

Санкт-Петербургский государственный университет

БАДМАЕВ Артем Тимурович

Выпускная квалификационная работа

***Построение вероятностной модели возникновения лесных пожаров на основе
природных факторов и антропогенной нагрузки***

Уровень образования: *бакалавриат*

Направление *50.03.01 «Искусства и гуманитарные науки»*

Основная образовательная программа *СВ.5045.2019 «Свободные искусства и науки»*

Профиль *«Компьютерные науки и искусственный интеллект»*

Научный руководитель:

доцент, кафедра проблем конвергенции
естественных и гуманитарных наук,
к.ф.-м.н., Князева Ирина Сергеевна

Рецензент:

ФГБУН Главная (Пулковская)
астрономическая обсерватория РАН,
к.ф.-м.н., Волобуев Дмитрий Михайлович

Санкт-Петербург

2023

Содержание

Введение	3
Глава 1. Моделирование риска лесных пожаров	8
1.1. Проблемы мониторинга и борьбы с лесными пожарами . . .	8
1.2. Современные подходы к моделированию	10
1.3. Применение систем моделирования в России	11
Глава 2. Построение модели	13
2.1. Постановка задачи	13
2.2. Переменные модели	13
2.3. Источники данных	15
2.4. Байесовская логистическая регрессия	17
2.5. Сверточная нейросеть U-Net	19
Глава 3. Анализ результатов	21
3.1. Регрессия	21
3.2. Нейросеть	23
Заключение	25
Список литературы	26

Введение

Данная работа посвящена прогнозированию возникновения лесных пожаров. Из года в год масштабные лесные пожары в России привлекают внимание множества людей и сопровождаются волной публикаций в прессе. Это особенно заметно в последние годы на фоне крупнейших в истории пожаров. По данным системы спутникового мониторинга ИСДМ-Рослесхоз в 2019 году пройденная огнем площадь приблизилась к рекорду 2012 года, превысив 15 млн га [1], а в 2021 был поставлен абсолютный рекорд за историю спутниковых наблюдений — площадь, пройденная огнем, превысила 18 млн га [2]. Притом последствия явлений таких масштабов не ограничиваются громкими заголовками в прессе и внушительными цифрами статистики. Оба раза густонаселенные районы Сибири накрыло дымом от пожаров на обширных территориях Красноярского края [3] и Якутии [4]. Прямой материальный ущерб насчитывал несколько миллиардов рублей [5, 6]. Притом это лишь фактически бухгалтерские оценки ущерба зарегистрированному имуществу, не включающие косвенные убытки, вред здоровью и долгосрочные последствия от рекордного сокращения лесного покрова [7].

Хотя в заявлениях официальных лиц периодически звучат тезисы о том, что лесные пожары являются нормальным природным явлением, а бороться с огнем в труднодоступных районах бессмысленно [8], исследования свидетельствуют об обратном. Крупные пожары приводят к выбросу в атмосферу больших объемов дыма и твердых частиц, которые могут значительно влиять на качество воздуха на населенных территориях, притом даже находящихся в сотнях и тысячах километрах от огня, представляя угрозу здоровью населения [9]. Кроме того, горение на большой площади и последующее уменьшение лесного покрова вносят дополнительный вклад в изменение климата. Дым и парниковые газы усиливают парниковый эффект, а уничтожение леса ухудшает способность планеты поглощать углекислый газ [10].

Несмотря на то, что лесные пожары действительно являются нормальной частью функционирования экосистемы тайги, последние десятилетия в результате антропогенных причин и изменения климата частота и масштабы пожаров растут [11]. Увеличивается пожароопасный сезон, а погода стано-

вится более засушливой, из-за чего снижается влажность топлива. Возгорания происходят чаще, а горение проходит интенсивнее. Масштаб пожаров значительно влияет на биоразнообразие и устойчивость экосистемы. И хотя влияние этих изменений не всегда однозначно негативное, предсказать дальнейший вектор развития при разбалансировке системы проблематично [11].

Расширение промышленного освоения северных районов Сибири также сказывается на состоянии экосистем и пожароопасности региона и ведет к их большей уязвимости [12]. Лесное хозяйство в нынешнем виде является дополнительной угрозой для лесов, так как одной из главных причин возгораний в тайге являются случайные и преднамеренные поджоги на лесозаготовках в силу слабого контроля за соблюдением правил пожарной безопасности [13]. Соответственно, изменения в антропогенных факторах дополнительно увеличивают риски лесных пожаров и требуют еще больших ресурсов для наблюдения и контроля за пожарной обстановкой [12]. В добавок, неутешительная динамика пожаров в России обещает еще и рост экономического ущерба для лесного фонда и лесного хозяйства.

Однако рост пожароопасности не единственная причина катастрофических масштабов лесных пожаров последнее десятилетие. Помимо факторов, влияющих на возникновение пожаров, важно учитывать пожароохранный режим и противодействие распространению огня, так как грамотные противопожарные меры позволяют смягчить последствия, минимизировать экономический ущерб и затронутую пожаром площадь. Однако немаловажным фактором успеха является своевременность реагирования. С вышедшим из под контроля пожаром крайне сложно бороться, ибо при масштабах лесов России для этого физически недостаточно ресурсов. Если пожар не удалось сдержать, то фокус обычно смещается на защиту населения, хозяйственных активов, инфраструктуры [14]. Поэтому важно своевременно выявлять и реагировать на возникающие очаги, пока они не успели выйти из под контроля.

Однако даже тушение таких небольших возгораний требует значительных усилий из-за обширной суммарной площади лесов и труднодоступности большинства из них, особенно в условиях нынешнего финансирования лесной охраны [15]. Сильная ограниченность в ресурсах не позволяет удачно

справляться с возросшим масштабом угроз. Это вынуждает выделять более приоритетные и менее приоритетные пожары. В России это разделение фактически фиксированное и отражено в делении на четыре зоны мониторинга, которые различаются способами слежения (наземный, авиационный, космический) и предписанными мерами реагирования на обнаруженный пожар [16].

Наибольшую критику вызывает так называемая «зона контроля», или зона космического мониторинга 2-го уровня, где осуществляется исключительно спутниковый мониторинг, а тушение пожаров не обязательно. Создание подобных зон оправданная мера, так как действительно, какие-то пожары целесообразно не тушить, тем самым высвобождая ресурсы для других [17]. Однако дефицит финансирования вынуждает делать зону контроля неоправданно большой. Именно там при неудачном сочетании погоды и отсутствия реагирования создаются условия для катастрофических пожаров масштаба 2019 и 2021 годов. Крупные пожары регулярно приводят к призывам о пересмотре зон контроля в связи с меняющимися природными и антропогенными условиями, что делает очаги даже в отдаленных районах более опасными [18].

Однако такие изменения требуют подробной оценки целесообразности тушения лесных пожаров на той или иной территории. Смысл деления на зоны — распределить ресурсы пожарной охраны, отдавая приоритет самым опасным пожарам, но притом тем, с которыми реально справиться. Тогда деление должно производиться с учетом потенциальных угроз, имеющихся ресурсов. Так как эти факторы очень изменчивы, будучи зависимыми от погоды, финансирования, поведения людей, то и прогнозирование должно быть достаточно гибким и отражать происходящие сдвиги. На основе подобных прогнозов уже будет осуществляться размещение противопожарных сил, которое позволит минимизировать время реагирования на обнаруженные очаги.

Создание такого прогноза — задача из области моделирования рисков. В качестве составляющих частей такого подхода выделяют моделирование угроз, подверженность риску, уязвимость и способность к противодействию [19]. В данном случае угрозы — это, собственно, лесной пожар. Подверженность риску — ценность имущества и активов, которым может быть нанесен ущерб, а уязвимость — собственно потенциальный ущерб при столкновении с угрозой. Способность к противодействию — потенциал для

уменьшения угрозы и связанного с ней ущерба.

Притом моделирование лесных пожаров — то есть угроз — можно разделить на два класса: модели возникновения и симуляции распространения. Первые обычно изучаются статистическим аппаратом и являются вероятностными моделями, то есть оценивающими вероятность возникновения пожара или достижения им тех или иных параметров. Тогда как с помощью симуляций строят подробные сценарии дальнейшего распространения огня на макроуровне. Классическим примером симуляции является Canadian Forest Fire Behavior Prediction System (CFFBPS) [20], российским аналогом которой является Информационная система дистанционного мониторинга (ИСДМ-Рослесхоз) [21]. Другими словами, модели возникновения оценивают вероятность возникновения нового пожара, тогда как симуляции распространения моделируют ожидаемое поведение уже существующего.

Вероятностные модели позволяют оценить вклад различных факторов в пожароопасность региона и инкорпорировать в предсказания неопределенность, неизбежно сопровождающую прогнозирование погоды и природных процессов. Так можно оценивать последствия изменения погоды и климата, влияние режима землепользования на вероятность возникновения пожара, оценивать эффективность мер противопожарной безопасности. Одну из таких вероятностных моделей пожарной активности предложили французские исследователи в рамках созданного ими байесовского фреймворка Firelihood [22].

Подход исследователей рассматривает возникновение пожаров как пространственно—временной пуассоновский процесс, где количество пожаров в день в одной ячейке пространственной сетки является целевой переменной и определяется распределением Пуассона. Параметры распределения в каждой пространственно-временной точке определялись обобщенной аддитивной моделью (Generalized Additive Model, GAM), что позволило учесть нелинейную зависимость параметров распределения от объясняющих переменных.

Притом альтернативным инструментом машинного обучения, который может учитывать нелинейные зависимости целевой переменной от объясняющих, являются нейронные сети, которые также успешно применялись для

моделирования пожарной активности. В частности, многослойный перцептрон был использован для аналогичной задачи моделирования вероятности возникновения лесных пожаров в Италии [23]. Рекуррентные нейросети применялись для моделирования динамики развития лесных пожаров в Канаде [24].

Однако задачу предсказания пожарной активности в регионе также можно рассматривать как задачу компьютерного зрения, если интерпретировать пространственную сетку с данными как многоканальное изображение. В таком случае, задачей модели будет сегментация этого изображения — определение границ «объектов», в нашем случае территорий с повышенном риском возникновения пожаров.

Задачу сегментации успешно решает архитектура сверточной нейросети U-Net [25], которая уже использовалась в сфере лесной охраны для сегментации контуров активных лесных пожаров на основе данных дистанционного зондирования Земли [26]. Кроме того, данная архитектура может быть использована в задачах попиксельной регрессии, то есть нахождении значений пикселей целевого изображения на основе значений соответствующих пикселей входных изображений [27].

Данная работа ставит своей целью реализовать модель предсказания пожарной активности на лесных территориях Сибири и Дальнего востока на основе архитектуры U-Net, а также сопоставить ее с обобщенной линейной моделью, чтобы оценить компромисс между интерпретируемостью и эффективностью моделирования.

Глава 1. Моделирование риска лесных пожаров

1.1 Проблемы мониторинга и борьбы с лесными пожарами

Географические условия и особенности экономического развития России ставят перед системой лесной охраны значительные вызовы. С одной стороны, большая площадь лесного покрова сама по себе создает условия для возникновения одновременно большого числа пожаров, их длительного горения и распространения. С другой, невысокая плотность населения и инфраструктуры ограничивают возможности по мониторингу и борьбе с лесными пожарами на обширных территориях тайги Сибири и Дальнего Востока, на которые приходится более 90% пройденной огнем площади [28].

Сочетание этих факторов привело к формированию системы лесной охраны в ее нынешнем виде. Отсутствие дорожной сети в отдаленных районах тайги делает невозможным наземное наблюдение и создание сети аэродромов для воздушного наблюдения, поэтому на значительной площади лесов осуществляется исключительно космический мониторинг [28]. Приоритет в тушении отдается пожарам, напрямую угрожающим населенным пунктам и объектам хозяйства, так как ресурсов Авиалесоохраны недостаточно для тушения всех обнаруженных пожаров [29].

Это, в свою очередь, делает возможным возникновение пожаров катастрофических масштабов на обширных территориях тайги. Так как региональные власти имеют право не тушить пожар в случае, если расходы на борьбу с ним превышают потенциальный ущерб, множество очагов в малонаселенных и отдаленных районах Сибири и Дальнего Востока остаются без внимания [17]. Несмотря на рекомендации Авиалесоохраны об определении границ «зон контроля» — территорий, на которых можно не осуществлять тушение — субъекты федерации определяют их по своему усмотрению, как правило расширяя за пределы рекомендуемых границ [29].

Власти регионов, на которых лежит ответственность за борьбу с пожарами на своей территории [29], руководствуются в первую очередь финансовыми и экономическими соображениями, вынуждены прибегать к такой практике в первую очередь по причине недостаточного финансирования [15].

Ограниченные ресурсы пожарной охраны, в частности авиацию и квалифицированный персонал, требуется держать наготове для борьбы с угрожающими населению пожарами, тем самым жертвуя малодоступными северными лесами [18].

Однако, если на этапе возникновения пожара в отдаленной северной тайге он действительно не представляет немедленной опасности, то в отсутствие всякого противодействия рискует достичь масштабов, которые будут представлять даже большую угрозу населенным пунктам и экономическим активам. Экономический анализ потенциального ущерба свидетельствует как раз о том, что вышедшие из-под контроля крупные пожары могут представлять большой риск, который, однако трудно оценить на этапе их возникновения [30].

Кроме того, несмотря на частую критику зон контроля, эксперты не отрицают необходимость ее существования как таковой [30, 31]. Возражения вызывают в первую очередь нынешние границы и площадь этих зон, тогда как само их наличие (только в ином виде) видится необходимым из-за уже упомянутой невозможности обеспечить авиационный мониторинг и своевременное реагирование на всей территории тайги [28, 30]. Так, Авиалесоохрана с 2015 года рекомендует определять зоны контроля по границе зоны космического мониторинга второго уровня — территорий, где регулярное патрулирование и обязательное тушение пожаров не предусмотрено изначально, и возможно лишь по решению органов власти субъекта [32]. Однако на деле власти регионов определяют зоны контроля в том числе на территориях зон космического мониторинга первого уровня [29], где Авиалесоохраной ранее предусматривалась борьба с пожарами средствами авиации [32].

Все вышеописанное означает, что задача космического мониторинга труднодоступных лесов останется актуальной для России в обозримом будущем, даже если будут приняты предложения о сокращении зоны контроля. Это же означает, что сохраняется проблема эффективного размещения и своевременного использования ограниченных ресурсов пожарной охраны для тушения потенциально опасных очагов возгорания. Так как в зонах космомониторинга второго уровня не планируется регулярное авиационное патрулирование, при принятии решений можно будет полагаться лишь на спут-

никовые данные. Притом при все еще большой площади данной зоны, велик шанс, что придется отдавать приоритет борьбе с одними, наиболее опасными пожарами, в ущерб тем, которые не представляют той же угрозы. И если ставить целью сокращение наносимого лесными пожарами ущерба, а не площади горения как таковой, то возникает необходимость моделирования рисков как эффективного способа прогнозирования ущерба [19, 28].

1.2 Современные подходы к моделированию

Моделирование рисков является объемной задачей, которую можно разделить на несколько компонентов: моделирование угроз (того, что наносит ущерб), уязвимостей (подверженности ущербу) и способности к противодействию (способности сократить ущерба) [19]. Моделирование угроз — в нашем случае это лесные пожары — ставит своей задачей установить связь между наблюдаемыми факторами, которые могут влиять на возникновение и распространение лесного пожара, и собственно самим пожаром. Оно является фундаментом оценки риска, так как угроза является непосредственным источником ущерба, тогда как остальные факторы риска (уязвимость и противодействие) лишь определяют, как степень угрозы транслируется в риск ущерба [33].

Модели угрозы могут предусматривать как физическое моделирование поведения пожара и являться объясняющими — описывающими физику процесса, — так и являться исключительно статистическими или предсказательными. В таком случае модели будут основаны не на представлениях о лежащих в основе возникновения лесного пожара природных явлениях и механизмах, а на методах анализа данных, и не будут обладать той же интерпретируемостью [19]. Подобные модели не используются для симуляции распространения конкретного пожара и являются менее детализированными, однако способны работать с большим объемом данных и моделировать значения агрегированных показателей, например, количество лесных пожаров или вероятность их возникновения на некоторой площади.

Эти преимущества становятся более актуальными в свете прогресса последних лет в области сбора данных. Новые инструменты космического

мониторинга (MODIS, VIIRS, LANDSAT) позволяют автоматически получать актуальные данные о возникающих пожарах и лесном покрове, а продвинутое климатические модели дают возможность сравнительно точного предсказания погоды в высоком пространственном разрешении [34], что открывает доступ к большому массиву данных. В этой связи для моделирования пожарной активности все больше применяются методы машинного обучения [19, 34]. Особняком стоят байесовские методы, которые позволяют моделировать вероятностную природу процессов [34].

1.3 Применение систем моделирования в России

В России также проводились исследования в области моделирования пожарной активности [35, 36], однако в целом область остается значительно недоисследованной, а предложенные в имеющихся работах методы значительно отстают от моделей, реализуемых в последние годы за рубежом [22, 23, 24]. У представленных российских работ временной масштаб моделирования ограничен годом или сезоном, тогда как изученные европейские модели способны предсказывать суточную активность, а также используют более продвинутые алгоритмы машинного обучения, включая глубокие нейронные сети.

Изыскания в этой сфере осложняются тем, что текущая главная российская система мониторинга — ИСДМ-Рослесхоз — предоставляет доступ преимущественно задействованным в пожарной охране и использовании лесного фонда ведомствам и частным компаниям [28]. Доступ к ресурсам системы для частных лиц, журналистов и исследователей слабо регламентирован и ограничен [29]. В открытом доступе информация предоставляется лишь в виде сводных отчетов по субъектам федерации и визуальных материалов (карт, графиков) в немашиночитаемом виде [37], что не позволяет сформировать на основе доступных широкой публике данных подходящий датасет. Кроме того, в последнее время наблюдается тенденция на ограничение доступа к этим и без того скудным данным: в 2022 году общественные деятели обращали внимание на то, что функция регистрации на веб-странице ИСДМ-Рослесхоз не работает [38]. При этом регистрация обязательная для получения «от-

крытых» данных.

Это вынуждает исследователей прибегать к открытым данным, основной массив которых безвозмездно предоставляют НАСА и Европейское космическое агентство, и создавать собственных инструментов обработки и анализа. Например, для независимой оценки площади лесных пожаров в России 2020 года российское отделение Гринпис проводило ручной анализ космических снимков прибора Sentinel-2, доступных благодаря сервису Sentinel Hub, и открытых данных американского инструмента VIIRS, доступных в рамках системы FIRMS [39].

Для большинства территорий за пределами США и ЕС самые продвинутые наборы предобработанных данных, которые наилучшим образом подходят для анализа и моделирования, как правило недоступны. Так, наиболее актуальные для анализа лесных пожаров данные о значениях индекса FWI и карты пройденных огнем территорий, создаваемые профильными организациями Евросоюза, доступны только для ограниченного набора стран Европы, Россия в которые не входит [40, 41]. Что касается российских ведомственных систем, то из-за ограниченного круга пользователей, независимые исследователи не имеют ни доступа к их продуктам, ни представления об их актуальных возможностях. Свидетельств использования системы для моделирования рисков в открытой информации Авиалесоохраны нет.

Глава 2. Построение модели

2.1 Постановка задачи

Целью данной работы является создание пространственно-временной модели возникновения лесного пожара с использованием доступных для территории России данных. Задачей модели является предсказание вероятности возникновения лесного пожара в течение в суток в пространственной ячейке с разрешением $0.1^\circ \times 0.1^\circ$. Сопоставляются качество предсказания и интерпретируемость двух моделей машинного обучения: байесовской обобщенной линейной модели и сверточной нейронной сети архитектуры U-Net.

2.2 Переменные модели

Общий вид переменных модели заимствуется из фреймворка Firelihood [22]. Оригинальное исследование в качестве предикторов использует погодный пожарный индекс FWI (Fire Weather Index) и относительную площадь лесного покрова (FA, Forest Area). Целевой переменной является количество активных пожаров площадью больше 1 га. Значения переменных рассчитывались посуточно для каждой ячейки пространственной сетки. В данной работе предикторы оставлены практически без изменений (различаются только источники данных и пространственное разрешение).

FWI (Fire Weather Index) — это канадский погодный индекс, отражающий предрасположенность погодных условий к возникновению и распространению лесного пожара. Индекс рассчитывается на основе показаний температуры, скорости ветра, относительной влажности воздуха и объема осадков. В основе расчета лежат физические модели, описывающие влияние влажности топлива на возникновение пожара, а также влажности топлива и скорости ветра на дальнейшее распространение огня. Данные о температуре, влажности и осадках, с учетом поправки на длину светового дня, определяют характеристики топлива, тогда как ветер. Кроме того, индекс является рекуррентным — его значение в определенный день зависит от значения предыдущего дня. Это позволяет отражать постепенные изменения во влажности топлива в зависимости от погодных условий.

Несмотря на то, что FWI является удобным и емким индикатором потенциальной пожарной активности, его сложно напрямую транслировать в метрику риска возникновения пожара. Во-первых, он учитывает только метеорологические факторы, но не берет во внимание наличие самого топлива. С одинаковым успехом можно рассчитать показатели FWI как для сравнительно прохладной и влажной, но очень лесистой северной тайги, так и для засушливой пустыни без леса в принципе. Из-за низкой влажности и высокой температуры значение индекса будет выше в пустыне, а не в тайге, даже несмотря на то, что в пустыне нечему гореть в принципе. Значит, при моделировании пожарной активности необходимо учитывать и характеристики топлива. Эту роль выполняет переменная FA — Forest Area. Ее значением является доля площади пространственной ячейки, которая покрыта плотной растительностью.

Однако в силу отсутствия в России открытой базы данных активных лесных пожаров и их характеристик оказалось невозможным напрямую заимствовать целевую переменную из опорной работы. Вместо моделирования количества пожаров, пришлось использовать показания прибора VIIRS, регистрирующего термические аномалии на поверхности Земли, также именуемые «термоточками». Термоточку нельзя интерпретировать как отдельный лесной пожар (крупный пожар может создавать сразу множество термоточек), поэтому предсказание количества термоточек, по аналогии с предсказанием числа пожаров в ячейке, видится менее релевантным.

По этой причине в качестве целевой переменной рассматривается альтернативный вариант: наличие одной или нескольких термоточек в ячейке. Несмотря на вероятность ложно-положительных показаний прибора VIIRS, наличие хотя бы одной активной термоточки в пространственной ячейке можно с достаточной уверенностью интерпретировать как наличие в ней активного лесного пожара. В то же время, количество термоточек в ячейке не поддается такой же уверенной интерпретации. Например, одно и то же количество термоточек в ячейке может означать как и множество пожаров в разных местах, так и один крупный пожар.

2.3 Источники данных

Отдельную сложность представляли сбор и обработка необходимых данных. Готовых датасетов с показаниями FWI или любого другого аналогичного индекса пожарной опасности для территории России нет, как нет и открытых датасетов площади лесного покрова. Поэтому, для расчета FWI использовались метеорологические данные и адаптированный для языка программирования Python [42] алгоритм вычисления [43]. Для работы с большими массивами многомерных данных (два пространственных и временное измерения) исходный код был переписан с использованием пакета для векторных вычислений NumPy [44].

Погодные данные для расчета FWI взяты из глобального датасета агрометеорологических показателей, доступного в рамках программы Евросоюза Copernicus Climate Change Service [45]. Датасет составлен на основе реанализа — совмещений наблюдений погоды и климатических моделей, экстраполирующих наблюдения в непрерывный набор данных. Пространственное разрешение датасета — $0.1^\circ \times 0.1^\circ$, временное разрешение — одни сутки. Из датасета были взяты показания среднесуточной температуры воздуха на высоте 2 м; среднесуточная скорость ветра на высоте 10 м; суточный уровень осадков; относительная влажность на высоте 2 м, замеренная в 6:00, 9:00, 12:00, 15:00 и 18:00 по местному времени. В дальнейшем для расчета FWI использовался минимальный суточный показатель относительной влажности.

Для подсчета площади лесного покрова использовались данные о классификации земель, полученные с помощью спутниковых наблюдений [46]. Данный датасет также является продуктом программы Copernicus Climate Change Service и представляет собой пространственную сетку с разрешением 300 м. Каждой ячейке присвоен определенный класс землепользования. Относительная площадь лесного покрова рассчитывалась как доля пространственных ячеек, соответствующих классам «Деревья», «Кустарники» и «Луга» относительно общего числа ячеек классификации разрешением 300 м, содержащихся в более крупной ячейке переменной FWI с разрешением $0.1^\circ \times 0.1^\circ$ (что соответствует как минимум разрешению 11 км \times 2.5 км для самых северных широт РФ) [22].

Данные о термических аномалиях взяты из открытого архива системы FIRMS, работающей в рамках программы NASA Earth Observing System Data and Information System [47]. Для унификации использовались исключительно показания прибора VIIRS S-NPP 375m, данные которого доступны с 2012 года по настоящее время. Это один из трех доступных радиоспектрометров NASA, наряду с более старым инструментом MODIS и более продвинутым VIIRS NOAA-20 375m. Выбор в пользу S-NPP обоснован более точными и качественными показаниями, в сравнении с MODIS [47] и в то же время достаточно длинным временным промежутком доступных данных, в отличие от NOAA-20, доступного лишь с 2020 года.

В дальнейшем термоточки нужно было перевести из координатного представления, где каждой точке соответствовали собственные координаты на непрерывной оси, в сетчатое, с дискретными показателями, соответствующее представлению остальных переменных. Значением каждой ячейки в таком представлении является количество термоточек в ее границах, либо наличие хотя бы одной термоточки. Для подсчета использовался функционал пакета для обработки географических данных GeoPandas, который позволяет работать с векторным представлением данных (к которым относятся точки в системе координат) [48], и пакета для обработки многомерных массивов Xarray, предоставляющий функционал для растрового представления (то есть регулярной сетки) [49]. Таким образом, данные о термических аномалиях были конвертированы из векторного в растровый формат.

Период использованных наблюдений: с 2017 по 2021 год. Использовались только данные с 1 мая по 30 ноября каждого года, что соответствует границам летне-осеннего пожароопасного сезона. Это сезон, когда не наблюдается значительного снежного покрова, а основная масса ландшафтных пожаров — это именно лесные пожары, генезис которых отличается от весенних пожаров, которые наиболее характерны для сельскохозяйственных земель и вызваны неосторожным обращением с огнем [50]. Каждому дню соответствовал многомерный пространственный массив, содержащий показания двух предикторов и целевой переменной. В качестве моделируемого региона была выбрана зона между 60° и 70° с. ш. и 130° и 140° в. д. Соответственно, пространственные массивы получили размерность 100×100 .

Ограничение размера исследуемого региона обусловлено оптимизацией размера датасета. Каждый пожароопасный сезон состоит из 184 дней, взят пятилетний период. Каждый день представляет из себя массив 100×100 . Итого, количество строк двумерного датасета определяется как $N_{row} = 184 \times 5 \times 100 \times 100$. То есть, полный датасет состоит из более чем 7 млн точек, что близко к пределу возможностей байесовских методов [22]. В то же время, это достаточно масштабный датасет изображений для обучения сверточной нейронной сети $N_{image} = 184 \times 5$ (подробнее о структуре данных применительно к каждому из методов см. п. 2.4 - 2.5) [25].

Регионом моделирования является территория юго-запада Республики Саха и севера Хабаровского края. Данный регион удобен для моделирования по той причине, что сочетает в себе разнообразный рельеф (равнины Ленского плато и горы Верхоянского хребта), крупные ландшафтные объекты (река Лена), варьирующуюся плотность лесного покрова, а также включает как отдаленные и малонаселенные территории, так и сравнительно освоенные районы к востоку от Якутска с развитой дорожной сетью. Выбранный регион позволяет обеспечить вариативность объясняющих переменных и пространственной структуры. Кроме того, территория Якутии является одной из самых пожароопасных, а также сочетает в себе различные зоны мониторинга Авиалесоохраны [51].

2.4 Байесовская логистическая регрессия

За основу линейной модели был взят фреймворк Firelihood, предложенный французскими исследователями для моделирования пожарной активности в северо-западном Средиземноморье [22]. Целевая переменная (количество пожаров больше 1 га) моделируется с помощью пуассоновского процесса — то есть, предсказанием модели является распределение Пуассона. Параметры распределения определяются обобщенной аддитивной моделью (Generalized Additive Model, GAM). Это значит, что параметры целевого распределения рассчитываются как сумма нелинейных функций предикторов. В базовом виде аддитивная модель в работе, описывающей Firelihood, была

задана следующим образом:

$$\log N_i \sim f_{FWI}(FWI_i) + f_{FA}(FA_i),$$

где N_i — количество пожаров в i -й ячейке, FWI_i и FA_i — соответствующие показатели переменных FWI и FA, а f_{FWI} и f_{FA} — нелинейные функции.

Однако в данной работе вместо GAM применяется обобщенная линейная модель (GLM), дабы получить более интерпретируемый результат и проверить, насколько качество предсказания линейной модели будет уступать GAM и насколько в моделировании необходимы нелинейные зависимости. Также вместо Пуассоновского процесса реализована бинарная байсовская логистическая регрессия, определяющая наличие пожара в ячейке. В регрессии использовались те же переменные, только умноженные на соответствующие коэффициенты β , а целевая переменная моделировалась с помощью распределения Бернулли:

$$Y \sim \text{Bernoulli}(\theta),$$

$$\text{logit } \theta = \beta_0 + \beta_{FWI} \cdot FWI_i + \beta_{FA} \cdot FA_i.$$

Модель была реализована на языке Python с помощью пакета для вероятностного программирования PyMC. Апостериорные распределения параметров модели семплировались методом Марковской цепи Монте-Карло (Markov Chain Monte-Carlo, MCMC) с помощью алгоритма NUTS (No-U-Turn Sampler) [52]. Переменные были заданы априорными распределениями следующим образом:

$$\beta_0 \sim \text{Normal}(-10, 5),$$

$$\beta_{FWI} \sim \text{Normal}(0, 1),$$

$$\beta_{FA} \sim \text{Normal}(2, 5).$$

В ходе обучения семплировалась цепочка из 5 тысяч шагов. Обучающей выборкой были показания для выбранного региона с 2017 по 2020 годы, включительно. Тестирование производилось на выборке 2021 года. Затем из апостериорного распределения целевой переменной были взяты выборки по 100 элементов.

2.5 Сверточная нейросеть U-Net

За основу сверточной нейросети (Convolutional Neural Network, CNN) взята архитектура U-Net, изначально предназначавшаяся для автоматического определения контуров опухолей на изображениях мозга, полученных в ходе МРТ [25]. В данном методе предсказание пожарной активности рассматривалось как задача компьютерного зрения. Если в линейной модели каждая пространственная ячейка не зависела от других, а предсказание осуществлялось исключительно на основе значений предикторов, то в данной нейронной сети одним элементом выборки являлась не одна пространственная ячейка в определенный день, а весь пространственный массив 100×100 . Два массива объясняющих переменных, объединенных в тензор размерности $2 \times 100 \times 100$ интерпретировались как двухканальное изображение 100×100 . Тогда целевая переменная рассматривалась в формате одноканального выходного изображения, значения каждого пикселя которого соответствуют наличию или отсутствию лесного пожара в соответствующей пространственной ячейке.

Применение CNN в данной задаче перспективно тем, что архитектура позволяет учитывать пространственные эффекты, не вводя при этом дополнительных пространственных переменных. В сверточных слоях нейросеть находит связи между соседними пикселями и каналами изображения, применяя к ним фильтры (ядра) — весовые коэффициенты значений пикселей. В слоях пулинга происходит уменьшение размерности. Так как обычно CNN состоит из нескольких последовательных процедур свертки и пулинга, это позволяет находить паттерны и взаимосвязи на разных масштабах. Кроме того, применение нелинейных функций активации (общее место всех нейросетей) создает нелинейные зависимости целевой переменной от предикторов [53].

То есть, сверточные нейросети обладают преимуществом GAM — нелинейностью. Но кроме того, уже на уровне своей архитектуры они учитывают пространственную структуру данных. К недостаткам такого подхода можно отнести вычислительную сложность обучения нейросетей, а также black-box архитектуру, то есть, обученная нейросеть мало поддается интерпретации. Притом, несмотря на объем необходимых вычислений, нейросети и суще-

ствующие программные средства для их создания хорошо справляются с обучением на очень больших выборках данных, тогда как методы и инструменты байесовского моделирования значительно сильнее ограничивают как сложность модели, так и размеры выборок.

U-Net — это одна из state-of-the-art архитектур сверточных нейросетей, которая завоевала свою популярность как нейросеть для сегментации, то есть определения контуров объектов на изображениях [25]. Также она успешно продемонстрировала себя в задаче попиксельной регрессии при обработке космических снимков [27]. Это позволяет предположить, что U-Net обладает потенциалом в предсказании пожарной активности в пространстве, так как фактически это также задача попиксельной регрессии.

Так как U-Net, состоит из 5 слоев сверток и пулинга. Так как в архитектуре используется окно пулинга 2×2 с шагом 2, то на выходе каждого слоя высота и ширина входного тензора сжимались в два раза. Всего таких слоев 5, то есть после всех сверток высота и ширина изображения должны сокращаться в 32 раза. Поэтому массивы оригинальных данных размеров 100×100 были обрезаны до 96×96 , чтобы в конечном итоге их можно было сжать до размерности 3×3 . Также было изменено число входных каналов до двух, соответствующих объясняющим переменным FWI и FA. В остальном оригинальная архитектура осталась без изменений.

Нейросеть была создана и обучена с помощью пакета PyTorch для языка Python [54]. В качестве функции потерь использовалась бинарная кросс-энтропия, а оптимизатором являлся AdamW [55]. Обучение длилось 30 эпох. Обучающая выборка состояла из показаний за 2017-2019 гг., валидационная — за 2020 год. Тестирование модели осуществлялось на выборке 2021 года.

Глава 3. Анализ результатов

3.1 Регрессия

В результате для каждой пространственной ячейке на выходе была получена выборка из 100 реализаций апостериорного распределения. На Рис. 1 один изображена динамика числа активных ячеек в регионе. На оси y находится количество ячеек, в которых находится хотя бы один лесной пожар, а по оси x дни пожароопасного сезона 2021 года. Синей линией обозначается усредненная сумма выборки пожаров. Закрашенная область обозначает 95% доверительный интервал. Оранжевой линией — реальное количество ячеек с зарегистрированными активными пожарами.

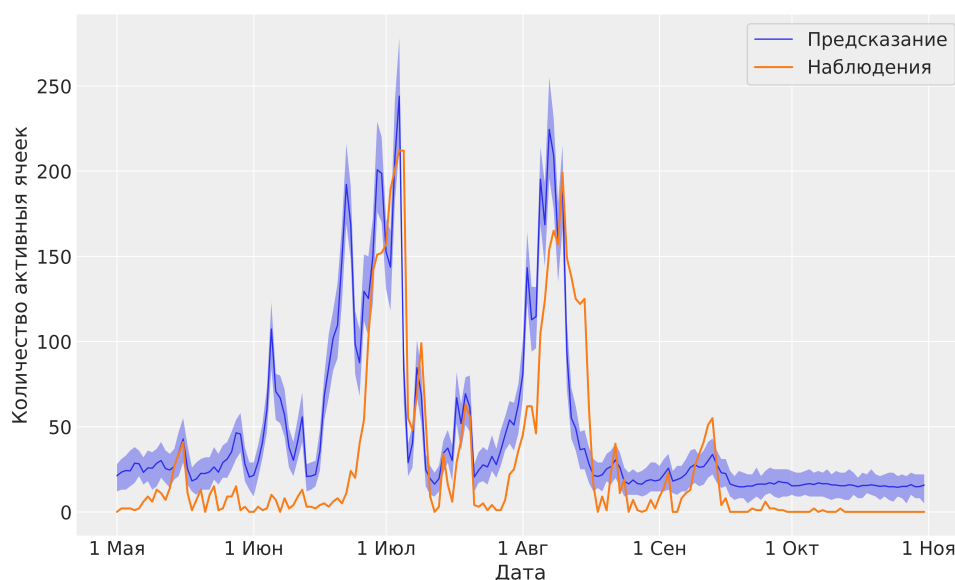


Рис. 1: Суточная динамика пожаров в регионе в 2021 году.

На графике предсказания модели в общих чертах повторяют динамику наблюдаемой величины. Однако видно, что большую часть периода низкой пожарной активности (первая половина июня, октябрь) предсказания модели стабильно превосходят наблюдения, которые также не укладываются в доверительный интервал.

Впрочем, если оценить пространственное распределение предсказаний, то даже в пиковый день предсказания модели значительно отличаются от реального географического распределения пожаров. На Рис. 2 слева изображе-

ны активные ячейки с активными лесными пожарами, а справа предсказанные моделью вероятности возникновения пожара. Как видно, предсказания лишь в очень общих чертах отражают пространственное распределение пожаров (увеличение по мере движения с севера на юг), а самые высокие значения вероятностей не совпадают с районами активных пожаров.

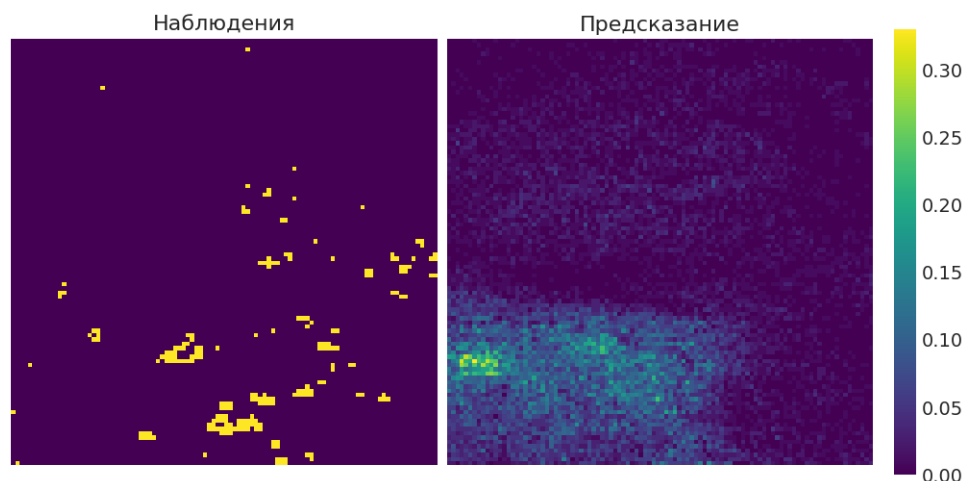


Рис. 2: Пространственное распределение активных пожаров и предсказаний GLM 4 июля.

Хотя регрессия достаточно удачно воспроизвела динамику пожаров в регионе в целом, она не показала тех же успехов в точечных предсказаниях лесных пожаров в каждой ячейке. Вероятно, на качестве модели сказались простота моделируемой линейной взаимосвязи и недостаток пространственного компонента. Также вероятностная природа модели не дала ожидаемых преимуществ, как как доверительные интервалы едва ли свидетельствуют о высокой неопределенности предсказаний. Модель делает предсказания уверенно, но делает их ошибочно.

Результаты свидетельствуют о том, что даже для качества упрощенной бинарной классификации может потребоваться нелинейность предсказания и учет пространственной корреляции. FWI и лесной покров недостаточны для уверенного предсказания, так как не учитывают неприродные факторы, влияющие на лесные пожары.

3.2 Нейросеть

Однако нейронная сеть продемонстрировала еще меньшие успехи. В ходе обучений приблизительно после 15-й эпохи функция потерь практически перестала убывать (Рис. 3), и после 30-й эпохи обучение было остановлено. Итоговое значение метрики $F1$ для категоризации каждого пикселя составило всего 0.038.

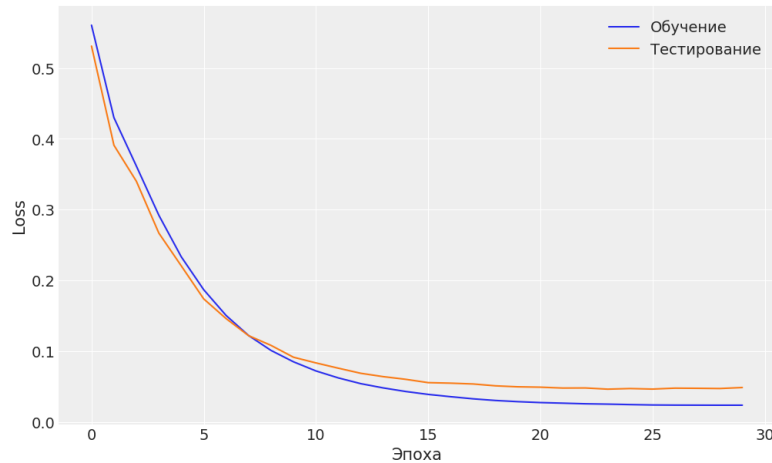


Рис. 3: Кривые обучения.

Основной проблемой предсказаний оказалось крайне низкие значений величин на выходе функции активации. На Рис. 4 изображена суточная предсказанная суточная динамика количества пожаров в зависимости от порога. Если значение пикселя оказывалось выше порогового, то данный пиксель определялся как активный. То есть, модель предсказывала в нем наличие лесного пожара. Как видно, уже при значении порога 0.2 модель не предсказывает ни одного пожара за год вообще. При меньших значениях динамика наблюдаемой величины воспроизводится в общих чертах, но очень грубо.

Пространственная динамика также моделируется неудовлетворительно (Рис. 5). Вероятно, на качестве U-Net сказывается довольно разреженная структура целевой переменной, которая редко образует большие и непрерывные «пятна» активных областей, что отличает данную задачу от классической сегментации изображения, когда нейросеть имеет дело со сравнительно более крупными областями изображений. Также используемые изображения данных обладали меньшим разрешением, чем используемые в задачах компьютерного

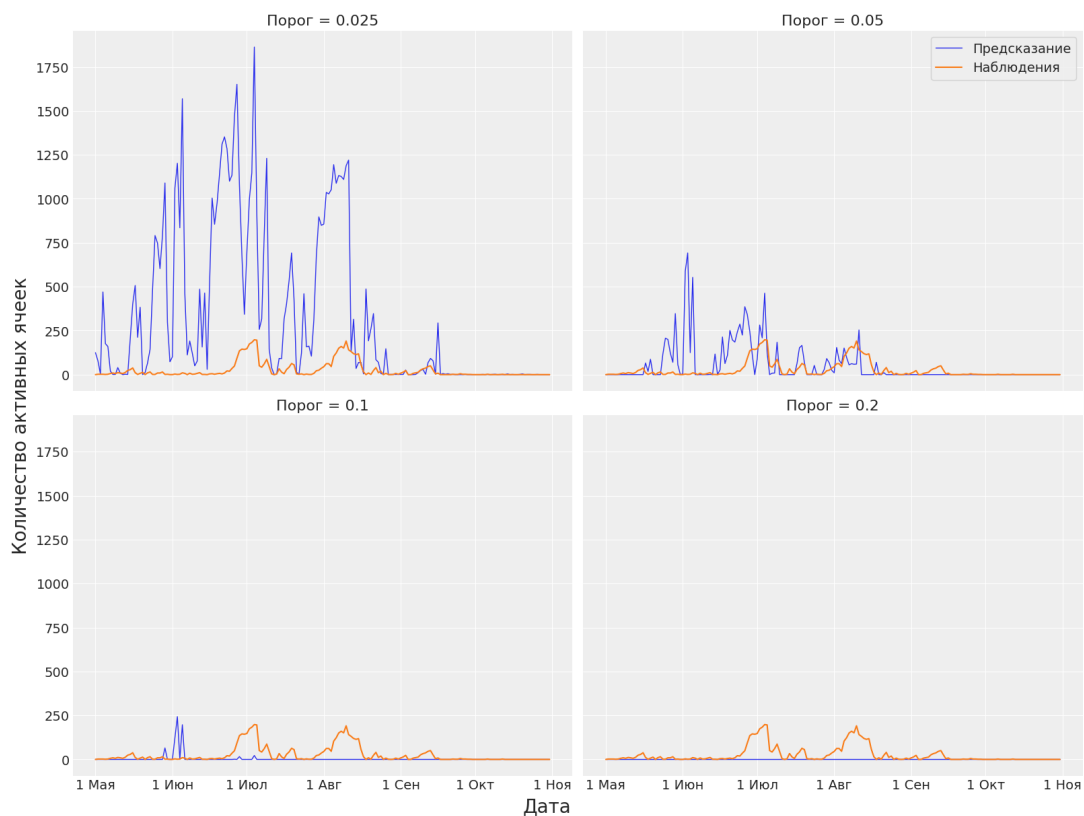


Рис. 4: Динамика активных пожаров и предсказаний U-Net в зависимости от значения порога.

зрения.

Не оправдались ожидания от потенциала пространственного моделирования. Модель показывает себя даже хуже, чем логистическая регрессия, не содержащая пространственного фактора в принципе. В том числе, нейросеть неудачно показывает себя в воспроизведении пространственной структуры.

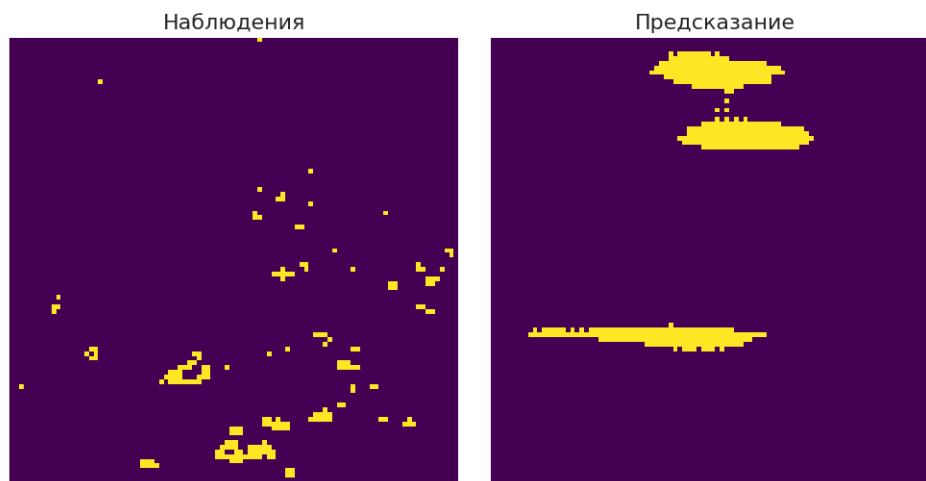


Рис. 5: Пространственное распределение активных пожаров и предсказаний U-Net при значении порога 0.05 4 июля.

Заключение

В ходе работы был собран массив необходимых для моделирования пожарной активности данных для территории России. Из метеорологических показателей, данных дистанционного зондирования и радиоспектрометрии был составлен пространственно-временной датасет значений погодного пожарного индекса FWI, относительной площади лесного покрова и пожарной активности. Были построены две модели машинного обучения для предсказания пожарной активности на территории Дальнего Востока: байесовская логистическая регрессия и сверточная нейронная сеть U-Net.

Впрочем, обе построенные модели показали невыдающиеся результаты. Байесовская логистическая регрессия справилась лишь с одной из подзадач — предсказанием активности на уровне региона. Однако и этот результат нельзя называть успешным, так как качество предсказания все равно осталось низким. Насколько известно, ранее подобные модели не реализовывались в рамках прогнозирования пожарной активности на территории России. При этом изученные в ходе работы методы основывались на иных источниках и представлениях данных, в значительной степени недоступных для России. В дальнейшем видится перспективным изучение сверточных нейросетей в решении указанной задачи, так как применение существующих архитектур не дает ожидаемых результатов.

Список литературы

- [1] 2019 год может стать рекордным по площади лесных пожаров. URL: <https://greenpeace.ru/news/2019/08/12/2019-god-mozhet-stat-rekordnym-po-ploshhadi-lesnyh-pozharov/>.
- [2] Куксин Григорий. 2021 год — полный рекордсмен по пожарам. URL: <https://greenpeace.ru/blogs/2021/09/17/2021-god-polnyj-rekordsmen-po-pozharam/>.
- [3] Дым над Сибирью: как живут люди в окружении лесных пожаров. URL: <https://www.bbc.com/russian/media-49410795>.
- [4] Дым от лесных пожаров вновь окутывает Сибирь, как в 2019 году. В Якутии горит около 5 млн гектаров леса. URL: <https://altapress.ru/zhizn/story/dim-ot-lesnih-pozharov-vnov-okutivaet-sibir-kak-v-godu-v-yakutii-gorit-okolo-mln-gektarov-lesa-290974>.
- [5] Пожары в России 2021 года вызвали рекордное сокращение лесного покрова на севере планеты. URL: <https://www.bbc.com/russian/news-61259933>.
- [6] В Якутии Ущерб От Лесных Пожаров в 2021 Году Превысил 3,7 Млрд Рублей. URL: <https://tass.ru/proisshestviya/13164579>.
- [7] В Минприроды Оценили Экономический Ущерб От Лесных Пожаров в России в 2021 Году. URL: <https://tass.ru/ekonomika/13265341>.
- [8] Красноярский губернатор назвал тушение ряда лесных пожаров бессмысленным. URL: <https://ria.ru/20190729/1556969481.html>.
- [9] Efimova Natalia V., Rukavishnikov Viktor S. Assessment of Smoke Pollution Caused by Wildfires in the Baikal Region (Russia). Т. 12, № 12. С. 1542. URL: <https://www.mdpi.com/2073-4433/12/12/1542>.
- [10] Wildfires, Global Climate Change, and Human Health / Rongbin Xu, Pei Yu, Michael J. Abramson [и др.]. Т. 383, № 22. С. 2173–2181.

- [11] Wildfires in the Siberian Taiga / Viacheslav I. Kharuk, Evgenii I. Ponomarev, Galina A. Ivanova [и др.]. Т. 50, № 11. С. 1953–1974. URL: <https://doi.org/10.1007/s13280-020-01490-x>.
- [12] Shvidenko A. Z., Schepaschenko D. G. Climate Change and Wildfires in Russia. Т. 6, № 7. С. 683–692. URL: <https://doi.org/10.1134/S199542551307010X>.
- [13] Гринпис получил первые результаты исследования причин лесных пожаров. URL: <https://greenpeace.ru/news/2021/08/10/greenpeace-poluchil-pervye-rezultaty-issledovaniya-prichin-lesnyh-pozharov/>.
- [14] Ярошенко Алексей. Алексей Ярошенко: «Важно не тушить пожары, а предотвратить их в следующем году». URL: <https://postnauka.ru/talks/102767>.
- [15] Аллахвердов Андрей. Greenpeace правительству: нужно больше денег на тушение пожаров. URL: <https://greenpeace.ru/blogs/2021/08/18/greenpeace-pravitelstvu-nuzhno-bolshe-deneg-na-tushenie-pozharov/>.
- [16] Информационная Система Дистанционного Мониторинга Лесных Пожаров Федерального Агентства Лесного Хозяйства РФ (Состояние и Перспективы Развития) / С.А. Барталев, Д.В. Ершов, Г.С. Коровин [и др.]. Т. 2, № 5. С. 419–429. URL: <https://aviales.ru/files/documents/2010/05/57.pdf>.
- [17] Куксин Григорий. Кто контролирует пожары в зонах контроля? (Спойлер: никто). URL: <https://greenpeace.ru/expert-opinions/2019/07/24/kto-kontroliruet-pozhary-v-zonah-kontrolja-spojler-nikto/>.
- [18] Аллахвердов Андрей. Зон контроля станет меньше. URL: <https://greenpeace.ru/blogs/2022/08/29/zon-kontrolja-stanet-menshe/>.
- [19] Oliveira Sandra, Rocha Jorge, Sá Ana. Wildfire Risk Modeling. Т. 23. С. 100274–100279. URL: <https://linkinghub.elsevier.com/retrieve/pii/S2468584421000465>.

- [20] Canadian Forest Fire Behavior Prediction (FBP) System. URL: <https://cwffis.cfs.nrcan.gc.ca/background/summary/fbp>.
- [21] Российская Система Спутникового Мониторинга Лесных Пожаров / Д.В. Ершов, Г.Н. Коровин, Е.А. Лупян [и др.]. Т. 1, № 1. С. 47–57. URL: <http://www.iki.rssi.ru/earth/trudi/0-05.pdf>.
- [22] Prediction of Regional Wildfire Activity in the Probabilistic Bayesian Framework of Firelihood / François Pimont, H       Fargeon, Thomas Opitz [и др.]. Т. 31, № 5. URL: <https://onlinelibrary.wiley.com/doi/10.1002/eap.2316>.
- [23] Estimating the Probability of Wildfire Occurrence in Mediterranean Landscapes Using Artificial Neural Networks / Mario Elia, Marina D’Este, Davide Ascoli [и др.]. Т. 85. С. 106474. URL: <https://linkinghub.elsevier.com/retrieve/pii/S0195925520304078>.
- [24] Liang Hao, Zhang Meng, Wang Hailan. A Neural Network Model for Wildfire Scale Prediction Using Meteorological Factors. Т. 7. С. 176746–176755.
- [25] Ronneberger Olaf, Fischer Philipp, Brox Thomas. U-Net: Convolutional Networks for Biomedical Image Segmentation. URL: <http://arxiv.org/abs/1505.04597>.
- [26] Active Fire Detection in Landsat-8 Imagery: A Large-Scale Dataset and a Deep-Learning Study / prefix=de useprefix=true family=Almeida Pereira, given=Gabriel Henrique, Andre Minoro Fusioka, Bogdan Tomoyuki Nassu [и др.]. Т. 178. С. 171–186. URL: <https://linkinghub.elsevier.com/retrieve/pii/S092427162100160X>.
- [27] Pixel-Wise Regression Using U-Net and Its Application on Pansharpening / Wei Yao, Zhigang Zeng, Cheng Lian [и др.]. Т. 312. С. 364–371. URL: <https://www.sciencedirect.com/science/article/pii/S0925231218307008>.
- [28] Ponomarev Evgeni I., Ivanov Valeri, Korshunov Nikolay. Chapter 10 - System of Wildfires Monitoring in Russia // Wildfire Hazards, Risks and Disasters /

под ред. John F. Shroder, Douglas Paton. Elsevier. С. 187–205. URL: <https://www.sciencedirect.com/science/article/pii/B9780124104341000105>.

- [29] Космический мониторинг лесных пожаров: история создания и развития ИСДМ-Рослесхоз / Р. В. Котельников, Е. А. Лупян, С. А. Барталев [и др.]. № 5. URL: <https://www.elibrary.ru/item.asp?id=39248600>.
- [30] Порфирьев Б. Н. О мнимой и реальной экономической эффективности борьбы с лесными пожарами в Сибири. Т. 49, № 11. С. 8–26. URL: <https://ecotrends.ru/index.php/eco/article/view/3913>.
- [31] Greenpeace подготовил предложения по пересмотру «зон контроля». URL: <https://greenpeace.ru/news/2019/08/15/greenpeace-podgotovil-predlozhenija-po-peresmotru-zon-kontrolja/>.
- [32] Федеральное агентство лесного хозяйства. Методические рекомендации по проведению космического мониторинга лесных пожаров на территории лесного фонда Российской Федерации. URL: https://aviales.ru/files/documents/2010/01/met_isdm2010_2.doc.
- [33] Scott Joe, Thompson Matthew, Calkin David. A Wildfire Risk Assessment Framework for Land and Resource Management. URL: <https://digitalcommons.unl.edu/usdafsfacpub/328>.
- [34] A Review of Machine Learning Applications in Wildfire Science and Management / Piyush Jain, Sean C.P. Coogan, Sriram Ganapathi Subramanian [и др.]. Т. 28, № 4. С. 478–505. URL: <https://cdnsiencepub.com/doi/full/10.1139/er-2020-0019>.
- [35] Янец П.К., Гададь С.Ж., Иванова С.А. Геоинформационное моделирование риска лесных пожаров в Республике Саха (Якутия). Т. 12. С. 37–42.
- [36] Звягинцева А. В., Яковлев Д. В. Перспективы Пространственного Анализа в Географических Информационных Системах Для Прогнозирования Риска Лесных Пожаров На Территории Воронежской Области - Гелиогеофизические Исследования. Т. 9. URL: <http://vestnik.geospace.ru/index.php?id=270>.

- [37] ФБУ «Авиалесохрана». Руководство пользователя информационных узлов ИСДМ-Рослесхоз. URL: https://nffc.aviales.ru/docs/isdm_docs/Open_interface.pdf.
- [38] Доступ к Системе Мониторинга Лесных Пожаров ИСДМ-Рослесхоз Через Портал Госуслуг и Единую Систему Идентификации и Аутентификации Не Работает Уже Семь с Половиной Месяцев. URL: 16 April 2022.
- [39] Оценка масштабов лесных пожаров 2020 года на территории России / И. В. Глушков, В. В. Лупачик, И. В. Журавлева [и др.]. Т. 4, № 2. URL: https://www.researchgate.net/profile/Ilna-Zhuravleva/publication/354124935_Assessment_of_landscape_fires_in_2020_in_Russia_of-landscape-fires-in-2020-in-Russia.pdf.
- [40] EFFIS - Data Request Form. URL: <https://effis.jrc.ec.europa.eu/apps/data.request.form/>.
- [41] Service Copernicus Climate Change. Fire Danger Indicators for Europe from 1970 to 2098 Derived from Climate Projections. URL: <https://cds.climate.copernicus.eu/doi/10.24381/cds.ca755de7>.
- [42] Van Rossum Guido, Drake Fred L. Python Reference Manual. CreateSpace.
- [43] Wang Y, Anderson K R, Suddaby R M. Updated Source Code for Calculating Fire Danger Indices in the Canadian Forest Fire Weather Index System.
- [44] Array Programming with NumPy / Charles R. Harris, K. Jarrod Millman, prefix=van der useprefix=true family=Walt, given=Stéfan J. [и др.]. Т. 585, № 7825. С. 357–362. URL: <https://www.nature.com/articles/s41586-020-2649-2>.
- [45] Boogaard H., Schubert J., De Wit A. [и др.]. Agrometeorological Indicators from 1979 to Present Derived from Reanalysis, Version 1.0. URL: <https://cds.climate.copernicus.eu/doi/10.24381/cds.6c68c9bb>.
- [46] Copernicus Climate Change Service, Climate Data Store. Land Cover Classification Gridded Maps from 1992

- to Present Derived from Satellite Observation. URL: <https://cds.climate.copernicus.eu/doi/10.24381/cds.006f2c9a>.
- [47] NASA FIRMS. NRT VIIRS 375 m Active Fire Product VNP14IMGD Distributed from NASA FIRMS. URL: <https://www.earthdata.nasa.gov/learn/find-data/near-real-time/firms/vnp14imgdlnrt>.
- [48] Geopandas/Geopandas: V0.13.0 / Kelsey Jordahl, prefix=den useprefix=false family=Bossche, given=Joris Van, Martin Fleischmann [и др.]. URL: <https://zenodo.org/record/7902652>.
- [49] Hoyer Stephan, Hamman Joe. Xarray: N-D Labeled Arrays and Datasets in Python. Т. 5, № 1. С. 10. URL: <https://openresearchsoftware.metajnl.com/articles/10.5334/jors.148>.
- [50] Ягодина Василиса, Аллахвердов Андрей. Весенние пожары в России. Антирейтинг регионов от Greenpeace. URL: <https://greenpeace.ru/stories/2020/06/02/vesennie-pozhary-v-rossii-antirejting-regionov-ot-greenpeace/>.
- [51] Схемы Расположения Зон Контроля в Субъектах Российской Федерации. URL: <https://aviales.ru/default.aspx?textpage=229>.
- [52] Hoffman Matthew D., Gelman Andrew. The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. URL: <http://arxiv.org/abs/1111.4246>.
- [53] Backpropagation Applied to Handwritten Zip Code Recognition / Y. LeCun, B. Boser, J. S. Denker [и др.]. Т. 1, № 4. С. 541–551. URL: <https://direct.mit.edu/neco/article/1/4/541-551/5515>.
- [54] Automatic Differentiation in PyTorch / Adam Paszke, Sam Gross, Soumith Chintala [и др.]. URL: <https://openreview.net/forum?id=BJJsrmfCZ>.

- [55] Loshchilov Ilya, Hutter Frank. Decoupled Weight Decay Regularization.
URL: <http://arxiv.org/abs/1711.05101>.