

## Shannon entropisi temelli düzeltmeli çeşitlilik ölçümleri ve Pla kestiricisinin üst güven sınır değerini kullanarak gerçek tür çeşitliliğinin hesaplanması

Kürşad Özkan<sup>a,\*</sup> 

**Özet:** Bu çalışma Pla kestiricisinin düzeltilmiş üst güven değeri ( $\hat{H}^e + 1,96\sigma_{\hat{H}}^*$ ) ve Shannon entropisi temelli düzeltmeli tür çeşitlilik (SDÇ) ölçümleri (Jackknife indisi ( $JK$ ), düzeltilmiş en büyük olabilirlik kestiricisi ( $MLEbc$ ), Chao ve Shen düzeltmeli çeşitlilik eşitliği ( $\hat{H}_{CS}$ ), Özkan düzeltmeli çeşitlilik kestiricisi ( $\hat{H}_O$ ), Bilgi teorisi temelli düzeltmeli çeşitlilik indisi ( $\hat{H}_{CWJ}$ ), Miller sapma düzeltme eşitliği ( $\hat{H}_{Miller}$ ) ve Grassberger sapma düzeltme eşitliği ( $\hat{H}_{\psi}$ )) kullanılarak gerçek tür çeşitliliğinin nasıl hesaplanabileceği bilgisini vermek amacıyla gerçekleştirilmiştir. Bu doğrultuda ilk olarak 5 toplumun  $JK$ ,  $MLEbc$ ,  $\hat{H}_{CS}$ ,  $\hat{H}_O$ ,  $\hat{H}_{CWJ}$ ,  $\hat{H}_{Miller}$ ,  $\hat{H}_{\psi}$  ve  $\hat{H}^e + 1,96\sigma_{\hat{H}}^*$  değerleri hesaplanmıştır. Daha sonra her bir toplum için hesaplanan  $\hat{H}^e + 1,96\sigma_{\hat{H}}^*$  değerinin altında kalan indis değerleri kullanılarak örnek toplumların gerçek tür çeşitlilik değerleri belirlenmiştir.

**Anahtar kelimeler:** Biyolojik çeşitlilik, Shannon indisi, Nadir tür, Negatif sapma, Düzeltilmiş tür çeşitliliği, Eksik envanter

## Computing true species diversity using Shannon entropy-based bias-corrected measures and upper confidence limit value of Pla estimator

**Abstract:** This study was conducted to illustrate how to estimate true species diversity using upper confidence limit value of Pla unbiased estimator ( $\hat{H}^e + 1,96\sigma_{\hat{H}}^*$ ) and Shannon entropy-based bias-corrected diversity (SCD) measures (i.e., Jackknife index ( $JK$ ), bias-corrected maximum likelihood estimator ( $MLEbc$ ), Chao and Shen bias-corrected estimator ( $\hat{H}_{CS}$ ), Özkan bias-corrected estimator ( $\hat{H}_O$ ), information theory-based bias-corrected estimator ( $\hat{H}_{CWJ}$ ), Miller index ( $\hat{H}_{Miller}$ ) and Grassberger index ( $\hat{H}_{\psi}$ )). In this regards, firstly  $JK$ ,  $MLEbc$ ,  $\hat{H}_{CS}$ ,  $\hat{H}_O$ ,  $\hat{H}_{CWJ}$ ,  $\hat{H}_{Miller}$ ,  $\hat{H}_{\psi}$  and  $\hat{H}^e + 1,96\sigma_{\hat{H}}^*$  values of five communities were calculated. Next for each community, true diversity values are calculated using the SCD values below  $\hat{H}^e + 1,96\sigma_{\hat{H}}^*$  value.

**Keywords:** Biodiversity, Shannon index, Rare species, Negative bias, Unbiased species diversity, Incomplete inventory


### 1. Giriş


Biyolojik çeşitlilik ekosistemlerin esnekliği, sağlığı, dinamizmi, verimliliği ve karbon depolama kapasitesi için temel göstergelerden biridir. Bu yüzden biyolojik çeşitlilik doğa bilimlerinin en önemli konuları içerisinde yer almaktadır. Peet (1974) tarafından ifade edildiği üzere biyolojik çeşitlilik, indisler kullanılarak hesaplanmaktadır. Biyolojik çeşitliliğin hesaplanması için geliştirilen veya önerilen birçok indis bulunmakta olup, bunların sayısı gün geçtikçe artmaktadır. Bu indisler tür çeşitlilik indisleri başlığı altında genel olarak, tür zenginlik indisleri, heterojenlik indisleri ve tür bolluk modelleri olmak üzere üç gruba ayrılmaktadır (Özkan, 2020). Tür çeşitliliğinin, heterojenlik indisleri ve tür bolluk modelleri ile hesaplanmasında temel girdiler tür sayısı ve türlere ait bolluk değerleridir. Tür sayısı ve onların bolluk değerine dayalı çeşitlilik hesaplamasında en fazla kullanılan indis ise Shannon entropisidir (Shannon, 1948).


Bir toplumu temsil etmek üzere örnekleme yapıldığında veya örnek alan alındığında; bu örnek alanın o toplumu tam olarak temsil ettiğini kabul etmek yanlıgılara sebep olabilir. Diğer bir ifade ile bir toplumu temsil etmek üzere örnekleme

veya örnek alan alındığında büyük ihtimalle bir negatif sapma söz konusudur. Böyle durumlar için tür çeşitlilik indisleri yetersiz kalmaktadır. Düzeltilmiş tür çeşitlilik indisleri bu yetersizliği ortadan kaldırmak için geliştirilmiştir. Düzeltilmiş tür çeşitlilik indisleri, düzeltilmiş tür zenginlik indisleri, Simpson indisine dayalı düzeltilmiş tür çeşitlilik indisleri ve Shannon entropisine dayalı düzeltilmiş tür çeşitlilik indisleri olmak üzere üç gruba ayrılmaktadır.

Shannon entropisi temelli düzeltilmiş tür çeşitlilik indisleri (SDÇ) düzeltilmiş tür zenginlik indislerinden ve Simpson indisi temelli düzeltilmiş tür çeşitlilik indislerinden daha fazla tercih edilmektedir. SDÇ indisleri grubunda Jackknife yöntemi ( $JK$ ) (Zahl, 1977), düzeltilmiş en büyük olabilirlik kestiricisi ( $MLEbc$ ) (Chao vd., 1993; Chao vd., 2000), Chao ve Shen düzeltilmiş çeşitlilik eşitliği ( $\hat{H}_{CS}$ ) (Chao ve Shen, 2003), Özkan düzeltilmiş çeşitlilik kestiricisi ( $\hat{H}_O$ ) (Özkan, 2020), bilgi teorisi temelli düzeltilmiş çeşitlilik indisi ( $\hat{H}_{CWJ}$ ) (Chao vd., 2013), Miller sapma düzeltme eşitliği ( $\hat{H}_{Miller}$ ) (Strobl, 2005; Miller, 1955) ve Grassberger sapma düzeltme eşitliği ( $\hat{H}_{\psi}$ ) (Bonachela vd., 2008) bulunmaktadır.

 <sup>a</sup> Isparta Uygulamalı Bilimler Üniversitesi, Orman Fakültesi, 32260, Isparta

 <sup>\*</sup> **Corresponding author** (İletişim yazarı): kursadozkan@isparta.edu.tr

 **Received** (Geliş tarihi): 24.05.2023, **Accepted** (Kabul tarihi): 10.07.2023



**Citation** (Atıf): Özkan, K., 2023. Shannon entropisi temelli düzeltilmiş çeşitlilik ölçümleri ve Pla kestiricisinin üst güven sınır değerini kullanarak gerçek tür çeşitliliğinin hesaplanması. Turkish Journal of Forestry, 24(3): 235-240. DOI: [10.18182/tjf.1302074](https://doi.org/10.18182/tjf.1302074)

Pla (2004) tarafından önerilen düzeltmeli çeşitlilik hesabı ( $\hat{H}^e$ ), Shannon entropisi temelli düzeltmeli tür çeşitlilik (SDÇ) indisleri grubuna dahil edilebilir. Ancak, Pla (2004)'nın önerdiği düzeltmeli çeşitlilik hesabı popüler bir yaklaşım değildir. Bunun muhtemel iki sebebi vardır. Birinci sebep,  $\hat{H}^e$  indisinin SDÇ formlarına göre genelde daha düşük kestirim değeri vermesidir. Bu durum  $\hat{H}^e$ 'in negatif sapma ile ortaya çıkan farkı kapatmada yetersiz kaldığı anlamına gelmektedir. Diğer sebep, önerilen bu yaklaşımın simülasyon ürünü esnek olmayan istatistiksel bir denkleme dayandırılmış olmasıdır. Bununla birlikte bu yöntemin sağladığı önemli bir avantaj vardır. Zira bu yaklaşımdan varyans değeri hesaplanabildiğinden, düzeltmeli çeşitlilik değerine eşlik edebilecek düzeltmeli alt ve üst güven düzeyi değerleri de ( $\hat{H}^e \pm 1,96\sigma_{\hat{H}^e}^*$ ) elde edilebilmektedir.

Negatif sapmadan kaynaklanan farkı kapatmama riskinin yüksek olmasından dolayı  $\hat{H}^e$  ve haliyle  $\hat{H}^e - 1,96\sigma_{\hat{H}^e}^*$  değerlerinin düzeltmeli çeşitlilik değerleri olarak kullanılmasına mesafeli durmak mantıklı bir seçenektir. Bununla birlikte  $\hat{H}^e + 1,96\sigma_{\hat{H}^e}^*$  değerini, SDÇ indislerinin sonuçlarını kontrol veya teyit etmek için kullanımı önerilebilir. Zira  $\hat{H}^e + 1,96\sigma_{\hat{H}^e}^*$  değeri  $\hat{H}^e$ 'in üst sınır değeri olduğundan, bunun negatif sapmadan kaynaklanan farkı kapatma ve hatta geçme ihtimali yükselmektedir.

O halde, bir örnek alan verisi için, düzeltilmeli üst güven düzeyi değeri, o örnek alan verisine uygulanan SDÇ indis grubu değerlerinin geçerliliği için kullanılabilir. Böylece SDÇ indislerinin olması gerekenden veya  $\hat{H}^e + 1,96\sigma_{\hat{H}^e}^*$  değerinden daha yüksek bir kestirim değeri pozitif sapmaya atfedilebilir ve pozitif sapma gösteren SDÇ indislerinin sonuçları geçersiz kılınabilir. Diğer bir ifade ile, bir örnek alan için SDÇ indislerine ait kestirimler eğer o örnek alan için elde edilen  $\hat{H}^e + 1,96\sigma_{\hat{H}^e}^*$  değerinden düşük ise geçerli olduğu kabul edilebilir ve örnek alanı temsil eden toplumun çeşitlilik değeri için kullanılabilir. Özetle, SDÇ indisleri içinde geçerli sonuçlar  $\hat{H}^e + 1,96\sigma_{\hat{H}^e}^*$  değerinin altında ise kabul edilebilir ve bu kabule uygun SDÇ indislerine ait sonuçların ortalaması ile

o örnek alanı en ideal temsil edecek düzeltilmiş çeşitlilik değeri elde edilebilir.

Bu çalışma, örneklem veya örnek alan verileri için Pla (2004)'nın yaklaşımı ile elde edilen düzeltmeli üst limit tür çeşitlilik değerini baz alarak, Shannon entropisi temelli düzeltmeli tür çeşitlilik indislerinin (SDÇ) geçerliliğini nasıl sorgulayacağımızı ve nihayetinde örneklem için en uygun düzeltmeli çeşitlilik değerini nasıl elde edebileceğimizi göstermek amacıyla gerçekleştirilmiştir.

## 2. Materyal ve yöntem

### 2.1. Çalışma materyali

Çalışmada materyal olarak Çizelge 1'deki verilen 5 toplum verisi kullanılmıştır. Bunlardan DT, NT, EW ve CP kod isimleri verilmiş 4 toplum verisi daha evvel Chao ve Shen (2003) ve Özkan (2020) tarafından kullanılmış olup, Janzen (1973a ve 1973b)'den alınmıştır. Beşinci toplum verisi (HP) ekstrem tür dağılıma örnek olması amacıyla yapay olarak üretilmiştir.

Çizelge 1'de  $f_k$  değerlerinin toplamı gözlenen tür sayısına denk gelmektedir ( $\sum f_k = S$ ).  $k$ , türlere ait birey sayılarını ifade etmekte ve  $k \times f_k$  değerlerinin toplamı toplam birey sayısını ( $N$ ) vermektedir. Örneğin DT toplumunda, tek bireyli 59 tür, iki bireyli 9 tür, 3 bireyli 3 tür, 4 bireyli 2 tür, 5 bireyli 2 tür, 6 bireyli 2 tür ve 11 bireyli 1 tür olmak üzere toplamda 78 tür ve 127 birey bulunmaktadır.

### 2.2. Shannon entropisi temelli düzeltmeli tür çeşitlilik indisleri

Bu makalede kullanılan Shannon entropisi temelli düzeltmeli tür çeşitlilik (SDÇ) indisleri Çizelge 2'de verilmiştir.

Çizelge 1. Örnek veri setleri (DT, NT, EW, CP verileri Janzen (1973a ve 1973b)'den alınmıştır.)

														Toplam
DT	$k$	1	2	3	4	5	6	11						127
	$f_k$	59	9	3	2	2	2	1						78
NT	$k$	1	2	3	5	7	10	14	16	18				170
	$f_k$	56	9	7	2	1	1	1	1	1				79
EW	$k$	1	2	3	5	6	11	16	21	25	26	35		170
	$f_k$	4	3	5	1	1	1	1	1	1	1	1		20
CP	$k$	1	2	3	4	5	9	11	14	20	30	65		198
	$f_k$	2	1	2	1	1	1	1	1	1	2	1		14
HP	$k$	1	3	5	8	14	41	62						236
	$f_k$	84	4	3	1	1	1	1						95

$k$ : türlere ait birey sayıları,  $f_k$ : gözlenen tür sayıları

Çizelge 2. Shannon entropisi temelli düzeltmeli tür çeşitlilik indisleri (SDÇ)

İndislerin formülleri	İndis no
$JK = n \ln n - (n - 1) \ln(n - 1) + \frac{1}{n} \sum_{k=2}^n f_k k^2 \ln \frac{(k-1)}{k} - \frac{1}{n} \sum_{k=2}^n f_k k \ln(k - 1)$	(1)
$MLEbc = \hat{H}_{MLE} + (\hat{S}_{ACE} - 1)/2n$	(2)
$\hat{H}_{CS} = -\sum_{i=1}^S \frac{\hat{p}_i \ln \hat{p}_i}{1-(1-\hat{p}_i)^n} I(A_i) = -\sum_{i=1}^S \frac{\hat{C}\hat{p}_i \ln \hat{C}\hat{p}_i}{1-(1-\hat{C}\hat{p}_i)^n} I(A_i)$	(3)
$\hat{H}_O = -\sum_{i=1}^S \frac{\hat{p}_k \ln \hat{p}_k}{1-(1-\frac{1}{S_{obs}})^{\sum_{i=1}^S f_k k + \sum_{i=1}^S f_k \hat{k}}} I(A_i) = -\sum_{i=1}^S \frac{\hat{p}_k \ln \hat{p}_k}{1-(1-\frac{1}{S_{obs}})^{\sum_{i=1}^S f_k \hat{k}}} I(A_i)$	(4)
$\hat{H}_{Cwj} = \sum_{1 \leq X_i \leq n-1} \frac{X_i}{n} [\psi(n) - \psi(X_i)] + \frac{f_1}{n} (1 - A)^{-n+1} \left\{ -\ln(A) - \sum_{r=1}^{n-1} \frac{1}{r} (1 - A)^r \right\}$	(5)
$\hat{H}_{Miller} = H(p) + \frac{S-1}{n}$	(6)
$\hat{H}_\psi = \sum_{i=1}^S \frac{X_i}{n} \left( \ln n - \psi(X_i) - \frac{(-1)^{X_i}}{X_i(X_i+1)} \right)$	(7)

Jackknife indisinde (*JK*) (Çizelge 2, indis 1) *n* toplumdaki toplam birey sayısını, *k* türlerin gözleme sayısını ifade etmektedir. *f<sub>k</sub>* birey sayılarının *k*. değerine denk gelen tür sayısıdır (Zahl, 1977).

Düzeltilmiş en büyük olabilirlik kestiricisinde (*MLEbc*) (Çizelge 2, indis 2) bulunan  $\hat{H}_{MLE}$  en büyük olabilirlik kestiricidir.  $\hat{S}_{ACE}$  bolluk verilerine dayalı örtü kestiricidir ve hesabı bol ve nadir görünen türlerin sayısına (*S<sub>abun</sub>*, *S<sub>rare</sub>*) dayanmaktadır (Chao vd., 1993; Chao vd., 2000).

$$\hat{H}_{MLE} = -\sum_{i=1}^S \hat{p}_i \ln \hat{p}_i = -\sum_{i=1}^S \frac{X_i}{n} \ln \left( \frac{X_i}{n} \right) \quad (8)$$

$$\hat{S}_{ACE} = S_{abun} + \frac{S_{rare}}{\hat{C}_{rare}} + \frac{f_1}{\hat{C}_{rare}} \hat{\gamma}^2 \quad (9)$$

Eşitlikte  $\hat{p}_i = X_i/n$ ,  $p_i$ 'nin *MLE*'si olmaktadır.  $\hat{C} = 1 - f_i/n$  ve  $\sum_{m=1}^n m f_m = n$  bağlantı eşitlikleri olup,  $\hat{C}_{rare} = 1 - f_1/\sum_{m=1}^k m f_m$  eşitliği ile hesaplanır. Genellikle  $\kappa=10$  olarak alınmaktadır (Chao vd., 2000).

Chao ve Shen (2003)'in önerdikleri sapma-düzeltilme çeşitlilik eşitliği ( $\hat{H}_{CS}$ ) (Çizelge 2, indis 3) Horvitz-Thompson düzeltmesi (Horvitz ve Thompson, 1952) temelinde inşa edilmiştir.  $\hat{H}_{CS}$  indisinde  $\hat{p}_i = \hat{C}\hat{p}_i$  ve  $\hat{C} = 1 - (f_1/n)$  olup, *f<sub>1</sub>* tek bireye sahip türlerin sayısını ifade etmektedir.  $\hat{p}_i = X_i/n$ ,  $p_i$ 'nin *MLE*'sidir.

Özkan (2020) tarafından önerilen sapma düzeltmeli çeşitlilik kestirimi ( $\hat{H}_O$ ) (Çizelge 2, indis 4) Von Neuman Entropisi'nin yarı dolanıklık yoğunluk matrisine dayanmaktadır. Bu yöntemde ilk olarak yarı dolanıklık yoğunluk matrisinden  $\lambda_{max}$  hesaplanmaktadır.  $\hat{H}_O$  hesabında *f<sub>k</sub>* türlerin birey sayısı (*k* = 1,2, ..., *m*) olup,  $\hat{k}$  bir türün birey sayısına eklenecek değeri ifade etmektedir. Diğer bir deyişle  $\hat{k}$  gözlenmeyen türlere atfen *k* kere gözlenen türün kestirilen birey sayıdır.  $\hat{k} = \lambda_{max} k/k^k$  olup, buradan  $\hat{k} = \hat{k} + k$  hesabı gerçekleştirilir ve böylece  $\hat{p}_k$  değerleri ( $\hat{p}_k = \hat{k}/\sum_{i=1}^S f_k \hat{k}$ ) belirlenebilir.  $\hat{p}_k$  *i*. tür için kullanılacak oransal değeri ifade etmektedir. Her *k* değeri için *i*. türe ait  $\hat{p}_k$  değerleri elde edildikten sonra sapma düzeltmeli çeşitlilik ( $\hat{H}_O$ ) hesaplanır.  $\hat{H}_O$  eşitliğinde (Çizelge 2, indis 4)  $I(A_i)$  gösterge fonksiyonudur. Eğer  $A_i$  doğru ise o zaman  $I(A_i) = 1$ , eğer yanlış ise  $I(A_i) = 0$  olmaktadır.

Bilgi teorisi temelli düzeltmeli çeşitlilik indisinde ( $\hat{H}_{Cwj}$ ) (Çizelge 2, indis 5), *X<sub>i</sub>* *i*. türün birey sayısını, *n* toplumdaki toplam birey sayısını, *f<sub>1</sub>* tek bireyli türlerin sayısını ve *f<sub>2</sub>* iki bireyli türlerin sayısını ifade etmektedir (Chao vd., 2013).  $\psi(x)$  digama fonksiyonu olup (Kokologiannaki ve Krasniqi, 2013), bu fonksiyonda yer alan  $\gamma = \lim_{m \rightarrow \infty} \left( \sum_{l=1}^m \frac{1}{l} - \ln m \right) = 0,577215665 \dots$  Euler sabitesi'dir (Choi ve Srivastava, 2000).  $\hat{H}_{Cwj}$  formülünde yer alan *A*'nın tespiti aşağıdaki eşitlikler ile gerçekleşmektedir.

$$A = \begin{cases} 2f_2/[(n-1)f_1 + 2f_2], & f_2 > 0 \\ 2/[(n-1)(f_1 - 1) + 2], & f_2 = 0, f_1 \neq 0 \\ 1, & f_1 = f_2 = 0, \end{cases} \quad (10)$$

Miller sapma düzeltme indisinde ( $\hat{H}_{Miller}$ ) (Çizelge 2, indis 6) *H(p)* Shannon entropisi ( $H(p) = -\sum_{i=1}^S p_i \log p_i$ ), *S* tür sayısı ve *n* toplam birey sayısıdır (Strobl, 2005; Miller,1955).

Grassberger (1988, 2003) tarafından önerilen sapma-düzeltilme eşitliği ( $\hat{H}_\psi$ ) (Çizelge 2, indis 7) Bonachela vd. (2008) tarafından özetlenerek verilmiştir.  $\hat{H}_\psi$  indisinde, daha önce  $\hat{H}_{Cwj}$  indisinde de açıklandığı üzere  $\psi(x)$ , digama fonksiyonu ve  $\gamma$ , Euler sabitesi'dir.

### 2.3. Pla düzeltmeli tür çeşitlilik kestirim metodu

SDÇ grubunda yer alabilecek bir diğer düzeltmeli çeşitlilik hesaplama yaklaşımı Pla (2004) tarafından önerilmiştir. Ancak Pla (2004)'nin önerdiği yaklaşım SDÇ indislerinden yapısal olarak farklıdır. Zira Pla (2004) tarafından önerilen yaklaşımda bootstrap toplumlarından Shannon indisine dayalı düzeltmeli tür çeşitliliği hesaplanmaktadır.

Hatırlanacağı üzere bir toplumun çeşitlilik hesabı için kullanılan Shannon indisi (Shannon, 1948) aşağıdaki gibidir.

$$H = -\sum_{i=1}^S p_i \ln p_i \quad (11)$$

Burada *H* ilgili toplum için doğru çeşitlilik değerini ifade etmektedir ve bu hesap, o toplum ile ilgili verinin eksiksiz

elde edildiği kabulüne dayanmaktadır. Bu durumda,  $S$  toplumdaki doğru tür sayısını ve  $p_i$ ,  $i$ . türün doğru nispi bolluk değerini ifade etmektedir.

Bir toplumu temsil etmek üzere örnekleme yapıldığında veya örnek alan alındığında az veya çok negatif bir sapma söz konusu olur. Bu durumda Shannon indisi için  $H$  yerine  $\hat{H}$  terminini kullanmak uygun olur (Pla, 2004).

$$\hat{H} = -\sum_{i=1}^S \hat{p}_i \ln \hat{p}_i \quad (12)$$

$\hat{H}$  eşitliğinde,  $\hat{S}$  toplumu temsil eden örnek alandaki tür sayısını ve  $\hat{p}_i$ , toplumu temsil etmek üzere alınan örnek alandaki  $i$ . türün bolluk değerinin ( $n_i$ ) o örnek alandan elde edilen toplam bolluk değerine ( $N$ ) oranını ( $\hat{p}_i = n_i/N$ ) ifade etmektedir. Haliyle,  $\hat{H}$  toplumun değil, toplumu temsil eden örnek alanın çeşitlilik değerini ifade etmektedir.

Shannon entropisine dayalı olarak bir topluma ait tür çeşitlilik değerini, diğer bir deyişle düzeltilmiş tür çeşitlilik değerini hesaplamak için daha önceden bahsi geçtiği üzere, Pla (2004) bootstrap toplumlarının kullanılmasını önermiştir.

Bu bağlamda öncelikle ana veri setinden üretilen her bir bootstrap toplumunun Shannon indisi ( $H_j^*$ ) değeri hesaplanır. Daha sonra hesaplanan  $H_j^*$  değerlerinin ortalaması ( $\bar{H}^*$ ) alınır.

$$\bar{H}^* = \frac{1}{B} \sum_{j=1}^B H_j^* \quad (13)$$

Varyans ( $\sigma_H^{*2}$ ) ise aşağıdaki eşitlik kullanılarak tespit edilir.

$$\sigma_H^{*2} = \frac{1}{B} \sum_{j=1}^B (H_j^* - \bar{H}^*)^2 \quad (14)$$

Simülasyon şartlarında, en iyi ampirik sapma kestirimine ve buradan indis nokta kestirimine ( $\hat{H}^e$ ) yönelik denklemler Pla (2004) tarafından üretilmiştir.  $\hat{H}^e$  değerine ulaşılmasını sağlayan bu denklemler aşağıda verilmiştir.

$$(\hat{H} - H) = 1,73(\bar{H}^* - \hat{H}) + 0,02\bar{H}^* - 0,0003\hat{S} \quad (15)$$

$$\hat{H}^e = 2,73\hat{H} - 1,75\bar{H}^* + 0,0003\hat{S} \quad (16)$$

Buradan  $\hat{H}^e \pm 1,96\sigma_H^*$  ile düzeltilmiş güven aralıkları elde edilir (Pla, 2004). Nihayetinde, işlem “eğer Shannon entropisi temelli  $i$ . düzeltilmeli tür çeşitlilik indisinin ( $SDC_i$ ) değeri  $< \hat{H}^e + 1,96\sigma_H^*$  ise, hesap kabul edilir aksi durumda reddedilir” koşullandırması ile son bulur.

Koşullandırma testinden geçen  $SDC_i$  indislerinin değerleri aşağıdaki formüle aktararak düzeltilmeli çeşitlilik değeri ( $\bar{H}_{bc}$ ) elde edilir.

$$\bar{H}_{bc} = \ln\left(\frac{\sum_{i=1}^T e^{SDC_i I(A)}}{\sum_{i=1}^T I(A)}\right) \quad (17)$$

Formülde  $T$  Shannon entropisi temelli düzeltilmeli tür çeşitlilik indislerinin toplam sayısını ve  $I(A)$  gösterge değerini ifade etmektedir. Eğer  $SDC_i$  grubunun  $i$ . indisine ait değer  $< \hat{H}^e + 1,96\sigma_H^*$  ise  $I(A) = 1$ , aksi durumda  $I(A) = 0$  olur.

Çalışmada bütün hesaplamalar bu amaçla hazırlanmış olan bir Microsoft Excel Spreadsheet programı kullanılarak gerçekleştirilmiştir.  $\hat{H}^e$ ,  $\hat{H}^e + 1,96\sigma_H^*$  hesapları her örnek veri seti için 500 bootstrap toplumu üretilerek gerçekleştirilmiştir.

### 3. Bulgular ve tartışma

Örnek toplumların hesaplanan  $SDC_i$  indis değerleri Çizelge 3’te, Pla düzeltilmeli tür çeşitlilik hesaplama sonuçları Çizelge 4’te verilmiştir.

HP, DT ve NT toplumlarının  $\hat{H}$  değerleri sırası ile 3,35;4,08;3,83 şeklindedir (Çizelge 4).  $SDC_i$  indislerine göre düzeltilmeli çeşitlilik değerleri HP toplumunda 3,73 ile 5,30 arasında, DT toplumunda 4,62 ile 5,11 arasında ve NT toplumunda 4,24 ile 4,62 arasında değişmektedir (Çizelge 3). EW ve CP toplumların  $\hat{H}$  değerleri sırası ile 2,41 ve 2,06 şeklindedir.  $SDC_i$  indislerinden elde edilen sonuçlara göre, düzeltilmiş çeşitlilik değerleri EW toplumu için 2,47 ile 2,52 arasında, CP toplumu için 2,09 ile 2,12 arasında değişmektedir.

Anlaşılabacağı üzere HP, DT ve NT toplumlarının  $\hat{H}$  değerleri ile düzeltilmeli çeşitlilik değerleri arasında önemli farklılıklar vardır. Oysaki, EW ve CP toplumların  $\hat{H}$  değerleri ile düzeltilmeli çeşitlilik değerleri arasında fark çok düşüktür. Bunun sebebi gayet açıktır. Şöyle ki; Çizelge 1’den görüleceği üzere, en fazla sayıda nadir tür sayısına, özellikle tek bireyli tür sayısına sahip toplumlar sırası ile HP, DT ve NT toplumlarıdır. EW ve CP toplumlarında ise nadir türlerin sayısı oldukça düşüktür. Bütün bunların sonucu olarak negatif sapma değeri en yüksek olan toplumlar HP, DT ve NT olmaktadır. EW ve CP toplumlarında ise negatif sapma değeri önemsenmeyecek kadar düşüktür, zira bu toplumlarda envanter büyük oranda tamamlanmıştır.

DT ve NT toplumunda  $JK$ ,  $\hat{H}_{CS}$ ,  $\hat{H}_O$ ,  $\hat{H}_{Miller}$  ve  $\hat{H}_\psi$  yöntemleri ile birbirlerine yakın kestirim değerleri elde edilmiştir. Bu toplumlarda en yüksek değerler bariz farklılıklarla  $MLEbc$  ve  $\hat{H}_{CWJ}$  indislerine aittir.

HP toplumunda en küçük kestirim değerleri  $JK$ ,  $\hat{H}_{Miller}$  ve  $\hat{H}_\psi$  tarafından elde edilmiştir.  $JK$ ,  $\hat{H}_{Miller}$  ve  $\hat{H}_\psi$  yöntemlerinden sırası ile 3,73;3,75;3,76 değerleri elde edilmiş olup, bu değerler birbirlerine oldukça yakındır. En büyük kestirim değerleri önemli farklılıklarla  $MLEbc$  (5,30) ve  $\hat{H}_{CWJ}$  (4,73) yöntemlerinden elde edilmiştir.  $\hat{H}_{CS}$  ve  $\hat{H}_O$  yöntemlerinden elde edilen değerler birbirlerine çok yakın olup, bunların değerleri  $JK$ ,  $\hat{H}_{Miller}$  ve  $\hat{H}_\psi$  grubuna ait değerler ile  $MLEbc$  ve  $\hat{H}_{CWJ}$  grubuna ait değerler arasında yer almaktadır.

EW ve CP toplumları için bütün yöntemlerin sonuçları birbirlerine çok yakın değerler vermiştir. Zira,  $SDC_i$  indisleri ile EW toplumu için elde edilen kestirim değerleri 2,47 ile 2,52 arasında, CP toplumu için 2,09 ile 2,12 arasında değişmektedir.

Görünen o ki, bir toplumda nadir tür sayısı ve nadir tür sayısının oranı arttıkça,  $SDC_i$  indisleri arasındaki fikir ayrılığı da kendini daha net hissettirmektedir. Bu durumda, bir toplumun düzeltilmeli çeşitlilik indisini temsil etmek için her zaman  $SDC_i$  indislerinin ortalama değerlerini almak doğru bir yaklaşım olmaz. Doğru yaklaşım,  $\hat{H}^e + 1,96\sigma_H^*$  eşik değerini kullanarak fikir ayrılığı en düşük düzeyde olan  $SDC_i$  indislerinin ortalamasını almaktır.

Çizelge 3 ve Çizelge 4’ten görüleceği üzere örnek toplumların Pla düzeltilmeli tür çeşitliliği ( $\hat{H}^e$ ) değerleri onların en küçük  $SDC_i$  değerlerinden bile daha küçüktür ya da o değerlere eşittir. Bu durum  $SDC_i$  indisleri ile kıyaslandığında  $\hat{H}^e$ ’nin negatif sapma değerini kapatmada yetersiz kaldığı göstermektedir. Örnek toplumların  $\hat{H}^e + 1,96\sigma_H^*$  değerleri ise bazı  $SDC_i$  indisleri değerlerinden daha

büyükür. Çizelge 5 SDC indislerinin gösterge deęerlerini vermektedir. DT, NT, EW, CP ve HP toplamlarının gösterge deęerleri ( $\hat{H}^e + 1,96\sigma_H^*$  deęerlerini geen indis sayısı) toplamı ( $\sum_{i=1}^T I(A)$ ) sırası ile 5;5;7;7;3 eklindeir (Çizelge

7). Çizelge 3 ve Çizelge 5'in hcre deęer arpımları her bir toplum iin ortalamaya girmeye hak kazanan SDC indislerine ait deęerleri iermektedir. Buradan her bir toplum iin elde edilen  $\sum_{i=1}^T e^{SDC_i} I(A)$  deęerleri Çizelge 7'nin ikinci satırında,  $\bar{H}_{bc}$  deęerleri Çizelge 7'nin son satırında verilmiřtir.

Çizelge 3. rnek toplamların Shannon entropisi temelli dzeltmeli tr eřitlilik (SDC) indislerine ait hesaplama sonuları

	$JK$	$\hat{H}_{CS}$	$MLEbc$	$\hat{H}_o$	$\hat{H}_{CWJ}$	$\hat{H}_{Miller}$	$\hat{H}_\psi$
DT	4,62	4,70	5,11	4,69	4,98	4,68	4,64
NT	4,24	4,30	4,62	4,34	4,48	4,29	4,25
EW	2,48	2,49	2,47	2,48	2,48	2,52	2,49
CP	2,10	2,09	2,09	2,09	2,10	2,12	2,10
HP	3,73	3,97	5,30	3,97	4,73	3,75	3,76

Çizelge 4. rnek toplamların Pla dzeltmeli tr eřitlilik hesaplama sonuları

	$\hat{H}$	$\bar{H}^*$	$\sigma_H^*$	$\hat{H}^e$	$\hat{H}^e - 1,96\sigma_H^*$	$\hat{H}^e + 1,96\sigma_H^*$
DT	4,08	3,73	0,08	4,62	4,47	4,77
NT	3,83	3,57	0,10	4,24	4,05	4,43
EW	2,41	2,35	0,08	2,47	2,32	2,61
CP	2,06	2,02	0,07	2,08	1,95	2,21
HP	3,35	3,13	0,12	3,69	3,46	3,92

Çizelge 5. SDC indis deęerlerinin rnek toplamlara gre belirlenen gstergedeęerleri ( $I(A)$ )

	$JK$	$\hat{H}_{CS}$	$MLEbc$	$\hat{H}_o$	$\hat{H}_{CWJ}$	$\hat{H}_{Miller}$	$\hat{H}_\psi$
DT	1	1	0	1	0	1	1
NT	1	1	0	1	0	1	1
EW	1	1	1	1	1	1	1
CP	1	1	1	1	1	1	1
HP	1	0	0	0	0	1	1

Çizelge 6. Çizelge 3 ve Çizelge 5'in hcre deęerleri arpımı

	$JK$	$\hat{H}_{CS}$	$MLEbc$	$\hat{H}_o$	$\hat{H}_{CWJ}$	$\hat{H}_{Miller}$	$\hat{H}_\psi$
DT	4,62	4,70	0	4,69	0	4,68	4,64
NT	4,24	4,30	0	4,34	0	4,29	4,25
EW	2,48	2,49	2,47	2,48	2,48	2,52	2,49
CP	2,10	2,09	2,09	2,09	2,10	2,12	2,10
HP	3,73	0	0	0	0	3,75	3,76

Çizelge 7. rnek toplamların gstergedeęerleri toplamı,  $\hat{H}^e + 1,96\sigma_H^*$  deęerinin altında kalan SDC indislerine ait stel deęer toplamları ve dzeltmeli eřitlilik deęerleri

	DT	NT	EW	CP	HP
$\sum_{i=1}^T I(A)$	5	5	7	7	3
$\sum_{i=1}^T e^{SDC_i} I(A)$	531,92	362,37	84,11	57,09	127,14
$\bar{H}_{bc}$	4,67	4,28	2,49	2,10	3,75

#### 4. Sonuç ve öneriler

“Giriş” kısmında da ifade edildiği üzere, tür çeşitliliğinde kullanılan çok sayıda indis vardır. Ancak hangi indisin biyolojik çeşitlilik hesaplamaları için uygun olduğu konusunda araştırmacılar arasında tam bir fikir birliği bulunmamaktadır. Literatüre göre biyolojik çeşitlilik hesaplarında araştırmacılar tarafından en fazla tür zenginliği, Simpson çeşitlilik indisi ve Shannon entropisi kullanılmaktadır.

Shannon entropisi çeşitlilik hesapları için en ideal formdur. Bununla birlikte örnekleme nadir türlerin varlığı Shannon entropisi ile gerçekleştirilen hesapların güvenilirliğini azaltmaktadır. Bu yüzden bazı araştırmacılar tarafından Shannon entropisini temel alan düzeltmeli çeşitlilik indisleri geliştirilmiştir. Araştırmacılar arasında düzeltmeli çeşitlilik indisleri içinde hangisinin gerçeğe daha yakın sonuç verdiği veya hangisinin kullanımının daha uygun olduğu konusunda da bir fikir birliği yoktur. Geleneksel yaklaşım, popüler olan bir düzeltmeli çeşitlilik indisinin seçilmesi veya örnekleme verilerinin birden fazla düzeltmeli çeşitlilik indisi ile hesaplanması ve değerlendirmesi şeklindedir. Ne var ki, birden fazla düzeltmeli tür çeşitlilik indisinin seçilerek kullanılması durumunda, elde edilen sonuçlar arasındaki uyumsuzluklara nasıl bir çözüm bulunacağı da belirsizdir.

Bir örnekleme veya örnek alan verisi için sadece tek bir Shannon entropisi temelli düzeltmeli çeşitlilik değeri istenirse, o zaman seçeneklerden biri Shannon entropisi temelli düzeltmeli çeşitlilik indislerine ait değerlerinin ortalama değerini kullanmak olabilir. Ne var ki, indislerin kestirim değerleri arasındaki farklılıkların büyük olması durumunda ortalama değerlerin temsil kabiliyeti için güvenilirlik azalır. Bununla birlikte, ortalama değerlerin temsil kabiliyetini yükseltmek için indislerin hesap çıktıları bir eşik değeri ile test edilebilir. Bu eşik değeri aynı amaca yönelik kabul edilebilir bir yaklaşım ile elde edilen düzeltilmiş üst limit çeşitlilik değeri olabilir. Pla (2004) tarafından önerilen yaklaşım ile düzeltilmiş üst sınır çeşitlilik değeri elde edilebilmektedir.

Bu çalışmada Çizelge 1’de verilen her örnekleme için eşik değeri olarak Pla (2004) yaklaşımı ile elde edilen düzeltilmiş üst sınır çeşitlilik değeri kullanılmış, kestirim yöntemleri ile bu eşik değerin altında elde edilen değerlerin ortalama değeri örnek alanların düzeltmeli çeşitlilik değeri olarak kabul edilmiştir. Böylece herhangi bir düzeltmeli çeşitlilik indisini seçme zorunluluğu olmadan gerçek tür çeşitliliğinin objektif bir yaklaşımla elde edilmesi sağlanmıştır. Elbette bu yaklaşımın geçerliliği gerçek ekolojik veriler kullanılarak test edilmeli, eğer uygulamasında eksiklikler varsa geliştirilmeli veya modifiye edilmelidir.

#### Açıklama

Bu makaledeki hesaplamalar için Microsoft Excel Spreadsheet programının oluşturulmasında yardımcı olan ISUBÜ Orman Fakültesi öğretim üyesi Doç. Dr. Ahmet MERT’e, makale metnini kontrol ederek eksiklerinin tamamlanmasına destek olan ISUBÜ Sütçüler Prof. Dr. Hasan GÜRBÜZ Meslek Yüksek Okulu öğretim üyesi Dr. Serkan ÖZDEMİR ve ISUBÜ Orman Fakültesi öğretim üyesi Dr. Ali ŞENOL’a teşekkür ederim.

#### Kaynaklar

- Bonachela, J.A., Hinrichsen, H., Munoz, M.A., 2008. Entropy estimates of small data sets. *Journal of Physics A: Mathematical and Theoretical*, 41(20): 202001.
- Chao, A., Ma, M.C., Yang, M.C., 1993. Stopping rules and estimation for recapture debugging with unequal failure rates. *Biometrika*, 80(1): 193-201.
- Chao, A., Hwang, W.H., Chen, Y.C., Kuo, C.Y., 2000. Estimating the number of shared species in two communities. *Statistica Sinica*, 10: 227-246.
- Chao A., Shen, T.J., 2003. Nonparametric estimation of Shannon’s index of diversity when there are unseen species in sample. *Environmental and Ecological Statistics*, 10: 429-442.
- Chao, A., Wang, Y.T., Jost, L., 2013. Entropy and the species accumulation curve: a novel entropy estimator via discovery rates of new species. *Methods in Ecology and Evolution*, 4(11): 1091-1100.
- Choi, J., Srivastava, H.M., 2000. Evaluation of higher-order derivatives of the gamma function. *Publikacije Elektrotehničkog Fakulteta. Serija Matematika*, 11: 9-18.
- Grassberger, P., 1988. Finite sample corrections to entropy and dimension estimates. *Physics Letters A*, 128(6-7): 369-373.
- Grassberger, P., 2003. Entropy Estimates from Insufficient Samplings. *ArXiv Physics, e-prints 0307138*.
- Janzen, D.H., 1973a. Sweep samples of tropical foliage insects: description of study sites, with data on species abundances and size distributions. *Ecology*, 54(3): 659-686.
- Janzen, D.H., 1973b. Sweep samples of tropical foliage insects: effects of seasons, vegetation types, elevation, time of day, and insularity. *Ecology*, 54(3): 687-708.
- Horvitz, D. G., Thompson, D. J., 1952. A generalization of sampling without replacement from a finite universe. *Journal of the American statistical Association*, 47(260), 663-685.
- Kokologiannaki, C.G., Krasniqi, V., 2013. Some properties of the k-gamma function. *Le Matematiche*, 68(1): 13-22.
- Miller G.A., 1955. A note on the bias of information estimates. In: *Information Theory In Psychology: Problems and Methods* (Ed: Quastler, H.), The Free Press, Glencoe Illinois, pp. 95-100.
- Özkan, K., 2020. A new proposed estimator for reducing bias due to undetected species. *Gazi University Journal of Science*, 33(1): 229-236.
- Peet, R.K., 1974. The measurement of species diversity. *Annual Review of Ecology and Systematics*, 5(1): 285-307.
- Pla, L., 2004. Bootstrap confidence intervals for the Shannon biodiversity index: a simulation study. *Journal of Agricultural, Biological, and Environmental Statistics*, 9(1): 42-56.
- Shannon, C.E., 1948. A mathematical theory of communication. *Bell System Technical Journal*, 27(3): 379-423.
- Strobl, C., 2005. Variable selection in classification trees based on imprecise probabilities. *ISIPTA ’05: Proceedings of the Fourth International Symposium on Imprecise Probabilities and Their Applications*, 20-23 July, Pittsburgh, USA, pp. 339-348.
- Zahl, S., 1977. Jackknifing an index of diversity. *Ecology*, 58(4): 907-913.