Bayesian and non-parametric estimations of the weed inventory in the cultivation of chrysanthemum using rigid grid quadrats

Estimación Bayesiana y no-paramétrica del inventario de arvenses en el cultivo de crisantemo usando cuadrados de una red rígida

Enrique Darghan[1], Sinha Surendra[2], Guido Plaza[1], and Julio Monroy[3]

ABSTRACT

Studies that involve the inventory of weeds are frequently carried out by students and professionals of the agricultural and/or environmental sciences with the principal objective of obtaining information on the distribution pattern, frequency, coverage, density or biodiversity of the species in a studied region. On many occasions, the only purpose consists of identifying those species that are considered important by farmers, perhaps because they are beneficial or because they compete with the crops in production. When the sampling of weeds is done using the quadrat method, some of the species which are present in the cultivated field of interest may not be sampled, meaning they will be absent in the final inventory. The principal objective of this article was to show how the quantity of weeds in an observed sample can be estimated Bayesian estimators, as well as non-parametric estimators, such as Chao 2, Jackknife, of the first and second order, and Bootstrap. The inventory estimation of the weeds using the Bayesian and classical proposals in the case of cultivation of chrysanthemum produced similar results, with 19 species in all of the estimators, except in the Mingoti estimators, which produced 18 weeds.

Key words: weed competition, sampling, biodiversity, cut flowers, statistical methods.

RESUMEN

Los estudios que involucran el inventario de arvenses son conducidos frecuentemente por estudiantes o profesionales de las ciencias agropecuarias y/o ambientales con el objetivo principal de obtener información acerca del patrón de distribución, frecuencia, cobertura, densidad o de la biodiversidad de las especies en una región de investigación. En muchas ocasiones el único propósito consiste en identificar aquellas especies que son consideradas importantes por los productores agropecuarios ya sea porque resultan benéficas o porque compiten con el cultivo en producción. Cuando se hace el muestreo de las arvenses usando el método del cuadrado, algunas de las especies que están presentes en el sembrado de interés podrían no ser muestreadas, por lo que estarían ausentes en el inventario final. El objetivo principal de este artículo es mostrar como la cantidad de especies arvenses en la muestra observada podría ser estimada al usar estimadores Bayesianos así como también estimadores clásicos tales como el Chao 2, Jackknife de primer y segundo orden y el Bootstrap. La estimación del inventario de arvenses utilizando las propuestas Bayesiana y no-paramétrica en el caso del cultivo de crisantemo rindieron resultados similares, con 19 especies en todos los estimadores excepto en el estimador Mingoti, el cual rindió 18 especies arvenses.

Palabras clave: competición de arvenses, muestreo, biodiversidad, flores de corte, métodos estadísticos.

Introduction

One of the environmental problems of major concern in the world in recent years is the loss of biodiversity as a consequence of human activities, whether it be in a direct way (overexploitation) or indirect way (habitat alteration). In a certain manner, communication media have impacted governments as well as society in general; to such an extent that, now, it is considered a priority to direct major efforts towards conservation program (Popescu, 2015). The base for an objective analysis of the biodiversity is correct evaluation and monitoring (Smith *et al.*, 2012). Biodiversity can be defined as "the variability among the living organisms of all the sources, including among others, the terrestrial, marine and other aquatic ecosystems, as well as the complex ecologies of which they form parts; this includes diversity within the species, between species and the ecosystems" (Moreno, 2000).

According to Gaston (1996) and Mingoti (2000), the number of species is the measurement most frequently used to evaluate biodiversity due to several reasons a) the abundance of the species reflects different aspects of the

biodiversity, b) In spite of this, there exist many approximations in order to define the concept of a species, its significance is widely understood (Aguilera and Silva, 1997), c) Besides certain groups, species are easily detectable and measurable and d) even when the taxonomical knowledge is not complete (especially for groups such as fungus, insects and other invertebrates in tropical zones), there exists much data that are available on the number of species.

Studies on the measurement of biodiversity are centered on the search for parameters in order to characterize it as a property emerging from the ecological communities (Izsák and Papp, 2000). However, communities are not isolated in a neutral surrounding. In every geographical unit, a variable number of communities is found; in order to understand the changes of the biodiversity in relation to the structure of the landscape, it may be highly useful to separate the alpha, beta and gamma components (Snedecor and Cochran, 1980) to measure and monitor the effects of human activities. The alpha diversity is the richness of a species of a particular community that is considered homogenous; the beta diversity is the level of change or replacement in the species composition between different communities in a landscape; and the gamma diversity is the richness of the species of the set of the communities which integrate a landscape resulting from the alpha diversity as well as the beta diversity (Heltshe and Forrester, 1983).

In the present research, the diversity alpha was considered from the point of view of the specific abundance of the species more so than the structure of the community (value of the importance of each species). The objective for obtaining these measures is to provide the reader non-parametric and Bayesian alternatives for the prediction of the number of species of weeds based on the point estimation of the same, which is obtained by a simple counting using the quadrat method with one rigid grid mapped on the crop.

Sampling with quadrats is generally used to quantitatively estimate the biodiversity of the species in a studied region. The objective of this sampling is generally related to the estimation of the abundance of species. A counting using quadrats provides a structured way to estimate the abundance of species to estimate the population size and/or assure the abundance of species and the diversity of a biotype. Quadrats provide a simple and reproducible method that is appropriate for achieving a broad set of statistical tests, making this methodology an ideal strategy for long term monitoring. Among the advantages of sampling with quadrats with a rigid grid, the non-destructiveness of this sampling stands out; it can be applied to a broad set of

habitats, can be repeated easily, which provides consistence to the sampling, provides a form to estimate abundance, does not require any special equipment, there are no overlapping quadrats such as may occur in sampling with quadrats thrown in the field and produces a set of robust data for the statistical analysis (Haas *et al.*, 2006).

The quadrat size may vary depending on the objective of the study and the following conditions: (i) the quantity and distribution of the species to be sampled, with species that are big or dispersed requiring different sized quadrats. (ii) The heterogeneity of the community in terms of the dispersion of the species, the quadrat should cover mainly a representative sample of the community.

According to Eleftheriou (2013), the determination of the number of samples to be used should be done with a pilot study of the area and, in this way, useful information for the analysis may be obtained. The number of quadrats for a reliable monitoring can be determined using a power analysis. The power analysis is a statistical technique that allows for the estimation of the number of samples required to detect the change level (Smith and van Belle, 1984). Also, the accumulated species curve can be used to assure when a population has been sufficiently sampled by a number of quadrats. The number of species accumulated is registered with each increment in the number of quadrats until a point is reached where all of the common species have been identified and a further increase in the number of quadrats would not lead to a significant increase in the number of species. Gamble (1984) gave a preliminary guide for the minimum number of samples as "that which, if it is duplicated, may yield only a 10% increase in the information" Kingsford and Battershill (1998) stated that 10 sampled quadrats within a discrete area will provide adequate precision for detecting changes in the complete community.

In order to achieve an appropriate coverage of a studied region, it may be appropriate to divide the area into compartments (such as a rigid grid) and take random sample from each compartment. For a better explanation of the different methods of random sampling, Kingsford and Battershill (1998) can be consulted.

Knowledge on the weeds in the cultivation of *Chrysanthemum* sp. is very important in order to implement strategies for the management of the cultivation directed primarily at its control in order to diminish the effects on the loss in yield as well as to keep the weeds from acting as a host that carries plagues and diseases that may reduce the quality of the flowers. The current study emerged from the necessity

to complement the descriptive work developed by Sánchez (2003), who carried out a counting of weeds in the cultivation of chrysanthemum and the simple description of the species found in an unit of production of the municipality Andrés Bello in Táchira State, Venezuela.

The application of these methods can help weed specialists as well as agronomists, botanists and others whose interest lies in this subject to obtain a predicted measurement of the number of species instead of a point estimation based only on the counting of the species found in the total of the quadrats mapped in the area. The procedure is simple to adopt if free software is used, which is illustrated herein, and, with only one optimization tool, such as Excel (Microsoft), the Bayesian predictions can be obtained.

Materials and methods

This research was carried out in the flower producing area of chrysanthemum in the municipality Andrés Bello in Táchira State, Venezuela, located at 1,150 m a.s.l. with a temperature range between 18 and 22ºC, classified as a humid, low mountainous forest. For the random sampling, the quadrat method (1 x 1 m) was used. A rigid grid method of survey was adopted on a real scale to establish the mapping units on the plots. The interest quadrats were placed on four points at random in plots, 1.20 m in width, in a representative unit of production in the zone of which four plots were selected at random for a total $n = 12$ quadrats; besides since the length of the plots were variables, N was taken as unknown, which did not affect the construction of estimators, since as it was said previously, none of the estimators depends on N.

The sampling area, 12 m were selected for which 12 random numbers were selected in the total of squares representing cultivation. The sowing was done in beds of 1.2 of width with 5 plots with a distance between plants of 12.5 cm. For the purpose of this weed study, only the measurement of the number of species found was done at the time when the growth of the chrysanthemum had 5 cm of inflorescence.

Non-parametric estimators

The non-parametric estimators were obtained manually and were contrasted with the values of software EstimateS v.13 (Colwell, 2013). In the case of the Bayesians estimators, the optimization tool solver of Excel was used.

Chao 2

The Chao2 ($\hat{S}_{C2}$) estimator is based on the incidence (Chao, 1984). This requires the presence-absence data of a species in a given sample and not of the abundance, which means if only the species is present, this estimator is based on the concept that the rare species carry the information of the species not sampled. The Chao 2 estimator is expressed as:

$$\hat{S}_{C2} = s' + n_1^2/2n_2 \tag{1}$$

where n_1 is the number of species that occurs only in a sample ("unique" species), n_2 is the number of species that occurs in exactly in two samples ("double" or "duplicated" species) and s' is the number of species observed in the quadrat samples. The corrected formula which is applied when the number of doubles is zero is

$$\hat{S}_{C2} = s' + \frac{n_1(n_1-1)\ (n-1)}{n(2n_2+1)} \tag{2}$$

where n is the number of quadrats that were mapped.

JackKnife of first and second order

The JackKnife method was developed initially as a generic non-parametric estimator of bias and standard error. The generalized equations of the JackKnife estimators of k-th order were derived by Gray and Schucany (1972) and the use of the JackKnife estimator in the estimation of abundance of the species dates back to the moment in which Zahl (Zahl, 1977), treated rectangular plots of vegetation as independent samples that could be fractionated for the estimation of its diversity. According to the manual of assessment of biodiversity of the OECD (2002), this procedure assumes that there exists a random sample of independent quadrats rather than a sample of individuals. The random selection of the quadrats is in fact a random sample of the space or region of investigation (Heltshe and Forrester, 1983). Moreover, the non- parametric JackKnife estimator does not assume any relationship between the species within a quadrat and does not make any assumption about the fundamental distribution of the species.

The JackKnife estimator of the first order is a function of the number of rare species that are found in the sample. Its calculation involves the number of species that are present only in one quadrat. The JackKnife estimator of the second order on the other hand takes into consideration the number of species found in only one quadrat and the number of species found in two quadrats. The theoretical basis of JackKnife is that the estimator of parameter of interest is obtained from n samples of size n-k, taking in account that each sample is generated by the elimination of k of the n original quadrats, where k =1 or k=2 depending on whether or not a JackKnife of the first order or the second

order, respectively, is being used. In the sequel we show the expressions which permitted to obtain the abundance estimation based on the procedure of JackKnife of first and second order respectively.

$$\hat{S}_{J1} = s' + n_1(n-1)/n; \quad \hat{S}_{J2} = s' + \frac{n_1(2n-3)(n-1)}{n} - \frac{n_2(n-2)^2}{n(n-1)} \quad (3)$$

Bootstrap

The bootstrap estimator was proposed by Efron around 1980 and is simply a procedure of quadrats with replacement. It should also be pointed out that this estimator does not assume any relationship between the species within a quadrat nor the suppositions about the statistical distribution of the species (Smith and van Belle, 1984). The expression which permitted to obtain the abundance estimator based on the Bootstrap procedure is given by

$$\hat{S}_B = s' + \sum_{k=1}^{s'} (1-p_k)^n \quad (4)$$

where p_k is the proportion of samples that contain k species.

Bayesians estimators

A pair of estimators are proposed as a solution of the problem of the estimation of abundance of the species from the point of view of the Bayesian approach. Suppose that the sample consists of n quadrats thrown at random in a study area where there exist N locations in which the quadrats might fall. For each quadrat, a certain number of different species of weeds can be reported within the cultivation that is being evaluated. At the end of the study, there will be s' different species in all the sampling performed. Each species could have appeared more than once. Let n_i be the number of species that appeared exactly in i sampled quadrats, $i=1,2,\ldots, n$. Therefore, $n_1+n_2+\ldots+n_n = s'$. Let S be the true value of the distinct species which could appear in an inventory of weeds. The true value of S is unknown and must be estimated by $\hat{S}$. The number s' observed in the sample is an estimator for S; however, it has been proven that s' subestimates S. Some estimators have been proposed to correct the bias of s' and other estimators and in this way obtain better results. Mingoti and Meeden (1992) and Mingoti (1999) proposed some Bayesian estimators that in general produced better results than s' and other estimators such as the aforementioned JackKnife, Bootstrap and Chao.

Mingoti and Meeden (empirical Bayesian)

Let n_1 be the number of distinct species that were found in one and only one quadrat in the sample. Then, the empirical estimator of Bayes for the true value of S is given by $\hat{S}_{MM}$, which is defined as:

$$\hat{S}_{MM} = s' + \frac{n_1}{n\alpha}(n+\beta-1)\left[1 - \frac{\Gamma(N+\beta)}{\Gamma(N+\alpha+\beta)} \frac{\Gamma(n+\alpha+\beta)}{\Gamma(n+\beta)}\right] \quad (5)$$

where $\Gamma(.)$ is the gamma function and the constants $\alpha>0$ and $\beta>0$ are the parameters of the beta distribution used as a prior distribution for the technical construction of estimator, which describes the probabilistic behavior of the unknown value of p_i, defined as the probability of that the species s_i may appear in a quadrat of the universe of possible quadrats, $i=1, 2,\ldots, S$. Given S, the probabilities $p_1, p_2, \ldots, p_S$ are assumed to be random variables, independent and identically distributed with the density function Beta. In order that the estimator $\hat{S}_{MM}$ be more attractive for practical uses, Mingoti and Meeden (1992) showed that the parameters α and β of the Beta distribution may be estimated considering that given the value of s', the random vector $(n_1, n_2, \ldots, n_n)$ has a multinomial distribution with parameters $(q_1, q_2,\ldots,q_n)$, with $0< q_x <1$, $x =1,2,\ldots,n$, y $q_1+q_2+\ldots+q_n=1$, where

$$q_x = \frac{\binom{n}{x}\Gamma(x+\alpha)\,\Gamma(n+\beta-x)}{\sum_{i=1}^{n}\binom{n}{i}\Gamma(i+\alpha)\,\Gamma(n+\beta-i)}, \quad (6)$$

with which the maximum likelihood estimators of the parameters (α, β) may be obtained by maximizing the likelihood function $f(n_1, n_2, \ldots n_n/s')$ with respect to α and β, where:

$$f(n_1, n_2, \ldots n_n/s') = \frac{s'}{(\Pi_{x=1}^{n} n_x!)}\left(\prod_{x=1}^{n}(q_x)^{n_x}\right). \quad (7)$$

Mingoti (Bayesian)

An admissible estimator for the true value of S: the number of different species in the population is given by $\hat{S}_M$, which is defined as:

$$\hat{S}_M = \frac{s' + Rq\gamma_0}{1-q\gamma_0}; \quad \gamma_0 = \frac{\Gamma(\alpha+\beta)\,\Gamma(n+\beta)}{\Gamma(\beta)\,\Gamma(n+\alpha+\beta)}, \quad 0<\gamma_0<1,\, \alpha>0,\, \beta>0 \quad (8)$$

The estimator $\hat{S}_{MM}$, under the assumption of the given value of S, the probabilities $p_1, p_2, \ldots, p_s$ are distributed identically and independently with the Beta density, with parameters (α, β), with $\alpha>0$ and $\beta>0$, where p_i is interpreted as in the case of empirical Bayesian estimator. The constants $R>0$, $0< q <1$, are related to the binomial negative distribution used as a prior distribution for the true value of S in the construction of the Mingoti estimator.

The q parameter represents the prior probability that any different species may appear in a particular quadrat, in

other words, the proportion of different species in the population. The γ_0 constant represents the difficulty in observing a particular species in the sampling procedure. Values of γ_0 near zero makes $\hat{S}_{MM}$ to tend to s', which implies the belief of the investigators that all the species were easy to obtain and that all the species will appear in the sample. Values of γ_0 near one describe a population where a large number of different species may be difficult to observe in the sample, in this case the value of s' will be much lower than the true value of *S*. The parameters α and β are obtained and interpreted in the same manner as the estimator $\hat{S}_{MM}$ so that γ_0 can be obtained easily. According to the "*ad hoc*" procedure suggested by Mingoti (1999), the parameters (*R,q*) can be estimated using the estimators $\hat{R}$ and $\hat{q}$, respectively, using $\hat{R} = n/s'$ and $\hat{q} = n_1/s'$; in the case of $\hat{R}$ is taken its whole part. If this final value is less than the unity, then $\hat{R} = 1$. It is important to emphasize that $\hat{S}_M$ does not depend on *N*; therefore, it can be used in the case of unknown or very large.

The Bayesian estimators presented in this article are also discussed in the given references and the steps for their construction appear with more details and, therefore, they will not be presented in this article. The principal objective

of the present research was to show how they can be applied in the agricultural and animal husbandry sciences for the estimation of weed inventories and in the environmental field for the estimation of the biodiversity.

Results

In Tab. 1, the species of the weeds found in the sampling and their frequency of appearance are presented and, even when the importance did not depend on the frequency but on the total of the sampled species, the description of each species suggested to the investigators that the use of the 1 m^2 quadrat was appropriate since some of the botanical characteristics of the species could imply the use of quadrats of a greater area, Besides due to the type of cultivation the size of 1 m^2 results more convenient (Mostacedo and Fredericksen, 2000). In many similar studies, this result appeared to be sufficient in order to evaluate the species of weeds in the crop; however, the present proposal will extract greater information than that presented in Tab. 1.

Table 2 shows the frequency of itemized appearance by the launching of the quadrats for all the species found, from this last table the useful values are selected in the

TABLE 1. Scientific names and frequency of appearance of the weeds based on the sampling.

Nº	Species	Count	Nº	Species	Count
1	*Paspalum notatum*	7	10	*Ageratum conyzoides*	5
2	*Paspalum conjugatum*	21	11	*Oxalis corniculata*	127
3	*Rottoboellia cochinchinensis*	35	12	*Polygonum hidropiperoides*	31
4	*Cynodon* spp.	39	13	*Rumex crispus*	21
5	*Eleusine indica*	8	14	*Portulaca oleracea*	2
6	*Heliotropium indicum*	7	15	*Brassica alba*	1
7	*Commelina nudiflora*	6	16	*Lepidium virginicum*	32
8	*Taraxacum officinale*	9	17	*Leonorus sibiricus*	10
9	*Bidens pilosa*	14	18	*Cyperus* spp.	15

TABLE 2. Presence of the species according to the sampled quadrats.

Quadrat	Species																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	0	2	1	6	1	1	1	1	1	2	11	2	3	0	0	2	0	1
2	0	0	4	9	1	1	0	2	3	0	8	2	2	0	0	2	0	1
3	1	3	2	4	1	0	1	0	1	0	11	1	2	1	0	2	0	2
4	1	2	3	2	0	2	1	0	2	0	6	5	1	1	0	6	2	1
5	0	2	0	5	1	0	0	0	2	0	8	2	2	0	1	2	0	2
6	0	2	1	3	1	1	0	0	0	1	10	3	3	0	0	5	2	0
7	2	2	5	1	1	0	0	0	2	1	12	4	2	0	0	1	1	1
8	1	2	3	0	0	1	2	0	1	0	10	4	0	0	0	2	1	1
9	0	2	3	2	0	0	0	1	0	1	12	3	2	0	0	2	1	2
10	0	2	5	3	0	0	1	1	0	0	14	1	1	0	0	0	1	1
11	0	0	4	2	1	1	0	1	2	0	11	3	1	0	0	5	2	2
12	2	2	4	2	1	0	0	3	0	0	14	1	2	0	0	3	0	1

In the sample of 12 quadrats (*n* = 12), a total of 18 different species appeared. (s' = 18).

TABLE 3. Distribution of the sampled species according to the number of quadrats.

s_i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
X_i	5	10	11	11	8	6	7	6	8	4	12	12	11	2	1	11	7	11

TABLE 4. Distribution of the observed species in exactly x quadrats.

x	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
n_x	0	1	1	0	1	1	2	2	2	0	1	5	2	0	0	0	0	0	0

calculation of each one of the estimators described in the introduction.

Table 3 shows the distribution of the observed species in the sampling. (X_i) according to the number of quadrats in which the species s_i (i=1,2,…,18) appeared. In order to obtain each value of X_i suffices to count the non-null entries of Tab. 2 for each species, for example, in species 11, zero did not appear in any of the 12 quadrats, hence X_i=12.

Table 4 presents the number of observed species in exactly x quadrats, in other words n_x (this is equivalent to count the cells for X_i which were repeated, and since some species could not appear in the quadrats, it was necessary to incorporate x=0; for example, 11 was observed in 5 cells and 0 did not appear in any of the cells of row X_i in the Tab. 3).

With Tab. 3 and 4, the estimates of the number of species considered were obtained in this article using EstimateS, available online at the address that appears in the references. By using the optimization tool, Solver of Excel, the maximum likelihood estimator for α and β was obtained: α=1.43 and β=0.80. Later, by substituting these estimators in Eq. 5, the estimation of the number of weeds was obtained based on the empirical Bayesian estimator.

For the Bayesian estimators, the estimated values of α and β of the Bayesian empiric estimator were used to obtain γ_0. The substitution of the estimators of the Beta distribution in Eq. 8 gave γ_0=0,025. In the case of the estimated values of the parameters (R, q), these were obtained from the equation R=[n/s']=[12/18]= 0 (floor function), with which was taken R=1 (suggested before) and from q=n_1/s'=1/18.

Table 5 presents the estimations obtained for the number of weeds species found in the cultivation of chrysanthemum in the studied region.

In the sequel each one of the calculation of non-Bayesian estimators are presented:

- Chao 2: from Tab. 4 are obtained the values of de n_1 and n_2, in this way the estimator Chao 2 not corrected

TABLE 5. Non-parametric and Bayesian estimations of the number of species.

Estimator	Estimation
Chao 2	18.50
Jackknife of first order	18.92
Jackknife of second order	18.99
Bootstrap	18.47
Mingoti and Meeden	18.64
Mingoti	18.03

for bias is: $\hat{S}_{C2} = s' + (n_1^2/2n_2) = 18 + (1^2/2(1)) = 18.50$ (by definition).

- Jackknife of the first order: from Tab. 4 the value of n_1 is obtained (species which appear in exactly one quadrat), in this way: $\hat{S}_{J1} = s' + n_1\left(\dfrac{n-1}{n}\right) = 18 + 1((12-1)/12) = 18.92$.

- Jackknife of second order: from Tab. 4, the values of n_1 (species which appear in exactly in one quadrat) and n_2 (species which appear in exactly two quadrats) are obtained:

$$\hat{S}_{J2} = s' + \left[\frac{n_1(2n-3)}{n} - \frac{n_2(n-2)^2}{n(n-1)}\right] = 18 + \left[\frac{1(2(12)-3)}{12} - \frac{1(12-2)^2}{12(12-1)}\right] = 18.99.$$

- Bootstrap: from Tab. 3 the values to be used in the following equation are obtained

$$\hat{S}_B = 18 + \left[\begin{array}{l} \left(1-\frac{5}{12}\right)^{12} + \left(1-\frac{10}{12}\right)^{12} + \left(1-\frac{11}{12}\right)^{12} + \left(1-\frac{11}{12}\right)^{12} + \left(1-\frac{8}{12}\right)^{12} + \left(1-\frac{8}{12}\right)^{12} + \\ \left(1-\frac{7}{12}\right)^{12} + \left(1-\frac{6}{12}\right)^{12} + \left(1-\frac{8}{12}\right)^{12} + \left(1-\frac{4}{12}\right)^{12} + \left(1-\frac{12}{12}\right)^{12} + \left(1-\frac{12}{12}\right)^{12} + \\ \left(1-\frac{11}{12}\right)^{12} + \left(1-\frac{2}{12}\right)^{12} + \left(1-\frac{1}{12}\right)^{12} + \left(1-\frac{11}{12}\right)^{12} + \left(1-\frac{7}{12}\right)^{12} + \left(1-\frac{11}{12}\right)^{12} \end{array}\right] = 18.47.$$

Once defined, the number of samples taken using the quadrats, non-parametric and Bayesian estimators was used for the estimation of the number of weeds species in the cultivation of chrysanthemum, obtaining in all the cases used similar estimations of the number of species, which are, 19 species in the most of the estimations, with the exception of the Bayesian, developed by Mingoti, with 18 species. The number of observed species was 18 with the sampling; however, based on the simultaneous appearance of some species (a fact considered in the estimations

which were used) and in the interest to recognize all the species which appeared in the sowing, a visual recognition was done in the cultivation area (0.2 ha) and a total of 19 species was found, a value estimated by all of the methods of estimation used in this case. With this additional study, validity of the estimation is given, as seen in this document.

Discussion

Reducing herbicide inputs in crops is a major objective in agriculture. The extensive and abusive use of herbicides has raised concerns about environmental safety and conservation of biodiversity on farmland. In this sense, the study of weeds plays an important role in the development of conservation policies in each region. Several modern methods have been proposed for this purpose, for example, the image analysis, sensors handling in the field of precision agriculture (Schepers and Holland, 2012), as well as classic and sophisticated Bayesian statistical methods (Rinella and Luschei, 2006).

The non-parametric Bootstrap and Bayesian estimators yielded a number of species, such as the number of observed species in the sampling with the quadrats, while the non-parametric estimators Jackknife of first and second order, the Chao 2, and the empirical Bayesian estimator yielded one more species in comparison with that observed in the sampling. With this, two-thirds (67%) of the estimations yield 19 species, one more than the species found with the quadrats, as illustrated in Tab. 3. By comparing these results with the exhaustive observation that was done in the study zone, an additional species was found that was not sampled with the quadrats, namely the "bledo" known by its scientific name as *Amaranthus* spp. With this additional species, the 19 observed species were obtained in all of the plots, with which the sampling with the quadrat permitted to find the totality of the species existing in the cultivation, and with the estimation it became possible to predict the existence of one more species, which in this case corresponded to "bledo" (pigweed). The more frequent specie was *Oxalis corniculata,* known as "pan de cuco" (Yellow sorrel with a 33.16% of appearance) followed by the genus *Cynodon* , with a 10.18% of the counting obtained within the quadrats. Masís and Madrigal (1994) presented a list of weeds where some genera and species appeared in the sampling, which was consistent with this research, moreover, López (2009) pointed out six of the 19 species listed as weeds in Chrysanthemum, however, were identified only weeds but is not any method for estimating species mentioned.

It is important to add that, in the case of the Bayesian estimation, the choice of the parameters (R, q, γ_0) reflected the knowledge of the investigator of the region where the research was carried out, however, an investigator could use past data to select the values at prior about R, q and γ_0, in spite of that in this case the proposal of estimation "*ad hoc*" suggested by Mingoti (1999) was used. Finally, it is well known in agriculture that the association or segregation that could exist between weed species due to which it is important to know the pattern of these associations (between weeds and between weeds and crops), since in obtaining of the Bayesian estimators, independence between species was assumed and as well as equiprobability that a species may appear in any of the selected quadrats (Monaco *et al.*, 2002).

Conclusions

The estimation of the number of species using the nonparametric and Bayesian methods produced values very close to each other when the quadrats were used as the sampling method; as a matter of fact, it may be observed how the Chao 2 estimate did not correct for bias, the jackknife and the empirical Bayesian estimators obtained approximately an estimation of 19 species, while the bootstrap and the Bayesian estimators estimated 18 species in the studied region, with a total of 12 quadrats sampled. The estimation of the weed species using any of these estimators was more informative not only in the necessary tabular requirements for obtaining estimators, which reflects the appearance of the species in each quadrat, the species which repeatedly go appearing with each quadrat, so that the process of nonparametric and Bayesian estimation results much more informative from botanical point of view and therefore very convenient at the moment of planning the control of weeds within the cultivation of economic importance.

Literature cited

Aguilera, M. and J.F. Silva. 1997. Especies y biodiversidad. Interciencia 22, 299-306.

Chao, A. 1984. Nonparametric estimation of the number of classes in a population. Scand. J. Statist. 11, 256-270.

Colwell, R.K. 2013. EstimateS: statistical estimation of species richness and shared species from samples, Version 9.1, user's guide and application. Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, CT.

Eleftheriou, A. 2013. Methods for the study of marine benthos. 4th ed. Wiley Blackwell, Chichester, UK. Doi: 10.1002/9781118542392

Gamble, J.C. 1984. Diving. pp. 99-139. In: Holme, N.A. and A.D. McIntyre (eds.). Methods for study of marine benthos. Blackwell Scientific Publications, Oxford, UK.

Gaston, K.J. 1996. Biodiversity: a biology of numbers and differences. Blackwell, Science, Oxford, UK.

Gray, H.L. and W.R. Schucany. 1972. The generalizad jackknife statistic. Marcel Dekker, New York, NY.

Haas, P.J., Y. Liu, and L. Stokes. 2006. An estimator of number of species from quadrat sampling. Biometrics 62, 135-141. Doi: 10.1111/j.1541-0420.2005.00390.x

Heltshe, J.F. and N.E. Forrester. 1983. Estimating species richness using the jackknife procedure. Biometrics 39, 1-11. Doi: 10.2307/2530802

Izsák, J. and L. Papp. 2000. A link between ecological diversity indices and measures of biodiversity. Ecol. Modell. 130, 151-156. Doi: 10.1016/S0304-3800(00)00203-9

Kingsford, M. and C. Battershill. 1998. Studying temperate marine environments: a handbook for ecologists. Canterbury University Press, Christchurch, NZ.

López M., N. 2009. Malezas asociadas a plantas ornamentales. Fitosanidad 13, 233-236.

Masís, C.E. and R. Madrigal. 1994. Lista preliminar de malezas hospedantes de trips (thysanoptera) que dañan al (*Chrysanthemum morifolium*) en el valle central de Costa Rica. Agron. Costarr. 18, 99-101.

Mingoti, S.A. and G. Meeden. 1992. Estimating the total number of distinct species using presence and absence data. Biometrics, 48, 863-75. Doi: 10.2307/2532351

Mingoti, S.A. 1999. Bayesian estimator for the total number of distinct species when quadrat sampling is used. J. Appl. Statist. 26, 469-483. Doi: 10.1080/02664769922359

Mingoti, S.A. 2000. A stepwise Bayesian estimator for the total number of distinct species in finite populations: sampling by elements. J. Appl. Statist. 27, 651-670. Doi: 10.1080/02664760050076461

Monaco, T.J., S.C. Weller, and F.M. Ashton. 2002. Weed science: principles and practices. 4th ed. John Wiley & Sons, New York, NY.

Moreno, C.E. 2000. Manual de métodos para medir la biodiversidad. Universidad Veracruzana, Xalapa, Mexico.

Mostacedo, B. and T.S. Fredericksen. 2000. Manual de métodos básicos de muestreo y análisis en ecología vegetal. BOLFOR, Santa Cruz de la Sierra, Bolivia.

OECD, Organisation for Economic Cooperation and Development. 2002. Handbook of biodiversity valuation: a guide for policy makers. Paris.

Popescu, O. 2015. United Nations decade on biodiversity: strategies, targets and action plans. Urban. Arhit. Constr. 6, 37-50.

Rinella, M.J. and E.C. Luschei. 2006. Hierarchical Bayesian methods estimate invasive weed impacts at pertinent spatial scales. Biol. Invasions 9, 545-558. Doi: 10.1007/s10530-006-9057-x

Sánchez, D. 2003. Inventario poblacional de arvenses en el cultivo de crisantemo (*Eranthemun* sp.) del Municipio Andrés Bello del Estado Táchira. Undergraduate thesis. Universidad Nacional Experimental del Táchira, San Cristobal, Venezuela.

Schepers, J.S. and K.H. Holland. 2012. Evidence of dependence between crop vigor and yield. Precis. Agric. 13, 276-284. Doi: 10.1007/s11119-012-9258-5

Smith, E.P. and G. Van Belle. 1984. Nonparametric estimation of species richness. Biometrics 40, 119-129. Doi: 10.2307/2530750

Smith, F.P., R. Gorddard, A.P.N. House, S. McIntyre, and S.M. Prober. 2012. Biodiversity and agriculture: production frontiers as a framework for exploring trade-offs and evaluating policy. Environ. Sci. Policy 23, 85-94. Doi: 10.1016/j.envsci.2012.07.013

Snedecor, G.W. and W.G. Cochran. 1980. Statistical methods. 7th ed. Iowa State University Press, Ames, IA.

Zahl, S. 1977. Jackknifing an index of diversity. Ecology 58, 907-913. Doi: 10.2307/1936227