BioClimLand, 2014
No. 1, 53–69

UDK

Geo-simulation approach to modelling spatial objects and its application to creating thermokarst lake model using remote sensing data

Y.M. Polishchuk[1], V.Y. Polishchuk[2]

[1] Institute of Petroleum Chemistry, SB RAS
[2] Institute of Monitoring Climate and Ecological Systems, SB RAS

The paper discusses various approaches to modelling complicated objects. As shown The geo-simulation approach is considered to be the most useful for modelling natural objects with spatial properties. The geo-simulation model of the dynamics of these fields is proposed on the basis of experimentally determined statistical properties of the thermokarst lake fields. An analysis of experimental data on the statistical properties of the fields of thermokarst lakes in the area of permafrost in Western Siberia was carried out using remote sensing for measurement of lake areas. The relationship between climatic changes and geo-cryological parameters was studied and a multiple regression equation to take account of temperature and precipitation changes in the model was determined. A geo-simulation software package was developed. This makes it possible to model the dynamics of the thermokarst lake fields, taking into account identified thermokarst and climate changes. The validity of the model and its accuracy in use were tested by comparing model results with experimental remote data. The results can be used for prediction of the thermokarst change dynamics in permafrost.

Keywords: geo-simulation modelling, climate changes, geo-simulation software, thermokarst lakes, permafrost

Introduction

The intensification of geo-cryological processes in the permafrost caused by climate change, accompanied by intensified impact on infrastructure and other buildings in Arctic regions, will increase geo-ecological risks and result in significant economic damage. Remedial measures require making valid predictions concerning changes in the state of the permafrost, taking into account global warming. The forecasting of changes in the permafrost caused by climate change is an extremely important problem. There is a need for techniques and tools of mathematical modelling which use experimental data concerning changes in the permafrost. Since it is hard to reach the West Siberian wetland, collecting data is impossible without the use of remote sensing.

Thermokarst lakes, which can be decoded easily in satellite images, are the most suitable indicators of geo-cryogenic surface changes in northern permafrost landscapes. Sets of lakes of different sizes characterise these landscapes. Thermokarst lake fields display noticeable changes in both space and time. Their boundaries' shape and water surface areas change, new lakes emerge, some lakes disappear, turning into khasyreis (hollows of drained lakes). It is necessary to determine statistical regularities of spatio-temporal changeability of these fields' properties, using the analysis of experimental data obtained by remote sensing, to develop techniques and tools for modelling thermokarst lake field dynamics.

Thermokarst processes can be modelled mathematically with analytical models based on theory. Matt has shown [1] that such models are efficient for studying processes in a single thermokarst lake, but unsuitable for modelling spatio-temporal changes of thermokarst lake fields. The methods of mathematical morphology developed by Victorov [2, 3]are of great importance here as they are designed to use analytical models for territory dynamics modelling. These methods enable long-term dynamics of the state of a territory to be predicted; but they are not designed for the study of the spatio-temporal changeability of fields of thermokarst lakes. At present, no models are available for studying the spatial-temporal changes of random thermokarst lake fields.

There is a need for a new approach to modelling the dynamics of thermokarst lakes' fields. This approach based on geo-simulation methodology was proposed in our previous work [4]. Methodological issues were not discussed in that work, nor in works of other authors. The aim of this paper is to present the methodological issues in a geo-simulation approach to modelling natural objects with spatial properties and its application via a mathematical model of the dynamics of thermokarst lakes' field.

1. Geo-simulation approach to modelling complex objects with spatial properties

1.1. Mathematical models and modelling

Complexity and specific features of natural objects make experimental analysis of such objects extremely costly in time and money. As a result, researchers of complex natural systems face a lack of experimental data. This problem could be solved by using theoretical information from previous research. Propositions of such kind required development of a new approach to natural system research, on the basis that a model should be regarded not only as a means for information storage about an object under but also as an instrument for research development and forecast. Thus, mathematical modelling becomes an independent research concept. According to Trusov [5], modelling is defined as a process of model construction and use.

A model of natural systems cannot be developed using only theoretical research; the whole range of information about a researched object should be used. Thus, to design a proper model a combination of theoretical and experimental methods and data should be applied. Computer experimental work, based on models of researched objects and using computer generated means, takes place in this framework.

There are two main types of computer-based experiments: computational and simulation ones. Both types apply mathematical models as "substitutes" for a real research object. But, there are differences also. A mathematical model used in a computational experiment is realized in equations (differential, algebraic, etc.) which are impossible to be solved in analytical form (formula enclosing parameter dependency). Besides, numerical methods and effective computational algorithms are to be used. In contrast, simulation experiments, not based on general research object theory, describe behaviour of a real object, using empirical data, and are realized in a set of algorithms, reflecting changes defined by some scenarios in conditions of a modelled object [6, 7]. Simulation experiments result not in numerical solutions of some equations of a model, but in implementation of computer generated processes imitating the behaviour of a real natural object. The accuracy of the task completed, generated by mathematical modelling, depends on the degree of coincidence between the model and a real object. Thus, development of valid models for real objects is an urgent problem.

There is no single definition of a model. The most common definitions are presented below. According to Trusov [5], a model is a substitute object (analogous to the real one) which substitutes for a real object in the course of the research, preserving its typical features critical for the research. This condition allows a researcher to study original properties under question. According to Schrader [8], a mathematical model is a set of some elements in relations determined by this set. Peschel defines a model is defined as a structure to store knowledge about an object [9]. According to Fleishman *et al* [10], a mathematical model of a complex object is a sign system, the properties of which are in such proximity relation to properties of a studied object, that computer experiments can obtain all the necessary information about the object's behaviour and properties in set conditions.

We shall now consider basic approaches to complex system model construction and summarise model classification on the principle of application of *a priori* information about an object and a researcher's involvement in the construction process. An approach may be either inductive or deductive. An inductive approach uses empirical information as a source for *a priori* knowledge about an object. A deductive approach uses theoretical knowledge for the same purpose. Depending on the knowledge chosen as a basis for model construction, conventional types of models may be either empirical (inductive) or analytical (deductive).

Complex natural object research tasks have shown that neither empirical nor theoretical knowledge alone is sufficient for model construction. Thus, a new inductive-deductive model construction approach has been formulated. This urges a model researcher to use both empirical and theoretical knowledge about an object's properties and its behaviour.

Since presentation formats of empirical and theoretical knowledge are substantially different, the main difficulty in model construction is to create a structure which can organize object property information extraction procedures from empirical data and theoretically determined patterns in the most efficient way. Thus, it is necessary when designing such models to focus on special strategies and procedures, capable of applying an increase in information (empirical and theoretical) in the model. As such, the task is rather informational than mathematical. Accordingly, this task requires application of computer science methods and means as well as information technology. The importance of computer use is higher in the process of model construction in comparison with its use in model application.

Today, semi-empirical models, based on inductive and deductive ones, are becoming widely used, replacing conventional types of models (analytical and empirical) applying conventional inductive and deductive approaches to model constructing. Simulation models are typical semi-empirical models used to record numerical information about a modelled object. A detailed account of these model types now follows.

1.2. Basic types of mathematical models

Analytical models use theoretical knowledge as an information source about an object. When constructing these, the modeller selects what he considers to be the most substantial phenomena and objects, their elements and relations. This makes such analytical models, due to their extreme simplicity, applicable only for "rough" explanation and qualitative description of studied phenomena or an object.

The researcher's role in such modelling process is central. Analytical models can be implemented without computer technology, but in case of high model equation complexity, when obtaining of dependencies explicitly is problematic, model application results can be obtained with the help of a computational experiment. Such results have a limited character due to a simplified object model description and, as a rule, they are not efficient enough for the tasks of numerical prediction of complex object properties.

Empirical models commonly implement an inductive approach, usually being empiric-statistical models. These are based on statistical information about a modelled object. They also depend on the application of well-developed methods of mathematical statistics (regression and correlation analyses, statistical distribution law hypothesis check, etc.) These methods are not designed to reveal cause-effect relations but can be used to identify factors of interrelation presence between indexes of studied object conditions and hypothesis check concerning these relations.

Simulation models are mathematical. They are constructed in a condition of insufficient object description. They are thus different from any conventional models of mathematical modelling. The ability to construct simulation models without sufficient information about an object makes simulation modelling more true to life. One may expect substantial expansion of this type's application in research of complex natural objects.

A simulation model is a mathematical model [6] representing the process of a modelled object functioning within time by imitating basic phenomena and processes, and preserving their logical structure and time sequences, which allows using initial data about object state at snapshots to estimate features of a real object. Object behaviour is presented in a model as a set of algorithms which actualize conditions emerging in a real object state and changing in accordance with defined scenarios. Prediction of object properties and object behaviour forecast within simulation modelling ideology is performed by means of simulation experiment with a model. Such experiment allows the researcher to "replay" various situations changing according to some set scenarios in a real object state. The "scenarios" of simulation experiment are defined as either formalized or non-formalized descriptions of experiment objectives and procedures.

1.3. Geo-simulation modelling and geo-simulation models

Simulation modelling is one of the most important mathematical modelling types. According to Moiseev and Svirezhev [11], simulation modelling is a research method which can build an approximate model of a studied object; the simulation model describes a real object with accuracy sufficient for current research. Kosolapova and Kovrov [12], and Low and Kelton [13] claim that simulation modelling is used to construct models in cases where, firstly, there is no analytical solution or this solution is very complex and requires huge computer capacity and, secondly, the amount of experimental data about a modelled object is insufficient for statistical method. In such a case a mathematical model is developed in simulation modelling.

Russian scientists, Berlyant *et al* [14], have made a major contribution to methodology development of spatial objects modelling. They have named this type mathematico-cartographical modelling. Tikunov [15] presents methodological issues of using this type of model for modelling spatial natural objects. Kuzmichenok [16], Kovalev [17], Serdutskaya and Yatsishin [18], Kulik and Yurofeev [19], and Timonin [20], have all studied spatio-distributed objects using methods of mathematico-cartographical modelling.

A group of authors have introduced the special terms for modelling spatial objects. Lawson and Denison [21], Wang [22], Poh-Chin et al [23], Zhao and Murayama [24] have introduced spatial modelling.

Royle and Dorazio [25], Lawson [26], and Sang and Gelfand [27] introduced hierarchical.

Polishchuk and Tokareva [28], Zhao and Murayama [24], and Polishchuk, Yu. M. and Polishchuk, V. Yu. [29] have introduced geo-simulation modelling.

All the above types of modelling aim to study spatial objects using spatial data, analysis of which is based on spatial analysis methods implemented with the help of modern geo-information systems (GIS-analysis). In our opinion, the most suitable general term for the mentioned types of modelling (mathematico-cartographical, spatial, hierarchical, geo-information, geo-simulation) is the term 'geo-simulation modelling' which is defined as a model creation and model application for objects with spatial structure. Problems of creating a geo-simulation model of thermokarst lake fields will be considered further.

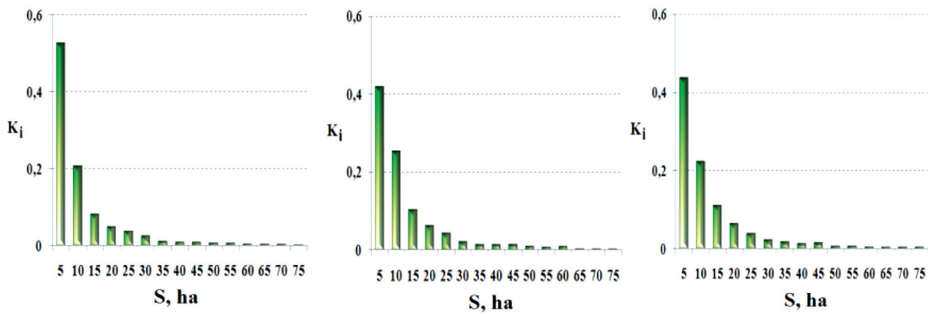
2. Creation of geo-simulation model of thermokarst lake fields

2.1. Experimental basis for geo-simulation modelling of spatial structure of thermokarst lakes' fields

Creation of a geo-simulation model of thermokarst lakes' fields requires knowledge of the basic properties of these fields, which can be obtained experimentally. Because of the inaccessibility of the northern territories of Siberia, thermokarst experimental studies were carried out by remote sensing. For remote study twenty-nine test sites (TS) were chosen in different zones of the West-Siberian permafrost (sporadic, discontinuous and continuous). Remote study of the shape of thermokarst lakes' boundaries was carried out via satellite images in our research [30]. Research conducted in test sites in sporadic, discontinuous and continuous permafrost showed that the error in estimating lakes' areas while replacing their real lakes boundaries by a circle is comparatively small (about 5%). It may serve as a reason to choose a circle as a model for a lake in geo-simulation modelling thermokarst lake fields. In addition, the formation of geo-simulation model of thermokarst lakes' fields in the form of a population of random circles requires experimental knowledge about the distribution of coordinates of lakes centres and the distribution of lakes sizes (areas).

To state the regularities for distribution of random coordinates of lakes, satellite images were used. Analysis of histograms of distribution of latitude and longitude values of location of lakes' centres given in Polishchuk, Yu. and Polishchuk, V., 2011 [30] showed that experimental regularities of distribution of coordinates of lakes' centres correspond to the law of uniform density according to criterion χ^2 with a probability of 95% [31, 32].

Histograms of distribution of lakes' number in accordance with their areas were built for all the test sites, located in different permafrost zones. For example, Fig. 1 represented the histograms for three test sites located in different permafrost zones. Here K_i – the relative number of lakes in each *i*-th interval of histogram, s – lake's area.



a *b* *c*

Fig. 1. Histograms of distribution of lakes in accordance with their areas: *a* – sporadic permafrost; *b* – discontinuous permafrost; *c* – continuous permafrost

Comparison of the diagrams shows that they have, in general, an exponential character by means of the experimental law of distribution, which makes it possible (for thermokarst lake fields to be modelled easily. We may choose a one-parameter exponential law to describe thermokarst lakes' distribution in accordance with areas in the following form:

$$y = \lambda \cdot e^{-\lambda s} \quad (1)$$

where λ – a parameter of distribution law.

The value of parameter λ can be determined with the help of experimental data. According to Wentzel [31], mathematical expectation of random quantity following the distribution law is determined in the form:

$$M(s) = \frac{1}{\lambda}.$$

Using as the estimate of mathematical expectation *M(s)* the value of average lake area estimated in accordance with experimental data in the following form:

$$\bar{S} = \frac{1}{n}\sum_{i=1}^{n} S_i,\ i = \overline{1,n},$$

we will find that parameter λ can be determined with the help of experimental data according to the formula:

$$\lambda = 1/\bar{S} \quad (2)$$

Testing correspondence of exponential law of lake area distribution given by Eq. (1) to experimental histograms shows that in all researched test sites this law corresponds to experimental data in accordance with criterion χ^2 with average probability 90%. Consequently, the stated law of lake distribution in form Eq. (1) according to areas does not contradict the experimental data. The analysis of the experimental distribution of lakes according to their areas shows that λ in all test sites in a continuous permafrost varies in the range of 0.037–0.071, and in a discontinuous permafrost – in the range of 0.034 – 0.086 with average values 0.054 and 0.060 respectively.

From the point of view of modelling thermokarst lakes, it is interesting to investigate the statistical relationship between changes in thermokarst lakes' areas and coordinates of their centres. Table 1 contains correlation matrices determining the degree of statistical relationship between changes in geographical coordinates of the lakes and their areas. Here we use the following agreed notations:

s – lakes' area;

x – geographical longitude;

y – geographical latitude.

Table 1. Correlation matrices of interconnection of changes in thermokarst lakes' areas (*s*) and centres' coordinates (*x,y*) in different permafrost zones

Discontinuous					Continuous			
random quantity	s	x	y		random quantity	s	x	y
s	1	–0,06	–0,09	xxxxxxx	s	1	–0,04	–0,03
x	–0,06	1	–0,06		x	–0,04	1	–0,08
y	–0,09	–0,06	1		y	–0,03	–0,08	1

Correlation analysis conducted for all test sites showed that the correlation coefficients between latitude and longitude do not exceed 0.1

These values display nearly absent correlated relationship between changing coordinates, which results in the conclusion about statistical independence of longitude and latitude changes. Correlation factors between values of areas and coordinates are within the interval –0.09 and +0.04, which also gives grounds for a conclusion about the statistical independence of changes in lakes' areas and their centres' location coordinates in different permafrost zones. Consequently, an important feature of thermokarst lakes' fields is statistical independence of changes in random quantities of centre coordinates and lakes' areas.

Accordingly, the following fundamental principles determining substantial properties of a model of spatial-temporal structure of thermokarst lakes' field can be formulated:

Lake coastline shapes can be represented by a circle equation with centres' coordinates x_i, y_i and area S_i (*i* – lake serial number).

Spatial changes in the position of centres of circles and their areas are statistically independent.

Random distribution of circle centres' coordinates x_i, y_i $(i =,\overline{1n})$ is governed by a uniform law.

Random distribution of number of circles over their areas conforms to the exponential law of distribution as in (1) with λ as a parameter.

Time changes in statistical properties of population of random circles and their dependence on climatic changes are determined by dependency of parameter λ on time and climatic characteristics in the following equation:

$$\lambda = f\ (T, P, t) \qquad (3)$$

where *T* – temperature, P – level of precipitation and *t* – time.

2.2. Geometric interpretation of model of spatial structure of thermokarst lake fields

A model of a thermokarst lake field is a population of random circles (Fig. 2), whose statistical characteristics correspond to the above principles (1–5). Figure 2 presents a geometrical interpretation of the model of thermokarst lake fields. Using number triple (*x, y, s*), representing the value of the centre of the circle coordinates and value of its area, coordinates of the points determining borders of each circle are calculated as follows:

$$x_k = R^* \cos\gamma + x, \ (4)$$
$$y_k = R^* \sin\gamma + y, \ (5)$$

where x and y – coordinates of the centre of the circle;

x_k and y_k – coordinates of k – th point on the circle;

γ – value of axial angle x and radius, directed from the centre of the circle into k –th point on the circle

R – radius of the circle, computed using the following formula $R = \sqrt{s/\pi}$;

where S – random value, distributed by exponential law and determining the area of the circle.

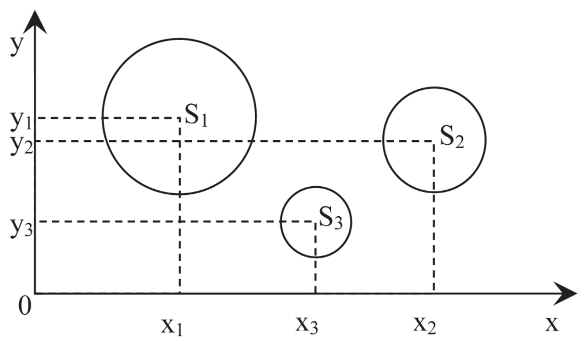


Fig. 2 Geometrical representation of thermokarst lake field model fragment as a population of random circles.

Consequently, major elements in the model description are characteristics of lake shapes, parameters of their random location on surface and random distribution of lakes over their size (areas). Analysis of experimental data on number change of thermokarst lakes, presented in the previous section, shows that the relative number of re-emerged lakes does not exceed 9% for the last three decades on the test site, and a relative number of disappeared lakes over the same period is a per cent share. Therefore, lake appearance and disappearance processes are not reflected in the developed geo-simulation model of thermokarst lakes.

2.3. Study of interrelation of climate and geo-cryological changes in permafrost and its accounting in the model

To consider climate change influence on properties of thermokarst lake fields, it is necessary to use geo-simulation modelling for the study of correlation of climatic and thermokast changes in the West Siberian permafrost zone. This requires the simultaneous analysis of data on area dynamics of thermokast lakes and time sequences of climatic indices provided by meteorological stations located on the geo-cryological changes test sites. Accordingly, location overlap (or at least close location) of test sites and meteorological stations is necessary.

The analysis of test sites and meteorological stations locations on the territory under study shows that, as a rule, they do not match. Meteorological stations are usually based on river and sea coasts or in an inhabited locality on hard-to-reach territories of the permafrost zone and thus are distant from test sites. Such location divergence of test sites and meteorological stations can cause considerable inaccuracy in the results of analysis of correlation of climatic and geo-cryological changes on the tested territory. Therefore, to analyse correlation of area change of thermokarst lakes and climatic indices (average annual temperature and precipitation level) an alternative approach was taken to obtain data on air temperature and precipitation. The approach is based on re-analysis of meteorological data which makes it possible to estimate the value of climatic indices in test sites.

Re-analysis (repeated meteorological analysis) [33] is a method to obtain meteorological information in set points of the tested territory. Re-analysis is based on "acquisition" of historical observational data from the net of meteorological stations, covering a continuous period and using a unified scheme of consistent data "acquisition" over the whole period of analysis. The main advantage of such meteorological data is uniform cover of the territory. Various systems [34, 35] of re-analysis are available. In the present work *Reanalysis ECMWF ERA-40, Reanalysis ECMWF ERA-INTERIM and APHRODITE JMA* systems were used to define climatic characteristics.

A brief description of re-analysis of data use procedure in the present research now follows. Average monthly values of air temperature and annual precipitation in test sites from 1970 until 2010 were obtained with the help of re-analysis. As such data are represented in a set of graphic files in the well-known *.png format, a template map locating test sites was developed to extract information necessary for climatic study. The template map was overlaid on the map of temperature and level of precipitation fields, obtained by re-analysis, and with the help of graphic editor MSPaint the template map was overlaid on the map of temperature and level of precipitation fields, obtained by re-analysis, a new image representing an overlap of test site location maps and temperature (or precipitation level) field was produced.

Such integrated maps helped to determine average monthly values of air temperatures and annual precipitation in the centre of each test site. The next stage was to calculate average value of air temperature for each test site monthly and annually. As a result, tables of time series of annual average value of air temperature and annual precipitation for each test site were obtained.

As shown in the previous section, features of thermokarst lake fields change under the conditions of contemporary climate warming. The most sensitive indicator of thermokarst processes influenced by temperature rise is thermokarst lakes' area, which is illustrated with the results of remote study in West Siberia [36, 37]. Bryksina *et al* [38] and Polishchuk *et al* [29] analyse the interconnection between thermokarst and climate change. This section will describe the results of analysis of interconnection between geo-cryological changes in thermokarst lake fields and climate features (air temperature and precipitation level) in the permafrost of West Siberia.

One can see the results of comparing time series of average lake area and average annual temperature in 29 test sites. According to the results of measuring lake areas with the help of satellite images, average lake areas were determined in each test site. The analysis of time series of total areas of thermokarst lakes made it possible to indicate trends in changing average area of the lakes. For example, in Fig. 3 there is a diagram of time dependence of average area of thermokarst lakes in TS – 13, displaying on average decrease in time in average lake area with coefficient of a linear trend $\alpha = -0{,}11$ha/year. A straight line in Fig. 3 presents the trend line.

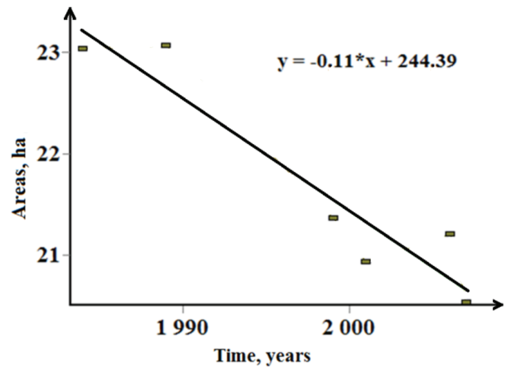


Fig. 3. Changes in the average area of thermokarst lakes in time

To study the interrelation between the changes of thermokarst lake areas and changes of air temperature and precipitation level we shall compare coefficients of a linear trend of time changes of average values of the lakes' areas and climate characteristics. It is the analysis of the data obtained for developing a model of thermokarst lake fields suitable for prediction that is of most interest. It is necessary to study temperature dependence of parameter λ, which determines the kind of law for thermokarst lakes' distribution in accordance with their areas, discussed in section 2.1. The data exist only for the years when cloudless images were taken, which made it possible to calculate the value of parameter λ.

The diagrams illustrating dependence of the value of parameter λ on average annual air temperature and precipitation level, presented in Fig. 4, show the dependence of parameter λ on climate indicators under study.

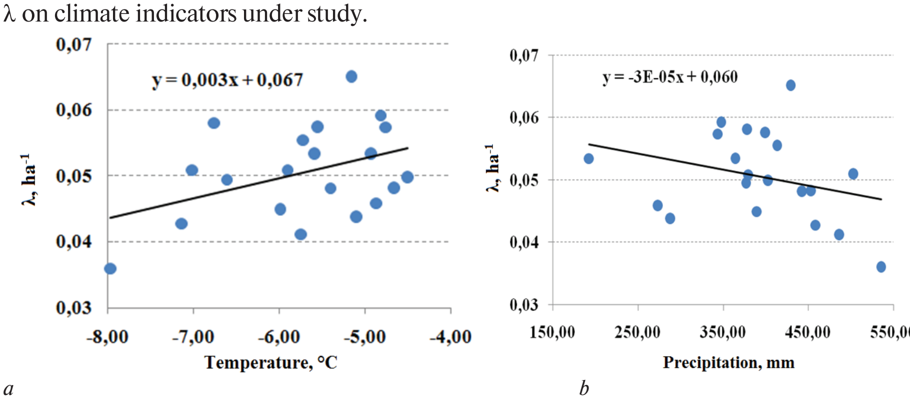


Fig. 4. Dependence of the values of parameter λ on annual rainfall and average annual air temperature

Figure 5 shows the time dependence of parameter λ. The equation of linear approximation of time dependence of λ shows the presence of dependence of parameter λ on time. This conclusion is important because taking into account time dependence of this parameter in modelling will allow the prediction of spatio-temporal changes in thermokarst lake fields under the conditions of climate change.

Previously the equation of dependence of parameter λ on time and climate features was introduced in implicit form (3). To develop a model of actual thermokarst lake dynamics it is necessary to define this dependence in explicit form. This was the reason for doing multidimensional regression analysis [32] of time series of the values of parameter λ and climate features in the West Siberian territory under study represented in Figs. 4 and 5.

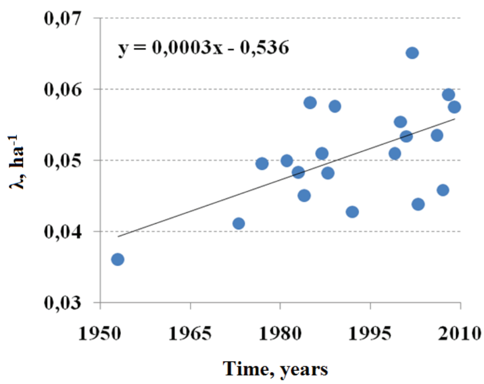


Fig. 5. Average value of parameter λ depending on time

The results of multidimensional regression analysis of the data on parameter λ and climate features can be presented as an equation of multiple regression in the form:

$$\lambda = c_0 + c_1{}^* x_1 + c_2{}^* x_2 + c_3{}^* x_3, \quad (6)$$

where x_1 – average annual air temperature, x_2 – precipitation level, x_3 – time, c_i – coefficients of regression equation, $i = 0,\dots,3$.

In the result of the regression analysis of time series of the values of parameter λ and climate features, the following values of regression equation coefficients were obtained (6):

$c_0 = -0{,}585$ ha^{-1}; $c_1 = 0{,}00062$ ha^{-1}/ºC; $c_2 = 0{,}000014$ ha^{-1}/mm; $c_3 = 0{,}00032$ ha^{-1}/year.

The stated regression dependence of parameter λ on time and climate changes is a basis for developing algorithms for modelling random thermokarst lake fields, discussed in the next section.

3. Geo-simulation modelling spatio-temporal structure of thermokarst lake fields

3.1. Algorithm for spatio-temporal modelling thermokarst lake fields

In a general case, mutual density of probabilities of random coordinates of centres and areas of circles imitating lakes in a mathematical model of random thermokarst lake fields can be presented in the form:

$$f(x, y, s) \quad (7)$$

where x and y – coordinates of circle centre in a model;

s – area of a circle imitating a lake.

Consequently, the totality of circles in the model of lake fields will be presented as a totality of groups of three random values (x, y, s). To develop an algorithm for modelling thermokarst lake fields, it is necessary to take into consideration statistical connections between changes in lakes' coordinates and their areas. The analysis of correlation matrix done in subsection 1.3 showed that random spatial changes in areas and coordinates of lake centres, in the same way as coordinates of lake centres between each other, are statistically independent. This is the reason why, when modelling thermokarst lake fields, mutual density of probabilities (7) can be presented in the form:

$$f(x, y, s) = f(x) \cdot f(y) \cdot f(s) \qquad (8)$$

where $f(õ)$ and $f(y)$ – densities of the probability of even distribution; $f(s)$ and value of model parameter λ are determined by (1) and (2) respectively.

Further to equation (7), the random-number sequence determining characteristics of location of circles' centres (x and y) is generated using the antenna of pseudo-random numbers distributed in accordance with the law of even distribution. And to form circles of random size whose areas are distributed according to the law conforming to equation (1) it is necessary to generate random-number sequences distributed in accordance with the demonstrative law conforming to equation [39]:

$$s_i = -\frac{1}{y}\ hz_j, \qquad (9)$$

where Z_j – numbers with even distribution in the interval (0,1), $j = 1,..., m$.

Consequently, together with using software generators for even distribution of pseudo-random numbers, the software realization of an imitation model of thermokarst lake fields includes creating a generator for pseudo-random number sequences, distributed in accordance with the demonstrative law. Then algorithms for modelling spatio-temporal structure of random thermokarst lake fields are considered.

We should consider a geo-simulation model of spatial structure of a thermokarst lake field $M_{\Pi}(t)$ that is a totality of circles and reflects the state of a thermokarst lake field at the moment of time t, whose geometrical representation is shown in Fig. 2. To model the dynamics of thermokarst lake fields, we should consider a general model of spatio-temporal structure of a thermokarst lake field in the form:

$$M_{\Pi B} = \{M_{\Pi B}(t_1),...,M_{\Pi}(t_j),..., M_{\Pi}(t_n)\}, \cdot j = 1,...,n \qquad (10)$$

which is a totality (time sequence) of geo-simulation models of a thermokarst lake field $M_{\Pi}(t_j)$, $j = 1,...,n$, where each model relates to a particular moment of time.

Figure 6 gives a visual presentation of the general model for spatio-temporal structure of thermokarst lake fields in the form of geo-information system (GIS) layers that relate to given time moments $t_1, t_2,..., t_n \in (t_1, t_n) \cdot \in$

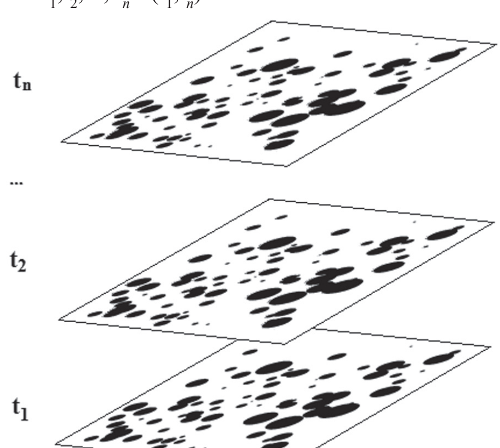


Fig. 6. Visual presentation of a general model of dynamics of thermokarst lake fields. Legend: t_i – time (year), $i = \overline{1,n}$

When modelling spatio-temporal structure of thermokarst lake fields it is important to take into consideration both time dependence and climate features (temperature, precipitation level). Accordingly, the dependence of parameter λ on time and climate features is determined by the equation of multiple regression in the form (6). This is the reason why equation (6) was used to develop an algorithm for numerical modelling dynamics of thermokarst lake fields.

The developed algorithm for modelling dynamics of thermokarst lake fields can be presented as follows:

the year of modelling is specified t_j, $j = 1,...,m$;

the areas (S_{MA}) of the model area (*MA*) under study are specified;

lake density (δ_{MA}) in *MA* is specified;

the number of circles within MA is determined in accordance with formula:

$$N_{MA} = S_{MA} \times \delta_{MA};$$

the centre of MA location in the map is specified;

parameter λ is determined in accordance with formula (6) for given values of temperature and time (year of modelling);

pseudo-random number is generated, distributed in accordance with uniform law;

using the number obtained at the previous step, a pseudo-random number is calculated according to formula (9) to characterize the value of circle area;

two pseudo-random numbers are generated, distributed in accordance with uniform law, determining the coordinates for circle centre location on the screen;

using the values of a number triple *(x, y, s)* obtained at previous steps 8 and 9, in accordance with equations (4) and (5) a circle is formed on the screen;

If the number of circles obtained is less than N_{MA}, determined at step 4, the algorithm repeats beginning with step 7, otherwise it is completed.

The given algorithm allows formation of a model of spatial structure for a given time moment $M_{\Pi}(t_j)$, where $j = 1,...,m$. To make a general model of dynamics of a thermokarst lake field by means of forming a time sequence of models $M_{\Pi}(t_j)$ for a given set of moments t_j ($i = 1,...,m$) the algorithm repeats for the number of times (m) needed.

Polishchuk (2013) presents a structural scheme of the developed software package for geo-simulation modelling thermokarst lake fields [40]; and the same paper discusses the software realization of the algorithm for geo-simulation modelling. Figure 7 presents an example of modelling thermokarst lakes' field.

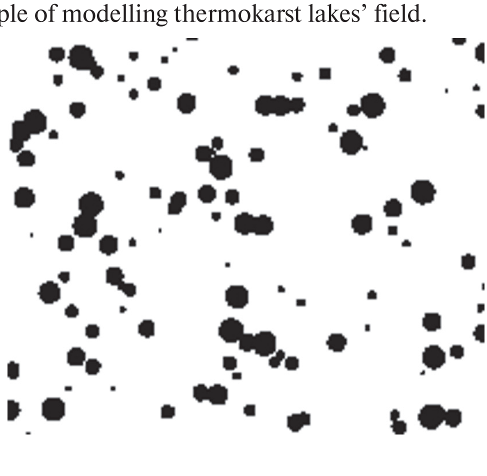


Fig. 7. Example of modelling thermokarst lakes' fields

The software package for geo-simulation modelling thermokarst lake fields is registered by Russian Federal Intellectual Property Service [41].

3.2. Studying model validity and modelling accuracy

The research is based on carrying out computer-based experiments. Model validity was studied by comparing average lake areas according to experimental and modelling data. Two data arrays containing 29 values corresponding to the number of TS were compared. The first data array contains values of thermokarst lake areas for each TS, averaged for the period under study, determined in accordance with satellite images. The second one – averaged for the period under study – contains values of model lake areas calculated with the help of the developed model. The second data array, in accordance with empirical data averaged for the period under study, contains average values of lake areas with the values of parameter λ being determined, as they were used in modelling.

Comparison of two selected arrays was done in accordance with the Student test. The formula used in calculation [42] to compare two small samples i.e. samples less than is:

$$t_\beta = \frac{\left|\overline{X} - \overline{Y}\right| * \sqrt{n * (n-1)}}{\sqrt{\sum \left[(x_i - y_i) - \left(\overline{X} - \overline{Y}\right)\right]^2}}$$

with the number of degrees of freedom $k = 2*(n–1)$,

where t_β – value of Student's test, used to determine confidence probability of groups' difference;

x_i and y_i –average values of lake areas in i-th TS, averaged for the period under study, according to experimental and modelling data respectively for i-th TS;

$\overline{X}$ and $\overline{Y}$ – arithmetic means for average values of lake areas according to experimental and modelling data, averaged for the period under study ;

n – size of experimental and modelling data samples.

During the research it was stated that with number of degrees of freedom $k = 56$ value of Student's test $t_\beta = 0,62$. According to Polishchuk [40], if $t_\beta = 0,62 < t_{\beta\_tab} = 2,003$ with $k = 56$, then for confidence probability 95 % we can reach a conclusion about the reliability of the similarity of samples.

Consequently, the developed model can be regarded as appropriate to the experimental data.

Accuracy of modelling dynamics of thermokarst lake fields was studied in the form of computer experiment on the model. Values of parameter λ in this case are calculated according to the multiple regression formula (6) using the data about average annual temperature and precipitation level determined for each TS by re-analysis. Then a model field is formed in accordance with the algorithm described above. The formed model thermokarst lake field is displayed on-screen (Fig. 7), then the array of values of modelled lake areas is saved as a file for analysis.

As mentioned above, calculated values of λ were used to generate model lake fields used to determine average values of model lake areas. Figure 8 shows the graph of time dependence of model area values with a solid line. A dashed line shows time dependence of average experimental values of lake areas. Vertical dashed sections indicate intervals between maximal and minimal values of experimental average values of lake areas. Figure 8 shows that the graph for average values of model areas in general keeps within the intervals, which may serve to indicate the reliability of modelling results.

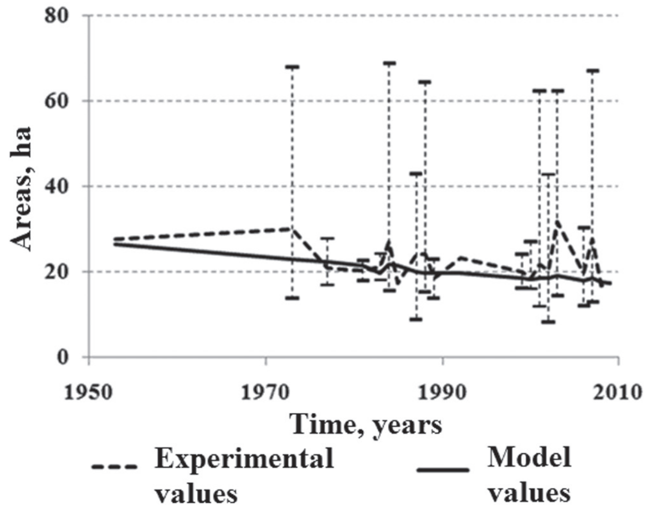


Fig. 8. Time dependences of average values of experimental and modelled average values of lakes' areas

To estimate the accuracy of modelling average thermokarst lake areas in the permafrost we shall determine root-mean-square deviation with Wentzel's formula [31]:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} (\frac{x_i - y_i}{x_i})^2}{n-1}},$$

where x_i and y_i – average values for experimental and modelled lake areas respectively; n – number of values, $i = \overline{1n}$.

Estimation showed that the error of determination of average values of lakes' areas on base of modelling with use of experimental data is 17 %. This may well be regarded as a suitable result of modelling thermokarst lake fields for predicting thermokarst lake fields' dynamics.

Acknowledgements

This research was supported by Russian Federal Programme "Kadry" (contract No 14.740.11.0409), ESA CAT-1 Project ID:5762 "Cryogenic processes of monitoring in Russian permafrost territories using radar and optical space images", RFBR Project No. 12-05-90833-mol_rf_nr. We also acknowledge financial support from a grant in accordance with Resolution of the Government of the Russian Federation № 220 dated April 09, 2010, under Agreement № 14.B25.31.0001 with the Ministry of Education and Science of the Russian Federation dated June 24, 2013 (BIO-GEO-CLIM).

References

[1] Matt A. 2011. Modelling thermokarst lakes dynamics and carbon flux Available online at: http://www.docstoc.com/docs/35351317/Methods_of_ thermokarst_lakes_modelling.

[2] Viktorov A.S. 2009. Mathematical Models of surface morphology in the geo-ecologic task decision. Presented at the annual session of the Scientific Council on Environmental Geoscience, Engineering Geology and Hydrogeology: Sergeevskie readings. Moscow, 23–24 March. (in Russian)
[3] Victorov A.S. 2006. The main problems of mathematical morphology. Moscow. Nauka. 252. (in Russian)
[4] Polishchuk Yu. M., Polishchuk V.Yu. 2011. Simulation modelling fields of thermokarst lakes in the permafrost // Information Systems and Technology. 1, 53–60. (in Russian)
[5] Trusov P.V. et al. 2005. Introduction to mathematical modelling. / Ed. P. V. Trusov. Moscow. Logos. 440. (in Russian)
[6] Maksimey I.V. 1988. Computer simulation. Moscow. Radio and Communications. 232. (in Russian)
[7] Shannon R. 1978. Simulation systems - the art and science. Moscow. Academic Press. 411. (in Russian)
[8] Schrader Yu. A., Sharov A.A. 1982. Systems and models. Moscow. Radio and communication. 152. (in Russian)
[9] Peschel M. 1981. Modelling of signals and systems. Moscow. Academic Press. 300. (in Russian)
[10] Fleishman B.S., Brusilovsky P.M., Rosenberg G.S. 1982. Methods of mathematical modelling of complex systems / System study. Moscow. Nauka. 65-79. (in Russian)
[11] Moiseev N.N., Svirezhev Yu. M. 1979. System analysis of dynamic processes of the biosphere. Conceptual model of the biosphere. Bulletin of the Academy of Sciences of USSR. 2. 47–54. (in Russian)
[12] Kosolapova L.G., Kovrov B.G. 1988. Evolution of populations: A discrete mathematical modelling. Novosibirsk. Bulletin of the Novosibirsk State University. 93. (in Russian)
[13] Low A.M., Kelton W.D. 2004. Simulation: Classic Computer Science. St. Petersburg-Kiev. Publishing Group BHV. 847. (in Russian)
[14] Berlyant A. M., Zhukov V. T., Tikunov V. S. 1976. Mathematics and cartographic modelling in the "creation - the use of maps"/ Geographical studies at Moscow University. Moscow. Moscow State University Press. 235-243. (in Russian)
[15] Tikunov V.S. 1997. Simulation in cartography: studies for high schools. Moscow. Bulletin of Moscow State University. 405. (in Russian)
[16] Kuzmichenok V.A. 2003. Mathematics and cartographic modelling of possible changes in water resources and glaciers of Kyrgyzstan to climate change. Bulletin of the Kyrgyz-Russian Slavic University. 6. 33–41. (in Russian)
[17] Kovalev T.M. 2008. Mathematical modelling of infection foci of tick-borne zoonoses in the Altai region. News of Altai University. 1. 58-62. (in Russian)
[18] Serdutskaya L.F., Yatsishin A.V. 2009. Technoecology: Mathematical and cartographic modelling. Moscow. Publishing house LIBROKOM. 232. (in Russian)
[19] Kulik K.N., Yurofeev V.G. 2010. Computer Mathematics and cartographic modelling agrolesolandshafts based on aerospace information. Reports of Russian Academy of Agricultural Sciences. 1. 52-54. (in Russian)
[20] Timonin S.A. 2010. Mathematics and Mapping and GIS modelling of demographic processes in the regions of the Russian Federation. Bulletin of Moscow State University, Ser. 5. Geography. 5 11-18. (in Russian)
[21] Lawson A.B., Denison D.G. 2002. Spatial cluster modeling / Boca Raton-London-New-York. CRC Press / Chapman and Hall. 287.
[22] Wang X. 2005. Integrating GIS, simulation models and visualization in traffic impact analysis. Computers, Environment and Urban Systems. 29. Issue 4. 471–496.
[23] Poh-Chin Lai, Fun-Mun So, Ka-Wing Chan. 2009. Spatial epidemiological approaches in disease mapping and analysis / Boca Raton-London-New-York. CRC Press / Tailor and Francis Group. 184.
[24] Zhao Y., Murayama Y. 2011. Urban dynamics analysis using spatial metrics geosimulation. In book: Y. Murayama, R. Thapa. (Eds.) Spatial analysis and modelling in geographical transformation process. Dordrecht-Heidelberg-New-York-London. (Springer, GeoJournal Library). 100. pp.153–168.
[25] Royle J. A., Dorazio R. M. 2008. Hierarchical modelling and inference in ecology. Oxford. Elsevier Sciences & Technologies. 444.
[26] Lawson A. B. 2009. Bayesian disease mapping: Hierarchical modelling in spatial epidemiology. Boca Raton. London. New-York: CRC Press / Chapman and Hall. 344.

[27] Sang H., Gelfand A.E. 2009. Hierarchical modelling for extreme values observed over space and time. Environmental and ecological statistics. 16(3). 407–426.
[28] Polishchuk Yu.M., Tokareva O.S. 2010. Geosimulation modelling of air pollution zones as a result of the gas burning in oil fields. Information Systems and Technology. 2. 39–46. (in Russian)
[29] Polishchuk V.Y., Bryksina N.A., Dneprovskaya V.P., Polishchuk Y.M. 2009. Analysis of the relationship of climatic and geo-cryologic changes in the permafrost of West Siberia. Presented at the 8th Siberian conference on climato-ecological monitoring, Tomsk, 8–10 October.
[30] Polishchuk V.Yu, Polishchuk Yu.M. 2012. Remote studies of variability of the shape of coastal boundaries of thermokarst lakes in the permafrost of West Siberia. Study of Earth from space. 1. 61–64. (in Russian)
[31] Wentzel E.S. 2002. Probability theory. Moscow. Vyschaya School. 575. (in Russian)
[32] Kremer N.S. 2003. Theory of probability and mathematical statistics: a textbook for high schools Moscow. UNITY-DANA. 573. (in Russian)
[33] Meteorological reanalysis. 2011. Available online at: http://en.wikipedia.org/wiki/Meteorological_reanalysis. (in Russian)
[34] European Center for Medium-Range Weather Forecast. 2011. Available online at: http://www.ecmwf.int
[35] Aphrodite's scope. 2011. Available online at: http://www.chikyu.ac.jp/precip/ scope/index.html
[36] Dneprovskaya V.P., Bryksina N.A., Polishchuk Yu.M. 2009. Study of changes of thermokarst lakes in the area of the discontinuous permafrost in Western Siberia on the base of satellite images. Study of Earth from Space. 4. 88-96. (in Russian)
[37] Smith L. C., Sheng Y., MacDonald G. M., Hinzman L. D. 2005. Disappearing Arctic Lakes. Science. 308(3). 14.
[38] Bryksina N. A., Polishchuk V. Yu., Polishchuk Yu. M. 2009. Study of relationship between climate change and thermokarst processes in the areas of continuous and discontinuous permafrost in Western Siberia. Bulletin of the Yugra State University. 3. 3-12. (in Russian)
[39] Buslenko N. P. 1968. Simulation of complex systems. Moscow. Nauka. 356. (in Russian)
[40] Polishchuk V. Yu. 2013. Software system of dynamics simulation of thermokarst lake fields in permafrost zones. Reports of TUSUR. 1. 125-128. (in Russian)
[41] Polishchuk V. Yu. 2011. Program simulation of space-time structure of the fields of thermokarst lakes. Registered Rospatent - Certificate No. 2011614293. (in Russian)
[42] Polishchuk V. Yu., Polishchuk Yu.M. 2013. Geo-simulation modelling fields of thermokarst lakes in zone of permafrost. Khanty-Mansiysk. Bulletin of Yugra State University. 129. (in Russian)