

**А. С. Соколов**

Гомельский государственный университет им. Ф. Скорины,  
г. Гомель, Республика Беларусь

Псковский государственный университет, г. Псков, Россия

E-mail: alsokol@tut.by

## **ФОРМАЛИЗОВАННЫЕ ПОКАЗАТЕЛИ СХОДСТВА ЭТНИЧЕСКОГО СОСТАВА ТЕРРИТОРИЙ**

*Использование формализованных показателей структуры социально-экономических систем лежит в основе создания их математических моделей и является важным инструментом их изучения как полисистемных и иерархически упорядоченных комплексов. Целью исследования является обзор количественных показателей сходства дескриптивных множеств и определение возможности их использования для оценки сходства этнической структуры различных территорий. Приведены методики расчёта сходства пар множеств — симметричных мер сходства, отличающиеся различной чувствительностью к относительному объёму пересечения множеств, и несимметричных мер сходства (мер включения множеств друг в друга), а также многоместных мер сходства  $n$  множеств. Показан способ графического отражения сходства с помощью треугольной диаграммы с осями, отражающими относительную численность населения этнических групп, а также прямоугольной диаграммы с осями, соответствующими факторам формирования этнического состава, выявленным в результате факторного анализа. Отражены результаты апробации методики расчёта сходства этнического состава территорий на примере Гродненской области Белоруссии: вычисления попарного сходства районов с использованием индексов сходства Чекановского и Ружички в 1989 и 2019 гг., показателей сходства этнического состава районов с этническим составом области в целом в 1989 и 2019 гг., показателя дисперсности этнического состава районов.*

**Ключевые слова:** дескриптивные множества, индекс сходства, мера включения, этнический состав, Гродненская область.

**Для цитирования:** Соколов А. С. Формализованные показатели сходства этнического состава территорий // Вестник Псковского государственного университета. Серия: Естественные и физико-математические науки. 2026. Т. 19. № 1. С. 36–52. DOI: <https://doi.org/10.37490/S0037480-8>

A. S. Sokolov

F. Skorina Gomel State University, Gomel, Belarus

Pskov State University, Pskov, Russia

E-mail: alsokol@tut.by

## FORMALIZED INDICATORS OF THE ETHNIC COMPOSITION SIMILARITY OF TERRITORIES

*The use of formalized indicators of structure of socio-economic systems underlies creation of their mathematical models and is an important tool for studying them as polysystemic and hierarchically ordered complexes. The aim of the research is to review quantitative indicators of similarity of descriptive sets and determine the possibility of their use to assess the similarity of the ethnic structure of different territories. The paper presents methods for calculating the similarity of pairs of sets: symmetric similarity measures, characterized by varying sensitivity to the relative volume of intersection of sets, and asymmetric similarity measures (measures of inclusion of sets within each other), as well as multiple similarity measures of n sets. A method for graphically reflecting similarity is shown using a triangular diagram with axes reflecting the relative population size of ethnic groups, as well as a rectangular diagram with axes corresponding to the factors shaping ethnic composition, identified as a result of factor analysis. The results of testing the methodology for calculating the similarity of the ethnic composition of territories are presented using the example of Grodno region of Belarus: calculations of pairwise similarity of districts using Chekanovsky and Ruzicka similarity indices in 1989 and 2019, indicators of similarity of the ethnic composition of districts with the ethnic composition of the region as a whole in 1989 and 2019, and the indicator of the ethnic composition dispersion of districts.*

**Keywords:** *descriptive sets, similarity index, inclusion measure, ethnic composition, Grodno region.*

**For citation:** Sokolov A. S. (2026), Formalized indicators of ethnic composition similarity of territories, *Vestnik Pskovskogo gosudarstvennogo universiteta. Seriya: Estestvennye i fiziko-matematicheskie nauki* [Bulletin of the Pskov State University. Series “Natural and physical and mathematical sciences”], vol. 19, no. 1, pp. 36–52. (In Russ.). DOI: <https://doi.org/10.37490/S0037480-8>

**Введение.** Математическое моделирование является одним из наиболее распространённых методов исследования в социально-экономической географии, позволяющим изучать различные аспекты структуры, связи, динамики и отношений исследуемых систем. Возможности их применения определяются широким набором инструментариев, позволяющих собирать, обрабатывать, анализировать, интерпретировать данные, выявлять закономерности и тенденции. Особенность географических исследований с привлечением математики — их логичность, многоаспектность, всесторонность охвата, системность и комплексность отображения объекта изучения [20]. В этнической географии и демографии применению математических моделей

способствуют большие объёмы числовых данных. Наряду с определением количественных мер разнообразия, параметров неоднородности территориального распределения, направления и тесноты связей, влияния отдельных факторов на функционирование изучаемых систем, группировки и т. д. важным аспектом изучения этнической структуры населения является определение количественных мер сходства различных территорий по этническому составу населения.

**Целью исследования** является определение формализованных показателей, с помощью которых возможно количественно оценить степень сходства этнической структуры различных территорий и их апобация.

**Теоретические предпосылки.** Известно, что множество возможных наблюдений у экономико-географов заведомо конечное и обычно небольшое, но о каждом из них есть много показателей, бывает, что даже больше, чем самих наблюдений. Среди важных для экономико-географов показателей особенное место занимают переменные, которые отражают разные уровни наших наблюдений [12]. Это обстоятельство позволяет представлять объекты изучения как полисистемные и иерархически упорядоченные комплексы. Методология полисистемного анализа и синтеза [19] предлагает алгоритмы создания разнообразных математических моделей объектов, основываясь на существовании множества их системных интерпретаций [20]. В основе математической модели лежат две ключевые операции: формализация — как отражение содержания модели и интерпретация — как перенос результатов моделирования на объект исследования [6]. Формализованное описание структуры исследуемых систем, в т. ч. территориальной, являясь следующим этапом их изучения после структуризации и параметризации, может включать вычисление показателей централизации и агломерации, территориальной концентрации и дифференциации, неравномерности структуры, пространственной корреляции и автокорреляции характеристик, компактности и дисперсности, контрастности, поляризации, стратификации, сопряжённости, пересекаемости и множество других [8]. Одним из таких показателей является сходство, позволяющее количественно оценить степень подобия исследуемой системы по отношению к другим системам того же или другого иерархического уровня или по отношению к самой себе в другие моменты или периоды времени. Показатели сходства были разработаны в рамках экологической науки [10] и в настоящее время очень широко применяются в геоботанике [11], флористике [3; 4], генетике [2], теории информации [1], лингвистике [18] и ряде других отраслей науки.

**Материалы и методика исследования.** С математической точки зрения, этническая структура любого региона, состоящая из  $n$  этнических групп (элементов) с определёнными количественными значениями их обилия (измеренными в одной и той же шкале, что позволяет найти сумму элементов), является дескриптивным множеством, т. е. конечным множеством, каждому элементу которого поставлено в соответствие неотрицательное число («вес») [13].

Таким образом, при изучении этнических структур можно применять методы изучения дескриптивных множеств, в частности, определять меры сходства — количественные безразмерные показатели сходства сравниваемых объектов (как в пространстве, так и во времени). К основным классам мер сходства относятся:

1. *Бинарные* (парные) меры сходства — используют для определения сходства двух множеств.

1.1. *Симметричные* меры сходства, отражающие величину сходства двух множеств и имеющие свойство симметричности  $S(A; B) = S(B; A)$ , т. е. не зависящие от порядка сравнения, дающие одинаковый результат вне зависимости от того, сравниваются ли множество  $A$  с  $B$  или  $B$  с  $A$ .

1.2. *Несимметричные* меры сходства (меры включения), отражающие степень близости одного множества по отношению к другому, результат их расчёта зависит от того, по отношению к какому множеству происходит сравнение, т. е.  $S(A; B) \neq S(B; A)$ . В частности, они могут применяться при определении отношения части-целого, т. е. степени сходства части системы и всей системы (например, сходство этнического состава административного района и области, в состав которой он входит).

2. *Многоместные* ( $n$ -арные) меры сходства — используют для определения сходства  $n$  множеств (например, степень дисперсности этнического состава административных районов области).

*Относительными* называют меры, результаты расчёта которых нормированы, то есть находятся в заданном диапазоне (чаще всего от 0 до 1).

К симметричным мерам сходства дескриптивных множеств относят [13; 15–17; 21]:

– индекс Чекановского:

$$I_{Ch} = \frac{2 \cdot \sum_{i=1}^n \min(a_i, b_i)}{\sum_{i=1}^n (a_i) + \sum_{i=1}^n (b_i)}$$

где  $n$  — количество этнических групп, одинаковое в обоих сравниваемых множествах (территориях), если этническая группа присутствует только на одной территории, то для второй территории соответствующий вес (показатель обилия) будет равным нулю;  $\min(a_i, b_i)$  — минимальное значение обилия  $i$ -ой этнической группы в двух множествах, т. е. область пересечения множеств  $A$  и  $B$ ;  $\sum(a_i)$ ,  $\sum(b_i)$  — сумма показателей обилия этнических групп соответственно первой и второй территорий.

Показателями обилия могут выступать как абсолютные (количество), так и относительные (доля от общей численности) величины. В первом случае индекс сходства будет учитывать разницу в численности различных этнических групп на двух территориях (или одной в различные моменты времени), во втором — только изменение их соотношения;

– индекс Браун–Бланке:

$$I_{BB} = \frac{\sum_{i=1}^n \min(a_i, b_i)}{\max(\sum_{i=1}^n a_i, \sum_{i=1}^n b_i)}$$

где  $\max(\sum_{i=1}^n a_i, \sum_{i=1}^n b_i)$  — максимальное значение суммы показателей суммы обилия этнических групп первой и второй территорий. Величиной, двойственной по отношению к индексу Браун–Бланке, является расстояние Юрцева [14], определяемое,

как дополнение меры Браун–Бланке до единицы:  $D = 1 - I_{BB}$ , или

$$D = \frac{\sum_{i=1}^n \max(a_i, b_i) - \sum_{i=1}^n \min(a_i, b_i)}{\sum_{i=1}^n \max(a_i, b_i)},$$

т. е. отношение объёма непересечения двух множеств к общему объёму этих множеств. Использование пары этих величин позволяет корректно анализировать отношения включения-сходства;

– индекс Шимкевича–Симпсона:

$$I_{ShS} = \frac{\sum_{i=1}^n \min(a_i, b_i)}{\min(\sum_{i=1}^n a_i, \sum_{i=1}^n b_i)},$$

– индекс Ружички:

$$I_R = \frac{\sum_{i=1}^n \min(a_i, b_i)}{\sum_{i=1}^n \max(a_i, b_i)};$$

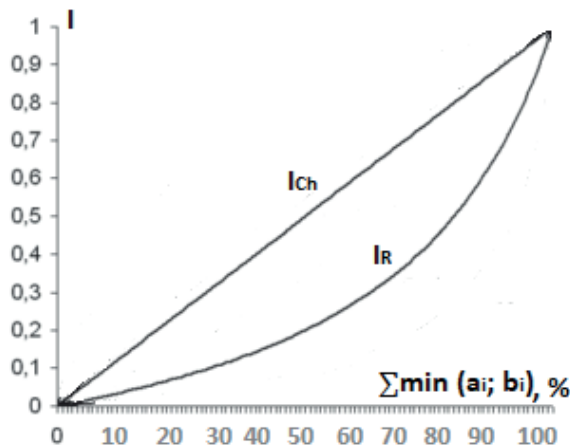
– индекс Отиаи:

$$I_O = \frac{\sum_{i=1}^n \min(a_i, b_i)}{\sqrt{\sum_{i=1}^n a_i \cdot \sum_{i=1}^n b_i}};$$

– индекс Кульчинского:

$$I_K = \frac{\sum_{i=1}^n \min(a_i, b_i)}{2} \left[ \frac{1}{\sum_{i=1}^n a_i} + \frac{1}{\sum_{i=1}^n b_i} \right].$$

Все эти индексы сходны, возрастают от нуля до единицы в зависимости от удельного веса области пересечения и, по сути, отличаются лишь разной чувствительностью к малым и большим значениям областей пересечения множеств по сравнению с объёмом этих множеств [9] (рис. 1).



**Рис. 1.** Изменения индексов Чекановского и Ружички при различных относительных значениях области пересечения множеств (составлено автором)

**Fig. 1.** Changes in the Chekanowski and Ruzicka indices for different relative values of the intersection space of sets (compiled by the author)

Если представить сравниваемые территории как точки в  $n$ -мерном пространстве (где  $n$  — количество этнических групп), координатами которых являются значения обилия по каждой этнической группе (например, доля от общего количества населения), то величину сходства можно рассчитать как эвклидово расстояние по обобщённой теореме Пифагора:

$$D_{\varepsilon} = \sqrt{\sum_{i=1}^n (a_i - b_i)^2}$$

Поскольку в данном случае верхний значений предел отсутствует, то возможно перейти к нормированному значению:

$$D'_{\varepsilon} = \frac{D_{\varepsilon}}{D_{\varepsilon}^{\max}}$$

где  $D_{\varepsilon}^{\max}$  при расчёте значений для двух множеств — максимально возможное значение  $D_{\varepsilon} = \sqrt{\sum_{i=1}^n (a_i + b_i)^2}$  (когда в двух множествах нет ни одной общей этнической группы, то есть множества  $A$  и  $B$  являются непересекающимися); при расчёте значений для  $n$  пар множеств (например сходство нескольких этнических структур с одной или сходство одной этнической структуры в разные моменты времени) в качестве  $D_{\varepsilon}^{\max}$  возможно принять максимальное значение в рассчитанной совокупности значений  $D_{\varepsilon}$ . В отличие от других величин, значение эвклидова расстояния возрастает с уменьшением сходства.

К наиболее распространённым мерам включения относятся [16]:

$$K(B; A) = \frac{\sum_{i=1}^n \min(a_i, b_i)}{\sum_{i=1}^n (a_i)}$$

и

$$K(A; B) = \frac{\sum_{i=1}^n \min(a_i, b_i)}{\sum_{i=1}^n (b_i)},$$

где  $K(B; A)$  — мера включения множества  $A$  в множество  $B$ ;  $K(A; B)$  — мера включения множества  $B$  в множество  $A$ .

Также предложены следующие меры включения [16]:

$$T(B; A) = \frac{\sum_{i=1}^n \min(a_i, b_i)}{2 \cdot \sum_{i=1}^n (a_i) - \sum_{i=1}^n \min(a_i, b_i)}$$

и

$$T(A; B) = \frac{\sum_{i=1}^n \min(a_i, b_i)}{2 \cdot \sum_{i=1}^n (b_i) - \sum_{i=1}^n \min(a_i, b_i)}.$$

Меры включения связаны с симметричными мерами сходства: вторые можно вычислить через средние значения первых. Так, индекс Чекановского представляет собой среднее гармоническое  $K(B; A)$  и  $K(A; B)$ . Поскольку эти меры включения представляют собой дроби с одинаковым числителем, а все члены этих дробей поло-

жительны, то среднее гармоническое легко можно вычислить через взятие медианты [17]:

$$\frac{\sum_{i=1}^n \min(a_i, b_i)}{\sum_{i=1}^n (a_i)} \oplus \frac{\sum_{i=1}^n \min(a_i, b_i)}{\sum_{i=1}^n (b_i)} = \frac{\sum_{i=1}^n \min(a_i, b_i) + \sum_{i=1}^n \min(a_i, b_i)}{\sum_{i=1}^n (a_i) + \sum_{i=1}^n (b_i)} = \frac{2 \cdot \sum_{i=1}^n \min(a_i, b_i)}{\sum_{i=1}^n (a_i) + \sum_{i=1}^n (b_i)},$$

где  $\oplus$  — знак взятия медианты.

Индекс Отиаи представляет собой среднее геометрическое  $K(B; A)$  и  $K(A; B)$ , индекс Кульчинского — их среднее арифметическое, индекс Браун–Бланке — максимальное значение из этих двух мер включения, индекс Шимкевича–Симпсона — минимальное значение из них (соответственно, меры Браун–Бланке и Шимкевича–Симпсона являются мерами квазисходства, т. к. не удовлетворяют всем критериям мер сходства). Индекс Ружички представляет собой среднее гармоническое  $T(B; A)$  и  $T(A; B)$  [17] и также легко вычисляется через взятие медианты.

Наиболее распространённая мера сходства  $n$  дескриптивных множеств предложена Л. Кохом в 1957 г. [15] и вычисляется для системы (региона), состоящей из  $n$  подсистем (подрегионов) по формуле:

$$K_T(x_1, \dots, x_n) = \frac{n(T-S)}{(n-1)T} = \frac{n\{\sum_{i=1}^n \sum_{j=1}^r \mu_{x_i}(x_j) - \sum_{j=1}^r \max[\mu_{x_1}(x_j), \dots, \mu_{x_n}(x_j)]\}}{(n-1) \sum_{i=1}^n \sum_{j=1}^r \mu_{x_i}(x_j)}$$

где  $\mu_{x_i}(x_j)$  — обилие  $j$ -ой этнической группы в  $i$ -ом подрегионе;  $n$  — количество подрегионов;  $r$  — количество этнических групп.

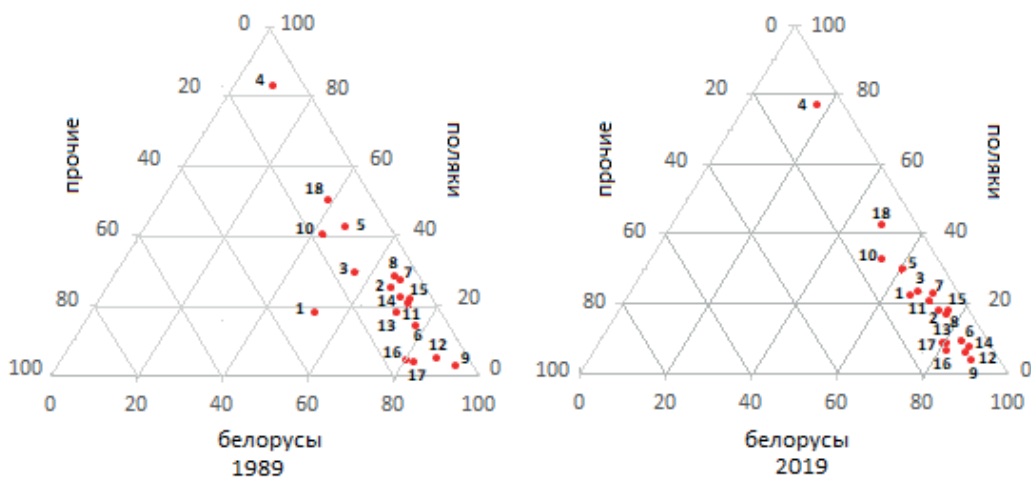
Наряду с величинами сходства можно вычислить величины различий, которые у относительных величин коэквивалентны и могут вычисляться путём вычитания соответствующих величин сходства от 1.

В некоторых случаях графически сходство этнического состава нескольких территорий можно показать с помощью треугольной диаграммы, методика создания которой в MS Excel изложена в [5]. Такая диаграмма строится по трём переменным (соответствующим трём осям), значения которых в сумме составляющим 100 %, т. е. третья переменная функционально зависит от первых двух. Такая диаграмма будет наглядной при наличии двух или трёх крупных этнических групп в составе населения территорий, существенно превышающих долю остального населения.

Апробация методов количественного определения сходства этнического состава было проведено на примере административно-территориальных единиц Гродненской области, Белоруссия (административных районов и г. Гродно). Источником данных об этническом составе стали статистические материалы<sup>1</sup>.

<sup>1</sup> Численность и основные социально-демографические характеристики населения Гродненской области по данным переписи населения 1999 года. Гродно: Грод. обл. упр. стат., 2000. 34 с.; Итоговые данные переписей населения Республики Беларусь. Нац. стат. Ком-т РБ. [Электронный ресурс]: URL: <https://census.belstat.gov.by> (дата обращения 22.01.2026).

**Результаты исследования.** Сходство этнического состава районов Гродненской области и г. Гродно в 1989 и 2019 гг. показаны с использованием треугольных диаграмм на рисунке 2. Эти диаграммы состоят из трёх осей — доля белорусов, доля поляков и доля прочих этнических групп (% от общей численности населения). Точка, обозначающая район, ставится на пересечении значений по этим осям. В 2019 г. по сравнению с 1989 г. видна тенденция «стягивания» точек к правому углу диаграммы, т. е. увеличения этнического сходства районов. Особенно заметно изменились координаты точки 1, соответствующей г. Гродно, где существенно увеличилась доля белорусов и уменьшилась доля прочих этнических групп при небольшом росте доли поляков.



**Рис. 2.** Треугольные диаграммы этнического состава административно–территориальных единиц Гродненской области в 1989 и 2019 гг., % (составлено автором)

**Fig. 2.** Triangular diagrams of the ethnic composition of administrative-territorial units of Grodno region in 1989 and 2019, % (compiled by the author)

Цифрами обозначены: 1 — г. Гродно; районы: 2 — Берестовицкий; 3 — Волковысский; 4 — Вороновский; 5 — Гродненский; 6 — Дядловский; 7 — Зельвенский; 8 — Ивьевский; 9 — Кореличский; 10 — Лидский; 11 — Мостовский; 12 — Новогрудский; 13 — Островецкий; 14 — Ошмянский; 15 — Свислочский; 16 — Слонимский; 17 — Сморгонский; 18 — Щучинский

Для каждой пары районов были рассчитаны меры сходства Чекановского и Ружички в 1989 и 2019 гг. (табл. 1, 2), по которым можно определить сходство различных районов, выделить районы и их группы с наибольшим и наименьшим сходством, динамику сходства во времени и разницу этой динамики между различными парами районов. Из таблиц видно, что этническое сходство районов увеличивается в постсоветском периоде, наибольшее увеличение сходства с 1989 по 2019 гг. зафиксированы для пары Гродненской — Сморгонский районы (на 0,34 по индексу Ружички). Эти же районы характеризуются наибольшей средней величиной увеличения сходства со всеми другими районами (на 0,15 по индексу Ружички), также высокое значение данной величины имеют Кореличский и Слонимский районы (на 0,13), а наименьшее значение — Ошмянский (на 0,04) и Зельвенский (на 0,05) районы.

Таблица 1

Индексы сходства Чекановского (над диагональю) и Ружички (под диагональю) между административно-территориальными единицами Гродненской области, 1989

Table 1

Chekanovsky (above diagonal) and Ruzicka (below diagonal) similarity indices between administrative-territorial units of Grodno region, 1989

|    | 1*   | 2    | 3    | 4    | 5    | 6    | 7    | 8    | 9    | 10   | 11   | 12   | 13   | 14   | 15   | 16   | 17   | 18   |
|----|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| 1  |      | 0,80 | 0,87 | 0,33 | 0,77 | 0,76 | 0,77 | 0,76 | 0,61 | 0,79 | 0,79 | 0,66 | 0,78 | 0,78 | 0,78 | 0,73 | 0,71 | 0,69 |
| 2  | 0,67 |      | 0,89 | 0,40 | 0,80 | 0,89 | 0,97 | 0,95 | 0,74 | 0,76 | 0,97 | 0,79 | 0,91 | 0,93 | 0,94 | 0,79 | 0,79 | 0,73 |
| 3  | 0,76 | 0,80 |      | 0,44 | 0,87 | 0,78 | 0,88 | 0,89 | 0,63 | 0,87 | 0,86 | 0,68 | 0,80 | 0,83 | 0,83 | 0,75 | 0,73 | 0,79 |
| 4  | 0,20 | 0,25 | 0,29 |      | 0,58 | 0,29 | 0,42 | 0,43 | 0,17 | 0,55 | 0,37 | 0,20 | 0,34 | 0,36 | 0,36 | 0,19 | 0,19 | 0,65 |
| 5  | 0,62 | 0,67 | 0,77 | 0,40 |      | 0,69 | 0,80 | 0,80 | 0,54 | 0,94 | 0,77 | 0,60 | 0,72 | 0,74 | 0,75 | 0,62 | 0,61 | 0,92 |
| 6  | 0,61 | 0,80 | 0,64 | 0,17 | 0,53 |      | 0,87 | 0,84 | 0,85 | 0,65 | 0,92 | 0,90 | 0,92 | 0,93 | 0,93 | 0,90 | 0,90 | 0,62 |
| 7  | 0,63 | 0,94 | 0,79 | 0,26 | 0,66 | 0,77 |      | 0,97 | 0,75 | 0,75 | 0,95 | 0,78 | 0,91 | 0,94 | 0,95 | 0,78 | 0,77 | 0,72 |
| 8  | 0,62 | 0,91 | 0,80 | 0,27 | 0,67 | 0,73 | 0,94 |      | 0,73 | 0,76 | 0,92 | 0,76 | 0,88 | 0,91 | 0,92 | 0,75 | 0,74 | 0,72 |
| 9  | 0,44 | 0,59 | 0,46 | 0,09 | 0,37 | 0,74 | 0,60 | 0,57 |      | 0,50 | 0,77 | 0,94 | 0,79 | 0,80 | 0,80 | 0,88 | 0,90 | 0,47 |
| 10 | 0,65 | 0,62 | 0,77 | 0,38 | 0,88 | 0,48 | 0,60 | 0,61 | 0,33 |      | 0,73 | 0,55 | 0,67 | 0,70 | 0,71 | 0,62 | 0,60 | 0,90 |
| 11 | 0,66 | 0,93 | 0,75 | 0,23 | 0,62 | 0,85 | 0,91 | 0,86 | 0,63 | 0,57 |      | 0,82 | 0,94 | 0,97 | 0,98 | 0,82 | 0,82 | 0,69 |
| 12 | 0,50 | 0,65 | 0,52 | 0,11 | 0,42 | 0,82 | 0,64 | 0,61 | 0,89 | 0,38 | 0,70 |      | 0,83 | 0,84 | 0,83 | 0,93 | 0,94 | 0,52 |
| 13 | 0,64 | 0,84 | 0,67 | 0,21 | 0,56 | 0,86 | 0,83 | 0,78 | 0,65 | 0,51 | 0,89 | 0,71 |      | 0,95 | 0,95 | 0,83 | 0,82 | 0,64 |
| 14 | 0,64 | 0,88 | 0,70 | 0,22 | 0,59 | 0,87 | 0,88 | 0,84 | 0,67 | 0,54 | 0,94 | 0,72 | 0,91 |      | 0,99 | 0,83 | 0,83 | 0,66 |
| 15 | 0,63 | 0,89 | 0,71 | 0,22 | 0,60 | 0,86 | 0,90 | 0,84 | 0,66 | 0,54 | 0,95 | 0,72 | 0,91 | 0,98 |      | 0,83 | 0,82 | 0,67 |
| 16 | 0,58 | 0,66 | 0,60 | 0,11 | 0,45 | 0,82 | 0,63 | 0,60 | 0,78 | 0,45 | 0,70 | 0,86 | 0,71 | 0,72 | 0,71 |      | 0,98 | 0,54 |
| 17 | 0,55 | 0,65 | 0,57 | 0,10 | 0,44 | 0,81 | 0,63 | 0,59 | 0,82 | 0,43 | 0,69 | 0,89 | 0,70 | 0,71 | 0,70 | 0,95 |      | 0,54 |
| 18 | 0,53 | 0,57 | 0,66 | 0,48 | 0,85 | 0,44 | 0,56 | 0,57 | 0,30 | 0,82 | 0,53 | 0,35 | 0,47 | 0,49 | 0,50 | 0,37 | 0,33 |      |

Примечание: \* — обозначения административно-территориальных единиц см. в рисунке 1.

Таблица 2

Индексы сходства Чекановского (над диагональю) и Ружички (под диагональю) между административно-территориальными единицами Гродненской области, 1989

Table 2

Chekanovsky (above diagonal) and Ruzicka (below diagonal) similarity indices between administrative-territorial units of Grodno region, 2019

|   | 1    | 2    | 3    | 4    | 5    | 6    | 7    | 8    | 9    | 10   | 11   | 12   | 13   | 14   | 15   | 16   | 17   | 18   |
|---|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| 1 |      | 0,95 | 0,98 | 0,43 | 0,92 | 0,81 | 0,95 | 0,87 | 0,76 | 0,87 | 0,91 | 0,78 | 0,83 | 0,79 | 0,90 | 0,83 | 0,84 | 0,80 |
| 2 | 0,91 |      | 0,96 | 0,42 | 0,89 | 0,86 | 0,98 | 0,92 | 0,81 | 0,84 | 0,96 | 0,83 | 0,87 | 0,84 | 0,95 | 0,86 | 0,88 | 0,78 |
| 3 | 0,96 | 0,92 |      | 0,44 | 0,93 | 0,82 | 0,96 | 0,88 | 0,76 | 0,88 | 0,92 | 0,79 | 0,83 | 0,79 | 0,90 | 0,83 | 0,85 | 0,81 |
| 4 | 0,28 | 0,26 | 0,28 |      | 0,51 | 0,30 | 0,43 | 0,37 | 0,25 | 0,55 | 0,39 | 0,27 | 0,31 | 0,29 | 0,39 | 0,27 | 0,30 | 0,63 |
| 5 | 0,85 | 0,80 | 0,86 | 0,35 |      | 0,75 | 0,89 | 0,81 | 0,69 | 0,95 | 0,85 | 0,72 | 0,77 | 0,73 | 0,84 | 0,76 | 0,78 | 0,88 |
| 6 | 0,68 | 0,76 | 0,69 | 0,18 | 0,75 |      | 0,86 | 0,91 | 0,94 | 0,70 | 0,90 | 0,97 | 0,95 | 0,97 | 0,91 | 0,95 | 0,96 | 0,64 |
| 7 | 0,91 | 0,97 | 0,93 | 0,28 | 0,89 | 0,75 |      | 0,92 | 0,80 | 0,85 | 0,96 | 0,83 | 0,86 | 0,83 | 0,94 | 0,85 | 0,87 | 0,78 |
| 8 | 0,77 | 0,85 | 0,78 | 0,23 | 0,81 | 0,83 | 0,85 |      | 0,86 | 0,76 | 0,96 | 0,88 | 0,90 | 0,89 | 0,97 | 0,89 | 0,91 | 0,70 |
| 9 | 0,61 | 0,67 | 0,62 | 0,14 | 0,69 | 0,89 | 0,67 | 0,75 |      | 0,64 | 0,84 | 0,97 | 0,91 | 0,97 | 0,86 | 0,93 | 0,91 | 0,58 |

|    |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
|----|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| 10 | 0,78 | 0,73 | 0,79 | 0,38 | 0,95 | 0,54 | 0,74 | 0,62 | 0,47 |      | 0,80 | 0,67 | 0,72 | 0,68 | 0,79 | 0,71 | 0,73 | 0,92 |
| 11 | 0,84 | 0,92 | 0,85 | 0,24 | 0,85 | 0,82 | 0,91 | 0,92 | 0,73 | 0,67 |      | 0,87 | 0,90 | 0,87 | 0,98 | 0,89 | 0,91 | 0,74 |
| 12 | 0,64 | 0,71 | 0,65 | 0,15 | 0,72 | 0,93 | 0,71 | 0,78 | 0,94 | 0,51 | 0,87 |      | 0,93 | 0,98 | 0,88 | 0,95 | 0,94 | 0,61 |
| 13 | 0,71 | 0,78 | 0,71 | 0,18 | 0,77 | 0,91 | 0,75 | 0,81 | 0,83 | 0,56 | 0,90 | 0,88 |      | 0,94 | 0,90 | 0,97 | 0,98 | 0,65 |
| 14 | 0,65 | 0,72 | 0,66 | 0,17 | 0,73 | 0,95 | 0,71 | 0,81 | 0,93 | 0,51 | 0,87 | 0,96 | 0,88 |      | 0,89 | 0,94 | 0,94 | 0,61 |
| 15 | 0,81 | 0,90 | 0,82 | 0,24 | 0,84 | 0,84 | 0,89 | 0,95 | 0,75 | 0,65 | 0,98 | 0,79 | 0,83 | 0,89 |      | 0,89 | 0,91 | 0,72 |
| 16 | 0,70 | 0,76 | 0,71 | 0,16 | 0,76 | 0,91 | 0,73 | 0,79 | 0,86 | 0,55 | 0,89 | 0,91 | 0,94 | 0,94 | 0,81 |      | 0,98 | 0,64 |
| 17 | 0,73 | 0,79 | 0,73 | 0,17 | 0,78 | 0,92 | 0,77 | 0,83 | 0,84 | 0,58 | 0,91 | 0,88 | 0,97 | 0,94 | 0,84 | 0,95 |      | 0,66 |
| 18 | 0,67 | 0,63 | 0,68 | 0,46 | 0,88 | 0,47 | 0,64 | 0,54 | 0,41 | 0,85 | 0,74 | 0,44 | 0,49 | 0,61 | 0,57 | 0,47 | 0,50 |      |

Такие же меры сходства были определены и для областей Белоруссии в 2019 г. (табл. 3). Наименьшее сходство имеют Гродненская и Гомельская области, а также Гродненская и Могилёвская области, наибольшее — г. Минск и Брестская область, г. Минск и Гомельская область, Минская и Могилёвская области.

Таблица 3

Индексы сходства Чекановского (над диагональю) и Ружички (под диагональю) между административно-территориальными единицами Белоруссии, 2019 г.

Table 3

Chekanovsky (above diagonal) and Ruzicka (below diagonal) similarity indices between administrative-territorial units of Belarus, 1989

| Области     | Брестская | Витебская | Гомельская | Гродненская | г.Минск | Минская | Могилёвская |
|-------------|-----------|-----------|------------|-------------|---------|---------|-------------|
| Брестская   |           | 0,94      | 0,98       | 0,79        | 0,99    | 0,98    | 0,97        |
| Витебская   | 0,89      |           | 0,95       | 0,79        | 0,95    | 0,93    | 0,93        |
| Гомельская  | 0,97      | 0,90      |            | 0,78        | 0,99    | 0,98    | 0,98        |
| Гродненская | 0,65      | 0,65      | 0,64       |             | 0,79    | 0,78    | 0,78        |
| г. Минск    | 0,98      | 0,90      | 0,98       | 0,65        |         | 0,98    | 0,97        |
| Минская     | 0,95      | 0,87      | 0,95       | 0,65        | 0,96    |         | 0,99        |
| Могилёвская | 0,94      | 0,86      | 0,95       | 0,64        | 0,95    | 0,98    |             |

Дальнейшим развитием исследования этнического сходства территорий является применение методов ординации [7] в том случае, когда возможно количественное определение значения факторов, влияющих на этнический состав населения территорий.

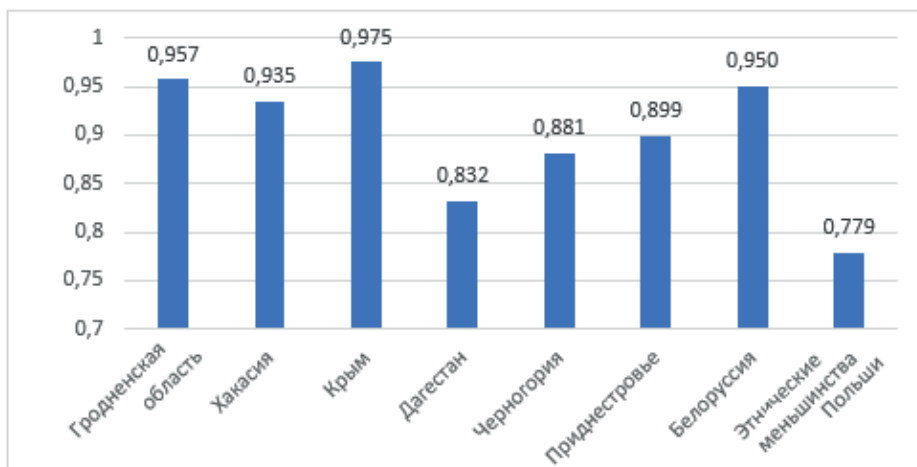
Также были рассчитаны меры включения  $K$  и  $T$  для районов Гродненской области по отношению к области в целом (рис. 3). Эти меры также в основном увеличились в различной степени за постсоветский период. Расчёт мер включения может стать основной для создания графов включения.



**Рис. 3.** Меры включения множеств  $A_i$  (районов Гродненской области и г. Гродно) в множество  $B$  (Гродненская область в целом)  $K(B;A)$  и  $T(B;A)$  (составлено автором)  
**Fig. 3.** Measures of inclusion of sets  $A_i$  (districts of Grodno region and the city of Grodno) into set  $B$  (Grodno region as a whole)  $K(B;A)$  and  $T(B;A)$  (compiled by the author)

Многоместные сходства подрегионов для различных территорий были рассчитаны по формуле Коха (рис. 4). Индекс Коха обладает низкой чувствительностью к большим относительным значениям пересечения множеств, а также сильно зависит от количества подрегионов. Поэтому можно определять и другие меры сходства этнического состава подрегионов в целом, например, средние значения индексов сходства всех пар подрегионов. Для районов Гродненской области в 2019 г. среднее значения индекса Чекановского составляет 0,80, индекса Ружички — 0,72, для регионов Белоруссии соответственно 0,91 и 0,85. Таким же образом можно рассчитать и средние

меры включения для всех пар «подрегион — регион». В случае, когда территория обладает высокой степенью моноэтничности, например, Польша, где в большинстве воеводств доля поляков превышает 99 %, а в остальных составляет не менее 95 %, можно рассчитывать меру сходства регионов не для всего населения, а например, только для этнических меньшинств (именно это значение показано для Польши на рис. 4) или для каких-либо других категорий населения.



**Рис. 4.** Мера сходства административно-территориальных единиц по формуле Коха (составлено автором)

**Fig. 4.** Measure of similarity of administrative-territorial units according to Koch's formula (compiled by the author)

В случае, когда общая сумма весов элементов множеств равна 100 %, значения индексов Барун-Бланке, Шимкевича-Симпсона, Отиаи и Кульчинского будут равны значению индекса Чекановского. Рассмотрим случай, когда она не будет равна 100 %. Возьмём в качестве весов не долю этнической группы от общего количества населения, а долю этнической группы, указавшей русский язык в качестве родного, умноженную на долю данной этнической группы в общей численности населения. Сумма весов элементов множества в этом случае будет равняться доле общего количества населения, указавшего русский язык в качестве родного (таким образом, сравнивается доля населения этнических групп с русским языком в качестве родного с учётом доли этих этнических групп в общей численности населения). На рисунке 5 показаны результаты расчёта различных индексов сходства для пяти пар районов, различающихся по степени сходства. Для сравнения динамики индексов в зависимости от величины удельного веса области пересечения множеств по сравнению с объёмом двух множеств по данным рисунка 5 был построен график (рис. 6), показывающий, что различие индексов Чекановского, Отиаи и Кульчинского более выражена при низкой степени сходства, индекс Ружички имеет наименьшие значения с сокращением разницы по мере увеличения сходства, а индексы Браун-Бланке и Шимкевича-Симпсона, как индексы квазисходства, могут иметь сложную динамику, резко отличаясь от тренда в случае полного включения одного множества в другое.

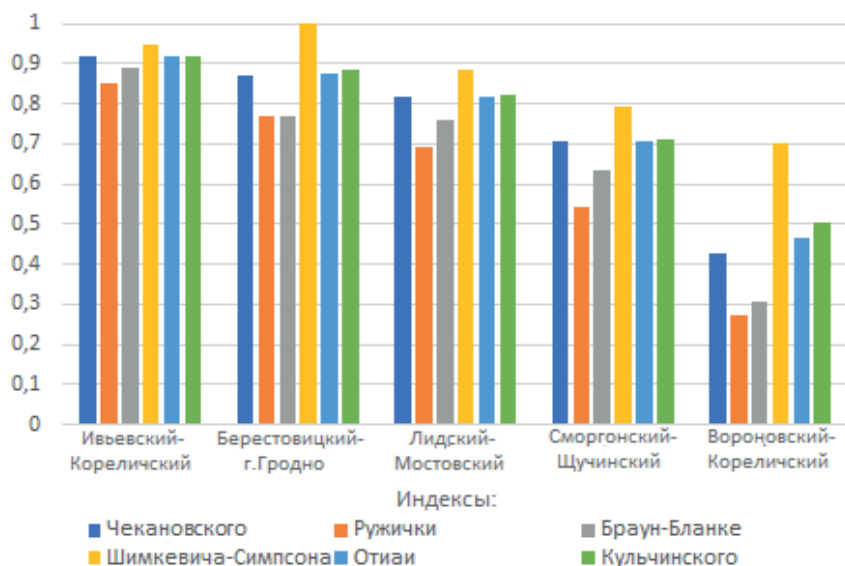


Рис. 5. Индексы сходства для пар районов Гродненской области (составлено автором)  
 Fig. 5. Similarity indices for pairs of districts of the Grodno region (compiled by the author)

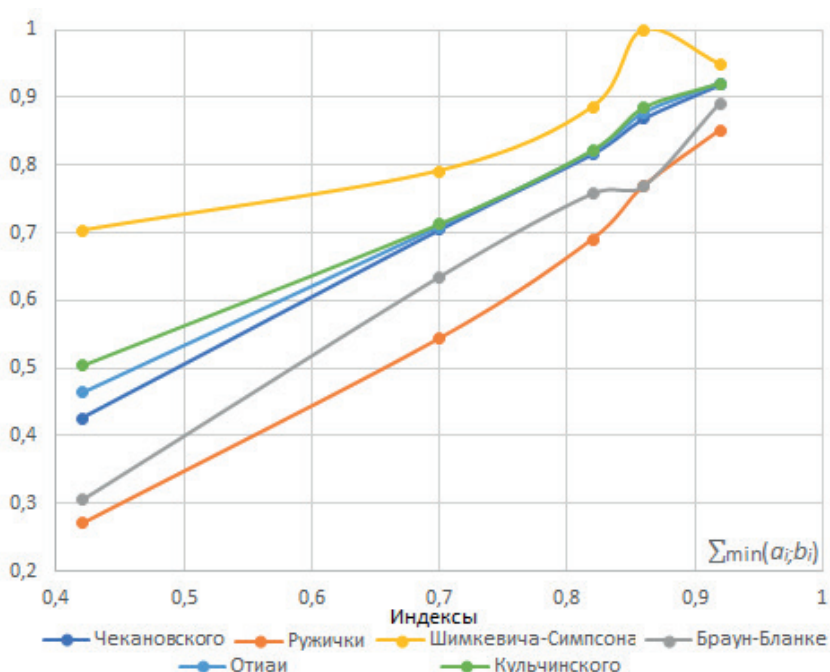
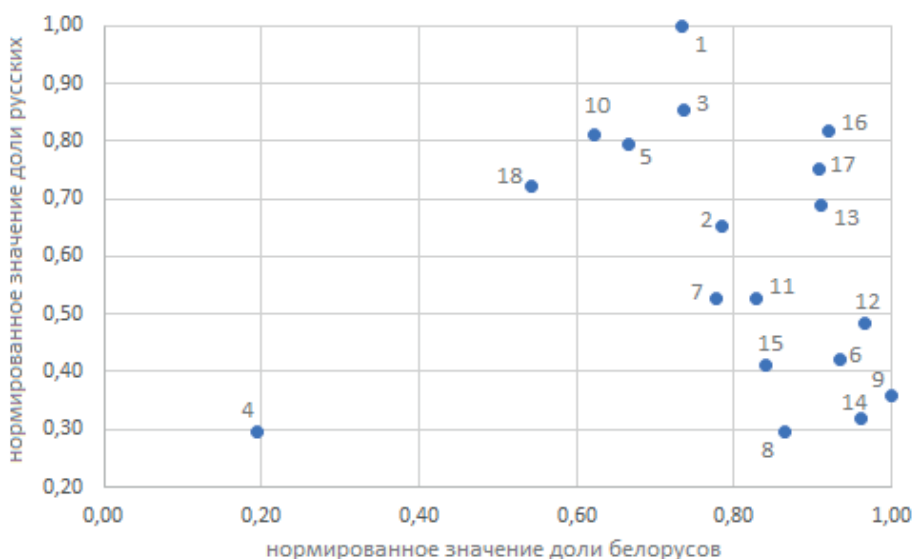


Рис. 6. Динамика индексов сходства в зависимости от относительного объёма пересечения множеств (составлено автором)  
 Fig. 6. Dynamics of similarity indices depending on the relative volume of intersection of sets (compiled by the author)

Факторный анализ показал, что на этническую структуру населения оказывают влияние два основных фактора — первый влияет на удельный вес поляков и белорусов, второй — русских и украинцев. В этом случае можно построить точечную диаграмму по двум осям, построенным по нормированным значениям доли, например, белорусов и русских (соответствующим интенсивности проявления двух выявленных факторов), на которой можно выделить группы районов со схожими значениями координат (рис. 7).



**Рис. 7.** Локализация административно-территориальных единиц Гродненской области в двухмерном пространстве факторов (обозначения административно-территориальных единиц см. рис. 1) (составлено автором)

**Fig. 7.** Localization of administrative-territorial units of Grodno region in a two-dimensional space of factors (for the designations of administrative-territorial units, see Fig. 1) (compiled by the author)

**Заключение.** Методы оценки сходства, разработанные в экологии, хоть и имеют ограничения, связанные с различием содержательной сущности объектов исследования, тем не менее при корректной интерпретации могут применяться в этнической географии. Использование количественных формализованных мер сходства этнического состава территорий позволяет определить этническую дистанцию между регионами или моментами времени, которая является важным компонентом этнической структуры территорий.

### Литература

1. Детков А. А., Воронина В. А., Гарифуллина Ю. В., Корепанов А. М., Вишнякова А. Ю. Сравнительный анализ метрик векторного расстояния растровых изображений // Вестник кибернетики. 2024. № 3. С. 22–30. DOI: 10.35266/1999-7604-2024-3-3.
2. Еленцева Д. С., Немцова Е. В. Молекулярно-генетический анализ сортов земляники садовой методом ISSR-PCR // Учёные записки Брянского гос. ун-та: естественные науки, 2025. № 2. С. 17–23.

3. *Ильина В. Н., Козловская О. В., Шукурова С. А., Фёдорова Ю. П.* Выявление особенностей флоры долинноводосборных геосистем при оценке их экологического состояния // *Известия ТулГУ. Естественные науки.* 2025. Вып. 1. С. 68–77. DOI: 10.24412/2071-6176-2025-1-68-77.
4. *Костина Н. В.* Применение индексов сходства и различия для районирования территорий на основе локальных флор // *Известия Самарского научного центра РАН.* 2013. Т. 15. № 3. С. 2160–2168.
5. *Крюкова И. Б.* К вопросу о построении треугольных диаграмм в программе Microsoft Excel при изучении пород-коллекторов Штокмановского месторождения // *Вести газовой науки.* 2011. № 3. С. 178–179.
6. *Лысенко А. В.* Математическое моделирование как метод исследования феномена автономизма в политической географии // *Геополитика и экогеодинамика регионов.* 2014. № 2. С. 129–135.
7. *Новаковский А. Б.* Методы ординации в современной геоботанике // *Вестник Института биологии Коми НЦ УрО РАН.* 2008. № 10. С. 2–8.
8. *Полян П. М.* Территориальные структуры — урбанизация — расселение: теоретические подходы и методы изучения. М.: Новый хронограф, 2014. 785 с.
9. *Родионов И. В., Созонтов А. Н.* О доверительном оценивании на основе количественных коэффициентов сходства // *Автоматика и телемеханика.* 2020. № 2. С. 157–172. DOI: 10.31857/S0005231020020105.
10. *Розенберг Г. С.* Поль Жаккар и сходство экологических объектов // *Самарская Лука: проблемы региональной и глобальной экологии.* 2012. Т. 21. № 1. С. 190–202.
11. *Розенберг Г. С.* По волнам моей памяти... (количественная фитоценология) // *Фиторазнообразие Восточной Европы.* 2024. Т. 18. № 3. С. 145–174. DOI: 10.24412/2072-8816-2024-18-3-145-174.
12. *Ростислав К. В., Синицын Н. А.* Новый количественный подход к экономико–географическим вопросам // *Теоретические и методические подходы в экономической и социальной географии: сб. статей.* М.: Геогр. ф-т МГУ, 2019. С. 46–60.
13. *Сёмкин Б. И.* Дескриптивные множества и их приложения // *Исследование систем.* Т. 1. Анализ сложных систем. Владивосток: ДВНЦ АН СССР, 1973. С. 83–94.
14. *Сёмкин Б. И.* Количественные показатели для оценки односторонних флористических связей, предложенных Б. А. Юрцевым // *Бот. ж.* 2007. Т. 92. № 4. С. 114–127.
15. *Сёмкин Б. И., Горшков М. В.* Об оценке сходства и различия в серии флористических и фитоценологических описаний // *Комаровские чтения.* Вып. 57. Владивосток: Дальнаука, 2010. С. 203–220.
16. *Сёмкин Б. И., Орешко А. П., Горшков М. В.* Об использовании биоинформационных технологий в сравнительной флористике. II. Меры включения дескриптивных множеств и их использование // *Бюллетень Ботанического сада–института ДВО РАН,* 2009. Вып. 4. С. 58–70.
17. *Сёмкин Б. И.* О связи между средними значениями двух мер включения и мерами сходства // *Бюллетень Ботанического сада–института ДВО РАН,* 2009. Вып. 3. С. 91–101.
18. *Синелева А. В.* Формальные методы оценки лексико-семантических процессов одноименных терминов (на материале терминов логики и философии) // *Вестник ЮУрГТПУ.* 2014. №5. С. 273–284.
19. *Черкашин А. К.* Полисистемный анализ и синтез. Приложение в географии. Новосибирск: Наука, 1997. 502 с.
20. *Черкашин А. К., Распутина Е. А.* Комплексная география как направление теоретических исследований и моделирования // *Географический вестник.* 2022. № 1 (60). С. 6–22. DOI: 10.17072/2079-7877-2022-1-6-22.
21. *Samdan A. M., Chimitov D. G.* Identification of genetic connections between the flora of the Alash plateau (the Republic of Tyva, Western Sayan) and the Altai-Sayan Mountain country // *Proceedings of the Moravia State Nature Reserve.* 2024. Vol. 35. P. 112–118. DOI: 10.24412/cl-31646-2686-7117-2024-35-112-118.

## References

1. *Detkov A. A., Voronina V. A., Garifullina Yu. V., Korepanov A. M., Vishnyakova A. Yu.* (2024), Comparative analysis of vector distance metrics for raster images, *Proceedings in Cybernetics*, no. 3, pp. 22–30. (In Russ.). DOI: 10.35266/1999-7604-2024-3-3.

2. Elentseva D. S., Nemtsova E. V. (2025), Molecular genetic analysis of garden strawberry varieties using the ISSR-PCR method, *Scientific notes of the Bryansk State University. Natural sciences*, no. 2, pp. 17–23. (In Russ.).
3. Ilyina V. N., Kozlovskaya O. V., Shukurova S. A., Fedorova Yu. P. (2025), Identification of features of the flora of valley-drainage geosystems when assessing their ecological state, *Proceedings of Tula State University. Natural sciences*, vol. 1, pp. 68–77. (In Russ.). DOI: 10.24412/2071-6176-2025-1-68-77.
4. Kostina N. V. (2013), Indexes of similarity and dissimilarity for territory zoning based on local floras, *Izvestiya of Samara Scientific Center of the Russian Academy of Sciences*, vol. 15, no.3, pp. 2160–2168. (In Russ.).
5. Kryukova I. B. (2011), On the issue of constructing triangular diagrams in Microsoft Excel when studying reservoir rocks of the Shtokman field, *Vesti Gazovoy Nauki* [Bulletin of Gas Science], no. 3, pp. 178–179. (In Russ.).
6. Lisenko A. V. (2014), Mathematical modeling as a method of studying the phenomenon of autonomism within political geography, *Geopolitics and Ecogeodynamics of regions*, no. 2, pp. 129–135. (In Russ.).
7. Novakovskiy A. B. (2008), Ordination methods in modern geobotany, *Bulletin of the Institute of Biology of Komi scientific center of the Russian Academy of Sciences*, no. 10, pp. 2–8. (In Russ.).
8. Polian P. M. (2014), *Territorial structures — urbanization — settlement: theoretical approaches and research methods*, Moscow, Novy Chronograph, 785 p. (In Russ.).
9. Rodionov I. V., Sozontov, A. N. (2020), On Confidence Estimation Based on Quantitative Similarity Coefficients, *Automation and Remote Control*, vol. 81, no. 2, pp. 320–332. DOI: 10.31857/S0005231020020105.
10. Rozenberg G. S. (2012), Paul Jacquard and the similarity of ecological objects, *Samara Luka: problems of regional and global ecology*, vol. 21, no. 1, pp. 190–202. (In Russ.).
11. Rozenberg G. S. (2024), According to the waves of my memory... (quantitative phytocoenology), *Phytodiversity of Eastern Europe*, vol. 18, no. 3, pp. 145–174. (In Russ.). DOI: 10.24412/2072- 8816-2024-18-3-145-174.
12. Rostislav K. V., Simitsyn N. A (2019), New quantitative approach to economic-geographical issues, *Theoretical and Methodological Approaches in Economic and Social Geography*, Moscow, 2019, pp. 46–60. (In Russ.).
13. Semkin B. I. (1973), Descriptive sets and their applications, *Systems Research. Vol. 1. Analysis of Complex Systems*, Vladivostok, pp. 83–94. (In Russ.).
14. Semkin B. I. (2007), Quantitative indicators for assessing one-sided floristic relationships proposed by B. A. Yurtsev, *Botanical Journal*, vol. 92, no. 4, pp. 114–127. (In Russ.).
15. Semkin B. I., Gorshkov M. V. (2010), About estimation of similarity and dissimilarity in series of floristic and phytocenotic lists, *Komarov's Lectures*, iss. 57, Vladivostok, Dalnauka, pp. 203–220. (In Russ.).
16. Semkin B. I., Oreshko A. P., Gorshkov M. V. (2009), On the use of bioinformation technologies in comparative floristic studies. II. Measures of inclusion of descriptive sets and their application, *Bulletin of the BGI FEB RAS*, vol. 4, pp. 58–70. (In Russ.).
17. Semkin B. I. (2009), On the relation between mean values of two measures of inclusion and measures of similarity, *Bulletin of the BGI FEB RAS*, vol. 3, pp. 91–101. (In Russ.).
18. Sineleva A. V. (2014), Formal evaluation methods of lexico-semantic processes same terms (on the material terms of logic and philosophy), *The Herald of South-Ural State Humanities-Pedagogical University*, no. 5, pp. 273–284. (In Russ.).
19. Cherkashin A. K. (1997), *Polysystemic Analysis and Synthesis: Applications in Geography*, Novosibirsk, Nauka, 502 p. (In Russ.).
20. Cherkashin A. K., Rasputina E. A. (2022). Complex geography as a direction of theoretical research and modeling, *Geographical Bulletin*, no. 1, pp. 6–22. DOI: 10.17072/2079-7877-2022-1-6-22. (In Russ.).
21. Samdan A. M., Chimitov D. G. (2024), Identification of genetic connections between the flora of the Alash plateau (the Republic of Tyva, Western Sayan) and the Altai-Sayan Mountain country, *Proceedings of the Mordovia State Nature Reserve*, vol. 35, pp. 112–118. DOI: 10.24412/cl-31646-2686-7117-2024-35-112-118.

*Сведения об авторе*

**Соколов Александр Сергеевич** — старший преподаватель кафедры экологии, Гомельский государственный университет им. Ф. Скорины, г. Гомель, Республика Беларусь; соискатель кафедры географии, Псковский государственный университет, г. Псков, Россия.

E-mail: alsokol@tut.by

ORCID: 0000-0003-4676-9600

*About the author*

**Aleksandr Sokolov**, Senior Lecturer of the Department of Ecology, F. Skorina, Gomel, State University; Gomel, Belarus; Applicant of the Department of Geography, Pskov State University, Pskov, Russia.

E-mail: alsokol@tut.by

ORCID: 0000-0003-4676-9600

*Поступила в редакцию 22.01.2026 г.*

*Поступила после доработки 05.03.2026 г.*

*Статья принята к публикации 23.03.2026 г.*

*Received 22.01.2026*

*Received in revised form 05.03.2026*

*Accepted 23.03.2026*