

Estimating True Species Richness from Braun Blanquet Scale

Kürşad ÖZKAN*

Isparta University of Applied Sciences, Forest Faculty, Isparta, TURKEY

*Corresponding Author: kursadozkan@isparta.edu.tr

Received Date: 17.06.2021

Accepted Date: 21.10.2021

Abstract

Aim of study: The goal of the study was to estimate true species richness from Braun Blanquet (BB) scale data.

Area of study: Yazılı Canyon Nature Park (YCNP) located in the Mediterranean Region of Turkey.

Material and methods: A bias-corrected approach was adapted based on the Good-Turning frequency formula to estimate true species richness (S_{est}) for 9 vegetation plots under three scenarios (Rare species are singletons: with 1/1 probability (Sc_1), with 1/2 probability (Sc_2), with 1/3 probability (Sc_3)).

Main results: The results indicate that with increasing uncertainty about the number of singletons, the difference between expected species richness and observed species richness decreases. To estimate the species richness of the plots taken from YCNP, scenario III (Sc_3) seems to be the best option due to existing maximum uncertainty concerning the number of singletons.

Highlights: All the proposed bias-corrected estimators have been developed by considering the abundance or the incidence-based data except for S_{st} . For employing S_{st} , all the data consists of the number of singletons (f_1) and super doubletons (f_{2+}). f_1 and f_{2+} can be obtained from BB scale because its "r" code usually corresponds to f_1 . However, some scientists prefer to use "r" in description of a few species. That causes an uncertainty about f_1 . Using S_{st} , this study offers an approach and a spreadsheet program to estimate true species richness even though the number of singletons is uncertain.

Keywords: Cover-abundance Scale, Prediction, Rare Species, Species Diversity, Enigmatic Statistic, Negative Bias

Braun Blanquet Skalasında Doğru Tür Zenginliğinin Kestirimi

Öz

Çalışmanın amacı: Bu çalışma Braun Blanquet (BB) skala verileri kullanarak doğru tür zenginliğinin nasıl kestirilebileceğini göstermek için gerçekleştirilmiştir.

Çalışma alanı: Türkiye'nin Akdeniz Bölgesi'nde bulunan Yazılı Kanyon Tabiat Parkı (YKTP)'dir.

Materyal ve yöntem: Çalışma materyali Braun Blanquet (BB) skalasına göre 9 örnek alanda kaydı yapılmış bitki verilerinden oluşmaktadır. Araştırmada tür zenginliğinin kestirimi için, üç farklı senaryoya göre (nadir türler; tek bireyli türlerdir (Sc_1), 1/2 oranında tek bireyli türlerdir (Sc_2), 1/3 oranında tek bireyli türlerdir (Sc_3)) Chao ve ark. (2017) tarafından geliştirilen Good-Turing frekans formülüne dayalı sapma-düzeltilme yaklaşımı (S_{est}) kullanılmıştır.

Temel sonuçlar: Elde edilen sonuçlara göre, tek bireyli tür sayıları ile ilgili belirsizlik arttıkça kestirilen tür sayısı ile gözlenen tür sayısı arasındaki fark azalmaktadır. YKTP'den alınan örnek alanlarda tek bireyli tür sayıları ile ilgili maksimum belirsizlik söz konusudur. Bu yüzden tür zenginliğini kestirmek için en ideal senaryo üçüncü senaryodur (Sc_3).

Araştırma vurguları: Negatif sapmayı azaltmak için birçok kestirici önerilmiştir. S_{st} hariç, önerilen diğer bütün kestiriciler bolluk veya rastlanma sıklığı verisini dikkate alınarak geliştirilmiştir. S_{st} 'in kullanılması için kullanılacak veri tek bireyli türlerin sayısı (f_1) ile iki ve daha fazla bireye sahip türlerin sayısından (f_{2+}) ibarettir. f_1 ve f_{2+} BB skalası'ndan elde edilebilir. Zira BB skalasında "r" genellikle tek bireyli türler için kullanılır. Bununla birlikte bazı bilim insanları "r" kodunu "bir veya birkaç tür" tanımı için kullanmaktadır. Bu durum f_1 değeri ile ilgili belirsizliğe sebep olmaktadır. Bu çalışma tek bireyli tür sayıları ile ilgili bir belirsizlik olsa bile, S_{st} üstünden tür zenginlik kestirimini sağlayabilecek bir yaklaşım ve excel programı sunmaktadır.

Anahtar Kelimeler: Örtü-bolluk Skalası, Kestirim, Nadir Türler, Tür Çeşitliliği, Enigmatik İstatistik, Negatif Sapma



Introduction

Biodiversity plays an essential role in maintaining ecological balance and the health of habitats. It is therefore one of the central topics in community ecology, conservation biology and historical biogeography (Pyšek et al., 2005; Pärtel et al., 2011; Chao & Chiu, 2016; Özkan, 2016).

A wide range of indices have been proposed for measuring biodiversity. Amongst those indices the oldest, simplest, and still most frequently used one is species richness (i.e., the number of the species present in an ecological unit or an assemblage). Although the mathematical properties of species richness are intuitive, its statistical properties are much worse than those of any other common diversity measure since it is insensitive to the individual counts of the species (Chao & Jost, 2012).

For biodiversity studies and assessments, the data have generally been taken from sample plots, ecological units or biological assemblages. However, it is nearly impossible to record all species with limited field efforts because of sampling limitation. Hence there are unseen species in almost all species inventories. Unseen species causes negative bias and therefore leads to underestimation of true species richness (Chao & Chiu, 2016). Numerous measures have been proposed to reduce the effect of the negative bias. Two types of data (i.e., abundance data and incidence data) have been recorded for majority of ecological or biological communities. All the bias-corrected measures have therefore been designed for such data types. However, abundance or incidence-based data type has not always been recorded for all community inventories.

In many basic and applied plant ecological studies, visual estimation methods have frequently been used. Amongst visual estimation methods Braun-Blanquet (BB) scale is the most popular method since it is faster, cheaper and easier to use when compared to the other botanical field survey methods (Chmura & Salachna, 2016). BB scale includes 7 degree (Table 1).

Table 1. BB cover-abundance scales (Westhoff & Van der Maarel, 1973)

The Codes	Cover Abundance
<i>r</i>	One or few individuals
+	Occasional and less than 5%
1	Abundant and with very low cover, or less abundant but with higher cover; in any case less than 5% cover of total area
2a	Very abundant and less than 5% cover
2b	5%–12.5% cover, irrespective of number of individuals
2c	12.5%–25% cover of total area, irrespective of number of individuals
3	25%–50% cover of total area, irrespective of number of individuals
4	50%–75% cover of total area, irrespective of number of individuals
5	75%–100% cover of total area, irrespective of number of individuals

Apart from "*r*" which represents integer data (usually singletons), the other scales represent neither abundance data nor incidence data, but cover-abundance data (Table 1). Therefore, BB scale data is unavailable in employing the bias corrected estimators designed for abundance and incidence data. In this context, there is only one available method, which was proposed by Chao et al. (2017). Using this novel method, true species richness can be estimated from the number of the observed species (S_{obs}) and singletons (f_1). BB scales provide both of two inputs if the code "*r*" is merely recorded for singletons.

From the given information above, it is no doubt that singletons are definitely described by "*r*" because there is only one code for singletons which is "*r*" in BB scales. But "*r*" does not always represent to singletons since "*r*" is sometimes used for a few species (i.e., for both one individual and a few individuals) in characterizing plant communities (Poore, 1955; Westhoff & Van der Maarel, 1973; Werger, 1974; Peet et al., 1998; Van der Maarel, 2007; Miller et al., 2014). In this context, a question arises

How can we estimate true species richness if "r" code does not solely represent to singletons? The lack of information about the numerical value of "r" code for a given ecological community means that there is an uncertainty about the number of singletons for that community.

Although we can't get information to make clear the situation about the "r" records in a sample plot, there is a solution to estimate true species richness. This solution involves generating random subsets of the data to estimate species richness.

The aim of the study is to propose an approach in estimation of true species richness when the number of singletons is uncertain. The approach is based on Chao et al. (2017) modification of the Enigmatic statistics resulting from the Good-Turing theory.

Materials and Methods

Data Example

The Turkish Lake district of the Mediterranean region is rich in plant species. One of the most important nature reserves of the district is Yazılı Canyon Nature Park. It covers an area of around 600 ha. In the Park, one hundred seven sample plots were taken using Braun-Blanquet method during the summertime of 2006. Amongst them, nine sample plots were selected to illustrate the proposed approach.

All the "r" codes were recorded for less than or equal to three individuals during the survey. The number of singletons in the sample plots is therefore unclear. The number of the observed species (S_{obs}) and the number of the r codes of each plot (P_i) are given in Table 2. S_{obs} ranges between 12 and 41. Sample plot 5 (P_5) was the richest in taxa to which the "r" code was assigned, whereas P_1 and P_3 included only one rare species.

Table 2. The number of the observed species (S_{obs}) and the number of rare (r) species of the selected sample plots from Yazili Canyon Nature Park in the Mediterranean Region, Turkey

Sample plot	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9
S_{obs}	41	19	24	19	41	31	14	18	12
r	1	6	1	2	9	7	4	8	6

Estimating S_{est} Using S_{obs} and f_1

Enigmatic statistics were introduced by Alan Turing and his colleague I.J. Good and used by them during World War II in breaking the German "Enigma" code. After the War, their frequency estimation was published by Good (Good, 2000; Good, 1953). The proposed method by Chao et al. (2017) is based on this estimation.

Let f_k represent the abundance-based frequency counts, i.e., the number of species that are recorded in exactly k times ($k = 0, 1, \dots, n$) in a sample plot. S_{obs} is the number of observed species ($S_{obs} = \sum_{k \geq 1} f_k$). f_1 and f_2 are singletons and doubletons respectively and, f_0 represents the number of undetected species.

As mentioned before, only "r" is available from BB cover-abundance scale in the estimation of true (expected) species richness. If "r" is only recorded for f_1 ($f_{Br} = f_1$) in the sample plot, as explained

by Chao et al. (2017), the statistical inference problem can be solved by using f_1 and the number of super-duplicates (i.e., the number of species observed at least twice in the sample plot).

In this context, instead of using $S_{obs} = f_1 + f_2 + f_3 + \dots + f_n$ for all species and $f_{2+} = f_2 + f_3 + \dots + f_n$ for super-duplicates (Chao et al. 2017), we might write $S_{obs} = f_{Br} + f_{Br<} + f_{B+} + f_{B1} + f_{B2a} + \dots + f_{B5}$ and, $f_{Br<+} = f_{Br<} + f_{B+} + f_{B1} + f_{B2a} + \dots + f_{B5} = S_{obs} - f_{Br}$ respectively. Here, $f_{Br<+} = f_{2+}$ and the number of the "r" codes equal to $f_{Br} + f_{Br<}$ where $f_{Br<} \neq f_1$ since $f_{Br} = f_1$. The terms correspond to BB cover-abundance codes.

Suppose that total number of individuals (n) and the number of individuals belonging to each species described by BB codes are known. Based on the binomial model, the general form for the expected value of f_k becomes employable.

$$E(f_k) = E(\sum_{i=1}^S I(X_i = k)) = \sum_{i=1}^S \binom{n}{k} \psi_i^k (1 - \psi_i)^{n-k}, k = 0, 1, \dots, n \quad (1)$$

Thus, we can use the lower bound of undetected species richness (Chao, 1987).

$$\hat{S}_{Chao1} = S_{est} = S_{obs} + m \frac{f_{Br}^2}{2f_2} \quad (2)$$

Where $f_{Br}^2 = f_1^2$ and $m = (n - 1)/n$. $m \approx 1$ when we assume that n is large enough.

The solution to the problem begins from this point because we need to estimate f_2 . The following equation is the key to the proposed approach by Chao et al. (2017).

$$f_{Br<+} \approx f_{Br} \left(\frac{f_{Br}}{2f_2} \right) \left[\frac{1}{2} \left(\frac{2f_2}{f_{Br}} \right)^2 + \frac{1}{3} \left(\frac{2f_2}{f_{Br}} \right)^3 + \dots + \frac{1}{n} \left(\frac{2f_2}{f_{Br}} \right)^n \right] \quad (3)$$

and Eq. 4 is the alternative equation given as

$$f_{Br<+} \approx h(f_2) = \frac{f_{Br}^2}{2f_2} \left[-\ln \left(1 - \frac{2f_2}{f_{Br}} \right) - \frac{2f_2}{f_{Br}} \right] \quad (4)$$

By solving Eq. 3 when n is known or Eq. 4 when n is large enough or tends to infinity (i.e., computation of f_2), the problem is solved.

Once f_2 is defined, expected species richness can be easily computed by

$$S_{est} = S_{obs} + f_0 \quad (5)$$

where $f_0 = f_{Br}^2/2f_2$. Further details about this method can be retrieved in the APP4.pdf of the article of Chao et. al. (2017).

Estimating S_{est} in the Case of Uncertainty

Various integer or non-integer counts have been used or proposed for the code "r" of BB scale (Poore, 1955; Westhoff & Van der Maarel, 1973; Campbell, 1974; Werger, 1974; Prieditis, 1997; Peet et al., 1998; Van der Maarel, 2007; Miller et al., 2014; Vahdati et al., 2014; Chmura & Salachna, 2016; Camiz et al., 2017). Numerical differences in description of code "r" in the literature reveal that although f_1 certainly corresponds to "r", "r" probably does not correspond to f_1 .

In this section, we focus on estimation of true species richness for such a case that might be called as uncertainty.

Suppose that an ecological community includes 100 species. 10 species have "r" code (i.e., rare species ("r" species)). The number of species is certain. It is equal to 100 ($S_{obs} = 100$) but it is uncertain whether "r" species represent singletons or not. Although the number of "r" species does not mean the number of singletons, the sample plot probably consists of singletons due to presences of "r" codes. It is however unclear how to decide the number of singletons (f_1) via the number of "r" species. If we do not know the number of singletons, how can we estimate true species richness? The solution to this problem is to employ the proposed method of Chao et al. (2017) by repeatedly generating data with replacement.

Based on the literature, we might assume that the number of individuals for "r" code generally varies between 1 and 3. In this case count 1 and count 3 can represent the minimum and the maximum interval values of any r code if the number of individuals of the species described by "r" code is unclear in an inventory card. In this context, random data processing is performed between 1 and 3. Thus many values can be produced by replications for each of "r" species and, expected species richness can be computed at each replication. Average of the replication results gives the values we would like to reach. Consequently, we can enable true species richness (S_{est}) with confidential interval (C.I.) by means of S_{obs} and the number of "r" species.

With regards to defining confidential interval (C.I.), population variance ($\hat{\sigma}^2$) should be computed.

Numerous estimators for population variance have been proposed (Chao, 1987; Kadilar & Çingi, 2007; Özel et al., 2014; Sarmah, 2017; Chao et. al., 2017; Çekim & Kadilar, 2020). To define $\hat{\sigma}^2$ ($var_{est}(\hat{N})$), the following equation was used in the present study (Chao, 1987).

$$\hat{\sigma}^2 = f_2 \left[\frac{1}{4} \left(\frac{f_1}{f_2} \right)^4 + \left(\frac{f_1}{f_2} \right)^3 + \frac{1}{2} \left(\frac{f_1}{f_2} \right)^2 \right] \quad (6)$$

For the lower bound, confidential interval (C.I.) based on a log transformation was chosen (Chao, 1987) whereas classic C.I. was employed for the upper bound.

$$[S + (S_{est} - S)/C, S_{est} + 1.96\hat{\sigma}] \quad (7)$$

Where C is defined by

$$C = e^{\{1.96[\ln(1+\hat{\sigma}^2/(S_{est}-S)^2)]^{0.5}\}} \quad (8)$$

The reason is that lower bound of classic confidential interval (C.I.) might be greater than S_{obs} . Log-transformed C.I. is therefore appropriate for the lower bound. Likewise, the upper bound of log transformed C.I. tends to give a greater value than expected. Therefore, Classic C.I. seems to be a better option for the upper bound. Modification to the lower or the upper bound of confidential interval is not a new notion (Eren et al., 2012).

Results

For the solutions (i.e. computing S_{est} with C.I.), a spreadsheet program was created and used. For readers or users, the program (BBrOzkanMert) is available from <http://kantitatifekoloji.net/takdivozkan> to facilitate all computations. The program is user friendly but not time consuming. To illustrate the application of the program, the data taken from $P9$ was selected (Table 2). The application window is displayed in Figure 1. The procedure is described in the following steps.

Step 1: Enter the counts of S_{obs} and "r" species into $B2$ and $B3$ cells.

Step 2: Enter the minimum and maximum counts for "r" species into $B6$ and $B7$ cells respectively. For instance, $min = 1$ and $max = 3$ might be entered, but it depends on the users.

Step 3: Enter the number of replications into $B9$ and run. $R = 200$ or larger is suggested ($R_{max} = 1000$).

The results appear in $D13 - D15$ cells. Replication results exist between column J and column Q (Figure 1).

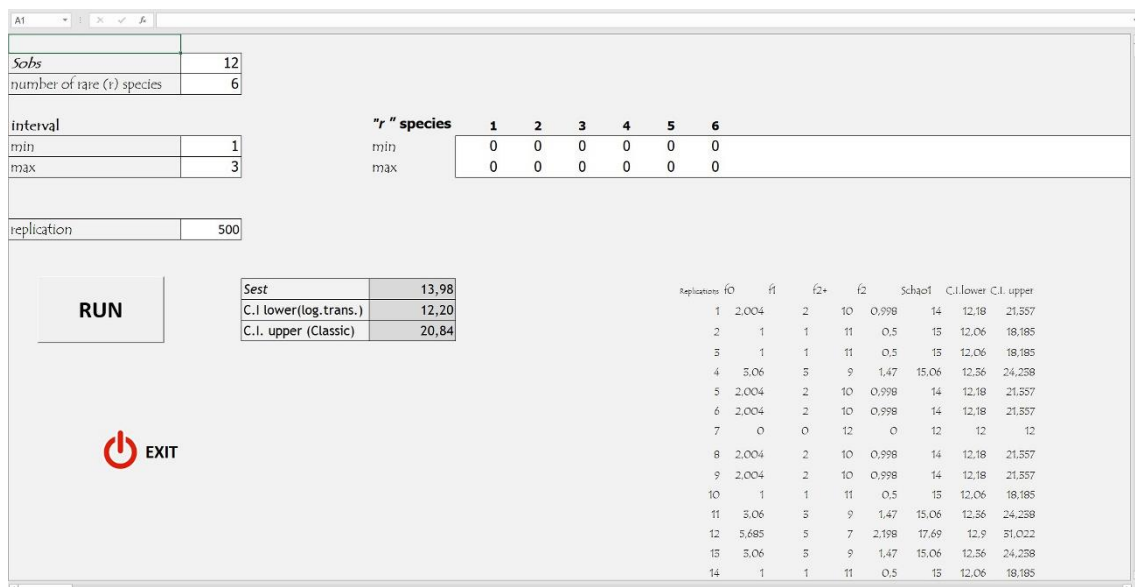


Figure 1. The menu of the program (min=1 and max=3)

The users might prefer to enter the minimum and maximum for each of the "r" species for two reasons.

1. Some of "r" species certainly represent singletons

2. Degrees of uncertainty for all or some of "r" species are different.

For such cases, the cells occupied by row 6 and row 7 starting from E column are used. To activate those cells, make inactive $B6$ and $B7$ cells ($B6 = B7 = 0$). To better

explain this, a hypothetical community example was given in the Appendix.

If the user ensures that all "r" species are singletons, then he/she should activate B_6 and 7 ($B_6 = B_7 = 1$).

With regards to the survey data of Yazılı Canyon Nature Park, there are three options for a rare species. Namely it is probably a singleton. If it is not a singleton, it is probably a doubleton. It is a tripleton, otherwise.

In this sense, to illustrate the approach proposed by the present study, we can establish three scenarios as follows.

Scenario I (Sc_1): Rare species are singletons with 1/1 probability ($f_{Br(1)} = f_1$)

Scenario II (Sc_2): Rare species are singletons with 1/2 probability ($f_{Br(1:2)}$)

Scenario III (Sc_3): Rare species are singletons with 1/3 probability ($f_{Br(1:3)}$)

The results of each sample plot for each scenario were obtained using the spreadsheet program ($R = 500$ for all applications) (Table 3).

Table 3. The results of expected species richness with C.I. considering $f_{Br(1)} = f_1$, $f_{Br(1:2)}$ and $f_{Br(1:3)}$

Scenarios	Sc_1			Sc_2			Sc_3		
	$f_{Br(1)} = f_1$	C.I.		$f_{Br(1:2)}$	C.I.		$f_{Br(1:3)}$	C.I.	
Sample plots	$S_{est(1)}$	lower	upper	$S_{est(1:2)}$	lower	upper	$S_{est(1:3)}$	lower	upper
P_1	42.00	41.06	47.15	41.48	41.03	43.98	41.32	41.02	42.98
P_2	25.32	20.12	38.74	21.99	19.38	30.67	21.02	19.22	27.92
P_3	25.00	24.06	30.19	24.51	24.03	27.13	24.34	24.02	26.13
P_4	21.00	19.18	28.33	20.01	19.08	24.43	19.72	19.05	23.05
P_5	50.10	43.06	65.85	45.56	41.72	56.48	44.01	41.39	52.68
P_6	38.09	32.39	52.00	34.48	31.48	43.95	33.35	31.27	40.88
P_7	18.14	14.57	28.90	15.99	14.20	22.97	15.34	14.12	20.63
P_8	27.37	19.95	44.79	22.18	18.62	32.74	20.86	18.36	29.33
P_9	19.53	13.31	35.82	15.21	12.41	24.50	13.98	12.20	20.84

As can be seen in Table 3, for a fixed sample plot, the lower and the upper C.I. values show a decreasing trend from $S_{est(1)}$ to $S_{est(3)}$. Such results are reasonable due to changing uncertainty concerning the number of singletons. In other words, the more uncertainty about the number of singletons it becomes, the smaller lower and upper C.I. values we obtain.

With regards to scenario I, P_5 has the maximum $S_{est(1)}$ with a value of 50.10 (Table 3). That is expected result because this sample plot consists of 41 species and, nine of them are rare species. P_1 has also the same number of observed species (Table 2).

However, its expected value remains almost the same since it has only one rare species. With regards to the number of observed species (S_{obs}), P_7 , P_8 and P_9 are the least species rich (Table 1). However, the

ordering of the sample plots from the largest to smallest is changed when the comparison is done by considering $S_{est(1)}$. For instance, $S_{est(1)}$ value of P_8 becomes greater than those of P_2 , P_3 and P_4 (Table 3).

The S_{est} results of Scenario II ($S_{est(1:2)}$) are smaller than those of Scenario I because this scenario assumes that a rare species is less likely to be a singleton. With regards to of $S_{est(1:2)}$, the ordering of the sample plots from largest to smallest is P_5 , P_1 , P_6 , P_3 , P_8 , P_2 , P_4 , P_7 and P_9 respectively. Scenario II and Scenario III does not indicate a considerable difference in terms of the ordering of sample plots. The change in ordering is solely between P_8 and P_2 . $S_{est(1:3)}$ values of the sample plots are the closest to their number of observed species when compared to the $S_{est(1)}$ and $S_{est(1:3)}$ values. That is expected result since a rare

species is a singleton with a 1/3 probability in accordance with the third scenario.

Consequently, we find the more uncertainty in the community data, the less information we gain.

Discussion and Conclusion

As reported by Magurran (2004), "Species richness is the simplest measure of diversity. Yet such simplicity is illusory". This illusory is due to negative bias. To overcome this problem, several bias-corrected estimators have been proposed (Smith & Van Belle, 1984; Shen & He, 2008; Gotelli & Chao, 2013; Chao et al., 2015). Those bias-corrected metrics have been designed for abundance or incidence data. However, such data collection is very costly and time-consuming for plant community inventory. On the contrary visual methods are faster and cheaper. BB scale is the most popular visual based data collection method and has frequently been employed by phytosociologist or vegetation ecologists (Chmura & Salachna, 2016). This method was initially developed in studies on the mountain ecosystems of Switzerland and on the Mediterranean ecosystems in France. Its pristine version inspired the work of Braun-Blanquet's students and followers, who applied the method on the European countries, Russia, Japan, North America, Pacific Islands, South America, North Africa, South Africa and West Australia (Guarino et al., 2018).

The unique integer (abundance) data in BB scale corresponds to "r" code. As mentioned in previously sections, some phytosociologists or plant ecologists use "r" only for singletons. However, the others prefer to use "r" for both singletons and/or a few species. Such description of "r" code means that we have unclear information how many singletons there are in a community or a sample plot.

More recently a remarkable method has been proposed by Chao et al. (2017). This method enables to estimate true species richness by using only the observed number of the species (S_{obs}) and the number of singletons (f_1).

The present study was conducted to indicate how to estimate true species richness by not only presenting this novel method (Chao et al., 2017) but also proposing an approach via that method when the number of singletons is uncertain.

To illustrate the proposed approach, the BB scale dataset collected from Yazılı Canyon Nature Park was used. The "r" codes were employed for the species that have equal or less than three individuals during the field survey. Three scenarios were established and, expected species richness of each sample plot was defined for each of the scenarios. The results indicate that with increasing uncertainty about the number of singletons, the difference between expected species richness and observed species richness decreases.

To estimate the species richness of the sample plots taken from Yazılı Canyon Nature Park, scenario III seems to be the best option due to existing maximum uncertainty concerning the number of singletons. If the "r" codes had been recorded for the species that had one or two species during the field survey in the Canyon, selection of the second scenario have been made sense.

Consequently, the proposed approach indicates that we can estimate the expected species richness under all conditions by selecting an appropriate scenario depends on the degree of uncertainty. Thus, it can be notified that all the BB scale data collected to this day by the ecologists and the phytosociologists are available in estimation of expected species richness owing to the method proposed by Chao et al. (2017) and its alternative version represented by the present study.

Acknowledgements

The author thanks to Assoc. Prof. Dr. Ahmet Mert for providing helpful in creating spreadsheet program.

Ethics Committee Approval

N/A

Conflict of Interest

The authors have no conflicts of interest to declare.

Funding

This The authors declared that this study has received no financial support.

References

- Camiz, S., Torres, P. & Pillar, V. D. (2017). Recoding and multidimensional analyses of vegetation data: a comparison. *Community Ecology*, 18(3), 260-279.
- Campbell, B. M. (1974). Phytosociological study of some forest patches on Table Mountain. University of Cape Town Honours project.
- Çekim, H. O. & Kadılar, C. (2020). In-type estimators for the population variance in stratified random sampling. *Communications in Statistics-Simulation and Computation*, 49(7), 1665-1677.
- Chao, A. (1987). Estimating the population size for capture-recapture data with unequal catchability. *Biometrics*, 783-791.
- Chao, A. & Jost, L. (2012). Coverage-based rarefaction and extrapolation: standardizing samples by completeness rather than size. *Ecology*, 93(12), 2533-2547.
- Chao, A. & Chiu, C. (2016). Nonparametric Estimation and Comparison of Species Richness, In: eLS. Wiley Online Reference in the Life Sciences. *John Wiley and Sons*, DOI: 10.1002/9780470015902.a0026329.
- Chao, A., Colwell, R. K., Chiu, C. H. & Townsend, D. (2017). Seen once or more than once: Applying Good–Turing theory to estimate species richness using only unique observations and a species list. *Methods in Ecology and Evolution*, 8(10), 1221-1232.
- Chao, A., Hsieh, T. C., Chazdon, R. L., Colwell, R. K. & Gotelli, N. J. (2015). Unveiling the species-rank abundance distribution by generalizing the Good-Turing sample coverage theory. *Ecology*, 96(5), 1189-1201.
- Chmura, D. & Salachna, A. (2016). The errors in visual estimation of plants cover in the context of education of phytosociology. *Chemistry-Didactics-Ecology-Metrology*, 21(1-2), 75-82.
- Eren, M. I., Chao, A., Hwang, W. H. & Colwell, R. K. (2012). Estimating the richness of a population when the maximum number of classes is fixed: a nonparametric solution to an archaeological problem. *PLoS One*, 7(5), e34179.
- Good, I. J. (1953). The population frequencies of species and the estimation of population parameters. *Biometrika*, 40(3-4), 237-264.
- Good, I. J. (2000). Turing's anticipation of empirical Bayes in connection with the cryptanalysis of the naval Enigma. *Journal of Statistical Computation and Simulation*, 66(2), 101-111.
- Gotelli, N. J. & Chao, A. (2013). Measuring and estimating species richness, species diversity, and biotic similarity from sampling data. *Encyclopedia of biodiversity, Elsevier*, 5, 195-211.
- Guarino, R., Willner, W., Pignatti, S., Attorre, F. & J. Loidi, J. (2018). Spatio-temporal variations in the application of the Braun-Blanquet approach in Europe. *Phytocoenologia*, 48(2), 239-250.
- Kadılar, C. & Çingi, H. (2007). Improvement in variance estimation in simple random sampling. *Communications in Statistics-Theory and Methods*, 36(11), 2075-2081.
- Magurran, A. E. (2004). Measuring biological diversity. *Blackwell Publishing: Oxford*. 8, 256, ISBN 0-632-05633-9.
- Miller, A. L., Diez, J. M., Sullivan, J. J., Wangen, S. R., Wisser, S. K., Meffin, R. & Duncan, R. P. (2014). Quantifying invasion resistance: the use of recruitment functions to control for propagule pressure. *Ecology*, 95(4), 920-929.
- Özel, G., Çingi, H. & Oğuz, M. (2014). Separate ratio estimators for the population variance in stratified random sampling. *Communications in Statistics-Theory and Methods*, 43(22), 4766-4779.
- Özkan, K. (2016). Application of information theory for an entropic gradient of ecological sites. *Entropy*, 18(10), 340.
- Pärtel, M., Szava-Kovats, R. & Zobel, M. (2011). Dark diversity: shedding light on absent species. *Trends in ecology & evolution*, 26(3), 124-128.
- Peet, R. K., Wentworth, T. R. & White, P. S. (1998). A flexible, multipurpose method for recording vegetation composition and structure. *Castanea*, 262-274.
- Poore, M. E. D. (1955). The use of phytosociological methods in ecological investigations: I. The Braun-Blanquet system. *Journal of Ecology*, 43(1), 226-244.
- Prieditis, N. (1997). Vegetation of wetland forests in Latvia: A synopsis. In *Annales Botanici Fennici, Finnish Zoological and Botanical Publishing Board*, 34, 91-108.
- Pyšek, P., Jarošík, V., Kropáč, Z., Chytrý, M., Wild, J. & Tichý, L. (2005). Effects of abiotic factors on species richness and cover in Central European weed communities. *Agriculture, Ecosystem and Environmental*, 109, pp. 1-8.
- Sarmah, C. K. (2017). Chao, jackknife and bootstrap estimators of species richness. *IJAMAA*, 12 (1), 7-15.
- Shen, T. J. & He, F. (2008). An incidence-based

richness estimator for quadrats sampled without replacement. *Ecology*, 89(7), 2052-2060.

- Smith, E. P. & Van Belle, G. (1984). Nonparametric estimation of species richness. *Biometrics*, 119-129.
- Vahdati, F., Mehrvarz, S. S., Naqinezhad, A. & Gholizadeh, H. (2014). How plant diversity features change across ecological species group? A case study of a temperate deciduous forest in northern Iran. *Biodiversitas*, 15(1), 31-38.
- Van der Maarel, E. (2007). Transformation of cover-abundance values for appropriate numerical treatment-Alternatives to the proposals by Podani. *Journal of Vegetation Science*, 18(5), 767-770.
- Werger, M. J. A. (1974). On concepts and techniques applied in the Zürich-Montpellier method of vegetation survey. *Bothalia*, 11(3), 309-323.
- Westhoff, V. & Van der Maarel, E. (1973). *The Braun-Blanquet approach*. In Classification of Plant Communities: Whittaker, R.H., Ed: Springer: Hague, 617-726, Netherlands.

S5) are rare (**r**) species. Two species (S1 and S2) are certainly singletons, and one species (S3) is singleton with 1/2 probability. The other species (S4 and S5) are singletons with 1/3 probability. To get the estimation results from the program for this hypothetical example, the values of 13 and 5 should be entered into **B3** and **B4** cells respectively (**B6 = B7 = 0**). The other values (for S1 and S2, E6=E7=F6=F7=1; for S3, G6=1 and G7=2; for S4 and S5, H6=I6=1 and H7=I7=3) should be also entered. After entering the number of replications into **B10**, the results are obtained from **D13**, **D14** and **D15** (Figure. A).

Appendix

Suppose that there is a sample plot including 13 species and, five of them (S1-

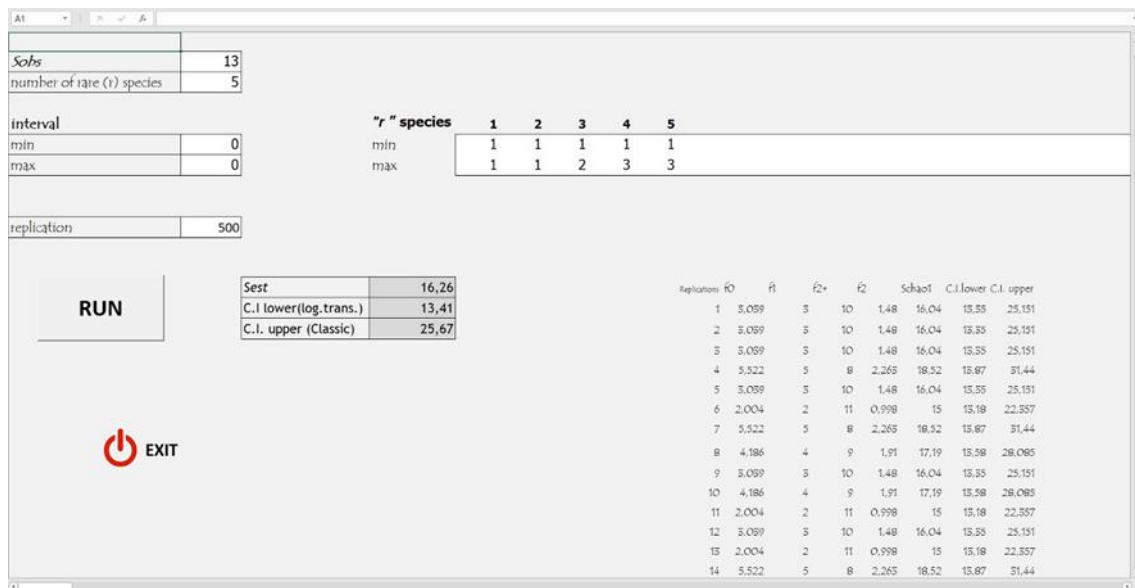


Figure. A. The results of the hypothetical community data