

Методика выбора рационального метода сбора и первичной обработки информации при проведении ВПН с учетом географических и социально-экономических особенностей регионов

В статье рассматриваются вопросы применения новых информационно-коммуникационных технологий при проведении Всероссийской переписи населения, связанные с внедрением мобильных устройств и сети Интернет, которые открывают новые возможности для повышения качества и своевременности сбора и первичной обработки статистической информации о населении.

В 2010 году в Российской Федерации проводилась Всероссийская перепись населения (далее – ВПН-2010), которая имела много отличий от предыдущих переписей. Были автоматизированы все технологические этапы организации ВПН-2010 – от разработки машиночитаемых бланков переписных вопросников до подготовки публикаций с итогами.

В ходе проведения пробной переписи населения 2018 года (далее – ППН-2018) Росстат впервые осуществил комбинирование применения сети Интернет и мобильных устройств для сбора сведений о населении. Так с 1 по 10 октября проводилась Интернет-перепись для всей страны, далее с 11 по 14 октября проводилась обработка сведений, собранных в ходе Интернет-переписи, затем с 15 по 31 октября 2018 года проводился опрос переписчиками в регионах ППН-2018 с применением планшетных компьютеров и машиночитаемых документов. Однако, как показал опыт проведения ППН-2018, один метод сбора сведений о населении не может быть одинаково эффективным во всех регионах России, вследствие различного отношения населения к способам сбора сведений, а также условий доступности информационно-коммуникационных технологий (ИКТ) для населения в каждом из них.

Полученный опыт свидетельствует о необходимости выбора для каждого региона метода сбора и первичной обработки статистической информации о населении, основанного на комбинировании интернет, мобильных и традиционных технологий и адаптивного к региональным условиям проведения переписи населения (удаленности и труднодоступности некоторых регионов, неготовности отдельных слоев населения к применению современных информационно-коммуникационных технологий для сбора сведений, доступности сети Интернет в различных регионах, различного отношения населения к методам проведения переписи). Решение этой проблемы является важным условием повышения полноты охвата населения, точности и достоверности результатов и снижения расходов на проведение переписи населения в каждом субъекте России.

Применение новых информационно-коммуникационных технологий, связанных с внедрением мобильных устройств и сети Интернет, открывают новые возможности для повышения качества и своевременности сбора и первичной обработки статистической информации о населении. Однако с развитием технологий проведения переписей растет количество альтернативных вариантов организации ее проведения. Все это обуславливает необходимость обоснования выбора рационального метода сбора и первичной обработки статистической

информации о населении, учитывающего региональные особенности развитости и доступности ИКТ, согласие населения к взаимодействию с органами государственной власти через Единый портал государственных услуг, доступности использования ИКТ в домашних хозяйствах. Обоснованность такого выбора связывается с использованием адекватных особенностям поставленной задачи методов и моделей принятия решений.

Цель. Разработка моделей выбора рационального метода сбора и первичной обработки информации Всероссийской переписи населения для каждого региона России с применением мобильных средств и Интернет-технологий, соответствующих имеющимся технологическим условиям, особенностям организации процессов проведения ВПН с учетом региональных условий.

Материалы и методы. В процессе решения поставленных задач использовались методы системного и статистического анализа, методы кластерного анализа, экспертных оценок и систем, теории нечетких множеств. Расчеты производились с использованием пакетов прикладных программ MS Excel, статистического пакета IBM SPSS STATISTICS.

Результаты. В научной статье разработана методика и модели обоснования рационального метода сбора и обработки первичной информации в каждом регионе России, основанная на их кластеризации по признакам, характеризующим уровень развитости и доступности ИКТ для населения в каждом регионе, его согласия к взаимодействию с органами государственной власти через Единый портал государственных услуг, доступности использования ИКТ в домашних хозяйствах и выявлении соответствия возможностей альтернативных методов сбора и первичной обработки информации о населении характеристикам готовности населения, переписного персонала и территориального органа государственной статистики к проведению переписи с применением современных ИКТ и ограничениям по затратам на проведение ВПН в каждом из них.

Заключение. В статье рассматривается методика, позволяющая обеспечить выполнение основной концепции проведения Всероссийской переписи населения – обеспечить максимальный охват населения при принятии обоснованных управленческих решений в части выбора метода сбора и первичной обработки информации ВПН в каждом регионе России, учитывающего региональные особенности развитости и доступности ИКТ, согласие населения к взаимодействию с органами государственной власти через Единый портал государственных услуг, доступности использования ИКТ в домашних хозяйствах.

Ключевые слова: методика, концепция, информационно-коммуникационные технологии, Всероссийская перепись населения, сеть Интернет, Единый портал государственных услуг, ЕПГУ, мобильные средства, кластеризация, экспертная система, нечеткие множества, модель принятия решений, качественная оценка, граф, выбор метода сбора и первичной обработки информации ВПН.

Expert System for the Selection of the Rational Method for Collecting and Processing Information During Russian Census Based on Geographical and Socio-Economic Characteristics of Regions

The article considers the issues of using new information and communication technologies (ICTs) when conducting the Russian Census related to the introduction of mobile devices and the Internet, which provide new opportunities for improving the quality and timeliness of the collecting and primary processing of statistical information about the population.

In 2010, the Russian Census was conducted in the Russian Federation, which had many differences from previous censuses. All the technological stages of Russian Census – 2010 were automated starting from the development of machine-readable forms of census questionnaires and up to the preparation of publications of its outcomes.

In the course of the trial population census held in 2018, ROSSTAT first combined the use of the Internet and mobile devices for collecting the information about the population. So, from October 1 to October 10, 2018 an online census was conducted for the whole country. Then, from October 11 to October 14, 2018 the information collected during the online census was processed, from October 15 to October 31, 2018, a survey was conducted by census takers in the regions of the trial population census – 2018 with the use of tablets and machine-readable documents. However, as the experience of the trial population census – 2018 demonstrated, one method of collecting information about the population cannot be equally efficient in all regions of Russia, due to the different attitude of the population to the methods of collecting information, as well as the conditions for the availability of information and communication technologies (ICTs) to the population in each region.

The experience gained indicates the need for each region to choose the method of collecting and primary processing of statistical information about the population for each region based on a combination of Internet, mobile and traditional technologies that is adaptive to the regional conditions of the population census (remoteness and inaccessibility of some regions, the unreadiness of individual segments of the population to use modern ICTs for collecting information, Internet accessibility in various regions, different attitude of the population to census methods). The solution of this problem is an important condition for increasing the coverage of the population, the accuracy and reliability of the results and reducing the cost of conducting a census in each subject of Russia.

The use of new information and communication technologies related to introduction of mobile devices and the Internet opens up new opportunities for improving the quality and timeliness of the collection and primary processing of statistical information about the population. However, with the development of census technologies, the number of alternative options for organizing it is growing. All this makes it necessary to justify the selection of a rational method of collecting and primary processing of statistical information about the

population based on the regional characteristics of the development and accessibility of ICT, the consent of the population to interact with public authorities through the Public Services Portal, availability of ICTs in households. The validity of this choice is associated with the use of methods and decision-making models that are adequate to the peculiarities of the task.

Purpose. Development of models for the selection of the rational method of collecting and primary processing of information of the Russian Census for each region of Russia using mobile devices and Internet technologies that correspond to the existing technological environment and the specificities of organizing Russian Census processes.

Materials and methods. In the process of performing the preset tasks, methods of system and statistical analysis, expert assessment and systems, the fuzzy-set theory and cluster analysis were used. Calculations were carried out using MS Excel application software packages and IBM SPSS STATISTICS statistical software package.

Results. The scientific article describes the developed expert system and models for substantiating a rational method of collecting and processing primary information in each Russian region based on their clustering according to the features characterizing the level of development and their accessibility to the population in each region, the population's consent to cooperate with governmental authorities through the Public Services Portal, accessibility of ICT use in households and identifying the compliance of possibilities provided by alternative methods of collecting and processing the information on the population to the characteristics of readiness of the population, census personnel and territorial state statistics agency for conducting a census using modern ICTs and the limitations on the cost of conducting Russian Census in each region.

Conclusion. The paper discusses the methodology that allows for the implementation of basic concept of the Russian Census, which is to ensure maximum coverage of the population while making informed management decisions on the choice of methods for collecting and primary processing of Russian Census information in each region of Russia, taking into account regional characteristics of the development and accessibility of ICTs, the consent of the population to interaction with public authorities through the Public Services Portal, access to ICTs in households.

Keywords: methodology, concept, information and communication technologies (ICTs), Russian Census, the Internet, Public Services Portal, mobile devices, clustering, expert system, fuzzy-set theory, decision making model, qualitative assessment, the selection of the method for collecting and primary processing information of the Russian Census.

Введение

Всероссийская перепись населения представляет собой планомерный, научно организованный и систематический сбор данных о явлениях и процессах общественной жизни

путем регистрации существенных признаков с целью получения в дальнейшем обобщающих характеристик этих явлений и процессов.

Основные научные подходы к формированию теории и организации проведения

переписей населения нашли отражение в работах зарубежных и отечественных ученых Т.А. Асмуса, Г. Беккера, С. Вобан, Э. Галлея, Л.М. Гохберга, А.А. Кауфмана, В.Г. Минашкина, В.С. Мхитаряна, В. Петти, П.В. Перегеева, А.В. Пеше-

хонова, И.П. Селивановского, А.И. Чупрова, А.А. Чупрова и др. В данных работах сформированы общие принципы организации и проведения переписей населения, заложены основы новых технологических подходов к сбору и первичной обработке статистической информации. Методологические положения по разработке методов сбора статистической информации о населении и автоматизации процессов ее обработки отражены в научных трудах таких отечественных ученых, как В.И. Борткевича, А.Г. Ковалевского, Н.С. Четверикова, М.А. Королева, В.В. Шуракова, В.П. Божко, Я.Л. Циписа и др. Опыт применения перспективных методов при проведении переписей населения в развитых странах дает возможность сделать ряд обобщающих выводов [1, 2].

Применение новых информационно-коммуникационных технологий, связанных с внедрением мобильных устройств и сети Интернет, открывают новые возможности для повышения качества и своевременности сбора и первичной обработки статистической информации о населении. Однако с развитием технологий проведения переписей растет количество альтернативных вариантов организации ее проведения. Все это обуславливает необходимость обоснования выбора рационального метода сбора и первичной обработки статистической информации о населении, учитывающего региональные особенности развитости и доступности ИКТ, согласие населения к взаимодействию с органами государственной власти через Единый портал государственных услуг, доступности использования ИКТ в домашних хозяйствах. Обоснованность такого выбора связывается с использованием адекватных особенностям поставленной задачи методов и моделей принятия решений.

Опыт последней ВПН-2010 показал необходимость модернизации подходов к проведению переписей населения, без которых уже невозможно достичь главной ее цели – получения полной и достоверной картины современного общества.

Это же показал опыт проведения ППН-2018, один метод сбора сведений о населении с применением планшетных компьютеров и сети Интернет не может быть одинаково эффективным во всех регионах России, вследствие различного отношения населения к способам сбора сведений, а также условий доступности информационно-коммуникационных технологий (ИКТ) для населения в каждом из них.

С учетом анализа международного опыта [3] и существующих практик Росстата [4] рассмотрены и усовершенствованы методы сбора и обработки информации переписи населения с применением современных ИКТ.

1. Последовательный метод.

Основан на разделении этапов самозаполнения населением электронных переписных вопросников (далее – ЭПЛ) при проведении Интернет-переписи и переписи населения с применением мобильных средств переписчиками в период обхода адресов проживания тех, кто не принял участие в Интернет-переписи (с контрольным обходом адресов, принявших участие в Интернет-переписи). Однако данный метод существенно ограничивает по времени период проведения Интернет-переписи и переписи путем обхода населения переписчиками ввиду того, что общий период проведения ВПН остается неизменным. При проведении ППН-2018 отводилось 10 дней на проведение Интернет-переписи, 4 дня на обработку итогов Интернет-переписи и 14 дней проведении «полевой» населения переписчиками.

2. Одновременный метод.

Предполагает одновременное проведение Интернет-переписи и «полевой» переписи населения переписчиками. В ходе информационно-разъяснительной компании всему населению предлагается принять участие в Интернет-переписи путем самозаполнения ЭПЛ на ЕПГУ. Одновременно проходит опрос населения переписчиками с применением мобильных устройств. Период переписи составляет месяц и контрольные мероприятия 4 дня. Однако данный метод может быть более затратным, так как необходим найм переписчиков для проведения опроса во всех адресах проживания населения. Уровень охвата населения Интернет-переписью является прогнозным.

3. Локализованный метод.

Предлагается разделение помещений (локализацию) на группы для прохождения Интернет-переписи и опроса путем обхода помещений переписчиками. Перед проведением переписи должна проводиться работа, по результатам которой выделяются наиболее перспективные, с точки зрения востребованности населением ЕПГУ и наличия сети Интернет, адреса. Населению, проживающему по данным адресам через электронную почту, зарегистрированную на Портале ЕПГУ, сразу предлагается пройти перепись путем самозаполнения ЭПЛ на ЕПГУ. Этот способ применим столичных регионах и в крупных городах, с высоким уровнем жизни и где население не охотно идет на контакт с переписчиками. Это дает возможность увеличить численность населения, самостоятельно принявшего участие в Интернет-переписи, к остальному населению идут переписчики по адресам проживания для сбора сведений с применением мобильных устройств. Период переписи составляет три недели и одна неделя – контрольные мероприятия.

4. Классический метод.

Метод сбора и обработки сведений о населении на классических машиночитаемых бумажных переписных листах (МЧД). Предлагается применять там, где нет возможности использовать современных ИКТ для сбора сведений о населении.

Однако с учетом социально-экономических особенностей ряда регионов (неразвитая ИКТ-инфраструктура для проведения Интернет-опроса, труднодоступные регионы, регионы в которых невозможно проведение ВПН в общее время для всей страны, неготовность определенной части респондентов принять участие в Интернет-переписи в единое время для всей страны или других причин) обоснован выбор в отдельном регионе России другого метода сбора и первичной обработки информации ВПН. Данное обстоятельство требует разработки математических моделей выбора методов сбора и первичной обработки информации ВПН с применением мобильных средств и Интернет-технологий с учетом многокритериальности решаемой задачи и необходимости уточнения социально-экономических характеристик регионов России.

Принимая во внимание тот факт, что каждый из методов сбора и первичной обработки информации имеет свои сильные и слабые стороны, целесообразно осуществить их выбор в определенном сочетании с позиции достижения высокого качества формирования статистической информации и с учетом выделенного бюджета для каждого региона России на основе кластеризации регионов России по характерным признакам, а также на основе анализа качественных характеристик готовности населения, переписного персонала и территориального органа государственной статистики с применением модели нечеткой оценки.

Методика выбора рационального метода сбора и обработки информации при проведении ВПН

Принимая во внимание тот факт, что каждый из методов сбора и первичной обработки информации ВПН имеет свои сильные и слабые стороны, целесообразно осуществить выбор определённого метода для каждого региона России с учетом анализа качественных факторов, учитывающих региональные особенности развитости и доступности ИКТ, согласие населения к взаимодействию с органами государственной власти через Единый портал государственных услуг,

доступности использования ИКТ в домашних хозяйствах.

Структурная блок-схема методики выбора рационального метода приведена на рис. 1.

Решение задачи [5–7] о выборе рационального метода сбора и первичной обработки информации о населении сводится к выбору одной или нескольких лучших альтернатив набора методов с помощью разработанной методики, которая позволяет выполнить кластеризацию регионов с учетом групп однородных признаков: развитости и доступности ИКТ для населения в каждом регионе, его согласия к взаимодействию с органами государственной власти через Единый

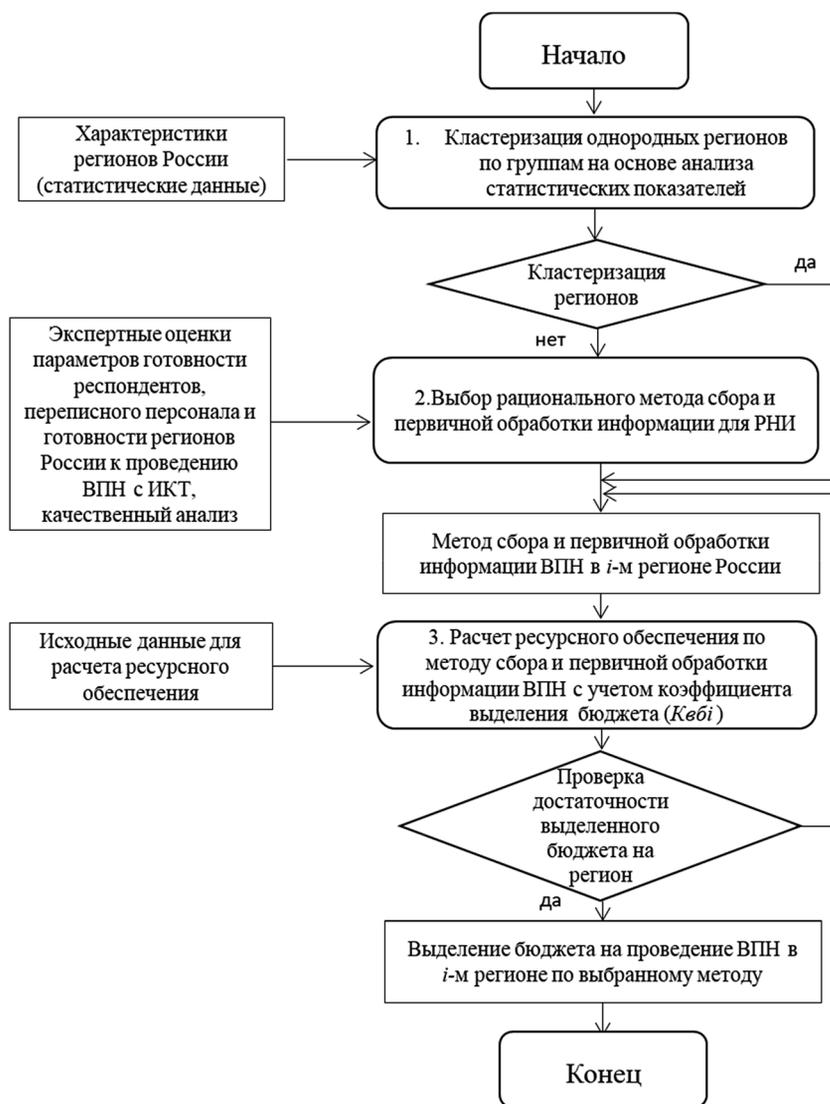


Рис. 1. Структурная блок-схема методики выбора рационального метода сбора и обработки информации при проведении ВПН в каждом регионе России

портал государственных услуг; доступности использования ИКТ в домашних хозяйствах и в дальнейшем возможности провести качественную оценку нечетких характеристик готовности населения, переписного персонала и территориального органа государственной статистики к проведению переписи населения с применением современных ИКТ; отношении населения к участию в переписи по определенному методу сбора и первичной обработки информации ВПН, а также провести количественную оценку бюджета на проведение ВПН в каждом регионе России.

На первом этапе методики выполняется кластеризация регионов России, по которым достаточно статистической информации для группировки по признакам развитости и доступности ИКТ для населения в каждом регионе, его согласия к взаимодействию с органами государственной власти через Единый портал государственных услуг и доступности использования ИКТ в домашних хозяйствах. Для каждой из данных групп регионов России определяется предпочтительный метод сбора и первичной обработки информации ВПН с применением мобильных средств и сети Интернет.

На втором этапе методики для регионов России, по которым недостаточно статистической информации для выбора предпочтительного метода по результатам кластеризации, определяется рациональный метод сбора и первичной обработки информации с применением модели принятия решений с учетом нечетких характеристик готовности населения, переписного персонала и территориального органа государственной статистики к проведению переписи населения с применением современных ИКТ, а также отношения населения к участию в переписи

по определенному методу сбора и первичной обработки информации ВПН.

На третьем этапе методики производится расчёт использования ресурсов (трудовых, временных, стоимостных) с учетом вероятности рисков отказа наемного персонала в отдельных регионах России, стоимостных затрат на развитие архитектуры информационно-вычислительной системы (ИВС), а также выполняется проверка достаточности выделенного бюджета на проведение переписи в каждом регионе России в сравнении с нормативными затратами по классическому методу проведения переписи населения. Если расчет ресурсов по выбранному методу в регионе превышает нормативные затраты по «классическому» методу проведения переписи населения, производится уточнение метода сбора и первичной обработки информации о населении.

Далее рассмотрим подробно каждый этап методики выбора рационального метода сбора и первичной обработки информации ВПН.

Кластеризации регионов России с учетом анализа однородных признаков

С учетом анализа статистических данных выборочного федерального статистического наблюдения по вопросам использования населением информационных технологий и информационно-телекоммуникационных сетей в работе обосновывается необходимость группировки ряда регионов по методам сбора и обработки первичной информации в зависимости от признаков развитости и доступности ИКТ для населения в каждом регионе, его согласия к взаимодействию с органами государственной власти через Единый портал государственных услуг и доступности использования ИКТ в домашних хозяйствах. Процесс анализа данных признаков регионов подразумевает кластеризацию регионов, где каждый кластер, который определяет применение метода сбора и обработки первичной информации, будет соответствовать набору приблизительно оди-



Рис. 2. Древоидная структура системы оценок (статистических данных) для кластеризации регионов России

наковых признаков регионов и набору близких для проживающего в них населения характеристик.

Из [5] известно, кластеризация – это группировка определённых объектов (явлений, фактов) на основе определенных данных, описывающих характерные свойства данных объектов, т.е. регионы России внутри своего кластера должны иметь общие характерные свойства, которые отличают их от регионов, которые вошли в другие кластеры. В этой связи предлагается кластеризация регионов России по двенадцати параметрам, характеризующим доступность к электронным государственным услугам, наличие доступа к сети Интернет, использование ИКТ в домашних хозяйствах, регистрацию респондентов на порталах государственных и муниципальных услуг, развитость и доступность ИКТ для населения в каждом регионе.

Исходя из этой постановки задачи, исходные данные (характерные свойства регионов России) для кластеризации целесообразно представить в виде совокупности оценок (статистических данных) по каждому из указанных факторов, показанных на рис. 2. Каждая группа статистических данных по i -му признаку состоит из группы значений по 82-регионам, которые собраны в 12-мерный вектор строку $x_i = [x_{i1}, x_{i2}, \dots, x_{in}]$, где i – признак, $i = 1, n, j = 1, n$ – регионы России.

$$X_n = \begin{pmatrix} x_{11}, x_{1j}, \dots, x_{1n} \\ x_{i1}, x_{ij}, \dots, x_{in} \\ \vdots \\ x_{N1}, x_{Nj}, \dots, x_{Nn} \end{pmatrix} \quad (1)$$

Древовидная структура системы оценок (статистических данных) для кластеризации регионов России показана на рис. 2.

Рассмотрено решение задачи распределения регионов России по кластерам с использованием статистического па-

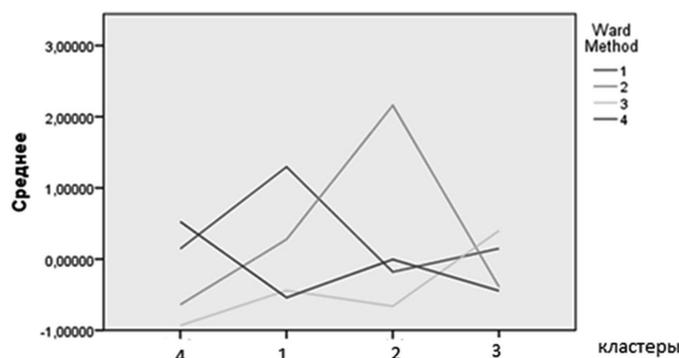


Рис. 3. Графики средних значений распределения кластеров регионов России (по вертикали – значения факторов, полученных в результате факторного анализа)

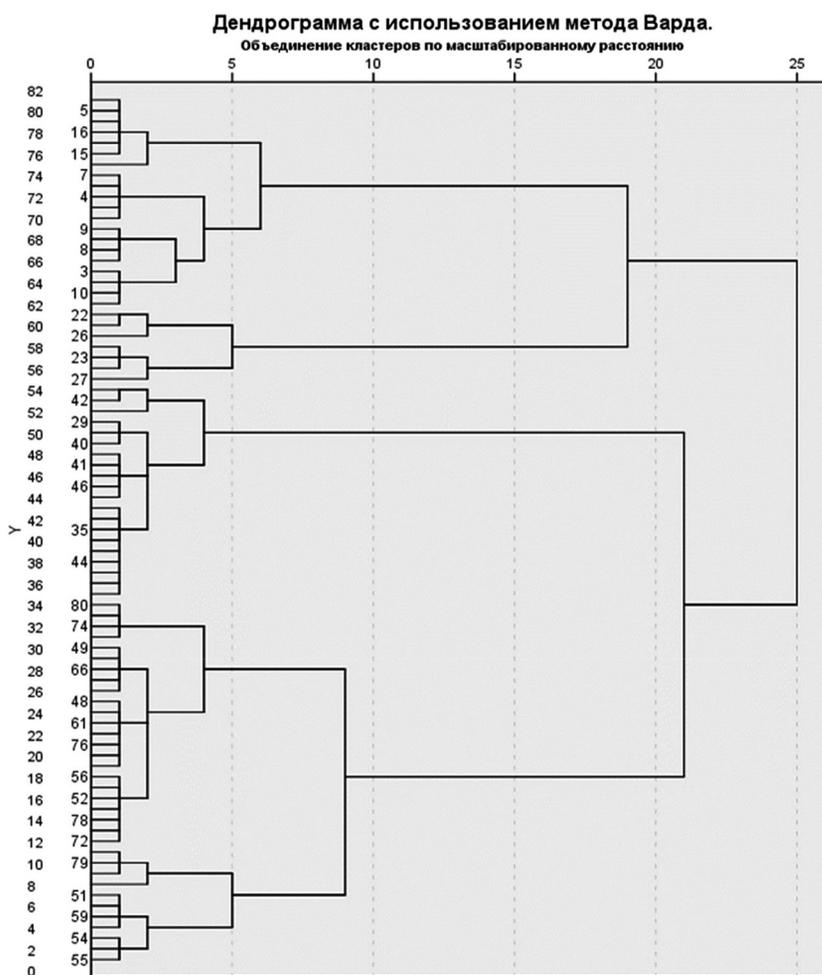


Рис. 4. Определение числа кластеров по методу Варда

кета SPSS.

Предварительно проведен факторный анализ (метод главных компонент) – по результатам которого система оценок приведена к новой системе координат (4 главных компоненты вместо исходных 12 переменных, указанных на рис. 2, которые объясняют 72%

дисперсии исходных данных).

Далее с помощью иерархической кластеризации (метод Варда) получено 4 кластера регионов России. На рис. 3 показан график средних значений распределения кластеров регионов России (для интерпретации названий кластеров).

На дендрограмме рис. 4 по-

Таблица 1

Результаты отнесения метода сбора в регионах России по выделенным признакам кластеров

Номер кластера	Характеристики регионов России	Количество регионов
1	Наиболее активно пользуются электронными Гос.услугами	20
2	Высокий уровень доступности и востребованности Интернета, низкая активность пользования электронными Гос.услугами	34
3	Наиболее высокий уровень технической/финансовой недоступности Интернета	7
4	Интернет не востребован, низкая активность пользования электронными Гос.услугами	14
5	Труднодоступные регионы/районы России.	7

Таблица 2

Результат соответствия метода сбора и первичной обработки информации каждому кластеру регионов России

Метод 1 (последовательный)	Метод 2 (параллельный)	Метод 3 (локализованный)	Метод 4 (классический)	Кластер
+				1
		+		2
	+			3
			+	4
			+	5

казано определение числа кластеров по методу Варда.

В табл. 1 приведены результаты отнесения регионов России по признакам иерархической кластеризации.

В каждом кластере включены регионы, наиболее близко расположенные в пространстве факторов, полученных в результате факторного анализа.

В табл. 2 по результатам кластеризации для каждого из регионов России представлен результат логического назначения каждому кластеру регионов предпочтительного метода сбора и первичной обработки информации ВПН.

В частности, кластеру № 1 соответствуют столичные регионы и крупные промышленные регионы России, в которых респонденты с одной стороны наиболее активно пользуются государственными услугами в электронном виде, а с другой – не желают общаться с переписчиками лично.

Кластеру 2 соответствуют развитые регионы высоким уровнем доступности и востребованности сети Интернет и низкой активностью пользова-

ния государственными услугами в электронном виде.

Кластеру 3 соответствуют регионы с наиболее высоким уровнем технической/финансовой недоступности сети Интернет.

Кластеру 4 соответствуют регионы, в которых Интернет не востребован, а также наблюдается низкая активность пользования государственными услугами.

Кластеру 5 относятся труднодоступные регионы/районы России, такие как Камчатский край, Магаданская обл., республика Якутия, Чукотский АО, Сахалинская обл., (а также в данный кластер попадут труднодоступные районы в других регионах, которые не попали ни в один из перечисленных выше кластеров и решение о применении того или иного метода сбора и первичной обработки информации при проведении ВПН принимается на основании качественной оценки параметров данных труднодоступных районов).

В случае, если регион России, по которому недостаточно статистической информации

для выбора предпочтительного метода по результатам кластеризации, применяется модель принятия решений для выбора метода сбора и первичной обработки информации ВПН на основе применения аппарата теории нечетких множеств. Данные регионы исключаются из кластеризации и называются регионы с недостаточной информацией для кластеризации (РНИ).

Модель принятия решений для выбора метода сбора и первичной обработки информации ВПН в регионе России

Задача принятия управленческих решений – одна из самых востребованных в любой прикладной области [7,9,10,19]. Ее решение сводится к набору правил для выбора одной или нескольких наиболее оптимальных альтернатив из некоторого множества данных.

В предыдущем разделе рассмотрена кластеризация регионов России по характерным свойствам на основе статистических данных. Однако не по все регионы можно однозначно отнести к той или иной группе при кластеризации в виду недостаточности исходных данных для этого. К этим регионам России (РНИ – регионы или районы с не достаточной информацией), предлагается применить модель принятия решений по выбору рационального метода сбора и первичной обработки информации ВПН на основе аппарата теории нечетких множеств, которая позволяет учесть характеристики готовности населения, переписного персонала и территориального органа государственной статистики на с учетом экспертной интерпретации данных анкетирования регионов.

Выделим основные факторы, на основании качественного анализа характеристик

которых будет построен выбор рационального метода сбора и первичной обработки информации ВПН в РНИ.

Выбор рационального метода имеет следующие ключевые факторы (подцели решения задачи):

А) – готовность населения (респондентов, зарегистрированных на Портале ЕПГУ) к участию в Интернет-переписи;

В) – готовность переписного персонала к проведению ВПН с планшетными компьютерами (ПК);

С) – готовность ТОГС (Территориального органа государственной статистики) к проведению ВПН с ПК.

Выбор метода сбора и первичной обработки информации ВПН в отдельном регионе России осуществляется на основе набора правил экспертной системы, интерпретирующей готовность региона к проведению ВПН с применением информационно-коммуникационных технологий.

Принципиальная модель принятия решений для выбора метода в виде графа «И-ИЛИ» представлена на рис. 5.

Граф «И-ИЛИ» позволяет отражать в виде структуры функциональные зависимости выбранных целевых переменных от выявленных в результате проведенного анализа факторов [10, 11, 18]. В этом графе каждая вершина-фактор, за исключением окончательных переменных, представляет собой целевую переменную, а определяющие его зависимые вершины-аргументы – связанные с ней подчиненные факторы.

Таким образом, анализ набора правил принятия решений в виде графа «И-ИЛИ» отражает в обобщенной форме процесс оценки факторов, влияющих на выбор рационального метода сбора и обработки информации при переписи населения в РНИ.

Представленной на рис. 5 модели принятия решений для выбора метода в виде графа «И-ИЛИ» соответствуют следующие подзадачи.

А) оценка готовности населения к участию в Интернет-переписи на основе следующих характеристик:

а.1 – удобство использования ЕПГУ (Единого Портала государственных услуг) для участия в Интернет-переписи населения;

а.2 – возможность прохождения Интернет-переписи на ЕПГУ;

а.3 – доверие к обработке конфиденциальных данных на ЕПГУ;

а.4 – сложность электронного вопросника переписи населения для самостоятельного заполнения респондентами на ЕПГУ;

а.5 – необходимость мотивации участия в переписи населения на ЕПГУ.

В) готовность переписного персонала к проведению ВПН с ПК на основе следующих характеристик:

б.1 – сложность электронного вопросника для проведения опроса населения на планшетном компьютере для переписного персонала;

б.2 – удобство интерфейса вопросника переписи населе-

ния на планшетном компьютере;

б.3 – опыт участия в других обследованиях Росстата на планшетных компьютерах;

б.4 – риски отказа респондентов от участия в переписи населения на планшетных компьютерах;

б.5 – риски по порче или утрате планшетных компьютеров в ходе проведения опроса.

С) готовность ТОГС к проведению ВПН с ПК на основе следующих характеристик:

с.1 – качество методического материала для обучения привлекаемого персонала для работы на планшетных компьютерах;

с.2 – сложность подбора персонала для проведения опроса населения на планшетных компьютерах;

с.3 – риски отказа переписного персонала в ходе проведения переписи населения.

с.4 – наличие труднодоступных районов, в которых применение планшетных компьютеров невозможно по объективным причинам (невозможность подзарядить ПК, транспортная недоступность, аномальные природные условия, или др. причины).

с.5 – безопасность процесса проведения переписи населения для персонала с планшетными компьютерами в регионе (отдельном районе).

Данные для качественных оценок параметров предлагается получать по результатам предварительного анкетирования:

– по параметру А) Оценка готовности респондентов к участию в Интернет-переписи – размещение анкеты на ЕПГУ за 3 месяца до проведения Интернет-переписи;

– по параметру В) Готовность переписного персонала к проведению ВПН с мобильными устройствами – проведение анкетирования штатными сотрудниками ТОГС привлекаемого персонала для проведения опросов населения



Рис. 5. Принципиальная модель набора правил принятия решений для выбора метода сбора и первичной обработки информации ВПН в РНИ

Таблица решений по оценке переменных А, В, С

Подцель	Параметры	Интервальное значение параметра	Коэффициент уверенности подцели (cf)
А. Готовность респондентов участвовать в Интернет-переписи	a1 – удобство использования ЕПГУ для участия в ИП населения	a1 = 0	A+= “готов” cf 0
		a1 > 0 и a1 < 0,6	A+= “готов” cf 10
		a1 ≥ 0,6 и a1 ≤ 1	A+= “готов” cf 20
	a2 – возможность прохождения ИП на ЕПГУ	a2 = 0	A+= “готов” cf 0
		a2 > 0 и a2 < 0,5	A+= “готов” cf 15
		a2 ≥ 0,5 и a2 ≤ 1	A+= “готов” cf 30
	a3 – доверие к обработке конфиденциальных данных на ЕПГУ	a3 = 0	A+= “готов” cf 0
		a3 > 0 и a3 < 0,6	A+= “готов” cf 10
		a3 ≥ 0,6 и a3 ≤ 1	A+= “готов” cf 20
	a4 – сложность электронного вопросника переписи населения для самостоятельного заполнения респондентами на ЕПГУ	a4 = 0	A+= “готов” cf 0
		a4 > 0 и a4 < 0,5	A+= “готов” cf 10
		a4 ≥ 0,5 и a4 ≤ 1	A+= “готов” cf 20
	a5 – необходимость мотивации участия в ИП на ЕПГУ	a5 = 0	A+= “готов” cf 30
		a5 > 0 и a5 < 0,5	A+= “готов” cf 15
		a5 ≥ 0,5 и a5 ≤ 1	A+= “готов” cf 0
В. Готовность переписного персонала к проведению ВПН с ПК	b1 – сложность электронного вопросника для проведения опроса населения на ПК для переписного персонала.	b1 = 0	B+= “готов” cf 0
		b1 > 0 и b1 < 0,6	B+= “готов” cf 10
		b1 ≥ 0,6 и b1 ≤ 1	B+= “готов” cf 20
	b2 – удобство интерфейса вопросника переписи населения на ПК	b2 = 0	B+= “готов” cf 0
		b2 > 0 и b2 < 0,5	B+= “готов” cf 15
		b2 ≥ 0 и b2 ≤ 1	B+= “готов” cf 30
	b3 – опыт участия в других обследованиях Росстата на ПК.	b3 = 0	B+= “готов” cf 0
		b3 > 0 и b3 < 0,3	B+= “готов” cf 20
		b3 ≥ 0,3 и b3 ≤ 1	B+= “готов” cf 40
	b4 – риски отказа респондентов от участия в переписи населения на ПК.	b4 = 0	B+= “готов” cf 30
		b4 > 0 и b4 < 0,7	B+= “готов” cf 15
		b4 ≥ 0,7 и b4 ≤ 1	B+= “готов” cf 0
	b5 – риски по порче или утрате ПК в ходе проведения опроса	b5 = 0	B+= “готов” cf 30
		b5 > 0 и b5 < 0,5	B+= “готов” cf 15
		b5 ≥ 0,5 и b5 ≤ 1	B+= “готов” cf 0
С. Готовность ТОГС к проведению ВПН с ПК	c1 – качество методического материала для обучения привлекаемого персонала для работы на ПК.	c1 = 0	C+= “готов” cf 0
		c1 > 0 и c1 < 0,6	C+= “готов” cf 15
		c1 ≥ 0,6 и c1 ≤ 1	C+= “готов” cf 30
	c2 – сложность подбора персонала для проведения опроса населения на ПК.	c2 = 0	C+= “готов” cf 0
		c2 > 0 и c2 < 0,5	C+= “готов” cf 20
		c2 ≥ 0,2 и c1 ≤ 1	C+= “готов” cf 40
	c3 – риски отказа переписного персонала в ходе проведения переписи населения.	c3 = 0	C+= “готов” cf 30
		c3 > 0 и c3 < 0,7	C+= “готов” cf 15
		c3 ≥ 0,7 и c3 ≤ 1	C+= “готов” cf 0
	c4 – наличие труднодоступных районов, в которых применение планшетных компьютеров невозможно по объективным причинам (не возможность подзарядить ПК, транспортная недоступность, аномальные природные условия, или др. причины).	c4 = 0	C+= “готов” cf 20
		c4 > 0 и c4 < 0,5	C+= “готов” cf 10
		c4 ≥ 0,5 и c3 ≤ 1	C+= “готов” cf 0
	c5 – безопасность процесса проведения переписи населения для персонала с ПК в регионе (отдельном районе).	c5 = 0	C+= “готов” cf 0
		c5 > 0 и c5 < 0,6	C+= “готов” cf 15
		c5 ≥ 0,6 и c5 ≤ 1	C+= “готов” cf 30

в каждом регионе России за 2 месяца до начала переписи;

– по параметру С) Готовность ТОГС к проведению ВПН с ИКТ – проведение анкетирования штатных сотруд-

ников, задействованных для организации проведения переписей населения в РНИ.

Предлагается использовать статическую экспертную систему, основанную на об-

работке лингвистических переменных для качественного оценивания факторов, представленных в графе «И-ИЛИ».

Оценки по перечисленным подцелям формируются

с помощью таблиц решений, в которой каждая строка есть правило, определяющее зависимость значений анализируемых характеристик от конъюнкции факторов.

Выбор рационального метода [19,20] сбора и первичной обработки информации при проведении ВПН в РНИ по факторам A, B, C осуществляется с помощью нечетких правил оценки коэффициента уверенности cf переменных подцелей, представленных в таблице решений 3. Интервальные значения оцениваемых параметров заданы экспертным методом. Значения оцениваемых параметров конкретных регионов получают путем качественной интерпретации на шкалах суммарных оценок респондентов, заполняющих анкеты.

Общая оценка готовности к проведению ВПН с ИКТ по подцелям $X \in \{A, B, C\}$ осуществляется по рекурсивной процедуре:

$$\begin{aligned} 1. CF_{рез_i}(X) &= CF(X_i) \\ 2. CF_{рез_i}(X) &= \\ &= \sum_1^n (CF_{рез_{i-1}}(X_i) + CF(X_i) - \\ &- CF_{рез_{i-1}}(X_i) * CF(X_i) / 100) \quad (2) \end{aligned}$$

где $x \in X$ – параметр x_i подцели X , $i = 1, n$, $CF(X_i)$ – коэффициент уверенности x_i параметра, $CF_{рез_i}(X)$ – итоговый коэффициент уверенности подцели X , с учетом i -го параметра.

Аналогично по представленной рекурсивной процедуре определяется коэффициент уверенности целевой перемен-

ной G – «Готовность региона к проведению ВПН с ИКТ» CF_g по полученным коэффициентам уверенности подцелей CF_a, CF_b, CF_c .

В результате с помощью модели принятия решений предлагается производить выбор метода сбора и первичной обработки информации ВПН в регионе России, для которого не был однозначно сделан выбор метода на предыдущем этапе методики.

Заключение

Вместо ранее применявшегося одного традиционного метода сбора информации о населении на машиночитаемых бумажных переписных листах [3] предлагается определять предпочтительный метод сбора и первичной обработки информации о населении в регионах России на основе результатов кластеризации по статистическим параметрам, характеризующим особенности взаимодействие населения с государственными органами власти и местного самоуправления, уровню использования ИКТ в домашних хозяйствах, регистрации респондентов на порталах государственных и муниципальных услуг, доступности для населения сети Интернет в каждом регионе.

В соответствии с концепцией проведения ВПН в условиях региональных различий, влияющих на качество сбора и первичной обработки информации ВПН с использованием мобильных средств и Интер-

нет-технологий возникает необходимость по оптимизации расходов на проведение ВПН на основе выбора рациональных методов сбора и обработки первичной информации, что позволяет повысить максимальный охват населения при проведении переписи населения и оперативно получить достоверную статистическую информацию о населении [17] с одновременной минимизацией затрат на организационно-техническую схему проведения переписи населения в каждом регионе России.

Методика выбора рационального метода сбора и первичной обработки информации Всероссийской переписи населения с применением современных ИКТ на основе кластеризации регионов по двенадцати параметрам, характеризующим доступность к электронным государственным услугам, наличие доступа к сети Интернет, использование ИКТ в домашних хозяйствах, регистрацию респондентов на порталах государственных и муниципальных услуг, развитость и доступность ИКТ для населения в каждом регионе и анализа качественных характеристик развитости и доступности ИКТ для населения, готовности респондентов к участию в Интернет-переписи, готовности переписного персонала и штатного персонала территориальных органов государственной статистики к проведению ВПН с применением современных ИКТ целесообразно использовать на этапе подготовки к ВПН.

Литература

1. Андреев Е.М. О точности результатов российских переписей населения и степени доверия к разным источникам информации // Вопросы статистики. 2012. № 11. С. 21–35.
2. Клупт Н. А., Никифоров О. Н., Альтернативные методы проведения переписей в России, применимы ли они? // Вопросы статистики. 2010. № 9. С. 3–8.
3. Методология переписи: основные результаты проведенного ЕЭК ООН опроса по

практике проведения странами переписей и первоначальные предложения в отношении рекомендаций КЕС по проведению цикла переписей 2020 года. Записка Целевой группы ЕЭК ООН по методологии переписи. Европейская экономическая комиссия. Конференция европейских статистиков. Группа экспертов по переписям населения и жилищного фонда. Пятнадцатое совещание. Женева, 30 сентября – 3 октября 2013 года Пункт 2 предварительной повестки дня.

4. Суринов, А.Е. Всероссийская перепись населения: опыт и перспективы. [Электрон. ресурс]. Доклад на научно-практической конференции, Росстат, 27–28 ноября 2012 г. Режим доступа: http://www.gks.ru/free_doc/new_site/perepis2010/croc/perepis_itogi1612.htm.

5. Болотова Л.С. Системы искусственного интеллекта: модели и технологии, основанные на знаниях. М.: Финансы и статистика, 2012.

6. Рыбина Г.В. Основы построения интеллектуальных систем. М.: Финансы и статистика (Инфра-М), 2010.

7. Тельнов Ю.Ф., Трёмбач В.М. Интеллектуальные информационные системы. М.: МЭСИ, 2010.

8. Винтизенко И.Г., Черкасов А.А., Математические методы в экономике // Вестник Адыгейского государственного университета. Серия 5: Экономика. 2011. № 2.

9. Андрейчиков А.В., Андрейчикова О.Н. Анализ, синтез, планирование решений в экономике. М.: Финансы и статистика, 2000. 368 с.

10. Манжула О.В. Экспертная система выбора методов сбора и обработки информации Всероссийской переписи населения с учетом географических и социально-экономических особенностей регионов // Информационное общество. 2016. № 3. С. 57–66.

11. Тельнов Ю.Ф. Реинжиниринг бизнес-процессов: компонентная методология. М.: «Финансы и статистика», 2004. 319 с.

12. Романов А.Н., Одинцов Б.Е. Информационные системы в экономике. 2-е издание. М.: Вузовский учебник, 2008.

13. Божко В.П., Лури А.В., Сычев Е.Б. Совершенствование процессов проведения статистических переписей и обследований. М.: МЭСИ, 2008. 129 с.

14. Власов С.А., Девятков В.В., Кобелев Н.Б. Имитационные исследования: от классических технологий до облачных вычислений // Пятая всероссийская научно-практическая конференция «Имитационное моделирование, теория и практика. СПб.: ОАО ЦТСС. 2011. Т. 1. С. 42–50.

15. Волкова В.Н., Денисов А.А. Теория систем и системный анализ. М.: Изд-во «Юрайт», серия «Университеты России», 2010. 679 с.

16. Волкова В.Н., Денисов А.А. Теория систем и системный анализ. М.: Изд-во «Юрайт», 2012. 610 с.

17. Всероссийский центр изучения общественного мнения. Итоги переписи 2010. Пресс-выпуск. 2010. № 1621. 1 с.

18. Дрогобыцкий И.Н. Системный анализ в экономике. Учебник для студентов вузов, обучающихся по специальностям «Математические методы в экономике», «Прикладная информатика», 2-е изд. М.: ЮНИТИ-ДАНА, 2012. 423 с.

19. Дубров А.М., Лагоша Б.А., Хрусталев Е.Ю. Моделирование рискованных ситуаций в экономике и бизнесе. М.: Финансы и статистика, 2000. 176 с.

20. Девятков Т. В. Некоторые вопросы создания систем автоматизации имитационных исследований // Прикладная информатика. 2010. № 5 (29). С. 102–116.

References

1. O sostoyanii sanitarno-epidemiologicheskogo blagopoluchiya naseleniya Respubliki Mariy El v 2018 godu: Doklad. Yoshkar-Ola: Upravleniye Federal'noy sluzhby po nadzoru v sfere zashchity prav potrebiteley i blagopoluchiya cheloveka po Respublike Mariy El = On the state of the sanitary-epidemiological well-being of the population of the Republic of Mari El in 2018: Report. Yoshkar-Ola: Office of the Federal Service for Supervision of Consumer Rights Protection and Human Well-Being in the Republic of Mari El, 2019. 227 p. (In Russ.)

2. Varaksin A.N. Statisticheskiye modeli regressionnogo tipa v ekologii i meditsine = Statistic models of regression type in ecology and medicine. Yekaterinburg: Goshchitsky; 2006. 255 p. (In Russ.)

3. Sitnikova O.O. Assessment of environmental risks to public health in medical and demographic processes during anthropogenic air pollution (on the example of the Belgorod region). Materialy VI Mezhdunarodnoy studencheskoy nauchnoy konferentsii «Studencheskiy nauchnyy forum» = Materials of the VI International Student Scientific

Conference “Student Scientific Forum”. 2014. (In Russ.)

4. Veremchuk L.V., Cherpak N.A., Gvozdenko T.A., Volkova M.V. Methodology for assessing the effect of air pollution on the formation of levels of the general incidence of bronchial asthma. Gigiyena i sanitariya = Hygiene and sanitation. 2015; 94; 3: 119–122. (In Russ.)

5. Kurkatov S.V., Tikhonova I.V., Ivanova O.Yu. Risk assessment of the effects of atmospheric pollution on the health of the population of Norilsk. Gigiyena i sanitariya = Hygiene and sanitation. 2015; 94; 2: 28–31. (In Russ.)

6. Kanatnikova N.V., Yegorova N.A. The effect of hardness of drinking water on the incidence of the population of the city of Orel. Gigiyena i sanitariya = Hygiene and sanitation. 2017; 96; 3: 235–240. (In Russ.)

7. Luchaninova V.N., Tsvetkova M.M., Veremchuk L.V., Krukovich Ye.V., Mostovaya I.D. The health status of children and adolescents and factors affecting its formation. Gigiyena i sanitariya = Hygiene and sanitation. 96; 6: 561–568. (In Russ.)

8. Malysh N.G., Doan S.I. The use of factor analysis in the study of the epidemic process of acute intestinal infections. *Gigiyena i sanitariya = Hygiene and sanitation*. 2017; 96; 6: 519–523. (In Russ.)
9. Kiku P. F., Izmaylova O.A. Gorborukova T.V., Anan'yev V.Yu. The influence of environmental and hygienic environmental factors on the spread of respiratory diseases in the population of Primorsky Krai. *Gigiyena i sanitariya = Hygiene and sanitation*. 2012; 5. (In Russ.)
10. Trifonova T.A., Martsev A.A. Evaluation of the effect of air pollution on the incidence of the population of the Vladimir region. *Gigiyena i sanitariya = Hygiene and sanitation*. 2015; 94; 4. (In Russ.)
11. Zaykova Z.A., Burdukovskaya A.V., Belykh A.I. Determination of priority adverse environmental factors. *Gigiyena i sanitariya = Hygiene and sanitation*. 2016; 95; 12. (In Russ.)
12. Izmaylova O.A., Kiku P.F., Yarygina M.V., Moreva V.G., Anan'yev V.Yu., Kosolapov A.B. Hygienic aspects of the spread of environmentally dependent diseases in children and adolescents of the Primorsky Territory. *Gigiyena i sanitariya = Hygiene and sanitation*. 2016; 95; 11. (In Russ.)
13. Shcherbo A.P. Okruzhayushchaya sreda i zdorov'ye: podkhody k otsenke riska = Environment and health: risk assessment approaches. Saint Petersburg: SPbMAPO; 2002. 376 p. (In Russ.)
14. Gorodskiye okruga i munitsipal'nyye rayony Respubliki Mariy El: statisticheskiy sbornik = Urban districts and municipal districts of the Republic of Mari El: statistical compilation. Yoshkar-Ola: Maristat; 2017. 266 p. (In Russ.)
15. Respublika Mariy El v tsifrakh: Kratkii statisticheskiy sbornik = The Republic of Mari El in numbers: A brief statistical compilation. Yoshkar-Ola: Maristat. 2019; 385 p. (In Russ.)
16. Katsnel'son B.A., Kosheleva A.A., Privalova L.I., Kuz'min S.V., Malykh O.L., Khal'fin R.A., Nikonov B.I., Ozkaynak KH., Shu Dzh. The effect of short-term increases in air pollution on population mortality. *Gigiyena i sanitariya = Hygiene and sanitation*. 2000; 1: 15-18. (In Russ.)
17. Korotkov P.A., Trubyanov A.B., Zagaynova Ye.A., Nikonorov K.N. Comparative analysis of models for assessing the environmental performance of large cities. *Sovremennyye problemy nauki i obrazovaniya = Modern problems of science and education*. 2015; 2: 328. (In Russ.)
18. Korotkov P.A., Zagaynova Ye.A. The relationship between the level of suicide prevalence and working hours. *Statistika i ekonomika = Statistics and Economics*. 2017; 14; 4: 41–53. (In Russ.)
19. Shimanovich G. Vneshniy dolg Belarusi: opyt postsotsialisticheskikh stran. Rabochiy material Issledovatel'skogo tsentra IPM WP/09/01 = External debt of Belarus: the experience of post-socialist countries. Work material of the IPM Research Center WP / 09/01. 2009; 24 p. (In Russ.)
20. Kopnova Ye.D., Rozental' O.M. Econometric analysis of ecological management of fish resources. *Prikladnaya ekonometrika = Applied Econometrics*. 2010; 2 (18): 90–100. (In Russ.)
21. Larin A.V., Tarunina Ye.N. Entrepreneurial activity and the level of economic development: a form of dependence. *Prikladnaya ekonometrika = Applied Econometrics*. 2015; 37 (1): 3–26. (In Russ.)
22. Verbik M. Models based on panel data. *Prikladnaya ekonometrika = Applied Econometrics*. 2006; 1: 94–135. (In Russ.)
23. Swamy P.A. V.B. and Arora S.S. The exact finite sample properties of the estimators of coefficients in the error components regression model. *Econometrica*. 1972; 40: 261–275.

Сведения об авторе

Олег Владимирович Манжула

Заместитель начальника управления
информационных ресурсов и технологий,
Федеральная служба государственной
статистики, Москва, Россия
Эл. почта: <mailto:manzhula@gks.ru>

Information about the author

Oleg V. Manzhula

Deputy Head of the Department of Information
Resources and Technologies,
Federal State Statistic Service,
Moscow, Russia
E-mail: <mailto:manzhula@gks.ru>