

Опыт прогнозирования массовых размножений сибирского шелкопряда *Dendrolimus sibiricus* Tschetv. по метеоданным

Д. А. ДЕМИДКО, А. А. ГОРОШКО, С. М. СУЛЬТСОН, Н. Н. КУЛАКОВА, П. В. МИХАЙЛОВ

Сибирский государственный университет науки и технологий
имени академика М. Ф. Решетнева
660037, Красноярск, Проспект имени газеты “Красноярский рабочий”, 31
E-mail: sawer-beetle@yandex.ru

Статья поступила 09.10.2023

После доработки 16.11.2023

Принята к печати 21.11.2023

АННОТАЦИЯ

Одним из важнейших этапов предотвращения дефолиации лесов является его прогнозирование. Значительную роль в формировании массовых размножений лесных дефолиаторов играет погода. Именно с использованием погодных данных в качестве предикторов был построен ряд моделей, более или менее успешно работающих. Мы предприняли такую попытку для сибирского шелкопряда (*Dendrolimus sibiricus* Tschetv.) – одного из наиболее опасных филлофагов южной тайги Сибири. Для модельной территории (южная тайга и подтайга Томской и Кемеровской областей, Красноярского края) построена модель на основе градиентного спуска (алгоритм XGBoost). Точность предсказания вспышек составила 0,952. Основными предикторами были температурные условия вегетационного периода за 4 и 5 лет до зафиксированного начала вспышки.

Ключевые слова: южная тайга, сибирский шелкопряд, массовые размножения, прогнозирование, погода, машинное обучение.

ВВЕДЕНИЕ

Дефолиация лесов насекомыми нередко охватывает до несколько миллионов гектар в год [Moulinier et al., 2011; Лямцев, 2014]. Сибирский шелкопряд *Dendrolimus sibiricus* (Tschetv.) (= *superans* (Butler)) [Kononov et al., 2016] (далее СШ) также способен к этому [Кондаков, 1974; Kharuk et al., 2018].

Его ареал занимает территорию бореальной зоны от Урала до Сахалина [Окунев, 1955; Рожков, 1963]. В южной тайге Сибири площади очагов могут достигать миллионов гектаров [Кондаков, 2002], значительная

часть поврежденных древостоев полностью или частично гибнет [Рожков, 1963; Кондаков, 1974; Гродницкий и др., 2002].

Мнение о влиянии погоды на подъем численности *D. sibiricus* высказано давно [Болдаруев, 1960; Коломиец, 1960] и подтверждено позже [Кондаков, 1974; Исаев и др., 2015; Суховольский и др., 2020]. Роль погоды в росте численности СШ до вспышечного уровня заключается в том, что она создает благоприятные условия для развития его гусениц. В таких условиях значительная часть популяции переходит с двухлетнего на однолетний цикл

развития (см. Биология сибирского шелкопряда) [Исаев и др., 2001; Кондаков, 2002], популяции наращивают свою численность за счет миграций бабочек в первичные очаги [Рожков, 1965] и эффекта группы [Исаев и др., 2001]. Как следствие резкого ускорения развития и роста численности, такие популяции уходят из-под контроля энтомофагов [Исаев и др., 2001]. Через несколько лет это с высокой вероятностью приводит к дефолиации, так как комплекс паразитоидов сибирского шелкопряда является одним из ключевых факторов сдерживания его численности, но действие этого фактора обладает значительной инерцией. Нарастание плотности популяции его естественных врагов отстает от роста численности СШ, и восстановить утерянный контроль они могут только через несколько лет на фазе максимума вспышки [Исаев и др., 2001, 2015]. Так, реализовавшейся в 1993 г. вспышке предшествовал подъем численности, начавшийся, по крайней мере, в 1990 г. [Кондаков, 2002].

Прямые наблюдения за популяциями *D. sibiricus* в условиях тайги затруднены, и при составлении прогнозов изменения его численности логично попытаться опереться на метеоданные. Однако единственный способ прогноза, основанный на использовании погодных условий [Кондаков, 1974], нестабилен [Тарасова, Волков, 2021]. Критерии для прогнозирования вспышек массового размножения предлагали и для других видов филлофагов [Leśniak, 1976; Пальникова и др., 2002; Ray et al., 2016; Möller et al., 2017; Bao et al., 2020; Li et al., 2020]. Наилучшие результаты были получены при использовании машинного обучения [Möller et al., 2017].

Машинное обучение определяется как систематическое обучение алгоритмов, в результате которого качество их работы возрастает по мере накопления опыта [Флах, 2015]. Использование этого подхода как инструмента прогнозирования крайне универсально и используется в таких разных отраслях, как, например, сейсмология [Sawant et al., 2023], медицина [Laios et al., 2020; Chen et al., 2022], энергетика [Gurgen, Dinh, 2022] и лесоводство [Peng, Karimi Sadaghiani, 2023]. Основной нашей задачей была разработка системы прогноза времени появления очагов массового размножения СШ с применением машинного обучения.

МАТЕРИАЛ И МЕТОДЫ

Район исследований. Для моделирования использовали данные о вспышках массового размножения в южно-таежных и подтаежных лесах Красноярского края, Томской и северной части Кемеровской области (рис. 1). В южной тайге значительные площади заняты древостоями с участием кормовых пород СШ – кедра сибирского *Pinus sibirica* Du Tour и пихты сибирской *Abies sibirica* Ledeb. [Леса СССР, 1969], встречаются такие древостои и в подтайге [Безруких и др., 2020]. Для этой территории типичны периодически возникающие вспышки массового размножения *D. sibiricus* (см. Данные о вспышках).

Климат достаточно, но не абсолютно однообразен. Среднегодовые температуры на уровне $-1,5...-0,6$ °С, средние температуры вегетационного периода $13,5-14,5$ °С. Безморозный период на западе длится 110–115 дней, на востоке – до 75 дней. Зимние температуры почвы достигают -8 °С. Количество осадков 400–500 мм в год. Континентальность климата растет к востоку [Леса СССР, 1969; Алпатыев и др., 1976]. Движение воздушных масс определяют преимущественно циклоны, приходящие с северного, южного и западного направлений, и отроги Азорского и Сибирского антициклонов. Увеличение температуры в целом определяется влиянием циклонов западного направления и антициклонов, рассматриваемых как отроги Азорского [Горбатенко и др., 2011; Ипполитов и др., 2011].

Биология сибирского шелкопряда. Гусеницы *D. sibiricus* лучше всего усваивают хвою лиственницы (*Larix* spp.), несколько хуже – *A. sibirica* и *P. sibirica* [Кириченко, Баранчиков, 2008]. В южной тайге вспышки происходят на пихте и кедре [Окунев, 1955; Коломиец, Майер, 1963; Кондаков, 2002], которые крайне неустойчивы к потере хвои [Журавлев, 1960; Кондаков, 1974; Гродницкий и др., 2002].

Имаго *D. sibiricus* не питаются [Рожков, 1963]. Лёт начинается с последней декады июня, пик – в середине июля, последние бабочки встречаются до начала августа [Прозоров, 1953; Рожков, 1963].

Сроки откладки яиц почти совпадают со сроками лёта [Рожков, 1963]. В лабораторном эксперименте показано, что эмбриональное развитие занимает от 10 дней до двух меся-

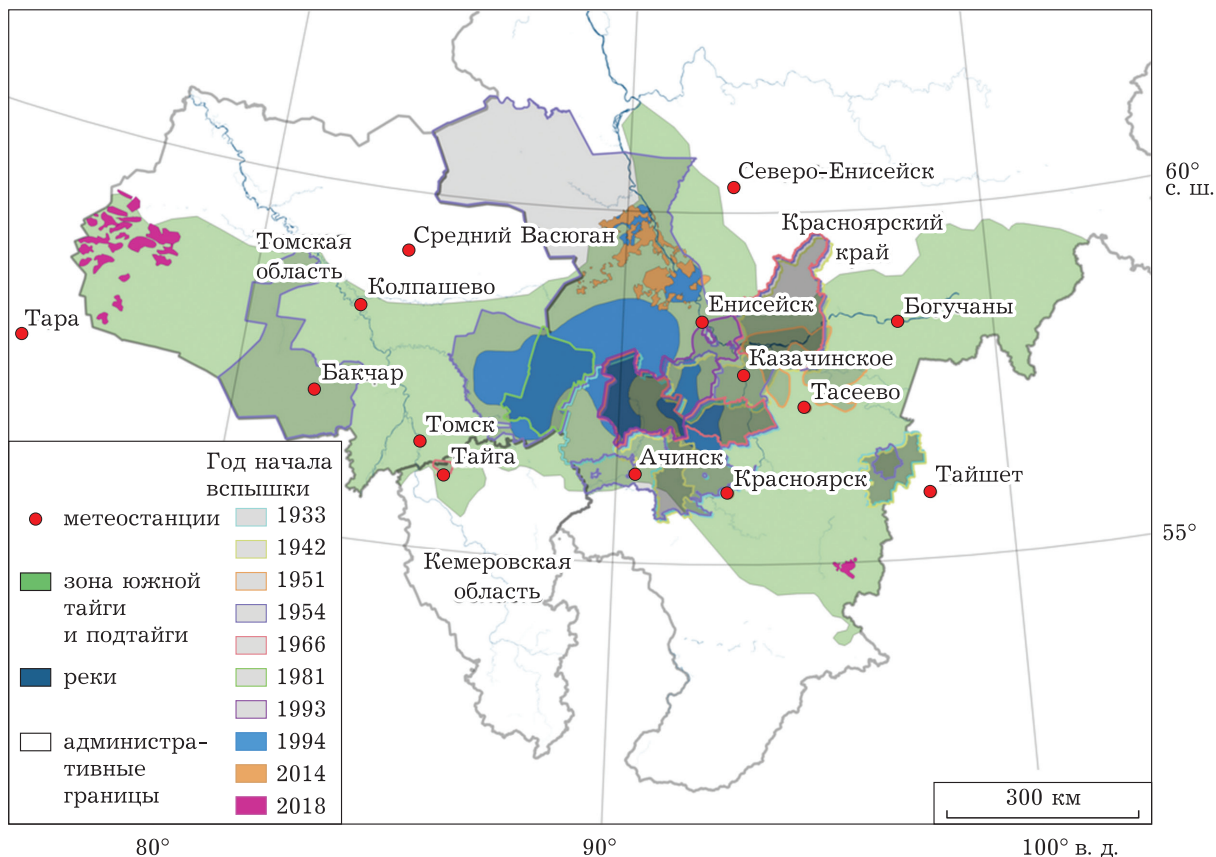


Рис. 1. Регион исследования, метеостанции и положение очагов массового размножения. Источники данных приведены в табл. 1

цев [Прозоров, 1953]. В природе же с откладки яиц до выхода гусениц в зависимости от локальных условий проходит от 13–17 [Прозоров, 1953] до 20–30 дней [Флоров, 1948].

Выход из яиц начинается с последних чисел июля, в массе – в первой половине августа. До окукливания проходит, как правило, шесть возрастов [Прозоров, 1953; Рожков, 1963]. Зимуют гусеницы под лесной подстилкой или мхом, спускаясь туда из крон очень поздно, порой уже после установления снежного покрова [Рожков, 1963]. Гусеницы на зимовке могут переносить температуры до $-16\text{ }^{\circ}\text{C}$ [Bulakhova et al., 2022], но чувствительны к переохлаждению и вымоканию [Костин, 1958; Коников и др., 1959; Рожков, 1963]. Подъем их в кроны начинается в конце апреля – начале мая, массовый – с первой половины мая. При $26\text{ }^{\circ}\text{C}$ они прекращают питание, а при $35\text{ }^{\circ}\text{C}$ наступает ведущий к гибели тепловой шок [Рожков, 1963]. Такая реакция на перегрев [Коников и др., 1959], ве-

роятно, приводит к наступлению летней диапаузы и, возможно, определяет длительность стадии личинки [Коников и др., 1965].

Как правило, для завершения развития гусеницам СШ нужно два года (две зимовки). В благоприятных условиях они окукливаются, перезимовав лишь однажды, а в неблагоприятных количество зимовок увеличивается до трех [Рожков, 1963]. По-видимому, условия жизни этой наиболее продолжительной стадии развития и определяют способность популяции СШ к качественному переходу от стабильного состояния к массовому размножению.

Подъем численности *D. sibiricus* до вспешного уровня связан с сильным воздействием модифицирующих факторов [Исаев и др., 2001], возможно, летней засухи [Исаев и др., 2015; Кондаков, 2002], а также температуры и осадков [Исаев и др., 2015]. Вероятно, благоприятные погодные условия приводят к массовому переходу гусениц на однолетний цикл развития [Кондаков, 2002; Исаев и др., 2015]

и уходу популяции из-под контроля энтомофагов [Исаев и др., 2001].

Данные о вспышках. Первая известная по литературным данным вспышка массового размножения СШ в исследованном регионе датируется 1914 г. [Галкин, 1960; Коломиец, Майер, 1963]. В работе использованы только массовые размножения, проходившие в период с непрерывными рядами метеонаблюдений хотя бы для одной метеостанции (табл. 1). Начало первой из таких вспышек датируется 1933 г. [Галкин, 1960; Кондаков, 1974].

Точная локализация территории, на которой происходили вспышки, основанная на анализе данных дистанционного зондирования Земли, возможна только для двух последних из них (1993 и 2014 гг.) [Харук и др., 2002; Kharuk et al., 2018]. В остальных случаях вспышки локализовывались по приведенным в литературных источниках названиям лесхозов, административных районов или рек, в бассейнах которых располагались очаги.

Для анализа связи погодных условий с началом массового размножения СШ выбирали по возможности такие метеостанции, которые находились на расстоянии не более 100 км. Если самая близкая метеостанция находилась на большем удалении, то для анализа выбирали несколько ближайших (максимальное расстояние от метеостанции до очага ~200 км).

Метеоданные и другие предикторы. Данные до 2021 г. о температуре, количестве осадков и высоте снежного покрова в суточ-

ном разрешении были получены из базы данных [ВНИИГМИ-МЦД, 2023] (табл. 2).

В качестве предикторов протестировано 112 характеристик погоды (табл. 3), рассчитанных на основе исходных данных. Дату начала и окончания периода с температурой выше 0,5 и 10 °С рассчитали по методу Педя [Педь, 1959]. Гидротермический коэффициент Селянинова (ГТК) рассчитывали по формуле [Vlăduț et al., 2017]:

$$\text{ГТК} = \frac{P}{\sum_{T > 10} T / 10}.$$

Также в качестве предикторов использовали широту, долготу станции и количество лет, прошедших с начала предыдущей вспышки. Первые два предиктора введены исходя из размеров исследуемой территории, где популяции *D. sibiricus* могли по-разному реагировать на отдельные метеоэлементы [Toigo et al., 2017; Суховольский и др., 2020]. Время после предыдущего массового размножения позволяло учитывать возможность нахождения популяции в стадии депрессии после предыдущей вспышки, когда подъем численности невозможен [Исаев и др., 2001].

Зависимой переменной было начало вспышки в году n или его отсутствие. Исходя из задачи заблаговременного прогнозирования массового размножения и из предположения, что триггером подъема численности являются погодные условия за несколько лет до его начала [Исаев и др., 2001], мы исполь-

Т а б л и ц а 1

Годы вспышек массового размножения СШ и метеостанции, из окрестностей которых известны очаги и для которых на начало вспышки были достаточно длинные ряды данных

Год начала вспышки	Метеостанции	Источник данных
1933	Енисейск, Красноярск	[Галкин, 1960; Кондаков, 1957]
1942	Ачинск, Богучаны, Енисейск, Красноярск	[Галкин, 1960; Кондаков, 1957; Прозоров, 1953]
1954	Ачинск, Бакчар, Богучаны, Енисейск, Колпашево, Красноярск, Средний Васюган, Тайга, Тайшет, Томск	[Галкин, 1961; Коломиец, Майер, 1963; Кондаков, 1974]
1966	Ачинск, Богучаны, Енисейск, Красноярск, Тайга, Томск	[Агеев, 2006; Кондаков, 1974]
1981	Ачинск, Тайга, Томск	[Kharuk et al., 2018]
1993	Ачинск, Богучаны, Енисейск, Казачинское	[Кондаков, 2002; Kharuk et al., 2018]
2014	Бакчар, Енисейск, Колпашево, Северо-Енисейский, Средний Васюган, Тайга, Тара, Томск	[Денисова и др., 2020; Kharuk et al., 2018]

Данные о метеостанциях, наблюдения которых использовали для прогнозирования

Метеостанция	Год начала наблюдений	Широта	Долгота
Ачинск	1929	56.28	90.52
Бакчар	1936	57.00	82.07
Богучаны	1930	58.38	97.45
Енисейск	1923	58.45	92.20
Казачинское	1959	57.68	93.28
Колпашево	1936	58.32	82.95
Красноярск	1925	56.00	92.80
Северо-Енисейск	1952	60.38	93.03
Средний Васюган	1929	59.22	78.23
Тайга	1936	56.07	85.26
Тайшет	1938	55.95	98.00
Тара	1936	56.90	74.38
Тасеево	1988	57.20	94.55
Томск	1926	54.45	84.97

зовали данные о погоде со сдвигом 4–7 лет. Также приходилось учитывать то, что ста-
дия нарастания численности может разли-

чатся по продолжительности [Исаев и др., 2001] (рис. 2). На практике это означает невоз-
можность составления прогноза с точностью

Характеристики погоды, изученные в качестве предикторов вспышек массового размножения СШ

Характеристика погоды	Период	Источник данных
Среднемесячная температура, °С	Январь – декабрь	[Флоров, 1948; Прозоров, 1953; Окунев, 1955; Аммосов, 1971; Кондаков, 2002; Kharuket et al., 2018]
Абсолютный минимум температуры за месяц, °С	Январь – декабрь	[Журавлев, 1960; Коломиец, 1960]
Абсолютный максимум температуры за месяц, °С	Январь – декабрь	[Коников и др., 1959; Bulakhova et al., 2022]
Время устойчивого перехода температуры через 0, 5 и 10 °С весной, день с начала года	Год	[Флоров, 1948; Прозоров, 1953]
Время устойчивого перехода температуры через 0, 5 и 10 °С осенью, день с начала года	Год	[Флоров, 1948; Прозоров, 1953; Журавлев, 1960]
Продолжительность периода между устойчивыми переходами температуры через 0, 5 и 10 °С, количество дней	Год	[Кондаков, 2002]
Сумма активных температур при пороге 0, 5 и 10 °С, °С	Март – октябрь	[Флоров, 1948; Окунев, 1955; Кондаков, 2002; Kharuk et al., 2018]
Сумма эффективных температур при пороге 5 и 10 °С, °С	Март – октябрь	[Прозоров, 1953; Кондаков, 2002]
Число дней с температурой выше 25, 30 и 35 °С	Год	[Коников и др., 1959]
Гидротермический коэффициент Селянинова	Июнь – август	[Окунев, 1955; Журавлев, 1960; Кондаков, 1974; Kharuk et al., 2018; Лямцев, 2019]
Сумма осадков за месяц, мм	Январь – декабрь	[Окунев, 1955; Коломиец, 1960; Чикидов, 2009; Kharuk et al., 2018]
Высота снежного покрова, см	Январь – май, октябрь – декабрь, гидрологический год	[Коломиец, 1960; Чикидов, 2009]

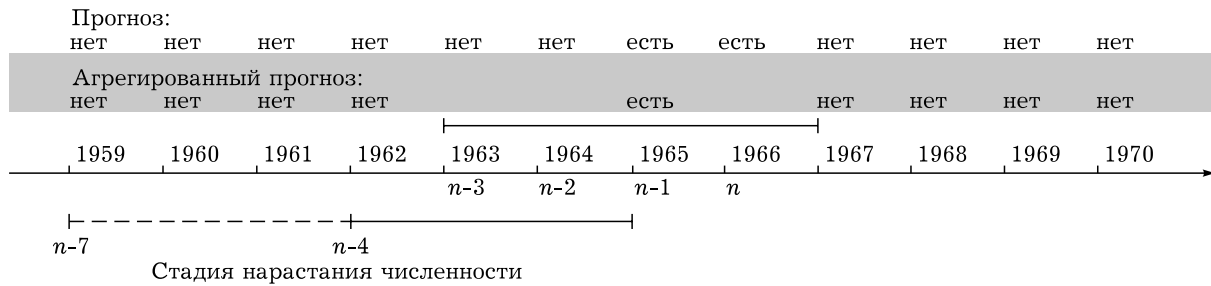


Рис. 2. Пример агрегации результатов прогноза (окрестности метеостанции Енисейск, массовое размножение, начавшееся в 1966 г.)

до года: в случае наступления благоприятных для массового размножения погодных условий первые признаки дефолиации могут проявиться раньше или позже. Наконец, зафиксированные в литературных источниках даты начала могли запаздывать по сравнению с реальными, особенно в случае наиболее ранних вспышек. Исходя из этих причин мы считали успешным прогноз в том случае, если он показывал начало вспышки не только в году n , но и в году $n - 1 \dots n - 3$ (см. рис. 2). Например, для начавшегося в 1954 г. массового размножения прогнозирование считали успешным и тогда, когда его начало предсказывалось для 1951–1953 гг.

С учетом сдвига данных количество предикторов было равно 451. Поэтому мы в два этапа выбрали наилучшие из них. Задачей первого этапа было уменьшение их количества для ускорения расчетов далее. Мы провели факторный анализ для смешанных данных (функция FAMD, пакет FactoMineR 2.7) [Lê et al., 2008] и выбрали предикторы, тесно связанные с теми же скрытыми переменными, что и зависимая. Второй этап проводили методом случайного леса (функция randomForest со сбалансированными выборками из пакета randomForest 4.7–1.1) [Liaw, Wiener, 2002], а по его результатам выбрали предикторы, исключение которых из анализа снижало коэффициент Джини на 0,4 и более, не имеющие тесной корреляции друг с другом, и значения которых для лет, предшествующих вспышкам массового размножения ($n - 4 \dots n - 7$), значимо при $p = 0,05$ (критерий Манна – Уитни) отличаются от остальных лет.

Моделирование. Моделирование проводили в среде Python 3.9.12. Объектами моделирования служили пары “год – метеостанция”. Каждая строка матрицы данных (всего

1052) содержала год, для которого проводили моделирование, код метеостанции, для которой были рассчитаны значения метеопараметров, значения выбранных на предыдущем этапе работы предикторов (т. е. метеозлементов для той же метеостанции за предшествующие годы) и значение булевой зависимой переменной. Погодную ситуацию, ведущую к началу массового размножения, условно называем “провоцирующей”, в противном случае – “спокойной”. Задачей классификации было построение модели, способной предсказать начало вспышки СШ в годы $n \dots n - 3$. Максимизируемой метрикой служила точность классификации, использование которой на предварительном этапе работы неожиданно показало лучший результат по сравнению с другими метриками.

Данные разделили на обучающую, валидационную и тестовую выборки в соотношении 60:20:20. Перед моделированием построили фиктивный классификатор, который относил каждый объект к одному из классов с учетом их частоты. Полученное для фиктивного классификатора значение точности далее использовали для оценки эффективности других моделей.

Мы протестировали три алгоритма классификации: градиентный бустинг (библиотеки XGBoost 1.7.3) [Chen, Guestrin, 2016] и LightGBM 3.3.5 [Ke et al., 2017]) и случайный лес (библиотека Scikit-learn 1.0.2 [Pedregosa et al., 2011]). В последнем случае мы отказались от алгоритма, реализованного в R [Liaw, Wiener, 2002], из-за более широкого набора используемых в Scikit-learn 1.0.2 гиперпараметров. Для оптимизации подбора значений гиперпараметров (табл. 4) использовали средства библиотеки Optuna 3.1.1 [Akiba et al., 2019]. Во всех случаях при обучении учитывали дисбаланс классов.

Гиперпараметры моделей, оптимизируемые в ходе их обучения, и значения, в пределах которых был проведен их подбор

Гиперпараметр	Случайный лес	XGBoost	LightGBM	Смысл гиперпараметра
n_estimators	10–500	10–500	10–500	Количество деревьев
criterion	entropy, gini			Функция оценки качества разбиения
max_depth	2–50	2–50	2–50	Максимальная глубина дерева решений
max_features	2–7			Количество признаков, используемых при построении дерева решений
max_samples, subsample	0,05–0,5	0,05–0,5	0,05–0,5	Доля объектов, используемых при построении дерева решений
class_weight	balanced, balanced_subsample			Метод балансировки классов
learning_rate		0,03–0,5	0,03–0,5	Скорость обучения модели

Для обученной модели рассчитывали точность по валидационной выборке. Модель, точность которой была наибольшей для всех классов, проверяли на тестовой выборке. Если результаты работы на валидационной и тестовой выборках отличались между собой незначительно, то констатировали хорошее качество обучения.

Работу лучшей модели проверяли на всем наборе данных. Вспышка в году n считалась успешно предсказанной, если модель хотя бы в одном случае показывала “провоцирующую” погодную ситуацию перед этим годом или перед $n \dots n - 3$ годами для данной метеостанции. Результаты для $n \dots n - 3$ лет для данной станции при оценке точности прогнозирования объединялись в одно наблюдение (см. рис. 2). Значение точности для таких агрегированных результатов рассчитывали отдельно.

Завершив обучение модели и проверку его качества, мы использовали ее для составления прогноза появления очагов массового размножения СШ в районе исследований для 2024–2026 гг.

Сочетания метеопараметров, которые были классифицированы выбранной моделью как “провоцирующие”, после нормализации разделили на однородные группы по методу Варда (библиотека SciPy) [Virtanen et al., 2020]. Значения отдельных метеоэлементов попарно сравнивали между выделенными кластерами и группой объектов со “спокойной” погодной ситуацией, используя критерий Манна – Уитни для сравнения центров распределения

и критерий Левене для сравнения отклонений от них.

Равномерность распределения метеостанций по кластерам исследовали с помощью критерия хи-квадрат.

РЕЗУЛЬТАТЫ

Выбор предикторов. По результатам факторного анализа отобрали 272 предиктора. Анализ методом случайного леса сократил их количество до 12. Из них критериям отсутствия корреляции с другими предикторами и наличия статистически значимых различий значений для “провоцирующих” и “спокойных” ситуаций отвечали восемь. Все они так или иначе описывали температурный режим за 4–5 лет до начала массового размножения: сумма активных температур мая (t.0.act.May.5) и июня (t.0.act.Jun.5) при пороге 0 °С, максимальная температура августа (t.max.Aug.5), минимальная температура сентября (t.min.Sep.5) и количество дней в году с температурой 25 °С и более (t.25.num.5) за пять лет до начала вспышки, сумма активных температур июня при пороге 0 °С (t.0.act.Jun.4), средние температуры сентября (t.mean.Sep.4), количество дней в году с температурой 25 °С и более (t.25.num.4) за четыре года до начала вспышки.

Выбор наилучшей модели. Точность классификации с помощью фиктивной модели, построенной для решения основной задачи, была равна 0,686 для валидационной выборки.

Вклад предикторов в модель XGBoost с наилучшими гиперпараметрами

Предиктор	Важность предиктора
t.25.num.5	0,138
t.0.act.May.5	0,137
t.0.act.Jun.4	0,137
t.max.Aug.5	0,133
t.min.Sep.5	0,126
t.0.act.Jun.5	0,126
t.25.num.4	0,115
t.mean.Sep.4	0,087

Модели, основанные на других алгоритмах, показали на этой выборке значительно лучший результат. Лучше всего себя показала модель, основанная на XGBoost (точность 0,876, learning_rate = 0,118, max_depth = 8, n_estimators = 458, subsample = 0,454), которую и использовали в дальнейшей работе. На тестовой выборке точность классификации для модели XGBoost составила 0,896. Наиболее важными для работы модели оказались предикторы, характеризующие погодную ситуацию за 5 лет до начала массового размножения (табл. 5).

Значение точности для результатов применения лучшей модели ко всему набору данных (табл. 6) после агрегации (см. абзац “Метеоданные и другие предикторы”) было равно 0,982.

Очень важно, что в большинстве случаев (24 из 33 истинно положительных прогнозов) положительные прогнозы идут друг за другом серией из двух-четырех, иногда с годичным разрывом между ними. Так, начавшейся в 1993 г. вспышке в окрестностях Енисейска предшествовала “провоцирующая” погодная ситуация в 1985–1986, 1986–1987, 1987–1988

и 1988–1989 гг. Еще в одном случае (окрестности Красноярска, вспышка 1933 г.) мы не можем подтвердить или опровергнуть наличие такой серии, так как в нашем распоряжении имеются непрерывные результаты наблюдений только с 1928 г. Лишь для восьми пар “год – метеостанция” перед началом массового размножения была зафиксирована только одна “провоцирующая” ситуация.

Из 13 ложноположительных прогнозов серия из двух “провоцирующих” ситуаций подряд наблюдалась только в одном случае (Колпашево, 1960–1961 и 1961–1962 г.). При этом восемь таких прогнозов были сделаны для тех лет, в которые массовое размножение произошло в других, удаленных от данной метеостанции регионах. Так, согласно результатам работы модели погодные условия благоприятствовали массовому размножению в 1941 г. в Томской области (по метеостанциям Томск и Средний Васюган). Однако эта вспышка, начавшаяся в 1942 г., прошла на территории Красноярского края (приблизительно на 400 км и более к востоку от Томска), где она была предсказана по данным всех четырех работавших на тот момент метеостанций (см. рис. 1).

Результат применения модели XGBoost с наилучшими гиперпараметрами к данным о времени начала массовых размножений СШ в районе исследований

		Прогноз начала вспышки	
		отрицательный	положительный
Начало вспышки в данном году возле данной станции	Не было	895	13
	Было	4	33

Характеристики погоды, типичные для “провоцирующих” и “спокойных” ситуаций (средние ± стандартные отклонения)

Станции	t.0.act.May.5	t.0.act.Jun.5	t.max.Aug.5	t.min.Sep.5	t.25.num.5	t.0.act.Jun.4	t.mean.Sep.4	t.25.num.4
“Провоцирующие”	249,9±45,7	443,0±57,9	27,4±2,1	-3,0±1,8	27,1±10,1	442,0±52,3	8,6±1,3	27,7±10,5
“Спокойные”	266,1±76,2	469,6±63,1	28,5±2,5	-3,7±2,5	33,9±11,7	470,8±64,3	8,2±2,0	34,1±11,6

Оптимистичные результаты работы лучшей модели с ретроспективными данными позволили нам попытаться дать прогноз появления очагов массового размножения СШ на ближайшие несколько лет. Согласно результатам работы модели, на всей исследованной территории вспышек в 2024–2026 гг. не ожидается.

Характеристики погоды, типичные для “провоцирующих” и “спокойных” ситуаций. Сравнение характеристик погоды в “провоцирующих” и “спокойных” ситуациях показало, что при $p \leq 0,05$ в среднем первые отличается более холодные май, летние месяцы и вегетационный сезон в целом, но при этом более теплый сентябрь (табл. 7). Однако кластерный анализ по методу Варда показал, что “провоцирующие” погодные ситуации разделяются на три хорошо обособленные группы (0, 1 и 2, соответственно 33, 41 и 19 объектов в каждой) (рис. 3).

Погодные ситуации, отнесенные к кластеру 0, от других кластеров отличают сравнительно теплые май и июнь за 5 лет до вспышки (рис. 3, а, б). Сентябрь за 5 лет и за 4 года до начала массового размножения также теплый (рис. 3, ж, з). При этом, однако, максимальные температуры августа, сумма активных температур июня и количество дней с температурами выше 25 °С в этом кластере невелики (рис. 3, в, д, е, з).

Погодная ситуация кластера 1 может быть охарактеризована как самая холодная (см. рис. 3). В определенной степени исключение представляет сумма активных температур июня, но и ее величина превосходит только таковую для кластера 0, будучи ниже по сравнению с кластером 2 и “спокойными” годами (см. рис. 3, е).

Май кластера 2 по сумме активных температур сопоставим со “спокойными” годами, на 23,8 °С уступая по этому показателю кластеру 0 ($p = 0,036$) (см. рис. 3, а). Лето и сезон за 5 лет до массового размножения в целом

холодные, сентябрь этого года теплый (см. рис. 3, б–д). В отличие от двух других кластеров здесь четвертый год перед массовым размножением теплый: июнь и сезон в целом сопоставимы со “спокойными” годами, средние температуры сентября выше и не отличаются значимо от температур кластера 0 (см. рис. 3, е–з).

В целом же значения метеоэлементов из разных кластеров со “спокойными” годами интерпретируются сходно с результатами, приведенными в табл. 7. За редкими исключениями сентябрь “провоцирующих” ситуаций из любого кластера в среднем теплее по сравнению со “спокойными” (см. рис. 3, г, ж), а весна, лето и сезон в целом холоднее (см. рис. 3, а–в, д, е, з).

Кроме того, следует отметить, что температурные характеристики мая и летних месяцев “спокойных” лет показывают значительно больший разброс по сравнению с “провоцирующими”, далеко отклоняясь от средних значений в область как минимума, так и максимума (см. рис. 3, а–в, е). Для температурных характеристик сентября в “спокойной” ситуации характерен хвост значений в область минимума (см. рис. 3, г, ж), для количества дней с температурой выше 25 °С – напротив, в область максимума (см. рис. 3, д, з). То же подтверждено и сравнением по критерию Левене (табл. 8): разброс значений предикторов в “провоцирующих” ситуациях практически всегда значимо ниже.

Проверка того, насколько равномерно метеостанции распределены по кластерам, показала отсутствие закономерностей ($p = 0,84$, критерий хи-квадрат).

ОБСУЖДЕНИЕ

Наиболее важным прикладным результатом является то, что с помощью машинного обучения возможно создать модель прогноза подъема численности СШ за несколько

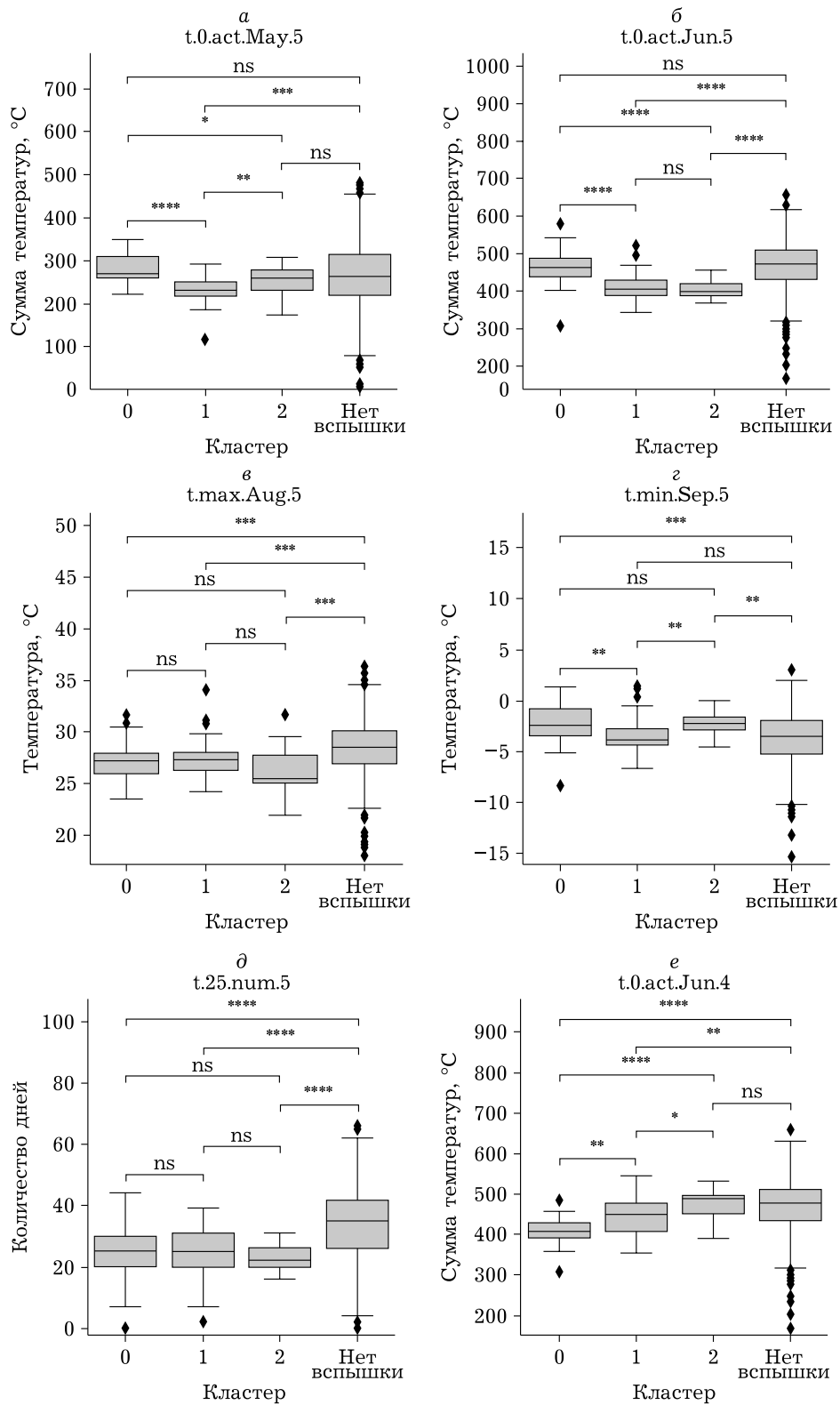


Рис. 3 (начало)

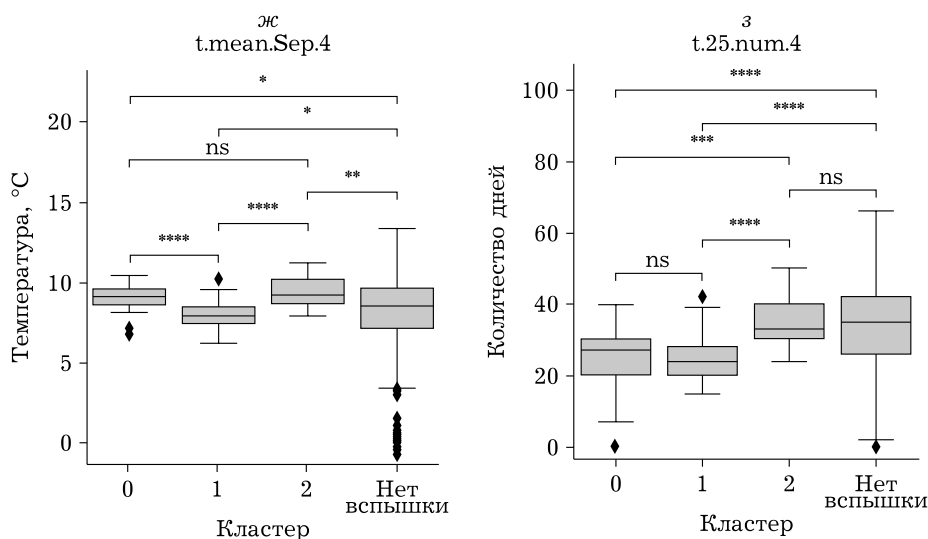


Рис. 3 (окончание). Значения метеопараметров для выделенных кластеров и для лет со “спокойной” ситуацией. Статистическая значимость различий по критерию Манна – Уитни: ns – различия незначимы, * – значимы на уровне 0,05, ** – на уровне 0,01, *** – на уровне 0,001, **** – на уровне 0,0001

лет до нанесения им ущерба. Важно, однако, оценить, насколько эта модель обоснована биологически.

Предположение о том, что в жизненном цикле насекомых имеется какая-либо ключевая стадия, воздействие на которую единственного модифицирующего фактора определяет переход от стабильного состояния к вспышечному [Болдаруев, 1960; Кондаков, 1974; Пальникова и др., 2002; Rouault et al., 2006], по-видимому, зачастую ошибочно. Согласно закону толерантности Шелфорда, благополучие вида определяется комплексом факторов, и повлиять на состояние популяции может отклонение от оптимума как одного из них, так и их комбинации [Shelford, 1911]. Из всего разнообразия потенциальных предикторов (см. табл. 3) с появлением оча-

гов массового размножения *D. sibiricus* тесно связаны те, которые описывают температурные условия во время периода активной жизнедеятельности этого филлофага (см. “Выбор предикторов”). Аналогичный вывод о преимущественной роли температуры в качестве модифицирующего фактора был сделан и ранее [Болдаруев, 1960].

То, что на начало вспышек лесных филлофагов и *D. sibiricus*, в частности, влияет комплекс характеристик погоды, предполагали давно [Коломиец, 1960; Аммосов, 1971; Leśniak, 1976]. В последнее время это было показано для разных видов неоднократно и убедительно [Ray et al., 2016; Möller et al., 2017; Bao et al., 2020; Demidko et al., 2022]. То, что все выбранные предикторы описывают температурные условия (см. табл. 5), обеспечивает

Т а б л и ц а 8

Стандартные отклонения значений метеопараметров для “спокойных” ситуаций и кластеров “провоцирующих”

Станции	t.10.act.May.5	t.0.act.Jun.5	t.max.Aug.5	t.min.Sep.5	t.25.num.5	t.0.act.Jun.4	t.mean.Sep.4	t.25.num.4
“Спокойные”	75,4	63,1	2,53	2,45	11,6	64,3	1,996	11,8
Кластер 0	33,3***	48,0#	1,77*	1,75#	9,58	35,1***	0,763***	8,91*
Кластер 1	29,1***	37,3***	1,86**	1,86**	8,67*	45,7	0,764***	6,87***
Кластер 2	34,8**	21,8***	2,42	1,06***	4,55**	36,7**	0,916**	7,38*

П р и м е ч а н и е. Уровни значимости различий по критерию Левене между “спокойными” и “провоцирующими” ситуациями: *** – 0,001, ** – 0,01, * – 0,05, # – 0,1.

надежность результатов моделирования в условиях редкой сети метеостанций и значительных расстояний от них до очагов массового размножения. Ранее было показано, что в условиях региона исследований изменения температур с мая по сентябрь тесно коррелированы на расстоянии по крайней мере 300 км [Исаев и др., 2015], что намного больше расстояний от очагов до метеостанций, характеризующих погоду в них (см. рис. 1, табл. 1).

Зависимость скорости развития от температуры имеет очень близкую форму для всех пойкилотермных животных [Детлаф, 2001], что показано, в частности, для СШ на стадии яйца и куколки [Прозоров, 1953]. Увеличение температуры вблизи от нижнего предела толерантности ведет к быстрому увеличению скорости развития, но затем ослабевает вплоть до замедления развития [Детлаф, 2001]. В мае, однако, температуры далеки от верхнего предела толерантности для гусениц *D. sibiricus*, равного 35 °С [Рожков, 1963], и влияние, даже статистически значимое, снижения на 36,6 °С (13 %) суммы активных температур мая с лагом 5 лет (см. рис. 3, а) для кластера 1 по сравнению со “спокойными” ситуациями выглядит контринтуитивным. Более того, этот факт противоречит наблюдениям, согласно которым рост температуры благоприятствует подъему гусениц в кроны [Рожков, 1963; Кондаков, 2002].

Мы не обнаружили в известной нам литературе данных о суммах активных температур, достаточных для выхода гусениц СШ с зимовки. Однако обработка результатов наблюдений из работы [Прозоров, 1953] позволяют утверждать, что эта сумма ≈65 °С. Между тем даже минимальные значения $t.0.act.May.5$ для кластера 1 превышают 100 °С (см. рис. 3, а). И если изменения сумм активных майских температур не улучшают условия для популяции *D. sibiricus* в “провоцирующих” ситуациях, то, во всяком случае, не ухудшают.

Таким образом, сам по себе сравнительно холодный май не оказывает какого-либо влияния на популяцию СШ, а более низкие суммы его активных температур – всего лишь следствие типичной для ситуации из кластера 1 прохладной погоды сезона в целом. Это косвенно подтверждается результатами моделирования численности *D. sibiricus*, пока-

завшими положительное влияние на его численность теплой погоды мая текущего сезона [Суховольский и др., 2020].

Наши результаты предположительно связаны с особенностями атмосферной циркуляции [Bai et al., 2021; Camarero et al., 2022]. Указание на отсутствие его влияния на колебания численности чешуекрылых [Westgarth-Smith et al., 2012] объясняется тем, что в данное исследование были включены только индифферентные виды, на динамику численности которых погода вообще влияет слабо [Исаев и др., 2001]. Понижение температуры мая–сентября в Сибири в значительной степени определяется увеличением Скандинавского индекса, которое вызывает ослабление западного переноса воздушных масс [Ипполитов и др., 2011]. Действительно, количество дней теплого сезона года с западным переносом в 1927–1966 гг. [Батталов, 1968, приложение 10] при сдвиге в 5 лет в среднем выше для “спокойных” ситуаций и сравнительно тепло-го кластера 0, хотя из-за небольшого размера выборок статистическую значимость этих различий подтвердить не удалось.

Менее неожидан факт, что подъему численности СШ способствует прохладная погода июня и августа (см. рис. 3, б, в, е), а также не слишком большое количество жарких дней в сезоне в целом (см. рис. 3, д, з), особенно явно при сдвиге в 5 лет (см. рис. 3, б, в, д). Значения метеопараметров, благоприятные для роста численности популяций *D. sibiricus* (кластеры 0–2), находятся внутри узкого по сравнению со “спокойной” ситуацией коридора (см. табл. 8). Сходная реакция на амплитуду колебаний температур показана для *Diprion pini*: его вспышкам массового размножения благоприятствуют низкие значения стандартных отклонений температур [Möller et al., 2017].

Тяготение к небольшому диапазону значений температуры согласуется с законом толерантности Шелфорда [Shelford, 1911], согласно которому неблагоприятно отклонение от оптимума в любую сторону. Интересно то, что средние значения условий, оптимальных для развития вспышки, несколько ниже тех, которые ему не способствуют. Объяснить это можно тем, что слишком сильные отклонения от оптимума вверх приводят к прекращению питания гусениц [Рожков, 1963], уходу их

в летнюю диапаузу и развитию по двухлетнему циклу [Коников и др., 1959; Коников и др., 1965; Баранчиков и др., 2010]. В случае же высоких максимальных температур августа возможно и увеличение смертности молодых гусениц [Rouault et al., 2006]. Подобная реакция на излишне высокие температуры известна и для других видов насекомых [Саулич, Мусолин, 2017]. Неблагоприятное влияние высоких температур августа на развитие яиц СШ маловероятно. Верхний предел толерантности этой стадии равен 35 °С [Прозоров, 1953] и почти никогда в природе не достигается.

Вместе с тем имеются данные, не согласующиеся с этим предположением. Массовому размножению *D. sibiricus* в Прибайкалье предшествовали сезоны с повышенной температурой мая и августа (сдвиг 1–3 года). Более того, превышение средних температур августа над среднегодовыми значениями не помешало подавляющему большинству гусениц СШ завершить развитие по однолетнему циклу [Болдаруев, 1960]. Но остроту этого противоречия сложно оценить, не зная количества дней с температурами, выше критических. Также можно предположить, что реакция *D. sibiricus* на погодные условия в условиях более континентального климата Прибайкалья с коротким вегетационным периодом (см. “Район исследований”) может отличаться от описанной здесь.

Наиболее очевидно объяснение положительного влияния теплой сентябрьской погоды на подъем численности *D. sibiricus* (см. рис. 3, 2, ж). Более высокие температуры сентября позволяют продлить период активности гусениц [Рожков, 1963], что ускоряет их развитие и позволяет перейти к одногодичному жизненному циклу. Характерно, что при сдвиге в 5 лет значение в качестве предиктора имеют именно минимальные суточные температуры (см. рис. 3, 2): это следует понимать как указание на риск гибели от раннеосенних заморозков многих молодых гусениц. То, что в сентябре наибольшее значение имеет отклонение от оптимума (см. табл. 8) в сторону нижнего предела толерантности, подтверждается анализом эмпирического распределения (см. рис. 3, 2, ж).

Наконец, нужно иметь в виду, что такие условия, в целом представляющие сочетание прохладного (но не холодного) лета и те-

плого сентября, повторяются на протяжении развития двух или более смежных поколений СШ. Следствием этого является ускорение развития за счет избегания летней диапаузы и более продолжительного питания перед зимовкой. В результате незадолго до начала вспышки два поколения *D. sibiricus* перед вспышкой проходят развитие преимущественно по годичному циклу [Кондаков, 1957, 2002], уходят из-под контроля паразитоидов [Исаев и др., 2001] и наносят ущерб южно-таежным и подтаежным древостоям.

При рассмотрении полученных результатов следует избегать интерпретации, согласно которой численность СШ в районе исследований модифицируют лишь выбранные нами для моделирования предикторы. Их использование лишь позволяет составить наиболее точный прогноз, но влияние на популяцию *D. sibiricus*, несомненно, оказывают и другие элементы погоды. Желательно в дальнейшем обсудить вклад влагообеспеченности [Кондаков, 1974; Leśniak, 1976; Кондаков, 2002; Haynes et al., 2014; Bao et al., 2020].

Кроме того, популяции лесных фитофагов, включая массово размножающиеся виды, функционируют в системе тритрофа, которая включает взаимодействующие популяции самих фитофагов, их естественных врагов и кормовых растений [Abdala-Roberts et al., 2019]. Помимо очевидного предположения о непосредственном влиянии погоды на фитофага, приходится также допускать возможность косвенного воздействия, реализуемого через другие компоненты тритрофа [Barbour, 1988; Kendall et al., 2005].

ЗАКЛЮЧЕНИЕ

Построение хорошо работающей модели заблаговременного предсказания массовых размножений сибирского шелкопряда, основанной на машинном обучении, вполне возможно. В нашем случае предикторами служили температурные характеристики мая–сентября за 4–5 лет до зарегистрированного начала вспышки. Модель градиентного бустинга, основанная на алгоритме XGBoost, давала точность прогноза 0,952. Основными биологическими механизмами, лежащими в основе этой модели, мы предполагаем изменение скорости развития гусениц из-за

предотвращения ухода в летнюю диапаузу и удлинения сезона питания осенью. В целом, температуры, благоприятные для реализации вспышки, находятся в более узких лимитах по сравнению с годами, после которых массовое размножение не реализовывалось. Также желательна рассмотреть возможность влияния температурных характеристик среды через другие элементы тритрофа.

Благодарности

Авторы выражают особую благодарность Красноярскому региональному центру коллективного пользования ФИЦ КНЦ СО РАН за предоставленное оборудование для обеспечения выполнения проекта “Фундаментальные основы защиты лесов от энтомо- и фитовредителей в Сибири” (№ FEFE-2020-0014).

Авторы благодарят директора Центра защиты леса Красноярского края В. В. Солдатову за предоставленный материал.

Авторы признательны своим коллегам Д. М. Данилиной и Д. И. Назимовой за ценные советы при написании работы.

Вклад авторов

Демидко Д. А. – концептуализация, хранение и обслуживание данных, формальный анализ, проведение исследования, методология, программное обеспечение, визуализация, создание черновика рукописи; Горошко А. А. – формальный анализ, валидация данных, визуализация, создание рукописи и ее редактирование; Сультсон С. М. – администрирование проекта, ресурсы, создание рукописи и ее редактирование; Кулакова Н. Н. – хранение и обслуживание данных, создание рукописи и ее редактирование; Михайлов П. В. – концептуализация, получение финансирования, ресурсы, руководство исследованием, валидация данных.

Финансирование

Работа выполнена в рамках государственного задания Минобрнауки России на выполнение коллективом научной лаборатории “Защита леса” проекта “Фундаментальные основы защиты лесов от энтомо- и фитовредителей в Сибири” (№ FEFE-2020-0014).

Соблюдение этических стандартов

В данной работе отсутствуют исследования человека или животных.

Конфликт интересов

Авторы данной работы заявляют, что у них нет конфликта интересов.

ЛИТЕРАТУРА

- Агеев А. А. Периодичность вспышек массового размножения сибирского шелкопряда и их влияние на формирование лесов Чулымско-Кетского междуречья: автореф. дис. ... канд. с.-х. наук. Красноярск, 2006. 21 с.
- Алпатьев А. М., Архангельский А. М., Подоплелов Н. Я., Степанов А. Я. Физическая география СССР (азиатская часть). М.: Высш. шк., 1976. 360 с.
- Аммосов Ю. Н. К вопросу о массовом размножении сибирского шелкопряда (*Dendrolimus superans sibiricus* Tshetv.) в Центральной Якутии // Проблемы защиты таежных лесов. Красноярск, 1971. С. 7–8.
- Баранчиков Ю. Н., Петько В. М., Бабичев Н. С. Пространственно-временная динамика возрастной структуры популяции сибирского шелкопряда в дефолированном лиственничнике // Энтомологические исследования в Северной Азии. М.: Т-во науч. изд. КМК, 2010. С. 239–240.
- Батталов Ф. З. Многолетние колебания атмосферных осадков и вычисление норм осадков. Л.: Гидрометеоиздат, 1968. 184 с.
- Безруких В. А., Хилиманюк А. А., Авдеева Е. В., Коротков А. А., Лигаева Н. А. Своеобразие растительности подтайги и лесостепей бореальной зоны Сибири // Хвойные бореальной зоны. 2020. Т. 38, № 3–4. С. 97–101.
- Болдаруев В. О. Итоги и перспективы изучения и истребления сибирского шелкопряда в Восточной Сибири // Материалы по проблеме сибирского шелкопряда. Новосибирск: Изд-во СО АН СССР, 1960. С. 3–10.
- ВНИИГМИ-МЦД. Всероссийский научно-исследовательский институт гидрометеорологической информации – мировой центр данных. URL: <http://meteo.ru/> (дата обращения: 23.09.2023).
- Галкин Г. И. Некоторые вопросы формирования резерваций и первичных очагов сибирского шелкопряда в лесах Красноярского края // Сибирский шелкопряд. Новосибирск: Изд-во СО АН СССР, 1960. С. 21–33.
- Галкин Г. И. Сибирский шелкопряд в лиственничных лесах Красноярского края // Лиственница и ее использование в народном хозяйстве СССР. М.: Центральное бюро технической информации лесной промышленности, 1961. С. 66–72.
- Горбатенко В. П., Ипполитов И. И., Логинов С. В., Поднебесных Н. В., Харюткина Е. В. Роль циркуляционных факторов в потеплении климата Сибири // Вестн. Том. гос. ун-та. 2011. № 346. С. 174–180.
- Гродницкий Д. Л., Разнобарский В. Г., Солдатов В. В., Ремарчук Н. П. Деградация древостоев в таежных шелкопрядниках // Сиб. экол. журн. 2002. Т. 9, № 1. Приложение. С. 3–12.
- Денисова Н. Б., Соболев А. А., Шипинская У. С. Результаты обследования очагов сибирского шелкопряда (*Dendrolimus sibiricus* Tschetv.) на территории Васюганского лесничества Томской области // Лесн. вестн. 2020. Т. 24, № 6. С. 65–72.
- Детлаф Т. А. Температурно-временные закономерности развития пойкилотермных животных. М.: Наука, 2001. 2011 с.
- Журавлев Г. П. Рекомендации по надзору за сибирским шелкопрядом в лесах Дальнего Востока. Хабаровск: ДальНИИЛХ, 1960. 33 с.
- Ипполитов И. И., Кабанов М. В., Логинов С. В., Поднебесных Н. В., Харюткина Е. В., Горбатенко В. П. Влияние атмосферной циркуляции на температурный ре-

- жим Сибири // Оптика атмосферы и океана. 2011. Т. 24, № 1. С. 15–21.
- Исаев А. С., Пальникова Е. Н., Суховольский В. Г., Тарасова О. В. Динамика численности лесных насекомых-филлофагов: модели и прогнозы. М.: Т-во науч. изд. КМК, 2015. 261 с.
- Исаев А. С., Хлебопрос Р. Г., Недорезов Л. В., Кондаков Ю. П., Киселев В. В., Суховольский В. Г. Популяционная динамика лесных насекомых. М.: Наука, 2001. 373 с.
- Кириченко Н. И., Баранчиков Ю. Н. Кормовые нормы гусениц сибирского шелкопряда на хвойных породах Сибири // Сиб. экол. журн. 2008. Т. 15, № 5. С. 709–716. [Kirichenko N. I., Baranchikov Yu. N. Withdrawn Food Rate for the Larvae of Siberian Moth on the Conifers of Siberia // Contemporary Problems of Ecology. 2008. Vol. 1, N 5. P. 709–716.]
- Коломиец Н. Г. Изучение сибирского шелкопряда в Западной Сибири и перспективы использования его паразитов // Сибирский шелкопряд. Новосибирск: Изд-во СО АН СССР, 1960. С. 61–73.
- Коломиец Н. Г., Майер Э. И. Важнейшие вредители лесов Томской области и меры борьбы с ними. Томск: Том. кн. изд-во, 1963. 35 с.
- Кондаков Ю. П. К вопросу биоэкологии сибирского шелкопряда (*Dendrolimus sibiricus* Tshtv) в пихтовых лесах Красноярского края // Уч. зап. Краснояр. гос. пед. ин-та. Красноярск, 1957. С. 144–153.
- Кондаков Ю. П. Закономерности массовых размножений сибирского шелкопряда // Экология популяций лесных животных Сибири. Новосибирск: Наука. Сиб. отд-ние, 1974. С. 206–265.
- Кондаков Ю. П. Массовые размножения сибирского шелкопряда в лесах Красноярского края // Энтомологические исследования в Сибири. Красноярск: КФ РЭО, 2002. С. 25–74.
- Конииков А. С., Платонова-Чернышева Л. В., Кондаков Ю. П., Зайцева А. И. Адаптация сибирского шелкопряда к условиям внешней среды. К характеристике факторов, определяющих численность сибирского шелкопряда // Уч. зап. Краснояр. пед. ин-та. Красноярск, 1959. С. 145–175.
- Конииков А. С., Платонова-Чернышева Л. В., Михайлова А. М. Значение диапаузы в пищевой адаптации гусениц сибирского шелкопряда // Исследования по защите лесов Сибири. М.: Наука, 1965. С. 51–52.
- Костин И. А. О вспышке массового размножения сибирского шелкопряда в горных лесах Восточного Казахстана // Тр. Ин-та зоологии АН КазССР. 1958. Т. VIII. С. 122–126.
- Леса СССР. Леса Урала, Сибири и Дальнего Востока / под ред. А. Б. Жукова. М.: Наука, 1969. 768 с.
- Лямцев Н. И. Очаги размножения непарного шелкопряда в Европейской России // Лесн. вестн. 2014. № 6. С. 78–85.
- Лямцев Н. И. Оценка и прогноз угрозы массового размножения сибирского шелкопряда в лесах Красноярского края // Изв. С.-Петербург. лесотехнической академии. СПб.: СПбГЛТА, 2019. С. 294–311.
- Окунев П. П. Географическое положение и зоны вредности сибирского шелкопряда // Геогр. сб. V. Географические вопросы лесного хозяйства. М. – Л.: Изд-во АН СССР, 1955. С. 210–222.
- Пальникова Е. Н., Свицерская И. В., Суховольский В. Г. Сосновая пяденица в лесах Сибири: экология, динамика численности, влияние на насаждения. Новосибирск: Наука, 2002. 231 с.
- Педь Д. А. Об определении дат устойчивого перехода температуры воздуха через определенные значения // Метеорология и гидрология. 1959. № 10. С. 38–39.
- Прозоров С. С. Сибирский шелкопряд в пихтовых лесах Сибири // Труды СибЛТИ. Красноярск: Краснояр. краевое изд-во, 1953. С. 93–132.
- Рожков А. С. Сибирский шелкопряд: Систематическое положение, филогения, распространение, экономическое значение, строение и образ жизни. М.: Изд-во АН СССР, 1963. 176 с.
- Рожков А. С. Массовое размножение сибирского шелкопряда и меры борьбы с ним. М.: Наука, 1965. 178 с.
- Саулич А. Х., Мусолин Д. Л. Летняя диапауза как особая сезонная адаптация насекомых: разнообразие форм проявления, механизмов контроля и экологическое значение // Энтотомол. обозрение. 2017. Т. 96, № 4. С. 665–703. [Saulich A. Kh., Musolin D. L. Summer diapause as a special seasonal adaptation in insects: diversity of forms, control mechanisms and ecological importance // Entomol. Rev. 2017. Vol. 96, N 4. P. 665–703.]
- Суховольский В. Г., Тарасова О. В., Ковалев А. В. Моделирование критических явлений в популяциях лесных насекомых // Журн. общ. биологии. 2020. Т. 81, № 5. С. 374–386. [Sukhovolsky V. G., Tarasova O. V., Kovaliev A. V. Modeling critical events in forest insects populations // J. Gen. Biol. 2020. Vol. 81, N 5. P. 374–386.]
- Тарасова О. В., Волков В. Е. Влияние погодных условий на развитие вспышек массового размножения сибирского шелкопряда в Средней Сибири // Сиб. лесн. журн. 2021. № 5. С. 49–59.
- Флах П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных. М.: ДМК-Пресс, 2015. 400 с.
- Флоров Д. Н. Вредитель сибирских лесов (сибирский шелкопряд). Иркутск: ОГИЗ Иркут. обл. кн. изд-во, 1948. 132 с.
- Харук В. И., Рэнсон К. Дж., Кузьмичев В. В., Буренина Т. А., Тихомиров А. Ю., Им С. Т. Съёмка “Landsat” в анализе шелкопрядников Южной Сибири // Исследования Земли из космоса. 2002. № 4. С. 79–90.
- Чикидов И. И. Роль климатических факторов в развитии очагов массового размножения сибирского шелкопряда в Центральной Якутии в 1998–2001 годах // Вестн. ЯГУ. 2009. Т. 6, № 3. С. 8–12.
- Abdala-Roberts L., Puentes A., Finke D. L., Marquis R. J., Montserrat M., Poelman E. H., Rasmann S., Sentis A., van Dam N. M., Wimp G., Mooney K., Björkman C. Tri-trophic interactions: bridging species, communities and ecosystems // Ecol. Lett. 2019. Vol. 22, N 12. P. 2151–2167.
- Akiba T., Sano S., Yanase T., Ohta T., Koyama M. Optuna: A Next-generation Hyperparameter Optimization Framework // Proc. of the 25th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining. Anchorage AK USA: Association for Computing Machinery. N. Y., 2019. P. 2623–2631.
- Bai M., Yao Q., Camarero J., Hu H., Dong Zh., Li Y., Zhou F., Chen X., Guoyang G., Cao X., Fang K. El Niño-Southern Oscillation modulates insect outbreaks in humid subtropical China: Evidences from tree rings and carbon isotopes // Dendrochronologia. 2021. Vol. 66. P. 125815.

- Bao Y., Na L., Han A., Guna A., Wang F., Liu X., Zhang J., Wang Ch., Tong S., Bao Y. Drought drives the pine caterpillars (*Dendrolimus* spp.) outbreaks and their prediction under different RCPs scenarios: A case study of Shandong Province, China // *Forest Ecology and Management*. 2020. Vol. 475. P. 118446.
- Barbour D. A. The Pine Looper in Britain and Europe // *Dynamics of Forest Insect Populations. Population Ecology* / Ed. A. A. Berryman. Springer: Boston MA, 1988. P. 291–308.
- Bulakhova N., Zhigul'skaya Z., Gashkova L., Berman D. On cold hardiness of the egg parasitoid wasp *Telenomus tetratomus* (Thomson, 1861) (Hymenoptera, Scelionidae) – a population regulator of the Siberian moth // *J. Hymenoptera Res.* 2022. Vol. 91. P. 27–39.
- Camarero J., Tardif J., Gazol A., Conciatori F. Pine processionary moth outbreaks cause longer growth legacies than drought and are linked to the North Atlantic Oscillation // *Sci. Total Environment*. 2022. Vol. 819. P. 153041.
- Chen S.-D., You J., Yang X.-M., Gu H.-Q., Huang X.-Y., Liu H., Feng J.-F., Jiang Y., Wang Y.-J. Machine learning is an effective method to predict the 90-day prognosis of patients with transient ischemic attack and minor stroke // *BMC Med. Res. Methodol.* 2022. Vol. 22, N 1. P. 195.
- Chen T., Guestrin C. XGBoost: A Scalable Tree Boosting System // *Proc. of the 22nd ACM SIGKDD Inter. Conf. on Knowledge Discovery and Data Mining KDD '16*. New York, N. Y.: Association for Computing Machinery, 2016. P. 785–794.
- Demidko D., Sultson S., Mikhaylov P., Verkhovets S. Influence of Weather Conditions and Climate Oscillations on the Pine Looper *Bupalus piniaria* (L.) Outbreaks in the Forest-Steppe of the West Siberian Plain // *Forests*. 2022. Vol. 13, N 1. P. 15.
- Gurgen A., Dinh N. T. Development and assessment of a reactor system prognosis model with physics-guided machine learning // *Nuclear Engineering and Design*. 2022. Vol. 398. P. 111976.
- Haynes K. J., Allstadt A. J., Klimetzek D. Forest defoliator outbreaks under climate change: effects on the frequency and severity of outbreaks of five pine insect pests // *Glob Change Biol.* 2014. Vol. 20, N 6. P. 2004–2018.
- Ke G., Meng Q., Finley T., Wang T., Chen W., Ma W., Ye Q., Liu T.-Y. LightGBM: A Highly Efficient Gradient Boosting Decision Tree // *Advances in Neural Information Processing Systems: Curran Associates, Inc.*, 2017. P. 3146–3154.
- Kendall B. E., Ellner S. P., McCauley E., Wood S. N., Briggs C. J., Murdoch W. W., Turchin P. Population Cycles in the Pine Looper Moth: Dynamical Tests of Mechanistic Hypotheses // *Ecol. Monogr.* 2005. Vol. 75, N 2. P. 259–276.
- Kharuk V. I., Im S. T., Yagunov M. N. Migration of the Northern Boundary of the Siberian Silk Moth // *Contemp. Probl. Ecol.* 2018. Vol. 11, N 1. P. 26–34.
- Kononov A., Ustyantsev K., Wang B., Mastro V., Fet V., Blinov A., Baranchikov Yu. Genetic diversity among eight *Dendrolimus* species in Eurasia (Lepidoptera: Lasiocampidae) inferred from mitochondrial COI and COII, and nuclear ITS2 markers // *BMC Genet.* 2016. Vol. 17, N 3. P. 157.
- Laios A., Katsenou A., Tan Y., Otiy M., Hutson R., Kaufmann A., Thangavelu A., Nugent D., de Jong D. Two-year prognosis estimation of advanced high grade serous ovarian cancer patients using machine learning // *Int. J. Gynecol. Cancer*. 2020. Vol. 30, Suppl 4. A67.2-A68
- Lê S., Josse J., Husson F. FactoMineR: An R Package for Multivariate Analysis // *J. Stat. Sof.* 2008. Vol. 25. P. 1–18.
- Leśniak A. Climatic and Meteorological Conditions of the Pine Moth (*Dendrolimus pini* L.) Outbreaks // *Ecol. Polska*. 1976. Vol. 24, N 4. P. 515–547.
- Li M., MacLean D., Hennigar C., Ogilvie J. Previous year outbreak conditions and spring climate predict spruce budworm population changes in the following year // *Forest Ecol. and Management*. 2020. Vol. 458. P. 117737.
- Liaw A., Wiener M. Classification and Regression by randomForest // *R News*. 2002. N 2. P. 18–22.
- Möller K., Hentschel R., Wenning A., Schröder J. Improved Outbreak Prediction for Common Pine Sawfly (*Diprion pini* L.) by Analyzing Floating 'Climatic Windows' as Keys for Changes in Voltinism // *Forests*. 2017. Vol. 8, N 9. P. 319.
- Moulinier J., Lorenzetti F., Bergeron Y. Gap dynamics in aspen stands of the Clay Belt of northwestern Quebec following a forest tent caterpillar outbreak // *Can. J. For. Res.* 2011. Vol. 41, N 8. P. 1606–1617.
- Pedregosa F., Varoquaux G., Gramfort A., Michel V., Thirion B., Grisel O., Blondel M., Prettenhofer P., Weiss R., Dubourg V., Vanderplas J., Passos A., Cournapeau D., Brucher M., Perrot M., Duchesnay E., Louppe G. Scikit-learn: Machine Learning in Python // *J. Mach. Learning Res.* 2011. N 12. P. 2825–2830.
- Peng W., Karimi Sadaghiani O. A review on the application of machine learning in production of woody biomass from natural and planted forests // *J. Renewable and Sustainable Energy*. 2023. Vol. 15, N 3. P. 032703.
- Ray D., Peace A., Moore R., Petr M., Grieve Y., Convery C., Ziesche T. Improved prediction of the climate-driven outbreaks of *Dendrolimus pini* in *Pinus sylvestris* forests // *Forestry: An Int. J. Forest Res.* 2016. Vol. 89, N 2. P. 230–244.
- Rouault G., Candau J.-N., Lieutier F., Nageleisen L.-M., Martin J.-C., Warzée N. Effects of drought and heat on forest insect populations in relation to the 2003 drought in Western Europe // *Ann. For. Sci.* 2006. Vol. 63, N 6. P. 613–624.
- Sawant S., Golegaonkar P., Gondane P., Gole R., Gole S., Gondkar A., Gorave A., Deshpande R. Earthquake prognosis using machine learning // *ITM Web Conf.* 2023. Vol. 56. P. 05017.
- Shelford V. Physiological animal geography // *J. Morphology*. 1911. Vol. 22, N 3. P. 551–618.
- Toïgo M., Barraquand F., Barnagaud J.-Y., Piou D., Jactel H. Geographical variation in climatic drivers of the pine processionary moth population dynamics // *Forest Ecol. and Management*. 2017. Vol. 404. P. 141–155.
- Virtanen P. et al., SciPy 1.0: fundamental algorithms for scientific computing in Python // *Nat. Methods*. 2020. Vol. 17, N 3. P. 261–272.
- Vlăduț A., Nikolova N., Licurici M. Aridity assessment within southern Romania and northern Bulgaria // *Hrvatski geografski glasnik*. 2017. Vol. 79, N 2. P. 5–26.
- Westgarth-Smith A. R., Roy D. B., Scholze M., Tucker A., Sumpter J. P. The role of the North Atlantic Oscillation in controlling U.K. butterfly population size and phenology // *Ecol. Entomol.* 2012. Vol. 37, N 3. P. 221–232.

Weather data-based prediction of Siberian moth *Dendrolimus sibiricus* Tschetv.: a case study

D. A. DEMIDKO, A. A. GOROSHKO, S. M. SULTSON, N. N. KULAKOVA, P. V. MIKHAYLOV

*Reshetnev Siberian State University of Science and Technology
31, Krasnoyarsky Rabochy ave., 660037, Krasnoyarsk, Russia
E-mail: sawer-beetle@yandex.ru*

The outbreak prediction is one of crucial components of forest pest management. Weather plays substantial role in the leaf-eating insects outbreaks establishment. The weather-based prediction models in this field are numerous and more or less precise. We attempted create such model for Siberian moth (*Dendrolimus sibiricus* Tschetv.) – one of most harmful defoliator in southern taiga of Siberia. For territory of interest (southern taiga and hemiboreal forests of Tomsk Oblast, Kemerovo Oblast and Krasnoyarsk Kray) the gradient boosting (XBGooost) model was created with accuracy 0,952. The temperatures of vegetation period 4th and 5th years before onset of outbreak are better predictors.

Key words: southern taiga, Siberian moth, outbreaks, prediction, weather, machine learning.