

УДК 576.8: 576.895.122

**ОПЫТ ИСПОЛЬЗОВАНИЯ СИНГУЛЯРНОГО  
СПЕКТРАЛЬНОГО АНАЛИЗА В ПАРАЗИТОЛОГИИ:  
ДИНАМИКА ЗАРАЖЕННОСТИ МОЛЛЮСКОВ  
HYDROBIA VENTROSA ПАРТЕНИТАМИ ТРЕМАТОД  
CRYPTOCOTYLE CONCAVUM И BUNOCOTYLE PROGENETICA  
НА БЕЛОМ МОРЕ**

© И. А. Левакин,<sup>1</sup> К. Е. Николаев, К. В. Галактионов

Зоологический институт РАН,  
Университетская наб., 1, С.-Петербург, 199034  
Санкт-Петербургский государственный университет,  
Университетская наб., 7/9, С.-Петербург, 199034

<sup>1</sup> E-mail: levakin2@gmail.com

Поступила 15.01.2013

В статье сравниваются результаты анализа паразитологических данных разными методами выявления структуры временных рядов — автокорреляционный анализ, спектральный анализ (Фурье) и сингулярный спектральный анализ (ССА). В качестве примера использованы материалы 10-летнего мониторинга за изменением зараженности литоральных моллюсков Белого моря *Hydrobia ventrosa* партенитами трематод *Cryptocotyle concavum* и *Bunocotyle progenetica*. Наилучшие результаты показало применение ССА, в то время как автокорреляционный анализ не выявил каких-либо достоверных закономерностей.

Многолетние данные по зараженности хозяев паразитами нередко носят сложный для анализа характер. Свойственное паразитам агрегированное распределение обуславливает высокую дисперсию временного ряда. Кроме того, агрегированное распределение может приводить к значительным колебаниям амплитуды периодических составляющих ряда. Влияние иррегулярных факторов, таких как, например, погодные условия, может вызывать варьирование не только амплитудных, но и частотных характеристик периодических составляющих (варьирование периода колебаний). Кроме того, сбор материала через фиксированные временные отрезки может оказаться невыполнимым условием. Наконец, длина анализируемых временных рядов обычно относительно невелика и редко превышает 30—40 значений.

Выделение достоверного тренда для подобных рядов при помощи регрессионного анализа проблематично, поскольку подбор адекватной моде-

ли тренда для коротких рядов с высокой дисперсией может быть невыполнимой задачей. В то же время знание модели тренда — необходимое условие при использовании метода наименьших квадратов. С другой стороны, выделение периодических составляющих для таких рядов при помощи спектрального (дискретное преобразование Фурье) либо автокорреляционного анализа, также проблематично вследствие варьирования периода и амплитуды циклической составляющей, малой длины ряда, а также возможной неравномерности временных отрезков между наблюдениями.

Использование метода сингулярного спектрального анализа (ССА) (Broomhead, King, 1986) предоставляет возможность одновременного выделения тренда и комплекса (квази)периодических составляющих из столь сложных для анализа иными методами временных рядов. Кроме того, ССА позволяет определить вклад каждой составляющей в дисперсию анализируемого ряда. ССА не требует стационарности ряда, знания модели тренда, сведений о наличии и характере периодических составляющих и пригоден для анализа коротких рядов. При этом индексы значений анализируемого ряда могут быть интерпретированы не только как дискретные моменты времени, но и как некоторые метки, имеющие линейно-упорядоченную структуру.

Вышеописанные свойства ССА привели к его широкому использованию в климатологии, геофизике, метеорологии и эконометрике. Подробное описание метода ССА содержится в монографии Н. Э. Голяндиной и др. (Golyandina et al., 2001). ССА по сути является адаптацией метода главных компонент для анализа одномерных данных — временного ряда. Математической основой обоих методов (ССА и компонентного анализа) служит сингулярное разложение прямоугольной матрицы. По-видимому, впервые в биологии идеи, лежащие в основе ССА, были использованы для анализа временных рядов численности планктона Колбруком (Colebrook, 1978) — на 8 лет раньше, чем ССА был предложен для анализа сигналов в физике Брумхедом и Кингом (Broomhead, King, 1986). К сожалению, в биологии этот подход не получил широкого распространения. К настоящему времени опубликовано менее 10 работ, использующих те или иные модификации этого метода (Colebrook, 1978; Ibanez, Dauvin, 1988; Наумов, 2006, 2007; Наумов и др., 2009; Халаман, Наумов, 2009; Levakin et al., 2012).

В настоящей статье мы предприняли попытку сравнить результаты анализа реальных паразитологических данных с использованием разных методов выявления структуры временных рядов — автокорреляционного анализа, спектрального анализа (СА), основанного на дискретном преобразовании Фурье, и ССА.

#### КРАТКОЕ ОПИСАНИЕ ИСПОЛЬЗОВАННЫХ МЕТОДОВ

При выполнении автокорреляционного анализа рассчитывают коэффициенты автокорреляции при различных сдвигах (лагах) и оценивают их достоверность. Максимальные достоверные коэффициенты при этом соответствуют периодам колебаний, содержащихся в анализируемом ряде (Лакин, 1990).

При проведении СА временной ряд  $x$  длины  $N$  представляют в виде суммы  $N/2$  гармонических колебаний:  $x_i = \sum_{j=1}^{N/2} (a_j \cos(\lambda_j i) + b_j \sin(\lambda_j i))$ , где  $\lambda_j = 2\pi j/N$ , а  $a_j$  и  $b_j$  — коэффициенты пропорциональности. Вклад каждой гармоники ( $w_j$ ) в дисперсию анализируемого ряда рассчитывают как сумму квадратов коэффициентов ( $w_j = a_j^2 + b_j^2$ ). Поскольку набор гармоник определяется только длиной исходного ряда, результатом СА служит набор коэффициентов ( $a_j$  и  $b_j$ ,  $1 \leq j \leq N/2$ ). Подобным образом может быть разложен только стационарный ряд, поэтому анализируемый ряд приводят предварительно к стационарному виду (Bloomfield, 2000).

ССА состоит из 4 шагов: вложение, сингулярное разложение, группировка и диагональное усреднение (Elsner, Tsonis, 1996; Golyandina et al., 2001; Голяндина, 2004).

**Вложение.** При проведении ССА исходного ряда  $x$  — столбца из  $N$  наблюдений  $(x_1, \dots, x_n, \dots, x_N)$  исследователь выбирает лаг (длину окна)  $L$  ( $1 < L < N$ ) и преобразует ряд (вектор)  $x$  в матрицу  $Y_{L \times K}$  ( $K = N - L + 1$ ). Матрица  $Y$  составлена из столбцов-фрагментов исходного ряда  $x$  длиной  $L$ , каждый фрагмент начинается с элемента ряда  $x$ , номер которого соответствует номеру столбца матрицы  $Y$  ( $y_{i,j} = x_{i+j-1}$ ). В настоящей работе при длине временного ряда  $N = 45$  (ниже), исходя из существующих рекомендаций ( $L \approx N/2$  — Golyandina et al., 2001), мы выбрали лаг  $L = 22$ .

**Сингулярное разложение.** Полученную матрицу  $Y$  подвергают сингулярному разложению ( $Y = U \times \Sigma \times V^T$ , где  $\Sigma$  — диагональная матрица, содержащая сингулярные числа матрицы  $Y$  в порядке убывания,  $U$  и  $V$  — матрицы соответствующих этим числам левых и правых сингулярных векторов, а  $T$  — символ транспонирования). Соответствующие друг другу сингулярные вектора ( $u_i$  и  $v_i$ ) и значения ( $\sigma_i$ ) называют  $i$ -той собственной тройкой матрицы  $Y$ .

**Группировка.** Собственные тройки  $i$  и  $j$  группируют, исходя из близости квадратов сингулярных чисел ( $\sigma_i^2 \approx \sigma_j^2$ ), сходства формы левых сингулярных векторов ( $u_i$  и  $u_j$  — рис. 1, А) и характера скаттер-диаграмм пар левых сингулярных векторов. Скаттер-диаграмма пары векторов образует замкнутые циклы вокруг центра координат, если эта пара векторов описывает один колебательный процесс (рис. 1, Б). По собственным тройкам строят матрицы ( $Y_i = \sigma_i \times u_i \times v_i^T$ ). Если тройки были сгруппированы (например,  $i$  и  $j$ ), то соответствующие им матрицы складывают ( $Y_k = Y_i + Y_j$ ). В результате из  $r$  собственных троек сингулярного разложения матрицы  $Y$  (где  $r$  — ее ранг) получают  $d$  матриц ( $d < r$ ), сумма которых равна исходной матрице ( $Y = Y_1 + \dots + Y_k + \dots + Y_d$ ).

**Диагональное усреднение.** По каждой из полученных матриц-слагаемых  $Y_k$  восстанавливают соответствующий ему временной ряд длины  $N$  —  $x_k$ , каждое значение которого получают, усредняя значения  $Y_k$  ( $y_{i,j}$ ) с равной суммой индексов ( $x_{i+j-1}$  равно среднему значению всех  $y_{i,j}$ , для которых  $i+j = \text{const}$ ) (эта операция по сути обратная вложению). Полученные ряды ( $x_k$ ), являющиеся независимыми составляющим исходного ряда, называют модами (их сумма равна исходному ряду:  $x = x_1 + \dots + x_k + \dots + x_d$ ).

Определение значимости вклада моды в формирование анализируемого ряда. Соотношение квадратов сингулярных чисел матриц  $Y_k$  принято интерпретировать как соотношение дисперсий мод. Если собственные тройки (например,  $i$  и  $j$ ) подвергали группировке, то как значение про-



Рис. 1. Пример сходства формы пары левых сингулярных векторов (ЛСВ) собственных троек, подлежащих группировке при проведении ССА (А) и соответствующая им скаттер-диаграмма (Б). В этом примере приведены третий и четвертый ЛСВ, полученные при анализе ряда динамики зараженности *H. ventrosa* партенитами *B. progeneretica* длиной  $N = 45$  с лагом  $L = 22$ .

Fig. 1. An example of a similarly-shaped pair of left singular vectors (LSV) of eigentriples subject to grouping during SSA (A) and the corresponding scatter plot (B). This example shows the third and the fourth LSV obtained during analysis of the series of dynamics of *B. progeneretica* parthenitae prevalence in *H. ventrosa* with a length of  $N = 45$  and a lag of  $L = 22$ .

порциональное вкладу моды в дисперсию исходного ряда рассматривают сумму квадратов сингулярных значений этих собственных троек ( $\sigma_i^2 + \sigma_j^2$  — Golyandina et al., 2001). Этот «классический» способ оценки вклада моды в дисперсию временного ряда представляется нам не вполне корректным (Levakin et al., 2012). Поэтому мы определяли вклад моды  $x_k$  в дисперсию временного ряда как величину, пропорциональную сумме квадратов отклонений элементов матрицы  $Y_k$  от среднего значения ( $SS_k = \sum_{i=1}^L \sum_{j=1}^K (\dot{y}_{i,j} - \tilde{y})^2$ , где  $\tilde{y} = (LK)^{-1} \sum_{i=1}^L \sum_{j=1}^K y_{i,j}$  — среднее значение элементов этой матрицы — Levakin et al., 2012). После этого значения  $SS_k$  нормировали ( $SS_k^* = SS_k / \sum_{i=1}^d SS_i$ ) и упорядочивали по убыванию значений. При определении значимости моды по упорядоченным значениям  $SS^*$  руководствовались стандартным при проведении компонентного анализа правилом «сломанной трости» («broken stick» rule — Jolliffe, 2002), полагая критерий Гутмана—Кайзера (Guttman, 1954; Kaiser, 1960) излишне либеральным.

## МАТЕРИАЛ И МЕТОДИКА

Материалом послужили данные многолетнего сезонного мониторинга зараженности литоральных моллюсков *Hydrobia ventrosa* партенитами трематод, который выполнялся в период с 1 августа 1999 г. по 30 августа 2011 г. в Сухой салме губы Чупа (Кандалакшский залив Белого моря) ( $66^{\circ}18'43''$  с. ш.,  $33^{\circ}40'26''$  в. д.) (рис. 2). Моллюсков собирали во время отлива при помощи стандартной гидробиологической рамки площадью  $0.025 \text{ м}^2$  с 5 точек (по 6 рамок с каждой) выбранного участка илистой литорали. Пробы промывали через сито с ячейей  $1 \text{ мм}^2$  и отбирали всех обнаруженных гидробий. Этих моллюсков затем вскрывали под стереомикроскопом, причем регистрировали наличие или отсутствие заражения и опреде-

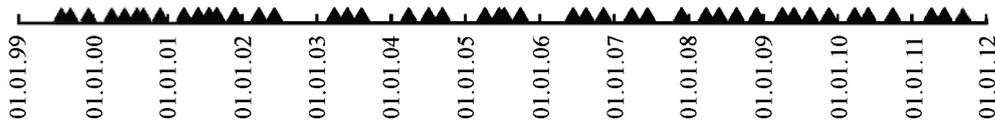


Рис. 2. Время сбора проб моллюсков *Hydrobia ventrosa* в Сухой салме (отмечено треугольниками).

Fig. 2. Sampling time of *Hydrobia ventrosa* snails in Sukhaya Salma (triangles).

ляли вид паразита. Для целей данного исследования объединяли пробы, собранные в один день со всех 5 точек полигона. Объем каждой такой объединенной пробы составлял 37—3199 особей *Hydrobia ventrosa*.

Анализу (автокорреляционному, СА и ССА) подвергли временные ряды зараженности *H. ventrosa* партенитами, наиболее часто встречающимися в районе сбора материала видов трематод — *Cryptocotyle concavum* и *Bunocotyle progenetica*. Мариты первого вида паразитируют в птицах (преимущественно в чайках), связанных трофически с морским побережьем. Роль второго промежуточного хозяина для *C. concavum* играют многие виды прибрежных рыб. Жизненный цикл *B. progenetica* моноксенный, в редях, паразитирующих в гидробиях, проходит весь морфогенез особей гермафродитного поколения вплоть до зрелых, продуцирующих яйца, марит. Дисперсия яиц этого вида осуществляется после гибели моллюска-хозяина.

При проведении спектрального анализа (Фурье) нагрузки периодограммы ( $w$ ) выражали в процентах. Рассчитанные по материалам мониторинга значения экстенсивности инвазии гидробий партенитами *C. concavum* и *B. progenetica*, расположенные в порядке взятия проб, формировали анализируемые временные ряды зараженности. Доверительные интервалы долей рассчитывали по точной формуле Фишера (Животовский, 1991; Sokal, Rohlf, 1995). Все доверительные интервалы приведены для 95%-ного уровня значимости.

Для сравнения результатов ССА и спектрального анализа сопоставляли значения коэффициентов корреляции сумм выделенных этими методами компонент временных рядов с исходными временными рядами.

Поскольку сингулярный спектральный анализ отсутствовал в доступных статистических пакетах, он был самостоятельно реализован И. А. Левакиным в виде VBA (Visual Basic for Application) макроса MS Excel. При проведении автокорреляционного и спектрального анализа использовали пакет Statistica 7.0.

## РЕЗУЛЬТАТЫ

Автокорреляционный анализ динамики зараженности *H. ventrosa* партенитами *C. concavum* и *B. progenetica* не выявил каких-либо достоверных автокорреляций. Вместе с тем была обнаружена слабая кросскорреляция ( $r = -0.39$ ,  $P < 0.05$ , лаг — 4) между этими рядами зараженности при сдвиге первого ряда относительно второго на 4 позиции.

ССА разложил ряды зараженности *H. ventrosa* партенитами *C. concavum* и *B. progenetica* на 16 мод (рис. 3). Динамика зараженности партенита-

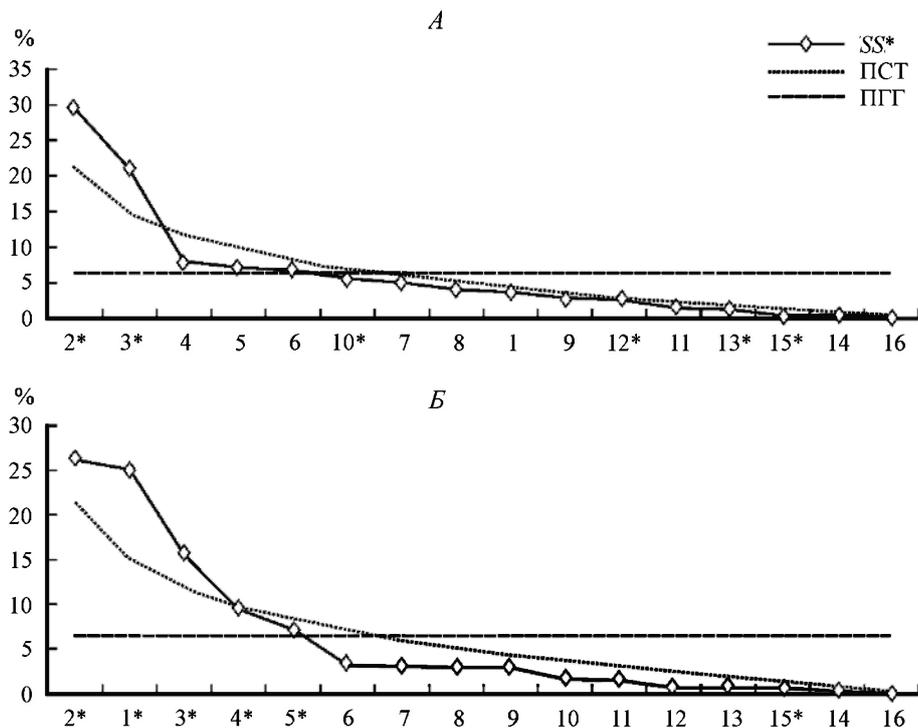


Рис. 3. Вклад мод ССА в дисперсию исходных рядов ( $SS^*$ ) динамики зараженности *H. ventrosa* партенитами *C. concavum* (А) и *B. progenetica* (Б).

Пунктиром обозначены критические значения согласно правилу «сломанной трости» (ПСТ) и правилу Гутмана—Кайзера (ПГК). Звездочкой отмечены моды, полученные группировкой собственных троек. Первая мода соответствует тренду временного ряда. По оси абсцисс — номера мод ССА, по оси ординат — доля объясненной дисперсии.

Fig. 3. Contribution of SSA modes to the dispersion of initial series ( $SS^*$ ) of the dynamics of *C. concavum* (A) and *B. progenetica* (B) parthenitae prevalence in *H. ventrosa*.

ми *C. concavum* содержала только 2 составляющие, значимые по правилу «сломанной трости». Тренд изменения зараженности *H. ventrosa* партенитами *C. concavum* не был выражен (рис. 3, А; 4, А). Динамика зараженности *H. ventrosa* партенитами *B. progenetica* содержала 3 составляющие, значимые по правилу «сломанной трости» (рис. 3, Б), включая тренд (рис. 4, А—В).

При проведении спектрального анализа (СА) временных рядов зараженности *H. ventrosa* партенитами *C. concavum* и *B. progenetica*, первые (нулевые для обоих рядов) значения были отброшены для достижения четного числа наблюдений (44). Периодограммы рядов зараженности гидробий партенитами двух анализируемых видов отличались (рис. 5). Для ряда зараженности *H. ventrosa* партенитами *C. concavum* наиболее характерны были 2 колебания: первому из них соответствовали гармоники с периодом в 2.4 и 2.6 наблюдений ( $w_{2.4} + w_{2.6} = 24.2\%$ ), а второму — гармоника с периодом колебаний в 8.8 наблюдений ( $w_{8.8} = 23.5\%$  — рис. 6, А). Присутствовало также колебание, которому соответствовали гармоники с периодом в 2.9 и 3.1 наблюдения ( $w_{2.9} + w_{3.1} = 10.6\%$  — рис. 6, В). Для ряда зараженности *H. ventrosa* партенитами *B. progenetica* наиболее характерно было

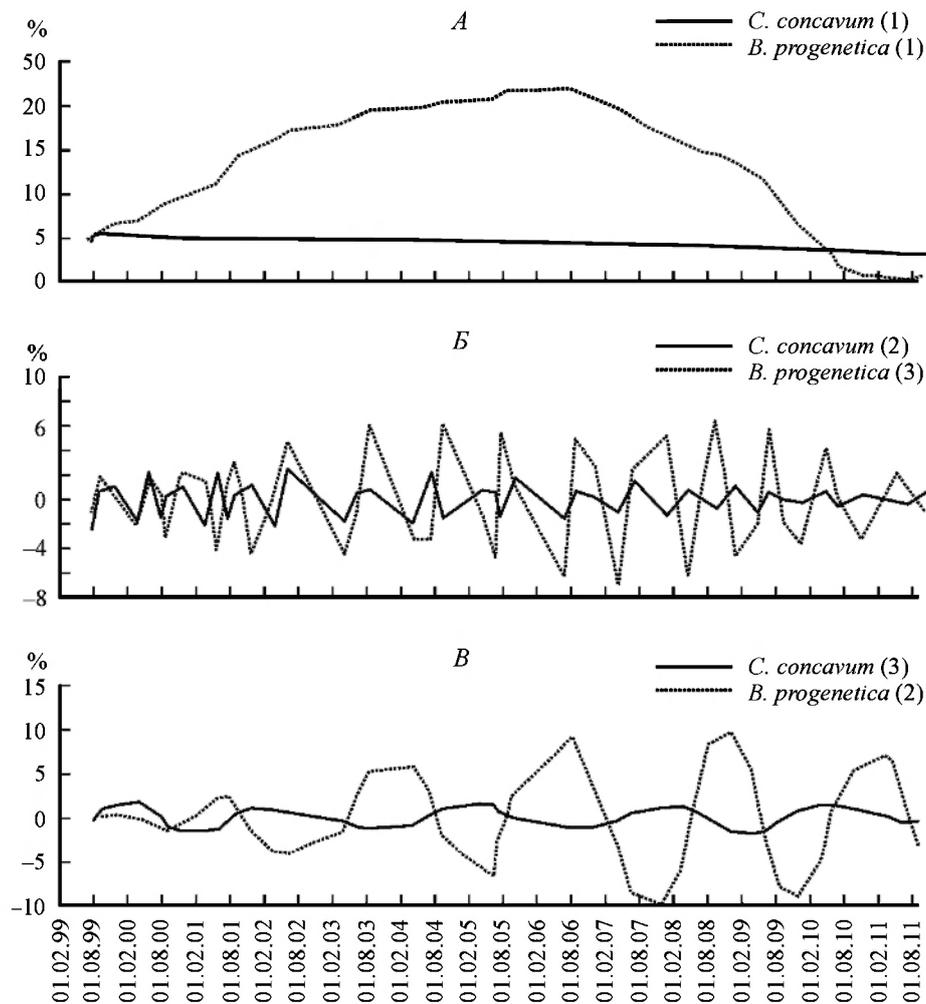


Рис. 4. Моды ССА динамики зараженности *H. ventrosa* партенитами *C. concavum* и *B. progenetica*, значимые по правилу «сломанной трости».

В скобках приведен номер моды ССА (пояснения в тексте). По оси абсцисс — время, по оси ординат — зараженность (А) и изменения зараженности (Б, В).

Fig. 4. SSA modes of dynamics of *C. concavum* and *B. progenetica* parthenitae prevalence in *H. ventrosa* snails significant according to the «broken stick» rule.

колебание с периодом в 44 наблюдения ( $w = 36.4\%$ ) и колебание, которому соответствовали гармоники с периодом в 7.3 и 8.8 ( $w_{7.3} + w_{8.8} = 21.7\%$ ) наблюдений (рис. 6, Б). Присутствовало также колебание с периодом в 3.1 наблюдения ( $w_{3.1} = 8.8\%$  — рис. 6, В).

«Низкочастотные» составляющие ряда зараженности *H. ventrosa* партенитами *C. concavum*, выделенные ССА (рис. 4, В) и СА (рис. 6, А), коррелировали между собой ( $r = 0.87$ ,  $P \ll 0.01$ ). То же справедливо и для «низкочастотных» составляющих ряда зараженности *H. ventrosa* партенитами *B. progenetica* ( $r = 0.68$ ,  $P \ll 0.01$ ). При этом «низкочастотные» моды ССА рядов зараженности *H. ventrosa* партенитами *C. concavum* и *B. progenetica* также коррелировали между собой ( $r = -0.52$ ,  $P < 0.01$ ).

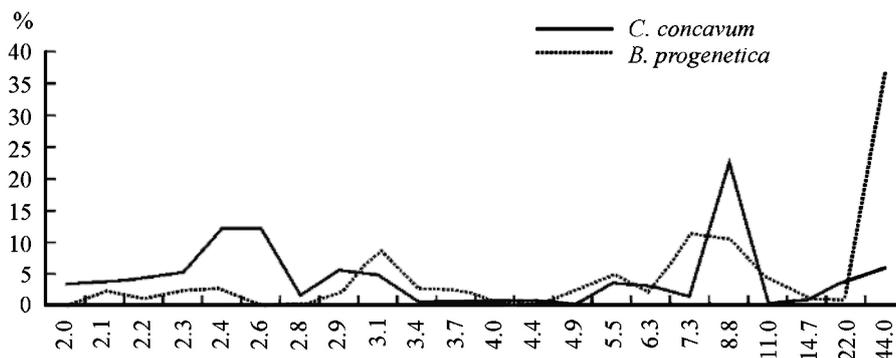


Рис. 5. Периодограммы спектрального анализа (Фурье) динамики зараженности *H. ventrosa* партенитами двух видов трематод.

По оси абсцисс — периоды гармоник, по оси ординат — нагрузка ( $w$ ), выраженная в %.

Fig. 5. Periodogram of Fourier spectrum analysis of the dynamics of prevalence of *C. concavum* and *B. progenetica* parthenitae in *H. ventrosa* snails.

тотные» гармоники, выделяемые СА для рядов зараженности *H. ventrosa* партенитами *C. concavum* и *B. progenetica*, не коррелировали ( $r = -0.07$ ,  $P > 0.05$ ).

«Высокочастотные» составляющие ряда зараженности *H. ventrosa* партенитами *C. concavum*, выделенные ССА (рис. 4, Б) и СА (рис. 6, А), коррелировали между собой ( $r = 0.91$ ,  $P \ll 0.01$ ). «Высокочастотные» составляющие ряда зараженности *H. ventrosa* партенитами *B. progenetica*, выделенные ССА (рис. 4, Б) и СА (рис. 6, В), также коррелировали между собой ( $r = 0.78$ ,  $P \ll 0.01$ ). «Высокочастотные» составляющие рядов зараженности *H. ventrosa* партенитами *C. concavum* и *B. progenetica* не коррелировали вне зависимости от вида анализа.

Тренд ССА (рис. 4, А) и гармоника СА с периодом 44 (рис. 6, Б) практически совпадали по форме для ряда зараженности *H. ventrosa* партенитами *B. progenetica* ( $r = 0.97$ ,  $P \ll 0.01$ ).

Сумма первых трех мод ССА (тренд и две значимые по правилу «сломанной трости» моды) временного ряда зараженности *H. ventrosa* партенитами *C. concavum* достоверно ( $p \ll 0.01$ ) коррелировала ( $r = 0.73$ ,  $r_s = 0.7$ ) с исходным рядом (рис. 7, А). Сумма пяти наиболее выраженных гармоник этого ряда, выделенных СА (с периодами в 2.4, 2.6, 8.8, 2.9 и 3.1 наблюдений — рис. 5), также достоверно ( $P < 0.01$ ) коррелировала ( $r = 0.57$ ,  $r_s = 0.52$ ) с исходным рядом (рис. 7, А). Сумма всех гармоник СА достоверно ( $P \ll 0.01$ ) коррелировала с исходным рядом ( $r = 0.84$ ,  $r_s = 0.83$ ), но несколько отличалась от него по форме.

Сумма первых трех мод ССА (значимых по правилу «сломанной трости») временного ряда зараженности *H. ventrosa* партенитами *B. progenetica* достоверно ( $P \ll 0.01$ ) коррелировала ( $r = 0.83$ ,  $r_s = 0.77$ ) с исходным рядом (рис. 7, Б). Сумма четырех наиболее выраженных гармоник этого ряда, выделенных СА (с периодами в 44, 7.3, 8.8 и 3.1 наблюдения — рис. 5), также достоверно ( $P \ll 0.01$ ) коррелировала ( $r = 0.7$ ,  $r_s = 0.68$ ) с исходным рядом (рис. 7, Б). Сумма всех гармоник СА достоверно ( $P \ll 0.01$ ) коррелировала с исходным рядом ( $r = 0.74$ ,  $r_s = 0.66$ ), но несколько отличалась от него по форме.

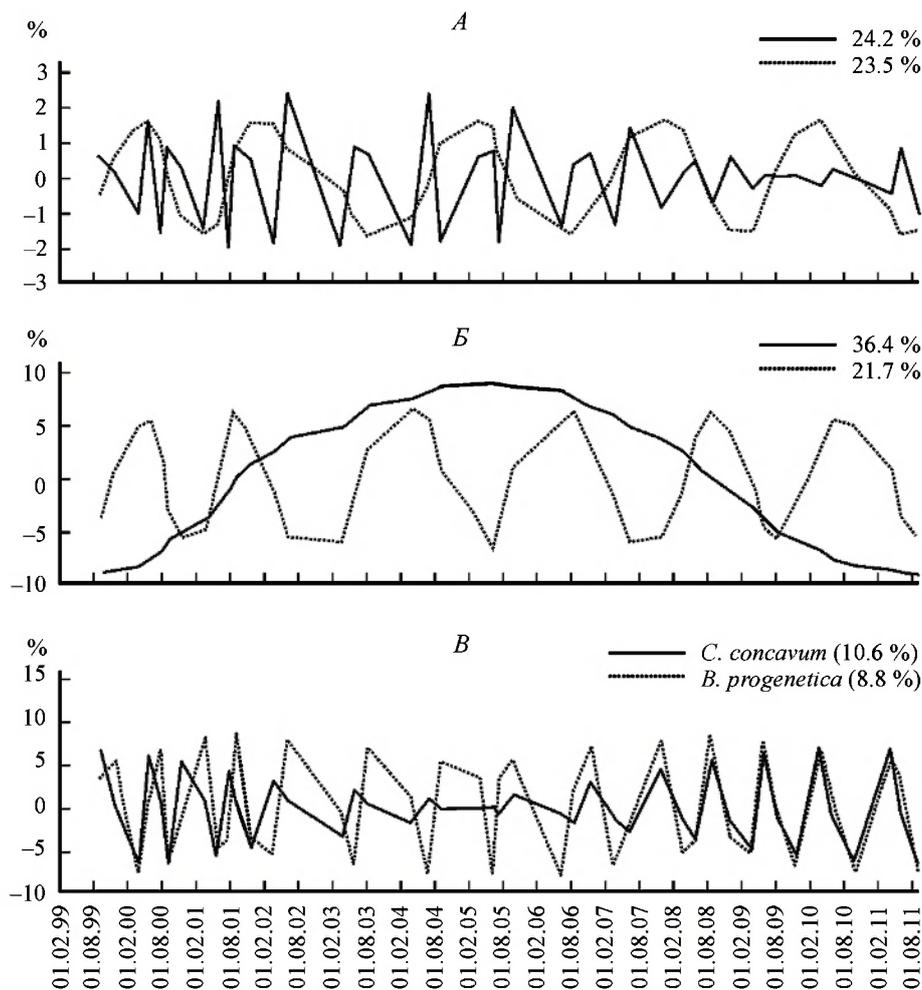


Рис. 6. Гармоники, выделенные спектральным анализом (Фурье) из рядов зараженности *H. ventrosa* партенитами *C. concavum* (А) и *B. progenetica* (Б), которым соответствовали наибольшие значения нагрузок (пояснения в тексте).

По оси абсцисс — время, по оси ординат — изменения зараженности.

Fig. 6. Harmonics, extracted with the help of Fourier spectrum analysis from the series of *C. concavum* (А) and *B. progenetica* (Б) parthenitae prevalence in *H. ventrosa*, to which the maximum values of the load corresponded.

## ОБСУЖДЕНИЕ

Автокорреляционный анализ оказался бесполезен при выявлении структуры временных рядов зараженности моллюсков *H. ventrosa* партенитами трематод. В то же время применение сингулярного спектрального анализа и спектрального анализа, основанного на дискретном преобразовании Фурье, позволило выявить в каждом временном ряду несколько независимых составляющих (рис. 4; 6). Важно заметить, что, несмотря на различия математического аппарата (сингулярное разложение матриц и дискретное преобразование Фурье), а также предпосылок, лежащих в

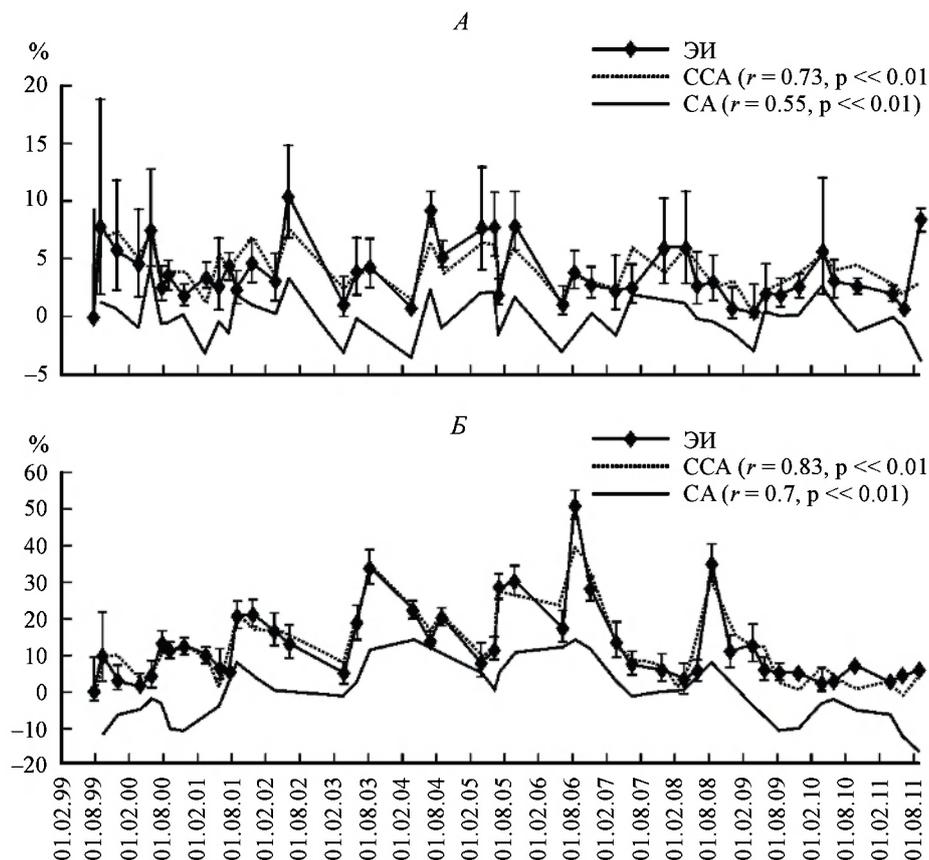


Рис. 7. Исходные ряды экстенсивности инвазии (ЭИ) *H. ventrosa* партенитами *C. concavum* (А) и *B. progenetica* (Б) и их реконструкции на основе значимых по правилу «сломанной трости» компонент сингулярного спектрального анализа (ССА) и наиболее значимых гармоник, выделяемых спектральным анализом (СА).

Доверительные интервалы ЭИ приведены для 95%-го уровня значимости. Для реконструированных рядов (ССА и СА) указаны коэффициенты корреляции с исходными рядами (пояснения в тексте). По оси абсцисс — время, по оси ординат — изменения зараженности.

Fig. 7. Initial series of prevalence of infection (PI) of *C. concavum* (A) and *B. progenetica* (B) in *H. ventrosa* snails and their reconstruction on the basis of SSA components signification according to the «broken stick» rule and on the basis of the most significant harmonics extracted by SSA.

основе этих двух методов, их использование привело к сходным результатам — выделило сходные составляющие. Независимое выделение сходных составляющих при анализе временных рядов разными методами свидетельствует в пользу реальности существования этих составляющих.

Несложно заметить, что составляющие, выделенные ССА, лучше описывают анализируемые ряды, чем составляющие, выделенные СА (рис. 7). В то время как большинство значений реконструированного по модам ССА ряда находится в пределах доверительных интервалов исходных значений зараженности, реконструированный по гармоникам СА ряд смещен относительно исходного ряда (рис. 7). Это закономерный результат, поскольку сумма всех мод ССА по определению совпадает с исходным рядом, в то время как перед проведением СА из анализируемого ряда вычи-

тают среднее значение, смещая ряд (в нашем случае вниз — рис. 7). Значительно более важным представляется то, что реконструированный по модам ССА ряд более скоррелирован с исходным рядом, чем ряд, реконструированный по гармоникам СА. Иными словами, моды ССА лучше отражают форму исходного ряда, чем гармоники СА. Этот результат также представляется закономерным по ряду причин.

Во-первых, перед проведением СА временной ряд приводится к стационарному виду и при выполнении этой операции искажение формы исходного ряда практически неизбежно. В частности, перед проведением СА из временного ряда удаляют тренд, в то время как корректное выделение тренда представляет собой достаточно сложную задачу. Однако это можно считать технической сложностью, решение которой сводится к аккуратному выделению тренда. Вторая причина носит фундаментальный характер: СА представляет любое колебание в виде суммы гармонических колебаний вне зависимости от характера исходного колебания. Вместе с тем колебательные процессы в биологических системах редко носят характер правильных гармоник. В отличие от СА ССА способен отразить особенности колебаний в выделяемых модах. Например, с ростом зараженности *H. ventrosa* партенитами *B. progenetica* увеличивается и амплитуда колебаний зараженности (рис. 7, Б). Выделяемые СА гармоники не отражают изменения амплитуды колебаний со временем (рис. 6, А, В), в то время как выделенные ССА моды отражают эти изменения (рис. 4, Б, В). Следует отметить, что гармоники СА являются абстрактными величинами, изменения которых отражают изменения анализируемого ряда. В то же время моды ССА выражаются в тех же единицах, что и исходный ряд, что облегчает интерпретацию результатов анализа.

При интерпретации полученных результатов необходимо учитывать то обстоятельство, что составляющие временных рядов выделяются использованными методами (ССА и СА) формально — по соответствию некоторым критериям, заложенным в математический аппарат метода (например, лежащие в основе мод ССА собственные тройки должны соответствовать условиям сингулярного разложения). Вследствие этого одному биологическому процессу может соответствовать несколько составляющих временного ряда, выделяемых ССА или СА. С другой стороны, оба этих метода не могут гарантировать, что выделенная ими составляющая временного ряда не будет являться результатом взаимодействия нескольких независимых процессов. Последнее замечание может быть особенно актуально при интерпретации результатов анализа коротких рядов, поскольку разрешающая способность ССА и СА в этом случае снижается. В случае ССА речь идет о снижении асимптотической делимости составляющих, описываемых различными закономерностями (подробнее см.: Golyandina et al., 2001). В случае СА существующая периодичность может распределиться между соседними гармониками, что приведет не только к сглаживанию периодограммы, но и снизит точность вычисления коэффициентов, являющихся результатом СА (по этой причине обычно не рекомендуют использовать СА для анализа коротких рядов). Несмотря на то что период наших наблюдений охватывает 10 лет, анализируемые ряды зараженности *H. ventrosa* партенитами трематод состоят всего из 45 значений, и замечание относительно снижения разрешающей способности ССА и СА при анализе

коротких рядов относится к ним в полной мере. Так, вследствие малой длины ряда половина выделяемых СА колебаний оказывается «размазана» по соседним гармоникам (рис. 5).

В отличие от СА, не позволяющего судить о трендовых составляющих временного ряда, ССА позволяет утверждать, что зараженность *H. ventrosa* партенитами *C. concavum* не претерпевала значимых долговременных изменений (рис. 4, А), обнаруживая лишь некоторую тенденцию к снижению (вклад тренда в дисперсию временного ряда не превышал 3.8 % — рис. 3, А). Следует отметить, что этот вывод нельзя сделать, рассматривая изменения среднегодовой зараженности гидробий этим паразитом, поскольку она значительно варьировала в период наблюдений от 1.9 (2009 г.) до 6.8 % (2002 г.) (рис. 7, А). В то же время он вполне соответствует реальной ситуации, которая на протяжении периода исследований была достаточно стабильной для трансмиссии *C. concavum* (Levakin et al., 2012). Зараженность *H. ventrosa* партенитами *B. progenetica* значимо изменялась в период проведения наблюдений (рис. 4, А), объясняя почти четверть дисперсии временного ряда (рис. 3, Б). Любопытно, что куполообразная форма тренда зараженности *H. ventrosa* партенитами *B. progenetica* позволила СА интерпретировать его как гармоническое колебание, чей период совпадает с периодом проведения наблюдений (рис. 6, Б). На настоящем этапе исследования мы не в состоянии однозначно трактовать выявленный тренд в динамике зараженности гидробий видом *B. progenetica*. Можно только отметить, что он напоминает характер протекания острой эпидемии вирусной или бактериальной природы (Swinton et al., 2002), а также динамику зараженности патогенным макропаразитом популяций животных с циклической динамикой численности (Tomkins et al., 2002). Учитывая, что *B. progenetica* — это летальный для гидробий паразит с моноксенным жизненным циклом (Левакин, 2004, 2005, 2008), такое совпадение вряд ли может быть случайным.

Изолировать сезонную составляющую временных рядов зараженности *H. ventrosa* партенитами трематод от случайных изменений зараженности (шума) не представляется возможным, поскольку длина ряда мала ( $N = 45$ ), а схема взятия проб (рис. 2) скорее соответствует временному ряду с пропусками, чем временному ряду с отклонениями от регулярного проведения наблюдений. Интервалы между наблюдениями значительно варьируют (от 25 до 293 дней), и в разные годы собиралось разное (от 2 до 5) количество проб (рис. 2). Вероятнее всего, сезонная составляющая (смешанная с шумом) содержится в «высокочастотных» составляющих временных рядов зараженности гидробий партенитами трематод (рис. 4, Б; 6, А, В). В пользу наличия регулярной компоненты в этой составляющей говорит чрезвычайное сходство высокочастотных составляющих, выделяемых разными методами анализа — шумы в разных видах анализа (ССА и СА) комбинировались бы в разные по форме составляющие.

«Низкочастотные» составляющие временных рядов зараженности, по-видимому, значительно лучше изолированы от шума (рис. 4, В; 6, А, Б). Возможно, эти колебания зараженности могут быть связаны с популяционной динамикой хозяина. Косвенно в пользу такого предположения может говорить совпадение периода этих колебаний с продолжительностью жизни хозяина (2—3 года, отдельные особи доживают до 4 лет — Горбу-

шин, 2000). Кроме того, максимумы этих составляющих для *B. progenetica* и *C. concavum* отстают друг от друга примерно на год (рис. 4, В; 6, А, Б). Подобное «запаздывание» роста зараженности *C. concavum* относительно *B. progenetica* хорошо согласуется с возрастной структурой зараженности гидробий партенитами этих трематод. Партениты *B. progenetica* заражают в первую очередь гидробий первого года жизни, а максимальная зараженность партенитами *C. concavum* наблюдается у гидробий второго года жизни. Это связано с тем, что инвазия партенитами *B. progenetica* в отличие от инвазии *C. concavum* значительно снижает жизнеспособность зараженного хозяина и соответственно резко уменьшает шансы зараженной особи дожить до следующего года (Левакин, 2004, 2005, 2008).

#### БЛАГОДАРНОСТИ

Работа выполнена при поддержке грантов РФФИ (№ 10-04-00430) и СПбГУ (№ 1.37.80.2011).

#### Список литературы

- Голяндина Н. Э. 2004. Метод «Гусеница»-SSA: анализ временных рядов. СПб.: Изд-во СПбГУ. 76 с.
- Горбушин А. М. 2000. Сравнительный морфофункциональный анализ взаимоотношений в системе моллюск—трематода. Паразитология. 34 (6): 502—511.
- Животовский Л. А. 1991. Популяционная биометрия. М.: Наука. 269 с.
- Лакин Г. Ф. 1990. Биометрия. М.: Высшая школа. 250 с.
- Левакин И. А. 2004. Влияние инвазии трематодами *Bunocotyle progenetica* (Hemiuroidea, Bunocotylinae) и *Cryptocotyle concavum* (Heterophyidae) на смертность морских литоральных моллюсков *Hydrobia ulvae* (Gastropoda, Prosobranchia) при воздействии экстремально высокой температуры. Паразитология. 38 (4): 352—358.
- Левакин И. А. 2005. Влияние инвазии трематодами *Bunocotyle progenetica* (Hemiuroidea) и *Cryptocotyle concavum* (Heterophyidae) на смертность морских литоральных моллюсков *Hydrobia ulvae* (Gastropoda: Prosobranchia) при замораживании. Паразитология. 39 (5): 407—413.
- Левакин И. А. 2008. Реализация моноксенного жизненного цикла *Bunocotyle progenetica* (Trematoda: Hemiuroidea, Bunocotylinae) в условиях Белого моря: Дис. ... канд. биол. наук. СПб.: Зоол. ин-т РАН. 171 с.
- Наумов А. Д. 2006. Двустворчатые моллюски Белого моря. Опыт эколого-фаунистического анализа. Исследования фауны морей 59 (67). СПб.: Зоол. ин-т РАН. 367 с.
- Наумов А. Д. 2007. Многолетние исследования литорального бентоса Белого моря в губе Чупа (Кандалакшский залив): сезонная и многолетняя динамика биомассы взморника *Zostera marina*. В кн.: Комплексные исследования процессов, характеристик и ресурсов российских морей Североевропейского бассейна. Апатиты: Изд-во Кольского научного центра РАН. 493—502.
- Наумов А. Д., Халаман В. В., Фокин М. В. 2009. Многолетняя динамика обилия некоторых литоральных полихет в двух небольших губах Кандалакшского залива (Белое море). Биология моря. 35 (5): 337—343.
- Халаман В. В., Наумов А. Д. 2009. Многолетняя динамика массовых видов полихет в сообществах обрастания Белого моря. Биология моря. 35 (6): 410—419.
- Broomhead D. S., King G. P. 1986. Extracting qualitative dynamics from experimental data. Physica D. 20: 217—236.
- Colebrook J. M. 1978. Continuous plankton records: zooplankton and environment, North-East Atlantic and North Sea, 1948—1975. Oceanologica Acta. 1: 9—23.

- Elsner J., Tsonis A. 1996. Singular Spectrum Analysis. A New Tool in Time Series Analysis. New York: Plenum Press. 177 p.
- Bloomfield P. 2000. Fourier Analysis of Time Series: An Introduction. New York: J. Wiley et Sons. 288 p.
- Golyandina N., Nekrutkin V., Zhigljavsky A. 2001. Analysis of Time Series Structure: SSA and Related Techniques. New York: Chapman et Hall/CRC. 306 p.
- Guttman L. 1954. Some necessary conditions for common factor analysis. *Psychometrika*. 19: 149—162.
- Ibanez F., Dauvin J.-C. 1988. Long-term changes (1977—1987) in a muddy fine sand *Abrar alba*—*Melinna palmate* community from Western English Channel: multivariate time-series analysis. *Mar. ecol. Prog. Ser.* 49: 65—81.
- Jolliffe I. T. 2002. Principal Component Analysis. Series: Springer Series in Statistics, 2nd ed, Vol. 29. New York: Springer. 518 p.
- Kaiser H. F. 1960. The application of electronic computer to factor analysis. *Educational and Psychological Measurement*. 20: 141—151.
- Levakin I.A., Nikolaev K.E., Galaktionov K.V. 2012. Long-term variation in trematode (Trematoda, Digenea) component communities associated with intertidal gastropods is linked to abundance of final hosts. *Hydrobiologia*. (doi:10.1007/s10750-012-1267-x).
- Sokal R. R., Rohlf F. J. 1995. Biometry, 3rd ed. New York: W. H. Freeman. 887 p.
- Swinton J., Woolhouse M. E. J., Begon M. E., Dobson A. P., Ferroglio E., Grenfell B. T., Guberti V., Hails R. S., Heesterbeek J. A. P., Lavazza A., Roberts M. G., White P. J., Wilson K. 2002. Microparasite transmission and persistence. In: *The Ecology of Wildlife Diseases* / Eds P.J. Hudson, A. Rizzoli, B. T. Grenfell, H. Heesterbeek, A. P. Dobson. New York: Oxford University Press. 83—101.
- Tompkins D. M., Dobson A. P., Arenberg P., Begon M. E., Cattadori I. M., Greenman J. V., Heesterbeek J. A. P., Hudson P. J., Newborn D., Pugliese A., Rizzoli A. P., Rosà R., Rosso F., Wilson K. 2002. Parasites and host population dynamics. In: *The Ecology of Wildlife Diseases* / Eds P. J. Hudson, A. Rizzoli, B. T. Grenfell, H. Heesterbeek, A. P. Dobson. New York: Oxford University Press. 45—62.

A CASE STUDY OF SINGULAR SPECTRUM ANALYSIS APPLICATION  
IN PARASITOLOGY: DYNAMICS OF PREVALENCE OF CRYPTOCOTYLE  
CONCAVUM AND BUNOCOTYLE PROGENETICA TREMATODE PARTHENITAE  
IN HYDROBIA VENTROSA SNAILS AT THE WHITE SEA

I. A. Levakin, N. E. Nikolaev, K. V. Galaktionov

*Key words:* singular spectrum analysis, auto-correlation analysis, Fourier spectrum analysis, parasitological monitoring, time series, trematodes, parthenitae, molluscan host, prevalence.

S U M M A R Y

In this study parasitological data were analyzed by different methods of revealing the structure of time series, namely. auto-correlation analysis (ACA), Fourier spectrum analysis (SA) and singular spectrum analysis (SSA), and the results of these analysis were compared (SSA makes it possible to present non-stationary time series as a sum of independent components and to determine the contribution of each component into the dispersion of the initial series — Golyandina et al., 2001). This case study was based on the result of 10-year-long monitoring of changes in the prevalence of *Cryptocotyle concavum* and *Bunocotyle progenetica* trematode parthenitae in intertidal snails *Hydrobia ventrosa* at the White Sea (in total, 45 observations). ACA did not reveal any statistically significant os-

cillations in the analyzed series. The application of SSA and SA allowed us to reveal at least two quasi-periodical components. In addition, SSA made it possible to reveal a significant dome-shaped trend in the prevalence of *B. progenerica* parthenitae, which were described by SA as an oscillation with a period equal to the duration of the study, as well as to give proof that there was no trend in the changes of *C. concavum* parthenitae prevalence. The components (modes) extracted by the SSA described the changes in the prevalence better than the harmonics extracted by the SA. In particular, SSA modes (contrary to SA harmonics) reflected that the amplitude of oscillations of the *B. progenerica* prevalence increased as the prevalence grew. The sums of SSA modes correlated more with initial prevalence series than the sums of SA harmonics. A possible interpretation of the trends and modes extracted by the SSA in the light of the transmission features of the investigated trematode species in the study area was proposed.

---