



КОМПЬЮТЕРНЫЕ ИССЛЕДОВАНИЯ И МОДЕЛИРОВАНИЕ

2013 Том 5 № 1

СОДЕРЖАНИЕ

МАТЕМАТИЧЕСКИЕ ОСНОВЫ И ЧИСЛЕННЫЕ МЕТОДЫ МОДЕЛИРОВАНИЯ

Виноградова П. В., Зарубин А. Г., Самусенко А. М. Метод Галёркина–Петрова для одномерных параболических уравнений высокого порядка в областях с меняющейся границей 3

ЧИСЛЕННЫЕ МЕТОДЫ И ОСНОВЫ ИХ РЕАЛИЗАЦИИ

Караваева А. С., Копысов С. П. Метод построения неструктурированных шестигранных сеток из объемных данных 11

МОДЕЛИ В ФИЗИКЕ И ТЕХНОЛОГИИ

Клименко А. А., Угольницкий Г. А. Подсистема «Разработчик» системы приема коммунальных платежей 25

Чернов И. А., Ивашко Е. Е., Никитина Н. Н., Габис И. Е. Численная идентификация модели дегидрирования в грид-системе на базе VOINC 37

АНАЛИЗ И МОДЕЛИРОВАНИЕ СЛОЖНЫХ ЖИВЫХ СИСТЕМ

Хрущев С. С., Абатурова А. М., Дьяконова А. Н., Устинин Д. М., Зленко Д. В., Федоров В. А., Коваленко И. Б., Ризниченко Г. Ю., Рубин А. Б. Моделирование белок-белковых взаимодействий с применением программного комплекса многочастичной броуновской динамики *ProKSim* 47

Устинина Д. М., Коваленко И. Б., Ризниченко Г. Ю., Рубин А. Б. Сопряжение различных методов компьютерного моделирования в комплексной модели фотосинтетической мембраны 65

Рисник Д. В., Левич А. П., Булгаков Н. Г., Бикбулатов Э. С., Бикбулатова Е. М., Еришов Ю. В., Конюхов И. В., Корнева Л. Г., Лазарева В. И., Литвинов А. С., Максимов В. Н., Мамихин С. В., Осипов В. А., Отюкова Н. Г., Поддубный С. А., Пырина И. Л., Соколова Е. А., Степанова И. Э., Фурсова П. В., Цельмович О. Л. Поиск связей между биологическими и физико-химическими характеристиками экосистемы Рыбинского водохранилища. Часть 1. Критерии неслучайности связи 83

МОДЕЛИ ЭКОНОМИЧЕСКИХ И СОЦИАЛЬНЫХ СИСТЕМ

Башкирцева И. А., Екатеринбург Е. Д., Рязанова Т. В., Сысолятина А. А. Математическое моделирование стохастических равновесий и бизнес-циклов модели Гудвина 107

COMPUTER RESEARCH AND MODELING

2013 Volume 5 Number 1

CONTENTS

MATHEMATICAL MODELING AND NUMERICAL SIMULATION

Vinogradova P. V., Zarubin A. G., Samusenko A. M. Galerkin–Petrov method for one-dimensional parabolic equations of higher order in domain with a moving boundary 3

NUMERICAL METHODS AND THE BASIS FOR THEIR APPLICATION

Karavaev A. S., Kopysov S. P. The method of unstructured hexahedral mesh generation from volumetric data 11

MODELS IN PHYSICS AND TECHNOLOGY

Klimenko A. A., Ougolnitsky G. A. Subsystem “Developer” as a part of the Retail Payment System 25

Chernov I. A., Ivashko E.E., Nikitina N.N., Gabis I. E. Numerical identification of the dehydrating model in a BOINC-based grid system 37

ANALYSIS AND MODELING OF COMPLEX LIVING SYSTEMS

Khrushchev S. S., Abaturova A. M., Diakonova A. N., Ustinin D. M., Zlenko D. V., Fedorov V. A., Kovalenko I. B., Riznichenko G. Yu., Rubin A. B. Multi-particle Brownian Dynamics software *ProKSim* for protein-protein interactions modeling 47

Ustinin D. M., Kovalenko I. B., Riznichenko G. Yu., Rubin A.B. Combination of different simulation techniques in the complex model of photosynthetic membrane 65

Risnik D. V., Levich A. P., Bulgakov N. G., Bikbulatov E. S., Bikbulatova E. M., Ershov Yu. V., Konyuhov I. V., Korneva L. G., Lazareva V. I., Litvinov A. S., Maksimov V. N., Mamihin S. V., Osipov V. A., Otyukova N. G., Poddubnii S. A., Pirina I. L., Sokolova E. A., Stepanova I. E., Fursova P. V., Cel'movich O. L. Searching for connections between biological and physico-chemical characteristics of Rybinsk reservoir ecosystem. Part 1. Criteria of connection nonrandomness 83

MODELS OF ECONOMIC AND SOCIAL SYSTEMS

Bashkirtseva I. A., Ekaterinchuk E. D., Ryazanova T. V., Sysolyatina A. A. Mathematical modeling of stochastic equilibria and business cycles of Goodwin model 107

УДК: 517.93

Метод Галёркина–Петрова для одномерных параболических уравнений высокого порядка в областях с меняющейся границей

П. В. Виноградова^{1,а}, А.Г. Зарубин², А. М. Самусенко¹

¹ Дальневосточный государственный университет путей сообщения,
Россия, 680021, г. Хабаровск, ул. Серышева, д. 47

² Тихоокеанский государственный университет,
Россия, 680035, г. Хабаровск, ул. Тихоокеанская, д. 136

E-mail: ^а vpolina17@hotmail.com

Получено 30 января 2013 г.

Исследуется начально-краевая задача для параболических уравнений высокого порядка в областях с переменной границей. Устанавливается возможность применения метода Галёркина–Петрова, и находятся асимптотические оценки скорости сходимости приближённых решений к точным.

Ключевые слова: начально-краевая задача, параболическое уравнение, метод Галёркина–Петрова, сходимость, скорость сходимости

Galerkin–Petrov method for one-dimensional parabolic equations of higher order in domain with a moving boundary

P.V. Vinogradova¹, A.G. Zarubin², A.M. Samusenko¹

¹ Far Eastern State Transport University, 680021, Khabarovsk, Serisheva 47, Russia

² Pacific National University, 680035, Khabarovsk, Tihookeanskaja 136, Russia

Abstract. – In the current paper, we study a Galerkin–Petrov method for a parabolic equations of higher order in domain with a moving boundary. Asymptotic estimates for the convergence rate of approximate solutions are obtained.

Keywords: initial boundary value problem, parabolic equation, Galerkin–Petrov method, convergence, convergence rate

Citation: *Computer Research and Modeling*, 2013, vol. 5, no. 1, pp. 3–9 (Russian).

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (проект 12-01-98506-р-восток-а).

1. Постановка задачи. Основной результат

Метод Галёркина–Петрова основан на выборе элементов двух координатных систем, причём приближённое решение находится в виде линейной комбинации по одной базисной системе, а невязка ортогональна другой базисной системе [Петров, 1940]. Если две системы связаны между собой посредством линейного оператора, то данный метод называется методом моментов. Исследованию данного проекционного метода посвящено достаточно большое количество работ, укажем, например, статьи [Дауговет, 1965; Вайникко, 1968], которые посвящены исследованию метода моментов решения краевых задач для линейных обыкновенных дифференциальных уравнений. Изучению метода моментов для линейных и квазилинейных операторных уравнений посвящены работы [Зарубин, 1978, 1987].

В данной статье исследуется метод моментов для линейных одномерных параболических уравнений в областях, граница которых меняется со временем.

Рассмотрим в R^2 криволинейную трапецию D с образующими параллельными оси Ox и ограниченными кривыми $x = \psi(t)$, $x = \phi(t)$, $0 \leq t \leq T$. Будем предполагать, что функции $\psi(t)$ и $\phi(t)$ непрерывно дифференцируемы на $[0, T]$, причём $0 < \mu \leq \phi(t) - \psi(t)$.

В области D рассмотрим следующую задачу:

$$\frac{\partial u}{\partial t} + (-1)^m \frac{\partial^{2m} u}{\partial x^{2m}} + \sum_{j=0}^r a_j(x, t) \frac{\partial^j u}{\partial x^j} = f(x, t), \quad 0 \leq r \leq 2m-1, \quad (1)$$

$$u(\psi(t), t) = u(\phi(t), t) = \dots = \frac{\partial^{m-1} u(\psi(t), t)}{\partial x^{m-1}} = \frac{\partial^{m-1} u(\phi(t), t)}{\partial x^{m-1}} = 0, \quad (2)$$

$$u(x, 0) = 0, \quad \psi(0) \leq x \leq \phi(0). \quad (3)$$

Для дальнейшего исследования задачи (1)–(3) нам потребуются пространства $W_2^{2m,1}(D)$, $L_p(D)$, определение которых можно найти, например, в [Ладыженская, Солонников, Уральцева, 1967].

Наряду с задачей (1)–(3) рассмотрим в цилиндре $\bar{Q} = \{(\xi, t), 0 \leq \xi \leq 1, 0 \leq t \leq T\}$ задачу

$$\frac{\partial v}{\partial t} + (-1)^m a(t) \frac{\partial^{2m} v}{\partial \xi^{2m}} + \sum_{j=0}^r b_j(\xi, t) \frac{\partial^j v}{\partial \xi^j} = f_1(\xi, t), \quad (4)$$

$$v(0, t) = v(1, t) = \dots = \frac{\partial^{m-1} v(0, t)}{\partial \xi^{m-1}} = \frac{\partial^{m-1} v(1, t)}{\partial \xi^{m-1}} = 0, \quad (5)$$

$$v(\xi, 0) = 0, \quad 0 \leq \xi \leq 1, \quad (6)$$

где функции $a(t)$, $b_j(\xi, t)$ непрерывны в \bar{Q} , причём $a(t) \geq \mu_1 > 0$ для всех $t \in [0, T]$. Пусть $f_1(\xi, t)$ принадлежит $L_2(Q)$, тогда задача (4)–(6) имеет решение $v(\xi, t)$ из $W_2^{2m,1}(Q)$ и оно единственно (см. [Ладыженская, Солонников, Уральцева, 1967]).

Предположим, что выполняются равенства

$$b_j(\xi, t) = \frac{a_j(\xi(\phi(t) - \psi(t)) + \psi(t), t)}{(\phi(t) - \psi(t))^j}, \quad j = 0, 1, 2, \dots, r, \quad a(t) = \frac{1}{(\phi(t) - \psi(t))^{2m}}, \quad (7)$$

$$f_1(\xi, t) = f(\xi(\phi(t) - \psi(t)) + \psi(t), t),$$

$\xi = \frac{x - \psi(t)}{\phi(t) - \psi(t)}$, $(\xi, t) \in Q$, и $v(\xi, t)$ — решение задачи (4)–(6) из пространства $W_2^{2m,1}(Q)$. Тогда

функция $u(x, t) = v\left(\frac{x - \psi(t)}{\phi(t) - \psi(t)}, t\right)$ принадлежит пространству $W_2^{2m,1}(D)$ и является решением задачи (1)–(3).

На пространстве $H_1 = W_2^{2m}(0,1) \cap \overset{\circ}{W}_2^m(0,1)$ определим линейные операторы $A(t)$ и $K(t)$ равенствами

$$A(t)z(\xi) = \frac{(-1)^m}{(\varphi(t) - \psi(t))^{2m}} \frac{d^{2m}z(\xi)}{d\xi^{2m}},$$

$$K(t)z(\xi) = \sum_{j=0}^r b_j(\xi, t) \frac{\partial^j z(\xi)}{\partial \xi^j}.$$

Тогда задача (4)–(6) примет вид

$$\frac{\partial v}{\partial t} + A(t)v + K(t)v = f_1(\xi, t), \tag{8}$$

$$v(\xi, 0) = 0. \tag{9}$$

Известно, что система функций $\{\sin(s\pi\xi)\}_{s=1}^\infty$ является полной ортогональной системой в $L_2(0,1)$. Рассмотрим функции $\phi_s(\xi)$, которые являются решением задачи

$$\frac{d^{2m}\phi_s(\xi)}{d\xi^{2m}} = \sin(s\pi\xi),$$

$$\phi_s(0) = \phi_s(1) = \dots = \frac{d^{m-1}\phi_s(0)}{d\xi} = \frac{d^{m-1}\phi_s(1)}{d\xi}.$$

Очевидно, что функции $\phi_s(\xi)$ можно выписать в явном виде.

Пусть P_n — ортопроектор в $L_2(0,1)$ на линейную оболочку функций $e_s(\xi) = \sin(s\pi\xi)$, $s = 1, \dots, n$. Согласно методу моментов, приближённое решение задачи (8)–(9) находим в виде конечной суммы $v_n(\xi, t) = \sum_{s=1}^n c_s(t)\phi_s(\xi)$, где неизвестные коэффициенты $c_s(t)$ являются решениями задачи Коши

$$P_n \frac{\partial v_n}{\partial t} + P_n A(t)v_n + P_n K(t)v_n = P_n f_1(\xi, t), \tag{10}$$

$$v_n(\xi, 0) = 0. \tag{11}$$

Тогда функцию $u_n(x, t) = v_n\left(\frac{x - \psi(t)}{\varphi(t) - \psi(t)}, t\right)$ назовём решением задачи (1)–(3), построенным по методу моментов, если выполнены равенства (7).

В дальнейшем через M_i будем обозначать различные положительные постоянные, не зависящие от n и t . Норму в $L_2(0,1)$ будем обозначать через $\|\cdot\|$.

Теорема 1. Пусть функция $f_1(\xi, t) \in L_2(Q)$, $f_1(\xi, 0) = 0$ и имеет производную $\frac{\partial f_1(\xi, t)}{\partial t}$ из $L_2(Q)$. Пусть функции $\phi(t)$, $\psi(t)$ непрерывно дифференцируемы на $[0, T]$, а функции $b_j(\xi, t)$ имеют непрерывную производную $\frac{\partial b_j(\xi, t)}{\partial t}$. Пусть коэффициенты $b_j(\xi, t)$ такие, что

$$(K(t)z, z)_{L_2(0,1)} \geq 0 \tag{12}$$

для любой функции $z(\xi)$ из H_1 . Тогда

$$\sup_{0 \leq t \leq T} \|v_n(\xi, t) - v(\xi, t)\| \leq M_1 n^{-m+\frac{1}{2}}. \quad (13)$$

Доказательство. Умножим (10) скалярно на $\frac{\partial^{2m} v_n(\xi, t)}{\partial \xi^{2m}}$ и проинтегрируем по t от 0 до ζ ($\zeta \leq T$), тогда

$$\frac{1}{2} \left\| \frac{\partial^m v_n(\xi, \zeta)}{\partial \xi^m} \right\|^2 + \mu_1 \int_0^\zeta \left\| \frac{\partial^{2m} v_n(\xi, \zeta)}{\partial \xi^{2m}} \right\|^2 dt \leq \int_0^\zeta \|K(t)v_n\| \left\| \frac{\partial^{2m} v_n}{\partial \xi^{2m}} \right\| dt + \|f_1\|_{L_2(Q)} \left(\int_0^\zeta \left\| \frac{\partial^{2m} v_n}{\partial \xi^{2m}} \right\|^2 dt \right)^{\frac{1}{2}}, \quad (14)$$

где $\mu_1 = (\sup(\phi(t) - \psi(t))^{2m})^{-1}$. Так как функции $b_j(\xi, t)$ непрерывны в Q , то

$$\|K(t)v_n\| \leq M_2 \|v_n\|_{W_2^r(0,1)}. \quad (15)$$

Применим к правой части мультипликативное неравенство (см. [Глушко, Крейн, 1960]) для пространств $W_2^r(0,1)$, $W_2^{2m}(0,1)$ и $L_2(0,1)$. Тогда из (15) получаем

$$\|K(t)v_n\| \leq M_3 \|v_n\|_{W_2^{2m}}^{\frac{r}{2m}} \|v_n\|^{1-\frac{r}{2m}} \leq M_4 \left\| \frac{\partial^{2m} v_n}{\partial \xi^{2m}} \right\|^{\frac{r}{2m}} \|v_n\|^{1-\frac{r}{2m}}. \quad (16)$$

Отсюда и из (14) следует оценка

$$\frac{1}{2} \left\| \frac{\partial^m v_n(\xi, \zeta)}{\partial \xi^m} \right\|^2 + \mu_1 \int_0^\zeta \left\| \frac{\partial^{2m} v_n(\xi, \zeta)}{\partial \xi^{2m}} \right\|^2 dt \leq M_5 \left(\int_0^\zeta \left\| \frac{\partial^{2m} v_n}{\partial \xi^{2m}} \right\|^{\left(1+\frac{r}{2m}\right)} \|v_n\|^{1-\frac{r}{2m}} dt \right) + \|f_1\|_{L_2(Q)} \left(\int_0^\zeta \left\| \frac{\partial^{2m} v_n}{\partial \xi^{2m}} \right\|^2 dt \right)^{\frac{1}{2}}.$$

Если к правой части применим неравенство Юнга и неравенство Фридрикса, то получим

$$\left\| \frac{\partial^m v_n(\xi, \zeta)}{\partial \xi^m} \right\|^2 \leq M_6 \left(\|f_1\|_{L_2(Q)}^2 + \int_0^\zeta \left\| \frac{\partial^{2m} v_n(\xi, \zeta)}{\partial \xi^{2m}} \right\|^2 dt \right).$$

Отсюда и из неравенства Гронуолла вытекает оценка

$$\sup_{0 \leq t \leq T} \left\| \frac{\partial^m v_n(\xi, t)}{\partial \xi^m} \right\| \leq M_7. \quad (17)$$

Рассмотрим задачу Коши:

$$P_n \frac{\partial w_n}{\partial t} + P_n A(t)w_n + P_n K(t)w_n = P_n \frac{\partial f_1(\xi, t)}{\partial t} - P_n A'(t)v_n - P_n K'(t)v_n, \quad (18)$$

$$w_n(\xi, 0) = 0. \quad (19)$$

Для $w_n(\xi, t)$ по аналогии с получением оценки (17) устанавливается неравенство

$$\sup_{0 \leq t \leq T} \left\| \frac{\partial^m w_n(\xi, t)}{\partial \xi^m} \right\| \leq M_8. \quad (20)$$

Так как $f_1(\xi, 0) = 0$ и $v_n(\xi, 0) = 0$, то из уравнения (10) вытекает тождество $P_n \frac{\partial v_n}{\partial t}(\xi, 0) = 0$.

Тогда

$$0 = \left(P_n \frac{\partial v_n}{\partial t}(\xi, 0), A \frac{\partial v_n}{\partial t}(\xi, 0) \right) = \left(\frac{\partial v_n}{\partial t}(\xi, 0), A \frac{\partial v_n}{\partial t}(\xi, 0) \right) = \left\| \frac{\partial^{m+1} v_n}{\partial \xi^m \partial t}(\xi, 0) \right\|^2 \geq M_9 \left\| \frac{\partial v_n}{\partial t}(\xi, 0) \right\|^2.$$

Итак, $\frac{\partial v_n}{\partial t}(\xi, 0) = 0$ для всех ξ из $[0, 1]$.

Коэффициенты $b_j(\xi, t)$ и функция $f_1(\xi, t)$ непрерывно дифференцируемы по t , поэтому решение $v_n(\xi, t)$ задачи (10)–(11) удовлетворяет тождествам

$$P_n \frac{\partial^2 v_n}{\partial t^2} + P_n A(t) \frac{\partial v_n}{\partial t} + P_n K(t) \frac{\partial v_n}{\partial t} = P_n \left(\frac{\partial f_1}{\partial t} - A'(t)v_n - K'(t)v_n \right),$$

$$v_n(\xi, 0) = 0, \frac{\partial v_n}{\partial t}(\xi, 0) = 0.$$

Отсюда и из (18) следует

$$P_n \frac{\partial z_n}{\partial t} + P_n A(t)z_n + P_n K(t)z_n = 0,$$

где $z_n = w_n(\xi, t) - \frac{\partial v_n(\xi, t)}{\partial t}$. Из последнего тождества вытекает оценка $\left\| w_n(\xi, t) - \frac{\partial v_n(\xi, t)}{\partial t} \right\| \leq 0$.

Таким образом, $w_n(\xi, t) = \frac{\partial v_n(\xi, t)}{\partial t}$. Последнее равенство и оценка (20) приводят к соотношению

$$\sup_{0 \leq t \leq T} \left\| \frac{\partial^{m+1} v_n(\xi, t)}{\partial \xi^m \partial t} \right\| \leq M_9. \tag{21}$$

Пусть $v(\xi, t)$ и $v_n(\xi, t)$ — решения задач (4)–(6) и (10)–(11). Тогда

$$\frac{\partial}{\partial t}(v - v_n) + A(t)(v - v_n) + K(t)(v - v_n) = (I - P_n) \left(f_1 - \frac{\partial v_n}{\partial t} - K(t)v_n \right). \tag{22}$$

Умножим (22) скалярно в $L_2(0, 1)$ на $(v - v_n)$ и проинтегрируем по t от 0 до ξ . Используя (12), имеем

$$\frac{1}{2} \|v - v_n\|^2 + \mu_1 \int_0^\xi \int_0^1 \left| \frac{\partial^m}{\partial \xi^m} (v - v_n) \right|^2 d\xi dt \leq \int_0^\xi \left((I - P_n) \left(f_1 - \frac{\partial v_n}{\partial t} - K(t)v_n \right), (v - v_n) \right) dt. \tag{23}$$

Из определения функций $\phi_s(\xi)$ следует, что для любой функции $z(\xi)$ из $L_2(0, 1)$ верно неравенство $\left\| (I - P_n) \left(\frac{d^{2m}}{d\xi^{2m}} \right)^{-1} z(\xi) \right\| \leq M_{10} n^{-2m+1}$. Отсюда и из оценок (23) и (16) получаем

$$\frac{1}{2} \|v - v_n\|^2 + \mu_1 \int_0^\xi \int_0^1 \left| \frac{\partial^m}{\partial \xi^m} (v - v_n) \right|^2 d\xi dt \leq M_{11} n^{-m+\frac{1}{2}} \int_0^\xi \left(\|f_1\| + \left\| \frac{\partial v_n}{\partial t} \right\| + \left\| \frac{\partial^{2m} v_n}{\partial \xi^{2m}} \right\| \right) \left\| \frac{\partial^m (v - v_n)}{\partial \xi^m} \right\| dt.$$

Если применить к правой части последнего соотношения ε -неравенство, а также оценку

$$a_0(x, t) = a_2(x, t) = a_3(x, t) = 0, a_1(x, t) = (t+1)e^{-x}, f(x, t) = tx(x^3 - 1),$$

тогда задача (1)–(3) примет вид

$$\begin{aligned} \frac{\partial u}{\partial t} + \frac{\partial^4 u}{\partial x^4} + (t+1)e^{-x} \frac{\partial u}{\partial x} &= tx(x^3 - 1), \\ u\left(-\frac{1}{9}t^2 + 1, t\right) &= u\left(\frac{1}{9}t^2 + 3, t\right) = \frac{\partial u}{\partial t}\left(-\frac{1}{9}t^2 + 1, t\right) = \frac{\partial u}{\partial t}\left(\frac{1}{9}t^2 + 3, t\right) = 0, \\ u(x, 0) &= 0, \psi(0) \leq x \leq \varphi(0), 0 \leq t \leq 1. \end{aligned}$$

В цилиндре $\bar{Q}=[0;1] \times [0;T]$ рас $\beta_k(\xi)$ являются решением следующей задачи:

$$\begin{aligned} \frac{d^4 \beta_k(\xi)}{d\xi^4} &= \sin(k\pi\xi), \\ \beta_k(0) = \beta_k(1) &= \frac{d\beta_k}{d\xi}(0) = \frac{d\beta_k}{d\xi}(1). \end{aligned}$$

Решения $\beta_k(\xi)$ имеют следующий вид:

$$\beta_k(\xi) = \frac{\sin(k\pi\xi)}{(\pi k)^4} - \frac{(1+(-1)^k)\xi^3}{(\pi k)^3} + \frac{(2+(-1)^k)\xi^2}{(\pi k)^3} - \frac{\xi}{(\pi k)^3}.$$

Составим невязку и запишем требование ортогональности в виде

$$\begin{aligned} \int_0^1 \left(\frac{\partial v}{\partial t} + \left(\frac{2}{9}t^2 + 2\right) \frac{\partial^4 v}{\partial \xi^4} + (t+1)e^{-\left(\xi\left(\frac{2}{9}t^2+2\right)-\frac{1}{9}t^2+1\right)} \frac{\partial v}{\partial \xi} - f_1(\xi, t) \right) \sin(m\pi\xi) d\xi &= 0, \\ m &= 1, \dots, n. \end{aligned}$$

Коэффициенты $\alpha_k(t)$ находятся как решения задачи Коши для системы дифференциальных уравнений:

$$\begin{aligned} \sum_{k=0}^n \frac{d\alpha_k(t)}{dt} \int_0^1 \beta_k(\xi) \sin(m\pi\xi) d\xi + \sum_{k=0}^n \left(\frac{2}{9}t^2 + 2\right)^{-4} \alpha_k(t) \int_0^1 \frac{d^4 \beta_k(\xi)}{d\xi^4} \sin(m\pi\xi) d\xi + \\ + \sum_{k=0}^n (t+1) \alpha_k(t) \int_0^1 e^{-\left(\xi\left(\frac{2}{9}t^2+2\right)-\frac{1}{9}t^2+1\right)} \frac{d\beta_k(\xi)}{d\xi} \sin(m\pi\xi) d\xi - \int_0^1 f_1(\xi, t) \sin(m\pi\xi) d\xi &= 0, \\ \alpha_k(0) &= 0, \\ m &= 1, \dots, n. \end{aligned}$$

Для нахождения приближённого решения данной задачи применим метод Рунге–Кутты. Вычислим значения $\alpha_k(t)$ в узлах сетки. Тогда приближенное решение исходной задачи в узлах сетки имеет вид

$$u_n(x, t_s) = \sum_{k=1}^n \alpha_k(t_s) \beta_k\left(\frac{x - \psi(t_s)}{\varphi(t_s) - \psi(t_s)}\right).$$

На рисунке 1 представлен график приближенного решения.

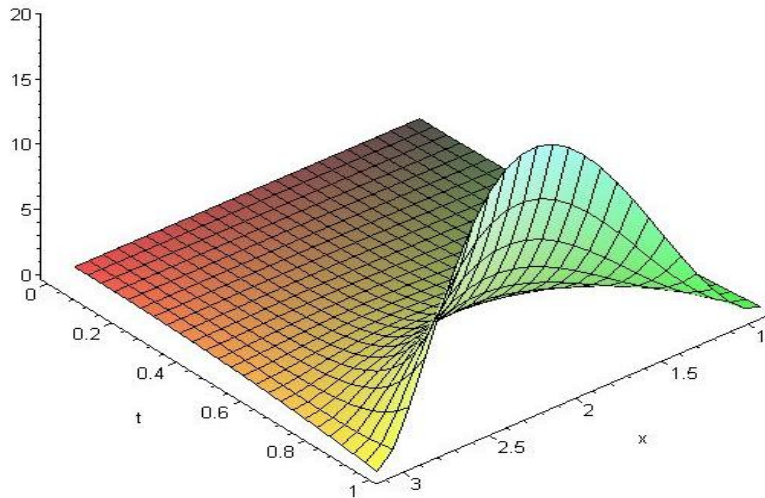


Рис. 1

Список литературы

- Вайникко Г. М.* О быстроте сходимости метода моментов для обыкновенных дифференциальных уравнений // Сибирский математический журнал. — 1968. — Т. 9, № 1. — С. 21–28.
- Глушко В. П., Крейн С. Г.* Неравенства для норм производных в пространствах с весом // Сибирский математический журнал. — 1960. — Т. 1, № 3. — С. 343–382.
- Дауговет И. К.* О методе моментов для обыкновенных дифференциальных уравнений // Сибирский математический журнал. — 1965. — Т. 6, № 1. — С. 70–75.
- Зарубин А. Г.* О методе моментов для одного класса нелинейных уравнений // Сибирский математический журнал. — 1978. — Т. 19, № 3. — С. 575–586.
- Зарубин А. Г.* Исследование проекционной процедуры Галёркина–Петрова методом дробных степеней // Доклады АН СССР. — 1987. — Т. 297, № 4. — С. 780–784.
- Ладыженская О. А., Солонников В. А., Уральцева Н. Н.* Линейные и квазилинейные уравнения параболического типа. — М.: Наука, 1967.
- Петров Г. И.* Применение метода Галёркина–Петрова к задаче об устойчивости течения вязкой жидкости // Прикладная математика и механика. — 1940. — Т. 4. — С. 1–13.

УДК: 004.925.8

Метод построения неструктурированных шестигранных сеток из объемных данных

А. С. Карavaев^а, С. П. Копысов^б

Институт механики УрО РАН,
Россия, 426067, г. Ижевск, ул. Т. Барамзиной, 34

E-mail: ^а karavaev-alexander@yandex.ru, ^б s.kopysov@gmail.com

Получено 22 января 2013 г.

Разработан метод преобразования поверхностного представления расчетной области в формат воксельных данных. Предложен алгоритм генерации расчетной шестигранной сетки на основе объемного формата данных.

Ключевые слова: неструктурированная шестигранная сетка, объемные данные, воксель, знакоопределенные поля расстояний, двойственное контурирование

The method of unstructured hexahedral mesh generation from volumetric data

A. S. Karavaev, S. P. Kopysov

Institute of Mechanics UB RAS 34 ul. T. Baramzinoy, Izhevsk, Russia 426067

Abstract. — The method of conversion of surface representation to the voxel data is developed. Voxel data based algorithm of hexahedral mesh generation is proposed.

Keywords: hexahedral mesh, volumetric data, voxel, signed distance fields, dual contouring

Citation: *Computer Research and Modeling*, 2013, vol. 5, no. 1, pp. 11–24 (Russian).

Работа выполнена при поддержке РФФИ (проект № 12-07-31114)

Введение

Существует несколько способов представления объектов в 3D-пространстве. Все их можно разделить на две группы: поверхностное, определяющее тело путем задания ограничивающей его поверхности, и объемное, содержащее всю информацию о внутренних точках объекта в пространстве.

В настоящее время поверхностное представление является распространенным способом задания геометрии расчетной области. Однако в приложениях, где рассматриваются структуры сложной формы (например моделирование в биомеханике, механике композиционных материалов) и когда возникает необходимость создания модели с внутренней структурой, использование воксельного представления данных является более предпочтительным. Также такой вид представления используется, когда геометрическая модель имеет неявное представление границы (фронт пламени, разделы фаз и т. д.). Для воксельных моделей (например медицинских данных со сканера магнитно-резонансной томографии) реализуется вывод любого сечения модели, что позволяет изучать любой срез данных.

Во многом алгоритмы построения расчетной сетки зависят от того, какой способ представления исходных данных они используют. Этим определяется граница применимости метода, точность и скорость работы, сложность практической реализации.

Разработано достаточное количество методов построения неструктурированных шестигранных сеток, большинство из них обладают небольшой степенью надежности, ограничением, накладываемым на применимость к определенному виду областей. Данные алгоритмы можно разделить на две группы: прямые методы — когда построение шестигранной сетки идет непосредственно внутри заданной границы; непрямые методы — когда чаще всего используется предварительно построенная тетраэдральная сетка с последующим преобразованием тетраэдральных ячеек в шестигранные либо осуществляется преобразование другого исходного дискретного представления области. Как правило, при построении расчетной сетки в качестве исходного описания области используют поверхностное представление (CAD-модель).

Среди прямых методов можно выделить методы регулярной сетки или методы суперпозиции, основанные на генерировании регулярной шестигранной сетки внутри объема с последующим созданием переходных ячеек, соединяющих регулярную сетку с четырехугольниками на границе [Schneiders et al., 1996]. Для этих методов часто возникает ситуация наличия вырожденных ячеек или ячеек плохого качества вдоль границы. При всей практичности такого подхода он является весьма трудоемким для реализации [Thompson et al., 1999].

Вторую группу составляют так называемые фронтальные алгоритмы, когда построение начинается от граничной четырехугольной сетки [Staten et al., 2005]. На каждой итерации проекцией к центру объема создается новый слой шестигранных ячеек. Такой способ обладает тем же недостатком, что и его двумерный аналог — сложным поиском соединения граней и ребер ячеек в случае пересечения двух и более фронтов.

Методы декомпозиции и блочной декомпозиции основаны на идее (ручной или полуавтоматической) декомпозиции тела на простые геометрические примитивы (параллелепипед, шар, цилиндр и др.), для каждого из которых построение расчетной сетки является простой задачей, с последующим объединением этих частей в исходный объект. Одним из способов автоматизации процесса декомпозиции на блоки является использование срединных поверхностей [Price, Armstrong, 1995]. Достоинством данной группы методов является их высокая скорость и качество получаемой сетки, однако они имеют высокую трудоемкость их использования для дискретизации сложных нестандартных областей.

Простейшим непрямым методом построения является алгоритм Эпштейна — пошаговое преобразование каждой тетраэдральной ячейки путем разбиения исходной на четыре новых шестигранника [Eppstein, 1996]. В этом случае модифицируются граничные ребра и получаемая

сетка полностью состоит из шестигранных элементов. Отметим низкое качество получаемых сеток, а также невозможность получения результирующей сетки, соразмерной с исходной, что делает его редко применяемым в практике вычислений.

Среди фронтальных непрямых методов можно выделить алгоритм H-Morph [Owen, Saigal, 2000], в котором трансформация ячеек тетраэдров в шестигранники начинается от так называемых граней подвижного фронта — замкнутой совокупности граней, отделяющей построенные шестигранные ячейки сетки от тетраэдральных. Далее в процессе работы фронт постоянно обновляется в соответствии с текущим состоянием сетки. Примеры работы данного алгоритма для двухмерного случая показывают хорошие результаты [Owen et al., 1998; Сковпень, 1998; Караваев и др., 2012], однако нужно отметить сложную практическую реализацию метода в 3D, где затраты построения сетки возрастают.

Выделим еще группу методов поэлементного преобразования тетраэдров на основе шаблонов, которые обеспечивают невырожденность ячеек. Однако, часто в результирующей сетке остаются тетраэдральные ячейки [Копысов, Пономарев, 2009]. Заслуживают внимания методы, использующие двойственное представление шестигранной сетки через ортогональные плоскости, определяющие связность ее ячеек [Murdoch et al., 1997]. Преимуществом данной группы методов является отсутствие вырожденных элементов, но накладываются требования построения качественной начальной дискретизации, что само по себе является трудоемкой задачей.

В данной работе представляется метод построения шестигранной сетки из объемных данных, основанный на алгоритме двойственного контурирования, используемого для построения изображения моделируемых тел.

Объемное представление данных

Объемные данные \mathcal{V} представляются в виде набора точек (вокселей) $v \in \mathcal{V}$, каждая из которых имеет некоторое скалярное значение, являющееся атрибутом данной точки в пространстве (плотность, концентрация, температура и т. д.). Объемные (воксельные) данные есть выборка значений функции на декартовой сетке координат, которая может быть записана как $\mathcal{V} = \{F(i, j, k) | i, j, k, \text{ где } i, j, k \text{ — индексы в } x, y, z \text{ декартовых координат сетки}\}$. Изоповерхность (граница объекта, неявная поверхность), соответствующая изозначению α , определяется как $I_F(\alpha) = \{(x, y, z) | F(x, y, z) = \alpha\}$.

В работе рассматривается представление воксельных данных для случая, когда сохраняется наименьшее расстояние от каждого вокселя до границы рассматриваемого тела с учетом знака. Это представление данных получило название «знакоопределенные поля расстояний». В данной статье описывается метод преобразования поверхностного представления расчетной области в формат знакоопределенных полей расстояния. На основе полученных данных строится шестигранная расчетная сетка с использованием алгоритма двойственного контурирования. Адаптация построенной сетки в областях, требующих дополнительного уточнения границы, производилась шаблонами «перестроения на 27 ячеек» Шнейдера.

Метод преобразования поверхностного представления расчетной области в формат воксельных данных

В качестве исходного и широко представленного на практике формата поверхностного представления использовался формат .STL (описание тела с помощью поверхностной треугольной сетки). Кроме того, в некоторых случаях такой формат является единственным, например, для данных, полученных путем сканирования. После преобразования создается файл формата .rawiv, содержащий знакоопределенные поля расстояний. В отличие от стандартного представления, где значение вокселя F не содержит знак (плотность, температура), расстояние является

знакоопределенным ($F > 0$, если воксель внутри тела, $F < 0$, если снаружи, $F = 0$ в случае расположения на границе). Функция $SD_{\mathcal{T}}(p)$, определяющая такое расстояние, получила название «знакоопределенная функция расстояния». Пусть $\Omega \subset R^3$ замкнутая область, ограниченная поверхностной треугольной сеткой \mathcal{T} , тогда для произвольной точки $p \in R^3$ данную функцию можно записать в виде

$$SD_{\mathcal{T}}(p) = S_{\mathcal{T}}(p) \cdot D_{\mathcal{T}}(p), \quad D_{\mathcal{T}}(p) = \inf_{n \in \mathcal{N}} \|n - p\|, \quad S_{\mathcal{T}}(p) = \begin{cases} +1, & \text{если } p \in \Omega; \\ -1, & \text{если } p \notin \Omega. \end{cases}$$

где \mathcal{N} — множество узлов сетки \mathcal{T} , n — любой узел \mathcal{T} , принадлежащий \mathcal{N} . Как видно, знакоопределенная функция расстояния состоит из произведения функции знака $S_{\mathcal{T}}(p)$ и функции наименьшего расстояния $D_{\mathcal{T}}(p)$. При построении метода преобразования поверхностной треугольной сетки использовался следующий алгоритм: в цикле обходится массив вокселей декартовой сетки и для каждого вокселя v находится наименьшее расстояние до поверхности объекта с учетом знака [Fuhrmann, 2007].

Функция $SD_{\mathcal{T}}(p)$ вычисляется в два этапа: нахождение наименьшего расстояния; определение знака вокселя. Для нахождения наименьшего расстояния был использован алгоритм проекции исходной точки на поверхностную треугольную сетку, описывающую объект. Основная идея его состоит в проекции исходной точку p_0 на плоскость, образованную треугольником Δ , состоящим из вершин p_1, p_2, p_3 (рис. 1). Если проекция p' попадает внутрь треугольника, то длина отрезка $[p_1, p']$ является искомой величиной. В противном случае ищется ближайшая вершина или ребро треугольника Δ до точки p' и вычисляется расстояние до этого элемента. Если предположить, что ребро e , образованное точками p_1 и p_3 , является ближайшим к точке p' , и p'' есть проекция точки p' на e , то расстояние до треугольника Δ определяется как $\sqrt{|p'p''|^2 + |p'p_0|^2}$.

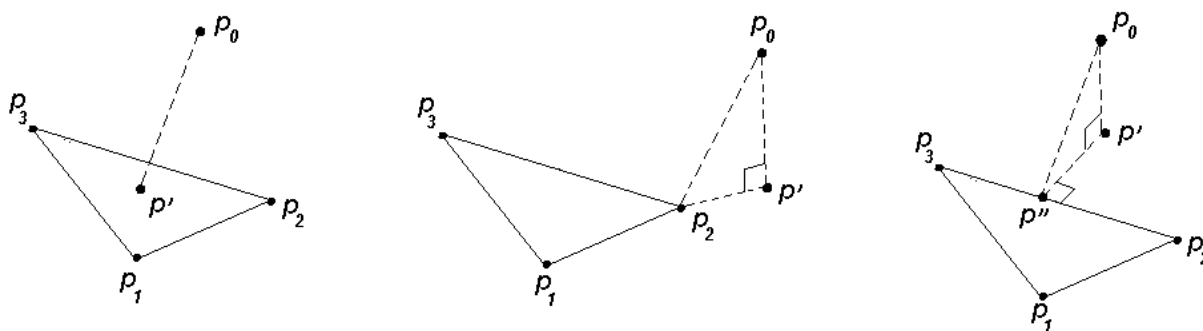


Рис. 1. Нахождение расстояния от точки p_0 до $\Delta p_1 p_2 p_3$

Положение произвольной точки p относительно тела определится с помощью нормалей заданных на его поверхности. Для формата .STL на поверхностной сетке заданы нормали треугольников. Нормаль ребра сетки вычисляется как нормализованная сумма нормалей прилежащих к нему треугольников. Пусть к ребру прилегают треугольники Δ_1 и Δ_2 , тогда нормаль ребра $\mathbf{n} = (\mathbf{n}_{\Delta_1} + \mathbf{n}_{\Delta_2}) / (\|\mathbf{n}_{\Delta_1} + \mathbf{n}_{\Delta_2}\|)$. Для вычисления нормали вершины сетки использовалась псевдонормаль, взвешенная по углам $\mathbf{n} = \sum_i \mathbf{n}_i \cdot \beta_i / \|\sum_i \mathbf{n}_i \cdot \beta_i\|$, где \mathbf{n}_i — нормали прилежащих к вершине треугольников, β_i — угол треугольника в рассматриваемой вершине.

Алгоритмы построения объемного изображения из воксельных данных

Самым известным и простым в реализации методом изображения объекта из объемных данных является алгоритм «марширующие кубы», предложенный Лоренсом [Lorensen, Cline,

1987]. В данном алгоритме осуществляется обход скалярного поля с выбором восьми соседних вокселей на каждой итерации. Далее определяются полигоны, необходимые для представления части изоповерхности, проходящей через данный куб. Совокупность данных полигонов и составляет поверхность объекта.

Наряду со своими преимуществами, такими как скорость и простота реализации, алгоритм имеет один существенный недостаток. Он плохо описывает острые углы рассматриваемой модели, малые углы пропускаются алгоритмом, вместо этого создаются полигоны, соединяющие точки пересечения с кубом декартовой решетки. Такой метод достаточно хорош для задания поверхности сферы, но дает плохую аппроксимацию параллелепипеда. Вторая особенность заключается в том, что сетка, генерируемая «марширующими кубами», является однородной, что приводит к неоправданным затратам памяти. Модификации алгоритма типа «расширенные марширующие кубы» требуют специальных вычислений при определении острого угла поверхности, а также дополнительный шаг переключения ребер для завершения аппроксимации в случае острых ребер границы.

Другую группу методов составляют так называемые двойственные методы. Алгоритм двойственного контурирования [Schaefer, Warren, 2002], создавался как метод построения изображения в компьютерной анимации, но благодаря своей простоте и эффективности получил широкое распространения в других областях компьютерного моделирования. Двойственное контурирование сохраняет острые углы объекта, а также генерирует поверхностную четырехугольную сетку. Основные особенности алгоритма:

- Использование структуры октодеревя для представления прямоугольной сетки. Каждый узел октодеревя содержит восемь вокселей, составляющих его вершины.
- Наряду с плотностью вокселя учитывается его нормаль к неявной поверхности. Нормаль точек пересечения с контуром поверхности определяется линейной интерполяцией соседних вокселей. Вершина сетки находится путем минимизации квадратичной функции погрешности $E(\mathbf{x}) = \sum_i (\mathbf{n}_i \cdot (\mathbf{x} - \mathbf{x}_i))^2$, где \mathbf{x}_i — вектор, содержащий координаты точки пересечения, а \mathbf{n}_i — ее единичная нормаль.
- Вычисление квадратичной функции погрешности для каждого куба декартовой сетки вокселей, пересекающего контур неявной поверхности вне зависимости от того, содержит он острый угол или нет. Благодаря этому вершины поверхности всегда располагаются внутри ячеек сетки. На рисунке 2, а белыми и черными квадратами изображены воксели расположенные внутри и снаружи тела соответственно, также показаны точки пересечения ребер решетки с контуром изоповерхности и их единичные нормали. На рисунках 2, б, в, г показаны варианты работы алгоритмов. Алгоритм «марширующие кубы» всегда располагает вершины на ребрах сетки, «расширенные марширующие кубы» генерируют вершину внутри ячейки для случая описания острого угла модели, двойственное контурирование всегда располагает вершины внутри ячеек сетки.

Исходя из возможности построения поверхностной четырехугольной сетки и использования структуры октодеревя, допускающего адаптацию сетки, метод был взят за основу при создании поверхностной однородной сетки предлагаемого алгоритма построения шестигранных ячеек.

Построение шестигранной сетки из объемного изображения

Предлагаемый алгоритм в качестве исходных данных использует знакоопределенные поля расстояний. Для хранения точек в пространстве и уровней детализации объекта используется структура октодеревя. Построения поверхностной четырехугольной сетки, а также выбор

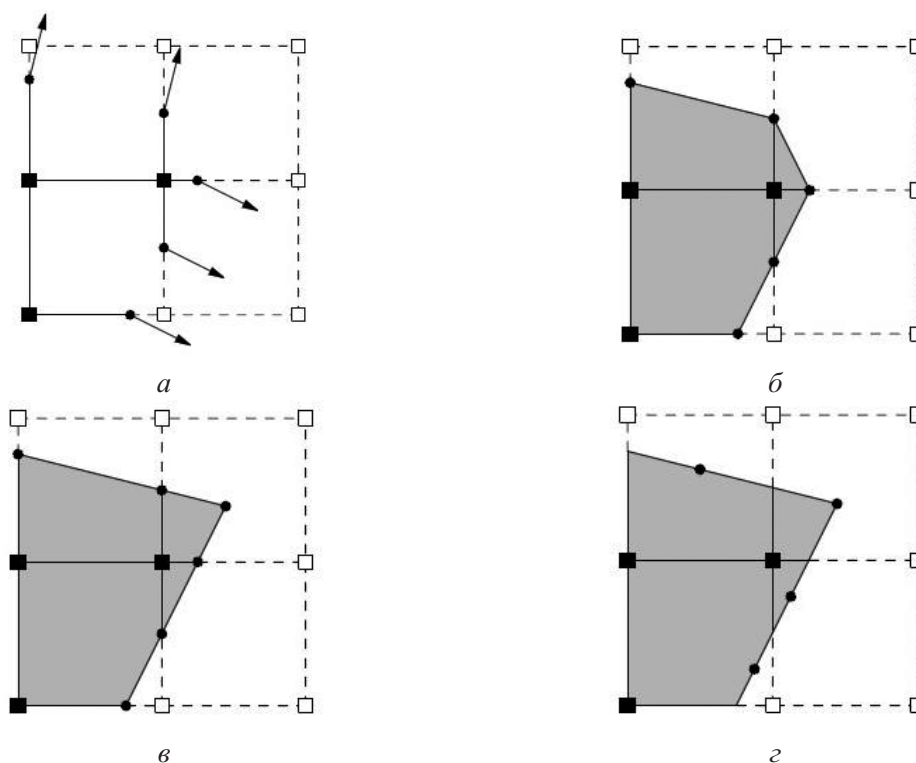


Рис. 2. Декартова сетка вокселей с нормальными к контуру поверхности (а). Поверхность построенная «марширующими кубами» (б). Поверхность, сгенерированная алгоритмом «расширенные марширующие кубы» (в). «Двойственное контурирование» (г)

начального уровня октодеревя выполняется на основе алгоритма двойственного контурирования, который позволяет сохранить острые углы и грани объекта. Основные этапы алгоритма следующие:

- Построение уровней октодеревя.
- Генерация однородной шестигранной сетки.
- Улучшение качества построенных ячеек.

Построение уровней октодеревя

Октодеревя — тип древовидной структуры данных, в которой у каждого внутреннего узла есть до восьми потомков. Как и в данном случае, деревья октантов чаще всего используются для разделения трехмерного пространства, рекурсивно разделяя его на восемь октантов. В алгоритме двойственного контурирования узел октодеревя хранит информацию о восьми вокселях, составляющих куб декартовой сетки. В каждом вокселе с индексами декартовой сетки (x_i, y_j, z_k) вычисляется градиент скалярного поля и нормаль равная нормализованному градиенту со знаком:

$$\mathbf{grad}F(x_i, y_j, z_k) = \frac{1}{2} \begin{pmatrix} F(x_{i+1}, y_j, z_k) - F(x_{i-1}, y_j, z_k) \\ F(x_i, y_{j+1}, z_k) - F(x_i, y_{j-1}, z_k) \\ F(x_i, y_j, z_{k+1}) - F(x_i, y_j, z_{k-1}) \end{pmatrix}, \quad \mathbf{n}(x, y, z) = -\frac{\mathbf{grad}F(x, y, z)}{|\mathbf{grad}F(x, y, z)|},$$

где $F(x_i, y_j, z_k)$ — значение вокселя (x_i, y_j, z_k) .

В алгоритме используется структура октодера с размерностью данных $dim \times dim \times dim$, которая удовлетворяет условию $2^m = dim - 1$, $m \in N$ и обеспечивает разделение подпространства строго кубической формы. Задаваемое число уровней октодера зависит от размерности исходных данных. В этом случае используется восходящий алгоритм, начинающий построение с самого нижнего уровня. При исходной размерности, равной $(2^k + 1)^3$, число уровней октодера равно k .

От выбранного уровня октодера зависит степень детализации изображаемого объекта, а следовательно точность аппроксимации границы физической модели. На рисунках 3, 4, 5, изображены примеры генерации изображения из файла объемных данных размерности $dim \times dim \times dim = 65 \times 65 \times 65$, при этом слева показаны узлы октодера, у которых хотя бы один воксель имеет значения больше нуля, именно в данных узлах располагаются вершины шестигранной сетки. Как видно, степень детализации изображения напрямую зависит от выбранного уровня октодера L , также от данного параметра зависит число узлов и ячеек построенной шестигранной сетки. Для уровня $L = 4$ число узлов сетки $N = 195$, а ячеек шестигранной формы — $M = 100$, при построении сетки на максимальном шестом уровне октодера число узлов возрастает до 57 980, а число шестигранников становится равным 52 074.

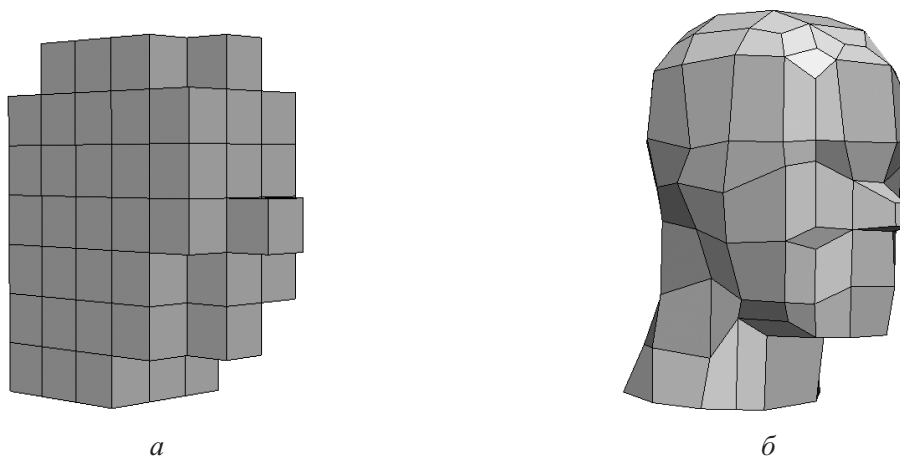


Рис. 3. $dim \times dim \times dim = 65 \times 65 \times 65$, уровень $L = 3$, число узлов сетки $N = 195$, ячеек — $M = 100$

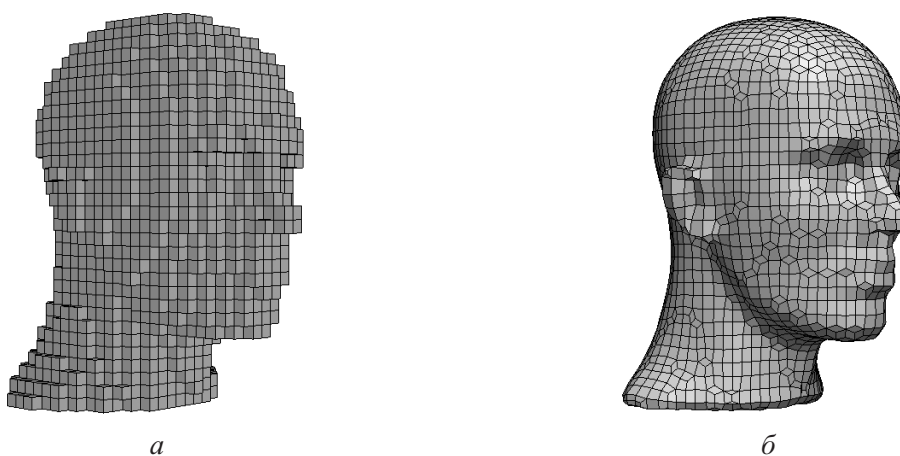


Рис. 4. $dim \times dim \times dim = 65 \times 65 \times 65$, уровень $L = 5$, число узлов сетки $N = 8073$, ячеек $M = 6561$

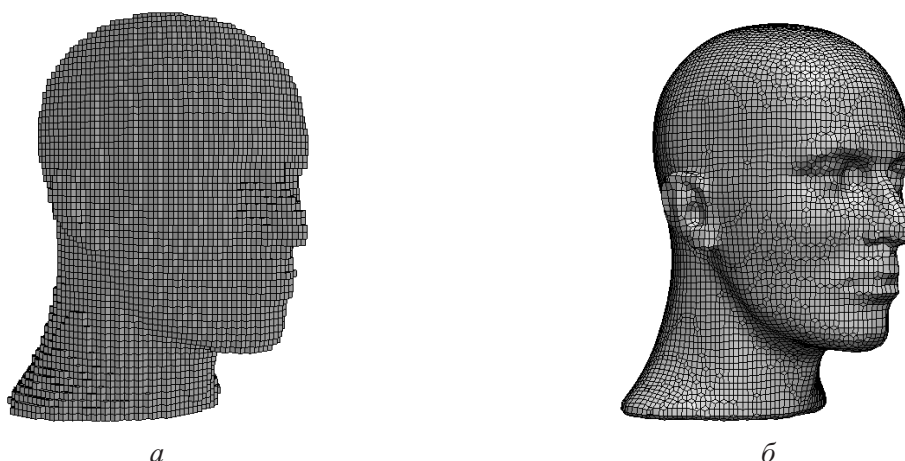


Рис. 5. $dim \times dim \times dim = 65 \times 65 \times 65$, уровень $L = 6$, число узлов сетки $N = 57980$, а ячеек $M = 52074$

Построение неструктурированной шестигранной сетки

Построение шестигранной сетки начинается с определения начального уровня октодеревя, в узлах которого будут вычисляться ее вершины. От выбора начального уровня октодеревя зависит степень детализации рассматриваемого объекта, а также число ячеек сетки. После этого для каждого узла октодеревя, у которого один и более вокселей имеет значение $F \geq 0$, вычисляется соответствующая вершина сетки. Далее найденные вершины восьми соседних узлов октодеревя соединяются, образуя шестигранник. Таким образом, механизм формирования шестигранных ячеек выглядит достаточно простым, однако остается задача нахождения координат вершин сетки.

В предлагаемом методе все вершины (узлы) сетки можно разделить на две группы: граничные — вершина находится внутри узла, имеющего воксели разного знака, и внутренние — вершина располагается в узле октодеревя, находящегося внутри рассматриваемого объекта. Внутренняя вершина сетки располагается в центре масс узла октодеревя и вычисляется как $p = (p_1 + p_7)/2$, где p_1, p_7 — координаты первого и седьмого вокселя рассматриваемого узла.

Для нахождения граничной вершины используется алгоритм двойственного контурирования. Данный метод для каждого граничного куба октодеревя генерирует вершину, удовлетворяющую минимуму квадратичной функции $E(\mathbf{x}) = \sum_i (\mathbf{n}_i \cdot (\mathbf{x} - \mathbf{x}_{p_i}))^2$, где пары $\mathbf{x}_{p_i}, \mathbf{n}_i$ — соответствующие точки пересечения ребер куба с изоповерхностью и единичные нормали данных точек. Функция $E(\mathbf{x})$ определяется в матричной форме $E(\mathbf{x}) = \mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x} - 2\mathbf{x}^T \mathbf{A}^T \mathbf{b} + \mathbf{b}^T \mathbf{b}$, где \mathbf{A} — матрица, составленная из нормалей точек пересечения \mathbf{n}_i , \mathbf{b} — вектор, элементы которого представлены в виде скалярного произведения $\mathbf{n}_i \cdot \mathbf{x}_{p_i}$. Минимум квадратичной функции находится решением переопределенной системы линейных уравнений $\mathbf{A} \mathbf{x} = \mathbf{b}$ алгоритмом Качмажа [Kaczmarz, 1993]. В случае если имеется множество решений минимума функции $E(\mathbf{x})$, она доопределяется так, чтобы вектор \mathbf{x} дополнительно минимизировал расстояние до центра масс точек c пересечения шестигранника с изоповерхностью. В таком случае выражение имеет вид $E'(\mathbf{x}) = E(\mathbf{x}) + (\mathbf{x} - \mathbf{x}_c)^2$. При этом к матрице добавляется одна строка, содержащая единицы, а к вектору \mathbf{b} добавляется сумма компонент центра масс.

Точки пересечения куба декартовой решетки с изоповерхностью определяются линейной интерполяцией по значениям соседних вокселей разного знака. Для координат вокселей \mathbf{x}_{p_1} и \mathbf{x}_{p_2} со скалярными значениями F_1, F_2 координаты точки пересечения \mathbf{x}_p вычисляются следующим

образом: $\mathbf{x}_p = \mathbf{x}_{p_1} - F_1 \cdot (\mathbf{x}_{p_2} - \mathbf{x}_{p_1}) / (F_2 - F_1)$. Нормаль в точке пересечения также вычисляется линейной интерполяцией по значениям нормалей вокселей на концах ребра, которому принадлежит данная точка.

Улучшение качества построенных ячеек

Одним из основных требований к расчетной сетке является невырожденность каждой ее ячейки. В данной работе условие невырожденности проверяется через положительность объемов десяти угловых тетраэдров, вписанных в шестигранник [Азаренок, 2008].

Первые пять тетраэдров изображены на рисунке 6, б – это четыре угловых тетраэдра при вершине ячейки, T_{1245} , T_{3427} , T_{6572} , T_{8754} , и внутренний – T_{2457} . Вторые пять состоят из четырех угловых тетраэдров T_{2316} , T_{4138} , T_{5861} , T_{7683} и одного внутреннего T_{1386} (рис. 6, в). Построенная

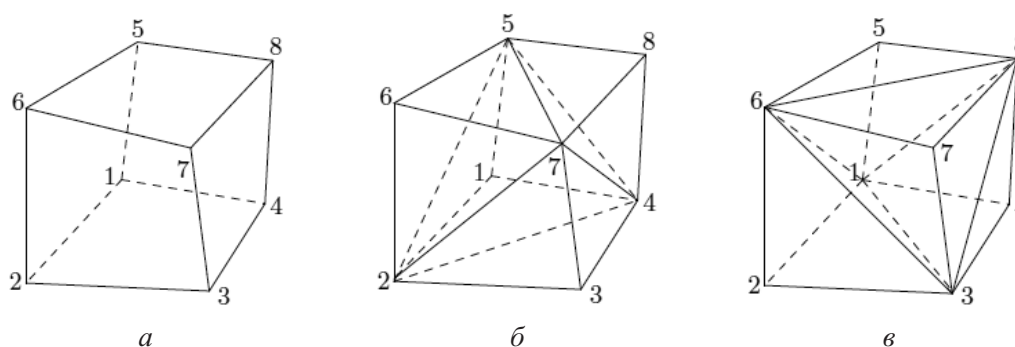


Рис. 6. Проверка невырожденности шестигранника

однородная шестигранный сетка может содержать вывернутые ячейки и в этом случае требуются дополнительные операции по ее оптимизации.

Улучшение качества построенных шестигранников осуществляется с помощью процедуры сглаживания по Лапласу [Canann et al., 1998], а также вставки дополнительного слоя граничных ячеек. Во время процедуры сглаживания координаты внутреннего узла \mathbf{x}_n сетки вычисляются следующим образом: $\mathbf{x}_n = \mathbf{x}_n + 1/k \sum_{i=1}^k (\mathbf{x}_{n_i} - \mathbf{x}_n)$, где \mathbf{x}_{n_i} – координаты соседних k узлов (с каждым из них сглаживаемый узел связан ребром). Новое положение граничного узла определяется движением по касательной плоскости без изменения формы граничного контура сетки.

При построении сетки могут возникать ситуации, когда ячейка содержит два и более четырехугольника на границе. В этом случае простая процедура сглаживания является малоэффективной. В работе был использован метод вставки дополнительного граничного слоя ячеек. В этом случае от каждого граничного четырехугольника создается новый шестигранник, что позволяет значительно улучшить форму граничных ячеек (рис. 7).

Адаптивное перестроение шестигранный сетки

При решении задач сеточными методами часто возникает необходимость уточнения полученного решения на выбранной подобласти или границ области. Это достигается алгоритмами перестроения начальной расчетной сетки. Необходимым условием такой операции является конформность сетки (согласованность узлов в соседних ячейках) и сохранение невырожденности ячеек после перестроения. В случае неструктурированной сетки особое внимание необходимо уделять выбору точек внутри перестраиваемых ячеек и обеспечению согласованного перехода между выбранными для перестроения ячейками и остальным множеством шестигранников.

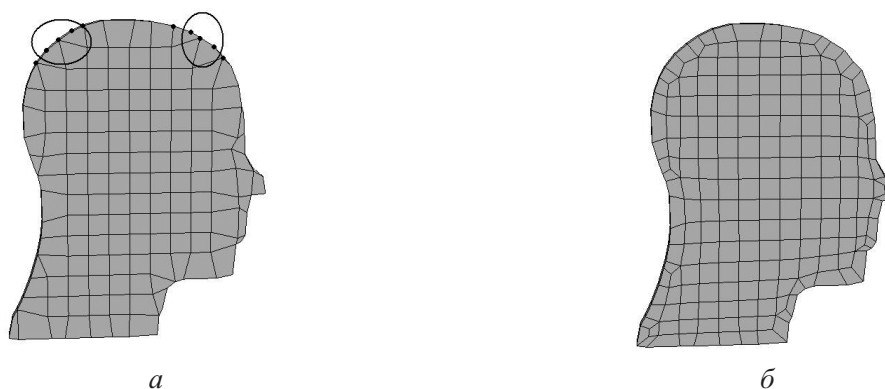


Рис. 7. Вставка буферного слоя ячеек

Основным способом перестроения шестигранной сетки являются шаблоны перестроения на 27 ячеек Шнейдера [Schneider, 1996]. Существенным преимуществом данного метода является независимость способа разбиения ячейки от ее соседей. В этом случае любой шестигранник, помеченный для перестроения, разбивается на 27 новых шестигранников, любой четырехугольник — на девять новых четырехугольников, любое ребро — на три новых ребра. На рисунке 8 показаны варианты разбиения шестигранника в зависимости от количества «помеченных» узлов, которые он имеет. Узел называется помеченным, если хотя бы один прилежащий к нему шестигранник должен быть перестроен.

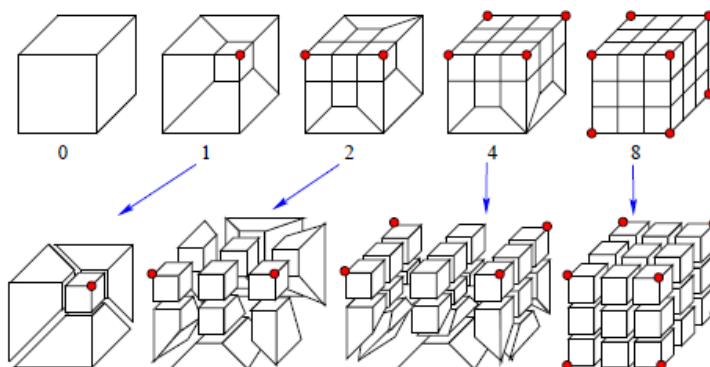


Рис. 8. Шаблоны перестроений

Рассматриваются четыре варианта расположения помеченных узлов для разбиения шестигранника, все остальные случаи конвертируются к рассматриваемым четырем шаблонам.

Уточнение границы осуществляется путем движения граничных узлов к точке пересечения с изоповерхностью. Точки пересечения более точно описывают границу тела, чем исходная вершина, полученная путем минимизации квадратичной функции. Движение граничных вершин осуществляется вдоль их нормалей, вычисленных трилинейной интерполяцией по значениям нормалей восьми окружающих вокселей (рис. 9).

Примеры построения расчетных сеток

В данном разделе представлены результаты работы алгоритма генерации расчетной шестигранной сетки из объемных (воксельных) данных. Исходные объемные данные получа-

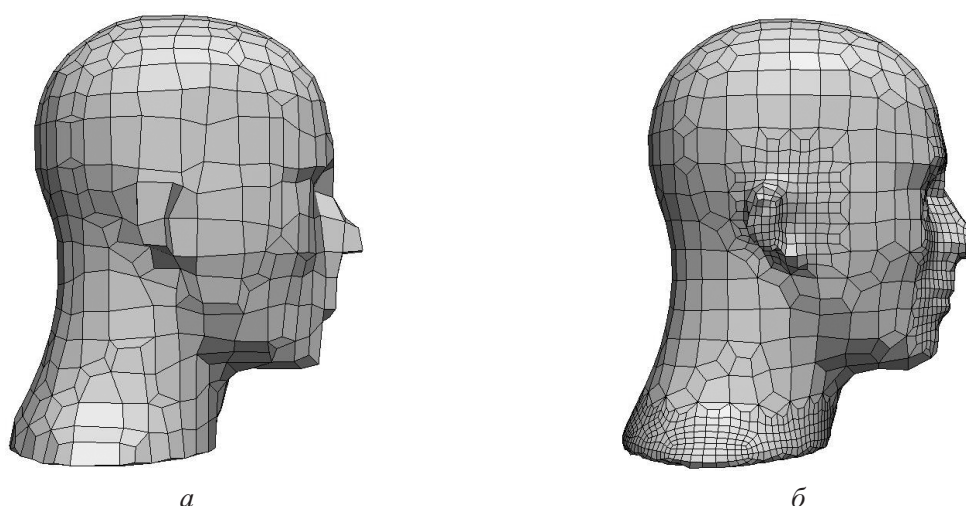
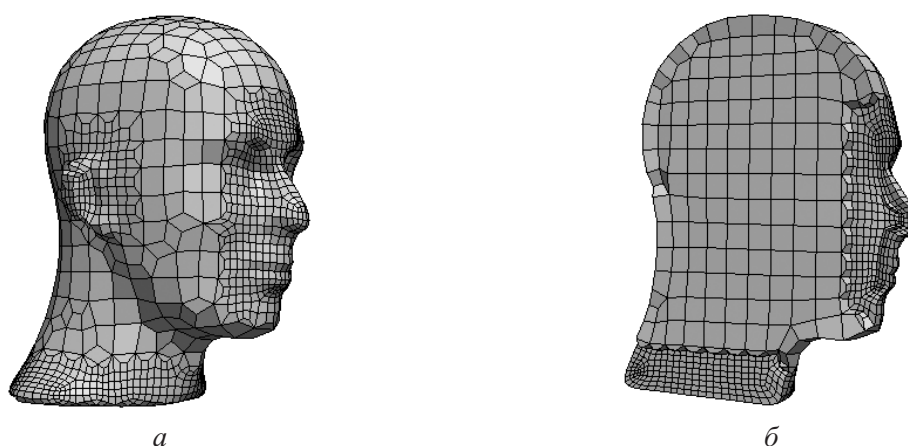


Рис. 9. Шестигранная сетка до (а) и после уточнение границы (б)

лись непосредственно из воксельного представления, а также путем конвертации начальной поверхностной сетки, описывающей геометрическую модель.

На рисунке 10 приведен пример сгенерированной шестигранной сетки. Размерность начальных составляла $65 \times 65 \times 65$, уровень октодеревя, выбранный для извлечения изоповерхности, а также генерации структурированной шестигранной сетки равнялся четырем. Сетка содержала $M = 814$ шестигранников и $N = 1\,209$ узлов, после адаптивного перестроения границы число шестигранников возросло до 11 980, число узлов стало равным 13 672. Каждая уточняемая ячейка сетки была разделена на 27 новых шестигранников. В таблице 1 приведены характеристики начальной однородной сетки, сетки после операций вставки граничного слоя ячеек и сглаживания, а также сетки после локального перестроения области.

Рис. 10. Область А. Шестигранная сетка построенная из файла объемных данных размерности $65 \times 65 \times 65$

На рисунке 11 представлен пример построения шестигранной сетки из данных, полученных путем преобразования поверхностной треугольной сетки, содержащей 255 треугольных ячеек. Объемные данные после преобразования содержали 65 вокселей вдоль каждой из осей, при этом для построения сетки был выбран пятый уровень октодеревя. Характеристики шестигранной сетки приведены в таблице 2. Все сетки, представленные в данной таблице, не содержат вырожденных ячеек.

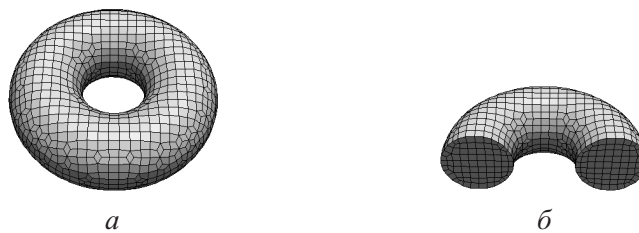


Рис. 11. Область Б. Сетка из шестигранных ячеек построена из поверхностной сетки, содержащей 225 треугольных ячеек

На рисунке 12 приведен пример сетки на области «В». Невырожденность ячеек сетки (см. таблицу 2) была обеспечена вставкой дополнительного слоя шестигранников вдоль границы цилиндрического выреза.

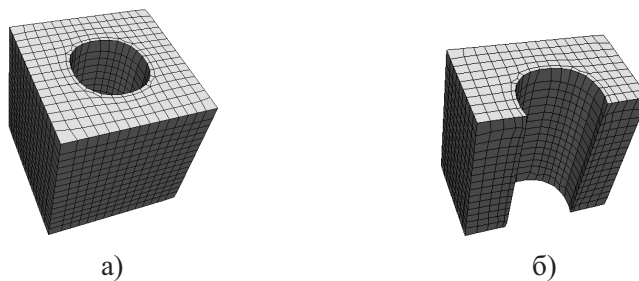


Рис. 12. Область В. Сетка построена из конвертированного файла формата .STL, содержащего 168 треугольных ячеек

Шестигранная сетка для области «Г» (Рис. 13) построена для размерности данных $65 \times 65 \times 65$, уровне октодера $L = 6$. В случае необходимости уточнения граничного контура потребуются исходные данные более высокого разрешения $129 \times 129 \times 129$, $257 \times 257 \times 257$ и т. д.

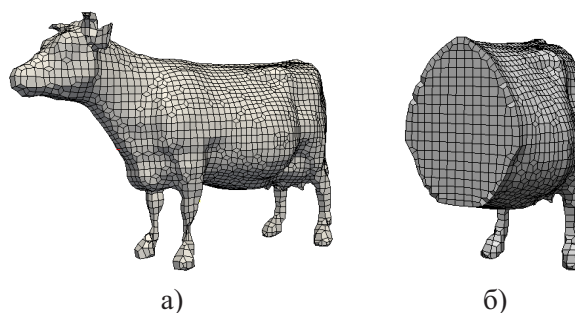


Рис. 13. Область Г. Сетка построена из конвертированного файла формата .STL, содержащего 5804 треугольных ячеек

Благодаря воксельной структуре данных метод построения сетки является достаточно простым в реализации и обладает небольшими вычислительными затратами. Программная реализация алгоритма осуществляется на облегченной структуре данных, что обеспечивает высокую интеграцию программы с приложениями в пакете FESstudio [Копысов и др., 2008].

Таблица 1. Характеристики шестигранных сеток, построенных на области А

Характеристика	Однородная сетка	Сетка после улучшения качества	Перестроенная сетка
Число узлов	1 209	1 915	13 672
Число ячеек	814	1 517	11 980
Макс. ребро	7.4	5.6	5.5
Мин. ребро	0.8	0.8	0.25
Аспектное соотношение	6.8	4.6	49.2
Макс. угол искажения четырехугольников	172.8°	85.6°	85.6°
Количество вырожденных шестигранников	55	0	0
<i>c</i> -валентность	8(501)	4(113), 6(150), 8(736), 10(208)	4(98), 6(1419), 8(7524), 10(516)

Таблица 2. Характеристики шестигранных сеток, построенных на областях Б, В и Г

Характеристика	Область Б	Область В	Область Г
Число узлов	4 720	4 042	14 774
Число ячеек	3 864	3 128	12 224
Макс. ребро	2.4	3.3	2.2
Мин. ребро	0.5	0.5	0.04
Аспектное соотношение	3.7	5.2	26.1
Макс. угол искажения четырехугольников	39.5°	76.8°	130.2°
<i>c</i> -валентность	4(216), 6(256), 8(2208), 10(448)	4(2), 6(60), 8(2128), 10(112)	4(538), 6(894), 8(7517), 10(1244)

Список литературы

- Азаренок Б. Н.* Вариационный метод построения гексаэдральных сеток с управляющей метрикой // Матем. моделирование. — 2008. — Т. 20. № 9. — С. 3–22.
- Караваев А. С., Копысов С. П., Пономарев А. Б.* Алгоритмы построения и перестроения неструктурированных четырехугольных сеток // Вычислительная механика сплошных сред. — 2012. — Т. 5. — С. 144–150.
- Копысов С. П., Новиков А. К., Пономарев А. Б., Рычков В. Н., Сагдеева Ю. А.* Программная среда построения расчетных моделей метода конечных элементов для параллельных распределенных вычислений // Информационные технологии. — 2008. — № 3. — С. 75–82.
- Копысов С. П., Пономарев А. Б.* Непрямой метод построения неструктурированных шестигранных сеток с преобладанием шестигранных элементов // В сб. «Актуальные проблемы, математики, механики, информатики». — Екатеринбург: ИММ УрО РАН, 2009. — С. 65–71.

- Сковпень А. В.* Реализация фронтального алгоритма построения нерегулярных четырехугольных сеток // ВАИТ. Серия: Математическое моделирование физических процессов. — 2005. — Вып. 1. — С. 9–30.
- Canann S. A., Tristano J. R. and Staten M. L.* An Approach to Combined Laplacian and Optimization-Based Smoothing for Triangular, Quadrilateral, and Quad-Dominant Meshes // 7th International Meshing Roundtable, Sandia National Labs. — 1998 — P. 479–494.
- Eppstein D.* Linear complexity hexahedral mesh generation // In Symposium on Computational Geometry. — 1996. — P. 58–67.
- Fuhrmann S.* Volume data generation from triangle meshes using the signed distance function // Darmstadt University of Technology. — Germany, 2007.
- Kaczmarz S.* Approximate solution of systems of linear equations // Internat. J. Control. — 1993. — Vol. 57, No. 6. — P. 1269–1271.
- Lorenson W., Cline H.* Marching Cubes: a high resolution 3D surface construction algorithm // Computer Graphics (SIGGRAPH 87 Proceedings). — 1987. — P. 163–169.
- Murdoch P., Benzley S. E., Blacker T. D., Mitchel S. A.* The spatial twist continuum: A connectivity based method for representing all-hexahedral finite element meshes // Finite Elements in Analysis and Design. — 1997. — No. 28. — P. 137–149.
- Owen S. J., Saigal S.* H-Morph: An Indirect Approach to Advancing Front Hex Meshing // International Journal for Numerical Methods in Engineering. — 2000. — P. 289–312.
- Owen S. J., Steven J., Staten M. L., Cannan S. A. and Saigal S.* Advancing Front Quadrilateral Meshing Using Triangle Transformations // 7th International Meshing Roundtable, Sandia National Labs. — 1998. — P. 409–428.
- Price M. A., Armstrong C. G.* Hexahedral mesh generation by medial surface subdivision: Part 1 // International Journal for Numerical Methods in Engineering. — 1995.
- Schneiders R., Schindler R., Weiler F.* Octree-based Generation of Hexahedral Element Meshes // 5th Annual International Meshing Roundtable. — 1996. — P. 205–216.
- Schaefer S. and Warren J.* Dual Contouring: The Secret Sauce // Technical Report 02-408, Department of Computer Science, Rice University. — 2002. — 6 p.
- Schneider R.* Refining quadrilateral and hexahedral meshes // Proc. 5th Int Conference on Numerical. Grid Generation In Computational Field. — 1996. — P. 679–688.
- Staten M. L., Owen S. J., Blacker T. D.* Unconstrained Paving and Plastering: A New Idea for All Hexahedral Mesh Generation // Proceedings, 14th International Meshing Roundtable. — Sandia National Laboratories:Springer-Verlag, 2005. — P. 399–416.
- Thompson J. F., Bharat K. S., Nigel P. W.* Hand book of grid generation. — New York: CRC Press, 1999. — 1200 p.

УДК: 517.977.5

Подсистема «Разработчик» системы приема коммунальных платежей

А. А. Клименко^а, Г. А. Угольницкий

Южный федеральный университет,
факультет механики, математики и компьютерных наук,
Россия, 344090, г. Ростов-на-Дону, ул. Мильчакова, 8а

E-mail: ^а antklim@gmail.com

Получено 20 июля 2012 г.

В работе рассматривается одна из ключевых подсистем приема коммунальных платежей «Разработчик». Описана разработанная система массового обслуживания, которая моделирует данную подсистему. Поставлена и решена задача о распределении ресурсов (в решении использовался модифицированный «венгерский» алгоритм). Приведено описание имитационной (агентной) модели данной подсистемы и результаты имитационных экспериментов.

Ключевые слова: система массового обслуживания, задача о распределении, модификация «венгерского» алгоритма, имитационное моделирование, агентная модель, имитационный эксперимент

Subsystem “Developer” as a part of the Retail Payment System

A. A. Klimenko, G. A. Ougolnitsky

Southern Federal University, Department of Applied Mathematics and Computer Science, 8a Milchakova str., Rostov on Don, 344090, Russia

Abstract. — In this paper we consider one of the core subsystems of the retail payment system named “Developer”. The Queuing System for modeling this subsystem was developed and information about it is provided. The task for the assignment problem was set up and solved (the modification of the Hungarian algorithm was used). Information about Agent Based Model for subsystem “Developer” and the results of the simulation experiments are given.

Keywords: queuing system, assignment problem, Hungarian algorithm, simulation modeling, agent based model

Citation: *Computer Research and Modeling*, 2013, vol. 5, no. 1, pp. 25–36 (Russian).

Введение

В последнее время бизнес теснее стал взаимодействовать с такой областью науки, как информационные технологии. Параллельно с внедрением информационных технологий в бизнес стала бурно развиваться область моделирования и прогнозирования бизнес-процессов, что определяет актуальность темы статьи.

Существует ряд работ, посвященных моделированию бизнес-процессов. Согласно [Vergidis et al., 2008], методы моделирования бизнес-процессов можно разделить на три типа: диаграммные модели, математические модели и языки моделирования бизнес-процессов. Первый тип – это простые диаграммы, изображающие бизнес-процесс, в большинстве случаев без использования стандартных обозначений [Havey, 2005]. Фалп и Шепперд [Phalp, Shepperd, 2000] отметили недостаток данного типа моделирования – анализ на основе использования этих моделей во многом зависит от квалификации аналитика. Второй тип соответствует моделям, имеющим математическое или иное формальное основание. Преимуществом таких моделей является то, что они могут быть проверены математически на предмет логичности и других свойств [Koubarakis, Plexousakis, 2002]. Недостатки формальных методов рассмотрены в работе [Hofacker, Vetschera, 2001]. Третий тип содержит искусственные языки, поддерживающие моделирование бизнес-процессов. К ним относятся BPMN [White, 2004], UML и UML2 [Джейсоул, 1973], XPDL [Елиферов, 2006], YAML [Aalst, Hofstede, 2003, Havey, 2005].

По теме имитационного моделирования известны работы Шрайбера Т. Дж. [Шрайбер, 1980], Борщева А. В. [Борщев, 2004], Карпова Ю. Г. [Карпов, 2005] и др.

В настоящей статье рассмотрена подсистема «Разработчик», которая входит в состав системы приема коммунальных платежей. Подсистема «Разработчик» является незаменимой, так как от нее зависит качество работы системы в целом. Для ее поддержания выделена отдельная группа программистов, численность которой может быть произвольной.

Система приема коммунальных платежей функционирует в ОАО КБ «Центр-Инвест» с 2008 года и представляет собой приложение с клиент-серверной архитектурой. В качестве СУБД используется IBM Informix, клиентская часть реализована в виде WEB-приложения на WEB-сервере Apache.

Целью работы является построение и исследование математических моделей подсистемы «Разработчик». Для достижения цели решены следующие задачи: формализация подсистемы «Разработчик», сбор и анализ данных о работе подсистемы, постановка задачи оптимизации, построение имитационной модели, выработка практических рекомендаций, направленных на улучшение производительности исследуемой подсистемы.

Анализ системы приема коммунальных платежей

Система приема коммунальных платежей работает в ОАО КБ «Центр-Инвест» на протяжении трех лет. Платежи принимаются в коммерческом банке. Пункты приема платежей сосредоточены по всему городу. Кассиры осуществляют прием наличности (оплата услуг связи, коммунальных услуг и т. п.) от клиентов круглосуточно. Согласно закону РФ о проведении банковских операций, деньги из кассы не могут быть напрямую перечислены поставщику услуг, требуется постобработка платежей, принятых кассиром. Поэтому система приема коммунальных платежей включает в себя модуль работников бэк-офиса, которые осуществляют постобработку платежей.

Ввиду сложности формализации процесса приема платежей и постобработки (связано с ограничениями, накладываемыми налоговым законодательством и нормативами ЦБ РФ) некоторые операции либо разрешение проблемных ситуаций проводятся вручную или с использованием специальных режимов с максимальным уровнем доступа. Все специальные режимы объединены в отдельный модуль администрирования.

При детальном рассмотрении общей системы приема и обработки коммунальных платежей в ее составе можно выделить подсистемы массового обслуживания, описывающие работу следующих процессов:

1. процесс приема платежа;
2. процесс обработки информации о платежах (например, процессы порождения платежных документов в централизованной базе данных платежей);
3. процесс поддержки пользователей при работе с прикладным программным обеспечением;
4. процесс разработки дополнительных модулей прикладной программы и оптимизации существующих.

Указанные подсистемы соответственно можно назвать «Кассир», «Центральная база данных», «Администратор» и «Разработчик». В данной статье подробно рассмотрена подсистема «Разработчик».

Система массового обслуживания «Разработчик»

Над системой работают от одного до N разработчиков. Задания, поступающие к разработчикам, характеризуются по нескольким критериям. Первый из них – тип задачи. Можно выделить три типа задач: исправление ошибок в существующем ПО, разработка новых функций, изменение существующих функций. В зависимости от типа задачи определяется приоритет ее выполнения.

С другой стороны, поступающие задачи различаются по степени их сложности либо объему работ. Объединение этих двух факторов обосновано тем, что единицей измерения работ являются человеко-часы, потраченные на задачу.

На основе вышеизложенного определим основные характеристики системы массового обслуживания, а затем определим, к какому типу она относится. Очевидно, что входным потоком здесь является поток заявок на разработку/доработку ПО. Обслуживающим прибором здесь выступает разработчик. Выходным потоком являются функционирующие части ПО, которые подверглись изменению. Пусть D – непустое множество приборов, $D = \{d_i, i = 1, \dots, n\}$. Каждый разработчик имеет накопитель заявок неограниченной емкости. Таким образом, возникает сеть массового обслуживания с N обслуживающими приборами.

Рассмотрим природу входного потока системы. Так как существует несколько типов задач, то имеем дело с неординарным входным потоком. Пусть в системе имеется $j = 1, \dots, H$ типов заявок. Так как интервалы времени между соседними заявками любого j -го класса являются случайными величинами, то получаем H случайных потоков. Пусть λ_{ij} — интенсивность потока заявок j -го типа в i -м обслуживающем приборе системы. Так как λ_{ij} не меняется во времени, то имеем дело со стационарным потоком. Для доказательства неизменности λ_{ij} были рассмотрены статистические данные, собранные за два года. Каждый год разобьем на 12 интервалов. В таблице 1 представлена интенсивность потока заявок в разные годы.

Вычислим среднюю интенсивность по каждому году. Получим значения 6 и 5,92.

Предположим, что интенсивность заявок неизменна, т. е. имеются незначительные отклонения от среднего значения при уровне значимости $\alpha = 0,05$. Для

Таблица 1. Интенсивность потока заявок в разные годы

Мес.	1-й год	2-й год
1	10	12
2	7	6
3	6	7
4	6	5
5	4	4
6	3	3
7	3	2
8	3	2
9	4	5
10	5	5
11	9	10
12	12	10

этого вычислим

$$\chi^2_{эм1} = \sum_{i=1}^{12} \frac{(n_{i1} - s_1)^2}{s_1} \approx 16,3, \quad \chi^2_{эм2} = \sum_{i=1}^{12} \frac{(n_{i2} - s_2)^2}{s_2} \approx 19,77,$$

где n_{i1} — интенсивность в i -м месяце первого года, n_{i2} — интенсивность в i -м месяце второго года, s_1 — среднее значение интенсивности первого года, s_2 — среднее значение интенсивности второго года.

По таблице критических распределений χ^2 по заданному уровню значимости $\alpha = 0,05$ и числу степеней свободы $k = 12 - 1$ находим критическую точку $\chi^2_{кр} = (0,05; 12) = 21,0$. Поскольку $\chi^2_{эм1} = 16,3 < 21,0 = \chi^2_{кр}$, $\chi^2_{эм2} = 19,77 < 21,0 = \chi^2_{кр}$, то нет оснований отвергать гипотезу об одинаковой интенсивности заявок.

Рассмотрим подробнее значения λ_{ij} для каждого месяца первого года и покажем, что они неизменны. Пусть $i, j = 1 \dots 3$. Рассмотрим 12 матриц размерности 3×3 :

$$\begin{bmatrix} 2 & 0 & 1 \\ 0 & 3 & 1 \\ 1 & 0 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \\ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix}, \begin{bmatrix} 3 & 1 & 0 \\ 0 & 4 & 1 \\ 0 & 0 & 3 \end{bmatrix}.$$

Вычислим матрицу средних значений λ_{ij} :

$$\begin{bmatrix} 1,67 & 0,83 & 0,08 \\ 0 & 2,08 & 0,17 \\ 0,08 & 0 & 1,83 \end{bmatrix}.$$

Снова воспользуемся критерием χ^2 для каждого λ_{ij} . Получим следующую матрицу эмпирических значений χ^2 :

$$\begin{bmatrix} 6,4 & 11 & 11,46 \\ 0 & 5,24 & 10 \\ 11,46 & 0 & 6,36 \end{bmatrix}.$$

Как видно, для заданного уровня значимости и количества степеней свободы ни для одного λ_{ij} эмпирическое значение χ^2 не превышает критическое. Следовательно, нет оснований отвергать гипотезу о неизменности интенсивности заявок λ_{ij} .

Эта система относится к системам без последствия, так как заявки поступают независимо друг от друга и момент поступления очередной заявки не зависит от поступления предыдущих заявок. По классификации дисциплин буферизации заявок во входном потоке, данный поток относится к беспriorитетным потокам. С точки зрения классификации дисциплин обслуживания имеем приоритетную систему одиночного типа, так как всякий раз на обслуживание назначается только одна задача. Следовательно, согласно обозначениям Кендалла получаем систему $\overline{M}_H / \overline{G}_H / N$. Характеристики по каждому классу заявок $j = 1, \dots, H$ идентичны СМО с однородным потоком. Подробное описание характеристик можно найти в [Алиев, 2009; Клейнрок, 1979; Willing, 1999].

Итак, определено, к какому типу систем массового обслуживания относится моделируемая система, и выявлены ее основные характеристики. Далее перейдем к постановке оптимизационной задачи для данной системы.

Оптимизация системы массового обслуживания «Разработчик»

Задачу оптимизации для данной системы можно сформулировать следующим образом: уменьшить время простоя разработчика, уменьшить время доработки. Ввиду использования в процессе разработки методологии экстремального программирования¹ долгосрочное планирование зачастую оказывается бесполезным. В данном случае наиболее подходящей является постановка задачи о назначениях.

Пусть x_1, x_2, \dots, x_n – разработчики, t_1, t_2, \dots, t_m — типы заданий, которые выполняются разработчиками, $f(x_i, t_j)$ – время выполнения i -м разработчиком j -го задания, $i = 1, \dots, n; j = 1, \dots, m$. Тогда необходимо найти допустимое назначение, у которого суммарное время выполнения работ будет минимальным. Наиболее эффективным методом решения данной задачи является «венгерский» алгоритм [Гинзбург, 2003]. Как известно, временная сложность оригинального алгоритма — $O(n^4)$. Однако лучший показатель временной сложности данного алгоритма – $O(n^3)$. Первый такой алгоритм описан в статье Диника и Кронрода (1969). Затем в начале 1970-х Эдмондс и Карп, а также Томидзава показали, что временная сложность может быть уменьшена до $O(n^3)$. В 1980-х были также представлены компьютерные реализации алгоритма (Буркард и Деригс, Джонкер и Волгенант, Карпането и др.) [Burkard, 2009]. Отличие приведенного ниже алгоритма заключается в том, что он позволяет решать задачи для прямоугольных матриц времени выполнения работ. В других модификациях, да и в оригинальном алгоритме рассматривается только квадратная матрица времени выполнения работ. Далее приведем решение поставленной задачи с использованием модифицированного «венгерского» алгоритма.

Сначала рассмотрим простой случай с n разработчиками, m задачами и значениями $n = m$. Построим исходную матрицу времени выполнения работ (таблица 2).

Таблица 2. Время выполнения работ. $C_{ij}, i = 1, \dots, n; j = 1, \dots, m$ – время выполнения i -м сотрудником j -го типа задания.

	1	2	3	...	m
1	C_{11}	C_{12}	C_{13}	$C_{1...}$	C_{1m}
2	C_{21}	C_{22}	C_{23}	$C_{2...}$	C_{2m}
...	$C_{...1}$	$C_{...2}$	$C_{...3}$		$C_{...m}$
N	C_{n1}	C_{n2}	C_{n3}	$C_{n...}$	C_{nm}

Так как в качестве решения необходимо знать, какой разработчик должен выполнять ту или иную работу, то решение можно записать в виде массива D :

$$D = \{d_1, d_2, \dots, d_n\}. \quad (1.1)$$

Здесь индекс d соответствует номеру работы, а значение d_j соответствует номеру разработчика, назначенного на j -ю работу. На каждом шаге алгоритма будем выбирать для каждой строки минимальное значение. Затем выберем наименьшее значение из минимумов по всем строкам. Если минимальные значения в разных строках совпадают, то берем значения, находящиеся в северо-западном углу матрицы (не всегда верно для матриц $2*2$). Записываем номер

¹ Экстремальное программирование (Extreme Programming, XP) — дисциплина разработки программного обеспечения и ведения бизнеса в области создания программных продуктов, которая фокусирует усилия обеих сторон (программистов и бизнесменов) на общих целях [Бек К., 2010].

строки минимального значения в соответствующую ячейку массива D . Удаляем строку и столбец, на пересечении которых находится найденный элемент. Повторяем шаг с новой матрицей.

Преимущество данного алгоритма над стандартным «венгерским» заключается в том, что на каждом шаге вычислительная сложность уменьшается, так как снижается размерность исходной матрицы, соответственно уменьшается количество элементов, подверженных обработке. Однако и данный алгоритм можно оптимизировать.

На первом шаге алгоритма для каждой строки матрицы можно выбрать минимальное значение и номер столбца, в котором находится выбранный элемент (на каждом шаге алгоритма значения элементов матрицы не меняются). Учтем замечания о выборе минимального элемента для матриц размерности 2×2 :

Шаг 1.

$$D = \underbrace{\{0, 0, \dots, 0\}}_m, \quad (1.2)$$

$$C = \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1m} \\ c_{21} & c_{22} & \dots & c_{2m} \\ & & \dots & \\ c_{n1} & c_{n2} & \dots & c_{nm} \end{bmatrix} \rightarrow M = \{M_1, M_2, \dots, M_n\}, \quad (1.3)$$

$$M_i = \left\{ \min(c_{i1}, c_{i2}, \dots, c_{im}) = \bar{m}, J = \{j : c_{ij} = \bar{m}\} \right\}. \quad (1.4)$$

Каждый массив M_i содержит минимальное значение и индексы столбцов, в которых они встречаются. Дальнейшая работа ведется с массивом M . Чтобы не зависеть от порядка следования элементов в массиве M и иметь возможность удалить элементы из массива, модифицируем структуру элемента массива (1.4). Добавим номер строки:

$$M_i = \left\{ i, \min(c_{i1}, c_{i2}, \dots, c_{im}) = \bar{m}, J = \{j : c_{ij} = \bar{m}\} \right\}. \quad (1.5)$$

Шаг 2.

Здесь и далее по тексту под элементом массива M будем понимать вложенный массив, который описывается формулой (1.5). Находим M_i с наименьшим \bar{m} и наименьшим значением i . Если элементов в массиве больше двух, то заносим i в массив D в ячейку с номером, совпадающим с наименьшим значением j , и обозначаем номер ячейки j^* . Затем M_i можно удалить из массива M , а все элементы со значением j , равным j^* , заменить на элемент с минимальным значением, полученный из столбцов матрицы с другими номерами. Перенумеровать оставшиеся элементы массива M с нуля и вернуться на начало шага 2.

На последней итерации, когда в массиве M останется два элемента, необходимо сделать проверку на пересечение значений j в оставшихся элементах и выбрать непересекающиеся.

Теперь рассмотрим случай $n > m$. Шаги 1 и 2 будут такими же, как и для равных значений. Отличие только в финальном шаге. На последней итерации в массиве M останется $n - m + 2$ элемента. На основе данных предыдущих итераций об обработанных i и j получаем итоговую матрицу размерности $(n - m + 2) \times 2$. Для каждой строки матрицы вычислим сумму первого элемента со вторыми элементами остальных строк. Индексы строк элементов, давших наименьшую сумму, заносятся в соответствующие позиции итогового массива D .

В случае $m > n$ алгоритм будет применен $\frac{m}{n} + 1$ раз. Первый шаг такой же, как и ранее. На втором шаге вычеркиваем самый левый столбец матрицы, в котором содержится наименьший элемент. Строку, в которой находился этот элемент, помечаем временно использованной. Далее выбираем минимальный элемент по оставшимся строкам матрицы. Возвращаемся на начало шага 2 и удаляем очередной столбец матрицы. Если останется только одна непомянутая как временно используемая строка матрицы, то выбираем в этой строке минимальный элемент m^* .

Удаляем из матрицы столбец с наименьшим номером, который содержит m^* . После этого снимаем пометки о временном использовании на строках и возвращаемся к началу шага 2. Данная итерация продолжается, пока не будут обработаны все столбцы матрицы.

Аналогично первым двум случаям исходную матрицу можно представить в виде массива:

$$C = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1m} \\ c_{21} & c_{22} & \cdots & c_{2m} \\ & & \cdots & \\ c_{n1} & c_{n2} & \cdots & c_{nm} \end{bmatrix} \rightarrow M = \{M_1, M_2, \dots, M_n\}. \quad (1.6)$$

В отличие от рассмотренных ранее алгоритмов, структура множества M_i будет иной:

$$M_i = \{i, u = \overline{0,1}, \min(c_{i1}, c_{i2}, \dots, c_{im}) = \overline{m}, J = \{j : c_{ij} = \overline{m}\}\}. \quad (1.7)$$

Здесь параметр u принимает значение 0, если множество M_i разрешено использовать, и 1, если M_i временно использовано.

Шаг 2 данного алгоритма отличается от рассмотренных для случаев $n = m, n > m$ тем, что множество M_i не удаляется из массива, а помечается как временно использованное. Данная пометка снимается со всех множеств, как только не найдено более свободных.

Таким образом, описан алгоритм для разового нахождения оптимального назначения. Однако более реалистичной является ситуация, когда разработку ведет группа специалистов и сама задача делится на подзадачи. Необходимо подобрать группу специалистов для максимально быстрого решения общей задачи. Иными словами, необходимо получить оптимальную последовательность задач о назначении.

Сначала рассмотрим случай, когда части основной задачи не зависят друг от друга и могут быть обработаны в любом порядке. Пусть имеется N разработчиков и M типов задач, и основная задача распадается на K подзадач:

$$T = \{C_k, k = \overline{1, K}\}, \quad (1.8)$$

где C_k представляют собой матрицы назначения.

Рассмотрим ситуацию $N = M = K = 3$. Тогда получаем три матрицы:

$$C_1 = \begin{bmatrix} c_{11}^1 & c_{12}^1 & c_{13}^1 \\ c_{21}^1 & c_{22}^1 & c_{23}^1 \\ c_{31}^1 & c_{32}^1 & c_{33}^1 \end{bmatrix} C_2 = \begin{bmatrix} c_{11}^2 & c_{12}^2 & c_{13}^2 \\ c_{21}^2 & c_{22}^2 & c_{23}^2 \\ c_{31}^2 & c_{32}^2 & c_{33}^2 \end{bmatrix} C_3 = \begin{bmatrix} c_{11}^3 & c_{12}^3 & c_{13}^3 \\ c_{21}^3 & c_{22}^3 & c_{23}^3 \\ c_{31}^3 & c_{32}^3 & c_{33}^3 \end{bmatrix}. \quad (1.9)$$

Применив первый шаг алгоритма, получим три массива для каждой из матриц:

$$M^1 = \{M_1^1, M_2^1, M_3^1\} M^2 = \{M_1^2, M_2^2, M_3^2\} M^3 = \{M_1^3, M_2^3, M_3^3\}. \quad (1.10)$$

Применяя алгоритм, получаем три массива решений:

$$D^1 = \{d_1^1, d_2^1, d_3^1\} D^2 = \{d_1^2, d_2^2, d_3^2\} D^3 = \{d_1^3, d_2^3, d_3^3\}. \quad (1.11)$$

Полученные результаты можно записать в виде матрицы. Такая запись даст возможность управлять последовательностью работ путем перестановок отдельных строк матрицы.

Таким образом, получено решение задачи о назначениях для задач сложного типа без зависимостей между подзадачами. Был рассмотрен случай для трех подзадач, но данный алгоритм действителен и для любого целого положительного K .

Далее усложним задачу и введем зависимости между подзадачами. Для описания последовательности работ построим орграф, вершинами которого являются подзадачи:

$$G = \{g_k, k = \overline{1, K}\}. \quad (1.12)$$

Пусть каждая вершина имеет вес, равный суммарному времени выполнения работ по подзадаче. Тогда решением задачи о назначениях с зависимыми подзадачами станет нахождение пути проходящего через все вершины графа. Суммарное время выполнения работ – сумма весов вершин графа.

В таблице 3 приведены результаты практической проверки времени работы алгоритма.

Таблица 3. Время работы алгоритма (в секундах) в зависимости от размерности матрицы. Каждое соотношение проверялось по 10 раз

	10	20	30	40	50	60	70	80	90	1600
1	0,016	0,004	0,008	0,005	0,013	0,018	0,007	0,008	0,014	9,327
2	0,004	0,004	0,005	0,005	0,006	0,006	0,007	0,009	0,009	9,442
3	0,010	0,004	0,004	0,007	0,005	0,006	0,007	0,007	0,005	9,426
4	0,004	0,004	0,11	0,005	0,012	0,008	0,012	0,014	0,018	9,282
5	0,003	0,004	0,004	0,005	0,006	0,011	0,007	0,011	0,012	9,412
6	0,005	0,004	0,005	0,005	0,012	0,012	0,009	0,010	0,019	9,434
7	0,005	0,004	0,006	0,004	0,005	0,006	0,016	0,015	0,015	9,382
8	0,004	0,004	0,005	0,005	0,010	0,007	0,009	0,012	0,009	9,533
9	0,004	0,006	0,005	0,011	0,005	0,006	0,010	0,008	0,017	9,541
10	0,004	0,006	0,005	0,005	0,006	0,006	0,008	0,008	0,013	9,714
Среднее	0,006	0,004	0,015	0,006	0,008	0,009	0,009	0,010	0,013	9,449

Алгоритм реализован на языке C, временная сложность алгоритма $O(n^3)$. Тесты проводились на ноутбуке Acer Aspire 5100. Следует также заметить, что значительную роль на увеличение времени обработки с увеличением размерности входной матрицы играет вывод результата. В таблице 3 приведено суммарное время работы (расчет и вывод).

Для примера работы алгоритма рассмотрим матрицу 4*4:

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \\ 1 & 2 & 3 & 4 \\ 4 & 3 & 3 & 4 \end{bmatrix}.$$

На первом шаге получаем массив M , содержащий четыре элемента (множества) M_0, M_1, M_2, M_3 :

$$M_0 = \{i = 0, \min = 1, J = [0]\};$$

$$M_1 = \{i = 1, \min = 1, J = [3]\};$$

$$M_2 = \{i = 2, \min = 1, J = [0]\};$$

$$M_3 = \{i = 3, \min = 3, J = [1, 2]\}.$$

Выбираем элемент M_0 массива M (так как минимальное значение по строке совпадает в трех строках матрицы M , то выбираем элемент с наименьшим номером строки i), результат заносим в массив D ($d_0 = j$). Далее работаем с матрицей 3*3, которую теперь описывает массив M , состоящий из трех элементов (элементы массива определяются по оставшимся строкам и столбцам матрицы M и нумеруются с нуля):

$$M_0 = \{i = 1, \min = 1, J = [3]\};$$

$$M_1 = \{i = 2, \min = 2, J = [1]\};$$

$$M_2 = \{i = 3, \min = 3, J = [1, 2]\}.$$

Так как M_0 снова имеет минимальное значение, то заносим его в массив D и продолжаем работать с матрицей 2×2 . Строим новый массив M (элементы массива определяются по оставшимся строкам и столбцам матрицы M и нумеруются с нуля):

$$M_0 = \{i = 2, \min = 2, J = [1]\};$$

$$M_1 = \{i = 3, \min = 3, J = [1, 2]\}.$$

Результаты обработки данного массива заносятся в массив $D = \{0, 2, 3, 1\}$. Таким образом, после данной итерации результирующий массив D содержит следующую информацию:

для выполнения задания 0 требуется сотрудник 0;

для выполнения задания 1 требуется сотрудник 2;

для выполнения задания 2 требуется сотрудник 3;

для выполнения задания 3 требуется сотрудник 1.

Итак, описана задача оптимизации, возникающая в подсистеме «Разработчик», и приведен метод ее решения.

Однако данный метод позволяет решать задачу статически, то есть ничего не известно о поведении данной системы во времени. Еще одним недостатком данного метода является невозможность определить взаимосвязь таких важных факторов, как время выполнения работ, сложность работы и стоимость исполнения работы. Добавление всех этих параметров к рассмотренной задаче увеличит в разы ее размерность, следовательно, и время расчетов. Время расчетов может быть настолько велико, что не уложится в приемлемые рамки. Решение такой задачи требует применения совершенно иных методов.

Имитационное моделирование подсистемы «Разработчик»

Одним из таких методов является имитационное моделирование, известными достоинствами которого являются гибкость и возможность моделирования систем на любом уровне детализации или абстракции. Первоначально имитационное моделирование определялось как дискретно-событийный способ описания и моделирования процессов, присущих системам массового обслуживания, системам с отказами и восстановлением элементов, дискретными производствами и т. д. [Бахвалов, 1997; Угольницкий, Тихонов, 2011; Шрайбер, 1980].

Наиболее подходящим инструментом для решения поставленной ранее задачи (распределение ресурсов во времени, описание связей различных факторов системы), является агентное моделирование².

Агентное моделирование является инструментом, при помощи которого возможно успешное моделирование сложных адаптивных систем. Агентное моделирование базируется на идее моделирования процессов «сверху-вниз»: в основе модели лежит набор основных элементов, из взаимодействия которых рождается обобщенное поведение системы. Агентами могут быть не только индивидуумы, но и группы индивидуумов, подсистемы и т. д. Применение данного подхода к моделированию наиболее удобно в случаях, когда представляют интерес характеристики поведения системы в целом, которые определяются как интегральные характеристики всей совокупности агентов [Борщев, 2004].

Данная модель позволит объединить решения, полученные на предыдущих этапах, для получения имитационной модели, максимально приближенной к реальности. С использованием формул, описывающих характеристики системы массового обслуживания, можно получить ограничения среды, в которой обитает агент. Решение задачи о назначениях, а также

² Агентное моделирование (agent-based modeling (ABM)) — метод имитационного моделирования, исследующий поведение децентрализованных агентов и то, как такое поведение определяет поведение всей системы в целом. При моделировании определяется поведение агентов на индивидуальном уровне, а глобальное поведение возникает как результат деятельности множества агентов.

модифицированный «венгерский» алгоритм использованы как составные части алгоритма, описывающего поведение агента.

Наиболее распространенным способом описания поведения агентов являются карты состояний (statecharts), которые представляют собой конечные автоматы с некоторыми дополнениями. Они описывают переходы между состояниями и события, инициирующие данные переходы, временные задержки и действия, совершаемые агентом на протяжении своей жизни. Агент может иметь несколько параллельно активных и взаимодействующих карт состояний, каждая из которых описывает какой-либо аспект его жизни [Arthur, 1991; Macal, North, 2005].

Для рассматриваемой подсистемы можно выделить следующих агентов: заявки на доработки, разработчики.

Простая схема работы заключается в том, что заявки попадают в систему, обслуживаются, задерживаются и покидают систему. Событие генерации очередной заявки будет соответствовать созданию нового агента. На рисунке 1 изображена схема жизненного цикла агента заявки.

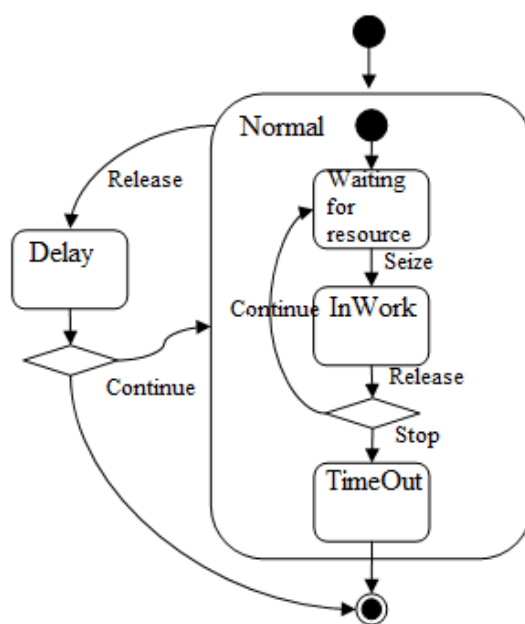


Рис. 1. Схема жизненного цикла агента заявки (см. пояснения в тексте)

Для каждой заявки можно выделить два основных состояния: в работе (на диаграмме — Normal) и отложено (на диаграмме — Delay).

Как только появляется новая заявка, она дожидается появления свободного ресурса для ее обработки. При появлении необходимого объема ресурсов заявка поступает в обработку. После обработки занятые ресурсы освобождаются, и заявка переходит на последний этап — верификация выполнения. В результате верификации определяется степень выполнения заявки и выявляется необходимость каких-либо доработок. В случае, когда необходимы доработки, заявка снова попадает в состояние ожидания свободных ресурсов. Из этого состояния заявка либо возвращается к рабочему циклу, либо завершается вовсе.

Поведение агентов, обрабатывающих заявки, можно описать тремя основными состояниями: свободен (Idle), занят (Busy), перерыв (TimeOut). Схема жизненного цикла агента обработчика заявок приведена на рисунке 2.

При появлении заявки обработчик переходит в состояние «занят» и выполняет работу. После завершения работ обработчик попадает в состояние «перерыв», так как моментально приступить к обработке новой заявки он не может.

Таким образом, завершён последний этап моделирования подсистемы «Разработчик». При помощи программного комплекса AnyLogic (данный ПК предоставляет возможности для

разработки моделей различных классов) были проведены имитационные эксперименты. Результаты имитационных экспериментов, проведенных с использованием построенной агентной модели, показали неэффективность существовавшего процесса разработки. Были даны рекомендации по разделению разработчиков на группы, специализирующиеся по видам деятельности: специализация на отчетах, клиентских интерфейсах и бизнес-логике. Помимо непосредственного процесса кодирования был взят в рассмотрение процесс проектирования как отдельный тип задач, соответственно проектировщики также рассматриваются как обслуживающий прибор.

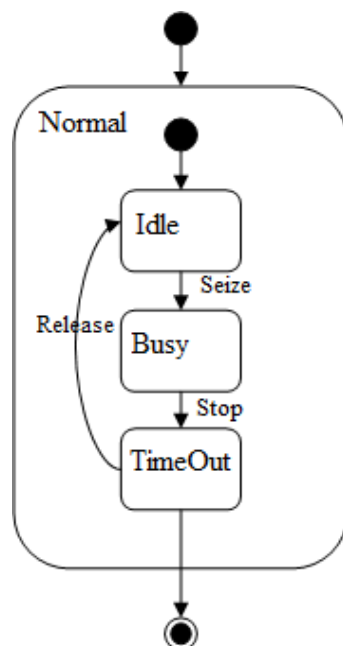


Рис. 2. Схема жизненного цикла агента обработчика заявок (см. пояснения в тексте)

Ранее эффективность оценивалась с использованием жестких планов: определенное число задач должно быть выполнено за некоторое время. Этот подход имел массу недостатков, в нем не учитывалась сложность задач и многие другие аспекты. Использование имитационных экспериментов позволило ввести гибкий график работ, предварительно рассчитывать необходимое для конкретной задачи время работы над ней и соответственно снизить количество часов переработок. Это напрямую отразилось на качестве продукта и моральном состоянии коллектива. При этом производительность труда увеличилась на десять процентов.

Список литературы

- Алиев Т. И. Основы моделирования дискретных систем // СПб.: СПбГУ ИТМО. — 2009. — С. 77–165.
- Бахвалов Л. Компьютерное моделирование — длинный путь к сияющим вершинам // Компьютерра. — 1997. — № 40 (217). — С. 26–36.
- Бек К. Экстремальное программирование // СПб.: Питер. — 2010.
- Борщев А. В. Практическое агентное моделирование и его место в арсенале аналитика // Exponenta Pro. — 2004. — № 3–4. — С. 38–47.
- Гинзбург А. И. Экономический анализ: Предмет и методы. Моделирование ситуаций. Оценка управленческих решений: учебное пособие // СПб.: Питер. — 2003. — 622 с.
- Джейсоул Н. К. Очереди с приоритетами / Пер с англ. М.: Мир. — 1973. — 279 с.

- Елиферов В. Г.* Управление качеством. Сказки, мифы и проза жизни // М.: Вершина. — 2006. — 296 с.
- Карпов Ю. Г.* Имитационное моделирование систем. Введение в моделирование с AnyLogic 5 / СПб.: БХВ-Петербург. — 2005. — 400 с.
- Клейнрок Л.* Теория массового обслуживания / Пер. с англ. Грушко И.И.; ред. Нейман В.И. — М.: Машиностроение. — 1979. — 432 с.
- Угольницкий Г., Тихонов С.* Имитационное моделирование бизнес-процессов: системы массового обслуживания / LAP Lambert Academic Publishing. — 2011. — 166 с.
- Шрайбер Т. Дж.* Моделирование на GPSS / Пер. с англ. Гаргера В. Л., Шмуйлович И. Л.; ред. Файнберг М. А. / М.: Машиностроение. — 1980. — 592 с.
- van der Aalst W. M. P., ter. Hofstede A. H. M.* YAWL: Yet another workflow language (revised version) // Queensland Univ. Technol., Brisbane, Australia: QUT Tech. Rep. - FIT-TR-2003-04. — 2003.
- Arthur, W. B.* Designing Economic Agents that Act Like Human Agents: A Behavioral Approach to Bounded Rationality // The American Economic Review. — 1991. — № 81 (2). — P. 353–359.
- Burkard, R. E.* Assignment problems/Rainer Burkhard, Mauro Dell’Amico, Silvano Martello // ISBN 978-0-898716-63-4. — Vol. 4. — P. 77.
- Havey M.* Essential business process modeling. — Sebastopol, CA: O’Reilly, 2005.
- Hofacker I., Vetschera R.* Algorithmical approaches to business process design // Comp. Oper. Res. — 2001. — Vol. 28. — P. 1253–1275.
- Koubarakis M., Plexousakis D.* A formal framework for business process modeling and design // Inf. Syst. — 2002. — Vol. 27. — P. 299–319.
- Macal C., North M.* Tutorial on agent based modeling and simulation / Proceedings of the 2005 Winter Simulation Conference, Center for Complex Adaptive Systems Simulations (CAS) // Argonne National Laboratory.
- Phalp K., Shepperd M.* Quantitative analysis of static models of processes // Syst. Software – 2000. — Vol. 52. — P. 105–112.
- Vergidis K. et al.* Business Process Analysis and Optimization: Beyond Reengineering // IEEE Transactions on Systems, Man, and Cybernetics. — 2008. — Part C. — Vol. 38, No. 1.
- Willing A.* A Short Introduction to Queuing Theory // Technical University Berlin. — July 21, 1999.
- White S.* Business Process Modeling Notation (BPMN): specification. — 2004. — Version 1.0. — May 3.

УДК: 51-72, 519.688

Численная идентификация модели дегидрирования в грид-системе на базе BOINC

И. А. Чернов^{1,a}, Е. Е. Ивашко^{1,b}, Н. Н. Никитина^{1,c}, И. Е. Габис^{2,d}

¹Институт прикладных математических исследований КарНЦ РАН,
Россия, 185910, г. Петрозаводск, ул. Пушкинская, 11

²Физический факультет СПбГУ,
Россия, 198504, г. Санкт-Петербург, Старый Петергоф, ул. Ульяновская, 3

E-mail: ^a IACHernov@yandex.ru, ^b ivashko@krc.karelia.ru, ^c nikitina@krc.karelia.ru, ^d igor.gabis@gmail.com

Получено 01 ноября 2012 г.

В работе рассматривается обратная задача определения по экспериментальным данным параметров модели выделения водорода из порошка гидрида металла. Методом слепого поиска в пространстве параметров установлено, что задача имеет многочисленные физически разумные решения. Решения задачи получены с помощью высокопроизводительного численного моделирования в грид-системе на базе платформы BOINC.

Ключевые слова: обратная задача, оценка параметров, математическое моделирование, вычислительные методы в физике, грид-системы, BOINC

Numerical identification of the dehydriding model in a BOINC-based grid system

I. A. Chernov¹, E. E. Ivashko¹, N. N. Nikitina¹, I. E. Gabis²

¹Institute of Applied Math Research, 11 Pushkinskaya street, Petrozavodsk, 185910, Russia

²Department of Physics, Saint-Petersburg State University, 3 Ulianovskaya Street, Petrodvorets, Saint-Petersburg, 198504, Russia

Abstract. — In the paper we consider the inverse problem of evaluating kinetic parameters of the model of dehydriding of metal powder using experimental data. The «blind search» in the space of parameters revealed multiple physically reasonable solutions. The solutions were obtained using high-performance computational modeling based on BOINC-grid.

Keywords: inverse problem, evaluation of parameters, mathematical modelling, computation methods in physics, grid systems, BOINC

Citation: *Computer Research and Modeling*, 2013, vol. 5, no. 1, pp. 37–45 (Russian).

Работа выполнена при поддержке гранта СПбГУ №11.38.68.2012.

Введение

Обратные задачи определения параметров математических моделей по измеряемым данным играют важную роль в физике. Способность модели воспроизводить наблюдаемые зависимости является важнейшим критерием ее адекватности. Модель, как правило, содержит несколько параметров, имеющих определенный физический смысл. Нелинейное взаимодействие различных элементарных процессов существенно обесценивает методы суперпозиции, призванные определить характеристики элементарных процессов по отдельности. Адекватная математическая модель эксперимента аппроксимирует не только отдельный экспериментальный результат, но целую серию, причем при близких значениях параметров (однако допустимый разброс для разных параметров может быть различным). Удачная модель опирается на описания элементарных процессов и общие законы природы (например законы сохранения) и потому содержит более одного параметра.

Общих методов решения обратных задач не существует. Обратные задачи, как правило, нелинейны даже для линейных моделей (уже потому, что решение линейного дифференциального уравнения в общем случае нелинейно) и некорректны [Aster et al, 2005; Тихонов, Арсенин, 1979]. Существующие методы обеспечивают в лучшем случае локальную сходимость, что оставляет открытым вопрос об альтернативных решениях. Различные попытки обеспечить единственность решения требуют, по существу, знания решения априори.

Критерием качества найденного решения является способность полученного набора параметров аппроксимировать серию экспериментальных кривых, а также физическая разумность, то есть соответствие ограничениям, налагаемым физическим смыслом. К очевидным ограничениям относится положительность величин, оценки сверху на параметры, имеющие смысл вероятности или доли, и т. п. Существуют и иные ограничения на параметры, связанные с физическими оценками. В ряде случаев оценки «мягкие» (например, если известно, что некоторая скорость существенно меньше скорости звука, но более точные оценки отсутствуют), приводящие к проверке маловероятных потенциальных решений.

Однако и в области допустимых значений параметров возможно несколько решений обратной задачи. Для определенности будем считать, что речь идет о минимизации среднеквадратичной невязки. Если параметров не слишком много, имеет смысл «слепой» поиск по сетке в области допустимых значений. Узлы, в которых обнаружены локально-минимальные значения функционала, могут служить начальными данными для тех или иных итерационных алгоритмов.

Описанная задача принадлежит к классу проблем, эффективно решаемых в рамках грид-систем и систем распределенных вычислений [Foster, Kesselman, 1999; Foster, Kesselman, Tuecke, 2001]:

1. отдельные подзадачи (расчеты модели в узлах сетки в пространстве параметров и вычисление среднеквадратичной невязки) никак не связаны между собой, могут решаться параллельно и не требуют синхронизации;
2. подзадачи выполняются быстро и не требовательны к вычислительной мощности;
3. объем входных данных невелик (вектор пространства параметров, массив экспериментальных значений пересылается однократно), а выходные данные — одно вещественное число (значение функционала), что не предъявляет высоких требований к скорости линий связи.

В настоящей работе мы рассматриваем слепой поиск решений задачи идентификации модели разложения гидрида алюминия методом наименьших квадратов в грид-системе. Гидриды металлов являются одним из перспективных способов хранения водорода, обеспечивая относительно высокую емкость водородных аккумуляторов при достаточно низком равновесном

давлении; кроме того, эндогенность реакции разложения гидрида приводит к снижению температуры при дегидрировании, что способствует пожаробезопасности [Тарасов, Лотоцкий, Яртысь, 2006]. Однако необходимо детальное понимание процессов формирования и разложения гидридов, в частности оценки кинетических констант.

В дальнейшем речь идет о выделении водорода из порошка гидрида алюминия [Gabis et al, 2012]. Алюминий является перспективным материалом в силу своего малого атомного веса и относительно высокого содержания водорода в гидриде (алане). Работа с порошком обусловлена особенностями синтеза гидрида алюминия методами мокрой химии. Одним из недостатков алана является длительный период инкубации — формирования зародышей фазы металла, служащих каналом облегченной десорбции водорода и, таким образом, обеспечивающих быстрое выделение водорода. Однако этот период может быть сокращен путем активации. Одним из способов активации является облучение порошка ультрафиолетовым светом. При длительной экспозиции можно упрощенно считать, что часть S_0 поверхности частиц, подвергшаяся облучению, соответствует тонкому слою фазы металла; при короткой экспозиции и перемешивании с целью облучить всю поверхность частиц считаем, что образуются зародыши новой фазы, занимающие некоторую долю поверхности S_0 .

Математическая модель

Экспериментальные данные по неактивированному алану получены методом барометрии. Небольшие количества (доли грамма) порошка $\alpha - AlH_3$ помещались в автоклав и вакуумировались. Затем автоклав соединялся с термостатированной камерой с калиброванным объемом и нагревался до заданной температуры. Нагревался только автоклав, объем которого пренебрежимо мал по отношению к объему камеры. Количество выделившегося водорода определялось по изменению давления в калиброванном объеме.

Активация ультрафиолетовым (УФ) светом проводилась как для выяснения деталей физических процессов, идущих во время инкубации дегидрирования, так и для сокращения ее длительности [Gabis et al, 2012]. В первой серии экспериментов алан, насыпанный тонким слоем на термостол в вакуумной камере, облучался через прозрачное для УФ окно. При этом частицы порошка алана засвечивались с одной стороны и могли экранировать нижние частицы. После окончания экспозиции УФ выключался, термостол нагревался до заданной температуры и происходило выделение водорода. Подробности постановки эксперимента можно найти в статье [Габис и др., 2012]

Во второй серии экспериментов алан при активации УФ помещался на вибрирующий стол, что позволило облучать его частицы со всех сторон. Эта процедура проводилась на атмосфере, затем образец засыпался в автоклав, и проводилось выделение водорода. СЭМ фотографии образцов, дегидрированных на 15% (см. [Gabis et al, 2012]) показывают наличие большого числа металлических зародышей, равномерно распределенных по поверхности частиц алана.

Методика и выбранный подход к математическому моделированию формирования и разложения гидридов достаточно подробно описаны в работе [Chernov, Gabis, 2011] (см. также библиографию там). Там же обосновано утверждение о малом влиянии формы частиц на кинетику и о допустимости одночастичного приближения — замены порошка одной частицей простой формы. Кратко опишем модельные допущения:

- частицы порошка сферические, рассматриваем одну частицу радиуса L ;
- предполагаем сценарий «сжимающегося ядра» — формирование слоя металла (корки) на поверхности частицы и его дальнейшее утолщение;

- сплошная корка занимает сферический слой, ограниченный парой концентрических сфер с радиусами L и $\rho < L$;
- формирующаяся корка состоит из зародышей симметричной формы и описывается в сферических координатах (r, φ, θ) декартовым произведением $\Omega \times [\rho, L]$, где Ω — произвольное измеримое множество на поверхности сферы; отношение его площади к поверхности частицы обозначим S .

При этих предположениях состояние частицы полностью описывается в каждый момент времени величинами ρ и S .

Физические допущения следующие: диффузию в фазе металла считаем быстрой в том смысле, что градиенты концентрации пренебрежимо малы; температура образца поддерживается постоянной; пренебрегаем реадсорбцией, то есть десорбция водорода идет в вакуум; десорбция с фазы металла идет непосредственно с границы раздела фаз (через поры); концентрация водорода c в фазе металла постоянна и мала по сравнению с принимаемой за единицу концентрацией в гидриде; размер частицы не меняется в ходе реакции. Пренебрегая концентрацией c , считаем поток десорбции, тем не менее, ненулевой величиной.

Плотности потоков десорбции постоянны, обозначим их J' для гидрида и J для металла. Полный поток десорбции тогда равен

$$J'(1 - S(t)) \cdot 4\pi L^2 + JS(t)4\pi\rho^2.$$

Учет радиального роста формируемой корки сводится к предположению, что доля ν потока десорбции с фазы металла обусловлена распадом ядра и, тем самым, приводит к движению границы ρ . Это движение описывается уравнением

$$\dot{\rho} = -\nu J, \quad \rho > 0. \quad (1)$$

Запись закона сохранения (полное количество водорода меняется только за счет десорбции с поверхности гидрида $4\pi L^2(1 - S)$ и поверхности металла $4\pi\rho^2 S$) приводит к уравнению

$$\dot{S}(t) \frac{L^3 - \rho^3}{3} = J'(1 - S(t))L^2 + (1 - \nu)JS(t)\rho^2. \quad (2)$$

Предполагаем, что $\nu = 1$, если $S = 1$, то есть при сформированной корке вся десорбция идет с ядра. При $S < 1$ возможно любое $\nu \in [0, 1]$. Отсутствие радиального роста означает $\nu = 0$.

Измеряемой в эксперименте величиной была доля прореагировавшего вещества, то есть относительный объем фазы металла. В модели это

$$f(t) = (1 - \check{\rho}^3(t))S(t), \quad \check{\rho}(t) = \rho(t)/L.$$

Модель (1)–(2) содержит следующие пять параметров: плотности потоков десорбции с двух фаз J и J' ; начальные данные для толщины корки ρ_0 и площади, занимаемой зародышами на начало десорбции S_0 (причем эта величина известна — равна нулю — в случае неактивированной десорбции); доля потока за счет радиального роста зародышей ν . Все параметры неотрицательны, кроме того, $S_0 \leq 1$, $\rho_0 \leq L$, $\nu \leq 1$. Таким образом, рассмотрение по N значений каждого параметра приводит к необходимости решения N^5 однотипных независимых задач. Кроме того, при вычислении f в эксперименте используется полное количество водорода, содержащиеся в частице изначально, неизвестное точно; в связи с этим экспериментальная кривая

может быть умножена на масштабный множитель s , ограниченный сверху и снизу (в пределах 1 ± 0.1). При расчете величина s выбиралась из условия минимизации среднеквадратичной невязки при неизменной модельной кривой $f(t)$. Таким образом, минимизируемый функционал

$$F = \min_{s \in [0.9, 1.1]} \int_0^{\bar{t}} (f(t) - s\hat{f}(t))^2 dt \cdot \left(\bar{t} \cdot \max_{t \in [0, \bar{t}]} \hat{f}(t) \right)^{-1},$$

$\hat{f}(t)$ — экспериментальные данные. Численно интеграл заменяется формулой трапеций с интерполяцией модельных значений. Нормировка обеспечивает безразмерность функционала и независимость от интервала времени \bar{t} (десорбция без активации протекает существенно дольше, а время эксперимента сильно варьируется в зависимости от температуры). Далее значения F выражены в процентах.

Уравнения модели несложны и могут быть проинтегрированы в элементарных функциях, однако оценки показывают, что численное интегрирование уравнений и вычисление функционала эффективнее расчета по сложной формуле; кроме того, такой подход сильно упрощает модификацию программного обеспечения в случае уточнения или изменения модели. Численный алгоритм типа предиктор-корректор реализован на языке Fortran-90/95 без использования нестандартных возможностей и библиотек, что обеспечивает работоспособность для большинства платформ.

Организация вычислительных экспериментов

Численное моделирование результатов экспериментов проводилось на локальной грид-системе, реализованной в ЦКП КарНЦ РАН «Центр высокопроизводительной обработки данных» [ЦКП КарНЦ РАН, 2012] на базе программной платформы BOINC (Berkeley Open Infrastructure for Network Computing) [Anderson, 2004]. BOINC является одной из наиболее популярных и активно развиваемых в настоящее время систем промежуточного программного обеспечения с открытым исходным кодом (LGPL) для организации проектов добровольных вычислений в сети Интернет, а также грид-систем в рамках локальной сети. Данные системы предполагают объединение в единую структуру неспециализированных вычислительных единиц — в частности персональных компьютеров, принадлежащих организации или отдельным лицам. При этом сравнительно небольшая мощность отдельных вычислительных узлов компенсируется их большим количеством и наличием высокоскоростных каналов связи между ними, таких как локальная коммуникационная сеть организации и Интернет.

Платформа BOINC позволяет выполнять в подобной распределенной вычислительной среде ресурсоемкие приложения, которые могут быть разделены на большое количество независимых подзадач (рабочих заданий). При этом не требуется, чтобы приложения имели открытый исходный код или были специальным образом адаптированы. BOINC имеет архитектуру «клиент-сервер»: сервер, или управляющий узел, генерирует рабочие задания и помещает их в очередь. Узлы-клиенты запрашивают новые задания при наличии на узле свободных ресурсов (процессорное время, оперативная и дисковая память), выполняют вычисления и отсылают результаты расчетов на сервер.

Грид-системы на базе платформы BOINC могут использоваться как для проведения ресурсоемких вычислительных экспериментов (см., например, [Ivashko, Nikitina, 2011]), так и для высокопроизводительной обработки данных (см. [Ivashko, Golovin, 2012]).

В качестве управляющего узла грид-сегмента КарНЦ РАН используется выделенный сервер. Для проведения вычислительных экспериментов клиентская часть была установлена на один настольный компьютер, два сервера, один управляющий и три вычислительных узла кластера

ЦКП КарНЦ РАН (на каждом узле кластера — два четырехъядерных процессора). Таким образом, при условии, что вычислительные узлы не заняты выполнением каких-либо программ и предоставляют в пользование BOINC все свои ресурсы, возможен одновременный расчет 43 вычислительных заданий. Использование BOINC-грид для выполнения серий однотипных экспериментов позволяет гибко управлять приоритетами задач, выполняемых на вычислительном кластере. При этом накладные расходы, вносимые грид-сетью, невелики, что позволяет достаточно эффективно задействовать ресурсы кластера.

Для выполнения расчетов использовалось однопоточное Linux-приложение, разработанное на языке Fortran-90/95. Так как приложение не адаптировалось для выполнения в BOINC-грид, для его запуска использовалось специальное приложение-обертка BOINC Wrapper [GenWrapper], организующее взаимодействие исходного приложения с BOINC-клиентом.

Результатом расчета является значение сеточного среднеквадратичного функционала; если оно не превосходит заданного порога, результат считается допустимым решением.

Показатели эффективности для двух описанных выше экспериментов — сравнение активированного и неактивированного разложения гидроксида алюминия (эксперимент № 1) и дегидрирование при пяти значениях постоянной температуры (эксперимент № 2) приведены в таблице 1. Для второго эксперимента общее число заданий составило 250 000, однако, поскольку в грид-системе наиболее эффективно выполняются вычислительные задачи, время счета которых существенно превышает время на передачу данных, каждое из 250 «пакетных» рабочих заданий последовательно решало 1 000 задач. Высокая точность расчета модели для первого эксперимента связана с большим количеством экспериментальных точек и призвана снизить влияние интерполяции.

Таблица 1. Показатели эффективности расчета

Характеристика	№ 1	№ 2
Число заданий	15 000	250
Число узлов сетки по времени	10^5	10^2
Общее время расчета на грид	6 ч. 13 мин.	66 мин.
Среднее время расчета рабочего задания	25.43 с.	12 мин. 3 с.
Время выполнения эксперимента на одном ПК	79 ч. 10 мин.	26 ч.

Результаты расчетов и обсуждение

Область параметров была ограничена исходя из физических оценок на значения параметров; пороговые значения определялись на основе эталонного решения, найденного стандартным методом наименьших квадратов.

Для первого эксперимента эталонный набор параметров: $\bar{J} = 1.025 \cdot 10^{-7}$ см/с, $\bar{J}' = 3.593 \cdot 10^{-9}$ см/с, $\bar{\rho}_0 = 0.82L$, $\bar{S}_0 = 0$ для случая без активации и $\bar{\rho}_0 = 0.85L$, $\bar{S}_0 = 0.09$ для активированной десорбции, $\bar{\nu} = 0$. Для плотностей потоков проверялось пять значений: \bar{J} , $(1 \pm 0.25)\bar{J}$ и $(1 \pm 0.5)\bar{J}$, аналогично для J' . Для ρ_0 — тоже пять: 0.5, 0.6, 0.7, $\bar{\rho}_0$, 0.9. Для S_0 — шесть: 0, 1/24, 1/12, 0.09, 1/8, 1/6, а для параметра ν , наименее известного, — десять: $\nu = 0.1i$, $i = 1 : 10$. Итого сетка состоит из 7 500 узлов, а расчет для двух случаев — активированного и неактивированного — приводит к решению 15 000 задач.

Для второго эксперимента эталонные значения \bar{J} и \bar{J}' выбраны для каждой температуры (поскольку скорости реакций существенно зависят от температуры). Сетки значений для всех

параметров — равномерные. Для плотностей потоков ρ_0 и ν проверялось десять значений: от $0.2\bar{J}$ до $2\bar{J}$, аналогично для J' , от 0.8 до 0.98 для ρ_0 и от 0 до 0.9 для ν . Для S_0 — пять: от 0 до 0.04. Итого сетка состоит из 50 000 узлов, а расчет для пяти экспериментальных кривых для разных температур приводит к решению 250 000 задач.

Опишем первый эксперимент. Пороговое значение F , при которых набор параметров считался допустимым решением, выбрано 6% без активации и 4% при активации. Из 15 тысяч кривых тест прошли 74, в том числе 9 кривых для неактивированной десорбции. Таким образом, мы имеем множество возможных решений задачи. Уменьшение их количества введением более жестких условий отбора не оправдано, поскольку при этом потеряются потенциальные решения, которые выглядят приемлемо — в отсутствие других решений они вполне удовлетворительны с физической точки зрения. Однако можно предложить некоторые дополнительные ограничения: потребовать радиальный рост ($\nu > 0$), для кривых с активацией необходимо $S_0 > 0$ (это результат активации), ограничить толщину корки снизу ($\rho_0 \geq 0.9L$), ограничить масштабный множитель ($s \geq 1$, для активированных кривых $s < 1.05$). при этом остается 9 решений (таблица 2), в том числе 4 (№№ 1–4) — для неактивированной десорбции.

Таблица 2. Оценки параметров для первого эксперимента

№	$10^7 J$	$10^9 J'$	ρ_0/L	ν	S_0	s
1	1.281	2.695	0.9	0.2	0.000	1.092
2	1.025	3.593	0.9	0.2	0.000	1.070
3	1.025	1.797	0.9	0.1	0.000	1.084
4	0.769	3.593	0.9	0.1	0.000	1.051
5	1.025	4.491	0.9	0.2	0.083	1.018
6	1.025	3.593	0.9	0.2	0.090	1.011
7	1.025	5.389	0.9	0.1	0.042	1.041
8	1.281	3.593	0.9	0.2	0.042	1.031
9	1.025	4.491	0.9	0.1	0.042	1.017

Пара (2,6) является решением наилучшего качества в том смысле, что соответствующие параметры аппроксимируют две экспериментальные кривые при одних и тех же значениях параметров (различие S_0 отражает влияние активации, а масштабный множитель s , вообще говоря, свой для каждого измерения). Однако пары (1,8) и (2,9) также представляют интерес (константа десорбции с поверхности гидрида играет роль лишь на начальной стадии процесса и определяется хуже).

Для второго эксперимента решений еще больше: пороговое значение функционала в 6% преодолевается более чем в 57 000 случаях, поэтому порог был понижен до 1.6%: он преодолевается для 2 598 наборов параметров из 250 000. Введение дополнительных ограничений ($S_0 > 0$, $\nu > 0$) позволяет уменьшить число решений; их распределение по температурам эксперимента приведено в таблице 3 (N — число решений, а $\min F$ — минимальное значение функционала).

Отметим, что понижение порога возможно только по завершении расчета, поскольку при пороговом значении 1.5% решений для некоторых температур не обнаруживается. Оценка 6% была получена на основе поиска решения для некоторых начальных значений параметров стандартным методом наименьших квадратов — таким образом, слепой поиск позволил обнаружить решения существенно лучшего качества, в общем случае недоступные локально сходящимся методам, если окрестность решения не известна априори. Число обнаруженных решений снижа-

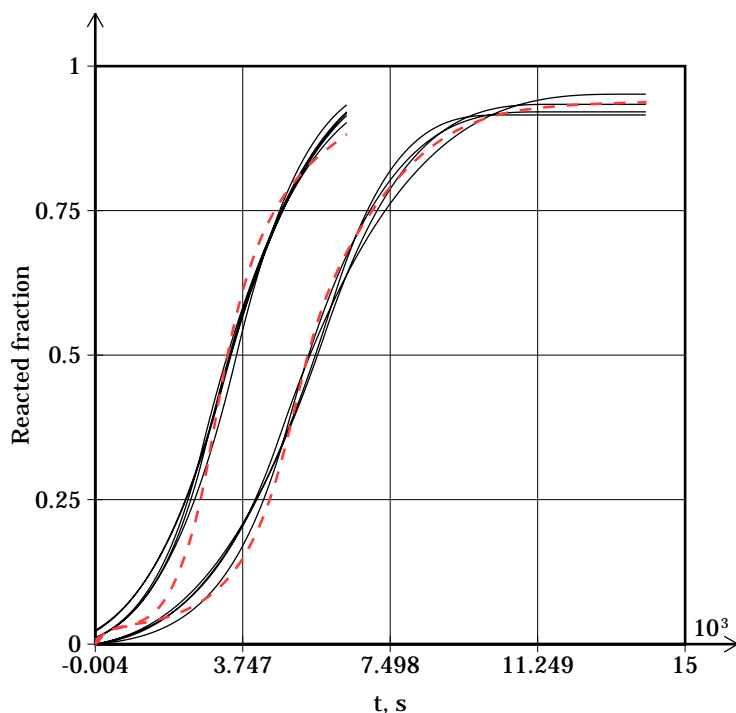


Рис. 1. Возможные решения задачи идентификации модели выделения водорода из активированного и неактивированного порошка гидрида алюминия; пунктирные линии — экспериментальные, сплошные — модельные аппроксимации.

Таблица 3. Распределение полученных решений по различным экспериментальным кривым

$T^{\circ}\text{C}$	80	90	100	110	120
N	1,173	750	2	7	2
$\min F$	0.9 %	1.1 %	1.6 %	1.6 %	1.2 %

ется при увеличении точности, но пропорционально растет и время расчета. Анализ полученных решений позволяет сделать ряд выводов:

- константа скорости десорбции с поверхности гидрида J' определяется однозначно: во всех полученных наборах ее значение постоянно, кроме кривых для $T = 80^{\circ}\text{C}$, где присутствует вдвое большее значение, однако в конце списка (при $F = 1.6\%$);
- константа скорости десорбции с поверхности металла J определяется достаточно точно: разброс значений — в пределах 20 %;
- аппроксимация лучше при малых значениях S_0 , точнее, при $S_0 \leq 0.02$;
- скорость радиального роста зародышей ν имеет тенденцию к росту по отношению к температуре: для 120°C $\nu \leq 0.2$ для всех найденных решений (при $\min F \leq 6\%$), для 100°C с уже $\nu < 0.5$, а при 80°C $\nu > 0.6$, причем значение 0.9 — в начале списка;
- начальная толщина корки ρ_0 слабо зависит от температуры: $\rho_0 \approx 0.95 \pm 0.01$.

Сделанные выводы и оценки были использованы при получении результатов, изложенных в работе [Gabis et al, 2012].

Заключение

В работе описано решение задачи оценивания параметров математической модели выделения водорода из порошка гидрида металла путем минимизации среднеквадратичного функционала методом слепого поиска на сетке в пространстве параметров с помощью грид-системы на базе BOINC. Описана схема организации вычислений с учетом специфики задачи (большое количество малосвязанных небольших подзадач), получены оценки параметров и качественные выводы. Использование BOINC-грид позволило существенно сократить время, необходимое для проведения вычислительных экспериментов. Показано, что задача имеет большое количество «физически разумных» решений, причем введение дополнительных ограничений возможно лишь после завершения расчетов.

Список литературы

- Габис И. Е., Елец Д. И., Кузнецов В. Г., Барабан А. П., Добротворский М. А., Добротворский А. М. Термо- и фотоактивация разложения гидрида алюминия // Ж-л физ. химии. — 2012. — Т. 11. — С. 1864–1870.
- Ивашко Е. Е., Головин А. С. Методы DATA MINING для анализа больших массивов данных в гетерогенной грид на базе BOINC. Труды международной суперкомпьютерной конференции «Научный сервис в сети Интернет: поиск новых решений». — 2012. — С. 196–199.
- Никитина Н. Н., Ивашко Е. Е. Организация квантовохимических расчетов с использованием пакета Firefly в гетерогенной грид-системе на базе BOINC // Вычислительные методы и программирование. Новые вычислительные технологии. — 2012. — Т. 13. — С. 8–12.
- Тарасов Б. П., Лотоцкий М. В., Яртысь В. А. Проблема хранения водорода и перспективы использования гидридов для аккумулирования водорода // Российский химический журнал. — 2006. — Т. 1, № 6. — С. 34–48.
- Тихонов А. Н., Арсенин В. Я. Методы решения некорректных задач. изд. 2-е. — М.: Наука, 1979. — 285 с.
- Центр высокопроизводительной обработки данных ЦКП КарНЦ РАН / Институт прикладных математических исследований Карельского научного центра РАН. 2012. URL: <http://cluster.krc.karelia.ru> (дата обращения: 1.11.2012).
- Anderson D. P. BOINC: A System for Public-Resource Computing and Storage // 5th IEEE/ACM International Workshop on Grid Computing. Pittsburgh, PA. — 2004. — P. 365–372.
- Aster R. C., Borchers B., Thurber C. H. Parameter Estimation and Inverse Problems. — Elsevier, 2005. — 303 p.
- Chernov I. A., Gabis I. E. Mathematical Modelling of Hydride Formation // Mathematical Modeling, Clustering Algorithms and Applications / Caroline L. Wilson. Mathematics Research Developments series. Nova publishers. — 2011. — P. 203–246.
- Foster I., Kesselman C. The Grid: Blueprint for a New Computing Infrastructure // Morgan Kaufmann, San Francisco. — 1999. — 675 p.
- Foster I., Kesselman C., Tuecke S. The Anatomy of the Grid: Enabling Scalable Virtual Organizations // International J. Supercomputer Applications. — 2001. — V. 15, N. 3. — P. 200–222.
- Gabis I. E., Voyt A. P., Chernov I. A., Kuznetsov V. G., Baraban A. P., Elets D. I., Dobrotvorsky M. A. Ultraviolet activation of thermal decomposition of α -alane // International Journal of Hydrogen Energy. — 2012. — V. 37. — P. 14405–14412.
- GenWrapper. A generic BOINC wrapper for legacy applications. URL: <http://genwrapper.sourceforge.net> (дата обращения: 1.11.2012).

УДК: 577.2; 51-76

Моделирование белок-белковых взаимодействий с применением программного комплекса многочастичной броуновской динамики *ProKSim*

С.С.Хрущев^а, А.М.Абатурова, А.Н.Дьяконова, Д.М.Устинин,
Д.В.Зленко, В.А.Федоров, И.Б.Коваленко^б, Г.Ю.Ризниченко, А.Б.Рубин

Московский государственный университет имени М. В. Ломоносова,
Биологический факультет, кафедра биофизики,
Россия, 119992, г. Москва, Ленинские горы, 1-12

E-mail: ^а styx@biophys.msu.ru, ^б ikovalenko78@gmail.com

Получено 5 декабря 2012 г.,
после доработки 20 февраля 2013 г.

Белок-белковые взаимодействия являются основой большинства биологических процессов. Компьютерное моделирование динамики связывания белков дает важную информацию для понимания механизмов их функционирования. Разработана компьютерная программа *ProKSim* (**P**rotein **K**inetics **S**imulator), предназначенная для моделирования взаимодействия макромолекул методом многочастичной броуновской динамики с учетом дальнедействующих электростатических взаимодействий. Проведено исследование диффузионно-столкновительных комплексов для трех пар белков: ферредоксин и ферредоксин:НАДФ⁺-редуктаза, пластоцианин и цитохром *f*, барназа и барстар. Исследована роль электростатических взаимодействий во взаимной ориентации молекул белков при образовании диффузионно-столкновительных комплексов.

Ключевые слова: многочастичная броуновская динамика, белок-белковые взаимодействия, механизмы молекулярного распознавания

Multi-particle Brownian Dynamics software *ProKSim* for protein-protein interactions modeling

S. S. Khrushchev, A. M. Abaturova, A. N. Diakonova, D. M. Ustinin,
D. V. Zlenko, V. A. Fedorov, I. B. Kovalenko, G. Yu. Riznichenko, A. B. Rubin

*Lomonosov Moscow State University, School of Biology, Biophysics Department
Leninskie Gory 1-12, 119992, Moscow, Russia*

Abstract. — Protein-protein interactions are of central importance for virtually every process in living matter. Modeling the dynamics of protein association is crucial for understanding their functionality. This paper proposes novel simulation software *ProKSim* (**P**rotein **K**inetics **S**imulator) for modeling of protein interactions by means of the multi-particle Brownian Dynamics. Effect of long-range electrostatic interactions on the process of transient encounter complex formation is numerically estimated. Investigation of transient encounter complex formation was performed for three pairs of proteins: ferredoxin and ferredoxin:NADP⁺-reductase, plastocyanin and cytochrome *f*, barnase and barstar.

Keywords: multi-particle Brownian Dynamics, protein-protein recognition, molecular recognition

Citation: *Computer Research and Modeling*, 2013, vol. 5, no. 1, pp. 47–64 (Russian).

Работа поддержана грантами РФФИ 11-04-01268-а, 12-07-33036-мол_а_вед и 12-04-31839-мол_а. Работа выполнена с использованием ресурсов суперкомпьютерного комплекса МГУ имени М.В. Ломоносова [Воеводин и др., 2012].

© 2012 Сергей Сергеевич Хрущев, Анна Михайловна Абатурова, Александра Никитична Дьяконова, Дмитрий Михайлович Устинин, Дмитрий Владимирович Зленко, Владимир Андреевич Федоров, Илья Борисович Коваленко, Галина Юрьевна Ризниченко, Андрей Борисович Рубин

1. Введение

Белок–белковые взаимодействия лежат в основе многих физиологических процессов, связанных с ферментативной активностью и ее регуляцией, электронным транспортом и др. Предсказание структуры белковых комплексов и скорости их образования путем компьютерного моделирования представляет собой сложную, до сих пор не решенную до конца проблему. Причиной этого является, в первую очередь, очень высокая вычислительная ресурсоемкость методов моделирования «из первых принципов», поэтому их применение на нынешнем этапе возможно лишь для самых простых систем (десятки атомов). Для более крупных систем приходится использовать те или иные эмпирические или полуэмпирические приближения, имеющие ограниченную область применимости. К числу наиболее часто используемых при исследовании взаимодействия биомакромолекул методов относятся молекулярный докинг и молекулярная динамика.

Процесс образования комплекса двух белковых молекул в растворе можно условно разделить на несколько стадий:

- 1) свободная диффузия молекул в растворе на большом расстоянии от других макромолекул,
- 2) сближение макромолекул и их взаимная ориентация за счет дальнедействующих электростатических взаимодействий с образованием предварительного (диффузионно-столкновительного) комплекса и
- 3) трансформация предварительного комплекса в финальный, т. е. в такую конфигурацию, в которой осуществляется биологическая функция.

Альтернативно диффузионно-столкновительный комплекс может распасться без образования финального комплекса. При трансформации предварительного комплекса в финальный происходят вытеснение молекул растворителя из белок-белкового интерфейса и конформационные изменения самих макромолекул. Важную роль в этом процессе играют гидрофобные взаимодействия и образование водородных связей и солевых мостиков. Различные авторы вкладывают разный смысл в понятие диффузионно-столкновительного (*encounter*) комплекса; обзор нескольких популярных точек зрения приведен в [Gabdoulline, Wade, 1999]. В настоящей работе используется буквальное понимание диффузионно-столкновительного комплекса как ансамбля структур, образующихся в процессе диффузии, в которых макромолекулы сближены друг с другом.

Разработана компьютерная программа *ProKSim* (Protein Kinetics Simulator), предназначенная для моделирования взаимодействия макромолекул в растворе и в клетке методом многочастичной броуновской динамики (БД) с учетом электростатических взаимодействий [Kovalenko et al., 2006; Kovalenko et al., 2011; Коваленко и др., 2009]. В модели молекулы белка рассматриваются как броуновские частицы, совершающие поступательное и вращательное движения в вязкой среде. Конформационная подвижность молекул не учитывается (молекулы рассматриваются как твердые тела), не рассматриваются также гидрофобные взаимодействия белков. В результате компьютерного моделирования может быть получена статистическая информация об образующихся диффузионно-столкновительных комплексах. Для заданного критерия образования белок-белковых комплексов могут быть получены кинетические параметры реакции — зависимость количества (концентрации) комплексов от времени и, следовательно, константа скорости реакции образования комплексов.

Структуры диффузионно-столкновительных комплексов, получаемые с помощью *ProKSim*, могут быть в дальнейшем исследованы другими методами, в частности, с помощью методов молекулярной динамики и квантовой химии, что позволит провести полную реконструкцию процесса взаимодействия белков.

2. Метод броуновской динамики

Метод броуновской динамики является разновидностью метода стохастической динамики и в применении к химической кинетике был развит в работе [Ermaк, McCammon, 1978]. В методе рассматривается решение стохастического уравнения движения (уравнения Ланжевена),

которое для поступательного движения имеет вид

$$m\dot{\mathbf{r}} = -\xi_{\text{tr}}\dot{\mathbf{r}} + \mathbf{F}(\mathbf{r}) + \mathbf{f}(t), \quad (1)$$

где m — масса частицы, \mathbf{r} — положение частицы, t — время, ξ_{tr} — коэффициент вязкого трения частицы в среде, $\mathbf{F}(\mathbf{r})$ — внешняя и $\mathbf{f}(t)$ — случайная силы. Случайная сила $\mathbf{f}(t)$ неявно учитывает тепловое движение молекул среды и их столкновения с макромолекулой и представляет собой дельта-коррелированную случайную величину (белый шум), математическое ожидание которой равно нулю, а автокорреляционная функция равна

$$\langle \mathbf{f}(t) \mathbf{f}(t') \rangle = 2kT\xi_{\text{tr}}\delta\delta(t-t'),$$

где k — постоянная Больцмана, T — температура, δ — символ Кронекера и $\delta(t)$ — дельта-функция Дирака.

Для получения численной схемы мы выполняем проекцию на каждую из координатных осей и почленно усредняем по времени уравнение (1). Для небольших белков в водной среде при нормальных условиях постоянная времени $\frac{m}{\xi} \approx 10^{-12}$ с. Использование шага по времени Δt ,

значительно превышающего $\frac{m}{\xi}$, позволяет пренебречь инерциальным членом ($\langle m\dot{\mathbf{r}} \rangle_{\Delta t} \approx 0$). Для смещений Δx вдоль каждой из осей координат получаем формулу

$$\Delta x = \frac{F_x \Delta t}{\xi_{\text{tr}}^x} + \sqrt{\frac{2kT\Delta t}{\xi_{\text{tr}}^x}} N(0,1),$$

где F_x — проекция внешней силы на соответствующую ось, ξ_{tr}^x — коэффициент вязкого трения для движения вдоль оси, $N(0,1)$ — нормально распределенная случайная величина с нулевым средним и единичной дисперсией.

Аналогичным образом выводятся уравнения для вращательных степеней свободы:

$$\Delta\alpha_x = \frac{M_x \Delta t}{\xi_{\text{rot}}^x} + \sqrt{\frac{2kT\Delta t}{\xi_{\text{rot}}^x}} N(0,1),$$

где $\Delta\alpha_x$, M_x и ξ_{rot}^x — угол поворота, момент внешней силы и коэффициент вязкого трения для вращения относительно оси, вокруг которой рассматривается вращение, соответственно.

Шаг по времени является параметром модели и по умолчанию равен 10^{-10} с. В качестве генератора псевдослучайных чисел используется алгоритм, предложенный [Panneton, L'Ecuyer, 2005].

3. Начальные и граничные условия

Модельная сцена представляет собой трехмерный реакционный объем, который может быть разделен на несколько компартментов. Каждый компартмент имеет форму прямоугольного параллелепипеда. Трехмерная модель молекул строится по данным из Protein Data Bank (PDB) [Berman et al., 2007]. Перед началом моделирования заданное количество молекул каждого белка распределяются определенным (по умолчанию случайным) образом в пределах отведенного им компартмента. Границы компартментов в зависимости от параметров модели могут быть либо периодическими (в случае столкновения с границей молекула проходит через нее, появляясь со стороны противоположной границы), либо неупругими (в случае столкновения молекула вплотную подходит к границе компартмента, с которой происходит столкновение).

4. Взаимодействие макромолекул с растворителем и между собой

Метод броуновской динамики учитывает взаимодействие частиц с молекулами растворителя за счет введения случайной силы и силы трения в уравнения движения. Взаимодействие между отдельными броуновскими частицами рассматривается в явном виде. Электростатические взаимодействия являются дальнедействующими и влияют на скорость сближения и взаимную ориентацию макромолекул в процессе диффузии. Контактные взаимодействия проявляются при сближении частиц на расстояние порядка суммы ван-дер-Ваальсовых радиусов входящих в них атомов и препятствуют взаимному проникновению молекул.

4.1. Гидродинамические взаимодействия

В модели растворитель рассматривается неявным образом как непрерывная среда с вязким трением; значение вязкости является параметром модели и по умолчанию принимается равным $10^{-3} \text{ кг}\cdot\text{м}^{-1}\text{с}^{-1}$, что соответствует вязкости воды при температуре 293 К. Гидродинамические взаимодействия между макромолекулами не учитываются.

В гидродинамических расчетах для оценки коэффициентов вязкого трения молекула белка рассматривается как удлинённый эллипсоид вращения с однородной по объёму плотностью и массой, равной массе молекулы. Центр масс эллипсоида совпадает с центром масс молекулы. Ориентация эллипсоида выбирается таким образом, чтобы его ось вращения совпадала по направлению с осью минимальной инерции молекулы; направления двух других осей выбираются произвольно в плоскости, перпендикулярной оси вращения. Определение необходимого направления оси вращения определяется методом последовательного перебора. Длины осей выбираются таким образом, чтобы момент инерции эллипсоида относительно оси вращения был равен минимальному моменту инерции молекулы. Для каждой молекулы вводится собственная ортогональная система координат, связанная с осями соответствующего эллипсоида вращения и в которой рассматривается движение макромолекулы. Размеры эллипсоида используются для вычисления коэффициентов вязкого трения макромолекулы в растворителе [Perrin, 1936]:

$$\xi_{\text{tr}}^a = \frac{8\pi\eta R(p^2 - 1)^{\frac{3}{2}}}{p^{\frac{1}{3}} \left((2p^2 - 1) \ln \left(p + (p^2 - 1)^{\frac{1}{2}} \right) - p(p^2 - 1)^{\frac{1}{2}} \right)},$$

$$\xi_{\text{tr}}^b = \frac{16\pi\eta R(p^2 - 1)^{\frac{3}{2}}}{p^{\frac{1}{3}} \left((2p^2 - 3) \ln \left(p + (p^2 - 1)^{\frac{1}{2}} \right) + p(p^2 - 1)^{\frac{1}{2}} \right)},$$

$$\xi_{\text{rot}}^a = \frac{16\pi\eta R^3 (p^2 - 1)^{\frac{3}{2}}}{3p \left(-\ln \left(p + (p^2 - 1)^{\frac{1}{2}} \right) + p(p^2 - 1)^{\frac{1}{2}} \right)},$$

$$\xi_{\text{rot}}^b = \frac{16\pi\eta R^3 (p^2 - 1)^{\frac{3}{2}} (p^2 + 1)}{3p \left((2p^2 - 1) \ln \left(p + (p^2 - 1)^{\frac{1}{2}} \right) - p(p^2 - 1)^{\frac{1}{2}} \right)},$$

где ξ_{tr}^a и ξ_{tr}^b — коэффициенты вязкого трения при движении вдоль (a) и перпендикулярно (b) осям вращения эллипсоида, ξ_{rot}^a и ξ_{rot}^b — соответствующие коэффициенты вязкого трения для вращательного движения, $R = (ab^2)^{\frac{1}{3}}$ — характерный размер эллипсоида вращения, $p = \frac{a}{b} > 1$ — отношение длин осей эллипсоида вращения (a и b), η — динамическая вязкость среды.

4.2. Электростатические взаимодействия

Для расчета электростатических взаимодействий используется предложенный [Ermak, McSammon, 1978] приближенный метод. Диэлектрические свойства макромолекул и растворителя, а также экранирующее влияние мобильных зарядов (ионов) учитываются с помощью уравнения Пуассона–Больцмана

$$\varepsilon_0 \nabla [\varepsilon(\mathbf{r}) \nabla \varphi(\mathbf{r})] = -\rho - \sum_i c_i^\infty z_i q_p e^{\left[\frac{z_i q_p \varphi(\mathbf{r})}{kT} \right]}, \quad (2)$$

где ε_0 — электрическая постоянная, $\varepsilon(\mathbf{r})$ — диэлектрическая проницаемость среды, $\varphi(\mathbf{r})$ — электростатический потенциал, ρ — объемная плотность фиксированных электрических зарядов (в молекуле белка), c_i^∞ — концентрация ионов в отсутствие электростатического поля, z_i — заряд иона, q_p — заряд протона, \sum_i обозначает суммирование по всем типам мобильных зарядов (ионов) в растворе [Князева и др., 2010].

Полагая $|z_i q_p \varphi(\mathbf{r})| < kT$, мы можем линеаризовать уравнение (2) подстановкой $e^{\left[\frac{z_i q_p \varphi(\mathbf{r})}{kT} \right]} \approx 1 - \frac{z_i q_p \varphi(\mathbf{r})}{kT}$:

$$\varepsilon_0 \nabla [\varepsilon(\mathbf{r}) \nabla \varphi(\mathbf{r})] = -\rho + \sum_i \frac{c_i^\infty z_i^2 q_p^2 \varphi(\mathbf{r})}{kT} - \sum_i c_i^\infty z_i q_p.$$

$\sum_i c_i^\infty z_i q_p = 0$ из соображений электронейтральности раствора.

Расчет потенциала электростатического поля проводится на кубической сетке по итерационной формуле [Ullmann, Knapp, 1999]

$$\varphi_0 = \frac{q_0 + \sum_j h \varepsilon_0 \varepsilon_j \varphi_j}{h^3 \kappa_0^2 + \sum_j h \varepsilon_0 \varepsilon_j}, \quad (3)$$

где φ_0 , φ_i — электростатический потенциал в ячейке, q_0 — суммарный электрический заряд в ячейке, ε_i — диэлектрическая проницаемость среды в ячейке, h — размер ячейки,

$\kappa_0 = \sqrt{\frac{2N_A q_p^2 I}{kT}}$ — обратная дебаевская длина для ионной оболочки, N_A — число Авогадро,

$I = \frac{1}{2} \sum_i c_i^\infty z_i^2$ — ионная сила раствора, \sum_j обозначает суммирование по ячейкам, контактирующим с данной ячейкой одной гранью. Шаг сетки является параметром модели и по умолчанию принимается равным 10^{-10} м.

Область для расчета потенциала задается как симметрично расположенный относительно начала координат прямоугольный параллелепипед, ребра которого параллельны осям собственной системы координат молекулы. Размеры области определяются как минимальные размеры, при которых любой атом белка расположен не ближе определенного расстояния (радиуса отсечения для электростатических взаимодействий) от границы области. Значение радиуса отсечения является параметром модели и по умолчанию принимается равным $3.5 \cdot 10^{-9}$ м. На границах области расчета потенциала электростатического поля его значение принимается равным нулю.

При расчете потенциала электростатического поля молекула белка рассматривается как набор ячеек кубической сетки с отличной от растворителя диэлектрической проницаемостью и пространственно-распределенными фиксированными зарядами. Принадлежность ячейки к области внутри белка определяется по ее доступности для молекул растворителя. В эту область входят

- (а) все ячейки, в центре которых нельзя разместить сферу заданного радиуса (по умолчанию – $1.4 \cdot 10^{-10}$ м [Connolly, 1983; Richards, 1977]) так, чтобы внутри нее не оказалось ни одной точки, являющейся координатой центра атома в файле структуры молекулы (PDB), а также
- (б) те ячейки, для которых условие (а) не выполнено, но которые принадлежат к замкнутой области, со всех сторон окруженной ячейками, для которых условие (а) выполнено.

Все остальные ячейки считаются заполненными растворителем с мобильными зарядами (ионами). Величина ионной силы раствора является параметром модели. Диэлектрическая проницаемость растворителя принимается равной 80; в ячейках, принадлежащих белку и контактирующих с растворителем, $\epsilon = 40$; в ячейках, принадлежащих белку и не контактирующих с растворителем, $\epsilon = 2$. Для ячеек растворителя учитывается экранирующее влияние мобильных зарядов (ионов).

Значения электрических зарядов на атомах белков могут быть прочитаны из файла структуры белка или дополнительного файла (поддерживается чтение файлов в формате PDB и PQR, а также файлов топологии программного пакета GROMACS [Hess et al., 2008]), либо рассчитаны автоматически по уравнению Хендерсона–Хассельбальха со стандартными значениями константы диссоциации pK_a для аминокислотных остатков [Dawson et al., 1986; Hasselbalch, 1917; Henderson, 1908] на основании заданного в конфигурационном файле значения pH раствора.

Расчет электростатического поля и его градиента проводится однократно для каждого типа молекул. Электростатическая сила и ее момент, действующие на каждую макромолекулу со стороны окружающих ее макромолекул, вычисляются на каждом шаге модели, исходя из заранее рассчитанных значений напряженности электростатического поля; при этом влияние диэлектрической проницаемости самой молекулы не учитывается. Это позволяет значительно увеличить скорость расчета, так как используются предварительно рассчитанные значения напряженности поля и не требуется повторный расчет напряженности электростатического поля на каждом временном шаге модели. Расчет электростатических взаимодействий для пар макромолекул, расстояние между всеми атомами которых заведомо превышает радиус отсечения электростатического взаимодействия, не производится.

4.3. Контактные взаимодействия

Общепринятым способом описания контактных взаимодействий в молекулярном моделировании является использование эмпирического потенциала Леннарда-Джонса («потенциала 6–12»). В то же время этот способ расчета контактных взаимодействий является весьма ресурсоемким и требует достаточно малого шага модели по времени (порядка 10^{-15} – 10^{-12} с), что плохо совместимо с методом броуновской динамики, в котором шаг по времени ограничен снизу (см. раздел 2). В *ProKSim* реализована более простая модель контактных взаимодействий как неупругих взаимодействий твердых тел. Для каждого белка определяется запрещенный объем, то есть область пространства, в которой не может находиться другая макромолекула [Northrup et al., 1987]. В том случае, если перемещение молекулы приводит к столкновению с другими молекулами или границами компартмента (если не используются периодические граничные условия), величина перемещения пропорционально уменьшается методом двоичного поиска по всем поступательным и вращательным степеням свободы. Расчет новых координат молекул и проверка их взаимопересечения осуществляется на каждом шаге модели последовательно для каждой молекулы.

При старте программы определяются радиусы сфер, вмещающих молекулы белков. Сначала проверка возможности взаимопересечения белковых молекул производится путем сравнения расстояния между их центрами масс с суммой радиусов вмещающих сфер. Для дальнейшей

более точной проверки столкновения макромолекул между собой в программе реализовано несколько алгоритмов:

- (а) Наиболее простой и очевидный (но и самый требовательный к вычислительным ресурсам) алгоритм использует представление каждого атома в виде шаров. Радиус шара равен ван-дер-ваальсовому радиусу для соответствующего химического элемента. Если два шара, относящиеся к разным макромолекулам, пересекаются, то такое взаимное расположение молекул считается невозможным и должно быть исправлено путем уменьшения величины перемещения молекулы в процессе работы модели, либо путем размещения молекулы в другом месте реакционного объема, если проверка производится на этапе начальной подготовки модельной сцены.
- (б) В другом варианте для проверки взаимопересечения молекул проводится проверка факта пересечения наборов ячеек сетки, используемой при расчете потенциала электростатического поля (см. рис. 1). Молекулы считаются столкнувшимися в том случае, если координаты центра поверхностной ячейки одного белка, то есть принадлежащей белку, но контактирующей с растворителем, либо, в зависимости от параметра модели, приповерхностной, то есть принадлежащей растворителю, но контактирующей с белком, оказываются внутри принадлежащей другому белку области. Этот вариант дает менее точные результаты, чем предыдущий, однако такая проверка требует значительно меньших вычислительных ресурсов.
- (в) Молекула белка представляется в виде набора шаров различных диаметров, приближенно описывающего форму макромолекулы [Kovalenko et al., 2006].

Для ускорения вычислений без ущерба для точности полученных результатов указанные алгоритмы могут использоваться совместно, например, вначале проводится грубая проверка взаимопересечения молекул по алгоритму (б), а более точная проверка с использованием координат атомов и их ван-дер-ваальсовых радиусов проводится только в том случае, если менее точный алгоритм предполагает возможность пересечения. Какие именно алгоритмы используются и в каком порядке, задается в конфигурационном файле программы.

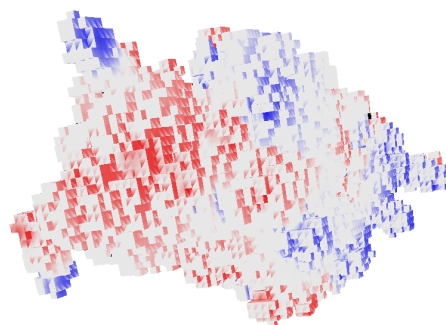


Рис. 1. Представление молекул белка в виде набора ячеек с ребром 10^{-10} м (алгоритм 4.4, б) на примере цитохрома *f* (см. далее). Цветом обозначено значение электростатического потенциала в ячейках от -0.3 В (красный) до $+0.3$ В (синий). Цветная версия изображения доступна в электронной версии журнала на сайте <http://crm.ics.org.ru/>

5. Анализ статистики взаимодействия макромолекул и кинетики образования белок-белковых комплексов

5.1. Статистика межмолекулярных взаимодействий

В программе *ProKSim* предусмотрена возможность сохранения взаимного расположения макромолекул в одном из стандартных форматов файлов, предназначенных для визуализации и анализа молекулярных структур. Возможна запись структур комплексов в файлы формата PDB (RCSB Protein Data Bank) или в платформенно-независимом формате файла для траекторий GROMACS XTC [Green et al., 1995]. Предусмотрено несколько вариантов отбора структур для последующего анализа:

- (а) Отбираются конфигурации, в которых расстояние между любыми двумя атомами, относящимися к разным молекулам, не превышает пороговое значение. Пороговое значение является параметром модели и по умолчанию принимается равным $5 \cdot 10^{-10}$ м.

- (б) Отбираются конфигурации, для которых электростатические взаимодействия приводят к взаимному притяжению молекул с энергией, превышающей (по модулю) пороговое значение. Пороговое значение является параметром модели и по умолчанию принимается равным $kT \approx 4.1 \cdot 10^{-21}$ Дж.
- (в) В конфигурационном файле программы задается набор расстояний между заранее выбранными точками в собственной системе координат молекулы. В том случае, если для проверяемых молекул заданное число расстояний (по умолчанию – все расстояния, указанные в файле конфигурации) меньше или равны указанным в конфигурационном файле значениям, считается, что взаимное расположение молекул удовлетворяет критерию.
- Проверка критериев производится на каждом шаге модели для каждой пары молекул.

5.2. Кинетика межмолекулярных взаимодействий

В программе предусмотрена возможность прямого моделирования протекания биохимических реакций. Для каждого из перечисленных в разделе 5.1 критериев образования предварительного (диффузионно-столкновительного) комплекса может быть задана вероятность образования финального комплекса. Реакция образования комплексов рассматривается как необратимая, прореагировавшие молекулы исключаются из дальнейшего цикла работы модели (их положение в реакционном объеме фиксируется; они участвуют в контактных и электростатических взаимодействиях с другими молекулами, но не могут участвовать в других реакциях).

Трансформация предварительного комплекса в финальный рассматривается как случайный процесс. Реакционное событие считается произошедшим в случае, если полученное от генератора равномерно распределенных в диапазоне $[0,1)$ псевдослучайных чисел значение оказывается меньшим или равным заданной в конфигурационном файле вероятности образования комплекса за 1 шаг модели. На основе полученных данных проводится оценка константы скорости бимолекулярной реакции взаимодействия белков из предположения равенства начальных концентраций реагентов и отсутствия конкурирующих реакций по формуле

$$k = \frac{c(t)}{c_0(c_0 - c(t))t}, \quad (4)$$

где c_0 — начальная концентрация каждого из реагентов и $c(t)$ — концентрация комплекса в момент времени t .

6. Получение структур диффузионно-столкновительных комплексов и анализ их конфигурации

Проведено исследование диффузионно-столкновительных комплексов, получаемых в результате диффузии с учетом дальнедействующих электростатических взаимодействий, для трех пар белков: ферредоксин и ферредоксин:НАДФ⁺-редуктаза, пластоцианин и цитохром *f*, барназа и барстар.

В кубический реакционный объем с периодическими граничными условиями и размером грани $1.6 \cdot 10^{-7}$ м помещали по 100 молекул каждого белка, что соответствует концентрации $4 \cdot 10^{-5}$ М. В качестве параметров модели использовали приведенные выше значения по умолчанию. Контактные взаимодействия учитывали по алгоритму 4.2, а. Для каждой пары белков проводили по два вычислительных эксперимента: в одном электростатические взаимодействия учитывались, в другом — не учитывались. Электростатические заряды на ионизированных аминокислотных остатках рассчитывали по уравнению Хендерсона–Хассельбальха для pH 7; ионную силу раствора брали равной 10^{-1} М. В каждом эксперименте по критериям 5.1, а и 5.1, б

отбирали взаимные положения белков для дальнейшего анализа. Таким образом, для каждой пары белков было получено три набора структур:

- (а) диффузионно-столкновительные комплексы, отобранные по критерию 5.1, а при диффузии белков без учета электростатических взаимодействий,
- (б) диффузионно-столкновительные комплексы, отобранные по критерию 5.1, а при диффузии белков с учетом электростатических взаимодействий,
- (в) структуры, отобранные по критерию 5.1(б) при диффузии белков с учетом электростатических взаимодействий.

Для каждой пары белков анализировали выборку из 20 000 случайно отобранных конфигураций комплексов из набора 6, б. В качестве меры сходства между комплексами было использовано среднеквадратичное отклонение положений α -углеродных атомов белков (RMSD) [Kabsch, 1976]. Были вычислены значения RMSD для всех пар конфигураций в каждой выборке и получена треугольная матрица расстояний 20000 x 20000. Для каждой выборки была определена типичная структура, то есть такая конфигурация, среднее значение RMSD которой по сравнению с остальными структурами в выборке минимально. Эта типичная структура в дальнейшем использовалась как эталон при расчете среднеквадратичного отклонения атомов в конфигурациях, отобранных по критерию 5.1, а для вычислительных экспериментов с учетом и без учета электростатических взаимодействий и 5.1, б для вычислительного эксперимента с учетом электростатических взаимодействий.

6.1. Неиерархический кластерный анализ структур комплексов белков

Для анализа представительности выбранной типичной структуры проводили кластерный анализ по методу, описанному в [Daura et al., 1999] и реализованный в утилите `g_cluster` программного пакета GROMACS. Процедура кластеризации состоит в том, что для всех отобранных комплексов подсчитывается количество сходных с ним. Два комплекса считаются схожими между собой в случае, если величина RMSD для них меньше заданного порогового значения (*расстояния кластеризации*). Кластер структур включает комплекс, имеющий наибольшее количество схожих с ним структур комплексов, и все эти структуры. После нахождения кластера все комплексы, входящие в него, исключаются из дальнейшего рассмотрения, и процедура кластеризации повторяется до тех пор, пока все комплексы не будут отнесены к какому-либо кластеру. Таким образом, в результате кластеризации каждая структура оказывается отнесенной к одному и только одному кластеру.

Для каждой выборки может быть определено такое минимальное значение расстояния кластеризации, начиная с которого все структуры относятся к одному кластеру. Если выборка состоит из случайно ориентированных друг относительно друга пар белков, отобранных по критерию 5.1, а, то при расстоянии кластеризации, лишь незначительно меньшем этого минимального значения, выборка будет разделена на два кластера, первый из которых будет включать почти все структуры комплексов, а второй — лишь одну или нескольких структур. При дальнейшем уменьшении расстояния кластеризации доля структур, входящих во второй кластер, будет возрастать, пока практически не сравняется с долей структур в первом кластере; дальнейшее уменьшение расстояния кластеризации приведет к быстрому росту числа кластеров. В том случае если взаимная ориентация белков в парах белков не полностью случайна и имеется какая-то предпочтительная ориентация, то зависимости числа кластеров и соотношения доли структур, относящихся ко второму и к первому кластеру, от расстояния кластеризации будут иными: на периферии первого кластера появится большое количество микрокластеров, каждый из которых включает лишь небольшую долю всех структур. Соотношение доли структур во втором и в первом кластере приблизится к единице лишь тогда, когда общее число кластеров будет достаточно большим. Таким образом, анализируя характер зависимости этих параметров от расстояния кластеризации, можно оценить, насколько адекватно описание выборки с помощью единственной типичной структуры комплекса.

6.2. Иерархический кластерный анализ структур комплексов белков

Для иерархической группировки структур по частоте встречаемости использовали однопараметрический метод классификации структур с использованием алгоритмов, описанных в [Ankerst et al., 1999; Elke et al., 2006; Sander et al., 2003]. Алгоритм был адаптирован с учетом специфики анализируемых данных. В этом методе под кластером понимается плотная (в смысле частоты встречаемости похожих структур) группа, отделенная от смежных групп менее плотными областями. Для уменьшения влияния случайного появления похожих структур в разреженных областях для каждой структуры вычисляется так называемое *центральное расстояние* CD (*core distance*). Это расстояние зависит от заданного исследователем параметра — минимального количества структур в кластере M — и определяется как значение RMSD данной структуры по сравнению M -той (по порядку увеличения RMSD) структурой из выборки. В данной работе принимали $M = 300$.

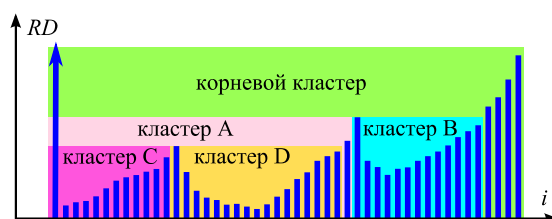


Рис. 2. Схематическое представление результата работы алгоритма поиска кластеров с помощью диаграммы зависимости расстояния доступности RD от номера структуры комплекса в списке i . Кластеры А и В вложены в корневой кластер, кластеры С и D — в кластер А. Цветная версия рисунка доступна в электронной версии журнала на сайте <http://ctm.ics.rg.ru/>

Задача кластеризации решается в два этапа. Вначале проводится упорядочивание структур: начиная с произвольно выбранной структуры (в нашей работе мы начинали со структуры с наименьшим центральным расстоянием), составляется список, в который на каждом шаге работы алгоритма добавляется та из отсутствующих в списке структур, которая наиболее похожа на любую из добавленных ранее. В качестве критерия добавления используется расстояние доступности RD (*reachability distance*). RD принимается равным минимальному значению RMSD между данной структурой и любой структурой, уже добавленной в список, если это значение больше CD данной структуры, либо равным соответствующему значению CD — в противном случае. На каждом шаге работы алгоритма RD пересчитывается для всех структур, еще не добавленных в список. При добавлении структуры в список значение RD сохраняется для дальнейшего анализа; для первой структуры оно принимается равным $+\infty$.

На втором этапе упорядоченный список используется для выделения кластеров структур; при этом кластеры являются вложенными друг в друга, и каждая структура относится как минимум к одному (корневому) кластеру. Проводится поиск локальных максимумов зависимости расстояния доступности структуры от ее номера в списке. Предшествующие локальному максимуму и следующие за ним подряд группы структур, имеющие расстояние доступности не более расстояния доступности локального максимума, выделяются в качестве отдельных подкластеров в том случае, если количество входящих в каждую из них структур не менее M . Структура комплекса, соответствующая локальному максимуму RD , считается принадлежащей к правому кластеру (рис. 2).

В результате применения этого алгоритма структуры комплексов белков оказываются расклассифицированными по иерархическому дереву, каждая ветвь которого соответствует более или менее обособленной группе схожих структур. Для каждой группы может быть определена своя типичная структура; можно сравнить группы по частоте (плотности) встречаемости похожих структур в группе и оценить степень обособленности двух близких групп. Группы структур (кластеры) могут быть вложены друг в друга, поэтому одна структура может принадлежать одновременно к нескольким кластерам.

7. Объекты исследования

7.1. Ферредоксин и ферредоксин:НАДФ⁺-редуктаза

Ферредоксин служит конечным акцептором электронов фотосистемы I. Он передает электроны в разные пути метаболизма, в том числе в путь фиксации углерода через ферредоксин:НАДФ⁺-редуктазу (ФНР), восстанавливающую никотинамиддинуклеотидфосфат (НАДФ⁺) до НАДФН, необходимого в цикле Кальвина. Ферредоксин содержит железо-серный кластер [2Fe-2S], в котором один из атомов железа меняет степень окисления между +2 и +3. В ФНР электроны распределяются по сопряженной системе ароматических связей флавинаденидинуклеотида (ФАД). Константа скорости образования комплекса ферредоксина и ФНР из циано-бактерии *Anabaena* составляет $k_a = 6.9 \cdot 10^8 \text{ M}^{-1} \text{ s}^{-1}$, константа диссоциации $K_d = 2.2 \cdot 10^{-6} \text{ M}$ [Hurley et al., 1999].

Взаимодействие белков ферредоксина и ФНР было исследовано [Kovalenko et al., 2010] методом многочастичной броуновской динамики. Были вычислены зависимости константы скорости образования комплексов белков от ионной силы для дикого типа и мутантов ФНР.

В настоящей модели рассматривается полученная методом рентгеноструктурного анализа (РСА) структура комплекса ферредоксина и ФНР из *Anabaena* (PDB ID: 1EWY) [Morales et al., 2000]. В качестве исходных данных для построения модели использовали цепь С (ферредоксин) и цепь А (ФНР) указанного комплекса. Недостающие в структуре легкие атомы были добавлены с помощью утилиты `pdb2gmx` программного комплекса GROMACS. Суммарные заряды белков в модели составляли -14.9 (ферредоксин) и -2.2 (ферредоксин:НАДФ⁺-редуктаза) элементарного заряда.

7.2. Пластоцианин и цитохром *f*

Пластоцианин является мобильным переносчиком электронов, передающим электрон с цитохромного *b₆f* комплекса на фотосистему I у всех высших растений и некоторых водорослей. Цитохром *f* — самая крупная субъединица цитохромного *b₆f* комплекса, содержащая сайт связывания пластоцианина. Окислительно-восстановительная активность пластоцианина обеспечивается наличием в его составе атома меди, меняющего степень окисления между +2 и +1. Цитохром *f* содержит гем с атомом железа, меняющим степень окисления между +2 и +3. Комплекс пластоцианина с цитохромом *f* находится в быстром равновесии с несвязанными формами белков, для взаимодействия окисленного пластоцианина из шпината *Spinacia oleracea* и восстановленного цитохрома *f* из турнепса *Brassica rapa* константа ассоциации комплекса $K_a = 7 \cdot 10^3 \text{ M}^{-1}$, константа скорости образования комплекса $k_a = 2 \cdot 10^8 \text{ M}^{-1} \text{ s}^{-1}$ [Kannt et al., 1996].

Взаимодействие белков пластоцианина и цитохрома *f* высших растений было исследовано методом БД в работах [Pearson, Gross, 1998; Rienzo et al., 2001; Ullmann et al., 1997]. Методом многочастичной БД было исследовано взаимодействие пластоцианина и цитохрома *f* в растворе [Kovalenko et al., 2006] и люмене тилакоида [Коваленко и др., 2008], в том числе с учетом электростатических взаимодействий с тилакоидной мембраной [Князева и др., 2010].

В настоящей модели рассматривается структура комплекса пластоцианина шпината и цитохрома *f* турнепса (PDB ID: 2PCF) [Ubbink et al., 1998], полученная методом спектроскопии ядерного магнитного резонанса (ЯМР). Распределение электростатических зарядов на геме взято из [Autenrieth et al., 2004]. Суммарные заряды белков в модели составляли -8.00 (пластоцианин) и -3.95 (цитохром *f*) элементарного заряда.

7.3. Барназа и барстар

Бактериальная РНК-аза (барназа) и ее ингибитор барстар являются одним из классических объектов при моделировании белок-белковых взаимодействий [Hoefling, Gottschalk, 2010];

Kieslich et al., 2010; Spaar et al., 2006; Wang et al., 2004; Ермакова, 2006], так как формируют (в среде с водородным показателем рН от 7 до 9) очень прочно связанный комплекс. Период полураспада этого комплекса составляет около двух дней (константа диссоциации комплекса $K_d = 10^{-14}$ М, константа скорости образования комплекса $k_a = 4 \cdot 10^8$ М⁻¹с⁻¹). В кислой среде (рН ниже 4.5) сродство белков друг к другу полностью исчезает [Schreiber, 2001]. Барназа является ядом, секретлируемым бактерией *Bacillus amyloliquefaciens* во внешнюю среду. Формирование устойчивого комплекса белков барназы и барстара внутри клетки бактерии является непременным условием выживания организма, так как белок чрезвычайно цитотоксичен из-за крайне высокой РНК-азной активности.

В модели рассматривали структуру комплекса барназы и барстара (PDB ID: 1BGS) [Guillet et al., 1993], полученную методом рентгеноструктурного анализа. В качестве исходных данных для построения модели использовали цепь А (барназа) и цепь Е (барстар) указанного комплекса. Недостающие в структуре легкие атомы были добавлены с помощью утилиты `pdb2gmx` программного комплекса GROMACS. Суммарные заряды белков в модели составляли +2.23 (барназа) и -5.89 (барстар) элементарного заряда.

8. Типичные структуры диффузионно-столкновительных комплексов

Проведено исследование диффузионно-столкновительных комплексов для трех пар белков: ферредоксин и ферредоксин:НАДФ⁺-редуктаза, пластоцианин и цитохром *f*, барназа и барстар. На рисунке 3 показаны результаты неиерархической кластеризации структур диффузионно-столкновительных комплексов, отобранных по критерию 5.1, а, с помощью метода, описанного в разделе 6.1. Вид приведенных графиков свидетельствует о том, что распределение конфигураций диффузионно-столкновительных комплексов ферредоксина и ферредоксин:НАДФ⁺-редуктазы и пластоцианина и цитохрома *f* значительно отличается от случайного и для них можно выбрать представительные типичные структуры. Эти структуры визуализированы на рис. 4, в и 5, б. Распределение структур диффузионно-столкновительных комплексов барназы и барстара соответствует случайному, и, следовательно, полученная типичная структура (рисунок 6, б) не может быть признана представительной.

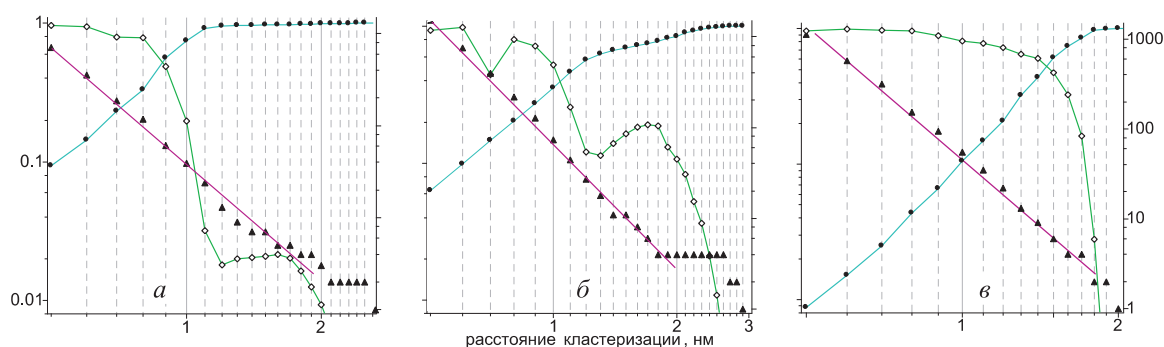


Рис. 3. Зависимости числа кластеров (\blacktriangle , по правой оси ординат), доли самого крупного кластера от всей выборки (\bullet , по левой оси ординат) и соотношения размеров второго и первого кластера (\diamond , по левой оси ординат) от расстояния кластеризации. Графики приведены в логарифмическом масштабе. (а) – ферредоксин и ферредоксин:НАДФ⁺-редуктаза, (б) – пластоцианин и цитохром *f*, (в) – барназа и барстар.

8.1. Ферредоксин и ферредоксин:НАДФ⁺-редуктаза

На рисунке 4 видно, что ориентация белков ферредоксина и ФНР в диффузионно-столкновительном комплексе в общем сходна с их ориентацией в кристалле (PDB ID: 1EWY). Следует отметить, что в кристаллической структуре на каждую молекулу ферредоксина приходится по две молекулы ФНР, и она оказывается зажатой между этими молекулами. Одна из этих

молекул (цепь А) сближена с железо-серным кластером ферредоксина (цепь С) флавиновым трициклом (расстояние $7.4 \cdot 10^{-10}$ м), другая (цепь В) – адениновым остатком (расстояние $1.1 \cdot 10^{-9}$ м). Расстояние между флавиновым трициклом цепи В и железо-серным кластером ферредоксина составляет $1.5 \cdot 10^{-9}$ м. В полученной типичной структуре диффузионно-столкновительного комплекса расстояние между флавиновым трициклом ФНР и железо-серным кластером ферредоксина составляет $2.3 \cdot 10^{-9}$ м. Среднеквадратичное отклонение положений α -углеродных атомов составляет $2.4 \cdot 10^{-9}$ м (как при сравнении с цепями А и С, так и при сравнении с цепями В и С).

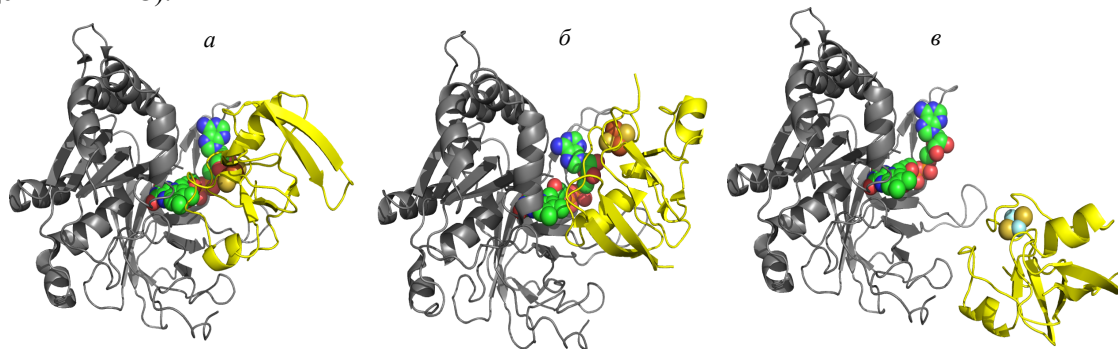


Рис. 4. Структура комплекса ферредоксина (желтый) и ФНР (серый), полученного методом РСА (PDB ID: 1EWY) (*a* – цепи А и С, *б* – цепи В и С), и типичная структура, полученная при анализе траекторий броуновской динамики (*в*). В виде сфер обозначены атомы, входящие в состав ФАД и железо-серного кластера [2Fe-2S]. Для визуализации структур использован PyMol [Schrödinger LLC, 2012]. Цветная версия изображения доступна в электронной версии журнала на сайте <http://crm.ics.org.ru/>

8.2. Пластоцианин и цитохром *f*

В типичной структуре диффузионно-столкновительного комплекса пластоцианина и цитохрома *f* (рис. 5) пластоцианин расположен вблизи экспериментально обнаруженного сайта связывания, однако ориентация его значительно отличается от экспериментально определенной. Расстояние между гемовым железом цитохрома *f* и ионом меди пластоцианина в экспериментально изученной структуре комплекса (PDB ID: 2PCF) составляет $1.1 \cdot 10^{-9}$ м, в то время как в полученном методом броуновской динамики диффузионно-столкновительном комплексе — $4.1 \cdot 10^{-9}$ м. Среднеквадратичное отклонение положения α -углеродных атомов от положения в комплексе, полученном методом ЯМР (PDB ID: 2PCF, модель 9), составляет $1.3 \cdot 10^{-9}$ м.

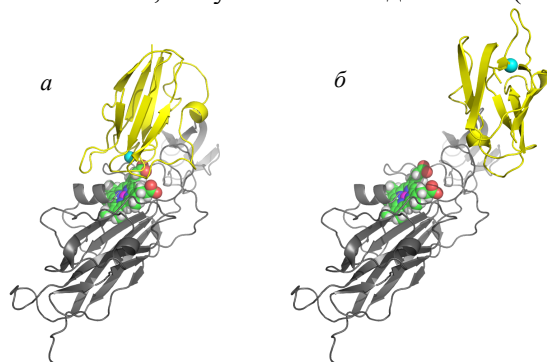


Рис. 5. Структура комплекса пластоцианина (желтый) и цитохрома *f* (серый), полученного методом ЯМР (PDB ID: 2PCF, модель 9) (*a*), и типичная структура, полученная при анализе траекторий броуновской динамики (*б*). В виде сфер обозначены ион меди и атомы, входящие в состав гема. Для визуализации структур использован PyMol [Schrödinger LLC, 2012]. Цветная версия изображения доступна в электронной версии журнала на сайте <http://crm.ics.org.ru/>

8.3. Барназа и барстар

На рисунке 6 показаны экспериментально полученные варианты ориентации белка барстара относительно барназы в кристалле (PDB ID: 1BGS) и типичная модельная структура диффузионно-столкновительного комплекса. Можно видеть, что эта структура не соответствует ни одному из вариантов расположения этих белков в кристалле. Более подробный анализ модель-

ной структуры не проводился, так как методами кластерного анализа было показано, что данная структура не может быть признана представительной для исследованной выборки. Использование иерархической классификации структур диффузионно-столкновительных комплексов, отобранных по критерию 5.1, *a*, с помощью метода, описанного в разделе 6.2, также не позволило выявить какие-либо конфигурации, частота встречаемости которых в выборке значительно отличалась от средней. В то же время анализ структур, отобранных по критерию 5.1, *б*, то есть таких конфигураций, в которых энергия притяжения между двумя молекулами превышает пороговое значение $kT \approx 4.1 \cdot 10^{-21}$ Дж, позволил выявить две обособленные группы энергетически выгодных взаимных расположений белков (рис. 7). Среди энергетически выгодных ориентаций наиболее часто встречаются структуры, схожие с приведенной на рисунке 7, *з*. Расположение барстара относительно барназы в этой структуре соответствует одному из физиологически незначимых положений в кристалле, хотя его ориентация отличается от экспериментально полученной (рис. 7, *б*; среднеквадратичное отклонение положений α -углеродных атомов — $1.1 \cdot 10^{-9}$ м). Вторая по встречаемости группа показана на рисунке 7, *в* и близка к физиологически значимому комплексу (рис. 7а; среднеквадратичное отклонение положений α -углеродных атомов — $1.3 \cdot 10^{-9}$ м). Полученные энергетически выгодные состояния находятся в согласии с результатами [Spraag et al., 2006]; использованный нами метод иерархической классификации структур позволил более наглядно представить эти состояния, так как использованный в [Spraag et al., 2006] метод позволяет только определить расположение центра масс одного белка относительно другого.

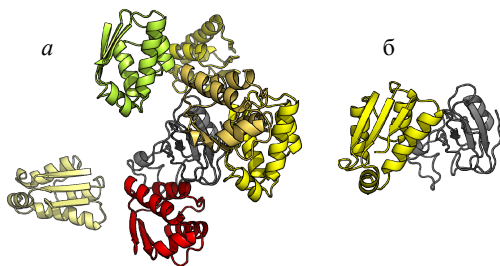


Рис. 6. *a* — варианты расположения барстара относительно барназы (серый) в комплексе, исследованном методом PCA (PDB ID: 1BGS). Красным показано физиологически значимое положение, различными оттенками желтого — альтернативные положения с учетом симметрии кристаллической решетки. *б* — неrepresentative типичная структура, полученная при анализе траекторий броуновской динамики. Для визуализации структур использован PyMol [Schrödinger LLC, 2012]. Цветная версия изображения доступна в электронной версии журнала на сайте <http://crm.ics.org.ru/>

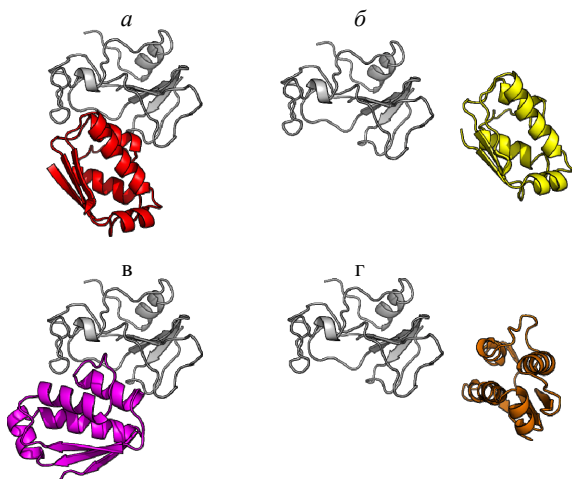


Рис. 7. *a, б* — варианты расположения барстара относительно барназы (серый) в комплексе, исследованном методом PCA (PDB ID: 1BGS). Красным показано физиологически значимое положение (*a*). *в, з* — соответствующие (*a*) и (*б*) группы энергетически выгодных ориентаций, полученных методом броуновской динамики. Для визуализации структур использован PyMol [Schrödinger LLC, 2012]. Цветная версия изображения доступна в электронной версии журнала на сайте <http://crm.ics.org.ru/>

9. Роль электростатических взаимодействий во взаимной ориентации белков в процессе диффузии

Для двух исследованных пар фотосинтетических электрон-транспортных белков ферредоксина и ферредоксин:НАДФ⁺-редуктазы и пластоцианина и цитохрома *f* получены предста-

вительные типичные структуры диффузионно-столкновительных комплексов. Показано, что в процессе диффузии с учетом электростатических взаимодействий мобильные переносчики электронов (ферредоксин и пластоцианин) подходят к соответствующим сайтам своих партнеров по взаимодействию, однако их ориентация может отличаться от ориентации в финальном комплексе (то есть конфигурации, в которой возможен перенос электрона). Для количественного определения вклада электростатических взаимодействий в формирование ансамбля структур диффузионно-столкновительного комплекса было проведено сравнение (по параметру среднеквадратичного отклонения положения α -углеродных атомов белков) всех отобранных в вычислительных экспериментах конфигураций диффузионно-столкновительных комплексов с соответствующими типичными структурами. График функции плотности полученных распределений показан на рис. 8. Видно, что распределение, полученное по критерию б, а и б, в. Полученные по методу наименьших квадратов веса соответствующих распределений приведены в таблице 1. Полученные веса дают количественную оценку роли далекодействующих электростатических взаимодействий для процесса сближения и взаимной ориентации белков в процессе диффузии.

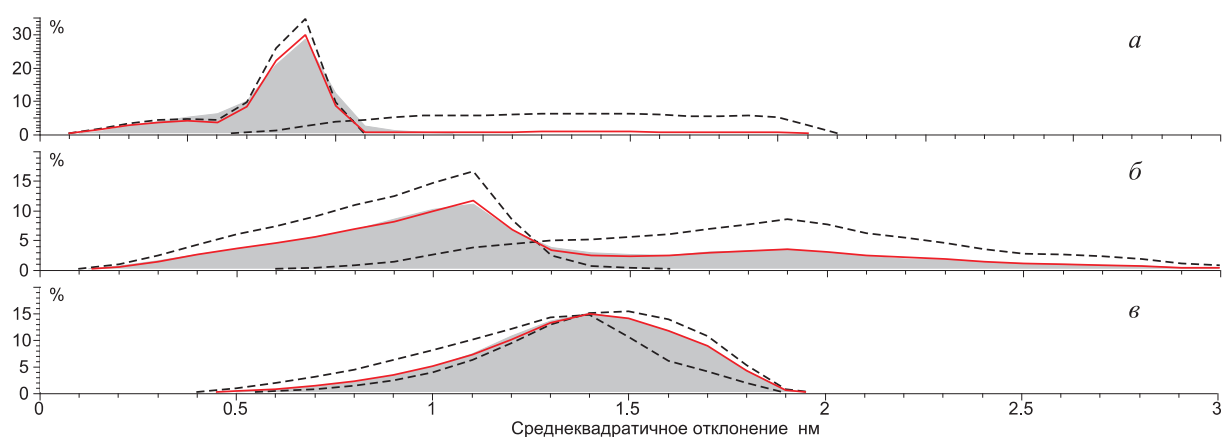


Рис. 8. Функция плотности распределения среднеквадратичного отклонения положения α -углеродных атомов диффузионно-столкновительных комплексов белков относительно типичной структуры. Пунктирная линия — диффузия без учета электростатических взаимодействий, отбор комплексов при расстоянии между атомами двух белков меньше $5 \cdot 10^{-10}$ м (б, а); точечная линия — диффузия с учетом электростатических взаимодействий, отбор комплексов при энергии электростатического притяжения $\geq kT$ (б, в). Полутоновая заливка — диффузия с учетом электростатических взаимодействий, отбор комплексов при расстоянии между атомами двух белков меньше $5 \cdot 10^{-10}$ м (б, б); красная сплошная линия — результат фитирования распределения б, б в виде взвешенной суммы распределений б, а и б, в, соответствующие веса приведены в таблице 1. а — ферредоксин и ферредоксин:НАДФ⁺-редуктаза, б — пластоцианин и цитохром *f*, в — барназа и барстар. Цветная версия графика доступна в электронной версии журнала на сайте <http://ctm.ics.org.ru/>

В представленной работе показана возможность применения метода броуновской динамики для определения роли далекодействующих электростатических во взаимной ориентации макромолекул в процессе диффузии и получения типичных структур диффузионно-столкновительных комплексов. Этот метод также широко применяется для изучения кинетики диффузионно-контролируемых реакций [Kovalenko et al., 2006 и др. работы]. Однако при сближении молекул в их взаимодействии существенную роль начинают играть те факторы, которыми мы пренебрегаем в методе броуновской динамики. К числу таких факторов относится внутримолекулярная подвижность, формирование водородных связей и солевых мостиков, а также гидрофобные взаимодействия. Все эти факторы участвуют в трансформации предварительных диффузионно-столкновительных комплексов в финальные (обладающие физиологической активностью, например, способствующие переносу электрона между активными центрами электрон-транспортных белков). Трансформация диффузионно-столкновительных комплексов, по-

лучаемых с помощью программного обеспечения *ProKSim*, может быть исследована с помощью методов молекулярной динамики и квантовой химии, что выходит за рамки представленной работы.

Таблица 1. Коэффициенты фитирования по методу наименьших квадратов функции плотности распределения б, в виде взвешенной суммы распределений б, а и б, в.

Белки	Вес функции плотности распределения в суперпозиции	
	б, а (электростатические взаимодействия)	б, в (без учета электростатических взаимодействий)
Ферредоксин и ферредоксин:НАДФ ⁺ -редуктаза	0.85	0.15
Пластоцианин и цитохром <i>f</i>	0.61	0.39
Барназа и барстар	0.30	0.70

Программное обеспечение *ProKSim* может быть загружено по адресу http://www.biophys.msu.ru/rus/science/complex_systems/ProKSim/. Исходный код может быть предоставлен по запросу.

10. Список литературы

- Воеводин В. В., Жуматий С. А., Соболев С. И., Антонов А. С., Брызгалов П. А., Никитенко Д. А., Стефанов К. С., Воеводин В. В. Практика суперкомпьютера «Ломоносов» // Открытые системы. — 2012. — № 7.
- Ермакова Е. А. Сравнительное изучение реакции взаимодействия *Bacillus umyloliquefaciens* рибонуклеазы (барназы) и *Bacillus intermedius* рибонуклеазы (биназы) с барстаром методом броуновской динамики // Биофизика. — 2006. — Т. 51, № 2. — С. 242–249.
- Князева О. С., Коваленко И. Б., Абатурова А. М., Ризниченко Г. Ю., Грачев Е. А., Рубин А. Б. Многочастичная модель диффузии и взаимодействия пластоцианина с цитохромом *f* в электростатическом поле фотосинтетической мембраны // Биофизика. — 2010. — Т. 55, № 2. — С. 259–268.
- Коваленко И. Б., Абатурова А. М., Громов П. А., Устинин Д. М., Грачев Е. А., Ризниченко Г. Ю., Рубин А. Б. Компьютерное моделирование образования комплекса между пластоцианином и цитохромом *f* в люмене тилакоида // Биофизика. — 2008. — Т. 53, № 2. — С. 261–270.
- Коваленко И. Б., Абатурова А. М., Ризниченко Г. Ю., Рубин А. Б. Новый метод компьютерного моделирования образования белок-белковых комплексов // Доклады Академии наук. — 2009. — Т. 427, № 5. — С. 696–698.
- Ankerst M., Breunig M. M., Kriegel H.-P., Sander J. OPTICS: Ordering Points To Identify the Clustering Structure // Proceedings of ACM SIGMOD, Philadelphia, PA., 1999. С. 49–60.
- Autenrieth F., Tajkhorshid E., Baudry J., Luthey-Schulten Z. Classical force field parameters for the heme prosthetic group of cytochrome *c*. // Journal of computational chemistry. — 2004. — Т. 25, № 13. — С. 1613–22.
- Berman H., Henrick K., Nakamura H., Markley J.L. The worldwide Protein Data Bank (wwPDB): ensuring a single, uniform archive of PDB data // Nucleic Acids Research. — 2007. — Т. 35. — Database issue. С. D301–D303.
- Connolly M. L. Analytical molecular surface calculation // J. Appl. Crystallogr. — 1983. — Т. 16. — С. 548–558.
- Daura X., Gademann K., Jaun B., Seebach D., Gunsteren W.F. van, Mark A.E. Peptide Folding: When Simulation Meets Experiment // Angewandte Chemie International Edition. — 1999. — Т. 38, № 1–2. — С. 236–240.

- Dawson R. M. C., Elliott D. C., Elliott W. H., Jones K. M.* Data for Biochemical Research (third edition). Oxford: Clarendon Press, 1986.
- Elke A., Böhm C., Kröger P.* DeLiClu: Boosting Robustness, Completeness, Usability, and Efficiency of Hierarchical Clustering by a Closest Pair Ranking // Proceedings of the 10th Pacific-Asian Conference on Advances in Knowledge Discovery and Data Mining., 2006. C. 119–128.
- Ermak D. L., McCammon J. A.* Brownian dynamics with hydrodynamic interactions // J. Chem. Phys. — 1978. — T. 69, № 4. — C. 1353–1360.
- Gabdoulline R. R., Wade R. C.* On the protein-protein diffusional encounter complex. // Journal of molecular recognition JMR. — 1999. — T. 12, № 4. — C. 226–234.
- Green D., Meacham K., Surridge M., Hoesel F. van, Berendsen H.* Methods and Techniques in Computational Chemistry: METECC-95 // Methods and Techniques in Computational Chemistry: METECC-95. : STEF, Cagliari, 1995. — C. 435.
- Guillet V., Laphorn A., Hartley R.W., Mauguen Y.* Recognition between a bacterial ribonuclease, barnase, and its natural inhibitor, barstar // Structure (London, England : 1993). — 1993. — T. 1, № 3. — C. 165–76.
- Hasselbalch K. A.* Die Berechnung der Wasserstoffzahl des Blutes aus der freien und gebundenen Kohlensäure desselben, und die Sauerstoffbindung des Blutes als Funktion der Wasserstoffzahl // Biochemische Zeitschrift. — 1917. — T. 78. — C. 112–144.
- Henderson L. J.* Concerning the relationship between the strength of acids and their capacity to preserve // Am. J. Physiol. — 1908. — T. 21. — C. 173–179.
- Hess B., Kutzner C., Van Der Spoel D., Lindahl E.* GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation // Journal of Chemical Theory and Computation. — 2008. — T. 4, № 3. — C. 435–447.
- Hoefling M., Gottschalk K. E.* Barnase-Barstar: From first encounter to final complex. // Journal of Structural Biology. — 2010. — T. 171, № 1. — C. 52–63.
- Hurley J. K., Hazzard J. T., Martinez-Julvez M., Medina M., Gomez-Moreno C., Tollin G.* Electrostatic forces involved in orienting Anabaena ferredoxin during binding to Anabaena ferredoxin:NADP⁺ reductase: site-specific mutagenesis, transient kinetic measurements, and electrostatic surface potentials // Protein Sci. 1999. T. 8. C. 1614–1622.
- Kabsch W.* A solution for the best rotation to relate two sets of vectors // Acta Crystallographica Section A. — 1976. — T. 32, № 5. — C. 922–923.
- Kannt A., Young S., Bendall D.S.* The role of acidic residues of plastocyanin in its interaction with cytochrome f // Biochim. Biophys. Acta. — 1996. — T. 1277. — C. 115–126.
- Kieslich C. A., Gorham R. D., Morikis D.* Is the rigid-body assumption reasonable?. Insights into the effects of dynamics on the electrostatic analysis of barnase-barstar // Journal of Non-Crystalline Solids. — 2010.
- Kovalenko I. B., Abaturova A. M., Diakonova A. N., Knyazeva O. S., Ustinin D. M., Khruschev S. S., Riznichenko G. Y., Rubin A. B.* Computer simulation of protein-protein association in photosynthesis // Math. Model. Nat. Phenom. — 2011. — T. 6, № 7. — C. 39–54.
- Kovalenko I. B., Abaturova A. M., Gromov P. A., Ustinin D. M., Grachev E. A., Riznichenko G. Y., Rubin A.B.* Direct simulation of plastocyanin and cytochrome f interactions in solution // Phys. Biol. — 2006. — T. 3. — C. 121–129.
- Kovalenko I. B., Diakonova A. N., Abaturova A. M., Riznichenko G. Y., Rubin A. B.* Direct computer simulation of ferredoxin and FNR complex formation in solution // Phys. Biol. — 2010. — T. 7, № 2. — C. 26001.
- Morales R., Kachalova G., Vellieux F., Charon M.H., Frey M.* Crystallographic studies of the interaction between the ferredoxin-NADP⁺ reductase and ferredoxin from the cyanobacterium Anabaena: looking for the elusive ferredoxin molecule. // Acta crystallographica. Section D, Biological crystallography. — 2000. — T. 56, № Pt 11. — C. 1408–12.
- Northrup S., Boles J., Reynolds J.* Electrostatic effects in the brownian dynamics of association and orientation of heme proteins // J Phys Chem. — 1987. — T. 91. — C. 5991–5998.

- Panneton F., L'Ecuyer P.* On the xorshift random number generators // ACM T. Model. Comput. S. — 2005. — Т. 15, № 4. — С. 346–361.
- Pearson Jr. D. C., Gross E. L.* Brownian Dynamics Study of the interaction between plastocyanin and cytochrome f // Biophysical Journal. — 1998. — Т. 75. — С. 2698–2711.
- Perrin F.* Mouvement Brownien d'un ellipsoïde (II). Rotation libre et dépolariation des fluorescences. Translation et diffusion de molécules ellipsoïdales // J. Phys. Radium. — 1936. — Т. 7, № 1. — С. 1–11.
- Richards F. M.* Areas, volumes, packing and protein structure. // Annual review of biophysics and bioengineering. — 1977. — Т. 6. — С. 151–76.
- De Rienzo F., Gabdoulline R. R., Menziani M. C., De Benedetti P. G., Wade R. C., Rienzo F., Benedetti P.* Electrostatic analysis and brownian dynamics simulation of the association of plastocyanin and cytochrome f // Biophys. J. — 2001. — Т. 81, № 3. — С. 3090–3104.
- Sander J., Qin X., Lu Z., Niu N., Kovarsky A.* Automatic Extraction of Clusters from Hierarchical Clustering Representations // Proceedings of the 7th Pacific-Asia Conference on Knowledge Discovery and DataMining (PAKDD), Seoul, Korea. : Springer-Verlag, 2003. — С. 75–87.
- Schreiber G.* Methods for Studying the Interaction of Barnase with Its Inhibitor Barstar // Nuclease Methods and Protocols (Methods in Molecular Biology) / под ред. C.H. Schein. : Humana Press, 2001. — С. 213–226.
- Schrödinger LLC.* The PyMOL Molecular Graphics System // 2012.
- Spaar A., Dammer C., Gabdoulline R.R., Wade R.C., Helms V.* Diffusional Encounter of Barnase and Barstar // Biophys. J. — 2006. — Т. 90, № 6. — С. 1913–1924.
- Ubbink M., Ejdebeck M., Karlsson B.G., Bendall D.S.* The structure of the complex of plastocyanin and cytochrome f, determined by paramagnetic NMR and restrained rigid-body molecular dynamics // Structure. — 1998. — Т. 6. — С. 323–335.
- Ullmann G. M., Knapp E.-W.* Electrostatic models for computing protonation and redox equilibria in proteins // Eur. Biophys. J. — 1999. — Т. 28, № 7. — С. 533–551.
- Ullmann G. M., Knapp E.-W., Kostic N.M.* Computational simulation and analysis of dynamic association between plastocyanin and cytochrome f. Consequences for the electron-transfer reaction // J. Am. Chem. Soc. — 1997. — Т. 119. — С. 42–52.
- Wang T., Tomic S., Gabdoulline R.R., Wade R.C.* How Optimal Are the Binding Energetics of Barnase and Barstar? // Biophysical Journal. — 2004. — Т. 87, № 3. — С. 1618–1630.

УДК: 519.8

Сопряжение различных методов компьютерного моделирования в комплексной модели фотосинтетической мембраны

Д. М. Устинин^а, И. Б. Коваленко, Г. Ю. Ризниченко, А. Б. Рубин

Московский государственный университет им. М. В. Ломоносова, Биологический факультет,
Россия, 119991, ГСП-1, г. Москва, Ленинские горы, МГУ

E-mail: ^а dmitry.ustinin@gmail.com

*Получено 18 декабря 2012 г.,
после доработки 01 марта 2013 г.*

Необходимость корректного учета деталей пространственной и функциональной организации клеточных структур требует поиска новых подходов к моделированию субклеточных процессов, в том числе первичных процессов фотосинтеза в тилакоидной мембране. Эти подходы должны интегрировать физические и биологические представления о конкретных механизмах, которые объединяются в общую картину на уровне компьютерной модели. В работе предлагается новый подход к моделированию, в котором воспроизводится трехмерная пространственная структура фотосинтетической мембраны. Разные стадии переноса зарядов при фотосинтезе моделируются с использованием разного математического аппарата и объединяются в единую компьютерную модель. Разработанные алгоритмы реализованы в виде программного комплекса, использующего параллельные вычисления на высокопроизводительных кластерах и графических процессорах.

Ключевые слова: компьютерное моделирование, броуновская динамика

Combination of different simulation techniques in the complex model of photosynthetic membrane

D. M. Ustinin, I. B. Kovalenko, G. Yu. Riznichenko, A. B. Rubin

Moscow State University, Faculty of Biology, MSU, GSP-1, Leninskiye Gory, Moscow, 119991, Russia

Abstract. — Complex geometric organization of subcellular structures, such as photosynthetic or mitochondrial membranes determines mechanism of electron and proton transfer processes. We propose new approach in modeling processes, where geometric shape of membranes is accurately taken into account. Different stages of charge transfer process are simulated following different approaches, which are integrated into a combined model. We implemented this model as software which utilize parallel computations on high-performance cluster and GPU for better performance.

Keywords: computer modeling, Brownian dynamics

Citation: *Computer Research and Modeling*, 2013, vol. 5, no. 1, pp. 65–81 (Russian).

Работа выполнена при финансовой поддержке РФФИ, гранты № 11-04-01019, № 11-04-01268, №12-07-00783, №12-07-33036

Введение

Фотосинтетическая система клетки представляет собой интересный объект для моделирования механизмов ее работы, что важно как для фундаментальной, так и для прикладной науки. Сейчас накоплен огромный массив информации о строении фотосинтетической системы. Известны структуры практически всех белков, участвующих в работе фотосинтетической цепи, пространственная организация фотосинтетической мембраны, доступны многочисленные экспериментальные методы изучения состояния системы (исследование флуоресценции, метод электронного парамагнитного резонанса и т.д.). Кроме того, возросшие мощности вычислительной техники позволяют строить подробные компьютерные модели, имитирующие процесс фотосинтеза и подробно учитывающие пространственную организацию системы.

Моделируя процессы фотосинтеза, необходимо учитывать два фактора. Во-первых, фотосинтетическая мембрана имеет сложную пространственную организацию (см. рис. 1). Во-вторых, характерные времена протекания отдельных стадий этих процессов значительно отличаются. Сначала рассмотрим более подробно структуру фотосинтетической мембраны.

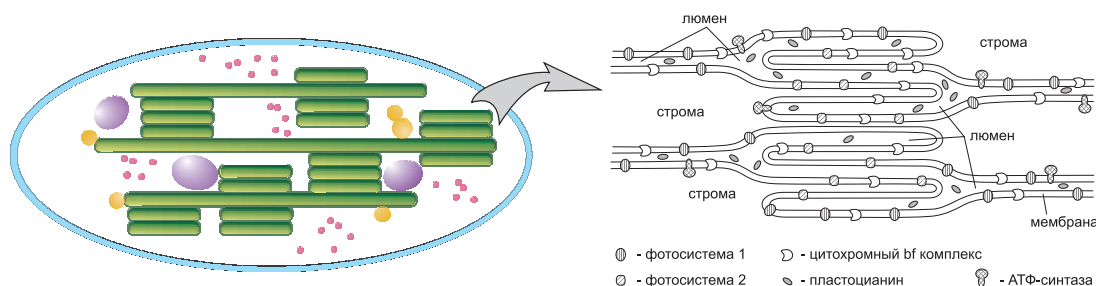


Рис. 1. Пространственная организация фотосинтетической мембраны хлоропласта

Она организована в виде большого количества стопок плоских дисков, называемых гранами (рис. 1). Каждая грана состоит из двойной липидной мембраны, в которую встроены белковые комплексы различной природы (фотосинтетические реакционные центры фотосистем 1 и 2, цитохромный b_6f комплекс, АТФ-синтазный комплекс). Внутреннее пространство гран между липидными мембранами называется люменом, внешнее пространство — стромой. Характерные размеры грани — диаметр около 500 нм, толщина мембран — 5 нм, толщина межмембранного пространства (люмена) — 5–10 нм.

Схема процессов, связанных с линейным транспортом электронов в фотосинтетической мембране, обеспечивающим синтез углеводов, может быть описана следующим образом. Комплекс фотосистемы II ($PSII$) под действием света расщепляет молекулы воды, при этом освобожденные протоны выделяются в люмен, а два электрона по цепочке внутренних переносчиков в фотосистеме II передаются на пристыкованный к ней пластохинон (PQ). После этого пластохинон в форме PQH^2 отстыковывается и диффундирует к цитохромному b_6f комплексу. Последующее окисление PQ приводит к восстановлению пластоцианина (Pc) — подвижного белкового переносчика, способного перемещаться в люмене (слое между двумя фотосинтетическими мембранами), процесс происходит с участием цитохромного b_6f комплекса в соответствии со схемой Q-цикла Митчелла (рис. 2).

Комплекс фотосистемы I (PSI) под действием света катализирует окисление пластоцианина (Pc) на люминальной поверхности мембраны тилакоида и восстановление ферредоксина (Fd) на стромальной стороне мембраны. После этого происходит диффузия восстановленного ферредоксина в строму, заканчивающаяся передачей электрона на $NADP^+$ через экспонированный в строму комплекс ферредоксин-NADPH-редуктазы (FNR).

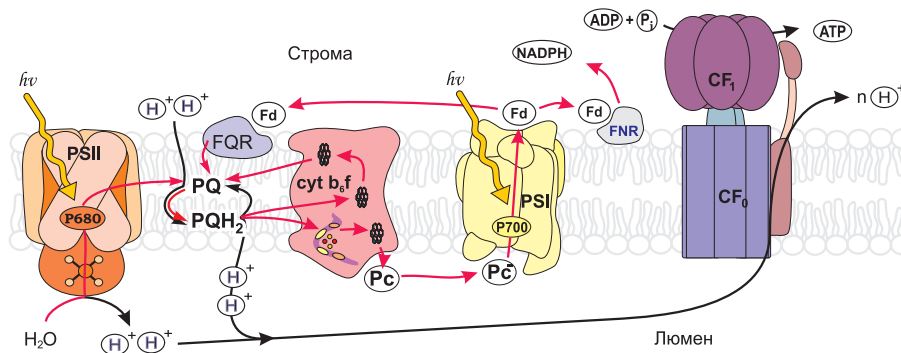


Рис. 2. Схема фотосинтеза. Пути переноса электронов (от H_2O до $NADP^+$, красные линии) и протонов (синие линии) при нециклическом электронном транспорте

Восстановленный $NADPH$ далее используется как источник электронов в цепи синтеза органических соединений (цикл Кальвина). Транспорт электрона также сопряжен с протонным транспортом. Протоны выделяются в люмен при расщеплении воды фотосистемой II, а также при работе Q-цикла Митчелла в цитохромном b_6f комплексе. В результате создается разность pH между люменом и стромой, за счет которой встроенные в мембрану комплексы АТФ-синтазы производят АТФ. Эти комплексы в хлоропласте расположены в основном на внешнем краю гранальных дисков, а также на ламеллах, выступающих в строму и соединяющих стопки гран друг с другом. Фотосистемы II сосредоточены в основном внутри гранальных дисков, цитохромные комплексы распределены равномерно по всей тилакоидной мембране. Таким образом, протоны закачиваются в люмен во внутренних областях граны, а выводятся в строму — на периферии. Следовательно, можно ожидать возникновения неоднородного распределения протонов в плоскости мембраны. Кроме того, после начала работы фотосинтетической цепи (например при включении освещения) синтез АТФ начинается не сразу, должно пройти некоторое время, пока в периферийных областях граны (где расположены АТФ-синтазы) накопится некоторое количество протонов, необходимое для синтеза АТФ.

Молекулы пластоцианина окисляют цитохром f и восстанавливают реакционный центр PSI , диффундируя в люминальном пространстве на довольно большие расстояния (сотни нм) и перенося электроны между гранальными и стромальными областями в тилакоидах [Мокронов, Гавриленко, 1992]; [Норе, 1993]; [Gross, 1996]. В нативном хлоропласте толщина люмена (4–10 нм [Mehta et al., 1999]) сравнима с размерами пластоцианина, а люмен заполнен выступающими частями трансмембранных мультиферментных комплексов [Норе, 2000], концентрация которых в тилакоидной мембране очень высока [Albertsson, 2001]. Как следствие, в люмене не обеспечиваются возможности свободной диффузии пластоцианина. В этом случае описание переноса электрона с помощью уравнений химической кинетики является весьма приближенным и не соответствует современным представлениям о гетерогенном характере мембраны. Необходимо поэтому рассматривать движение пластоцианина и других мобильных переносчиков в неоднородной среде, где диффузия затруднена и возможны столкновения с другими крупными молекулами. Рассмотрим основные методы моделирования фотосинтетических процессов.

Методы моделирования процессов в фотосинтетической мембране

Можно выделить несколько стадий транспорта электронов в фотосинтетической мембране, отличающихся своими механизмами, характерными временами и пространственными масштабами.

1. Разделение зарядов и перенос внутри белковых комплексов. Характерные времена лежат в диапазоне от 10^{-12} до 10^{-4} сек. Для моделирования используются кинетические модели, базирующиеся на аппарате обыкновенных дифференциальных уравнений. Переменными являются вероятности состояния комплекса [Stirbet et al., 1998]; [Ризниченко и др., 2009], [Лебедева и др., 2000]. При этом взаимодействие с подвижными переносчиками описывается законом действующих масс.
2. Диффузия подвижных переносчиков электрона в люмене и строме. Характерное время процесса определяется коэффициентом диффузии мобильных переносчиков. Например, для пластоцианина в люмене коэффициент диффузии равен 10^{-6} см²/с. Средний квадрат перемещения броуновской частицы связан со временем по формуле

$$\langle r^2 \rangle = Dt, \quad (1)$$

где D — коэффициент диффузии.

Характерное расстояние между белковыми комплексами на мембране — порядка 15 нм. По формуле (1) получаем, что среднее расстояние между комплексами переносчик будет преодолевать за время порядка $2.5 \cdot 10^{-7}$ с. Метод броуновской динамики позволяет моделировать электростатическое взаимодействие между белками и связывание одного подвижного переносчика с мультиферментным комплексом [Ullmann et al., 1997]; [Pearson, Gross, 1998]; [Ubbink et al., 1998]; [De Rienzo et al., 2001]; [Gross, Pearson, 2003]. В последние годы мы развиваем метод многочастичного прямого моделирования на основе броуновской динамики, позволяющий моделировать диффузию мобильных переносчиков в люмене и взаимодействие с пигмент-белковыми комплексами с учетом реалистичной геометрии белков и электростатических взаимодействий между ними [Riznichenko et al., 2010], [Коваленко и др., 2009], [Абатурова и др., 2009].

3. Диффузия протонов в люмене и их взаимодействие с буферными группами. Коэффициент диффузии протонов в воде равен 10^{-4} см²/с. Оценивая по формуле (1), получаем характерное время переноса протона от комплекса к комплексу, равное $2.5 \cdot 10^{-9}$ с. Таким образом, движение протонов на два порядка быстрее, чем движение мобильных переносчиков. Следовательно, непосредственное включение протонов в многочастичную модель при сохранении той же точности потребует уменьшения временного шага вычислений на два порядка, что неэффективно с вычислительной точки зрения. Более перспективно моделировать перенос протонов с помощью дифференциальных уравнений в частных производных [Вершубский и др., 2007].

В системах, где существует несколько процессов, протекающих с сильно отличающимися скоростями, можно использовать принцип узкого места и применить теорему Тихонова для уменьшения размерности задачи. Однако, для этого требуется наличие четкой иерархии времен. При моделировании фотосинтетической системы как целого, с учетом пространственной организации, это сделать сложно, поскольку времена различных стадий электронного транспорта сравнимы между собой. Вместе с тем, сложно смоделировать всю систему в целом в рамках одного подхода — описание с помощью кинетических уравнений не учитывает пространственной организации системы, многочастичные модели не рассматривают процессы внутри комплексов, системы уравнений в частных производных не учитывают дискретный характер системы. Для того чтобы решить эти проблемы, в настоящей работе мы предлагаем подход, при котором разные стадии в рамках единой модели рассчитываются разными методами. Перенос электрона внутри комплекса описывается системой обыкновенных дифференциальных уравнений, где переменными являются вероятности состояний комплекса. Перенос электрона мобильными переносчиками моделируется методом многочастичной броуновской динамики, где каждая молекула

переносчика и каждый белковый комплекс моделируется как объект в трехмерном пространстве. Диффузия протонов описывается системой дифференциальных уравнений в частных производных относительно концентраций протонов в люмене и строме. Решение уравнений диффузии протонов осуществляется сеточными методами или методом частиц. Комплексы АТФ-синтазы также моделируются как отдельные объекты со своими внутренними состояниями. Целью является построение модели, сочетающей детальный учет пространственной организации мембраны, степень детализации при рассмотрении отдельных стадий и скорость вычислений, достаточную для моделирования процессов на одной или нескольких гранах хлоропласта (порядка 10 000 белковых комплексов и переносчиков электрона) с общей длительностью порядка секунд. Современные средства параллельных вычислений — высокопроизводительные кластеры с обычными и графическими процессорами — позволяют обеспечить достаточное быстродействие, при условии наличия эффективных параллельных алгоритмов. Основная сложность — обеспечить сопряжение разных методов моделирования отдельных стадий: переноса внутри комплексов, движения мобильных переносчиков, транспорта протонов. Трудность состоит в том, что для описания каждой из этих стадий используется своя степень детализации и свой математический аппарат. Так, например, каждый белковый комплекс рассматривается как твердое тело с некоторым распределением электрического заряда при моделировании движения мобильных переносчиков, а при моделировании переноса в комплексе тот же белок рассматривается как вектор состояний внутренних переносчиков электрона. При построении единой модели нужно согласовать между собой эти подходы. Для этого предлагается следующая схема функционирования модели.

Пусть в некоторый момент времени нам известно состояние всех моделируемых объектов. Выполняются следующие операции:

1. Моделирование перемещения переносчиков путем решения уравнение Ланжевена. Рассчитываются новые координаты переносчиков на следующем шаге броуновской модели. Размеры объектов, рассматриваемых в модели, лежат в интервале от 1 до 1 000 нм, т. е. в пределах от размера одного белка-переносчика до размера грани хлоропласта. Для того чтобы аккуратно учесть столкновения с другими объектами, среднее перемещение белка-переносчика за один шаг модели должно быть примерно на порядок меньше его линейных размеров (2 нм). Используя формулу (1), получаем, что временной шаг, обеспечивающий достаточную точность моделирования перемещения мобильных переносчиков, определяется их коэффициентом диффузии (порядка 10^{-6} см²/с), и составляет 10^{-8} с. Дальнейшее уменьшение шага не имеет смысла, т. к. не приводит к заметному увеличению точности моделирования, хотя время расчета при этом сильно увеличится.
2. Расчет переходов между внутренними состояниями для каждого из белковых комплексов согласно матрице вероятностей перехода, рассчитываемой через решение системы дифференциальных уравнений. Определяется, как изменятся внутренние состояния всех комплексов на следующем шаге модели броуновского движения.
3. Если изменение состояния некоторых комплексов связано с выделением протонов в люмен или их утечкой, пересчитываются соответствующие плотности вероятности для протонов. Решая уравнение движения для протонов, рассчитываем изменение концентрации за время, равное одному шагу многочастичной модели.

Методы и алгоритмы

Движение мобильных переносчиков моделируется на трехмерной сцене (рис. 3), в которую включены строма (внутреннее пространство хлоропласта вне фотосинтетической мембраны), двойная мембрана тилакоида и межмембранное пространство — люмен. Встроенные в мембрану мультиферментные белковые комплексы и мобильные переносчики электрона описываются

как твердые тела с определенной геометрической формой (определяемой по структуре из Protein Data Bank), расположенные в трехмерном пространстве. Каждый объект в модели обладает набором свойств, таких как координаты, геометрическая форма и размеры, наличие или отсутствие переносимых электронов, параметры взаимодействия с другими объектами.

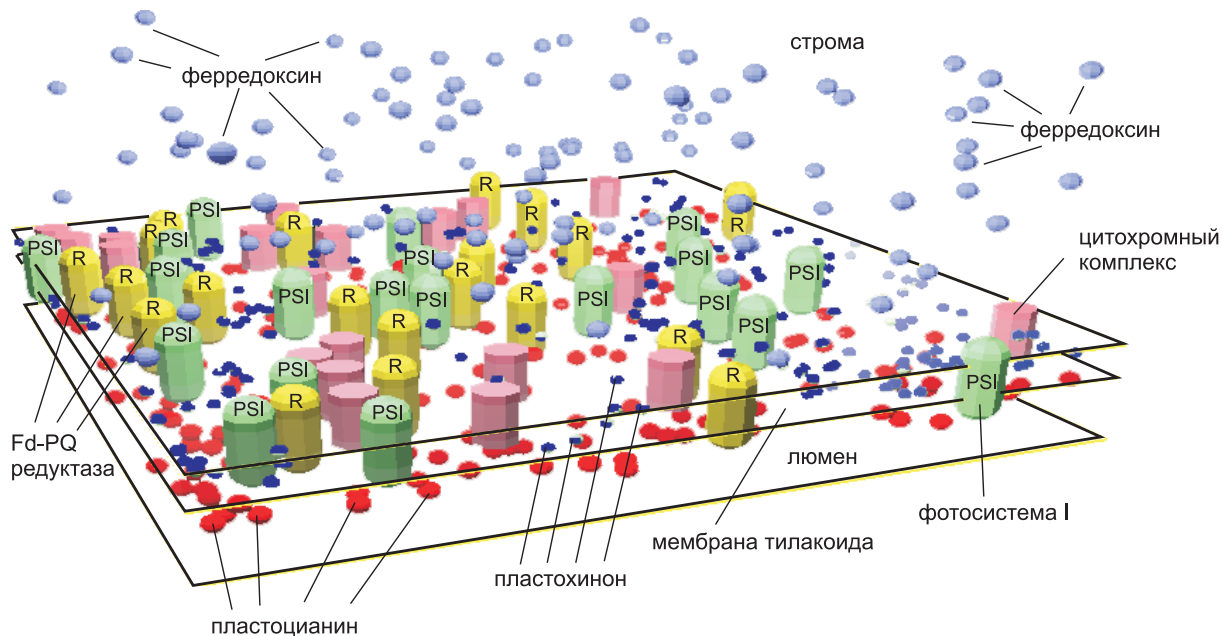


Рис. 3. Модельная сцена, представляющая участок фотосинтетической мембраны

Алгоритм формирования сцены и генерации псевдослучайного расположения комплексов на мембране подробно описан в нашей предыдущей работе [Устинин и др., 2010].

Моделирование движения мобильных переносчиков

Для моделирования движения Pc , Fd , PQ в пространствах соответствующих компартментов использовали математический аппарат описания броуновского движения с учетом геометрических ограничений, налагаемых сформированной модельной сценой. Правомерность рассмотрения движения белковых переносчиков как броуновских частиц обусловлена их глобулярностью. Предполагается, что движение частицы происходит в вязкой среде под действием случайной силы, возникающей из-за столкновений с молекулами среды. Для описания такого процесса можно использовать уравнение Ланжевена, описывающее изменение каждой координаты со временем под действием случайной силы

$$\xi \frac{dx}{dt} = f(t). \quad (2)$$

Здесь $f(t)$ — случайная сила, распределенная по Гауссу с нулевым средним и дисперсией, равной $2kT\xi$, k — постоянная Больцмана, T — температура, ξ — коэффициент трения в среде, вычисляемый (в предположении о сферичности частицы) по формуле $\xi = 6\pi\eta a$, где η — вязкость среды, a — радиус частицы. Уравнение решается численно для каждой из частиц. Если выбрать шаг по времени, равный Δt , в разностном виде это уравнение будет записано как

$$\Delta x = \frac{f(t)\Delta t}{\xi}. \quad (3)$$

Таким образом, перемещение частицы за один шаг на расстояние Δx является случайной величиной, распределенной по Гауссу с нулевым средним и дисперсией, равной

$$\langle \Delta x^2 \rangle = \frac{2kT\Delta t^2}{\xi}. \quad (4)$$

Итак, Δx можно представить как

$$\Delta x = \sqrt{\frac{2kT}{\xi}} N(0, 1) \Delta t. \quad (5)$$

Здесь $N(0, 1)$ — случайная величина, имеющая стандартное нормальное распределение. Если теперь вычислить средний квадрат перемещения за время t , получим

$$\langle r^2 \rangle = \langle x^2 + y^2 + z^2 \rangle = 3 \left(\frac{2kT}{\xi} \Delta t \right) t = Dt. \quad (6)$$

Здесь $D = \frac{6kT\Delta t}{\xi}$ — коэффициент диффузии. Итак, перемещение за один шаг и коэффициент диффузии связаны следующим соотношением, которое и используется в программе для расчета:

$$\Delta x = \sqrt{\frac{D\Delta t}{3}} N(0, 1). \quad (7)$$

Шаг по времени подбирается таким образом, чтобы корень из дисперсии перемещения частицы на каждом шаге (среднее перемещение частицы на каждом шаге) был порядка одной десятой диаметра мобильного переносчика. Такой выбор шага обеспечивает приемлемые точность вычислений и время расчета. На боковых границах области моделирования использованы тороидальные (периодические) граничные условия, также учитывалось отражение частиц от физических поверхностей, включая мембрану и белковые комплексы. Каждый из участников движения может переносить или не переносить электрон, что при визуализации динамики системы изображается условным изменением цвета частицы.

Для генерации случайных чисел использовался генератор L'Esuyer с периодом 2^{191} . В версии алгоритма для графических процессоров использовался параллельный генератор Mercenne-Twister.

Учет столкновений и взаимодействие с белковыми комплексами

Моделирование столкновений и взаимодействия с другими частицами подробно описано в работе [Устинин и др., 2010]. Стыковка мобильных переносчиков с белковыми комплексами описывается через эффективный радиус взаимодействия r и вероятность посадки p . У каждого комплекса есть один или несколько сайтов посадки, к которым могут пристыковаться определенные типы переносчиков.

Если на данном шаге переносчик подошел к соответствующему сайту посадки на комплексе ближе чем на расстояние r , то разыгрывается случайное число, равномерно распределенное на $[0; 1]$, и если оно меньше, чем величина вероятности стыковки p , считается, что переносчик состыковался с комплексом. Если стыковка не произошла, переносчик продолжает движение. После стыковки переносчик остается неподвижным на сайте посадки до тех пор, пока не произойдет передача электрона или случайная отстыковка. Радиус взаимодействия r и вероятность стыковки p являются параметрами модели. Радиус взаимодействия выбирается по данным о дальности действия электростатических сил между комплексами и переносчиками. Этот радиус составляет примерно 8 нм. Для определения параметров модели были использованы результаты более подробного моделирования с явным учетом электростатических взаимодействий [Коваленко и др., 2009]. Вероятность стыковки прямому определению не поддается и оценивается по

экспериментальным данным. Для того чтобы оценить вероятность стыковки, используются данные о скоростях реакций отдельных переносчиков и комплексов в растворе. Например, отдельно рассматривается реакция пластоцианина с цитохромом f . По измеренной экспериментальной кривой оценивается константа реакции, затем параметр p — вероятность стыковки пластоцианина с цитохромом — подбирается так, чтобы константа реакции, вычисленная по результатам моделирования, совпадала с экспериментальной. Аналогично, по данным экспериментов в растворе определяются параметры стыковки для пар ферредоксин — фотосистема I , пластоцианин — фотосистема I .

Моделирование переноса электрона внутри белковых комплексов

Представление белковых комплексов как элементарных объектов многочастичной модели не позволяет подробно учитывать процессы переноса заряда внутри комплексов фотосистем и цитохрома. Однако эти процессы необходимо учитывать, чтобы корректно смоделировать сигналы, получаемые в экспериментах по изучению флуоресценции, данные электронного парамагнитного резонанса и спектральных измерений. Эти сигналы зависят от состояния промежуточных переносчиков электрона в фотосистемах и цитохромном комплексе. Следовательно, требуется более подробное рассмотрение комплексов и их состояний. Для моделирования процессов внутри комплексов разработан удобный аппарат описания системами обыкновенных дифференциальных уравнений, где переменными выступают вероятности состояний комплекса [Rubin, Riznichenko, 2009]. Проблема заключается в том, чтобы состыковать многочастичную модель движения мобильных переносчиков с моделью переноса внутри комплексов. Количество комплексов в многочастичной модели велико, поэтому решать для каждого из них систему уравнений для перехода между внутренними состояниями на каждом шаге многочастичной модели невозможно, да и не имеет смысла. На самом деле, интересно только, в каком состоянии окажется каждый комплекс в конце очередного шага многочастичной модели, на каком переносчике находится заряд, произошел ли перенос заряда от мобильного переносчика или на него, произошел ли акт флуоресценции и т. п. Число промежуточных переносчиков электрона, моделируемых для каждого комплекса, конечно, и каждый переносчик может находиться в ограниченном числе состояний (обычно в двух — с электроном либо без электрона).

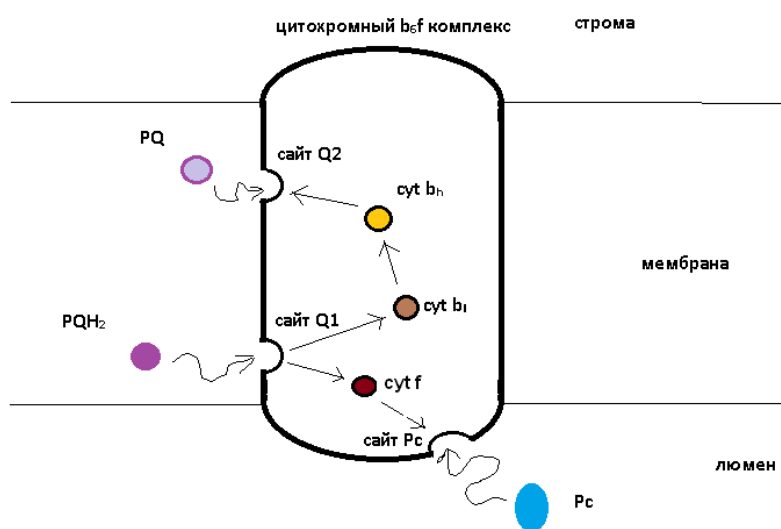


Рис. 4. Перенос электрона в цитохромном b_6f комплексе

В качестве примера рассмотрим модель цитохромного b_6f комплекса (рис. 4). Мы выделяем следующие промежуточные переносчики электрона: цитохром f , низкопотенциальный гем цитохрома $b - b_h$, высокопотенциальный гем цитохрома $b - b_l$. Каждый из этих переносчиков может находиться в одном из двух состояний — восстановленном (с электроном) или окисленном (без электрона). Сайты посадки пластохинона могут находиться в следующих состояниях:

1. переносчик не пристыкован;
2. пристыкован окисленный пластохинон;
3. пристыкован семихинон с одним электроном;
4. пристыкован восстановленный пластохинол PQH_2 с двумя электронами.

Итого по четыре возможных состояния на каждый из двух сайтов посадки пластохинона. Аналогично, сайт посадки пластоцианина может находиться в одном из трех возможных состояний:

1. пластоцианин не пристыкован;
2. пристыкован окисленный пластоцианин;
3. пристыкован восстановленный пластоцианин.

Таким образом, состояние комплекса «пластоцианин–цитохром b_6f –пластохинон» описывается вектором из 6 состояний каждого промежуточного переносчика электрона и сайтов посадки Pc и PQ . Общее количество возможных состояний комплекса можно вычислить, перемножив количество возможных состояний каждого элемента: 4 (люминальный сайт посадки PQ) \times 4 (стромальный сайт посадки PQ) \times 3 (сайт посадки Pc) \times 2 ($Cyt f$) \times 2 (гем b_h) \times 2 (гем b_l). Итого, получаем 384 различных возможных состояния комплекса. Перенос электрона с одного промежуточного переносчика на другой описывается сменой состояния комплекса. Например, на очередном шаге многочастичной модели движения переносчиков была зафиксирована стыковка восстановленного пластохинола PQH_2 с люминальным сайтом посадки цитохромного комплекса. Вектор состояния комплекса будет выглядеть следующим образом: (люминальный сайт посадки PQ — в состоянии 4 (пристыкован PQH_2), все остальные промежуточные переносчики — в состоянии 0). Если за следующий шаг модели броуновской динамики произошел перенос электрона с PQH_2 на цитохром f , вектор состояния изменится следующим образом: (люминальный сайт посадки PQ — в состоянии 3, $Cyt f$ — в состоянии 1, все остальные промежуточные переносчики останутся в состоянии 0). Для того чтобы смоделировать такой переход, нужно предварительно рассчитать вероятность этого перехода за время, равное одному шагу модели броуновской динамики. Для этого нужно записать систему обыкновенных дифференциальных уравнений, описывающих реакции переноса электрона внутри комплекса. Константы соответствующих реакций можно найти в литературе по кинетическим моделям [Chow, Nore, 2004]. Для каждого из возможных начальных состояний системы задаются начальные условия для дифференциальных уравнений. Так, для цитохромного комплекса получается 384 варианта начальных условий. Решая систему уравнений, получаем для каждого из этих начальных условий распределение состояний системы в конце заданного отрезка времени (например, шага многочастичной модели). Получив такие распределения для всех начальных состояний, мы вычисляем матрицу переходов размерности 384×384 элемента. На главной диагонали матрицы стоят вероятности остаться в текущем состоянии, выше главной диагонали — вероятности прямых реакций, ниже — обратных.

Имея такую матрицу, можно быстро моделировать результат переноса электрона внутри комплекса за один шаг броуновской динамики для каждого комплекса. Для расчета матрицы

перехода была разработана программа автоматической записи и решения системы дифференциальных уравнений по кинетической схеме переходов внутри комплекса. Для решения системы использовался метод Рунге–Кутты 4-го порядка. Решая систему уравнений с начальными условиями, в которых i -е состояние в начальный момент имеет вероятность 1, а остальные состояния — вероятность 0, получали вероятность состояния j через время, равное одному шагу многочастичной модели. Матрица перехода рассчитывается один раз для каждого типа комплексов перед началом вычислений по полной модели. В начале каждого шага многочастичной модели определяется, в каком состоянии находится каждый комплекс (с учетом пристыкованных переносчиков), разыгрывается случайное число от 0 до 1 и по соответствующей строке матрицы переходов определяется, в каком состоянии комплекс окажется в следующий момент. Если этот переход сопровождается флуоресценцией (при моделировании фотосистемы 2), имитируется испускание соответствующего кванта. Такой подход позволяет без существенной потери скорости вычислений детально учесть переходы внутри каждого комплекса и корректно смоделировать сигналы с внутренних акцепторов электрона, измеряемые в эксперименте. Таким образом, стыковка многочастичной модели движения переносчиков и кинетической модели переноса внутри комплексов осуществляется путем задания набора возможных состояний для каждого комплекса. Состояние комплекса определяется наличием пристыкованных мобильных переносчиков на сайтах посадки и наличием электронов на внутренних акцепторах. За один шаг многочастичной модели рассчитывается изменение состояния каждого комплекса, которое происходит двумя путями — через стыковку/расстыковку мобильного переносчика и через переходы между внутренними акцепторами, которые рассчитываются через матрицу перехода.

Вычислительный эксперимент

В качестве примера было проведено моделирование экспериментальной ситуации, описанной в работе [Chow, Nore, 2004]. Исследовался перенос электрона от фотосистемы 2 к цитохромному комплексу и затем на фотосистему 1 через пластоцианин. В экспериментах листья огурцов сначала выдерживались в темноте, затем подвергались миллисекундной лазерной вспышке. После этого в течение пяти секунд измерялось восстановление акцептора $P700$ в фотосистеме I, а также восстановленность гема b_h и цитохрома f в цитохромном комплексе (рис. 5 и 6). В численном эксперименте была сделана попытка воспроизвести эту экспериментальную ситуацию. Была сгенерирована модельная сцена, содержащая фотосистемы 2 в гранальной области, фотосистемы I в стромальной области и цитохромные комплексы, равномерно распределенные между обоими областями. Предполагалось, что сначала восстановленные переносчики электронов в системе отсутствуют, таким образом имитировалось длительное выдерживание системы в темноте. Далее моделировалась вспышка в 1 мс, затем свет выключался и рассчитывалось количество восстановленных $cyt\ b$ и $cyt\ f$ в системе. Результаты моделирования показаны на рисунках 5 и 6. Численный эксперимент показал хорошее совпадение с экспериментальными данными.

Моделирование протонного транспорта

Рассматривая процесс протонного транспорта в фотосинтетической системе, мы учитываем процессы, приводящие к изменению активности во внутритилакоидном пространстве. К ним относится выделение протонов в люмен, латеральная диффузия и утечка протонов. Включать протоны как частицы в модель движения мобильных переносчиков нецелесообразно, т. к. существенно отличаются скорости перемещения (коэффициент диффузии протонов в воде порядка 10^{-4} см²/с, на два порядка больше, чем коэффициент диффузии белков-переносчиков), и для моделирования каждого протона понадобится более мелкий шаг по времени. Поэтому, мы используем аппарат дифференциальных уравнений в частных производных, предложенный в работе [Вершубский и др., 2007]. При этом переменными в уравнениях являются концентрации

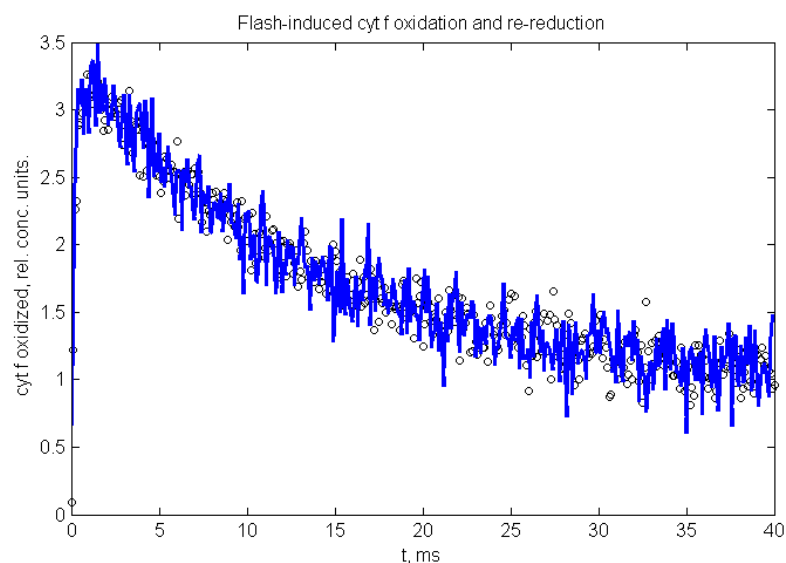


Рис. 5. Кривые окисления и последующего восстановления цитохрома *f* после вспышки. Черные точки — эксперимент [Chow, Норе, 2004]. Синяя сплошная линия — результаты моделирования

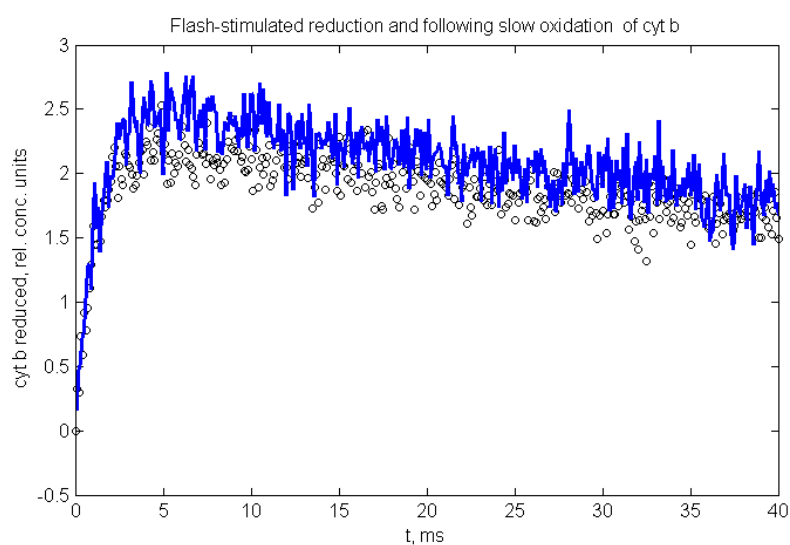


Рис. 6. Экспериментальная кривая восстановления и последующего окисления цитохрома *b*. Черные точки — эксперимент [Chow, Норе, 2004]. Синяя сплошная линия — результаты моделирования

протонов в люмене и строме, зависящие от координат и времени. Однако в живой биологической системе в реальных условиях общее количество протонов в одной грани тилакоида невелико и измеряется десятками, а иногда и единицами штук на грану. Поэтому возникает вопрос о правомерности использования самого понятия концентрации при столь малом количестве частиц. В настоящей работе мы используем интерпретацию, предложенную в работе [Тихонов, Блюменфельд, 1985], согласно которой речь идет не о концентрации, а о плотности вероятности распределения протонов: т. е., переменная в уравнении интерпретируется как плотность вероятности обнаружить протон в данной точке пространства.

Совмещение дискретной модели движения переносчиков с диффузионной моделью переноса протонов осуществляется следующим образом. Задается трехмерная сетка, покрывающая область модельной сцены, в которой нас интересует перенос протонов. Ширина ячеек сетки примерно соответствует диаметру белковых комплексов — порядка 10 нм. При этом в каждую ячейку попадает не более одного белкового комплекса — потенциального источника протонов. Для каждого протона в системе задается сеточная функция, определяющая вероятность обнаружить его в данной ячейке. В конце каждого шага многочастичной модели анализируются переходы между состояниями комплексов, произошедшие за этот шаг. Некоторые из этих переходов сопровождаются выделением протонов в люмен. В случае, если такой переход зафиксирован у некоторого комплекса, мы добавляем в систему новый протон. Функция распределения вероятности для него задается следующим образом: вероятность равна 1 в ячейке, соответствующей комплексу, через который произошло выделение протона, и нулю во всех остальных ячейках. Затем на каждом шаге многочастичной модели мы пересчитываем распределение вероятности всех протонов, решая для каждого из них конечно-разностную систему уравнений, описывающую диффузию и взаимодействие с буфером (10).

Стыковка многочастичной модели с конечно-разностной моделью диффузии протонов также необходима при моделировании утечки через АТФ-синтазу. Она выполняется следующим образом: для каждого комплекса АТФ-синтазы известны его координаты, и, следовательно, ячейка сетки, в которой он находится. Для каждого протона мы знаем вероятность обнаружить его в данной ячейке. В конце шага многочастичной модели для каждой АТФ-синтазы и для каждого протона разыгрывается случайное число, равномерно распределенное на отрезке $[0; 1]$. Если выпавшее число меньше, чем вероятность нахождения данного протона в ячейке вблизи АТФ-синтазы, моделируется, утечка протона. Протон исключается из рассмотрения в люмене и добавляется для рассмотрения в строме, а для АТФ-синтазы симулируется поворот подвижной субъединицы γ на 120 градусов. Когда произойдет три таких поворота, будет синтезирована молекула АТФ.

Модель выделения протонов в люмен

Плотность вероятности для каждого свободного протона в люмене определяется переменной N , заданной на прямоугольной сетке размерности $NX \times NY \times NZ$, причем размер ячейки выбирается порядка среднего диаметра белковых комплексов (10 нм). Поскольку толщина люмена тоже порядка 10 нм, при моделировании переноса протонов в люмене достаточно одной ячейки по оси z (перпендикулярной плоскости мембраны). Фактически диффузия протонов в люмене моделируется на двумерной сетке. Характерный диаметр грани — 500 нм, так что размерность сетки для одной грани составит $50 \times 50 \times 1$ ячейку.

Выделение протонов происходит в результате индуцированного светом разложения воды в ФС2, а также за счет окисления пластохинола (QH_2) цитохромным b_6f комплексом. В многочастичной модели эти события описываются через изменение состояний комплексов к конце каждого временного шага модели. Зная интенсивность света, можно рассчитать вероятность того, что какой-либо комплекс фотосистемы 2 поглотит квант света за время, равное длительности шага многочастичной модели. Рассчитать эту вероятность можно по формуле

$$p = \frac{I\sigma t}{h\nu}, \quad (8)$$

где I — интенсивность света (порядка 1000 Вт/м^2 в солнечный день), σ — эффективное сечение поглощения, t — длительность шага многочастичной модели, h — постоянная Планка, ν — частота поглощаемого света.

При интенсивности света, соответствующей яркому солнечному дню, и длительности шага модели в 100 наносекунд, вероятность p — величина порядка 0.0001. В конце каждого шага многочастичной модели для каждого из комплексов фотосистемы 2 разыгрывается случайное число из равномерного распределения на $[0; 1]$. Если оно оказалось меньше величины вероятности поглощения p , то считается, что данный комплекс фотосистемы 2 поглотил на данном шаге квант света. Так как мы знаем координаты комплекса, мы также знаем ячейку сетки для плотности вероятности протонов, в которую этот комплекс попадает.

Теперь добавляем в систему новый протон. Его функция плотности вероятности будет выглядеть следующим образом: в ячейке сетки, в которой находится выделившийся комплекс, плотность вероятности нахождения протона в данной ячейке равна:

$$\Delta H = \frac{1}{dV(1 + KB)}. \quad (9)$$

Здесь dV — объем ячейки, K — константа равновесия для реакции связывания протонов с буферными группами, B — концентрация буферных групп.

Во всех остальных ячейках вероятность обнаружить данный протон равна нулю (т. к. он только что был выделен и еще не успел никуда продиффундировать).

В этой формуле учтено взаимодействие ионов водорода с буферными группами, фиксированными на внутренней и внешней поверхностях тилакоидной мембраны. Физически буферные группы представляют собой выступающие в люмен заряженные аминокислотные остатки в белковых цепях многочисленных комплексов, встроенных в мембрану. Протоны могут временно связываться с этими остатками, что существенно замедляет их диффузию в люмене и примерно в 1000 раз уменьшает концентрацию свободных протонов в водной фазе люмена. В связи с этим, необходимо учитывать наличие буферных групп при моделировании диффузии протонов.

Модель латеральной диффузии протонов

Моделируется диффузия в люмене (межмембранном пространстве). Геометрически рассматриваемая область представляет собой прямоугольный параллелепипед, где верхняя и нижняя его стороны отображают плоскости фотосинтетической мембраны, через которую происходит утечка протонов.

В рассматриваемой нами системе ионы водорода могут диффундировать в водной фазе внутритилакоидного объема. Значение рН в строме (pH_s) в данной версии модели принимается постоянным.

Для моделирования процесса латерального переноса протонов во внутритилакоидном пространстве используется система дифференциальных уравнений, учитывающая связывание с буферными группами:

$$\frac{\partial H}{\partial t} = D_H \nabla^2 H - k_1 H(B_0 - c) + k_2 c + J, \quad (10)$$

$$\frac{\partial c}{\partial t} = k_1 H(B_0 - c) - k_2 c, \quad (11)$$

$H(r, t)$ — плотность вероятности нахождения протона в данной точке люмена; $c(r, t)$ — плотность распределения комплексов «протон–буферная группа»; B_0 — максимальная концентрация буферных групп (в данной версии модели предполагается не зависящей от координат и времени); k_1 , k_2 — константы связывания и распада; D_H — коэффициент диффузии свободных протонов в воде; J — сумма притока от фотосистем 2 и цитохромных b_6f комплексов и оттока через АТФ-синтазу.

Граничные условия задаются как непроницаемые стенки на верхней и нижней плоскостях люмена и периодические — на боковых границах. Численное решение поставленной задачи выполнялось методом конечных разностей.

Моделирование утечки протонов через АТФ-синтазу и синтез АТФ

Каждый комплекс АТФ-синтазы моделируется как отдельный объект в трехмерном пространстве, расположенный на мембране тилакоида. На каждом шаге многочастичной модели мы имеем дело с вероятностью обнаружить какой-либо протон в ячейке, соседней с выбранной АТФ-синтазой. Если на данном шаге выпадает случайное число, меньшее этой вероятности, мы полагаем, что протон в данной ячейке присутствует и с некоторой вероятностью p_f (которая является параметром модели) может произойти утечка протона через данную АТФ-синтазу. При этом имитируется поворот подвижной субъединицы γ («ротора») на 120 градусов внутри неподвижной субъединицы α («статора»). Когда через АТФ-синтазу проходят три протона, происходит синтез одной молекулы АТФ. На данном этапе мы не учитывали, сколько в системе АДФ и свободного фосфата, предполагая, что синтез АТФ лимитируется наличием протонного градиента.

Программная реализация

В качестве платформы для программной реализации были выбраны C++, как удобный объектно-ориентированный язык программирования. Для реализации параллельных вычислений использовались среда MPI для вычислений на кластерах, и среда программирования CUDA C для вычислений на графических процессорах. Основные вычислительные блоки, которые подвергались распараллеливанию — генерация случайных чисел, моделирование броуновского движения частиц, решение уравнения диффузии для протонов методом частиц.

Вычислительный эксперимент. Моделирование синтеза АТФ

Исследование переноса протонов в плоскости мембраны при фотосинтезе производилось в следующем вычислительном эксперименте. В гранальной области мембраны (справа-снизу на рисунке 7а) располагались фотосистемы 2 (зеленые точки на рисунке), в стромальной области (слева-сверху) — АТФ-синтазы и фотосистемы 1. Цитохромные комплексы (красные точки) располагались равномерно по всей плоскости мембраны. Сначала система некоторое время выдерживалась в темноте, затем включалось освещение. При включении света фотосистемы 2 начинают поглощать кванты, происходит разложение воды и выделение протонов в люмен. Когда новый протон выделяется в люмен водоразлагающим комплексом фотосистемы 2 или цитохромным комплексом, его плотность вероятности представляет собой дельта-функцию, локализованную в пространстве вблизи комплекса, выделившего протон. На рисунке 7а это можно видеть как пики в правой части графика. Со временем протон начинает диффундировать в пространстве, график его плотность вероятности становится более пологим и напоминает колокол гауссианы. В гранальной части люмена количество протонов возрастает, они начинают диффундировать в стромальную часть. По мере накопления протонов создается градиент между стромальной и гранальной частями (рис. 7b). Когда вероятность обнаружить протоны в стромальной области становится заметной величиной, начинается синтез АТФ, связанный с утечкой протонов через АТФ. Зависимость количества синтезированной АТФ от времени показана на рисунке 8а. Сначала можно наблюдать лаг-период, когда вероятность обнаружить протон в стромальной области мала. Затем скорость синтеза возрастает, и кривая выхода АТФ становится линейной. На рисунке 8b показано суммарное количество протонов в люмене. Оно сначала нарастает, затем приток от фотосистем 2 и цитохромных комплексов сравнивается с утечкой через АТФ-синтазы, и распределение протонов в люмене стабилизируется.

Учет неоднородного распределения комплексов на мембране и моделирование диффузии переносчиков электрона и протонов позволяют смоделировать эффект возникновения задержки между включением освещения и началом синтеза АТФ. Задержка при этом обуславливается

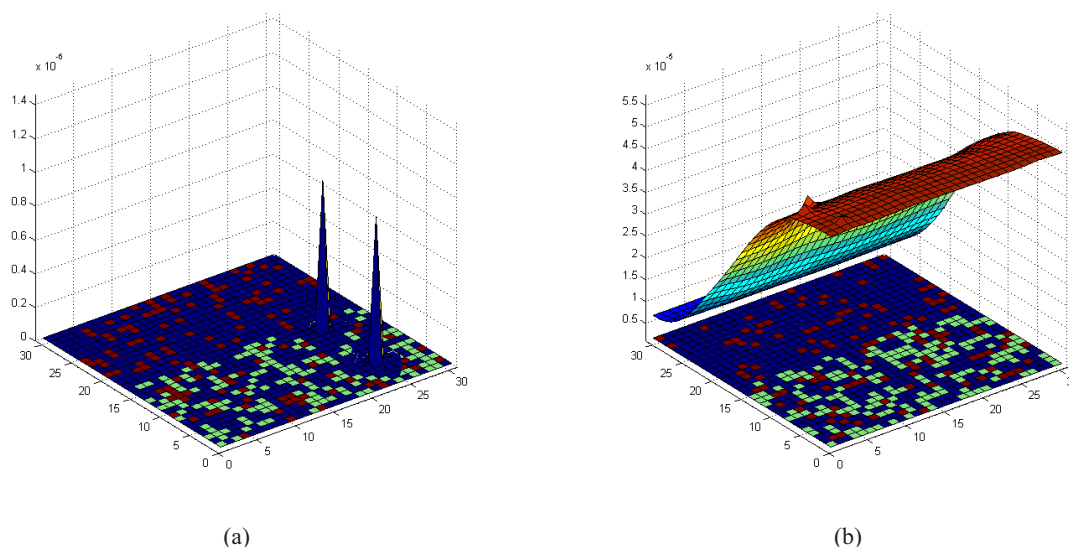


Рис. 7. Распределение плотности вероятности протонов в плоскости мембраны: а — в начале освещения; б — через 50 мс после начала освещения. Зеленые точки отмечают положение фотосистем 2 в плоскости мембраны, красные отмечают положение цитохромных комплексов. Пики соответствуют положению фотосистем или цитохромных комплексов, через которые произошло выделение протонов на текущем шаге модели. Через 50 мс после начала освещения сформировался градиент концентрации между гранальной областью (справа внизу), содержащей фотосистемы 2, и стромальной (слева вверху), не содержащей их

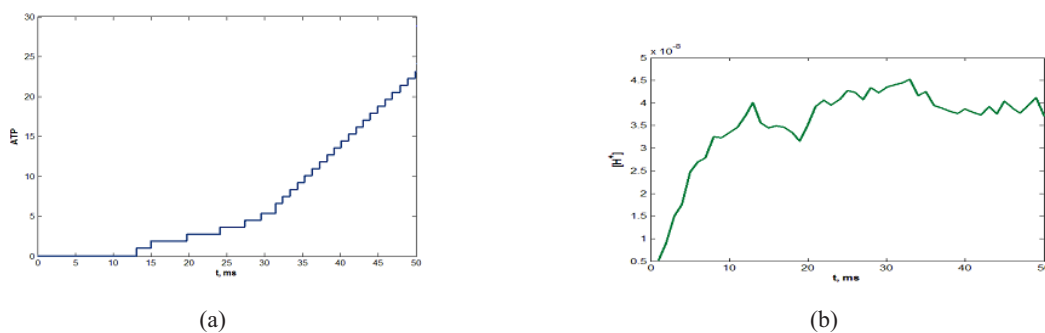


Рис. 8. а — с ростом рН в стромальной области АТФ-синтаза начинает производить АТФ. Сначала скорость синтеза лимитируется количеством доступных протонов, затем — временем поворота субъединиц. б — общее количество протонов в люмене сначала нарастает, а со временем выходит на постоянный уровень

несколькими факторами — конечной скоростью диффузии переносчиков электрона из гранальной части в стромальную, достаточно медленными характерными временами переноса электрона в цитохромном комплексе, сопряженного с протонным транспортом, а также конечной скоростью диффузии протонов в люмене. В результате накопление протонов в стромальной части, где находятся АТФ-синтазы, происходит с задержкой, что и приводит к задержке начала синтеза АТФ. Модель позволяет одновременно учитывать эти факторы в едином пространственно-временном масштабе.

Выводы

Предложен новый метод моделирования, позволяющий адекватно учесть пространственную структуру биологических энергопреобразующих мембран. В многочастичном подходе к моделированию используется различный математический аппарат для моделирования разных процессов в мембране — броуновскую динамику для движения мобильных переносчиков, уравнения химической кинетики — для реакций внутри белковых комплексов, уравнения в частных производных — для моделирования диффузии протонов. Объединение этих подходов в единую модель позволило найти компромисс между требованиями к аккуратности моделирования отдельных стадий, протекающих на разных пространственных и временных масштабах, и скорости расчетов, приемлемой для современной вычислительной техники. Программная реализация метода позволяет эффективно использовать возможности современных параллельных вычислений.

Список литературы

- Абатурова А. М., Коваленко И. Б., Ризниченко Г. Ю., Рубин А. Б. Исследование образования комплекса флаводоксина и фотосистемы I методами прямого многочастичного компьютерного моделирования // *Компьютерные исследования и моделирование*. — 2009. — Т. 1. — С. 85–91.
- Вершубский А. В., Приклонский В. И., Тихонов А. Н. Математическая модель электронного и протонного транспорта в фотосинтетических системах оксигенного типа // *Российский химический журнал*. — 2007. — Т. LI.
- Коваленко И. Б., Абатурова А. М., Ризниченко Г. Ю., Рубин А. Б. Новый метод компьютерного моделирования образования белок-белковых комплексов // *Доклады Академии наук*. — 2009. — Т. 427. — С. 696–698.
- Лебедева Г. В., Беляева Н. Е., Ризниченко Г. Ю. и др. Кинетическая модель фотосистемы II высших растений // *Журнал физической химии*. — 2000. — Т. 74. — С. 1874–1883.
- Мокроносов А.Т., Гавриленко В.Ф. Фотосинтез. Физиолого-экологические и биохимические аспекты. — Москва: Изд-во МГУ, 1992.
- Ризниченко Г. Ю., Беляева Н. Е., Коваленко И. Б., Рубин А. Б. Математическое и компьютерное моделирование первичных процессов фотосинтеза // *Биофизика*. — 2009. — Т. 54. — С. 16–33.
- Тихонов А. Н., Блюменфельд Л. А. Концентрация водородных ионов в субклеточных частицах: физический смысл и методы определения // *Биофизика*. — 1985. — Т. 30. — С. 527–537.
- Устинин Д. М., Коваленко И. Б., Грачев Е. А. и др. Метод прямого многочастичного компьютерного моделирования фотосинтетической электронно-транспортной цепи // *Динамические модели процессов в клетках и субклеточных наноструктурах*. — 2010. — С. 241–262.
- Albertsson P. A. A quantitative model of the domain structure of the photosynthetic membrane // *Trends in Plant Science*. — 2001. — Vol. 6. — Pp. 349–354.
- Chow W., Hope A. B. Kinetics of reactions around the cytochrome bf complex studied in intact leaf disks // *Photosynthetic Research*. — 2004. — Vol. 81. — Pp. 153–163.
- De Rienzo F., Gabdouliline R. R., Menziani M. C. et al. Electrostatic analysis and Brownian dynamics simulation of the association of plastocyanin and cytochrome f // *Biophysical Journal*. — 2001. — Vol. 81. — P. 3090–3104.
- Gross E. L. Plastocyanin: structure, location, diffusion and electron transfer mechanisms. In *Biochemistry and Molecular Biology of Plants*, edited by R. Jones. — Kluwer Academic Publishers, 1996. — Pp. 413–429.
- Gross E. L., Pearson Jr. D. C. Brownian Dynamics Simulations of the Interaction of Chlamydomonas Cytochrome f with Plastocyanin and Cytochrome c6 // *Biophysical Journal*. — 2003. — Vol. 85. — Pp. 2055–2068.

- Hope A. B. The chloroplast cytochrome bf complex: a critical focus on function // *Biochim. Biophys. Acta.* — 1993. — Vol. 1143. — Pp. 1–22.
- Hope A. B. Electron transfers amongst cytochrome f, plastocyanin and photosystem I: kinetics and mechanisms // *Biochim. Biophys. Acta.* — 2000. — Vol. 1456. — Pp. 5–26.
- Mehta M., Sarafis V., Critchley C. Thylakoid membrane architecture // *Australian Journal of Plant Physiology.* — 1999. — Vol. 26. — Pp. 709–716.
- Pearson D. C., Gross E. L. Brownian dynamics study of the interaction between plastocyanin and cytochrome f // *Biophysical Journal.* — 1998. — Vol. 75. — Pp. 2698–2711.
- Riznichenko G. Y., Kovalenko I. B., Abaturova A. M. et al. New direct dynamic models of protein interactions coupled to photosynthetic electron transport reactions // *Biophysical Reviews.* — 2010. — Vol. 2. — Pp. 101–110.
- Rubin A. B., Riznichenko G. Y. Models of primary processes in a photosynthetic membrane // *Advances in Photosynthesis and respiration.* — 2009. — Vol. 29. — Pp. 151–176.
- Stirbet A., Govindjee, Strasser B. J., Strasser R. J. Chlorophylla Fluorescence Induction in Higher Plants: Modelling and Numerical Simulation // *Journal of Theoretical Biology.* — 1998. — Vol. 193. — Pp. 131–151.
- Ubbink M., Ejdeback M., Karlsson G. B., Bendall D. S. The structure of the complex of plastocyanin and cytochrome f, determined by paramagnetic NMR and restrained rigid-body molecular dynamics // *Structure.* — 1998. — Vol. 6. — Pp. 323–335.
- Ullmann G. M., Knapp E.-W., Kostic N. M. Computational Simulation and Analysis of Dynamic Association between Plastocyanin and Cytochrome f. Consequences for the Electron-Transfer Reaction // *Journal of American Chemical Society.* — 1997. — Vol. 119. — Pp. 42–52.

УДК: 574.52: 57.045

Поиск связей между биологическими и физико-химическими характеристиками экосистемы Рыбинского водохранилища.

Часть 1. Критерии неслучайности связи

Д. В. Рисник^{1,a}, А. П. Левич^{1,b}, Н. Г. Булгаков^{1,c}, Э. С. Бикбулатов²,
Е. М. Бикбулатова², Ю. В. Ершов², И. В. Конюхов¹,
Л. Г. Корнева², В. И. Лазарева², А. С. Литвинов², В. Н. Максимов¹,
С. В. Мамихин³, В. А. Осипов¹, Н. Г. Отюкова², С. А. Поддубный²,
И. Л. Пырина², Е. А. Соколова², И. Э. Степанова², П. В. Фурсова¹,
О. Л. Цельмович²

¹Биологический факультет Московского государственного университета имени М. В. Ломоносова,
Россия, 119234, г. Москва, Ленинские горы, д. 1, стр. 12

²Институт биологии внутренних вод имени И. Д. Папанина РАН,
Россия, 152742, Ярославская обл., Некоузский р-н, п. Борок, ИБВВ РАН

³Факультет почвоведения Московского государственного университета имени М. В. Ломоносова
Россия, 119234, г. Москва, Ленинские горы, д. 1, стр. 12

E-mail: ^abiant3@mail.ru, ^bapl@chronos.msu.ru, ^cbulgakov@chronos.msu.ru

*Получено 20 июня 2012 г.,
после доработки 4 сентября 2012 г.*

На основании данных по содержанию пигментов фитопланктона, интенсивности флуоресценции проб и некоторыми физико-химическим характеристикам вод Рыбинского водохранилища проведен поиск связи между биологическими и физико-химическими характеристиками. Исследованы стандартные методы статистического анализа (корреляционный, регрессионный), методы описания связи между качественными классами характеристик, основанные на отклонении исследуемого распределения характеристик от независимого распределения. Предложен метод поиска оптимальных границ качественных классов по критерию максимума коэффициентов связи.

Ключевые слова: флуоресценция, фитопланктон, пигменты, хлорофилл, коэффициент Юлла, коэффициент Пирсона, поиск связи, Рыбинское водохранилище

Searching for connections between biological and physico-chemical characteristics of Rybinsk reservoir ecosystem. Part 1. Criteria of connection non-randomness

D. V. Risnik¹, A. P. Levich¹, N. G. Bulgakov¹, E. S. Bikbulatov², E. M. Bikbulatova²,
Yu. V. Ershov², I. V. Konyuhov¹, L. G. Korneva², V. I. Lazareva², A. S. Litvinov²,
V. N. Maksimov¹, S. V. Mamihin³, V. A. Osipov¹, N. G. Otyukova², S. A. Poddubnii²,
I. L. Pirina², E. A. Sokolova², I. E. Stepanova², P. V. Fursova¹, O. L. Cel'movich²

¹ Moscow State University, Faculty of Biology, MSU, 1-12, Leninskiye Gory, Moscow, 119991, Russia

² I.D. Papanin Institute for biology of inland waters Russian Academy of Sciences, IBIW RAS, Borok, Nekouzskii District, Yaroslavl'skii region, 152742, Russia

³ Moscow State University, Faculty of Soil Science, MSU, 1-12, Leninskiye Gory, Moscow, 119991, Russia

Abstract. — Based on contents of phytoplankton pigments, fluorescence samples and some physico-chemical characteristics of the Rybinsk reservoir waters, searching for connections between biological and physico-chemical characteristics is working out. The standard methods of statistical analysis (correlation, regression), methods of description of connection between qualitative classes of characteristics, based on deviation of the studied characteristics distribution from independent distribution, are studied. A method of searching for boundaries of quality classes by criterion of maximum connection coefficient is offered.

Keywords: fluorescence, phytoplankton, pigments, chlorophyll, Yule coefficient, Pearson coefficient, searching for connection, Rybinsk reservoir

Citation: *Computer Research and Modeling*, 2013, vol. 5, no. 1, pp. 83–105 (Russian).

Работа частично поддержана грантами РФФИ №№ 13-04-01027а, 12-07-00580а.

© 2013 Дмитрий Владимирович Рисник, Александр Петрович Левич, Николай Гурьевич Булгаков, Эрнст Саяфнурович Бикбулатов, Екатерина Максимовна Бикбулатова, Юрий Викторович Ершов, Иван Владимирович Конохов, Людмила Генриховна Корнева, Валентина Ивановна Лазарева, Александр Сергеевич Литвинов, Виктор Николаевич Максимов, Сергей Витальевич Мамихин, Владимир Алексеевич Осипов, Наталья Георгиевна Отыукова, Сергей Артурович Поддубный, Инна Логиновна Пырина, Евгения Александровна Соколова, Ирина Эрнстовна Степанова, Полина Викторовна Фурсова, Ольга Леонидовна Цельмович

Предлагаем читателю серию статей, посвященных поиску подходов к описанию взаимосвязей между полученными в натурных измерениях экологическими характеристиками. В первой статье подробно сформулированы проблемы, возникающие при анализе натурных данных, в деталях описаны исходные материалы и предложены критерии, указывающие на степень неслучайности связей между характеристиками. Во второй статье испытаны меры связи, предлагаемые детерминационным анализом, исследующим прогностические связи вида «из того, что характеристика X имеет значения в диапазоне A , следует, что характеристика Y имеет значения в диапазоне B ». В третьей статье серии рассмотрен частный, но важный в прикладном отношении случай поиска взаимосвязей между биологическими и физико-химическими характеристиками — поиск границ классов качества вод на основе установления экологических норм для биологических индикаторов состояния вод и для факторов среды.

1. Постановка проблемы

В контролируемых условиях лабораторных экспериментов «хорошо организованные» зависимости между биологическими и физико-химическими характеристиками имеют вид однозначных функций, поддающихся корреляционному, регрессионному и другим видам статистического анализа (рис. 1).

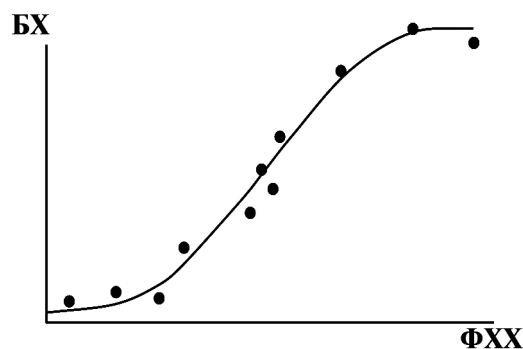


Рис. 1. Зависимости между биологическими (БХ) и физико-химическими (ФХХ) характеристиками в лабораторных опытах

В природных экосистемах на биологические характеристики одновременно действует множество факторов среды, среди которых только часть представлена в экспедиционных измерениях. Биологическая характеристика становится функцией многих физико-химических переменных. Зависимость между переменными в этом случае имеет вид «плохо организованного», «размытого» облака точек (рис. 2).

Корреляционный анализ «размытых» зависимостей обычно демонстрирует сравнительно низкую и незначимую величину связей. Многомерный регрессионный анализ также осложнен рядом обстоятельств [Налимов, Чернова, 1965]. Если даже в некоторых редких случаях регрессионный анализ приводит к получению высокого множественного коэффициента корреляции, то остается нерешенной задача о поиске тесноты связи между биологической переменной и каждым из факторов, поскольку частные коэффициенты корреляции остаются незначительными. Следует добавить [Рисник, Рыбка, 2011], что применение множественного регрессионного анализа очень часто осложнено недостаточным количеством данных, причем недостаточность определяется фактором с самым малочисленным количеством наблюдений, сколь бы много их ни было для других переменных.

Применение многих статистических процедур предполагает, что исходные данные обладают какими-либо априорными статистическими свойствами, например, распределены по гауссовскому, пуассоновскому или другому закону. Однако для реальных экологических данных это предположение выполняется очень редко.

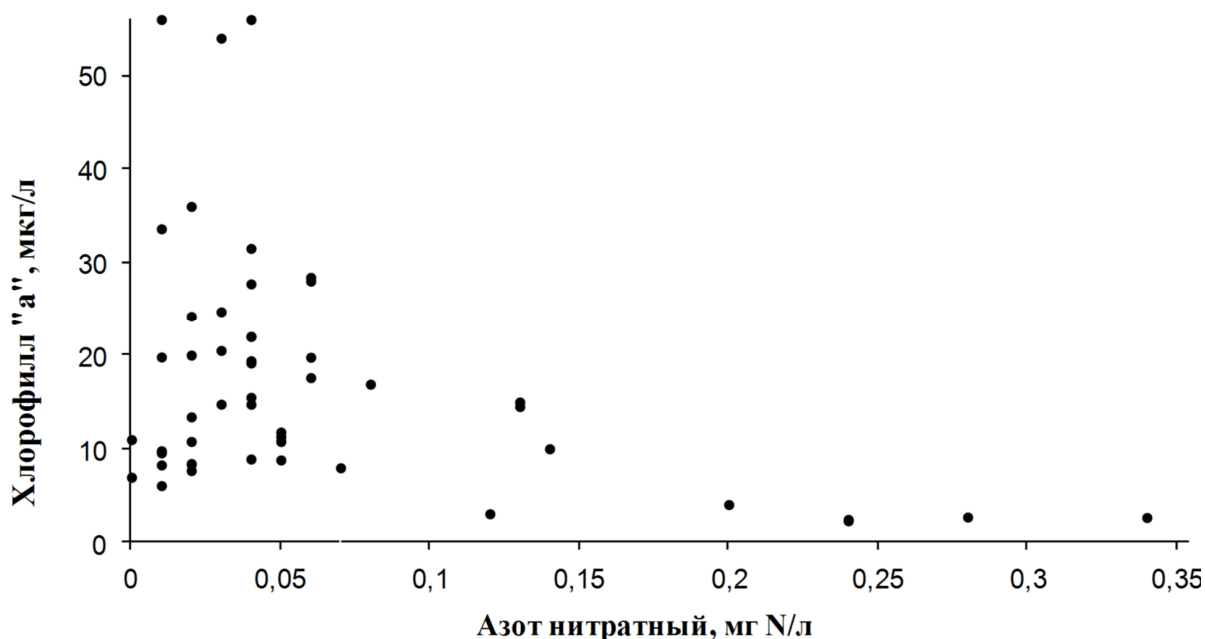


Рис. 2. Пример типичной зависимости между значениями биологической и физико-химической характеристик для Рыбинского водохранилища (июнь–август 2010 года)

Для анализа натуральных данных необходим метод поиска взаимосвязи между переменными, позволяющий выявлять корреляции, скрытые при рассмотрении парных зависимостей между характеристиками. Один из методов анализа «плохо организованных» данных — переход от количественных переменных к их качественным классам. Такими классами могут быть низкие, средние и высокие значения; благополучные и неблагополучные; допустимые и недопустимые значения и т. п. После выделения качественных классов возможен поиск корреляций и других видов связи уже между качественными классами различных переменных.

Применение анализа качественных переменных, в свою очередь, сталкивается по крайней мере с двумя трудностями. Во-первых, возникает проблема выбора объективного критерия для выделения качественных классов: какие значения считать высокими и какие низкими, какие допустимыми и какие недопустимыми. Обычно границы между качественными классами вводят в большой степени субъективно: например, диапазон измерения характеристики делят на равные интервалы (в линейной или логарифмической шкале) или экспертным образом назначают высокие и низкие или другие значения. Субъективность выбора границ ставит под сомнение обоснование всех последующих процедур установления связей. Вторая трудность вызвана упомянутым выше неустраняемым *in situ* влиянием на биологические характеристики всех факторов среды и состоит в том, что некоторое значение биологической характеристики может быть вызвано не исследуемым фактором, а каким-либо другим, действующим одновременно с ним.

Данная работа посвящена обсуждению следующих исследовательских проблем.

Возможно ли не субъективное, а основанное на четких критериях выделение качественных классов экологических характеристик? Если классы значений выделены, то какими методами следует оценивать связи между качественными переменными?

Предложено решение следующих исследовательских задач.

Найти связи между количественными значениями биологических и физико-химических характеристик экосистемы Рыбинского водохранилища с помощью корреляционного и регрессионного статистического анализа. В случаях, когда из-за «размытости» зависимостей эти виды анализа не обнаруживают связи, предложить методы поиска связей между качественными переменными и подходы к строгому обоснованию выделения качественных классов.

В качестве объективного обоснования для выделения качественных классов количественных переменных предложен экстремальный принцип: оптимальное разделение на классы соответствует экстремуму (минимуму или максимуму) некоторого количественного критерия. В качестве критериев в предлагаемой серии работ рассмотрены:

- Минимум суммарного отклонения «точек» класса от среднего для класса значения (метод k -средних в кластерном анализе, где k — количество качественных классов).
- Максимум коэффициента ассоциации Юла или коэффициента контингенции Пирсона, характеризующих неслучайность связи и аналогичных коэффициенту корреляции между количественными переменными в случае качественных переменных (для дихотомических признаков коэффициент Пирсона совпадает с соответствующим коэффициентом корреляции).
- Максимум прогностических коэффициентов Гуттмана, Валлиса и Чеснокова, описывающих степень достоверности утверждения «если характеристика X имеет значение в классе A , то характеристика Y имеет значение в классе B ».

Полученные в работе методы и подходы позволяют ответить на целый ряд вопросов экологических исследований:

- Какие физико-химические характеристики Рыбинского водохранилища влияют на биологические и в какой степени?
- Где проходят границы между высокими и низкими или благополучными и неблагополучными значениями для биологических характеристик Рыбинского водохранилища?
- Как объективно обосновать границы классов качества вод, устанавливаемых по биологическим показателям?
- Где проходят границы между допустимыми и недопустимыми (с точки зрения качества среды) значениями физико-химических характеристик, т. е. как объективно обосновать границы классов качества вод, устанавливаемых по физико-химическим показателям?

2. Экспериментальная часть

В качестве биологических характеристик использовали содержание пигментов и показатели флуоресценции фитопланктона, в качестве физико-химических — содержание минеральных и органических веществ, показатели кислотности, щелочности, утилизации кислорода, электропроводность, прозрачность, цветность, температуру воды и воздуха. Первичные данные, используемые для анализа, были получены в результате измерения этих характеристик в поверхностных и батометрических пробах в различных районах Рыбинского водохранилища в июне–августе 2010 года в ходе пяти рейсов научно-исследовательского судна «Академик Топчиев» Института биологии внутренних вод им. И. Д. Папанина РАН (ИБВВ РАН). Первичные данные опубликованы коллективом соавторов [Данные совместных измерений..., 2011]. Станции и даты отбора проб приведены на рисунке 3 и в таблице 1.

Сотрудниками ИБВВ РАН получены следующие показатели:

1. физические характеристики (температура воды и воздуха, электропроводность, прозрачность, цветность);
2. значения кислотности (рН) и щелочности;
3. содержание минеральных веществ (Ca^{2+} , Mg^{2+} , Na^+ , K^+ , Cl^- , SO_4^{2-} , HCO_3^- , сумма ионов);
4. содержание биогенных элементов (NO_3^- , NO_2^- , NH_4^+ , общий азот, PO_4^{3-} , общий фосфор);
5. содержание различных фракций органических веществ (ХПК, БПК₅, органический углерод во взвеси, нефтепродукты);
6. содержание хлорофиллов «а», «b», «с», феопигментов и чистого хлорофилла «а» (с поправкой на феопигменты), соотношение каротиноидов и хлорофилла, установленное по отношению поглощения экстракта пигментов при длине волны 480 и 664 нм (E_{480}/E_{664}).

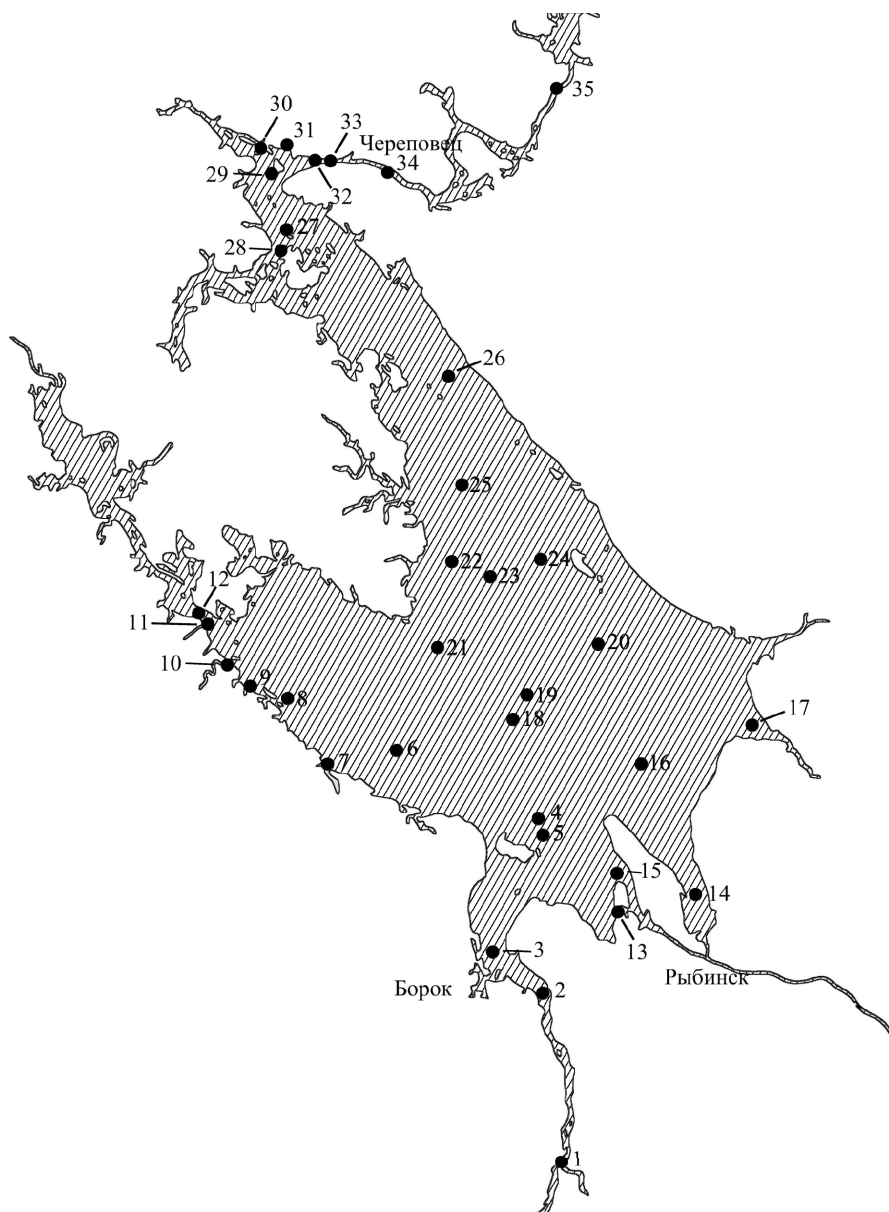


Рис. 3. Схема расположения станций отбора проб. Расшифровка номеров станций приведена в таблице 1

Сотрудниками Биологического факультета МГУ имени М. В. Ломоносова получены данные по интенсивностям флуоресценции проб.

Определение солевого состава вод выполнено аккредитованным Аналитическим центром ИБВВ (аттестат аккредитации: РОСС RU.0001.512040) в соответствии с методикой выполнения измерений массовых концентраций катионов калия, натрия, лития, магния, кальция, аммония, стронция, бария [Методика выполнения измерений..., 2007] и с методикой выполнения измерений массовых концентраций хлорид-, сульфат- и фторид-ионов [Методика выполнения измерений..., 2009]. Определение содержания биогенных элементов, органических веществ, нефтепродуктов, кислотности и щелочности проводили в той же аккредитованной лаборатории с применением соответствующих аттестованных методов.

Таблица 1. Станции и даты отбора проб. На рисунке 3 станции обозначены номерами из первого столбца данной таблицы

№	Станция	Рейс 1	Рейс 2	Рейс 3	Рейс 4	Рейс 5
1	русло Волги у г. Мышкин					27.08
2	русло Волги у с. Глебово		28.06			
3	русло Волги у бывшего с. Коприно	15.06	28.06	11.07	10.08	24.08
4	русло Волги у затопленного г. Молога	15.06	28.06	11.07	10.08	24.08
5	русло Мологи у буя № 81			11.07		
6	русло Мологи против п. Брейтово	15.06	28.06	10.07	10.08	24.08
7	р. Сить у п. Брейтово		28.06			25.08
8	русло Мологи у о. Первомайка			10.07		25.08
9	устье ручья Тереха против о. Первомайка		29.06			25.08
10	русло Мологи против устья р. Себла		29.06			25.08
11	русло Мологи против устья р. Лама		29.06			25.08
12	русло Мологи у с. Противье			10.07		25.08
13	Югский залив		30.06			
14	русло Шексны у с. Волково			09.07		27.08
15	русло Волги у п. Каменники		30.06			27.08
16	русло Шексны у затопленного с. Всехсвятское		30.06	10.07		26.08
17	устье р. Ухра		30.06	09.07		
18	Молого-Шекснинское междуречье у затопленного с. Наволок	15.06	28.06	10.07	10.08	24.08
19	Наволок-2			10.07		
20	против с. Измайлово	15.06	28.06	09.07	10.08	24.08
21	против Центрального мыса		29.06			25.08
22	правая пойма Шексны близ убежища Средний Двор			09.07		
23	русло Шексны против убежища Средний Двор	15.06	28.06	09.07	10.08	24.08
24	левая пойма Шексны близ убежища Средний Двор			09.07		
25	русло Шексны у затопленного с. Ягорба		29.06	07.07, 09.07		25.08
26	русло Шексны против с. Мякса		29.06	08.07, 09.07		26.08
27	русло Шексны у затопленного с. Любец		29.06	08.07		26.08
28	мелководье к югу от с. Любец		29.06			26.08
29	русло Шексны у о. Ваганиха			08.07		
30	устье р. Суда у турбазы Торово			07.07		26.08
31	устье р. Кошта					26.08
32	сброс промышленных стоков г. Череповец			07.07		
33	ниже устья р. Ягорба			07.07		
34	русло Шексны у п. Кабачино			08.07		
35	р. Шексна в 30 км выше г. Череповец			06.07		

Хлорофиллы «а», «b», «с» определяли стандартным спектрофотометрическим методом [Lorenzen, Jeffrey, 1980; ГОСТ, 1990], феопигменты и «чистый» хлорофилл «а» — по Лоренцену [Lorenzen, 1967] с модификацией процедуры отбора проб и проведения анализов применительно к исследованиям ИБВВ в Рыбинском водохранилище [Пырина, 2000]. Величину E_{480}/E_{664} находили как отношение оптических плотностей экстракта в области наибольшего вклада поглощения каротиноидов и хлорофилла [Burkholder et al., 1959; Пырина, Сигарева, 1986]. Общую ошибку определения исследованных характеристик рассчитывали методом оценки косвенных погрешностей [Зайдель, 1974; Пырина и др., 1993]. Измерения быстрой флуоресценции фитопланктона проводили согласно «Методике измерений обилия и индикации изменения со-

стояния фитопланктона в природных водах флуоресцентным методом» (аттестат №01.00225/205-66-11, ФР.1.39.2011.11246) [Маторин, Осипов, Рубин, 2012]. В отобранных пробах бортовым флуориметром «МЕГА-11» в 3–6-кратных повторностях определяли фоновую интенсивность флуоресценции F_0 общее (все реакционные центры фотосистемы 2, возбуждаемой светом в коротковолновой части спектра, находятся в «открытом» состоянии) и максимальную интенсивность переменной флуоресценции F_m общее (все реакционные центры фотосистемы 2 находятся в «закрытом» состоянии) (рис. 4).

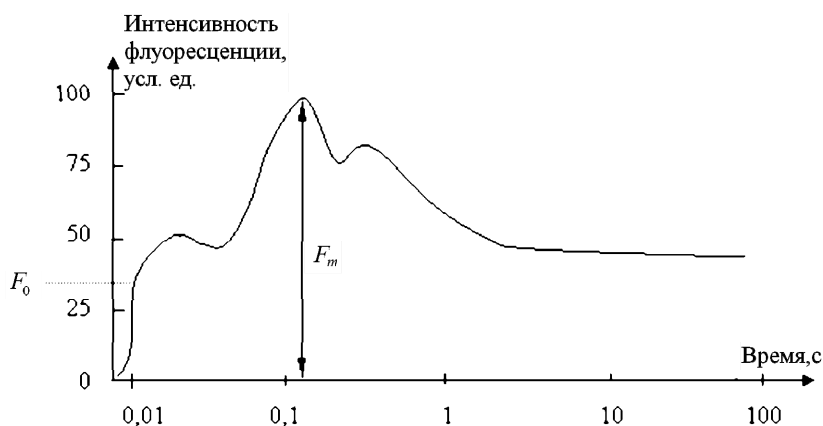


Рис. 4. Типичная кривая индукции флуоресценции фитопланктона

В связи с высоким содержанием в водах Рыбинского водохранилища растворенных органических веществ, вносящих значительный вклад в общую интенсивность флуоресценции, для корректного определения интенсивности флуоресценции фитопланктона необходимо определение интенсивности флуоресценции растворенной органики. Для учета указанного вклада исходя из предположения о том, что этот вклад достаточно постоянен для всей акватории Рыбинского водохранилища, на некоторых станциях во время второго и третьего рейсов, помимо анализа первичных проб, был произведен анализ их фильтратов. Во втором и третьем рейсах фильтрат был получен в результате смешения фильтратов проб с горизонтов 0–2 м, 2–6 м, 6 м–дно как побочный продукт в процессе определения хлорофилла стандартным методом. Однако оказалось, что вклад органических веществ существенно изменялся в зависимости от места отбора проб, поэтому в четвертом и пятом рейсах интенсивность флуоресценции фильтрата определяли для проб со всех станций и горизонтов 0–2 м, 2–6 м и 6 м–дно. После измерения интенсивности флуоресценции пробы фильтровали через трех-четырёхслойный фильтр (ватмановский фильтр JF/F с диаметром пор 0,7 мкм), F_0 фильтрата и F_m фильтрата определяли в трехкратной повторности. Величина F_0 фильтрата должна быть равна значению F_m фильтрата, поскольку разность между F_0 и F_m связана исключительно с различиями в состоянии реакционных центров в фотосистеме 2.

В анализе исследовали следующие показатели интенсивности флуоресценции фитопланктона: F_0 общее, F_m общее, F_0 фильтрата, F_0 фитопланктона = F_0 общее – F_0 фильтрата, F_m фитопланктона = F_m общее – F_m фильтрата.

Таблица 2 дает представление о количестве проанализированных проб. Во всех таблицах значения характеристик имеют следующие размерности: температура — градусы Цельсия; электропроводность — мкСм/см; прозрачность — см по диску Секки, рН — безразмерные единицы; щелочность — мг/л; $C_{орг.}$ — мг С/л во взвеси; цветность — градусы по Сг – Со шкале; БПК₅, ХПК — мг O₂/л; нефтепродукты — мг/л; NH₄⁺, NO₂⁻, NO₃⁻, N_{общ.} — мг N/л; PO₄³⁻, P_{общ.} — мг P/л; Ca²⁺, Mg²⁺, Na⁺, K⁺, Cl⁻, SO₄²⁻, HCO₃⁻, сумма ионов — мг/л; пигменты — мкг/л; E₄₈₀/E₆₆₄ — относительные единицы; интенсивности флуоресценции — относительные безразмерные единицы.

Таблица 2. Число станций (с) и измерений (и) экологических характеристик

Характеристики	Номер рейса										Всего	
	1		2		3		4		5			
	с	и	с	и	с	и	с	и	с	и	с	и
Ca ²⁺	6	6	19	19	0	0	6	6	22	22	53	53
Mg ²⁺	6	6	19	19	0	0	6	6	22	22	53	53
Na ⁺	6	6	19	19	0	0	6	6	22	22	53	53
K ⁺	6	6	19	19	0	0	6	6	22	22	53	53
Cl ⁻	6	6	19	19	0	0	6	6	22	22	53	53
SO ₄ ²⁻	6	6	19	19	0	0	6	6	22	22	53	53
HCO ₃ ⁻	6	6	19	19	0	0	6	6	22	22	53	53
Сумма ионов	6	6	19	19	0	0	6	6	22	22	53	53
NH ₄ ⁺	6	6	6	6	0	0	6	6	7	7	25	25
NO ₂ ⁻	6	6	6	6	20	20	6	6	19	19	57	57
NO ₃ ⁻	6	6	6	6	20	20	6	6	19	19	57	57
N _{общ.}	6	6	6	6	20	20	6	6	19	19	57	57
PO ₄ ³⁻	6	6	6	6	20	20	6	6	19	19	57	57
P _{общ.}	6	6	6	6	20	20	6	6	19	19	57	57
Цветность	6	6	11	11	24	24	6	6	20	20	67	67
БПК ₅	6	6	6	6	20	20	6	6	19	19	57	57
ХПК	6	6	6	6	20	20	6	6	19	19	57	57
Температура воды	6	23	12	31	25	97	6	16	20	20	69	187
Электропроводность	6	23	7	26	25	97	6	16	20	20	64	182
Прозрачность	6	6	12	12	25	25	6	6	20	20	69	69
Температура воздуха	0	0	0	0	23	23	0	0	20	20	43	43
pH	6	6	6	6	20	20	6	6	19	19	57	57
Щелочность	6	6	6	6	20	20	6	6	19	19	57	57
Нефтепродукты	5	5	5	5	5	5	5	5	18	18	38	38
C _{орг.} во взвеси	6	6	6	6	18	18	6	6	19	19	55	55
Пигменты фитопланктона	6	23	11	43	23	71	6	16	12	12	58	165
Показатели интенсивности флуоресценции	6	23	20	64	26	99	6	16	23	23	81	225

3. Методы анализа связей между экологическими характеристиками

В работе использованы как стандартные методы статистического анализа количественных характеристик, так и методы выявления связей между качественными классами по таблицам сопряженности. Расчеты проводили с помощью специально созданных программ, написанных на языке Visual Basic for Application для MS Excel.

3.1. Корреляционный и регрессионный анализ

Существует несколько шкал для качественной интерпретации результатов корреляционного анализа [Chaddock, 1925; Голубков, 1998 и другие]. Из их анализа следует, что одни и те же значения коэффициентов корреляции различные специалисты интерпретируют по-разному. Объективный выбор между шкалами, по-видимому, невозможен, поэтому при интерпретации коэффициентов корреляции необходимо и достаточно сослаться на выбранную шкалу.

Множественный регрессионный анализ позволяет определить вес вклада каждого фактора (коэффициенты в уравнении регрессии) и исследовать зависимость биологической характеристики от рассчитанного по уравнению регрессии показателя, учитывающего все включенные в анализ факторы. В данных по Рыбинскому водохранилищу для измеренных биологических характеристик часто отсутствуют те или иные данные по физико-химическим характеристикам, что осложняет регрессионный анализ.

3.2. Коэффициенты связи между качественными классами

Анализ связей в таблицах сопряженности рассмотрен на примере выделения двух качественных классов для каждой из характеристик (табл. 3). Обозначения из этой таблицы использованы в дальнейших формулах.

Таблица 3. Таблица сопряженности для двух характеристик X и Y с двумя классами значений каждой («высокие значения «и» низкие значения»). Символы n_a , n_b , n_c и n_d обозначают количества наблюдений в соответствующих ячейках таблицы

	Низкие значения характеристики X	Высокие значения характеристики X
Высокие значения характеристики Y	n_a	n_b
Низкие значения характеристики Y	n_c	n_d

3.2.1. Степень неслучайности связи

В качестве критерия неслучайности связи между характеристиками обычно применяют коэффициент ассоциации Юла $Q = \frac{n_a n_d - n_b n_c}{n_a n_d + n_b n_c}$ и коэффициент контингенции Пирсона

$P = \frac{n_a n_d - n_b n_c}{\sqrt{(n_a + n_b)(n_b + n_d)(n_d + n_c)(n_c + n_a)}}$ [Миркин, 1980]. Коэффициент Юла характеризует

одностороннюю связь, т. е. для максимальности связи, когда $Q = 1$ или $Q = -1$, достаточно, чтобы была пуста хотя бы одна ячейка таблицы сопряженности. Коэффициент Пирсона характеризует двустороннюю связь. Для максимальности связи, когда $P = 1$ или $P = -1$, необходимо, чтобы одновременно были пусты ячейки « b » и « c » или « a » и « d ». Если $|P| \geq 0.5$, считают, что надежно установлена двусторонняя связь [Паниотто, Максименко, 1982].

Коэффициент контингенции Пирсона равен коэффициенту корреляции соответствующих дихотомических признаков, рассматриваемых как количественные [Миркин, 1980].

В случае более чем двух качественных классов неслучайность связи измеряют коэффициентами, основанными на величине «хи квадрат» (коэффициенте квадратической сопряженности), например, коэффициентом $\Phi = \sqrt{\chi^2 / N}$, где χ^2 — коэффициент квадратической сопряженности, описанный подробнее в разделе 3.2.2.

3.2.2. Значимость связей

Обнаружение приемлемой степени связи между характеристиками по величине коэффициента связи еще не говорит о том, что найденная связь не случайна на исследуемой выборке, тогда как в генеральной совокупности связи между характеристиками нет.

Проверка любой статистической гипотезы сводится к следующему. По выборочным значениям случайной величины подсчитывается некоторая величина — статистика критерия. При допущении, что распределение вероятностей используемой статистики критерия в условиях

справедливости проверяемой гипотезы известно, определяется вероятность появления вычисленного значения статистики. На основе принципа значимости устанавливается уровень значимости — наибольшее значение вероятности, несовместимое с признанием случайности экспериментально вычисленного значения статистики критерия. Событие называется значимым (а не случайным), если теоретическая вероятность его случайного появления меньше, чем принятый уровень значимости. Уровнем значимости определяется критическое значение статистики критерия. Как правило, если значение статистики критерия, вычисленное по экспериментальным данным, больше критического, то гипотеза отклоняется на выбранном уровне значимости. В противном случае она признается не противоречащей результатам наблюдений и не отвергается [Кобзарь, 2006, с. 20 и 202].

Сформулируем нулевую гипотезу для случая независимости характеристик X и Y . В этом контексте независимость означает, что доля общего числа индивидуумов в строке, принадлежащая произвольному, но фиксированному столбцу, одна и та же для всех строк и что доля общего числа индивидуумов в столбце, принадлежащая произвольной, но фиксированной строке, одна и та же для всех столбцов [Афифи, Эйзен, 1982]. При одинаковых долях наблюдений в ячейках по отношению к общему числу наблюдений, и, следовательно, одинаковых коэффициентах связи, значимость увеличивается с ростом общего числа наблюдений. То есть при заметной и высокой связи и низком числе наблюдений, связь может быть незначима, что свидетельствует о необходимости добавления в анализ дополнительных наблюдений. А при слабой и отсутствующей связи, но большом числе наблюдений, связь может быть значимой, что часто неверно трактуют как наличие связи, хотя связь может быть настолько слаба, что фактически ее следует трактовать как отсутствие связи.

Для оценки отклонения исследуемой таблицы сопряженности от сформулированной выше нулевой гипотезы принято использовать [Миркин, 1980, Афифи, Эйзен, 1982, Айвазян и др.,

1985] критерий «хи квадрат»: $\chi^2 = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}}$, где f_{ij} — число наблюдений в ячейке с коор-

динатами i -тая строка, j -тый столбец; e_{ij} — расчетная вероятность числа наблюдений в ячейке с координатами i -тая строка, j -тый столбец, равная $f_i f_j / N$, где f_i — число наблюдений в i -той строке, f_j — число точек в j -том столбце таблицы сопряженности, N — общее число наблюдений [Айвазян и др., 1985].

Для определения значимости связи проводили сравнение расчетной величины χ^2 с табличным значением при уровне значимости $\alpha = 0.05$ и числе степеней свободы $n = (i - 1)(j - 1)$. Если расчетная величина χ^2 больше табличной, то гипотеза об отсутствии связи отвергается. То есть это свидетельствует о значимости связи с уровнем значимости $\alpha = 0.05$. Если расчетная величина χ^2 меньше табличной, то гипотеза об отсутствии связи не отвергается. Это означает, что для установления значимой связи с уровнем значимости $\alpha = 0.05$ необходимо большее число наблюдений. В нашем случае исследования двух классов качества по обоим характеристикам и уровне значимости $\alpha = 0.05$ критическое значение критерия χ^2 равно 3.84.

Применение критерия χ^2 возможно при условии, что ячеек с расчетной $e_{ij} < 5$ не более 20 % от всех ячеек [Дубина, 2006]. При этом общее число наблюдений в таблице должно быть больше 40. Для расчета критерия χ^2 при числе наблюдений менее 40, но более 20 применяли поправку Йетса на непрерывность [Афифи, Эйзен, 1982]. Критерий χ^2 с поправкой Йетса применим только для таблиц сопряженности с двумя классами качества по обоим характеристикам.

3.3. Методы поиска экстремальных качественных связей

Метод k -средних кластерного анализа. Метод k -средних [Айвазян и др., 1989] позволяет задать число качественных классов, необходимых исследователю. В нашем случае были выделены два класса: высоких и низких значений биологических и физико-химических характеристик. Алгоритм метода k -средних стремится минимизировать суммарное квадратичное откло-

нение точек кластера от средней арифметической величины μ_i его значений: $\sum_{i=1}^k \sum_{x_j \in S_i} (x_j - \mu_i)^2$,

где k — число кластеров, S_i — полученные кластеры. Поиск границ проводили отдельно для биологических и физико-химических характеристик, поскольку совместный анализ распределения не позволяет получить четкие границы классов.

Метод максимизации коэффициентов связи (метод МКС). Алгоритм метода состоит в переборе всевозможных положений границ как для биологической, так и для физико-химической характеристики и в выборе таких двух границ, для которых коэффициент связи максимален. Алгоритм включает несколько дополнительных условий:

- 1) Количество наблюдений в ячейках таблицы сопряженности должно быть достаточно представительным, чтобы результат поиска был достоверным. Представительность можно описать критерием $ПР = n_i / N$, где n_i — количество наблюдений в ячейке i , N — общее количество наблюдений. Каждая из представительностей должна быть больше заданного параметра поиска $ПР_{\min}$ (обычно $ПР_{\min}$ варьирует в диапазоне 0.15–0.25). При поиске прямой связи ограничение по минимальной представительности накладывают на ячейки «с» и «b», при поиске обратной связи — на ячейки «a» и «d». В данной работе была принята минимальная представительность, равная 0.2.
- 2) Достоверность результатов поиска может быть обеспечена, если общее число наблюдений N не слишком мало: $N > N_{\min}$, где N_{\min} — еще один параметр поиска (обычно его выбирают в пределах от 30 до 80). Кроме того, для того чтобы судить о значимости связи по критерию χ^2 , необходимо учитывать минимальное количество наблюдений для применимости данного критерия. Для двух классов качества как по биологической, так и по физико-химической характеристике минимальное число наблюдений составляет 20 (минимальное число наблюдений для применимости поправки Йетса на непрерывность), для трех и более классов минимальное число наблюдений составляет 40 [Афифи, Эйзен, 1982].
- 3) Поскольку значимость установленной связи характеризуют превышением величиной χ^2 ее табличного значения, уровень значимости по критерию χ^2 также составляет условие поиска, задаваемое исследователем (как правило, используют уровень значимости $\alpha = 0.05$ –0.10).

4. Результаты поиска связей

Рассчитаны коэффициенты корреляции между биологическими и физико-химическими характеристиками (табл. 4).

Проведено разбиение диапазонов измерения экологических характеристик на качественные классы низких и высоких значений по методу k -средних кластерного анализа. Для полученных классов рассчитаны коэффициенты Юла (табл. 5) и Пирсона (табл. 6).

Минимальная представительность для числа наблюдений в строках и столбцах таблицы сопряженности была равна 0.25.

Стоит отметить, что использование коэффициентов ассоциации в чистом виде плохо поддается трактовке, поэтому в таблице 5 представлен не только коэффициент ассоциации, но и ячейка, в связи с пустотой которой этот коэффициент оказался высок. Приведем пример трактовки полученного результата: все высокие (более 20.5 мкг/л) значения концентрации хлорофилла «а» сопряжены с низкой (менее 160 см) прозрачностью, а высокие значения прозрачности (более 160 см) сопряжены с низкими (менее 20.5 мкг/л) значениями концентрации хлорофилла.

Таблица 4. Коэффициенты корреляции между биологическими и физико-химическими характеристиками. Обозначения: значимые связи выделены жирным шрифтом. А — общий хлорофилл «а» (вместе с феопигментами); В — хлорофилл «b»; С — хлорофилл «с»; F — феопигменты; АЧ — «чистый» хлорофилл «а» (с поправкой на феопигменты); E_{480}/E_{664} — показатель соотношения каротиноидов и хлорофилла; F_m общее — максимальная интенсивность переменной флуоресценции в пробе без поправки на интенсивность флуоресценции фильтрата (растворенных органических веществ); F_0 общее — фоновая интенсивность флуоресценции без поправки на интенсивность флуоресценции фильтрата; F_0 фильтрата — интенсивность флуоресценции фильтрата; F_m фито — максимальная интенсивность переменной флуоресценции фитопланктона в пробе (F_m фито = F_m общее - F_0 фильтрата); F_0 фито — фоновая интенсивность флуоресценции фитопланктона в пробе (F_0 фито = F_0 общее - F_0 фильтрата). Размерности всех характеристик приведены в разделе 2

Физико-химические характеристики	Биологические характеристики										
	Показатели интенсивности флуоресценции					Хлорофилл				F	$\frac{E_{480}}{E_{664}}$
	F_m общ	F_m фито	F_0 общ	F_0 фито	F_0 фильтрата	A	B	C	АЧ		
Ca ²⁺	0.39	0.41	0.33	0.35	-0.05	0.39	-0.25	0.01	0.42	0.35	-0.44
Mg ²⁺	0.28	0.28	0.23	0.22	-0.03	0.43	-0.44	0.11	0.31	0.42	-0.37
Na ⁺	0.10	0.15	0.00	0.07	-0.26	0.24	-0.31	-0.03	0.20	0.23	-0.37
K ⁺	-0.05	0.00	-0.13	-0.01	-0.25	-0.04	-0.01	0.22	-0.04	-0.03	-0.05
Cl ⁻	0.00	0.07	-0.13	-0.01	-0.41	0.01	0.20	-0.22	0.15	-0.02	-0.29
SO ₄ ²⁻	-0.10	-0.10	-0.08	-0.12	0.00	0.13	-0.04	-0.06	0.05	0.13	0.11
HCO ₃ ⁻	0.36	0.37	0.30	0.32	-0.05	0.34	-0.29	0.06	0.34	0.32	-0.40
Сумма ионов	0.35	0.36	0.29	0.30	-0.07	0.40	-0.32	0.04	0.39	0.38	-0.43
NH ₄ ⁺	0.15	0.19	0.02	0.13	-0.43	0.33	0.17	-0.16	0.28	0.32	-0.25
NO ₂ ⁻	0.04	0.11	-0.08	0.09	-0.28	-0.25	0.02	-0.10	-0.06	-0.27	0.07
NO ₃ ⁻	-0.17	-0.13	-0.19	-0.14	-0.19	-0.39	0.07	-0.17	-0.19	-0.40	0.25
N _{общ}	0.15	0.15	0.16	0.14	0.11	0.44	0.13	-0.06	0.28	0.43	-0.30
PO ₄ ³⁻	0.02	0.09	-0.08	0.05	-0.25	0.10	0.08	-0.03	0.19	0.08	-0.28
P _{общ}	0.27	0.32	0.12	0.24	-0.14	0.61	0.01	-0.19	0.47	0.58	-0.59
БПК ₅	-0.13	-0.18	-0.06	-0.18	0.11	-0.13	-0.16	0.48	-0.33	-0.07	0.36
ХПК	0.29	0.21	0.38	0.23	0.40	0.22	0.19	-0.16	0.17	0.22	-0.12
C _{орг} во взвеси	0.00	-0.05	0.06	-0.04	0.16	0.35	-0.12	-0.02	0.12	0.36	-0.07
Нефтепродукты	0.00	0.04	0.00	0.02	-0.01	0.18	-0.06	-0.06	0.09	0.19	-0.04
pH	-0.04	-0.07	0.06	-0.04	0.18	-0.18	0.02	-0.03	-0.13	-0.18	0.12
Щелочность	0.01	-0.01	0.02	0.00	-0.02	-0.17	0.31	-0.20	0.07	-0.21	0.05
Цветность	0.18	0.08	0.28	0.07	0.34	0.04	0.25	0.00	0.13	0.02	0.06
Прозрачность	-0.43	-0.44	-0.37	-0.44	-0.16	-0.57	-0.18	0.08	-0.47	-0.55	0.37
Электропроводность	0.29	0.29	0.25	0.29	0.10	0.27	0.12	-0.10	0.22	0.25	-0.26
Температура воды	0.15	0.08	0.21	0.08	0.29	0.24	-0.13	0.22	0.11	0.26	-0.21
Температура воздуха	-0.03	-0.14	0.14	-0.13	0.35	-0.47	0.23	-0.27	-0.20	-0.49	0.24

Таблица 5. Коэффициенты ассоциации Юла для таблиц сопряженности с границами между качественными классами, установленными при помощи метода k -средних. Обозначения те же, что в таблице 4

Физико-химические характеристики	Биологические характеристики										
	Показатели интенсивности флуоресценции					Хлорофилл				F	$\frac{E_{480}}{E_{664}}$
	F_m общ	F_m фито	F_0 общ	F_0 фито	F_0 фильтрата	A	B	C	AЧ		
Ca ²⁺	0.68 a	0.54 a	0.28 c	0.71 a	0.41 a	0.17 a	0.13 a	1.00 a	-0.02 a	1.00 a	-0.27 a
Mg ²⁺	0.32 a	0.21 a	-0.03 c	0.37 a	0.37 a	0.43 a	-1.00 b	1.00 a	0.25 a	0.12 a	-0.48 b
Na ⁺	-0.02 b	-0.12 b	-0.42 c	-0.09 b	-0.24 b	0.27 a	-0.39 b	-1.00 b	0.25 a	-0.18 b	-1.00 b
K ⁺	0.00 b	-0.19 b	-0.50 b	-0.05 b	-0.48 b	-0.06 b	-0.04 b	0.18 b	-0.24 b	0.18 b	-0.21 b
Cl ⁻	0.23 d	0.22 d	-0.15 d	0.18 d	-0.75 b	0.09 a	0.59 a	-0.30 b	0.08 a	0.42 a	-0.62 b
SO ₄ ²⁻	-0.57 b	-0.49 b	-0.18 d	-0.48 b	-0.48 b	0.18 a	-0.45 b	-0.24 b	0.17 a	0.47 a	-0.10 b
HCO ₃ ⁻	0.66 a	0.56 a	0.24 c	0.69 a	0.43 a	0.16 a	-0.27 a	0.12 a	-0.05 a	0.12 a	-0.48 b
Сумма ионов	0.58 a	0.45 a	0.21 c	0.61 a	0.47 a	0.16 a	-0.27 a	1.00 a	-0.05 a	0.12 a	-0.48 b
NH ₄ ⁺	0.33 d	0.68 d	0.56 d	0.56 d	-1.00 b	0.85 d	-1.00 b	-1.00 b	0.85 d	0.60 d	-0.60 b
NO ₂ ⁻	0.27 a	0.23 b	0.22 a	0.38 d	-0.59 b	0.17 b	-0.27 b	-0.15 b	0.17 b	0.50 d	0.00 b
NO ₃ ⁻	-0.26 b	-0.18 b	-1.00 b	-0.22 b	-0.55 b	-0.50 b	0.27 b	-1.00 b	-0.50 b	0.26 d	0.00 b
N _{общ}	0.60 a	0.32 b	0.66 a	0.26 d	-0.05 d	0.34 a	0.41 a	0.02 a	0.34 a	0.27 a	-0.16 a
PO ₄ ³⁻	0.18 b	-0.11 b	0.04 b	0.19 d	-0.68 b	0.34 b	0.16 b	-1.00 b	0.34 b	0.43 d	-0.71 b
P _{общ}	0.92 a	0.74 a	1.00 a	0.61 a	-0.51 c	0.69 a	0.52 a	-0.31 b	0.69 a	0.64 a	-0.85 b
БПК ₅	-1.00 b	-1.00 b	-1.00 b	-0.14 b	1.00 d	0.08 b	-0.23 b	0.89 a	0.08 b	-0.37 b	0.59 b
ХПК	1.00 a	0.54 a	1.00 a	0.48 a	0.29 a	0.31 a	0.20 a	-0.19 b	0.31 a	0.13 b	-0.05 b
C _{орг} во взвеси	0.08 b	-0.03 b	-0.50 b	-0.23 b	0.29 d	0.52 a	-0.03 b	-1.00 b	0.52 a	0.39 a	-0.15 b
Нефтепродукты	-0.25 b	-0.50 b	-0.81 b	-0.63 b	-1.00 b	0.50 d	-1.00 b	0.63 b	-1.00 b	0.14 d	-1.00 b
pH	0.10 a	0.33 a	-0.34 a	0.19 a	0.66 a	-0.45 b	-0.26 b	-1.00 b	-0.45 b	-0.43 b	0.33 b
Щелочность	0.53 a	0.69 a	-0.19 a	0.42 a	0.53 a	-0.02 a	0.32 a	-1.00 b	-0.02 a	0.31 a	0.00 a
Цветность	0.24 a	0.13 a	-0.08 a	0.27 a	0.69 d	-0.01 a	0.19 a	0.23 a	0.11 a	0.55 a	0.42 a
Прозрачность	-0.53 b	-0.43 c	-0.77 b	-0.34 c	0.42 a	-1.00 b	-0.61 b	0.20 b	-1.00 b	-1.00 b	0.32 b
Электропроводность	0.72 b	0.62 b	0.08 d	0.49 d	0.10 d	0.32 d	0.30 b	0.06 b	0.57 d	0.10 b	-0.55 b
Температура воды	0.25 a	0.20 a	0.54 a	0.34 a	0.55 a	0.81 a	-0.23 a	0.09 a	0.68 a	0.72 a	-0.73 b
Температура воздуха	-0.08 a	-0.04 a	-0.30 a	0.07 a	0.86 a	-0.78 b	1.00 a	-0.35 a	-0.78 b	0.15 a	0.36 a

Таблица. 6. Коэффициенты контингенции Пирсона для таблиц сопряженности с границами между качественными классами, установленными при помощи метода *k*-средних. Обозначения те же, что в таблице 4

Физико-химические характеристики	Биологические характеристики										
	Показатели интенсивности флуоресценции					Хлорофилл				F	$\frac{E_{480}}{E_{664}}$
	$F_{т\text{ общ}}$	$F_{т\text{ фито}}$	$F_{0\text{ общ}}$	$F_{0\text{ фито}}$	$F_{0\text{ фильтра}}$	A	B	C	AЧ		
Ca ²⁺	0.33	0.27	0.13	0.36	0.19	0.07	0.04	0.20	-0.01	0.2	-0.10
Mg ²⁺	0.15	0.10	-0.01	0.18	0.18	0.20	-0.46	0.24	0.11	0.03	-0.19
Na ⁺	-0.01	-0.06	-0.21	-0.04	-0.12	0.13	-0.12	-0.26	0.12	-0.05	-0.35
K ⁺	0.00	-0.08	-0.23	-0.02	-0.20	-0.02	-0.01	0.05	-0.09	0.05	-0.06
Cl ⁻	0.10	0.10	-0.06	0.08	-0.35	0.04	0.20	-0.09	0.04	0.12	-0.23
SO ₄ ²⁻	-0.29	-0.25	-0.09	-0.24	-0.24	0.09	-0.14	-0.07	0.08	0.14	-0.03
HCO ₃ ⁻	0.34	0.30	0.12	0.37	0.21	0.07	-0.09	0.03	-0.02	0.03	-0.19
Сумма ионов	0.28	0.23	0.10	0.31	0.23	0.07	-0.09	0.24	-0.02	0.03	-0.19
NH ₄ ⁺	0.11	0.27	0.20	0.20	-0.25	0.43	-0.19	-0.17	0.43	0.23	-0.23
NO ₂ ⁻	0.12	0.11	0.07	0.19	-0.32	0.07	-0.10	-0.03	0.07	0.26	0.00
NO ₃ ⁻	-0.08	-0.07	-0.16	-0.08	-0.23	-0.19	0.11	-0.17	-0.19	0.12	0.00
N _{общ}	0.30	0.15	0.24	0.12	-0.02	0.14	0.15	0.00	0.14	0.13	-0.06
PO ₄ ³⁻	0.08	-0.05	0.01	0.09	-0.36	0.15	0.06	-0.18	0.15	0.21	-0.25
P _{общ}	0.49	0.41	0.31	0.34	-0.27	0.34	0.22	-0.07	0.34	0.36	-0.38
БПК ₅	-0.23	-0.29	-0.13	-0.05	0.34	0.02	-0.06	0.38	0.02	-0.13	0.23
ХПК	0.39	0.23	0.21	0.22	0.13	0.14	0.08	-0.04	0.14	0.06	-0.02
C _{орг} во взвеси	0.03	-0.01	-0.14	-0.11	0.14	0.24	-0.01	-0.24	0.24	0.2	-0.06
Нефтепродукты	-0.07	-0.16	-0.36	-0.22	-0.35	0.18	-0.12	0.20	-0.10	0.05	-0.14
pH	0.04	0.13	-0.10	0.08	0.30	-0.17	-0.09	-0.16	-0.17	-0.19	0.13
Щелочность	0.20	0.29	-0.05	0.19	0.25	-0.01	0.11	-0.39	-0.01	0.14	0.00
Цветность	0.11	0.06	-0.03	0.14	0.40	-0.01	0.07	0.05	0.05	0.15	0.16
Прозрачность	-0.27	-0.21	-0.32	-0.16	0.21	-0.39	-0.18	0.04	-0.34	-0.18	0.12
Электропроводность	0.33	0.27	0.03	0.20	0.04	0.14	0.09	0.01	0.27	0.04	-0.15
Температура воды	0.08	0.07	0.26	0.14	0.26	0.41	-0.07	0.01	0.27	0.33	-0.34
Температура воздуха	-0.03	-0.02	-0.11	0.03	0.56	-0.44	0.29	-0.16	-0.44	0.07	0.11

Проведено разбиение диапазонов изменения экологических характеристик на качественные классы низких и высоких значений методом максимизации коэффициентов связи. Границы классов, полученные при максимизации коэффициентов Юла и Пирсона, приведены соответственно в табл. 7 и 8.

Таблица 7. Границы качественных классов, установленные на основании максимизации абсолютного значения коэффициента ассоциации Юла. В числителе — граница качественных классов для биологической характеристики, в знаменателе — граница качественных классов для физико-химической характеристики. Границы приведены только для значимых по критерию χ^2 связей. Обозначения те же, что в таблице 4

Физико-химические характеристики	Биологические характеристики										
	Показатели интенсивности флуоресценции					Хлорофилл				F	$\frac{E_{480}}{E_{664}}$
	F_m общ	F_m фито	F_0 общ	F_0 фито	F_0 фильтрата	A	B	C	AЧ		
Ca ²⁺	<u>2.25</u> 1.22	<u>0.50</u> 1.50	<u>0.97</u> 1.35		<u>0.42</u> 1.35	<u>10.0</u> 1.50					
Mg ²⁺	<u>2.25</u> 0.49	<u>0.50</u> 0.61	<u>0.97</u> 0.49	<u>0.70</u> 0.49	<u>0.42</u> 0.49						
Na ⁺	<u>2.25</u> 0.11	<u>1.98</u> 0.11	<u>1.06</u> 0.11	<u>0.70</u> 0.11	<u>0.26</u> 0.11						
Cl ⁻					<u>0.48</u> 0.09						
SO ₄ ²⁻		<u>0.79</u> 0.19			<u>0.37</u> 0.26						
HCO ₃ ⁻	<u>2.25</u> 1.59	<u>0.50</u> 1.98	<u>0.97</u> 1.68	<u>0.70</u> 1.59							
Сумма ионов	<u>2.25</u> 3.64	<u>0.50</u> 4.54	<u>0.97</u> 4.00		<u>0.42</u> 4.00						
NH ₄ ⁺											<u>1.52</u> 0.07
NO ₂ ⁻			<u>0.71</u> 0.00	<u>0.32</u> 0.01	<u>0.31</u> 0.01	<u>8.45</u> 0.00		<u>0.94</u> 0.00			
NO ₃ ⁻			<u>0.97</u> 0.07		<u>0.48</u> 0.05	<u>16.9</u> 0.08			<u>12.7</u> 0.08	<u>5.80</u> 0.08	
N _{общ}	<u>0.96</u> 0.88	<u>0.47</u> 0.88	<u>0.50</u> 1.00	<u>0.19</u> 0.88	<u>0.49</u> 0.77	<u>8.00</u> 1.00		<u>0.94</u> 1.00	<u>6.40</u> 1.00	<u>6.70</u> 0.88	<u>0.97</u> 0.77
PO ₄ ³⁻					<u>0.41</u> 0.04	<u>15.5</u> 0.02			<u>11.5</u> 0.02		<u>1.44</u> 0.04
P _{общ}	<u>1.89</u> 0.06	<u>1.47</u> 0.06	<u>0.94</u> 0.06	<u>0.52</u> 0.06	<u>0.26</u> 0.06	<u>10.0</u> 0.09	<u>0.51</u> 0.07	<u>2.10</u> 0.08		<u>2.20</u> 0.09	<u>1.44</u> 0.09
БПК ₅	<u>0.96</u> 1.82	<u>0.50</u> 1.82	<u>0.51</u> 1.82	<u>0.52</u> 1.55	<u>0.42</u> 2.59	<u>8.00</u> 2.90			<u>6.40</u> 2.90	<u>5.50</u> 1.51	<u>0.97</u> 1.48
ХПК	<u>1.89</u> 33.1	<u>1.47</u> 33.1	<u>0.88</u> 31.2	<u>0.52</u> 33.1	<u>0.31</u> 45.0		<u>0.35</u> 31.8	<u>0.91</u> 45.0			<u>1.29</u> 33.8
C _{орг} во взвеси					<u>0.31</u> 1.30	<u>22.0</u> 0.80		<u>0.94</u> 1.10			
pH	<u>1.64</u> 8.30	<u>1.07</u> 8.30	<u>0.72</u> 8.20	<u>0.47</u> 8.30	<u>0.41</u> 7.94		<u>0.02</u> 8.06		<u>6.40</u> 8.00		<u>1.02</u> 8.30
Щелочность		<u>1.12</u> 85.4	<u>0.70</u> 85.4	<u>0.34</u> 85.4	<u>0.39</u> 90.3	<u>8.93</u> 97.6	<u>0.00</u> 90.3	<u>1.90</u> 114.7	<u>7.30</u> 100.0	<u>5.40</u> 90.3	<u>1.14</u> 97.6
Цветность	<u>0.86</u> 70.0	<u>0.50</u> 70.0	<u>0.50</u> 70.0	<u>0.18</u> 70.0	<u>0.24</u> 60.0	<u>14.8</u> 70.0	<u>0.02</u> 65.0	<u>0.94</u> 70.0	<u>6.40</u> 75.0		<u>1.18</u> 55.0
Прозрачность	<u>1.89</u> 160.0	<u>1.47</u> 160.0	<u>0.94</u> 160.0	<u>0.52</u> 160.0	<u>0.49</u> 160.0	<u>21.0</u> 160.0	<u>0.35</u> 160.0	<u>0.96</u> 130.0	<u>6.40</u> 110.0	<u>7.27</u> 160.0	<u>0.92</u> 160.0
Электропроводность	<u>1.24</u> 158.3	<u>0.93</u> 159.8	<u>0.75</u> 161.9	<u>0.33</u> 159.8	<u>0.26</u> 155.9	<u>3.05</u> 155.5		<u>0.69</u> 156.2	<u>2.05</u> 154.8	<u>1.30</u> 174.4	<u>0.95</u> 160.0
Температура воды	<u>0.67</u> 23.8	<u>0.35</u> 23.8	<u>0.44</u> 24.5	<u>0.12</u> 23.8	<u>0.42</u> 18.6	<u>3.57</u> 24.5	<u>0.10</u> 18.2	<u>0.70</u> 25.0	<u>2.80</u> 24.5	<u>5.80</u> 18.0	<u>0.92</u> 18.0
Температура воздуха	<u>1.24</u> 24.2	<u>0.76</u> 25.5	<u>0.53</u> 20.4	<u>0.38</u> 24.2	<u>0.52</u> 18.9	<u>11.2</u> 25.5	<u>0.05</u> 25.5				<u>0.89</u> 24.2

Таблица 8. Границы качественных классов, установленные на основании максимизации абсолютного значения коэффициента контингенции Пирсона. В числителе — граница качественных классов для биологической характеристики, в знаменателе — граница качественных классов для физико-химической характеристики. Границы приведены только для значимых по критерию χ^2 связей. Обозначения те же, что в таблице 4

Физико-химические характеристики	Биологические характеристики										
	Показатели интенсивности флуоресценции					Хлорофилл				F	$\frac{E_{480}}{E_{664}}$
	F_m общ	F_m фито	F_0 общ	F_0 фито	F_0 фильтрата	A	B	C	AЧ		
Ca ²⁺	<u>1.94</u> 1.52	<u>1.66</u> 1.52	<u>0.88</u> 1.52	<u>0.52</u> 1.50	<u>0.42</u> 1.35	<u>7.70</u> 1.23	<u>0.49</u> 1.35	<u>1.84</u> 1.22	<u>5.50</u> 1.25	<u>7.40</u> 1.23	<u>0.85</u> 1.45
Mg ²⁺	<u>1.94</u> 0.54	<u>1.66</u> 0.54	<u>0.88</u> 0.54	<u>0.59</u> 0.54	<u>0.42</u> 0.49	<u>7.70</u> 0.49	<u>0.05</u> 0.48	<u>1.90</u> 0.54	<u>7.60</u> 0.47	<u>7.40</u> 0.49	<u>0.85</u> 0.59
Na ⁺	<u>2.25</u> 0.11	<u>1.98</u> 0.11	<u>1.06</u> 0.11	<u>0.70</u> 0.11	<u>0.26</u> 0.11						<u>0.85</u> 0.11
Cl ⁻					<u>0.24</u> 0.09			<u>1.90</u> 0.07			
SO ₄ ²⁻	<u>1.57</u> 0.24	<u>1.23</u> 0.24	<u>0.50</u> 0.29	<u>0.47</u> 0.24	<u>0.37</u> 0.26			<u>1.65</u> 0.29			<u>0.87</u> 0.23
HCO ₃ ⁻	<u>1.94</u> 1.89	<u>1.66</u> 1.89	<u>0.88</u> 1.88	<u>0.59</u> 1.89	<u>0.42</u> 1.68	<u>7.70</u> 1.59	<u>0.49</u> 1.64	<u>2.10</u> 1.76	<u>5.50</u> 1.43	<u>7.40</u> 1.76	<u>0.87</u> 1.59
Сумма ионов	<u>1.94</u> 4.40	<u>1.66</u> 4.40	<u>0.88</u> 4.30	<u>0.58</u> 4.40	<u>0.42</u> 4.00	<u>7.70</u> 3.70	<u>0.49</u> 3.90	<u>2.50</u> 3.70	<u>4.30</u> 4.50	<u>12.8</u> 4.40	<u>0.85</u> 4.40
NH ₄ ⁺					<u>0.31</u> 0.07						<u>1.52</u> 0.07
NO ₂ ⁻	<u>1.24</u> 0.006		<u>0.71</u> 0.004	<u>0.32</u> 0.006	<u>0.31</u> 0.005	<u>8.50</u> 0.004		<u>0.94</u> 0.001			
NO ₃ ⁻			<u>0.97</u> 0.05		<u>0.51</u> 0.04	<u>16.9</u> 0.07		<u>2.50</u> 0.03	<u>1.60</u> 0.03	<u>8.00</u> 0.05	<u>1.52</u> 0.03
N _{общ}	<u>0.96</u> 0.88	<u>0.47</u> 0.88	<u>0.50</u> 0.88	<u>0.19</u> 0.88	<u>0.49</u> 0.77	<u>14.8</u> 0.88	<u>0.35</u> 0.88	<u>1.13</u> 1.00	<u>6.70</u> 0.88	<u>10.8</u> 0.88	<u>0.97</u> 0.80
PO ₄ ³⁻	<u>1.57</u> 0.019				<u>0.41</u> 0.029	<u>14.8</u> 0.024			<u>5.00</u> 0.024	<u>10.6</u> 0.018	<u>1.14</u> 0.019
P _{общ}	<u>1.64</u> 0.079	<u>1.23</u> 0.079	<u>0.97</u> 0.079	<u>0.42</u> 0.073	<u>0.27</u> 0.070	<u>11.8</u> 0.068	<u>0.51</u> 0.070	<u>1.01</u> 0.073	<u>3.80</u> 0.061	<u>9.60</u> 0.068	<u>1.01</u> 0.079
БПК ₅	<u>1.01</u> 1.82	<u>0.50</u> 1.82	<u>0.51</u> 1.82	<u>0.52</u> 1.55	<u>0.42</u> 2.59	<u>8.30</u> 2.66		<u>1.17</u> 1.92	<u>3.80</u> 1.60	<u>8.30</u> 2.66	<u>0.90</u> 1.86
ХПК	<u>1.58</u> 33.3	<u>1.47</u> 34.5	<u>0.72</u> 40.5	<u>0.52</u> 34.5	<u>0.42</u> 39.7	<u>14.8</u> 34.5	<u>0.35</u> 31.8	<u>0.91</u> 43.2	<u>4.80</u> 33.8	<u>10.70</u> 34.5	<u>1.29</u> 33.8
C _{орг} во взвеси	<u>0.96</u> 1.20	<u>0.48</u> 1.20	<u>0.71</u> 1.20	<u>0.18</u> 1.00	<u>0.31</u> 1.30	<u>14.8</u> 1.10		<u>1.13</u> 1.20		<u>10.70</u> 1.10	
pH	<u>0.96</u> 8.20	<u>1.07</u> 8.30	<u>0.72</u> 8.2	<u>0.38</u> 8.3	<u>0.41</u> 8.0		<u>0.02</u> 8.1			<u>7.30</u> 8.00	<u>1.14</u> 8.3
Щелочность		<u>1.12</u> 85.4	<u>0.70</u> 85.4	<u>0.34</u> 85.4	<u>0.36</u> 90.3	<u>8.90</u> 97.6	<u>0.00</u> 90.3	<u>1.90</u> 114.7	<u>5.4</u> 90.3	<u>8.30</u> 100.0	<u>1.14</u> 97.6
Цветность	<u>1.58</u> 65	<u>0.50</u> 70	<u>0.71</u> 65	<u>0.37</u> 65	<u>0.49</u> 65	<u>14.8</u> 70	<u>0.02</u> 65	<u>0.92</u> 65		<u>9.80</u> 70	<u>1.18</u> 55
Прозрачность	<u>1.24</u> 150	<u>0.87</u> 150	<u>0.97</u> 120	<u>0.31</u> 150	<u>0.49</u> 160	<u>16.9</u> 130	<u>0.47</u> 140	<u>0.96</u> 130	<u>5.9</u> 130	<u>11.5</u> 140	<u>0.90</u> 115
Электропроводность	<u>1.24</u> 162	<u>0.93</u> 162	<u>0.65</u> 162	<u>0.37</u> 168	<u>0.26</u> 156	<u>10.8</u> 162		<u>1.73</u> 169	<u>1.3</u> 174	<u>2.90</u> 155	<u>0.99</u> 160
Температура воды	<u>0.67</u> 23.8	<u>0.35</u> 23.8	<u>0.47</u> 23.8	<u>0.12</u> 23.8	<u>0.36</u> 20.8	<u>4.10</u> 18.4	<u>0.16</u> 18.4	<u>1.10</u> 24.5	<u>3.8</u> 18.4	<u>3.30</u> 18.4	<u>1.32</u> 18.9
Температура воздуха	<u>1.24</u> 24.2	<u>0.76</u> 25.5	<u>0.53</u> 20.4	<u>0.38</u> 24.2	<u>0.35</u> 20.4	<u>11.2</u> 25.5	<u>0.05</u> 25.5			<u>7.30</u> 26	<u>0.89</u> 24.2

Для полученных классов рассчитаны сами коэффициенты Юла (табл. 9) и Пирсона (табл. 10). Оценена значимость всех рассчитанных коэффициентов.

Таблица 9. Коэффициенты ассоциации Юла для таблиц сопряженности с границами между качественными классами, установленными на основании максимизации абсолютного значения коэффициента ассоциации. Обозначения те же, что в таблице 4. Прочерки означают, что границы качественных классов не найдены в связи с невыполнением условий на минимальную представительность ячеек

Физико-химические характеристики	Биологические характеристики										
	Показатели интенсивности флуоресценции					Хлорофилл				F	$\frac{E_{480}}{E_{664}}$
	F_m общ	F_m фито	F_0 общ	F_0 фито	F_0 фильтрата	A	B	C	AЧ		
Ca ²⁺	1.00 a	1.00 d	0.84 a	1.00 a	0.68 a	1.00 d	-1.00 b	1.00 a	1.00 d	1.00 d	-1.00 b
Mg ²⁺	1.00 a	1.00 d	1.00 a	1.00 a	0.80 a	1.00 d	-1.00 b	1.00 a	1.00 d	1.00 d	-1.00 c
Na ⁺	0.77 a	1.00 a	0.77 a	1.00 a	-0.80 c	0.58 d	-1.00 b	1.00 a	0.53 d	1.00 d	-1.00 b
K ⁺	—	—	—	—	-0.70 c	0.51 a	—	1.00 a	0.51 a	-0.53 c	-0.24 b
Cl ⁻	-0.44 c	-0.48 c	-0.62 c	-0.48 c	-1.00 b	-0.51 c	0.66 d	-1.00 c	-0.58 c	1.00 d	-1.00 b
SO ₄ ²⁻	1.00 d	-0.75 c	1.00 d	0.70 d	-0.69 b	-0.50 c	0.70 d	1.00 d	-0.58 c	-1.00 c	1.00 a
HCO ₃ ⁻	1.00 a	1.00 d	0.85 a	1.00 a	-0.73 c	1.00 a	-1.00 b	1.00 a	1.00 d	1.00 a	-1.00 c
Сумма ионов	1.00 a	1.00 d	0.86 a	1.00 a	0.73 a	1.00 d	-1.00 b	1.00 a	1.00 d	1.00 d	-1.00 b
NH ₄ ⁺	-0.43 c	0.59 a	-0.67 c	0.59 a	-1.00 c	-1.00 c	1.00 a	-0.72 c	-1.00 c	1.00 d	-1.00 b
NO ₂ ⁻	-0.67 c	-0.49 c	-0.50 b	0.55 d	-0.88 c	-0.66 c	0.45 d	-1.00 c	-0.70 c	0.50 d	0.44 d
NO ₃ ⁻	0.42 a	0.49 a	-1.00 b	0.54 a	-0.88 b	-1.00 b	0.51 d	-1.00 b	-1.00 b	-1.00 b	1.00 d
N _{общ}	0.73 d	0.87 d	1.00 d	0.80 d	0.67 a	1.00 d	-0.60 c	1.00 d	1.00 d	0.92 a	-0.90 c
PO ₄ ³⁻	0.60 a	0.57 a	-0.48 b	0.48 a	-0.88 b	0.71 a	-0.40 b	0.38 d	0.71 a	1.00 d	-1.00 b
P _{общ}	1.00 a	1.00 a	1.00 a	1.00 a	-1.00 c	1.00 d	0.78 a	-0.69 b	1.00 d	1.00 d	-1.00 b
БПК ₅	0.67 d	0.78 d	0.84 d	0.75 a	0.93 d	1.00 d	0.51 d	1.00 a	1.00 d	0.78 a	-0.79 c
ХПК	1.00 a	1.00 a	1.00 a	1.00 a	1.00 d	1.00 a	0.80 a	1.00 d	1.00 a	1.00 a	-0.77 c
C _{орг} во взвеси	0.73 d	-1.00 b	1.00 d	-1.00 b	1.00 d	1.00 a	-0.44 c	0.80 d	0.72 a	0.57 a	-0.43 b
Нефтепродукты	-0.26 b	1.00 d	0.35 d	0.57 d	0.57 d	0.52 d	-0.50 b	-0.40 b	0.61 d	0.63 d	-0.52 b
pH	-0.77 b	-0.85 b	0.58 d	-0.77 b	0.79 a	-0.70 c	0.62 d	-0.54 c	-1.00 c	-1.00 c	0.79 d
Щелочность	0.57 a	0.69 a	0.73 a	0.63 a	0.59 a	-0.87 c	0.83 a	-1.00 b	-0.86 c	0.72 a	0.71 a
Цветность	1.00 d	1.00 d	1.00 d	0.75 d	0.93 d	0.73 d	0.55 d	1.00 d	1.00 d	-0.73 c	0.64 a
Прозрачность	-1.00 b	-1.00 b	-1.00 b	-1.00 b	-0.75 b	-1.00 b	-0.85 b	-0.89 c	-1.00 c	-1.00 b	1.00 d
Электропроводность	0.81 a	0.81 a	0.67 a	0.80 a	-0.68 c	-0.74 c	-0.36 c	-0.49 c	-0.89 c	0.53 d	-0.81 c
Температура воды	0.80 d	0.62 d	0.89 d	0.58 d	0.85 a	1.00 d	-0.72 c	0.86 d	1.00 d	1.00 a	-1.00 c
Температура воздуха	-0.64 b	-0.72 b	0.85 d	-0.66 b	1.00 a	-0.86 c	0.88 a	-0.71 c	-1.00 b	-0.68 c	0.88 d

Таблица 10. Коэффициенты контингенции Пирсона для таблиц сопряженности с границами между качественными классами, установленными на основании максимизации абсолютного значения коэффициента контингенции. Обозначения те же, что в таблице 4. Прочерки означают, что границы качественных классов не найдены в связи с невыполнением условий на минимальную представительность ячеек

Физико-химические характеристики	Биологические характеристики										
	Показатели интенсивности флуоресценции					Хлорофилл				F	$\frac{E_{480}}{E_{664}}$
	F_m общ	F_m фито	F_0 общ	F_0 фито	F_0 фильтрата	A	B	C	AЧ		
Ca ²⁺	0.49	0.49	0.47	0.47	0.29	0.56	-0.42	0.43	0.49	0.49	-0.64
Mg ²⁺	0.39	0.39	0.40	0.35	0.32	0.48	-0.50	0.49	0.41	0.54	-0.77
Na ⁺	0.29	0.37	0.29	0.37	-0.32	0.26	-0.33	0.29	0.33	0.26	-0.47
K ⁺	—*	—	—	—	-0.24	0.18	—	0.33	-0.23	0.18	-0.09
Cl ⁻	-0.21	-0.22	-0.25	-0.22	-0.80	0.19	0.36	-0.55	0.30	-0.21	-0.35
SO ₄ ²⁻	-0.44	-0.37	0.31	-0.33	-0.37	-0.25	0.35	0.46	-0.34	-0.31	0.47
HCO ₃ ⁻	0.52	0.52	0.50	0.48	0.32	0.48	-0.45	0.41	0.44	0.44	-0.57
Сумма ионов	0.46	0.46	0.45	0.42	0.34	0.52	-0.45	0.44	0.43	0.49	-0.70
NH ₄ ⁺	0.20	0.29	-0.34	0.29	-0.75	-0.33	0.43	-0.34	0.45	-0.37	-0.50
NO ₂ ⁻	0.26	0.23	-0.27	0.29	-0.56	-0.32	0.22	-0.30	0.26	-0.27	0.20
NO ₃ ⁻	0.21	0.23	-0.31	0.23	-0.42	-0.43	0.24	-0.64	0.49	-0.48	-0.54
N _{общ}	0.37	0.40	0.42	0.39	0.28	0.60	0.29	0.40	0.53	0.60	-0.54
PO ₄ ³⁻	0.28	0.23	-0.24	0.25	-0.47	0.36	-0.17	0.17	0.29	0.37	-0.44
P _{общ}	0.56	0.52	0.58	0.57	-0.32	0.60	0.39	0.33	0.60	0.64	-0.68
БПК ₅	0.37	0.44	0.48	0.27	0.66	0.38	0.27	0.39	0.31	0.49	-0.32
ХПК	0.42	0.40	0.58	0.42	0.63	0.55	0.32	0.31	0.40	0.58	-0.45
C _{орг} во взвеси	0.28	0.28	0.29	0.35	0.33	0.42	-0.23	0.34	0.28	0.39	-0.21
Нефтепродукты	0.12	0.31	0.17	0.19	-0.24	0.24	-0.25	-0.20	0.34	0.29	-0.23
pH	0.31	-0.37	0.32	-0.30	0.44	-0.27	0.34	-0.21	-0.27	-0.37	0.41
Щелочность	0.25	0.29	0.37	0.29	0.30	-0.42	0.51	-0.39	0.33	-0.44	0.39
Цветность	0.40	0.34	0.53	0.26	0.55	0.36	0.26	0.40	-0.25	0.34	0.29
Прозрачность	-0.48	-0.55	-0.51	-0.51	-0.25	-0.61	-0.44	-0.52	-0.51	-0.58	0.62
Электропроводность	0.41	0.40	0.32	0.39	-0.24	0.28	-0.17	-0.19	0.19	-0.31	-0.37
Температура воды	0.30	0.24	0.34	0.23	0.49	0.76	-0.30	0.32	0.47	0.73	-0.43
Температура воздуха	-0.35	-0.41	0.53	-0.36	0.61	-0.54	0.59	-0.33	-0.37	-0.54	0.55

Таблицы 8 и 10 можно интерпретировать в отношении каждой пары из биологической и физико-химической характеристики. Например, для общего хлорофилла «а» выявлена прямая двусторонняя связь с концентрацией Ca²⁺ в воде (мера связи в виде коэффициента контингенции составляет 0.56). Низкие (менее 1.23 мг-экв/л) концентрации кальция сопряжены с низким (менее 7.7 мкг/л) содержаниями общего хлорофилла «а». Высокие концентрации кальция (более 1.23 мг-экв/л) сопряжены с высоким (более 7.7 мкг/л) содержанием общего хлорофилла «а». Поскольку связь двусторонняя, верно и обратное утверждение о том, что с низким содержанием хлорофилла «а» сопряжены низкие концентрации кальция, а с высоким, соответственно, высокие, что характеризуется тем же коэффициентом контингенции 0.56.

В случаях значимой связи, но низкого коэффициента контингенции, можно использовать коэффициент ассоциации Юла. Например, для содержания хлорофилла «с» и азота нитритов

выявлена обратная связь (коэффициент контингенции « -0.30 »). В тоже время коэффициент ассоциации для этой пары равен « -1 ». Имея дополнительную информацию о том, что в таблице сопряженности наиболее пуста ячейка «с», можно утверждать, что все низкие концентрации азота нитритов (менее 0.001 мкг N/л) сопряжены с высоким (более 0.94 мкг/л) содержанием хлорофилла «с», и низкое (менее 0.94 мкг/л) содержание хлорофилла «с» сопряжено с высокими (более 0.001 , мкг N/л) концентрациями азота нитритов. Поскольку коэффициент ассоциации — односторонняя мера связи, обратные утверждения не верны.

Исходя из таблиц 8 и 10, можно давать односторонние интерпретации качественных связей, например, низкие значения концентраций хлорофилла «а» (менее 10 мкг/л) сопряжены с низким содержанием общего фосфора (менее 0.09 мг P/л) и кальция (менее 1.5 мг-экв/л), а высокие содержания фосфора (более 0.09 мг P/л) и кальция (более 1.5 мг-экв/л) сопряжены с высокими (более 10 мкг/л) концентрациями хлорофилла «а» (мера связи в виде коэффициента ассоциации Юла равна 1).

5. Обсуждение результатов

5.1. Корреляционный и множественный регрессионный анализ

Как видно из таблице 4, при помощи корреляционного анализа данных выявлено сравнительно мало значимых связей. Максимальный коэффициент корреляции $R = 0.61$, соответствующий умеренной связи по шкале Чеддока, был выявлен между содержанием общего хлорофилла «а» и общего фосфора. Соответствующий данной связи коэффициент детерминации составил $R^2 = 0.37$, что говорит о том, что вариации значений общего хлорофилла «а» на 37% могут быть объяснены вариациями содержания общего фосфора.

Регрессионный анализ позволил определить вес вклада каждого фактора (коэффициенты в уравнении регрессии) и исследовать зависимость биологических характеристик от рассчитанного по уравнению регрессии показателя, учитывающего все включенные в анализ факторы. Для значений биологических характеристик часто отсутствуют те или иные данные по физико-химическим характеристикам. В данном случае для анализа были использованы 225 значений показателей интенсивности флуоресценции, 152 значения содержания хлорофилла и показателя его соотношения с каратиноидами и от 38 до 187 значений показателей физико-химических характеристик. Таким образом, применение множественного регрессионного анализа было осложнено недостаточным количеством данных. Для найденных регрессионных зависимостей связь между переменными практически не была обнаружена. Наибольший коэффициент детерминации (0.35) был получен для хлорофилла «а» и концентраций Mg^{2+} , Na^+ , K^+ , Cl^- , SO_4^{2-} , уравнение нелинейной регрессии: $\ln A = 0.02Mg^{2+} + 0.06Na^+ - 0.09K^+ - 0.29Cl^- + 0.61SO_4^{2-} - 0.72$.

Полученные при помощи корреляционного и регрессионного анализа результаты могут свидетельствовать как об отсутствии связи, так и о том, что данные плохо организованы. В связи с этим были выделены качественные классы характеристик и проанализированы связи между классами.

5.2. Коэффициенты Юла и Пирсона для метода k -средних кластерного анализа

Как видно из таблицы 7, результаты поиска связей для качественных классов, выделенных методом k -средних кластерного анализа, в целом несколько хуже, чем результаты поиска связей при помощи корреляционного анализа (выявлено меньше значимых связей, большинство абсолютных величин коэффициентов Пирсона ниже абсолютных величин коэффициентов корреляции). Однако коэффициенты ассоциации Юла (табл. 5), характеризующие одностороннюю связь, т. е. степень пустоты одной из ячеек таблицы сопряженности, в некоторых случаях дают хорошие результаты. Тем не менее ввиду малого числа найденных значимых связей метод

k -средних для отдельного выделения границ качественных классов индикаторов и факторов можно считать нецелесообразным.

5.3. Границы качественных классов, полученные методом МКС при максимизации коэффициентов Юла и Пирсона

Границы качественных классов при максимизации коэффициента Юла и Пирсона методом МКС приведены соответственно в таблицах 7 и 8. Границы классов для биологических характеристик, установленные по коэффициентам Юла, равны (с учетом 20 % погрешности измерений) аналогичным границам, установленным по коэффициентам Пирсона в 74 % случаев, в 10 % случаев границы по коэффициентам Юла больше границ по коэффициентам Пирсона, в 16 % случаев — меньше. Границы классов для физико-химических характеристик, установленные по коэффициентам Юла, равны аналогичным границам по коэффициентам Пирсона в 86% случаев, в 7 % случаев границы по коэффициентам Юла больше границ по коэффициентам Пирсона, в 7 % случаев — меньше. Таким образом большая часть границ классов биологических и физико-химических характеристик приблизительно совпадают.

5.4. Значения коэффициентов, полученные методом МКС при максимизации коэффициентов Юла и Пирсона

При максимизации коэффициента Юла было получено 149 значимых связей, при максимизации коэффициента Пирсона — 200, в то время как при расчете коэффициентов корреляции только 81 из 275 возможных связей. Тем самым метод МКС помог выявить значительно больше значимых связей, чем коэффициент корреляции.

Поскольку большая часть коэффициентов Пирсона значима даже в тех случаях, когда не было выявлено значимой связи по коэффициентам корреляции, метод МКС выделения границ между качественными классами выявляет взаимосвязи между характеристиками и находит значимые связи, не обнаруженные при помощи корреляционного анализа. Напомним, что для таблиц сопряженности с двумя классами качества по обеим характеристикам для заданного общего числа наблюдений критерий χ^2 возрастает пропорционально коэффициенту контингенции Пирсона. Из этого следует, что нахождение большего (чем при максимизации коэффициента контингенции) числа значимых связей для таких таблиц сопряженности невозможно. Если результаты двустороннего поиска не устраивают исследователя из-за низких коэффициентов Пирсона, целесообразно провести односторонний поиск исходя из максимизации коэффициента Юла (табл. 9).

6. Заключение

Анализ исходных данных показал, что зависимости между характеристиками не могут быть полностью описаны коэффициентами корреляции или регрессии. Применение метода k -средних для разбиения характеристик на качественные классы нецелесообразно, т. к. выявлено меньше значимых связей, чем при расчете коэффициентов корреляции, большинство абсолютных величин коэффициентов Пирсона ниже абсолютных величин коэффициентов корреляции. Применение метода МКС для поиска двух качественных классов по обеим характеристикам позволило как выявить значимые связи, не обнаруживаемые при помощи корреляционного анализа, так и лучше описать большинство выявленных связей. Ввиду того, что методы детерминационного анализа представляются более перспективными, поскольку характеризуют не отклонение взаимного распределения характеристик от случайности, а возможность прогноза качественных значений одной характеристики по качественным значениям другой, в дальнейших работах планируется применение метода МКС с использованием коэффициентов связи, взятых из детерминационного анализа, с введением большего числа классов качества и с использованием мер связи для ранжированных признаков.

Список литературы

- Афффи А., Эйзен С.* Статистический анализ: Подход с использованием ЭВМ. Пер. с англ. — М.: Мир, 1982. — 488 с.
- Айвазян С. А., Бухштабер В. М., Енюков И. С., Мешалкин Л. Д.* Прикладная статистика: Классификация и снижение размерности: Справ. изд. — М.: Финансы и статистика, 1989. — Том 3. — 607 с.
- Айвазян С. А., Енюков И. С., Мешалкин Л. Д.* Прикладная статистика: Исследование зависимостей: Справ. изд. — М.: Финансы и статистика, 1985. — Том 2. — 487 с.
- Голубков Е. П.* Маркетинговые исследования: теория, методология и практика. — М.: Финпресс, 1998. — 416 с.
- ГОСТ 17.1.04.02.90. Вода. Методика спектрофотометрического определения хлорофилла «а». — М.: Гос. комитет СССР по охране природы, 1990. — 15 с.
- Данные совместных измерений биологических и физико-химических характеристик экосистемы Рыбинского водохранилища / *Бикбулатов Э. С., Бикбулатова Е. М., Булгаков Н. Г., Еришов Ю. В., Конюхов И. В., Копылов А. И., Корнева Л. Г., Лазарева В. И., Левич А. П., Литвинов А. С., Масленникова Т. С., Митропольская И. В., Осипов В. А., Отюкова Н. Г., Поддубный С. А., Поромов А. А., Пырина И. Л., Рисник Д. В., Соколова Е. А., Степанова И. Э., Цельмович О. Л.* — М.: МАКС Пресс, 2011. — 67 с.
- Дубина И. Н.* Математические основы эмпирических социально-экономических исследований: учебное пособие. — Барнаул: Изд-во Алт. ун-та, 2006. — 263 с.
- Зайдель А. М.* Ошибки измерений физических величин. — Л.: Наука, 1974. — 108 с.
- Кобзарь А. И.* Прикладная математическая статистика. Для инженеров и научных работников. — М.: ФИЗМАТЛИТ, 2006. — 816 с.
- Маторин Д. Н., Осипов В. А., Рубин А. Б.* Методика измерений обилия и индикации изменений состояния фитопланктона в природных водах флуоресцентным методом. Теоретические и практические аспекты. Учебно-методическое пособие. М.: ПКЦ Альтекс, 2012. 138 с.
- Методика выполнения измерений массовых концентраций катионов калия, натрия, лития, магния, кальция, аммония, стронция, бария в пробах питьевых, природных, сточных вод методом капиллярного электрофореза с использованием системы капиллярного электрофореза «Капель». ПНД Ф 14.1:2:4.167-2000. — М., 2007. — 27 с.
- Методика выполнения измерений массовых концентраций хлорид-ионов, нитрит-ионов, сульфат-ионов, нитрат-ионов, фторид-ионов и фосфат-ионов в пробах природных, питьевых и очищенных сточных вод с применением системы капиллярного электрофореза «Капель». ПНД Ф 14.1:2:4.157-99. — М., 2009. — 32 с.
- Миркин Б. Г.* Анализ качественных признаков и структур. — М.: Статистика, 1980. — 319 с.
- Налимов В. В., Чернова Н. А.* Статистические методы планирования экстремальных экспериментов. — М.: Наука, 1965. — 340 с.
- Паниотто В. И., Максименко В. С.* Количественные методы в социологических исследованиях. — Киев, Наукова думка, 1982. — 272 с.
- Пырина И. Л.* Многолетние исследования содержания пигментов фитопланктона Рыбинского водохранилища // Биология внутренних вод. 2000. — № 1. — С. 37–44.
- Пырина И. Л., Сигарева Л. Е.* Содержание пигментов фитопланктона в Рыбинском водохранилище в различные по гидрометеорологическим условиям годы (1972–1976) // Биология и экология водных организмов. — Л.: Наука, 1986. — С. 65–89.
- Пырина И. Л., Сметанин М. М., Сметанина Т. Л.* Статистический подход к оценке концентраций пигментов и биомассы фитопланктона // Оценка продуктивности фитопланктона. — Новосибирск: Наука, 1993. — С. 30–44.

- Рисник Д. В., Рыбка К. Ю.* О методе поиска сопряженностей между биологическими и физико-химическими характеристиками для натуральных данных на примере экосистемы Рыбинского водохранилища // Материалы Всероссийской научно-практической конференции «Мологский край и Рыбинское водохранилище». — М.: МАКС Пресс, 2011. — С. 169–175.
- Burkholder P. R., Burkholder L. M., Rivero J. A.* Chlorophyll «a» in some corals and marine plants // *Nature*. 1959. Vol. 183, № 4671. P. 1338–1339.
- Chaddock R. E.* Principles and methods of statistics – Boston: Houghton Mifflin Company, 1925. 471 pp.
- Lorenzen C. J.* Determination of chlorophyll and pheopigments: spectrophotometric equations // *Limnology and Oceanography*. 1967. V. 12, № 2. P. 343–346.
- Lorenzen C. J., Jeffrey S. W.* Determination of chlorophyll in seawater // *UNESCO Technical Papers in Marine Sciences*. Paris: UNESCO, 1980. № 35. P. 3–20.

УДК: 519.8

Математическое моделирование стохастических равновесий и бизнес-циклов модели Гудвина

И. А. Башкирцева^a, Е. Д. Екатеринчук^b, Т. В. Рязанова^c,
А. А. Сысолятина^d

Уральский федеральный университет,
Россия, 620083, г. Екатеринбург, пр. Ленина, д. 51

E-mail: ^a irina.bashkirtseva@usu.ru, ^b ek.ekaterinchuk@gmail.com, ^c tatyana.ryazanova@usu.ru,
^d anna.sysolyatina@usu.ru

Получено 21 ноября 2012 г.,
после доработки 20 декабря 2012 г.

В работе рассматривается модель экономической динамики Гудвина, находящаяся под воздействием случайных возмущений. Проведен полный параметрический анализ равновесий и циклов детерминированной системы. Исследованы вероятностные свойства аттракторов стохастической системы с использованием техники функций стохастической чувствительности и метода прямого численного моделирования. Обсуждается явление генерации стохастических бизнес-циклов в зоне, где исходная детерминированная модель имеет лишь устойчивые равновесия.

Ключевые слова: модель Гудвина, бизнес циклы, случайные возмущения, функция стохастической чувствительности, индуцированные шумом переходы

Mathematical modeling of stochastic equilibria and business cycles of Goodwin model

I. A. Bashkirtseva, E. D. Ekaterinchuk, T. V. Ryazanova, A. A. Sysolyatina

Ural Federal University, 51, Lenina avenue, Ekaterinburg, 620083, Russia

Abstract. — The Goodwin dynamical model under the random external disturbances is considered. A full parametrical analysis for equilibria and cycles of deterministic model is developed. We study probabilistic properties of stochastic attractors using stochastic sensitivity functions technique and numerical methods. A phenomenon of the generation of stochastic business cycles in the zones of stable equilibria is discussed.

Keywords: Goodwin's model, business cycle, random perturbation, stochastic sensitivity function, noise-induced transitions

Citation: *Computer Research and Modeling*, 2013, vol. 5, no. 1, pp. 107–118 (Russian).

Исследование выполнено при поддержке Министерства образования и науки Российской Федерации, соглашение № 14.А18.21.0364 и АВЦП № 1.1099.2011, грантом РФФИ № 12-01-31210 и при финансовой поддержке молодых ученых УрФУ в рамках реализации программы развития УрФУ.

1. Введение

Основные экономические и финансовые показатели, такие как валовый внутренний продукт, уровень производства, уровень безработицы, курсы валют, банковские процентные ставки, курсы ценных бумаг и т. д. повсеместно демонстрируют сложную динамику, сочетающую как временную стабилизацию, так и большеамплитудные, зачастую апериодические колебания. Традиционно в экономической динамике исследовались флуктуации вблизи равновесий для линейных моделей [Scarth, 1996; Gandolfo, 1997]. Однако линейные приближения правомерны только для малоамплитудных колебаний и не могут описывать сложные характеристики большеамплитудных и апериодических колебаний. Здесь для понимания природы сложных экономических процессов стали использоваться нелинейные динамические модели и математическая теория бифуркаций [Puu, 1989; Rosser, 1991; Lorenz, 1993; Thomas, 2005; Zhang, 2005].

В сложной экономической динамике наряду с порядком возникает и хаос. Неустойчивые периодические орбиты как основа хаотических аттракторов рассматривалась в ряде работ [Hilborn, 1994; Szabo, 2000; Chian, 2007; Jakimowicz, 2009]. Одним из новых активно развивающихся направлений современной теоретической экономики является эконофизика [Романовский, Романовский, 2012; Чернавский и др., 2011], использующая развитые методы нелинейной динамики и теории вероятности.

В работах [Трубецков, 2004; Трубецков, 2006] продемонстрировано, как канонические примеры известных систем нелинейной динамики могут использоваться в математическом моделировании колебательных процессов в экономике. Показано, как математический аппарат теории бифуркаций систем дифференциальных уравнений может быть использован в анализе динамики базовых экономических показателей.

Предельные циклы являются основным элементом теории нелинейных динамических систем и ключом к пониманию происхождения и природы колебательных явлений в экономике. Одной из первых моделей макроэкономических циклов является модель Гудвина [Goodwin, 1951]. Эта модель и ее различные модификации до сих пор эффективно применяются при исследовании механизмов развития экономических процессов, характеризующихся цикличностью.

Детальный анализ модели Гудвина и его свойств был проведен в работе [Strotz, 1953]. Обобщенная форма этой модели была дана Лоренцем [Lorenz, 1993]:

$$\ddot{x}(t) + A(x(t))\dot{x}(t) + B(x(t)) = O^*(t),$$

где x — отклонение дохода от равновесия, $A(x)$ — четная функция с условиями $A(0) < 0$, $A''(0) > 0$, $B(x)$ — нечетная функция с условием $B(0) = 0$, $O^*(t)$ — функция издержек.

В своих исследованиях Lorenz и Nusse [Lorenz, Nusse, 2002] предложили конкретный вид функций $A(x)$ и $B(x)$

$$\ddot{x} + a \frac{x^2 - 1}{x^2 + 1} \dot{x}(t) - bx + cx^3 = O^*(t).$$

Функцию $O^*(t)$ можно трактовать как некоторое внешнее воздействие. В случае, когда это воздействие отсутствует ($O^*(t) \equiv 0$), в этой системе наблюдается устойчивый предельный цикл. Если $O^*(t)$ — периодическая функция, то модель демонстрирует переходы от регулярных колебаний к хаотическим [Lorenz, Nusse, 2002]. Различные модификации модели Гудвина исследовались в работах (см. например [Yoshida, 2007; Li, 2011; Cao, 2011]). Несомненный интерес представляет анализ модели, когда $O^*(t)$ является случайной функцией.

В данной статье рассматривается стохастическая модель экономической динамики Гудвина, задаваемая дифференциальным уравнением второго порядка

$$\ddot{x} + a \frac{x^2 - 1}{x^2 + 1} \dot{x} - bx + cx^3 = \varepsilon \dot{w}(t), \quad (1)$$

где $a > 0$, $b > 0$, $c > 0$, w — стандартный винеровский процесс, ε — интенсивность шума.

Раздел 2 данной статьи посвящен анализу соответствующей детерминированной модели ($\varepsilon = 0$), построена бифуркационная диаграмма, подробно описаны параметрические зоны для равновесий и исследована устойчивость аттракторов модели.

В разделе 3 приведен краткий обзор аппарата функции стохастической чувствительности (ФСЧ).

В разделе 4 подробно изучается стохастическая модель Гудвина. Здесь с помощью аппарата ФСЧ исследуются стохастические аттракторы модели.

В разделе 5 описан эффект вызванной шумом генерации бизнес-циклов в зоне параметров, где исходная детерминированная система имеет лишь устойчивые равновесия. Это явление приводит к качественному изменению формы плотности распределения (P -бифуркации [Arnold, 1998]).

2. Анализ детерминированной модели

Запишем детерминированную модель Гудвина (1) ($\varepsilon = 0$) в виде системы

$$\begin{cases} \dot{x} = y, \\ \dot{y} = -a \frac{x^2 - 1}{x^2 + 1} y + bx - cx^3. \end{cases} \quad (2)$$

В системе (2) существуют три равновесия $M_0(0, 0)$, $M_1\left(\sqrt{\frac{b}{c}}, 0\right)$, $M_2\left(-\sqrt{\frac{b}{c}}, 0\right)$. Они лежат на оси Ox и имеют абсциссы, не зависящие от параметра a . Равновесия M_1 , M_2 симметричны относительно M_0 .

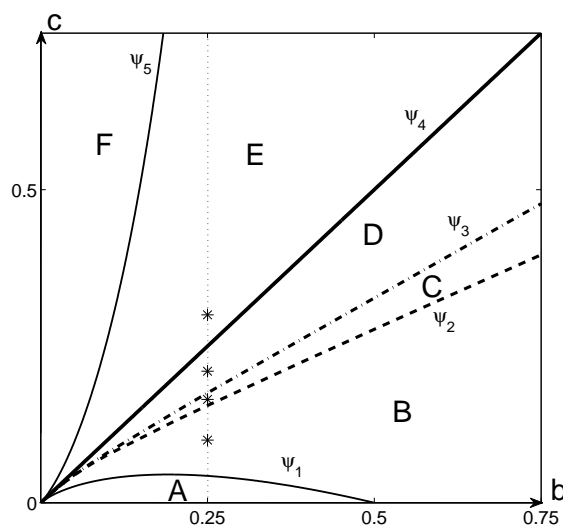
Классический анализ устойчивости равновесий основывается на системе первого приближения, матрица которой для (2) имеет вид

$$F = \begin{pmatrix} f'_x(\bar{x}, \bar{y}) & f'_y(\bar{x}, \bar{y}) \\ g'_x(\bar{x}, \bar{y}) & g'_y(\bar{x}, \bar{y}) \end{pmatrix},$$

где \bar{x}, \bar{y} — координаты равновесия и

$$\begin{aligned} f(x, y) &= y, & g(x, y) &= -a \frac{x^2 - 1}{x^2 + 1} y + bx - cx^3, \\ f'_x(x, y) &= 0, & g'_x(x, y) &= -a \frac{4xy}{(x^2 + 1)^2} + b - 3cx^2, \\ f'_y(x, y) &= 1, & g'_y(x, y) &= -a \frac{x^2 - 1}{x^2 + 1}. \end{aligned}$$

Проведенный анализ показывает, что равновесие $M_0(0, 0)$ при любых значениях параметров является седлом. Это означает, что взятое за точку отсчета равновесие переменной дохода всегда неустойчиво. Такая неустойчивость является внутренней причиной, порождающей в исследуемой экономической системе другие типы функционирования, связанные с нетривиальными равновесиями M_1 , M_2 и возможными автоколебаниями вокруг них. Математический анализ таких режимов связан с исследованием типов фазовых портретов и их бифуркаций. Локальные фазовые портреты равновесий M_1 , M_2 имеют один тип. Их параметрическое описание представлено в сводной бифуркационной диаграмме (рис. 1).

Рис. 1. Бифуркационная диаграмма при $a = 2$

Здесь для $a = 2$ представлены следующие зоны: А — устойчивый узел, $B \cup C \cup D$ — устойчивый фокус, Е — неустойчивый фокус, F — неустойчивый узел. Три границы, разделяющие зоны, имеют следующие аналитические представления:

$$\psi_1(a, b) = \frac{b}{a^2 - 8b}(a^2 + 8b - 4a\sqrt{2b}),$$

$$\psi_4(a, b) = b,$$

$$\psi_5(a, b) = \frac{b}{a^2 - 8b}(a^2 + 8b + 4a\sqrt{2b}).$$

Граница $c = \psi_2(a, b)$ (пунктирная линия) соответствует бифуркационной кривой, отвечающей жесткому рождению полуустойчивого цикла, а граница $c = \psi_3(a, b)$ — разделению одного неустойчивого цикла на два. Функции $\psi_2(a, b)$ и $\psi_3(a, b)$, не имеющие аналитического представления, были найдены численными методами. При фиксированном значении параметра $b = 0.25$ и увеличении параметра c происходят следующие изменения фазового портрета:

- $A \rightarrow B$: два устойчивых узла преобразуются в фокусы,
- $B \rightarrow C$: вокруг равновесий жестко рождается полуустойчивый цикл, который расщепляется на внешний устойчивый и внутренний неустойчивый, разделяющий бассейны притяжения устойчивого цикла и равновесий M_1 и M_2 ,
- $C \rightarrow D$: неустойчивый цикл расщепляется на два неустойчивых меньшего размера, охватывающих равновесия M_1 и M_2 соответственно,
- $D \rightarrow E$: оба неустойчивых цикла сливаются с равновесиями M_1 и M_2 , которые теряют устойчивость, и единственным аттрактором системы остается устойчивый цикл,
- $E \rightarrow F$: неустойчивые фокусы преобразуются в неустойчивые узлы, цикл остается устойчивым.

На рисунке 2 для характерных значений параметров, отмеченных звездочками на рисунке 1, представлены фазовые портреты, где жирной сплошной линией изображен устойчивый цикл, пунктиром — неустойчивые циклы, тонкими сплошными — фазовые траектории, кружками — устойчивые равновесия, окружностями — неустойчивые равновесия.

На рисунке 3 (сверху) демонстрируется зависимость размера аттракторов системы (2) от параметра c . Здесь представлены x -координаты аттракторов: сплошная жирная линия отражает

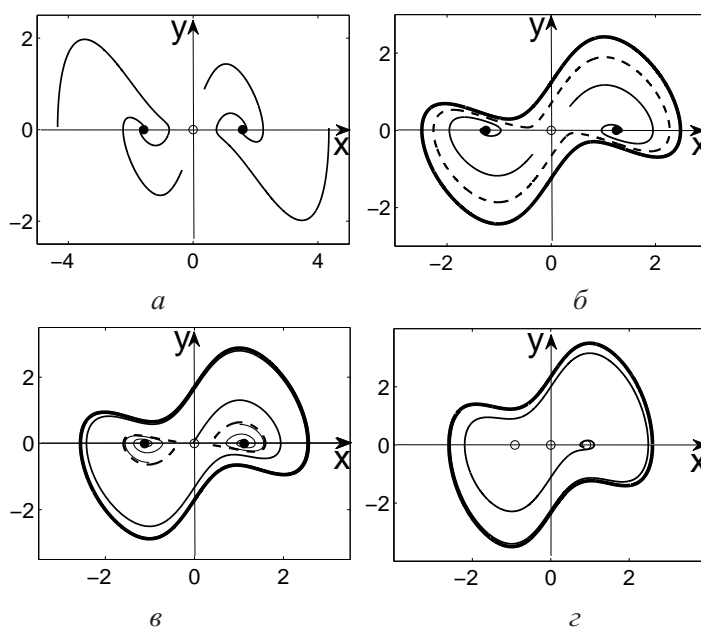


Рис. 2. Фазовые траектории системы для $a = 2, b = 0.25$ при а) $c = 0.1$, б) $c = 0.16$, в) $c = 0.2$ и г) $c = 0.3$

изменение координат устойчивых равновесий, точки — неустойчивых равновесий, тонкая сплошная — экстремальных значений устойчивого цикла, пунктир — экстремальных значений неустойчивых циклов. Бифуркационными значениями являются $c_1 = 0.155332, c_2 = 0.174635$ и $c_3 = 0.25$. Значение c_1 соответствует рождению полуустойчивого цикла, c_2 — расщеплению неустойчивого цикла на два, а c_3 — потере устойчивости равновесий M_1, M_2 . Также можно заметить, что устойчивый цикл практически не меняет своего размера по оси Ox .

На рисунке 3 (снизу) представлены зависимости характеристических показателей Ляпунова для детерминированных аттракторов системы (2) при изменении параметра c и фиксированных $a = 2, b = 0.25$. Здесь сплошной жирной линией выделен характеристический показатель устойчивых равновесий, точками — неустойчивых равновесий, тонкой сплошной линией — устойчивого цикла, пунктиром — неустойчивых циклов.

Как видим из рисунке 3 (снизу), при стремлении c к c_3 слева степень устойчивости равновесий M_1, M_2 уменьшается. При увеличении c от c_1 сначала происходит резкое увеличение степени устойчивости предельного цикла, а затем значение его характеристического показателя стабилизируется.

С экономической точки зрения полученные результаты означают следующее. В данной системе динамика дохода существенно зависит как от параметра c , так и от начальных значений динамических переменных $x_0, y_0 = \dot{x}_0$. Действительно, при $0 < c < c_1$ в зависимости от начальных значений доход стремится к одному из двух нетривиальных равновесных значений, связанных с M_1 или M_2 . В зоне $c_1 < c < c_3$ при определенных начальных условиях возможен еще один, третий тип динамики — устойчивые колебания фиксированной частоты и амплитуды. При $c > c_3$ устойчивые равновесные режимы исчезают и в системе остается только устойчивый бизнес-цикл.

Функционирование любой экономической системы всегда сопровождается случайными возмущениями как внешними (аддитивными), так и внутренними (параметрическими). Присутствие случайного фактора всегда приводит к деформации режимов динамического поведения исходной детерминированной модели.

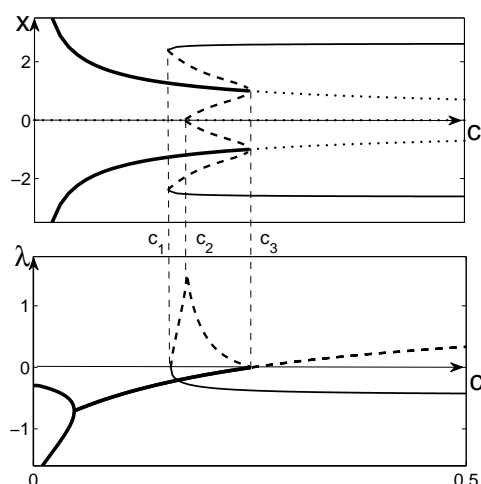


Рис. 3. Экстремумы по x и характеристический показатель аттракторов для $a = 2, b = 0.25$ при изменении c

В данной работе исследование случайных воздействий на динамику модели Гудвина опирается на технику функций стохастической чувствительности.

3. Аппарат функции стохастической чувствительности

Стандартной моделью динамической системы со случайными возмущениями является стохастическая система уравнений Ито [Хасьминский, 1969]

$$dx = f(x) dt + \varepsilon \sigma(x) dw(t). \quad (3)$$

Здесь $f(x)$ — достаточно гладкая n -вектор-функция, $\sigma(x)$ — достаточно гладкая $n \times n$ -матричная функция, задающая зависимость случайных возмущений от состояния системы, $w(t)$ — n -мерный стандартный винеровский процесс, ε — параметр интенсивности возмущений. Предполагается, что соответствующая детерминированная система (3) ($\varepsilon = 0$) имеет экспоненциально устойчивый аттрактор.

В результате действия шумов случайные траектории системы (3) покидают детерминированный аттрактор и формируют некоторый стохастический аттрактор со стационарным распределением $\rho(x, \varepsilon)$. Функция $\rho(x, \varepsilon)$ удовлетворяет стационарному уравнению Фоккера-Планка-Колмогорова (ФПК). Непосредственное использование этого уравнения уже для двумерных систем является технически трудной задачей. В этих обстоятельствах, широко используются асимптотики [Вентцель, Фрейдлин, 1979], основанные на функции квазипотенциала $v(x) = -\lim_{\varepsilon \rightarrow 0} \varepsilon^2 \ln \rho(x, \varepsilon)$. В случае малых шумов с помощью квазипотенциала можно записать асимптотику стационарной плотности как

$$\rho(x, \varepsilon) \approx K \cdot \exp\left(-\frac{v(x)}{\varepsilon^2}\right).$$

Конструктивный подход к аппроксимации квазипотенциала вблизи детерминированного аттрактора рассматривался в [Мильштейн, Ряшко, 1995].

В простейшем случае, когда аттрактором является точка покоя \bar{x} , для квазипотенциала используется квадратичная аппроксимация $v(x) \approx \frac{1}{2}(x - \bar{x}, W^{-1}(x - \bar{x}))$. Эта аппроксимация позволяет

представить асимптотику стационарной плотности в форме нормального распределения

$$\rho(x, \varepsilon) \approx K \exp\left(-\frac{(x - \bar{x}, W^{-1}(x - \bar{x}))}{2\varepsilon^2}\right)$$

с ковариационной матрицей $\varepsilon^2 W$.

Положительно определенная матрица W — матрица стохастической чувствительности — является решением алгебраического уравнения

$$FW + WF^T = -S, \tag{4}$$

где

$$F = \frac{\partial f}{\partial x}(\bar{x}), \quad S = GG^T, \quad G = \sigma(\bar{x}).$$

Эта матрица характеризует разброс случайных траекторий системы (3) вокруг равновесия \bar{x} .

Теперь рассмотрим случай, когда детерминированным аттрактором является предельный цикл, задаваемый T -периодическим решением $x = \xi(t)$. Пусть Π_t — гиперплоскость, ортогональная циклу в точке $\xi(t)$. В этом случае с помощью соответствующей квадратичной аппроксимации квазитенциала вблизи цикла для сечения Пуанкаре Π_t можно записать экспоненциальную гауссовскую асимптотику

$$\rho_t(x, \varepsilon) = K \exp\left(-\frac{(x - \xi(t))^T W^+(t)(x - \xi(t))}{2\varepsilon^2}\right)$$

со средним значением $m_t = \xi(t)$ и ковариационной матрицей $D(t, \varepsilon) = \varepsilon^2 W(t)$.

Матрица $W(t)$, играющая роль функции стохастической чувствительности цикла, является решением краевой задачи

$$\dot{W} = F(t)W + WF^T(t) + P(t)S(t)P(t), \quad W(t+T) = W(t), \quad W(t)r(t) = 0. \tag{5}$$

Здесь

$$F(t) = \frac{\partial f}{\partial x}(\xi(t)), \quad S(t) = G(t)G^T(t), \quad G(t) = \sigma(\xi(t)), \quad r(t) = f(\xi(t)), \quad P(t) = P_{r(t)}, \quad P_r = I - \frac{rr^T}{r^T r}.$$

Система (5), благодаря экспоненциальной устойчивости цикла, имеет единственное решение [Bashkirtseva, Ryashko, 2004].

В случае цикла на плоскости матрицы $W(t)$ и $P(t)$ имеют ранг, равный единице, и представимы в виде

$$W(t) = m(t)P(t), \quad P(t) = p(t)p^T(t).$$

Здесь $p(t)$ — нормированный вектор, ортогональный касательному вектору $f(\xi(t))$, а $m(t) > 0$ — T -периодическая скалярная функция, задающая разброс (дисперсию) пучка по нормали к циклу.

Функция $m(t)$ [Bashkirtseva, Ryashko, 2000; Башкирцева, Перевалова, 2007] удовлетворяет краевой задаче

$$\dot{m} = a(t)m + b(t), \quad m(0) = m(T) \tag{6}$$

с T -периодическими коэффициентами

$$a(t) = p^T(t)(F^T(t) + F(t))p(t), \quad b(t) = p^T(t)S(t)p(t).$$

Функция $m(t)$ определяет локальную стохастическую чувствительность цикла в точке $\xi(t)$. Удобной характеристикой стохастического цикла в целом является коэффициент стохастической чувствительности $M = \max_{(0,T)} m(t)$.

4. Стохастическая модель Гудвина

Для изучения влияния случайных возмущений на динамику модели Гудвина перейдем к стохастической системе

$$\begin{cases} \dot{x} = y, \\ \dot{y} = -a\frac{x^2-1}{x^2+1}y + bx - cx^3 + \varepsilon\dot{w}, \end{cases} \quad (7)$$

где w — стандартный винеровский процесс, моделирующий случайные внешние воздействия, ε — интенсивность шума. Траектории системы (7) под действием случайных возмущений покидают детерминированный аттрактор (равновесия или цикл) и формируют вокруг него некоторый пучок. При увеличении интенсивности шума разброс случайных состояний вокруг детерминированного аттрактора увеличивается. В данной статье представлены результаты исследования влияния шума на аттракторы системы и вызванные этим шумом качественные изменения динамики модели. Здесь и далее зафиксируем два параметра детерминированной системы $a = 2$ и $b = 0.25$.

В первую очередь рассмотрим случай стохастического равновесия на интервале изменения параметра $0 < c < c_3$. На рисунке 4 при одном и том же значении интенсивности $\varepsilon = 0.01$ представлены стохастические траектории системы (7) вблизи равновесия M_1 для $c = 0.22$ (слева) и $c = 0.05$ (справа). Здесь экономическая система демонстрирует случайные флуктуации малой амплитуды вокруг нетривиальных равновесий. Как легко заметить, разброс случайных состояний существенно зависит от параметра c . Аппарат функции стохастической чувствительности позволяет аналитически описать этот эффект, полученный численно.

Для стохастических равновесий M_1 и M_2 модели (7) матрица стохастической чувствительности, найденная из (4), имеет диагональный вид

$$W = \begin{pmatrix} \frac{c+b}{4ab(b-c)} & 0 \\ 0 & \frac{c+b}{2a(b-c)} \end{pmatrix}.$$

На интервале $0 < c < c_3$ структурной устойчивости равновесий, зависимость диагональных элементов $w_1 = \frac{c+b}{4ab(b-c)}$ и $w_2 = \frac{c+b}{2a(b-c)}$ матрицы W от параметра c представлена на рисунке 5. Как видно, стохастическая чувствительность равновесий при приближении параметра c к бифуркационному значению $c_3 = 0.25$ монотонно возрастает и стремится к бесконечности.

При $c > c_1$ вокруг детерминированного устойчивого цикла формируется пучок случайных состояний. Ширина этого пучка вдоль цикла неравномерна. Это означает, что стохастическая чувствительность отдельных участков цикла может существенно меняться. Функция стохастической чувствительности позволяет детально описать это явление. На рисунке 6 представлены стохастический цикл для $c = 0.2$, $\varepsilon = 0.1$ и соответствующий ему график функции стохастической чувствительности $m(t)$, полученный численным решением задачи (6). Заметим, что на цикле есть два наиболее чувствительных участка: на графике $m(t)$ им соответствуют высокие узкие пики. Как видим, здесь доход демонстрирует сложное колебательное движение, сочетающее детерминированную и стохастическую компоненты.

На рисунке 7 представлен график зависимости коэффициента стохастической чувствительности M от параметра c . Аналогично случаю стохастического равновесия здесь наблюдается неограниченный рост стохастической чувствительности при стремлении параметра c к значению c_1 справа. При увеличении параметра c коэффициент чувствительности M монотонно стремится к нулю, что свидетельствует об уменьшении стохастической чувствительности цикла.

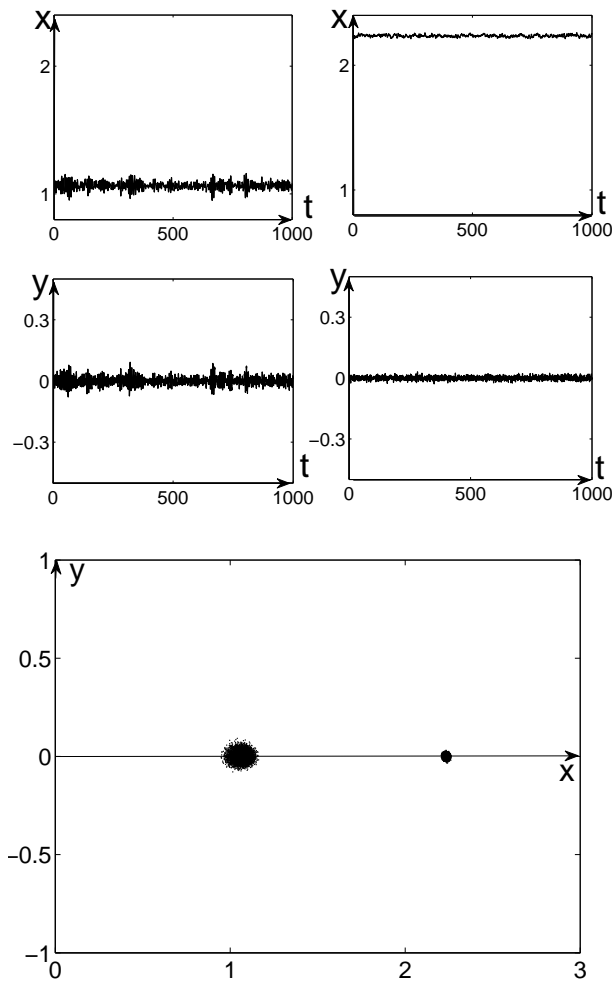


Рис. 4. Стохастическое равновесие при $a = 2$, $b = 0.25$ и $\varepsilon = 0.01$ для $c = 0.22$ (слева) и $c = 0.05$ (справа)

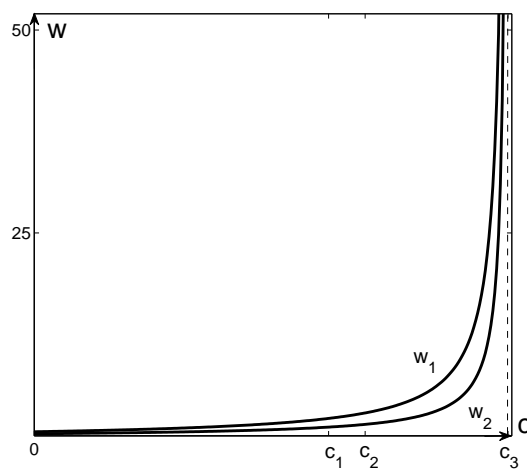


Рис. 5. Коэффициент стохастической чувствительности равновесия при $a = 2$, $b = 0.25$

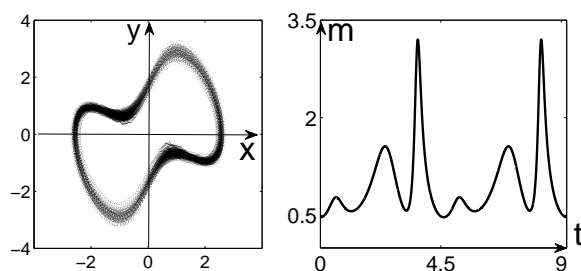


Рис. 6. Стохастический цикл для $a = 2$, $b = 0.25$, $c = 0.2$, $\varepsilon = 0.1$ и функция стохастической чувствительности цикла

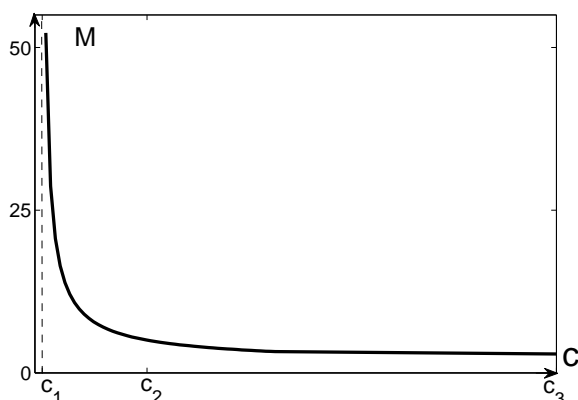


Рис. 7. Коэффициент стохастической чувствительности цикла $a = 2$, $b = 0.25$ при изменении c

5. Индуцированные шумом переходы

Рост интенсивности случайных возмущений приводит не только к количественным изменениям, сопровождающимся увеличением разброса случайных состояний вокруг детерминированных аттракторов, но и может порождать различные качественные эффекты. Рассмотрим зону параметров, где детерминированная система (7) имеет в качестве аттракторов лишь два устойчивых равновесия. На рисунках 8, *a*, *б* для $c = 0.15$ представлены случайные состояния и траектории $x(t)$ системы (2) под действием шумов различной интенсивности: $\varepsilon = 0.05$ (слева) и $\varepsilon = 0.3$ (справа).

В рассматриваемом случае при увеличении интенсивности шума наряду с увеличением разброса случайных состояний вокруг M_1 и M_2 наблюдаются и качественные изменения. Эти изменения сопровождаются деформацией стационарной плотности распределения (рис. 8, *в*). При малом шуме, случайные состояния локализуются вблизи детерминированных равновесий M_1 и M_2 (рис. 8, *a* слева). В этом случае функция распределения имеет два острых пика. При дальнейшем увеличении интенсивности шума форма графика стационарной плотности преобразуется в кратероподобную. Случайные траектории под действием шума покидают бассейны притяжения детерминированных равновесий и проходят через области фазовой плоскости, достаточно удаленные от этих равновесий (рис. 8, *a* справа). При этом случайные состояния концентрируются вдоль некоторой замкнутой кривой, форма которой близка к форме детерминированного цикла, возможного лишь при $c > c_1$. На рисунке 8, *б* показаны траектории переменной x по времени для двух этих режимов. В случае малого шума случайные траектории концентрируются

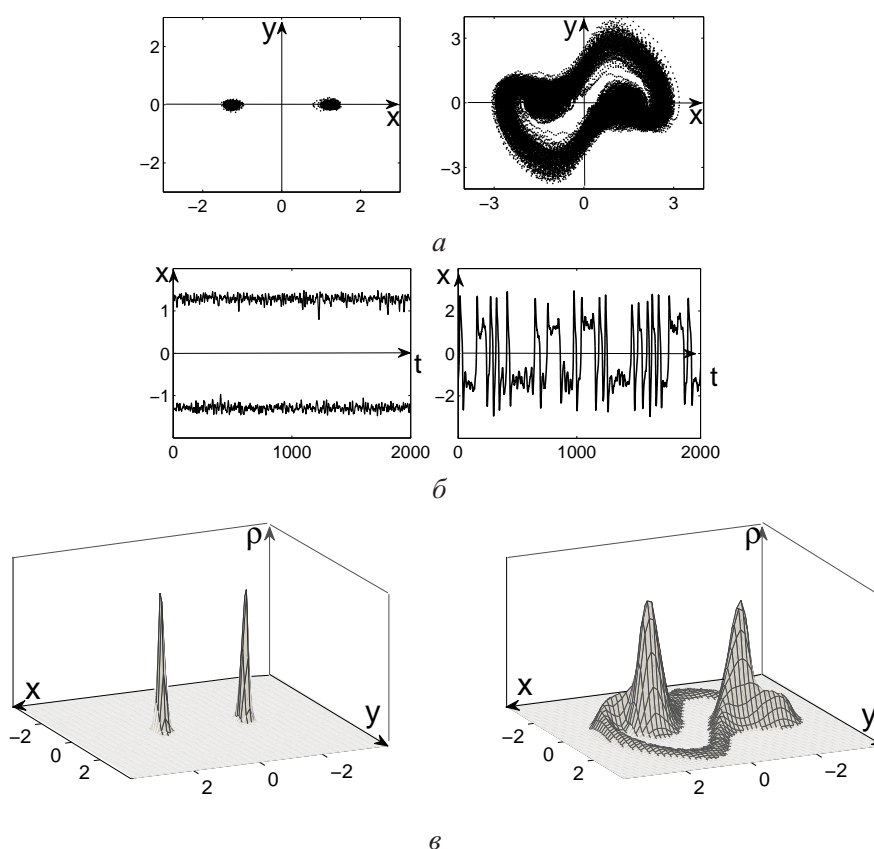


Рис. 8. Индуцированный шумом переход при $a = 2$, $b = 0.25$, $c = 0.15$ при $\varepsilon = 0.05$ (слева) и $\varepsilon = 0.3$ (справа)

вокруг детерминированных равновесий, переходов между которыми нет (рис. 8, б). Увеличение шума приводит к переходам случайных траекторий между равновесиями M_1 и M_2 .

Можно сказать, что при достаточно большом шуме в модели Гудвина сочетаются колебания малой амплитуды вокруг устойчивых равновесий и колебания большой амплитуды — стохастический цикл. Такое качественное изменение стохастической динамики (P -бифуркация) может быть интерпретировано как вызванная шумом генерация бизнес-цикла в зоне, где исходная детерминированная модель имеет лишь устойчивые равновесия.

Авторы признательны Ряшко Л. Б. за плодотворные обсуждения материалов работы.

Список литературы

- Баширцева И. А., Первалова Т. В. Анализ стохастических аттракторов при бифуркации точка покоя — цикл // Автоматика и телемеханика. — 2007. — № 10. — С. 53–69.
- Вентцель А. Д., Фрейдлин М. И. Флуктуации в динамических системах под действием малых случайных возмущений. — М.: Наука, 1979.
- Мильштейн Г. Н., Ряшко Л. Б. Первое приближение квазипотенциала в задачах об устойчивости систем со случайными невырожденными возмущениями // Прикл. математика и механика, 1995. — Т. 59, № 1. — С. 53–63.
- Романовский М. Ю., Романовский Ю. М. Введение в эконофизику: статистические и динамические модели. — ИКИ, 2012.
- Трубецков Д. И. Введение в синергетику. Хаос и структуры. — М: Едиториал УРСС, 2004.

- Трубецков Д. И.* Канонические модели нелинейной динамики в экономике // Известия вузов. Прикладная нелинейная динамика, 2006. — Т. 14, № 2. — С. 75–93.
- Хасьминский Р. З.* Устойчивость систем дифференциальных уравнений при случайных возмущениях их параметров. — М.: Наука, 1969.
- Чернавский Д. С., Старков Н. И., Малков С. Ю., Косе Ю. В., Щербаков А. В.* Об эконофизике и её месте в современной теоретической экономике // УФН, 2011. — Vol. 181:7. — С. 767–773.
- Arnold L.* Random Dynamical Systems. — Berlin–Heidelberg: Springer-Verlag, 1998.
- Bashkirtseva I., Ryashko L.* Stochastic sensitivity of 3D-cycles // Mathematics and Computers in Simulation, 2004. — V. 66. — P. 55–67.
- Bashkirtseva I. A., Ryashko L. B.* Sensitivity analysis of the stochastically and periodically forced Brusselator // Physica A, 2000. — V. 278. — P. 126–239.
- Cao J., Jiang H.* Stability and Hopf bifurcation analysis on Goodwin model with three delays // Chaos, Solutions and Fractals, 2011. — No. 44. — P. 613–618.
- Chian A. C. L.* Complex Systems Approach to Economic Dynamics. — Berlin–Heidelberg: Springer-Verlag, 2007.
- Gandolfo G.* Economic dynamics. — Berlin: Springer, 1997.
- Goodwin R. M.* The nonlinear accelerator and the persistence of business cycles // Econometrica, 1951. — Vol. 19, No. 1. — P. 1–17.
- Hilborn R. C.* Chaos and nonlinear dynamics: an introduction for scientists and engineers. — New York: Oxford University Press, 1994.
- Jakimowicz A.* Catastrophes and chaos in business cycle theory. — Proceedings of the 4th Polish symposium on Econo- and Sociophysics, Rzeszow, Poland, May 2009.
- Li S., Li Q., Li J., Feng J.* Chaos prediction and control of Goodwin's nonlinear accelerator model // Nonlinear analysis: Real world applications, 2011. — No. 12. — P. 1950–1960.
- Lorenz H. W.* Nonlinear dynamical economics and chaotic motion, 2nd ad. — Berlin–Heidelberg–New York: Springer, 1993.
- Lorenz H. W., Nusse H. E.* Chaotic attractors, chaotic saddles, and fractal basin boundaries: Goodwin's nonlinear accelerator model reconsidered // Chaos, Solutions and Fractals, 2002. — No. 13. — P. 957–965.
- Puu T.* Nonlinear economic dynamics. Lecture notes in economics and mathematical systems. — Berlin: Springer-Verlag, 1989. — Vol. 336.
- Rosser J. B.* From catastrophe to chaos: a general theory of economic discontinuities. — Boston: Kluwer Academic Publishers, 1991.
- Rosser J. B.* Chaos theory before Lorenz // Nonlinear dynamics psychology and life sciences. 2009. — Vol. 13, No. 3. — P. 257–269.
- Scarth W. M.* Macroeconomics: an introduction to advanced methods. — Toronto: Dryden, 1996.
- Strotz R. H., McAnulty J. C., Naines J. B.* Goodwin's nonlinear theory of the business cycle: An electro-analog solution // Econometrica, 1953. — Vol. 21, No. 3. — P. 390–411.
- Szabo K. G., Lai Y.-C., Tel T., Grebogi C.* Topological gap filling at crisis // Phys Rev E, 2000. — No. 61. — P. 5019–5032.
- Thomas L., Reitz S., Samanidou E., editors* Nonlinear dynamics and heterogeneous interacting agents. Lecture notes in economics and mathematical systems. — Berlin: Springer, 2005. — Vol. 550.
- Yoshida H., Asada T.* Dynamic analysis of policy lag in a Keynes-Goodwin model: Stability, instability, cycles and chaos // Journal of Economic Behavior and Organization, 2007. — Vol. 62, No. 3. — P. 441–469.
- Zhang W.-B.* Differential equations, and chaos in economics // Series on Advances in Mathematics for Applied Sciences. World Scientific, 2005. — Vol. 68.