

УДК: 528.854+519.253

DOI: 10.35595/2414-9179-2021-2-27-368-378

П.А. Украинский¹

ЭМПИРИЧЕСКИЙ МЕТОД ОПРЕДЕЛЕНИЕ ОПТИМАЛЬНОГО РАЗМЕРА СЛУЧАЙНОЙ ТОЧЕЧНОЙ ВЫБОРКИ ДЛЯ ОЦЕНКИ ПЛОЩАДЕЙ ЗЕМЕЛЬНЫХ УГОДИЙ ПО КОСМИЧЕСКИМ СНИМКАМ

АННОТАЦИЯ

Перспективным экспресс-методом для оценки площадей земельных угодий по космическим снимкам является использование случайной точечной выборки. Этот метод позволяет получить значения площадей без сплошного картографирования земельных угодий. При этом точность оценки площадей зависит от размера выборки. В представленной работе описан метод эмпирического определения оптимального размера выборки. Для использования этого метода необходимо выбрать ключевой участок с существующим эталонным слоем угодий. Далее производится многократная генерация выборок разного размера, по которым оценивается площадь угодий. Полученные площади сравниваются с эталонными и вычисляется относительная ошибка измерения. Анализ среднего значения и разброса ошибок при разных размерах выборки позволяет выявить тот момент, когда при увеличении размера выборки перестаёт существенно снижаться ошибка. Такой размер выборки является оптимальным. Предложенный метод опробован на примере Калачской возвышенности. Проанализирован размерный ряд от 100 до 3000 точек выборки на ключевой участок (величина выборки в ряду увеличивается на 100 точек). Для каждого элемента этого ряда было сгенерировано по 1000 выборок соответствующего размера. Проанализировано влияние размера выборки на общую относительную ошибку оценки площадей. Анализ показал, что для исследованного ключевого участка оптимальным является размер выборки равный 1000 точек (1,1 точка/км²). При таком размере выборки общая относительная ошибка определения площадей в среднем составила 4,0 %, а максимальная ошибка – 9,9%. Схожую точность следует ожидать при таком же размере выборки для других возвышенностей в лесостепной и степной зоне Восточно-Европейской равнины.

КЛЮЧЕВЫЕ СЛОВА: земельные угодья, измерение площади, случайная выборка, ошибки измерений, метод локтевого сгиба.

Pavel A. Ukrainskiy²

EMPIRICAL METHOD FOR ESTIMATION OF THE OPTIMUM SIZE OF RANDOM POINT SAMPLES FOR ASSESSMENT AREAS OF LAND COVER FROM SPACE IMAGES

ABSTRACT

A promising fast method for estimating land cover areas from satellite imagery is the use of random point sampling. This method allows you to obtain area values without spatially

¹ Белгородский государственный национальный исследовательский университет, Федерально-региональный центр аэрокосмического и наземного мониторинга объектов и природных ресурсов, ул. Победы, д. 85, 308015, Белгород, Россия; *e-mail*: pa.ukrainski@gmail.com

² Belgorod State National Research University, Federal and Regional Centre for aerospace and ground monitoring of objects and natural resources, Pobedy str., 85, 308015, Belgorod, Russia; *e-mail*: pa.ukrainski@gmail.com

continuous mapping of land areas. The accuracy of the area estimate by this method depends on the sample size. The presented work describes a method for empirically finding the optimal sample size. To use this method, you must select a key site for which a reference land cover exists. For the key site, we perform multiple generation of samples of different sizes. Further, using these samples, we estimate the area of land cover. Comparison of the obtained areas with the reference areas allows you to calculate the measurement error. Analysis of the mean and the range of errors for different sample sizes allows us to identify the moment when the error ceases to decrease significantly with an increase in the sample size. This sample size is optimal. We tested the proposed method on the example of the Kalach Upland. The size range from 100 to 3000 sampling points per key site is analyzed (the size of the sampling in the row increases by 100 points). For each element of this row, we created 1000 samples of the corresponding size. We then analyzed the effect of sample size on the overall relative error in area estimates. The analysis showed that for the investigated key site the optimal sample size is 1000 points (1.1 points / km²). With this sample size, the overall relative error in determining areas was 4.0% on average, and the maximum error was 9.9%. Similar accuracy should be at the same sample size for other uplands in the forest-steppe and steppe zones of the East European Plain.

KEYWORDS: land cover, area measurement, random sample, measurement error, elbow method.

ВВЕДЕНИЕ

Изучение распределения земельных угодий – одна из самых распространённых задач в практике использования космических снимков [Thyagarajan, 2019]. Обычно решение этой задачи подразумевает картографирование земельных угодий на основе визуального или автоматизированного дешифрирования снимков. Но иногда интерес могут представлять только площади угодий, а карта создаётся исключительно как промежуточный продукт, необходимый для измерения площадей. В таких случаях можно обойтись без сплошного картографирования изучаемой территории. Вместо этого для оценки площади угодий достаточно сделать случайную точечную выборку [Gallego, 2006].

При использовании точечной выборки в каждой точке по космическому снимку определяют угодье. Затем суммируют количество точек для каждого угодья и, зная общую площадь исследуемой территории, по пропорции вычисляют площади угодий. Если космические снимки дешифрируются визуальным способом, то использование случайной выборки играет роль экспресс-метода, который позволяет существенно сократить затраты времени. При автоматизированном дешифрировании применение такой метод позволяет избежать искажения площадей, вызванного эффектом смешанных пикселей. Точечную выборку имеет смысл использовать при наличии на снимке множества малых по площади объектов, размер которых превышает размер пикселя не более чем в 10–20 раз.

Перед использованием случайной точечной выборки для оценки площадей угодий требуется ответить на два вопроса. Первый вопрос – это как генерировать точки случайной выборки? Второй вопрос – сколько точек должно быть в выборке, чтобы полученная оценка площадей имела приемлемую точность? Первый вопрос не представляет большого интереса для специалистов в области геоинформатики. В большинстве ГИС-программ реализованы инструменты для генерации слоёв случайно расположенных точек. Стоит отметить лишь то, что качество работы этих инструментов в отдельных программах вызывает споры [Van Niel et al., 2003; Barry, 2011]. Но в случае сомнений можно воспользоваться альтернативой в виде программ для статистического анализа, где генерация случайных выборок реализована на более серьёзном уровне.

Второй вопрос является ключевым при выборочной оценке площадей угодий. Ведь чем больше точек в выборке, тем точнее результат. Однако размер выборки невозможно

увеличивать до бесконечности. Необходимо найти оптимальный размер. В этой работе мы предлагаем способ эмпирического определения оптимального размера случайной точечной выборки.

МАТЕРИАЛЫ И МЕТОДЫ ИССЛЕДОВАНИЯ

Если с помощью случайной точечной выборки по космическим снимкам определяются площади земельных угодий на значительных территориях, то оптимальный размер выборки можно определить опытным путём на небольшом ключевом участке. Он должен отражать разнообразие угодий и сложность их пространственной структуры, присущие всей исследуемой территории в целом. Для ключевого участка проводится эксперимент, состоящих из пяти этапов. На первом этапе выполняется картографирование земельных угодий на основе дешифрирования космических снимков (может быть заменено использованием уже существующих пространственных данных об угодьях). В результате получается слой угодий с полигональной геометрией, который в дальнейшем будет использоваться как эталон. На его основе рассчитываются площади угодий, с которыми будут сравниваться площади, оценённые с помощью точечных выборок.

На втором этапе выбирается ряд размеров выборки. Для каждого элемента в этом ряду выполняется множественная генерация случайных выборок соответствующего размера. На третьем этапе слои точечных выборок накладываются на полигональный эталонный слой. В атрибутивные таблицы точечных слоёв извлекаются названия угодий из эталонного слоя. В каждой из множества выборок суммируются число точек с различными угодьями. Затем по пропорции количество точек пересчитывается в площади угодий.

На четвёртом этапе для каждой выборки вычисляется общая относительная ошибка измерения площадей угодий. Расчёт проводится по формуле 1.

$$E = \frac{\sum_{i=1}^n |a_i - s_i|}{A} \cdot 100 \quad (1)$$

В формуле 1 использованы следующие обозначения: E – относительная ошибка, A – сумма площадей всех угодий (она же общая площадь ключевого участка); a_i – эталонная площадь угодья i ; s_i – площадь угодья i , рассчитанная на основе точечной выборки; n – количество угодий на ключевом участке. Умножение на 100 производится чтобы выразить относительную ошибку в процентах.

На пятом этапе анализируется, как изменяется ошибка в зависимости от размера выборки. Для каждого размера выборки вычисляется описательная статистика ошибок – среднее, максимальное и минимальное значение, стандартное отклонение ошибки. Далее строится график, по вертикальной оси которого откладываются характеристики ошибок, а по горизонтальной – размер выборки. Полученные кривые на графике будут иметь вогнутую форму, т.к. с увеличением размера выборки ошибки до определённого предела интенсивно уменьшаются, а затем начинают уменьшаться уже постепенно. Точка сгиба указывает на оптимальный размер выборки. Для разных характеристик ошибки точка сгиба может наблюдаться при разных размерах выборки. В таком случае следует считать оптимальным самый большой размер выборки из полученных вариантов.

Описанный выше графический метод в статистике известен как метод «локтевого сгиба». Графики с «локтевым сгибом» используют в кластерном анализе для определения оптимального числа кластеров, в преобразовании главных компонент для определения оптимального числа главных компонент, в факторном анализе для определения числа извлекаемых факторов. При визуальном анализе графиков точка сгиба не всегда однозначно идентифицируется, и интерпретация графиков может носить субъективный характер [Украинский, 2020]. Но в настоящее время уже существуют количественные

методы определения точки «локтевого сгиба» на кривой. Мы предлагаем использовать анализ кривых методом UIK (Unit Invariant Knee), описанный в 2016 году Demetris Christopoulos¹. В дополнение к методу «локтевого сгиба» также полезным может быть построение коробчатых диаграмм (боксплотов) относительной ошибки для разных размеров выборки. Этот способ является хорошей визуализацией существующего разброса значений.

Чтобы продемонстрировать работу описанного выше способа определения оптимального размера случайной точечной выборки, мы выбрали ключевой участок (рис. 1) расположенный на северо-востоке Калачской возвышенности [Мильков и др., 1972].

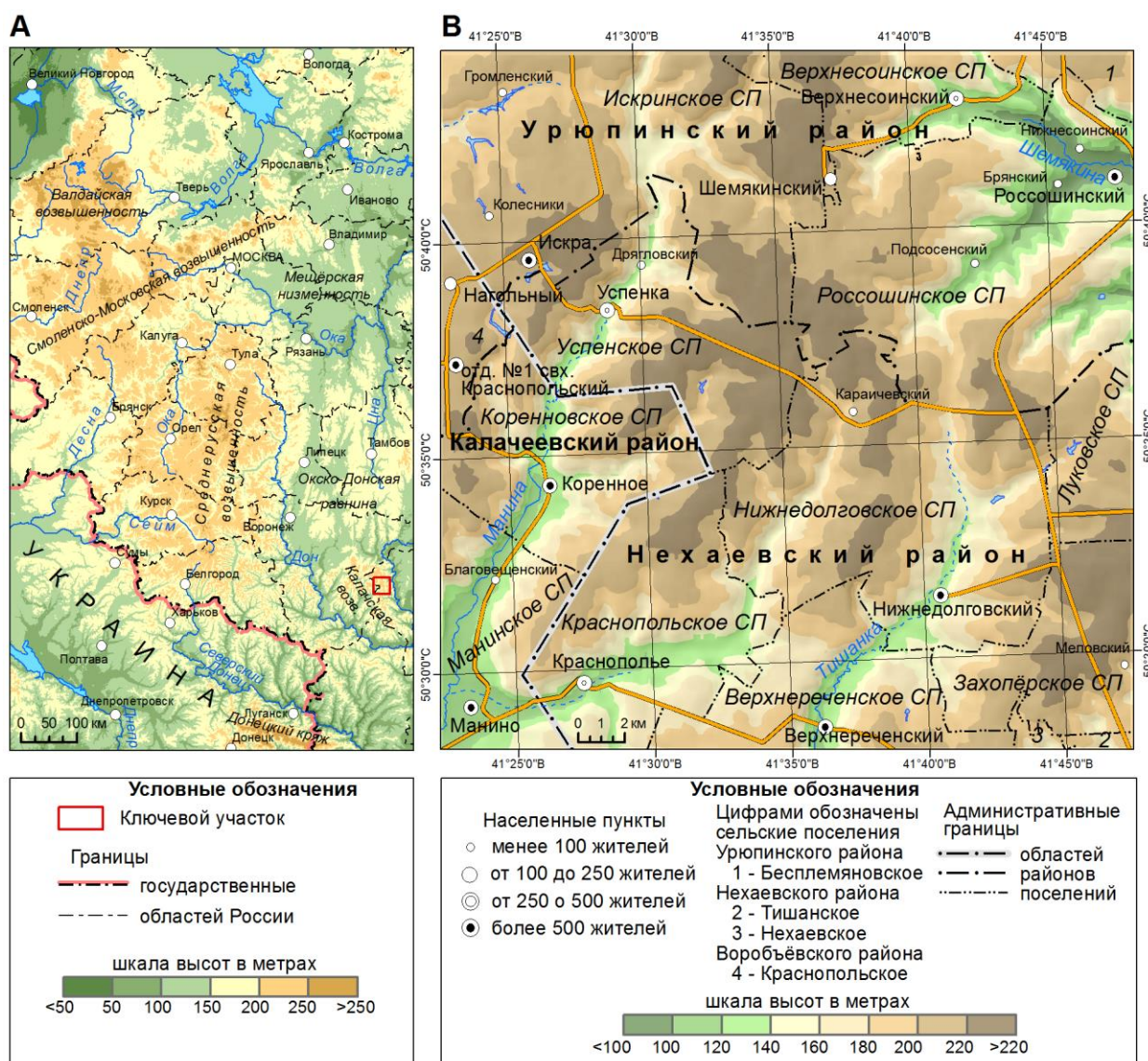


Рис. 1. Схема расположения (А) и карта (В) исследуемого участка
 Fig. 1. Location (A) and map (B) of the study area

¹ Christopoulos D.T. Introducing unit invariant knee (UIK) As an objective choice for elbow point in multivariate data analysis techniques. SSRN, 2016. URL: <https://ssrn.com/abstract=3043076> (Available at 30.04.2021). DOI: 10.2139/ssrn.3043076.

Пространственная структура угодий на ключевом участке характерна для возвышенностей в лесостепной и степной зоне Русской равнины (Калачская, Средне-русская, Приволжская возвышенность). Участок представляет собой квадрат размером 30 на 30 км, находящийся на стыке Воронежской (Воробьевский и Калачеевский район) и Волгоградской области (Урюпинский и Нехаевский район) (рис. 1).

По снимкам со спутников Sentinel-2A и Sentinel-2B на ключевом участке были закартографированы (отвекторизованы) земельные угодья. Использовалась серия безоблачных снимков, снятых в разные сезоны 2019–2020 года (табл. 1). Необходимость набора разносезонных снимков обусловлена тем, что для различных угодий есть сезоны, в которые они лучше или хуже распознаются. Кроме того, сезонная динамика изображения является одним из дешифровочных признаков земельных угодий.

Снимки получены из веб-сервиса Land Viewer¹. С помощью этого сервиса снимки в режиме онлайн были обрезаны по экстенду исследуемого участка, и в формате geotiff выгружены отдельные каналы снимков (2, 3, 4, 8, 11 и 12 каналы). При дешифрировании снимков использовалось несколько комбинаций каналов Sentinel-2. Комбинация каналов 4-3-2 обеспечивает отображение объектов в натуральных цветах. В комбинации каналов 8-11-4 лучше всего выделяются водоёмы. В комбинации каналов 12-11-4 хорошо распознаются искусственные поверхности в населённых пунктах и открытая поверхность почвы на полях. Комбинации каналов 8-11-2 и 8-4-3 эффективны при дешифрировании растительности.

Табл. 1. Космические снимки, использованные для картографирования земельных угодий на ключевом участке

Table 1. Satellite imagery used to map land cover at the key site

Дата съёмки	Спутник	Номер сцены	Номер тайла
07.03.2019	Sentinel-2B	N0207/R121	37UFC
31.03.2020	Sentinel-2B	N0209/R121	37UFC
30.04.2020	Sentinel-2B	N0209/R121	37UFC
4.07.2020	Sentinel-2A	N0209/R121	37UFC
07.10.2020	Sentinel-2B	N0209/R121	37UFC
01.12.2020	Sentinel-2A	N0209 R121	37UFC

Предварительная обработка снимков выполнена в программе ENVI 4.8. В ней отдельные каналы снимков были объединены в многоканальные файлы и пространственное разрешение приведено к 10 м/пиксель. Векторизация земельных угодий выполнена в программе ArcGIS 10.5. В результате был создан векторный слой (шейп-файл) угодий с полигональной геометрией. Для него был осуществлен поиск и исправление топологических ошибок двух типов – накладываются полигоны и пробелов между полигонами. В дальнейшем полученный слой рассматривался нами в качестве эталона. По нему были рассчитаны площади угодий, используемые для определения ошибки выборочных оценок площадей.

При дешифрировании все многообразие угодий было сведено к семи классам: пашня, сады, травянистая растительность, леса, застройка, водоёмы, болота. К застроенным территориям на ключевом участке были отнесены сельская застройка вместе огородами, животноводческие фермы, асфальтированные автомобильные дороги. В класс травостоев

¹ Land Viewer. URL: <https://eos.com/landviewer> (Available at 30.04.2021)

включены различные угодья с травянистой растительностью – пастбища и сенокосы (как используемые, так и неиспользуемые), залежи, участки травянистой растительности на месте исчезнувших населённых пунктов, лесные поляны и вырубki, на которых ещё не началось возобновление древесной растительности.

В состав класса лесов включены все леса, в том числе и не входящие в лесной фонд, а также заросли кустарников. Сомкнутые древостои на заболоченных территориях (в основном ольшаники) отнесены тоже к лесам, а не к болотам. Также к лесам причислены лесополосы.

В состав пашни при дешифрировании включались только те участки, на которых фактически возделываются сельскохозяйственные культуры (то есть все участки, кроме залежей). К садам при дешифрировании причислялись только промышленные сады – насаждения многолетних плодовых культур, возделываемые агропредприятиями. Сады на приусадебных участках в населённых пунктах были отнесены к застроенным территориям. В состав класса водоемов включены озера, искусственные водоемы (пруды) и водотоки (реки). Болота на ключевом участке представлены низинными осоковыми и тростниковыми болотами.

После векторизации слоя угодий дальнейшая обработка данных проводилась с помощью языка программирования R 3.4.4¹ в интегрированной среде разработки RStudio 1.1.453. Необходимость использования R вызвана тем, что предлагаемая нами методика поиска оптимального размера точечной выборки основана на последовательной генерации многих тысяч случайных выборок. При использовании ГИС-программ этот процесс будет занимать слишком много времени. А в R имеются удобные функции автоматизации однотипных повторяющихся процессов.

Импорт шейп-файла угодий в R произведён с помощью библиотек GDAL/OGR, реализованных в дополнительном пакете *rgdal*². Для поиска оптимального размера выборки была проанализирована серия из 30 размеров. Размер выборки в этой серии увеличивается на 100 точек. Самая маленькая выборка включает 100 точек, самая большая – 3000 точек. Для каждого размера было сгенерировано 1000 выборок. На этом этапе использовались функции из дополнительного пакета *sp* [Pebesma et al., 2005; Bivand et al., 2013]. Функция *spsample* применялась для генерации случайных точечных выборок, функция *over* – для извлечения по точкам выборки названия угодий из полигонального слоя угодий. Для определения точек сгиба на графике характеристик относительной ошибки использовался анализ кривых методом UK, реализованным в дополнительном пакете *inflection*³.

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ И ИХ ОБСУЖДЕНИЕ

По результатам дешифрирования космических снимков для исследуемой территории получены следующие площади земельных угодий. Леса занимают 102,0 га, травостой – 217,4 га, сады – 0,9 га, пашня – 542,2 га, застройка – 26,7 га, болота – 9,1 га, водоемы – 1,7 га. Эти значения были использованы в качестве эталонных для расчёта относительной ошибки выборочной оценки площадей земельных угодий. Распределение угодий в пространстве показано на рисунке 2. Видно, что пространственная структура

¹ R Core Team (2018). The R project for statistical computing. Web resource: <https://www.R-project.org/> (accessed 30.04.2021).

² R. Bivand, T. Keitt, B. Rowlingson. *Rgdal: bindings for the geospatial data abstraction library*. R package version 1.3-6. URL: <https://CRAN.R-project.org/package=rgdal> (Available at 30.04.2021).

³ D.T. Christopoulos (2019). *inflection: Finds the Inflection Point of a Curve*. R package version 1.3.5. URL: <https://CRAN.R-project.org/package=inflection> (Available at 30.04.2021).

угодий отличается сильной мозаичностью. Такая картина характерна в целом для всех возвышенностей в лесостепной и степной зоне Восточно-Европейской равнины.

График зависимости величины ошибки от размера выборки показан на рисунке 3. На нем видно, что с увеличением числа точек в выборке повышается точность определения площадей угодий. Чем больше выборка, тем меньше средняя величина ошибки и тем меньше размах возможных значений ошибки. Так при 100 точках выборки ошибка в среднем составляет 12,3%, максимальная ошибка равна 34,5%, а стандартное отклонение ошибки равно 1,3%. При 3000 точках выборки средняя ошибка снижается до 2,3%, максимальная ошибка составляет 5,9%, а стандартное отклонение ошибки равно 1,0%.

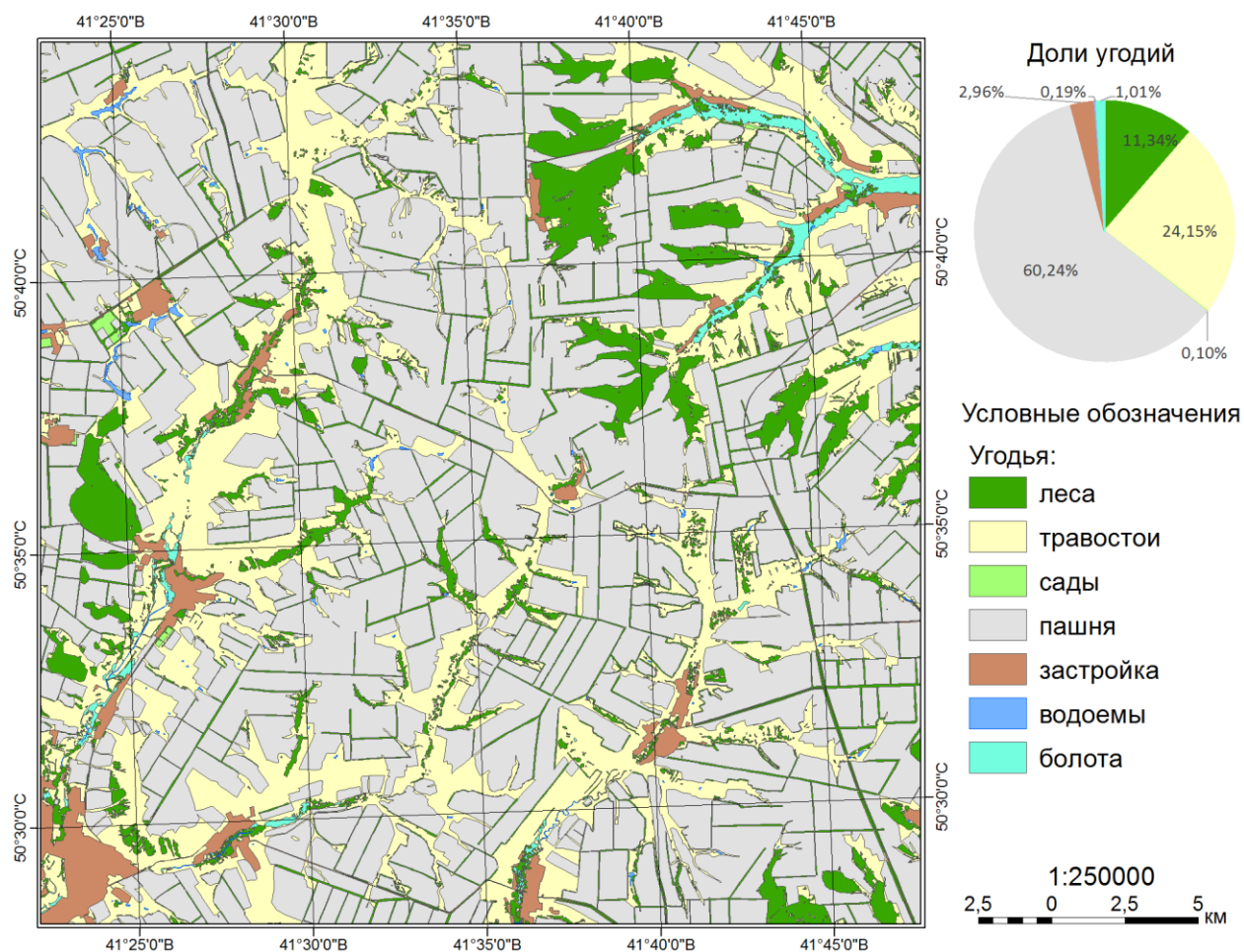


Рис. 2. Карта земельных угодий ключевого участка, составленная по результатам дешифрирования снимков со спутников Sentinel-2A/B

Fig. 2. Map of land cover of the key site, created based on the analysis of images from the Sentinel-2A/B satellites

С увеличением размера выборки ошибка оценки площадей угодий уменьшается неравномерно. Вначале происходит резкое уменьшение, затем оно сменяется на плавное. Наблюдается характерная форма кривых на графике – в виде так называемого «локтевого сгиба». При этом точка сгиба для кривых среднего значения ошибки и стандартного

отклонения ошибки различается не чётко. Она расположена где-то между размерами выборки в 500 и 1000 точек. Для максимального значения ошибки ситуация более однозначная. Здесь перегиб кривой на графике четко наблюдается при размере выборки в 1000 точек.

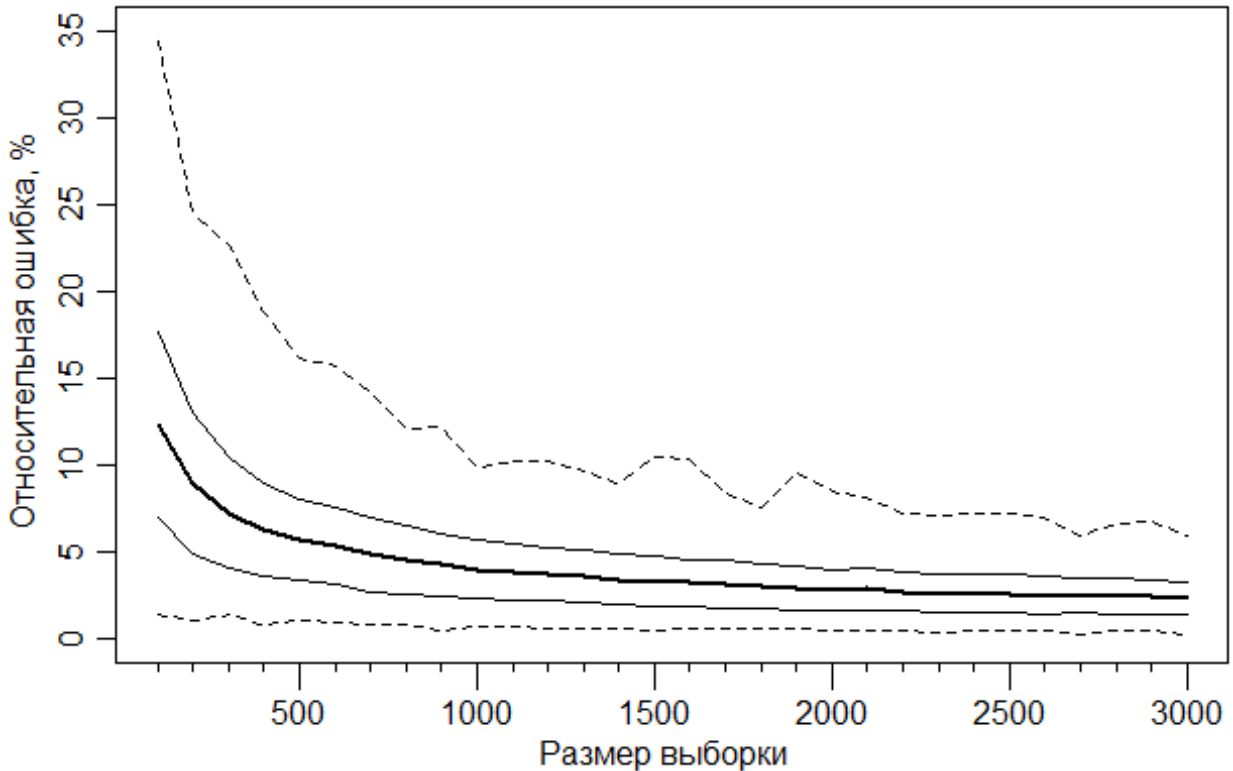


Рис. 3. Зависимость относительной ошибки определения площадей угодий от размера выборки. Толстая линия – среднее значение, тонкая линия – стандартное отклонение, пунктирная линия – минимальное и максимальное значение

Fig. 3. Dependence of the relative error in the assessment of land cover areas on the sample size. The thick line is the mean, the thin line is the standard deviation, and the dashed line is the minimum and maximum values

В качестве альтернативного метода графического анализа были построены боксплоты (коробчатые диаграммы, диаграммы размахов). Использовался вариант графика с насечками, показывающими доверительный интервал для медианы (рис. 4). Если доверительные интервалы боксплотов не перекрываются, то медианы статистически значимо отличаются.

На рисунке 4 можно наблюдать что медианы соседних боксплотов значимо отличаются при размерах выборки от 100 до 1000 точек. А начиная с размера в 1000 точек доверительные интервалы медиан начинают перекрываться. Это можно рассматривать как дополнительный аргумент в пользу того, что 1000 точек является оптимальным размером выборки.

Анализ с помощью UIK-метода показал, что для кривой среднего значения точкой сгиба является размер выборки равный 700 точек, для кривой стандартного отклонения –

500 точек, для кривой максимального значения – 1000 точек, для кривой минимального значения – 900 точек.

Помимо показателей описательной статистики ошибок, показанных на рисунки 3, с помощью УИК-метода были проанализированы медиана, межквартильных размах и среднее абсолютное отклонение ошибок. Для кривой медианы ошибок точкой сгиба является размер выборки в 700 точек, для межквартильного размаха – 700 точек, для среднего абсолютного отклонения ошибок – 500 точек.

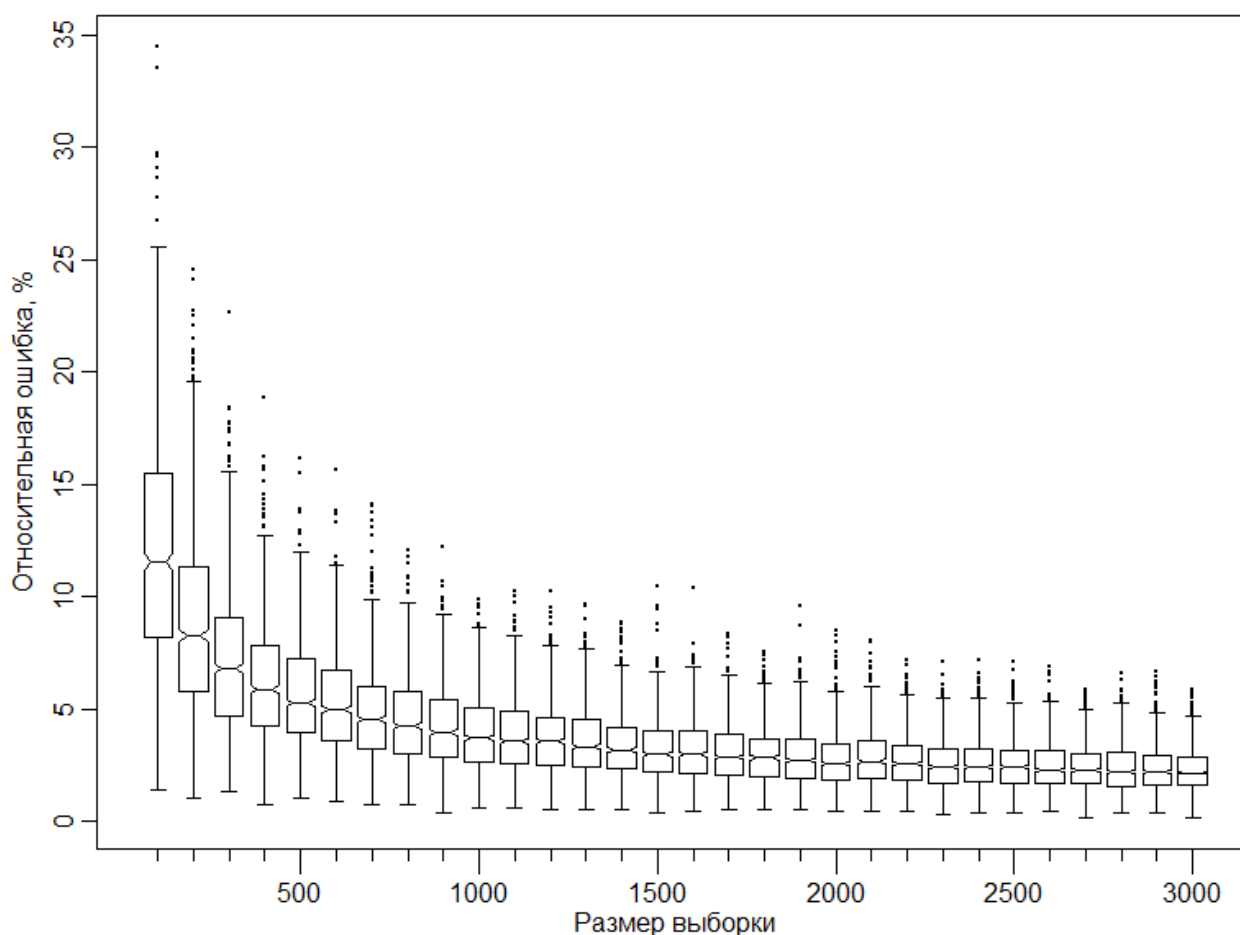


Рис. 4. Диаграммы размахов (боксплоты) значений относительной ошибки определения площадей угодий

Fig. 4. Boxplots of the values of the relative error in determining the areas of land cover

Поскольку самым большим значением точки сгиба, полученным с помощью УИК-анализа кривых ошибок, является 1000 точек, то это значение следует принять в качестве оптимального размера выборки. В пересчёте на единицу площади это составляет 1,1 точка/км². При таком размере выборки общая относительная ошибка определения площадей в среднем составила 4,0 %, а максимальная ошибка – 9,9%, минимальная ошибка – 0,6%, стандартное отклонение ошибок – 1,7%. Схожую точность следует ожидать при таком же размере выборки для других возвышенностей в лесостепной и степной зоне Восточно-Европейской равнины.

ВЫВОДЫ

Использование случайной точечной выборки позволяет оценить по космическим снимкам площади земельных угодий без выполнения сплошного картографирования. При правильно подобранном размере выборки достижима высокая точность определения площадей. При изучении крупных территорий размер выборки можно определить эмпирическим путём, поставив эксперимент на локальном ключевом участке для которого есть эталонный слой земельных угодий. Для этого необходимо произвести многократную генерацию случайных выборок, в точках которых по эталонному слою определяются угодья. Полученные данные позволят проследить, как меняется ошибка оценки площади в зависимости от размера выборки. Такой эксперимент будет слишком трудоёмок при использовании ГИС-программ. Поэтому необходимы средства, позволяющие эффективно автоматизировать вычислительные процессы. В качестве такого средства хорошо зарекомендовала себя среда для статистических вычислений R с дополнительным пакетом *sp*.

БЛАГОДАРНОСТИ

Исследование выполнено при финансовой поддержке РФФ, грант 20-67-46017.

ACKNOWLEDGEMENTS

The study was funded by the Russian Science Foundation, grant No 20-67-46017.

СПИСОК ЛИТЕРАТУРЫ

1. Мильков Ф.Н., Ахтырцева Н.И., Ахтырцев Б.П. Калачская возвышенность Воронеж: Издательство Воронежского университета, 1972. 178 с.
2. Украинский П.А. Выбор оптимального порядка соседства для разделения пространственного точечного образа на кластерную и шумовую составляющую (на примере анализа размещения античных поселений на Керченском полуострове). *ИнтерКарто. ИнтерГИС*. 2020. Т. 26. Ч. 4. С. 257–265. DOI: 10.35595/2414-9179-2020-4-26-257-265.
3. Barry S.C. How much impact does the choice of a random number generator really have? *International Journal of Geographical Information Science*. 2011. V. 25. No 4. P. 523–530. DOI: 10.1080/13658810903093185.
4. Bivand R.S., Pebesma E., Gomez-Rubio V. *Applied spatial data analysis with R*. New York: Springer-Verlag, 2013. 405 p. DOI: 10.1007/978-1-4614-7618-4.
5. Gallego F.J. Review of the main remote sensing methods for crop area estimates. *ISPRS Archives*, 2006. V. XXXVI. No 8/W48. P. 65–70.
6. Pebesma E.J., Bivand R.S. Classes and methods for spatial data in R. *R News*, 2005. V. 5. No 2. P. 9–13.
7. Thyagarajan K.K., Vignesh T. Soft computing techniques for land use and land cover monitoring with multispectral remote sensing images: a review. *Archives of Computational Methods in Engineering*, 2019. V. 26. No 2. P. 275–301. DOI: 10.1007/s11831-017-9239-y.
8. Van Niel K.P., Laffan S.W. Gambling with randomness: the use of pseudorandom number generators in GIS. *International Journal of Geographical Information Science*, 2003. V. 17. No 1. P. 49–68. DOI: 10.1080/713811743.

REFERENCES

1. *Barry S.C.* How much impact does the choice of a random number generator really have? *International Journal of Geographical Information Science*. 2011. V. 25. No 4. P. 523–530. DOI: 10.1080/13658810903093185.
 2. *Bivand R.S., Pebesma E., Gomez-Rubio V.* *Applied spatial data analysis with R*. New York: Springer-Verlag, 2013. 405 p. DOI: 10.1007/978-1-4614-7618-4.
 3. *Gallego F.J.* Review of the main remote sensing methods for crop area estimates. *ISPRS Archives*, 2006. V. XXXVI. No 8/W48. P. 65–70.
 4. *Milkov F.N., Akhtyrseva N.I., Akhtyrsev B.P.* *Kalachskaya Upland Voronezh*: Voronezh University Publishing House, 1972. 178 p.
 5. *Pebesma E.J., Bivand R.S.* Classes and methods for spatial data in R. *R News*, 2005. V. 5. No 2. P. 9–13.
 6. *Thyagarajan K.K., Vignesh T.* Soft computing techniques for land use and land cover monitoring with multispectral remote sensing images: a review. *Archives of Computational Methods in Engineering*, 2019. V. 26. No 2. P. 275–301. DOI: 10.1007/s11831-017-9239-y.
 7. *Ukrainskiy P.A.* The choice of the optimal order of the neighborhood for separation a spatial point pattern into a cluster and noise component (by the example of analysis of location of antique settlements in the Kerch Peninsula). *InterCarto. InterGIS*. 2020. V. 26 (4). P. 257–265.
 8. *Van Niel K.P., Laffan S.W.* Gambling with randomness: the use of pseudorandom number generators in GIS. *International Journal of Geographical Information Science*. 2003. V. 17. No 1. P. 49–68. DOI: 10.1080/713811743.
-