

Прогнозирование послепожарного отпада древостоев на основе данных дистанционного зондирования и машинного обучения

Е. Г. Швецов

Красноярский научный центр СО РАН, Красноярск, 660036, Россия
E-mail: eugeneshvetsov11@yandex.ru

Значительное количество и площади пожаров растительности, ежегодно возникающих на территории Сибири, определяет необходимость разработки средств оценки и прогнозирования последствий пирогенного воздействия на леса. В работе рассматриваются возможности применения открытых каталогов данных и методов машинного обучения для прогнозирования участков послепожарного отпада древостоя, а также оценки значимости признаков, определяющих величину отпада. С применением набора тематических спутниковых продуктов и метода машинного обучения RF (*англ.* Random Forest, случайный лес) для территорий Красноярского края, Республик Хакасия и Тыва выполнена оценка значимости набора признаков при прогнозировании доли отпада древостоя после пирогенного воздействия, а также разработана соответствующая модель. В работе использованы признаки, описывающие как лесорастительные и топографические условия местности, так и увлажнённость растительных горючих материалов, а также характеристики горения и величина изменений спектральных свойств поверхности, вызванных воздействием пожаров. Общая точность модели составила 0,84, а величина показателя F_1 -score — 0,76. Наибольшую значимость при прогнозировании доли послепожарного отпада продемонстрировал разностный индекс dNBR (*англ.* Differenced Normalized Burn Ratio). В целом, признаки, характеризующие изменения отражательных свойств поверхности после пирогенного воздействия (спектральные индексы dNBR, dBAI (*англ.* Differenced Burn Area Index), а также отражательная способность в отдельных диапазонах спектра), определяли более 50 % вариабельности в оценках послепожарного отпада древостоя.

Ключевые слова: пожары растительности, дистанционные данные, машинное обучение, Random Forest, Сибирь

Одобрена к печати: 21.03.2025

DOI: 10.21046/2070-7401-2025-22-3-171-181

Введение

Лесные пожары — один из наиболее значимых факторов, влияющих на динамику растительности, углеродный цикл и биоразнообразие в лесах (Барталев и др., 2015; Russian..., 2020). При этом на Сибирь приходится до 70 % от общего числа пожаров, возникающих на территории России ежегодно, и до 90 % площади, пройденной огнём (Kharuk et al., 2021). Воздействие пожаров на леса определяется рядом факторов, в том числе видом и интенсивностью пожара, породным составом и возрастной структурой древостоев, почвой и другими условиями. Результатом воздействия пожаров, в зависимости от их характеристик и исходных лесорастительных условий, может являться как стимуляция естественного возобновления лесов, так и гибель древостоев. Данные о состоянии повреждённых пожарами насаждений, в том числе о площади погибших насаждений, необходимы как для оценки экологических и экономических последствий пожаров, так и для решения задач послепожарного восстановления лесов.

Спутниковый мониторинг даёт возможность оценивать степень пирогенной нарушенности и послепожарную динамику растительности на обширных территориях. В настоящее время существует ряд продуктов, отражающих изменения растительных покровов, среди которых можно отметить GFC (*англ.* Global Forest Change) (<https://glad.earthengine.app/view/global-forest-change>) (Hansen et al., 2013), GLCLUC (*англ.* Global Land Cover and Land Use Change) (<https://glad.umd.edu/dataset/GLCLUC2020>) (Potapov et al., 2020), сервис «Bera-Pro»

(<http://pro-vega.ru/maps/>) (Loupian et al., 2022), а также продукт MODIS (англ. Moderate Resolution Imaging Spectroradiometer) Land Cover Type (MCD12Q1, <https://lpdaac.usgs.gov/products/mcd12q1v061/>) (Friedl et al., 2010).

Вместе с тем подобные продукты, основанные на анализе временных рядов (Hansen et al., 2013), характеризуются значительным временным лагом между датой окончания пожара и появлением информации о гибели древостоя, вызванной пожаром. В то же время известно значительное число работ (Барталев и др., 2010; Delcourt et al., 2021), в которых участки, где имела место гибель древостоев, связываются с величиной спектральных индексов, чьи значения могут быть оценены вскоре после окончания пожара. В качестве таких индексов часто используются разностные индексы dNDVI (англ. Differenced Normalized Difference Vegetation Index, разностный нормализованный вегетационный индекс) или dNBR (англ. Differenced Normalized Burn Ratio, разностный нормализованный индекс гарей) (Escuin et al., 2008), рассчитывающиеся как разность индексов NDVI/NBR до и после пожара.

В последнее десятилетие, вместе с ростом вычислительных мощностей, различные методы машинного обучения находят всё большее применение в задачах обработки данных дистанционного зондирования (Florath, Keller, 2022; Zhong et al., 2021). В частности, использование метода случайного леса (англ. Random Forest — RF) для анализа спектральных индексов демонстрировало хорошие результаты при картировании степени пирогенного воздействия на растительность (Collins et al., 2018; Gibson et al., 2020).

Основная цель работы — оценка точности прогнозирования послепожарного отпада древостоев с помощью методов машинного обучения и данных, полученных до и непосредственно после пожара. Рассматривались следующие аспекты: 1) выявление признаков, оказывающих наиболее существенное влияние на точность прогнозирования; 2) разработка модели для прогнозирования участков послепожарной гибели древостоев.

Материалы и методы

Район исследования включал территории Красноярского края, Республик Хакасия и Тыва общей площадью около $2,5 \cdot 10^6$ км² (рис. 1, см. с. 173). Согласно использованной в работе карте растительности, доступной через сервис «Вега» (ИКИ РАН, Москва, <http://pro-vega.ru/maps/>) (Барталев и др., 2016), преобладающими лесообразующими породами являются светлохвойные породы, включая лиственничные (*Larix sibirica*) (55 % лесной площади района исследования) и сосновые (*Pinus sylvestris*) (11 %) древостои, а также темнохвойные древостои (*Pinus sibirica*, *Abies sibirica*) (20 %). На лиственничные породы (*Betula spp.*, *Populus tremula*) приходится около 14 % лесной площади.

Поведение пожара регулируется тремя основными факторами (погодой, топографией и горючими материалами). Состав и запас растительных горючих материалов в сочетании с особенностями рельефа и погодой влияют на длительность горения и интенсивность тепловыделения, что в конечном итоге определяет степень пирогенной нарушенности и последствия пожара. Для формирования обучающих и тестовых наборов данных использовались входные переменные, описывающие как лесорастительные и топографические условия местности, так и увлажнённость горючих материалов. Также применялись данные ДЗЗ, представляющие характеристики горения и величину изменений спектральных свойств поверхности, вызванных воздействием пожаров. Список переменных, использованных для анализа, приведён в табл. 1.

В качестве целевой переменной была взята оценка отпада древостоя согласно продукту глобального изменения лесного покрова GFC (Hansen et al., 2013).

Предварительная обработка данных включала этап растеризации для векторных данных (база данных пожаров растительности), а также перепроецирование растровых данных. Помимо этого, выполнялась фильтрация с целью исключения из рассмотрения некорректных данных, для чего использовались флаги качества соответствующих продуктов. Также для продукта отражательной способности поверхности формировались помесечные композиты,

рассчитываемые как средние значения отражательной способности в течение месяца. Для расчёта разностных спектральных индексов использовались значения отражательной способности поверхности в первый месяц после окончания пожара (послепожарное значение) и в тот же месяц предыдущего года (допожарное значение).

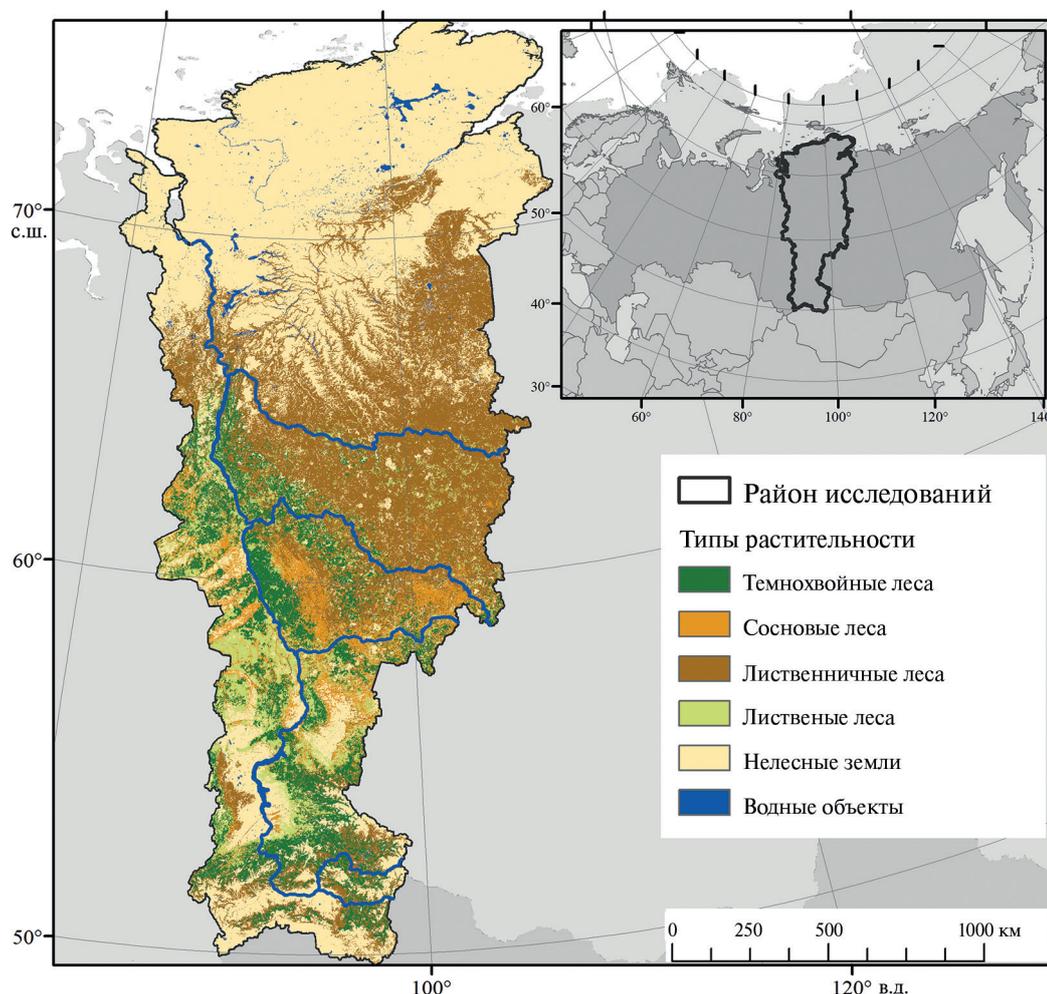


Рис. 1. Район исследований. Преобладающие древесные породы на территории района исследования показаны согласно карте растительности ИКИ РАН (<http://pro-vega.ru/maps/>) (Барталев и др., 2016). На вставке указано расположение района исследования. Цифрами отмечены преобладающие породы: 1 — темнохвойные леса; 2 — леса с преобладанием сосны; 3 — лиственничные леса; 4 — лиственные леса; 5 — нелесные земли

Таблица 1. Список переменных, использованных в анализе

Название	Описание
Тип леса	Категориальный признак согласно работе (Барталев и др., 2016)
Тип почвы	Категориальный признак согласно статье (Stolbovoi, Savin, 2002)
Район	Категориальный признак согласно публикации (Приказ..., 2014)
Высота	Высота над уровнем моря (в м), согласно ASTER GDEM (<i>англ.</i> Advanced Spaceborne Thermal Emission and Reflection Radiometer Global Digital Elevation Model) (Abrams et al., 2020)
Экспозиция	Категориальный признак (С, СВ, В, ЮВ, Ю, ЮЗ, З, СЗ), согласно ASTER GDEM (Abrams et al., 2020)
Крутизна	Крутизна склона (в град), согласно ASTER GDEM (Abrams et al., 2020)

Название	Описание
Месяц	Календарный месяц, на который приходилась большая часть времени действия пожара
Длительность	Длительность пожара (дни)
FRP	Суммарное значение радиационной мощности за время действия пожара (в МВт). Расчёт по данным продукта тепловых аномалий MODIS (Giglio et al., 2020)
SPEI-0	Значение индекса засушливости SPEI (<i>англ.</i> Standardized Precipitation Evapotranspiration Index) (Vicente-Serrano et al., 2010) в месяц пожара по данным SPEI Global Drought Monitor
SPEI-1	Значение индекса засушливости SPEI в предшествующий пожару месяц
SPEI-2	Среднее значение индекса засушливости SPEI за два месяца, предшествующих пожару
dBAI	Разность значений индекса BAI (<i>англ.</i> Burn Area Index) до и после пожара. Расчёт по данным продукта отражательной способности MODIS (Vermote et al., 2023)
dNDVI	Разность значений индекса NDVI до и после пожара. Расчёт по данным продукта отражательной способности MODIS
dEVI	Разность значений индекса EVI (<i>англ.</i> Enhanced Vegetation Index) до и после пожара. Расчёт по данным продукта отражательной способности MODIS
dNBR	Разность значений индекса NBR до и после пожара. Расчёт по данным продукта отражательной способности MODIS
dMSAVI	Разность значений индекса MSAVI (<i>англ.</i> Modified Soil Adjusted Vegetation Index) до и после пожара. Расчёт по данным продукта отражательной способности MODIS
dRed	Разность значений в красном канале до и после пожара. Расчёт по данным продукта отражательной способности MODIS
dNIR	Разность значений в ближнем инфракрасном канале до и после пожара. Расчёт по данным продукта отражательной способности MODIS
dSWIR	Разность значений в коротковолновом инфракрасном канале до и после пожара. Расчёт по данным продукта отражательной способности MODIS
Отпад древостоя	Послепожарный отпад древостоя по данным продукта глобального изменения лесного покрова GFC (Hansen et al., 2013)

Таким образом, все переменные, задействованные в дальнейшем анализе, имели формат растровых ГИС-слоёв (геоинформационная система), приведённых к единой системе координат, которая соответствовала продукту изменения лесного покрова (система координат WGS 1984 (*англ.* World Geodetic System 1984)). Пространственное разрешение сформированных растров при этом было в 10 раз меньше, чем в случае продукта GFC ($0,0025^\circ$ вместо $0,00025^\circ$). Таким образом, каждая ячейка выходного растра содержала 100 пикселей продукта GFC, что позволило характеризовать степень послепожарного отпада в каждом пикселе результирующего растра числом от 0 до 100. Поскольку вызванная пожарами гибель древостоя может наблюдаться в течение нескольких лет после пирогенного воздействия, использовалось пороговое значение в три года, согласно подходу, использованному ранее для выделения нарушений, связанных с воздействием лесных пожаров (Krylov et al., 2014).

Для отнесения каждой ячейки выходного растра к классу «отпад» или «отсутствие отпада» древостоя (её бинарная классификация) было использовано пороговое значение 70 %, т.е. вся ячейка классифицировалась как отпад древостоя, если гибель древостоя согласно продукту GFC отмечалась на 70 % или более её площади. Данное пороговое значение было получено путём сопоставления данных GFC с независимым продуктом усыхания древостоев, доступным через сервис «Вега» (ИКИ РАН, Москва, <http://pro-vega.ru/maps/>).

В работе были использованы материалы за 20-летний срок наблюдений, охватывающий период с 2003 по 2022 г. Как отмечалось ранее (Барталев и др., 2017; Швецов, 2024), пожары, приводящие к гибели лесов, наблюдаются, главным образом, в летние месяцы. В связи с этим

в анализе использовались данные только за июнь – август каждого года. В результате был сформирован набор данных, где наборам входных признаков (табл. 1) соответствовали значения целевой переменной (процент гибели древостоя). Общий объём данных, использованных для разработки модели, составлял около 220 тыс. пикселей.

Попиксельная классификация сформированных растров и прогнозирование гибели древостоев выполнялась с применением метода RF (Breiman, 2001). Разработка RF-модели выполнялась на языке Python (версия 3.11) с использованием библиотеки машинного обучения Scikit-learn (версия 1.3.2). Параметры модели (количество деревьев решений, максимальная глубина и т.д.) соответствовали значениям, применяемым в библиотеке Scikit-learn по умолчанию, так как вариация этих параметров на проверочной выборке (~10 % от общего объёма) не привела к существенному росту точности классификации. С помощью рекурсивного исключения признаков (англ. Recursive Feature Elimination) выполнялась оптимизация числа признаков, используемых для построения модели. Обучающая выборка включала 70 % от общего объёма, а ещё 20 % задействовалось для финального тестирования модели.

Для оценки точности модели применялись функции библиотеки Scikit-learn (модуль metrics). Выполнялся расчёт матрицы ошибок, а также набора стандартных метрик, используемых в задачах оценки качества классификации. Такие метрики включали в себя ОА (англ. Overall Accuracy), Precision, Recall и F_1 -score (Douzas et al., 2019).

Результаты и обсуждение

Относительная значимость признаков показана на рис. 2. Данные рис. 2 получены по результатам 10 запусков процедуры построения модели. Наибольшая относительная значимость среди использованных признаков выявлена для индекса dNBR (28 %). В целом можно отметить, что на признаки, характеризующие изменение спектральных свойств поверхности в результате воздействия пожара, суммарно приходилось около 56 % от относительной значимости по всем признакам, а на лесорастительные и топографические признаки — около 20 %.

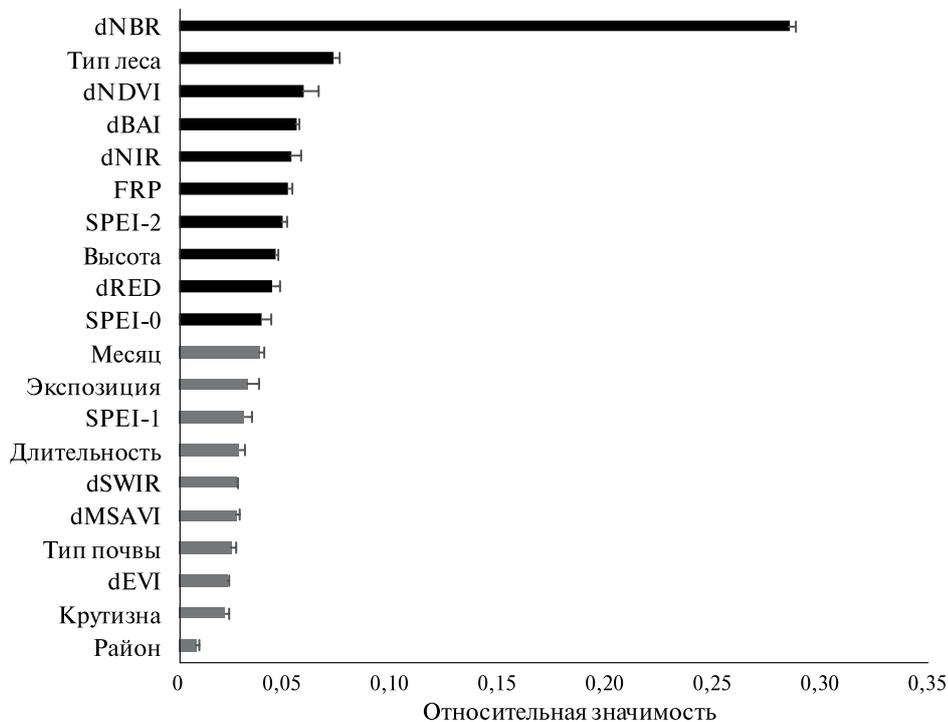


Рис. 2. Относительная значимость признаков (среднее значение и стандартное отклонение) в порядке убывания по результатам 10 запусков процедуры построения модели. Чёрным цветом выделены наиболее значимые признаки, использованные в финальной версии модели

Наиболее значимыми среди них оказались тип леса (7,3 %) и высота над уровнем моря (4,5 %). Значимым признаком оказался индекс засушливости SPEI как во время действия пожара, так и в предшествующий период (SPEI-0 и SPEI-2 — 3,8 и 4,8 % соответственно). Признак SPEI-1, вероятно, учитывался в признаке SPEI-2, и значимость его была невелика. Относительная значимость этих трёх признаков составила 11 %. Среди признаков, связанных с характеристиками горения, наибольшую значимость демонстрировало суммарное значение FRP (5,1 %). Относительная значимость этой группы признаков составляла около 11 %.

Ранее для Сибири была показана связь индекса NBR со степенью пирогенной нарушенности территории (Барталев и др., 2010; Delcourt et al., 2021), что подтверждает преобладающее влияние данного признака на результаты классификации. Также можно отметить достаточно высокую значимость индексов dBAI и dNDVI, которые также достаточно широко применяются для оценки пирогенной нарушенности (Cuevas-González et al., 2009; Furniss et al., 2020). Ещё следует обратить внимание на то, что использованные спектральные индексы демонстрируют мультиколлинеарность на уровне $R^2 = 0,5...0,65$. Хотя в использованной модели RF влияние мультиколлинеарности на результат меньше, чем, например, в случае линейных моделей, при выборе моделью одного из таких признаков значимость остальных, тем не менее, может снижаться.

Для упрощения итоговой модели из анализа были исключены 10 признаков, показавших наименьшую значимость (отмечены серым на *рис. 2*). Общая точность модели при использовании всего набора признаков составила 0,87, а после исключения наименее значимых признаков — 0,84. Метрики точности классификации приведены в *табл. 2*.

Таблица 2. Метрики точности итоговой классификации

Класс	Модель		Точность производителя
	Непородосменные	Породосменные	
Непородосменные	26 095	3 596	0,88
Породосменные	3 144	10 657	0,78
Точность пользователя	0,89	0,75	
Общая точность	0,84		
Precision	0,75		
Recall	0,78		
F_1 -score	0,76		

На *рис. 3* (см. с. 177) показаны зависимости доли отпада древостоев от величины 10 признаков с наибольшей относительной значимостью. Можно отметить, что для большинства признаков наблюдается нелинейная связь с величиной послепожарного отпада древостоя. Исключением являются индексы dNBR и dNDVI, где зависимость близка к линейной (см. *рис. 3а, в*), о чём также сообщалось ранее (Барталев и др., 2010; Швецов, 2022). Наибольшая доля послепожарного отпада древостоя наблюдалась в лесах с преобладанием лиственных и темнохвойных древостоев (30–40 %) (см. *рис. 3б*). Более высокая доля послепожарного отпада древостоев, характерная для лесов региона с преобладанием темнохвойных пород и лиственницы, отмечалась также в ранее опубликованных работах (Барталев и др., 2015; Пономарёв, Швецов, 2024). Кроме того, можно отметить, что эти типы древостоев преобладают в горных районах на юге рассмотренного региона, чем может объясняться рост доли отпада на высотах более 1400 м (см. *рис. 3и*).

Влияние погодных условий, описываемых с помощью индекса засушливости SPEI, на степень пирогенного воздействия и, как следствие, на долю послепожарного отпада древостоя показана на *рис. 3ж и л*. Доля отпада древостоя демонстрировала снижение по мере роста индекса SPEI, что соответствует росту увлажнённости территории (см. *рис. 3ж, л*) (Vicente-Serrano et al., 2010).

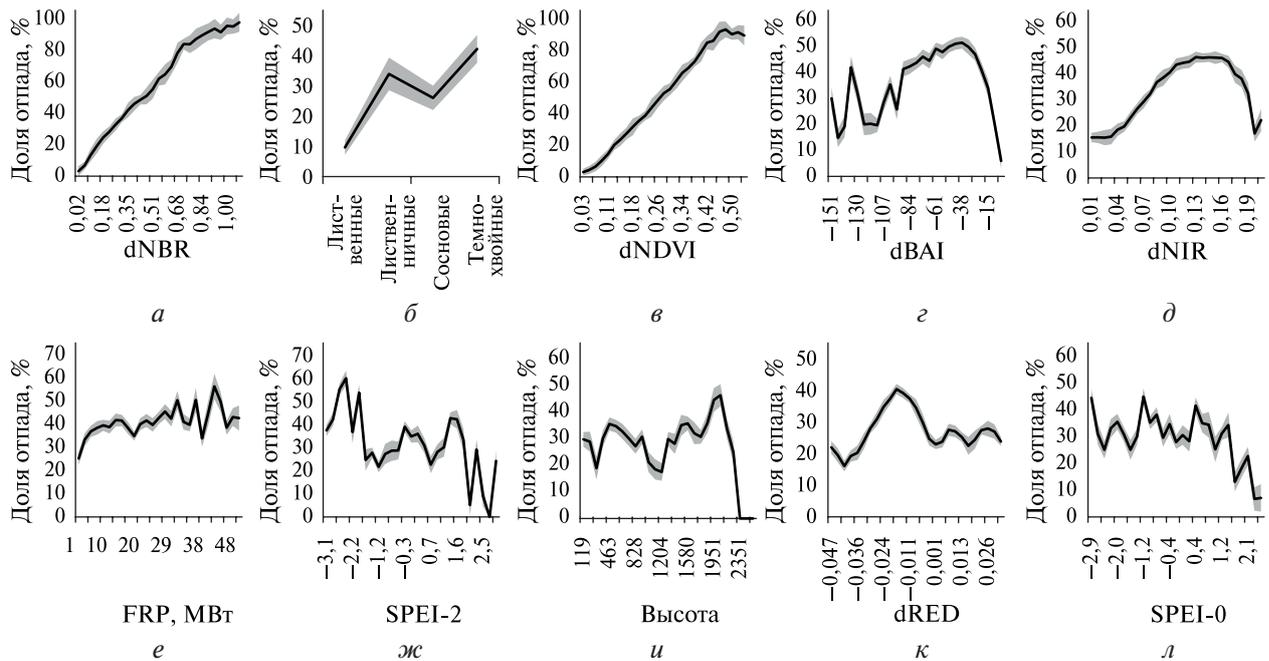


Рис. 3. Доля отпада древостоев для 10 признаков с наибольшей относительной значимостью. Графики расположены в порядке убывания значимости признаков (см. рис. 2). Чёрная линия соответствует среднему значению, серая область — стандартному отклонению

Заключение

Для территории Средней Сибири, включая территории Красноярского края, Республик Хакасия и Тыва, с использованием спутниковых данных и методов машинного обучения проведён анализ значимости факторов, влияющих на величину послепожарного отпада древостоя. Также была выполнена оценка точности прогнозирования отпада с применением модели случайного леса. Результаты показали удовлетворительную общую точность (0,84), величина F_1 -score составила 0,76. Наибольшую относительную значимость при прогнозировании доли послепожарного отпада демонстрировал разностный индекс dNBR. Помимо него к числу значимых признаков были отнесены спектральные индексы dNDVI и dBAI, а также изменения отражательных свойств поверхности в ближнем инфракрасном и красном диапазонах спектра. В целом признаки, характеризующие изменения отражательных свойств поверхности, определяли более 50 % вариабельности в оценках послепожарного отпада. Породный состав древостоя также оказался значимым признаком при прогнозировании величины отпада; темнохвойные и лиственничные леса характеризовались более высоким уровнем гибели древостоев. Можно отметить и значимое влияние погодных условий, определяющих влажность растительных горючих материалов как во время действия пожара, так и в предшествующий пожару период. Дальнейшее повышения точности прогнозирования, вероятно, возможно за счёт дополнительной оптимизации признаков, а также применения других алгоритмов, в том числе нейронных сетей.

Работа выполнена в рамках государственного задания № FWES-2025-0025 (ФИЦ КНЦ СО РАН).

Литература

1. Барталев С.А., Егоров В.А., Крылов А.М. и др. Исследование возможностей оценки состояния поврежденных пожарами лесов по данным многоспектральных спутниковых измерений // Современные проблемы дистанционного зондирования Земли из космоса. 2010. Т. 7. № 3. С. 215–225.

2. *Барталев С. А., Стыщенко Ф. В., Егоров В. А., Лупян Е. А.* Спутниковая оценка гибели лесов России от пожаров // *Лесоведение*. 2015. № 2. С. 83–94.
3. *Барталев С. А., Егоров В. А., Жарко В. О. и др.* Спутниковое картографирование растительного покрова России. М.: ИКИ РАН, 2016. 208 с.
4. *Барталев С. А., Стыщенко Ф. В., Хвостиков С. А., Лупян Е. А.* Методология мониторинга и прогнозирования пирогенной гибели лесов на основе данных спутниковых наблюдений // *Современные проблемы дистанционного зондирования Земли из космоса*. 2017. Т. 14. № 6. С. 176–193. <https://doi.org/10.21046/2070-7401-2017-14-6-176-193>.
5. *Пономарёв Е. И., Швецов Е. Г.* Сопоставление оценок отпада древостоев Сибири после воздействия пожаров по дистанционным данным // *Современные проблемы дистанционного зондирования Земли из космоса*. 2024. Т. 21. № 6. С. 213–223. <https://doi.org/10.21046/2070-7401-2024-21-6-213-223>.
6. Приказ Министерства природных ресурсов и экологии Российской Федерации «Об утверждении Перечня лесорастительных зон Российской Федерации и Перечня лесных районов Российской Федерации» № 367 от 18.08.2014 (ред. от 02.08.2023).
7. *Швецов Е. Г.* Исследование влияния мощности теплоизлучения лесных пожаров на степень повреждения лесов на территории юга средней Сибири по спутниковым данным // *Современные проблемы дистанционного зондирования Земли из космоса*. 2022. Т. 19. № 5. С. 136–146. <https://doi.org/10.21046/2070-7401-2022-19-5-136-146>.
8. *Швецов Е. Г.* Оценка послепожарной гибели лесов в Сибири по спутниковым данным и ее связь с характеристиками пожаров // *Сибирский эколог. журн.* 2024. № 4. С. 550–560. <https://doi.org/10.15372/SEJ20240403>.
9. *Abrams M., Crippen R., Fujisada H.* ASTER Global Digital Elevation Model (GDEM) and ASTER Global Water Body Dataset (ASTWBD) // *Remote Sensing*. 2020. V. 12. No. 7. Article 1156. <https://doi.org/10.3390/rs12071156>.
10. *Breiman L.* Random Forests // *Machine Learning*. 2001. V. 45. P. 5–32. <https://doi.org/10.1023/A:1010933404324>.
11. *Collins L., Griffioen P., Newell G., Mellor A.* The utility of Random Forests for wildfire severity mapping // *Remote Sensing of Environment*. 2018. V. 216. P. 374–384. <https://doi.org/10.1016/j.rse.2018.07.005>.
12. *Cuevas-González M., Gerard F., Balzter H., Riaños D.* Analysing forest recovery after wildfire disturbance in boreal Siberia using remotely sensed vegetation indices // *Global Change Biology*. 2009. V. 15. P. 561–577. <https://doi.org/10.1111/j.1365-2486.2008.01784.x>.
13. *Delcourt C. J. F., Combee A., Izbicki B. et al.* Evaluating the differenced Normalized Burn Ratio for assessing fire severity using Sentinel-2 imagery in Northeast Siberian larch forests // *Remote Sensing*. 2021. V. 13. No. 12. Article 2311. <https://doi.org/10.3390/rs13122311>.
14. *Douzas G., Bacao F., Fonseca J., Khudinyan M.* Imbalanced learning in land cover classification: Improving minority classes' prediction accuracy using the geometric SMOTE algorithm // *Remote Sensing*. 2019. V. 11. No. 24. Article 3040. <https://doi.org/10.3390/rs11243040>.
15. *Escuin S., Navarro R., Fernández P.* Fire severity assessment by using NBR (Normalized Burn Ratio) and NDVI (Normalized Difference Vegetation Index) derived from LANDSAT TM/ETM images // *Intern. J. Remote Sensing*. 2008. V. 29. P. 1053–1073. <https://doi.org/10.1080/01431160701281072>.
16. *Florath J., Keller S.* Supervised machine learning approaches on multispectral remote sensing data for a combined detection of fire and burned area // *Remote Sensing*. 2022. V. 14. No. 3. Article 657. <https://doi.org/10.3390/rs14030657>.
17. *Friedl M. A., Sulla-Menashe D., Tan B. et al.* MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets // *Remote Sensing of Environment*. 2010. V. 114. P. 168–182. <https://doi.org/10.1016/j.rse.2009.08.016>.
18. *Furniss T. J., Kane V. R., Larson A. J., Lutz J. A.* Detecting tree mortality with Landsat-derived spectral indices: Improving ecological accuracy by examining uncertainty // *Remote Sensing of Environment*. 2020. V. 237. Article 111497. <https://doi.org/10.1016/j.rse.2019.111497>.
19. *Gibson R., Danaher T., Hehir W., Collins L.* A remote sensing approach to mapping fire severity in south-eastern Australia using Sentinel 2 and random forest // *Remote Sensing of Environment*. 2020. V. 240. Article 111702. <https://doi.org/10.1016/j.rse.2020.111702>.
20. *Giglio L., Schroeder W., Hall J. V., Justice C. O.* MODIS Collection 6 active fire product user's guide. 2020. 63 p.
21. *Hansen M. C., Potapov P. V., Moore R. et al.* High-resolution global maps of 21st-century forest cover change // *Science*. 2013. V. 342. P. 850–853. <https://doi.org/10.1126/science.1244693>.
22. *Kharuk V. I., Ponomarev E. I., Ivanova G. A. et al.* Wildfires in the Siberian taiga // *Ambio*. 2021. V. 50. P. 1953–1974. <https://doi.org/10.1007/s13280-020-01490-x>.
23. *Krylov A., McCarty J. L., Potapov P.* Remote sensing estimates of stand replacement fires in Russia, 2002–2011 // *Environmental Research Letters*. 2014. V. 9. Article 105007. <https://doi.org/10.1088/1748-9326/9/10/105007>.

24. Loupian E. A., Bourtsev M. A., Proshin A. A. et al. Usage experience and capabilities of the VEGA-Science system // *Remote Sensing*. 2022. V. 14. No. 1. Article 77. <https://doi.org/10.3390/rs14010077>.
25. Potapov P., Li X., Hernandez-Serna A. et al. Mapping and monitoring global forest canopy height through integration of GEDI and Landsat data // *Remote Sensing of Environment*. 2020. V. 253. Article 112165. <https://doi.org/10.1016/j.rse.2020.112165>.
26. Russian forests and climate change. What science can tell us / eds. Leskinen P., Lindner M., Verkerk P.J. et al. V. 11. European Forest Institute, 2020. 140 p. <https://doi.org/10.36333/wsctu11>.
27. Stolbovoi V., Savin I. Land resources of Russia — maps of soil characteristics, version 1 (GGD601). Boulder, Colorado, USA: National Snow and Ice Data Center, 2002. <https://doi.org/10.7265/vk5t-nh20>.
28. Vermote E. F., Roger J. C., Ray J. P. MODIS surface reflectance user's guide. Collections 6 and 6.1. Version 1.7. Washington, DC, USA: National Aeronautics and Space Administration, 2023, 40 p. <https://doi.org/10.5067/MODIS/MOD09A1.006>.
29. Vicente-Serrano S. M., Beguería S., López-Moreno J. I. A multi-scalar drought index sensitive to global warming: The standardized precipitation evapotranspiration index // *J. Climate*. 2010. V. 23. P. 1696–1718. <https://doi.org/10.1175/2009JCLI2909.1>.
30. Zhong S., Zhang K., Bagheri M. et al. Machine learning: New ideas and tools in environmental science and engineering // *Environmental Science and Technology*. 2021. V. 55. P. 12741–12754. <https://doi.org/10.1021/acs.est.1c01339>.

Predicting post-fire forest mortality using remote sensing data and machine learning

E. G. Shvetsov

Krasnoyarsk Science Center SB RAS, Krasnoyarsk 660036, Russia
E-mail: eugeneshvetsov11@yandex.ru

Large number of wildfires and burned area that occur annually in Siberia determine the need to develop tools for assessing and predicting the effects of wildfires on forests. The study uses open access data sets and machine learning methods to predict areas of post-fire forest mortality, as well as to assess the significance of features that determine the amount of tree mortality. Using a set of thematic satellite products and the Random Forest method for the territory of Krasnoyarsk Krai, Republics of Khakassia and Tyva, the significance of a set of features in predicting the proportion of forest stand loss after fire was assessed, and a Random Forest model was developed. The work used features describing both forest and topographic conditions of the territory, moisture content of forest fuels, as well as characteristics of wildfires and magnitude of changes in spectral properties of the surface caused by wildfires. The overall accuracy of the model was 0.84, and the F_1 -score was 0.76. The dNBR index showed the greatest relative importance in predicting the proportion of post-fire tree mortality. In general, the features characterizing changes in the surface reflective properties after fire (spectral indices dNBR, dBAI, as well as reflectivity in several spectral ranges) determined more than 50 % of variability in the estimates of post-fire tree mortality.

Keywords: wildfires, remote sensing, machine learning, Random Forest, Siberia

Accepted: 21.03.2025

DOI: 10.21046/2070-7401-2025-22-3-171-181

References

1. Bartalev S. A., Egorov V. A., Krylov A. M. et al., The evaluation of possibilities to assess forest burnt severity using multi-spectral satellite data, *Sovremennyye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2010, V. 7, No. 3, pp. 215–225 (in Russian).
2. Bartalev S. A., Stytsenko F. V., Egorov V. A., Loupian E. A., Satellite-based assessment of Russian forest fire mortality, *Lesovedenie*, 2015, V. 2, pp. 83–94 (in Russian).

3. Bartalev S.A., Egorov V.A., Zharko V.O. et al., *Sputnikovoe kartografirovaniye rastitel'nogo pokrova Rossii* (Land cover mapping over Russia using Earth observation data), Moscow: IKI RAN, 2016, 208 p. (in Russian).
4. Bartalev S.A., Stytsenko F.V., Khvostikov S.A., Loupian E.A., Methodology of post-fire tree mortality monitoring and predictions using remote sensing data, *Sovremennyye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2017, V. 14, No. 6, pp. 176–193 (in Russian), <https://doi.org/10.21046/2070-7401-2017-14-6-176-193>.
5. Ponomarev E.I., Shvetsov E.G., Comparison of estimates of post-fire forest mortality in Siberia based on remote sensing data, *Sovremennyye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2024, V. 21, No. 6, pp. 213–223 (in Russian), <https://doi.org/10.21046/2070-7401-2024-21-6-213-223>.
6. *Prikaz Ministerstva prirodnykh resursov i ehkologii Rossiiskoi Federatsii "Ob utverzhdenii Perechnya lesorastitel'nykh zon Rossiiskoi Federatsii i Perechnya lesnykh raionov Rossiiskoi Federatsii"* (Order of the Ministry of Natural Resources of Russian Federation "On approval of the List of forest growth zones of the Russian Federation and the List of forest regions of the Russian Federation"), No. 367, 18.08.2014 (as amended on 02.08.2023) (in Russian).
7. Shvetsov E.G., Study of the influence of Fire Radiative Power of forest fires on forest disturbance degree in southern regions of Central Siberia using satellite data, *Sovremennyye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2022, V. 19, No. 5, pp. 136–146 (in Russian), <https://doi.org/10.21046/2070-7401-2022-19-5-136-146>.
8. Shvetsov E.G., Assessment of post-fire forest loss in Siberia using satellite data and its relationship with characteristics of fires, *Contemporary problems of ecology*, 2024, V. 17, No. 4, pp. 488–496, <https://doi.org/10.1134/S1995425524700288>.
9. Abrams M., Crippen R., Fujisada H., ASTER Global Digital Elevation Model (GDEM) and ASTER Global Water Body Dataset (ASTWBD), *Remote Sensing*, 2020, V. 12, No. 7, Article 1156, <https://doi.org/10.3390/rs12071156>.
10. Breiman L. Random Forests, *Machine Learning*, 2001, V. 45, pp. 5–32, <https://doi.org/10.1023/A:1010933404324>.
11. Collins L., Griffioen P., Newell G., Mellor A., The utility of random forests for wildfire severity mapping, *Remote Sensing of Environment*, 2018, V. 216, pp. 374–384, <https://doi.org/10.1016/j.rse.2018.07.005>.
12. Cuevas-González M., Gerard F., Balzter H., Riaños D., Analysing forest recovery after wildfire disturbance in boreal Siberia using remotely sensed vegetation indices, *Global Change Biology*, 2009, V. 15, P. 561–577, <https://doi.org/10.1111/j.1365-2486.2008.01784.x>.
13. Delcourt C.J.F., Combee A., Izbicki B. et al., Evaluating the differenced Normalized Burn Ratio for assessing fire severity using Sentinel-2 imagery in Northeast Siberian larch forests, *Remote Sensing*, 2021, V. 13, No. 12, Article 2311, <https://doi.org/10.3390/rs13122311>.
14. Douzas G., Bacao F., Fonseca J., Khudinyan M., Imbalanced learning in land cover classification: Improving minority classes' prediction accuracy using the geometric SMOTE algorithm, *Remote Sensing*, 2019, V. 11, No. 24, Article 3040, <https://doi.org/10.3390/rs11243040>.
15. Escuin S., Navarro R., Fernández P., Fire severity assessment by using NBR (Normalized Burn Ratio) and NDVI (Normalized Difference Vegetation Index) derived from LANDSAT TM/ETM images, *Intern. J. Remote Sensing*, 2008, V. 29, pp. 1053–1073, <https://doi.org/10.1080/01431160701281072>.
16. Florath J., Keller S., Supervised machine learning approaches on multispectral remote sensing data for a combined detection of fire and burned area, *Remote Sensing*, 2022, V. 14, No. 3, Article 657, <https://doi.org/10.3390/rs14030657>.
17. Friedl M.A., Sulla-Menashe D., Tan B. et al., MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets, *Remote Sensing of Environment*, 2010, V. 114, pp. 168–182, <https://doi.org/10.1016/j.rse.2009.08.016>.
18. Furniss T.J., Kane V.R., Larson A.J., Lutz J.A., Detecting tree mortality with Landsat-derived spectral indices: Improving ecological accuracy by examining uncertainty, *Remote Sensing of Environment*, 2020, V. 237, Article 111497, <https://doi.org/10.1016/j.rse.2019.111497>.
19. Gibson R., Danaher T., Hehir W., Collins L., A remote sensing approach to mapping fire severity in south-eastern Australia using Sentinel 2 and random forest, *Remote Sensing of Environment*, 2020, V. 240, Article 111702, <https://doi.org/10.1016/j.rse.2020.111702>.
20. Giglio L., Schroeder W., Hall J.V., Justice C.O., *MODIS Collection 6 active fire product user's guide*, 2020, 63 p.
21. Hansen M.C., Potapov P.V., Moore R. et al., High-resolution global maps of 21st-century forest cover change, *Science*, 2013, V. 342, pp. 850–853, <https://doi.org/10.1126/science.1244693>.
22. Kharuk V.I., Ponomarev E.I., Ivanova G.A. et al., Wildfires in the Siberian taiga, *Ambio*, 2021, V. 50, pp. 1953–1974, <https://doi.org/10.1007/s13280-020-01490-x>.
23. Krylov A., McCarty J.L., Potapov P., Remote sensing estimates of stand replacement fires in Russia, 2002–2011, *Environmental Research Letters*, 2014, V. 9, Article 105007, <https://doi.org/10.1088/1748-9326/9/10/105007>.

24. Loupian E. A., Bourtsev M. A., Proshin A. A. et al., Usage experience and capabilities of the VEGA-Science system, *Remote Sensing*, 2022, V. 14, No. 1, Article 77, <https://doi.org/10.3390/rs14010077>.
25. Potapov P., Li X., Hernandez-Serna A. et al., Mapping and monitoring global forest canopy height through integration of GEDI and Landsat data, *Remote Sensing of Environment*, 2020, V. 253, Article 112165, <https://doi.org/10.1016/j.rse.2020.112165>.
26. *Russian forests and climate change. What science can tell us*, Leskinen P., Lindner M., Verkerk P.J. et al. (eds.), V. 11, European Forest Institute, 2020, 140 p., <https://doi.org/10.36333/wsctu11>.
27. Stolbovoi V., Savin I., *Land resources of Russia — maps of soil characteristics, version 1 (GGD601)*, Boulder, Colorado, USA: National Snow and Ice Data Center, 2002, <https://doi.org/10.7265/vk5t-nh20>.
28. Vermote E. F., Roger J. C., Ray J. P., *MODIS surface reflectance user's guide. Collections 6 and 6.1. Version 1.7*. Washington, DC, USA: National Aeronautics and Space Administration, 2023, 40 p., <https://doi.org/10.5067/MODIS/MOD09A1.006>.
29. Vicente-Serrano S. M., Beguería S., López-Moreno J. I., A multi-scalar drought index sensitive to global warming: The standardized precipitation evapotranspiration index, *J. Climate*, 2010, V. 23, pp. 1696–1718, <https://doi.org/10.1175/2009JCLI2909.1>.
30. Zhong S., Zhang K., Bagheri M. et al., Machine learning: New ideas and tools in environmental science and engineering, *Environmental Science and Technology*, 2021, V. 55, pp. 12741–12754, <https://doi.org/10.1021/acs.est.1c01339>.