

Atf İçin: Yılmaz M, Tosunoğlu F, Demirel MC, 2021. Taşkın Frekans Analizinde Klasik Yöntemler ve Alternatif Bir Parametre Tahmin Yönteminin Karşılaştırılması. İğdır Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 11(2): 1243-1254.

To Cite: Yılmaz M, Tosunoğlu F, Demirel MC, 2021. Comparison of Classical Methods and an Alternative Parameter Estimation Method in Flood Frequency Analysis. . Journal of the Institute of Science and Technology, 11(2): 1243-1254.

Taşkın Frekans Analizinde Klasik Yöntemler ve Alternatif Bir Parametre Tahmin Yönteminin Karşılaştırılması

Muhammet YILMAZ^{1*}, Fatih TOSUNOĞLU¹, Mehmed Cüneyd DEMİREL²

ÖZET: Taşkın frekans analizi su yapılarının tasarımı, taşkın kontrolünün planlanması ve projelendirilmesinde oldukça önemlidir. Taşkın frekans analizinin önemli adımlarından biri veriler için uygun bir dağılım modelinin ve parametrelerinin tahmin edilmesidir. Bu çalışmada, olasılık dağılım fonksiyonlarının parametrelerini tahmin etmek için optimizasyon tabanlı DEoptim algoritması önerilmiş ve geleneksel maksimum olabilirlik yöntemi (MLM) ve momentler yöntemi (MOM) ile karşılaştırılmıştır. Doğu Karadeniz Havzası üzerinde bulunan 2 adet istasyona ait yıllık maksimum akım verileri kullanılmıştır. Taşkın tekerrür debilerinin tahmini için Gamma, Weibull, Log-Normal, Logistic ve Normal olasılık dağılımları kullanılmıştır. En uygun olasılık dağılımını belirlemek için Kolmogorov-Smirnov (KS) testi, ortalama hata karesi (OHK), Kling-Gupta etkinlik katsayısı (KGE), ortalama mutlak hata (OMH), Nash-Sutcliffe etkinlik katsayısı (NSE) ve Akaike bilgi kriteri (AIC) testleri kullanılmıştır. Sonuçlara göre, DEoptim yönteminin MLM ve MOM yöntemine kıyasla olasılık dağılım fonksiyonu parametrelerini tahmin etmede daha etkili bir yöntem olduğu görülmüştür. Son olarak, en uygun olasılık modelleri kullanılarak farklı tekerrür periyotlarında (5, 10, 20, 50, 100, 200 ve 500) meydana gelebilecek taşkın tahminleri hesaplanmış ve sonuçlar tartışılmıştır.

Anahtar Kelimeler: Doğu Karadeniz Havzası, maksimum olabilirlik metodu, DEoptim, olasılık dağılımları, momentler yöntemi

Comparison of Classical Methods and an Alternative Parameter Estimation Method in Flood Frequency Analysis

ABSTRACT: Flood frequency analysis is very important in the design of water structures, planning and projecting of flood control. One of the important steps of flood frequency analysis is to estimate a suitable distribution model and its parameters for the data. In this study, optimization based DEoptim algorithm is proposed to estimate the parameters of probability distribution functions and is compared with the traditional maximum likelihood method (MLM) and Method of moments (MOM). Annual maximum flow data of 2 stations on the Eastern Black Sea Basin are used. Gamma, Weibull, Log-Normal, Logistic and Normal probability distributions were used to determine flood recurrences flows. To determine the most appropriate probability distribution, Kolmogorov-Smirnov (KS) test, mean square error (RMSE), Kling-Gupta efficiency coefficient (KGE), mean absolute error (MAE), Nash-Sutcliffe efficiency coefficient (NSE) and Akaike information criterion (AIC) tests were used. According to the results, it was seen that the DEoptim method is a more effective method in estimating the probability distribution function parameters compared to the MLM and MOM methods. Finally, using the most appropriate probability models, flood estimates that may occur in different recurrence periods (5, 10, 20, 50, 100, 200 and 500) were calculated and the results were discussed.

Keywords: Eastern Black Sea Basin, Maximum likelihood method, DEoptim, probability distributions, Method of moments

¹ Muhammet YILMAZ ([Orcid ID: 0000-0002-9844-6654](https://orcid.org/0000-0002-9844-6654)), Fatih TOSUNOĞLU ([Orcid ID: 0000-0002-8423-1089](https://orcid.org/0000-0002-8423-1089))
Erzurum Teknik Üniversitesi, Mühendislik ve Mimarlık Fakültesi, İnşaat Mühendisliği Bölümü, Erzurum, Türkiye

² Mehmed Cüneyd DEMİREL ([Orcid ID: 0000-0003-4402-906X](https://orcid.org/0000-0003-4402-906X)), İstanbul Teknik Üniversitesi, İnşaat Mühendisliği, Bölümü, İstanbul, Türkiye

***Sorumlu Yazar/Corresponding Author:** Muhammet YILMAZ, muhammet.yilmaz@erzurum.edu.tr

GİRİŞ

Taşkın olayı, dünya çapında mal ve can kaybına neden olan en tehlikeli doğal afetlerden biridir. Taşkın oluşumu ve değişkenliği birçok faktöre bağlı olduğu için karmaşık ve dinamik bir süreçtir. Son yıllarda, küresel ısınmadan kaynaklı olarak yağış yoğunluğu, sıklığı ve türü değişmiştir. Bununla birlikte şehirleşmenin beraberinde getirdiği geçirimsiz alanlardaki artış ve bitki örtüsündeki azalış taşkın olaylarını artırmaktadır. Bu iki ana faktörün sonucu olarak iklim değişikliği ve insan faaliyetlerinin birleşimi, dünyanın birçok bölgesinde taşkın olaylarında önemli değişikliklere neden olmuştur (Bai ve ark., 2016; Baykal ve Terzi, 2017).

Taşkın olayı nedeniyle meydana gelen ekonomik kayıplar, çevresel zararlar ve sosyal kırılganlık riskini tahmin etmek ve azaltmak için, taşkın debileri ile buna karşılık gelen gerçekleşme olasılığı arasındaki ilişkiyi anlamak temel bir adımdır. Taşkın koruma yapıları ve diğer su yapılarının projelendirilmesi aşamasında taşkın büyüklüğünün, zamanlamasının ve sıklığının güvenilir bir şekilde tahmin edilmesi oldukça önemlidir (Seçkin ve Topçu, 2016).

Literatürde, taşkın debilerini tahmin etmek için kullanılan en yaygın iki yöntem noktasal ve bölgesel taşkın frekans analizi teknikleridir. Noktasal taşkın frekans analizi, bir noktaya ait ölçülmüş veriler için uygun bir olasılık dağılım fonksiyonu (ODF) belirlenmesi işlemidir ve belirli bir alanın tasarım taşkın tahmininde en doğrudan yöntemdir (Rahman ve ark., 2013; Hassan ve ark., 2019). Taşkın frekans analizinin ilk adımı uygun bir ODF seçim işlemidir. Tasarım taşkın debisinin fazla tahmin edilmesi, tasarlanan yapının maliyetinde artışlara neden olabilir, az tahmin edilmesi ise yüksek sel hasarı maliyetlerine neden olabilir (Seçkin ve ark., 2011). Bununla birlikte, her taşkın frekans analizi için ortak bir ODF bulunmadığı için, en iyi dağılımın seçimi hidrologlar için hala bir problemdir (Can ve Tosunoğlu, 2013). Dünya çapında maksimum akımların olasılık modellemesini ele alan birçok çalışma vardır (Wallis ve Wood, 1985; Stedinger ve Griffis, 2008; Haddad ve Rahman, 2011; Mlynski ve ark., 2019). Ülkemizde de taşkın debilerinin tahminleri ile ilgili çeşitli çalışmalar mevcuttur. Haktanır ve ark., (1990), Türkiye akarsu havzalarında bulunan 112 adet akım gözlem istasyonuna ait en büyük debi değerlerine uyan dağılım modellerini (Gumbel, 2 ve 3 parametrelili Log-Normal, 3 parametrelili Gamma, Log-pearson tip III ve Log-Lojistik) incelemişlerdir. Bu dağılımların uygunluklarını Ki-kare ve Kolmogorov-Smirnov testleri ile kontrol etmişlerdir. Sonuç olarak incelenen istasyonlar için tek bir dağılımın önerilemeyeceğini fakat Gumbel dağılımın, Log-Pearson Tip III ve Log-Lojistik gibi dağılımlar kadar güvenilir tahminler vermediği sonucuna varmışlardır. Aşık ve ark., (2004) tarafından yapılan çalışmada Gediz havzası akarsularındaki 11 istasyonda taşkın frekans analizi için Normal, Log-Normal II, Log-Normal III, Gamma II, Log-Pearson Tip III ve Gumbel dağılımları kullanılarak modeller geliştirilmiş ve modellerin uygunlukları Kolmogorov-Smirnov uygunluk testi ile sınanmıştır. Sonuçlara göre bu istasyon verilerine Log-Pearson Tip III dağılımının en iyi uyumu gösterdiği tespit edilmiştir. Saka ve ark., (2013), çalışmada Doğu Karadeniz Havzası akarsularında bulunan 23 istasyona ait maksimum akımların frekans modellemesi için Log-Normal, Gumbel ve Log-Pearson Tip III dağılımlarını kullanmış ve bu dağılımların uygunluğunu Ki-Kare ve Kolmogorov-Smirnov testleri ile kontrol etmişlerdir. Uygunluk test sonuçlarına göre veriler için en uygun dağılımın Log-Normal olduğu sonucuna varmışlardır.

Uygun bir ODF seçildikten sonra dağılımın parametrelerinin tahmin edilmesi gerekmektedir. Uygun bir parametre tahmin yöntemi bulmak frekans analizinde önemli bir adımdır. ODF' lerin parametrelerini tahmin etmenin geleneksel yöntemleri şunlardır; 1) momentler yöntemi (MOM), 2) maksimum olasılık yöntemi (MLM), 3) minimum mesafe yöntemi (MMY) ve 4) olasılık ağırlıklı momentler (OAM). Farklı araştırmacılar, ODF' lerin parametrelerini tahmin etmek için bu yöntemlerden

birini veya kombinasyonunu uygulamıştır (Strupczewski ve ark., (2002); Opere ve ark., 2006; Heidarpour ve ark., 2017). Bir parametre tahmin yönteminin seçilmesi kullanıcıya bağlıdır. Her tahmin yönteminin kendi artıları ve eksileri vardır. Geleneksel yöntemler, kullanılan olasılık dağılım modelin fonksiyonları hakkında birkaç yaklaşım, basitleştirme veya türev bilgisi gerektirebilir ve sonuçlar yerel optimum çözümlere yakınsayabilir. Böylece, bir taşkın frekans analizi problemlerine en uygun çözümleri elde etmek için yeni optimizasyon yöntemlerini keşfetmeye ve uygulamaya daha fazla ihtiyaç duyulmaktadır.

DEoptim algoritması, karmaşık matematiksel işlemlerden kaçınmak ve ayrıca mühendislik ve finans modellerine doğru çözümler sunmak için ilk olarak Storn ve Price (1997) tarafından geliştirilmiştir. DEoptim algoritması R programında mevcuttur (Mullen ve ark., 2011) ve bir çok farklı optimizasyon problemde uygulanmıştır (Borner ve ark., 2007; Cao ve ark., 2009; Mullen ve ark., 2011). Yapılan çalışmalara bakıldığı zaman, ODF' lerin parametrelerini tahmin etmek için DEoptim yönteminin uygulanmasıyla ilgili literatürde yayınlanmış bir çalışma bulunmamaktadır. ODF' lerin parametre tahmini için klasik yöntemleri kullanmayan çok az çalışma yapılmıştır. Örneğin, Kasiviswanathan ve ark., (2017) çalışmasında dağılım parametrelerini belirlemek için genetik algoritmayı kullanmışlardır. Xu ve ark., (2018), Çin'in Hanjiang Nehrinde taşkın frekansı analizi için iki parametrelili Log-Normal dağılımını kullanmış ve dağılımın parametrelerini tahmin etmek için Metropolis-Hastings optimizasyonu (AM-HO) algoritması ile Bayesian yaklaşımını kullanmışlardır.

Bu çalışmanın amacı, taşkın frekans analizinde kullanılan ODF' lerin parametrelerinin tahmini için DEoptim ile klasik MLM ve MOM yöntemlerini karşılaştırmaktır. 'Çalışma ayrıca bootstrap metodunu kullanarak belirli bir dağılım için taşkın miktarlarındaki güven aralıklarını tahmin etmeyi amaçlamaktadır. Bunun için Doğu Karadeniz Havzası' da bulunan 2 adet akım gözlem istasyonuna ait yıllık maksimum akım verileri kullanılmıştır ve bu verilere Gamma, Weibull, Log-Normal, Logistic ve Normal uygulayarak en uygun dağılım fonksiyonu belirlenmeye çalışılmıştır. Kullanılan ODF' lerin parametrelerinin tahmini için DEoptim, MLM ve MOM yöntemi kullanılmıştır. Kolmogorov-Smirnov (KS) testi, ortalama hata karesi (OHK), Kling-Gupta etkinlik katsayısı (KGE), ortalama mutlak hata (OMH), Nash-Sutcliffe etkinlik katsayısı (NSE) ve Akaike bilgi kriteri (AIC) gibi uygunluk testleri en uygun modeli seçmek ve dağılım parametrelerinin tahmininde DEoptim, MLM ve MOM yöntemlerini karşılaştırmak için kullanılmıştır. En uygun olasılık modelleri seçildikten sonra farklı tekerrür periyotlarında (5, 10, 20, 50, 100, 200 ve 500) meydana gelebilecek taşkın debilerini hesaplanmıştır. Son olarak, geri dönüş periyotlarındaki taşkın miktarının % 95 güven aralıkları bootstrap yöntemi ile hesaplanmıştır.

MATERYAL VE METOT

Parametre Tahmin Yöntemleri

ODF' nin parametrelerinin tahmini işlemi, frekans analizi ile tasarım taşkın değerinin belirlenmesinde en etkili faktörlerden biridir. Olasılık dağılımının parametrelerini tahmin etmek için bir dizi yöntem kullanılabilir. Bu çalışmada, önceki çalışmalarda en çok benimsenen yöntemlerden MLM ve MOM adlı klasik tahmin yöntemleri kullanılmıştır. Daha sonra, DEoptim yöntemi tanıtılmış ve dağılım parametrelerini tahmin etmek için kullanılmıştır.

MLM, özellikle büyük örnekli veriler için frekans analizinde kullanılan en etkili parametre tahmin yöntemlerinden biridir. Bu metot, elde bulunan gözlenmiş serinin gerçekleşmesi durumunda olabilirliği en yüksek seri değerleri elde edilmesi prensibine dayanmaktadır. Gözlenmiş her bir değer için ODF ler yazılır ve hepsi birbiriyle çarpılırsa sonuçta bu serinin olabilme ihtimali elde edilmiş olur. Birbirinden bağımsız örnek serilerden olabilme fonksiyonunu maksimum yapan parametre değerleri bu

dağılımda orjinal seriye en uygun parametreler olmaktadır. Belirli bir örneğin olabilirlik fonksiyonun maksimum yapan değerler ile logaritması alınmış olabilirlik fonksiyonun maksimum değerleri birbirinin aynısıdır. Bundan dolayı çarpımın türevini almak yerine aşağıdaki denklemde verilen toplamlar fonksiyonu olan $Log (ODF)$ ' nin türevini almak daha uygun olur.

$$Log (ODF) = \sum_{i=1}^N f(\varepsilon_i, x_1, x_2, \dots \dots \dots x_m) \quad (1)$$

Burada, m parametresi bilinmemektedir. Eşitliğin kısmi türevi alınır ve elde edilen denklemlerin hepsi sıfıra eşitlenir. Denklemleri sıfır yapan m parametre değerleri MLM yöntemine göre aranan tahminler olur (Seçkin, 2009).

Momentler yöntemi kolay bir yöntem olduğu için hidrolojik çalışmalarda sıklıkla kullanılmaktadır. Bu metot da, $Y_1 \dots \dots Y_N$ den oluşan bir örnek düşünelim bir a noktasında N . Örneğin momenti aşağıdaki gibi ifade edilebilir;

$$M_n = \frac{1}{N} \sum_{i=1}^N (Y_i - a)^n \quad (2)$$

Denklemde N örnek sayısını gösterir. Merkezi moment dikkate alınır, a örneğin ortalama değeri ile değiştirilir. Herhangi bir dağılım (veya model) aşağıdaki gibi ifade edilen örneğe uydurulur;

$$Y_i = f(Y_{i-1}, Y_{i-2}, \dots \dots, \alpha_1, \alpha_2, \dots \dots \dots \alpha_m) + \varepsilon_t \quad (3)$$

Burada, ε modelin hatasını gösterir, $\alpha_1, \alpha_2, \dots \dots \dots \alpha_m$ ise modelin parametreleridir. Model parametreleri aşağıdaki denklem ile bulunur.

$$M_n = \frac{1}{N} \sum_{i=1}^N (Y_i - a)^n = \frac{1}{N} \sum_{i=1}^N (f(Y_{i-1}, Y_{i-2}, \dots \dots, \alpha_1, \alpha_2, \dots \dots \dots \alpha_m) - a)^n \quad (4)$$

m , parametre sayısını ifade eder (fatih hoca makale).

Differential Evolution (DEoptim) Storn ve Price (1997) tarafından sunulan bir sezgisel arama algoritmasıdır. Karmaşık optimizasyon problemlerine güvenilir çözümler sağlamak için global bir optimizasyon algoritması olarak geliştirilmiştir (Storn ve Price, 1997). DEoptim minimizasyon problemlerinin çözümünde güçlü bir araca sahip olduğu için farklı alanlarda araştırmacılar tarafından uygulanmıştır (Borner ve ark., (2007); Cao ve ark., (2009); Jiang ve Ma (2018)). DEoptim, R istatistik yazılımında DEoptim paketi olarak mevcuttur. Bu çalışmada, DEoptim fonksiyonu her olasılık dağılım fonksiyonunun teorik ve ampirik kümülatif dağılım fonksiyonları (CDF) arasındaki hataları en aza indirmek için kullanılmıştır. Bu işlem için amaç fonksiyonu olarak NSE etkinlik kriteri kullanılmıştır. Bilindiği gibi NSE' nin değeri 1'e ne kadar yakınsa modelin doğruluğu da o kadar yüksek olur. Fakat DEoptim algoritması ise NSE değerini 0 yaklaşmasında iyi bir model elde edileceğini gösterir. Bu yüzden NSE fonksiyonunu aşağıdaki denklemdeki gibi revize edilmiştir.

$$NSE_{mutlak} = \frac{\left| \sum_{i=1}^n (q_s - q_g)^2 \right|}{\left| \sum_{i=1}^n (q_g - \bar{q}_g)^2 \right|} \quad (5)$$

Burada, q_g ve q_s olasılık dağılım fonksiyonlarının teorik ve ampirik CDF ' sini temsil eder. Sezgisel arama algoritması, amaç işlevinin en iyi değerini belirlemek için her modelde 3000 kez çalıştırılmıştır.

Olasılık Dağılım Fonksiyonları ve Uygunluk Testleri

Bu çalışmada frekans analizinde yaygın olarak kullanılan Gamma, Weibull, Log-Normal, Logistic ve Normal dağılımları kullanılmıştır. Bu dağılımlar için CDF ler Çizelge 1 de (Tosunoğlu ve Kahya, 2020) verilmiştir. En uygun dağılımı belirlemek amacıyla Kolmogorov–Smirnov (KS) testi, ortalama hata karesi (OHK), Kling–Gupta etkinlik katsayısı (KGE), ortalama mutlak hata (OMH), Nash–Sutcliffe etkinlik katsayısı (NSE) ve Akaike bilgi kriteri (AIC) gibi uyum iyiliği kriterleri uygulanmıştır. Bu kriterler aşağıdaki denklemlerde verildiği gibi hesaplanmıştır.

$$KS = \max |Fx_i - Fx_i^*| \quad (6)$$

Burada Fx_i belirlenen olasılık dağılımına göre elde edilen CDF' yi temsil eder. Fx_i^* ise ampirik CDF dir.

$$OHK = \sqrt{OKH} = \sqrt{\frac{1}{n} \sum_{i=1}^n (q_g - q_s)^2} \quad (7)$$

$$OMH = \frac{1}{n} \sum_{i=1}^n |q_g - q_s| \quad (8)$$

$$NSE = 1 - \frac{\sum_{i=1}^n (q_s - q_g)^2}{\sum_{i=1}^n (q_g - \bar{q}_g)^2} \quad (9)$$

$$AIC = n \log(OKH) + 2p \quad (10)$$

Burada, n veri serisinin uzunluğudur. q_g orijinal zaman serisi ve q_s ise tahmin değerleridir. p kullanılan dağılımdaki parametre sayısıdır.

$$KGE = 1 - \sqrt{(r - 1)^2 + \left(\frac{\sigma_s}{\sigma_g} - 1\right)^2 + \left(\frac{\mu_s}{\mu_g} - 1\right)^2} \quad (11)$$

Bu denklemden, r gözlem ve tahmin değerleri arasındaki lineer korelasyon değeridir. σ_g ve σ_s sırasıyla gözlem ve tahminlerin standart sapma değeridir. μ_g ve μ_s sırasıyla gözlem ve tahminlerin ortalama değeridir.

Olasılık dağılım fonksiyonlarının tümü için bu değerler hesaplandıktan sonra en düşük OHK, OMH, KS ve AIC değerlerine ve en yüksek NSE ve KGE değerlerine sahip model en uygun modeldir.

Çeşitli Dönüş Periyotları İçin Yapılan Akım Tahminleri ve Bootstrap Yöntemi Kullanılarak Güven Aralıklarının Belirlenmesi

Taşkın frekans analizinin amaçlarından biri de taşkın miktarını ve belirli bir geri dönüş dönemi için belirsizliğini tahmin etmektir. Uygun dağılım seçildikten sonra, farklı geri dönüş dönemlerine (T) karşılık gelen taşkın debi değerleri tahmin edilebilir. Belirli bir geri dönüş süresi için debi değeri, aşağıdaki denklemle basitçe elde edilebilir.

$$F = F(y_T) = 1 - \frac{1}{T} \quad (12)$$

Dağılım fonksiyonun ($F(y_T)$) tersi alınarak $y_T=y(F)$, y_T doğrudan elde edilebilir.

Taşkın debilerinin hesaplanmasından sonra tahminlerin güven aralıkları da değerlendirilir. Farklı görülme sıklıklarına ait taşkın debi değerlerinin güven aralığını tahmin etmek için bootstrap yöntemini kullanıldı (Davison v.d., 1986). Bootstrap yöntemi, güven aralığı hesaplama gibi parametrik olmayan tahmin problemlerinde kullanılan güvenilir ve basit bir yöntemdir.

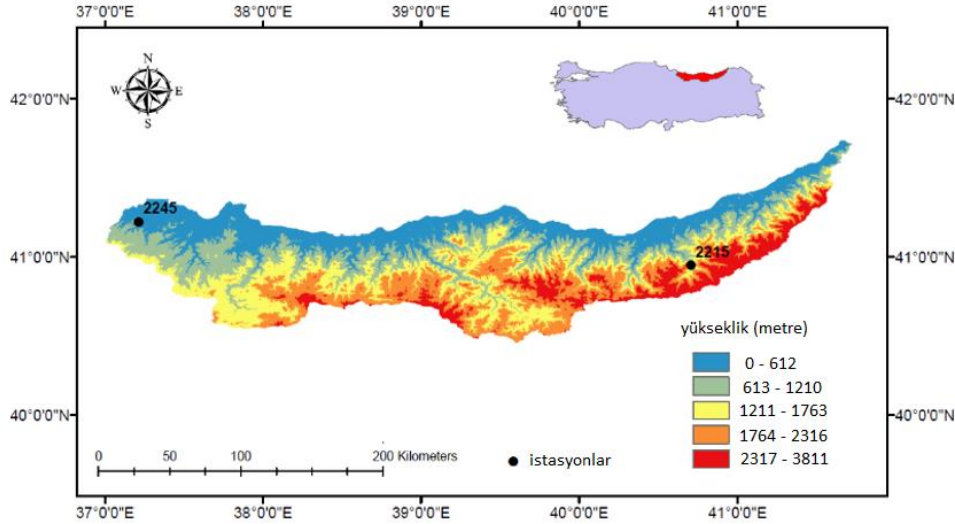
Çizelge 1. Kullanılan olasılık dağılım fonksiyonlarının genel tanımı (Can ve Tosunoğlu, 2013)

Dağılım Tipi	Olasılık Yoğunluk Fonksiyonu	Parametreler
Weibull (WBL)	$f(x) = \frac{k}{\alpha} \left(\frac{x}{\alpha}\right)^{k-1} \exp\left[-\left(\frac{x}{\alpha}\right)^k\right]$	k = şekil parametresi (k>0) α = ölçek parametresi (α>0)
II Parametrelili Lognormal (LN2)	$f(x) = \frac{1}{x\sigma_y\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma_y^2}(\log x - \mu_y)^2\right]$	μ _y = ortalama log. (μ _y >0) σ _y = standart sapma log. (σ _y >0)
Gamma (GAM)	$f(x) = \frac{x^{k-1}}{\alpha^k \Gamma(k)} \exp\left[-\frac{x}{\alpha}\right]$	k = şekil parametresi (k>0) α = ölçek parametresi (α>0) Γ = gamma fonksiyonu
Normal (NORM)	$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$	μ = ortalama (konum parametresi) σ = standart sapma (ölçek parametresi) (σ>0)
Logistic (LOGIS)	$f(x) = \frac{1}{\alpha} \exp\left(\frac{x-\xi}{\alpha}\right) \left[1 + \exp\left(\frac{x-\xi}{\alpha}\right)\right]^{-2}$	ξ = konum parametresi α = ölçek parametresi (α>0)

Çalışma Alanı ve Veriler

Doğu Karadeniz Havzası; Türkiye'nin kuzey doğu kıyısında bulunmaktadır. Havza güneyde Doğu Karadeniz Dağları ve kuzeyde Karadeniz ile çevrilidir. Havza alanı Gürcistan sınırından başlayıp Samsun şehrinin doğusunda Terme çayı ile son bulmaktadır. Toplam havza alanı yaklaşık olarak 24077 km²'yi kapsamaktadır (Bayazıt ve Avcı, 1997). Havza alanı bol miktarda yağış almaktadır ve Türkiye'nin dört mevsim yağış alan tek bölgesidir. Bu bölgenin çalışma alanı olarak seçilmesinin nedenleri; havzanın toprak yapısından dolayı yoğun ve ani yağışlardan sonra hızlı bir şekilde sel olayının meydana gelmesi ve havzanın keskin vadilere ve önemli su potansiyeline sahip olması nedeniyle orta ve küçük hidroelektrik potansiyeli açısından büyük avantajlara sahip olmasıdır. Havzada meydana gelen ani sel ve bunun yol açtığı enkaz can kaybı, yaralanma ve maddi zararlara neden olmaktadır. Fakat aynı zamanda havza, hidroelektrik enerji ihtiyacını sağlaması bakımından Türkiye için oldukça önemlidir. 2018 yılında Türkiye'nin hidroelektrik kurulu gücünün yaklaşık% 10'unu sağlamıştır. Son zamanlarda HES vb. birçok hidrolik yapının inşasının tamamlanmasından sonra, bölge Türkiye'nin ekonomik olarak uygulanabilir hidroelektrik enerji potansiyelinin % 18,8'ini karşılayabilecektir (Kangal ve Akçay, 2019).

Bu çalışmada Devlet Su İşleri'nin işletmiş olduğu 2 adet akım gözlem istasyonuna (AGİ) ait yıllık maksimum akım değerleri kullanılmıştır. Bölgede bu istasyonların seçilmelerinin nedeni, uzun gözlem sürelerine sahip oldukları için istatistiki açıdan inceleme yapılmasına imkan sağlamalarıdır. Çalışmada kullanılan AGİ'lerin yerleri ve çalışma alanı Şekil 1 de verilmiştir. Bu AGİ'lere ait bilgiler ise Çizelge 2'de verilmiştir.



Şekil 1. Çalışma alanı ve kullanılan istasyonlar

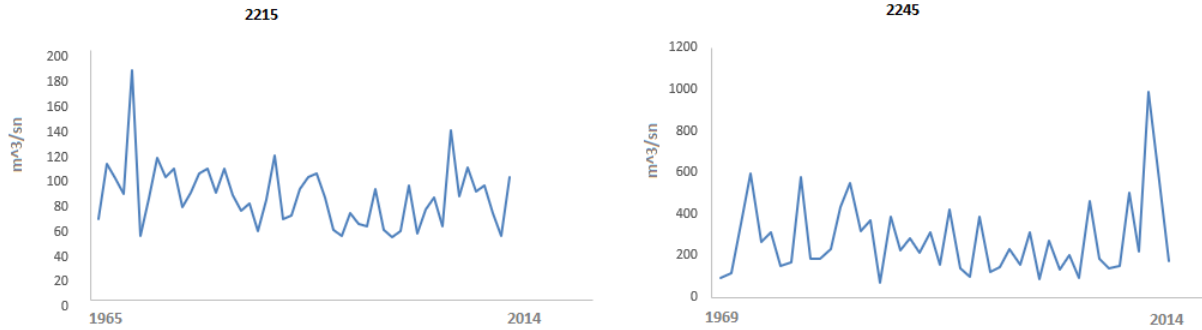
Çizelge 2. Çalışmada kullanılan istasyonlara ait bilgiler

İstasyon Adı	No	Gözlem Aralığı (Yıl)	Yağış Alanı (km ²)	Kot (m)
Dereköy	2215	1965-2014	445.2	942
Gökçeli	2245	1969-2014	232.80	66.00

BULGULAR VE TARTIŞMA

Kullanılan istasyonların maksimum akımlarına seçilen olasılık dağılımlarını uygulamadan önce verilerin durağanlığı araştırılmalıdır. Bu amaçla verilere parametrik olmayan Mann-Kendall test uygulanmıştır ve sonuçlar Çizelge 3 de verilmiştir. Sonuçlara göre hesaplanan z değerinin mutlak değeri %5 anlam seviyesindeki kritik değer olan 1.96 dan küçük olduğundan dolayı verilerde istatistiki açıdan pozitif veya negatif trend gözlenmediği için veriler durağan kabul edilmiştir. Ayrıca veriler için çizilen zaman serileri ve korelogramlar Şekil 2 ve 3 de verilmiştir. Şekil 3 de görüldüğü gibi otokorelasyon sayıları %95 güven sınırları içinde kalmaktadır. Verilerin bağımsızlığı ve durağanlığı kanıtlandıktan sonra, zaman serilerine GAM, WBL, LN2, LOGIS ve NORM dağılımları uygulanmıştır. Dağılımın parametrelerini tahmin etmek için MLM, MOM ve DEoptim yöntemleri kullanılmıştır. Dağılımların performanslarını ve tahmin yöntemlerini karşılaştırmak için KS, OHK, KGE, OMH, NSE ve AIC testleri dikkate alınmıştır. Verilere en iyi uyan dağılımlar ve onların parametreleri ve en uygun parametre yöntemi Çizelge 4 de verilmiştir. Sonuçlara göre 2215 istasyonu için WBL dağılımı, 2245 istasyonu için LN2 dağılımı en uygun dağılımlardır. Uyum iyiliği test sonuçlarına göre, DEoptim yöntemi her iki istasyon için de olasılık dağılım fonksiyonları parametrelerini tahmin etmede MLM ve MOM yöntemlerinden daha iyi performans göstermektedir. Kullanılan 5 dağılımın her biri için hesaplanan uyum iyiliği test sonuçları Çizelge 5 de verilmiştir. Test sonuçlarına göre, GAM dağılımı 2215 istasyonu için MLM ve MOM yöntemlerine göre en iyi dağılımdır. Fakat DEoptim yöntemine göre ise WBL dağılımı en iyi dağılım olarak gözükmektedir. DEoptim yöntem MLM ve MOM yöntemlerine göre daha iyi test değerleri ile üstün sonuçlar sağladığından dolayı bu istasyon için en iyi dağılım WBL dağılımı seçilmiştir. 2245 istasyonunda ise her üç tahmin yöntemine göre de en iyi dağılım LN2 seçilmiştir. Fakat, Çizelge 5 de görüleceği üzere DEoptim yöntemi daha iyi test değerleri göstermiştir. Parametre tahmin yöntemlerinin performansının görsel olarak değerlendirilmesi için teorik kümülatif dağılım fonksiyonlarının ampirik olanlarla karşılaştırılmasını gösteren Olasılık-Olasılık (P-P) grafikleri

kullanılmıştır (Şekil 4). DEoptim yöntemi tahminlerinin, MLM ve MOM yöntemlerinden daha iyi bir uyum sağladığı görülmektedir.



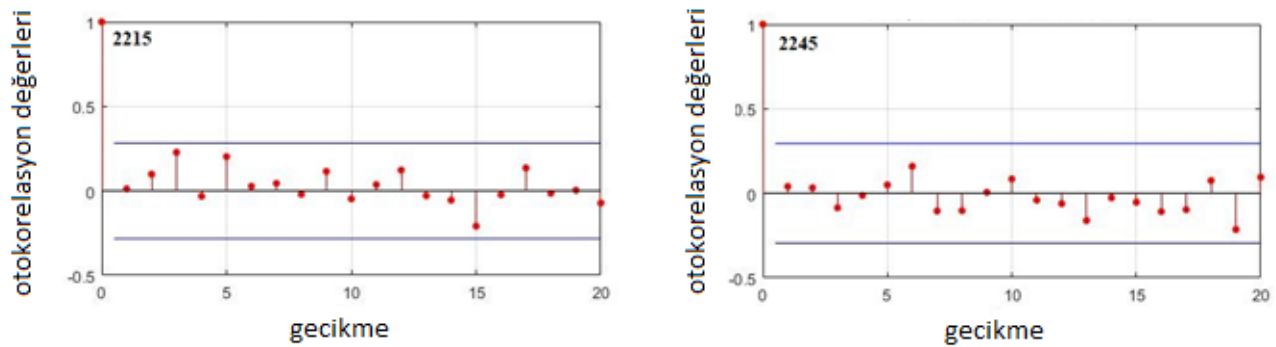
Şekil 2 Kullanılan istasyonlar için zaman serileri

Çizelge 3. Çalışmada kullanılan veriler için Mann Kendall test sonuçları

No	MK Test İstatistiği	Z değeri	Trend
2215	-216	-1.80	–
2245	-36	-0.33	–

Çizelge 4 Verilere uyan en iyi dağılımlar ve tahmin yöntemi

No	En iyi dağılım	En iyi tahmin yöntemi
2215	WBL $\alpha = 89.73079$ $k = 3.915$	DEoptim
2245	LN2 $\mu_y = 4.38353$ $\sigma_y = 0.30439$	DEoptim



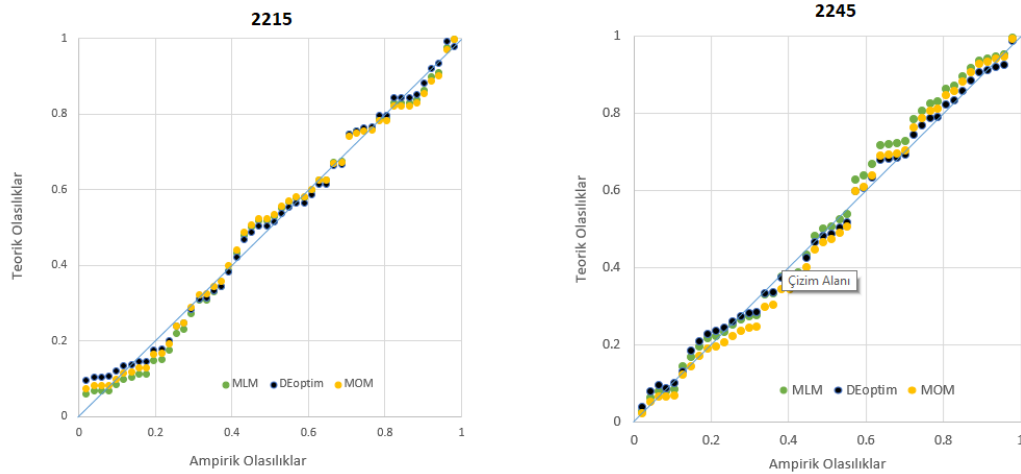
Şekil 3. İstasyonlara ait otokorelasyon değerleri

En iyi olasılık dağılımının seçilmesinden sonra, taşkın frekansı analizinin ana amaçlarından biri, hidrolojik uygulamada çeşitli tekerrür aralıkları için taşkın miktarlarını tahmin etmektir. Bu çalışmada, sırasıyla 0,2, 0,1, 0,05, 0,02, 0,01, 0,005 ve 0,002 olasılık düzeyine eşit olan 5, 10, 20, 50, 100, 200 ve 500 yıllık geri dönüş dönemleri için taşkın debileri hesaplanmış ve Çizelge 6' da verilmiştir. Özellikle daha yüksek geri dönüş dönemleri için, taşkın miktarının güven aralığı, yapıların yönetimi, planlaması ve tasarımında önemli bilgiler sağlayabilir. Çalışma da güven aralıklarını elde etmek için bootstrap yöntemi kullanılmıştır. Her bir geri dönüş periyodu için hesaplanan güven aralıkları Çizelge 6 'da

sunulmuştur. DEoptim yaklaşımı geleneksel MLM ve MOM dan daha üstün performans gösterdiğinden, taşkın miktarları ve güven aralıkları, yalnızca DEoptim tarafından belirlenen dağılımlar için hesaplanmıştır. Örneğin tasarımcılar veya mühendisler, 2215 istasyonu için 100 yıllık geri dönüş süresine karşılık gelen taşkın miktarı ile ilgilenebilirler. Bu miktar WBL dağılımı kullanılarak 132.54 m³/s olarak tahmin edilmiştir. Bootstrap yöntemine göre ise %95 güven aralıkları 120.79 – 144.66 olarak hesaplanmıştır.

Çizelge 5. Kullanılan istasyonlar için model seçim kriter sonuçları

MLM						
2215		GAM	WBL	LN2	LOGIS	NORM
	KS	0.066	0.104	0.072	0.065	0.072
	RMSE	0.031	0.043	0.036	0.031	0.035
	KGE	0.975	0.887	0.961	0.968	0.950
	MAE	0.026	0.034	0.030	0.026	0.031
	NSE	0.988	0.977	0.984	0.988	0.985
	AIC	-147.243	-132.318	-140.373	-146.250	-141.471
MOM						
2215		GAM	WBL	LN2	LOGIS	NORM
	KS	0.057	0.288	0.071	0.069	0.072
	RMSE	0.029	0.167	0.034	0.037	0.035
	KGE	0.988	0.562	0.981	0.958	0.950
	MAE	0.024	0.150	0.028	0.031	0.031
	NSE	0.989	0.650	0.985	0.983	0.985
	AIC	-149.504	-73.629	-142.524	-139.622	-141.471
DEoptim						
2215		GAM	WBL	LN2	LOGIS	NORM
	KS	0.056	0.077	0.063	0.081	0.073
	RMSE	0.029	0.026	0.033	0.028	0.027
	KGE	0.989	0.991	0.987	0.987	0.990
	MAE	0.024	0.021	0.026	0.023	0.022
	NSE	0.989	0.991	0.987	0.990	0.991
	AIC	-149.531	-154.001	-144.523	-150.810	-153.644
MLM						
2245		GAM	WBL	LN2	LOGIS	NORM
	KS	0.120	0.096	0.077	0.161	0.148
	RMSE	0.058	0.045	0.034	0.080	0.076
	KGE	0.914	0.960	0.935	0.914	0.918
	MAE	0.045	0.038	0.028	0.066	0.065
	NSE	0.957	0.974	0.986	0.919	0.927
	AIC	-109.550	-119.481	-131.417	-96.840	-98.916
MOM						
2245		GAM	WBL	LN2	LOGIS	NORM
	KS	0.112	0.094	0.074	0.161	0.148
	RMSE	0.055	0.045	0.035	0.080	0.076
	KGE	0.919	0.955	0.930	0.914	0.918
	MAE	0.046	0.038	0.029	0.066	0.065
	NSE	0.962	0.974	0.985	0.919	0.927
	AIC	-111.754	-119.763	-130.026	-96.840	-98.916
DEoptim						
2245		GAM	WBL	LN2	LOGIS	NORM
	KS	0.062	0.080	0.045	0.117	0.113
	RMSE	0.031	0.038	0.023	0.056	0.057
	KGE	0.984	0.974	0.988	0.966	0.966
	MAE	0.027	0.033	0.019	0.049	0.050
	NSE	0.988	0.982	0.994	0.961	0.960
	AIC	-134.351	-126.575	-147.441	-111.186	-110.646



Şekil 4. 2215 ve 2245 istasyonları için Olasılık-Olasılık (P-P) grafikleri

Çizelge 6. En iyi parametrelili dağılımlar için taşkın tahmin miktarları ve %95 güven aralıkları

No	Dağılım	İstatistik	Dönüş Periyodu (Yıl)						
			5	10	20	50	100	200	500
2215	WBL	Alt	94.50	102.99	109.71	116.56	120.79	124.51	129.08
		Fit	101.33	111.04	118.76	127.13	132.54	137.37	143.09
		Üst	108.14	118.86	128.14	138.05	144.66	150.60	157.84
2245	LN2	Alt	318.45	414.59	508.90	638.25	742.67	848.06	1002.68
		Fit	400.85	538.31	686.72	903.24	1084.30	1281.64	1569.53
		Üst	498.09	687.07	901.40	1233.26	1505.49	1843.87	2322.04

SONUÇ

Bu çalışma, Doğu Karadeniz Havzası'ndaki iki akım gözlem istasyonuna ait yıllık maksimum akım değerlerini kullanarak taşkın frekansı analizi için uygun bir olasılık dağılımı ve parametre tahmin yönteminin seçimini sunmaktadır. En uygun olasılık dağılım modelini ve parametre tahmin prosedürünü belirlemek için yaygın olarak kullanılan altı model seçim kriteri (KS, OHK, KGE, OMH, NSE ve AIC) kullanılmıştır. Çalışmada taşkın frekans analizi GAM, WBL, LN2, LOGIS ve NORM, dağılımları kullanılarak yapılmıştır. DEoptim yöntemi kullanılan dağılımların parametrelerini tahmin etmek için kullanılmış ve klasik MLM ve MOM yöntemleri ile karşılaştırılmıştır. Ayrıca taşkın miktarlarının güven aralıklarının belirlenmesi için DEoptim yöntemi ile belirlenen en iyi dağılımlar kullanılmıştır. Çeşitli uyum iyiliği kriterleri değerlendirildikten sonra sonuçlar, DEoptim yönteminin klasik MLM ve MOM'dan daha iyi performansa sahip olduğunu göstermiştir. Bu nedenle, önerilen yaklaşımla daha doğru taşkın frekans modelleri ve taşkın tahminleri elde edilebilir. Ek olarak, 2215 istasyonu için en iyi dağılımının WBL, 2245 istasyonu için ise en iyi dağılımın LN2 olduğu sonucuna varılabilir. Bu çalışmanın taşkın frekans analizinde kullanılacak yeni bir parametre tahmin yöntemi ve gelecekteki çalışmalar için bazı önemli çıkarımlar sağladığı düşünülmektedir.

TEŞEKKÜR

Çalışmada kullanılan akış verilerini sağlayan Devlet Su İşleri Genel Müdürlüğü'ne içtenlikle teşekkür ederiz. DEoptim kodu <https://cran.r-project.org/web/packages/DEoptim/> adresinde mevcuttur.

Çıkar Çatışması

Makale yazarları aralarında herhangi bir çıkar çatışması olmadığını beyan ederler.

Yazar Katkısı

Yazarlar makaleye eşit oranda katkı sağlamış olduklarını beyan eder.

KAYNAKLAR

- Aşık Ş, Akkuzu, E, Karataş BS, 2004. Gediz Havzası Akarsuları Yıllık Toplam Akımlarına En Uygun Olasılık Dağılım Modelinin Belirlenmesi. Ege Üniv. Ziraat Fak. Derg, 41 (3):177-184.
- Bai P, Liu X, Liang K, Liu C, 2016. Investigation of changes in the annual maximum flood in the Yellow River basin, China. Quaternary International, 392: 168-177
- Bayazit M, Avcı I. 1997. Water Resources of Turkey: Potential, Planning, Development and Management. Water Resources Development, 13(4): 443-452.
- Baykal T, Terzi Ö, 2017. Küçük Aksu Çayı Taşkın Frekans Analizi. Cumhuriyet Sci. J, 38(4): 639-646
- Börner J, Higgins SI, Kantelhardt J, Scheiter S, 2007. Rainfall or price variability: What determines rangeland management decisions? A simulation-optimization approach to South African savannas. Agricultural Economics, 37(2-3): 189-200.
- Can I, Tosunoğlu F, 2013. Estimating T-year flood confidence intervals of rivers in Coruh basin, Turkey. Journal of Flood Risk Management, 6(3): 186-196.
- Cao R, Vilar JM, Devia A, 2009. Modelling consumer credit risk via survival analysis. Sort-Statistics and Operations Research Transactions, 33(1): 3-30.
- Davison AC, Hinkley DV, Schechtman E, 1986. Efficient Bootstrap Simulation. Biometrika, 73(3), 555-566.
- Rahman AS, Rahman A, Zaman MA, Haddad K, Ahsan A, Imteaz M, 2013. A study on selection of probability distributions for at-site flood frequency analysis in Australia. Natural Hazards, 69(3): 1803-1813.
- Haktanir T, Özcan Z, Çapar ÖF, 1990. Türkiye akarsularının taşkın pikleri frekans analizi . Su Mühendisliği Problemleri Kongresi, 1-4 Ocak 1990, Antalya.
- Hassan MU, Hayat O, Noreen Z, 2019. Selecting the best probability distribution for at-site flood frequency analysis; a study of Torne River. Sn Applied Sciences, 1(12).
- Heidarpour B, Saghafian B, Yazdi J, Azamathulla HM, 2017. Effect of Extraordinary Large Floods on at-site Flood Frequency. Water Resources Management, 31(13): 4187-4205.
- Hu L, Nikolopoulos EI, Marra F, Anagnostou EN, 2020. Sensitivity of flood frequency analysis to data record, statistical model, and parameter estimation methods: An evaluation over the contiguous United States. Journal of Flood Risk Management, 13(1).
- Jiang ZH, MA WC, 2018. Integrating Differential Evolution Optimization to Cognitive Diagnostic Model Estimation. Frontiers in Psychology, 9.
- Kankal M, Akçay F, 2019. Doğu Karadeniz Havzasi Hidroelektrik Enerji Durumunun İncelenmesi. Ömer Halisdemir Üniversitesi Mühendislik Bilimleri Dergisi, 8(2): 892-901.
- Kasiviswanathan KS, HE, JX, TAY JH, 2017. Flood frequency analysis using multi-objective optimization based interval estimation approach. Journal of Hydrology, 545: 251-262.
- Mlynski D, Walega A, Stachura T, Kaczor G, 2019. A New Empirical Approach to Calculating Flood Frequency in Ungauged Catchments: A Case Study of the Upper Vistula Basin, Poland. Water, 11(3).
- Mullen KM, Ardia D, Gil DL, Windover D, Cline J, 2011. DEoptim: An R Package for Global Optimization by Differential Evolution. Journal of Statistical Software, 40(6): 1-26.
- Opere AO, Mkhanda S, Willems P, 2006. At site flood frequency analysis for the Nile Equatorial basins. Physics and Chemistry of the Earth, 31(15-16): 919-927.

- Haddad K, Rahman A, 2011. Selection of the best fit flood frequency distribution and parameter estimation procedure: a case study for Tasmania in Australia. *Stochastic Environmental Research and Risk Assessment*, 25(3): 415-428.
- Saka F, Yüksek Ö, Yıldırım MS, Baş K, Anılan T, Kaplan H, 2013. Yıllık Maksimum Akımların Bazı Olasılık Dağılımlarına Uygunluğunun Ki-Kare Ve Kolmogorov-Smirnov Testleriyle Belirlenmesi. *Taşkın ve Heyelan Kongresi*, 24-26 Ekim 2013, Trabzon.
- Seçkin N, 2009. L-Momentlere Dayalı Gösterge-Sel Metodu ile Bölgesel Taşkın Frekans Analizi. *Çukurova Üniversitesi Fen Bilimleri Enstitüsü, Doktora Tezi (Basılmış)*.
- Seckin N, Haktanır T, and Yurtal R, 2011. Flood frequency analysis of Turkey using L-moments method. *Hydrological Processes*, 25(22): 3499-3505.
- Seckin N, Topçu E, 2016. Adana ve çevre illerde gözlenen yıllık maksimum yağışların bölgesel frekans analizi. *Journal of the Faculty of Engineering and Architecture of Gazi University* 31(4): 1049-1062
- Stedinger JR, Griffis, VW, 2008. Flood frequency analysis in the United States: Time to update. *Journal of Hydrologic Engineering*, 13(4): 199-204.
- Storn R, Price K, 1997. Differential evolution - A simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, 11(4): 341-359.
- Strupczewski WG, Singh VP, Weglarczyk S, 2002. Asymptotic bias of estimation methods caused by the assumption of false probability distribution. *Journal of Hydrology*, 258(1-4): 122-148.
- Tosunoğlu F, Kahya E, 2020. Definition of the best probability distribution functions for annual minimum flows in the rivers of the Upper Euphrates River Basin, Turkey. *IOP Conference Series: Materials Science and Engineering*.
- Wallis JR, Wood EF, 1985. Relative Accuracy of Log Pearson-Iii Procedures. *Journal of Hydraulic Engineering-Asce*, 111(7): 1043-1056.
- Xu WT, Jiang C, Yan L, Li L, Liu S, 2018. An Adaptive Metropolis-Hastings Optimization Algorithm of Bayesian Estimation in Non-Stationary Flood Frequency Analysis. *Water Resources Management*, 32(4): 1343-1366.