

УДК 597.553.2-154.343 : 57.087.1

**ПРОГНОЗИРОВАНИЕ ПОДХОДОВ ЛОСОСЕВЫХ (НА ПРИМЕРЕ КИЖУЧА ЗАПАДНОЙ КАМЧАТКИ) С ИСПОЛЬЗОВАНИЕМ МОДЕЛЕЙ ЭКСТРАПОЛЯЦИИ ВРЕМЕННЫХ РЯДОВ И МОДЕЛЕЙ «ЗАПАС–ПОПОЛНЕНИЕ»****М.Г. Фельдман, Е.А. Шевляков, Ж.Х. Зорбиди**

Ст. н. с., зав. лаб., вед. н. с., Камчатский научно-исследовательский институт рыбного хозяйства и океанографии  
683000 Петропавловск-Камчатский, Набережная, 18  
Тел., факс: (4152) 41-27-01, (4152) 42-07-74  
E-mail: feldman.m.g.@kamniro.ru

ТИХООКЕАНСКИЕ ЛОСОСИ, КИЖУЧ, ПРОГНОЗИРОВАНИЕ, ВРЕМЕННЫЕ РЯДЫ, АВТОРЕГРЕССИЯ, «ЗАПАС–ПОПОЛНЕНИЕ», ARIMA, МОДЕЛЬ РАСПРЕДЕЛЕННОГО ЛАГА

В данной работе рассматривается проблема прогнозирования подходов тихоокеанских лососей. В качестве модельной популяции взято стадо кижуча *Oncorhynchus kisutch* Западной Камчатки. В регрессионном анализе использовались и сравнивались модели двух классов: модели временных рядов и модели зависимости пополнения популяции от нерестового запаса. Предлагается представлять прогноз либо в виде произведения плотностей вероятности прогнозов по различным моделям, либо как средневзвешенный результат плотностных прогнозов из нескольких регрессионных моделей.

**FORECASTING THE RUNS OF PACIFIC SALMON (FOR EXAMPLE – A COHO SALMON OF THE WESTERN KAMCHATKA) ON THE BASE OF “STOCK–RECRUITMENT” MODELS AND TIME SERIES EXTRAPOLATION SIMULATIONS****M.G. Feldman, E.A. Shevlyakov, Zh.Kh. Zorbidi**

Senior scientist, head of dep., leading scientist, Kamchatka Research Institute of Fisheries and Oceanography  
683000 Petropavlovsk-Kamchatsky, Naberezhnaya, 18  
Tel., fax: (4152) 41-27-01, (4152) 42-07-74  
E-mail: feldman.m.g.@kamniro.ru

PACIFIC SALMON, COHO SALMON, FORECASTING, TIME SERIES, AUTOREGRESSION, “STOCK–RECRUITMENT”, ARIMA, DISTRIBUTED LAG

The paper provides analysis of the issues of forecasting runs of Pacific salmon. The stock used as a basis in simulation was the West Kamchatkan stock of coho salmon *Oncorhynchus kisutch*. Simulations of two classes were compared and used in regression analysis: of time series and of the relationship between the spawning stock and the recruitment. It is suggested to perform the forecast either as a product of probability density of forecasts of different models or as an average weighted result of density-based forecasts of several regression models.

Прогнозирование динамики численности рыб, в частности тихоокеанских лососей, представляет собой важную проблему рыбного хозяйства и имеет достаточно долгую историю. Основоположником исследований популяционной динамики в рыбном хозяйстве по праву считается Ф.И. Баранов. Его также называют «дедушкой рыбных популяционных исследований» (Quinn, 2003). Наиболее известной и важной его работой является статья «К вопросу о биологических основаниях рыбного хозяйства» (1918), в которой автор впервые применяет дифференциальное счисление для моделирования динамики численности рыбного запаса. Решая уравнение:

$$dN / dt = -ZN,$$

где  $N$  — численность рыб, а  $Z=F+M$  — коэффициент смертности, складываемый из коэффициентов

промысловой  $F$  и естественной смертности  $M$ , было получено:

$$N_{t+1} = N_t e^{-Zt}.$$

Известно, что один из переводов данной работы получили Р. Бивертон и С. Холт (по данным Anderson, 2002) в 1943 г., а другой был сделан У. Рикером в 1945 г. (по данным Quinn, 2003).

В дальнейшем У. Рикер (Ricker, 1954) предложил решать задачу определения численности будущего подхода путем расчета пополнения исходя из родительского нерестового запаса и представил первую математическую модель зависимости пополнения от родительского запаса. В данном случае исходное дифференциальное уравнение для численности поколения имело вид:

$$dN / N = -(q + pS)dt, \quad (1)$$

где  $S$  — исходный нерестовый запас,  $q$  и  $p$  — неза-

висимая и зависимая от  $S$  компоненты смертности. Учитывая, что исходная численность поколения равна численности родительского запаса, умноженной на плодовитость —  $N_0 = fS$ , решая уравнение (1), получаем:

$$R = aSe^{-bS}, \quad (2)$$

где  $S$  — нерестовой запас,  $R$  — ожидаемое пополнение,  $a$  и  $b$  — коэффициенты депенсационной и компенсационной смертности соответственно.

Графиком модели (2) является куполообразная кривая, показывающая снижение пополнения при высоких значениях нерестового запаса.

Немного позже, в 1957 г., Р. Бивертон и С. Холт (Бивертон, Холт, 1969) предложили альтернативную модель, где коэффициент смертности  $Z$  зависит не от родительского запаса  $S$ , как в (1), а от текущего значения численности  $N$ :

$$dN / N = -(q + pN)dt \quad (3)$$

Решая уравнение (3) относительно численности пополнения для любого момента времени  $t$ , получим:

$$R = aS / (1 + (a/b)S) \quad (4)$$

При этом  $a$  — максимум  $R/S$  при низком запасае,  $b$  — максимум  $R$  при высоком запасае (Хилборн, Уолтерс, 2001). В отличие от кривой Рикера, кривая Бивертон-Холта представляет собой непрерывно возрастающую к горизонтальной асимптоте функцию.

В дальнейшем происходит развитие теории: Д. Кушинг (Cushing, 1973), Дж. Полик (Paulik, 1973) и Р. Деризо (Deriso, 1980) предлагают свои модели и вводят в них степенные параметры.

В 1982 г. Дж. Шепард (Sheperd, 1982), обобщая ранее полученные результаты, приходит к выводу о единой структуре всех предложенных моделей:

$$R = aSf(S/K^c),$$

где  $f(S/K^c)$  — функция убывания, зависящая от запаса, и  $K^c$  — пороговая биомасса, при превышении которой в запасае начинают преобладать зависящие от плотности процессы. В качестве такой функции Дж. Шепард, стремясь получить более гибкую кривую, использовал степенной параметр  $c$  и предложил:

$$f(S/K^c) = 1/(1+S/K^c)^c, \text{ соответственно пополнение}$$

будет равно:

$$R = aS / (1 + S / K^c)^c \quad (5)$$

При  $c < 1$  уравнение (5) аналогично модели Кушинга, при  $c = 1$  принимает вид модели Бивертон-Холта, а при  $c > 1$  модель становится аналогичной зависимости Рикера.

Более подробно с историей развития моделирования согласно теории «запас–пополнение» можно ознакомиться в работах В.К. Бабаяна (1988) и Е.А. Криксунова (1995).

Несмотря на зачастую довольно приблизительные результаты, получаемые по моделям «запас–пополнение», теория, с одной стороны, позволила установить безопасные пределы промыслового изъятия, с другой стороны, дала мощный толчок рыбохозяйственной статистике — сбору первичного материала, обязательному определению возрастного состава рыб в уловах, определению плодовитости, оценке нерестового запаса и т. п.

Таким образом, при решении задачи определения численности будущего подхода не обойтись без моделей, основанных на теории «запас–пополнение». В данной работе мы хотим сравнить результаты, полученные с помощью моделей этого класса, с другими подходами. В качестве таковых были выбраны модели, относящиеся к методам экстраполяции временного ряда и допустимые в прогнозировании динамики рыбных популяций (Бабаян, 2000), а именно модель авторегрессии и проинтегрированного скользящего среднего (*autoregression and integrated moving average — ARIMA*) и модель распределенного лага (*distributed lag model — DLM*).

Первая из них была предложена Дж. Боксом и Г. Дженкинсом (1974) и позволяет выделить в исследуемом временном ряду регулярные компоненты (тренд и циклические компоненты). Данный метод чрезвычайно популярен в эконометрических исследованиях, и практика подтвердила его мощь и гибкость (Hoff, 1983; Pankratz, 1983; Vandaele, 1983). Модель можно представить в следующем виде:

$$\Delta^d X_t = c + \sum_{i=1}^p a_i \Delta^d X_{t-i} + \sum_{j=1}^q b_j \varepsilon_{j-1} + \varepsilon_t, \quad (6)$$

где  $X_t$  — временной нестационарный ряд,  $\Delta^d$  — оператор разности временного ряда порядка  $d$  (последовательное взятие  $d$  раз разностей первого порядка — сначала от временного ряда, затем от полученных разностей первого порядка, затем от второго порядка и т. д.),  $c$  — константа,  $a_i$  и  $b_j$  — параметры авторегрессии и скользящего среднего,  $\varepsilon_t$  — ошибки модели.

Модель распределенного лага можно представить в общем случае как:

$$y_t = c + b_0 x_t + b_1 x_{t-1} + \dots + b_n x_{t-n} + \varepsilon_t, \quad (7)$$

где  $y_t$  — зависимая исследуемая переменная,  $x_t$  —

независимая объясняющая переменная,  $c$  — константа,  $\varepsilon_t$  — ошибки модели.

Применительно к популяциям кижуча Западной Камчатки, проблема прогнозирования подходов данного вида рыб встает достаточно остро.

Сравнивая динамику заполнения нерестилищ производителями кижуча с его общим подходом (рис. 1), можно отметить парадоксальную ситуацию: начиная с середины 2000-х годов, при снижающемся пропуске производителей кижуча на нерестилища, его подходы начали расти, также возрастает и интенсивность промысла. Следовательно, увеличивается и эффективность воспроизводства этой рыбы, причем многократно, особенно если учесть при этом и значительное нелегальное изъятие (Запорожец и др., 2007). Действительно, согласно теоретическим моделям зависимости пополнения от запаса, эффективность воспроизводства (показатель отношения числа рекрутов на единицу производителя) растет при снижении плотности запаса. Однако при этом общее

количество потомков *снижается*. Таким образом, объяснить значительный рост эффективности воспроизводства можно с помощью одновременного влияния двух факторов: улучшения условий внешней среды, способствующих более высокой выживаемости в посткатадромный период жизни (Шевляков и др., 2013), и недоучета производителей на нерестилищах.

В пользу первого фактора косвенно говорит, например, одновременный рост подходов других видов тихоокеанских лососей, начиная с 2006 г. В частности, у западнокамчатской горбуши подходы уверенно растут в четные годы (рис. 2); у кеты р. Большой (одной из важнейших промысловых рек Западной Камчатки) промышленный вылов также значительно увеличивается в этот же период (Заварина, 2010); значительно возросли возвраты нерки (Дубынин, 2012). При этом снижения заполнения нерестилищ у этих видов не отмечается, что также свидетельствует в пользу второго фактора — недоучета производителей кижуча. Основная при-

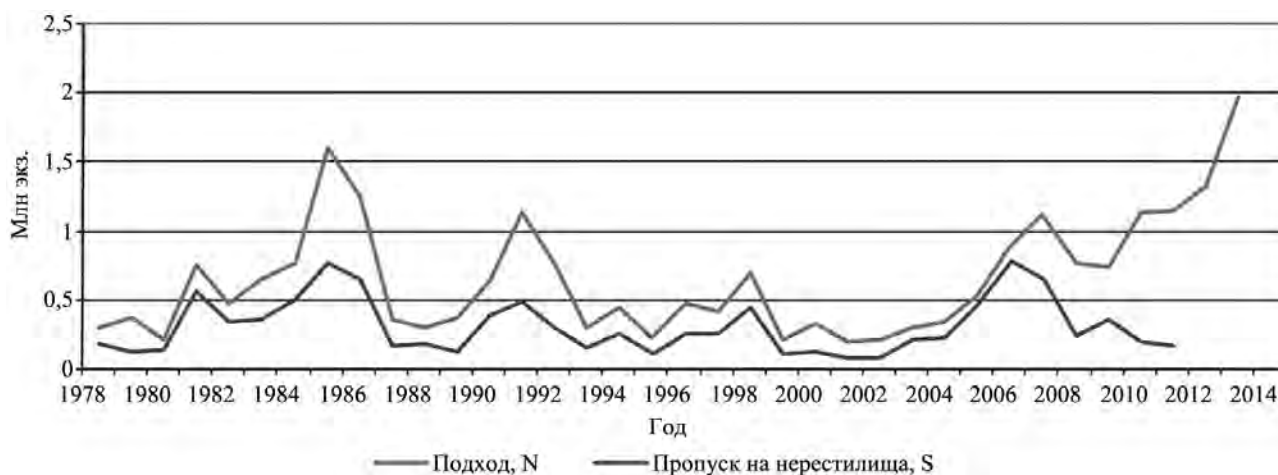


Рис. 1. Динамика подходов кижуча и его пропуска на нерестилища на Западной Камчатке



Рис. 2. Динамика подходов горбуши и ее пропуска на нерестилища на Западной Камчатке

чина такой ситуации заключается, конечно, в резком сокращении финансирования авиаучетных работ во второй половине 2000-х гг. Здесь приходится констатировать, что авиаучет кижуча — наиболее «больной» вопрос для мониторинга заходов лососей на Камчатке: сроки его нерестового хода самые поздние и растянутые. Как правило, в период проведения авиаучетных работ именно для кижуча нехватка финансирования имеет определяющее значение и обуславливает недоучет его производителей на нерестилищах. С 2006–2007 гг. общий уровень финансирования авиаучетных работ, и до этого времени дефицитный, снизился по отношению к предыдущему периоду в 1,5 раза, что резко сократило учет подходов кижуча в пользу более востребованных промыслом видов.

Таким образом, заниженная оценка числа зашедших в реки производителей влечет за собой достаточно важные последствия в оценке будущих

подходов в сторону занижения. Эту проблему можно решить двумя способами: первый — вносить экспертные корректировки в оценку родительского запаса, второй — использовать модели прогнозирования численности, не привязанные к этой величине.

Итак, цель данной работы — дать прогноз подхода лососевых, применяя несколько различных моделей, принадлежащих к разным классам. В качестве примера используются данные по популяции кижуча Западной Камчатки. Предполагается, что получив схожие либо одинаковые оценки будущих подходов, используя при этом несколько различных методов моделирования, мы будем иметь более обоснованную окончательную оценку подхода.

#### МАТЕРИАЛ И МЕТОДИКА

Исходным материалом для моделирования являлись следующие временные ряды (рис. 1–5):



Рис. 3. Колебания долевого участия возрастной группы 2+ у кижуча



Рис. 4. Колебания относительной величины R/S (количества потомков на производителя) у горбуши Западной Камчатки

- общего подхода кижуча (численность —  $N_{киж}$ );
- пропущенных на нерест производителей кижуча (родители —  $S_{киж}$ ) и обусловленного ими пополнения (потомки —  $R_{киж}$ );
- долевого участия возрастных когорт кижуча в пополнении (доли возрастов —  $C^{2+}, C^{3+}, C^{4+}$ );
- соотношения родителей и потомков горбуши (коэффициент воспроизводства  $R_{зобр}/S_{зобр}$ ).

Взаимосвязь исходных данных, методов разведочного анализа (автокорреляции и кросскорреляции), преобразования и регрессии, а также анализа ошибок, показана на схеме (рис. 6).

В работе рассматриваются четыре регрессионные модели прогнозирования подходов кижуча, которые можно разбить на два основных класса:

- модели, основанные на экстраполяции временного ряда,
- модели, основанные на теории зависимости пополнения от родительского нерестового запаса.

Из первого класса использовались модели авторегрессии и проинтегрированного скользящего среднего отклонений (ARIMA), с помощью которых строили модельный ряд подходов кижуча ( $N$ ) и расщепления возрастных когорт кижуча в пополнении. Модель была предложена Боксом и

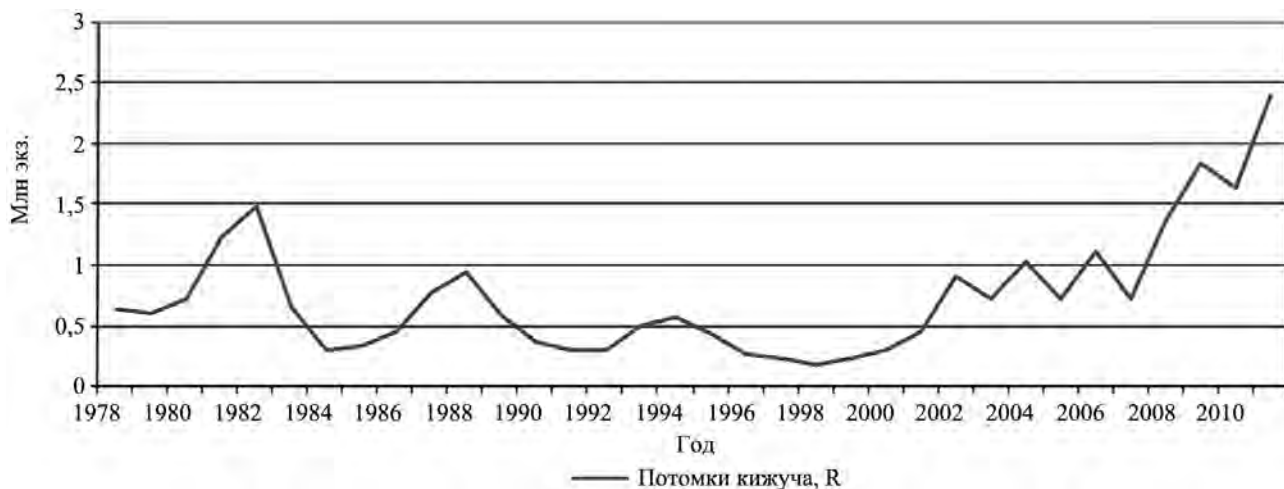


Рис. 5. Пополнение кижуча Западной Камчатки в динамике

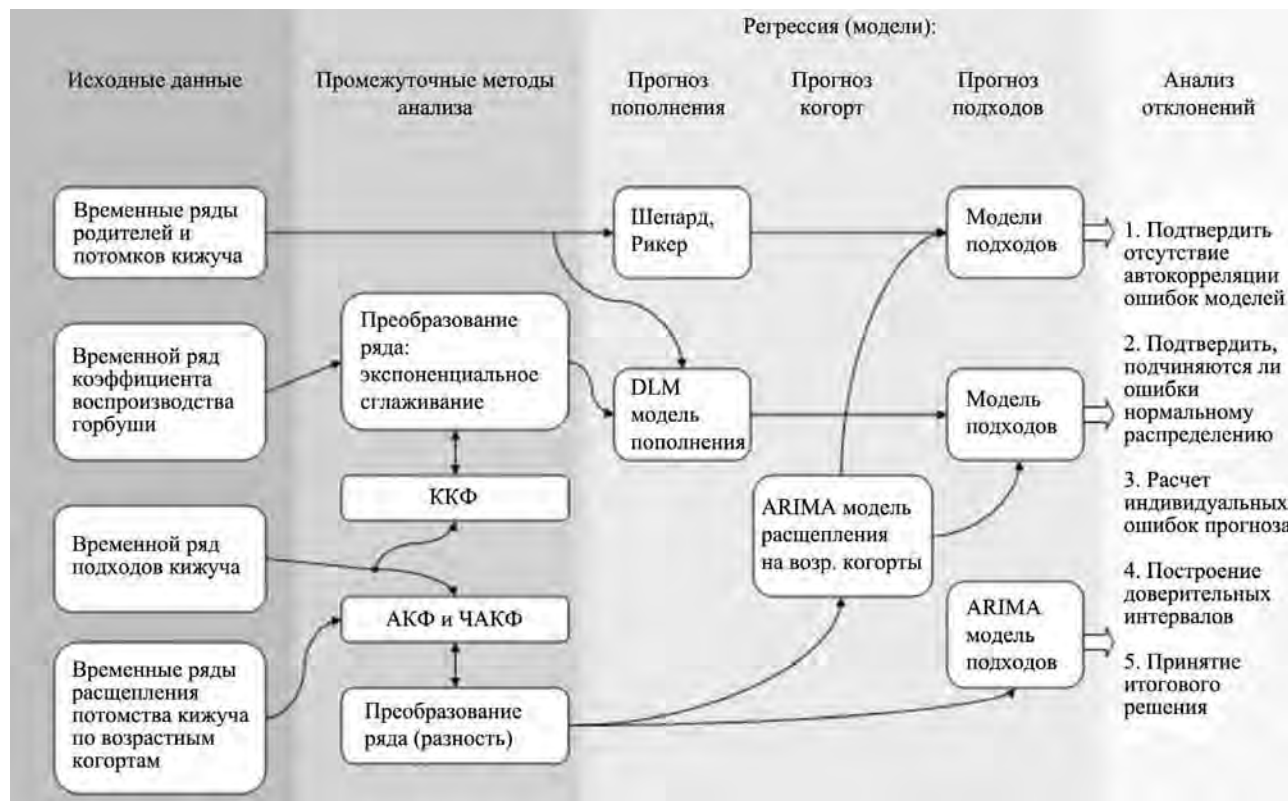


Рис. 6. Схема проведения исследований

Дженкинсом (1974) и включает в себя параметры авторегрессии и скользящего среднего. Имеется три типа параметров модели: параметры авторегрессии ( $p$ ), порядок разности ( $d$ ), параметры скользящего среднего ( $q$ ). В обозначениях Бокса и Дженкинса модель записывается как  $ARIMA(p, d, q)$ . Например, модель  $(0, 1, 2)$  содержит  $0$  параметров авторегрессии ( $p$ ) и  $2$  параметра скользящего среднего ( $q$ ), которые вычисляются для ряда после взятия разности с лагом  $1$ .

Для идентификации модели  $ARIMA$  (определения количества параметров) крайне важен предварительный анализ автокорреляций исходного и преобразованного с помощью взятия разности (дифференцированного) временного ряда. Автокорреляции представляют в виде автокорреляционной и частной автокорреляционной функций (АКФ и ЧАКФ).

АКФ показывает статистическую взаимосвязь между случайными величинами из одного ряда, но взятых со сдвигом по времени. Максимумы на автокорреляционной функции соответствуют циклическим колебаниям.

ЧАКФ является более углубленной версией обычной автокорреляционной функции. Ее отличительной особенностью является исключение корреляционной зависимости между наблюдениями внутри лагов, т. е. частная автокорреляционная функция на каждом лаге отличается от обычной автокорреляционной функции на величину удаленных автокорреляций с меньшими временными лагами. Следовательно, частная автокорреляционная функция более точно характеризует автокорреляционные зависимости внутри временного ряда.

Визуальный анализ АКФ и ЧАКФ помогает определить, какой процесс (авторегрессии или скользящего среднего) преобладает, а также количество параметров этих процессов. Если, например, АКФ имеет синусоидальный вид или экспоненциально убывает, то преобладает процесс авторегрессии. Если, наоборот, такой вид имеет ЧАКФ, то имеет место процесс скользящего среднего. В большинстве случаев имеют место смешанные процессы, а чтобы получить модель временного ряда, хватает одного-двух параметров авторегрессии и (или) одного-двух параметров скользящего среднего. Основные сочетания параметров  $p$  и  $q$  и вид соответствующих АКФ и ЧАКФ определены в (Бокс, Дженкинс, 1974).

Из класса моделей временных рядов, кроме  $ARIMA$ , рассматривается также модель распределенного лага ( $DLM$ ), с помощью которой моделируется пополнение кижуча. Эта модель учитывает наличие разных промежутков времени между разными частями зависимой и независимой переменной, где в качестве ряда независимой переменной выступает ряд отношения между потомками и родителями горбуши ( $R_{горб}/S_{горб}$ ). Об основных предпосылках выбора этой объясняющей переменной для оценки пополнения кижуча будет сказано ниже, в подразделе, рассматривающем эту модель.

Из второго класса использовали достаточно популярные модели Рикера и Шепарда.

Аппроксимация модельных значений к фактическим проводилась путем минимизации суммы квадратов отклонений. После проводились оценка значимости уравнения регрессии (с помощью анализа дисперсий) и оценка значимости модельных коэффициентов (по  $t$ -критерию Стьюдента).

Кроме регрессионного анализа, для большинства моделей использовались промежуточные и разведывательные статистические методы, такие как: преобразование рядов, автокорреляционный и кросскорреляционный анализы.

Для удобства и ускорения расчетов использовались программы:

- MS Excel (предварительная аппроксимация, графика);
- Statistica 6.0 (регрессионный анализ моделей  $DLM$ , Рикера и Шепарда, анализ ошибок);
- EViews 8 (анализ автокорреляций, моделирование и оценка моделей  $ARIMA$ , графика).

Анализ отклонений модельных данных от фактических проводился по следующим показателям:

- имеется ли автокорреляция в отклонениях модельных данных от фактических: если автокорреляция наблюдается, значит какой-либо фактор не охватывается моделью и ее необходимо доработать;
- распределению отклонений модели от фактических данных по частотному ряду. Иными словами, анализируется вид распределения ошибок, проверяется гипотеза, принадлежит ли это распределение к нормальному.

Для подтверждения гипотезы о нормальном распределении ошибок использовались следующие критерии:

– Колмогорова-Смирнова с поправкой Лиллиефорса (Lilliefors, 1967), при расчете которого находится максимальное отклонение  $d$  между выборочной и теоретической интегральными функциями распределения и проверяется нулевая гипотеза о нормальном распределении выборки;

– Жарка-Бера (Jarque, Vega, 1981), в котором сравниваются коэффициенты асимметрии и эксцесса между выборочным и нормальным распределением. Чем ближе распределение ошибок к нормальному, тем меньше данная статистика отличается от нуля. При достаточно большом значении статистики вероятность нулевой гипотезы о нормальности распределения будет мала, и тогда будет основание ее отвергнуть.

Если в отклонениях модели отсутствует автокорреляция и частотное распределение ошибок близко к нормальному, то мы можем заключить, что модель в целом адекватно описывает фактический ряд наблюдений.

Изначально предполагалось, что те модели для одного и того же фактического ряда, которые будут признаны статистически значимыми и адекватными, должны дать достаточно близкие прогнозы; как минимум, их доверительные интервалы должны перекрываться.

Таким образом, следующим шагом анализа является необходимость построения доверительного интервала прогнозных значений на заданном уровне вероятности. Для этого рассчитывались стандартные ошибки точечных прогнозов, учитывающие как неопределенности самой модели регрессии (стандартная ошибка регрессии), так и неопределенности отдельных значений факторных переменных (Брюков, 2011):

$$SE_{ind} = S \cdot \sqrt{(1 + X_t^T (X^T X)^{-1} X_t)}, \quad (8)$$

где  $X$  — матрица исходных значений факторных переменных по всему временному ряду;

$X^T$  — транспонированная матрица исходных значений факторных переменных по всему временному ряду;

$X_t$  — матрица-столбец значений факторных переменных для момента времени  $t$ ,

$X_t^T$  — транспонированная матрица-столбец значений факторных переменных для момента времени  $t$ ,

$S$  — стандартное отклонение уравнения регрессии.

Доверительный интервал точечного прогноза будет равен:

$$CI_{1-\alpha} = Y + tSE_{ind}, \quad (9)$$

где  $t$  — аргумент функции Лапласа, при  $\alpha=0,05$ ,  $t=1,96$ .

После вычисления доверительных интервалов прогнозных значений остается последний этап верификации модели — это сравнение плотности вероятности прогноза по ней с аналогичным показателем у других моделей. Если плотность вероятности прогноза такой модели не пересекается с другими плотностями (на заданном уровне доверия), то, следовательно, вес прогноза такой модели наименьший. В общем, мы проводим сравнительный анализ прогнозов по различным моделям и должны принять решение об окончательном прогнозе. Сделать это можно следующими способами:

– первый: найти средневзвешенный прогноз, т. е. сначала присвоить веса каждому прогнозу, затем найти их средневзвешенное значение;

– второй: перемножить плотности вероятности прогнозов и, следовательно, найти плотность вероятности того, что все прогнозы верны.

## РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЙ

### Авторегрессионная модель подходов кижуча (ARIMA).

Идентификация модели авторегрессии и скользящего среднего начинается с анализа АКФ и ЧАКФ. Автокорреляционная функция исходного ряда имеет синусоидальный вид с полупериодом около 11 лет, частная автокорреляционная функция имеет значимую корреляцию на единичном лаге и не имеет на других (рис. 7).

Так как исходный временной ряд имеет восходящий тренд, преобразуем его в стационарный путем взятия разности (дифференцирования ряда) первого порядка  $\Delta^1 N_t = N_t - N_{t-1}$ , где  $N_t$  — ряд подходов кижуча в год  $t$  (обозначим как ряд  $D(N,1)$ ),

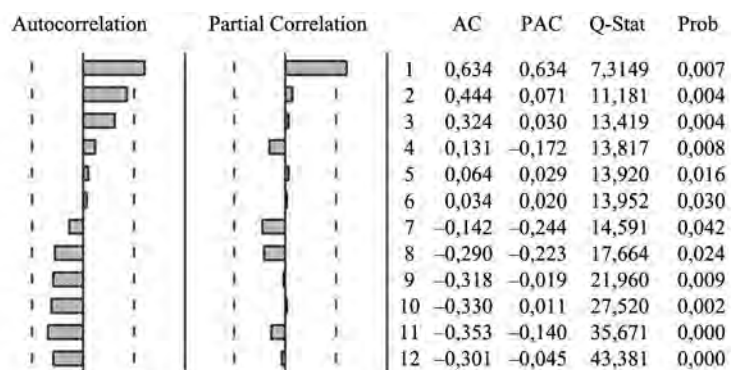


Рис. 7. АКФ и ЧАКФ исходного ряда подходов кижуча

и получим также коррелограммы преобразованного ряда. Однако АКФ и ЧАКФ не имеют значимых корреляций (рис. 8).

Возьмем еще раз разность между соседними членами ряда  $D(N,1)$ , получим ряд  $D(N,2)(\Delta^2 N_t = \Delta^1 N_t - \Delta^1 N_{t-1})$  и проанализируем коррелограммы (рис. 9). На них уже присутствуют значимые корреляции на втором лаге. Фактически значимые результаты автокорреляций ряда разности второго порядка по сравнению аналогичными для ряда с разностью первого порядка говорят о наличии нелинейного тренда в исходном ряду подходов.

Так как вид АКФ и ЧАКФ до третьего лага мало отличаются друг от друга, а на следующих лагах корреляции уменьшаются и беспорядочны, то сложно сказать, какой процесс преобладает: авторегрессии второго порядка или скользящего среднего этого же порядка. Поэтому в дальнейшем анализе идентификация модели *ARMA* идет методом перебора моделей с количеством параметров авторегрессии  $p$  и скользящего среднего  $q$ , показанных в таблице 1.

В результате наиболее значимой моделью осталась модель авторегрессии второго порядка и скользящего среднего третьего порядка:

$$\Delta^2 N_t = a \times \Delta^2 N_{t-2} + b \times \varepsilon_{t-3} + \varepsilon_t,$$

где  $a$  — параметр авторегрессии второго порядка,  $b$  — параметр скользящего среднего третьего порядка,  $\varepsilon_t$  — ошибка.

В EViews данная модель записывается как:

$$D(N,2) = AR(2) + MA(3).$$

Результаты регрессионного анализа и оценки значимости коэффициентов показаны в (табл. 2–4). Сравним фактические и модельные коррелограммы (рис. 10).

Таблица 1. Варианты моделей *ARMA*

$q \backslash p$	0	1	2	3
0	–	–	+	–
1	–	–	+	–
2	+	+	+	+
3	–	–	+	–

Таблица 2. Регрессионная статистика

Множ. R	0,836
R-квадрат	0,697
Норм. R-квадрат	0,665
Стандартная ошибка	0,185
Наблюдения	11

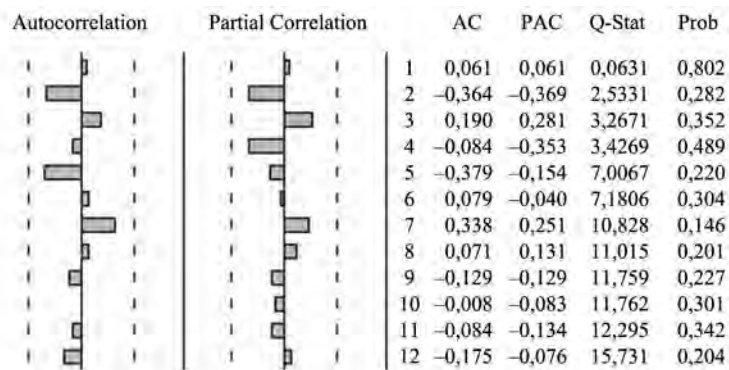


Рис. 8. АКФ и ЧАКФ ряда первой разности подходов кижуча  $D(N,1)$

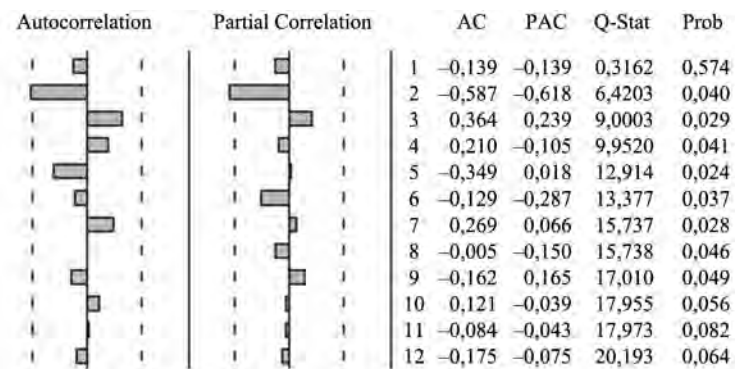


Рис. 9. АКФ и ЧАКФ ряда разности второго порядка подходов кижуча  $D(N,2)$

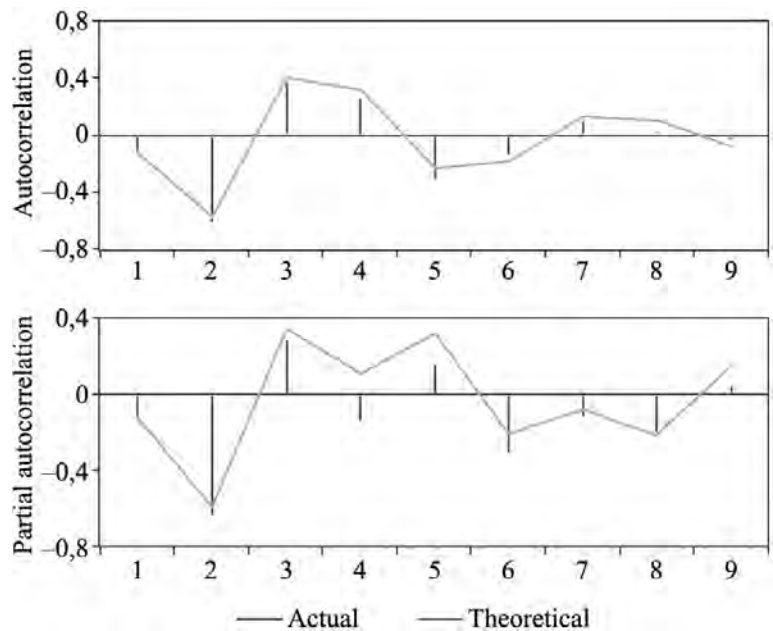


Рис. 10. Фактические и модельные коррелограммы ряда разности второго порядка подходов кижуча  $D(N,2)$



Как видно из рисунков, теоретические коррелограммы достаточно хорошо описывают фактические значения как АКФ, так и ЧАКФ.

Сравним визуально ряд и модельный ряд  $D(N,2)$ , а также ряд ошибок на рис. 11. Модель достаточно хорошо описывает фактический ряд.

Далее суммируем (интегрируем) полученный модельный ряд  $D(N,2)$  два раза, это обратная операция к взятию разности ряда (дифференцированию). В итоге получаем модельный ряд подходов (рис. 12). Собственно, после этой процедуры модель  $ARMA$  превращается в модель  $ARIMA$ .

Анализ остатков показывает адекватность модели. Гипотеза о нормальном распределении остатков не отвергается, значимых автокорреляций в ряду остатков не наблюдается (рис. 13–14). В заключение рассчитаем стандартные ошибки индивидуальных прогнозов по формуле (8), с помощью которых определим доверительный интервал (95%) точечных прогнозов, в том числе и на 2014 год.

Запомним интервальный прогноз на 2014 год ( $2,59 \pm 0,43$  млн экз.) и перейдем к следующей модели.

**Модель распределенного лага (DLM).** Применительно к задаче расчета пополнения методом распределенных лагов в общем виде модель можно записать как:

$$R_t = \beta_0 \times x_t + \beta_1 \times x_{t-1} + \beta_2 \times x_{t-2} + \dots + \beta_n \times x_{t-n} \quad (10)$$

где  $R_t$  — пополнение кижуча: количество потомков от производителей в год  $t$ ,

$\beta_{0...n}$  — коэффициенты,

$x_{t...t-n}$  — независимые объясняющие переменные.

В качестве объясняющей переменной  $x$  был выбран коэффициент воспроизводства западнокамчатской горбуши  $R/S$  (отношение потомков к производителям, показывает величину потомков на одного производителя). Причины выбора такой переменной следующие.

1. Как численно доминирующий вид, горбуша оказывает большое влияние на

жизнь других видов лососей и, в конечном виде, на их распределение и численность. Отношений «хищник–жертва» между различными видами тихоокеанских лососей не наблюдается, но так как горбуша служит основным источником биогенов в речной экосистеме, то влияние это не прямое, а косвенное, выражающееся через кормовую базу. Отнерестившиеся производители горбуши являются основным источником биогенов в речной экосистеме и через трофические связи оказывают

Таблица 3. Дисперсионный анализ

	df	SS	MS	F	Значимость F
Регрессия	2	0,755	0,377	10,925	0,004
Остаток	9	0,311	0,035		
Итого	11	1,065			

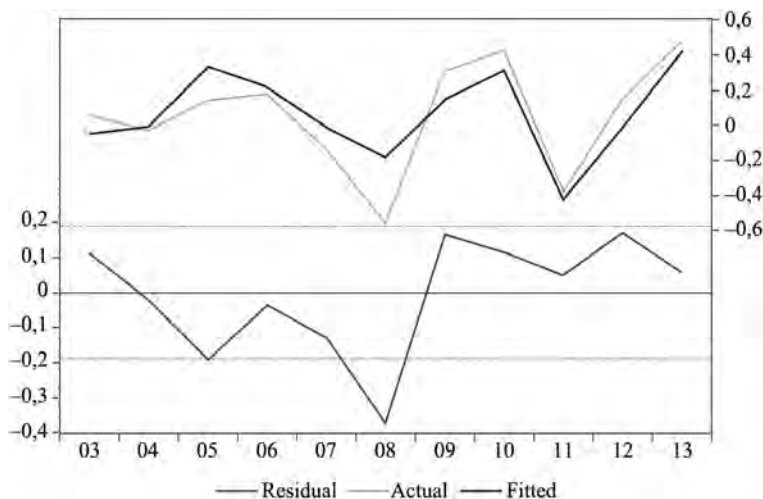


Рис. 11. Фактический и модельный временные ряды разности второго порядка подходов кижуча  $D(N,2)$  и ошибки модели  $ARMA$

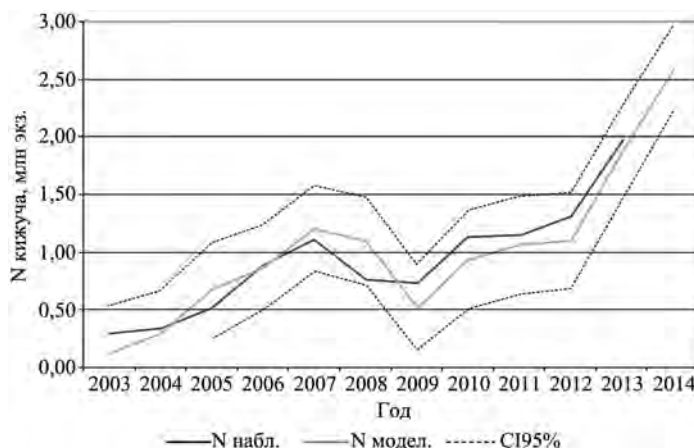


Рис. 12. Фактический и модельный временные ряды ( $ARIMA$ ) подходов кижуча  $N$ , с доверительным интервалом прогнозных значений ( $\alpha=0,05$ )

Таблица 4. Оценка параметров регрессии

Переменная	Коэффициенты	Стандартная ошибка	t-статистика	P-значение	Нижние 95%	Верхние 95%
AR(2)	-0,570	0,237	-2,399	0,040	-1,107	-0,032
MA(3)	0,878	0,074	11,798	$8,9 \times 10^{-7}$	0,710	1,047

влияние на выживаемость молоди тихоокеанских лососей на всех стадиях онтогенеза речного периода жизни (Бирман, 1985; Заварина, Шевляков, 2004; Шевляков, Заварина, 2004). Факты, свидетельствующие о влиянии горбуши на численность и прочие биологические показатели других видов лососей, отмечались неоднократно (Бирман, 1985; Бугаев, 1995, 2000; Bugayev, Dubynin, 2000; Шунтов, Темных, 2004).

2. Чередование численности пополнения кижуча в четные и нечетные годы, наблюдающееся в последнее время (рис. 5, 2002–2008 гг.). Такой феномен можно объяснить влиянием смены у горбуши депрессивной линии нечетных лет и продуктивной линии четных лет.

3. Показатель  $R/S$  у горбуши можно представить как интегральный показатель, обобщающий условия окружающей среды. Условно его высокие значения соответствуют режиму благоприятных экологических факторов, а низкие — соответственно, неблагоприятных.

4. Продолжительность цикла жизни горбуши два года, что минимум на два года меньше, чем у кижуча (основные возрастные группы которого на Западной Камчатке 2+ и 3+). Это может дать заблаговременные оценки пополнения кижуча, а следовательно, даст возможность прогнозировать.

5. Кросскорреляционный анализ между временными рядами пополнения кижуча  $R_{киж}$  и коэффициента воспроизводства горбуши  $R/S$  (рис. 15) показывает значимые корреляции на лагах с запаздыванием и говорит о нетривиальных отношениях между двумя этими рядами.

Следующим шагом анализа является преобразование исходных данных по коэффициенту воспроизводства горбуши. В некоторые годы (1992, 1999 и 2009) данный параметр (количество потомков на производителя) достигает неправдоподобно высоких значений (рис. 16).

Это можно объяснить недоучетом производителей на нерестилищах, особенно в малочисленные нечетные годы депрессивной линии горбуши. При не-

больших подходах производителей ошибка их учета нарастает экспоненциально. Подробно об этом говорится в работе (Шевляков, Маслов, 2011).

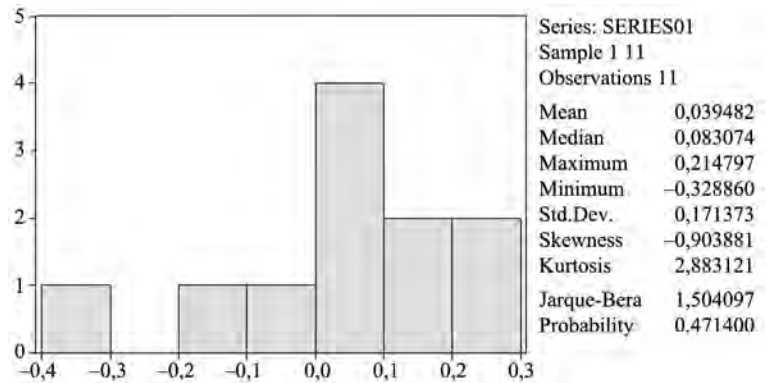


Рис. 13. Гистограмма остатков ( $d=0,16953$ , Колмогорова-Смирнова тест:  $p>0,2$ ; Лиллиефорс тест:  $p>0,2$ ; Гипотеза о нормальности распределения остатков не отвергается. Вероятность нормального распределения — 47% по Жарку-Бера)

	Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
1			0,150	0,150	0,3210	0,571
2			-0,113	-0,138	0,5232	0,770
3			0,211	0,262	1,3211	0,724
4			-0,388	-0,551	4,4005	0,355
5			-0,399	-0,125	8,1963	0,146
6			0,008	-0,133	8,1982	0,224
7			-0,048	0,170	8,2807	0,308
8			-0,013	-0,136	8,2894	0,406
9			0,071	-0,203	8,6489	0,470
10			0,022	-0,138	8,7156	0,559

Рис. 14. АКФ и ЧАКФ остатков модели

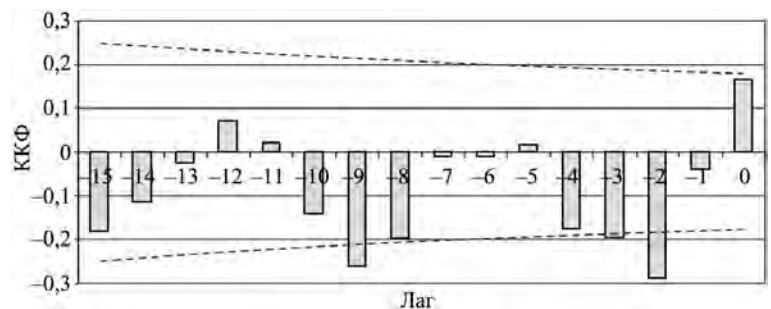


Рис. 15. Кросскорреляционный анализ между временными рядами пополнения кижуча  $R_{киж}$  и коэффициента воспроизводства горбуши  $R/S$

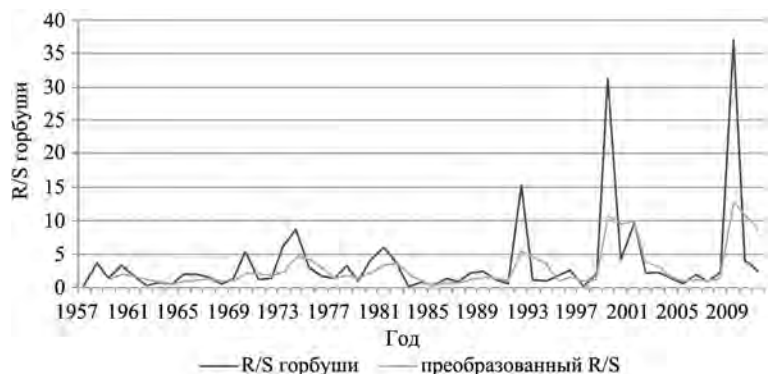


Рис. 16. Исходный и преобразованный ряды коэффициента воспроизводства горбуши  $R/S$

Было принято решение преобразовать ряд  $R/S$  горбуши с помощью экспоненциального сглаживания, по формуле:

$$\bar{r}_t = \frac{r_t + r_{t-1} / k + r_{t-2} / k^2}{3}, \quad (11)$$

где  $r_t$  — коэффициент воспроизводства горбуши  $R_t/S_t$ ,

$\bar{r}_t$  — преобразованный коэффициент воспроизводства горбуши,

$k$  — параметр, определяющий вес переменной  $r$  на лагах  $-1$  и  $-2$ . На нулевом лаге —  $k^0 = 1$ . Таким образом, со временем его значение изменяется в геометрической прогрессии. Этот параметр определяется одновременно с коэффициентами  $b_n$  при минимизации суммы квадратов отклонений приведенного ниже уравнения (12).

Итоговая модель зависимости пополнения кижуча от коэффициента воспроизводства горбуши:

$$R_{t\text{киж}} = b_0 \times \bar{r}_t + b_1 \times \bar{r}_{t-1} + b_2 \times \bar{r}_{t-2} + \dots + b_n \times \bar{r}_{t-n} \quad (12)$$

Преобразованный ряд коэффициентов воспроизводства горбуши при оптимальном для итоговой модели  $k$  показан на рис. 16. При минимизации квадратов отклонений (ошибок) модели от реальных данных на различных временных отрезках, выяснилось, что модель, начиная с 1993 года, достаточно хорошо описывает временной ряд численности поколений кижуча (рис. 17).

На высоком уровне значимости оказались коэффициенты на лагах  $-12, -10, -9, -7$  и  $0$ . Остальные коэффициенты незначимы, и соответствующие лаги были исключены из модели. Регрессионная статистика и значимость уравнения регрессии показаны в таблицах 5–7. Данные показатели очень высокие. Причина несоответствия модели фактическому ряду до 1993 г. заключается в следующем: в 1978 г. был запрещен японский дрейфтерный промысел в 200-мильной экономической зоне (Курмазов, 2001), вследствие чего, как минимум до этого года, данные по уловам

горбуши и кижуча (следовательно, и данные по подходам к местам нереста) являются неполными. Однако если из 1993 г. вычесть максимальный лаг объясняющей переменной ( $-12$ ), мы получим 1981 год. Вспомним, что преобразование ряда коэффициента воспроизводства горбуши шло методом сглаживания по трем последним годам. Иными словами, объясняющая переменная для 1981 года  $\bar{r}_{1981}$  несет в себе информацию и за предыдущие два года. Вычитая их, мы получим 1979 год, первый год с новыми условиями промысла и отличающейся статистикой.

На рис. 18 показаны фактические и смоделированные данные общей численности пополнения кижуча Западной Камчатки с 95%-м доверительным интервалом. Численность пополнения кижуча от родителей 2010 г. составит  $R_{2010} = 1,64$  млн экз., а от родителей 2011 г. —  $R_{2011} = 2,395$  млн экз.

Моделирование процентных долей в возрастных когортах потомков кижуча. Моделирование численности потомков — это только половина

Таблица 5. Регрессионная статистика

Множ. R	0,998
R-квадрат	0,992
Норм. R-квадрат	0,893
Стандартная ошибка	0,072
Наблюдения	15

Таблица 6. Дисперсионный анализ

	df	SS	MS	F	Значимость F
Регрессия	6	10,637	1,773	311,355	$6,5 \times 10^{-10}$
Остаток	9	0,051	0,006		
Итого	15	10,689			

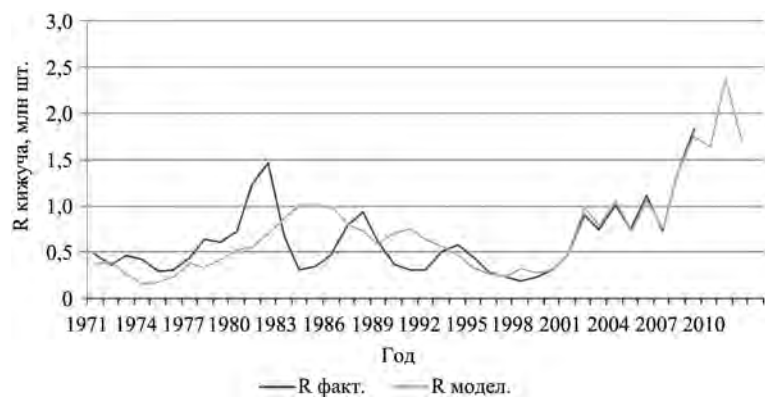


Рис. 17. Фактическая и модельная численность поколений кижуча Западной Камчатки (модель DLM)

Таблица 7. Оценка параметров регрессии

Переменная	Коэффициенты	Стандартная ошибка	t-статистика	P-значение	Нижние 95%	Верхние 95%
-12	0,109	0,019	5,783	0,000	0,066	0,152
-10	0,117	0,015	7,954	0,000	0,084	0,150
-9	0,058	0,008	7,002	0,000	0,039	0,077
-7	0,047	0,009	5,346	0,000	0,027	0,067
0	-0,028	0,006	-4,638	0,001	-0,042	-0,014
k	1,351	0,202	6,688	0,000	0,894	1,808

решения задачи прогнозирования численности будущего подхода рыбы. Необходимо также определить доли возрастных групп в будущем подходе, и сделать это можно различными путями. Одним из самых простых путей является взятие средних значений долей за последние несколько лет. Такой подход является оправданным, когда с течением времени доли возрастных групп мало изменчивы и остаются на стабильном уровне. В противном случае, когда за рассматриваемый период мы наблюдаем достаточно сильные колебания возрастных долей, а также можем предположить цикличность, лучше воспользоваться моделированием возрастного состава. У кижуча Западной Камчатки нерестовое стадо преимущественно (в среднем 98% от подхода за период с 1970 по 2009 года) состоит из особей возрастов 2+ и 3+ (соответственно третий и четвертый года жизни) (Зорбиди, 2010), поэтому в нашем случае достаточно смоделировать временной ряд какой-либо одной группы. Возьмем ряд изменения доли возраста 2+ (рис. 3). После визуального анализа можно предположить, что ряд представляет собой не случайные колебания. Действительно, мы видим разные режимы: двухлетние циклы в 1974–1978, 1990–1997 гг., которые чередуются с 5-, 6-летними циклами в 1979–1989 гг. и 3-, 4-летними циклами в 1997–2009 гг. Поэтому для моделирования возьмем ряд, начиная с 1997 г. Используем модель *ARIMA*.

Так же, как и в показанной выше модели подходов *ARIMA*, сначала приведем ряд к стационарному, путем взятия разности первого порядка и анализа АКФ и ЧАКФ (рис. 19).

Автокорреляции имеют значимые уровни на втором лаге, и их вид весьма напоминает аналогичные автокорреляции для ряда подходов кижуча, что не случайно, ведь доля возрастной группы 2+ (как и других возрастных групп) — это доля именно от его подхода. В данном случае нельзя определенно сказать, какой процесс преобладает: авторегрессии 2-го порядка или скользящего сред-

него 2-го порядка, поэтому так же, как и при моделировании подходов, для идентификации модели необходимо рассмотреть варианты моделей, показанные в таблице 1.

В итоге наиболее подходящей адекватной моделью, отвечающей условиям стационарности и обратимости, остается модель вида:

$$D(C^{2+}, 1) = AR(2) + MA(1).$$

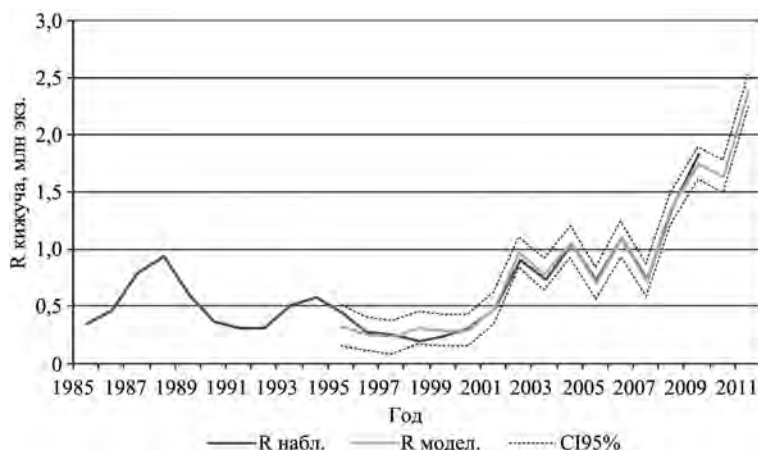
Статистическая значимость уравнения регрессии и коэффициентов, а также анализ ошибок, показаны в таблицах 8–10.

Таблица 8. Регрессионная статистика

Множественный R	0,9027
R-квадрат	0,8148
Стандартная ошибка	7,08
Наблюдения	10

Таблица 9. Дисперсионный анализ

	df	SS	MS	F	Значимость F
Регрессия	2	1339,84	669,92	13,356	0,003
Остаток	8	401,28	50,16		
Итого	10	1741,12			

Рис. 18. Фактическая и модельная численность поколений кижуча Западной Камчатки (модель *DLM*) с доверительным интервалом прогнозных значений ( $\alpha=0,05$ )

	Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
1			-0,181	-0,181	0,4992	0,180
2			-0,708	-0,765	8,9140	0,012
3			0,343	-0,052	11,115	0,011
4			0,412	-0,070	14,676	0,005
5			-0,280	0,226	16,551	0,005
6			-0,309	-0,154	19,233	0,004
7			0,228	-0,045	20,984	0,004
8			0,126	-0,397	21,655	0,006
9			-0,206	-0,167	24,026	0,004
10			0,005	-0,101	24,028	0,008
11			0,068	-0,080	24,811	0,010

Рис. 19. АКФ и ЧАКФ ряда первой разности доли возрастной группы 2+ ряд  $D(C^{2+}, 1)$ 

Таблица 10. Оценка параметров регрессии

Переменная	Коэффициенты	Стандартная ошибка	t-статистика	P-значение	Нижние 95%	Верхние 95%
AR(2)	-0,8537	0,2280	-3,745	0,0057	-1,3794	-0,3280
MA(1)	-0,9371	0,0577	-16,233	0,0000	-1,0703	-0,8040

Сравним фактические и модельные коррелограммы, а также фактический и модельный временные ряды долевого участия возрастной группы 2+ (рис. 20–21).

Анализ ошибок этой модели (рис. 22–23) показывает возможность опровержения гипотезы о нормальном распределении остатков на уровне  $\alpha=0,1$  (по Лиллиефорсу). Другие тесты эту гипотезу не опровергают. Автокорреляционный анализ тоже «на грани»: довольно высокое значение частной функции на 4-м лаге, хотя и не выходящее за пределы доверительного интервала.

Теперь, используя модель DLM пополнения, выполненную на первом этапе анализа, и модель ARIMA процентных долей в расщеплении поколений на возраста, выполненных на втором этапе, можно спрогнозировать общий подход кижуча Западной Камчатки в 2014 г. Для этого надо перемножить смоделированные значения пополнения и соответствующие им доли возрастов:

$$N_{2014} = C_{2011}^{2+} \times R_{2011} + C_{2010}^{3+} \times R_{2010}$$

Фактический подход, его модельные значения и 95%-й доверительный интервал индивидуальных прогнозов покажем на графике (рис. 24).

Анализ ошибок не опровергает гипотезу о нормальном распределении остатков. Вероятность гипотезы о нормальном распределении по тесту Жарка-Бера  $p=0,997$  (рис. 25–26).

Как и в предыдущей модели, запомним полученный результат: прогноз подхода кижуча в 2014 году — **2,18±0,27 млн экз.**

Следующие две модели, которые мы рассмотрим, принадлежат к классу моделей зависимости пополнения от нерестового запаса.

**Модель Рикера.** В качестве уравнения, описывающего зависимость пополнения от запаса, используется формула Рикера (2).

Как и в предыдущей модели DLM, численность подхода рассчитывается с помощью модельных значений пополне-

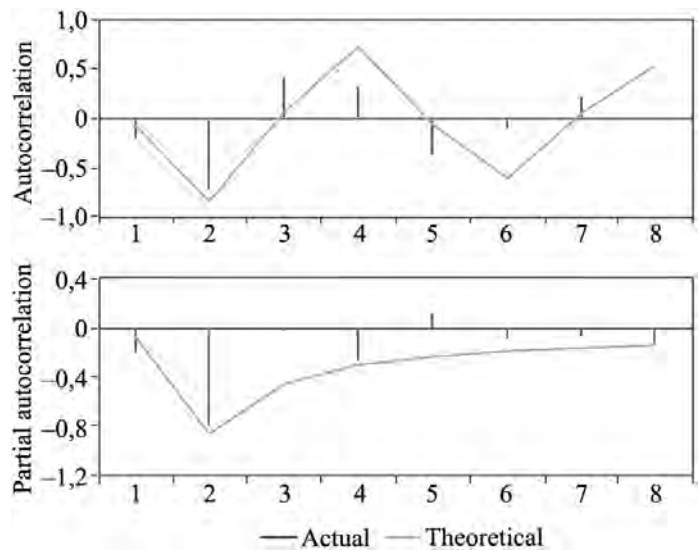


Рис. 20. Фактические и модельные коррелограммы ряда  $D(C^{2+}, 1)$

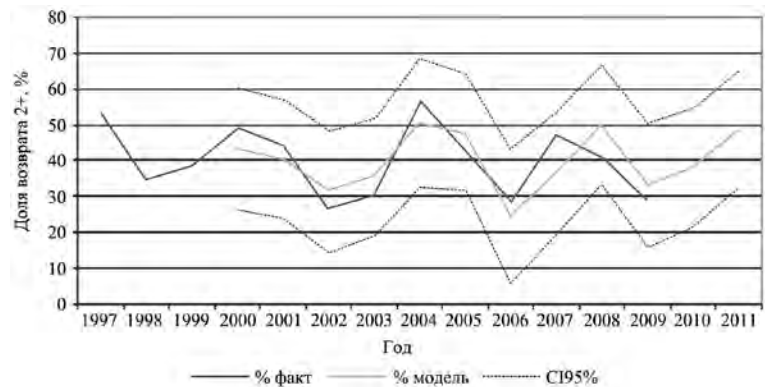


Рис. 21. Фактический и модельный ряды изменения доли возрастной группы 2+ у кижуча

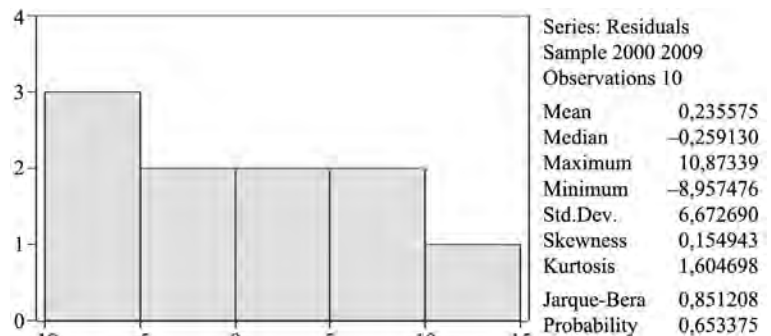


Рис. 22. Гистограмма остатков ( $d=0,24656$ , Колмогорова-Смирнова тест:  $p>0,2$ ; Лиллиефорс тест:  $p<0,1$ ; гипотеза о нормальности распределения отвергается на уровне  $\alpha=0,1$ ; по тесту Жарка-Бера гипотеза о нормальности распределения не отвергается)

	Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
1	-0,176	-0,176	0,4146			
2	-0,418	-0,464	3,0403			
3	0,230	-0,052	3,9438	0,047		
4	-0,241	-0,483	5,1039	0,078		
5	-0,121	-0,222	5,4571	0,141		
6	0,334	-0,146	8,8121	0,066		
7	0,131	0,193	9,4958	0,091		
8	-0,174	-0,041	11,304	0,079		
9	-0,064	0,008	11,800	0,107		

Рис. 23. АКФ и ЧАКФ остатков модели

ния  $R$ , возрастов 2+ и 3+ и их долей в подходе (используется модель *ARIMA* возрастного состава):

$$N_t = C^{2+}_{t-3} \times R_{t-3} + C^{3+}_{t-4} \times R_{t-4}. \quad (13)$$

Основной трудностью при моделировании стал оптимальный выбор интервала временных рядов нерестового запаса и пополнения кижуча. Разброс точек в облаке данных на диаграмме зависимости  $R$  от  $S$  (рис. 27) достаточно большой для выбора однозначной и статистически значимой модели. Даже для промежутка данных в последние десять лет (2000–2009 гг.) параметры модели будут явно не стационарными (рис. 28). Поэтому набор исходных данных для регрессионного анализа пришлось значительно сократить до последних пяти лет наблюдений (2004–2009 гг.). Обусловлено это скачкообразным изменением параметров модели, начиная с 2008 года. На рис. 28 хорошо видно, что линии регрессии № 1–3 изменяются незначительно, однако следующие две линии № 4 и № 5, характерные включением данных 2008 и 2009 годов соответственно, значительно отличаются от предыдущих. Максимум куполообразной кривой за два года вырастает почти в два раза, иными словами, почти вдвое повышается эффективность воспроизводства.

Результаты регрессионного анализа показаны в таблицах 11–13. Несмотря на короткий временной ряд исходных данных величиной в пять лет, модель и ее параметры  $a$  и  $b$  получились статистически значимыми с высокой степенью достоверности.

Модельные значения подхода на фоне фактических данных показаны на рис. 29. Напомним, что фактические значения 2004–2008 гг. не учитывались в модели, а модельные значения для этих лет показаны с помощью обратной экстраполяции. Для этого промежутка видим, что модельные значения всегда выше фактических, однако стоит заметить, что тенденции возрастания и убывания численности практически не изменяются.

Ограниченный пул исходных данных существенно расширяет границы доверительного интервала. Однако с другой

Таблица 11. Регрессионная статистика

Множ. R	0,899
R-квадрат	0,809
Стандартная ошибка	0,225
Наблюдения	5

Таблица 12. Дисперсионный анализ

	SS	df	MS	F	Значимость F
Регрессия	8,5109	2	4,2555	83,992	0,0023
Остаток	0,1519	3	0,0506		
Итого	8,6629	5			

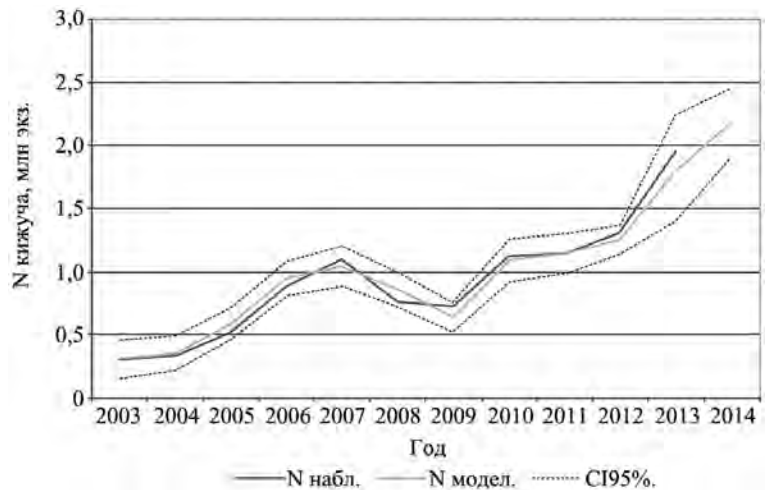


Рис. 24. Фактический и модельный ряды подходов кижуча (модель *DLM*), с доверительным интервалом прогнозных значений ( $\alpha=0,05$ )

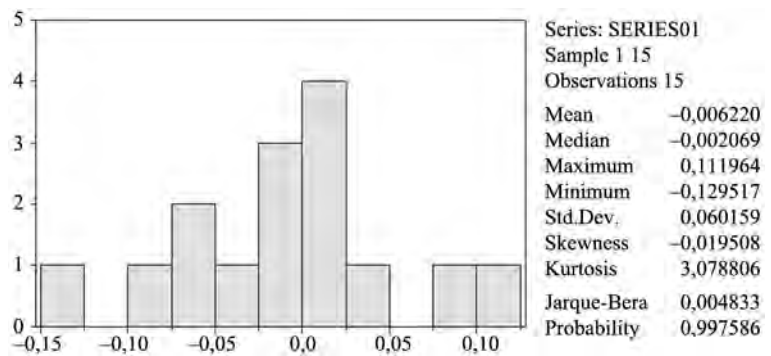


Рис. 25. Гистограмма остатков ( $d=0,15664$ , Колмогорова-Смирнова тест:  $p>0,2$ ; Лиллиефорс тест:  $p>0,2$ ; гипотеза о нормальности распределения остатков не отвергается. Вероятность нормального распределения — 99% по Жарку-Бера)

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0,203	0,203	0,7505	0,386
		2	-0,097	-0,144	0,9346	0,627
		3	-0,246	-0,207	2,2238	0,527
		4	0,111	0,216	2,5113	0,643
		5	0,132	0,020	2,9531	0,707
		6	-0,117	-0,217	3,3379	0,765
		7	-0,414	-0,304	8,7883	0,268
		8	-0,169	-0,013	9,8298	0,277
		9	-0,010	-0,125	9,8344	0,364
		10	-0,014	-0,210	9,8440	0,454
		11	-0,174	-0,132	11,777	0,381
		12	0,044	0,169	11,942	0,450

Рис. 26. АКФ и ЧАКФ остатков модели

стороны, анализ ошибок не отвергает гипотезу о нормальности распределения остатков, и модель можно признать адекватной (рис. 30–31).

С этими оговорками примем к дальнейшему рассмотрению прогноз подхода кижуча на 2014 год, составивший  $1,72 \pm 0,62$  млн экз.

**Модель Шепарда.** Следующей моделью из класса «запас–пополнение» рассмотрим модель Шепарда, вида:

$$R = aS / (b + S^2). \quad (14)$$

В отличие от модели (5), мы предлагаем принять параметр  $c$  постоянно равным 2 (константе), исходя из следующих соображений. Во-первых, мы понизим количество параметров в модели, а тем самым повысим ее статистическую значимость, ведь наблюдений у нас всего пять. Во-вторых, такую модель (как и модель Рикера) можно будет представить в линейном виде, что облегчает регрессионный анализ и вычисление необходимых статистик. И в-третьих, есть возражения теоретического плана: если параметр  $c$  не равен целым 0, 1 или 2 (например,  $c=3,65$ ), то невозможно аналитически объяснить, — что означает нерестовый запас  $S$  в степени 3,65. Однако рассуждения на эту тему выходят за рамки данной работы.

А численность будем рассчитывать, как и в предыдущей модели, по формуле (13). Результаты регрессионного анализа и анализа ошибок представлены в таблицах 14–16 и на рис. 32–33. Статистическая значимость уравнения регрессии выше, чем у модели Рикера, однако параметр  $b$  статистически незначим. С другой стороны, доверительный интервал индивидуальных прогнозных значений несколько уже, чем по модели Рикера. Как и в предыдущем случае, при анализе отклонений модели, гипотезы о нормальном распределении остатков не отвергаются.

Таблица 13. Оценка параметров регрессии

Переменная	Коэффициенты	Стандартная ошибка	t-статистика	P-значение	Нижние 95%	Верхние 95%
a	17,468	3,916	4,461	0,021	5,006	29,929
b	3,814	0,552	6,906	0,006	2,056	5,572

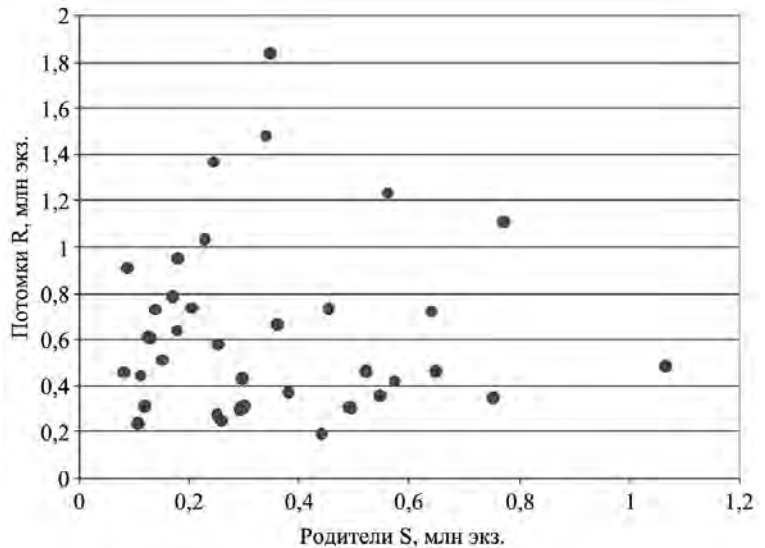


Рис. 27. Пополнение кижуча в зависимости от пропущенных на нерест рыб (1971–2009 гг.)

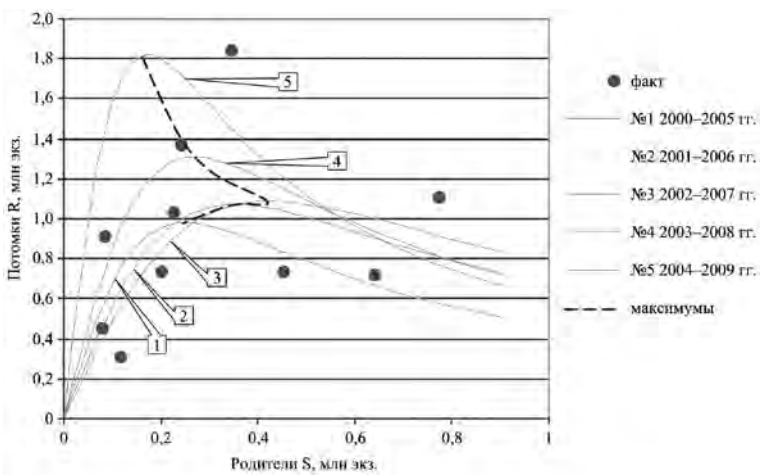


Рис. 28. Данные по родителям и потомкам кижуча Западной Камчатки за период 2000–2009 гг. и проведенные к ним кривые Рикера за пятилетние промежутки

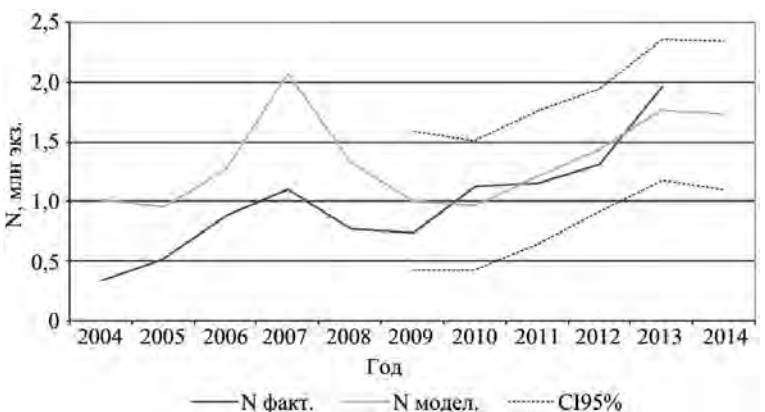


Рис. 29. Фактический и модельный ряд подходов кижуча (по модели Рикера), с доверительным интервалом прогнозных значений ( $\alpha=0,05$ )

Таблица 14. Регрессионная статистика

Множ. R	0,931
R-квадрат	0,866
Стандартная ошибка	0,188
Наблюдения	5

Таблица 15. Дисперсионный анализ

	df	SS	MS	F	Значимость F
Регрессия	8,556	2	4,278	120,582	0,0014
Остаток	0,106	3	0,036		
Итого	8,663	5			

Таблица 16. Оценка параметров регрессии

Переменная	Коэффициенты	Стандартная ошибка	t-статистика	P-значение	Нижние 95%	Верхние 95%
a	0,6080	0,1137	5,347	0,0128	0,2461	0,9698
b	0,0228	0,0199	1,146	0,3348	-0,0405	0,0862

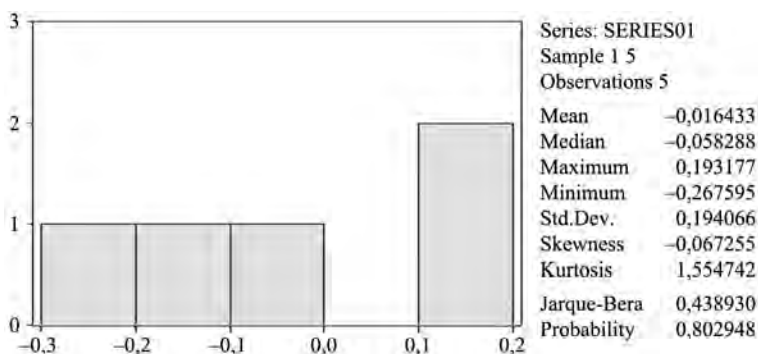


Рис. 30. Гистограмма остатков ( $d=0,2235$ , Колмогорова-Смирнова тест:  $p>0,2$ ; Лиллиефорс тест:  $p>0,2$ ; гипотеза о нормальности распределения остатков не отвергается — 80% по Жарку-Бера)

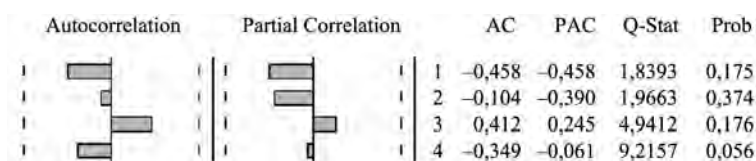


Рис. 31. АКФ и ЧАКФ остатков модели

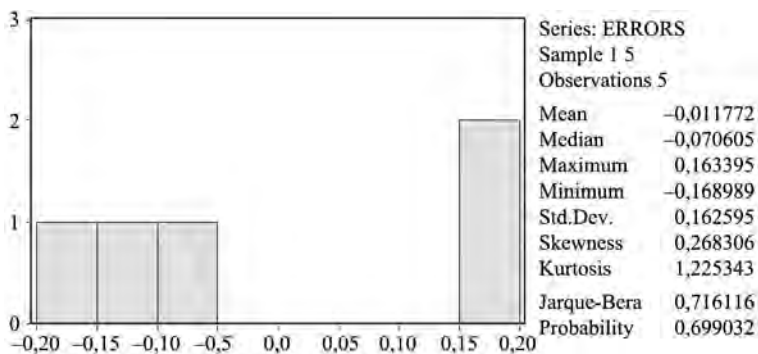


Рис. 32. Гистограмма остатков ( $d=0,2551$ , Колмогорова-Смирнова тест:  $p>0,2$ ; Лиллиефорс тест:  $p>0,2$ ; гипотеза о нормальности распределения остатков не отвергается — 69% по Жарку-Бера)

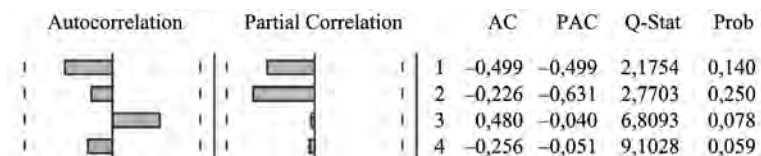


Рис. 33. АКФ и ЧАКФ остатков модели

Модельные значения подходов кижуча в сравнении с фактическими данными показаны на рис. 34. Прогноз на 2014 г. по этой модели составляет **2,16±0,52 млн экз.**

## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Получив прогнозы по четырем рассмотренным моделям, определим, какая модель является лучшей. Для этого сравним их статистики в сводной таблице 17.

Практически по всем статистикам, лучшей является модель *DLM*. С прогнозом по этой модели с точностью до 0,2 млн экз. совпадает прогноз, данный по модели Шепарда, однако она имеет более широкий интервальный прогноз и худшие статистики, что, как мы помним, обусловлено небольшим набором наблюдений, однако по статистикам она в основном лучше, чем ее «одноклассник», модель Рикера. Наихудший результат по стандартным ошибкам регрессии и прогноза показывает модель Рикера, кроме того, она единственная показывает снижение численности подхода в 2014 г., по сравнению с 2013 г.

Чтобы дать средневзвешенную оценку интервального прогноза на 2014 год, необходимо найти средневзвешенные оценки точечного прогноза и его стандартного отклонения. Для этого будем ранжировать модели по 4-балльной системе, где наивысший балл будет получать модель с наилучшей статистикой. Балльные оценки показаны в табл. 18.

В результате наивысший балл набирает модель *DLM* (28 баллов), затем следуют модели Шепарда (16 баллов) и *ARIMA* (15 баллов), наименьший результат у модели Рикера (11 баллов). Набранное количество баллов будет служить весовым коэффициентом вклада каждой модели. Рассчитаем средневзвешенный прогноз:



$$N_{срвзв.} = \frac{W_{ARIMA} \cdot N_{ARIMA} + W_{DLA} \cdot N_{DLA} + W_{Рикер} \cdot N_{Рикер} + W_{Шепард} \cdot N_{Шепард}}{W_{ARIMA} + W_{DLA} + W_{Рикер} + W_{Шепард}}, \quad (15)$$

где  $N$  — прогнозы, а  $W$  — весовые коэффициенты для каждой модели. Получаем:

$$N_{срвзв.} = \frac{15 \cdot 2,59 + 28 \cdot 2,18 + 11 \cdot 1,72 + 16 \cdot 2,16}{15 + 28 + 11 + 16} = 2,19 \text{ млн экз.}$$

Далее вычисляем средневзвешенное стандартное отклонение:

$$s_{срвзв.} = \sqrt{\frac{\sum_{i=1}^n (N_i - N_{срвзв.})^2}{n-1}}$$

где  $n$  — общее количество баллов,

$$s_{срвзв.} = \sqrt{\frac{4,8162}{70-1}} = 0,2642$$

Затем средневзвешенную стандартную ошибку (где  $n$  — количество моделей):

$$SE = \frac{s_{срвзв.}}{\sqrt{n}} = \frac{0,2642}{\sqrt{4}} = 0,1321$$

Находим доверительный интервал при  $\alpha=0,05$ :

$$CI_{0,95} = 0,1321 \cdot 1,96 = 0,259$$

Следовательно, средневзвешенный прогноз подхода кижуча на 2014 год составит:

$$N_{2014} = 2,19 \pm 0,259 \text{ млн экз.}$$

Другим путем обобщения полученных результатов является представление полученных интервальных прогнозов в виде плотностей вероятности. Имея точечные прогнозы и их доверительные интервалы  $CI_{0,95}$ , предполагаем, что точечный прогноз представляет собой среднее, приближенное к математическому ожиданию нормального распределения его ошибок, и представляет собой первый параметр нормального распределения. Вторым параметром является стандартное отклонение от среднего, и в нашем случае оно равно стандартной ошибке прогноза (табл. 17).

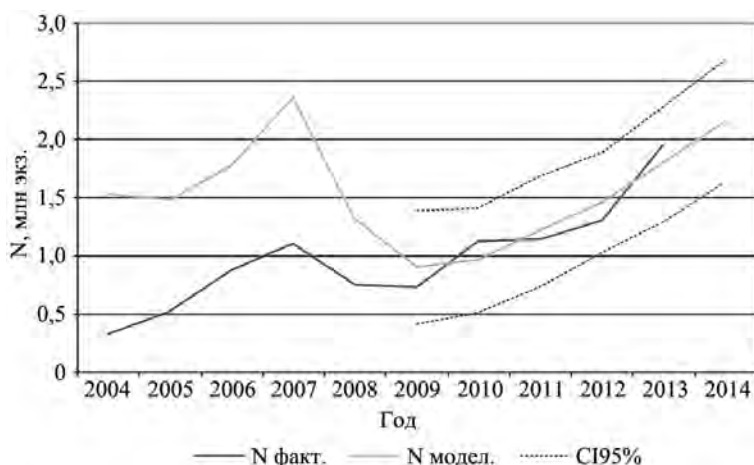


Рис. 34. Фактический и модельный ряд подходов кижуча (по модели Шепарда), с доверительным интервалом прогнозных значений ( $\alpha=0,05$ )

Таблица 17. Основные статистики рассмотренных моделей

Модель	Статистическая значимость уравнения регрессии		Станд. ошибка регрессии	$R^2$	Станд. ошибка прогноза	Макс. $p H_0$ коэф.	Информ. критерий Шварца $BIC$	Информ. критерий Акаике $AIC$
	F	$p$						
ARIMA	10,92524	0,0040	0,1858	0,70	0,1896	0,0310	-3,1306	-4,0137
DLM	311,355	$6,5 \times 10^{-10}$	0,0716	0,99	0,1407	0,0012	-4,5959	-30,619
Рикер	83,9920	0,0023	0,2251	0,81	0,3177	0,0210	-2,8496	0,7227
Шепард	120,5815	0,0014	0,1884	0,87	0,2658	0,3348	-3,2058	-1,0586

Таблица 18. Набранные моделями баллы

Модель	Стат. значимость уравнения регрессии	Станд. ошибка регрессии	$R^2$	Станд. ошибка прогноза	Стат. значимость коэф. регрессии	$BIC$	$AIC$	Всего
ARIMA	1	3	1	3	2	2	3	15
DLM	4	4	4	4	4	4	4	28
Рикер	2	1	2	1	3	1	1	11
Шепард	3	2	3	2	1	3	2	16

Используя эти два параметра, мы можем получить функции нормального распределения и представить их как плотности вероятности прогнозов по каждой модели (рис. 35).

Учитывая, что все полученные прогнозы являются только различными (отчасти совпадающими) оценками одного и того же реального будущего

го подхода, можно выделить область, где все четыре оценки будут учитываться; иными словами, все данные прогнозы сбудутся. В теории вероятность одновременного наступления событий определяется как произведение вероятностей каждого из них. Графически на рис. 36 произведение вероятностей прогнозов показано вместе с плотностью

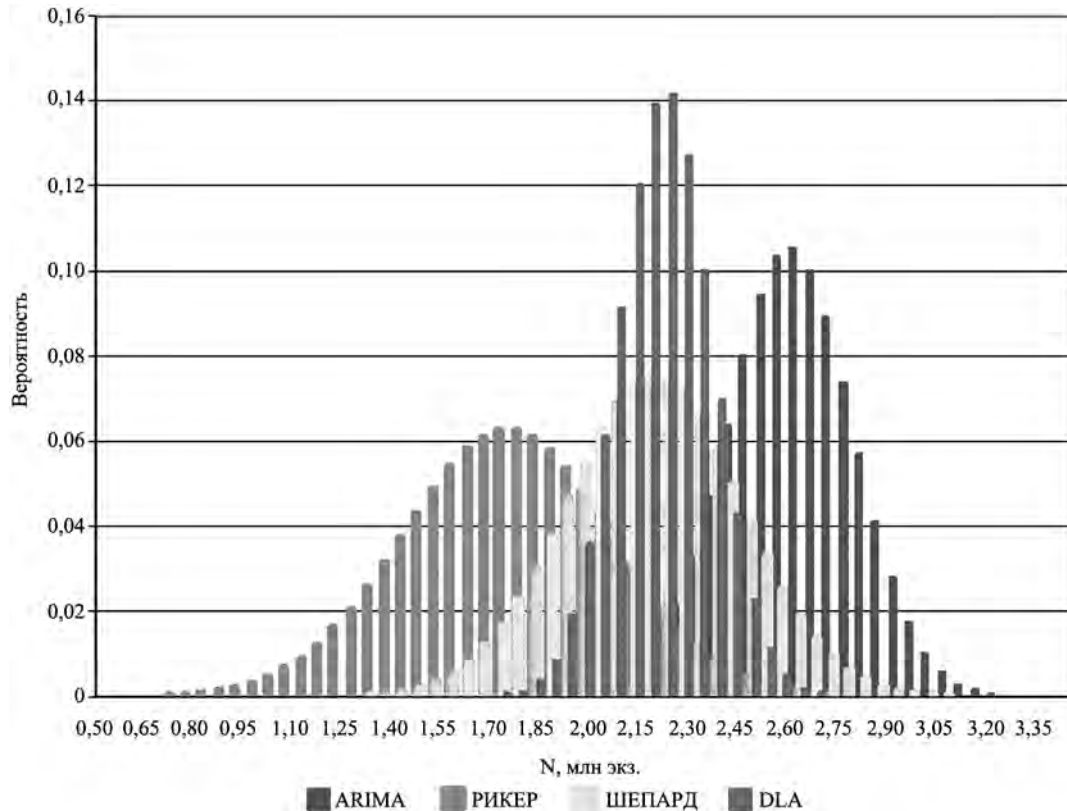


Рис. 35. Плотности вероятности прогнозов подхода кижуча Западной Камчатки на 2014 г.

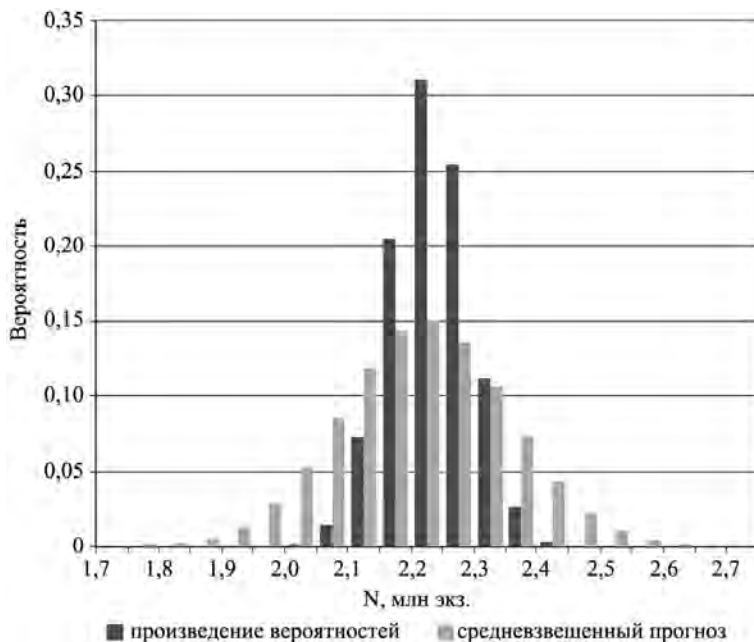


Рис. 36. Обобщенные прогнозы по четырем моделям: как произведение плотностей вероятности и как средневзвешенный прогноз

вероятности ранее вычисленного средневзвешенного прогноза. А прогноз, определенный как произведение вероятностей четырех моделей, составит (при  $\alpha=0,05$ ):

$$N_{2014} = 2,21 \pm 0,125 \text{ млн экз.}$$

В итоге можно резюмировать, что при применении двух способов обобщения результатов по рассмотренным моделям мы получили приблизительно один и тот же точечный прогноз (с различием в 0,02 млн экз.) подхода кижуча в 2014 году: 2,18–2,20 млн экз. При этом доверительные интервалы отличаются: средневзвешенный доверительный интервал практически в два раза шире, чем аналогичный у произведения плотностей вероятности. Объясняется это тем, что произведение плотностей вероят-

ности имеет более жесткие ограничительные рамки (все прогнозы верны и сбываются). Средне-взвешенный же плотностной прогноз, напротив, допускает возможность того, что один из двух прогнозов по крайним моделям неверен (модели *ARIMA* или Рикера) и является более подходящим при осторожной политике управления рыбными запасами.

## ЗАКЛЮЧЕНИЕ

Предложенный подход к решению задачи прогнозирования численности тихоокеанских лососей не является строго регламентированным. В нем должны соблюдаться основные моменты: применение нескольких прогностических моделей, обязательный анализ отклонений модельных данных от фактического материала, сравнение моделей между собой и выбор окончательного прогноза. Во всем остальном исследователю даются широкие возможности по применению различных видов прогностических моделей, методов разведывательного анализа, процедур сглаживания и преобразования временных рядов, выбора статистических критериев и методов аппроксимации, определения выбросов из данных или длины временного ряда. Во всех этих случаях не обойтись также без должного опыта и определенной интуиции. С другой стороны, мы против применения в прогнозе какой-либо одной модели без сравнения ее результатов с показателями альтернативных моделей. Также интересно будет узнать, насколько полученный нами прогноз подхода западнокамчатского кижуча оправдается.

## СПИСОК ЛИТЕРАТУРЫ

*Бабаян В.К.* 1988. Математические методы теории рыболовства (модели изолированных популяций) // Рыб. хоз-во. Обзоры по информационному обеспечению общесоюзных научно-технических программ. М.: ЦНИИТЭИРХ. 76 с.

*Бабаян В.К.* 2000. Предосторожный подход к оценке общего допустимого улова (ОДУ): Анализ и рекомендации по применению. М.: ВНИРО, 192 с.

*Баранов Ф.И.* 1918. К вопросу о биологических основаниях рыбного хозяйства // Изв. отдела рыбоводства и науч.-промысл. Исследований. № 1 (1). С. 84–128.

*Бивертон Р., Холт С.* 1969. Динамика численности промысловых рыб. Пер. с англ. М.: Пищ. пром-сть. 248 с.

*Бирман И.Б.* 1985. Морской период жизни и вопросы динамики стад тихоокеанских лососей. М.: Агропромиздат. 208 с.

*Бокс Дж., Дженкинс Г.* 1974. Анализ временных рядов, прогноз и управление. М.: Мир. Вып. 1–2. 604 с.

*Брюков В.Г.* 2011. Как предсказать курс доллара. Эффективные методы прогнозирования с использованием Excel и EViews. М.: ЦИПСИР. 272 с.

*Бугаев В.Ф.* 1995. Азиатская нерка (пресноводный период жизни, структура локальных стад, динамика численности). М.: Колос. 464 с.

*Бугаев В.Ф.* 2000. Мир реки. Петропавловск-Камчатский: Северная Пацифика. 32 с.

*Дубынин В.А.* 2012. Об оптимальности производителей нерки на нерестилищах бассейна р. Озерной в современный период / Матер. Всерос. науч. конф., посвящ. 80-летию юбилею ФГУП «КамчатНИРО». Петропавловск-Камчатский: КамчатНИРО. С. 302–308.

*Заварина Л.О., Шевляков Е.А.* 2004. Возможный механизм формирования цикличности урожайных поколений кеты на северо-восточном побережье Камчатки / Сохранение биоразнообразия Камчатки и прилегающих морей. Матер. V науч. конф. Петропавловск-Камчатский. С. 52–55.

*Заварина Л.О.* 2010. О динамике биологических показателей и тенденциях изменения численности кеты (*Oncorhynchus keta*) р. Большой (Юго-Западная Камчатка) // Исслед. водн. биол. ресурсов Камчатки и сев.-зап. части Тихого океана: Сб. науч. тр. КамчатНИИ рыб. хоз-ва и океанографии. Вып. 18. С. 38–57.

*Запорожец О.М., Шевляков Е.А., Запорожец Г.В., Антонов Н.П.* 2007. Возможности использования данных о нелегальном вылове тихоокеанских лососей для реальной оценки их запасов // Вопр. рыболовства. Т. 8. № 3 (31). С. 471–483.

*Зорбиди Ж.Х.* 2010. Кижуч азиатских стад. Петропавловск-Камчатский: КамчатНИРО. 306 с.

*Криксунов Е.А.* 1995. Теория пополнения и интерпретация динамики популяций рыб // Вопр. ихтиологии. Т. 35. № 3. С. 302–321.

*Курмазов А.А.* 2001. Международно-правовые условия освоения морских биологических ресурсов в Тихом океане // Мир. океан: использование биол. ресурсов. М.: ВИНТИ. С. 24–41.

*Хилборн Р., Волтерс К.* 2001. Количественные методы оценки рыбных запасов. Выбор, динамика, неопределенность. СПб.: Политехника. 228 с.

- Шевляков Е.А., Заварина Л.О.* 2004. Об особенностях динамики численности и методиках прогнозирования запасов кеты *Oncorhynchus keta* W. (Salmonidae) Западной Камчатки // Исслед. водн. биол. ресурсов Камчатки и сев.-зап. части Тихого океана: Сб. науч. тр. КамчатНИИ рыб. хоз-ва и океанографии. Вып. 7. С. 181–186.
- Шевляков Е.А., Маслов А.В.* 2011. Реки, определяющие воспроизводство тихоокеанских лососей на Камчатке, как реперы для оценки заполнения нерестового фонда // Изв. Тихоокеан. НИИ рыб. хоз-ва и океанографии. Т. 164. С. 114–139.
- Шевляков Е.А., Дубынин В.А., Зорбиди Ж.Х., Заварина Л.О., Попова Т.А., Артюхина Н.Б., Горин С.Л., Коваль О.О.* 2013. Современное состояние лососевого реки Большой (Западная Камчатка): воспроизводство, промысел, управление // Изв. Тихоокеан. НИИ рыб. хоз-ва и океанографии. Т. 174. С. 3–37.
- Шунтов В.П., Темных О.С.* 2004. Взгляд на лососевую путину 2004 через призму итогов изучения и промысла лососей в 2003 г. // Рыбн. хоз-во № 2. С. 26–27.
- Anderson E.D.* (Ed.). 2002. The Raymond J.H. Beverton lectures at Woods Hole, Massachusetts. Three Lectures on Fisheries Science Given May 2–3, 1994. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-F/SPO-54. 161 p.
- Bugayev V.F., Dubynin V.A.* 2000. Factors influencing abundance of sockeye salmon (*Oncorhynchus nerka*) from the Ozernaya River, Southwest Kamchatka // Recent Changes in Ocean Production of Pacific Salmon J.H. Helle, Y. Ishida, D. Noakes and V. Radchenko (ed.). North Pac. Anadromous Fish Com. Bull. 2. Vancouver, Canada. P. 181–189.
- Cushing D.H.* 1973. The dependence of recruitment on parent stock // J. Fish. Res. Board. Canada. V. 30. P. 1965–1975.
- Deriso R.B.* 1980. Harvesting strategies and parameter estimation for an age-structured model // Can. J. Fish. Aquat. Sci. V. 37. P. 268–282.
- Hoff J.C.* 1983. A practical guide to Box-Jenkins forecasting. Belmont, Calif.: Lifetime Learning Publications. 316 p.
- Jarque C.M., Bera A.K.* 1981. Efficient tests for normality, homoscedasticity and serial independence of regression residuals: Monte Carlo evidence // Economics Letters. V. 7 (4). P. 313–318.
- Lilliefors H.* 1967. On the Kolmogorov-Smirnov test for normality with mean and variance unknown // Journal of the American Statistical Association/ V. 62. № 318 (Jun., 1967). P. 399–402.
- Pankratz A.* 1983. Forecasting with univariate Box-Jenkins models: Concepts and cases. N.Y.: Wiley. 562 p.
- Paulik G.J.* 1973. Studies of the possible form of the stock-recruitment curve // Rapp. P. – V. Reun. Cons. Perm. Intern. Explor. Mer. V. 164. P. 302–315.
- Quinn T.J. II.* 2003. Ruminations on the development and future of population dynamics models in fisheries // Natural Resource Modeling 16 (4). P. 341–392.
- Ricker W.* 1954. Stock and recruitment // Journal of the fisheries research board of Canada. V. 11. № 5. P. 559–623.
- Sheperd J.G.* 1982. A versatile new stock-recruitment relationship for fisheries, and the construction of sustainable yield curves // J. Cons. Intern. Explor. Mer. V. 40 (1). P. 67–75.
- Vandaele W.* 1983. Applied time series and Box-Jenkins models. London: Academic Press. 417 p.