

Корректировка оценок запаса в когортной модели с помощью фильтра Калмана и данных учетных съеомок

- Модели с возрастной структурой являются приоритетными для оценки запаса

Важные разработки в классе когортных моделей

VPA (Fry, 1949; Gulland, 1965)

CA (Pope, 1972)

SVPA (Pope, 1974; Pope, Shepherd, 1982)

CAGEAN (Deriso et al., 1985)

ADAPT (Gavaris, 1988)

CASA (Sullivan et al., 1990)

Stock Synthesis (Methot, 1990)

XSA (Shepherd, 1992)

ISVPA (Vasilyev, Kizner, 1998)

Joint Time Series Analysis (Gudmundsson, 1994)

ОСНОВНЫЕ СООТНОШЕНИЯ В КОГОРТНЫХ МОДЕЛЯХ

Динамика численности и вылова в когортных моделях

$$(1) \quad n_{a+1,t+1} = (1 - \varphi_{at})sn_{at}$$

$$(2) \quad c_{at} = \varphi_{at}s^{\Delta}n_{at}$$

n_{at} — численность возрастного класса a в год t в запасе;
 c_{at} — численность возрастного класса a в год t в улове;
 φ_{at} — коэффициент эксплуатации для возрастного класса a в год t ;

s — коэффициент естественной выживаемости;

Δ — доля года, когда промысла нет.

Обратный расчет коэффициентов эксплуатации в когортных моделях

$$(3) \quad \varphi_{at} = \frac{\varphi_{a+1,t+1}}{\alpha_{at} + \varphi_{a+1,t+1}},$$

$$\alpha_{at} = \frac{c_{a+1,t+1}}{sc_{at}}.$$

$$\varphi_{a+1,t+1} = \frac{\alpha_{at}\varphi_{at}}{1 - \varphi_{at}}$$

Обратное решение корректно, т.к. $\varphi_{at} < 1$ для всех $a = \overline{1, A-1}$ и $t = \overline{1, T-1}$, а φ_{aT} и φ_{At} заданы.

Специальное представление сепарабельных коэффициентов эксплуатации (Васильев, 2001):

$$(4) \quad \varphi_{sep_{at}} = sel_a f_t;$$

$$f_t = \sum_a \varphi_{at};$$

f_t — фактор года t в
коэффициентах эксплуатации;

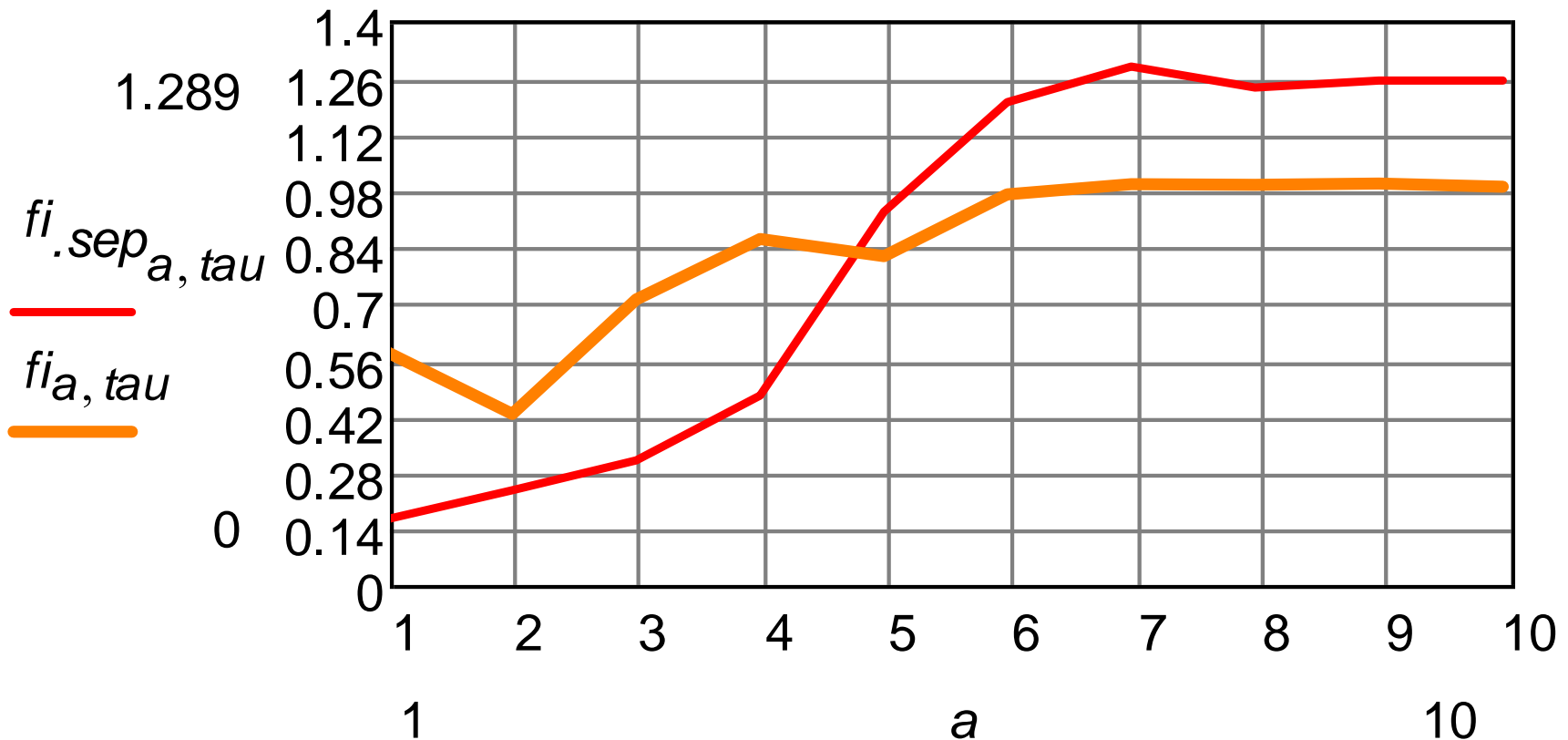
$$sel_a = \frac{\sum_t \varphi_{at}}{\sum_t f_t};$$

sel_a — фактор возраста a в
коэффициентах эксплуатации;

$$\sum_a sel_a = 1.$$

- Специальное представление ($\sum_a sel_a = 1$) гарантирует, что $f_t = \sum_a \varphi_{sep_{at}}$.
- Однако в этом случае корректность решения нарушается, т.к. появляются $\varphi_{sep_{at}} > 1$.

Иллюстрация к некорректности решений при сепарабельном представлении коэффициентов эксплуатации

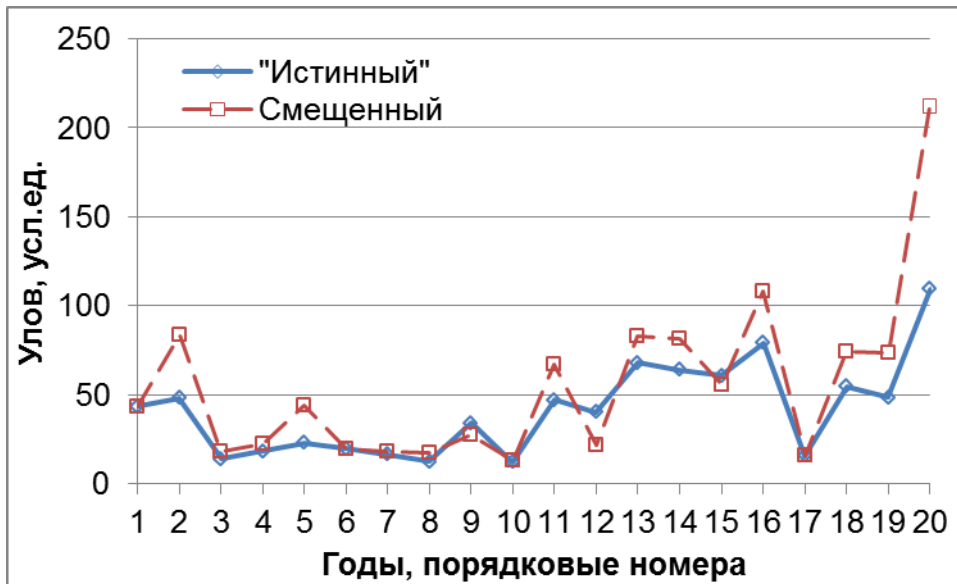


- Если данные о вылове содержат систематическую ошибку, то оценки запаса в когортной модели будут смещены.
- Например, если имеется неучтенный вылов, то оценки запаса будут регулярно занижаться.

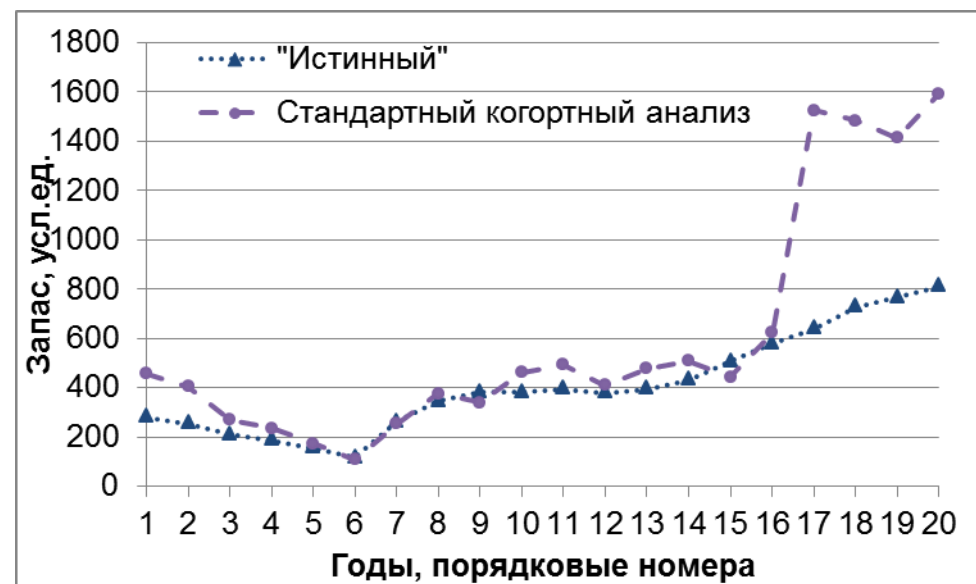
ПРОБЛЕМА С ДАННЫМИ О ВЫЛОВЕ

Данные по уловам содержат только шум высокой интенсивности

Данные



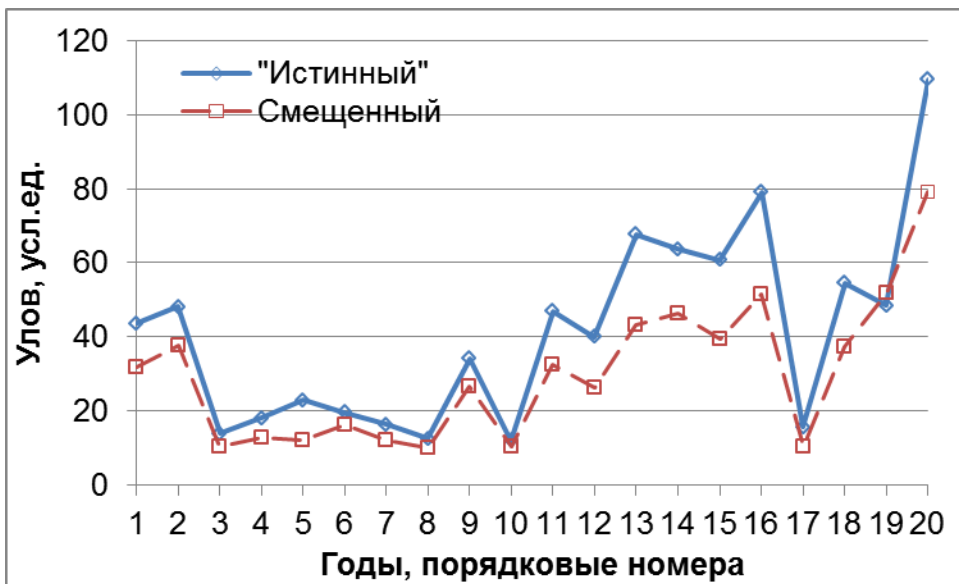
Оценка когортной модели



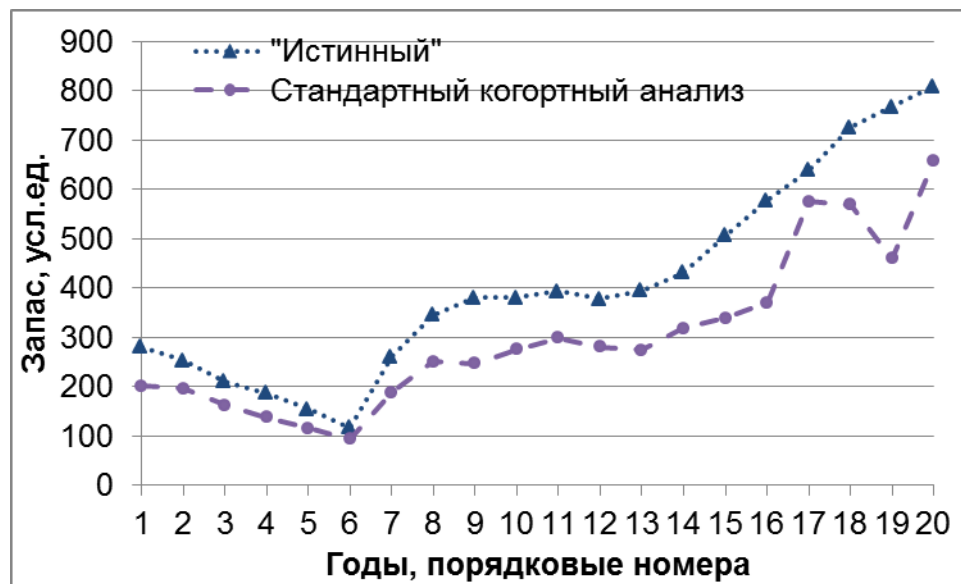
Шум логнормальный

Данные по уловам содержат помимо шума слабой интенсивности систематическое занижение

Данные



Оценка когортной модели



Шум логнормальный

- Для устранения смещения в оценках запаса можно попытаться использовать данные наблюдений за индексами запаса, как это делается, например, в когортных моделях типа ADAPT (Gavaris, 1988).

ОДИН ИЗ СПОСОБОВ РЕШЕНИЯ ПРОБЛЕМЫ С ДАННЫМИ О ВЫЛОВЕ

Уравнения наблюдения:

$$(5) \quad I_{it} = H_{it}N_t + \varepsilon_{it};$$

$$H_{1,t} = s\lambda \sum_{a=1}^A m_a p_{a,t}; \quad H_{2,t} = qs^{\Delta}; \quad H_{3,t} = ks^{\Delta_s}.$$

I_{it} — индексы запаса, $i = 1, 2, 3, \dots$;

ε_{it} — гауссовские шумы наблюдений с неизвестными дисперсиями σ_i^2 ;

$$p_{at} = \frac{n_{at}}{N_t};$$

q, k — коэффициенты улавливаемости для промысла и съемки;

Δ_s — время до начала съемки в долях года;

λ — относительная популяционная плодовитость;

$\{m_a\}_{a=1}^A$ — огиба половозрелости;

- Поскольку уравнения наблюдения оперируют с запасом без привязки к возрасту, то для настройки когортной модели ее целесообразно свести к производственной модели.
- Если допустить, что искажение в данных о вылове пропорционально распределяются по всем возрастным классам, то когортная модель сводится к производственной модели с пополнением.

СВЕДЕНИЕ КОГОРТНОЙ МОДЕЛИ К ПРОДУКЦИОННОЙ

Уравнение процесса

$$(6) \quad N_{t+1} = sN_t - s^{1-\Delta}C_t + R_t + \varepsilon_{N_t}$$

$$N_t = \sum_a n_{at};$$

$$C_t = \sum_a c_{at};$$

$R_t = n_{1,t}$ — пополнение

ε_{N_t} — гауссовский шум с неизвестной дисперсией σ_N^2 .

Уравнения (1)-(2):

$$(1) \quad n_{a+1,t+1} = (1 - \varphi_{at})sn_{at} = sn_{at} - s^{1-\Delta}c_{at}$$

$$(2) \quad c_{at} = \varphi_{at}s^\Delta n_{at}$$

суммированием всех переменных по индексу a сводятся к уравнению процесса (6) в виде стохастической продукционной модели с пополнением.

- Однако настройка когортной модели по индексам запаса осложняется тем, что эти данные также содержат ошибки, хотя, как правило, не систематические.
- Подходящим инструментом для корректировки оценок когортной модели по индексам запаса является фильтр Калмана, поскольку он разделяет шум в модели процесса и ошибку наблюдений в общей неопределенности оценки.
- ФК широко применяется в когортных моделях (Sullivan, 1992; Schnute, 1994; Gudmundsson, 1998).

ОСНОВЫ РАБОТЫ ФИЛЬТРА КАЛМАНА (ФК)



Rudolf Emil Kalman

(Kalman, 1960; Kalman, Bucy, 1961)

Уравнение коррекции оценки

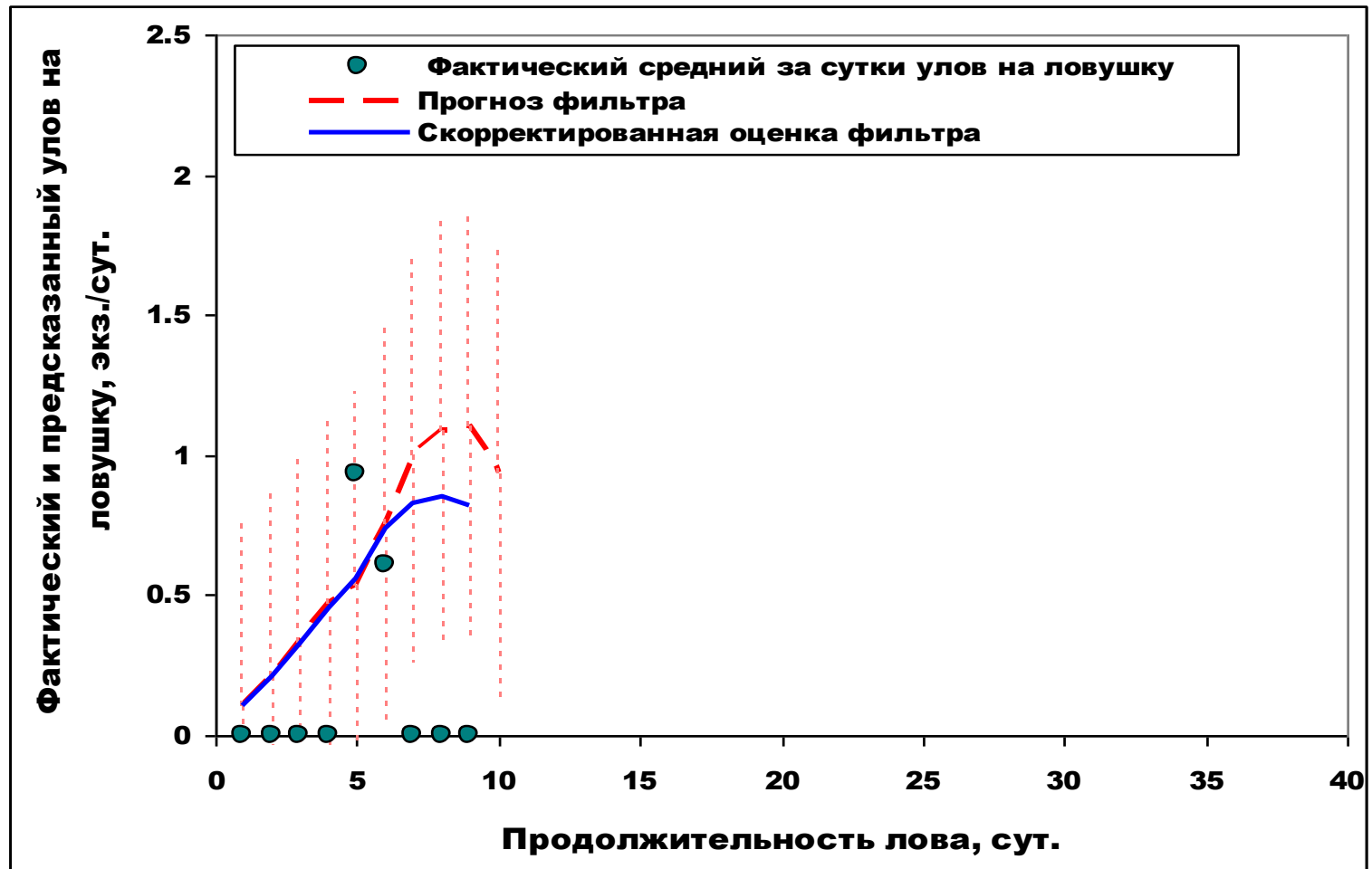
$$(7) \quad N_{cor_t} = w_t N_t^+ + (1 - w_t) y_t.$$

$$w_t \rightarrow 1; \quad N_{cor_t} \rightarrow N_t^+;$$

$$w_t \rightarrow 0; \quad N_{cor_t} \rightarrow y_t.$$

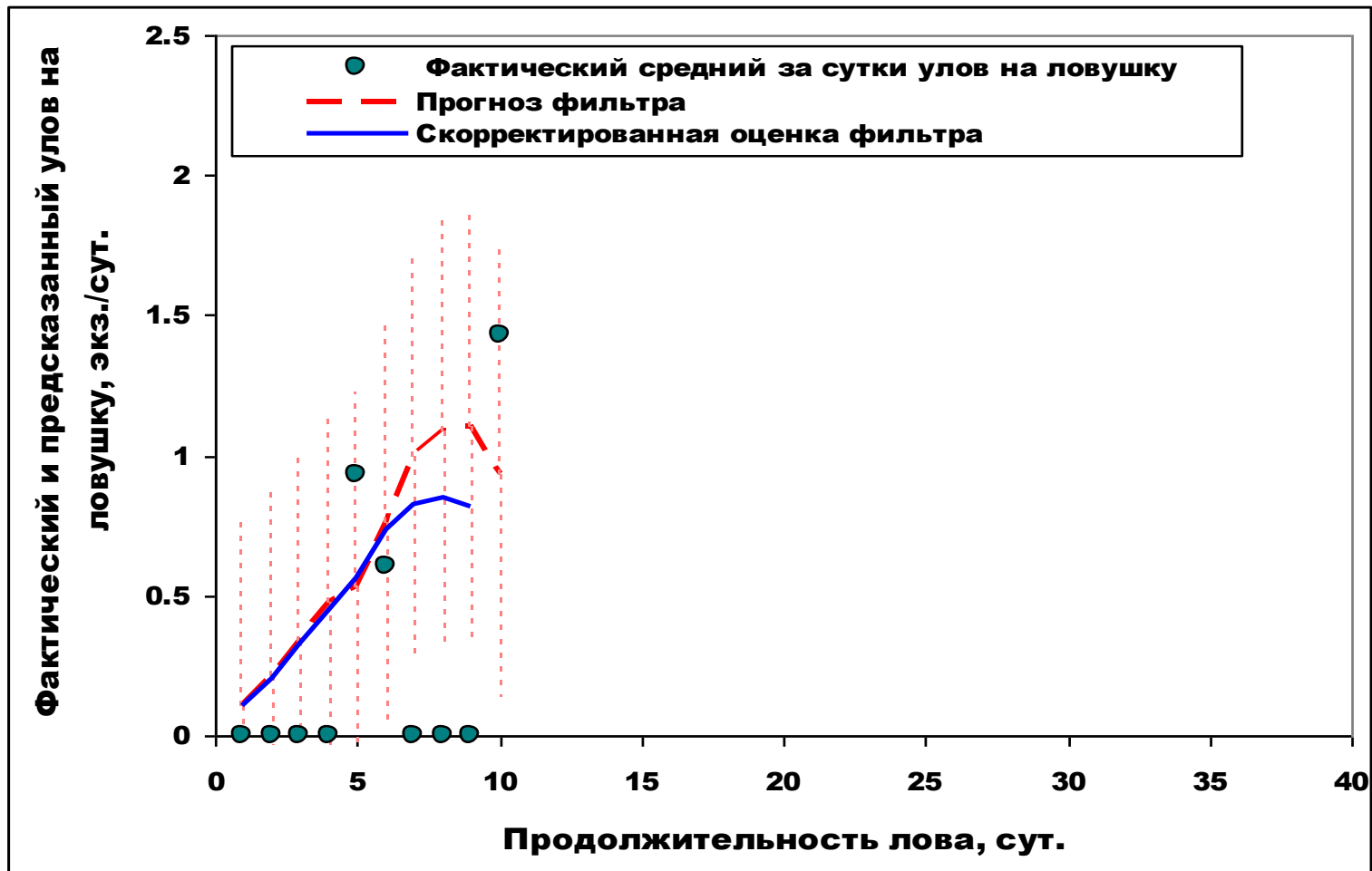
ПРИНЦИП РАБОТЫ ФИЛЬТРА КАЛМАНА

Прогноз индекса численности на момент времени $t=10$ сут.



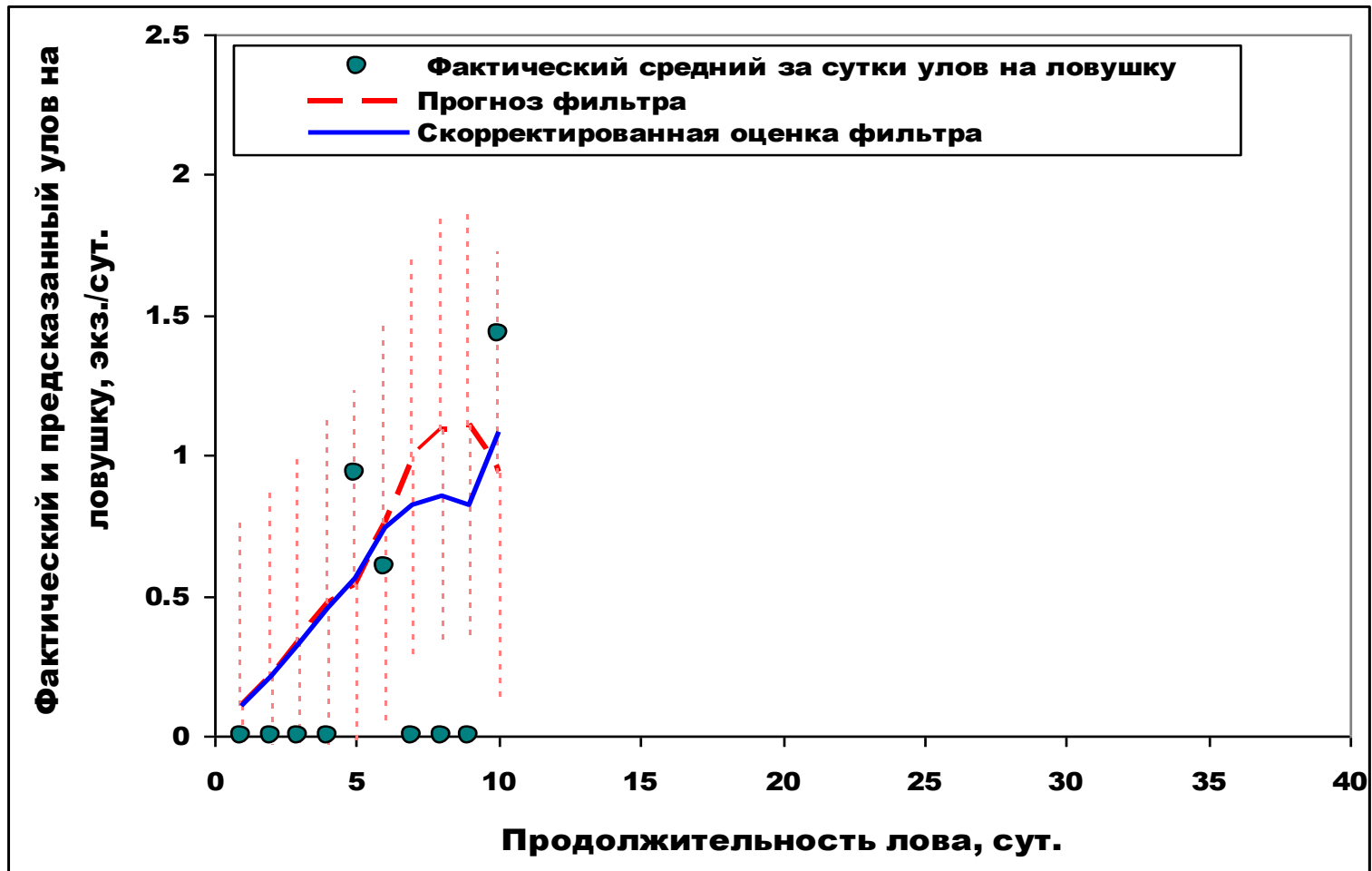
ПРИНЦИП РАБОТЫ ФИЛЬТРА КАЛМАНА

Наблюдение индекса численности на момент времени $t=10$ сут.



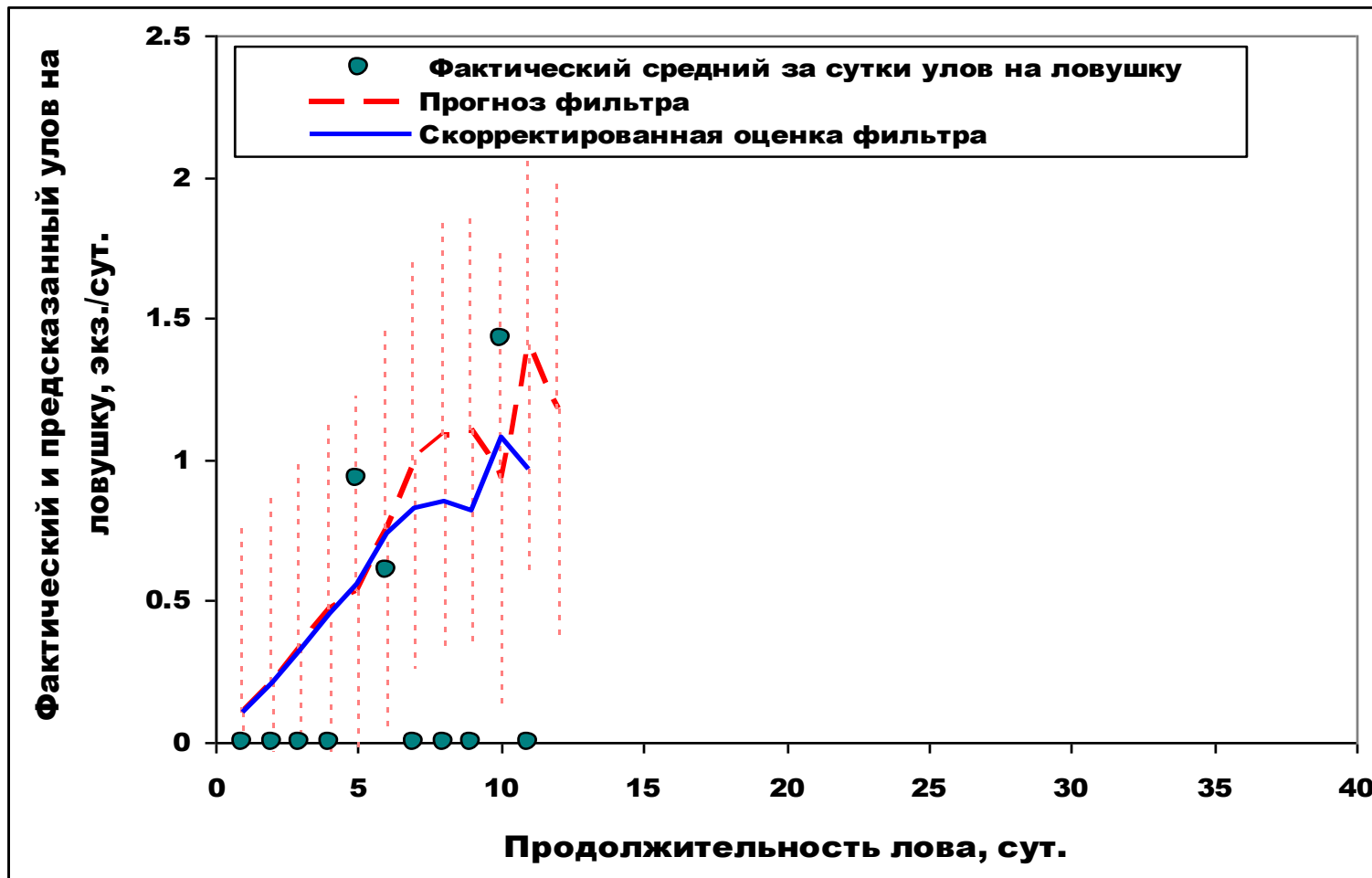
ПРИНЦИП РАБОТЫ ФИЛЬТРА КАЛМАНА

Коррекция индекса численности на момент времени $t=10$ сут.



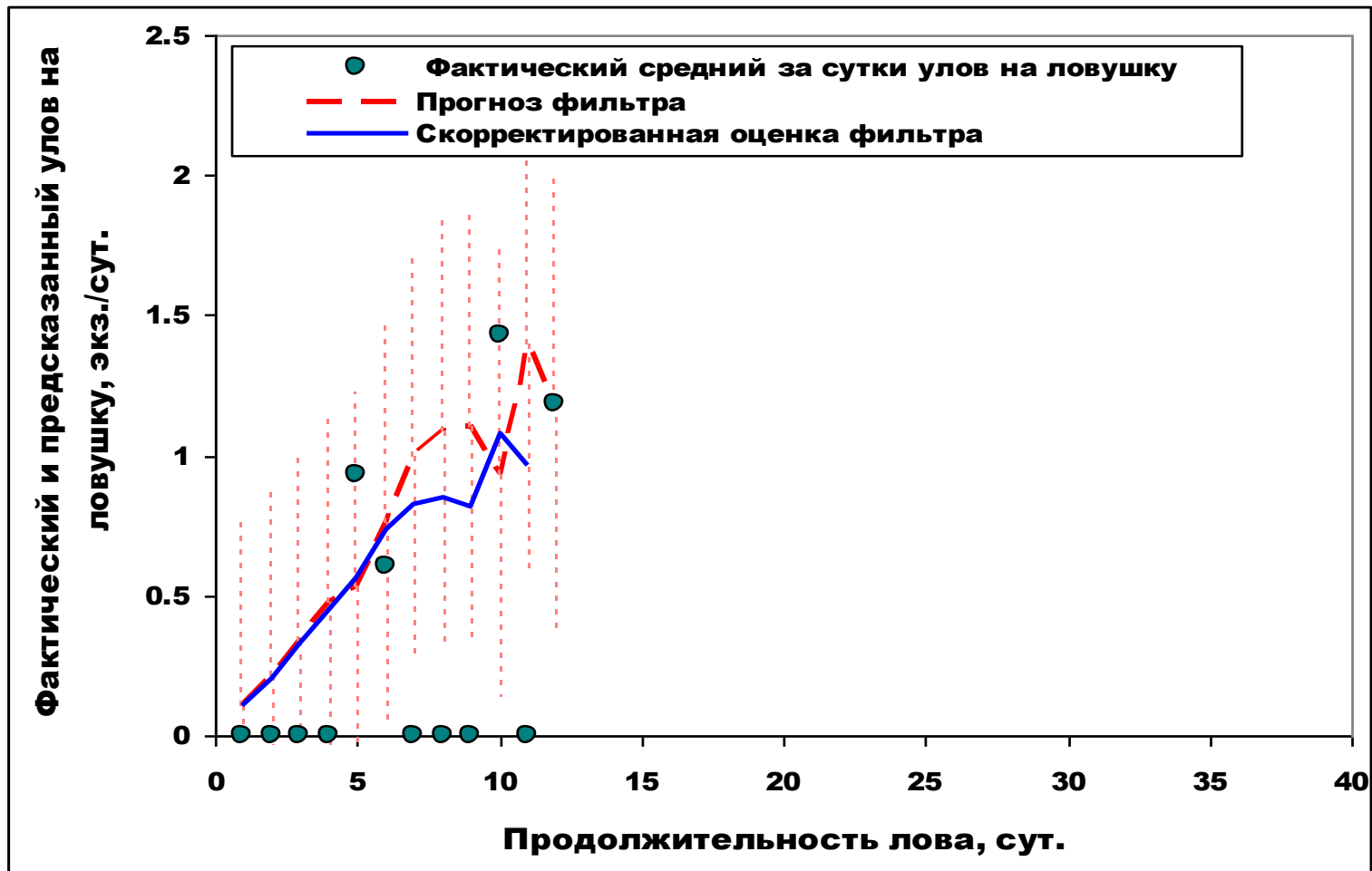
ПРИНЦИП РАБОТЫ ФИЛЬТРА КАЛМАНА

Прогноз индекса численности на момент времени $t=12$ сут.



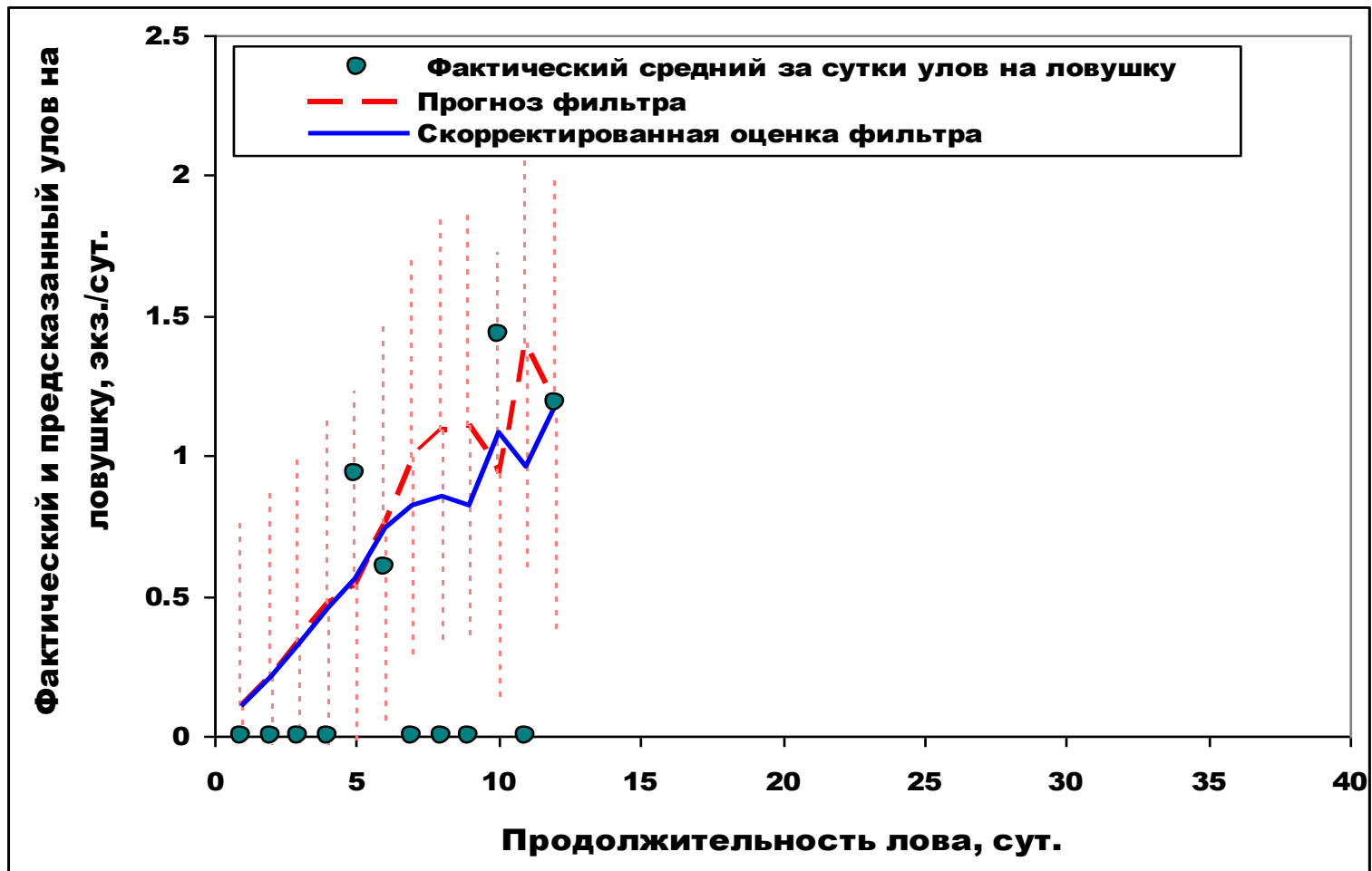
ПРИНЦИП РАБОТЫ ФИЛЬТРА КАЛМАНА

Наблюдение индекса численности на момент времени $t=12$ сут.



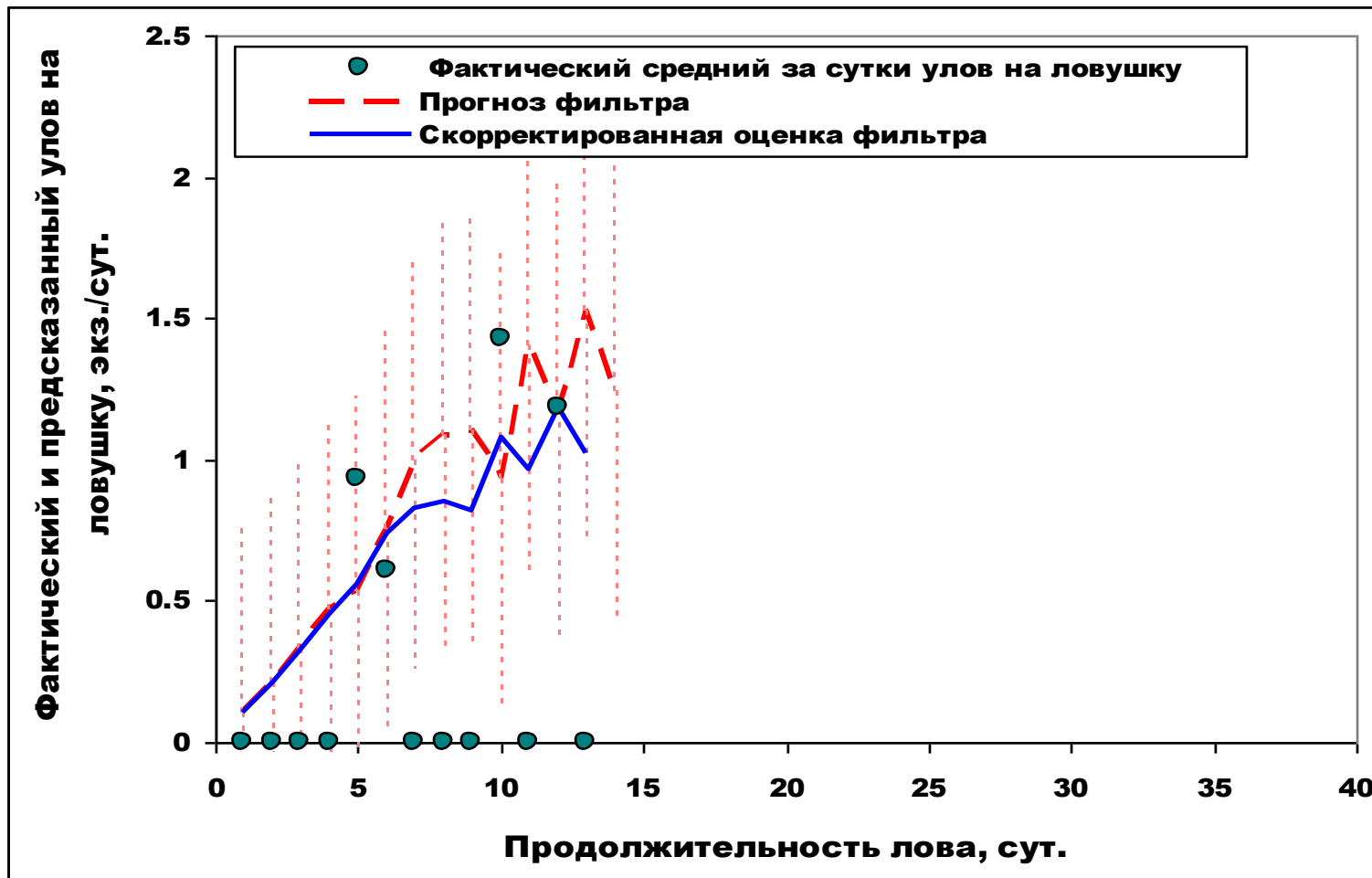
ПРИНЦИП РАБОТЫ ФИЛЬТРА КАЛМАНА

Коррекция индекса численности на момент времени $t=12$ сут.



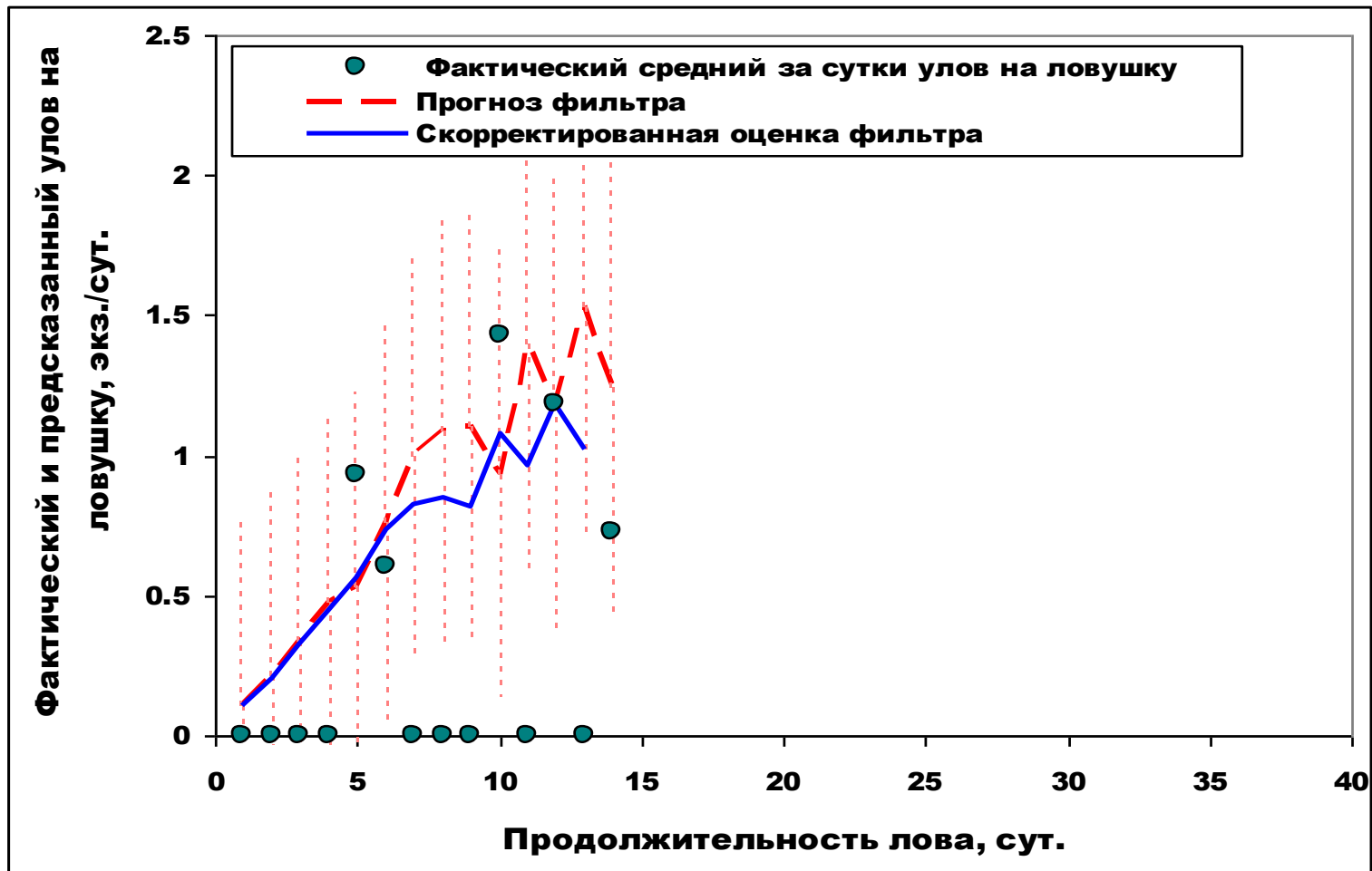
ПРИНЦИП РАБОТЫ ФИЛЬТРА КАЛМАНА

Прогноз индекса численности на момент времени $t=14$ сут.



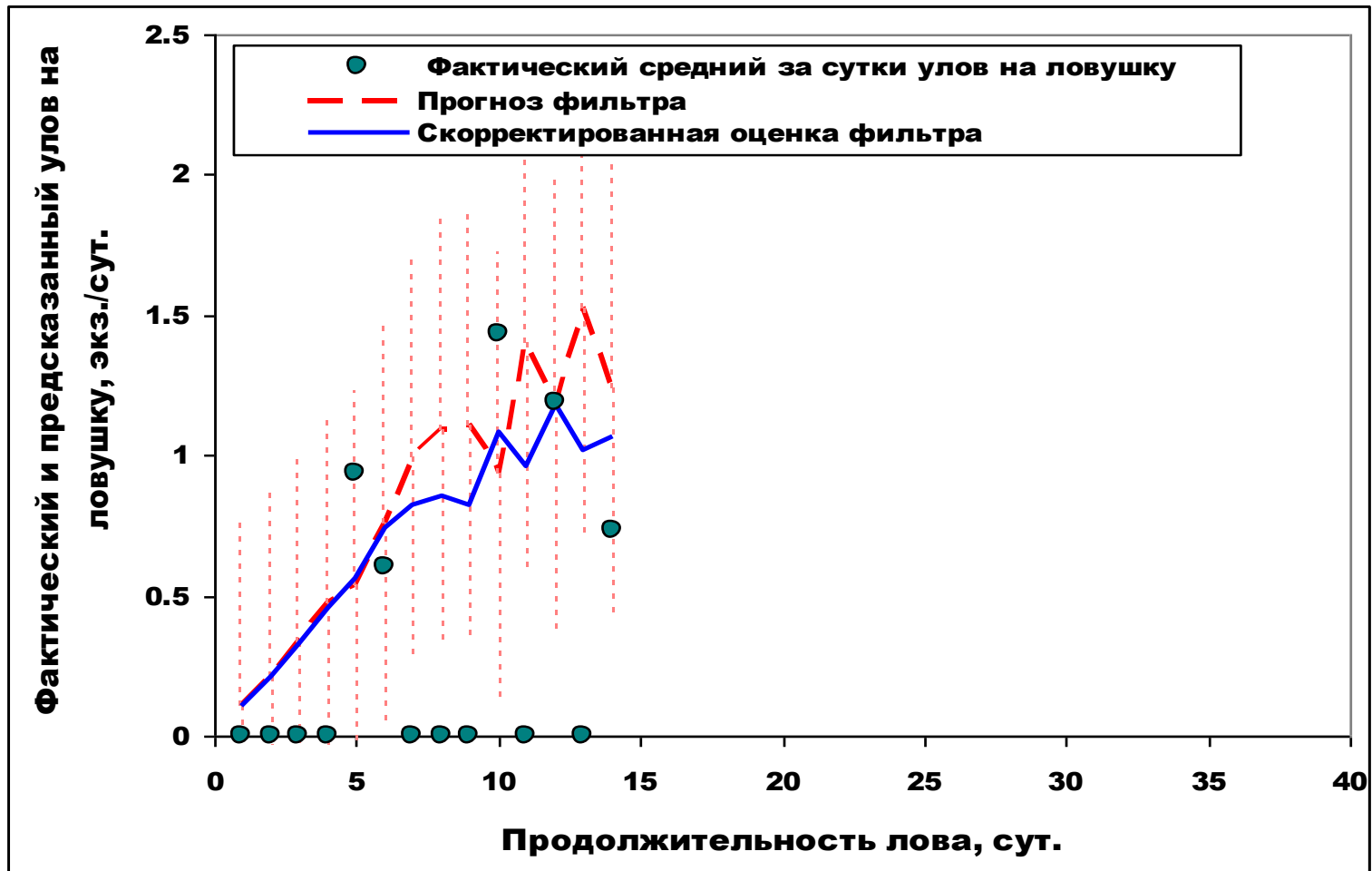
ПРИНЦИП РАБОТЫ ФИЛЬТРА КАЛМАНА

Наблюдение индекса численности на момент времени $t=14$ сут.



ПРИНЦИП РАБОТЫ ФИЛЬТРА КАЛМАНА

Коррекция индекса численности на момент времени $t=14$ сут.



Вес для коррекции оценки

$$(8) \quad w_t = \frac{\prod_i s_i^2}{D_t},$$

$$\prod_i s_i^2 = (s_1 s_2 s_3 \dots)^2;$$

s_i^2 — дисперсия ошибок в наблюдениях из i -того источника данных;

D_t — дисперсия обновляющего процесса.

В случае трех источников данных:

$$(9) \quad D_t = P_t^+ \left[(H_{i,t} s_j s_k)^2 + (H_{j,t} s_i s_k)^2 + (H_{k,t} s_i s_j)^2 \right] + (s_i s_j s_k)^2,$$

$$i, j, k = 1, 2, 3; \quad i \neq j; \quad i \neq k; \quad j \neq k.$$

P_t^+ — ковариация ошибки прогноза

Уравнения (8)-(9) получены из требования статистической эффективности оценки, т.е. ее несмещенности и состоятельности (Kalman, 1960),

Коэффициент усиления Калмана

$$(10) \quad K_{i,t} = \frac{P_t^+ H_{i,t} (s_j s_k)^2}{D_t},$$

Из уравнений (8)-(9)

$$(8) \quad w_t = \frac{\Pi_i s_i^2}{D_t},$$

$$(9) \quad D_t = P_t^+ \left[(H_{i,t} s_j s_k)^2 + (H_{j,t} s_i s_k)^2 + (H_{k,t} s_i s_j)^2 \right] + (s_i s_j s_k)^2,$$

легко получить тождество:

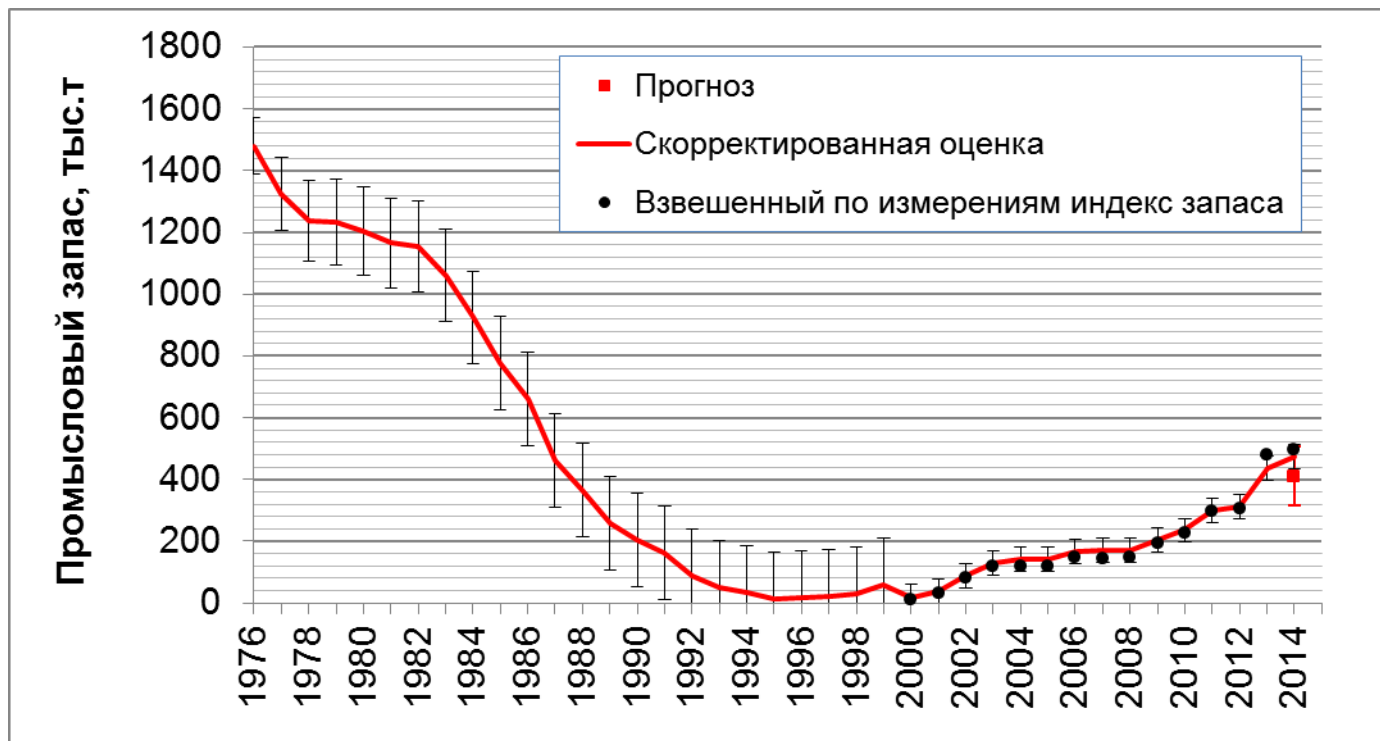
$$(11) \quad \sum_i K_{i,t} H_{i,t} + w_t \equiv 1$$

Уравнение (11) показывает, каким должен быть вес в уравнении коррекции

$$(7) \quad N_{cor_t} = w_t N_t^+ + (1 - w_t) y_t,$$

а уравнения (8)-(10) – как его вычислить через параметры фильтра и модели.

Коррекция оценки запаса на примере результатов работы когортной модели с ФК (КАФК) для минтая подзоны ВС



Коррекция прогноза на 2014 год по индексам запаса (данные ихтиопланктонной и траловой НИС-съемки, СРУЕ из отчетов промысловой статистики)

- Ряды данных научных съемок часто имеют пробелы. Необходимо убедиться, что пробелы в наблюдениях не нарушают работу фильтра.

НЕКОТОРЫЕ ОСОБЕННОСТИ РАБОТЫ ФК С РЯДАМИ ДАННЫХ, СОДЕРЖАЩИХ ПРОПУСКИ

Оценки ФК для рядов с пропущенными значениями

Пусть $I_{i,t} = 0$; $i = 1, 2, 3$. Если $H_{i,t} = 0$, то пропущенные значения на оценки не влияют.

Из уравнений (9)-(10):

$$(9) \quad w_t = \frac{\prod_i s_i^2}{D_t},$$

$$(10) \quad D_t = P_t^+ \left[(H_{i,t} s_j s_k)^2 + (H_{j,t} s_i s_k)^2 + (H_{k,t} s_i s_j)^2 \right] + (s_i s_j s_k)^2,$$

следует:

$$w_t = \frac{s_i^2 (s_j s_k)^2}{s_i^2 \left\{ P_t^+ \left[(H_{j,t} s_k)^2 + (H_{k,t} s_j)^2 \right] + (s_j s_k)^2 \right\}} = \frac{(s_j s_k)^2}{P_t^+ \left[(H_{j,t} s_k)^2 + (H_{k,t} s_j)^2 \right] + (s_j s_k)^2}$$

$$K_{i,t} = \frac{P_t^+ H_{i,t} (s_j s_k)^2}{D_t} = 0;$$

$$K_{j,t} = \frac{P_t^+ H_{j,t} (s_i s_j)^2}{D_t} = \frac{P_t^+ H_{j,t} s_k^2}{P_t^+ \left[(H_{j,t} s_k)^2 + (H_{k,t} s_j)^2 \right] + (s_j s_k)^2}.$$

Практический интерес представляет задача получения оценок фильтра из источников с пропущенными значениями.

Уравнения ФК для когортной модели

Начальные значения:

$$N_{cor_1} = N_1,$$

где N_1 — оценка запаса, полученная с помощью когортной модели;
ковариация ошибки оценки:

$$P_1 = \frac{\prod_i s_i^2}{\left[(H_{i,1} s_j s_k)^2 + (H_{j,1} s_i s_k)^2 + (H_{k,1} s_i s_j)^2 \right] + \prod_i s_i^2}.$$

Вычисления осуществляются в цикле по $t = \overline{1, T}$, где T — период лет наблюдений за запасом. Каждая итерация состоит из 4 основных уравнений:

прогноз

$$N_{t+1}^+ = s N_{cor_t} - s^{1-\Delta} C_t + R_{t+1},$$

ковариация ошибки прогноза

$$P_{t+1}^+ = s^2 P_t + \sigma_N^2.$$

скорректированная оценка

$$N_{cor_t} = w_t N_t^+ + \sum_i K_{i,t} I_{i,t},$$

ковариация ошибки коррекции

$$P_t = w_t P_t^+.$$

Целевая функция и веса значимости для источников наблюдений

Целевая функция для вектора оптимизируемых параметров модели

$$\boldsymbol{\theta} = \{s; \Delta; W; \beta; a_{50}; q; k; \Delta_S; \lambda; a_R; \sigma_N; \sigma_E; \sigma_Y; \sigma_S\}:$$

$$L(\boldsymbol{\theta} | Data) = \sum_t \sum_i w'_{i,t} \left(\ln \frac{y_{i,t}}{N_t^+} \right)^2.$$

$w'_{i,t}$ — веса значимости источников данных

- В общем случае используемые для настройки модели источники информации об индексах запаса не равноценны по достоверности.

Тестирования модели КАФК выполнялось с помощью вычислительных экспериментов с рядами данных, сгенерированных с помощью операционной модели.

УРАВНЕНИЯ ОПЕРАЦИОННОЙ МОДЕЛИ

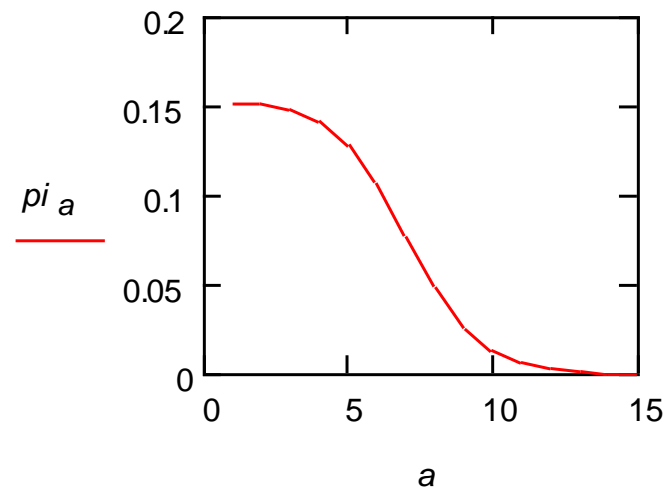
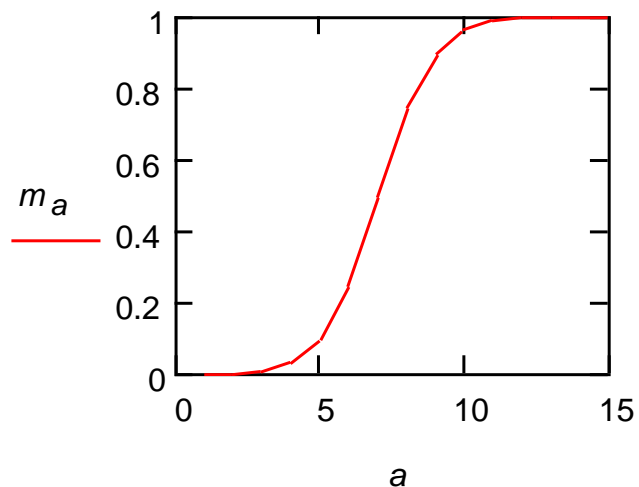
Начальное распределение в запасе

$$N_1 := 10^9 \quad \pi'_a := 1 - \frac{1}{1 + \exp[0.8 \cdot (7 - a)]} \quad n_{a,1} := \pi'_a \cdot N_1$$

Огиба созревания

$$m_a := \frac{1}{1 + \exp[1.1 \cdot (7 - a)]}$$

$$\pi_a := \frac{\pi'_a}{\left(\sum_a \pi'_a \right)}$$



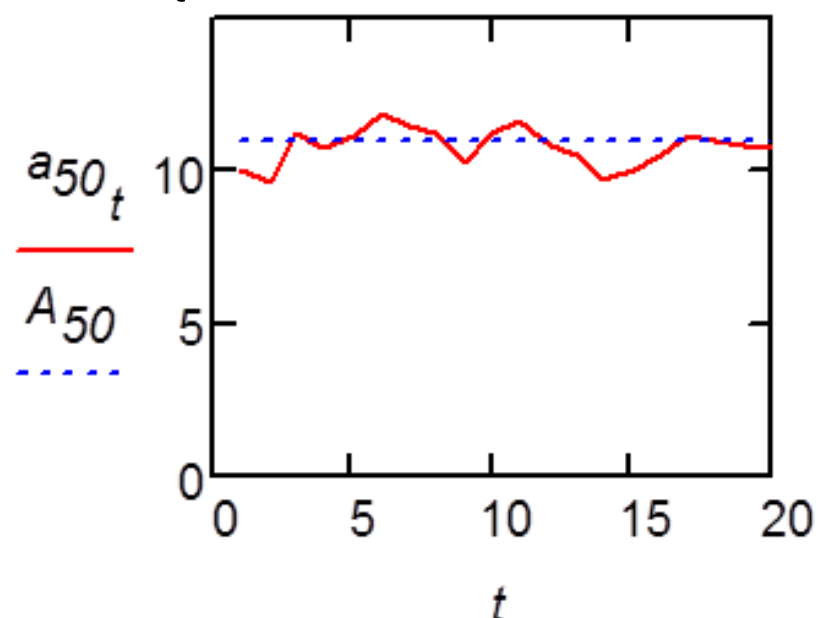
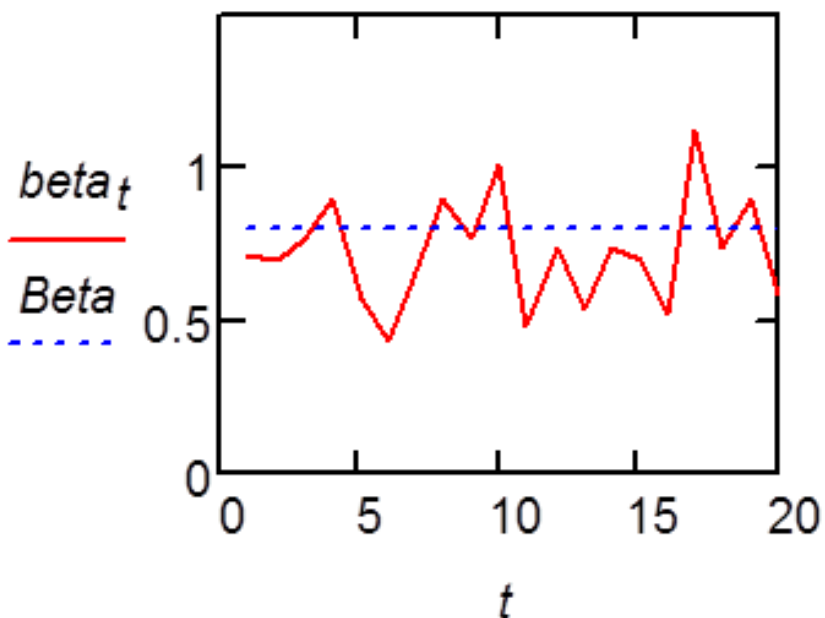
Коэффициенты изъятия

$$Beta := 0.8$$

$$A_{50} := a_R + 5$$

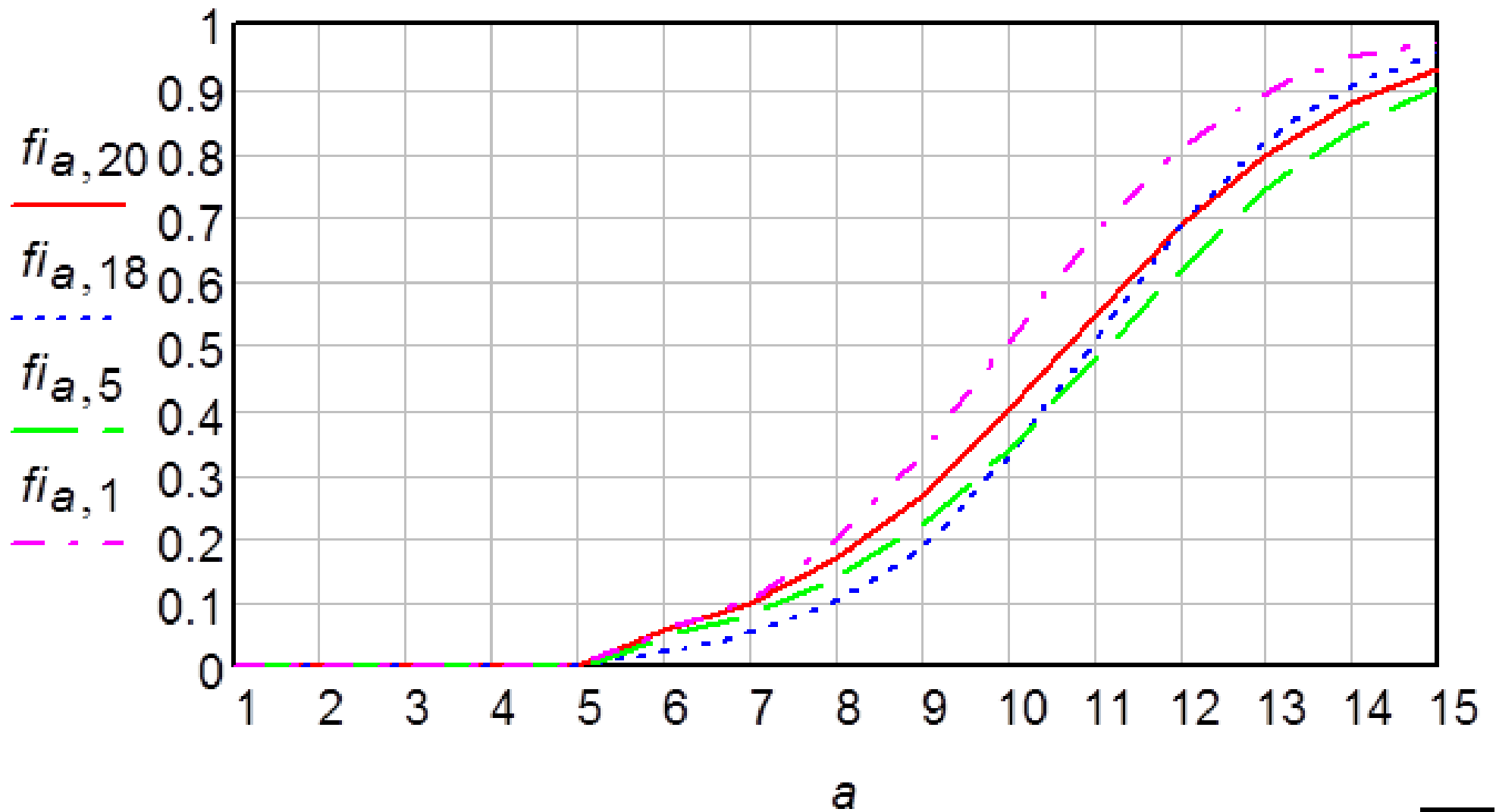
$$beta_t := Beta + rnorm(T, 0, 0.2)_t$$

$$a_{50}_t := A_{50} + rnorm(T, 0, 0.5)_t$$



$$f_{a,t} := \begin{cases} 0 & \text{if } a < a_R \\ \frac{1}{1 + \exp[beta_t \cdot (a_{50}_t - a)]} & \text{otherwise} \end{cases}$$

Коэффициенты изъятия



Генерация динамики запаса

$$s \equiv 0.7$$

$$\text{delta} \equiv 0.2$$

$$s'_R \equiv 2 \cdot 10^{-5}$$

$$a_m := 4$$

$$a_R := 6$$

$$N_t := \sum_{a=a_R}^A n_{a,t}$$

$$R_t := n_{a_R,t}$$

$$\begin{pmatrix} n \\ E \end{pmatrix} :=$$

for $t \in 1 .. T - 1$

for $a \in 1 .. A - 1$

$$n_{a+1,t+1} \leftarrow (1 - fi_{a,t}) \cdot s \cdot n_{a,t}$$

$$E_t \leftarrow 0.5 \cdot \text{lambda} \cdot s \cdot \left[\sum_{a=a_m}^A (m_a \cdot n_{a,t}) \right]$$

$$n_{1,t+1} \leftarrow s'_R \cdot E_t$$

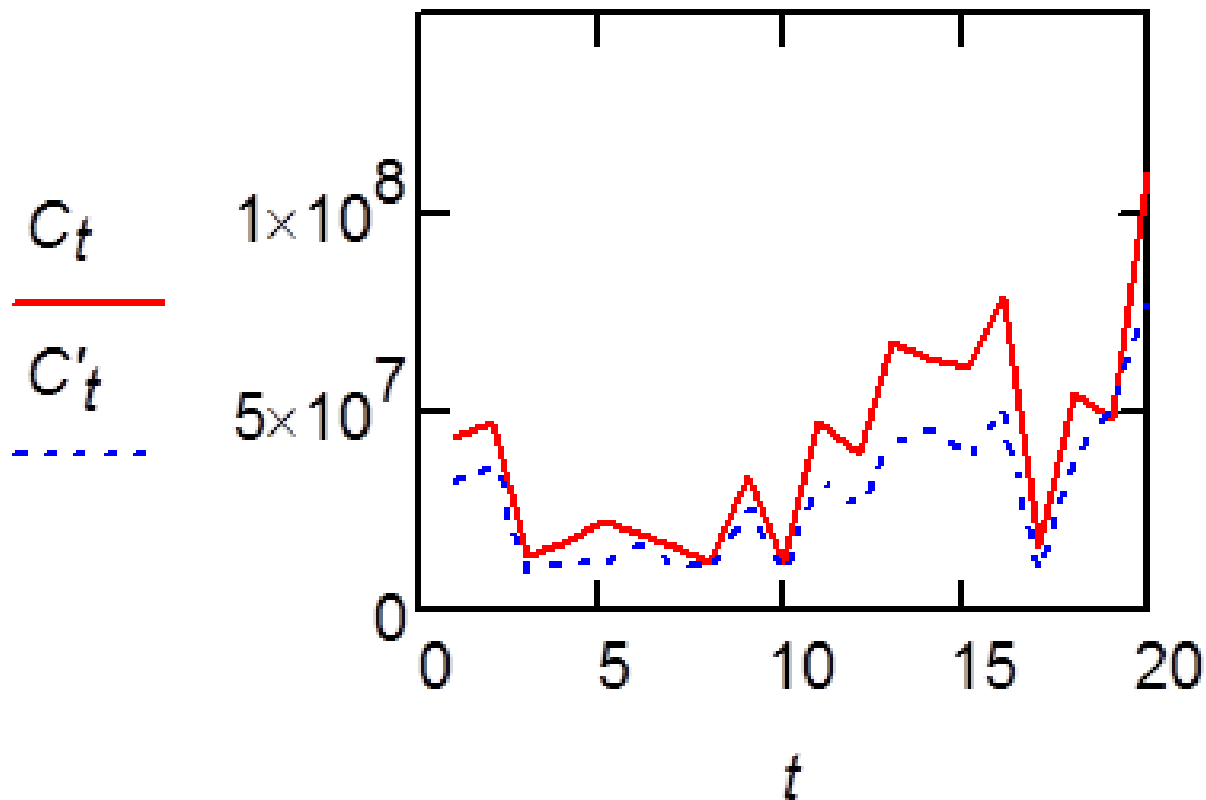
$$E_T \leftarrow 0.5 \cdot \text{lambda} \cdot s \cdot \left[\sum_{a=a_m}^A (m_a \cdot n_{a,T}) \right]$$

$$\begin{pmatrix} n \\ E \end{pmatrix}$$

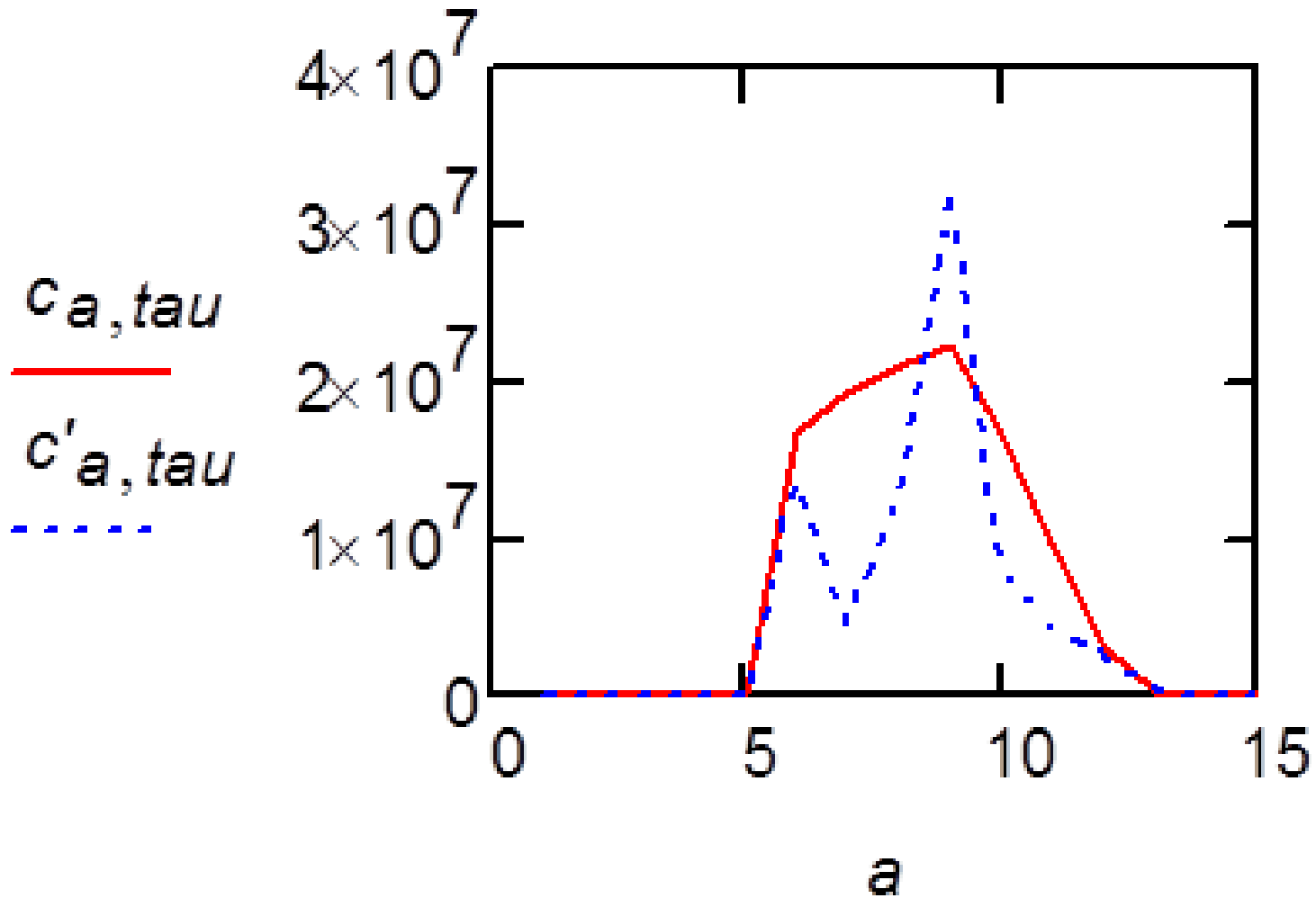
Генерация данных по уловам

$$c_{a,t} := f_{i_{a,t}} \cdot s^{\text{delta}} \cdot n_{a,t} \quad c'_{a,t} := 0.7 \cdot c_{a,t} \cdot \text{rlnorm}(T, 0, 0.4)_t$$

$$C_t := \left(\sum_a c_{a,t} \right) \quad C'_t := \left(\sum_a c'_{a,t} \right)$$



Возрастные составы уловов



Генерация «истинных» и «наблюдаемых» индексов запаса

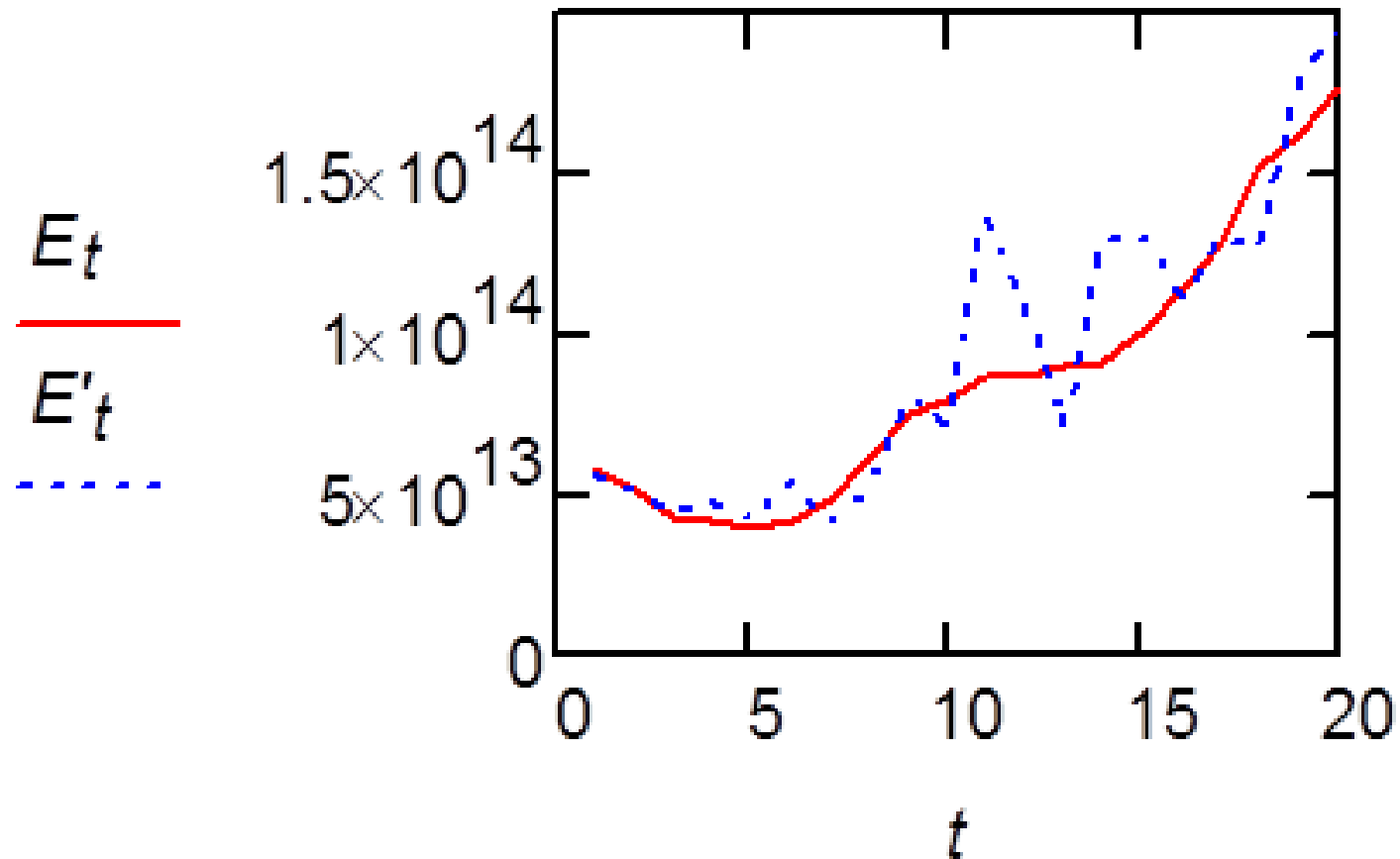
$$q := 2 \cdot 10^{-4} \quad k := 10^{-6} \quad \text{delta}_s := 0.5$$

$$Y_t := q \cdot s^{\text{delta}} \cdot N_t \quad S_t := k \cdot s^{\text{delta}_s} \cdot N_t$$

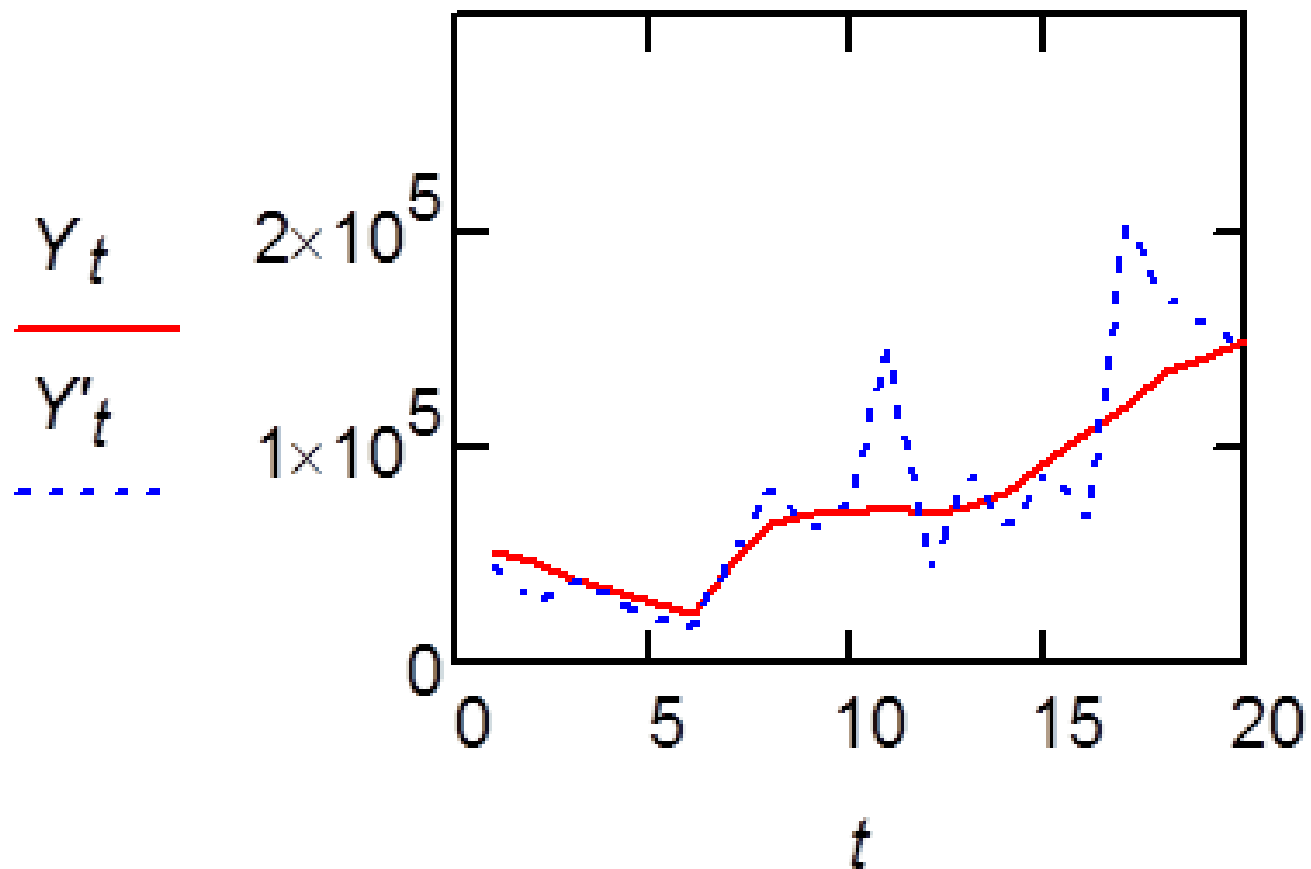
$$E'_t := E_t \cdot \text{rlnorm}(T, 0, 0.2)_t \quad Y'_t := 0.99 \cdot Y_t \cdot \text{rlnorm}(T, 0, .3)_t$$

$$S'_t := 1 \cdot S_t \cdot \text{rlnorm}(T, 0, .2)_t$$

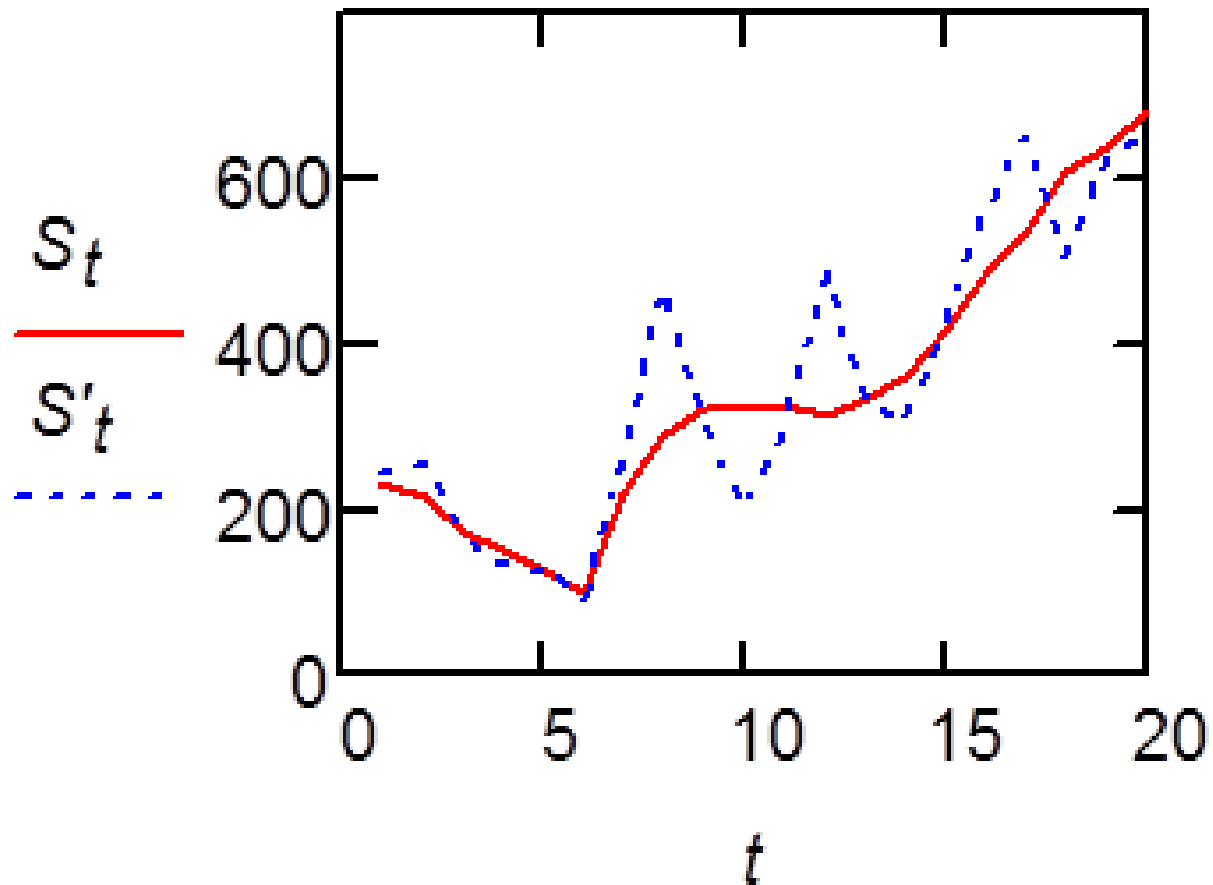
Данные по ихтиопланктонному индексу запаса



Данные по CPUЕ



Данные по учетному траловому индексу запаса

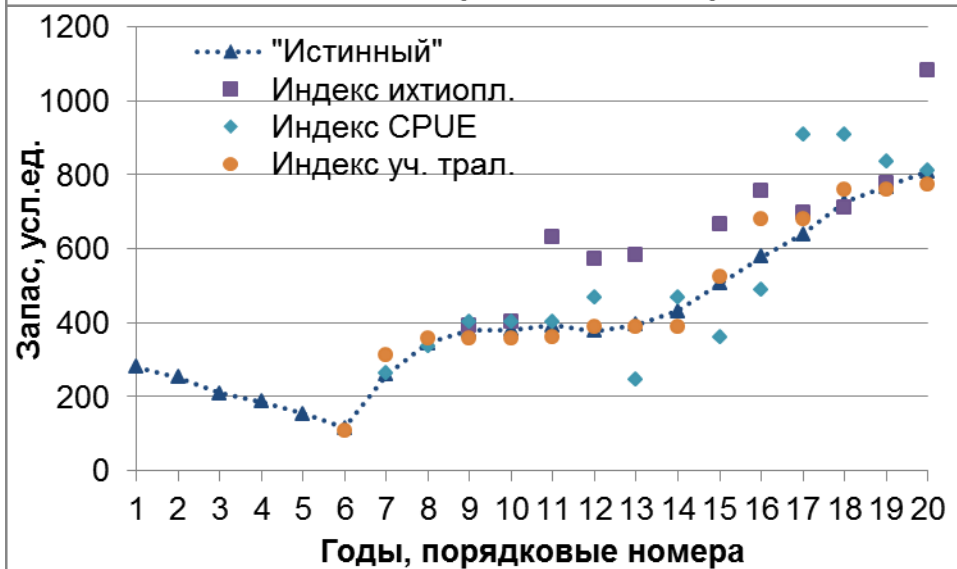
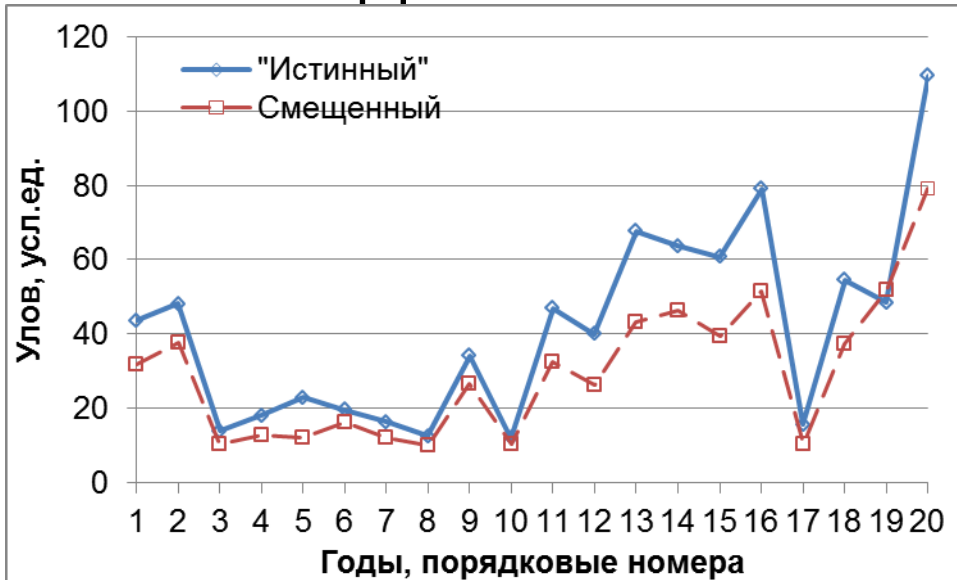


Результаты вычислительных экспериментов показали, что если использовать при настройке когортной модели данные об индексах запаса из независимых источников (траловые и ихтиопланктонные съемки, промышленные уловы на усилие), включающие только случайные ошибки, то ФК дает несмещенные оценки запаса.

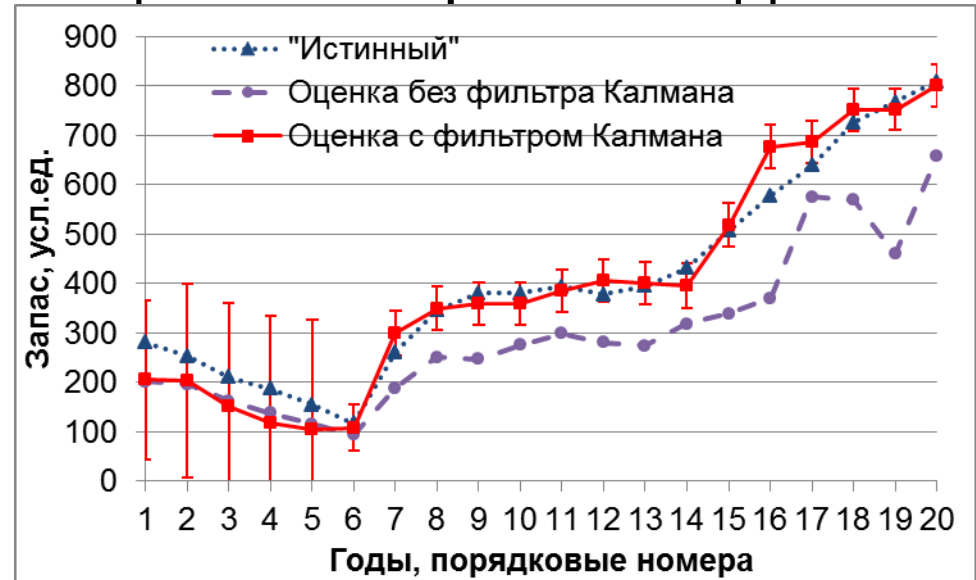
РЕЗУЛЬТАТЫ ВЫЧИСЛИТЕЛЬНОГО ЭКСПЕРИМЕНТА

Результаты работы ФК

Данные



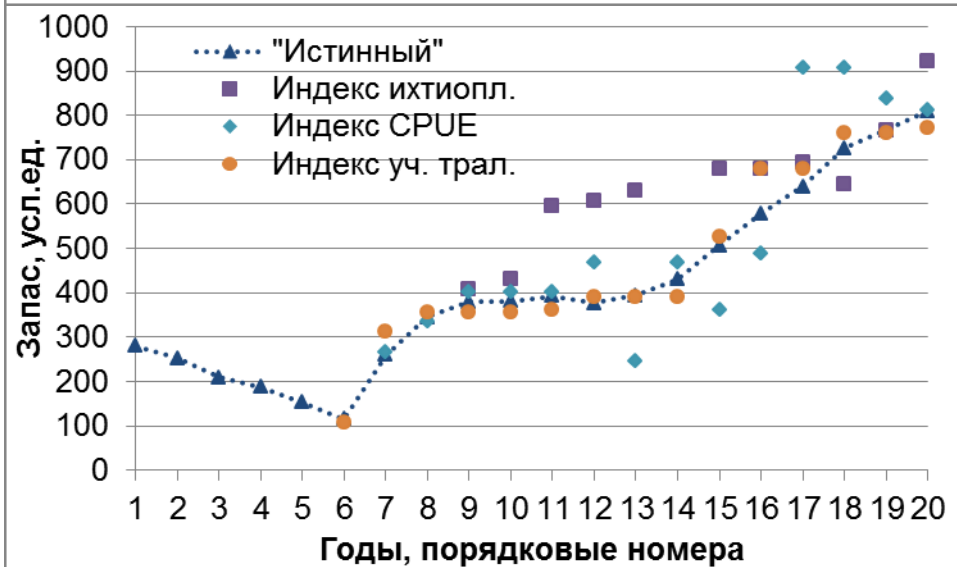
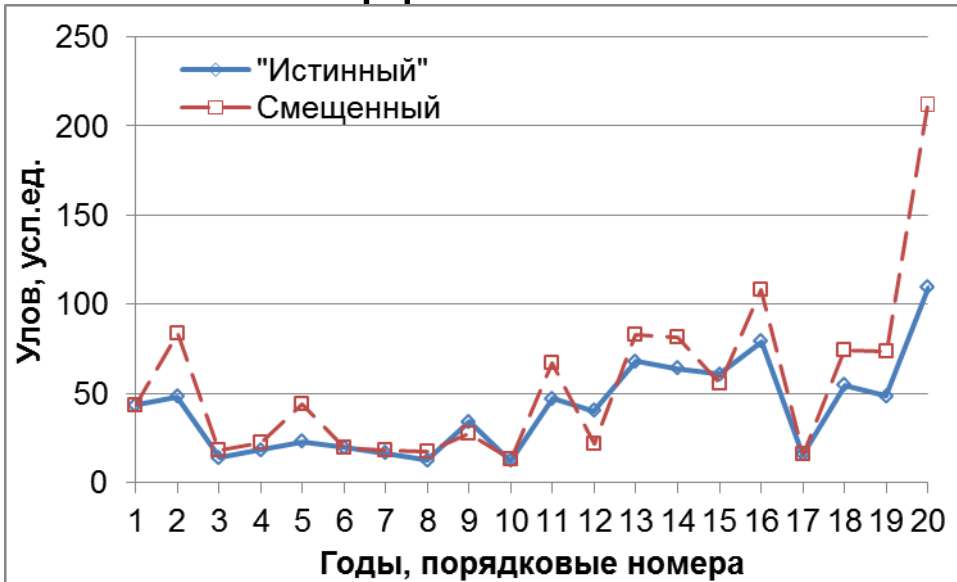
Оценка когортной модели



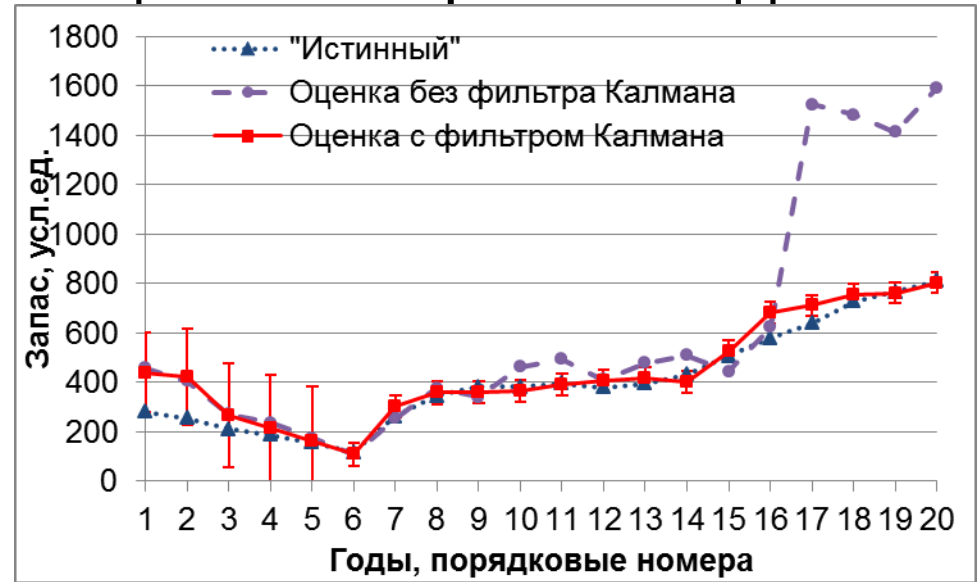
- Данные по уловам содержат помимо логнормального шума слабой интенсивности систематическое занижение.
- Наблюдения за годы №1-5 отсутствуют.

Результаты работы ФК

Данные



Оценка когортной модели



- Данные по уловам содержат только логнормальный шум высокой интенсивности.
- Наблюдения за годы №1-5 отсутствуют.

Взвешенный индекс запаса с коэффициентами усиления Калмана в качестве весов может служить эталоном для стандартизации CPUЕ.

ФК И СТАНДАРТИЗАЦИЯ ИНДЕКСА ЗАПАСА

Статистические веса для индексов запаса

Уравнение (11)

$$\sum_i K_{i,t} H_{i,t} + w_t \equiv 1$$

показывает, что $K_{i,t} H_{i,t}$ — это вес масштабированного к уровню запаса t -го наблюдения из i -го источника.

$y_{i,t} = \frac{I_{i,t}}{H_{i,t}}$ — индекс запаса, масштабированный к уровню запаса.

Уравнение коррекции :

$$\begin{aligned} N_{cor_t} &= w_t N_t^+ + (1 - w_t) y_t = w_t N_t^+ + \sum_i K_{i,t} H_{i,t} y_t \\ &= w_t N_t^+ + \sum_i K_{i,t} I_{i,t}, \end{aligned}$$

Стандартизированные индексы запаса

$$y_{w_t} = \sum_i w'_{i,t} y_{i,t}$$

$$w'_{i,t} = \frac{K_{i,t} H_{i,t}}{\sum_i K_{i,t} H_{i,t}}, \quad i = 1, 2, 3.$$

$$I_{w_t} = \sum_i w''_{i,t} I_{i,t}$$

$$w''_{i,t} = \frac{K_{i,t}}{\sum_i K_{i,t}}, \quad i = 1, 2, 3.$$

Стандартизированный индекс запаса

