

**Entropi, Yapay Sinir Ağları ve Geçiş Teoremlerinin Akım Gözlem
Ağının ve Su Kaynakları Sistemlerinin Planlanmasında Birlikte
Kullanılması ve Kızılırmak Havzasına Uygulanması**

Proje No: 102I068

**Prof.Dr. A. Ünal ŞORMAN
Nermin ŞARLAK**

Haziran 2008
ANKARA

Önsöz

Bu çalışma Tübitak kurumunca desteklenen 102I068 no'lu proje kapsamında hazırlanmıştır. Ülkemizde mevcut yağış ve/veya akım ölçüm ağlarının performansının ve elde edilen verilerin niteliğinin, ciddi bir şekilde gözden geçirilip değerlendirilmesi gerekmektedir. Bunun için bilgi kuramında tanımlı Entropi metodundan faydalanılacaktır. Bu metot, gözlem ağında varolan akım gözlem istasyonlarını taşıdıkları bilgi miktarına göre istasyon kapatma kararı vermede kullanılabilirdiği gibi aynı zamanda yeni örnekleme istasyonlarının sayısını ve yerini belirlemede de kullanılabilir. Literatürde genel olarak bu metod normal ve log-normal dağılım için geliştirilmiş olmasına karşın, bu çalışmada dağılımın normal dağılım olmaması durumundaki Entropi metodu üzerinde durulacaktır. Böylece seçilen dağılım tipinin Entropi metodu üzerindeki etkisini gözlemlene fırsatı bulunabilecektir.

Planlama ve karar vermede kullanılan metotlardan sağlıklı sonuçlar alınabilmesi için elde uzun süreli verilere ihtiyaç duyulmaktadır. Bu amaçla, eldeki kısa süreli verilerin zaman serisi analizinde mevcut olan teknikler kullanılarak uzatılması zorunlu hale gelmektedir. Sentetik akım serisi üretmede başka bir deyişle hidrolojik zaman serilerini modellemede pek çok stokastik model bulunmaktadır. Bu modeller pek çok varsayım üzerine kurulmuşlardır. Bunların başında da değişkenler arasında lineer ilişki olması gerekliliği gelmektedir. Hidrolojik veriler için bu varsayım çoğunlukla geçerli değildir. Bu da bizi yapay sinir ağları gibi doğrusal olmayan (non-linear) modellerini kullanmaya teşvik etmektedir. Bu çalışmada sentetik aylık akım verilerinin yapay sinir ağları metodu için kurulacak olan istatistiksel bir model ile türetilip bu iki modelden elde edilecek sentetik verilerin kıyaslanması amaçlanmaktadır.

Ölçüm ağında ölçüm yapılması gerekli olan istasyonlara entropi metodu ile karar verdikten ve bu istasyonlara ait veriler yapay sinir ağları modeli geliştirilerek uzatıldıktan sonra geçiş metodunun uygulanması amaçlanmaktadır. Bu metot belirli bir kesme seviyesinin üstündeki(taşkın) veya altındaki(kuraklık) değerleri baz almaktadır. Uygulanacak olan geçiş eoremi metodolojisi ile belirli bir kesme seviyesinin üzerinde veya altında kalan değerlerin sayısının, süresinin, hacminin ve ardışık durumlar arasındaki sürelerin beklenen değerlerini bulabilmek mümkün olabilmektedir. Bu parametreler de su kaynaklarının planlanmasında oldukça önemli parametreler olduğundan bundan sonra yapılacak olan su kaynağı sistemlerinin planlanmasında yararlı olacağı aşikardır.

Yukarıda kısaca bahsedilen amaçların Kızılırmak havzasındaki akım gözlem istasyonlarına uygulanması planlanmaktadır. Bilindiği üzere Kızılırmak nehri Türkiye'nin en uzun nehirlerinden biridir ve bu özelliği itibari ile bünyesinde pek çok su kaynağı yapısı barındırmaktadır.

Sonuç olarak bu çalışmada, yukarıda bahsi geçen üç ayrı metodun birbirleri ile ilişkilendirilerek Kızılırmak havzasındaki akım gözlem istasyonu verilerine uygulanması amaçlanmaktadır.

İÇİNDEKİLER

TABLO LİSTESİ	iv
ŞEKİL LİSTESİ	v
ÖZET	vi
ABSTRACT	vii
1. GİRİŞ	1
2. ENTROPİ METODU	2
2.1 Yöntem	2
2.1.1. Çoklu değişken durumunda Entropi kavramı	2
2.1.2 Ölçüm ağı tasarım problemlerine Entropi metodunun uygulanması	3
3. YAPAY SİNİR AĞLARI	6
4. STOKASTİK SÜREÇLER İÇİN GEÇİŞ TEOREMİ (CROSSING THEORY)	8
5. KIZILIRMAK HAVZASINA UYGULAMA	10
5.1 Entropi Metodunun Uygulanması	12
5.1.1 Harmancıoğlu Metodunun Uygulanması	12
5.1.2 Markus Metodunun Uygulanması	13
5.1.2.1 Markus metodunun uygulaması (Normal ve Log-Normal Dağılımları için):	13
5.1.2.2 Markus metodunun uygulaması (Gamma Dağılımı için):	14
5.2 Yapay Sinir Ağları Modelinin Kızılırmak Havzasına Uygulanması	15
5.3 Geçiş Teoreminin Uygulanması	16
6. ÖZET VE SONUÇLAR	19
7. KAYNAKLAR	21
EK A: NORMAL DAĞILIM İÇİN GEÇİŞ TEOREMİ	23

TABLO LİSTESİ

Tablo

5.1. Normal Dağılıma göre İstasyonların Seçilmesi	12
5.2. Log-normal Dağılıma göre İstasyonların Seçilmesi	12
5.3. Normal Dağılım İçin Aktarılan Bilgi, S(i), Alınan Bilgi, R(i) ve Net Bilgi, N(i) Değerleri	13
5.4. Lognormal Dağılım İçin Aktarılan Bilgi, S(i), Alınan Bilgi, R(i) ve Net Bilgi, N(i) Değerleri	14
5.5. Gamma Dağılımı İçin Aktarılan Bilgi, S(i), Alınan Bilgi, R(i) ve Net Bilgi, N(i) Değerleri	14
5.6. EIE 1501 İstasyonu Yıllık Akım Verilerinin İstatistiksel Özeti	16
6.1. Farklı Dağılımlarda Harmancıoğlu ve Markus Metotları için İstasyon Sıralamaları	20

ŞEKİL LİSTESİ

Şekil	
3.1. Üç birimli bir öğrenme ağına sahip yapay sinir ağı	6
3.2 Bir birim ve onun fonksiyonları	8
4.1 Sürekli rasgele sürecin geçiş istatistikleri	9
5.1 Kızılırmak havzası ve akım gözlem istasyonlarının yerleri	11
5.2.Normal ve lognormal dağılımlar için minimum tekrar eden bilgiye dayalı istasyon sıralaması	13
5.3.Normal, log-normal ve gamma dağılımları için ölçüm ağındaki istasyonların net bilgi miktarlarına göre sıralanması	15
5.4 EIE 1501 istasyonunun ANN modeli kullanılarak modelden elde edilen ve gözlemlenmiş aylık veri seti	16
5.5. EIE 1501 standardize akım verilerinin histogramı	17
5.6 EIE 1501 istasyonu için ölçülmüş ve modelle elde edilmiş beklenen geçiş sayıları; $E[N]$	17
5.7 EIE 1501 istasyonu verileri için belli kesme seviyelerindeki (h) yukarı geçiş noktalarının ölçülmüş ve teorik beklenen süresi, $E[l_h^+]$	18
5.8 EIE 1501 istasyonu verileri için belli kesme seviyelerindeki (h) ardışık yukarı geçiş noktalarının ölçülmüş ve teorik beklenen süresi, $E[B_h^+]$	18
5.9 EIE 1501 istasyonu verileri için belli kesme seviyelerinin üzerindeki ölçülmüş ve teorik toplam logaritmik hacim miktarları, $E[A_h^+(T)]$	19

ÖZET

Su kaynakları planlanmasında büyük önemi olan hidrolojik verilerin toplanarak değerlendirilmesine olanak sağlayacak nesnel ölçütün, bilgi kuramında tanımlanan “entropi kavramı” olabileceği çeşitli araştırmacılarca öne sürülmektedir. Entropi metodu, hidrolojik verilerin ölçülmesinde kullanılan akım gözlem istasyonlarının taşıdıkları bilgi miktarına göre istasyon kapatma kararı vermede kullanılabilirdiği gibi aynı zamanda yeni örnekleme istasyonlarının sayısını ve yerini belirlemede de kullanılabilir. Bu çalışmada, Kızılırmak Havzasında yer alan akım gözlem istasyonlarına iki farklı entropi metodu uygulanarak ölçüme devam edilmesi gereken istasyonlara karar verilmesi planlanmıştır. Ayrıca dağılım tiplerinin entropi metodu üzerindeki etkileri de araştırılmıştır. Serilere en uygun dağılım tipi olarak log-normal dağılım seçilmiştir. Bu varsayım altında EIE 1501, EIE 1541 ve EIE 1536 ölçüme devam etmesi gerekli istasyonlar olarak belirlenmiştir.

Kullanılan modellerden elde edilen serilerin güvenilir olması hem ekonomi hem de verimlilik açısından oldukça önemlidir. Bugüne kadar yapılmış olan zaman serisi modellerinde normal dağılıma uyma varsayımı kabul edilmiştir. Bu çalışmada, son yıllarda pek çok çalışmada geçerlilik bulan yapay sinir ağları (YSA) gibi doğrusal olmayan bir modelin aynı havzada yer alan akım gözlem istasyon verileri için kurulması planlanmıştır. Yapılan çalışmada YSA modelinin Kızılırmak havzasında yer alan istasyonların sentetik seri türetilmesinde oldukça başarılı olduğu sonucuna varılmıştır.

Yukarıda açıklanan süreçlerden sonra geçiş (crossing) metodunun uygulanması amaçlanmıştır. Su kaynakları sistemleri (özellikle barajlar), suyun fazla geldiği zamanlarda biriktirme yapmak, eksik olduğu zamanlarda da haznesinde biriktirdiği suyla ihtiyacı karşılamak amaçlı yapılmaktadır. Bu maksatla, taşkın veya kuraklık süreleri, hacimleri ve ardışık pikler arasındaki süreler su kaynakları sistemlerini planlama aşamasında önem kazanacağından aynı akım gözlem istasyonu verilerine yukarıda belirtilen değerlerin bulunmasını sağlayan geçiş modelinin uygulaması yapılmıştır.

Anahtar Sözcükler : Entropi, Yapay sinir ağları, Geçiş teoremi

ABSTRACT

The “entropy concept” defined in the information theory by several researchers is claimed to be the objective criterion which makes possible to obtain and assess the hydrologic data which has great importance in water resources engineering. Entropy method may possibly be employed for both discontinuing the water monitoring stations according to the information contained and determining the location and number of the streamflow measuring stations. In the current study, it is aimed to determine which stations should be continued to measure the runoff in the Kizilirmak Basin. Furthermore, the effect of the distribution type on Entropy method is also investigated in this study. The log-normal distribution is selected as an appropriate distribution for each series. Therefore, EIE 1501, EIE 1541 and EIE 1536 stations are obtained to necessary stations to continue measure under this assumption.

The reliability of the synthetic series obtained from the models used has great importance for both economy and efficiency. The time series model, which has been used up to now, has accepted various assumptions. One of them is normality assumption. In this study, a non-linear model such as artificial neural network (ANN) which are valid for most of the studies are planned to be used for the data collection from the existing stations. In this study, it is concluded that ANN model is very successfully to derive synthetic series which occur in Kizilirmak Basin.

The crossing method is aimed to be used after the processes previously explained. Water resources systems especially dams are constructed to store water if it is more than demand and use the stored water when it is less. Therefore flood and drought periods, volume and the time between the sequent peaks will gain importance. As a consequence of the referred the data used to obtain these parameters by means of a crossing theory.

Keywords: Entropy, Artificial neural network, Crossing theory

1.GİRİŞ

Toplanan verilerin herbiri doğal sistemden gelen sinyalleri gösteren rasgele bir sayı dizisi niteliği taşımaktadır. Planlamacının görevi ise, bu diziden azami ölçüde bilgi edinmek, su kaynakları sistemlerinin tasarımı ve işletimi için gerek duyduğu parametreleri belirlemektir. Hidrolojik süreç hakkında gözlemlenmiş çok sayıda veri olmasına karşın verilerin getirdiği bilginin az olması veri fazlalığına sebep olmaktadır. Bunun sonucunda ölçüm yapan kuruluşları (DSİ, EİE, DMI, Köy Hizmetleri vb.) ölçümlere devam etmenin gerekliliği ve elde edilen verilerin yararlılığı konusunda tereddüte düşmektedir. Ölçüm maliyetinin yüksekliğinin yanı sıra bu tür problemlerinde ortaya çıkması, kuruluşların ölçüm istasyonlarının kapatılması kararını vermeye dahi götürebilmektedir(Özkuş,S.,1996).

Şimdiye kadar yapılmış olan çalışmalar göstermiştir ki, entropi metodu mevcut akım gözlem ağlarında iyi sonuçlar vermektedir. Bu çalışmalara göre entropi metodu var olan gözlem yerlerini kaldırma, gözlemlerin örnek alma aralığının azaltılması gibi kararları almada kullanılabilirdiği gibi aynı zamanda yeni örnekleme istasyonlarının sayısını ve yerini seçmede de kullanılabilir. Çoklu istasyonlar için Entropi yöntemi literatürde normal, log-normal dağılımlar için uygulanmıştır. Bu çalışmada, entropi yönteminin çoklu istasyonlar için diğer dağılımlara uygulanması amaçlanmasına rağmen multivariate dağılım fonksiyonlarının bulunmasında karşılaşılan zorluk neticesinde sadece gamma dağılımına uygulanabilmiştir. Farklı dağılımlara uygulanamadığından dolayı, çalışmada farklı dağılımlar uygulandığında iki farklı entropi yönteminden elde edilen sonuçlar kıyaslanmıştır. İki farklı entropi yönteminden biri Harmancıoğlu (1981) tarafından, diğeri Markus (2003) tarafından geliştirilmiştir.

Planlama ve karar vermede kullanılan metodlardan sağlıklı sonuçlar alınabilmesi için elde uzun süreli verilere ihtiyaç duyulmaktadır. Bu amaçla, eldeki kısa süreli verilerin zaman serisi analizinde mevcut olan teknikler kullanılarak uzatılması zorunlu hale gelmektedir. Sentetik akım serisi üretmede başka bir deyişle hidrolojik zaman serilerini modellemede pek çok stokastik model bulunmaktadır. Bu modeller pek çok varsayım üzerine kurulmuşlardır. Bunların başında da değişkenler arasında lineer ilişki olması gerekliliği gelmektedir. Hidrolojik veriler için bu varsayım çoğunlukla geçerli değildir. Bu da bizi yapay sinir ağları gibi doğrusal olmayan (non-linear) modellerini kullanmaya teşvik etmektedir. Bir diğer varsayım da serilerin normal dağılıma uyması gerekliliğidir. Hidrolojik veriler için bu varsayım da geçerli olmamaktadır. Bunun için ya seriler çeşitli dönüşüm (transformasyon) işlemleri ile normal dağılım haline getirilip mevcut zaman serisi modelleri kullanılır ya da bu çalışmada yapılması planlandığı gibi serilerin normal dağılıma uymaması durumunda geçerli olacak zaman serisi modelleri geliştirme yoluna gidilir. Bu çalışmada sentetik aylık akım verilerinin yapay sinir ağları metodu ve normal olmayan dağılımlar için kurulacak olan istatistiksel bir model ile türetilip bu iki modelden elde edilecek sentetik verilerin kıyaslanması amaçlanmaktadır.

Ölçüm ağında ölçüm yapılması gerekli olan istasyonlara entropi metodu ile karar verdikten ve bu istasyonlara ait veriler yapay sinir ağları modeli kurularak uzatıldıktan sonra crossing metodunun uygulanması amaçlanmaktadır. Bu metod belirli bir kesme seviyesinin üstündeki(taşkın) veya altındaki(kuraklık) değerleri baz almaktadır. Uygulanacak olan crossing teoremi metodolojisi ile belirli bir kesme seviyesinin üzerinde veya altında kalan değerlerin sayısının, süresinin, hacminin ve ardışık durumlar

arasındaki sürelerin beklenen değerlerini bulabilmek mümkün olabilmektedir. Bu parametreler de su kaynaklarının planlanmasında oldukça önemli parametreler olduğundan bundan sonra yapılacak olan su kaynağı sistemlerinin planlanmasında yararlı olacağı aşikardır.

Yukarıda kısaca bahsedilen amaçların Kızılırmak havzasındaki akım gözlem istasyonlarına uygulanması planlanmaktadır. Bilindiği üzere Kızılırmak nehri Türkiye'nin en uzun nehirlerinden biridir ve bu özelliği itibari ile bünyesinde pek çok su kaynağı yapısı barındırmaktadır. Bu amaçla Kızılırmak Havzasında bilgi alış verişi bakımından önemli olduğu düşünülen ve Elektrik İşleri Etüt İdaresi (EİE) teşkilatının ilgili birimi ile yapılan sözlü görüşmelerin sonucunda beş akım gözlem istasyon verisinin bu çalışmada kullanılması kararlaştırılmıştır.

Sonuç olarak bu çalışmada, yukarıda bahsi geçen üç ayrı metodun birbirleri ile ilişkilendirilerek Kızılırmak havzasındaki akım gözlem istasyonu verilerine uygulanması amaçlanmaktadır.

2. ENTROPİ METODU

Ülkemizde mevcut yağış ve/veya akım ölçüm ağlarının performansının ve elde edilen verilerin niteliğinin, ciddi bir şekilde gözden geçirilip değerlendirilmesi gerekmektedir. Bu nedenle bu çalışmanın amaçları olarak aşağıdaki hususlar ele alınacaktır;

- a) Bir gözlem ağının sağlayacağı bilgi, konuma bağlı olarak değiştiği gözönüne alınarak gözlem ağı tasarımının gerçekleştirilmesi
- b) Bu değerlendirmeyi yapabilmek için yapılacak gözlemin getireceği bilgiyi ölçülebilir bir büyüklük ile ifade edilmesi.

Bunun için bilgi kuramında tanımlı entropi metodundan faydalanılacaktır. Bu metod, gözlem ağında varolan akım gözlem istasyonlarını taşıdıkları bilgi miktarına göre istasyon kapatma kararı vermede kullanılabildiği gibi aynı zamanda yeni örnekleme istasyonlarının sayısını ve yerini belirlemede de kullanılabilmektedir. Literatürde genel olarak bu metod normal ve log-normal dağılım için geliştirilmiştir.

2.1. Yöntem

Bu çalışmada Harmancıoğlu tarafından geliştirilen çoklu değişkenli entropi yöntemi üzerinde durulacaktır. Tek ve iki değişken olması durumundaki entropi kavramlarının tanımını literatürde bulmak mümkündür.

2.1.1 Çoklu değişken durumunda Entropi kavramı

İki değişken arasında stokastik bağımlılık olması durumunda marjinal ve toplam entropi değerlerinde azalma olmaktadır. Bağımlı değişkenlerin sayısı arttıkça bu durum geçerliliğini korumaktadır.

Entropi kavramlarında ki herhangi bir ifadeyi bulabilmek için analiz edilen verileri en iyi ifade ettiği düşünülen olasılık dağılım tipinin belirlenmesi gerekir. Eğer dağılım çoklu değişkenli (multivariate) doğrusal olmayan normal dağılım olarak belirlenmiş ise X olayının ortak (joint) entropisi aşağıdaki gibi ifade edilebilir;

$$H(X) = (M/2) \ln 2\pi + \ln |C| / 2 + M/2 - M \ln(\Delta x) \quad (2.1)$$

Buradaki M değişken sayısını ifade etmektedir ve logaritma e tabanında olduğu içinde entropi birimi napierdir. Δx kesikleme aralığını ifade etmekte olup 1 olarak alındığında ifadeden düşmektedir (Özkul,S.,1996). Eğer gözlenmiş değerlerin logaritması alınırsa denklem 2.1 lognormal dağılmış değişkenler içinde kullanılabilir. Kovaryans matrisi, değişkenlerin kendi içinde oluşturdukları, içsel bağımlılığın ölçütü olan otokovaryans değerleri ile değişkenlerin birbirleri ile olan bağımlılıklarının ölçütü olan çapraz kovaryans değerlerini içerir ve boyutu $M \times M$ dir;

$$C = \begin{bmatrix} C_{11} & C_{12} & \dots & C_{1M} \\ C_{21} & C_{22} & \dots & C_{2M} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ C_{M1} & C_{M2} & \dots & C_{MM} \end{bmatrix} \quad (2.2)$$

Öteleme miktarlarının da dahil edildiği kovaryans matrisi aşağıdaki gibi ifade edilebilir ve boyutu $[M \times (K+1)] [M \times (K+1)]$ dir:

$$C = \begin{bmatrix} C_{11}(0) & C_{11}(K) & \dots & \dots & C_{1M}(0) & C_{1M}(-K) \\ C_{11}(K) & C_{11}(0) & \dots & \dots & C_{1M}(K) & C_{1M}(0) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ C_{M1}(0) & C_{M1}(-K) & \dots & \dots & C_{MM}(0) & C_{MM}(K) \\ C_{M1}(K) & C_{M1}(0) & \dots & \dots & C_{MM}(K) & C_{MM}(0) \end{bmatrix} \quad (2.3)$$

Kovaryans matrisi (C), tanımı itibari ile pozitif, simetrik ve kare bir matristir. Eldeki değişkenin süresinin yeterince uzun olmadığı durumlarda, entropi değerlerinde negatif ifadeler rastlanması mümkündür. Kovaryans matrisinin determinantının çok küçük olması buna sebep olmaktadır ve bu durum da çıkan sonuçları yorumlamayı zorlaştırmaktadır.

2.1.2 Ölçüm ağı tasarım problemlerine iki farklı Entropi metodunun uygulanması

Harmancıoğlu Metodu:

Bu metodun amacı, gözlem istasyonlarının yeri ve sayısının uygun bir şekilde seçilmesiyle tekrar eden bilgi miktarını minimuma indirmektir. Bu amaçla, en önemli (en belirsiz) değişkenin olduğu gözleme noktasının konumunun önceden bilinmesi gerekir. Bunun için var olan gözlem istasyonları belirsizliği veya bilgisizliği azaltma derecesine göre sınıflandırılabilir. Bu sınıflandırma listesindeki ilk istasyon, değişken hakkında belirsizliğin en fazla olduğu yerdir. Onu takip eden diğer ölçüm noktaları belirsizliği azaltmayı sağlamaktadır. Sıralama listesine en son giren istasyon ise en az bilgi getiren istasyondur.

En iyi istasyon sayısı ve konumunu seçmek için aşağıdaki yöntem sırası ile takip edilmelidir:

- a) Gözönüne alınan ölçüm ağındaki her bir istasyonun veri dizisi, X_m olarak ifade edilir. Ölçüm ağında M adet istasyon olduğu farzedilirse X_m ifadesindeki m değeri ($m=1, \dots, M$) istasyonların sıra numarasını gösterir.
- b) Her bir istasyonda ölçülen akım verilerinin olasılık yoğunluk fonksiyonları belirlenir. Çarpık dağılımların çoklu değişkenli olasılık yoğunluk fonksiyonlarını tarif etmek güç olduğu için, normal ve log-normal dağılım seçilmiş olması durumlarında geçerli olan toplam entropi $H(X_1, \dots, X_m)$ veya bağımlı değişkenlerin toplam belirsizlik miktarları hesaplanır.
- c) Her bir istasyondaki değişkenin marjinal entropisi (2.1) denkleminde M yerine 1 yazılarak hesaplanır. En yüksek marjinal entropi değerini veren istasyon birinci öncelikli istasyon X_1 olarak seçilir. İstasyonun konumu itibari ile en fazla belirsizlik o istasyonda oluşmaktadır. Bu da bu istasyonda ölçüm yapmayı zorunlu hale getirmektedir.
- d) Seçilen birinci öncelikli istasyon, ölçüm ağında bulunan diğer istasyonların her biri ile ikili olarak düşünülüp, aralarında en az tekrar eden bilgiye sahip olan istasyon ikinci öncelikli istasyon X_2 olarak seçilir. Genelde bu istasyon farklı bir kolda yer aldığı gibi, seçilen birinci öncelikli istasyona uzak bir yerde de yer alabilmektedir. Çünkü mesafe arttıkça istasyonlar arasındaki tekrar eden bilgi azalmaktadır. İkili istasyonlar için en az tekrar eden bilgi miktarı aşağıdaki gibi ifade edilir:

$$\min \{T(X_1, X_2)\} = \min \{H(X_1) - H(X_1/X_2)\} \quad (2.4)$$

buradaki şartlı entropi;

$$H(X_1/X_2) = H(X_1, X_2) - H(X_2) \quad (2.5)$$

denklem 2.5 denklem 2.4'de yerine yazılırsa tekrar eden bilgi ifadesi aşağıdaki hali alır:

$$T(X_1, X_2) = H(X_1) - H(X_1/X_2) \quad (2.6)$$

Bu işlemler, en az bilgi tekrarı ve en yüksek ortak entropiye (belirsizliğe) sahip istasyon çiftini belirlemeyi sağlar. Buna göre, belirlenmiş istasyonlar birlikte kullanıldığında en fazla bilgi miktarını üretmektedirler.

- e) Belirlenmiş olan ikili istasyonlar ile ölçüm ağındaki diğer istasyonların her biri için toplam entropi, şartlı entropi ve tekrar eden bilgi miktarları en az tekrar eden bilgi miktarına sahip üçlü yapıyı belirlemek için hesap edilir. Aşağıdaki durumu sağlayan istasyon üçüncü öncelikli istasyon olarak seçilir:

$$\min \{T(X_1, X_2, X_3)\} = \min \{H(X_1, X_2) - H(X_1, X_2/X_3)\} \quad (2.7)$$

- f) Aynı işlemler en az tekrar eden bilgi miktarına sahip $4, 5, \dots, j$. öncelikli istasyonları seçmek için tekrar edilir. En genel hali ile minimum tekrar eden bilgi aşağıdaki gibi ifade edilir:

$$\min \{ T(X_1, \dots, X_{j-1}, X_j) \} = \min \{ H(X_1, \dots, X_{j-1}) - H(X_1, \dots, X_{j-1} / X_j) \} \quad (2.8)$$

Buradaki X_1 birinci öncelikli istasyon ve X_j j. öncelikli istasyonu ifade etmektedir.

Markus Metodu:

Yang Y. ve arkadaşlarının 1994'te öne sürdükleri aktarılan bilgi miktarına benzer olarak, Markus ve arkadaşları 2003'te X değişkenindeki bilgi azalmasını, $R(X,Y)$ olarak aşağıdaki gibi tarif etmişlerdir:

$$R(X, Y) = \frac{T(X, Y)}{H(X)} \quad (2.9)$$

Bu değer, ayrıca Y biliniyorsa X'in belirsizliğindeki azalma veya Y'den X tarafından *alınan bilgi* olarak ifade edilebilir. Benzer şekilde X'den Y'ye *aktarılan bilgi* aşağıdaki ifade edilebilir:

$$S(X, Y) = \frac{T(X, Y)}{H(Y)} \quad (2.10)$$

İki değişken X ve Y arasındaki ilişkiyi (2.9) ve (2.10) denklemleri tarif etmektedir. Bu sebeple bu ifadeler çoklu istasyonlu gözlem ağı içinde uygulanabilmektedir. Denklem (2.9) ve (2.10) kullanılarak i istasyonun aldığı ve aktardığı bilgi aşağıdaki gibi ifade edilmiştir:

$$R(i) = R(X(i), \hat{X}(i)) \quad (2.11)$$

Buradaki $X(i)$; i istasyonundaki veri setini ifaderken $\hat{X}(i)$; $X(i)$ istasyonunun geri kalan diğer istasyonlarla çoklu lineer regresyon (multiple linear regression) denklemi ile bulunan veri setini ifade etmektedir:

$$\hat{X}(i) = a(i) + \sum_{j=1}^{M-1} b_j(i) * Y(j) \quad (2.12)$$

denklemdaki $Y(j)$; i istasyonundan geri kalan istasyonlardan elde edilen veri matrisini, $a(i)$ ve $b_j(i)$; i istasyonu ve geri kalan istasyonlar arasındaki regresyon parametrelerini, M; gözlem ağındaki istasyon sayısını ifade etmektedir. Bu ifadelerde farklı alanlardaki istasyon verileri arasında lineer veya lineere yakın bir ilişki olduğu varsayımı vardır. Benzer şekilde:

$$S(i) = S(X(i), \hat{X}(i)) \quad (2.13)$$

Bu çalışmada entropi kavramı en yüksek $S(i)$ ve $R(i)$ miktarlarına sahip istasyonları belirlemek için kullanılmıştır. $R(i)$ değeri diğer istasyonlara göre fazla miktarda olan istasyon en fazla bilgi alan istasyon olarak tanımlanmaktadır. Başka bir deyişle, fazla miktarda bilgi aktaran istasyon daha büyük $S(i)$ değerine sahip olduğundan daha önemlidir ve ölçüm ağında kapatılmaması gereken istasyondur. Sonuç olarak aktarılan net bilgi miktarı $N(i)$, $S(i)$ ve $R(i)$ arasındaki farktır:

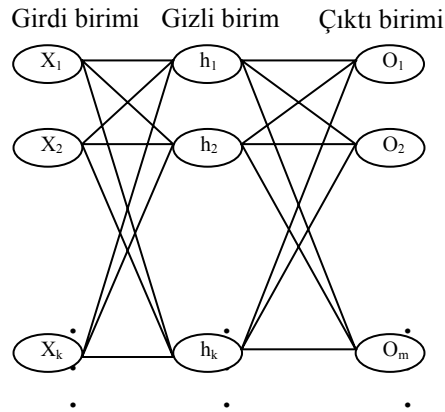
$$N(i) = S(i) - R(i) \quad (2.14)$$

Bu metotta positif $N(i)$ değerine sahip istasyon bölgesel analizde gözlem yapılması açısından önemli olan istasyonlar olarak göz önüne alınmaktadır. Eğer gözlem ağındaki istasyonların ekonomik veya başka sebeplerle kapatılması gerekse, positif $N(i)$ değerine sahip istasyonlar negatif $N(i)$ değerine sahip istasyonlara göre ölçüm ağında kalması gerekli istasyonlar olarak bu metodla belirlenebilmektedir (Markus, M ve ark., 2003).

3. YAPAY SİNİR AĞLARI

Bir yapay sinir ağında girdi, gizli ve çıktı birimleri olmak üzere üç farklı birim bulunmaktadır. Her birim bir çok nörondan oluşmakta olup birimler aralarında ağırlık kümeleri ile bağlanmaktadır. Bağlanma şekli ve her kısımdaki nöron sayısı deneme yanılma yolu ile bulunmaktadır. Aynı kısımdaki nöronlar arasında iletişim olmasına izin verilmemektedir. Nöronlar girdiyi başlangıç girdilerinden yada ara bağlantılardan almaktadırlar. Geriye doğru hata yayılması iki etaptan oluşmaktadır: Çıktı birimindeki çıktı bilgi sinyali hesaplamak için girdi nöronlarındaki verilen girdi bilgisini ileriye dağıtan etap ile çıktı birimindeki gözlemlenmiş ve hesaplanmış bilgi sinyalleri arasındaki farklara dayanarak bağlantı kuvvetlerinin değiştirilmesiyle yapılan geriye doğru dağıtım etabı (Raman, H. at all; 1995).

Bir eğitim sürecinin başında, bağlantı kuvvetlerine rasgele değerler verilir. Öğrenme algoritmasında her iterasyonda eğitim başarı ile tamamlanana kadar kuvvet değiştirilmektedir. İterasyon süreci bir sonuç elde ettiğinde bağlantı kuvvetleri, eğitim sürecinde kullanılan örneklerdeki mevcut bilgiyi elde eder ve saklar.



Şekil 3.1. Üç birimli bir öğrenme ağına sahip yapay sinir ağı

Yeni bir girdi grubu sunulduğunda, ileriye doğru besleme ile yapay sinir ağının bağlantı kuvvetlerindeki öğrenilmiş ve saklanan bilgi sayesinde bir çıktı grubu elde edilir. Bir girdi, bir gizli ve çıktı biriminden oluşan üç birimli bir öğrenme ağına sahip yapay sinir ağı Şekil 3.1'de gösterilmiştir.

Her birinin girdi biriminde, x_i , $i=1, \dots, k$ girdi değerleri, gizli birimde h_i , $i=1, \dots, k$ gizli değerleri, çıktı biriminde de O_n , $n=1, \dots, m$ çıktı değerleri kümesi bulunan toplam N adet girdi grubu bulunmaktadır. Gizli katmanının girdi değerlerinin seçilmesi YSA modeli için oldukça zor bir görevdir. Bugüne kadar girdi değerlerinin seçimi için teorik bir yöntem olmamasından dolayı girdi değerleri deneme yanılma yöntemi ile

belirlenmektedir. Girdi değerleri gizli birimdeki ilk ara bağlantı ağırlıkları, w_{ij} , $j=1, \dots, h$, ile çarpılmakta ve sonuçlar i indeksi boyunca toplanmakta ve gizli birimlerin girdileri olmaktadır (Raman and Sunilkumar, 1995). Ağırlıkların bulunması için adım, sigmoid, lineer gibi pek çok aktivasyon fonksiyonu vardır. Bunlar arasında sigmoid fonksiyonu en çok kullanılan bir yöntemdir. Girdi değerleri, x_i gizli katmandaki ara bağlantı ağırlıkları ile çarpılmakta w_{ij} , $j=1, \dots, k$ ve çarpım değerleri toplanarak girdi katmanının girdisini oluşturmaktadır:

$$H_j = \sum_{i=1}^k w_{ij} x_i \quad j=1, \dots, k \quad (3.1)$$

H_j , j . gizli birimdeki girdi değeri, w_{ij} , i . süreç elemanından j . süreç elemanına iletilen ağırlıktır. Her bir gizli birimdeki değerler sigmoid fonksiyonunu kullanarak gizli birimin çıktısını elde ederek çıktı birimin girdi değerini oluşturmaktadır. Matematiksel olarak ifadesi (Raman and Sunilkumar, 1995):

$$HO_j = f(H_j) = \frac{1}{1 + e^{-H_j}} \quad (3.2)$$

H_j , birimin girdisi, $f(H_j)$ birimin çıktısıdır. Çıktı, HO_j izleyen katmanın girdisini oluşturur ve süreç çıktı katmanına ulaşmaya kadar devam eder. Bu süreç ileri aktivasyon akımı olarak tanımlanır. Sigmoid fonksiyonu 0 ve 1 arasında sınırlı olduğu için, girdi verileri YSA modeli uygulanmadan önce standardize edilmelidir. Standardize işlemi aşağıdaki formülle yapılmıştır:

$$x_s = (x / 1.24 x_{max}) + 0.1 \quad (3.3)$$

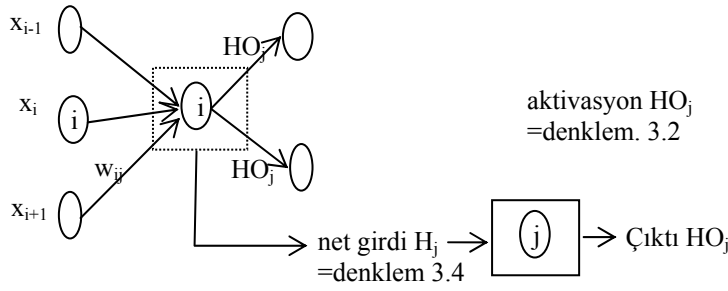
eşitlikteki x_s = standardize değişkenleri; ve x_{max} = her bir değişkenin maksimum değeridir. Verileri standardize ettikten sonra, YS ağı eğitim setindeki veriler kullanılarak eğitilir. Literatürde, eğitim seti örnek setinin %75'i seçilerek elde edilmektedir. Sonra öğrenme oranı ve momentum katsayısı belirlenmektedir.

Süreçte girdi sinyalleri işlenip alınarak diğer girdilere aşağıdaki formül kullanılarak çıktı sinyali olarak gönderilir:

$$IO_n = \sum_{i=1}^h w_{jn} HO_i \quad n=1, \dots, m \quad (3.4)$$

girdi verileri çıktı verileri, O_n elde edilene kadar öğrenme oranı, momentum katsayısı ve sigmoid fonksiyonu kullanılarak işlenir. Bir birim ve onun fonksiyonları Şekil 3.2'deki gibi resmedilebilmektedir.

Öğrenme süreci veya ardışık ağırlık uyarlamaları geriye yayımlı öğrenme algoritması ile tamamlanmaktadır. Çıktı katmanındaki O_n , hedef değeri T_n ile aynı olamamaktadır.



Şekil 3.2 Bir birim ve onun fonksiyonları

Geriye yayımlı öğrenme algoritması, öğrenme sürecinde çıktı katmanındaki çıktı değeri, O_n ve bulunması amaçlanan değer, T_n arasındaki farkların karelerinin ortalaması alınarak iteratif olarak minimize edilmesini amaçlamaktadır (Dikmen, 2001).

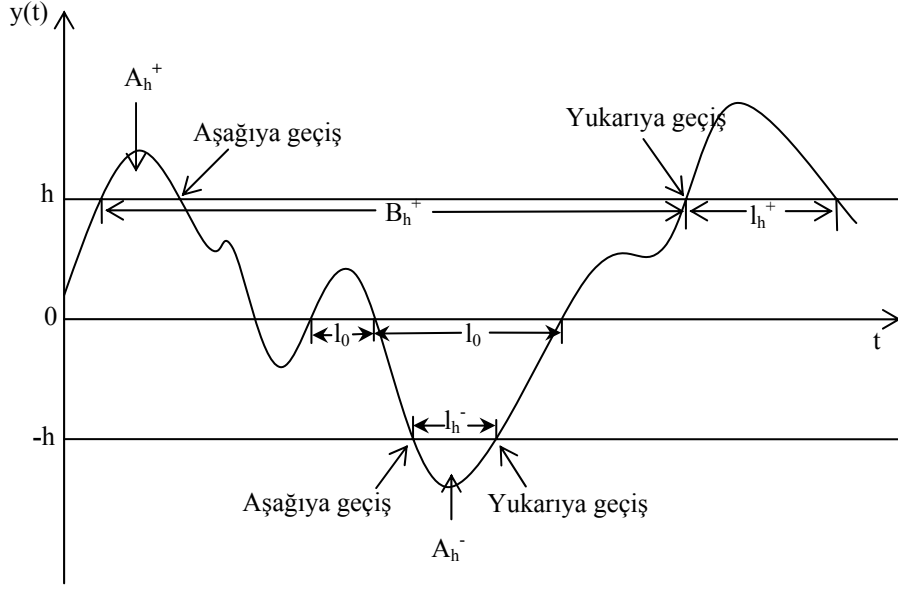
Test etme süreci sistemin nasıl öğrendiğini belirlemedeki tek yoldur. Bu sebepten, sistem performansını ölçmede ortalama hata karelerinin karekökü (RMSE) kullanılmıştır. RMSE aşağıdaki formül ile hesaplanmaktadır:

$$RMSE = \sqrt{\frac{1}{N} \sum_{n=1}^m (T_n - O_n)^2} \quad (3.5)$$

4. STOKASTİK SÜREÇLER İÇİN GEÇİŞ TEOREMİ (CROSSING THEORY)

Taşkın ve kuraklık toplumun sosyal, ekonomik ve politik faaliyetlerini etkileyen iki önemli hidrolojik olaydır. Bu sebepten, bu iki konu deneysel ve teorik olarak uzun zamandır pek çok araştırmacı tarafından incelenmektedir. Yapılan çalışmaların amacı, geçmiş deneyimlere dayanarak geleceğe yönelik tahmin yapabilen modeller geliştirmektir.

Bu çalışmalardan biri geçiş teoremidir. Bu teori Rice tarafından 1945’de geliştirilmiştir. Geçiş teoremi rasgele süreçteki değerlerin belli bir kesme seviyesinin altında veya üstünde kalması ile ilgilidir. Zaman serisinde ilgilenilen ve Şekil 5.1’ de gösterilen çoğu istatistiksel özellik geçiş teoremindeki basit tekniklerle tahmin edilebilmektedir.



Şekil 4.1 Sürekli rasgele sürecin geçiş istatistikleri

$y(t)$; $(0,t)$ aralığında standardize edilmiş gözlemlenmiş sürekli stokastik süreci (continuous-time stochastic process) ifade etmektedir (Nordin and Rosbjerg, 1970):

$$y(t) = (Q(t) - m) / \sigma \quad (4.1)$$

denklemdaki:

$Q(t)$: t zamanındaki akım değerini
 m : ortalama debiyi, ve
 σ : akımın standart sapmasını

ifade eder. Yukarıya geçiş (upcrossing), sürecin pozitif eğimle belli bir kesme seviyesini aşan noktaları gösterirken aşağıya geçiş (downcrossing) belli bir kesme seviyesini negatif eğimle geçen noktaları göstermektedir. Belli bir h seviyesi için hidrolojinin ilgilendiği diğer nicelikler (quantities):

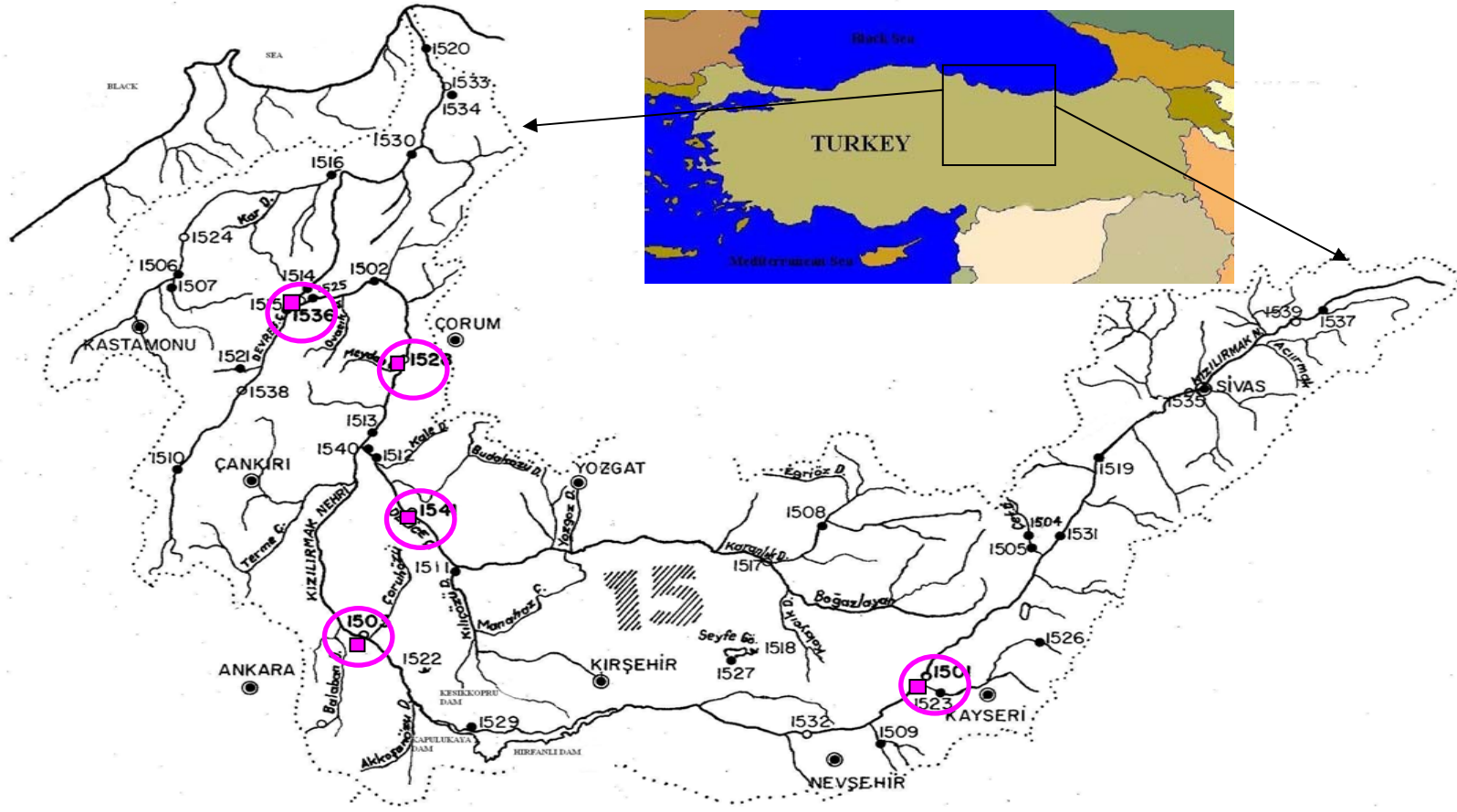
- 1) $y(t)$ sürecinde belli bir h seviyesini geçen noktalarının sayısı, N ;
- 2) belli bir h seviyesini aşan noktalar arasındaki zaman, l_h^+ ;
- 3) ortalama değerin altında ve üstünde kalan geçişler arasındaki zaman, l_0 ;
- 4) belli bir h seviyesini aşan ardışık noktalar arasında kalan zaman, B_h^+ ;
- 5) belli bir h seviyesinin üstünde kalan alan ise A_h^+ olarak ifade edilmiştir. Benzer olarak N, l_h^- ve A_h^- nitelikleri h seviyesinin altında kalan noktalar için de tanımlamak mümkündür (Şekil 5.1'de görüldüğü gibi).

$y(t)$ süreci akım değerini temsil ettiğinden N ; $(0,t)$ aralığındaki taşkın sayısını, l_h^+ ve B_h^+ taşkın süresi ve tekrür aralığını ve A_h^+ depolama kapasitesini aşan fazla hacmi ifade etmektedir (Desmond and Guy, 1991). Kuraklık için gerekli olan nicelikler ise; $(l_h^-$ ve $A_h^-)$ olarak ifade edilebilir. Akımın azlığı durumunda gerekli olan su miktarının

depodan, yeraltı suyundan veya diğer kaynaklardan karşılanması gerektiğinde bu parametrelerin bilinmesi önem arz etmektedir. Normal ve log-normal dağılım için geçiş teoremi formülleri literatürde mevcuttur ve Ek A' da sunulmuştur (Desmond and Guy, 1991).

5. KIZILIRMAK HAVZASINA UYGULAMA

Yukarıda bahsedilen metotları uygulamak için Kızılırmak Havzasında Kızılırmak nehri üzerinde uzun yıllar gözlem değerine sahip akım gözlem istasyonları seçildi. Seçilen istasyonlar Elektrik İşleri Etüd İdaresi tarafında işletilmekte olup 1955 yılından itibaren gözlenmektedirler. Bu istasyonlar: EIE 1501, EIE 1503, EIE 1541, EIE 1528 ve EIE 1536 olup 41 yıl ortak gözlem değerine sahiptir (bkz. Şekil 5.1). Seçilen istasyonlardan EIE1503, EIE 1528 ve EIE 1536 istasyonları Hirfanli, Kesikkörü and Kapulukaya barajlarının mansabında yer aldıklarından bu istasyonların akım değerlerinin doğal akım haline dönüştürülmesi gerekmektedir. EIE 1501 istasyonu barajların membasında yer aldığından bahsi geçen ve barajların mansabında yer alan istasyonların doğal akım değerleri EIE 1501 istasyon verileri kullanılarak bulunmuştur. Bunun için baraj yapılmadan önceki istasyon akım verileri arasındaki mevsimsel korelasyon denklemleri kullanılarak baraj yapıldıktan sonraki akım değerleri doğal hale getirilmeye çalışılmıştır. Bulduğumuz doğal akımların doğru olup olmadığını kontrol etmek için DSI'den barajlara giren çıkan akımlar alınarak barajlardaki depolama miktarları bulunmuştur. Doğal akım değerlerinden ölçülmüş akım değerleri çıkarılarak hesaplamalara göre beklenen depolama miktarları ile DSI'den alınan depolama miktarları arasındaki korelasyon incelenmiştir. Bu iki değişken arasında önemli bir korelasyon değeri ($R^2 = 0.85$) elde edildiğinden doğal hale getirilmiş akım değerlerinin doğru olduğu varsayılmıştır. Diğer istasyon EIE 1541 no'lu istasyon Delice kolunda yer almakta olup herhangi bir kontrol yapısından etkilenmemektedir.



Şekil 5.1 Kızılırmak havzası ve akım gözlem istasyonlarının yerleri

5.1 Entropi Metodunun Uygulaması

5.1.1 Harmancıođlu Metodunun Uygulaması

Harmancıođlu metodunda tanımlanan metodoloji normal ve log-normal dađılımları için Kızılırmak Havzasındaki 5 akım gözlem istasyonu yıllık akım verilerine uygulanmıştır. Her bir istasyonda ölçülen deđişkenlerin ölçüm aralıkları farklı olmasına rağmen 1955 ve 1995 yılları arasındaki her istasyon için ortak olan dönem gözönüne alınmıştır. Tekrar eden bilginin minimum olması kriterine göre istasyonlar normal, log-normal dađılımları için akım ađında kalması gerekli istasyonlar olarak sıralanmışlardır. Sıralamadaki ($r=5$) ölçüme devam etmesi mutlak gerekli olan istasyonu, ($r=1$) ise ölçüme devam etmesi gerekli olmayan en fazla tekrar eden bilgiye sahip istasyonu göstermektedir.

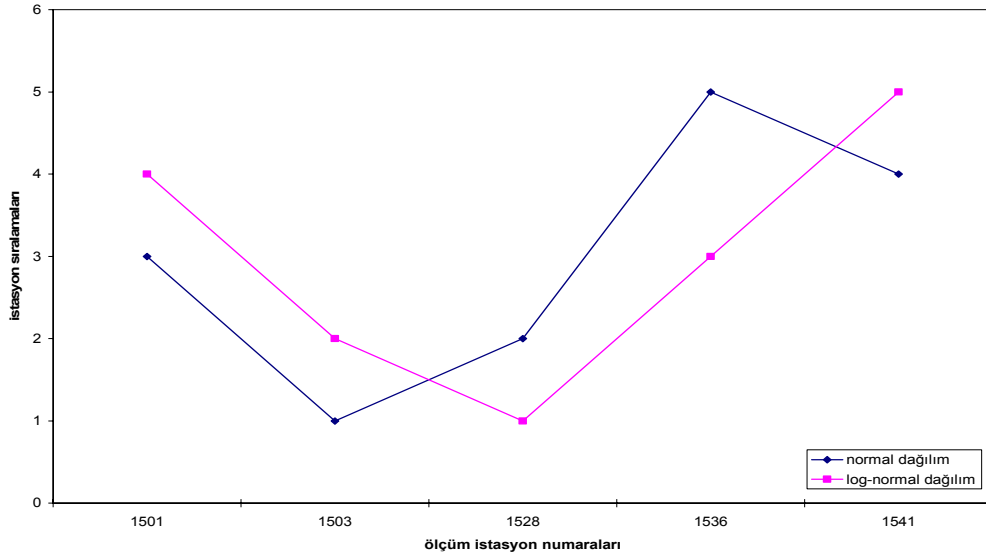
Tablo 5.1. Normal Dađılıma göre İstasyonların Seçilmesi

İstasyon numarası	Eklenen istasyonlar	Marjinal Entropi	Ortak Entropi	Şarh Entropi	Tekrar eden bilgi	Sıra (r)
	(M)	(napier)	(napier)	(napier)	(napier)	-
1536	1	8.74	8.74	-	-	5
1541	2	5.88	12.82	6.94	1.795	4
1501	3	7.56	18.06	10.50	2.319	3
1528	4	8.47	22.89	14.42	3.630	2
1503	5	7.97	25.83	17.85	5.039	1

Normal ve log-normal dađılım için seçilen istasyonlar Tablo 5.1 ve Tablo 5.2’de sırasıyla gösterilmiş ve Şekil 5.2 aynı dađılımlar için istasyonların ölçüm sıralamalarını göstermek için çizilmiştir.

Tablo 5.2. Log-normal Dađılıma göre İstasyonların Seçilmesi

İstasyon numarası	Eklenen istasyonlar	Marjinal Entropi	Ortak Entropi	Şarh Entropi	Tekrar eden bilgi	Sıra (r)
	(M)	(napier)	(napier)	(napier)	(napier)	-
1541	1	-2.08	-2.08	-	-	5
1501	2	-2.58	-5.79	-3.22	1.14	4
1536	3	-3.06	-10.40	-7.34	1.54	3
1503	4	-2.88	-17.12	-14.24	3.84	2
1528	5	-3.08	-30.10	-27.03	9.91	1



Şekil 5.2. Normal ve lognormal dağılımlar için minimum tekrar eden bilgiye dayalı istasyon sıralaması

5.1.2 Markus Metodunun Uygulaması

5.1.2.1 Markus metodunun uygulaması (Normal ve Log-Normal Dağılımları için):

Bu çalışmada her bir istasyonun yıllık akım verilerinin normal ve log-normal dağıldığı kabul edilerek Markus metodu uygulanmıştır. Entropiye dayalı bu metot aktarılan bilgi için marjinal entropi ve ortak entropi ifadeleri kullanılarak uygulanmıştır. i istasyonu tarafından alınan toplam bilgiyi ($R(i)$) hesaplamak için denlem 2.9 ile 2.12 kullanılırken, denklemler 2.10 ve 2.13 ise i istasyonu tarafından aktarılan ($S(i)$) bilgiyi hesaplamak için kullanılmıştır. Sonuç olarak i istasyonu için net bilgi ($N(i)$) denklemler 2.14 kullanılarak elde edilmiştir. Bilgi aktarım parametre ($S(i)$, $R(i)$ ve $N(i)$) değerlerine göre seçilen istasyon verileri için önem seviyesine göre istasyon sıralamaları Tablo 5.3'de normal dağılım için Tablo 5.4'de log-normal dağılım için verilmiştir.

Tablo 5.3. Normal Dağılım İçin Aktarılan Bilgi, $S(i)$, Alınan Bilgi, $R(i)$ ve Net Bilgi, $N(i)$ Değerleri

İstasyon No	İstasyon Numarası	bilgi			Sıra (r)		
		aktarılan $S(i)$	alınan $R(i)$	net $N(i)$	aktarılan $S(i)$	alınan $R(i)$	net $N(i)$
(i)	(M)	(napier)	(napier)	(napier)	-	-	-
1501	1	0.92290	0.92280	0.00015	2	2	2
1503	2	0.92506	0.92380	0.00125	3	3	5
1541	3	0.91814	0.91777	0.00037	1	1	4
1528	4	1.23465	1.23497	-0.00032	5	5	1
1536	5	1.08959	1.08941	0.00018	4	4	3

Düşük sıralamadaki istasyonlar (örneğin $r=1$), bilgi aktarımında daha az öneme sahip ve gözlem ağından çıkartılması en muhtemel istasyon olarak, yüksek sıralamadaki istasyonlar (örneğin $r=5$) gözlem ağına ölçüme devam etmesi gereken, yüksek bölgesel

değer içeren istasyonlar olduğuna işaret etmektedir. Net bilgi değerinin negatif olması, bilgi alımının bilgi aktarımından daha baskın olduğunu gösterir.

Tablo 5.4. Lognormal Dağılım İçin Aktarılan Bilgi, S(i), Alınan Bilgi, R(i) ve Net Bilgi, N(i) Değerleri

İstasyon No	İstasyon Numarası	bilgi			Sıra (r)		
		aktarılan	alınan	net	aktarılan	alınan	net
		S(i)	R(i)	N(i)	S(i)	R(i)	N(i)
(i)	(M)	(napier)	(napier)	(napier)	-	-	-
1501	1	-96.048	-96.909	0.1391	5	5	4
1503	2	-358.211	-358.569	-0.3584	3	3	2
1541	3	-118.050	-117.499	0.5511	4	4	5
1528	4	-2320.400	-2320.924	-0.5242	1	1	1
1536	5	-1892.092	-1892.379	-0.2871	2	2	3

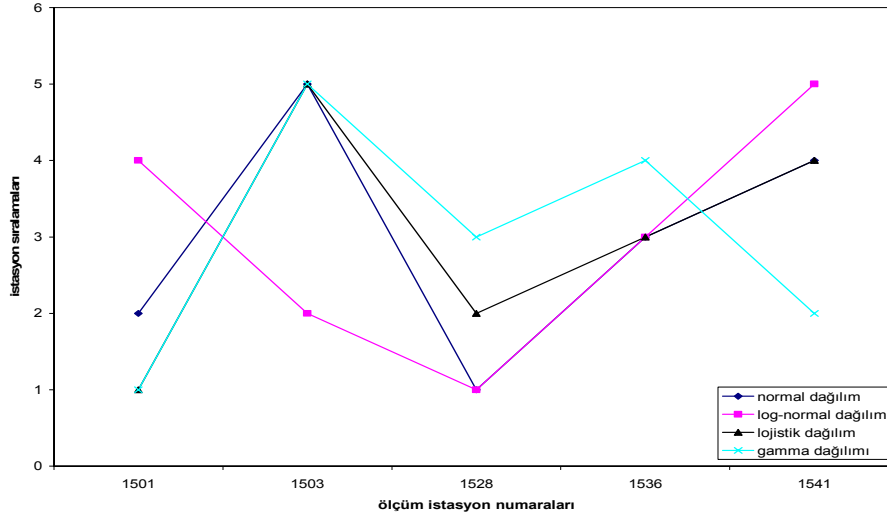
5.1.2.2 Markus metodunun uygulaması (Gamma Dağılımı için):

Gamma dağılımının uygulaması normal ve log-normal dağılımlar için geçerli olan yöntemin aynısıdır. Sadece kullanılacak olan formüllerde gamma dağılım fonksiyonunun kullanılması gerekir. Bilgi parametreleri S (i), R (i) ve N (i) Tablo 5.5'da gamma dağılımı için gösterilmiştir.

Table 5.5. Gamma Dağılımı İçin Aktarılan Bilgi, S(i), Alınan Bilgi, R(i) ve Net Bilgi, N(i) Değerleri

İstasyon No	İstasyon Numarası	bilgi			Sıra (r)		
		aktarılan	alınan	net	aktarılan	alınan	net
		S(i)	R(i)	N(i)	S(i)	R(i)	N(i)
(i)	(M)	(napier)	(napier)	(napier)	-	-	-
1501	1	0.634890	0.635020	-0.000121	1	1	1
1503	2	0.690330	0.690240	0.000098	2	2	5
1541	3	0.720000	0.720150	-0.000096	3	3	2
1528	4	1.047992	1.047986	0.000006	5	5	3
1536	5	0.993540	0.993520	0.000025	4	4	4

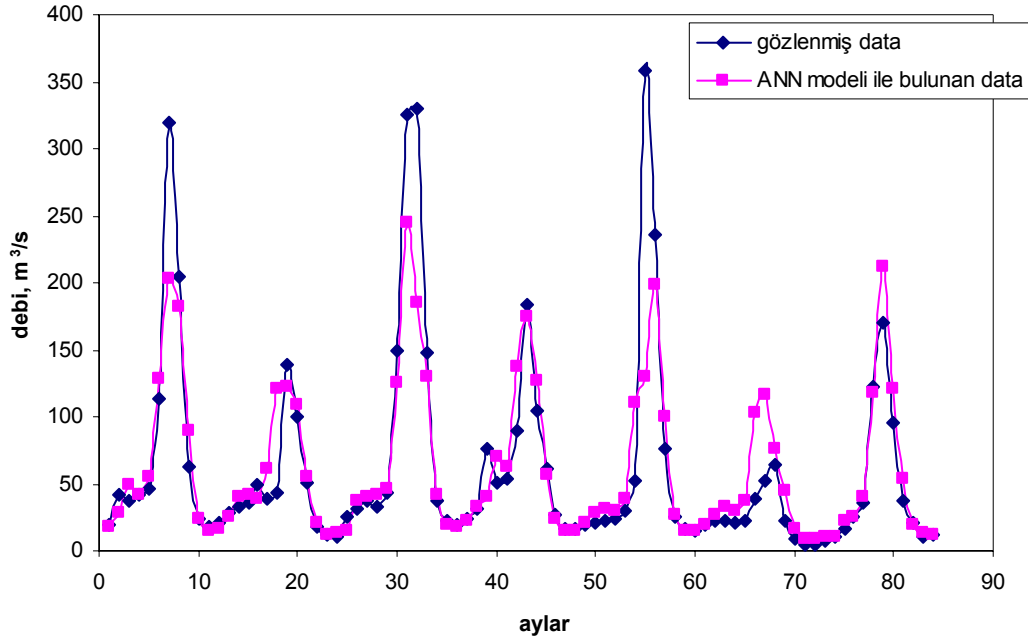
Gamma dağılımı için Markus metodu uygulandığında, EIE 1528 istasyonu gözlem ağında ölçülmesi en gerekli istasyon olarak, akım gözlem ağının en membasındaki istasyon olan EIE 1501 istasyonu ise bilgi transferi bakımından en önemsiz istasyon olarak belirlenmektedir.



Şekil 5.3. Normal, log-normal ve gamma dağılımları için ölçüm ağındaki istasyonların net bilgi miktarlarına göre sıralanması

5.2 Yapay Sinir Ağları Modelinin Kızılırmak Havzasına Uygulaması

Kızılırmak Havzasında yer alan EIE 1501 akım gözlem istasyonu 1955-1995 yılları arasındaki aylık akım verilerine yapay sinir ağları modeli geliştirilmiş ve 1995-2002 yılına kadar olan aylık akım verileri tahmin edilmeye çalışılmıştır. EIE 1501 istasyonu verilerinin seçilmesinin nedeni elde 2002 yılına kadar olan aylık verilerin olması ve hiç bir su yapısından etkilenmemesinden dolayıdır. Diğer istasyonlar içinde model kurulması ve akımın tahmin edilmesi mümkün olmasına karşın burada sadece EIE 1501 için yapılan çalışma verilmiştir. Yapay sinir ağları modelini kurmak için neuralware paket programından faydalanılmıştır. Neuralware paket programında girdi verileri ve çıktı verileri belirlenmiştir. Gözönüne alınan veriler aylık akım olduğundan öncelikle akım verilerinde mevcut olan mevsimsellik etkisi giderilmiştir. Bunun için literatürde en yaygın şekilde kullanılan yöntem tercih edilmiştir. Bunun için mevsimsel aylık ortalama ve standart sapma değerleri belirlenmiş ve aylık akım değerlerinden aynı aya ait ortalama değer çıkartılıp standart sapma değerine bölünmüştür. Bu şekilde aylık akım değerlerindeki mevsimsellik etkisi giderilmeye çalışılmıştır. Yapay sinir ağı modelinde sigmoid fonksiyonundan faydalanılmıştır. Bu sebepten mevsimsellik etkisi giderilmiş aylık akım verileri [0,1] aralığında tariflenmiştir. Paket program yardımı ile model kurulmuş ve 1995-2002 yılları arasındaki aylık akım verileri tahmin edilmiştir. Şekil 6.4'de modelden elde edilen ve gözlemlenmiş akım verileri gösterilmiştir.



Şekil 5.4 EIE 1501 İstasyonunun ANN modeli kullanılarak modellenen elde edilen ve gözlemlenmiş aylık veri seti

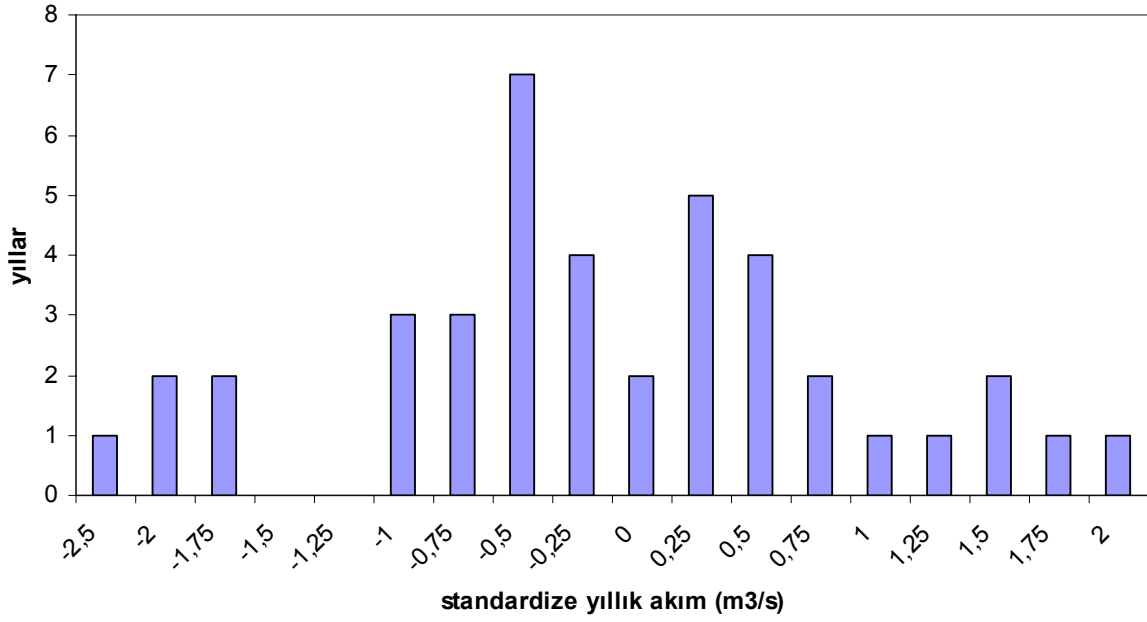
İstatistiksel olarak model performansını ölçmek için hata karelerinin karekökü olarak ifade edilen RMSE değeri elde edilmiştir. EIE 1501 istasyonu için RMSE değeri 38.58 olarak hesaplanmıştır (Sarlak,2005).

5.3 GEÇİŞ TEOREMİNİN UYGULAMASI

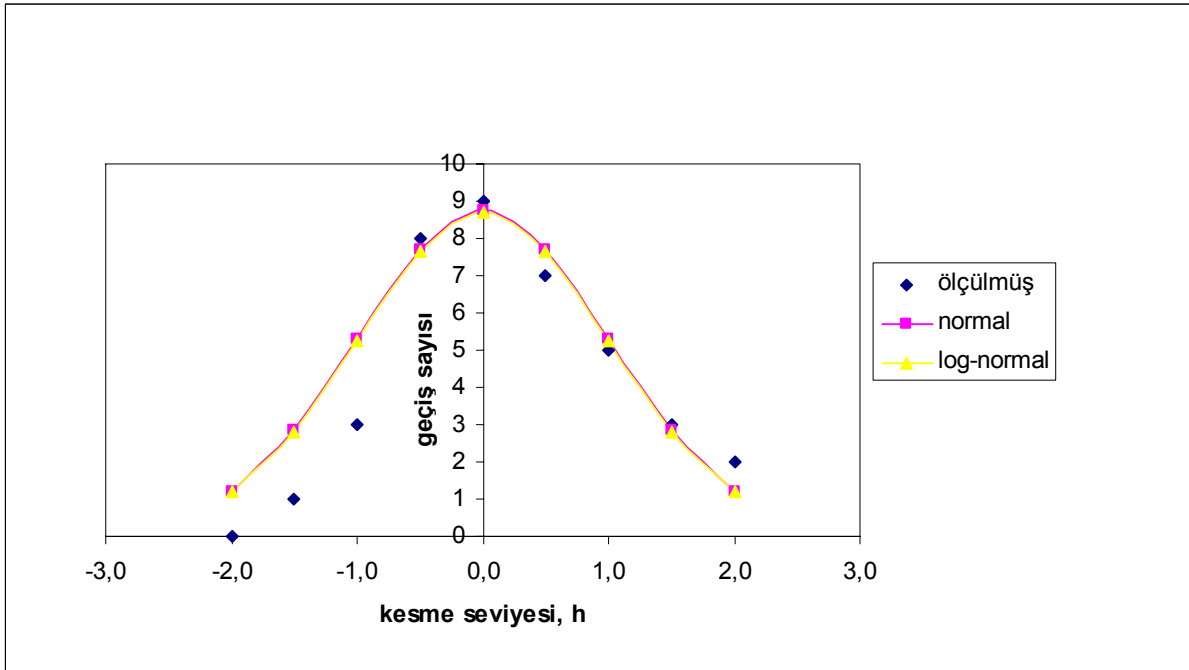
Geçiş teoremi Kızılırmak havzasında yer alan EIE 1501 istasyonu standardize edilmiş yıllık akım verilerine uygulanmıştır. Data setinin bazı istatistiksel özellikleri Tablo 5.7’de özetlenmiştir. Data setinin histogramı Şekil 5.6’da serinin çarpıklığını görmek için çizilmiştir. Bu şekle göre serinin önemli miktarda çarpık olmadığı gözlemlenmiştir. Bu sebepten normal dağılımın kullanılabilmesine karar verilmiştir. Çıkarılan bu sonucu, Şekil 5.7’den Şekil 5.10’a kadar çizilen ölçülmüş değerler gözönüne alınarak bulunan geçiş davranışları ile normal dağılım varsayımı altında teorik olarak bulduğumuz geçiş davranışlarının yakın sonuçlar vermesi doğrulamıştır. Aynı şekiller üzerinde log-normal dağılım için bulunan teorik değerlerde gösterilmiştir.

Tablo 5.6. EIE 1501 istasyonu yıllık akım verilerinin istatistiksel özeti

İstatistikler	EIE 1501
	değerler
Kayıt süreci	1955-1995
Ortalama akım, m ³ /s	0.00
Standart sapma, m ³ /s	1.00
Çarpıklık katsayısı	-0.0548
Birinci mertebe korelasyon katsayısı	0.10
Birinci mertebe korelasyon katsayısı (log-data)	0.11

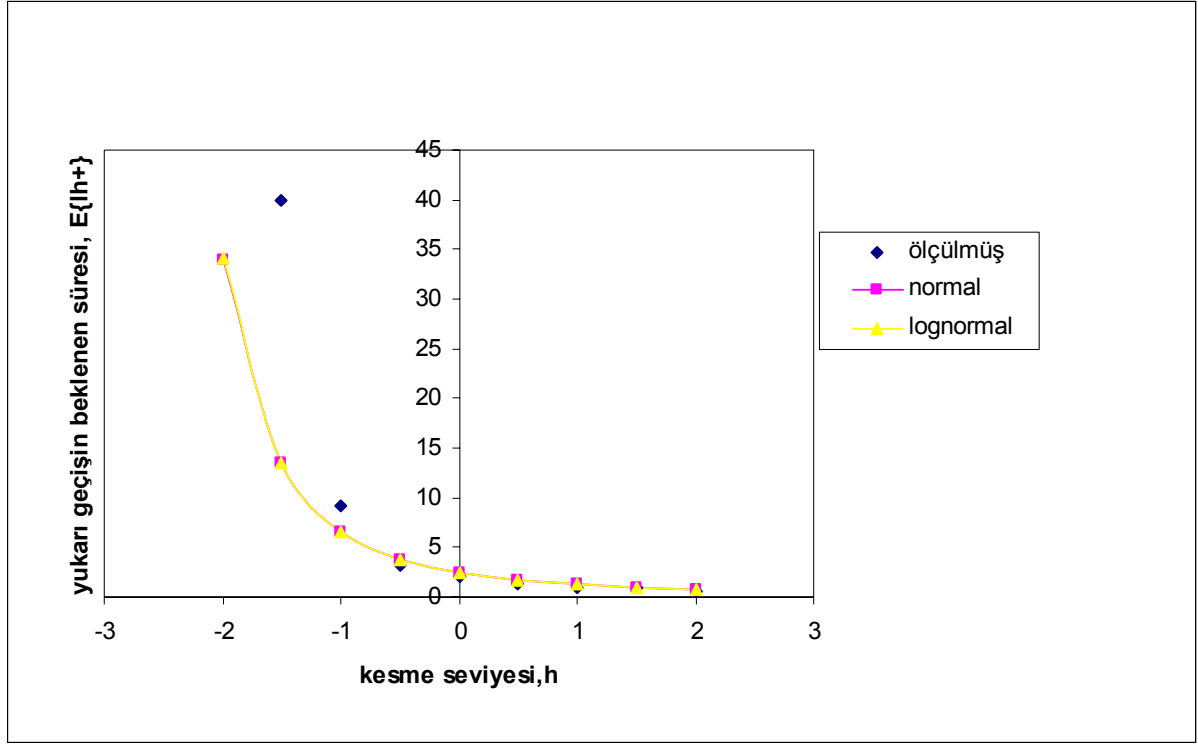


Şekil 5.5. EIE 1501 standardize akım verilerinin histogramı

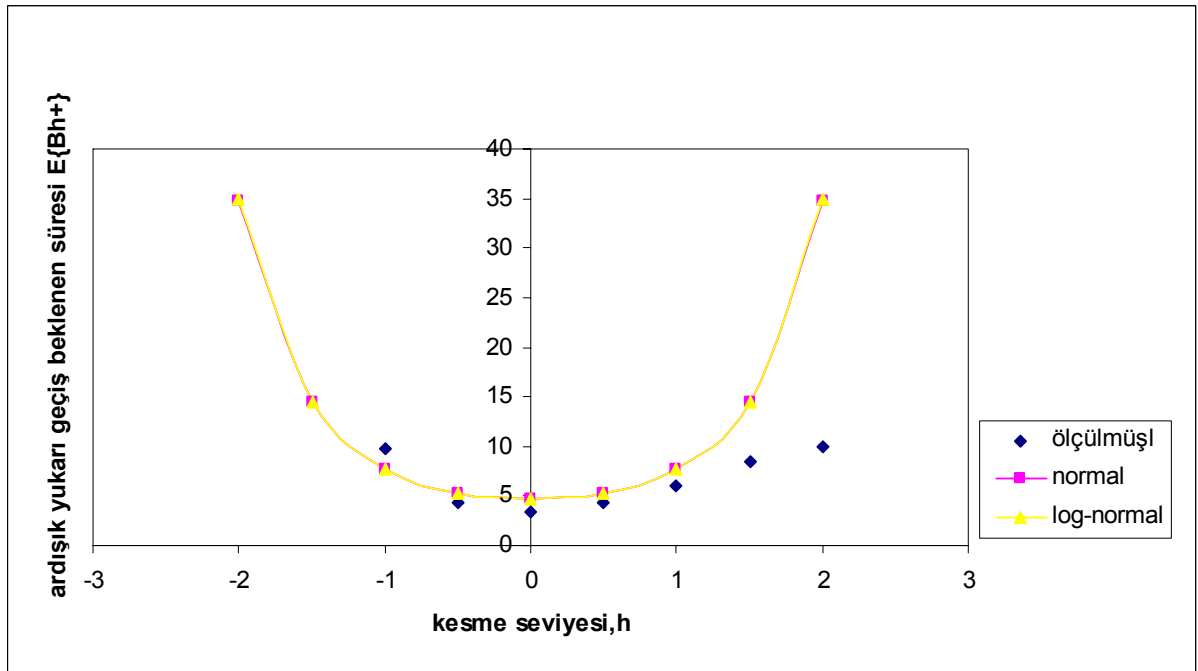


Şekil 5.6 EIE 1501 istasyonu için ölçülmüş ve modelle elde edilmiş beklenen geçiş sayıları; $E[N]$

Şekil 5.6'dan görüleceği gibi, ortalama değerden küçük değerler için bulunan ve teorik değerlerde normal ve lognormal dağılım varsayımları altında sapmalar gözlenmektedir. Bu sapmalar ihmal edilen çarpıklık değerinden kaynaklanmaktadır. Seride mevcut olan çarpıklık nedeniyle ortalamanın altındaki değerlerde teorik olarak bulunan değerlerden sapmalar oluşabilmektedir.



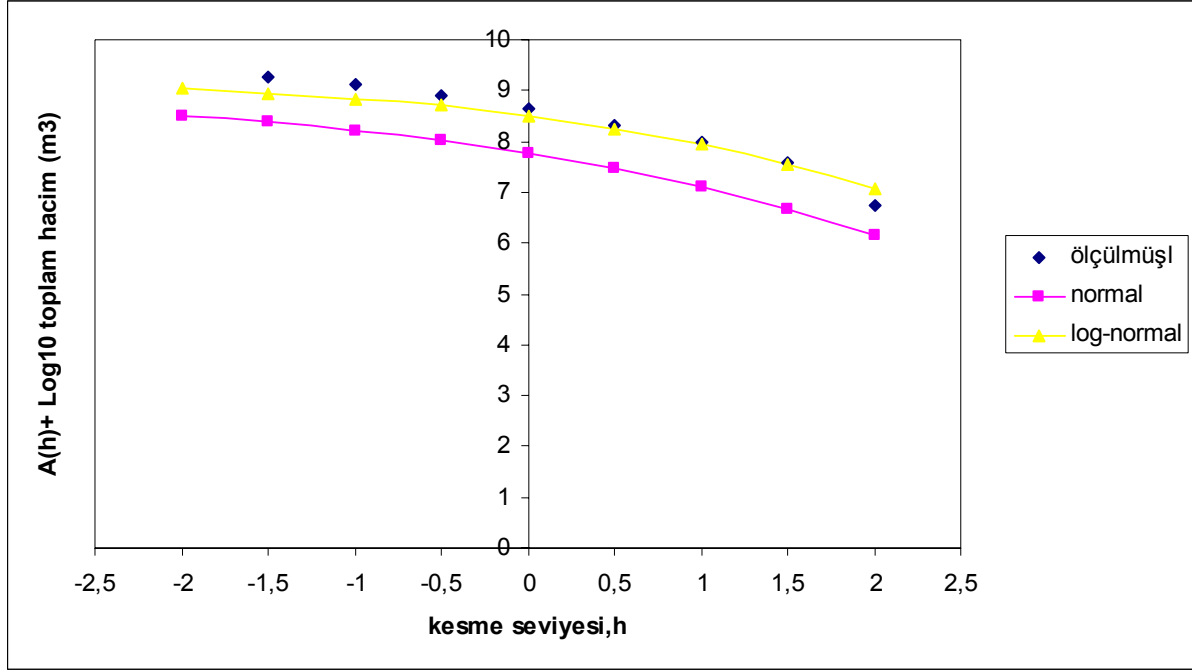
Şekil 5.7 EIE 1501 istasyonu verileri için belli kesme seviyelerindeki (h) yukarı geçiş noktalarının ölçülmüş ve teorik beklenen süresi, $E[h^+]$



Şekil 5.8 EIE 1501 istasyonu verileri için belli kesme seviyelerindeki (h) ardışık yukarı geçiş noktalarının ölçülmüş ve teorik beklenen süresi, $E[B_h^+]$

Teorik olarak bulunan ardışık yukarı geçiş noktalarının beklenen süresi ölçülmüş değerlerle bulunanların üstünde çıkmıştır. Bunun çarpıklık katsayısından kaynaklandığı düşünülmektedir. Aynı durum alan hesabında da kendini göstermiştir.

Buna karşılık hacim hesabında log-normal dağılımdan elde edilen değerlerin kesme seviyesinin üstünde gözlenmiş datadan elde edilen değerlere oldukça yakın olduğu görülmektedir. Gözlem hataları da göz önüne alındığında geçiş teoremi ile teorik olarak bulunan sonuçların güvenli olduğu sonucuna varılabilir. Bu da su kaynaklarının planlanmasında bu teoremin kullanımının yararlı olacağı sonucuna bizi götürmektedir.



Şekil 5.9 EIE 1501 istasyonu verileri için belli kesme seviyelerinin üzerindeki ölçülmüş ve teorik toplam logaritmik hacim miktarları, $E[A_h^+(T)]$

6. ÖZET VE SONUÇLAR

İki entropi metodu Kızılırmak havzasında yer alan akım gözlem istasyonu verilerine uygulanarak istasyonlar önem seviyelerine göre sıralanmışlardır. Harmancıoğlu yöntemi minimum tekrar eden bilgi miktarlarına göre istasyon seçilmesi esasına dayanmaktadır. Seçilen ilk istasyon en fazla önem seviyesine sahip istasyondur. Yani, seçilen istasyon ölçüm ağına kalmalı ve ölçüm yapmaya devam etmelidir. Markus yönteminde de istasyonlar sıralanmaktadır. Düşük sıra numarası ölçüm ağına en az önemli istasyon olup bu metota göre ölçüme devam etmemesine karar verilecek istasyonların başında yer almaktadır. Yüksek sıra numarası ölçüm ağına mutlak kalması gerekli istasyonu ifade etmektedir. Farklı dağılım varsayımları altında istasyon önem sıralamaları Harmancıoğlu ve Markus yöntemi için Tablo 6.1'de dağılım tiplerinin entropi metodu üzerindeki etkilerini göstermek için özetlenmişlerdir.

Log-normal dağılımı varsayımı altında iki metotla da EIE 1501 ve EIE 1541 istasyonları ölçüme devam etmesi gereken istasyonlar olarak belirlenmişlerdir. Her iki istasyonda ölçüm ağının en membasında ve farklı kollarda yer aldığından en yüksek belirsizliğe sahip olmaları beklenen bir durumdur. Bir başka deyişle, elde edilen sonuç diğer dağılımlar için bulunan sıralamalardan daha makuldür.

En az tekrar eden veya taşıdıkları net bilgi miktarlarına göre EIE 1503 istasyonu üçüncü öncelikli istasyon olarak seçilmesi beklenirken, her iki model de log-normal dağılım varsayımı altında EIE 1536 istasyonu üçüncü öncelikli istasyon olarak seçilmektedir. EIE 1536 gözönüne alınan ölçüm ağının en mansabında yer aldığı düşünülürse ölçüm yapmaya devam etmesi gereken istasyon olarak seçilmesi doğaldır.

Çalışmanın başlangıcında amacımız veri setine en uygun dağılım tipinin bulunması ve iki farklı entropi yönteminin uygulanmasıydı. Fakat bazı dağılımların çoklu dağılım fonksiyonlarının olmaması çalışmayı kısıtlamıştır. Buna rağmen farklı dağılımlar için Entropi yöntemlerinden elde edilen sonuçların farklı çıktığı gözlemlenmiştir. Bu da eldeki seriye en uygun dağılımın bulunarak entropi yöntemlerinin uygulanmasının elde edilecek sonuçların güvenliği açısından oldukça önemli olduğunu vurgulamaktadır.

Bu çalışmada, log-normal dağılımın diğer dağılımlarla kıyaslandığında eldeki akım serileri için en uygun dağılım olduğu istatistiksel testler yardımı ile bulunmuştur. Bu çalışmada Kolmogorov-Smirnov ve Shapiro-Wilk testleri uygulanmış ve serilerin log-normal dağılıma daha iyi uyduğu bulunmuştur. İki farklı metotla log-normal dağılım için aynı sıralamanın bulunması bunu doğrulamaktadır. Bu sonuca göre akım ağında EIE 1503 ve EIE 1528 nolu istasyonların kapatılmasına karar verilebilir. Bilindiği gibi akım gözlem istasyonlarının işletilmesi oldukça pahalıdır. Bu bakımdan yeniden yapılması gerekli istasyon gözlem ağlarına Entropi yönteminin uygulanması ve bilimsel olarak hangi istasyonların önemli olduğuna karar verilmesi Entropi metodu ile mümkün olmaktadır.

Tablo 6.1. Farklı Dağılımlarda Harmancıoğlu ve Markus Metotları İçin İstasyon Sıralamaları

İstasyon no.	Minimum tekrar eden bilgiye dayalı sıralama (Harmancıoğlu)		Net bilgi aktarımına dayalı sıralama (Markus)		
	Dağılım tipi		Dağılım tipi		
(i)	Normal	Log- normal	Normal	Log-normal	Gamma
1501	3	4	2	4	1
1503	1	2	5	2	5
1541	4	5	4	5	2
1528	2	1	1	1	3
1536	5	3	3	3	4

Çalışmanın ikinci kısmında gözlenmiş verileri en iyi şekilde uzatabilen stokastik modeller üzerinde durulmuştur. Stokastik modellerden yapay sinir ağları modeli ele alınmıştır. ANN modelinden elde edilen sonuçların korelasyonun daha yüksek olduğu aylık akım verilerinde daha iyi sonuç elde ettiği gözlemlenmiştir (Sarlak, 2005). Baraj planlanmasında kullanılacak olan tahmini akım değerlerini bulmak için kurulan modellerin önemli olduğu göz önüne alınırsa kullanılacak modellerin güvenilirliğinin test edilmesi gerekmektedir. Bu çalışmada hem görsel (Şekil 5.2) hem de istatistiksel olarak (RMSE=38.58) model test edilmiştir. Hata değerinin sıfıra yakın olması modelin

mükemmel olduğunu göstermektedir. Ancak, gözlem hatalarını, doğal ortam gözönüne alındığında aslında istatistiksel yöntemlerden elde edilen sonuçların hiçbir zaman sıfır veya sıfıra bir hata ile bulunamayacağı açıktır. Bu bakımdan elde edilen hata miktarının azımsanmayacak kadar iyi olduğu düşünülmektedir.

Çalışmanın üçüncü kısmında entropi yönteminde ölçülmesi gerekli olan istasyonlara geçiş teoreti uygulanmıştır. Bu çalışmada yalnızca EIE 1501 nolu istasyon verilerinden elde edilen sonuçlar sunulmuştur. Geçiş teorisi, beklendiği gibi, normal ve log normal dağılımları varsayımı altında uç değerlerin istatistiklerini bulmada oldukça faydalı bir yöntemdir. Bu çalışmada normal ve log-normal dağılımları varsayımı ile geçişlerin sayısı, kesme seviyelerinin altında veya üstünde kalan süreleri ve yukarıda bahsedilen diğer istatistiksel özelliklerin teorik değerleri bulunmuş ve gözlenmiş veriler ile elde edilen sonuçlarla kıyaslanmıştır. Standardize edilen değerler kullanıldığından normal ve log normal dağılım için bulunan teorik değerler üst üste çakıştığından geçiş teoreminde dağılımların etkisi incelenememiştir. Yalnız hacim hesabında dağılımların etkisi gözlemlenebilmiş ve log-normal dağılımdan elde edilen sonuçların daha iyi olduğu gözlemlenmiştir. Bu da serimizin istatistiksel olarak gösterildiği gibi normal dağılıma uymadığından kaynaklanmaktadır. Geçiş teoremi ile elde edilen belirli bir kesme seviyesini elimizdeki serinin kaç kez aşdığı, ardışık kesme seviyesini aşan (veya aşmayan) değerler arasındaki zamanı, kesme seviyesini aşan (veya aşmayan) değerler arasında geçen zamanı ve en önemlisi kesme seviyesini aşma (veya aşmama) hacmini bulabilmekteyiz. Bu da bize son günlerde çağımızın büyük bir sorunu olan kuraklık çalışmalarında kullanılabilecek bir yöntem olduğunu göstermektedir.

Projenin başlangıcında amacımız serilere en uygun dağılımları bulup yukarıdaki yöntemlerin entegre edilmesiydi. Ancak Entropi ve geçiş teoremlerinin normal ve log-normal dağılım dışındaki dağılımlara uygulanmalarında bazı zorluklarla karşılaşmıştır. Bu sebepten sözkonusu yöntemler yalnızca bahsi geçen dağılımlara uygulanabilmişlerdir. Böylelikle Entropi yöntemi ile seçilen gözlem ağındaki istasyonların bilgi bakımından hangisinin daha iyi olduğuna karar verilerek, YSA modeli ile elde edilen serilerin modellenmesi yapılmış ve geçiş teoremi ile kuraklık ve taşkın için önemli parametreler elde edilmeye çalışılmıştır. Elde edilen sonuçlar proje kapsamında verilmeye çalışılmıştır.

7. KAYNAKLAR

1. Desmond A. F. And Guy B.T., Crossing Theory for Non-gaussian Stochastic Processes with an Application to Hydrology, *Water Resources Research*, Vol. 27, No. 10, p. 2791-2797, October, (1991).
2. Dikmen, İ., Strategic Decision Making in Construction Companies: an Artificial Neural Network Based Decision Support System for International Market Selection. (Doktora Tezi), ODTU, Ankara (2001).
3. Harmancıoğlu N., Measuring the Information Content of Hydrological Process by the Entropy Concept, *Ege Üniversitesi, İnşaat Fakültesi Dergisi*, Atatürk'ün 100. Doğum Yılı Özel Sayısı. p. 13-40, İzmir, (1981).
4. Markus M., Knapp H. V. and Tasker G.D, Entropy and Generalized Least Square Methods in Assessment of the Regional Value of Streamgages, *Journal of Hydrology* 283, p. 107-121, (2003).

5. Nordin C. F. And Rosbjerg D.M, Applications of Crossing Theory in Hydrology, *Bull. Int. Assoc.Sci. Hydrol.*, 15, 27-43, (1970).
6. Özkul S., *Space / time Design of Water Quality Monitoring Networks by the Entropy Method*, (Doktora tezi), Dokuz Eylül Üniversitesi, İnşaat Mühendisliği Fakültesi, İzmir, (1996).
7. Raman, H. and Sunilkumar, N., Multivariate Modeling of Water Resources Time Series Using Artificial Neural Networks. *Hydrologic Sciences*, 40(2), 145-163 (1995).
8. Rice, S. O., Mathematical Analysis of Random Noise. *Bell System Tech. J.* 24, 46-56, (1945).
9. Sarlak, N., *Evaluation and Modeling of Streamflowdata: Entropy Method, Autoregressive Models with Asymmetric Innovations and Artificial Neural Networks*, (Doktora tezi), ODTU, Ankara, (2005).
10. Yang, Y. and Burn, D.H., An Entropy Approach to Data Collection Network Design. *Journal of Hydrology* 157, 307-324, (1994).

EK A: NORMAL DAĞILIM İÇİN GEÇİŞ TEOREMİ

$y(t)$, örneğin t zamanındaki akım değerli sürecinin sıfır ortalamalı kararlı normal dağılıma uyduğu varsayalım. Orjinde spektral momentlerle otokovaryans fonksiyonları arasındaki standart ilişki aşağıdaki gibi tanımlanmaktadır:

$$\lambda_{2i} = (-1)^i \psi^{(2i)}(0) = (-1)^i \sigma^2 \rho^{(2i)}(0) \quad (1)$$

denklemdaki $\sigma^2 = \lambda_0 = \text{Var}(y(t))$.

$y(t)$ sürecinin belli bir kesme seviyesinin aşan değerlerinin beklenen sayısı Rice (1945) tarafından verilmiştir:

$$E[N] = \frac{T}{2\pi} \left(\frac{\lambda_2}{\lambda_0} \right)^{1/2} \exp\left(-\frac{1}{2} h^2\right) = \mu_h T, \quad (2)$$

buradaki h standardize edilmiş kesme seviyesini göstermektedir.

$$\mu_h = \frac{1}{2\pi} \left(\frac{\lambda_2}{\lambda_0} \right)^{1/2} \exp(-h^2 / 2) \quad (3)$$

μ_h ; birim zamandaki ortalama yukarı geçiş (upcrossing) oranıdır.

Sürecin belli kesme seviyelerinin üzerinde yukarı geçişin beklenen süresi;

$$E[I_h^+] = \mu_h^{-1} \text{pr}\{y(0) > h\} \quad (4)$$

olarak tarif edilmiştir ve $\text{pr}\{y(0) > h\} = \frac{1}{\sqrt{2\pi}} \int_h^\infty e^{-\frac{1}{2}(x)^2} dx$.

B_h^+ ; ardışık yukarı geçişler arasındaki beklenen zaman denklem (5) ile hesaplanmaktadır:

$$E[B_h^+] = \mu_h^{-1} \quad (5)$$

$y(t)$ sürecinin h seviyesinin üstünde kalan alan $TZ(T)$ 'dir ve beklenen değeri; $E\{Z_1(T)\}$;

$$E\{Z_1(T)\} = \int_h^\infty (x - h) f_y(x) dx \quad (6)$$

şeklinde hesaplanmaktadır. $y(t)$ sürecinin h seviyesi üzerinde kalan alan A_h^+ olarak ifade edildiği belirtilmiştir. A_h^+ 'ın olasılık dağılımını elde etmek kolay olmadığından beklenen değeri denklem (7) ile yaklaşık olarak normal dağılım için hesaplanabilmektedir (Nordin ve Rosberg ;1970):

$$E\{A_h^+\} = \frac{1}{\mu_h} \int_h^\infty (x - h) f_y(x) dx \quad (7)$$

Log-normal Dağılım İçin Geçiş Teoremi

Lognormal süreç, $y(t)$ 'nin ortalaması; m , varyansı; σ^2 ve otokorelasyon fonksiyonu; $\rho_y(\tau)$ olan kararlı normal dağılıma sahip olduğu varsayılarak tariflenmektedir. Bu durumda yeni süreç denklem (8)'deki gibi gösterilir:

$$\gamma(t) = e^{y(t)} \quad (8)$$

$\gamma(t)$ marjinal lognormal dağılıma sahiptir. Desmond ve Guy (1991) $\gamma(t)$ 'nin ortalamasını; $m_\gamma = \exp(m + \sigma^2)$ ve varyansını; $\sigma_\gamma^2 = (e^{\sigma^2} - 1)\exp(2m + \sigma^2)$ olarak tariflerken, korelasyon katsayısını; $\rho_\gamma(\tau)$ denklem (9)'daki gibi hesaplamışlardır:

$$\rho_\gamma(\tau) = \frac{\exp\{\sigma^2 \rho_y(\tau) - 1\}}{\exp(\sigma^2) - 1} \quad (9)$$

Log-normal dağılım için yukarıdaki denlemler kullanıldığı gibi $y(t)$ değerlerinin logaritması alınarak normal dağılım için geçerli formüller de kullanılabilir.

TÜBİTAK PROJE ÖZET BİLGİ FORMU

Proje No: 102I068
Proje Başlığı: Entropi, Yapay Sinir Ağları ve Geçiş Teoremlerinin Akım Gözlem Ağının ve Su Kaynakları Sistemlerinin Planlanmasında Birlikte Kullanılması ve Kızılırmak Havzasına Uygulanması
Proje Yürütücüsü ve Araştırmacılar: Prof. Dr. A. Ünal ŞORMAN Arş.Gör. Nermin ŞARLAK
Projenin Yürütüldüğü Kuruluş ve Adresi: Orta Doğu Teknik Üniversitesi, ODTU İnşaat Mühendisliği Bölümü, Su Kaynakları Lab., ANKARA
Destekleyen Kuruluş(ların) Adı ve Adresi: Tubitak Elektrik İşleri ve Etüt İdaresi
Projenin Başlangıç ve Bitiş Tarihleri: 01.01.2003~01.07.2005
Öz (en çok 70 kelime) Entropi yöntemi ile seçilen gözlem ağındaki istasyonların bilgi bakımından hangisinin daha iyi olduğuna karar verilerek, YSA modeli ile eldeki serilerin modellenmesi yapılmış ve geçiş teoremi ile kuraklık ve taşkın için önemli parametreler elde edilmeye çalışılmıştır. Söz konusu yöntemler Kızılırmak Havzasındaki seçilen akım gözlem istasyonlarına uygulanmıştır. Böylece seçilen akım gözlem istasyonlarından EIE 1501, EIE 1541 ve EIE 1536 istasyonlarının ölçümüne devam etmesi gerekli istasyonlar olarak seçilerek sentetik seri türetimi bu seriler için YSA modeli ile yapılmış ve geçiş teoremi ile kurak ve ıslaklık ile ilgili parametreler elde edilmiştir.
Anahtar Kelimeler: Entropi, Yapay sinir ağları, Geçiş teoremi
Projeden Çıkan Yayınlar: Sarlak, N., and Sorman, A.U., (2006), "Evaluation and Selection of Streamflow Network Stations Using Entropy Methods", <u>Turkish Journal of Engineering and Environmental Sciences</u>, The Scientific and Technical Research Council of Turkey (TÜBİTAK), 30, 91-100.