, ortalama etrafında daha çok veya daha az doğrusallık gösterirken, uç değerlerde yumuşak bir doğrusal olmama özelliği taşımaktadır. Histogram eşitleme yönteminde ise veriler büyüklüklerine göre sıralanır ve her bir değer sıra numarası ile değiştirildikten sonra, sıra numaraları 0-1 aralığına dönüştürülür. Hem sürekli hem de kesikli değişkenler için kullanışlıdır [75].

Maksimum-minimum normalizasyon ölçütleri (4.7) eşitliği ile verilmektedir.

$$X'_i = \frac{X_i - X_{min}}{X_{maks} - X_{min}} \text{ kar ölçütü; } X'_i = \frac{X_{maks} - X_i}{X_{maks} - X_{min}} \text{ zarar ölçütü} \quad (4.7)$$

Burada,  $X'$  normalize edilmiş değerleri,  $X$  değişken değerlerini,  $X_{min}$  en küçük değişken değerini ve  $X_{maks}$  en büyük değişken değerini ifade etmektedir. Maksimum-minimum normalizasyonu kar ölçütünde, değişken değerleri, küçükten büyüğe aralarındaki oran değişmeden 0-1 aralığına yerleştirilir. Zarar ölçütünde ise büyükten küçüğe aralarındaki oran değişmeden 0-1 aralığına yerleştirilir. Kar ölçütünde minimum değişken değeri “0” değerini alırken, maksimum değişken değeri “1” değerini almaktadır. Zarar ölçütünde ise minimum değişken değeri “1” değerini alırken, maksimum değişken değeri “0” değerini almaktadır [76].

Bir  $X$  değişkeninin standart Z-skor değeri:

$$Z = \frac{X - \mu}{\sigma} \quad (4.8)$$

Burada,  $\mu$  anakütlenin ortalamasını ve  $\sigma$  standart sapmasını ifade etmektedir [73], [75].

#### 4.1.4 Hipotez Testleri

Hipotezlerin toplumdan seçilen bir örneklem yardımıyla incelenmesine “hipotezlerin testi” denir. Hipotez testinde iki hipotez karşılaştırılır. Karşılaştırılan hipotezlerden birine sıfır hipotezi ( $H_0$ ), diğerine ise karşıt hipotez ( $H_1$ ) denir. Bu hipotezlerden biri diğerinin alternatifidir. Bu nedenle hipotezler öyle kurulmalıdır ki, sıfır hipotezi kabul edildiğinde, alternatif hipotez reddedilmelidir. Sıfır hipotezinin reddedilmesi halinde ise, karşıt hipotez kabul edilmelidir.

Hipotez testlerinde, her zaman test edilen hipotez sıfır hipotezidir. Sıfır hipotezi, değişkenin önceden belirlenmiş değerinin değişmediğini ifade etmektedir. Sıfır hipotezine ilişkin bir karar verebilmek için örneklemin test edilmesi gerekir [72].

$X$  bir rastgele örneklem,  $X'$  örneklem aralığı ve  $R$  ret bölgesi ( $R \subset X'$ ) olmak üzere, eğer  $X \in R$  ise sıfır hipotezi reddedilmekte, değilse kabul edilmektedir [77].

$$X \in R \Rightarrow H_0 \text{ ret}; X \notin R \Rightarrow H_0 \text{ kabul} \quad (4.9)$$

Hipotez testlerinde anlamlılık düzeyinin veya güven düzeyinin belirtilmesi gerekir. Olasılık yoğunluk fonksiyonunun sağında ve solunda eşit kısımlar halinde ayrılan  $\alpha$  olasılığına, testlerde anlamlılık düzeyi adı verilir (Şekil 4.3). Anlamlılık düzeyi  $\alpha$ , güven

düzeyi ise  $(1 - \alpha)$ 'dır. Normal dağılımlı bir değişkenin  $(1 - \alpha)$  güven düzeyi ile içinde bulunduğu aralığın olasılık değeri;  $\mu$  ortalama değerine göre simetrik sınırlar ya da iki yanlı güven sınırları kullanılarak,

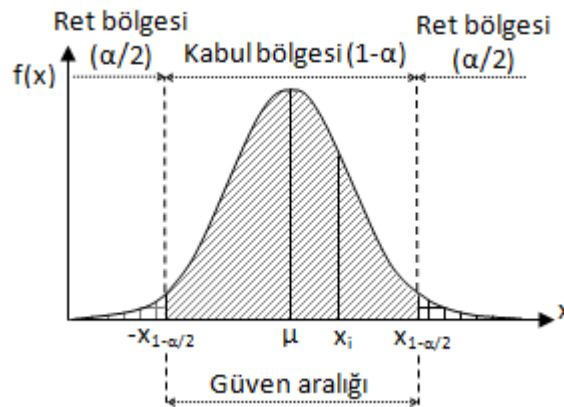
$$P(-x_{1-\alpha/2} < x < x_{1-\alpha/2}) = \int_{-x_{1-\alpha/2}}^{x_{1-\alpha/2}} f(x) dx = 1 - \alpha \quad (4.10)$$

Tek yanlı güven sınırı kullanılarak,

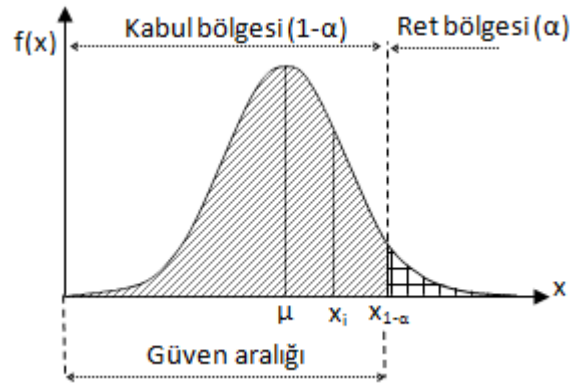
$$P(x < x_{1-\alpha}) = \int_0^{x_{1-\alpha}} f(x) dx = 1 - \alpha \quad (4.11)$$

eşitliklerinden elde edilir [70].

İki veya tek yanlı güven aralığından hangisinin seçilmesi gerektiği, probleme göre değişir. Seçilen anlamlılık düzeyi ise olasılık dağılımında kabul ve red bölgelerinin belirlenmesini sağlar. Test büyüklüğü ( $x_i$ ), güven düzeyi  $(1 - \alpha)$  için verilen iki yanlı  $(-x_{1-\alpha/2}$  ve  $x_{1-\alpha/2})$  güven aralığında veya tek yanlı  $(0$  ve  $x_{1-\alpha})$  güven aralığında ise, sıfır hipotezi kabul edilir (kabul bölgesi). Test büyüklüğü ( $x_i$ ) güven aralığı dışında bir değer alıyorsa, sıfır hipotezi reddedilir ve karşıt hipotez kabul edilir (ret bölgesi). Güven düzeyi  $(1 - \alpha)$  için genellikle 0.95 ya da 0.99 olasılık değerleri öngörülür. Buna göre anlamlılık düzeyi ( $\alpha$ ); 0.05 ya da 0.01 olur. İki ve tek yanlı hipotez testlerinde güven aralığı ile belirlenen kabul ve ret bölgeleri Şekil 4.3 ve Şekil 4.4'de gösterilmiştir.



Şekil 4.3 İki yanlı güven düzeyi için  $x$  test büyüklüğüne ilişkin kabul ve ret bölgeleri [70]

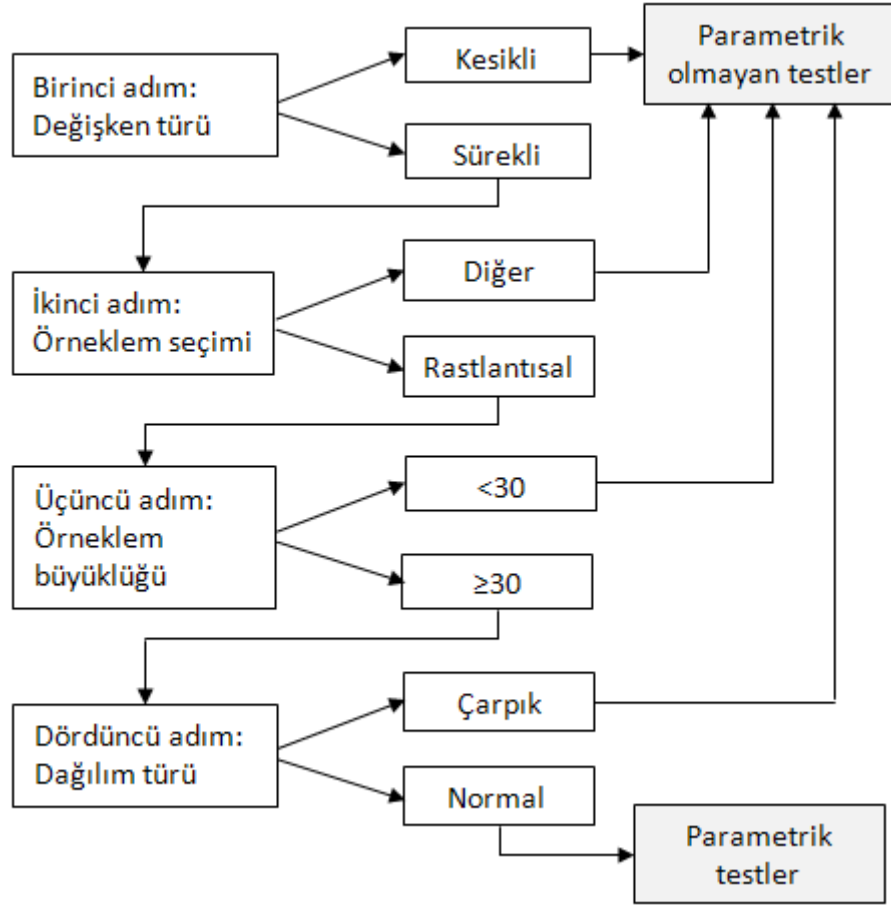


Şekil 4.4 Tek yanlı güven düzeyi için  $x$  test büyüklüğüne ilişkin kabul ve ret bölgeleri [70]

Hipotez testleri, parametrik ve parametrik olmayan hipotez testleri olarak ikiye ayrılmaktadır. İstatistiksel analiz yapılmadan önce, değişkenlerin kesikli (sınıflama ve sıralama ölçeği) ya da sürekli (eşit aralıklı ve oranlı ölçek) olup olmadığına, örneklem seçimine, örneklem büyüklüğüne ve dağılımın türüne bakılmalıdır.

Parametrik testler (t-testi, varyans analizi gibi) modelden rastlantısal çekilen örnekleme ilişkin varsayımlar üretir. Parametrik testlerin hemen hemen hepsinin uygulanabilmesi için, örneklemin yeterince büyük olması ( $\geq 30$ ) ve verilerin normal dağılıma uyması gerekir.

Parametrik olmayan testler ise, bu gibi katı gereksinimlere ihtiyaç duymaz. Parametrik olmayan testler; sınıflama ve sıralama ölçeği için ya da normal dışı dağılım gösteren veriler için daha uygundur. Çok küçük örneklem için ( $< 30$ ) ve verinin parametrik testlerin varsayımlarına uygun olmadığı durumlarda, parametrik olmayan testler daha kullanışlıdır. Parametrik olmayan testlerin dezavantajı ise parametrik testlerden daha az duyarlı olmalarıdır [71], [72].



Şekil 4.5 Test türünün belirlenmesi [78]

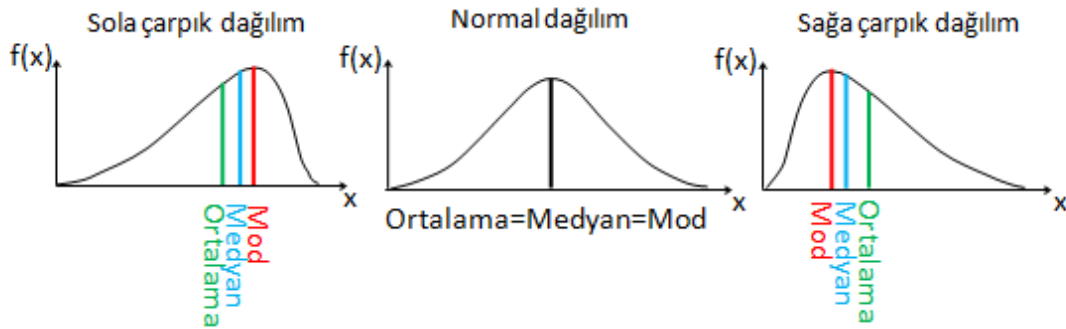
#### 4.1.5 Normal Dağılıma Uygunluğun Araştırılması

Özniteliklerin (akarsu uzunluğu vd.) normal dağılımlı olup olmadığı aşağıdaki yöntemlerle belirlenebilir.

- Çarpıklık (skewness) ve basıklık (kurtosis) katsayısı
- Histogram grafiği, kantil-kantil grafiği (Quantile-Quantile (Q-Q) plot) ve kutu grafiği (box plot)
- Kolmogrov-Smirnov uygunluk (iyi uyum) tek örneklem testi

Tam simetri durumunda aritmetik ortalama, mod ve medyan birbirine eşit olup, çarpıklık değeri sıfır olacaktır. Eğer bu eşitlik bozulursa dağılım çarpık hale gelecektir. Çarpıklık, dağılımın ortalama etrafında simetriden ne kadar saptığını yani, verilerin simetrisini belirleyen ölçüttür. Çarpıklık arttıkça mod ve aritmetik ortalama birbirinden uzaklaşacaktır. Şekil 4.6'da görüldüğü gibi, eğer ortalama medyandan küçük ise,

gözlemlerin değerlerinin dağılımı sola (negatif) çarpık; büyük ise, sağa (pozitif) çarpık olur [71].



Şekil 4.6 Normal ve çarpık dağılımlar

#### 4.1.5.1 Çarpıklık ve Basıklık Katsayısı

Fisher'in çarpıklık katsayısı, normal dağılıma uygunluğun araştırılmasında kullanılabilir [79], [80]. Bu değer pozitif çıkması verilerin sağa çarpık olduğunu, negatif çıkması ise sola çarpık olduğunu gösterir. Fisher'in çarpıklık katsayısı (4.12) eşitliği ile verilmektedir [80].

$$Fisher'ın\ çarpıklık\ katsayısı = \frac{n \sum_{i=1}^n (x_i - \mu)^3}{(n-1)(n-2)\sigma^3} \quad (4.12)$$

Basıklık, normal dağılım eğrisinin ne kadar sivri (leptokurtic) veya basık (platykurtic) olduğunu gösterir. Basıklık değeri hesabının amacı, değişkenlerin ortalama etrafında nasıl bir dağılım gösterdiğini ortaya koymaktır. Tam bir çan eğrisinin basıklık değeri sıfırdır. Sıfıra yakın bir basıklık, normal dağılıma yakın bir biçime sahiptir. Basıklık değeri, pozitif ise eğri normale göre daha sivri, negatif ise daha basıktır. Fisher'ın basıklık katsayısı (4.13) eşitliği ile verilmektedir [80].

$$Fisher'ın\ basıklık\ katsayısı = \frac{n(n+1) \sum_{i=1}^n (x_i - \mu)^4}{(n-1)(n-2)(n-3)\sigma^4} - \frac{3(n-1)^2}{(n-2)(n-3)} \quad (4.13)$$

Çarpıklık katsayısının, çarpıklık standart hatasına ( $SH_c$ ) (eşitlik 4.14) bölünmesi, çarpıklık değerini; basıklık katsayısının, basıklık standart hatasına ( $SH_b$ ) (eşitlik 4.15) bölünmesi ise basıklık değerini vermektedir. Elde edilen değerler (%5 anlamlılık düzeyine göre), ortalamadan  $\pm 1.96$  standart sapma aralığında ise veriler normale

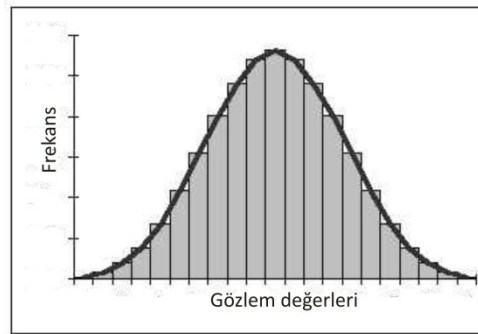
yakın dağılmaktadır. Başka bir deyişle, değerler  $\pm 2$  aralığında ise ilgili değişken normal dağılımlı kabul edilmektedir [71].

$$SH_{\zeta} = \sqrt{6n(n-1)/(n-2)(n+1)(n+3)} \quad (4.14)$$

$$SH_b = 2SH_{\zeta} \sqrt{(n^2-1)/(n-3)(n+5)} \quad (4.15)$$

#### 4.1.5.2 Histogram Grafiği, Kantil-Kantil Grafiği ve Kutu Grafiği

Çubuk veya histogram grafikler, gözlemlerin ne kadar sıklıkla tekrar ettiğini gösteren grafiklerdir. Yatay eksen boyunca genellikle sistematik bir biçimde aralıklar belirtilir ve her aralıkla ilgili frekansı temsil etmek üzere dikey çubuklar yerleştirilir. Sürekli değişkene ait çubukların tepe noktaları birleştirilerek elde edilen eğri dağılım hakkında bilgi verir. Histogram eğrisinin tam simetrik olması, verinin tam normal dağılımlı olduğunu göstermektedir (Şekil 4.7).

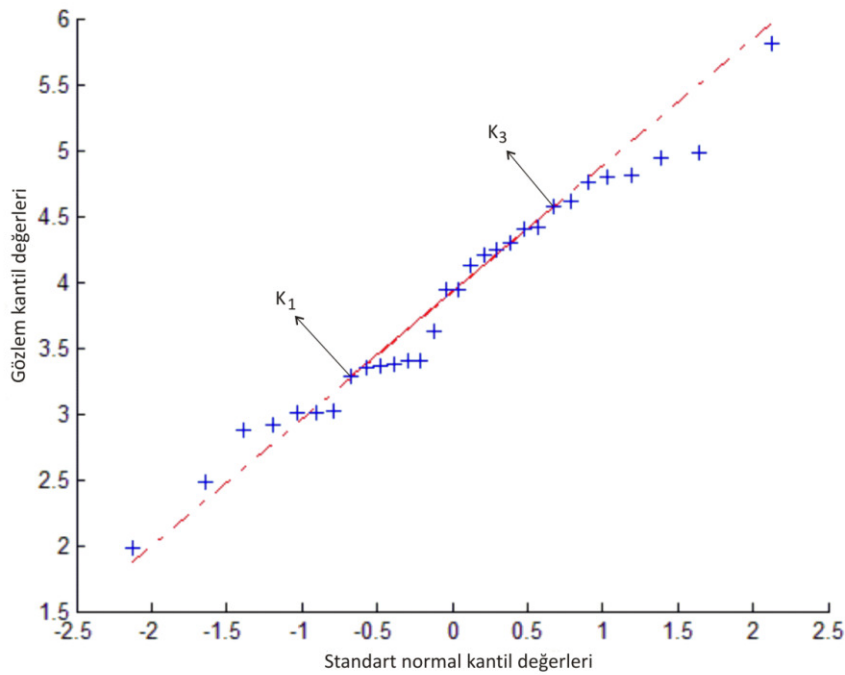


Şekil 4.7 Normal dağılımlı histogram

Q-Q grafiği, gözlem değerlerine ait kantil değerler ile normal dağılıma ait teorik kantil değerleri karşılaştırır. Kantil (quantile), küçükten büyüğe doğru veya büyükten küçüğe doğru sıralanmış veriyi eşit parçalara ayıran değerlerdir. Veri setini iki eşit parçaya ayıran ve tam orta sırada yer alan kantile medyan denir. Veri setini dört eşit parçaya ayıran kantillere ise dörtte birlik veya kartil (quartile) denir. Şekil 4.8’de görünen Q-Q grafiğinde, düşey eksen gözlem kantil değerlerini, yatay eksen standart normal dağılıma ait kantil değerleri, “+” işareti ise her bir gözlem değerini göstermektedir. Birinci dörtte birlik ( $K_1$ ) ve üçüncü dörtte birlik ( $K_3$ ) değerleri sürekli bir çizgi ile birleştirilir ve uç değerlere kadar kesikli bir çizgi ile uzatılır. Veri normal dağılımlı ise, değerlerin bu doğru üzerinde veya etrafında toplanması gerekir.

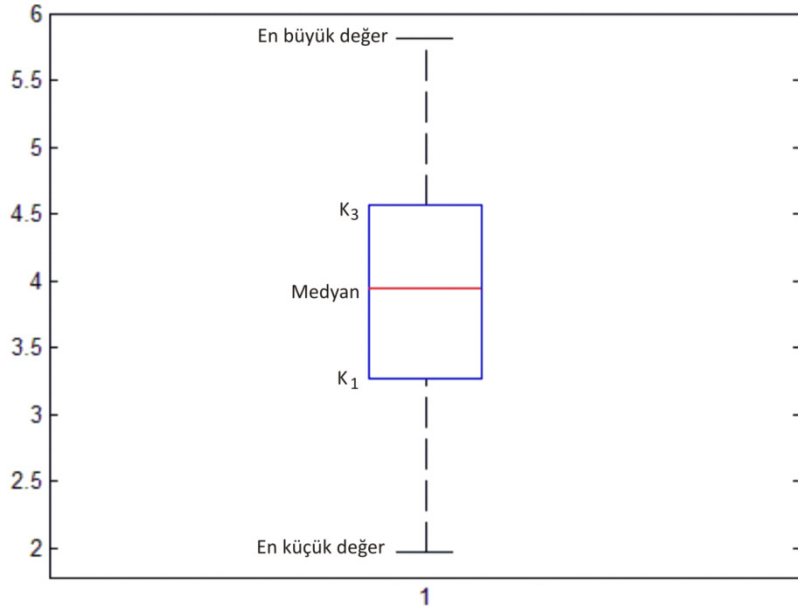
30 adet akarsuya ait “km” birimindeki uzunluklar; en küçük gözlem, birinci dördte birlik ( $K_1$ ), medyan, üçüncü dördte birlik ( $K_3$ ) ve en büyük gözlem değerlerinin belirlenmesi için, küçükten büyüğe doğru sıralanarak aşağıda verilmiştir. Şekil 4.8’de görünen Q-Q grafiği normal dağılımlıdır.

[1.978; 2.479; 2.881; 2.92; 3.007; 3.013; 3.018; 3.276; 3.343; 3.359; 3.374; 3.396; 3.407; 3.627; 3.937; 3.945; 4.12; 4.199; 4.25; 4.3; 4.401; 4.413; 4.571; 4.613; 4.76; 4.799; 4.815; 4.942; 4.975; 5.809]



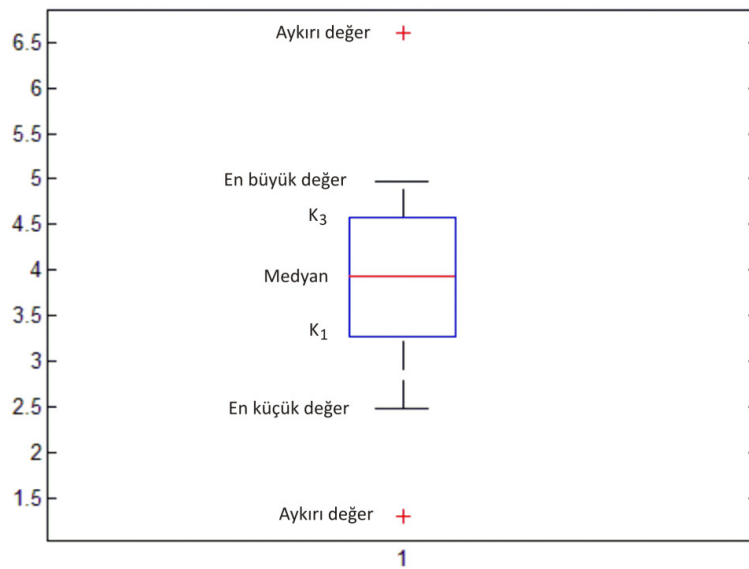
Şekil 4.8 Normal dağılımlı verilerin Q-Q grafiği

Diğer bir grafik yöntem, kutu grafiğidir [81]. Kutu grafiği, verinin aykırı değeri (outlier) olmayan en küçük değer ile en büyük değerler arasında bir kutu ve iki çizgiden (whisker) oluşur. Kutu,  $K_1$  ve  $K_3$  arasına yerleştirilir. Kutunun her iki tarafında, kutudan itibaren aykırı değer olmayan en küçük ve en büyük değere kadar çizgiler (Şekil 4.9’da kesikli çizgi) çizilir. Kutunun içinde medyanın yeri de işaretlenir. Aykırı değerler;  $(K_1 - 1.5 \times (K_3 - K_1))$ ’den küçük veya  $(K_3 + 1.5 \times (K_3 - K_1))$ ’den büyük değerler olarak kabul edilir. Kutu grafiği ile dağılım yorumlanabilir. Normal dağılımlı 30 tane akarsuya ait uzunluklar için verilen kutu grafiği Şekil 4.9’da gösterilmiştir.



Şekil 4.9 Normal dağılım ve kutu grafik

Aykırı değerler kutu grafiği dışında “+” işareti ile gösterilir. Şekil 4.10’da aykırı değerler;  $(3.276 - 1.5 \times (4.571 - 3.276))$ ’den küçük ve  $(4.571 + 1.5 \times (4.571 - 3.276))$ ’den büyük değerlerdir. Yani, 1.33 değerinden küçük ve 6.51 değerinden büyük değerler aykırı değerlerdir. En küçük (1. sırada) ve en büyük (30. sırada) değerler yerine, aykırı değerler (örneğin, 1.3 ve 6.6) kullanılırsa; en küçük değer (2. sırada), en büyük değer (29. sırada) ve aykırı değerler (1. ve 30. sırada), Şekil 4.10’da görüldüğü gibi bir kutu grafiği meydana getirir.



Şekil 4.10 Kutu grafikte aykırı değerlerin gösterimi

#### 4.1.5.3 Kolmogorov-Smirnov Uygunluk Tek Örneklem Testi

Örneklem setinin dağılımı ile kuramsal (beklenen) dağılım arasındaki uyumu araştıran testlere uygunluk veya iyi uyum (goodness-of-fit) testleri adı verilir [82]. Dağılımın normal dağılıma uygunluğunu test etmek amacıyla kullanılan iki yaygın test yöntemi: Kolmogorov-Smirnov (K-S) yöntemi ve Shapiro Wilks (W istatistiği) yöntemidir. Shapiro Wilks yöntemi, genellikle gözlem sayısının 50'nin altında olduğu durumlarda kullanılır. K-S, en sık kullanılan parametrik olmayan uygunluk testidir [83]. Test, gözlem frekans değerleri ile kuramsal frekans değerlerin kümülatif dağılımları arasındaki maksimum farka dayanmaktadır. Gözlem frekans değerlerinin kümülatif dağılımı, kuramsal bir kümülatif dağılıma yeterince yakın değilse, bu kuramsal dağılama uygun değildir [82]. K-S tek örneklem testinde hipotezler şöyle kurulur.

$H_0$ : Gözlem frekans değerleri, beklenen frekans değerlerine uygundur. Gözlemler normal dağılımlıdır.

$H_1$ : Gözlem frekans değerleri, beklenen frekans değerlerine uygun değildir. Fark önemlidir. Gözlemler normal dağılımlı değildir.

Bu tezde, gözlem değerlerinin ( $x$ ) dağılımı, standart normal dağılım ( $\mu = 0, \sigma = 1$ ) ile karşılaştırılmıştır. Gözlemler standart normal dağılıma uygunsa sıfır hipotezi, değilse karşıt hipotez kabul edilir.

Test büyüklüğü ( $D$ ); gözlem değerlerinin ve kuramsal değerlerin kümülatif nispi frekansları (bir aralıktaki frekansın, toplam frekansa oranı) arasındaki mutlak farkın en büyüğüdür.

$$D = \max(|F(x) - G(x)|) \quad (4.16)$$

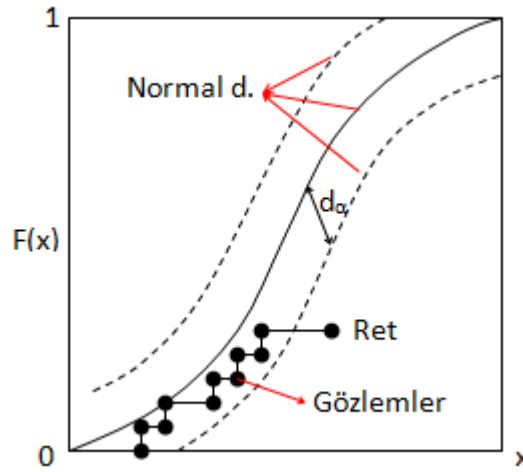
Burada,  $F(x)$  gözlem kümülatif nispi frekans değerini, yani  $n$  adet gözlem içerisinde belirli bir  $x$  değerinden küçük ya da eşit olanların oranını;  $G(x)$  standart normal dağılım (beklenen) kümülatif nispi frekans değerini göstermektedir. Çizelge 4.1'de görünen Smirnov'un kritik tablo değerleri ( $d_\alpha$ ), 35 gözlemden fazla örneklem için ( $n > 35$ ), örneklem ve yığın (kitle veya ana kütle) kümülatif dağılımları mutlak farklarının en büyük değerlerini verir [82]. Test istatistiği tek yanlı testlerde kolayca uygulanabilir [83]. Test büyüklüğü ( $D$ ) ile karşılaştırılarak hipotezler hakkında karar verilir. Eğer

$d_\alpha > D$  ise sıfır hipotezi kabul edilir, yani gözlemler normal dağılımlıdır. Eğer  $d_\alpha < D$  ise sıfır hipotezi reddedilir, yani gözlemler normal dağılımlı değildir. Şekil 4.11’de, kuramsal kümülatif dağılımın (sürekli çizgi),  $d_\alpha$  aralığında sağında ve solunda eğriler (kesikli çizgi) görülmektedir. Eğer  $D$  bu bantın dışına çıkarsa  $\alpha$  anlamlılık düzeyinde sıfır hipotezi reddedilir.

Çizelge 4.1 Smirnov’un kritik tablo değerleri [82]

| Anlam Düzeyleri ( $\alpha$ ) |                         |                         |                         |                         |
|------------------------------|-------------------------|-------------------------|-------------------------|-------------------------|
| 0.20                         | 0.15                    | 0.10                    | 0.05                    | 0.01                    |
| $\frac{1.07}{\sqrt{n}}$      | $\frac{1.14}{\sqrt{n}}$ | $\frac{1.22}{\sqrt{n}}$ | $\frac{1.36}{\sqrt{n}}$ | $\frac{1.63}{\sqrt{n}}$ |

Şekil 4.11’de,  $d_\alpha$  aralığında ve yığının kuramsal kümülatif dağılım (standart normal dağılım) eğrisinin sağında ve solunda yer alan kuramsal kümülatif dağılım eğrileri gösterilmiştir. Eğer  $d_\alpha < D$  ise, gözlem değerlerinin kümülatif nispi frekansı, bu bantın dışında çıkar yani  $\alpha$  anlamlılık düzeyinde sıfır hipotezi reddedilir.

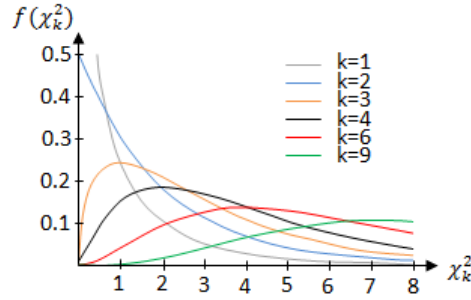


Şekil 4.11 K-S testinin kümülatif dağılım fonksiyonunda gösterimi [82]

#### 4.1.6 Khi-Kare Bağımsızlık Testi

İlk kez [84] tarafından ortaya atılan khi-kare ( $\chi^2$ ), parametrik olmayan testler içinde en çok kullanılan testlerden biridir. İki veya daha fazla rastgele kesikli değişken arasında ilişki bulunup bulunmadığını incelemek için bağımsızlık testi olarak kullanılır.

Bağımsız rastgele değişkenler  $Z_1, \dots, Z_k$  olmak üzere,  $V = \sum_{i=1}^k Z_i^2$ ;  $k$  serbestlik dereceli  $\chi^2$  dağılımı göstermektedir.  $\chi^2$  test büyüklüğünün bulunabilmesi için serbestlik derecesinin bilinmesi gerekmektedir. Şekil 4.12’de farklı serbestlik derecelerine göre  $\chi^2$  dağılımları gösterilmiştir. Çizelge 4.2’de ise  $\chi^2$  dağılımı için farklı serbestlik derecesi ve anlamlılık düzeyine göre kritik  $\chi^2$  değerleri gösterilmiştir. Büyük serbestlik derecesi ( $k > 10$ ) için normal dağılıma yaklaşır [70].



Şekil 4.12 Farklı serbestlik derecelerine göre  $\chi^2$  dağılımları [70]

Çizelge 4.2 Farklı anlamlılık düzeylerinde serbestlik derecesine göre kritik tablo değerleri

| Serbestlik derecesi | $\chi^2_{k,1-\alpha}$ kritik tablo değerleri |      |      |      |       |       |       |       |       |       |
|---------------------|----------------------------------------------|------|------|------|-------|-------|-------|-------|-------|-------|
| 1                   | 0.02                                         | 0.06 | 0.15 | 0.46 | 1.07  | 1.64  | 2.71  | 3.84  | 6.64  | 10.83 |
| 2                   | 0.21                                         | 0.45 | 0.71 | 1.39 | 2.41  | 3.22  | 4.60  | 5.99  | 9.21  | 13.82 |
| 3                   | 0.58                                         | 1.01 | 1.42 | 2.37 | 3.66  | 4.64  | 6.25  | 7.82  | 11.34 | 16.27 |
| 4                   | 1.06                                         | 1.65 | 2.20 | 3.36 | 4.88  | 5.99  | 7.78  | 9.49  | 13.28 | 18.47 |
| 5                   | 1.61                                         | 2.34 | 3.00 | 4.35 | 6.06  | 7.29  | 9.24  | 11.07 | 15.09 | 20.52 |
| 6                   | 2.20                                         | 3.07 | 3.83 | 5.35 | 7.23  | 8.56  | 10.64 | 12.59 | 16.81 | 22.46 |
| 7                   | 2.83                                         | 3.82 | 4.67 | 6.35 | 8.38  | 9.80  | 12.02 | 14.07 | 18.48 | 24.32 |
| 8                   | 3.49                                         | 4.59 | 5.53 | 7.34 | 9.52  | 11.03 | 13.36 | 15.51 | 20.09 | 26.12 |
| 9                   | 4.17                                         | 5.38 | 6.39 | 8.34 | 10.66 | 12.24 | 14.68 | 16.92 | 21.67 | 27.88 |
| 10                  | 4.86                                         | 6.18 | 7.27 | 9.34 | 11.78 | 13.44 | 15.99 | 18.31 | 23.21 | 29.59 |
| Anlamlılık Düzeyi   | 0.90                                         | 0.80 | 0.70 | 0.50 | 0.30  | 0.20  | 0.10  | 0.05  | 0.01  | 0.001 |

$\chi^2$  dağılımının olasılık yoğunluk fonksiyonu şöyledir [77]:

$$f(x) = \frac{x^{(k/2)-1} e^{-x/2}}{2^{(k/2)} \Gamma(k/2)} \quad k > 0 \quad (4.17)$$

$x$ :  $\chi^2$  test büyüklüğü,

$k$ : Serbestlik derecesi,

$\Gamma(k/2)$ : Gama fonksiyonudur.

$\chi^2$  bağımsızlık testinin uygulanabilmesi için gözlem sonuçlarının kategorik olması gerekmektedir [85].

Hipotezler:

$H_0$ = Rastgele değişkenler birbirinden bağımsızdır.

$H_1$ = Rastgele değişkenler birbirine bağımlıdır.

$\chi^2$  test büyüklüğü (4.18) eşitliği ile hesaplanır.

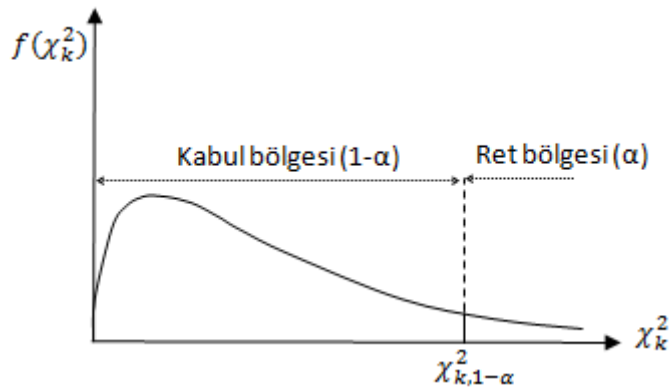
$$\chi^2 = \sum_{i=1}^m \sum_{j=1}^n \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \quad (4.18)$$

Buna göre;

$\chi^2 > \chi_{k,1-\alpha}^2$  ise  $H_0$  hipotezi red,  $H_1$  hipotezi kabul edilir.

$\chi^2 < \chi_{k,1-\alpha}^2$  ise  $H_0$  hipotezi kabul,  $H_1$  hipotezi reddedilir (Şekil 4.13).

Burada,  $\alpha$  anlamlılık düzeyi,  $E_{ij}$  beklenen (expected) frekans değeri,  $O_{ij}$  gözlem (observed) frekans değeri ve  $k$  serbestlik derecesidir [71], [77].



Şekil 4.13 Tek yanlı güven düzeyi için  $\chi^2$  kabul ve ret bölgeleri

$\chi^2$  test büyüklüğü, gözlem frekans değerleri ile beklenen frekans değerleri arasındaki bağımlılığı gösteren bir ölçüttür. Bu değerler birbirinden uzaklaştıkça,  $\chi^2$  test büyüklüğünün ret bölgesine girme olasılığı artacaktır.  $\chi^2$  test büyüklüğünün ret bölgesindeki yeri, değişkenlerin bağımlılık derecesini gösterir.  $\chi^2$  test büyüklüğü, ret bölgesi sınırından ne kadar uzakta ise, değişkenler arasındaki bağımlılık o kadar fazladır [86].

$\chi^2$  testinde şu hususları gözden kaçırmamak gerekir.

Eğer  $k = 1$  ise (4.19) eşitliğindeki gibi “Yates Düzeltmesi” tavsiye edilmektedir.

$$\chi^2 = \sum_{i=1}^m \sum_{j=1}^n \frac{(|O_{ij} - E_{ij}| - 0.5)^2}{E_{ij}} \quad (4.19)$$

Ayrıca aşağıdaki şartlar oluştuğunda, örneklem çapı artırılmalı veya rastgele değişkenlerin mevcut düzeyinde birleştirme yapılmalıdır [72].

- Eğer  $k = 1$  ise, beklenen frekans değerinin 5’ten küçük olmaması gerekir.
- Eğer  $k > 1$  ise, beklenen frekansların %20’sinden fazlasının 5’ten küçük olmaması gerekmektedir.

Kontenjans (contingency) tablosu; bağımsızlık testi yapılacak iki kategorik veriden birinci kategorik veriye ait gözlem frekanslarının satırlar halinde, ikinci kategorik veriye ait gözlem frekanslarının ise sütunlar halinde gösterildiği bir tablodur. Kontenjans tablosu terimi ilk kez Pearson tarafından 1904 yılında kullanılmıştır. Tablodaki satır sayısı ( $r$ ) ve sütun sayısı ( $c$ ) ile gösterilirse ( $r \times c$ )’lik bir kontenjans tablosu elde edilir. Bu şekilde herhangi bir sıradaki elemanla herhangi bir sütundaki eleman arasındaki ilişkinin (bağımlılık ya da bağımsızlığın) incelenmesi amaçlanır. Bunun için her sıra ve sütundaki elemanlara ait beklenen frekansların ( $E_{ij}$ ), gözlem frekanslarıyla ( $O_{ij}$ ) karşılaştırılması gerekir. Beklenen değer (4.20) eşitliği ile hesaplanır [71], [72], [85].

$$E_{ij} = \frac{i. \text{ satır toplamı} \times j. \text{ sütun toplamı}}{\text{Toplam gözlem sayısı}} \quad (4.20)$$

Serbestlik derecesi ise (4.21) eşitliği ile hesaplanır.

$$k = (r - 1)(c - 1) \quad (4.21)$$

İki rastgele kesikli değişken arasındaki ilişkiyi ölçmek için *Cramer'in V* katsayısı kullanılabilir. Bu katsayı  $\chi^2$  değerini normalize etmektedir. 0-1 aralığında değer alır. "0" ilişkinin olmaması, "1" ilişkinin çok güçlü olmasını ifade eder. *Cramer'in V* katsayısı ( $\varphi_c$ ) (4.22) eşitliği ile hesaplanır.

$$\varphi_c = \sqrt{\frac{\chi^2}{N(m-1)}} \quad (4.22)$$

$N$  örneklem sayısı ve  $m$  kontenjans tablosunun satır ve sütun sayılarından küçük olanıdır.

Orijinal 1:24,000 ölçekli Pomme de Terre akarsu ağından, 1:100,000 ölçekli akarsu ağının elde edilebilmesi için seçilen ve elenen akarsular, akarsu düzeylerine göre Çizelge 4.3'deki kontenjans tablosunda gösterilmiştir. Buna göre seçim ile akarsu düzeyi arasındaki ilişki,  $\alpha = 0,01$  anlamlılık düzeyinde test edilerek *Cramer'in V* katsayısı hesaplanmaktadır.

$H_0$ = Seçim ile akarsu düzeyi arasında ilişki yoktur.

$H_1$ = Seçim ile akarsu düzeyi arasında ilişki vardır.

Çizelge 4.3 Kontenjans tablosu

| Akarsu Düzeyleri |        |        |        |        |        |        |        |
|------------------|--------|--------|--------|--------|--------|--------|--------|
| Seçim            | Düzy 4 | Düzy 5 | Düzy 6 | Düzy 7 | Düzy 8 | Düzy 9 | Toplam |
| Seçme (1)        | 8      | 237    | 327    | 138    | 23     | 0      | 733    |
| Eleme (0)        | 0      | 263    | 609    | 559    | 160    | 16     | 1607   |
| Toplam           | 8      | 500    | 936    | 697    | 183    | 16     | 2340   |

Serbestlik derecesi:  $k = (2 - 1)(6 - 1)=5$

$\chi^2$  kritik tablo değeri:  $\chi^2_{5,1-\alpha} = 15.09$

$\chi^2$  test büyüklüğü, kritik tablo değerinden büyükse,  $H_0$  hipotezi reddedilir ve karşıt hipotez ( $H_1$ ) kabul edilir. Yani, seçim ile akarsu düzeyi arasında ilişki vardır. Beklenen frekansların hesaplanması gerekmektedir. Çizelge 4.4'de gözlem değerleri ve beklenen değerler bir arada gösterilmiştir.

Çizelge 4.4 Gözlenen değerler ve beklenen değerler

| Akarsu Düzeyleri |                                                                                  |                                                |                                                |                                                |                                                |                                            |        |
|------------------|----------------------------------------------------------------------------------|------------------------------------------------|------------------------------------------------|------------------------------------------------|------------------------------------------------|--------------------------------------------|--------|
| Seçim            | Düzy 4                                                                           | Düzy 5                                         | Düzy 6                                         | Düzy 7                                         | Düzy 8                                         | Düzy 9                                     | Toplam |
| Seçme (1)        | O <sub>11</sub> =8<br>E <sub>11</sub> =<br>(733x8)/2340<br>E <sub>11</sub> =2.5  | O <sub>12</sub> =237<br>E <sub>12</sub> =156.6 | O <sub>13</sub> =327<br>E <sub>13</sub> =293.2 | O <sub>14</sub> =138<br>E <sub>14</sub> =218.3 | O <sub>15</sub> =23<br>E <sub>15</sub> =57.3   | O <sub>16</sub> =0<br>E <sub>16</sub> =5   | 733    |
| Eleme (0)        | O <sub>21</sub> =0<br>E <sub>21</sub> =<br>(1607x8)/2340<br>E <sub>21</sub> =5.5 | O <sub>22</sub> =263<br>E <sub>22</sub> =343.4 | O <sub>23</sub> =609<br>E <sub>23</sub> =642.8 | O <sub>24</sub> =559<br>E <sub>24</sub> =478.7 | O <sub>25</sub> =160<br>E <sub>25</sub> =125.7 | O <sub>26</sub> =16<br>E <sub>26</sub> =11 | 1607   |
| Toplam           | 8                                                                                | 500                                            | 936                                            | 697                                            | 183                                            | 16                                         | 2340   |

$$\chi^2 = \sum_{i=1}^m \sum_{j=1}^n \frac{(O_{ij} - E_{ij})^2}{E_{ij}} =$$

$$\frac{(8-2.5)^2}{2.5} + \frac{(237-156.6)^2}{156.6} + \frac{(327-293.2)^2}{293.2} + \frac{(138-218.3)^2}{218.3} + \frac{(23-57.3)^2}{57.3} + \frac{(0-5)^2}{5} + \frac{(0-5.5)^2}{5.5} + \frac{(263-343.4)^2}{343.4} + \frac{(609-642.8)^2}{642.8} + \frac{(559-478.7)^2}{478.7} + \frac{(160-125.7)^2}{125.7} + \frac{(16-11)^2}{11} = 163.54$$

$$\chi^2 > \chi_{5,1-\alpha}^2$$

O halde, H<sub>0</sub> hipotezi reddedilir. Yani, seçim ile akarsu düzeyi arasında ilişki vardır.

*Cramer'in V* katsayısı:

$\varphi_c = 0.26$  değeri zayıf bir ilişki göstermektedir.

Akarsu ağının model genelleştirmesinde, kullanılacak özniteliklerin ağırlıkları eşit olmamalıdır. Özniteliklere dair ilişki değerleri kullanılarak öznitelikler ağırlıklandırılabilir (Bölüm 5.2.2).

## 4.2 Kümeleme ve Sınıflandırma Yöntemleri

Kümelemenin ardındaki temel düşünce, nesnelerin (değişkenlerden oluşan vektörler  $V = [x_k; k = 1, 2, 3, \dots, p]$ ) kendi aralarındaki benzerlikleri göz önüne alınarak gruplandırılması işlemidir. Başka bir deyişle, birbirine benzeyen nesneleri ayırma işlemidir. Kümeleme, genellikle geometrik yakınlık (geometric proximity) ilkesine

dayanan benzerlik ölçütleri ile yapılır. Yaygın olarak kullanılan benzerlik ölçütü (4.23) eşitliği ile hesaplanan Öklid benzerliğidir.

$$d(i, j) = \sqrt{\sum_{k=1}^p (x_{ik} - x_{jk})^2} \quad (4.23)$$

$d(i, j)$ : i. vektörü ile j. vektörü arasındaki uzaklık

$x_{ik}$ : i. vektördeki k. değişken değeri

$x_{jk}$ : j. vektördeki k. değişken değeri

$p$ : Vektörlerdeki değişken sayısı

Diğer benzerlik ölçütleri Manhattan ve Minkowski uzaklıklarıdır [74].

Kümeleme yöntemleri genellikle hiyerarşik yöntemler, bölümlenmeli yöntemler, yoğunluğa dayalı algoritmalar, grid temelli algoritmalar ve yapay sinir ağları olarak gruplandırılmaktadır.

Hiyerarşik olmayan kümeleme yöntemleri arasında en çok kullanılan yöntemlerden biri, k-ortalama kümeleme yöntemidir. Verilen  $k$  sayıda kümeye göre bölümlenme yapar. Yapay sinir ağları yöntemlerinden biri olan Kendini Düzenleyen Haritalar (KDH) (Self Organizing Maps-SOM) yöntemi de k-ortalama kümeleme yönteminde olduğu gibi sabit sayıda küme sayısı ve aynı amaç fonksiyonunu (benzerliğin en yüksek düzeye çıkarılması) kullanır. Fakat topolojik düzen içinde bir bölümlendirme ile kümeleme yapmaktadır [54].

Sınıflandırmanın kümelemeden farkı, önceden tanımlanmış sınıflar olmasıdır. En önemli sınıflandırma yöntemleri; karar ağaçları, regresyon, k-en yakın komşu, denetimli (supervised) bir yapay sinir ağı olan ileri beslemeli geri yayımlı (feed-forward back-propagation) çok katmanlı ağlar ve bir makine öğrenmesi yöntemi olan Destek Vektör Makineleri (DVM) (Support Vector Machines-SVM)'dir.

### 4.3 Yapay Zeka

Yapay zeka (Artificial Intelligence), insanların düşünme yöntemlerini analiz ederek bunların benzeri yapay yönergeleri geliştirmeye çalışan bir araştırma alanıdır. Günümüzde bilgisayarların, matematiksel olarak formülasyonu kurulamayan ve

özölmesi mümkün olmayan problemleri, sezgisel yöntemler ile özmesi için alışılmaktadır. Bilgisayarı, hem olaylar hakkında karar verebilen, hem de olaylar arasındaki ilişkileri öğrenebilen özelliklerle donatan ve bu yeteneklerin gelişmesini sağlayan alışmalar, “yapay zeka” alışmaları olarak bilinmektedir. Yapay zekanın en temel özelliğı bilgiye dayalı olarak karar verebilmesi ve eldeki bilgiler ile olayları öğrenecek, sonraki olaylar hakkında karar verebilmesidir. Yapay zeka teknolojilerinden önemli olanlarına kısaca değinmek faydalı olacaktır [87].

**Uzman Sistemler:** Bir problemi, o problemin uzmanlarının özdüğü gibi özebilen bilgisayar programları geliştiren teknolojidir. Bir uzman sistemin 4 temel elmanı vardır [87].

- a) Bilginin temin edilmesi: Uzmanlık alanı ile ilgili bilgilerin toplanması, derlenmesi ve bilgisayarın anlayacağı şekle dönüştürölmesidir.
- b) Bilgi tabanı: Bilgiler genellikle kurallar (Eğer ... ise O zaman...şeklinde), bilgi atıları ve prosedürlerden oluşur.
- c) ıkarım mekanizması: Bilgi tabanında bulunan bilgileri arayan, filtreleyen, yorumlayan ve sonuçlar ıkaran mekanizmadır.
- d) Kullanıcı ara birimi: Uzman sistemi kullanan kişiler ile uzman sistemin iletişimini sağlar.

**Genetik Algoritmalar:** Optimizasyon problemlerinin özümünde kullanılmaktadır. Biyolojik evrimin genetik sürecine dayanır. Doğal seleksiyon, mutasyon ve tekrar üretim işlem adımlarından oluşmaktadır. Popölasyonlar, doğal seleksiyon ilkesine göre gelişmiştir. Bu süreç örnek alınarak, problemlerin özölmesinde arzu edilen sonucu üretecek özelliklerin kalıtım yolu ile başlangıç özömlerinden elde edilen yeni özömlere, onlardan daha sonraki özömlere geçtiğı kabul edilmektedir. Algoritma, popölasyon adı verilen ve kromozomlarla temsil edilen bir özüm kümesi ile başlamaktadır. Bir kromozom, temsil ettiğı özüm hakkında bilgi içermelidir. Popölasyondaki özömler doğal seleksiyon ile seçilir. Bu seçimde uygunluğu yüksek olan özömlün yeni nesle aktarılma ihtimali de daha yüksektir. Yeni nesil aprazlama yolu ile seçilen kromozomlardan oluşur. aprazlama neticesinde farklı özömlere ulaşmak bazen zor olmaktadır. Yeni özüm aramanın kolaylaştırılması ve aramanın

yönünü değiştirmek amacı ile bir kromozomun değiştirilmesi (mutasyon) gerekmektedir. Yeni çözümün katılımı ve en kötü çözümün popülasyondan çıkarılmasıyla en iyi çözüm bulunur [87], [88].

**Bulanık Mantık:** Bulanık küme teorisine dayanmaktadır. Kesin olarak bilinmeyen olaylar hakkında karar vermek için uzmanlar; normal, yüksek, düşük, yaklaşık gibi kavramları kullanmaktadırlar. Bulanık mantık teknolojisi, bilgisayarın da bu gibi durumlarda karar verebilmesi için geliştirilmiş teknolojidir [87].

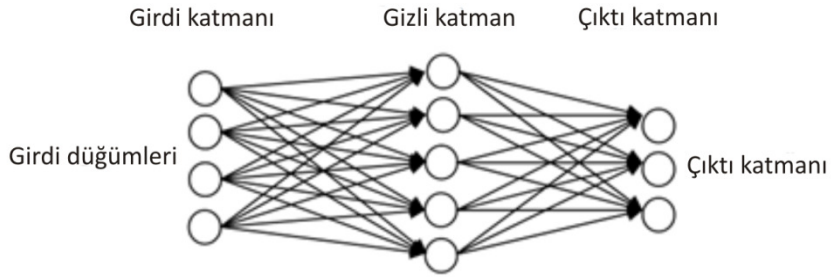
**Makine Öğrenimi:** Bilgisayarın bir olay ile ilgili bilgileri ve tecrübeleri öğrenerek, gelecekte oluşacak benzeri olaylar hakkında kararlar verebilmesi ve problemlere çözümler üretebilmesidir [87]. Makine öğrenmesinde kullanılan iki önemli öğrenme paradigması; Yapay Sinir Ağları (YSA) ve istatistiksel öğrenme teorisidir.

#### 4.3.1 Yapay Sinir Ağları

İlk yapay sinir ağı modeli 1943 yılında, bir nörolog olan Warren McCulloch ile bir matematikçi olan Walter Pitts tarafından gerçekleştirilmiştir. 1949 yılında Hebb, “The Organization of Behavior: A Neuropsychological Theory” isimli kitabında, öğrenme ile ilgili temel teoriyi ele almıştır. YSA, insan beyninden esinlenerek geliştirilmiş ağırlıklı bağlantılar aracılığıyla birbirine bağlanan ve her biri kendi belleğine sahip işlem elemanlarından oluşan paralel ve dağıtılmış bilgi işleme yapılarıdır. Yapay sinir ağları; matematik, bilgisayar teknolojisi, robot bilimi, mühendislik, tıp, finans, endüstri, bulanık veya var olan basit yöntemlerle çözülemeyen problemlerin çözümünde, doğrusal olmayan sistemlerde başarıyla uygulanmaktadır [89].

Hesaplamalı sinir ağları (computational neural networks) olarak da bilinen YSA’nın artık biyolojik sinir ağlarının simülasyonu olarak görülmediğinin belirtilmesi gerekir. Günümüzde birçok uygulamada, YSA’nın biyolojik sinir ağlarından esinlenerek oluşturulduğu düşüncesinden uzaklaşmıştır. YSA ve biyolojik sinir ağlarının ortak noktası, bilginin tek bir yerde değil, paralel ve yayılmış bir yapıda depolanması ve yeni bilginin ağına büyük bir kısmını etkileyen değişikliklerle öğrenilmesidir (öğrenme kuralı). Aslında YSA, çoğunlukla kara kutu (black box) işlemleri olarak değerlendirilmektedir. Yani sistem içerisinde ne olduğu bilinmemektedir. YSA’nın genel yapısı, girdi ve çıktı

düğüm noktaları kümelerinden oluşmaktadır. Bu düğümler; nöron, sinir hücresi, işlem elemanı veya hesaplama birimi olarak da adlandırılmaktadır. Girdi nöronlarına sunulan çok değişkenli veri, nöronlar arasındaki bağlantılarla ilişkilendirilmiş ağırlıklara göre işlenerek çıktı nöronları aktive edilir. Sinir ağları, genellikle bu ağırlıkların hesaplanması ile ilgilidir. Bunun için sinir ağlarının, bir veya daha fazla nöron içeren gizli katmanı bulunur (Şekil 4.14).



Şekil 4.14 Denetimli ileri beslemeli sinir ağları [54]

Ağırlıkların hesaplanması süreci ağın eğitimi olarak bilinir. Eğitim boyunca, bu nöronlara gelen bağlantılardaki ağırlıklar önceden tanımlanmış bir fonksiyona göre toplanarak hesaplanır. Belli bir eşik fonksiyonu kullanılarak sonuçlar elde edilmektedir.

Eğitim boyunca nöron ağırlıklarına yapılan dengeleme ileri doğru beslenir (feed forward) veya bir önce gelen katmandan etkilenir. İleri beslemeli ağlarda kullanılan geri yayılım (back propogation) önemli bir kavramdır. Nöron ağırlıklarının dengelenmesi ile hataların (bilinen çıktı ve ağın çıktısı arasındaki fark) nasıl minimize edildiği ifade edilir.

YSA, iki temel öğrenme stratejisi kullanmaktadır. Bunlar; denetimli (supervised) ve denetimsiz (unsupervised) öğrenmedir. Denetimli durumda, çıktı değerleri veya sınıfı belli olan çok değişkenli veriyi içeren girdi değerleri, eğitim boyunca ağa sunulmaktadır (girdi ve çıktı çiftleri şeklinde veri tipi). Eğitim boyunca, gizli katmanların ağırlıkları uygun bir çıktı değeri yakalanana dek iteratif olarak dengelenir. Eğitim tamamlandıktan sonra, çıktı değeri bilinmeyen yeni girdiler ağa sunulur ve çıktı değerleri tahmin edilir (ağın test edilmesi). Denetimli YSA, sınıflandırma problemlerinde kullanılmaktadır.

Denetimsiz öğrenmede ise girdi vektörlerinin sınıfları önceden bilinmez. Çıktı nöronları, belli benzerlik fonksiyonlarını temel alarak girdi vektörleri için yarışır ve kazanan

nöronların ağırlıkları, ilgili girdi nöronların ağırlıklarına göre dengelenir. Yarışmaya dayanan öğrenme prosedürü nedeniyle, oldukça benzer olan girdi nöronları, benzer çıktı nöronlarının dengelenmesini sağlar. Aynı zamanda girdi verisindeki farklılıklar (dissimilarities) vurgulanır hale gelir. Denetimsiz YSA, kümeleme problemlerinde kullanılmakta ve çok değişkenli veri içindeki ilişkileri bulmaktadır [54]. Kendini Düzenleyen Haritalar en yaygın kullanılan denetimsiz YSA yöntemidir.

#### 4.3.2 Kendini Düzenleyen Haritalar

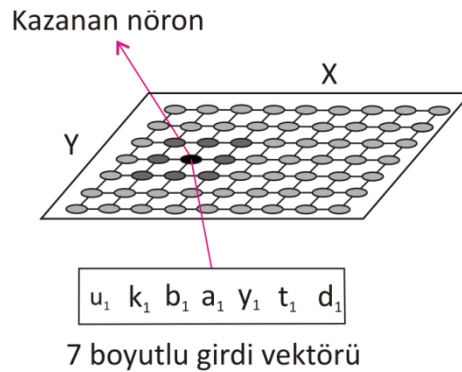
Kendini Düzenleyen Haritalar (KDH) (Self Organizing Maps); matematik, bilgisayar teknolojisi, robot bilimi, mühendislik, tıp, finans, endüstri, coğrafi bilgi sistemleri, demografik verilerin analizi, karar destek sistemleri, kartografik genelleştirme, uzaktan algılama verilerinin kümelenmesi, öznitelik uzayında yörünge (trajectory) analizi, iklim ve meteoroloji gibi alanlarda kullanılabilmektedir.

Teknolojinin gelişimiyle beraber, coğrafi bilgi sistemlerinde kullanılan veri hacmi hızla artmaktadır. Verinin geometrik ve öznitelik uzayında büyümesi, çok boyutlu verinin işlenmesini gerektirmektedir. Artan veri miktarı karşısında, geleneksel yöntemler başarısız olmakta, ilişkileri ve anlamı araştırırken bir takım engeller ortaya çıkartmaktadır. Büyük veri setlerine uygulanan yoğun hesaplama işlemlerinde Kohonen haritası olarak da bilinen KDH yöntemi kullanılabilmektedir. Kohonen'in 1982 yılında ilk kez ortaya attığı KDH ağırları rağbet görmüştür [90]. KDH ağlarının, coğrafi birçok konuda daha yaygın uygulanabilmesi için, KDH ağının karmaşık bilgisayar işlemlerinin ve sonuçlarının daha fazla tartışılması gerekmektedir.

Amacına ve eğitim parametrelerine bağlı olarak bir kümeleme yöntemi olan KDH'nin boyut azaltma işlevi de vardır. Bir vektör  $n$  adet değişkenden oluşmuşsa,  $n$ -boyutlu olarak ifade edilir. Başlıca boyut azaltma yöntemleri; Temel Bileşenler Analizi (Principal Component Analysis) ve Çok Boyutlu Ölçekleme (Multi-Dimensional Scaling) yöntemidir. KDH boyut azaltma işleminde,  $n$ -boyutlu veri topolojik bir düzende daha az boyutlu hale getirilmektedir. *Akarsu uzunluğu, kıvrımlılık, bağlılık bakımından merkezlik, arasında olma bakımından merkezlik, yakınlık bakımından merkezlik, akarsu tipi ve akarsu düzeyi* değişkenlerinden/özniteliklerinden oluşan yedi boyutlu

nesnelerin/vektörlerin iki boyutlu düzlemde topolojik bir düzende kümelerle temsil edilmesi bir boyut azaltma işlemidir [54].

KDH; yarışmacı öğrenmeye dayanan, gizli katman olmayan, girdi ve çıktı katmanlarından oluşan denetimsiz sinir ağlarını içerir. KDH, doğrusal olmayan ve büyük boyutlu girdi vektörlerini ( $x_n$ ), daha az boyutlu bir uzaya (çıkı katmanına) izdüşürmektedir. (Şekil 4.15). Her nöron, girdi verisi ile aynı boyutlu nöron ağırlığı veya referans vektörü diye adlandırılan bir vektör ile ilişkilidir.

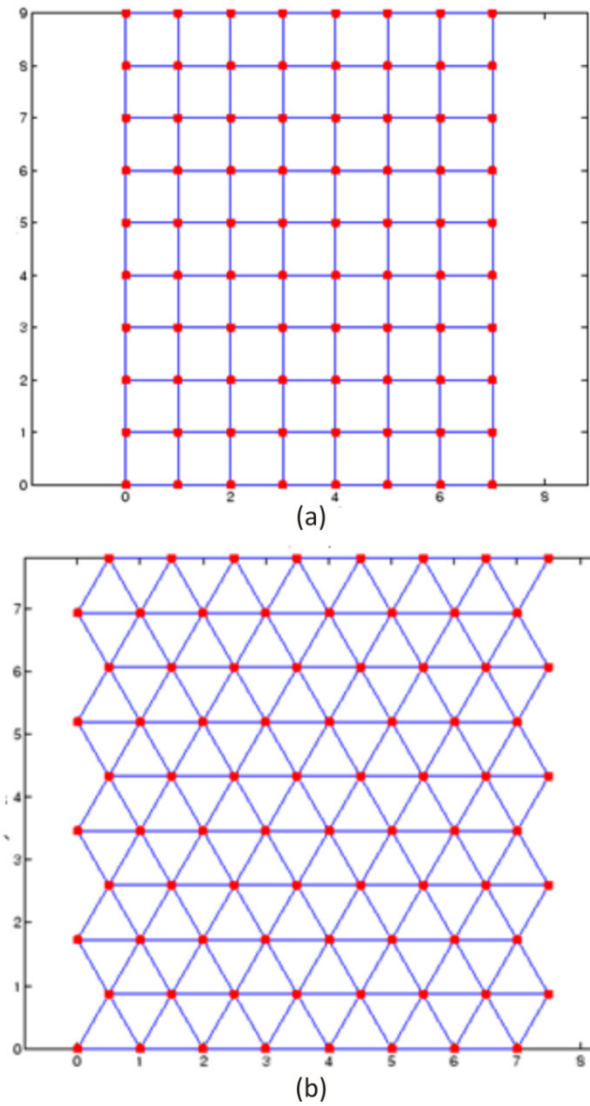


Şekil 4.15 KDH’de girdi ve çıktı katmanları

KDH yöntemi, çok fazla çeşitlilik gösteren veri setlerine uygulanabilmektedir. Girdi verisi, genellikle nicel öznitelikleri içeren  $n$ -boyutlu vektörler formundadır.  $N$ -boyutlu vektörlerin ön işlenmesi, diğer sinir ağlarındaki tipik işlemlere benzemektedir. Dikkat edilmesi gereken iki ana konudan biri verinin dağılımı, diğeri ise her öznitelik için değer aralığıdır. Her değişkenin, 0-1 aralığına ya da önceden tanımlanmış bir aralığa normalize edilmesi iyi bir düşüncedir. Her değişkenin ortalaması sıfır ve varyansı eşit olacak biçimde normalize edilmesi de uygun olacaktır. Bu nedenle, KDH için maksimum-minimum normalizasyonu (eşitlik 4.7) veya Z-skor standardizasyonu (eşitlik 4.8) uygun olacaktır [54], [75], [90].

KDH’nin oluşturulmasında birinci basamak; KDH’nin büyüklüğünün ve topolojisinin belirlenmesidir. KDH’nin büyüklüğü;  $X, Y$  yönlerinde kullanılacak nöron sayısı olarak verilir. Örneğin,  $X$  için 2,  $Y$  için 3 nöron ile KDH büyüklüğü belirlenirse, KDH toplam 6 adet nöron içermektedir. KDH boyutu 100 x 100 ise; 10,000 adet nöron içermektedir. KDH topolojisi kurulurken çoğunlukla iki yöntem kullanılmaktadır: Kare topolojide her

nöron birbirine dört komşu nöron ile bağlıdır. Altıgen topolojide ise her nöronun 6 adet komşusu bulunmaktadır. Altıgen topoloji daha yaygın olarak kullanılmaktadır [54], [91].



Şekil 4.16 a) Kare topoloji ve b) altıgen topoloji

KDH büyüklüğü, şekli ve topolojisi kümeleme doğruluğuna (girdi ve referans vektörleri arasındaki benzerlik miktarına) göre deneme yanılma yolu ile belirlenebileceği gibi; büyüklüğü [91]'e göre, şekli ise [92]'e göre belirlenebilir.

[91]'e göre, KDH'de en uygun nöron sayısı (4.24) eşitliği ile hesaplanabilir.

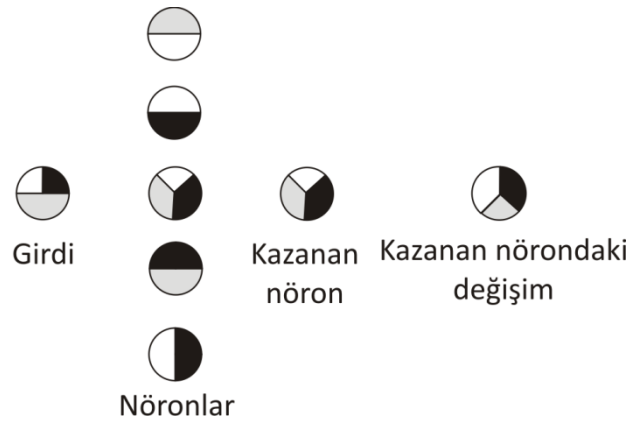
$$5\sqrt{n} \quad (4.24)$$

Burada,  $n$  girdi vektörlerinin sayısıdır.

[92] tarafından, kısa kenarı uzun kenarının en az yarısı kadar olan KDH şekli tavsiye edilmektedir. [93] tarafından, bu özellik “bire yarım kuralı” olarak adlandırılmaktadır. Örneğin, 62 vektörden oluşan bir girdi verisi için, (4.24) eşitliğine göre, KHD nöron sayısı 39 ve bire yarım kuralına göre en uygun KDH büyüklüğü 4 x 10 olabilir.

KDH yönteminde girdi vektörleri önceden bilinen sınıflara karşılık gelmemektedir. Eğitim öncesi nörona rastgele veya algoritmanın eğitimine yardımcı olarak (eğitim süresinin kısaltılması için) doğrusal bir tahmin ile atanan referans vektörleri (ağırlıklar), girdi vektörlerine olan benzerliklerine göre rekabet etmektedir.

Ağın eğitimi; her bir girdi vektörüne en çok benzeyen yani kazanan (best matching unit) nöronun bulunması, kazanan nöronun ve bu nörona komşu olan nöronların ağırlıklarının güncellenmesi şeklinde yapılan iteratif işlemleri içermektedir. Kısaca, rekabeti kazanan nöronlar ve komşularına ait ağırlıklar, ilgili girdi vektörlerine göre güncellenmektedir. Şekil 4.17’de, girdi vektörüne en çok benzeyen (kazanan) nöronun ağırlığının değişmesiyle, nöronun girdi vektörüne benzeme çabası gösterilmiştir.



Şekil 4.17 Kazanan nöronun girdi vektörüne göre değişimi [94]

İki tip eğitim bulunmaktadır. Bunlar; kademeli (incremental) ve yığınsal (batch) eğitimidir. Kademeli eğitimde, girdi vektörleri ağa tek tek sunularak ağın ağırlıkları güncellenir. Fakat yığınsal eğitimde, bütün girdi vektörleri bir arada ağa sunularak ağın ağırlıkları güncellenmektedir.

Kademeli eğitime göre işlem adımları:

- Rastgele veya doğrusal bir tahmin ile belirlenen referans vektörleri (ağırlıklar) nöronlara atanır (başlangıç durumu).

- b) Bir girdi vektörü ile referans vektörleri arasındaki benzerlikler hesap edilir.
- c) Kazanan nöron tespit edilir.
- d) Kazanan nöron ve tüm komşularının referans vektörleri güncellenir (ağın eğitimi).
- e) Kümeleme, kararlı bir yapı kazanana kadar iterasyona devam edilir.

Girdi verisi  $i = (1, 2, \dots, m)$  olmak üzere, n-boyutlu  $x_i$  vektörlerini içermektedir.

$$x_i = [\varepsilon_1, \dots, \varepsilon_n]^T \in R^n \quad (4.25)$$

Her bir nöron, bir  $m_i$  referans vektörüne sahiptir.

$$m_i = [\mu_1, \dots, \mu_n]^T \in R^n \quad (4.26)$$

Eğitim boyunca, bir  $x$  vektörü; bir minimum mesafe veya maksimum benzerlik ölçütünü sağlayan  $m_c$  referans vektörünü bulmak için, tüm  $m_i$  vektörleri ile karşılaştırılmaktadır. Öklid uzaklığı en çok kullanılan benzerlik ölçütüdür.

$$\|x - m_c\| = \min_i \{\|x - m_i\|\} \quad (4.27)$$

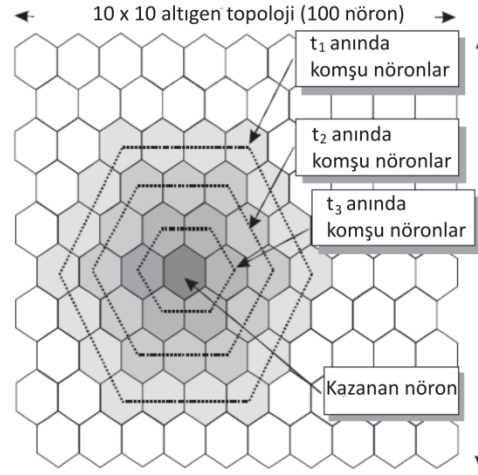
Kazanan nöron ve komşu nöronları (4.28) eşitliği ile güncellenir.

$$m_i(t + 1) = m_i(t) + \alpha(t)h_{ci}(t)[x(t) - m_i(t)] \quad (4.28)$$

Öğrenme katsayısı  $\alpha$ , 0-1 aralığında bir girdi parametresidir ve işlem zamanına ( $t$ ) bağlı olarak giderek azalmaktadır. Öğrenme katsayısı genellikle başlangıçta yaklaşık 0.1 seçilir [94]. Eğitimi etkileyen temel parametrelerden biri; nöron vektörlerini güncellemek için bir mesafe-ağırlıklı model olan komşuluk fonksiyonudur ( $h_{ci}$ ). Doğrusal ve Gauss modelleri en popüler fonksiyonlardır. Gauss modeli:

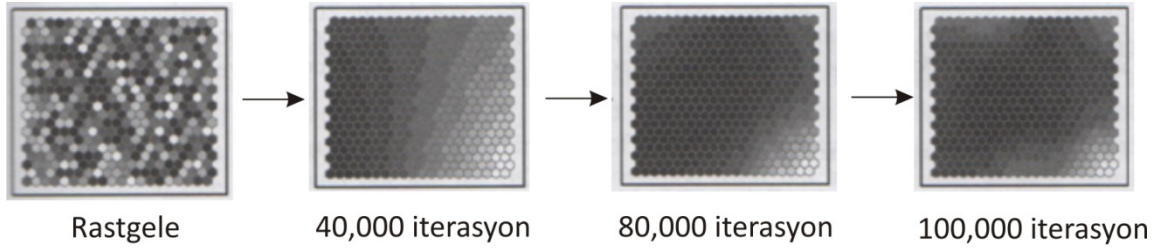
$$h_{ci}(t) = e^{-d_{ci}^2/2\gamma_i^2(t)} \quad (4.29)$$

Görüldüğü gibi komşuluk fonksiyonu, kazanan nöron ile komşuları arasındaki topolojik mesafe ( $d_{ci}$ ) ve  $t$  işlem zamanına bağlıdır. Maksimum  $d_{ci}$  komşuluk yarıçapını verir. Gauss modelinde komşuluk yarıçapı ( $\gamma$ ) çekirdek (kernel) genişliğidir ve zaman içerisinde azalmaktadır (Şekil 4.18). KDH eğitimi, önceden belirlenen eğitim döngü sayısı ( $t_{max}$ ) tamamlanınca durmaktadır ( $m_i(t + 1) = m_i(t)$ ,  $t = \infty$ ).



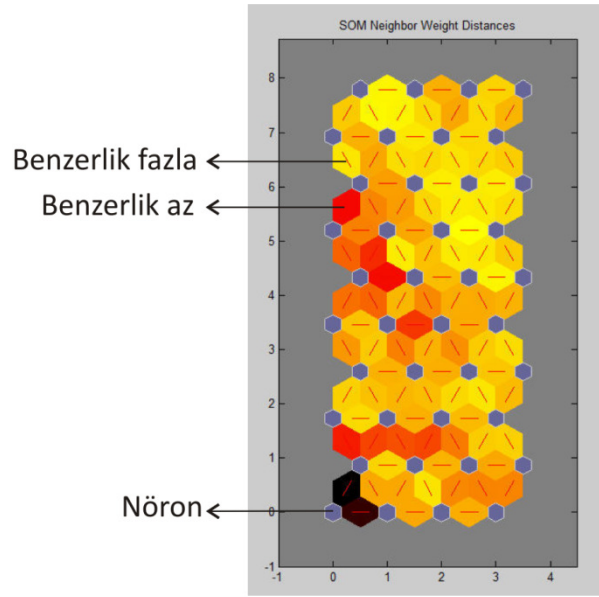
Şekil 4.18 Eğitim boyunca azalan komşuluk yarıçapı [52]

KDH'nin görselliği, genellikle bileşen düzlemler (component planes) ve U-matrisi [95] ile sağlanmaktadır. Bileşen düzlemler için en çok kullanılan yaklaşım; renkli veya gri gölgelendirmedir. Eğitim boyunca KDH, Şekil 4.19'da görüldüğü gibi yavaş yavaş kararlı bir yapıya kavuşmaktadır. Şekil 4.19'da, 20x20 boyutlu bir KDH'nin dört iterasyon zamanı gösterilmiştir. Eğitim, rastgele atanmış ağırlıklar ile başlamaktadır. 100,000 iterasyon sonunda eğitim tamamlanmaktadır [54].

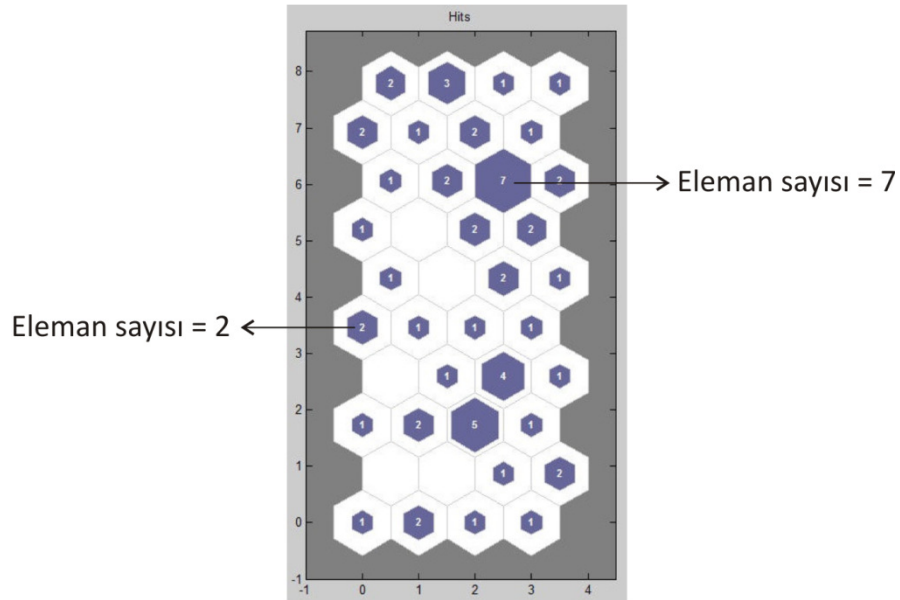


Şekil 4.19 KDH eğitimi [54]

Şekil 4.20'de U-matrisi gösterilmiştir. Mavi renkte altıgen biçiminde gösterilen nöronlar (kümeler) kırmızı renkte çizgiler ile komşularına bağlanmaktadır. Komşu nöronlar birbirlerine daha çok benzemektedir. Çizgileri içeren bölgelerin renkleri, nöronlar arasındaki benzerliği göstermektedir. Koyu renkler, komşu kümeler arasındaki benzerliğin daha az olduğunu; açık renkler ise benzerliğin daha fazla olduğunu göstermektedir. Şekil 4.21'de nöronlar; eleman sayısına göre değişik büyüklüklerde altıgenlerle gösterilmiştir.

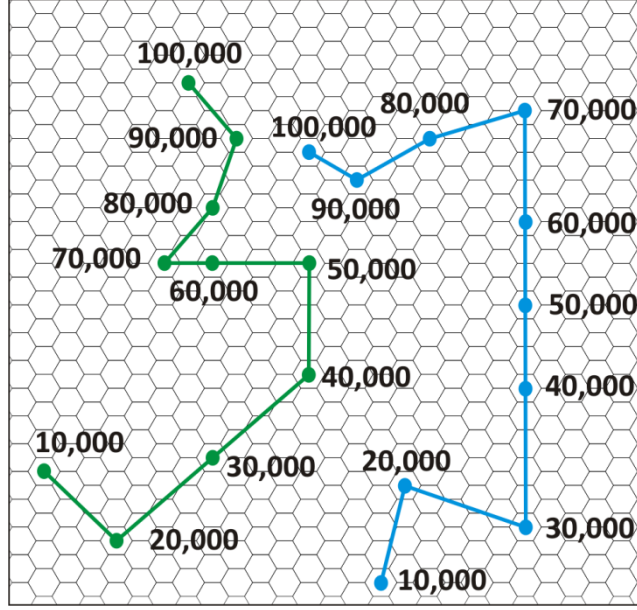


Şekil 4.20 U-matrisinde nöronlar arası benzerlik

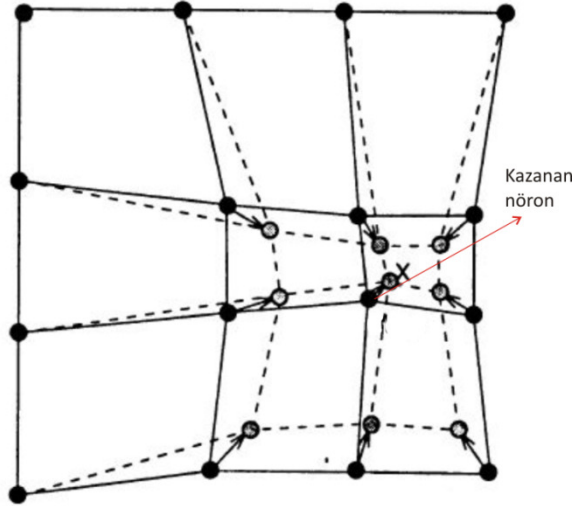


Şekil 4.21 Küme eleman sayıları

Şekil 4.22’de, toplam 100,000 iterasyondan oluşan bir KDH ağının eğitim sürecinde, iki vektörün (yeşil ve mavi) hareketleri gösterilmiştir. Başlangıçta her bir nörona 0-1 aralığında rastgele ağırlıklar atanmaktadır. Nöron ağırlıkları her döngüde güncellenmektedir. Bu nedenle, her 10,000 iterasyon sonunda bu iki girdi vektörü kendilerine en çok benzeyen nöronlara taşınmakta ve sonunda kararlı bir yapıya kavuşmaktadır [54]. Şekil 4.23’de, kazanan nöronun ve komşularının girdi vektörü  $x$ ’e göre ağırlıklarını değiştirmesi sonucu aldıkları yeni konumlar gösterilmiştir [75], [91].



Şekil 4.22 Toplam 100,000 döngüden oluşan bir KDH ağının eğitiminde her 10,000 döngü sonunda iki girdi vektörünün konumları [54]



Şekil 4.23 Kazanan ve komşu nöronların güncelleme öncesi (devamlı çizgi) ve sonrası (kesikli çizgi) durumları [75], [91]

1:24,000 ölçekli bir akarsu ağına ait *uzunluk*, *kıvrımlılık*, *merkezlik değerleri* (*bağlılık*, *arasında olma*, *yakınlık*), *düzey ve tip*; öznitelikleri veya değişkenleri, bu öznitelikler ise her bir nesneye ait girdi vektörlerini oluşturmaktadır. Bir akarsu ağında, akarsu sayısı (gözlem sayısı) kadar girdi vektörü bulunur. Vektörler girdi matrisinin satırlarını, değişkenler ise sütunlarını oluşturmaktadır.

Aşağıdaki örnekte KDH'nin çalışma prensibi anlatılmaktadır. Akarsu ağındaki dört adet akarsuya ait yedi boyutlu girdi vektörlerinin (Çizelge 4.5) 3x3 boyutlu KDH ile

kümelenmesi ele alınmaktadır. Öncelikle, girdi vektörlerinin normalizasyonu gerçekleştirilir. Daha sonra, 9 adet nörona algoritma tarafından otomatik olarak -1 ile +1 aralığında rastgele referans vektörleri atanır (Çizelge 4.6 ve Şekil 4.24a). Öğrenme katsayısı ( $\alpha$ ) 0.1, çekirdek genişliği ( $\gamma$ ) ise 1 olarak alınmıştır.

Çizelge 4.5 Akarsulara ait özniteliklerden oluşan girdi vektörleri

| Vektör | Uzunluk (km) | Kıvrımlılık | Bağıllık | Arasında olma | Yakınlık | Düzyey | Tip |
|--------|--------------|-------------|----------|---------------|----------|--------|-----|
| 1      | 1.55         | 0.95        | 2        | 0.38          | 0.19     | 4      | 1   |
| 2      | 0.71         | 0.96        | 1        | 0.00          | 0.16     | 4      | 0   |
| 3      | 1.02         | 0.96        | 1        | 0.00          | 0.14     | 4      | 0   |
| 4      | 7.29         | 0.86        | 11       | 0.36          | 0.13     | 4      | 1   |

Çizelge 4.6 Rastgele belirlenen 9 adet referans (ağırlık) vektörü

| Nöron    | Rastgele belirlenen referans vektörleri |       |       |        |       |       |        |  |
|----------|-----------------------------------------|-------|-------|--------|-------|-------|--------|--|
| x=1; y=1 | 0.9                                     | 0.737 | 0.91  | -0.925 | 0.849 | 0.629 | 0.269  |  |
| x=1; y=2 | -0.690                                  | 0.315 | 0.791 | 0.782  | 0.24  | 0.065 | 0.863  |  |
| x=1; y=3 | 0.302                                   | 0.503 | 0.157 | 0.421  | 0.502 | 0.703 | 0.695  |  |
| x=2; y=1 | 0.575                                   | 0.175 | 0.584 | 0.104  | 0.008 | 0.466 | 0.641  |  |
| x=2; y=2 | -0.147                                  | 0.702 | 0.603 | 0.793  | 0.267 | 0.835 | 0.575  |  |
| x=2; y=3 | 0.27                                    | 0.343 | -0.47 | -0.091 | 0.824 | 0.802 | 0.814  |  |
| x=3; y=1 | 0.352                                   | 0.629 | 0.829 | 0.324  | 0.819 | 0.692 | -0.923 |  |
| x=3; y=2 | 0.199                                   | 0.128 | 0.193 | 0.4    | 0.218 | 0.5   | 0.75   |  |
| x=3; y=3 | -0.234                                  | 0.512 | 0.434 | 0.868  | 0.458 | 0.412 | 0.474  |  |

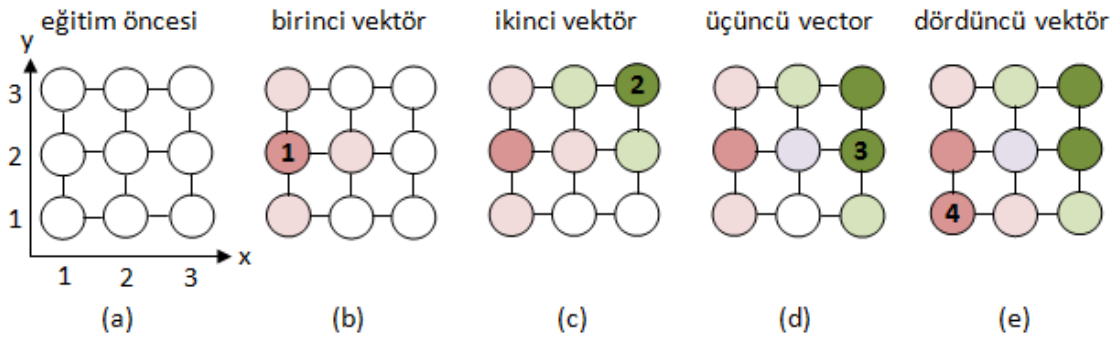
Girdi vektörlerinden (eşitlik 4.8) ile hesaplanan Z-skor standart matris (normalizasyon):

$$\begin{bmatrix} -0.351 & 0.37 & -0.36 & 0.919 & 1.353 & 0 & 0.866 \\ -0.619 & 0.592 & -0.566 & -0.865 & 0.077 & 0 & -0.866 \\ -0.521 & 0.532 & -0.566 & -0.865 & -0.451 & 0 & -0.866 \\ 1.491 & -1.493 & 1.493 & 0.811 & -0.979 & 0 & 0.866 \end{bmatrix}$$

Birinci vektör ağı sunulur. Öklid benzerliği (eşitlik 4.23) kullanılarak birinci vektöre en çok benzeyen referans vektörü belirlenir. Birinci vektöre en çok benzeyen yani en yakın vektör,  $(x=1; y=2)$  konumunda bulunan nörondur (Çizelge 4.7). (4.28) eşitliği kullanılarak, bu nöronun ve komşularının ağırlıkları, birinci girdi vektörüne göre değiştirilir (Şekil 4.24b). Çizelge 4.7 ve Çizelge 4.8 karşılaştırılırsa, girdi vektörü ile kazanan nöron ve komşuları arasındaki Öklid uzaklığı azalmakta yani bu nöronlar girdi vektörüne daha çok benzemektedir.

Çizelge 4.7 Birinci vektörün nöronlara olan Öklid uzaklıkları

| Nöron                        | Uzaklık      |
|------------------------------|--------------|
| $x=1; y=1$                   | 3.017        |
| <b><math>x=1; y=2</math></b> | <b>0.908</b> |
| $x=1; y=3$                   | 1.778        |
| $x=2; y=1$                   | 1.723        |
| $x=2; y=2$                   | 1.319        |
| $x=2; y=3$                   | 2.539        |
| $x=3; y=1$                   | 2.588        |
| $x=3; y=2$                   | 1.547        |
| $x=3; y=3$                   | 1.223        |



Şekil 4.24 Örnek KDH eğitimi [54]

Çizelge 4.8 Kazanan nöron ve komşularının ağırlıklarının değişimi

| Nöron           | Referans vektörleri |              |              |               |              |              |              |  | Uzaklık      |
|-----------------|---------------------|--------------|--------------|---------------|--------------|--------------|--------------|--|--------------|
| <b>x=1; y=1</b> | <b>0.824</b>        | <b>0.715</b> | <b>0.911</b> | <b>-0.787</b> | <b>0.776</b> | <b>0.591</b> | <b>0.305</b> |  | <b>2.834</b> |
| <b>x=1; y=2</b> | <b>-0.656</b>       | <b>0.321</b> | <b>0.803</b> | <b>0.839</b>  | <b>0.18</b>  | <b>0.059</b> | <b>0.863</b> |  | <b>0.817</b> |
| <b>x=1; y=3</b> | <b>0.262</b>        | <b>0.495</b> | <b>0.203</b> | <b>0.477</b>  | <b>0.45</b>  | <b>0.66</b>  | <b>0.705</b> |  | <b>1.554</b> |
| x=2; y=1        | 0.575               | 0.175        | 0.584        | 0.104         | 0.008        | 0.466        | 0.641        |  | 1.723        |
| <b>x=2; y=2</b> | <b>-0.159</b>       | <b>0.682</b> | <b>0.622</b> | <b>0.827</b>  | <b>0.229</b> | <b>0.784</b> | <b>0.593</b> |  | <b>1.239</b> |
| x=2; y=3        | 0.27                | 0.343        | -0.47        | -0.091        | 0.824        | 0.802        | 0.814        |  | 2.539        |
| x=3; y=1        | 0.352               | 0.629        | 0.829        | 0.324         | 0.819        | 0.692        | -0.923       |  | 2.588        |
| x=3; y=2        | 0.199               | 0.128        | 0.193        | 0.4           | 0.218        | 0.5          | 0.75         |  | 1.547        |
| x=3; y=3        | -0.234              | 0.512        | 0.434        | 0.868         | 0.458        | 0.412        | 0.474        |  | 1.223        |

İkinci vektör ağı sunulur. Bu vektöre en çok benzeyen (x=3; y=3) konumunda bulunan nöron rekabeti kazanır (Çizelge 4.9). Bu nöronun ve komşularının ağırlıkları ikinci girdi vektörüne göre değiştirilir (Çizelge 4.10) (Şekil 4.24c).

Çizelge 4.9 İkinci vektörün nöronlara olan Öklid uzaklıkları

| Nöron           | Uzaklık      |
|-----------------|--------------|
| x=1; y=1        | 3.085        |
| x=1; y=2        | 2.643        |
| x=1; y=3        | 2.454        |
| x=2; y=1        | 2.553        |
| x=2; y=2        | 2.524        |
| x=2; y=3        | 2.537        |
| x=3; y=1        | 2.505        |
| x=3; y=2        | 2.363        |
| <b>x=3; y=3</b> | <b>2.341</b> |

Çizelge 4.10 Kazanan nöron ve komşularının ağırlıklarının değişimi

| Nöron           | Referans vektörleri |              |               |               |              |              |              | Uzaklık      |
|-----------------|---------------------|--------------|---------------|---------------|--------------|--------------|--------------|--------------|
| x=1; y=1        | 0.824               | 0.715        | 0.911         | -0.787        | 0.776        | 0.591        | 0.305        | 3.085        |
| x=1; y=2        | -0.656              | 0.321        | 0.803         | 0.839         | 0.18         | 0.059        | 0.863        | 2.643        |
| x=1; y=3        | 0.262               | 0.495        | 0.203         | 0.477         | 0.45         | 0.66         | 0.705        | 2.454        |
| x=2; y=1        | 0.575               | 0.175        | 0.584         | 0.104         | 0.008        | 0.466        | 0.641        | 2.553        |
| x=2; y=2        | -0.159              | 0.682        | 0.622         | 0.827         | 0.229        | 0.784        | 0.593        | 2.524        |
| <b>x=2; y=3</b> | <b>0.216</b>        | <b>0.358</b> | <b>-0.494</b> | <b>-0.081</b> | <b>0.740</b> | <b>0.753</b> | <b>0.712</b> | <b>2.383</b> |
| x=3; y=1        | 0.352               | 0.629        | 0.829         | 0.324         | 0.819        | 0.692        | -0.923       | 2.505        |
| <b>x=3; y=2</b> | <b>0.149</b>        | <b>0.156</b> | <b>0.129</b>  | <b>0.380</b>  | <b>0.170</b> | <b>0.470</b> | <b>0.652</b> | <b>2.220</b> |
| <b>x=3; y=3</b> | <b>-0.272</b>       | <b>0.520</b> | <b>0.304</b>  | <b>0.789</b>  | <b>0.356</b> | <b>0.371</b> | <b>0.340</b> | <b>2.107</b> |

Üçüncü vektör ağı sunulur. Bu vektöre en çok benzeyen (x=3; y=2) konumunda bulunan nöron rekabeti kazanır (Çizelge 4.11). Bu nöronun ve komşularının ağırlıkları üçüncü girdi vektörüne göre değiştirilir (Çizelge 4.12) (Şekil 4.24d).

Çizelge 4.11 Üçüncü vektörün nöronlara olan Öklid uzaklıkları

| Nöron           | Uzaklık      |
|-----------------|--------------|
| x=1; y=1        | 2.938        |
| x=1; y=2        | 2.839        |
| x=1; y=3        | 2.560        |
| x=2; y=1        | 2.560        |
| x=2; y=2        | 2.715        |
| x=2; y=3        | 2.369        |
| x=3; y=1        | 2.577        |
| <b>x=3; y=2</b> | <b>2.310</b> |
| x=3; y=3        | 2.325        |

Çizelge 4.12 Kazanan nöron ve komşularının ağırlıklarının değişimi

| Nöron           | Referans vektörleri |              |              |              |              |              |               | Uzaklık      |
|-----------------|---------------------|--------------|--------------|--------------|--------------|--------------|---------------|--------------|
| x=1; y=1        | 0.824               | 0.715        | 0.911        | -0.787       | 0.776        | 0.591        | 0.305         | 2.938        |
| x=1; y=2        | -0.656              | 0.321        | 0.803        | 0.839        | 0.18         | 0.059        | 0.863         | 2.839        |
| x=1; y=3        | 0.262               | 0.495        | 0.203        | 0.477        | 0.45         | 0.66         | 0.705         | 2.560        |
| x=2; y=1        | 0.575               | 0.175        | 0.584        | 0.104        | 0.008        | 0.466        | 0.641         | 2.560        |
| <b>x=2; y=2</b> | <b>-0.181</b>       | <b>0.673</b> | <b>0.532</b> | <b>0.749</b> | <b>0.181</b> | <b>0.736</b> | <b>0.504</b>  | <b>2.551</b> |
| x=2; y=3        | 0.216               | 0.358        | -0.494       | -0.081       | 0.740        | 0.753        | 0.712         | 2.369        |
| <b>x=3; y=1</b> | <b>0.299</b>        | <b>0.623</b> | <b>0.726</b> | <b>0.277</b> | <b>0.735</b> | <b>0.650</b> | <b>-0.919</b> | <b>2.421</b> |
| <b>x=3; y=2</b> | <b>0.082</b>        | <b>0.194</b> | <b>0.029</b> | <b>0.297</b> | <b>0.097</b> | <b>0.423</b> | <b>0.500</b>  | <b>2.079</b> |
| <b>x=3; y=3</b> | <b>-0.288</b>       | <b>0.521</b> | <b>0.233</b> | <b>0.714</b> | <b>0.300</b> | <b>0.348</b> | <b>0.267</b>  | <b>2.184</b> |

Dördüncü vektör ağa sunulur. Bu vektöre en çok benzeyen (x=1; y=1) konumunda bulunan nöron rekabeti kazanır (Çizelge 4.13). Bu nöronun ve komşularının ağırlıkları dördüncü girdi vektörüne göre değiştirilmektedir (Çizelge 4.14) (Şekil 4.24e).

Çizelge 4.13 Dördüncü vektörün nöronlara olan Öklid uzaklıkları

| Nöron           | Uzaklık      |
|-----------------|--------------|
| <b>x=1; y=1</b> | <b>2.558</b> |
| x=1; y=2        | 3.596        |
| x=1; y=3        | 3.082        |
| x=2; y=1        | 2.705        |
| x=2; y=2        | 3.598        |
| x=2; y=3        | 2.953        |
| x=3; y=1        | 3.416        |
| x=3; y=2        | 3.055        |
| x=3; y=3        | 3.510        |

Çizelge 4.14 Kazanan nöron ve komşularının ağırlıklarının değişimi

| Nöron           | Referans vektörleri |              |              |               |              |              |              | Uzaklık      |
|-----------------|---------------------|--------------|--------------|---------------|--------------|--------------|--------------|--------------|
| <b>x=1; y=1</b> | <b>0.891</b>        | <b>0.494</b> | <b>0.901</b> | <b>-0.806</b> | <b>0.848</b> | <b>0.532</b> | <b>0.361</b> | <b>2.303</b> |
| <b>x=1; y=2</b> | <b>-0.526</b>       | <b>0.211</b> | <b>0.803</b> | <b>-0.729</b> | <b>0.260</b> | <b>0.055</b> | <b>0.863</b> | <b>3.378</b> |
| x=1; y=3        | 0.262               | 0.495        | 0.203        | 0.477         | 0.45         | 0.66         | 0.705        | 3.082        |
| <b>x=2; y=1</b> | <b>0.631</b>        | <b>0.074</b> | <b>0.60</b>  | <b>0.038</b>  | <b>0.098</b> | <b>0.438</b> | <b>0.655</b> | <b>2.541</b> |
| x=2; y=2        | -0.181              | 0.673        | 0.532        | 0.749         | 0.181        | 0.736        | 0.504        | 3.598        |
| x=2; y=3        | 0.216               | 0.358        | -0.494       | -0.081        | 0.740        | 0.753        | 0.712        | 2.953        |
| x=3; y=1        | 0.299               | 0.623        | 0.726        | 0.277         | 0.735        | 0.650        | -0.919       | 3.416        |
| x=3; y=2        | 0.082               | 0.194        | 0.029        | 0.297         | 0.097        | 0.423        | 0.500        | 3.055        |
| x=3; y=3        | -0.288              | 0.521        | 0.233        | 0.714         | 0.300        | 0.348        | 0.267        | 3.510        |

Vektörler ağı sunuldukça ilgili nöron ağırlıkları güncellenmeye devam eder. Örnekte, (x=2; y=2) konumunda bulunan nöron (mor), oluşan iki kümenin farklı yönlerde doğru çekilmesi ile kümeler arasında bir ayraç haline gelmiştir. Sonuç olarak nesne uzayında, Çizelge 4.15’de Öklid uzaklıkları verilen bu dört vektör, iki ayrı küme (pembe ve yeşil renkli nöronlar) oluşturmaktadır (1 ve 4; 2 ve 3).

Böylece kümeler, topolojik ilişkilerine göre 2 boyutlu uzayda yerleşirler. Akarsu ağına model genelleştirmesi için herhangi bir kümeleme yönteminin kullanımı, hangi kümelerin seçileceği sorununu ortaya çıkarmaktadır. Bu bağlamda, KDH yönteminin sağladığı topolojik ilişkiler kullanılarak bu sorun çözülebilir (Bölüm 5.3).

Çizelge 4.15 Öklid uzaklıkları verilen dört vektör ve oluşan iki ayrı küme

|          | Öklid uzaklığı |              |              |              |
|----------|----------------|--------------|--------------|--------------|
| Nöronlar | 1. vektör      | 2. vektör    | 3. vektör    | 4. vektör    |
| x=1; y=1 | 2.866          | 3.158        | 3.004        | <b>2.303</b> |
| x=1; y=2 | <b>0.920</b>   | 2.653        | 2.820        | 3.378        |
| x=1; y=3 | 1.670          | 2.454        | 2.560        | 3.082        |
| x=2; y=1 | 1.825          | 2.630        | 2.619        | 2.541        |
| x=2; y=2 | 1.265          | 2.362        | 2.551        | 3.598        |
| x=2; y=3 | 2.485          | 2.383        | 2.369        | 2.953        |
| x=3; y=1 | 2.548          | 2.352        | 2.421        | 3.416        |
| x=3; y=2 | 1.627          | 1.996        | <b>2.079</b> | 3.055        |
| x=3; y=3 | 1.350          | <b>1.970</b> | 2.184        | 3.510        |

#### 4.3.3 İstatistiksel Öğrenme Teorisi

Bu çalışmada, istatistiksel öğrenme teorisine dayalı sınıflandırma yöntemlerinden biri olan Destek Vektör Makineleri (DVM) kullanılmıştır. İstatistiksel öğrenme teorisi, Vapnik ve Chervonenkis tarafından 1960'lı yıllarda geliştirilmiştir. Veriden öğrenme problemi, çıktı vektörlerini (bağımlı değişkenler) mümkün olan en iyi şekilde tahmin eden  $f(x, w)$  fonksiyonunun seçilmesidir. Burada  $x$  girdi vektörü (bağımsız değişkenler),  $w$  parametre vektörüdür.

Sınıflandırma probleminde, bilinen çıktı vektörleri öğrenme makinesinin sınıf etiketleridir. 1:24,000 ölçekli akarsu ağının karakteristik nesnelerinden elde edilen girdi vektörlerinden oluşan eğitim setinin, orijinal 1:100,000 ölçekli akarsu ağına göre seçme (+1) ve eleme (-1) sınıf etiketleri belirlenir. Amaç, akarsu ağını seçme ve eleme olarak ikiye ayırabilen mümkün olan en iyi tahmin fonksiyonunun  $f(x, w)$  bulunmasıdır.

Tahmin fonksiyonu  $f(x, w)$ 'nın kalitesini ölçmek için  $L(y, f(x, w))$  ile gösterilen kayıp fonksiyonu (lost function) kullanılmalıdır. Bu kayıp fonksiyonu, öğrenme makinesi tarafından üretilen yakınsama (çıktı) değeri ile sınıf etiketi arasında bir kayıp değeri

tanımlar. İstatistiksel araştırmalarda, farklı uygulama alanları için birçok kayıp fonksiyonu kullanılmıştır [94], [96].

İkili sınıflandırma problemi için  $y = \{0,1\}$  olur ve tahmin fonksiyonu  $f(x, w)$ , 0 veya 1 değerlerinden birisini alabilir. İkili sınıflandırma problemi için kayıp fonksiyonu aşağıdaki gibi tanımlanır.

$$L(y, f(x, w)) = \begin{cases} 0, y = f(x, w) \\ 1, y \neq f(x, w) \end{cases} \quad (4.30)$$

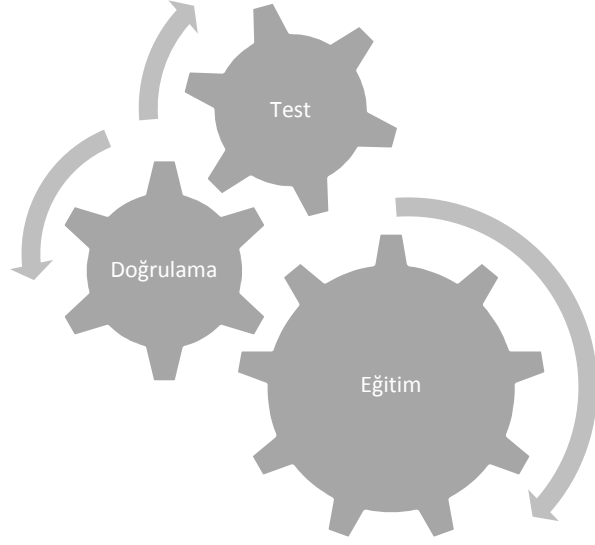
Geleneksel regresyon problemlerinde kullanılan en yaygın iki kayıp fonksiyonu;  $x$  bilinen girdi değeri,  $y$  bilinen çıktı değeri,  $f(x, w)$  tahmin fonksiyonu olmak üzere, (4.31) ve (4.32) eşitliklerinde verilen ortalama mutlak hata ve ortalama karesel hatadır [96].

$$L_1(y, f(x, w)) = \frac{1}{n} \sum_{i=1}^n |y_i - f(x_i, w_i)| \quad (4.31)$$

$$L_2(y, f(x, w)) = \frac{1}{n} \sum_{i=1}^n (y_i - f(x_i, w_i))^2 \quad (4.32)$$

Uygun modelin seçilmesi ve fazla/eksik (overfit/underfit) uyumun araştırılması için veri seti; eğitim (training), test ve doğrulama (validation) olmak üzere üçe ayrılır (Şekil 4.25) ve kayıp fonksiyonları; eğitim, test ve doğrulama setlerinde hesaplanmaktadır. Doğrulama seti, modelin eğitiminde kullanılmaz ve eğitim süresince modelin yüksek uyum (overfit) veya eksik uyum (underfit) problemini araştırmak için kullanılır. Model seçimi ise test hatalarının karşılaştırılması ile mümkündür. Doğrulama ve test setinde meydana gelen hata, genelleme hatası (generalization error); eğitim setinde meydana gelen hata ise deneysel hata (empirical error)'dır. Doğrulama seti, genelleme hatasının tahmin edilmesinde kullanılmaktadır [94].

Öğrenme makinesinin deneysel riskini minimumda tutmayı amaç edinen tümevarımsal metotlar, Deneysel Risk Minimizasyonu prensibiyle ifade edilirler. Deneysel risk, eğitim seti için yanlış sınıflandırılma olasılığını temsil etmektedir. Beklenen risk ise bilinmeyen dağılımdaki doğrulama ve test verisi için yanlış sınıflandırılma olasılığını temsil eder. Beklenen risk değerini minimumda tutmaya çalışan tümevarımsal öğrenme makinesi metotları ise Yapısal Risk Minimizasyonu (YRM) prensibindeki öğrenme metotları olarak bilinir. Beklenen risk değeri (4.33) eşitliği ile ifade edilebilir [94], [96], [97], [98].



Şekil 4.25 Veri setinin eğitim, test ve doğrulama olarak bölünmesi

$$R(w) = \int L(y, f(x, w))p(x)dx \quad (4.33)$$

Burada,  $R(w)$  beklenen riski,  $p(x)$  olasılık dağılımını,  $L(y, f(x, w))$  özel bir kayıp fonksiyonunu göstermektedir.

Gerekli modellerin oluşturulabilmesi ancak yapısal prosedürlerin tanıtılması, YRM tümevarımsal prensibi ve Vapnik-Chervonenkis (VC) teorisinin geliştirilmesi sayesinde başarılmıştır. Bu kavramlar istatistiksel öğrenme teorisinin ve dolayısıyla DVM öğrenme metodunun da temelini oluşturmaktadır.

YRM prensibi ve VC teorisindeki temel fikir; çok sayıda aday model arasından beklenen risk veya genelleme hatasını minimum yapacak oranda karmaşıklığa sahip, aynı zamanda deneysel riski ve yakınsama fonksiyonlarının karmaşıklığını minimum yapan, güçlü bir genelleme yeteneğine sahip olan modelin seçilmesidir. Böylece, seçilen optimum öğrenme makinesi, verilen eğitim kümesi için en uygun kapasiteye sahip olacaktır.

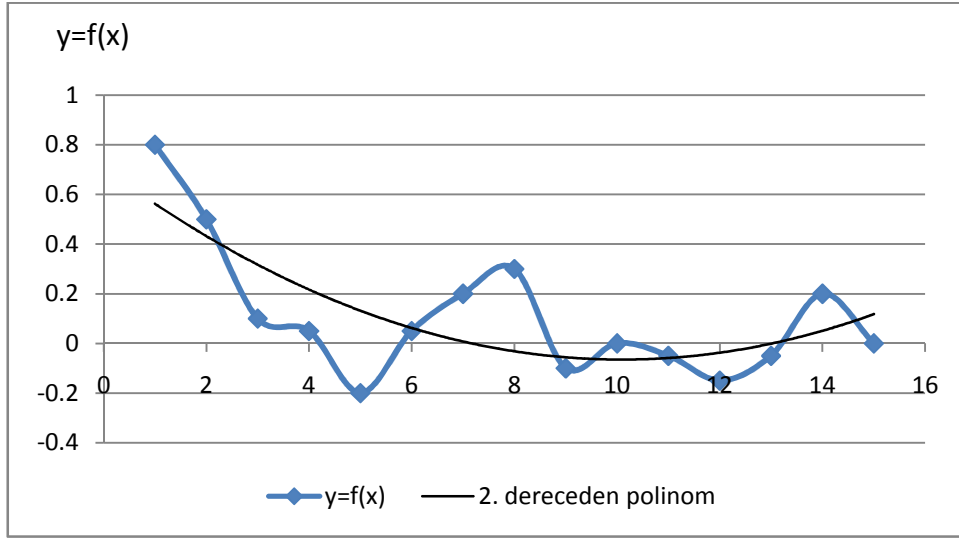
Bu fikir, sınıflandırma veya regresyon işlemleri gerçekleştirilirken meydana gelen fazla uyum ve eksik uyum problemlerinin analizinde de gözlenebilir. Eğer amaca uygun ve küçük bir veri kümesi ile çalışılıyorsa, fazla uyum durumunun oluşması beklenir. Eğer basit bir yakınsama fonksiyonu yetersiz bir şekilde kullanılıyorsa, veri ile uygun yakınsama fonksiyonu arasındaki sapma çok büyük olabilir. Bu takdirde de eksik uyum

durumu oluşabilir. Bu nedenle, deneysel riskin veya deneysel hatanın minimum tutulması, beklenen riskin veya genelleme hatasının küçük olmasını garanti edemez. Eksik uyum ve fazla uyum problemleri arasında bir denge her zaman mevcuttur. Bu denge noktası, düzenleyici bir parametre yardımıyla sağlanabilir. İstatistiksel öğrenme teorisinde, uygun fonksiyonun seçildiği fonksiyonlar kümesinin karmaşıklığı (kapasitesi), VC boyutu kullanılarak tanımlanır. YRM prensibini temel alan istatistiksel öğrenme teorisi aynı anda hem deneysel riski hem de VC boyutunu minimum tutarak öğrenme makinesinin genelleme yeteneğini kontrol edebilir [96].

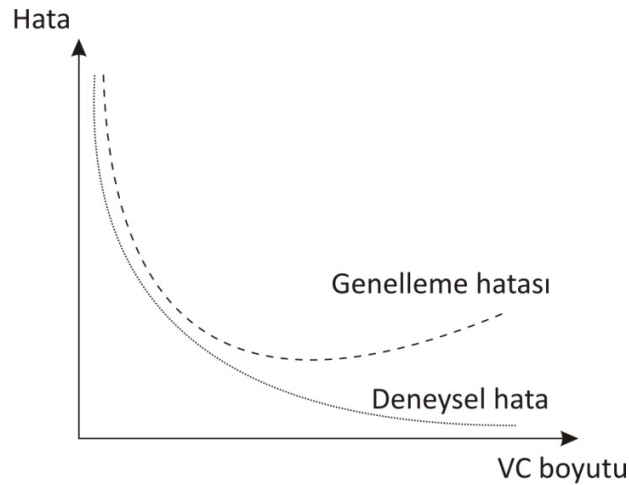
VC boyutu,  $f(x, w)$  fonksiyonu tarafından sınıflara ayrılabilen nokta sayısı olarak da tanımlanabilir. Bir sınıflandırıcının daha çok veriyi sınıflandırabilmesi daha fazla kapasiteye sahip olmasını gerektirir. Fakat çok yüksek kapasite fazla uyum sorunu doğuracaktır. Kapasite, genellikle parametre sayısı ile ilişkilidir. Fakat bazı istisnai durumlar söz konusu olabilmektedir. Örneğin,  $f(x) = \sin(wx)$  sadece bir parametreye ( $w$ ) sahiptir ancak sinüzoidal bir fonksiyon olduğu için yüksek kapasitelidir [94].

Genellikle, deneysel hata eğitim setinde minimize edilmektedir. Ancak, deneysel hatada global minimuma erişilirse, model eğitim setine muhtemelen çok fazla uyum gösterecek ve bilinmeyen veri üzerindeki yanlış tahmin sayısı artacaktır (fazla uyum). Diğer taraftan, model yeterince optimize edilmezse veya tahmin fonksiyonu çok basitse, veri model tarafından yeterince temsil edilemeyecek ve eksik uyum gelişecektir. Şekil 4.26'da, eğitim verisi noktalarından geçen iki eğri kullanılarak model tahmin edilmeye çalışılmaktadır. Amaç,  $\{x_i, y_i\}$  eğitim setinde  $n$ -boyutlu ( $n$  adet öznitelik veya bağımsız değişken) bir girdi vektörü olan  $x_i$  ile  $y_i$  çıktı vektörlerini tahmin eden fonksiyonun bulunması ve öğrenmenin gerçekleştirilmesidir. Şekil 4.26'da, düz çizgi ile gösterilen 2. dereceden polinom, eğitim seti için çok basit; noktalardan geçen eğri ise çok karmaşıktır. Düz çizgi ile gösterilen eğri, eksik uyum problemi; kesikli çizgi ile gösterilen eğri ise fazla uyum problemi ortaya çıkarmaktadır. Bir başka deyişle, düz çizgi ile gösterilen 2. dereceden polinomun tahmin etme gücü çok zayıf olacaktır. Kesikli çizgi ile gösterilen fonksiyonun karmaşıklığı (kapasitesi) yani VC boyutu çok yüksek olduğundan, model eğitim setine çok fazla uyum gösterecek ve test setinde hiç görmediği örnekleri tahmin etme gücü azalacaktır. Şekil 4.27'de gösterildiği gibi, VC

boyutunun artması ile beraber eğitim setinde deneysel hata sıfıra yaklaşırken, test setinde genelleme hatası bir denge noktasından sonra yükselmektedir [94].



Şekil 4.26 Regresyon problemine bir örnek [94]



Şekil 4.27 Model karmaşıklığının (VC boyutunun) artması ile genelleme ve deneysel hatanın değişimi [94]

Veri setinin eğitim, test ve doğrulama setlerine ayıramayacak kadar küçük olması durumunda, istatistiksel yeniden örneklemeyle dayalı teknikler (özellikle “jackknife” yöntemleri) kullanılmaktadır. Jackknife yöntemleri, örnekleme (veri setinde) sistematik olarak bir gözlemi ya da gözlem setini dışarıda bırakarak standart hatayı yeniden hesaplamaktadır. En popüler yöntem, çapraz doğrulama (cross validation)’dır. Birini dışarıda bırakarak çapraz doğrulama (leave-one-out cross validation) ve k-kat çapraz doğrulama (k-fold cross validation) yöntemleri, çapraz doğrulamaya örnek olarak

verilebilmektedir. Birini dışarıda bırak yönteminde, her seferinde bir nesneye ait veri dışarıda bırakılarak eğitim gerçekleştirilmektedir. Dışarıda bırakılan veri test edilmektedir. Böylece her veri eğitim ve test setlerinde kullanılabilir. K-kat yönteminde ise veri,  $k$  adet eşit gruba ayrılmaktadır. Bir grup test, geriye kalan  $(k - 1)$  grup ise eğitim verisi olarak kullanılır. Bu durum  $-k$  kez tekrarlanarak- her verinin bir kez eğitim bir kez de test verisi olarak kullanılabilmesini sağlamaktadır. Hata değerlerinin ortalaması alınarak, modelin doğruluk değeri belirlenmektedir. K-kat çapraz doğrulama yöntemi, bilgisayar kapasitesinin kullanımı açısından daha uygun olmaktadır.

K-kat çapraz doğrulama işlem adımları aşağıdaki gibi sıralanabilir [94]:

- a) Veri setinin rastgele olarak  $k$  parçaya ayrılması,
- b) Veri setinde ilk parçanın dışarıda bırakılması,
- c) Kalan  $(k - 1)$  verinin eğitilmesi,
- d) Dışarıda bırakılan parçanın test edilmesi ve hatanın eşitlik (4.31) veya (4.32)'e göre hesaplanması,
- e) Tüm parçalar için b-d adımlarının tekrar edilmesi ve
- f)  $k$  adet hatanın ortalamasının hesaplanması.

Bulanık modeller, sinir ağları, gauss ağları ve DVM gibi metotlar, bazı koşulları sağlayan  $m$  adet sabit fonksiyonun birleşimi olarak temsil edilebilirler. Bu nedenle, doğrusal, parametrik ve  $m$  adet sabit fonksiyonun birleşiminden oluşan modeller için VC boyutu,  $m + 1$  olarak düşünülür. Bu bakımdan, lineer modellerde optimum fonksiyonun bulunması, lineer olmayan modellere göre daha kolaydır. Doğrusal fonksiyon yapısına sahip öğrenme makinelerinin VC boyutu kolaylıkla hesaplanabilirken, doğrusal olmayan fonksiyonlar için hesaplanması oldukça zordur [96].

İstatistiksel öğrenme teorisine göre, eğitim verisi ile tutarlı olan minimum kapasiteli fonksiyonların bulunması, daha iyi genelleme yeteneği sağlamaktadır. Destek Vektör Makineleri, bu düşünceyle, istatistiksel öğrenme teorisi çatısı altında geliştirilmiş bir öğrenme yöntemidir [94].

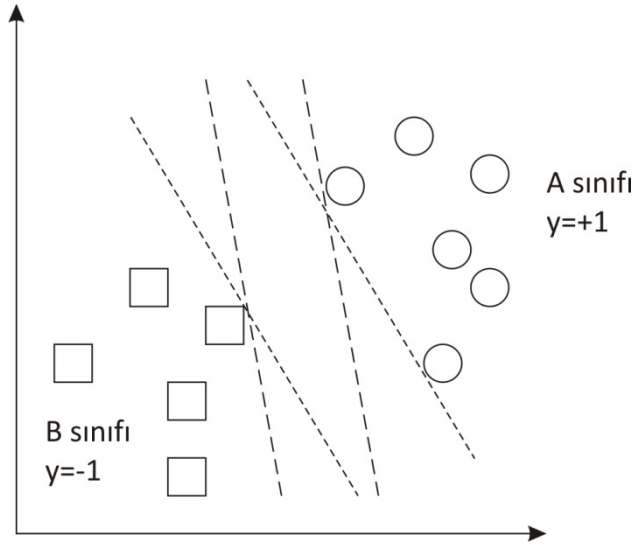
#### 4.3.4 Destek Vektör Makineleri

Destek Vektör Makineleri (DVM) (Support Vector Machines), 1990'lı yılların başında bulunan güçlü bir sınıflandırma ve tümevarımsal bir makine öğrenmesi yöntemidir [99]. Veriyi birbirinden ayırmak için en uygun fonksiyonun tahmin edilmesi esasına dayanır [74]. Vapnik ve Chervonenkis tarafından geliştirilmiş olan istatistiksel öğrenme teorisinden türetilmiş bir yöntemdir [100]. DVM algoritmalarının amacı, sınıflandırma yaparken iki sınıfı birbirinden ayıracak en uygun hiper düzlemin (ayırt etme yüzeyinin) bulunmasıdır. Yazı tanıma, nesne tanıma, ses tanıma ve yüz tanıma gibi uygulamalarda kullanılmaktadır. Uzaktan algılama verilerinin sınıflandırmasında da sıklıkla kullanılmaktadır [34]. Bu yöntem, istatistiksel öğrenme teorisine dayalı olarak, eğitim verileri ve çıktıları arasında doğrusal olmayan bir bağımlılık fonksiyonu bulmaya çalışmaktadır. DVM, doğrusal olmayan fonksiyon yardımıyla, orijinal girdi uzayından, doğrusal olarak kolayca sınıflandırma yapabileceği daha büyük boyutlu nitelik uzayına dönüşüm yapar [96]. İkili sınıflandırma probleminde,  $y$  bağımlı değişkenler (+1 veya -1 olarak) etiketlenirler. Amaç,  $y$  bağımlı veya sınıf değişkeniyken, etiketleme fonksiyonuna yakınsayan bir  $f$  modelini oluşturmaktır. Bunu yerine getirmek için +1 ve -1 sınıflarına en iyi ayıran doğru, düzlem ya da hiper düzlem oluşturulur [99]. Akarsu ağının model genelleştirmesinde de, akarsu ağı benzer biçimde seçme ve eleme olarak iki sınıfa ayrılabilir.

DVM eğitim setinin  $\{x_i, y_i\}$  biçiminde olduğunu varsayalım [94].

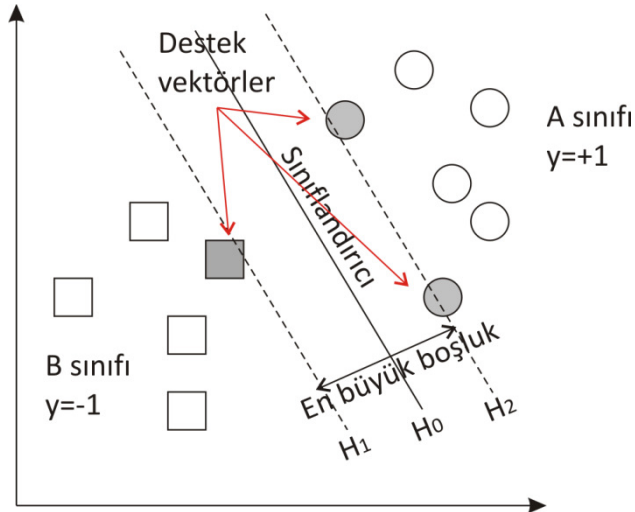
$$y_i = \begin{cases} +1 & \text{eğer } x_i \in A \\ -1 & \text{eğer } x_i \in B \end{cases} \quad (4.34)$$

Şekil 4.28'de, verilerin birbirinden farklı biçimlerde doğrusal olarak ayrılabilceği gösterilmiştir. Veri farklı ve çok sayıda doğru ile ayrılabilir. Çok boyutlu uzayda bu doğruların yerini hiper düzlemler alacaktır [74]. İkili sınıflandırmada, bir doğrusal karar yüzeyi olarak adlandırılan doğru, düzlem veya hiper düzlem nesneleri iki gruba ayırır [99].



Şekil 4.28 İki boyutlu uzayda doğrusal olarak ayrılabilen veriler [74]

Problemin en uygun çözümü için, birbirinden en uzak iki hiper düzlem belirlenmelidir. En uzak iki hiper düzlemin ortasında yer alan hiper düzleme *en uygun hiper düzlem* adı verilmektedir. Şekil 4.29'da,  $H_1$  ve  $H_2$  hiper düzlemlerinin ortasını oluşturan  $H_0$  hiper düzlemi, veriyi ikiye ayıran en uygun doğrusal hiper düzlemdir [74].  $H_1$  ve  $H_2$  hiper düzlemleri üzerinde yer alan ve  $H_0$  hiper düzlemine en yakın noktalara destek vektörler denir. Şekil 4.29'da, destek vektörler gri renkte gösterilmiştir.  $H_1$  ve  $H_2$  hiper düzlemlerinin aralarındaki uzunluğa en büyük boşluk veya marjin ( $m$ ) denir.



Şekil 4.29 Doğrusal olarak ayrılabilen veriler arasındaki en büyük boşluk [74]

DVM’de temel fikir; en uygun hiper düzlemlerin geçirilmesini sağlayacak destek vektörlerinin belirlenmesi ve sonuç olarak genelleme hatasının minimize edilmesidir [94].

$H_0$  düzlemi (doğrusal karar yüzeyi) şu şekilde ifade edilebilir:

$$H_0: wx + b = 0 \quad (4.35)$$

$H_1$  hiper düzlemi:

$$H_1: wx + b = 1 \quad (4.36)$$

$H_2$  hiper düzlemi:

$$H_2: wx + b = -1 \quad (4.37)$$

Burada,  $w$  ağırlık vektörünü,  $b$  ise sabit bir sayıyı göstermektedir. Amaç,  $w$  parametresini bulmaktır. (4.35) eşitliği ile bilinen hiper düzlemin üst ve alt tarafında kalan noktalar (4.38) ve (4.39) eşitlikleri ile belirlenmektedir.

$$wx_i + b \geq +1 \text{ ise } y_i = +1 \quad (4.38)$$

$$wx_i + b \leq -1 \text{ ise } y_i = -1 \quad (4.39)$$

(4.38) ve (4.39) eşitlikleri birleştirilerek (4.40) eşitliğine dönüştürülebilir.

$$y_i(x_iw + b) - 1 \geq 0 \quad (4.40)$$

Bir  $x_p$  destek vektörü ile  $H_0: wx + b = 0$  ile gösterilen en uygun hiper düzlem arasındaki dik uzaklık  $d$ , (4.41) eşitliği ile hesaplanır.

$$d = \frac{|wx_p + b|}{\|w\|} \quad (4.41)$$

O halde, problemin en uygun çözümü için yani boşluğun maksimizasyonu için ağırlık vektörünün minimizasyonu gerekmektedir. Bu durumda, optimizasyon problemi aşağıdaki gibi ifade edilebilir.

$$\text{Min: } \frac{1}{2} \|w\|^2 \text{ ve Kısıt: } y_i(x_iw + b) - 1 \geq 0 \quad (4.42)$$

(4.42) eşitliği doğrusal olmayan bir optimizasyon problemidir. Bu problem Lagrange çarpanları kullanılarak çözülebilmektedir.  $\alpha_i$  Lagrange çarpanları olmak üzere, Lagrange fonksiyonu:

$$L(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^l \alpha_i y_i (x_i w + b) + \sum_{i=1}^l \alpha_i \quad (4.43)$$

(4.43) eşitliğinde  $w$  ve  $b$ 'ye göre kısmi türevler alınarak sıfıra eşitlenirse:

$$w = \sum_{i=1}^l \alpha_i y_i x_i \text{ ve } \sum_{i=1}^l \alpha_i y_i = 0 \quad (4.44)$$

Bu ifadelere Karush-Kuhn-Tucker koşulları adı verilmektedir. O halde  $L(w, b, \alpha)$  Lagrange fonksiyonunun duali şu şekilde yazılabilmektedir:

$$L(w, b, \alpha) = \sum_i \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j x_i^T x_j \quad (4.45)$$

En uygun hiper düzlemi bulmak için (4.44) eşitliğinde verilen Lagrange fonksiyonunun duali maksimize edilerek  $\alpha_i$  bulunur ve destek vektörler belirlenir.  $\alpha_i \neq 0$  sağlayan noktalar destek vektörleridir. Sınıflandırıcının çıktı değerleri (4.46) eşitliği ile belirlenir.

$$f(x) = \text{sign}(\sum_i \alpha_i y_i x_i^T x) \quad (4.46)$$

$$\text{sign}(z) = \begin{cases} -1 & \text{eğer } z < 0 \\ 0 & \text{eğer } z = 0 \\ 1 & \text{eğer } z > 0 \end{cases}$$

Çözümün karmaşıklığı veri sayısına bağlı değil, destek vektör sayısına bağlıdır. Eğer veriler doğrusal bir hiper düzlem ile ayıramıyorsa, bazı noktalar hatalı olarak hiper düzlemin yanlış tarafına düşmektedir. Verilerin doğrusal bir düzlemle ayıramama durumunda, negatif olmayan ve hataları ifade eden  $\xi_i$  gevşek değişkenlerinin (slack variables) optimizasyon modeline eklenmesi sağlanarak soruna çözüm aranır.

$$x_i w + b \geq +1 - \xi_i \text{ ise } y_i = +1 \quad (4.47)$$

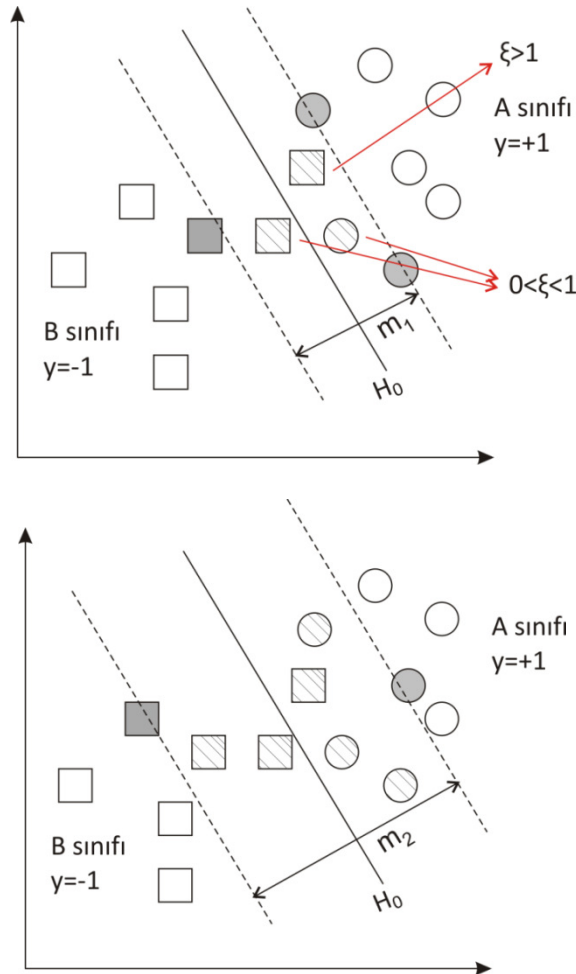
$$x_i w + b \leq -1 + \xi_i \text{ ise } y_i = -1$$

$$\xi_i > 0 \quad \forall i$$

$\xi_i$  gevşek değişkenler, hiper düzlemin yanlış tarafında sınıflandırılan noktalardır. Amaç fonksiyonu:

$$\min \frac{1}{2} \|w\|^2 + C \sum_i \xi_i \quad (4.48)$$

haline dönüşmektedir. Sınıflandırma hatasını etkileyen ceza (penalty) parametresi  $C$  ( $C > 0$ ), kullanıcıya bağlı bir parametredir ve yanlış sınıflandırılan nokta sayısı ile sınıflandırıcının karmaşıklığı arasındaki ilişkiyi kontrol altında tutmaktadır. Eğer  $C$  küçük seçilirse, ideal pozisyonda olmayan birçok veriye izin verilir. Aksi takdirde, ideal pozisyonda olmayan çok az sayıda veri ortaya çıkar.  $C$  için optimal değer, VC teorisi tarafından analitik olarak belirlense de, pratikte çapraz doğrulama kullanılarak belirlenmektedir. Şekil 4.30'da, kullanıcıya bağlı bir parametre olan  $C$ 'nin alacağı değerlere göre hiper düzlemlerin değişimi gösterilmiştir. Burada  $\xi_i > 1$  olan veriler hiper düzlemin yanlış tarafında kalan, yani doğrusal olarak ayrılmayı önleyen gözlem değerleridir.  $0 < \xi_i < 1$  ise, hiper düzlemin doğru yanında yer alan, ancak marjin bölgesi ( $m$ ) içinde kalan gözlem değerlerini ifade eder [74], [94].

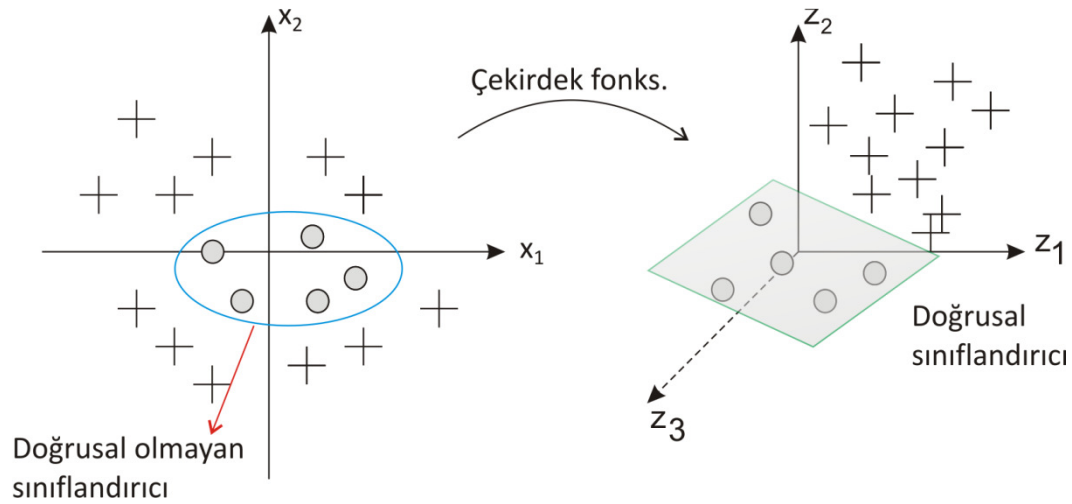


Şekil 4.30  $C$ 'nin alacağı değerlere göre marjin [74]

Doğrusal olarak verilerin ayırt edilemediği durumlarda, verileri doğrusal olarak sınıflandırma yapabileceği daha büyük boyutlu uzaylara taşıyacak çekirdek fonksiyonları kullanılabilmektedir. Şekil 4.31’de görüldüğü gibi,  $X = (x_1, x_2)$  değişkenleri 2 boyutlu uzayda ( $X = \mathbb{R}^2$ ) bir elips ile ikiye ayrılabilir. Bir çekirdek fonksiyon  $\phi$ , (4.49) eşitliğiyle 3 boyutlu uzaya ( $Z = \mathbb{R}^3$ ) dönüştürülebilir. Böylece veriler bir doğrusal sınıflandırıcı ile ayrılabilir.

$$Z = (z_1, z_2, z_3) = \phi(x) = (x_1^2, \sqrt{2}x_1x_2, x_2^2) \quad (4.49)$$

Daha büyük boyutlu bir uzaydaki doğrusal bir sınıflandırıcı, orijinal uzaydaki doğrusal olmayan bir sınıflandırıcıya karşılık gelmektedir (Şekil 4.31).



Şekil 4.31 İki boyutlu uzayda doğrusal olmayan bir sınıflandırıcı ile ayrılabilen verilerin bir çekirdek fonksiyonu ile üç boyutlu uzaya taşınması ve bir doğrusal sınıflandırıcı ile ayrılabilmesi [77]

Genel çekirdek fonksiyonu yapısı (4.50) eşitliği ile verilmektedir.

$$K(x_i, x_j) = \phi(x_i)^T \phi(x_j) \quad (4.50)$$

Aşağıda belirtilen çekirdek fonksiyonlar uygulamalarda sıkça kullanılmaktadır [77].

$$\text{Doğrusal: } K(x_i, x_j) = x_i^T x_j \quad (4.51)$$

$$\text{Polinom: } K(x_i, x_j) = (x_i^T x_j + a)^r \quad (4.52)$$

Radyal tabanlı fonksiyon (Radial Basis Function):

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\gamma^2}\right), \gamma > 0 \quad (4.53)$$

$$\text{Sigmoid: } K(x_i, x_j) = \tanh(ax_i^T x_j + b) \quad (4.54)$$

Radyal tabanlı fonksiyonda; Kernel parametresi ( $\gamma$ ), sınıflandırıcının yumuşaklığını kontrol etmektedir. Uygun  $\gamma$  değeri,  $C$  parametresindeki gibi çapraz doğrulama sonuçlarına göre kullanıcı tarafından belirlenmelidir [94].

Lagrange dual fonksiyonundaki (eşitlik 4.45)  $x_i^T x_j$  çarpımını daha büyük boyutlu özellik uzayı için yeniden tanımlamak gerekmektedir.  $x_i^T$  ifadesinin dönüşümü için  $\phi(x_i)^T$  ve  $x_j$  değişkeninin dönüşümü için  $\phi(x_j)$  fonksiyonları kullanılır. O halde daha büyük boyutlu özellik uzayı için (4.55) eşitliği şu şekilde olacaktır:

$$L(w, b, \alpha) = \sum_i \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j \phi(x_i)^T \phi(x_j) \quad (4.55)$$

Bu durumda dual optimizasyon problemi aşağıdaki gibi ifade edilebilir:

$$\text{Max: } L(w, b, \alpha) = \sum_i \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j \phi(x_i)^T \phi(x_j) \quad (4.56)$$

$$\text{Kısıt: } \sum_{i=1}^l \alpha_i y_i = 0 \quad \alpha_i \geq 0$$

O halde sınıflandırıcı aşağıdaki şekilde olmalıdır [74], [94].

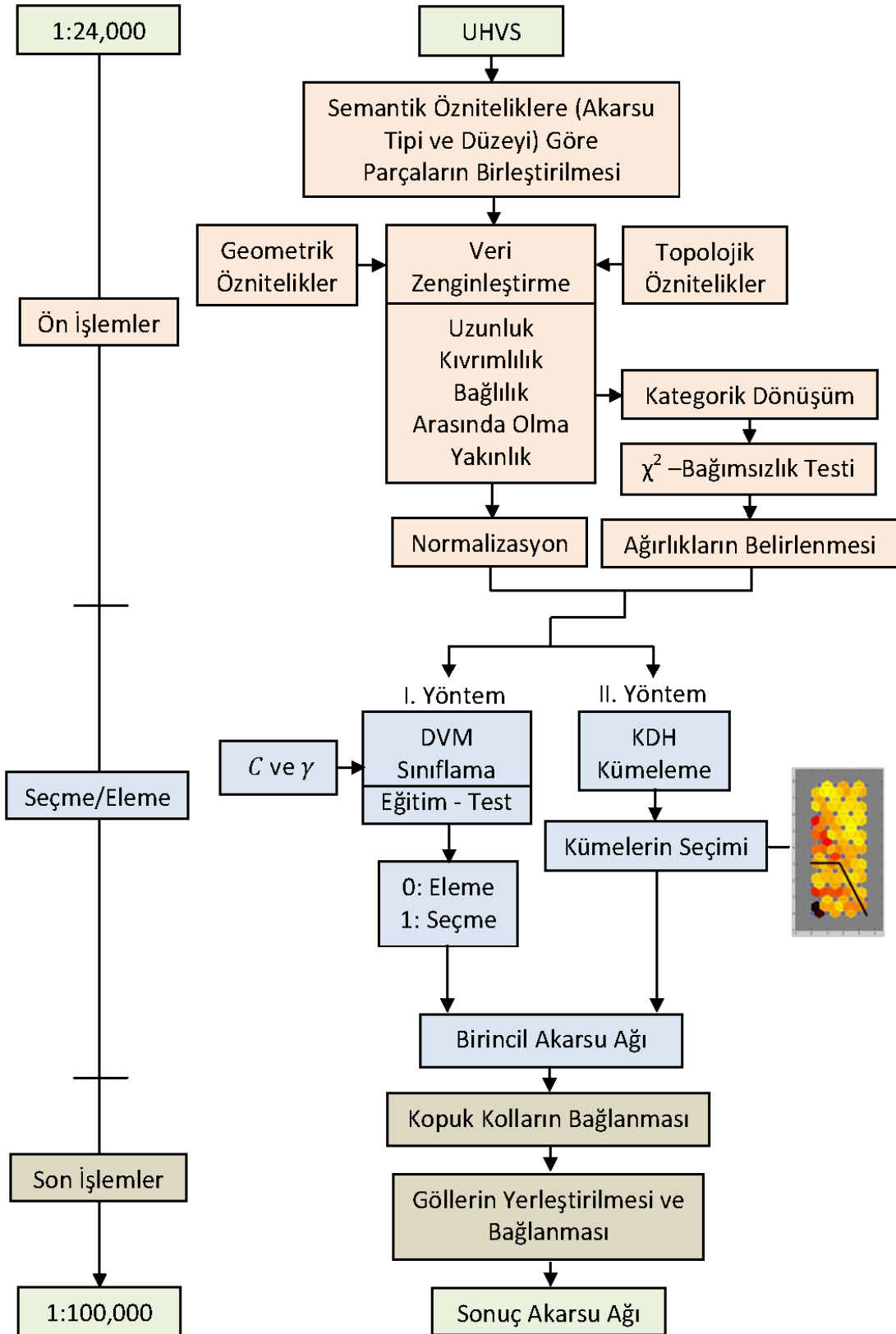
$$f(x) = \text{sign}\left(\sum_i \alpha_i y_i K(x_i, x_j)\right) \quad (4.57)$$

## BÖLÜM 5

---

### UYGULAMA

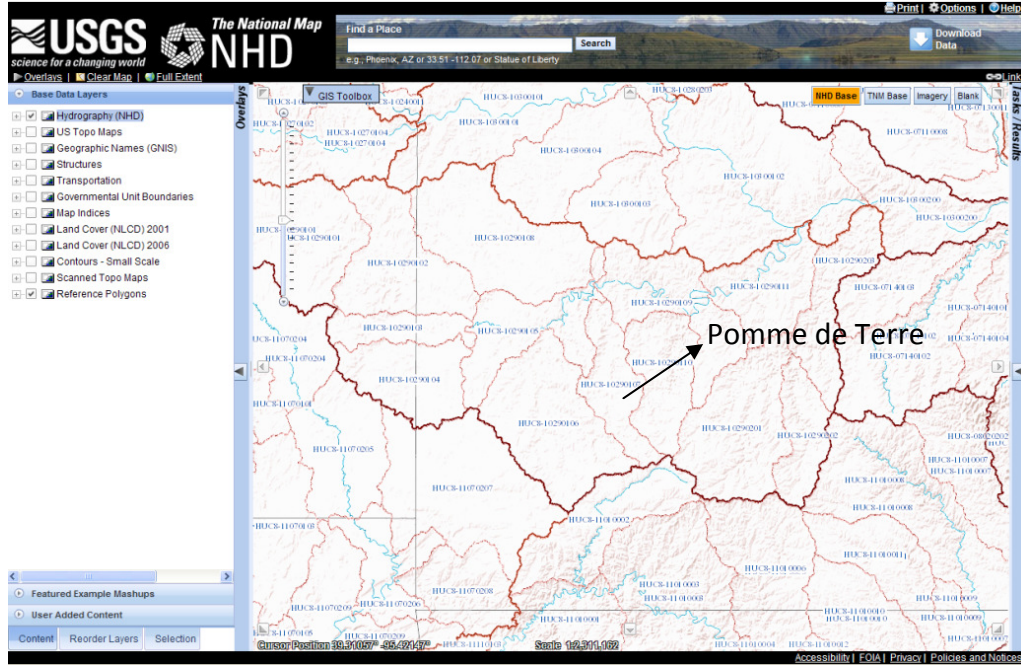
Önerilen yöntemlerin amacı, denetimsiz bir sınıflandırma yaklaşımı olan KDH ve istatistiksel öğrenme teorisine dayanan bir yaklaşım olan DVM kullanarak, hedef ölçekte gösterilecek olan nesnelerin seçimini sağlamaktır. Girdi matrisi, akarsu nesnelerinin karakteristik özniteliklerinin bir setidir. Bu nedenle, öncelikle uygun öznitelikler ve ağırlıkları belirlenmiştir. Daha sonra, seçme işlemi KDH yöntemiyle tam otomatik olarak; DVM yöntemiyle ise yarı otomatik olarak gerçekleştirilmiştir. KDH ile elde edilen kümelerden hangilerinin seçileceğini belirlemek için yeni bir yöntem geliştirilmiştir. Metodolojiye ait akış şeması Şekil 5.1’de gösterilmiştir.



Şekil 5.1 Metodoloji akış şeması

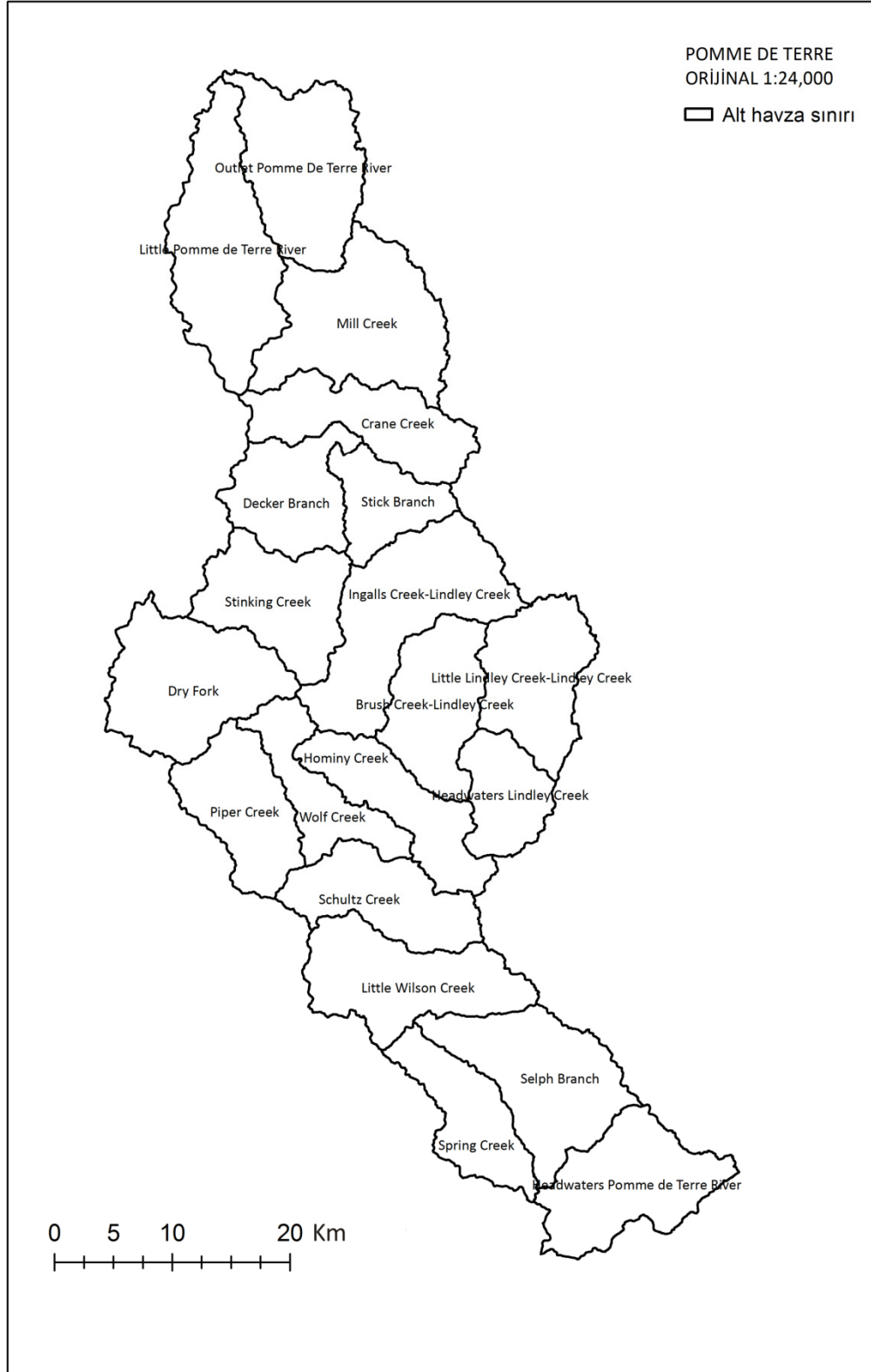
## 5.1 Veriler ve Çalışma Alanı

Bu çalışmada, veri kaynağı USGS tarafından geliştirilen Ulusal Hidrografik Veri Setleridir. Veriler; 1:24,000 ölçekli NHDH0207, NHDH1029 ve 1:100,000 ölçekli NHDM0207, NHDM1029 isimli 4 adet settir. Ayrıca çalışma alanına ait (1/3) saniye çözünürlüklü (yaklaşık 10 metre) ve 2.17 metre duyarlıklılı (RMSE: Root Mean Square Error-Karesel Ortalama Hata) sayısal yükseklik modelleri kullanılarak hidrojeomorfometrik analizler yapılmıştır. USGS UHVS veri tabanında bulunan veri setlerine web üzerinden “USGS TNM 2.0 Viewer” ile erişilmiştir (Şekil 5.2). Veri setleri, Pomme de Terre (PT) ve South Branch Potomac (SBP) adlı havzaları içermektedir. Havzalar doku, kaya tipi ve jeolojik yapı bakımından farklı topografik özellikler göstermektedir. Farklı akarsu ağı desenleri içerdikleri için bu havzalar tercih edilmiştir.

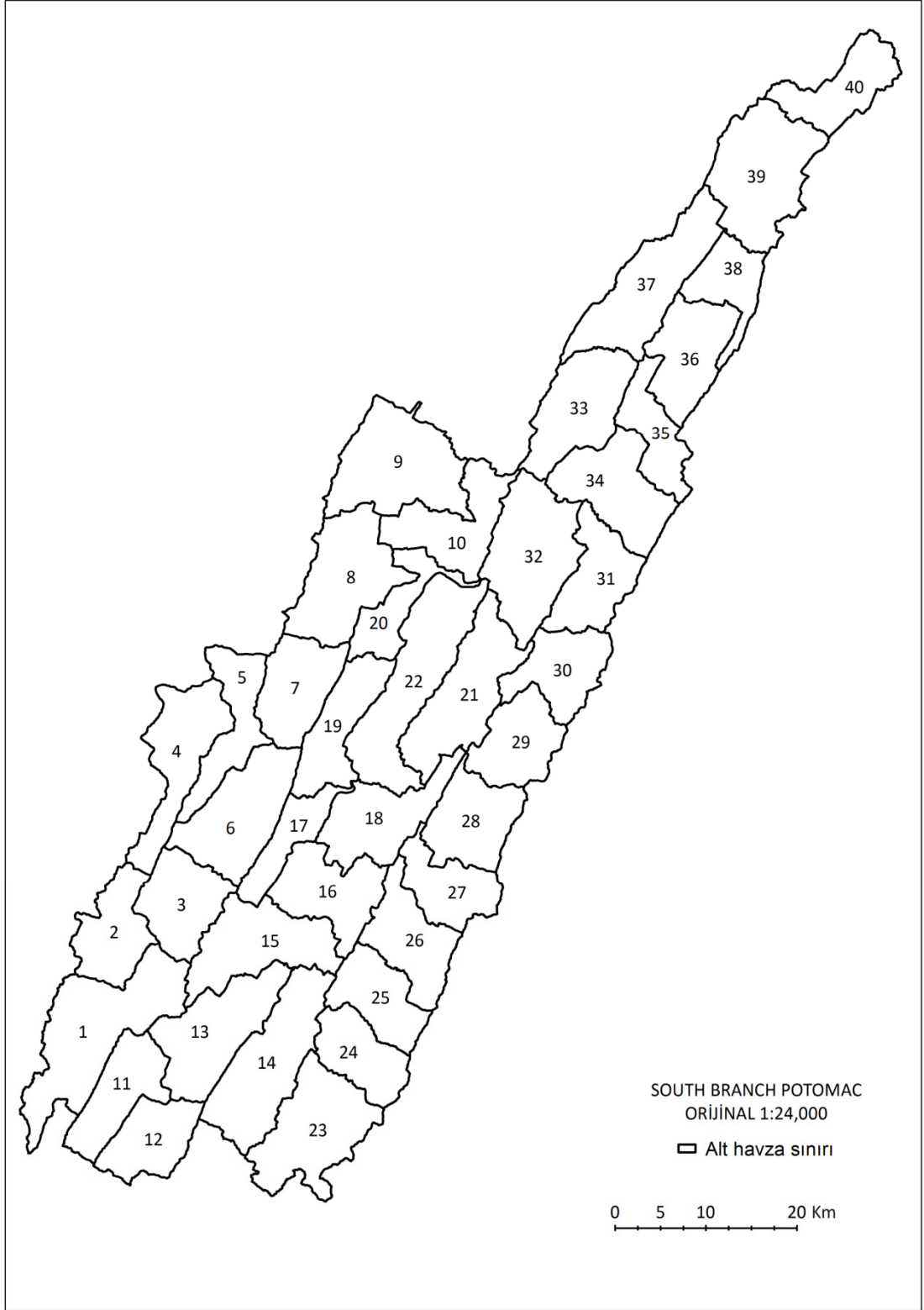


Şekil 5.2 USGS TNM 2.0 Viewer

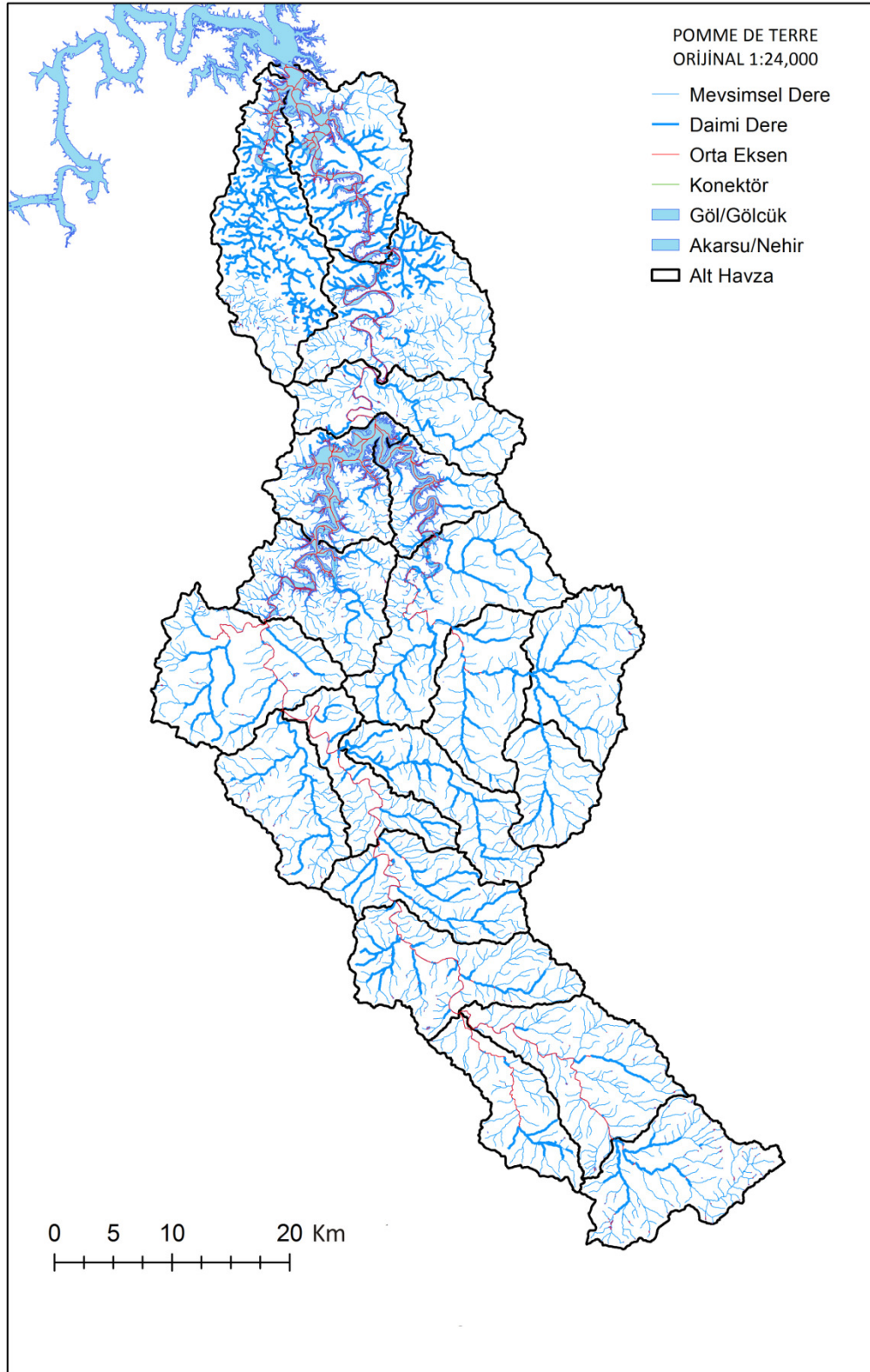
PT, ağaçsı desen ve SBP, kafes ve dikdörtgen desen göstermektedir. [101] ve [102] tarafından, havzalar dağlık (SBP) ve tepelik (PT) olarak sınıflandırılmıştır. PT, 20 alt havzadan (Şekil 5.3); SBP ise 40 alt havzadan (Şekil 5.4) oluşmaktadır. SBP'nin havza isimleri veri tabanında bulunmamaktadır. Bu nedenle numara verilmiştir (Şekil 5.4). Şekil 5.5 ve 5.6'da, PT ve SBP için orijinal 1:24,000 ölçekli akarsu ağları; Şekil 5.7 ve 5.8'de, PT ve SBP için orijinal 1:100,000 ölçekli akarsu ağları gösterilmiştir. Bu tezde, coğrafi analizler ArcGIS 10 yazılımında gerçekleştirilmiştir.



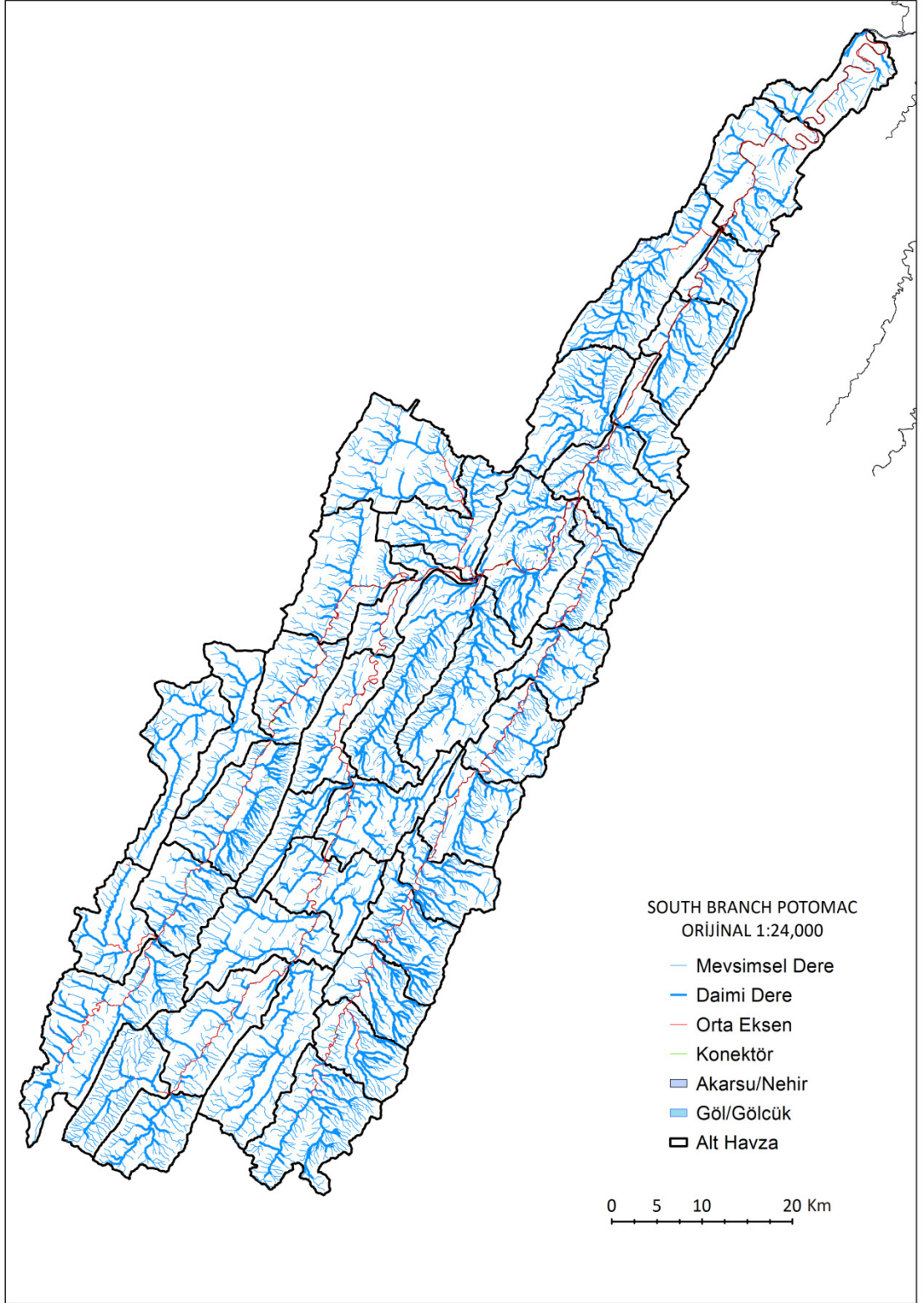
Şekil 5.3 PT alt havza sınırları



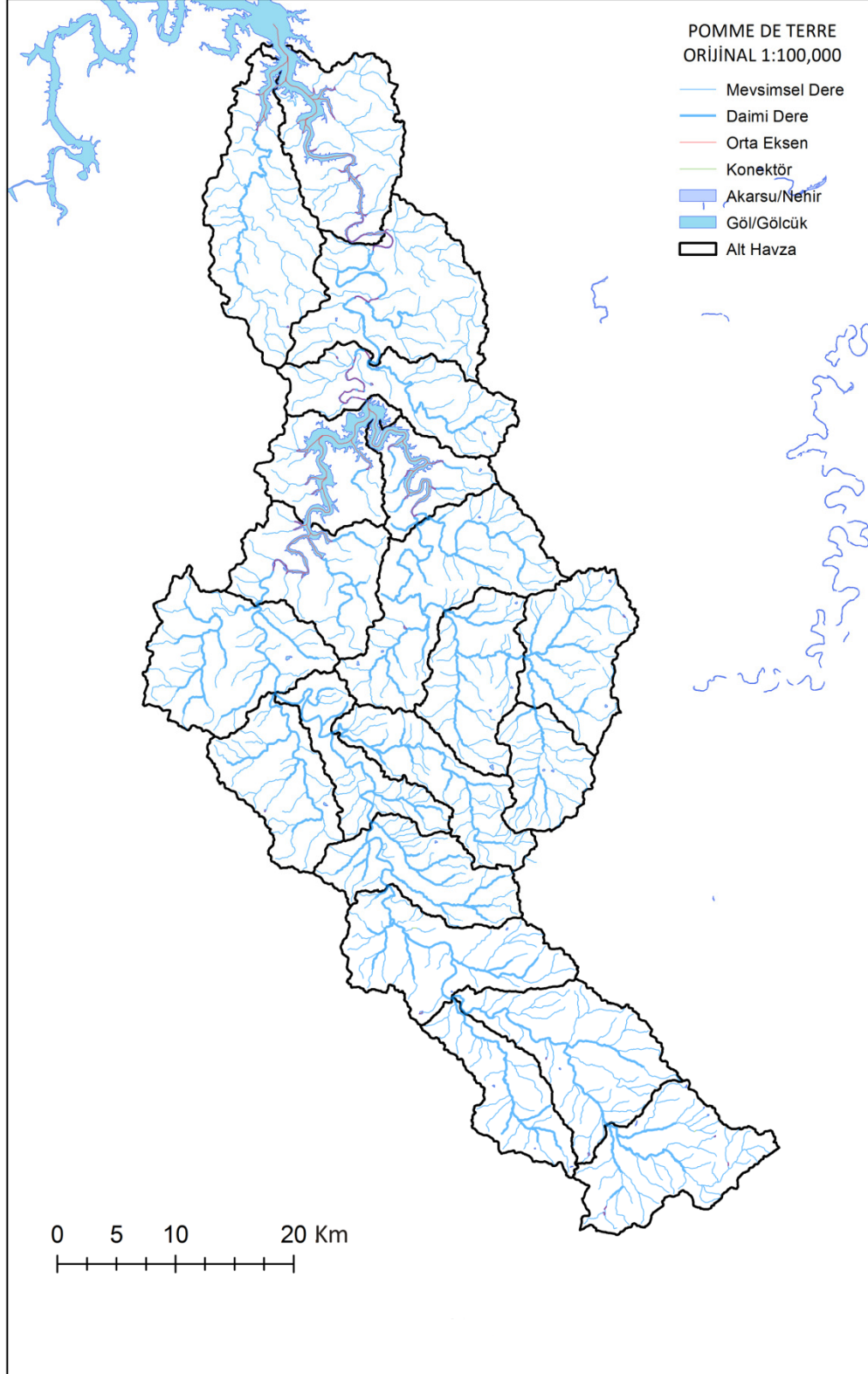
Şekil 5.4 SBP alt havza sınırları



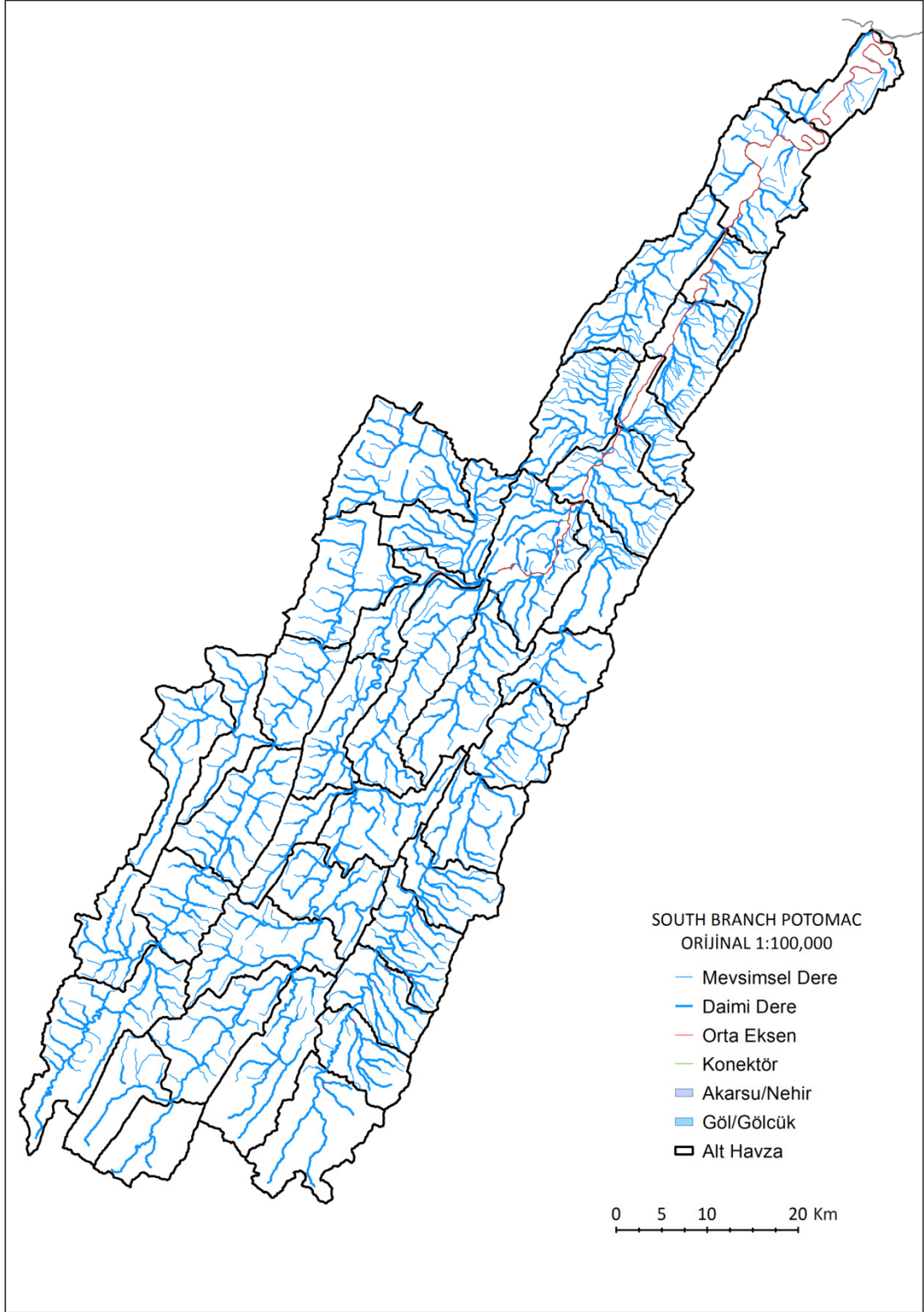
Şekil 5.5 Orijinal 1:24,000 ölçekli PT'nin 1:400,000 ölçeğinde gösterimi



Şekil 5.6 Orijinal 1:24,000 ölçekli SBP'nin 1:500,000 ölçeğinde gösterimi



Şekil 5.7 Orijinal 1:100,000 ölçekli PT'nin 1:400,000 ölçeğinde gösterimi



Şekil 5.8 Orijinal 1:100,000 ölçekli SBP'nin 1:500,000 ölçeğinde gösterimi

Çizelge 5.1'de, havzaların; eyalet, havza numaraları ve fizyografik bölüm bilgisi verilmiştir. Akarsu ağları ve havzalarına ait jeomorfometrik parametreler [60], Çizelge

5.2’de verilmiştir. SBP, diğer havzaya kıyasla, 1:24,000 ölçeğinde toplam akarsu uzunluğu ve 10 metre aralıklı yükseklik eğrilerinin toplam uzunluğu daha fazla, daha büyük akarsu yoğunluğu, rölyef, eğim ve engebe değerine sahip, daha geniş olan havzadır.

Çizelge 5.1 Çalışma bölgesine ait eyalet, havza no ve fizyografik bölüm bilgisi

| Alt havza adı | Eyalet        | Havza No | Fizyografik Bölüm                                 |
|---------------|---------------|----------|---------------------------------------------------|
| PT            | Missouri      | 10290107 | İç dağlık bölgesi Ozark Platosu                   |
| SBP           | Batı Virginia | 02070001 | Appalachian dağlık bölgesi vadi ve sırt bölümleri |

Çizelge 5.2 Akarsu ağları ve havzalarına ait jeomorfometrik parametreler

| Morfometrik Parametreler                                         | PT       | SBP       |
|------------------------------------------------------------------|----------|-----------|
| Alan (km <sup>2</sup> )                                          | 2188.93  | 3832.05   |
| Çevre (km)                                                       | 350.21   | 470.53    |
| Havzanın zirve noktası (m)                                       | 468.62   | 1482.56   |
| 1:24000 ölçeğinde toplam akarsu uzunluğu (km)                    | 3011.25  | 5516.07   |
| Akarsu yoğunluğu                                                 | 1.38     | 1.44      |
| Havzada 10 m aralıklı yükseklik eğrilerinin toplam uzunluğu (km) | 16121.31 | 122574.56 |
| Rölyef (m)                                                       | 253.31   | 1322.37   |
| Engebe değeri                                                    | 0.35     | 1.90      |
| Melton’ın engebe değeri                                          | 0.005    | 0.021     |
| Havzanın eğimi                                                   | 0.07     | 0.32      |

Bu çalışmada önerilen yöntemler, sadece daimi ve mevsimsel derelerin seçme/eleme işleminde kullanılmıştır. Alan nesneler, çizgisel nesne genelleştirmesine göre sonradan genelleştirilebileceği için UHVS’de bulunan alan nesnelerin orta eksenleri girdi verisinde kullanılmamıştır.

## 5.2 Ön İşlemler

Ön işlem, 4 ana aşamadan meydana gelmektedir.

- **Parçaların birleştirilmesi:** Genelleştirmede, bir akarsu kolunun tek bir parça olarak ele alınması gerekir. Başka bir deyişle, akarsu koluna ait parçaların ayrı ayrı elenmesi ağda kopukluğa neden olacaktır [103]. Akarsu kollarının kesişim noktalarını birleştiren akarsu parçalarının kümelenmesi, seçme/eleme işleminde kısıtlamaya yol açmaktadır [104]. Bu nedenle, akarsu kollarına ait parçalar, akarsu düzeyi ve akarsu tipine (mevsimsel ve daimi) göre birleştirilmiştir.
- **Veri zenginleştirme:** Bölüm 5.2.1’de açıklanmıştır.
- **Ağırlıkların belirlenmesi:** Bölüm 5.2.2’de açıklanmıştır.
- **Normalizasyon:** Girdiler (4.8) eşitliğinde açıklanan Z-skor standardizasyonu kullanılarak normalize edilmiştir.

### 5.2.1 Çalışmada Kullanılan Özniteliklerin Belirlenmesi

Ağ bağlantılarını ve hidrografik anlamı karakterize edecek geometrik, topolojik ve semantik öznitelikler seçilmiştir. Bu öznitelikler Çizelge 5.3’de verilmiştir.

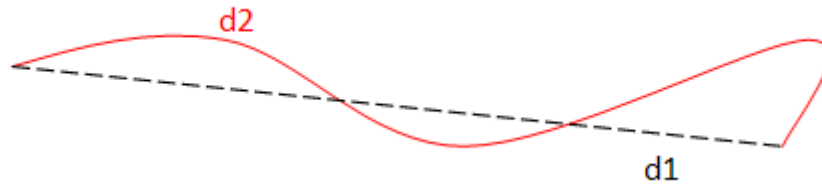
Çizelge 5.3 Kullanılan özniteliklerin tipleri ve değer aralıkları

|           | Öznitelik                          | Tip                       | Aralık                   |
|-----------|------------------------------------|---------------------------|--------------------------|
| Geometrik | Uzunluk (m)                        | Oranlı ölçek (sürekli)    | $0 < x < +\infty$        |
|           | Kıvrımlılık                        | Oranlı ölçek (sürekli)    | $0 < x \leq 1$           |
| Topolojik | Bağlılık Bakımından Merkezlik      | Oranlı ölçek (sürekli)    | $0 < x < +\infty$        |
|           | Arasında Olma Bakımından Merkezlik | Oranlı ölçek (sürekli)    | $0 \leq x < +\infty$     |
|           | Yakınlık Bakımından Merkezlik      | Oranlı ölçek (sürekli)    | $0 \leq x < +\infty$     |
| Semantik  | Akarsu Düzeyi                      | Sıralama ölçeği (kesikli) | $0 < x < +\infty$        |
|           | Akarsu Tipi                        | İkili (kesikli)           | 0: mevsimsel<br>1: daimi |

### 5.2.1.1 Geometrik Öznitelikler

**Uzunluk:** Her bir akarsu nesnesinin uzunluğudur.

**Kıvrımlılık:** Her bir akarsuyun başlangıç ve bitiş noktalarının koordinatları belirlenerek, iki nokta arasındaki uzunluk (Öklid uzunluğu) hesaplanmış ve Öklid uzunluğunun akarsuyun uzunluğuna oranı ile kıvrımlılık değerleri bulunmuştur. Şekil 5.9’da görüldüğü gibi  $d1/d2$  oranı ile kıvrımlılık değeri hesaplanmıştır.



Şekil 5.9 Kıvrımlılık özniteliği

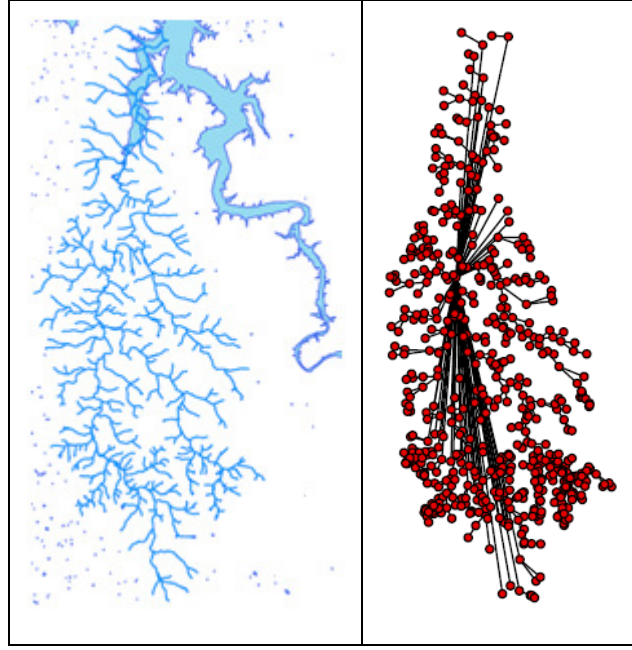
### 5.2.1.2 Topolojik Öznitelikler

Günümüzde, ağ yapısı oluşturan nesnelerin (akarsu ağı, yol ağı, vb.) topolojik olarak modellenmesi çizge (graph) üzerinden sağlanmaktadır. Topolojik model, çizgenin düğüm noktaları ve kenarları ile oluşturulmaktadır [17].

Çizge, düğümler ve bu düğümleri birbirine bağlayan kenarlardan oluşan bir tür ağ yapısıdır. Çizge,  $G(V,E)$  matrisi ile tanımlanırsa, matris  $V=\{V_1, V_2, ..., V_n\}$  düğüm ve  $E=\{V_i, V_j\}$  kenar vektörlerinden oluşmaktadır. Çizgeden elde edilen  $V=\{V_1, V_2, ..., V_n\}$  düğüm ve  $E=\{V_i, V_j\}$  kenar vektörlerinden merkezlik değerleri hesaplanır.

Topolojik özniteliklerin belirlenmesi için öncelikle akarsu ağından oluşturulan çizge oluşturulmuştur. Çizgede her bir nesne (ana akarsular ve kolları) düğüme, her bir kavşak noktası ise kenara dönüşmektedir (Şekil 5.10).

Bağlantı kenar düğüm bilgileri, Fatih Gülgen tarafından MapInfo ortamında MapBasic dilinde yazılan bir betik (script) ile belirlenmiş ve ArcGIS 10’da görselleştirilmiştir.

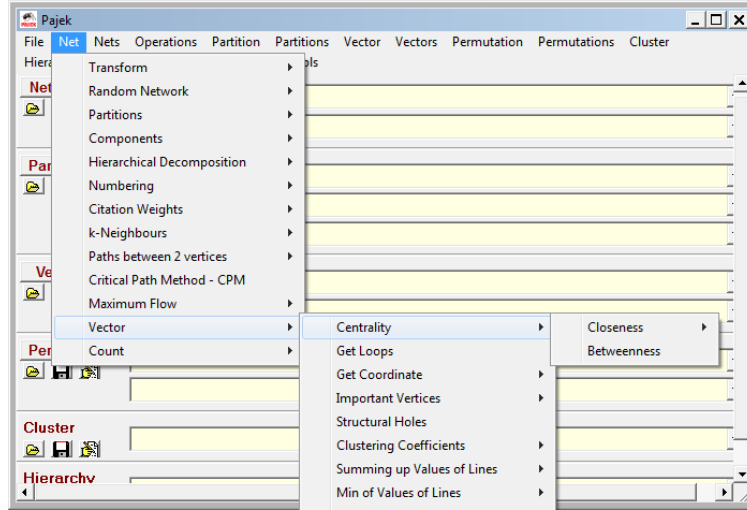


Şekil 5.10 Bir akarsu ağı (solda) ve bağlantı çizgesi (sağda)

Merkezlilik (centrality) kavramı ilk kez 1948 yılında Bavelas tarafından sosyal ağlarda insan iletişimi konulu bir uygulama yapmak amacıyla ortaya atılmıştır. Bavelas, merkezlilik ve grup içi iletişim arasında bir ilişki olduğu hipotezi ile yola çıkmıştır. Merkezlilik kavramı, sosyal ağların önemli bir yapısal özelliğidir. Örneğin, bir yıldızın ortasında yer alan nokta, diğer noktalara göre daha önemli bir merkezlilik değeri taşımaktadır. Bir ağdaki noktaların merkezlilik değerini tespit eden üç önemli ölçüt aşağıda verilmektedir [105].

- Bağlılık bakımından merkezlilik
- Arasında olma bakımından merkezlilik
- Yakınlık bakımından merkezlilik

Merkezlilik değerleri bağlantı çizgesinden Pajek yazılımı ile elde edilmiştir. Pajek, ağ analizi ve görselleştirmesi sağlayan bir yazılımdır [106]. Şekil 5.11’de Pajek yazılımı arayüzü gösterilmiştir.



Şekil 5.11 Pajek yazılımı arayüzü

**Bağlılık bakımından merkezlik (Degree centrality):** Bir akarsuya bağlı akarsuların sayısı veya çizgede bu akarsuya karşılık gelen düğüm ile bağlantılı düğüm sayısıdır. Verilen bir  $p_k$  düğümünün bağlılık bakımından merkezlik değeri  $C_D(p_k)$ , (5.1) eşitliği ile hesaplanmaktadır [17], [52], [105].

$$C_D(p_k) = \sum_{i=1}^n a(p_i, p_k) \quad (5.1)$$

$p_i$  ve  $p_k$  bir kenar ile bağlı ise,  $a(p_i, p_k) = 1$ ; bağlı değilse 0 değerini alır.

Farklı çizgelerde bulunan benzer düğümlerin karşılaştırılabilmesi için (5.1) eşitliğindeki değerin normalize edilmesi gerekmektedir. Eleman sayısı  $n$  olan yıldız şeklinde bir ağ söz konusu olduğunda, merkezde yer alan düğüme  $(n - 1)$  adet düğüm bağlıdır. Düğümün bağlılık bakımından merkezlik değerinin en çok 1 olmasının sağlanması amacıyla yıldız şeklinde ağlar düşünülerek, (5.1) eşitliği düğüm sayısının bir eksiğine  $(n - 1)$  bölünerek tüm değerler 0-1 aralığına normalize edilmektedir [105].

$$C_D'(p_k) = \frac{\sum_{i=1}^n a(p_i, p_k)}{n-1} \quad (5.2)$$

**Yakınlık bakımından merkezlik (Closeness centrality):** Çizge üzerinde iki düğüm arasındaki en kısa güzergahta (shortest linking) bulunan kenar sayısına (number of edges) bağlı olarak değişen bir ölçüttür. Çizgede herhangi iki düğüm noktası olan  $p_i$  ve  $p_k$  noktaları arasındaki en kısa güzergahta bulunan kenar sayısı  $d(p_i, p_k)$  ile gösterildiğinde,  $p_k$  noktasının yakınlık bakımından merkezlik değeri  $C_C(p_k)$ , (5.3) eşitliği ile hesaplanmaktadır [17], [52], [105].

$$C_C(p_k) = \frac{1}{\sum_{i=1}^n d(p_i, p_k)} \quad (5.3)$$

Yakınlık bakımından merkezlik ölçütü (5.4) eşitliği ile normalize edilir [105].

$$C'_C(p_k) = \frac{n-1}{\sum_{i=1}^n d(p_i, p_k)} \quad (5.4)$$

**Arasında olma bakımından merkezlik (Betweenness Centrality):** Bir akarsudan diğer akarsulara olan en kısa güzergahlarda, arada uğranılması zorunlu olan akarsu için arasında olma bakımından merkezlik ölçütü kullanılabilir. Çizgede herhangi iki düğüm noktası olan  $p_i$  ve  $p_j$  noktaları arasındaki en kısa güzergahların sayısı ve bu düğümler arasında  $p_k$  noktasından geçen en kısa güzergahların sayısı hesaplanarak bulunan  $p_k$  noktasının arasında olma bakımından merkezlik değeri  $C_B(p_k)$ , (5.5) eşitliğinde verilmiştir [17], [52], [105].

$$C_B(p_k) = \sum_i^n \sum_j^n \frac{g_{ij}(p_k)}{g_{ij}} \quad (5.5)$$

$g_{ij}$ :  $i$  ve  $j$  düğümleri arasındaki en kısa güzergah sayısı

$g_{ij}(p_k)$ : Bir  $p_k$  noktasından geçilmesi zorunlu olması durumunda,  $p_i$  ve  $p_j$  düğümleri arasındaki en kısa güzergah sayısı

Arasında olma bakımından merkezlik ölçütü (5.6) eşitliği ile normalize edilir [105].

$$C'_B(p_k) = \frac{2C_B(p_k)}{n^2 - 3n + 2} \quad (5.6)$$

### 5.2.1.3 Semantik Öznitelikler

**Akarsu Tipi:** UHVS'de akarsu nesne tipleri, mevsimsel ve daimi olarak belirlenmiştir. Buna göre, ikili (binary) sayı sistemine uygun olarak, mevsimsel ise sıfır, daimi ise bir değeri atanmıştır.

**Akarsu düzeyi:** UHVS'deki akarsu düzeyleri Bölüm 3.3'de açıklanmıştır.

### 5.2.1.4 Özniteliklerin Normal Dağılıma Uygunluğunun Araştırılması

Verinin normal dağılıma sahip olup olmadığı, genellikle birçok istatistiksel analizin bir arada değerlendirilmesi ile ortaya çıkmaktadır. Çizelge 5.4 ve 5.5'deki sürekli

özniteliklere ait aritmetik ortalama, mod ve medyan birbirine eşit değildir. Fisher'ın basıklık ve çarpıklık değeri,  $\pm 2$  aralığının (0.05 anlamlılık düzeyinde) dışında olduğundan, öznitelikler normal dağılımlı değildir. Şekil 5.12, 5.13, 5.14 ve 5.15'deki grafiklerden de verilerin normal dağılımlı olmadığı görülmektedir.

Çizelge 5.4 PT'ye ait istatistikler

|                         | Uzunluk (km) | Kıvrımlılık | Arasında | Yakınlık | Bağılılık |
|-------------------------|--------------|-------------|----------|----------|-----------|
| Sayı                    | 2340         | 2340        | 2340     | 2340     | 2340      |
| Ortalama                | 1.287        | 0.897       | 0.003    | 0.022    | 2.105     |
| Medyan                  | 0.943        | 0.912       | 0.000    | 0.024    | 2.000     |
| Mod                     | 0.007        | 1.000       | 0.000    | 0.027    | 1.000     |
| Std. sapma (s)          | 1.511        | 0.077       | 0.018    | 0.005    | 2.313     |
| Varyans                 | 2.282        | 0.006       | 0.000    | 0.000    | 5.352     |
| Çarpıklık katsayısı     | 4.831        | -1.610      | 8.511    | -0.693   | 6.439     |
| Çarpıklığın std. hatası | 0.051        | 0.051       | 0.051    | 0.051    | 0.051     |
| Çarpıklık değeri        | 94.725       | -31.569     | 166.882  | -13.588  | 126.255   |
| Basıklık katsayısı      | 38.331       | 4.138       | 80.257   | -0.937   | 62.261    |
| Basıklığın std. hatası  | 0.101        | 0.101       | 0.101    | 0.101    | 0.101     |
| Basıklık değeri         | 379.515      | 40.970      | 794.624  | -9.277   | 616.445   |
| Minimum                 | 0.007        | 0.457       | 0.000    | 0.011    | 1.000     |
| Maksimum                | 18.856       | 1.000       | 0.272    | 0.030    | 37.000    |

Çizelge 5.5 SBP'ye ait istatistikler

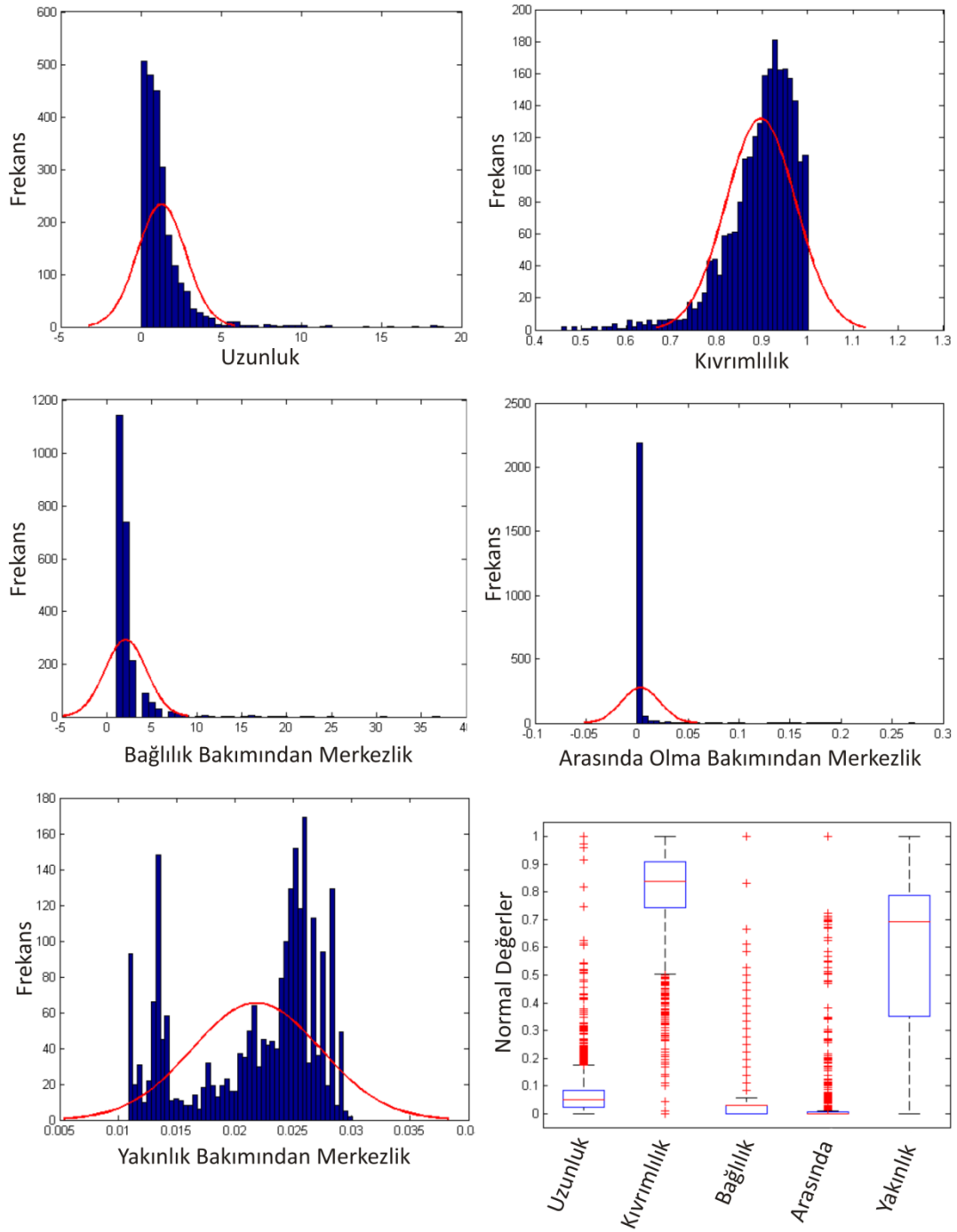
|                         | Uzunluk (km) | Kıvrımlılık | Arasında | Yakınlık | Bağılılık |
|-------------------------|--------------|-------------|----------|----------|-----------|
| Sayı                    | 3951         | 3951        | 3951     | 3951     | 3951      |
| Ortalama                | 1.396        | 0.908       | 0.001    | 0.084    | 2.144     |
| Medyan                  | 1.106        | 0.927       | 0.000    | 0.085    | 2.000     |
| Mod                     | 0.005        | 1.000       | 0.000    | 0.111    | 1.000     |
| Std. sapma (s)          | 1.531        | 0.074       | 0.010    | 0.017    | 2.364     |
| Varyans                 | 2.343        | 0.006       | 0.000    | 0.000    | 5.587     |
| Çarpıklık katsayısı     | 7.673        | -1.846      | 31.932   | -1.035   | 9.762     |
| Çarpıklığın std. hatası | 0.039        | 0.039       | 0.039    | 0.039    | 0.039     |
| Çarpıklık değeri        | 196.744      | -47.333     | 818.769  | -26.538  | 250.308   |
| Basıklık katsayısı      | 101.984      | 5.683       | 1355.216 | 1.023    | 161.231   |
| Basıklığın std. hatası  | 0.078        | 0.078       | 0.078    | 0.078    | 0.078     |
| Basıklık değeri         | 1307.487     | 72.659      | 17374.56 | 13.115   | 2067.064  |
| Minimum                 | 0.005        | 0.342       | 0.000    | 0.027    | 1.000     |
| Maksimum                | 30.905       | 1.000       | 0.497    | 0.142    | 57.000    |

Öznitelikler için test büyüklüğü ( $D$ ), Smirnov'un kritik tablo değerleri ( $d_\alpha$ ) ve  $\alpha = 0.05$  anlamlılık düzeyinde K-S uygunluk tek örneklem testi Çizelge 5.6'da verilmiştir. Sıfır hipotezi (gözlemler normal dağılımlıdır) reddedildiğinden ( $d_\alpha < D$ ), öznitelikler normal dağılımlı değildir.

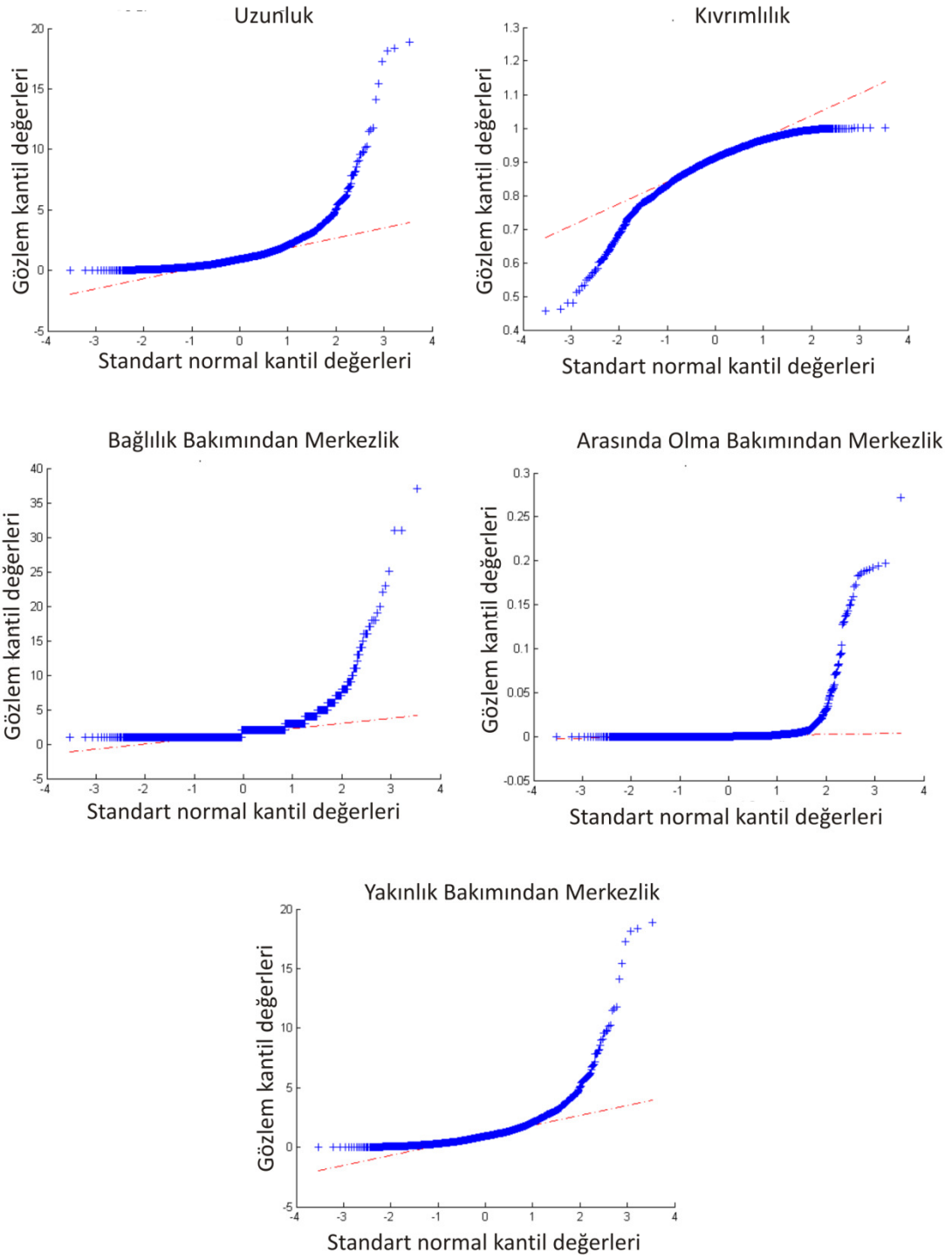
Çizelge 5.6 Kolmogorov-Smirnov uygunluk tek örneklem testi ( $\alpha = 0.05$ )

|     |            | Sürekli Öznitelikler |             |           |          |          |
|-----|------------|----------------------|-------------|-----------|----------|----------|
|     |            | Uzunluk              | Kıvrımlılık | Bağılılık | Arasında | Yakınlık |
| PT  | $d_{0.05}$ | 0.0280               | 0.0280      | 0.0280    | 0.0280   | 0.0280   |
|     | D          | 0.1984               | 0.0951      | 0.3215    | 0.4252   | 0.1733   |
|     | $H_0$      | Ret                  | Ret         | Ret       | Ret      | Ret      |
| SPB | $d_{0.05}$ | 0.0216               | 0.0216      | 0.0216    | 0.0216   | 0.0216   |
|     | D          | 0.1863               | 0.1100      | 0.3143    | 0.4543   | 0.1401   |
|     | $H_0$      | Ret                  | Ret         | Ret       | Ret      | Ret      |

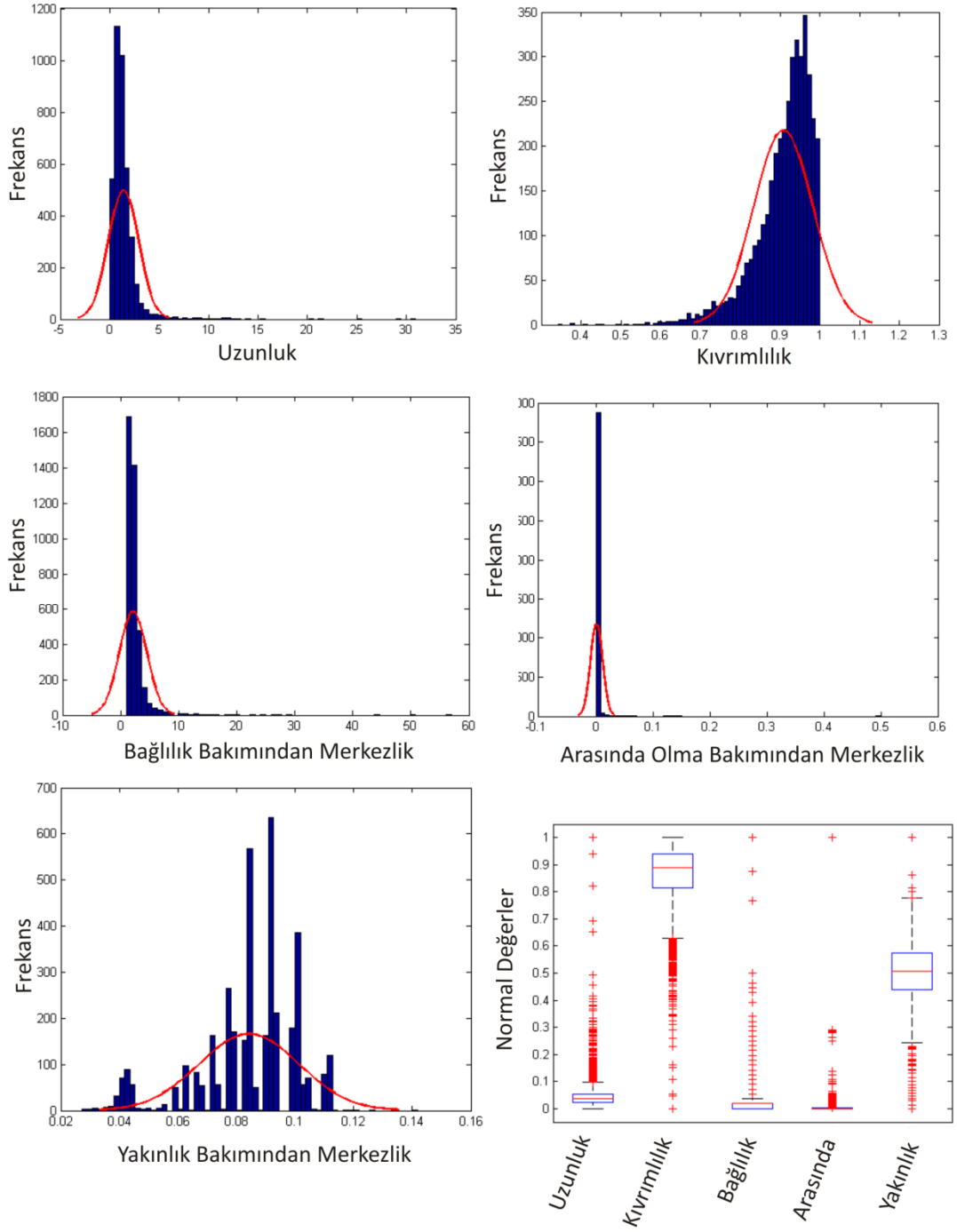
Bu sonuçlara göre, veriler normal dağılımlı olmadığından ve veri setinde kesikli veriler olduğundan öznitelik ağırlıklarının belirlenmesinde parametrik olmayan hipotez testlerinin kullanımı uygun olacaktır.



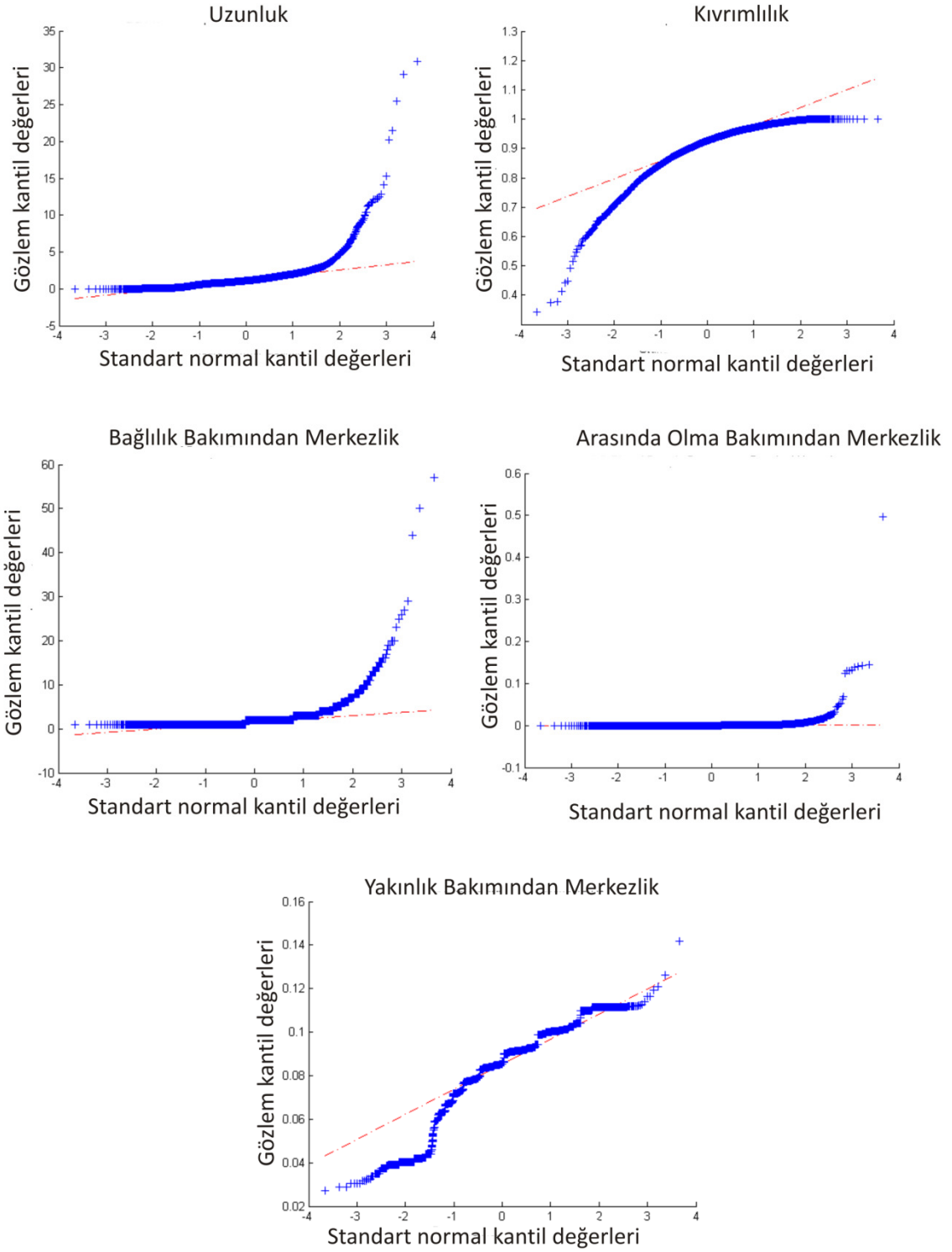
Şekil 5.12 PT'ye dair histogram ve kutu grafikleri



Şekil 5.13 PT'ye dair kantil-kantil grafikleri



Şekil 5.14 SBP'ye dair histogram ve kutu grafikleri



Şekil 5.15 SBP'ye dair kantil-kantil grafikleri

### 5.2.2 Ağırlıkların Belirlenmesi

Her bir öz nitelik normalize edildikten sonra, coğrafi bilginin önem derecesine göre ağırlıklandırılabilir [107]. [52] tarafından, KDH yöntemi kullanılarak yapılan yol ağı genelleştirme çalışmasında, öz nitelikler deneysel (empirically) olarak

ağırlıklandırılmıştır. Bunun dışında, değişkenlerin ağırlıklandırıldığı çeşitli KDH ve SVM uygulamaları da vardır [90], [107].

Bu çalışmada, her özniteliğin seçme/eleme işlemindeki ağırlığının (etkinliğinin) farklı olması gerektiği düşüncesiyle, ağırlıklandırma yöntemleri araştırılmıştır. Kullanılan verinin normal dağılımlı olmaması nedeniyle, parametrik olmayan  $\chi^2$  bağımsızlık testinin uygun olacağı tespit edilmiştir. Böylece, her bir öznitelik ile orijinal 1:100,000 ölçekli akarsu ağlarından elde edilen seçme/eleme (0: seçim; 1: eleme) bilgisi arasında ilişki olup olmadığı belirlenmiştir.  $\chi^2$  bağımsızlık testinde, Bölüm 4.1.6'da açıklandığı gibi, eğer ilişki varsa  $H_0$  hipotezi reddedilmektedir.

$H_0$ : Öznitelik ile orijinal 1:100,000 arasında ilişki yoktur.

$H_1$ : Öznitelik ile orijinal 1:100,000 arasında ilişki vardır.

$\chi^2$  bağımsızlık testi kategorik (kesikli) veri için kullanılmaktadır. Bu nedenle, sürekli değişkenler kategorik değişkenlere dönüştürülmelidir [85]. Bu bağlamda, uzunluk özniteliği, 1:100,000 ölçekli akarsu ağları için USGS UHVS standartlarına [69] uygun olarak iki kategoriye ( $<1.6$  km ve  $\geq 1.6$  km) ayrılmıştır. Diğer sürekli öznitelikler (kıvrımlılık, bağlılık, yakında olma ve arasında olma) ise standart sapmalarına göre kategorilere ayrılmıştır.

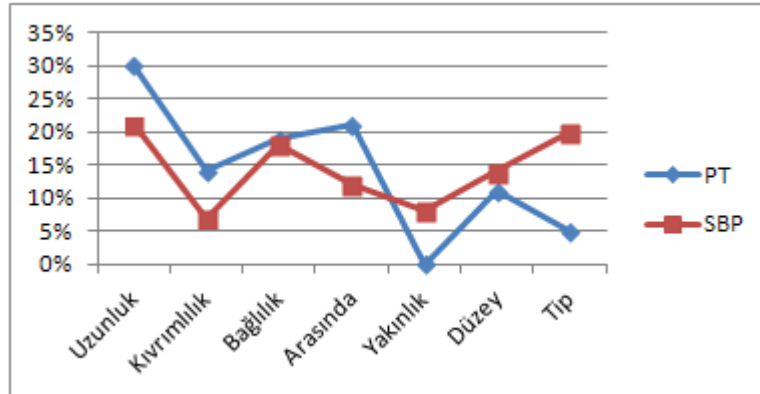
0.05, 0.01 and 0.001 anlamlılık düzeylerinde yapılan test sonucunda, PT akarsu ağında yakınlık bakımından merkezlik değerinin, seçimle ilişkisiz olduğu belirlenmiştir. Bu nedenle, bu öznitelik PT akarsu ağının kümeleme ve sınıflandırmasında kullanılmamıştır.

$\chi^2$  test büyüklükleri, *Cramer'in V* ilişki katsayıları ( $\varphi_c$ ) ile toplam ağırlık bir olacak şekilde normalize edilmiştir: Her bir öznitelik için hesaplanan *Cramer'in V* katsayısı, tüm özniteliklerin katsayı değerlerinin toplamına bölünerek ağırlıklar bulunmuştur.

PT ve SBP için serbetlik dereceleri ( $k$ ), kritik tablo değerleri ( $\chi^2_{0,01}$ ),  $\chi^2$  test büyüklükleri ( $\chi^2_{hes}$ ),  $H_0$  hipotezi, *Cramer'in V* ilişki katsayıları ( $\varphi_c$ ) ve ağırlıklar ( $w$ ) Çizelge 5.7'de verilmiştir. Ağırlıklara dair grafik ise Şekil 5.16'da gösterilmiştir. Özniteliklerin jenerik ağırlıklarının belirlenebilmesi için, ülke genelinde birçok karakteristik akarsu ağının istatistiksel olarak değerlendirilmesi gerektiği unutulmamalıdır. Bu çalışmada hesaplanan ağırlıklar, sadece çalışma alanı için geçerlidir.

Çizelge 5.7 Özniteliklerin hesaplanan ağırlıkları

|     |                 | Uzunluk | Kıvrımlılık | Arasında | Yakınlık | Bağlılık | Akarsu | Tip    |
|-----|-----------------|---------|-------------|----------|----------|----------|--------|--------|
| PT  | $k$             | 1       | 4           | 2        | 3        | 2        | 5      | 1      |
|     | $\chi^2_{0.01}$ | 6.64    | 13.28       | 9.21     | 11.34    | 9.21     | 15.09  | 6.64   |
|     | $\chi^2_{hes}$  | 1134.86 | 226.31      | 543.98   | 8.4      | 421.72   | 163.54 | 27.35  |
|     | $H_0$           | Ret     | Ret         | Ret      | Kabul    | Ret      | Ret    | Ret    |
|     | $\varphi_c$     | 0.696   | 0.311       | 0.482    | -        | 0.425    | 0.264  | 0.108  |
|     | $w$             | %30     | %14         | %21      | -        | %19      | %11    | %5     |
| SBP | $k$             | 1       | 4           | 2        | 5        | 3        | 4      | 1      |
|     | $\chi^2_{0.01}$ | 6.64    | 13.28       | 9.21     | 15.09    | 11.34    | 13.28  | 6.64   |
|     | $\chi^2_{hes}$  | 694.25  | 81.77       | 204.37   | 96.62    | 459.57   | 314.11 | 614.46 |
|     | $H_0$           | Ret     | Ret         | Ret      | Ret      | Ret      | Ret    | Ret    |
|     | $\varphi_c$     | 0.419   | 0.144       | 0.227    | 0.156    | 0.341    | 0.282  | 0.394  |
|     | $w$             | %21     | %7          | %12      | %8       | %18      | %14    | %20    |



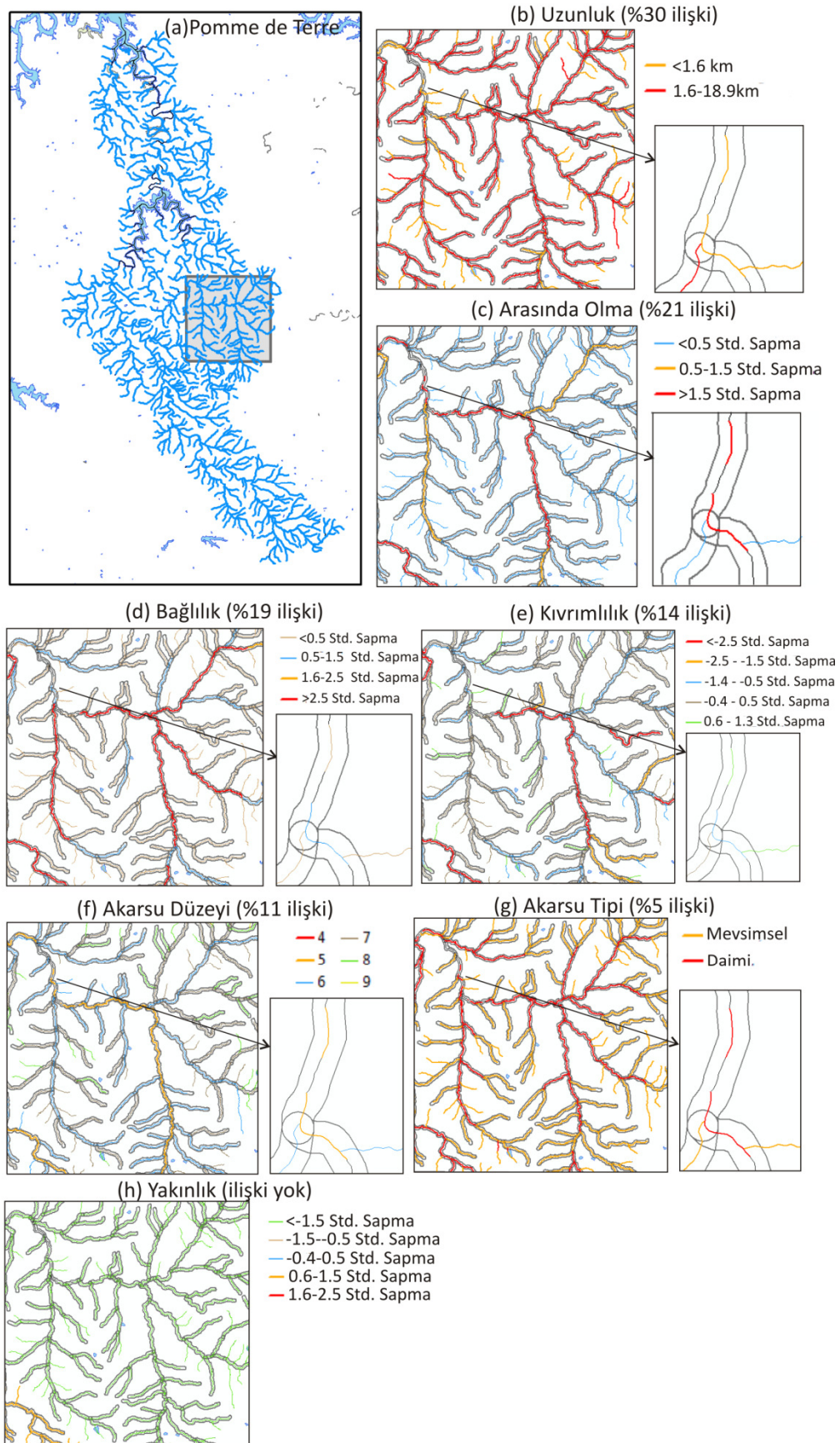
Şekil 5.16 Ağırlıklara dair grafik

Şekil 5.17a ve 5.18a'da, orijinal 1:100,000 ölçekli PT ve SBP ağları gösterilmiştir. Seçme/eleme sonuçlarının daha iyi analiz edilebilmesi için bu ağların belli kısımları çok daha büyük ölçekte ve orijinal 1:24,000 ölçekli akarsu ağları ile üst üste bindirilerek verilmiştir (Şekil 5.17b-h ve 5.18b-h). 1:100,000 ölçekli akarsu ağlarının, 1:24,000

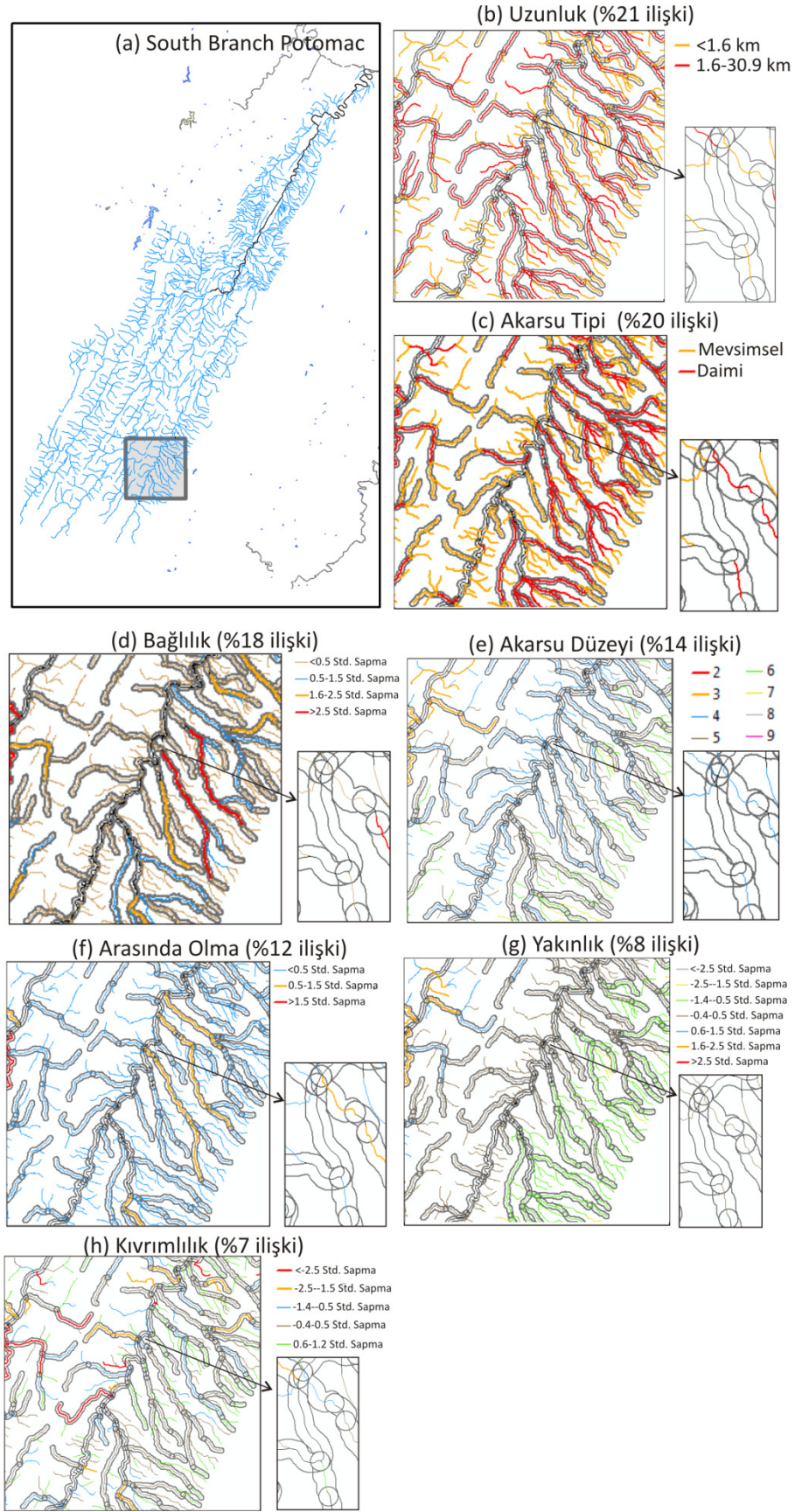
ölçekli akarsu ağlarından ayırt edilebilmesi için, 1:100,000 ölçekli akarsu ağları tampon bölgeler ile gösterilmiştir. İşaret tablolarında her bir özniteliğe dair kategori bilgileri verilmiştir. Şekillerdeki siyah renkli nesneler, orta eksenlerdir. Orta eksenler kategorik nesneler olmadığı için işaret tablolarında gösterilmemiştir.

Orijinal 1:100,000 ölçekli ağlar, öznitelikler bakımından incelendiğinde aşağıdaki sonuçlar tespit edilmiştir.

- 1.6 km'den uzun kolların çoğu seçilmiştir. Uzunluk, en etkili özniteliktir (Şekil 5.17b ve 18b).
- Az kıvrımlı olanlara göre çok kıvrımlı olanlar tercih edilmiştir (Şekil 5.17e ve 5.18h).
- Az bağlantılı olanlara göre çok bağlantılı olanlar tercih edilmiştir (Şekil 5.17d ve 5.18d).
- Ağda sürekliliği sağlayacak olanlar seçilmiştir (Şekil 5.17c ve 5.18f).
- Kolların ana akarsuya mesafeleri genellikle göz önüne alınmamıştır (Şekil 5.17h ve 5.18g).
- Akarsu tipi, desene bağlı olarak göz önüne alınmıştır (Şekil 5.17g ve 5.18c).



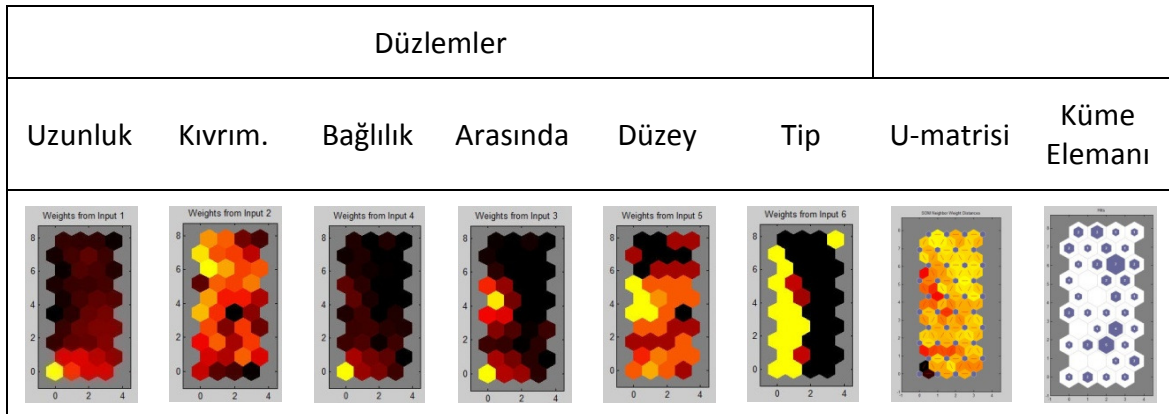
Şekil 5.17 Orijinal 1:100,000 ölçekli PT akarsu ağında öznitelik ilişkileri



Şekil 5.18 Orijinal 1:100,000 ölçekli SPB akarsu ağında öznitelik ilişkileri

### 5.3 Kümelerin Seçimi için Geliştirilen Yöntem

Kendini Düzenleyen Haritalar için Matlab R2009b Sinir Ağları aracı kullanılmıştır. Her alt havzadaki akarsu yoğunluğunu ve desenini korumak için, PT ve SBP havzalarında bulunan toplam 60 alt havzaya ait akarsu ağları ayrı ayrı kümelendirilmiştir. KDH büyüklüğü ve şekli hariç, kullanılan parametreler tüm alt havzalarda aynıdır. Eğitim parametreleri (komşuluk fonksiyonu, kernel genişliği, öğrenme katsayısı, topoloji ve uzaklık kriteri) literatürde yaygın olarak kullanıldığı gibidir [54], [90], [91], [92], [93], [94]. Yığılma eğitimi; Gauss komşuluk fonksiyonu, altıgen topoloji ve Öklid mesafesi kullanılmıştır. Öğrenme katsayısı olarak 0.1 değeri alınmıştır. 100,000 eğitim döngüsünün yeterli olduğu görülmüştür. KDH büyüklüğü (nöron sayısı), [91]’e göre (4.24) eşitliği ile hesaplanmıştır. KDH’nin şekli ise “Bire Yarım Kuralı”na göre belirlenmiştir [92], [93]. Şekil 5.19’da, PT’nin alt havzalarından biri olan Brush isimli akarsu ağına ait her bir özneliğe göre kümeleme (bileşen düzlemler), tüm özneliklere göre kümeleme (U-matrisi) ve küme eleman sayıları gösterilmiştir. Buna göre, uzunluk ve bağlılık bakımından merkezlik değerleri birbirine benzer biçimde kümelendirilmiştir. Bunun nedeni, uzun akarsuya bağlanan kol sayısı, kısa akarsuya bağlanan kol sayısından daha fazla olduğundan, uzun akarsuyun bağlılık bakımından merkezlik değeri daha büyüktür. Bu sayede, her iki öznelik (uzunluk ve bağlılık) benzer biçimde kümelendirilmiştir. Şekil 5.19’daki hem bağlılık hem de uzunluk düzlemlerinin sol alt köşelerinde bulunan kümeler (sarı), en uzun ana akarsuları içermektedir. Tüm alt havzalarda (60 adet) yapılan KDH kümeleme sonuçlarına ait U-Matrisleri EK-A’da verilmiştir.



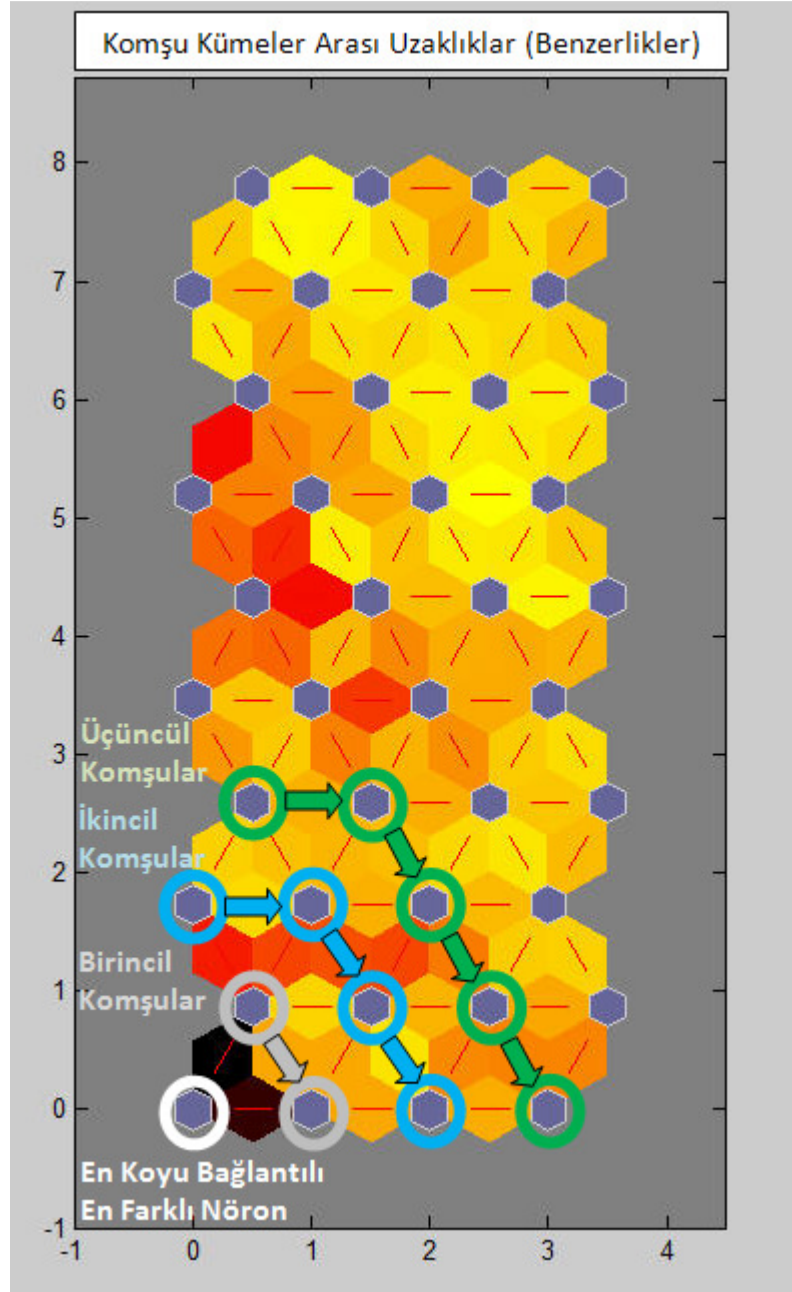
Şekil 5.19 Brush alt havzasında KDH eğitimi sonunda oluşan düzlemler, U-matrisi ve küme eleman sayıları

Bu çalışmada, hedef 1:100,000 ölçekli akarsu ağı için seçilecek nesnelerin belirlenmesi amacıyla, U-matrisinde komşuluk ilişkilerine göre yerleşmiş olan kümelerin (nöron) kademeli olarak seçildiği bir yöntem geliştirilmiştir.

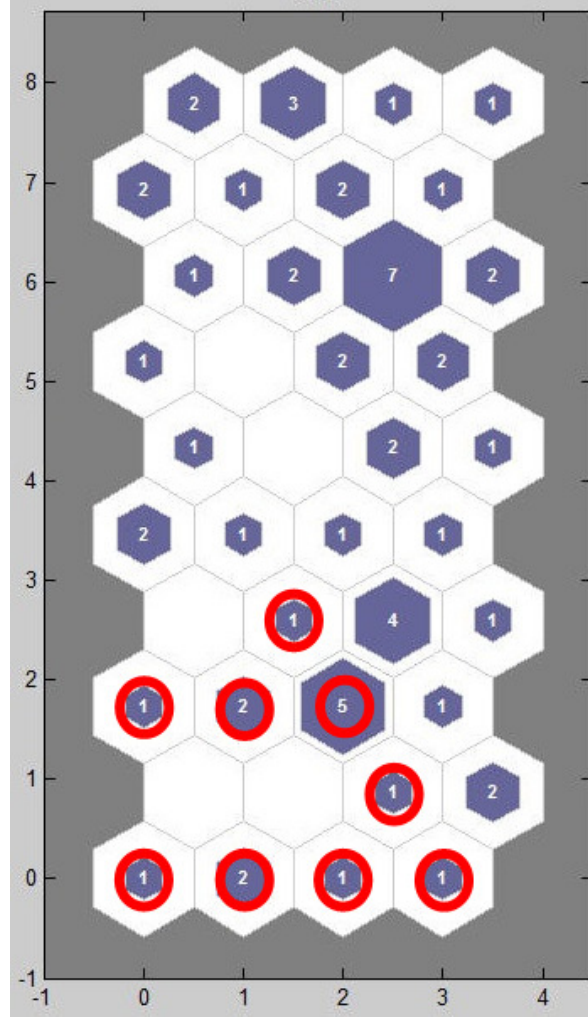
Öncelikle, Matlab R2009b Sinir Ağları tarafından temel renklerle (RGB: Red Green Blue) toplamalı senteze göre renklendirilmiş olan U-matrisindeki bağlantıların gri değerleri  $(0.2989 \times R + 0.5870 \times G + 0.1140 \times B)$  bağıntısı ile hesaplanmıştır. Şekil 5.20’de görünen en koyu gri değerine sahip bağlantı, diğerlerinden en uzak iki kümeyi bağlamaktadır. Bu iki kümeden de U-matrisinin kenar(lar)ına en yakın olanı, diğerlerinden en farklı olanıdır. Bu şekilde ilk küme seçilmiştir. Bu küme, en uzun olan ana akarsu gibi en farklı nesneleri içermektedir. Böylesi kümeler genellikle U-matrisin köşesinde yer almaktadır.

Sonraki adımda, bu kümenin birincil komşuları ele alınmıştır. Birincil komşular U-matrisinin kenarına en yakın olandan başlanarak saat ibresi yönünde sırayla seçim setine eklenmiştir. Her ekleme işleminden sonra seçim setindeki toplam eleman sayısı hesaplanmış ve Radikal Kanun’a göre belirlenen sayı ile karşılaştırılmıştır. Radikal Kanun’a göre belirlenen sayı aşıldıkça ekleme işlemine devam edilmiştir. Töpfer’in Radikal Kanunu’nun çizgi işaretlere uygulanan (1.3) eşitliği, seçilecek kümelerin, dolayısıyla nesnelerin sayısını belirlemede eşik değeri olarak kullanılmıştır. Aynı işlem, saat ibresinin tersi yönünde tekrarlanmıştır. Böylece iki seçim seti ortaya çıkmıştır. Bunlardan hangisinin eleman sayısı, Radikal Kanun’a göre belirlenen sayıya daha yakın ise o seçim seti, sonuç seçim seti olarak kabul edilmiştir. Şekil 5.21’de seçilen kümelerin (kırmızı daireler) eleman sayıları gösterilmiştir.

Birincil komşulardaki toplam eleman sayılarının, Radikal Kanun’a göre belirlenen sayıdan küçük olması durumunda, aynı yaklaşımla ikincil, üçüncül vd. komşular ele alınmıştır.



Şekil 5.20 U-matrisinde kümelerin seçimi için geliştirilen yöntem



Şekil 5.21 U-matrisinde seçilen kümelerin eleman sayıları

#### 5.4 Destek Vektör Makineleri ile Uygulama

DVM için Matlab R2009b DVM fonksiyonları kullanılarak, k-kat çapraz doğrulama işlemini de içeren bir kod yazılmıştır. DVM yöntemi, ceza ( $C$ ) ve kernel ( $\gamma$ ) parametrelerine bağlıdır.  $C$ 'nin kontrolsüz biçimde değişmesi modelin genelleme hatasını olumsuz yönde etkileyeceğinden, k-kat çapraz doğrulamanın kullanımı, modelin hem eğitim hem de test setlerinde eğitilebilmesi açısından uygun olacaktır. DVM uygulamasında en uygun hiper düzlem, en karakteristik olduğu düşünülen ve ortalama nesne sayısına sahip bir alt havzanın sınıf etiketleri (eğitim seti) kullanılarak, 10-kat çapraz doğruluk değerlerine göre deneysel olarak belirlenen ceza ( $C$ ) ve kernel ( $\gamma$ ) parametrelerine göre elde edilmiştir (Çizelge 5.8). Bu tezde, deneysel olarak belirlenen  $C$  ve  $\gamma$  parametrelerinin en iyi kombinasyonu, onları üssel olarak değiştiren grid taraması ile sağlanmıştır [108].

Çizelge 5.8 Ceza ve kernel parametrelerine göre 10-kat çapraz doğruluk değerleri

| Alt Havza | $C$        | $\gamma$ | Hatalı seçim yüzdesi | Doğru seçim yüzdesi |
|-----------|------------|----------|----------------------|---------------------|
| PT        | $10^{-10}$ | 1        | 75                   | 0.25                |
|           | 0.02       | 1        | 18                   | 0.82                |
|           | 0.2        | 1        | 17                   | 0.83                |
|           | 1          | 1        | 16                   | 0.84                |
|           | 1.5        | 1        | 16                   | 0.84                |
|           | 1.5        | 1.5      | 16                   | 0.84                |
|           | 2          | 1        | 16                   | 0.84                |
|           | 5          | 1        | 15                   | 0.85                |
|           | 5          | 3        | 15                   | 0.85                |
|           | 7          | 1        | 16                   | 0.84                |
|           |            |          |                      |                     |
| SBP       | 0.002      | 1        | 22                   | 0.78                |
|           | 0.02       | 0.2      | 33                   | 0.67                |
|           | 0.02       | 1        | 23                   | 0.77                |
|           | 0.2        | 1        | 20                   | 0.80                |
|           | 1          | 1        | 26                   | 0.74                |
|           | 1.5        | 1        | 24                   | 0.76                |
|           | 3          | 1        | 27                   | 0.73                |
|           | 3          | 3        | 20                   | 0.80                |
|           | 5          | 10       | 19                   | 0.81                |
|           | 5          | 20       | 21                   | 0.79                |
|           | 16         | 16       | 19                   | 0.81                |
|           | 20         | 20       | 21                   | 0.79                |
|           | 20         | 1        | 18                   | 0.72                |

En uygun hiper düzlem, en karakteristik olduğu düşünülen ve ortalama nesne sayısına sahip bir adet alt havza kullanılarak belirlenmiş, diğer alt havzalar belirlenen en uygun hiper düzleme göre ikili olarak (seçme/eleme) sınıflandırılmıştır. Çizelge 5.8’de, doğruluk değerlerine göre en uygun DVM parametreleri, PT için  $C = 5$  ve  $\gamma = 1$ ; SBP için  $C = 5$  ve  $\gamma = 10$  değerleridir.

### 5.5 Son İşlemler

Göllerin genelleştirilmesi çalışma kapsamı dışında tutulmuştur. Bu nedenle göllerin orta eksenleri girdi verisinde kullanılmamıştır. Orta eksenler, model akışını sağlamak için alan nesnelerden türetilen yapay nesneler olduğundan ayrıca değerlendirilmelidir. Bu çalışmada, orijinal 1:100,000 ölçekli haritalardaki göller, KDH ve DVM yöntemleriyle seçilen birincil akarsu ağına yerleştirilmiştir. Bununla birlikte, eğer göllere bağlanan ya da göllerden çıkan akarsular birincil akarsu ağında bulunmuyorsa, bunlar da ilave edilir. Son olarak, eğer birincil akarsu ağında kopukluklar (bağılantısız akarsular) varsa, bağlantıyı sağlayacak ilgili akarsular da seçilmiştir. Böylece, sonuç akarsu ağı hem alt havza bazında, hem de tüm havzada sürekli hale getirilmiştir.

## BÖLÜM 6

### BULGULAR

Örtüşen çizgi katsayısı (ÖÇK; The coefficient of line correspondence - CLC), yeryüzünde aynı nesneyi temsil eden iki çizgi setinin ne kadar örtüştüğünü hesaplamaktadır [43].

$$\text{ÖÇK} = D_u / (E_u + O_u + D_u) \quad (6.1)$$

Burada  $D_u$  (Denk gösterim), orijinal 1:100,000'lik ve türetilen 1:100,000'lik veri setindeki örtüşen çizgilerin, orijinal 1:100,000'likteki toplam uzunluğudur.  $E_u$  (Eksik gösterim), orijinal 1:100,000'lik veri setinde bulunan, fakat türetilen 1:100,000'lik veri setinde olmayan çizgilerin toplam uzunluğudur.  $O_u$ , türetilen 1:100,000'lik veri setinde olan, fakat orijinal 1:100,000'lik veri setinde olmayan çizgilerin toplam uzunluğunun ( $F_u$ : Fazla gösterim), uzunluk büyüme faktörüne ( $LEF$ : Length Expansion Factor) oranıdır ( $O_u = F_u / LEF$ ). Türetilen 1:100,000'lik veri setindeki çizgiler, orijinal 1:24,000'den geldiği için, orijinal 1:100,000'dekiler gibi basitleştirilmiş değildir. Uzunluk büyüme faktörü, türetilen 1:100,000'lik veri setinde olan, fakat orijinal 1:100,000'lik veri setinde olmayan çizgilerin, bir anlamda, sanal olarak basitleştirilmesi için kullanılmaktadır. Böylece,  $D_u$  ve  $E_u$  gibi  $O_u$  da, 1:100,000'lik veri setindeki basitleştirilmiş çizgilerin toplam uzunluğu olmaktadır [43]. Uzunluk büyüme faktörü, hem orijinal 1:24,000, hem de orijinal 1:100,000 ölçekli veri setinde mevcut olan aynı akarsuların (her iki havzada (PT ve SBP) seçilen homojen dağılımlı 30'ar adet akarsu) uzunluklarının birbirlerine oranlarının aritmetik ortalaması şeklinde hesaplanmıştır. PT ve SBP için sırasıyla 1.03 ve 1.01 değerleri elde edilmiştir. Orijinal ve önerilen yöntemler ile türetilen haritaların karşılaştırılması sonucu hesaplanan ÖÇK değerleri,

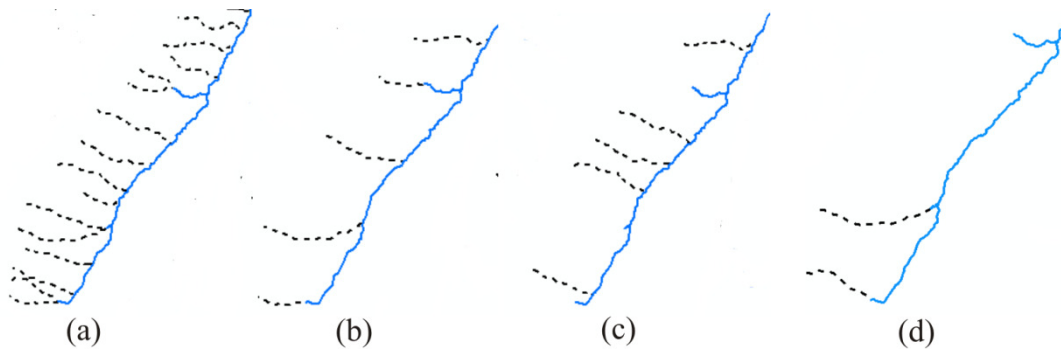
Çizelge 6.1’de gösterilmiştir. ÖÇK değerinin “1” çıkması, orijinal ve türetilen akarsu ağlarının tam olarak örtüştüğü anlamını taşır.

Çizelge 6.1 Önerilen yöntemler ile türetilen akarsu ağlarına ait ÖÇK değerleri

| Yöntem | Havza | $D_u$ (km) | $E_u$ (km) | $O_u$ (km) | ÖÇK  |
|--------|-------|------------|------------|------------|------|
| KDH    | PT    | 1550.77    | 365.45     | 146.00     | 0.75 |
|        | SBP   | 2360.28    | 633.74     | 410.25     | 0.69 |
| DVM    | PT    | 1536.46    | 379.75     | 203.60     | 0.72 |
|        | SBP   | 2460.86    | 533.15     | 583.58     | 0.69 |

ÖÇK değerlerine göre, her iki yöntem, ağaçsı desene sahip PT’de -kafes ve dikdörtgen desene sahip SBP’ye göre- daha iyi sonuç vermiştir. KDH yöntemi ile PT’den türetilen akarsu ağı, orijinal 1:100,000’likteki akarsu ağı ile daha iyi örtüşmektedir. SBP’de her iki yöntemle türetilen akarsu ağları aynı oranda orijinalleri ile örtüşmektedir.

Kafes ve dikdörtgen desenli ağlarda -ağaçsı desene sahip ağlara göre- daha kötü sonuçların alınmasının nedeni, Şekil 6.1’de görüldüğü gibi, birbirine benzer kolların bulunmasıdır. Bu kolların merkezlik değerleri (bağlılık, arasında olma, yakınlık) ve akarsu düzeyleri aynıdır. Bu durumda, kümeleme ve sınıflandırma sadece akarsu tipi, uzunluk ve kıvrımlılık değerlerine bağlı olmaktadır. Bu nedenle, seçilen akarsular orijinal akarsular ile daha az örtüşmektedir.



Şekil 6.1 a) Orijinal 1:24,000, b) orijinal 1:100,000, c) KDH 1:100,000 ve d) DVM 1:100,000 ölçeğinde kafes desenli benzer kollar (Siyah kesikli çizgi: Mevsimsel dereler, mavi çizgi: Daimi dereler)

DVM uygulamasında, PT'de Crane alt havzası ve SBP'de 5 nolu alt havza (Şekil 5.3 ve 5.4), eğitim seti olarak kullanıldığından orijinale çok yakındır. Eğitimde kullanılan eğitim setine (desen ve öznitelik değerleri açısından) benzeyen ağlarda orijinal ağlara yakın değerler elde edilmiştir.

KDH yönteminin uygulandığı önerilen yaklaşımda, Töpfer'in Radikal Kanunu'nun çizgi işaretlere uygulanan (1.3) eşitliği, seçilecek kümelerin, dolayısıyla nesnelerin sayısını belirlemede eşik değer olarak kullanılmıştır. Çizelge 6.2'de, 1:24,000 ve 1:100,000 ölçeğinde orijinal ve türetilen akarsu ağlarına ait çeşitli değerler ve Töpfer'in Radikal Kanunu'ndan elde edilen değerler (çizgi işaretler için) verilmektedir. Parantez içinde, seçilen nesnelerin yüzdeleri verilmektedir.

Kartograf tarafından, orijinal 1:100,000 ölçeğinde birçok göl elenmiş ve alan nesneler çizgi nesnelere dönüştürülmüş olduğundan, Töpfer'in Radikal Kanunu'na göre seçilecek nesne sayısı, orijinal 1:100,000 ölçeğindeki nesne sayısı ile karşılaştırılmamalıdır.

Töpfer sayısına, KDH yöntemi ile türetilen nesne sayısı, DVM yöntemi ile türetilen nesne sayısından daha yakındır. Buna karşın, DVM ile türetilen akarsu ağının yoğunluğu, orijinal akarsu ağının yoğunluğuna daha yakındır. Bunun nedeni, KDH yaklaşımında, küme sayısının Töpfer'in Radikal Kanunu'na göre belirlenmesi; buna karşın DVM yönteminde bir eşik değer kullanılmaması, havzadan seçilen eğitim verisi kullanılarak eğitilmesidir.

Türetilen akarsu ağlarında minimum akarsu uzunluğu, ekranda görüntülenebilecek minimum uzunluk olan 30m'den daha büyüktür [109].

Çizelge 6.2 1:24,000 ve 1:100,000 ölçeğinde orijinal ve 1:100,000 ölçeğinde türetilen PT ve SBP akarsu ağlarına ait nicel bilgiler ve Töpfer'in Radikal Kanununa göre hesaplanan eşik değerler

| Yöntem   | Havza | Ölçek              | Nesne Sayısı | $\Sigma$ Uzunluk (km) | Min. Uzunluk (km) | Nesne Sayısı<br>(Uzunluk $\geq 1.6$ km) | Yoğunluk<br>(km/km <sup>2</sup> ) |
|----------|-------|--------------------|--------------|-----------------------|-------------------|-----------------------------------------|-----------------------------------|
| Orijinal | PT    | Orijinal 1:24,000  | 2340 (%100)  | 3011.25 (%100)        | 0.007             | 572                                     | 1.38                              |
|          |       | Orijinal 1:100,000 | 619 (%26)    | 1916.22 (%64)         | 0.08              | 507                                     | 0.87                              |
|          |       | Töpfer 1:100,000   | 562 (%24)    | -                     | -                 | -                                       | -                                 |
|          | SBP   | Orijinal 1:24,000  | 3952 (%100)  | 5516.07 (%100)        | 0.005             | 1093                                    | 1.44                              |
|          |       | Orijinal 1:100,000 | 1067 (%27)   | 2994.01( %54)         | 0.05              | 682                                     | 0.78                              |
|          |       | Töpfer 1:100,000   | 948 (%24)    | -                     | -                 | -                                       | -                                 |
| KDH      | PT    | Önerilen 1:100,000 | 632 (%27)    | 1547.71 (%51)         | 0.20              | 378                                     | 0.71                              |
|          | SBP   | Önerilen 1:100,000 | 1014 (%26)   | 2416.07 (%44)         | 0.08              | 607                                     | 0.63                              |
| DVM      | PT    | Önerilen 1:100,000 | 722 (%31)    | 1608.92 (%53)         | 0.05              | 398                                     | 0.73                              |
|          | SBP   | Önerilen 1:100,000 | 1220 (%31)   | 2723.84 (%49)         | 0.10              | 702                                     | 0.71                              |

Türetilen akarsu ağlarını, semantik açıdan nesne tipi ve akarsu düzeyi bazında incelemek faydalı olacaktır. Çizelge 6.3'te, orijinal ağlarda daimi ve mevsimsel dere sayıları verilmektedir. PT için orijinal 1:24,000 ve 1:100,000 ölçekli haritalarda, daimi ve mevsimsel dere oranlarının yaklaşık aynı olduğu, fakat SBP için 1:100,000 ölçeğinde daimi derelerin mevsimsel derelere oranla arttığı görülmektedir. Başka bir deyişle, PT'de daimi ve mevsimsel dereler aynı oranda elenmişken; SBP'de mevsimsel dereler, daimi derelere göre daha fazla elenmiştir. Bunun nedeni, SBP'de mevsimsel derelerin hiyerarşik sıralamada altlarda yer almasıdır. Başka bir deyişle, mevsimsel derelerin gövdeye (ana akarsuya) uzak olmalarıdır.

Çizelge 6.3. Orijinal 1:24,000 ve 1:100,000 ölçeğinde PT ve SBP akarsu ağlarına ait akarsu tipi ile ilgili nicel bilgiler

|          |       | Akarsu Tipi       |             |              |                    |            |              |
|----------|-------|-------------------|-------------|--------------|--------------------|------------|--------------|
|          |       | Orijinal 1:24,000 |             |              | Orijinal 1:100,000 |            |              |
| Yöntem   | Havza | Daimi             | Mevsimsel   | Toplam       | Daimi              | Mevsimsel  | Toplam       |
| Orijinal | PT    | 579<br>%25        | 1761<br>%75 | 2340<br>%100 | 112<br>%22         | 507<br>%78 | 619<br>%100  |
|          | SBP   | 901<br>%23        | 3051<br>%77 | 3952<br>%100 | 509<br>%48         | 558<br>%52 | 1067<br>%100 |

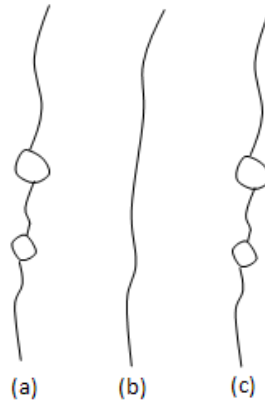
Çizelge 6.4'te, KDH yöntemine göre türetilen, Çizelge 6.5'te ise DVM yöntemine göre türetilen daimi ve mevsimsel dere sayıları verilmektedir. 1:100,000 ölçekli orijinal ve türetilen ağlardaki denk nesne sayısı  $D_s$ , orijinal ağda olup, türetilen ağda olmayan (eksik gösterilen) nesne sayısı  $E_s$  ve türetilen ağda olup, orijinal ağda olmayan (fazladan gösterilen) nesne sayısı ise  $F_s$  ile ifade edilmektedir. Şekil 6.3, 6.4, 6.5 ve 6.6'da bu değerlere ait sütun grafikler gösterilmiştir.

Çizelge 6.4, 6.5, 6.7 ve 6.8'de görünen  $x \equiv y$  (örneğin,  $201 \equiv 97$ ), önerilen yöntemlerle 1:100,000 ölçeğinde türetilen ağdaki  $x$  nesne sayısının, orijinal 1:100,000 ölçekli ağdaki  $y$  nesne sayısına denk olduğunu ifade etmektedir. Bunun iki nedeni vardır:

Birincisi, göllerin elenmesi ve alan-çizgi dönüşümü gibi genelleştirme işlemleri nedeniyle orijinal 1:100,000 ölçekli ağdaki nesneler, 1:24,000 ölçeğindeki nesnelerden

sayı ve geometri bakımından farklı olabilmektedir. Bununla birlikte, önerilen yöntemlerde göller ele alınmadığından ve alan-çizgi dönüşümü gibi kartografik genelleştirme işlemleri uygulanmadığından, orijinal 1:100,000 ölçekli ağdaki ve önerilen yöntemlerle türetilen ağdaki aynı nesneler, farklı sayı ve geometrilerde karşımıza çıkabilmektedir. Şekil 6.2a'da, orijinal 1:24,000 ölçeğinde üç çizgi ve iki alan geometriden oluşan gösterim; Şekil 6.2b'de, orijinal 1:100,000 ölçeğinde tek bir çizgi ile gösterilmişken; Şekil 6.2c'de, önerilen yöntemlere göre türetilen ağda -orijinal 1:24,000 ölçeğindeki gibi- üç çizgi ve iki alan geometri ile gösterilmiş olabilir. Önerilen yöntemlerin sonuçları, orijinal 1:100,000 ölçeğindekiyle karşılaştırıldığından, bu durumda, önerilen yöntemlere göre türetilen ağdaki üç nesne, orijinal 1:100,000 ölçeğindeki tek çizgi nesneye denk kabul edilmektedir.

İkincisi, NHD veri tabanındaki orijinal 1:24,000 ve 1:100,000 ölçekli ağlarda, nesne tipi bakımından bazı tutarsızlıklar göze çarpmaktadır. Şekil 5.5 ve 5.7'de Little Pomme de Terre alt havzasında görünen 1:24,000 ölçeğinde daimi dereler, 1:100,000 ölçeğinde mevsimsel derelere dönüşmüş ve semantik açıdan yanlış bir gösterim ortaya çıkmıştır (Alt havza isimleri: Şekil 5.3).



Şekil 6.2 a) Orijinal 1:24,000, b) orijinal 1:100,000, c) türetilen 1:100,000

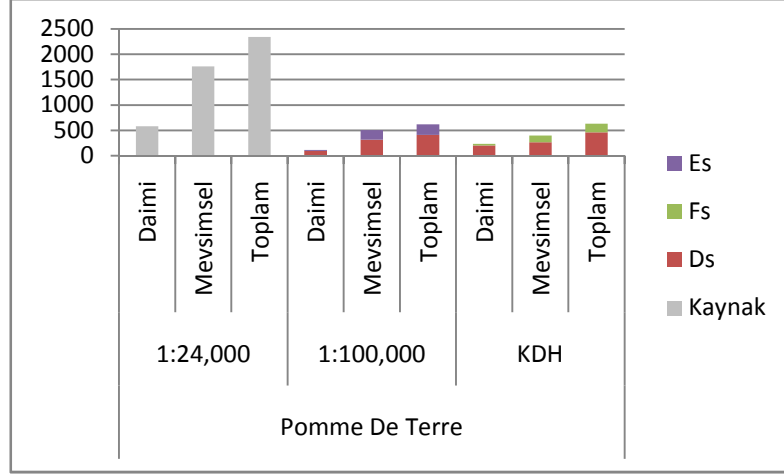
Şekil 6.3, 6.4, 6.5 ve 6.6'da görünen grafiklerdeki denk nesne sayılarından, orijinal ve türetilen ağlar, nesne tipi açısından genellikle uyumludur..

Çizelge 6.4 KDH yöntemiyle türetilen 1:100,000 ölçeğinde akarsu ağlarının, orijinal akarsu ağları ile akarsu tipi açısından karşılaştırılması

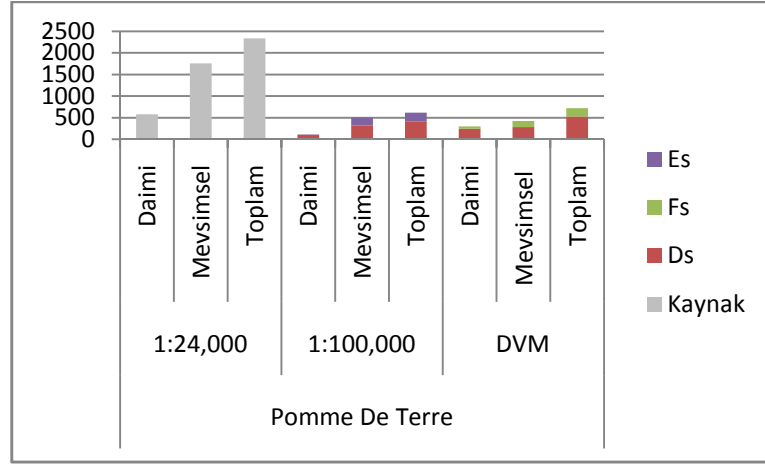
| Yöntem | Havza | Karşılaştırma  | Akarsu Tipi |            |              |
|--------|-------|----------------|-------------|------------|--------------|
|        |       |                | Daimi       | Mevsimsel  | Toplam       |
| KDH    | PT    | Tümü           | 236<br>%37  | 396<br>%63 | 632<br>%100  |
|        |       | D <sub>s</sub> | 201≡97      | 262≡314    | 463≡411      |
|        |       | F <sub>s</sub> | 35          | 134        | 169          |
|        |       | E <sub>s</sub> | 15          | 193        | 208          |
|        | SBP   | Tümü           | 608<br>%60  | 406<br>%40 | 1014<br>%100 |
|        |       | D <sub>s</sub> | 469≡432     | 202≡187    | 671≡619      |
|        |       | F <sub>s</sub> | 139         | 204        | 343          |
|        |       | E <sub>s</sub> | 77          | 371        | 448          |

Çizelge 6.5 DVM yöntemiyle türetilen 1:100,000 ölçeğinde akarsu ağlarının, orijinal akarsu ağları ile akarsu tipi açısından karşılaştırılması

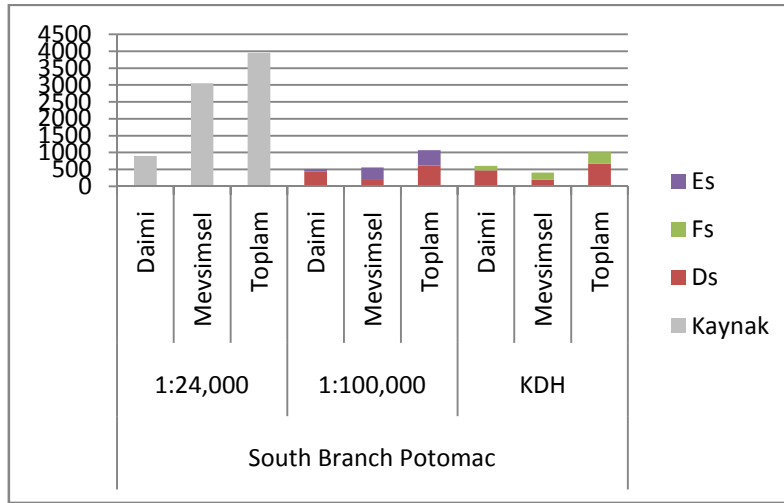
| Yöntem | Havza | Karşılaştırma  | Akarsu Tipi |            |              |
|--------|-------|----------------|-------------|------------|--------------|
|        |       |                | Daimi       | Mevsimsel  | Toplam       |
| DVM    | PT    | Tümü           | 300<br>%42  | 422<br>%58 | 722<br>%100  |
|        |       | D <sub>s</sub> | 246≡102     | 281≡317    | 527≡419      |
|        |       | F <sub>s</sub> | 54          | 141        | 195          |
|        |       | E <sub>s</sub> | 10          | 190        | 200          |
|        | SBP   | Tümü           | 712<br>%58  | 508<br>%42 | 1220<br>%100 |
|        |       | D <sub>s</sub> | 541≡469     | 292≡254    | 833≡723      |
|        |       | F <sub>s</sub> | 171         | 216        | 387          |
|        |       | E <sub>s</sub> | 40          | 304        | 344          |



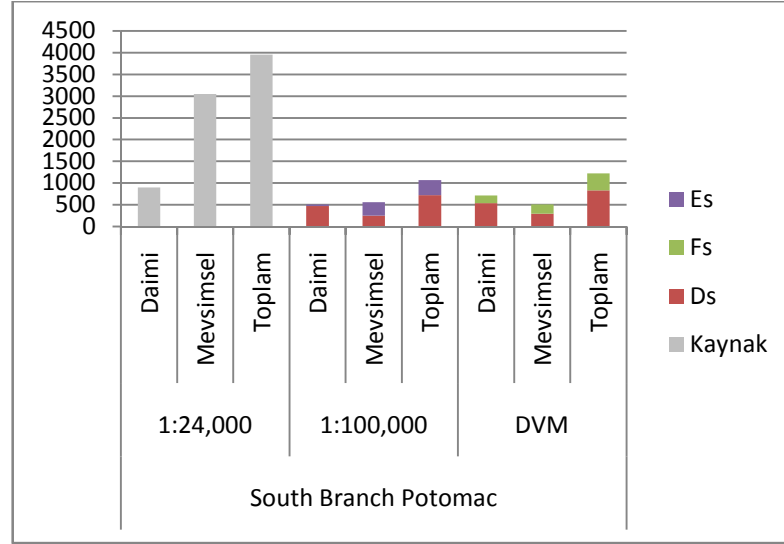
Şekil 6.3 Orijinal PT’de ve KDH’ye göre türetilen haritada daimi ve mevsimsel dere sayısı ile ilgili sütun grafik



Şekil 6.4 Orijinal PT’de ve DVM’ye göre türetilen haritada daimi ve mevsimsel dere sayısı ile ilgili sütun grafik



Şekil 6.5 Orijinal SBP’de ve KDH’ye göre türetilen haritada daimi ve mevsimsel dere sayısı ile ilgili sütun grafik



Şekil 6.6 Orijinal SBP’de ve DVM’ye göre türetilen haritada daimi ve mevsimsel dere sayısı ile ilgili sütun grafik

Hiyerarşi sınıflarına (akarsu düzeylerine) göre orijinal akarsu sayıları, Çizelge 6.6’da verilmektedir. Çizelge 6.7 ve 6.8’de ise KDH ve DVM yöntemlerine göre türetilen ağlara ait akarsu düzeylerine göre nesne sayıları verilmiştir. Seçim, Çizelge 6.7 ve 6.8’deki yüzdelere göre, her iki yöntemde de yüksek hiyerarşiden düşük hiyerarşiye doğru azalmaktadır (en yüksek hiyerarşi “2”, en düşük hiyerarşi “9”). Türetilen ağlarda bazı hiyerarşi düzeylerinde (örneğin, PT için 7. düzeyde) eleme diğer hiyerarşi düzeylerindekiyle göre daha fazla gerçekleştirilmiştir.

DVM yöntemi ile türetilen ağın nesne sayısı, KDH ile üretilen ağın nesne sayısından fazladır (Çizelge 6.2). Fazla nesne sayısı genellikle orta hiyerarşi düzeyinde bulunmaktadır. PT için DVM ile türetilen ağın 5. ve 6. düzeydeki nesne sayıları, KDH ile türetilen ağın aynı düzeylerindeki nesne sayılarından daha fazladır. SBP için DVM ile türetilen ağın 4. düzeyindeki nesne sayıları ise KDH ile türetilen ağın aynı düzeylerindeki nesne sayılarından daha fazladır. Bu durum daha düşük düzeylerde ortaya çıksaydı, yöntemin geliştirme için uygunluğu azalabilirdi. Sonuç olarak 1:100,000 ölçekli türetilen ağlar, hiyerarşi bakımından orijinal ağlar ile örtüşme ve eğilim açısından benzerdir. Çizelge 6.6, 6.7 ve 6.8’e ait grafikler, Şekil 6.7, 6.8, 6.9 ve 6.10’da verilmektedir. Orijinal ve türetilen ağlara ait haritalar, Şekil 6.11, 6.12, 6.13 ve 6.14’de verilmiştir. KDH ile türetilen ağların, DVM ile türetilen ağlara ve orijinal ağlara göre alt havza bazında akarsu yoğunluğunu daha iyi koruduğu anlaşılmaktadır.

Çizelge 6.6 Orijinal 1:24,000 ve 1:100,000 ölçeğinde PT ve SBP akarsu ağlarına ait akarsu düzeyi ile ilgili nicel bilgiler

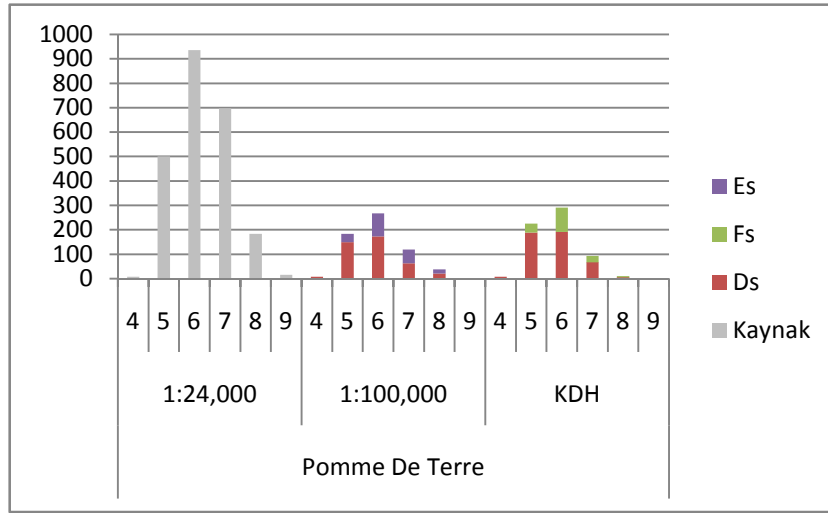
| Yöntem   | Havza | Ölçek              | Akarsu düzeylerine göre nesne sayıları |            |            |            |            |            |           |          |
|----------|-------|--------------------|----------------------------------------|------------|------------|------------|------------|------------|-----------|----------|
|          |       |                    | 2                                      | 3          | 4          | 5          | 6          | 7          | 8         | 9        |
| Orijinal | PT    | Orijinal 1:24,000  | -                                      | -          | 8          | 500        | 936        | 697        | 183       | 16       |
|          |       | Orijinal 1:100,000 | -                                      | -          | 8<br>%100  | 184<br>%37 | 268<br>%29 | 119<br>%17 | 38<br>%21 | 2<br>%12 |
|          | SBP   | Orijinal 1:24,000  | 5                                      | 358        | 1386       | 1458       | 619        | 114        | 11        | 1        |
|          |       | Orijinal 1:100,000 | 2<br>%40                               | 180<br>%50 | 473<br>%34 | 318<br>%22 | 67<br>%11  | 7<br>%6    | -         | -        |

Çizelge 6.7 KDH yöntemiyle türetilen 1:100,000 ölçeğinde akarsu ağlarının, orijinal akarsu ağları ile akarsu düzeyi açısından karşılaştırılması

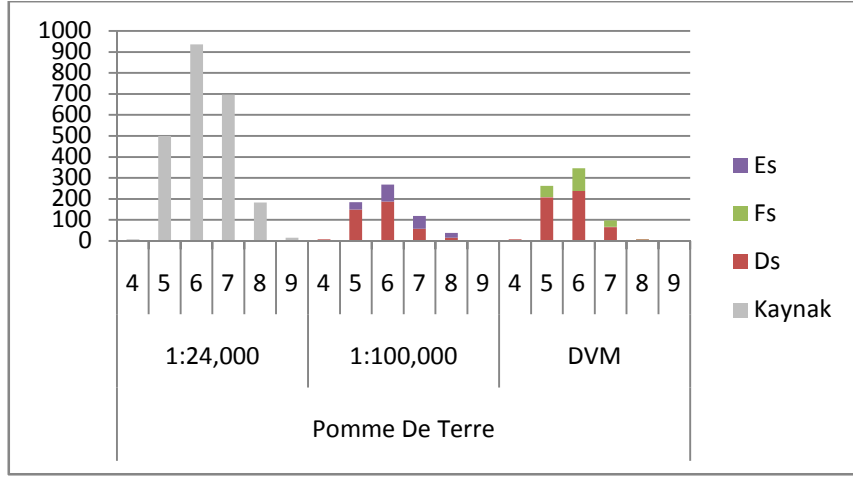
| Yöntem | Havza | Ölçek          | Akarsu düzeylerine göre nesne sayıları |             |             |             |             |           |          |         |
|--------|-------|----------------|----------------------------------------|-------------|-------------|-------------|-------------|-----------|----------|---------|
|        |       |                | 2                                      | 3           | 4           | 5           | 6           | 7         | 8        | 9       |
| KDH    | PT    | Tümü           | -                                      | -           | 8<br>%100   | 226<br>%45  | 292<br>%31  | 94<br>%13 | 11<br>%6 | 1<br>%6 |
|        |       | D <sub>s</sub> | -                                      | -           | 8≡8         | 189<br>≡150 | 192<br>≡173 | 68≡63     | 6≡21     | 0       |
|        |       | F <sub>s</sub> | -                                      | -           | 0           | 37          | 100         | 26        | 5        | 1       |
|        |       | E <sub>s</sub> | -                                      | -           | 0           | 34          | 95          | 56        | 17       | 2       |
|        | SBP   | Tümü           | 4<br>%80                               | 208<br>%58  | 510<br>%37  | 257<br>%18  | 31<br>%5    | 4<br>%3   | -        | -       |
|        |       | D <sub>s</sub> | 4≡2                                    | 148<br>≡125 | 360<br>≡297 | 140<br>≡162 | 18≡21       | 1≡1       | -        | -       |
|        |       | F <sub>s</sub> | 0                                      | 60          | 150         | 117         | 13          | 3         | -        | -       |
|        |       | E <sub>s</sub> | 0                                      | 55          | 176         | 156         | 46          | 6         | -        | -       |

Çizelge 6.8 DVM yöntemiyle türetilen 1:100,000 ölçeğinde akarsu ağlarının, orijinal akarsu ağları ile akarsu düzeyi açısından karşılaştırılması

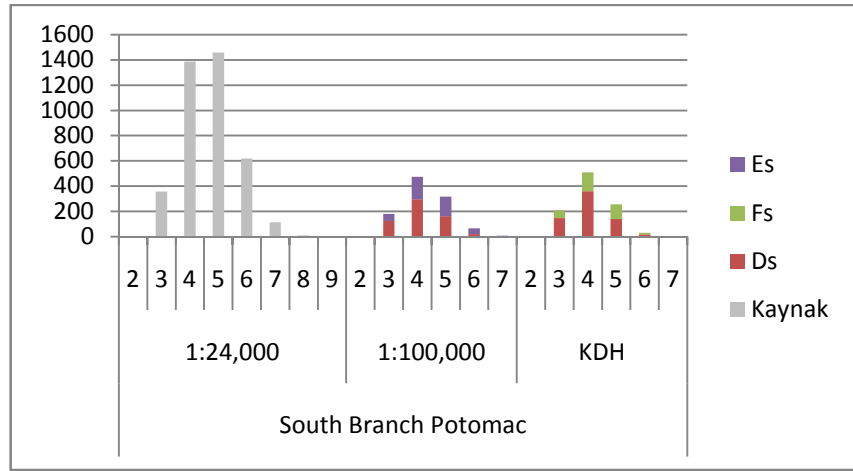
|        |       |                | Akarsu düzeylerine göre nesne sayıları |             |             |             |             |           |         |         |
|--------|-------|----------------|----------------------------------------|-------------|-------------|-------------|-------------|-----------|---------|---------|
| Yöntem | Havza | Ölçek          | 2                                      | 3           | 4           | 5           | 6           | 7         | 8       | 9       |
| DVM    | PT    | Tümü           | -                                      | -           | 8<br>%100   | 262<br>%52  | 346<br>%37  | 97<br>%14 | 8<br>%4 | 1<br>%6 |
|        |       | D <sub>s</sub> | -                                      | -           | 8≡8         | 208<br>≡150 | 238<br>≡188 | 66≡58     | 7≡15    | 0       |
|        |       | F <sub>s</sub> | -                                      | -           | 0           | 54          | 108         | 31        | 1       | 1       |
|        |       | E <sub>s</sub> | -                                      | -           | 0           | 34          | 80          | 61        | 23      | 2       |
|        | SBP   | Tümü           | 5<br>%100                              | 254<br>%71  | 645<br>%46  | 273<br>%19  | 40<br>%6    | 3<br>%3   | -       | -       |
|        |       | D <sub>s</sub> | 4≡2                                    | 183<br>≡146 | 454<br>≡349 | 168<br>≡182 | 23≡29       | 1≡1       | -       | -       |
|        |       | F <sub>s</sub> | 1                                      | 71          | 191         | 105         | 17          | 2         | -       | -       |
|        |       | E <sub>s</sub> | 0                                      | 34          | 124         | 136         | 38          | 6         | -       | -       |



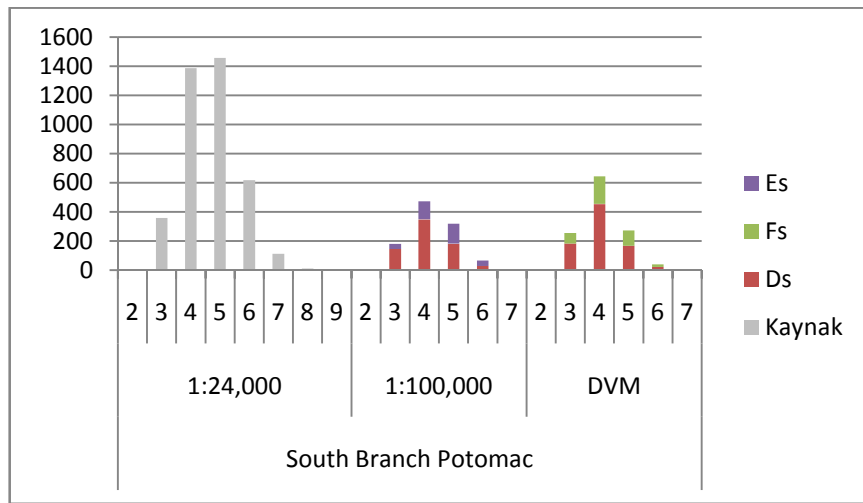
Şekil 6.7 Orijinal PT’de ve KDH’ye göre türetilen haritada akarsu düzeyleri ile ilgili sütun grafik



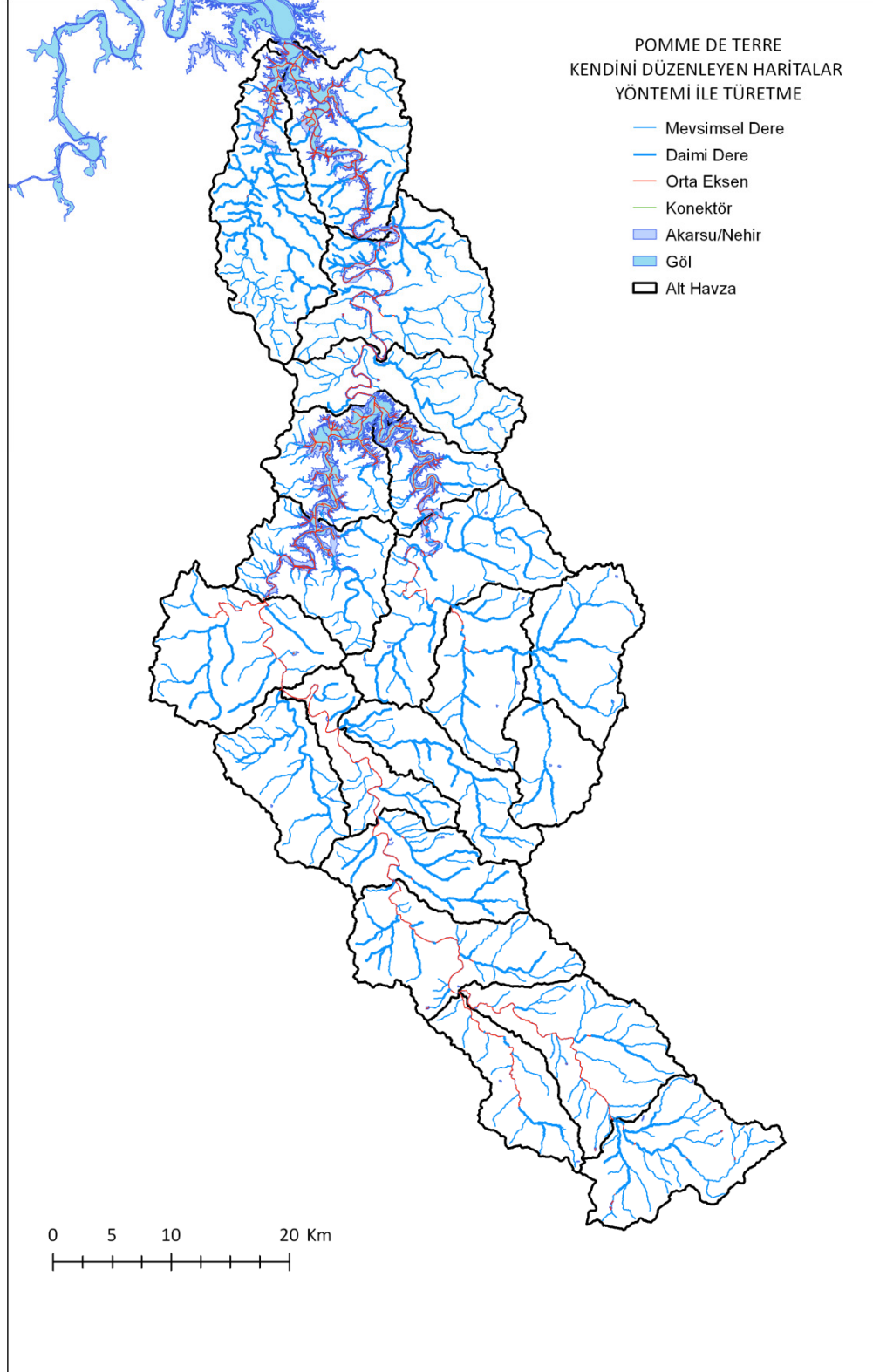
Şekil 6.8 Orjinal PT’de ve DVM’ye göre türetilen haritada akarsu düzeyleri ile ilgili sütun grafik



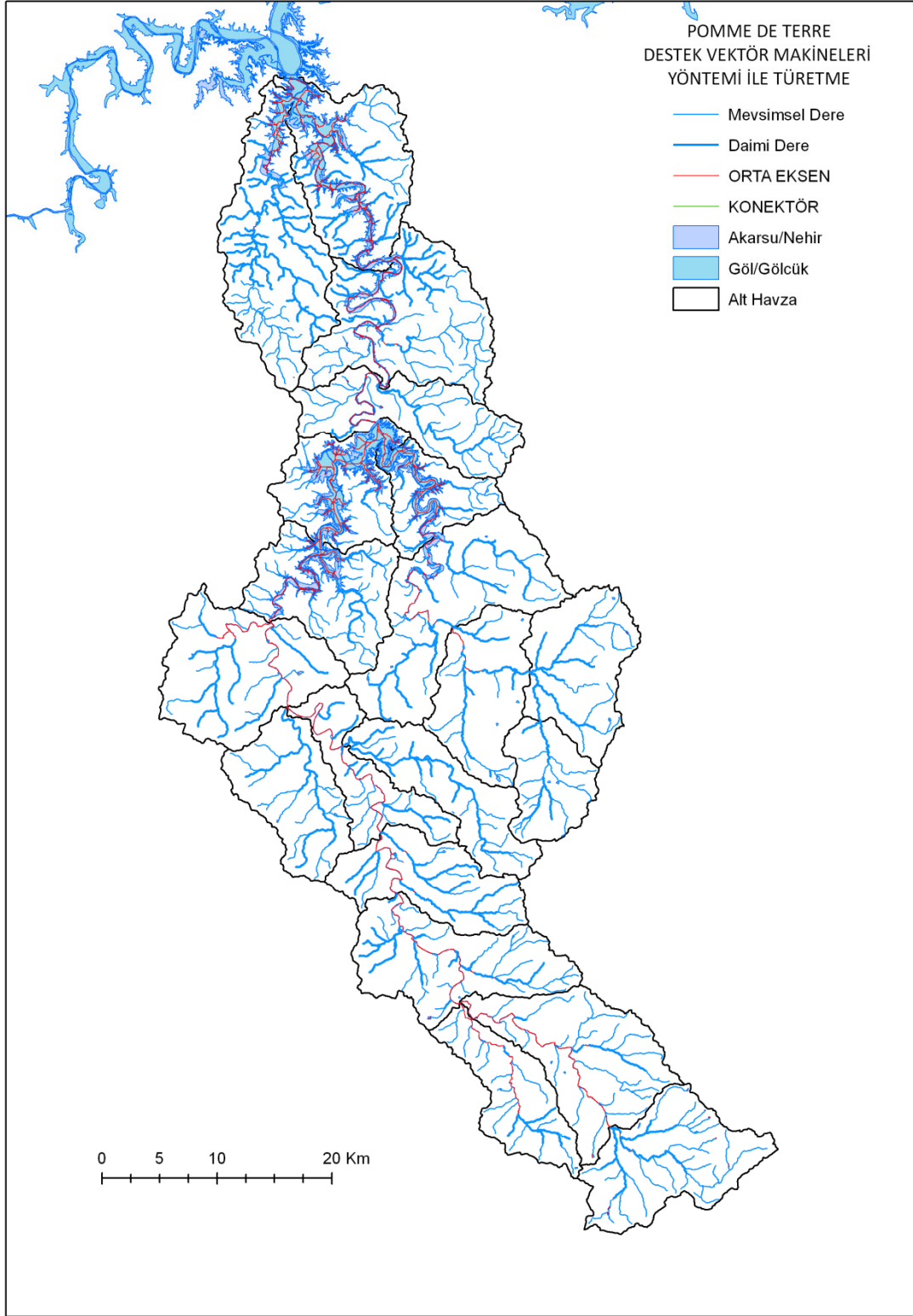
Şekil 6.9 Orjinal SBP’de ve KDH’ye göre türetilen haritada akarsu düzeyleri ile ilgili sütun grafik



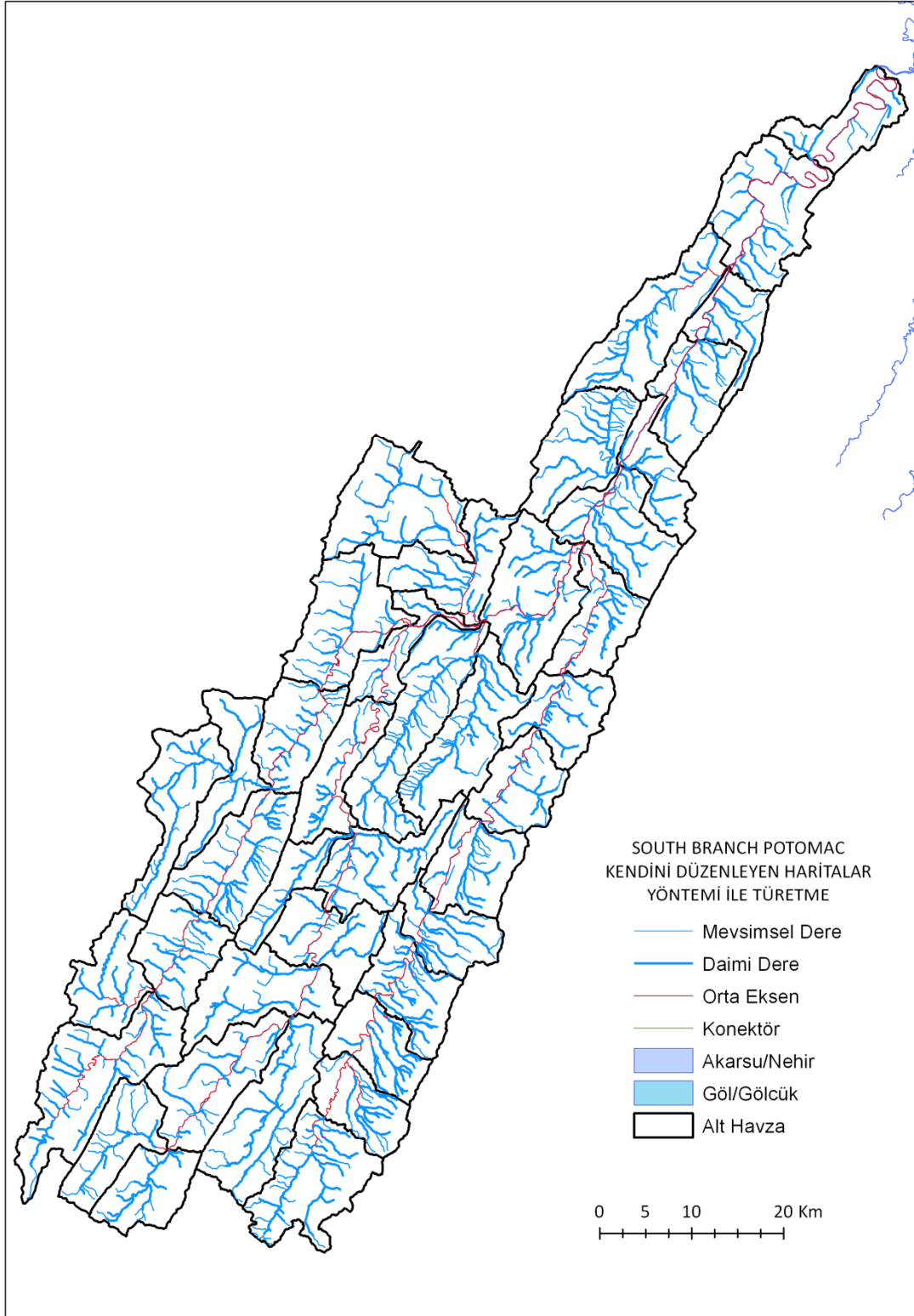
Şekil 6.10 Orjinal SBP’de ve DVM’ye göre türetilen haritada akarsu düzeyleri ile ilgili sütun grafik



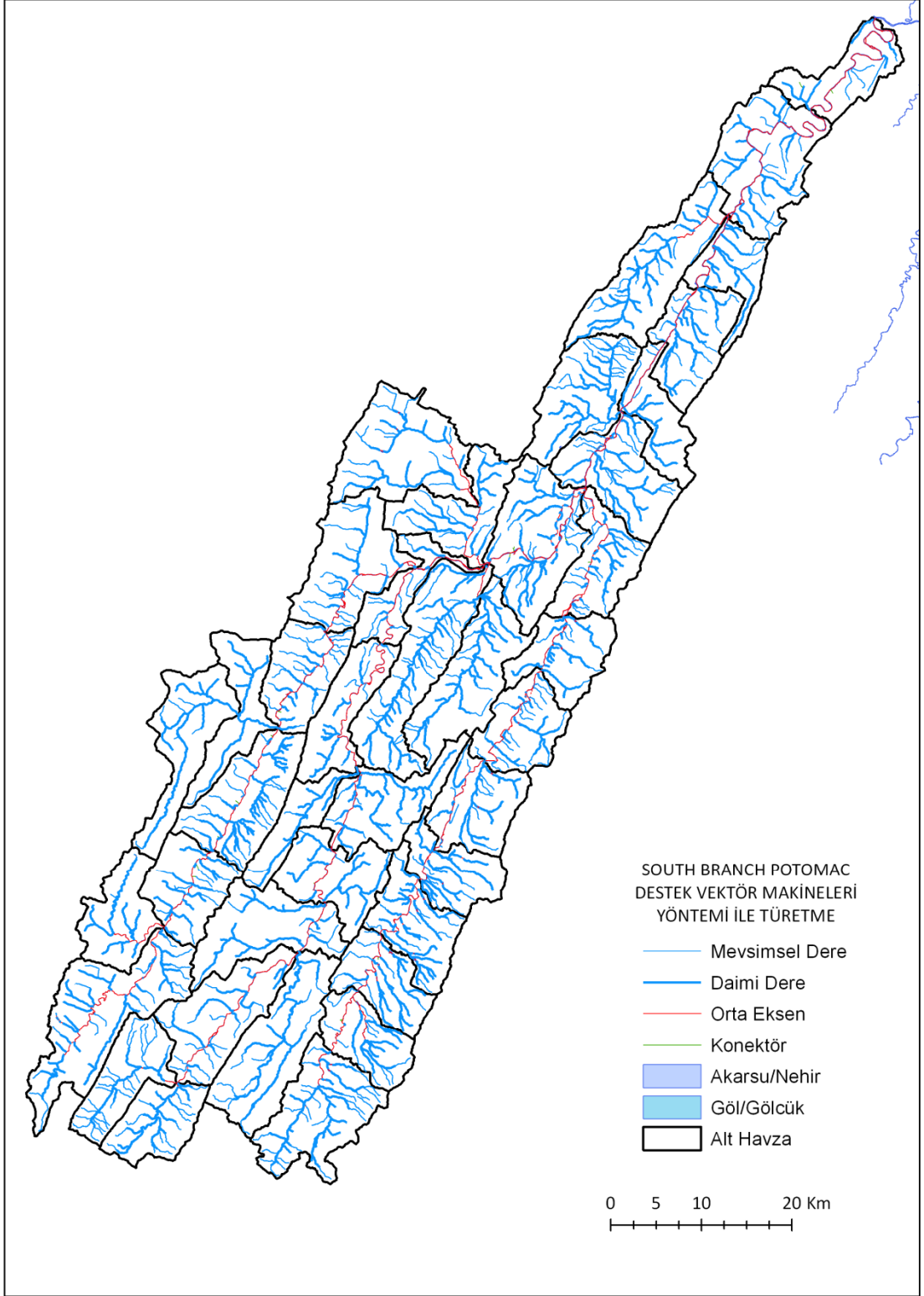
Şekil 6.11 Kendini düzenleyen haritalar yöntemiyle türetilen 1:100,000 PT akarsu ağı  
(1:400,000 ölçekli gösterim)



Şekil 6.12 Destek vektör makineleri yöntemiyle türetilen 1:100,000 PT akarsu ağı  
(1:400,000 ölçekli gösterim)



Şekil 6.13 Kendini düzenleyen haritalar yöntemiyle türetilen 1:100,000 SBP akarsu ağı  
(1:500,000 ölçekli gösterim)



Şekil 6.14 Destek vektör makineleri yöntemiyle türetilen 1:100,000 SBP akarsu ağı  
(1:500,000 ölçekli gösterim)

### SONUÇ VE ÖNERİLER

Bu çalışmada geliştirilen KDH temelli yöntem ve metodoloji ile akarsu ağının model genelleştirme işlemi tam otomatik olarak gerçekleştirilebilmektedir. Oysaki DVM, ceza ve kernel parametrelerine bağımlıdır. Ayrıca, sınıflandırmanın doğruluğu, eğitim setinin karakteristik deseni içerip içermediğine bağlı olduğu için, doğru eğitim setinin belirlenmesi bu yöntemdeki en kritik aşamadır ve sonucu doğrudan etkilemektedir. Günümüzde eğitim seti kullanıcı tarafından belirlenmektedir ve bu nedenle tam otomatik bir yöntem değildir. Bu bağlamda, eğitim setinin otomatik olarak belirlenmesi araştırmaya açık bir konudur.

Çalışmanın başlıca özellikleri aşağıdaki gibi özetlenebilir.

- Uzunluk, kıvrımlılık, bağlılık bakımından merkezlik, arasında olma bakımından merkezlik, yakınlık bakımından merkezlik, akarsu düzeyi ve akarsu tipi genelleştirmede seçme/eleme için en uygun öznitelikler olarak belirlenmiştir.
- Öznitelikler istatistiksel olarak analiz edilmiştir.
- İki farklı desen için  $\chi^2$  bağımsızlık testi ile normal dağılımlı olmayan özniteliklerin orijinal 1:100,000 ölçekli akarsu ağı ile ilişkileri belirlenmiş ve ağırlıkları hesaplanmıştır.
- Hidrografik genelleştirmede seçme/eleme işlemi için KDH ve DVM yöntemleri ilk defa uygulanmıştır.
- KDH ile akarsu ağının model genelleştirmesinde “kaç adet kol seçilmelidir?” sorununun yanıtı Radikal Kanun ile belirlenmiştir.

- “KDH’de hangi kümeler seçilmelidir?” sorusunun yanıtı için U-matrisinde komşuluk ilişkilerine göre yerleşmiş olan kümelerin kademeli olarak seçildiği yeni bir yöntem geliştirilmiştir.
- DVM yönteminde eğitim verisi olarak, en karakteristik olduğu düşünülen ve ortalama nesne sayısına sahip bir alt havza kullanılmış ve en uygun hiper düzlem belirlenmiştir. Diğer alt havzalar, en uygun hiper düzleme göre ikili (seçme/eleme) olarak sınıflandırılmıştır.
- Her iki yöntemle elde edilen ağlarda kopukluklar (bağılantısız akarsular) olması durumunda, bağlantıyı sağlayacak ilgili akarsular da seçilmiştir. Böylece sonuç akarsu ağı hem alt havza bazında, hem de tüm havzada sürekli hale getirilmiştir.

Sonuç olarak, akarsu ağı seçme/eleme işleminde, aşağıda sıralanan problemler ve subjektif konular aydınlığa kavuşmuştur.

- Hangi öznelitlikler kullanılmalıdır?
- Öznelitlikler nasıl ağırlıklandırılmalıdır?
- Kümelerden/akarsulardan hangileri seçilmelidir?

Önerilen yöntemlerle elde edilen akarsu ağlarının niteliği şöyle değerlendirilebilir.

- ÖÇK değerlerine göre, her iki yöntem, ağaşı desene sahip PT’de -kafes ve dikdörtgen desene sahip SBP’ye göre- daha iyi sonuç vermiştir.
- KDH yöntemi ile PT’den türetilen akarsu ağı, orijinal 1:100,000 ölçekli akarsu ağı ile daha iyi örtüşmektedir. SBP’de her iki yöntemle türetilen akarsu ağları aynı oranda orijinali ile örtüşmektedir.
- KDH yöntemi ile türetilen nesne sayısı -DVM yöntemi ile türetilen nesne sayısına göre- Radikal Kanun’a göre belirlenen nesne sayısına daha yakındır.
- Orijinal ve türetilen ağlar, nesne tipi açısından genellikle uyumludur.
- Her iki yöntemle elde edilen ağlar, hiyerarşi bakımından orijinal 1:100,000 ölçekli ağlar ile benzer özelliktedir. Aynı düzeyde akarsulardan yaklaşık aynı oranda seçim yapılmıştır.

- KDH’de kümelerin seçimi için geliştirilen yöntem alt havza temelli olduğundan, her alt havzada akarsu yoğunluğu ve deseni korunmaktadır.
- Sonuç akarsu ağı hem alt havza bazında hem de tüm havzada sürekli, kopukluklar içermemektedir.
- Akarsu tipleri orijinal 1:24,000 ölçekli akarsu ağlarından elde edildiğinden, türetilen ağlar semantik açıdan tutarlıdır.

Seçme/eleme sonuçları akarsu ağı desenlerine bağlıdır. Gelecekte, KDH ve DVM yöntemlerinin akarsu ağının model genelleştirmesindeki yetileri, farklı desenler için değişik öznitelikler ilave edilerek artırılabilir. Bu durumda özniteliklerin ağırlıklandırılması için yeni bir çalışmaya da gereksinim duyulacaktır.

Bu çalışma ülkemiz verileri ile tekrarlanırsa, gerekli öznitelikler (uzunluk, kıvrımlılık, bağlılık, arasında olma, yakınlık bakımından merkezlik değerleri, akarsu düzeyi ve tipi) türetildiği takdirde, ülkemiz için geçerli ağırlıklar (KDH için) ve eğitim seti (DVM için) belirlenebilir ve böylece bu yöntemler ülkemizde de kullanılabilir.

Bu yöntemlerin diğer ölçek geçişleri (1:25,000-1:50,000, 1:25,000-1:100,000, 1:25,000-1:250,000, 1:50,000-1:100,000, 1:50,000-1:250,000, 1:100,000-1:250,000) için de kullanılabilmesi için, bu çalışmanın her biri için ayrı ayrı tekrarlanması gerekmektedir.

## KAYNAKLAR

---

- [1] Başaraner, M.A., (2005). Nesne Yönelimli Coğrafi Bilgi Sistemi Ortamında Orta Ölçekli Topografik Haritalar için Bina ve Yerleşim Alanlarının Otomatik Genelleştirmesi, Doktora Tezi, YTÜ Fen Bilimleri Enstitüsü, İstanbul.
- [2] Sarjakoski, L.T., (2007). "Conceptual Models of Generalisation and Multiple Representation", 11-35; Derleyen: Mackaness, W.A., Ruas, A. ve Sarjakoski, L.T., (2007). Generalization of Geographic Information: Cartographic Modelling and Applications, Elsevier, Amsterdam.
- [3] Smaalen, J.W.N. van, (1996). "Spatial Abstraction Based on Hierarchical Re-classification", Cartographica, 33(1): 65-73.
- [4] Weibel, R., (1997). "Generalization of Spatial Data – Principles and Selected Algorithms", 1340: 99-152; Derleyen: Van Kreveld, M., Nievergelt, J., Ross, Th., Widmayer, P., Algorithmic Foundations of GIS Lecture Notes in Computer Science, Springer-Verlag, Berlin.
- [5] Jones, C., (1997). Geographic Information Systems and Computer Cartography, Addison Wesley Longman, Harlow, 319.
- [6] Jaakkola, O., (1998). "Multi-Scale Categorical Data Bases with Automatic Generalization Transformations Based on Map Algebra", Cartography and Geographic Information Systems, 25(4): 195-207.
- [7] Brassel, K. ve Weibel, R., (2002). "Generalization", 2; Derleyen: Anson, R.W. ve Ormeling, F.J., (2002). Basic Cartography for Students and Technicians, ikinci baskı, Butterworth-Heinemann, Londra.
- [8] Sester, M., (2008). "Self-Organizing Maps for Density-Preserving Reduction of Objects in Cartographic Generalization", 107-120; Derleyen: Agarwal, P. ve Skupin, A., (2008). Self-Organizing Maps Applications in Geographic Information Science, John Wiley & Sons, İngiltere.
- [9] Uçar, D., Bildirici, İ.Ö. ve Uluğtekin, N., (2003). "Coğrafi Bilgi Sistemlerinde Model Genelleştirmesi Kavramı ve Geometri ile İlişkisi", TUJK Coğrafi Bilgi Sistemleri ve Jeodezik Ağlar Çalıştayı, Konya.

- [10] Joao, E.M., (1998). Causes and Consequences of Map Generalisation, Taylor & Francis Ltd., Londra, 35-37.
- [11] Robinson, A.H., Arthur, J. L., Muehrcke P. C., Kimerling A. J. ve Guptill, S. C., (1995). Elements of Cartography, Altıncı baskı, Wiley, ABD.
- [12] McMaster, R.B. ve Shea, K.S., (1992). "Generalization in Digital Cartography", Association of American Geographers, Washington DC.
- [13] Mackaness, W.A., (2007). "Understanding geographic space", 1-10; Derleyen: Mackaness, W.A., Ruas, A. and Sarjakoski, L.T., (2007). Generalization of Geographic Information: Cartographic Modelling and Applications, Elsevier, Amsterdam.
- [14] Beard, K., (1991). "Theory of the Cartographic Line Revisited", Cartographica, 4(28): 32-58.
- [15] Regnauld, N. ve McMaster R. B., (2007). "A Synoptic View of Generalisation Operators", 37-66; Derleyen: Mackaness, W.A., Ruas, A. and Sarjakoski, L.T., (2007). Generalization of Geographic Information: Cartographic Modelling and Applications, Elsevier, Amsterdam.
- [16] Bank, E., (1998). Büyük Ölçekli Haritalardan Bilgisayar Destekli Genelleştirme ile Küçük Ölçekli Harita Üretimi, Doktora Tezi, YTÜ Fen Bilimleri Enstitüsü, İstanbul.
- [17] Gülgen, F., (2009). Yerleşim İçi Yol Ağı Genelleştirmesinde Yeni bir Seçme/Eleme Yaklaşımı, Doktora Tezi, YTÜ Fen Bilimleri Enstitüsü, İstanbul.
- [18] Töpfer, F. ve Pillewiser, W., (1966). "The Principles of Selection", Cartographic Journal, 3(1): 10-16.
- [19] Dutton, G., (1999). "Scale, Sinuosity, and Point Selection in Digital Line Generalization", Cartography and Geographic Information Science, 26(1): 33-53.
- [20] Kilpelainen, T., (1999). Map Generalisation in the Nordic Countries, Reports of The Finnish Geodetic Institute, Finland.
- [21] Staufenbiel W., (1973). Zur Automation der Generalisierung topographischer Karten mit besonderer Berücksichtigung grossmaßstäbiger Gebäudedarstellungen, Doktora Tezi, Hannover Üniversitesi.
- [22] Kilpelainen, T., (1995). "Updating Multiple Representation Geodata Bases by Incremental Generalization", Geo-Information-Systeme, 8(4): 13-18.
- [23] Regnauld, N. ve Mackaness, W., (2006). "Creating a Hydrographic Network from a Cartographic Representation: A Case Study Using Ordnance Survey MasterMap Data", International Journal of Geographical Information Science, 20(6): 611-631.
- [24] Ai, T., Liu, Y. ve Chen, J., (2006). "The hierarchical watershed partitioning and data simplification of river network", 11: 617-632; Derleyen: Riedl, A., Kainz, W., Elmes, G.A., (2006), Progress in Spatial Data Handling, Springer.

- [25] Çetinkaya, B., (2006). Topografik Haritaların Üretiminde Eş Yükseklik Eğrileri, Akan Su ve Su İletim Hatları Coğrafi Verilerinin Otomasyon Süreçleri, Doktora Tezi, İTÜ Fen Bilimleri Enstitüsü, İstanbul.
- [26] HGK, (1999). 1:50,000 Ölçekli Haritaların Tahviline Ait Teknik Talimat, Harita Genel Komutanlığı, Ankara.
- [27] HGK, (1981). Genel Kartografya, Harita Genel Komutanlığı, Ankara.
- [28] HGK, (1964). Tahvil Talimatı (1:100,000), Harita Genel Komutanlığı, Ankara.
- [29] Gökgöz, T., (1999). Yükseklik Eğrilerinin Basitleştirilmesinde Yeni Bir Yaklaşım, Doktora Tezi, YTÜ Fen Bilimleri Enstitüsü, İstanbul.
- [30] Bildirici, İ.Ö., (2000). 1:1000 - 1:25,000 Ölçek Aralığında Bina ve Yol Objelerinin Sayısal Ortamda Kartografik Genelleştirmesi, Doktora Tezi, İTÜ Fen Bilimleri Enstitüsü, İstanbul.
- [31] Avcı, M., (2009). 1:25,000 - 1:100.000 Ölçek Aralığında Yol Objelerinin Seçme - Eleme İşlemlerinin Otomasyonu, Doktora Tezi, Selçuk Üniversitesi Fen Bilimleri Enstitüsü, Konya.
- [32] Doğru, A.Ö., (2009). Çoklu Gösterim Veritabanları Kullanılarak Araç Navigasyon Haritası Tasarımı için Kartografik Yaklaşımlar, Doktora Tezi, İTÜ Fen Bilimleri Enstitüsü, İstanbul.
- [33] Aslan, S., (2011). Orta Ölçek Aralığında Binaların Kartografik Genelleştirmesi Doktora Tezi, Selçuk Üniversitesi Fen Bilimleri Enstitüsü, Konya.
- [34] Bilgi, S., (2012). Çok Ölçekli Kartografik Gösterimlerde Mekansal Bilginin Nicelik Analizi, Doktora Tezi, İTÜ Fen Bilimleri Enstitüsü, İstanbul.
- [35] Horton, R.E., (1945). "Erosional Development of Streams and Their Drainage Basins; Hydrophysical Approach to Quantitative Morphology", Bulletin of the Geological Society of America, 56(3): 275-370.
- [36] Strahler, A.N., (1957). "Quantitative Analysis of Watershed Geomorphology" Trans. American Geophys. Union. 38: 913-920.
- [37] Shreve, R.L., (1966). "Statistical Law of Stream Numbers" Journal of Geology, 74: 17-37.
- [38] Wolf, G.W., (1988). "Weighted Surface Networks and Their Application to Cartographic Generalization", 199-212; Derleyen: Barth, W., Visualization Technology and Algorithm, Springer-Verlag, Berlin.
- [39] Richardson, D., (1994). "Generalisation of Spatial and Thematic Data Using Inheritance and Classification and Aggregation Hierarchies", Advances in GIS Research 2, Taylor and Francis, 957-972.
- [40] Thomson, R.C. ve Brooks, R., (2000). "Efficient Generalization and Abstraction of Network Data Using Perceptual Grouping", 5th International Conference on GeoComputation, 23-25 Ağustos, Greenwich.

- [41] Itzhak, E., Yoeli P. ve Doytscher Y., (2001). "Analytic Generalization of Topographic and Hydrologic Data and its Cartographic Display - Intermediate Results", ICC, Beijing, Çin.
- [42] Touya, G., (2007). "River Network Selection Based on Structure and Pattern Recognition" 23rd International Cartographic Conference, 4-9 Ağustos, Moskova, Rusya.
- [43] Stanislawski, L.V., (2009). "Feature Pruning by Upstream Drainage Area to Support Automated Generalization of the United States National Hydrography Dataset", Computers, Environment and Urban Systems, 33: 325-333.
- [44] Şen, A. ve Gökgöz, T., (2012). "Clustering Approaches for Hydrographic Generalization", GIS Ostrava 2012, 23-25 Ocak, Ostrava, Çek Cumhuriyeti.
- [45] Kadmon, N., (1972). "Automated Selection of Settlements in Map Generalization", The Cartographic Journal, 9(2): 93-98.
- [46] Stenhouse, H., (1979). "Selection of Towns on Derived Maps", The Cartographic Journal, 16(1): 30-39.
- [47] Fitzsimons, D.E., (1985). "Base Data on Thematic Maps", The American Cartographer, 12(1): 57-61.
- [48] Mackaness, W.A., (1995). "Analysis of Urban Road Networks to Support Cartographic Generalization", Cartography and Geographic Information Systems, 22(4): 306-316.
- [49] Ruas, A., (1998). "O-O Constraints Modelling to Automate Urban Generalisation Process", SDH'98, Vancouver, 225-235.
- [50] Ruas, A., (1999). "The Role of Meso Level for Urban Generalisation", Workshop on Progress in Automated Map Generalisation, ICA, Ottawa.
- [51] Thomson, R.C. ve Richardson, D.E. (1999). "The 'Good Continuity' Principle of Perceptual Organisation Applied to the Generalisation of Road Networks", 19th International Cartographic Conference, Ottawa, Canada, 1215-1225.
- [52] Jiang, B. ve Harrie, L., (2004). "Selection of Streets from a Network Using Self-Organizing Maps", Transactions in GIS, 8(3): 35-350.
- [53] Gülgen, F. ve Gökgöz, T., (2011). "A Block-Based Selection Method for Road Network Generalization", International Journal of Digital Earth, 4(2): 133-153.
- [54] Skupin, A. ve Agarwal, P., (2008), "Introduction: What is a Self-Organizing Map?", 1-20; Derleyen: Agarwal, P. ve Skupin, A., (2008). Self-Organizing Maps: Applications in Geographic Information Science, John Wiley & Sons, İngiltere.
- [55] Stanislawski, L.V., (2008). "Development of a Knowledge-Based Network Pruning Strategy for Automated Generalisation of the United States National Hydrography Dataset", 11th ICA Workshop on Generalisation and Multiple Representation, Montpellier, Fransa.
- [56] DeBarry P.A., (2004). Watersheds: Processes, Assessment and Management, Wiley, New Jersey.

- [57] Erkek, C. ve Ağırlioğlu, N., (2006). Su Kaynakları Mühendisliği, İTÜ, İstanbul.
- [58] Howard, A.D., (1967). "Drainage Analysis in Geologic Interpretation a Summation." The American Association of Petroleum Geologists Bulletin. 51:11.
- [59] Mazur, E.R. ve Castner, H.W., (1990). "Horton's Ordering Scheme and the Generalisation of River Networks", The Cartographic Journal, 27: 104-112.
- [60] Ferentinou, M., Karymbalis, E., Charou, E., ve Sakellariou, M., (2011). "Using Self Organising Maps in Applied Geomorphology"; Derleyen: Mwasiagi, J.I., (2011). Self Organizing Maps - Applications and Novel Algorithm Design, ISBN: 978-953-307-546-4.
- [61] Horton, R.E., (1932). "Drainage Basin Characteristics" Transactions of the American Geophysical Union, 13: 350-61.
- [62] Strahler, A.N., (1952). "Hypsometric (Area-Altitude) Analysis of Erosional Topography" Bulletin of the Geological Society of America; 63: 1117-1142.
- [63] Patton, P.C., ve Baker, V.R., (1976). "Morphometry and Floods in Small Drainage Basins Subject to Diverse Hydrogeomorphic Controls" Water Resources Research, 12: 941-952.
- [64] Melton, M.A., (1965). "The Geomorphic and Palaeoclimatic Significance of Alluvial Deposits in Southern Arizona" Journal of Geology, 73: 1-38.
- [65] Chorley, R.J., (1972). Spatial Analysis in Geomorphology, Mathuen and Co. Ltd., Londra.
- [66] USGS The National Map, <http://nationalmap.gov/>, 06 Mayıs 2013.
- [67] USGS National Hydrography Dataset, <http://nhd.usgs.gov/>, 06 Mayıs 2013.
- [68] ISO 8211, (1997). Spatial Data Transfer Standard (SDTS), ABD.
- [69] Standards for National Hydrography Dataset, US Environmental Protection Agency, US Departmant of the Interior, US Geological Survey.
- [70] Demirel, H., (2003). Dengeleme Hesabı, YTÜ Basım Yayın Merkezi, İstanbul.
- [71] Kalaycı, Ş., (2010). SPSS Uygulamalı Çok Değişkenli İstatistik Teknikleri, 5. Baskı, Asil Yayın Dağıtım Ltd. Şti., Ankara.
- [72] Çil, B., (2008). İstatistik, 6. Baskı, Detay Yayıncılık, Ankara.
- [73] Kenny, D.A., (1987). For the Social and Behavioral Sciences, Little, Brown and Company, ABD.
- [74] Özkan, Y., (2008). Veri Madenciliği Yöntemleri, Papatya Yayıncılık, İstanbul.
- [75] Vesanto, J., Himberg, J., Alhoniemi, E. ve Parhankangas, J., (1999). "Self-organizing map in Matlab: the SOM toolbox", Matlab DSP Conference, Finland, 16-17 Kasım, 35-40.
- [76] Başaraner, M., (2011). "A Zone-Based Iterative Building Displacement Method Through the Collective Use of Voronoi Tessellation, Spatial Analysis and Multicriteria Decision Making", Boletim de Ciencias Geodesicas, 17(2).

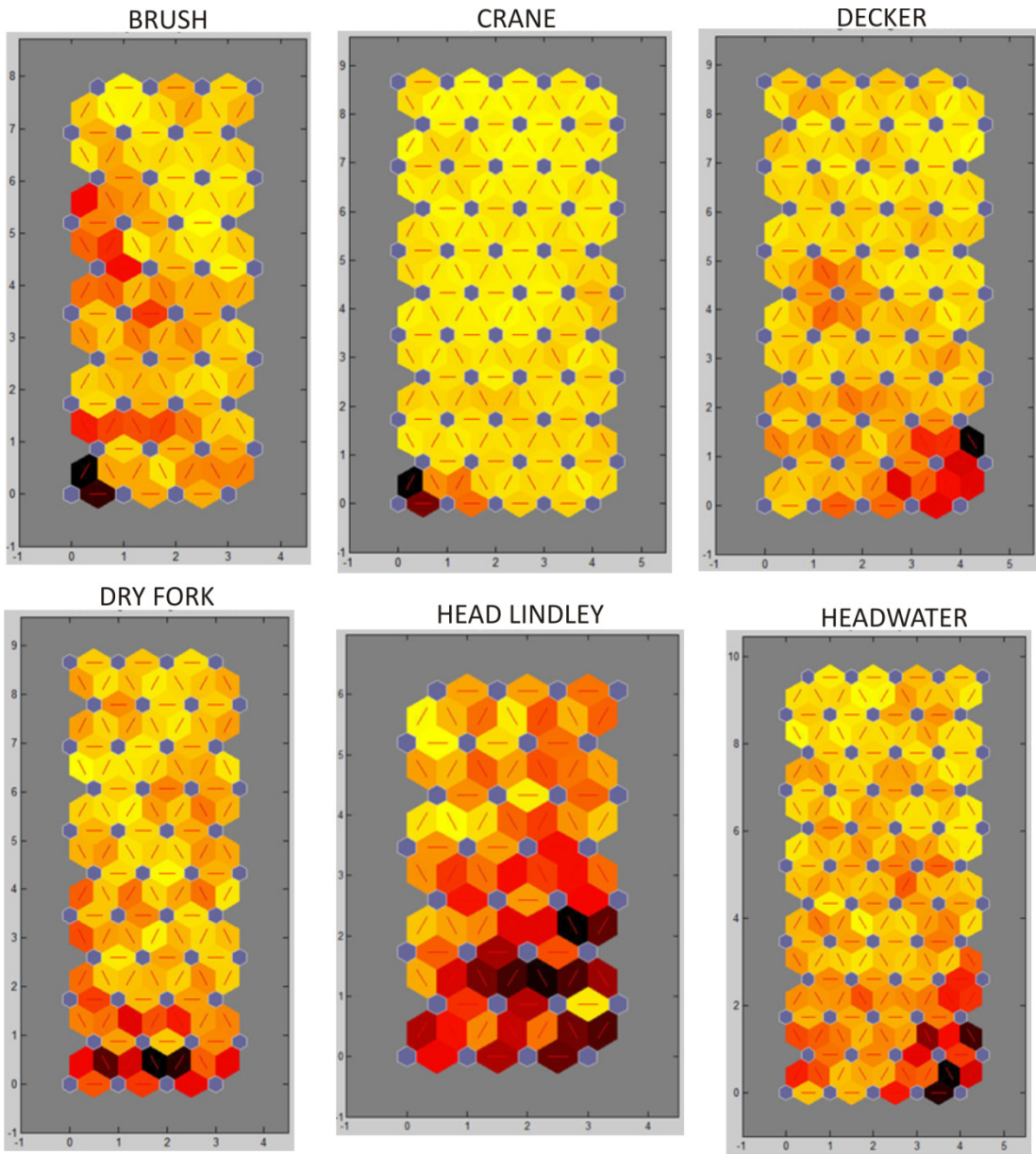
- [77] Wasserman, L., (2004). All of Statistics a Concise Course in Statistical Inference, Springer, 149-375.
- [78] Turan, İ., (2012). Temel İstatistik, <http://temelistatistik.files.wordpress.com/2012/05/spss-4.pdf>, 06 Mayıs 2013.
- [79] Pearson, K., (1895). "Contributions to the Mathematical Theory of Evolution, II: Skew Variation in Homogeneous Material", Philosophical Transactions of the Royal Society of London, 186: 343-414.
- [80] Weinberg S.L., (2008). Statistics Using SPSS: An Integrative Approach, 2 edition, Cambridge University Press.
- [81] Tukey, J.W., (1977). Exploratory Data Analysis. Addison-Wesley.
- [82] Massey, F.J., 1951. "The Kolmogorov-Smirnov Test for Goodness of Fit" Journal of the American Statistical Association. 46(253): 68–78.
- [83] Arnold, T.B. ve Emerson J.W., (2011). "Nonparametric Goodness-of-fit Tests for Discrete Null Distributions" The R Journal, 3(2): 34-39.
- [84] Pearson, K., (1900). "On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling", Philosophical Magazine, Series 5 50(302): 157–175.
- [85] LeBlanc, D., (2004). Statistics: Concepts and Applications for Science, Jones & Bartlett Learning, 339-347.
- [86] Freedman, D., Pisani, R. ve Purves, R., (1978). Statistics, Norton&Company, ABD, 472.
- [87] Öztemel E., (2006). Yapay Sinir Ağları, Papatya yayıncılık, 2. baskı, İstanbul.
- [88] Wilson, I.D., Ware, J.M. ve Ware, J.A., (2003). "A Genetic Algorithm Approach to Cartographic Map Generalisation", Computers in Industry, 52: 291–304.
- [89] Elmas, Ç., (2003). Yapay Sinir Ağları, Seçkin Yayınevi, Ankara.
- [90] Kohonen, T., (2001). Self Organizing Maps. 3. Baskı, Springer.
- [91] Vesanto, J., (2005). SOM Implementation in SOM Toolbox, <http://www.cis.hut.fi/projects/somtoolbox/documentation/somalg.shtml>, 06 Mayıs 2013.
- [92] Kohonen, T., (1990). "The self-organizing map." Proceedings IEEE, 78(9): 1464-1480.
- [93] Wendel, J., ve Battenfield, B.P., (2010). "Formalizing guidelines for building Self-Organizing Maps." Short paper proceedings of GIScience, Zurich, Switzerland.
- [94] Ratle, F., Pozdnoukhov, A., Demyanov, V., Timonin, V., and Savelieva E., (2008). "Spatial Data Analysis and Mapping Using Machine Learning Algorithms", 95-148; Derleyen: Kanevski, M., (2008). Advanced Mapping of

Enviromental Data: Geostatistics, Machine Learning and Bayesian Maximum Entropy, ISTE ,Wiley, 95-148.

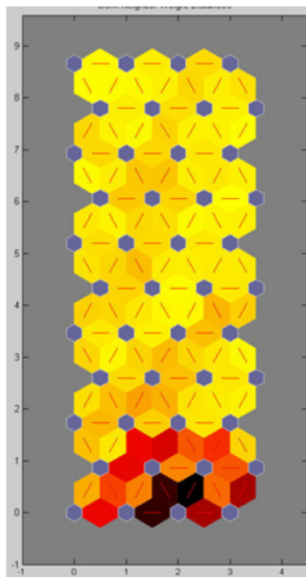
- [95] Ultsch, A., ve Siemon, H., (1989). Technical Report 329, University of Dortmund, Germany.
- [96] Çomak, E., (2008). Destek Vektör Makinelerinin Etkin Eğitimi için Yeni Yaklaşımlar, Doktora Tezi, Selçuk Üniversitesi Fen Bilimleri Enstitüsü, Konya.
- [97] Vapnik, N.V., (1995). The Nature of Statistical Learning Theory, Springer-Verlag, NY.
- [98] Vapnik N.V., (1998). Statistical Learning Theory, John Wiley & Sons, Toronto, CA.
- [99] Erdoğan, B.E., (2012). "Prediction of Bankruptcy Using Support Vector Machines: An Application to Bank Bankruptcy", Journal of Statistical Computation and Simulation, Taylor & Francis, 1-13.
- [100] Vapnik N.V. ve Chervonenkis A.Y., (1991). "The Necessary and Sufficient Conditions for Consistency in the Empirical Risk Minimization Method", Pattern Recognition and Image Analysis, 1(3), 283-305.
- [101] Stanislawski, L., Finn, M., Usery, E.L., ve Barnes, M., (2007). "Assessment of a Rapid Approach for Estimating Catchment Areas for Surface Drainage Lines", ACSM-IPLSA-MSPS 2007, St. Louis Missouri.
- [102] Brewer, C.A., Battenfield, B.P., ve Usery, E.L., (2009). "Evaluating Generalizations of Hydrography in Differing Terrains for The National Map of the United States" 24th International Cartographic Conference (ICC2009), Santiago, Chile.
- [103] Stanislawski, L.V., ve Savino S., (2011). "Pruning of Hydrographic Networks: A Comparison of Two Approaches" 14th ICA/ISPRS Workshop on Generalisation and Multiple Representation, 2011, Paris.
- [104] Şen, A. ve Gökgöz, T., (2011). "Hydrographic Selection/Elimination in Automated Generalization by Using Artificial Neural Networks", (Poster), 25th International Cartographic Conference, 3-8 July 2011, Paris, France.
- [105] Freeman, L.C., (1978). "Centrality in Social Networks: Conceptual Clarification", Social Networks, Elsevier, Netherlands, 1: 215-239.
- [106] Batagelj, V. ve Mrvar, A., (2008). Networks/Pajek: Program for Large Networks Analysis, <http://pajek.imfm.si/doku.php>, 06 Mayıs 2013.
- [107] Baçao, F., Lobo, V. ve Painho, M., (2008). "Applications of Different Self Organizing Map Variants to Geographical Information Science problems", 28-29; Derleyen: Agarwal, P., and Skupin, A., (2008). Self-Organising Maps: Applications in Geographic Information Science, Wiley, İngiltere.
- [108] Hsu, C.-W., Chang, C.-C. Ve Lin, C.-J., (2003), "A Practical Guide to Support Vector Classification", Tech. Rep., Department of Computer Science, National Taiwan University, Taipei.

- [109] Spiess, E., (1995). "The Need for Generalization in a GIS Environment";  
Derleyen: Müller, J.C., Lagrange, J.P., and Weibel, R., GIS and Generalization,  
Methodology and Practise, Taylor & Francis, Londra.

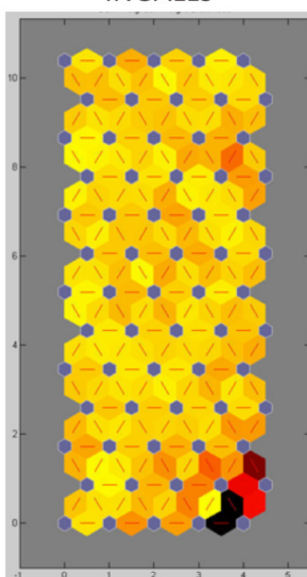
## ALT HAVZALARIN KÜMELEMESİNDE U-MATRİSLERİ



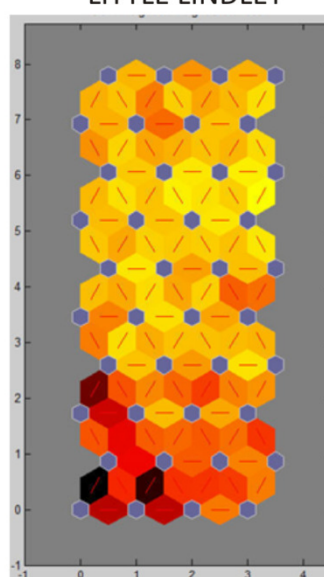
HOMINY



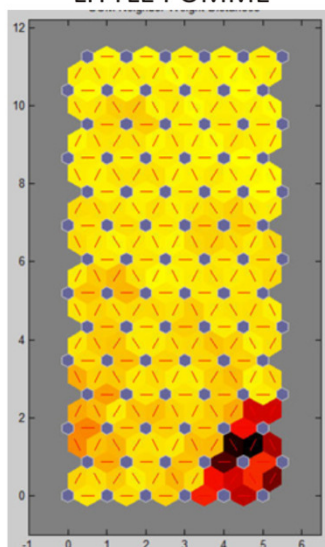
INGALLS



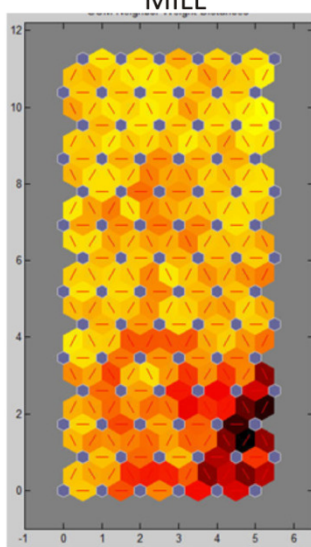
LITTLE LINDLEY



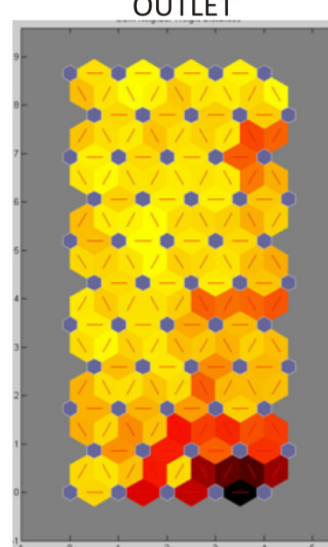
LITTLE POMME



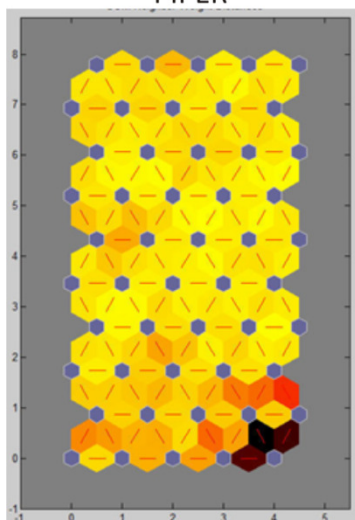
MILL



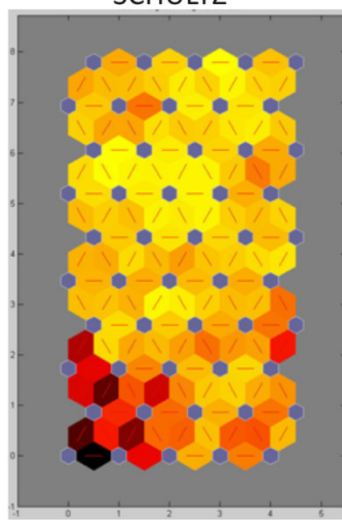
OUTLET



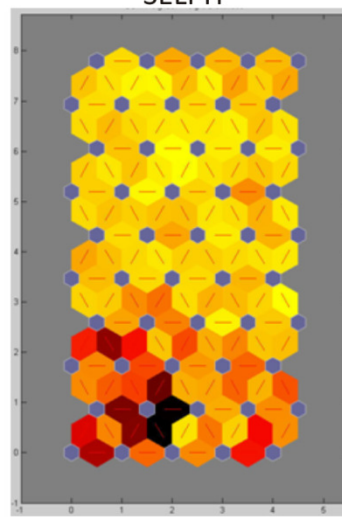
PIPER

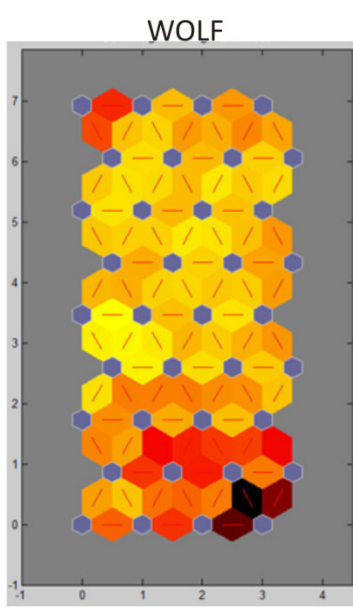
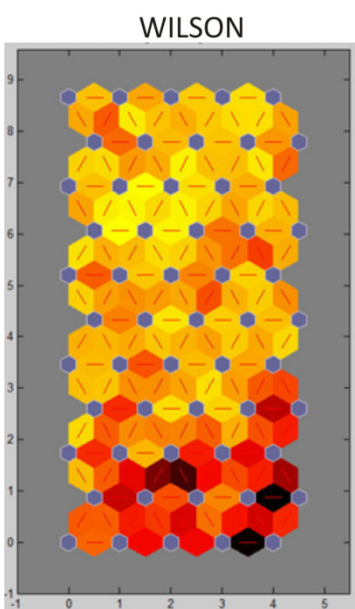
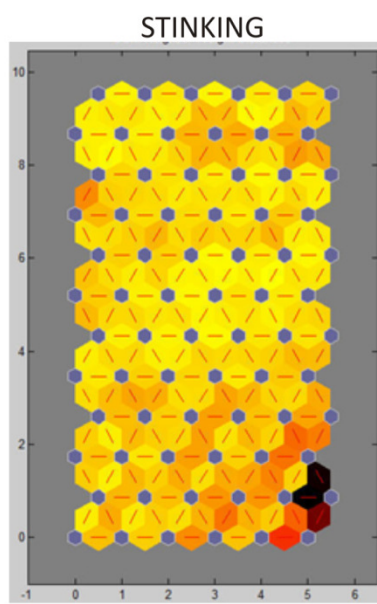
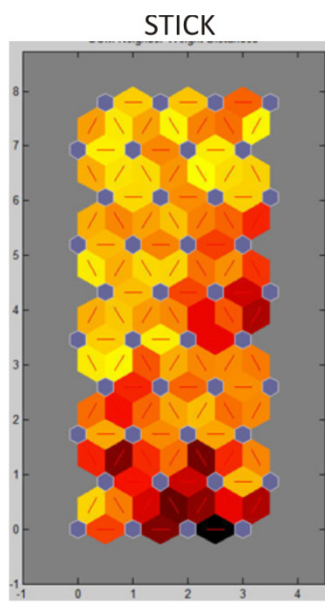
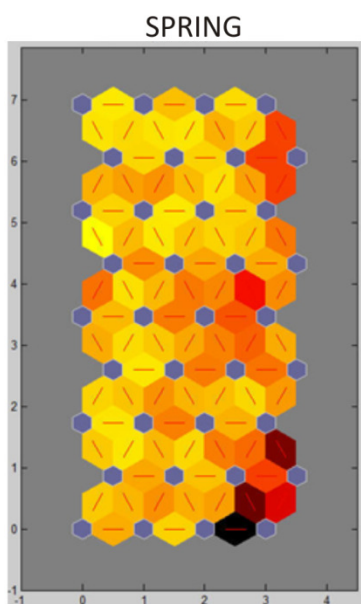


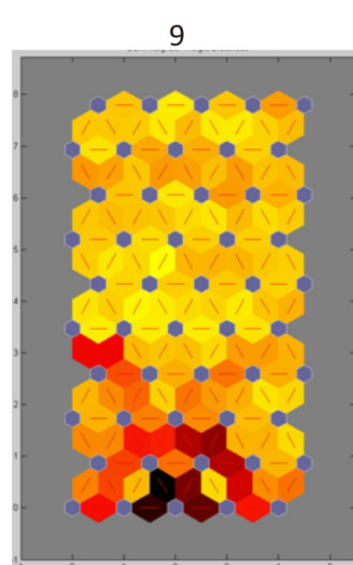
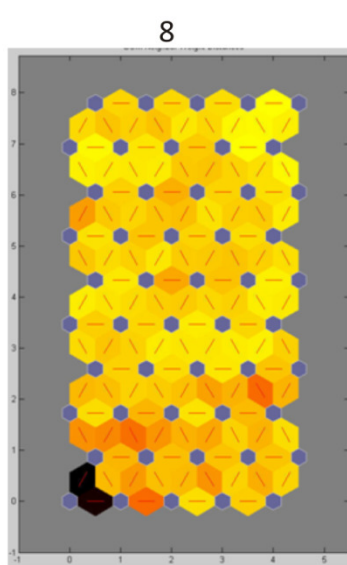
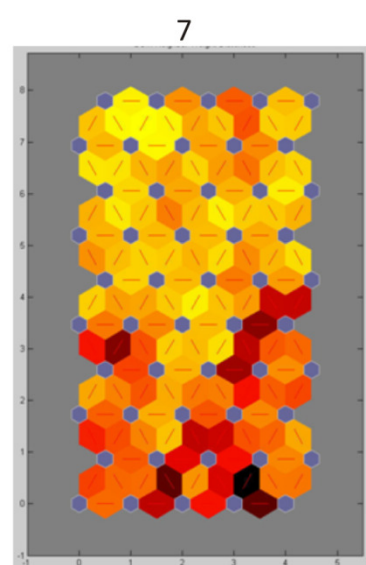
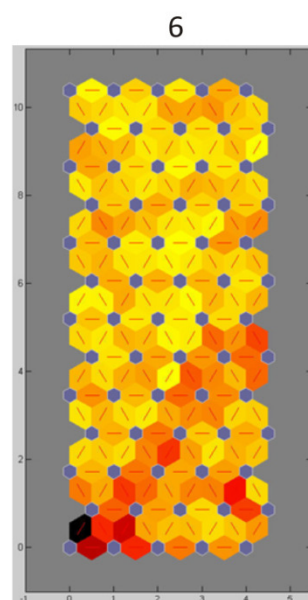
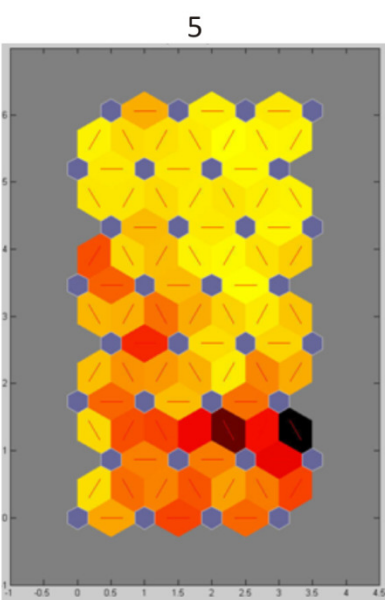
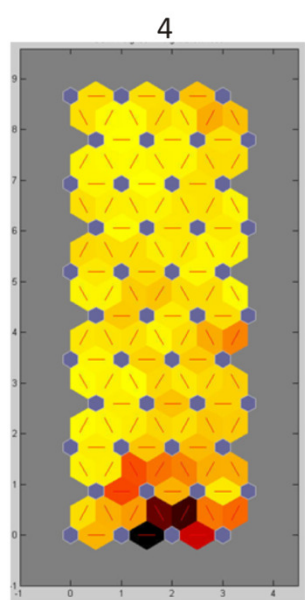
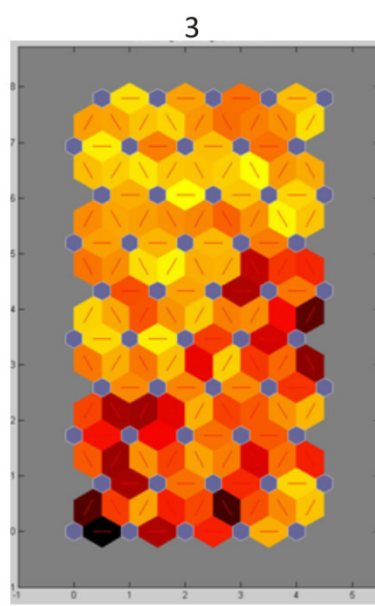
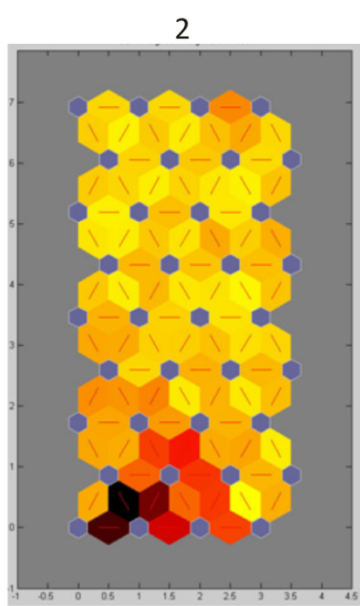
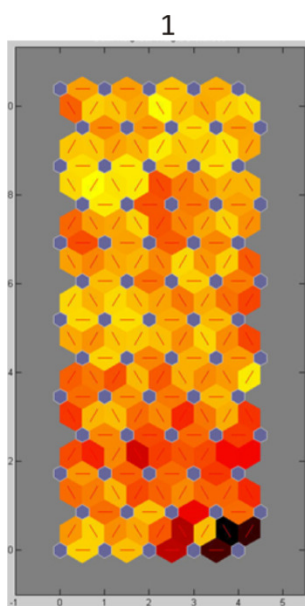
SCHULTZ



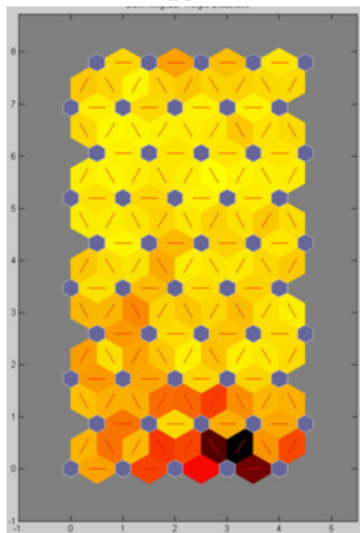
SELPH



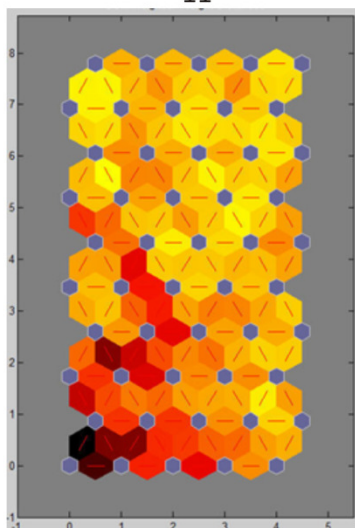




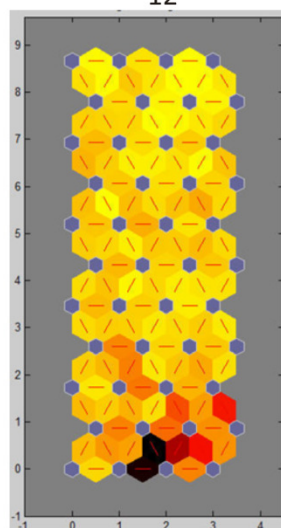
10



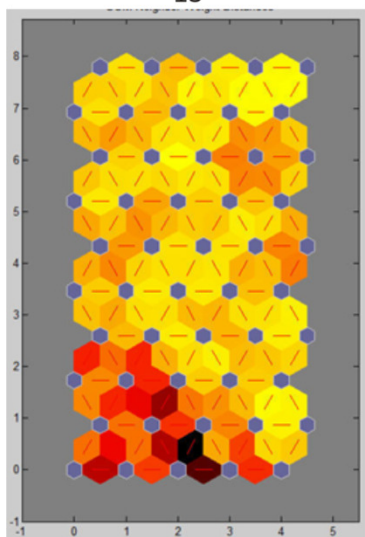
11



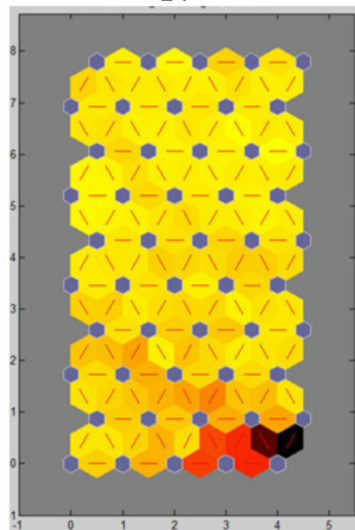
12



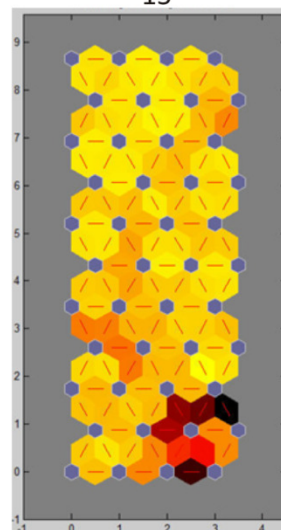
13



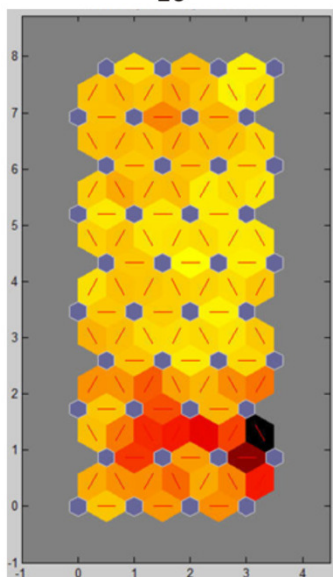
14



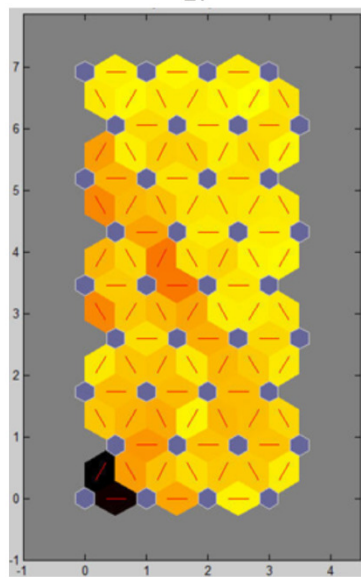
15



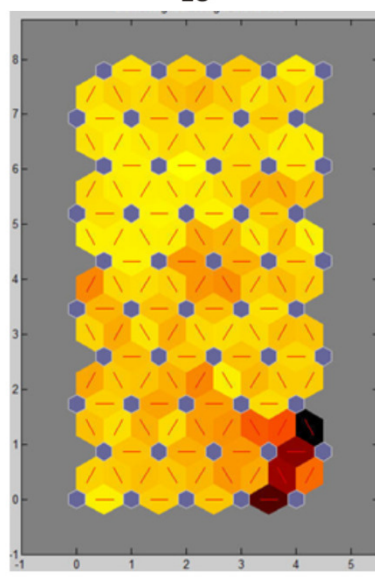
16

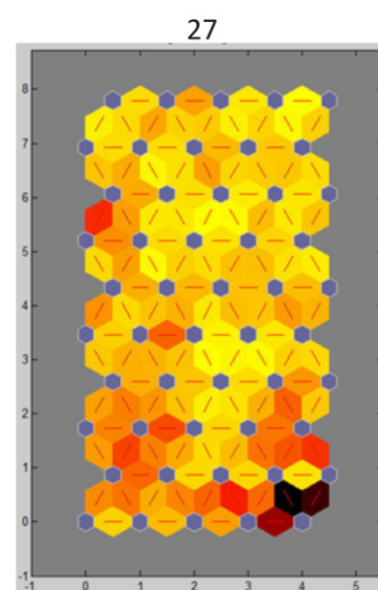
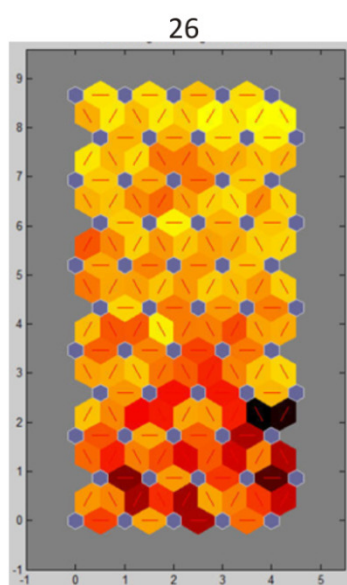
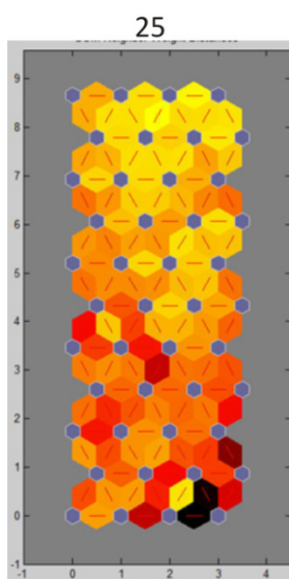
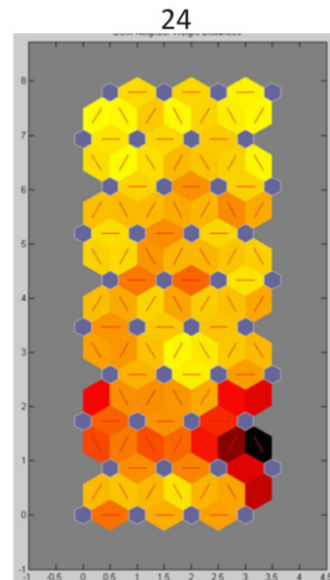
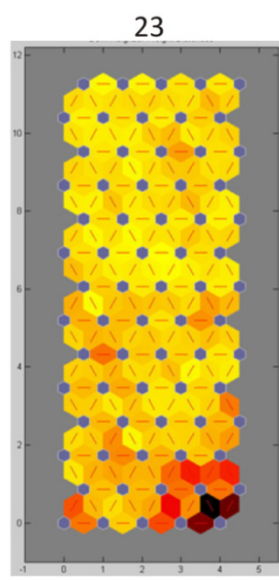
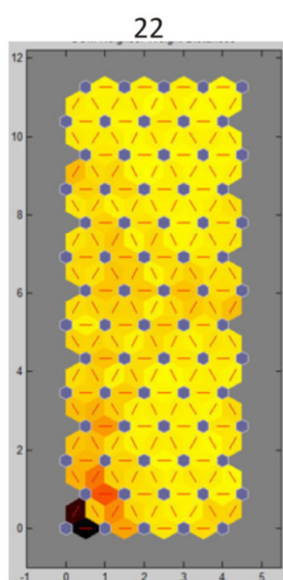
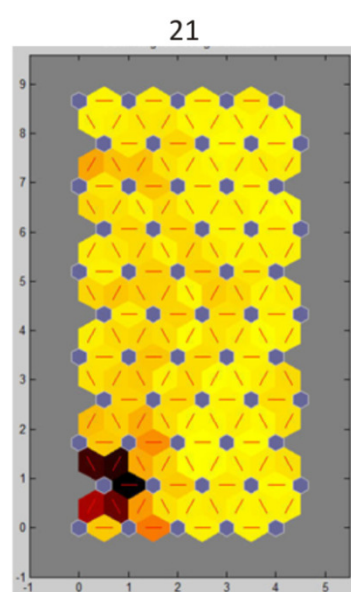
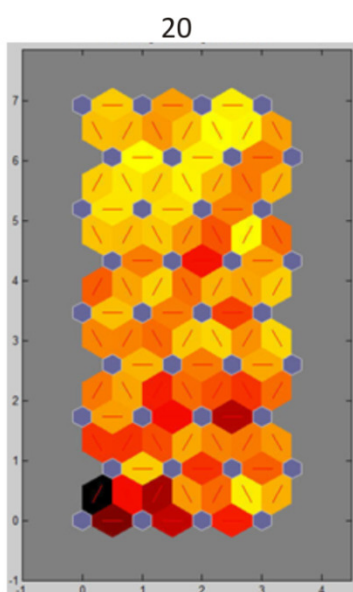
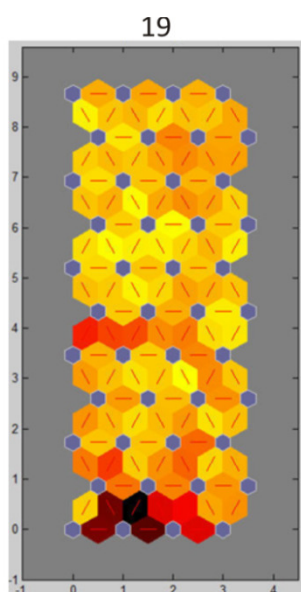


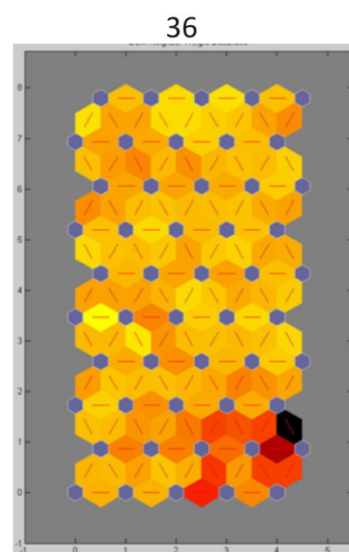
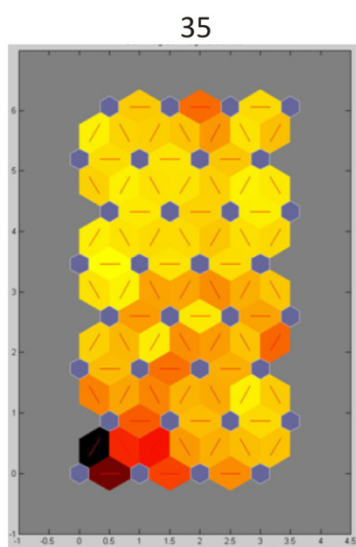
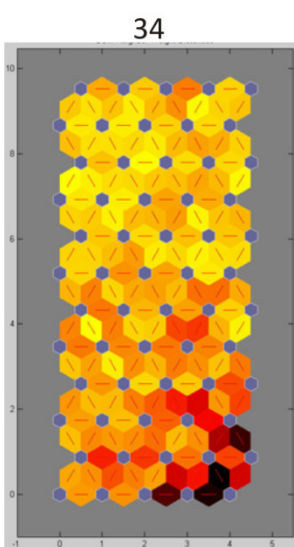
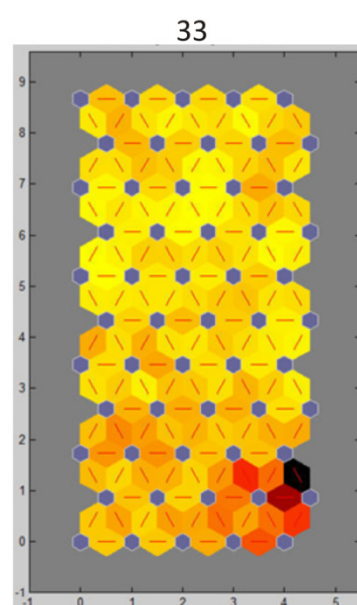
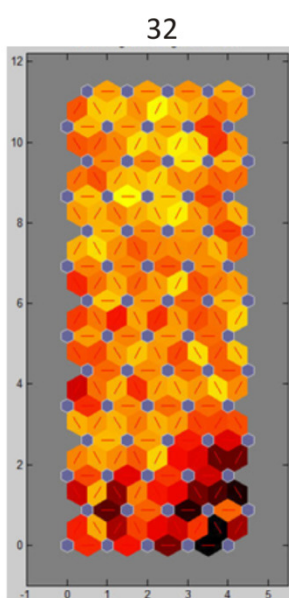
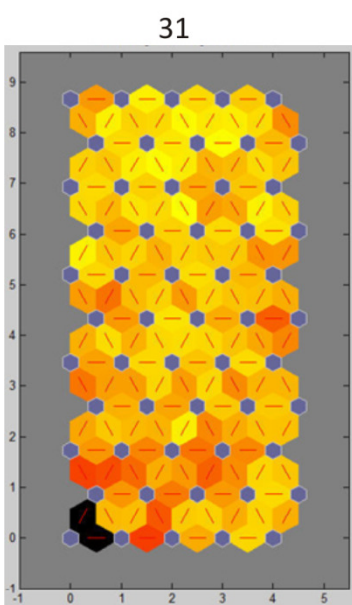
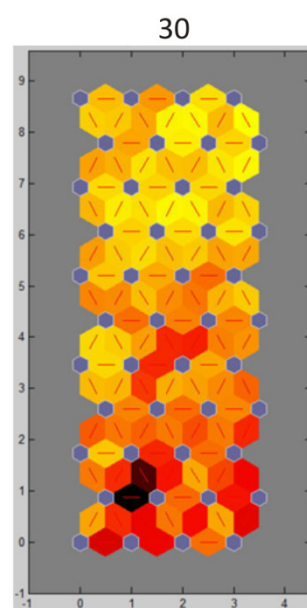
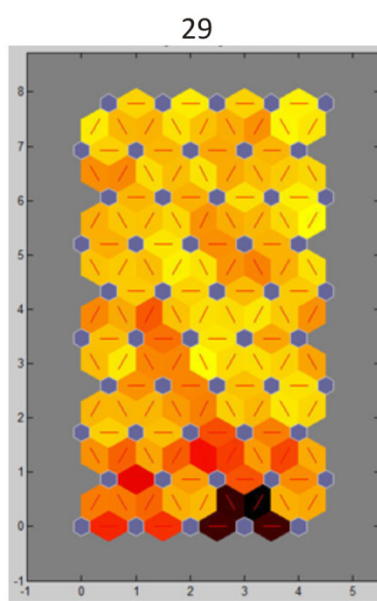
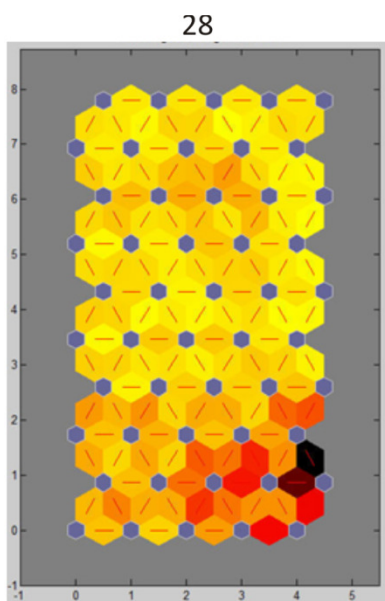
17



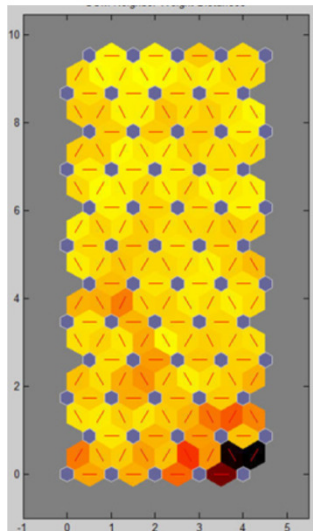
18



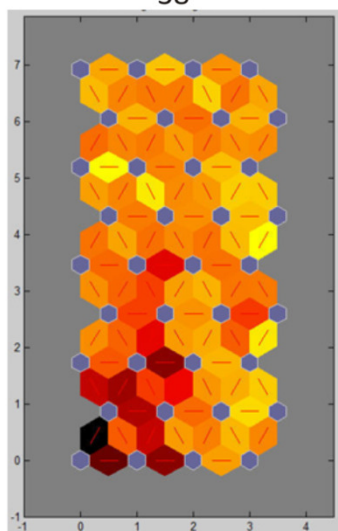




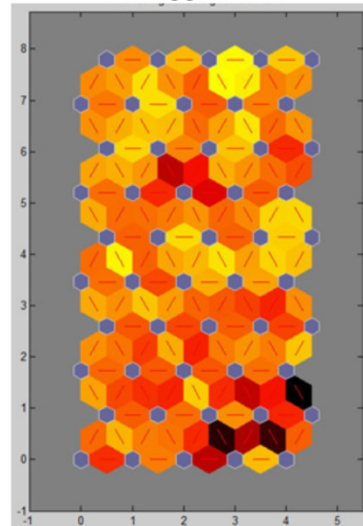
37



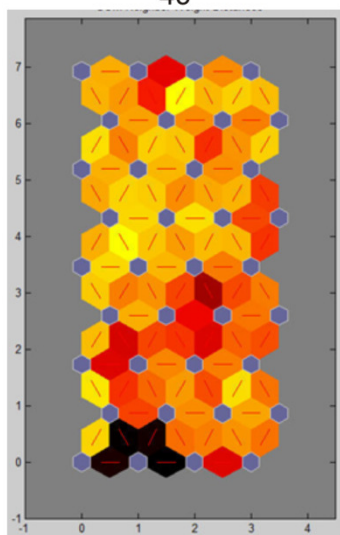
38



39



40



## ÖZGEÇMİŞ

### KİŞİSEL BİLGİLER

**Adı Soyadı** : Alper Şen  
**Doğum Tarihi ve Yeri** : 20.04.1979 - Konya  
**Yabancı Dili** : İngilizce  
**E-posta** : alpersen@yildiz.edu.tr

### ÖĞRENİM DURUMU

| Derece    | Alan                   | Okul/Üniversite              | Mezuniyet Yılı |
|-----------|------------------------|------------------------------|----------------|
| Y. Lisans | Uzaktan Alg. ve CBS    | Yıldız Teknik Üniversitesi   | 2007           |
| Lisans    | Harita Mühendisliği B. | Yıldız Teknik Üniversitesi   | 2004           |
| Lise      | Fen - Matematik        | Konya Gazi Lisesi            | 1997           |
| Ortaokul  | -                      | Konya Karatay Anadolu Lisesi | 1994           |

### İŞ TECRÜBESİ

| Yıl          | Firma/Kurum                | Görevi              |
|--------------|----------------------------|---------------------|
| 2009 - Devam | Yıldız Teknik Üniversitesi | Araştırma Görevlisi |
| 2005 - 2008  | STFA Harita A.Ş.           | Harita Mühendisi    |
| 2004         | Nizamoğlu Mühendislik      | Harita Mühendisi    |

## **YAYINLARI**

### **Makale**

1. Sen, A., Gumusay, M.U., Kavas, A. ve Bulucu, U., (2008). "Programming an Artificial Neural Network Tool for Spatial Interpolation in GIS - A Case Study for Indoor Radio Wave Propagation of WLAN", Sensors, 8(9): 5996-6014.

### **Bildiri**

1. Şen, A. ve Gökgöz, T., (2013). "Kümelemede Normalleştirmenin Etkisi", TMMOB Harita ve Kadastro Mühendisleri Odası, 14. Türkiye Harita Bilimsel ve Teknik Kurultayı, 14-17 Mayıs, Ankara.
2. Şen, A. ve Gökgöz, T., (2012). "Hidrografik Genelleştirmede Seçme/Eleme İşlemi için Yapay Sinir Ağı Yöntemlerinin Uygulanabilirliği", YTÜ İnşaat Fakültesi Bilim Şenliği, 21-22 Haziran, İstanbul.
3. Onden, I., Sen, C.G. ve Sen, A., (2012). "Integration of Integer Programming with GIS Analyzing Abilities for Determining the Convenience Levels of Retail Stores", World Conference on Business, Economics and Management, 4-6 Mayıs, Antalya.
4. Sen, A. ve Gokgoz, T., (2012). "Clustering Approaches for Hydrographic Generalization", GIS Ostrava 2012, 23-25 Ocak, Ostrava, Çek Cumhuriyeti.
5. Sen, A. ve Gokgoz, T., (2011). "Hydrographic Selection/Elimination in Automated Generalization by Using Artificial Neural Networks", 25th International Cartographic Conference, 3-8 Temmuz, Paris, Fransa.
6. Sen,A., Onden, I., Gokgoz, T. ve Sen, C., (2011). "A GIS Approach to Fire Station Location Selection", Geoinformation for Disaster Management (Gi4DM 2011), 3-8 Mayıs, Antalya.
7. Gumusay, M.U., Sen, A., Bulucu, U., Kavas, A., (2007). "Electromagnetic Coverage Calculation in GIS", 5th International Symposium on Mobile Mapping Technology (MMT'07), 28-31 Mayıs, Padua, İtalya.