Journal of Environmental Sciences Vol. 13, No. 2, pp. 247—251, 2001

ISSN 1001—0742
CN11—2629/X

Article ID: 1001-0742(2001)02-0247-06 **CLC number**: Q142.2;X16 **Document code**: A

Spatial distribution analysis on climatic variables in northeast China*

SHANG Zong-Bo, GAO Qiong, YANG Dian-An, YANG Zheng-Yu

(Laboratory of Quantitative Vegetation Ecology, Institute of Botany, The Chinese Academy of Sciences, Beijing 100093, China. E-mail: shangzb@sohu.com)

Abstract: Information ecology is a new research area of modern ecology. Here describes the spatial distribution analysis methods of four sorts of climatic variables, i.e. temperature, precipitation, relative humidity and sunshine fraction in Northeast China. First, digital terrain models was built with large-scale maps and vector data. Then trend surface analysis and interpolation method were used to analyze the spatial distribution of these four kinds of climatic variables at three temporal scale: (1) monthly data; (2) mean monthly data of thirty years, and (3) mean annual data of thirty years. Ecological information system were used for graphics analysis on the spatial distribution of these climatic variables.

Key words: spatial distribution analysis; climatic variables; trend surface analysis; information ecology

Introduction

We are at the dawn of a new informational era. The advance of computer technology, the progress of informational theories, the improvements of research methods and experimental techniques, and the urgent needs for collection and treatments of ecological and environmental information have led us into this information ecology era. In this new era, it has been made necessary for using remote sensing (RS), global positioning systems (GPS), and geographic information system (GIS) to collect and deal with the information, using quantitative methods to analyze the information, using mechanistic modellings to simulate the ecosystems, and using informational systems to make decision and management. Information ecology (IE)(Zhang, 1990a; 1990b; 1997; Gao, 1994) has become an important divarication of modern ecology, which focus on five major research areas: information analysis and treatments, quantitative methods, graphics analysis and treatments, ecological modellings, and ecology information system (EIS).

EIS is a microcomputer software developed by Laboratory of Quantitative Vegetation Ecology, the Chinese Academy of Sciences. EIS can finish major functions of traditional geographic information system (GIS; Bian, 1996), as database and data treatments, and graphic information analysis. EIS has its independence and superiority on ecological information analysis, quantitative methods, graphics analysis, simulation modelling, and decision-making. EIS has a description language, ecological information description language (EIDL), which reinforce the function of information analysis. EIS has been used in many studies, such as impact of global change on terrestrial ecosystem(Zhang, 1993a; 1990c; Zheng, 1997), impact of global change on Tibetan Plateau(Zhang, 1996), and Northeast China transect (NECT)(Zhang, 1993b; 1997c). Grid data were commonly used in EIS, for the convenience of quantitative analysis and simulation modelling. EIS will prove to be a powerful tool for modern ecology research, such as macroscopic ecology (landscape ecology, territory ecology, and global ecology), biodiversity study, sustainable development, and simulation modelling.

Here, the spatial analysis methods and the formation processing of climate data were described in detail, with trend surface analysis (TSA) method and interpolation method, and EIS was used to analyze the spatial distribution of Northeast China.

1 Methods for spatial distribution analysis

Topography factor is very important for macroscopic ecology study. With large-scale relief maps, EIS can be used to convert vector data into grid data. Digital elevation modelic(DEM; Ke, 1992)was built for grid data formations of topography factors, as mean elevation, slope gradient, slope aspect and shading limit. Climatic variables as radiation, temperature, precipitation, relative humidity, and sunshine fraction are common used in ecological study. Grid radiation data can be calculated with meteorological(climatic) equations (Wang, 1987; Gao, 1988), with the data of latitude, elevation, slope gradient, slope aspect, and shading limit. Other climatic data usually come from meteorological observation data. Since meteorological observation stations are scattered distantly in China, it is of

Foundation item: The National Natural Science Foundation of China(No. 39899370, No. 39770133, and No. 39725006)

great significance for grid data formation, and spatial distribution regulation analysis of the climatic variables. Interpolation method is widely used for interpolating the climatic variables, which has its benefits for simplicity and convenience, and vices for disability for analysis the spatial regulations of scattered observation data. TSA is a statistical model for describing the spatial (or spatial-temporal) distribution of certain factors. It uses space coordinates as independent variables, such as abscissa and ordinate for two-dimensional TSA, or longitude, latitude and elevation for three-dimensional TSA (Chorley, 1968; Gittins, 1968; Pan, 1984)). It can use space-time coordinates as independent variables, such as abscissa, ordinate and time for three-dimensional TSA too.

1.1 Theoretical consideration

In a large-scale, the spatial distribution of climatic variables are impacted by many environmental factors, which can be summed up into three types by their scale and impact degree (Wang, 1987): (1) large-scale topography factors, as latitude, distance from sea, influence of great mountains, plateaus or basins, or atmospheric circulation backgrounds; (2) medium and small-scale topography factors, such as slope gradient, slope aspect, hypsography, and shading limit; (3) vegetative and microclimate factors. Here longitude, latitude and elevation are selected as large-scale topography factors, slope gradient and slope aspect are selected as medium and small-scale topography factors, and vegetation is selected as vegetative and microclimatic factors.

But spatial-temporal distribution regulations of climatic variables have their non-determinacy, for the regulations will always be fluctuated by random environment. It is found that in a certain bound, with the short of temporal scale, the functions of random events will "cover up" the spatial-temporal regulations greatly. So the spatial-temporal distribution of climatic variables can be described as

$$Y_t = q_t\{\varphi,\lambda,H,\alpha,\beta,V\cdots\} + \sigma(t), \tag{1}$$

where Y_t is the spatial-temporal distribution of climatic variables, t is the time, φ is the latitude, λ is the longitude, H is the elevation, α is the slope gradient, β is the slope aspect, V is the vegetative factor, and σ is the function of random events.

To decrease the function of random events, we integrated Y_t by time and got

$$Y_p = \frac{\int_{t_1}^{t_2} q_t\{\varphi,\lambda,H,\alpha,\beta,V\cdots\}\mathrm{d}t}{t_2 - t_1} + \frac{\int_{t_1}^{t_2}\sigma \mathrm{d}t}{t_2 - t_1}, \tag{2}$$

where Y_p is the average spatial distribution of climate variables in a certain period (t_1 to t_2), such as the mean monthly temperature of one year, or mean annual precipitation of thirty years. In a certain area, with the extension of the period, the functions of random events decrease sharply, and the accuracy of simulation increase greatly. So the Equation (2) can be converted into

$$Y_p = q_p\{\varphi,\lambda,H,\alpha,\beta,V\cdots\} + \sigma_p, \tag{3}$$

Latitude, longitude and elevation can be regarded as large-scale topography factors, and slope gradient, slope aspect and vegetative factor can be regarded as local environmental factors. Presuming that the functions of large-scale topography factors and local environmental factors can be separated from each other for some certain variables, such as temperature, precipitation, relative humidity and sunshine fraction, we can get

$$Y_p = f_p\{\varphi,\lambda,H\} + g_p\{\alpha,\beta,V\cdots\} + \sigma_p, \tag{4}$$

where f_p is the function of large-scale topography, and g_p is function of local environmental factors.

1.2 Trend surface analysis

Trend surface analysis (TSA)(Chorley, 1968; Gittins, 1968) is a special multivariate statistical analysis, which uses statistical model to describe the spatial(or spatial-temporal) distribution of variables. Here gives the major theory of TSA.

Supposing a dot in a three-dimensional space, with coordinate of (x_i, y_i, z_i), and observation data of w_i, we used Trend surface model (TSM) to simulate spatial distribution of such dots aggregate

$$W = GA + \varepsilon, \tag{5}$$

where W is the matrix of observation data, and ε is the residual matrix, both are vector-matrix of $N\times 1$; N is the number of observation data; G is the vector-matrix($N\times q$); A is the matrix of coefficients($q\times 1$), and q is determined by s, rank of TSA.

$$q = (s+1)(s+2)(s+3)/6. \tag{6}$$

Here gives D, one row vector of G, which has been omitted the subscripts for code of observation dot

$$\begin{aligned} D = \{&1,z,z^2\cdots,z^s,y,yz,yz^2\cdots yz^{s-1}\cdots,y^{s-1},y^{s-1}z,y^s,x,xz,xz^2\cdots,xz^{s-1},xy,xyz,\\ &\cdots xyz^{s-2}\cdots,xy^{s-1},x^2,x^2z\cdots,x^2z^{s-2},x^2y,x^2yz\cdots,x^2yz^{s-3}\cdots,x^{s-1}y,x^s\}. \end{aligned} \tag{7}$$

We use the method of generalized inverse matrix to calculate the TSM. With Gram-Schmidt' method of orthogonalization, we can calculate the Moore-Penros's generalized inverse matrix G^+ of real matrix G, and get the shortest least squares solution $\tilde{A}$ of A

$$\tilde{A} = G^{+} Z. \tag{8}$$

Then we can get the simulated values and the residuals, and the degree of accuracy can be expressed as

$$C = \left[1 - \frac{\sum_{i=1}^{N}(w_i - \hat{w}_i)^2}{\sum_{i=1}^{N}(w_i - \bar{w})^2}\right] \times 100\% , \tag{9}$$

where C is the degree of accuracy, w_i is the observed data, $\hat{w}_i$ is the simulated result, and $\bar{w}$ is the mean observed data of the territory.

1.3 Method for grid data formation with TSA and interpolation method

With TSA, we can get the function of large-scale topography easily, as we has built DEM and calculated latitude, longitude and the elevation for the three independent variables, and use climate variables (temperature, precipitation, relative humidity, and sunshine fraction) for dependent variables. All these data can be collected from the scattered meteorological observation stations. First of all, the independent variables should be converted in order to avoid overflowing (each independent variable should be divided by multiple of 10 to be less than 1). As we get the trend surface model, the simulated results and the residuals are available. Simulated results of each grid can be calculated as

$$\hat{Y}_{i,j} = D_{i,j}\tilde{A}_T. \tag{10}$$

where $\hat{Y}_{i,j}$ is the simulated result of climate variables, and $\tilde{A}_T$ is the coefficient vector.

The residuals can be regarded as the synthetical functions of local environmental factors and random events, which cannot be explained by large-scale topography analysis. Residuals are interpolated into each grid, to get fully information about spatial distribution of climate variables.

2 Data set and preprocessing

The following data sets were used in this research: (1) monthly meteorological data of 114 meteorological stations that cover Northeast China, including mean temperature, precipitation, relative humidity and sunshine fraction of thirty years (from 1951 to 1980); (2) monthly meteorological data of 134 meteorological stations that cover Northeast China, from 1981 to 1990, including mean temperature, precipitation, relative humidity and sunshine fraction of each month; (3) 1:500000 relief map. All the maps and data were provided by Laboratory of Quantitative Ecology.

Some computer programs were developed on microcomputer, with Visual C++ 5.0, to calculate the monthly and annual meteorological data, to calculate the DEM and TSM, to interpolate the residuals, and to form the grid climate data. EIS was used to convert the vector data into grid data, including elevation data, soil sorts data, and vegetation data. All the grid data were saved partitively into maps. Each map is 3° × 2° large(longitude × latitude), and covers 96 × 96 dots. EIS were used to graphics analyze the spatial distribution regulations of each environmental factor.

3 Results and discussion

With 1:500000 relief map, EIS were used to convert vector data into grid data. Then DEM were built to calculate topography grid data, including slope gradient, slope aspect, and shading limit.

Table 1 Simulated results of TSM for three scale and four climatic variables

Item	Scale	Period	Number of TSM	N	Average C, %	R^2
Mean temperature	Ⅰ	1981 – 1990	12 × 10	16063	77.807	0.9877
	Ⅱ	1951 – 1980	12	1368	96.445	0.9969
	Ⅲ	1951 – 1980	1	114	97.109	0.9711
Precipitation	Ⅰ	1981 – 1990	12 × 10	15275	50.868	0.8117
	Ⅱ	1951 – 1980	12	1350	83.943	0.9551
	Ⅲ	1951 – 1980	1	114	86.486	0.8675
Relative humidity	Ⅰ	1981 – 1990	12 × 10	16055	63.034	0.8235
	Ⅱ	1951 – 1980	12	1368	83.501	0.9053
	Ⅲ	1951 – 1980	1	114	85.345	0.8535
Sunshine fraction	Ⅰ	1981 – 1990	12 × 10	16074	55.719	0.7395
	Ⅱ	1951 – 1980	12	1368	78.384	0.8861
	Ⅲ	1951 – 1980	1	114	84.249	0.8426

Notes: Ⅰ, Ⅱ and Ⅲ are the three temporal scale: (1) monthly data every year, (2) monthly data of thirty years, and (3) annual data of thirty years. N is the total effective sampled data

Monthly meteorological data were used to build TSM at three temporal scales: (1) monthly data (from 1981 to

1990);(2) annual data of thirty years (from 1951 to 1980) and (3) monthly data of thirty years (from 1951 to 1980), for four climatic variables: (1) temperature; (2) precipitation; (3) relative humidity and (4) sunshine fraction. Form 1 shows the simulated results and the degree of accuracy of TSM. The simulated results of these four sorts of climatic variables are remarkable, and the TSM simulation results of the last temporal scale are the best ones.

With TSM and DEM, grid data of these four climatic variables were calculated for three temporal scale, and the residuals were interpolated into each grid to get the full information of each climatic element. Fig 1, 2, 3, and 4 show the average spatial distribution regulation of these four climatic variables for thirty years (from 1951 to 1980).

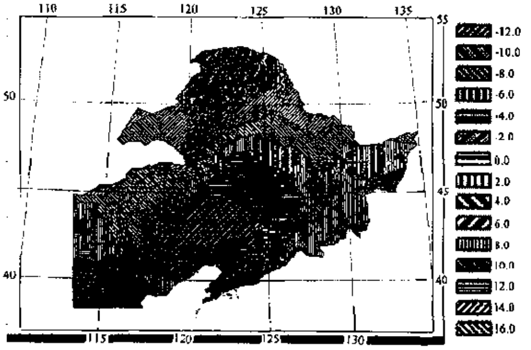

Fig. 1 Spatial distribution of mean annual temperature in Northeastern China

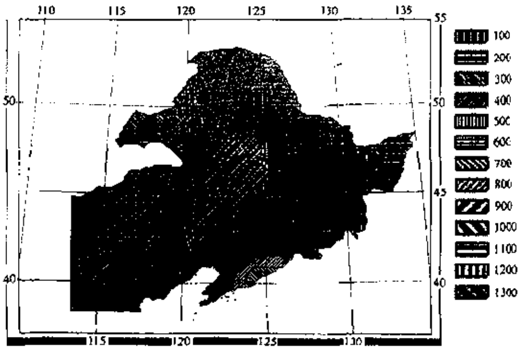

Fig. 2 Spatial distribution of mean annual precipitation in Northeastern China

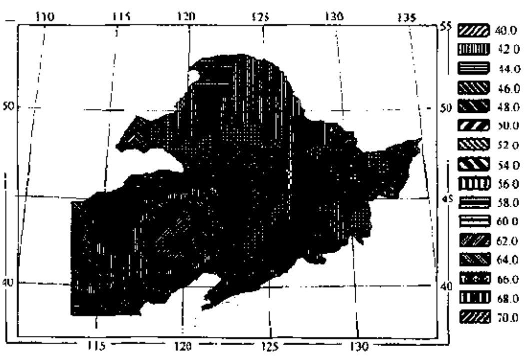

Fig. 3 Spatial distribution of mean annual relative humidity in Northeastern China

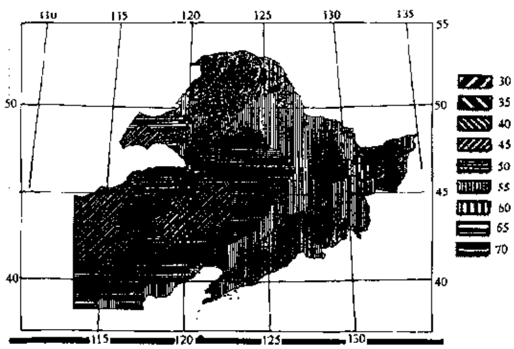

Fig. 4 Spatial distribution of mean annual sunshine fraction in Northeastern China

4 Conclusion

As a study of information ecology, we analyzed the spatial distribution of four types of climatic variables (temperature, precipitation, relative humidity, and sunshine fraction) and formed the database of Northeast China. EIS was used to convert the vector data into grid data with large-scale relief maps. DEM were built to analyze the topography characters. TSA is proved to be an excellent method for separating influence of large-scale topography and local environmental factors on spatial distribution of climatic factors. The simulated results of TSM of temperature, precipitation, relative humidity and sunshine fraction prove that the spatial distribution regulations of these four climatic variables are remarkable. TSA is an excellent method for separating the influence of large-scale topography and local environment on spatial distribution of climatic factors. TSM and interpolation method were used to form grid data of these four climatic variables. EIS is a powerful tool for modern ecology study, and it is used here to graphics analysis of these four types of ecological factors.

Acknowledgements: We would like to express our sincere appreciation to Laboratory of Quantitative Vegetation Ecology, which provided the climatic data, investigation data of soil sorts data and vegetation, relief map and grid elevation data.

References:

Bian F L, 1996. Theories and methods of geographic information system[M]. Beijing: Topographical Publishing House.

Chorley R J, Haggett P, 1968. Spatial analysis: A reader in statistical geography[M](Ed. by Berry B J L, Marble D F). New Jersey: Prentice-Hall Inc., Englewood Cliffs. 195—217.
Gao Q, 1994. Science and Technology Policy and Management[J]. 7:89—94.
Gao G D, Lu Y R, 1988. Climatology[M]. Beijing: Meteorological Publishing House.
Gittins R, 1968. J Ecology[J]. 56:845—869.
Ke Z Y, He J B, Chi T H, 1992. Digital terrain models[M]. Beijing: China Scientific and Technical Publishing House.
Pan E P, Zhou Y Q, 1984. Multivariate statistic programs collection (for DJS-21 computer)[M]. Beijing: Atomic Energy Publishing House.
Wang Y S, 1987. Atmospheric physics[M]. Beijing: Meteorological Publishing House.
Zhang X S, 1990a. Chinese Bulletin of Botany[J], 7(4): 1—6.
Zhang X S, 1990b. A new divarication of 90's Ecology-Information ecology bioscience information[J]. Bioscience Information, 2(3): 101—103.
Zhang X S, Gao Q *et al*., 1997a. Study on information ecology(Volume 1)[M]. Beijing: Science Press.
Zhang X S, Yang D A, 1993a. Journal of Environmental Sciences (China)[J], 5(2): 130—149.
Zhang X S, Yang D A, 1990c. Journal of Environmental Sciences (China)[J], 2(4): 95—109.
Zhang X S, Yang D A, 1996. Climate change and plants in East Asia(Ed. by Omasa K *et al*.)[M]. Springer-Verlag. 67—74.
Zhang X S, Yang D A, 1993b. China Quaternary Sciences[J], 2:157—169.
Zhang X S, Gao Q, Yang D A *et al*., 1997c. Acta Botanica Sinica[J], 39(9): 785—799.
Zheng Y R, Zhou G S, Zhang X S *et al*., 1997. Acta Botanica Sinica[J], 39(9): 837—840.

(Received for review March 6, 2000. Accepted October 18, 2000)