
МЕТОДИКА РЕГИОНАЛЬНЫХ ИССЛЕДОВАНИЙ

УДК 911.3.001

ПРОБЛЕМА ГЕНЕРАЛИЗАЦИИ ДАННЫХ В ИССЛЕДОВАНИЯХ ПРОСТРАНСТВЕННОЙ НЕОДНОРОДНОСТИ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ ЯВЛЕНИЙ НА РАЗНЫХ МАСШТАБНЫХ УРОВНЯХ

© 2019 г. Е. И. Шевчук^{1*}, П. Л. Кириллов^{2**}, А. Н. Петросян^{1***}

¹ *Национальный исследовательский университет
«Высшая школа экономики», Москва, Россия*

² *Московский государственный университет имени М.В. Ломоносова,
географический факультет, Москва, Россия*

* e-mail: egor.shevchuk@gmail.com

** e-mail: linard@mail.ru

*** e-mail: artur29031@mail.ru

В статье рассматриваются особенности применения полимасштабного подхода в географических исследованиях пространственной неоднородности социально-экономических явлений. В качестве примера анализируются социально-экономические показатели для разных уровней пространственной организации в России – муниципальных образований, субъектов федерации и экономических районов. Установлено, что оперирование низовыми уровнями пространственной организации, в соответствии с основами теории статистики, приводит к увеличению фиксируемой неоднородности. При этом отмечается, что сопоставление территориальных объектов на разных масштабных уровнях приводит к частичной деформации информации об уровне неоднородности. Возникновение подобных погрешностей объясняется континуальностью географического пространства и невозможностью однозначного определения истинных географических границ. В качестве инструмента оценки полимасштабной неоднородности предлагается коэффициент генерализации – сравнение статистических мер неоднородности на разных масштабных уровнях. Использование коэффициента позволяет выделять масштабные уровни с наибольшим разнообразием территориальных объектов. Кроме того, с его помощью возможна оценка вклада в показатели неоднородности действительной географической дифференциации. На примере оценки неоднородности для сетки микрорайонирования по Е.Е. Лейзеровичу показано, что истинная неоднородность может значительно отличаться от расчетной. Предполагается использование коэффициента для решения прикладных задач, например, оценки правомерности проведения типологий или прогнозирования на различных иерархических уровнях.

Ключевые слова: пространственная неоднородность, пространственный масштаб, генерализация, проблема изменяющегося масштаба, полимасштабный подход.

DOI: 10.5922/1994-5280-2019-3-1

Введение и постановка проблемы. Внедрение математико-статистических методов в практику экономико-географических исследований повышает объективность процесса научного познания и получаемых результатов. Однако особенности географического пространства не позволяют полностью переносить значения показателей при

обобщении, которое является неотъемлемой частью научной абстракции [1]. Возникающая потеря информации осложняется границами территориальных единиц, в рамках которых осуществляется сбор статистических данных [21].

При исследовании пространственной неоднородности важнейшим методом

становится рассмотрение явлений на разных масштабных уровнях. Британский географ С. Оупеншоу показал, что агрегирование показателей при переходе от одного масштаба к другому также нарушает валидность величин и не позволяет выявить истинные закономерности и связи между пространственными объектами [25]. Потеря информации о неоднородности становится существенной и, строго говоря, требует адекватной количественной оценки.

Теория полимасштабного подхода в отечественной географии в наибольшей степени раскрыта в работах А.И. Трейвиша [16], тогда как его практическое применение встречается в публикациях С.С. Артоболевского, П.Я. Бакланова и А.И. Трейвиша [17], В.Л. Бабурина [14], Т.Г. Нефедовой [12]. Тем не менее, количественная оценка перехода между масштабными уровнями исследований до сих пор отсутствует, а проблема генерализации данных при подобных исследованиях почти не раскрыта.

Для России актуальность данной темы повышается в связи с появлением в открытом доступе данных и относительно длинных временных рядов высокой степени объективности для разных иерархических уровней административно-территориального деления, что расширяет возможности полимасштабного анализа. Необходимость конкретизации полимасштабного подхода становится особенно важной для нашей страны в условиях обострения территориальных диспропорций экономического развития [7].

В этой связи необходимо рассмотреть инструменты исследований пространственной дифференциации социально-экономических явлений с учетом фактора генерализации показателей на разных масштабных уровнях. Существует возможность количественно оценить правомерность обобщения (или наоборот – деагрегирования) данных при переходе между масштабными уровнями с помощью коэффициента генерализации – отношения значений показателей пространственной неоднородности более высокого и низкого уровней. Данный коэффициент может быть использован в качестве дополнительного обоснования полимасштабного подхода и являться количественным подтверждением географической интуиции исследователей.

Результаты исследования и их обсуждение. Определение и методы оценки пространственной неоднородности.

Пространственная неоднородность является центральным понятием для географической науки. За рубежом зачастую именно она, или пространственная дифференциация, считается предметом или объектом изучения географии [23].

В отечественной социально-экономической географии наиболее полноценно понятие неоднородности раскрыто в рамках системно-структурного подхода, представленного, в частности, концепцией территориальных структур. Значительный вклад в ее формирование в отечественной социально-экономической географии внесли И.М. Маергойз [11] и его ученики П.М. Полян, А.И. Трейвиш [16], а также Л.И. Василевский [17], В.М. Гохман, А.А. Минц, В.С. Преображенский [4] и др.

Согласно методологии системных исследований, изучение системы невозможно без ее членения на объективно существующие компоненты, которые обладают между собой связями различного генезиса [20]. Другими словами, выделяются структуры – «инвариантные аспекты системы, ее отдельные, независимые срезы» [13]. Разные типы структурных отношений, т. е. собственно характер и источник связей, требуют выявления нескольких типов структуризации. П.М. Полян и А.И. Трейвиш указывают на необходимость выделения по крайней мере трех подобных типов структуризации применительно к пространству:

- по функциональным отношениям – связность и различие по задачам, выполняемым в системе;
- по атрибутивным отношениям – физическая связность в географическом пространстве и во времени;
- по иерархическим отношениям – связность внутри и между географическими масштабами исследований (полимасштабность) [13].

Каждая из подобных структур требует формализации посредством задания определенных параметров и их последующего измерения. Многомерность географического пространства обеспечивает большое разнообразие сторон, или по П.М. Поляну и Л.И. Василевскому, «аспектов параметров географических объектов», которые потен-

циально требуют формализации [2]. Среди них авторы выделяют территориальную дифференциацию – «разнообразие наполняющих территорию явлений и объектов, проявляющееся в их рельефном чередовании, смежности, сочетаемости» [2].

Каждый из структурных параметров может быть применен без исключения ко всем типам выделяемых структур. Так, иерархической дифференциацией может считаться неоднородность объектов разных масштабных уровней. Именно его и стоит назвать пространственной мозаичностью. В отличие от атрибутивной дифференциации мозаичность основана на объективно существующих иерархических связях элементов внутри системы. Данный факт отличает географический подход к понятию «мозаичность» от механистического в региональной экономике [5]. Для географии важны структурные особенности изучаемых пространственных объектов, тогда как региональная экономика, изучая структуры хозяйства, рассматривает пространство как вмещалище территориальных объектов.

Необходимо отметить, что попытка конкретизировать понятийный аппарат географии путем создания отдельных специальных терминов кажется излишней. Разумно отождествление многими авторами понятий территориальной «неоднородности» и «дифференциации». Термины «гетерогенность» и «мозаичность» также могут быть использованы наряду с более общим понятием «неоднородность» как уточнение последнего.

Особенности полимасштабного подхода в географических исследованиях.

Понятие «масштаба» в отечественной социально-экономической географии остается в значительной мере нераскрытым. В зависимости от смысла, вкладываемого в определение масштаба в контексте конкретных исследований, в западной географии чаще всего встречается выделение его четырех видов [22]:

– *географический* – масштаб действительного проявления исследуемых явлений на местности (т.е. зона их воздействия на окружающую среду);

– *операционный* – уровень, на котором изучается данное явление;

– *относительный* – минимально различимый для данных явлений уровень исследу-

ований, относительно которого происходит обобщение информации на более высоких иерархических уровнях;

– *картографический* – отношение размеров объекта на местности и карте.

Проблема описания пространственных явлений по территориальным ячейкам сводится к операции выделения операционного масштаба из географического. В таком случае масштаб представляется как уровень в иерархии пространственных объектов. Представление пространственных масштабов как иерархических структур позволяет продемонстрировать способы системного соподчинения явлений. Выделяются два противоположных приема соотнесения явлений (рис. 1):

– *Агрегирование* – сведение эффектов низовых уровней для получения релевантной характеристики явления на высшем уровне иерархии [6]. Функционально эта связь может быть представлена следующим образом:

$$y_k = \sum_{i=k}^n y_i \quad (1)$$

где y_k – значение показателя на уровне организации k , более высоком, чем уровень $k(n)$ на n иерархических уровнях; y_i – значения характеристик на низовых уровнях организации.

– *Декомпозиция* – разложение значения показателя при переходе на низовые уровни на сумму значений по дробным территориальным единицам и некоторой величины, демонстрирующей эффект от перехода на подчиненные уровни организации [6]. Подобный подход для пространственных исследований на основе системного подхода предложили использовать американские ученые У. Тоблер и Х. Моэльринг [23]. В универсальном виде эту зависимость можно представить следующим образом:

$$y_{k(n)} = \sigma + \sum_{i=k}^n \varepsilon_i \quad (2)$$

где $y_{k(n)}$ – значение показателя для территориальной ячейки уровня пространственной организации $k(n)$, низового по отношению к уровню k на n иерархических уровнях; σ – среднее значение показателя по стране; $\sum_{i=k}^n \varepsilon_i$ – сумма величин, показывающих эффект от перехода с уровня пространственной

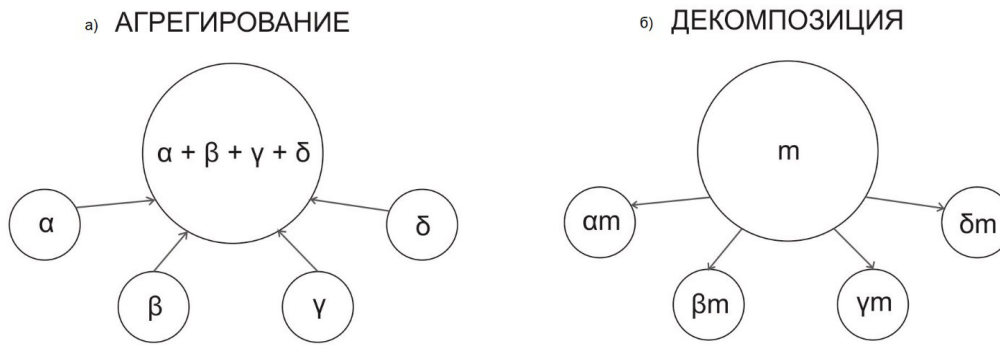


Рис. 1. Концептуальные принципы процедур межуровневого агрегирования и декомпозиции.
Источник: составлено авторами.

организации k на более низкие уровни, n – количество уровней пространственной организации в системе.

Например, для России формула (2) имела бы следующий вид:

$$y_{i(jk)} = \sigma + \varepsilon_i + \varepsilon_j + \varepsilon_k \quad (3)$$

где $y_{i(jk)}$ – значение показателя для муниципалитета k субъекта j экономического района (макрорегиона) i ; σ – среднее значение показателя по стране; ε – эффекты от перехода на более низкие уровни территориальной организации.

Движение между иерархическими уровнями в любом направлении подразумевает перенос информации о пространственных объектах. Переход с одного уровня территориальной организации на другой означает необходимость генерализации или детализации явлений¹ [3].

Таким образом, работа с пространственными данными при полимасштабном подходе требует выбора такого пространственного масштаба, для которого возможно грамотное отображение свойств частного и совокупности посредством сохранения описательных статистических характеристик при переходе между уровнями.

В контексте пространственных данных теоретические основы данной проблемы были оформлены британским ученым С. Оупеншоу, в начале 1980-х гг., который

сформулировал ее как *проблему изменяющегося масштаба* (англ. 'modifiable areal unit problem') [25]. Проще всего объяснить механизмы ее формирования на конкретном примере (рис. 2). Для каждого масштабного уровня (графики 2.а, 2.б, 2.с) характерны одинаковые значения среднего арифметического по всем объектам ($\mu = 7,5$). Однако очевидно, что движение вниз по иерархическим уровням (от графика 2.с к 2.а) приводит к увеличению пространственной дифференциации – дисперсия σ^2 увеличивается от 0 до 9,75. Таким образом, сетка агрегирования формирует различия в уровне оцениваемой пространственной дифференциации, которые могут приводить к неверным выводам о степени неоднородности территории.

Причиной подобных «игр цифр» Оупеншоу называет континуальность географического пространства и следующую из этого возможность его дробления на бесконечное множество непересекающихся территориальных единиц. Другими словами, территориальные ячейки в социально-экономических исследованиях искажают поле распределения явлений и выполняют в первую очередь функцию систематизации пространства. Таким образом, при изменении уровней исследования, т.е. масштаба, преобразуется и начертание границ, в результате чего данные таксоны могут не обладать валидностью описательных статистик для полимасштабного сравнения.

¹ Стоит отметить, что единый термин для процесса обобщения данных с точки зрения относительного пространства в географических исследованиях отсутствует: в западной географии оно соответствует понятию «scaling» [22], что можно дословно перевести как «выбор масштаба» (масштабирование); в российской географической науке ему соответствует концепция полимасштабности.

(c) $\mu - 7.5$ $\sigma^2 - 0$	7.5	7.5
	7.5	7.5

(b) $\mu - 7.5$ $\sigma^2 - 1.75$	6	7
	9	8
	6	6
	9	9

(a) $\mu - 7.5$ $\sigma^2 - 9.75$	4	8	12	2
	6	12	6	10
	2	10	8	4
	10	8	10	8

Рис. 2. Схематическое изображение проблемы изменяющегося масштаба
Источник: [20].

Количественная оценка генерализации пространственных явлений.

Проблема количественной оценки полимасштабности слабо освещена в отечественной и зарубежной научной географической литературе. Многие работы вплотную подходят или непосредственно имеют дело с различием показателей на иерархических уровнях, однако ограничиваются констатацией данного факта. Возможно, подобное пренебрежение можно объяснить кажущейся очевидностью данного вопроса² или отсутствием практического применения методики. Данный тезис может быть оспорен, если рассмотреть погрешности, возникающие при неверной обработке статистических данных или выборе мер неоднородности с недостаточной степенью валидности [15]. Наконец, как будет показано далее, измерение генерализации имеет прикладное применение.

Решение проблемы перехода между масштабами связано с определением уровня,

позволяющего показать релевантное с точки зрения математической статистики и логичное с позиций географической интуиции исследователей разнообразие географических объектов. Наиболее логичным вариантом, на наш взгляд, является анализ отклонений значений мер неоднородности описательных характеристик для разных уровней масштаба.

Таким образом, в качестве *коэффициента географической генерализации* предлагается использование меры, описывающей степень агрегирования величин между объектами одного уровня и между иерархическими уровнями

$$K_{gen} = \frac{\mu_j}{\mu_i} \quad (4)$$

где μ – значения показателя неоднородности на данном таксономическом уровне i и его более низком порядке j .

Дополнительное исследование, включая обзор прикладных экономико-географических работ, показали, что при расчете коэффициента генерализации наиболее уместно применение в качестве показателей неоднородности коэффициента Джини и энтропийного индекса Тейла. Данные показатели обладают сопоставимостью результатов вне зависимости от шкал³ и позволяют взвешивать результат на величину размера генеральной совокупности. Особое преимущество коэффициента Джини заключается в возможности графического отображения неравенства кривой Лоренца и оценке неравенства для рядов, сгруппированных в интервалах. Энтропийные меры, в свою очередь, меньше искажаются «выбросами» и имеют более строгое соответствие теоретическим принципам (аксиомам) математико-статистических мер неоднородности [8].

Для сложных систем, при сопоставлении показателей низовых уровней с наиболее высоким или отсутствии необходимости анализа факторов генерализации расчет коэффициента можно привести к следующему виду, который основан на математико-статистических методах определения вариации признака:

$$K_{gen} = \frac{I}{A_i} \sqrt{\frac{\sum_j (x_j - A_i)^2}{n}} \quad (5)$$

² Предположение по этому поводу в своей докторской диссертации высказывает А.И. Трейвиш [16]: «Работы на эту тему [масштаба] редки, хотя каждому профессионалу есть что сказать, и он делает это, но чаще всего мимоходом».

³ Для индекса Тейла – при соотношении абсолютной величины меры с количеством элементов в выборке.

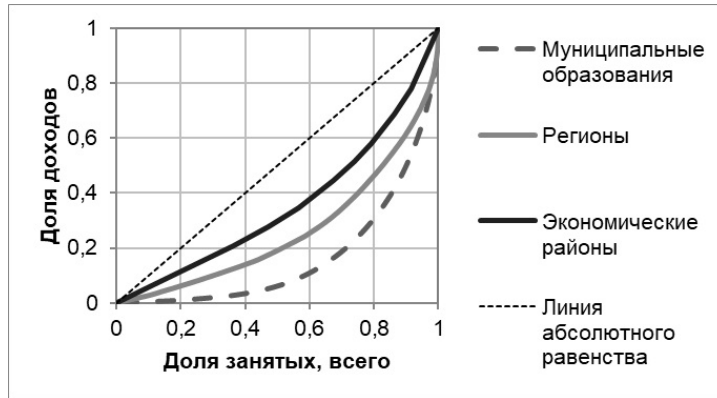


Рис. 3. Кривые Лоренца для распределения территориальных единиц разных масштабных уровней в Российской Федерации по уровню среднедушевых и суммарных заработных плат работников организаций, 2013 г. Составлено авторами по данным [26, 27].

где A_i – значение описательной характеристики на таксономическом уровне i , x_j – значение описательной характеристики субъекта таксономического уровня j , являющегося частным уровнем i , n – число объектов на таксономическом уровне j .

Для демонстрации актуальности использования коэффициента генерализации далее в данной работе будут приведены конкретные примеры его использования на практике.

Применение коэффициента генерализации в прикладных исследованиях.

1. *Определение наиболее репрезентативного иерархического уровня.* Репрезентативность метода была оценена для данных об уровне заработных плат работников организаций в 2013 г.⁴ для иерархии уровней пространственной организации в России (экономические районы, субъекты федерации, первый уровень муниципального устройства – городские округа, муниципальные районы и внутригородские территории городов федерального значения) по данным Федеральной службы государственной статистики [26, 27].

В качестве гипотезы предположим, что для России характерна высокая территориальная дифференциация по доходам населения, однако до конца неясно, на каком уровне ее отображение будет наиболее репрезента-

тивным. В решении данной задачи поможет расчет коэффициента генерализации – сравнение для разных уровней масштаба мер неоднородности (индекса Тейла и коэффициента Джини) по размеру среднемесячной заработной платы работников организаций. В качестве иллюстрации различий коэффициента Джини на разных уровнях иерархии были построены кривые Лоренца (рис. 3).

Расчеты показывают, что наибольшая гетерогенность характерна для уровня муниципальных образований (табл. 1). Это естественным образом вытекает из основ теории вероятностей и теории систем – чем дробнее устроена система, тем ниже вероятность встретить в пределах более мелкой территориальной ячейки некоторое значение показателя, т. е. возрастает энтропия, мера хаоса системы [15, 18].

Однако значения коэффициентов генерализации для трех масштабных уровней показывают, что переход с регионального на муниципальный уровень формально дает меньший вклад в меру пространственной неоднородности (1,4 для коэффициента Джини и 1,9 для индекса Тейла), чем движение от экономических районов к регионам (1,7 и 2,5 соответственно). Данный результат позволит привлечь внимание исследователей к необходимости ее более тщательного анализа. Фиксирование основной части пространствен-

⁴ Выбор связан с наибольшей репрезентативной выборкой значений показателя на уровне муниципальных образований.

Таблица 1. Результаты расчета мер неоднородности и коэффициента генерализации (на примере уровня заработных плат работников организаций для разных масштабных уровней в России, 2013 г.)

Коэффициент Джини		Коэффициенты генерализации		Индекс Тейла		Коэффициенты генерализации	
Среднемесячная заработная плата работников организаций, 2013 г.							
Экономические районы	0,289	1,657		Экономические районы	0,029	2,483	
Субъекты федерации	0,478	1,408	1,779	Субъекты федерации	0,072	1,903	3,939
Муницип. образования	0,674			Муницип. образования	0,137		
Случайное распределение по субъектам	0,269			Случайное распределение по субъектам	0,018		
Микрорайоны (Лейзерович)	0,697	1,535		Микрорайоны (Лейзерович)	0,086	1,410	
Случайное распределение микрорайонов	0,454			Случайное распределение микрорайонов	0,061		

Составлено авторами по данным: [26, 27].

ной неоднородности, т. е. преобладающих тенденций дифференциации, достигается за счет перехода на региональный уровень. Оперирование муниципальными образованиями хотя и позволяет выделить локальные особенности неравенства, может напротив даже осложнить выявление общих тенденций. Так, можно предположить, что при переходе на сетку муниципалитетов некоторые пространственно смежные территории или муниципалитеты с одинаковой численностью населения или функциями (например, административные центры субъектов федераций) имеют близкие значения показателя.

С другой стороны, расчет коэффициента сам по себе может быть источником важных выводов. На основе полученных результатов можно предположить, что меньшее увеличение дифференциации на низовом уровне связано с тем, что территории одного и того же типа (например, близкие друг другу по численности населения) в разных регионах во многом похожи по уровню жизни населения, тогда как сами регионы в целом могут отличаться в зависимости от того, насколько широко представлены данные типы территорий.

Таким образом, коэффициент генерализации может являться как количественным подтверждением правомерности использования в географических исследованиях тех или иных показателей пространственной неравномерности, так и источником нового знания – базы для дальнейших исследований.

2. *Определение степени географической дифференциации.* Возрастание значений мер неоднородности при увеличении объема выборки означает, что уровень неоднородности определяется не только географическими факторами дифференциации, но и особенностями использования математико-статистических методов. Возникает необходимость вычлнить возрастание неравномерности, создаваемое реальным увеличением географической мозаичности объектов пространства.

Проведем своеобразную проверку распределения величин на статистическую значимость, взяв для примера уровень муниципальных образований. Неравномерность распределения может быть признана статистически значимой в том случае, если

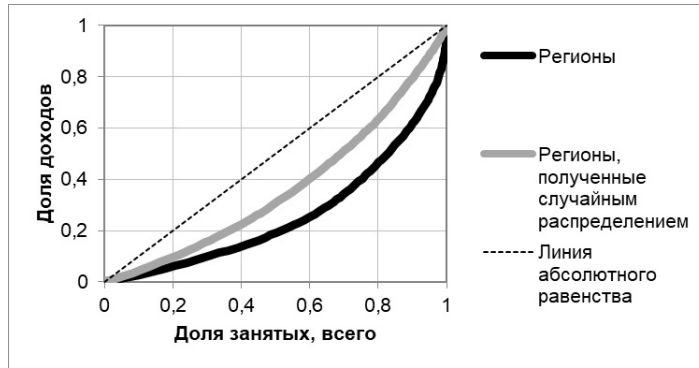


Рис. 4. Кривые Лоренца для распределения субъектов Российской Федерации и «идеальных» субъектов, полученных случайным распределением муниципалитетов, по уровню среднемесячной заработной платы работников, 2013 г. Составлено авторами по данным: [26, 27].

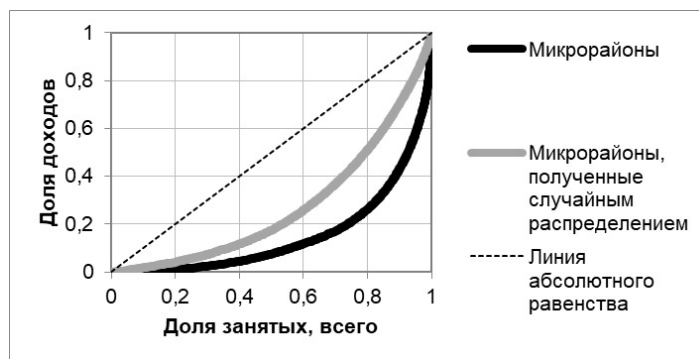


Рис. 5. Кривые Лоренца для распределения экономических микрорайонов России (по Е.Е. Лейзеровичу) и «идеальных» микрорайонов, полученных случайным распределением муниципалитетов по уровню среднемесячной заработной платы работников, 2013 г. Составлено авторами по данным: [26, 27].

статистическая нулевая гипотеза (равенство сравниваемых совокупностей) отвергается, т.е. вероятность получения подобного распределения выше, чем у случайного [15].

Сравним две системы: фактическую и условную, для которой значения описательной характеристики в каждой ячейке присваивается случайным образом из набора фактических значений. Для апробации методики используем аналогичные данные по уровню заработной плат работников организаций на уровне муниципальных образований России. Муниципалитеты случайным образом распределяются между идеальными территориальными единицами верхнего уровня. В нашем случае 51 субъект Российской Федерации имеет по 31 муниципалитету, а оставшиеся – по 30. В результате получим некоторую условно

идеальную картину распределения, на которую влияют только математические факторы (рис. 4, табл. 1 – четвертая строка).

Коэффициент генерализации в данном случае рассчитывается путем сравнения неоднородности для случайных идеальных регионов и реального распределения. Подобный показатель показывает степень действительной географической дифференциации территорий. Он позволяет скорректировать значения меры пространственной неоднородности, чтобы та иллюстрировала только действительное разнообразие. Чем выше коэффициент генерализации, тем разнороднее территориальные единицы.

Подобные рассуждения и расчеты могут все равно показаться излишне теоретизированными и не имеющими практического

применения. В опровержение, рассмотрим пример с микрорайонированием России, разработанным отечественным географом Е.Е. Лейзеровичем. Используем классическую схему районирования, предложенную ученым в 1980-е гг. [10].

Микрорайоны являются низовым уровнем, на котором проявляются устойчивые внутренние функциональные связи [9, 19]. Подобная территориальная социально-экономическая (по Лейзеровичу, «хозяйственная») система в отличие от макро- и мезорайонов отличается максимальной целостностью и однородностью.

Рассчитаем средневзвешенные заработные платы на душу населения для каждого из микрорайонов и оценим реальную степень неоднородности, объясняемую дифференциацией в специализации, наличием устойчивых территориальных связей и определенного набора факторов производства.

Степень дифференциации экономических микрорайонов по уровню доходов в сетке Е.Е. Лейзеровича оказывается крайне высокой (коэффициент Джини – 0,7), что, кажется, подтверждает качество районирования, отражающего высокую степень внутренней однородности районов при различии их между собой. Однако для того, чтобы строго утверждать это, необходимо сопоставить распределение со случайным. Условно идеальные микрорайоны были получены путем объединения муниципалитетов России на 406 групп по 6 муниципальных образований и 15 групп по 7.

Значение коэффициента Джини для случайного распределения и коэффициент генерализации (табл. 1 – последние 2 строки) демонстрируют, что истинная дифференциация не настолько сильна, как может показаться на первый взгляд. Математический фактор значительно влияет на величину меры неоднородности, которая в действительности нуждается в корректировке.

Полученные результаты не следует трактовать как оценку качества районирования Е.Е. Лейзеровича. Территориальная организация общества в постсоветский период претерпела значительные изменения. В этой связи, сетка районирования середины 1980-х гг. не может быть использована для полноценного отображения дифференциации в показателях 2013 г. Данный пример весьма наглядно демонстрирует необходимость взвешенного

подхода к анализу показателей неоднородности – все они нуждаются в проверке и грамотном использовании.

Выводы. Данная работа может считаться еще одним подтверждением ценности полимасштабного подхода в географических исследованиях. Только сопоставление характеристик объектов на разных уровнях позволяет однозначно идентифицировать как системные отношения между ними, так и их уникальность. Однако, в отличие от других исследований, мы считаем необходимым акцентировать внимание не только на результаты применения полимасштабного подхода, но и на его методологию.

Установлено, что сопоставление иерархических уровней приводит к потере значимой информации о неоднородности пространства. Анализ теоретических основ понятий «дифференциации пространства» и «географического масштаба» позволил выделить основную причину возникновения неточностей – возможность дробления пространства на бесконечное количество территориальных единиц. Идентификация подобных неточностей возможна с помощью введения количественной характеристики полимасштабных измерений. Предлагается расчет коэффициента генерализации, который представляет собой сравнение значений математико-статистических мер неоднородности на соподчиненных иерархических уровнях.

В приведенном примере генерализация данных (размер заработных плат) при переходе от уровня муниципальных образований на уровень субъектов Федерации оказывается большей, чем при переходе от субъектов Федерации к экономическим районам (коэффициент генерализации по индексу Джини 1,408 и 1,657 соответственно). Такие результаты позволяют обратить внимание, что неоднородность при переходе к более дробным территориальным ячейкам возрастает более медленным темпом и предположить, что причиной этому является наличие значительного количества муниципалитетов со схожими значениями показателя (например, пространственно смежные территории или муниципалитеты с одинаковой людностью).

С помощью расчетов на основе статистических данных для разных уровней территориальной организации России по-

казано, что увеличение дробности территориальной сетки приводит к росту значений показателей неоднородности, который не связан с реальным увеличением пространственной дифференциации. Предложена ранее не используемая в социально-экономической географии методика определения величины географической дифференциации. Коэффициент генерализации применен для сравнения фактического распределения со случайным. Таким образом, например, было установлено, что реальная неоднородность по уровню заработных плат для сетки экономических микрорайонов по Е.Е. Лейзеровичу для 1990 г. значительно меньше (в 1,5 раза), чем без учета математических особенностей расчета.

Предполагаются и другие возможности практического применения коэффициента генерализации. В нашем примере он позволяет определить, что переход от масштаба экономических районов к регионам дает больший вклад в оценки дифференциации, чем от последних к муниципалитетам (1,657 и 1,408 для коэффициента Джини соответственно). Кроме того, в прикладных исследованиях показатель может быть использован для решения следующих задач, нерассмотренных в работе:

1. Проверка значимости проведения типологий территорий – путем сравнения коэффициентов генерализации для разных

вариантов группировок и случайного распределения значений показателей по территориальным ячейкам.

2. Оценка репрезентативности административно-территориального деления – аналогичная процедура при сопоставлении разных сценариев делимитации границ территориальных единиц на одном масштабном уровне.

3. Прогнозирование социально-экономических показателей на низовых уровнях пространственной организации – определение будущих значений показателей для низовых территориальных единиц на основе сопоставления прогнозов уровня коэффициента генерализации и показателя на более высоком уровне пространственной организации.

Данные задачи, на наш взгляд, должны рассматриваться отдельно в контексте особенностей каждого прикладного исследования. Использование коэффициента генерализации при прогнозировании требует более подробного рассмотрения, так как включает в себя также и анализ временной генерализации данных.

Использование коэффициента генерализации, на наш взгляд, может способствовать расширению использования сравнительно-географического метода, и полимасштабности, в частности в географических исследованиях. Тем самым возможно получение новых, более точных выводов о специфике пространственной организации общества.

СПИСОК ЛИТЕРАТУРЫ

1. Алаев Э.Б. Социально-экономическая география: Понятийно-терминологический словарь. М.: Мысль, 1983. 350 с.
2. Василевский Л.И., Полян П.М. Системно-структурный подход и экономическая география // Системные исследования: Ежегодник 1978. М., 1978. С. 200–242.
3. Гмурман В.Е. Теория вероятностей и математическая статистика: учебное пособие для вузов. М.: Высшая школа, 2005. 479 с.
4. Гохман В.М., Минц А.А., Преображенский В.С. Системный подход в географии // Теоретическая география: Вопросы географии. Сб. 88. М.: Мысль, 1971. С. 65–75.
5. Гранберг А.Г. Основы региональной экономики: учебник для вузов. М.: Изд. дом ГУ ВШЭ, 2001. 496 с.
6. Гребнев М.И., Шульц Д.Н. Статистический метод агрегирования производственных функций // Экономика и математические методы. 2016. Т. 52, № 2. С. 112–128.
7. Зубаревич Н.В. Региональная проекция нового российского кризиса // Вопросы экономики. 2015. № 4. С. 37–52.
8. Кислицына О.А. Неравенство в распределении доходов и здоровья в современной России. М.: РИЦ ИСЭПН, 2005. 376 с.
9. Лейзерович Е.Е. Теория и практика экономического районирования. Курс лекций. М.: Изд-во Российского открытого ун-та, 1994. 72 с.
10. Лейзерович Е.Е. Экономические микрорайоны СССР: опыт типологического исследования: автореф. дисс. ... докт. геогр. наук: 11.00.02 / Ин-т географии АН СССР. М., 1990. 52 с.
11. Маергойз И.М. Территориальная структура хозяйства. Новосибирск: Наука, 1986. 302 с.
12. Нефедова Т.Г. Территориальная организация сельскохозяйственной деятельности в Европейской части современной России: дисс. ... докт. геогр. наук: 25.00.24 / Ин-т географии РАН. М., 2004. 287 с.
13. Полян П.М., Трейвиш А.И. Территориальные структуры в науке и практике. М.: Знание. 1988. 46 с.

14. Пространство циклов: Мир – Россия – регион / Под ред. В.Л. Бабурин, П.А. Чистякова. М.: Изд-во ЛКИ, 2007. 320 с.
15. Теория статистики: Учебник / Под ред. Г.Л. Громыко. М.: ИНФРА-М, 2005. 476 с.
16. *Трейвиш А.И.* Географическая полимасштабность развития России: город, район, страна и мир: дисс. ... докт. геогр. наук: 25.00.24 / Ин-т географии РАН. М., 2006. 309 с.
17. *Трейвиш А.И., Артоболевский С.С., Бакланов П.Я.* Пространство и развитие России: полимасштабный анализ // Вестник РАН. 2009. Т. 79, № 2. С. 101–112.
18. *Шеннон К.* Работы по теории информации и кибернетике. М.: Изд-во иностр. литературы, 1963. 830 с.
19. *Шувалов В.Е.* Районирование в российской социально-экономической географии: современное состояние и направления развития // Региональные исследования. 2015. № 3. С. 19–29.
20. *Щедровицкий Г.П.* Проблемы методологии системного исследования. М.: Знание, 1964. 436 с.
21. *Burt J.E., Barber G.M., Rigby D.L.* Elementary statistics for geographers. Guilford Press, 2009. 658 p.
22. *Marceau D.J.* The scale issue in the social and natural sciences // Canadian Journal of Remote Sensing. 1999. Vol. 25, № 4. P. 347–356.
23. *Mayhew S.* A dictionary of geography. Oxford University Press, USA, 2015. 554 p.
24. *Moellering H., Tobler W.* Geographical variances // Geographical Analysis. 1972. Vol. 4, № 1. P. 34–50.
25. *Openshaw S.* The modifiable areal unit problem // Concepts and techniques in modern geography. 1984. 86 p.
26. База данных показателей муниципальных образований Федеральной службы государственной статистики. URL: <http://www.gks.ru/dbscripts/munst/> (дата обращения: 20.03.2019).
27. Федеральная служба государственной статистики. Официальная статистика. Реальная среднемесячная начисленная реальная заработная плата работников по субъектам Российской Федерации за 2000–2018 гг. URL: http://www.gks.ru/wps/wcm/connect/rosstat_main/rosstat/ru/statistics/wages/ (дата обращения: 20.03.2019).

Об авторах

Шевчук Егор Игоревич – студент магистерской программы «Мировая экономика» факультета Мировой экономики и мировой политики Национального исследовательского университета «Высшая школа экономики», г. Москва.

Кириллов Павел Линардович – кандидат географических наук, ведущий научный сотрудник кафедры экономической и социальной географии России МГУ имени М.В. Ломоносова, г. Москва

Петросян Артур Нельсонович – студент магистерской программы «Демография» факультета социальных наук Национального исследовательского университета «Высшая школа экономики», г. Москва

Для цитирования:

Шевчук Е.И., Кириллов П.Л., Петросян А.Н. Проблема генерализации данных в исследованиях пространственной неоднородности социально-экономических явлений на разных масштабных уровнях // Региональные исследования. 2019. № 3. С. 4–15.

DOI: 10.5922/1994-5280-2019-3-1

Data generalization for spatial socio-economic disparities on different research scales

E. I. Shevchuk¹, P. L. Kirillov^{2}, A. N. Petrosian^{1***}**

¹ National Research University Higher School of Economics, Moscow, Russia

² Lomonosov Moscow State University, Faculty of Geography, Moscow, Russia

* e-mail: egor.shevchuk@gmail.com

** e-mail: linard@mail.ru

*** e-mail: artur29031@mail.ru

This paper reviews aspects of utilizing various research scales (polyscale method) in spatial heterogeneity studies. Analysis is based on socio-economic data for different subdivision levels in Russia – municipalities, regions and economic regions ('rayon'). Operating on lower levels of spatial organization, in accordance with fundamentals of statistics theory, is determined as the source of an increase in the ratio of observed heterogeneity. Yet, the loss of information on the rate of heterogeneity is observed when applying polyscale method. Such inaccuracies occur owing to the continuity of space and artificial nature of geographical borders. Generalization ratio which defines the difference

of heterogeneity indices on spatial scales is proposed as a measure of polyscale heterogeneity. The ratio allows to distinguish the scale levels with the highest variety of spatial units, along with the degree of actual geographical contribution to the heterogeneity indicators. Calculations example on data for microregions, as proposed by E. Leizerovich, shows considerable differences between actual and assessed heterogeneity. Application of the method in research studies is proposed, e.g. typology assessment or multi-level forecasting.

Key words: spatial heterogeneity, spatial scaling, generalization, modifiable areal unit problem, polyscale approach.