

# Marine: summary ecological statistics

Version 1.0



This specification was prepared by Shane Geange in 2017.

## Contents

Synopsis .....	2
Ecological measures .....	3
Abundance, frequency and density .....	3
Metrics of diversity .....	4
Sex ratios and age (or size) structure .....	9
Biomass .....	11
Catch per unit effort (CPUE) .....	11
Calculating summary ecological statistics from marine monitoring data .....	13
Measures of location and spread .....	13
Response ratios and effect sizes .....	18
Presenting data .....	21
Presenting data in tables .....	21
Presenting data in figures .....	23
What statistical analysis should I use? .....	26
How should I present statistical results as written text? .....	28
References and further reading .....	30

### Disclaimer

This document contains supporting material for the Inventory and Monitoring Toolbox, which contains DOC's biodiversity inventory and monitoring standards. It is being made available to external groups and organisations to demonstrate current departmental best practice. DOC has used its best endeavours to ensure the accuracy of the information at the date of publication. As these standards have been prepared for the use of DOC staff, other users may require authorisation or caveats may apply. Any use by members of the public is at their own risk and DOC disclaims any liability that may arise from its use. For further information, please email [biodiversitymonitoring@doc.govt.nz](mailto:biodiversitymonitoring@doc.govt.nz)

## Synopsis

This component of the marine module is a compilation of basic topics in statistics that are intended to be used as a supplement to DOC's Biodiversity Assessment Framework (Lee et al. 2005) and the methodologies presented within the marine module of the Inventory and Monitoring Toolbox<sup>1</sup>.

The Biodiversity Assessment Framework uses indicators such as species composition and diversity or contaminants to indicate how healthy an ecosystem is. Each indicator is informed by several measures that are in turn informed by data elements. For example, the species composition and diversity indicator may be informed by a measure of change in species diversity. The first part of this document details how to calculate the following measures:

- Abundance, frequency and density
- Metrics of diversity
- Sex ratios and age (or size) structure
- Biomass
- Catch per unit effort (CPUE)

The second part of this document details how to calculate summary ecological statistics from data generated by many of the monitoring methodologies described elsewhere in the marine module. These summary statistics include measures of location and spread, and response ratios and effect sizes. Although response ratios and effect sizes are frequently under-used in the reporting of ecological results, they are particularly important. Unlike *P*-values that merely report the probability of a difference between groups occurring due to chance, response ratios and effect sizes provide an estimate of the magnitude of difference between groups, which is often the variable most of interest to ecologists and conservation managers.

The last section of the document provides guidance on the appropriate uses of tables and figures, and discusses some approaches for ensuring they are clear and easy to interpret. The document concludes with advice on selecting statistical analyses appropriate for study objectives, and how to present statistical results.

Although there are equations presented within this document, these have been kept to a minimum. Data elements are illustrated with appropriate examples, and we have tried to achieve a balance between being concise and being thorough. For more detailed information on statistical analyses, this document should be read and used alongside more traditional statistical texts or primary literature, which are identified throughout this document.

---

<sup>1</sup> <http://www.doc.govt.nz/our-work/biodiversity-inventory-and-monitoring/marine/>



## Ecological measures

### Abundance, frequency and density

*Abundance*, *frequency* and *density* are ecological measures referring to the representation of a single species within a particular ecosystem or sampling unit.

#### Abundance

In ecology, abundance is typically measured as the number of individuals per sample. For example, the abundance of Sp.1 in Transect 2 of Table 2 is 3 individuals. A variety of sampling methods can be used to measure abundance, and may include transects, quadrats, time counts or pots.

How abundances of different species are distributed within an ecosystem is referred to as relative species abundances (see '[Metrics of diversity](#)' below).

#### Frequency

Frequency refers to the number of samples within which a particular outcome occurs. For example, in Table 1 the frequency with which more than 10 snapper occur in the 16 transects outside the marine reserve is 13 (or 81%).

Table 1. Simulated snapper (*Pagrus auratus*) abundances inside and outside of a marine reserve, as measured on replicate 500 m<sup>2</sup> transects.

Site	Transect	Protected	Abundance	Site	Transect	Protected	Abundance
1	1	outside	10	5	1	inside	14
1	2	outside	14	5	2	inside	17
1	3	outside	12	5	3	inside	16
1	4	outside	11	5	4	inside	18
2	1	outside	15	6	1	inside	17
2	2	outside	12	6	2	inside	18
2	3	outside	13	6	3	inside	16
2	4	outside	12	6	4	inside	17
3	1	outside	13	7	1	inside	16
3	2	outside	12	7	2	inside	15
3	3	outside	16	7	3	inside	18
3	4	outside	10	7	4	inside	14
4	1	outside	9	8	1	inside	15
4	2	outside	11	8	2	inside	19
4	3	outside	14	8	3	inside	17
4	4	outside	11	8	4	inside	13

*This simulated data set is used in this document to illustrate the calculation of different measures and summary statistics.*



The frequency with which a species occurs within a sample is typically correlated with abundance. However, when abundance is high and frequency is low, species are considered locally (or sporadically) abundant. Conversely, when frequency is high but abundance is low, species are considered widely distributed but not numerically dominant.

## Density

Density is a measure of abundance per unit area or volume, and is calculated as:

$$\text{Density} = \frac{\text{abundance}}{\text{area}}$$

The unit of area is typically mm<sup>2</sup>, cm<sup>2</sup>, m<sup>2</sup>, 10's m<sup>2</sup>, 100's m<sup>2</sup> or km<sup>2</sup>. For example, the density of snapper inside the marine reserve for the first 500 m<sup>2</sup> transect in Table 1 is:

$$\frac{14}{500} = 0.028 \text{ indiv. m}^{-2}$$

Because density reports the total number of individuals per unit area, it is usually more meaningful than reporting abundances.

## Metrics of diversity

Diversity refers to the representation of a number of different species (or types of species) within a particular ecosystem or sampling unit, and is commonly expressed as either *species richness*, *relative abundance*, *species diversity*, *phylogenetic diversity*, or *functional diversity*.

### Species richness

Species richness is a simple count of species represented in one site (or sample) and does not take into account relative species abundances. For example, in Table 2 the species richness in Transect 1 is five, and species richness in Transect 2 is two.

### Relative abundance

If the abundance of different species in a community is recorded, it is invariably found that some species are rare, and others are abundant. This feature of ecological communities is found independent of the taxonomic group or area investigated.

Relative abundance is calculated as the percent composition of an organism of a particular kind (e.g. species, sex, age class) relative to the total number of organisms in the sample, calculated as:

$$p_i = \frac{N_i}{N}$$

where  $N_i$  is the abundance of the  $i$ -th species in the sample and:

$$N = \sum_{i=1}^S N_i$$



with  $S$  the total number of species in the sample.

For example, if there are 1717 individuals of Sp.1 in a sample of 21,457 individuals, the relative abundance of Sp.1 is:

$$\frac{1,717}{21,457} = 0.08$$

That is, Sp.1 represents 8% of all individuals in the sample.

Commonly, relative abundance is visualised as either rank/abundance plots or  $k$ -abundance curves. The rank/abundance plot (Figure 1A) ranks species in sequence from most to least abundant along the  $x$ -axis. The  $k$ -abundance curve (Figure 1B) shows the cumulative abundance of species (the cumulative abundance of the  $k$ -th most abundant species plus all more abundant species), expressed as a proportion of all individuals in the community.

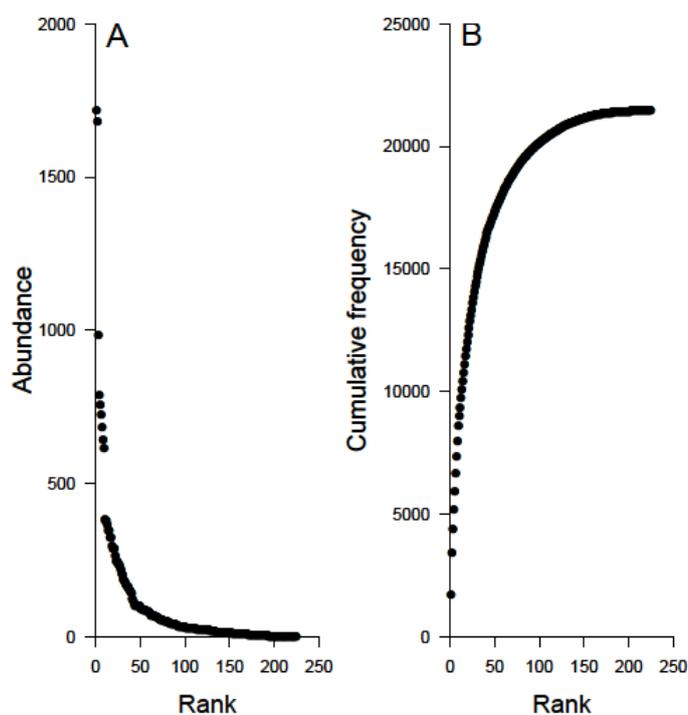


Figure 1. (A) Rank abundance plot and (B)  $k$ -dominance curves for a simulated sample of reef fishes comprising 21,457 individuals from 225 species. The abundance of the highest ranked species in this example is 1717 individuals.

## Species diversity

Species diversity concurrently considers the number of different species that are represented within a given site or sample (i.e. *species richness*) and how similar the *relative abundances* of each species are. Typically, species diversity is quantified as either *alpha diversity* or *beta diversity*.

Diversity is often measured because high diversity is perceived as synonymous with ecosystem health. In general, diverse communities are believed to have increased stability, increased productivity, and resistance to invasion and other disturbances.



## Alpha diversity

Alpha diversity measures the number and proportion in which each species is represented in one site (or sample). A sample will have high alpha diversity when there is a high number of species and their abundances are similar, and low alpha diversity when there are few species, one of which is numerically dominant. Commonly, alpha diversity is described using the Shannon index.

The Shannon index ( $H$ ) reflects the differences in the abundance (numbers of individuals) of each species and can be calculated as:

$$H = - \sum_{i=1}^R p_i \ln p_i$$

where  $p_i$  is the relative abundance of species  $i$  in the total population ( $R$ ) at any one location, and  $\ln$  is the natural log.

For example, in Transect 1 of Table 2 the relative abundance ( $p_i$ ) of each species is 0.2;  $p_i \ln p_i$  is  $-0.32$ , the sum of  $p_i \ln p_i$  is  $-1.61$ , and alpha diversity ( $H$ ) is 1.61. Alpha diversity of Transect 2 is 0.56.

Table 2: Abundances of five species in two simulated 500 m<sup>2</sup> transects;  $p_i$  is the relative abundance of species  $i$  for a given transect, and  $\ln$  is the natural log.

Species	Transect 1			Transect 2		
	Abundance	$p_i$	$p_i \ln p_i$	Abundance	$p_i$	$p_i \ln p_i$
Sp.1	1	0.2	-0.322	3	0.75	-0.216
Sp.2	1	0.2	-0.322	0	0	0
Sp.3	1	0.2	-0.322	0	0	0
Sp.4	1	0.2	-0.322	1	0.25	-0.347
Sp.5	1	0.2	-0.322	0	0	0
<b>Sum</b>	5	1	1.61	4	1	0.56

## Beta diversity

Beta diversity measures the turnover of species between sites (or samples) in terms of the gain or loss of species. Many different measures of beta diversity have been introduced, and the most appropriate measure to use will depend on the particular ecological questions or management issues being addressed. In lieu of presenting numerous indices of beta diversity here, we suggest you consult Anderson et al. (2001) for approaches to determining beta diversity in relation to specific questions, including:

- Species turnover between samples
- Species turnover between samples along an environmental or stress gradient
- The rate of species turnover along spatial, temporal or environmental gradients
- Rates of turnover for different groups of species or taxa



- Variation among communities from a set of samples
- Community variation among *a priori* groups
- Community variation along environmental gradients

## Phylogenetic and functional diversity

Simple diversity indices only consider species identity and assume all different species are equally different (e.g. the difference between two species of triplefin is the same as the difference between a triplefin and a sea anemone). In contrast, indices of phylogenetic and functional diversity account for differences in species attributes (e.g. two species of triplefin are more similar to each other than they are to a sea anemone). Phylogenetic and functional diversities have very similar reasoning, with the major difference being the use of either phylogenetic or functional trait data to describe difference between species.

### Phylogenetic diversity

If two data sets have identical numbers of species and equivalent patterns of species diversity, but differ in the diversity of taxa to which the species belong, the most phylogenetically varied data set is the more diverse. As long as there are well resolved phylogenetic trees for the species of interest (e.g. Figure 2), it is possible to measure phylogenetic diversity.

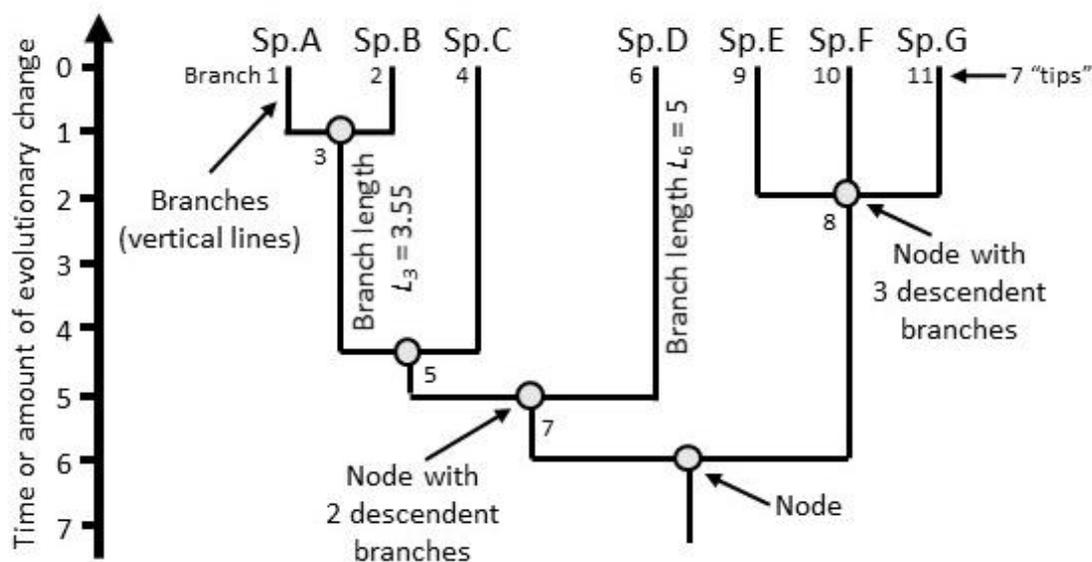


Figure 2. The components of a phylogenetic tree required to calculate metrics of phylogenetic diversity.

There are a number of common metrics used to calculate phylogenetic diversity, with the most appropriate measure depending on the particular ecological questions or management issues being addressed. In lieu of presenting numerous indices of phylogenetic diversity here, we list the most commonly used indices and suggest consulting Vellend et al. (2011), Warwick & Clarke (1995), Winter et al. (2013) and the references cited below for the following phylogenetic indices.



### Phylogenetic distinctiveness of single species

- **Taxonomic distinctiveness (TD).** Topology based; species values are calculated as the reciprocal of the number of nodes between the species and the tree root (Vane-Wright et al. 1991).
- **Evolutionary distinctiveness (ED).** Topology based; species values are calculated as the sum of values per branch (tip to root). The branch value is its length divided by the number of descendant species (Isaac et al. 2007).

### Phylogenetic richness of communities

- **Phylogenetic diversity (PD).** Calculated as the sum of branch lengths between root and tips for a community. PD is mathematically related to species richness and can be used as a complementary measure by identifying added evolutionary information by additional species (Faith 1992).

### Phylogenetic distinctiveness of communities to explore ecological processes

- **Average Taxonomic Distinctiveness (AvTD).** Calculated as the sum of all branch lengths connecting two species averaged across all species representing the mean distance between two randomly chosen species (Warwick & Clarke 1998).
- **Mean nearest taxon distance (MNTD).** Calculated as the mean of the branch lengths connecting each species to its closest relative. MNTD reflects the phylogenetic structure of the tips of the tree (Webb 2000).

### *Functional diversity*

Functional diversity analysis uses species traits data to describe differences between species in a sample.

The positive relationship between ecosystem functioning and species richness is often attributed to the greater number of functional groups found in richer assemblages. Petchey & Gaston (2002, 2006) proposed a method for quantifying functional diversity. It is based on total branch length of a dendrogram, which is constructed from species trait values (e.g. Figure 3). A community of species with different trait values (e.g. filter feeders, deposit feeders, fluid feeders, bulk feeders and ram feeders) will have a higher functional diversity than a community of equal species diversity but where the species are functionally similar (e.g. all filter feeders). One important consideration is that only those traits linked to the ecosystem process of interest are used. Thus, a study focusing on the feeding mode of benthic organisms would exclude traits such as sex ratio that are not related to this function, but traits such as feeding apparatus, feeding method/behaviour and diet should be included. With standard clustering algorithms, a dendrogram is then constructed.



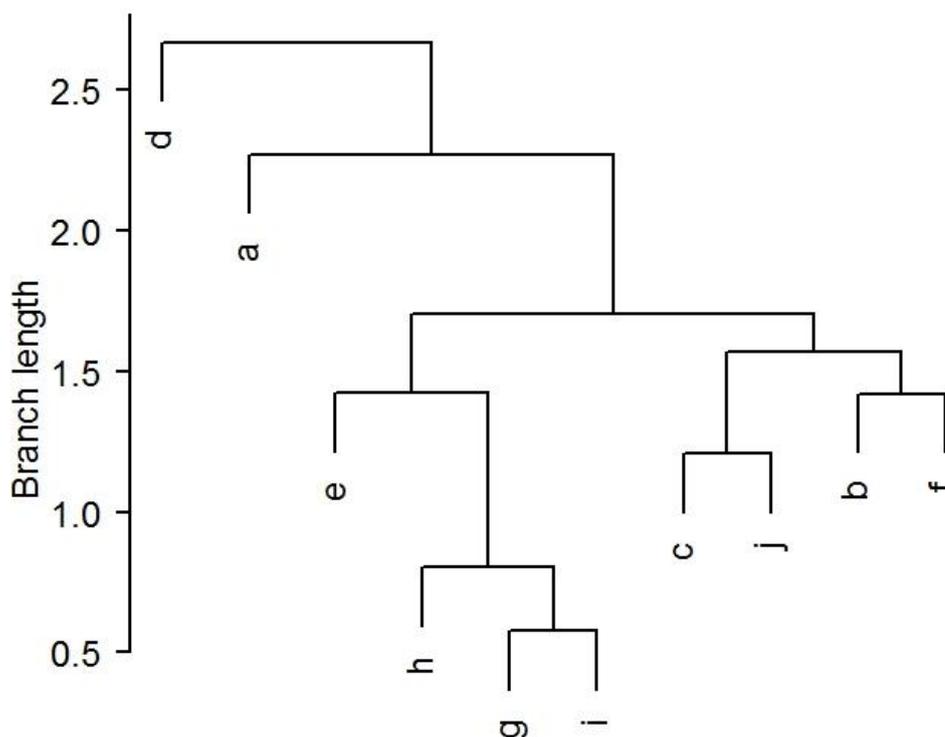


Figure 3. An example functional dendrogram depicting the functional distance between 10 species (species a–j) based on a Euclidean distance matrix for three function traits. Functional diversity is calculated as the sum of the branch lengths within the functional dendrogram, and in this example equals 16.3. See Petchey & Gaston (2002) for more information, and <http://www.thetrophiclink.org/resources/calculating-fd/> for the worked example.

An alternative method to assess functional diversity is to evaluate the relative occurrence of functional traits within a community based on diversity and redundancy of functional traits (see 'Marine: functional trait surveys for benthic organisms'—doccm-2733380<sup>2</sup>).

## Sex ratios and age (or size) structure

Sex ratios and age (or size) structure of populations are often important tools for wildlife management and conservation programmes because a disruption in the proportion of males to females can dramatically affect the reproductive success of a population, and age (or size) structure of a population is a good predictor for population growth, decline or stability.

### Sex ratios

Sex ratios can have important impacts on reproductive success and population dynamics. The sex ratio is the ratio of males to females in a sample (or population) and can be expressed in several ways. For example, in a population of 90 individuals, of which 30 are male, the sex ratio can be expressed as the ratio of males to females (1:2), the proportion of males (0.3) or the percentage of males (30%).

<sup>2</sup> <http://www.doc.govt.nz/Documents/science-and-technical/inventory-monitoring/im-toolbox-marine-functional-trait-surveys-for-benthic-organisms.pdf>



In most species, sex ratio varies according to the age profile of the population, and can be divided into four subdivisions:

- Primary sex ratio at fertilisation
- Secondary sex ratio at birth
- Tertiary sex ratio in reproductively mature adults (also called operational sex ratio)
- Quaternary sex ratio in post-reproductive adults

### Age (or size) structure

Many populations have overlapping generations where individuals of more than one generation coexist producing a distinct age structure that is a good predictor for population growth, decline or stability.

Age structure simply refers to the relative numbers of individuals of each age in a sample (or population) at a given point in time. However, often age is not easily measured, so stage or size classes are used instead. Therefore, age structure can be defined in several ways, including:

- Time (e.g. years) = age class
- Life-history stage (e.g. juvenile, adult) = stage class
- Size classes (e.g. > or < minimum legal size) = size class

Commonly, age structure is visualised using a histogram. For example, Figure 4 illustrates the size structure of pāua (*Haliotis iris*) from within reserve and adjacent non-reserve sites relative to minimum legal size.

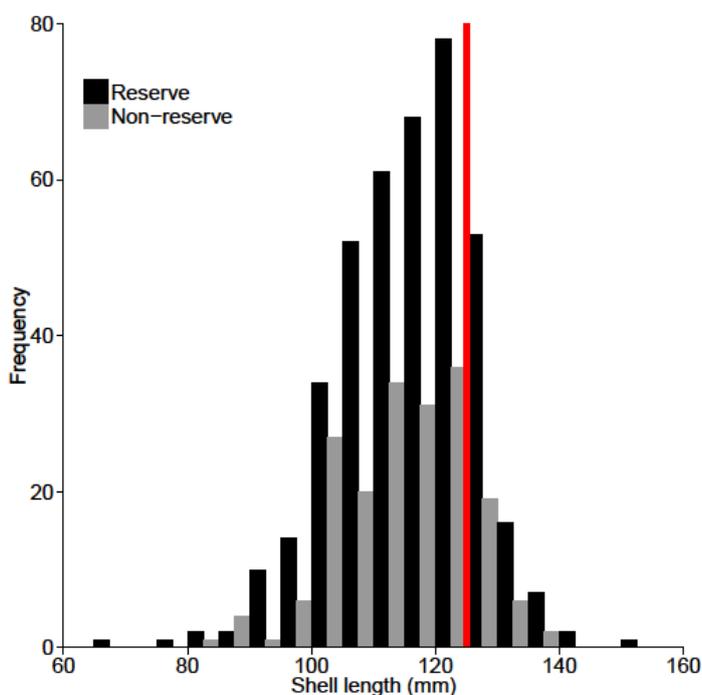


Figure 4. Simulated size structure of pāua (*Haliotis iris*) within a marine reserve ( $n = 402$  individuals) and at adjacent non-reserve sites ( $n = 187$  individuals). The red vertical line represents minimum legal size.



## Biomass

Biomass is a measure of the mass of living biological organisms in a given area at a given time. Biomass can refer to species biomass, which is the mass of one or more species, or to community biomass, which is the mass of all species in the community. Biomass is typically expressed as the average mass per unit area (e.g. per transect), or as the total mass in the community (e.g. within an entire marine reserve).

Biomass can be measured as the natural mass of organisms *in situ* (e.g. blue cod biomass might be calculated as the total wet weight fish would have if they were taken out of the water and weighed) or in terms of dried organic mass (e.g. the biomass of seaweed is often calculated as the total weight it would have if it was taken out of the water and dried).

In the simplest situation, biomass of a single sample is estimated as:

$$B = N \times \bar{w}$$

where  $B$  = estimated biomass (kg),  $N$  = estimated abundance, and  $\bar{w}$  = mean weight of individuals in the sample (kg). In this equation, abundance can be estimated as described earlier, and mean weight is estimated from a random sub-sample representative of the size- or age-groups contained in the abundance sample (Anderson & Neumann 1996).

For example, in Table 2 the abundance of snapper in Transect 1 at Site 1 is 14 individuals. Assuming the mean weight of snapper is 3.2 kg, the mean biomass of snapper in Transect 1 at Site 1 is:

$$\hat{B} = 14 \times 3.2 = 44.8 \text{ kg}$$

For fish, visual estimates of standard length of fishes from surveys can be converted to weight using the following length–weight conversion:

$$\bar{w} = a \times SL^b$$

where  $a$  and  $b$  are constants for allometric growth and  $SL$  is standard length in mm. For some species,  $a$  and  $b$  constants can be region- or sex-specific. For example, different constants are used for lobster (*Jasus edwardsii*) fisheries. The  $a$  and  $b$  constants for allometric growth for many fish species are freely available at [www.fishbase.org](http://www.fishbase.org). Where  $a$  and  $b$  constants for a given species do not exist, parameters from a species belonging to the same genus may be used.

## Catch per unit effort (CPUE)

In conservation biology, catch per unit effort (CPUE) is an indirect measure of species abundance, with changes in CPUE interpreted as representing changes in abundance. Indices of CPUE assume a constant and proportional relationship between CPUE and abundance, which is expressed by Harley et al. (2001) as:

$$CPUE = \frac{C}{E} = q \times N$$



where  $C$  = catch,  $E$  = effort,  $q$  = catchability, and  $N$  = abundance. Catchability can be defined as the fraction of a fish stock collected per unit effort as:

$$q = \frac{C/N}{E}$$

where the term  $C/N$  represents the individual capture probability with total effort ( $E$ ).

Temporal variation in  $q$  must be small relative to the magnitude of change that a monitoring programme aims to detect for CPUE data to effectively describe trends in abundance. However, there are many factors that can cause variation in  $q$  (e.g. abundance, variation in sampling gear, habitat, season) and their combined effects can be difficult to isolate, undermining the utility of CPUE data to describe abundance. Thus, understanding how  $q$  varies through time and across habitat conditions for each species is essential for interpreting trends in CPUE.

One approach to estimate  $q$  is to tag individuals at a location, and then conduct standardised sampling to obtain estimates of catchability. However, this is not always logistically or financially feasible, and often  $q$  is assumed to be constant, especially in studies where sampling sites, sampling gear and the season in which sampling occurs is standardised. In these instances, the goal is to estimate relative abundance, so catchability, strictly defined, is not a factor, and CPUE is calculated as:

$$CPUE = \frac{C}{E}$$

Results are typically expressed as catch per effort, where effort comprises a sampling unit and a time interval. For example, CPUE from a lobster potting survey may be expressed as kg of lobsters per pot per day.

CPUE data tend not to be normally distributed (i.e. data are not symmetrically distributed around the mean), so care should be taken when applying parametric tests (Hubert & Fabrizio 2007).



## Calculating summary ecological statistics from marine monitoring data

### Measures of location and spread

When we report on monitoring data, we typically summarise our data using summary statistics, which can be divided into *measures of location* and *measures of spread*. *Measures of location* illustrate where the majority of data can be found and include means, medians and modes. Conversely, *measures of spread* describe how variable the data are around the measure of location, and include sample standard deviation, variance and standard errors.

#### Measures of location

##### *The mean*

One of the most common ways to summarise data is to use the mean of the observations (also known as the arithmetic mean or the average). The mean is calculated as the sum of the observations ( $Y_i$ ) divided by the number of observations ( $n$ ) and is denoted by  $\bar{Y}$ :

$$\bar{Y} = \frac{\sum_{i=1}^n Y_i}{n}$$

For the data presented in Table 1, the mean numbers of snapper per transect are  $\bar{Y}_1 = 12.18$  for transects outside the reserve and  $\bar{Y}_2 = 16.25$  for transects inside the reserve.

##### *The median and mode*

Two other measures of location—the median and mode—are often used to summarise data. The median is defined as the value in a set of observations that has an equal number of observations above and below it. For an odd number of observations, the median is the central observation. For an even number of observations, the median is the midway point between the  $(n/2)$  and  $[(n/2)+1]$  observation.

The mode is the value of observations that occurs most frequently in the sample. The mode can be read easily off a histogram of data as the peak.

For the abundance of snapper inside the marine reserve as presented in Table 1, the median is 16.5 and the mode is 17.



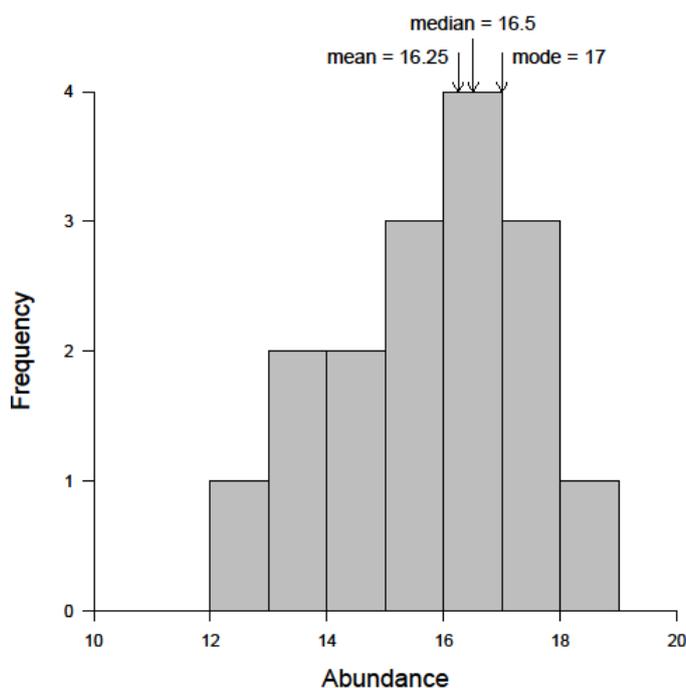


Figure 5. Histogram of the simulated abundance of snapper per 500 m<sup>2</sup> transect inside the marine reserve as presented in Table 1 ( $n = 16$  transects), illustrating the mean, median and mode.

### *When to use each measure of location*

The mean is the most commonly used measure of location because it can be easily used to test hypotheses. The median and mode better describe the location of the data when distribution of observations cannot be fit to a normal probability distribution, or when there are extreme outliers. In symmetric distributions, the mean, median and mode are all equal. In asymmetric distributions, the mean occurs towards the largest tail of the distribution, the mode occurs towards the heaviest part of the distribution and the median occurs between the two, as in Figure 5.

### Measures of spread

Because there is variation in nature, and because there is a limit to the precision with which we can make measurements, we must also quantify the spread or variability of our observations.

### *The variance and standard deviation*

The variance of the mean is a measure of how far the observations in our sample differ from the mean. We can calculate the variance ( $s^2$ ) of the mean as:

$$s^2 = \frac{1}{n-1} \sum (Y_i - \bar{Y})^2$$

For the abundance of snapper inside the marine reserve as presented in Table 1, the variance is 2.87.



The standard deviation ( $s$ ) of a sample of observations is defined as the square root of the variance. The square root transformation ensures that the units of standard deviation are the same as the units of the mean:

$$s = \sqrt{\frac{1}{n-1} \sum (Y_i - \bar{Y})^2}$$

For the abundance of snapper inside the marine reserve as presented in Table 1, the standard deviation is 1.69.

### *The standard error of the mean*

Another measure of spread is the standard error of the mean ( $se_{\bar{Y}}$ ), which is calculated by dividing the sample standard deviation by the square root of the sample size:

$$se_{\bar{Y}} = \frac{s}{\sqrt{n}}$$

For the abundance of snapper inside the marine reserve as presented in Table 1, the standard error of the mean is 0.42.

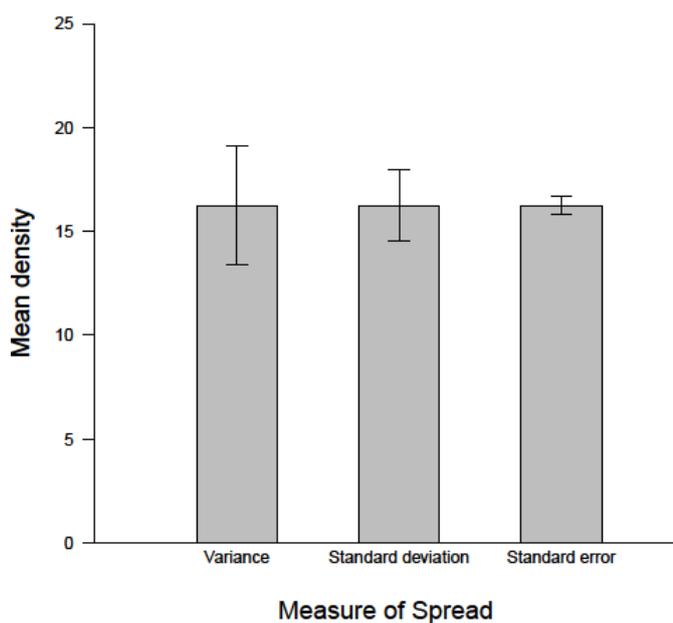


Figure 6. Bar chart showing the mean for the abundance of snapper inside a marine reserve as presented in Table 1 ( $n = 16$  transects) along with error bars indicating the sample variance (left bar), standard deviation (middle bar) and standard error (right bar).

*Note: Although the standard deviation is always bigger than the standard error, variance can be smaller than the standard deviation when the standard deviation is between, but not equal to, 0 and 1.*

*Note: Figure captions should always provide the sample sizes and indicate clearly what has been used to construct the error bars.*

### *When to use the standard deviation and standard error of the mean*

The standard deviation is a descriptive statistic, and the standard error of the mean ( $se$ ) is an inferential statistic.

The standard deviation should be used to describe the variation observed in a sample. For example,  $\bar{Y} = 100$ ,  $s = 25$  suggests a much more variable sample than does  $\bar{Y} = 100$ ,  $s = 5$ .



Conversely, the standard error of the mean describes the standard deviation of the population from which the sample is drawn, adjusted by the amount of certainty we gain about our mean estimate from the sample size. For example, in a comparison of males ( $m$ ) and females ( $f$ ),  $\bar{Y}_m = 100$ ,  $se = 2$  and  $\bar{Y}_f = 120$ ,  $se = 1.5$  would allow an inference that the population mean value  $Y$  is greater among females than among males. Unlike the standard deviation, the standard error decreases with increasing sample size.

As long as you report the sample size in your text, figure or figure caption (e.g. Figure 6), readers can compute the standard error of the mean from the sample standard deviation or vice versa.

## Quantiles

Another way to illustrate the spread of a distribution is to report its quantiles. In presentations of statistical data, we most commonly report upper and lower quantiles (the values for the 25th and 75th percentiles), and the upper and lower deciles (the values for the 10th and 90th percentiles). For example, an observation being reported as being in the 90th decile of a sample of observations means that 90% of the scores are lower than the one being reported and 10% are higher.

Unlike the standard error of the mean and the standard deviation, the values of quantiles do not depend on the values of the mean. When distributions are asymmetric or contain outliers (extreme data points not characteristic of the distribution they were drawn from), box plots of quantiles can portray the distribution of the data more accurately than conventional plots of means and standard errors (e.g. Figure 7).

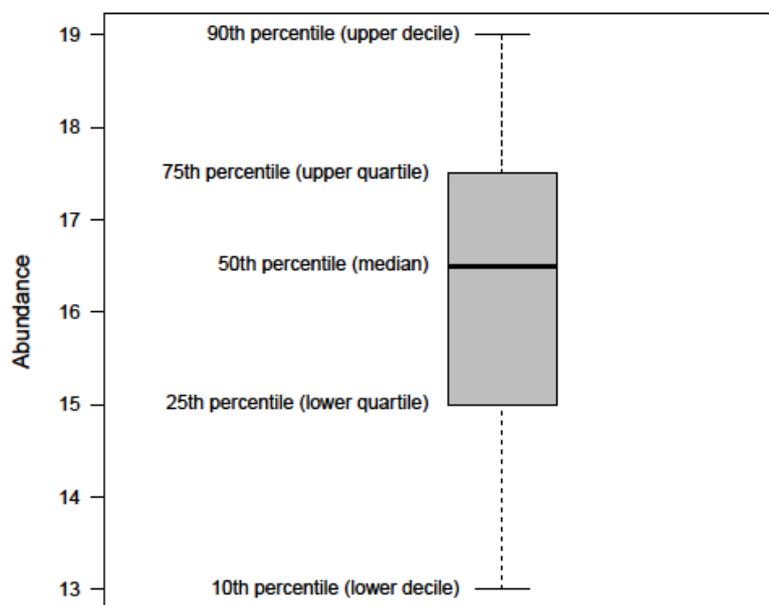


Figure 7. Boxplot illustrating quantiles of data for the abundance of snapper inside a marine reserve as presented in Table 1 ( $n = 16$  transects). The line indicates the 50th percentile (median), and the box encompasses 50% of the data, from the 25th to the 75th percentile. The dashed vertical lines extend from the 10th to the 90th percentile.



## Confidence intervals

The standard error can be used to construct a confidence interval (CI) around the mean. For a normally distributed random variable, approximately 67% of observations occur within  $\pm 1$  standard error of the mean, and approximately 96% of observations occur within  $\pm 2$  standard errors of the mean. Therefore, we can use the standard error to create a 95% CI:

$$P(\bar{Y} - 1.96se_{\bar{Y}} \leq \mu \leq \bar{Y} + 1.96se_{\bar{Y}}) \approx 0.95$$

The CI represents the probability that the true mean of the population ( $\mu$ ) from which samples were drawn falls within the CI. That is, 95% of the time a CI calculated this way will contain the true mean of the population. Therefore, if you carried out your sampling 100 times, and created 100 CIs, approximately 95 of them would contain the true mean of the population, and 5 would not.

For the abundance of snapper inside the marine reserve as presented in Table 1 and Figure 9, the CIs around the mean are  $15.42 \leq 16.25 \leq 17.08$ . Therefore, we are 95% confident that the true mean falls within these bounds. Or, if we sampled the same marine reserve 100 times, 95 of the calculated CIs would contain the true population mean abundance (Figure 8).

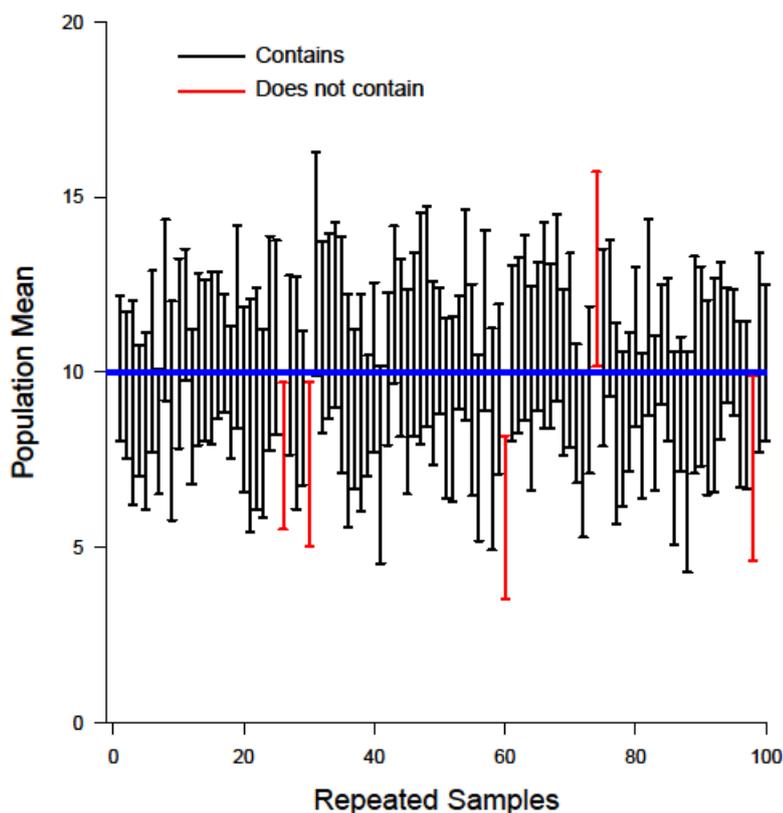


Figure 8. Illustration of the concept of CIs and resampling. The blue line is the true population mean. If resampled 100 times, 95 of the calculated CIs would contain the true population mean (black CIs) while 5 CIs would not contain the true population mean (red CIs).



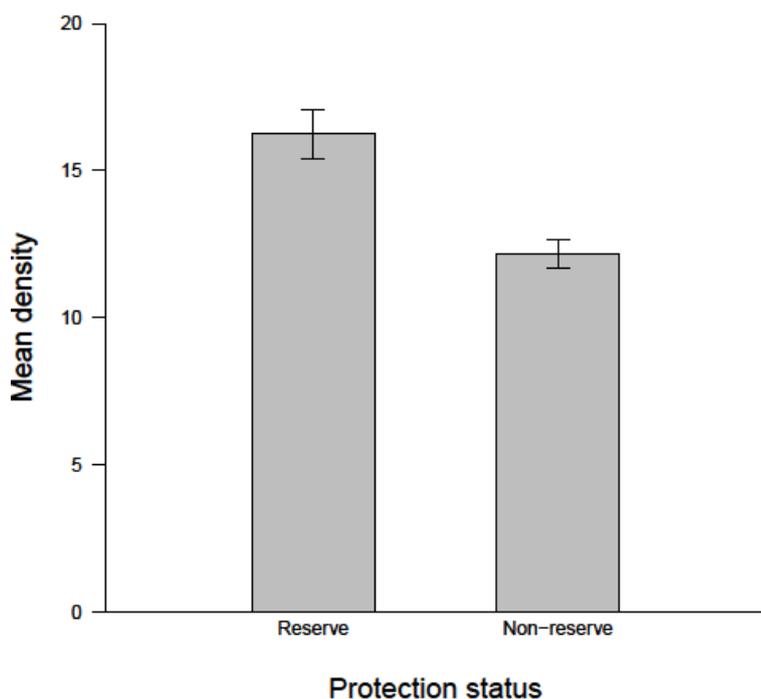


Figure 9. Mean ( $\pm$  95% CIs) density of snapper on 500 m<sup>2</sup> transects from reserve and non-reserve areas ( $n = 16$  transects per area).

## Response ratios and effect sizes

Often, we are most interested in the magnitude of difference between populations or treatments, which can be expressed as *response ratios* and *effect sizes*. These are easy to calculate, readily understood and are valuable for quantifying the effectiveness of a particular intervention, relative to some comparison—for example, the effectiveness of marine protection in increasing the abundance of snapper relative to non-protected areas.

By placing the emphasis on the most important aspect of an intervention—the size of the effect—*response ratios* and *effect sizes* are important tools in reporting and interpreting effectiveness, and facilitate the interpretation of the substantive, as opposed to the statistical, significance of a result.

### Response ratios

The response ratio, calculated as the ratio of some measured quantity in two groups, is commonly used as a measure of effect because it quantifies proportional change. Examples of response ratios include relative yield, relative density and relative competitive intensity. The response ratio ( $R$ ) is simply calculated as:

$$R = \frac{\bar{Y}_1}{\bar{Y}_2}$$

where  $\bar{Y}_1$  is the mean of group 1 and  $\bar{Y}_2$  is the mean of group 2.

For response ratios, computations of standard error, standard deviation and CIs are conducted on a log scale to maintain symmetry in the analysis (see Hedges et al. 1999 for an explanation). These are then converted back to the original scale. The log response ratio is calculated as:

$$\ln(R) = \ln(\bar{Y}_1) - \ln(\bar{Y}_2)$$

where  $\ln$  is the natural log. If the variances in the two groups are approximately equal, then the variance of  $\ln(R)$  can be calculated using the pooled within-groups standard deviation (see Hedges et al. 1999 for testing the statistical significance of variance components):

$$s_{\ln(R)}^2 = s_{\text{pooled}}^2 \left( \frac{1}{n_1(\bar{Y}_1)^2} + \frac{1}{n_2(\bar{Y}_2)^2} \right)$$

where  $s_{\text{pooled}}$  is the pooled standard deviation. An approximate 95% CI for the log response ratio is given by:

$$P \left( \ln(R) - 1.96 \sqrt{s_{\ln(R)}^2} \leq \mu \leq \ln(R) + 1.96 \sqrt{s_{\ln(R)}^2} \right) \approx 0.95$$

We then convert computations of the log response ratio and log upper or lower CIs back to the original scale using:

$$\exp(\delta)$$

where  $\delta$  is either the log response ratio or the log upper or lower CIs.

For example, the log response ratio for the mean density of snapper inside the marine reserve relative to outside the marine reserve as presented in Table 1 is:

$$\ln(R) = \ln(16.25) - \ln(10.19) = 0.47$$

with an associated variance of:

$$s_{\ln(R)}^2 = s_{\text{pooled}}^2 \left( \frac{1}{n_1(\bar{Y}_1)^2} + \frac{1}{n_2(\bar{Y}_2)^2} \right) = 1.905^2 \times \left( \frac{1}{16 \times 16.25^2} + \frac{1}{16 \times 10.188^2} \right) = 0.003$$

and a 95% confidence interval of:

$$P(0.47 - 1.96\sqrt{0.003} \leq \mu \leq 0.47 + 1.96\sqrt{0.003}) = P(0.36 \leq \mu \leq 0.58) \approx 0.95$$

After conversion back to the original scale, the response ratio is:

$$R = \exp(0.47) = 1.60$$

with a 95% upper confidence interval (UCI) and lower confidence interval (LCI) of:

$$LCI = \exp(0.36) = 1.43$$

$$UCI = \exp(0.58) = 1.78$$

The density of snapper inside the marine reserve is therefore  $1.43 \leq 1.60 \leq 1.78$  times greater than the density of snapper outside the marine reserve (Figure 9).



## Effect sizes

Like response ratios, the effect size is a quantitative measure of the difference between populations or treatments, with larger absolute values indicating a stronger effect.

The effect size ( $d$ ) is the standardised mean difference between two groups, calculated as:

$$d = \frac{\bar{Y}_1 - \bar{Y}_2}{s}$$

where  $\bar{Y}_1$  is the mean of group 1,  $\bar{Y}_2$  is the mean of group 2 and  $s$  is the standard deviation. The standard deviation is estimated from group 2 or from a 'pooled' value from both groups. Group 1 is the group of interest (e.g. marine reserve sites) and group 2 is a 'control' group to which group 1 is being compared (e.g. non-reserve sites).

One feature of an effect size is that it can be directly converted into statements about the overlap between the two samples in terms of a comparison of percentiles. For example, if we were to compare the abundance of snapper inside a marine reserve with the abundance of snapper outside a marine reserve as presented in Table 1, we would calculate an effect size as:

$$d = \frac{16.25 - 12.19}{1.91} = 2.13$$

With a calculated effect size of 2.1, the mean abundance of snapper within the marine reserve would be 2.1 standard deviations greater than the mean abundance of snapper outside the marine reserve.

Table 3 shows conversions of effect size into percentiles. Hence, an effect size of 2.1 converts to 98%, meaning the mean abundance in the marine reserve exceeds the densities of 98% of the samples from the non-reserve area.

Table 3. Conversions of effect sizes into percentiles.

Effect size	Percentage of group 2 samples below average sample in group 1	Effect size	Percentage of group 2 samples below average sample in group 1
0.0	50%	0.9	82%
0.1	54%	1.0	84%
0.2	58%	1.2	88%
0.3	62%	1.4	92%
0.4	66%	1.6	95%
0.5	69%	1.8	96%
0.6	73%	2.0	98%
0.7	76%	2.5	99%
0.8	79%	3.0	99.9%

We can also calculate 95% CIs for effect sizes. If the CI includes zero, there is no statistically significant difference between the two means. Alternatively, if the CI does not include zero, then



there is a statistical difference between the means at the 5% significance level (see '[Confidence intervals](#)' for an explanation).

The 95% CIs for an effect size can be calculated as:

$$P(d - 1.96s_d \leq \mu \leq d + 1.96s_d) = 0.95$$

where  $d$  is the effect size estimate and  $s_d$  is the standard deviation of the effect size.

As per Hedges & Olkin (1985, p. 86), the standard deviation is calculated as:

$$s_d = \sqrt{\frac{n_{g1} + n_{g2}}{n_{g1} \times n_{g2}} + \frac{d^2}{2(n_{g1} + n_{g2})}}$$

where  $n_{g1}$  and  $n_{g2}$  are the numbers of samples in groups 1 and 2, respectively.

For example, the effect size of 2.13 for the snapper example above has an associated standard deviation of:

$$s_d = \sqrt{\frac{16 + 16}{16 \times 16} + \frac{2.13^2}{2(16 + 16)}} = 0.44$$

and a standard error of

$$se_d = \frac{s_d}{\sqrt{n}} = \frac{0.44}{\sqrt{32}} = 0.08$$

The CI is calculated by adding and subtracting 1.96 times the standard error to/from the mean ( $2.13 \pm 1.96 \times 0.08$ ), giving a 95% CI of  $2.13 \pm 0.22$ . Because the CI does not include 0, there is a statistically significant difference between the mean abundance of snapper inside and outside the marine reserve, with mean abundance inside the marine reserve exceeding the densities of 98% of the samples from the non-reserve area.

## Presenting data

### Presenting data in tables

Tables are the format in which most numerical data are initially stored and analysed and are likely to be what you use to organise data collected during monitoring and research. However, when writing up your work you will have to make a decision about whether a table is the best way of presenting the data, or if the data would be easier to understand presented as a graph or chart.

#### When to use tables

Tables are an effective way of presenting data when:



- You want to show how a single category of information varies when measured at different points (in time or space). For example, a table would be an appropriate way of showing how annual visitor numbers vary between different marine reserves (different points in space), although note that a bar chart may be a more easily interpretable way of presenting this data.
- The data set contains relatively few numbers. This is because it is very hard for a reader to assimilate and interpret lots of numbers presented in a table.
- The precise value is crucial to your argument and a graph would not convey the same level of precision—for example, when it is important that the reader knows that the result was 9.27 and not 9.22.

## Table design

Since tables consist of rows and columns of information, it is important to consider how the data are arranged between the two. Most people find it easier to identify patterns in numerical data by reading down a column rather than across a row. This means that you should plan your row and column categories to ensure that the patterns you wish to highlight are revealed in the columns. It is also easier to interpret the data if they are arranged according to their magnitude so there is numerical progression down the columns, although this may not always be possible.

If there are several columns or categories of information, a table can appear complex and become hard to read. It also becomes more difficult to list the data by magnitude since the order that applies to one column may not be the same for others. In such cases you need to decide which column contains the most important trend, and this should be used to structure the table. If the columns are equally important, it is often better to include two or more simple tables rather than using a single more complex one.

To ensure that tables are clear and easy to interpret, the following design issues need to be considered:

Do	Don't
Ensure that tables have an accompanying title above them that describes the content and explains any abbreviations or symbols so that a reader can understand the content without needing to consult the accompanying text.	Do not rely on the text of your manuscript to explain the contents of your table.
Use the text to focus on the significance or key points of your tables.	Do not repeat the contents of tables within the text.
Present values and details consistently in tables and text (e.g. abbreviations, group names, treatment names).	Do not use different values or abbreviations in text and tables.
Present any species name or acronyms in full where they first appear in each table.	Do not use acronyms or abbreviations in tables without giving them in full at their first use.
Ensure all words or symbols appearing in your tables are large enough for easy reading.	Do not use illegible (small) font sizes, symbols or line weights that make reading difficult.



Do	Don't
Use serif-free fonts because they have better legibility.	Do not use serif fonts because they are harder to read.
Present numbers in their simplest form—this may mean rounding up values to avoid the use of excessive decimal places (e.g. $P < 0.001$ ).	Do not use excessive decimal places when reporting numbers (e.g. $P = 0.000027$ ).
Ensure that different material is presented in tables and figures.	Do not repeat the same material in figures and tables.

## Presenting data in figures

Data visualisation is the presentation of data in a pictorial or graphical format. The visual presentation of data makes it easier to grasp difficult concepts or identify patterns. Choosing the correct type of figure to communicate the key points of your data is essential. For example, line graphs can be used to show continuous data over time; bar charts can be used to compare groups; scatter plots can be used to show how change in one variable is related to change in another variable; and histograms can show data within intervals.

Well-prepared figures help you present complex data in a concise and visually appealing manner, as well as enabling readers to get a quick overview of your findings. Therefore, it is essential to ensure that figures are flawless, effective, and attractive. When preparing figures to visualise your data, follow these general guidelines:

Do	Don't
Ensure that figures have an accompanying caption and are self-explanatory and can be understood independent of text.	Do not rely on the text of your manuscript to explain the contents of your figures.
Use the text to focus on the significance or key points of your figures.	Do not repeat the contents of figures within the text.
Use fonts consistently in all figures.	Do not use a mixture of fonts in figures.
Present values and details consistently in figures and text (e.g. abbreviations, group names, treatment names).	Do not use different values or abbreviations in text and figures.
Present any species names or acronyms in full in the caption of each figure.	Do not use acronyms or abbreviations in figures without giving them in full in the figure caption.
Titles or labels not necessary for understanding the figure should be removed and explained in the caption.	Do not include titles or labels in figures that are not necessary for interpreting the figure.
Ensure all words or symbols appearing in your figures are large enough for easy reading.	Do not use illegible (small) font sizes, symbols or line weights that make reading difficult.
Use solid symbols for plotting data if possible (unless data overlap or there are multiple symbols).	Do not use small dotted lines, thin lines, multiple levels of grey shading, or stippling.
Ensure that your data fill your entire plot.	Do not extend scales or axes beyond the range of the data plotted.



Do	Don't
Ensure that differentiation between symbols and bars in figures is sufficient so they appear distinct from one another.	Do not use symbols or shading schemes that do not allow different treatments or groupings to be differentiated from one another.
Ensure legends are kept as simple as possible and are positioned so they do not enlarge the figure, or obscure data within the figure.	Do not use needlessly complicated legends, or crowd data in the figure with the legend.
Use fewer numbers on the axes to achieve an uncluttered look.	Do not crowd the x- and y-axis with unnecessary numbers.
Use a slightly smaller font size on the x- and y-axis than for the axis labels.	Do not use a larger size font for the x- and y-axis than for the axis labels.
Make figures as simple as possible.	Do not use gridlines and boxes unless necessary.
Use serif-free fonts because they have better legibility.	Do not use serif fonts because they are harder to read.
Ensure that different material is presented in tables and figures.	Do not repeat the same material in figures and tables.

A summary of commonly used types of figures for reporting ecological data are included below, although see Kelly et al. 2005<sup>3</sup> for a more detailed discussion of visualisation methods for ecological data.

- **Bar charts** are used to display and compare the number, frequency or other measure (e.g. mean) for different discrete categories of data (e.g. Figure 10A). They are useful for displaying data that are classified into nominal or ordinal categories. Nominal data are categorised according to descriptive or qualitative information such as functional group, sex or protection status. Ordinal data are similar but the different categories can also be ranked—for example, the number of aggressive displays by territorial fish grouped according to the intensity of the display (low, medium and high).
- **Histograms** are bar charts that represent a frequency distribution of interval data. The height of each of the bars equals the frequency of the observation identified on the x-axis. Histograms are a useful tool for visualisation of the distribution of values of a variable (e.g. Figure 10B). They are one of the first graphs looked at to see whether a variable follows a normal distribution or has a skewed distribution. They are also useful for comparing the distributions of two independent variables. For example, Figure 4 shows the size distributions of pāua within a marine reserve and at adjacent non-reserve sites.
- **Boxplots** are used to illustrate the distribution of variables within a data set (e.g. Figure 10C). They are a standardised way of displaying the distribution of data based on the median, upper and lower quantiles, upper and lower deciles, and any outliers.
- **Line graphs** are usually used to show time series data—that is, how one or more variables vary over a continuous period of time (e.g. Figure 10D). An example of the type of data that can be presented using line graphs is the abundance of lobsters within a marine reserve from annual surveys. Line graphs are particularly useful for identifying patterns and trends in the data such as seasonal effects, large changes and turning points. As well as showing

<sup>3</sup> <http://www.doc.govt.nz/Documents/science-and-technical/docts32entire.pdf>



time series data, line graphs can also be appropriate for displaying data that are measured over other continuous variables, such as distance. For example, a line graph could be used to show how the density of lobster varies with increasing distance from the centre of a marine reserve into adjacent non-reserve areas. However, it is important to consider whether the data have been collected at sufficiently regular intervals so that estimates made for a point lying halfway along the line between two successive measurements would be reasonable. In a line graph, the x-axis represents the continuous variable (e.g. year or distance from the initial measurement) whilst the y-axis has a scale and indicates the measurement. Several data series can be plotted on the same line chart, and this is particularly useful for analysing and comparing the trends in different data sets (e.g. for comparing estimates of abundance from reserve and non-reserve areas).

- **Scatterplots** are two-dimensional plots used to illustrate bivariate data. The predictor (independent) variable is usually placed on the x-axis, and the response (dependent) variable is usually placed on the y-axis (e.g. Figure 10E). Each point represents the measurements of these two variables for each observation.

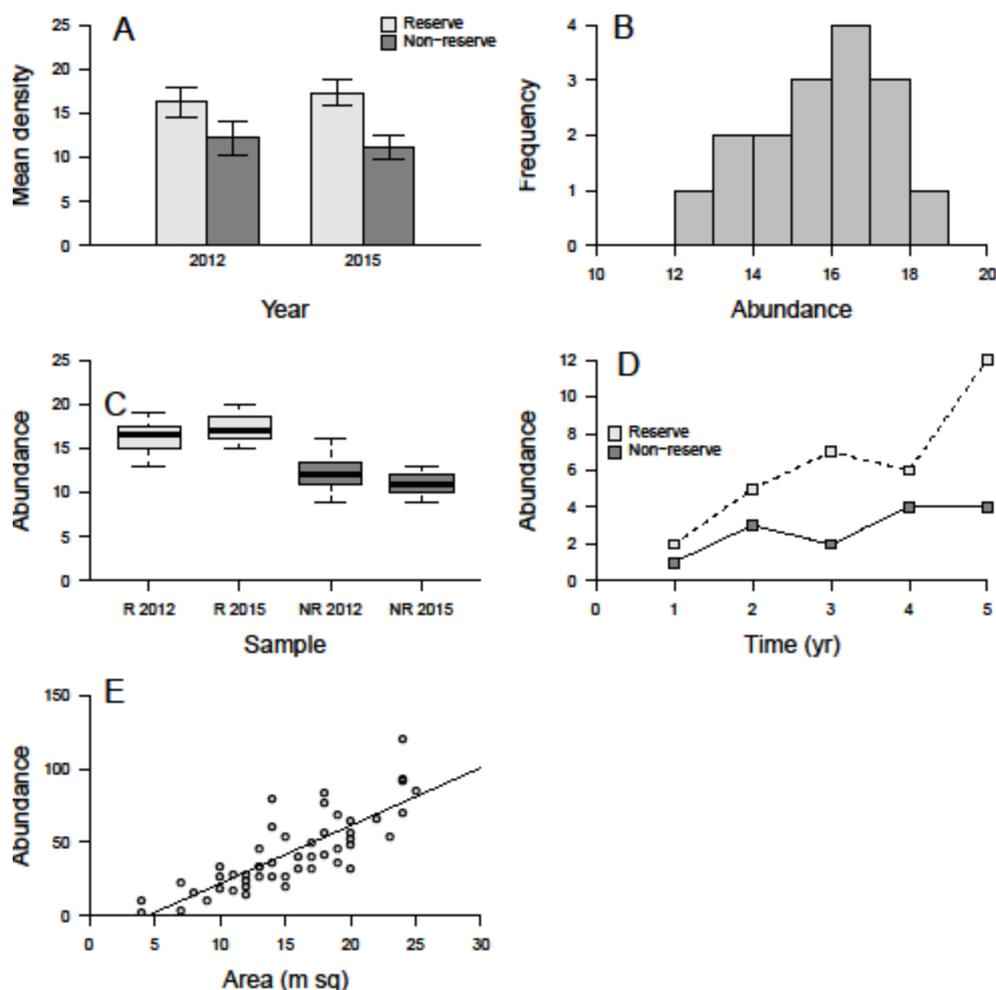


Figure 10. Examples of commonly used types of figures for reporting ecological data: (A) bar chart, (B) histogram, (C) boxplot, (D) line graph, and (E) scatterplot.



## What statistical analysis should I use?

The type of analysis most applicable to the data will largely depend on the objectives of the study, and whether additional supporting information (such as physical conditions or habitat variables) has been recorded or is available.

Although there are many different types of statistical analysis, choosing the correct analytical approach for your situation can be a complicated task. Table 4 provides general guidelines for choosing a statistical analysis and differentiates between a number of common analyses based on the number and nature of dependent variables (sometimes referred to as outcome variables), and the nature of your independent variables (sometimes referred to as predictors). For example, if we were to measure the response of snapper abundance to protection status, snapper abundance would be the dependent variable and protection status would be the independent variable with two levels (protected versus non-protected).

Table 4. Choosing the appropriate statistical test (adapted from Institute for Digital Research and Education 2017).

Number of dependent variable(s)	Number of independent variable(s) (IVs)	Nature of dependent variable(s)*	Statistical analysis
1	0 IV (1 population)	interval & normal	one-sample <i>t</i> -test
		ordinal or interval	one-sample median
		categorical (2 categories)	binomial test
	1 IV with 2 levels	categorical	chi-squared goodness-of-fit
		interval & normal	two independent sample <i>t</i> -test
		ordinal or interval	Wilcoxon–Mann–Whitney test
	1 IV with 2 or more levels	categorical	chi-squared test Fisher's exact test
		interval & normal	one-way analysis of variance (ANOVA)
		ordinal or interval	Kruskal–Wallis test
	1 IV with 2 levels	categorical	chi-squared test
		interval & normal	paired <i>t</i> -test
		ordinal or interval	Wilcoxon signed ranks test
1 IV with 2 or more levels	categorical	McNemar test	
	interval & normal	one-way repeated measures ANOVA	
	ordinal or interval	Friedman test	
2 or more IVs with 2 or more levels	categorical	repeated measures logistic regression	
	interval & normal	factorial ANOVA	
	ordinal or interval	ordered logistic regression	
		categorical	factorial logistic regression



Number of dependent variable(s)	Number of independent variable(s) (IVs)	Nature of dependent variable(s)*	Statistical analysis	
	1 interval IV	interval & normal	correlation	
		interval & normal	simple linear regression	
		ordinal or interval	non-parametric correlation	
		categorical	simple logistic regression	
	1 or more interval IVs and/or 1 or more categorical IVs	interval & normal	multiple regression	
			analysis of covariance	
		categorical	multiple logistic regression	
			discriminant analysis	
	2+	1 IV with 2 or more levels	interval & normal	one-way multivariate analysis of variance (MANOVA)
		2+ IVs	interval & normal	multivariate multiple linear regression
0 IV		interval & normal	factor analysis	
2 sets of 2+	0 IV	interval & normal	canonical correlation	

\* A **categorical** variable is one that has two or more categories, but there is no intrinsic ordering to the categories. For example, sex of lobsters is a categorical variable (male or female). A purely categorical variable is one that simply allows you to assign categories but they cannot be clearly ordered. If the categories have a clear ordering, then that variable would be an **ordinal** variable. An example would be aggressive displays by territorial reef fish, with three categories (low, medium and high). In addition to being able to classify aggression, you can order the categories. A high score means more aggression than a medium score, and that is more than a low score, but the difference between low and medium may not be the same as that between medium and high. An **interval** variable is similar to an ordinal variable, except that the intervals between the values of the interval variable are equally spaced. For example, suppose you have a variable such as shell length of pāua that is measured in millimetres; the difference between a shell length of 100 mm and 90 mm is the same as the difference between 90 mm and 80 mm. For some analyses, the assumption is that the distribution of the sample means (*t*-tests and ANOVAs) or the residuals (regression analysis) conform to a **normal** distribution.

The statistical analyses presented in Table 4 should not be construed as hard and fast rules. Usually your data could be analysed in multiple ways, each of which could yield legitimate answers. Useful texts that can help you choose and conduct statistical analyses most appropriate for your data and then interpret the resulting outputs include:

- Anderson, R.O.; Neumann, R.M. 1996: Length, weight, and associated structural indices. Pp. 447–482 in Murphy, B.R.; Willis, D.W. (Eds): Fisheries techniques. 2nd edition. American Fisheries Society, Bethesda, Maryland.
- Clark, M.; Randal, J.A. 2004: A First Course in Applied Statistics: With Applications in Biology, Business and the Social Sciences. Pearson Education New Zealand, Auckland.
- Crawley, M.J. 2012: The R book. John Wiley & Sons, West Sussex.
- Gotelli, N.J.; Ellison, A.M. 2004: A primer of ecological studies. Sinauer, Sunderland, Massachusetts.
- Quinn, G.P.; Keough, M.J. 2002: Experimental design and data analysis for biologists. Cambridge University Press, Cambridge.



- Sokal, R.R.; Rohlf, F.J. 1995: Biometry: the principals and practice of statistics in biological research. WH Freeman and Company, New York.
- Stevens, M.H. 2009: A Primer of Ecology with R. Springer Science & Business Media, New York.

## How should I present statistical results as written text?

Sampling designs, experimental designs, data-collection protocols, precision of measurements, sampling units, experimental units, and sample sizes must be clearly described. Reported statistical information usually includes the sample size, measure of location (e.g. mean) and some measure of spread (standard deviation (SD) or standard errors (SEs)) or specified confidence intervals (CIs), although this may not be necessary or possible in all instances, especially for unusual statistics.

When reporting the results of statistical tests, you should follow these general guidelines:

- If a statistics program or program package was used, a complete citation (including version number) should be given. If necessary, the author should indicate which procedure within a package was used and which method within a procedure was chosen (e.g. 'All statistical analyses were conducted in R 2.7.0 (R Development Core Team 2007). We used the package segmented 0.2-7.1 (Muggeo 2004) for piecewise regressions.')
- The specific statistical procedure must always be stated (e.g. analysis of variance (ANOVA), two-sample *t*-test, generalised linear mixed model (GLMM)).
- Report the descriptive statistics, such as means and standard deviations (e.g. mean = 15.2, *SD* = 6.32).
- When presenting results such as  $a \pm b$ , always indicate if *b* is a standard deviation, standard error or CI.
- If a CI is to be used, give the lower and upper limits, as these can be asymmetrical around the estimate.
- The methods section should indicate the CI used (e.g. 90%, 95% or 99%).
- If conclusions are based on an ANOVA or a regression analysis, information sufficient to permit the construction of the full analysis of variance table (at least degrees of freedom, the structure of *F*-ratios, and *P*-values) must be presented or be clearly implicit (e.g. ANOVA:  $F_{1,24} = 1.22$ ,  $P > 0.001$ ).
- When reporting a significant difference between two conditions, indicate the direction of the difference using (for example) response ratios or effect size and associated CIs (e.g. snapper density within the reserve was 1.595 times greater than density outside the reserve (95% CI = 1.432–1.777)).
- Define all symbols, abbreviations, and acronyms the first time they are used. Use leading zeroes with all numbers < 1, including probability values (e.g.  $P < 0.001$ ).
- Units of measure should conform to the International System of Units (SI).<sup>4</sup>

<sup>4</sup> <https://www.nist.gov/sites/default/files/documents/pml/div684/fcdc/sp330-2.pdf>



- Test statistics and *P*-values should be rounded to three decimal places. All statistical symbols that are not Greek letters should be italicised (*M*, *SD*, *N*, *t*, *P*, etc.).

## Examples

The following examples illustrate how to report statistics in the text. Please pay attention to issues of italics and spacing.

**Mean and standard deviation** are most clearly presented in parentheses. For example:

The average age of marine reserves used in the analysis was 7.22 years (*SD* = 3.45).

**Percentages** are also most clearly displayed in parentheses with no decimal places. For example:

Nearly half (49%) of the lobsters sampled were female.

**Chi-squared statistics** are reported with degrees of freedom and sample size in parentheses, the Pearson chi-squared value (rounded to two decimal places), and the significance level. For example:

The number of lobsters that were larger than minimum legal size did not differ by gender (*Chisq*<sub>(1,90)</sub> = 0.89, *P* = 0.35).

***t*-tests** are reported with the degrees of freedom as a subscript followed by the *t* statistic (rounded to two decimal places) and the significance level. For example:

There was a significant difference in the density of snapper between reserve (mean = 4.2 100 m<sup>-2</sup>, *SE* = 1.3) and non-reserve (mean = 2.2 100 m<sup>-2</sup>, *SE* = 0.84) areas: *t*-test: *t*<sub>8</sub> = 2.89, *P* = 0.02).

**ANOVAs** (both one-way and two-way) are reported like the *t*-test, but there are two degrees-of-freedom numbers to report. First, report the between-groups degrees of freedom, then report the within-groups degrees of freedom (separated by a comma). After that, report the *F* statistic (rounded off to two decimal places) and the significance level. For example:

There was a statistically significant difference in the density of snapper (*F*<sub>1,30</sub> = 40.65, *P* < 0.05) between reserve (mean = 16.25 indiv. 100 m<sup>-2</sup>, *SE* = 0.42) and non-reserve sites (mean = 12.19 indiv. 100 m<sup>-2</sup>, *SE* = 0.48).

**Correlations** are reported with the degrees of freedom (which is the number of observations minus 2) in parentheses and the significance level. For example:

The two variables were strongly correlated, *r*(55) = 0.59, *P* < 0.001.

**Regression** results are reported as the slope, along with the *t*-test and the corresponding significance level. You should also report the percentage of variance explained along with the corresponding *F*-test. For example:



Kelp cover explained a significant proportion of variance in butterflyfish abundance ( $R^2 = 0.649$ ,  $F_{1,14} = 25.93$ ,  $P < 0.001$ ) and had a significant negative relationship with the abundance of butterflyfish ( $\beta = 5.434$ ,  $SE = 1.067$ ,  $P < 0.001$ ).

## References and further reading

- Anderson, D.R.; Link, W.A.; Johnson, D.H.; Burnham, K.P. 2001: Suggestions for presenting the results of data analysis. *The Journal of Wildlife Management* 65(3): 373–378. <http://www.jstor.org/stable/3803088>
- Anderson, R.O.; Neumann, R.M. 1996: Length, weight, and associated structural indices. Pp. 447–482 in Murphy, B.R.; Willis, D.W. (Eds): *Fisheries techniques*. 2nd edition. American Fisheries Society, Bethesda, Maryland.
- Clark, M.; Randal, J.A. 2004. *A first course in applied statistics: with applications in biology, business and the social sciences*. Pearson Education New Zealand, Auckland.
- Crawley, M.J. 2012: *The R book*. John Wiley & Sons, West Sussex.
- Faith, D.P. 1992: Conservation evaluation and phylogenetic diversity. *Biological Conservation* 61: 1–10. [https://doi.org/10.1016/0006-3207\(92\)91201-3](https://doi.org/10.1016/0006-3207(92)91201-3)
- Goodman, L.A. 1960: On the exact variance of products. *Journal of the American Statistical Association* 55: 708–713.
- Gotelli, N.J.; Ellison, A.M. 2004: *A primer of ecological studies*. Sinauer, Sunderland, Massachusetts.
- Harley, S.J.; Myers, R.A.; Dunn, A. 2001: Is catch-per-unit-effort proportional to abundance? *Canadian Journal of Fisheries and Aquatic Sciences* 58: 1760–1772. <https://doi.org/10.1139/f01-112>
- Hedges, L.; Olkin, I. 1985: *Statistical methods for meta-analysis*. Academic Press, New York.
- Hedges, L.V.; Gurevitch, J.; Curtis, P.S. 1999: The meta-analysis of response ratios in experimental ecology. *Ecology* 80: 1150–1156. [https://doi.org/10.1890/0012-9658\(1999\)080\[1150:TMAORR\]2.0.CO;2](https://doi.org/10.1890/0012-9658(1999)080[1150:TMAORR]2.0.CO;2)
- Hubert, W.A.; Fabrizio, M.C. 2007: Relative abundance and catch per unit effort. In Guy, C.S.; Brown, M.L. (Eds): *Analysis and interpretation of freshwater fisheries data*. American Fisheries Society, Bethesda, Maryland.
- Institute for Digital Research and Education. 2017. Choosing the correct statistical test in SAS, STATA, SPSS and R. <https://stats.idre.ucla.edu/other/mult-pkg/whatstat/>
- Isaac, N.J.B.; Turvey, S.T.; Collen, B.; Waterman, C.; Baillie, J.E.M. 2007: Mammals on the EDGE: conservation priorities based on threat and phylogeny. *PLoS ONE* 2(3): e296. <https://doi.org/10.1371/journal.pone.0000296>



- Kelly, D.; Jasperse, J.; Westbrooke, I. 2005: Designing graphs for data analysis and presentation. *Department of Conservation Technical Series 32*. 68 p.
- Lee, W.; McGlone, M; Wright, E. 2005: Biodiversity Inventory and Monitoring: A review of national and international systems and a proposed framework for future biodiversity monitoring by the Department of Conservation. Landcare Research Contract Report: LC0405/122. Department of Conservation, Wellington.  
[http://www.landcareresearch.co.nz/publications/researchpubs/biodiv\\_inventory\\_system\\_review\\_framework.pdf](http://www.landcareresearch.co.nz/publications/researchpubs/biodiv_inventory_system_review_framework.pdf)
- Petchey, O.L.; Gaston, K.J. 2002: Functional diversity (FD), species richness and community composition. *Ecology Letters* 5: 402–411. <https://doi.org/10.1046/j.1461-0248.2002.00339.x>
- Petchey, O.L.; Gaston, K.J. 2006: Functional diversity: back to basics and looking forward. *Ecology Letters* 9: 741–758. <https://doi.org/10.1111/j.1461-0248.2006.00924.x>
- Quinn, G.P.; Keough, M.J. 2002: Experimental design and data analysis for biologists. Cambridge University Press, Cambridge.
- Sokal, R.R.; Rohlf, F.J. 1995: Biometry: the principals and practice of statistics in biological research. WH Freeman and Company, New York.
- Stevens, M.H. 2009: A primer of ecology with R. Springer Science & Business Media, New York.
- Vane-Wright, R.I.; Humphries, C.J.; Williams, P.H. 1991: What to protect? Systematics and the agony of choice. *Biological Conservation* 55(3): 235–254. [https://doi.org/10.1016/0006-3207\(91\)90030-D](https://doi.org/10.1016/0006-3207(91)90030-D)
- Vellend, M.; Cornwell, W.K.; Magnuson-Ford, K.; Mooers, A.Ø. 2011: Measuring phylogenetic biodiversity. Pp 194–207 in Magurran, A.E.; McGill, B.J. (Eds): Biological diversity: frontiers in measurement and assessment. Oxford University Press, Oxford.
- Warwick, R.M.; Clarke, K.R. 1995: New 'biodiversity' measures reveal a decrease in taxonomic distinctness with increasing stress. *Marine Ecology Progress Series* 129: 301–305.  
<http://www.int-res.com/articles/meps/129/m129p301.pdf>
- Warwick, R.M.; Clarke, K.R. 1998: Taxonomic distinctness and environmental assessment. *Journal of Applied Ecology* 35: 532–543. <https://doi.org/10.1046/j.1365-2664.1998.3540532.x>
- Webb, C.O. 2000: Exploring the phylogenetic structure of ecological communities: an example for rain forest trees. *The American Naturalist* 156: 145–155. <https://doi.org/10.1086/303378>
- Winter, M.; Devictor, V.; Schweiger, O. 2013: Phylogenetic diversity and nature conservation: Where are we? *Trends in Ecology & Evolution* 28: 199–204. <https://doi.org/10.1016/j.tree.2012.10.015>

