"Complexity–entropy" diagrams and their application to the study of coal tectonic disturbance

Olga Malinnikova[1*], *Dmitry* Uchaev[2], *Denis* Uchaev[2], *Vasiliy* Malinnikov[2], and *Catherine* Ulyanova[1]

[1] Research Institute of Comprehensive Exploitation of Mineral Resources of Russian Academy of Sciences, Laboratory of Geotechnological Processes, 111020 Moscow, Russia
[2] Moscow State University of Geodesy and Cartography, Department of Space Monitoring and Ecology, 105064 Moscow, Russia

Annotation. It is investigated the possibility of using so–called "complexity–entropy" diagrams for the quantitative description of the degree of coal disturbance using coal images obtained by means of scanning electron microscope (SEM). These diagrams plot structural complexity measure (vertical axis) versus entropy measure (horizontal axis) for distribution of probability given in some way on the image. In this paper, the values of both measures were calculated on the basis of the shearlet transform, and the Jensen divergence was used as the basic divergence for calculating the complexity measure. All calculations were performed for more than 140 images of coal specimens with various degrees of disturbance, obtained from the quiet zone of the seam and the outburst zone. As a result of research, it was found that two-dimensional distributions for measures of complexity and entropy in most cases are informative data sets for differentiating coals by degree of complexity. Moreover, such characteristics of these distributions as mathematical expectation and, to a less degree, mode can be used as simple quantitative descriptors of coals with various degrees of disturbance. These characteristics can be used to show the closeness of the spatial structure for the analyzed coal specimens to strictly periodic or absolutely chaotic ones. On the basis of the obtained results, conclusions about the possibility to separate coals according to the degree of their outburst hazard were done.

1 Introduction

Natural coal is a product of a complex chain of transformations of heteromolecular substances of plant remains under the influence of biological, chemical and tectonic factors at high temperatures and pressures. Coal seams may contain a large amount of gas, mainly methane, most of which is in the absorbed state in the coal substance [1]. The complex irregular structure of the pore space of fossil coals is largely determined by the distribution of voids (pores, microcracks) filled with gas by their number, size and directions. The sizes

---

* Corresponding author: olga_malinnikova@mail.ru

of these voids vary in rather wide limits: from 0.3 nm to 10 cm, and the directions depend on the dip and strike of rock layers and movement of material particles in coal seams [2]. It is known that the outburst–hazardous coal seams generally have a more disturbed structure, which is often a superposition of several systems of exogenous and endogenous fracture, and have increased microporosity. Thus, the study and quantitative assessment of natural and man–made disturbance of coal can contribute to the understanding of the mechanism of rapid release of sorbed methane, the participation of methane in the destruction of coal and improve the methods of preliminary prediction of dangerous gas–dynamic phenomena in the development of coal seams [3].

In general, the processes of multiple destruction of fossil coal (accumulation of destructions in coals) are the processes of spatial self–organization of their structure. The formation of the main macrofracture in the process of destruction of coal seams is preceded by the process of development (origin, movement, growth and aggregation) of microdefects (pores, microfractures, dislocations, etc.), which is stochastic in nature [4]. Not so long ago, the stochastic behavior of the system could be explained only by the influence of random forces. However, in recent years, the chaotic behavior of nonlinear deterministic systems has become widely known. Chaos is a rather unusual form of behavior of a deterministic system in a steady state. Although the evolution of chaotic system is uniquely determined by dynamic laws and no random forces acting on her, the dynamics of the system in some area of the phase space is stochastic [5]. Chaos easily occurs in many natural and living systems where nonlinearity exists. In this regard, the question arises: does the image of the surface structure of the coal show the result of a chaotic or random process? One of the methods for solving this problem is the use of "complexity–entropy" diagrams, proposed in [6], which are planes of values of a measure of complexity (vertical axis) depending on the corresponding values of the entropy of the probability distribution (horizontal axis). When constructing the "complexity–entropy" diagrams, the key issue is the choice of the method of forming the measure. In the case of studying the disturbance of coals by their SEM–images, methods of forming a measure that take into account the high degree of anisotropy of the surface of the coal samples are interesting.

Over the past twenty years, various methods have been proposed to recognize anisotropic objects, among them: directional wavelets, complex wavelets, contourlets, curvlets, etc. In turn, Donoho D., Labate D. and Kutinek G. [7–11] suggested a slightly different approach to the analysis of anisotropic components based on shearlet–transformation. Unlike wavelets or curvlets, the system of shearlets is built in the class of affine systems and has the ability to recognize the directionality due to the additional shift parameter. Shearlets have characteristics that favorably distinguish them from a number of similar functions: a finite number of generating functions; optimal representation of the anisotropic characteristics of the analyzed data; fast algorithmic implementation; unified approach to the decomposition of continuous and discrete data [12]. For the first time, the use of a shearlet transform to create "complexity–entropy" diagrams was proposed by Brazhe A. in [13]. In this paper, we studied the possibility of using "complexity–entropy" diagrams based on shearlet transform to quantitatively describe the degree of disturbance of coals by their SEM–images.

2 Methodology of research

To produce modified shearlets, a dilation matrix A_a and a shear matrix S_s are used:

$$A_a = \begin{pmatrix} a & 0 \\ 0 & \sqrt{a} \end{pmatrix}, a \in \widecheck{Y}^+, \ S_s = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}, s \in \widecheck{Y}$$

Thus, given scale a, shear s, and translation t, it is possible define the "mother" shearlet function

$$\psi_{a,s,t} = a^{-3/4}\psi\left[A_a^{-1}S_s^{-1}\left(x-t\right)\right],$$

and on its basis a continuous shearlet–transformation of a digital image, as a convolution of the image with scaled, sheared and translated copies of the "mother" shearlet function.

The discrete system of shearlets associated with shearlet ψ, is the set of functions

$$SH(\psi) = \left\{\psi_{j,k,m} = a_j\psi\left(A_{2^j}^{-1}S_{j,k}^{-1}\left(x-t_m\right)\right) : j,k \in \widecheck{y}, m \in \widecheck{y}^2\right\}.$$

In general, the scale parameter is selected from the set $\left\{a_j\right\}_{j\in\widecheck{y}} \subset \widecheck{Y}^+$. The shear parameters $\left\{s_{j,k}\right\}_{k\in\widecheck{y}} \subset \widecheck{Y}$ are chosen dependent on j, so that the direction of interest varies according to the scale. Finally, the translation parameter t_m is chosen from $c_1\widecheck{y}+c_2\widecheck{y}$ for some $(c_1,c_2) \in \widecheck{Y}$, which provides the necessary flexibility of a construction for many applications. For the image size $M\times N$ it is accepted to determine the discrete values of the parameters as follows [14]:

$$a_j = 2^{-2j} = 4^{-j}, \ j = 0,...,j_0-1,$$

$$s_{j,k} = k2^{-j}, \ -2^j \le k \le 2^j$$

$$t_m = \left(\frac{m_1}{M},\frac{m_2}{N}\right), m \in I,$$

where $j_0 = \lfloor 0.5\log_2 N\rfloor$, $I = \left\{(m_1,m_2)\,|\,m_1 = 0,1,...,M-1, m_2 = 0,1,...,N-1\right\}$. This leads to the following shearlet transform [13]:

$$\psi_{j,k,m}(x) = \psi_{a_j,s_{j,k},t_m}(x) = a^{-3/4}\psi\left[A_{a_j,1/2}^{-1}S_{s_{j,k}}^{-1}\left(x-t_m\right)\right].$$

Thus, a discrete shearlet transform is defined as a mapping from the original image to a set of shearlet coefficients:

$$SH(f): f \ a \ \left\langle f,\psi_{j,k,m}\right\rangle.$$

It was shown in [13] that because the translation grid (dimensionality m) is scale independent and redundant, the shearlet coefficients can also be represented as $SH(f) = \left\{S_{j,k}(x,y)\right\} = \left\{S_i(x,y)\right\}$, where $S_i(x,y)$ are images, obtained by convolution of the function $f(x,y)$ with directional filters of different spatial scales of the same size as the function $f(x,y)$.

3 Entropy and complexity measures based on shearlet transform

Both entropy and complexity measures are defined as functionals of some probability distribution $P = \{P_i, i = 1,...,L\}$. Intuitively, the energy of shearlet coefficients $E_i(x,y) = S_i^2(x,y)$ describes how the corresponding scale and orientation is represented at a given location of the image $f(x,y)$. Starting from $E_i(x,y)$ we can define local shearlet features distributions in the image as follows [13]:

$$P_i(x,y) = \frac{E_i^*(x,y)}{\sum_{j,k}E_{j,k}^*(x,y)},$$

where $E_i^*(x,y) = \left(K_{\sigma_j} * E_i\right)(x,y)$, K_{σ_j} denotes a Gaussian kernel with a standard deviation $\sigma_j = \sigma_0 2^{j_0 - j - 1}$.

Based on the obtained probability distribution $P = \{P_i, i = 1,...,L\}$ Shannon entropy $H[P]$ and Shannon normalized entropy $H^N[P]$ can be defined as

$$H[P] = -\sum_i P_i \log_2 P_i \,,$$

$$H^N[P] = H[P] / \mathrm{H}[P_e] \,,$$

where $H[P_e] = H_{\max} = \log_2 L$ is the entropy of an equiprobable (uniform) probability distribution P_e, in which all shearlet–features are represented with equal probability $P_i = 1/L$.

As a measure of complexity in this work, we use the statistical measure of complexity $C[P]$ proposed by Lopez–Ruiz et al. [6] and defined as

$$C[P] = Q_{\mathrm{JS}}[P, P_e] H^N[P] \,,$$

where $Q_{\mathrm{JS}} = J[P, P_e] / J_{\max}$ is a measure of disequilibrium describing the distance between the observed P and uniform P_e statistical distributions by means of the distance measure based on Jensen–Shannon divergence:

$$J[P, P_e] = H\left[\frac{P + P_e}{2}\right] - \frac{1}{2}(H[P] + H[P_e]) \,.$$

Clearly, $J[P, P_e] = 0$ if $P = P_e$ and is maximal, $J_{\max} = -0.5\left\{\left[(L+1)\log_2(L+1)\right]\big/L - 2\log_2 2L + \log_2 L\right\}$, when only one feature, say m th, is present, while all others are absent: $P_i = 1 \,|\, i = m$, $P_i = 0 \,|\, i \neq m$.

A measure of complexity defined in this way quantifies both randomness and degree of spatial correlations in the data, and therefore there is a range of admissible complexity measure values $[C^{\min}, C^{\max}]$ for each entropy value H [6, 15, 16]. For a distance measure based on Jensen–Shannon divergence, the lower bound is found by a family of distributions where one of the outcomes P_i has probability $P_i = 1/L,...,1$ and the rest P_j have uniform probabilities $P_j = (1 - P_i)\big/(L - 1) \,|\, j \neq i$. The upper bound is formed by distributions with $n < L$ outcomes having probabilities $P_i = 1/n, i = 1,...,n$ and the rest having zero $P_j = 0$, $j = n + 1,...,L$ [13].

4 "Complexity–entropy" diagrams

In this paper, the above–mentioned interrelated measures of entropy and complexity of digital images were used to construct of "complexity–entropy" diagrams. Thus, points on these diagrams show values $C\{P(x,y), H^N[P(x,y)]\}$ calculated for each pixel of the digital image, and lines show values of the lower $C^{\min}$ and upper $C^{\max}$ bounds of the range of allowable values of the measure of complexity C. For construction of summary "complexity–entropy" diagrams, showing results of the analysis of a set of images, instead of the values $C\{P(x,y), H^N[P(x,y)]\}$, modal values $C\{P(x,y), H^N[P(x,y)]\}$ for each image from the set were shown in these diagrams.

To implement an algorithm for constructing the above diagrams, a computer program in Python was developed using the free program *shearlexity* [17], standard Python libraries *scipy* and *matplotlib* for scientific calculations and visualization, as well as fast discrete transform library *PyShearlets* [18].

The interpretation of these diagrams is based on the following intuitive considerations: zero entropy and complexity correspond to a completely regular structure, and high entropy and zero complexity correspond to a completely random spatially independent noise [19, 20].

5 Source data

In this study, imaging of coal specimen surfaces was performed using scanning electron microscopes JEOL JSM 5910–LV and Jeol–6610–LV. The spatial resolution of the microscopes is more than 10 and 100 nm for secondary and reflected electrons, respectively. Low–energy secondary electrons are used in imaging surface topography. Natural–shape coal specimens were placed in a work camera via a gate. In the mode of registration of secondary electrons, the work camera was vacuumized (with $>10^{-6}$ mm hg vacuum). Secondary electrons were recorded by a standard detector, which a type of a sweeping–field photomultiplier tube connected to scintillator.

As the source data for this study was IPKON collection of coal specimens from the Zapolyarnaya mine (Vorkuta) and the Kirov mine (Leninsk–Kuznetsk), obtained from outburst–nonhazardous zones and outburst zones. For our research, we analyzed microstructure of coal surface in the images magnified 1000 times which showed coal grains with a characteristic size from 0.5 to a few microns. Methane can desorb from such grains, diffuse and flow in fractures as free gas [21].

Processed regions of these images were same as in [22], which further allowed to compare obtained results with the results from [22].

6 Research results

In Figs. 1, *d* and 2, *d*, examples of typical "complexity–entropy" diagrams constructed for selected regions of test coal surface images (Figs. 1, *a* and 2, *a*) are shown. These diagrams are based on values for measures of local entropy H and complexity C, which are demonstrated in Figs. 1, *b*, *c* and 2, *b*, *c*. As can be seen from these diagrams, the distribution of values $C\{P\ (x,y), H^N[P\ (x,y)]\}$ obtained for the image of the surface of coals from the quiet zone is more compact than the corresponding distribution of values $C\{P\ (x,y), H^N[P\ (x,y)]\}$ for the image of the surface of coals from the outburst zone. It was also found that most of the processed regions of coal images have smooth unimodal distributions of values for local Shannon entropies and Jensen–Shannon complexity (Figs. 1, *e*, *f* and 2, *e*, *f*). This observation suggests the possibility of using the predominant values of the selected measures of entropy and complexity as quantitative descriptors of the degree of order and structure of the coal surface regions. At the same time, taking into account the nature of the analyzed distributions, it seems statistically correct to use the mean or modal values of these distributions as such values. Both variants were analyzed in the context of the problem of coal separability from the quiet (outburst non-hazardous) zone and the outburst zone by their SEM–images. As a result, it was found that both modal and mean values of entropy and complexity measures indicate a random (not chaotic) nature of the distribution of anisotropic properties for most of the analyzed images regions with a

sufficiently close scatter of values: $H \in [0,62 \div 0,94]$, $C \in [0,09 \div 0,31]$ and $H \in [0,65 \div 0,96]$, $C \in [0,08 \div 0,33]$ for the mean and modal values, respectively.
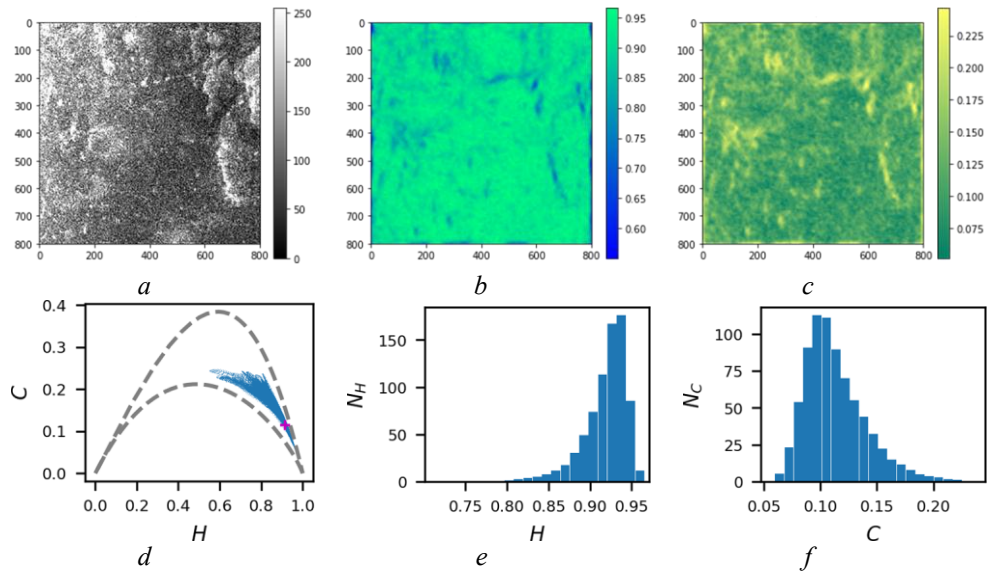

Fig. 1. Images of local entropy H (*b*) and complexity C (*c*), the corresponding histograms (*e* and *f*), and the "complexity–entropy" diagram (*d*) built for the region of the image of coal from the quiet zone (*a*). A cross in the "complexity–entropy" diagram shows the mean value of the measures of complexity and entropy.
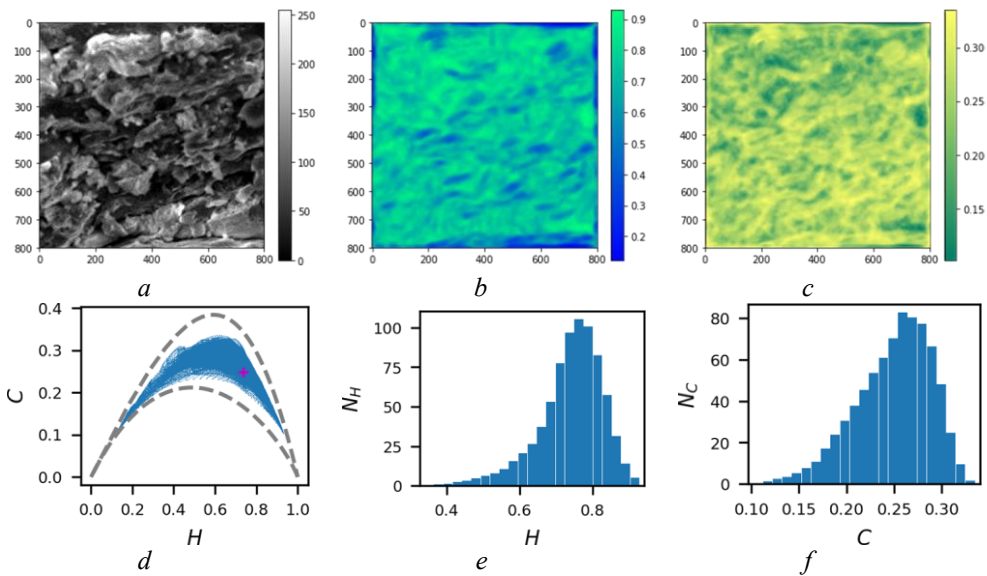

Fig. 2. Images of local entropy H (*b*) and complexity C (*c*), the corresponding histograms (*e* and *f*), and the "complexity–entropy" diagram (*d*) built for the region of the image of coal from the outburst zone (*a*). A cross in the "complexity–entropy" diagram shows the mean value of the measures of complexity and entropy.

At the same time, the images of coals from the outburst zone have in most cases a higher mean local complexity and a lower mean local entropy compared to the images of coals from the quiet zone of the seam (Fig. 3, *a*). Modal values, as seen in Fig. 3, *b*, practically do not find this separability.
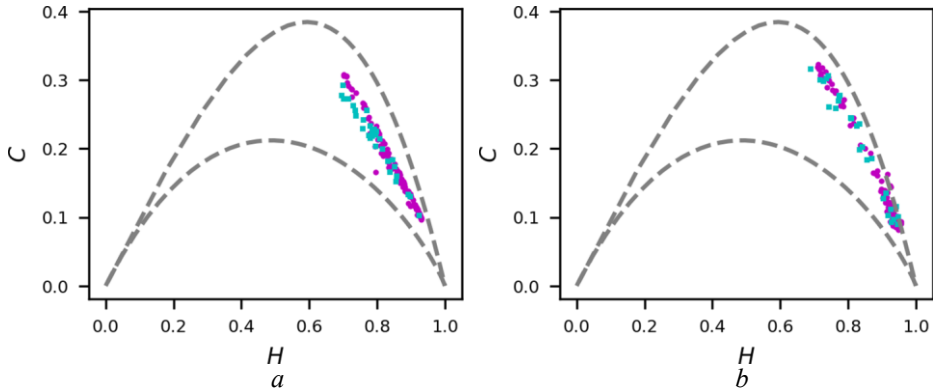

Fig. 3. The distribution of mean (*a*) and modal (*b*) local entropy H and complexity C values, calculated for SEM–images of coals from the outburst zone (■) and the quiet zone of the seam (●).

Comparing the obtained results with the results of multifractal analysis [22, 23] it can be noted that both approaches indicate a more complex surface structural organization for coals from the outburst zone compared with coals from the quite zone of the seam. This is manifested by the mean values of measures of local statistical complexity and in the spectra of fractal dimensions. The discovered regularities require further study and confirmation on a larger series of test samples to reveal the potential, which, as our studies have shown, is inherent in the described methods of SEM–image processing.

5 Conclusions

The following conclusions can be drawn from the present study:

- In most cases, the two-dimensional distributions for measures of complexity and entropy are informative data sets for differentiating coals by degree of complexity. Moreover, such characteristics of these distributions as mathematical expectation and, to a less degree, mode can be used as simple quantitative descriptors of coals with various degrees of disturbance. These characteristics can be used to show the closeness of the spatial structure for the analyzed coal specimens to strictly periodic or absolutely chaotic ones.

- By analyzing more than 140 test images of coal specimens from the quite zone of the seam and the outburst zone, it was found that most of the studied specimens exhibit a random (not chaotic) nature of the distribution of anisotropic properties.

- In most cases, images of coals from the outburst zone have a higher mean local complexity and a lower mean local entropy compared to images of coals from the quite zone of the seam.

References

1. Y.N. Malyshev, K.N. Trubeckoj, A.T. Ajruni, *Fundamental applied methods for solving the problem of coal bed methane* (Academy of Mining Sciences, Moscow, 2000).
2. A.V. Astahov, A.A. Belyj, D.L. Shirochin, Solid fuel chemistry, 4, 15–24 (2000).
3. O.N. Malinnikova, V.N. Odintsev, D.V. Uchaev, Dm.V. Uchaev, Proc. of the All–Russian scientific conference "Geodynamics and stress state of the Earth's interior", dedicated to the 80th anniversary of academician M.V. Kurleni, **2**, 93–99 (2011).
4. V.N. Zakharov, O.N. Malinnikova, Journal of Mining Institute, **210**, 43–52 (2014).
5. L.O. Kirichenko, Radio Electronics and Informatics, 3, 131–134 (1998).
6. R. López–Ruiz, H.L. Mancini, X. Calbet, Physics Letters A., **209**(5), 321–326 (1995).
7. D.L. Donoho, G. Kutyniok, *SampTA'09*, 2009.
8. K. Guo, G. Kutyniok, D. Labate, *Sparse multidimensional representations using anisotropic dilation and shear operators*, In: *Wavelets and Splines* (Nashboro Press, Nashville, 2006).
9. G. Kutyniok, D. Labate, J. Wavelet Theory and Appl. 1, 1–10 (2007).
10. D. Labate, G. Easley, W. Lim, Applied Computational Harmonic Analysis, **25**, 25–46 (2008).
11. D. Labate, W.–Q. Lim, G. Kutyniok, G. Weiss, SPIE Proc. **5914**, 254–262 (2005).
12. M.A. Kuraka, K.V. Simonov, N.O. Kudrya, Educational resources and technologies, **2**(14), 186–192 (2016).
13. A. Brazhe, Phys. Rev. E., 97(6), 061301 (2018).
14. S. Häuser, G. Steidl, Fast Finite Shearlet Transform, arXiv:1202.1773 [math] (2012).
15. O.A. Rosso, H.A. Larrondo, M.T. Martin, A. Plastino, M.A. Fuentes, Phys. Rev. Lett., **99**(15), 154102 (2007).
16. M.T. Martin, A. Plastino, O.A. Rosso, Physics Letters A., **311**(2), 126–132 (2003).
17. https://github.com/abrazhe/shearlexity/.
18. https://github.com/grlee77/PyShearlets.
19. P.W. Lamberti, M.T. Martin, A. Plastino, O.A. Rosso, Physica A: Statistical Mechanics and its Applications. **334**(1), 119–131 (2004).
20. L. Zunino, H.V. Ribeiro, Chaos, Solitons & Fractals, **91**, 679–688 (2016).
21. O.N. Malinnikova, V.N. Odintsev, V.A. Trofimov, Methane/GIAB, SI 11, 189–204 (2009).
22. V. Malinnikov, V. Zakharov, D. Uchaev, Dm. Uchaev, O. Malinnikova, *E3S Web of Conferences*, **56**, 01020 (2018).
23. O. Malinnikova, Dm. Uchaev, D. Uchaev, V. Malinnikov, *E3S Web of Conferences*, (2019), this volume.