Известия ТИНРО

2024 **Том 204, вып. 4. С. 1018–1034.**

Izvestiya TINRO, 2024, Vol. 204, No. 4, pp. 1018–1034.

Краткое сообщение

УДК 639.28:004.5
DOI: 10.26428/1606-9919-2024-204-1018-1034
EDN: TREJKR

ПРИМЕНЕНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ВОССТАНОВЛЕНИЯ РАЗМЕРНО-ПОЛОВОГО СОСТАВА УЛОВОВ КРАБА-СТРИГУНА ОПИЛИО

И.С. Черниенко, А.Г. Слизкин, В.В. Кулик*
Тихоокеанский филиал ВНИРО (ТИНРО),
690091, г. Владивосток, пер. Шевченко, 4

Аннотация. Выявлены статистические связи между условиями окружающей среды и пространственным распределением размерно-половых групп краба-стригуна опилио в Западно-Беринговоморской зоне. На основе найденных закономерностей выполнено восстановление значений абсолютных траловых уловов размерно-половых групп, отсутствующих в базах данных. Для формирования входных переменных и оценки статистических связей использовались методы машинного обучения, которые могут быть применены к другим единицам запаса донных гидробионтов.

Ключевые слова: Берингово море, краб-стригун опилио *Chionoecetes opilio*, состав уловов, условия обитания гидробионтов, машинное обучение

Для цитирования: Черниенко И.С., Слизкин А.Г., Кулик В.В. Применение методов машинного обучения для восстановления размерно-полового состава уловов краба-стригуна опилио // Изв. ТИНРО. — 2024. — Т. 204, вып. 4. — С. 1018–1034. DOI: 10.26428/1606-9919-2024-204-1018-1034. EDN: TREJKR.

Short message

Application of machine learning methods to restore size-sex composition in catches of snow crab

Igor S. Chernienko*, Aleksei G. Slizkin, Vladimir V. Kulik*****
*–*** Pacific branch of VNIRO (TINRO), 4, Shevchenko Alley, Vladivostok, 690091, Russia
* Ph.D., leading researcher, igor.chernienko@tinro.vniro.ru, ORCID 0000-0002-6410-0081
** Ph.D., leading researcher, aleksei.sleezkin@tinro.vniro.ru, ORCID 0009-0006-9368-0771
*** Ph.D., head of laboratory, vladimir.kulik@tinro.vniro.ru, ORCID 0000-0003-0920-5312

Abstract. Impact of environmental conditions on spatial distribution of size-sex groups is described statistically for snow crab *Chionoecetes opilio*. Based on the relationships identified, absolute values of catches are calculated for each such group. Machine learning approach is

* *Черниенко Игорь Сергеевич, кандидат биологических наук, ведущий научный сотрудник, igor.chernienko@tinro.vniro.ru, ORCID 0000-0002-6410-0081; Слизкин Алексей Гаврилович, кандидат биологических наук, ведущий научный сотрудник, aleksei.sleezkin@tinro.vniro.ru, ORCID 0009-0006-9368-0771; Кулик Владимир Владимирович, кандидат биологических наук, заведующий лабораторией, vladimir.kulik@ tinro.vniro.ru, ORCID 0000-0003-0920-5312.*

© Черниенко И.С., Слизкин А.Г., Кулик В.В., 2024

implemented for the feature engineering and statistical relationships evaluation. The approach can be adopted for other benthic stocks.

Keywords: Bering Sea, snow crab, *Chionoecetes opilio*, catch composition, habitat environment, machine learning

For citation: Chernienko I.S., Slizkin A.G., Kulik V.V. Application of machine learning methods to restore size-sex composition in catches of snow crab, *Izv. Tikhookean. Nauchno-Issled. Inst. Rybn. Khoz. Okeanogr*., 2024, vol. 204, no. 4, pp. 1018–1034. (In Russ). DOI: 10.26428/1606-9919-2024-204-1018-1034. EDN: TREJKR.

Введение

На протяжении многих лет рыбохозяйственная наука занимается сбором данных о состоянии запасов водных биологических ресурсов на Дальнем Востоке. Полученные материалы ложатся в основу научно обоснованных рекомендаций по эксплуатации этих ресурсов. В ТИНРО разработано несколько баз данных (БД), используемых для хранения материалов научно-исследовательских работ. В БД «Морская биология» содержатся материалы, собранные в ходе исследований, выполненных с 1950-х гг. по настоящее время. Эта БД позволяет быстро извлекать данные об обилии и составе уловов гидробионтов, а также об условиях их обитания и привести все сведения в форму, пригодную для дальнейшего использования, например как входные данные для моделей динамики обилия запаса. Однако часть материалов в БД отсутствует. Такая ситуация связана с тремя основными причинами. Первая состоит в значительном объеме первичных материалов. Перенос данных в таблицы БД требует больших затрат времени. При этом необходим предварительный анализ первичных материалов, включающий поиск и исправление ошибок, фильтрацию, приведение данных к единому формату. Данные могут храниться в различном виде, поскольку специалисты их создавали исходя из собственных соображений, и поэтому они зачастую неструктурированы. Материалы, собранные до широкого применения персональных компьютеров, приходится переносить в БД с бумажных носителей. С этим связана вторая причина — слабая степень применения в прошлом, да и в настоящее время, формализованной схемы фиксации первичных материалов. В большей степени это касается данных наблюдений на промысле и ловушечных съемок и отчасти данных по беспозвоночным, собранным в ходе траловых съемок. Материалы научных траловых съемок, в первую очередь по рыбам, уже давно вносятся в траловые карточки и специально разработанные таблицы. Наконец, третья причина — частичная утрата первичных материалов.

Если данные по промысловым видам рыб практически полностью перенесены в БД «Морская биология», то промысловых видов крабов это коснулось в меньшей степени. Для оценки и прогнозирования запасов крабов с использованием моделей на основе функциональных групп [Ильин, Иванов, 2015, 2018; Федотов, Черниенко, 2022] необходимы сведения об уловах особей определенных размеров. В БД же для ряда лет, особенно за 1960–1990-е гг., имеются только данные об общих уловах без разделения по каким-либо группам. Материалы биологического анализа могут вовсе отсутствовать.

Можно попытаться восстановить численность размерно-половых групп в уловах, зная их относительную численность. Простейший способ — вычислить средние доли групп по имеющимся данным и умножить их на общую численность. Однако такой подход слишком огрубляет результаты и таким образом снижает эффективность использования полученной в ходе многолетних исследований информации. Известно, что пространственное распределение крабов, находящихся на различных стадиях жизненного цикла, связано с условиями окружающей среды [Слизкин, Сафронов, 2000]. Использование таких связей для восстановления численности размерно-половых групп могло бы дать более корректные результаты.

Приняв во внимание вышесказанное, мы поставили перед собой следующую цель — восстановить абсолютные значения индексов численности различных размерно-

половых групп, используя статистические связи пространственного распределения особей на разных стадиях жизненного цикла с факторами окружающей среды. В качестве примера был выбран запас краба-стригуна опилио зоны Западно-Беринговоморской.

Материалы и методы

Сбор и подготовка данных

Исходными данными послужили результаты донных траловых съемок, выполненных с 1970 по 2021 г. у западного побережья Берингова моря (рис. 1). Из данных траловых карточек, хранящихся в БД «Морская биология», была получена информация о географических координатах, глубинах, придонной температуре и общих уловах краба-стригуна опилио на каждой станции.

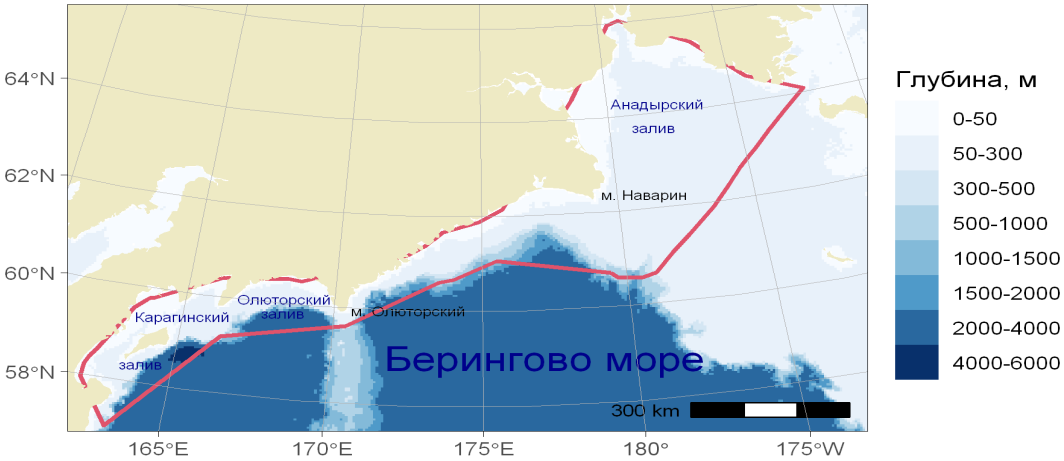

Рис. 1. Район исследований, 1970–2021 гг.
Fig. 1. The study area in 1970–2021

Использованы данные 25 траловых съемок (табл. 1), в ходе которых выполнено 2260 станций, а также материалы 25674 имевшихся в доступе биологических анализов.

Для съемок, по которым данные промеров отсутствовали, использовали файлы, подготовленные для оценки запаса в программах MapDesigner [Поляков, 1995] и «КартМастер» [Бизиков и др., 2007]. Файлы содержали данные об уловах на усилие, сгруппированные следующим образом: промысловые самцы с шириной карапакса, превышающей 100 мм (MC); предрекруты — самцы с шириной карапакса от 80 до 100 мм (MP); молодые самцы с шириной карапакса менее 80 мм (MJ); самки с икрой на плейоподах (FF) и самки без икры (FJ). В ряде случаев деление было более грубым — коммерческие самцы с шириной карапакса 100 мм и более (CM), прочие самцы (CJ) с шириной карапакса менее 100 мм, самки (F).

В случаях, когда в наличии имелись данные о биологическом анализе, размерно-половой состав делили на группы MC, MP, MJ, FF и FJ и пересчитывали их уловы, умножая общее число крабов в улове на станции на доли этих групп.

Данные об окружающей среде включали топографические характеристики рельефа дна, состав грунта, а также состав бентоса и климатические статистические характеристики воды у дна. Состав бентоса получили из материалов бентосных съемок (табл. 2). Бентос западной части Берингова моря был достаточно тщательно исследован гидробиологами ТИНРО, чтобы сделать вывод о стабильности его состава и обилия по большинству таксономических групп (за редкими исключениями, связанными среди прочего с влиянием случайности при облове мозаично расположенных локальных поселений таких видов) [Надточий и др., 2007, 2017а, б; Надточий, Колпаков, 2022]. Характеристики морфологии морского дна получены из цифровой модели рельефа (ЦМР) GEBCO_2023 (https://doi.org/10.5285/f98b053b-0cbc-6c23-e053-6c86abc0af7b),

Таблица 1

Научно-исследовательские траловые съемки, использованные в работе

Table 1

List of research trawl surveys used in the study

Год	Судно	Сроки	Широта	Долгота	Глубина, м
1970	СРТМ «Космический»	04.04–25.07	56°18′N–64°34,5′N	162°19,5′E–173°58′W	21–430
1971	СРТМ «Космический»	13.04–18.08	56°44,4′N–62°29′N	162°41′E–176°7′W	20–550
1972	НПС «Профессор Дерюгин»	03.05–08.07	59°30′N–61°55′N	168°4′E–175°51′W	65–590
	СРТМ «Пеламида»	15.07	56°37′N–65°26′N	162°13′E–174°56′W	8–740
1975	СРТМ 8–449	25.05–25.08	60°3′N–63°19′N	167°5′E–173°15′W	32–125
	СРТМ «Пеламида»	15.07	59°46′N–63°35′N	167°5′E–178°42′W	15–600
1989	НИС «Профессор Леванидов»	16.06–24.07	60°34′N–64°30′N	172°32′E–177°41,3′W	32–1000
	БМ «Барановичи»	24.06–18.07	60°49′N–64°22′N	179°55′E–175°2,3′W	55–284
	РТМС «Гиссар»	31.10–20.11	60°36′N–62°5′N	179°8′E–179°58′E	90–1125
1996	СРТМ «Шурша»	23.08–10.09	61°0′N–65°15′N	179°33′E–173°1,5′W	35–223
	НИС «Профессор Кагановский»	30.08–23.09	60°44′N–62°0′N	163°20,3′E–176°3,9′W	117–266
1998	НИС «Профессор Кагановский»	21.08–08.10	57°2′N–63°40′N	163°19′E–177°12,6′W	49–505
1999	БАТМ «Березина»	12.06–13.08	60°58′N–62°21′N	162°29,4′W–175°54,6′W	84–200
	НИС «ТИНРО»	21.08–29.10	59°48′N–64°20′N	170°39′E–176°40,02′W	37–762
2001	НИС «Профессор Кагановский»	15.07	61°30′N–62°40′N	179°54′E–173°58,1′W	87–154
	НИС «Профессор Кагановский»	19.08–31.10	57°58′N–65°13′N	162°12′E–180°0′	22–500
2005	НИС «ТИНРО»	07.08–12.10	60°23′N–66°2,8′N	171°6,1′E–180°0′	21–773
	РС «Щапино»	10.08–30.09	59°0′N–65°58′N	170°0′E–180°0′	14–265
2019	НИС «Профессор Леванидов»	16.07–08.08	58°3,4′N–62°57,9′N	162°16,6′E–180°0′	18–355
	НИС «Профессор Леванидов»	29.07–01.08	61°20,7′N–62°4,1′N	173°22,3′E–176°3,9′E	31–35
2020	НИС «Песков»	04.08–05.10	58°1,5′N–64°41,5′N	162°15′E–180°0′	10–204
	НИС «Профессор Кагановский»	08–25.09	59°4,2′N–61°45,3′N	165°52,8′E–176°40,1′E	22–952
	НИС «ТИНРО»	15–22.09	60°41′N–62°39,8′N	177°6′E–180°0′	98–919
2021	НИС «Профессор Кагановский»	31.08–17.09	61°0,5′N–65°14,1′N	174°16,3′E–180°0′	23–281
	СРТМ «Валерий Маслаков»	15.09–10.10	57°58,7′N–63°21,3′N	162°29,4′E–180°0′	20–370

данные о статистических характеристиках (минимум, максимум, среднее и разброс) придонной температуры, солености, скорости течений, концентрациях нитратов, фосфатов, силикатов, растворенного кислорода, железа, хлорофилла, обилия фитопланктона, первичной продукции и интенсивности света у дна — из системы Bio-ORACLE [Tyberghein et al., 2012; Assis et al., 2018].

Таблица 2

Научно-исследовательские дночерпательные съемки, использованные в работе

Table 2

List of research bottom sampler surveys used in the study

Год	Судно	Сроки	Широта	Долгота	Глубина, м
1983	БМРТ «Мыс Юноны»	31.08–19.09	57°49,8′N–60°15′N	162°7,8′E–166°30′E	20–190
1984	БМРТ «Мыс Тихий»	04–11.06	49°30′N–50°46,8′N	155°43,2′E–157°10,8′E	30–235
1985	БМРТ «Мыс Тихий»	31.08–25.09	61°5,4′N–65°22,2′N	179°0′E–173°0′W	13–250
2001	НИС «Профессор Кагановский»	11.08–15.10	57°49,8′N–65°40,2′N	162°7,2′E–169°51′W	20–250
2005	НИС «ТИНРО»	10.08–14.10	59°49,2′N–65°16,2′N	170°28,8′E–175°0′W	18–201
2012	НИС «Профессор Кагановский»	28–31.08	59°40,2′N–60°18,6′N	166°34,8′E–169°31,2′E	51–270

Конструирование признаков

Под конструированием признаков (Feature engineering) понимают выбор формы представления и преобразования необработанных данных с целью формирования входных переменных [Müller, Guido, 2016]. Конструирование признаков может включать множество различных процедур — от простого использования исходной информации, как есть, до получения некоторых производных величин [Черниенко, Черниенко, 2021], вплоть до сложных модельных расчетов, позволяющих на основе спутниковых данных получить динамические характеристики водных масс [Kulik et al., 2022].

В качестве исходных сведений использованы координаты с десятичными долями градусов, даты тралений, состав грунтов, придонные температуры на траловых станция в градусах Цельсия, общие уловы крабов на час траления, уловы крабов по размерно-половым группам на час траления, размерно-половой состав крабов из траловых уловов в долях единицы, плотности поселений донных организмов по данным траловых и дночерпательных съемок в граммах на квадратный метр, глубины в метрах, статистические характеристики гидрологических параметров.

В БД «Морская биология» координаты хранятся в формате градусов с десятичными долями. Долготу Восточного полушария отсчитывают в восточном направлении от Гринвичского меридиана, долготу Западного полушария, соответственно, к западу. При этом градусы западной долготы записывают со знаком «минус». Использование такого обозначения создаст разрыв по 180-му меридиану, этот разрыв может серьезно исказить результаты вычислений. Для того чтобы получить непрерывную однонаправленную величину, градусы западной долготы преобразовали в градусы на сфере:

$$Lon = 360 - Lon_w, \quad (1)$$

где Lon — долгота в градусах с десятичными долями на сфере; Lon_w — градусы западной долготы с десятичными долями.

Процессы, связанные с жизненным циклом организмов, обитающих в умеренных и крайних широтах, такие как линька, нерест, миграции и т.д., обычно имеют ярко выраженный циклический характер. В связи с этим даты тралений пересчитывали на

день года и вычисляли циклические компоненты дня года как переменные, связанные со временем:

$$d_{sin} = sin2\pi \frac{D}{D_Y},$$
$$d_{cos} = cos2\pi \frac{D}{D_Y}, \quad (2)$$

где D — порядковый номер дня года; D_Y — количество дней в году (365 или 366). Использование такого преобразования предотвращает разрыв между началом и концом года.

Пропуски в данных бентосных съемок заполняли с помощью процедуры множественного заполнения пропусков последовательными уравнениями (MICE) [Кулик, Горюнов, 2022], основанными на модели случайного леса [Breiman, 2001]. Среди заполняемых в MICE пропусков были плотности бентосных организмов и коды типов грунта.

На основе данных ЦМР рассчитывали морфометрические характеристики дна: восточно-западная и северо-южная составляющие экспозиции, наклона, вогнутость морского дна, кривизна в плане и профиле. Расчеты выполнены в ГИС SAGA версии 7.2.0. Зональная статистика со всех новых слоев ГИС собрана взвешенно по площади снятия ячеек. Площадь дна рассчитана в модуле для ГИС SAGA на основе ЦМР [Макрофауна..., 2014]. Сетка агрегирования значений образована гексагонами шириной примерно 3,5 км, высотой 3,0 км и площадью 7,8 км^2 в равновеликой азимутальной проекции Ламберта с центром по координатам 63^o с.ш. 169^o в.д. Площади ячеек, обрезанных границей района исследований, были меньше 7,8 км^2, поэтому биомасса рассчитывалась перемножением модельной плотности донных организмов в каждой ячейке на ее реальную площадь с учетом рельефа, которая у целых гексагонов в местах с наибольшим уклоном достигала 8 км^2. Значения слоев для ГИС из системы Bio-ORACLE сняты центроидами полученных гексагонов или полигонов. В рамках выбранной полигональной сетки усреднены плотности бентосных организмов.

Группа переменных, связанная с условиями окружающей среды, включала 162 параметра (рельеф дна, гидрохимическая статистика, плотности поселения видов бентоса и т.д.), для снижения размерности данной группы выполнили анализ главных компонент. Как видно на рис. 2, большая часть дисперсии может быть объяснена двумя главными компонентами (ГК).

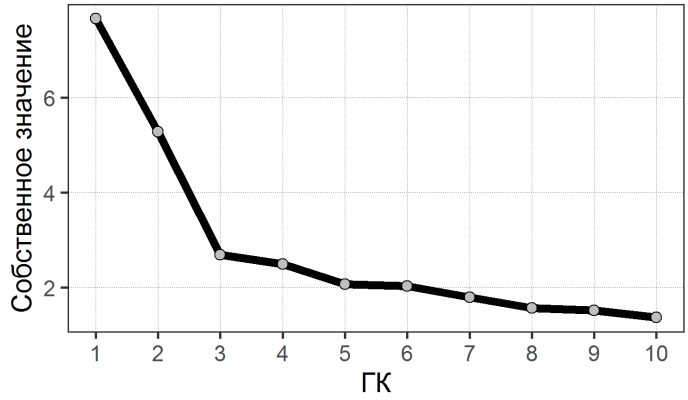

Рис. 2. Собственные значения главных компонент
Fig. 2. Eigenvalues of the principal components

В качестве объясняющей переменной использовали также общий улов крабов на станции, экз./час траления.

Таким образом, среди предикторов можно выделить группу статических (координаты, глубины, переменные, связанные со средой обитания) и динамических (придонная температура, циклические компоненты года, улов на час траления на станции) переменных.

Объясняющие переменные различной физической природы, выраженные в разных единицах измерения, имеют широкий размах значений. Это отрицательно влияет на работу многих алгоритмов машинного обучения. Для того чтобы привести входные переменные к единому масштабу, их нормировали, вычтя среднее значение и разделив на стандартное отклонение.

Переменные отклика формировали на основе уловов размерно-половых групп MC, MP, MJ, FF и FJ. Уловы CM, CJ и F использовали для валидации. Для каждой станции рассчитывали доли наблюдавшихся на них размерно-половых групп. Доли обозначали в соответствии с обозначением групп как p_{MC}, p_{MP}, p_{MJ}, p_{FF}, p_{FJ}, p_{CM}, p_{JM}, p_F. При этом следует отметить, что $p_{MC} + p_{MP} + p_{MJ} + p_{FF} + p_{FJ} = 1$, $p_{CM} + p_{CJ} + p_F = 1$, $p_{CM} = p_{MC}$, $p_{JM} = p_{MP} + p_{MJ}$ и $p_F = p_{FF} + p_{FJ}$ (табл. 3).

Таблица 3

Описание переменных, использованных в модели

Table 3

Description of the variables used in the model

Переменная	Обозначение	Единицы
Предиктор		
Широта	*Lat*	Градусы с десятичными долями
Долгота	*Lon*	Градусы с десятичными долями
Глубина	*D*	Метр
Главная компонента 1	GC1	
Главная компонента 2	GC2	
День года, синус	d_{sin}	Радиан
День года, косинус	d_{cos}	Радиан
Придонная температура	*T*	Градусы Цельсия
Улов на час траления	*U*	Экземпляр
Отклик		
Доля промысловых самцов	p_{MC}	Доля единицы
Доля предрекрутов	p_{MP}	«
Доля молодых самцов	p_{MJ}	«
Доля самок с икрой	p_{FF}	«
Доля самок без икры	p_{FJ}	«
Доля промысловых самцов*	p_{CM}	Доля единицы, $p_{CM} = p_{MC}$
Доля непромысловых самцов*	p_{JM}	Доля единицы, $p_{JM} = p_{MP} + p_{MJ}$
Доля самок*	p_F	Доля единицы, $p_F = p_{FF} + p_{FJ}$

* Используется как дополнительная величина при верификации.

Определение модели

Доли размерно-половых групп по определению являются связанными величинами, что приводит к необходимости использовать методы регрессии с многомерным откликом. В настоящей работе применяется многослойный перцептрон (multilayer perceptron, MLP) [Murtagh, 1991], для которого линейный выход преобразован функцией softmax.

Многослойный перцептрон, также известный как искусственная нейронная сеть, представляет собой систему уравнений, моделирующих соединенные и взаимодействующие между собой элементарные процессы, называемые искусственными нейронами. Искусственный нейрон — система, принимающая на вход один или более параметров с некоторыми коэффициентами (весом), суммирующих и преобразующих их некоторой функцией, называемой функцией активации:

$$y = \varphi(\Sigma w_i x_i), \tag{3}$$

где x_i — входные переменные; y — выход; w_i — вес; φ — функция активации.

Нейроны объединяются в слои. В случае однослойного перцептрона количество нейронов идентично числу входных переменных, которые преобразуются и подаются

на выход. При этом выходов может быть более одного. В случае многослойного перцептрона в его схеме присутствуют также промежуточные (скрытые) слои. Выходы промежуточного слоя служат входами для последующего.

За счет своей структуры MLP может моделировать сложные нелинейные связи. Использовали перцептрон с двумя промежуточными слоями. Первый промежуточный слой содержал 18, второй — 36 нейронов. Промежуточные слои имели кусочно-линейную функцию активации ReLu [Fukushima, 1969]:

$$\varphi(x) = \begin{cases} 0: x < 0 \\ 1: x \geq 0 \end{cases}. \quad (4)$$

Линейный выход преобразовывался многопеременной логистической функцией (Softmax) [Bridle, 1990].

$$\sigma(z)_i = \frac{e^{z_i}}{\sum_{k=1}^{K} e^{z_k}}, \quad (5)$$

где z — вектор размерности K; σ — вектор той же размерности с единичной суммой. В нашем случае $K = 5$, по числу групп. Сумма переменных на выходе функции равна единице и, таким образом, с ее помощью может быть имитировано распределение долей размерно-половых групп на станции.

Настройка модели

Исходный набор данных случайным образом разделили на обучающий (80 % от исходного объема) и тестовый (20 % от исходного объема). Данные с делением размерно-полового состава на группы CM, CJ, F использовали как отложенный (валидационный) набор данных.

Обучающий набор применяли для обучения перцептрона. Тестовый набор служил для проверки того, насколько результаты зависят от конкретных данных, поданных на вход алгоритма. Валидационный набор не содержал данных о съемках, использованных для обучения, имитируя применение модели к новым данным [Harrington, 2012].

Обучали перцептрон, минимизируя среднеквадратичную ошибку (mean squared error, MSE) истинных и оцененных моделью данных:

$$MSE = \frac{1}{n}\sum(y_i - \hat{y}_i)^2, \quad (6)$$

где y — истинные значения; $\hat{y}$ — значения, оцененные моделью.

Для минимизации целевой функции использовали алгоритм ADAM (http://arxiv.org/abs/1412.6980).

Верификация результатов

Используя обученную модель, рассчитывали доли размерно-половых групп на тестовых станциях и наносили их на точечную диаграмму. То же самое делали для валидационного набора, предварительно пересчитав доли размерно-половых групп так, как описано выше. Для истинных и расчетных долей вычисляли коэффициенты корреляции Пирсона.

Используя расчетные доли размерно-половых групп, пересчитывали наблюденные значения абсолютных общих уловов в оценке абсолютных уловов размерно-половых групп на час траления. Уловы на час траления пересчитывали на плотность [Руководство..., 1979]:

$$n_i = \frac{U_i}{КУ 1{,}82 v w_{tr} 10^{-6}}, \quad (7)$$

где i — индекс размерно-половой группы; U_i — улов на час траления на i-й станции; n_i — плотность размерно-половой группы на станции, экз./км^2; *КУ* — коэффициент уловистости; v — скорость траления, уз.; w_{tr} — ширина раскрытия трала, м. *КУ* приняли равным единице.

После этого выполняли интерполяцию пространственного распределения на регулярную сетку и оценку распределения численности размерно-половых групп как по наблюденным, так и по расчетным данным при помощи двумерного β-сплайна [Lee

et al., 1997] на площади, автоматически оконтуренной с помощью алгоритма alphahull [Rodriguez-Casal, 2007]. После этого выполняли оценку численности:

$$N_i = \sum_{c=1}^{c=NC} S_c n_{i,c}, \quad (8)$$

где N_i — численность i-й размерно-половой группы; NC — число ячеек в регулярной сетке; c — номер ячейки регулярной сетки; S_c — площадь ячейки регулярной сетки; $n_{i,c}$ — плотность i-й размерно-половой группы в ячейке c.

Следует отметить, что *КУ*, методика интерполяции и оконтуривания отличаются от традиционно используемых в Тихоокеанском филиале ВНИРО, поэтому указанные здесь оценки численности размерно-половых групп отличаются от приведенных в обоснованиях ОДУ.

Для расчетов использовали язык сценариев R с дополнительными библиотеками, для работы с данными — пакет data.table (https://CRAN.R-project.org/package=data.table). Для построения графиков — пакет ggplot2 [https://CRAN.R-project.org/package=ggOceanMaps]. Для расчета зональной статистики всех новых слоев ГИС применяли пакет exactextractr (https://CRAN.R-project.org/package=exactextractr). Определение архитектуры перцептрона и его настройку выполняли с использованием библиотеки torch (https://CRAN.R-project.org/package=torch), интерполяцию значений долей размерно-половых групп — с помощью библиотеки MBA (https://CRAN.R-project.org/package=MBA). Для работы с пространственными данными и картирования задействовали библиотеки alphahull (https://github.com/beatrizpateiro/alphahull), sf (https://CRAN.R-project.org/package=sf), raster (https://CRAN.R-project.org/package=raster), geosphere (https://CRAN.R-project.org/package=geosphere) и ggOceanMaps (https://CRAN.R-project.org/package=ggOceanMaps).

Результаты и их обсуждение

На рис. 3 изображена динамика целевой функции в процессе обучения модели. Сплошная линия показывает изменение целевой функции для обучающего набора данных, пунктирная — для тестового. С 396-й итерации целевая функция для тестовой выборки начинает расти, в то время как для обучающей продолжает падать. Это свидетельствует о том, что с этого момента модель переобучается, для обучающей выборки подгонка приближается к идеальной, в то время как для прочих данных качество подгонки снижается и модель перестает работать. Таким образом, на 396-й итерации обучение модели прекратили и использовали ее для дальнейших расчетов.

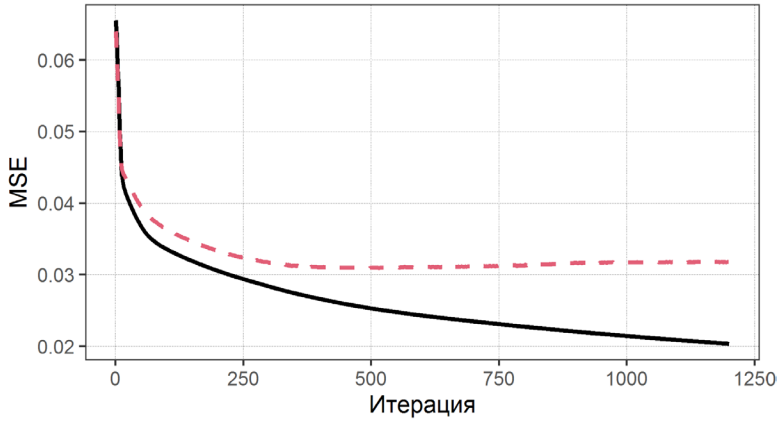

Рис. 3. Динамика изменения целевой функции в ходе обучения MLP: *сплошная линия* — для обучающей выборки, *пунктирная* — для тестовой

Fig. 3. Change of the objective function during MLP learning: *solid line* is for the training dataset, *dashed line* is for the test dataset

Для тестового и валидационного наборов данных получена отчетливая связь между наблюдениями и оценкой (табл. 4). Как видно на рис. 4 и 5, модель позволяет получить гораздо более адекватные оценки долей размерно-половых групп по сравнению с простым осреднением. Вместе с тем расчетные значения сглаживают наблюдения.

Таблица 4

Коэффициенты корреляции между наблюденными и расчетными долями размерно-половых групп для тестовой выборки
(ρ — коэффициент корреляции, p — уровень значимости)

Table 4

Correlation between observed and estimated portions of size-sex groups for the test sample
(ρ — correlation coefficient, p — significance level)

Группа	ρ	p
Тестовый набор		
MC	0,65	<< 0,001
MP	0,71	<< 0,001
MJ	0,72	<< 0,001
FF	0,44	<< 0,001
FJ	0,53	<< 0,001
Валидационный набор		
CM	0,47	<< 0,001
CJ	0,50	<< 0,001
F	0,20	<< 0,001

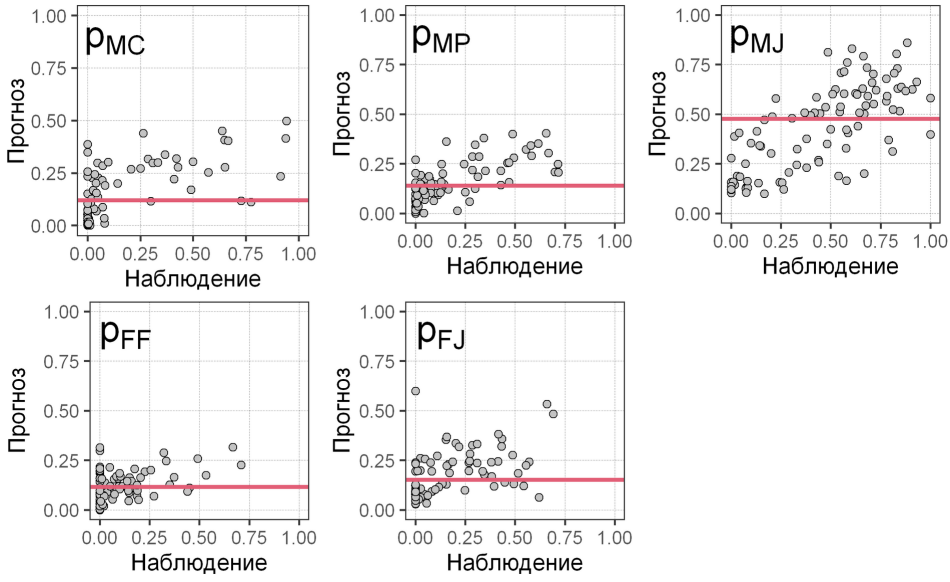

Рис. 4. Наблюденные и предсказанные моделью значения долей размерно-половых групп на станциях в тестовой выборке: *красная горизонтальная линия* — среднее значение доли размерно-половой группы

Fig. 4. Observed and modeled portions of size-sex groups by stations in the test sample. *Red horizontal line* shows the mean portion of a size-sex group

Наилучшие результаты получены для молодых самцов (MJ). Это объясняется преобладанием данной группы в уловах и, как следствие, наиболее полным охватом всего диапазона значений предикторов для нее. Сравнительно низкая предсказательная способность в отношении половозрелых самок (FF), по всей видимости, связана с тем, что данная группа включает особей с неодинаковой степенью развития икры — от недавно отложенной до выпущенных личинок. Самки на разных стадиях развитиях икры могут иметь различные требования к условиям окружающей среды [Слизкин,

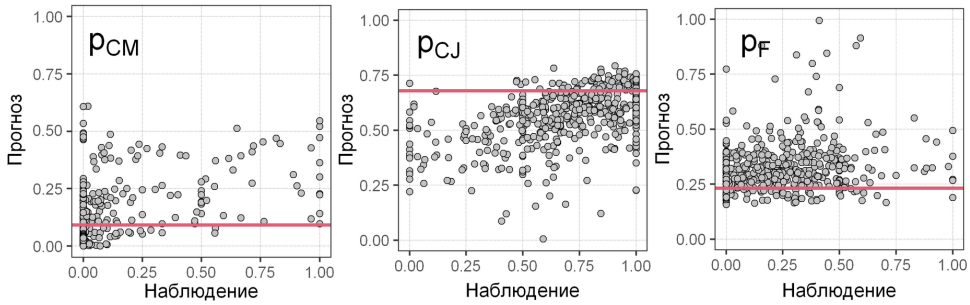

Рис. 5. Наблюденные и предсказанные моделью значения долей размерно-половых групп на станциях в валидационной выборке: *красная горизонтальная линия* — среднее значение доли размерно-половой группы

Fig. 5. Observed and modeled portions of size-sex groups by stations in the validation sample. *Red horizontal line* shows the mean portion of a size-sex group

Сафронов, 2000]. Кроме того, данная группа в целом хуже доступна для облова тралами, в связи с чем индексы ее численности подвержены наибольшим искажениям.

Аналогичным образом можно объяснить и результаты, полученные для валидационной выборки. Непромысловые самцы это, в сущности, сумма долей предрекрутов и молодых самцов, в то время как самки содержат как половозрелых самок на различных стадиях, так и неполовозрелых.

Покажем для примера распределение уловов промысловых самцов на час траления для съемки 2001 г. (рис. 6). В валидационном наборе съемка 2001 г. наиболее полно охватывает район исследований. Как видно, восстановленные уловы на усилие достаточно хорошо аппроксимируют наблюдения.

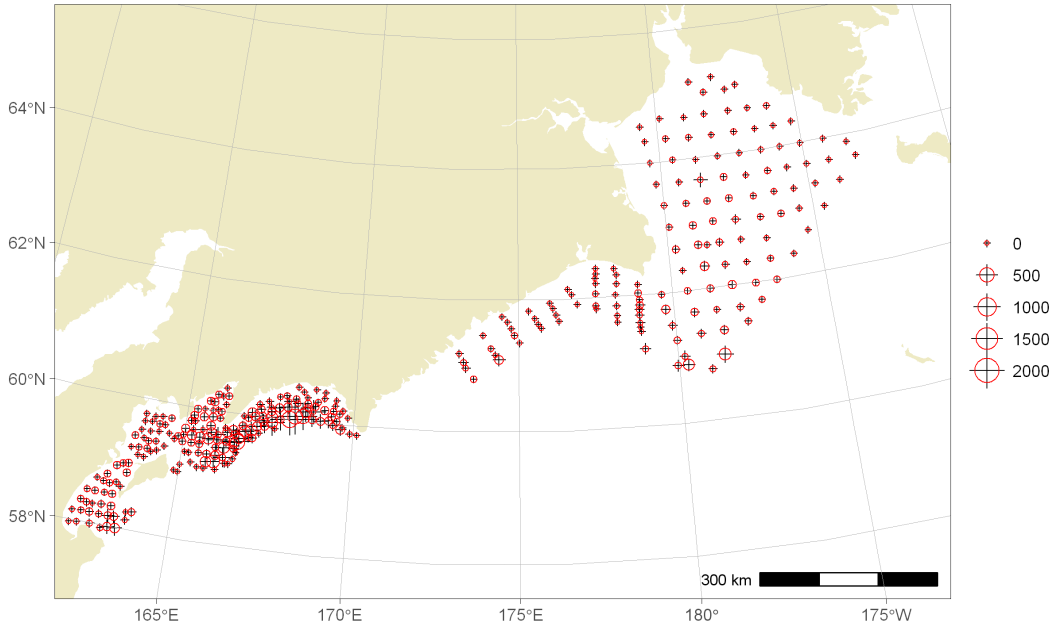

Рис. 6. Распределение уловов на усилие промысловых самцов краба-стригуна опилио (экз./час траления) в 2001 г.: + — наблюденные значения, ○ — восстановленные значения

Fig. 6. Spatial distribution of catches per unit effort for commercial-sized males of snow crab in 2001, ind. per hour of trawling: + — observed values, ○ — reconstructed values

На картах распределения численности размерно-половых групп, построенных по наблюденным и восстановленным данным (рис. 7), видно, что особенности рас-

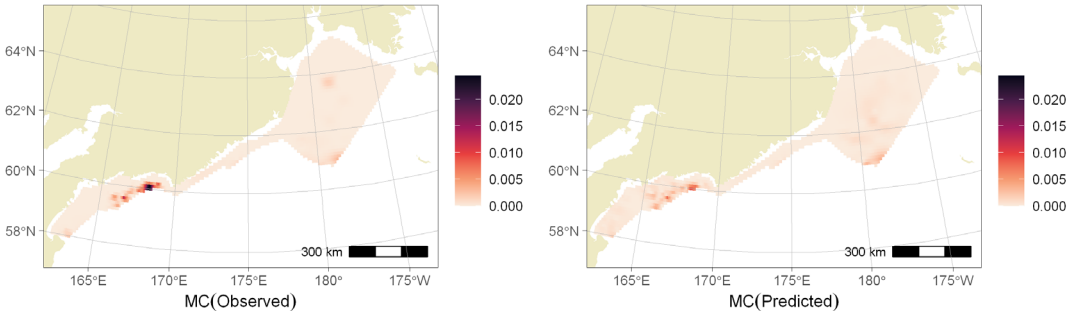

Рис. 7. Распределение плотности скоплений промысловых самцов (млн экз./км^2) по материалам учетной съемки 2001 г.: **слева** — по наблюденным, **справа** — по восстановленным данным

Fig. 7. Density distribution of commercial males on the data of 2001 survey, 10^6 ind./km^2: **left panel** — observed values, **right panel** — reconstructed values

пределения в целом воспроизводятся. При этом карта плотности распределения промысловых самцов, построенная на основе восстановленных уловов при одних и тех же параметрах интерполяции, получилась более «размазанной», т.е. более сглаженной. Численность на основе наблюдений оценена в 19,21 млн экз., на основе восстановленных уловов — 19,60 млн экз.

В табл. 5–6 и на рис. 8–9 приведены оценки численности размерно-половых групп для съемок, вошедших в обучающий и валидационный наборы. Видно, что

Таблица 5

Оценки численности по наблюденным и предсказанным моделью индексам уловов размерно-половых групп для обучающего набора данных, млн экз.

Table 5

Estimates of the snow crab abundance based on observed and modeled indices of size-sex groups for the training dataset, 10^6 ind.

Год	N_{St}	S	MC_O	MS_E	MP_O	MP_E	MJ_O	MJ_E	FF_O	FF_E	FJ_O	FJ_E
1970	98	151,83	1,39	0,51	2,38	2,58	74,38	70,68	1,00	9,58	34,30	37,21
1971	109	58,62	1,02	1,62	0,24	1,30	7,04	7,56	1,83	1,08	7,43	5,41
1972	77	159,11	1,78	1,61	1,28	3,05	271,69	240,66	26,59	18,42	74,25	94,87
1975	244	71,01	0,09	0,10	0,57	0,42	3,26	3,29	1,28	0,70	0,07	0,58
2019	202	81,61	3,28	5,64	3,03	3,86	14,07	18,03	24,27	8,60	2,96	8,24
2020	328	195,88	10,18	10,96	24,71	27,40	242,79	256,55	19,44	17,84	143,86	136,84
2021	341	176,31	10,94	9,66	13,56	14,08	109,36	80,25	16,68	7,08	15,11	14,24

Примечание. Здесь и в табл. 6: N_{St} — число станций; S — обследованная площадь; нижний индекс «O» — наблюдения; нижний индекс «E» — предсказания модели.

Таблица 6

Оценки численности по наблюденным и предсказанным моделью индексам уловов размерно-половых групп для валидационного набора данных, млн экз.

Table 6

Estimates of the snow crab abundance based on observed and modeled indices of size-sex groups for the validation dataset, 10^6 ind.

Год	N_{St}	S	CM_O	CM_E	CJ_O	CJ_E	F_O	F_E
1989	139	100,61	2,58	0,84	77,86	77,95	33,36	31,64
1996	39	93,15	0,03	0,97	763,20	696,84	142,70	203,50
1998	59	96,39	1,38	4,01	77,32	74,49	53,08	40,68
1999	135	133,54	4,60	3,00	851,14	781,57	142,54	213,32
2001	323	192,60	19,21	19,60	926,22	1063,34	514,81	374,27
2005	440	174,36	4,22	8,64	350,57	441,99	168,26	129,60

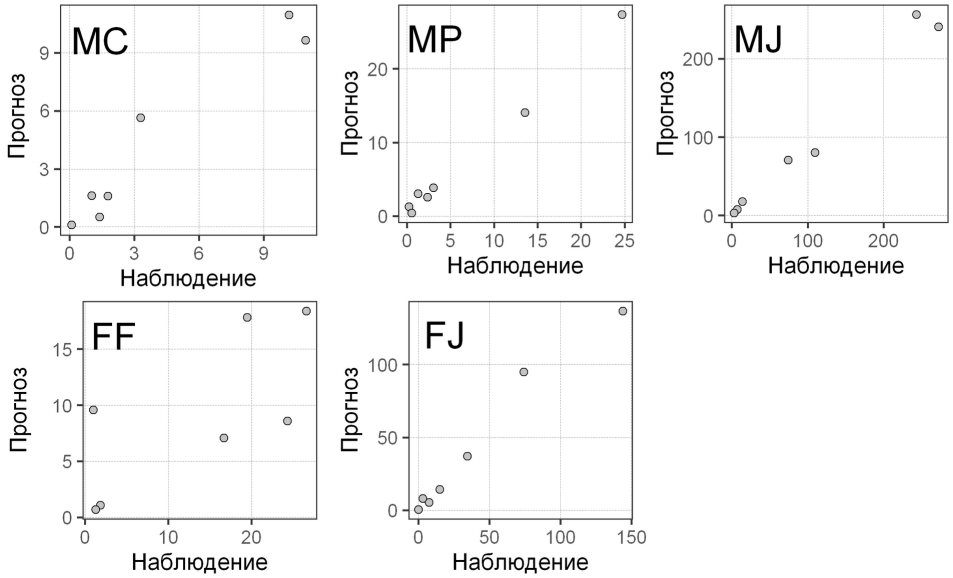

Рис. 8. Сравнение численности, оцененной по наблюденным и предсказанным моделью индексам уловов размерно-половых групп, для обучающего набора данных

Fig. 8. Comparison of the snow crab abundance estimations based on observed and modeled indices of size-sex groups catch for the training dataset

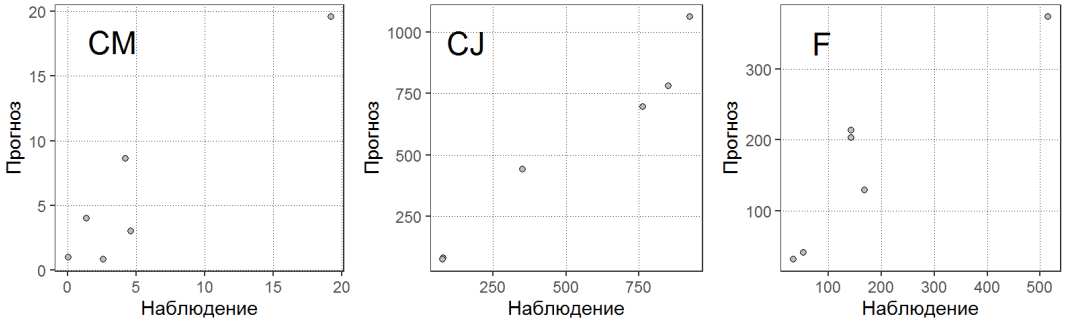

Рис. 9. Сравнение численности, оцененной по наблюденным и предсказанным моделью индексам уловов размерно-половых групп, для валидационного набора данных

Fig. 9. Comparison of the snow crab abundances estimations based on observed and modeled indices of size-sex groups catch for the validation dataset

в ряде случаев могут происходить значительные расхождения. Степень расхождения наблюдения и прогноза показывает связь с общей численностью группы, а также обследованной площадью и степенью охвата района обитания. Для высокой численности отмечается практически полное совпадение, в то время как в области малой численности может быть расхождение наблюдения и прогноза в 2–3 раза. Для наименьшей оценки численности промысловых самцов (1996 г.) расхождение составило 32 раза. Кроме того, довольно значительные расхождения наблюдаются для половозрелых самок.

Интересно отметить, что оценки численности предрекрутов получились довольно точными. Мы не считаем расхождения в области низких численностей критичными для оценки и прогнозирования запаса с использованием модели. Это искажение вряд ли будет иметь большее значение, чем ошибка наблюдения в принципе. Кроме того, при низких значениях плотностей скоплений крабов индекс численности, полученный с помощью трала, также будет иметь значительную ошибку.

Заключение

Полученный результат имеет как теоретическое, так и практическое значение. С одной стороны, достаточно надежно оценены связи между распределением различных групп крабов, что может позволить сделать нетривиальные заключения об их биологии при соответствующей интерпретации. Метод может быт использован для моделирования связи условий обитания крабов и, вероятно, других организмов с ростом и переходом на различные стадии жизненного цикла. С другой стороны, применение описанного метода позволит довольно точно оценить численность и биомассу при условии достаточного охвата района обитания.

С практической точки зрения представленная работа позволила получить метод заполнения пропусков в данных о размерно-половом составе и определить границы его применимости. Метод будет демонстрировать существенные искажения при недостаточном охвате района обитания и при малой численности.

Заявленную нами цель, а именно восстановление абсолютной численности индексов размерно-половых групп крабов, считаем выполненной. Отмеченные нами ограничения метода побуждают к дальнейшему его совершенствованию. Прежде всего необходимо искать дополнительные подходы к конструированию предикторов и экспериментировать со структурой самой модели. Вместе с тем примененный подход может быть уже сейчас приложен и к другим единицам запаса донных гидробионтов.

Резумируя, необходимо подчеркнуть, что никакая, даже самая совершенная, модель не заменит добросовестного и регулярного сбора материала в ходе экспедиционных исследований. Именно полевые и экспедиционные работы всегда были и навсегда останутся краеугольным камнем в рациональной эксплуатации водных биологических ресурсов.

Благодарности (ACKNOWLEDGEMENTS)

Выражаем искреннюю благодарность сотрудникам ТИНРО, ВНИРО и его филиалов, на протяжении десятилетий добросовестно собиравших материалы в процессе научно-исследовательских работ и наблюдений на промысле. Мы глубоко признательны начальнику отдела промысловой статистики и баз данных Н.Н. Герасимову за его титанический труд, связанный с ведением баз данных ТИНРО.

The authors are heartily grateful to their colleagues in TINRO and other branches of VNIRO, who for decades conscientiously collected the materials on snow crab in research surveys and aboard fishing vessels, and particularly thankful to N.N. Gerasimov, head of the Fishery Statistics and Databases department, for his titanic labour in maintaining the databases of the Pacific branch of VNIRO (TINRO).

Финансирование работы (FUNDING)

Исследование не имело спонсорской поддержки.

The study was not supported by sponsors.

Соблюдение этических стандартов (COMPLIANCE WITH ETHICAL STANDARDS)

Все применимые международные, национальные и/или институциональные принципы использования животных были соблюдены. Авторы заявляют, что у них нет конфликта интересов.

All applicable international, national and/or institutional and ethical guidelines are implemented. The authors declare that they have no conflict of interest.

Список литературы

Бизиков В.А., Гончаров С.М., Поляков А.В. Географическая информационная система «Картмастер» // Рыб. хоз-во. — 2007. — № 1. — С. 96–99.

Ильин О.И., Иванов П.Ю. К оценке состояния запасов краба-стригуна бэрди Камчатско-Курильской подзоны // Исслед. водн. биол. ресурсов Камчатки и сев.-зап. части Тихого океана. — 2018. — Вып. 50. — С. 27–33. DOI: 10.15853/2072-8212.2018.50.27-33.

Ильин О.И., Иванов П.Ю. Об одном модельном подходе к оценке состояния запасов камчатского краба *Paralithodes camtschaticus* западнокамчатского шельфа // Изв. ТИНРО. — 2015. — Т. 182. — С. 38–47. DOI: 10.26428/1606-9919-2015-182-38-47.

Кулик В.В., Горюнов М.И. Применение метода машинного обучения для оценки биомассы трески в Северо-Курильской зоне // Изв. ТИНРО. — 2022. — Т. 202, вып. 4. — С. 1002–1014. DOI: 10.26428/1606-9919-2022-202-1002-1014. EDN: IAVNBZ.

Макрофауна бентали западной части Берингова моря: таблицы встречаемости, численности и биомассы. 1977–2010 / В.П. Шунтов, И.В. Волвенко, В.В. Кулик, Л.Н. Бочаров; под ред. В.П. Шунтова и Л.Н. Бочарова. — Владивосток : ТИНРО-центр, 2014. — 803 с.

Надточий В.А., Будникова Л.Л., Безруков Р.Г. Некоторые результаты бонитировки бентоса в российских водах дальневосточных морей: состав и количественное распределение (Охотское море) // Изв. ТИНРО. — 2007. — Т. 149. — С. 310–337.

Надточий В.А., Колпаков Н.В. Макрозообентос Олюторского залива (Беринговo море) четверть века спустя: состав, распределение, сообщества // Изв. ТИНРО. — 2022. — Т. 202, вып. 1. — С. 161–171. DOI: 10.26428/1606-9919-2022-202-161-171.

Надточий В.А., Колпаков Н.В., Корнейчук И.А. Распределение таксонов макрозообентоса — потенциальных индикаторов уязвимых морских экосистем в западной части Берингова моря. 2. Чукотский и корякский районы // Изв. ТИНРО. — 2017а. — Т. 190. — С. 177–195. DOI: 10.26428/1606-9919-2017-190-177-195.

Надточий В.А., Колпаков Н.В., Корнейчук И.А. Распределение таксонов макрозообентоса — потенциальных индикаторов уязвимых морских экосистем в западной части Берингова моря. 1. Анадырский район // Изв. ТИНРО. — 2017б. — Т. 189. — С. 156–170. DOI: 10.26428/1606-9919-2017-189-156-170.

Поляков А.В. Программа построения карт распределения запаса и планирования съемки. — М. : ВНИРО, 1995. — 46 с.

Руководство по изучению десятиногих ракообразных Decapoda дальневосточных морей / сост. В.Е. Родин, А.Г. Слизкин, В.И. Мясоедов и др. — Владивосток : ТИНРО, 1979. — 59 с.

Слизкин А.Г., Сафронов С.Г. Промысловые крабы прикамчатских вод : моногр. — Петропавловск-Камчатский : Северная Пацифика, 2000. — 180 с.

Федотов П.А., Черниенко И.С. Динамика численности синего краба (*Paralithodes platypus* Brandt, 1849) в северо-западной части Берингова моря // Изв. ТИНРО. — 2022. — Т. 202, вып. 2. — С. 332–342. DOI: 10.26428/1606-9919-2022-202-332-342. EDN: DAYGOT.

Черниенко Э.П., Черниенко И.С. Информационное сопровождение промысла японской скумбрии *Scomber japonicus* в тихоокеанских водах Российской Федерации // Изв. ТИНРО. — 2021. — Т. 201, вып. 2. — С. 390–399. DOI: 10.26428/1606-9919-2021-201-390-399.

Assis J., Tyberghein L., Bosch S. et al. Bio-ORACLE v2.0: Extending marine data layers for bioclimatic modelling // Global Ecol Biogeogr. — 2018. — Vol. 27, Iss. 3. — P. 277–284. DOI: 10.1111/geb.12693.

Breiman L. Random forests // Mach. Learn. — 2001. — Vol. 45, Iss. 1. — P. 5–32. DOI: 10.1023/A:1010933404324.

Bridle J.S. Probabilistic Interpretation of Feedforward Classification Network Outputs, with Relationships to Statistical Pattern Recognition // Neurocomputing. — 1990. — Vol. 68. — P. 227–236.

Fukushima K. Visual Feature Extraction by a Multilayered Network of Analog Threshold Elements // IEEE Trans. Syst. Sci. Cyber. — 1969. — Vol. 5, № 4. — P. 322–333. DOI: 10.1109/TSSC.1969.300225.

Harrington P. Machine Learning in Action. — N.Y. : Manning, 2012. — 384 p.

Kulik V.V., Prants S.V., Uleysky M.Yu., Budyansky M.V. Lagrangian characteristics in the western North Pacific help to explain variability in Pacific saury fishery // Fish. Res. — 2022. — Vol. 252. — 106361. DOI: 10.1016/j.fishres.2022.106361.

Lee S., Wolberg G., Shin S.Y. Scattered data interpolation with multilevel B-splines // IEEE Trans. Visual. Comput. Graphics. — 1997. — Vol. 3, № 3. — P. 228–244. DOI: 10.1109/2945.620490.

Müller A.C., Guido S. Introduction to machine learning with Python: a guide for data scientists. 1st edition. — Sebastopol, CA : O'Reilly Media, Inc, 2016. — 376 p.

Murtagh F. Multilayer perceptrons for classification and regression // Neurocomputing. — 1991. — Vol. 2, № 5–6. — P. 183–197. DOI: 10.1016/0925-2312(91)90023-5.

Rodriguez-Casal A. Set estimation under convexity type assumptions // Annales de l'Institut Henri Poincare (B) Probability and Statistics. — 2007. — Vol. 43, № 6. — P. 763–774. DOI: 10.1016/j.anihpb.2006.11.001.

Tyberghein L., Verbruggen H., Pauly K. et al. Bio-ORACLE: a global environmental dataset for marine species distribution modelling // Global Ecology and Biogeography. — 2012. — Vol. 21, № 2. — P. 272–281. DOI: 10.1111/J.1466-8238.2011.00656.X.

References

Bizikov, V.A., Goncharov, S.M., and Polyakov, A.V., The geographical informational system CardMaster, *Rybn. Khoz.*, 2007, no. 1, pp. 96–99.

Ilyin, O.I. and Ivanov, P.Yu., To the stock abundance assessment of Tanner crab in the Kamchatka-Kurile subzone, *Issled. Vodn. Biol. Resur. Kamchatki Sev.-Zapadn. Chasti Tikhogo Okeana*, 2018, vol. 50, pp. 27–33. doi 10.15853/2072-8212.2018.50.27-33

Ilyin, O.I. and Ivanov, P.Yu., On one model approach to stock assessment for red king crab *Paralithodes camtschaticus* on the shelf of West Kamchatka, *Izv. Tikhookean. Nauchno-Issled. Inst. Rybn. Khoz. Okeanogr.*, 2015, vol. 182, pp. 38–47. doi 10.26428/1606-9919-2015-182-38-47

Kulik, V.V. and Goryunov, M.I., Application of the machine learning method to estimate the biomass of pacific cod in the North Kuril zone, *Izv. Tikhookean. Nauchno-Issled. Inst. Rybn. Khoz. Okeanogr.*, 2022, vol. 202, no. 4, pp. 1002–1014. doi 10.26428/1606-9919-2022-202-1002-1014. EDN: IAVNBZ.

Shuntov, V.P., Volvenko, I.V., Kulik, V.V., and Bocharov, L.N., *Makrofauna bentali zapadnoi chasti Beringova morya: tablitsy vstrechaemosti, chislennosti i biomassy. 1977–2010* (Benthic Macrofauna of the Western Part of the Bering Sea: Occurrence, Abundance, and Biomass. 1977–2010), Shuntov, V.P. and Bocharov, L.N., Eds., Vladivostok: TINRO-Tsentr, 2014.

Nadtochy, V.A., Budnikova, L.L., and Bezrukov, R.G., Some results of benthos valuation in Russian waters of the Far Eastern Seas: composition and quantitative distribution (Okhotsk Sea), *Izv. Tikhookean. Nauchno-Issled. Inst. Rybn. Khoz. Okeanogr.*, 2007, vol. 149, pp. 310–337.

Nadtochy, V.A., and Kolpakov, N.V., Macrozoobenthos of the Olyutorsky Bay(Bering Sea) a quarter of century later: composition, distribution, communities, *Izv. Tikhookean. Nauchno-Issled. Inst. Rybn. Khoz. Okeanogr.*, 2022, vol. 202, no. 1, pp. 161–171. doi 10.26428/1606-9919-2022-202-161-171

Nadtochy, V.A., Kolpakov, N.V., and Korneichuk, I.A., Distribution of macrozoobenthic taxa — potential indicators of vulnerable marine ecosystems in the western part of Bering Sea. 2. Chukotka and Koryak districts, *Izv. Tikhookean. Nauchno-Issled. Inst. Rybn. Khoz. Okeanogr.*, 2017, vol. 190, pp. 177–195. doi 10.26428/1606-9919-2017-190-177-195

Nadtochy, V.A., Kolpakov, N.V., and Korneichuk, I.A., Distribution of macrozoobenthic taxa — potential indicators of vulnerable marine ecosystems in the western part of Bering Sea. 1. Anadyr Bay area, *Izv. Tikhookean. Nauchno-Issled. Inst. Rybn. Khoz. Okeanogr.*, 2017, vol. 189, pp. 156–170.

Polyakov, A.V., *Programma postroyeniya kart raspredeleniya zapasa i planirovaniya s"yemki* (The program for the construction of stock distribution maps and survey planning), Moscow: VNIRO, 1995.

Rodin, V.E., Slizkin, A.G., Myasoedov, V.I., Barsukov, V.N., Miroshnikov, V.V., Zgurovskii, K.A., Kanarskii, O.A., and Fedoseev, V.Ya., *Rukovodstvo po izucheniyu desyatinogikh rakoobraznykh Decapoda dal'nevostochnykh morei* (Guide to the Study of Decapods Crustaceans, Decapoda, in Far Eastern Seas), Vladivostok: TINRO, 1979.

Slizkin, A.G. and Safronov, S.G., *Promyslovye kraby prikamchatskikh vod* (Commercial Crabs of Kamchatkan Coastal Waters), Petropavlovsk-Kamchatsky: Severnaya Patsifika, 2000.

Fedotov, P.A. and Chernienko, I.S., The Population Dynamics of the Blue Crab (*Paralithodes platypus* Brandt, 1849) in the Northwestern Bering Sea, *Russ. J. Mar. Biol.,* 2022, vol. 48, no. 7, pp. 671–677. doi 10.1134/S1063074022070069

Chernienko, E.P. and Chernienko, I.S., Information support for chub mackerel Scomber japonicus fishery in the Pacific waters of the Russian Federation, *Izv. Tikhookean. Nauchno-Issled.Inst. Rybn. Khoz. Okeanogr.,* 2021, vol. 201, no. 2, pp. 390–399. doi 10.26428/1606-9919-2021-201-390-399

Assis, J., Tyberghein, L., Bosch, S., Verbruggen, H., Serrao, E.A., and Clerck, O., Bio-ORACLE v2.0: Extending marine data layers for bioclimatic modeling, *Global Ecol Biogeogr.,* 2018, vol. 27, no. 3, pp. 277–284. doi 10.1111/geb.12693

Breiman, L., Random forests, *Mach. Learn.*, 2001, vol. 45, no. 1, pp. 5–32. doi 10.1023/A:1010933404324

Bridle, J.S., Probabilistic Interpretation of Feedforward Classification Network Outputs, with Relationships to Statistical Pattern Recognition, *Neurocomputing*, 1990, vol. 68, pp. 227–236.

Fukushima, K., Visual Feature Extraction by a Multilayered Network of Analog Threshold Elements, *IEEE Trans. Syst. Sci. Cyber.*, 1969, vol. 5, no. 4, pp. 322–333. doi 10.1109/TSSC.1969.300225

Harrington, P., *Machine Learning in Action*, New York: Manning, 2012.

Kulik, V.V., Prants, S.V., Uleysky, M.Yu., and Budyansky, M.V., Lagrangian characteristics in the western North Pacific help to explain variability in Pacific saury fishery, *Fish. Res.*, 2022, vol. 252, 106361. doi 10.1016/j.fishres.2022.106361

Lee, S., Wolberg, G., and Shin, S.Y., Scattered data interpolation with multilevel B-splines, *IEEE Trans. Visual. Comput. Graphics*., 1997, vol. 3, no. 3, pp. 228–244. doi 10.1109/2945.620490

Müller, A.C. and Guido, S., *Introduction to machine learning with Python: a guide for data scientists*, Sebastopol, CA: O'Reilly Media, Inc, 2016.

Murtagh, F., Multilayer perceptrons for classification and regression, *Neurocomputing*, 1991, vol. 2, no. 5–6, pp. 183–197. doi 10.1016/0925-2312(91)90023-5

Rodriguez-Casal, A., Set estimation under convexity type assumptions, *Annales de l'Institut Henri Poincare (B) Probability and Statistics*, 2007, vol. 43, no. 6, pp. 763–774. doi 10.1016/j.anihpb.2006.11.001

Tyberghein, L., Verbruggen, H., Pauly, K., Troupin, C., Mineur, F., and De Clerck, O., Bio-ORACLE: a global environmental dataset for marine species distribution modeling, *Global Ecology and Biogeography*, 2012, vol. 21, no. 2, pp. 272–281. doi 10.1111/j.1466-8238.2011.00656.x

The GEBCO_2023 Grid — a continuous terrain model of the global oceans and land, NERC EDS British Oceanographic Data Centre NOC. https://doi.org/10.5285/f98b053b-0cbc-6c23-e053-6c86abc0af7b.

Kingma, D.P., Ba, J., *Adam: A Method for Stochastic Optimization.* http://arxiv.org/abs/1412.6980. Cited July 30, 2024.

data.table: Extension of `data.frame`. https://CRAN.R-project.org/package=data.table. Cited July 30, 2024.

Vihtakari, M., *ggOceanMaps: Plot Data on Oceanographic Maps using «ggplot2».* https://CRAN.R-project.org/package=ggOceanMaps. Cited July 30, 2024.

Baston, D., *exactextractr: Fast Extraction from Raster Datasets using Polygons*. https://CRAN.R-project.org/package=exactextractr. Cited July 30, 2024.

Tensors and Neural Networks with «GPU» Acceleration. https://CRAN.R-project.org/package=torch. Cited July 30, 2024.

Finley, A., Banerjee, S., Hjelle, Ø., *MBA: Multilevel B-Spline Approximation*. https://CRAN.R-project.org/package=MBA. Cited July 30, 2024.

Pateiro-Lopez, B., Rodriguez-Casal, A., *alphahull: Generalization of the convex hull of a sample of points in the plane.* https://github.com/beatrizpateiro/alphahull. Cited July 30, 2024.

Pebesma, E., *sf: Simple Features for R.* https://CRAN.R-project.org/package=sf. Cited July 30, 2024.

Hijmans, R.J., *raster: Geographic Data Analysis and Modeling.* https://CRAN.R-project.org/package=raster. Cited July 30, 2024.

Hijmans, R.J., *geosphere: Spherical Trigonometry.* https://CRAN.R-project.org/package=geosphere. Cited July 30, 2024.

Поступила в редакцию 12.09.2024 г.

После доработки 30.10.2024 г.

Принята к публикации 10.12.2024 г.

The article was submitted 12.09.2024; approved after reviewing 30.10.2024; accepted for publication 10.12.2024