Global Ecology & Biogeography (2003) **12**, 177–179

ECOLOGICAL SOUNDING

A tribute to Claude Shannon (1916–2001) and a plea for more rigorous use of species richness, species diversity and the 'Shannon–Wiener' Index

IAN F. SPELLERBERG* and PETER J. FEDOR‡

**Isaac Centre for Nature conservation, Lincoln University, New Zealand. E-mail: Spelleri@lincoln.ac.nz; ‡Department of Ecosozology, Faculty of Natural Sciences, Comenius University, Slovakia. E-mail: fedor@fns.uniba.sk*

ABSTRACT

In the literature, the terms species richness and species diversity are sometimes used interchangeably. We suggest that at the very least, authors should define what they mean by either term. Of the many species diversity indices used in the literature, the Shannon Index is perhaps most commonly used. On some occasions it is called the Shannon–Wiener Index and on other occasions it is called the Shannon–Weaver Index. We suggest an explanation for this dual use of terms and in so doing we offer a tribute to the late Claude Shannon (who passed away on 24 February 2001).

Key words diversity indices, Shannon Index, Shannon–Wiener Index, Shannon–Weaver Index, species richness, species diversity.

INTRODUCTION

There is much research on 'diversity' at different scales and levels of biological organization. Most of this research is at the species level (which may contribute to an apparently common perception that biological diversity equals variety at the species level of biological organization). Species diversity occurs at different spatial scales and in a recent paper, Whittaker *et al*. (2001) provided a much-needed solution to the confusion surrounding the terms of alpha, beta and gamma diversity.

In this paper, we confine our discussion to the use of the terms species richness and species diversity and also to the use of both 'Shannon–Wiener' and 'Shannon–Weaver' as an index of species diversity.

As has been noted previously by some authors (for example Krebs (1999)) the Shannon–Wiener function is sometimes mislabelled the Shannon–Weaver function. However, no explanation is suggested by Krebs as to why it is sometimes mislabelled. A second aim, therefore, was to explore why the Shannon Index of diversity is sometimes referred to as the Shannon–Wiener Index and on other occasions referred to as the Shannon–Weaver Index . The two aims of this paper have something in common: a plea for more rigorous use of these terms.

Correspondence: Ian F. Spellerberg, Isaac Centre for Nature conservation, Lincoln University, New Zealand, E-mail: Spelleri@lincoln.ac.nz

SPECIES RICHNESS AND SPECIES DIVERSITY

The terms diversity, species diversity and species richness are sometimes used in confusing ways. In some papers the term 'species richness' is used in the title and in the text it is assumed to mean the number of species, but this may not be made clear (Bruun, 2000). By way of contrast, 'diversity' is sometimes used in the title but in the text the data seem to refer to the number of species only or the number of other taxa (Schedvin *et al*., 1994).

A brief search of the journals *Biodiversity Letters*, *Diversity and Distributions* and *Global Ecology and Biogeography* revealed that in some papers 'diversity' or 'species richness' is used in the title but species richness and diversity may be used almost interchangeably in the text (Cotgreave & Harvey, 1994; Honnay *et al*., 1999; Harrison *et al*., 2000; Brown, 2001; Heaney, 2001; Lobo *et al*., 2001).

We were pleased to find that in some papers the term 'species richness' was used to refer to the number of species (Kessler *et al*., 2001) and in other papers a clear distinction was made between 'species richness' and 'species diversity' (Sax, 2002).

Although species richness is a measure of variety of species, we suggest that 'species richness' should be used to refer to the number of species (in a given area or in a given sample). It is not uncommon for species diversity to be expressed in the

form of a species diversity index. It is suggested that 'species diversity' could be retained for use in this context, that is as an expression or index of some relation between number of species and number of individuals (Spellerberg, 1991).

We are not suggesting that species diversity can be only one measure in the sense of number of species and number of individuals. Rather, we simply suggest that it would be helpful to the reader if the title of research papers used terms exactly as they are used in the main part of the text. That is, if the research is about species richness only (which we agree is a measure of variety or diversity), then the title should refer to species richness. We also suggest that rather than using the terms species richness and species diversity interchangeably, it is helpful to distinguish between these two terms.

THE 'SHANNON' INDEX OF SPECIES DIVERSITY

Several indices of species diversity are used in the large amount of literature on biological diversity and ecological monitoring. A commonly used index is that referred to as 'Shannon's Index' or '*H*'.

This Index is based on communication theory and stems from a common question in communication: how to predict the next letter in a message or communication? The uncertainty is measured by the Shannon Function '*H*'. This is the measure corresponding to the entropy concept defined by:

$$H = -\sum_{i=1}^{n} pi \ln pi$$

In the literature, the 'Shannon Index' is sometimes referred to as the 'Shannon–Weaver' Index (see, for example Poole, 1974; Niklaus *et al*., 2001) and sometimes as the 'Shannon–Wiener' Index (see for example Hixon & Brostoff, 1983; Sax, 2002). When references from current research reports on species diversity indices are tracked back via the cited references, it is not unusual for Shannon to be listed in the bibliography. Information in the BIOSIS literature database for 1993–2002 suggests that the term 'Shannon' is used far more (1105 times) than either 'Shannon–Wiener' (214) or Shannon–Weaver (165). It appears that 'Shannon–Wiener' is used slightly more than Shannon–Weaver.

That is, a common reference to 'Shannon's Function H' for species diversity is the following, for which there is more than one reprint (1949–1971):

> Shannon, C.E. & Weaver, W. (1949) *The mathematical theory of communication*. The University of Illinois Press, Urbana, 117pp.

To add to the mistakes in the literature, some researches have attributed this publication to Shannon & Wiener (see for example Burchfield, 1993). Shannon & Weaver's small book contains two sections or two reports, which are republished versions of previous reports from about 14 years previously:

1 'The mathematical theory of communication' by Claude E. Shannon, Bell Telephone Laboratories. This paper is reprinted from the *Bell System Technical Journal*, July and October, 1948. This is the same paper with some corrections and additional references.

2 'Recent contributions to the mathematical theory of communication' by Warren Weaver, the Rockefeller Foundation. This paper had not previously been published in this form but a condensed version appeared in *Scientific American* in July 1949.

CLAUDE SHANNON (1916–2001)

According to Sloan & Wyner (1993), Claude Shannon became interested in mathematical sciences at an early age. He graduated in mathematics as well as in electrical engineering. In 1936 Shannon was a research assistant in the Department of Electrical Engineering at the Massachusetts Institute of Technology. He then later spent a short time in advanced study at Princeton University. In 1941 he commenced work for the Bell Telephone laboratories in New Jersey where he was to spend 15 years among a well-respected scientific community. Following his research on information communication he summarized his ideas (Shannon, 1948). It was this paper in the *Bell System Technical Journal* that there was a reference to Shannon's Mathematical Theory of Communication, based on notes contained in studies by Nyquist (1924, 1928) and Hartley (1928). It includes the first form of the present Shannon expression as:

$$H = -K\sum_{i=1}^{n} pi \log pi$$

where K is a positive constant. This expression has a central role in information theory as a measure of information, choice and uncertainty (Shannon, 1948).

In 1949, Shannon published this information in the jointly authored book *The mathematical theory of communication* (Shannon & Weaver, 1949). Warren Weaver was a mathematician and the text in the book is very similar to the original published papers.

Thus, Shannon first published an account of the entropy 'H' in 1948. Weaver builds on this in 1949 in the second part of the above book.

In his paper Shannon acknowledges the fact that 'communication theory is heavily indebted to the mathematician Norbert Wiener for much of its basic philosophy and theory' and cites several of his publications that refer to basic cybernetics (for example Wiener, 1939; see also Wiener, 1948, 1949). Shannon also refers to earlier work, including that of Boltzmann. It seems, therefore, that at about this time

there occurred an evolution of ideas about the principles of statistical physics.

We suggest that the 'mislabelling' of the Shannon Index 'H' (as referred to by Krebs, 1999) has come about partly because of the joint authorship of Shannon & Weaver's book, which has led to a belief that these two authors can be attributed to the Index. That is, the Shannon index is sometimes called the Shannon & Weaver Index.

In fact, in the late 1940s Shannon had built on the work of Wiener. That being the case, it seems preferable to refer to 'H' (the species diversity index) as the 'Shannon Index' or the 'Shannon & Wiener Index'.

Had Weaver's name been anything else and not similar to 'Wiener', this confusion may not have arisen. If the book published in 1949 had been by Shannon and 'Smith' then perhaps there would have been less confusion.

In summary, we suggest that the mislabelling of the Shannon Index arises from the following:

1 A frequently cited reference to the Shannon Index is the book by Shannon & Weaver (1949).

2 There has been an assumption that Shannon and Weaver developed the function 'H'.

3 The names Wiener and Weaver are similar (sometimes Wiener is spelt incorrectly as Weiner).

4 It is not uncommon in publications for authors to quote and re-quote references without going back to the original source of the material. This is, we believe, a common fault in ecological research.

REFERENCES

Brown, J.H. (2001) Mammals on mountainsides: elevational patterns of diversity. *Global Ecology and Biogeography*, **10**, 101–109.

Bruun, H.H. (2000) Deficit in community species richness as explained by area and isolation of sites. *Diversity and Distributions*, **6**, 129–135.

Burchfield, M.L. (1993) Using pitfall traps to investigate species diversity. *Journal of Biological Education*, **27**, 217–219.

Cotgreave, P. & Harvey, P.H. (1994) Associations among biogeography, phylogeny and bird species diversity. *Biodiversity Letters*, **2**, 46–55.

Harrison, S., Viers, J.H. & Quinn, J.F. (2000) Climatic and spatial patterns of diversity in serpentine plants of California. *Diversity and Distributions*, **6**, 153–161.

Hartley, R.V.L. (1928) Transmission of information. *Bell System Technical Journal*, **7**, 535–563.

Heaney, L.R. (2001) Small mammal diversity along elevational gradients in the Philippines: an assessment of patterns and hypotheses. *Global Ecology and Biogeography*, **10**, 15–39.

Hixon, M.A. & Brostoff, W.N. (1983) Damselfish as keystone species in reverse: intermediate disturbance and diversity of reef algae. *Science*, **220**, 511–513.

Honnay, O., Vereecken, P.E.H. & Hermy, M. (1999) The role of patch area and habitat diversity in explaining native plant species richness in disturbed suburban forest patches in northern Belgium. *Diversity and Distributions*, **5**, 129–141.

Kessler, M., Herzog, S.K., Fjeldsa, J. & Bach, K. (2001) Species richness and endemism of plant and bird communities along two gradients of elevation, humidity and land use in the Bolivian Andes. *Diversity and Distributions*, **7**, 61–77.

Krebs, C.J. (1999) *Ecological methodology*. Addison Wesley Longman, Menlo Park, CA.

Lobo, J.M., Lumaret, J.-P. & Jay-Robert, P. (2001) Diversity, distinctiveness and conservation status of the Mediterranean coastal dung beetle assemblage in the Regional Natural Park of the Camargue (France). *Diversity and Distributions*, **7**, 257–270.

Niklaus, P.A., Leadley, P.W., Schmid, B. & Korner, Ch (2001) A long-term field study on biodiversity X elevated CO_2 interactions in grassland. *Ecological Monographs*, **71**, 341–356.

Nyquist, H. (1924) Certain factors affecting telegraph speed. *Bell System Technical Journal*, **3**, 324–346.

Nyquist, H. (1928) Certain topics in telegraph transmission theory. *Transactions of the American Institute of Electrical Engineers*, **47**, 617.

Poole, R.W. (1974) *An introduction to quantitative ecology*. McGraw-Hill, New York.

Sax, D.F. (2002) Equal diversity in disparate species assemblages: a comparison of native and exotic woodlands in California. *Global Ecology and Biogeography*, **11**, 49–57.

Schedvin, N.K., Cook, S.P. & Thornton, I.W.B. (1994) The diversity of bats on the Krakatau Islands in the early 1990s. *Biodiversity Letters*, **2**, 87–92.

Shannon, C.E. (1948) A mathematical theory of communication. *Bell System Technical Journal*, **27**, 379–423.

Shannon, C.E. & Weaver, W. (1949) *The Mathematical Theory of Communication*. University of Illinois Press, Urbana.

Sloane, N.J.A. & Wyner, A.D. (1993) *Claude Elwood Shannon: collected papers*. Institute of Electrical and Electronic Engineers Press, New York.

Spellerberg, I.F. (1991) *Monitoring ecological change*. Cambridge University Press, Cambridge.

Whittaker, R.J., Willis, K.J. & Field, R. (2001) Scale and species richness: towards a general, hierarchical theory of species diversity. *Journal of Biogeography*, **28**, 453–470.

Wiener, N. (1939) The ergodic theorem. *Duke Mathematical Journal*, **5** [page numbers omitted in the paper].

Wiener, N. (1948) *Cybernetics*. Wiley, New York.

Wiener, N. (1949) *The interpolation, extrapolation, and smoothing of stationary time series*. Wiley, New York.

BIOSKETCHES

Ian Spellerberg's research is centred around the ecological basis of nature conservation and includes biogeography, ecological evaluation, ecological monitoring and the ecology of roads and traffic.

Peter Fedor deals with systematic entomology (Orthoptera, Thysanoptera), the ecology of insect communities and their bioindication potential in environmental practice, especially to evaluate anthropogenic impact. He works as a teacher of invertebrate zoology, bioindication and biomonitoring.