

## БАЙЕСОВСКОЕ ОЦЕНИВАНИЕ ВРЕМЕННЫХ РЯДОВ ЧИСЛЕННОСТИ И УЛОВОВ ПО ВОЗРАСТАМ С ИСПОЛЬЗОВАНИЕМ МЕТОДА ЧАСТИЦ И НЕПАРАМЕТРИЧЕСКИХ РАСПРЕДЕЛЕНИЙ

© 2023 г. И.И. Шевченко

*Тихоокеанский филиал Всероссийского научно-исследовательского  
института рыбного хозяйства и океанографии (ТИНРО),  
Владивосток, 690091*

*E.mail: igor.shevchenko@tinro-center.ru*

Поступила в редакцию 25.06.2022 г.

В статье рассматривается задача моделирования достаточно длинных временных рядов численности запаса и уловов по возрастам. В качестве основы модели выступают уравнения выживания и улова в логарифмической форме в предположениях о сепарабельности коэффициентов промысловой и естественной смертности в зависимости от возраста и года промысла, а также о представлении соответствующих составляющих в виде функций определённого вида с настраиваемыми по имеющимся данным параметрами. На этапе построения базовой детерминистической модели параметры подбираются таким образом, чтобы исходные и сгенерированные с её применением значения логарифмов численности запаса и уловов были в определенном смысле близки. На втором этапе логарифмы численности когорт и уловов считаются случайными величинами. Их условные распределения оцениваются по невязкам, которые вычисляются как разности между исходными и модельными данными. В соответствующей скрытой марковской модели, представляющей динамику некоторой когорты, логарифмы численности когорт являются состояниями, а логарифмы численности уловов – наблюдениями. С помощью этой модели строятся оценки плотностей апостериорных функций распределений предсказанных и сглаженных логарифмов численности когорт. Применяется метод частиц (многоэтапный метод Монте-Карло), а оценивание функций плотности случайных величин по выборкам проводится с применением базовой функции (ядра). С использованием разработанного пакета Fishmetica в среде Julia продемонстрирована возможность численной реализации методики на примере тестового набора данных. Обсуждаются некоторые модификации методики, включая сравнение различных базовых моделей на основе оценок правдоподобия их параметров.

*Ключевые слова:* динамика промысловых запасов, уравнения выживания когорт, уравнения уловов, скрытые марковские модели, последовательное байесовское оценивание, многоэтапный метод Монте-Карло, аппроксимация функций плотности с применением базовой функции (ядра).

### ВВЕДЕНИЕ

Рассмотрим базовый набор переменных и связей между ними, предназначенный для описания изменения во времени некоторого промыслового запаса; например: Quinn, Deriso, 1999. Пусть  $T$  – число последовательных лет,

для которых доступны данные,  $J$  – число учетных возрастов. Через  $n_t^j$  и  $c_t^j$  будем обозначать численность запаса и улова в год с индексом  $t \in 1:T$  и для возраста с индексом  $j \in 1:J$ . Введём базовые и производные коэффициенты:

- $m_t^j$  – естественной смертности,
- $f_t^j$  – промысловой смертности,
- $z_t^j = m_t^j + f_t^j$  – полной смертности,
- $s_t^j = \exp(-z_t^j)$  – выживания,
- $e_t^j = \frac{f_t^j}{z_t^j} (1 - \exp(-z_t^j))$  – пере-  
счёта уловов (изъятия).

В этих обозначениях уравнения выживания и уравнения уловов в логарифмической форме имеют вид:

$$\begin{aligned} \ln n_{t+1}^{j+1} &= \ln s_t^j + \ln n_t^j, t \in 1:T-1, \\ j \in 1:J-1, \\ \ln c_t^j &= \ln e_t^j + \ln n_t^j, t \in 1:T, j \in 1:J, \end{aligned} \quad (1)$$

и при заданных граничных условиях описывают механизм взаимосвязи численности запаса и уловов через искомые коэффициенты естественной и промысловой смертности. Это линейные уравнения относительно логарифмов численности запаса и уловов, но соответствующие базовые коэффициенты входят в них нелинейно.

Ряд наблюдений динамики промыслового запаса за  $T$  лет с числом возрастов  $J$  содержит данные о  $(T+J-1)$  когортах. При  $T \geq J-1$  среди них имеется ровно  $T-J+1$  когорт с полным набором возрастов; например: Megrey, 1989; Шевченко, 2017. Пусть  $\mathcal{C}$  – некоторая когорта, для которой определены данные для  $l$  индексов возрастов из диапазона от  $a_1$  до  $a_l$ . При её описании применяются переменные  $N_a$  (численность когорты) и  $C_a$  (численность улова), а также коэффициенты:

- $M_a$  – естественной смертности,
- $F_a$  – промысловой смертности,
- $Z_a = M_a + F_a$  – полной смертности,

- $S_a = \exp(-Z_a)$  – выживания,
- $E_a = \frac{F_a}{Z_a} (1 - \exp(-Z_a))$  – пере-

счета уловов.

Для  $\mathcal{C}$  уравнения (1) переписываются в виде:

$$\begin{aligned} \ln N_{a+1} &= \ln S_a + \ln N_a, a \in a_1:a_l-1, \\ \ln C_a &= \ln E_a + \ln N_a, a \in a_1:a_l. \end{aligned} \quad (2)$$

Граничные условия для численности могут задаваться при некотором  $a \in a_1:a_l$ .

Структурированная по возрастам динамика рыбных запасов оценивается по данным, полученным в ходе экспедиционных исследований и на промысле, с привлечением дополнительной информации и математических моделей. Модели в пространстве состояний разбиваются на классы в зависимости от предположений об ошибках обработки состояний и/или наблюдений. Их параметры считаются случайными величинами. При необходимости состояния также включаются в вектор неизвестных параметров. Оценка распределений параметров осуществляется на основе максимизации функций правдоподобия, например: Megrey, 1989; Schnute, 1994; de Valpine, Hastings, 2002; de Valpine, 2002.

В Тихоокеанском филиале ВНИРО (ТИНРО) по некоторым запасам наряду с данными по численности уловов были накоплены временные ряды точечных оценок численности запасов по возрастам, полученных на основе лабораторных методик, в которых не использовались математические модели. Это, например, ряды оценок численности запасов и уловов берингоморского минтая за период 1962–2015 гг. с диапазоном возрастов 1–17 и тихоокеанской трески за период – 2001–2014 гг. с диапазоном возрастов 1–13. Для них значительная часть первичных данных

не может быть восстановлена. В связи с изменением требований к обоснованию прогнозов возникла задача разработки простых моделей, воспроизводящих такие временные ряды. Построение моделей, базирующихся на уравнениях типа (1) или (2), требует выдвижения предположений о статистической вариативности. В рассматриваемой информационной ситуации сформулировать какие-либо разумные предположения невозможно. Поэтому было решено разработать специализированную методику. Она включает в себя создание базовой детерминистической модели, относительно которой характеризуется статистическая вариативность. Распределения разностей исходных данных и значений, сгенерированных по базовой модели, определяет скрытую марковскую модель. На её основе можно проводить сглаживание и формировать предсказания для рядов логарифмов численности когорт при заданных логарифмах численности уловов.

В первом разделе кратко описываются оба этапа методики. На первом этапе в качестве основы выступают уравнения (1). Для сокращения числа параметров предполагается сепарабельность коэффициентов промысловой и естественной смертности в зависимости от возраста и года. Также предполагается, что соответствующие сомножители представлены в виде функций определённого вида. Для базовой модели наборы значений параметров соответствуют найденным численными методами локальным экстремумам, поскольку в общем случае нелинейные задачи глобальной оптимизации с ограничениями неразрешимы, например: Нестеров, 2007. Исходные ряды значений логарифмов численности когорт и уловов удовлетворяют построенным уравнениям только приближенно. На втором этапе они рассма-

триваются как реализации случайных величин. Условные распределения, описывающие степень доверия, с которой эти величины удовлетворяют базовым модельным уравнениям, оцениваются по невязкам, которые получаются при подстановке исходных данных в полученные уравнения. В соответствующей марковской модели, представляющей динамику некоторой когорты, логарифмы численности когорт трактуются как скрытые состояния, а логарифмы численности уловов выступают в качестве наблюдений (измерений). Эта модель используется для оценки плотностей апостериорных распределений логарифмов численности когорт по возрастам. Вспомогательной является задача фильтрации – реконструкции распределений состояний при получении нового наблюдения. Предсказание (экстраполяция) – это реконструкция без получения новых наблюдений, а сглаживание представляет собой реконструкцию при наличии всех доступных наблюдений. При расчётах оценок используется метод частиц (многоэтапный метод Монте-Карло), например: Kitagawa, 1987; Gordon et al., 1993; Sarkka, 2013. В настоящей работе аппроксимация функций плотности случайных величин по выборкам проводилась с применением базовой функции (ядра) (Gramacki, 2018) в отличие от предыдущих работ (Шевченко, 2019, 2020), где рассматривалась такая же постановка задачи и использовались предположения о стандартных типах распределений, параметры которых подбирались с применением метода наименьших квадратов.

Во втором разделе приведены примеры расчётов байесовских оценок для рядов, сформированных на основе набора данных для тестирования моделей динамики запасов из ВНИРО. С применением различных пакетов в среде Julia

(Bezanson et al., 2017) построены оценки функций плотностей апостериорных распределений и их некоторых статистических характеристик.

В заключительном разделе предложены модификации методики, включая оценивание функции правдоподобия и сравнение различных базовых моделей, которые могут, например, соответствовать разным локальным экстремумам или показателям степени близости исходных и модельных данных.

#### МАТЕРИАЛ И МЕТОДИКА

В качестве исходных данных рассматриваются временные ряды оценок численности когорт и уловов по возрастам за последовательность лет. Для построения модели динамики численности когорт в такой информационной ситуации предлагается методика, состоящая из двух этапов. Сначала рассчитываются параметры детерминистической модели, которая в определенном смысле наилучшим образом воспроизводит имеющиеся ряды. Затем формируются скрытые марковские модели отдельных когорт. Логарифмы численности когорт являются состояниями, а логарифмы численности уловов – наблюдениями. Считается, что аддитивные ошибки в уравнениях состояний и наблюдений распределены как невязки при подстановке исходных данных в оптимизированные по параметрам варианты уравнений (1). Для оценки функций плотности логарифмов численности когорт при фильтрации, предсказании и сглаживании применяются известные численные методы.

#### *Параметрическая идентификация механизмов взаимосвязи*

Матрицы  $m$  и  $f$  коэффициентов естественной и промысловой смертности являются неизвестными и подлежат оценке по имеющимся данным. Число

их элементов совпадает с числом исходных данных, представленных матрицами численности запаса  $\hat{n}$  и уловов  $\hat{c}$ . Чтобы обойтись меньшим количеством неизвестных параметров, используются явные функциональные зависимости. Так коэффициенты естественной  $m_t^j$  и промысловой  $f_t^j$  смертности представлены в виде произведения двух сомножителей. Для коэффициентов естественной смертности  $m_t^j$  один из них ( $u_t^j$ ) оценивает неблагоприятность для когорты определенного возраста условий выживания в некоторый заданный год, а второй ( $m^j$ ) зависят только от возраста. Реальное сокращение числа неизвестных параметров, относящихся к естественной смертности, достигается за счёт того, что учёт влияния факторов неблагоприятности через  $u_t^j$  ограничивается только младшими возрастными группами. Зависимость  $m^j$  от возраста описывается U-образной функцией, которая убывает на начальном участке, затем является постоянной для среднего диапазона возрастов и возрастает в конечной части. Для коэффициентов промысловой смертности  $f_t^j$  один сомножитель ( $\sigma^j$ ), который имеет вид сигмиды, связан с возрастом, а второй ( $f_t$ ) определяется интенсивностью промысла в конкретном году; детальную информацию в (Шевченко, 2017).

При подборе параметров, можно использовать любой эффективно вычисляемый показатель степени близости исходных данных и значений, сгенерированных по модели. Если  $\hat{u}_t^j, \hat{m}^j, \hat{\sigma}^j, \hat{f}_t$  – это значения параметров, минимизирующих выбранный показатель близости исходных и модельных значений, то соответствующие элементы исходных матриц  $\hat{n}$  и  $\hat{c}$ , которые являются оценками элементов матриц  $\hat{n}$  и  $\hat{c}$ , удовлетворяют при заданных граничных условиях (точным) соотношениям (см. (1)):

$$\begin{aligned} \ln \hat{n}_{t+1}^{j+1} &= \ln \hat{s}_t^j + \ln \hat{n}_t^j, t \in 1:T-1, \\ j &\in 1:J-1, \\ \ln \hat{c}_t^j &= \ln \hat{e}_t^j + \ln \hat{n}_t^j, t \in 1:T, j \in 1:J, \end{aligned} \quad (3)$$

где

$$\begin{aligned} \hat{s}_t^j &= \exp(-\hat{z}_t^j) \\ \hat{e}_t^j &= \frac{f_t^j}{z_t^j} \left(1 - \exp(-\hat{z}_t^j)\right), \\ \hat{z}_t^j &= \hat{u}_t^j \hat{m}^j + \hat{f}_t^j \hat{\sigma}^j, t \in 1:T, j \in 1:J. \end{aligned}$$

Подставим в (3) исходные данные:

$$\begin{aligned} \ln \check{n}_{t+1}^{j+1} &= \ln \hat{s}_t^j + \ln \check{n}_t^j + \omega_t^j, t \in 1:T-1, \\ j &\in 1:J-1, \\ \ln \check{c}_t^j &= \ln \hat{e}_t^j + \ln \check{n}_t^j + \nu_t^j, t \in 1:T, j \in 1:J, \end{aligned} \quad (4)$$

и предположим, что аддитивные невязки  $\omega_t^j$  и  $\nu_t^j$  являются выборками независимых случайных величин  $\omega$  и  $\nu$ .

Для каждой из  $(T+J-1)$  когорт, включающей  $l$  возрастов, соответствующие уравнения (2) относительно случайных величин  $x$  и  $y$ , представляющих логарифмы численности когорт и уловов, с учетом аддитивных случайных ошибок  $\omega$  и  $\nu$ , принимают вид:

$$\begin{aligned} x_{a+1} &= \hat{f}_a(x_a) + \omega, a \in a_1:a_l-1, \\ y_a &= \hat{g}_a(x_a) + \nu, a \in a_1:a_l, \end{aligned} \quad (5)$$

где

$$\begin{aligned} \hat{f}_a(x_a) &= \ln \hat{S}_a + x_a, a \in a_1:a_l-1, \\ \hat{g}_a(x_a) &= \ln \hat{E}_a + x_a, a \in a_1:a_l, \end{aligned}$$

$\hat{S}_a, \hat{E}_a$  – оптимальные значения коэффициентов  $S_a, E_a$ .

*Оценка функции плотности вероятности по выборке*

Функция плотности случайной величины может быть оценена по выборкам значений, а затем использована для генерации выборок с соответствующи-

ми распределениями. Альтернативой подбору параметров для общеизвестных распределений по выборке является использование непараметрических методов (Hardle et al., 2004). В рамках одного из практических подходов к построению приближения для функции плотности  $p(z)$  от непрерывной переменной  $z$  по выборке её значений  $z^{(1)}, \dots, z^{(n)}$  с весами  $w^{(1)}, \dots, w^{(n)}$ ,  $\sum_{i=1}^n w^{(i)} = 1$ , является оценивание плотности с использованием базовой функции (ядра)  $K$  (kernel density estimation, KDE). С каждым значением  $z^{(i)}$  из выборки связывается «небольшой» выброс плотности на интервале сглаживания длины  $h > 0$ , представляемый симметричной и гладкой функцией ядра  $K$ . Результирующая плотность получается суммированием всех таких выбросов

$$p(z) \approx \frac{1}{h} \sum_{i=1}^n w^{(i)} K\left(\frac{z^{(i)}-z}{h}\right).$$

Выбор значения  $h$  является критическим, а конкретный вид ядра  $K$  является менее значимым, например: Gramacki, 2018.

#### *Байесовское оценивание*

При изложении материалов, связанных с байесовским оцениванием, используются стандартные обозначения. Так функция плотности распределения случайной величины  $x$  записывается как  $p(x)$ . Совокупность, содержащая  $x_0, \dots, x_k$ , представляется как  $x_{0:k}$  и т.д. Предполагается, что все модельные случайные переменные абсолютно непрерывны, т.е. их распределения могут быть выражены через функции плотности.

Байесовская сеть определяется набором вершин, ассоциированных со случайными переменными, и набором дуг между ними. При этом все переменные в байесовской сети упорядочиваются в соответствии с имеющимися причинно-следственными зависимостями: сначала

причинные, затем скрытые и после них результирующие. В байесовской сети отсутствуют входящие из одной вершины в другую дуги, если соответствующие случайные величины независимы или условно независимы при заданной третьей величине. Любая переменная условно независима от всех последующих переменных, если заданы все её родительские вершины. Случайная величина в любой вершине байесовской сети зависит только от заданных значений переменных в родительских вершинах, например: Russell, Norvig, 2010.

В модели динамики логарифмы численности когорты – это скрытые (неизмеряемые) состояния, а логарифмы численности уловов – это наблюдения (измерения). Связь логарифмов численности когорты и уловов (5) представляется скрытой марковской моделью с функциями плотности распределений  $p(x_{a_1})$  и  $p(x_{a+1}|x_a)$ ,  $p(y_a|x_a)$  для всех значений  $a$ . Условные вероятности, интерпретируются как степени уверенности в том, что одна величина принимает какое-либо конкретное значение при заданном значении другой в соответствии с определенными соотношениями. Модель является примером байесовской сети и схематически изображена на рисунке 1.

Байесовское оценивание для скрытой марковской модели может трактоваться как задача статистического об-

ращения, где неизвестные значения временного ряда восстанавливаются по зашумленным наблюдениям. В этом случае цель анализа – оценить скрытые состояния по наблюдениям, т.е. построить апостериорные распределения состояний при полученных наблюдениях, например: Sarkka, 2013.

*Фильтрация*

При фильтрации оцениваются плотности распределения  $p(x_a|y_{a_1:a})$ ,  $a \in a_1:a_l$ . Чтобы выразить  $p(x_{a+1}|y_{a_1:a+1})$  через  $p(x_a|y_{a_1:a})$ , можно, например, воспользоваться выражением из (Gordon et al., 1993):

$$p(x_{a+1}|y_{a_1:a+1}) = \frac{p(x_{a+1}|y_{a_1:a})p(y_{a+1}|x_{a+1})}{p(y_{a+1}|y_{a_1:a})}, \quad (6)$$

где плотности  $y_{a+1}|x_{a+1}$  определяются моделями наблюдения, а плотности одношагового предсказания  $x_{a+1}|y_{a_1:a}$  и знаменатель (6) рассчитываются как маргинальные плотности:

$$p(x_{a+1}|y_{a_1:a}) = \int p(x_{a+1}|x_a) \times p(x_a|y_{a_1:a}) dx_a,$$

$$p(y_{a+1}|y_{a_1:a}) = \int p(y_{a+1}|x_{a+1}) \times p(x_{a+1}|y_{a_1:a}) dx_{a+1}.$$

Поскольку аналитические вычисления плотностей апостериорных распределений возможны только для специальных случаев, применяются ре-

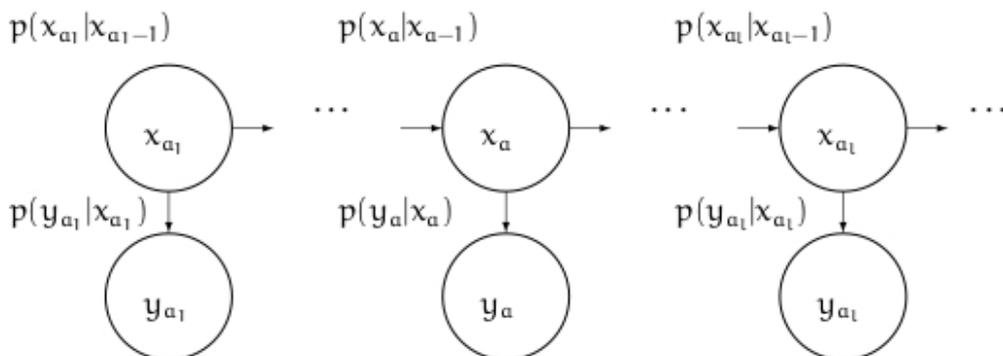


Рис. 1. Динамическая байесовская сеть для когорты,  $p(x_{a_1}|x_{a_1-1}) = p(x_{a_1})$ .

курсивные процедуры, которые по выборкам априорных позволяют получать выборки апостериорных величин. Пусть  $\{x_a^{(i)}\}_{i \in 1:N}$  – выборка из  $N > 0$  значений случайной величины с плотностью распределения, аппроксимирующей  $p(x_a|y_{a_1:a})$  для некоторого индекса  $a$ . Рассмотрим задачу генерирования выборки  $\{x_{a+1}^{(i)}\}_{i \in 1:N}$  из  $N$  значений случайной величины с плотностью распределения, аппроксимирующей  $p(x_{a+1}|y_{a_1:a+1})$ . Её можно решить, например, применяя метода бутстрепа, который состоит из одношагового предсказания и последующего уточнения, например: Kitagawa, 1987; Gordon et al., 1993; Doucet et al., 2000; Sarkka, 2013.

Процесс начинается при  $a = a_1$  с генерации выборки  $\{x_{a_1}^{(i)}\}_{i \in 1:N}$  по распределению  $p(x_{a_1})$  с весами  $w_{a_1}^{(i)} = 1/N$ . Для формирования предсказания на  $(a + 1)$ -ом шаге,  $a \in a_1:a_l$ , для каждого  $i \in N$  сначала генерируются значения  $\omega_a^{(i)} \sim p(\omega)$ , а затем по имеющимся  $x_a^{(i)}$  и  $\omega_a^{(i)}$  вычисляются элементы  $x_{a+1}^{(i)}$  через уравнение состояния (см. (5)):

$$x_{a+1}^{(i)} = \hat{f}_a(x_a^{(i)}) + \omega_a^{(i)}.$$

Приближение для плотности одношагового предсказания имеет вид:

$$p(x_a|y_{a_1:a-1}) \approx \sum_{i=1}^N \tilde{w}_a^{(i)} \delta_{x_a^{(i)}}(x_a), \quad (7)$$

Где  $\tilde{w}_a^{(i)} = p(\omega_a^{(i)})$  – веса для одношагового предсказания.

На этапе уточнения по модели наблюдения при известном  $y_a$  вес каждого предсказанного значения  $x_{a+1}^{(i)}$  меняется на (Gordon et al., 1993):

$$w_a^{(i)} = \frac{p(v_a^{(i)})}{\sum_{j=1}^N p(v_a^{(j)})}, \quad (8)$$

где  $v_a^{(i)} = y_a - \hat{g}_a(x_a^{(i)})$ ,  $y_a$  – соответствующее наблюдение (логарифм численности улова в возрасте  $a$ ),  $a \in a_1:a_l$ ; см. (5). По  $\{x_a^{(i)}\}_{i=1}^N$  с уточненными весами (8) можно генерировать выборки случайной величины, плотность распределения которой при  $N \rightarrow +\infty$  приближает  $p(x_a|y_{a_1:a})$ :

$$p(x_a|y_{a_1:a}) \approx \sum_{i=1}^N w_a^{(i)} \delta_{x_a^{(i)}}(x_a). \quad (9)$$

Описанный метод имеет вычислительную сложность  $O(lN)$ , например: Gordon et al., 1993; Doucet et al., 2000.

### Предсказание

Предсказание представляет собой фильтрацию без добавления новых наблюдений и заключается в оценке  $p(x_{a_l+k}|y_{a_1:a_l})$  при  $a_l < J$  для  $k \in 1:J - a_l$ . Процедура для предсказания может базироваться на рекурсивном соотношении:

$$p(x_{a_l+k+1}|y_{a_1:a_l}) = \int p(x_{a_l+k+1}|x_{a_l+k}) \times p(x_{a_l+k}|y_{a_1:a_l}) dx_{a_l+k}, k \geq 0.$$

Выборка на  $(k + 1)$ -шаге,  $k \geq 0$ , генерируется по выборке  $x_{a_l+k}^{(i)}$  на предыдущем шаге по соотношению (5):

$$x_{a_l+k+1}^{(i)} = \hat{f}_{a_l+k}(x_{a_l+k}^{(i)}) + \omega_{a_l+k}^{(i)}, \quad (10)$$

где  $\omega_{a_l+k}^{(i)} \sim p(\omega)$ . Пересчёт весов осуществляется по формуле:

$$w_{a_l+k+1}^{(i)} = \frac{w_{a_l+k}^{(i)} p(\omega_{a_l+k}^{(i)})}{\sum_{j=1}^N w_{a_l+k}^{(j)} p(\omega_{a_l+k}^{(j)})}, i \in 1:N. \quad (11)$$

Генерация по (10) и (11) начинается с  $k = 0$ , когда выборка значений  $\{x_{a_l}^{(i)}\}_{i \in 1:N}$  и весов  $\{w_{a_l}^{(i)}\}_{i \in 1:N}$  для приближения  $p(x_{a_l}|y_{a_1:a_l})$  определяется в результате применения процедуры

филътрации, описанной в предыдущем разделе.

Функция плотности для  $k$ -ого предсказания может быть записана как, например: Doucet et al., 2000:

$$p(x_{a_l+k} | y_{a_1:a_l}) = \int p(x_{a_l} | y_{a_1:a_l}) \prod_{n=1}^{n=k} p(x_{a_l+n} | x_{a_l+n-1}) dx_{a_l:a_l+k-1}, k \geq 1.$$

Заменяв  $p(x_{a_l} | y_{a_1:a_l})$  на приближение из предыдущего раздела, получаем для  $k \geq 2$ :

$$p(x_{a_l+k} | y_{a_1:a_l}) = \sum_{i=1}^N w_{a_l}^{(i)} \int p(x_{a_l+1} | x_{a_l}^{(i)}) \prod_{n=2}^{n=k} p(x_{a_l+n} | x_{a_l+n-1}) dx_{a_l+1:a_l+k-1}.$$

При оценке интеграла можно использовать уравнения перехода (5).

Итак, для  $k = 0$ :

$$p(x_{a_l} | y_{a_1:a_l}) \approx \sum_{i=1}^N w_{a_l}^{(i)} \delta_{x_{a_l}^{(i)}}(x_{a_l}),$$

и для  $k \geq 1$  (см. (11)):

$$p(x_{a_l+k} | y_{a_1:a_l}) \approx \sum_{i=1}^N w_{a_l+k}^{(i)} \delta_{x_{a_l+k}^{(i)}}(x_{a_l+k}), k \geq 1,$$

где  $x_{a_l+k}^{(i)} \sim p(x_{a_l+k} | x_{a_l+k-1}^{(i)})$ . Описанный метод имеет вычислительную сложность  $O(kN)$ , если  $\{w_{a_l}^{(i)}\}_{i \in 1:N}$  и  $\{x_{a_l}^{(i)}\}_{i \in 1:N}$  вычислены предварительно, например: Kitagawa, 1987; Doucet et al., 2000.

### Сглаживание

Под сглаживанием понимается оценка  $p(x_a | y_{a_1:a_l})$  для  $a \in a_1:a_l$ . Алгоритм сглаживания в заданном интервале (fixed-interval smoothing) (Doucet et al., 2000), позволяющий избегать повторных выборок, базируется на соотношении из (Kitagawa, 1987):

$$p(x_a | y_{a_1:a_l}) = \int p(x_a | x_{a+1}, y_{a_1:a_l}) dx_{a+1}. \tag{12}$$

Поскольку  $x_a$  не зависит от  $y_{a+1:a_l}$  при заданном  $x_{a+1}$ , с применением формулы Байеса, получаем:

$$p(x_a | x_{a+1}, y_{a_1:a_l}) = p(x_a | x_{a+1}, y_{a_1:a}, y_{a+1:a_l}) = \frac{p(x_{a+1} | x_a) p(x_a | y_{a_1:a})}{p(x_{a+1} | y_{a_1:a})}.$$

Перепишем теперь плотность совместного распределения  $x_a$  и  $x_{a+1}$  при заданном  $y_{a_1:a_l}$ :

$$p(x_a, x_{a+1} | y_{a_1:a_l}) = \frac{p(x_{a+1} | x_a) p(x_a | y_{a_1:a})}{p(x_{a+1} | y_{a_1:a})} p(x_{a+1} | y_{a_1:a_l}).$$

Таким образом, (12) может быть представлено в виде:

$$p(x_a | y_{a_1:a_l}) = p(x_a | y_{a_1:a}) \int \frac{p(x_{a+1} | x_a) p(x_{a+1} | y_{a_1:a_l})}{p(x_{a+1} | y_{a_1:a})} dx_{a+1}. \tag{13}$$

где знаменатель подинтегрального выражения рассчитывается по (7).

Пусть сгенерирована выборка  $x_a^{(i)}$  и по (8) вычислены веса  $w_a^{(i)}$ , определяющие приближенные плотности отфильтрованных состояний (9). Определим новые веса  $w_{a|a_l}^{(i)}$  для приближений плотностей сглаженных состояний:

$$p(x_a | y_{a_1:a_l}) \approx \sum_{i=1}^N w_{a|a_l}^{(i)} \delta_{x_a^{(i)}}(x_a), a \in a_1:a_l, \tag{14}$$

следующим образом. Для  $a = a_l$  полагаем

$$w_{a_l|a_l}^{(i)} = w_{a_l}^{(i)},$$

а затем в обратном порядке (от  $a = a_l - 1$  до  $a = a_1$ ) вычисляем

$$w_{a|a_l}^{(i)} = \sum_{j=1}^N w_{a+1|a_l}^{(j)} \frac{w_a^{(i)} p(x_{a+1}^{(j)} | x_a^{(i)})}{\sum_{k=1}^N w_a^{(k)} p(x_{a+1}^{(j)} | x_a^{(k)})}, i \in 1:N.$$

В соответствии с уравнениями перехода, например,  $p(x_{a+1}^{(j)} | x_a^{(i)})$  рассчитывается как  $p(\omega_a^{(ij)})$  для  $\omega_a^{(ij)} = x_{a+1}^{(j)} - \hat{f}(x_a^{(i)})$ .

Для того, чтобы обосновать описанный алгоритм, отметим, что в силу (9):

$$p(x_{a+1} | y_{a_1:a}) \approx \sum_{i=1}^N w_a^{(i)} p(x_{a+1} | x_a^{(i)}). \quad (15)$$

Кроме того,

$$\begin{aligned} \int \frac{p(x_{a+1} | x_a) p(x_{a+1} | y_{a_1:a_l})}{p(x_{a+1} | y_{a_1:a})} dx_{a+1} &\approx \\ \approx \int \frac{p(x_{a+1} | x_a)}{p(x_{a+1} | y_{a_1:a})} \sum_{i=1}^N w_{a+1|a_l}^{(i)} \delta_{x_{a+1}^{(i)}}(x_{a+1}) & \\ = \sum_{i=1}^N w_{a+1|a_l}^{(i)} \frac{p(x_{a+1}^{(i)} | x_a)}{p(x_{a+1}^{(i)} | y_{a_1:a})}. & \quad (16) \end{aligned}$$

Подставив теперь (9), (15) и (16) в (13), получаем приближенное равенство (14). Для описанного алгоритма сглаживания необходимо хранить выборки и веса для всех возрастов, а также значения функции плотности  $p(x_{a+1}^{(j)} | x_a^{(i)})$ . Его вычислительная сложность оценивается как  $O(lN^2)$ , например: Doucet et al., 2000.

## РЕЗУЛЬТАТЫ

При проведении тестовых расчётов по описанной методике матрицы численности запаса  $\tilde{n}$  и уловов  $\tilde{c}$  были сформированы на основе данных за 1974–2003 гг. ( $T = 30$ ) для 16 учетных возрастов ( $J = 16$ ) из первого из трёх наборов данных ВНИРО, которые предназначены для тестирования моделей динамики запасов. Отличные от единицы коэффициенты неблагоприятности, допускались для трёх младших возрастов. Детерминистическая модель включает 129 параметров, значения которых определялись в результате минимизации показателя близости, равного сумме модулей невязок между исходны-

ми и модельными данными (Шевченко, 2017). При оценках функций плотности распределения число частиц  $N$  полагалось равным 1000. При  $a_1 = 1$  функции плотности задавались с помощью функций Дирака, сосредоточенных в соответствующих заданных точках. Поскольку функции  $\hat{f}_{a_l+k}(\cdot)$ ,  $k > 0$ , неизвестны, при проведении расчётов плотностей предсказанных распределений полагалось, что  $\hat{f}_{a_l+k}(\cdot) = \hat{f}_{a_l}(\cdot)$ ,  $k > 0$ .

Для выполнения расчётов по описанным алгоритмам фильтрации, предсказания и сглаживания пакет Fishmetica был пополнен соответствующими функциями. Численная оптимизация основывалась на процедурах из библиотеки NLOpt. Гладкая аппроксимация плотностей по заданным выборкам проводилась через вызовы функции kde из пакета KernelDensity с опциями по умолчанию, а именно с плотностью нормального распределения в качестве ядра и интервалом сглаживания, вычисляемым по эмпирическому правилу Сильвермана, например: Gramacki, 2018. Оценка статистических характеристик и построение графиков выполнялись с применением функций из пакетов StatsBase, Distributions, Gadfly и др.

В анализируемом наборе содержатся данные о 45 когортах. Первая и последняя когорты содержат только один возраст, вторая и предпоследняя – два и т. д. Все 16 возрастов входят в 15 когорт. В данных наблюдаются аномалии, которые заключаются в возрастании численности когорт при переходе в последний учётный возраст – 16 лет. Когорты разбиваются на несколько групп в зависимости от диапазонов значений логарифмов численности когорт. Наименее продуктивная группа состоит только из одной когорты (с номером 19); см. рисунок 2.

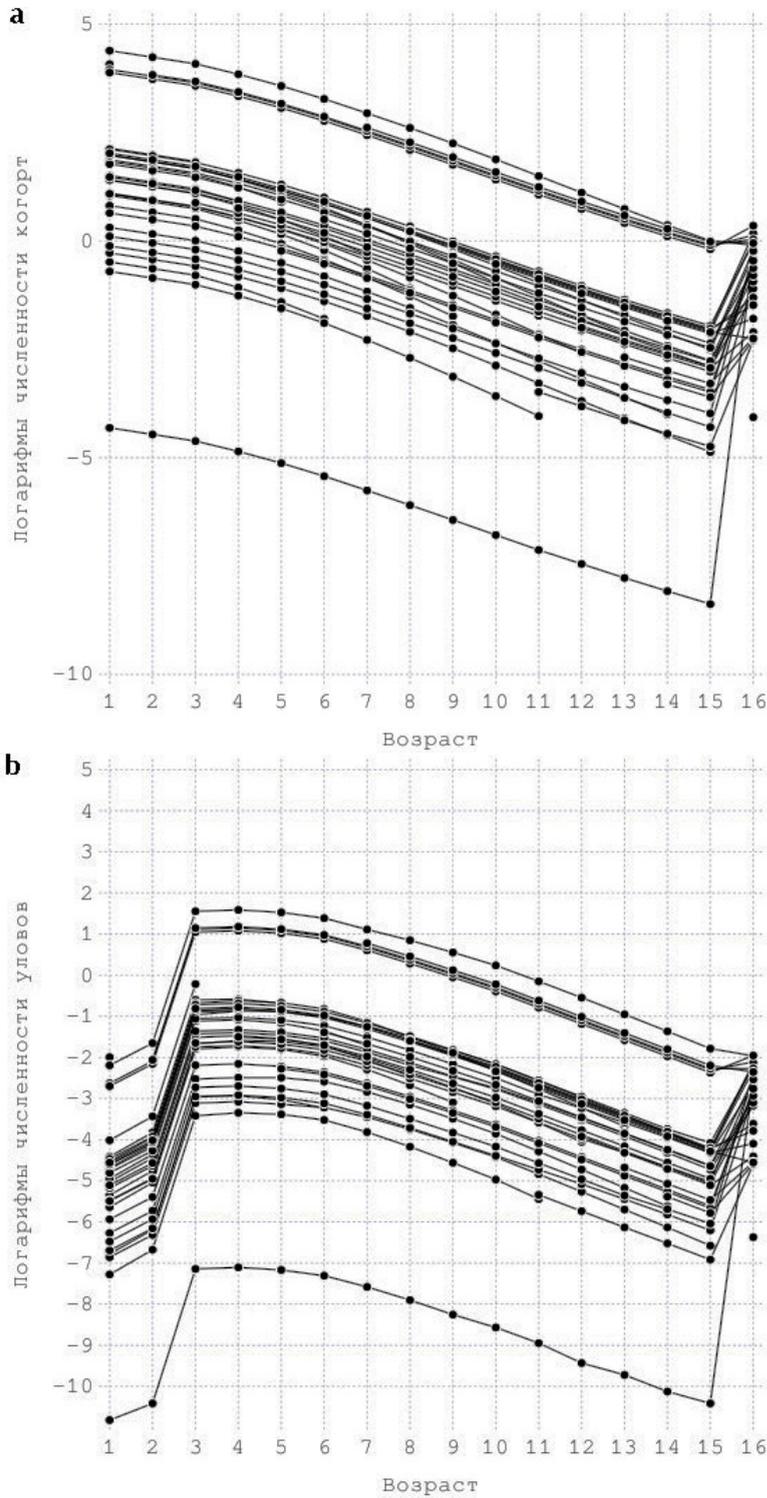
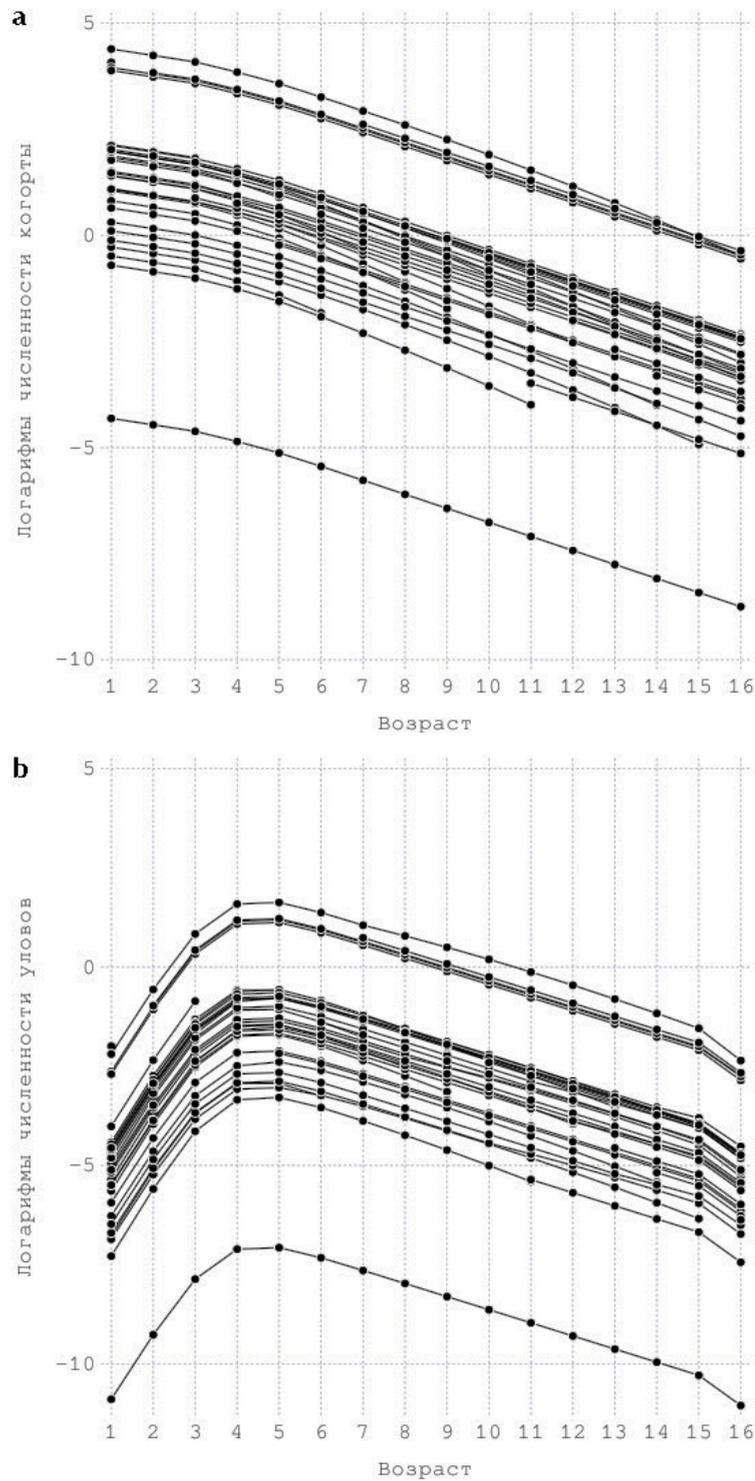


Рис. 2. Исходные данные по численности когорт (а) и уловов (б) по возрастам.

На рисунке 3 представлены логарифмы численности когорт и уловов по возрастам, рассчитанные по модели первого этапа. Приближения функций плотностей  $w$  и  $v$  (см. (5)), по-

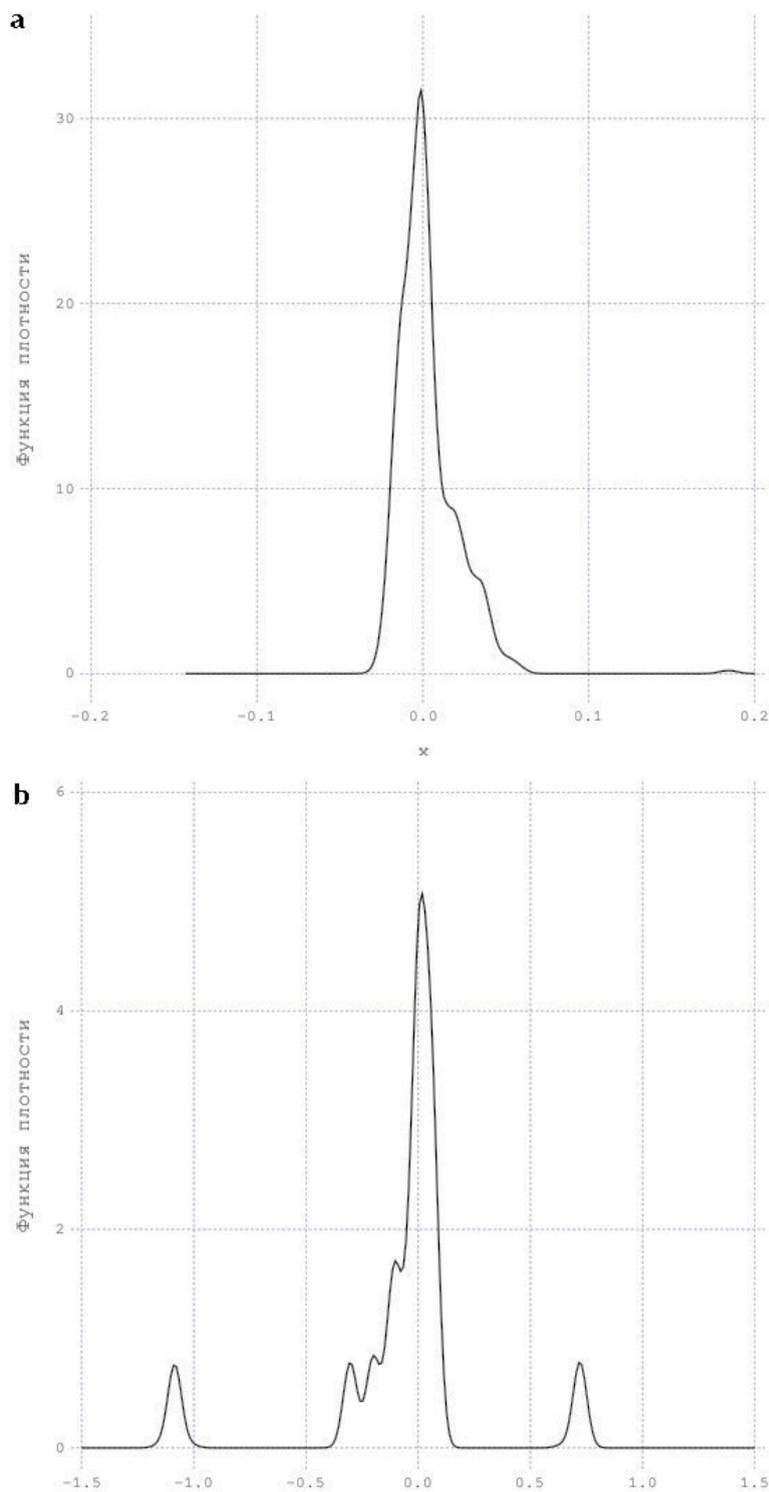
строенные по невязкам между исходными данными и значениями, полученными по детерминистической модели первого этапа, приведены на рисунке 4.



**Рис. 3.** Логарифмы численности когорты (а) и уловов (б) по возрастам, оцененные по модели первого этапа.

Результаты оценок апостериорных плотностей распределений и их статистических характеристик демонстрируются на примере когорты 20, содержащей все 16 возрастов, и когорты 36,

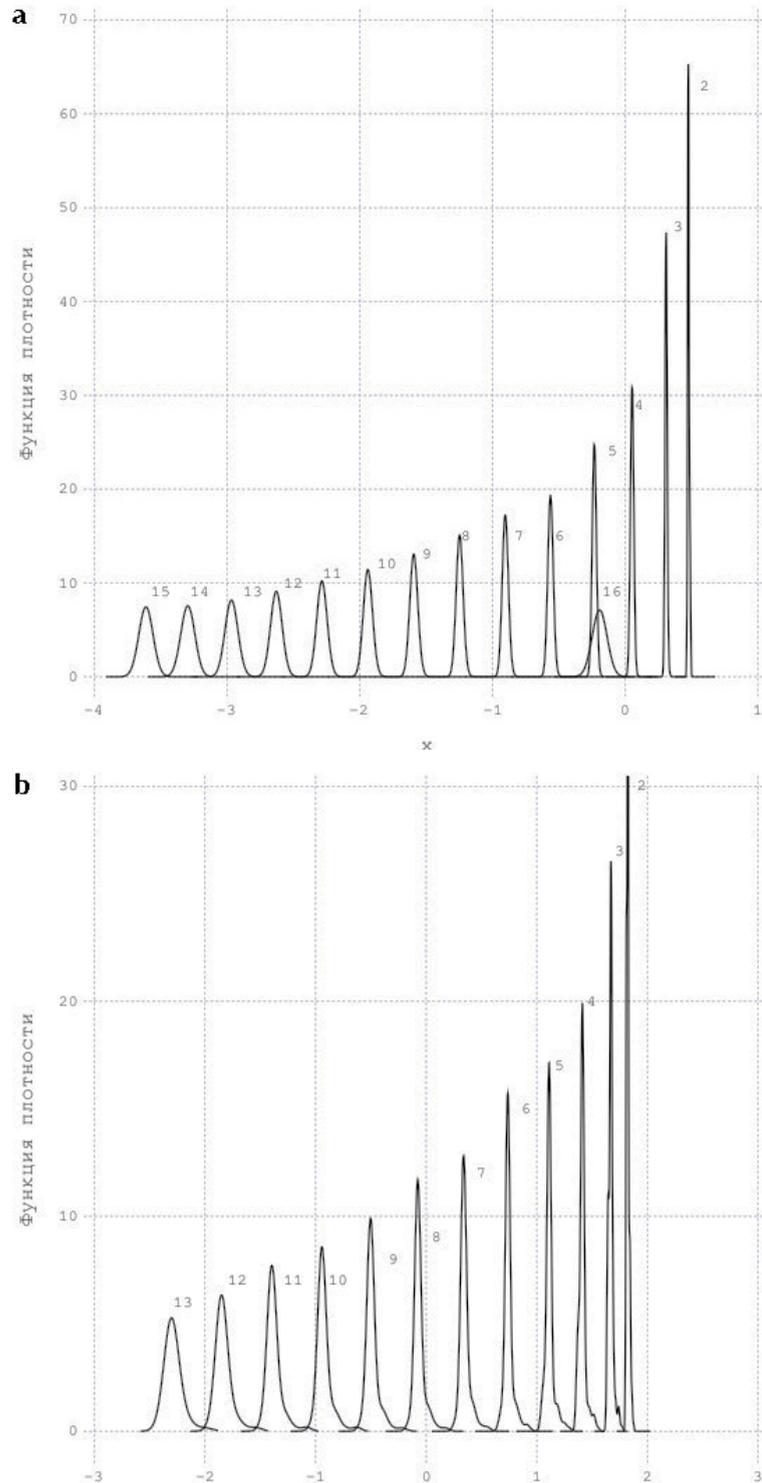
содержащей первые 10 возрастов. На рисунке 5 (а) показаны приближения для сглаженных (возраст 2–16) плот-



**Рис. 4.** Приближения функций плотности ошибок обработки состояний  $\omega$  (а) и наблюдений  $\nu$  (б).

ностей логарифмов численности когорты 20, а на рисунке 5 (б) – для сглаженных (возраст 2–10) и предсказанных (возраст 11–13) плотностей логарифмов численности когорты 36. В таблицах 1 и 2 собраны значения медиан и медианных абсолютных отклонений от медиан, характеризующих средние и

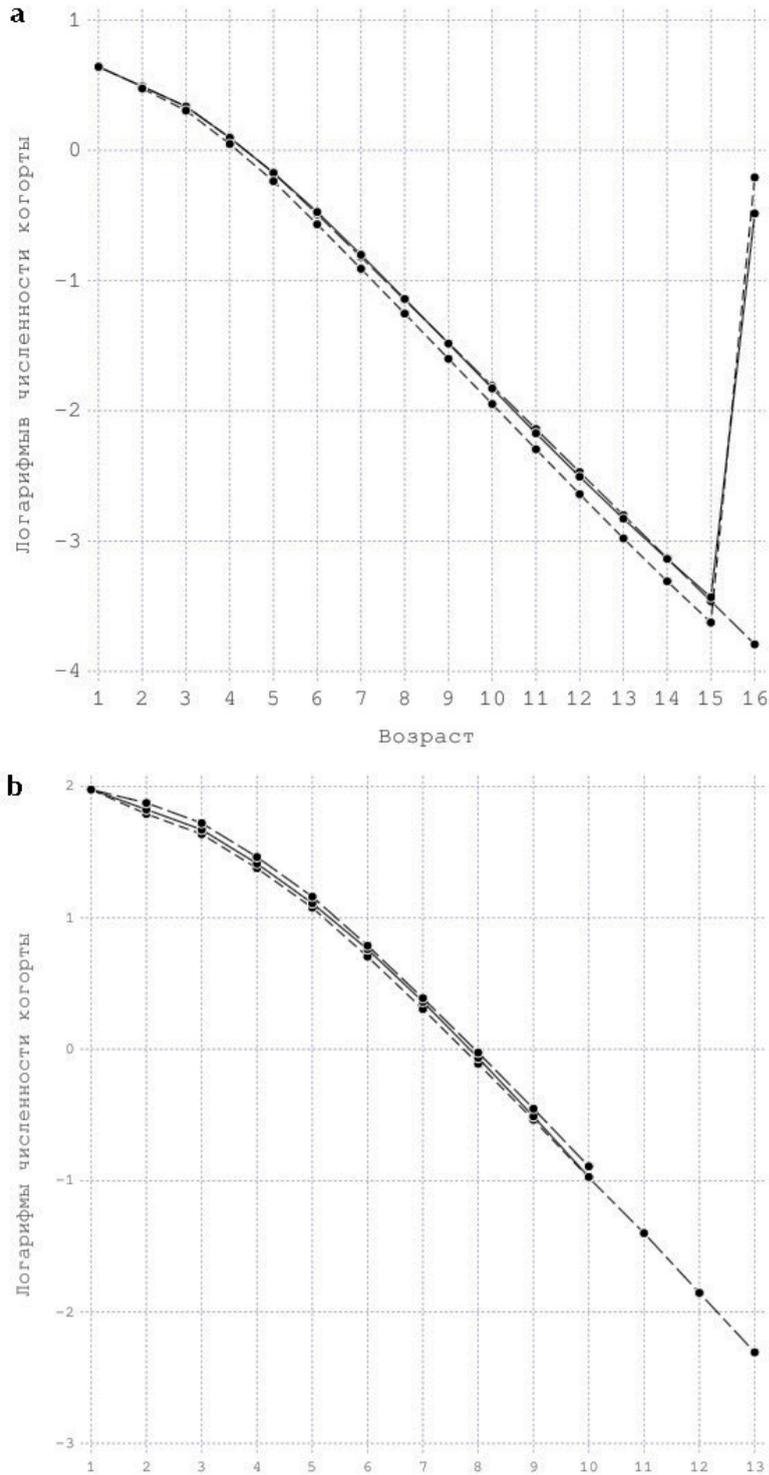
рифмов численности когорты 36. В таблицах 1 и 2 собраны значения медиан и медианных абсолютных отклонений от медиан, характеризующих средние и



**Рис. 5.** Приближения функций плотности сглаженных (сплошная линия) и предсказанных (пунктир) логарифмов численности для когорты 20 (а) и 36 (б) по возрастам.

отклонения от средних для этих плотностей. Статистические оценки рассчитаны по приближениям для функций, но могут быть оценены непосред-

ственно по сгенерированным выборкам. На рисунке б показаны исходные логарифмы численности, результаты их моделирования на первом этапе, а



**Рис. 6.** Исходные логарифмы численности (сплошная линия), результаты их моделирования на первом этапе (длинный пунктир), медианы для сглаженных (короткий пунктир) и предсказанных (чередующийся короткий и длинный пунктир) функций плотности логарифмов численности для когорт 20 (а) и 36 (б) по возрастам.

также характеристики средних значений для приближенных сглаженных

и предсказанных функций плотности для когорт 20 и 36.

**Таблица 1.** Исходные логарифмы численности, их значения по модели первого этапа и характеристики средних и отклонений от средних значений для плотностей сглаженных логарифмов численности для когорты 20 по возрастам

Возраст	Данные	Значения по модели первого этапа	Медианы сглаженной плотности	Медианные абсолютные отклонения от медиан
2	0,489267	0,489267	0,474213	0,0043015
3	0,336265	0,336265	0,305045	0,00671957
4	0,0962643	0,0962643	0,0481502	0,00948742
5	-0,173735	-0,73444	-0,237414	0,012531
6	-0,473739	-0,490369	-0,568431	0,0164324
7	-0,802181	-0,818518	-0,909328	0,0165735
8	-1,14109	-1,14866	-1,25409	0,0192269
9	-1,48386	-1,47911	-1,601	0,0232779
10	-1,82803	-1,80962	-1,94793	0,0259134
11	-2,1727	-2,14014	-2,29551	0,0306497
12	-2,50654	-2,47066	-2,64014	0,0330563
13	-2,82836	-2,80118	-2,97918	0,0364911
14	-3,13707	-3,13169	-3,30899	0,0393642
15	-3,43207	-3,46221	-3,62515	0,037014
16	-0,484947	-3,79272	-0,207941	0,039498

**Таблица 2.** Исходные логарифмы численности, их значения по модели первого этапа и характеристики средних и отклонений от средних значений для плотностей сглаженных логарифмов численности для когорты 36 по возрастам

Возраст	Данные	Значения по модели первого этапа	Медианы сглаженной/предсказанной плотности	Медианные абсолютные отклонения от медиан
2	1,82417	1,82417	1,82014	0,00852048
3	1,67076	1,67076	1,66622	0,0123917
4	1,41277	1,41277	1,40703	0,0165417
5	1,11076	1,11103	1,10705	0,0181214
6	0,760764	0,738236	0,73468	0,0175327
7	0,360942	0,338696	0,335027	0,0233144
8	-0,0661291	-0,075776	-0,0810629	0,0270546
9	-0,511693	-0,502806	-0,507583	0,0310077
10	-0,972385	-0,941957	-0,947523	0,0389295
11			-1,39964	0,0463516
12			-1,85087	0,0541913
13			-2,30544	0,058154

## ОБСУЖДЕНИЕ

В работе описана методика, в соответствии с которой в два этапа конструируются модели временных рядов оценок численности когорт и уловов при дополнительных предположениях о функциональных зависимостях, которые описывают коэффициенты естественной и промысловой смертности. Метод частиц и приближение функций плотности с помощью ядра позволяют численно оценивать функции плотности апостериорных распределений. Применение методики продемонстрировано на тестовых данных из ВНИРО. Результаты численных расчётов на этом наборе показывают, что усреднённые значения логарифмов численности когорт, рассчитанные по построенным моделям, в целом воспроизводят характер поведения численности когорт в соответствии с данными, которые содержатся в исходных таблицах, включая аномалии.

Зависимость получаемых результатов от числа используемых частиц  $N$  не исследовалась. Кроме того, не проводилась проверка различных опций для функции  $kde$ , с помощью которой формировались приближения функций плотности по выборкам. Разработанная методика является достаточно ресурсоёмкой с вычислительной точки зрения. Так для тестового набора расчёты оценок плотностей и их статистических характеристик для 45 когорт занимали время, далеко выходящее за границы рекомендаций для рыбохозяйственных моделей, например: Cotter et al., 2004.

Отметим также, что при описании распределений невязок между исходными данными и значениями, сгенерированными по детерминистической модели, предполагалась взаимная независимость по годам и возрастам. Возможно

оценивание их распределения отдельно для каждого возраста.

Далее, поскольку на этапе построения базовой модели задачи оптимизации решаются численными методами, показатели близости исходных данных и значений могут оцениваться любым эффективным образом, например, как взвешенная сумма максимумов модулей невязок по логарифмам численности и уловам. На первом этапе численные решения нелинейных оптимизационных задач с ограничениями определяют наборы значений параметров, обеспечивающие только локальные экстремумы. Кроме того, оптимизируемые функции зачастую являются «плоскими» в районах экстремумов (существенные изменения параметров приводят к незначительному изменению самой функции). Поэтому могут возникать конкурирующие между собой различные варианты детерминистической модели. Можно расширить описанную методику, добавив сравнение оценок правдоподобия альтернативных наборов параметров моделей, соответствующих разным показателям близости и/или локальным экстремумам. Функции правдоподобия можно рассчитывать, например, как  $p(x_{a_1:a_l}, y_{a_1:a_l})$  или  $p(y_{a_1:a_l})$  (Schnute, 1994; Kitagawa, 1996; Doucet et al., 2000; de Valpine, 2002).

И наконец, следует учитывать (например: Schnute, 1994), что попытки применения любой методики для построения модели, соответствующей имеющимся данным, могут приводить к неудачам не только по техническим причинам, но из-за того, что сделанные предположения о механизмах взаимосвязи величин не соответствуют реальности или сами данные получены со слишком большой погрешностью.

СПИСОК ЛИТЕРАТУРЫ

- Нестеров Ю.Е.* Методы выпуклой оптимизации. М.: Изд-во МЦНМО, 2010. 281 с.
- Шевченко И.И.* Моделирование промысловых запасов при известных оценках возрастной структуры популяций и уловов // *Вопр. рыболовства*. 2017. Т. 18. № 4. С. 507–519.
- Шевченко И.И.* Моделирование промысловых запасов при известных оценках возрастной структуры популяций и уловов. II // *Вопр. рыболовства*. 2019. Т. 20. № 2. С. 152–163.
- Шевченко И.И.* Применение методов частиц при байесовском оценивании временных рядов численности и уловов по возрастам // *Вопр. рыболовства*. 2020. Т. 21. № 2. С. 235–249.
- Bezanson J., Edelman A., Karpinski S. et al.* Julia: A fresh approach to numerical computing // *SIAM Rev. Soc. Ind. Appl. Math.* 2017. № 59. P. 65–98.
- Cotter A. J. R., Burt L., Paxton C. G. M. et al.* Are stock assessment methods too complicated? // *Fish. Fish.(Oxf)*. 2004. V. 5. № 3. P. 235–254.
- deValpine P.* Review of methods for fitting time-series models with process and observation error and likelihood calculations for nonlinear, non-Gaussian state-space models // *Bull. Mar. Sci.* 2002. V. 70. № 2. P. 455–471.
- deValpine P., Hastings A.* Fitting population models incorporating process noise and observation error // *Ecol. Monogr.* 2002. V. 72. № 1. P. 57–76.
- Doucet A., Godsill S., Andrieu C.* On sequential Monte Carlo sampling methods for Bayesian filtering // *Stat. Comput.* 2000. V. 10. № 3. P. 197–208.
- Gordon N.J., Salmond D.J., Smith A. F.M.* Novel approach to nonlinear/non-Gaussian Bayesian state estimation // *IEE Proceedings F (Radar and Signal Processing)*. 1993. V. 140. P. 107–113.
- Gramacki A.* Nonparametric kernel density estimation and its computational aspects. Cham: Springer, 2018. 171 p.
- Hardle W., Muller M., Sperlich S. et al.* Nonparametric and semiparametric models. V. 1. Cham: Springer, 2004.
- Kitagawa G.* Monte Carlo filter and smoother for non-Gaussian nonlinear state space models // *J. Comput. Graph.Stat.* 1996. V. 5. № 1. P. 1–25.
- Kitagawa G.* Non-Gaussian state-space modeling of nonstationary time series // *J. Am. Stat. Assoc.* 1987. V. 82. № 400. P. 1032–1041.
- Megrey B.A.* Review and comparison of age-structured stock assessment models // *Amer. Fish. Symp.* 1989. V. 6. P. 8–48.
- Quinn T. J., Deriso R. B.* Quantitative fish dynamics. Oxford University Press, USA, 1999. 542 p.
- Russell S. J., Norvig P.* Artificial Intelligence: A Modern Approach. Upper Saddle River: Prentice Hall, 2010. 1152 p.
- Sarkka S.* Bayesian filtering and smoothing. Cambridge: Cambridge University Press, 2013. 254 p.
- Schnute J. T.* A general framework for developing sequential fisheries models // *Can. J. Fish. Aquat.Sci.* 1994. V. 51. № 8. P. 1676–1688.

**USING SEQUENTIAL PARTICLE METHODS  
AND NON-PARAMETRIC DISTRIBUTIONS  
IN BAYESIAN EVALUATIONS OF ABUNDANCE  
AND CATCH AT AGE TIME SERIES**

**I. I. Shevchenko**

*The Pacific branch of the Russian Federal Research Institute  
of Fisheries and Oceanography, Vladivostok, 690091*

We describe an approach to analyzing time series for two variables connected through the state model with abundance and catch data sets and cohort and catch equations as an example. First, we create a deterministic model with parameters that maximizes the closeness of given data and data generated by a model. Then, we obtain cohort stochastic models using the difference between initial and modeled data. They are represented as hidden Bayesian models with abundances as states and catches as observations. Using these models, one can evaluate posterior densities and calculate averages, deviations, etc. As a general matter, the recursive equations met by posterior densities have no analytic solutions. We describe several particle methods that may be used for density approximations and following calculations of their statistical quantities. All generated sample densities are smoothed with non-parametric kernel density estimation. The Fishmetica package was extended with functions for generating samples and weights for filtering, predicting and smoothing densities. Numerical simulations were conducted for a test data set. Several extensions of the approach are proposed including an additional option for comparing the basic models with the use of a likelihood function.

*Key words:* state space models, hidden Bayesian models, kernel density estimation, filtering, prediction, smoothing, likelihood of model parameters.