Effects of Scale on Interpretation and Management of Sediment and Water Quality (Proceedings of a Boulder Symposium, July 1995). IAHS Publ. no. 226, 1995.

Deliberation on scale theory

LIU SUXIA
Department of Hydrology, Institute of Geography, Chinese Academy of Sciences, Beijing 100101, China

MO XINGGUO
Agro-ecosystem Experimental Station, Institute of Geography, Chinese Academy of Sciences, Beijing 100101, China

LIU CHANGMING
Shijiazhuang Institute of Agricultural Modernization, Chinese Academy of Sciences, PO Box 185, Shijiazhuang 050021, China

Abstract Experiments have demonstrated a lognormal distribution for scale parameters of soil water characteristics. Based on the distribution theory of random variables in statistics, some scale parameters not of lognormal distribution are identified in this paper. Representative Elementary Length (REL) is proposed to evaluate dominant processes across land scale units affected by different precipitation inputs. The implication is explored by synergetic theory. With distance between precipitation stations ranging from less than REL (threshold) to larger than REL, a length variable controls whether a catastrophe occurs and whether the system changes from dependence to independence. Semi-variance theory is used to calculate REL and is successfully applied to the design of a remote precipitation station network of the Yanghe Reservoir Basin of China.

INTRODUCTION

Scale theory was developed from dimensionlessness and was first proposed by Miller & Miller (1955) for applying the similarity principle to non-saturated soil water movement dynamics analysis. It was expected to be able to simulate field soil water movement in the laboratory. Although this expectation was only partially realized, it has found a wide application in studies of spatial variability.

Scale theory is based on the similar media concept of fluid dynamics of porous media. That is, if there is similarity between any two points within a study area in terms of volume and shape, and if the geometry distribution of the soil porosity and the soil moisture are the same, the two points are said to have similarity. All of the soil physics parameters can be expressed by a scale factor that represents the scale change of microscopic characteristics. In other words, the microscopic characteristics of one medium are equal to the microscopic characteristics of another medium multiplied by a constant, i.e., a scale factor. Although the soil moisture values are the same when there is geometric similarity, the characteristics of soil water movement are different at different points because the curvature radii are probably different.

Based on the interface concept (Liu, 1993), this paper deduces the scale method in a new way; some considerations on the distribution of scale parameters are given. A new scale concept of Representative Elementary Length (REL), which is well explained by synergetic theory, is proposed to evaluate the dominant processes across land scale units affected by different precipitation inputs. The concept of REL is applied successfully to the design of a remote precipitation station network in the Yanghe Reservoir Basin of China.

THE DEDUCTION OF SCALE METHOD

Based on interfacial theory (Liu, 1993), a new deduction of the scale method that does not need an arbitrary characteristic length (Nielson & Erh, 1973) to introduce the scale factor is proposed in this paper. The suction or matrix, ψ, is simplified as the capillary force at the interface between soil water and air inside the capillaries:

$$\psi = \frac{2\sigma}{r}\cos\theta \tag{1}$$

in which σ is surface tension, r is capillary radius, and θ is the touch angle between soil water and the wall of a soil capillary. Within the soil, σ and θ can be regarded as constants. From equation (1), at any two points or for any two similar media, i and j, there exists:

$$\psi_{mi} r_i = \psi_{mj} r_j \tag{2}$$

and then:

$$\psi_{mi} r_i = \psi_m r \tag{3}$$

in which ψ_m is the average of point matrices and r is the average of point radii. If $\alpha_i = r_i/r$, which is called the scale factor, then:

$$\psi_m = \alpha_i \psi_{mi} \tag{4}$$

The direct definition of α_i as r_i/r is the main difference between this method and other deductive methods. Within an area, the values of α_i at points are different from each other, which has been identified to have lognormal distribution by field data simulation using the Monte-Carlo Method (Luxmoore & Sharma, 1980).

THE DISTRIBUTION OF SCALE PARAMETERS

Based on scale theory, some scale parameters are formulated as:

$$K_i = \alpha_i^2 \hat{K} \tag{5}$$

$$f_i = \alpha_i^2 \hat{f} \tag{6}$$

$$S_i = \alpha_i^{1/2} \hat{S} \tag{7}$$

$$A_i = \alpha_i^2 \hat{A} \tag{8}$$

in which K, f, S and A are hydraulic conductance, infiltration rate, and the two parameters of the Philip infiltration formula, respectively. According to Nielson, *et al.* (1986), Smith & Hebbert (1979) and Zhang & Cundy (1989), all of the above parameters also are of lognormal distribution based on their experiments. A lognormal distribution, however, can not be identified by the distribution principle of the function of random variables of statistics. Specifically, if a random variable, β, is of normal distribution, its probability density function (pdf) is:

$$f_\beta(y) = \frac{1}{\sigma_y\sqrt{2\pi}}\exp(-(y-\mu_y)^2/(2\sigma_y^2)) \qquad -\infty < y < \infty \tag{9}$$

in which σ is the mean square deviance, μ_y is the mean and y is the value taken for β. As the scale factor, α, is of lognormal distribution, the probability density function of α is:

$$f_\alpha(x) = \frac{1}{x\sigma_y\sqrt{2\pi}}\exp(-(\ln x-\mu_y)^2/(2\sigma_y^2)) \qquad 0 < x < \infty \tag{10}$$

in which x is the value taken for α. To identify the distribution of the above scale parameters, two other random variables are introduced as $\xi = C_\xi\alpha^2$ and $\eta = C_\eta\alpha^{1/2}$; C_ξ and C_η are constants. It is assumed that z_ξ is the value taken for ξ. From $z_\xi = C_\xi x^2$, we get $x = g_\xi^{-1}(z_\xi) = \pm(z_\xi/C_\xi)^{1/2}$. Here g_ξ^{-1} is the opposite function. Because x is non-negative, we take its positive value. That is:

$$x = g_\xi^{-1}(z_\xi) = \sqrt{z_\xi/C_\xi} \tag{11}$$

Based on the distribution principle of the function of random variables of statistics (East China Technical University of Water Resources, 1980), there is:

$$\begin{aligned} f_\xi(z_\xi) &= f_\alpha[g_\xi^{-1}(z_\xi)]\left|\frac{d}{dz_\xi}g_\xi^{-1}(Z_\xi)\right| = f_\alpha(\sqrt{z_\xi/C_\xi})\left|\frac{d}{dz_\xi}\sqrt{z_\xi/C_\xi}\right| \\ &= \frac{1}{\sqrt{z_\xi/c_\xi}\,\sigma_y\sqrt{2\pi}}\exp\{-[½(\ln z_\xi - \ln C_\xi)-\mu_y]^2/(2\sigma_y^2)\}\cdot ½\sqrt{z_\xi/C_\xi} \\ &= \frac{1}{2\sigma_y\sqrt{2\pi}}\exp\{-[\ln(\sqrt{z_\xi/C_\xi})-\mu_y]^2/(2\sigma_y^2)\} \end{aligned} \tag{12}$$

Equation (12) is not a pdf of a lognormal distribution, and it is seen from this pdf that hydraulic conductance, K, infiltration rate, f, and the parameter, A, are not all of lognormal distribution. From $z_\eta = C_\eta x^{1/2}$, we get $x = (z_\eta/C_\eta)^2 = g_\eta^{-1}(z_\eta)$. Then:

$$\begin{aligned} f_\eta(z_\eta) &= f_\alpha[g_\eta^{-1}(z_\eta)]\left|\frac{d}{dz_\eta}g_\eta^{-1}(z_\eta)\right| \\ &= f_\alpha[(z_\eta/C_\eta)^2]\left|\frac{d}{dz_\eta}[(z_\eta/C_\eta)^2]\right| \\ &= \frac{2}{z_\eta\sigma_y\sqrt{2\pi}}\exp\{-[\ln(z_\eta)^2-\ln(C_\eta)^2-\mu_y]^2/(2\sigma_y^2)]\} \end{aligned} \tag{13}$$

Equation (13) also is not a pdf of lognormal distribution. Therefore, parameter S in the Philip infiltration formula also is not of lognormal distribution.

Quite a few field research results (Nielson, *et al.*, 1986; Smith & Hebbert, 1979; Zhang & Cundy, 1989) have shown that the distribution of all the above parameters is as the same as that of scale factor α, i. e., of lognormal distribution. What we can get from the contradiction is that the scale theory is waiting for further study.

Considering the contradiction, we propose a new concept of Representative Elementary Length, REL, to solve a real case scale problem in the Yanghe Reservoir Basin. REL is not a substitute for the former scale concepts, but is simple and easy to use.

REPRESENTATIVE ELEMENTARY LENGTH (REL)

Concept

The problem of how to get average values from point data is often faced by us and we often divide a study area into sub-areas according to points in which data are available. With the points increasing, the distance, D, between the points decreases. With the number of divided sub-areas increasing gradually, a sub-area ultimately reaches a point where differences within the sub-area become meaningless and a scale exists that we define as Representative Elementary Length (REL) (Liu, 1993). If the sub-area scale is smaller than REL, the spatial variability within the sub-area is evident and it is impractical to get the average using the weight for point data. If the scale is larger than REL, the representation of point data for average is questionable.

Wood *et al.* (1988) proposed the concept of REA (Representative Elementary Area). They found that if the area of sub-basins is less than 1 km^2, the relation between precipitation and runoff is strongly affected by topography, soil and spatial variability of rainfall intensity. If the areas of sub-basins are larger than 1 km^2, however, classic statistics can be used to study the spatial variability of the areas, and the basin responses can be simulated by simple models. In fluid dynamics of porous media, the concept of REV (Representative Elementary Volume) is often used to deduce macroscopic equations from microscopic equations by averaging (Baveye & Sposito, 1984). The meaning of REL is similar to REA or REV, all of which have a synergetic implication. REL is the most explicit and can be calculated by semivariance.

The synergetic explanation

Synergetics was proposed by Haken (1977). According to this theory, when a natural process is continuing and further progress is indeterminant, it always proceeds between some collective motions. There always exists spontaneous and irregular independence of motion in sub-systems, and the motion of one sub-system is synergetically related by other sub-systems. With the change of control variables, *CV* (here, the distance between points or the areas of sub-basins), there is always a threshold (here, REA, REV, or REL). If a system approaches this threshold, independence becomes weaker and correlation becomes stronger. At the threshold a catastrophe occurs from disorder to order, or *vice versa*.

Calculation by semivariance

REL can be calculated by semivariance theory. In any two points of space, x and $x(D)$ (D is distance between the points), if the values of some characteristics of the points (precipitation, in our following real case) are $Z(x)$ and $Z(x + D)$, the semivariance function, $\gamma(D)$ (Journel & Huijbregts, 1978), reads:

$$\gamma(D) = (\tfrac{1}{2})var[Z(x) - Z(x+D)] \tag{14}$$

in which var$[Z(x) - Z(x + D)]$ is the variance of the difference between $Z(x)$ and $Z(x + D)$. It can be estimated by:

$$\gamma^*(D) = 1/[2N(D)]\sum_{i=1}^{N(D)}[Z(x_i) - Z(x_i + D)]^2 \tag{15}$$

in which $\gamma^*(D)$ is the estimated semivariance, and $N(D)$ is the pair number decided by distance D, assuming that the random function, $Z(x)$, is stationary. The relation between the autocorrelation coefficient, R, and semivariance, γ, is:

$$\gamma(D) = \sigma^2[1 - R(D)] \tag{16}$$

where $R = 0$, $\gamma(D_{max}) = \sigma^2$, and D_{max} is REL.

Case study

A real case study involves a remote precipitation station network design of the Yanghe Reservoir Basin, Hebei Province, China. As the cost of building a remote precipitation station is high, it is better to choose as few remote rainfall stations as possible. An inadequate network of stations, however, gives bad representation and therefore an optimum choice was generated by REL. Precipitation data from 1968 through 1988 were available for 11 stations within the basin. Using stationary analysis (Liu, 1993), the spatial distribution of precipitation in the basin was shown to be stationary. Dividing the precipitation data into two groups, $Z(x_i)$ and $Z(x_i + D)$, and applying equation (15), we developed a relation between γ and D (Table 1; Fig. 1). The semivariance model was simulated as:

$$\gamma(D) = 13.91D^{1.355} \qquad \text{(the unit of } D \text{ is 5.17 km)} \tag{17}$$

from which REL, i.e., D_{max}, was calculated as 21.4 km based on equation (16). It is seen from Fig. 1 that there is a threshold between 20 and 30 km.

Notwithstanding the sample number of 11 being not enough for statistical analysis, the REL calculated here is meaningful to network design of remote precipitation stations

Table 1 Relation between semivariance function, $\gamma(D)$, and distance, D, between stations.

D(5.17 km)	1	2	3	4	5	6	7
$\gamma(D)$	13.65	37.75	64.16	84.23	100.43	208.8	179.9

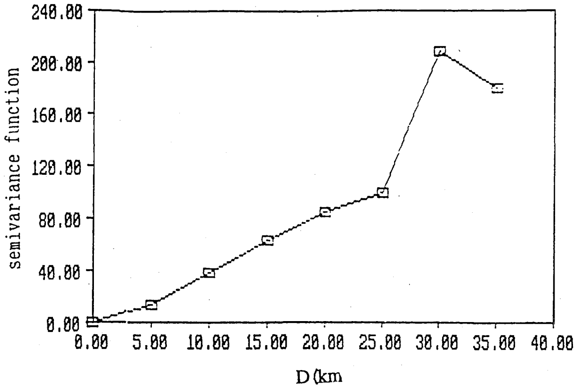

Fig. 1 Graph showing the semivariance function of precipitation, Yanghe Reservoir Basin.

of this basin. Based on the REL and using step-by-step regression analysis, principle components analysis, grey system theory and an isohyetal map, 8 of 11 precipitation stations were chosen (Fig. 2). The results were adopted by the project *Diverting Qinglong River Water to Tide Over Water Shortage* of Qinhuangdao City.

CONCLUSIONS

Considerations on scale theory are given. Scale parameters of hydraulic conductance,

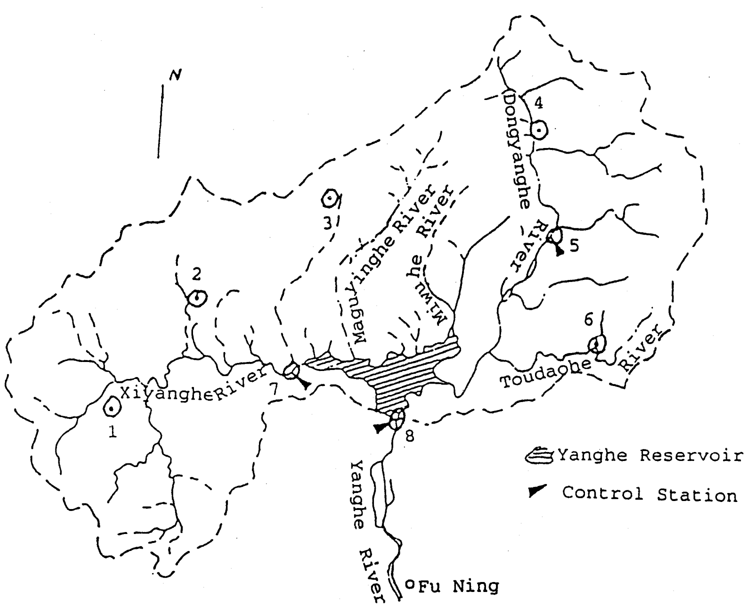

Fig. 2 Map of the Yanghe Reservoir Basin with the eight sites chosen for remote precipitation stations.

infiltration rate, and Philip infiltration parameters are shown to be not of lognormal distribution, which is opposite to earlier field research results.

A new scale concept, Representative Elementary Length, is proposed. Its synergetic implication with other similar concepts of REA and REV is explored. It is easily calculated by semivariance theory and was successfully applied to network design of remote precipitation stations in the Yanghe Reservoir Basin of the Peoples Republic of China.

Acknowledgement The work was partially supported by the Chinese Natural Science Foundation and the Foundation of the project of Diverting Qinglong River Water to Tide Over Water Shortage of Qinhuangdao City.

REFERENCES

Baveye, P. & Sposito, G. (1984) The operational significance of the continuum hypothesis in the theory of water movement through soils and aquifers. *Wat. Resour. Res.* **20**, 521-530.

East China Technical University of Water Resources (1980) *Statistics and Probability in Hydrology* (in Chinese). Hydroelectric Press, Beijing.

Haken, H. (1977) *Synergetics*. Springer-Verlag, New York.

Journel, A. G. & Huijbregts, J. C. (1978) *Mining Geostatistics*. Academic Press, San Diego, California.

Liu Suxia (1993) Mass transfer on and across the groundwater/surface water interface (in Chinese). PhD Thesis, Institute of Geography, Chinese Academy of Sciences, Beijing.

Luxmoore, R. J. & Sharma, M. L. (1980) Runoff response to soil heterogeneity: experiment and simulation comparison for two contrasting watersheds. *Wat. Resour. Res.* **16**, 675-684.

Miller, E. E. & Miller, R. D. (1956) Physical theory for capillary flow phenomena. *J. Appl. Phys.* **27**, 324-332.

Nielson, D. R. & Erh, K. T. (1973) Spatial variability of field-measured soil water properties. *Hilgardia* **42**, 215-259.

Nielson, D. R., van Genuchten, M. Th. & Bigger, J. W. (1986) Water flow and solute transport processes in the unsaturated zone. *Wat. Resour. Res.* **22**(9), 89-108.

Smith, R. E. & Hebbert, R. H. B. (1979) A Monte Carlo analysis of the hydrologic effects of spatial variability of infiltration. *Wat. Resour. Res.* **15**, 419-429.

Wood, E. F., Sivapalan, M. & Beven, K. (1990) Similarity and scale in catchment storm response. *Rev. Geophys.* **28**(1).

Zhang, W. & Cundy, T. Y. (1989) Modelling of two dimensional overland flow. *Wat. Resour. Res.* **25**(9), 2019-2035.