World Academy of Science, Engineering and Technology
International Journal of Agricultural and Biosystems Engineering
Vol:3, No:8, 2009

Application the Statistical Conditional Entropy Function for Definition of Cause-and-Effect Relations during Primary Soil Formation

Vladimir K. Mukhomorov

***Abstract*—**Within the framework of a method of the information theory it is offered statistics and probabilistic model for definition of cause-and-effect relations in the coupled multicomponent subsystems. The quantitative parameter which is defined through conditional and unconditional entropy functions is introduced. The method is applied to the analysis of the experimental data on dynamics of change of the chemical elements composition of plants organs (roots, reproductive organs, leafs and stems). Experiment is directed on studying of temporal processes of primary soil formation and their connection with redistribution dynamics of chemical elements in plant organs. This statistics and probabilistic model allows also quantitatively and unambiguously to specify the directions of the information streams on plant organs.

***Keywords*—**Chemical elements, entropy function, information, plants.

I. Introduction

WE previously [1,2] reported the results of our long-term research on patterns of rapid abiogenic parent material transformation in bioinert soil–like bodies under the action of plants root systems with their associated microflora and breakdown products during intensive year-round cultivation of plants in a controllable conditions. The integrity and interrelation of natural objects and phenomena underlie formation of the entire diversity of soils with their inherent information content [2]. When we studying the information function of soils, information theory methods permit quantitative estimation of the relative diversity of structures of the multitude of elements and monitoring of changes of these structures with time. This approach seems to be very useful for analysis of the evolution of diversity measures in the case of physical modeling of primary soil formation reproduced in a model soil-plant complex with permanent intense cultivation of plants in regulated agroecosystem.

We established experimentally [1] the considerable changes in physical, chemical and biogenic characteristics of root-inhabited media (RM) which based on granite crushed stone and other mineral materials occur in the case of long-term year-round cultivation of plants (spring wheat and tomato). The chemical elements composition of plants, as well as the biochemical composition of the organic matter formed in RM, and the accompanying microbiotic complex also changes. The most important consequence of this is a multiple increase of activity of biochemical processes by transformation of mineral substratum during primary pedogenesis. It has been noted that these changes in the RM-plant system have much in common with the evolutionary processes involved in primary soil formation under natural conditions with coarse mineral RM being subjected to intense biogenic erosion.

The purpose of this study was quantitative description of the response of the RM-plant system to complex evolutionary processes occurring in the course of long-term cultivation of higher plants (23 life cycles of wheat and tomato). Duration of life cycles are 75 days. The following operations were carried out while growing the plants. After each life cycle of the plants, first- and second-order roots were removed, the RM was composted for 20-30 days, and them combined acid-alkali regeneration was carried out by the method developed in [6]. After the twelfth life cycle we made interchange the positions of wheat and tomato. A number of fundamental studies [3,4] dealt with the problem of the close relationship between the element composition of plants and the soils on which they grow at natural conditions. At the same time, the biogenic accumulation of chemical elements at the case of intense year-round plants cultivation of modeling of soil-plant systems in a regulated agroecosystem has been poorly studied. However this problem is very interesting from both theoretical and practical points of view. We studied the regular dependence of the relative redistribution, formation of the relative diversity of chemical elements in plant organs (CaO, K_2O, P_2O_5, SO_3, Na_2O, SiO_2, Al_20_3, Fe_2O_3, MgO, Cl, MnO, ZnO) and its cause-and-effect relations occurring in plant during primary pedogenesis in the course of the evolution of the RM-plant system that played an important role in mineral nutrition of plants.

II. Information Model

It is important for solution of the practical problems to know the structure of the object being studied, consist of a set of components interacting with one another. Therefore, the information approach seems promising in this case. As a quantitative estimation of the dynamics of changes in the RM-plant system Shannon's entropy of information source of a finite ensemble of events $H = -\sum_{i=1}^{n} P_i \log P_i$ may be used. This value characterizes the diversity of the system's states under

V. K. Mukhomorov is with the Agrophysical Institute, 14 Grazhdanskii ave., St.-Petersburg,, 195220 Russia (e-mail: vmukhomorov@mail.ru).

the additional conditions $0 \le P_i \le 1$ and $\sum_{i=1}^{n} P_i = 1$, where n is the number of events-characters of the sequence of states that determine the space of their possible discrete values and P_i is the probability that the ith event-character will occur (the measure of the definiteness of the event realization). The P_i values are determined from the experimental data on the chemical elements composition of the tissues of plant organs. As our investigations showed, the entropy function sufficiently accurately characterizes the states of a complex multicomponent system. Since information and statistical entropies are proportional to one another, function H serves also as a diversity measure of the system or its structural ordering. According to our experimental data, the dynamics of the chemical elements in the RM – plant system does not have a random or chaotic character in time but obeys certain evolutionary regularities (Fig. 1). Fig. 1 shows, that during initial soil formation the increase a diversity of chemical elements in plant roots are accompanied by reduction of a diversity in reproductive bodies and on the contrary. The systems must have a definite structure, whose suitability is determined by its function in the overall system.

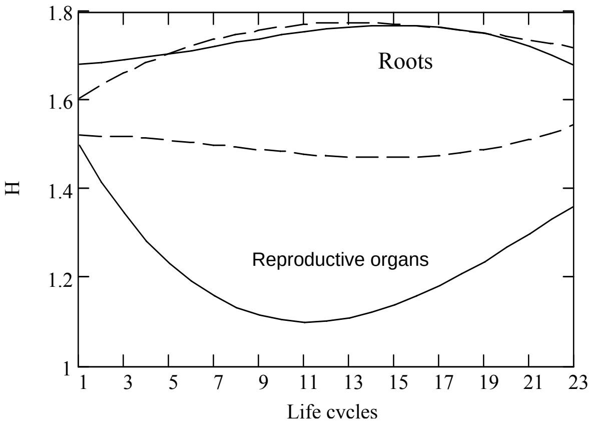

Fig.1 Dynamics (trends) of information entropy function H for chemical elements composition
– – – spring wheat, —— tomato

The mathematical approach using the entropy function for investigating of dynamic processes of multicomponent systems and dependent on the set of probabilities $P_i(t)$ of occurrence of a chemical element on a set of the investigated series of elements makes it possible not only to specify the direction of the interrelation between the distribution (i.e., mutual structuredness or heterogeneity of components of the systems) of chemical elements in individual plant organs, but also allow us to determine quantitatively a measure of cause-and-effect relations in the dynamic processes characterized by the function $H(t)$. When studying a complex phenomenon, such as functioning of the RM – plant system it is important to investigate the problem of cause-and-effect interrelations in it that including redistributions of chemical elements between plant organs. It is especially important at primary soil formation when there is a destruction of minerals and mineral elements act in plants. It is necessary to note, that traditional methods of processing experimental data by the regression analysis, correlation analysis, variance analysis, spectral analysis, etc., allow us establish only correspondences of the investigated processes. Without attraction the investigator's intuition or researcher's knowledge these methods do not allow to specify unequivocally a relationship of cause-and-effect between events. The presence of a correlation between time lines only point to no more as to a high degree of accompaniment in the development of the studied factors in time. However this correlation speaks nothing about cause-and-effect relations in the dynamics of the factors. It is known, that the high measure of closeness of a relation between levels in certain cases can be obtained even in the absence of cause-and-effect relations between corresponding phenomena. Stable trends in their development are sufficient for this, that is, a possible autocorrelation within each series. Processes may be correlated, but at the same time do not have cause-and-effect, since a correlation is a necessary, but not a sufficient condition for this. For clearing up of the relationships of cause-and-effect, the experimental material that is obtained by us on the distribution of chemical elements in wheat and tomato plant organs has been treated by methods of information theory.

Using the results of [6,7], we introduce the definition of causality: $\gamma = I(Y|X)/I(X|Y)$, where $I(Y|X) = H(Y|X)/H(X)$ is the independence of event $Y[Y_j]$ from event $X[X_i]$: $I(X|Y) = H(X|Y)/H(Y)$ is the independence of event $X[X_i]$ from event $Y[Y_j]$, and it is obvious, that $0 \le I \le 1$. If $I(Y|X) = 1$ then event Y does not depend on event X. For $I(Y|X) = 0$ the process Y is a one-valued function of process X. The conditional entropy of two random processes X and Y is defined by expression [5,6]:

$$H(X \mid Y) = -\sum_{j=1}^{k} P(Y_j) \sum_{i=1}^{n} P(X_i \mid Y_j) \ln P(X_i \mid Y_j)$$

where $P(X_i|Y_j)$ is the conditional probability, i.e., the probability of the ith value of event X under the condition of $Y = Y_j$; $P(Y_j)$ is an absolute probability jth level of realization, $H(X) = -\sum_{i=1}^{n} P_i(X) \ln P_i(X)$ is the unconditional entropy of the process $X(t)$. Here $X_i(t)$ and $Y_j(t)$ are the set of realization levels of events X and Y; k and n are the number of realization levels. Obviously, the conditional entropy is less then or equal to the unconditional entropy: $0 \le H(X|Y) \le H(X)$ and satisfies the conditions of reversibility of information $H(Y) - H(Y|X) = H(X) - H(X|Y)$.

Regarding the events X and Y as random processes having a statistically - probabilistic character and using the parameter γ as a quantitative measure of the cause-and-effect relation, we can reveal the directions of the interrelation between the relative content of chemical elements in a tissues of the plant organs and also estimate the quantitative measure of this interrelation. When $\gamma > 1$, event Y does not depend of event X. When $\gamma = 0$, the coupling between events is extremely irreversible and the causal relationships between X and Y are rigorously determined, i.e., event X is the cause of event Y. When $\gamma = 1$, processes X and Y depend equally with each other, i.e., cause-and-effect relationships are absent. The region of normal causality between events X and Y is determined by inequalities: $\gamma < 1$, $\beta < 1$, $\alpha < 1$ (Fig. 2) and the region of reversed causality by the inequalities: $\gamma > 1$, $\beta > 1$, $\alpha > 1$. Here the following values are used: $\alpha = H(Y)/H(X)$ is the unconditional asymmetry, $\beta = H(Y|X)/H(X|Y)$ is the conditional asymmetry of random processes X and Y. If the value $\gamma < 1$, then X is the cause, and Y is the effect.

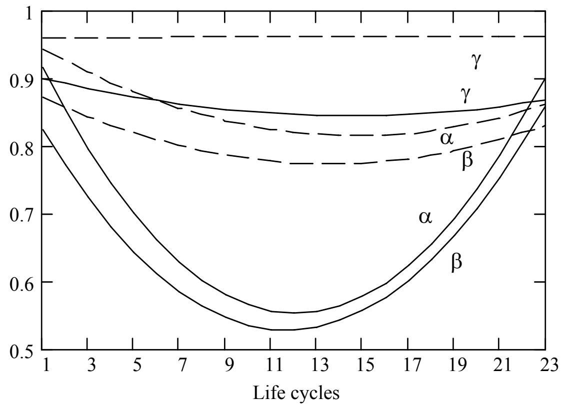

Fig. 2 The parameters $\alpha(t)$, $\beta(t)$ and $\gamma(t)$ the normal of cause-and-effect relations. The direction of information stream: roots (X) → reproductive organs (Y). – – – spring wheat, —— tomato. Arrow specify a direction of the information stream

III. RESULTS

The entropy approach makes it possible not only to establish the time dependence of causal relations between processes but also to estimate their asymmetry quantitatively. Fig. 2 shows the time trend dependences $\gamma(t)$, $\beta(t)$, and $\alpha(t)$ establishing cause-and-effect relations between the heterogeneity dynamics of the chemical elements composition in tissues of tomato and wheat plant roots and reproductive organs. Fig. 2 shows, that for tomato and wheat plants for the roots and reproductive organs the condition of normal causality is satisfied. Thus, change in structural heterogeneity of the chemical elements composition in roots (X) is the cause of the change in structuredness of reproductive organs (Y).

The dependence $\gamma(t)$ opens an opportunity to trace change of a causal relationship in time. The lower $\gamma(t)$ with respect to the horizontal line that the stronger the causal relation between dynamic processes $X[X_i]$ and $Y[Y_j]$. It is seen from Fig. 3 (the results of the analysis are given only for tomato plants) the parameter $\gamma(t)$ changes noticeably in time, i.e., during microevolution of primary pedogenesis in the RM-plant system, the causal relation in relative distribution of chemical elements between plant organs increases with increasing number of life cycles. However on a measure of increase with use RMs the causality weakens. These dependences are kept also for wheat qualitatively, however they are more intensive for tomato plants. Hence, the diversity of the chemical elements composition of the plant reproductive organs is secondary with respect to heterogeneity of the element composition in the plants roots, and the character of the cause-and-effect relations is preserved in time for both wheat and tomato plants and it has the same qualitative behavior in this period.

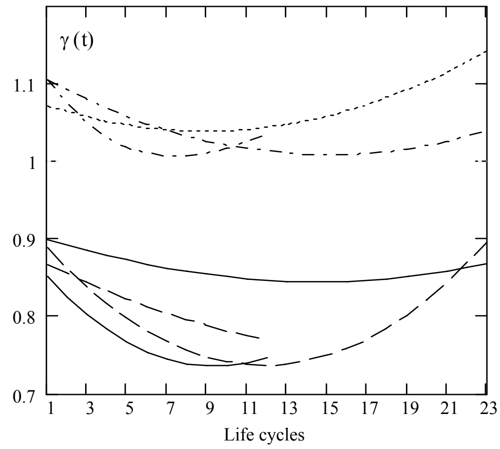

Fig. 3 Dynamics of cause-and-effect relations $\gamma(t)$ for the tomato plants. Tomato cultivated on granite crushed stone. · · · interrelation of the elements composition: roots (X) ← stems (Y), – · – interrelation of the elements composition: roots (X) ← leafs (Y). —— interrelation of the elements composition: roots (X) → fruits (Y). – – –interrelation of the elements composition: leaves (X) → fruits (Y). The curves until twelve life cycles are for tomato cultivated on zeolite

The entropy approach establishes a causal relation between the distribution of chemical elements in plant organs in the following direction: leaves (X) → fruits (Y). This direction of a causal relation does not contradict the known procedure of foliar fertilization of plants. At the same time, as it is seen from Fig. 3, a cause-and-effect relation roots (X) → stems (Y) is absent. The general character of the dynamics of the cause-and-effect relation roots → fruits and leaves → fruits for tomato plants cultivated on zeolite during twelve life cycles remains the same as when growing plants on granite crushed stone; however, in a quantitative respect this relation is strengthen. It must be pointed out that established direction of the change in the distribution of the chemical elements in plant organs is a reflection of the development of primary pedogenesis in a mineral RM under the action of biogenic factors.

It is known, the information entropy function determines also the contents of information in an object, which can variate in accordance with changes in the values of the function $H(t)$. According to the capacity theorem in a communications system [7] the upper limit of the rate of receiving information from source X to the receiver Y is determined by the relation [6]: $\sup C_{X \to Y} = R(Y \mid X) / H(Y)$, where $R(Y|X) = H(Y) - H(Y|X)$ is the information transmission capacity, $\delta t > 0$ is the representative transmission time of the information. The communication channel is any natural or artificial system in which an initial X and final state Y can be distinguished [6].

Taking into account, that the lower limit of information transmission is related to the rate by the relation

$\inf \xi_{X \to Y} = (\sup C_{X \to Y})^{-1}$, and accordingly for reverse transition $\inf \xi_{Y \to X} = (1 - I(X \mid Y))^{-1}$. The final difference $\Delta\xi = \inf \xi_{X \to Y} - \inf \xi_{Y \to X}$ defines the direction of information diffusion. If the transmission of the information $\Delta\xi > 0$ is implemented from X to Y, and if $\Delta\xi < 0$ the transmission of information is in the opposite direction $Y \to X$. After certain transformations of these formulas we obtain the relation for the information diffusion rat

$$C = -\frac{(\gamma - I(Y \mid X))(1 - I(Y \mid X)}{\delta t \cdot I(Y \mid X)(1-\gamma)} \quad .$$

Since the inequalities $\gamma > I(Y|X)$, $0 \le I(Y|X) \le 1$ the sign of the quantity C is determined by the difference in the denominator of the last relation. For the cause-and-effect relations (Y is the cause, X is the effect) the quantity $C < 0$ and for $C > 0$: X is the cause and Y is the effect. Figure 4 demonstrates the dynamics $C(t)$ and the direction of information streams: roots – reproductive organs, roots - stems for wheat and tomato plants. In the first case, information is directed from the plant roots to the reproductive organs, whereas in the second case, mainly from the stems to the roots.

The analysis of our experimental results has shown, that the entropy function describing a collective state of composition of the chemical elements in the plant organs is subordinated to the stringent rule.

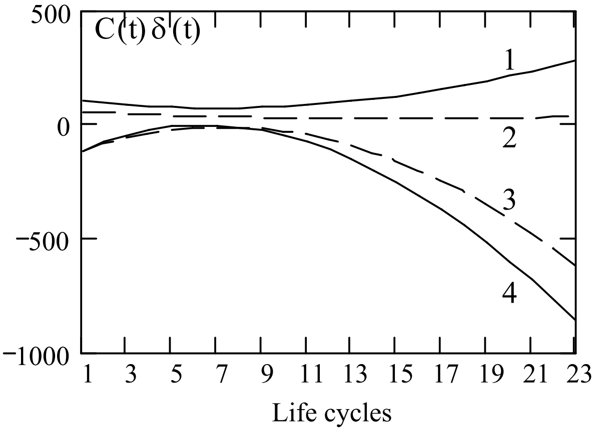

Fig. 4 Dynamics of the information streams $C(t)$.
1, 2 – roots (X) → reproductive bodies (Y). 3, 4 - roots (X) ← stems (Y). ——— spring wheat, – – – – tomato

The values of the information entropy for various tissues of the plants organs are in strictly specific region and these values do not overlap among themselves. We obtained (Table I) the following sequence of inequalities for values of the entropy function of the chemical elements composition in plant organs: H(roots) > H(leaves) > H(stems) > H(reproductive organs). In our experiment this sequence of the inequalities does not depend on a botanical kind of a plant and using the kind of mineral substratum (Figs. 5A, 5B), and points at nonidentity of collective states of the chemical elements in the various plant organs.

TABLE I
INFORMATION ENTROPY FUNCTION H (IN *NAT* UNITS) FOR CHEMICAL ELEMENTS IN PLANTS ORGANS

Roots	Reproductive organs	Stems	Leafs
Tomato			
Substrate - Granite crushed stone			
1.72±0.02	1.28±0.05	1.47±0.05	1.61±0.01
Substrate - Granite crushed stone after regeneration			
1.70±0.02	1.25±0.04	1.49±0.04	1.62±0.02
Spring wheat			
Substrate - Granite crushed stone			
1.76±0.04	1.50±0.02	1.56±0.02	*
Substrate - Granite crushed stone after regeneration			
1.79+0.03	1.51±0.02	1.55±0.02	*
Tomato			
Substrate - Zeolite			
1.87±0.05	1.32±0.07	1.52±0.04	1.60±0.01
Spring wheat			
Substrate - Zeolite			
1.87±0.02	1.55±0.02	1.66±0.02	*
Cucumber			
Substrate - Crushed granite			
1.88±0.02	1.37±0.03	*	*
Substrate - Granite crushed stone after regeneration			
1.94±0.03	1.29±0.03	*	*

* The experiment has not been realized.

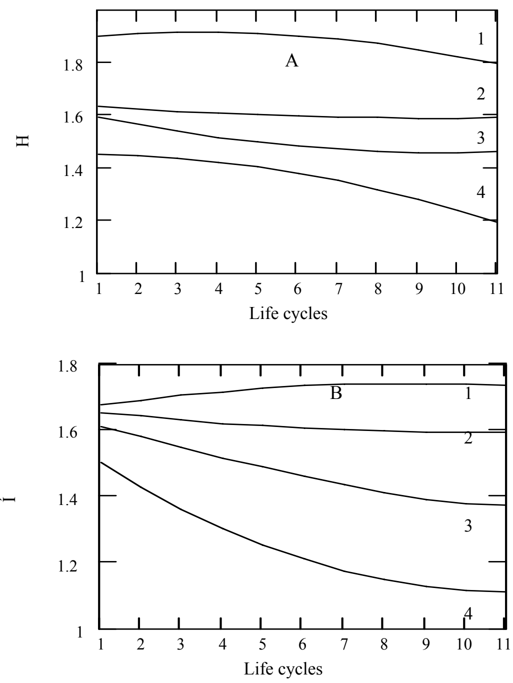

Fig. 5 Dynamics of entropy function for chemical elements composition of the plant tissues. A – tomato cultivated on granite crush stone, B – tomato cultivated on zeolite.1 - roots, 2 – leafs, 3 – steams, 4 – fruits

Great important in processes of primary pedogenesis is the effect of the vital functions of plants, of the organic matter forming in this case, and of the microbiotic community developing in mineral substrate, which is simultaneously combined with the transforming effect of living matter inseparably related to the planet's soil covering [2]. As a result, mutual intensification of the effect of the related processes of formation of organic matter and accumulation of living microorganisms on weathering and transformation of initial abiogenic RM occurs which happens especially intensively during year-round growing of plants. The most important consequence of this is a multiple increase of activity of chemical process of exogenous transformation of minerals during primary pedogenesis and change in the dynamics of uptake of chemical elements into plants. Under the effect of organic matter of the RM forming, and vital functions of the plant root system, chemical elements pass from a mineral state into a cation-anion labile state and become available to plants.

The information approach allows to establish connection between formed organic matter in RM and chemical element structure of plant tissues. Figs. 6 and 7 show a conjugation of a collective state of organic matter components composition and a collective state of chemical elements in root tissues.

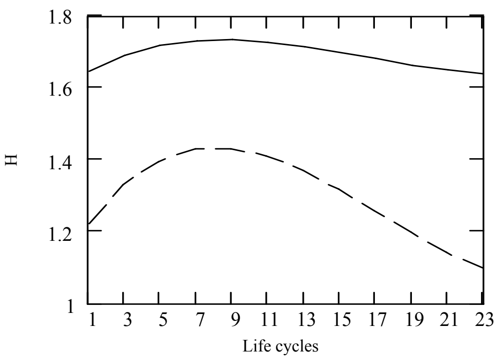

Fig. 6 Dynamics (trends) entropy function *H*.
—— chemical elements composition. Tomato. Roots.
– – – organic matter formed in the mineral substrate.
Statistics: Correlation coefficient $R = 0.97$, $F = 145 > F_{0.95;1,10}^{(cr)} = 4.84$

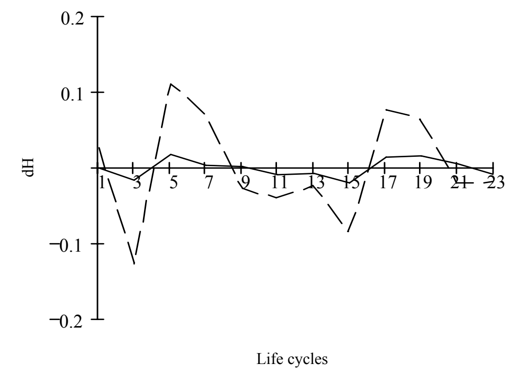

Fig. 7 Oscillations of the entropy functions *dH* around trends.
—— chemical elements composition. Tomato. Roots.
– – – organic matter formed in the mineral substrate.
Statistics: Association coefficient $\Phi = 0.7$, $\chi^2 = 5.7 > \chi_{cr}^2 = 3.8$

The time series we obtained on the basis of experimental data had a long-term component (trend), against the background of which variations developed in the diversity of systems. It can be see from Fig. 7 that the temporal variations of function *dH* for the organic matter and chemical elements composition of roots change synchronously. The sequence for *H* and *dH* forms a stable cause-and-effect connection between the organic matter diversity and the chemical elements composition in the roots. Obviously, the formed organic substance is connected with formation of chemical elements structure of plants and promotes erosion of a mineral substratum.

IV. Conclusion

As a result of the investigations, cause-and-effect relations in the dynamics of redistribution of a set of chemical elements between plant organs during primary pedogenesis were established for the first time. During intensive cultivation of tomato and wheat plants on an initially abiogenic RM, and increase of heterogeneity of chemical elements in the root system leads to a decrease of heterogeneity in the reproductive plant organs and vice versa. At the same time, a change in the relative content of chemical elements in plant roots does not have a causal relation to the change in heterogeneity in plant stems. The proposed entropy approach applied to an analysis of mutual distributions of chemical elements between plants organs also makes it possible to indicate the direction of information streams in plants during a succession of biogeocenoses.

Acknowledgment

This research was sponsored by the Russian Foundation for Basic Research (Grant 05-04-49542).

References

[1] E. I. Ermakov, V.K. Mukhomorov. “Evolution of diversity measures as a reflection of the processes of primary soil formation in a model soil-plant system” Doklady - Biochemistry and Biophysics. vol. 379, no. 6, pp. 297-301, Nov. 2001.
[2] G.V. Dobrovol’skii, E. D. Nikitin. Funktzii pochv v biosphere i ekosistemakh (Ekologicheskoe znachenie pochv). (Soil Functions in Biosphere and Ecosystems: The Ecological Significance of Soils), Moscow: Nauka, 1990, ch. 6. (Russian Ed.).
[3] E. I. Ermakov, L. M. Anikina, L. A. Chaikovskaya. Inventor’s Certificate No. 1303063, “Method of combined chemical sterilization and regeneration of soil substitutes”, Byull. Izobr., 1987, no. 14.
[4] V. I. Vernadskii. Khimicheskii sostav zhivogo veshchestva v svyazi s khimiei zemnoi kory (Chemical Composition of Living Matter in the Context of Chemistry of the Earth’s Crust). Moscow: Nauka, 2001, ch. 2, 3. (Russian Ed.).
[5] V. B. Il’in. Elementnyi khimicheskii sostav rastenii (Chemical-Elements Composition of Plants). Novosibirsk: Nauka, 1985, p.23. (Russian Ed.).
[6] S. M. Korotaev. “O vozmozhnosti prichinnogo analiza geophizicheskih processov (On the possibility of the causal analysis of the geophysical processes” Geomagnetiz’m i Aeronomiya (Geomagnetism and Aeronomy), vol.32, no. 1, pp.27-33, Jan. 1992. (Russian Ed.).
[7] C. E. Shannon. “A mathematical theory of communications” Bell System Tech. J., vol. 27, no.3, pp.379-423, 1948.