

Научная статья / Original article

УДК 597.552.511(639.2.053)

doi:10.15853/2072-8212.2022.65.42-51



РЕЗУЛЬТАТЫ ИСПОЛЬЗОВАНИЯ КЛИМАТО-ОКЕАНОЛОГИЧЕСКИХ ПРЕДИКТОРОВ ДЛЯ МОДЕЛИ СЛУЧАЙНОГО ЛЕСА (RANDOM FOREST) В ЦЕЛЯХ ПРОГНОЗИРОВАНИЯ ДИНАМИКИ ЧИСЛЕННОСТИ ПОДХОДОВ ГОРБУШИ ЗАПАДНОЙ КАМЧАТКИ

Фельдман Марк Геннадиевич, Бугаев Александр Викторович, Тепнин Олег Борисович

Камчатский филиал Всероссийского научно-исследовательского института рыбного хозяйства и океанографии (КамчатНИРО), Петропавловск-Камчатский, Россия, feldman.m.g@kamniro.ru, bugaev.a.v@kamniro.ru, tenpin.o.b@kamniro.ru

Аннотация. Совершенствование прогнозирования подходов горбуши является одной из наиболее актуальных проблем управления лососевыми запасами Камчатки. Перспективным методом прогнозирования представляется алгоритм случайного леса деревьев решений (Random Forest) с использованием итеративного отбора наиболее значимых климато-океанологических предикторов. Определены три достоверно влияющих предиктора для модели случайного леса. Ошибка прогноза при этом составила 28%. Для верификации полученных результатов использованы альтернативные методы отбора наилучших признаков. Полученные результаты показали возможность использования исследуемых климато-океанологических предикторов для моделирования прогнозов динамики численности горбуши Западной Камчатки.

Ключевые слова: горбуша Западной Камчатки, динамика численности подходов, прогнозирование, классификация, моделирование, случайный лес, дерево решений

Для цитирования: Фельдман М.Г., Бугаев А.В., Тепнин О.Б. Результаты использования климато-океанологических предикторов для модели случайного леса (Random Forest) в целях прогнозирования динамики численности подходов горбуши Западной Камчатки // Исследования водных биологических ресурсов Камчатки и северо-западной части Тихого океана. 2022. № 65. С. 42–51X.

RESULTS OF USING CLIMATE-OCEANOLOGICAL PREDICTORS IN A RANDOM FOREST MODEL FOR THE FORECAST OF STOCK ABUNDANCE DYNAMICS OF PINK SALMON RUNS IN WESTERN KAMCHATKA

Mark G. Feldman, Alexandr V. Bugaev, Oleg B. Tepnin

Kamchatka Branch of Russian Federal Research Institute of Fisheries and Oceanography (KamchatNIRO), Petropavlovsk-Kamchatsky, Russia, feldman.m.g@kamniro.ru, bugaev.a.v@kamniro.ru, tenpin.o.b@kamniro.ru

Abstract. An improvement of the forecast of pink salmon runs is one of the most actual problems in management of salmon stocks of Kamchatka. Algorithm of Random Forest with the use of iterative selection of most important climate-oceanological predictors seem to be promising. Three authentically influential predictors have been found for the model of Random Forest at the error of the forecast of 28%. Verification of the obtained results used alternative methods of selection of the best predictors. Obtained results have demonstrated possibility of the use of considered climate-oceanological predictors in modeling the forecasts of stock abundance dynamics of pink salmon in Western Kamchatka.

Keywords: pink salmon of Western Kamchatka, stock abundance dynamics of runs, forecasting, classification, modeling, random forest

For citation: Feldman M.G., Bugaev A.V., Tepnin O.B. Results of using climate-oceanological predictors in a Random Forest model for the forecast of stock abundance dynamics of pink salmon runs in Western Kamchatka // The researches of the aquatic biological resources of Kamchatka and of the north-west part of the Pacific Ocean. 2022. Vol. 65. P. 42–51. (In Russian)

На Камчатке воспроизводится крупнейший по численности запас горбуши в бассейне Северной Пацифики. В начале XXI в. (2001–2021 гг.) в регионе в среднем добывали около 132 тыс. т, что составляло приблизительно 53% от общей среднемноголетней добычи вида на Дальнем Востоке России (~ 247 тыс. т). При этом в по-

следнее десятилетие (2012–2021 гг.) уровень вылова горбуши Камчатки достиг в среднем 186 тыс. т, что составило более 67% от дальневосточного показателя среднемноголетнего вылова.

Западное побережье Камчатки является одним из основных центров воспроизводства

камчатской горбуши. В XXI в. в урожайные годы (четная линия лет воспроизводства) здесь в среднем ежегодно добывали около 106 тыс. т вида. На восточном побережье Камчатки в урожайные годы (нечетная линия лет воспроизводства) добывали около 119 тыс. т. Это составляет приблизительно 47% и 53% соответственно. Таким образом, западное побережье продуцирует порядка 50% от общего запаса высокоурожайных поколений горбуши Камчатки. Следовательно, прогнозирование динамики численности запасов горбуши Западной Камчатки является важнейшей практической задачей для всего Дальневосточного рыбохозяйственного бассейна.

В настоящее время прогнозы численности горбуши Камчатки строятся на различных методах математического моделирования, а также данных учетных траловых съемок, ежегодно проводимых во время осенней откочевки молоди в бассейнах Охотского и Берингова морей (Фельдман, Шевляков, 2015; Фельдман, 2020; Фельдман, Бугаев, 2021). В Охотском море информация о численности молоди горбуши дополняется данными результатов генетической идентификации о региональном происхождении рыб в смешанных выборках. Это необходимо для вычисления доли горбуши, имеющей западно-камчатское происхождение. Соответственно, оставшаяся доля молоди в основном будет представлять южноохотоморские центры воспроизводства вида: Сахалин, Южные Курильские о-ва и бассейн р. Амур. Наличие фактора смешивания молоди различного регионального происхождения в бассейне Охотского моря значительно усложняет прогнозирование динамики запасов горбуши Западной Камчатки с преимущественным использованием только данных учетных траловых съемок.

Поэтому для западнокамчатской горбуши методы математического моделирования прогнозов численности могут иметь определяющее значение. Моделирование процессов динамики запасов имеет многофакторный характер, включающий, помимо исходных данных о численности объекта, набор предикторов, отражающих влияние среды на онтогенез и конечные продукционные показатели рыб. В связи с этим для создания прогностических моделей актуален поиск наиболее значимых климато-океанологических факторов (предикторов), так как основной рост и формирование численности горбуши происходят во вре-

мя морского/океанического нагула. Отбор статистически значимых предикторов возможен с помощью использования модели случайного леса деревьев решений (Random Forest). В настоящее время данный метод многомерного моделирования прогнозов динамики численности горбуши показал высокий уровень результативности для запаса вида, воспроизводящегося в реках Северо-Восточной Камчатки (Фельдман, 2020).

Целью работы является оценка значимости климато-океанологических предикторов для моделирования прогнозов динамики численности горбуши Западной Камчатки.

МАТЕРИАЛ И МЕТОДИКА

В качестве материала использованы данные о численности родителей (численность производителей на нерестилищах (нерестовый запас)) и потомков (численность возвратов (подходов) производителей) горбуши Западной Камчатки по наблюдениям в 1982–2019 гг. Сбор первичной информации выполнен сотрудниками Северо-Восточного территориального управления Росрыболовства (промысловая статистика) и КамчатНИРО (данные о заполнении нерестилищ и оценки численности подходов).

Методической основой работы послужил классификационный анализ, выполненный с помощью метода Random Forest (случайный лес деревьев решений) (Breiman, 2001). Зависимые переменные обозначены тремя уровнями стратифицированной модели типа «запас–пополнение» (рис. 1), ранее разработанной для горбуши западного побережья Камчатки (Фельдман и др., 2019).

В ходе анализа временных рядов некоторая часть наблюдений была исключена из анализа, так как из-за относительно низкой численности родителей и потомков данные переменные статистически недостоверны. К ним относятся наблюдения нечетной генеративной линии горбуши Западной Камчатки 90-х гг. XX в. (1993, 1995 и 1999 гг.). В XXI в. исключены наблюдения 2009, 2011 и 2013 гг. Таким образом, всего было определено 6 лет наблюдений, которые интерпретированы как «точки выброса». Следовательно, набор использованных в работе данных составил 32-летний период. Общий временной ряд наблюдений был разделен на три уровня, характеризующих эффективность воспроизводства:

1-й уровень (high) включает 9 лет наблюдений;

2-й уровень (medium) включает 11 лет наблюдений;

3-й уровень (low) включает 12 лет наблюдений (табл. 1).

В качестве предикторов для модели случайного леса были выбраны климатические индексы, которые отражают воздействие изменчивости метеорологических и океанологических условий на воспроизводство и последующее формирование численности горбуши Западной Камчатки во время морского/океанического нагула:

PDO — индекс Тихоокеанской декадной осцилляции (Mantua et al., 1997; Mantua, Hare, 2002). Данные взяты с сайта Национального управления океанических и атмосферных исследований (NOAA, США) (<https://www.ncdc.noaa.gov/teleconnections/pdo>);

WP — Западно-Тихоокеанский индекс циклонической активности (Linkin, Nigam, 2008). Данные взяты с сайта NOAA (<https://www.cpc.ncep.noaa.gov/data/teledoc/wp.shtml>);

АО — индекс Арктической осцилляции (Thompson, Wallace, 1998). Данные взяты с сайта NOAA (https://www.cpc.ncep.noaa.gov/products/precip/CWlink/daily_ao_index/ao.shtml);

aSST — аномалия температуры морской поверхности. Данные получены из акваторий морского нагула горбуши Западной Камчатки на раннем и заключительном этапах морского периода жизни в восточной части Охотского моря. Ранее было показано, что показатели aSST являются климатическими индикатора-

ми, отражающими влияние температурного режима вод на формирование численности тихоокеанских лососей как непосредственно Камчатки, так и в целом Дальнего Востока России (Бугаев, Тепнин, 2015; Бугаев и др., 2018, 2021).

В настоящей работе переменные aSST ориентированы исключительно на горбушу Западной Камчатки, поэтому здесь задействованы среднемесячные показатели из следующих районов:

а) квадрат 22 (средние координаты: 52°50' с. ш. 155°00' в. д.) — расположен у юго-западного побережья Камчатки, в соответствии с разбивкой районов, представленной в работе А.В. Бугаева с соавторами (2018), для бассейна Северной Пацифики (рис. 2). Для обозначения переменных из этого района используется аббревиатура SEA;

б) прибрежная зона Западной Камчатки: район, ограниченный координатами 51°00'–58°00' с. ш. и 153°00'–157°00' в. д. Для обозначения переменных из этого района используется аббревиатура COAST.

В работе для детализации переменных принято трехзначное обозначение предикторов. Первые три буквы соответствуют названию месяца (на английском языке). Далее следует цифра, обозначающая основные этапы жизненного цикла горбуши (1 — год нереста, 2 — год ската). В конце — обозначение используемого в работе климатического фактора: PDO, AO, WP, SEA, COAST.

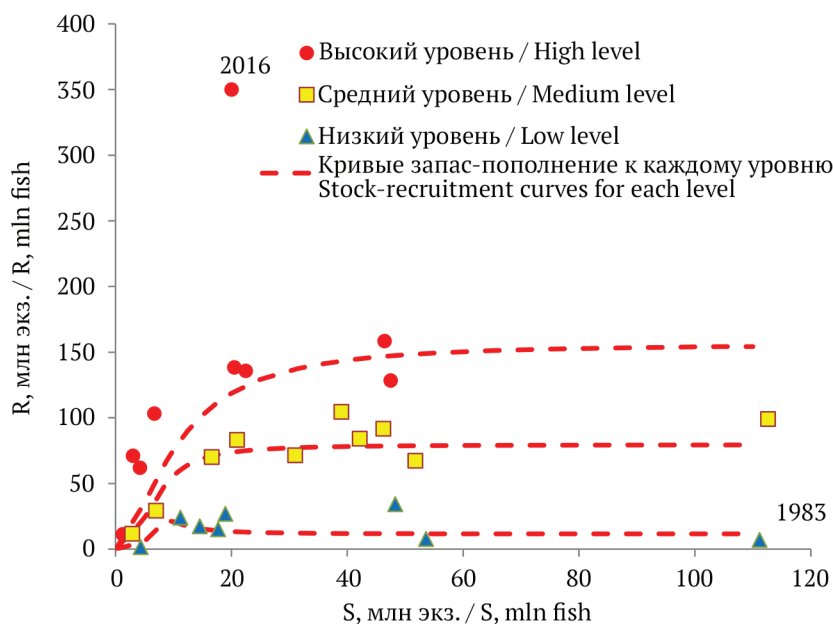


Рис. 1. Стратифицированная модель «запас–пополнение» для горбуши Западной Камчатки по данным 1982–2019 гг.: S — нерестовый запас, R — возврат (подход)
Fig. 1. Stratified model “stock–recruitment” for pink salmon of Western Kamchatka on the data for 1982–2019: S — spawning stock, R — return (run)

Таблица 1. Информация о численности родителей и потомков горбуши Западной Камчатки и уровнях эффективности воспроизводства, определенных с помощью стратифицированной модели «запас–пополнение» по данным 1982–2019 гг.

Table 1. Data on the abundance of parents and progenies of pink salmon in Western Kamchatka and levels of reproduction efficiency, found with the use of stratified model “stock-recruitment” on the data for 1982–2019

Год Year	Численность, млн экз. / Abundance, mln fish		Уровень Level
	Родители (нерестовый запас) Parents (spawning stock)	Потомки (возврат) Progeny (return)	
1982	16,57	62,46	medium
1983	111,16	7,01	low
1984	48,27	34,21	low
1985	4,34	1,57	low
1986	18,96	26,83	low
1987	1,29	1,30	medium
1988	14,12	30,39	low
1989	1,04	2,53	high
1990	14,53	17,49	low
1991	1,42	0,61	low
1992	6,68	103,24	high
1993	0,48	0,57	–
1994	81,09	80,41	medium
1995	0,50	0,87	–
1996	47,49	128,39	high
1997	0,68	0,08	low
1998	42,12	84,08	medium
1999	0,04	1,37	–
2000	20,96	83,31	medium
2001	1,22	11,43	high
2002	46,20	91,79	medium
2003	11,19	23,98	low
2004	51,74	67,39	medium
2005	17,71	15,19	low
2006	31,00	71,49	medium
2007	3,64	4,25	low
2008	38,95	104,52	medium
2009	0,12	5,05	–
2010	46,43	158,55	high
2011	0,83	1,57	–
2012	33,56	8,31	low
2013	0,21	4,60	–
2014	3,49	71,20	high
2015	2,85	11,76	medium
2016	20,00	350,00	high
2017	4,20	68,00	high
2018	112,60	99,18	medium
2019	15,00	229	high

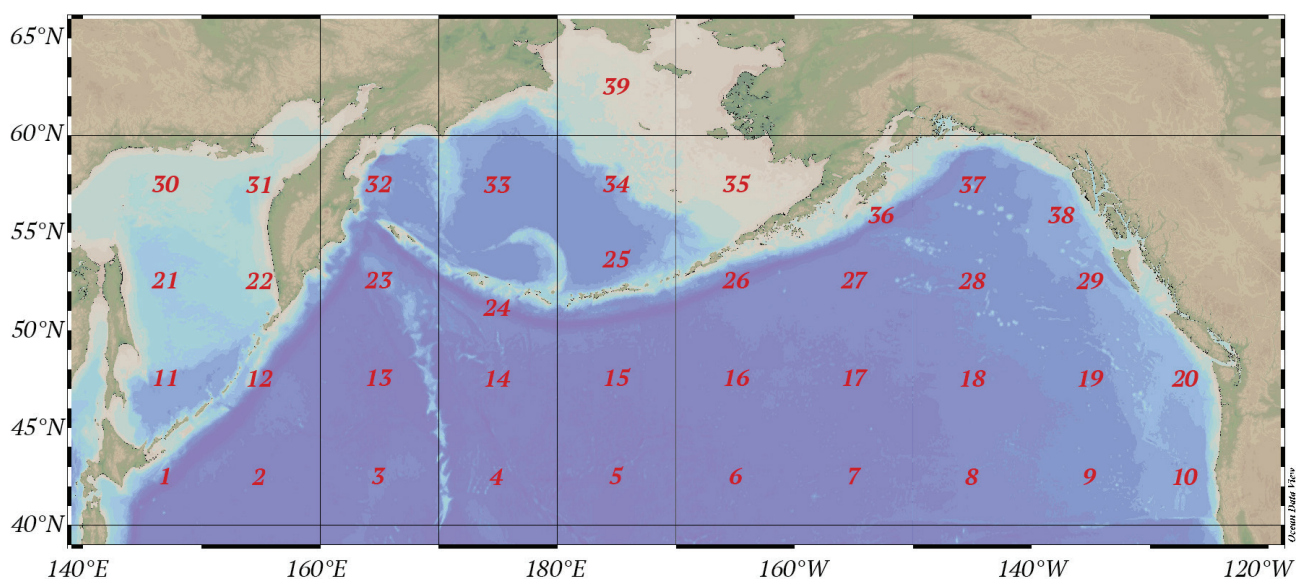


Рис. 2. Карта-схема бассейна Северной Пацифики и районов, для которых рассчитывались среднемесячные значения аномалий температуры поверхностного слоя вод (aSST). Цифрами обозначены картографические трапеции (5×10°), являющиеся зонами расчета параметра (Бугаев и др., 2018)

Fig. 2. Schematic map of the North Pacific Ocean and districts for which monthly averages of the abnormal sea surface temperatures (aSST) were calculated. The numbers indicate cartographic trapezoids (5×10°), which are the areas to calculate the parameter (Бугаев et al., 2018)

Методика отбора значимых предикторов совпадает с использованным ранее итеративным методом исключения незначимых признаков для горбуши Северо-Восточной Камчатки, реализованным в среде *R* (Фельдман, 2020). В качестве сторонних методов для верификации полученного результата использован метод Борута (Boruta) (Kursa, Rudnicki, 2010; Kursa, 2020) с одноименным пакетом в среде *R*, а также пакет Random Forest Explainer (<https://cran.rstudio.com/web/packages/randomForestExplainer/vignettes/randomForestExplainer.html>).

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Первым этапом была проведена оценка значимости климатических предикторов по воздействию на изменчивость численности возвратов (подходов) горбуши Западной Камчатки с учетом двухгодичной периодичности жизненного цикла вида. В ходе итеративного исключения малозначимых признаков (рис. 3), наилучший средний из 1000 вариантов случайного леса результат был достигнут на 47-й итерации с двумя используемыми признаками на итерации.

Средняя доля ошибок вне выборки (ООВ) на 47-й итерации составила 28%, с размахом от 22% до 44%. Распределение ошибки не вполне отвечает законам нормальности распределения переменных и смещено в меньшую сторону (рис. 4). Большинство вариантов случайного леса (567) имеет среднюю ошибку ООВ в 28%.

По встречаемости признаков в совокупности 1000 моделей к наиболее важным предикторам можно отнести три. На первом месте по важности в моделях случайного леса в подавляющем большинстве случаев (99,3%) находится показатель *aSST* в квадрате 22 в июне, который относится к году нереста поколения (обозначен как Jun1SEA). Данный период совпадает со сроками преднерестовых миграций горбуши в прибрежных водах Охотского моря непосредственно перед заходом родительских стад в нерестовые реки (табл. 2). Возможно, что условия нагула в этот период являются ключевыми для закладки физиологических параметров производителей для дальнейшего эффективного нереста.

Следует отметить, что модель случайного леса не зафиксировала наличия значимых связей в отношении зависимости возвратов поко-

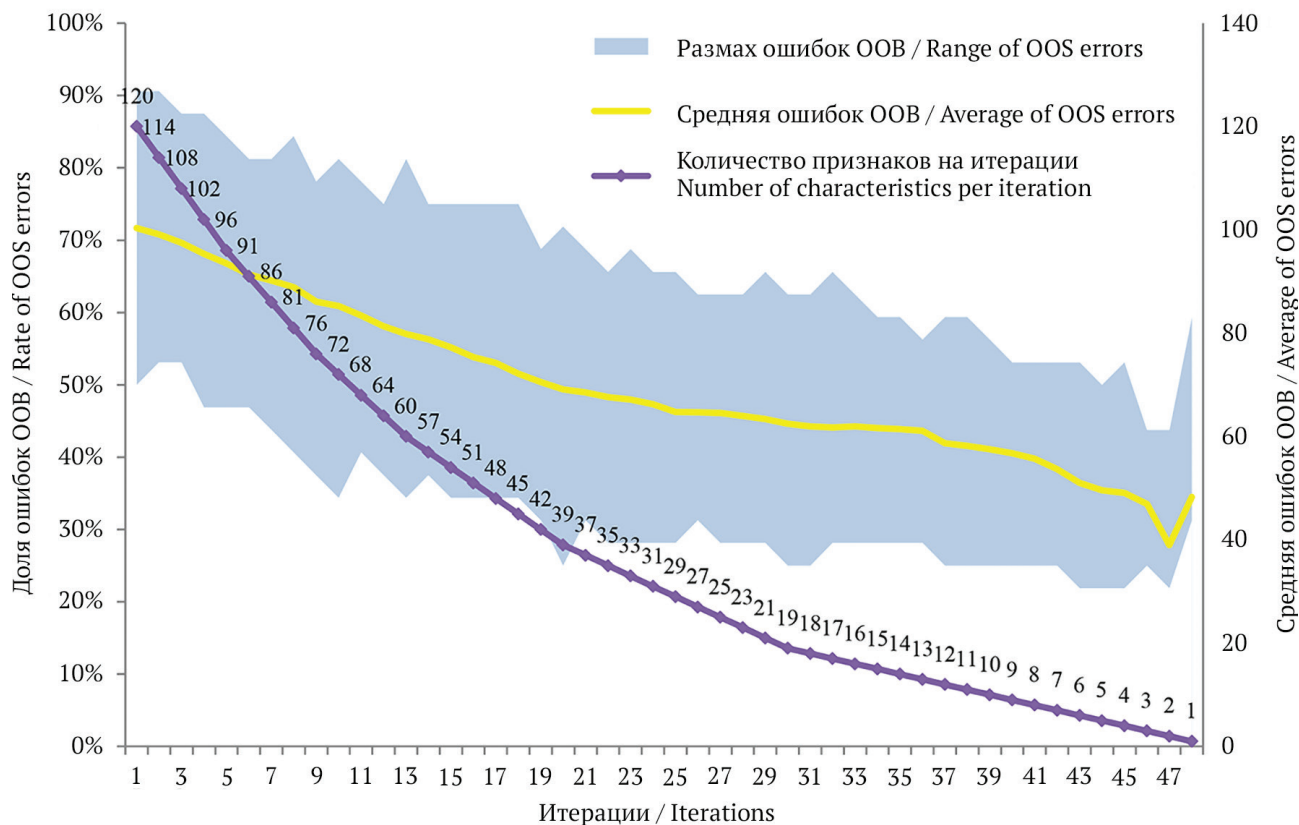


Рис. 3. Динамика изменения средней доли ошибок вне выборки (ООВ) и их размаха для 1000 вариантов случайного леса по мере исключения наименее значимых предикторов на каждой итерации
Fig. 3. Dynamics of the mean out-of-sample (OOS) error rate changes and ranges for 1000 random forest decisions in the course of exclusion of the least significant predictors at every iteration

лений горбуши от изменчивости aSST в зонах осеннего нагула молоди стад Западной Камчатки в бассейне Охотского моря. Подобные взаимосвязи ранее были зафиксированы с применением прямой корреляционной зависимости (Бугаев и др., 2018, 2021). Однако уточним, что их характер также был относительно слабо выражен. Вероятно, использование классификационных методов требует привлечения комплекса дополнительных данных о состоянии термического режима вод и других природных условий в зонах осеннего нагула молоди горбуши Западной Камчатки.

Второе и третье места достаточно близко по значимости разделяют предикторы, характеризующие метеорологические и океанологические условия морского/океанического нагула горбуши: Западно-Тихоокеанский индекс циклонической активности в январе (Jan1WP — поколение нереста) и индекс Арктической осцилляции в феврале (Feb2AO — поколение ската). В первом случае потенциальное воздействие на условия нереста родительского поколения могут оказывать климатические условия в зимний период. Во втором случае — влияние

климатических условий перед выклевом мальков горбуши из нерестовых бугров.

К сожалению, предиктор aSST в прибрежной зоне Западной Камчатки во время ската молоди горбуши (обозначенный как Apr2COAST) оказался малозначимым на уровне статистической ошибки. Вероятно, данный показатель имеет слишком динамичную природу, так как связан с активными приливно-отливными течениями в прибрежной зоне. Это не позволяет адекватно определить временной диапазон подбора первичных переменных aSST для получения точной индикаторной информации. Поэтому требуется дальнейшая отработка методики по применению данного предиктора в целях прогнозирования динамики запасов горбуши Западной Камчатки.

Для верификации полученных результатов использовали альтернативные методы отбора наилучших признаков. Первым из таких был метод Борута в одноименном пакете для языка R. В данном случае также выделены три подтвержденных значимых предиктора: Jun1SEA, Feb2AO и Jan1WP (рис. 5). Два признака (Apr2COAST, Jul1AO), так называемые чувствительные, почти совпадают с верхней границей для отвергнутых признаков (ShadowMax).

В статье описывается R-пакет Борута, который позволяет реализовать новый алгоритм выбора функций с поиском всех соответствующих переменных. Алгоритм разработан в качестве оболочки для алгоритма классификации случайных лесов. Он итеративно удаляет функции статистических тестов, которые менее актуальны для случайных проб. Пакет Борута предоставляет удобный интерфейс для алгоритма.

Следующим из альтернативных методов был применен пакет Random Forest Explorer, который с помощью сравнения различных метрик значимости признаков дает несколько иные результаты. По таким показателям важности признаков, как среднее уменьшение точности прогноза при исключении данной переменной (ось x), среднее уменьшение неопределенности Джини (мера однородности узлов и

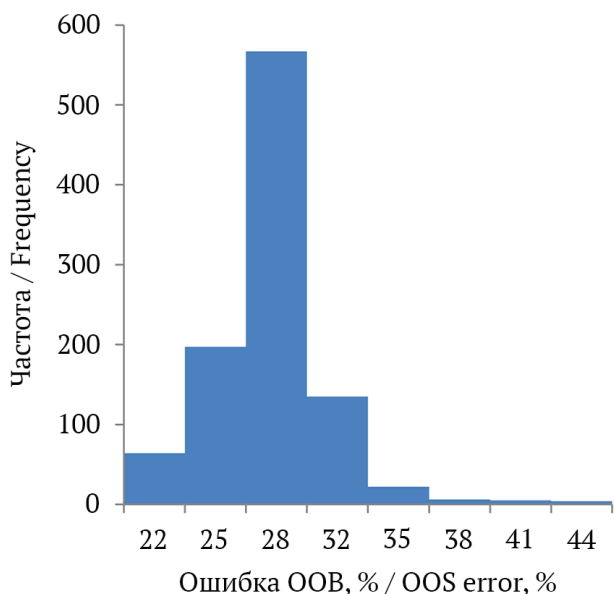


Рис. 4. Распределение средней ошибки OOB по 1000 вариантов случайного леса
Fig. 4. Distribution of the mean OOS error over 1000 random forest decisions

Таблица 2. Важность признаков в 1000 моделей случайного леса на 47-й итерации с двумя признаками
Table 2. Importance of the predictors in 1000 random forest decisions at the 47th iteration with two characteristics

Признак / Predictor	1 место / 1 st position	2 место / 2 nd position
Jun1SEA	993	7
Jan1WP	7	542
Feb2AO	0	436
Apr2COAST	0	13
Apr2AO	0	1
Jul1AO	0	1

листьев — ось y) и средней минимальной глубины признака в дереве (чем выше признак, тем он более значим, обозначено размером маркеров), абсолютными лидерами являются два признака: Jun1SEA и Feb2AO (рис. 6). Предиктор Jan1WP хоть и входит в пятерку наиболее значимых признаков, но не является определяющим. И если в предыдущих методах Jan1WP и Feb2AO разделяли второе и третье места, то в данном анализе признак Jan1WP с большим отрывом находится на третьем месте совместно с признаками Apr2WP и Mar2AO, зна-

чимостью которых предыдущие методы вообще не показали.

Ошибка прогноза по совокупности ансамбля из 1000 моделей случайного леса на 47-й итерации составила 28%, т. е. из 32 наблюдений 9 определены неправильно (табл. 3). Необходимо отметить, что ошибочные прогнозы поколений 1998 и 2008 гг. разделяются почти поровну между высоким и низким уровнями (в первом случае 60 на 40%, во втором 56 на 44%). По сути, в реальности наблюдается средний уровень. Поскольку алгоритм отдает предпочтение уровню с простым

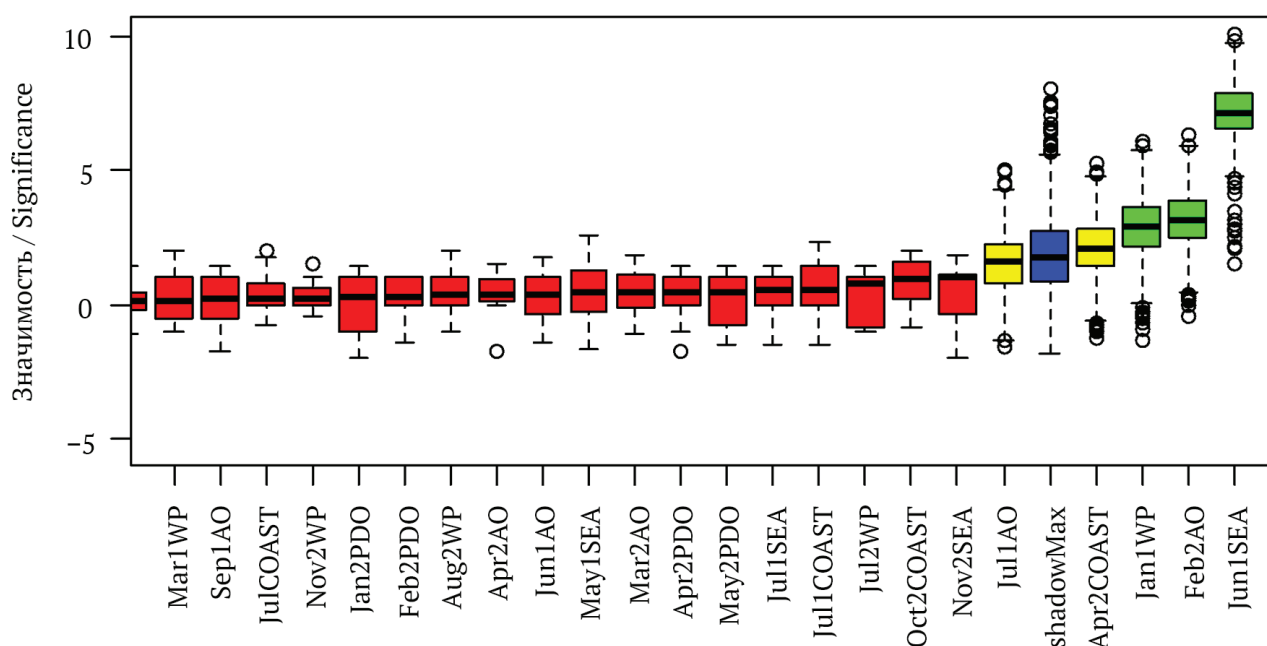


Рис. 5. Анализ значимости признаков с помощью метода Борута: зеленый цвет — подтвержденные признаки, желтый — чувствительные признаки, красный цвет — неподтвержденные признаки, круги — верхняя граница теневых, заведомо незначимых искусственных признаков (ShadowMax), бокс — стандартная ошибка, усы — стандартное отклонение, линия — средний показатель

Fig. 5. Analysis of the significance of characteristics using the Borut method: green color — confirmed characteristics, yellow color — sensitive characteristics, red color — unconfirmed characteristics, circles — the upper limit of shadow, obviously insignificant artificial characteristics (ShadowMax), boxing — standard error, whiskers — standard deviation, line — the average

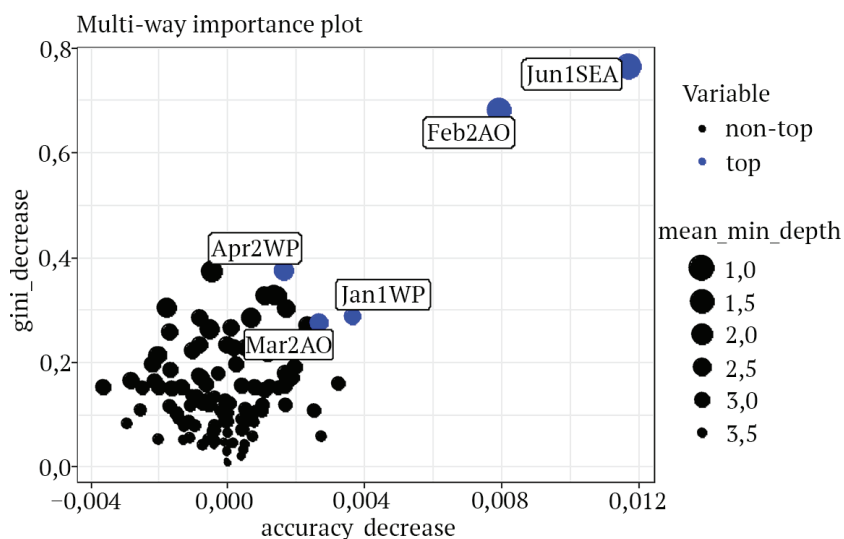


Рис. 6. Многоплановый график значимости признаков (предикторов) по трем метрикам (пояснение в тексте)

Fig. 6. Multidimensional graph of the significance of characteristics (predictors) for three metrics (explanation in the text)

большинством голосов и «не знает», что средний уровень занимает промежуточное положение между низким и высоким, полагаем, что исследователь вправе в таких случаях при прогнозировании принять окончательное решение и выбрать средний уровень. Соответственно, тем самым ошибка прогноза уменьшится примерно до 20%.

С помощью дерева решений или диаграммы рассеивания по трем важнейшим признакам можно давать предварительные прогнозы с заблаговременностью более одного года (рис. 7). Полученное дерево имеет три уровня деления и пять конечных листов, ошибка классификации в этом случае составила 12,5%, в том числе

Таблица 3. Результаты голосования по 1000 вариантов моделей случайного леса на 47-й итерации по данным 1982–2019 гг.

Table 3. Results of voting for over 1000 decisions of the random forest model at the 47th iteration on the data for 1982–2019

Год / Year	Уровень / Level	High, %	Medium, %	Low, %	Прогноз / Forecast
1982	medium	0,0	52,6	47,4	medium
1983	low	0,0	58,7	41,3	medium*
1984	low	0,9	0,0	99,1	low
1985	low	0,0	100,0	0,0	medium*
1986	low	0,1	1,3	98,6	low
1987	medium	0,0	99,9	0,1	medium
1988	low	43,5	0,0	56,5	low
1989	high	43,8	0,0	56,2	low*
1990	low	55,1	0,0	44,9	high*
1991	low	44,4	0,6	55,0	low
1992	high	79,4	3,5	17,1	high
1994	medium	1,5	98,5	0,0	medium
1996	high	44,1	0,5	55,4	low*
1997	low	0,0	0,0	100,0	low
1998	medium	60,8	0,0	39,2	high*
2000	medium	0,0	99,9	0,1	medium
2001	high	75,7	0,0	24,3	high
2002	medium	0,0	100,0	0,0	medium
2003	low	0,0	100,0	0,0	medium*
2004	medium	0,0	36,3	63,7	low*
2005	low	3,3	0,0	96,7	low
2006	medium	0,0	100,0	0,0	medium
2007	low	1,4	0,0	98,6	low
2008	medium	56,4	0,0	43,6	high*
2010	high	98,7	0,0	1,3	high
2012	low	0,1	0,0	99,9	low
2014	high	99,7	0,1	0,2	high
2015	medium	0,0	96,1	3,9	medium
2016	high	99,4	0,4	0,2	high
2017	high	94,8	3,8	1,4	high
2018	medium	0,0	100,0	0,0	medium
2019	high	99,3	0,0	0,7	high

Примечание. * — ошибка прогноза. Note. * – is an error of the forecast.

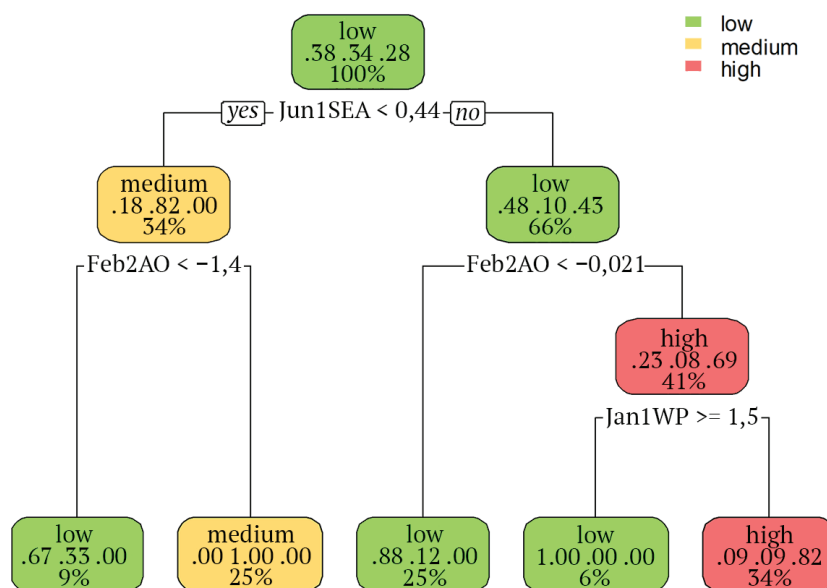


Рис. 7. Классификация данных с помощью дерева решений по трем важнейшим признакам
Fig. 7. Data classification using a decision tree on the three most important predictors

два наблюдения ошибочно отнесены к уровню high и два наблюдения к уровню low.

ЗАКЛЮЧЕНИЕ

Проделанная работа позволила выявить три значимых климато-океанологических фактора (предиктора), использование которых позволяет строить с помощью метода случайного леса прогностические модели динамики численности запасов горбуши Западной Камчатки с достаточной заблаговременностью. Конечно, данные факторы не являются окончательным решением проблемы. Фактически, построенная модель представляет собой конгломерат из двух фракций простых деревьев решений, каждая из которых основана на двух предикторах. При этом один из них присутствует в обеих фракциях (Jun1SEA), а второй характерен только для одной фракции (Jan1WP и Feb2AO). Тем не менее в голосовании деревьев решений по некоторой части обучающего набора обе фракции дают противоречивые прогнозы.

Тестовая ошибка вне выборки (ООВ) наиболее велика в случае прогноза низкочисленных поколений (40%), однако минимальна для прогноза высокочисленных поколений (13%). Для среднего уровня ошибка составляет 32%. Данное обстоятельство обязывает исследователей принимать окончательное прогнозное решение, ориентируясь на дополнительные материалы и индикаторы.

Тем не менее с помощью дерева решений или диаграммы рассеивания по двум обозначенным фракциям предикторов можно давать адекватные прогнозы с примерно годичной заблаговременностью. Результаты данных исследований включены в прогнозы динамики численности запасов горбуши Западной Камчатки на 2023 г.

СПИСОК ИСТОЧНИКОВ

Бугаев А.В., Тепнин О.Б. 2015. Продуктивность тихоокеанских лососей: влияние термических условий вод в период первой зимы в бассейне Северной Пацифики // Тр. ВНИРО. Т. 158. С. 89–111.

Бугаев А.В., Тепнин О.Б., Радченко В.И. 2018. Климатическая изменчивость и продуктивность тихоокеанских лососей Дальнего Востока России // Исслед. водн. биол. ресурсов Камчатки и сев.-зап. части Тихого океана. Вып. 49. С. 5–50.

Бугаев А.В., Фельдман М.Г., Тепнин О.Б., Коваль М.В. 2021. Аномалии температуры поверхности воды в западной части Северной Пацифи-

ки — потенциальный климатический предиктор прогнозирования тихоокеанских лососей Камчатки // Вопр. рыболовства. Т. 22. № 4. С. 46–62.

Фельдман М.Г. 2020. Использование метода Random Forest в целях прогнозирования подходов горбуши северо-востока Камчатки // Исслед. водн. биол. ресурсов Камчатки и сев.-зап. части Тихого океана. Вып. 59. С. 76–96.

Фельдман М.Г., Бугаев А.В. 2021. Современные принципы управления запасами горбуши *Oncorhynchus gorbuscha* Камчатского края (динамика численности, прогнозирование, регулирование промысла) // Вопр. рыболовства. Т. 22. № 4. С. 86–95.

Фельдман М.Г., Шевляков Е.А. 2015. Выживаемость камчатской горбуши как результат совокупного воздействия плотностной регуляции и внешних факторов среды // Изв. ТИНРО. Т. 182. С. 88–114.

Фельдман М.Г., Шевляков Е.А., Артюхина Н.Б. 2019. Оценка ориентиров пропуска в реки Западной Камчатки // Исслед. водн. биол. ресурсов Камчатки и сев.-зап. части Тихого океана. Вып. 52. С. 50–78.

Breiman L. 2001. Random Forest // Machine learning. Vol. 45 (1). P. 5–32.

Kursa M. 2020. Boruta for those in a hurry. Electronic article. 6 p. (<https://cran.r-project.org/web/packages/Boruta/vignettes/inahurry.pdf>)

Kursa M., Rudnicki W. 2010. Feature Selection with the Boruta Package // Journal of Statistical Software. Vol. 36 (11). P. 2–12.

Linkin M.E., Nigam S. 2008. The North Pacific Oscillation – West Pacific Teleconnection Pattern: Mature – Phase Structure and Winter Impacts // Journal of Climate. Vol. 21 (9). P. 1979–1997.

Mantua N., Hare S., Zhang Y., Wallace J., Francis R.A. 1997. Pacific interdecadal climate oscillation with impacts on salmon production, Bull. Amer. Meteor. Soc. № 78. P. 1069–1079.

Mantua N.J., Hare S.R. 2002. The Pacific Decadal Oscillation // Journal of Oceanography. Vol. 58. P. 35–44.

Thompson D.W.J., Wallace J.M. 1998. The Arctic Oscillation signature in the wintertime geopotential height and temperature fields // Geophysical Research Letters. № 25. P. 1297–1300.

REFERENCES

Bugaev A.V., Tepnin O.B. Productivity of Pacific salmon: effects of water temperature conditions during first wintering in the North Pacific basin. *Trudy VNIRO*, 2015, vol. 158, pp. 89–111. (In Russian)

Bugaev A.V., Tepnin O.B., Radchenko V.I. Climate variability and Pacific salmon productivity in Russian Far East. *The researches of the aquatic biological resources of Kamchatka and of the north-west part of the Pacific Ocean*, 2018, vol. 49, pp. 5–50. (In Russian)

Bugaev A.V., Feldman M.H., Tepnin O.B., Koval M.V. Water surface temperature anomalies in the western North Pacific – potential climatic predictor of Pacific salmon runs in Kamchatka. *Problems of Fisheries*, 2021, vol. 22, № 4, pp. 46–62. (In Russian)

Feldman M.H. Using the decisions of the random forest algorithm for the purposes of forecasting pink salmon runs on North-Eastern Kamchatka. *The researches of the aquatic biological resources of Kamchatka and of the north-west part of the Pacific Ocean*, 2020, vol. 59, pp. 76–96. (In Russian)

Feldman M.H., Bugaev A.V. Modern principles of pink salmon (*Oncorhynchus gorbuscha*) stocks management in the Kamchatka Region (population dynamics, forecasting, fishing regulation). *Problems of Fisheries*, 2021, vol. 22, no. 4, pp. 86–95. (In Russian)

Feldman M.G., Shevlyakov E.A. Survival of Kamchatka pink salmon as result of combined influence of density regulation and environmental factors. *Izvestia TINRO*, 2015, vol. 182, pp. 88–114. (In Russian)

Feldman M.G., Shevlyakov E.A., Artukhina N.B. Evaluation of the Pacific salmon spawning escapement parameters for the river basins of West Kamchatka. *The researches of the aquatic biological resources of Kamchatka and of the north-west part of the Pacific Ocean*, 2019, vol. 52, pp. 50–78. (In Russian)

Breiman L. Random Forests. *Machine Learning*, 2001, no. 45 (1), pp. 5–32.

Kursa M. Boruta for those in a hurry. Online article, 2020, 6 p. (<https://cran.r-project.org/web/packages/Boruta/vignettes/inahurry.pdf>)

Kursa M., Rudnicki W. Feature Selection with the Boruta Package. *Journal of Statistical Software*, 2010, vol. 36 (11), pp. 2–12.

Linkin M.E., Nigam S. The North Pacific Oscillation – West Pacific Teleconnection Pattern: Mature – Phase Structure and Winter Impacts. *Journal of Climate*, 2008, vol. 21 (9), pp. 1979–1997.

Mantua N., Hare S., Zhang Y., Wallace J., Francis R.A. Pacific interdecadal climate oscillation with impacts on salmon production. *Bull. Amer. Meteor. Soc.*, 1997, № 78, pp. 1069–1079.

Mantua N.J., Hare S.R. The Pacific Decadal Oscillation. *Journal of Oceanography*, 2002, vol. 58, pp. 35–44.

Thompson D.W.J., Wallace J.M. The Arctic Oscillation signature in the wintertime geopotential height and temperature fields. *Geophysical Research Letters*, 1998, № 25, pp. 1297–1300.

Информация об авторах

М.Г. Фельдман — канд. биол. наук, вед. н. с. Камчатского филиала ВНИРО (КамчатНИРО)

А.В. Бугаев — док. биол. наук, зам. руководителя Камчатского филиала ВНИРО (КамчатНИРО)

О.Б. Тепнин — зав. сектором Камчатского филиала ВНИРО (КамчатНИРО)

Information about the authors

Mark G. Feldman – Ph. D. (Biology), Leading Researcher (KamchatNIRO)

Alexandr V. Bugaev – Dr. of Science (Biology), Deputy Director (KamchatNIRO)

Oleg B. Tepnin – Head of Division (KamchatNIRO)

Статья поступила в редакцию: 27.06.2022

Одобрена после рецензирования: 21.10.2022

Статья принята к публикации: 23.10.2022