MAXIMUM ENTROPY METHOD AND SEISMIC FREQUENCY – MAGNITUDE RELATION

L.H. Feng and G.Y. Luo

Department of Geography, Zhejiang Normal University, Jinhua 321004, China

Email: fenglh@zjnu.cn

ABSTRACT:

Entropy is a state function. The entropy increase principle tells us that under isolated or adiathermal conditions, the spontaneous development of a system from a state of non-equilibrium to a state of equilibrium is a process of entropy increase, in which a state of equilibrium corresponds to a state of maximum entropy. When the system is in a state of equilibrium, it is also at its most chaotic and disordered. The occurrence of earthquakes can be classified as a random event and can be described using entropy. Earthquakes occur in the most disordered way, indicating that entropy has reached its maximum value, so we can use the Maximum Entropy Method to determine the distribution of earthquakes that occur within a certain area curing a particular period of time. Results show that the formula representing the relationship between seismic frequency and magnitude (based on data and experience) is in fact a negative exponential distribution under given restraints and supposing seismic entropy is set as the maximum value. Therefore, we can theoretically explain the origin of the relationship between seismic frequency and magnitude.

KEYWORDS: Maximum Entropy Method, disorder, distribution, relationship between seismic frequency and magnitude

1. INTRODUCTION

In 1854, Clausius referred to the ratio of heat absorbed by working substances to temperature in reversible processes as “entropy”. Thereafter, during a period of over 100 years, the concept of entropy was widely applied in various scientific fields (Guo and Garland 2006; Emanuel 1995; Goltz and Bose 2002). As a state function, entropy has profound meaning: in thermodynamics, it is the measurement of unavailable energy; in statistical physics, it is the measurement of the number of microscopic states of a system; in informationism, it is the measurement of the uncertainty of random events (Wang 1988).

Based on research into seismic data, Gutenberg and Richter proposed the following Seismic Frequency – Magnitude Formula (Feng 1998):

$$\lg n = a - bM \tag{1}$$

In this equation, M is magnitude; n is seism times between M and $M+dM$, i.e. differential times (accumulated times n' is the times of seism with magnitude $\geqslant M$); a and b are undetermined parameters. Although equation (1) is a statistical equation calculated based on existing data and experience, it clearly demonstrates the distribution form (i.e. distribution function) of the seism frequency and magnitude. For a long time, however, there has been little research into the reasons why the relationship conforms to this function (Li *et al*. 1979). This article probes into the relationship between seismic frequency and magnitude on the basis of the Maximum Entropy Method (Akin 2006; Zhou and Gan 2006).

2. THE CONCEPT OF ENTROPY AND THE MAXIMUM ENTROPY METHOD

During the 18th century, research performed on the basic relationship between energy and substance by physicians led to the discovery of the Second Law of Thermodynamics. Clausius discovered that entropy (S) is a state function and can be used to present the Second Law:

$$dS \geq dQ / T \tag{2}$$

Where dQ represents primary heat and T is absolute temperature. This equation tells us that in a solitary system, the entropy of reversible processes remains unchanged while irreversible processes increase in entropy. As a result, all irreversible processes, including processes tending towards balance, increase in entropy.

In 1896, Boltzmann associated entropy with the accessible microstate number Ω, and endowed explicit statistical meaning to the entropy function, resulting in the creation of the Boltzmann Equation:

$$S = k \ln \Omega \tag{3}$$

In this equation, k is the Boltzmann constant. As indicated by the equation, the more the accessible microstate numbers, the higher the entropy value of the system; so entropy can be viewed as a quantitative description of the randomness (disorder degree) of molecular thermodynamic movement within the system. Boltzmann's elaboration of entropy offers a compromise between macroscopic value (S) and microscopic value (Ω), giving momentum to the development of thermodynamics and statistical physics.

In 1948, Shannon introduced Boltzmann's entropy concept into information theory, using entropy to quantitatively describe the uncertainties inherent in information and establishing the foundation of modern information theory. Shannon gave a new definition to entropy from a different point of view:

$$H = -C \sum_{i=1}^{n} p_i \ln p_i \tag{4}$$

The value H in this formula is referred to as information entropy; p_i is the probability of the occurrence of random events, and C is a constant. Equation (4) shows the relationship between information and entropy and solves the problem of how to quantitatively describe information. It also expands the concept of entropy, which has evolved greatly since being solely a concept of thermodynamics, now combining elements of the fields of statistical physics and information theory and gradually penetrating into many aspects of the natural and even social sciences (Nanda and Paul 2006; Nicholson *et al*. 2000).

It can be concluded from equation (4) that entropy H is a function of p_i. Thus, under given experimental conditions, there exists a distribution under which H is at its maximum value. This is the most common distribution and has dominant probability. This distribution is known as the "Most Probable Distribution". To sum up, the Maximum Entropy Method involves choosing, under given restrictions, the distribution under which entropy is the maximum possible value in all compatible distributions.

Based on the Maximum Entropy Method and by applying the Lagrangian Undetermined Multiplicator Method, we can calculate the distribution at maximum entropy (Zhang and Ma 1992). Let random variable x be $x_1, x_2, \cdots, x_n$, with corresponding probabilities are $p_1, p_2, \cdots, p_n$. When $C = 1$, the Entropy Formula (4) of the discrete random function can be simplified as:

$$H = -\sum_{i=1}^{n} p_i \ln p_i \tag{5}$$

This equation requires that:

$$\sum_{i=1}^{n} p_i = 1 \qquad p_i \geq 0 \tag{6}$$

Now we consider the issue from the opposite side. If the probability p_i is an unknown value, and we know that entropy H has reached its maximum value, we must consider the requirements that p_i is required to meet. The answer to this question is related to specific restrictive conditions; different conditions result in different answers.

First of all, we know that the value of H is restricted by the function $\sum_{i=1}^{n} p_i = 1$, i.e. only when the sum of all probabilities is equal to 1 can H reach its maximum value. Secondly, another restrictive condition is often that the average value of a certain function $f(x)$ of variable x is known. If there are m restrictive conditions ($m < n$), where all restrictive conditions are known functions of x, $f_1(x), f_2(x), \cdots, f_m(x)$, and all have the pre-set average values $F_1, F_2, \cdots, F_m$, i.e.:

$$F_k = \sum_{i=1}^{n} f_k(x_i) p_i \qquad k = 1,2,\cdots,m (m < n) \tag{7}$$

Then the problem becomes finding, under the restrictive conditions (6) and (7), at which values of p_i does the entropy H in (5) reach its maximum value. To answer this question and by applying the Lagrange Multiplier Method, we can introduce Multipliers α and β_k to form a new function, i.e. there is a linear relation between H and constants α, β_k and F_k:

$$H-\alpha-\beta_1F_1-\beta_2F_2-\cdots-\beta_mF_m$$

From equations (5), (6) and (7), we can see that:

$$H-\alpha-\sum_{k=1}^{m}\beta_kF_k=-\sum_{i=1}^{n}p_i\ln p_i-\alpha\sum_{i=1}^{n}p_i-\sum_{k=1}^{m}\beta_k\sum_{i=1}^{n}f_k(x_i)p_i$$

$$=\sum_{i=1}^{n}p_i\ln\left\{\frac{1}{p_i}\exp\left[-\alpha-\sum_{k=1}^{m}\beta_kf_k(x_i)\right]\right\}$$

By use of the inequation $\ln x\le x-1$, the above becomes:

$$H\le\sum_{i=1}^{n}p_i\left\{\frac{1}{p_i}\exp\left[-\alpha-\sum_{k=1}^{m}\beta_kf_k(x_i)\right]-1\right\}+\alpha+\sum_{k=1}^{m}\beta_kF_k$$

To find the maximum value of *H*, the above formula must be converted to an equation, then p_i must satisfy:

$$p_i=\exp\left[-\alpha-\sum_{k=1}^{m}\beta_kf_k(x_i)\right]\qquad i=1,2,\cdots,n \tag{8}$$

Using equations (6) and (8), we can re-write the equation as:

$$\alpha=\ln\left\{\sum_{i=1}^{n}\exp\left[-\sum_{k=1}^{m}\beta_kf_k(x_i)\right]\right\}$$

Let $Z=e^{\alpha}$:

$$Z=\sum_{i=1}^{n}\exp\left[-\sum_{k=1}^{m}\beta_kf_k(x_i)\right] \tag{9}$$

Z is called the Partition Function. Thus equation (8) becomes:

$$p_i=\left\{\exp\left[-\sum_{k=1}^{m}\beta_kf_k(x_i)\right]\right\}\Big/Z \tag{10}$$

In order to calculate the value β_k, we substitute (10) into constraint equation (7) as follows:

$$F_k=\sum_{i=1}^{n}\left\{f_k(x_i)\exp\left[-\sum_{k=1}^{m}\beta_kf_k(x_i)\right]\right\}\Big/Z \tag{11}$$

In equation (11), both F_k and $f_k(x_i)$ are known, while the *m* β values $(\beta_1,\beta_2,\cdots,\beta_m)$ are unknown. *M* equations could get *m* β values, thus we can calculate the value of p_i when entropy is at its maximum value. The above formula used in calculating discrete conditions can also be used in calculations involving continuous conditions.

3. THE SEISMIC FREQUENCY – MAGNITUDE RELATIONSHIP OBTAINED USING THE MAXIMUM ENTROPY METHOD

The principle of entropy increase shows that, under isolated or adiathermal conditions, the spontaneous development of a system from a state of non-equilibrium to a state of equilibrium is a process of entropy increase, in which a state of equilibrium corresponds to the maximum entropy. When the system is in a state of equilibrium, it is also at its most chaotic and disordered.

The occurrence of earthquakes can be classified as a random event and can be described using entropy. Earthquakes occur in the most disordered way, indicating that entropy has reached the maximum value, so we could use the Maximum Entropy Method to determine the distribution of earthquakes in a certain area within a certain period of time.

Let M_0 be threshold magnitude and $\overline{M}$ mean magnitude, then equation (6) becomes: $\int_{M_0}^{\infty} p(M)dM = 1$, and equation (7) becomes a relation, i.e. m=1.

$$\overline{M} = \int_{M_0}^{\infty} Mp(M)dM \tag{12}$$

We can now calculate partition function $Z = [\exp(-\beta M_0)]/\beta$ and substitute the result into equation (10), so

$$p(M) = \beta \exp[-\beta(M - M_0)] \tag{13}$$

By use of (12), we can calculate that $\beta = 1/(\overline{M} - M_0)$. In this equation, $p(M) = n/N$ (N is the total number of earthquakes). We can then substitute these values as follows (13):

$$n = \frac{N}{\overline{M} - M_0} \exp\left(-\frac{M - M_0}{\overline{M} - M_0} \right) \tag{14}$$

We know that magnitude complies with the negative exponential distribution (Figure 1). If we use logarithms to present equation (3), and calculate the logarithm of (14), we see that:

$$\lg n = \lg \frac{N}{\overline{M} - M_0} + \frac{M_0}{\overline{M} - M_0} - \frac{1}{\overline{M} - M_0} M \tag{15}$$

Through comparison with (1), we see that:

$$\begin{cases} a = \lg \dfrac{N}{\overline{M} - M_0} + \dfrac{M_0}{\overline{M} - M_0} \\ b = \dfrac{1}{\overline{M} - M_0} \end{cases} \tag{16}$$

This shows that the relationship between seismic frequency and magnitude, as calculated based on data and experience, is in fact a negative exponential distribution under given restraints when seismic entropy is at its maximum value. Therefore, we can theoretically explain the origin of the relationship between seismic frequency and magnitude.

Figure 1 shows the relationship between seismic differential times and magnitudes in Ningxia and its surrounding area from 1991 to 2005 (Ningxia Seismological Bureau 2006). As we can see, the result is a negative exponential distribution. Figure 2 shows the relationship between seismic differential time and magnitude in the semilog coordinate system. In order to get the optimum combination of straight lines and points, the least-squares method and maximum likelihood method can be used to estimate regression coefficients *a* and *b*. The straight-line in Figure 2 is the best fitting line obtained through the least-squares method: $\lg n = 5.0573 - 0.8518M$, correlation coefficient *R*=0.98.

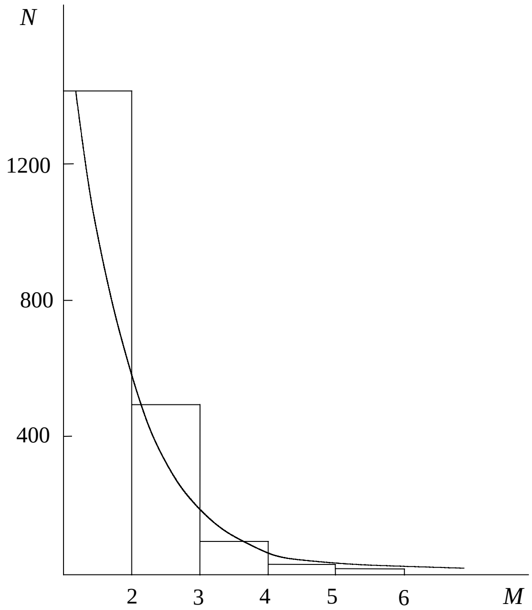

Figure 1 Relation between seismic differential times and magnitudes of Ningxia and its surrounding areas in 1991–2005

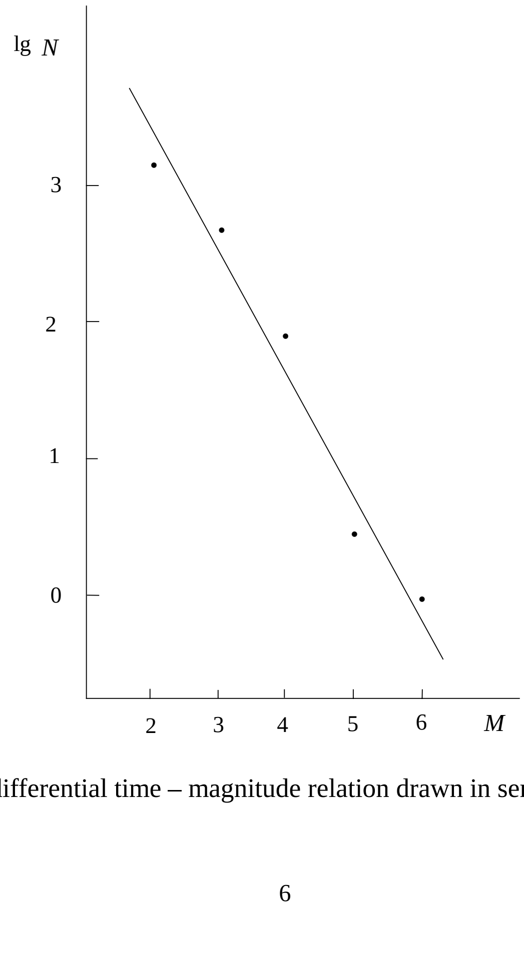

Figure 2 Seismic differential time – magnitude relation drawn in semilog coordinate system

4. CONCLUSION

The successful pervasion of entropy proves its tremendous momentum and wide scope of application. Earthquakes are random events characterized by a large degree of uncertainty, so entropy can play an important role in seismic description. Earthquakes themselves are endowed with rich information. The introduction of the concept of Entropy can be used to establish a new theoretical foundation for the measurement of the uncertainty inherent in seismic information and is helpful in seeking new practical methods. We can draw two conclusions based on the above analysis: (1) Earthquakes occur everywhere and in the most disordered way, indicating that seismic entropy has reached its maximum value. We can determine the distribution of magnitudes in certain areas within a certain period of time by use of the Maximum Entropy Method. (2) The formula representing the relationship between seismic frequency and magnitude developed by Gutenberg and Richter based on data and experience is in fact a negative exponential distribution under given restraints when seismic entropy has reached its maximum value.

In this article we have deduced the formula representing the relationship between seismic frequency and magnitude using the Maximum Entropy Method, provided a theoretic explanation of the process and further understanding of the formula.

ACKNOWLEDGMENTS

This work was supported by National Natural Science Foundation of China (No. 40771044).

REFERENCES

Akin, H. (2006). The topological entropy of nth iteration of an additive cellular automata. *Applied Mathematics and Computation* **174**: **2**, 1427-1437.

Emanuel, K.A. (1995). The behavior of a simple hurricane model using a convective scheme based on subcloud-layer entropy equilibrium. *Journal of the Atmospheric Sciences* **52**, 3960-3968.

Feng, L.H. (1998). Application test of matter element analysis in earthquake forecast. *Acta Seismologica Sinica* **11**: **6**, 745-750.

Goltz, C. and Bose, M. (2002). Configurational entropy of critical earthquake populations. *Geophysical Research Letters* **29**, 51-51.

Guo, L.F. and Garland, M. (2006). The use of entropy minimization for the solution of blind source separation problems in image analysis. *Pattern Recognition* **39**, 1066-1073.

Li, Q.L., Yu, L. and Hao, B.L. (1979). *Space-time scanning for relation of seismic frequency – magnitude*. Seismological Press, Beijing, pp1-22.

Nanda, A.K. and Paul, P. (2006). Some results on generalized residual entropy. *Information Sciences* **176**, 27-47.

Nicholson, T., Sambridge, M. and Gudmundsson, O. (2000). On entropy and clustering in earthquake hypocentre distributions. *Geophysical Journal International* **142**, 37-51.

Ningxia Seismological Bureau. (2006). *Seismological Directory in Ningxia*. People Press in Ningxia, Yinchuan, pp172-274.

Wang, B. (1988). Talking Entropy. *Entropy and Cross Science*. Meteorological Press, Beijing, pp19-22.

Zhang, X.W. and Ma, L. (1992). *Entropy Meteorology*. Meteorological Press, Beijing, pp114-125.

Zhou, S.M. and Gan, J.Q. (2006). Constructing accurate and parsimonious fuzzy models with distinguishable fuzzy sets based on an entropy measure. *Fuzzy Sets and Systems* **157**: **8**, 1057-1074.