ISSN: 2054-7404
Archives of Business Research; Vol. 14 No. 07 (2026) (12-37)
https://doi.org/10.14738/abr.1407.11909

Definition of the Concept of Force in Geopolitics

K.E. Plokhotnikov

1. Faculty of Physics, M.V.Lomonosov Moscow State University, Leninskie Gory, Moscow 119991 Russia
2. Financial University under the Government of the Russian Federation, address: 49 Leningradsky Prospekt, 125993, Moscow, Russia

Abstract: The paper introduces the concept of force in geopolitics. Force acts as a measure of confrontation between low-ranking and high-ranking clusters, associations, and aggregates of land regions on the Earth's surface. The force is calculated both for the globe as a whole and for individual states. Geopolitics is determined based on the author's previous work, which took into account climatic data in terms of temperature and precipitation, terrain, and traffic costs for transporting one conventional unit of cargo from one point to another on the entire globe. Individual traffic cost profiles have been calculated for each of the territories of the selected population. According to these profiles, the procedure of association (clusterization) in the association of individual territories is carried out. The slicing of territories with subsequent clustering is considered in the context of counting the number of low-ranking and high-ranking clusters and assessing force as measures of confrontation between low-ranking and high-ranking clusters. The highest form of manifestation of large-scale clusters was the so-called "imperial towers" or simply "towers". There were 24 such towers, all of them identified and mapped on the current political map of the world. The first top 5 territories with the highest force parameter were as follows: Antarctica, Russia, Canada, USA, China.

Keywords: force, climate, relief, traffic, mathematical model, cluster analysis, statistical experiment.

INTRODUCTION

Earlier in the author's work [1], a mathematical model of geopolitics was presented. The central concept of "habitat capacity" was introduced in the model. This concept made it possible to separate the political conjuncture represented by peoples, their territories of residence, as well as current political preferences from geopolitics, which is determined by objective geophysical circumstances. Geopolitics is the climate, topography, logistics specifics of global commodity flows, and the geopolitical situation in terms of "sea-continent", i.e. everything that makes up the material set of conditions accompanying the existence of the inhabitants of the Earth. This complex, to a large extent, mediates the behavior of the population from a political point of view. The model of geopolitics clarifies and generalizes the generally accepted geopolitical classification of territories in terms of orientation and positioning either at sea or on the continent. A mathematical model of minimizing transportation costs is formulated for an arbitrary number of points on the Earth's surface acting as logistics centers.

In this paper, the concept of "force" is introduced and discussed. A similar concept was introduced by the author earlier [2,3] and was not considered in the context of geopolitics. Force served as a measure of confrontation between supporters of two mutually exclusive goals in global politics. These goals were introduced in the normative model of global history, they were named [4], “the kingdom of freedom” and “the kingdom of necessity” respectively. Force, as a measure of the confrontation between supporters of two global meta-historical goals, is important in terms of assessing the measure of sustainability of both the global political system and its individual components.

As it turns out later, the analogues of the kingdoms of freedom and necessity in geopolitics are groups of territories positioned on the “sea” and “continent”, respectively. In geopolitical terms, force acts as a measure of the “sea – continent” confrontation.

PRESENTATION OF INITIAL DATA

In the author's previous works, a mathematical model of what is called differently in different disciplines is constructed: habitat capacity, living space, etc.

Let us determine the density of the habitat capacity, ρ, which is a function of such climatic characteristics as the average annual temperature, T, and the average annual precipitation, Q. The following functional dependence has been considered in the author's previous works:

$$\rho = \rho(T, Q) = e^{a_1 + a_2 |T/T_{\text{comfort}} - 1|^{\alpha} + a_3 |Q/Q_{\text{comfort}} - 1|^{\beta}}, \tag{1}$$

where $a_1, a_2, a_3, \alpha, \beta, T_{\text{comfort}}, Q_{\text{comfort}}$ — some constants.

The constants in formula (1) were estimated taking into account real climate data taken from the website [5]. Suitable data were selected for the average annual temperature [6] and precipitation [7] (data dating back to 1995), they were presented in the form of 360×720 matrices, which corresponds to a latitude and longitude resolution of 0.5^0.

Figure 1 shows the distribution of the average annual temperature, and Figure 2 shows the average annual precipitation in the form of three – dimensional surfaces in latitude – longitude coordinates.

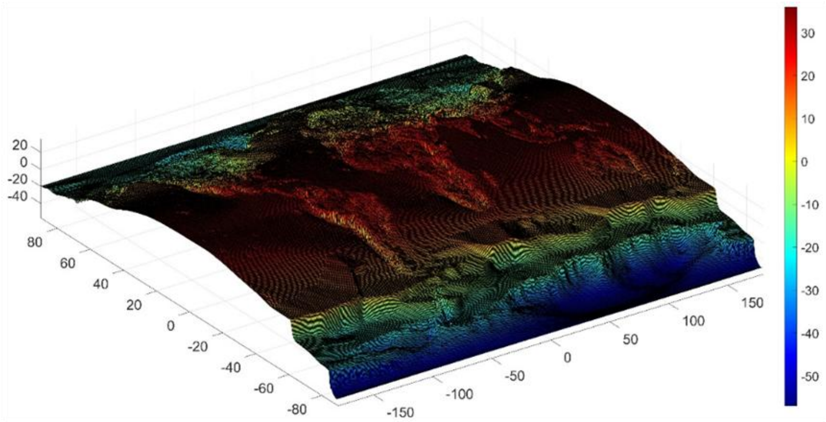


Figure 1: Distribution of the average annual surface temperature in latitude – longitude coordinates (in degrees Celsius)

The comfortable values of temperature, $T_{comfort}$ and precipitation, $Q_{comfort}$ were found using a weighted average procedure, where the Earth's population density, given on the resource [8], acted as weights. Based on the materials of the specified resource, the surface of the decimal logarithm of population density (as of 2000), which was expressed in units of $people/km^2$, is plotted in Figure 3. It follows from the map that the density varies from 0 to 10^4 $people/km^2$. It turned out that $T_{comfort} = 18.5775^0\text{C}$, $Q_{comfort} = 1200.2\ mm$.

The parameters a_1, a_2, and a_3 in formula (1) were calculated earlier in the author's previous work, based on the fact that formula (1) defines a regression model for calculating the density of habitat capacity. As a result, it turned out that $a_1 = 3.4189$, $a_2 = -1.6624$, $a_3 = -1.4789$. The values of the α and β parameters were also selected. It was assumed that $\alpha = 0.2$, $\beta = 0.2$.

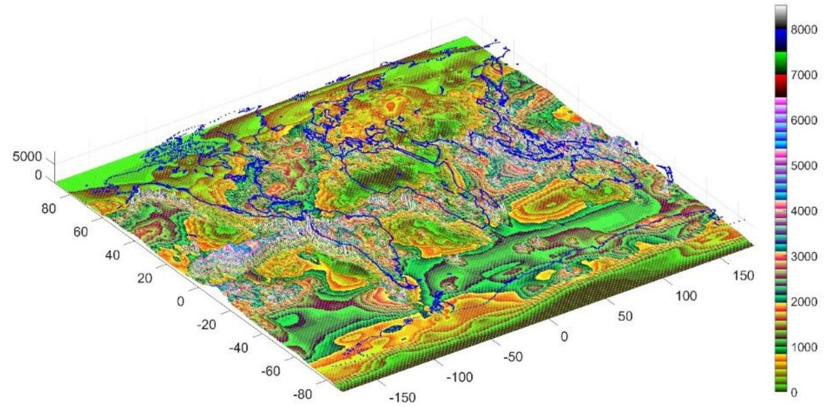


Figure 2: Distribution of average annual precipitation as a surface in latitude – longitude coordinates (in millimeters)

The final surface of the habitat capacity density, ρ, depending on latitude and longitude is shown in Figure 4. The range of variation in habitat capacity density was $[0,12.83]$. Note that the absolute values of the density of the habitat capacity are not important, only their relations at different points of the Earth's surface are important.

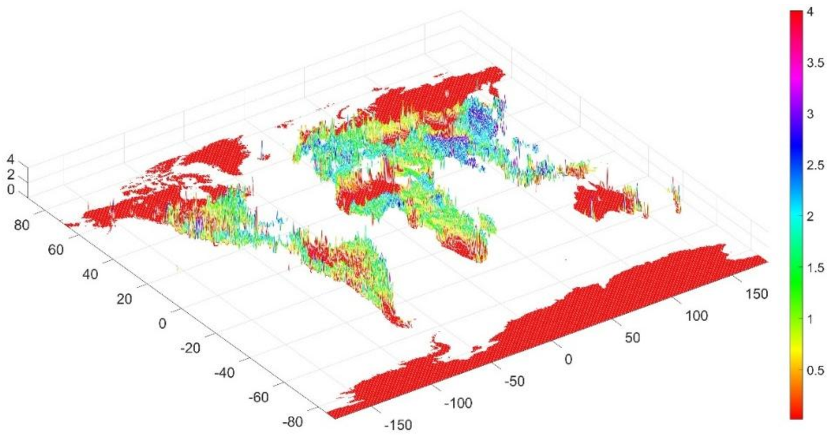


Figure 3: Decimal logarithm of population density expressed in units of $people/km^2$

Comparing the population density surfaces (Figure 3) and the carrying capacity density (Figure 4) as a whole, their topographic similarity is noteworthy. This closeness was possible because the population density was logarithmized and subsequently found to be visually comparable to the carrying capacity density. This is also confirmed by the correlation coefficient estimates for the logarithmized population density and carrying capacity density, Pearson's, Kendall's, and Spearman's values of 0.5789, 0.7610, and 0.9058, respectively.

The Spearman index has a particularly high correlation coefficient, which is 0.9058. Considering that the Spearman correlation coefficient is an indicator of the relationship of rank variables, the desired estimate of the density of habitat capacity (1), without taking into account the absolute values of the indicator, quite adequately describes the current population density of the Earth.

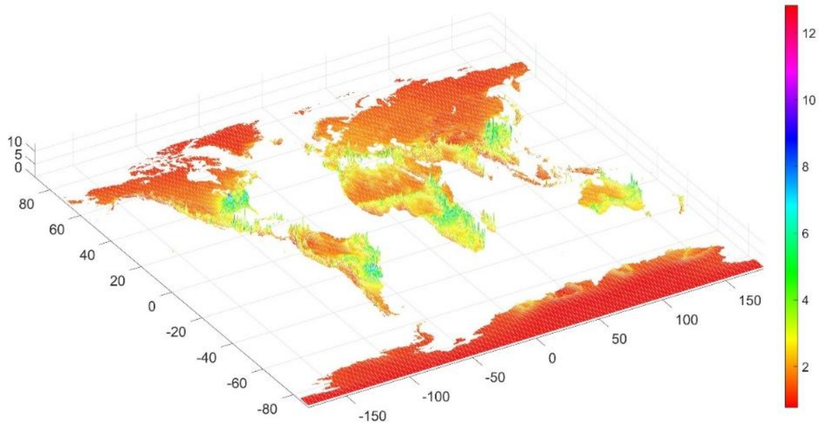


Figure 4: Distribution of habitat capacity density in latitude – longitude coordinates

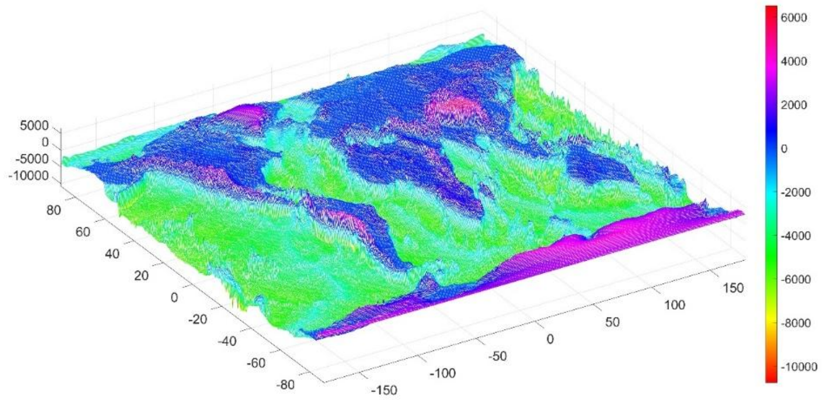


Figure 5: Relief of the Earth's surface relative to sea level in latitude – longitude coordinates

As already noted in the author's previous work, peak population density values are associated with minimizing transportation costs for moving goods within the Earth's surface. Transportation costs are determined by the energy costs of moving a conventional unit of cargo from one point to another. When calculating transportation costs, it is necessary to take into account the relief of the Earth's surface. Let's introduce a notation for the height matrix (in meters), $Z(360 \times 720)$ above sea level [9]. Figure 5 shows the surface of the relief of the Earth's surface without taking into account the oceans in latitude – longitude coordinates.

ARBITRARY CUTTING OF THE LAND AREA OF THE EARTH'S SURFACE

Taking into account the carrying capacity of the habitat, we "cut" the land area of the Earth's surface. Cutting refers to defining N land fragments. Generally, the number of fragments can vary from one to infinity, i.e., $N = 1, 2, \ldots$ For $N = 1$, the entire land area of the Earth is considered to be a single fragment. This is a trivial case and will not be considered further, since with such a cut, we have no transport costs associated with exchange between territorial fragments. For this reason, we will henceforth assume that $N \geq 2$.

It should be noted that the arbitrary division into N fragments for $N \geq 2$ is a highly ambiguous procedure, as an infinite number of divisions with the same number of fragments, N, can exist. Against this backdrop of ambiguity in land area divisions, it is necessary to identify some stable, statistically invariant configurations of land area fragments. The aforementioned statistical search for invariant land area configurations involves the use of statistical trials or the Monte Carlo method.

Let's formulate an algorithm for randomly distributing N points on the land surface, taking into account the carrying capacity of the habitat. We will define Voronoi polygons for each point. Based on these polygons, one of the possible partitions of the Earth's land surface will be constructed. In this paper, the algorithm for randomly distributing points on the land surface will differ significantly from the one discussed in the author's previous work.

Let's define N points within the habitat on the Earth's surface with coordinates $z_i = (\varphi_i, \lambda_i)$, $i = 1, \ldots, N$, where φ_i is the latitude, λ_i is the longitude of the i-point. In the future, we assume that latitude and longitude change accordingly, or within $[-90^0, 90^0]$, $[-180^0, 180^0]$, either $-$ $[-\pi/2, \pi/2]$, $[-\pi, \pi]$, when degrees or radians are selected when measuring angles. We will assume that all points are localized within the land area, which we will designate as *Land*.

Let each point z_i correspond to a certain area ω_i. As a result of these considerations, the global capacity of the habitat,

$$U = \iint_{Land} d\varphi d\lambda \cdot \cos\varphi \cdot \rho(T(\varphi, \lambda), Q(\varphi, \lambda)), \tag{2}$$

can be rewritten in the form:

$$U = \sum_{i=1}^{N} u_i \, , u_i = \iint_{\omega_i} d\varphi d\lambda \cdot \cos\varphi \cdot \rho. \tag{3}$$

Taking into account (2), (3), it is clear that $Land = \bigcup_{i=1}^{N} \omega_i$, i.e., the set of regions $\{\omega_1, \ldots, \omega_N\}$ defines a certain subdivision into subdomains of the entire land surface of the earth.

Let's take into account the fact that our data is determined on the Earth's surface with a resolution of 0.5^0. Assuming that $\varphi_i = \frac{\pi}{180}(-89.75 + 0.5(i-1))$, $\lambda_j = \frac{\pi}{180}(-179.75 + 0.5(j-1))$, $i = 1, \ldots, 360$, $j = 1, \ldots, 720$, a grid with nodes in at points $g_{i,j} = (\varphi_i, \lambda_j)$, $i =$

$1,\ldots,360$; $j = 1,\ldots,720$. There will be N_{Sphere} total of such nodes $360 \times 720 = 259'200$. Among all the nodes of the N_{Sphere} grid, we select those that relate to land. Let the value N_{Land} denote the number of nodes of the grid, which refers to land.

Let's introduce notation for coordinates of points on the land surface: $z_k = (\varphi_k, \lambda_k)$, $k = 1,\ldots,N_{Land}$. Using formula (1), we calculate the density of habitat capacity ρ_k, $k = 1,\ldots,N_{Land}$ in each of the selected land points. Let's determine the probabilities of random selection of each of the land points according to the following discrete distribution:

$$p_k = \frac{\rho_k \cos\varphi_k}{\sum_{l=1}^{N_{Land}} \rho_l \cos\varphi_l},\ k = 1,\ldots,N_{Land}. \qquad (4)$$

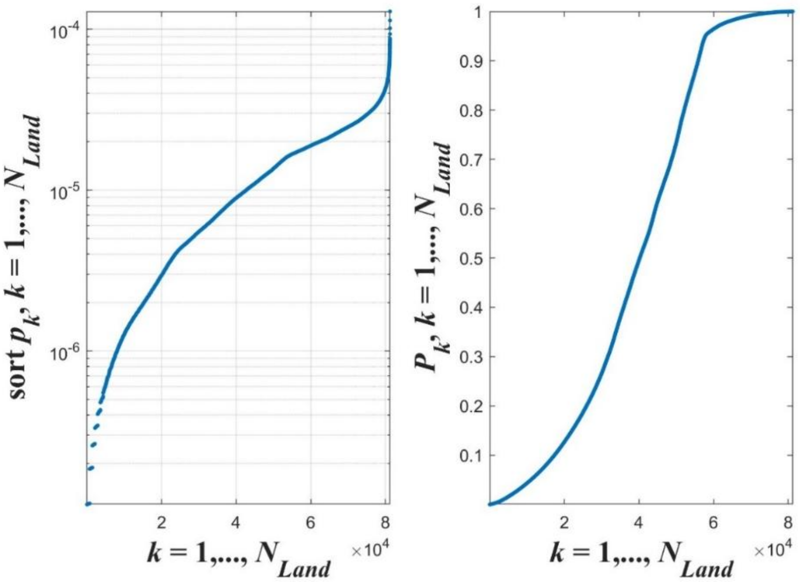


Figure 6: Discrete probability distribution (left graph), cumulative probability distribution (right graph)

The distribution (4) is indeed a discrete probability distribution, since $p_k \geq 0$, $k = 1,\ldots,N_{Land}$, and $\sum_{k=1}^{N_{Land}} p_k = 1$. Figure 6 (left graph, logarithmic scale on the ordinate) shows the distribution of the ranked probabilities (4).

It turned out that $N_{Land} = 81'192$, while the ratio $N_{Land}/N_{Sphere} \cong 0.3132$ is somewhat larger than the known ratio of land and ocean surfaces, which is equal to 0.292. This difference can be attributed to the use of a relatively coarse grid of longitudes and latitudes with a resolution of 0.5^0.

Taking into account (4), we can use the Monte Carlo method to simulate the random positioning of N points within the land area. To refine the random selection procedure, we define the cumulative probability of the discrete random variable (4) as a sequence $P_k = \sum_{l=1}^{k} p_l$, $k = 1,\ldots,N_{Land}$, where $P_1 = p_1$,…, $P_{N_{Land}} = 1$. Figure 6 (right graph) shows the distribution of the cumulative probability of the discrete random variable (4) as a graph of the sequence $\{P_1,\ldots,P_{N_{Land}}\}$

Let's choose a uniformly random number ξ from the interval $[0,1]$, and let's assign it the point number $k(\xi) \in \{1,\ldots,N_{Land}\}$ according to the following formula:

$$k(\xi) = \begin{cases} 1, \ \xi < P_1; \\ 2, \ P_1 \le \xi < P_2; \\ \dots \\ N_{Land}, P_{N_{Land}-1} \le \xi < P_{N_{Land}}. \end{cases} \tag{5}$$

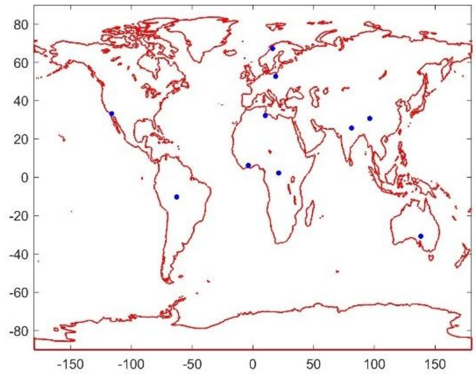


Figure 7a: Sample size $N = 10$

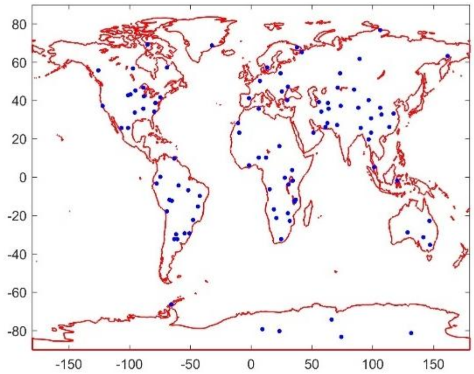


Figure 7b: Sample size $N = 10^2$

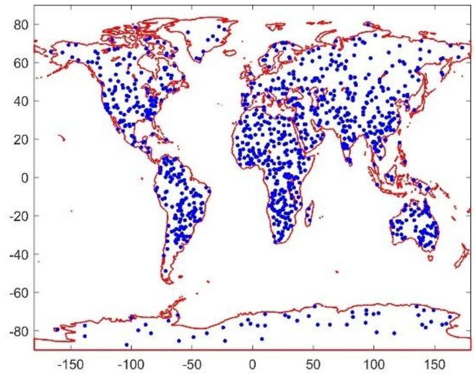


Figure 7c: Sample size $N = 10^3$

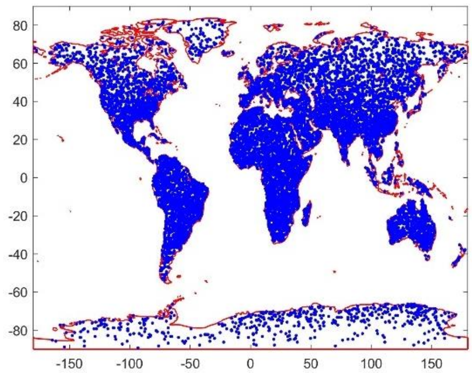


Figure 7d: Sample size $N = 10^4$

As a result, taking into account (4), (5), an algorithm for generating a random set of N points $\{z_{k_1}, \dots, z_{k_N}\}$ from the set of available $\{z_1, \dots, z_{N_{Land}}\}$ is defined. In the general case, among a sample of N points $\{z_{k_1}, \dots, z_{k_N}\}$, some of them may be repeated. Considering that in the future the sample of points will not exceed the value of 10^4, i.e. $N \le 10^4$, and $N_{Land} = 81'192 \gg 10^4 \ge N$, we will ignore possible repetitions of points in the sample.

Figure 7 shows examples of random samples of land points with lengths of 10, 10^2, 10^3, and 10^4, respectively. All four figures (7a, 7b, 7c, and 7d) were generated using the same algorithm (4) and (5).

Figure 8 shows an example of a random distribution using the Monte Carlo method within the framework of the procedure (4), (5) for $N = 3 \cdot 10^3$ points on the land surface. The corresponding Voronoi polygons are also shown, taking into account the coastline and poles. The random points are represented by blue dots, and the Voronoi polygons are marked with random colors.

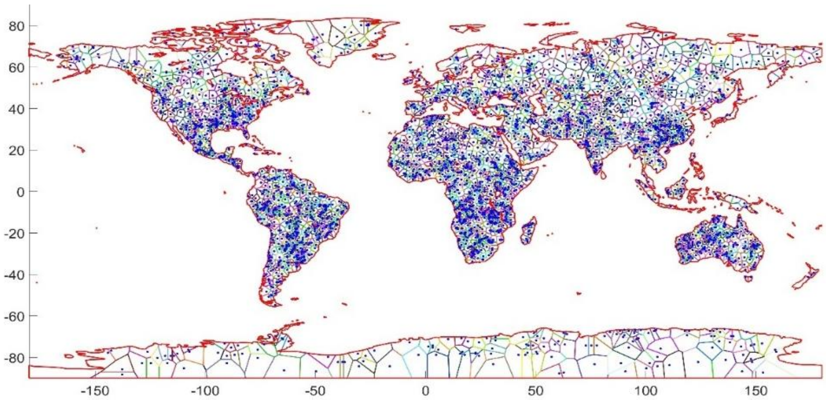


Figure 8: Map of the Earth's surface with a random set of $N = 3 \times 10^3$ points within the land and Voronoi polygons

When constructing Voronoi polygons in latitude – longitude coordinates, a feature associated with the poles occurred. Four additional "fictitious" points were introduced: $z_{N+1} = \chi(90^0, 180^0)$, $z_{N+2} = \chi(-90^0, 180^0)$, $z_{N+3} = \chi(-90^0, -180^0)$, $z_{N+4} = \chi(90^0, -180^0)$, where $\chi = \text{const} > 0$. In the calculations, the parameter χ was chosen equal to 5. Then, Voronoi polygons were found based on the extended set of points $z_i = (\varphi_i, \lambda_i)$, $i = 1, \dots, N$, $N+1$, $N+2$, $N+3$, $N+4$. In this procedure, finite Voronoi polygons were guaranteed to include the coastline, while unbounded polygons were placed in a fictitious region.

In the calculation, the results of which are shown in Figure 8, were calculated: the share of the area of the Voronoi polygons, S_{Land}, as well as the habitat capacity, U_{Land}, on land. It turned out that $S_{Land} = 0.2910$, $U_{Land} = 6.2179$. Thus, such an indicator as the proportion of land on the Earth's surface, $S_{Land} = 0.2910$, turned out to be close to the known indicator.

DETERMINATION OF TRAFFIC BETWEEN TERRITORIES

Following the previous work of the author, we will determine the traffic between the territories in connection with the energy costs of it. We will introduce a "generalized distance" between territories, which is associated with taking into account the terrain and the costs of overcoming it due to traffic. If, among other things, in the previous work the author solved the problem of minimizing transport costs, then in this work the problem of combining territories into groups according to the type of similar formats of traffic costs in the global trade exchange is solved. It is along this path that the definition and calculation of force is expected.

First, we will introduce cost-free traffic. Let us define the matrix $\{\alpha_{i,j}\}$, $i, j = 1, \dots, N$, the elements of which characterize the transfer of a fraction of the i-th capacity of the habitat, u_i, to the j-th capacity of the habitat, u_j in connection with traffic. The matrix $\{\alpha_{i,j}\}$, $i, j = 1, \dots, N$ acts as a measure of the exchange of habitat capacity between a set of territories associated with points $z_i = (\varphi_i, \lambda_i)$, $i = 1, \dots, N$ on the Earth's surface. According to the definition, it follows that

$$\sum_{j=1}^{N} \alpha_{i,j} = 1, \tag{6}$$

where $\alpha_{i,i}$ is the fraction of the i-th habitat capacity that is not subject to traffic, i.e. it is the fraction that remains inside the i-th territory. The set of habitat capacities of territories u_i, $i = 1, \ldots, N$ acts as a certain resource, which in the context of traffic is considered universal and additive.

Now, when moving part of the habitat capacity from point i to point j, the share of the resource $\gamma_{i,j}$ is lost due to traffic, we assume that $0 \le \gamma_{i,j} \le 1, i \ne j;\ i,j = 1, \ldots, N$. We assume that there are no costs within the logistics point, i.e. $\gamma_{i,i} = 0, i = 1, \ldots, N$. The matrix $\{\gamma_{i,j}\}, i,j = 1, \ldots, N$ is called the transport cost matrix. Taking into account (6), we will make up the functional, Tr of all transport costs:

$$Tr = 2\sum_{i=1}^{N}\sum_{j=1, j\ne i}^{N}\alpha_{i,j}\gamma_{i,j}u_i. \tag{7}$$

For further study of the functional of transport costs (7), additional clarifications of the type of matrices $\{\alpha_{i,j}\}, \{\gamma_{i,j}\}, i,j = 1, \ldots, N$ are necessary. Let us consider the so-called gravity model, well-known in the theory of traffic flows [10], which introduces a certain "generalized distance", $d_{i,j}$ ($d_{i,j} \ge 0$) between the i-th and j-th territories. Note that the distance matrix $\{d_{i,j}\}$, $i,j = 1, \ldots, N$, generally speaking, is not symmetric, i.e. $d_{i,j} \ne d_{j,i}$, whereas by definition it is assumed that $d_{i,i} = 0$, $i = 1, \ldots, N$. Note that the condition $d_{i,j} = 0$ does not necessarily mean that the points z_i and z_j coincide in physical space.

Let's choose the exponential dependence of the coefficients of the matrices $\{\alpha_{i,j}\}$, $\{\gamma_{i,j}\}$, $i,j = 1, \ldots, N$ on the generalized distance, then we can write the following representations:

$$\alpha_{i,j} = \frac{e^{-\beta d_{i,j}}}{\sum_{k=1}^{N} e^{-\beta d_{i,k}}},\ \gamma_{i,j} = \tfrac{1}{2}r(1 - e^{-\beta_2 d_{i,j}}), \tag{8}$$

where β, β_2, r are some non-negative parameters. By direct verification, one can verify that the matrix $\{\alpha_{i,j}\}$, $i,j = 1, \ldots, N$ in the form (8) fulfills condition (6). According to (8), there are no transport costs for $d_{i,j} = 0$, i.e. $\gamma_{i,j} = 0$. Finally, for $\beta_2 \to \infty$ and for $d_{i,j} \ne 0$, it follows that $\gamma_{i,j} \to \frac{1}{2}r$, i.e. the share of transport costs becomes a constant value equal to $\frac{1}{2}r$. From the last remark it follows that $0 \le r \le 2$.

Substitute (8) into (7), then we find the following expression for the total cost of transportation:

$$Tr = r\sum_{i=1}^{N}[1 - \sum_{k=1}^{N} e^{-(\beta+\beta_2)d_{i,k}} / \sum_{k=1}^{N} e^{-\beta d_{i,k}}]u_i. \tag{7'}$$

The following properties, which are directly verifiable, are characteristic of transportation costs in the form (7′). First, when the distance between the points becomes zero, i.e. $d_{i,j} = 0, i,j = 1, \ldots, N$, there are no transport costs, $Tr = 0$. Secondly, when the distance between points tends to infinity, i.e. $d_{i,j} \to \infty, i \ne j; i,j = 1, \ldots, N$, transport costs due to the selected dependencies (8) also tend to zero, $Tr \to 0$. Note that a distance equal to zero or tending to infinity between a pair of points does not mean that the points merge or diverge to infinity.

Let us consider transport costs in the form (7′) as a function of the parameter β_2, i.e. $Tr = Tr(\beta_2)$. In this case, it is obvious that $Tr(0) \equiv 0$. Now let $\beta_2 \to \infty$, then $Tr(\beta_2) \to Tr(\infty) = r\sum_{i=1}^{N}[1 - 1/\sum_{k=1}^{N} e^{-\beta d_{i,k}}]u_i$. For the last functional of transport costs, it is obvious that the global minimum, equal to zero, is achieved only when the distances between points tend to infinity. It is the last version of the functional that will be considered further. In this case, the points will not be able to gather together; they will "repel" each other and fill the maximum of the habitat, since they are "locked" on the surface of the earth's sphere.

In the author's previous work, the condition for the limit transition $\beta_2 \to \infty$ was called the "*minimax transport doctrine*", which is deciphered according to the formula: *minimum transport costs with maximum filling of the capacity of the habitat*.

Taking into account the minimax transportation doctrine, the global traffic costs (7′) can be rewritten as:

$$Tr = \sum_{i,j=1}^{N} Tr_{i,j}, \; Tr_{i,j} = ru_i \mathrm{A}_{i,j}, \tag{7''}$$

where $\mathrm{A}_{i,j} = \alpha_{i,j}(1 - \delta_{i,j})$ $i, j = 1, \dots, N$, $\delta_{i,j}$ is the Kronecker delta, i.e. $\delta_{i,i} = 1$, $\delta_{i,j} = 0$, when $i \neq j$. Note that, according to (7″), $Tr_{i,i} = 0$, since $\mathrm{A}_{i,i} = 0$, $i = 1, \dots, N$. It is obvious that the elements of the matrices $\{\mathrm{A}_{i,j}\}$, $i, j = 1, \dots, N$ coincide with the elements of the matrix $\{\alpha_{i,j}\}$, $i, j = 1, \dots, N$ except for the diagonal elements, where there are zeros.

The elements of the transport cost matrix $\{Tr_{i,j}\}$, $i, j = 1, \dots, N$ in (7″) have a transparent physical meaning. They represent the portion of the habitat capacity of the i-th territory that is spent on traffic to the j-th territory. Note that the habitat capacities of all land territories $\{u_i\}$, $i = 1, \dots, N$ act as a universal resource providing traffic between territories. As a result, for each i-th territory, we can construct a traffic cost vector $Tr_{i,:} = (Tr_{i,1}, \dots, Tr_{i,N})$. It remains to compare the vectors $Tr_{1,:}$,…, $Tr_{N,:}$ with each other and gain an understanding of how the territories are connected. Let us now present the method for constructing the matrix of generalized distances $\{d_{i,j}\}$, $i, j = 1, \dots, N$.

CALCULATION OF THE GENERALIZED DISTANCE MATRIX

To determine the algorithm for calculating the distance matrix $\{d_{i,j}\}$, $i, j = 1, \dots, N$, we will express a number of physical considerations about the energy costs of moving one conventional unit of cargo weight from point $z_1 = (\varphi_1, \lambda_1)$ to point $z_2 = (\varphi_2, \lambda_2)$. Let's define the route of cargo movement in the form of a line: $\varphi = \varphi(s)$, $\lambda = \lambda(s)$, $0 \leq s \leq l$. The line parameterization argument is the length of the line, s, varying from zero to its maximum value, l, equal to the length of the route, while it is assumed that $\varphi_1 = \varphi(0)$, $\lambda_1 = \lambda(0)$ and $\varphi_2 = \varphi(l)$, $\lambda_2 = \lambda(l)$.

Let's assume that the route is completely located on the land surface at the beginning. In this case, the energy cost of moving one conventional unit of cargo weight consists of three characteristic movement options: 1) vertical movement of the load upward due to the features of the relief; 2) vertical lowering of the load due to the features of the relief; 3) movement of the load along an inclined surface. The first two points of the movement options describe the energy costs of moving the load "up and down". The last point is characterized mainly by the energy costs of overcoming rolling friction in such modes of transport as road and rail. Let $Z = Z(\varphi, \lambda)$ be the relief of the land surface, then for the chosen route of movement of the conventional unit of mass of the cargo, the function $Z(s) = Z(\varphi(s), \lambda(s))$, $0 \leq s \leq l$ can be written.

Taking into account the above physical considerations, we write down the formula for calculating the generalized distance along the selected trajectory between a pair of points z_1 and z_2, d_{earth}:

$$d_{\text{earth}} = g_1 \int_0^l \frac{dZ}{ds}\kappa\left(\frac{dZ}{ds}\right)ds - g_2 \int_0^l \frac{dZ}{ds}\kappa\left(-\frac{dZ}{ds}\right)ds + g_3 \int_0^l \sqrt{1-\left(\frac{dZ}{ds}\right)^2}ds, \tag{9}$$

where $\kappa(t) = 1,\ t \geq 0;\ \kappa(t) = 0,\ t < 0$ is the so-called "unit" function. The parameters g_1, g_2, and g_3 characterize the contribution of each type of motion to the movement of one conventional unit of mass of the load. The integrals included in (9) are called transport integrals in the author's previous work.

Now let's assume that the route of movement of one conventional unit of cargo mass lies entirely at sea, i.e. the points of departure, arrival, and all other points of the route lie on the surface of the sea. In this case, the energy costs of moving one unit of weight of a conditional cargo by sea are associated with overcoming viscous friction. To calculate the distance along the selected trajectory between a pair of points z_1 and z_2, d_{sea}, we can use the formula (9). We assume that water transport moves along a horizontal surface, for which we can assume that $dZ/ds = 0$. Taking into account the last integral in (9) and performing elementary integration, we write the corresponding transport integral in the form:

$$d_{\text{sea}} = g_4 l, \tag{10}$$

where g_4 is a non-negative parameter that takes into account the averaged features of viscous friction in the aquatic environment of combined water transport.

Note that an arbitrary route between the point of departure and destination can be divided into stages of movement only by land or only by sea. Applying either formula (9) or formula (10) to each stage and adding the values obtained, we will find the total distance between a pair of points. The distances calculated by formulas (9) and (10) are not distances in the usual sense of the word. Rather, they act as effective distances, which can always be measured by calculating the average energy for moving one conventional unit of cargo mass from the point of departure to the destination.

Considerations on the choice of parameter values g_1, g_2, g_3, and g_4 were described in detail in the author's previous work. We believe that

$$g_1 = 1,\ g_2 = 1,\ g_3 = 10^{-3},\ g_4 = 1.3639 \cdot 10^{-3}. \tag{11}$$

To calculate the generalized distance (9), (10) between an arbitrary pair of points on the Earth's surface $z_1 = (\varphi_1, \lambda_1)$, $z_2 = (\varphi_2, \lambda_2)$, we construct a complete circle of the Earth passing through a given pair of points. We assume that a complete circle lies in a plane passing through the center of the Earth. Let's define two single-length vectors $\boldsymbol{n}_1, \boldsymbol{n}_2$, which point to a pair of selected points $z_1 = (\varphi_1, \lambda_1)$, $z_2 = (\varphi_2, \lambda_2)$, then

$$\boldsymbol{n}_k = (\cos\lambda_k \cos\varphi_k, \sin\lambda_k \cos\varphi_k, \sin\varphi_k), k = 1,2. \tag{12}$$

Let the vector $\boldsymbol{n} = (n_x, n_y, n_z)$ of unit length point to an arbitrary point of a complete circle passing through the selected pair of points. It is clear that vector $\boldsymbol{n}$ lies in the plane formed by vectors $\boldsymbol{n}_1, \boldsymbol{n}_2$. Given (12), after simple transformations, we find

$$\boldsymbol{n} = \boldsymbol{n}(s) = \frac{\sin(\theta - s)}{\sin\theta}\boldsymbol{n}_1 + \frac{\sin s}{\sin\theta}\boldsymbol{n}_2, \tag{13}$$

where θ is the angle between two vectors $\boldsymbol{n}_1, \boldsymbol{n}_2$, and $(\boldsymbol{n}_1, \boldsymbol{n}_2) = \cos\theta$. According to (13), it is obvious that $\boldsymbol{n}(0) = \boldsymbol{n}_1$ and $\boldsymbol{n}(\theta) = \boldsymbol{n}_2$. The angle θ can be calculated using the formula:

$$\theta = \begin{cases} \pm \arccos(\boldsymbol{n}_1, \boldsymbol{n}_2), (\boldsymbol{n}_1, \boldsymbol{n}_2) \geq 0; \\ \pm\pi \mp \arccos|(\boldsymbol{n}_1, \boldsymbol{n}_2)|, (\boldsymbol{n}_1, \boldsymbol{n}_2) < 0; \end{cases} \tag{14}$$

in further calculations, the upper sign was selected.

Considering (13) and assuming that $\boldsymbol{n}(s) = (\cos\lambda(s)\cos\varphi(s), \sin\lambda(s)\cos\varphi(s), \sin\varphi(s))$, find the parametric recording the full circle in the coordinates of the "latitude - longitude":

$$\varphi = \varphi(s) = \arcsin(n_z) = \arcsin[\tfrac{\sin(\theta - s)}{\sin\theta}\sin\varphi_1 + \tfrac{\sin s}{\sin\theta}\sin\varphi_2], \tag{15}$$

$$\lambda = \lambda(s) = \operatorname{arctg}\frac{n_y}{n_x} = \operatorname{arctg}\frac{\sin(\theta - s)\sin\lambda_1\cos\varphi_1 + \sin s\sin\lambda_2\cos\varphi_2}{\sin(\theta - s)\cos\lambda_1\cos\varphi_1 + \sin s\cos\lambda_2\cos\varphi_2}. \tag{15′}$$

Note that formula (15′) is correct when $n_x > 0$. In the other two cases: 1) $n_x < 0, n_y > 0$; 2) $n_x < 0, n_y < 0$ to the expression in (15′) it is necessary to add and subtract π, respectively.

Let's construct a vector of unit length $\boldsymbol{n}_{dc}$, which points to the intersection point of the full circle with the date line. We assume that the vector $\boldsymbol{n}_{dc}$ lies in the same plane as the vectors $\boldsymbol{n}_1, \boldsymbol{n}_2$. To do this, taking into account (15′), it is necessary to solve the equation $\lambda(s) = \pm\pi$ relative to s. We will need three root $r_{dc} \in \{s_{dc}, s_{dc} \pm \pi\}$, where $s_{dc} = \operatorname{arctg}\frac{\sin\theta\sin\lambda_1\cos\varphi_1}{\cos\theta\sin\lambda_1\cos\varphi_1 - \sin\lambda_2\cos\varphi_2}$. Either the first root is selected, or the second and third, with $n_x(r_{dc}) < 0$. Between the second and third roots, the choice is made in favor of the minimum modulo value. In this case, the desired vector, $\boldsymbol{n}_{dc}$, pointing to the date line, is determined according to the equation $\boldsymbol{n}_{dc} = \boldsymbol{n}(r_{dc})$.

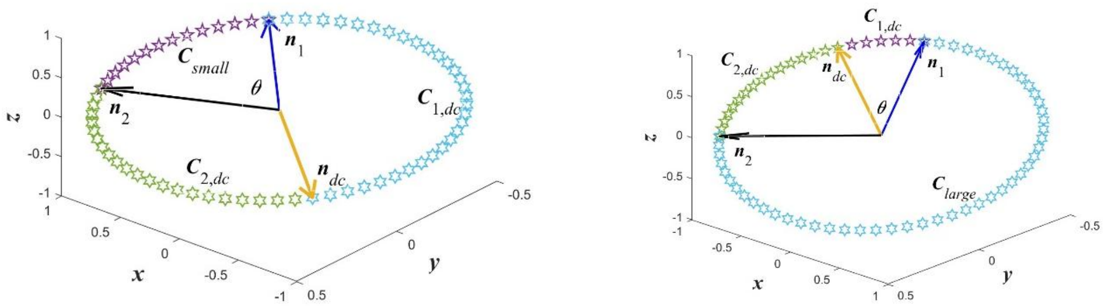


Figure 9a: The date change vector falls on a large circle fragment

Figure 9b: The date change vector falls on a small circle fragment

The small and large fragments of a circle passing through a given pair of points and the center of the Earth are described by formulas (12) — (15′), when the parameter s takes values from the segments $[0, \theta]$ and $[\theta, 2\pi]$ (or $[-2\pi + \theta, 0]$), respectively. Figure 9 shows a graphical example of the positioning of two random vectors $\boldsymbol{n}_1, \boldsymbol{n}_2$ and the corresponding date vector $\boldsymbol{n}_{dc}$. The small and large fragments of the circle are marked with markers in the form of a set of pentagrams and hexagrams, respectively. Examples of two cases are given when the date change vector falls on large (Figure 9a) and small (Figure 9b) fragments of a circle, respectively. From the point of view of the parameter s, the date change vector $\boldsymbol{n}_{dc}$ falls on a large fragment of the circle when $r_{dc} \notin [0, \theta]$ and, conversely, falls on a small fragment of the circle when $r_{dc} \in [0, \theta]$.

The vectors $\boldsymbol{n}_1, \boldsymbol{n}_2$, and $\boldsymbol{n}_{dc}$ divide the great circle into three parts. Let's assemble the small and large fragments of a large circle from the three specified parts. Let C, C_{small}, and C_{large} denote the sets of points lying on the entire circle, as well as on its small and large fragments, respectively. In this case, by definition, $C = C_{small} \cup C_{large}$, with $C_{small} = \{\boldsymbol{n}(s) | s \in [0, \theta]\}$, $C_{large} = \{\boldsymbol{n}(s) | s \in [\theta, 2\pi]\}$. Taking into account Figure 9, two cases of assembling small and large fragments of the entire circle from these three parts should be considered:

$$\begin{cases} C_{small}, C_{large} = C_{1,dc} \cup C_{2,dc}, r_{dc} \notin [0, \theta]; \\ C_{small} = C_{1,dc} \cup C_{2,dc}, C_{large}, r_{dc} \in [0, \theta]; \end{cases} \quad (16)$$

where $C_{1,dc}$, $C_{2,dc}$ – the arc of C between pairs of vectors $\boldsymbol{n}_1, \boldsymbol{n}_{dc}$ and $\boldsymbol{n}_2, \boldsymbol{n}_{dc}$ respectively.

Figure 10 shows an example of positioning $N = 50$ points on a land surface. The points are randomly selected according to the algorithm of section 3, i.e., taking into account the density of the habitat capacity. Taking into account the spherical geometry, all the $\frac{50 \times 49}{2} = 1'225$ large circles connecting each pair of points. In addition, the circles are marked in terms of their passage over land (solid black line) and at sea (dotted blue line).

According to formula (9), in order to calculate the effective distance for moving one conventional unit of mass of cargo on the Earth's surface, it is important to know the derivative of the relief $Z = Z(\varphi, \lambda)$ along the route, which we describe with a certain trajectory. For example, suppose that a route is laid between a pair of points $z_1 = (\varphi_1, \lambda_1)$, $z_2 = (\varphi_2, \lambda_2)$ such that $\varphi = \varphi(s), \lambda = \lambda(s)$, where s is a parameter describing the traveled path from the starting point. In this case, it is obvious that $\frac{dZ}{ds} = \frac{\partial Z}{\partial \varphi}\frac{d\varphi}{ds} + \frac{\partial Z}{\partial \lambda}\frac{d\lambda}{ds}$.

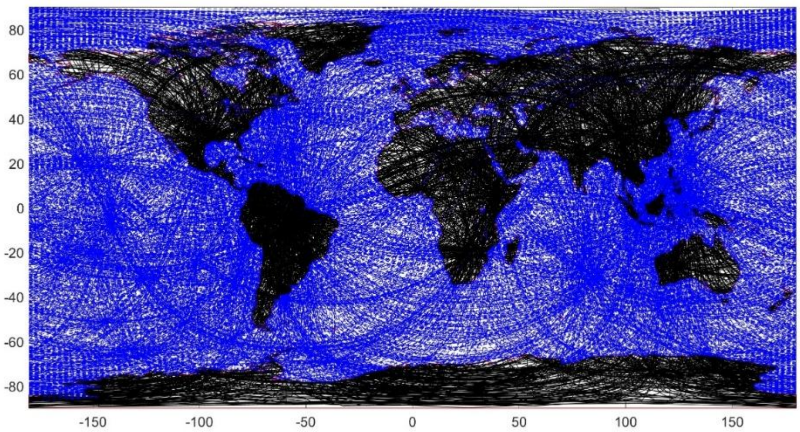


Figure 10: Projection of large circles connecting $N = 50$ points in pairs onto a latitude – longitude rectangle

As a route between a pair of points on the surface of the earth's sphere, let's choose sequentially both fragments (small and large) of a complete circle passing through a pair of points. A suitable trajectory is presented in the form of formulas (15), (15′). It remains to find the derivatives. After simple calculations, we get:

$$\frac{d\varphi}{ds}=\frac{-\cos(\theta-s)\sin\varphi_1+\cos s\sin\varphi_2}{\sqrt{\sin^2\theta-[\sin(\theta-s)\sin\varphi_1+\sin s\sin\varphi_2]^2}},\qquad (17)$$

$$\frac{d\lambda}{ds}=\frac{\sin\theta\sin(\lambda_2-\lambda_1)\cos\varphi_1\cos\varphi_2}{\sin^2\theta-[\sin(\theta-s)\sin\varphi_1+\sin s\sin\varphi_2]^2}.\qquad (17')$$

Let's introduce the distance matrix $D=\{d_{i,j},i,j=1,\dots,N\}$ between all pairs of points. Let us define similar distance matrices found for small $D_{\text{small}}=\{d_{\text{small},i,j},i,j=1,\dots,N\}$ and large $D_{\text{large}}=\{d_{\text{large},i,j},i,j=1,\dots,N\}$ fragments of arcs of corresponding complete circles. Considering the prospect of minimizing traffic costs, let's assume that the desired distance matrix D is the element-wise minimum of a pair of distance matrices D_{small} and D_{large}, i.e. $D=\min(D_{\text{small}},D_{\text{large}})=\{d_{i,j}=\min(d_{\text{small},i,j},d_{\text{large},i,j})\}$, $i,j=1,\dots,N$.

For the calculation of the transport integrals in (9), we need the partial derivatives $\frac{\partial Z}{\partial\varphi}$ and $\frac{\partial Z}{\partial\lambda}$, which we calculate using finite differences with a resolution of relief 0.5^0 according to the formula:

$$\left(\frac{\partial Z}{\partial\varphi}\right)_{i,j}=\frac{60}{\pi}(Z_{i+1,j}+Z_{i+2,j}-Z_{i-1,j}-Z_{i-2,j}),\qquad (18)$$

where $i=3,\dots,357$; $j=1,\dots,720$ and

$$\left(\frac{1}{\cos\varphi}\frac{\partial Z}{\partial\lambda}\right)_{i,j}=\frac{60}{\pi\cos\varphi_i}(Z_{i,j+1}+Z_{i,j+2}-Z_{i,j-1}-Z_{i,j-2}),\qquad (18')$$

where $i=1,\dots,360$; $j=1,\dots,720$, while assuming in (18′) a periodic continuation of the index j with a period of 720.

The ordinary derivatives $\frac{d\varphi}{ds}$ and $\frac{d\lambda}{ds}$ were found according to (17), (17′). Otherwise, the fragment of a complete circle connecting a pair of points was divided into parts running separately over land and sea. For each part, a suitable finite-difference grid was constructed with the number of nodes adapted to the length Δs of the selected part of the complete circle according to the formula $2+[n\Delta s]$. In the latter formula, the parameter n was usually chosen in the vicinity of 15, and $[\dots]$ was understood to be the integer part of the number.

Figure 11 shows an example in which $N=350$ points were randomly located within the land according to the algorithm of section 3. Some of the minimal fragments of large circles passing through the corresponding pair of points are constructed. In the current calculation, the values of the elements of the distance matrix D varied in the range $[1.19\cdot 10^{-5},0.0092]$. Note that the relief matrix Z used in formulas (18) and (18′) was divided by the radius of the Earth. For the elements of the symmetric distance matrix $D_{sym}=\frac{1}{2}(D+D^T)$, the average value and standard deviation were found, they turned out to be 0.0029 and 0.0015, respectively.

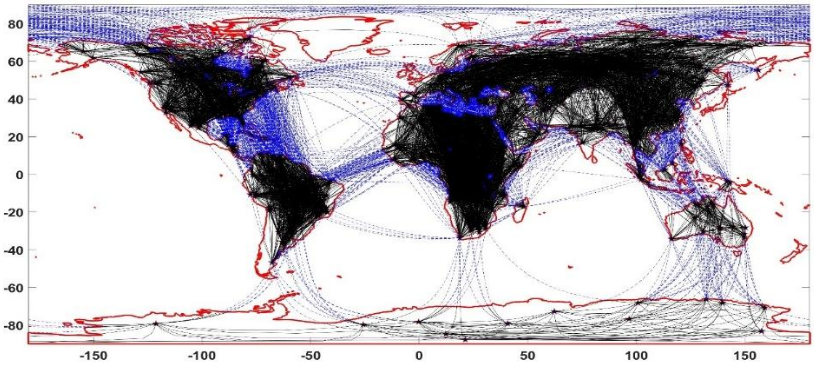


Figure 11: An example of positioning $N = 350$ points (markers in the form of pentagrams) within the land, as well as small fragments of large circles

Note that almost always $d_{small,i,j} \leq d_{large,i,j}$, $i,j = 1,\dots,N$. Taking into account all $\frac{1}{2}N(N-1)$ possible binary connections, this inequality was not fulfilled in $\cong 2.81\%$ of cases.

CALCULATION AND ANALYSIS OF THE TRANSPORTATION COST MATRIX

The formula (7″) contains an expression for calculating the amount of transportation costs $\{Tr_{i,j}\}$, $i,j = 1,\dots,N$. This expression includes an undefined parameter β. To evaluate it, we introduce a set of vectors $Tc_i = Tr_{i,:}$, $i = 1,\dots,N$, which characterize the individual traffic cost profile of each of the regions (Voronoi polygons). The set of vectors $Tc_1,\dots,Tc_N$ are the rows of the transport cost matrix $\{Tr_{i,j}\}$, $i,j = 1,\dots,N$. Let's calculate the matrix of correlation coefficients $\{Cr_{i,j}\}$, $i,j = 1,\dots N$ vectors Tc_i, $i = 1,\dots,N$, i.e. $Cr_{i,j} = Corr(Tc_i, Tc_j)$, $i,j = 1,\dots,N$, where the function $Corr(Tc_i, Tc_j)$ returns the Pearson correlation coefficient between a pair of vectors Tc_i and Tc_j.

Taking into account that the Pearson correlation coefficient is invariant with respect to the linear transformation of the data vectors, it turns out that the matrix of correlation coefficients does not depend on the product ru_i, $i = 1,\dots,N$ and can be represented as: $Cr_{i,j} = Corr(\mathrm{A}_{i,:}, \mathrm{A}_{j,:})$, $i,j = 1,\dots,N$. According to the definition in (7″), $\mathrm{A}_{i,j} = \alpha_{i,j}(1-\delta_{i,j})$, where the vectors $\mathrm{A}_{i,:}$, $i = 1,\dots,N$ are understood as the rows of the matrix $\{\mathrm{A}_{i,j}\}$, $i,j = 1,\dots,N$.

We assume the future prospect of uniting individual regions into larger associations. The criterion for forming large territorial associations is the proximity of individual traffic cost profiles. We will consider the Pearson correlation coefficient as a measure of proximity. We will calculate the number of binary correlations $N_{corr} = \sum_{1=i<j=N}(Cr_{i,j} \geq\ 0.7)$ exceeding the threshold of 0.7, above which the correlation is considered high. We will calculate and plot the value of N_{corr} as a function of the parameter β. It turns out that the function $N_{corr} = N_{corr}(\beta)$ has a maximum. In other words, the set of vectors $Tc_1,\dots,Tc_N$ is maximally correlated at some value $\beta = \beta_{max}$, at which the number of high binary correlation coefficients reaches a maximum $N_{corr,max} = \max_{\beta} N_{corr}(\beta)$.

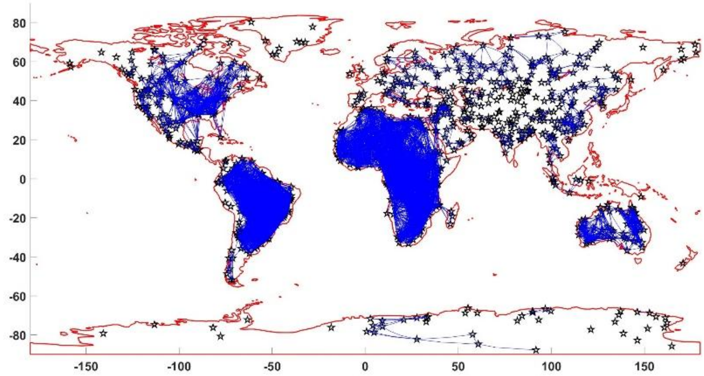


Figure 12a: Map of highly correlated pairs of points

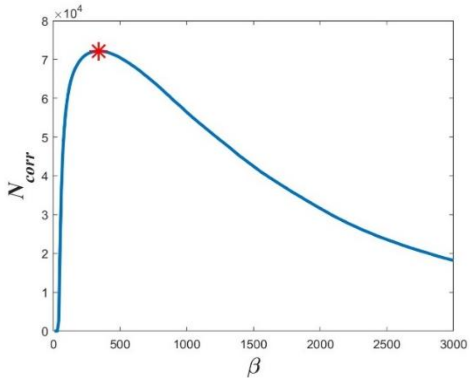


Figure 12b: Graph of the dependence of the function $N_{corr}\ (\beta)$ on β

Figure 12a shows a map of highly correlated pairs of points, they are connected by lines. For these pairs, the correlation coefficients exceeded the value of 0.92. Figure 12b shows a graph of the dependence of the $N_{corr}(\beta)$ function on β. In the presented calculation, it was assumed that $N = 10^3$, $\beta_{max} = 340$, $N_{corr,max} = 72'179$ were found. The optimal value of the parameter β_{max} is marked in Figure 12b with a red marker in the form of a star. Note that the total binary correlations are $\frac{1}{2}N(N-1)|_{N=10^3} = 499'500$. Relative number $\frac{N_{corr,max}}{\frac{1}{2}N(N-1)}$

The percentage of highly correlated relationships was 14.45%.

According to the map in Figure 12a, the proximity of individual traffic costs of selected points in the following regions is clearly expressed in decreasing order: Africa, South America, North America, etc.

To aggregate land points in an association, we will use cluster analysis. As a measure of the proximity of a pair of points according to traffic profiles, we choose the so-called "Pearson distance" $\rho_{i,j} = 1 - Cr_{i,j}$, $i,j = 1,\ldots,N$. The Pearson distance takes values from the segment $[0,2]$, while the Pearson distance is smaller the closer the correlation coefficient is to unity.

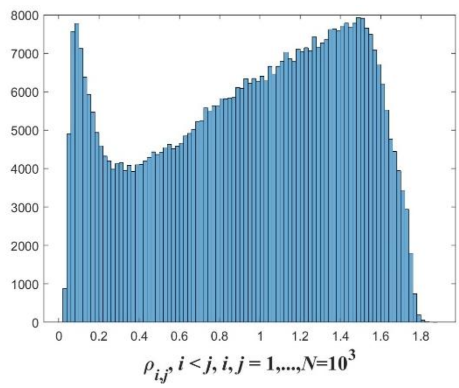


Figure 13a: Histogram of Pearson distances

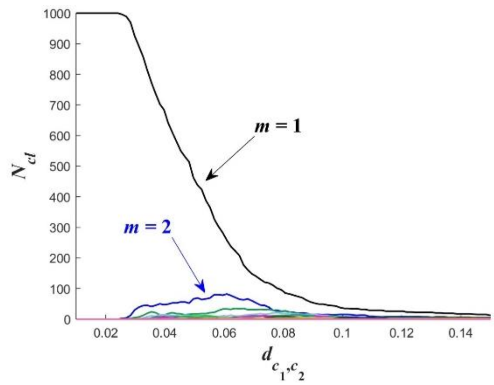


Figure 13b: Dependence $N_{cl} = N_{cl}(m, d_{c_1,c_2})$

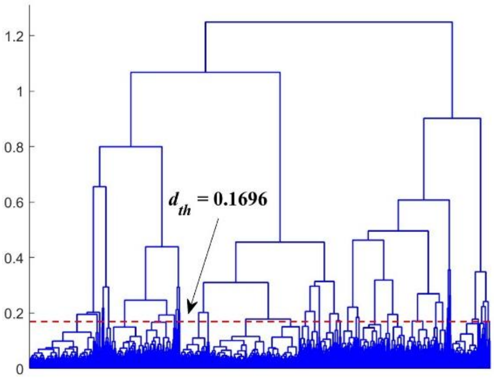


Figure 13c: Dendrogram of cluster analysis

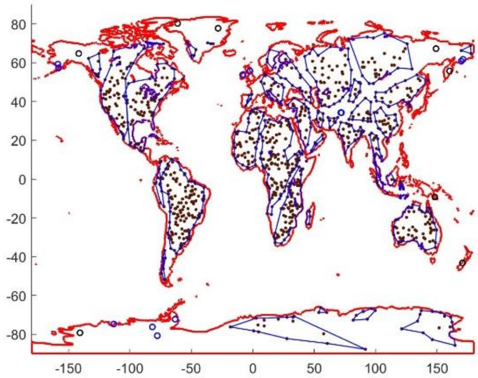


Figure 13d: An example of using cluster analysis

Figure 13a shows an example of a histogram of the distribution of values of binary Pearson distances $\rho_{1,2}, \rho_{1,3}, \ldots, \rho_{N-1,N}$ at $N = 10^3$. It is clearly seen that the binary Pearson distance is bimodal.

As is known from the theory (algorithm) of cluster analysis [11], within the framework of the agglomerative cluster formation procedure, they start with individual observations, which are considered to be initial atomic clusters. Further, in our case, using the Pearson distance, a pair is found, the distance between which is minimal, these clusters are combined into a new composite cluster. Then the procedure is repeated, and not only atomic clusters can participate in it, but also composite ones. The distance d_{c_1,c_2} between composite clusters $c_1 = \{i_1, \ldots, i_r\}$, $c_2 = \{j_1, \ldots, j_t\}$ is calculated using the average formula $d_{c_1,c_2} = \frac{1}{rt}\sum_{k=1}^{r}\sum_{l=1}^{t}\rho_{i_k,j_l}$, where r, t is the number of atomic clusters in each of the two components. As a result, step by step we come to a single, composite cluster. In the future, the cluster rank m will be understood as the number of atomic clusters that are part of it.

Let's introduce a function of the number of clusters $N_{cl} = N_{cl}(m, d_{c_1,c_2})$, which depends on two variables: 1) m is the rank of the cluster and 2) d_{c_1,c_2} is the distance between the clusters. It is clear that both the function N_{cl} itself and the variable m can take values from the set $\{1,2,\ldots,N\}$. Figure 13b shows an example of the function $N_{cl} = N_{cl}(m, d_{c_1,c_2})$, constructed for the case $N = 10^3$. The individual graphs in Figure 13b correspond to the dependence of N_{cl} on the distance d_{c_1,c_2} for different cluster ranks, $m = 1,2,\ldots,N = 10^3$.

Let's determine the number of one- and two-rank clusters $N_{1,2} = N_{cl}(1, d_{c_1,c_2}) + N_{cl}(2, d_{c_1,c_2})$. The allocation of single- and double-rank clusters is due to the fact that it is impossible to build a bordering polygon for each of them. From this point of view, their contribution will be considered small, while the number of high-ranking clusters should be as noticeable as possible. Let's assume that the number of one- and two-rank clusters is $f_r N$, i.e. $N_{1,2} = f_r N$. In the future, we will choose the following value of the fraction, f_r of one- and two-rank clusters $-$ 0.0135, i.e. $f_r = 0.0135$ or as a percentage of 1.35%.

Figure 13c shows an example of a cluster analysis dendrogram for the case $N = 10^3$. There is also a straight dotted line defined by the distance between the clusters $d_{th} = 0.1696$. The distance d_{th} is found from the condition that $N_{1,2} = f_r N$, $f_r = 0.0135$. Taking into account the graph in Figure 13b, the dependence of the number of peer clusters on

d_{c_1,c_2} is a monotonically decreasing function, so you can always find the appropriate value of d_{th} for a given value of f_r.

Fixing the distance d_{th} determines the desired cluster classification of land surface points. Figure 13d shows the result of applying cluster analysis to the traffic cost matrix at the cut-off value of the intercluster distance $d_{th} = 0.1696$, as well as at $N = 10^3$. In Figure 13d, there are clusters of very different ranks, starting with clusters of rank one, two (markers in the form of circles). Large-scale clusters with three or more atomic clusters are identified using bordering polygons.

DEFINITION OF THE CONCEPT OF "FORCE" IN GEOPOLITICS

To determine the concept of force, several calculations must be performed, similar to those summarized in Figure 13d. Each calculation involves using cluster analysis to classify points on the land surface, randomly distributed according to the procedure in section 3. Clearly, each calculation will yield its own clustering map. It is necessary to somehow find the average clustering map.

Let S calculations of the transportation cost matrix $Tr^{(s)}$, $s = 1, \dots, S$, be performed for some N. In each of the S calculations, N points on the land surface are positioned randomly according to the algorithm of section 3. We apply the clustering procedure Cl according to section 6 to each of the transportation cost matrices $\{Tr^{(1)}, \dots, Tr^{(S)}\}$. We choose some value of the parameter f_r that determines the proportion of the number of one- and two-rank clusters in relation to all selected points N. Applying the clustering procedure to the transportation cost matrix $Tr^{(s)}$ generates a set of clusters $\{\mathrm{K}_1^{(s)}, \dots, \mathrm{K}_{\kappa_s}^{(s)}\}$ of the original points, i.e. $Cl: Tr^{(s)} \to \{\mathrm{K}_1^{(s)}, \dots, \mathrm{K}_{\kappa_s}^{(s)}\}$.

Let us depict a set of clusters $\{\mathrm{K}_1^{(s)}, \dots, \mathrm{K}_{\kappa_s}^{(s)}\}$, $s = 1, \dots, S$, together with the coastline. Clusters of ranks one and two are depicted as markers in the form of circles. Clusters of rank three and higher are depicted using bordering polygons. Other points on the land surface are depicted as markers in the form of dots. Figure 14 shows an example of a map on which all possible clusters are concentrated. It was assumed that $f_r = 0.0135$, $N = 10^3$, $S = 24$. The cluster of lines in Figure 14 clearly demonstrates the presence of large-rank clusters.

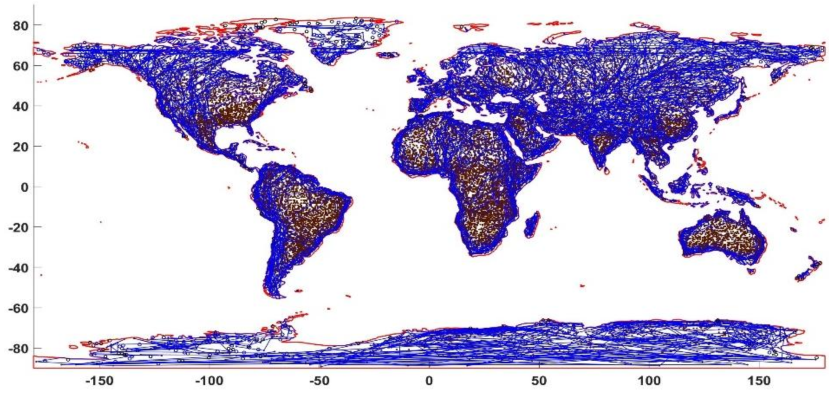


Figure 14: Cluster map, in the form of a set of bordering polygons

For a more detailed cutting of the land area, it is necessary to make a number of assumptions and calculations. In section 3, as part of the global procedure for marking the surface of the latitude – longitude globe, a 360×720 grid was selected. It turned out that the number of grid nodes, N_{Land}, located on land, was $N_{Land} = 81'192$. Let $z_l = (\varphi_l, \lambda_l)$ be the position of the l-th point on the land surface, $l = 1, \dots, N_{Land}$. For each cluster $\mathrm{K}_\kappa^{(s)}$, whose rank is three or more, we construct a bordering polygon $B[\mathrm{K}_\kappa^{(s)}]$, $\kappa = 1, \dots, \kappa_s$, $s = 1, \dots, S$. Due to the choice of a small value of the parameter f_r, clusters of rank 1 and 2 will be ignored, since their contribution is small. A visually noticeable number of one- and two-rank clusters are concentrated in the polar regions.

For each land point z_l, $l = 1, \dots, N_{Land}$, we calculate the weight w_l, which consists of the number of occurrences of the l-th point in all clusters in the form of polygons of all statistical experiments on estimating transport cost matrices with subsequent use of cluster analysis, i.e.

$$w_l = \sum_{s=1}^{S} \sum_{\kappa=1}^{\kappa_s} (z_l \in B[\mathrm{K}_\kappa^{(s)}]),\ l = 1, \dots, N_{Land}, \tag{19}$$

where parentheses $(z_l \in B[\mathrm{K}_\kappa^{(s)}])$ return one, if the inclusion of $z_l \in B[\mathrm{K}_\kappa^{(s)}]$ is true, and zero if the opposite is true, i.e. $z_l \notin B[\mathrm{K}_\kappa^{(s)}]$.

The set of weights $W = \{w_1, \dots, w_{N_{Land}}\}$ will make it possible to separate regions that, on the one hand, are close to each other in terms of their individual traffic cost profiles, on the other hand, the conservative component in the form of membership memory will be taken into account in the set of statistical tests S. In formula (19), the contribution of the l-th point to each of the clusters is the same and equal to one.

Figure 15 shows two identical maps of the set of weights W in planar and spatial formats. When calculating the weight according to formula (19), it was assumed that $f_r = 0.0135$, $N = 10^3$, $S = 24$, while ignoring the contribution of single- and double-rank clusters. It is visually visible that the maps in Figure 14 and in Figure15 are close. Clusters of similar traffic cost profiles in Figure 15b are clearly expressed in the form of corresponding "towers" with flat tops.

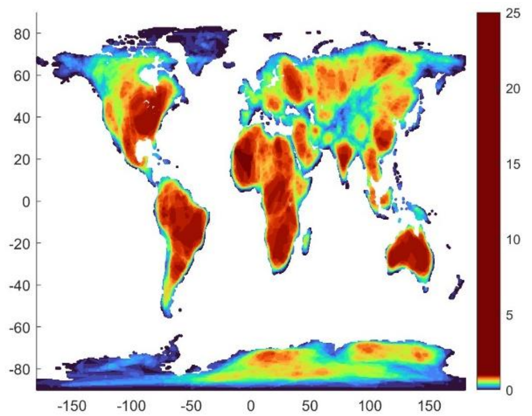


Figure 15a: Map of a set of weights in a flat format

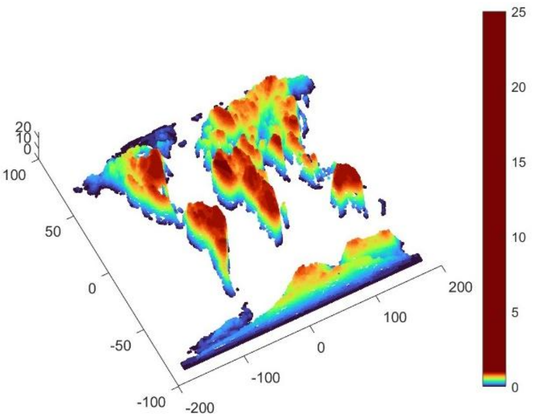


Figure 15b: Map of the set of weights in spatial format

Let's divide the set of weights (19) by the maximum value, i.e. we introduce the normalization $W \to W/\max_{1\le l\le N_{Land}} w_l$. After normalization, the weight $W=\{w_l\}$, $l=1,\dots,N_{Land}$ will vary in the range $[0,1]$ and will not directly depend on the number of statistical experiments S.

The ratio of certain territories to low-ranking and high-ranking ones is characterized by a weight normalized for a segment of $[0,1]$. When $w_l \sim 0$, it is assumed that the l-th territory acts as a low-ranking territory. With $w_l \sim 1$, it is assumed that the l-th territory is included in a high-ranking cluster. Within the framework of these remarks, it is possible to determine the force, F, as a weighted average sum of the squares of deviations from the weighted average, i.e., the weighted average variation of the weights, namely,

$$F=\frac{\sum_{l=1}^{N_{Land}}\cos\varphi_l\cdot(w_l-\overline{w})^2}{\sum_{l=1}^{N_{Land}}\cos\varphi_l},\ \overline{w}=\frac{\sum_{l=1}^{N_{Land}}\cos\varphi_l\cdot w_l}{\sum_{l=1}^{N_{Land}}\cos\varphi_l}, \tag{20}$$

where $\overline{w}$ is the weighted average of the set of weights $W=\{w_l\}$, $l=1,\dots,N_{Land}$. The presence of the cosine of latitude in (20) takes into account the sphericity of the Earth's surface.

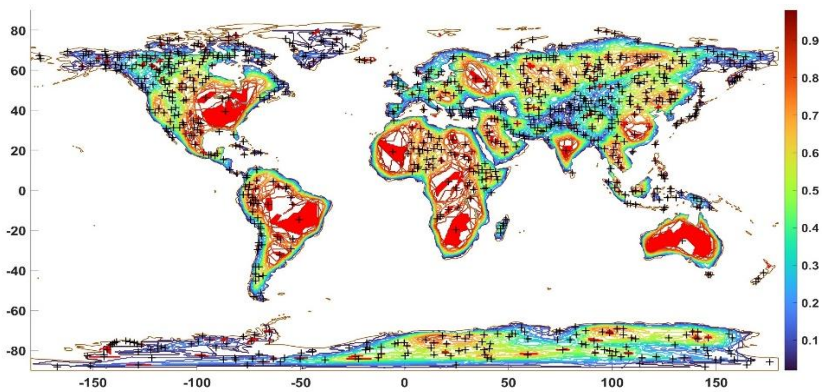


Figure 16: Level lines, local maxima (red dots) and centers of accumulation of local maxima (black pluses) of the set of weights

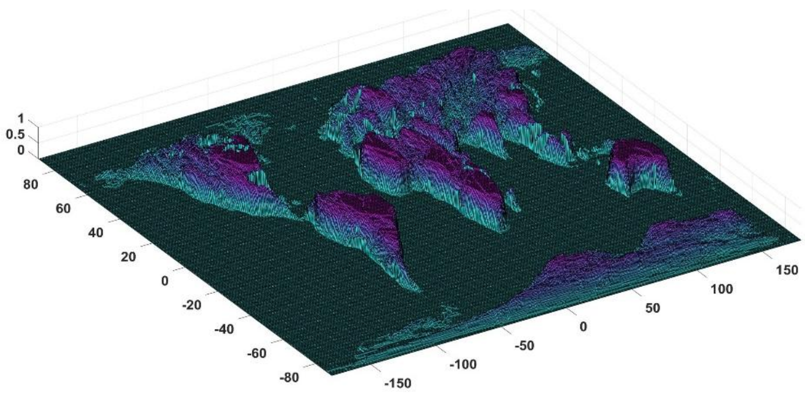


Figure 17: Three-dimensional image of the normalized weight function

In the future, keeping in mind the clustering procedure called k-means, the value (20) could be called a global single-cluster force estimate. It measures the absence or presence of consolidation of atomic clusters into large-scale associations due to the proximity and range of individual traffic cost profiles. The force F was calculated for the entire globe, and it turned out to be equal to $F \cong 7063.55$.

Figure 16 shows the lines of the normalized weight level. Markers in the form of red dots mark the local weight maxima. Clusters of red dots characterize the geometry of the profiles in the form of a plateau. In Figure 16, 5881 local maxima are marked with red dots, and $N_{lm} = 891$ black markers in the form of "pluses" indicate the centers of clusters of local maxima.

The presence of the geometry of the normalized weight function in the form of plateau profiles is clearly visible on the map in Figure 17, similar to the map in Figure 16, but executed in three-dimensional format.

Let's take the k-means method as a basis. We assume that the number of clusters is equal to the number of cluster centers of local maxima $N_{lm} = 891$. As the distance between a pair of points, consider the distance on the Earth's sphere in the form of an angle θ, the formula for calculating which is given in (14). Let's find the smallest distance from the current point to the points of the selected set of centers $N_{lm} = 891$ local maxima. The found minimum distance allows you to assign the current point to the corresponding point of the center of the local maxima.

Figure 18 shows two maps in which all N_{Land} land points are classified according to the specified procedure. In Figure 18a, and individual polygons of uniformity in the number $N_{lm} = 891$ are filled in with random colors. In Figure 18b, individual polygons of uniformity are shaded with a special palette color, taking into account the weighted average value of the weight within each of the polygons.

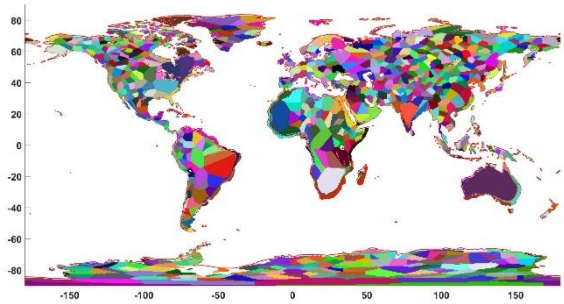


Figure 18a: Polygons of uniformity colored with random colors

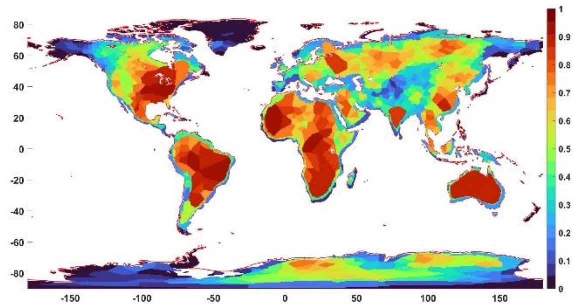


Figure 18b: The polygons of uniformity are colored with a color from the palette, taking into account the weight value

Roughly speaking, in Figure 18b, the red-brown color indicates regions that prefer a high-ranking affiliation, and the blue-purple color indicates a low-ranking affiliation. It should also be noted that in many places the coastline adjacent to high-ranking regions is colored blue, which corresponds to belonging to low-ranking clusters. This contrast makes a noticeable contribution to the force parameter.

Let's calculate the total quadratic deviation of the points from the cluster centers, f. Let's assume that as a result of applying the above procedure there are $\mathrm{K}_1, \mathrm{K}_2, \dots, \mathrm{K}_k$

clusters, in each of which there are $n_1, n_2, \ldots, n_k$ points, then the total quadratic deviation of the points from the cluster centers, or the force in our terminology, can be calculated using the formula:

$$f = \sum_{i=1}^{k} \frac{\sum_{j=1}^{n_i} \cos\varphi_j \cdot (w_{i,j} - \bar{w}_i)^2}{\sum_{j=1}^{n_i} \cos\varphi_j}, \ \bar{w}_i = \frac{\sum_{j=1}^{n_i} \cos\varphi_j \cdot w_{i,j}}{\sum_{j=1}^{n_j} \cos\varphi_j}, \tag{21}$$

where $w_{i,j}$ is the weight of the j-th point of the i-th cluster, $j = 1, \ldots, n_i$, $i = 1, \ldots, k$, $\bar{w}_i$ is the weighted average center of the i-th cluster. In the calculations, the results of which are shown in Figure 18, it was assumed that $k = N_{lm} = 981$, while it turned out that $f = 511.36$. The latter value is 13.81 times less than the single-cluster estimate of $F = 7063.55$. Thus, the global force parameter can be significantly reduced, taking into account the local characteristics of traffic costs.

As a result, it can be stated that the force parameter, as a measure of the confrontation between low-ranking and high-ranking clusters, is an objective characteristic. The force is built into the infrastructure of traffic costs, it is generated by the entire geophysical complex: the climate in terms of temperature and precipitation, terrain and traffic.

Let's build a map of the level line $w = 0.5$, which separates the low-ranking ($w < 0.5$) and large-scale clusters ($w > 0.5$). Of particular interest are large-scale clusters, which are few and can be enumerated. In Figure 19a, all polygons with a weight w exceeding 0.5 are highlighted. In Figure 19b, 19 of 152 polygons are selected that contain points with a weight $w > 0.5$. The polygons are ranked in descending order of area and renumbered. The selected 19 polygons account for 99.23% of the total area, while they are spatially separated from each other.

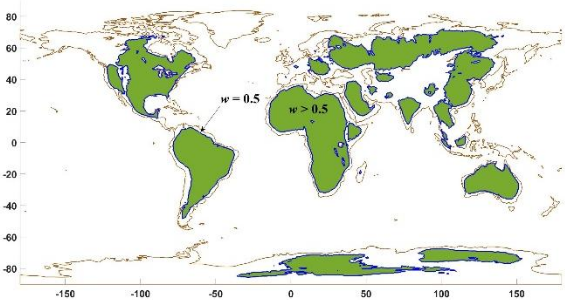


Figure 19a: Map of the level line $w = 0.5$

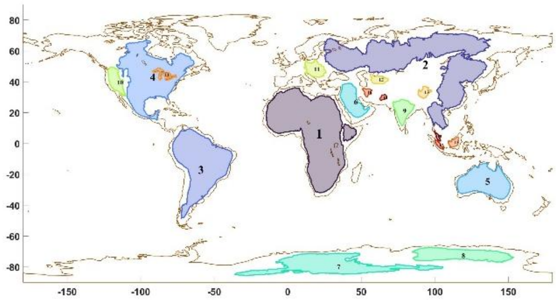


Figure 19b: Large-scale polygons ranked in descending order of area

In Figure 19b, all 19 regions in descending order of area are as follows: 1) Africa, 2) Russia + China + Southeast Asia, 3) South America, 4) USA + Canada + Mexico, 5) Australia, 6) Arabian Peninsula, 7-8) Antarctica, 9) Hindustan, 10) California + Oregon, 11) Eastern Europe, 12) Central Asia, 13) Tibet (China), 14) Great Lakes, 15) Kalimantan, 16) Iran, 17) Sumatra, 18) Afghanistan + Pakistan, 19) Malacca (Malaysia). All other regions that are not highlighted in Figure 19 belong to low-rank clusters.

Thus, we can speak of a multitude of low-ranking clusters, most of which are positioned "maritimely", and a small number of high-ranking clusters that are positioned "continentally". There is a tension between them, measured by force. Maritime and continental positioning are terms characteristic of geopolitics.

IMPERIAL TOWERS AND THE MAGNITUDES OF THE FORCE OF STATES

Let's align the level lines of the normalized weight function W in Figure 16 with the state borders of countries on the current political map of the world [12]. Figure 20 shows the desired map, where level lines and corresponding designations highlight the "shelves" of vertices of the normalized weight of land points. From now on, we will call these shelves "imperial towers" or simply "towers". We introduce the following designations for the towers: **T1**, **T2.1**, ..., **T17**; they are colored yellow by the level lines of the weight function. We will combine some towers into groups, taking into account the current political map of the world. In this case, we will obtain the following identification of all towers.

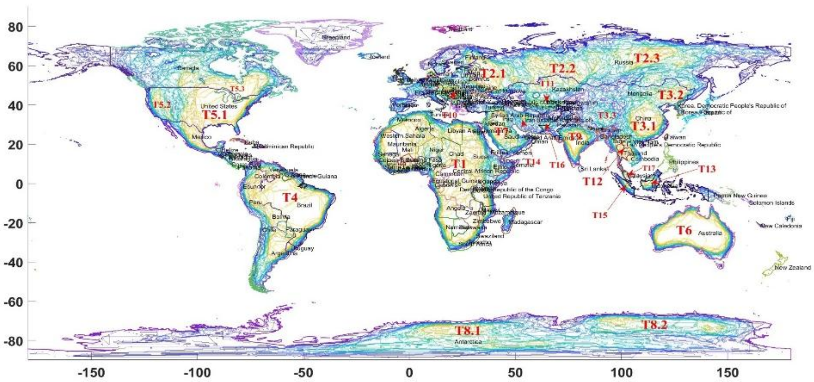


Figure 20: Positioning of imperial towers on the political map of the world

- **T1** — Africa. One of the largest towers by area.
- **T2.1**, **T2.2**, **T2.3** — Russian Federation. Three towers are located within the Russian Federation. They are partially located within the Central Federal District, the Ural Federal District (with an extension into Kazakhstan), and the northern part of the Far Eastern Federal District.
- **T3.1**, **T3.2**, **T3.3** — People's Republic of China. There are three towers located within the PRC. The first is located near the provinces of Hubei, Hunan, and Jiangxi; the second is located near the provinces of Jilin and the Inner Mongolia Autonomous Region; and the third is in Tibet.
- **T4** — South America. One big tower.
- **T5.1**, **T5.2**, **T5.3** — North America. Three towers were found in North America. The first is located primarily in the United States, with some in Mexico and Canada. The second tower is located on the California-Oregon border. The third tower is located on the Great Lakes.
- **T6** — Australia. One big tower.
- **T7** — Arabian Peninsula. One big tower.

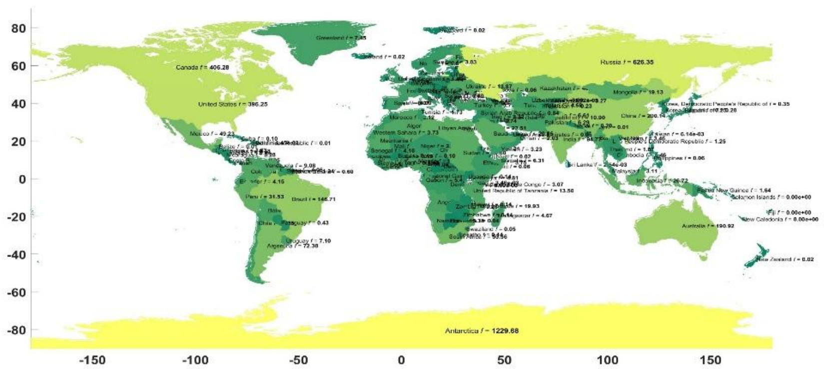
- **T8.1**, **T8.2** — Antarctica. Two towers are visible.
- **T9** — Hindustan. One big tower.
- **T10** — Eastern Europe. The tower is located in Austria and Hungary.
- **T11** — Central Asia. A tower in Uzbekistan and Tajikistan.
- **T12** — Southeast Asia. The tower is located in Thailand, Cambodia, and Laos.
- **T13** — Kalimantan.
- **T14** — Iran.
- **T15** — Sumatra.
- **T16** — One tower on the border of Afghanistan and Pakistan.
- **T17** — Malacca.

Taking into account the above list and Figure 20, 24 imperial towers are clearly visible. These towers contrast with all the other low-ranking clusters, which are depicted by the blue level lines. Note that approximately half of the towers are politically inactive or only slightly active; their current contribution to global politics is not commensurate with their size and potential influence.

Let the boundaries of K states be defined. Let's calculate the force parameter, f_k, $k = 1, \dots, K$ for each of the states in the current policy. Let's find those points, $z_{k,l} = (\varphi_{k,l}, \lambda_{k,l})$, $l = 1, \dots, l_k$, which belong to the k-th state, determine their weight, $w_{k.l}$, $l = 1, \dots, l_k$, $k = 1, \dots, K$. The points $z_{k,l}$, $l = 1, \dots, l_k$, $k = 1, \dots, K$ were selected from the entire set of land points, z_n, $n = 1, \dots, N_{Land}$ with a resolution of half a degree. Following (21), we write down the formula of the weighted average variation for calculating the magnitude of the strength of each of the states:

$$f_k = \frac{\sum_{l=1}^{l_k} \cos\varphi_{k,l} \cdot (w_{l,k} - \bar{w}_k)^2}{\sum_{l=1}^{l_k} \cos\varphi_{k.l}}, \quad \bar{w}_k = \frac{\sum_{l=1}^{l_k} \cos\varphi_{k,l} \cdot w_{k,l}}{\sum_{l=1}^{l_k} \cos\varphi_{k,l}}, \tag{22}$$

where $\bar{w}_k$ is the weighted average of the weight of the k-th state, $k = 1, \dots, K$.



<u>Figure 21</u>**: The political map of the world with the names of states and the corresponding values of force**

Figure 21 shows a political map of the world with the names of states and the corresponding values of power. The force values of the states are calculated using the formula (22). It is curious that the dominant strength value of 1229.68 is Antarctica, which, as is known, has the status of a continent. The Russian Federation is in second place with a value of 626.35. Three imperial towers, designated above as **T2.1**, **T2.2**, and **T2.3**, contribute to this force value. Unexpectedly, Canada came in third place with a force value of 406.28. The United States, with a force value of 396.25, was in fourth place with a minimal margin from Canada. Finally, China ranked fifth with a value of 200.14.

Figure 22 shows a fragment of the histogram of calculating the force of states according to the formula (22). The ranking was carried out as the force value decreased, leaving the first 25 most significant force values.

CONCLUSION

The paper presents the results of mathematical modeling and computational experiment to determine and estimate the magnitude of the force of all states on Earth from the point of view of geopolitics. The final force estimates are ranked in descending order of force values in the sequence: Antarctica, Russia, Canada, USA, China, Australia, Brazil, India, Argentina, South Africa, etc.

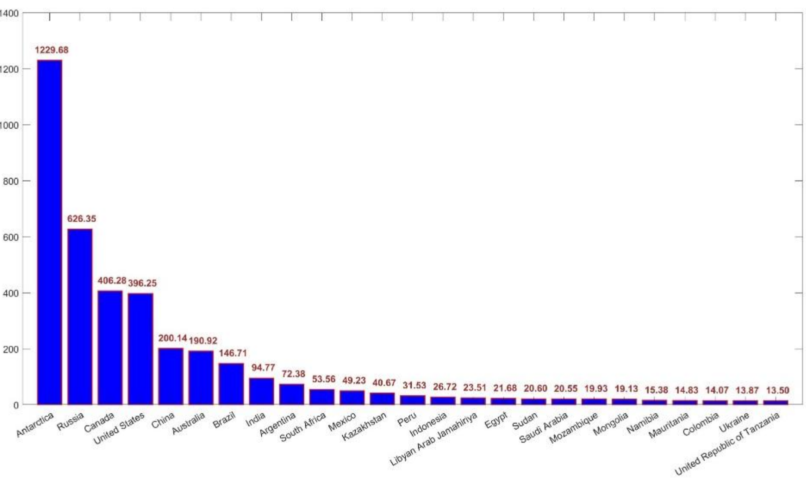


Figure 22: Histogram of the distribution of force by country (top 25)

Force acts as a measure of confrontation between groups of low-ranking and high-ranking clusters of land territories. Large- and low-ranking associations, clusters, and associations are understood from a geopolitical point of view, i.e. these are those territories where individual global traffic cost profiles are close. Traffic is understood as the exchange of a universal resource, which is interpreted in the author's previous work as the capacity of the habitat. Since traffic costs can be anything, they require non-trivial optimization. The paper develops and presents a set of algorithms for estimating traffic costs, taking into account the climate in terms of temperature and precipitation, terrain, as well as land and sea traffic. In the presented mathematical model, the initial regionalization of the Earth's surface is based on a grid with a step of half a geographical degree.

The locations of large-scale clusters on the land surface have been identified, and there are a small number of them, namely 24. They all face off against a multitude of low-ranking clusters. The global spatial structure of the mutual positioning of large- and small-rank clusters resembles a fractal structure in some places. Due to the special importance of the small number of high-ranking clusters, they received special names “imperial towers” or simply “towers”.

The African continent acts as one large tower. Three towers have been discovered on the territory of the Russian Federation (they are partially inscribed in the Central Federal District, the Ural Federal District and the North of the Far Eastern Federal District). There are three towers each in North America and China.

It should be noted that approximately half of the towers are currently either inactive or excluded from global politics. The identified towers serve as an important part of the objective context for the functioning of the human population. Towers are the geopolitical counterpart to the liberal “freedom” of low-ranking groups. As the human population develops the geopolitical infrastructure of its habitats, it must also navigate the conflict, which is described by the parameter of force, the calculation of which is presented in this paper.

REFERENCES

1. Plokhotnikov K.E. Mathematical model of geopolitics, e-book. — Society for Science and Education. Promoting Education and Research, United Kingdom, 2018. — 66p., http://scholarpublishing.org/sse/wp-content/uploads/2018/08/ABR-4819.pdf
2. Plokhotnikov K.E. Awareness of force in global politics: a mathematical model// Consciousness and physical reality. 1999. Vol. 4, No. 2. pp.27-38.
3. Plokhotnikov K.E. Mathematical modeling of global social dynamics/a course of lectures. — Moscow: Flinta, 2018. — 388p.
4. Plokhotnikov K.E. Normative Model of Global History// Global Journal of Science Frontier Research: A Physics and Space Science, 2022, Vol.22, Issue 5, Version 1.0, pp.33–59; https://globaljournals.org/GJSFR_Volume22/5-Normative-Model-of-Global.pdf (DOI: 10.17406/GJSFR) https://doi.org/10.34257/GJSFRAVOL22IS5PG33
5. http://climate.geog.udel.edu
6. http://climate.geog.udel.edu/html_pages/README.lw2.html
7. http://climate.geog.udel.edu/html_pages/README.lw.html
8. https://neo.gsfc.nasa.gov/view.php?datasetId=SEDAC_POP
9. The terrain data was taken from the Mapping Toolbox section of the MATLAB package.
10. Erlander S. The Gravity Model in Transportation Analysis: Theory and Extensions / S. Erlander, N. F. Stewart. – Utrecht: VSP, 1990. – 226 p. – (Topics in transportation). – ISBN 9067640891.
11. James G., Witten D., Hastie T., Tibshirani R. An Introduction to Statistical Learning. 2nd ed. Springer, 2021. 607 p.
12. https://lischoir.carto.com/tables/tm_world_borders_simpl_0_3/public