How different is different? Measuring diversity

C. E. Cloete
*Department of Construction Economics,
University of Pretoria, South Africa*

Abstract

Resilience of a city can be defined as the capacity of the system to experience shocks while retaining essentially the same function, structure, feedbacks and identity without shifting into a different regime. One of the factors that enhances resilience is diversity. It is therefore of particular interest to develop a measure that can compare the diversity of a city before and after a natural or man-made shock, or that can be used to indicate the difference in diversities of different cities.

The paper suggests a possible approach to the problem by firstly elaborating on the concept of diversity and the use of an entropy-based measure for the diversity of a system and secondly by proposing a measure for the degree of difference between two diversities. The deduction of the entropy-based measure is general, and therefore in principle applicable to the study of the diversity of cities as well.
Keywords: resilience, diversity, difference, entropy, cities.

1 Introduction

The concept of sustainable development has become an integral part of the discourse on the future of human society and concepts like complexity, sustainability, adaptability and resilience have been widely explored particularly since their advent in the 1980s. A science of sustainability necessarily requires interdisciplinary collaboration between different theoretical and applied scientific disciplines. It is disconcerting that while the social sciences contribute about a third of the total output in terms of total number of publications, the contribution of urban studies *per se* is less than 4 per cent [1].

In an attempt to place the study of the diversity of cities on a more quantitative foundation, the application of the concept of a general separation

WIT Transactions on Ecology and The Environment, Vol 179, © 2013 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/SC130011

index as originally developed in chemistry was investigated. It is proposed that the measure can be applied to the measurement of the diversity of cities, to the difference between diversities of different cities and also to the difference in diversity of a city before and after a shock.

2 Diversity

In spite of it being a relatively new concept and in danger of becoming just another buzzword, resilience has gained considerable prominence in ecological, biological, and also social sciences. Resilience of any system is determined by, among others, the diversity in that system [2–14]. In addition to the degree of connectedness within the system and the tightness of feedback, the continued sustainability of a complex system is enhanced by having sufficient diversity in the system's components. Consideration of a number of resilient systems, including engineering systems, led Fiksel [5] to conclude that "Characteristics such as diversity and adaptability may not have an obvious relationship to system performance but may contribute to the system's longevity and ultimate success."

The term diversity can be used to describe the distribution of differences among the members of a unit with respect to a common unit. A variety of measures of diversity has been proposed in different disciplines. The different types that had variously been proposed have been categorized by Page [15] and by Harrison and Klein [16]. These categorizations are briefly summarized below.

Page [15] differentiates among five types of measures: variation, entropy, distance, attribute and population measures. Entropy measures consider the number of types and the distribution among those types. The so-called Shannon entropy is a special case of the class of generalized entropy functions (cf. [15: 69]).

Harrison and Klein [16] distinguish three fundamental types of diversity constructs: separation, variety, and disparity. The three types differ in their substance, shape, maxima, and implications. *Separation* describes differences among unit members in their position on a horizontal continuum. *Variety* describes differences among unit members from different categories, reflecting access to unique sources of knowledge. *Disparity* describes differences among unit members in their portion of a valued resource. The measurement of variety can be operationalized by means of the Blau index or by means of entropy.

Harrison and Klein [16: 22] conclude their analysis of diversity by urging "...future investigators to specify the diversity types they are studying, and to align them with specific, appropriate operationalizations. By systematically asking and answering "what's the difference?" management scholars may reveal a clearer, more cumulative understanding of diversity in organizations."

This exhortation is equally applicable to the study of resilience of cities and the effect of diversity on sustainability. The following sections propose that diversity in two cities or in one city at different times may be compared by using entropy as a separation criterion, analogous to the use of entropy as a separation criterion in chemical processes (e.g. [17, 18]).

3 Entropy

As pointed out by Glucina and Mayumi [19], the debate about the relevance of thermodynamics to economics, such as those put forward by Young [20, 21] and Khalil [22, 23], have been based on misunderstandings of entropy and the second law of thermodynamics [24]. Also, some attempts to bridge the two fields have been based on similar misunderstandings [25–27]. "Entropy may fairly be called one of the great buzzwords of twentieth-century science. The very abstractness and obscurity of the term evokes in laymen an aura of mystery and arcane knowledge." [28: 22]. It is therefore necessary to clarify what entropy is and what it is not.

Entropy is related to the number of possible distributions in a system. A system with a relatively high number of possible distributions is said to have high entropy, and likewise, a smaller number of possible distributions corresponds to lower entropy.

The entropy measure $H= - \sigma p_i \log p_i$ is being used with increasing frequency in the analysis of business and economic data [29, 30]. It is, however, simply another measure of dispersion which can be related to the moments of the probability function. Its virtues stem from its decomposition and interpretative properties [31]. It captures distributions over types, where types are things with the same attributes. Definition of an attribute is not necessarily unique; it can vary according to the question asked [25].

4 Entropy of the separation process

Separation of entities can be considered as a reduction of the entropy of a mixture of the entities. The entropy S_o of a mixture can be expressed as [32]:

$$S_o = - kN \sum \alpha_i \ln \alpha_i \tag{1}$$

where N is the (sufficiently large) number of entities, k is the analog of the Boltzmann constant, α_i denotes *i*'s proportion in the population and ln is the natural logarithm (base *e*).

We also have:

$$\sum_i^n \alpha_i = 1 \tag{2}$$

where $i \in \{1,2, ..., n\}$ denotes the types in the population,
$\alpha_i = m_i / M$,
m_i denotes the number of types in the population

and $M = \sum_i^n m_i$ is the size of the total population.

The decrease in entropy following the separation process is given by [32: 44]:

$$\Delta S = -kN\left(S_o - \sum_{1}^{p} \gamma_j \sum_{1}^{n} \beta_{ij} \ln\beta_{ij}\right) \leq 0 \tag{3}$$

where γ_j is the fraction of the j-product
and β_{ij} the content of the i-th component in the latter.

From eq. (3) we can see that in the case of complete separation (Figure 1(a)) $\Delta S = S_o$, i.e. the entropy after separation is $S = 0$. In the case of no separation, $\Delta S = 0$, $S = S_o$ (Figure 1(b)) [32: 45].

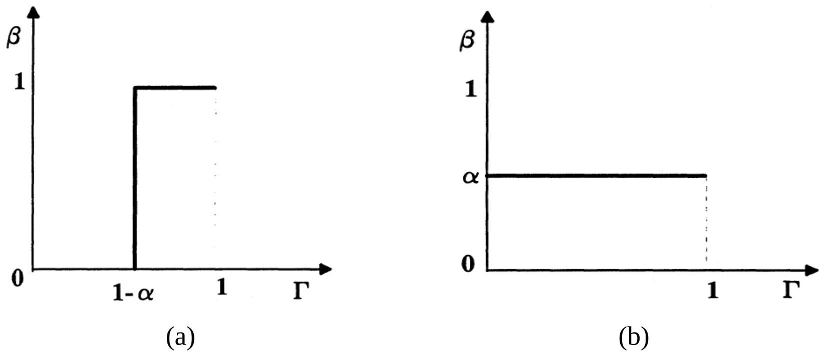

Figure 1: (a) The case of a complete separation. (b) The case of no separation.

In this respect it is intriguing to consider the Similarity Principle postulated by Lin [23]. If all the other conditions remain constant, the higher the similarity among the components of an ensemble (or a considered system) is, the higher value of entropy of the mixture or the assemblage or any other structure will be, the more stable the mixture or the assemblage will be, and the more spontaneous the process leading to such a mixture or assemblage will be.

The similarity Z can be easily understood when two items A and B are compared: if A and B are distinguishable (minimal similarity), Z=0. If they are indistinguishable (maximal similarity), Z=1. Lin argues that the entropy of mixing or assembling increases continuously with the increase in the similarity (Figure 2).

This statement is supported by the following argument:
From the well-known inequality

$$-\sum_{i}^{n} \alpha_i \ln \alpha_i \leq \ln n \tag{4}$$

and the general entropy expression

$$S_o = -\sum_{i}^{n} \alpha_i \ln \alpha_i \tag{5}$$

the condition for the maximum entropy must be the indistinguishability among the w components.

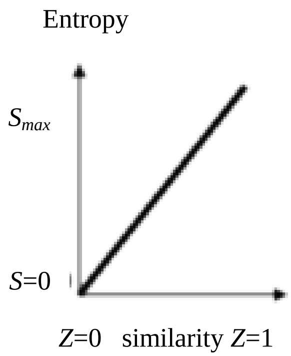

Figure 2: Relationship between entropy and similarity [23].

From eq. (4),

$$Z = S/S_{max} = -\left[\sum_{i}^{n} \alpha_i \ln \alpha_i\right] / \ln n \qquad (6)$$

defines a similarity index, and entropy increases continuously with the property similarity of the w subsystems. The similarity depends directly on the similarity among the considered components.

The increase in entropy when m substances are mixed is given by (e.g. [34, 35])

$$\Delta S = -R \sum_{i=1}^{w} n_i \ln y_i \qquad (7)$$

where any attractive or repulsive forces are neglected. n_i is the number of component i and y_i is its fraction in the mixture. Equation (1) can be re-written as

$$\Delta S = S_2 - S_1$$

where

$$S_2 = (\sum n_i) \ln (\sum n_i)$$

and

$$S_1 = \sum n_i \ln(\sum n_i) \qquad (8)$$

can be regarded as the entropy of the completely mixed and completely pure states, respectively. The latter may here be regarded as a convenient reference state. For the present purpose, the coefficient R is arbitrary and is conveniently set equal to unity.

5 Entropy as a measure of diversity

Following the arguments of De Clerk and Cloete [17] it will now be shown that the concept of entropy may be used to formulate a suitable index for the difference between distribution across types. This may be done in a completely abstract manner. Consider, for instance, the distributions in Figure 3. If the fractions were cut at the indicated point and collected separately, the entropy in the jth vessel would be given by

$$S_j = (\sum_i n_{ij}) \ln (\sum_i n_{ij}) \tag{9}$$

where n_{ij} is the number of units of component i in the jth vessel. We will use the convention that the jth region contains the largest fraction of the jth component.

This definition is not complete, however, since the reference state relative to which the entropy is measured has not yet been defined. The definition of such a state is arbitrary but for the purposes of characterizing separation it is convenient to take it relative to the state in which all the components of the jth region have been completely separated. This is illustrated schematically as State A in Fig. 3.

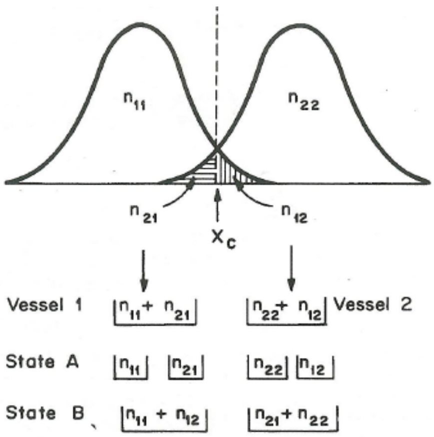

Figure 3: Illustration of the entropy of separation.

The reason for not using State B as reference is also apparent from this figure; the extra entropy change in going from A to B is not physically significant in the separation sense. The expression for the total entropy then becomes

$$S = \sum_j S_j \tag{10}$$

where

$$S_j = (\sum_i n_{ij}) \ln (\sum_i n_{ij}) - (\sum_i n_{ij} \ln n_{ij}) \tag{11}$$

Usually the extent of separation as such is characterized, and the number of units separated in the process is regarded as a separate problem. Since the present analysis is concerned with the former goal, it is convenient to introduce the concept of specific entropy S_j' of a region j. This is defined by the entropy per unit of the region and follows from Eq. (5) as

$$-S_j' = \frac{(\sum_i n_{ij} \ \ln n_{ij})}{(\sum_i n_{ij})} - \ln(\sum n_{ij}) \tag{12}$$

with
$$S_j' = \sum_j S_j' \tag{13}$$

The expression for specific entropy given in Eq. (13) may be inconvenient to use in practice due to difficulties in its measurement, mathematical manipulation, or both. To circumvent such difficulties functions can be introduced which are in a 1-1 correspondence with the specific entropy and which exhibit extrema at identical values of the independent variables. Such functions will be termed resolution functions, and will be illustrated here by means of a binary mixture.

For a binary mixture Eq. (13) becomes

$$\begin{aligned} -S' &= \frac{n_{11}\ln n_{11} + n_{21}\ln n_{21}}{n_{11}+n_{21}} - \ln(n_{21}+n_{11}) \\ &+ \frac{n_{22}\ln n_{22} + n_{12}\ln n_{12}}{n_{11}+n_{21}} - \ln(n_{22}+n_{12}) \end{aligned} \tag{14}$$

Equation (14) may be re-written as:

$$S' = S_1' + S_2'$$

with

$$-S_1' = \frac{1}{1+\eta_1}\ln\left(\frac{1}{1+\eta_1}\right) + \frac{\eta_1}{1+\eta_1}\ln\left(\frac{\eta_1}{1+\eta_1}\right) \tag{15}$$

and

$$-S_2' = \frac{1}{1+\eta_2}\ln\left(\frac{1}{1+\eta_2}\right) + \frac{\eta_2}{1+\eta_2}\ln\left(\frac{\eta_2}{1+\eta_2}\right) \tag{16}$$

where $\eta_1 = \eta_{21}/\eta_{11}$ and $\eta_2 = \eta_{12}/\eta_{22}$ are the impurity fractions (5) of the respective regions.

S_j' is seen to be a function only of the impurity fraction of the *j*th region. The task in formulating a resolution function for S_j' is therefore to find a suitable function of η_j which is in a 1-1 correspondence with S_j'.

Equation (15) may be simplified by noting that as η_j decreases,

$$S_j' \rightarrow -\eta_j \ ln\, \eta_j \tag{17}$$

This is an acceptable approximation for $0 < \eta_j < 0.1$.

Equation (17) suggests two suitable resolution functions, viz.

$$M_j = \eta_j \tag{18}$$

and

$$I_j = -log_{10}\, \eta_j \tag{19}$$

M_j is merely the impurity ratio as originally suggested by Glueckauf [36]; I_j will be termed the purity index and, since it has been found to be mathematically particularly convenient, merits further discussion.

The fact that I_j and S_j' are functionally related does not imply that this also applies to $S' = \sum S_j'$ and $I = \sum I_j$ since a specific value for I is not uniquely related to a value of S'. This does not disqualify I as an index for separation; the relative merits of S' and I depend on their relation to the significant factors involved in the separation.

The direct link of S' with thermodynamics promises a more generally applicable role for the specific entropy in separation science; the purity index, on the other hand, is mathematically more convenient. Both I and S' are, however, defined on the same basis, viz., the η_js, and can be formally put into 1-1 correspondence by considering the quantities defined by setting all the η_js equal to each other.

6 Conclusions and recommendations

The analysis presented above has demonstrated the potential of the entropy concept as a measure of diversification in general. It should be possible to apply this measure to determine the difference in diversity of a city before and after a shock or to indicate the difference in diversities of different cities. Care should be taken, however, to formulate the appropriate question as the definition of an attribute is not necessarily unique – it is a phenomenological issue, determined by the investigator [37–39].

It is also suggested that the entropy measure should be weighted by incorporating the dispersion of frequency classes (i.e. dispersion-related state-values) used in the entropy calculation, possibly weighting of entropy by squared deviations from the mean and entropy weighted by absolute deviations from the mean [40, 41].

An additional potentially fruitful line of enquiry could be the application of the Similarity Principle of Lin [33] referred to above to socio-ecological systems [42], to neighborhood homogeneity and cohesion in sustainable community development [43], to community resilience [44] and to cohesion in cities [45].

References

[1] Bettencourt, Luis M.A. and Kaur, J. The Evolution and Structure of Sustainability Science. *Proc. Nat. Acad. Sci. USA*, 108: 19540–19545, 2011.

[2] Davoudi, S. Resilience: A bridging concept or a dead end? *Planning Theory and Practice* 13(2): 299–333, 2012.

[3] Walker, B.H. and Salt, D. Resilience thinking. Sustaining ecosystems and people in a changing World, Washington D.C.: Island Press, 2006.

[4] deGeus, A. *The Living Company*. Harvard Business School Press: Cambridge, 1997.

[5] Fiksel, J. Designing Resilient, Sustainable Systems. *Environ. Sci. Technol.* 37: 5330–5339, 2003.

[6] Sanders, T. Irene, Complex Systems Thinking and New Urbanism, in *New Urbanism and Beyond: Designing Cities for the Future*. Tigran Haas (ed.), New York: Rizzoli, pp. 275–279, 2008.

[7] Sendzimir, J., *Critical examination of resilience theories and measures.* Report to CAVES project. Deliverable 6, EU 6th Framework, 2006.

[8] Holling, C.S. Understanding the Complexity of Economic, Ecological and Social Systems, *Ecosystems*, 4, pp. 390–405, 2001.

[9] Reggiani, A., de Graff, T. and Nijkamp, P., Resilience: An Evolutionary Approach to Spatial Economic Systems, *Networks and Spatial Economics*, 2, pp. 211–229, 2002.

[10] Meadows, Donella H. *Thinking in Systems*. Vermont: White River Junction: Chelsea Green Publishing, 2008.

[11] Pike, A., Dawley, S. and Tomaney, J. Resilience, Adaptation and Adaptability, *Cambridge Journal of Regions, Economy and Society*, 3(1): 59–70, 2010.

[12] Holling, C.S., Resilience and Stability of Ecological Systems, *Annual Review of Ecology and Systematics*, **4,** 1–23, 1973.

[13] McGlade, J., Murray, R., Baldwin, J., Ridgway, K. and Winder, B., Industrial Resilience and Decline: A Co-evolutionary Framework, in Garnsey, E. and McGlade, J. (eds.), *Complexity and Co-Evolution: Continuity and Change in Socio-economic Systems*, Cheltenham: Edward Elgar, 147–176, 2006.

[14] Vale, L.J. and Campanella. T.J., (eds.) *The Resilient City*, Oxford University Press: New York, 2005.

[15] Page, S.E. *Diversity and Complexity*. Princeton University Press, Oxford and Princeton, 2011.

[16] Harrison, D.A. and Klein, K.J. What's the Difference? Diversity constructs as separation, variety or disparity in organizations. *Academy of Management Review*, 2007.

[17] De Clerk, K. and Cloete, C.E. Entropy as a general separation criterion. *Separation Science* 6(5): 627–635, 1971.

[18] Stewart, G.H. On the measurement and evaluation of separation. *Separation Science and Technology* 13(3): 201–213, 1978.

[19] Glucina, M.D. and Mayumi, K. Connecting thermodynamics and economics Well-lit roads and burned bridges. *Ann. N.Y. Acad. Sci.* 1185: 11–29, 2010.
[20] Young, J.F. 1994. Entropy and natural resource scarcity: a reply to the critics. *J. Environ. Econ. Manage*. 26: 210–213.
[21] Young, J.T. 1991. Is the entropy law relevant to the economics of natural resource scarcity? *J. Environ. Econ. Manage.* 21: 169–179.
[22] Khalil, E. 1990. Entropy law and exhaustion of natural resources: is Nicholas Georgescu-Roegen's paradigm defensible? *Ecol. Econ*. 2: 163–178.
[23] Khalil, E.L. 1991. Entropy law and Nicholas Georgescu-Roegen's paradigm: a reply. *Ecol. Econ.* 3: 161–163.
[24] Faber, M. *et al*. 1996. The use of the entropy concept in ecological economics. In *Ecological Economics: Concepts and Methods:* 115–135, Edward Elgar, Cheltenham.
[25] Lozada, G.A. 2004. Entropy and the economic process. In *Encyclopaedia of Energy,* Vol. 2: 472–477, Elsevier.
[26] Gillet, S.L. 2005. Entropy and its misuse, I. Energy, free and otherwise. *Ecol. Econ*. 56: 58–57.
[27] Lozada, G.A. 2006. Entropy, free energy, work, and other thermodynamic variables in economics. *Ecol. Econ*. 56: 71–78.
[28] Corning, P.A. and S.J. Kline. 1998. Thermodynamics, information and life revisited, part I: 'to be or entropy'. *Syst. Rese. Behav. Sci.* 15: 273–295.
[29] Horowitz, A.R. and Horowitz, I. The real and illusory virtues of entropy-based measures for business and economic analysis. *Decision Sciences* 7(1): 121–136, 1976.
[30] Salat, Serge, with the collaboration of Françoise Labbe and Caroline Nowacki. *Cities and Forms. On Sustainable Urbanism.* CSTB Urban Morphology Laboratory, Paris, 2011.
[31] Singh, V.P. Entropy Theory and its Application in Environmental and Water Engineering. John Wiley & Sons, Ltd, Chichester, UK, 2013.
[32] Lin, I.J. and Katyshev, A. Physical interpretation of the separation criteria. *Magnetic and Electrical Separation,* 7: 37–50, 1995.
[33] Lin, S.-K. Correlation of entropy with similarity and symmetry. *J. Chem. Inf. Comp. Sci.* 36: 367–376, 1996.
[34] Denbigh, K. *The Principles of Chemical Equilibrium.* Cambridge: Cambridge University Press, 1966
[35] Atkins, P. 1994. *The 2nd Law: Energy, Chaos, and Form*. Scientific American Books.
[36] Glueckauf, E. *Trans. Faraday Soc.*, 51: 34, 1955.
[37] Read, S. Meaning and Material: Phenomenology, Complexity, Science and "Adjacent Possible" Cities. In Portugali, J. Meyer, H. Stolk, E and Tan, E (eds.). *Complexity Theories of Cities Have Come of Age.* Heidelberg: Springer-Verlag, pp. 105–128, 2012.
[38] Norgaard, R.B. The case for methodological pluralism. *Ecological Economics,* 1: 37–57, 1989.

[39] Heylighen, F., Cilliers, P. and Gershenson, C., Complexity and Philosophy, in: Jan Bogg and Robert Geyer (eds.), *Complexity, Science and Society*, Radcliffe Publishing: Oxford, 2007.
[40] Nawrocki, D.N. and Harding, W.H. State-value weighted entropy as a measure of investment risk. *Applied Economics* 18: 411–419, 1986.
[41] Adami, C. The use of information theory in evolutionary biology. *Annals of the New York Academy of Sciences,* 1256: 49–65, 2012.
[42] Resilience Alliance. 2010. Assessing resilience in social-ecological systems: Workbook for practitioners. Version 2.0. Online: http://www.resalliance.org/3871.php
[43] Chau-kiu Cheung and Kwan-kwok Leung. Neighborhood homogeneity and cohesion in sustainable community development. *Habitat International* **35***:* 564–572, 2011.
[44] Seymoar, Nola-Kate, *Strengthening community resilience.* Paper written for the Province of British Columbia, Community, Aboriginal and Women's Services, with support from Western Economic Diversification Canada and the International Centre for Sustainable Cities, abridged version Oct 2007.
[45] Buck, N., Social cohesion in cities. In Buck, N, Gordon, I, Harding, A & Turok, I (eds.), *Changing cities: Rethinking urban competitiveness, cohesion and governance*. Palgrave MacMillan: New York (USA), 44–61, 2005.