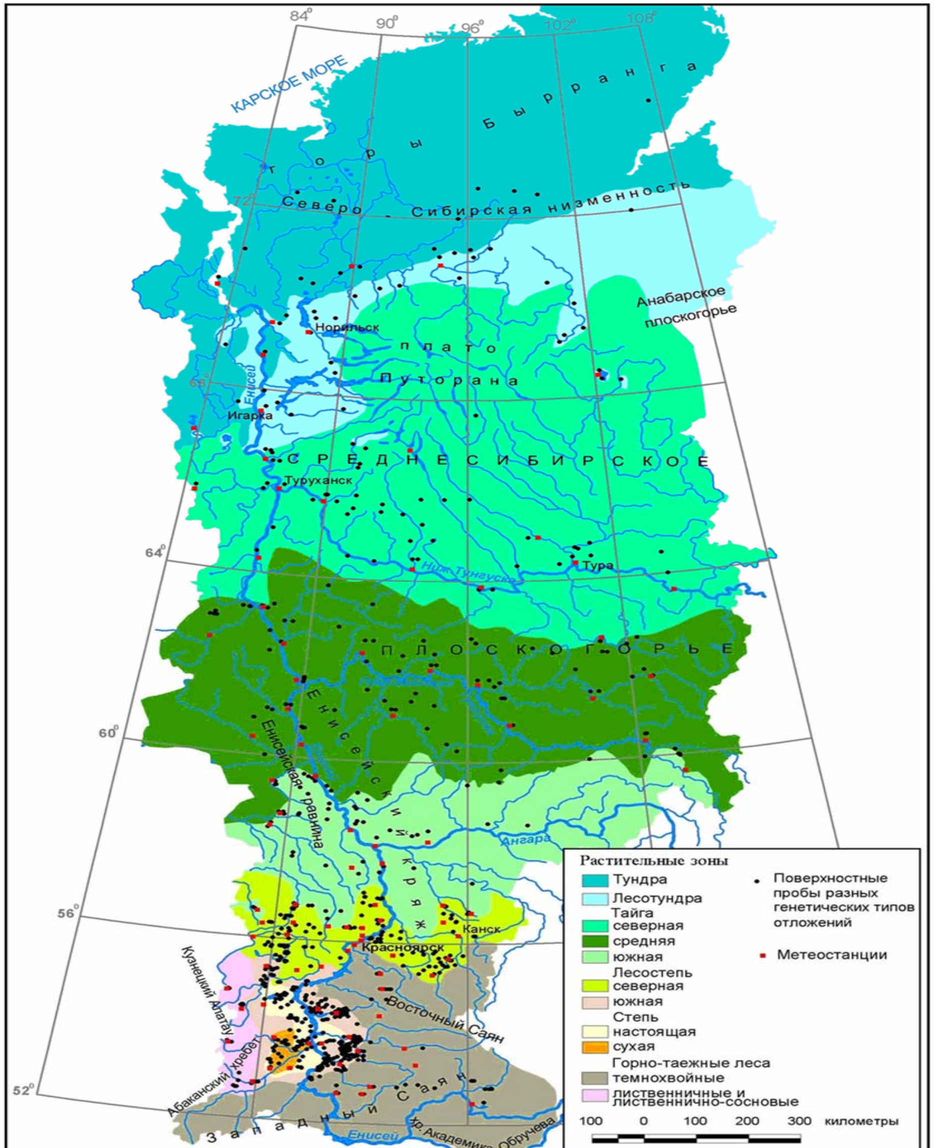
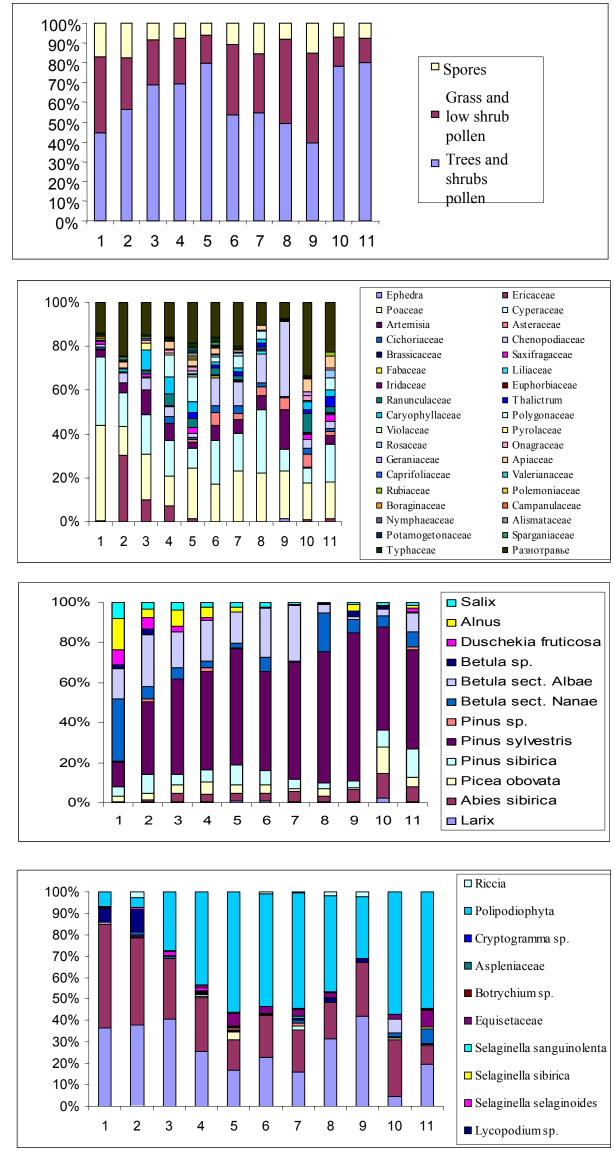
Using nonlinear regression models for reconstructing Holocene climate of Yenisey-close Siberia basing on palynological data

Yamskikh G.Yu. , Shchemel A.L.***

**Siberian Federal University, Krasnoyarsk, Russia **Institute of Biophysics SB RAS*



Location of different types of sedimental probes and weather stations



Typical structure of spore-pollen spectrum

Regional level of paleogeographical forecasts and correlation is not well developed yet. However they are very important for the understanding of mechanisms of global changes. Intracontinental territories have specific mechanisms of landscape formation. Paleoclimatic reconstructions for these areas are based on mathematical simulation. Major point of such simulation is quantitative interpretation of interdependence between vegetation and climatic changes.

It is almost always impossible to make correct interpretations without correlations between spore-pollen data from sediments and recent samples. This is considered to be a reliable criteria for correlations between modern and past climate and vegetation.

Paleoclimate and vegetation reconstructions in the intracontinental areas of Yenisei Siberia are specific because of the variability of factors which form spore-pollen spectra. Among them are atmosphere circulation and mountain-intermountain relief. These features determine peculiarities of climate differentiation and vegetation distribution. Specific group of landscapes are forming in the depressions - intermountain depression landscapes. Considering complexity of natural processes this group of landscapes can be compared with mountain areas.

Correlation analysis was used at the initial stages of investigation. This allowed to estimate pare correlation coefficients between recent spore-pollen spectra and modern climate characteristics. Major aim of these analyses is revealing reguliarities between specific modern climate characteristics and spore-pollen spectra. Aiming to do that 794 recent samples from all the major genetic types of sediments from different landscapes were analyzed.

Recent (surface) samples were divided into the eleven landscape zones and subzones. Their distribution among landscape zones and sediments is presented at upper-left figure. These samples were correlated with fourteen climate elements. Meteorological data were compiled from 95 meteorological stations. Each spore-pollen spectrum is characterized by 70 components.

Maximal possible quantity of spectrum components –70 of initially 83 were used for investigation. Only rare elements were rejected. Each sample were correlated by 14 climate elements. Major aim is construction of universal model for intracontinental areas. That is why absolute height (Alt), geographical latitude (Lat) and geographical longitude (Long) were included in the models as factors.

Then a nonlinear model which utilizes equation of a following form was built:

$$f_{at} = \Delta_a + C_a \sum_{s=1}^{Size} \sin(\varphi_{as} + \sum_{i=1}^{I} u_{si} x_{it})$$

where *a* is an element of climate, *s* is a row of a matrix of model's coefficient, *i* is a number of component of a sporo-pollen specter, *t* is a number of sample component, *Size* is a number of rows of matrix of the model, *I* is the quantity of the components of sporo-pollen specters, *x* is the value of the *i*-th element of the *t*-th spectre, *u* is a coefficient from the matrix, *C* is the coefficient used for normalization of each component of climate, *f* is a climate's component, Δ are coefficients used to centralize *f*, are phase coefficients.

To find coefficient of model we used constrained variant of Quazi-Newton optimization while norm was quadratic. Chosen form of dependences allowed calculate derivates very fast even when coefficients were quite a number and also perform regularization constraint fast as well.

$$0 \le \varphi_s \le 2\pi, \qquad C^2 \times \sum_{s,i} u_{si}^2 \le P$$

where *P* is a chosen level of smoothness.

Such form allows to easily determine relevance of the input parameters:

$$A_i = \frac{\sum_s u_s^2}{\sum_{s,i} u_{si}^2}$$

The same equations were used as discriminating functions to provide nonlinear classification of vegetation zones.
So, we have got 14 equations to calculate each element of climate basing on spore-pollen spectrum, and 11 discriminating functions for classification vegetation zones.

Climate element	Average temperature,° C		temperature					Humidity		Precipitation, mm				
Simulation's accuracy	July	January	Warm period	Cold Period	Mean annual temperature,° C	Duration of period without severe frost, days	Sum of active temperature >10°C	Absolute, mbar	Relative, %	Annual sum	Warm period	Cold period	April, May, June	Coefficient of aridness
Correlation coefficient	0.97	0.99	0.99	0.99	0.99	0.85	0.98	0.99	0.85	0.96	0.93	0.94	0.94	0.83
Mean square error	0.014	0.027	0.011	0.019	0.012	0.31	2.26	0.0003	0.085	2.07	1.46	1.29	0.58	0.009
Absolute mean error, %	1.51	2.66	2.88	2.45	13.07	5.26	2.99	1.42	1.97	7.00	6.83	13.36	6.08	11.52