

## Оценивание качества идентификации состояния лесной растительности на примере заказника «Озеро Щучье»

В. А. Зеленцов, В. Ф. Мочалов

Санкт-Петербургский федеральный исследовательский центр РАН  
Санкт-Петербург, 199178, Россия  
E-mail: vicavia@yandex.ru

Рассматриваются вопросы совместного и одновременного решения задач обработки материалов мультиспектральной космической съёмки и оценивания качества результатов обработки при идентификации состояния лесной растительности. Подчёркивается роль обоснованного формирования для этих задач состава исходных данных о принадлежности лесной растительности тому или иному классу. Представлена методика, включающая предварительное и финальное оценивание качества обработки материалов съёмки. Ключевым элементом методики является новый подход к определению на предварительном этапе множества элементарных площадок (пикселей) фрагмента анализируемой сцены для их использования в качестве исходных данных при обработке материалов съёмки. Выбор элементарных площадок основан на применении нечёткой кластеризации и количественном анализе степени принадлежности каждого пикселя одному из классов идентифицируемых элементов лесного ландшафта. Значение степени принадлежности используется для определения состава элементарных площадок, требующих уточнения, в том числе проведением наземных обследований. Финальная обработка данных и оценивание качества результатов обработки для всей анализируемой сцены выполняются с использованием уточнённых исходных данных. Анализ качества обработки по предложенной методике продемонстрирован на примере задачи классификации состояния елового леса по классам пожарной опасности на территории заказника «Озеро Щучье» в Ленинградской области. В качестве материалов съёмки используются мультиспектральные данные космического аппарата Sentinel-2, для их обработки применены алгоритмы расчёта различных вегетационных индексов. Для обоснованного выбора состава исходных данных выполнена предварительная нечёткая кластеризация фрагмента сцены из 400 пикселей, определены 39 пикселей, принадлежность которых нуждается в уточнении, выполнена обработка всей сцены на основе уточнённых значений. Показано, что применение методики позволяет улучшить финальные значения показателей качества обработки, существенно сократить временные затраты на подготовку исходных данных, а также дать рекомендации по выбору алгоритмов обработки, обеспечивающих наилучшее качество идентификации. В рассмотренной задаче это алгоритмы расчёта NDVI (*англ.* Normalized Difference Vegetation Index) и ARVI (*англ.* Atmospherically Resistant Vegetation Index). Наиболее эффективной областью использования методики является решение задач обработки мультиспектральных данных применительно к относительно небольшим участкам территории, поскольку она позволяет максимально полно учесть локальные особенности и конкретные спектрально-отражательные характеристики исследуемых элементов ландшафта.

**Ключевые слова:** материалы мультиспектральной космической съёмки, показатели качества результатов обработки, исходные данные, алгоритмы обработки, идентификация состояния лесной растительности, нечёткая кластеризация, наземные обследования, классы пожарной опасности

Одобрена к печати: 18.12.2025

DOI: 10.21046/2070-7401-2026-23-2-173-187

### Введение

При идентификации состояния лесной растительности для последующего соотнесения оценки с классами пожарной опасности могут быть применены различные методы и алгоритмы обработки мультиспектральных данных, включая алгоритмы машинного обучения, нейросетевые алгоритмы и алгоритмы, предусматривающие расчёт вегетационных индексов (Liu et al., 2023; Schowengerdt, 2007; Wang, Yao, 2024). При этом недостаточно исследованным остаётся вопрос об оценивании качества результатов, полученных с использованием того

или иного метода обработки. Вместе с тем наличие количественной оценки качества, сопровождающей результаты обработки, является важным фактором, определяющим уверенность пользователя в выборе того или иного решения по управлению состоянием лесов (Зеленцов, Мочалов, 2024а). В общем случае показатели качества результатов идентификации состояния лесной растительности представляют собой метрики, позволяющие оценить, насколько результаты обработки материалов съёмки соответствуют эталонным данным или реальному состоянию растительности, отражённому в контрольных выборках.

Качество результатов обработки и значения соответствующих показателей (метрик) зависят от информативности материалов мультиспектральной съёмки, используемых алгоритмов расчёта и программного обеспечения, а также от состава исходных данных, необходимых для работы алгоритмов. Исходные данные, как правило, включают в себя значения характеристик определённого множества элементарных площадок (пикселей), расположенных в границах анализируемой сцены. Причём эти данные используются как в самих алгоритмах расчёта, так и при формировании контрольной выборки и оценивании качества результатов расчётов. Проведённый анализ показывает, что именно от выбора элементарных площадок в значительной степени зависят как качество результатов обработки материалов космической съёмки, так и сама возможность достоверного оценивания этого качества. Так, авторы базы данных пространственного распределения древесных пород на тестовом полигоне «Ляльский», Республика Коми (Мыльникова и др., 2025), по этой причине столкнулись с трудностями «оценивания точности классификации породного состава крон» на основе построения «матрицы ошибок». Исследования по определению сравнительных характеристик алгоритмов обработки материалов съёмки на основе применения различных вегетационных индексов при оценке состояния растительного покрова (Комаров и др., 2021; Саворский и др., 2016), как правило, содержат только качественные оценки преимущества того или иного алгоритма.

В исследованиях, предусматривающих оценивание качества результатов обработки, приводятся примеры использования различных подходов к формированию совокупности исходных данных. Так, для 1200 тестовых площадок на территории одной из провинций Китая с различными известными видами лесных растительных сообществ, учитываются данные, накопленные в течение нескольких лет (Liu et al., 2024). Также распространён способ, когда в качестве источника достоверных исходных данных для контрольной выборки принимаются результаты интерактивной обработки материалов аэрокосмической съёмки сверхвысокого (несколько сантиметров) пространственного разрешения (Ferreira et al., 2019). Однако при использовании таких подходов оценка качества результатов обработки остаётся трудно решаемой задачей. Причиной этого является необходимость регулярного контроля тестовых площадок на больших по площади территориях или привлечение экспертов, которые, в свою очередь, могут вносить ошибки в подготавливаемые данные.

Наиболее распространённым стал способ формирования исходных данных с использованием баз данных или библиотек спектральных отражательных характеристик (Мальшев, Фомин, 2019) типовых элементов ландшафта. Однако такой способ не всегда позволяет обеспечить сопоставление результатов измерений с данными космической съёмки на конкретной анализируемой территории в условиях сезонной изменчивости. Наиболее достоверная база исходных данных может быть получена наземными измерениями на всей площади тестовых площадок синхронно с выполнением космической съёмки. Но такой метод считается чрезвычайно затратным и не всегда может быть реализован.

В настоящей работе рассматривается методика оценивания качества результатов идентификации состояния лесной растительности, основанная на новом подходе к формированию массива исходных данных о состоянии лесных насаждений. В качестве исходных данных для оценивания результатов идентификации используются характеристики ограниченного множества элементарных площадок (пикселей), выбранных специальным образом в пределах анализируемой сцены — части условного кадра материалов съёмки. Процедура выбора состава элементарных площадок основана на нечёткой кластеризации и предварительном оценивании степени принадлежности каждого пикселя тому или иному классу лесной растительно-

сти. При этом характеристики элементарных площадок определяются максимально синхронно со временем проведения съёмки.

### Исходные данные и методика оценивания

Предлагаемая методика разрабатывается применительно к классификации состояния одного из видов лесных насаждений — елового леса, подверженного усыханию.

Классы пожарной опасности ельника, соответствующие Приказу Рослесхоза (Приказ..., 2011), приведены в *табл. 1*.

Таблица 1. Идентифицируемые виды состояния ельника

Класс пожарной опасности	Вид состояния ельника
I — очень высокий	Сухостой в стадии вымирания растительности
II — высокий	Сухостой высокой стадии отмирания растительности
III — средний	Сухостой средней стадии отмирания растительности
IV — слабый	Сухостой начальной стадии, или отсутствуют признаки усыхания

Исходные данные, необходимые для оценивания качества обработки материалов съёмки, включают: 1) множество элементарных площадок (пикселей), имеющих строгую координатную привязку и расположенных в границах анализируемой сцены (части условного кадра материалов съёмки), а на предварительном этапе — в границах ограниченного фрагмента сцены; 2) спектральные отражательные характеристики лесной растительности этих площадок; 3) семантическое описание идентифицируемых элементов лесного ландшафта. Количество и состав пикселей фрагмента сцены определяется содержанием тематической задачи обработки и должно обеспечивать выполнение свойства репрезентативности. Репрезентативность означает соответствие статистических показателей спектральных характеристик пикселей выбранного фрагмента сцены аналогичным показателям пикселей в пределах всей сцены. В данном случае в ходе дальнейшего описания будем рассматривать суть предлагаемого подхода на примере первоначального разделения лесной растительности на два класса — условно здоровый (соответствует классу IV пожарной опасности) и условно усыхающий (объединяет классы I, II и III) ельник с последующим разделением условно усыхающего леса на три класса.

Выбор показателей для оценивания качества обработки в общем случае определяется положениями статистической теории проверки гипотез (Bradley, 1997). В разрабатываемой методике учитывается тот факт, что на практике часто встречаются задачи семантической сегментации элементов лесного ландшафта для несбалансированного набора исходных данных, а также задачи, в которых следует учитывать различную цену возможных ошибок идентификации элементов ландшафта. В этом случае при оценивании качества результатов обработки не всегда удобно опираться на традиционные показатели в виде ошибок первого и второго рода (Горелик, Скрипкин, 2004), обобщённых показателей точности результатов обработки материалов съёмки (*англ.* accuracy) (Maxwell et al., 2021a, b), или вероятности верного распознавания отдельного элемента ландшафта (Congalton, Green, 2009). В ряде исследований показано, что более приемлемыми следует считать следующие базовые показатели: TP (*англ.* True Positive), TN (*англ.* True Negative), FP (*англ.* False Positive), FN (*англ.* False Negative) (Зеленцов, Мочалов, 2024б; Хасты и др., 2020; Zhao et al., 2025), которые в рамках проводимых исследований могут быть интерпретированы следующим образом: TP — количество фактически положительных образцов (пикселей), которые верно идентифицированы как положительные, характерные для усыхающего леса; TN — количество фактически отрицательных образцов, которые верно идентифицированы как отрицательные, характерные для здорового леса; FP — количество фактически отрицательных образцов, которые ошибочно идентифициро-

ваны как положительные; FN — количество фактически положительных образцов, которые ошибочно идентифицированы как отрицательные.

Более полную характеристику качества обработки материалов съёмки можно получить с помощью обобщённых показателей, формируемых на основе базовых. Формулы для расчёта обобщённых показателей приведены в *табл. 2* (Maxwell et al., 2021a).

Таблица 2. Обобщённые показатели качества

Показатель	Формула для расчёта
Обобщённый показатель качества	$E = \frac{TP + TN}{TP + TN + FP + FN}$
Полнота	$REC = \frac{TP}{TP + FN}$
Точность	$PRE = \frac{TP}{TP + FP}$
Мера $F_{\beta}$ , комплексный показатель	$F_{\beta} = \frac{(1 + \beta^2) \cdot PRE \cdot REC}{\beta^2 \cdot PRE + REC}$ , $\begin{cases} \beta < 1 \rightarrow PRE \\ \beta > 1 \rightarrow REC \end{cases}$

Примечание: OA (*англ.* Overall Accuracy) — обобщённый показатель качества, REC (*англ.* Recall) — полнота, PRE (*англ.* Precision) — точность;  $\beta$  — любое действительное число.

Структурная схема предлагаемой методики приведена на *рис. 1*.

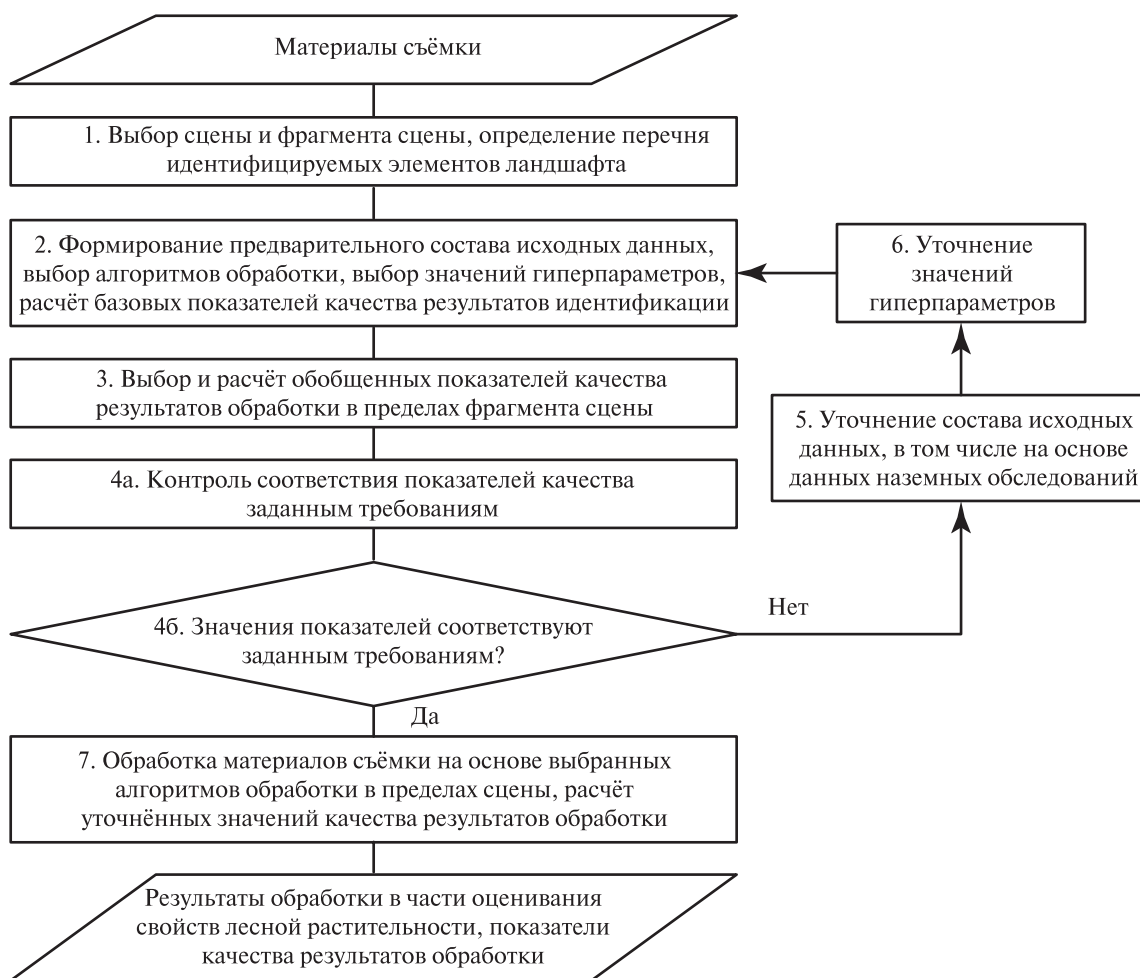


Рис. 1. Структурная схема методики оценивания качества идентификации состояния лесной растительности

Одна из основных особенностей методики заключается в том, что при её использовании вначале, на основе данных предварительного оценивания качества результатов обработки материалов космической съёмки в рамках фрагмента сцены, должен быть уточнён состав исходных данных, применяемых в дальнейшей обработке, а также определены основные параметры алгоритмов обработки материалов съёмки, включая пороговые значения индексов и значения гиперпараметров НР (*англ.* Hyperparameter), обеспечивающих распределение пикселей по множествам классов идентифицируемых элементов ландшафта. Финальная обработка данных и оценивание качества результатов обработки для всей анализируемой сцены осуществляются с использованием уже уточнённых параметров и на основе уточнённого состава исходных данных.

Для определённости ниже при описании методики анализируется лишь ограниченный состав алгоритмов обработки мультиспектральных данных, а именно алгоритмы, предусматривающие расчёт вегетационных индексов (Schowengerdt, 2007) NDVI (*англ.* Normalized Difference Vegetation Index), EVI (*англ.* Enhanced Vegetation Index) и ARVI (*англ.* Atmospherically Resistant Vegetation Index).

Цифрами на схеме (см. *рис. 1*) обозначены основные этапы реализации методики.

На этапе 1 оператором формируется фрагмент сцены в пространстве  $(x, y)$ , включающий элементарные площадки (пиксели) с результатами измерений спектральных отражательных характеристик, составляющих множество  $R = \{r_i(x, y, \lambda_l)\}$ ,  $r_i = [r_i^1, r_i^2, \dots, r_i^l, \dots, r_i^L]^T$ , где  $l$  — номер спектрального канала с длиной волны  $\lambda_l$ ;  $L$  — количество спектральных каналов;  $I$  — количество пикселей. На основе анализа содержания решаемой тематической задачи обработки определяется состав кластеров, соответствующий перечню идентифицируемых элементов ландшафта.

На этапе 2 на основе применения математического аппарата нечёткой кластеризации FCM (*англ.* fuzzy means) (Демидова, Коняева, 2008; Bezdek, 1981; Mochalov et al., 2020) в пределах фрагмента сцены формируется предварительный набор исходных данных распределением множества  $R$  по кластерам:

$$g : R \times K \rightarrow R_1, R_2, \dots, R_k, U_{i,k},$$

где  $K$  — множество кластеров;  $U_{i,k}$  — матрица, отражающая степень принадлежности  $i$ -го пикселя  $k$ -му кластеру,  $k = 1, 2, \dots, N$ ;  $N$  — количество кластеров;  $g$  — оператор нечёткой кластеризации.

Далее в рамках 2-го этапа выбираются алгоритмы обработки (Комаров и др., 2021; Рашка, 2017; Саворский и др., 2016; Brovkina et al., 2017); определяются предварительные пороговые значения выбранных индексов и гиперпараметров; рассчитываются предварительные значения базовых показателей качества результатов обработки материалов съёмки.

На этапе 3 с учётом анализа содержания тематической задачи обработки осуществляется выбор и расчёт обобщённых показателей качества результатов обработки материалов съёмки для определённого на 1-м этапе фрагмента сцены.

На этапе 4 проводится анализ достигаемых значений показателей качества для фрагмента сцены и контроль соответствия показателей качества заданным требованиям на обработку. В случае невыполнения требований осуществляется выполнение этапов 5 и 6.

На этапе 5 уточняется состав исходных данных. Уточнение основано на анализе степени принадлежности пикселей  $U_{i,k}$  рассматриваемому кластеру с учётом критерия Стьюдента (Иванов, 1990; Мочалов и др., 2023; Hollander, Wolfe, 1999). Окончательное решение по уточнению состава исходных данных принимается на основе проведения, при необходимости, дополнительных наземных обследований. Порядок проведения наземных обследований приведён в публикациях (Зеленцов, Мочалов, 2024а; Зеленцов и др., 2025).

По итогам наземных обследований существуют следующие варианты работы с исходными данными: подтвердить принадлежность элементарной площадки к рассматриваемому кластеру; присвоить принадлежность элементарной площадки к другому кластеру; исключить элементарную площадку из состава исходных данных, как не относящуюся ни к одному из

рассматриваемых кластеров. В представленной методике решение принимается только на основе данных наземных обследований. Вместе с тем для получения дополнительных объективных данных о состоянии лесной растительности могут быть предусмотрены варианты проведения оперативной съёмки, например, с помощью беспилотных летательных аппаратов.

На этапе 6 с учётом обновлённого состава исходных данных уточняются пороговые значения гиперпараметра и пороговые значения анализируемых индексов. Можно отметить, что для сбалансированных исходных данных варьирование значением гиперпараметра обеспечивает возможность расчёта более строгих оценок качества алгоритмов обработки материалов съёмки. Например, кривые или кривые ошибок ROC (*англ.* Receiver Operating Characteristic) (Рашка, 2017) строятся на основе варьирования значений гиперпараметра. В настоящей работе эти вопросы не рассматриваются.

На этапе 7 осуществляется обработка материалов съёмки в пределах сцены и определяются финальные значения показателей качества идентификации элементов лесной растительности для всей анализируемой сцены.

### Пример расчёта и анализа показателей качества обработки

Порядок расчётов по предложенной методике рассмотрим на примере задачи идентификации состояния елового леса и последующей его классификации по классам пожарной опасности на территории Государственного природного заказника «Озеро Щучье», Ленинградская область.

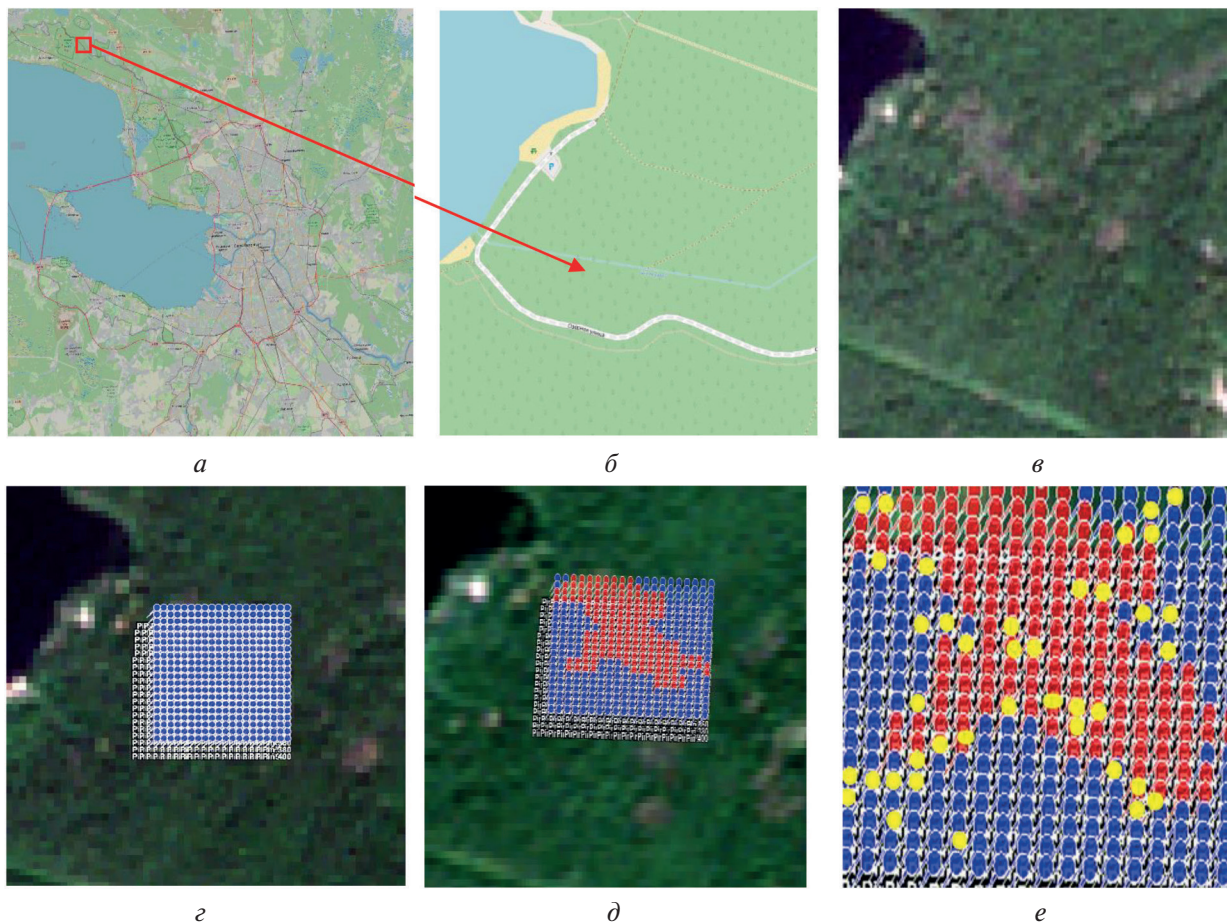


Рис. 2. Территория, выбранная для оценивания степени пожарной опасности лесных насаждений на северо-западе от Санкт-Петербурга (а); часть территории заказника «Озеро Щучье» (б); сцена по данным Sentinel-2 (в); 400 пикселей для формирования исходных данных в пределах фрагмента сцены (г); результаты кластеризации исходных данных (д); точки для проведения наземных обследований (е)

Использовались материалы мультиспектральной космической съёмки, полученные с помощью космического аппарата Sentinel-2 17 июня 2023 г., после проведения атмосферной коррекции в один из наиболее информативных периодов для оценивания свойств лесной растительности.

Этап 1. Выбор сцены, фрагмента сцены, а также перечня идентифицируемых элементов ландшафта осуществлён в соответствии с требованием репрезентативности и обусловлен результатами предварительных наземных обследований (Зеленцов и др., 2025), а также материалами публикации (Киселев и др., 2025). На *рис. 2а, б, в* (см. с. 178) приведено пространственное расположение сцены, на *рис. 2г* 400 пикселей синего цвета обозначают фрагмент сцены.

Этап 2. Результаты формирования предварительного набора исходных данных на основе нечёткой кластеризации приведены на *рис. 2д*. По результатам кластеризации 400 пикселей из общего объёма исходных данных выделены два подмножества пикселей: 235 пикселей для условно здорового леса, окрашенных синим цветом, и 165 — для условно усыхающего леса, окрашенных красным цветом (см. *рис. 2д*). Решение задачи разделения на классы методом визуального анализа, как можно увидеть на *рис. 2в*, крайне затруднительно. Для выбора гиперпараметров и расчёта базовых показателей качества результатов идентификации построены гистограммы, представленные на *рис. 3*.

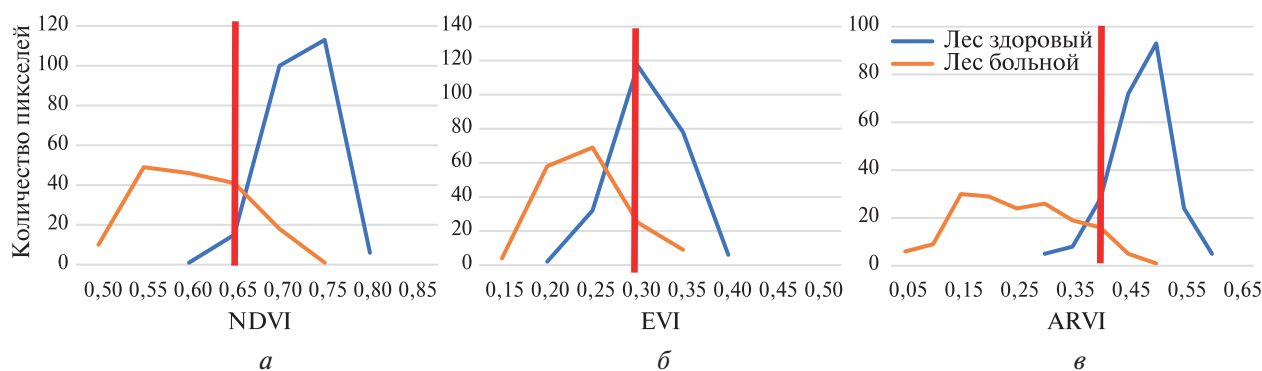


Рис. 3. Гистограммы вегетационных индексов

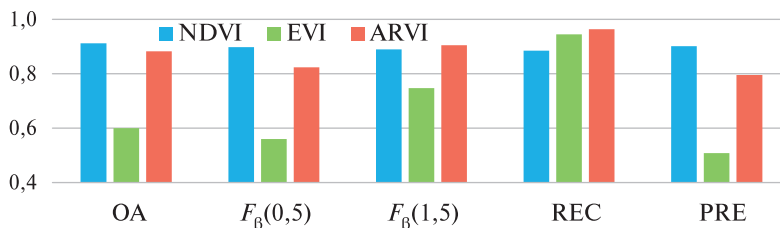
На *рис. 3* по осям абсцисс отложены значения индексов, а по осям ординат — количество пикселей, для которых характерны соответствующие значения индексов. Красные прямые обозначают гиперпараметры. Значения гиперпараметров при первичном расчёте и настройке параметров алгоритмов обработки, как правило, определяются на основе анализа гистограмм экспертным способом в зависимости от типа тематической задачи обработки и используемых показателей качества. При необходимости может осуществляться тонкая настройка гиперпараметров (Рашка, 2017). Значения гиперпараметров показаны на *рис. 3* и в *табл. 3*.

Результаты расчёта базовых показателей качества приведены в *табл. 3* (строки 2–5).

Таблица 3. Показатели качества результатов идентификации для различных алгоритмов

№	Показатели	Алгоритм расчёта		
		NDVI	EVI	ARVI
1	HP	0,65	0,30	0,40
2	TP	146	156	159
3	TN	219	84	194
4	FP	16	151	41
5	FN	19	9	6

Этап 3. Итоги расчёта обобщённых показателей качества результатов обработки в пределах фрагмента сцены представлены на *рис. 4*.

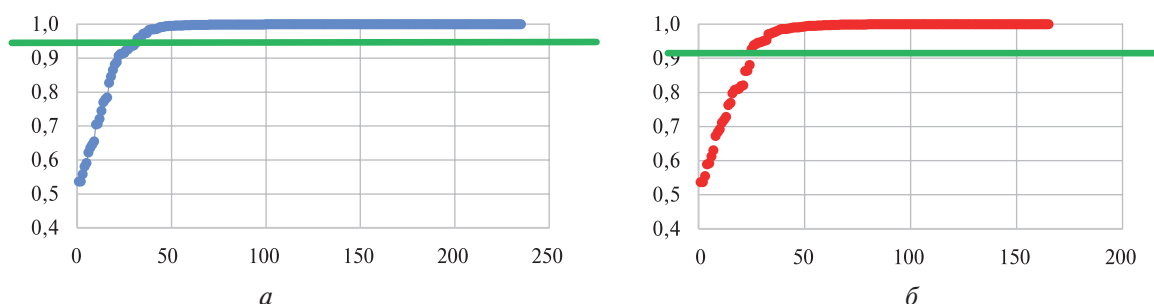


*Рис. 4.* Обобщённые показатели качества результатов обработки

Из данных, представленных на *рис. 4*, видно, что обобщённые показатели качества для различных алгоритмов существенно различаются, при этом наилучшие значения обобщённых показателей качества характерны для алгоритмов, предусматривающих расчёт индексов NDVI и ARVI. Таким образом, применение методики даёт возможность также обосновать выбор алгоритмов для решения конкретной задачи.

Этап 4. По итогам контроля соответствия показателей качества результатов обработки возможны два варианта выполнения дальнейших работ. В случае невыполнения требований реализуются этапы 5 и 6. В случае выполнения — этап 7.

Этап 5. Проводится уточнение состава исходных данных. На *рис. 5* приведены диаграммы степени принадлежности для всех элементарных площадок из предварительного набора исходных данных (см. *рис. 2д*) к двум рассматриваемым классам лесной растительности.



*Рис. 5.* Степень принадлежности 235 пикселей кластеру условно здорового леса (*а*) и 165 пикселей кластеру условно усыхающего леса (*б*)

На *рис. 5* по осям абсцисс отложены условные номера пикселей из множества, приведённого на *рис. 2г*, по оси ординат — значения их степени принадлежности кластерам условно здорового и усыхающего леса. Зелёные прямые обозначают пороговые значения, отделяющие пиксели, требующие внимания при уточнении состава исходных данных.

Расчёт порогового значения  $X$  осуществляется с использованием критерия Стьюдента (Иванов, 1990; Mochalov et al., 2020):

$$X = \bar{x} - 2t_{\infty} \frac{S}{\sqrt{\nu}},$$

где  $\bar{x}$  — среднее арифметическое степени принадлежности рассматриваемому кластеру;  $S$  — среднееквадратическое отклонение степени принадлежности;  $\nu$  — количество пикселей в кластере (число степеней свободы в таблице критических значений критерия Стьюдента);  $t_{\infty}$  — критическое значение двустороннего  $t$ -критерия Стьюдента при уровне значимости  $\alpha = 0,1$ , означающем 90%-й доверительный уровень, для заданного числа степеней свободы (Иванов, 1990).

Выбор доверительного уровня определяется содержанием задачи, заключающейся в обоснованном определении количества пикселей элементарных площадок, на которых необходимо выполнить наземные обследования.

В *табл. 4* приведены параметры и результаты расчёта пороговых значений степени принадлежности, представленных на *рис. 5*.

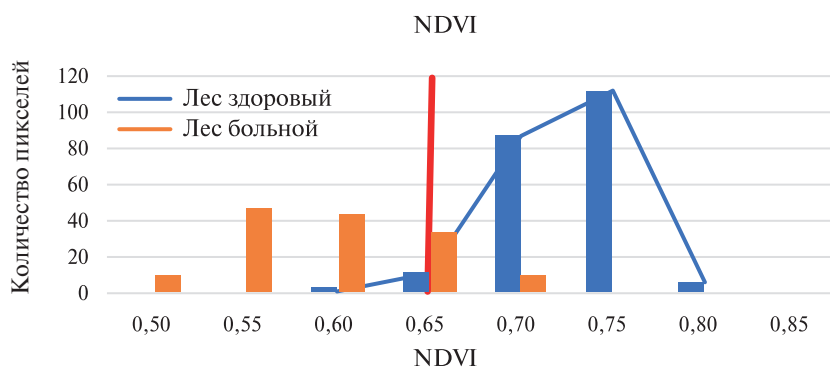
*Таблица 4.* Результаты расчёта порогового значения степени принадлежности

Кластеры	$v$	$\bar{x}$	$S$	$t_{\infty}$	$X$	$n$
Условно здоровый лес	235	0,969	0,090	1,653	0,95	21
Условно усыхающий лес	165	0,954	0,106	1,658	0,92	18

Количество пикселей, требующих контроля, обозначено в *табл. 4* символом  $n$ .

В приведённом примере количество пикселей, заслуживающих внимания для проведения наземных обследований, равно 39. Расположение элементарных площадок для проведения наземных обследований показано на *рис. 2е* в виде жёлтых точек. Обоснованное сокращение объёма наземных обследований почти на 90 % и, соответственно, снижение временных затрат на подготовку достоверных исходных данных являются одним из достоинств предлагаемой методики.

Этап 6. На основе уточнённого на этапе 5 состава исходных данных построена гистограмма для индекса NDVI (*рис. 6*), обеспечивающая возможность уточнения значения параметров обработки.



*Рис. 6.* Гистограмма индекса NDVI на основе уточнённого состава исходных данных

По итогам выполнения этапов 2–6 методики на основе анализа данных, приведённых на *рис. 3а* и *б*, определены пороговые значения индекса NDVI, учитываемые при идентификации четырёх стадий усыхания ельника и находящиеся в интервале от 0,5 до 0,7 (*табл. 5*). В общем случае оценивание качества результатов идентификации рассматриваемых четырёх стадий усыхания лесной растительности и расчёт индекса для разделения этих стадий могут быть проведены с полным выполнением всех этапов предлагаемой методики. В данном примере демонстрируется возможность комбинированного подхода к классификации состояния лесной растительности, при котором для финального разделения усыхающего ельника на классы применён достаточно широко используемый приём (Schowengerdt, 2007), основанный на расчёте пороговых значений NDVI. Общий принцип определения пороговых значений при этом заключается в нахождении, на первом шаге, границ диапазона изменения индекса NDVI, которые являются строго индивидуальными и зависят от конкретных условий произрастания лесной растительности, характеристик съёмочной аппаратуры и многих других факторов, и в разделении, на втором шаге, этого диапазона на интервалы, соответствующие анализируемым классам состояния растительности, с использованием эмпирических или статистических подходов. Расчёт пороговых значений индекса из найденного диапазона [0,5; 0,7] в рассматриваемом иллюстративном примере выполнен на основе эмпирического подхода

посредством равномерного распределения индекса из этого диапазона по интервалам, соответствующим стадиям усыхания ельника.

Таблица 5. Пороговые значения индекса NDVI

Класс пожарной опасности	Состояние ельника	Пороговые значения индекса
I — очень высокий	Сухостой в стадии вымирания растительности	$0,50 < NDVI \leq 0,55$
II — высокий	Сухостой высокой стадии отмирания растительности	$0,55 < NDVI \leq 0,60$
III — средний	Сухостой средней стадии отмирания растительности	$0,60 < NDVI \leq 0,65$
IV — слабый	Сухостой начальной стадии	$0,65 < NDVI \leq 0,70$

Уточнённые показатели для анализируемой сцены приведены в табл. 6.

Таблица 6. Показатели качества идентификации

Показатели качества	TP	FP	TN	FN	REC	PRE	OA	$F_{\beta(0,5)}$	$F_{\beta(1,5)}$
На предварительном наборе исходных данных	146	16	219	19	0,88	0,90	0,91	0,90	0,89
На уточнённом наборе исходных данных	134	12	205	10	0,93	0,92	0,94	0,92	0,92

Основной вывод по табл. 6 состоит в подтверждении обоснованности сформированного набора исходных данных, поскольку обобщённые показатели качества для всей сцены не ухудшились по сравнению с их значениями на предварительном этапе расчётов для фрагмента сцены. Более того, для данного примера наблюдается улучшение обобщённых показателей качества результатов обработки до 5 %.

Этап 7. Результаты обработки материалов съёмки в пределах сцены приведены на рис. 7, слева в пределах сцены представлен фрагмент карты с результатами идентификации четырёх классов лесной растительности, отражающих степени её усыхания. В рамках прямоугольника справа показано соответствующее пространственное размещение исходных данных, сформированных с помощью предлагаемой методики в пределах фрагмента сцены, обозначенного красным контуром.

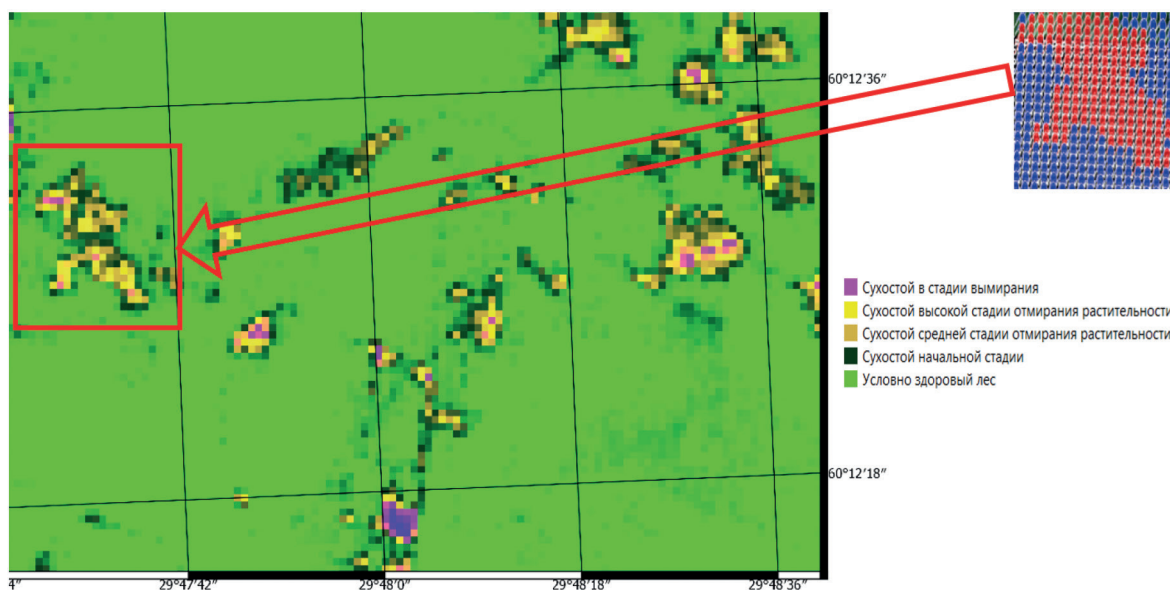


Рис. 7. Результаты идентификации состояния лесных насаждений в пределах сцены

## Заключение

Наличие количественной оценки качества получаемых результатов обработки материалов мультиспектральной космической съёмки является важным фактором, определяющим уверенность пользователя в высокой достоверности проведённой классификации состояния лесной растительности, оценке степени пожарной опасности лесной растительности и правильности выбора того или иного решения по управлению состоянием лесов на конкретном участке местности.

Описанная методика позволяет реализовать совместное решение как задач оценивания качества результатов идентификации состояния лесной растительности, так и задач настройки параметров алгоритмов обработки материалов мультиспектральной космической съёмки. Привлечение математического аппарата нечёткой кластеризации позволяет на этапе подготовки и уточнения исходных данных для обработки материалов съёмки обоснованно выбирать объём обучающей и контрольной выборки и при этом выделять нетипичные объекты (с наименьшей степенью принадлежности к каждому из рассматриваемых кластеров). Тем самым уточняется и существенно снижается объём необходимых дополнительных наземных обследований. В рассмотренном примере этот объём составил 10 % от размера анализируемого фрагмента сцены.

Кроме того, как показано в приведённом примере расчёта, на базе предложенного подхода может быть проведён сравнительный анализ качества результатов работы различных алгоритмов применительно к решению конкретной тематической задачи на конкретном участке территории. В частности, применение методики к решению задачи категорирования лесной растительности по классам пожарной опасности с использованием алгоритмов расчёта вегетационных индексов позволило не только оценить показатели качества обработки, но и провести сравнительный анализ результатов работы алгоритмов и сделать обоснованный вывод о целесообразности применения алгоритмов расчёта индексов NDVI и ARVI для этой задачи.

При этом необходимо иметь в виду, что, хотя предлагаемая методика носит достаточно общий характер, получаемые выводы об эффективности работы алгоритмов, о пороговых значениях параметров обработки в каждой конкретной задаче, не являются универсальными. Эти выводы могут меняться для разных участков местности, временных интервалов исследования, водно-почвенных условий и др. По мнению авторов, методику наиболее целесообразно применять при решении вопросов обработки данных, мониторинга и управления применительно к относительно небольшим участкам территории, в том числе при решении высокоответственных или требующих больших затрат задач, поскольку она позволяет детально учесть локальные, присущие только анализируемым участкам леса, особенности и конкретные спектрально-отражательные характеристики исследуемых элементов ландшафта в совершенно определённые отрезки времени.

Следует также отметить, что в настоящее время существенно возрастают потребности в решении задач обоснованного выбора исходных данных для обучающей и контрольной выборки в связи с интенсивным развитием моделей обработки на основе искусственных нейронных сетей и методов машинного обучения. Поэтому актуальным становится дальнейшее развитие исследований по затронутому в статье направлению.

Исследование выполнено за счёт гранта Российского научного фонда № 25-29-00684, <https://rscf.ru/project/25-29-00684/>.

## Литература

1. Горелик А. Л., Скрипкин В. А. Методы распознавания: учеб. пособие для вузов. 4-е изд., испр. М.: Высш. шк., 2004. 261 с.
2. Демидова Л. А., Коняева Е. И. Кластеризация объектов с использованием FCM-алгоритма на основе нечетких множеств второго типа и генетического алгоритма // Вестн. РГРТУ. 2008. № 4. Вып. 26. 9 с.

3. Зеленцов В. А., Мочалов В. Ф. (2024а) Пример оценивания качества результатов автоматизированной идентификации лесной растительности // 9-я Всероссийская научно-техн. конф. «Леса России: политика, промышленность, наука, образование»: материалы конф. СПб.: СПбГЛТУ, 2024. С. 46–49.
4. Зеленцов В. А., Мочалов В. Ф. (2024б) Показатели и алгоритмы оценивания качества результатов идентификации состояния прилегающих территорий // Интеллектуальные технологии на транспорте. 2024. № 3(39). С. 73–82. DOI: 10.20295/2413-2527-2024-339-73-82.
5. Зеленцов В. А., Мочалов В. Ф., Мухаметов Д. И. Методика наземных обследований лесной растительности для организации автоматизированной обработки материалов мультиспектральной космической съёмки // 10-я Всероссийская научно-техн. конф. «Леса России: политика, промышленность, наука, образование»: материалы конф. СПб.: СПбГЛТУ, 2025. С. 78–80.
6. Иванов В. С. Основы математической статистики: учеб. пособие для вузов. М.: Физкультура и спорт, 1990. 176 с.
7. Киселев А. В., Тронин А. А., Крицук С. Г. и др. Многолетняя динамика состояния ельников заказника «Озеро Щучье» по спутниковым данным // Современные проблемы дистанционного зондирования Земли из космоса. 2025. Т. 22. № 2. С. 173–185. DOI: 10.21046/2070-7401-2025-22-2-173-185.
8. Комаров А. А., Кирсанов А. Д., Малашин С. Н. Сравнительная характеристика различных вегетационных индексов при оценке состояния растительного покрова кормовых трав // Изв. Санкт-Петербургского гос. аграрного ун-та. 2021. № 2(63). С. 18–29. DOI: 10.24412/2078-1318-2021-2-18-29.
9. Малышев В. Б., Фомин Б. Н. Децентрализованная база данных спектральных характеристик и параметров объектов земной поверхности // Современные проблемы дистанционного зондирования Земли из космоса. 2019. Т. 16. № 2. С. 53–65. DOI: 10.21046/2070-7401-2019-16-2-53-65.
10. Мочалов В. Ф., Григорьева О. В., Спесивцева К. А. Оценка характеристик леса по данным отечественной космической съёмки // 8-я Всероссийская научно-техн. конф. «Леса России: политика, промышленность, наука, образование»: материалы конф. СПб.: СПбГЛТУ, 2023. С. 79–82.
11. Мьльникова Т. А., Боровлёв А. Ю., Елсаков В. В., Шанов В. М. Формирование базы данных пространственного распределения древесных пород на тестовом полигоне «Ляльский» (Республика Коми) по материалам БПЛА-съёмок // Современные проблемы дистанционного зондирования Земли из космоса. 2025. Т. 22. № 1. С. 116–130. DOI: 10.21046/2070-7401-2025-22-1-116-130.
12. Приказ Рослесхоза «Об утверждении классификации природной пожарной опасности лесов и классификации пожарной опасности в лесах в зависимости от условий погоды» № 287 от 05.07.2011.
13. Рашка С. Python и машинное обучение: пер. с англ. А. В. Логунова. М.: ДМК Пресс, 2017. 418 с.
14. Саворский В. П., Кашицкий А. В., Константинова А. М. и др. Возможности анализа гиперспектральных индексов в информационных системах дистанционного мониторинга семейства «Созвездие-Вега» // Современные проблемы дистанционного зондирования Земли из космоса. 2016. Т. 13. № 3. С. 28–45. DOI: 10.21046/2070-7401-2016-13-3-28-45.
15. Хасты Т., Тибришани Р., Фридман Д. Основы статистического обучения. Интеллектуальный анализ данных, логический вывод и прогнозирование. 2-е изд. М.: Изд-во «Вильямс», 2020. 768 с.
16. Bezdek J. C. Pattern recognition with fuzzy objective function algorithms. N. Y.: Springer, 1981. 272 p. DOI: 10.1007/978-1-4757-0450-1.
17. Bradley A. P. The use of the area under the ROC curve in the evaluation of machine learning algorithms // Pattern Recognition. 1997. V. 30. Iss. 7. P. 1145–1159. DOI: 10.1016/S0031-3203(96)00142-2.
18. Brovkina O., Cienciala E., Zemek F. et al. Composite indicator for monitoring of Norway spruce stand decline // European J. Remote Sensing. 2017. V. 50. Iss. 1. P. 550–563. DOI: 10.1080/22797254.2017.1372697.
19. Congalton R. G., Green K. Assessing the accuracy of remotely sensed data. Principles and practices. 2<sup>nd</sup> ed. CRC Press, 2009. 209 p.
20. Ferreira M. P., Wagner F. H., Aragão L. E. et al. Tree species classification in tropical forests using visible to shortwave infrared WorldView-3 images and texture analysis // ISPRS J. Photogrammetry and Remote Sensing. 2019. V. 149. P. 119–131. DOI: 10.1016/j.isprsjprs.2019.01.019.
21. Hollander M., Wolfe D. Nonparametric statistical methods. 2<sup>nd</sup> ed. N. Y.: Wiley, 1999. 787 p.
22. Liu Y., Ren C., Liang J. et al. A robust index based on phenological features to extract sugarcane from multisource remote sensing data // Remote Sensing. 2023. V. 15. Iss. 24. Article 5783. DOI: 10.3390/rs15245783.
23. Liu P., Ren C., Wang Z. et al. Evaluating the potential of Sentinel-2 time series imagery and machine learning for tree species classification in a mountainous forest // Remote Sensing. 2024. V. 16. Iss. 2. Article 293. DOI: 10.3390/rs16020293.
24. Maxwell A. E., Warner T. A., Guillén L. A. (2021a) Accuracy assessment in convolutional neural network-based deep learning remote sensing studies — Part 1: Literature review // Remote Sensing. 2021. V. 13. Iss. 13. Article 2450. DOI: 10.3390/rs13132450.

25. Maxwell A. E., Warner T. A., Guillén L. A. (2021b) Accuracy assessment in convolutional neural network-based deep learning remote sensing studies — Part 2: Recommendations and best practices // *Remote Sensing*. 2021. V. 13. Iss. 13. Article 2591. DOI: 10.3390/rs13132591.
26. Mochalov V., Grigorieva O., Zhukov D. et al. Remote sensing image processing based on modified fuzzy algorithm // *Artificial Intelligence and Bioinspired Computational Methods: Proc. 9<sup>th</sup> Computer Science On-line Conf.* 2020. Ser.: *Advances in Intelligent Systems and Computing*. 2020. V. 1225. P. 563–572. DOI: 10.1007/978-3-030-51971-1\_46.
27. Schowengerdt R. *Remote Sensing, models and methods for image processing*. 3<sup>rd</sup> ed. L.: Academic Press, 2007. 560 p.
28. Wang B., Yao Y. Mountain vegetation classification method based on multichannel semantic segmentation model // *Remote Sensing*. 2024. V. 16. Iss. 2. Article 256. DOI: 10.3390/rs16020256.
29. Zhao B., Mi Y., Sun R., Wu C. Semantic segmentation method of residential areas in remote sensing images based on cross-attention mechanism // *Remote Sensing*. 2025. V. 17. Iss. 18. Article 3253. DOI: 10.3390/rs17183253.

## Assessing the quality of forest vegetation condition identification using Shchuchye Ozero Nature Reserve as an example

V. A. Zelentsov, V. F. Mochalov

*Saint Petersburg Federal Research Center RAS, Saint Petersburg 199178, Russia*  
*E-mail: vicavia@yandex.ru*

The article examines aspects of joint and simultaneous processing of multispectral satellite imagery and assessment of processing results quality for the task of forest vegetation condition identification. The role of a well-founded composition of initial data on forest vegetation classification for these purposes is emphasized. A methodology is presented that includes preliminary and final assessments of imagery processing quality. The key element of the method is a new approach to determining at a preliminary stage a set of elementary sites (pixels) of the analyzed scene fragment for their use as initial data when processing satellite images. The selection of elementary sites is based on the use of fuzzy clustering and quantitative analysis of the degree of each pixel membership in one of the classes of forest landscape identifiable elements. The degree of membership is used to determine the composition of elementary sites that require clarification, including through field surveys. Final data processing and processing quality assessment for the entire analyzed scene are performed using refined initial data. The analysis of processing quality using the proposed methodology is demonstrated by the example of classifying spruce forest condition in fire hazard classes in Lake Shchuchye Nature Reserve in Leningrad Region. Sentinel-2 multispectral data were used in this case, and algorithms for calculating various vegetation indices were applied. To justify the selection of the initial data, preliminary fuzzy clustering of a 400-pixel scene fragment was performed. Thirty-nine pixels requiring clarification were identified, and the entire scene was processed using these clarified values. It is shown that the application of this methodology improves the final processing quality indicators, significantly reduces the time spent on preparing the initial data, and also helps formulate recommendations for selecting processing algorithms that ensure the best identification quality. In the considered example, these are the algorithms for calculating the Normalized Difference Vegetation Index (NDVI) and the Atmospherically Resistant Vegetation Index (ARVI). The most effective area of application of the methodology is solving multispectral data processing problems for relatively small areas of territory, since it allows for the most complete consideration of local features and specific spectral-reflective characteristics of landscape elements.

**Keywords:** multispectral satellite imagery, processing quality indicators, initial data, processing algorithms, forest vegetation condition identification, fuzzy clustering, field surveys, fire hazard classes

Accepted: 18.12.2025

DOI: 10.21046/2070-7401-2026-23-2-173-187

## References

1. Gorelik A. L., Skripkin V. A., *Metody raspoznavaniya: uchebnoe posobie dlya vuzov* (Recognition methods: A textbook for universities), 4<sup>th</sup> ed., corr. Moscow: Vysshaya shkola, 2004, 261 p. (in Russian).
2. Demidova L. A., Konjaeva E. I., Clustering of objects using the FCM algorithm based on fuzzy sets of the second type and the genetic algorithm, *Vestnik of Ryazan State Radio Engineering University*, 2008, No. 4, Iss. 26, 9 p. (in Russian).
3. Zelentsov V. A., Mochalov V. F. (2024a), An example of evaluating the quality of the results of automated identification of forest vegetation, *9-ya Vserossiiskaya nauchno-tehnicheskaya konferentsiya "Lesa Rossii: politika, promyshlennost', nauka, obrazovanie": materialy konferentsii* (Proc. 9<sup>th</sup> All-Russian Scientific and Technical Conf. "Forests of Russia: Politics, Industry, Science, Education"), Saint Petersburg: SPbGLTU, 2024, pp. 46–49 (in Russian).
4. Zelentsov V. A., Mochalov V. F. (2024b), Indicators and algorithms for assessing the quality of the results of identification of the condition of adjacent territories, *Intellektual'nye tehnologii na transporte*, 2024, No. 3(39), pp. 73–82 (in Russian), DOI: 10.20295/2413-2527-2024-339-73-82.
5. Zelentsov V. A., Mochalov V. F., Mukhametov D. I., Methodology of ground surveys of forest vegetation for the organization of automated processing of multispectral satellite survey materials, *10-ya Vserossiiskaya nauchno-tehnicheskaya konferentsiya "Lesa Rossii: politika, promyshlennost', nauka, obrazovanie": materialy konferentsii* (Proc. 10<sup>th</sup> All-Russian Scientific and Technical Conf. "Forests of Russia: Politics, Industry, Science, Education"), Saint Petersburg: SPbGLTU, 2025, pp. 78–80 (in Russian).
6. Ivanov V. S., *Osnovy matematicheskoi statistiki: uchebnoe posobie dlya vuzov* (Fundamentals of mathematical statistics: A textbook for universities), Moscow: Fizkul'tura i sport, 1990, 176 p. (in Russian).
7. Kiselev A. V., Tronin A. A., Kritsuk S. G. et al., Long-term dynamics of spruce forests of Lake Shchuchye Reserve based on satellite data, *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2025, V. 22, No. 2, pp. 173–185 (in Russian), DOI: 10.21046/2070-7401-2025-22-2-173-185.
8. Komarov A. A., Kirsanov A. D., Malashin S. N., Comparative characteristics of various vegetation indices (VI) when the vegetation cover state of forage grasses assessing, *Izvestiya of Saint-Petersburg State Agrarian University*, 2021, No. 2(63), pp. 18–29 (in Russian), DOI: 10.24412/2078-1318-2021-2-18-29.
9. Malyshev B. V., Fomin B. N., Decentralized database of spectral characteristics and parameters of Earth surface features, *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2019, V. 16, No. 2, pp. 53–65 (in Russian), DOI: 10.21046/2070-7401-2019-16-2-53-65.
10. Mochalov V. F., Grigor'eva O. V., Spesivtseva K. A., Assessment of forest characteristics based on national space survey data, *8-ya Vserossiiskaya nauchno-tehnicheskaya konferentsiya "Lesa Rossii: politika, promyshlennost', nauka, obrazovanie": materialy konferentsii* (Proc. 8<sup>th</sup> All-Russian Scientific and Technical Conference "Forests of Russia: Politics, Industry, Science, Education"), Saint Petersburg: SPbGLTU, 2023, pp. 79–82 (in Russian).
11. Mylnikova T. A., Borovlev A. Yu., Elsakov V. V., Shchanov V. M., Formation of a spatial database of tree species distribution at Lyalsky test site (Komi Republic) based on UAV data, *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2025, V. 22, No. 1, pp. 116–130 (in Russian), DOI: 10.21046/2070-7401-2025-22-1-116-130.
12. *Prikaz Rosleskhoza "Ob utverzhdenii klassifikatsii prirodnoi pozharnoi opasnosti lesov i klassifikatsii pozharnoi opasnosti v lesakh v zavisimosti ot uslovii pogody"* (Order of the Federal Forestry Agency "On approval of the classification of natural fire danger of forests and classification of fire danger in forests depending on weather conditions"), 05.07.2011, No. 287 (in Russian).
13. Raschka S., *Python Machine Learning*, 1<sup>st</sup> ed. Packt Publishing, 2015, 454 p.
14. Savorskiy V. P., Kashnickiy A. V., Konstantinova A. M. et al., Capabilities of hyperspectral indices analysis of the Vega-Constellation remote monitoring information systems, *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2016, V. 13, No. 3, pp. 28–45 (in Russian), DOI: 10.21046/2070-7401-2016-13-3-28-45.
15. Hastie T., Tibshirani R., Friedman J., *The elements of statistical learning. Data mining, inference, and prediction*, 2<sup>nd</sup> ed., New York: Springer, 2009, 767 p.
16. Bezdek J. C., *Pattern recognition with fuzzy objective function algorithms*, New York: Springer, 1981, 272 p., DOI: 10.1007/978-1-4757-0450-1.
17. Bradley A. P., The use of the area under the ROC curve in the evaluation of machine learning algorithms, *Pattern Recognition*, 1997, V. 30, Iss. 7, pp. 1145–1159, DOI: 10.1016/S0031-3203(96)00142-2.
18. Brovkina O., Cienciala E., Zemek F. et al., Composite indicator for monitoring of Norway spruce stand decline, *European J. Remote Sensing*, 2017, V. 50, Iss. 1, pp. 550–563, DOI: 10.1080/22797254.2017.1372697.
19. Congalton R. G., Green K., *Assesing the accuracy of remotely sensed data. Principles and practices*, 2<sup>nd</sup> ed., CRC Press, 2009, 209 p.

20. Ferreira M. P., Wagner F. H., Aragão L. E. et al., Tree species classification in tropical forests using visible to shortwave infrared WorldView-3 images and texture analysis, *ISPRS J. Photogrammetry and Remote Sensing*, 2019, V. 149, pp. 119–131, DOI: 10.1016/j.isprsjprs.2019.01.019.
21. Hollander M., Wolfe D., *Nonparametric statistical methods*, 2<sup>nd</sup> ed., New York: Wiley, 1999, 787 p.
22. Liu Y., Ren C., Liang J. et al., A robust index based on phenological features to extract sugarcane from multisource remote sensing data, *Remote Sensing*, 2023, V. 15, Iss. 24, Article 5783, DOI: 10.3390/rs15245783.
23. Liu P., Ren C., Wang Z. et al., Evaluating the potential of Sentinel-2 time series imagery and machine learning for tree species classification in a mountainous forest, *Remote Sensing*, 2024, V. 16, Iss. 2, Article 293, DOI: 10.3390/rs16020293.
24. Maxwell A. E., Warner T. A., Guillén L. A. (2021a), Accuracy assessment in convolutional neural network-based deep learning remote sensing studies — Part 1: Literature review, *Remote Sensing*, 2021, V. 13, Iss. 13, Article 2450, DOI: 10.3390/rs13132450.
25. Maxwell A. E., Warner T. A., Guillén L. A. (2021b), Accuracy assessment in convolutional neural network-based deep learning remote sensing studies — Part 2: Recommendations and best practices, *Remote Sensing*, 2021, V. 13, Iss. 13, Article 2591, DOI: 10.3390/rs13132591.
26. Mochalov V., Grigorieva O., Zhukov D. et al., Remote sensing image processing based on modified fuzzy algorithm, In: *Artificial Intelligence and Bioinspired Computational Methods: Proc. 9<sup>th</sup> Computer Science Online Conf. 2020*, Ser.: Advances in Intelligent Systems and Computing, 2020, V. 1225, pp. 563–572, DOI: 10.1007/978-3-030-51971-1\_46.
27. Schowengerdt R., *Remote Sensing, models and methods for image processing*, 3<sup>rd</sup> ed., London: Academic Press, 2007, 560 p.
28. Wang B., Yao Y., Mountain vegetation classification method based on multi-channel semantic segmentation model, *Remote Sensing*, 2024, V. 16, Iss. 2, Article 256, DOI: 10.3390/rs16020256.
29. Zhao B., Mi Y., Sun R., Wu C., Semantic segmentation method of residential areas in remote sensing images based on cross-attention mechanism, *Remote Sensing*, 2025, V. 17, Iss. 18, Article 3253, DOI: 10.3390/rs17183253.