Acta Palaeobot. 31 (1,2): 289–294, 1991

PATTERN RECOGNITION AND MANY–DIMENSIONAL ANALYSIS FOR PALEOGEOGRAPHYCAL RECONSTRUCTIONS OF HOLOCENE

G. F. BUKREYEVA

Institute of Geology and Geophysics, Novosibirsk 630090, USSR

ABSTRACT. Three metods are described of quantitative estimation of paleoclimate elements, which are based on the search for the informative system of spore–pollen spectra components via purposive iteration classification, step–by–step regression analysis and method of main components.
It is suggested to determine the type of vegetation zone employing discriminant functions, taking in account spectrum informative components, preliminarily revealed by pattern recognition. The periodicity of paleoclimate and vegetation changes has been studied quantitatively by the methods of one–dimensional and many–dimensional analysis of palynological and climatic series of observations.

KEY WORDS: many dimensional analysis, paleoclimate, palynology, Holocene

There is a number of problems in palynology, which are solved by mathematical methods. There are, particularly, quantitative estimation of paleoclimate elements, determination of vegetation zone types and study of periodicity structure in climate and vegetation changes (Bukreyeva 1987).

The essence of the problem on quantitative estimation of paleoclimate is the folloving. There is a number of spore–pollen spectra (SPS) of surface samples of the region studied, which spectra reflect all the variety of presented three vegetation zones, and also the data of meteorological stations observations in the same region, presented as average quantitative observations of many years, that is: average monthly temperature, annual precipitation sum and so on. Each recent spore–pollen spectrum is corresponded to those quantitative features of modern climate elements, which are obtained by the meteorological station, the nearest to the place, where the surface samples were taken. The obtained connections (a linear regression model) between quantitative estimation of climate element and component contence in recent spore–pollen spectra may be used to calculate paleoclimate elements of Holocene and Pleistocene, i.e. of geological Past, which judging by the contence of fossil spore–pollen spectra, is characterized by the vegetation and, consequently, by the climate similar to the modern one.

Proceeding from the structure of factual data, which are the observation of great number of features (SPS components and climate elements), the most suitable mathematical method of their consideration is the method of many–dimensional statistical analysis complected with the pattern recognition.

The main problem as considering paleoclimate investigations, is the choice of SPS components informative system relative to the studied climate element in the whole complex considered, minimalization of their initial number being obligatory.

In the method of quantitative estimation of paleoclimate elements via informative systems of SPS components (Bukreyeva et al. 1984, 1986) the choice of optimum combination of spores and pollen types for including them into calculation equations which are the best to determine this or that climatic index, is presented by one of pattern recognition methods: purposive iterative classification. The informative system of SPS components is derived from the initial spore–pollen matrix via estimation of informative wieght of each SPS component relative to a concrete climate element. SPS components included into the system, are employed as arguments in calculating regression equation for quantitative estimation of the corresponding climatic element.

Construction of regression models of paleoclimate elements quantitative estimation employing the method of step–by–step regression analysis is realized as follows. Only those SPS components are also included in regression equation, which essentially influence the given climate element (Bukreyeva 1988). Here, though, this influence is estimated by the value of correlation partial coefficient. At each step that SPS component is introduced into regression equation, which presents the greatest partial correlation with the climate element under consideration.

For the quantitative estimation of paleoclimate elements by the method of main components (Bukreyeva 1988), the indication space is minimized proceeding from the supposition, that the correlation of indications means their redundancy, and reducing of redundant indications to that free of redundancy means diminishing of dimension of the initial indication space. That is realized by revealing of similar (double) indications. From each chosen group of indications only one may be left under consideration (without any loss of essential information). Formed in such a way set will consist of independent indications, whose number is essentially smaller than the initial one.

A comparative analysis has been done of estimations of July temperatures presented according to each of three methods using the training initial many–dimensional matrix, compiled of the data of Issyk–Kul hollow. It made it clear, that the some number of SPS components has been included in calculation models – four (of the total number 28). The models, constructed by step–by–step regression analysis, display a good approximation to factual data. But it is not doubtless, that they ensure obtaining of stable appropriateness. The models constructed by the method of main components are more stable comparing with the models of step–by–step regression, as the main components include the maximum variability, existing in the training excerpt. But in this case the informative system is searched irrespectively of the purpose (climate element) and therefore it is a set of essential SPS components for a given, closed by itself, palynological matrix. The model formed by the system of SPS informative components, chosen employing purposive iterative classification, distinguish by a stable appropriateness, which is due to the principle of indications choise according to the measure of their similarity and difference relative to the purpose – the studied climate element.

The problem of determination of vegetation zone type by the contence of fossil spore–pollen spectra is solved by the method of discriminant analysis. As the training excerpt, the same palynological matrix is used as in the previous problem, and as purposive indication – the vegetation zone type. At the first stage of mathematical consideration it is searched (with the help of pattern recognition program) the informative system of SPS components, including the existing differences of zones and similarity inside the zone. Having exposed the derived system to the discriminant analysis, one obtains a set of discriminant functions, which makes it possible to determine by the structure of fossil SPS the type of corresponding vegetation zone, with probability estimation of the derived results.

The method of periodicity quantitative estimation on the palynological data is based on the idea of fossilization in geological sections of climate and vegetation vibrational changes. The fossil spore–pollen spectra may be considered as their materialized consequence. They are a total result of interation of vibrational motions of different orders, which appear simultaneously and coinside (short-period cycles with long–period ones).

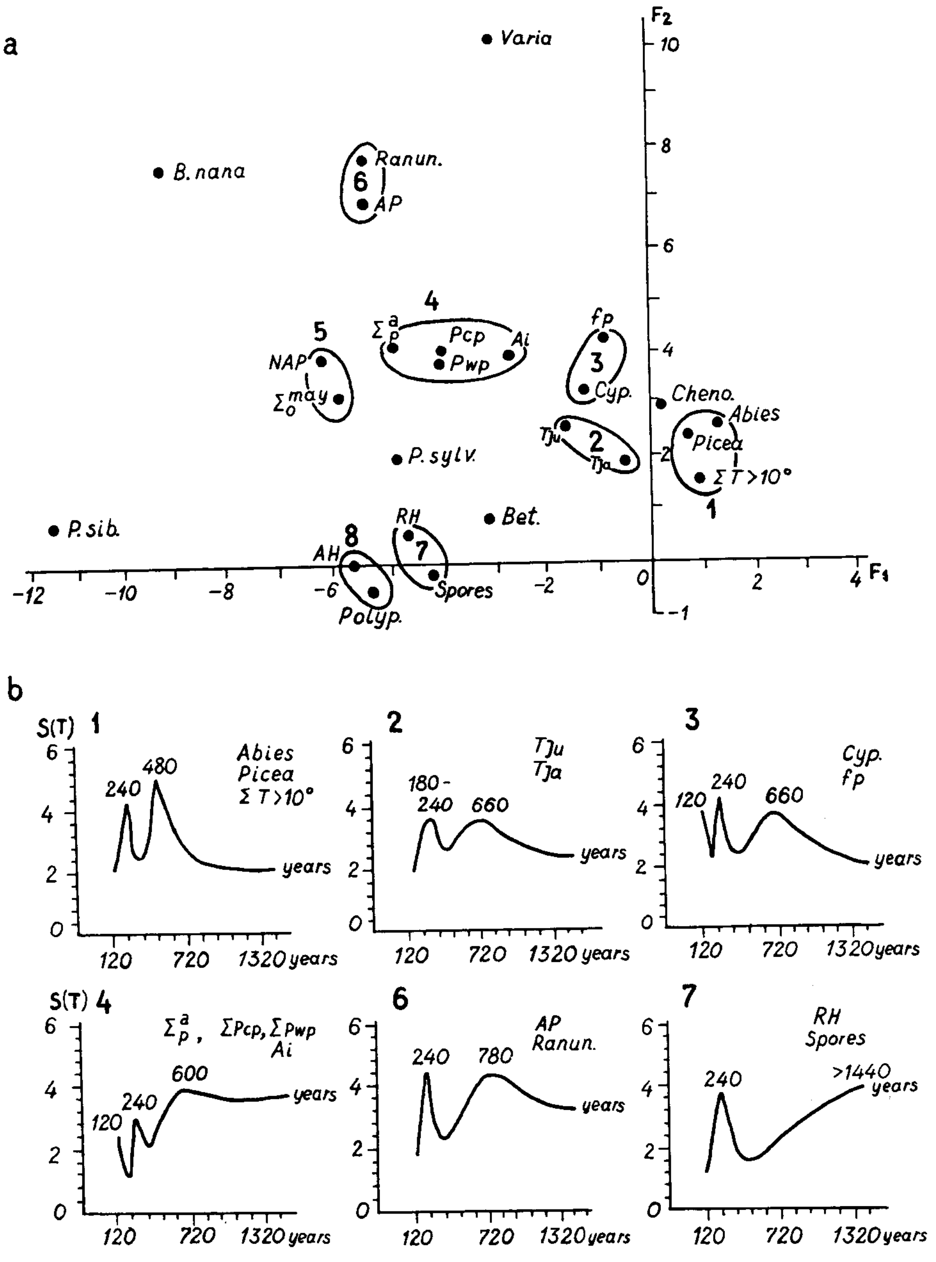
A harmonic model expressed in the form of Fourier series is taken as a dynamic model describing the process of climate and vegetation change. Every harmonic component of this model characterizes vibrational motion with its own unique period and amplitude. So, the solution of this problem reduces to the estimation of the parameters of the process dynamic model.

The study of periodicity over one time series is realized by one–dimensional, and over several ones – by many–dimensional spectral analysis.

As the initial data, discrete time series serve, which reflect the variation of SPS components structure over the vertical geological section and of quantitative estimations of paleoclimate elements. The optimum data can be obtained from the geological section without breaks of precipitation acumulation, rather completely and evenly over all logitude characterized by spore–pollen spectra and having determinations of the rocks absolute age. These requirements are frequently satisfied by peatbogs, lacustrine and marine deposits of Holocene age.

Many–dimensional spectra analysis includes the estimation of time series, as well as the calculations of mutual spectrum functions, coherency and difference of phases estimating the correlation of one series with another for each frequency. Its essence has been represented in details in the book (Jenkins & Watts 1972). It may be supplemented with the procedure of classification of the obtained time series spectra by the method of main components. It will display those SPS components and paleoclimate elements, which vary in time with the same periodicity.

Thus, the many–dimensional analysis of time series spectra, reflecting the change of 14 SPS components and 11 paleoclimat elements in the section of peatbog of Late Holocene age (Southern forest–steppe zone of Minusinsk hollow) showed that they are compounded into eight classes. It turned out, that the similar structure of periodicity



have: a) containing of spruce and *Abies* and the annual sum of temperatures above 10° C; b) total containing in spectrum of herbaceous plants pollen and precipitation sum in May; c) total containing of spores and relative humidity, etc. (Fig. 1a).

From the stucture of integral spectra, which are the generalized characteristics of

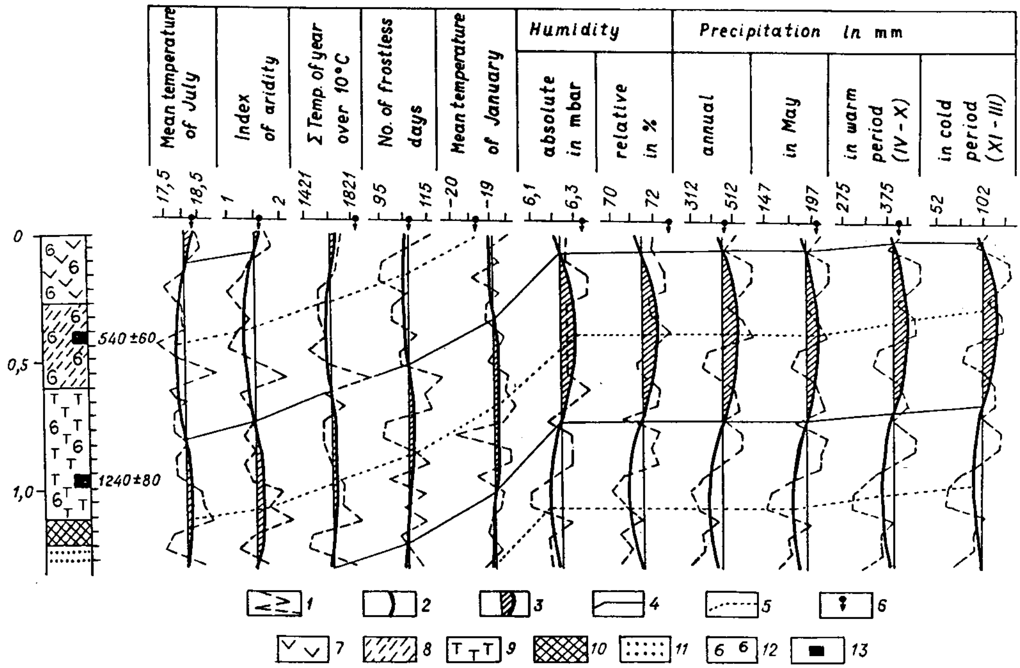


Fig. 2. Manifestation of tendency in paleoclimate elements variety of Late Holocene (basing on the date of peatbog Znamenskoye in Southern Minusinsk hollow): 1 – initial series; 2 – the graph of function of the main harmonics; 3 – period with maximum value of climatic index; 4 – the curve fixing the half-periods boundaries; 5 – the curve connecting maximum and minimum values of the main harmonic of climatic time series; 6 – recent quantitative value of climate elements; 7 – equisetum peat; 8 – hypnum peat; 9 – intercalation of sedge and equisetum peat; 10 – sandy loam with peat formation; 11 – sand; 12 – mollusk shells; 13 – the site of sampling for radiocarbon analysis.

classes, one may conclude, that each initial series contain one or two dominant periods (the total duration of the series being equal 1440 years). The periodicity have been displayed of 120, 240, 480, 600–660 and more than 1440 years (Fig. 1b).

Determination of harmonic component parameters allows one to find out that the dominant periods or their superpositions are timed to geochronological scale. It is

←

Fig. 1. Compiling of classes (1–8) with similar periodicty structure of the initial spectra of palynological and time series of peat deposits of Late Holocene age in the peatbog Znamenskoye of Southern Minusinsk hollow (a) and integral spectra of displayed classes (b): F1, F2 – coordinate axes of the first and second basic compònent; S(t) – spectral density of process; abbreviated component terms of spore–pollen spectra: AP, NAP, SP – the total content of aborescent pollen, herbaceous plants and spores in spectrum respectively; P. sib – *Pinus sibirica*, P. sylv – *Pinus sylvestris*, Bet – *Betula*, B. nana – *Betula nana*, Cheno. – *Chenopodiaceae*, Cyp. – *Cyperaceae*. Ranun.– *Ranunculaceae*, Varia – diff. herbage, Polyp. – *Polypodiaceae*; abbreviated names of climate elements: T_{Ju} – July temperatyre, T_{Ja} – January temperature, ΣT 10° – annual sum of temperature over 10°, fp – duration of frostless period, $\Sigma^a p$ – average annual sum of precipitates, Σpcp – amount of precipitation during cool period of the year (XI–III months), Σpwp – amount of precipitation during warm period of the year (IV–X months), $\Sigma^{May}{}_p$ – amount of precipitates in May, JA – index of aridity, AH – absolute humidity, RH – relative humidity

known, that the essential information is contained in the main harmonics with the period equal to the initial series duration. In the reconstructed climatic curves it revealed the tendency to the increase of precipitation fall, increase humidity and climate growing colder, which took place in the studied region 700 years ago. But 100 years ago the temperatures of July increased again, the quantity of precipitation decreased, aridity index increased (Fig. 2).

Amplitudes and initial phases of dominant harmonics make it also possible to write an equation for climate element prognose, basing on its variability during a long period of time, emounting to thousands of years.

REFERENCES

BUKREYEVA G. F., VOTAKH M. R. & BISHAEV A. A. 1984. The metods of paleoclimate determinations by informative systems of features of spore and pollen spectra. Geol. Geofiz., 6: 16–28 (in Russian).

—, — & — 1986. Determination of paleoclimats basing on palynological data by purposive iterative classification and regression analysis. Novosybirsk, Institute of Geology and Geophysics publishing, (in Russian).

BUKREYEVA G. F. 1987. Application of mathematical methods to the palynological investigations. Izv. Vsesojuzn. Geograf. Obshchestva, 119 (5): 427–433 (in Russian).

— 1988. Quantitative estimation of paleoclimate elements by method of main components and step–by–step regression analysis (on palynological data). Novosybirsk; Institute of Geology and Geophysics.

JENKINS G. & WATTS D. 1972. Spectral analysis and its application. Moscow, "Mir", 1972, issue 1. 316 p.; issue 2.287 p. (in Russian).